SLAC-TN-07-023

Study of K_S Production With The *BABAR* Experiment

Thomas J. Colvin

Office of Science, Science Undergraduate Laboratory Internship (SULI)

The Ohio State University

Stanford Linear Accelerator Center

Stanford, CA

August 24, 2007

Prepared in partial fulfillment of the requirements of the Office of Science, Department of Energy's Science Undergraduate Laboratory Internship under the direction of Jochen Dingfelder at the BaBar, Stanford Linear Accelerator Center.

Participant: ______________________

Signature

Research Advisor: ______________________

Signature

Work supported in part by US Department of Energy contract DE-AC02-76SF00515

TABLE OF CONTENTS

ABSTRACT

Study of K_S Production With The *BABAR* Experiment. THOMAS J. COLVIN (The Ohio State University, Columbus OH, 43210) JOCHEN DINGFELDER (BaBar, Stanford Linear Accelerator Center, Stanford, CA 94025)

We study the inclusive production of short-lived neutral kaons (K_S) with the *BABAR* experiment at the Stanford Linear Accelerator Center. The study is based on a sample of 383 million $B\overline{B}$ pairs produced in e^+e^- collisions at the $\Upsilon(4S)$ resonance, in which one B meson has been fully reconstructed. We select a clean sample of K_S mesons and compare kinematic spectra for data and simulation. We find that the simulation overestimates the total production rate of K_S and we see differences in the shape of the K_S momentum spectra. We derive correction factors for different momentum intervals to bring the simulation into better agreement with the observed data.

INTRODUCTION

The *BABAR* experiment [1] at the Stanford Linear Accelerator Center (SLAC) studies the decays of B mesons, which are bound states of a b-quark and a light quark. Its goals are studying CP-violation in the B system and the properties of the electroweak interaction, as well as measurements of rare B-meson decays and much more. SLAC operates as a B meson factory, producing B and $\overline{B}$ meson pairs from e^+e^- collisions that are tuned to the $\Upsilon(4S)$ resonance ($e^+e^- \rightarrow \Upsilon(4S) \rightarrow B\overline{B}$). In addition to the actual data taken from *BABAR*, there are also computer simulations which model the production of particles from collisions as well as modeling their reconstruction and identification in the detector. The simulations allow us to better understand the data that *BABAR* physically takes. Thus, we wish to ensure that the simulation is as accurate as possible. It is the goal of this study to examine the production of neutral kaons with *BABAR* and to compare the simulated data with the actual data, look for discrepencies, and suggest corrections to the simulation.

Neutral kaons come in two physically observable states, which are symmetric or anti-symmetric superpositions of K^0 and $\bar{K}^0$. One state is long-lived, the other short-lived. The neutral kaons with the longer lifetime are called K-longs (K_L) and have a mean lifetime of 51.2 ns [2]. Those with the shorter lifetime are called K-shorts (K_S) and they have a mean lifetime of 89.6 ps [2]. It is the specific focus of this study to investigate the production rates and kinematic spectra of the K_S particle from decaying B mesons.

Since we are only interested in looking at K_S particles which come from B decays, events in the data with a $B\overline{B}$ pair in them are selected by "tagging" one of the B mesons. Having received data that contain only tagged events, we then want to obtain a clean sample of the K_S which are reconstructed through their decays into $\pi^+\pi^-$ pairs. This involves removing the noise generated by combinatorial errors in the reconstruction of the B and K_S decays, as well as removing the remaining noise from non-$B\overline{B}$ events. With this clean sample, we can then confidently study the rate of production and the spectra of K_S. Various points of interest are the momentum, and angular distributions, as well as the decay length and lifetime of K_S. Any discrepancies in these distributions between the computer simulation and the actual data taken are reflections of faults in the simulation. After correcting for the known shortcomings in the reconstruction efficiency of the K_S, the problems must then lie in the simulated production of the particles.

While the main goal of the study is to measure and understand the production of K_S and to improve the computer simulations, the findings have a broader applicability. K_L mesons are produced at the same rate as K_S and a study of K_S production thus allows us to draw direct conclusions in K_L production. K_L mesons escape the *BABAR* detector without completely depositing their energy and are rarely fully detected. This poses a serious problem, for instance, when trying to reconstruct B decays with a neutrino in the final state. Neutrinos leave no traces in the detector whatsoever and they can only be inferred from missing momentum and energy in the event. Since the K_L particles leave the detector carrying some energy with them, they obscure the amount of missing energy and momentum that would otherwise be attributed to the neutrinos. Thus, knowing how many K_S particles are in a sample will indicate how many K_L particles are expected as well, thereby allowing one to correct for their effects on the reconstruction of the neutrinos.

By performing this inclusive study (we do not care what other particles are producted) of K_S from B decays, we can improve our knowledge of K_S production and determine the extent to which the simulations and experiments agree. Adjustments can then be made to the K_S production rates and spectra in the simulation to bring the computer models into better agreement with the physical data.

MATERIALS AND METHODS

The $B\overline{B}$ *Data Sample*

For this analysis, we use data collected by the *BABAR* detector between 1996 and 2006 (runs 1-5) as well as data generated by Monte Carlo simulation[3]. Since we study K_S from B decays, the data set which we examine contains only those events in which one B-meson was fully reconstructed[4]. The B meson is reconstructed from purely hadronic decay modes: $B \rightarrow D^{(*)} + X$, where $X = n\pi + mK + rK_S^0 + q\pi^0$ with $n + m + r + d < 6$. More than 1000 such modes are used, however many do not yield a high purity for B's. To reduce the probability of poorly reconstructed events polluting our sample, we require purity greater than 70% and in the case of multiple tag-B candidates, we choose the one with the highest purity.

In order to be able to compare the amount of K_S from B decays produced in data with those in Monte Carlo, we must determine the number of $B\overline{B}$ events in the data and simulation. This can be done by fitting a special function to the reconstructed mass spectrum of the tagged B meson. The B mass is reconstructed as $m_{ES} = \sqrt{(\frac{\sqrt{s}}{2})^2 - \overrightarrow{p_B}^2}$, where the B energy has been substituted by $\frac{\sqrt{s}}{2}$ and $\sqrt{s}$ is the center-of-mass energy of the e^+e^- collision. The subscript "ES" stands for "energy substituted". The function that we use to fit the m_{ES} spectrum is the sum of an Argus and a Crystal Ball function:

$$f_{Bmass}(x) = f_{Argus}(x) + f_{CrystalBall}(x; \alpha, n, \bar{x}, \sigma),$$

where

$$f_{Argus}(x; S, m_0, c, p, \Delta x) = S \cdot (x - \Delta x) \cdot (1 - \frac{x}{m_0}^2)^p \cdot e^{c(1-\frac{x}{m_0}^2)},$$

and

$$f_{CrystalBall}(x; N, \alpha, n, \bar{x}, \sigma) = N \cdot \begin{cases} exp(-\frac{(x-\bar{x})^2}{2\sigma^2}) & \text{for } \left|\frac{x-\bar{x}}{\sigma}\right| < \alpha \\ A \cdot (B - \frac{x-\bar{x}}{\sigma})^{-n} & \text{for } \left|\frac{x-\bar{x}}{\sigma}\right| \geq \alpha \end{cases}$$

with

$$A = (\frac{n}{|\alpha|})^n \cdot exp(-\frac{|\alpha|^2}{2}) \quad \text{and} \quad B = \frac{n}{|\alpha|} - |\alpha|\,.$$

The Crystal Ball function[5] represents the events with a well-reconstructed tagged B meson. The Argus function[6] represents the remaining background from non-$B\overline{B}$ events (continuum events from processes $e^+e^- \rightarrow q\bar{q}$ where q=u,d,s,c) and to some extent the combinatorial background, which comes from $B\overline{B}$ events where particles are incorrectly matched together to produce something that looks like a B meson. By summing them together we get a function which models both the well-reconstructed $B\overline{B}$ and background events.

