

A search for the decay of a B meson into a kaon and a tau lepton pair at the $BABAR$ experiment

Racha Cheaib

Physics Department

McGill University, Montreal

March 15, 2016

A thesis submitted to McGill University in partial fulfillment of the
requirements for the degree of Doctor of Philosophy

©Racha Cheaib, 2016

Acknowledgements

I am grateful to the *BABAR* collaboration, physicists and engineers, for the quality and precision of the *BABAR* detector and *BABAR* dataset. I would like to thank my supervisor, Steven Robertson, for his continuous help through all stages of this analysis. Also, special thanks to the Leptonic Analysis Working Group and members of my review committee for their feedback and contribution. Finally, a special thanks to my family for their continuous support and encouragement. This thesis is dedicated to Ali and Fatima!

Abstract

The flavour changing neutral current (FCNC) process , $B^+ \rightarrow K^+ \tau^+ \tau^-$, is highly suppressed in the Standard Model (SM). This decay is forbidden at tree level and only occurs at lowest order via one-loop diagrams. $B^+ \rightarrow K^+ \tau^+ \tau^-$ thus has the potential to provide a stringent test of the SM and a fertile ground for new physics searches. Contributions due to virtual particles in the loop allow one to probe, at relatively low energies, new physics at large mass scales. We search for the rare FCNC process $B^+ \rightarrow K^+ \tau^+ \tau^-$ using data collected by the *BABAR* detector at the SLAC National Accelerator Laboratory. The *BABAR* data sample corresponds to a total integrated luminosity, at the energy of the $\Upsilon(4S)$ resonance, of 424.4 fb^{-1} and 471 million $B\bar{B}$ pairs. For this search, hadronic B_{tag} reconstruction is employed, where one B is exclusively reconstructed via one of many possible hadronic modes. The remaining decay products in an event are then attributed to the signal B , on which the search for $B^+ \rightarrow K^+ \tau^+ \tau^-$ is performed. Each τ is required to decay leptonically, into either an electron or a muon and the lepton neutrinos. Furthermore, a multi-variate analysis technique (neural network) is used to select for signal events and suppress dominant background modes. No significant signal is observed. The resulting branching fraction is measured to be $\mathcal{B}(B^+ \rightarrow K^+ \tau^+ \tau^-) = 1.31_{-0.61}^{+0.66}(\text{stat.})_{-0.25}^{+0.35}(\text{sys.}) \times 10^{-3}$, which is consistent with zero at the 1.9σ level, with an upper limit of 2.25×10^{-3} at the 90% confidence level.

Résumé

Le processus de courant neutre avec changement de saveur (“flavour changing neutral current” ou FCNC) $B^+ \rightarrow K^+ \tau^+ \tau^-$ est fortement réduit dans le modèle standard (SM). Cette désintégration n’est pas permise au niveau des diagrammes en arbre et se produit seulement à l’ordre le plus bas via les diagrammes à une boucle. Le processus $B^+ \rightarrow K^+ \tau^+ \tau^-$ fournit donc un test rigoureux du SM et un terrain fertile pour la quête d’une nouvelle physique. Des contributions dues aux particules virtuelles présentes dans la boucle permettent de sonder, à des énergies relativement basses, une nouvelle physique à des échelles de masse élevées. Nous recherchons le processus FCNC rare $B^+ \rightarrow K^+ \tau^+ \tau^-$ dans les données récoltées par le détecteur *BABAR* au SLAC National Accelerator Laboratory. L’échantillon de données de *BABAR* correspond à une luminosité intégrée totale, pour l’énergie de résonance du $\Upsilon(4S)$, de 424.4 fb^{-1} et à 471 millions de paires $B\bar{B}$. Pour cette recherche, nous utilisons la reconstruction hadronique B_{tag} , où un B est reconstruit exclusivement via les modes de désintégration hadroniques. Les produits de désintégration restants d’un événement sont attribués au signal du méson B , pour lequel la recherche de $B^+ \rightarrow K^+ \tau^+ \tau^-$ est effectuée. On demande à ce que chaque τ se désintègre de façon leptonique, soit un électron ou un muon, accompagnés les neutrinos leptoniques associés. En outre, on utilise une technique d’analyse multivariable (réseau neuronal) pour sélectionner des événements de signal et supprimer les modes dominants en bruit de fond. Nous n’observons pas de signal significatif. Le rapport de branchement résultant mesuré est de $\mathcal{B}(B^+ \rightarrow K^+ \tau^+ \tau^-) = 1.31_{-0.61}^{+0.66}(\text{stat.})_{-0.25}^{+0.35}(\text{sys.}) \times 10^{-3}$, ce qui est consistant avec zéro pour un niveau de signification statistique de 1.9σ , avec une limite supérieure de 2.25×10^{-3} .

Preface

This thesis is submitted for the degree of Doctor of Philosophy in experimental high energy physics. The dataset used in this thesis is the property of the international *BABAR* collaboration. This thesis also uses software tools developed by the *BABAR* collaboration, such as the skimming code for the hadronic B_{tag} reconstruction. The data and MC ntuples available at McGill University are also used in this thesis and have been produced by members of the *BABAR* group at McGill, specifically Dr. Steven Robertson.

The $B^+ \rightarrow K^+ \tau^+ \tau^-$ analysis described herein is, to the best of the author's knowledge, original work. The signal selection, background estimate, validation and closure tests are developed and implemented by the author of the thesis. No part of this thesis has been previously published.

Contents

1	Introduction	1
2	Theory	3
2.1	The Standard Model	3
2.1.1	Leptons and quarks	4
2.1.2	Gauge bosons	5
2.2	SuperSymmetry and the Two-Higgs-Doublet model	10
2.3	$B^+ \rightarrow K^+ \tau^+ \tau^-$: significance and motivation	11
2.3.1	$B^+ \rightarrow K^+ \tau^+ \tau^-$ in the Standard Model	12
2.3.2	$B^+ \rightarrow K^+ \tau^+ \tau^-$ beyond the Standard Model	18
2.3.3	Related measurements: $B \rightarrow K^{(*)} \ell^+ \ell^-$	23
3	The <i>BABAR</i> experiment	28
3.1	PEP-II	28
3.2	The <i>BABAR</i> detector	29
3.2.1	Silicon vertex tracker	32
3.2.2	Drift chamber	33
3.2.3	Detector of internally reflected Cherenkov light	36
3.2.4	Electromagnetic calorimeter	38
3.2.5	Instrumented flux return	41
3.2.6	Trigger system	43

3.3	Event reconstruction	44
4	Analysis Tools	46
4.1	Hadronic \mathbf{B}_{tag} reconstruction method	46
4.2	Analysis software	48
4.3	Event reconstruction	49
4.4	<i>BABAR</i> dataset	51
4.5	Background Monte Carlo	51
4.6	Signal Monte Carlo	55
5	Signal Selection	57
5.1	\mathbf{B}_{tag} cuts	60
5.1.1	m_{ES} cut	60
5.1.2	Purity cut	61
5.1.3	Continuum background suppression	63
5.2	B_{sig} cuts	68
5.2.1	E_{miss}	68
5.2.2	Track multiplicity and PID	69
5.2.3	π^0 veto	71
5.2.4	γ , J/ψ , and D^0 vetos	72
5.2.5	s_B cut	73
5.3	Peaking background	74
5.4	Multilayer perceptron neural network	75
5.4.1	Angular discriminating variables	76
5.4.2	Kinematic discriminating variables	79
5.4.3	Calorimeter discriminating variables	80
5.4.4	MLP neural network output	83
6	Uncorrected signal and background MC yields	88

7	Background estimation	93
7.1	m_{ES} sideband substitution	93
7.2	Combinatorial background estimate	101
7.3	Peaking background estimate	106
7.4	Discriminating variables after m_{ES} sideband substitution	109
7.5	$B^+ \rightarrow D^0 \ell^+ \bar{\nu}_\ell$ study: MLP neural network validation	112
8	Branching Fraction Calculation	118
8.1	Signal reweighting	119
8.2	Final ϵ_{sig} and N_{bkg}	120
8.3	Branching fraction results	121
8.4	Branching fraction closure test	127
9	Systematic Uncertainties	132
9.1	B_{tag} yield correction	133
9.2	Continuum Likelihood Cut	138
9.3	Theoretical Uncertainty	139
9.4	Particle Identification cuts	142
9.5	π^0 veto	147
9.6	s_B cut	150
9.7	MLP cut	150
9.8	Summary of systematics	153
10	Results	155
10.1	Discussion	159
11	Conclusion	168
	Appendices	viii

A	Discriminating variables	ix
B	Region above D meson in $M_{K^+l^-}$ distribution	xv
C	Kaon Momentum after MLP cut	xvii
D	Combinatorial B^+B^- and $B^0\bar{B}^0$ components in m_{ES} distribution	xix
E	Discriminating Variables of Validation Test	xxi
E.1	Discriminating variables before m_{ES} sideband substitution	xxi
E.2	Discriminating variables after m_{ES} sideband substitution	xxv
F	Comment on correlations	xxviii
G	Discriminating Variables before and after MLP cut	xxxii
G.1	Discriminating variables before the final MLP cut	xxxii
G.2	Discriminating variables after the final MLP cut	xxxvii

List of Figures

2.1	The Cabbibo-Kobayashi-Maskawa matrix and its different components. . . .	7
2.2	One of the 3 unitary triangles that result from the CKM matrix [29].	8
2.3	Differential branching fraction for $B \rightarrow X_s \ell^+ \ell^-$ for $l = \tau$ (solid and dashed) and $l = e$ (dotted and dash-dotted) with and without the long-distance contributions [11].	15
2.4	τ polarization asymmetry as a function of \hat{s} , with (black line) and without (dashed curve) long-distance contributions [11].	17
2.5	Sensitivity of the τ polarization asymmetry to the Wilson coefficients as a function of \hat{s} . From top to bottom, the asymmetry is shown if $C_{10}(M_W) \rightarrow -C_{10}(M_W)$ (dashed, top curve), $C_{9,10}(M_W) \rightarrow -C_{9,10}(M_W)$ (dash-dotted), $C_9(M_W) \rightarrow -C_9(M_W)$ (dotted), and $C_{7,8}(M_W) \rightarrow -C_{7,8}(M_W)$ (long-dashed, bottom curve). The solid distribution is the SM [11].	18
2.6	New physics contributions in $b \rightarrow s$ transitions: a) charged Higgs (H^-), b) squark and chargino, c) gluino and neutralino.	19
2.7	Additional Feynman diagrams due to the exchange of the neutral Higgs bosons in 2HDM. The wavy lines represent the propagators of charged bosons W^\pm and H^\pm , while the dashed lines represent the neutral bosons H^0, h^0 and A^0 [45].	20
2.8	$\mathcal{B}(B \rightarrow K \tau^+ \tau^-)$ with long distance effects for each of mass set 1, 2, and 3, respectively. Curve 4 is the Standard Model prediction. Top-left: $\tan \beta = 1$, top-right: $\tan \beta = 20$, bottom $\tan \beta = 30$ [45].	20

2.9	Longitudinal τ polarization asymmetry as a function of p^2 without long distance effects for each of mass set 1, 2, and 3, respectively. Curve 4 is the Standard Model prediction. Top-left: $\tan \beta = 1$, top-right: $\tan \beta = 30$, bottom $\tan \beta = 50$ [45].	21
2.10	Transverse τ polarization asymmetry as a function of p^2 without long distance effects for each of mass set 1, 2, and 3, respectively. Top-left: $\tan \beta = 1$, top-right: $\tan \beta = 30$, bottom $\tan \beta = 50$ [45].	22
2.11	Total branching fraction for $B \rightarrow K \ell^+ \ell^-$ and $B \rightarrow K^* \ell^+ \ell^-$, as measured by <i>BABAR</i> , CDF, and Belle and compared with predictions from Ali <i>et al.</i> and Zhong <i>et al.</i> models [4].	24
2.12	Partial branching fraction results of <i>BABAR</i> [4], CDF [6], Belle [5] and LHCb [58] [59] for $B \rightarrow K \ell^+ \ell^-$ (top) and $B \rightarrow K^* \ell^+ \ell^-$ (bottom) in bins of $s \equiv m_{\ell^+ \ell^-}^2$. The magenta lines show the SM theoretical predictions from Ali <i>et al.</i> model [55]. The vertical yellow bands are the vetoed regions of the J/ψ and $\psi(2S)$ resonances.	25
2.13	Standard Model branching fraction with form factors calculated using unquenched lattice QCD[57], compared to experimental results from <i>BABAR</i> [4], CDF [6], Belle [5], and LHCb [58] [59]. The vertical yellow bands are the vetoed regions of the J/ψ and $\psi(2S)$ resonances.	26
2.14	A_{CP} (left) and $R_{K^{(*)}}$ (right) for $B \rightarrow K \ell^+ \ell^-$ and $B \rightarrow K^* \ell^+ \ell^-$ modes as a function of s . The vertical yellow bands are the vetoed regions of the J/ψ and $\psi(2S)$ resonances [4].	27
2.15	$A_I^{K^{(*)}}$ for $B \rightarrow K \ell^+ \ell^-$ and $B \rightarrow K^* \ell^+ \ell^-$ modes as a function of s . The vertical yellow bands are the vetoed regions of the J/ψ and $\psi(2S)$ resonances [4]. . .	27
3.1	A schematic diagram of the SLAC linac and the PEP-II storage ring [64]. . .	29
3.2	The <i>BABAR</i> detector: longitudinal cross section [22].	31
3.3	The end view of the <i>BABAR</i> detector [22].	31

3.4	Longitudinal view of the SVT, where the roman numerals stand for the different types of the silicon detectors [65].	33
3.5	Schematic layout of drift cells for the four innermost layers [67].	34
3.6	Side view of the <i>BABAR</i> drift chamber. Dimensions are shown in mm [66]. . .	35
3.7	Measurement of dE/dx in the DCH as a function of the track momenta. The curves show the Bethe-Bloch predictions determined from selected control samples [67].	36
3.8	Imaging principle and the transport of photon signals in the DIRC [67]. . . .	37
3.9	Cherenkov angle as a function of the particle's momentum [68].	38
3.10	Longitudinal cross section of the EMC (top half is shown here) showing the arrangement of the 56 crystal rings. Dimensions are given in mm [67].	39
3.11	Schematic of a CsI crystal with its readout electronics [67].	40
3.12	Energy resolution of the EMC as a function of energy, determined using various processes. The fit is given by Eq. (3.1) $a = 2.32 \pm 0.30\% \sqrt{\text{GeV}}$ and $b = 1.85 \pm 0.12\%$ [67].	41
3.13	Diagram of the different sections of the IFR (barrel and end doors) and the shape of its modules [67].	41
3.14	Cross section of a planar Resistive Plate Chamber [67].	43
4.1	Hadronic B_{tag} reconstruction approach.	47
4.2	m_{ES} distribution of B_{tag} candidates in data (points), B^+B^- (brown) and $B^0\bar{B}^0$ + continuum background (yellow), with signal and sideband regions highlighted. This distribution contains events with a <i>charged</i> B_{tag} candidate and three signal-side tracks. These events have also passed the purity and continuum likelihood cut, which will be discussed in detail in the upcoming sections.	53

5.1	m_{ES} distribution after sub-skin level cuts. Signal MC (cocktail) distribution is shown in red. The data distribution (black points) is overlaid on the background MC distributions (color-filled). $q\bar{q}$ refers to $q = u, d, s$	61
5.2	The purity value, determined for each event depending on the corresponding B_{tag} decay mode. Signal MC (<i>generic</i>) distribution is shown in red. The data distribution (black points) is overlaid on the background MC distributions (color-filled).	62
5.3	Schematic diagram of a $B\bar{B}$ (left) and continuum (right) event.	63
5.4	Event shape variables used to calculate the continuum likelihood ratio. Signal MC (<i>generic</i>) distribution is shown in red. The data distribution (black points) is overlaid on the background MC distributions (color-filled). $q\bar{q}$ refers to $q = u, d, s$	66
5.5	Output of the continuum multivariate likelihood for generic $B^+ \rightarrow K^+ \tau^+ \tau^-$ signal, background MC, and data. Signal MC (<i>generic</i>) distribution is shown in red. The data distribution (black points) is overlaid on the background MC distributions (color-filled). $q\bar{q}$ refers to $q = u, d, s$	67
5.6	Invariant mass distributions of signal-side track combinations. Signal MC (<i>cocktail</i>) distribution is shown in red. The data distribution (black points) is overlaid on the background MC distributions (color-filled). $q\bar{q}$ refers to $q = u, d, s$	73
5.7	s_B distribution after all above cuts for data, background and signal Monte Carlo. Signal MC (<i>cocktail</i>) distribution is shown in red. The data distribution (black points) is overlaid on the background MC distributions (color-filled). $q\bar{q}$ refers to $q = u, d, s$	74
5.8	Schematic diagram of the angle between the two leptons in the di-tau frame, $\cos \theta_{l+l-}$, for a signal (left) and background (right) event.	76

5.9	Schematic diagram of the angle between the kaon and the oppositely charged lepton in the di-tau frame, $\cos \theta_{K+l^-}$, for a signal (left) and background (right) event.	77
5.10	Schematic diagram of the angle between B_{sig} and the oppositely charged lepton in the CM frame, $\cos \theta_{B+l^-}$, for a signal (left) and background (right) event.	78
5.11	Angular variables used in the MLP neural network: angle between two leptons in di-tau frame (top-left), angle between K and oppositely charged lepton in di-tau frame (top-right), angle between B_{sig} and oppositely charged lepton in CM frame (bottom-left), and angle between the K recoil vector and the lepton with low momentum in CM frame (bottom-right). Signal MC (<i>cocktail</i>) distribution is shown in red. The data distribution (black points) is overlaid on the background MC distributions (color-filled). $q\bar{q}$ refers to $q = u, d, s$	79
5.12	Momentum of lepton, with charge opposite to that of the kaon, in the di-tau frame. Signal MC (<i>cocktail</i>) distribution is shown in red. The data distribution (black points) is overlaid on the background MC distributions (color-filled). $q\bar{q}$ refers to $q = u, d, s$	80
5.13	Calorimeter variables used in the MLP neural network. Signal MC (<i>cocktail</i>) distribution is shown in red. The data distribution (black points) is overlaid on the background MC distributions (color-filled). $q\bar{q}$ refers to $q = u, d, s$	83
5.14	MLP neural network output for electron, muon, electron-muon and all three modes combined.	86
5.15	Barlow upper limit as a function of MLP cut, calculated assuming $N_{\text{obs}} \approx N_{\text{bkg}}$.	87
6.1	Signal efficiency as a function of s_B for all 3 modes combined.	92
7.1	m_{ES} distribution after applying the sub-skim, continuum likelihood, purity and track multiplicity cuts. The data distribution (black points) is overlaid on the background MC distributions (color-filled). $q\bar{q}$ refers to $q = u, d, s$	94

7.2	E_{extra} distribution after applying sub-skim, m_{ES} , continuum likelihood, purity and track multiplicity cuts and before m_{ES} sideband substitution. The background distribution is shown filled with blue, while the data is overlaid as black dots.	95
7.3	Combinatorial background estimate (yellow) of the E_{extra} distribution calculated by scaling the sideband data with R_{comb} . The full E_{extra} distribution in the m_{ES} signal region is shown (green) to highlight the fraction of combinatorial events.	97
7.4	m_{ES} distribution of peaking data (black) vs. peaking B^+B^- MC (blue) after applying sub-skim, continuum, purity, and track multiplicity cuts.	98
7.5	E_{extra} distribution of data (black) vs. background MC (blue), where the background MC consists of the combinatorial contribution, determined using sideband data, and a peaking B^+B^- component without C_{yield} correction. . . .	99
7.6	E_{extra} distribution after the m_{ES} sideband substitution and peaking background correction. The agreement between data (black) and background MC (blue) is much improved.	99
7.7	Ratio of data and MC yield in the E_{extra} distribution after the m_{ES} sideband substitution.	100
7.8	R_{comb} as a function of cut in the signal selection.	102
7.9	m_{ES} distribution of $B^0\bar{B}^0$ background after various cuts in the signal selection. .	103
7.10	m_{ES} distribution of $c\bar{c}$ background after various cuts in the signal selection. .	104
7.11	C_{yield} as a function of cut.	107
7.12	m_{ES} distribution of B^+B^- background after various cuts in the signal selection.	108
7.13	Angular variables used in the MLP neural network, after applying the s_B cut and m_{ES} sideband substitution.	110
7.14	Calorimeter variables used in the MLP neural network, after applying the s_B cut and m_{ES} sideband substitution.	111

7.15	Lepton Momentum in the di-tau frame, used in the MLP neural network, after applying the s_B cut and m_{ES} sideband substitution.	112
7.16	B^+B^- MC (red) vs. data (points with error bars) after applying the control sample selection.	113
7.17	MLP neural network output for both the B^+B^- (red) and the data control sample (points with error bars).	114
7.18	MLP neural network output for total MC (blue) and data (points with error bars).	114
7.19	Event yields as a function of MLP cut.	115
7.20	Ratio of the event yields as a function of MLP cut.	116
7.21	R_{comb} and C_{yield} as a function of cut flow in control sample selection.	117
8.1	Predicted Standard Model contribution to dB_τ/dq^2 calculated using unquenched lattice QCD approach [57].	119
8.2	Reweighted generic signal MC distribution to the unquenched lattice QCD result. The reweighted signal MC is scaled down by $\sim 10\%$ to distinguish the different curves.	120
8.3	Branching fraction as a function of N_{obs} for each of the electron (top-left), muon (top-right), and electron-muon (bottom) modes.	124
8.4	m_{ES} distribution of <i>mock</i> data sample. The <i>mock</i> data sample (black points) is overlaid on the background MC distributions (color-filled). $q\bar{q}$ refers to $q = u, d, s$	128
8.5	R_{comb} (left) and C_{yield} (right) distributions using the <i>mock</i> data sample. . .	128
8.6	MLP output of <i>mock</i> data sample. Signal MC (<i>generic</i>) distribution is shown in red (left) and is appended to the <i>mock</i> data sample (right). The <i>mock</i> data sample (black points) is overlaid on the background MC distributions (color-filled). $q\bar{q}$ refers to $q = u, d, s$	130

9.1	Gaussian distribution resulting from the random fluctuation of R_{comb} with a Gaussian generator of $\mu = 0.21$ and $\sigma = 0.018$	134
9.2	Resulting C_{yield} determined using the anti-correlation with R_{comb}	135
9.3	2D plot of R_{comb} and C_{yield} displaying a linear anti-correlation.	135
9.4	Fit to $\Delta\epsilon_{sig}$ after fluctuating R_{comb} with a Gaussian random generator and determining the anti-correlated C_{yield}	136
9.5	Continuum likelihood ratio, after applying the s_B cut and before the MLP cut, before (left) and after (right) the m_{ES} sideband substitution.	139
9.6	s_B distribution of cocktail signal MC after reweighting to the unquenched lattice QCD calculation. The signal MC before and after reweighting is shown in red and blue respectively.	140
9.7	s_B distribution resulting from the dispersion quark model, with and without long-distance (peaking) contributions from the $\psi(2S)$ resonance [90].	141
9.8	Cocktail signal s_B distribution after reweighting to the dispersion quark model. The signal MC before and after reweighting is shown in red and blue, respectively.	141
9.9	Performance plot for the kaon BDT tight selector as a function of the lab frame momentum . Data and MC efficiency for positive (left) and negative (center) kaons. The ratio of the data to MC efficiency $\epsilon_{data}/\epsilon_{MC}$ is shown in the right hand plot[22].	143
9.10	Performance plot for the electron KM tight selector as a function of the lab frame momentum. Data and MC efficiency for positive (left) and negative (center) electronss. The ratio of the data to MC efficiency $\epsilon_{data}/\epsilon_{MC}$ is shown in the right hand plot [22].	144

9.11	Performance plot for the muon BDT loose selector as a function of the lab frame momentum. Data and MC efficiency for positive (left) and negative (center) muons. The ratio of the data to MC efficiency $\epsilon_{data}/\epsilon_{MC}$ is shown in the right hand plot [22].	144
9.12	Lab frame momentum of kaon in all 3 signal modes, after applying kaon PID.	145
9.13	Lab frame momentum of electron in the Electron mode, after applying the electron PID cut.	145
9.14	Lab frame momentum of muon in the Muon mode, after applying the muon PID cut.	146
9.15	Lab frame momentum of electron in the Electron-Muon mode, after applying PID cuts.	146
9.16	Lab frame momentum of muon in the Electron-Muon mode, after applying PID cuts.	147
9.17	Background MC π^0 mass fit.	148
9.18	Fit of π^0 mass distribution in data.	149
9.19	MLP output distribution in sideband region before (left) and after (right) m_{ES} sideband substitution.	151
9.20	MLP output distribution for data (points) and total MC (red) after applying the $B^+ \rightarrow D^0 l \nu_\ell$, $D^0 \rightarrow K^+ \pi^-$ selection.	152
9.21	Ratio of data to total MC yields, in each bin of the MLP output distribution.	152
9.22	Fit to the ratio of data yield to total MC yield in each bin of the MLP output distribution.	153
10.1	MLP output distributions with unblinded data and signal MC.	157
10.2	MLP output distributions with unblinded data after m_{ES} sideband substitution.	158
10.3	Ratio of data to MC yield for each bin of the MLP output distribution. . . .	160
10.4	m_{ES} distribution of each mode after the MLP neural network cut.	161
10.5	R_{comb} (left) and C_{yield} (right) for each mode as a function of cut.	162

10.6	Peaking data vs. peaking MC after MLP neural network cut.	163
10.7	Angular variables of the electron-muon mode, used in the MLP neural network, after applying the final cut in the analysis.	165
10.8	Calorimeter variables of the electron-muon mode, used in the MLP neural network, after applying the final cut in the analysis.	166
10.9	s_B distribution (left) and kinematic variable of the electron-muon mode (right), after applying the final cut in the analysis. The latter is used in the MLP neural network.	167
A.1	Kaon Momentum in CM frame for cocktail (red) and generic (blue) signal MC samples.	x
A.2	E_{extra} for cocktail (red) and generic (blue) signal MC samples.	x
A.3	Invariant mass of kaon with oppositely charge lepton in lab frame for cocktail (red) and generic (blue) signal MC samples.	xi
A.4	E_{miss} in di-tau frame for cocktail (red) and generic (blue) signal MC samples.	xi
A.5	Momentum of lepton with charge opposite to kaon for cocktail (red) and generic (blue) signal MC samples.	xii
A.6	$\cos \theta_{\ell+\ell-}$ in di-tau frame for cocktail (red) and generic (blue) signal MC samples.	xii
A.7	$\cos \theta_{K+l-}$ in di-tau frame for cocktail (red) and generic (blue) signal MC samples.	xiii
A.8	$\cos \theta_{B+l-}$ in CM frame for cocktail (red) and generic (blue) signal MC samples.	xiii
A.9	$\cos \theta_{(B_{sig}-K)l_{low}}$ in CM frame for cocktail (red) and generic (blue) signal MC samples.	xiv
B.1	MLP output for each of the electron (top-left), muon (top-right) and electron-muon (bottom) modes, when the MLP neural network is trained and tested in the region $M_{K+\ell-} < 1.8 \text{ GeV}/c^2$	xvi
C.1	Kaon Momentum in CM frame (left) and s_B distribution (right) after MLP cut for both signal and background MC.	xviii

D.1	Combinatorial B^+B^- (blue) and $B^0\bar{B}^0$ (red) distribution in the continuum likelihood sideband region after requiring a neutral and charged B_{tag} candidate, respectively.	xx
E.1	Calorimeter variables used in the MLP neural network validation test. The data distribution (black points) is overlaid on the background MC distributions (color-filled). $q\bar{q}$ refers to $q = u, d, s$	xxii
E.2	Momentum of lepton, with charge opposite to that of the kaon, in the di-tau frame used in the MLP neural network validation test. The data distribution (black points) is overlaid on the background MC distributions (color-filled). $q\bar{q}$ refers to $q = u, d, s$	xxiii
E.3	Angular variables used in the MLP neural network validation test: angle between two leptons in di-tau frame (top-left), angle between Kaon and oppositely charged lepton in di-tau frame (top-right), angle between B_{sig} and oppositely charged lepton in CM frame(bottom-left), and angle between the Kaon recoil vector and the lepton with low momentum in CM frame(bottom-right). The data distribution (black points) is overlaid on the background MC distributions (color-filled). $q\bar{q}$ refers to $q = u, d, s$	xxiv
E.4	Calorimeter variables used in the MLP neural network validation test. The data distribution (black points) is overlaid on the corrected B^+B^- MC + sideband data (color-filled). $q\bar{q}$ refers to $q = u, d, s$	xxv
E.5	Momentum of lepton, with charge opposite to that of the kaon, in the di-tau frame used in the MLP neural network validation test. The data distribution (black points) is overlaid on the corrected B^+B^- background + sideband data (color-filled). $q\bar{q}$ refers to $q = u, d, s$	xxvi

E.6	Angular variables used in the MLP neural network validation test: angle between two leptons in di-tau frame (top-left), angle between Kaon and oppositely charged lepton in di-tau frame (top-right), angle between B_{sig} and oppositely charged lepton in CM frame(bottom-left), and angle between the Kaon recoil vector and the lepton with low momentum in CM frame(bottom-right). The data distribution (black points) is overlaid on the corrected B^+B^- background + sideband data (color-filled). $q\bar{q}$ refers to $q = u, d, s$	xxvii
F.1	Given $N_{\text{sig}} = 4$, the smear in ΔBF , when N_{bkg} is fluctuated using a Gaussian random generator with $\mu = N_{\text{bkg}}$ and $\sigma = 8.26\% \times N_{\text{bkg}}$. Here, $\Delta BF = BF_{\text{trial}} - BF_{\text{actual}}$, where BF_{trial} is the trial central value given a fluctuation in N_{bkg} and BF_{actual} is the actual value of the branching fraction for $N_{\text{sig}} = 4$	xxx
F.2	Given $N_{\text{sig}} = 4$, the smear in ΔBF , when ϵ_{sig} is fluctuated using a Gaussian random generator with $\mu = \epsilon_{\text{sig}}$ and $\sigma = 8.26\% \times \epsilon_{\text{sig}}$	xxx
F.3	Given $N_{\text{sig}} = 4$, the smear in ΔBF , when N_{bkg} and ϵ_{sig} are fluctuated.	xxx
F.4	σ as a function N_{sig} when N_{bkg} is fluctuated.	xxxi
F.5	σ as a function N_{sig} when ϵ_{sig} is fluctuated.	xxxi
F.6	σ as a function N_{sig} when both N_{bkg} and ϵ_{sig} are fluctuated.	xxxi
G.1	Angular variables used in the MLP neural network before applying the final cut in the analysis, before (left) and after (right) m_{ES} sideband substitution.	xxxiii
G.2	Angular variables used in the MLP neural network before applying the final cut in the analysis, before(left) and after (right) m_{ES} sideband substitution.	xxxiv
G.3	Calorimeter variables used in the MLP neural network before applying the final cut in the analysis, before (left) and after (right) m_{ES} sideband substitution.	xxxv
G.4	Kinematic variables used in the MLP neural network before applying the final cut in the analysis, before (left) and after(right) m_{ES} sideband substitution.	xxxvi

G.5	Angular variables used in the MLP neural network after applying the final cut in the analysis, before (left) and after (right) m_{ES} sideband substitution. . .	xxxviii
G.6	Angular variables used in the MLP neural network after applying the final cut in the analysis, before(left) and after (right) m_{ES} sideband substitution. . . .	xxxix
G.7	Calorimeter variables used in the MLP neural network after applying the final cut in the analysis, before (left) and after (right) m_{ES} sideband substitution.	xl
G.8	Kinematic variables used in the MLP neural network after applying the final cut in the analysis, before (left) and after (right) m_{ES} sideband substitution.	xli

List of Tables

2.1	List of quarks and leptons with their intrinsic properties [18].	5
2.2	Integrated branching fractions for $l = e, \mu$ and τ in the total and high dilepton mass regions, where \hat{s} is the normalized q^2 and $x \equiv m_\tau^2/m_b^2$ [11].	15
2.3	List of mass values for each of the Higgs particles used to determine the branching fraction and τ polarization plots in 2HDM [45].	19
4.1	Luminosity and B -counting values of the <i>BABAR</i> dataset [75].	51
4.2	Generic Background Monte Carlo Information	54
4.3	Number of generated and skimmed events for each run of each type of the generic background Monte Carlo.	54
4.4	Number of generated and skimmed events for each signal Monte Carlo sample used in this analysis.	56
5.1	List of PID selectors used to identify the signal-side tracks.	69
5.2	Value of MLP cut applied for each mode.	85
6.1	Number of raw events in $B \rightarrow K^\pm \tau^+ \tau^-$ generic and cocktail signal Monte Carlo after each selection cut.	89
6.2	Number of normalized events in background Monte Carlo after each selection cut.	89
6.3	Efficiencies in $B \rightarrow K^\pm \tau^+ \tau^-$ generic and cocktail signal Monte Carlo after each selection cut.	90

6.4	Normalized efficiencies $\times 10^{-2}$ (%) in background Monte Carlo and data after each selection cut.	90
6.5	Partial efficiencies (%) in $B \rightarrow K^\pm \tau^+ \tau^-$ generic and cocktail signal Monte Carlo after each selection cut.	91
6.6	Partial efficiencies (%) in background Monte Carlo and data after each selection cut.	91
8.1	The final values for $B \rightarrow K^+ \tau^+ \tau^-$ with statistical and systematic uncertainties.	120
8.2	$\mathcal{B}(B^+ \rightarrow K^+ \tau^+ \tau^-, \tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau, \tau^- \rightarrow e^- \bar{\nu}_e \nu_\tau)$ values with upper and lower Barlow and Feldmann-Cousins limits.	125
8.3	$\mathcal{B}(B^+ \rightarrow K^+ \tau^+ \tau^-, \tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau, \tau^- \rightarrow \mu^- \bar{\nu}_\mu \nu_\tau)$ values with upper and lower Barlow and Feldmann-Cousins limits.	125
8.4	$\mathcal{B}(B^+ \rightarrow K^+ \tau^+ \tau^-, \tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau, \tau^- \rightarrow \mu^- \bar{\nu}_\mu \nu_\tau)$ values with upper and lower Barlow and Feldmann-Cousins limits.	126
8.5	Result of branching fraction closure test with <i>mock</i> data + signal sample. . .	131
9.1	Values of N , N_\pm , and δ_\pm used to calculate the B_{tag} yield systematic uncertainty on the background estimate.	138
9.2	List of PID systematic uncertainties for each charge of each particle type, calculated as a function of the particle's momentum in a specific signal mode.	143
9.3	Systematic uncertainty due to the π^0 veto, evaluated by fluctuating the π^0 mass window by $\pm\delta$	149
9.4	Summary of systematics uncertainties on the signal efficiency and background estimate (except for the uncertainty due to theoretical model) common for all 3 signal channels.	154
9.5	Summary of lepton PID systematics for each of the 3 signal channels. These uncertainties are applied on the signal efficiency and background estimate. .	154

10.1	Efficiency (ϵ_{sig}), background estimate (N_{bkg}), and number of observed events (N_{obs}) for each of the electron, muon, and electron-muon modes. ϵ_{sig} and N_{bkg} include statistical and systematic errors, while N_{obs} has a statistical error only.	156
10.2	$\mathcal{B}(B^+ \rightarrow K^+ \tau^+ \tau^-)$ central value, 90% upper and lower limits for each mode separately determined using the Barlow method.	156
B.1	The final values for $B \rightarrow K^+ \tau^+ \tau^-$	xvi

Chapter 1

Introduction

The current understanding of the basic constituents of matter in our universe is embedded in the Standard Model (SM). This model has succeeded as a low energy effective theory and its predictions are in agreement with numerous experimental tests [1]. Yet, even with its high level of consistency with experimental measurements, many questions are left unanswered. For instance, while the discovery of the Higgs boson [2] is consistent with the SM, the mass of this boson is much lighter than the Planck scale, which leads to the hierarchy problem [3] and the need for a more complete theory of particle physics. Currently, various experiments worldwide are working towards unprecedented high energies and luminosities in order to break the boundaries and discover new physics.

