University of Glasgow | Experimental Particle Physics

GLAS-PPE/2009-25
23rd November 2009

Department of Physics and Astronomy
Experimental Particle Physics Group
Kelvin Building, University of Glasgow,
Glasgow, G12 8QQ, Scotland
Telephone: +44 (0)141 330 2000 Fax: +44 (0)141 330 5881

A Monte Carlo simulation free method of measuring lifetimes using event-by-event acceptance functions at LHCb

M. Gersabeck[1], V. V. Gligorov[1], J. Imong[2],
J. Rademacker[2]

[1] CERN, Geneva and University of Glasgow, Glasgow, G12 8QQ, Scotland
[2] University of Bristol

Abstract

A set of innovative methods and tools for precision lifetime and lifetime-difference measurements in hadronic B decays at LHCb is presented. All methods are purely data-driven and Monte Carlo simulation independent, a particularly important feature if lifetime measurements are to be made in the early period of LHCb's data taking. The methods and tools are shown to work in detailed simulation studies, including both Toy and Full Monte Carlo simulation studies of possible systematic biases in the measurements.

1 Introduction

In this note we describe LHCb's potential for measuring B-lifetimes in fully hadronic decay modes in a way which takes account of the lifetime bias in LHCb's impact-parameter based trigger without recourse to Monte Carlo simulations.

The method described here has been previously used at CDF to measure the B^+ lifetime in the decay mode $\mathrm{B}^+ \to \bar{\mathrm{D}}^0\pi^+$ as a proof of principle [1]. Its application to LHCb is described in detail in [2, 3]. A significant advantage for LHCb is the absence of an upper impact parameter cut, which significantly reduces the statistical precision at CDF [4] and also complicates the measurement.

We will give a brief overview of the method (henceforth "the swimming method"), but focus our attention on progress that has been made since our previous report [2], in particular:

- Significant improvement of the acceptance function calculation which is now based on the real trigger software, rather than an emulation, made possible by a special interface to the high Level Trigger (HLT);
- Application of a new fitting method, likelihood subtraction;
- Extension to further decay modes ($\mathrm{B} \to \mathrm{hh}$ and $\mathrm{D} \to \mathrm{hh}$) with a fitting method specifically tailored for multi-signal environments;

2 Lifetimes

2.1 The Role of lifetime measurements in hadronic decays

B-hadron lifetimes, parameters of fundamental importance in their own right, gain special significance due to the precise predictions of Heavy Quark Expansion (HQE) [5, 6, 7, 8]. Precision lifetime measurements, and in particular, precision measurements of the ratios of B-hadron lifetimes, provide a testing ground for this theoretical tool that is frequently relied upon for relating experimental observables to parameters of the CKM matrix. The status of the current measurements is covered in Section 2.2; the rest of this section will give a brief overview of some other areas where lifetime measurements play an important role.

The relative width difference between the long and short lived CP eigenstate of the $\mathrm{B}^0_\mathrm{s} - \bar{\mathrm{B}}^0_\mathrm{s}$ system is predicted to be $\frac{\Delta\Gamma_s}{\Gamma_s} \sim \mathcal{O}(10\%)$ and is currently only weakly constrained. Combined with a measurement of the mass difference between those two states, this parameter could be sensitive to new physics. The most precise measurement [9] of the lifetime difference at LHCb uses the mode $\mathrm{B_s} \to \mathrm{J}/\psi\phi$ [9]. However, this measurement benefits significantly from a precision measurement of the average B lifetime in modes such as $\mathrm{B_s} \to \mathrm{D_s}\pi$. Similarly, a measurement of the CKM angle γ from the decay mode $\mathrm{B_s} \to \mathrm{D_s K}$ benefits significantly [10] from a simultaneous fit to $\frac{\Delta\Gamma_s}{\Gamma_s}$ using the decay mode $\mathrm{B_s} \to \mathrm{D_s}\pi$.

The lifetime difference can also be measured by combining the lifetime measured in decays to flavour-specific states, such as $\mathrm{B_s} \to \mathrm{D_s}\pi$, with those from pure CP eigenstates. One example is the decay mode $\mathrm{B_s} \to \phi\phi$, which requires a similar angular analysis as $\mathrm{B_s} \to \mathrm{J}/\psi\phi$ to differentiate between the CP-even and CP-odd component. While statistically not competitive with $\mathrm{B_s} \to \mathrm{J}/\psi\phi$, it is of specific interest since the decay is mediated by a $\mathrm{b} \to \mathrm{s}$ penguin, which is sensitive to new physics [11] and could affect the lifetime measurement. Due to Standard Model CP violation, the decay $\mathrm{B_s} \to \mathrm{KK}$ measures only approximately the CP-even lifetime; the deviation from the pure CP even lifetime is an interesting parameter that is sensitive to new physics. In both $\mathrm{B_s} \to \phi\phi$ and $\mathrm{B_s} \to \mathrm{KK}$, the lifetime measurement is an ideal first step towards the measurements of CP violating phases in time-dependent decay rate asymmetries.

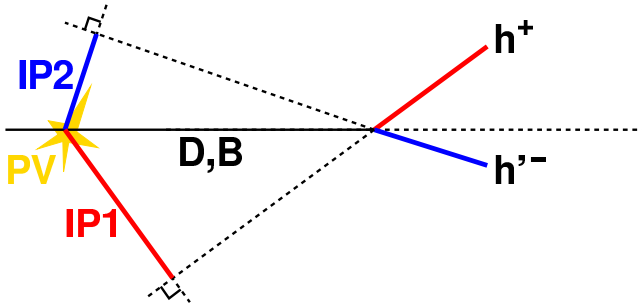
2.2 Status

The latest HFAG [12] averages for the lifetimes and lifetime ratios of B hadrons are given in Table 1, alongside current [13] theoretical predictions. In particular, in the case of the B_s to B^0 lifetime ratio, there is currently an interesting tension between experiment and theory, which could be resolved one way or another by bringing the experimental precision below the percent level. Similarly, for the Λ_b to B^0 lifetime ratio, bringing the experimental measurement to the percent level might stimulate improvements in the theoretical predictions, especially if the improved measurements hinted at tension between theory and experiment.

Table 1: The latest HFAG[12] averages for the lifetimes and lifetime ratios of B hadrons (labeled exp.), and the latest theoretical predictions for the ratios [13].

B hadron	Lifetime (exp.)	Ratio wrt. the B^0 (exp.)	Ratio wrt. the B^0 (theory)
B^0	1.530 ± 0.008 ps		
B^+	1.639 ± 0.009 ps	1.073 ± 0.008	1.063 ± 0.027
B_s	$1.478^{+0.020}_{0.022}$ ps	0.966 ± 0.015	1.00 ± 0.01
Λ_b	1.379 ± 0.051 ps	0.901 ± 0.034	0.88 ± 0.05

Figure 1: Given the 3-momenta of all particles in the decay, the cut on the impact parameter of the decay products translates directly into a cut on the lifetime of the primary particle. This is illustrted for a two body decay, where IP1 corresponds to h^+ and IP2 to h^-.



3 Determining lifetime acceptances in a Monte Carlo simulation free manner

The swimming method has been described in detail elsewhere [1, 2, 3]; the salient points will be recapitulated below for the reader's convenience.

3.1 The Basic idea

Taking a given event and keeping every kinematic aspect of it fixed, except for the decay time of the primary particle, an upper and a lower impact parameter cut directly translate into cuts on the decay-length and hence on the lifetime of the decaying particle, as illustrated for the case of a two-body decay in Figure 1. A more complex scenario is given in Figure 2. The figure illustrates the direct geometric relationship between impact parameter and decay length, and hence lifetime. The impact parameters correspond to the distance between the primary vertex and the point where the backwards extensions of the tracks hit the dashed lines perpendicular to them. Where an individual track passes the IP requirements, the corresponding perpendicular dashed line is solid (in same colour as the corresponding track). Whenever 2 tracks pass the IP requirements, the acceptance as a function of time (plotted at the bottom) is set to one, otherwise zero.

Key to this method is that none of the kinematics needed to translate from an impact parameter cut to a cut on the decay time, have themselves any dependence on the lifetime of the primary particle.

3.2 The signal PDF ignoring measurement errors and other detector effects.

We can write the probability to find an event with decay time t_i as the product of the probability to find t_i given that t_i is between $t_{\min i}$ and $t_{\max i}$ and the probability that t_i is constrained to lie within those limits:

$$\begin{aligned} \mathrm{P}(t_i, [t_{\min i}, t_{\max i}]) &= \mathrm{P}(t_i | t_i \in [t_{\min i}, t_{\max i}]) \cdot \mathrm{P}([t_{\min i}, t_{\max i}]) \\ &= \frac{\frac{1}{\tau} e^{\frac{-t_i}{\tau}}}{\int_{t_{\min i}}^{t_{\max i}} \frac{1}{\tau} e^{\frac{-t'_i}{\tau}} dt'} \cdot \mathrm{P}([t_{\min i}, t_{\max i}]) \\ &= \frac{\frac{1}{\tau} e^{\frac{-t_i}{\tau}}}{e^{\frac{-t_{\min i}}{\tau}} - e^{\frac{-t_{\max i}}{\tau}}} \cdot \mathrm{P}([t_{\min i}, t_{\max i}]) \,, \end{aligned} \tag{1}$$

Figure 2: Given the 3-momenta of all particles in the decay, and the decay lengths of particles down the decay chain (here a D^0), the requirements of the hadronic trigger that two particles pass the IP cut translates into an acceptance of one or more intervals.

