CDF/ANAL/BOTTOM/CDFR/7867
September 26, 2005
Version 1.0

Combined Measurement of the Λ_b^0 Lifetime in $\Lambda_b^0 \to J/\psi\Lambda^0$

Konstantin Anikeev, Jonathan Lewis, Pat Lukens, Robyn Madrak, Ting Miao, Rick Tesarek, Slawek Tkaczyk

Fermi National Laboratory

Luis Labarga Echeverría, Juan Pablo Fernández

Universidad Autónoma de Madrid/ CIEMAT

Elliot Lipeles, Mark Neubauer, Frank Würthwein

University of California, San Diego

John Alison, Joe Boudreau, Chunlei Liu

University of Pittsburgh

Abstract

In this note, we present a blinded measurement of the Λ_b^0 lifetime in the exclusive decay channel $\Lambda_b^0 \to J/\psi\Lambda^0$ with $\Lambda^0 \to p\pi^-$ where we expect the uncertainties to be 9.1% statistical and $^{+3.5\%}_{-2.3\%}$ systematic. We also present a measurement of the B^0 lifetime in $B^0 \to J/\psi K_s^0$ with $K_s^0 \to \pi^+\pi^-$, as a reference mode, finding $c\tau = 450.6\ ^{+15.0}_{-14.5}$(stat.) $^{+13.1}_{-6.3}$(syst.)μm. The analysis is based upon $\sim 370 pb^{-1}$ of luminosity collected up to the August 2004 shutdown.

Contents

1 Introduction

In b-flavored hadron decays, all b-hadrons have the same lifetime via weak transition $b \to Wq$ ($q = c, u$) of the b-quark if one considers the other quarks comprising the hadron a mere idle spectators. In actuality, non-trivial spectator quark effects lead to differences in the lifetime of b-hadrons, in disagreement with the naive spectator quark model. Lifetime measurements of b-hadrons are now precise enough to demonstrate a clear hierarchy of b-hadron lifetimes that can be used to quantitatively test theoretical predictions.

Currently, the hierarchy of heavy hadron lifetimes are calculated in the framework of the Heavy Quark Expansion (HQE). The HQE expresses the decay rate of a heavy hadron into a final state as an (operator) expansion in both $\Lambda_{\rm QCD}/m_b$ and $\alpha_s(m_b)$. The results of these calculations are in good quantitative agreement with observed hierarchy of b-meson lifetimes and in qualitative agreement with charmed hadron lifetime measurements. The only remaining problem has been in the ratios of heavy meson and baryon lifetimes. Early theoretical predictions for $\tau(\Lambda_{\rm b}^0)/\tau({\rm B}^0)$ were higher than the observed ratio by approximately 2σ, suggesting a discrepancy that might be exposing a problem with the calculations.

Recent next-to-leading order (NLO) calculations of perturbation QCD including both $1/m_b$ and sub-leading $1/m_b$ corrections to spectator effects have significantly reduced the $\tau(\Lambda_{\rm b}^0)/\tau({\rm B}^0)$ discrepancy. These calculations [1] yield

$$\tau(\Lambda_{\rm b}^0)/\tau({\rm B}^0) = 0.86 \pm 0.05$$

while the current world average lifetime ratio from the Heavy Flavor Averaging Group (HFAG) [2] is

$$\tau(\Lambda_{\rm b}^0)/\tau({\rm B}^0) = 0.803 \pm 0.047,$$

which are consistent at the 0.8σ level.

In this note, we present measurements of the $\Lambda_{\rm b}^0$ and ${\rm B}^0$ lifetime in the exclusively reconstructed decay modes $\Lambda_{\rm b}^0 \to {\rm J}/\psi\Lambda^0$ and ${\rm B}^0 \to {\rm J}/\psi{\rm K}_{\rm s}^0$, with ${\rm K}_{\rm s}^0 \to \pi^+\pi^-$, $\Lambda^0 \to {\rm p}\pi^-$, and ${\rm J}/\psi \to \mu^+\mu^-$. Auxiliary measurements of the ${\rm B}^+$ lifetime and ${\rm B}^0$ lifetime in other modes are also presented. The basic strategy pursued in using all of these measurements to arrive at the relevant physics quantity of interest, the lifetime ratio between $\Lambda_{\rm b}^0$ and ${\rm B}^0$ as discussed in Section 2. This work builds upon previous work documented in [3, 4, 5, 6, 7].

Lifetime measurements in exclusively reconstructed decays allow for precise measurement of the $\beta\gamma$ factor used to convert the measured decay length into a proper decay length. It is worth pointing out that the PDG average for the $\Lambda_{\rm b}^0$ lifetime is dominated entirely by measurements for which the $\Lambda_{\rm b}^0$ decay is not fully reconstructed. We believe our measurement, as well as a similar measurement by D0[8], add new and unique information about the $\Lambda_{\rm b}^0$ lifetime even if in itself is not yet competitive with the PDG world average.

The document is organized as follows. Section 2 discusses the overall analysis strategy, thus providing a roadmap for the remainder of the document. Section 3 describes our candidate selection. Sections 4 and 5 describe the fitting procedure as well as fit results. Section 6 details a long list of systematics checks, introduced by an introduction, and summarized in Section 6.11. Section 7 completes the paper with a summary and conclusion, remaining work to be done.

2 Analysis Strategy

The physics goal of this analysis is the best possible measurement of the lifetime ratio $\tau(\Lambda_b^0)/\tau(B^0)$, as well as the Λ_b^0 lifetime itself. We choose the final states $\Lambda_b^0 \to J/\psi\Lambda^0$ and $B^0 \to J/\psi K_s^0$ because they are fully reconstructed decays which allow for complete determination of the proper decay length and have very similar decay topology ($J/\psi + V^0$) so that many detector systematics cancel in the ratio. In addition, the selection makes use of the dimuon trigger which simplifies the lifetime analysis compared to the hadronic trigger because there are no impact parameter requirements at the trigger level.

In practice, a CDF measurement of the B^0 lifetime is at present not competitive with the HFAG 2004 world average of $\tau = 1.534 \pm 0.013$ ps ($c\tau = 460 \pm 4\mu$m) due to limited statistics. A lifetime ratio determination based on CDF measurements alone is thus not likely to be the best way of using the present data, as we expect to be statistics rather than systematics limited. The fact that systematics cancel in the ratio of CDF measurements alone is thus merely of academic interest at this point. The most important physics result is the Λ_b^0 lifetime measurement.

We use the $B^0 \to J/\psi K_s^0$ decay as control mode to validate the analysis procedure, as well as to study the majority of systematic errors. To avoid any statistical bias on the lifetime measurement, we kept the Λ_b^0 blind from a lifetime perspective until the analysis procedure was set.

There are two separate kinds of backgrounds in our sample. Prompt background consists of a J/ψ and a V^0 produced at the collision point. Long-lived backgrounds come from a J/ψ from a b-hadron decay and a V^0 either from a b-hadron in the event or from the primary vertex. In Section 6.2 we discuss the degree to which the origin of the V^0 can be determined in data from vertexing information. These two backgrounds effect this analysis very differently. Prompt background has a very different proper decay length distribution than the signal. Long-lived background on the other hand looks very similar to the signal because the J/ψ comes from a b-hadron and all b-hadrons have similar decay lengths.

Optimization of our sample selection is motivated by toy Monte Carlo studies of how the statistical error on the fitted lifetime depends upon background size and composition. The detailed selection is then optimized using BGenerator Monte Carlo samples for both $B^0 \to J/\psi K_s^0$ and $\Lambda_b^0 \to J/\psi\Lambda^0$ signals and "far" sidebands for the background. Because

we use the sidebands to determine the level and shape of the background in our lifetime fit, we define "far" sidebands which are outside of the mass region used for the fit to avoid any potential statistical bias on the lifetime. When feasible we use the $\Lambda_b^0 \to J/\psi\Lambda^0$ optimized cuts for both signal modes so that the reference mode is treated in the same way as the signal.

Our fitting procedure is the extract the signal lifetime using a maximum likelihood fit to the proper decay length and mass event information. We validate our lifetime fitting procedure on both BGenerator signal Monte Carlo and Pythia $b\bar{b}$ Monte Carlo sample where the b-quark decay leads to $J/\psi \to \mu^+\mu^-$ in the final state. Both samples are fully simulated using the standard run-dependent MC production procedure [9]. In both cases, the identical reconstruction code and selection applied to the data is employed. The signal MC validates that fitting for the lifetime of realistic simulation and reconstruction of $B^0 \to J/\psi K_s^0$ and $\Lambda_b^0 \to J/\psi\Lambda^0$ decays yields results consistent with the generated lifetimes. Fits to reconstructed Pythia $b \to J/\psi X$ MC validate that the signal lifetimes can be fit in the presence of long-lived backgrounds very similar to that in the data.

Once our fitting procedure has been successfully applied to realistic Monte Carlo, we fit the $J/\psi \to \mu^+\mu^-$ triggered data reconstructed as $B^0 \to J/\psi K_s^0$ for lifetime. We verify that our lifetime fitter with model parameters determined from the fit are statistically unbiased by generating toy datasets of same size as the data and verifying that the $c\tau$ fit residual is consistent with a unit Gaussian. The systematics on our B^0 lifetime measurement are determined and numerous "sanity checks" are performed. This procedure is then repeated for the Λ_b^0 lifetime.

We use the lifetimes of other b-meson modes to constrain our knowledge of the overall distance scale of the CDF detector by comparing our measured lifetimes with the HFAG world averages for these b-meson decays. These auxiliary modes include

$B^0 \to \psi(2S)K_s^0$, with $\psi(2S) \to \mu^+\mu^-$, $K_s^0 \to \pi^+\pi^-$
$B^0 \to \psi(2S)K_s^0$, with $\psi(2S) \to J/\psi\pi^+\pi^-$, $J/\psi \to \mu^+\mu^-$, $K_s^0 \to \pi^+\pi^-$
$B^0 \to J/\psi K^{*0}$, with $J/\psi \to \mu^+\mu^-$, $K^{*0} \to K^+\pi^-$
$B^0 \to \psi(2S)K^{*0}$, with $\psi(2S) \to \mu^+\mu^-$, $K^{*0} \to K^+\pi^-$
$B^0 \to \psi(2S)K^{*0}$, with $\psi(2S) \to J/\psi\pi^+\pi^-$, $J/\psi \to \mu^+\mu^-$, $K^{*0} \to K^+\pi^-$
$B^+ \to J/\psi K^+$, with $J/\psi \to \mu^+\mu^-$
$B^+ \to \psi(2S)K^+$, with $\psi(2S) \to \mu^+\mu^-$
$B^+ \to \psi(2S)K^+$, with $\psi(2S) \to J/\psi\pi^+\pi^-$, $J/\psi \to \mu^+\mu^-$
$B^+ \to J/\psi K^{*+}$, with $J/\psi \to \mu^+\mu^-$, $K^{*+} \to K_s^0\pi^+$

In order to avoid systematic errors that are blown up by limited statistics of control samples, we distinguish two types of systematic error investigations. There are those checks for which we know before doing them that they can not possibly result in a statistically meaningful conclusion because they depend on the statistics of the data sample itself. We refer to such tests as "feel good" tests, and report on them here for completeness.

The second type of tests are those that were done to assess a quantitative systematic error on our procedures. For those tests, we estimate beforehand the precision of the systematic test. We thus make no a posteriori choices when deciding on what enters the systematic error.

Our strategy was to set a threshold of 3 standard deviations (σ) for feel good tests. If a feel good test deviated by more than 3σ, we consider this a serious issue that requires additional study. The choice of this three sigma guideline is of course arbitrary. It is motivated by the desire not to be too distracted by arbitrary fluctuations. We are looking at serveral tens of tests, many of which are on variables which should not influence the lifetime if life was just.

3 Candidate Selection

In this section, we describe our b-hadron reconstruction procedure and candidate selection. This selection is then applied to the $J/\psi \to \mu^+\mu^-$ triggered data to determine the signal yields. Comparisons of sideband-subtracted signal distributions with our BGenerator signal MC are also presented.

3.1 Dataset

The dataset used in this analysis is xpmm0d, which is the compressed data from the 5.1 production pass of dimuon triggers rich in $J/\psi \to \mu^+\mu^-$ and $\psi(2S)$ decays. This data spans the period of February 4, 2002 (run 138425) to August 22, 2004 (run 186598) which is the run before the Summer 2004 shutdown. According to the Data File Catalog (DFC), the total offline luminosity for this dataset is 496 pb^{-1}, however not all runs had an operational SVX detector that we require for good beamline determination so that the actual luminosity of runs passing our good run selection will be considerably lower than this.

3.2 Calibrations

In all cases, calibration pass 13A is used in the analysis. This pass is based upon 5.3.1 silicon alignment tables and the appropriate calibration pass for Gen5 physics analyses (pre-5.3.4). It is also the same calibration pass used in production of xpmm0d, up to differences in calorimeter and TOF calibration irrelevant for this analysis.

3.3 Track Refitting

All tracks are refit according to the procedure [10] blessed by the BPAK group for 5.3.3-based analyses. This involves using a Kalman fitter within the TrackRefitter class to correct for energy losses resulting from different particles traversing detector material.

The track covariance matrix scale factors were taken directly from [10] as well as the B-field magnitude of 1.41296 T calibrated on J/ψ decays. No L00 silicon information is used in this analysis and any L00 hits are dropped before tracks are refit.

Since K^0_s and Λ^0's have long lifetimes ($c\tau$ = 2.7 and 7.9 cm, respectively), their decay particles can traverse a significantly different amount of material depending upon where the decay occurs, unlike most other tracks that originate inside the beampipe. To properly account for effect on the energy loss, we use the setInnerRadius() method of the TrackRefitter class to ignore hits inside of the vertex radius and set the starting point of the refit and material integration.

3.4 Muon Selection

We select `defMuons` which are CMU, CMP, CMUP, or CMX (muon type ≤ 7). For muons with CMU or CMX information, we require that χ^2_x for the muon stub position is less than 9. This cut is 99% efficient for CMX muons [11].

3.5 Good runs

For good run selection, we use a slightly modified version of the sql script provided by the BPAK group [12]. The only difference between our selection and the standard selection is that we remove the requirement on `CAL_OFFLINE` and `SVT_STATUS`. There is no reason to make any requirement on these bits, since our analysis does not use calorimeter data or SVT. With this good run selection, we obtain an offline integrated luminosity of 370.9 $\pm$ 21.8 pb^{-1}.

```
SELECT RUNNUMBER, sum(LUM_INTEGRAL_OFFLINE),  sum(LUM_INTEGRAL_ONLINE)
FROM Run_Status, FILECATALOG.CDF2_RUNSECTIONS
WHERE
Run_Status.RUNNUMBER = FILECATALOG.CDF2_RUNSECTIONS.RUN_NUMBER
AND
Run_Status.RUNCONTROL_STATUS = 1
AND
Run_Status.SHIFTCREW_STATUS = 1
AND
Run_Status.CLC_STATUS = 1
AND
Run_Status.L1T_STATUS = 1
AND
Run_Status.L2T_STATUS = 1
AND
Run_Status.L3T_STATUS = 1
AND
Run_Status.COT_OFFLINE = 1
```

```
AND
(Run_Status.SVX_OFFLINE = 1 OR
((Run_Status.SVX_OFFLINE Is Null) AND Run_Status.SVX_STATUS = 1))
AND
(Run_Status.CMU_OFFLINE = 1 OR
((Run_Status.CMU_OFFLINE Is Null) AND Run_Status.CMU_STATUS = 1))
AND
(Run_Status.CMP_OFFLINE = 1 OR
((Run_Status.CMP_OFFLINE Is Null) AND Run_Status.CMP_STATUS = 1))
AND
(RUNNUMBER<=179056 OR RUNNUMBER>=182843 OR (RUNNUMBER>=180954 AND RUNNUMBER<=181190))
AND
(RUNNUMBER<184062 OR RUNNUMBER>184208)
AND
RUNNUMBER <=186598
GROUP BY RUNNUMBER
ORDER BY RUNNUMBER ASC
/
QUIT
```

3.6 Reconstruction Procedure

All the tracks used for both the J/ψ and the V^0 are required to satisfy the following basic requirements:

- A helix fit and a physical error.
- At least five hits in each of two axial and two stereo superlayers.

3.6.1 J/ψ and ψ(2S)

Muon candidates are derived from the defMuon bank and are required to have at least 3 $r-\phi$ silicon (SVXII + ISL) hits and $p_T > 1.5$ GeV each. The tracks are refit according to the procedure described in Section 3.3.

For $J/\psi \to \mu^+\mu^-$ and $\psi(2S) \to \mu^+\mu^-$ candidates, pairs of opposite charged muons are fit to a common vertex in 3D[1] using CTVMFT [13]. We require the fit to converge and Prob(χ^2) > 0.1%. Candidates are selected if the invariant mass satisfies the following:

$$3.014 \text{ GeV/c}^2 < M_{\mu^+\mu^-} < 3.174 \text{ GeV/c}^2 \qquad (J/\psi \to \mu^+\mu^-)$$

$$3.643 \text{ GeV/c}^2 < M_{\mu^+\mu^-} < 3.723 \text{ GeV/c}^2 \qquad (\psi(2S) \to \mu^+\mu^-)$$

[1] unless otherwise stated, all vertex fit probabilities are 3D χ^2 probabilities

For $\psi(2S) \to J/\psi\pi^+\pi^-$ candidates, we require that a $J/\psi \to \mu^+\mu^-$ candidate make a vertex with two oppositely-charge pions and Prob(χ^2) > 0.0001% with the constraint that $M_{\mu^+\mu^-} = 3.096916$ GeV/c^2. Following the selection described in [15], we also require the di-pion mass to satisfy

$$0.35 \text{ GeV/c}^2 < M_{\pi^+\pi^-} < 0.61 \text{ GeV/c}^2$$

The lower limit is motivated by the know dipion mass for $\psi(2S) \to J/\psi\pi^+\pi^-$ decays and the upper limit is the maximum allowed dipion mass for this decay. Our $\psi(2S) \to J/\psi\pi^+\pi^-$ candidates are selected if the invariant mass satisfies

$$3.676 \text{ GeV/c}^2 < M_{\mu^+\mu^-\pi^+\pi^-} < 3.696 \text{ GeV/c}^2$$

3.6.2 K^0_s and Λ^0

Our V^0 candidates are constructed by first selecting opposite sign combinations of tracks and fitting them to a common vertex using CTVMFT. To reduce the number of track pairs considered for further analysis, we make a loose requirement on the 2D decay length $L_{xy} > 0.10$ cm, which cuts less than 0.1% of true V^0's The tracks for these candidates are then refit using the appropriate mass assignments for K^0_s and Λ, dropping hits inside of the vertex radius according to the procedure described in Section 3.3. The vertices are then refit using the refit tracks with the appropriate mass assignments and requiring Prob(χ^2) > 0.1%. Candidates are selected if the invariant mass satisfies the following:

$$0.473 \text{ GeV/c}^2 < M_{\pi^+\pi^-} < 0.523 \text{ GeV/c}^2 \qquad (K^0_s \to \pi^+\pi^-)$$

$$1.107 \text{ GeV/c}^2 < M_{p\pi^-} < 1.125 \text{ GeV/c}^2 \qquad (\Lambda^0 \to p\pi^-)$$

Two oppositely-charge tracks that make a good vertex at large (> 0.1 cm) L_{xy} are dominated by real K^0_s and Λ^0 decays. We can greatly reduce cross-contamination between K^0_s and Λ^0 by vetoing K^0_s candidates that make Λ^0 candidates with appropriate swapped mass assignments for the pion tracks and vice versa.

Some fraction of real K^0_s decays will satisfy our Λ^0 selection criteria and contibute to a valid $\Lambda^0_b \to J/\psi\Lambda^0$ candidate. Similarly, some true Λ^0's will contibute to selected $B^0 \to J/\psi K^0_s$ candidates. We can greatly reduce cross-contamination between K^0_s and Λ^0 by vetoing K^0_s candidates that make Λ^0 candidates with appropriate swapped mass assignments for the daughter tracks and vice versa. Specifically, we veto K^0_s candidates with

$$1.1085 \text{ GeV/c}^2 < M_{(\pi\to p)\pi} < 1.1235 \text{ GeV/c}^2 \qquad (\Lambda^0 \text{ veto})$$

where $\pi \to p$ represents a swap in particle hypothesis of the higher momentum π^+ (π^-) with an anti-proton (proton), and veto Λ^0 candidates with

$$0.48175 \text{ GeV/c}^2 < M_{(p\to\pi)\pi} < 0.51125 \text{ GeV/c}^2 \qquad (K^0_s \text{ veto})$$

where the track associated with the proton in the Λ^0 candidate is swapped with a pion mass hypothesis.

Additional cuts (e.g. p_T, mass window) are made when the K^0_s and Λ^0 candidates are combined with a J/ψ or $\psi(2S)$ to make a b-meson candidate. These additional variables and their cut values are described in Section 3.6.4 and Section 3.7, respectively.

In addition to the K^0_s and Λ^0 candiates just described, a parallel set is generated using COT parent tracks for the V^0 candidates. These are generated by removing the silicon hits from defTracks using the standard tracking procedure: track_util::removeSilicon. Note that these include both the COT parents from outside-in tracks and the found COT tracks from the inside-out procedure.

3.6.3 K^{*0} and K^{*+}

Since a number of our b-hadron decay modes involve K^{*0} and K^{*+} resonances, we briefy describe their reconstruction here.

We construct our $K^{*0} \to K^+\pi^-$ candidates by considering all two track combinations, under $\pi - K$ mass hypothses, that make a common vertex with $\text{Prob}(\chi^2) > 0.1\%$. Additional cuts on K^{*0} p_T and mass are applied with values determined from optimization of the $B^0 \to J/\psi K^{*0}$ mode.

For $K^{*+} \to K^0_s\pi^+$ candidates, we combine a K^0_s candidate with a pion track and require $\text{Prob}(\chi^2) > 0.1\%$. We use identical K^0_s selection in this mode as is determined from optimization on the $B^0 \to J/\psi K^0_s$ mode. We also use the same mass window for K^{*+} as that determined for K^{*0} in optimization of the $B^0 \to J/\psi K^{*0}$ mode, since the K^{*0} and K^{*+} are very close in mass ($\sim$4 MeV).

3.6.4 *b*-hadrons

Our *b*-hadron candidates are constructed by vertexing together combinations of J/ψ, $\psi(2S)$, V^0 candidates just described and, in some cases, other tracks in the event. The

following b-hadron modes are currently considered:

$$
\begin{aligned}
&\mathrm{B}^0 \to \mathrm{J}/\psi \mathrm{K}^0_\mathrm{s} \quad \text{, with } \mathrm{J}/\psi \to \mu^+\mu^- \\
&\mathrm{B}^0 \to \psi(2\mathrm{S})\mathrm{K}^0_\mathrm{s} \text{ , with } \psi(2\mathrm{S}) \to \mu^+\mu^-,\ \mathrm{K}^0_\mathrm{s} \to \pi^+\pi^- \\
&\mathrm{B}^0 \to \psi(2\mathrm{S})\mathrm{K}^0_\mathrm{s} \text{ , with } \psi(2\mathrm{S}) \to \mathrm{J}/\psi\pi^+\pi^-,\ \mathrm{J}/\psi \to \mu^+\mu^-,\ \mathrm{K}^0_\mathrm{s} \to \pi^+\pi^- \\
&\mathrm{B}^0 \to \mathrm{J}/\psi \mathrm{K}^{*0} \quad \text{, with } \mathrm{J}/\psi \to \mu^+\mu^-,\ \mathrm{K}^{*0} \to \mathrm{K}^+\pi^- \\
&\mathrm{B}^0 \to \psi(2\mathrm{S})\mathrm{K}^{*0} \text{, with } \psi(2\mathrm{S}) \to \mu^+\mu^-,\ \mathrm{K}^{*0} \to \mathrm{K}^+\pi^- \\
&\mathrm{B}^0 \to \psi(2\mathrm{S})\mathrm{K}^{*0} \text{, with } \psi(2\mathrm{S}) \to \mathrm{J}/\psi\pi^+\pi^-,\ \mathrm{J}/\psi \to \mu^+\mu^-,\ \mathrm{K}^{*0} \to \mathrm{K}^+\pi^- \\
&\mathrm{B}^+ \to \mathrm{J}/\psi \mathrm{K}^+ \quad \text{, with } \mathrm{J}/\psi \to \mu^+\mu^- \\
&\mathrm{B}^+ \to \psi(2\mathrm{S})\mathrm{K}^+ \text{, with } \psi(2\mathrm{S}) \to \mu^+\mu^- \\
&\mathrm{B}^+ \to \psi(2\mathrm{S})\mathrm{K}^+ \text{, with } \psi(2\mathrm{S}) \to \mathrm{J}/\psi\pi^+\pi^-,\ \mathrm{J}/\psi \to \mu^+\mu^- \\
&\mathrm{B}^+ \to \mathrm{J}/\psi \mathrm{K}^{*+} \text{ , with } \mathrm{J}/\psi \to \mu^+\mu^-,\ \mathrm{K}^{*+} \to \mathrm{K}^0_\mathrm{s}\pi^+ \\
&\Lambda^0_\mathrm{b} \to \mathrm{J}/\psi \Lambda^0 \quad \text{, with } \mathrm{J}/\psi \to \mu^+\mu^-,\ \Lambda^0 \to \mathrm{p}\pi^-
\end{aligned}
$$

In each fit involving a J/ψ or $\psi(2\mathrm{S})$, we introduce mass constraints on the daughter tracks of 3.096916 GeV/c^2 and 3.68596 GeV/c^2, respectively. No other mass constraints are used in our baseline analysis. In modes involving V^0, we introduce the additional constraint that the V^0 momentum points back to the J/ψ or $\psi(2\mathrm{S})$ vertex in 3 dimensions, to significantly reduce combinatorial backgrounds. In order to study the impact of the various vertex fitting techniques on our results, we also include in our studies the same candidates with the following kinematic fit and V^0 silicon hit variations:

- a 2-dimensional pointing constraint with no V^0 mass constraint
- a 2-dimensional pointing constraint with V^0 mass constraint
- a 3-dimensional pointing constraint with V^0 mass constraint
- no V^0 pointing constraint with V^0 mass constraint
- no V^0 pointing constraint with no V^0 mass constraint
- the standard 3-dimensional pointing constraint and no V^0 mass constraint with the COT parent track V^0 set described above.

For each b-hadron mode, there are cuts on the mass and $\mathrm{p_T}$ of the non-charmonium resonance (K^0_s, Λ^0, K^{*0}, K^{*+}) or simply the $\mathrm{p_T}$ of the kaon in $\mathrm{B}^+ \to \mathrm{J}/\psi\mathrm{K}^+$, $\mathrm{B}^+ \to \psi(2\mathrm{S})\mathrm{K}^+$ modes. For V^0 modes, there is an additional cut on the $\mathrm{L_{xy}}$ signficance ($\mathrm{L_{xy}}/\sigma_{\mathrm{L_{xy}}}$) of the V^0 as measured from the J/ψ vertex. This cut on the significance of a non-zero decay length takes advantage of the long V^0 lifetime to suppress backgrounds. Finally, we also make cuts on the b-hadron $\mathrm{p_T}$ and Prob(χ^2). The values used for all of these cuts are determined from an optimization procedure described in the next section involving the $\mathrm{B}^0 \to \mathrm{J}/\psi\mathrm{K}^0_\mathrm{s}$, $\Lambda^0_\mathrm{b} \to \mathrm{J}/\psi\Lambda^0$, $\mathrm{B}^0 \to \mathrm{J}/\psi\mathrm{K}^{*0}$, and $\mathrm{B}^+ \to \mathrm{J}/\psi\mathrm{K}^+$ b-hadron decay modes.

3.7 Cut Optimization

3.7.1 General Considerations

In optimization of our lifetime analysis, we are ultimately trying to minimize the total uncertainty on our measurement of the B^0 and Λ_b^0 lifetimes (and their ratio). While it remains to be demonstrated, we anticipate that the statistical error on our lifetimes will dominate over systematics given the sample sizes involved at our current integrated luminosity. Therefore, we focus on minimizing the statistical uncertainty, $\sigma_{c\tau}$, on our fitted lifetimes.

In the absence of background, our optimization would simply involve maximizing the sample (signal) size. Of course, we do have backgrounds in our B^0 and Λ_b^0 reconstruction and their effect on $\sigma_{c\tau}$ depends upon how much their proper decay length (PDL) distributions resemble the lifetime of the signal we are trying to measure.

To study how $\sigma_{c\tau}$ depends upon background size and composition, we fit for lifetime in toy datasets generated from a model of the PDL observed in the actual data. Using the B^0 selection in [7] and PDL model used in previous analyses (see [16], for example), we determine the model parameters consistent with a fraction of the data reconstructed as $B^0 \to J/\psi K_s^0$. The $\sigma_{c\tau}$ resulting from the fit, averaged over many toy datasets randomly drawn from the same model, is recorded for each variation of the generated background. We studied two variations in background:

- $N_{\text{background}}$ **scan**: For constant number of signal events and background composition, vary the total number of background events.

- f_{prompt} **scan**: For constant number of signal and total background events, vary the relative fraction of prompt and non-zero lifetime background.

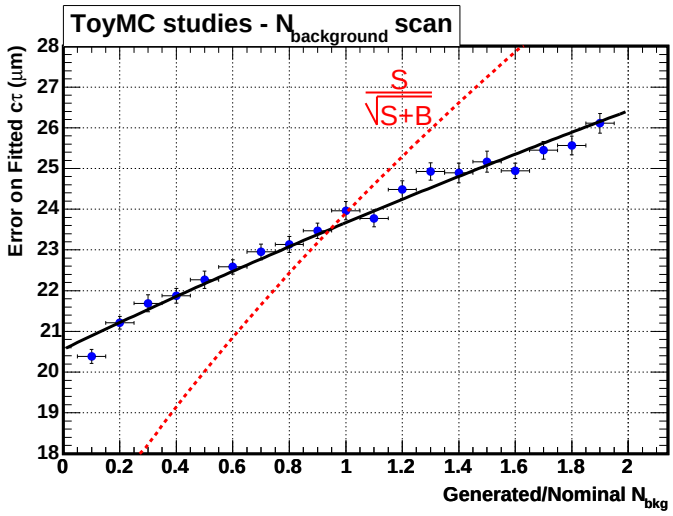
along with the $\sigma_{c\tau}$ scaling for $S/\sqrt{S+B}$ one would maximize for the significance of a signal in the presence of background.

One can draw several conclusions from an analysis of Figure 1:

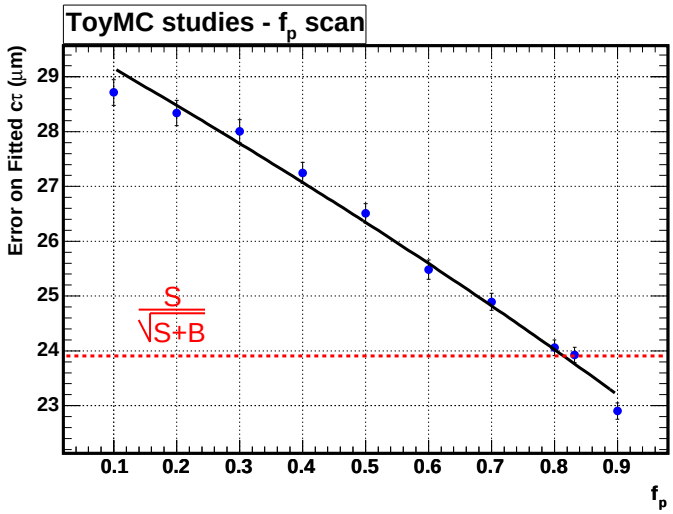
1. For constant S, $\sigma_{c\tau}$ and $S/\sqrt{S+B}$ follow the same trend on the total number of background events (i.e. as B increases, $\sigma_{c\tau}$ increases). The slopes are different, however, with $\sigma_{c\tau}$ less strongly dependent upon B than in $S/\sqrt{S+B}$. This is because the majority of background ($\sim$80% in this selection) is prompt rather than long-lived. The prompt background has a PDL shape which does not resemble the signal (i.e. has small statistical covariance) in contrast to the positive lifetime background.

2. For constant S and B, $\sigma_{c\tau}$ increases for increasing fraction of positive-lifetime background, in contrast to $S/\sqrt{S+B}$ which has no dependence (by construction).

Consideration of these results motivates the optimization of our selection. We conclude that, while not completely optimal, simple optimization of signal significance ($S/\sqrt{S+B}$)



(a)



(b)

Figure 1: Based upon toy Monte Carlo of a model for $B^0 \to J/\psi K^0_s$ candidates in the data, $\sigma_{c\tau}$ versus (a) the fractional change in the total number of background events and (b) relative fraction of zero-lifetime events in the overall background. Also shown in red is the error scaling for $S/\sqrt{S+B}$ fixed to match the toy MC $\sigma_{c\tau}$ at the nominal point.

is a reasonable approach to optimizing our measurement of the lifetimes. Since most of the background is prompt rather than positive lifetime, keeping signal is preferred over rejecting *total* background as compared with $S/\sqrt{S+B}$. It is also clear that cuts which cut preferentially the positive-lifetime background are preferred over indiscriminate background cuts and Figure 1(b) tells us that we should seek out such cuts.

The reader may be wondering at this point why we do not optimize our selection strictly upon the $\sigma_{c\tau}$ results in Figure 1(a). In fact, we have fit this data to the following parameterization

$$\frac{1}{\sigma_{c\tau}} = \frac{S}{\sqrt{S + \alpha B_{\rm prompt} + \beta B_{\rm expo}}} \tag{1}$$

where $B_{\rm prompt}$ and $B_{\rm expo}$ are the number of prompt and non-zero lifetime background events in a mass window of $\pm 12\sigma$, respectively, and $\alpha = 0.05$ and $\beta = 0.21$. We have actually implemented an optimization procedure where we fit, at each optimization step, the PDL distribution in the far sidebands to determine $B_{\rm prompt}$ and $B_{\rm expo}$ and maximize (using properly scaled signal Monte Carlo for S) Equation 1. We found that in practice it lead to very similar results as the signal significance optimization at the expense of a significantly more complex and time-consuming procedure involved many thousand lifetime fits. One exception was in b-meson p_T threshold, which the lifetime-based optimization pushed toward significantly lower threshold compared to the signal significance optimization. However, this difference is washed out by our uncertainty on the Λ_b^0 production spectrum coupled with our desire to keep the $\Lambda_b^0 \to J/\psi\Lambda^0$ and $B^0 \to J/\psi K_s^0$ selection as similar as possible. This issue is addressed in more detail in the next section.

In the end, we used the signal significance optimization procedure for selection and the lifetime-based optimization results were used as a guide (accept more signal, look for ways to supress long-lived background). We feel that as a matter of principle the most important reason for optimization in this analysis is not to fine tune to the most optimal point in cut space which will be imperfect anyway due to inaccuracies in modeling of the signal but rather to use it to convince others that we choose our cuts in an unbiased fashion.

3.7.2 Signal Significance Optimization Results

In this section, we present our optimization of the cuts in the selection described in Section 3.6. We choose to optimize the following cuts:

$B^0 \to J/\psi K_s^0$ ($\Lambda_b^0 \to J/\psi\Lambda^0$):

- $K_s^0(\Lambda^0)$ $L_{xy}/\sigma_{L_{xy}}$ from J/ψ vertex (L_{xy} significance)
- $K_s^0(\Lambda^0)$ p_T
- $K_s^0(\Lambda^0)$ mass window

- $B^0(\Lambda_b^0)$ Prob(χ^2)
- $B^0(\Lambda_b^0)$ p_T

$B^0 \to J/\psi K^{*0}$:

- K^{*0} p_T
- K^{*0} mass window
- B^0 Prob(χ^2)
- B^0 p_T

$B^+ \to J/\psi K^+$:

- K^+ p_T
- B^+ Prob(χ^2)
- B^+ p_T

Since $\Lambda_b^0 \to J/\psi\Lambda^0$ is the focus of our analysis and has smaller statistics than $B^0 \to J/\psi K_s^0$, we first optimize the $\Lambda_b^0 \to J/\psi\Lambda^0$ selection and try to keep analogous cuts for $B^0 \to J/\psi K_s^0$ as similar as possible. The $B^0 \to J/\psi K^{*0}$ and $B^+ \to J/\psi K^+$ modes are independently optimized. The Monte Carlo used for signal is single b-quark generation via BGenerator. For b-mesons modes, the generated b-quark spectrum corresponds to the measurement of the inclusive J/ψ spectrum as measured in [19]. Although it contributes to the inclusive J/ψ spectrum, the Λ_b^0 p_T is less well determined than that for b-mesons. However, properties of the Λ_b^0 production have been determined as part of the measurement in [20]. Their measured spectrum is actually quite a bit softer than the measured inclusive J/ψ spectrum. We use the measured Λ_b^0 spectrum provided to us by Eiko Yu as it relates to their analysis to generate the signal Monte Carlo used for this analysis and use the inclusive J/ψ spectrum as a cross-check.

Our optimization procedure is as follows. We reconstruct our signal Monte Carlo and data and apply the basic selection as described in Section 3.6. We find the values for the cuts to be optimized that maximize $S^2/(S+B)$, where S is the number of Monte Carlo events in 2σ of the signal peak (normalized by the ratio selected signal events in the data and Monte Carlo) and B is the number of background events predicted in the 2σ mass window using the data "far" sidebands. In $S^2/(S+B)$, S is always muliplied by the ratio of signal events observed in the data to that in the Monte Carlo to properly weight the relative signal and background sample sizes in the optimization. We define as the "far" sideband regions:

For $B^0 \to J/\psi K_s^0$:

- $5.07\ \mathrm{GeV/c^2} < \mathrm{M}_{\mu\mu\pi\pi} < 5.15\ \mathrm{GeV/c^2}$ **OR** $5.41\ \mathrm{GeV/c^2} < \mathrm{M}_{\mu\mu\pi\pi} < 5.49\ \mathrm{GeV/c^2}$

For $\Lambda_b^0 \to J/\psi\Lambda^0$

- $5.41\ \mathrm{GeV/c^2} < \mathrm{M}_{\mu\mu\pi\pi} < 5.50\ \mathrm{GeV/c^2}$ **OR** $5.74\ \mathrm{GeV/c^2} < \mathrm{M}_{\mu\mu\pi\pi} < 5.83\ \mathrm{GeV/c^2}$

We iterate this procedure until we would no longer choose a different value for the cuts. At each iteration step, all of the cuts are applied at their current optimal value except the cut being optimized (i.e. N-1 optimization).

Figures 2-6 show the final optimization iteration for $\Lambda_b^0 \to J/\psi\Lambda^0$ using the $\Lambda_b^0 \to \Lambda_c\pi$ spectrum Monte Carlo. Note that all optimization variables show a clear peak in $S/\sqrt{S+B}$ except for the Λ_b^0 p_T which is nearly flat near the lower threshold of our optimization. We choose to fix the Λ_b^0 p_T optimal cut at 4.0 GeV which is a reasonable threshold somewhere between the optimal determined from the $\Lambda_b^0 \to \Lambda_c\pi$ spectum and inclusive J/ψ spectrum results. The chosen optimal values are:

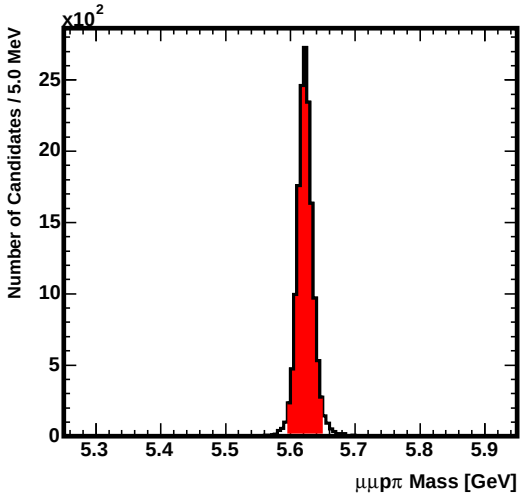
- Λ^0 $L_{xy}/\sigma_{L_{xy}}$ from J/ψ vertex = 4.0
- Λ^0 $p_T = 2.6$ GeV/c
- Λ^0 mass window = $\pm 9\ \mathrm{GeV/c^2}$
- Λ_b^0 Prob(χ^2) = 10^{-4}
- Λ_b^0 $p_T = 4.0$ GeV/c

Figures 7-11 show the final optimization iteration for $\Lambda_b^0 \to J/\psi\Lambda^0$ using the inclusive J/ψ spectrum Monte Carlo and Λ_b^0 p_T fixed to 4.0 GeV/c in the optimization. The optimal values are:

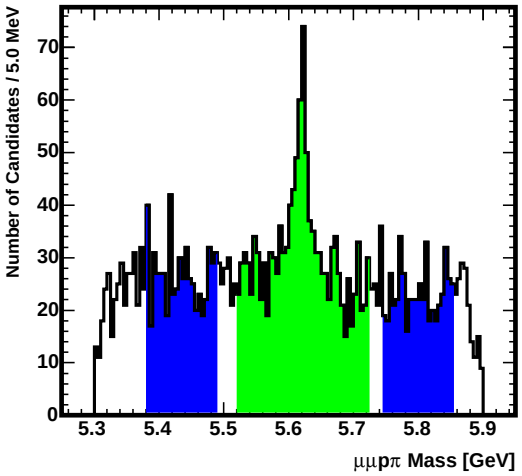
- Λ^0 $L_{xy}/\sigma_{L_{xy}}$ from J/ψ vertex = 3.0
- Λ^0 $p_T = 2.8$ GeV/c
- Λ^0 mass window = $\pm 9\ \mathrm{GeV/c^2}$
- Λ_b^0 Prob(χ^2) = 10^{-4}
- Λ_b^0 $p_T = 4.0$ GeV/c (fixed)

These are very similar to the optimal thresholds using the $\Lambda_b^0 \to \Lambda_c\pi$ spectrum Monte Carlo. We conclude that our $\Lambda_b^0 \to J/\psi\Lambda^0$ optimal cuts via this procedure are not strongly sensitive to the choice of Λ_b^0 p_T spectrum (except for the Λ_b^0 p_T spectrum itself!).

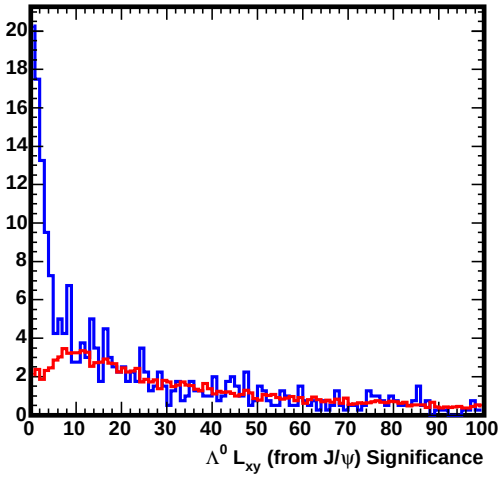
Figures 12-16 show the final optimization iteration for $B^0 \to J/\psi K_s^0$ with B^0 p_T fixed to 4.0 GeV/c in the optimization. We fixed the B^0 p_T to 4.0 GeV/c in the spirit of keeping



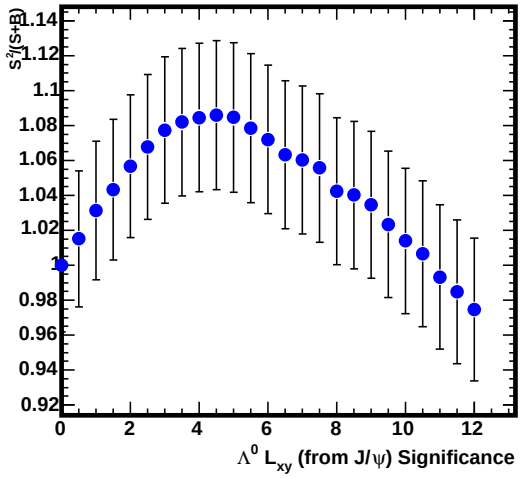
(a) Λ_b^0 mass in $\Lambda_b^0 \to J/\psi\Lambda^0$ in Monte Carlo. In red is the 2σ region used in the optimization.



(b) Λ_b^0 mass in data. In blue are the sideband regions used in the optimization. In green is the region used to fit for the lifetime.

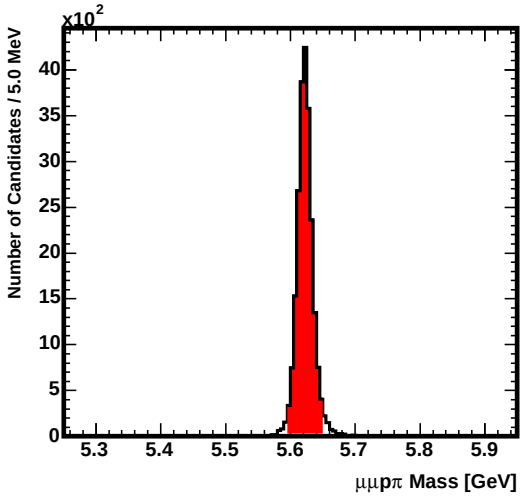


(c) Distribution of cut being optimized. In red is signal region from MC and blue is data sidebands, with normalization set by the relative signal yield in data and Monte Carlo.

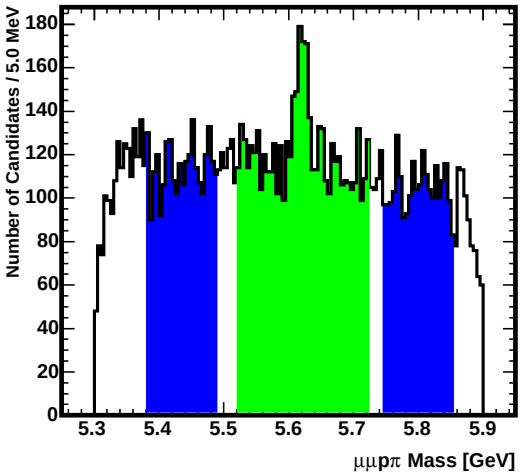


(d) Signal significance versus cut being optimized. The arrow indicates the optimal cut value.

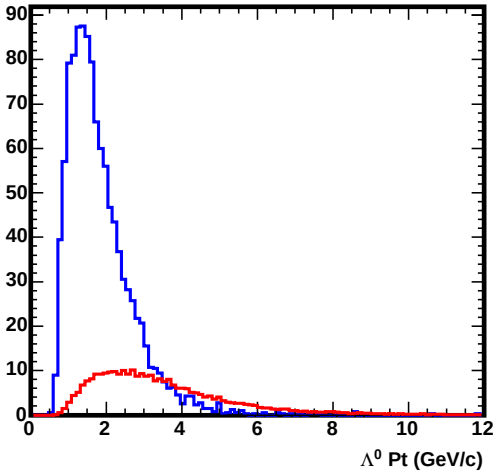
Figure 2: $\Lambda_b^0 \to J/\psi\Lambda^0$ optimization of Λ^0 $L_{xy}/\sigma_{L_{xy}}$ from the J/ψ vertex, with all other cuts applied. Measured $\Lambda_b^0 \to \Lambda_c\pi$ spectrum from [20] used in Monte Carlo generation.



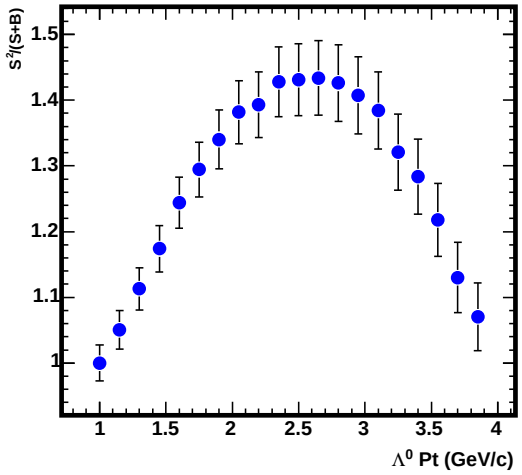
(a) Λ_b^0 mass in $\Lambda_b^0 \to J/\psi\Lambda^0$ in Monte Carlo. In red is the 2σ region used in the optimization.



(b) Λ_b^0 mass in data. In blue are the sideband regions used in the optimization. In green is the region used to fit for the lifetime.

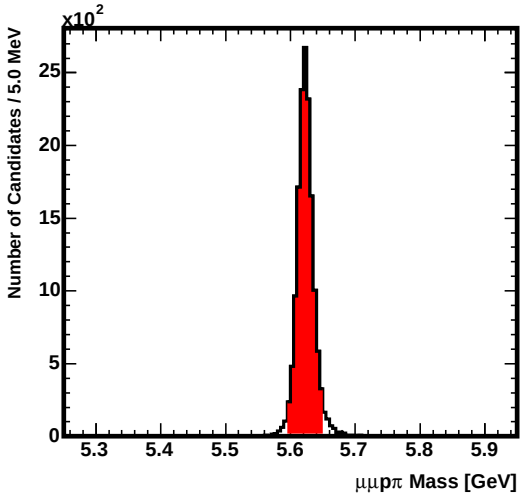


(c) Distribution of cut being optimized. In red is signal region from MC and blue is data sidebands, with normalization set by the relative signal yield in data and Monte Carlo.

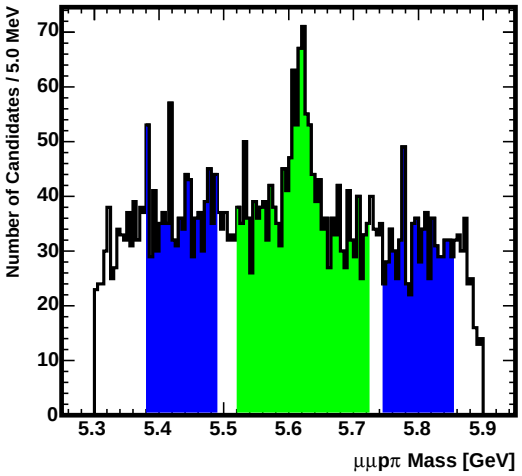


(d) Signal significance versus cut being optimized. The arrow indicates the optimal cut value.

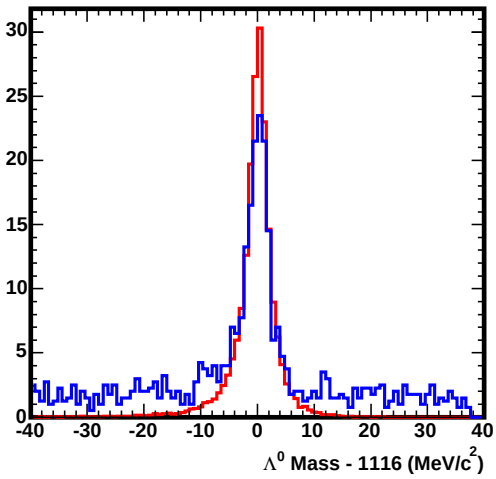
Figure 3: $\Lambda_b^0 \to J/\psi\Lambda^0$ optimization of Λ^0 p_T, with all other cuts applied. Measured $\Lambda_b^0 \to \Lambda_c\pi$ spectrum from [20] used in Monte Carlo generation.



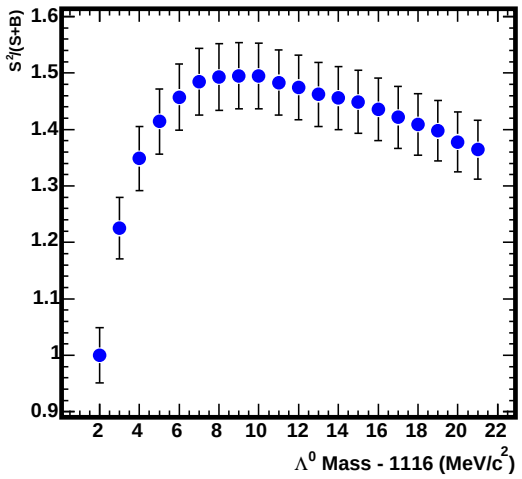
(a) Λ_b^0 mass in $\Lambda_b^0 \to J/\psi\Lambda^0$ in Monte Carlo. In red is the 2σ region used in the optimization.



(b) Λ_b^0 mass in data. In blue are the sideband regions used in the optimization. In green is the region used to fit for the lifetime.

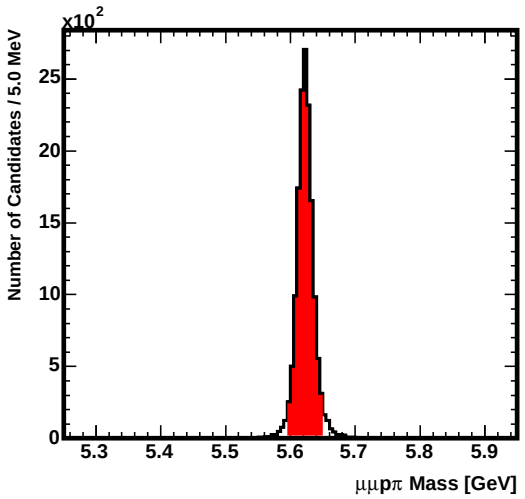


(c) Distribution of cut being optimized. In red is signal region from MC and blue is data sidebands, with normalization set by the relative signal yield in data and Monte Carlo.

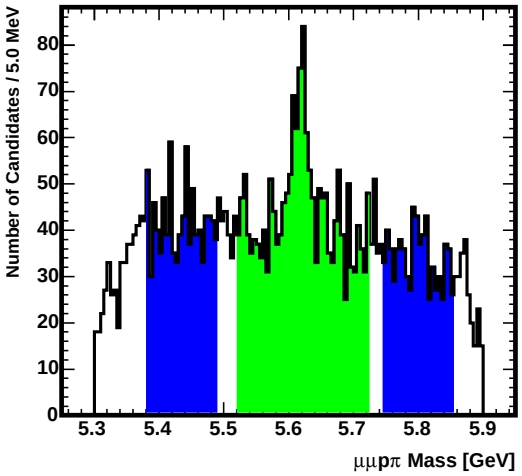


(d) Signal significance versus cut being optimized. The arrow indicates the optimal cut value.

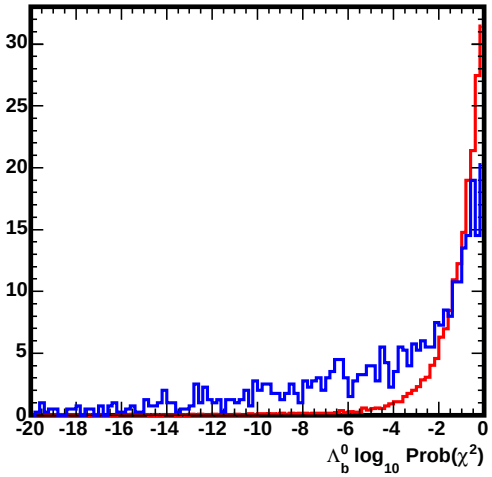
Figure 4: $\Lambda_b^0 \to J/\psi\Lambda^0$ optimization of Λ^0 mass window, with all other cuts applied. Measured $\Lambda_b^0 \to \Lambda_c\pi$ spectrum from [20] used in Monte Carlo generation.



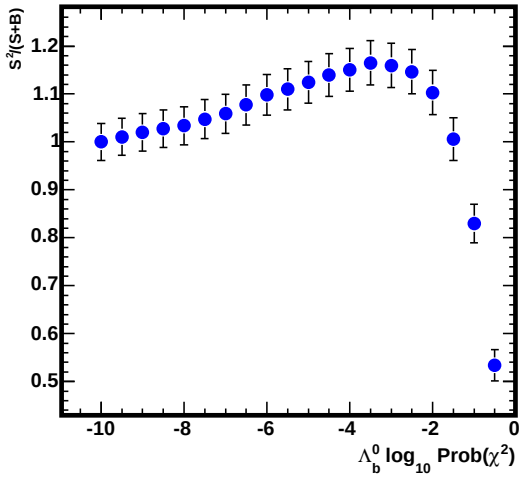
(a) Λ_b^0 mass in $\Lambda_b^0 \to J/\psi\Lambda^0$ in Monte Carlo. In red is the 2σ region used in the optimization.



(b) Λ_b^0 mass in data. In blue are the sideband regions used in the optimization. In green is the region used to fit for the lifetime.

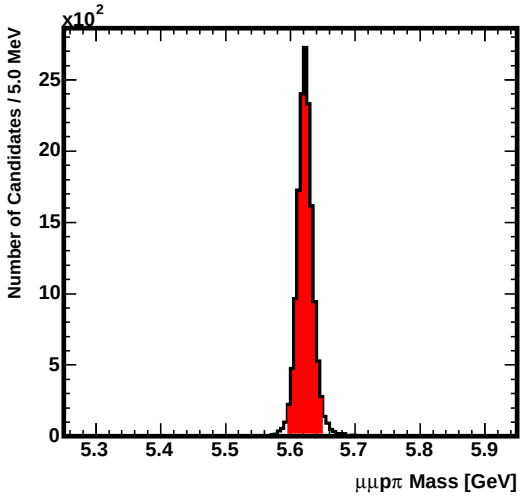


(c) Distribution of cut being optimized. In red is signal region from MC and blue is data sidebands, with normalization set by the relative signal yield in data and Monte Carlo.

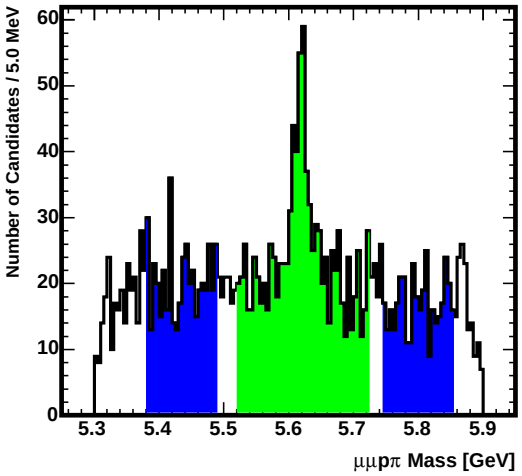


(d) Signal significance versus cut being optimized. The arrow indicates the optimal cut value.

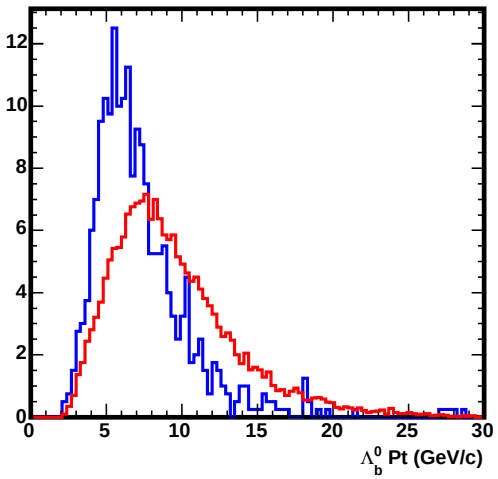
Figure 5: $\Lambda_b^0 \to J/\psi\Lambda^0$ optimization of Prob(χ^2) for Λ_b^0 vertex fit, with all other cuts applied. Measured $\Lambda_b^0 \to \Lambda_c\pi$ spectrum from [20] used in Monte Carlo generation.



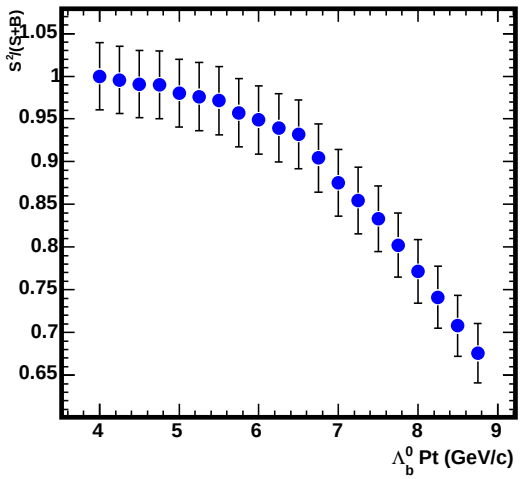
(a) Λ_b^0 mass in $\Lambda_b^0 \to J/\psi\Lambda^0$ in Monte Carlo. In red is the 2σ region used in the optimization.



(b) Λ_b^0 mass in data. In blue are the sideband regions used in the optimization. In green is the region used to fit for the lifetime.

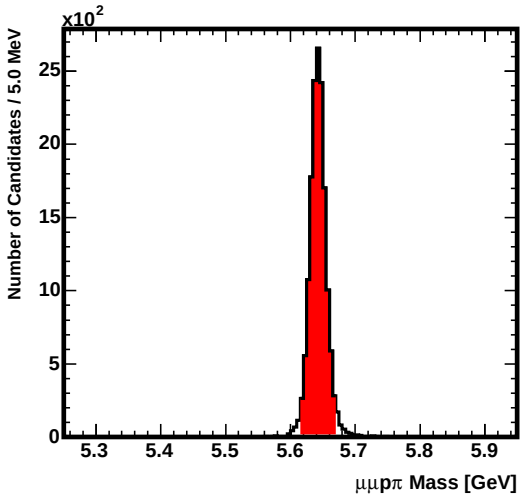


(c) Distribution of cut being optimized. In red is signal region from MC and blue is data sidebands, with normalization set by the relative signal yield in data and Monte Carlo.

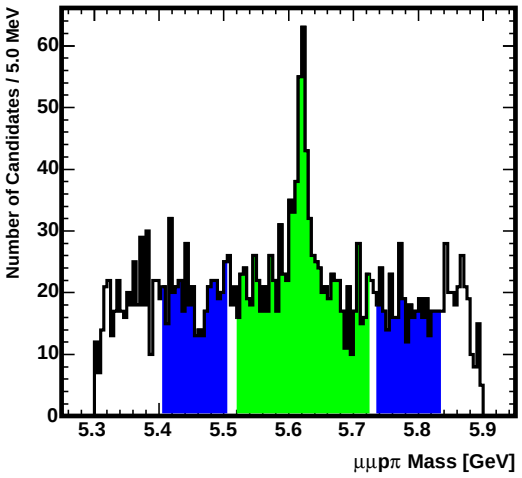


(d) Signal significance versus cut being optimized. The arrow indicates the optimal cut value.

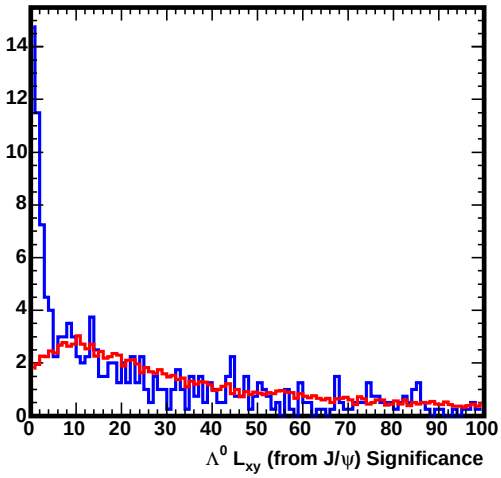
Figure 6: $\Lambda_b^0 \to J/\psi\Lambda^0$ optimization of Λ_b^0 p_T, with all other cuts applied. Measured $\Lambda_b^0 \to \Lambda_c\pi$ spectrum from [20] used in Monte Carlo generation.