The aim of this thesis is to push forward the ongoing hunt for physics beyond the Standard Model. The flavor-changing neutral current process, $B^+ \rightarrow K^+ \tau^+ \tau^-$, is highly suppressed in the SM and only occurs, at lowest order, via one-loop diagrams. Virtual particles can enter in the loop and thus deviate the branching fraction from its SM expectation. $B^+ \rightarrow K^+ \tau^+ \tau^-$ is therefore a stringent test of the SM and a promising window to what may lie beyond. Searches for $B^+ \rightarrow K^+ \ell^+ \ell^-$ where $\ell = e$ or μ have been previously performed by *BABAR* [4] and other collider experiments [5]-[7]. Some of the results were found to be consistent with SM theoretical predictions, while others exhibit tension [8, 9]. $B^+ \rightarrow K^+ \tau^+ \tau^-$ is the third

generation equivalent of $B \rightarrow K^{(*)}\ell^+\ell^-$ and thus investigation of this mode offers exciting possibilities. Its SM branching fraction is comparable to its electron and muon counterparts [11] and the presence of a third generation lepton increases the sensitivity to specific new physics scenarios, such as couplings to the neutral Higgs boson in the two-Higgs-doublet model [10, 11, 12]. Furthermore, the recent measurement of $B \rightarrow D^{(*)}\tau\nu_\tau$ by *BABAR* [13], Belle [14] and LHCb [15] showed a significant deviation from the SM expectations. This suggests that third generation leptons may hold the key to the new physics we are looking for.

The goal of this thesis is to measure the $B^+ \rightarrow K^+ \tau^+\tau^-$ branching fraction and discover a signal yield in excess of the SM prediction. If a signal yield is not observed, an upper limit on the branching fraction will be determined. Before discussing the details of this search, Chapter 2 of the thesis introduces the Standard Model and its possible extensions. An overview of the *BABAR* detector and the PEP-II storage ring is provided in Chapter 3. Chapter 4 lists the analysis tools used for this search, including data and Monte Carlo samples as well as the hadronic B_{tag} reconstruction method. The signal selection is described in detail in Chapter 5, and the resulting efficiencies are listed in Chapter 6. Furthermore, the expected background estimate and the branching fraction calculation are presented in Chapters 7 and 8, respectively. Finally, the systematic uncertainties are examined in Chapter 9, and the final results are discussed in Chapter 10.

Chapter 2

Theory

Over the past few decades, the Standard Model (SM) has succeeded in predicting and explaining a wide range of experimental results to the limit of available experimental precision. After the discovery of the Higgs boson, the theory is considered a successful low-energy effective theory as each of its particles has been found. However, there are many shortcomings to the SM. For instance, the current theory does not explain the baryonic asymmetry of the universe or the hierarchy problem[3]. It also does not contain any viable dark matter candidates and is incompatible with general relativity. Nevertheless, experimental evidence of physics beyond the SM is yet to be found. Below is a brief description of the Standard Model and its most relevant extensions.

2.1 The Standard Model

In any quantum field theory, the Lagrangian defines the dynamics of the theory, while reflecting the internal symmetries of the system. The Standard Model is based on quantum field theory with the gauge symmetry $SU(3)_C \times SU(2)_L \times U(1)_Y$. This gauge group includes the symmetry associated with the strong force and quantum chromodynamics, $SU(3)_C$. The second part represents the weak and electromagnetic forces and their symmetry groups $SU(2)_L \times U(1)_Y$. The latter two forces are said to be unified into the electroweak force [16].

Each kind of particle in the Standard Model is described in terms of a dynamical field, while the forces are represented by gauge fields which act on the dynamical fields [17]. Thus, in the Standard Model, there are two classes of particles: the fermions which are the building blocks of matter and consist of spin-1/2 quarks and leptons, and the spin-1 gauge bosons which mediate the interactions between the leptons and quarks. These are discussed in greater detail below.

2.1.1 Leptons and quarks

The basic components of matter are spin-1/2 fermions which obey the Pauli exclusion principle [17]. These fermions are divided into two groups according to the forces with which they interact: the quarks and the leptons. Furthermore, there are six types of leptons and quarks, called flavours, divided into three generations. The leptons are the electron, e^- , the muon, μ^- , and the tau, τ^- , with an electromagnetic (EM) charge of $Q=-1$ and their corresponding neutrinos: ν_e , ν_μ , and ν_τ , with an EM charge $Q=0$. On the other hand, the six flavours of quarks are : u , d , c , s , b and t and carry a fractional charge of either $Q=+\frac{2}{3}$ or $Q=-\frac{1}{3}$. These particles and their intrinsic properties are summarized in Table 2.1.

The three generations of quarks and leptons are usually written as $SU(2)_Y$ doublets and are listed below from the first (left) to the third (right) generation:

$$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix} \quad \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix} \quad \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix} \quad (2.1)$$

$$\begin{pmatrix} u \\ d \end{pmatrix} \quad \begin{pmatrix} c \\ s \end{pmatrix} \quad \begin{pmatrix} t \\ b \end{pmatrix} \quad (2.2)$$

Each pair of particles in a doublet has the same properties, except for mass and charge. Furthermore, except for the neutrinos, each member of a generation has a greater mass than the previous one, and due to baryon number conservation, the first generation is stable.

Leptons do not interact via the strong force, while quarks interact via all fundamental forces of the SM (weak, strong, and electromagnetic). In addition to charge, spin and mass,

Leptons			
Particle	Charge	Antiparticle	Mass (MeV/ c^2)
e^-	-1	e^+	0.51
ν_e	0	$\bar{\nu}_e$	$< 15 \times 10^{-6}$
μ^-	-1	μ^+	105.7
ν_μ	0	$\bar{\nu}_\mu$	< 0.17
τ^-	-1	τ^+	1.78×10^3
ν_τ	0	$\bar{\nu}_\tau$	< 24

Quark			
u	+2/3	\bar{u}	$2.3^{+0.7}_{-0.5}$
d	-1/3	\bar{d}	$4.8^{+0.5}_{-0.3}$
c	+2/3	\bar{c}	$(1.27 \pm 0.025) \times 10^3$
s	-1/3	\bar{s}	95 ± 5
t	+2/3	\bar{t}	$(1.73 \pm 0.9) \times 10^5$
b	-1/3	\bar{b}	$4.18 \pm 0.03 \times 10^3$

Table 2.1: List of quarks and leptons with their intrinsic properties [18].

quarks also have an additional quantum number called color [17]. This color charge allows quarks to engage in the strong interaction, resulting in color neutral composite particles known as hadrons. Due to a phenomenon called color confinement, quarks, unlike leptons, cannot be isolated singularly. Instead, they join together to form two possible types of hadrons: mesons, composed of quark and anti-quark pairs, or baryons, made up of three quarks or three antiquarks. An example of a baryon is the proton ($p^+ = uud$) or neutron ($n = udd$). Mesons include kaons ($K^+ = u\bar{s}$, $K^{*0} = d\bar{s}$), pions ($\pi^+ = u\bar{d}$, $\pi^0 = u\bar{u}$, $d\bar{d}$ or $s\bar{s}$), B ($B^+ = u\bar{b}$ or $B^0 = b\bar{d}$) and the $\Upsilon(4S)$ ($b\bar{b}$). In addition, pentaquark [20] states have recently been discovered and are comprised of four quarks and one antiquark bound together in a color-neutral state.

2.1.2 Gauge bosons

Interactions between the different quarks and leptons, via the electroweak and strong force, are mediated by spin-1 gauge bosons. The number of gauge bosons is equal to the number of generators of a specific symmetry group, and therefore there are 8+3+1 gauge bosons [16].

In the strong sector, these are the eight gluons, $g_\alpha = 1 \dots 8$, which correspond to the eight generators of $SU(3)_C$. Furthermore, the gauge field B_μ corresponds to the $U(1)_Y$ group and the three fields $W_\mu^{1,2,3}$ correspond to the generators of the $SU(2)_L$ group. After electroweak symmetry breaking, which will be discussed later in this section, the gauge fields B_μ and $W_\mu^{1,2,3}$ manifest themselves as the four gauge bosons, W^\pm , Z^0 and γ [21].

The photon is massless and acts as the exchange particle of the electromagnetic force between two charged particles. The fact that the photon has zero mass implies that the electromagnetic interaction has an infinite range. Furthermore, gluons are also massless neutral bosons and are the carriers of the strong force. However, unlike the photon which is electrically neutral, gluons also carry color charge. This implies that gluons participate in the strong interaction, in addition to mediating it. Because of the strength of gluon interactions and their self-interacting nature, the effective range of the strong interaction is not infinite. Instead, it has a range of about 10^{-12} cm, which corresponds to the size of the lightest hadron. Finally, the W^\pm and Z^0 bosons are heavy with masses of 80.4 and 91.2 GeV/ c^2 respectively [18]. Because of the massive W^\pm and Z^0 bosons, the weak interaction has a very short range, 10^{-16} cm. These bosons mediate weak interactions amongst all SM particles, and in the case of W^\pm , amongst particles of different flavours. Such processes are often referred to as flavour-changing charged-current weak interactions, where exchanging a W^\pm can lead, for instance, to the coupling of an up-type quark ($Q=+2/3$) to a down-type quark ($Q=-1/3$). Also, because the mass eigenstates of quarks are not the weak eigenstates, these interactions, referred to as quark mixing, can also occur between the different generations of quarks, such as the coupling of an up-type quark to a strange-type quark.

Flavour-changing charged-current weak interactions are favoured in the SM and occur at tree level. The strength of such decays is governed by elements of the Cabbibo-Kobayashi-Maskawa matrix (CKM) matrix [18], which is shown in Fig. 2.1. The CKM matrix relates the mass eigenstates of the quarks involved to their weak eigenstates. Its diagonal elements relate quarks within the same generation and are thus very close to 1. Off-diagonal elements

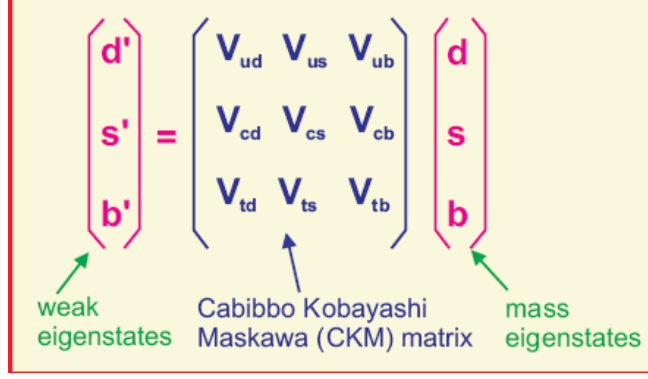


Figure 2.1: The Cabbibo-Kobayashi-Maskawa matrix and its different components.

relate quarks between different generations and vary in strength [18]. Furthermore, the CKM matrix is unitary by definition which allows for such relations between its different elements:

$$\sum_i V_{ij} V_{ik}^* = \delta_{jk} \quad (2.3)$$

$$V_{ud} V_{ub}^* + V_{cd} V_{cb}^* + V_{td} V_{tb}^* = 0 \quad (2.4)$$

These relations can be expressed in terms of unitarity triangles, whose side and angles can be experimentally determined from a large number of tree and loop-level decays, as shown in Fig. 2.2. Over the past decade or so, the *BABAR* [22] and *Belle* [23] experiments, along with other experiments, measured the sides and angles of such triangles and indeed found unitarity [24]. Furthermore, unlike charged-current interactions, which are mediated by W^\pm bosons, neutral-current interactions conserve flavour in the SM and are mediated by a Z^0 or γ . Flavour-changing neutral current interactions are forbidden at tree level. They only occur at loop level, as shown in Fig. 2.3, and are thus highly suppressed in the SM. Such decays are interesting since new physics contributions could enter into the loop and deviate observables, such as the branching fraction, from their SM expectation.

Weak decays only take place between left-handed particles or right-handed anti-particles, and thus the weak force violates parity symmetry [25]. In fact, it was experimentally observed, initially in K^\pm decays [26], that the weak interactions are CP-violating, where CP

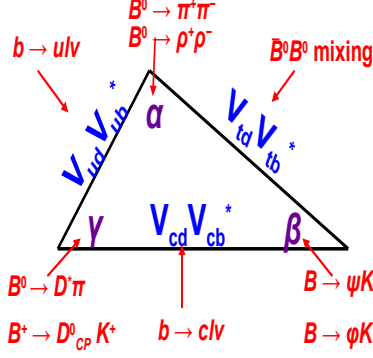


Figure 2.2: One of the 3 unitary triangles that result from the CKM matrix [29].

is the combination of the charge conjugation (C) operation and the parity (P) operation. The combined CP operation transforms a left-handed particle state into its corresponding right-handed antiparticle. This implies that, under the weak force, a particle will not react the same way if it is interchanged with its antiparticle and its parity is swapped. Within the SM, CP-violation in the weak sector arises from the presence of one irreducible complex phase in the CKM matrix. While CP-violation in the weak sector can contribute to the baryonic asymmetry of the universe, it is found to be insufficient to account for the large observed discrepancy [18]. Nevertheless, the precision measurements of the *BABAR* [22] and Belle [23] experiments proved Kobayashi and Maskawa's theory about the CKM matrix and the origin of CP-violation in weak decays. This led to these theorists being awarded the Nobel Prize in physics in 2008 [27]. On the other hand, the strong force does not appear to violate CP, a puzzling fact known as the strong CP problem [28]. The Standard Model's inability to account for the matter-antimatter asymmetry and the unexplained absence of CP-violation in the strong sector further promotes the need for a more encompassing model of particle physics.

The fact that the W^\pm and Z^0 bosons are massive implies that $SU(2)_L \times U(1)_Y$ is not an exact symmetry. These particles are accurately represented by an $SU(2)$ gauge theory, describing massless bosons, as with the photon for the $U(1)$ gauge group. However, because these bosons are observed to be massive, the symmetry of the Lagrangian must be broken

in the ground state of the system. In the SM, symmetry breaking is achieved with the Higgs mechanism, where the symmetry in the form $SU(3)_C \times SU(2)_L \times U(1)_Y$ is spontaneously broken to $SU(3)_C \times U(1)_{(EM)}$. This is done by introducing a complex $SU(2)$ doublet of scalar fields, the Higgs field, which permeates all space and has a potential with a nonzero expectation value [21]. The Higgs field induces spontaneous symmetry breaking in three of the four generators of the group $SU(2)_L \times U(1)_Y$. Spontaneous symmetry breaking of a gauge group necessitates the presence of a Goldstone boson [30]. With the Higgs mechanism, three of the four components of the resulting Goldstone boson mix with the W^\pm and Z^0 bosons, rendering them massive. The photon, which is the generator of the residual $U(1)_{(EM)}$, remains massless [21].

The last component of the Goldstone boson is a spin-0 Higgs particle. Collider experiments have been on the hunt for the Higgs boson for the past 4 decades. In 2012, the Large Hadron Collider (LHC) announced the discovery of a particle with a mass of 125 GeV/ c^2 [2]. Since then, the particle has been shown to have properties consistent with the SM Higgs boson [31], confirming the discovery of this long awaited particle. The Higgs boson is the first elementary scalar particle to be discovered in nature and the final missing component of the SM.

The Higgs field is also responsible for the masses of SM fermions. This is achieved by introducing Yukawa coupling terms, which describe the interaction between a scalar and Dirac field, to the Lagrangian. Yukawa couplings have the form $-\lambda_n \bar{\psi}_{m,L} \tilde{\phi} \psi_{n,R}$, where ϕ and ψ are the scalar and Dirac fields (leptons and quarks) respectively, m and n are the indices for the different generations, and L and R indicate the handedness of fermions (L : left-handed, R : right-handed) [21]. The scalar Higgs field ϕ can be expressed as $\tilde{\phi} = \phi - \phi_0$, where ϕ_0 is the non-zero vacuum expectation value, and included into the Lagrangian. Doing so, the mass terms of the fermions can be easily identified in the form $\lambda_n \phi_0 \bar{\psi} \psi$, where $\lambda_n \phi_0$ is determined as the fermion mass [21]. Because neutrinos do not have a right-handed partner, they cannot acquire mass through Yukawa couplings and thus should be massless in the SM.

The discovery of neutrino mass [32] is strong evidence of physics beyond the SM.

2.2 SuperSymmetry and the Two-Higgs-Doublet model

Various new physics models have been proposed, with a potential solution to shortcomings of the Standard Model. However, in many cases, experimental verification of their predictions requires access to energy scales that are beyond the limits of current colliders. One of the more popular ideas of physics beyond the SM, which can be within potential reach of current and future experiments, is SuperSymmetry (SUSY). Experiments at the high energy frontier, such as ATLAS at the LHC [33], are currently on the hunt for evidence of SuperSymmetry. Below is a brief introduction to SuperSymmetry and its implications in terms of the new physics models we are looking for.

SuperSymmetry is established by introducing commutator relationships between fermions and bosons. A supersymmetry transformation turns a fermionic state into a bosonic state and vice versa. The irreducible representation of the supersymmetry algebra is a supermultiplet, containing both fermion and boson states [34].

There are various supersymmetric extensions to the SM, the most basic of which is the Minimal SuperSymmetric Standard Model, MSSM. In this model, the only additional particles are the superpartners of those in the Standard model. Thus, every spin 1/2 lepton or quark has a spin-0 superpartner, and every spin-1 boson has a spin 1/2 superpartner. If supersymmetry was an exact symmetry of nature, a bosonic superpartner should have the same mass as its fermion counterpart and vice versa. However, superpartners have not been detected in current experiments, which implies that, if they exist, they must be more massive than SM particles. This pushes the possible SUSY scale to high energies, where it is less evident how SUSY can solve the hierarchy problem. It also implies that supersymmetry is a broken symmetry and this can be achieved by introducing soft supersymmetry breaking operators into the MSSM Lagrangian [34].

Even though associating a superpartner to each Standard Model particle is fairly straightforward, more than one Higgs superpartner is required. The presence of only one “Higgsino” would result in a gauge anomaly of the electroweak symmetry. Having one fermionic partner to the Higgs boson yields a nonzero contribution to the otherwise vanishing traces of the weak isospin and weak hypercharge. A vanishing trace is required for the cancellation of the gauge anomaly [10]. To avoid this, two Higgs supermultiplets are introduced, with contributions that cancel and thus resolve the gauge anomaly. This is the idea behind the two-Higgs-doublet model (2HDM), which appears as a natural feature of supersymmetric models [10]. It is also a non-SUSY minimal extension of the SM in which the Higgs sector is extended such that there are 2 Higgs doublets, H_1 and H_2 . In Type-II 2HDM, H_1 has an expectation value v_1 and it is from this doublet that the down-type quarks and leptons acquire their mass through Yukawa couplings [10]. Up-type quarks acquire their mass from Yukawa couplings to the other doublet H_2 which has a vacuum expectation value of v_2 . Within this model, there are five Higgs fields: two neutral scalars H^0 and h^0 , one neutral pseudoscalar A^0 , and two charged scalars H^+ and H^- . The interactions of the Higgs with fermions and quarks are dependent on the free parameter $\tan \beta = v_2/v_1$, which has been constrained, for instance, by studies of $B - \bar{B}$ and $K - \bar{K}$ mixing, $b \rightarrow s \gamma$ and $b \rightarrow c \tau \nu_\tau$ decays [35].

2.3 $B^+ \rightarrow K^+ \tau^+ \tau^-$: significance and motivation

A Feynman diagram of $B^+ \rightarrow K^+ \tau^+ \tau^-$ is shown in Fig. 2.3. By searching for such rare decays, one can test the predictions and parameters of the SM while simultaneously investigating the possibility of new physics contributions. Observables, such as the branching fraction, can be determined in the SM using an effective field theory approach.

An effective field theory (EFT) is a quantum field theory valid to a chosen energy scale [36]. Given a quantum field theory with energy scale E_0 , one can examine physics at energy $E \ll E_0$ using an EFT approach. This can be done by selecting an energy

¹The charge conjugate mode, $B^- \rightarrow K^- \tau^+ \tau^-$, is also implied.

λ , such that $\lambda \leq E_0$ and then dividing the fields in the path integral into high and low energy, or frequency, parts $\phi = \phi_H + \phi_L$. After integrating out the high-energy part, the result is an effective Lagrangian, which can be expressed in terms of an infinite sum of local operators, O_i , multiplied by coupling constants C_i , which are referred to as the Wilson Coefficients. The expression of the Lagrangian in terms of operators and coefficients is referred to as operator product expansion (OPE). In general, the Lagrangian includes all operators allowed by the symmetries of the system. To evaluate this infinite sum, dimensional analysis is used to determine the contribution of a given operator to an observable at low energy. Only operators and Wilson coefficients with a contribution $\gg 1$ are retained in the effective Lagrangian [37].

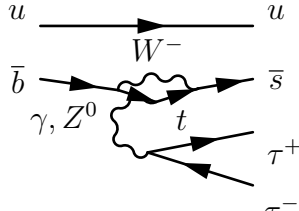


Fig 2.3: Feynman diagram for $B^+ \rightarrow K^+ \tau^+ \tau^-$.

In the upcoming section, the application of EFT to $B \rightarrow X_s \ell^+ \ell^-$ is discussed to derive the relevant SM branching fraction and explain the decay's sensitivity to the SM and beyond. Here, $B \rightarrow X_s \ell^+ \ell^-$ refers to the inclusive decay mode, where $\ell = e, \mu, \tau$ and X_s is any strange meson. $B^+ \rightarrow K^+ \tau^+ \tau^-$ is the exclusive decay mode where the meson system is specified to be a charged kaon and the leptons are specified to be taus.

2.3.1 $B^+ \rightarrow K^+ \tau^+ \tau^-$ in the Standard Model

Using the method of OPE and the approach of EFT, the decay amplitude of $B \rightarrow X_s \ell^+ \ell^-$ can be separated into two parts: the short-distance (high energy) physics which is described by the Wilson coefficients and the long-distance (low energy) contributions contained in the operator matrix elements [38]. The electroweak effective Hamiltonian of $B \rightarrow X_s \ell^+ \ell^-$ is

written as:

$$H = \frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \sum_{i=1}^{10} C_i(\mu) O_i(\mu) \quad (2.5)$$

where V_{ij} are the relevant CKM matrix elements and C_i are the corresponding Wilson coefficients. O_i are a complete set of renormalized operators involving the fields that govern $b \rightarrow s$ transitions [39]. These include 6 four-quark operators $O_1 - O_6$, the electromagnetic operator C_7 , the chromomagnetic operator C_8 , and the vector and axial-vector contributions to the exclusive $b \rightarrow s \ell^+ \ell^-$ denoted by O_9 and O_{10} respectively. The operators are usually parameterized in terms of form factors, which can be calculated using various theoretical models. The exact expressions for these operators can be found in [40].

The Wilson coefficients are calculated perturbatively at the electroweak scale and then evolved down to the renormalization scale μ , where here $\mu \sim m_b$. For this specific decay, $C_{1,3-6}(M_W) = 0$, $C_2(M_W) = -1$, and C_{7-10} are given by the Inami-Lim functions [41]. The differential branching fraction for $B \rightarrow X_s \tau^+ \tau^-$ is thus given by:

$$\begin{aligned} \frac{dB(B \rightarrow X_s \tau^+ \tau^-)}{d\hat{s}} = & \\ B(B \rightarrow X l \nu) \frac{\alpha^2}{4\pi^2} \frac{|V_{tb} V_{ts}^*|^2}{|V_{cb}|^2} \frac{(1 - \hat{s})^2}{f(z) \kappa(z)} \left[1 - \frac{4x}{\hat{s}} \right]^{1/2} & \left([|C_9^{eff}|^2 - |C_{10}|^2] 6x \right. \\ & + [|C_9^{eff}|^2 - |C_{10}|^2] \left[(\hat{s} - 4x) + (1 + \frac{2x}{\hat{s}})(1 + \hat{s}) \right] \\ & \left. + 12C_7 \text{Re} C_9^{eff} (1 + \frac{2x}{\hat{s}}) + \frac{4|C_7|^2}{\hat{s}} (1 + \frac{2x}{\hat{s}})(2 + \hat{s}) \right) \end{aligned} \quad (2.6)$$

where $\hat{s} \equiv q^2/m_b^2$, $x \equiv m_\tau^2/m_b^2$, and $z \equiv m_c/m_b$ [11]. The branching fraction is scaled to that of the decay $B \rightarrow X l \nu$ to remove systematic uncertainties associated with the CKM matrix elements, as well as other factors. Furthermore, $f(z)$ and $\kappa(z)$ represent the phase space and QCD corrections to the decay respectively [42]. The SM-based branching fraction is model-dependent in the leading order, where the scale dependence of $C_{7,9}$ yields large uncertainties. This effect is reduced by including the next-to-leading logarithmic correction to $C_9(\mu)$ while

keeping only the leading logarithms in the remaining Wilson coefficients. This yields an effective value of $C_9^{eff}(\mu) = C_9(\mu) + Y(\mu, q^2)$, where the function Y contains the contributions from the one-loop matrix element of the four-quark operators. The numerical estimates of these coefficients at $\mu = m_b$, taking $m_t = 175 \text{ GeV}$, $m_b = 4.87 \text{ GeV}$ and $\alpha_s(M_Z) = 0.118$, are [48] :

$$\begin{aligned} C_7(\mu = (m_b)_{+m_b}^{-m_b/2}) &= -0.312_{+0.034}^{-0.059}, \\ C_9(\mu = (m_b)_{+m_b}^{-m_b/2}) &= 4.21_{-0.40}^{+0.31}, \\ C_{10}(\mu = m_b) &= -4.55. \end{aligned} \tag{2.7}$$

Here, the scale dependence of C_7 and C_9 is clearly visible in $\sim 10\%$ and $\sim 20\%$ deviations when μ is varied between $m_b/2 \leq \mu \leq 2m_b$.

In addition, $B \rightarrow X_s \ell^+ \ell^-$ is also affected by long distance contributions. These arise predominantly due to J/ψ and $\psi(2S)$ resonances through the decay chain $B^+ \rightarrow K^+ \psi(2S)/J/\psi, \psi(2S)/J/\psi \rightarrow \ell^+ \ell^-$. Such resonant contributions result in significant interference effects with the short-distance physics as well as the dispersive part of the resonance. However, well-chosen mass cuts in the $\ell^+ \ell^-$ spectrum can cleanly separate such contributions from the short distance physics. For $B^+ \rightarrow K^+ \tau^+ \tau^-$, only the $\psi(2S)$ resonance contributes because the mass of the τ leptons pushes the \hat{s} distribution, where $\hat{s} = q^2/m_B^2$ is the normalized q^2 , into high regions in the $\ell^+ \ell^-$ mass.