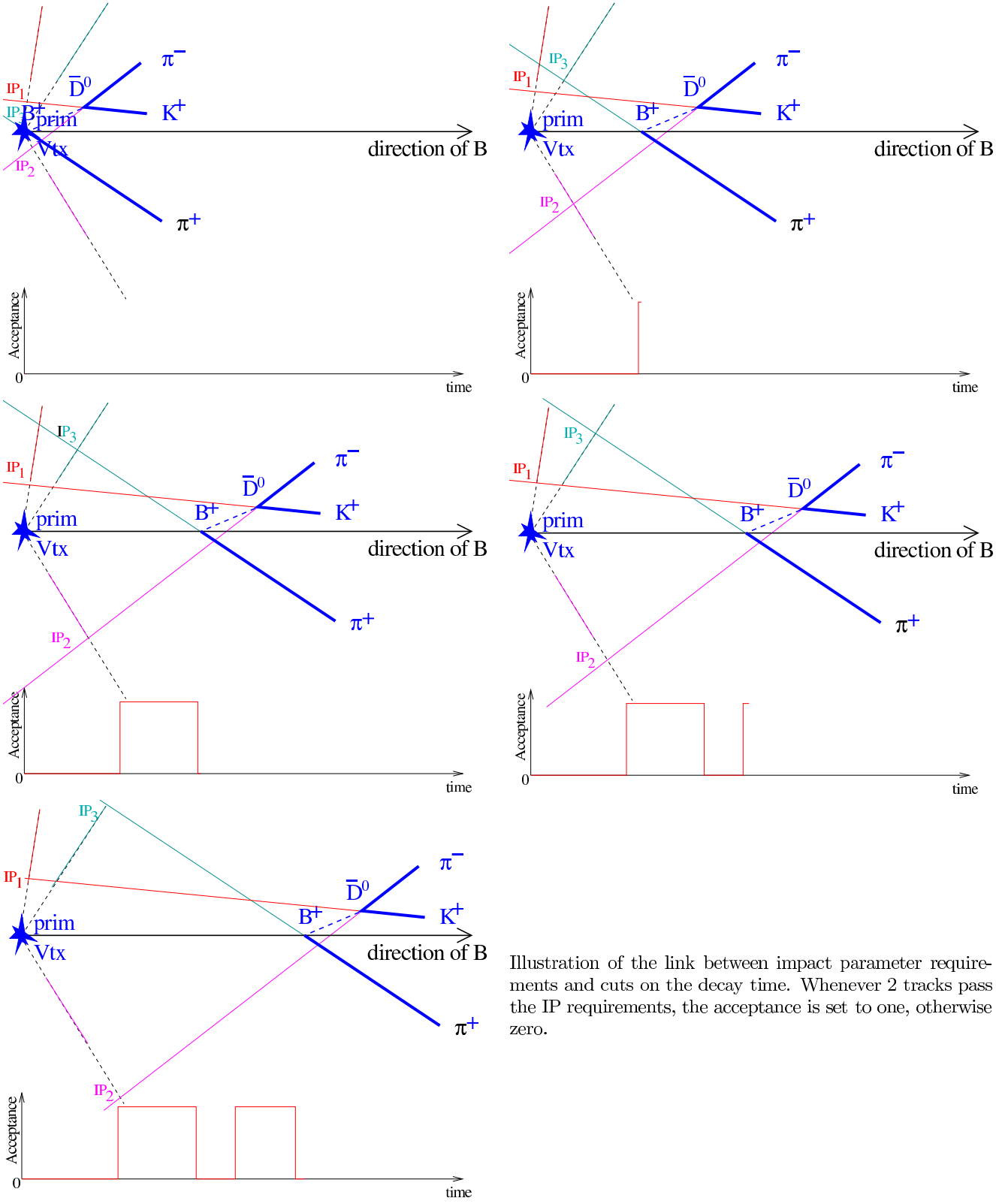


Illustration of the link between impact parameter requirements and cuts on the decay time. Whenever 2 tracks pass the IP requirements, the acceptance is set to one, otherwise zero.

where $t_{\min}$ and $t_{\max}$ are the minimum and maximum decay times which could have been recorded by the experiment. For a series of measurements, the probabilities for each measured time t_i can be multiplied to give the likelihood for the mean decay time τ. The limits $t_{\min i}$ and $t_{\max i}$ can be calculated easily from the kinematics of each decay. In general it will be difficult to calculate $\mathrm{P}(t_{\min i}, t_{\max i})$. However, $\mathrm{P}(t_{\min i}, t_{\max i})$ depends only on the impact parameter cut, and the kinematics of the decay – the momenta of the particles, and possibly the decay lengths some long-lived particles within the decay chain, like the $\mathrm{D_s}$ in $\mathrm{B_s^0} \to \mathrm{D_s}\pi-$ but not on the lifetime of the primary decaying particle itself. So in the log-likelihood, the sum over the $\log(\mathrm{P}(t_{\min i}, t_{\max i}))$ is simply a constant that can be ignored. The total log-likelihood function for a set of N "ideal" decays (no measurement uncertainties, background, etc) is given by:

$$\begin{aligned} \log \mathcal{L} &= -N \log(\tau) \\ &\quad - \sum_{i=1}^{N} \left(\frac{t_i}{\tau} + \log\left(e^{-t_{\min i}/\tau} - e^{-t_{\max i}/\tau} \right) \right) \end{aligned} \tag{2}$$

where the index i labels the event, each of which has its measured decay time t_i and minimum and maximum decay times $t_{\min i}$ and $t_{\max i}$.

Note that the only difference to the the likelihood function without an impact parameter cut is the term:

$$\log \mathcal{L}_{\mathrm{ip}} = -\sum_{i=1}^{N} \log\left(e^{-t_{\min i}/\tau} - e^{-t_{\max i}/\tau} \right) \tag{3}$$

The upper lifetime cut has a dramatic effect on the precision with which the lifetime can be measured. Finding an event with lifetime t contains less information, if the range of possible values for t is already restricted due to lifetime cuts. The effect is quite significant. For example, an upper lifetime cut at twice the B lifetime loses only 14% of events. However, the statistical error of the measurement is increased by a factor of 2, equivalent to a signal loss of 75%. This is discussed in more detail elsewhere [4]. The present design of the LHCb HLT, described in Section 4, does not envisage such a cut being used.

For more complicated decay geometries, the impact parameter cuts on the decay products can translate into a series of disjoint time-intervals which changes the correction term to:

$$\log \mathcal{L}_{\mathrm{ip}} = -\sum_{i=1}^{N} \log\left(\sum_{j=1}^{n_i} e^{-\frac{t_{\min ij}}{\tau}} - e^{-\frac{t_{\max ij}}{\tau}} \right) \tag{4}$$

where i labels the events and j labels the allowed time-intervals for each event.

3.3 Measurement Errors

The likelihood function in Equation 2 is derived for the ideal case of an exact time measurement and impact parameter cut. Any real measurement will have an uncertainty on both.

3.3.1 The signal PDF for an "offline trigger", with measurement errors

In the presence of measurement errors the acceptance is still a top-hat function (or a combination of them), but now as a function of measured decay time, rather than true decay time. Nothing changes in the illustrations in Figure 2, except that the acceptance for the event is plotted as a function of the measured proper time.

We can write the probability density to measure a decay time t_0 (given the IP cut and the decay kinematics that relate the impact parameter to the time measurement for any given decay) as an integral over all true decay times t in terms of the following functions:

- The probability density that a particle decays with true decay time t, given its mean life it τ,

$$\frac{1}{\tau} e^{\frac{-t}{\tau}} \Theta(t).$$

- The probability density that, given the true decay time t and measurement uncertainty of σ_t, the measured decay time is t_0

$$\frac{1}{\sqrt{2\pi}\sigma_t} e^{-\frac{(t-t_0)^2}{2{\sigma_t}^2}}.$$

- The acceptance as a function of the *measured* decay time t_0 for the given decay kinematics.

$$A_{\rm ip}\left(t_0\right).$$

In terms of these parameters, the total probability density is:

$$P\left(t_0, A_{\rm ip}\left(t_0\right)\right) = \frac{\int_0^\infty \frac{1}{\tau} e^{\frac{-t}{\tau}} \frac{1}{\sqrt{2\pi}\sigma_t} e^{-\frac{(t-t_0)^2}{2{\sigma_t}^2}} A_{\rm ip}\left(t_0\right)\, dt}{\int_{-\infty}^\infty \int_0^\infty \frac{1}{\tau} e^{\frac{-t}{\tau}} \frac{1}{\sqrt{2\pi}\sigma_t} e^{-\frac{(t-t_0')^2}{2{\sigma_t}^2}} A_{\rm ip}\left(t_0\right)\, dt\, dt_0'}. \tag{5}$$

The acceptance $A_{\rm ip}\left(t_0\right)$, is given by

$$A_{\rm ip}\left(t_0, \ldots\right) = \sum_{\substack{i=\text{all}\\ \text{intervals}}} \left(\Theta(t_0 - t_{\min\, i}) - \Theta(t_0 - t_{\max\, i})\right), \tag{6}$$

where Θ is the Heaviside function.