Figure 1 shows the result of the m_{ES} fits for data and Monte Carlo. The χ^2 per degree of freedom for the data nd MC fits are 17.1 and 48.8 respectively. Since the ideal case would have a χ^2 of 1, our goodness of fit in these cases is poor. It is known that in order to improve the χ^2, for example, an additional contribution to the fit function is necessary to more accurately model the combinatorial background. However, this complicated fit is out of the scope of this study and the accuracy achieved by the present fit is sufficient for this analysis. To determine the total number of events with a well-reconstructed B ($N_{B\overline{B}}$), we define a signal region in m_{ES}: $5.275 < m_{ES} < 5.284$ MeV/c^2. $N_{B\overline{B}}$ can then be determined in two different ways:

- $N_{B\overline{B}} = \int_{SignalRegion} CrystalBall$

– by integration of the Crystal Ball function over the m_{ES} signal region.

- $N_{B\overline{B}} = \int_{SignalRegion} m_{ES} Spectrum - \int_{SignalRegion} Argus$
  - by subtracting from the number of events in the signal region of the m_{ES} spectrum the integral of the Argus function over the m_{ES} signal region.

We choose to use the first method for our analysis and will later assign the difference between the two methods as systematic error on $N_{B\overline{B}}$.

Because there are more well-reconstructed events in the simulation than in the data, due to the fact that more events have been generated in the simulation, it will be necessary when comparing K_S spectra, and to determine the K_S production rate in data and MC, to scale the simulation histograms down to match the size of the data sample for a meaningful comparison. This scale factor can be found by taking the ratio of the determines number of $B\overline{B}$ events in both samples. The number of $B\overline{B}$ events and the scale factor are given in Table 1.

K_S *Selection*

The K_S candidates are reconstructed in the decay channel $K_S \to \pi^+\pi^-$. This decay channel has a large branching fraction of $(68.95 \pm 0.14)\%$ and is experimentally easy to reconstruct. These two pions from the K_S decay are reconstructed from charged tracks in the detector that do not come from the reconstructed B which we tagged earlier. To ensure that our sample of K_S is as clean as possible we apply additional selection criteria using the following variables:

- d_{3D}: The flight length is the distance from the primary vertex (e^+e^- interaction point) to the vertex where the pions are found to have originated from. This is how far the K_S particle travelled in 3 dimensions. In calculating d_{3D}, we can ignore the flight length of the B meson from the primary vertex because it is on the order of 500μm and thus nominal compared to the K_S flight length.

- Insist that $d_{3D} > 2$mm.

- α_{xy}: The angle between the reconstructed momentum vector of the K_S and the flight length vector in the xy-plane.
  - Insist that $\mathrm{Cos}(\alpha_{xy}) > 0.9992$.
- $VtxProb$: The vertex probability is the probability that the vertex the pions originate from is correctly reconstructed, or more precisely, the χ^2 probability of the vertex fit.
  - Insist that $VtxProb > 10^{-6}$
- $m_{\pi\pi}$: The reconstructed $\pi^+\pi^-$ invariant mass.
  - This cut is discussed in the following section.

Figure 2 shows how cutting on the first three quantities changes the $m_{\pi\pi}$ distribution.

With this sample we can begin to make comparisons between K_S production rates and spectra for data and simulation by inspecting the following six kinematic variables for the K_S:

- p^*: The momentum of the K_S in the rest frame of the $\Upsilon(4S)$.
- $Cos(\theta)$: The cosine of the angle (θ) between the K_S direction and the electron beam as seen from the $\Upsilon(4S)$ rest frame.
- d_{3D}: the 3-dimensional distance between the interaction point and the decay vertex in the lab frame.
- d_{xy}: the 2-dimensional distance between the interaction point and the decay vertex in the lab frame.
- The time that it takes for the K_S to decay in the lab frame, $\tau = \frac{d_{3D} \cdot m}{p_{lab}}$, where:
  - m = the mass of the K_S (0.4976 GeV/c^2)[2];

– p_{lab} = the momentum of the K_S in the lab frame;

- The corrected K_S decay time, $\Delta\tau = (d_{3D} - d_{min}) \cdot \frac{m}{p_{lab}}$ where $d_{min} = 2$ mm.

Since we want to study the production of K_S in data and MC, we need to correct for known differences in the reconstruction efficiencies for $K_S \to \pi\pi$ decays between data and MC. The efficiency correction factors (recommended by the *BABAR* collaboration[7]) in Table 2 are applied as weights to the reconstructed K_S candidates in simulation. Upon the spectra of these variables after correction for reconstruction efficiencies, we perform two background subtractions: one from the background in the B mass distribution and the other from the background in the K_S mass distribution.

K_S *Background Subtraction*

First we perform a sideband subtraction based on the reconstructed $\pi^+\pi^-$ invariant mass, $m_{\pi\pi}$, to remove fake and wrongly reconstructed K_S candidates. We define a $m_{\pi\pi}$ signal region with a full-width of 8 MeV/c^2 which is centered around the mass peak as determined by the Particle Data Group[2]. The sidebands are chosen to lie equidistant from the peak and themselves each have a width of 40 MeV/c^2. A 10 MeV/c^2 gap is left on both sides between the signal and sideband regions so that any tails from the mass distribution of true K_S do not penetrate into the sidebands (Figure 3). The factor by which we weight the sidebands to subtract the contribution of fake K_S in the signal region is simply defined as:

$wtSide = \frac{SignalWidth}{TotalSideWidth} = 0.1.$

To perform the actual subtraction for any given kinematic variable, we take the spectrum in the signal region and subtract from it the spectrum of the sideband region multiplied by the scaling factor of $wtSide$. The kinematic variables before and after sideband subtraction are shown in Figure 4. All following distributions will be shown for the K_S signal region after sideband subtraction.

B *Background Subtraction*

Since we only study K_S from B decays, we need to remove the remaining non-$B\overline{B}$ background and the combinatorial background caused by wrongly reconstructed tag-B mesons. We again make a fit of the reconstructed tag-B mass distribution with a function comprised of an Argus and Crystal Ball sum. Unlike before where we fit the tagged B mass of the entire sample, here we use the sample after the K_S sideband subtraction. The range of the m_{ES} signal region was chosen earler and contains the peak of the $CrystalBall$ ($5.275 < m_{ES} < 5.284$ GeV/c^2). The sideband is chosen far away from the signal region where only the $Argus$ function contributes to the fit ($5.22 < m_{ES} < 5.25$ GeV/c^2). The result of the fits is shown in Figure 5. The scale factor for the sidebands is defined as:

$wt_{mES} = \frac{\int_{signalregion} Argus}{\int_{sidebandregion} Argus}$.

As with the previous sideband subtraction, we plot the signal region minus the sidebands that are scaled down by the factor wt_{mES} (Figure 6). This produces spectra that are representative of the spectra for well-reconstructed K_S from B decays. All future plots will employ this sideband subtraction as well as the K_S subtraction.

RESULTS

In order to compare the data and MC spectra, the MC spectra must first be scaled down according to the number of $B\overline{B}$ events found in the data. The scale factor for this was computed in Table 1 and found to be $wt_{MC} = 0.321314$. Figures 7 and 8 compare the spectra of the data against those of the scaled MC. We observe that the total production rate of K_S is overestimated in simulation by 17.2%. Please note that the flight length distributions in Figure 8 are cut off at 5 cm, which means that about 25% of the total rate is not shown. The discrepancy between data and MC is shown to be less in this region than for the total rate, we infer then that the rate discrepancy between data and MC gets larger for longer flight distances.

The ratios for both the regular and corrected decay times are relatively flat. The shape of the center-of-mass momentum spectra in the intermediate momentum range is rather similar in data and MC. However, both the high and low momentum regions show disagreement. The ratio of $Cos(\theta)$ is also relatively flat until $Cos(\theta) \rightarrow 1$, where MC shows a higher production than data.

We can calculate the efficiencies of the K_S selection as a function of the studied kinematic variables by taking the ratio of the reconstructed and fully subtracted MC spectra to the generated m_{ES} subtracted MCtruth spectra. The efficiency plots for the center-of-mass K_S momentum and polar angle are shown in Figure 9. The average efficiency is 0.249 (indicated as red line). From the efficiency spectra, we can calculate the efficiency-corrected spectra. These are found by dividing the reconstructed and fully subtracted data spectra by the efficiency spectra obtained from simulation. Figure 10 shows the efficiency-corrected data spectra for the center-of-mass K_S momentum and polar angle.