The differential branching fraction for $B \rightarrow X_s \ell^+ \ell^-$ is shown in Fig. 2.3 with and without the long distance contributions for both $l = e$ or τ [11]. As can be seen, the massive leptons impose an upper limit of $\sim 1.5 \text{ GeV}/c$ on the kaon momentum, in the lab frame, and thus the \hat{s} distribution for $B \rightarrow X_s \tau^+ \tau^-$ only dominates the high dilepton mass region. The integrated branching fractions for $l = e^+, \mu^+$ and τ^+ are given in Table 2.2 for both the total and high dilepton mass regions [11]. It should be noted that the branching fraction of $B \rightarrow X_s \tau^+ \tau^-$ is comparable to its light lepton counterparts in the high dilepton mass region above the $\psi(2S)$ resonance. According to Ref. [43], the exclusive branching fraction for $B \rightarrow K \tau^+ \tau^-$

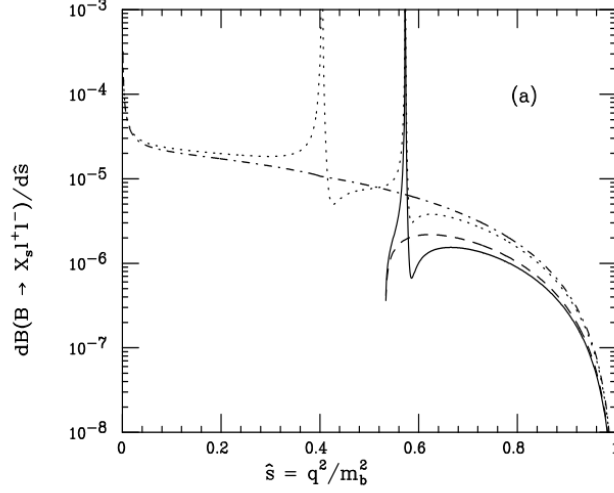


Figure 2.3: Differential branching fraction for $B \rightarrow X_s \ell^+ \ell^-$ for $l = \tau$ (solid and dashed) and $l = e$ (dotted and dash-dotted) with and without the long-distance contributions [11].

is calculated, using heavy meson chiral perturbation theory, to be $\approx 50-60\%$ of the inclusive decay. This leads to a branching fraction of $\approx 2 \times 10^{-7}$.

l	$4x \leq \hat{s} \leq 1$	$0.6 \leq \hat{s} \leq 1$
e	1.2×10^{-5}	8.5×10^{-7}
μ	1.0×10^{-5}	8.5×10^{-7}
τ	5.4×10^{-7}	4.3×10^{-7}

Table 2.2: Integrated branching fractions for $l = e, \mu$ and τ in the total and high dilepton mass regions, where \hat{s} is the normalized q^2 and $x \equiv m_\tau^2/m_b^2$ [11].

Within the SM, FCNC processes such as $B^+ \rightarrow K^+ \tau^+ \tau^-$ are important because they allow for the testing of QCD and its available techniques (perturbative QCD, heavy quark effective theory HQET, lattice-QCD, QCD sum-rules), which are directly applicable here. In addition, by measuring the decay rate of $B^+ \rightarrow K^+ \tau^+ \tau^-$, one can constrain the Wilson coefficients as well as the operator matrix elements. As shown in equation 2.6, the C_7 , C_9^{eff} and the C_{10} enter into the differential branching fraction and thus a precise measurement of the decay rate is essential in accurately determining these parameters. As will be discussed in section 2.3.3, the $B^+ \rightarrow K^{(*)} \ell^+ \ell^-$ branching fractions, where $l = e$ or μ , had previously

been measured and used to constrain the relevant Wilson coefficients.

Distributions such as the lepton pair invariant mass spectrum and the lepton pair forward-backward asymmetry, in addition to the total rate, are also essential in separating the short distance physics from the long-distance contributions. One distribution that is unique to $B^+ \rightarrow K^+ \tau^+ \tau^-$, when compared to its light-lepton counterparts, is the tau polarization asymmetry. This is defined as [45]:

$$P_i(p^2) = \frac{\frac{d\Gamma}{dp^2}(\vec{n} = \vec{e}_i) - \frac{d\Gamma}{dp^2}(\vec{n} = -\vec{e}_i)}{\frac{d\Gamma}{dp^2}(\vec{n} = \vec{e}_i) + \frac{d\Gamma}{dp^2}(\vec{n} = -\vec{e}_i)} \quad (2.8)$$

where $\frac{d\Gamma}{dp^2}$ is the differential decay rate for a given spin direction \vec{n} of the τ^- lepton and \vec{n} is a unit vector in the τ^- rest frame. Here, p^2 is the q^2 , the invariant mass of the lepton pair. The i denotes the 3 components of the tau polarization (longitudinal, transverse, and normal) defined by the following 3 unit vectors:

$$\begin{aligned} \vec{e}_L &= \frac{\vec{p}_1}{|\vec{p}_1|}, \\ \vec{e}_N &= \frac{\vec{p}_K \times \vec{p}_1}{|\vec{p}_K \times \vec{p}_1|}, \\ \vec{e}_T &= \vec{e}_N \times \vec{e}_L \end{aligned} \quad (2.9)$$

where \vec{p}_l and \vec{p}_K are the 3 momenta vectors of the ℓ^- and the K , respectively, in the center-of-mass (CM) of the $\ell^+ \ell^-$ system.

The τ polarization asymmetry can be determined by studying the momentum distributions of its decay products. These have a large dependence on the spin polarization of the final state lepton and thus the τ 's helicity can be easily established. The different components of the polarization, P_L, P_T , and P_N , involve different combinations of the Wilson coefficients C_7, C_9^{eff} , and C_{10} and thus contain independent information on the structure of the SM [45].

A distribution of the longitudinal τ polarization asymmetry, P_L , is shown in Fig. 2.4 with and without the long distance contributions. The sensitivity of this distribution to

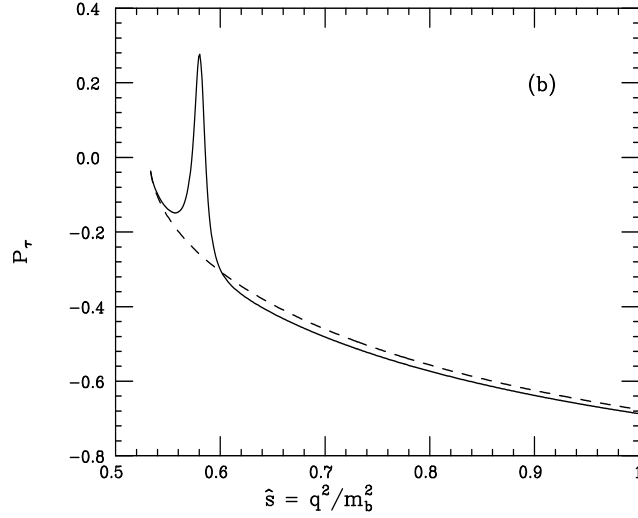


Figure 2.4: τ polarization asymmetry as a function of \hat{s} , with (black line) and without (dashed curve) long-distance contributions [11].

the Wilson coefficients is shown in Fig. 2.5 [11]. The sign of C_{7-10} is flipped while keeping their magnitudes constant and the difference in distribution is clearly visible. The highest sensitivity is for any combination of sign changes in C_9 and C_{10} , mainly because the operators O_9 and O_{10} dominate the decay in the high \hat{s} region.

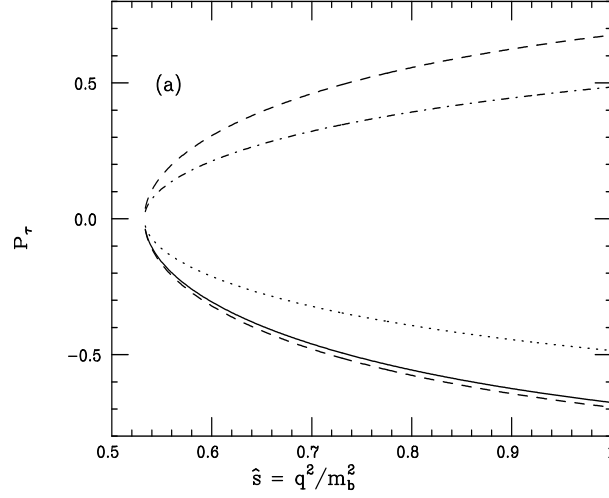


Figure 2.5: Sensitivity of the τ polarization asymmetry to the Wilson coefficients as a function of \hat{s} . From top to bottom, the asymmetry is shown if $C_{10}(M_W) \rightarrow -C_{10}(M_W)$ (dashed, top curve), $C_{9,10}(M_W) \rightarrow -C_{9,10}(M_W)$ (dash-dotted), $C_9(M_W) \rightarrow -C_9(M_W)$ (dotted), and $C_{7,8}(M_W) \rightarrow -C_{7,8}(M_W)$ (long-dashed, bottom curve). The solid distribution is the SM [11].

2.3.2 $B^+ \rightarrow K^+ \tau^+ \tau^-$ beyond the Standard Model

$B^+ \rightarrow K^+ \tau^+ \tau^-$ is also a vital search for physics beyond the SM because new physics contributions could enter into the loop and alter the decay rate. These could include a charged Higgs boson or a supersymmetric particle as shown in Fig. 2.6. Such contributions could affect the total branching fraction, as well as the kinematic and angular distributions of the final lepton pair. An important contribution to this decay is from neutral Higgs bosons in two-Higgs-doublet models. The lepton-lepton-Higgs vertices are proportional to the mass of the lepton and thus in the case of the τ lepton, contributions from the neutral Higgs bosons are significant [45]. Fig. 2.7 shows the additional Feynman diagrams corresponding to the exchange of the neutral Higgs bosons, H^0, h^0 and A^0 . Taking into account these contributions, the effective Hamiltonian can be written as [46]:

$$H_{eff} = \frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \left(\sum_{i=1}^{10} C_i(\mu) O_i(\mu) + \sum_{i=1}^{10} C_{Q_i}(\mu) Q_i(\mu) \right) \quad (2.10)$$

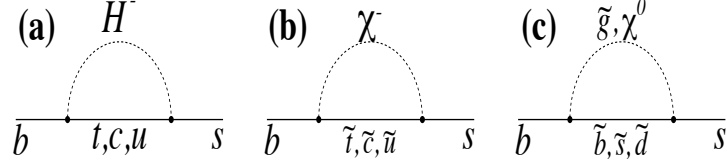


Figure 2.6: New physics contributions in $b \rightarrow s$ transitions: a) charged Higgs (H^-), b) squark and chargino, c) gluino and neutralino.

	m_{h^0} (GeV)	m_{H^\pm} (GeV)	m_{H^0} (GeV)	m_{A^0} (GeV)
mass set-1	80	200	150	100
mass set-2	250	300	100	350
mass set-3	100	400	200	150

Table 2.3: List of mass values for each of the Higgs particles used to determine the branching fraction and τ polarization plots in 2HDM [45].

Here, the first set of operators in the brackets describe the $b \rightarrow s \ell^+ \ell^-$ effective Hamiltonian in the SM. The second set of operators, $C_{Q_i}(\mu)Q_i(\mu)$, corresponds to the contribution of the neutral Higgs bosons. Furthermore, contributions from the charged Higgs bosons of 2HDM are also present and are taken into account by modifying the Wilson Coefficients C_{Q_i} , without introducing any new operators.

Fig. 2.8 shows the branching fraction as a function of the squared momentum of the $\tau^+ \tau^-$ pair (q^2), referred to in [45] as p^2 , for different values of $\tan \beta$ [45]. Lines 1, 2, 3 indicate the different sets of mass values of the Higgs particles, as given in Table 2.3. For each mass set, contributions from neutral Higgs bosons deviate the branching fraction from the SM expectation, for high values of $\tan \beta$. The same is true for the P_L and P_T distributions shown in Figs. 2.9 and 2.10.

Additional sources of new physics and their effect on the $B^+ \rightarrow K^+ \tau^+ \tau^-$ branching fraction and the kinematic distributions of the $\tau^+ \tau^-$ pair are also discussed in Refs. [47]-[54]. In conclusion, $B^+ \rightarrow K^+ \tau^+ \tau^-$ is an interesting mode because it provides both a precision test of the SM and a search for new physics. However, because of the current size of the *BABAR* data sample, only the latter can be tested in this analysis. Although the

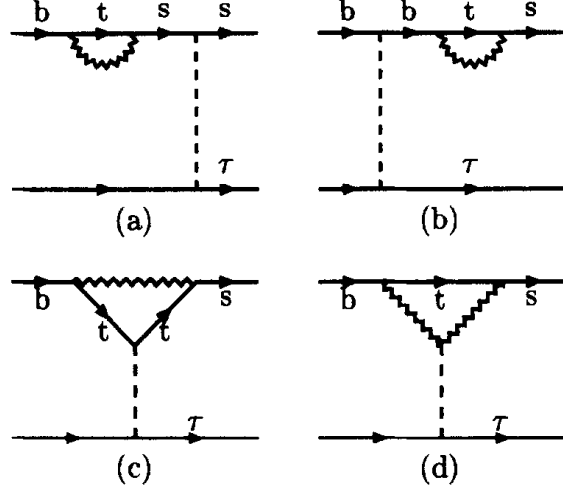


Figure 2.7: Additional Feynman diagrams due to the exchange of the neutral Higgs bosons in 2HDM. The wavy lines represent the propagators of charged bosons W^\pm and H^\pm , while the dashed lines represent the neutral bosons H^0 , h^0 and A^0 [45].

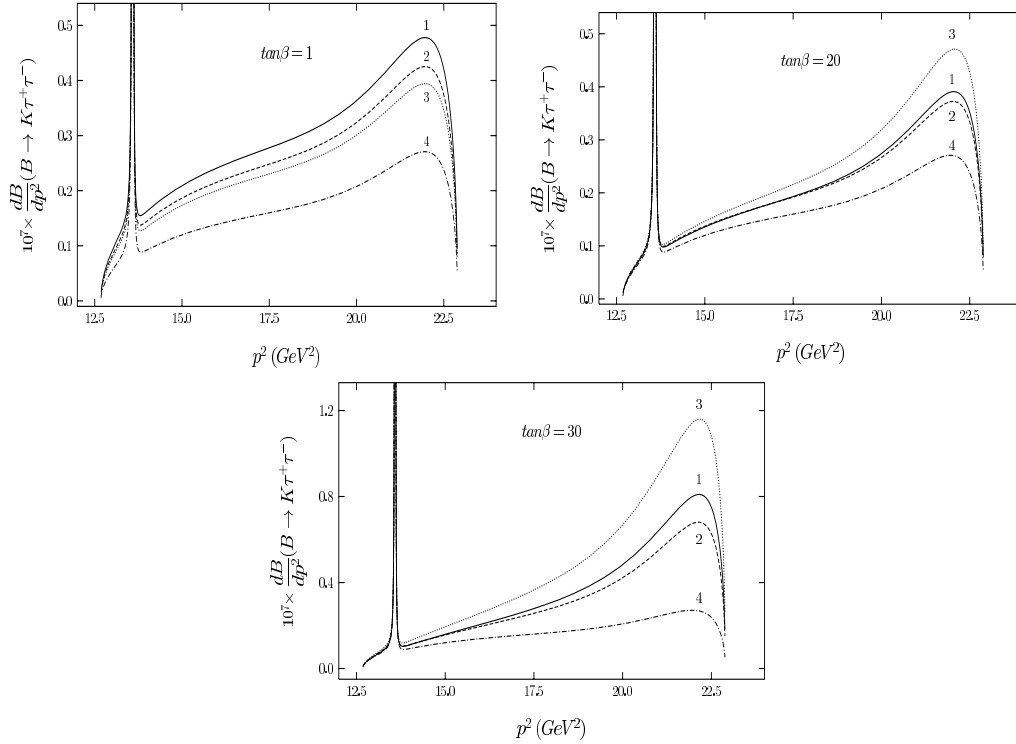


Figure 2.8: $\mathcal{B}(B \rightarrow K \tau^+ \tau^-)$ with long distance effects for each of mass set 1, 2, and 3, respectively. Curve 4 is the Standard Model prediction. Top-left: $\tan \beta = 1$, top-right: $\tan \beta = 20$, bottom $\tan \beta = 30$ [45].

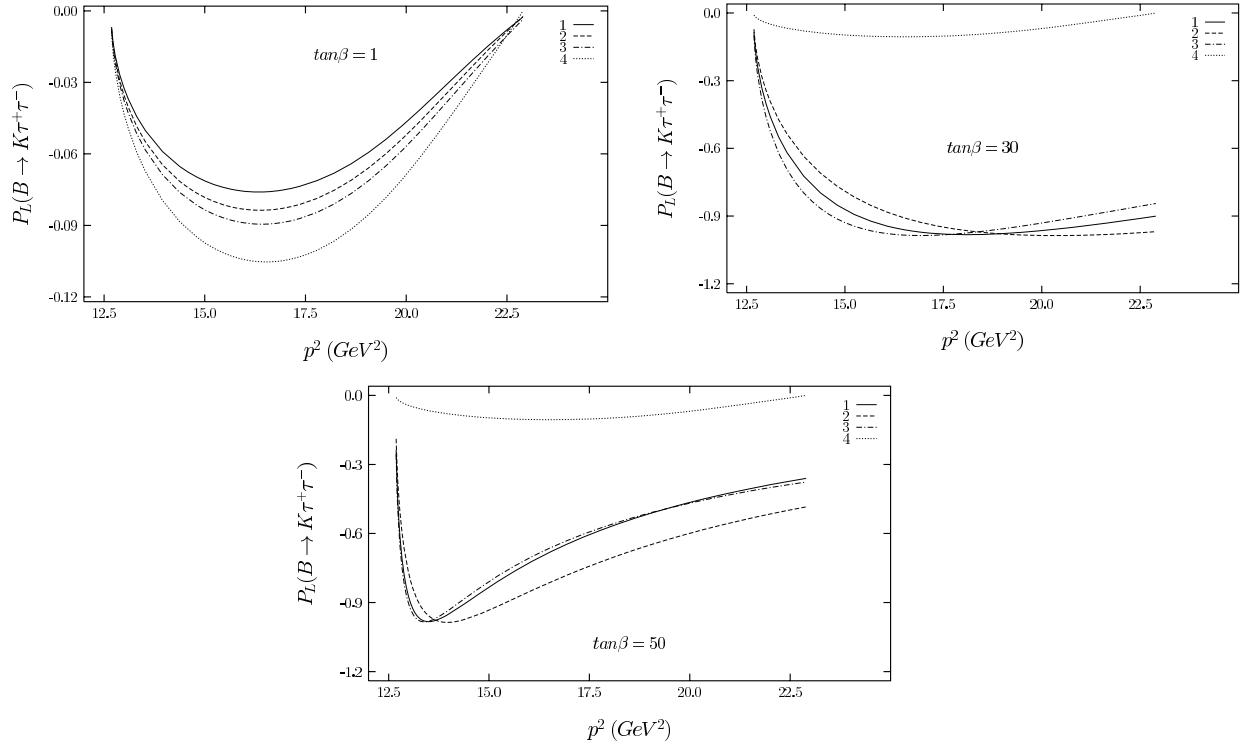


Figure 2.9: Longitudinal τ polarization asymmetry as a function of p^2 without long distance effects for each of mass set 1, 2, and 3, respectively. Curve 4 is the Standard Model prediction. Top-left: $\tan\beta = 1$, top-right: $\tan\beta = 30$, bottom $\tan\beta = 50$ [45].

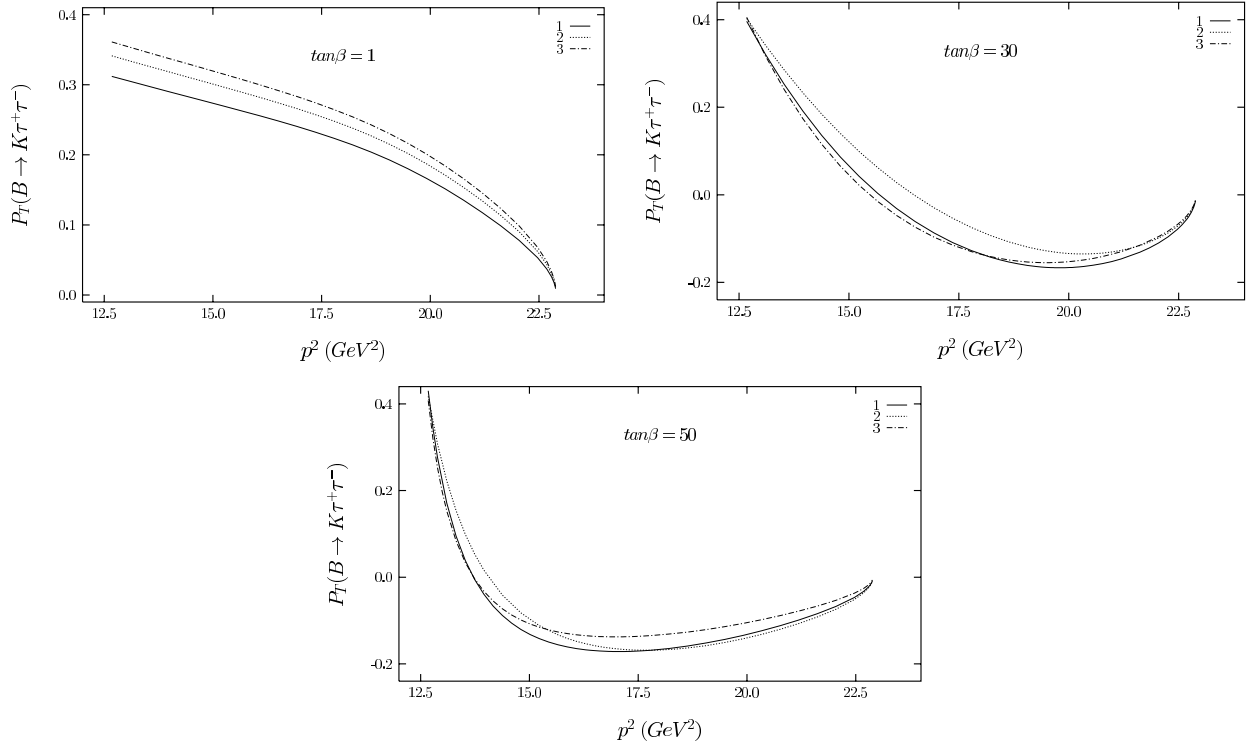


Figure 2.10: Transverse τ polarization asymmetry as a function of p^2 without long distance effects for each of mass set 1, 2, and 3, respectively. Top-left: $\tan \beta = 1$, top-right: $\tan \beta = 30$, bottom $\tan \beta = 50$ [45].

BABAR data sample is not sufficiently large to expect sensitivity to $B^+ \rightarrow K^+ \tau^+ \tau^-$ decays occurring at the SM rate, the branching fraction can be constrained using the present search. Furthermore, deviation from the SM expectation, if found, can be a clear hint of new physics. Observables like the τ polarization asymmetry, while very interesting, cannot be measured in the current study because of the limited statistics.

2.3.3 Related measurements: $B \rightarrow K^{(*)} \ell^+ \ell^-$

$B \rightarrow K^{(*)} \ell^+ \ell^-$, where $\ell = e$ or μ , has previously been measured by *BABAR* [4], as well as other collider experiments, specifically CDF [6], Belle [5] and LHCb [7]. In addition to the total branching fraction, rate asymmetry measurements were also made. These rate asymmetries allow for the cancellation of the large theoretical uncertainties associated with the form factors and thus increases this mode's sensitivity to new physics. The total branching fractions, excluding long distance contributions, for $B \rightarrow K \ell^+ \ell^-$ and $B \rightarrow K^* \ell^+ \ell^-$ were measured by *BABAR* to be $(4.7 \pm 0.6 \pm 0.2) \times 10^{-7}$ and $(10.2^{+1.4}_{-1.3} \pm 0.5) \times 10^{-7}$ [4]. As shown in Fig. 2.11, these results are in agreement with measurements from Belle [5], CDF [6], and in agreement with SM theoretical predictions of Ali *et al.* [55] and Zhong *et al.* [56]. The $B \rightarrow K \ell^+ \ell^-$ and $B \rightarrow K^* \ell^+ \ell^-$ partial branching fractions, measured by *BABAR* in bins of dilepton mass squared $s \equiv m_{\ell^+ \ell^-}^2$, are also shown in Fig. 2.12. The results are found to be consistent with the SM, Belle [5], CDF [6] and LHCb [7], where the latter only include $B \rightarrow K^* \mu^+ \mu^-$. Furthermore, recent calculations of the differential branching fractions, $dB(B \rightarrow K \ell^+ \ell^-)/dq^2$ for $\ell = e, \mu, \tau$, where $q^2 = s = (p_B - p_{K^{(*)}})^2$, have also been made for the first time using unquenched lattice QCD form factors [57]. The result of these calculations is shown in Fig. 2.13 and compared to the Belle, CDF, *BABAR* and more recent LHCb results [58] [59]. Agreement with the SM and all aforementioned experiments is found over the full q^2 region. Finally, the LHCb experiment recently published the most precise result

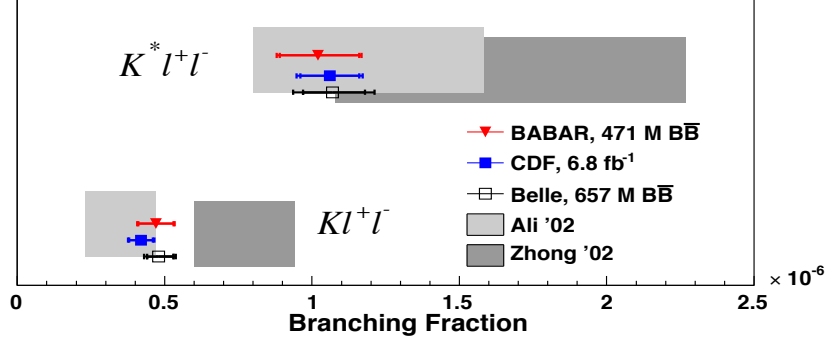


Figure 2.11: Total branching fraction for $B \rightarrow K \ell^+ \ell^-$ and $B \rightarrow K^* \ell^+ \ell^-$, as measured by *BABAR*, *CDF*, and *Belle* and compared with predictions from *Ali et al.* and *Zhong et al.* models [4].

on the total integrated branching fractions in the $\mu^+ \mu^-$ channel [61]:

$$\begin{aligned}
\mathcal{B}(B^+ \rightarrow K^+ \mu^+ \mu^-) &= (4.29 \pm 0.07(stat) \pm 0.21(sys)) \times 10^{-7}, \\
\mathcal{B}(B^+ \rightarrow K^0 \mu^+ \mu^-) &= (3.27 \pm 0.34(stat) \pm 0.17(sys)) \times 10^{-7}, \\
\mathcal{B}(B^+ \rightarrow K^{*+} \mu^+ \mu^-) &= (9.24 \pm 0.93(stat) \pm 0.67(sys)) \times 10^{-7}.
\end{aligned} \tag{2.12}$$

In addition to the differential branching fractions, *BABAR* also measured the direct CP asymmetry, the lepton flavor ratio and the CP-averaged isospin asymmetry of $B \rightarrow K \ell^+ \ell^-$ and $B \rightarrow K^* \ell^+ \ell^-$. These are given by the following equations, respectively:

$$\begin{aligned}
A_{CP}^{K(*)} &= \frac{B(\bar{B} \rightarrow \bar{K}^{(*)} \ell^+ \ell^-) - B(B \rightarrow K^{(*)} \ell^+ \ell^-)}{B(\bar{B} \rightarrow \bar{K}^{(*)} \ell^+ \ell^-) + B(B \rightarrow K^{(*)} \ell^+ \ell^-)}, \\
A_I^{K(*)} &= \frac{B(B^0 \rightarrow K^{(*)0} \ell^+ \ell^-) - r_\tau B(B \rightarrow K^{(*)+} \ell^+ \ell^-)}{B(B^0 \rightarrow K^{(*)0} \ell^+ \ell^-) + r_\tau B(B \rightarrow K^{(*)+} \ell^+ \ell^-)}, \\
R_{K(*)} &= \frac{B(B \rightarrow K^{(*)} \mu^+ \mu^-)}{B(B \rightarrow K^{(*)} e^+ e^-)}.
\end{aligned} \tag{2.14}$$

where $r_\tau \equiv \tau_{B^0}/\tau_{B^+} = 1/(1.071 \pm 0.009)$ is the ratio of the B^0 and B^+ lifetimes. In the SM, A_{CP} is expected to be of $\mathcal{O}(10^{-3})$, while $R_{K(*)}$ is expected to be of order unity. Contributions from new physics could enhance the CP asymmetry and increase $R_{K(*)}$ by up to 10%. Fig. 2.14 shows the *BABAR* results for A_{CP} and $R_{K(*)}$ as a function of s . These asymmetries were

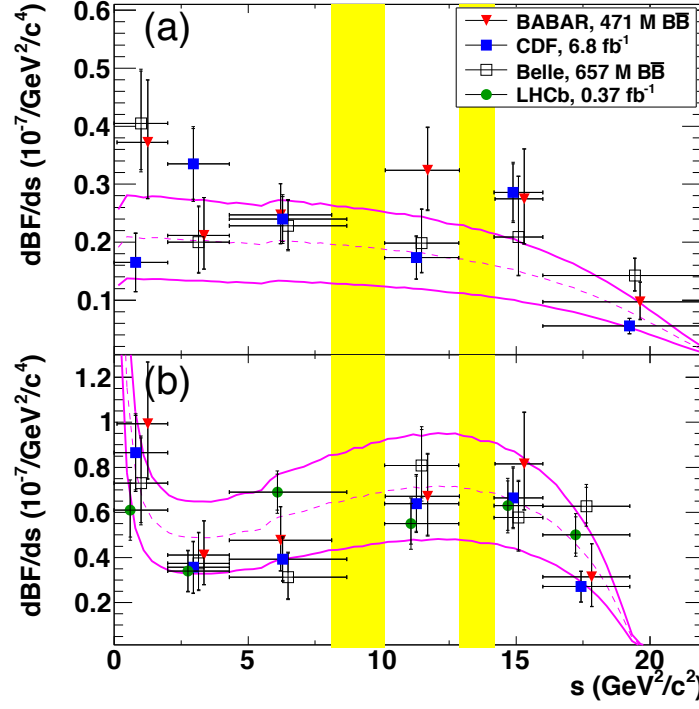


Figure 2.12: Partial branching fraction results of *BABAR* [4], *CDF* [6], *Belle* [5] and *LHCb* [58] [59] for $B \rightarrow K \ell^+ \ell^-$ (top) and $B \rightarrow K^* \ell^+ \ell^-$ (bottom) in bins of $s \equiv m_{\ell^+ \ell^-}^2$. The magenta lines show the SM theoretical predictions from Ali *et al.* model [55]. The vertical yellow bands are the vetoed regions of the J/ψ and $\psi(2S)$ resonances.

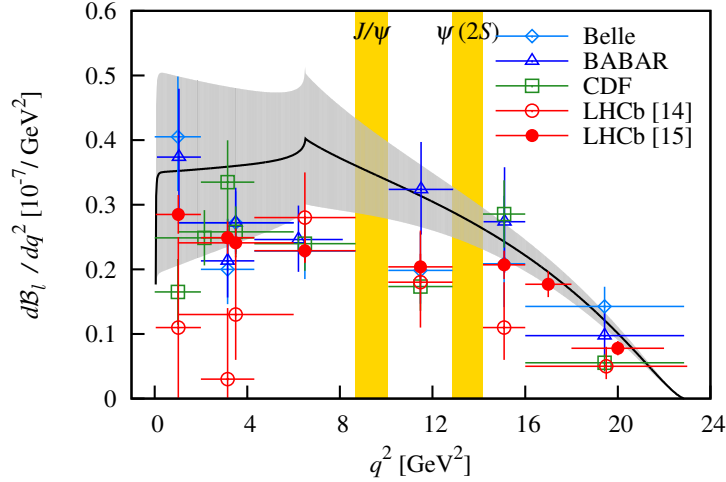


Figure 2.13: Standard Model branching fraction with form factors calculated using unquenched lattice QCD[57], compared to experimental results from *BABAR* [4], CDF [6], Belle [5], and LHCb [58] [59]. The vertical yellow bands are the vetoed regions of the J/ψ and $\psi(2S)$ resonances.

found to be consistent with the SM. Similar measurements have been done by Belle [5] and LHCb [7][58][59], where the latter found a 2.6σ deviation from the SM expectation in their measurement of $R_{K^{(*)}}$ [60]. Furthermore, $A_I^{K^{(*)}}$ is expected to be of the order of a few percent in the SM. *BABAR* [4] and Belle [5] also measured $A_I^{K^{(*)}}$, as shown in Fig. 2.15, and found agreement with the SM at the 1.2σ and 2.1σ level for the K and K^* modes, respectively. Finally, LHCb recently measured the isospin asymmetry for $B \rightarrow K \mu^+ \mu^-$ and $B \rightarrow K^* \mu^+ \mu^-$ [61] and also found consistency with the SM.