3.4 Including Background

3.4.1 Introduction

So far, we have described a Monte Carlo simulation free method to correct for the trigger bias which works on signal data alone. Including background makes the situation considerably more complicated, because the basic trick we applied doesn't quite work anymore. In our PDF, we calculate the probability density to find a lifetime *given* the efficiency function $A_{\rm trig}$, calculated from the decay kinematics that translate the trigger cuts into different lifetime cuts on an event by event basis. The argument was that those kinematics do not themselves depend on the lifetime, and the corresponding term in the PDF can be ignored. Mathematically: In the expression

$$P(t, \text{kin}) = P(t|\text{kin})P(\text{kin}) \tag{7}$$

we can ignore $P(\text{kin})$ because it is a simple factor and $\frac{\mathrm{d}}{\mathrm{d}\tau}P(\text{kin}) = 0$. However, if we add background, the full expression is (where $P(s)$ is the signal probability and $P(b) = 1 - P(s)$ the background probability)

$$P(t, A_{\rm trig}) = P(s)P(t|A_{\rm trig}, s)P(A_{\rm trig}|s) + P(b)P(t|A_{\rm trig}, b)P(A_{\rm trig}|b) \tag{8}$$

Now the efficiency-function terms, $P(A_{\rm trig}|s)$ and $P(A_{\rm trig}|b)$, only factor out if they are the same for signal and background. If they are different for signal and background, ignoring these factors is equivalent to getting the signal fraction wrong in the fit, which is more obvious if we re-write Equation 8 as

$$\begin{aligned} P(t, A_{\rm trig}) &= P(s)P(A_{\rm trig}|s)P(t|A_{\rm trig}, s) + P(b)P(A_{\rm trig}|b)P(t|A_{\rm trig}, b) &(9)\\ &= \left\{P(s|A_{\rm trig})P(t|A_{\rm trig}, s) + P(b|A_{\rm trig})P(t|A_{\rm trig}, b)\right\} P(A_{\rm trig}) &(10)\end{aligned}$$

So we can either, as in Equation 9, fit the probability to find a given acceptance function, or at least, as in Equation 10 calculate an event-by-event signal probability based on the acceptance function. The last term in Equation 10, describing the total probability to get the given acceptance function (whether it's signal or background), does indeed factor out and can be ignored, but if we ignore the kinematics altogether, we will get the event-by-event signal fractions wrong and hence the wrong fit result.

The same problem shows up for anything that changes our PDF event-by-event, be it the event-by-event acceptance functions, or event-by-event lifetime errors. The latter is the example used by Giovanni Punzi when he discusses this effect in [14].

3.4.2 Specific backgrounds

In order to correct for the Punzi effect described above, it is necessary to have not only an event-by-event acceptance function but also an event-by-event probability of this event being signal given this acceptance function. This procedure is discussed further in Sections 7 and 8 for the cases of $B^0_{d,s} \to hh$, where $h = (\pi, K)$ and $B_s \to D_s^\pm \pi$. An alternative strategy to avoid the Punzi effect is likelihood subtraction which is discussed in Section 7.

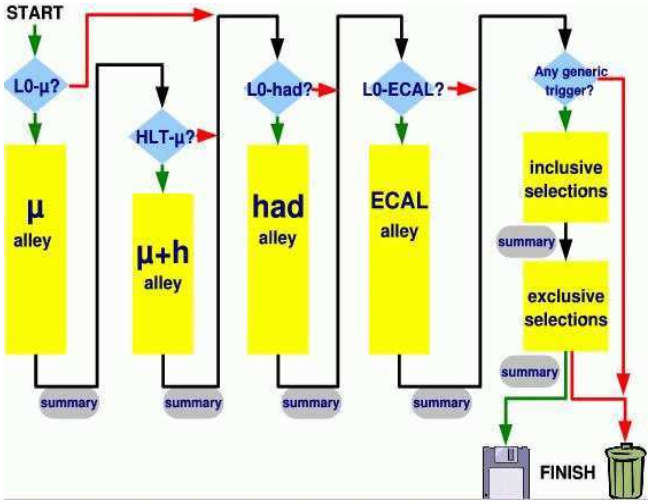


Figure 3: The conceptual layout of the LHCb HLT. The "alleys" refer to the HLT1 stage, while the inclusive and exclusive selections refer to the HLT2 stage of the trigger.

4 The LHCb trigger

The LHCb trigger has recently undergone a major redesign from the structure described in the Trigger TDR [15] of 2003. At the time of writing, no public references exist for the new trigger; a brief overview of those parts most relevant to the present study must therefore be given. The LHCb trigger is rapidly evolving in response to ongoing optimization efforts, and this evolution will continue and intensify once real data is available for optimization and commissioning work. Nonetheless, it is expected that the design described here will be deployed without major architectural changes when LHCb begins data taking.

The LHCb trigger consists of three parts, utilising increasingly sophisticated event reconstructions to discriminate signal from background and reduce the LHCb bunch crossing rate of 40 MHz to a data taking rate of 2 kHz. Of these, the first part is the L0 hardware trigger, while the second and third parts are software triggers collectively referred to as the High Level Trigger (HLT). The L0 trigger uses information from the Calorimeters and Muon chambers to select events with high transverse energy deposits, or high p_T muons, or a dimuon combination close to the J/ψ mass or above the B mass, and reduce the event rate to 1 MHz. Its structure has not been altered from the Trigger TDR, and it is not lifetime biasing. The HLT, on the other hand, uses lifetime biasing information to discriminate between signal and background. It consists of two distinct stages, HLT1 and HLT2.

4.1 HLT1

Not all LHCb subdetectors are read out at the L0 stage of the trigger, and there is no attempt to reconstruct tracks[1)] in the detector. Once the L0 has reduced the data rate to 1 MHz it is possible to read out information from all the LHCb subdetectors for use in the High Level Trigger. The first stage of this trigger is called HLT1 and its purpose is to confirm the L0 trigger decision using the additional information now available. An event may pass the L0 trigger because

1. A hadron with high transverse energy has been detected in the hadronic calorimeter.
2. An electron with high transverse energy has been detected in the electron calorimeter.
3. A photon with high transverse energy has been detected in the electron calorimeter.
4. A π^0 with high transverse energy has been detected in the electron calorimeter.
5. A high p_t muon or dimuon pair has been reconstructed in the muon chambers.

The HLT1 trigger attempts to find the candidate which triggered the L0 and match it to an object (track or calorimeter deposit) reconstructed using the full detector information now available. This process is divided according to the type of L0 decision, as shown in Figure 3. For example, an event in which the L0 triggered

[1)] Except in the special case of muons, and then only the muon chambers are used.

because of a large E_T deposit in the hadron calorimeter would pass through the "hadron alley" in HLT1, while an event in which the L0 triggered because of a dimuon pair would pass through the "muon alley" in HLT1.

Different signal types will therefore trigger in different ways. For example, a $B_d \to D^+\pi^-$ event is likely to pass HLT1 through the hadron alley, while a $B_d \to \rho^+\rho^-$ event could pass through both the hadron and the ECAL alley, depending on whether the π^0 or π^+ daughter of the ρ^+ triggered at L0. It is also important to notice that even though a signal type has a preferred mechanism for passing the trigger, it may still trigger in other ways, for example because of the decay products of the "other" B in the event. On the other hand, the L0-confirmation procedure makes it less likely that the event could trigger because of a random high energy track coming from the primary vertex.

From the point of view of determining the acceptance, the HLT structure poses no special difficulties, as long as every part of the trigger is subjected to the swimming procedure. For each event, the L0 decision (which is lifetime independent) fully determines the subsequent path taken by the event through the HLT. All that is therefore needed is to swim the signal candidate while preserving this same path through the trigger. The current study will apply the swimming method to one specific HLT1 alley (the hadron alley), and ignore the rest of the HLT. Swimming the entire HLT will simply involve repeating this procedure for all other alleys and selection within the trigger.

4.1.1 The HLT1 Hadron Alley

An event enters the HLT1 hadron alley if it passed the L0 trigger due to a hadron with high transverse energy being detected in the hadronic calorimeter. Subsequently, the following takes place

1. The HLT1 track reconstruction is used to reconstruct 2D tracks (in the **r-Z** plane) using information from the VELO. In order for the HLT1 to pass, one of these tracks must match the calorimeter "object" which triggered the L0.

2. If such a 2D VELO track is found, it is upgraded to a 3D VELO track by adding information from the ϕ sensors. The 3D VELO track must have an impact parameter of greater than 0.1 mm with respect to any of the primary vertices in the event in order to survive to the next stage.

3. If such a track is found, the forward reconstruction is run to upgrade the 2D VELO tracks to 3D forward tracks, and to add momentum information to them. The matching track must have a transverse momentum greater than 2.5 GeV, as measured in HLT1, in order for the event to pass the trigger.

4. In order to pass the HLT1 hadron alley, an event now requires either that the track which matched the L0 "object" also has a transverse momentum greater than 5.0 GeV as measured in HLT1 (single hadron trigger), or that an additional track exists which forms a good vertex with this matching track (dihadron trigger). In the latter case, the additional track must have a transverse momentum greater than 1.0 GeV and an impact parameter of greater than 0.1 mm with respect to any of the primary vertices in the event.

The lifetime biasing cuts in the above sequence are the impact parameter cuts applied to the tracks, as well as a "pointing" cut which is applied on the reconstructed secondary vertex in the case of the dihadron trigger, requiring its momentum to point back towards the reconstructed primary vertex.

4.2 HLT2

The HLT1 stage reduces the 1 MHz output of the L0 to approximately 30 kHz. The HLT2 stage is designed to reduce it to 2 kHz, the rate at which data can be written to disk. It consists of a series of inclusive and exclusive reconstruction algorithms. The difference between selecting events within HLT2 and selecting them offline are the time constraints which exist within the trigger. These mean that "fast" algorithms must be used to reconstruct the event, which result in poorer resolutions, for example on the momenta and impact parameters of tracks. The HLT2 does not concern us further here, but the techniques described in this note can also be used to measure HLT2 acceptances.

5 Practical aspects of the method

As described in Section 3, the swimming method requires that the trigger decision is determined for each lifetime of the reconstructed B meson. Since the B comes from a primary vertex, this can be done in two ways: either by moving all the final state tracks associated with the decay of the reconstructed B and re-reconstructing the event, or by moving the primary vertex and recomputing all lifetime biasing variables (impact parameters, flight distances, etc.) with respect to the new primary vertex. The second approach is significantly simpler and has

been chosen here. It assumes that tracks from the B are not used in the reconstruction of the primary vertex; this is almost, but not exactly, the case, and it will introduce a systematic error into the method[2)]. The associated uncertainty is expected to be negligible compared to other sources of error in the lifetime measurements.