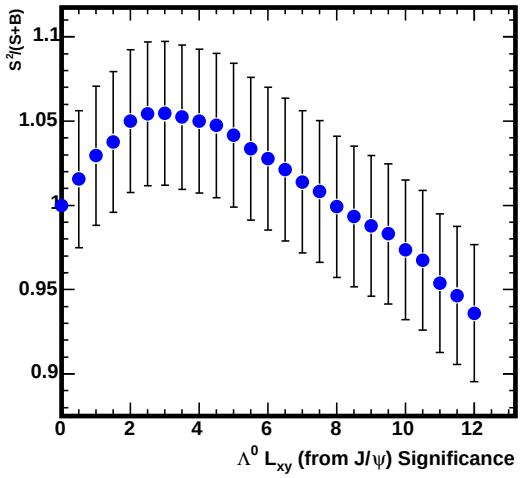
(a) Λ_b^0 mass in $\Lambda_b^0 \to J/\psi\Lambda^0$ in Monte Carlo. In red is the 2σ region used in the optimization.



(b) Λ_b^0 mass in data. In blue are the sideband regions used in the optimization. In green is the region used to fit for the lifetime.

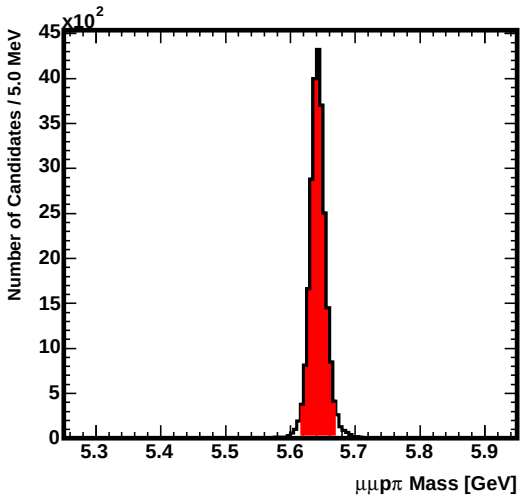


(c) Distribution of cut being optimized. In red is signal region from MC and blue is data sidebands, with normalization set by the relative signal yield in data and Monte Carlo.

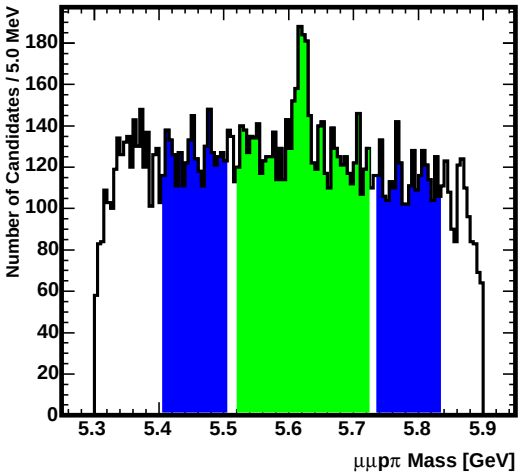


(d) Signal significance versus cut being optimized. The arrow indicates the optimal cut value.

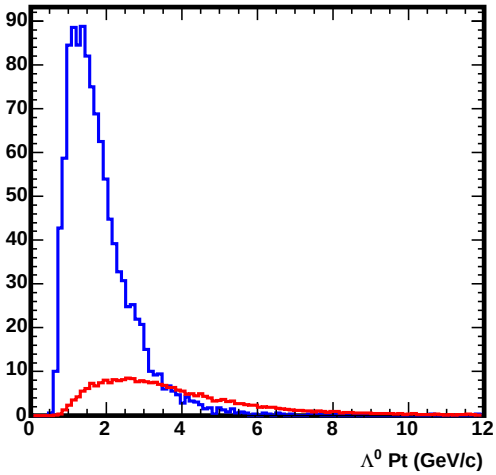
Figure 7: $\Lambda_b^0 \to J/\psi\Lambda^0$ optimization of Λ^0 $L_{xy}/\sigma_{L_{xy}}$ from the J/ψ vertex, with all other cuts applied. Measured inclusive $b \to J/\psi X$ spectrum from [19] used in Monte Carlo generation.



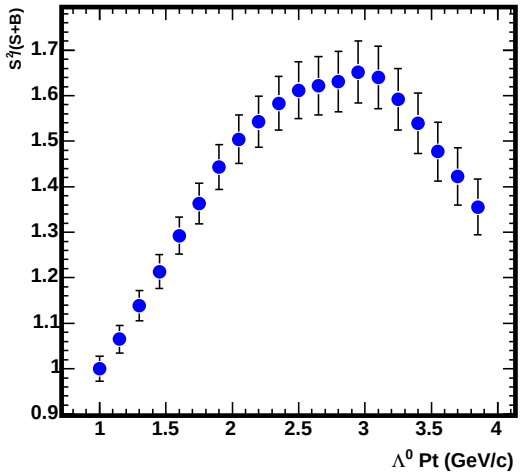
(a) Λ_b^0 mass in $\Lambda_b^0 \to J/\psi\Lambda^0$ in Monte Carlo. In red is the 2σ region used in the optimization.



(b) Λ_b^0 mass in data. In blue are the sideband regions used in the optimization. In green is the region used to fit for the lifetime.

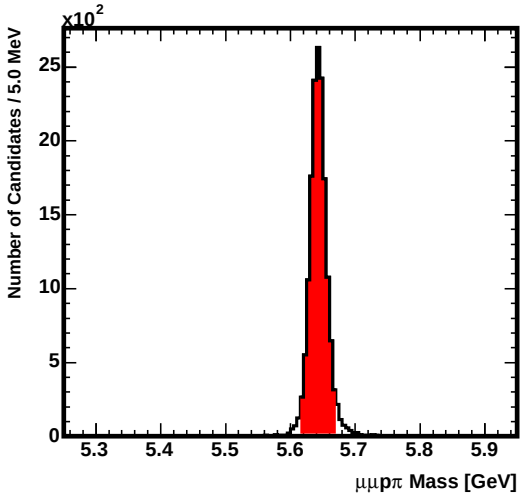


(c) Distribution of cut being optimized. In red is signal region from MC and blue is data sidebands, with normalization set by the relative signal yield in data and Monte Carlo.

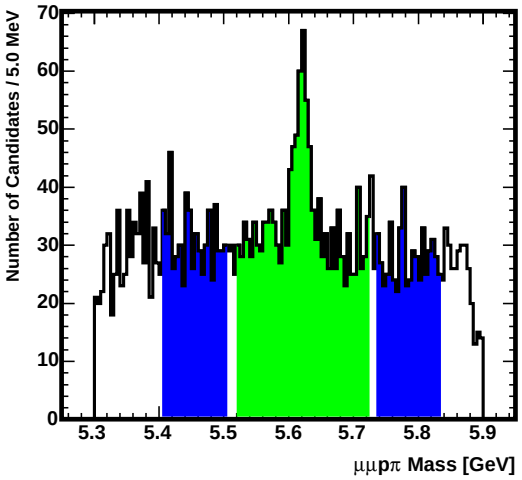


(d) Signal significance versus cut being optimized. The arrow indicates the optimal cut value.

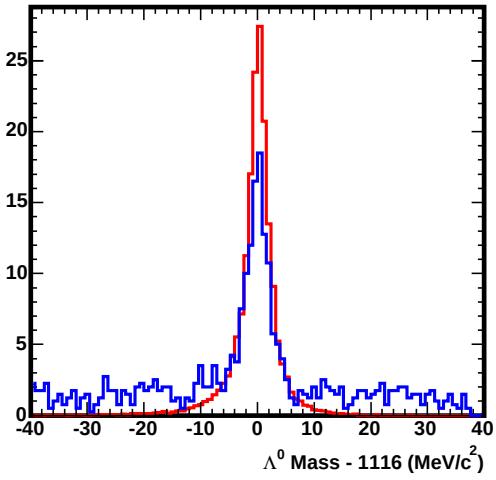
Figure 8: $\Lambda_b^0 \to J/\psi\Lambda^0$ optimization of Λ^0 p_T, with all other cuts applied. Measured inclusive $b \to J/\psi X$ spectrum from [19] used in Monte Carlo generation.



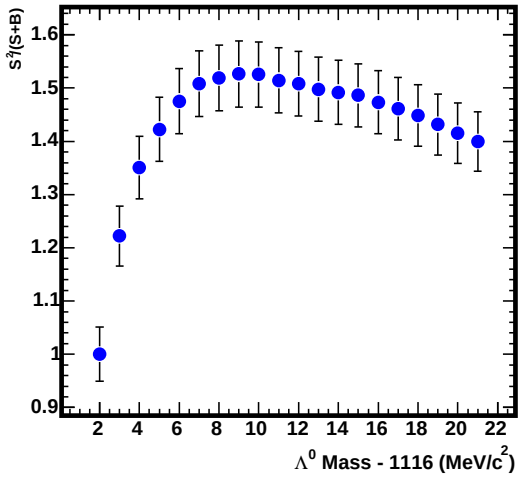
(a) Λ_b^0 mass in $\Lambda_b^0 \to J/\psi\Lambda^0$ in Monte Carlo. In red is the 2σ region used in the optimization.



(b) Λ_b^0 mass in data. In blue are the sideband regions used in the optimization. In green is the region used to fit for the lifetime.

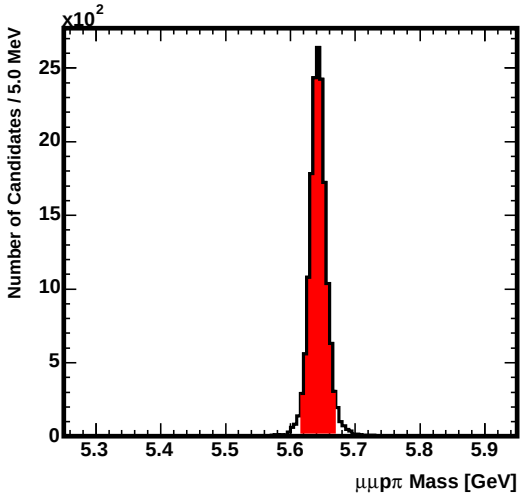


(c) Distribution of cut being optimized. In red is signal region from MC and blue is data sidebands, with normalization set by the relative signal yield in data and Monte Carlo.

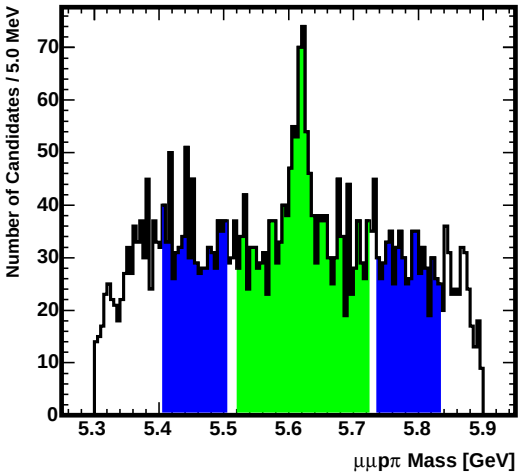


(d) Signal significance versus cut being optimized. The arrow indicates the optimal cut value.

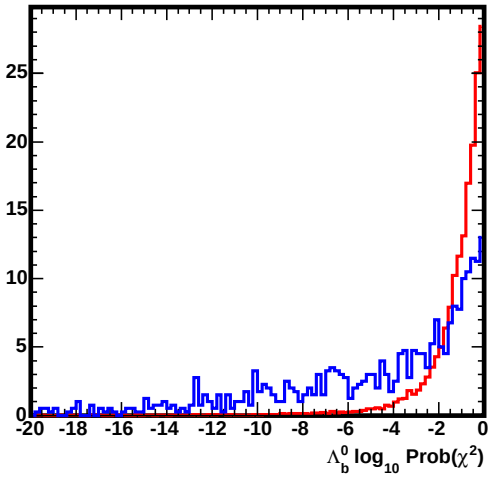
Figure 9: $\Lambda_b^0 \to J/\psi\Lambda^0$ optimization of Λ^0 mass window, with all other cuts applied. Measured inclusive $b \to J/\psi X$ spectrum from [19] used in Monte Carlo generation.



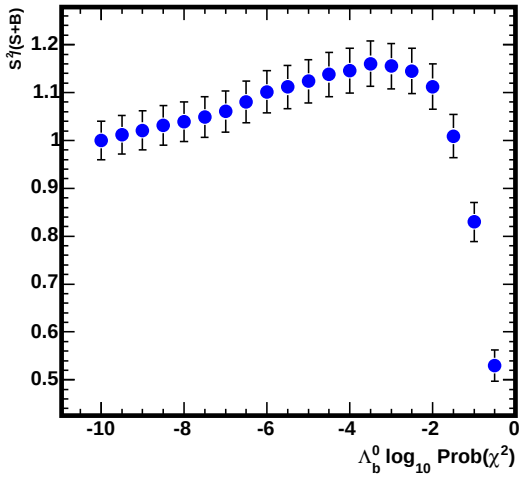
(a) Λ_b^0 mass in $\Lambda_b^0 \to J/\psi\Lambda^0$ in Monte Carlo. In red is the 2σ region used in the optimization.



(b) Λ_b^0 mass in data. In blue are the sideband regions used in the optimization. In green is the region used to fit for the lifetime.

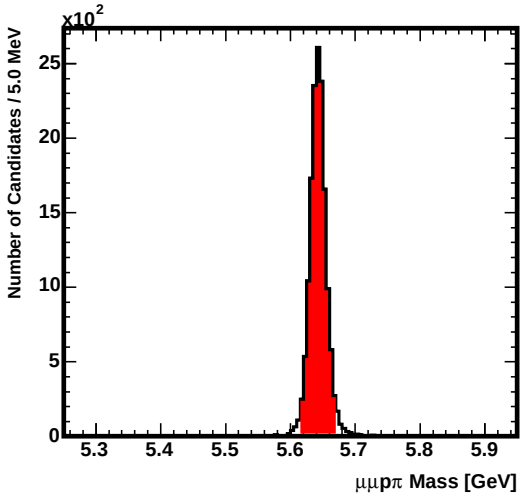


(c) Distribution of cut being optimized. In red is signal region from MC and blue is data sidebands, with normalization set by the relative signal yield in data and Monte Carlo.

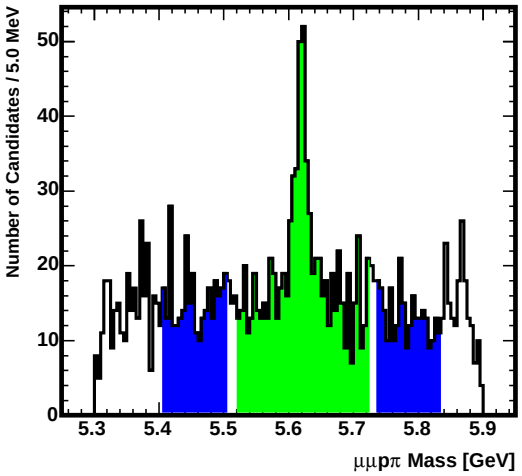


(d) Signal significance versus cut being optimized. The arrow indicates the optimal cut value.

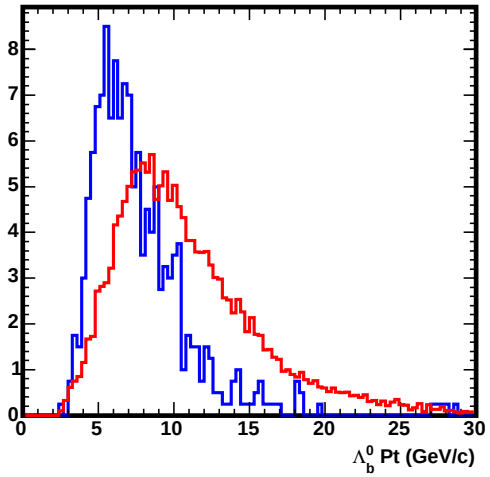
Figure 10: $\Lambda_b^0 \to J/\psi\Lambda^0$ optimization of Prob(χ^2) for Λ_b^0 vertex fit, with all other cuts applied. Measured inclusive $b \to J/\psi X$ spectrum from [19] used in Monte Carlo generation.



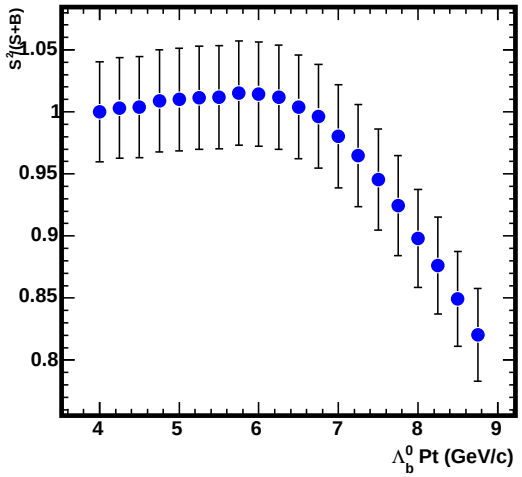
(a) Λ_b^0 mass in $\Lambda_b^0 \to J/\psi\Lambda^0$ in Monte Carlo. In red is the 2σ region used in the optimization.



(b) Λ_b^0 mass in data. In blue are the sideband regions used in the optimization. In green is the region used to fit for the lifetime.

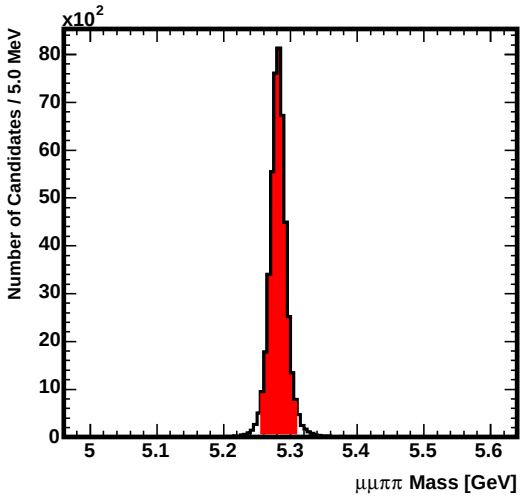


(c) Distribution of cut being optimized. In red is signal region from MC and blue is data sidebands, with normalization set by the relative signal yield in data and Monte Carlo.

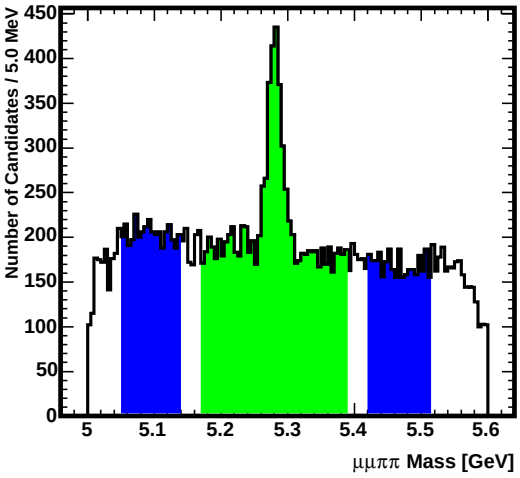


(d) Signal significance versus cut being optimized. The arrow indicates the optimal cut value.

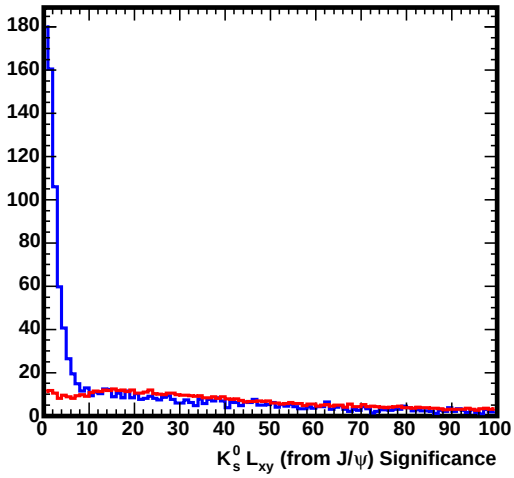
Figure 11: $\Lambda_b^0 \to J/\psi\Lambda^0$ optimization of Λ_b^0 p_T, with all other cuts applied. Measured inclusive $b \to J/\psi X$ spectrum from [19] used in Monte Carlo generation.



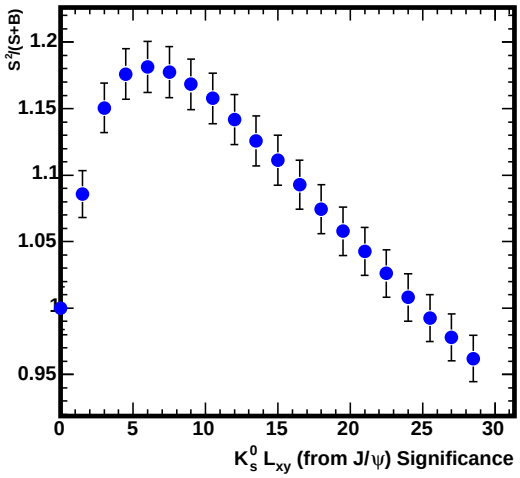
(a) B^0 mass in $B^0 \to J/\psi K^0_s$ in Monte Carlo. In red is the 2σ region used in the optimization.



(b) B^0 mass in data. In blue are the sideband regions used in the optimization. In green is the region used to fit for the lifetime.

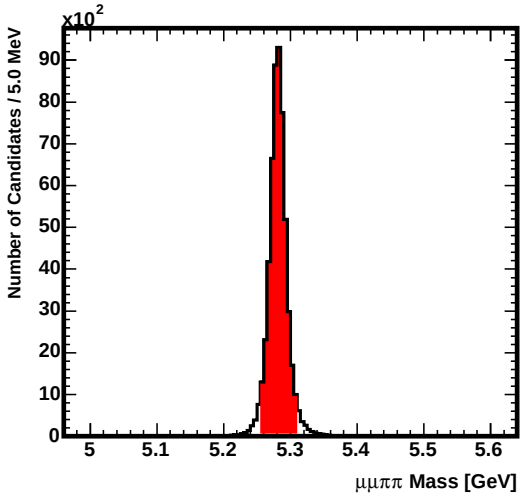


(c) Distribution of cut being optimized. In red is signal region from MC and blue is data sidebands, with normalization set by the relative signal yield in data and Monte Carlo.

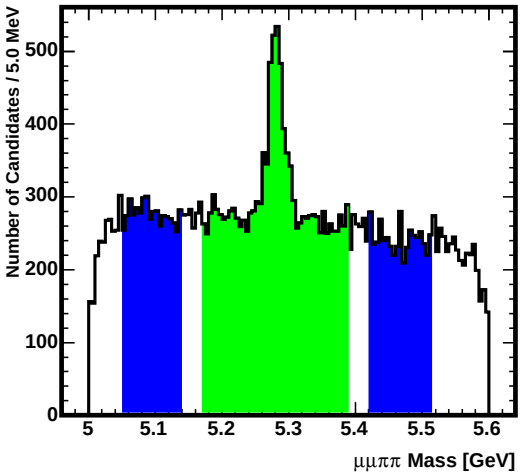


(d) Signal significance versus cut being optimized. The arrow indicates the optimal cut value.

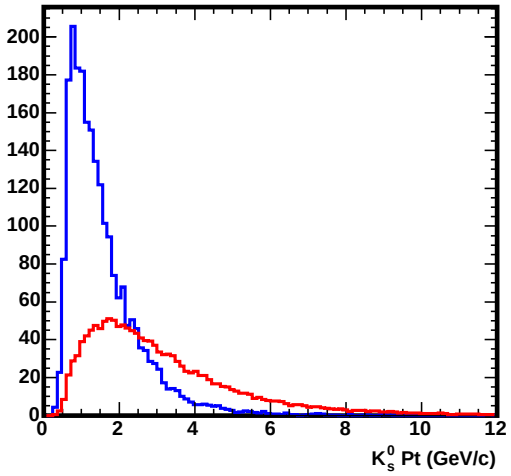
Figure 12: $B^0 \to J/\psi K^0_s$ optimization of K^0_s $L_{xy}/\sigma_{L_{xy}}$ from the J/ψ vertex, with all other cuts applied. Measured inclusive $b \to J/\psi X$ spectrum from [19] used in Monte Carlo generation.



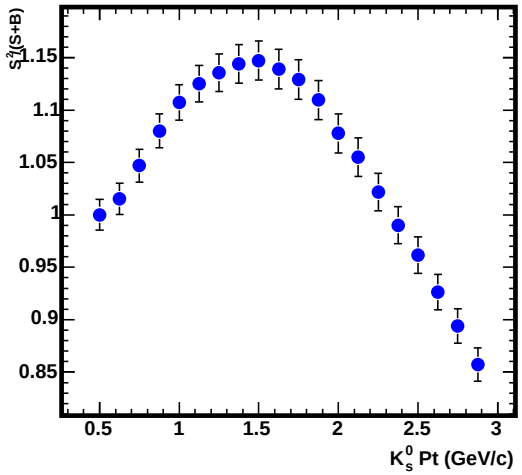
(a) B^0 mass in $B^0 \to J/\psi K^0_s$ in Monte Carlo. In red is the 2σ region used in the optimization.



(b) B^0 mass in data. In blue are the sideband regions used in the optimization. In green is the region used to fit for the lifetime.

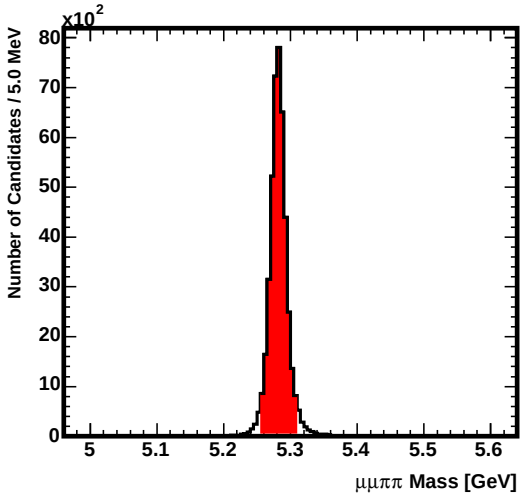


(c) Distribution of cut being optimized. In red is signal region from MC and blue is data sidebands, with normalization set by the relative signal yield in data and Monte Carlo.

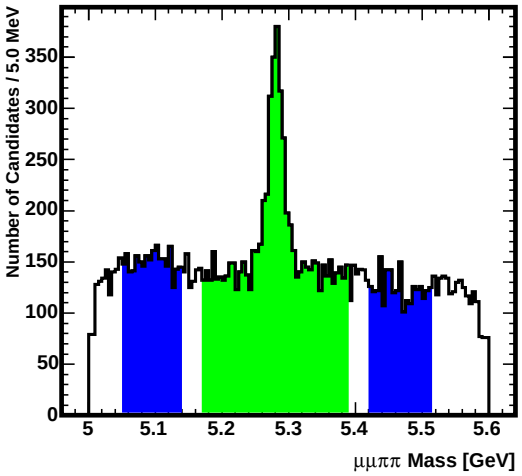


(d) Signal significance versus cut being optimized. The arrow indicates the optimal cut value.

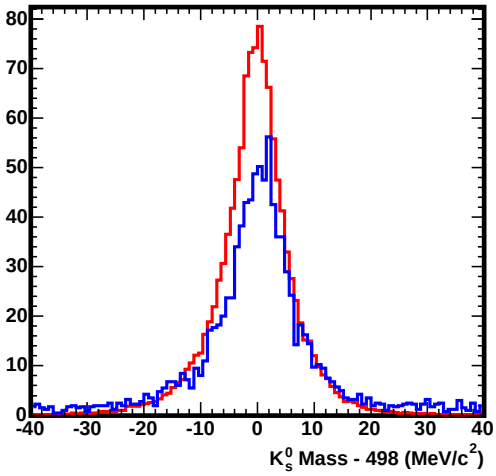
Figure 13: $B^0 \to J/\psi K^0_s$ optimization of K^0_s p_T, with all other cuts applied. Measured inclusive $b \to J/\psi X$ spectrum from [19] used in Monte Carlo generation.



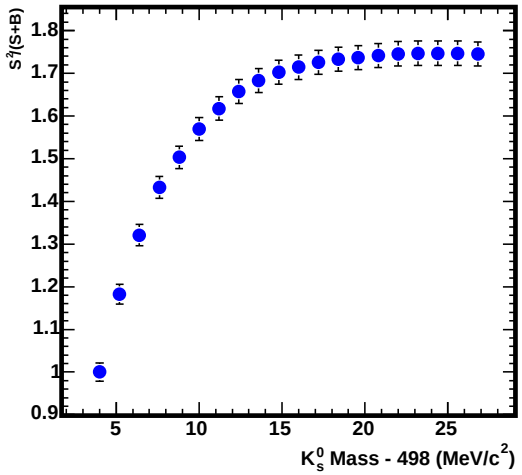
(a) B^0 mass in $B^0 \to J/\psi K^0_s$ in Monte Carlo. In red is the 2σ region used in the optimization.



(b) B^0 mass in data. In blue are the sideband regions used in the optimization. In green is the region used to fit for the lifetime.

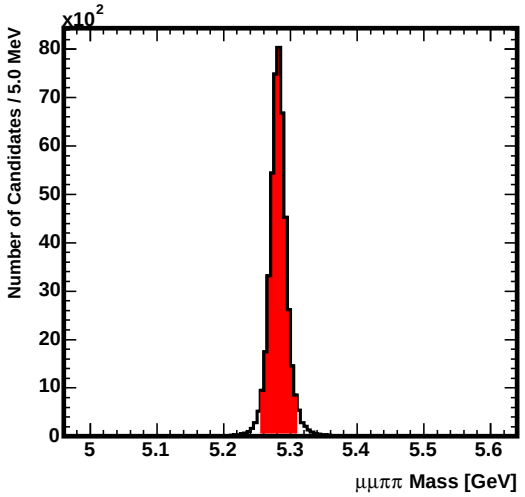


(c) Distribution of cut being optimized. In red is signal region from MC and blue is data sidebands, with normalization set by the relative signal yield in data and Monte Carlo.

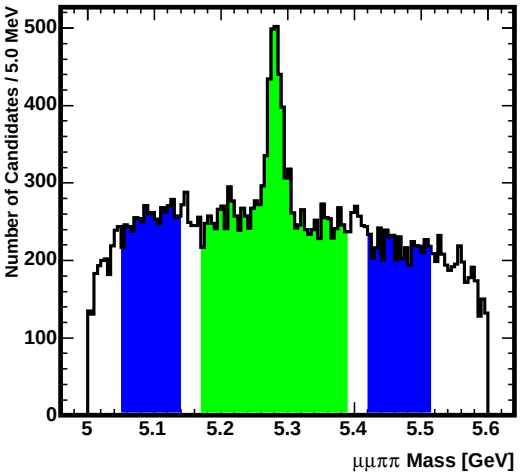


(d) Signal significance versus cut being optimized. The arrow indicates the optimal cut value.

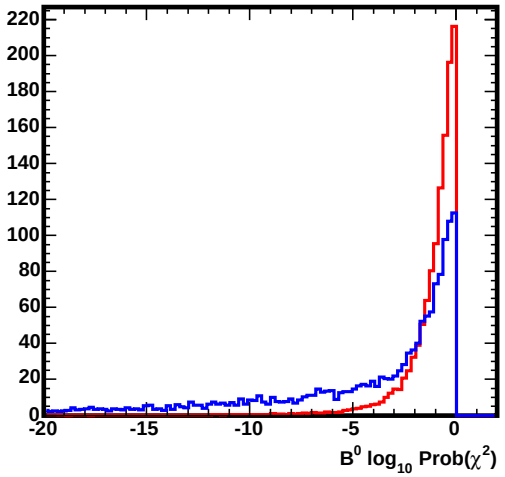
Figure 14: $B^0 \to J/\psi K^0_s$ optimization of K^0_s mass window, with all other cuts applied. Measured inclusive $b \to J/\psi X$ spectrum from [19] used in Monte Carlo generation.



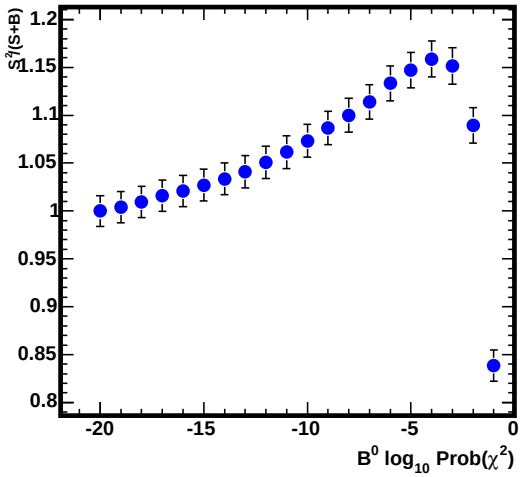
(a) B^0 mass in $B^0 \to J/\psi K^0_s$ Monte Carlo. In red is the 2σ region used in the optimization.



(b) B^0 mass in data. In blue are the sideband regions used in the optimization. In green is the region used to fit for the lifetime.

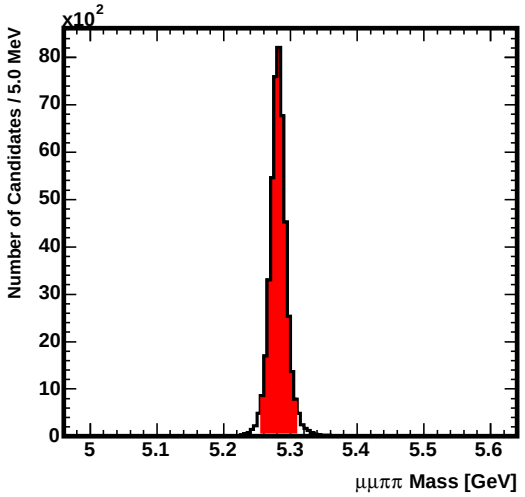


(c) Distribution of cut being optimized. In red is signal region from MC and blue is data sidebands, with normalization set by the relative signal yield in data and Monte Carlo.

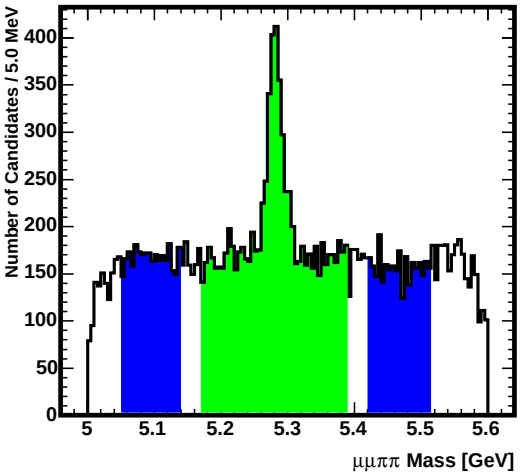


(d) Signal significance versus cut being optimized. The arrow indicates the optimal cut value.

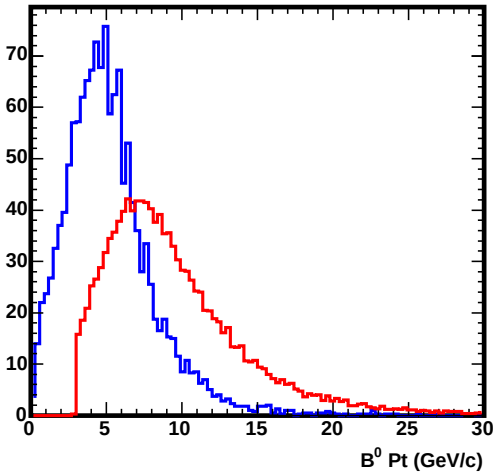
Figure 15: $B^0 \to J/\psi K^0_s$ optimization of Prob(χ^2) for B^0 vertex fit, with all other cuts applied. Measured inclusive $b \to J/\psi X$ spectrum from [19] used in Monte Carlo generation.



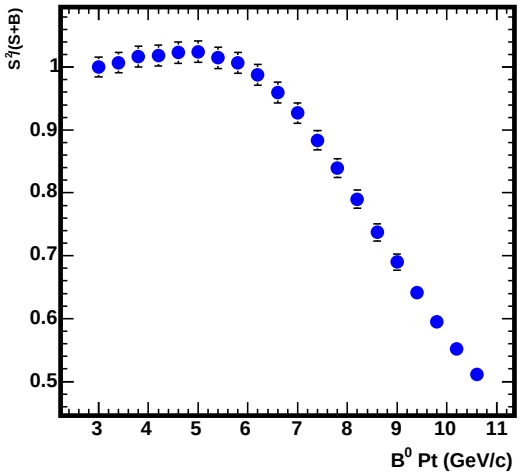
(a) B^0 mass in $B^0 \to J/\psi K^0_s$ in Monte Carlo. In red is the 2σ region used in the optimization.



(b) B^0 mass in data. In blue are the sideband regions used in the optimization. In green is the region used to fit for the lifetime.



(c) Distribution of cut being optimized. In red is signal region from MC and blue is data sidebands, with normalization set by the relative signal yield in data and Monte Carlo.



(d) Signal significance versus cut being optimized. The arrow indicates the optimal cut value.

Figure 16: $B^0 \to J/\psi K^0_s$ optimization of B^0 p_T, with all other cuts applied. Measured inclusive $b \to J/\psi X$ spectrum from [19] used in Monte Carlo generation.

the Λ_b^0 and B^0 selection similar and also because the lifetime-based optimization points to a looser selection that accepts more signal at the expense of even more background[2]. The chosen optimal values are:

- K_s^0 $L_{xy}/\sigma_{L_{xy}}$ from J/ψ vertex = 6.0
- K_s^0 $p_T = 1.5$ GeV/c
- K_s^0 mass window = ± 25 GeV/c^2
- B^0 Prob(χ^2) = 10^{-4}
- B^0 $p_T = 4.0$ GeV/c (fixed)

The $B^0 \to J/\psi K^{*0}$ and $B^+ \to J/\psi K^+$ selection are separately optimized without any of the variables fixed so that this optimization is directly applicable to full analysis of these modes. Figure 17-20 show the final optimization iteration for $B^0 \to J/\psi K^{*0}$. The chosen optimal values are:

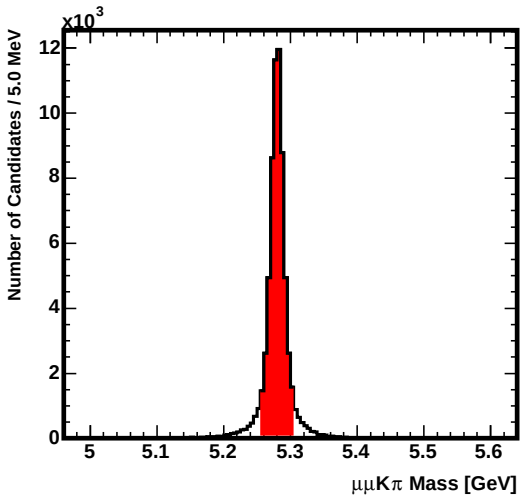
- K^{*0} $p_T = 3.0$ GeV/c
- K^{*0} mass window = ± 65 GeV/c^2
- B^0 Prob(χ^2) = 10^{-3}
- B^0 $p_T = 6.4$ GeV/c

Figures 21-23 show the

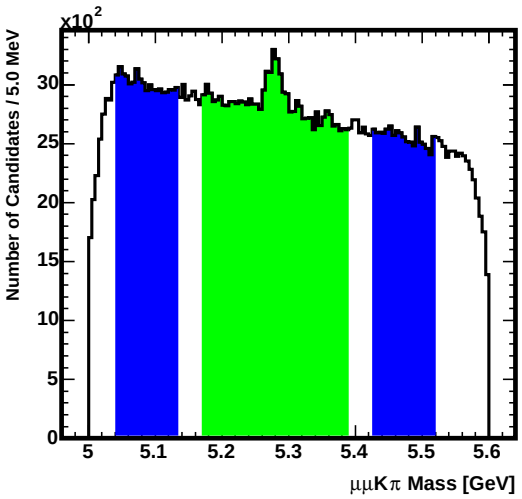
Table 1 summarizes the cut values with maximum $S^2/(S+B)$ for $B^0 \to J/\psi K_s^0$ and $\Lambda_b^0 \to J/\psi \Lambda^0$ based upon our signal Monte Carlo and data sidebands. Also shown are the cut values we choose to apply for this analysis. We have chosen the Λ_b^0 and Λ^0 p_T cuts slight looser than optimal because we do not trust the Monte Carlo to produce the correct p_T distributions. show the final optimization iteration for $B^0 \to J/\psi K^{*0}$. The chosen optimal values are:

- K^+ $p_T = 2.0$ GeV/c
- B^0 Prob(χ^2) = 10^{-3}
- B^0 $p_T = 6.2$ GeV/c

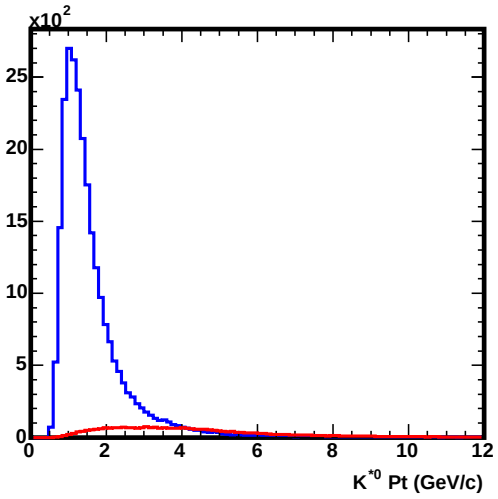
[2]recall from Section 3.7.1 that this is because the majority of background is prompt rather than long-lived



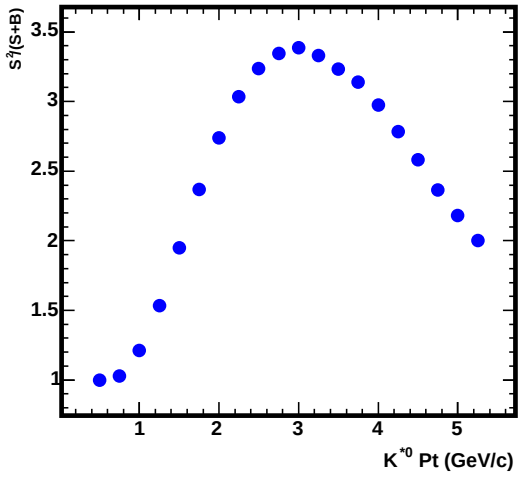
(a) B^0 mass in $B^0 \to J/\psi K^{*0}$ in Monte Carlo. In red is the 2σ region used in the optimization.



(b) B^0 mass in data. In blue are the sideband regions used in the optimization. In green is the region used to fit for the lifetime.

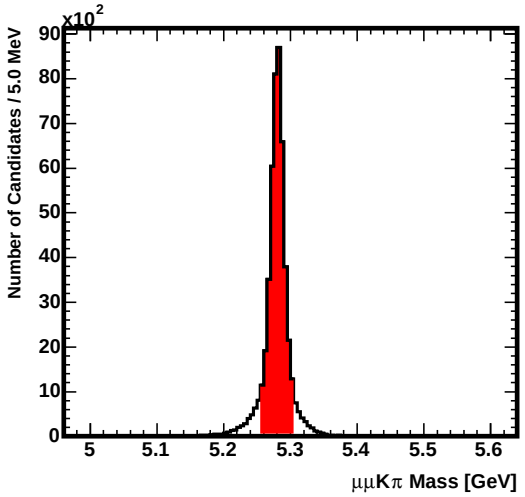


(c) Distribution of cut being optimized. In red is signal region from MC and blue is data sidebands, with normalization set by the relative signal yield in data and Monte Carlo.

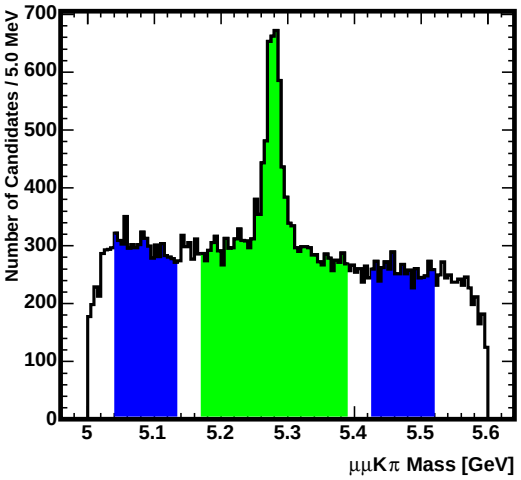


(d) Signal significance versus cut being optimized. The arrow indicates the optimal cut value.

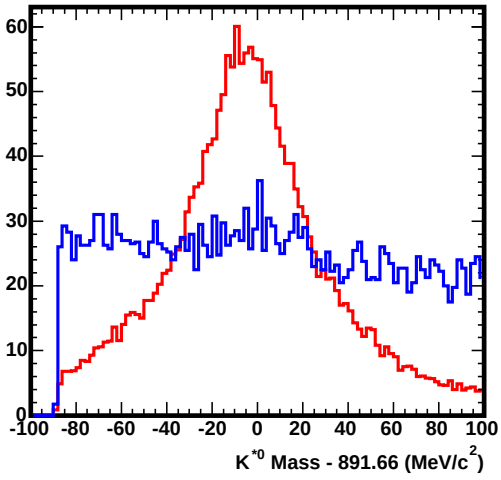
Figure 17: $B^0 \to J/\psi K^{*0}$ optimization of K^0_s p_T, with all other cuts applied. Measured inclusive $b \to J/\psi X$ spectrum from [19] used in Monte Carlo generation.



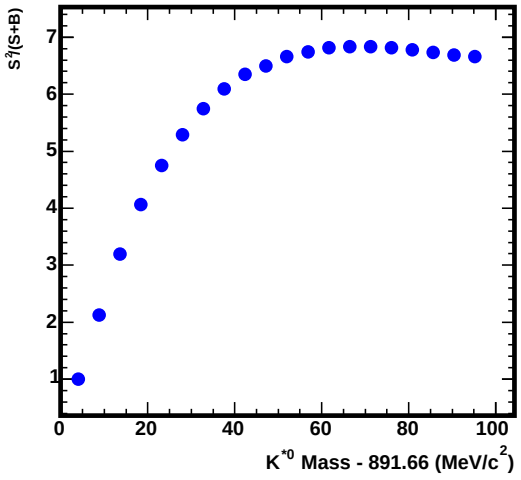
(a) B^0 mass in $B^0 \to J/\psi K^{*0}$ in Monte Carlo. In red is the 2σ region used in the optimization.



(b) B^0 mass in data. In blue are the sideband regions used in the optimization. In green is the region used to fit for the lifetime.

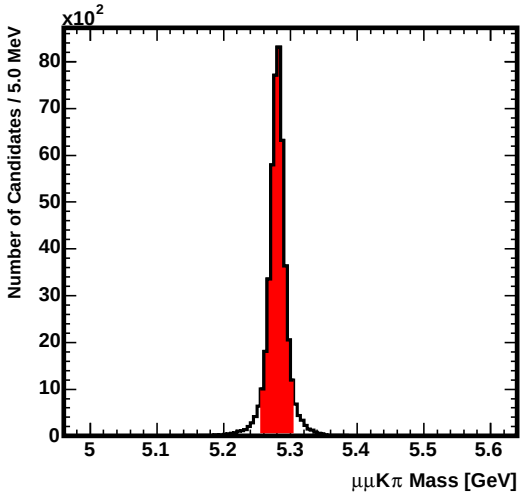


(c) Distribution of cut being optimized. In red is signal region from MC and blue is data sidebands, with normalization set by the relative signal yield in data and Monte Carlo.

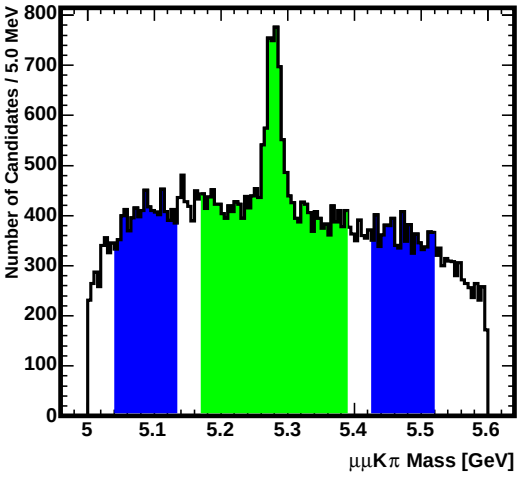


(d) Signal significance versus cut being optimized. The arrow indicates the optimal cut value.

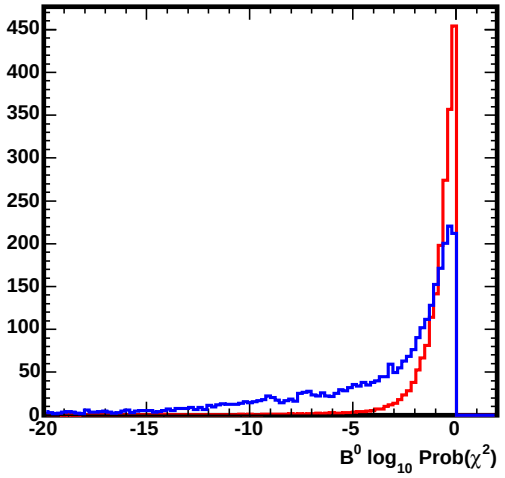
Figure 18: $B^0 \to J/\psi K^{*0}$ optimization of K^0_s mass window, with all other cuts applied. Measured inclusive $b \to J/\psi X$ spectrum from [19] used in Monte Carlo generation.



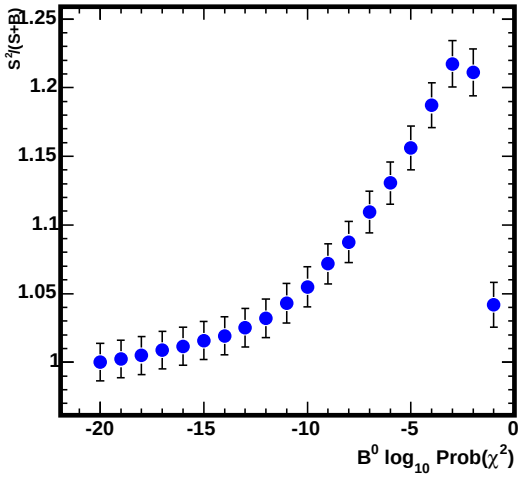
(a) B^0 mass in $B^0 \to J/\psi K^{*0}$ Monte Carlo. In red is the 2σ region used in the optimization.



(b) B^0 mass in data. In blue are the sideband regions used in the optimization. In green is the region used to fit for the lifetime.

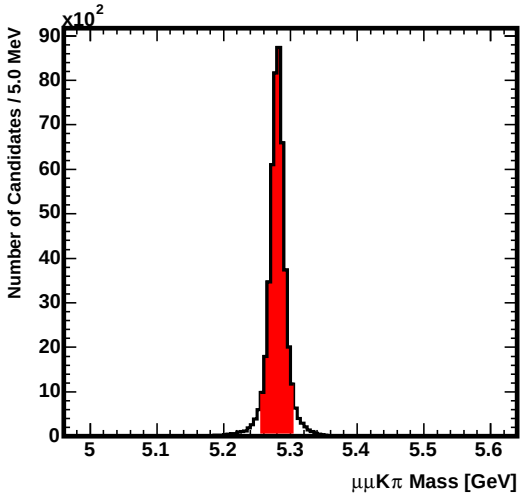


(c) Distribution of cut being optimized. In red is signal region from MC and blue is data sidebands, with normalization set by the relative signal yield in data and Monte Carlo.

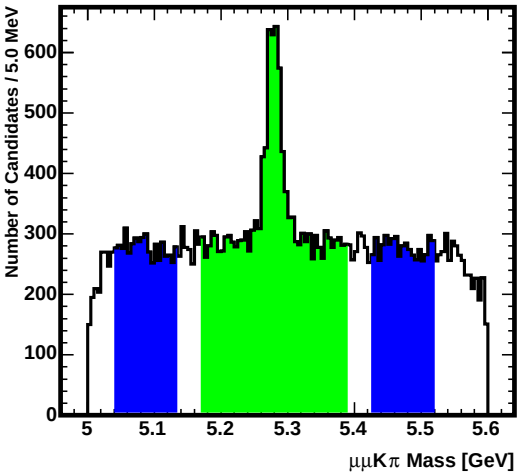


(d) Signal significance versus cut being optimized. The arrow indicates the optimal cut value.

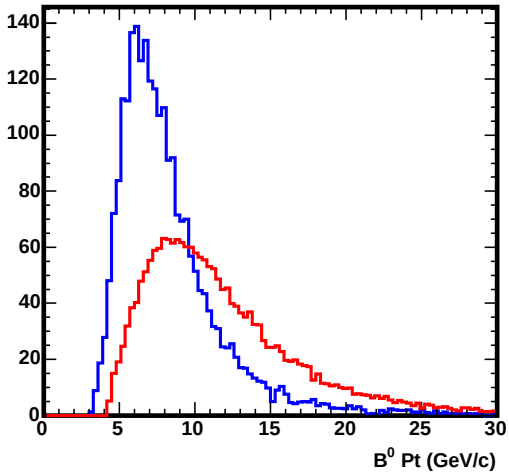
Figure 19: $B^0 \to J/\psi K^{*0}$ optimization of Prob(χ^2) for B^0 vertex fit, with all other cuts applied. Measured inclusive $b \to J/\psi X$ spectrum from [19] used in Monte Carlo generation.



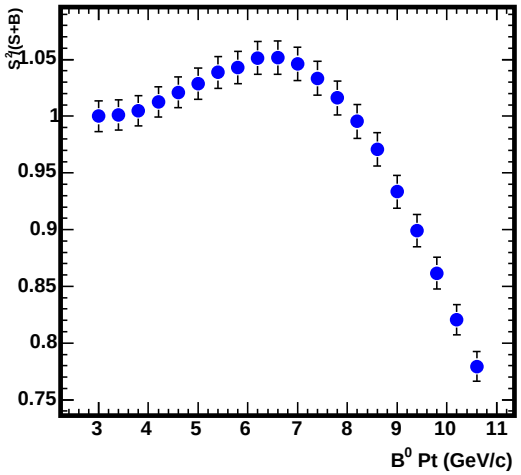
(a) B^0 mass in $B^0 \to J/\psi K^{*0}$ in Monte Carlo. In red is the 2σ region used in the optimization.



(b) B^0 mass in data. In blue are the sideband regions used in the optimization. In green is the region used to fit for the lifetime.

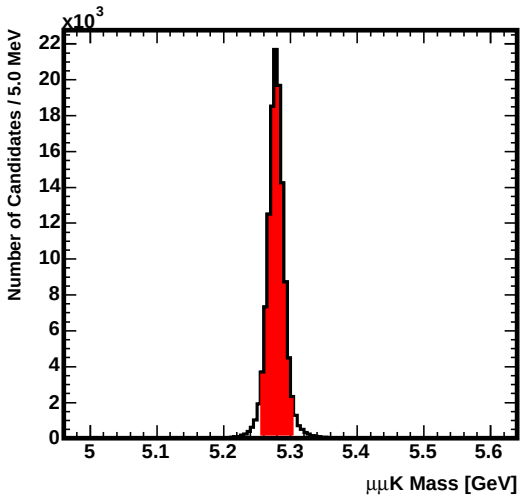


(c) Distribution of cut being optimized. In red is signal region from MC and blue is data sidebands, with normalization set by the relative signal yield in data and Monte Carlo.

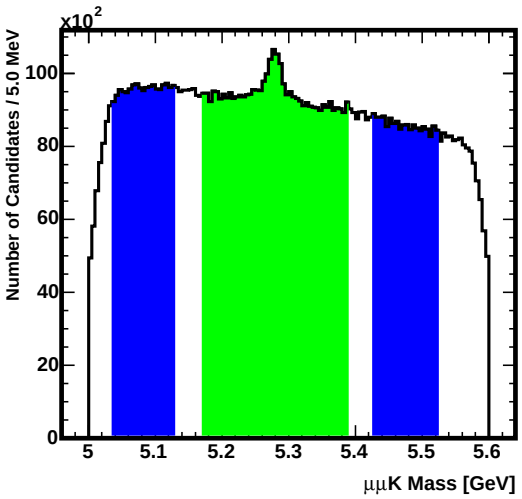


(d) Signal significance versus cut being optimized. The arrow indicates the optimal cut value.

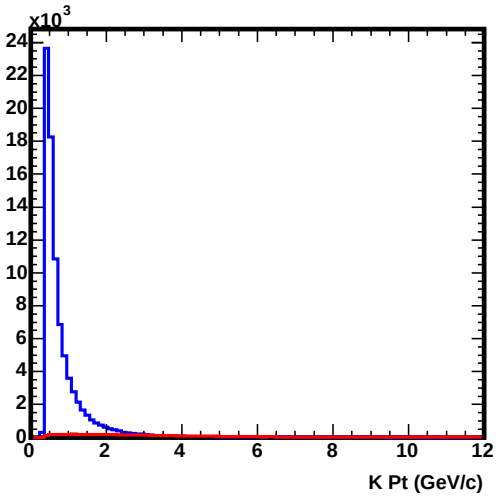
Figure 20: $B^0 \to J/\psi K^{*0}$ optimization of B^0 p_T, with all other cuts applied. Measured inclusive $b \to J/\psi X$ spectrum from [19] used in Monte Carlo generation.



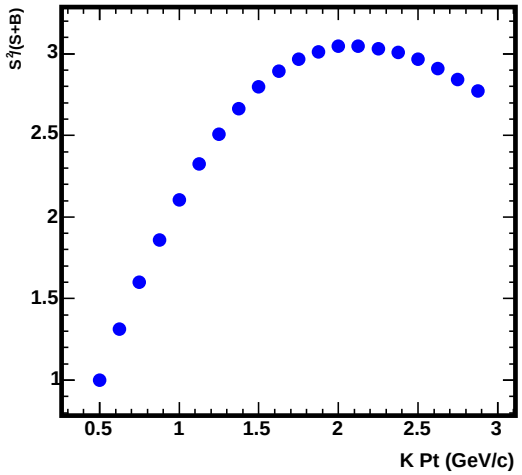
(a) Kaon p_T in $B^+ \to J/\psi K^+$ in Monte Carlo. In red is the 2σ region used in the optimization.



(b) B^+ mass in data. In blue are the sideband regions used in the optimization. In green is the region used to fit for the lifetime.

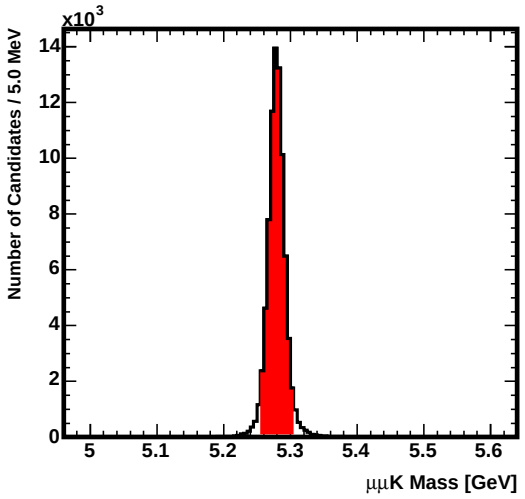


(c) Distribution of cut being optimized. In red is signal region from MC and blue is data sidebands, with normalization set by the relative signal yield in data and Monte Carlo.

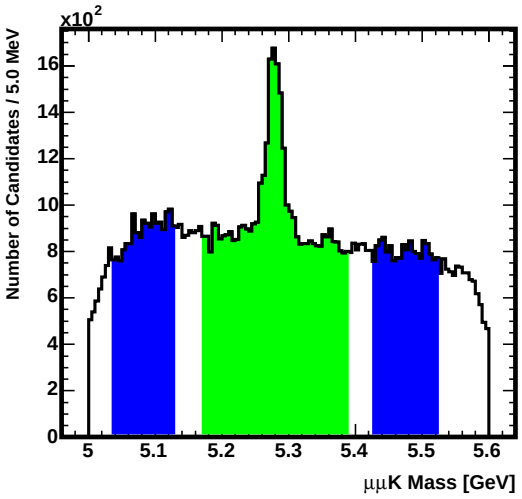


(d) Signal significance versus cut being optimized. The arrow indicates the optimal cut value.

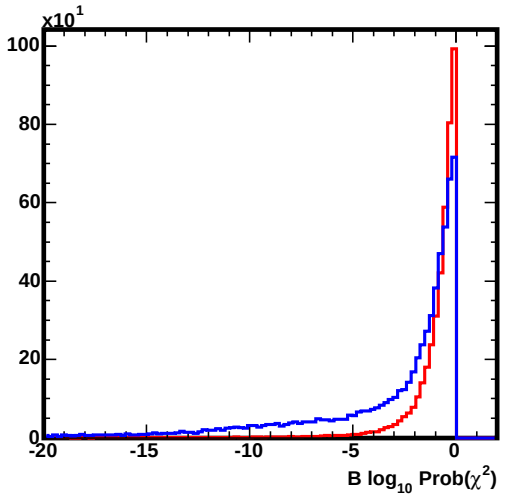
Figure 21: $B^+ \to J/\psi K^+$ optimization of K^0_s mass window, with all other cuts applied. Measured inclusive $b \to J/\psi X$ spectrum from [19] used in Monte Carlo generation.



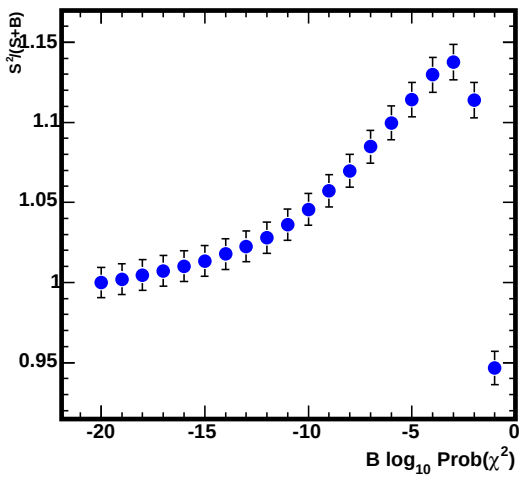
(a) B^+ mass in $B^+ \to J/\psi K^+$ Monte Carlo. In red is the 2σ region used in the optimization.



(b) B^+ mass in data. In blue are the sideband regions used in the optimization. In green is the region used to fit for the lifetime.

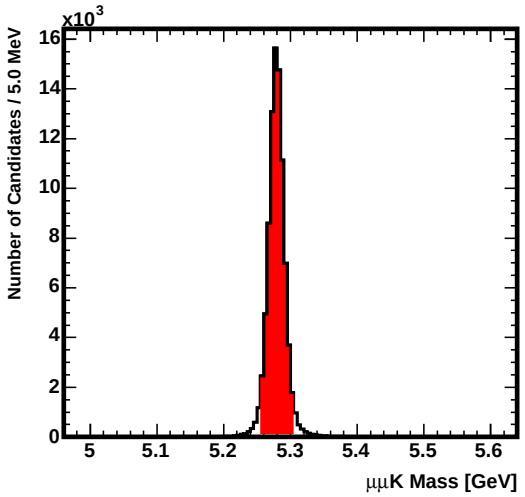


(c) Distribution of cut being optimized. In red is signal region from MC and blue is data sidebands, with normalization set by the relative signal yield in data and Monte Carlo.