DISCUSSION AND CONCLUSIONS

We find the K_S production rate as the number of K_S detected divided by the number of $B\overline{B}$ events. The results are shown in Table 4, where the uncertainties quoted are statistical, taken from the error of the Crystal Ball fit. The sytematic uncertainty can be assigned as the difference between the yield calculated from the Crystal Ball fit and the yield calculated from the Argus-subtracted integral. I find this systematic uncertainty to be 0.03% absolute for data and less than 0.02% absolute for Monte Carlo. The table of relative yield in K_S production rate confirms what we noted earlier: that MC has a higher rate of production than data.

Besides the overall rate difference, the MC overestimates the K_S yield in data especially for low and high momentum regions. We divide the momentum spectra into three regions and find the average of the ratio of data and MC yields in each region (Table 5). We suggest to use these averages as correction factors in their respective regions to bring the MC momentum spectra into better agreement with the data.

Compared to a previous study in *BABAR* [8], this study uses a larger K_S data sample and thus has a better statistical precision and makes use of the latest improvements in the *BABAR* reconstruction software. Another improvement over the previous study is that known differences in the K_S reconstruction efficiency have properly been accounted for. The results obtained in this study will be useful for many future *BABAR* analyses that depend on a reliable simulation of K_S production.

ACKNOWLEDGMENTS

I would like to thank my mentor Jochen Dingfelder, as well as Wells Wulsin for all of their help and guidance. I would also like to thank Vera Luth for her guidance. I am also grateful to SULI and the DOE for funding my research.

REFERENCES

[1] B. Aubert *et al.* (*BABAR* Coll.), Nucl. Inst. Meth. **A479**, 1 (2002).

[2] W.-M. Yao *et al.* (Particle Data Group), J. Phys. G **33**, 1 (2006).

[3] S. Agostinelli *et al.* (GEANT4 Coll.), Nucl. Instrum. Methods Phys Res. Sect. A **506**, 250 (2003).

[4] D. del Re, S. Grancagnolo, R. Faccini, A. Sarti, G. Denardo, *Semi-Exclusive B reconstruction* BAD #271 *BABAR* Internal Analysis Document (2001).

[5] Crystal Ball Collaboration, T. Skwarnicki, Report No. DESY F31-86-02 (unpublished).

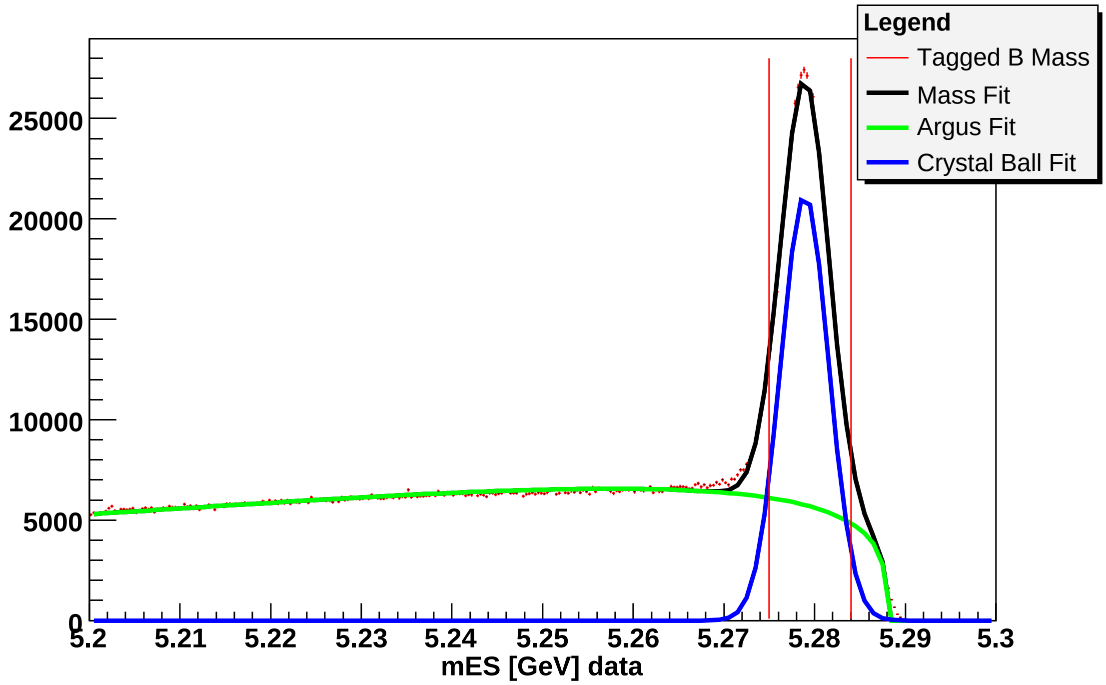
[6] ARGUS Collaboration, H. Albrecht *et al.*, Z. Phys. C **48**, 543 (1990).

[7] K_s Reconstruction Efficiency Study, http://www.slac.stanford.edu/BFROOT/www/Physics/TrackEff
2006.html

[8] W. Wulsin and J. Dingfelder, K_s *Production Spectra* BAD #1642 *BABAR* Internal Analysis Document (2006).

FIGURES

Figure 1: m_{ES} distribution in data and MC.
The red points with error bars indicate simulated events in data or MC. The vertical red lines indicate the chosen m_{ES} signal region.



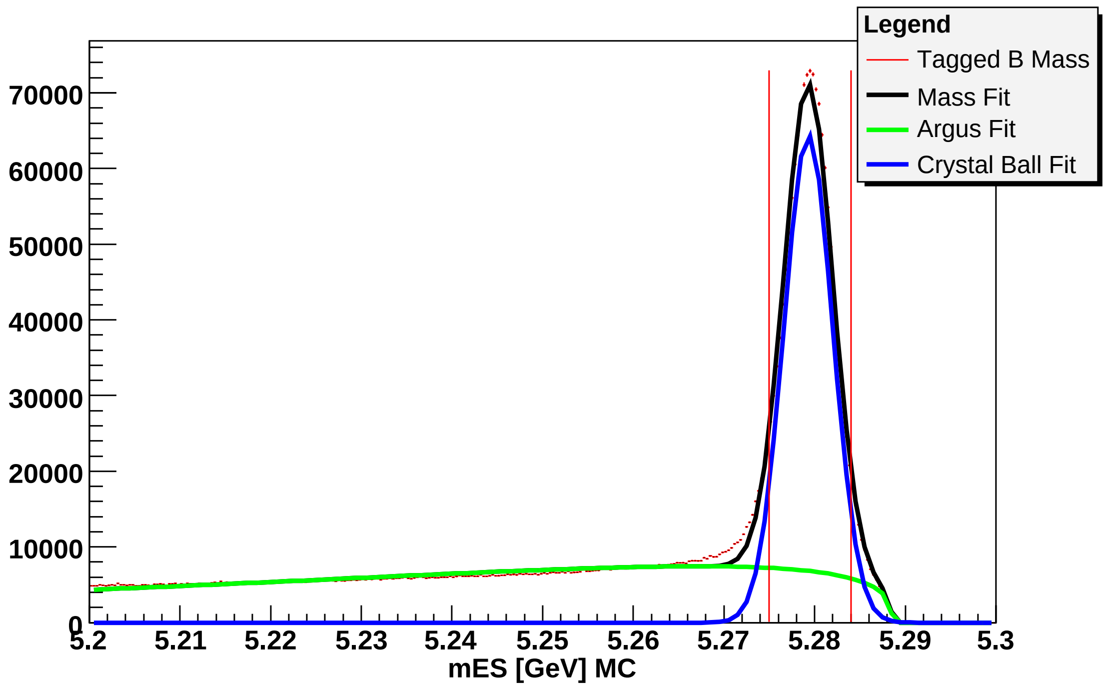


Figure 2: $M_{\pi\pi}$ distribution in MC demonstrating K_S cuts.
(top) d_{3D} cut only, (middle) d_{3D} and $VtxProb$ cuts, (bottom) all three cuts