In addition to determining the trigger decision at any given lifetime, the swimming method also requires a way of choosing which lifetimes the trigger decision is computed for, so that the acceptance function is determined with the required degree of precision. The implementation of these will now be considered in turn.

5.1 Determining the trigger decision at a given B lifetime

From a technical point of view, the HLT1 trigger is written as a collection of options files which define the sequence of reconstruction and selection steps taken in order to accept or reject an event. Within HLT1 each alley has its own options file, and there are also common options files which define the parameters of the HLT1 reconstruction. What makes swimming the HLT1 trigger particularly trivial is that all steps in the trigger sequence allow the user to define where they should look for input data (tracks, vertices, etc.). In principle, therefore, all that has to be done is to move the primary vertices and tell every step in the HLT1 hadron alley to use these new primary vertices through the relevant input locations. In addition, however, it is also necessary to restrict the trigger to evaluate the decision of tracks from the B alone, since if the event is triggered by tracks which do not come from the B, it will pass the trigger irrespective of the B lifetime. Such events are said to be triggered independently of signal and provide an unbiased sample for lifetime measurements, albeit one which is around an order of magnitude smaller in size than the lifetime biased sample. Therefore, the trigger decision is determined as follows:

1. Find HLT1 2D VELO tracks matched to the final state daughters of the reconstructed B candidate. These are the tracks which the trigger decision will be evaluated on, but they will not be moved in any way.

2. Move the primary vertices in the direction of the reconstructed B momentum to achieve the required B lifetime.

3. Set each subsequent step in the HLT1 hadron alley, as defined in Section 4.1.1, to use the moved primary vertices and evaluate the trigger decision only on those tracks matched to the final state daughters of the reconstructed B candidate.

5.2 Determining the single event acceptance with an arbitrary precision

In order to determine the acceptance function for any single event, it is necessary to retrieve the trigger decision for this event at many possible B lifetimes, a potentially time consuming task. It is therefore desirable to minimize the number of lifetimes for which the trigger decision has to be evaluated. Because the acceptance function is a sequence of top-hat functions, it is only necessary to determine its turning points with the given precision, since the areas between the turning points are zero or one by definition. An iterative technique is therefore adopted. Firstly, the B lifetime is scanned between limits given by the size of the VELO detector in large, equally spaced, steps in order to find the turning points. Next, the region around each turning point is scanned in finer steps to refine the position of the turning point, and this step is repeated as many times as required to achieve the desired precision. The output of the swimming is a list of turning points which fully define the acceptance function for a given B candidate. This procedure is repeated for all B candidates in the event.

5.3 Dealing with events containing multiple primary vertices

Events with more than one primary vertex pose a potential problem for the swimming, insofar as it is possible for tracks from the candidate B to fire the trigger on the wrong primary vertex (i.e. not the one the B came from).

For events with more than one primary vertex it should in theory suffice to apply the same method as for a single PV, i.e. to move all of them coherently in the same direction which is determined by the PV from which the decay originates. This so-called best PV is chosen as the PV where the direct connection between the PV position and the decay vertex matches best with the momentum of the decaying particle.

As events with multiple PVs make up only about ten per cent of all events, they have been ignored in the studies presented here. LHCb's event yields mean that lifetime measurements will be systematics limited anyhow, and using only those events which have one primary vertex will improve the signal purity, further

[2)] The precise size of this error has not been investigated but it is expected to be small, since the total number of tracks in the event is an order of magnitude greater than the number of final state tracks from the B candidate.

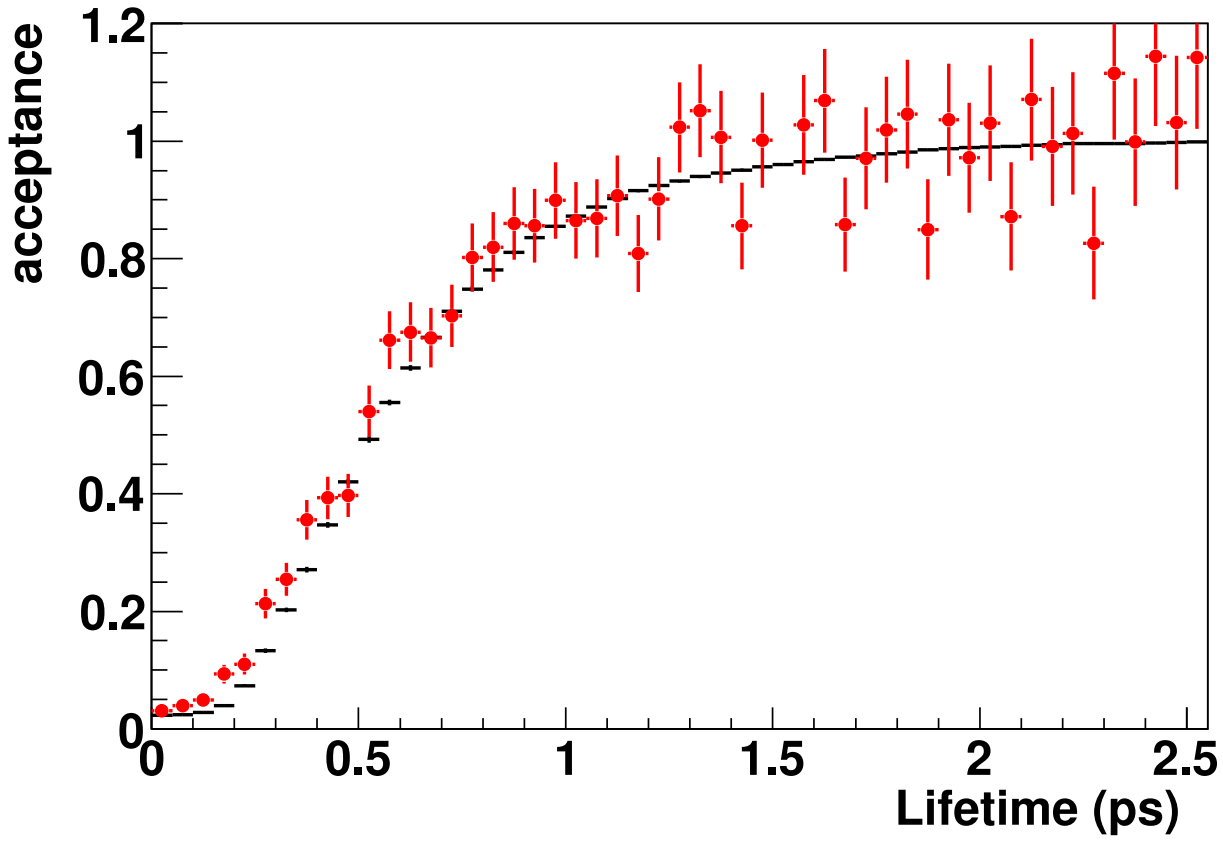


Figure 4: Average acceptance function as determined by the HLT swimming method in black (small markers), and measured lifetime distribution divided by an exponential in red (large markers).

justifying this approach. First attempts to include multiple PV events show that it is possible in principle, however more detailed studies are needed to understand the particularities of these events and any possible biases arising from rejecting them.

6 The trigger acceptance for $B \to hh$ events

6.1 MC tests

As a first test prior to attempting a lifetime fit with the output of the HLT swimming algorithm the results have been used to estimate an average acceptance function. In addition, this will be of particular importance even for measurements that do not involve a lifetime fit.

To get an average acceptance function, the event-by-event acceptance intervals have simply been added up and normalised. For the evaluation of this acceptance function the on-line reconstructed lifetime distribution has been divided by an exponential, again with the normalisation such that it results in an acceptance function between 0 and 1.

The result of this comparison using $1.4M$ $B \to \pi\pi$ events is shown in Figure 4. The red crosses denote the measured lifetime distribution divided by an exponential while the black symbols show the average acceptance function as determined by the HLT swimming method. It can be seen that the shape of the acceptance turn-on at small lifetimes is reproduced, as is the rise towards the plateau starting at around 2 ps. The discrepancy between the two curves is believed to be a result of events which triggered on the daughter particles of the signal B but could **also** have triggered independently of these particles. Such events are lifetime-unbiased, since they can trigger without the presence of the signal B, but were treated as if they triggered on the signal B and therefore as if they were lifetime biased. This would explain why the observed acceptance (in red) is systematically higher than the acceptance obtained from the swimming (black). Further studies are required to confirm or disprove this hypothesis.

For this comparison, only events that were reconstructed correctly (as determined by MC truth) have been considered. When using an offline selection the increased number in mis-identified decays due to large

combinatorics leads to an increase in events observed at small lifetimes and thus to difficulties when making this comparison.

7 Lifetime fits for single-signal environments

7.1 Likelihood Subtraction

The fit for the B lifetime is performed using a modified background subtraction. In this method the same likelihood is applied to a mixed sample of signal and background events, and to a sample of pure background events which are used to correct the fit result.

If the signal likelihood is applied to a sample of events containing some background contribution then the presence of the background will bias the signal lifetime obtained using this likelihood.

The contribution to the likelihood made by the background events can be quantified by calculating the likelihood on a sample of pure background events, in this case obtained from the sidebands of the mass distribution. Then, if the number of background events under the mass peak is known, the contribution to the likelihood from background in the sample can simply be subtracted off to leave the likelihood for signal events only. The likelihood incorporating event-by-event lifetime and acceptance is given below:

$$lnL = \sum_{i=1,N_{peak}} \ln P_{Sig}(t_i \mid \alpha_i) - \frac{n_B}{N_B} \sum_{j=1,N_B} \ln P_{Sig}(t_j \mid \alpha_j) \tag{11}$$

Where t_i is the measured lifetime, α_i is the measured acceptance, N_{peak} is the number of events under the signal peak, n_B is the number of background events in the sample, and N_B is the number of sideband events.