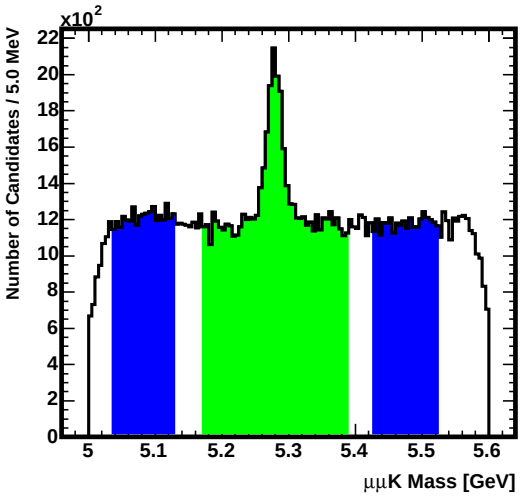


(d) Signal significance versus cut being optimized. The arrow indicates the optimal cut value.

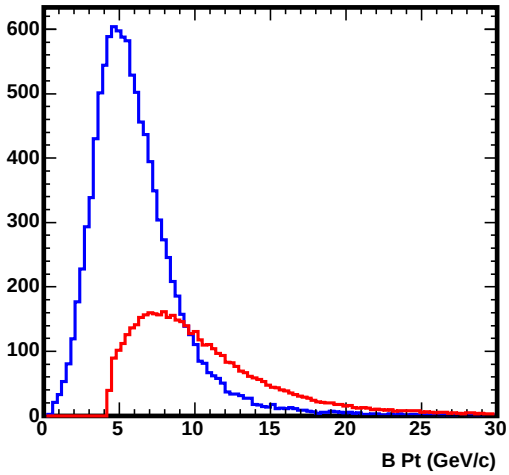
Figure 22: $B^+ \to J/\psi K^+$ optimization of Prob(χ^2) for B^+ vertex fit, with all other cuts applied. Measured inclusive $b \to J/\psi X$ spectrum from [19] used in Monte Carlo generation.



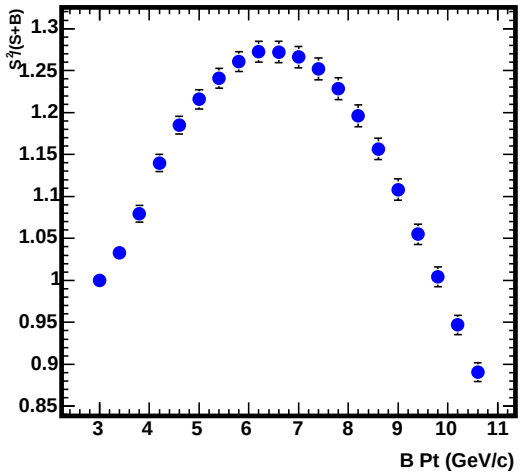
(a) B^+ mass in $B^+ \to J/\psi K^+$ in Monte Carlo. In red is the 2σ region used in the optimization.



(b) B^+ mass in data. In blue are the sideband regions used in the optimization. In green is the region used to fit for the lifetime.



(c) Distribution of cut being optimized. In red is signal region from MC and blue is data sidebands, with normalization set by the relative signal yield in data and Monte Carlo.



(d) Signal significance versus cut being optimized. The arrow indicates the optimal cut value.

Figure 23: $B^+ \to J/\psi K^+$ optimization of B^+ p_T, with all other cuts applied. Measured inclusive $b \to J/\psi X$ spectrum from [19] used in Monte Carlo generation.

	Cut value with maximum $S^2/(S+B)$				
Cut	$J/\psi\Lambda^0$ (Λ_b^0 excl.)	$J/\psi\Lambda^0$ (b incl.)	$J/\psi K_s^0$	$J/\psi K^{*0}$	$J/\psi K^+$
V^0 $L_{xy}/\sigma_{L_{xy}}$ from J/ψ	4.0	3.0	6.0	-	-
$V^0/K^{*0}/K^+$ p_T (GeV/c)	2.6	2.8	1.5	3.0	2.0
V^0/K^{*0} M_{win} (MeV/c^2)	± 9	± 9	± 25	± 65	-
b-hadron Prob(χ^2)	10^{-4}	10^{-4}	10^{-4}	10^{-3}	10^{-3}
b-hadron p_T (GeV/c)	4.0	4.0†	4.0†	6.4	6.2

Table 1: Optimized cut values with maximum $S^2/(S+B)$ for $\Lambda_b^0 \to J/\psi\Lambda^0$, $B^0 \to J/\psi K_s^0$, $B^0 \to J/\psi K^{*0}$, $B^+ \to J/\psi K^+$ based upon signal Monte Carlo and data sidebands. For Λ_b^0, "Λ_b^0 excl." and "b incl." refer input spectrum used for the signal Monte Carlo generation (see tex for details). † Fixed to this value throughout the optimization procedure.

	Cut Quantity	Value	
Muon Tracks	N($r-\phi$) SVX hits	≥ 3	
	$p_T(\mu)$(GeV)	≥ 1.5	
Non-Muon Tracks	N(COT-axial)*	≥ 2	
	N(COT-stereo)*	≥ 2	
Muons	Muon Type	CMU,CMP,CMUP,CMX (type ≤ 7)	
	χ^2_x for CMU, CMX	≤ 9	
J/ψ	mass (GeV)	3.104 $\leq$m($\mu\mu$)$\leq$ 3.174	
	2 track fit $P(\chi^2)$	> 0.001	
		for $B^0 \to J/\psi K^0_s$	for $\Lambda^0_b \to J/\psi\Lambda^0$
K^0_s, Λ	L_{xy}(cm)	≥ 0.1	$\geq$0.1
	2 track fit $P(\chi^2)$	$>$0.001	$>$0.001
	mass(GeV)	$0.472 \leq$ m ($\pi\pi$) ≤ 0.523	$1.107 \leq$ m($p\pi$) ≤ 1.125
	p_T(GeV)	$\geq$1.5	≥ 2.6
	$L_{xy}/\sigma_{L_{xy}}$	≥ 6.0	≥ 4.0
	Λ/K^0_s veto (GeV)	$1.1085 \leq$ m($p\pi$) ≤ 1.1235	$0.482 \leq$ m($\pi\pi$) ≤ 0.511
B^0, Λ^0_b	4 track fit $P(\chi^2)$	> 0.0001	> 0.0001
	p_T(GeV)	$\geq$4.0	$\geq$4.0

* N(COT-axial/stereo) refers the the number of axial/stereo superlayers with ≥ 5 hits.

Table 2: All cuts used in reconstruction of $B^0 \to J/\psi K^0_s$ and $\Lambda^0_b \to J/\psi\Lambda^0$

3.8 Selection Summary

A summary of all selection cuts used to reconstruct $B^0 \to J/\psi K^0_s$ and $\Lambda^0_b \to J/\psi\Lambda^0$ is shown in Table 2.[3]

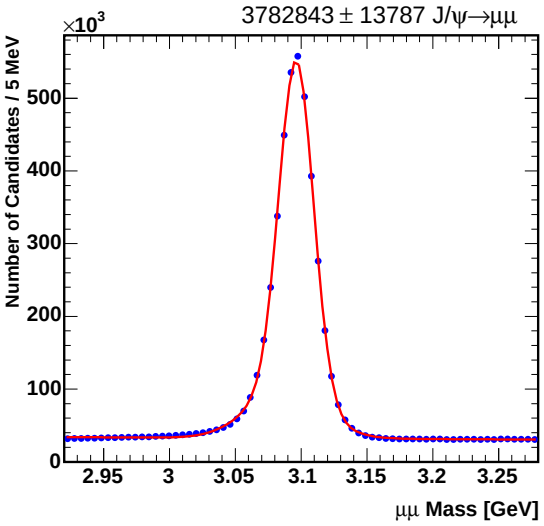
3.9 Selection Applied to Data

In the previous sections of this chapter, we presented our procedure for arriving at our selection. With this selection now set, we apply it to the dimuon-triggered dataset and in this section show the resulting signal yields.

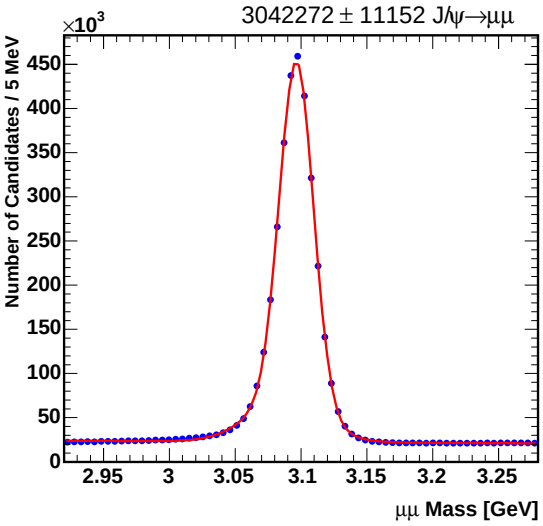
3.9.1 J/ψ and ψ(2S) Yields

Figure 24 shows the charmonium signals from our selection. We include the invariant mass distribution with and without requirement of 3 r $-\phi$ silicon hits on the muon tracks. The masses are fit to two Gaussian and the background to a linear function.

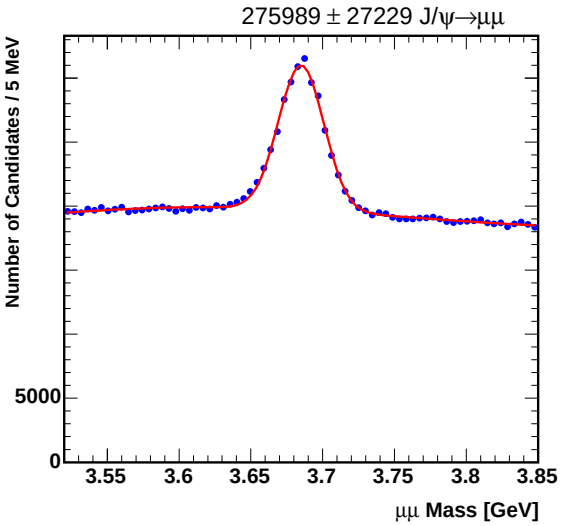
[3]In addition, before fitting the lifetime, we require that -0.2 $\leq c\tau$(cm) ≤ 0.4, $\sigma_{c\tau} \leq 100$ μm, and $\sigma_m \leq$ 0.02 GeV.



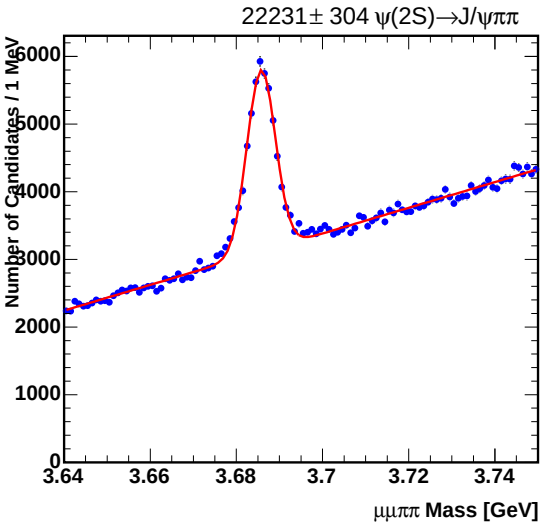
(a) J/ψ $\rightarrow$ $\mu^+\mu^-$ candidates (no silicon hits requirement)



(b) J/ψ $\rightarrow$ $\mu^+\mu^-$ candidates (requiring at least 3 r $-$ ϕ silicon hits per muon)

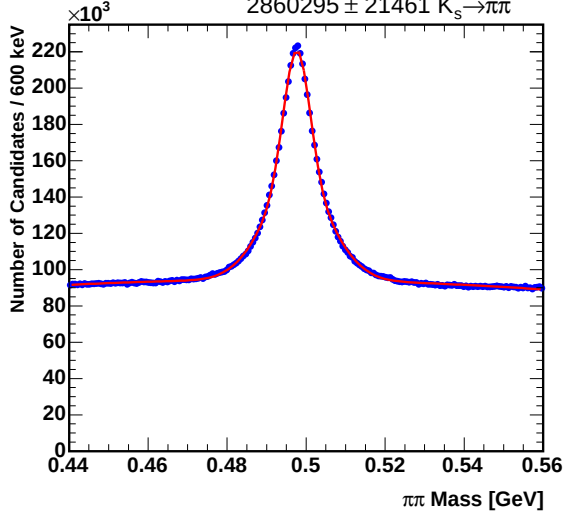


(c) ψ(2S) $\rightarrow$ $\mu^+\mu^-$ candidates (requiring at least 3 r $-$ ϕ silicon hits per muon)

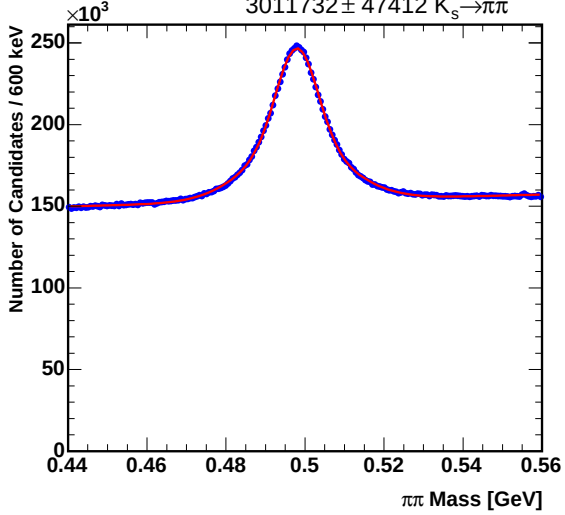


(d) ψ(2S) $\rightarrow$ J/$\psi\pi^+\pi^-$ candidates

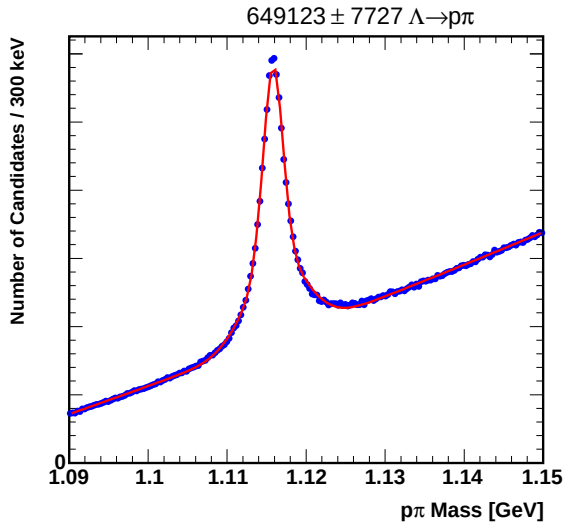
Figure 24: Charmonium signals



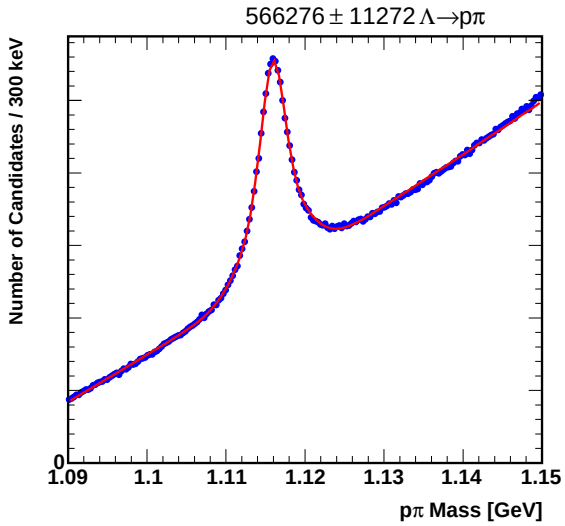
(a) $K^0_s \to \pi^+\pi^-$ candidates



(b) $K^0_s \to \pi^+\pi^-$ (COT-parent tracks) candidates



(c) $\Lambda^0 \to p\pi^-$ candidates



(d) $\Lambda^0 \to p\pi^-$ (COT-parent tracks) candidates

Figure 25: V^0 signals

3.9.2 K^0_s and Λ^0 Yields

Figure 25 shows the V^0 signals resulting from our selection. We show both the defTracks and COT parent tracks versions of these distributions for illustration purposes. Recall that comparing lifetime results with and without silicon hits on the V^0 tracks is an important check that silicon hits on V^0 tracks does not bias the b-hadron lifetimes. Note that the V^0 signal with COT parent tracks has both worse mass resolution and large background compared to deTracks with can have silicon hits attached to the tracks.

Note that these are the V^0 yields for V^0 produced in association with a dimuon pair as determined by the trigger rather than all V^0's produced in an unbiased way in CDF. Therefore, these are a combination of V^0's from b-hadrons and promptly produced V^0's that overlap with a dimuon candidate. Its also worth commenting that yields shown in Figure 25 are for scale and not to be taken seriously, since there are significant systematics

Decay mode	Signal yield
$J/\psi \to \mu^+\mu^-$	$\sim$ 3.8M
$J/\psi \to \mu^+\mu^-$ ($\geq$3 r $-$ ϕ silicon hits)	$\sim$ 3.0M
$\psi(2S) \to \mu^+\mu^-$	$\sim$ 276k
$\psi(2S) \to J/\psi\pi^+\pi^-$	$\sim$ 22k
$K^0_s \to \pi^+\pi^-$	$\sim$ 2.9M
$K^0_s \to \pi^+\pi^-$ (COT parent)	$\sim$ 3.0M
$\Lambda^0 \to p\pi^-$	$\sim$ 649k
$\Lambda^0 \to p\pi^-$ (COT parent)	$\sim$ 566k
$B^0 \to J/\psi K^0_s$ $(J/\psi \to \mu^+\mu^-)$	1225 $\pm$ 53
$B^0 \to \psi(2S)K^0_s$ $(\psi(2S) \to \mu^+\mu^-,\ K^0_s \to \pi^+\pi^-)$	85 $\pm$ 17
$B^0 \to \psi(2S)K^0_s$ $(\psi(2S) \to J/\psi\pi^+\pi^-,\ J/\psi \to \mu^+\mu^-,\ K^0_s \to \pi^+\pi^-)$	49 $\pm$ 9
$B^0 \to J/\psi K^{*0}$ $(J/\psi \to \mu^+\mu^-,\ K^{*0} \to K^+\pi^-)$	1938 $\pm$ 69
$B^0 \to \psi(2S)K^{*0}$ $(\psi(2S) \to \mu^+\mu^-,\ K^{*0} \to K^+\pi^-)$	105 $\pm$ 24
$B^0 \to \psi(2S)K^{*0}$ $(\psi(2S) \to J/\psi\pi^+\pi^-,\ J/\psi \to \mu^+\mu^-,\ K^{*0} \to K^+\pi^-)$	76 $\pm$ 13
$B^+ \to J/\psi K^+$ $(J/\psi \to \mu^+\mu^-)$	4209 $\pm$ 106
$B^+ \to \psi(2S)K^+$ $(\psi(2S) \to \mu^+\mu^-)$	372 $\pm$ 44
$B^+ \to \psi(2S)K^+$ $(\psi(2S) \to J/\psi\pi^+\pi^-,\ J/\psi \to \mu^+\mu^-)$	170 $\pm$ 19
$B^+ \to J/\psi K^{*+}$ $(J/\psi \to \mu^+\mu^-,\ K^{*+} \to K^0_s\pi^+)$	332 $\pm$ 24
$\Lambda^0_b \to J/\psi\Lambda^0$ $(J/\psi \to \mu^+\mu^-,\ \Lambda^0 \to p\pi^-)$	194 $\pm$ 23

Table 3: Summary of signal yields in xpmm0d

in the way one fits the mass distribution.

3.9.3 *b*-hadron Yields

Figures 26-30 show the b-hadron signals resulting from our selection. The $B^0 \to \psi(2S)K^0_s$ signals shown in Figures 29(a) and 29(b) are first observations at the Tevatron, to our best knowledge.

3.9.4 Yield Summary

Table 3 summarizes the yields we obtain with our selection applied to the xpmm0d dataset.

3.10 Monte Carlo Comparisons

Although the Monte Carlo does not play a central role in our lifetime analysis, it is used in optimization and important checks on the analysis procedure. Therefore, we check to see

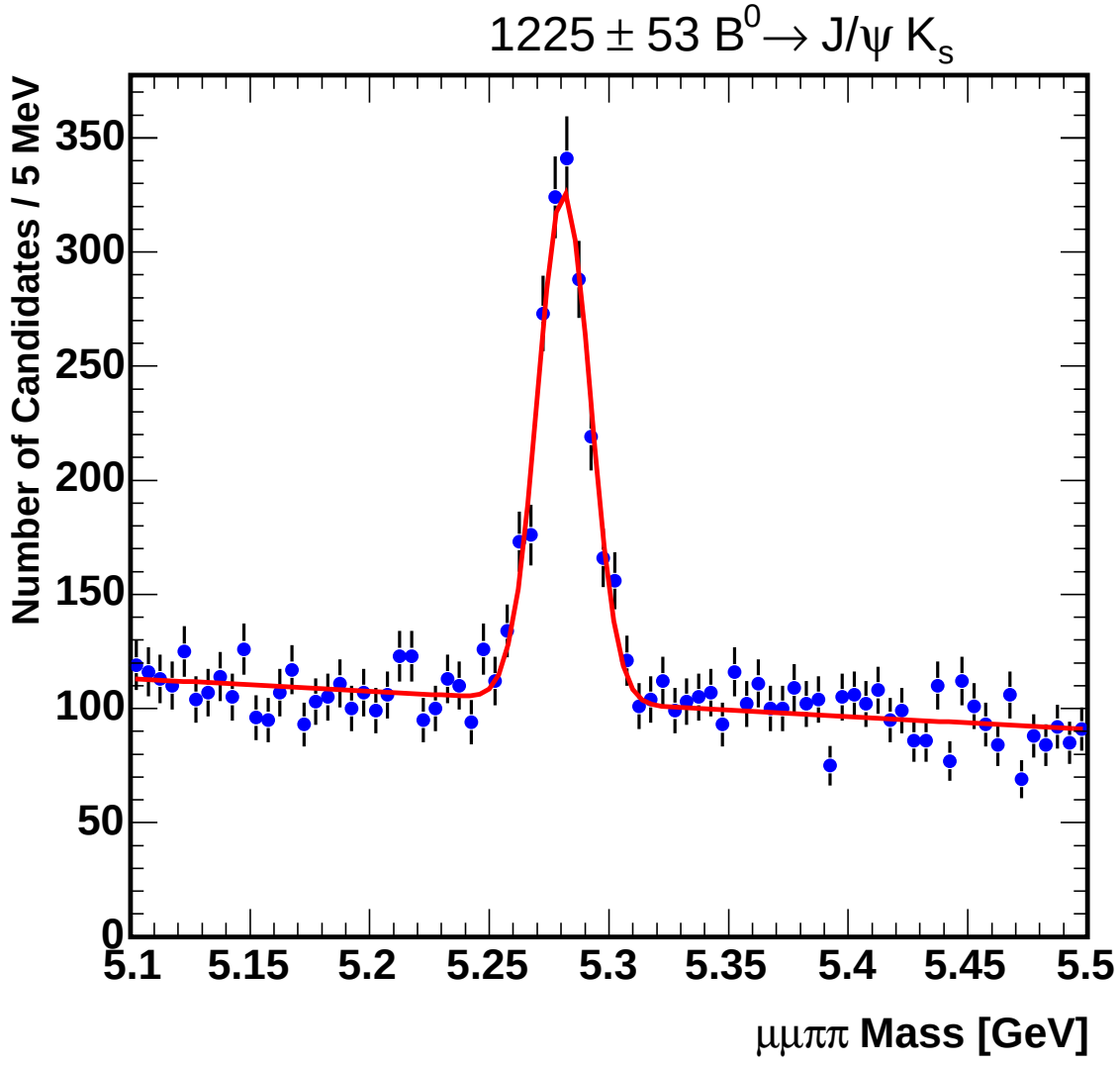


Figure 26: $B^0 \rightarrow J/\psi K^0_s$ Candidates

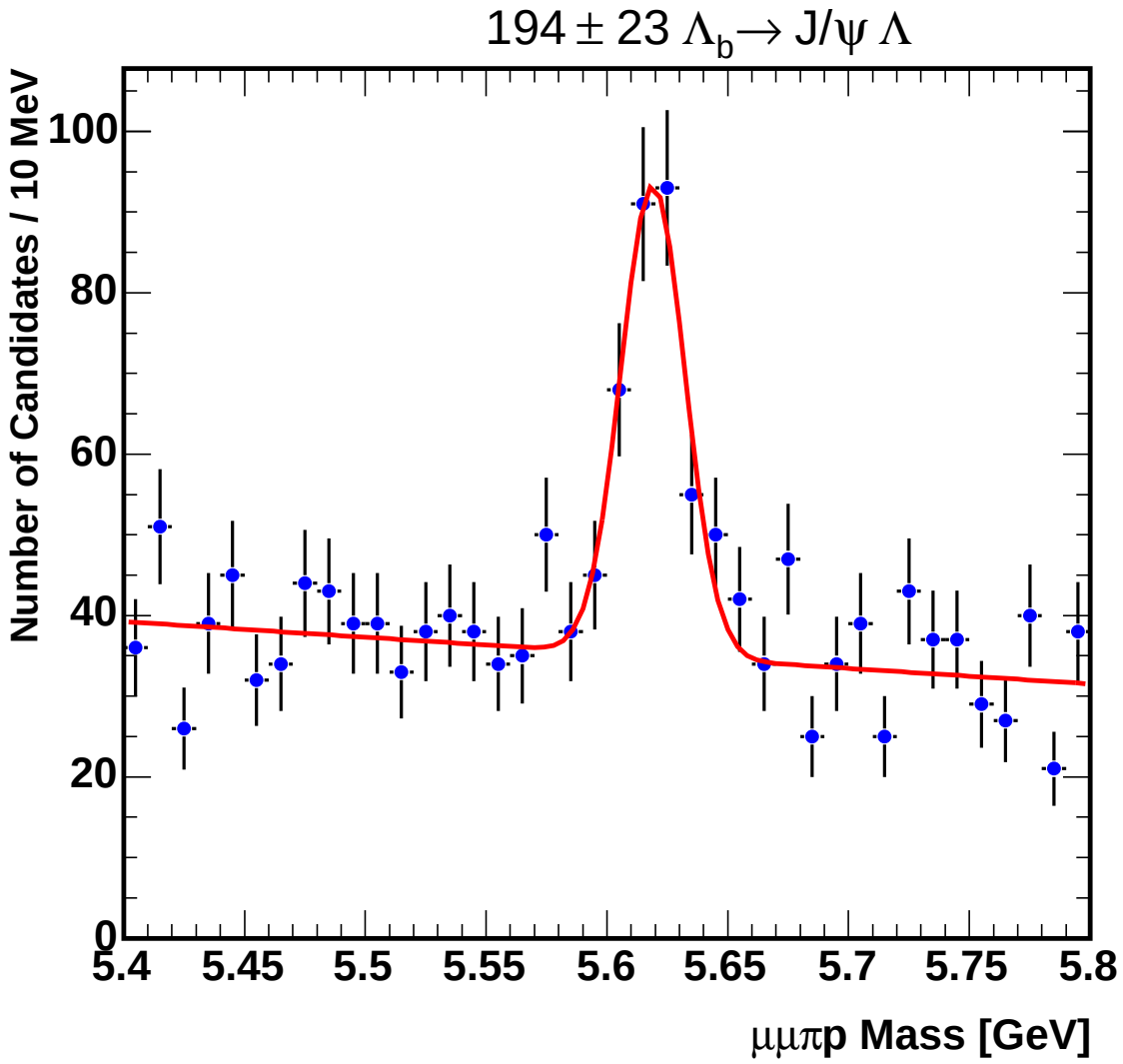


Figure 27: $\Lambda_b^0 \rightarrow J/\psi\Lambda^0$ Candidates

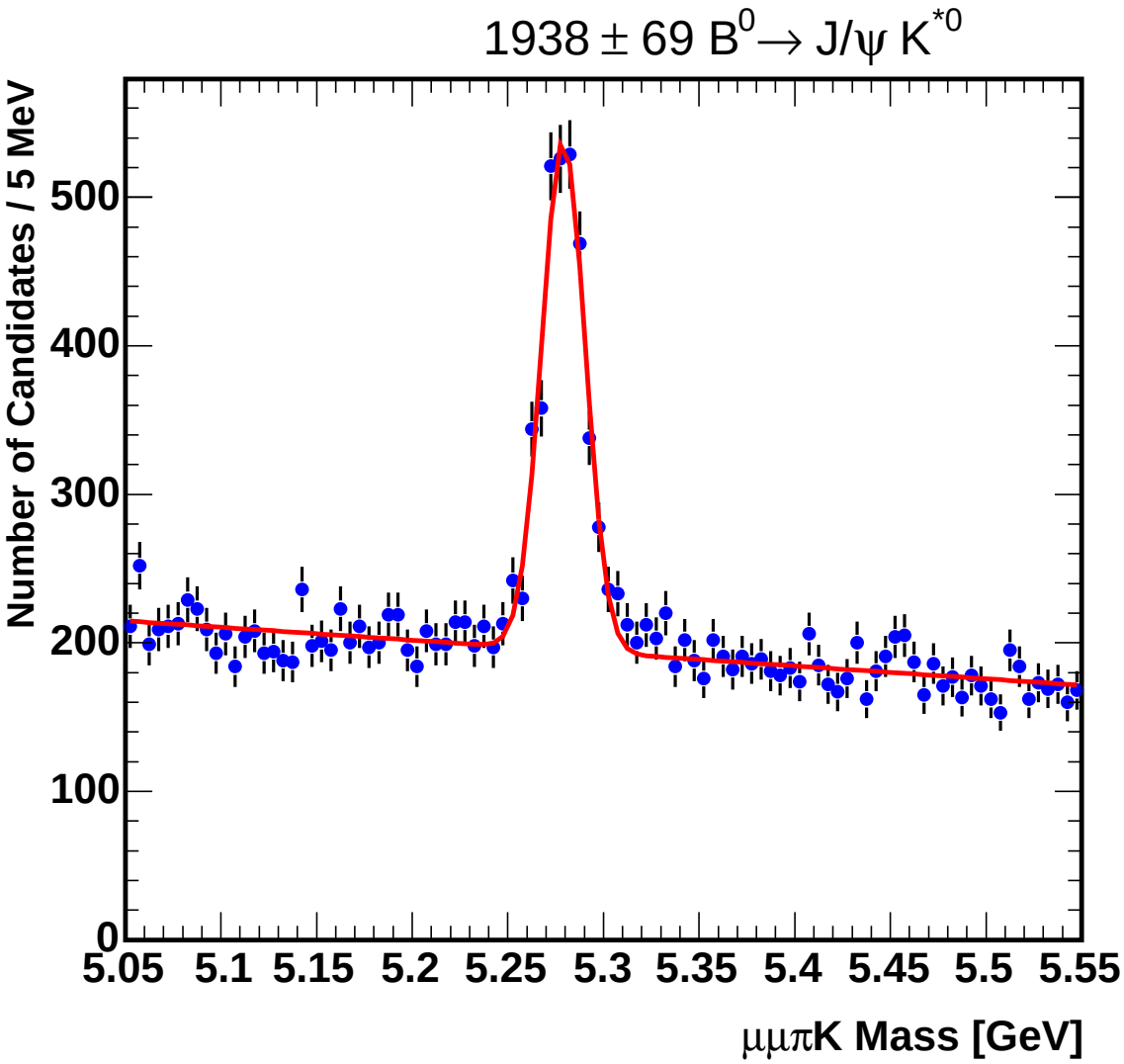
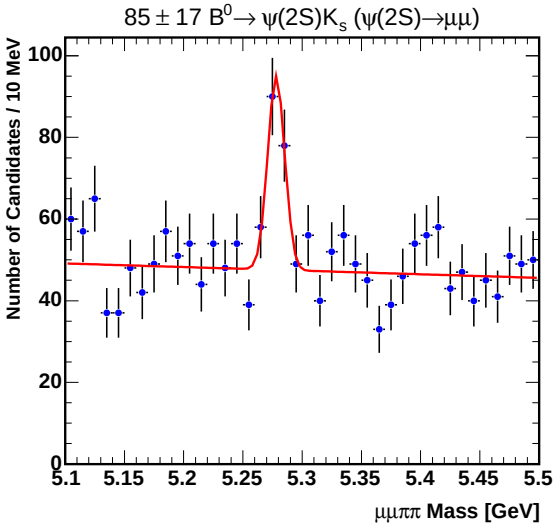
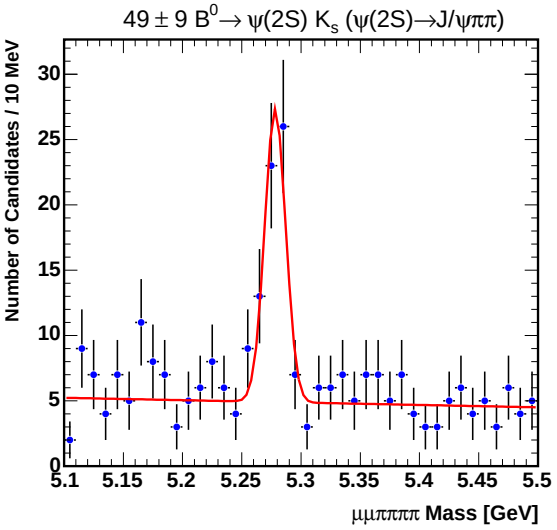


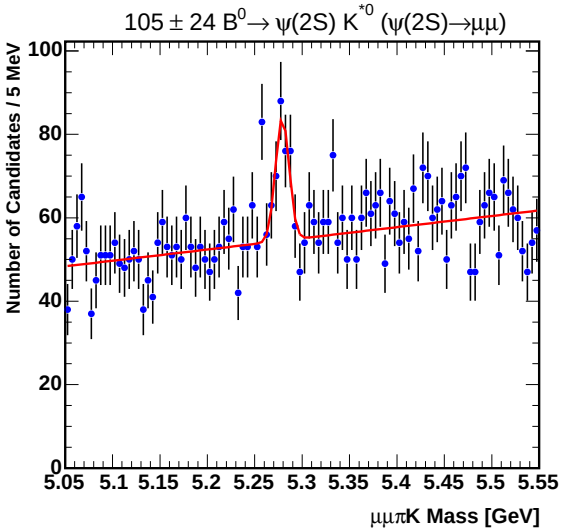
Figure 28: $B^0 \rightarrow J/\psi K^{*0}$ Candidates



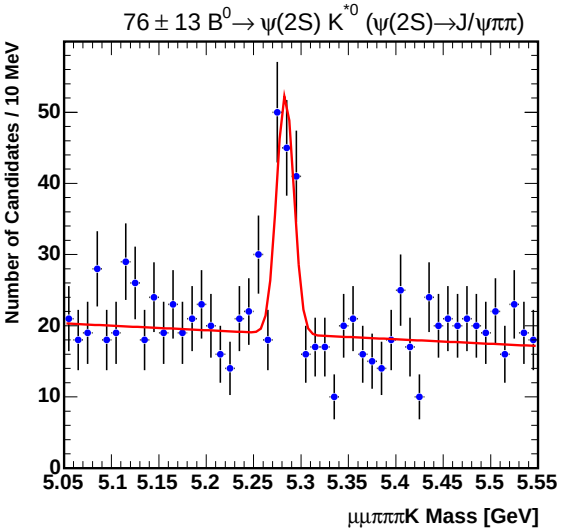
(a) $B^0 \to \psi(2S)K^0_s$ ($\psi(2S) \to \mu^+\mu^-$) Candidates



(b) $B^0 \to \psi(2S)K^0_s$ ($\psi(2S) \to J/\psi\pi^+\pi^-$) Candidates

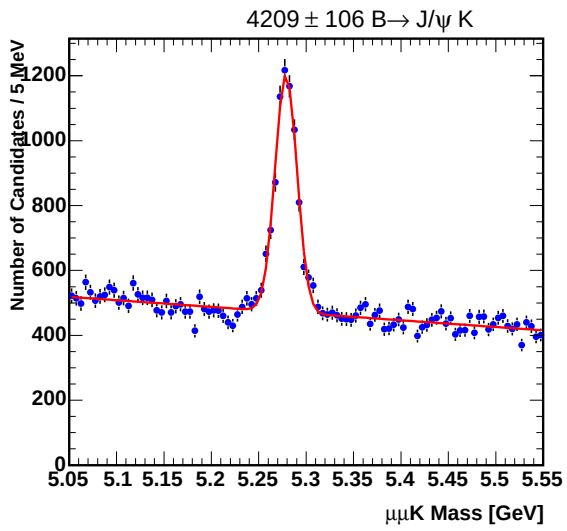


(c) $B^0 \to \psi(2S)K^{*0}$ ($\psi(2S) \to \mu^+\mu^-$) Candidates

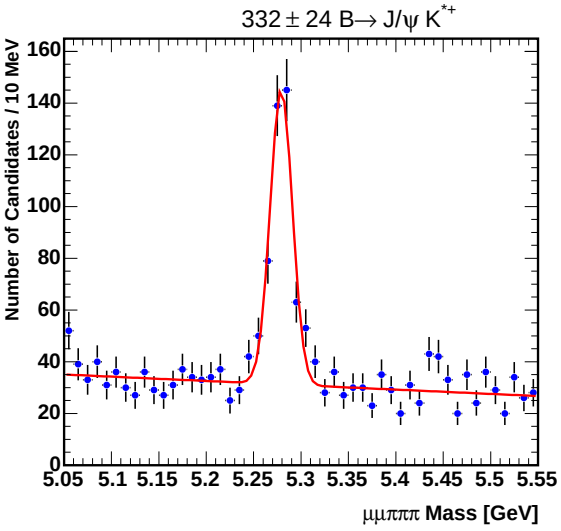


(d) $B^0 \to \psi(2S)K^{*0}$ ($\psi(2S) \to J/\psi\pi^+\pi^-$ Candidates

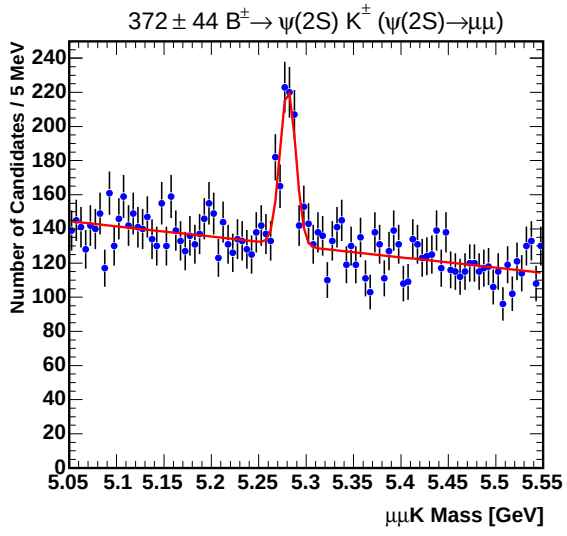
Figure 29: Signals from a variety of B^0 decay modes



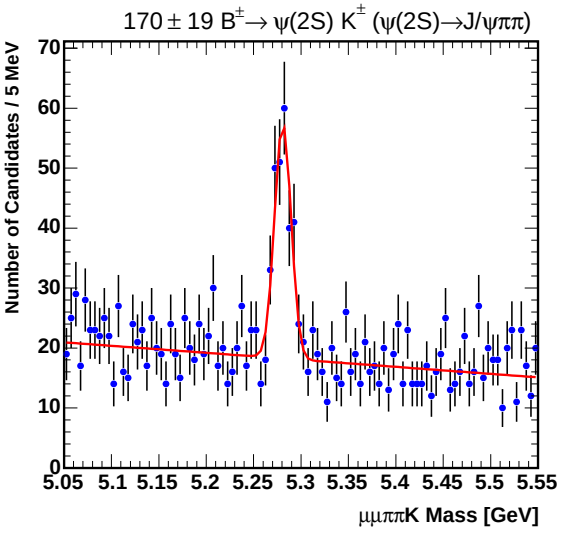
(a) $B^+ \to J/\psi K^+$ Candidates



(b) $B^+ \to J/\psi K^{*+}$ ($K^{*+} \to K^0_s \pi^+$) Candidates



(c) $B^+ \to \psi(2S)K^+$ ($\psi(2S) \to \mu^+\mu^-$) Candidates



(d) $B^+ \to \psi(2S)K^+$ ($\psi(2S) \to J/\psi\pi^+\pi^-$) Candidates

Figure 30: Signals from a variety of B^+ modes

how well our Monte Carlo models the data under identical reconstruction and selection procedure.

Figures 31-35 present comparison of sideband-subtracted signal distributions with Monte Carlo of the $B^0 \to J/\psi K^0_s$ and $\Lambda^0_b \to J/\psi \Lambda^0$ signals. Each of these plots shows three curves, sideband subtracted data (black data points with statistical error bars), signal Monte Carlo (blue line), and sideband data (red line). In addition, each plot shows the χ^2 comparison between sideband subtracted signal and signal BGenerator Monte Carlo. The agreement overall is quite good. Note that the B^0 and (and Λ^0_b) fit probabilities are far from flat, in contrast to all the subresonance fit probabilities. Though, that lack of flatness is reproduced by the Monte Carlo. Section 6.2 discusses the V^0 pointing systematics in some detail. The pull distributions shown there are clearly not unit gaussians, and thus explain the B^0 fit probability not being flat.

The Λ^0_b p_T spectrum deserves some extra focus. Figure 36 shows a comparison of the Λ^0_b p_T spectrum of our $\Lambda^0_b \to J/\psi \Lambda^0$ candidate in data to the Monte Carlo generated using the $\Lambda^0_b \to \Lambda_c \pi$ spectrum (Figure 36(a)) and the inclusive J/ψ spectrum (Figure 36(b)). Both spectra are good fits to our data and we do not current have the statistical precision to prefer one over the other.

4 Lifetime Fitting

In this section, we describe our procedure for determining the lifetime of the b-hadron signals in our data. This involves determining the proper decay length of the reconstructed b-hadron candidate and using a maximum likelihood fitting technique to extract the lifetime from the data. In addition to presenting of method for determine the lifetime, we also present some important tests to validate the fitting procedure.

4.1 Proper Decay Length Calculation

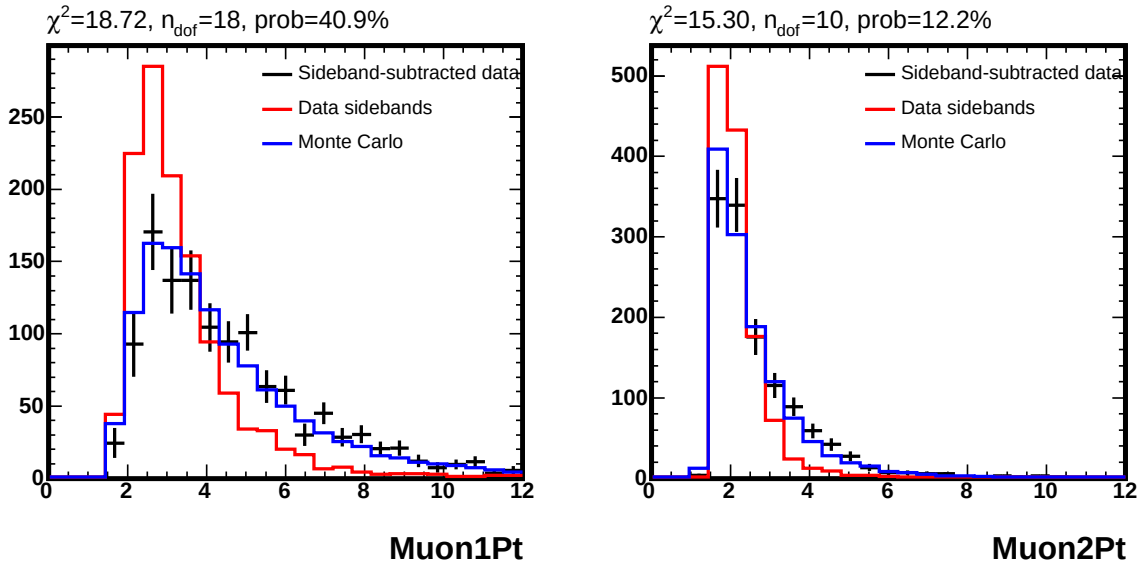
We determine the lifetime of a b-hadron by measuring the distance traveled by each b-hadron candidate in a plane transverse to the beam direction, L_{xy}, and then apply a correction for the Lorentz boost. To be specfic, the transverse decay length, L_{xy}, is given by

$$L_{xy} = \mathbf{L} \cdot \mathbf{p_T} / p_T \tag{2}$$

where $\mathbf{L}$ is the vector pointing from the primary to secondary (b-hadron) vertex and p_T is the transverse momentum vector of the b-hadron candidate. The proper decay length (PDL) for the b-hadron candidate is then given by

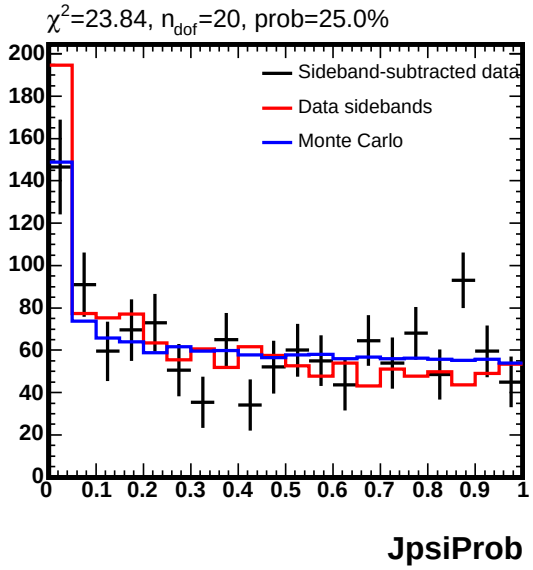
$$PDL = \frac{L_{xy}}{(\beta\gamma)^b_T} = L_{xy} \frac{c\ M_b}{p_T} \tag{3}$$

where $(\beta\gamma)^b_T$ and M_b are the transverse boost and mass of the b-hadron, respectively. We use the PDG mass for M_b in the PDL calculation.

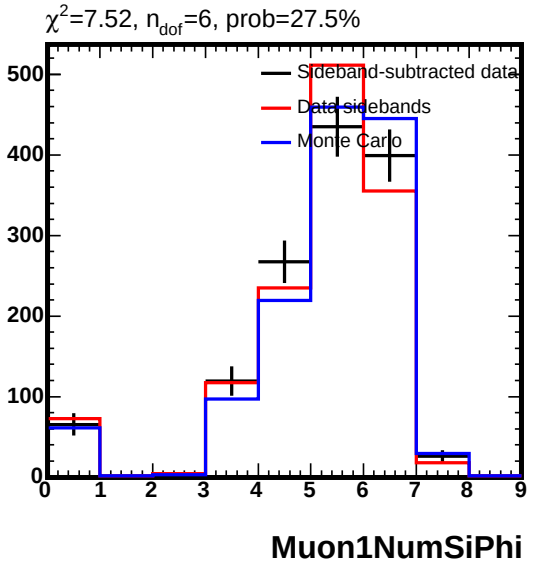


(a)

(b)

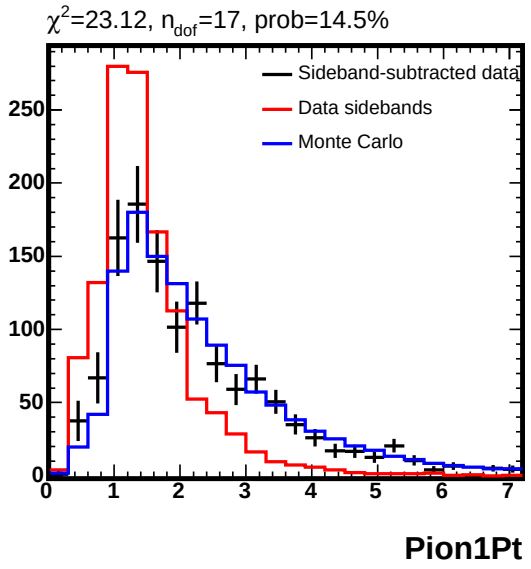


(c)

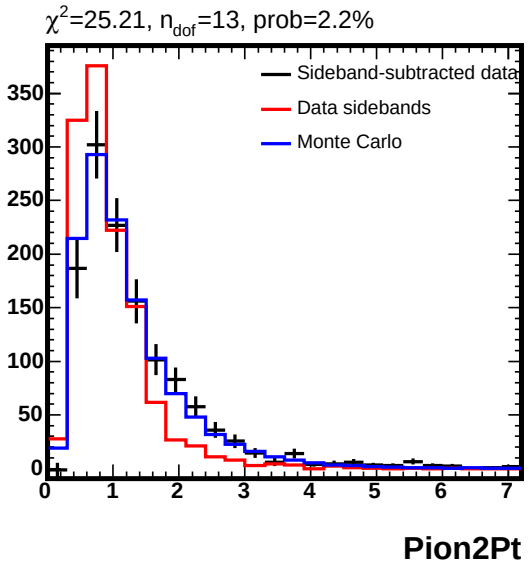


(d)

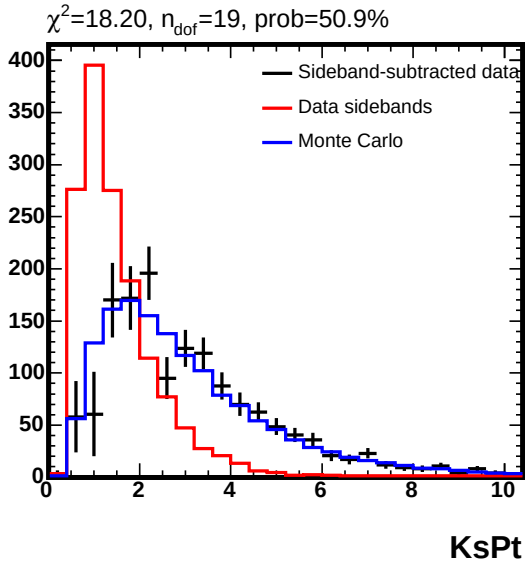
Figure 31: Data/Monte Carlo comparisons.



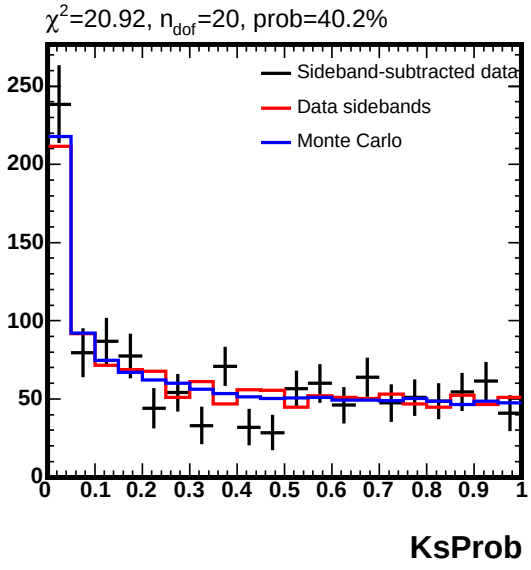
(a)



(b)

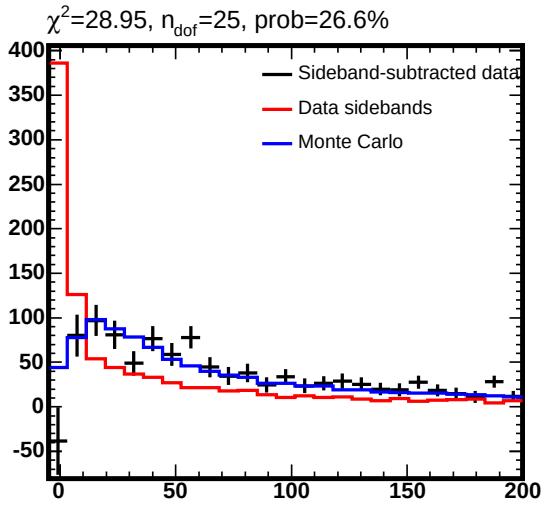


(c)

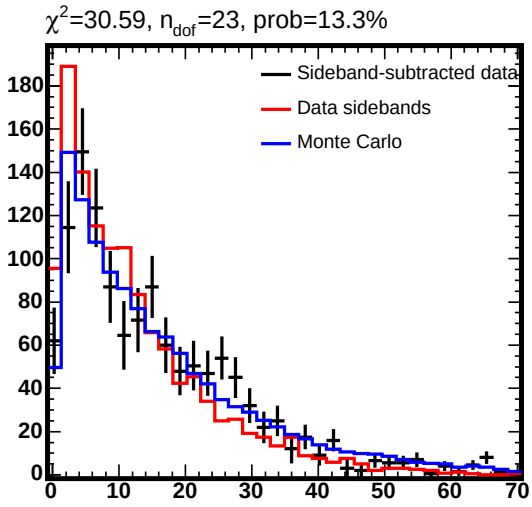


(d)

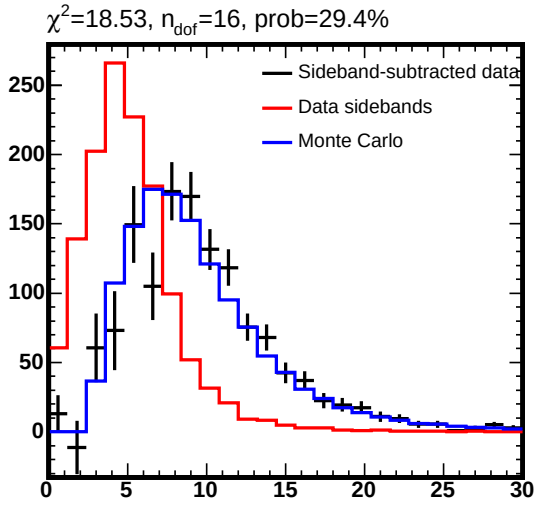
Figure 32: Data/Monte Carlo comparisons.



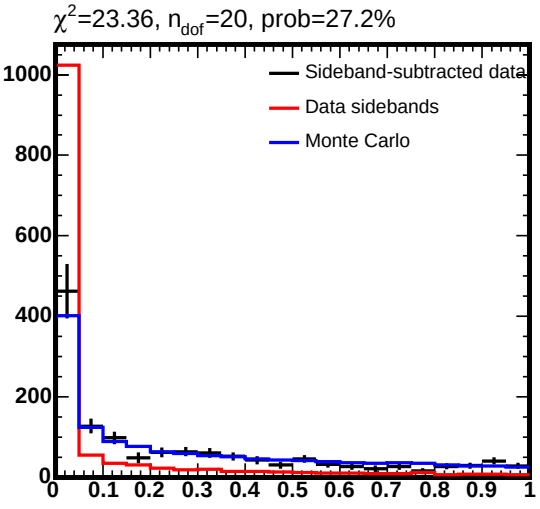
(a)



(b)



(c)



(d)

Figure 33: Data/Monte Carlo comparisons.

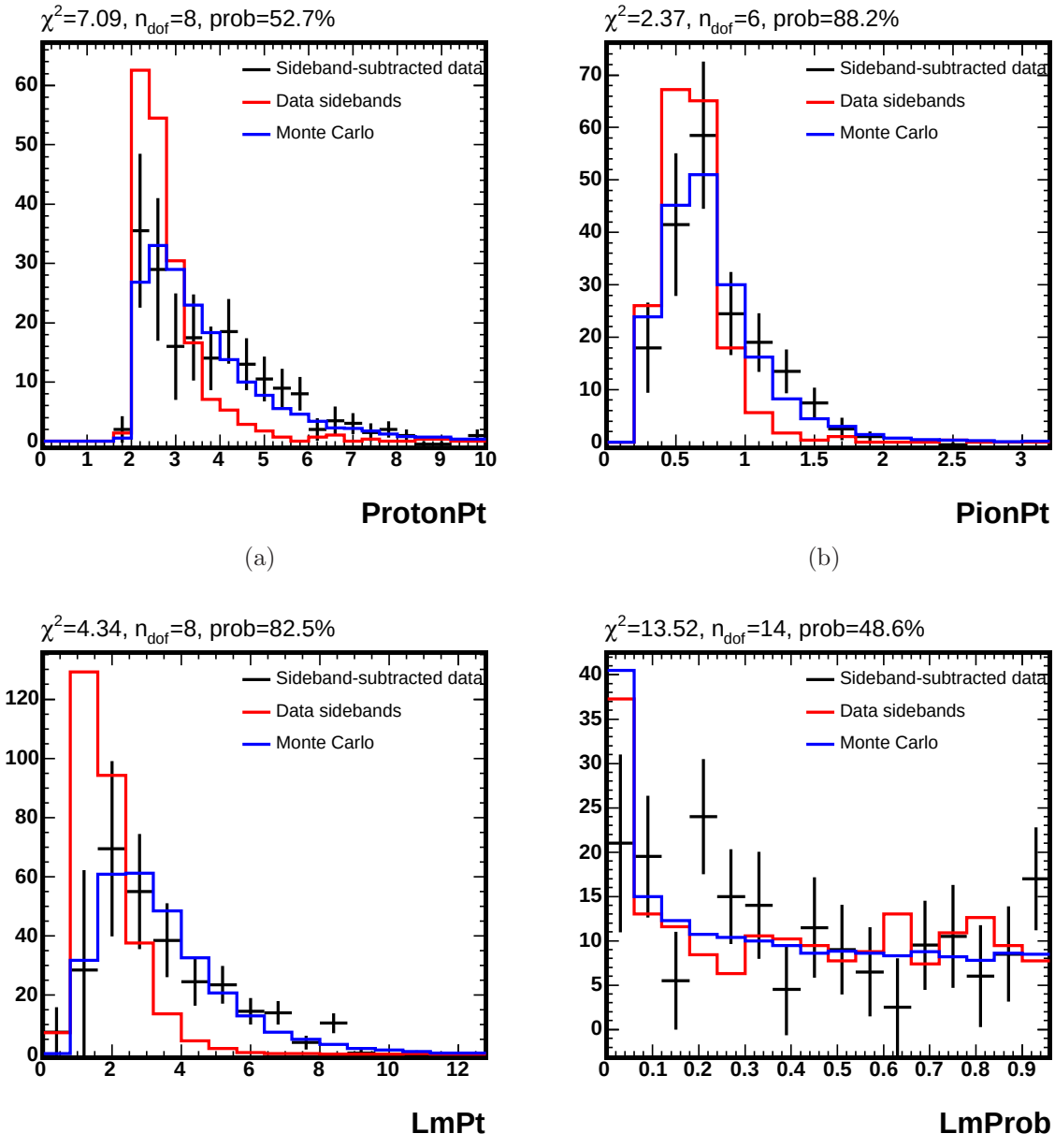
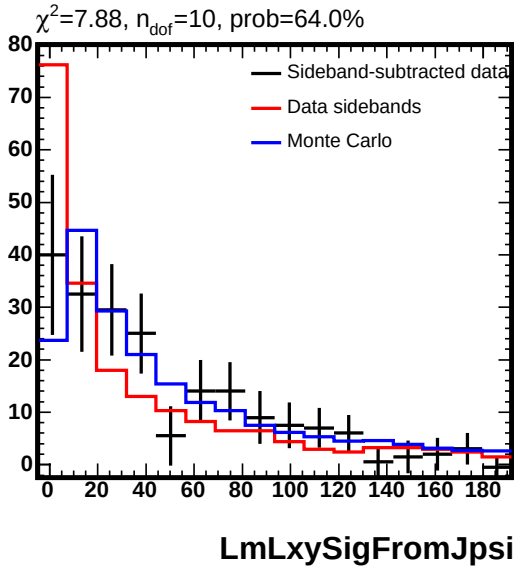
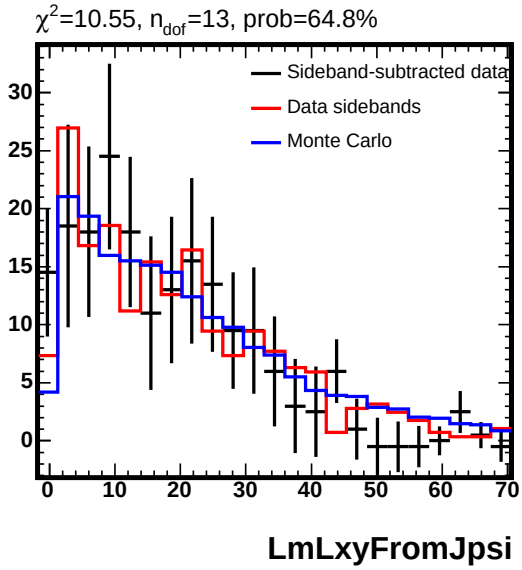


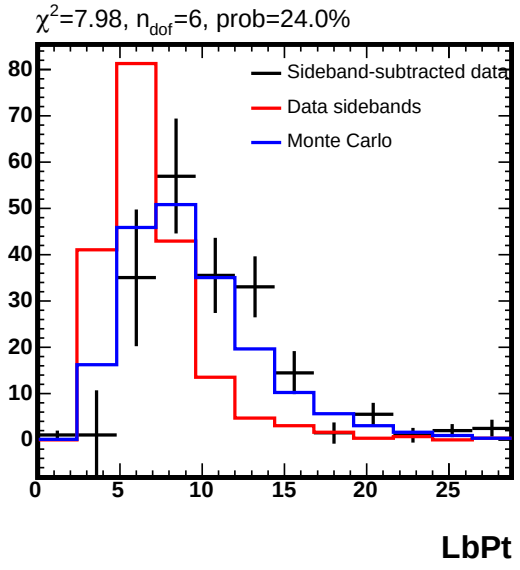
Figure 34: Data/Monte Carlo comparisons.



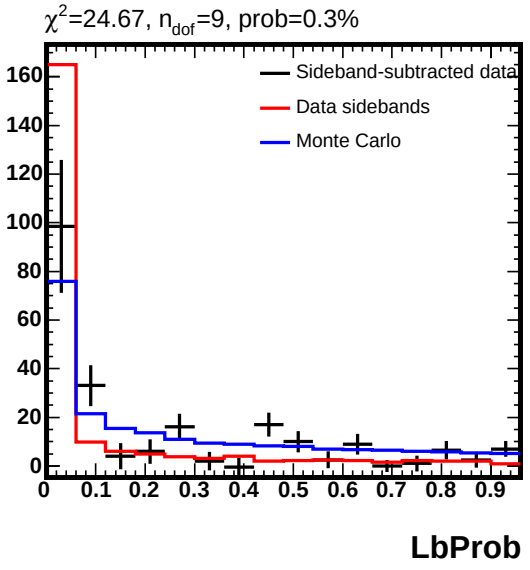
(a)



(b)

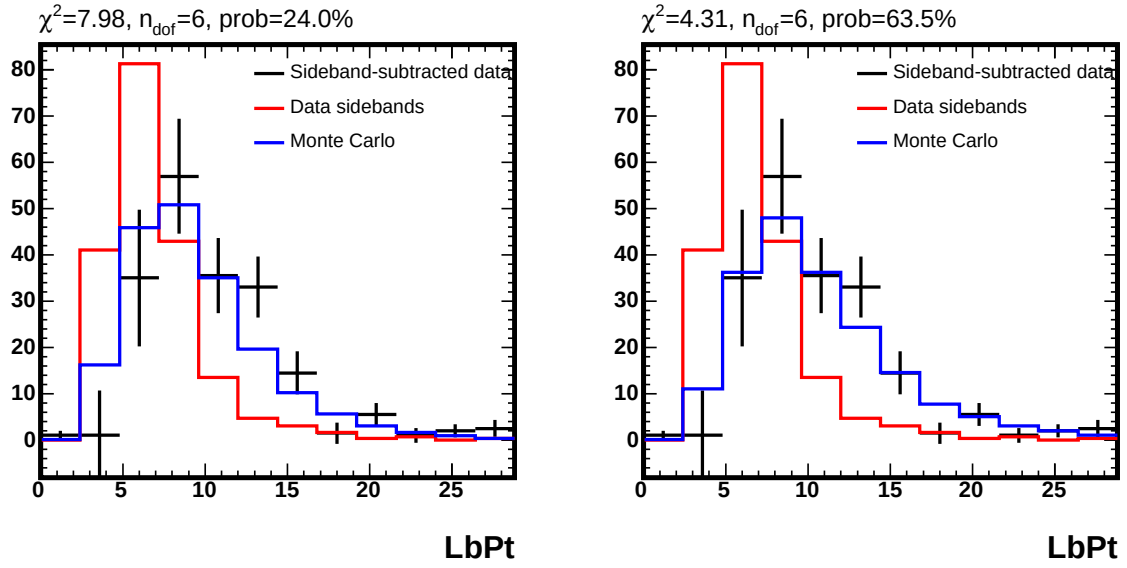


(c)



(d)

Figure 35: Data/Monte Carlo comparisons.