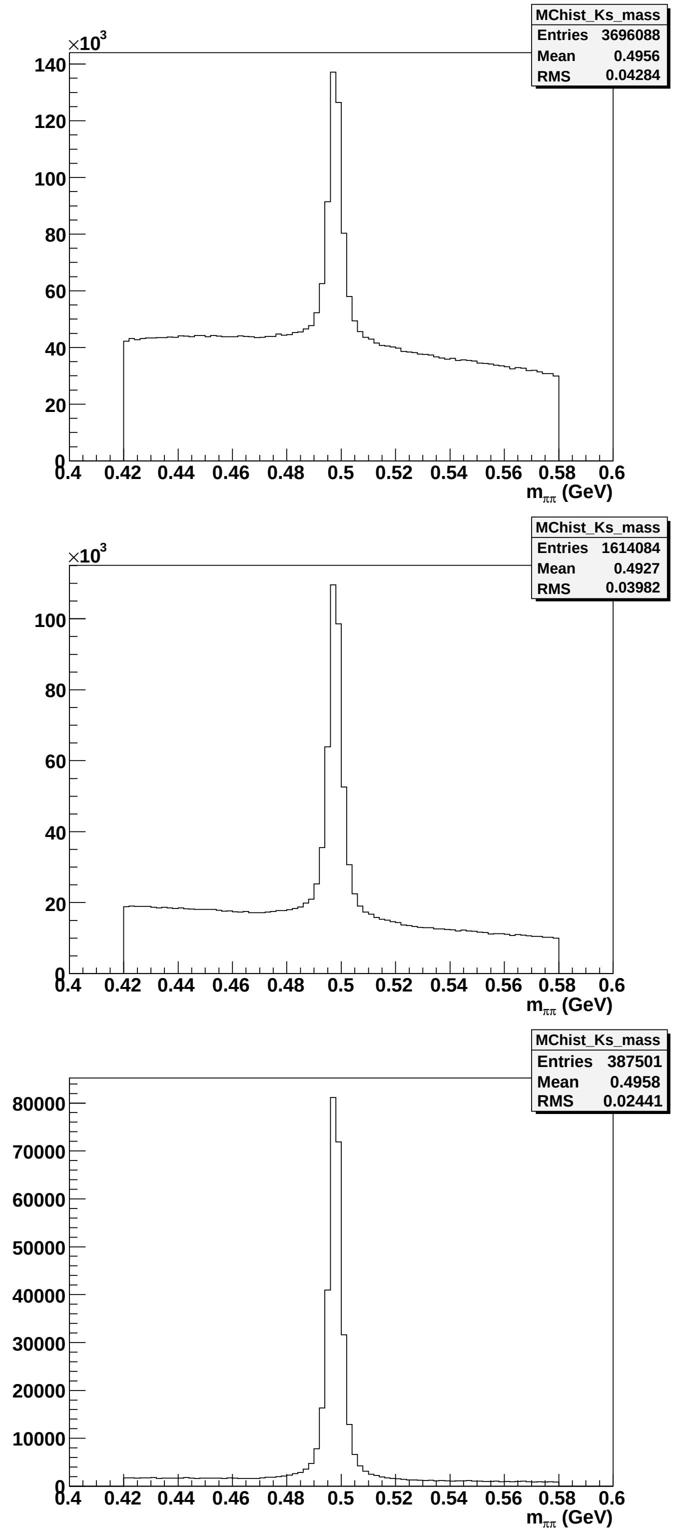
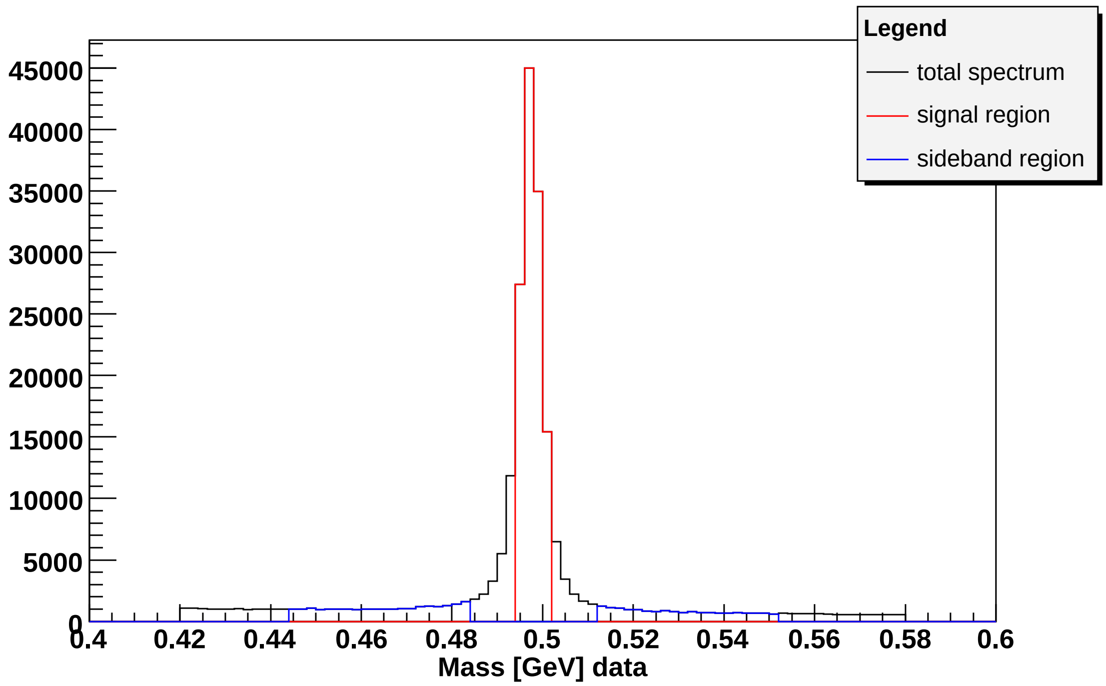


Figure 3: $M_{\pi\pi}$ Distribution In Data And MC.