7.2 Obtaining the correct error estimate

The uncertainty in the fit result returned by Minuit is in fact an underestimate of the actual uncertainty. Minuit calculates the error from the shape of the likelihood curve only, and because the fit method corrects for the background contribution to the likelihood, it also corrects for the contribution to the error from background events, leaving only the error from signal. The true error must also contain a contribution from background events, and from the sideband events used in the correction

$$lnL = \sum_{i=1}^{N_S} lnP_{Sig}(t_i;\tau) + \sum_{j=1}^{n_B} lnP_{Sig}(t_j;\tau) - \frac{n_B}{N_B} \sum_{k=1}^{N_B} lnP_{Sig}(t_k;\tau), \tag{12}$$

where N_S indicates signal events, n_B background events, and N_B background events from the sideband region.

Figure 5 provides an illustration of the problem. The total likelihood curve (shown in green), is what Minuit sees when it does the minimisation. The error returned by Minuit is a measure of how 'sharp' the total likelihood curve is in the region about the fit result. But in this case the total likelihood is the sum of signal and background components. The problem is that statistical fluctuations of these components close to the minimum likelihood can affect the fit result. On average the sum of these fluctuations will be zero. However, their magnitude will be non-zero and provides an additional uncertainty that must be taken into account when calculating the error on the fit result.

Although it is obviously not possible to distinguish between signal and background on an event-by-event basis, it does not matter for the sum, and they can be treated separately in this way to find their contribution to the overall error. To find the error we expand the expression about $\tau = \tau_{fit}$ and set $(\partial/\partial\tau)lnL$ to zero

$$N_S {\sigma_S}^{-1}(\tau - \tau_{fit}) = \sum_{i=1}^{N_S} \frac{\partial}{\partial\tau} \delta lnP_{Sig}(t_i;\tau) + \sum_{j=1}^{n_B} \frac{\partial}{\partial\tau} \delta lnP_{Sig}(t_j;\tau) - \frac{n_B}{N_B} \sum_{k=1}^{N_B} \frac{\partial}{\partial\tau} \delta lnP_{Sig}(t_k;\tau), \tag{13}$$

where

$${\sigma_S}^{-1} = \left\langle \left(\frac{\partial lnP_{Sig}}{\partial\tau} \right)^2 \right\rangle_S - \left\langle \frac{\partial lnP_{Sig}}{\partial\tau} \right\rangle_S^2. \tag{14}$$

This describes the shift in the signal lifetime due to fluctuations from the individual terms in the likelihood. On average this shift will sum to zero. However, the contribution to the error comes from the square of these fluctuations, which do not.

For the sum over signal events

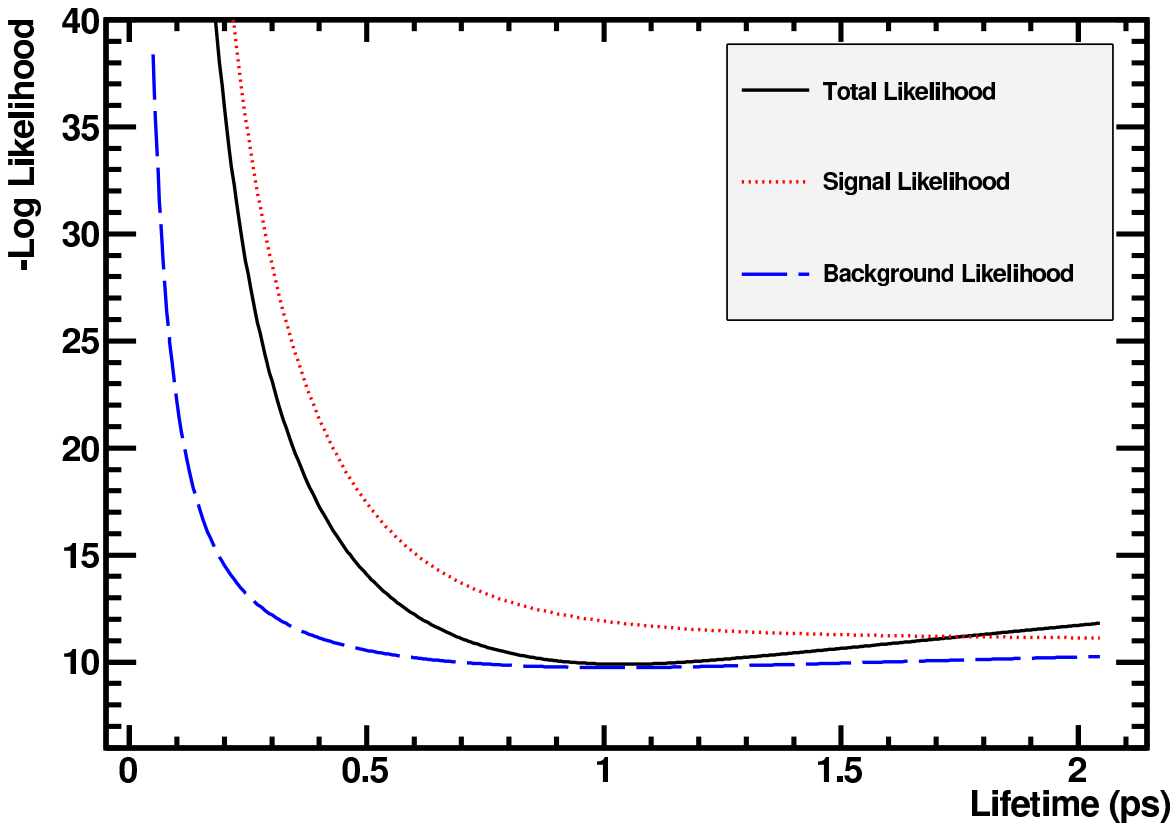


Figure 5: Plot of likelihood components.

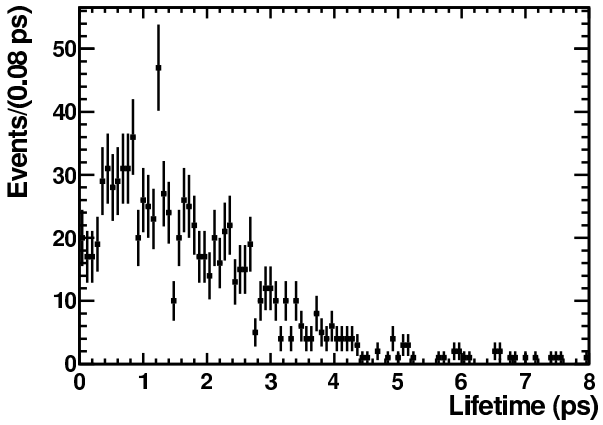


Figure 6: $B_s^0 \to D_s^- \pi^+$ signal lifetime after the trigger.

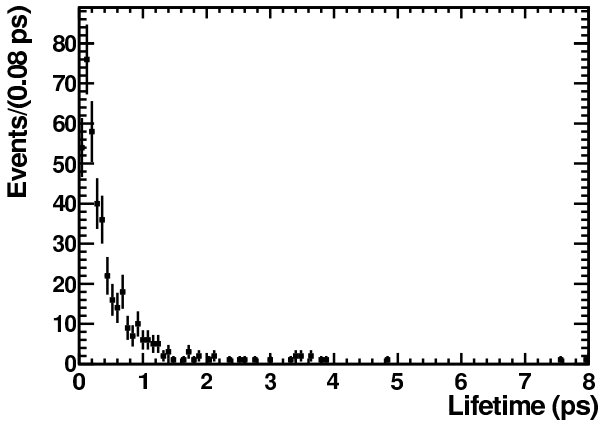


Figure 7: $B_s^0 \to D_s^- \pi^+$ background lifetime after the trigger.

$$\sum_{i=1}^{N_S}\sum_{i'=1}^{N_S}\left(\frac{\partial}{\partial\tau}\delta lnP_{Sig}(t_i;\tau)\right)\left(\frac{\partial}{\partial\tau}\delta lnP_{Sig}(t'_i;\tau)\right) = \tag{15}$$

$$\left\langle\left(\frac{\partial\delta lnP_{Sig}}{\partial\tau}\right)^2\right\rangle_S - \left\langle\frac{\partial\delta lnP_{Sig}}{\partial\tau}\right\rangle_S^2 = N_S{\sigma_S}^{-1}, \tag{16}$$

and for the sum over background

$$\sum_{j=1}^{N_B}\sum_{j'=1}^{N_B}\left(\frac{\partial}{\partial\tau}\delta lnP_{Sig}(t_j;\tau)\right)\left(\frac{\partial}{\partial\tau}\delta lnP_{Sig}(t'_j;\tau)\right) = \tag{17}$$

$$\left\langle\left(\frac{\partial\delta lnP_{Sig}}{\partial\tau}\right)^2\right\rangle_B - \left\langle\frac{\partial\delta lnP_{Sig}}{\partial\tau}\right\rangle_B^2 = N_B{\sigma_B}^{-1}, \tag{18}$$

So from looking at the square of the fluctuations we find the error is given by

$$\left\langle(\tau-\tau_{fit})^2\right\rangle = \frac{1}{N_S}\sigma_S + \frac{n_B}{{N_S}^2}\left(1+\frac{n_B}{N_B}\right)\frac{{\sigma_S}^2}{\sigma_B}. \tag{19}$$

7.3 Results

7.3.1 Determination of n_B from mass fit

The term $\frac{n_B}{N_B}$ in the likelihood gives the weight of the background subtraction. N_B is the number of background events in the sideband, which can be obtained simply by counting. n_B is the number of background events under the mass peak. This number is more difficult to calculate as there is both signal and background in the peak, and we cannot know on an event-by-event basis to which category a particular event belongs. To obtain n_B we fit to the mass distribution for the signal fraction f_{sig}, and $(1-f_{sig})\times N_{\rm events}$ then gives the total number of background events, which along with the background description used in the mass fit gives us an estimate of n_B. It follows that if the fit result for f_{sig} is biased we will obtain an incorrect estimate of n_B, which will also bias the lifetime result from the likelihood subtraction. It is important then that the mass fit, or whichever method is used to find n_B gives an unbiased result. It is also important that the sideband sample contains only background. If the mass of the events does not provide a good discriminator between signal and background, it may be the case that other variables need to be used to obtain a pure background sample in order for the subtraction to work correctly.