(a) Comparison of data and MC generated using $\Lambda_b^0 \to \Lambda_c \pi$ spectrum

(b) Comparison of data and MC generated using inclusive J/ψ spectrum

Figure 36: Data/Monte Carlo comparisons for Λ_b^0 p_T.

In order to minimize the effect of the V^0 tracking on the lifetime measurement, we construct the PDL using the J/ψ vertex for the distance measurements (i.e. $L_{xy} = L_{xy}^{J/\psi}$). The V^0 reconstruction then enters only via the momentum and direction of the b-hadron (see Figure 37). The primary vertex is the (x,y) of the time-dependent SVX beamline evaluated at z0 given by the average z0 of the J/ψ muon tracks.

Since the we use the non-V^0 b-hadron decay modes to validate the $B^0 \to J/\psi K_s^0$ and $\Lambda_b^0 \to J/\psi \Lambda^0$ lifetime procedure, we also use the J/ψ vertex in the PDL determination for these modes.

4.2 Fit Model

We use a unbinned maximum likelihood fit to our b-hadron candidate PDL and mass data to extract the lifetime. The model used to fit the data is motivated by Monte Carlo studies and previous work in CDF on lifetimes. There is an important difference, however, between our fitting procedure and all other previous CDF lifetime analyses of which we are aware. This issue is addressed in Section 4.7.

The overall PDF used to fit the data is a sum of signal and background contributions, weighted by the fraction of signal, f_s, and background, f_b, present in the data. If the signal and background functions are each normalized (i.e. PDF's) and we use the fact

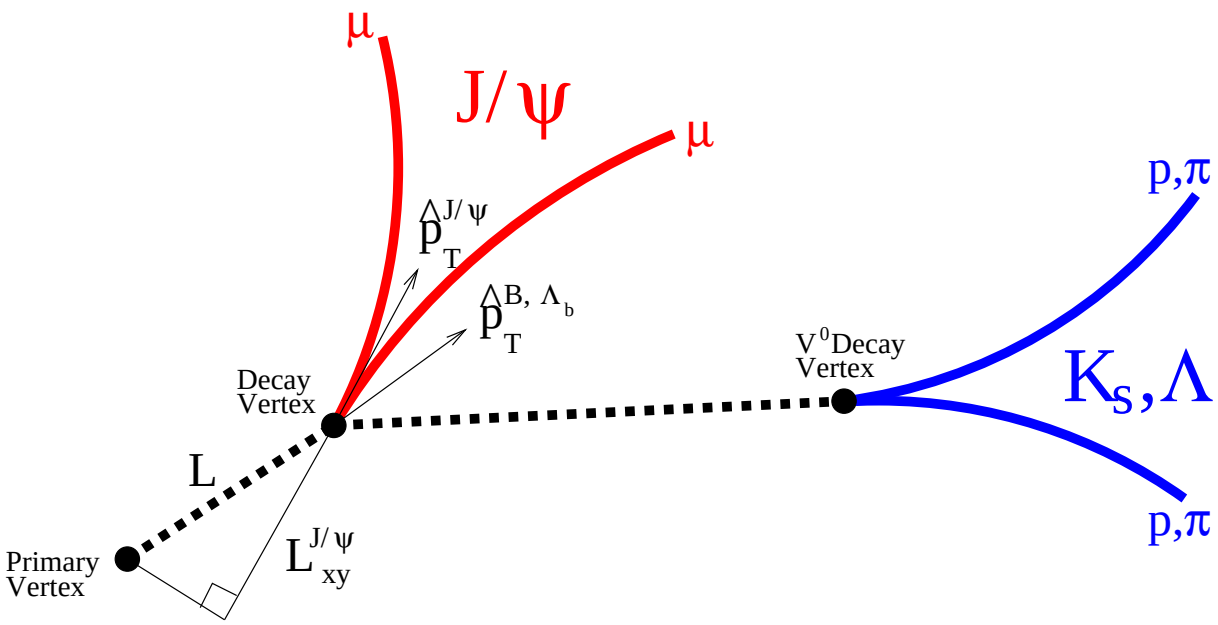


Figure 37: Decay topology in $b \to \mathrm{J}/\psi \mathrm{V}^0$ for PDL calculation

that $\mathrm{f_b} + \mathrm{f_s} = 1$, the overall model is a properly normalized PDF given by

$$\mathcal{P}_{tot}(\lambda_\mathrm{i}, \sigma_\mathrm{i}^\lambda, \mathrm{m_i}, \sigma_\mathrm{i}^\mathrm{m}|\tilde{\xi}) = (1 - \mathrm{f_b})\mathcal{P}_\mathrm{sig} + \mathrm{f_b}\mathcal{P}_\mathrm{bkg} \tag{4}$$

where (the subscript i represents the i^{th} candidate)

$$\begin{aligned}
\lambda_\mathrm{i} &= \text{measured proper decay length} \\
\sigma_\mathrm{i}^\lambda &= \text{error on measured proper decay length} \\
\mathrm{m_i} &= \text{measured mass} \\
\sigma_\mathrm{i}^\mathrm{m} &= \text{error on measured mass} \\
\mathcal{P}_{sig}, \mathcal{P}_{bkg} &= \text{ signal and background PDF's} \\
\mathrm{f_b} &= \text{total background fraction} \\
\vec{\xi} &= \mathrm{f_b} \text{ and the set of fit parameters upon which } \mathcal{P}_\mathrm{sig} \text{ and } \mathcal{P}_\mathrm{bkg} \text{ depend}
\end{aligned} \tag{5}$$

In general, the signal and background PDFs are products of PDFs for lifetime, $\mathcal{P}_{sig,bkg}^\lambda$, and mass $\mathcal{P}_{sig,bkg}^m$,

$$\mathcal{P}_{sig,bkg} = \mathcal{P}_{sig,bkg}^\lambda(\lambda_\mathrm{i}, \sigma_\mathrm{i}^\lambda|\tilde{\alpha}_\mathrm{sig,bkg})\mathcal{P}_\mathrm{sig,bkg}^\mathrm{m}(\mathrm{m_i}, \sigma_\mathrm{i}^\mathrm{m}|\tilde{\beta}_\mathrm{sig,bkg}) \tag{6}$$

where $\vec{\alpha}_{sig,bkg}$ and $\vec{\beta}_{sig,bkg}$ are the set of parameters that specify the PDF's for signal and background. Using Bayes Theorem, we can rewrite a given PDF as

$$\mathcal{P}(x, \sigma^x|\vec{\gamma}) = \mathcal{P}(x|\sigma^x, \vec{\gamma})P(\sigma^x|\vec{\gamma}) \tag{7}$$

so that Equation 6 now becomes

$$\mathcal{P}_{sig,bkg} = \mathcal{P}^{\lambda}_{sig,bkg}(\lambda_{\rm i}|\sigma^{\lambda}_{\rm i},\tilde{\alpha}_{\rm sig,bkg})\mathcal{P}^{\sigma^{\lambda}}_{\rm sig,bkg}(\sigma^{\lambda}_{\rm i}|\tilde{\alpha}_{\rm sig,bkg})\mathcal{P}^{\rm m}_{\rm sig,bkg}({\rm m_i}|\sigma^{\rm m}_{\rm i},\tilde{\beta}_{\rm sig,bkg})\mathcal{P}^{\sigma^{\rm m}}_{\rm sig,bkg}(\sigma^{\rm m}_{\rm i}|\tilde{\beta}_{\rm sig,bkg}) \tag{8}$$

It is *very* important to note that if, and only if, any of these PDF's are the same for signal and background can they be safely factored out of Equation 8 as an overall constant in the likelihood. To do otherwise, the fitting procedure will have biases in its parameters, as pointed out by Giovanni Punzi in [17]. More on this later. In the mean time, we demonstrate using the error distributions from our data that one can safely ignore that difference in mass error, but not PDL error.

Figure 38 shows a comparison of background and sideband-subtracted signal mass error distributions for $\rm B^0 \to J/\psi K^0_s$, $\rm \Lambda^0_b \to J/\psi \Lambda^0$, $\rm B^0 \to J/\psi K^{*0}$, and $\rm B^+ \to J/\psi K^+$ decays. Note that that signal and background distributions are very nearly statistically indistinguishable, implying that the mass error PDF can be safetly factored out of our overall PDF for the lifetime fit. This is not the case, however, for the PDL error where signal and background are clearly different in shape, as shown in Figure 39.

This means that we can now simplify Equation 8 to

$$\mathcal{P}_{sig,bkg} = \mathcal{P}^{\lambda}_{sig,bkg}(\lambda_{\rm i}|\sigma^{\lambda}_{\rm i},\tilde{\alpha}_{\rm sig,bkg})\mathcal{P}^{\sigma^{\lambda}}_{\rm sig,bkg}(\sigma^{\lambda}_{\rm i}|\tilde{\alpha}_{\rm sig,bkg})\mathcal{P}^{\rm m}_{\rm sig,bkg}({\rm m_i}|\sigma^{\rm m}_{\rm i},\tilde{\beta}_{\rm sig,bkg}) \tag{9}$$

Our overall PDF is now a simultaneous fit to PDL, PDL error, mass, and mass error data using the sum of signal and background PDF's that are each products of independent PDL, PDL error, and mass PDF's. These PDF's are described in turn in the following sections.

4.3 Proper Decay Length PDF

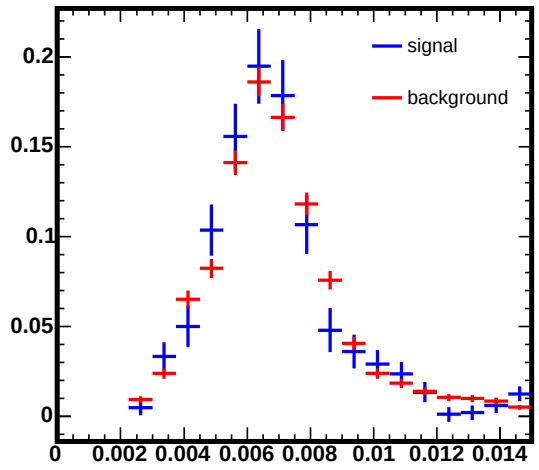
In this section we describe $\mathcal{P}^{\lambda}_{sig}$ and $\mathcal{P}^{\lambda}_{bkg}$, the PDF's for the lifetime part of the signal and background.

4.3.1 Signal

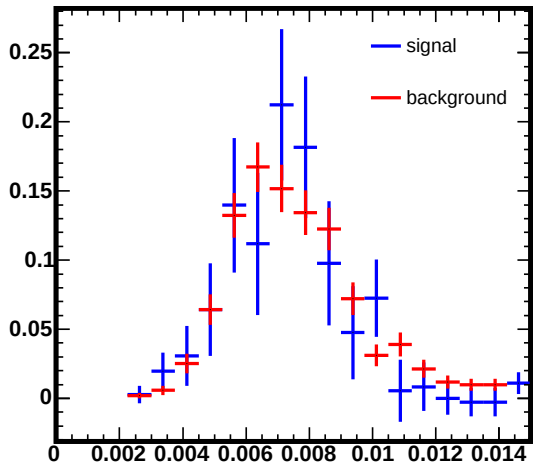
The signal we are trying to measure is the lifetime of a particle, $c\tau$, from its observed proper decay length and error. Therefore, the signal PDF is an exponential decay convoluted with the detector resolution (response) function.

From the shape of the zero-lifetime background, one can see that the resolution function is reasonably well approximated by a single Gaussian, although some non-Gaussian tails are evident in the data. The functional form we use as our signal PDF is an exponential convoluted with a Gaussian resolution function ($\mu = 0$),

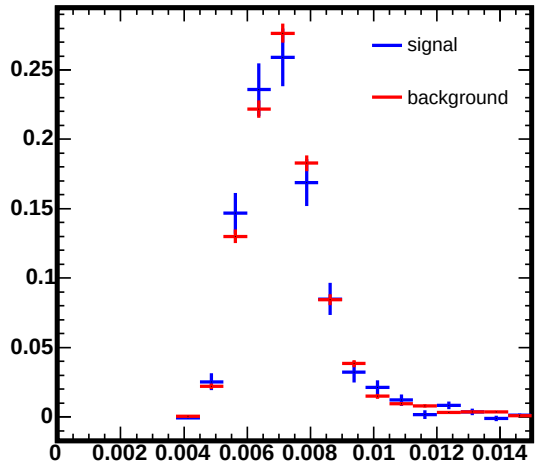
$$\mathcal{P}^{\lambda}_{sig}(\lambda_{\rm i},\sigma^{\lambda}_{\rm i}|\tilde{\alpha}_{\rm sig}) = {\rm E}(\lambda_{\rm i}|{\rm c}\tau) \otimes {\rm G}(\lambda_{\rm i},\sigma^{\lambda}_{\rm i}|{\rm s}) \tag{10}$$



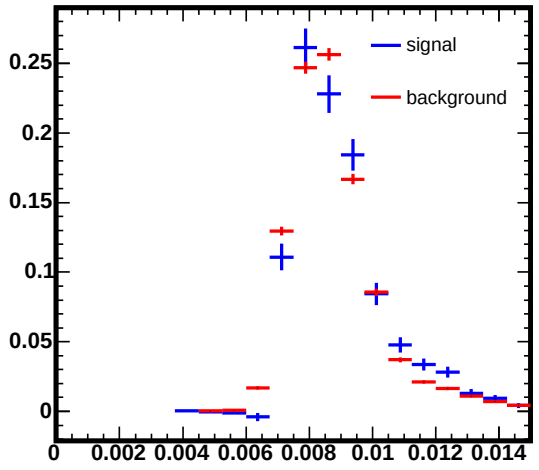
(a) $B^0 \to J/\psi K^0_s$



(b) $\Lambda^0_b \to J/\psi \Lambda^0$

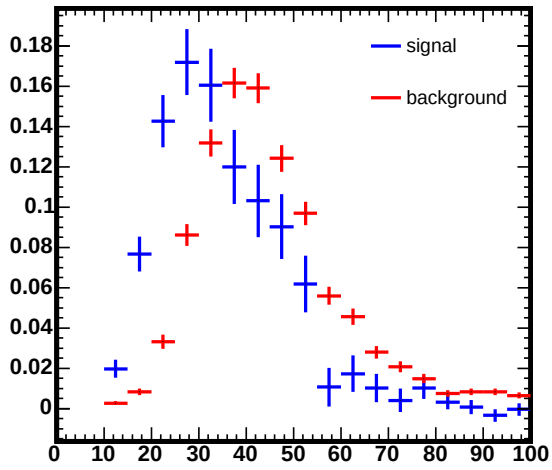


(c) $B^0 \to J/\psi K^{*0}$

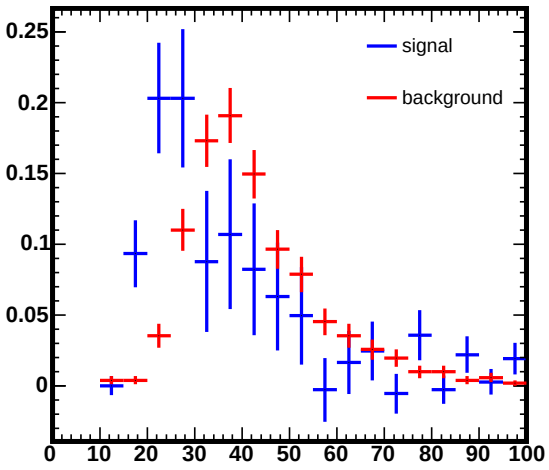


(d) $B^+ \to J/\psi K^+$

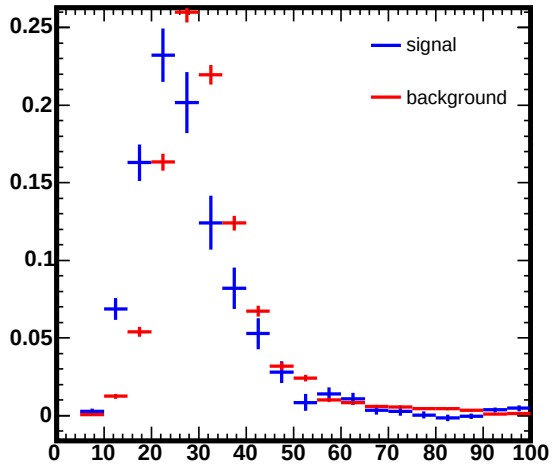
Figure 38: Comparison of background and sideband-subtracted signal mass error distributions



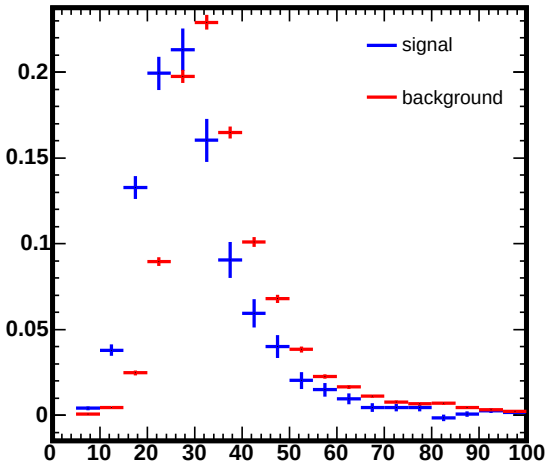
(a) $B^0 \to J/\psi K^0_s$



(b) $\Lambda^0_b \to J/\psi \Lambda^0$



(c) $B^0 \to J/\psi K^{*0}$



(d) $B^+ \to J/\psi K^+$

Figure 39: Comparison of background and sideband-subtracted signal PDL error distributions

where s is an overall scale factor on the $c\tau$ errors and

$$E(\lambda_i|c\tau) = \begin{cases} \frac{1}{c\tau} e^{-\lambda_i/c\tau}, & c\tau \geq 0 \\ 0, & \text{otherwise} \end{cases} \tag{11}$$

$$G(\lambda_i, \sigma_i^\lambda|s) = \frac{1}{\sqrt{2\pi} s\sigma_i^\lambda} \exp\left[\frac{-(\lambda_i)^2}{2(s\sigma_i^\lambda)^2}\right] \tag{12}$$

Sparing the reader the tedious algebra, this convolution results in the following analytic expression

$$\mathcal{P}_{sig}^\lambda(\lambda_i, \sigma_i^\lambda|c\tau, s) = \frac{1}{2c\tau} \exp\left[\frac{(s\sigma_i^\lambda)^2}{2(c\tau)^2} - \frac{\lambda_i}{c\tau}\right] \text{erfc}\left(\frac{s\sigma_i^\lambda}{\sqrt{2}c\tau} - \frac{\lambda_i}{\sqrt{2}s\sigma_i^\lambda}\right) \tag{13}$$

See Appendix A for more details on this function and its use in this analysis.

4.3.2 Background

In addition to the signal we are trying to measure, there are basically three other contributions to the proper decay length distribution:

1. Zero-lifetime backgrounds
   These are prompt J/ψ's combined with random V^0 tracks and contribute by far the largest background. Since they are effectively δ-functions at $c\tau = 0$, their shape actually determines the Gaussian resolution function convoluted with the signal exponential.[4]

2. Non-zero lifetime backgrounds
   There are numerous contributions here, but a large contribution should be from inclusive $B \to J/\psi$ combined with a random V^0.

3. Non-Gaussian tails on resolution function

The lifetime background function is the sum of a zero lifetime background ($\delta(c\tau = 0)$), negative-tail exponential, and two positive-tail exponentials, all convoluted with a Gaussian resolution function:

For $\lambda_i \geq 0$:

$$\mathcal{P}_{bkg}^\lambda(\lambda_i, \sigma_i^\lambda|s, \lambda_-, f_-, \lambda_+, f_+, \lambda_{++}, f_{++}) = (1 - f_- - f_+ - f_{++}) \times \left(\frac{1}{\sqrt{2\pi} s\sigma_i^\lambda} \exp\left[\frac{-(\lambda_i)^2}{2(s\sigma_i^\lambda)^2}\right] + \frac{f_+}{\lambda_+} e^{\frac{-\lambda_i}{\lambda_+}} + \frac{f_{++}}{\lambda_{++}} e^{\frac{-\lambda_i}{\lambda_{++}}}\right) \otimes G(\lambda_i, \sigma_i^\lambda|s) \tag{14}$$

[4]This assumes that the shape of the resolution function is not a strong function of distance from the primary vertex at the few hundred μm-level.

For $\lambda_{\rm i} < 0$:

$$\mathcal{P}^{\lambda}_{bkg}(\lambda_{\rm i}, \sigma^{\lambda}_{\rm i}|{\rm s}, \lambda_{-}, {\rm f}_{-}, \lambda_{+}, {\rm f}_{+}, \lambda_{++}, {\rm f}_{++}) = (1 - {\rm f}_{-} - {\rm f}_{+} - {\rm f}_{++}) \times \left(\frac{1}{\sqrt{2\pi} s \sigma^{\lambda}_{\rm i}} \exp\left[\frac{-(\lambda_{\rm i})^2}{2(s\sigma^{\lambda}_{\rm i})^2} \right] + \frac{{\rm f}_{-}}{\lambda_{-}} e^{\frac{-\lambda_{\rm i}}{\lambda_{-}}} \right) \otimes G(\lambda_{\rm i}, \sigma^{\lambda}_{\rm i}|{\rm s}) \quad (15)$$

4.4 Proper Decay Length Error PDF

Its difficult to deduce a well-motivated, "theoretical" model of the PDL (or mass) error that a good fit to the observed distributions shown in Figures 39 and 38. However, we do find that we can reasonably fit the PDL error distributions for both signal and background for all the modes by our old friend, the exponential convoluted with a Gaussian ($\mu \neq 0$),

$$\mathcal{P}^{\sigma^{\lambda}}_{sig,bkg}(\sigma^{\lambda}_{\rm i}|\lambda_{\rm psig,bkg}, \sigma_{\rm psig,bkg}, \mu_{\rm psig,bkg}) = \frac{1}{2\lambda_{\rm psig,bkg}} \exp\left(\frac{\sigma^2_{\rm psig,bkg}}{2\lambda^2_{\rm psig,bkg}} - \frac{\sigma^{\lambda}_{\rm i} - \mu_{\rm psig,bkg}}{\lambda_{\rm psig,bkg}} \right) \times {\rm erfc}\left(\frac{\sigma_{psig,bkg}}{\sqrt{2}\lambda_{psig,bkg}} - \frac{\sigma^{\lambda}_{\rm i} - \mu_{\rm psig,bkg}}{\sqrt{2}\sigma_{psig,bkg}} \right) \quad (16)$$

as shown in Figures 40 and 41.

It is worth noting that this is purely an empirical model that is convenient because it is a smooth function used throughout the fitter for the signal and background PDL functions. Also, the fits are not a perfect parameterization of the PDL error distributions. The PDL error modeling is considered as part of our systematic uncertainty.

4.5 Mass PDF

In this section we describe $\mathcal{P}^{m}_{sig}$ and $\mathcal{P}^{m}_{bkg}$, the PDF's for the mass part of the signal and background.

4.5.1 Signal

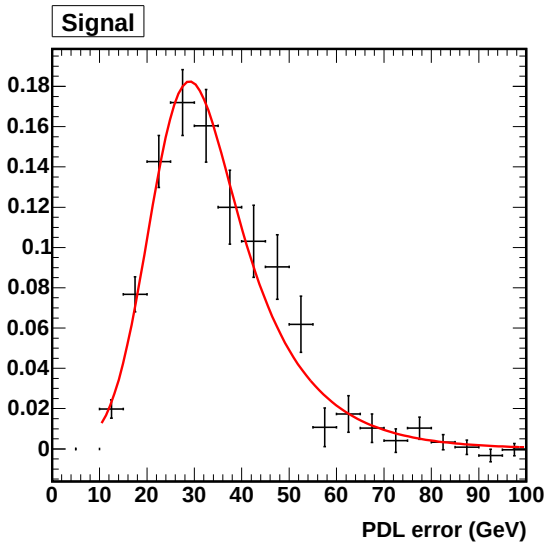
The mass signal function we use is a single Gaussian centered at mass M with a width given by a scaling parameter, $s_{\rm M}$ of the event mass errors

$$\mathcal{P}^{m}_{sig}({\rm m_i}, \sigma^{\rm m}_{\rm i}|{\rm M}, \sigma_{\rm M}) = \frac{1}{\sqrt{2\pi} {\rm s_M} \sigma^{\rm m}_{\rm i}}) \exp\left[\frac{-({\rm m_i} - {\rm M})^2}{2({\rm s_M}\sigma^{\rm m}_{\rm i})^2} \right] \quad (17)$$

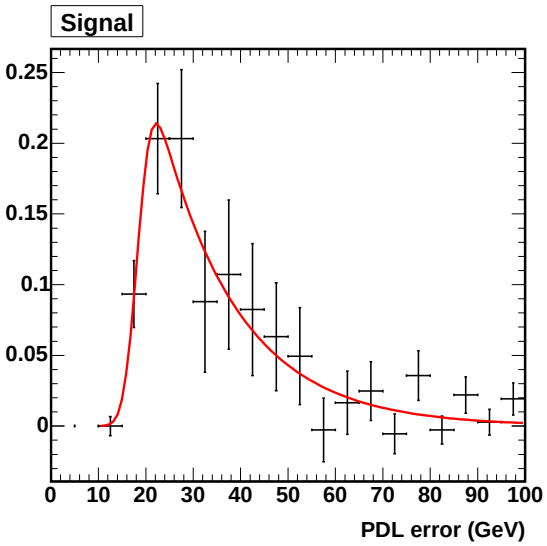
4.5.2 Background

We use a linear mass shape as our mass background model

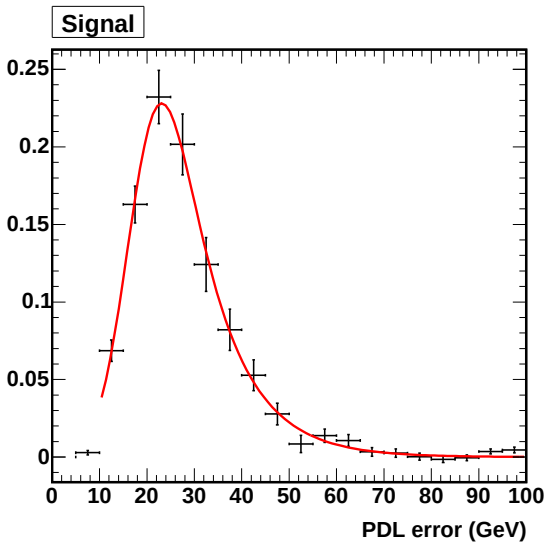
$$\mathcal{P}^{m}_{bkg}({\rm m_i}|{\rm C}_0) = \left(\frac{2}{{\rm M}^2_{\rm high} - {\rm M}^2_{\rm low}} - \frac{2{\rm C}_0}{{\rm M}_{\rm high} - {\rm M}_{\rm low}} \right) \quad (18)$$



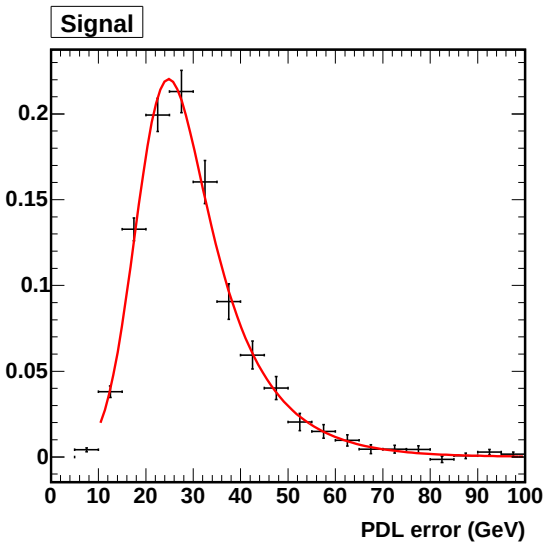
(a) $B^0 \to J/\psi K^0_s$



(b) $\Lambda^0_b \to J/\psi \Lambda^0$

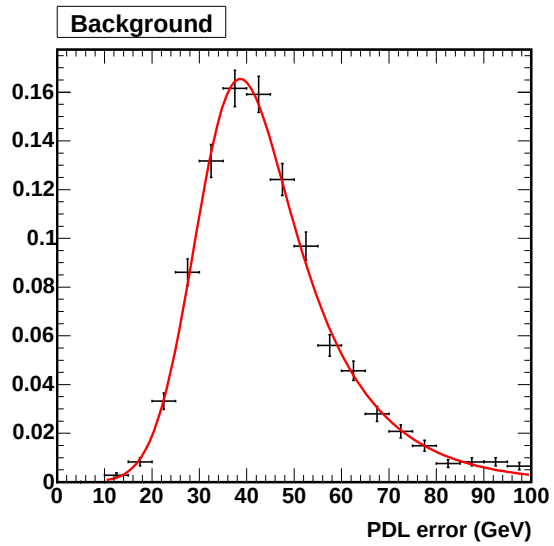


(c) $B^0 \to J/\psi K^{*0}$

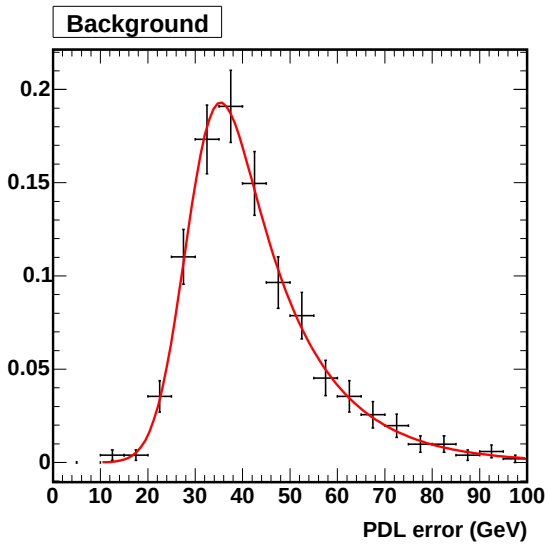


(d) $B^+ \to J/\psi K^+$

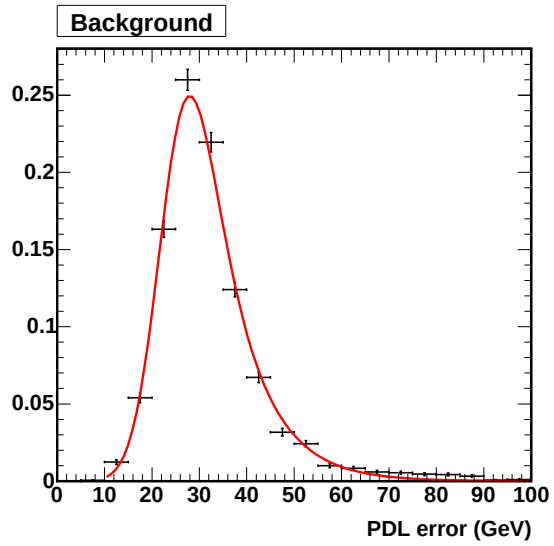
Figure 40: Fits of PDL error model to sideband-subtracted signal PDL error distributions



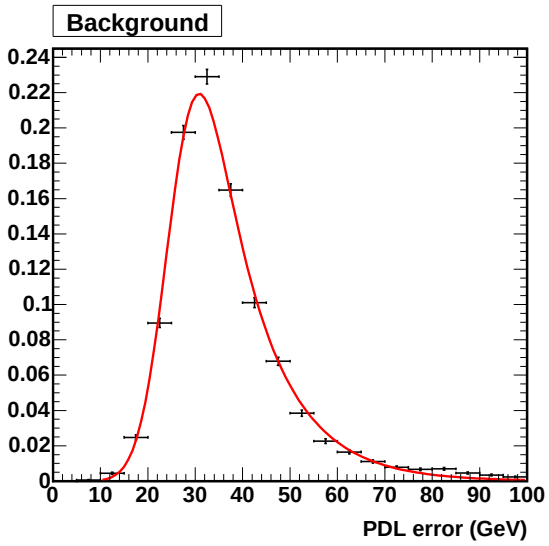
(a) $B^0 \to J/\psi K^0_s$



(b) $\Lambda^0_b \to J/\psi \Lambda^0$



(c) $B^0 \to J/\psi K^{*0}$



(d) $B^+ \to J/\psi K^+$

Figure 41: Fits of PDL error model to background PDL error distributions from data sidebands

with C_0 representing the sole fit parameter after normalization over the mass window (M^2_{low}, M^2_{high}) .

4.5.3 Fit Range

Our baseline fit is over a restricted range of the data. Specifically, the fit all of the modes over the following ranges

- Proper decay length: -2000 μm $\leq$ PDL $\leq$ 4000 μm
- Proper decay length error: 0 μm $\leq$ PDL error $\leq$ 100 μm
- Mass:
  - For B^0 and B^+ modes: 5.170 GeV/c^2 $\leq$ Mass $\leq$ 5.390 GeV/c^2
  - For $\Lambda^0_b \rightarrow J/\psi\Lambda^0$: 5.521 GeV/$c^2$ $\leq$ Mass $\leq$ 5.721 GeV/c^2
- Mass Error: 0 μm $\leq$ Mass error $\leq$ 20 MeV

Recall from Figure 38 and Figure 39 that these are loose upper limit requirements on the mass and PDL errors, respectively. However, they avoid using data with very large errors that are likely to be mis-estimately anyway.

4.6 Validation

Our baseline fit is now an 18 parameter maximum likelihood fit to PDL, PDL error, mass, and mass error data using the sum of signal and background PDF's that are each products of independent PDL, PDL error, and mass PDF's. Given the relative complexity of the fitting procedure and data reconstruction, its important to have checks that validate the fitting procedure used to extract the lifetimes. Our checks involve toy Monte Carlo datasets and Monte Carlo reconstructed in an identical way as the data.

4.6.1 Toy Monte Carlo

Here we check our fitting procedure by generating toy Monte Carlo data sets and fitting them in exactly the same way as we fit the data. We perform two different toy Monte Carlo tests:

1. Generate and fit a single toy data set with many more events than in the actual data. In particular, this involved a toy data set 100 times the size of our real data. We expect our fit results to be consistent with our generated model parameters within their statistical uncertainties and that our fit probabilities should be reasonable.

2. Generate and fit a large number of toy data sets, each of size equal to the number of events in the data. We expect the distribution of pulls (i.e. (fitted_i − $\text{generated})/\text{error}_\text{i}$) to be consistent with a unit Gaussian.

In both cases, the toy data results from random generation of events using the model and model parameters resulting from our baseline fit to the $B^0 \to J/\psi K^0_s$ data (to be presented in Section 5.1). In this way, we can use these studies both to demonstrate the validity of the fitting procedure and to determine our lifetime systematic from fitter bias. The toy data sets are fit in an identical way as our baseline fit to the data, including the way the parameters are seeded.

Table 4 summarizes the results of the single toy Monte Carlo run using 100 times the size of our $B^0 \to J/\psi K^0_s$ dataset. Figure 42 shows the fit projections in PDL, PDL error, and mass. Figure 43 shows the variable-binned overlay for goodness-of-fit determination and differences between data and the model. Note from Table 4 that difference in the fit and generated parameters are reasonable and the fit probability for each projection is good.

We more precisely measured the bias underlying our fit the $B^0 \to J/\psi K^0_s$ data by making a large number (9600) of statistically-independent toy Monte Carlo runs with the exact same statistics as in our data. This is an actual measure of our fitter bias (to the extent that we correctly toy our model, which is adequately demonstrated by Table 4 and Figures 42 and 43) since the bias on the fit parameters are in principal sample-size dependent. For each of our 9600 toy data sets, we fit the toy data and record the fitted value and error of each model parameter. The distribution of fitted value (minus generated value) and "pull" defined as ($\text{fit}_i - \text{generated})/\text{error}_i$ are shown in Figures 44-52.

Most of the parameters have pull distributions consistent with unit Gaussians. The fit parameter which we are dependent upon for this analysis is the signal $c\tau$, which is consistent with a unit gaussian indicating the parameter is largely unbiases with errors that make sense. There is a very small bias (0.16 μm) toward low values due to the asymmetry of the likelihood function for this sample size, which we consider acceptable.

4.6.2 Fits to Signal Monte Carlo

In this section, we fit the reconstructed BGenerator Monte Carlo for $B^0 \to J/\psi K^0_s$ and $\Lambda^0_b \to J/\psi \Lambda^0$ using the lifetime fitting procedure just described. Any problems in reconstruction which bias the signal and are modeled by the Monte Carlo will show up and a discrepency between the fitted and generated lifetimes.

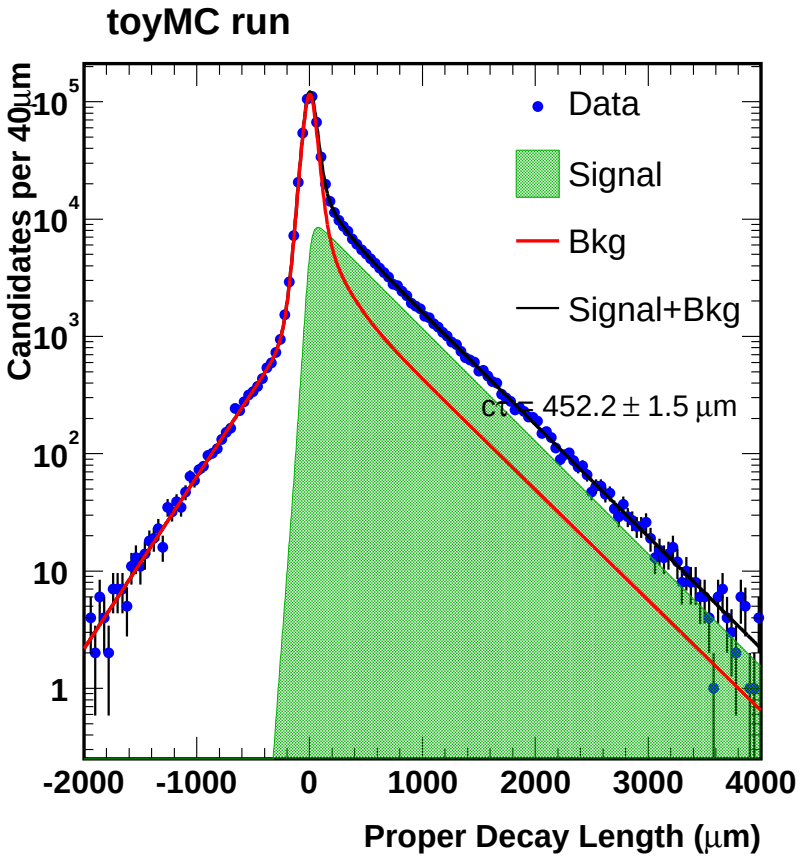
Figures 53 and 54, we show the PDL fit projections and variable-binned overlay for $B^0 \to J/\psi K^0_s$ and $\Lambda^0_b \to J/\psi \Lambda^0$ and in Table 5 we summarize the numerical results. We find the Monte Carlo is actually not a great fit to a signal gaussian PDL resolution function (at least for $B^0 \to J/\psi K^0_s$ but rather is a better fit to 2 gaussian resolution functions. However, the fitted lifetime is completely unaffected by adding this extra gaussian on a scale we care about ($<< 1\mu$m).

Parameter	Generated	Fit	Units	Difference (σ)
$c\tau$	450.6147	452.1647 ± 1.4865	μm	+1.0
s	1.2518	1.2544 ± 0.0021		+1.2
λ_+	141.1789	130.7306 ± 4.5531	μm	−2.3
f_+	0.0771	0.0728 ± 0.0020		−2.2
λ_{++}	472.2418	461.6518 ± 5.0945	μm	−2.1
f_{++}	0.0930	0.0981 ± 0.0020		+2.6
λ_-	303.7694	295.7545 ± 3.3220	μm	−0.8
f_-	0.0314	0.0311 ± 0.0004		−0.8
f_b	0.7843	0.7848 ± 0.0007		+0.7
M	5.28120	5.28116 ± 0.00004	GeV/c^2	−1.0
s_M	1.7613	1.7633 ± 0.0052		+0.4
C_0	7.0684	7.3600 ± 0.5729	GeV/c^2	+0.5
λ_{psig}	13.2106	13.2744 ± 0.0764	μm	+0.8
σ_{psig}	5.7931	5.7463 ± 0.0363	μm	−1.3
μ_{psig}	21.8063	21.7147 ± 0.0511	μm	−1.8
λ_{pbkg}	13.7683	13.7337 ± 0.0409	μm	−0.8
σ_{pbkg}	7.3250	7.3312 ± 0.0221	μm	+0.3
μ_{pbkg}	31.4021	31.4285 ± 0.0307	μm	
PDL χ^2 Probability:		80.6%		
PDL error χ^2 Probability:		13.4%		
Mass χ^2 Probability:		77.4%		

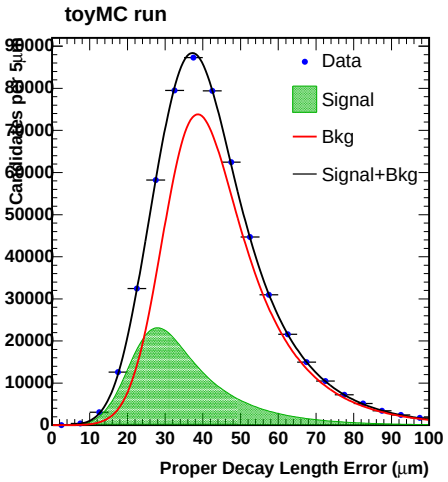
Table 4: Summary of fit to toy Monte Carlo of baseline model fit to the data, with the 100 times the toy data as in the actual data.

	$B^0 \to J/\psi K^0_s$		$\Lambda^0_b \to J/\psi \Lambda^0$	
Parameter	baseline	2 Gaus	baseline	2 Gaus
$c\tau$	466.6 ± 2.3 μm	466.9 ± 2.3 μm	369.2 ± 3.0μm	369.4 ± 3.0μm
s	1.30 ± 0.02	1.02 ± 0.03	1.26 ± 0.03	1.04 ± 0.04
s_{g2}	–	3.5 ± 0.4	–	3.1 ± 0.5
f_{g2}	–	0.038 ± 0.01	–	0.043 ± 0.02
PDL χ^2 Prob:	0.0%	54.7%	53.1%	82.9%

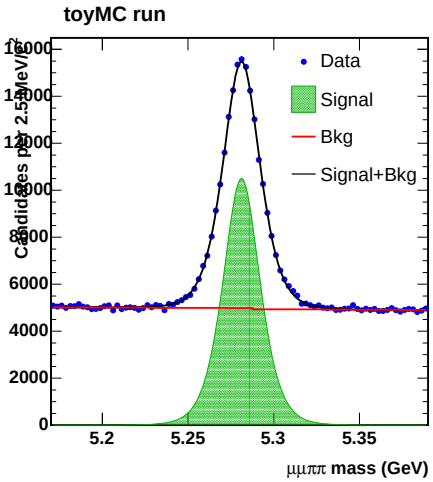
Table 5: Summary of lifetime fits to signal Monte Carlo. The generated lifetimes were 464 μm for B^0 and 368 μm for Λ^0_b.



(a) PDL projection

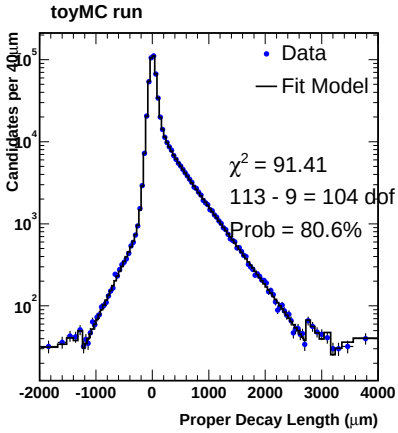


(b) PDL error projection

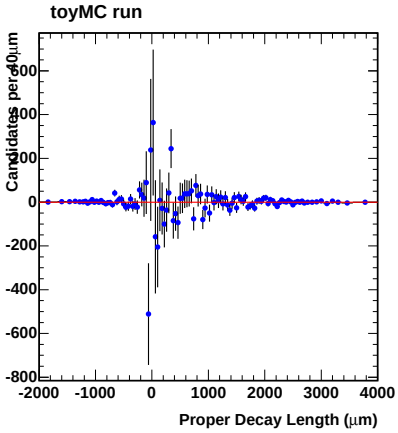


(c) Mass projection

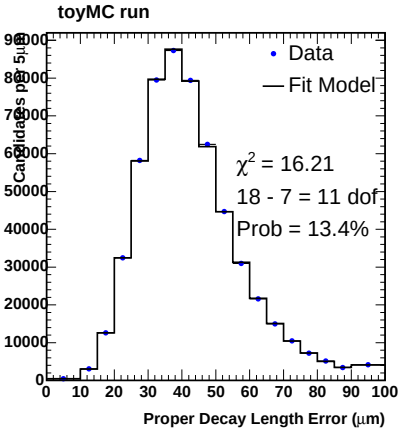
Figure 42: Fit to toy Monte Carlo of baseline model fit to the data, with the 100 times the toy data as in the actual data.



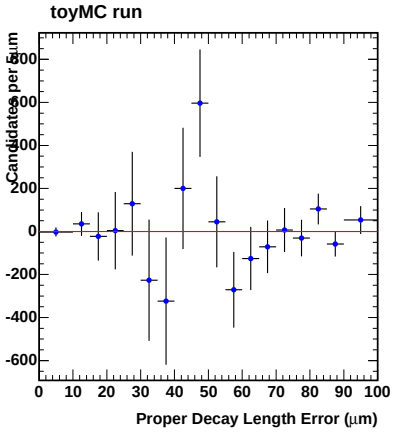
(a) Variable-binned PDL projection



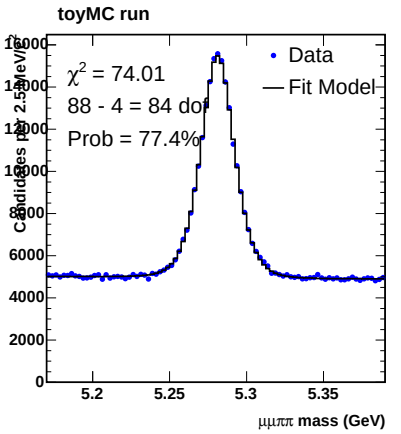
(b) Data-Model PDL projection



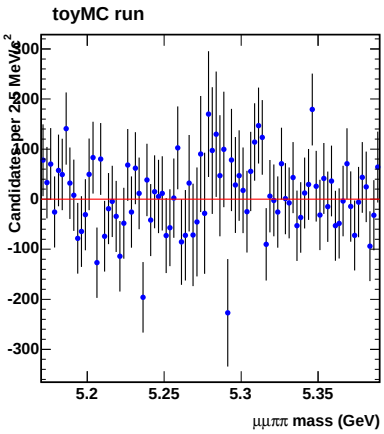
(c) Variable-binned PDL error projection



(d) Data-Model PDL error projection

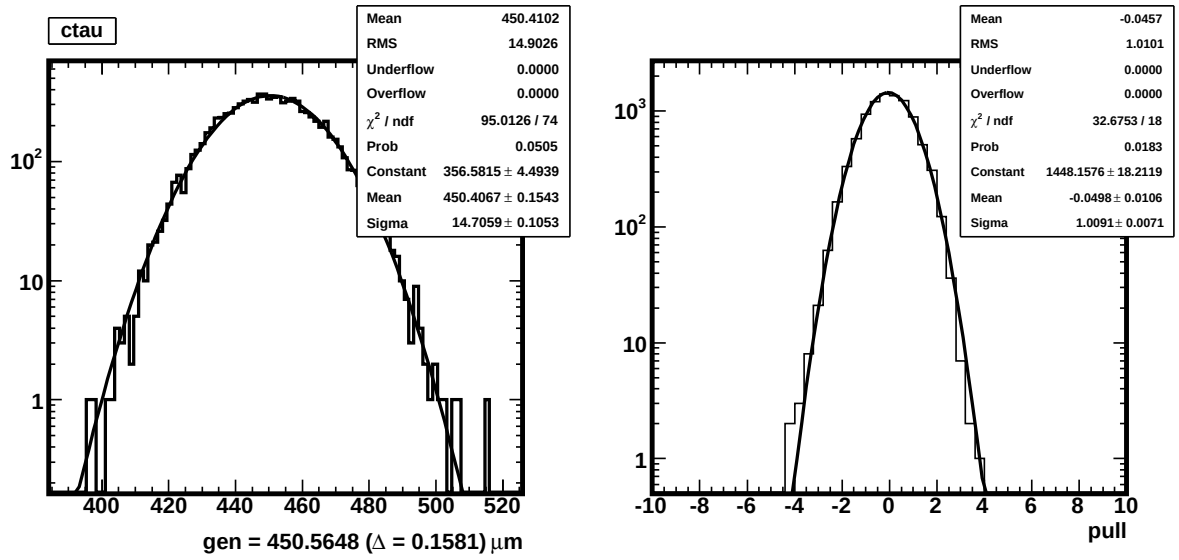


(e) Variable-binned mass projection



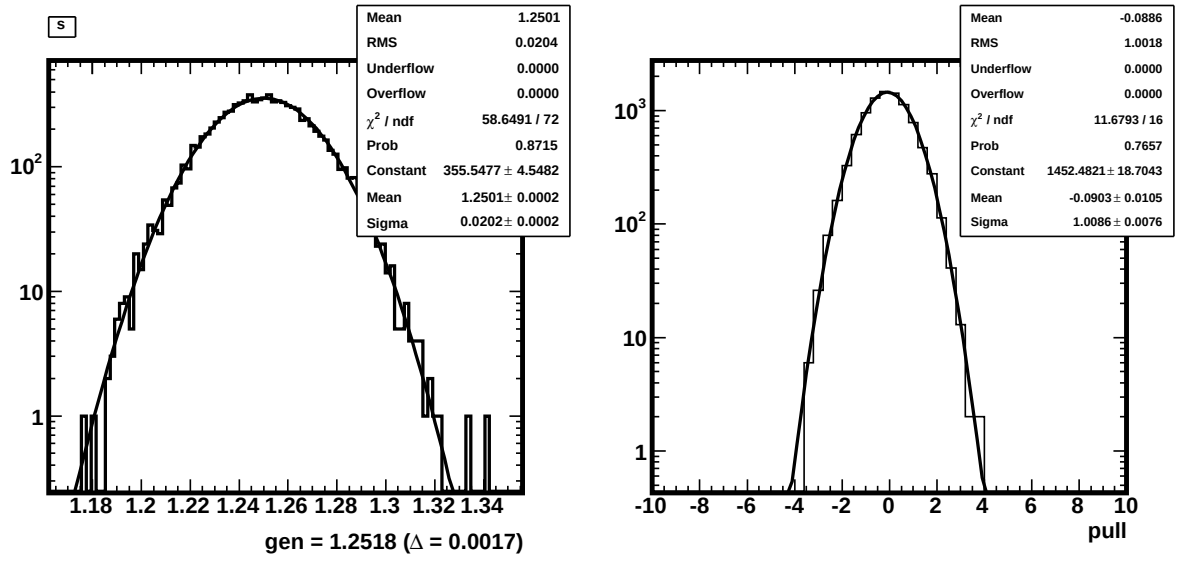
(f) Data-Model mass projection

Figure 43: Fit to toy Monte Carlo of baseline model fit to the data, with the 100 times the toy data as in the actual data.



(a) Fitted $c\tau$

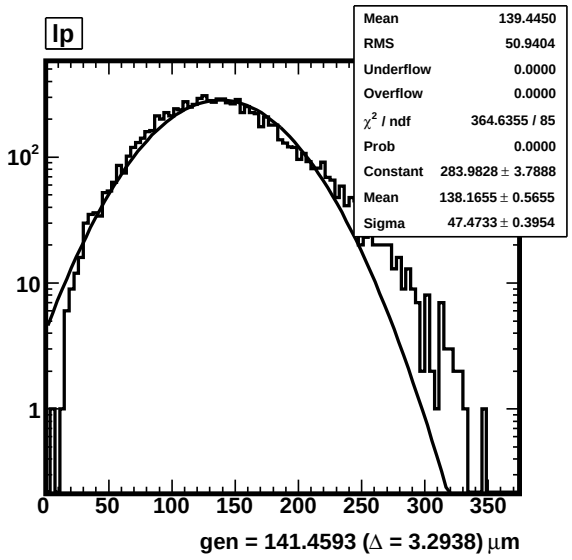
(b) Pull distribution for $c\tau$



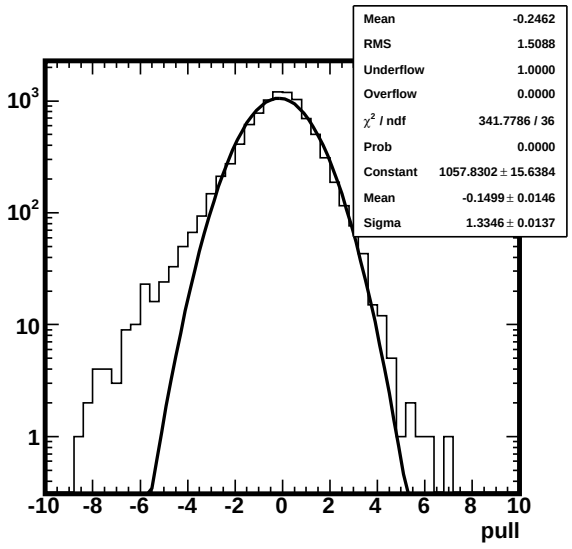
(c) Fitted s

(d) Pull distribution for s

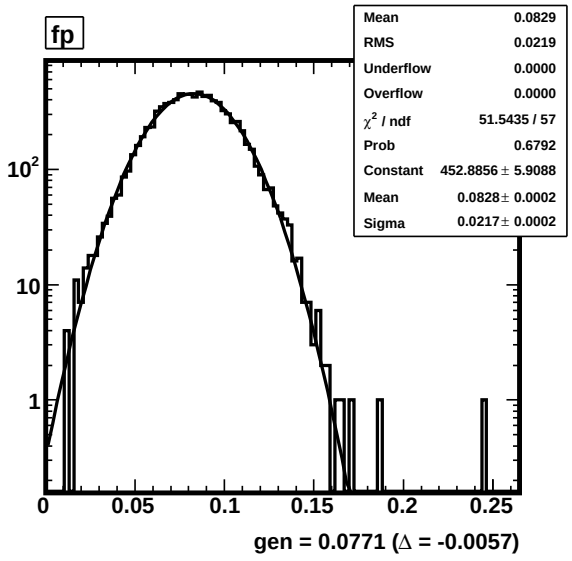
Figure 44: Pull distributions from fits to 9600 toy Monte Carlo data sets of the same size and model composition as baseline fit to actual data.



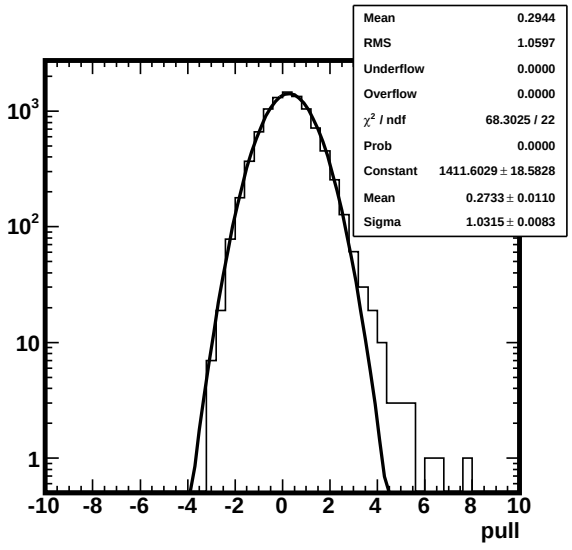
(a) Fitted λ_+



(b) Pull distribution for λ_+

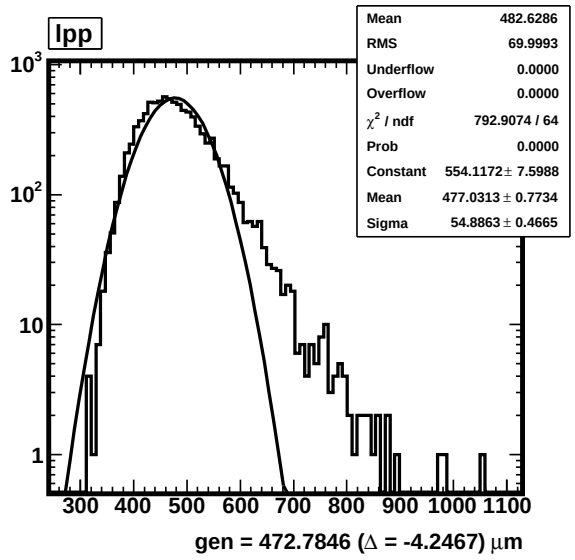


(c) Fitted f_+

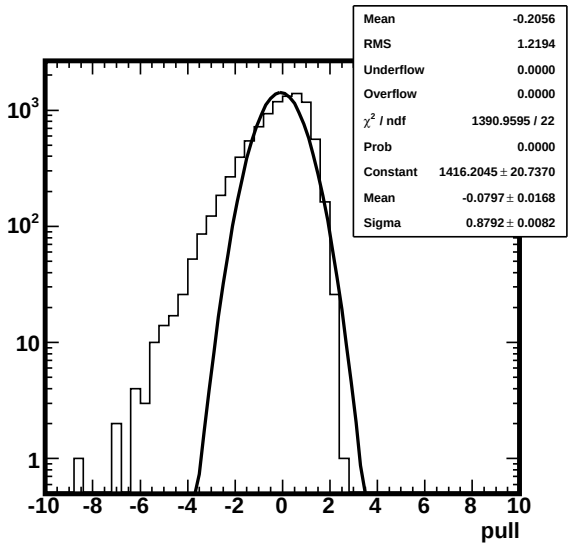


(d) Pull distribution for f_+

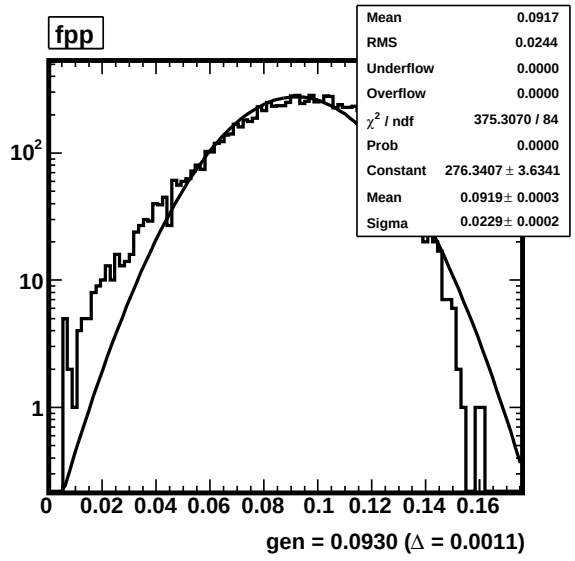
Figure 45: Pull distributions from fits to 9600 toy Monte Carlo data sets of the same size and model composition as baseline fit to actual data.



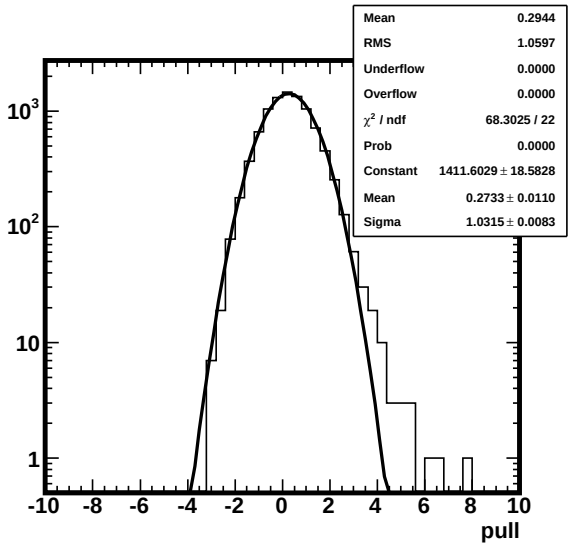
(a) Fitted λ_{++}



(b) Pull distribution for λ_{++}

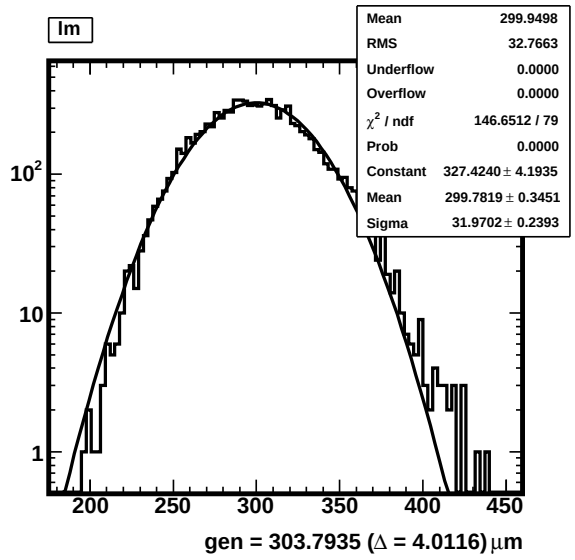


(c) Fitted f_{++}

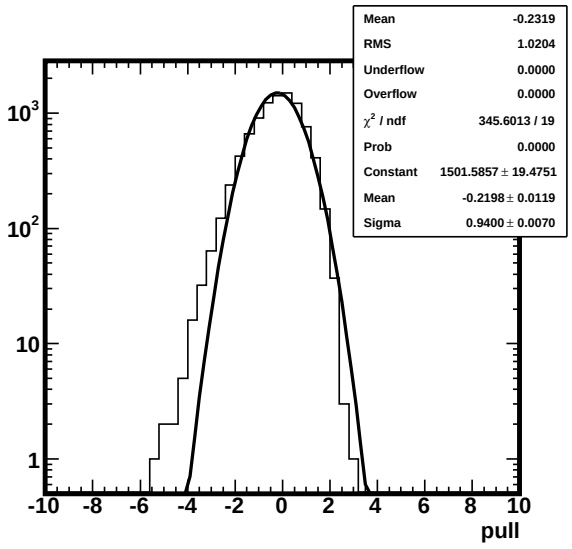


(d) Pull distribution for f_{++}

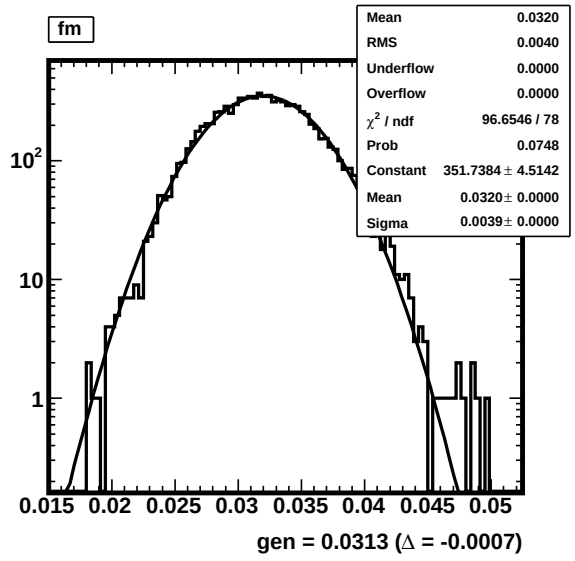
Figure 46: Pull distributions from fits to 9600 toy Monte Carlo data sets of the same size and model composition as baseline fit to actual data.