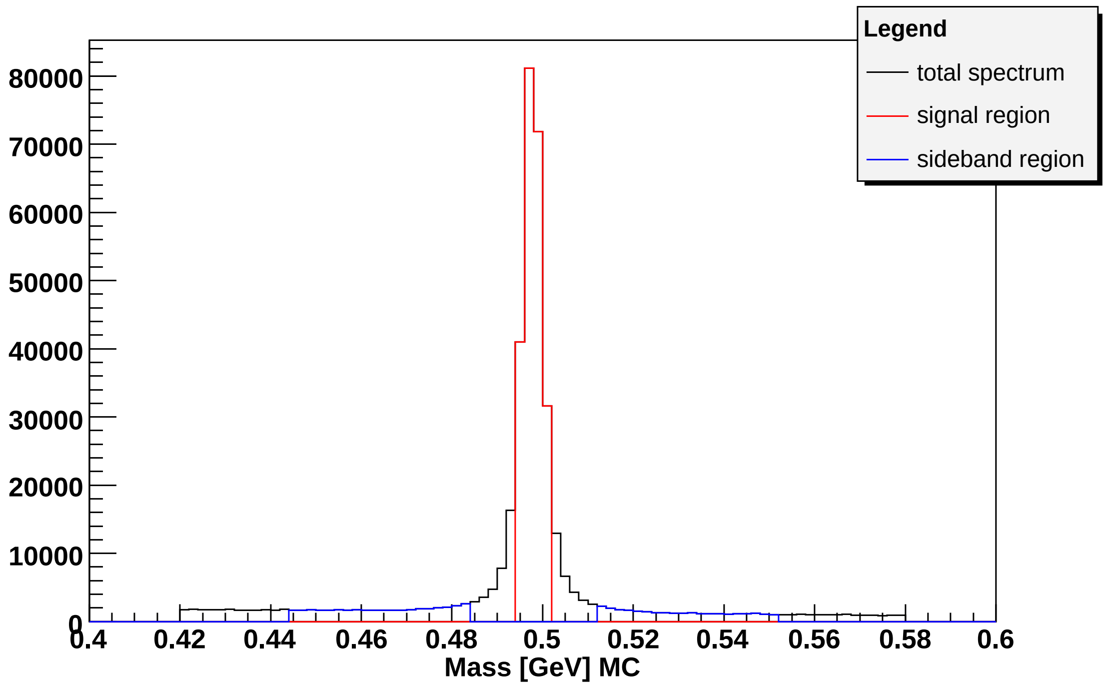


Figure 4: Kinematic Spectra demonstrating K_S sideband subtraction

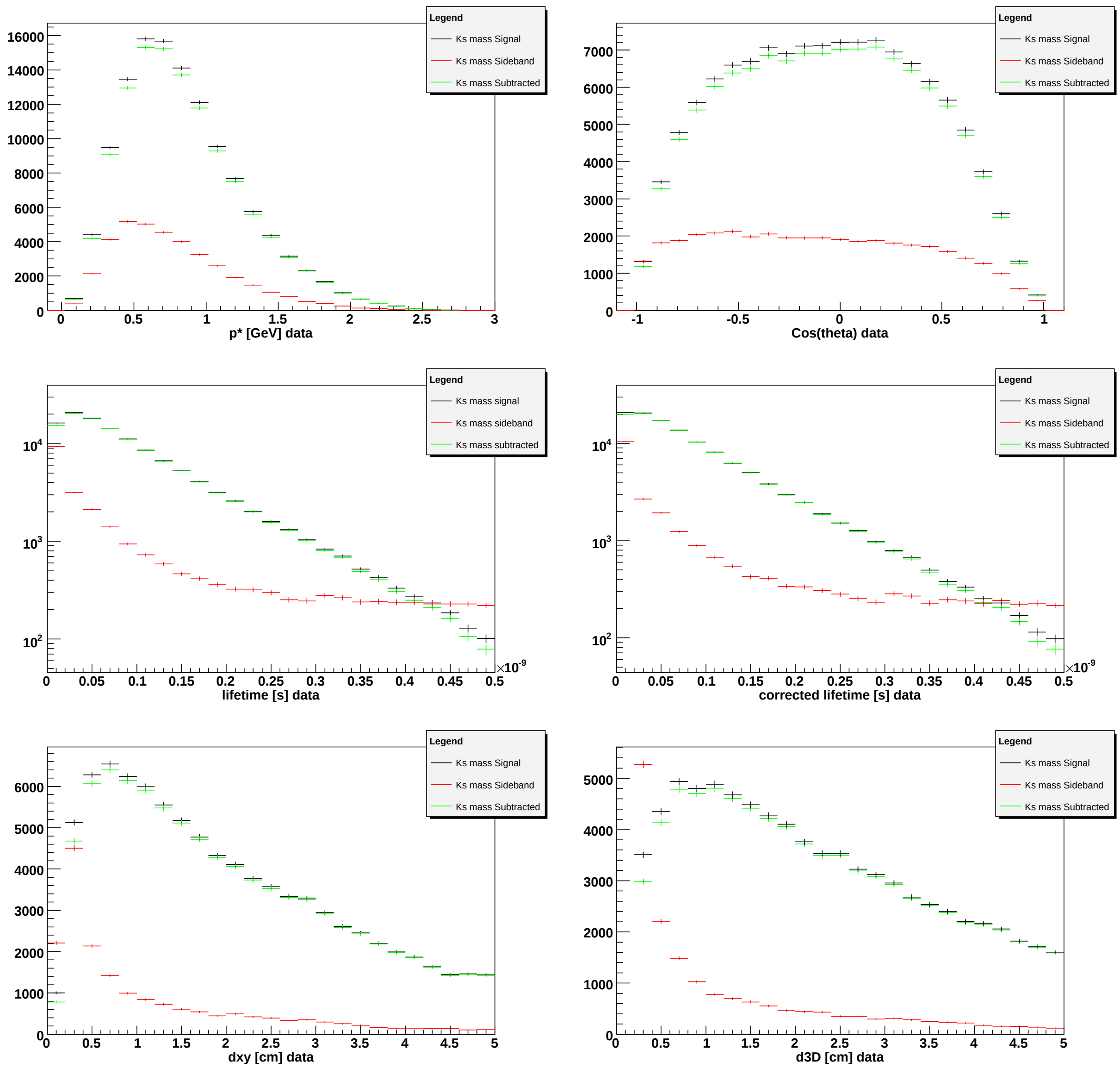
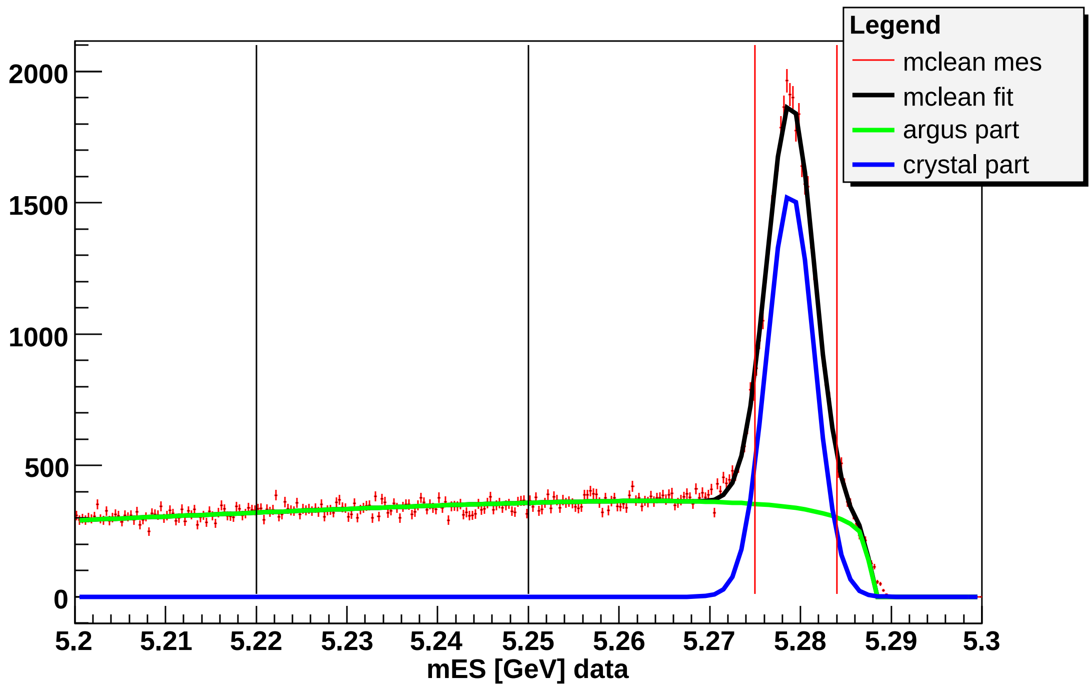


Figure 5: K_S sideband-subtracted m_{ES} distribution in data and MC
The red points with error bars indicate simulated events in data or MC. The vertical red lines indicate the chosen m_{ES} signal region. The vertical black lines indicate the chosen sideband region.