7.3.2 Lifetime fit

Once the signal fraction has been determined and the number of background events estimated, the lifetime fit can proceed.

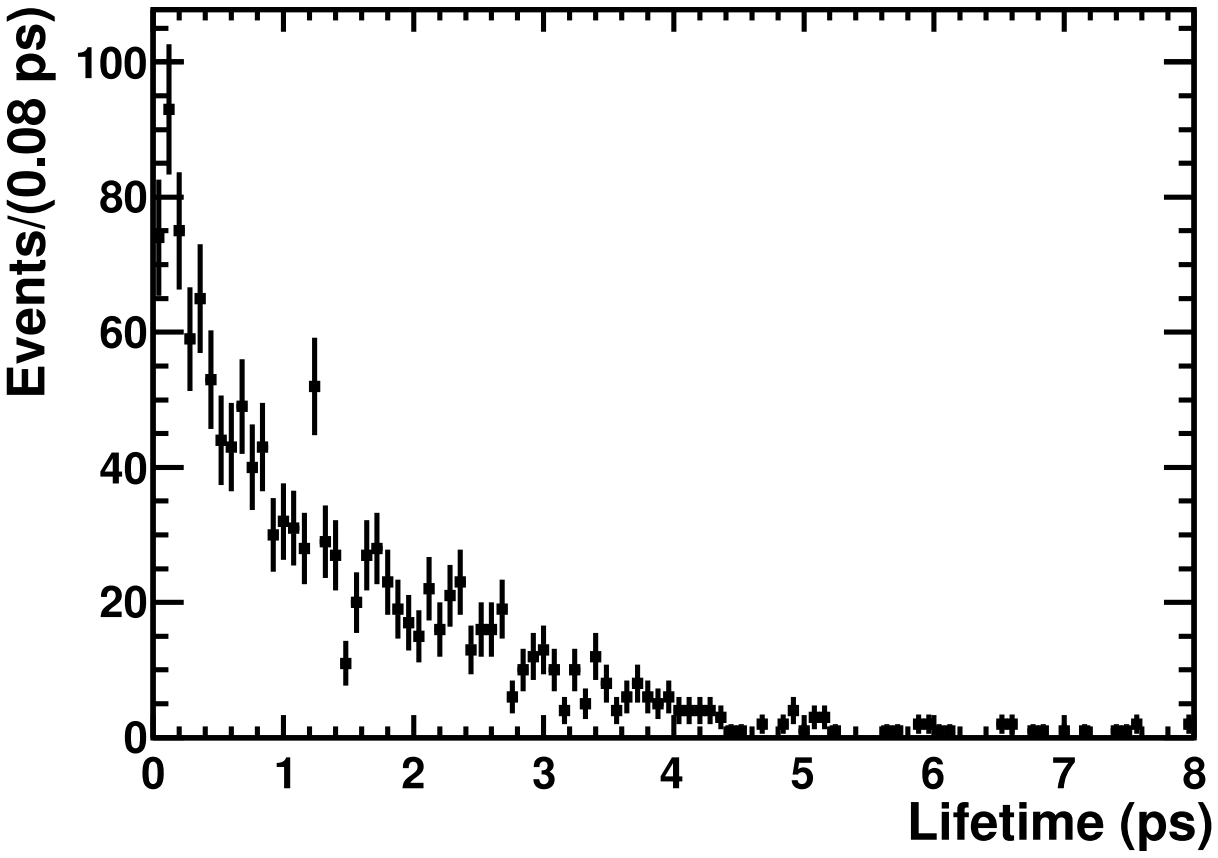


Figure 8: $B_s^0 \to D_s^- \pi^+$ signal + background lifetime after the trigger.

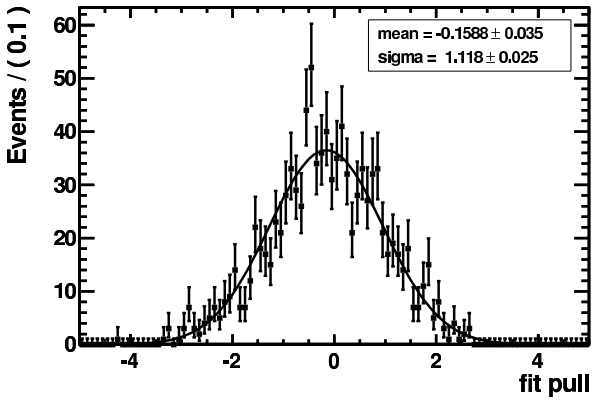


Figure 9: Pull distributions with the uncorrected error.

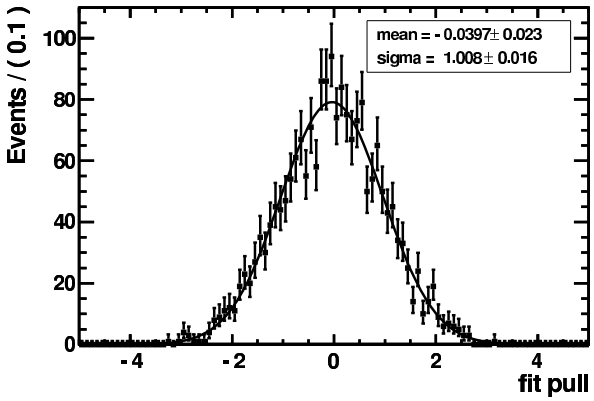


Figure 10: Pull distributions with the corrected error.

Figures 6 and 7 show the signal and background lifetimes after the trigger, and Figure 8 shows the sum of both. Although the signal and background do not of course have an identical lifetime, they can both be modelled with the same type of distribution: exponential lifetimes with event-by-event acceptance effects and Gaussian errors. It is then possible to fit both signal and background with the same function, as they are characterised by different values of the same parameters. This is important as background events that are extremely unlikely (probability close to zero) have large logarithms, contributing heavily to the likelihood. When the signal likelihood is applied to the background events in the correction, this can result in the subtraction term becoming too large. If the signal model provides a poor description of the background, these contributions will dominate and the fit will likely fail.

The fit returns a lifetime of 1.47ps $\pm$ 0.15ps. The generator value for the B lifetime is 1.534ps.

7.3.3 Pull Studies

A pull study was performed with 1000 toy Monte Carlo simulation experiments, each consisting of 10000 events, to check for any bias in the fit result or the corresponding error. The toys included both signal and background events, with a generated lifetime of 1.5ps for signal and 0.5ps for background. For simplicity only one acceptance interval was generated for each event. The position and duration of an acceptance window were normally distributed, with different mean values for signal and background. The pull distributions are shown in Figures 9 and 10, with the uncorrected error, and the corrected error respectively.

7.3.4 Effect of biases in the signal fraction

We recall that the fit proceeds in two stages, firstly the mass is fitted to find the signal fraction, and then the lifetime is fitted using the mass fit result to weight the background subtraction in the likelihood. One question that arises with this method is: to what extent does a bias in the signal fraction have on the fit? To answer this question a toy study was performed with 20 pulls, each consisting of 500 toy experiments of 10000 events each. The signal fraction was set by hand in each toy to $n\sigma$ from the true value, with n taking integer values between -10 and 10, and σ being the error on the mass fit result.

Figures 11 and 12 show the resulting distributions for the mean and width of the pulls versus the offset in the signal fraction. The plot of the pull mean shows that there is a negative correlation between the signal fraction offset and the fit bias. The reason for this becomes clear when we consider the term in the likelihood that weights the background subtraction, n_B/N_B. If the signal fraction is too high, we underestimated n_B and the magnitude of the subtraction is too small. A background effect then remains that biases the fit towards lower lifetimes. Conversely, if the signal fraction is too low, we overestimate n_B and the magnitude of the subtraction is too large. We then overcompensate for the background, leading to a bias towards longer lifetimes. The plot of the pull width appears largely uncorrelated with the signal fraction offset.

8 Lifetime fits for multi-signal environments

8.1 A fitting model for multi-signal environments

In order to maximise the use of information present in each event a fitter has been designed that combines the lifetime measurements with those of other variables which can be used to distinguish signal and background

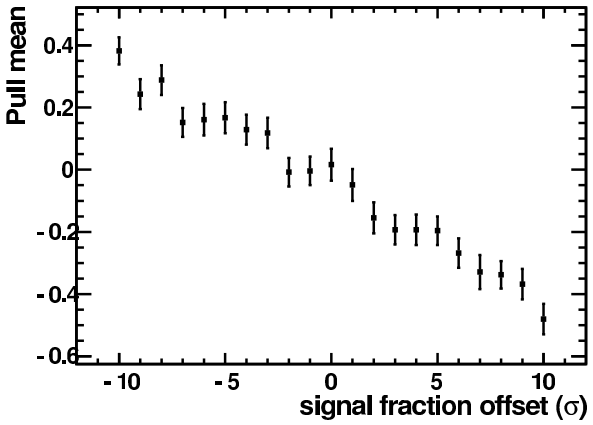


Figure 11: Pull mean as a function of the signal fraction offset.