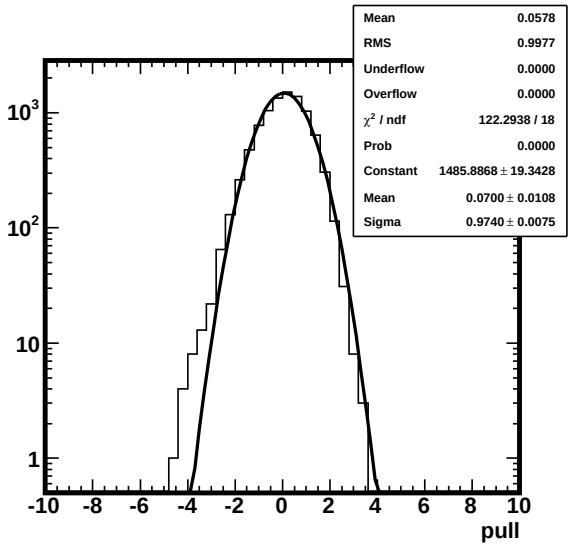
(a) Fitted λ_-



(b) Pull distribution for λ_-

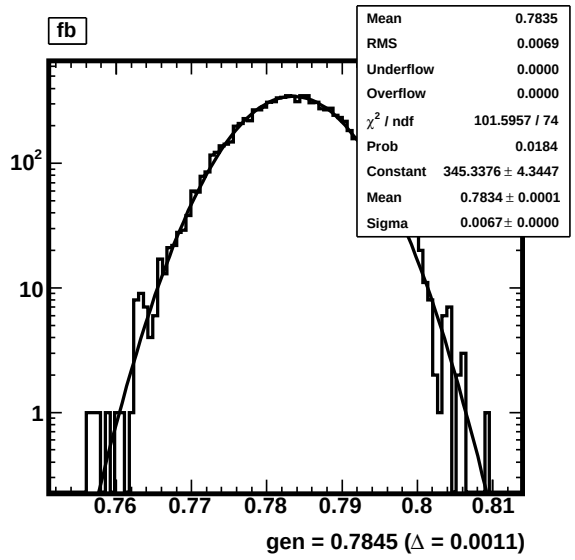


(c) Fitted f_

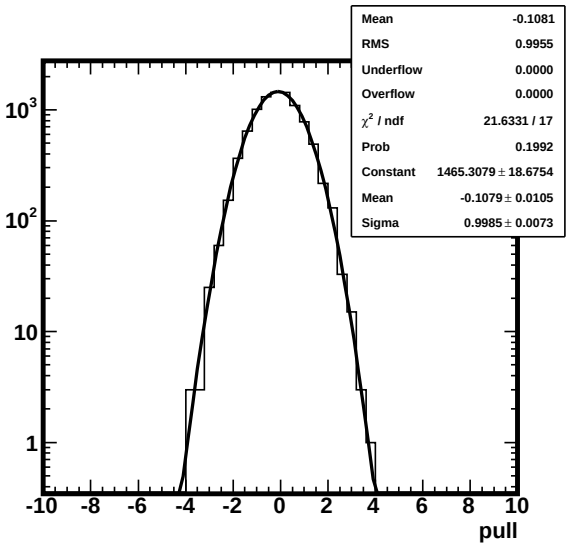


(d) Pull distribution for f_

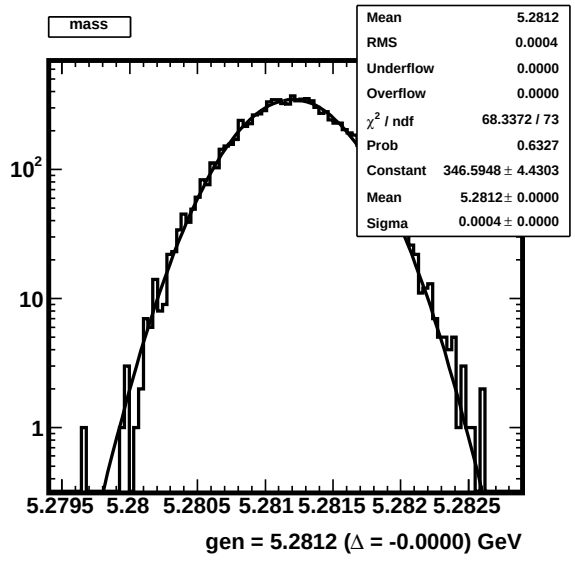
Figure 47: Pull distributions from fits to 9600 toy Monte Carlo data sets of the same size and model composition as baseline fit to actual data.



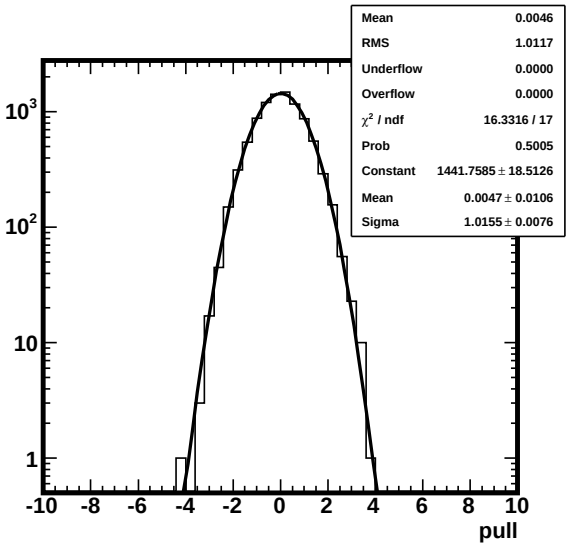
(a) Fitted f_b



(b) Pull distribution for f_b

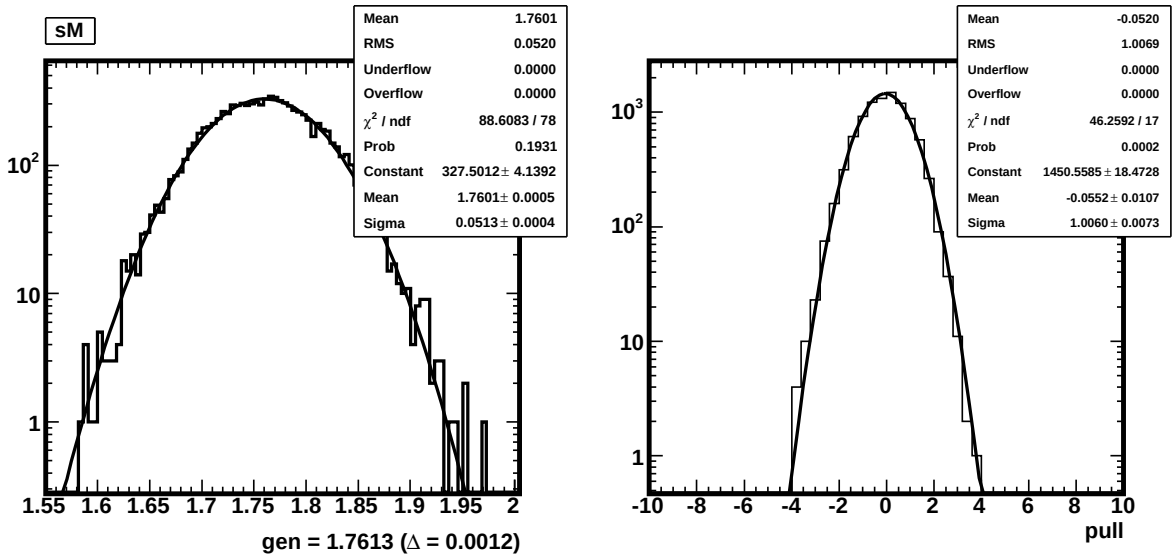


(c) Fitted M



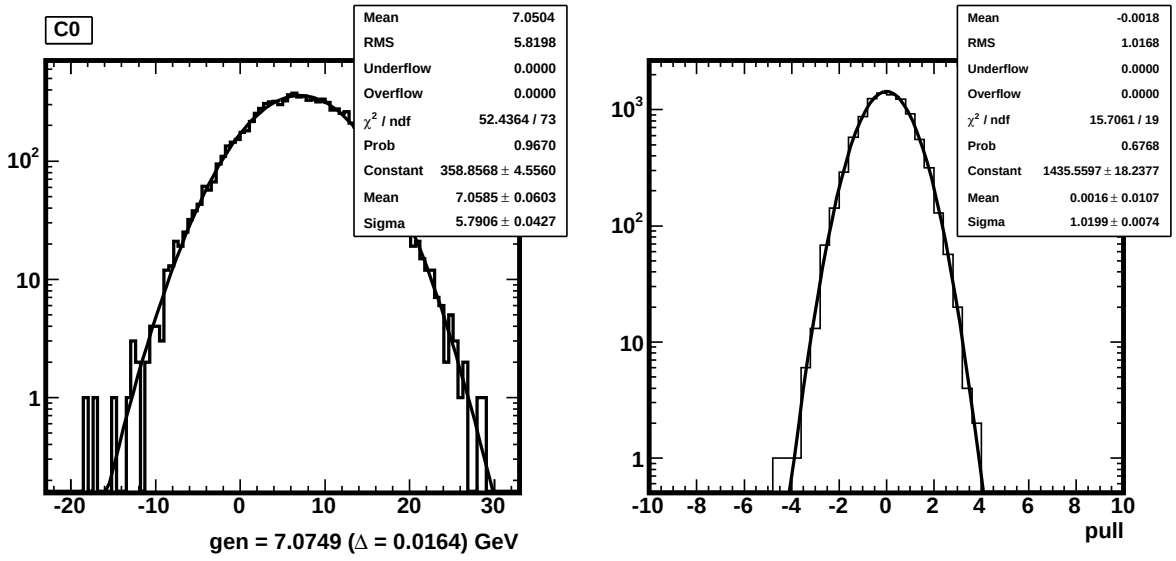
(d) Pull distribution for M

Figure 48: Pull distributions from fits to 9600 toy Monte Carlo data sets of the same size and model composition as baseline fit to actual data.



(a) Fitted s_M

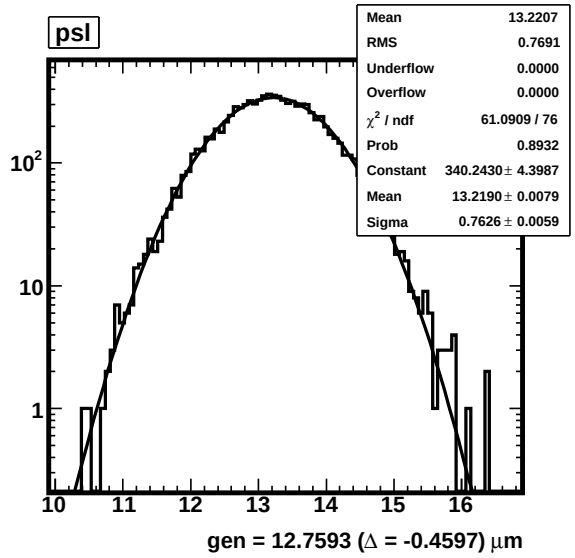
(b) Pull distribution for s_M



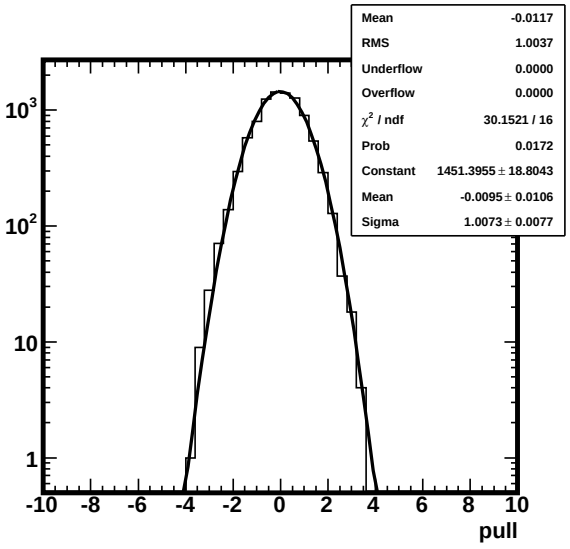
(c) Fitted C_0

(d) Pull distribution for C_0

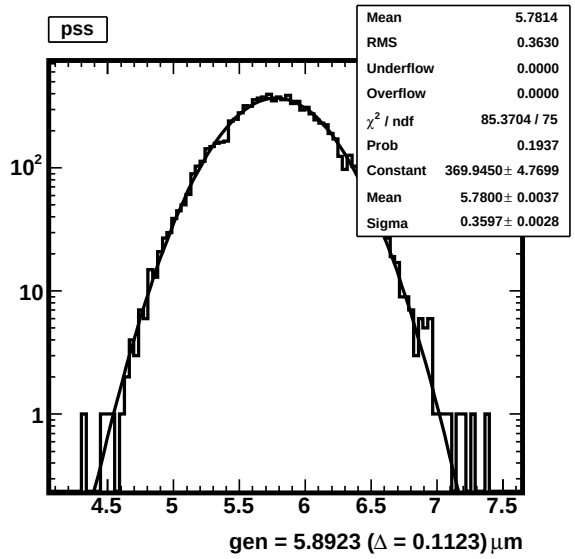
Figure 49: Pull distributions from fits to 9600 toy Monte Carlo data sets of the same size and model composition as baseline fit to actual data.



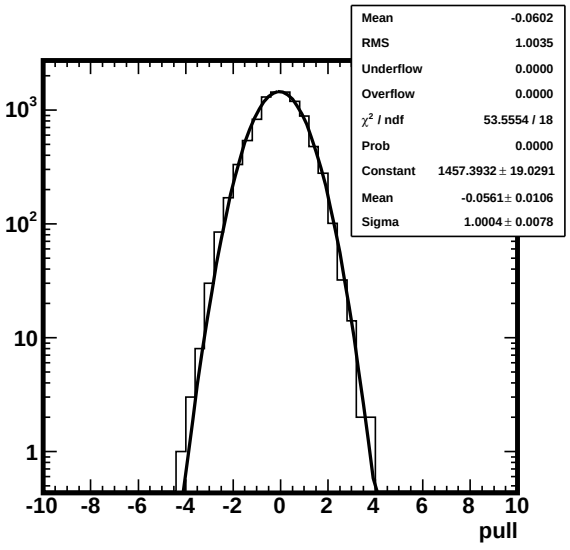
(a) Fitted λ_{psig}



(b) Pull distribution for λ_{psig}

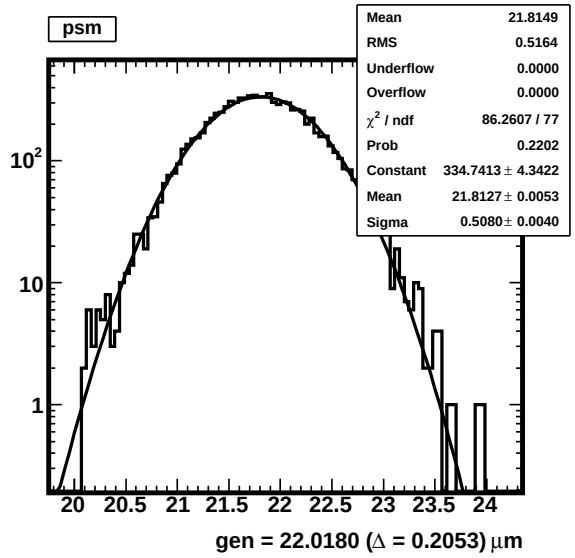


(c) Fitted σ_{psig}

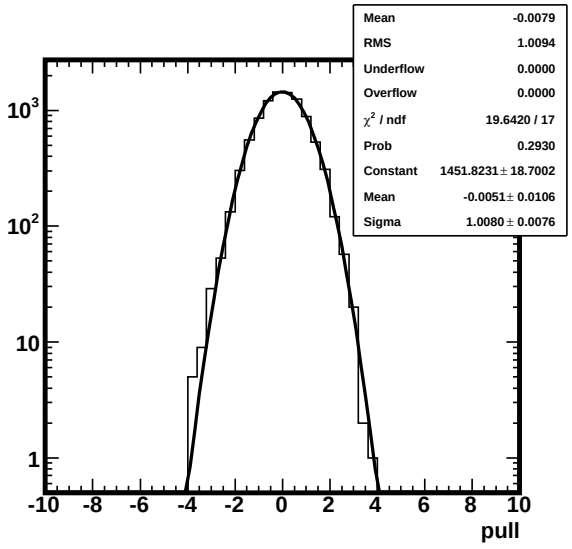


(d) Pull distribution for σ_{psig}

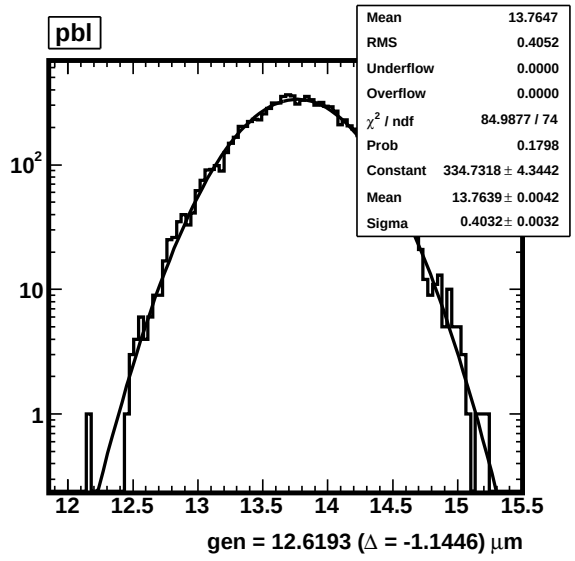
Figure 50: Pull distributions from fits to 9600 toy Monte Carlo data sets of the same size and model composition as baseline fit to actual data.



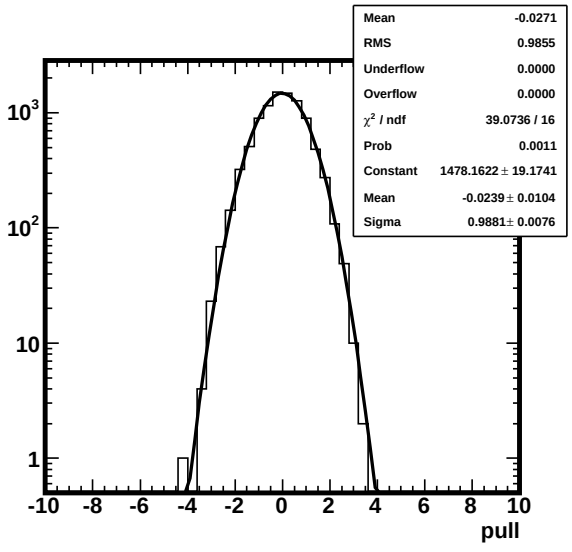
(a) Fitted μ_{psig}



(b) Pull distribution for μ_{psig}

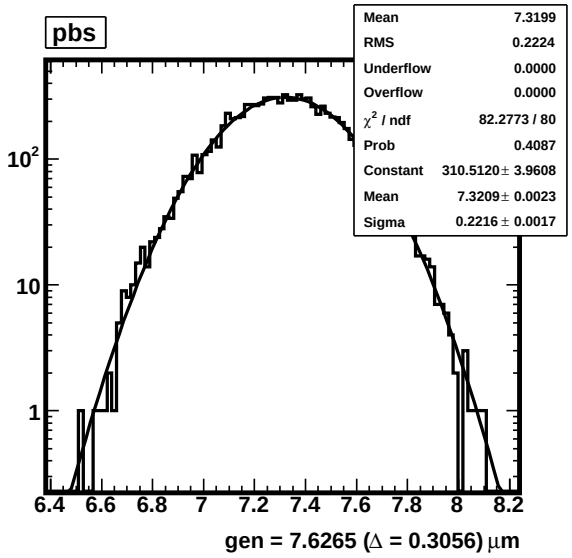


(c) Fitted λ_{pbkg}

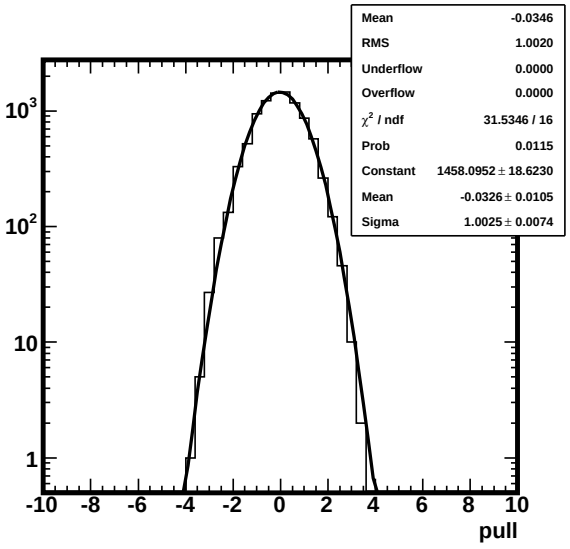


(d) Pull distribution for λ_{pbkg}

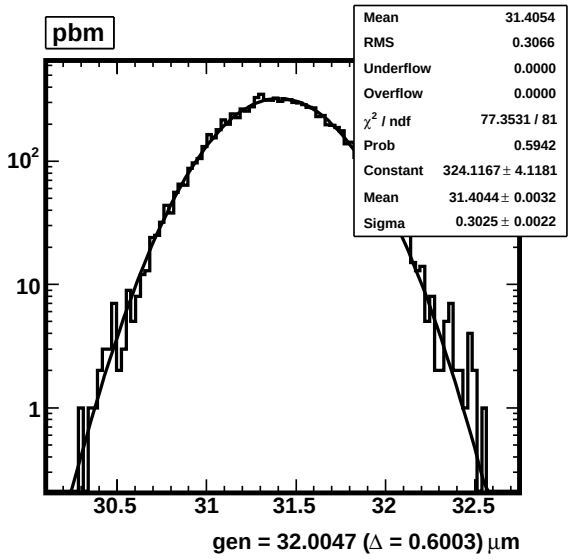
Figure 51: Pull distributions from fits to 9600 toy Monte Carlo data sets of the same size and model composition as baseline fit to actual data.



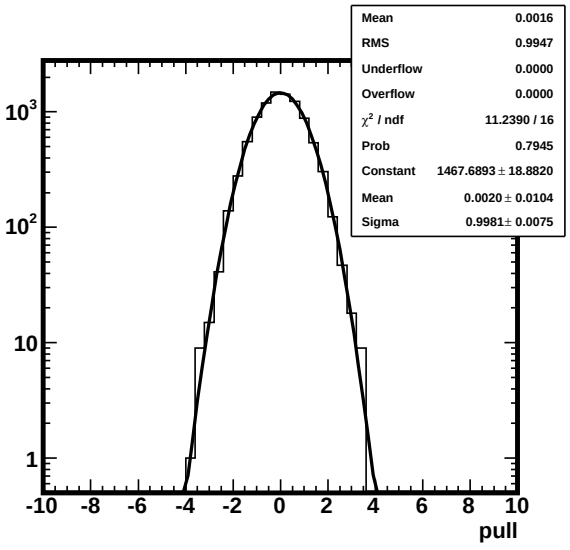
(a) Fitted σ_{pbkg}



(b) Pull distribution for σ_{pbkg}



(c) Fitted μ_{pbkg}



(d) Pull distribution for μ_{pbkg}

Figure 52: Pull distributions from fits to 9600 toy Monte Carlo data sets of the same size and model composition as baseline fit to actual data.

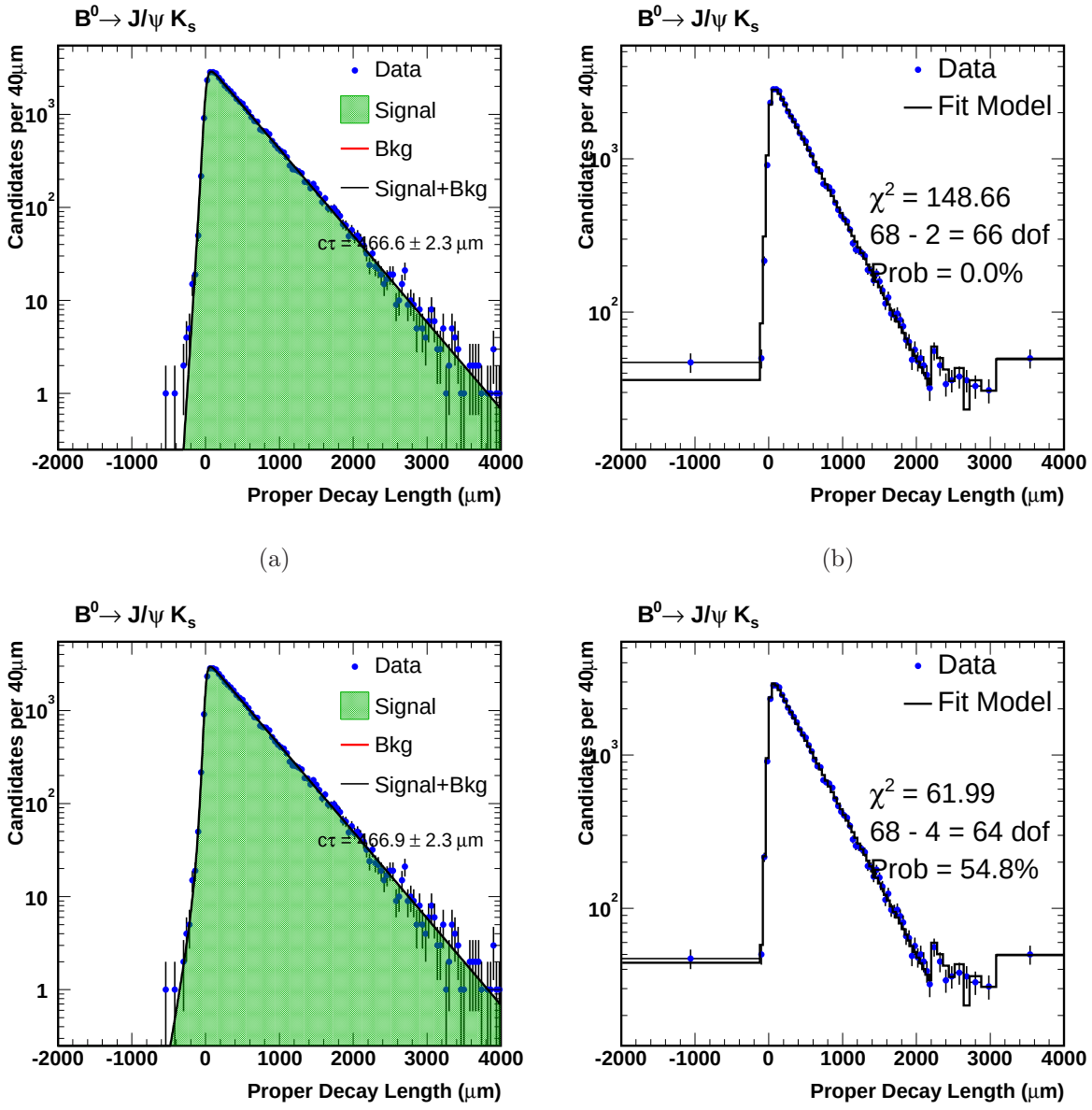


Figure 53: Fits to $B^0 \to J/\psi K^0_s$ signal Monte Carlo

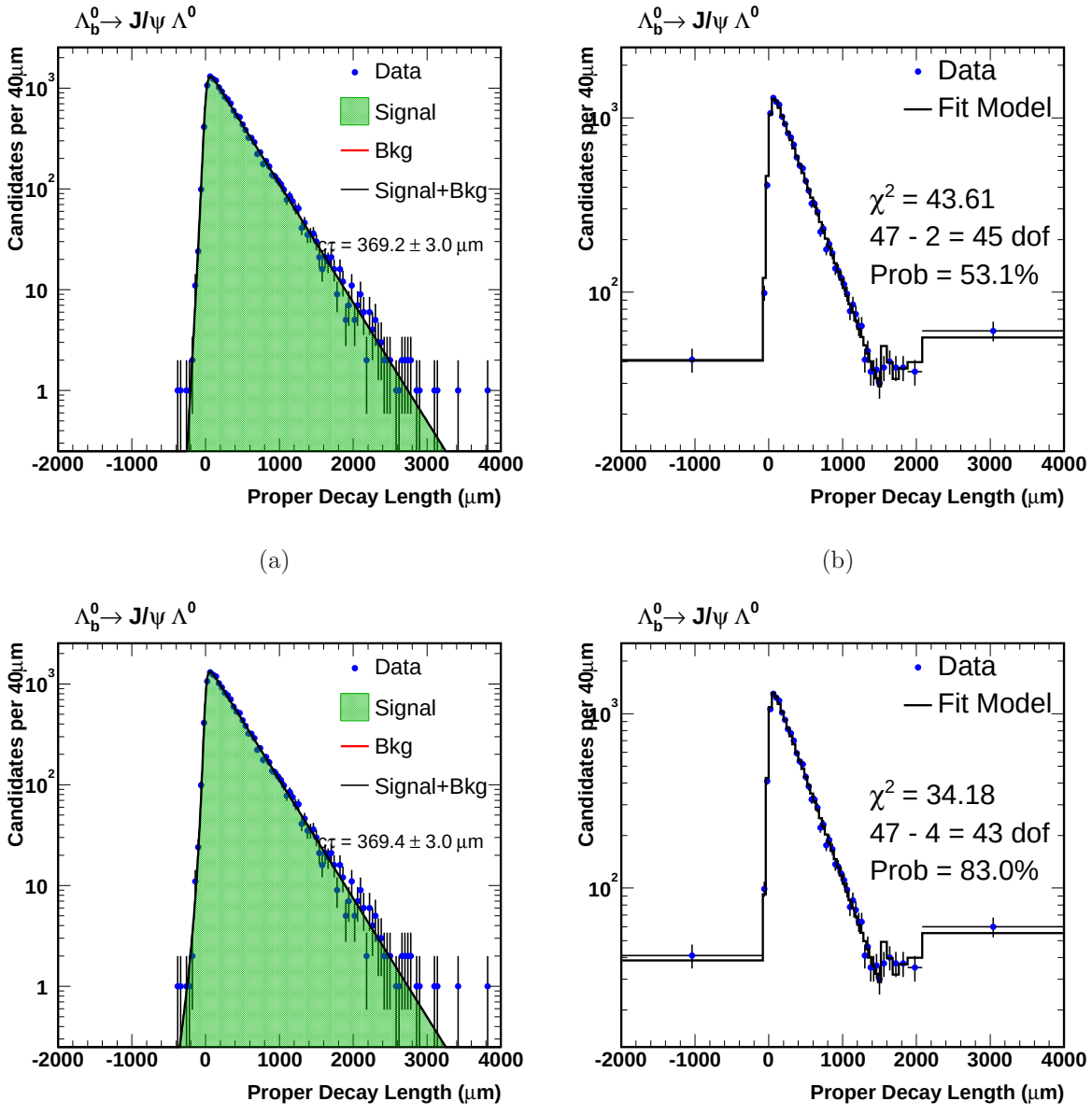
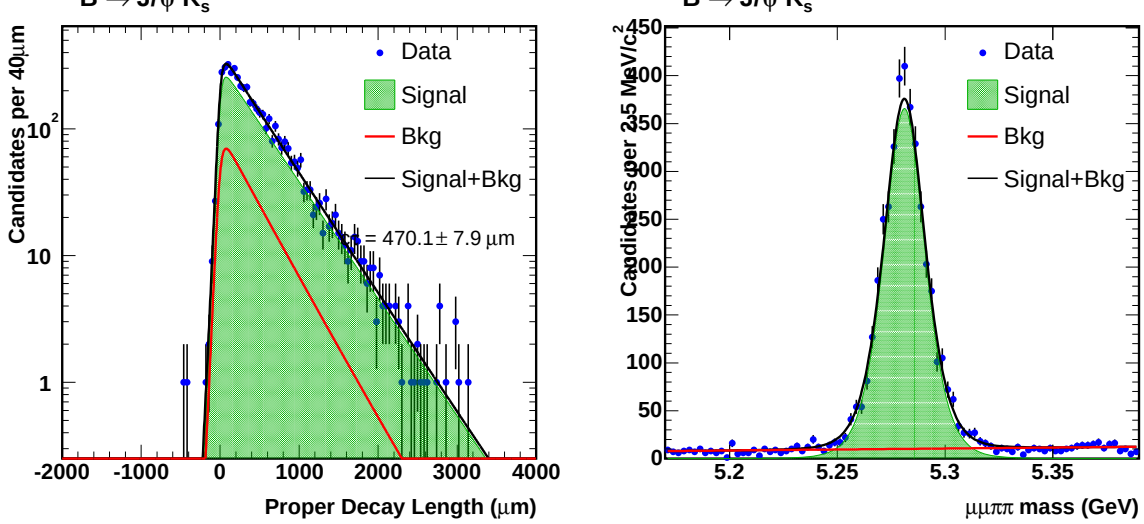


Figure 54: Fits to $\Lambda_b^0 \to J/\psi\Lambda^0$ signal Monte Carlo



(a) PDL projection in lifetime fit

(b) Mass projection in lifetime fit

Figure 55: Lifetime fit to $b \to J/\psi X$ Monte Carlo reconstructed as $B^0 \to J/\psi K^0_s$

Mode	No PDL error fit	Baseline fit with PDL error	Toy MC expection
$B^0 \to J/\psi K^0_s$	444.7 ± 14.7 μm	450.6 ± 14.7 μm (+5.9μm)	+5.6μm
$B^0 \to J/\psi K^{*0}$	460.1 ± 11.6 μm	468.4 ± 11.8 μm (+8.3μm)	+5.4μm
$B^+ \to J/\psi K^+$	492.0 ± 8.4 μm	498.1 ± 8.4 μm (+6.1μm)	+3.9μm
$\Lambda^0_b \to J/\psi \Lambda^0$	638.0 ± 60.6 μm	644.2 ± 59.8 μm (+1.0%)	-

Table 6: Comparison of the observed shift in measured lifetime with that predicted by the studies in [18].

4.6.3 Fit to $b \to J/\psi X$ Monte Carlo

To test the response of the fitter to long-lived background we fit our Pythia $b \to J/\psi X$ Monte Carlo sample reconstructed as $B^0 \to J/\psi K^0_s$. The results are shown in Figure 55(a). We fit a lifetime of $470.1 \pm 7.9 \mu$m which is consistent with the generated lifetime of 464μ.

4.7 Comments on the "Punzi-Effect" for Lifetimes

Here we compare the predicted effect of the Punzi shift using toy Monte Carlo studies from [18] with the observed shifts when the PDL error model is included in the fit to data. Table 6 summarizes the results. Note that the observed shifts are in the same direction and within a few microns of the shifts expected from the toy Monte Carlo. There is also a statistical component to the differences since the PDL error fit uses data all data not just that determined from the sidebands and sideband-substracted signal region, as in the toy Monte Carlo.

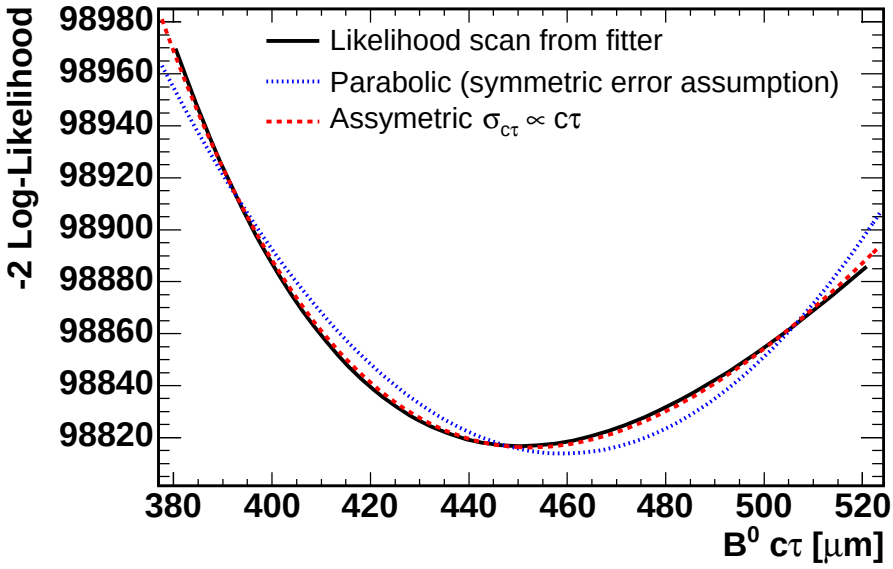


Figure 56: Likelihood as a function of $c\tau$.

5 Fit Results

5.1 B^0 Lifetime in $B^0 \to J/\psi K^0_s$

The fit results for the $B^0 \to J/\psi K^0_s$ mode in data including the full Minos errors are shown in Table 7. We find B^0 $c\tau = 451 \pm 15 \mu$m, statistical error only. The fit gives a good description the proper decay length, mass and PDL error distributions as seen graphically in Figures 57 and 58 and by the probabilities of the projections.

This result is 0.7σ below the PDG average, 460.5 ± 4.2. This statistical significance is assessed by evaluating the difference in $-2\ln\mathcal{L}|$ at the minimum with $c\tau$ free and at the minimum when guassian constrianing $c\tau$ to the PDG value. This fully takes into account both the uncertainty on the PDG value and the assymetry in the shape of the likelihood function. Figure 56 shows a scan of the minimum $-2\ln\mathcal{L}|$ as a function $c\tau$. A clear assymetry can be observed. Since the statistical error on the width of an exponential should be proportional to is width which is the lifetime, we expect the uncertainty to be proportional to the lifetime. The red (dashed) line in Figure 56 shows this assumption and the blue (dotted) line shows a symmetric, parabolic, error. Clearly the asymmety is well described by this proportionality.

5.2 Other b-meson Modes

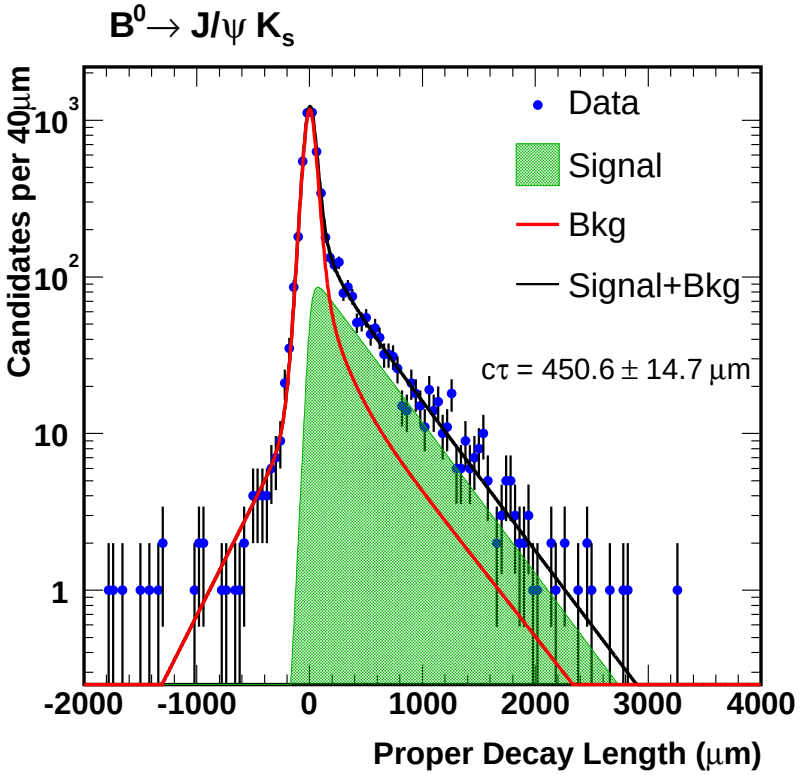
We also measure the fit for the lifetimes in the modes

Model parameters:					
Parameter	Fit value	$\pm 1\sigma$ Sym Err	-1σ Minos Err	+1σ Minos Err	Units
$c\tau$	451	15	-14	15	μm
s	1.252	0.023	-0.023	0.023	
λ_+	141	31	-30	37	μm
f_+	0.077	0.019	-0.020	0.019	
λ_{++}	472	49	-47	58	μm
f_{++}	0.093	0.017	-0.018	0.018	
λ_-	304	36	-34	40	μm
f_-	0.0314	0.0045	-0.0043	0.0048	
f_b	0.7843	0.0070	-0.0070	0.0069	
M	5.2812	0.0004	-0.0004	0.0004	GeV/c^2
s_M	1.761	0.059	-0.058	0.060	
C_0	7.1	5.7	-5.7	5.7	GeV/c^2
λ_{psig}	13.21	0.78	-0.78	0.81	μm
σ_{psig}	5.79	0.40	-0.40	0.42	μm
μ_{psig}	21.81	0.56	-0.56	0.57	μm
λ_{pbkg}	13.77	0.40	-0.40	0.41	μm
σ_{pbkg}	7.33	0.21	-0.21	0.22	μm
μ_{pbkg}	31.40	0.30	-0.30	0.30	μm

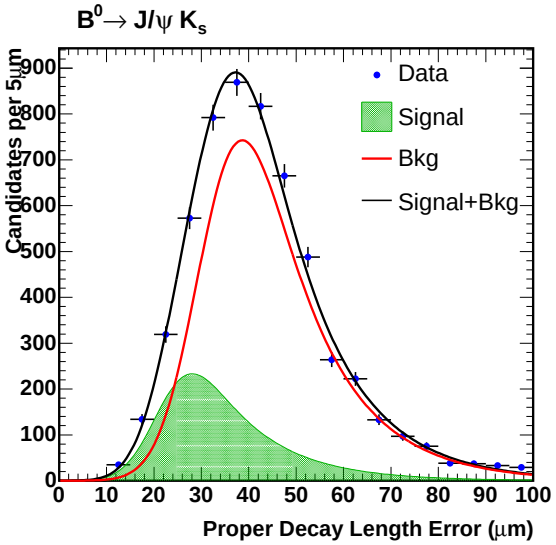
Yields:		
Component	Yield	Error
N_{signal}	1212	39
$N_{background}$	4408	39
N_{prompt}	3110	31
N_-	138	20
N_+	340	83
N_{++}	410	74

Goodness-of-fit:		
Projection	χ^2/dof	χ^2 Probability
PDL:	37.7/26	6.5%
Mass:	80.3/84	59.4%

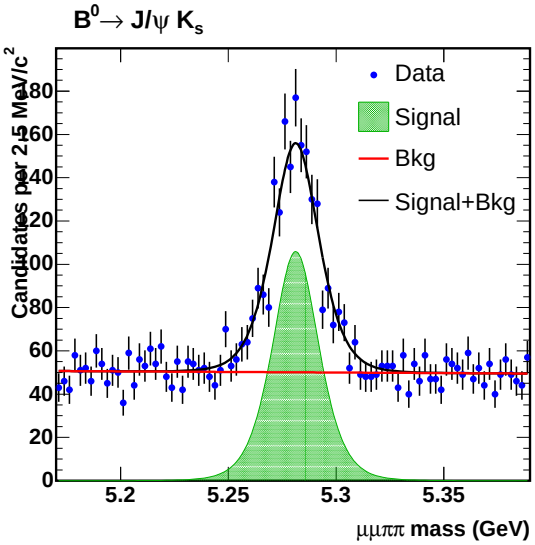
Table 7: Fit results for $B^0 \to J/\psi K^0_s$ on the data.



(a) PDL projection in lifetime fit

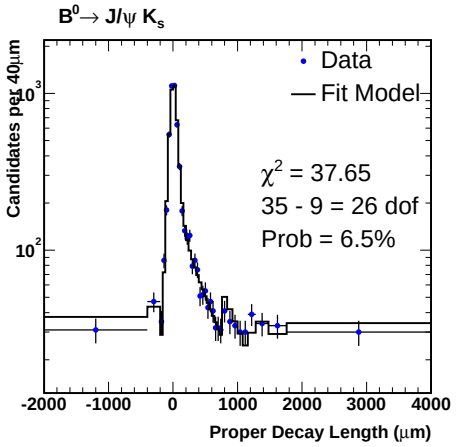


(b) PDL error projection in lifetime fit

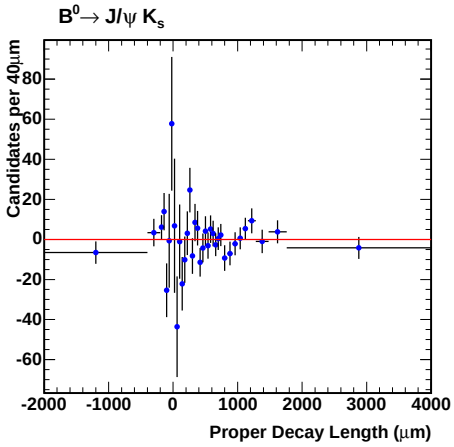


(c) Mass projection in lifetime fit

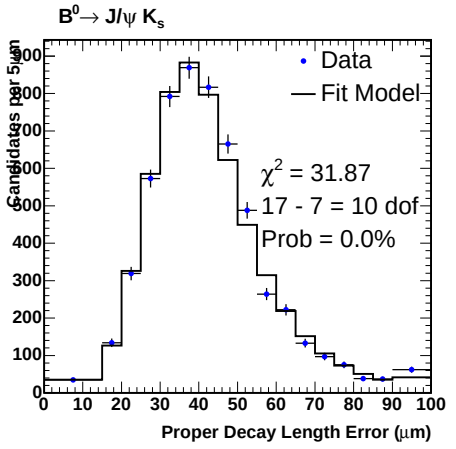
Figure 57: Projections of the $B^0 \to J/\psi K^0_s$ fit in data.



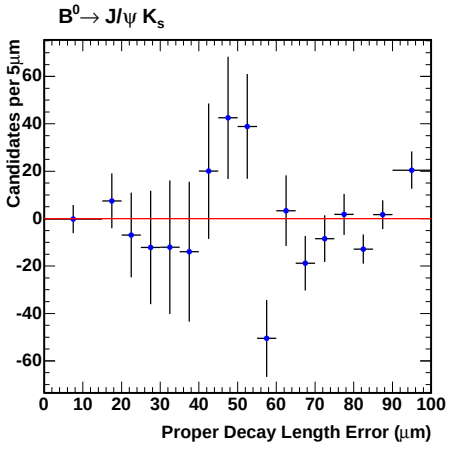
(a) Variable-binned PDL projection



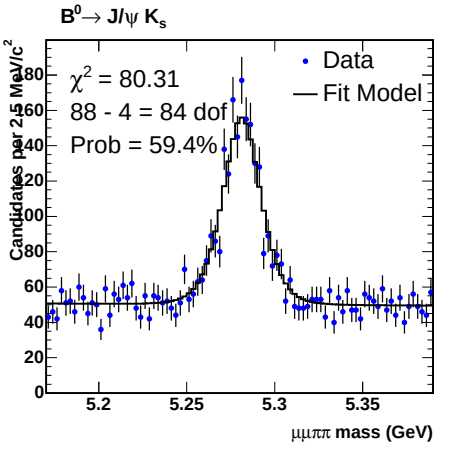
(b) Data-Model PDL projection



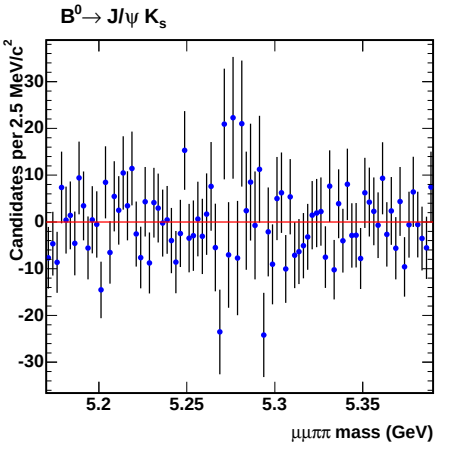
(c) Variable-binned PDL error projection



(d) Data-Model PDL error projection



(e) Variable-binned mass projection



(f) Data-Model mass projection

Figure 58: Goodness of fit tests for the $B^0 \to J/\psi K^0_s$ fit in data.

- $B^0 \to J/\psi K^{*0}$,
- $B^+ \to J/\psi K^+$,
- $B^0 \to \psi(2S) K^0_s$ ($\psi(2S) \to \mu^+\mu^-$),
- $B^0 \to \psi(2S) K^0_s$ ($\psi(2S) \to J/\psi\pi^+\pi^-$),
- $B^0 \to \psi(2S) K^{*0}$ ($\psi(2S) \to \mu^+\mu^-$),
- $B^0 \to \psi(2S) K^{*0}$ ($\psi(2S) \to J/\psi\pi^+\pi^-$),
- $B^+ \to J/\psi K^{*+}$ ($K^{*+} \to K^0_s\pi^+$),
- $B^+ \to \psi(2S) K^+$ ($\psi(2S) \to \mu^+\mu^-$),
- $B^+ \to \psi(2S) K^+$ ($\psi(2S) \to J/\psi\pi^+\pi^-$).

This purpose of these fits is two-fold one is the validate the procedure on the higher statistics modes and the other is to constrain the CDF length scale by calibrating with physics. The projections of the individual shifts are shown in Figures 59 through 64 and are summarized in Table 9 and Figure 5.4. We find that the fit model gives good agreement with data and that $c\tau$ fit results are consistent with the PDG values. The $B^+ \to J/\psi K^{*+}$ mode is particularly interesting because it also contains a K^0_s and has reasonable statistics.

5.3 Λ^0_b Lifetime

Finally we present the **blinded** results for the $\Lambda^0_b \to J/\psi\Lambda^0$ mode. The blinding procedure allows us to fit the data and assess the agreement of the data with the fit model without knowing the final fit value. In Table 8, we shoud the blinded fit results. The projections of this fit (which will not change with the unblinding) are shown in Figures 65 and 66. We find a good agreement between the data and the fit model.

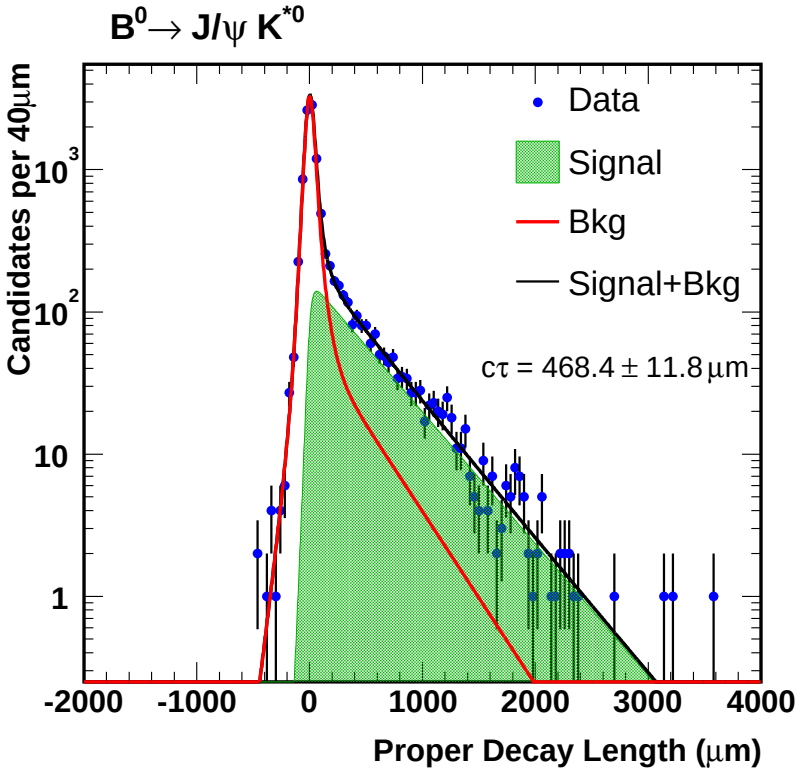
5.4 Summary of Lifetime Results

Table 9 and Figure 5.4 summarize all of the data fits. We find that all the fits agree well with the PDG. The Λ^0_b results is not on the plot because it is still blind.

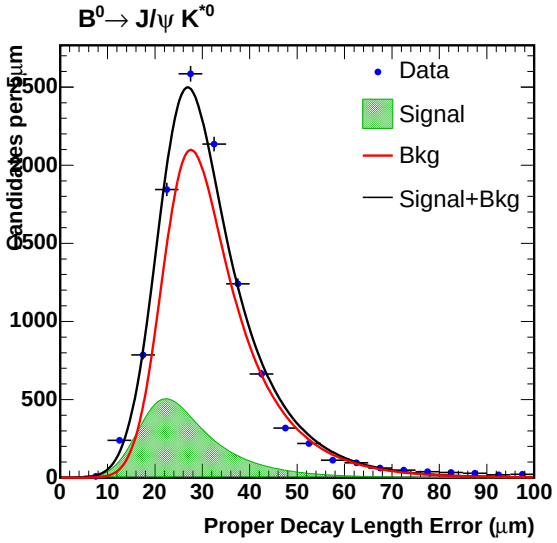
6 Sanity Checks and Systematics

6.1 Fitter Bias

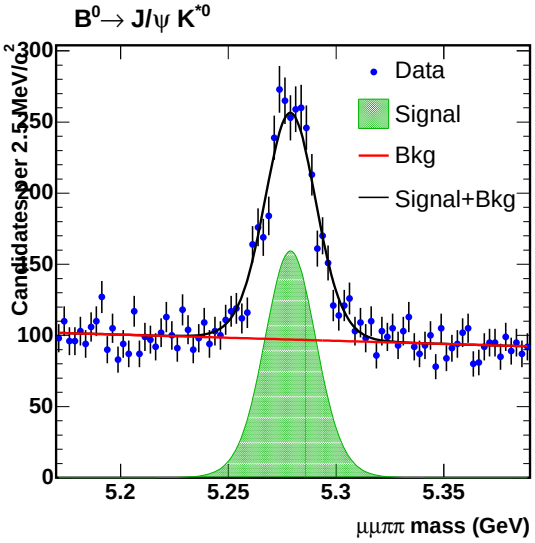
The lifetime due to fitting procedure in $B^0 \to J/\psi K^0_s$ has been measured using 9600 toy Monte Carlo datasets in Section 4.6.1. This study was done toying the same model



(a) PDL projection in lifetime fit for $B^0 \to J/\psi K^{*0}$ candidates

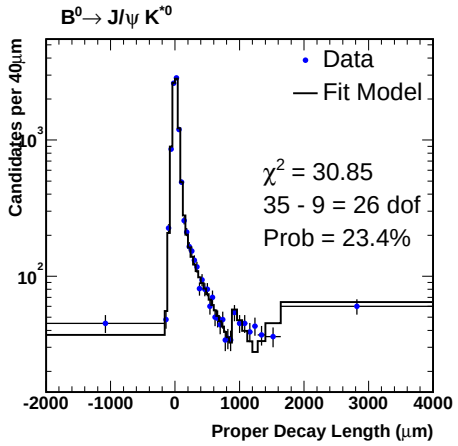


(b) PDL error projection in lifetime fit for $B^0 \to J/\psi K^{*0}$ candidates

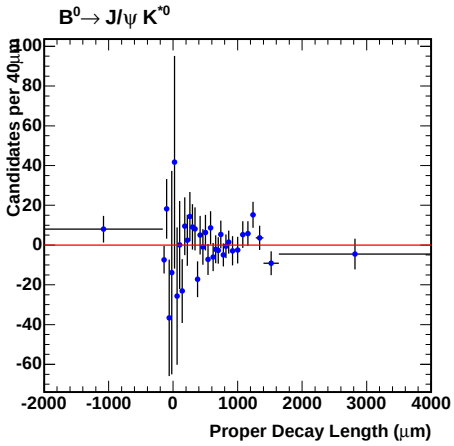


(c) Mass projection in lifetime fit for $B^0 \to J/\psi K^{*0}$ candidates

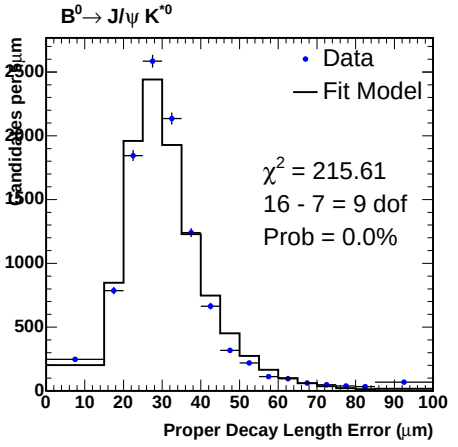
Figure 59: Projections of the $B^0 \to J/\psi K^{*0}$ fit in data.



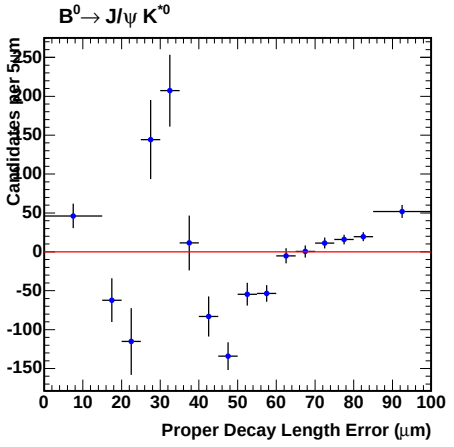
(a) Variable-binned PDL projection



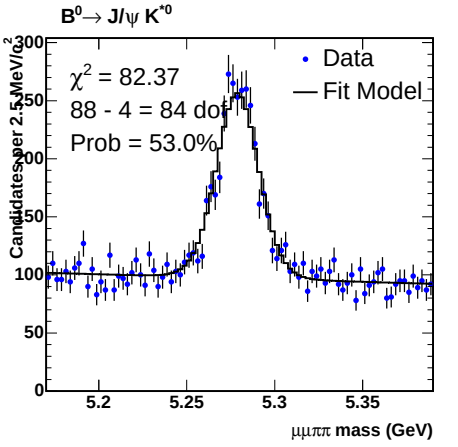
(b) Data-Model PDL projection



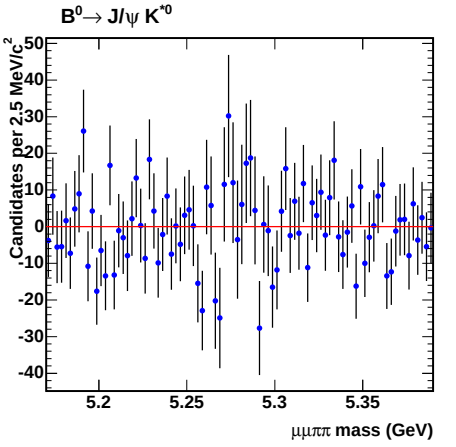
(c) Variable-binned PDL error projection



(d) Data-Model PDL error projection

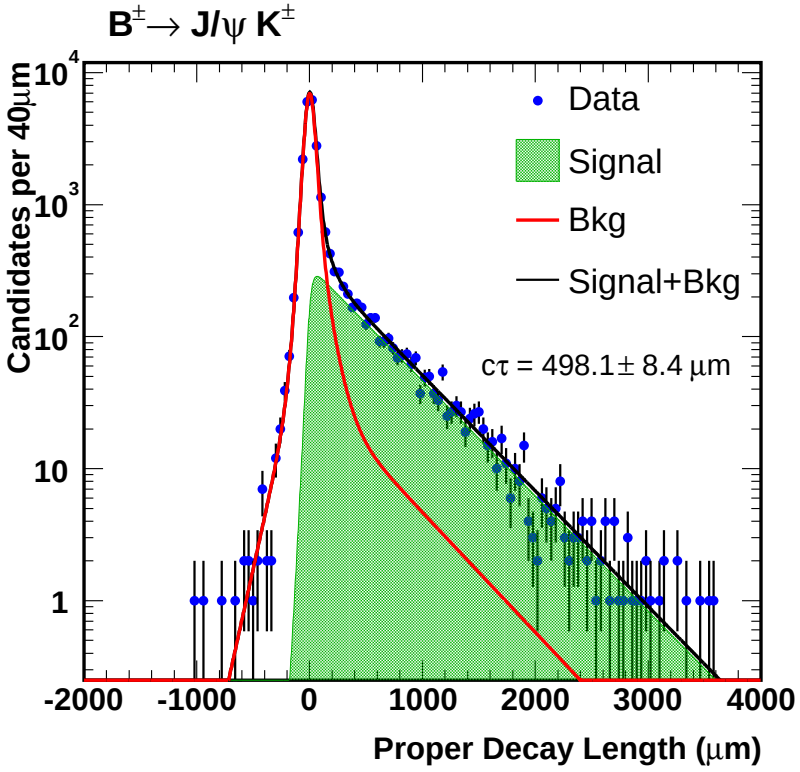


(e) Variable-binned mass projection

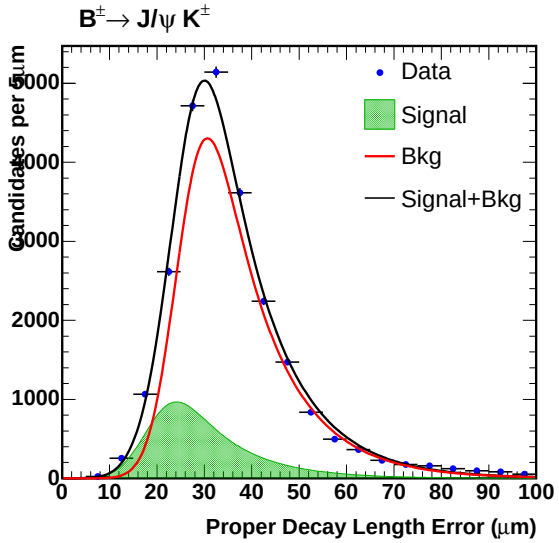


(f) Data-Model mass projection

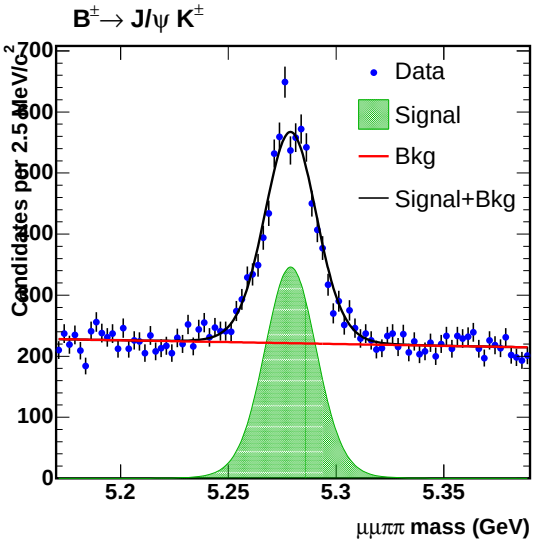
Figure 60: Goodness of fit tests for the $B^0 \to J/\psi K^{*0}$ fit in data.