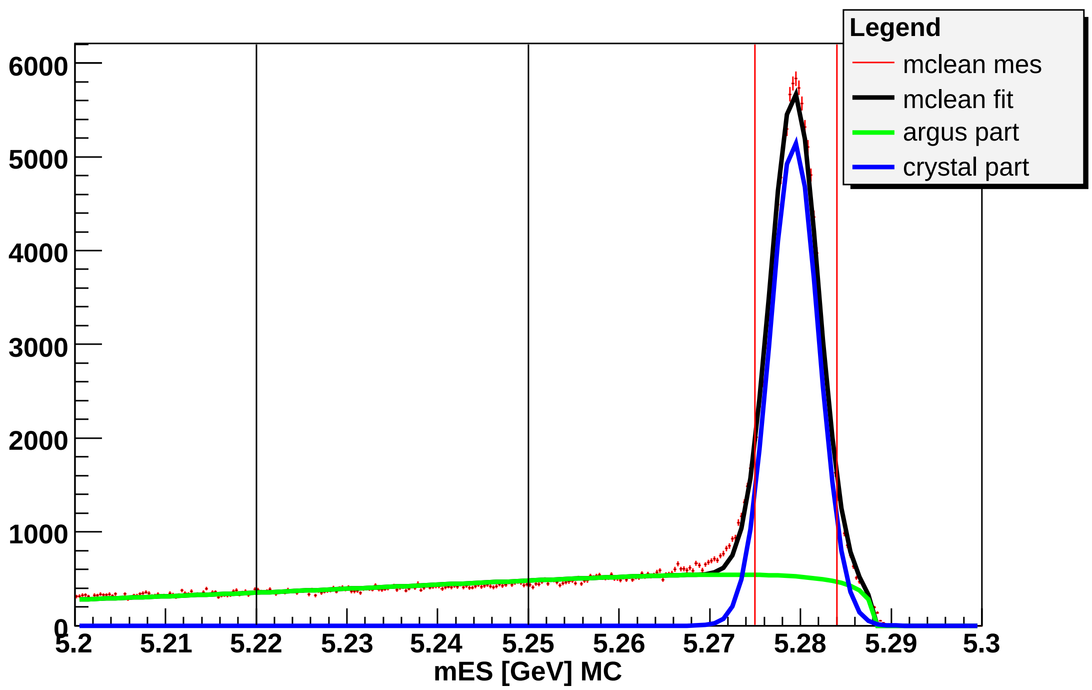


Figure 6: Kinematic Spectra demonstrating B sideband subtraction

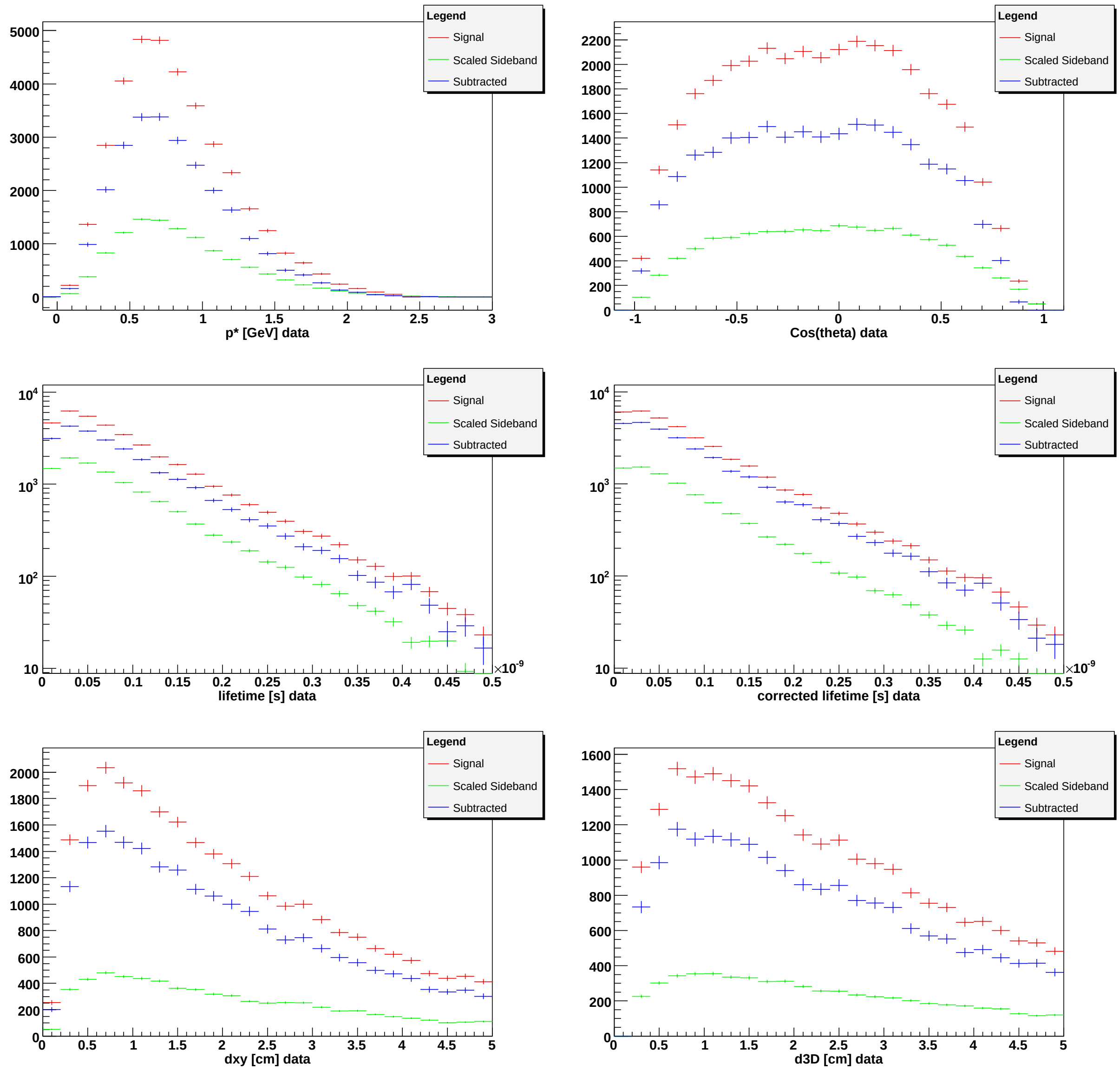


Figure 7: Comparing data and MC
Kinematic Spectra (left) and data over MC ratio plots (right)

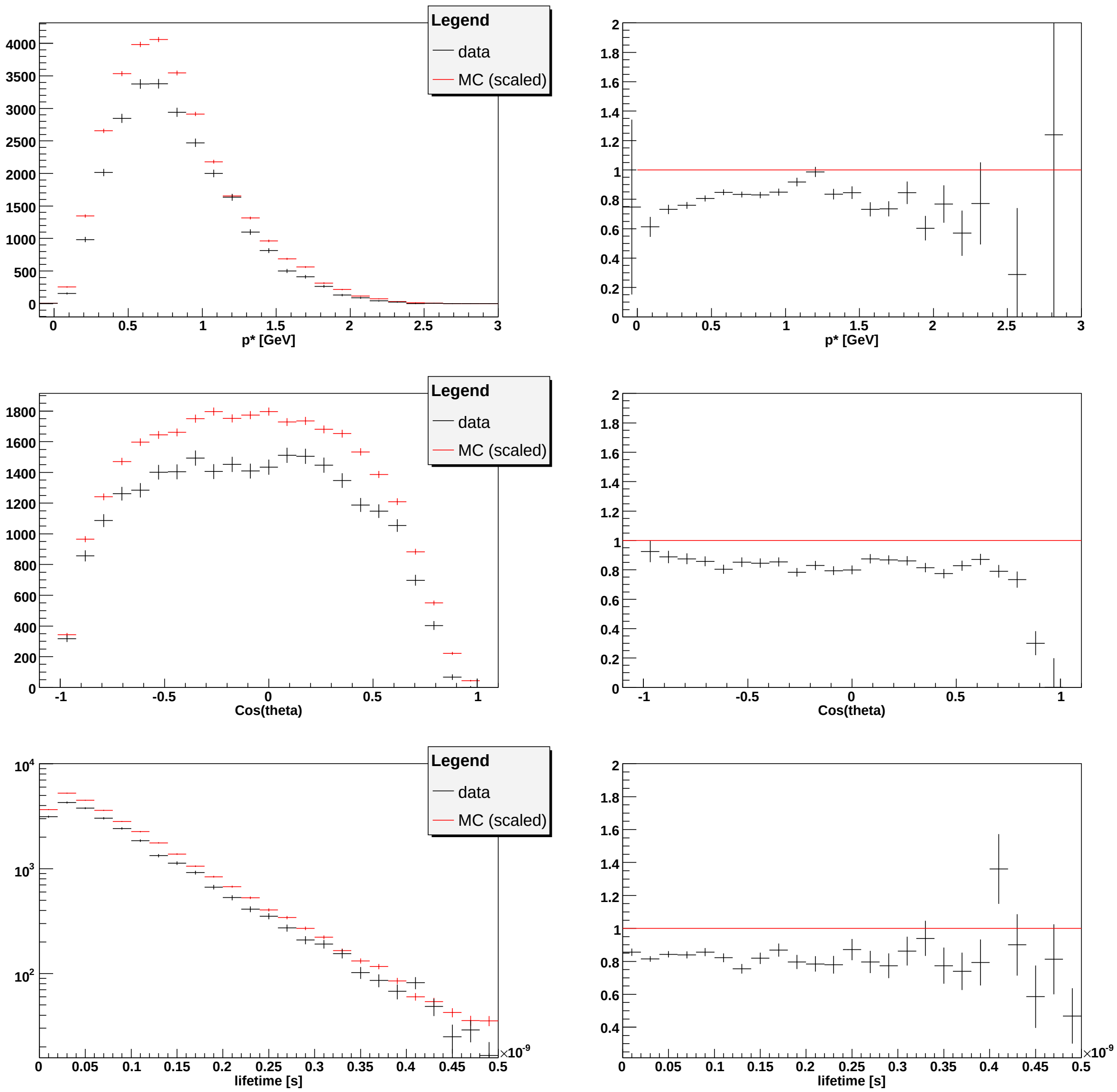


Figure 8: Comparing data and MC
Kinematic Spectra (left) and data over MC ratio plots (right)