In conclusion, $B \rightarrow K \ell^+ \ell^-$ has been extensively studied by *BABAR* and other experiments. So far, deviations from the SM expectations have not been significant. Furthermore, preliminary results on $B \rightarrow K \tau^+ \tau^-$ have been previously presented, but not published [62]. While the experimental precision is expected to be lower than its electron and muon counterparts, the potential sensitivity of $B \rightarrow K \tau^+ \tau^-$ to new physics is in some cases much larger, particularly in models with enhanced couplings to third generation or mass.

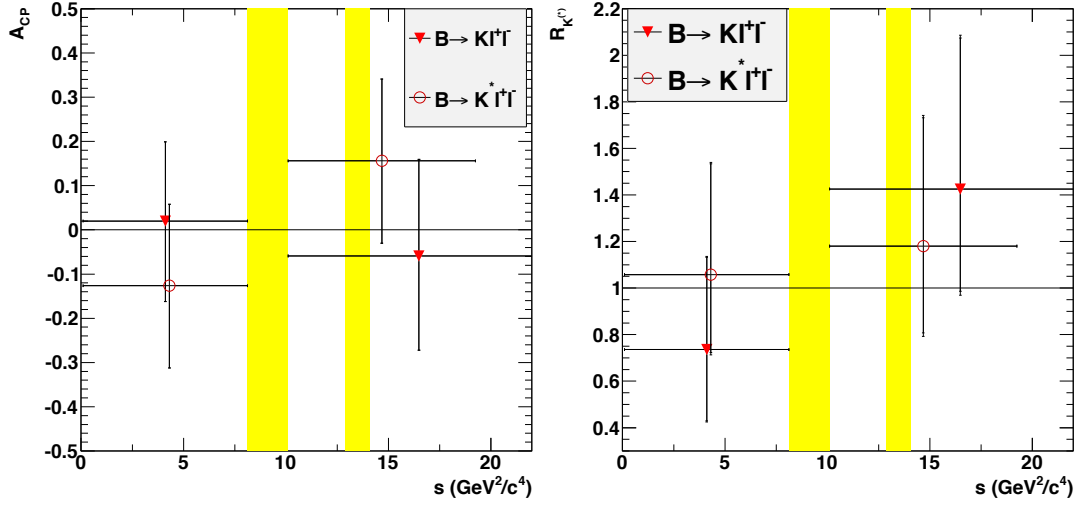


Figure 2.14: A_{CP} (left) and $R_{K^{(*)}}$ (right) for $B \rightarrow K \ell^+ \ell^-$ and $B \rightarrow K^* \ell^+ \ell^-$ modes as a function of s . The vertical yellow bands are the vetoed regions of the J/ψ and $\psi(2S)$ resonances [4].

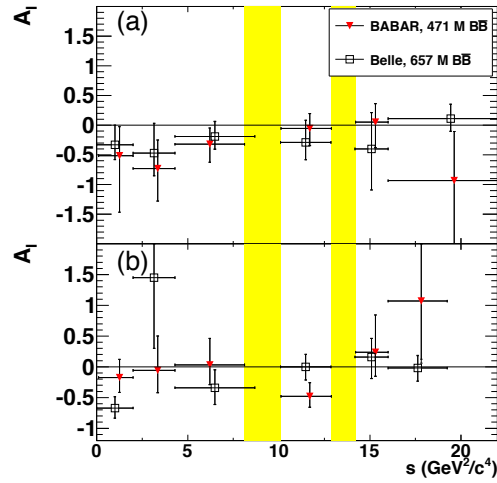


Figure 2.15: $A_I^{K^{(*)}}$ for $B \rightarrow K \ell^+ \ell^-$ and $B \rightarrow K^* \ell^+ \ell^-$ modes as a function of s . The vertical yellow bands are the vetoed regions of the J/ψ and $\psi(2S)$ resonances [4].

Chapter 3

The *BABAR* experiment

The *BABAR* experiment, located at the SLAC National Accelerator Laboratory, was designed to study CP-violation in B mesons, make precision measurements of SM decays, and search for new physics processes. The detector operated on the PEP-II asymmetric e^+e^- collider, which collided electrons and positrons at a center-of-mass energy of 10.58 GeV. After ten years of data collection, the *BABAR* experiment reached a total integrated luminosity of 530.8 fb^{-1} . Approximately $470.97 \times 10^6 \text{ } B\bar{B}$ events have been recorded, allowing for a wide range of precision measurements in the flavour sector of the SM. The PEP-II storage ring and the *BABAR* detector are described in detail below.

3.1 PEP-II

The PEP-II facility consisted of two vertically-stacked storage rings, as shown in Fig. 3.1. The high energy ring (HER) stored electrons at 9 GeV, while the low energy ring (LER) stored positrons at 3.1 GeV [22]. The two beams were injected into PEP-II from the SLAC linear collider after being accelerated to their collision energy. Furthermore, collisions occurred at a single interaction point (IP), where the upper LER was brought down into the plane of the HER. The energy difference between the beams was why the PEP-II ring was considered an asymmetric collider. It was designed to operate at a luminosity of 3×10^{33}

$\text{cm}^{-2} \text{ s}^{-1}$ and a CM energy of 10.58 GeV, corresponding to the mass of the $\Upsilon(4S)$ resonance. The $\Upsilon(4S)$ resonance is just above the production threshold for $B\bar{B}$ and has a lifetime of $3.3 \times 10^{-23} \text{ s}$ [18]. Therefore, after a collision, the $\Upsilon(4S)$ almost instantly decayed into a $B\bar{B}$ pair, which were produced with a momentum of $\sim 320 \text{ MeV}/c$ in the CM frame. The asymmetry in the beam energies resulted in a Lorentz boost to the $\Upsilon(4S)$ resonance of $\beta\gamma = 0.56$. This allowed each B , which has a lifetime of $1.64 \times 10^{-12} \text{ s}$ [18]¹, to travel a small distance ($\sim 260 \mu\text{m}$) before it decayed and thus separation of the two B decay vertices was possible. Furthermore, the relative decay times of each $B\bar{B}$ pair was determined allowing for accurate measurements of time-dependent CP violation.

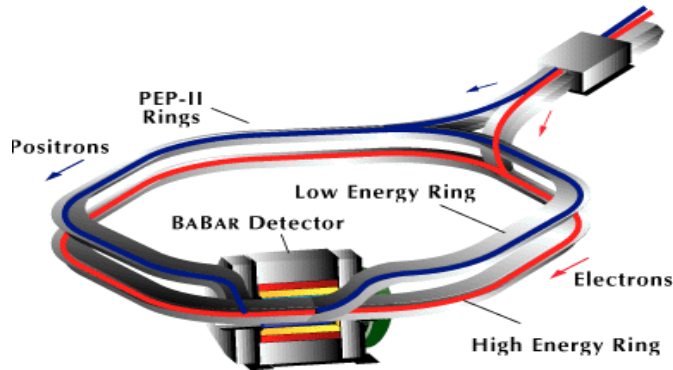


Figure 3.1: A schematic diagram of the SLAC linac and the PEP-II storage ring [64].

3.2 The *BABAR* detector

The need for precision measurements of CP-violation and SM decays placed stringent requirements on the performance of the *BABAR* detector. To achieve the desired physics objectives, the detector was required to have very good reconstruction efficiency for tracks ($p > 60 \text{ MeV}/c$) and photons ($E > 20 \text{ MeV}$), excellent vertex resolution ($< 60 \mu\text{m}$) as well as energy and angular resolution, efficient and accurate e, μ identification, a kaon-pion separation ($> 3\sigma$), and a light composition of its inner components to minimize scattering of charged

¹This is for charged B mesons, neutral B mesons have a lifetime of $1.52 \times 10^{-12} \text{ s}$ [18].

particles.

The *BABAR* detector was built around the PEP-II interaction region, and was offset by 0.37 m relative to the beam-beam interaction point, to increase the geometric acceptance of the boosted $\Upsilon(4S)$ system. The detector consisted of a silicon vertex tracker (SVT), drift chamber (DCH), detector of internally reflected Cherenkov light (DIRC), and an electromagnetic calorimeter (EMC). These were surrounded by a superconducting solenoid, which was designed for a magnetic field of magnitude 1.5T. In addition, an instrumented flux return (IFR) surrounded the magnet and allowed for the detection of muons and neutral hadrons. A schematic diagram of the longitudinal cross section of the detector and its end view are shown in Fig. 3.2 and Fig. 3.3, respectively [22]. The polar angle coverage extended from 350 mrad in the forward direction and 400 mrad in the backward direction, defined relative to the HER. As shown in both diagrams, the right-handed coordinate system was such that the positive z -axis runs along the beam line in the direction of the LER, the positive y -axis pointed upwards, and the positive x -axis pointed away from the center of the PEP-II storage rings. The detector subsystems are described in more detail in the following sections.

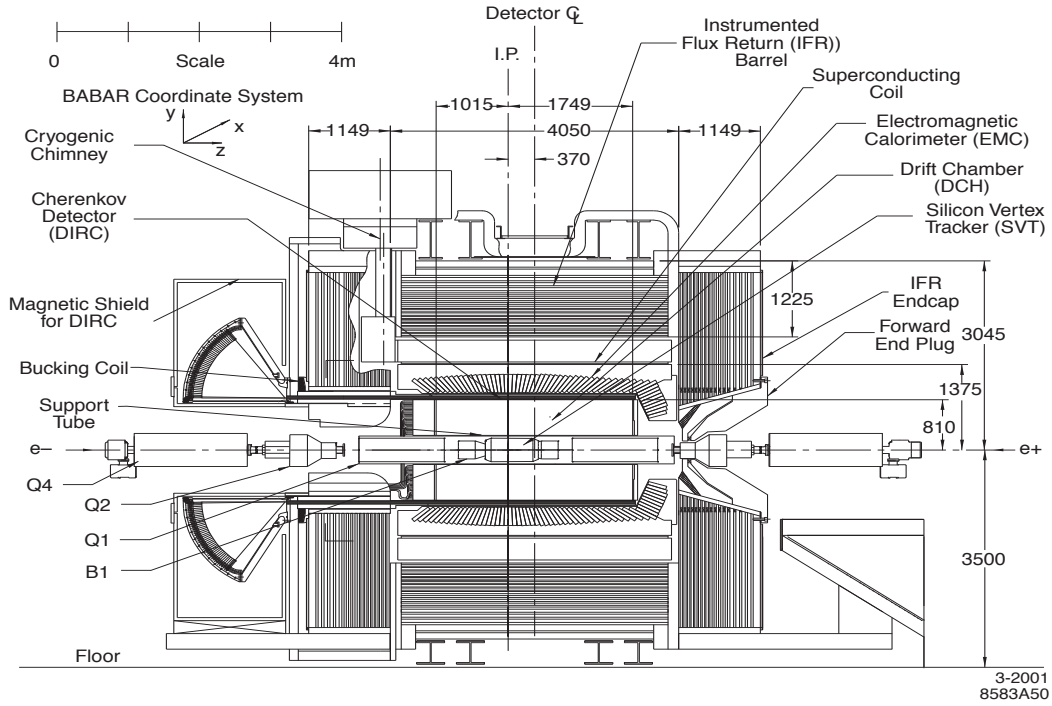


Figure 3.2: The *BABAR* detector: longitudinal cross section [22].

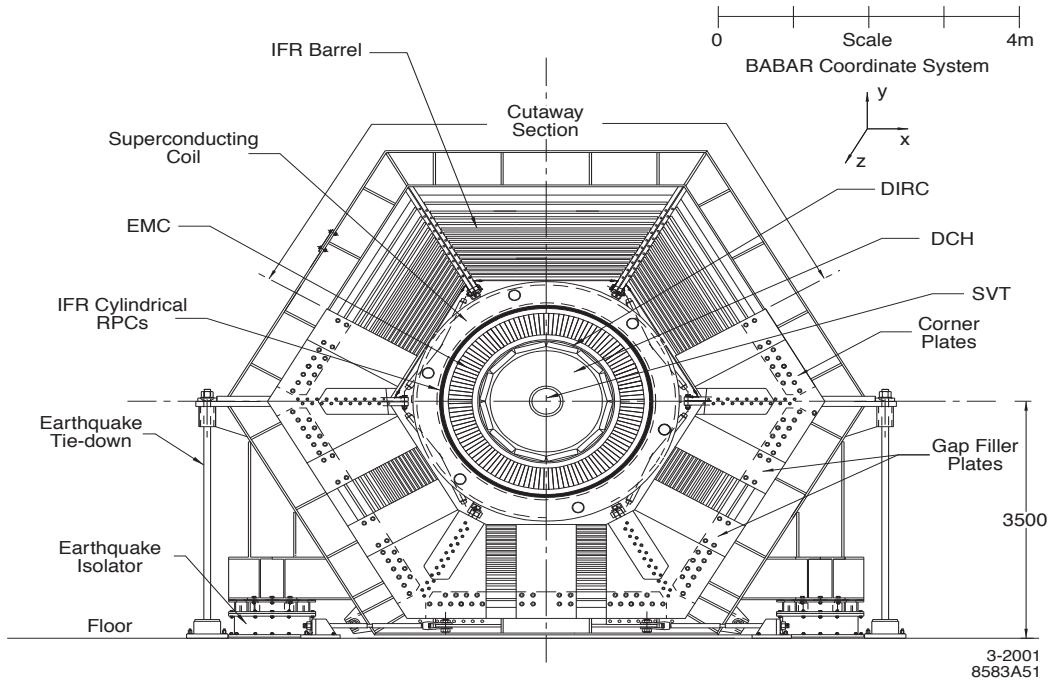


Figure 3.3: The end view of the *BABAR* detector [22].

3.2.1 Silicon vertex tracker

The main purpose of the SVT was to measure the position and angle of charged particles, thus allowing for precise reconstruction of their trajectories and decay vertices. Charged particles traversing the silicon strips deposited energy according to the Bethe-Bloch formula [18]. The deposited energy excited the electrons in the silicon atoms, creating electron-hole pairs. These then drifted to the nearest electrode, where the signal was collected and amplified, thus allowing position determination.

In the *BABAR* experiment, the SVT was required to have high vertex resolution ($\sim 80\mu\text{m}$) to separate the decay positions of the two B mesons, and high radiation tolerance to handle the expected luminosity from the PEP-II storage rings. Furthermore, because B decay products have an average momentum of less than $1\text{ GeV}/c$, many tracks did not reach the *BABAR* drift chamber due to the applied magnetic field. This implied that the SVT must be capable of stand-alone particle identification of such charged particles [65].

The *BABAR* SVT was positioned closest to the interaction region, within the support tube - a cylindrical structure with a 20 cm radius designed to support the beam pipe. The SVT provided an angular coverage from 20.1° to 150.2° in polar angle. It was composed of 5 cylindrical layers of double-sided silicon strip detectors, which were assembled in modules with readout at each end. The three inner layers consisted of 6 modules each and primarily provided position and angle information with a spatial resolution of $10\text{-}15\text{ }\mu\text{m}$ for perpendicular tracks. The outer two layers had 16 and 18 modules respectively, and provided coordinates and angle measurements that linked the SVT to the DCH with a resolution of $\sim 30\text{-}40\mu\text{m}$. Each SVT module consisted of several silicon detectors and was divided into two electrically independent halves which were read out by electronics at the forward and backward ends [22].

A schematic of the SVT longitudinal section is shown in Fig. 3.4. As can be readily seen, silicon wafers farthest from the IP were bent into a “lampshade” arrangement to reduce the angle of incidence in the forward and backward regions and thus improve the resolution [65]. In total, 340 double-sided silicon strip detectors were used with 6 different wafer geometries

and approximately 150,000 readout channels. Furthermore, signals were collected such that the output width of the pulse (ToT) is a quasi-logarithmic function of the collected charge. Measurements of ToT value yielded the pulse height and thus the ionization, dE/dx , in the silicon sensor. Up to ten measurements of dE/dx were obtained per charged particle trajectory, otherwise referred to as track, and a 60% truncated mean dE/dx was calculated for charged particles with signal from at least 4 strips in the SVT .

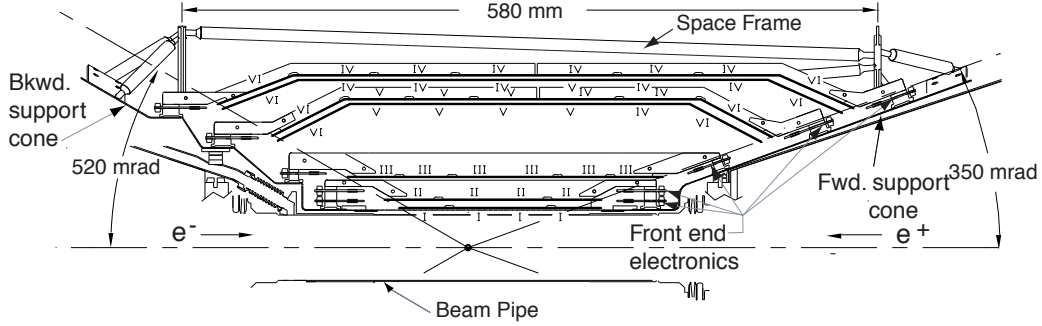


Figure 3.4: Longitudinal view of the SVT, where the roman numerals stand for the different types of the silicon detectors [65].

3.2.2 Drift chamber

The main component of a drift chamber is a gas-filled volume containing field wires, used to maintain an electric field, and sense wires, used to detect ionization electrons. A charged particle passing through the gas will interact with the atoms and nuclei and deposit energy, producing free electrons and ions. In the presence of an electric field, these electrons will drift to the positively charged electrode, undergoing repeated collisions with the gas molecules. The electric field is strong enough, such that the ionization electrons can acquire enough energy to knock additional free electrons from the gas molecules. The additional electrons can then also cause more ionization, producing an avalanche with an exponentially increasing number of electrons. This avalanche will then arrive to the sense wires in the form of a measurable current, which is proportional to the original number of created ions. The drift

velocity of an electron is dependent on the electric field and therefore a measurement of the time it took for an electron to arrive at the cathode allows for a determination of the distance of the original source particle from the electrode and thus a position measurement. Furthermore, by combining measurements from multiple sense wires, the trajectory (track) of the charged particle can be determined.

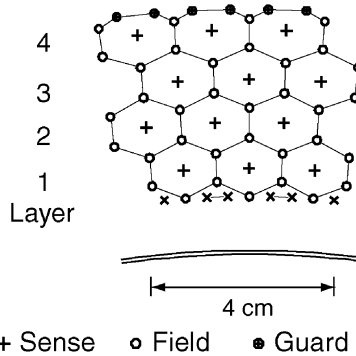


Figure 3.5: Schematic layout of drift cells for the four innermost layers [67].

The *BABAR* drift chamber was the second component of the tracking system. Its main purpose was to measure the momentum and energy loss, dE/dx , of charged particles. Some particles originated outside the volume of the SVT, such as K_s^0 mesons, and thus relied solely on information from the drift chamber for identification and reconstruction. To meet the physics requirements of the experiment, the *BABAR* DCH provided a spatial resolution better than $140 \mu\text{m}$, averaged over the cells. For separation between K and π mesons, with energies up to $700 \text{ MeV}/c^2$, the required resolution for dE/dx measurements, combined over all hits, was $\sim 7\%$ [66].

At *BABAR*, the DCH was made up of 40 layers of small hexagonal cells, defined by field wires, with a low density gas (helium:isobutane = 80%:20%). A schematic of the four innermost layers of the DCH is shown in Fig. 3.5. The cells allowed for 40 position and dE/dx measurements for tracks with transverse momenta, p_t , greater than $180 \text{ MeV}/c$. Furthermore, “stereo” wires in 24 of the 40 layers were oriented at small angles to the z -axis to provide longitudinal position information. Fig. 3.6 shows a schematic diagram of the drift chamber’s

longitudinal cross-section. The layers were combined into a 3 m long cylindrical structure surrounding the silicon vertex tracker, with a radius of 0.8 m and centre offset by 370 mm from the IP. The inner cylindrical wall of the DCH was thin to minimize multiple scattering and allow for track matching with the SVT. For this reason, all the read-out electronics were mounted on the backward endplate of the chamber.

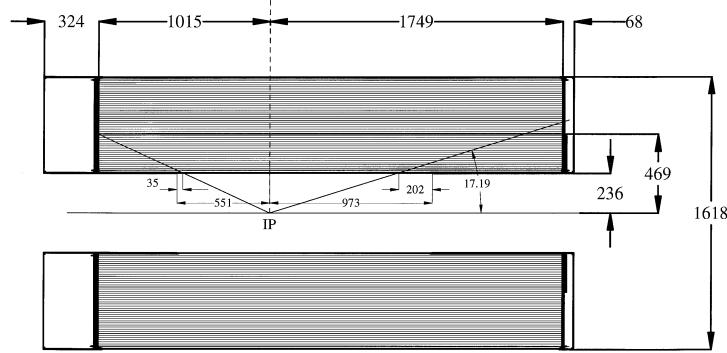


Figure 3.6: Side view of the *BABAR* drift chamber. Dimensions are shown in mm [66].

The DCH was designed to provide a measurement of the drift time and integrated charge for every sense wire with a signal. With the gas mixture used, there were on average 22 primary and 44 total ionization clusters per cm of track length [67]. The position of the primary ionization was determined from the timing of the leading edge of the amplified signal. The measured time and drift distance relation was established using samples of e^+e^- and $\mu^+\mu^-$ events, and thus with a time measurement, the position of the track was determined.

Furthermore, the energy loss, dE/dx , per track was derived from the measurement of the total charge deposit in every drift cell. Fig. 3.7 shows the measured dE/dx in the DCH as a function of the particle's momenta, made using large samples of beam background events. As can be readily seen, the measurements agree with the predicted Bethe-Bloch curves, also shown in Fig. 3.7.

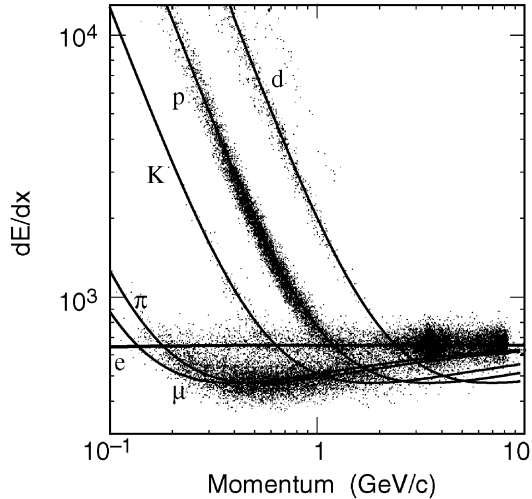


Figure 3.7: Measurement of dE/dx in the DCH as a function of the track momenta. The curves show the Bethe-Bloch predictions determined from selected control samples [67].

3.2.3 Detector of internally reflected Cherenkov light

The DIRC was a novel type of ring-imaging Cherenkov detector, which served as a particle identification (PID) system with a fast signal response and a tolerance to high backgrounds. High quality particle identification of kaons and pions over a wide range of solid angle and momentum is necessary for flavour-tagging and CP-violation studies. The *BABAR* DIRC was required to provide K/π separation of $\sim 4\sigma$ within the momentum range 500 MeV/ c to 4.5 GeV/ c . Furthermore, because the PID system was surrounded by an electromagnetic calorimeter, the DIRC was thin and uniform in radiation lengths to minimize degradation of the calorimeter energy resolution [22].

The *BABAR* DIRC consisted of 4.9 m long bars, made of synthetic fused silica with a rectangular cross section of 1.7×4.3 cm. A charged particle with velocity v traversing the silica bar with index of refraction $n \sim 1.474$ generated a cone of Cherenkov photons if $v > c/n$, where c is the speed of light in vacuum. The generated cone has a half angle θ_c with respect to the particle direction, and thus $\cos \theta_c = 1/\beta n$, where $\beta = v/ct$. Fig. 3.8 shows a schematic of the imaging principle in the DIRC. The generated photons were transported to each end of the silica bar through successive total internal reflection. Furthermore, a mirror

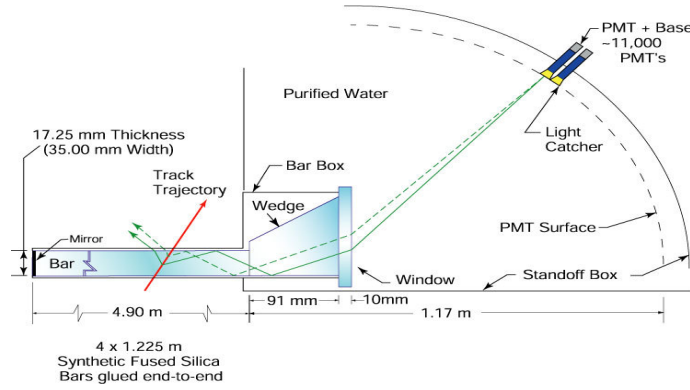


Figure 3.8: Imaging principle and the transport of photon signals in the DIRC [67].

was placed at the front end of each bar to reflect incident photons to the backward end, where a detection apparatus was set up.

Because the bars had a rectangular cross section and were finished to optical precision, reflections off the surface do not affect the magnitude of the Cherenkov angle, except for up-down left-right ambiguities. A fused silica wedge was glued to the end of the bar, as shown in Fig. 3.8, such that rays with an odd or even number of reflections are directed to the same point in the photon detection plane. The photons exited the bar and emerged into a water-filled expansion region known as the stand-off box, where the Cherenkov image is allowed to expand. Purified water was used because it has an index of refraction ($n \approx 1.34$) close to that of the silica and thus total internal reflection at the interface is reduced. The stand-off box was composed of a stainless-steel cone and cylinder along with 12 sectors of photomultiplier tubes (PMTs) and about 6000 litres of pure water [67]. Each of these sectors consisted of 896 PMTs, which lay on an almost toroidal surface. The reflected photons were detected by these densely packed PMTs, located a distance of 1.2 m from the bar end.

The expected Cherenkov light pattern was a cone section, with the opening angle being the Cherenkov production angle. Images of the Cherenkov rings were reconstructed from the position and time of arrival of the photons at the PMTs. Using the Cherenkov angle, the velocity of a charged particle was determined. Combining this with the momentum measurements of the tracking system, the mass of the particle was calculated, thus revealing its

identity [22]. The performance of the *BABAR* DIRC is shown in Fig. 3.9. As can be readily seen, K/π separation was achieved up to a momentum of 4.5 GeV.

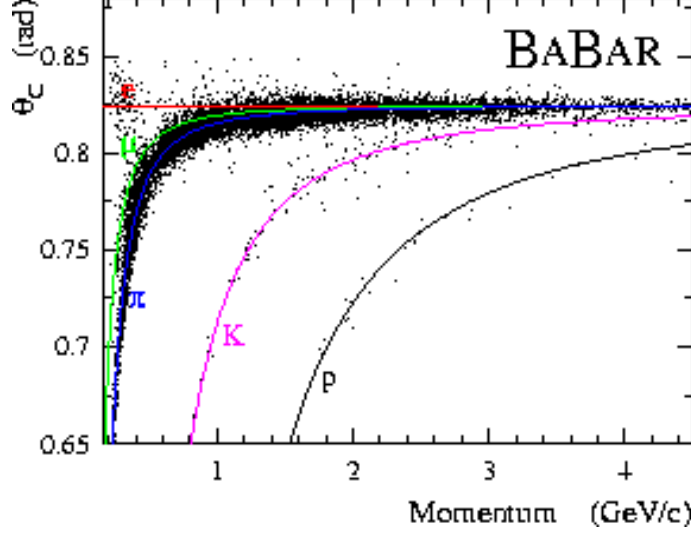


Figure 3.9: Cherenkov angle as a function of the particle's momentum [68].

3.2.4 Electromagnetic calorimeter

The EM calorimeter was designed to measure electromagnetic showers from photons and charged particles, with energy from 20 MeV to 9 GeV. This range allowed for the detection of low energy π^0 and η mesons and high energy electrons and photons. By identifying electrons, the EMC contributed to the flavour tagging of B mesons in semi-leptonic decays and measurements of rare B and D meson decays, as well as QED processes. To allow for such studies, an EMC energy resolution of ~ 1 to 2% was required along with an angular resolution of a few mrad. To meet these requirements, a hermetic, total absorption calorimeter composed of a finely segmented array of thallium-doped Cesium Iodide crystals was used. The energy resolution of a homogeneous crystal calorimeter is given by

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus b, \quad (3.1)$$

where E and σ_E are the energy, in GeV, of a photon and its RMS error, respectively. a is the energy-dependent term which arises primarily due to electronic noise and photon statistics. It is dominant at low energies. At high energies, the term b arises due to the non-uniformity in the light collection, leakage or absorption in the crystal material, and uncertainties in the calibration [67].

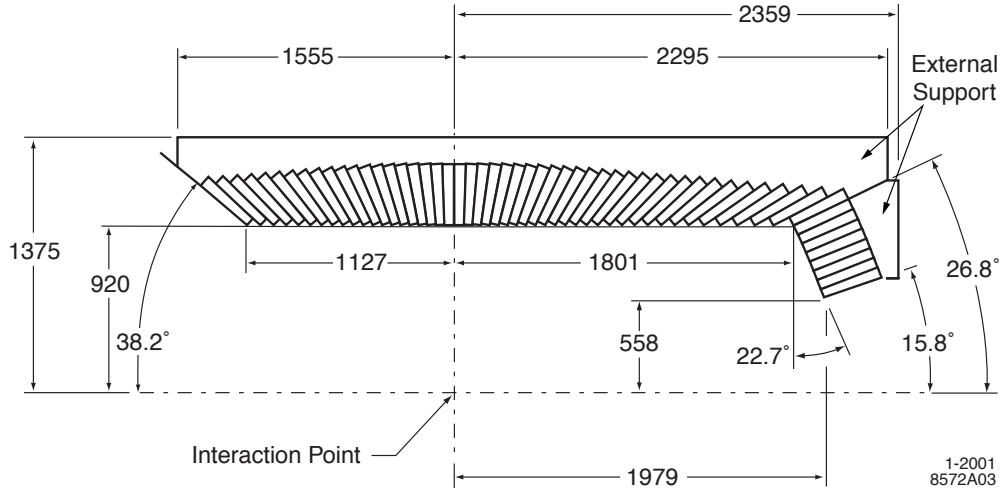


Figure 3.10: Longitudinal cross section of the EMC (top half is shown here) showing the arrangement of the 56 crystal rings. Dimensions are given in mm [67].

At *BABAR*, thallium-doped CsI was chosen because of its high light yield (50,000 γ /MeV), small Molière radius (3.8 cm) and other properties which allowed for excellent energy and angular resolution, as well as shower containment within the compact design. The EMC consisted of two sections: a barrel and a forward end-cap, thus covering a polar angle from 15.8° to 141.8° and the full azimuthal angle. There were 5,760 crystals in the barrel structure, arranged in 48 distinct rings, whereas the end-cap carried 820 crystals in 8 such rings. The typical area of a crystal is $4.1 \times 4.7 \text{ cm}^2$ for the front face and $6.1 \times 6.7 \text{ cm}^2$ for the back face, while the transverse dimension varied to provide the required hermetic coverage. A longitudinal cross section of the EMC and the layout of these rings is shown in Fig. 3.10. Furthermore, the crystals, shown in Fig. 3.11, had a tapered trapezoidal structure, with

their length increasing in the forward direction in order to prevent leakage from increasingly high energy particles. They were inserted into trapezoidal modules, which were bonded to an aluminum strong back and mounted to the external support. Each crystal was read out individually, using a pair of silicon PIN diodes. These were in turn connected to a preamplifier.