(a) PDL projection in lifetime fit for $B^+ \to J/\psi K^+$ candidates

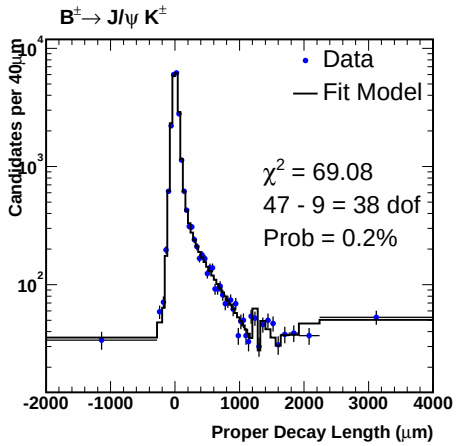


(b) PDL error projection in lifetime fit for $B^+ \to J/\psi K^+$ candidates

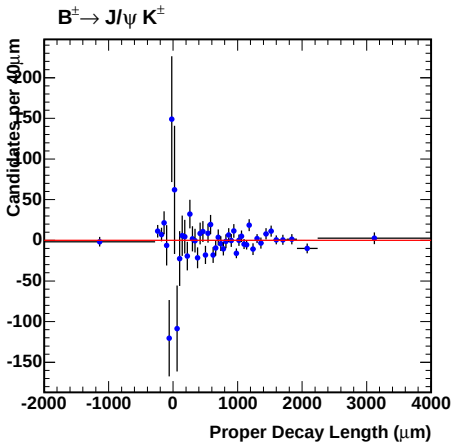


(c) Mass projection in lifetime fit for $B^+ \to J/\psi K^+$ candidates

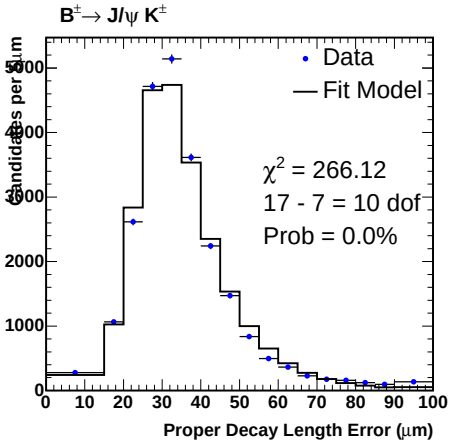
Figure 61: Projections of the $B^+ \to J/\psi K^+$ fit in data.



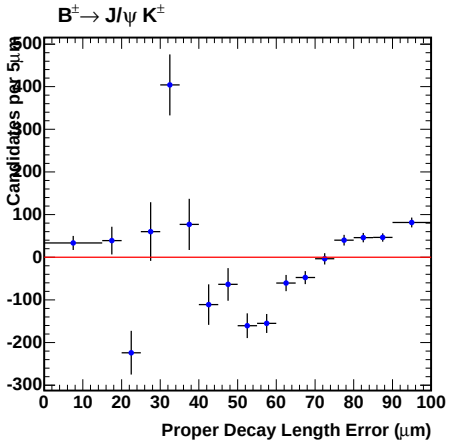
(a) Variable-binned PDL projection



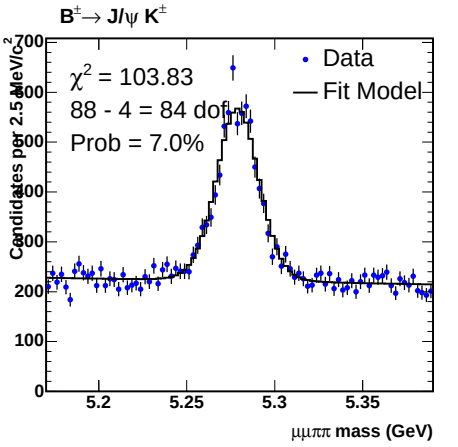
(b) Data-Model PDL projection



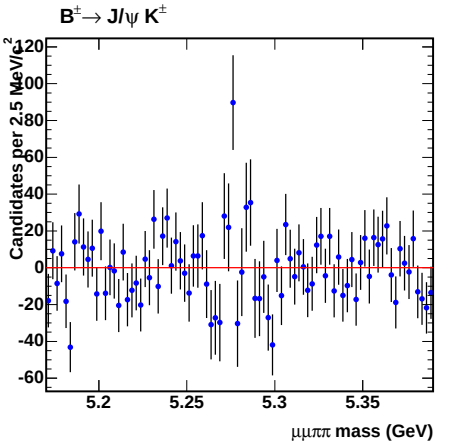
(c) Variable-binned PDL error projection



(d) Data-Model PDL error projection

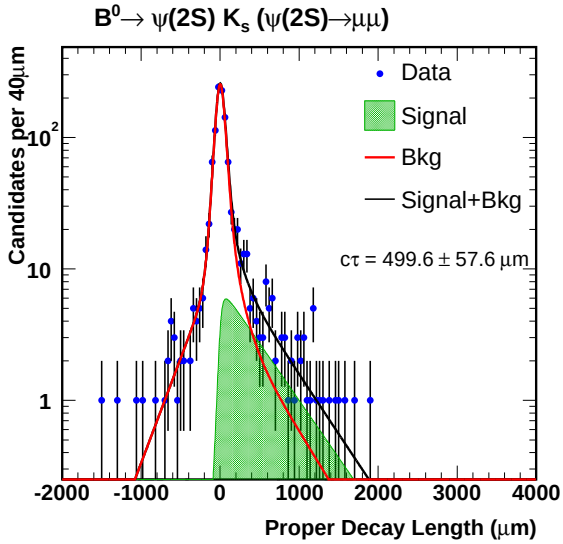


(e) Variable-binned mass projection

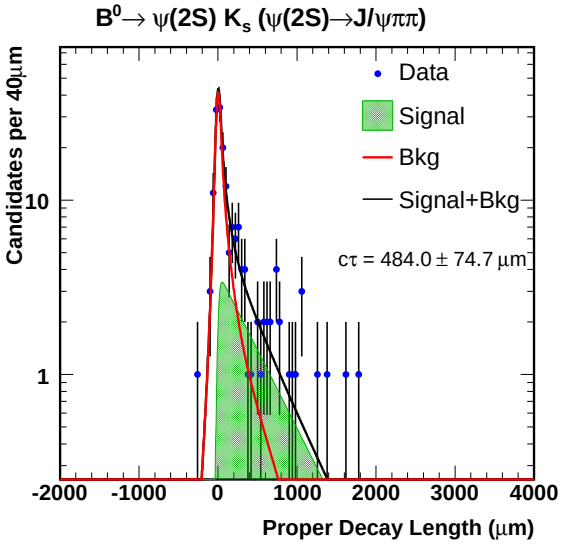


(f) Data-Model mass projection

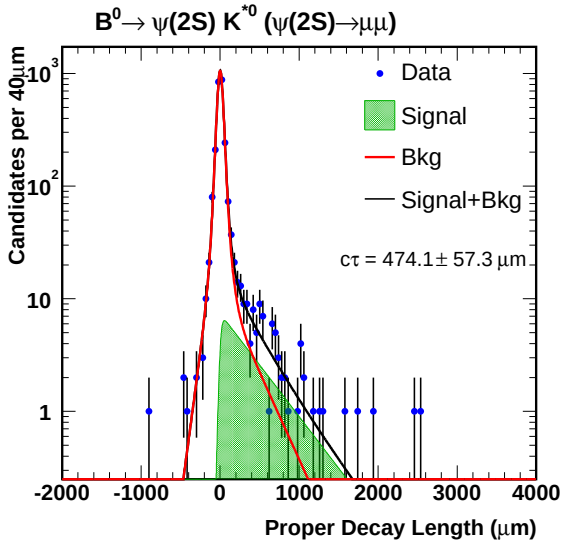
Figure 62: Goodness of fit tests for the $B^+ \to J/\psi K^+$ fit in data.



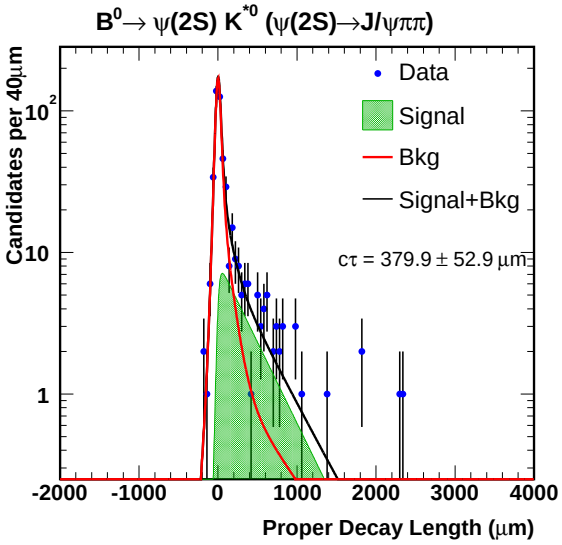
(a) PDL projection in lifetime fit for $B^0 \to \psi(2S)K^0_s$ ($\psi(2S) \to \mu^+\mu^-$) candidates



(b) PDL projection in lifetime fit for $B^0 \to \psi(2S)K^0_s$ ($\psi(2S) \to J/\psi\pi^+\pi^-$) candidates

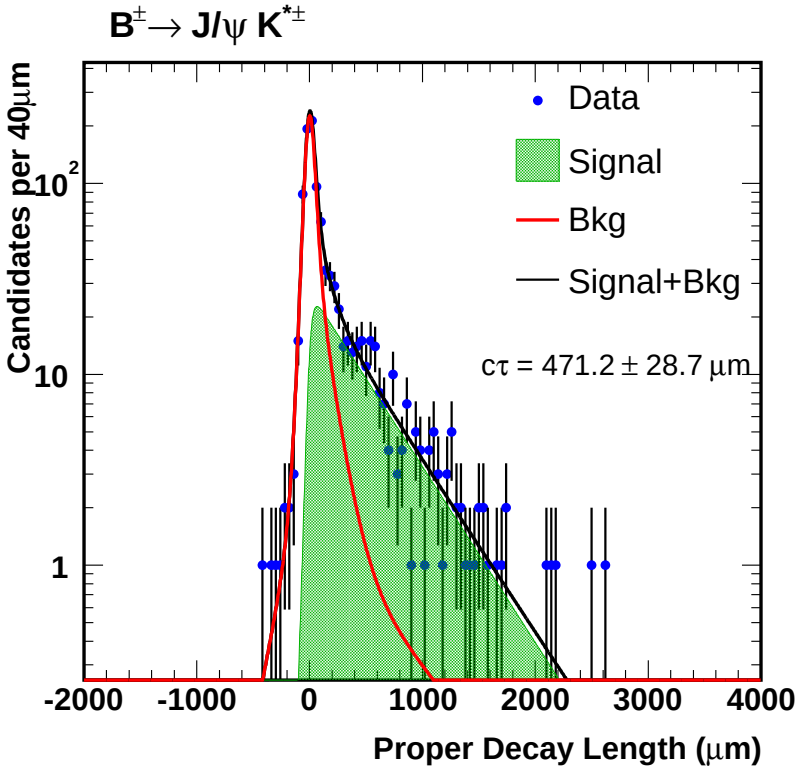


(c) PDL projection in lifetime fit for $B^0 \to \psi(2S)K^{*0}$ ($\psi(2S) \to \mu^+\mu^-$) candidates

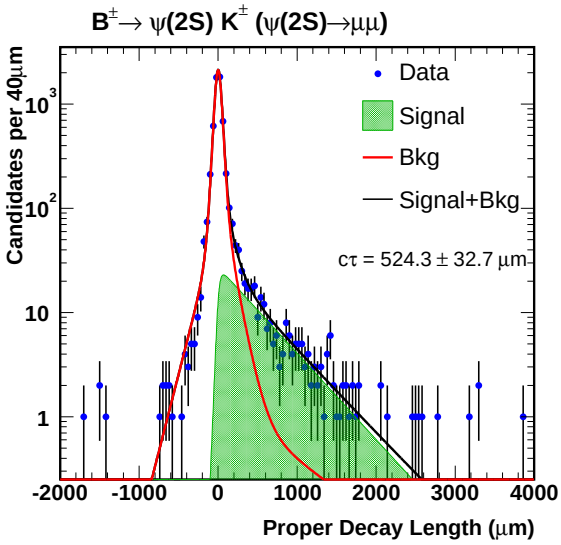


(d) PDL projection in lifetime fit for $B^0 \to \psi(2S)K^{*0}$ ($\psi(2S) \to J/\psi\pi^+\pi^-$) candidates

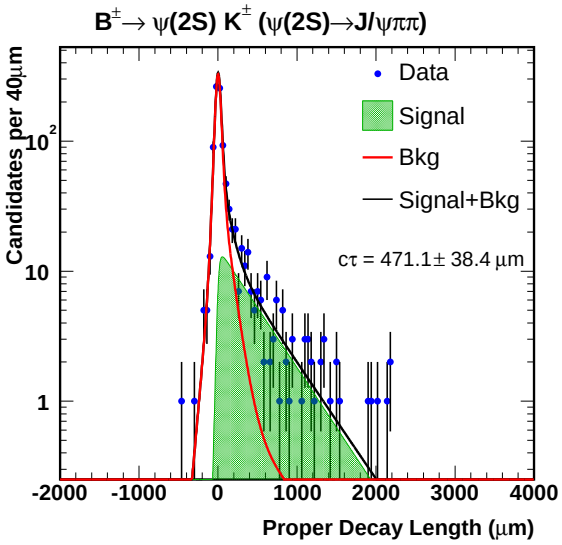
Figure 63: Lifetime projections from the B^0 modes reconstructed with $\psi(2S)$



(a) PDL projection in lifetime fit for $B^+ \to J/\psi K^{*+}$ ($K^{*+} \to K^0_s \pi^+$) candidates

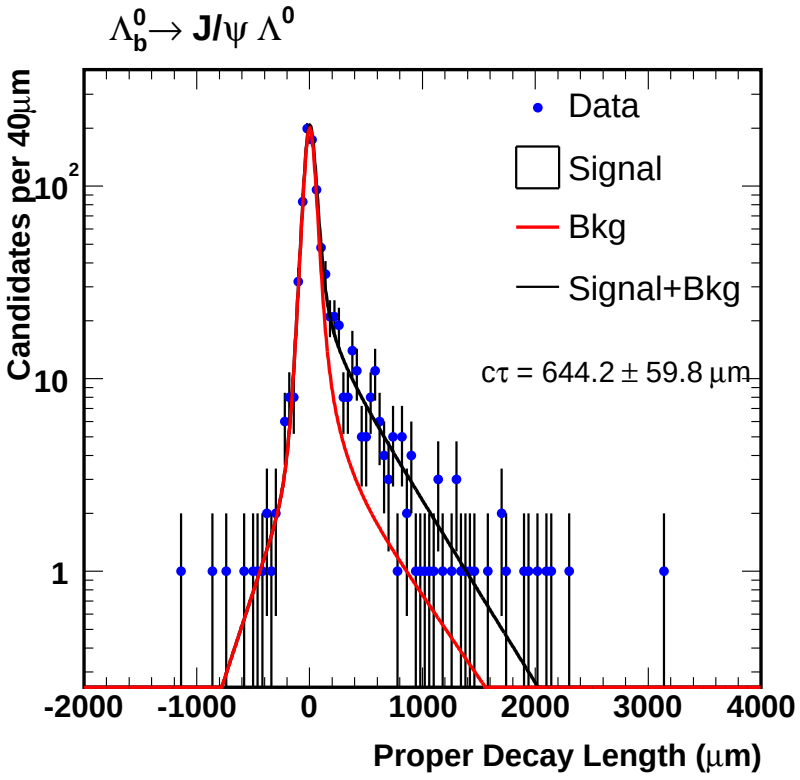


(b) PDL projection in lifetime fit for $B^+ \to \psi(2S)K^+$ ($\psi(2S) \to \mu^+\mu^-$) candidates

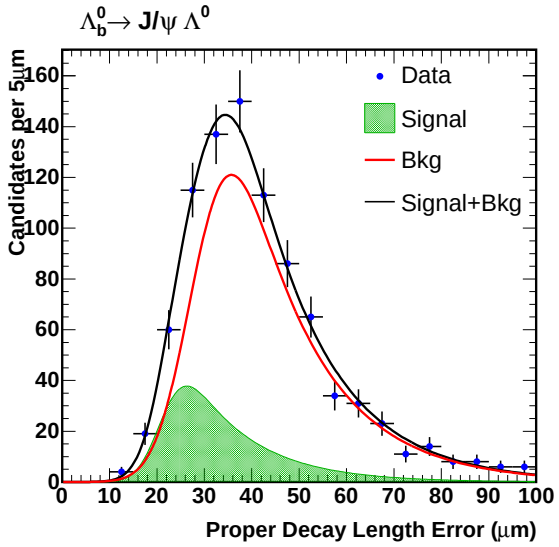


(c) PDL projection in lifetime fit for $B^+ \to \psi(2S)K^+$ ($\psi(2S) \to J/\psi\pi^+\pi^-$) candidates

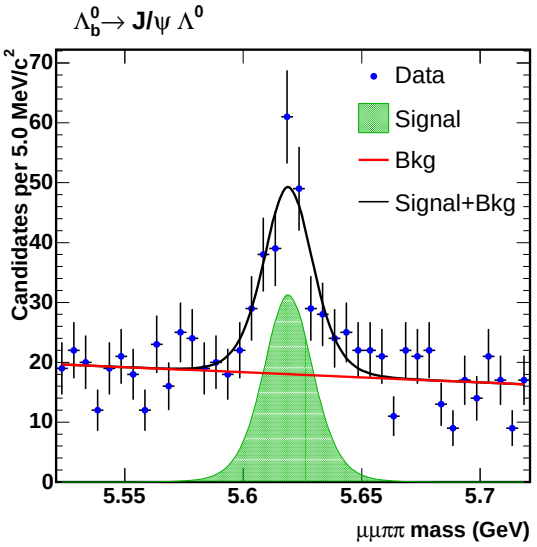
Figure 64: Lifetime projections of the B^+ modes reconstructed with $\psi(2S)$ or $K^{*+} \to K^0_s \pi^+$



(a) PDL projection in lifetime fit for $\Lambda_b^0 \to J/\psi\Lambda^0$ candidates

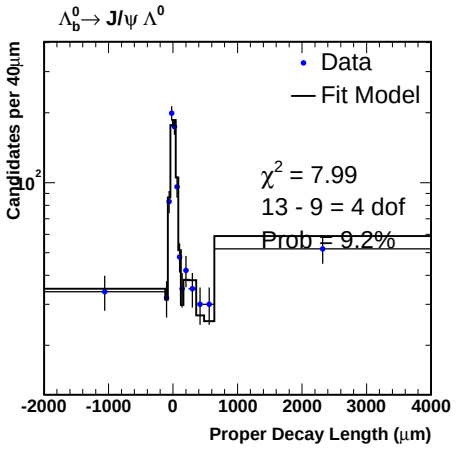


(b) PDL error projection in lifetime fit for $\Lambda_b^0 \to J/\psi\Lambda^0$ candidates

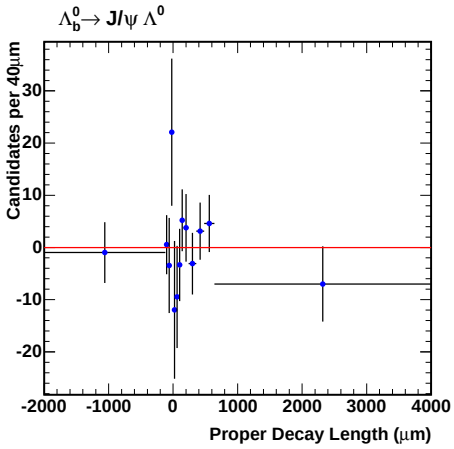


(c) Mass projection in lifetime fit for $\Lambda_b^0 \to J/\psi\Lambda^0$ candidates

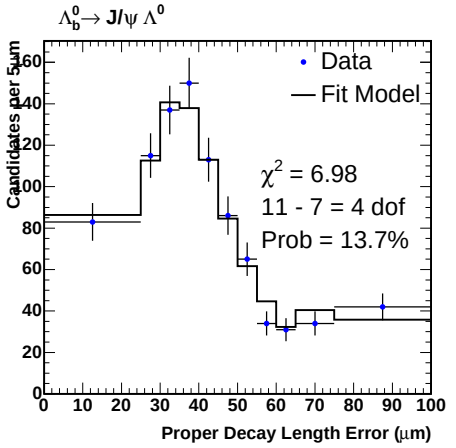
Figure 65: Projections of the $\Lambda_b^0 \to J/\psi\Lambda^0$ fit in data.



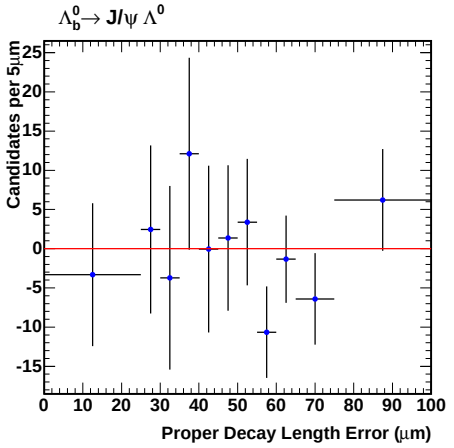
(a) Variable-binned PDL projection



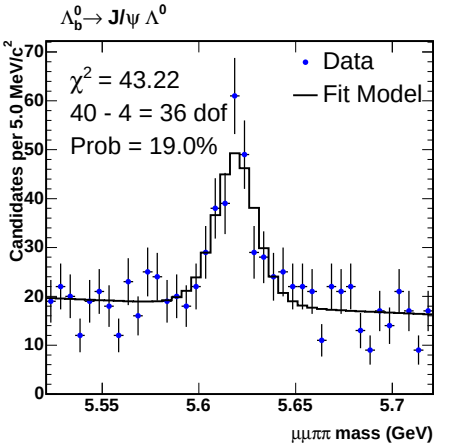
(b) Data-Model PDL projection



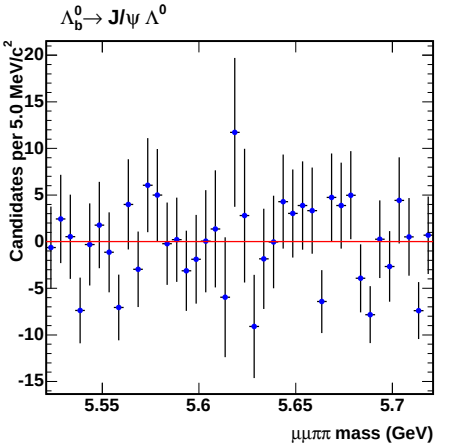
(c) Variable-binned PDL error projection



(d) Data-Model PDL error projection



(e) Variable-binned mass projection



(f) Data-Model mass projection

Figure 66: Goodness of fit tests for the $\Lambda_b^0 \to J/\psi\Lambda^0$ fit in data.

Model parameters:					
Parameter	Fit value	$\pm 1\sigma$ Sym Err	-1σ Minos Err	+1σ Minos Err	Units
$c\tau$	644	60	-58	63	μm
s	1.226	0.057	-0.057	0.058	
λ_+	112	61	-45	118	μm
f_+	0.077	0.038	-0.041	0.038	
λ_{++}	507	118	-101	177	μm
f_{++}	0.094	0.030	-0.044	0.031	
λ_-	244	57	-49	69	μm
f_-	0.050	0.014	-0.013	0.016	
f_b	0.808	0.017	-0.017	0.017	
M	5.6190	0.0009	-0.0009	0.0009	GeV/c^2
s_M	1.50	0.12	-0.12	0.13	
C_0	32	19	-19	18	GeV/c^2
λ_{psig}	13.5	2.0	-1.9	2.1	μm
σ_{psig}	4.04	0.78	-0.71	0.88	μm
μ_{psig}	21.2	1.0	-1.0	1.1	μm
λ_{pbkg}	15.8	1.1	-1.0	1.1	μm
σ_{pbkg}	6.43	0.50	-0.49	0.51	μm
μ_{pbkg}	28.54	0.68	-0.68	0.69	μm

Yields:		
Component	Yield	Error
N_{signal}	171	15
$N_{background}$	719	15
N_{prompt}	492	12
N_-	36	10
N_+	56	28
N_{++}	68	21

Goodness-of-fit:		
Projection	χ^2/dof	χ^2 Probability
PDL:	7.99/4	9.2%
Mass:	43.22/36	19.0%

Table 8: Fit results for $\Lambda_b^0 \to J/\psi\Lambda^0$ on the data.

Mode	Fitted $c\tau$ (μm)
$B^+ \to J/\psi K^+$	498.1 ± 8.4
$B^+ \to J/\psi K^{*+}$ ($K^{*+} \to K^0_s \pi^+$)	471.2 ± 28.7
$B^+ \to \psi(2S) K^+$ ($\psi(2S) \to J/\psi \pi^+ \pi^-$)	471.1 ± 38.4
$B^+ \to \psi(2S) K^+$ ($\psi(2S) \to \mu^+ \mu^-$)	524.3 ± 32.7
$B^0 \to J/\psi K^{*0}$	468.4 ± 11.8
$B^0 \to \psi(2S) K^{*0}$ ($\psi(2S) \to J/\psi \pi^+ \pi^-$)	379.9 ± 52.9
$B^0 \to \psi(2S) K^{*0}$ ($\psi(2S) \to \mu^+ \mu^-$)	474.1 ± 57.3
$B^0 \to J/\psi K^0_s$	450.6 ± 14.7
$B^0 \to \psi(2S) K^0_s$ ($\psi(2S) \to J/\psi \pi^+ \pi^-$)	484.0 ± 74.7
$B^0 \to \psi(2S) K^0_s$ ($\psi(2S) \to \mu^+ \mu^-$)	499.6 ± 57.6
$\Lambda^0_b \to J/\psi \Lambda^0$ (**blinded value**)	643.5 ± 59.7

Table 9: Fit results for $B^0 \to J/\psi K^0_s$ on the data.

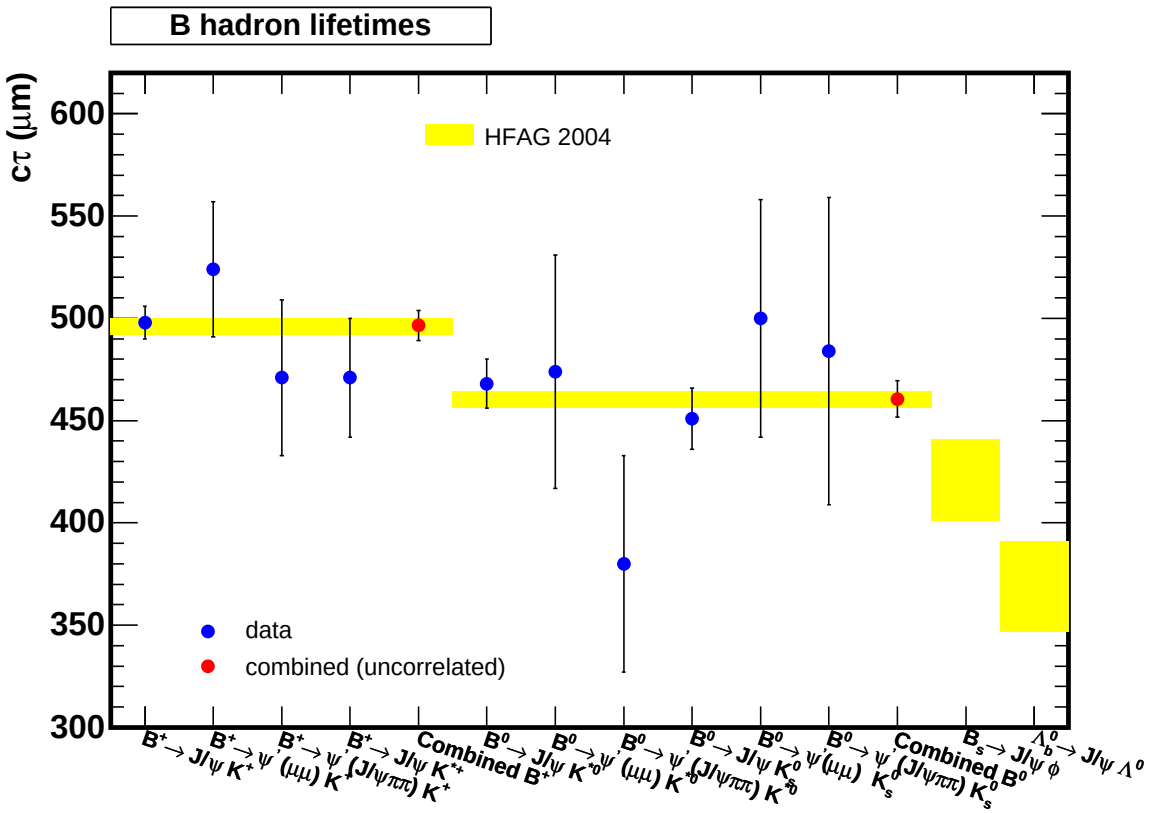


Figure 67: Summary of lifetime fits to b-meson modes. HFAG 2004 shown in yellow

parameters as our baseline $B^0 \to J/\psi K^0_s$ fit to data, so this is a measure of the lifetime fitting bias. We quote a 0.2μm systematic for $B^0 \to J/\psi K^0_s$.

The Λ^0_b fitting bias will be completed using the same procedure one we unblind the fit.

6.2 V^0 Pointing

6.3 Studies of V^0 Pointing

6.3.1 The importance of V^0 tracking

Much of the tracking code has been optimized to be efficient and unbiased for tracks originating close to the interaction point. This analysis depends on using reconstructed tracks originating from V^0 decays that can have vertices as far as tens of centimeters from the beam line. It is therefore important to consider possible biases that originate from the V^0 tracking.

In order to a priori minimize any such effect, we have chosen to the J/ψ vertex instead of the b-hadron vertex in the lifetime fit. This means that it is impossible for the V^0 tracking to pull the $c\tau$ used in the lifetime fit. The V^0 momentum does appear in the b-hadron momentum in the $c\tau$ calculation, but a significant bias in the momentum would make it hard to reconstruct a b-mass peak and would still have a very small effect on the measured $c\tau$.

The V^0 however is used in the selection. The V^0 pointing constraint both in two and, even more so, in three dimensions is effective in suppressing background. A bias in the V^0 tracking could in principal create a variation in the efficiency as a function of the b-hadron $c\tau$ and thereby bias the measurement. It should be noted that not all V^0 reconstruction biases will bias the measurement. It must be one that is specifically a function of the b-hadron $c\tau$. If the V^0-bias is not dependent on the b-hadron $c\tau$, it cannot create a slope in the b-hadron $c\tau$ efficiency and will therefore not enter into the measurement.

6.3.2 The relevant variables

Since we are most concerned about the V^0 pointing constraint biasing the two-dimensional b-hadron $c\tau$, we consider the variable which describes the difference between the related L_{xy}, measured by the J/ψ vertex (used in the lifetime fit) and that determined by the combination of the b flight direction and the V^0 vertex (which only effects selection),

$$\Delta L_{xy} \equiv L^b_{xy}(V^0) - L^b_{xy}(J/\psi). \tag{19}$$

Figure 68 diagrams this variable. This variable contains the pointing information used in the two-dimensional pointing constraint and is oriented to have a meaningful sign and scale. For the three-dimensional pointing we also study z_0 of the V^0, the z-displacement of the V^0 from the J/ψ vertex at the point of closest approach.

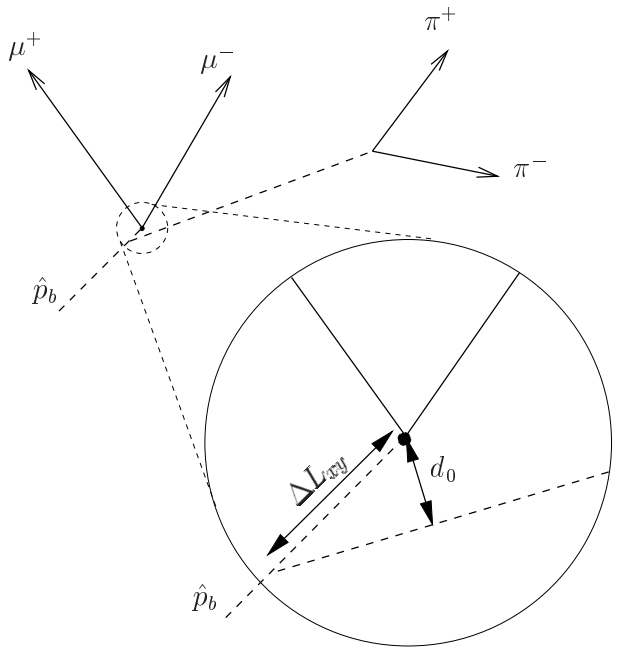


Figure 68: Diagram of the definition of the ΔL_{xy} variable.

All of the ΔL_{xy} and z_0 studies are conducted using the same event selection except the pointing constraint in the vertex fit is not used. The ΔL_{xy} and z_0 distributions are therefor not baised by the selection and correspond to the distributions the pointing constraint vertex fits are actually seeing.

In Figure 69, we show the ΔL_{xy} distribution for the $B^0 \to J/\psi K_s^0 \to J/\psi X$ MC sample. Notice the background is displaced from zero because it is J/ψ from a b-hadron combined with a K_s^0 from another vertex (other b or prompt). The signal is well centered at zero, indicating no overall bias. The narrow and broad gaussians are from K_s^0 particles with and without silicon hits respectively. Also shown in Figure 69 is the pull distribution for ΔL_{xy}.

In Figure 70 are the analogous distributions in data. Note that the large prompt background component is centered at zero (both the J/ψ and the K_s^0 are from the primary vertex). Also the widths are substantially larger ($\approx 50\%$), but the pull distribution is consistent indicating that the tracking and vertex fitting correctly describe the difference (although still at scale factor of $\approx 30\%$).

In Figure 71 are the same distributions for the Λ_b^0 BGen Monte Carlo (the $b \to J/\psi X$ Monte Carlo sample does not have sufficient statistics to be useful). And finally, in Figure 72 are the corresponding data distributions.

There is no evidence of an overall bias in the ΔL_{xy} distributions or pulls for either K_s^0 or Λ modes. The pull distributions are however not unit gaussians. This is probably a

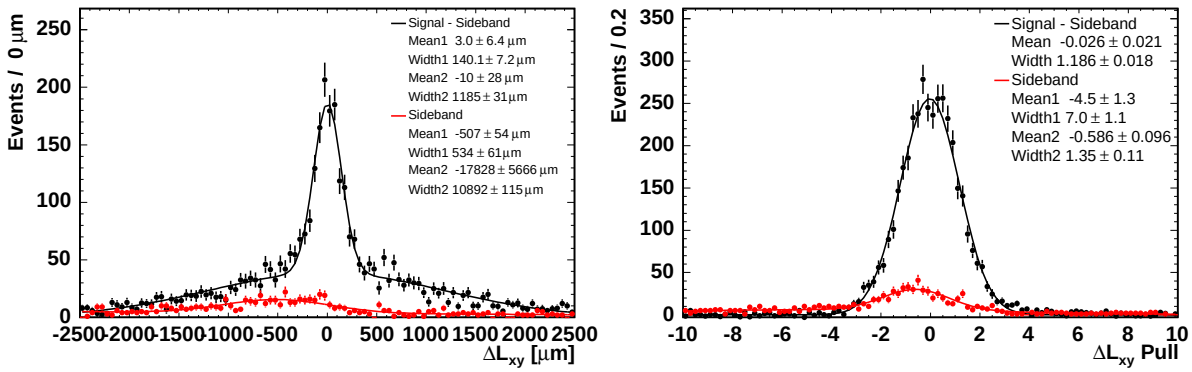


Figure 69: Distribution of ΔL_{xy} and the related pull distribution for the $B^0 \to J/\psi K^0_s$ mode in the $b \to J/\psi X$ Monte Carlo sample

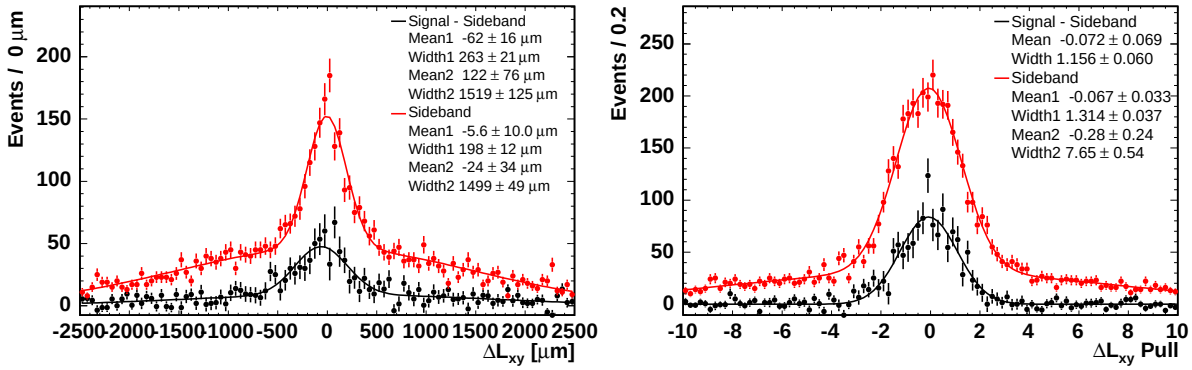


Figure 70: Distribution of ΔL_{xy} and the related pull distribution for the $B^0 \to J/\psi K^0_s$ mode in the data.

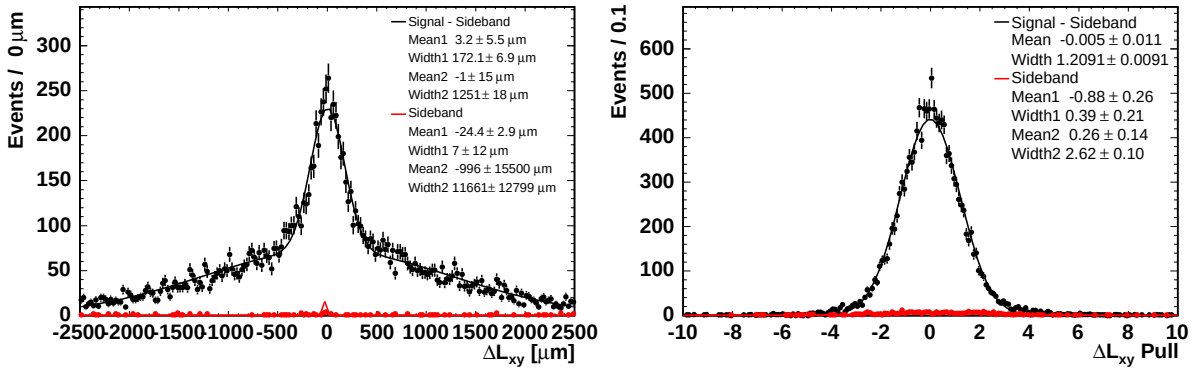


Figure 71: Distribution of ΔL_{xy} and the related pull distribution for the $\Lambda^0_b \to J/\psi \Lambda^0$ mode in the Λ^0_b BGen Monte Carlo.

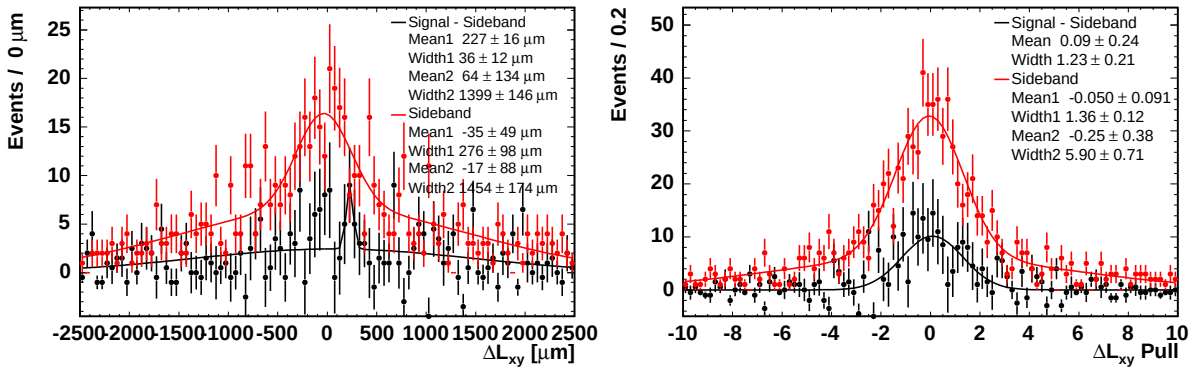


Figure 72: Distribution of ΔL_{xy} and the related pull distribution for the $\Lambda_b^0 \to J/\psi\Lambda^0$ mode in the data.

major contribution to the failure of the vertex probability distribution to be flat.

6.3.3 V^0 Pointing Dependence on b-hadron $c\tau$

In this section, we look for and seek to place limits on any variation of the V^0 pointing parameters on the $c\tau$ of the b-hadron ($c\tau_b$). Any dependence would lead directly to a bias on the measured $c\tau$ through a variation of the efficiency due to the vertex probability cut.

In Figure 73, we show the ΔL_{xy} pull distribution for the $B^0 \to J/\psi K_s^0$ mode in the data in four bins of $c\tau_b$. We only show the pulls because the statistics are not sufficient to observer separately the two gaussians in the ΔL_{xy} distribution. In each of the four bins the signal is fit to a single gaussian. The results of the fits are plotted in Figure 74. We fit the $c\tau$ dependence of the pull distributions using the mean $c\tau$ for a $460\mu m$ exponential in that bin as the center. We find no statistically significant slope. The analogous plots for the $\Lambda_b^0 \to J/\psi\Lambda^0$ mode are shown in Figures 75 and 76. Again no statistically significant slopes are seen.

All the same plots above can also be made for the z_0 of the K_s^0 and z_0 of the λ parameters and are shown in Figures 77 to 82.

6.3.4 Derivation of a the V^0-Pointing Systematic Error

At this point, we have fit the ΔL_{xy} and z_0 pull distributions as a function of the proper decay length. These fits describe the shapes of distributions on which we are cutting indirectly through the b-hadron vertex probablity. This is a five dimensional χ^2 probability, where the dimensions are ΔL_{xy}, z_0, the miss distance of the two K_s^0 tracks, miss distance of the two J/ψ tracks, and the J/ψ mass.

We present first a calculation of the bias due to these cuts using the simplifying assumption that a direct cut is made on only one of the ΔL_{xy} and z_0 guassian distributions. We will call the cut value A. Smaller A corresponds to a tighter cut and hence a larger

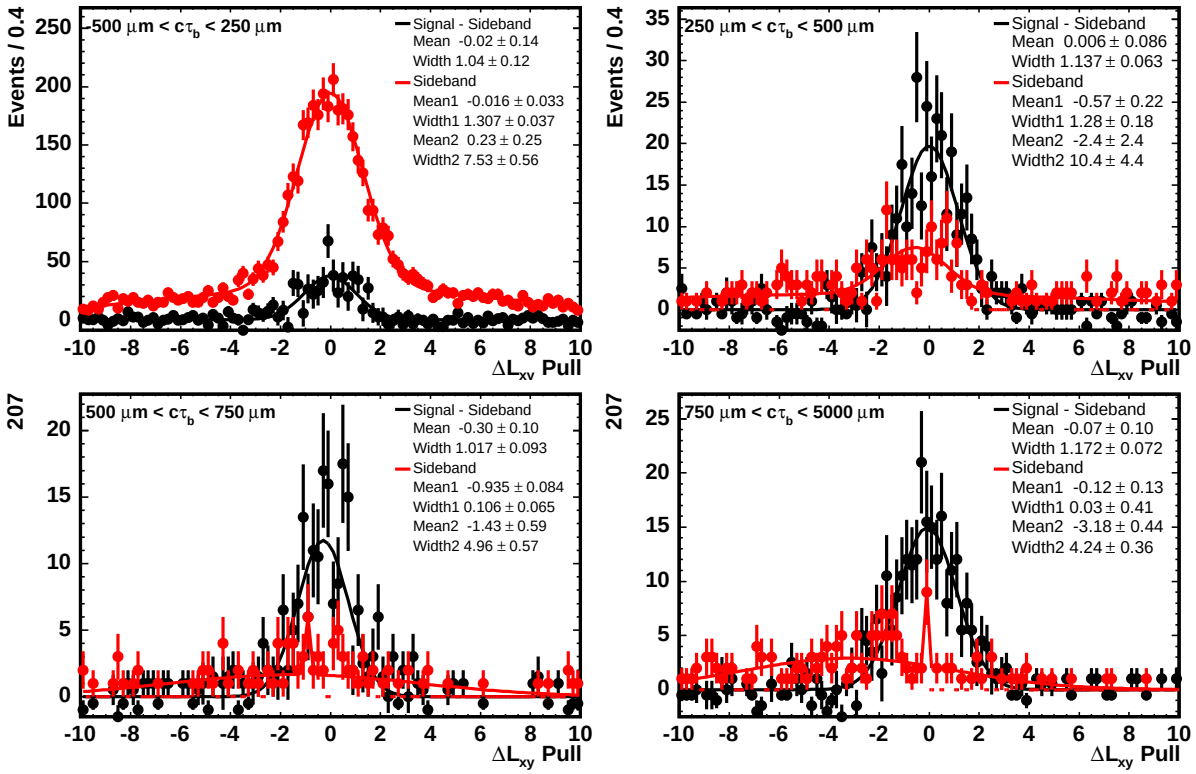


Figure 73: Distribution of ΔL_{xy} pull in four bins of b-hadron $c\tau$ for the $B^0 \to J/\psi K^0_s$ mode in the data.

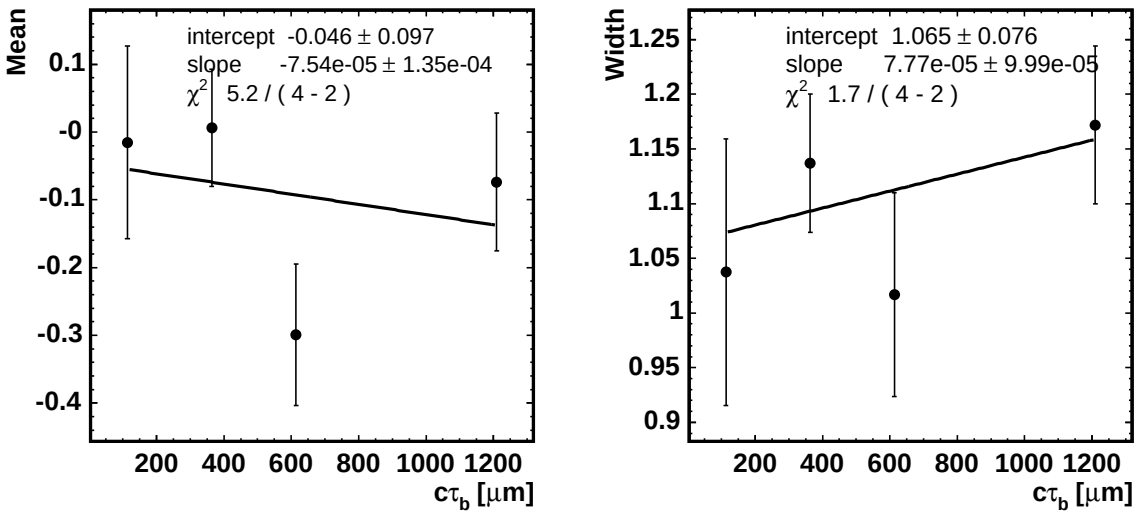


Figure 74: Results of gaussian fits to the ΔL_{xy} pull in the four bins shown in Figure 73

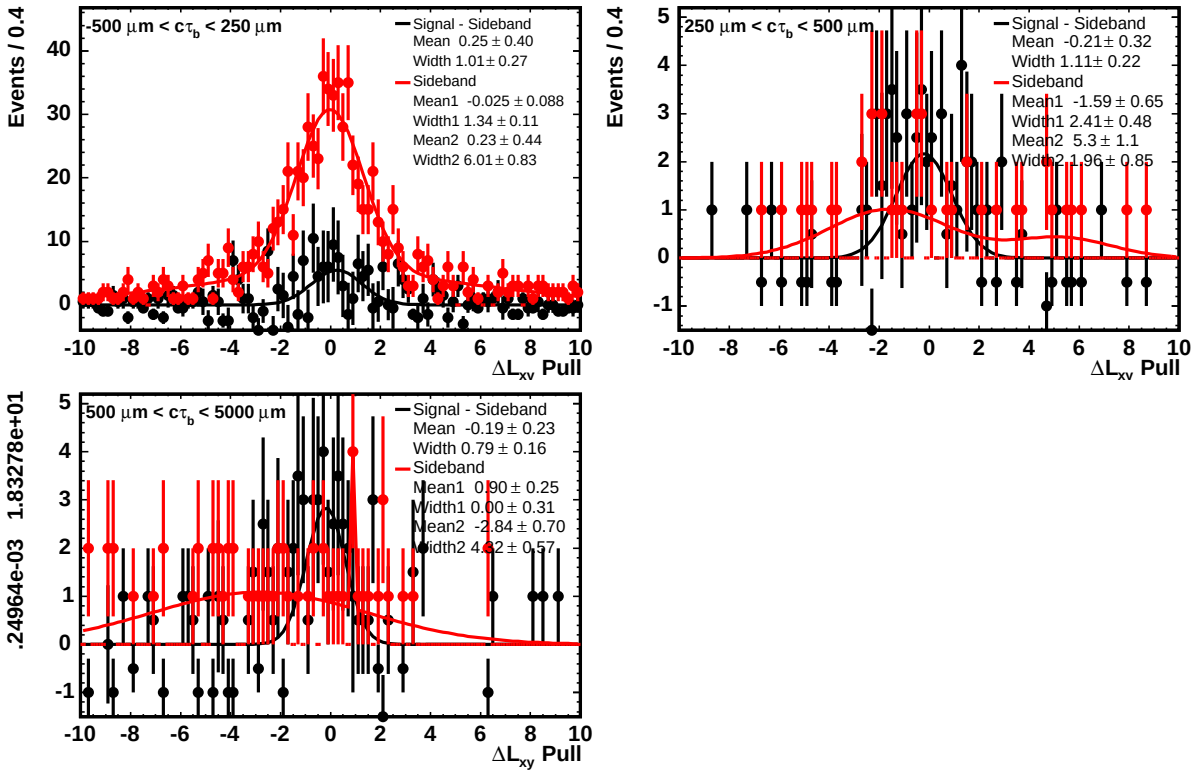


Figure 75: Distribution of ΔL_{xy} pull in three bins of b-hadron $c\tau$ for the $\Lambda_b^0 \to J/\psi\Lambda^0$ mode in the data.

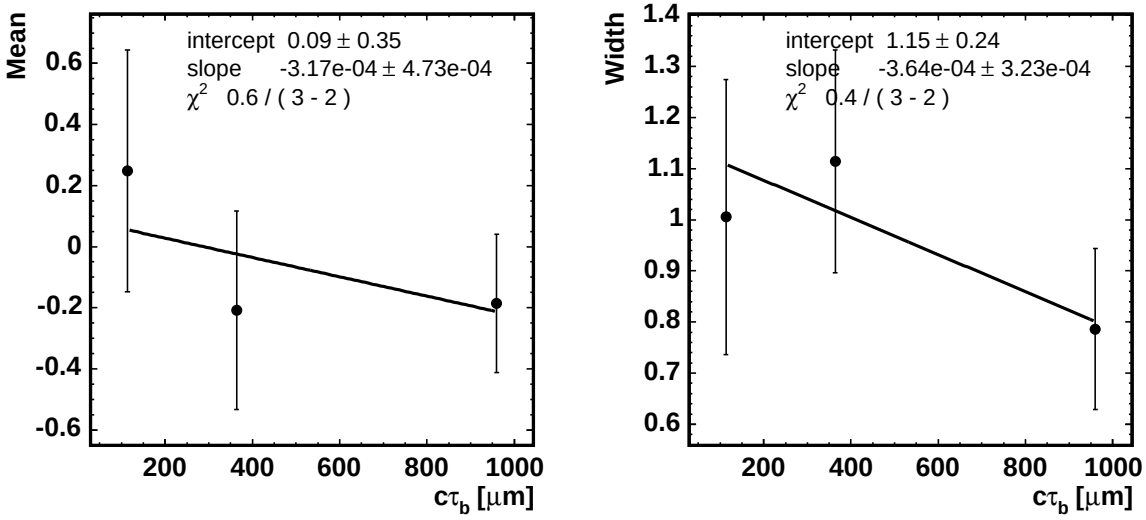


Figure 76: Results of gaussian fits to the ΔL_{xy} pull in the four bins shown in Figure 75

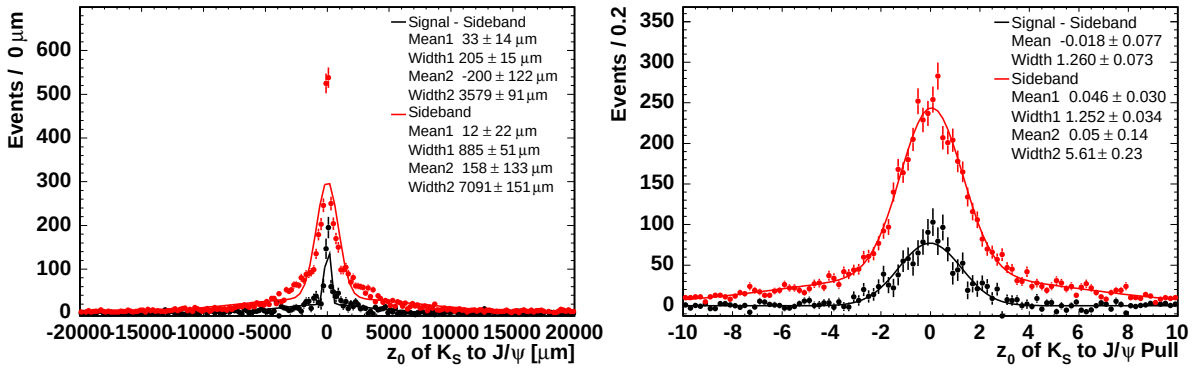


Figure 77: Distribution of z_0 of the K^0_s and the related pull distribution for the $B^0 \to J/\psi K^0_s$ mode in the data.

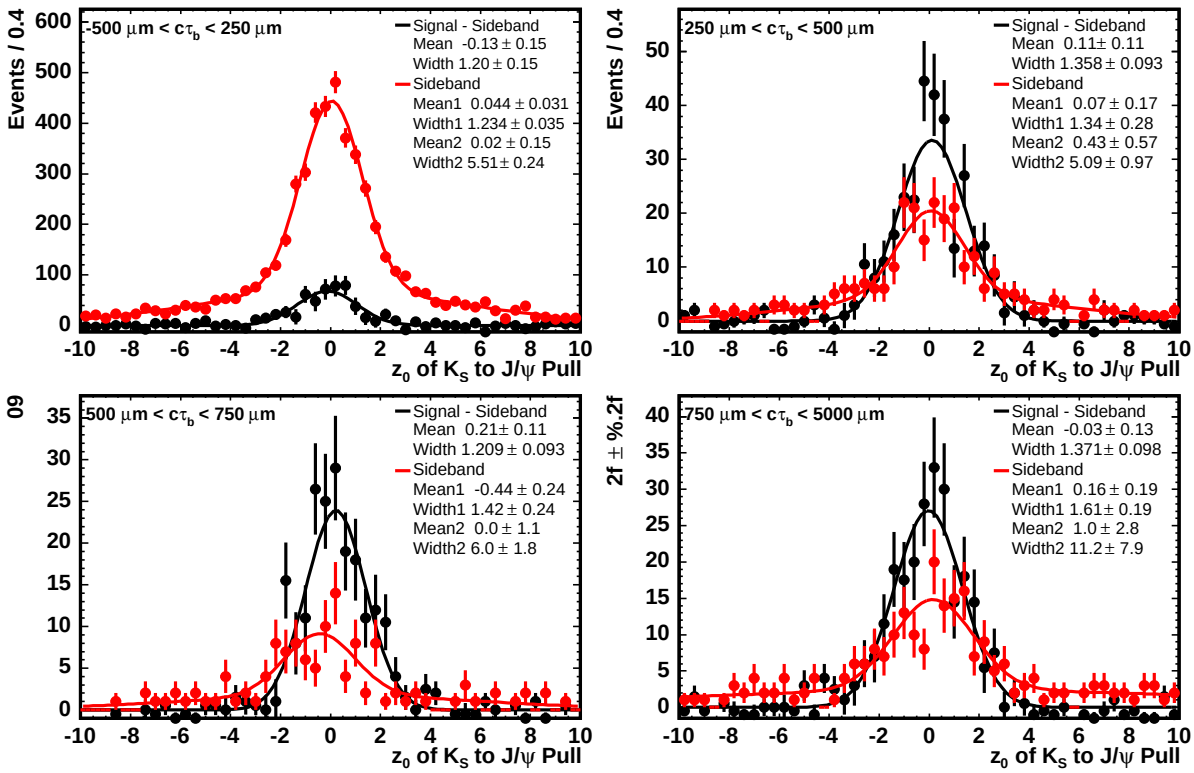


Figure 78: Distribution of z_0 of the K^0_s pull in four bins of b-hadron $c\tau$ for the $B^0 \to J/\psi K^0_s$ mode in the data.

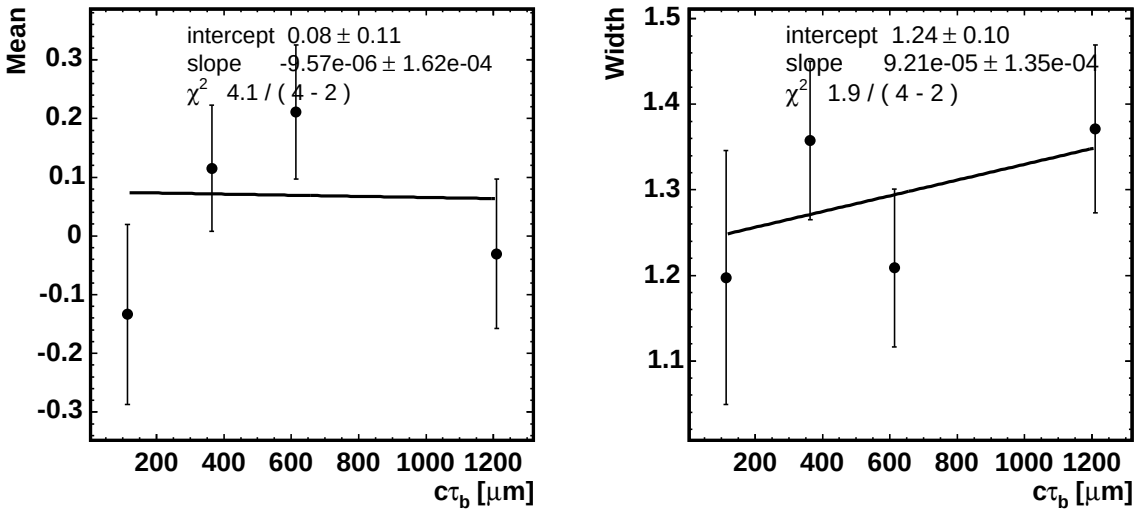


Figure 79: Results of gaussian fits to the z_0 of the K^0_s pull in the four bins shown in Figure 78

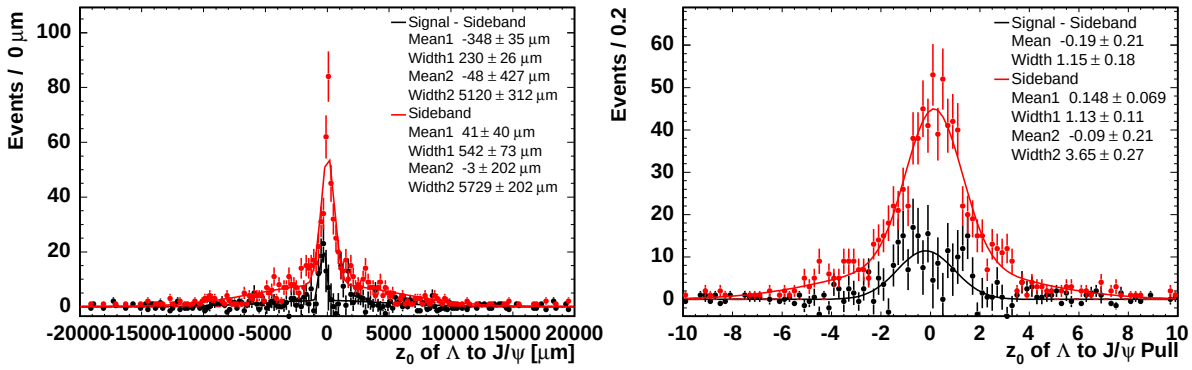


Figure 80: Distribution of z_0 of the Λ and the related pull distribution for the $B^0 \to J/\psi K^0_s$ mode in the data.

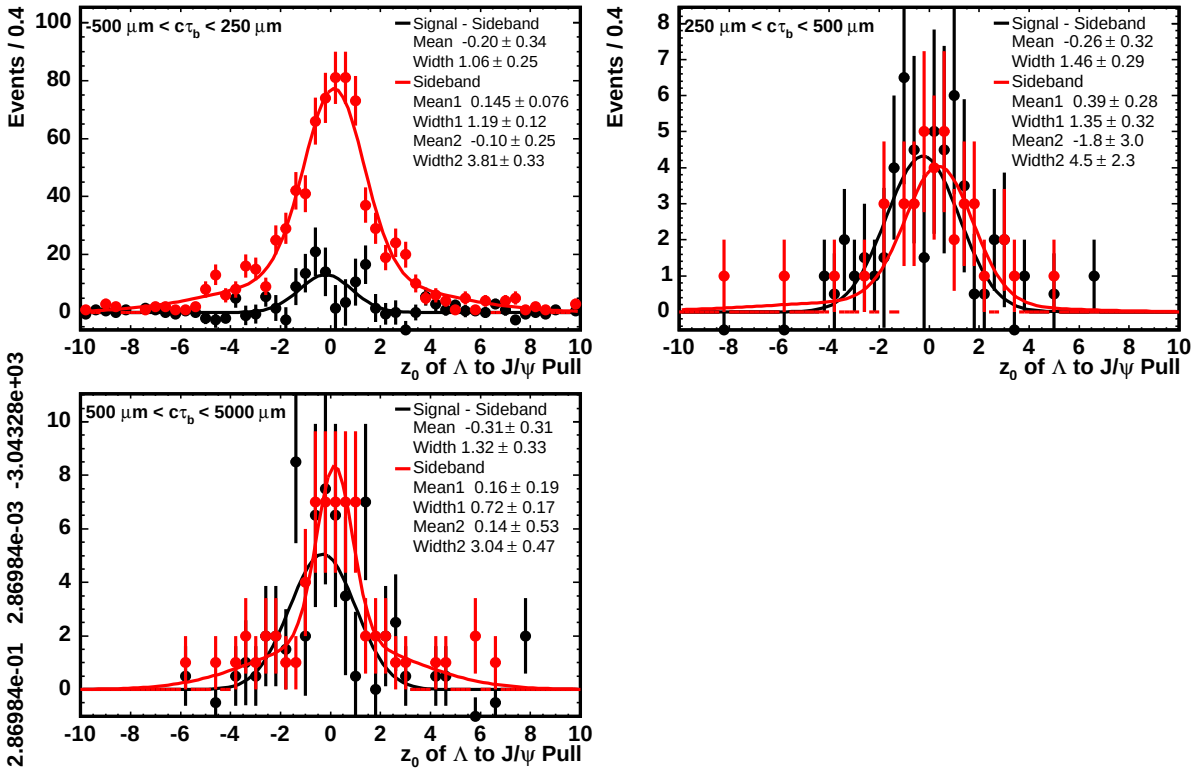


Figure 81: Distribution of z_0 of the Λ pull in three bins of b-hadron $c\tau$ for the $B^0 \to J/\psi K^0_s$ mode in the data.