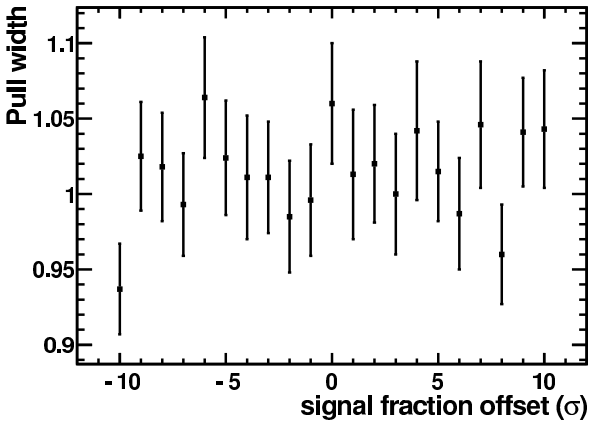


Figure 12: Pull width as a function of the signal fraction offset.

channels. The most likely candidates for these discriminating variables are invariant mass and PID but also the use of kinematical variables can be envisaged.

For several use cases it will be advantageous to be able to fit several signal channels at once. From a theoretical point of view lifetime ratios are of great importance, so to fit them directly is desirable. In addition, environments like $B_{(s)} \to h^+h^-$ decays have a number of signal channels that (partially) overlap in mass and PID which motivates a simultaneous fit of all components.

The full probability density for each event can be written as

$$f(t, \mathbf{X}) = f(t|\mathbf{X}) \times f(\mathbf{X}), \tag{20}$$

with the measured lifetime t and $\mathbf{X}$ a vector of discriminating variables. The total probability density is split in a lifetime probability density that is expressed under the condition of a set of discriminating variables and the probability density of observing this set of variables. To avoid the effect of correlations between the two components, the actual fit is split in two steps. First, the relative fractions of the various signal and background components are determined by fitting $f(\mathbf{X})$. Thereafter, using these fractions as input, the fit of $f(t|\mathbf{X})$ is used to obtain the lifetime results.

8.1.1 Fitting the signal fractions

The total probability density of observing a certain set of variables that distinguish different event classes can be written as

$$f(\mathbf{X}) = \sum_{class} f(\mathbf{X}|class) \times P(class), \tag{21}$$

where $f(\mathbf{X}|class)$ is the PDF of $\mathbf{X}$ for a given class and $P(class)$ is the probability of an event being of that class, i.e. $P(class) = N_{class}/N_{tot}$ which will be a fit parameter. Then the signal or background probability for any event is given by

$$P(class|\mathbf{X}) = \frac{f(\mathbf{X}|class) \times P(class)}{f(\mathbf{X})}. \tag{22}$$

This definition requires models for $f(\mathbf{X}|class)$, i.e. the PDF for any discriminating variable to be used for any class of events present in the fit. The main discriminating variable will always be the invariant mass of the reconstructed daughters of a certain decay. In this case, a Gaussian (or double-Gaussian or Crystal Ball function depending on the decay) can be used for the signal peak while an exponential or linear function can be used for the combinatoric background. The result of the fit of the signal fractions is automatically fed into the lifetime fit.

8.1.2 Fitting the lifetimes

The lifetime part can be split up in signal and background components

$$f(t|\mathbf{X}) = \sum_{class} f(t|class) \times P(class|\mathbf{X}), \tag{23}$$

where $f(t|class)$ are the different lifetime models for signal(s) and background(s), while $P(class|\mathbf{X})$ determines the probability of this event belonging to a certain class as described in Section 8.1.1.

The various lifetime models are defined as follows. The signal lifetime model is defined as an exponential with the mean lifetime τ as a free parameter.

$$f(t|sig) = N \times \frac{1}{\tau} e^{-t/\tau}. \tag{24}$$

The normalisation of the signal lifetime model contains the information about the event by event acceptance function as determined by the HLT swimming method. It is essentially the integral over the exponential function inside the accepted time intervals. The same lifetime model can be used for other signal channels or known specific background modes.

For background from combinatoric events it is very difficult to estimate a proper model parametrisation for the lifetime probability density. Therefore, an approach has been developed that does not assume the shape of this probability density function (p.d.f.). For each event, a so-called kernel function is used to represent the measured lifetime. This approach follows earlier implementations of this technique at LEP [16] and has been first introduced to LHCb to model the forward-backward asymmetry in $B_d^0 \to K^* \mu\mu$ decays [17]. The sum of all kernel functions, weighted by the event's probability of being background, then describes the time probability density for background events. Hence, for a pure background sample the time p.d.f. would be computed as

$$f(t|bg) = \frac{\sum_i Gauss(t, t_i, \sigma) \times P(bg|X_i)}{\underbrace{\sum_i P(bg|X_i)}_{N_{bg}}}. \tag{25}$$

Here, a Gaussian is used as a kernel function. Its width must not be confused with the measurement error; it is mainly a parameter to achieve a smooth distribution for the p.d.f. In the current implementation, the width of the kernel functions increases linearly with the measured lifetime to account for the smaller number of large lifetime events.

In case of a data sample that comprises both signal and background events (as is usually the case) the sum in Equation 25 would run over the full sample and hence the background p.d.f. would get a contribution from signal events. To avoid this the signal contribution is subtracted from the total p.d.f. using the known shape of the signal p.d.f. to yield a pure background p.d.f. The resulting function is

$$f(t|bg) = \frac{\sum_i (Gauss(t, t_i, \sigma) - \sum_{class \neq bg} f(t|class) \times P(class|X_i))}{\sum_i P(bg|X_i)}. \tag{26}$$

To enhance the contribution of regions with a cleaner background sample in the computation of the background time p.d.f. the probability for each event of being background is used as a weight. The resulting formula is

$$f(t|bg) = \frac{\sum_i ([Gauss(t, t_i, \sigma) - \sum_{class \neq bg} f(t|class) \times P(class|X_i)] \times P(bg|X_i))}{\sum_i P(bg|X_i)^2}. \tag{27}$$

A challenge this method presents originates in the subtraction of probability densities when evaluating the background p.d.f. For vastly wrong signal lifetimes this can lead to negative values in certain regions of lifetime (see Figure 13). However, even when using the correct signal lifetime negative values are not excluded. They can occur due to statistical fluctuations that may not be fully smoothed out by the kernel approach in regions of lifetime where the signal fraction is high and hence the subtraction naturally produces values close to zero.

As no lifetime region is physically excluded for background events, setting the background p.d.f. to zero if the subtraction yields negative values is not an option. Instead, if the resulting background p.d.f. has a fraction of the total p.d.f. of less than a small value ϵ, all possible fractions between ϵ and $-\infty$ are continuously mapped to be inside the interval $[\epsilon, 0)$. The mapping uses the diverging behaviour of the tangent function to uniquely assign a positive value to every value between ϵ and $-\infty$. Using this method a stable behaviour of the fitter is obtained.

Since the value of the kernel functions of all events has to be evaluated for each event this results in a quadratic time dependence of the fit on the number of events. Therefore, a lookup table has been introduced which evaluates the kernel functions at discrete points before the start of the fit. A non-equidistant binning is used to allow an efficient coverage of a time range up to about 500 times the lifetime of a B meson. This reduces the time dependence to become linear on the number of events.

8.2 Toy Monte Carlo Simulation Studies

To evaluate the performance of this fitter a range of toy Monte Carlo simulation studies has been performed. It was attempted to reproduce the distributions as obtained from the full LHCb simulation as well as possible.

The generated time distributions for signal and background events follow an exponential convoluted with a Gaussian resolution function. They also simulate a lifetime bias by applying an acceptance function with a rising edge at positive lifetimes. The signal masses are simulated as single Gaussians, whereas the background has a flat distribution in mass.

The full fit has been tested with 1000 toy experiments of 15000 events each. The fraction of each of the two signal classes was simulated to be 1/3, corresponding to 5000 events on average. One signal was simulated with a lifetime of 1.5 ps and the second signal with a lifetime ratio of 0.9, hence a lifetime of 1.35 ps. The background was simulated with on average 90% of the events having a lifetime of 0.5 ps and 10% of the events having a lifetime of 10 ps. The two signal mass peaks were simulated to have a resolution of 25 MeV/c^2 at 5280 MeV/c^2 and 5360 MeV/c^2, respectively. The mass range for the background was simulated flat between 5000 MeV/c^2 and 5800 MeV/c^2.

The signal fractions are obtained in an initial fit and have been tested to be unbiased. They are fixed as input to the lifetime fit. The fit results for all lifetime fits are shown in Figure 14. The fit of the lifetime of the first signal returned an average of 1.502 ps with an average error of 0.024 ps. The ratio of the lifetimes of the second signal w.r.t. the first was on average measured to be 0.900 with an average error of 0.021. This yields pulls with means of 0.08 ± 0.03 and -0.01 ± 0.03, for the two quantities respectively, and widths of 0.98 ± 0.02 for both parameters.

The correlation was on average -0.738 with a spread of 0.005. The high correlation can be understood as the lifetime ratio has contributions both from the second independent lifetime and from the first lifetime, i.e. the other fit parameter. As the lifetime ratio is the value with a higher theoretical interest, its fit has been preferred to fitting two lifetimes.

9 Conclusion

We have presented a set of innovative methods and tools for precision lifetime and lifetime-difference measurements in hadronic B decays at LHCb. All methods are purely data-driven and Monte Carlo (MC) simulation independent, a particularly important feature if lifetime measurements are to be made in the early period of LHCb's data taking. We demonstrated that these methods and tools work in detailed simulation studies. In particular:

- We introduced a method for an MC-independent correction of the trigger-induced bias in lifetime measurements using fully hadronic final states. We have developed publicly available tools that implement this method in the LHCb software environment and demonstrated their functionality.
- The likelihood subtraction technique presented is not restricted to lifetime measurements. It removes the need to model the background distribution in unbinned fits and resolves complications such as the so-called Punzi effect [14]. We showed how parameter errors in this method can be calculated directly from data, without recourse to simulations or background models.
- We introduced a kernel-based method for fits in multi-signal environments such as $B \to hh$. This method does not rely on a background model, and is designed to deal with cases where no pure background sample exists, a situation that constitutes a problem for many other methods including likelihood subtraction.