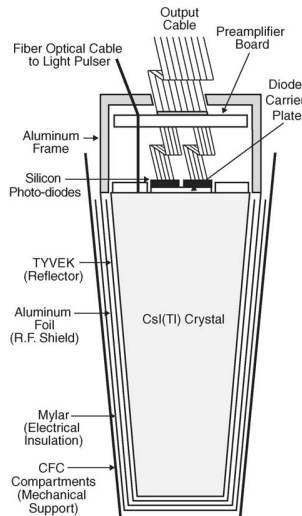


Figure 3.11: Schematic of a CsI crystal with its readout electronics [67].

A typical EM shower spreads over many adjacent crystals, depositing clusters of energy. Pattern recognition algorithms were developed to distinguish single from merged clusters and charged from neutral particles [67]. A cluster was associated with a charged particle, as reconstructed by the inner tracking system, by measuring the distance between the particle's trajectory and the cluster's centroid. If it was consistent, then the cluster was assigned to the charged particle. Otherwise, the cluster was assumed to originate from a neutral particle. The energy resolution of the *BABAR* EMC is shown in Fig. 3.12 as a function of energy.

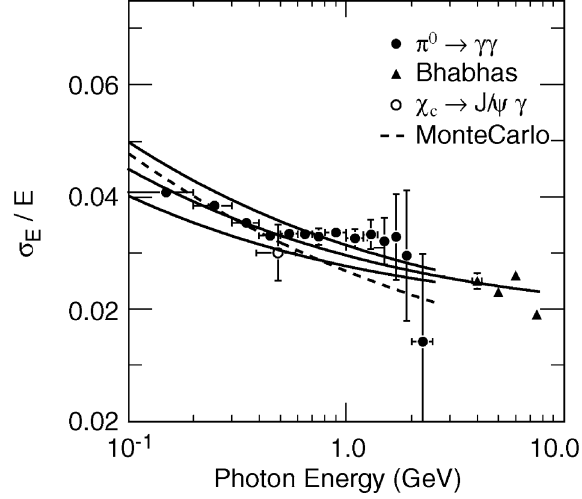


Figure 3.12: Energy resolution of the EMC as a function of energy, determined using various processes. The fit is given by Eq. (3.1) $a = 2.32 \pm 0.30\% \sqrt{\text{GeV}}$ and $b = 1.85 \pm 0.12\%$ [67].

3.2.5 Instrumented flux return

Muons and neutral hadrons, which are essential for flavour-tagging and CP-violation studies, were identified using the Instrumented Flux Return (IFR). To achieve its physics goals, the IFR was required to be efficient, with a large solid angle coverage and a background rejection for muons with momenta down to $< 1 \text{ GeV}/c$ [67]. A schematic diagram of the IFR structure is shown in Fig. 3.13.

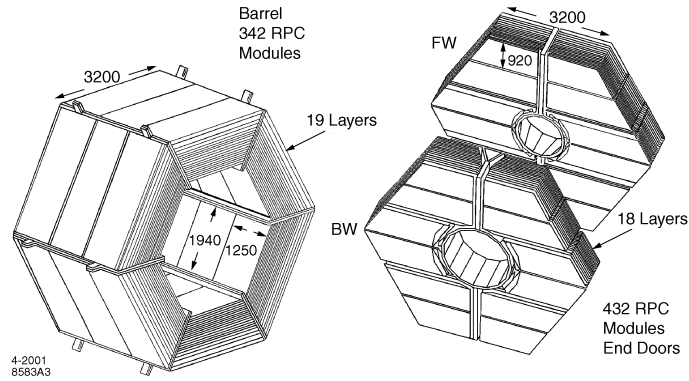


Figure 3.13: Diagram of the different sections of the IFR (barrel and end doors) and the shape of its modules [67].

The magnet flux return steel in the barrel and the two end-caps was segmented into

layers and functioned as both a muon filter and a hadron absorber. Each gap between the absorbers was filled with a layer of single gap resistive plate chambers (RPCs). RPCs are gaseous detectors composed of two oppositely charged parallel-plates. The plates are made of a highly resistive plastic material and separated by a gas volume. Fig. 3.14 displays the cross section of a planar RPC and its voltage connections. The outer surfaces of the parallel plates were coated with conductive graphite and connected to ~ 8 kV potential. Muons passing through the RPC ionized the gas, causing an avalanche of free electrons. This formed a signal which was then picked up by external electrodes made of aluminum strips after a small but precise time delay. The pattern of hit strips gave a quick measure of the muon momentum. In general, muons penetrated more layers than neutral hadrons. Therefore, if hits in multiple IFR layers were present and can be linked to a charged particle track in the SVT and DCH, then the track in question was more likely a muon or a charged hadron. Otherwise, the particle was a neutral hadron interacting with the steel of the magnetic flux return.

In total, there were 19 RPC layers in the barrel, 18 in the end-caps, and two additional cylindrical RPC's inserted between the EMC and the magnet cryostat to detect particles leaving the calorimeter. Furthermore, due to efficiency losses and rapid aging, the Barrel RPC's were replaced by Limited Streamer Tubes (LST's) during the summer of 2006. The latter were composed of silver plated wires which collect the free charge in a CO_2 based gas mixture.

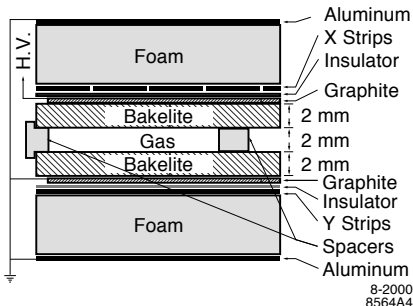


Figure 3.14: Cross section of a planar Resistive Plate Chamber [67].

3.2.6 Trigger system

The *BABAR* trigger system was responsible for selecting events of interest resulting from the e^+e^- collisions in the PEP-II storage ring. At *BABAR*'s design luminosity, beam-induced backgrounds were very high, with a rate of ~ 20 kHz each for one or more tracks with $p_t > 120$ MeV/ c in the DCH or at least one EMC cluster with $E > 100$ MeV [67]. The goal of the trigger system was to reject these background events while keeping the total event rate below 120 Hz. Furthermore, the total trigger efficiency was expected to exceed 99% for $B\bar{B}$ events and 95% for continuum events.

To achieve these goals, the trigger system operated as a sequence of two independent levels, one hardware (called L1) and one software (called L3). The first level, L1, used track information from the DCH and IFR, along with cluster information from the EMC to trigger on events. It thus consisted of 4 subsystems: the Drift Chamber Trigger (DCT), the Electromagnetic Calorimeter Trigger (ECT), the Instrumented Flux Return Trigger (IFT), and the Global Level Trigger (GLT). The DCT and ECT received information from their associated detector subsystems and forwarded it to the GLT, in the form of trigger *primitives*, which were summary data on the position and energy of particles. The GLT then processed all trigger primitives, generated the L1 trigger and forwarded it to Fast Control and Timing System (FCTS). The DCT and EMT both satisfied the full trigger requirements independently, providing a sufficient level of redundancy. The IFT was mainly used to trigger on

$\mu^+\mu^-$ and cosmic ray events.

Based on both the complete event and L1 trigger information, the L3 software algorithms determined whether or not an event was stored for processing. In addition to physics filters, the L3 trigger also vetoed Bhabha events and identified special categories of events, which were needed for calibration and luminosity measurements. After the L3 filtering, the L1 triggers were reduced by a factor of ~ 10 . Its output rate was limited to 120 Hz, ~ 90 Hz for physics events and ~ 30 Hz for events of special categories, such that it did not overload the downstream storage and processing capacity.

3.3 Event reconstruction

The *BABAR* reconstruction software [69] is organized as a set of Modules, which apply a sequential processing of the data in an Event [76]. The reconstruction packages are used for processing and particle identification at the subdetector level, as well as pattern recognition, and fitting.

Charged particle trajectories (tracks) are reconstructed separately in the SVT and DCH using different algorithms. The separately-found DCH and SVT tracks are then projected and matched. Those that match are combined into a single track and placed in the output list of good tracks. SVT or DCH tracks that fail to merge into a single track are also copied to the output list. These tracks in the output list are then fitted with the mass hypothesis of a pion. If a track is merged, its parameters are based on the weighted average of the two input tracks from the SVT and DCH [76]. Furthermore, the DIRC reconstruction provides, for each track that intersects it, an estimate of the Cherenkov angle and its error. This information is then used in different particle identification (PID) algorithms to formulate a particle hypothesis for each track.

A cluster is the energy deposit caused by particle interactions with the EMC. The reconstruction of clusters starts from a set of adjacent crystals, with the sum of their deposited

energy above a threshold of 20 MeV. A cluster can be split into bumps, where a bump represents the fraction of the cluster deposited by a single particle interaction. Bumps are identified as local maxima within the cluster. In the case of one local maximum, clusters and bumps are identical. Furthermore, bumps are also distinguished from shower fragments that manifest themselves as additional local maxima in the cluster. All bumps and clusters are copied as candidates, with a calibrated energy and position assuming a photon particle hypothesis. In addition, calorimeter clusters are matched geometrically to nearby charged tracks, to separate between charged and neutral particles.

Additional particle lists are also created from the primitive track and cluster lists. These include additional information such as particle ID or particle vertexing. Furthermore, specific candidate lists are also formed using various combinations of tracks or clusters, such as π^0 lists created by combining pairs of clusters or B meson lists formed using multiple track and cluster combinations, various PID inputs and mass hypotheses, and kinematic fits of entire decay trees.

Chapter 4

Analysis Tools

The goal of this analysis is to measure the branching fraction of the SM suppressed decay $B^+ \rightarrow K^+ \tau^+ \tau^{-1}$ using data from the *BABAR* experiment. Signal Monte Carlo samples are used to select a series of cuts that can separate signal and background events. The latter are studied using generic background Monte Carlo samples. To reduce the potential for experimental bias, the analysis is carried out blinded, which means that access to the data in the signal region is only possible after developing and optimizing the entire signal selection. Furthermore, a method called hadronic B_{tag} reconstruction is implemented in order to separate between the two B daughters, such that signal events can be selected with high purity.

4.1 Hadronic B_{tag} reconstruction method

At *BABAR*, electron and positron beams collide at a CM energy of 10.58 GeV. This corresponds to the mass of $\Upsilon(4S)$, which almost exclusively decays into $B\bar{B}$ pairs. To measure $B^+ \rightarrow K^+ \tau^+ \tau^-$, hadronic B_{tag} reconstruction is applied. Using this approach, the decay of one of the two B mesons is reconstructed exclusively using hadronic modes, as shown in Fig. 4.1. This B is referred to as the B_{tag} . The remaining information in the event, charged

¹The charge conjugate mode, $B^- \rightarrow K^- \tau^+ \tau^-$ is also implied.

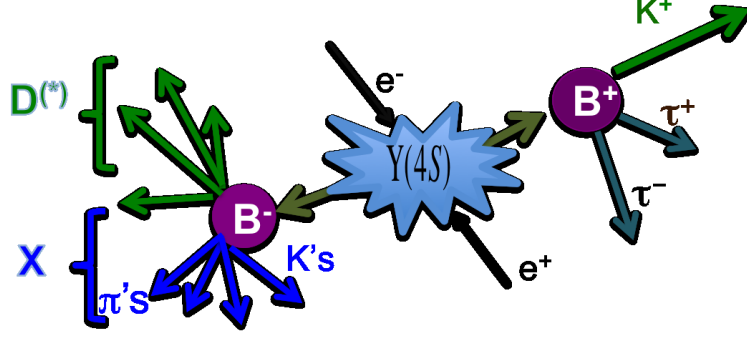


Figure 4.1: Hadronic B_{tag} reconstruction approach.

particle trajectories and energy deposits, are attributed to the signal B , B_{sig} , on which the search for $B^+ \rightarrow K^+ \tau^+ \tau^-$ is applied.

There are many advantages to the hadronic B_{tag} reconstruction method. First, because B_{tag} is fully reconstructed, its four-momentum is known and thus that of the B_{sig} can be easily calculated since the total CM energy is precisely known. The rest frame of B_{sig} is also easily found, which allows for kinematic constraints on the B_{sig} daughters and the calculation of discriminating variables that are crucial in selecting for a rare B decay. In addition, hadronic B_{tag} reconstruction is ideal for decays with missing energy. This is true because any missing energy in the event can be associated with B_{sig} . In this analysis, the signal $B^+ \rightarrow K^+ \tau^+ \tau^-$ decay has a large missing energy contribution, due to the τ neutrinos and lepton neutrinos that result from the τ daughter decays. On the other hand, because the hadronic branching fraction of B mesons decay is of $\mathcal{O}(1\%)$, the signal efficiency is lower using hadronic B_{tag} reconstruction as opposed to other tagging methods. However, for a rare decay search with missing energy, the hadronic B_{tag} reconstruction is the best approach to extract signal events with high purity.

4.2 Analysis software

Simulated Monte Carlo (MC) signal and background events are used to develop a signal selection and study potential backgrounds. These events are generated with EvtGen [70] for B decays to exclusive final states and Jetset [71] for generic continuum and inclusive B decay simulation. The detector response is then simulated using Geant4 [72], which includes a detailed model of the *BABAR* detector and its response.

Software in *BABAR* is organized in releases, each with hundreds of packages used to perform specific tasks. In terms of *BABAR* software, the data and Monte Carlo samples used for this analysis are generated using physics release R26a. Recently, release 26.0.1 has also been added to the B^+B^- , $B^0\bar{B}^0$, and $c\bar{c}$ Monte Carlo samples. The result is a significantly larger sample for these MC types, which amounts to almost 10 times the data luminosity. Events in all MC samples are then reconstructed using the BSemiExclAdd skim, which is a skimming code that applies the hadronic B_{tag} reconstruction. The packages used to analyze the reconstructed events are: BRecoilUser SHR_151006, BRecoilTools SHR_151006, and BetaPID SHR_151006.

These packages are configured for analyses with a reconstructed B^\pm or B^0 . Custom ntuples are then produced from MC skim samples at SLAC and copied over to McGill University for ROOT-based analyses [73]. The ntuples contain all the relevant information on the B_{sig} daughter tracks and clusters. Information on the B_{tag} daughters is excluded, along with all B_{sig} events with more than 9 daughter tracks or 12 daughter clusters. However, the ntuples do contain all Truth MC information about the B_{tag} decay, which is information on what was generated at the simulation level for each event. Furthermore, information on the reconstructed B_{tag} candidate itself, such as its charge, energy substituted mass, and purity value, is also included in the ntuples.

4.3 Event reconstruction

B mesons decay hadronically, with a relatively large branching fraction, into charmed mesons. Therefore, the B_{tag} is reconstructed, by the BRecoiUser software package, via the decays $B \rightarrow D^{(*)0}X$ and $B \rightarrow D^{(*)\pm}X$. Here, the D is a seed meson, the reconstruction of which is described below, and the X is a combination of kaons and pions, with total charge of ± 1 , such that $X = n_1\pi + n_2K + n_3K_s^0 + n_4\pi^0$ with $n_1 + n_2 \leq 5$, $n_3 \leq 2$, $n_4 \leq 2$, and $n_1 + n_2 + n_3 + n_4 + n_5 \leq 5$. The D seed is reconstructed using the following modes:

- $D^{*\pm} \rightarrow D^0 \pi^\pm$
- $D^{*0} \rightarrow D^0\pi^0, D^0\gamma$
- $D^0 \rightarrow K^\pm\pi^\mp, K^\pm\pi^\mp\pi^0, K^\pm\pi^\mp\pi^+\pi^-, K_s^0\pi^+\pi^-, K_s^0\pi^+\pi^-\pi^0, K^+K^-, \pi^+\pi^-\pi^0, \pi^+\pi^-, K_s^0\pi^0$
- $D^\pm \rightarrow K_s^0\pi^\pm, K_s^0\pi^\pm\pi^0, K_s^0\pi^\pm\pi^+\pi^-, K^\pm\pi^+\pi^-, K^\pm\pi^+\pi^-\pi^0, K^+K^-\pi^\pm, K^+K^-\pi^\pm\pi^0$
- $D_s^{*\pm} \rightarrow D_s^\pm\gamma$
- $D_s^{*\pm} \rightarrow \phi\pi^\pm, K_s^0K^\pm$
- $K_s^0 \rightarrow \pi^+\pi^-, \pi^0 \rightarrow \gamma\gamma$, and $\phi \rightarrow K^+K^-$

A D_s^{*+} seed and a J/ψ seed are also used in this skim: $B \rightarrow D_s^{*+}X$ and $B \rightarrow J/\psi X$, where $J/\psi \rightarrow \ell^+\ell^-$ ($\ell = e$ or μ). In total, there are 2968 B decay modes used to reconstruct a B_{tag} candidate.

To ensure a properly reconstructed B_{tag} , a cut on the energy substituted mass, m_{ES} , and ΔE are applied. The former is defined as the mass of the B_{tag} , calculated using the CM energy instead of that of the B meson. This is done in order to avoid resolution uncertainties associated with measuring the energy of the B_{tag} , by exploiting the fact that the CM energy is precisely known. m_{ES} is given by the following equation:

$$m_{\text{ES}} = \sqrt{E_{\text{CM}}^2 - \vec{p}_{B_{\text{tag}}}^2} \quad (4.1)$$

Here, E_{CM} is half the total colliding energy and $\vec{p}_{B_{\text{tag}}}$ is the 3-momentum of the reconstructed B_{tag} , in the CM frame. ΔE is the difference between the CM energy of the reconstructed B_{tag} candidate and the CM beam energy. It is given by the equation:

$$\Delta E = E_{\text{CM}} - E_{B_{\text{tag}}} \quad (4.2)$$

A properly reconstructed B_{tag} must have a m_{ES} consistent with the mass of the B meson and thus within the range $5.18 < m_{\text{ES}} < 5.30 \text{ GeV}/c^2$. Furthermore, ΔE should be in the range $-0.2 < \Delta E < 0.2 \text{ GeV}$.

Combining the D or J/ψ seed meson with kaons and pions can lead to more than one B_{tag} candidate passing the BSemiExclAdd skim. In this case, tighter cuts on m_{ES} and ΔE are applied according to the “BestB” selection: $5.20 < m_{\text{ES}} < 5.30$ and $-0.12 < \Delta E < 0.12$. If more than one B candidate passes the “BestB” selection, the B_{tag} decay mode with the highest purity will be chosen. Here, purity is determined from MC studies and is defined as the expected fraction of properly reconstructed B_{tag} candidates with $m_{\text{ES}} > 5.27 \text{ GeV}/c^2$ for any given mode. This purity is calculated by the authors of the BSemiExcl and BSemiExclAdd skims [74], and is referred to as the “high multiplicity” purity. It is calculated using a fit to the m_{ES} distribution, where an ARGUS [78] function accounts for the mis-reconstructed B_{tag} candidates and a Crystal Ball [78] function is used to model the peaking component of the m_{ES} distribution and is composed of properly reconstructed B mesons. This “high multiplicity” purity is different from the “low multiplicity” purity that is calculated specifically for this analysis and is further discussed in section 5.1.2. Finally, if more than one B_{tag} candidates passes the “BestB” selection and have the same purity value, the B_{tag} candidate that has the lower $|\Delta E|$ is chosen.

4.4 *BABAR* dataset

This analysis uses on-peak *BABAR* data, recorded at a CM energy corresponding to the $\Upsilon(4S)$ resonance, with a total integrated luminosity of 424.43 fb^{-1} [75]. The data sample is divided into 6 runs, each run corresponding to a different set of detector conditions. The luminosity and resulting number of $B\bar{B}$ pairs, also referred to as B -counting [76], for each run are summarized in Table 4.1. The errors quoted here are statistical only.

Run #	Data Set	Luminosity (fb^{-1})	B -counting ($\times 10^6$)
1	Lumi-BSemiExcl-Run1-OnPeak-R24c	20.374	22.556 ± 0.005
2	Lumi-BSemiExcl-Run2-OnPeak-R24c	61.322	68.439 ± 0.008
3	Lumi-BSemiExcl-Run3-OnPeak-R24c	32.279	35.751 ± 0.006
4	Lumi-BSemiExcl-Run4-OnPeak-R24c	99.606	111.430 ± 0.003
5	Lumi-BSemiExcl-Run5-OnPeak-R24c	132.371	147.620 ± 0.012
6	Lumi-BSemiExcl-Run6-OnPeak-R24c	78.308	85.173 ± 0.009

Table 4.1: Luminosity and B -counting values of the *BABAR* dataset [75].

4.5 Background Monte Carlo

Generic background Monte Carlo plays an important role in testing the validity of the signal selection and understanding potential sources of background. There are five different types of background MC: B^+B^- , $B^0\bar{B}^0$, $c\bar{c}$, $q\bar{q}$ ($q=u,d,s$) (also referred to as uds), and $\tau^+\tau^-$. Each type is split into six runs, to correspond with the detector conditions that were present during data collection. The number of generated and skimmed events for each of the MC types is shown in Table 4.2, along with the equivalent cross-section and normalization.

According to how these events enter into the B_{tag} sample, the background MC can be further divided into three categories:

- Peaking MC background: These are B^+B^- MC events where a $B\bar{B}$ pair is formed from the $\Upsilon(4S)$ decay and a charged B_{tag} candidate has been properly reconstructed. The

reconstructed B_{tag} candidates have a peaking distribution within the m_{ES} signal region, which for the purpose of this analysis is defined between $5.27 \text{ GeV}/c^2$ and $5.29 \text{ GeV}/c^2$.

- Combinatorial MC background: B^+B^- and $B^0\bar{B}^0$ MC events where an $\Upsilon(4S)$ decays into a $B\bar{B}$ pair, but a properly reconstructed B_{tag} candidate has not been formed. Such events do not have a peaking distribution in the m_{ES} signal region.
- Continuum MC background: $c\bar{c}$, $q\bar{q}$ ($q=u,d,s$), and $\tau^+\tau^-$ MC events where e^+e^- do not collide to form a $\Upsilon(4S)$ resonance but instead $e^+e^- \rightarrow q\bar{q}$ where $q = u, d, s, c$ or $e^+e^- \rightarrow \ell^+\ell^-$ where $\ell = e^\pm, \mu^\pm, \tau^\pm$. Such events consist of tracks and clusters that can be combined to form a B_{tag} candidate. However, the resulting B_{tag} candidate is not a real B meson. The mass distribution of such events usually dominates the m_{ES} sideband region, which is defined between 5.21 and $5.26 \text{ GeV}/c^2$.

Fig. 4.2 shows the signal and sideband regions of a sample m_{ES} distribution of data and MC background, after a set of preliminary cuts. Because this is a charged B decay, only B^+B^- MC has a peaking distribution in the m_{ES} signal region.

The B^+B^- , $B^0\bar{B}^0$, and $c\bar{c}$ MC have almost ten times the statistics of the data, while the uds and $\tau^+\tau^-$ MC have about 4 times the number of data events. $c\bar{c}$ and $\tau^+\tau^-$ MC are separated from the other $e^+e^- \rightarrow f\bar{f}$ MC, where f is a lepton or quark. This is done because the $c\bar{c}$ MC can result in properly reconstructed D seeds, while the $\tau^+\tau^-$ MC can decay hadronically and thus lead to mis-reconstructed B_{tag} candidates. However, the contribution of the $\tau^+\tau^-$ MC is quite small and is almost negligible throughout the signal selection.

Furthermore, as shown in Table 4.2, each MC type is weighted such that it matches the data luminosity for the corresponding run. These weights are calculated as follows:

$$\text{weight} = \frac{\mathcal{L}_{\text{data}}}{\mathcal{L}_{\text{MC}}} \quad \text{where} \quad \mathcal{L}_{\text{MC}} = \frac{N_{\text{events}}}{\sigma_{\text{MC}}} \quad (4.3)$$

where $\mathcal{L}_{\text{data}}$ and \mathcal{L}_{MC} are the data and Monte Carlo luminosities respectively, N_{events} is the number of generated events per run for each Monte Carlo type and σ_{MC} is the cross-section

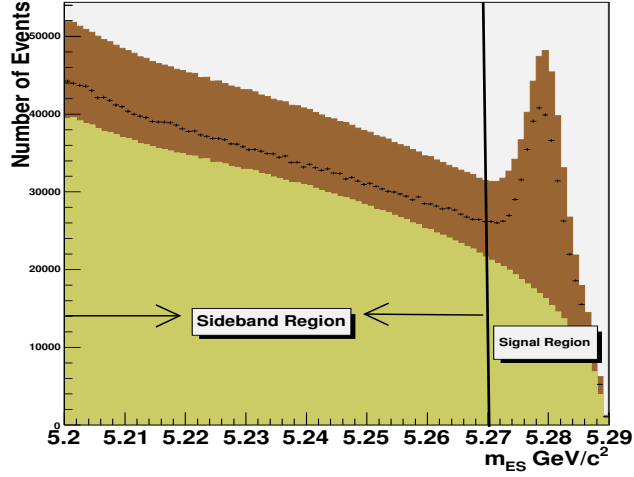


Figure 4.2: m_{ES} distribution of B_{tag} candidates in data (points), B^+B^- (brown) and $B^0\bar{B}^0$ + continuum background (yellow), with signal and sideband regions highlighted. This distribution contains events with a *charged* B_{tag} candidate and three signal-side tracks. These events have also passed the purity and continuum likelihood cut, which will be discussed in detail in the upcoming sections.

of the process modeled in each Monte Carlo sample.

The cross-sections of the continuum Monte Carlo are obtained from the *BABAR* physics book [76]. The B^+B^- and $B^0\bar{B}^0$ cross sections are calculated assuming an equal number of generated charged and neutral $B\bar{B}$ events. The B -counting value of each run, given in Table 4.1, is thus divided by 2 and the cross section is calculated as the ratio of the B -counting value to the data luminosity:

$$\sigma = \frac{B_{\text{count}}}{2 \times \mathcal{L}_{\text{data}}} \quad (4.4)$$

This way $B\bar{B}$ Monte Carlo is effectively normalized to the B -counting yields in data. In all the figures of this document, each Monte Carlo background sample is scaled by the appropriate weight to match the data luminosity.

Table 4.2: Generic Background Monte Carlo Information

Run #	Generated Events($\times 10^6$)	Skimmed Events	Skim Efficiency (%)	Cross-Section (pb)	Normalization weight
<i>B⁺B⁻ Monte Carlo SP-1235</i>					
1	113.877	42851548	37.630 ± 0.007	548	0.099
2	340.106	127153897	37.387 ± 0.004	551	0.101
3	176.806	67104199	37.954 ± 0.005	547	0.101
4	556.454	210706801	37.866 ± 0.003	553	0.100
5	724.256	270883359	37.401 ± 0.003	551	0.102
6	431.176	163900008	38.012 ± 0.003	539	0.099
<i>B⁰\bar{B}^0 Monte Carlo SP-1237</i>					
1	113.501	39955536	35.203 ± 0.007	548	0.099
2	349.964	122262783	34.936 ± 0.004	551	0.098
3	180.262	64089470	35.554 ± 0.005	547	0.099
4	553.458	195903930	35.396 ± 0.003	553	0.101
5	761.07	265477026	34.882 ± 0.002	551	0.097
6	429.68	152373995	35.462 ± 0.003	539	0.099
<i>c\bar{c} Monte Carlo SP-1005</i>					
1	266.961	69604868	26.073 ± 0.004	1300	0.100
2	797.386	208169539	26.106 ± 0.002	1300	0.101
3	439.931	116096984	26.390 ± 0.003	1300	0.097
4	1297.32	346202217	26.686 ± 0.002	1300	0.101
5	1677.53	445755971	26.572 ± 0.001	1300	0.104
6	1017.42	278228302	27.346 ± 0.002	1300	0.101
<i>uds Monte Carlo SP-998</i>					
1	166.591	29124036	17.482 ± 0.003	2090	0.258
2	482.575	84856167	17.584 ± 0.003	2090	0.269
3	276.381	49077341	17.757 ± 0.003	2090	0.247
4	755.839	136791199	18.098 ± 0.002	2090	0.279
5	1071.84	194428338	18.140 ± 0.001	2090	0.261
6	647.762	121893210	18.818 ± 0.002	2090	0.255
<i>$\tau^+\tau^-$ Monte Carlo SP-3429</i>					
1	74.665	59325	0.0795 ± 0.0003	940	0.259
2	215.775	182738	0.0847 ± 0.0003	940	0.270
3	117.694	100824	0.0857 ± 0.0003	940	0.261
4	360.242	335622	0.0932 ± 0.0002	940	0.263
5	474.008	469996	0.0992 ± 0.0001	940	0.266
6	363.346	384601	0.0106 ± 0.0002	940	0.204

Table 4.3: Number of generated and skimmed events for each run of each type of the generic background Monte Carlo.

4.6 Signal Monte Carlo

Signal Monte Carlo samples are essential for the development of a proper signal selection, as well as calculating an accurate signal efficiency. In these samples, $\Upsilon(4S)$ decays into a B^+B^- pair, where one B decays generically according to its modeled branching fractions [18] and the other B decays specifically via the signal mode, $B^\pm \rightarrow K^\pm \tau^+\tau^-$. In this analysis, two types of signal MC samples are used: “generic” and “cocktail”, as outlined in Table 4.4. These are generated using the BTOSLLBALL model [55], which utilizes a Light Cone Sum Rule (LCSR) approach to estimate the form factors that enter into the $B^+ \rightarrow K^+ \tau^+\tau^-$ exclusive decay. Other theoretical methodologies have been developed for the evaluation of these form factors and include: constituent quark models [86], lattice QCD [87], and approaches based on the heavy quark symmetry [88] and analytical constraints [89]. Each approach is valid in a certain s_B region, where $s_B = m_{\tau^+\tau^-}^2/m_B^2$ is the normalized invariant mass of the $\tau^+\tau^-$ pair, and extrapolation procedures are usually employed to obtain the form factors in the full kinematical region. For instance, lattice QCD simulations are limited to high s_B regions because the daughter light quark s cannot move fast enough on the lattice and is therefore constrained to low momentum regions [90]. On the other hand, the LCSR approach makes explicit use of the large energy of the final state meson at small values of the momentum transfer to the leptons, and is thus valid in the low s_B region [54]. Because more accurate calculations of the form factors were recently made using unquenched lattice Quantum Chromodynamics, the signal MC distribution is also reweighted according to Ref. [57] to determine the final signal efficiency. This is discussed in detail in section 8.1.

As previously mentioned, the hadronic B_{tag} reconstruction decreases the final signal efficiency. Therefore, a major obstacle in achieving a proper signal selection in this analysis is the resulting low statistics of the generic signal MC after surviving the BSemiExclAdd skim. This is why cocktail samples are also generated and used in this analysis. In these samples, the other B (B_{tag}) is required by the generator to decay to a specific hadronic decay to ensure the event passes the hadronic B_{tag} reconstruction. The selected hadronic decay is

Mode	Theoretical Model	Number of Generated Events	Number of Skimmed Events	Skim Efficiency(%)
Generic $B^\pm \rightarrow K^\pm \tau^+ \tau^-$				
11521	BTOSLLBALL	3308000	703960	21.3
Cocktail: $B^\pm \rightarrow K^\pm \tau^+ \tau^-; B^\mp \rightarrow D^0 \pi^\mp, D^0 \rightarrow K^\mp \pi^\pm$				
11520	BTOSLLBALL	3563000	2021621	56.7

Table 4.4: Number of generated and skimmed events for each signal Monte Carlo sample used in this analysis.