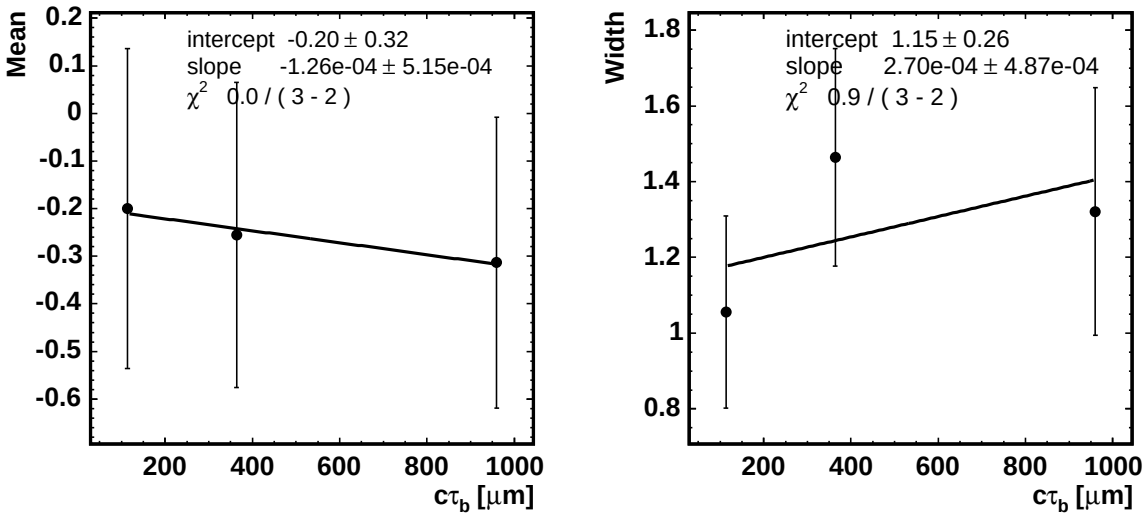


Figure 82: Results of gaussian fits to the z_0 of the Λ pull in the three bins shown in Figure 81

systematic error. Our final calculation of the bias uses a toy Monte Carlo with realistic assumptions about the other dimensions. The vertex probability cut is 10^-4 which for five degrees of freedom corresponds roughly to a $\chi^2 < 25$.

The efficiency of the a cut on a guassian is

$$\epsilon(c\tau) = \int_{-A}^{A} \frac{1}{\sqrt{2\pi}\sigma(c\tau)} e^{-\frac{(x-\mu(c\tau))^2}{2\sigma(c\tau)^2}} dx \tag{20}$$

where μ and σ are the fit results from the previous section and are functions of the proper decay length. This can be written as

$$\epsilon = \frac{1}{2}\big[erf(\frac{A+\mu}{\sqrt{2}\sigma}) + erf(\frac{A-\mu}{\sqrt{2}\sigma})\big]. \tag{21}$$

The equation combined with the fits from the previous section give us efficiency curves with errors.

A calcuation of the bias due to the efficiency can be made using a no background approximation (keep in mind the background shape uncertainty is dealt with as a fit model systematic). Without background, one can calculate the measured cτ of a particle as

$$\mathrm{c}\tau_M = \frac{\int_0^\infty x\epsilon(x)\frac{1}{\mathrm{c}\tau_\mathrm{T}} e^{\frac{-x}{\mathrm{c}\tau_\mathrm{T}}} dx}{\int_0^\infty \epsilon(x)\frac{1}{\mathrm{c}\tau_\mathrm{T}} e^{\frac{-x}{\mathrm{c}\tau_\mathrm{T}}} dx} \tag{22}$$

where $\mathrm{c}\tau_M$ is the measured cτ and $\mathrm{c}\tau_T$ is the true cτ. Using Equations 21 and 22, we can calculate bias with errors from the fits in Section 6.3.3. Errors are propagated using by repeating the calculation with slopes and intercepts of the pull distribution variables randomly drawn from appropriate gaussians. Table 10 summarizes the information from the vertex pull fits and the derived cτ bias with error. The result is present for A =3,4, and 5, showing the progressive decrease sensitivity to the $\Delta\mathrm{L_{xy}}$ and z_0 pull distributions as the cut is loosened.

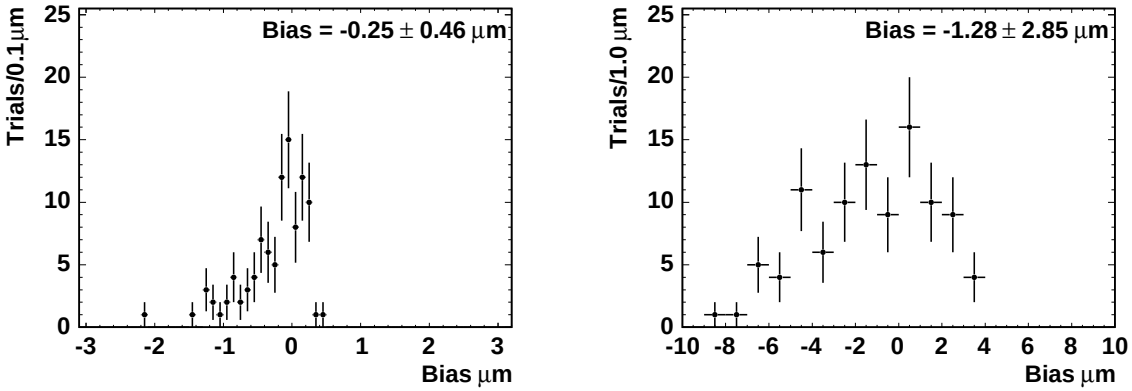
In order to include the effect of cutting on the combined 5-dimensional χ^2 probability instead of the individual variables, we use toy Monte Carlo. The proceedure used is:

1. Choose slopes and intercepts of the $\Delta\mathrm{L_{xy}}$ and z_0 pull distribution variables according to the fits from the previous section including errors.

2. Calculate the bias using Monte Carlo integration:

   (a) Choose a cτ from a $460\mu m$ exponential (the final uncertainty is too small for the difference in Λ_b^0 and B^0 lifetimes to be important).

   (b) Choose $\Delta\mathrm{L_{xy}}$ and z_0 χ^2 contributions according to the gaussians determined by the slopes and intercepts from Step 1.



(a) The applied Prob(χ^2) $< 10^{-4}$ cut

(b) A tighter Prob(χ^2) $< 10^{-2}$ cut

Figure 83: Distribution of pointing biases due to variations of the ΔL_{xy} and z_0 pull distribution within the ranges allowed the pull distribution fits.

(c) Choose three gaussian distributed random numbers for the other three variables in the vertex probablity (we use zero mean guassians with widths of 1.2 in order to include approximate scale factors for the tracking errors).

(d) Sum the χ^2 contributions from the previous two steps to get a total vertex probability

(e) Calculate the mean of the events passing the probability $< 10^{-4}$ cut, and from that the resulting bias.

This is repeated to get a bias distribution from which the pointing constraint systematic is determined. The results of this procedure are show in Figure 83 for both the $< 10^{-4}$ cut and a tighter $< 10^{-2}$ cut. Comparison of the our cut and the tighter cut show that the bias is small largely because our cut is very loose, not because the pull distributions are very well constrained. Because of the negative tail, we conservatively choose a systematic of $1\mu m$.

6.4 COT to SVX Matching Study

The broad guassians in Figures 69 to 72 show that the COT resolution on V^0 pointing is $\approx 2000\mu$m which is not sufficient to resolve the b-hadron liftetime. The SVX (narrow gaussians) however does have sufficient resolution, so it is critical that the matching of COT V^0s to SVX V^0s not be biased as a function of the b-hadron lifetime. In Figure 84 we show the efficiency for the B^0 candidates found using COT-only tracks to be found using defTracks. Also shown in Figure 84 is the rate at which the candidates found in defTracks have SVX hits on the higher momentum of the two pions. There are no statistically significant slopes.

	ΔL_{xy}	z_0
Mean	-0.046 ± 0.097	0.08 ± 0.11
Mean Slope	$(-0.75 \pm 1.34) \times 10^{-4}/\mu m$	$(-0.10 \pm 1.64) \times 10^{-4}/\mu m$
Width	1.065 ± 0.076	1.24 ± 0.10
Width Slope	$(7.77 \pm 9.99) \times 10^{-5}/\mu m$	$(-0.92 \pm 1.35) \times 10^{-4}/\mu m$
Bias ($A = 3$)	$-1.85 \pm 1.99\ \mu m$	$0.47 \pm 2.27\ \mu m$
Bias ($A = 4$)	$-0.36 \pm 0.51\ \mu m$	$0.01 \pm 0.49\ \mu m$
Bias ($A = 5$)	$-0.05 \pm 0.09\ \mu m$	$-0.02 \pm 0.09\ \mu m$
Bias (Prob(χ^2) $< 10^{-2}$)	$-1.28 \pm 2.85\ \mu m$	
Bias (Prob(χ^2) $< 10^{-4}$)	$-0.25 \pm 0.46\ \mu m$	

Table 10: Systematic uncertianties due the the V^0 pointing.

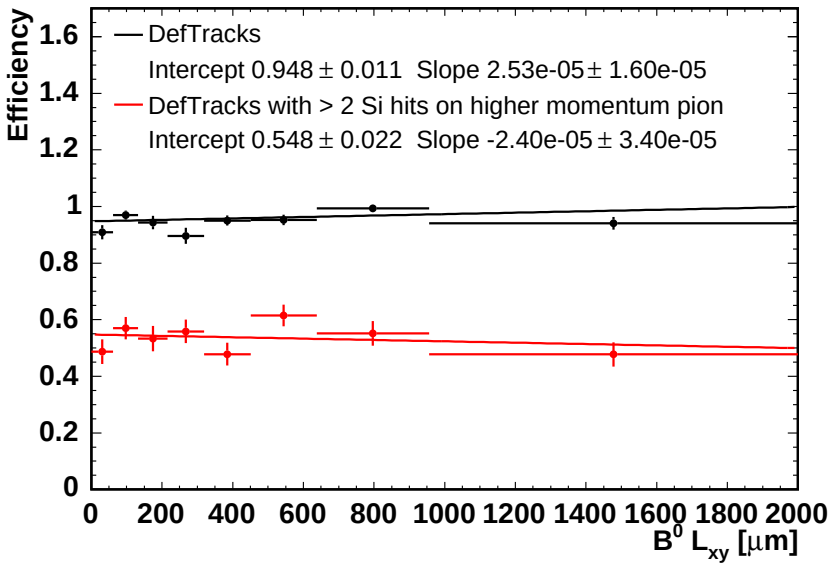


Figure 84: Probablities for a COT-only candidate to be found in the DefTracks sample (black line) and for those found to have SVX hits (red line).

6.5 Fit Model

The fit model systematics are in general derived from variations of the fit model. The various potential systematics are addresses by seperate variations. **Because the analysis is still blinded the Λ_b^0 systematics are on the blinded result, and will be scaled when the unblinding is done.**

6.5.1 $c\tau$ Resolution

The model of the $c\tau$ resolution in the nominal fit is based on the vertex fitter and ultimately tracking software estimate of the resolution on the vertex position. The fit includes an overall scale factor for this estimate. This means that the $c\tau$ error pull distribution is now modeled by a gaussian of arbitrary width instead of a unit gaussian. In order to add an additional degree of freedom, we add a second scale factor so that the pull distribution is modeled by the sum of two gaussians. This adds two degrees of freedom to the fit, the width and relative normalization of the additional gaussian. The resulting shift is -3.0 μm for the B^0 $c\tau$ and -2.1 μm for the Λ_b^0. It should be noted that this improves the χ^2 probability for the proper decay length distribution, particularly for the backward tail. Based on these number we assign systematic errors of 3.0 μm and 2.1 μm respectively. It should be noted that this does not account for the distribution of $c\tau$ resolutions which is discussed in Section 6.5.6.

6.5.2 Mass Signal

Similar to the $c\tau$ model, we use the vertex fit mass error scaled to describe the signal mass distribution. We also add a second gaussian to this scale factor to assess a systematic error. The resulting shifts are 0.8 μm for the B^0 $c\tau$ and -1.7 μm for the Λ_b^0. We also consider using a standard gaussian for the mass distribution, ignoring the vertex fit errors. The resulting shifts are -1.8 μm for the B^0 $c\tau$ and -1.9 μm for Λ_b^0. Based on these number we assign systematic errors of 1.8 μm and 1.9 μm respectively.

6.5.3 Mass Background

The nominal shape of the mass background is described line with varying slope. To assess a systematic associated with this assumption we fixed the shape to a constant. The resulting shift for both modes is less than 0.1 μm. Based on these number we assign a systematic error of 0.1 μm

6.5.4 PDL Background

One of the most ad-hoc assumptions made in the nominal fit is the description of the lifetime distribution of the long-lived background. This distribution arises from a complicated mixed of the b-hadron lifetimes, the momentum spectrum of real and fake V^0

candidates, and the resolutions of the vertexing information used to suppress the background. Because of this uncertainty we consider a wide variety of background models. All the models consist of a sum of exponentials and one delta function a $c\tau = 0$ for the prompt background. The prompt background is always convoluted with the scaled gaussian resolution from the vertex fit. Several variations on whether the exponentials are convoluted with the scale factor are considered. The resulting shifts are show in Table 6.5.4. The notation is that E_- is a exponential for negative $c\tau$, E_+, E_{++}, and E_{+++} are exponentials for positive $c\tau$, and $\otimes G$ indicates which terms are convoluted with a gaussian. Based on these number we assign a systematic error of 0.6 μm for B^0 and 5.5 μm for Λ_b^0.

6.5.5 Fits Using Sideband PDL Background Shapes

One model assumption not varied in the previous section is that the shape of the background is independent of the candidate mass. In order to assess the uncertainty related to this assumption, we fit using only one side of the sideband at a time. Specifically for B^0 mode in the region $5.2375 < m_{\rm cand} < 5.390$ we see a shift of 0.2 μm and for the region $5.170 < m_{\rm cand} < 5.3225$ we see a shift of -1.6 μm. For the Λ_b^0 in the region $5.521 < m_{\rm cand} < 5.651$GeV we see a shift of 14.7 μm and for the region $5.591 < m_{\rm cand} < 5.721$GeV we see a shift of 6.4 μm. Because these shifts contain a large statisical uncertainty, we use the average shifts as an estimates of the systematics. This gives 0.9 μm for the B^0 and 10.6 μm for the Λ_b^0.

6.5.6 PDL Error Model

We observe from our fits to the data that our modeling of the PDL error distribution is not actually a good fit to the data. We estimate the systematic uncertainty on our PDL error model by generating toy data (10 times the size of the actual data) for $B^0 \to J/\psi K_s^0$ using the *true* PDL error distributions for signal and background from the data and then fitting with our parameterized model of the PDL error. We observe a shift of -3.6μm and quote a systematic uncertainty of $\pm 3.6\mu$m due to our PDL error modeling.

6.5.7 Mass Error "Model"

Recall that we made the observation that we could factor out the mass error PDF from our overall PDF because the signal and background distributions were statistically indistinguishable. We assign a systematic on this assumption in the exactly same way as done in Section 6.5.6 using the toy Monte Carlo to generate with the actual signal and background distributions from the data but fit the data assumming they are the same. We observe a shift of -0.6μm and quote a systematic uncertainty of $\pm 0.6\mu$m due to our assumption in mass error modeling.

Variation	$B^0 \to J/\psi K^0_s$ Fitted $c\tau$ (μm)	Shift (μm)
$c\tau$ Res 2 Scales	447.6 ± 14.8	-3.0
Mass Res 2 Scales	451.4 ± 14.6	0.8
Mass Res Fixed Gaus	448.8 ± 14.7	-1.8
Constant Bkg Mass	450.6 ± 14.7	0.0
$(E_- + E_+ + \delta(0)) \otimes G$	450.8 ± 14.6	0.2
$(E_- + E_+ + E_{++} + E_{+++} + \delta(0)) \otimes G$	450.6 ± 14.7	0.0
$E_- + E_+ + (\delta(0)) \otimes G$	450.5 ± 14.6	-0.1
$E_- + E_+ + E_{++} + \delta(0) \otimes G$	450.0 ± 14.7	-0.6
$E_- + E_+ + E_{++} + E_{+++} + \delta(0) \otimes G$	450.0 ± 14.7	-0.6
Low Sideband Only	450.8 ± 15.2	0.2
High Sideband Only	449.0 ± 15.0	-1.6

Variation	$\Lambda^0_b \to J/\psi \Lambda^0$ Fitted $c\tau$ (μm)	Shift (μm)
$c\tau$ Res 2 Scales	642.1 ± 59.7	-2.1
Mass Res 2 Scales	642.6 ± 59.9	-1.7
Mass Res Fixed Gaus	642.3 ± 62.2	-1.9
Constant Bkg Mass	644.2 ± 59.8	0.0
$(E_- + E_+ + \delta(0)) \otimes G$	649.8 ± 58.2	5.5
$(E_- + E_+ + E_{++} + E_{+++} + \delta(0)) \otimes G$	644.3 ± 54.9	0.0
$E_- + E_+ + (\delta(0)) \otimes G$	648.3 ± 58.1	4.0
$E_- + E_+ + E_{++} + \delta(0) \otimes G$	642.0 ± 59.6	-2.2
$E_- + E_+ + E_{++} + E_{+++} + \delta(0) \otimes G$	642.0 ± 59.7	-2.2
Low Sideband Only	658.9 ± 65.6	14.7
High Sideband Only	650.6 ± 64.6	6.4

Table 11: Fit model variations used to assess the data modeling systematics. **Note the value for Λ^0_b is the scaled blinded value**

6.6 Primary Vertex Determination

In this analysis, the coordinates of the primary vertex are determined by evaluating the time-dependent beamline at $z = \bar{z}_0(\mu^+, \mu^-)$, where $\bar{z}_0(\mu^+, \mu^-)$ is the average z_0 of the two muon tracks from the J/ψ. We have also fit the B^0 lifetime using an alternative method of primary vertex determination. We evaluate the beamline position at $z = z_v$, where z_v is the z coordinate of the primary vertex closest to the z coordinate of the J/ψ vertex. In the latter case the B^0 lifetime is 0.2 μm smaller than when using standard method.

6.7 Alignment systematics

Several factors can lead to systematic uncertainties on the lifetime measurement. There could be a shift or pull in the fitter we use to extract the lifetime. The background parameterization may be a bad match to the background distribution in data. Uncertainties in the alignment of the silicon detector can lead to scale or shape distortions.

Misalignments of the silicon system can be either random displacements of individual detectors (the so called internal misalignment) or collective motions of the detector as a whole (global misaligment).

6.7.1 Internal alignment systematics

The type of internal displacement that effects lifetime measurement most strongly is a radial dilation or contraction of the system. Such an effect can be due to a bowing of silicon detectors, which is known to occur in CDF. The size of the effect is conservatively estimated at 50 microns.

One can derive a simple yet robust estimation of the effect of a 50 micron bowing on lifetime measurement by noting that the fractional change to the position of the first silicon layer, at 2.44 cm, of 0.2%; if such a distortion is present it will shift the value of a measured lifetime by the same fraction, which, for a particle with a $c\tau$ of 500 μm is 1 μm.

The true internal alignment systematic error was obtained in CDF-note 6387 by a procedure in which different misalignments (+/-50 μm in raidus and 50 μm bow outward/inward on all ladders in the silicon system) were put into the alignment files, and the full simulation of Monte Carlo was done with the new alignment files. The maximum deviation obtained was 2 μm and was the assigned systematic uncertainty coming from internal alignment.

6.7.2 Global alignment systematics

A global misalignment can exist either in the position of the center of the SVX with respect to the COT (translation), or in a relative rotation.

Since the beam position is most sensitive to a translation, we estimate the maximum global misaligment (offset) by plotting the differences between COT and SVX beam coordinates for a given run. Figure 85 shows x(Cot Beam)-x(Svx Beam) and y(Cot Beam)-y(Svx Beam) as a function of run number. We estimate a maximum misalignment in x or y of ≈ 30 μm.

A rotation of the SVX with respect to the COT can generate a "false impact parameter", i.e. for measurement layers i, $\phi_i \rightarrow \phi_i + \delta/r_i \Rightarrow d_0 \rightarrow d_0 + \delta$ [14]. We estimate this false impact parameter from distribtions of the impact parameter (with respect to the SVX beam) of COT muon tracks from J/ψ. Figures 86(a), 86(b), and 86(c) show these distributions for all muon tracks, positive, and negative muon tracks, with the cores of the distribution fit to a double gaussian. Figure 6.7.2 shows the same distributions, but with the means and RMS's of the histograms displayed. The means of the distributions are shown in Table 12. The difference in results for positive tracks and negative tracks is possibly due to unacounted-for energy loss (though the energy loss correction is done). We estimate the false impact param bias to be approximately 20 μm. With the COT inner radius of 40 cm, we arrive at a rotation in ϕ of 50 μrad.

Track Charge	mean (μm) fit gaussian	mean (μm) arithmetic
All	17 $\pm$ 1	16 $\pm$ 4
positive	32 $\pm$ 3	38 $\pm$ 6
negative	4 $\pm$ 1	-7 $\pm$ 6

Table 12: Means of impact parameter distributions for COT-only muon tracks from J/ψ.

The global alignment systematics are obtained by a procedure in which the different misalignments discussed above (+1mm in z, +100 μm in x, +30 μm in y and rotation of the whole silicon system with respect to the COT) were put into the alignment files, and the full simulation of Monte Carlo was done with the new alignment files.

- 160050 1 GOOD: best (version used for reconstruction).
- 160050 1 TEST: moved spaceframe +100 μ in x direction w.r.t. 160050 1 GOOD
- 160050 2 TEST: moved spaceframe +1 mm in z direction w.r.t. 160050 1 GOOD
- 160050 8 TEST: rotated spaceframe around z by 50 μrad w.r.t. 160050 1 GOOD
- 160050 9 TEST: rotated spaceframe around z by -50 μrad w.r.t. 160050 1 GOOD

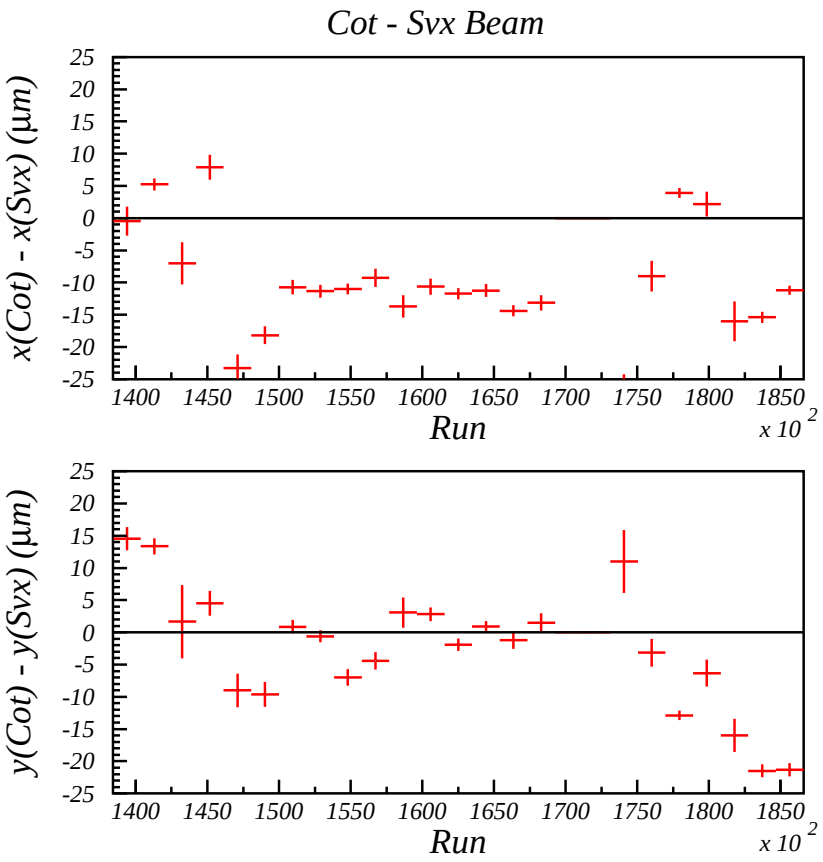


Figure 85: Difference between COT and SVX beam coordinates (cm), for runs satisfying our good run selection

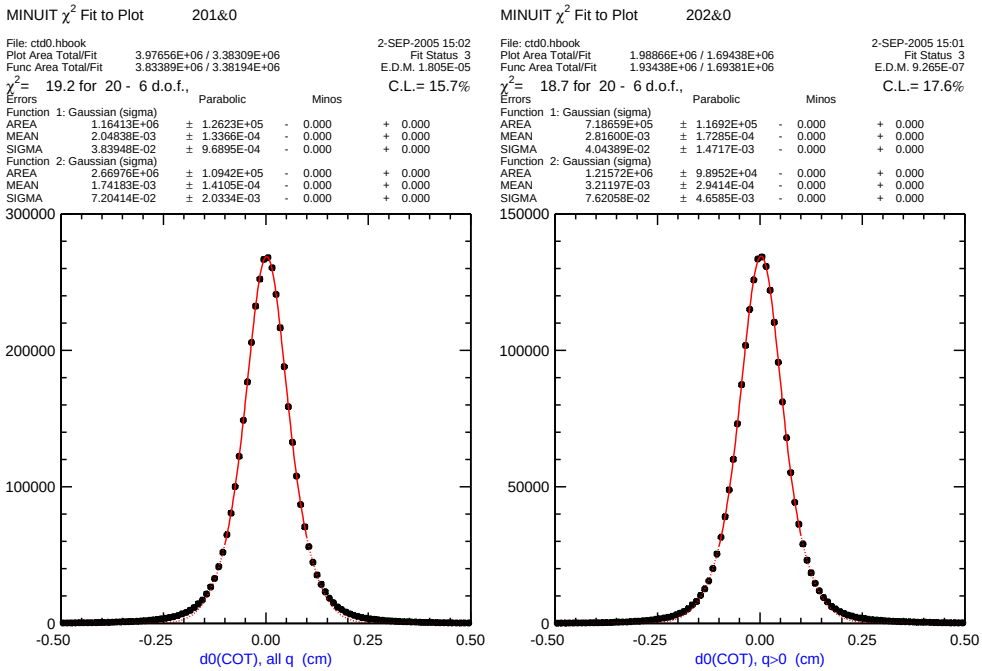
- 160050 12 TEST: moved spaceframe -30 μ in y direction w.r.t. 160050 1 GOOD
- 160050 13 TEST: moved spaceframe +30 μ in y dir. w.r.t. 160050 1 GOOD

The different realistic Monte Carlo samples were obtained using Bgen in 5.3.4, produced with passname 17 and a dimuon trigger requirement of two muons with $p_T > 1.4$ GeV/c. The results are listed in table 13. The ratio of $J/\psi\ \Lambda$ to $J/\psi\ K^0_s$ lifetime is also added in the table.

All fits use the nominal procedure (simultaneous fit for Mass and Lifetime, J/ψ vertex). The first row shows the fitted $c\tau$ (in μm) obtained with the best alignment (the one used for our nominal results). Rows 2 to 7 show the values of the fitted lifetime, when the relevant quantities were obtained with each alignment version. All lifetime values are also expressed in μm.

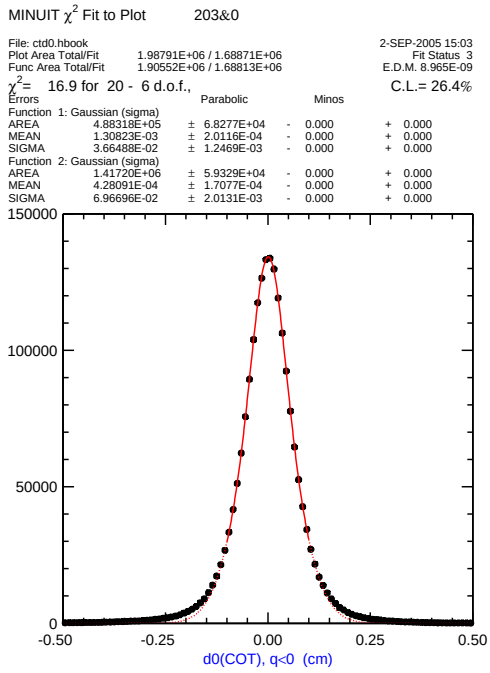
All the channels have the behavior expected from the type of distorsion introduced on each alignment version (alignments shifted up/down yield lower/higher lifetimes).

The variation we observe in this test may overlaps with the one coming from the background model. The observed variation would then be an upper bound for this error. This upper bound is already small compared to other sources, so we assign $^{+5}_{-2},^{+7}_{-1}\,\mu$m as the systematic uncertainty coming from the global alignment for the $\Lambda_b \to J/\psi\,\Lambda$ and $B^0 \to J/\psi\,K^0_s$ modes respectively.



(a) d_0 for COT tracks from J/ψ muons, μ^+ and μ^-

(b) d_0 for COT tracks from J/ψ muons, μ^+



(c) d_0 for COT tracks from J/ψ muons, μ^-

Figure 86: d_0 distribution for COT tracks from J/ψ muons

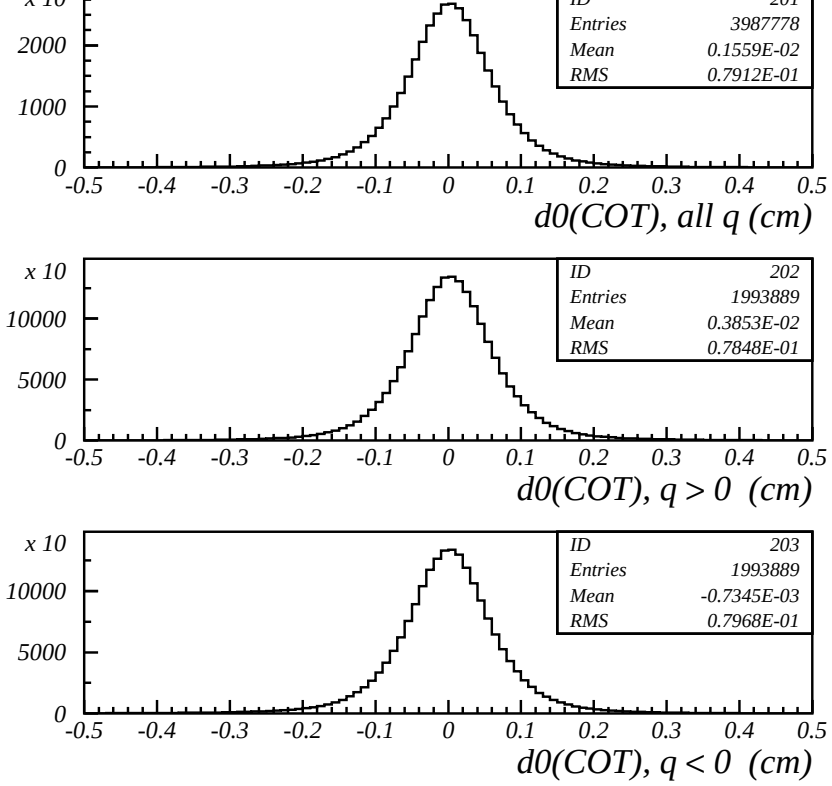


Figure 87: Distributions of d_0 for COT muon tracks from J/ψ, showing mean and RMS

6.8 Other b-Meson Modes

In the present section we determine the "size of the CDF micron" as well as any possible bias in the L_{xy} measurement of the J/ψ vertex. We do so by fitting the $B^0 \to J/\psi\,\mathrm{K}^{*0}$ and $B^+ \to J/\psi\,K^+$ modes. The modes were reconstructed analogously to the V^0 modes in that the PDL was constructed from the J/ψ vertex information, not from the B vertex. The selection used was the same as in the CDF blessed result [16]. Our results are consistent to within $\approx 3\mu$m, which we consider reasonable given the difference in the reconstruction and the systematic error on the $c\tau$ resolution function (Section 6.5.1).

In addition to being a high-level check of analysis aside from V^0's, we can also use the difference between these result and the PDG values to assign an uncertainty to the CDF length scale. That is how do we know a μm in the CDF software is really a μm. This works for the Λ_b^0, because the statistical error is much larger than the errors in the PDG and on the non-V^0 modes. The resulting length scale uncertainties are $13 \pm 9\mu$m and $4 \pm 11\mu$m. Combining these we get $10 \pm 7\mu$ from which we assess a 10μm systematic error. This error covers biases from both the CDF length scale and the treatment of the J/ψ half of the decay.

6.9 Alternative Lifetime Fitting Techniques

6.9.1 Mass Fit in Lifetime Bins

We have fit the B^0 lifetime using a method which is insensitive to our ability to model the shape of the background ct distribution. The only shape fitting is that to the $\mu\mu\pi\pi$

Alignment	$J/\psi\ \Lambda$	$J/\psi\ K_s^0$	ratio
160050 1 GOOD (best)	370.6 ± 3.6	465.3 ± 2.1	0.796 ± 0.011
± 160050 1 TEST (+100 μm x-shift)	365.5 ± 4.0	461.6 ± 1.8	0.792 ± 0.012
160050 2 TEST (+1 mm z-shift)	365.3 ± 4.1	463.2 ± 1.8	0.789 ± 0.012
160050 8 TEST (+50μrad z-rot)	368.4 ± 4.3	462.5 ± 2.5	0.797 ± 0.014
160050 9 TEST (-50μrad z-rot)	366.7 ± 4.3	463.5 ± 2.5	0.791 ± 0.014
160050 12 TEST (-30 μm y-shift)	368.9 ± 4.2	456.8 ± 2.4	0.807 ± 0.013
160050 13 TEST (+30 μm y-shift)	371.8 ± 4.2	463.0 ± 2.4	0.803 ± 0.013

Table 13: Fitted $c\tau$ values and ratios of all modes for various alignment versions. The first row shows the values obtained with the alignment used for the nominal fit. All numbers are in microns. The input generated value was 370 μ for the Λ_b and 464 μ for the B^0.

invariant mass distribution, for which we use a gaussian and a first-order polynomial.

We divide the data into 5 bins of proper decay length. For each bin, we fit the invariant mass distribution to extract the number of signal events. If the bin boundaries are ct_i and ct_{i+1}, the number of events decaying between ct_i and ct_{i+1} is

$$N_{tot} f_i = N_{tot}(e^{-ct_i/c\tau} - e^{-ct_{i+1}/c\tau}) \tag{23}$$

We then minimize χ^2 to fit N_{tot} and $c\tau$:

$$\chi^2 = \sum_i \left(\frac{N_{tot} f_i - N_i}{\sigma_{N_i}} \right)^2 \tag{24}$$

The binning and number of signal events per bin are shown in Table 14. The individual invariant mass distribution fits used to obtain these numbers are shown in Figure 88. The contents of the table are displayed graphically in Figure 89 along with the the projected fit. For the lifetime, we obtain a value which is 3.6 μm smaller than that from the standard maximum likelihood fit.

Since this method uses no knowledge or supposition about the shape of the background ct distribution, and since the results of this method agree with those of the maximum likelihood fitter, this indicates that we do not have a problem modelling the background in the case of the maximum likelihood fitting.

6.9.2 Sideband-subracted Lifetime Fit

Another method to determine the lifetime without depending on the background fit model is a background substracted fit to just the lifetime distribtuion. After a sideband subtraction, we fit the $B^0 \to J/\psi\, K_s^0$ data to the function

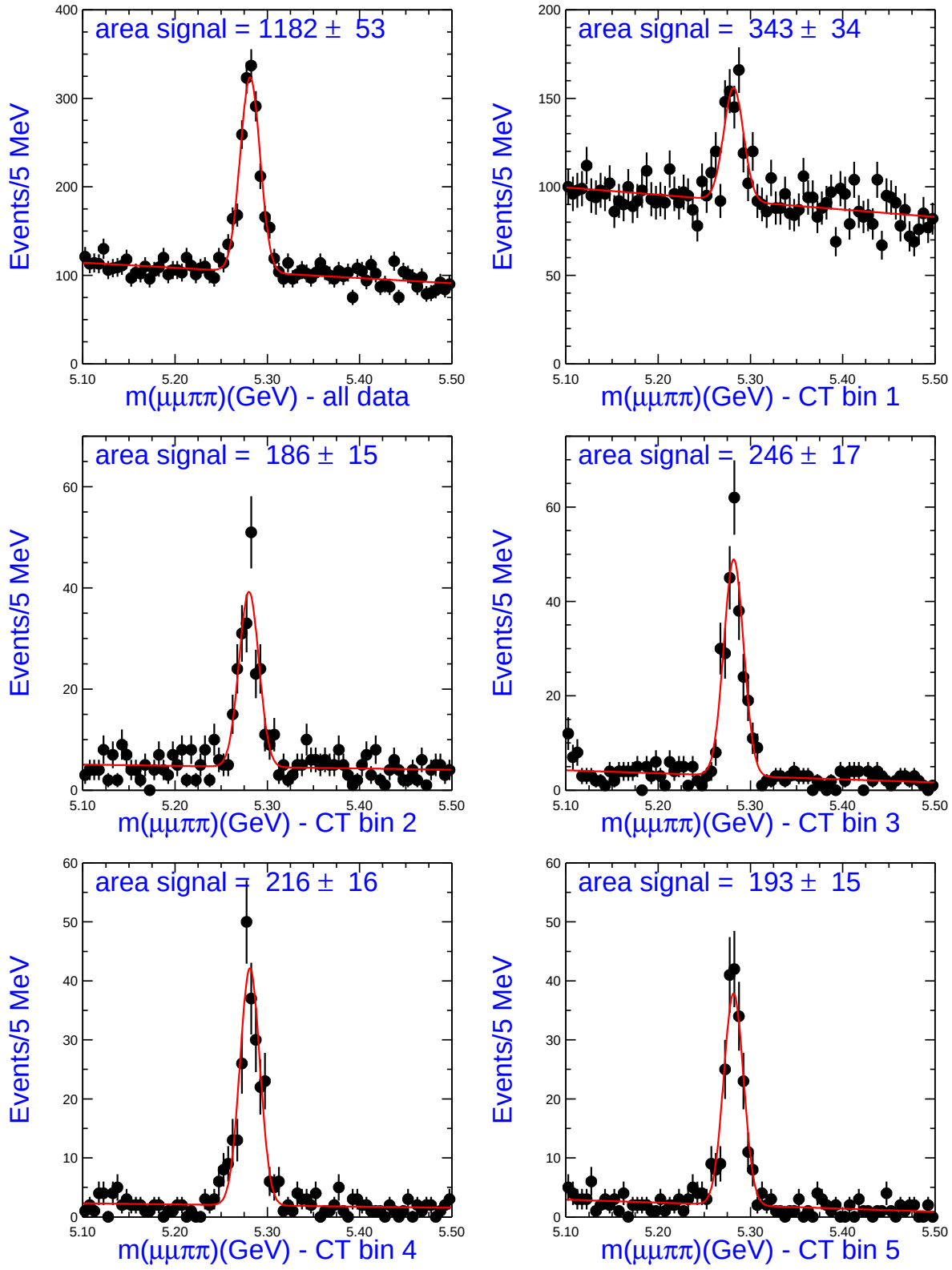


Figure 88: Mass distribution fits for determination of number of events per *ct* bin, used in the simple lifetime fitting method.

ct plot bin	ct range (cm)	N events
1	$ct \leq 0.015$	343 ± 34
2	$0.015 < ct \leq 0.027$	186 ± 15
3	$0.027 < ct \leq 0.048$	246 ± 17
4	$0.048 < ct \leq 0.08$	216 ± 16
5	$0.08 < ct \leq 0.7$	193 ± 15

Table 14: Number of $B^0 \to J/\psi\, K_s^0$ events for given bin in proper decay length.

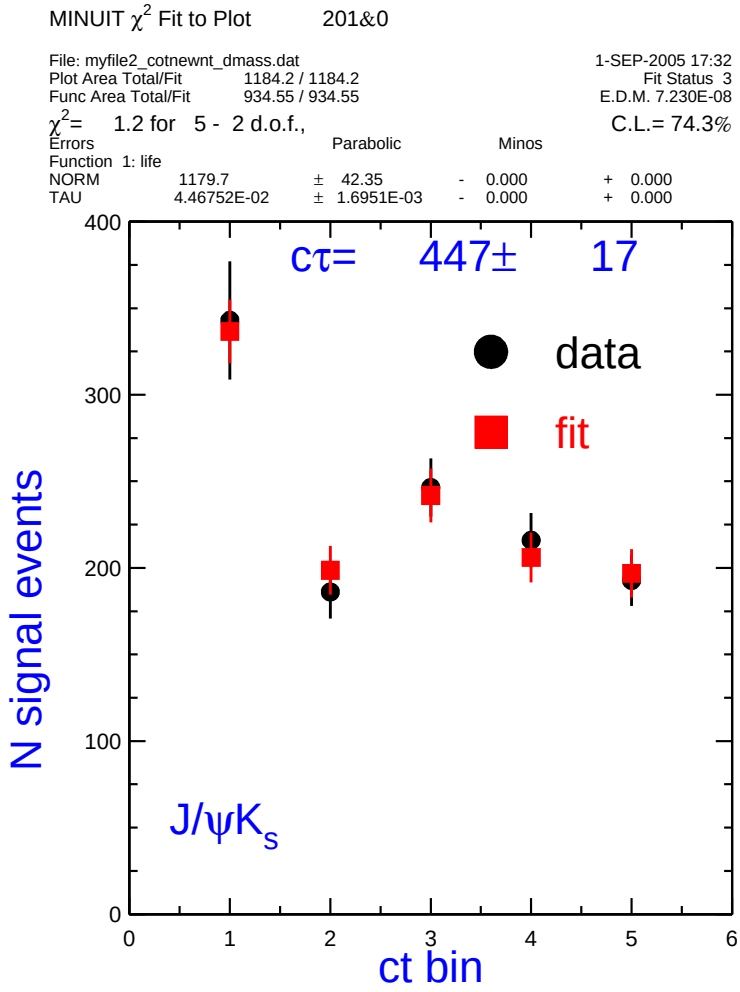


Figure 89: Determination of B^0 lifetime in $B^0 \to J/\psi\, K_s^0$ with alternative fitting method

$$\mathcal{P}^{\lambda}_{sig}(\lambda|c\tau,\sigma_{C\tau}) = \frac{1}{2c\tau} exp\left[\frac{\sigma^2_{C\tau}}{2(c\tau)^2} - \frac{\lambda}{c\tau}\right] erfc(\frac{\sigma_{C\tau}}{\sqrt{2}c\tau} - \frac{\lambda}{\sqrt{2}\sigma_{C\tau}})$$

where λ is the measured proper decay length. The $\sigma_{C\tau}$ determined here is an average $\sigma_{C\tau}$. The event by event $\sigma^{\lambda}_{\rm i}$ are not used, so this method is also not sensitive to the Punzi effect bias. The results are shown in comparision to the nominal results in Figure 90.

6.9.3 Moment Lifetime Extraction

After a sideband subtraction the lifetime can be determined by simply taking the first moment, $c\tau_M$, of the measured proper decay length, λ, distribution:

$$\begin{aligned} c\tau_M &= \sum_i w_i(m_i)\lambda_{\rm i} \\ &= \int \lambda\ E(\lambda - x)\ G(x)\ dx\ d\lambda \\ &= \int \lambda\ 1/c\tau_T\ exp(-\lambda/c\tau_T) + \int xG(x)dx \\ &= c\tau_T, \end{aligned}$$

where $w_i(m_i)$ is event-by-event weight that performs the sideband subtraction. The second line in the above equation expresses this moment as the moment of the convolution of the lifetime exponential $E(\lambda)$ with the resolution gaussian, $G(x)$. The third line is dervied using the fact that a moment of a convolution is the sum of moments of the two convolved functions. Finally the last line is arrived at assuming that the resolution is unbiased (zero moment). This shows that the starting moment is an estimator of the true $c\tau$, $c\tau_T$. The results of this method are also shown in comparision to the nominal results in Figure 90.

6.10 B^0 Lifetime Dependence Sanity Checks

6.10.1 Run Range

Figure 91 shows a histogram of runs corresponding to events passing all of our cuts, including our good run selection. To check for any run dependence of the B^0 lifetime (and thus possibly the $\Lambda^0_{\rm b}$ lifetime), we divide the data into the five run ranges shown in Table 15, and fit the B^0 lifetime. The bin boundaries are chosen to correspond to shutdowns/accesses and the state of the COT (i.e. compromised running or recovered), as is also shown in Figure 91.

Figure 92(a) shows the fitted B^0 lifetime in each run range. In addition, we do the same study using COT-only tracks for the K^0_s, in which case we obtain similar results. Dependence of lifetime on run range for the COT-only case is shown in Figure 92(b).

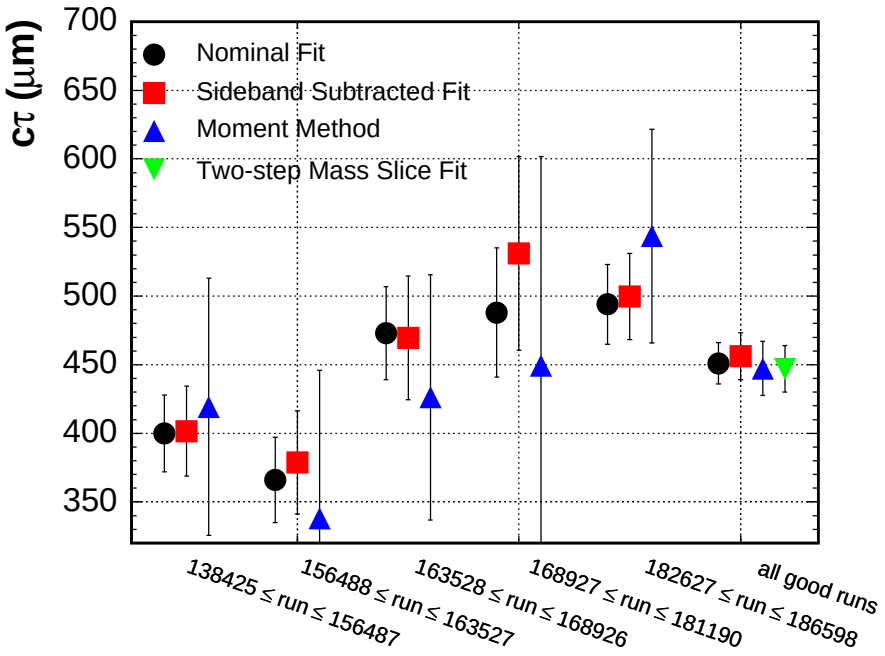


Figure 90: Summary of the alternative fitting methods described in Section 6.9.

Bin	Run Range	Integrated Luminosity (pb^{-1})	B^0 Yield
1	$138425 \leq$ run ≤ 156487	76.5 ± 4.5	227 ± 28
2	$156488 \leq$ run ≤ 163527	57.8 ± 3.4	191 ± 31
3	$163528 \leq$ run ≤ 168926	73.5 ± 4.3	268 ± 33
4	$168927 \leq$ run ≤ 181190	51.2 ± 3.0	157 ± 49
5	$182627 \leq$ run ≤ 186598	111.8 ± 6.6	342 ± 28

Table 15: Run ranges, and their corresponding luminosity and B^0 yield, in which we study the data

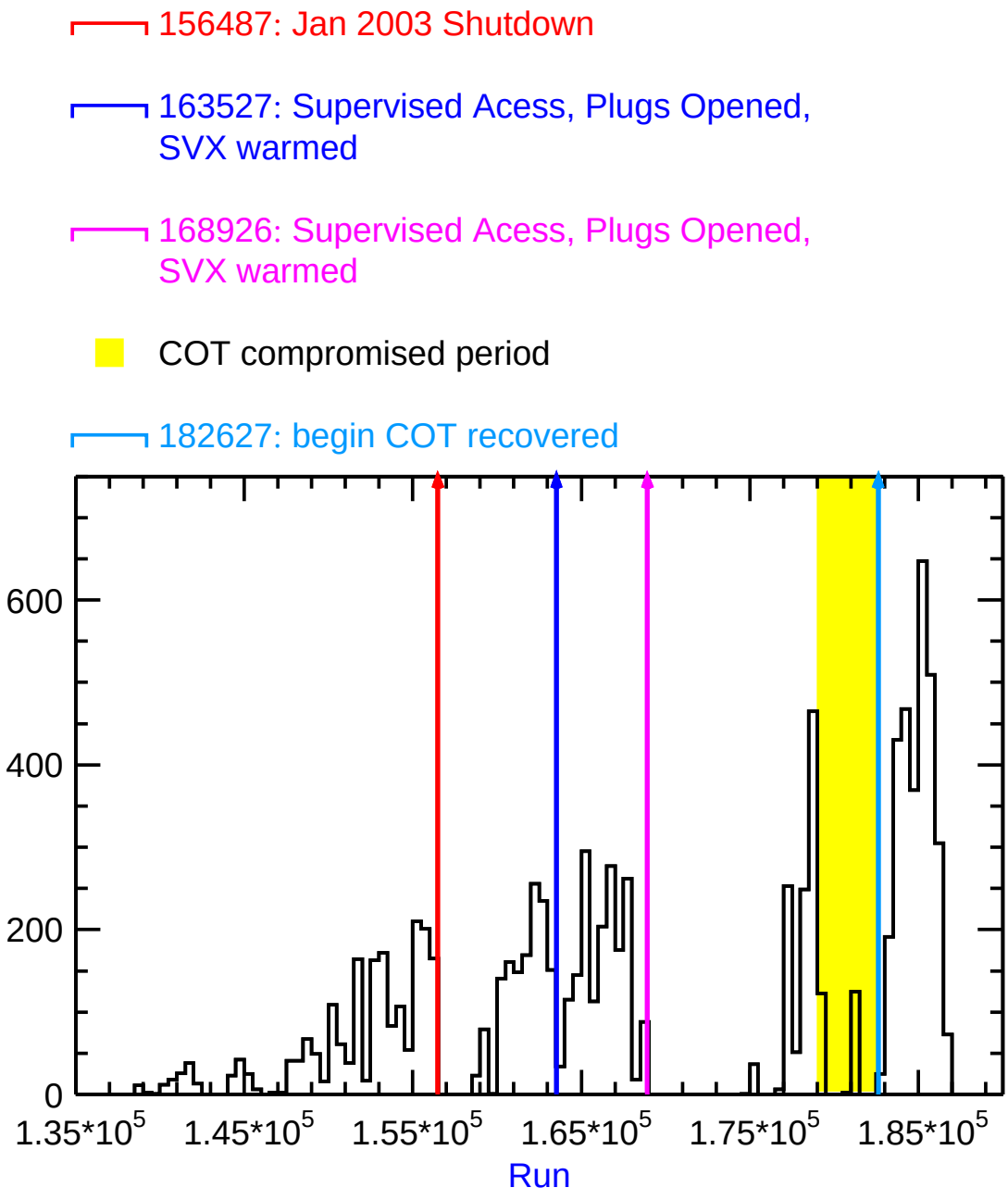
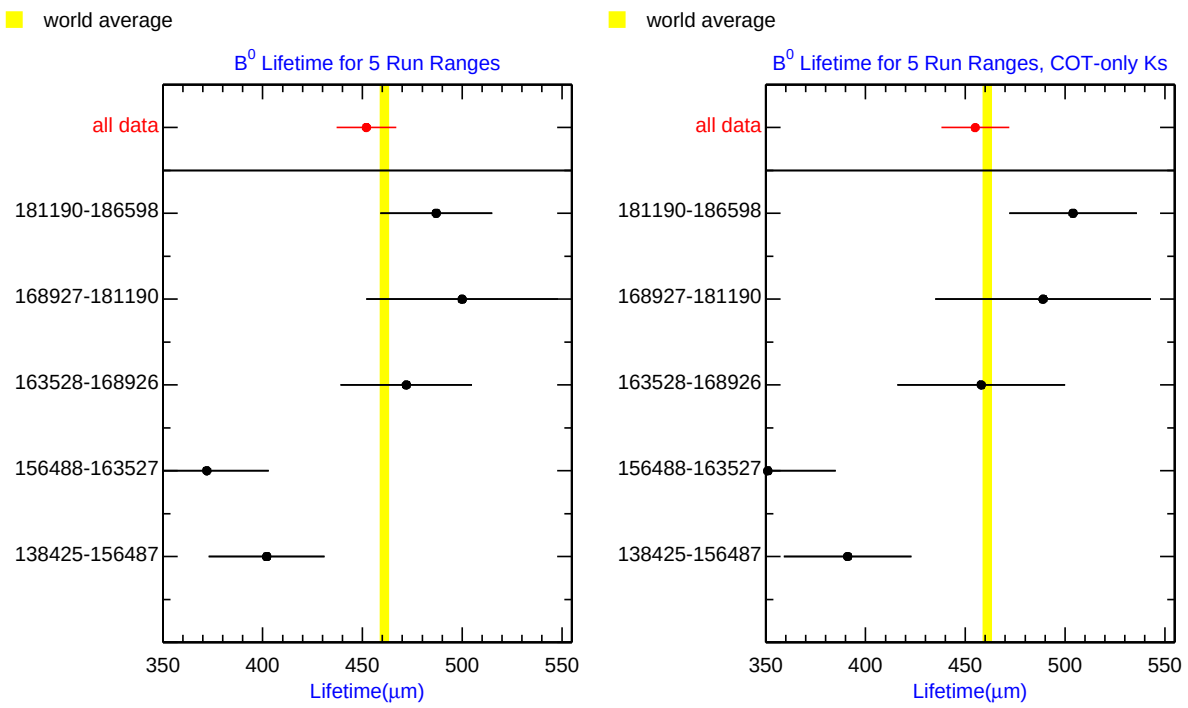


Figure 91: Histogram of runs corresponding to events which pass all of our cuts.



(a) B^0 lifetime for various run ranges

(b) using COT-only tracks for the K^0_s

Figure 92: Run range dependence for the $B^0 \to J/\psi\, K^0_s$ mode using the default reconstruction and using COT-only K^0_ss.

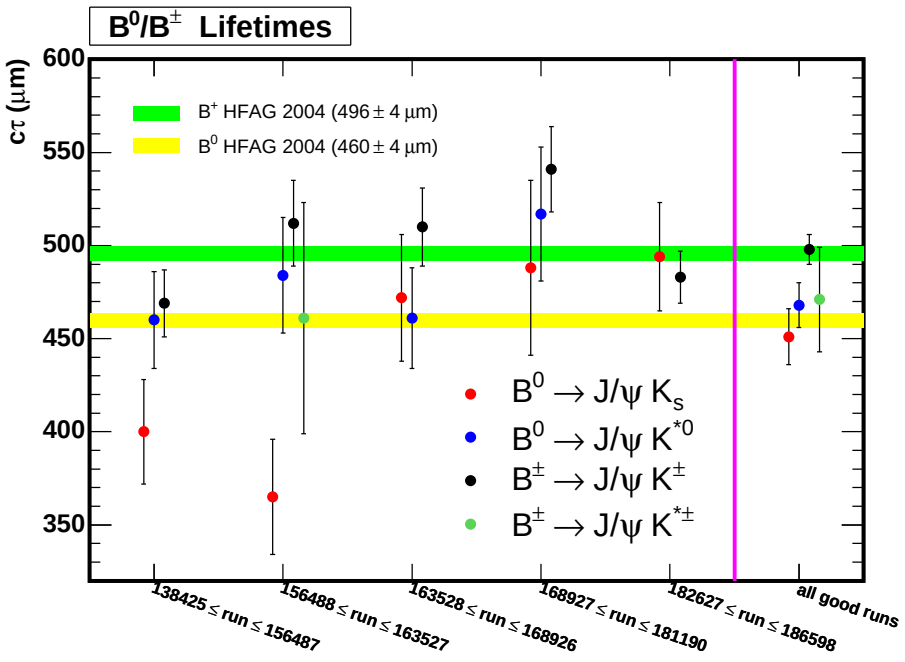


Figure 93: Run range dependence of the lifetimes measured in $B^0 \to J/\psi\, K^0_s$, $B^0 \to J/\psi\, \mathrm{K}^{*0}$, $B^+ \to J/\psi\, K^+$ and $B^+ \to J/\psi\, K^{*+}$.

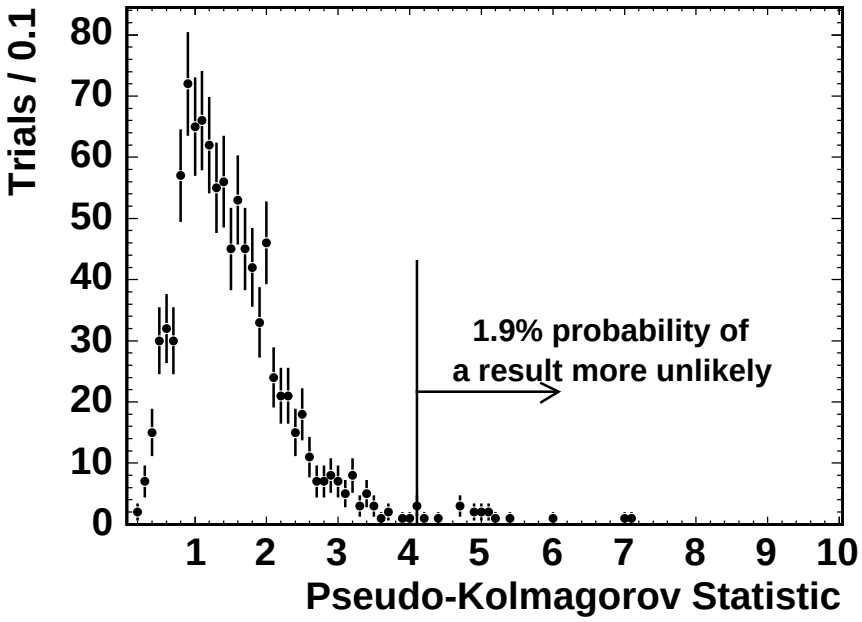


Figure 94: Distribution of the $D_{\max}$ statistic drawing subsamples from $B^0 \to J/\psi\, K^0_s$ data.

6.10.2 Probability of Observed Run Range Varation

The observed run dependence has a Prob(χ^2) of 4.3%. The χ^2 statistical however does not include the order of the bins in the probability calculation. We need to know what the probility of the first two bins being low is. In order to assess this we choose a pseudo-Kolmogorov statistic defined as:

$$D_{\max} = \big| \max_{1 \le n \le 5} \sum_{i=1}^{n} \sigma_i \big|,$$

where σ_i^2 is the signed significance of the deviation from the full data fit. Larger $D_{\max}$ corresponds to less likely results. We calculate probability distribution of $D_{\max}$ by randomly dividing the full data into 5 bins of corresponding size to the run ranged bins. Figure 94 shows the resulting distribution of $D_{\max}$ for sets of the randomly divide data which is the probability distribution of $D_{\max}$. Using this distribution we get a probablity of observing the run dependence observed in the $B^0 \to J/\psi\, K^0_s$ data of 1.9%.

6.10.3 Fit Range

The fit result should not depend strongly on the choice of mass and $c\tau$ windows for the fit. Variations in these windows however a dependent on statistical fluctuations of the events being added and subtracted. Because events on the tails of exponentials can play a large role in a likelihood fit, it is important that the results do not change dramatically as the $c\tau$ window is varied. The results of the window variations are shown in Table 6.10.3.

	$B^0 \to J/\psi\, K_s^0$	
Variation	Fitted $c\tau$ (μm)	Shift (μm)
-1000 $< c\tau <$ 3000 μm	448.2 $\pm$ 14.6	-2.4
-1500 $< c\tau <$ 3500 μm	450.6 $\pm$ 14.7	0.0
-1500 $< c\tau <$ 4500 μm	449.9 $\pm$ 14.8	-0.7
-1000 $< c\tau <$ 5000 μm	449.9 $\pm$ 14.8	-0.7
5.150 $<$ mass $<$ 5.410 GeV	448.8 $\pm$ 14.7	-1.8
5.155 $<$ mass $<$ 5.405 GeV	450.1 $\pm$ 14.7	-0.5
5.160 $<$ mass $<$ 5.400 GeV	449.4 $\pm$ 14.7	-1.2
5.175 $<$ mass $<$ 5.385 GeV	450.2 $\pm$ 14.7	-0.4
5.180 $<$ mass $<$ 5.380 GeV	450.8 $\pm$ 14.8	0.2
5.185 $<$ mass $<$ 5.375 GeV	451.6 $\pm$ 14.8	1.0
5.190 $<$ mass $<$ 5.370 GeV	453.1 $\pm$ 14.9	2.5

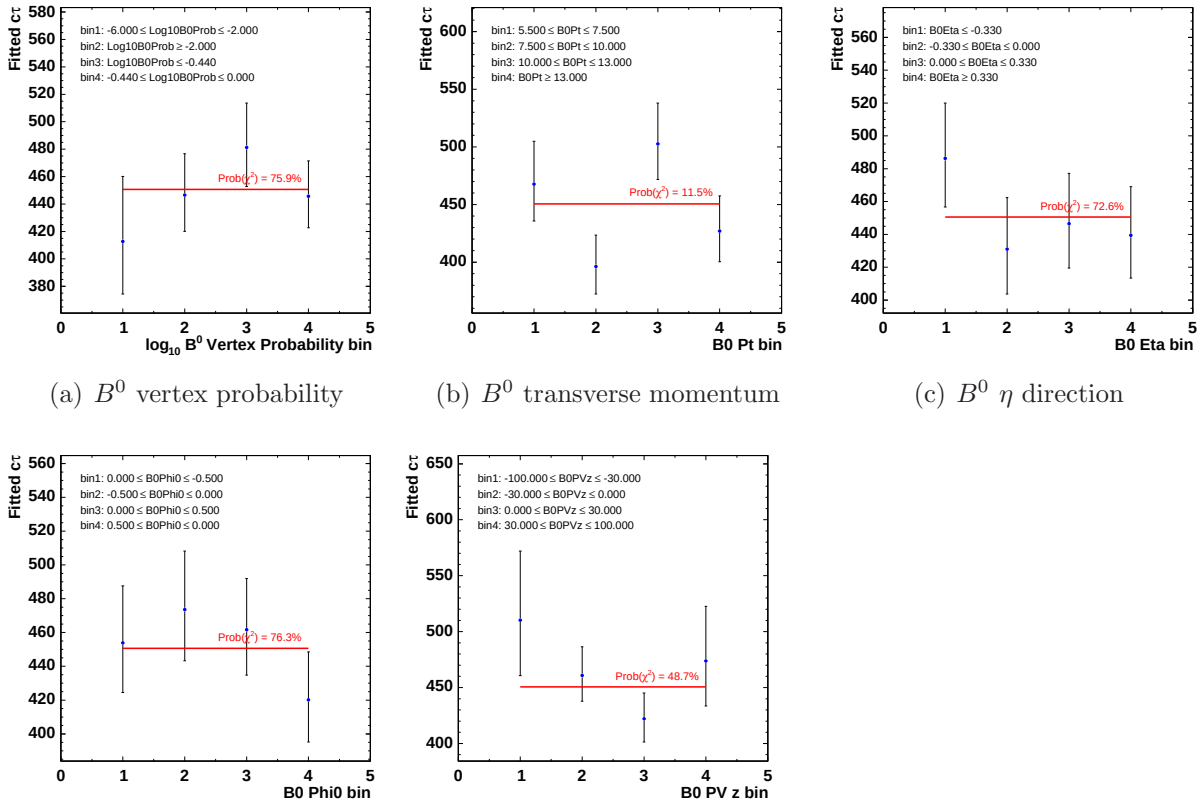
Table 16: Effect of variations of the fit region.