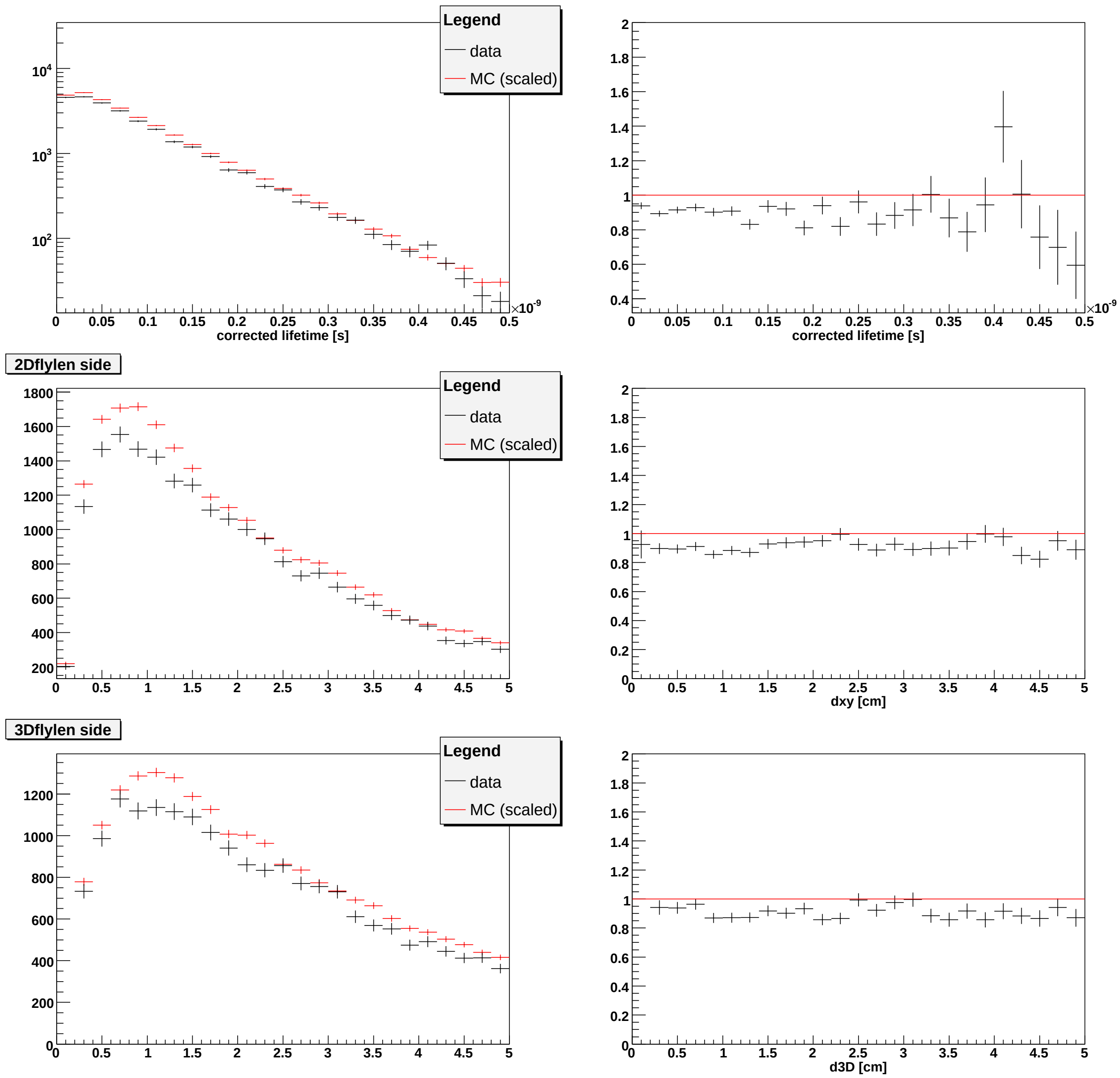
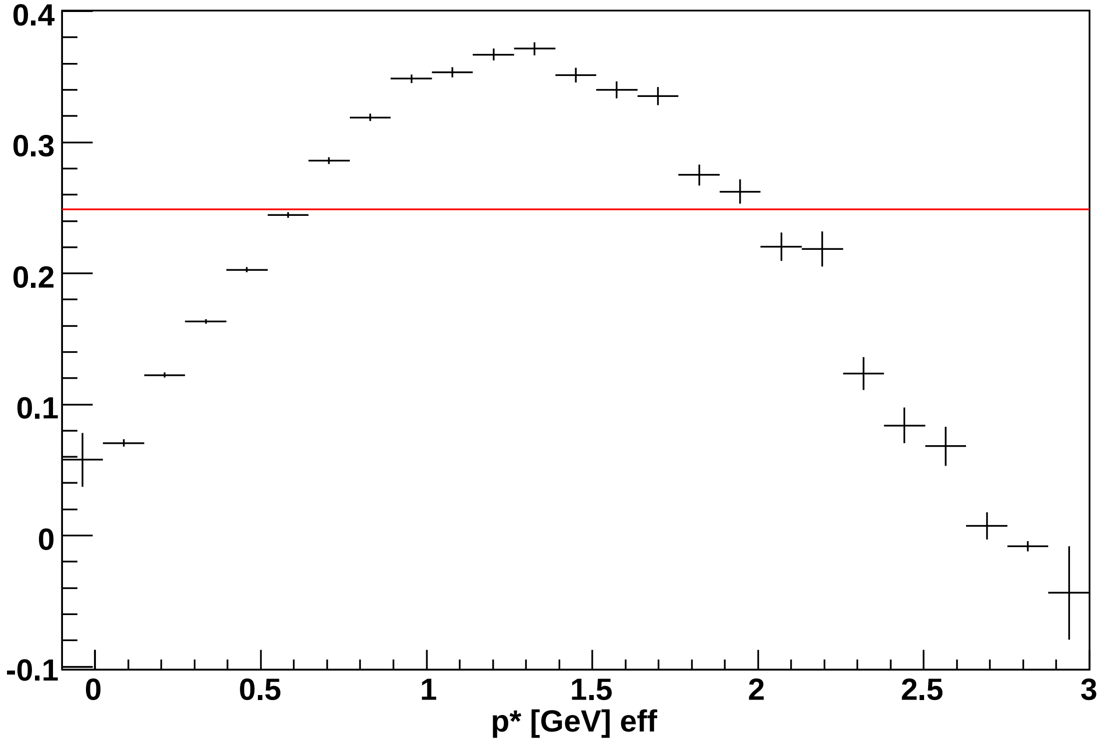


Figure 9: Efficiency Plots
Center-of-mass momentum (left) and cos(θ) (right)