$B^\pm \rightarrow D^0 \pi^\pm, D^0 \rightarrow K^\pm \pi^\mp$. The skim efficiency of the cocktail sample, shown in Table 4.4, is more than twice that of the generic MC after the BSemiExclAdd skim. Furthermore, a large fraction of the cocktail events surviving the BSemiExclAdd skim populate the m_{ES} signal region and thus pass the $B^+ \rightarrow K^+ \tau^+ \tau^-$ signal selection. This leads to about two orders of magnitude improvement in the final signal efficiency of the cocktail signal sample as compared to the generic signal sample.

Chapter 5

Signal Selection

A signal selection is applied to select $B^+ \rightarrow K^+ \tau^+ \tau^{-1}$ events and suppress potential backgrounds. In this analysis, only leptonic tau decays will be considered: $\tau^+ \rightarrow \bar{\nu}_\tau e^+ \nu_e$ or $\tau^+ \rightarrow \bar{\nu}_\tau \mu^+ \nu_\mu$. This will result in three signal decay modes with either e^+e^- , $\mu^+\mu^-$, or $e^+ \mu^-$ in the final state, along with the associated lepton neutrinos. The signal selection consists of B_{tag} and B_{sig} cuts. The former ensure that only events with a properly reconstructed B_{tag} are selected and are thus related to detector activity corresponding to the hadronic B_{tag} reconstruction. B_{sig} cuts are applied to identify $B^+ \rightarrow K^+ \tau^+ \tau^-$ events and are related to detector activity and event information after fully reconstructing the hadronic B_{tag} candidate. Because each τ decays into a pair of neutrinos and a lepton, B_{sig} cannot be fully reconstructed since neutrinos cannot be identified by the *BABAR* detector. However, with the hadronic B_{tag} reconstruction, any missing energy in the event must be attributed to the signal B and consequently can be ascribed to the undetected neutrinos. The B_{sig} and B_{tag} cuts are further discussed in section 5.1 and section 5.2, respectively.

To reduce the sample size, a set of cuts is applied to signal, background, and data samples to generate ntuples that are much faster to process throughout the development of the signal selection. These will be referred to as *sub-skim* cuts and involve both the B_{sig} and B_{tag} sides:

¹Charge-conjugate modes, $B^- \rightarrow K^- \tau^+ \tau^-$, are also included where the τ also decays to lepton final states.

- Event should include between 0 and 9 signal side tracks and 0 and 12 signal-side clusters.
- Event should include a properly reconstructed B_{tag} candidate, meeting the “BestB” requirements.
- B_{tag} must be charged.
- Charge of the B_{tag} must be opposite to the sum charge of signal-side tracks.

The *sub-skim* cuts have effectively 100% “marginal” efficiency. These cuts do not reject any events in the signal sample that would not have been rejected by other B_{sig} cuts anyway.

Here is an overview of the B_{tag} cuts:

- The event must contain one properly reconstructed charged B_{tag} , meeting the “BestB” requirements.
- The charge of the B_{tag} must be opposite to the sum charge of signal-side tracks.
- The energy-substituted mass (m_{ES}) of the B_{tag} must be greater than 5.27 GeV/ c^2 .
- The low-multiplicity purity of the B_{tag} must be greater than 0.4.
- The continuum likelihood ratio, for which event shape variables are used to separate between a B -meson decay and a continuum decay, should be greater than 0.5.

Here is an overview of cuts applied on the B_{sig} side:

- The missing energy in the event should be greater than zero.
- The number of signal-side tracks must be 3.
- One track must meet the requirements of kaon particle identification, and should have a charge opposite to that of the B_{tag} .

- Two tracks must meet the requirements of either electron or muon particle identification.
- The two identified leptons must be oppositely charged.
- Events with one or more reconstructed π^0 candidates are vetoed.
- Events with a photon conversion or J/ψ candidate are vetoed: the invariant mass of the lepton pair must not be less than 50 MeV/ c^2 and must not lie within the nominal mass region of a J/ψ .
- Events with a D^0 candidate are vetoed: the invariant mass of the sum of the lepton and oppositely charged kaon must not correspond to the mass of a D^0 .
- The s_B of the event must lie within the kinematically allowed region to account for the presence of a $\tau^+\tau^-$ pair.
- The output of the MLP neural network, where a set of discriminating variables are used in a multi-variate analysis technique (TMVA) to separate between background and signal events, should lie above a specific value that has been optimized separately for each mode.

A detailed description of the cuts used for this analysis is presented in the upcoming sections. Unless otherwise indicated, the plots show the *cocktail* signal MC distributions, scaled to represent a branching fraction of 10^{-3} . This is done because the high statistics in the cocktail sample allow for a better discrimination between signal and background. For plots with generic signal MC, the assumed branching fraction is 10^{-1} . Furthermore, Appendix A includes the distributions of the generic signal sample and a comparison of the two samples. The final signal efficiency is calculated using the generic signal MC to avoid any bias that might result from the use of a signal sample with a low multiplicity and cleanly reconstructed B_{tag} decay mode.

5.1 B_{tag} cuts

Every event is required to possess one B_{tag} candidate that passes the “BestB” requirements. Furthermore, because this is a search for a charged B decay, $B^+ \rightarrow K^+ \tau^+ \tau^-$, the charge of the B_{tag} must be ± 1 . The charge of the B_{tag} must be opposite to the sum charge of all tracks on the signal side. This ensures that there are no missing tracks and that the B_{sig} and the B_{tag} are oppositely charged daughters of the $\Upsilon(4S)$ in the event.

5.1.1 m_{ES} cut

A properly reconstructed B_{tag} will have a reconstructed invariant mass consistent with that of a B meson. The m_{ES} of a B_{tag} , calculated using Eq. 4.1, is thus required to be greater than $5.27 \text{ GeV}/c^2$. Fig. 5.1 shows a plot of the m_{ES} distribution for signal and background Monte Carlo after applying the sub-skim cuts. The signal, shown in red, and the B^+B^- MC show a clear peak in the region between 5.27 and $5.29 \text{ GeV}/c^2$. The peak is accompanied by a threshold function, which dominates the m_{ES} sideband region and drops to zero near the kinematic endpoint, $m_{\text{ES}} = 5.29 \text{ GeV}/c^2$. This threshold function is parametrized with an ARGUS [78] function and represents the continuum and mis-reconstructed $B\bar{B}$ background. The data distribution is also shown as dots with error bars, representing the statistical uncertainty. As can be seen, the background MC overestimates the data at this stage in the signal selection. This is a known feature of the BSemiExclAdd skim, mainly due to the large uncertainties in the hadronic branching fractions of the B meson. However, the disagreement improves at later stages in the analysis, and is compensated for by applying the m_{ES} sideband substitution (see section 7.1).

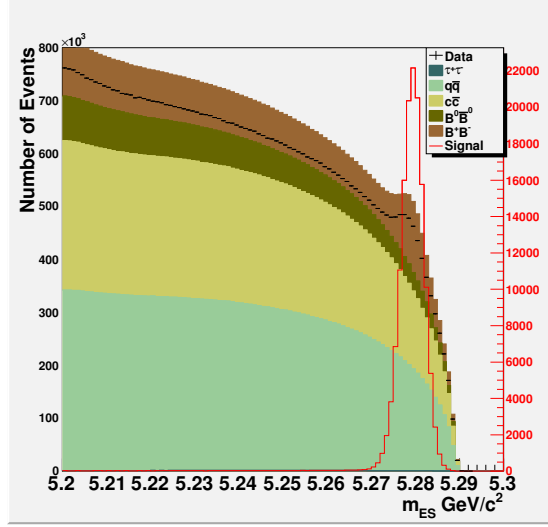


Figure 5.1: m_{ES} distribution after sub-skim level cuts. Signal MC (cocktail) distribution is shown in red. The data distribution (black points) is overlaid on the background MC distributions (color-filled). $q\bar{q}$ refers to $q = u, d, s$.

5.1.2 Purity cut

Purity is defined as the fraction of properly reconstructed B_{tag} 's within a given B_{tag} decay mode. As discussed in 4.3, B_{tag} candidates are reconstructed using 1077 various modes in the BSemiExclAdd skim. After reconstruction, some modes do not result in a clean B_{tag} sample. This is why the concept of purity is introduced. It is a measure of how well a B_{tag} candidate can be reconstructed using a certain hadronic mode.

The purity discussed here is referred to as the “low-multiplicity” purity and is defined in an environment that is more relevant to the signal decay mode. It is calculated using the B^+B^- and $B^0\bar{B}^0$ Monte Carlo samples after requiring 3 tracks or less and no more than 13 clusters per event on the B_{sig} side. The MC samples are also required to have the correct B_{tag} charge, neutral for $B^0\bar{B}^0$ and charged for B^+B^- , non-zero missing energy, and $m_{\text{ES}} > 5.273$. The fraction of properly reconstructed B_{tag} 's is calculated using Truth MC information, as has been done in Ref. [77]. The purity of a specific decay mode is determined by truth-matching the particle type of the tracks that were used to reconstruct the B_{tag} to its actual daughters. To do so, the number of required daughter kaons, pions, and K_s^0 is determined for each

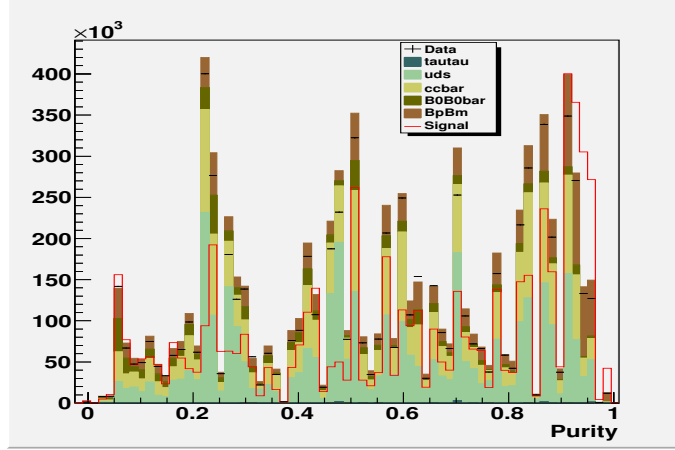


Figure 5.2: The purity value, determined for each event depending on the corresponding B_{tag} decay mode. Signal MC (*generic*) distribution is shown in red. The data distribution (black points) is overlaid on the background MC distributions (color-filled).

B_{tag} decay mode. The numbers are then compared to the information obtained from Truth MC information on the particle types of each B_{tag} daughter that were actually produced in the event. A match implies a properly reconstructed B_{tag} . This procedure is applied on an event-by-event basis.

Fig. 5.2 shows the purity distribution for *generic* signal, background MC and data². The peaking structure implies that certain B_{tag} decay modes are more likely to yield B_{tag} candidates that pass the “BestB” requirements. As can be seen, the signal peaks at higher values of the purity, whereas the background populates the whole region, with more events at lower values. For the purpose of this analysis, a purity cut of > 0.4 is applied, removing more than 50% of the combinatorial background while retaining 70% of signal and peaking B^+B^- MC. Cutting on the purity thus implies removing B_{tag} decay modes, which have a high proportion of combinatorial background due to misreconstruction. By using the low multiplicity purity, the B_{tag} modes that introduce most of the background events into the signal topology of this analysis are rejected.

²The cocktail sample cannot be used here because only a single B_{tag} mode is simulated: $B^+ \rightarrow D^0 \pi^+$, $D^0 \rightarrow K^+ \pi^-$. Hence, all cocktail events have the same value of purity

5.1.3 Continuum background suppression

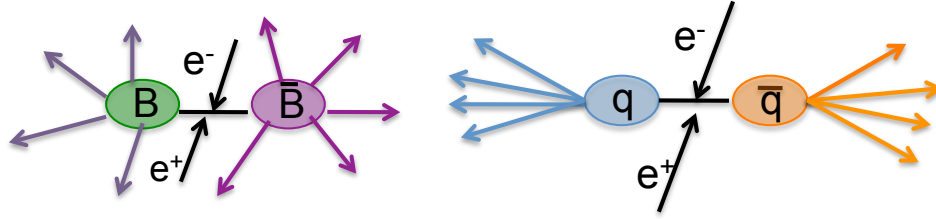


Figure 5.3: Schematic diagram of a $B\bar{B}$ (left) and continuum (right) event.

Continuum events have a different event topology as compared to $B\bar{B}$ events, as shown in Fig. 5.3. When $\Upsilon(4S)$ decays into a $B\bar{B}$ pair, the particles produced have a combined mass very close to that of the $\Upsilon(4S)$ and thus very small momentum. The B mesons decay isotropically, while almost at rest. On the other hand, the daughter quarks or leptons in a continuum event have a much smaller mass and therefore higher momentum. The decay products in $e^+e^- \rightarrow q\bar{q}$ or $\ell^+\ell^-$ events will have a more preferred direction and the resulting event topology is more co-linear. A mis-reconstructed B_{tag} candidate in a continuum event can be distinguished from that in a $B\bar{B}$ event using this difference in event topology. To do so, a multivariate likelihood is used, composed of six event shape variables: R2All, magnitude of the thrust axis and its z -component (Thrust_z), $\cos\theta_Y$, $\cos\theta_B$, $\cos\theta_{p_{\text{miss}}}$.

The event shape variables are shown in Fig. 5.4 and are discussed in greater detail below:

- R2All: This variable is defined as the ratio of the second to the zeroth Fox-Wolfram [79] moment using all charged and neutral particles in the event. It quantifies the “jettiness” of an event ranging between zero and one, with zero being more isotropic and one being more jet-like or collimated. As can be readily seen in Fig. 5.4a, signal and $B\bar{B}$ Monte Carlo events peak at lower values of R2All than continuum events.
- Magnitude of the thrust axis: The thrust axis is defined as the axis which maximizes the sum of the longitudinal momenta of an event’s decay products. The value of the thrust for $B\bar{B}$ events is small, since the B daughters do not have a preferred direction.

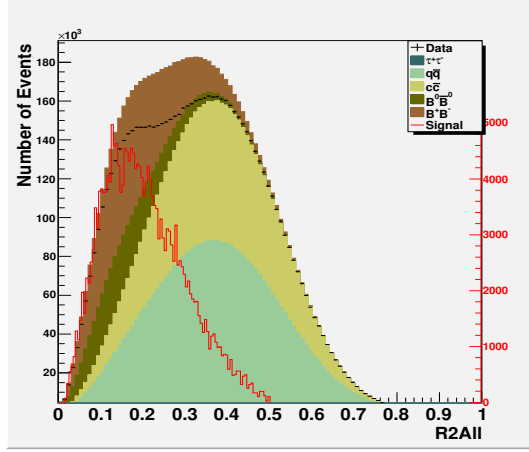
On the other hand, in continuum events, the decay products are boosted due to the high momentum of the daughter leptons or quarks. Continuum events thus have higher values of the thrust magnitude when compared to $B\bar{B}$ events, as shown in Fig. 5.4b.

- $|\cos \theta_T|$: $|\cos \theta_T|$ is defined as the cosine of the angle between the B_{tag} 's thrust axis and that of the B_{sig} in the CM frame. The decay products in an $e^+e^- \rightarrow q\bar{q}$ or $\ell^+\ell^-$ event are more likely to be back-to-back, and are usually mis-reconstructed into a B_{tag} and B_{sig} candidate. Fig. 5.4c shows that continuum events are strongly peaked at 1, whereas isotropic $B\bar{B}$ events have a smoother distribution.
- Thrust_z : It is the z -component of the thrust axis. The isotropic topology of a $B\bar{B}$ event implies that the decay products are more likely found near the central region of the detector. Continuum events, however, tend to produce high momentum jets that are more likely to travel at small angles with the beam axis. Thrust_z , will thus have higher values for $e^+e^- \rightarrow q\bar{q}$ events, as can be seen in Fig. 5.4d.
- $\cos \theta_{p_{\text{miss}}}$: This is the angle of the missing energy calculated as $\cos \theta_{p_{\text{miss}}} = p_{z_{\text{miss}}} / \vec{p}_{\text{miss}}$. The missing energy of signal events originate from neutrinos, which are produced isotropically. However, missing energy in continuum events is more likely to be the result of high-momentum jets, which travel at very small angles with the beam pipe and are often missed. This implies that the angle of the missing energy also displays a more dominant peak at ± 1 for continuum events, as shown in Fig. 5.4e.
- $\cos \theta_B$: $\cos \theta_B$ is defined as the cosine of the angle between the CM momentum of the B_{tag} and the beam axis. The distribution of B mesons from the decay of a spin-1 $\Upsilon(4S)$ is proportional to $\sin^2(\theta_B)$ and the resulting $\cos \theta_B$ distribution shows a higher peak at zero. Continuum events, on the other hand, display a flat distribution as shown in Fig. 5.4f.

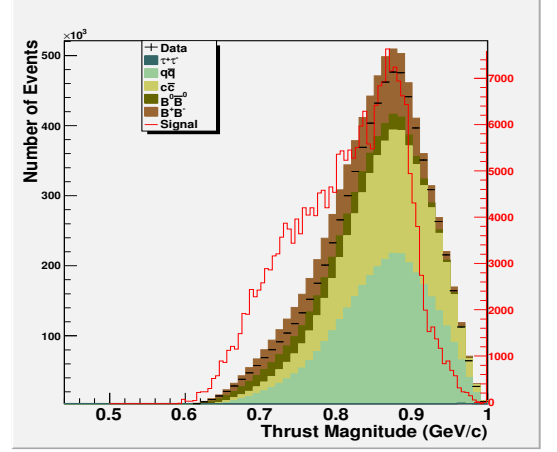
The six event-shape variables are then used in a multi-variate likelihood approach according to the following equation:

$$\mathcal{L} = \frac{\prod_i P_B(x_i)}{\prod_i P_B(x_i) + \prod_i P_q(x_i)} \quad (5.1)$$

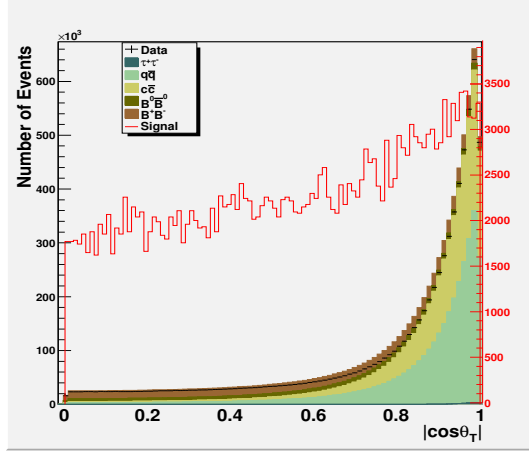
where $P(x_i)$ are probability density functions that describe six event shape variables for $B\bar{B}$ ($P_B(x_i)$) and continuum ($P_q(x_i)$) events. $P_B(x_i)$ is obtained using $B^+ \rightarrow K^+ \tau^+ \tau^-$ generic signal MC, while $P_q(x_i)$ is determined using $c\bar{c}$, uds and $\tau^+ \tau^-$ background MC. The event shape variables used in this multivariate approach are relatively uncorrelated. Before calculating both $P_B(x_i)$ and $P_q(x_i)$, a few cuts are applied to ensure proper $B\bar{B}$ reconstruction. An event is required to have one charged B_{tag} with $m_{\text{ES}} > 5.27 \text{ GeV}/c^2$. Also, an event must have exactly 3 signal side tracks with a sum charge opposite to that of the B_{tag} . The output of the multivariate likelihood is shown in Fig. 5.5, with $B\bar{B}$ and signal events strongly peaked at 1 and continuum events peaking at zero. A cut at 0.5 is applied to get rid of more than 75% of the continuum events with mis-reconstructed B mesons while retaining 80% of signal MC and $B\bar{B}$ events.



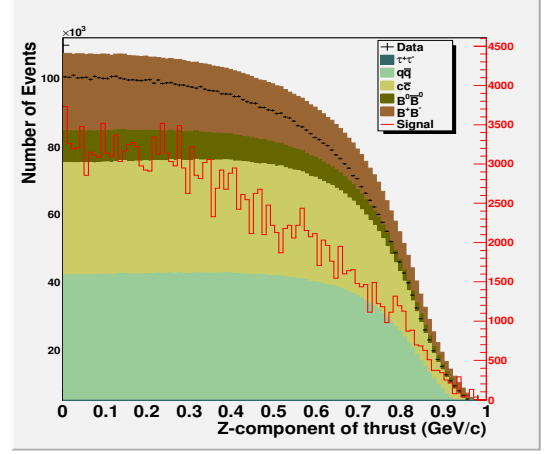
(a) R2All distribution.



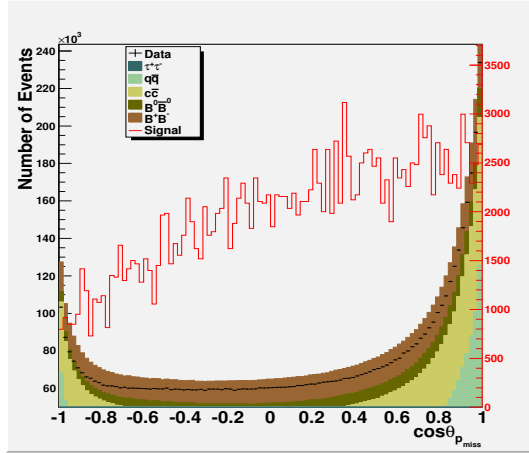
(b) Thrust magnitude distribution.



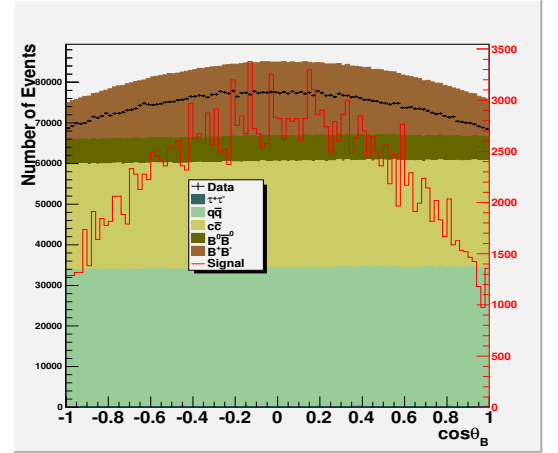
(c) $\cos \theta_T$ distribution.



(d) $Thrust_z$ distribution.



(e) $\cos \theta_{p_{miss}}$ distribution.



(f) $\cos \theta_B$ distribution.

Figure 5.4: Event shape variables used to calculate the continuum likelihood ratio. Signal MC (*generic*) distribution is shown in red. The data distribution (black points) is overlaid on the background MC distributions (color-filled). $q\bar{q}$ refers to $q = u, d, s$.

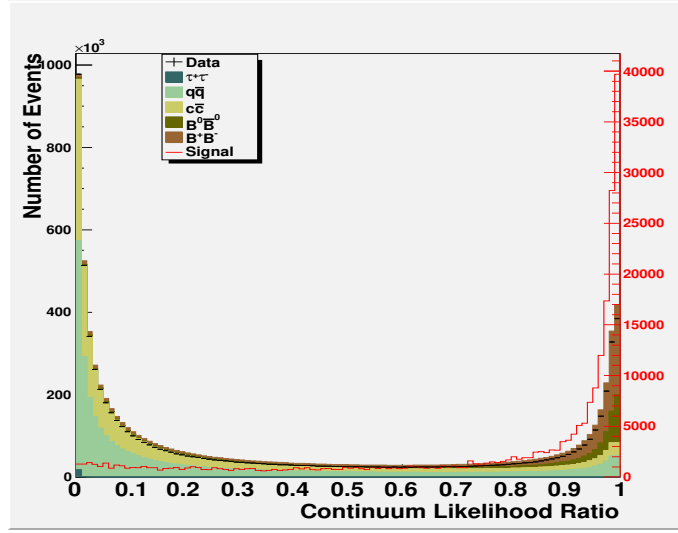


Figure 5.5: Output of the continuum multivariate likelihood for generic $B^+ \rightarrow K^+ \tau^+ \tau^-$ signal, background MC, and data. Signal MC (*generic*) distribution is shown in red. The data distribution (black points) is overlaid on the background MC distributions (color-filled). $q\bar{q}$ refers to $q = u, d, s$.

5.2 B_{sig} cuts

After requiring a properly reconstructed B_{tag} and applying continuum likelihood suppression, the following cuts are applied to select for $B^+ \rightarrow K^+ \tau^+ \tau^-$ events. As previously mentioned, only leptonic tau decays will be considered. The resulting three modes are:

- Electron Mode: $B^+ \rightarrow K^+ \tau^+ \tau^-$, $\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$, $\tau^- \rightarrow e^- \bar{\nu}_e \nu_\tau$
- Muon Mode: $B^+ \rightarrow K^+ \tau^+ \tau^-$, $\tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau$, $\tau^- \rightarrow \mu^- \bar{\nu}_\mu \nu_\tau$
- Electron-Muon Mode: $B^+ \rightarrow K^+ \tau^+ \tau^-$, $\tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau$, $\tau^- \rightarrow \mu^- \bar{\nu}_\mu \nu_\tau$

Before discussing the details of the signal selection, a brief discussion of the terminology used is in order. As previously noted, clusters refer to energy deposits in the EMC. Only clusters with energy greater than 20 MeV in the lab frame are considered in the data and MC samples. Clusters with smaller energies are disregarded. Furthermore, particle trajectories are denoted as tracks and thus, as previously mentioned, every event will have anywhere between zero and nine tracks. Information on the momentum and energy of the tracks is used for calculating variables like the E_{miss} or s_B , as will be discussed in section 5.2.1 and 5.7. Furthermore, particle identification algorithms [76] are applied on each track to determine whether or not it is a kaon, pion, electron, muon or proton.

5.2.1 E_{miss}

The signal modes discussed above have four neutrinos per decay, and thus typically have a large amount of missing energy. As previously mentioned, one of the main reasons for using the hadronic B_{tag} reconstruction is that there is no ambiguity about the source of missing energy in the event, as by definition the B_{tag} modes do not have undetected particles. The missing energy is attributed exclusively to B_{sig} . The missing energy four-vector is defined as the four-momentum of the B_{sig} , calculated by subtracting the four-momentum of the B_{tag} from that of the $\Upsilon(4S)$ in the CM frame, minus the four-momentum of all signal-side

tracks and clusters. Because of the neutrinos in this analysis, all signal candidate events are required to have $E_{miss} > 0$, where here the requirement is placed on the energy component of the four-vector.

5.2.2 Track multiplicity and PID

Exactly 3 tracks are required to account for the kaon and τ daughter charged leptons. The tracks must satisfy particle identification requirements of a kaon, electron or muon. Within the *BABAR* framework, particle identification of tracks is achieved using a set of software tools called PID selectors. These have been developed and optimized such that information from different parts of the detector is combined and used in various multivariate or cut-based techniques to establish a criteria for distinguishing a specific particle type. Furthermore, each PID selector has different selection levels, referred to as tightness levels, which vary according to the requirements that are applied on the variables used in the PID selector. The PID selectors used for this analysis are outlined in Table 5.1 and are applied in the order shown. Thus, each track in the event is passed through the kaon PID selector first. If it passes, then it is determined to be a kaon. Otherwise, it is passed through the electron selector. Tracks that fail the electron PID selector are passed through the muon one. Failing all 3 selectors, a track is passed through the pion PID selector. Furthermore, each track is assigned the nominal mass of its respective particle type: 493.7 MeV/ c^2 for a charged kaon, 0.51 MeV/ c^2 for an electron, 105.7 MeV/ c^2 for a muon, and 139.6 MeV/ c^2 for a charged pion.

Particle	PID selector	Tightness level
Kaon	PidKaonBDTSlector	Tight
Electron	PidElectronKMSelector	Tight
Muon	PidMuonBDTSelector	Loose
Pion	PidKMPionSelector	Tight
	!PIDElectronKMSelector	Loose
	!PIDBDTFakeRateMuon	Very Loose

Table 5.1: List of PID selectors used to identify the signal-side tracks.

The PID selector used for kaon identification is based on a Decision Tree [80] multivariate technique. A decision tree is a multivariate analysis tool that employs a set of discriminating variables to distinguish each particle type. It is composed of many nodes, the first of which divides a single sample of events into 2 classes (is or is not a kaon). After the first node, successive layers use different information from the detector subsystems to further determine whether or not the track in question is a kaon. The BDT Kaon selector refers to a Bagged [81] Decision Tree selector which specializes in kaon-pion separation. It identifies kaons using dE/dx information from the SVT and DCH below $p < 700 \text{ MeV}/c$ and Cherenkov angle measurements from the DIRC at higher momenta [76]. With the selected tightness level, K^\pm mesons are selected with an efficiency greater than 85% and with approximately 1% misidentification probability for pions and muons [77].

The Kalanand Mishra (KM) electron selector uses a technique called Error Correcting Output Code [82] (ECOC) to distinguish between kaon, pions, protons and electrons. It combines multiple binary classifiers, each trained differently, into a multiclass classifier to identify electrons with $\sim 90\%$ efficiency. Electrons are identified primarily using information from the EMC. The ratio of the measured energy of a calorimeter shower to the measured momentum of the corresponding charged track, E/p , provides good separation between electrons and other charged particles. Furthermore, the spatial distribution of the deposited energy is different between electrons, muon and hadrons, and is also used to distinguish electrons [76].

Muons are distinguished from pions using information from the IFR such as the penetration depth in the iron or the transverse size of the deposited cluster [76]. The PID selector used for muon identification is also a Decision Tree classifier. It uses 30 input variables to provide good separation between muons and pions. BDT muon selectors are tuned to give a constant muon ID efficiency as a function of the muon momentum, and the selected tightness level chosen for this analysis corresponds to a muon efficiency of 80%.

Finally, a track that fails the BDTTightKaon, KMTightElec, and BDTLMuon selector is

tagged as a pion if it passes the pion KMTight selector. A track identified as a pion should also fail the loose requirement by the KM electron selector and the very loose requirement by the BDTFakeRateMuon selector. The FakeRate muon selectors are tuned to give a constant pion mis-ID rate. At the loose level, the BDTFakeRate selector has a pion misID rate of 5% [77]³.

As previously mentioned, a signal event should have one kaon and either two electrons, two muons, or one electron and one muon. Furthermore, to ensure charge conservation, the charge of the kaon must be opposite to that of the B_{tag} , while the two leptons are required to be oppositely charged.

5.2.3 π^0 veto

A signal event should ideally include zero additional clusters, but only $\sim 20\%$ do not have any extra energy deposits. Background events, with extra clusters, can pass the above track multiplicity requirements and PID requirements. Specifically, B meson decays to charmed mesons, such as a D^0 or D^+ , can have a similar final state as a signal event with an additional π^0 candidate. Such background decay modes have a relatively high SM branching fraction [18], where here the K and π^0 are usually daughters of the $D^{(*)}$. To suppress such backgrounds, a π^0 veto is applied. Here, a veto implies rejecting a class of events that satisfy a specific criteria. This is in contrast to the standard signal selection, where events that fail a certain criteria are rejected.