6.10.4 B^0 Prob(χ^2)

The b-hadron vertex probability is the main point at which we rely on the pointing properties the V^0. The contributions the the vertex probability are J/ψ muon vertex probability and mass, the V^0 vertex probability, and V^0 pointing constraint (our nominal fit uses a three-dimension pointing constraint). Figure 95(a) show the dependence of the fitted $c\tau$ on the B^0 vertex probability for the $B^0 \to J/\psi\, K_s^0$ mode.

6.10.5 B^0 P_T

The dependence of the fitted $c\tau$ on the B^0 P_T is shown in Figure 95(b). There is no particular reason to worry about such a dependence, but a broad range of the tracking and background features vary as a function of the B^0 P_T. In particular the background levels change significantly.



(a) B^0 vertex probability

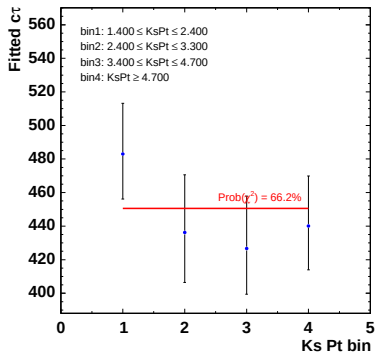
(b) B^0 transverse momentum

(c) B^0 η direction

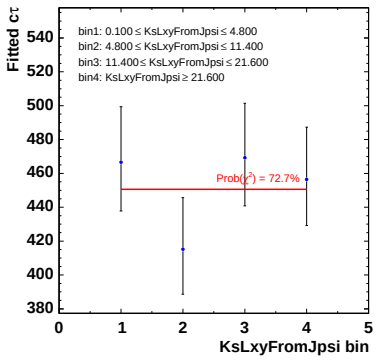
(d) B^0 ϕ direction

(e) z-position of the B^0 production vertex

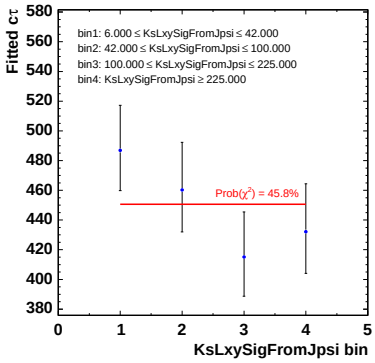
Figure 95: Dependence of the fitted $c\tau$ for the $B^0 \to J/\psi\, K^0_s$ mode on various quantities



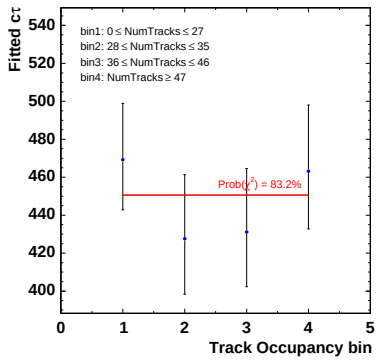
(a) K_s^0 transverse momentum



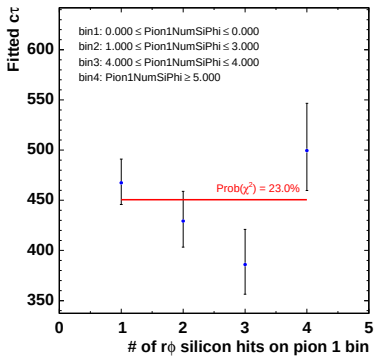
(b) K_s^0 $\mathrm{L_{xy}}$ measured from the J/ψ vertex



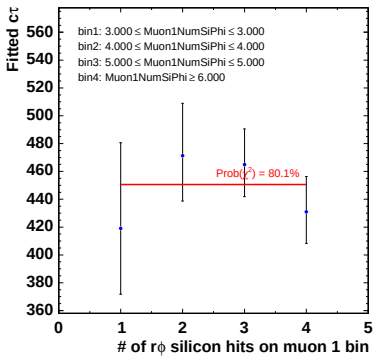
(c) Significance of the K_s^0 $\mathrm{L_{xy}}$ measured from the J/ψ vertex



(d) Total number of tracks in the event



(e) Number of $r\phi$ silicon hits on the first pion



(f) Number of $r\phi$ silicon hits on the first muon

Figure 96: Dependence of the fitted $c\tau$ for the $B^0 \to J/\psi\, K_s^0$ mode on various quantities (con't).

6.10.6 B^0 η

The scan the η dimension (shown in Figure 95(c) probes potential unexpected dependence on these kinematics.

6.10.7 B^0 ϕ^0

The CDF detector was designed to be approximately cylindrically symmetric, however there are a few real world effect that could cause a dependence. The most worrisome are the fact that the beam is not centered in the detector and the COT wire aging effected primarily wires are a the bottom of the detector (see Reference [3] for a more detailed study). The previous blessed measurement at CDF [7] an effect of marginal statistical significance was seen. Figure 95(d) shows the scan of fitted $c\tau$ as a function of B^0 ϕ^0 with no statistically significant effect.

6.10.8 B^0 Primary Vertex z-position

The slope of the beam is taken care of in the reconstruction and is too small to create an appreciable effect anyway. The scan of the B^0 Primary Vertex z-position (shown in Figure 95(e)) is just a cross-check for any unexpected problems in particular alignment.

6.10.9 K^0_s P_T

Several track topologies change as a function of the K^0_s P_T, in particular the opening angle between the two tracks. Also fake K^0_s background is dominantly at low K^0_s P_T. The scan of the fitted $c\tau$ as a function of the K^0_s P_T is shown in Figure 96(a).

6.10.10 K^0_s $\mathrm{L_{xy}}$ and $\mathrm{L_{xy}}/\sigma_{\mathrm{L_{xy}}}$ from J/ψ Vertex

The $\mathrm{L_{xy}}$ and $\mathrm{L_{xy}}/\sigma_{\mathrm{L_{xy}}}$ are discriminants for rejecting fake V^0 candidates. The V^0 pointing properties are also dependent on the $\mathrm{L_{xy}}$ because of the change in the extrapolation length from the V^0 vertex to the J/ψ vertex. Scans for these quantities are shown in Figures 96(b) and 96(c), respectively.

6.10.11 Track Occupancy

The tracking efficiency and systematics are very likely dependent on the track occupancy (which also correlates strong with the instantaneous luminosity). In Figure 96(d), we show dependence of the fitted $c\tau$ on the track occupancy for the $B^0 \to J/\psi\, K^0_s$ mode. There is no statistically significant correlation.

6.10.12 K_s^0 r-ϕ Silicon Hits

When the V^0 tracks have silicon hits, the pointing resolutions are substantially smaller than those without. The choice made above unambiguously divides the candidates into two non-overlapping subsets. A scan of the fitted $c\tau$ as a function of the number of silicon hits on the higher momentum pion track is shown in Figure 96(e). No statistically significant effects are observed. Since the number of silicon hits on the tracks primarily depends on the V^0 vertex position, the number of silicon hits on the two tracks is highly correlated.

6.10.13 J/ψ r-ϕ Silicon Hits

The PDL resolution is directly dependent on the number of silicon hits. Although the resolution is not a large contribution to the statistical uncertainty, it must be accounted for correctly. Furthermore tracks with low numbers of silicon hits are less constrained and more likely to pick up wrong hits, contributing to non-gaussian tails. A scan of the fitted $c\tau$ as a function of the number of silicon hits on the higher momentum muon track is shown in Figure 96(f). No statistically significant effects are observed.

6.10.14 Mass and PDL Error Cut

In order to remove clearly mismeasured events loose proper decay length error and mass error cuts are applied. The nominal values for these cuts are 100 μm for the proper decay length and 20 MeV for the mass. Table 6.10.14 shows fit results when these cuts are varied.

6.10.15 COT-only tracking for Ks

As described in Section 6.4, the addition of silicon hits to K_s^0 tracks introduces the potential for bias. A cross check analysis has be performed with COT-only tracks. A $4.1\mu m$ shift is observed. Because the two samples used to determine this shift are highly correlated the statistical signifacance of this result is not obvious. In order to determine the correlation of the two results, we use a bootstrap technique. The events in the two samples are correlated by run and event number creating a superset of events some of which are in both samples, some of which are in only one or the other. The number of instances of the events in the superset are then Poisson fluctated about the nominal 1 instance of that event outputing into two samples. This procedure is then repeat many times fitting the resulting samples in each iteration. Figure 97 shows the resulting correlation of the COT-only and DefTracks reconstructions and the expected difference which has an RMS of $10.4\mu m$. The observed shift can then be said to be $4.1 \pm 10.4\mu m$ showing the two methods are statistically consistent.

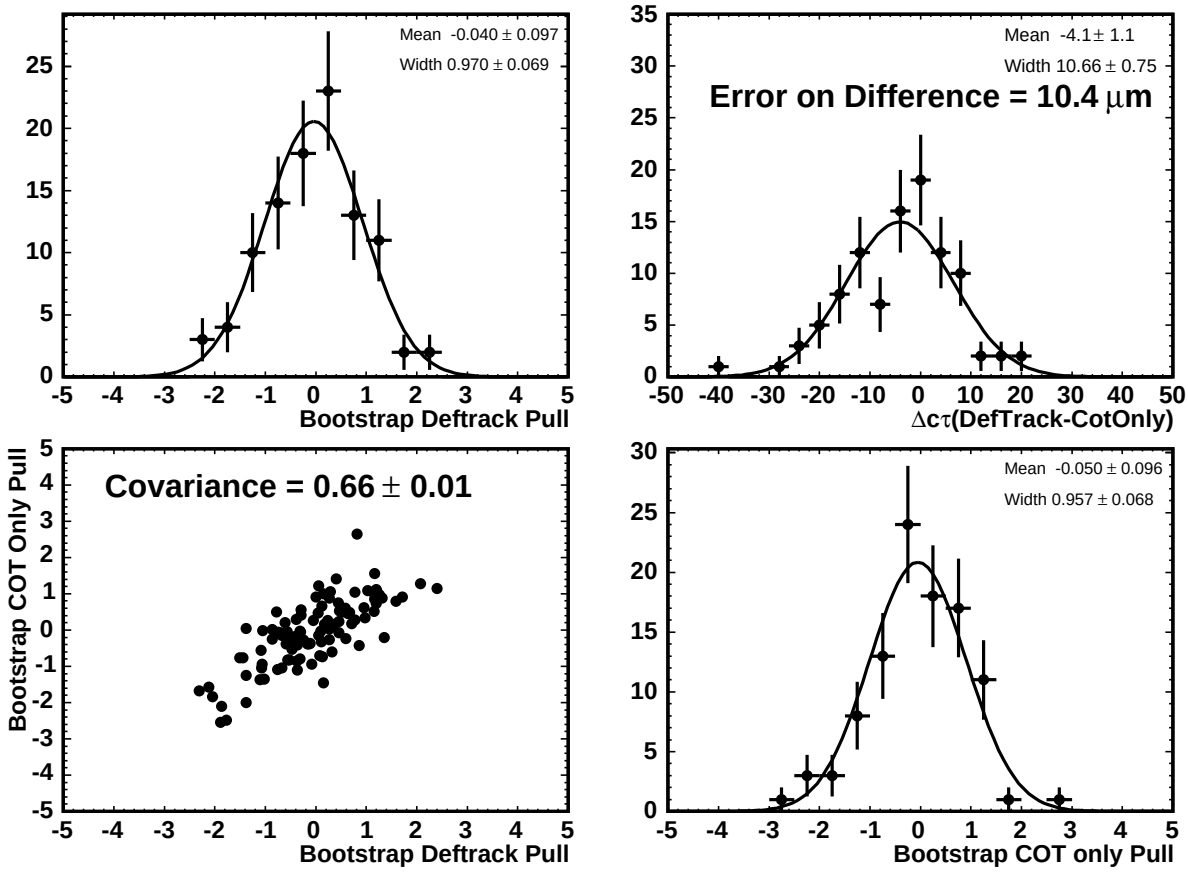


Figure 97: Results of the bootstraping procedure: one dimensional pull distributions for DefTrack (top-left) and COT-only (bottom-right), the two dimensional scatter plot (bottom-left), and the difference distribution (top-left). The one dimensional pull distributions are unit gaussians, demonstrating that the bootstrap sample have the expected correlation with the original results. The RMS of the difference distributions gives the statistical uncertainty on the difference between the two methods.

Proper Decay	$B^0 \to J/\psi\, K^0_s$		$\Lambda^0_b \to J/\psi\, \Lambda^0$	
Length Error (μm)	Fitted $c\tau$ (μm)	Shift (μm)	Fitted $c\tau$ (μm)	Shift (μm)
70	448.9 ± 14.8	-1.7	649.1 ± 61.3	4.8
80	448.5 ± 14.7	-2.1	640.5 ± 60.7	-3.7
90	448.8 ± 14.7	-1.8	638.3 ± 60.0	-5.9
110	449.8 ± 14.7	-0.8	644.9 ± 59.8	0.7
120	449.7 ± 14.7	-0.9	646.7 ± 59.9	2.5
130	449.6 ± 14.7	-1.0	647.4 ± 59.9	3.2

	$B^0 \to J/\psi\, K^0_s$		$\Lambda^0_b \to J/\psi\, \Lambda^0$	
Mass Error (MeV)	Fitted $c\tau$ (μm)	Shift (μm)	Fitted $c\tau$ (μm)	Shift (μm)
10	443.9 ± 15.1	-6.7	643.2 ± 59.4	-1.0
15	451.9 ± 14.9	1.3	646.8 ± 60.9	2.5
25	450.9 ± 14.7	0.3	643.8 ± 59.7	-0.4
30	450.9 ± 14.7	0.3	644.1 ± 59.8	-0.1
no cut	450.5 ± 14.7	-0.1	643.5 ± 59.7	-0.7

Table 17: Dependence of the Result on the Mass and Proper Decay Length Error Requirements.

6.10.16 Variations of the b-Hadron Vertex Fit Constraints

The b-hadron vertex probability used in the selection, is a combination of several pieces of information, including the V^0 pointing constraint, the V^0 and J/ψ vertex χ^2 and the J/ψ mass. In order to check that these do not excessively influence the result, we have fit with many different combinations of the constraints. The results are summarized in Table 6.10.16. Keeping in mind that the backgrounds and the statistical uncertainty are increased as the various constraints are released the agreement is good.

6.10.17 Variations on PDL Calculation

:

Two variations have been made in the calculation of the proper decay length. First, the B^0 vertex was used instead of nominal the J/ψ vertex. Second, the mass candidate mass was used in place of the PDG mass for the b-hadron in PDL calculation (see Equation 3 in Section 4.1). The results of these two variations are shown in Table 6.10.17.

As we see, our results effectively unchanged due to these changes in the PDL calculation.

Constraints			$B^0 \to J/\psi\, K^0_s$		$\Lambda^0_b \to J/\psi\, \Lambda^0$	
Pointing	V^0 Mass	COT-only	Fitted $c\tau$ (μm)	Shift (μm)	Fitted $c\tau$ (μm)	Shift (μm)
2d	X		458.9 ± 15.4	8.3	642.0 ± 60.8	-2.3
3d	X		453.4 ± 14.9	2.8	634.4 ± 60.5	-9.8
None	X		466.8 ± 17.2	16.2	576.2 ± 64.9	-68.1
2d			451.2 ± 15.1	0.6	660.1 ± 61.4	15.8
3d			450.6 ± 14.7	0.0	644.2 ± 59.8	0.0
None			461.6 ± 17.0	11.0	619.9 ± 66.8	-24.4
2d	X	X	464.6 ± 18.0	14.0	617.7 ± 65.9	-26.5
2d		X	458.2 ± 18.0	7.6	619.3 ± 63.9	-24.9
3d	X	X	460.9 ± 17.0	10.3	611.2 ± 62.9	-33.0
3d		X	455.0 ± 16.8	4.4	618.9 ± 61.0	-25.3

Table 18: Dependence of the result on alternative b-vertex constrains.

Variation	Fitted $c\tau$ (μm)
B^0 vertex	451.1 ± 14.6
candidate mass	450.9 ± 14.8

Table 19: Dependence of the fitted B^0 $c\tau$ on choices made in the proper decay length calculation.

6.10.18 Stability Against Adding Long-lived Backgrounds

6.11 Summary of Systematics

The systematics are summarized in Table 20.

Source	$B^0 \to J/\psi\, K^0_s$ $[\mu m]$	$\Lambda^0_b \to J/\psi\, \Lambda^0$ $[\mu m]$	ratio
Fitter Bias	0.2	0.05%	
Fit Model:			
$c\tau$ Resolution	3.0	0.3%	
Mass Signal	1.8	0.3%	
Mass Background	0.1	0.0%	
PDL Background	0.6	0.9%	
Mass-dependent PDL Background	0.9	1.6%	
PDL Error Modeling	3.6	0.6%	
Mass Error Modeling	0.6	0.09%	
Primary Vertex Determination	0.2	0.05%	
Alignment systematics	$^{+7}_{-2}$	$^{+5}_{-3}$ $(^{+1.3}_{-0.8}\%)$	$^{+.008}_{-.010}$
V^0 Pointing	1.0	1.0 (0.2%)	0.002
CDF Length Scale	$^{+9.7}_{-2.8}$	$^{+9.7}_{-2.8}$ $(^{+2.6\%}_{-0.8\%})$	-
Total	$^{+13.1}_{-6.3}$	$^{+3.5\%}_{-2.3\%}$	

Table 20: Systematic Uncertainties

7 Summary and Conclusions

This note describes a blinded measurement of the Λ^0_b lifetime. The full procedure including all systematic studies has been validated on the $B^0 \to J/\psi\, K^0_s$ mode where we find the lifetime to be $c\tau = 450.6\ ^{+15.0}_{-14.5}(stat.)\ ^{+13.1}_{-6.3}(syst.)\mu$m. This is consistent with more precise the PDG world average for B^0 lifetime, $c\tau = 460 \pm 4$. The uncertainty is statistically dominated, with the systematic error dominated by the uncertainty on the length scale in the CDF detector as measured with $B^+ \to J/\psi\, K^+$ and $B^0 \to J/\psi\, \mathrm{K}^{*0}$, as well as the systematics due the COT-SVX alignment. A large number of studies have been performed to search for potential biases; none have been found. We have also performed the full systematic analysis on the Λ^0_b data using a blinded fitter and are currently prepared to unblind the result after review by the B-group. We expect the uncertainties on the Λ^0_b lifetime to be 9.1% statistical and $^{+3.5\%}_{-2.3\%}$ systematic.

References

[1] F.Gabbiani, A.Onishchenko, and A.Petrov, Phys. Rev. D **70**, 094031 (2004).

[2] Heavy Flavor Averaging Group, *Averages of b-hadron Properties as of Summer 2004*, hep-ex/0412073.

[3] R. Madrak, et al, "Update on the Measurement of the Λ_b^0 Lifetime in the Decay Mode $\Lambda_b^0 \to J/\psi\,\Lambda$" CDF Note 7520

[4] E. Lipeles, M. Neubauer, F. Wuerthwein, "Measurement of the Λ_b^0 Lifetime in $\Lambda_b^0 \to J/\psi\,\Lambda$" CDF Note 7503.

[5] R. Madrak, et al, "Update on the Measurement of the Λ_b^0 Lifetime in the Decay Channel $\Lambda_b^0 \to J/\psi\,\Lambda$" CDF Note 7353.

[6] R. Madrak, et al, "Measurement of the Λ_b^0 Lifetime in the Decay Mode $\Lambda_b^0 \to J/\psi\,\Lambda$"

[7] "Measurement of the Λ_b^0 Lifetime in the Decay Mode $\Lambda_b^0 \to J/\psi\,\Lambda$", Ph.D. Thesis by R. Madrak, Harvard (2004).

[8] "Measurement of the Λ_b^0 lifetime in the decay $\Lambda_b^0 \to J/\psi\,\Lambda^0$ with the D0 detector", *Phys.Rev.Lett.* 94 (2005) 102001.

[9] Saverio D'Auria, *et al.*, *Run Dependent MC Production from Summer 2004*, CDF Note 7055.

[10] M. Campanelli and E. Gerchtein, *Calibration of the Momentum Scale for Kalman Refitter using J/ψ Events*, CDF Note 6905 (2004).

[11] D. Glenzinski, et al, *Kinematic-Dependent Matching Cut for Low-p_T CMX Muons*, CDF Note 6835 (2004).

[12] `http://www-cdf.fnal.gov/internal/physics/bottom/bpak/goodrun/goodrun.sql`

[13] J. Marriner, *Secondary Vertex Fit with Mass and Pointing Constraints (CTVMFT)*, CDF Note 1996.

[14] W. Ashmanskas and A. Mukherjee, *CTC Alignment for Run Ib*, CDF Note 4456 (1998).

[15] The CDF Collaboration, *Observation of $B^+ \to \psi(2S)K^+$ and $B^0 \to \psi(2S)K^{*0}$ decays and measurement of B-meson branching fractions into J/ψ and $\psi(2S)$ final states*, *Phys.Rev.* **D58**, 072001 (1998) (CDF Note 4425).

[16] K. Anikeev, *et.al.*, *B Meson Lifetime Measurements Using Exclusively Reconstructed Decays* $B \to J/\psi\, X$ *in Run II Data*, CDF Note 6266 (2003). Updated CDF Note 6550 (2003).

[17] Giovanni Punzi, "Comments on Likelihood fits with variable resolution", e-print *physics/0401045*

[18] J. Alison, *et.al.*, *Estimating the Size of the Punzi Effect on Lifetime Measurements in the* $B^0 \to J/\psi\, K^0_s$ *channel*, CDF Note 7782.

[19] D. Acosta, *et.al*, The CDF Collaboration, *Measurement of the* J/ψ *Meson and b-Hadron Production Cross Sections in ppbar Collisions at* $\sqrt{(}s) = 1960$ *GeV*, *Phys.Rev.* **D71**, 032001 (2005).

[20] G. Apollinari, *et.al*, *Observation of* $\Lambda^0_{\rm b} \to \Lambda_{\rm c}\pi$ *and the measurement of* $(f_{\Lambda^0_{\rm b}}/f_d)({\rm BR}(\Lambda^0_{\rm b} \to \Lambda_{\rm c}\pi)/{\rm BR}(\bar{\rm B}^0 \to {\rm D}^+\pi^-))$, CDF Note 6396.

A Exponential Convoluted with a Gaussian

In our analysis, we use in numerous places (e.g. PDL background model, PDL error modeling) the exponential convoluted with a Gaussian resolution function

$$\mathcal{C}(x)) \equiv E(x|\lambda) \otimes G(x|\sigma,\mu) = \int_{-\infty}^{+\infty} E(\overline{x}|\lambda)G(x-\overline{x}|\sigma,\mu)d\overline{x} \tag{25}$$

where

$$G(x|\sigma,\mu) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{26}$$

and

$$E(\lambda_{\rm i}|c\tau) = \begin{cases} \frac{1}{c\tau} e^{-\lambda_{\rm i}/c\tau}, & c\tau \geq 0 \\ 0, & \text{otherwise} \end{cases} \tag{27}$$

$$\mathcal{C}(x) = \frac{1}{2\lambda} \exp\left(\frac{\sigma^2}{2\lambda^2} - \frac{x-\mu}{\lambda}\right) \text{erfc}\left(\frac{\sigma}{\sqrt{2}\lambda} - \frac{x-\mu}{\sqrt{2}\sigma}\right) \tag{28}$$

Use of Equation 28 in maximum likelihood fitting often requires integration of this equation for purposes of proper normalization. The example of this in our analysis is the PDL error PDF which is fit in a restricted range in PDL error (see 4.5.3). Fortunately, the definite integral of Equation 28 can be computely analytically. The general result is

$$\mathcal{I}(a,b) \equiv \int_a^b \mathcal{C}(x)dx = \lambda\left[\mathcal{C}(a) - \mathcal{C}(b)\right] + \frac{1}{\sqrt{\pi}} \int_{\frac{a-\mu}{\sqrt{2}\sigma}}^{\frac{b-\mu}{\sqrt{2}\sigma}} e^{-z^2} dz \tag{29}$$

For our specific case of the PDL error where we wish to normalize the function between $a = -\infty$ and $b = m$, where m is the maximum PDL error, we get

$$\mathcal{I}(m) = -\lambda\mathcal{C}(m) + \frac{1}{2}\left[1 + \text{erf}(m)\right] \tag{30}$$

Note that $\mathcal{I}(\infty) = 1$, as expected.

In practical use of this analytic form for the convolution in our fitter a numerical accuracy problem arises for large arguments of the exponential and erfc functions. In this case, the exponential tends toward infinity and the erfc function toward zero but the product remains finite and well-behaved. For example, one will encounter large arguments when a data point has large $c\tau$ error compared to the proper decay length ($\sigma_{\rm i}^{\lambda}/c\tau \gg 1$) in the convolution equation which leads to numerical inaccuracies that cause the fit to have trouble converging. To avoid this problem, we use the asymptotic expansion of erfc(ξ) for large ξ

$$\text{erfc}(\xi) = \frac{\text{e}^{-\xi 2}}{\xi\sqrt{\pi}} \left[1 + \sum_{\text{n}=1}^{\infty} (-1)^{\text{n}} \frac{(2\text{n})!}{\text{n}!(2\xi)^{2\text{n}}}\right] \tag{31}$$

In practice, we use the first three terms of the sum for the convolution when $\sigma_{\rm i}^{\lambda}/c\tau > 700$.

B Detailed Comparisons of Analyses

The analysis was done collaboratively, but seperately, by three different groups: FNAL, Madrid, and UCSD. The reconstruction procedure, cuts, and lifetime fitters do the same thing, but each implementation was with completely independent code.

Though every effort was made to assure that the the analysis was done in exactly the same way in all groups. some slight differences remain. We address this in the following sections.

B.1 $B^0 \to J/\psi\, K^0_s$

Figures 98 and 99 show the yields and lifetimes obtained by the FNAL and Madrid groups. These are to be compared with the corresponding results presented in the main part of the note (UCSD).

Yields and lifetimes obtained by each group are listed in Table 21.

Figures 100 through 102 show distributions of the differences in various quantities obtained by each group (for common events). The event overlap is 96%.

Group	B^0 Yield	B^0 Lifetime (μm)
FNAL	1230 $\pm$ 54	451.6 $\pm$ 14.7
Madrid	1221 $\pm$ 54	448.6 $\pm$ 14.7
UCSD	1225 $\pm$ 53	450.6 $\pm$ 14.7

Table 21: Results obtained independently by each group, for B^0

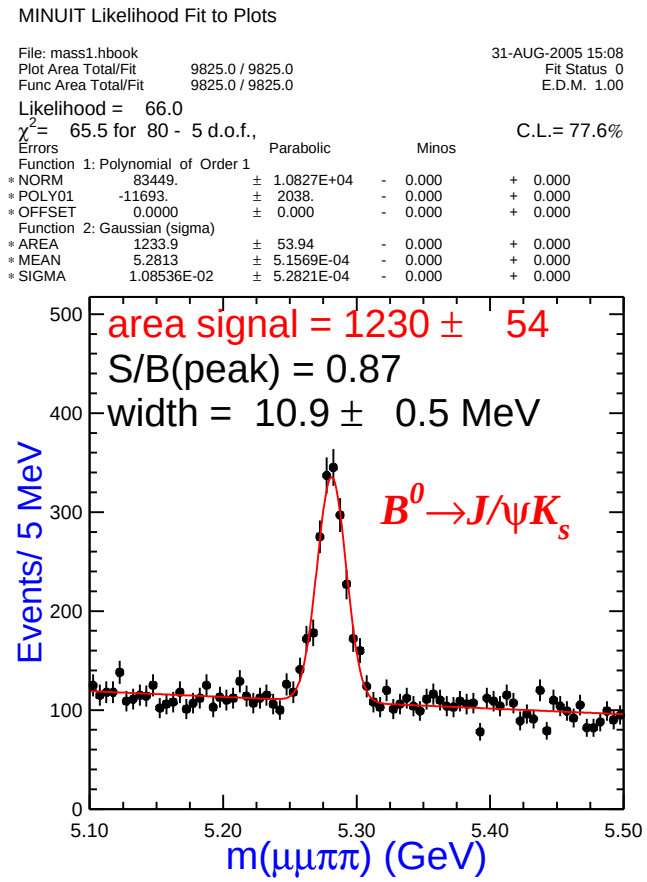
B.2 $\Lambda^0_{\rm b} \to J/\psi\, \Lambda^0$

The results for $\Lambda^0_{\rm b}$, analagous to those for B^0, discussed in the previous section, are shown in Table 22 and Figures 103(a) and 103(b).

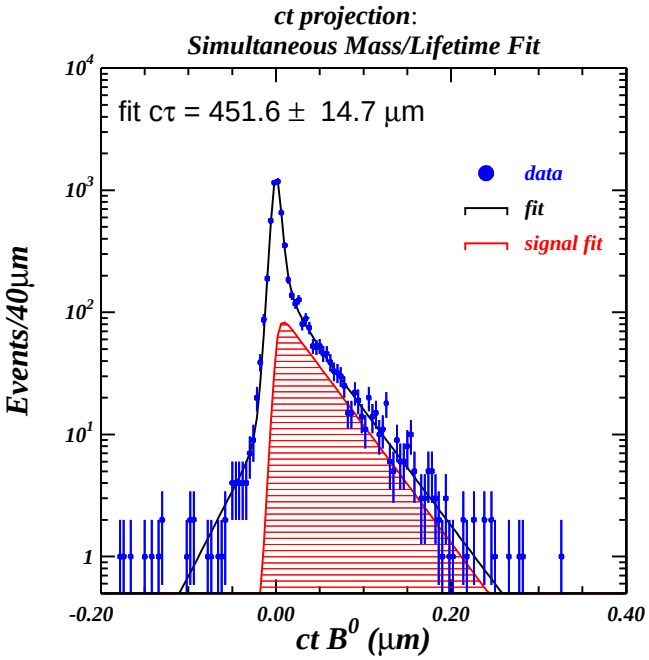
The difference plots for common events are shown in Figures 104 through 106. Event overlap is 95%.

Group	$\Lambda^0_{\rm b}$ Yield	$\Lambda^0_{\rm b}$ Lifetime (μm)
FNAL	181 $\pm$ 23	
Madrid	189 $\pm$ 23	
UCSD	194 $\pm$ 23	

Table 22: Results obtained independently by each group, for $\Lambda^0_{\rm b}$.

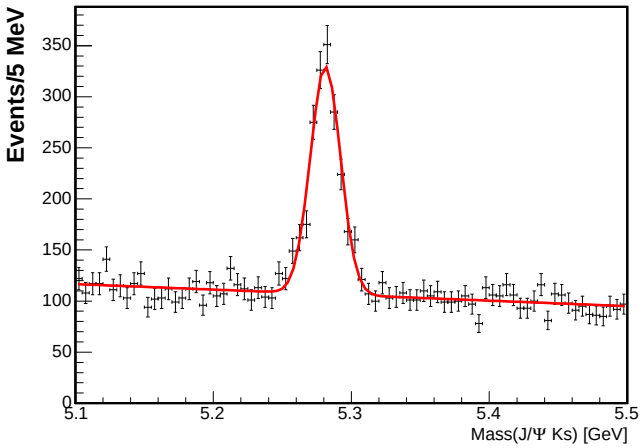


(a) Event yield in $B^0 \to J/\psi\, K_s^0$, FNAL group

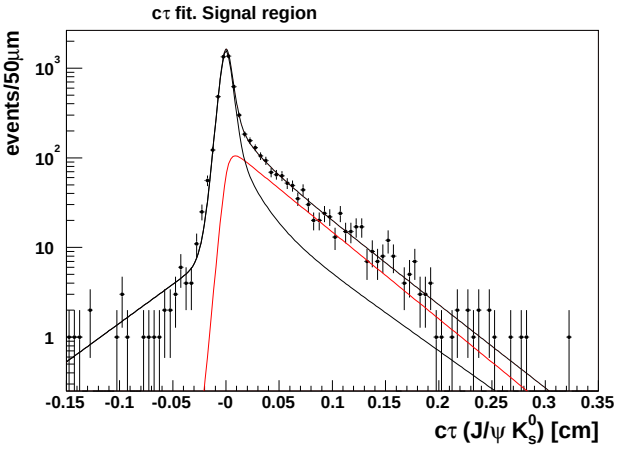


(b) Lifetime in $B^0 \to J/\psi\, K_s^0$, FNAL group

Figure 98: Mass and Lifetime projections for $B^0 \to J/\psi\, K_s^0$ from the FNAL group

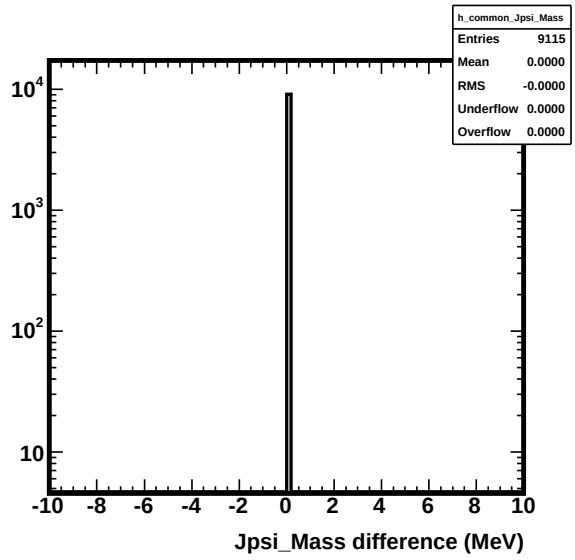


(a) Event yield in $B^0 \to J/\psi\, K_s^0$, Madrid group

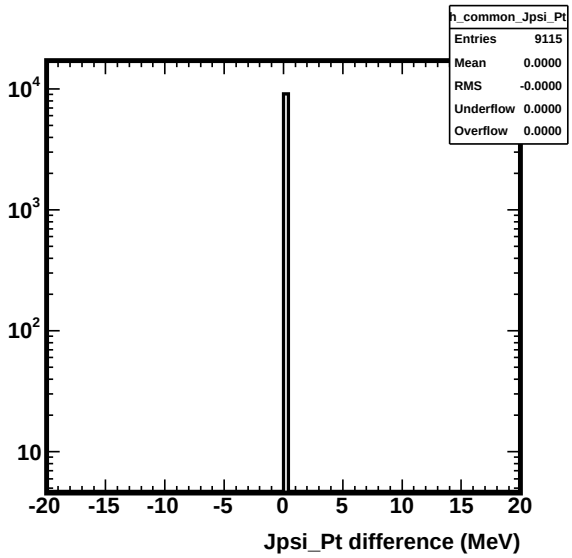


(b) Lifetime in $B^0 \to J/\psi\, K_s^0$, Madrid group

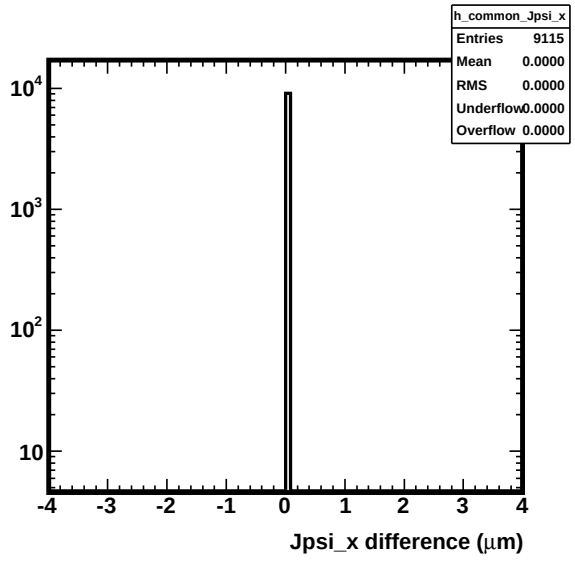
Figure 99: Mass and Lifetime projections for $B^0 \to J/\psi\, K_s^0$ from the Madrid group



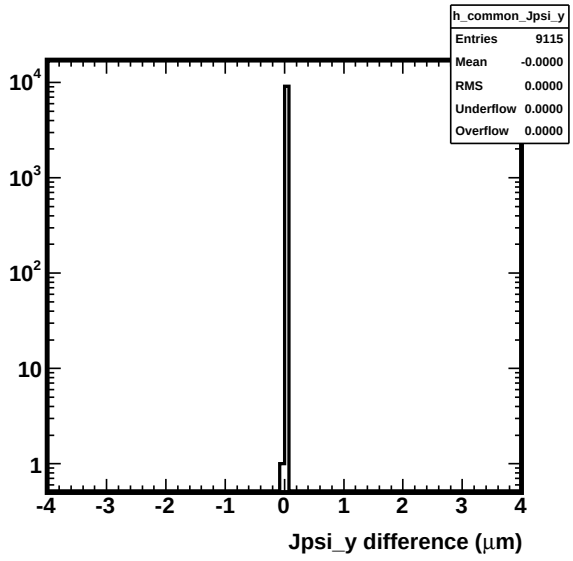
(a) J/ψ Mass



(b) J/ψ P_T

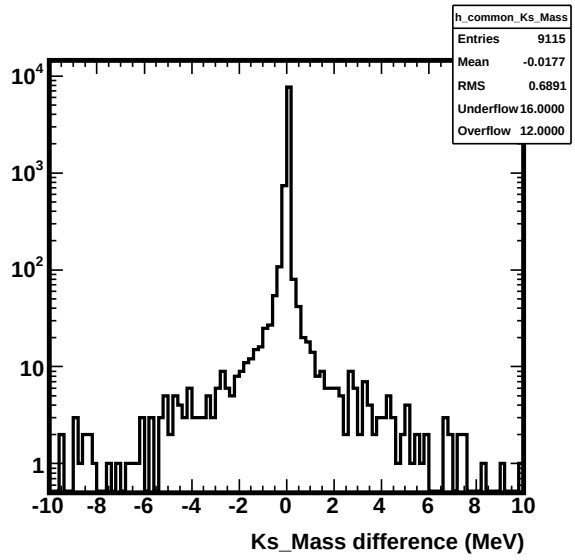


(c) J/ψ x-vertex position

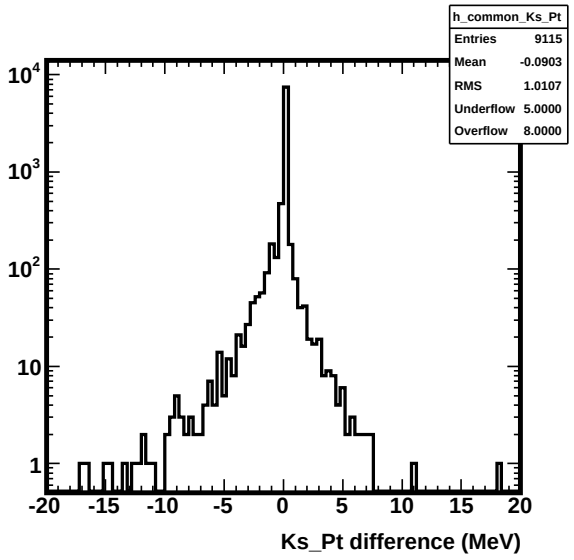


(d) J/ψ y-vertex position

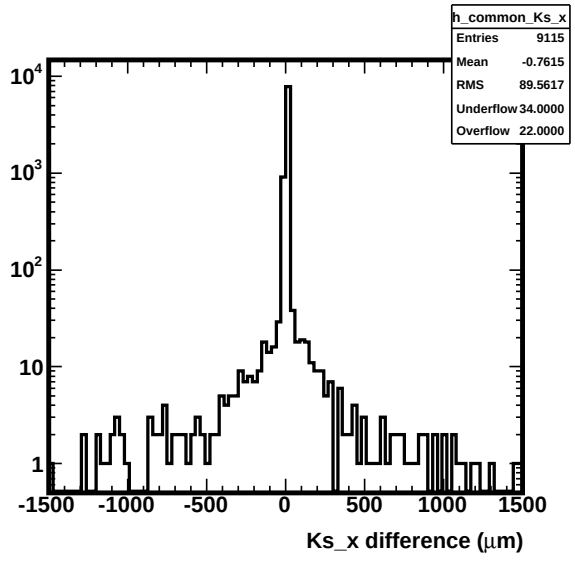
Figure 100: Comparision of the J/ψ properties in the $B^0 \rightarrow J/\psi K_s^0$ overlap events between FNAL and UCSD



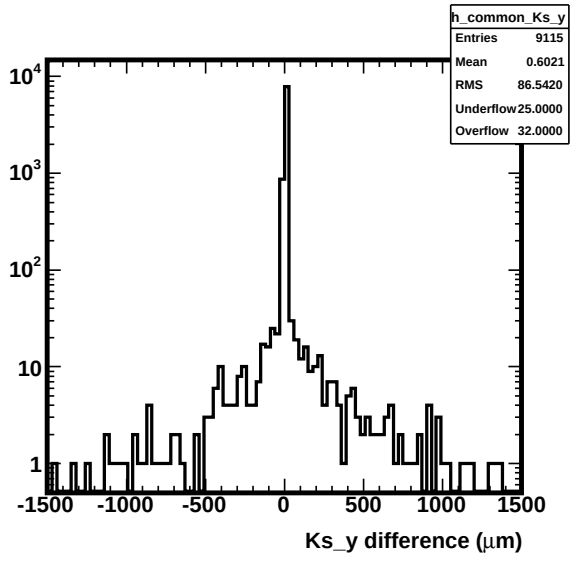
(a) K_s^0 Mass



(b) K_s^0 P_T

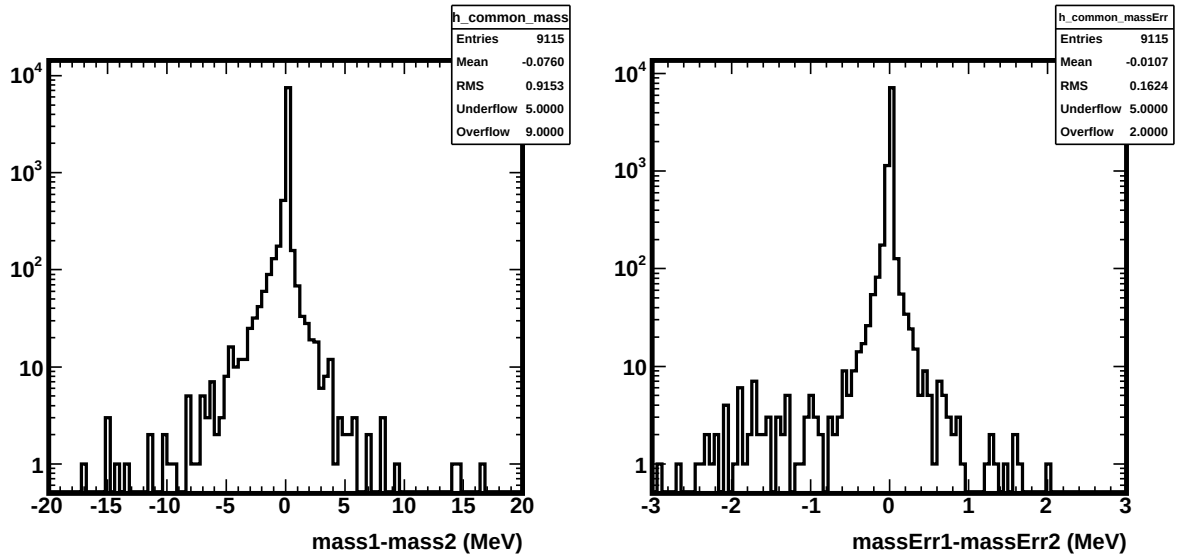


(c) K_s^0 x-vertex position



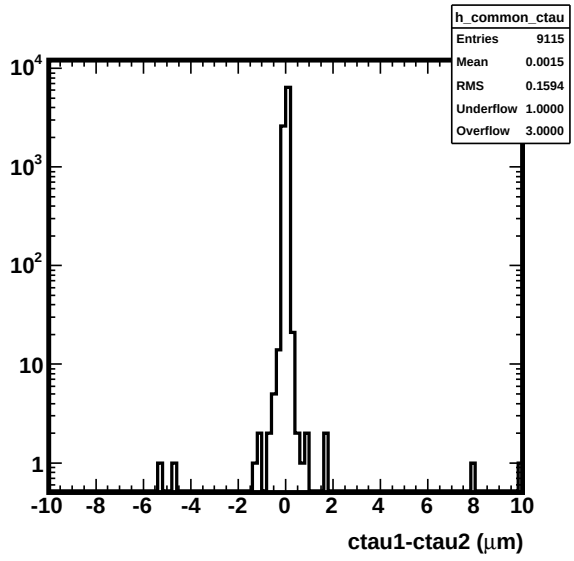
(d) K_s^0 y-vertex position

Figure 101: Comparision of the K_s^0 properties in the $B^0 \rightarrow J/\psi\, K_s^0$ overlap events between FNAL and UCSD

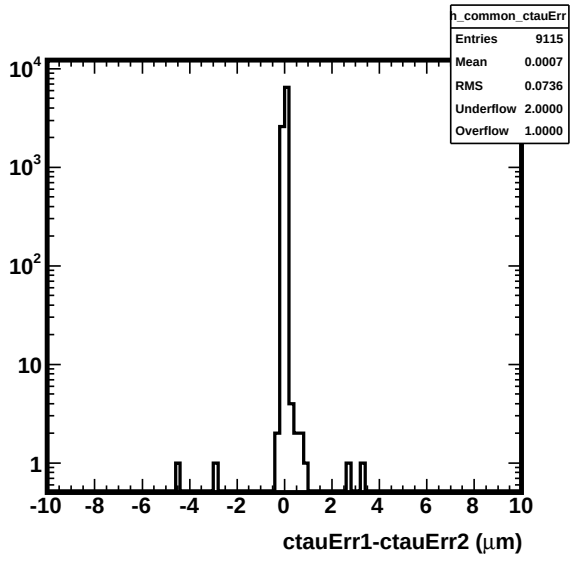


(a) B^0 Mass

(b) B^0 Mass Error

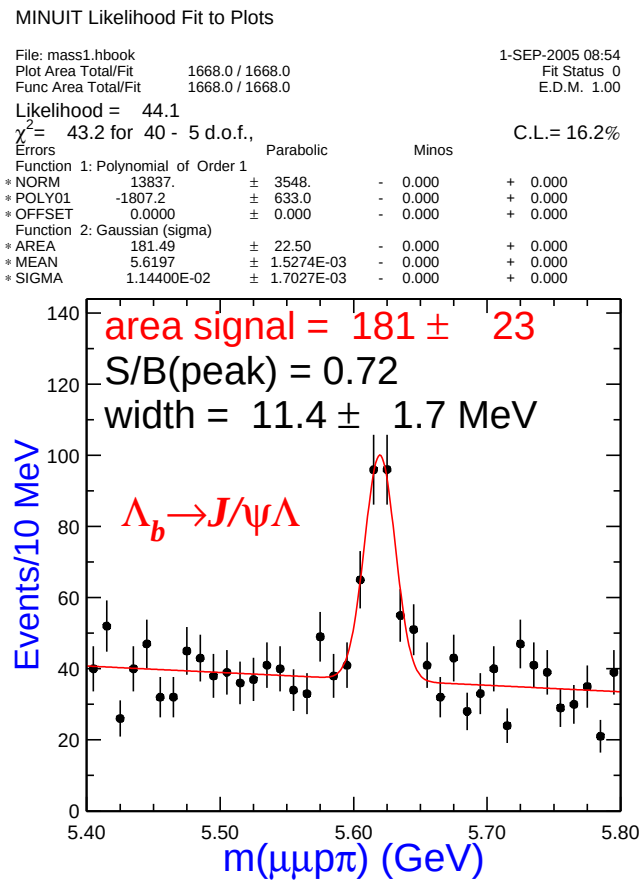


(c) B^0 $c\tau$

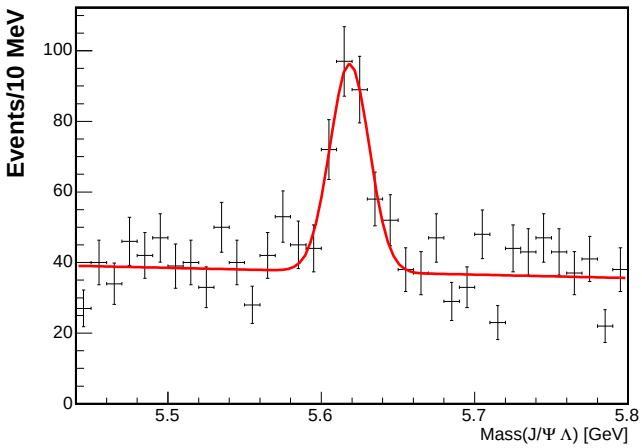


(d) B^0 $c\tau$ Error

Figure 102: Comparision of the B^0 properties in the $B^0 \to J/\psi\, K^0_s$ overlap events between FNAL and UCSD

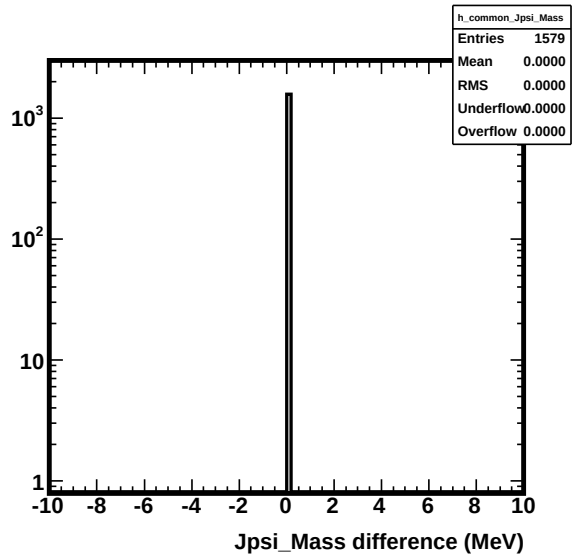


(a) Event yield in $\Lambda_b^0 \to J/\psi\,\Lambda^0$, FNAL group

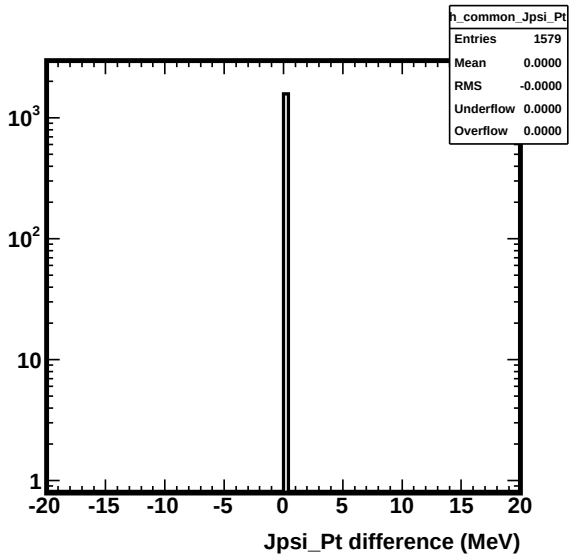


(b) Event yield in $\Lambda_b^0 \to J/\psi\,\Lambda^0$, Madrid group

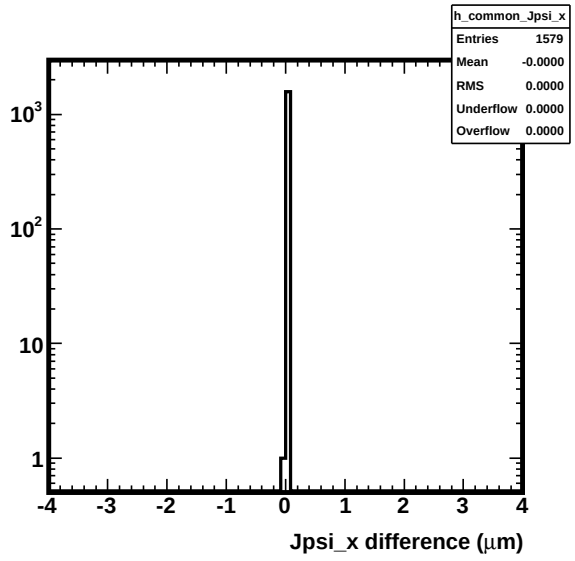
Figure 103: Mass distributions for the FNAL and Madrid $\Lambda_b^0 \to J/\psi\,\Lambda^0$ reconstructions



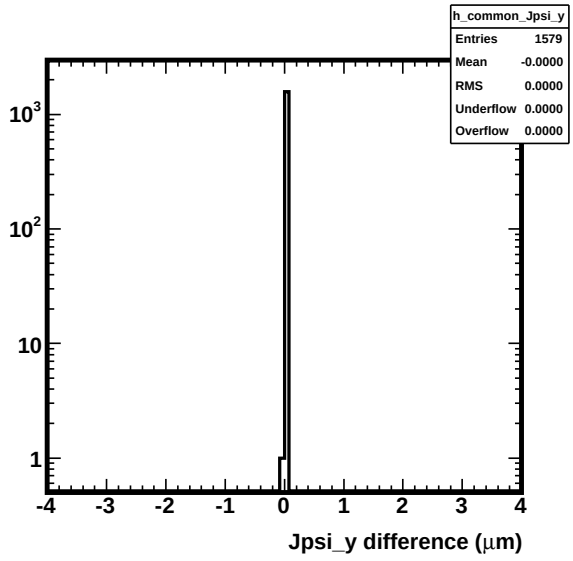
(a) J/ψ Mass



(b) J/ψ P_T

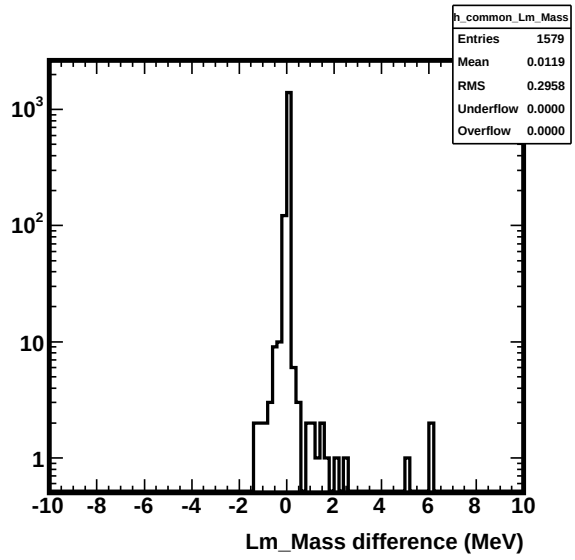


(c) J/ψ x-vertex position

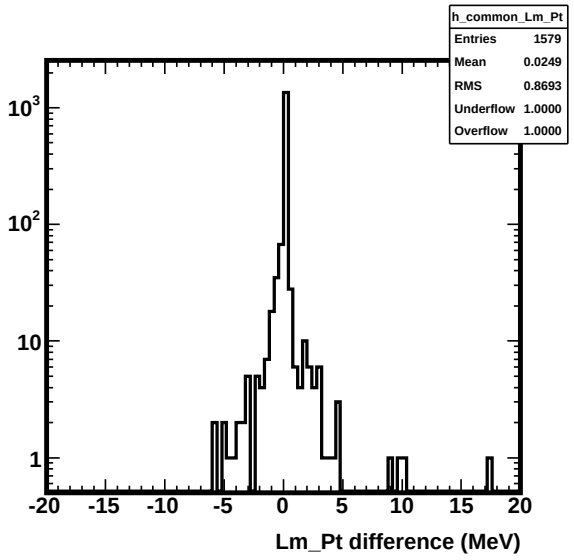


(d) J/ψ y-vertex position

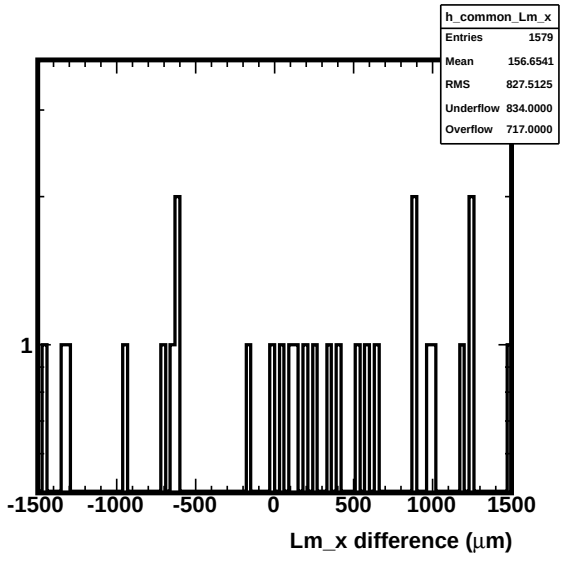
Figure 104: Comparision of the J/ψ properties in the $\Lambda_b^0 \to J/\psi\,\Lambda^0$ overlap events between FNAL and UCSD



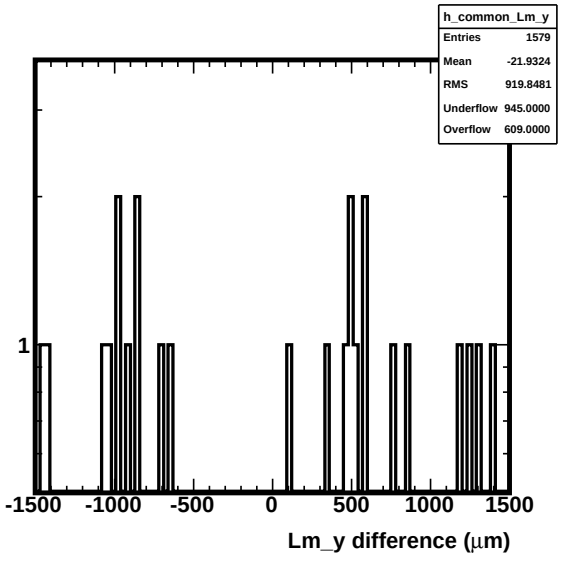
(a) Λ Mass



(b) Λ P_T

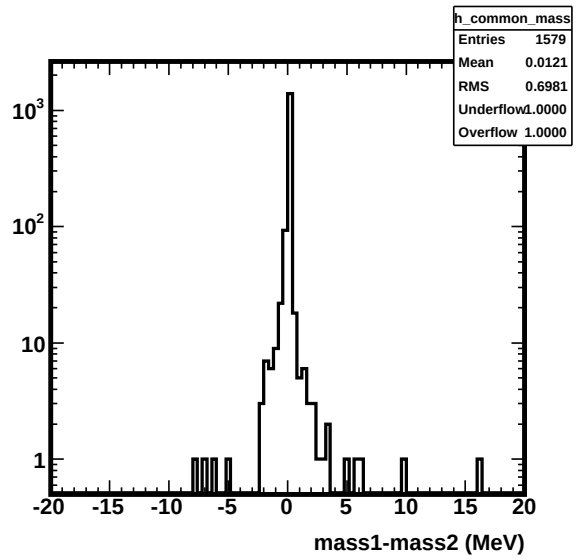


(c) Λ x-vertex position

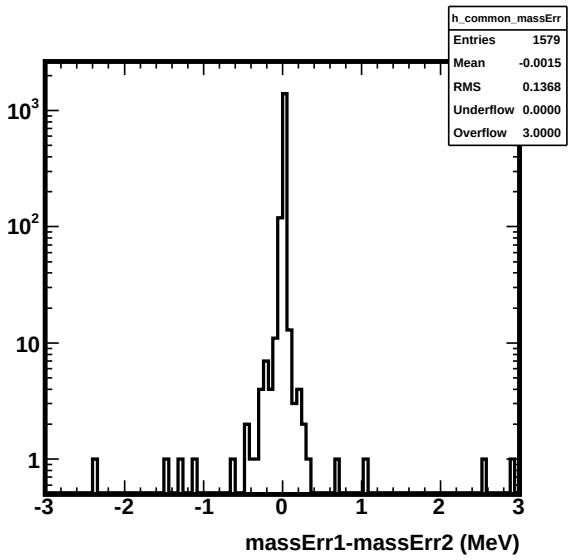


(d) Λ y-vertex position

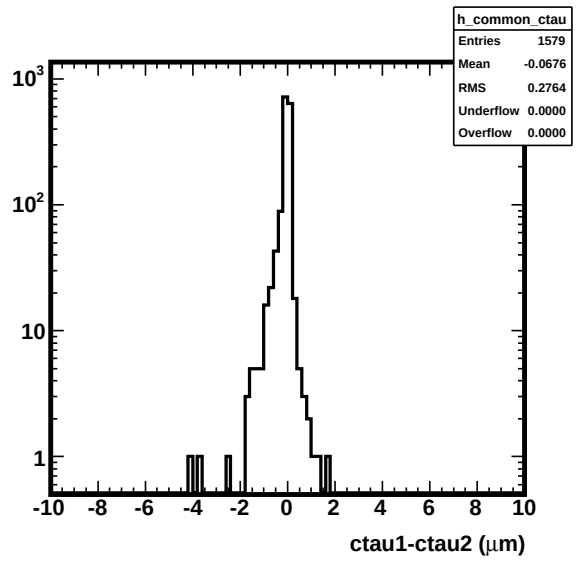
Figure 105: Comparision of the Λ properties in the $\Lambda_b^0 \to J/\psi\,\Lambda^0$ overlap events between FNAL and UCSD



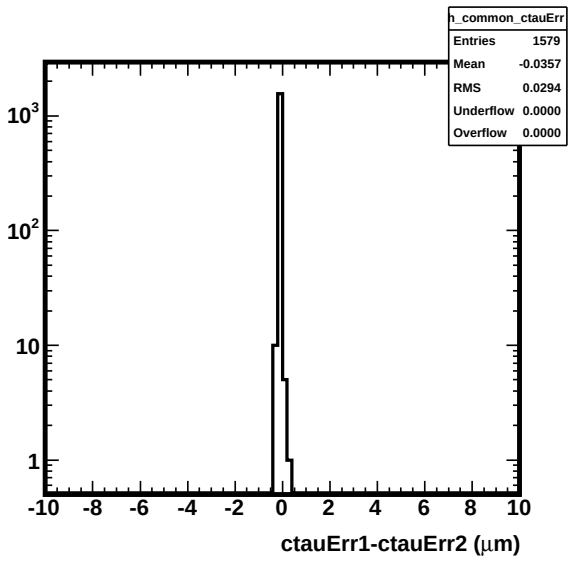
(a) Λ_b^0 Mass



(b) Λ_b^0 Mass Error



(c) Λ_b^0 $c\tau$



(d) Λ_b^0 $c\tau$ Error

Figure 106: Comparision of the Λ_b^0 properties in the $\Lambda_b^0 \to J/\psi\,\Lambda^0$ overlap events between FNAL and UCSD