10 Acknowledgments

This study and note benefited from a great deal of help from our colleagues. Particularly, we would like to thank Jose Hernando Morata for helping implement the swimming method in the trigger, Chris Parkes for helpful discussions about the fitting method in multi-signal environments, and Tim Gershon, Vincenzo Vagnoni, and Ulrich Kerzel whose comments greatly improved the substance of this note.

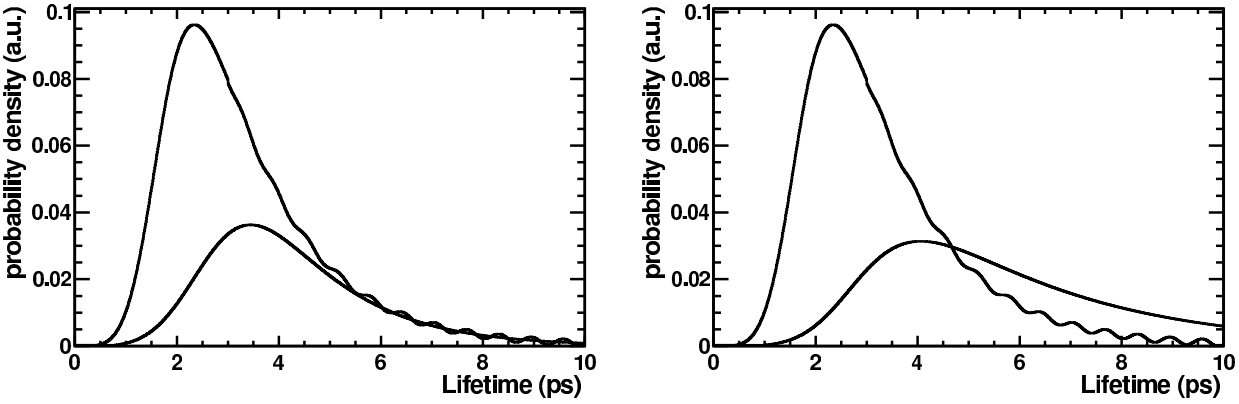


Figure 13: Example for situations resulting in negative values for the background pdf when subtracting the signal contribution (red, dotted) from the total distribution (black) due to statistical fluctuations (left) or wrong fit values of the signal lifetime (right). The label "a.u." stands for arbitrary units.

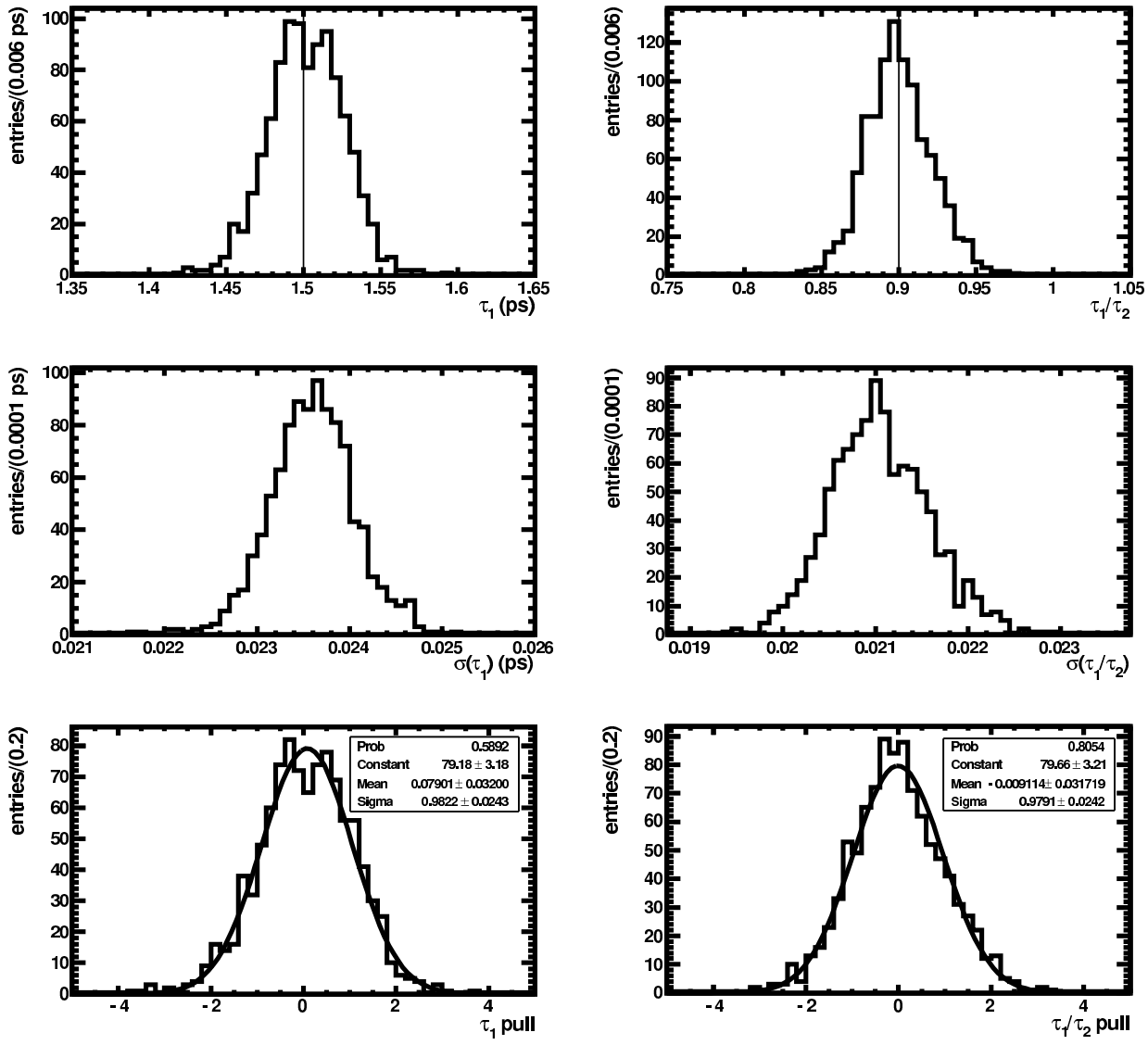


Figure 14: Fit values (top, line marks input value), their errors (middle), and pulls (bottom) for lifetime of first signal (τ_1, left) and lifetime ratio (τ_1/τ_2, right).

References

[1] Sneha Malde. Simulation free measurement of the B+ lifetime using decays selected using displaced tracks. FERMILAB-THESIS-2009-06.

[2] V. V. Gligorov and J Rademacker. Monte Carlo Independent Lifetime Fitting at LHCb in Lifetime Biased Channels. Technical Report CERN-LHCb-2007-053, 2007.

[3] V. V. Gligorov. Measurement of the CKM angle γ and B meson lifetimes at the LHCb detector. PhD thesis, 2008. CERN-THESIS-2008-44.

[4] J. Rademacker. Reduction of statistical power per event due to upper lifetime cuts in lifetime measurements. *Nucl. Instrum. Meth.*, A570:525, 2007.

[5] M. Shifman. Quark-hadron duality. Technical Report TPI-MINN-00/44, hep-ph/0009131, 2000.

[6] N. Uraltsev. Heavy-quark expansion in beauty and its decays. Technical Report UND-HEP-98-BIG1, hep-ph/9804275, 1998.

[7] A. Petrov. Lifetimes of heavy hadrons. Technical Report WSU-HEP-0407, hep-ph/0408093, 2004.

[8] F. Gabbiani, A. Onishchenko, and A. Petrov. Spectator effects and lifetimes of heavy hadrons. Technical report, 2004.

[9] P Clarke, C McLean, and A Osorio-Oliveros. Sensitivity studies to β_s and $\Delta\Gamma_s$ using the full $B_s \to J/\psi\phi$ angular analysis at LHCb. Technical Report CERN-LHCb-2007-101, Jul 2007.

[10] S. Cohen, M. Merk, and E. Rodrigues. $\gamma + \phi_s$ sensitivity studies from combined $B_s^0 \to D_s^{\mp}\pi^{\pm}$ and $B_s^0 \to D_s^{\mp}K^{\pm}$ samples at LHCb. Technical Report CERN-LHCb-2007-041, 2007.

[11] S Amato, J McCarron, F Muheim, B Souza de Paula, and Y Xie. Lhcb's sensitivity to new cp-violating phases in the decay $b_s \to \phi\phi$. Technical Report LHCb-2007-047. CERN-LHCb-2007-047, CERN, Geneva, May 2007.

[12] E. Barberio et al. Averages of b-hadron and c-hadron Properties at the End of 2007. 2008. hep-ex/0808.1297.

[13] A. Lenz. Theoretical status of B_s-mixing and lifetimes of heavy hadrons. *Nucl. Phys. Proc. Suppl.*, 177-178:81–86, 2008.

[14] G. Punzi. Comments on Likelihood fits with variable resolution. Number physics/0401045 in Phystat 2003. SLAC, 2003.

[15] The LHCb Collaboration. LHCb Trigger System Technical Design Report. Technical report, 2003. CERN-LHCC-2003-031.

[16] K.S. Cranmer. Kernel estimation in high-energy physics. *Comput. Phys. Commun.*, 136:198–207, 2001.

[17] F. Marinho. A non-parametric method to estimate the Forward-Backward Asymmetry from the $B_d \to K^*\mu^+\mu^-$. Technical Report LHCb-2009-004. CERN-LHCb-2009-004, CERN, Geneva, Jan 2009.