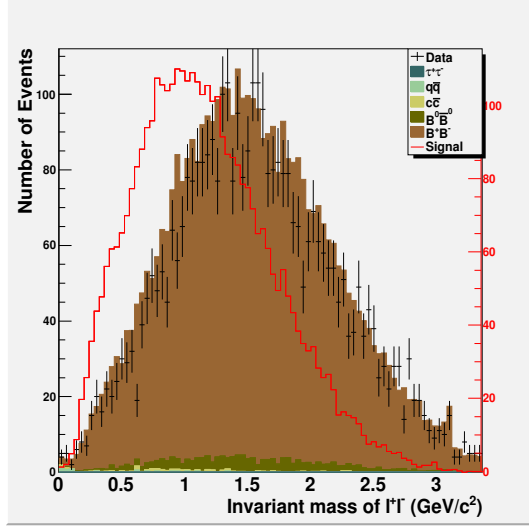
To apply a π^0 veto, π^0 candidates are first reconstructed in each event via the decay $\pi^0 \rightarrow \gamma\gamma$ by combining pairs of clusters in the event. To pass the reconstruction, these clusters must have a lab energy greater than 50 MeV and a lateral moment [76] ranging between 0 and 0.8. The lateral moment is calculated from the energy distribution of the electromagnetic shower and many “junk” clusters have a lateral moment of exactly zero [77]. Furthermore, the sum energy of the $\gamma\gamma$ combination should be greater than 100 MeV and

³Pion PID will be necessary when investigating peaking backgrounds and evaluating a validation test for the signal selection.

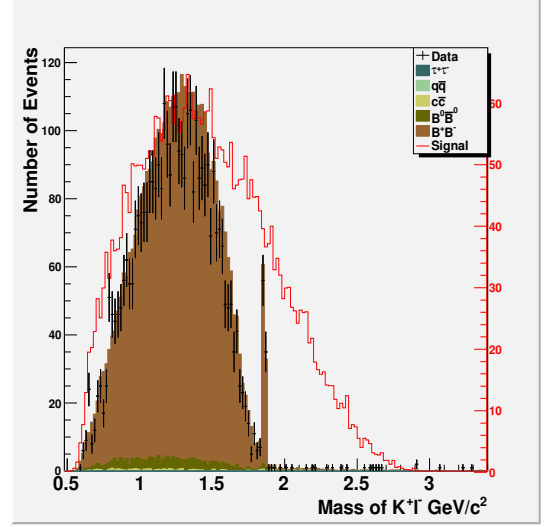
the invariant mass should lie within 0.1 and 0.16 GeV/c^2 . Any event with a pair of clusters that satisfy these requirements is considered to have a π^0 candidate and is thus rejected.

5.2.4 γ , J/ψ , and D^0 vetos

At this point, it is necessary to investigate potential sources of background. One way to do so is to look at different mass combinations of the tracks in the event. Fig. 5.6a shows the invariant mass of the lepton pair for both background and cocktail signal Monte Carlo. One background peak can be easily noted, at about $3.10 \text{ GeV}/c^2$. This is the J/ψ resonance which decays via $J/\psi \rightarrow e^+e^-$ or $J/\psi \rightarrow \mu^+\mu^-$. Even though the $\psi(2S)$ also decays into an e^+e^- or $\mu^+\mu^-$ pair, a peak is not clearly defined in the plot. Fig. 5.6a also shows events with an invariant mass of zero, which correspond to photon conversions: $\gamma \rightarrow e^+e^-$. Furthermore, another interesting mass combination is the sum of the kaon with the *oppositely* charged lepton. The invariant mass distribution is shown in Fig. 5.6b, where a D^0 peak is clearly visible. This peak is a result of pions from $D^0 \rightarrow K^+ \pi^-$ being misidentified as muons and thus forming a D^0 candidate when joined with the oppositely charged kaon. To get rid of all such events, γ , J/ψ and D^0 vetoes are applied. To veto photon conversions, the invariant mass of the combination of the electron with any other oppositely charged track in the event must be greater than $50 \text{ MeV}/c^2$. Furthermore, the invariant mass of the lepton pair in a signal event should lie outside the J/ψ mass region. If $3.00 < m_{\ell^+\ell^-} < 3.194 \text{ GeV}/c^2$, then the event is vetoed. Finally, the invariant mass of the $K^\pm \ell^\mp$ sum should be less than $1.80 \text{ GeV}/c^2$ or greater than $1.90 \text{ GeV}/c^2$. Otherwise, the event is considered to have a D^0 candidate and is thus discarded. After applying these vetoes, more than 90% of signal and background events are retained, excluding uds where only 70% of events pass these cuts. Thus, the γ , D^0 , and J/ψ vetoes do not have a large impact on either the signal efficiency nor the background levels. Nevertheless, because of the evident peaks in the plots shown, it is important that they are applied.



(a) Invariant mass of the lepton pair.



(b) Invariant mass of the kaon and oppositely charged lepton.

Figure 5.6: Invariant mass distributions of signal-side track combinations. Signal MC (*cocktail*) distribution is shown in red. The data distribution (black points) is overlaid on the background MC distributions (color-filled). $q\bar{q}$ refers to $q = u, d, s$.

5.2.5 s_B cut

The normalized q^2 distribution, known as s_B , is given by the following equation:

$$s_B = \frac{q^2}{m_B^2} = \frac{(p_{B_{\text{sig}}} - p_K)^2}{m_B^2}, \quad (5.2)$$

where $p_{B_{\text{sig}}}$ is the four-momentum vector of the signal B , p_K is the four-momentum vector of the kaon, and m_B is the mass of B_{sig} . As shown in Fig. 5.7, the s_B signal distribution only populates values of s_B higher than 0.45. This is due to the mass of the τ leptons, which limits the phase space available to the kaon. On the other hand, background events populate the entire s_B region. Thus, signal events are required to lie within the kinematically permitted region: $s_B > 0.45$.

The cut-off at $s_B = 0.45$ is a kinematic boundary. Therefore, this cut is not affected by the choice of the theoretical model used in generating the signal MC or the signal reweighting discussed in section 4.6.

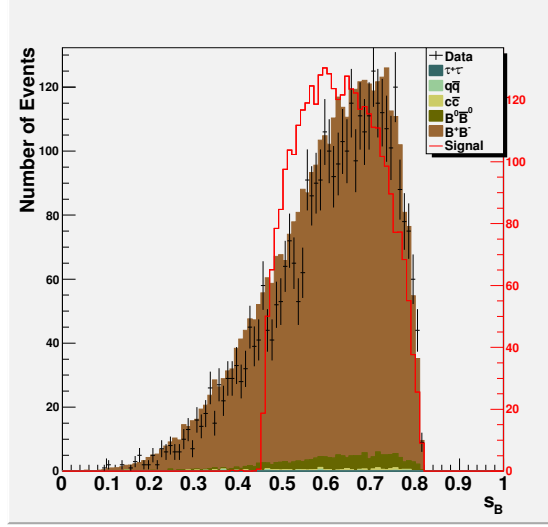


Figure 5.7: s_B distribution after all above cuts for data, background and signal Monte Carlo. Signal MC (*cocktail*) distribution is shown in red. The data distribution (black points) is overlaid on the background MC distributions (color-filled). $q\bar{q}$ refers to $q = u, d, s$.

5.3 Peaking background

As shown in Fig. 5.7, the MC predicts that the background at this stage of the selection is dominated by B^+B^- events. Other background sources, such as $B^0\bar{B}^0$ and $c\bar{c}$, are highly suppressed. Using Truth MC information, the type of B^+B^- events are determined to be mainly $B \rightarrow D^{(*)}\ell\bar{\nu}_\ell$, where $\ell = e^\pm$ or μ^\pm , and with the D^0 decaying into one of several possible modes, in particular semileptonic final states such as $D^0 \rightarrow K \ell\bar{\nu}_\ell$. Such events have the same final state as the signal: a charged kaon, two oppositely charged leptons, and a missing energy component from the lepton neutrinos. If the background decay mode includes a D^* , then the D^* usually decays first into a D^0 or D^\pm and a π^0 or γ . Furthermore, a small portion of the background events include real τ 's, $B \rightarrow D^{(*)}\tau\nu_\tau$, where the τ further decays via lepton or hadron modes. These events form a large background component that peaks in the m_{ES} distribution related to this stage of the signal selection, and thus need to be further suppressed to properly isolate signal events.

5.4 Multilayer perceptron neural network

As mentioned in section 5.3, there is a large peaking background component which has a very similar final state as the signal events. To suppress this background, a multilayer perceptron (MLP) neural network, using 8 discriminating variables, has been trained to separate between signal and peaking background. A MLP neural network is a multivariate analysis technique inspired by biological neural networks [83]. It is composed of interconnected layers of nodes (or artificial neurons) which receive and transmit signals from one another. In general, a neural network consists of one input layer, an output layer, and a set of hidden layers. Each node collects all the input information, determines a net signal or decision and transmits the resulting output to the next layer. The connections between the neurons are weighted with different values, and the network propagates the signal through the layers to produce the final output [83]. The final output allows for the classification of a certain event as either signal or background.

For this analysis, the MLP neural network is selected because it provided optimal and stable results, using the given variables. Other TMVA techniques were tested, and the MLP neural network was found to give 10-30% better separation between signal and background, as well as 2-13% improved signal efficiency at a 30% background rejection rate. The neural network is trained using cocktail signal events and B^+B^- background events. The samples are randomly split in half for training and validation. The discriminating variables are chosen by narrowing down a larger list of possible input variables and identifying a set of 8 variables that are the most effective in separating signal and background. The cocktail signal is used because of its high statistics. At this point, the generic signal sample has a low number of surviving events and the signal-to-background differences in a variable distribution cannot be clearly defined. Nevertheless, Appendix A shows the distribution of generic signal as compared to cocktail signal for each of the discriminating variables discussed below.

The discriminating variables are shown in Fig. 5.11, Fig. 5.12 and Fig. 5.13. They are split into 3 categories: angular, kinematic and calorimeter discriminating variables. The angu-

lar variables refer to angles between signal-side tracks, calculated using the four-momentum vectors of the B_{sig} and its daughters. Kinematic variables are related to tracking information, while calorimeter variables are determined using cluster information in the ntuples. Furthermore, a new frame of reference is defined for some of these discriminating variables: the di-tau reference frame. It is the rest frame of the combination of the two tau's in a $B^+ \rightarrow K^+ \tau^+ \tau^-$ decay, and is calculated by subtracting the four-vector of the K from that of B_{sig} . The discriminating variables are described below.

5.4.1 Angular discriminating variables

- $\cos \theta_{l+l-}$: This is defined as the angle between the two leptons in the di-tau frame. Fig. 5.8 shows a schematic diagram of the angle in a $B^+ \rightarrow K^+ \tau^+ \tau^-$ signal event compared to a $B \rightarrow D^0 \ell \bar{\nu}_\ell$ event. In a $B \rightarrow D^0 \ell \bar{\nu}_\ell$ event, the two leptons are more likely back-to-back. The distribution in Fig. 5.11a shows a clear peak at -1 for background events, with a much lower concentration at higher values of $\cos \theta_{l+l-}$. Signal events also peak at -1 , but, in addition, have a more uniform distribution covering the full region between $+1$ and -1 . This is because signal events do not have such a strong preferred direction for the angle between the two leptons.

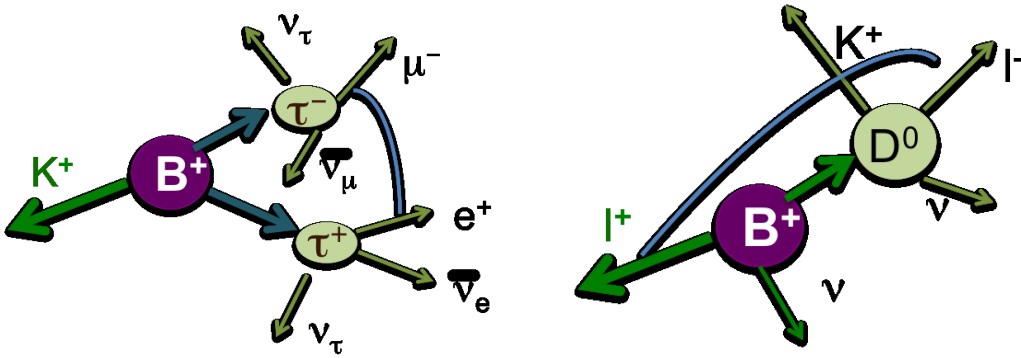


Figure 5.8: Schematic diagram of the angle between the two leptons in the di-tau frame, $\cos \theta_{l+l-}$, for a signal (left) and background (right) event.

- $\cos \theta_{K+l-}$: This is defined as the angle between the kaon and the oppositely charged

lepton in the di-tau frame. A schematic diagram, shown in Fig. 5.9, depicts the difference between a $B^+ \rightarrow D^0 \ell \bar{\nu}_\ell$ event and a signal event. In the former, the K is a daughter of the B while the lepton is a τ daughter, which implies that there is no strong correlation between the direction of the two particles. However, for a background event, the K and the oppositely charged lepton originate from the same particle, the D meson. Because this is a 3-body decay, the K and the oppositely charged lepton are not always back-to-back. Therefore, the angle between the two tends to be less than 90° or $\cos \theta_{K+l^-} > 0$, as can be seen in the background distribution of Fig. 5.11b. The signal distribution, on the other hand, is relatively uniform implying a less significant correlation between the direction of the K and the oppositely charged lepton in a $B^+ \rightarrow K^+ \tau^+ \tau^-$ event.

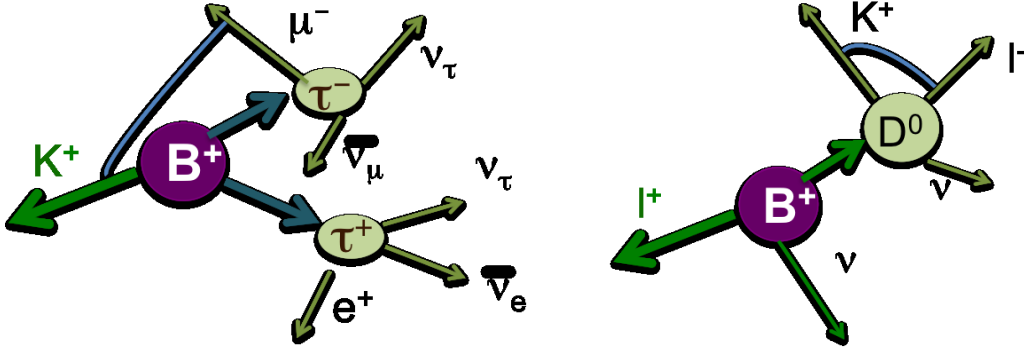


Figure 5.9: Schematic diagram of the angle between the kaon and the oppositely charged lepton in the di-tau frame, $\cos \theta_{K+l^-}$, for a signal (left) and background (right) event.

- $\cos \theta_{B+l^-}$: This is defined as the angle between B_{sig} and the oppositely charged lepton in the CM frame. As shown in Fig. 5.11c, the angle between B_{sig} and the oppositely charged lepton tends to dominate high values of the spectrum, with an evident peak at around 0.9 and a smaller one at -0.9 . This implies that the two particles are either aligned together, shown in Fig. 5.10, or almost back-to-back in the CM frame of a background event. This behaviour is not so evident for a signal event, where the role of the same charged or oppositely charged lepton can be easily interchanged and thus

the correlation is much less prominent.

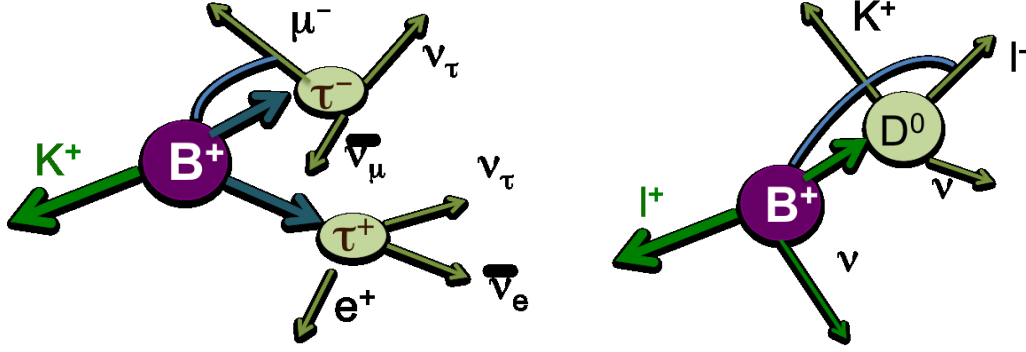
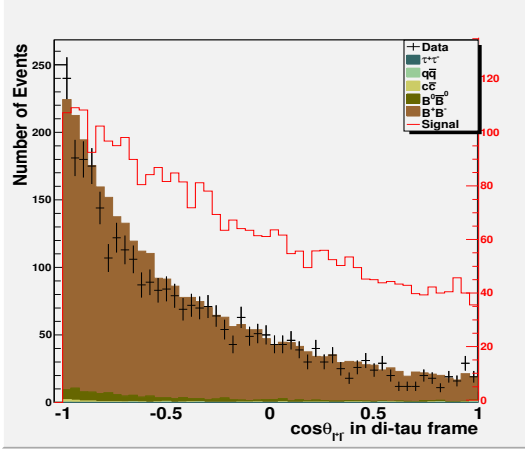
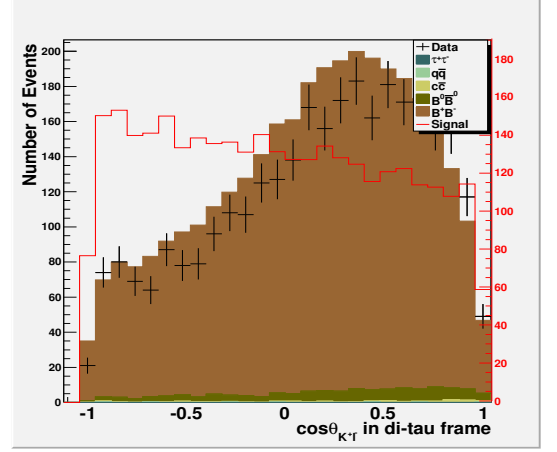


Figure 5.10: Schematic diagram of the angle between B_{sig} and the oppositely charged lepton in the CM frame, $\cos \theta_{B+l^-}$, for a signal (left) and background (right) event.

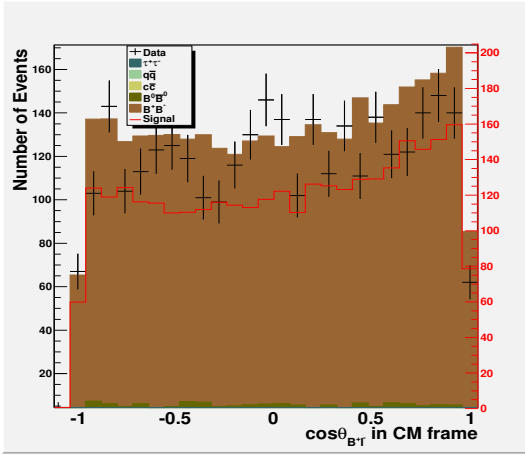
- $\cos \theta_{(B_{\text{sig}}-K)l_{\text{low}}}$: This is defined as the angle between the vector recoiling against the K and the low momentum lepton in the CM frame. Fig. 5.11d shows the distribution of $\cos \theta_{(B_{\text{sig}}-K)l_{\text{low}}}$ for signal and background events. $B^+ \rightarrow D^{(*)} \ell \bar{\nu}_\ell$, $D^{(*)} \rightarrow K \ell \bar{\nu}_\ell$ events show a peaking behaviour around zero. The low-momentum lepton is the daughter of the D meson in the event. On the other hand, in a signal event, it is not clear which tau daughter will have lower momentum. It is thus hard to distinguish any angular correlation for $B^+ \rightarrow K^+ \tau^+ \tau^-$ events, which is why the signal distribution is relatively uniform over the entire region.



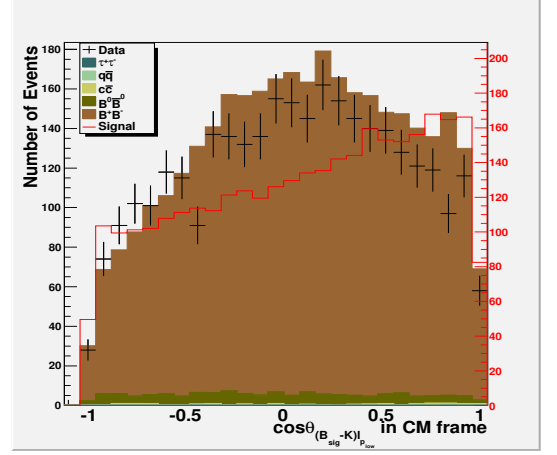
(a) $\cos \theta_{l+l-}$ in di-tau frame.



(b) $\cos \theta_{K+l-}$



(c) $\cos \theta_{B+l-}$ in center-of-mass frame.



(d) $\cos \theta_{(B_{\text{sig}}-K)l_{p_{\text{low}}}}$ in CM frame.

Figure 5.11: Angular variables used in the MLP neural network: angle between two leptons in di-tau frame (top-left), angle between K and oppositely charged lepton in di-tau frame (top-right), angle between B_{sig} and oppositely charged lepton in CM frame (bottom-left), and angle between the K recoil vector and the lepton with low momentum in CM frame (bottom-right). Signal MC (*cocktail*) distribution is shown in red. The data distribution (black points) is overlaid on the background MC distributions (color-filled). $q\bar{q}$ refers to $q = u, d, s$.

5.4.2 Kinematic discriminating variables

The kinematic discriminating variable used in the neural network is the momentum of the lepton with charge opposite to that of a kaon, in the di-tau frame. In a background event, this lepton is more likely the daughter of the D and is thus recoiling against the K . In a

signal event, it is not easy to distinguish between the two lepton daughters of the $\tau^+\tau^-$ pair. As can be seen in Fig. 5.12, more background events have lower lepton momentum than signal events.

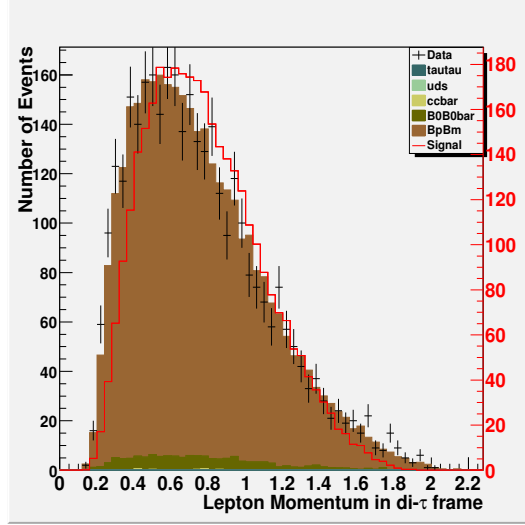


Figure 5.12: Momentum of lepton, with charge opposite to that of the kaon, in the di-tau frame. Signal MC (*cocktail*) distribution is shown in red. The data distribution (black points) is overlaid on the background MC distributions (color-filled). $q\bar{q}$ refers to $q = u, d, s$.

5.4.3 Calorimeter discriminating variables

- Missing energy in di-tau frame: The definition of missing energy here differs from the previous requirement of a non-zero E_{miss} . The latter is calculated by taking the energy component of the four-vector obtained by subtracting the four-vector of all signal-side tracks and clusters from the B_{sig} four-vector. The missing energy variable used in the MLP neural network is the energy component of the four-vector calculated using the following equation:

$$p_{miss} = p_{B_{sig}} - p_K - p_{\ell^+\ell^-} \quad (5.3)$$

where $p_{B_{sig}}$, p_K , and $p_{\ell^+\ell^-}$ are the four-vectors of the B_{sig} , kaon, and lepton sum, respectively, all in the di-tau frame. The two definitions of missing energy should yield the same value for signal events with zero additional clusters, except for the difference

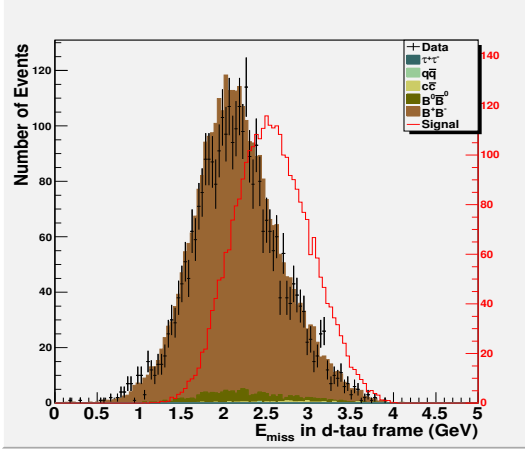
in reference frames. However, this is not necessarily true for background events. As can be evidently seen in Fig. 5.13a, the missing energy is higher for signal events than background events. Signal events have 4 lepton neutrinos contributing to the total missing energy, whereas $B \rightarrow D^{(*)}\ell\bar{\nu}_\ell$ events have a smaller multiplicity of neutrinos.

- E_{extra} : This is defined as the sum of the energies of all clusters with individual energy greater than 50 MeV in the CM frame. The 50 MeV cut-off is selected such that clusters resulting from accelerator beam backgrounds or detector noise, which are expected to dominate at low energies (below ~ 50 MeV), are excluded [77]. These low energy clusters are also not well modeled in the simulation.

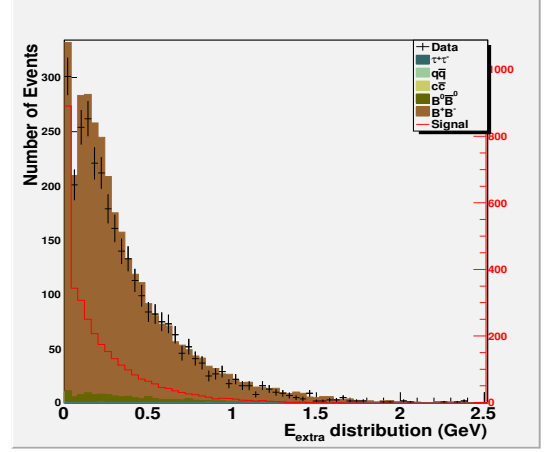
Ideally, signal events should have an E_{extra} of zero. However, as can be seen in Fig. 5.13b, in addition to a peak at zero, many signal events also have nonzero values of the E_{extra} variable. These additional clusters can result from cluster “fragments” or “split-offs”, which are typically the result of hadronic showers of tracks on either the B_{tag} or B_{sig} side. Such “fragments” can be mis-reconstructed and assigned as multiple clusters in the signal event. Furthermore, daughter clusters of the B_{tag} can be mis-reconstructed as B_{sig} daughters and also yield a non-zero E_{extra} . These mis-assigned clusters are generally low in energy and thus still allow the B_{tag} candidate to pass the “BestB” selection. Finally, even at energies greater than 50 MeV, noise from the beam or detector, as well as cosmic rays, still contribute to the non-zero E_{extra} distribution of signal events. On the other hand, the background distribution has a large concentration of events with E_{extra} greater than zero, as compared to signal. This extra energy in the peaking B^+B^- component is mainly due to γ or π^0 candidates, which are daughters of a D , D^* , or the B meson itself. An example of such a background event is $B^+ \rightarrow D^{*0} \ell^+ \nu_\ell$, $D^{*0} \rightarrow D^0 \pi^0$, $D^0 \rightarrow K \ell \nu_\ell$, where the π^0 candidate has an invariant mass or individual photon energy such that the event is not rejected by the π^0 veto.

- Invariant mass of K^+l^- : Fig. 5.13c shows the distribution of the invariant mass of

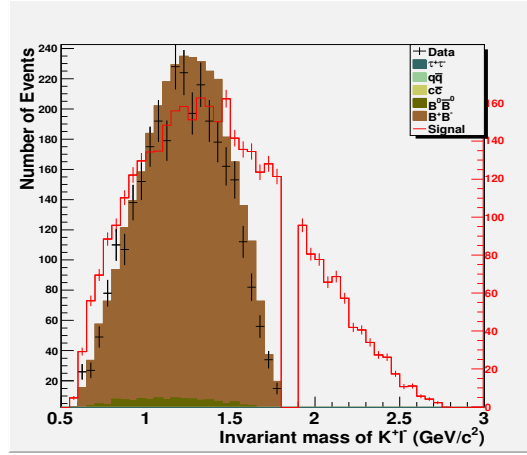
the kaon with the oppositely charged lepton, calculated in the lab frame. Background events have the distribution concentrated below the D meson mass. This is because the kaon and the oppositely charged lepton are decay daughters of the D in most peaking B^+B^- events. However, the same is not true for signal events where the distribution goes to values greater than $3.0 \text{ GeV}/c^2$. Furthermore, as can be readily seen, the region above the D meson mass has a very low number of background events. A direct cut on this variable has been considered for the signal selection and is discussed in Appendix B. The approach with the MLP neural network is found to result in an improved suppression of the background levels.



(a) Missing energy in the di-tau frame.



(b) E_{extra} distribution in the CM frame.



(c) Invariant mass of the kaon and oppositely charged lepton sum.

Figure 5.13: Calorimeter variables used in the MLP neural network. Signal MC (*cocktail*) distribution is shown in red. The data distribution (black points) is overlaid on the background MC distributions (color-filled). $q\bar{q}$ refers to $q = u, d, s$.

5.4.4 MLP neural network output

The MLP neural network is trained for each of the three signal modes (e^+e^- , $\mu^+\mu^-$, $e^+\mu^-$) separately. Cocktail signal MC is used as signal and B^+B^- MC as background. The signal and background events are randomly split in half for training and testing the neural network. Signal events are assigned a weight of 1, whereas B^+B^- background events are assigned their normalization weight, to ensure that the background types are appropriately represented.

The output for each mode is shown in Fig. 5.14. The discrimination between signal and background is evident. The final step in the signal selection is to apply a cut on the MLP neural network output.

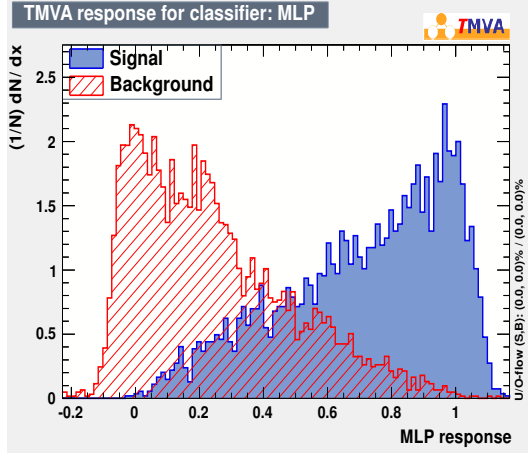
Since the sensitivity of this analysis is far from the SM expectation, the anticipated outcome is that no signal will be present. Consequently, the final cut is optimized so as to yield the most stringent branching fraction upper limit. This choice will simultaneously not impact the ability to observe a signal if present. To select the appropriate MLP cut, the upper limit of the branching fraction is calculated using trial values of the cut value in the MLP output signal region, which is defined in the range: MLP output > 0.5 . The cuts are applied between an output of 0.5 and 0.7 in intervals of 0.1, and between an output of 0.7 and 1.0 in intervals of 0.05. The upper limit is calculated for each given cut using the Barlow method [93], which is further discussed in section 8, assuming $N_{obs} \approx N_{bkg}$. The result is shown in Fig. 5.15 for each of the three modes. The final cut for each mode is chosen such that the resulting branching fractions corresponds to the minima of the second order polynomial used to fit the output in Fig. 5.15. The chosen cut values are summarized in Table 5.2.