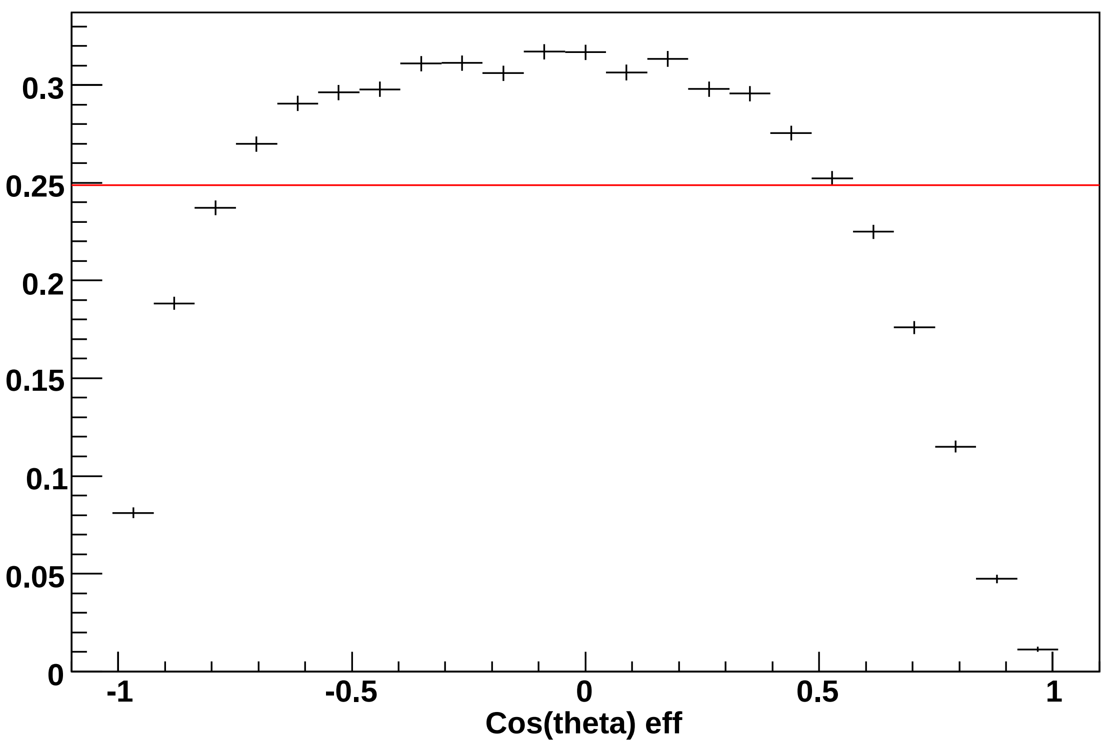
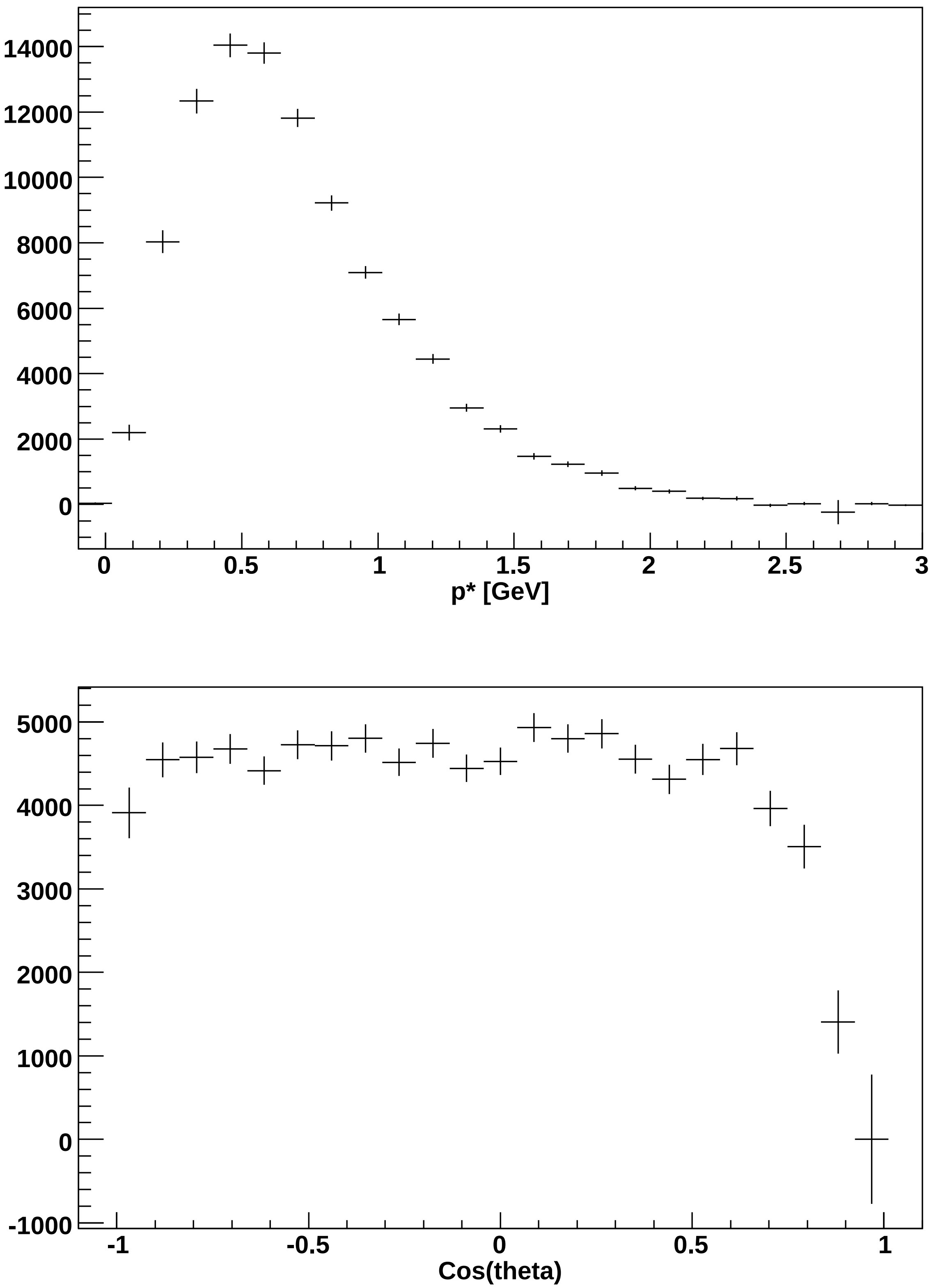


Figure 10: Efficiency-Corrected Plots
Center-of-mass momentum (left) and $\cos(\theta)$ (right)



TABLES

Table 1: Number of $B\overline{B}$ events before K_S selection

	A	B	C	D
$N_{B\overline{B}} =$	$\int_{m_{ES}sig} CrystalBall$	$\int_{m_{ES}sig} m_{ES}Spectrum$	$\int_{m_{ES}sig} Argus$	B - C
Data	382664	532716	152149	380567
MC	1190930	1363830	182549	1181281
wt_{MC} = D/MC	0.321314	–	–	0.322164

Table 2: Correction Factors for K_S reconstruction, applied to K_S in simulation

	Correction for K_S reconstruction efficiency	Correction for signal window cut	Total Correction (Product)
Run1	1.004 ± 0.025	1.000	1.004
Run2	1.009 ± 0.038	0.970	0.979
Run3	1.001 ± 0.018	0.974	0.975
Run4	0.983 ± 0.012	0.961	0.945
Run5	0.994 ± 0.011	0.950	0.944

Table 3: Number of K_S Candidates

	A	B	C	D = B/C
$N_{K_S} =$	$\int_{signal} CrystalBall$	$\int_{signal} Argus$	$\int_{sideband} Argus$	wt_{mES}
m_{ES} interval	$5.275 < m_{ES} < 5.284$	$5.275 < m_{ES} < 5.284$	$5.22 < m_{ES} < 5.25$	
Data	27669	9100	30848	0.294996
MC	94878	14069	37678	0.373399

Table 4: Relative yield of K_S from B decays

$N_{B\overline{B}} = \int_{m_{ES} sig}$	$CrystalBall$	$m_{ES}Spectrum - Argus$
Data	$(7.23 \pm .06)\%$	$(7.20 \pm 0.06)\%$
MC	$(7.97 \pm 0.04)\%$	$(7.95 \pm 0.04)\%$

Table 5: Correction Factors

Average Correction Factors for K_S MC production in three K_S momentum intervals.

Region of p^*	Yield ratio = data/MC
$0.0 < p^* < 0.4$ GeV/c^2	77.05%
$0.4 < p^* < 1.5$ GeV/c^2	85.13%
$1.5 < p^* < 3.0$ GeV/c^2	76.55%