The MLP cut is the final cut in the signal selection. At this point, the signal efficiency is calculated using the generic signal Monte Carlo, while the background estimate is obtained using the m_{ES} sideband substitution approach, discussed in section 7. Before applying the MLP cut, data is not explicitly blinded. The high levels of background render it impossible to identify signal events in data before the MLP cut. Furthermore, access to data at earlier stages in the analysis is necessary to study the level of data-MC agreement. However, the output of the MLP neural network in data is explicitly blinded to avoid any bias in the selection of the final cut in this analysis. Furthermore, to ensure that the MLP cut does not affect the sensitivity to new physics contributions, plots of the kaon momentum and s_B are shown in Appendix C after the full signal selection, with the data blinded. The background MC distributions of the kaon momentum and s_B are uniform, and thus the MLP cut does

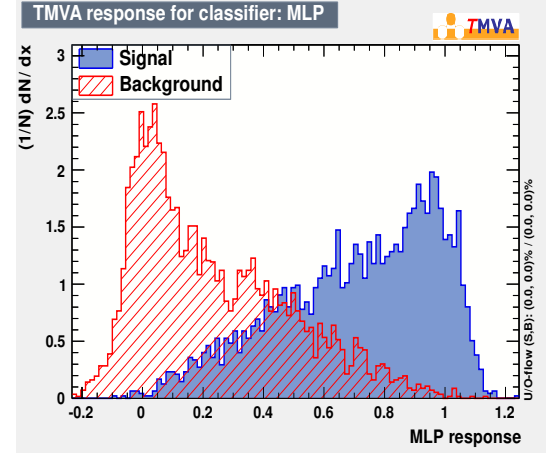
not appear to limit the phase space in an otherwise kinematically accessible region.

Mode	MLP Cut
Electron	> 0.7
Muon	> 0.7
Electron-Muon	> 0.75

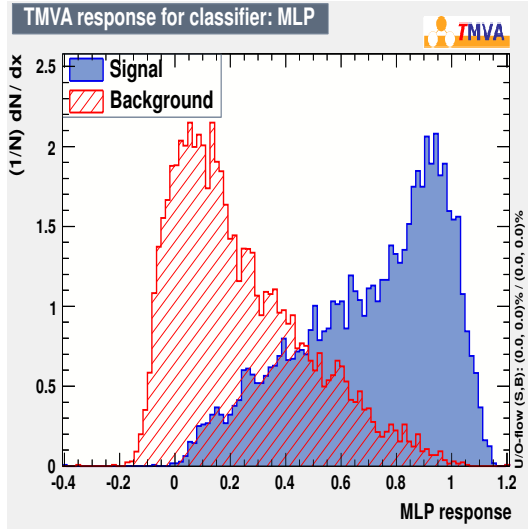
Table 5.2: Value of MLP cut applied for each mode.



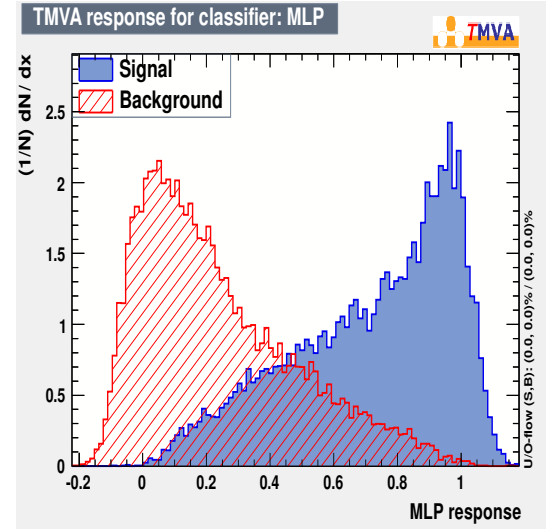
(a) Electron Mode.



(b) Muon mode.

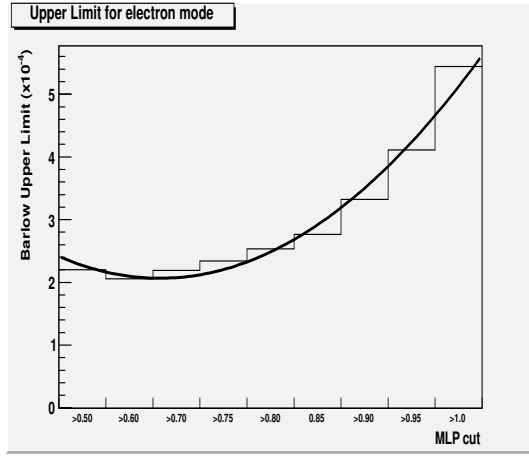


(c) Electron-muon mode.

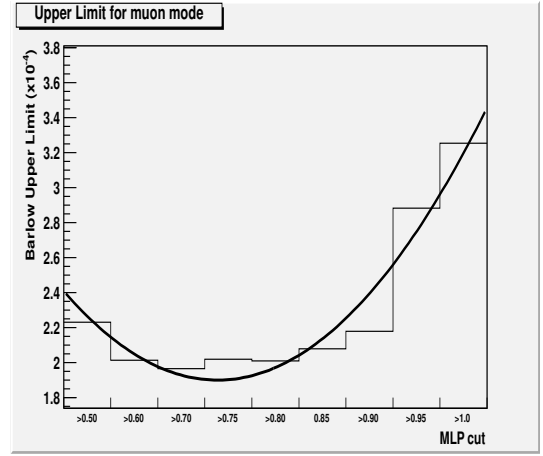


(d) All 3 modes combined.

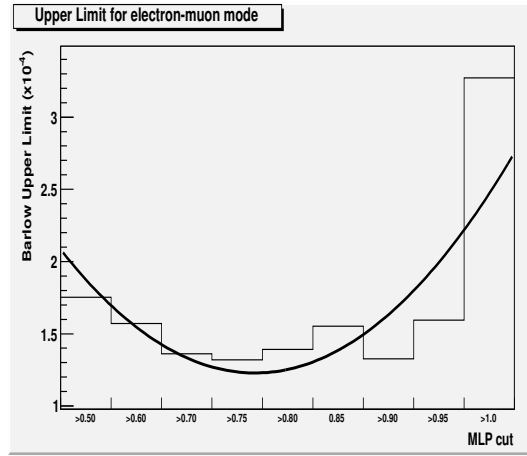
Figure 5.14: MLP neural network output for electron, muon, electron-muon and all three modes combined.



(a) Electron Mode.



(b) Muon mode.



(c) Electron-muon mode.

Figure 5.15: Barlow upper limit as a function of MLP cut, calculated assuming $N_{obs} \approx N_{bkg}$.

Chapter 6

Uncorrected signal and background MC yields

The event yield, efficiencies and partial efficiencies of each of the signal, cocktail and background Monte Carlo, along with the data, are given in the tables below. The signal efficiencies and background numbers shown here are the raw yields, before the m_{ES} sideband substitution (see section 7). The background MC event yields are weighted. The efficiencies are calculated as $N_{\text{passed}}/N_{\text{generated}}$, where N_{passed} is the number of events passing a specific cut and $N_{\text{generated}}$ is the number of events generated for a specific MC sample. The partial efficiencies are shown to compare the effect of different cuts on the signal efficiency and on the background levels. These are given by $\frac{N_{i+1}}{N_i}$, where N_i is the event yield after a certain cut, i , and N_{i+1} is the event yield after the subsequent cut, $i + 1$.

Furthermore, the plot of the final signal efficiency as a function of s_B is also shown in Fig. 6.1, for all 3 modes combined. The signal efficiency, within the limited statistics, is relatively uniform in the kinematically allowed range of s_B . This is important to ensure that, with this signal selection, a $B^+ \rightarrow K^+ \tau^+ \tau^-$ signal can be observed in the full s_B range. Furthermore, this also implies that the result of the signal selection is relevant regardless of the nature and kinematics of possible new physics contributions.

Cut	Generic	Cocktail
Generated Events	3.31×10^6	3.56×10^6
B_{tag} Cuts	1.71×10^4	1.11×10^6
E_{miss} Cut	1.69×10^4	1.1×10^6
Continuum Likelihood Cut >0.5	1.37×10^4	7.02×10^5
Purity Cut >0.4	9.92×10^3	7.02×10^5
Three Tracks	6.59×10^3	5.04×10^5
Kaon PID	4.09×10^3	3.58×10^5
Kaon Charge Cut	3.96×10^3	3.54×10^5
Two lepton Cut	371	3.69×10^4
π^0 Veto	306	2.97×10^4
D^0, γ , and J/ψ veto	284	2.80×10^4
S_B Cut	284	2.80×10^4
Leptonic e^+e^-	76	7.81×10^3
Leptonic $\mu^+\mu^-$	71	6.42×10^3
Leptonic $e^+\mu^-$	137	1.38×10^4
MLP cut	158	1.54×10^4

Table 6.1: Number of raw events in $B \rightarrow K^\pm \tau^+ \tau^-$ generic and cocktail signal Monte Carlo after each selection cut.

Cut	B^+B^-	$B^0\bar{B}^0$	$c\bar{c}$	uds	$\tau^+\tau^-$	data
Generated Events	2.3×10^9	2.4×10^9	5.5×10^9	3.4×10^9	1.6×10^9	–
B_{tag} Cuts	1.5×10^6	6.2×10^5	2.6×10^6	3.4×10^6	1.9×10^4	7.5×10^6
E_{miss} Cut	1.3×10^6	5.7×10^5	2.1×10^6	2.4×10^6	1.8×10^4	5.8×10^6
Continuum Likelihood Cut >0.5	1.1×10^6	4.7×10^5	4.8×10^5	5.9×10^5	32.8	2.3×10^6
Purity Cut >0.4	7.8×10^5	2.3×10^5	2.9×10^5	3.6×10^5	25.7	1.45×10^6
Three Tracks	3.1×10^5	6.1×10^4	1×10^5	1.3×10^5	4.7	5.0×10^5
Kaon PID	1.1×10^5	1.39×10^4	2.3×10^4	2.3×10^4	0.5	1.5×10^5
Kaon Charge Cut	1.0×10^5	1.1×10^4	1.8×10^4	1.73×10^4	0.3	1.27×10^5
Two lepton Cut	8.1×10^3	334	97	33.3	0	7.4×10^3
π^0 Veto	4.02×10^3	144	19.5	5.3	0	3.7×10^3
D^0, γ , and J/ψ veto	3.9×10^3	138	17.9	3.7	0	3.6×10^3
S_B Cut	3.27×10^3	123	14.6	2.4	0	3.1×10^3
Leptonic e^+e^-	883	31.6	3.5	0.8	0	838
Leptonic $\mu^+\mu^-$	766	28.4	3.8	1.0	0	697
Leptonic $e^+\mu^-$	1.6×10^3	63.4	7.24	0.5	0	1.6×10^3
MLP cut	1.6×10^2	8.9	2.3	1.1	0.0	–

Table 6.2: Number of normalized events in background Monte Carlo after each selection cut.

Cut	Generic $\times 10^{-1}$ (%)	Cocktail $\times 10^{-3}$ (%)
B_{tag} Cuts	516.66 ± 0.40	311.02 ± 0.34
E_{miss} Cut	511.82 ± 0.39	307.33 ± 0.34
Continuum Likelihood Cut > 0.5	414.24 ± 0.35	197.12 ± 0.26
Purity Cut > 0.4	299.94 ± 0.30	197.10 ± 0.26
Three Tracks	199.18 ± 0.25	141.43 ± 0.21
Kaon PID	123.64 ± 0.19	100.40 ± 0.18
Kaon Charge Cut	119.86 ± 0.19	99.39 ± 0.18
Two lepton Cut	11.22 ± 0.06	10.35 ± 0.05
π^0 Veto	9.25 ± 0.05	8.33 ± 0.05
D^0, γ , and J/ψ veto	8.59 ± 0.05	7.87 ± 0.05
S_B Cut	8.59 ± 0.05	7.86 ± 0.05
Leptonic e^+e^-	2.30 ± 0.03	2.19 ± 0.02
Leptonic $\mu^+\mu^-$	2.15 ± 0.03	1.80 ± 0.02
Leptonic $e^+\mu^-$	4.14 ± 0.04	3.87 ± 0.03
MLP cut	4.78 ± 0.38	4.32 ± 0.03

Table 6.3: Efficiencies in $B \rightarrow K^\pm \tau^+ \tau^-$ generic and cocktail signal Monte Carlo after each selection cut.

Cut	B^+B^-	$B^0\bar{B}^0$	$c\bar{c}$	uds	$\tau^+\tau^-$	data
B_{tag} Cuts	645.2 ± 0.5	259.2 ± 0.3	476.0 ± 0.3	1001.4 ± 0.5	11.6 ± 0.1	1110.4 ± 0.4
E_{miss} Cut	565.0 ± 0.5	237.3 ± 0.3	382.2 ± 0.3	711.9 ± 0.5	11.4 ± 0.1	860.7 ± 0.4
Continuum > 0.5	454.6 ± 0.4	196.0 ± 0.3	87.0 ± 0.1	174.6 ± 0.2	0.0 ± 0.0	341.7 ± 0.2
Purity Cut > 0.4	331.0 ± 0.4	94.4 ± 0.2	53.5 ± 0.1	105.1 ± 0.2	0.0 ± 0.0	215.3 ± 0.2
Three Tracks	131.8 ± 0.2	25.3 ± 0.1	18.3 ± 0.1	37.3 ± 0.1	0.0 ± 0.0	7.5 ± 0.1
Kaon PID	47.0 ± 0.1	5.84 ± 0.1	4.2 ± 0.0	6.7 ± 0.0	0.0 ± 0.0	21.7 ± 0.1
Kaon Charge Cut	43.6 ± 0.1	4.52 ± 0.0	3.3 ± 0.0	5.1 ± 0.0	0.0 ± 0.0	18.9 ± 0.1
Two lepton Cut	3.5 ± 0.0	0.1 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	1.1 ± 0.0
π^0 Veto	1.7 ± 0.0	0.1 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.6 ± 0.0
D^0, γ , and J/ψ veto	1.7 ± 0.0	0.1 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.5 ± 0.0
S_B Cut	1.4 ± 0.0	0.1 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.5 ± 0.0
Leptonic e^+e^-	0.4 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.1 ± 0.0
Leptonic $\mu^+\mu^-$	0.3 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.1 ± 0.0
Leptonic $e^+\mu^-$	0.7 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.2 ± 0.0
MLP cut	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	—

Table 6.4: Normalized efficiencies $\times 10^{-2}$ (%) in background Monte Carlo and data after each selection cut.

Cut	Generic	Signal
B_{tag} Cuts	0.51 ± 0.0	31.2 ± 0.0
E_{miss} Cut	99.1 ± 1.1	98.8 ± 0.1
Continuum Likelihood Cut >0.5	80.9 ± 0.9	64.1 ± 0.1
Purity Cut >0.4	72.4 ± 1.0	100.0 ± 0.2
Three Tracks	66.4 ± 1.1	71.8 ± 0.1
Kaon PID	62.1 ± 1.2	71.8 ± 0.1
Kaon Charge Cut	96.9 ± 2.1	100 ± 0.2
Two lepton Cut	9.4 ± 0.5	10.4 ± 0.1
π^0 Veto	82.5 ± 6.4	80.5 ± 0.6
D^0, γ , and J/ψ veto	92.8 ± 7.6	94.4 ± 0.8
S_B Cut	100.0 ± 0.0	100.0 ± 0.0
Leptonic e^+e^-	26.8 ± 3.5	27.8 ± 0.4
Leptonic $\mu^+\mu^-$	25.0 ± 4.1	22.9 ± 0.3
Leptonic $e^+\mu^-$	48.2 ± 7.1	49.2 ± 0.8
MLP cut	55.6 ± 6.5	55.0 ± 0.6

Table 6.5: Partial efficiencies (%) in $B \rightarrow K^\pm \tau^+ \tau^-$ generic and cocktail signal Monte Carlo after each selection cut.

Cut	B^+B^-	$B^0\bar{B}^0$	$c\bar{c}$	uds	$\tau^+\tau^-$	data
B_{tag} Cuts	0.10 ± 0.0	0.03 ± 0.00	0.05 ± 0.00	0.10 ± 0.00	0.001 ± 0.00	—
E_{miss} Cut	87.6 ± 0.1	91.6 ± 0.2	80.3 ± 0.1	71.1 ± 0.2	98.1 ± 1.0	77.5 ± 0.0
Continuum Likelihood Cut >0.5	80.4 ± 0.1	82.6 ± 0.2	22.8 ± 0.0	24.5 ± 0.1	0.2 ± 0.0	39.7 ± 0.0
Purity Cut >0.4	72.8 ± 0.1	48.2 ± 0.1	61.5 ± 0.1	60.2 ± 0.3	78.3 ± 20.6	63.0 ± 0.1
Three Tracks	39.8 ± 0.1	26.8 ± 0.1	34.1 ± 0.1	35.5 ± 0.2	18.4 ± 3.3	34.6 ± 0.1
Kaon PID	35.7 ± 0.1	23.1 ± 0.2	23.1 ± 0.2	18.0 ± 0.1	9.9 ± 21.2	29.1 ± 0.1
Kaon Charge Cut	92.8 ± 0.4	77.4 ± 1.0	79.1 ± 0.8	76.0 ± 0.8	56.9 ± 34.6	87.3 ± 0.3
Two lepton Cut	8.0 ± 0.1	3.1 ± 0.2	0.5 ± 0.1	0.2 ± 0.0	0.0 ± 0.0	5.8 ± 0.1
π^0 Veto	49.4 ± 1.0	43.0 ± 4.3	20.1 ± 5.0	15.8 ± 7.4	0.0 ± 0.0	50.8 ± 1.0
D^0, γ , and J/ψ veto	96.1 ± 2.2	96.3 ± 11.4	91.7 ± 30.0	70.3 ± 47.3	0.0 ± 0.0	96.0 ± 2.3
S_B Cut	84.5 ± 2.0	89.2 ± 11.1	81.5 ± 28.8	64.7 ± 53.8	0.0 ± 0.0	86.0 ± 2.1
Leptonic e^+e^-	27.0 ± 1.0	25.6 ± 5.1	24.1 ± 14.3	32.9 ± 43.0	0.0 ± 0.0	27.1 ± 1.1
Leptonic $\mu^+\mu^-$	23.4 ± 1.2	23.0 ± 6.0	26.1 ± 19.3	44.8 ± 62.5	0.0 ± 0.0	22.5 ± 1.2
Leptonic $e^+\mu^-$	49.5 ± 2.1	51.4 ± 11.6	49.7 ± 31.4	22.3 ± 36.1	0.0 ± 0.0	50.4 ± 2.3
MLP cut	4.9 ± 0.4	7.2 ± 2.6	15.8 ± 11.9	45.8 ± 78.2	0.0 ± 0.0	—

Table 6.6: Partial efficiencies (%) in background Monte Carlo and data after each selection cut.

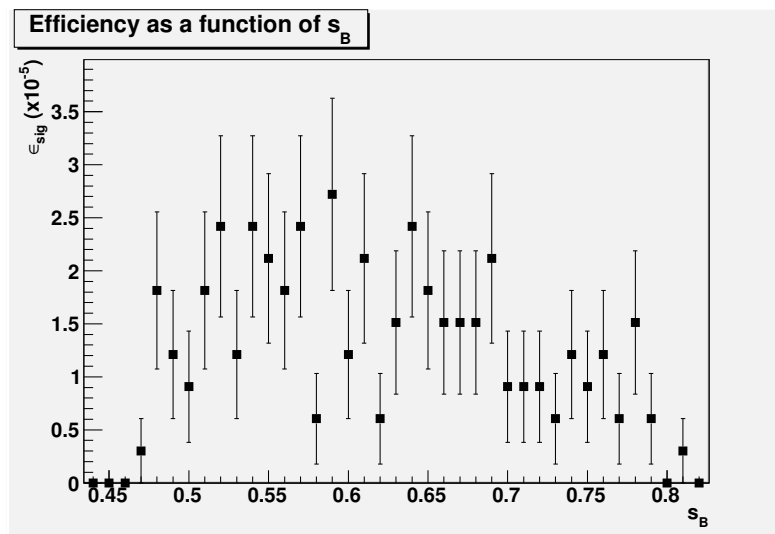


Figure 6.1: Signal efficiency as a function of s_B for all 3 modes combined.

Chapter 7

Background estimation

The number of surviving background events, after the full signal selection, is determined primarily using background MC. As can be readily seen in previous plots, after the BSemiEx-clAdd skim, the MC overestimates the data. This is mainly due to the large uncertainties associated with the hadronic B meson branching fractions and the poor simulation of certain high multiplicity modes in the MC. A method called m_{ES} sideband substitution is used to correct known differences between data and MC, before determining the final background estimate. Furthermore, because the signal efficiency determination is also dependent on the MC, the signal efficiency is also corrected by the m_{ES} sideband substitution.

7.1 m_{ES} sideband substitution

Using this method, the m_{ES} distribution, shown in Fig. 7.1, is divided into two regions: “signal” ($5.27 < m_{\text{ES}} < 5.29 \text{ GeV}/c^2$) and “sideband” ($5.20 < m_{\text{ES}} < 5.26 \text{ GeV}/c^2$). Furthermore, the Monte Carlo background is divided into two types:

- Combinatorial background: $c\bar{c}$, $\tau^+\tau^-$, $q\bar{q}$ where $q = u, d, s$, $B^0\bar{B}^0$ and B^+B^- events with mis-reconstructed B_{tag} candidates.
- Peaking background: B^+B^- events with correctly reconstructed B_{tag} candidates, which

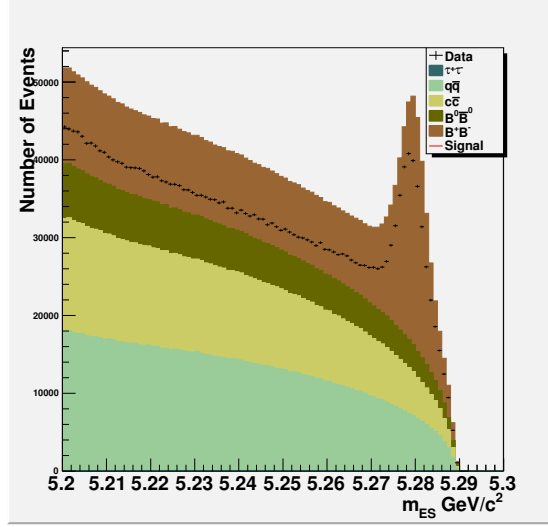


Figure 7.1: m_{ES} distribution after applying the sub-skim, continuum likelihood, purity and track multiplicity cuts. The data distribution (black points) is overlaid on the background MC distributions (color-filled). $q\bar{q}$ refers to $q = u, d, s$.

do not decay via $B^+ \rightarrow K^+ \tau^+ \tau^{-1}$.

The goal of this approach is to estimate the combinatorial background in the m_{ES} signal region directly using data from the sideband region. This way any systematic error associated with the modeling of continuum or mis-reconstructed B_{tag} events in the MC background is eliminated and the sideband data distribution is used instead. The peaking background distribution, consisting of B^+B^- , is then further corrected to match the peaking data. This is done to correct for discrepancies introduced by the BSemiExclAdd skim, which result from differences, for instance, in the modeled branching fractions of the many hadronic decay modes used in the hadronic B_{tag} reconstruction. Both the combinatorial and peaking contributions are then added to calculate the total background estimate. Doing so, the total background estimate consists of sideband data and B^+B^- MC only. The other MC types are only used to determine the shape of the combinatorial component of the m_{ES} distribution and its normalization. The advantage of this approach is to reduce the reliance on MC simulation. Also, it allows for an improved agreement between data and Monte Carlo before

¹This method is described for the case of a charged B_{tag} . For neutral modes, the $B^0\bar{B}^0$ and B^+B^- backgrounds are swapped.

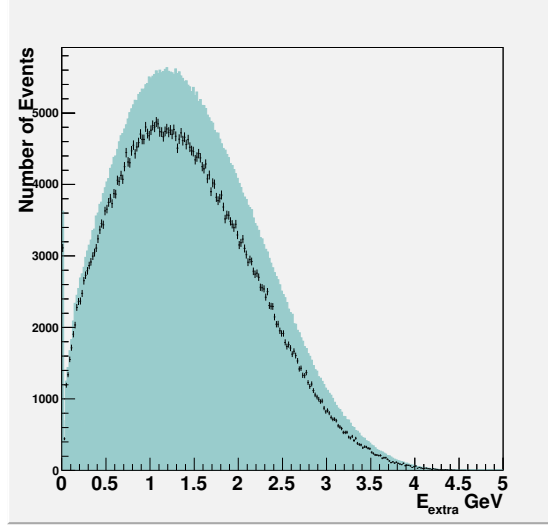


Figure 7.2: E_{extra} distribution after applying sub-skim, m_{ES} , continuum likelihood, purity and track multiplicity cuts and before m_{ES} sideband substitution. The background distribution is shown filled with blue, while the data is overlaid as black dots.

unblinding and thus a more accurate estimate of the background levels and signal efficiency.

Fig. 7.2 shows the E_{extra} distribution after applying the sub-skim level cuts, as well as the purity, continuum likelihood, and track multiplicity cuts. As can be readily seen, the MC overestimates the data and this is a familiar discrepancy resulting from the BSemiExclAdd skim. By implementing the m_{ES} sideband substitution, this discrepancy can be corrected. Any residual data/MC differences after the m_{ES} sideband substitution are then accounted for with systematic uncertainties.

First, the combinatorial background is estimated using data from the m_{ES} sideband region. This sideband data is scaled by a Combinatoric Ratio, R_{comb} , which is calculated using MC simulation as the ratio of *combinatorial* events in the m_{ES} signal region to events in the sideband region. R_{comb} is given by the following equation:

$$R_{comb} = \frac{N_{MC}^{signal}}{N_{MC}^{sideband}} \equiv \frac{N_{\tau^+\tau^-}^{sig} + N_{q\bar{q}}^{sig} + N_{B^0\bar{B}^0}^{sig} + (N_{B^+B^-}^{side} \times R_{B^0\bar{B}^0})}{N_{\tau^+\tau^-}^{side} + N_{q\bar{q}}^{side} + N_{B^0\bar{B}^0}^{side} + N_{B^+B^-}^{side}} \quad (7.1)$$

where here N_{MC}^{sig} is the number of events in the m_{ES} signal region for a certain MC type, N_{MC}^{side} is the number of events in the m_{ES} sideband region and $R_{B^0\bar{B}^0}$ is the ratio calculated for

the $B^0\bar{B}^0$ MC only. Because this is a charged mode, B^+B^- MC background consists of both a peaking and combinatorial (non-peaking) contribution. The former has been discussed in section 5.3, whereas the latter generally consists of mis-reconstructed B_{tag} candidates. It is hard to separate the combinatoric contribution from the peaking one for this MC type and thus the combinatorial shape of B^+B^- in the m_{ES} signal region is modeled using $B^0\bar{B}^0$ MC. This is done since charged and neutral B mesons have similar kinematic properties and thus tend to get mis-reconstructed in similar ways, regardless of charge. The combinatorial shape of the $B\bar{B}$ MC is further discussed in Appendix D. The combinatoric component of B^+B^- is thus determined by multiplying $N_{B^+B^-}^{\text{side}}$ by $R_{B^0\bar{B}^0}$, as shown in the equation 7.1. Here, $R_{B^0\bar{B}^0}$ is given by

$$R_{B^0\bar{B}^0} \equiv \frac{N_{B^0\bar{B}^0}^{\text{sig}}}{N_{B^0\bar{B}^0}^{\text{side}}} \quad (7.2)$$

and the combinatorial component of B^+B^- in the m_{ES} signal region is thus

$$N_{B^+B^-}^{\text{sig}}(\text{combinatorial}) = N_{B^+B^-}^{\text{side}} \times R_{B^0\bar{B}^0} = N_{B^+B^-}^{\text{side}} \times \frac{N_{B^0\bar{B}^0}^{\text{sig}}}{N_{B^0\bar{B}^0}^{\text{side}}} \quad (7.3)$$

The sideband data is then scaled by R_{comb} to obtain the combinatorial background estimate. R_{comb} is thus considered the scale factor that gives the correct normalization to the sideband data, such that an accurate estimate of the combinatorial background is made. Fig. 7.3 shows the combinatorial component of the E_{extra} distribution determined by scaling the sideband data by R_{comb} . This is shown in contrast to the full E_{extra} distribution, combinatorial + peaking, before the m_{ES} sideband substitution.

After estimating the combinatorial background, the peaking background is then determined. In this analysis, the peaking component consists of B^+B^- MC only, after subtracting the combinatoric part. As previously mentioned, the assumption is that the combinatorial B^+B^- component in the signal region has the same shape as that of $B^0\bar{B}^0$. Therefore, the

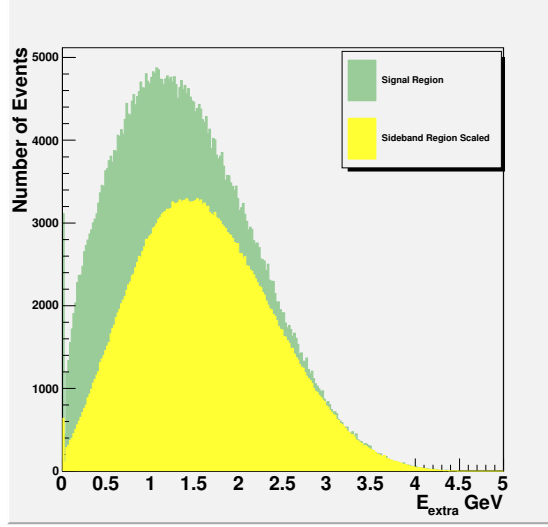


Figure 7.3: Combinatorial background estimate (yellow) of the E_{extra} distribution calculated by scaling the sideband data with R_{comb} . The full E_{extra} distribution in the m_{ES} signal region is shown (green) to highlight the fraction of combinatorial events.

peaking B^+B^- component is given by:

$$N_{B^+B^-}^{pkg} = N_{B^+B^-}^{sig} - (N_{B^+B^-}^{side} \times R_{B^0\bar{B}^0}) \quad (7.4)$$

where the superscript pkg stands for peaking component. This peaking B^+B^- component is then scaled by the B_{tag} Yield Correction, C_{yield} , to match the peaking data. The latter is calculated by subtracting the combinatorial background estimate, as given by the equation:

$$N_{data}^{pkg} = N_{data}^{sig} - (N_{data}^{side} \times R_{comb}) \quad (7.5)$$

Fig. 7.4 shows the m_{ES} distribution of the peaking data and peaking B^+B^- components after applying the above equations. In addition, Fig. 7.5 shows the total E_{extra} distribution when the peaking B^+B^- contribution is added to the combinatorial component. As can be readily seen, the peaking MC overestimates the peaking data in the m_{ES} distributions and the resulting E_{extra} distribution also displays this effect. To correct for this, the peaking B^+B^- is scaled by C_{yield} . C_{yield} is calculated as the ratio of the event yields in peaking data

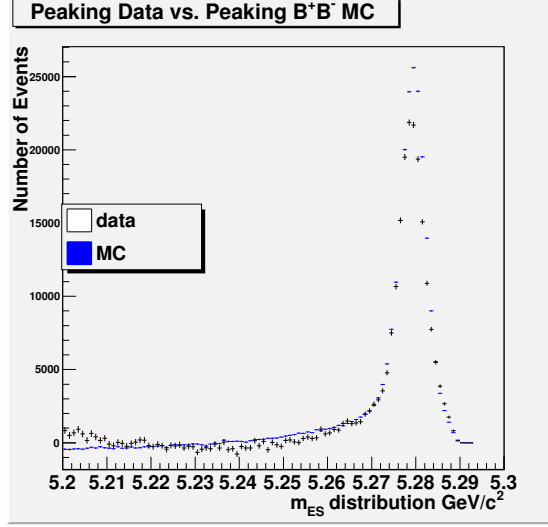


Figure 7.4: m_{ES} distribution of peaking data (black) vs. peaking B^+B^- MC (blue) after applying sub-skim, continuum, purity, and track multiplicity cuts.

to peaking B^+B^- :

$$C_{yield} = N_{data}^{pkg} / N_{B^+B^-}^{pkg} \quad (7.6)$$

Fig. 7.6 shows the final E_{extra} distribution after determining the combinatorial component from sideband data and scaling the peaking B^+B^- to match the peaking data. The agreement between data and MC is considerably enhanced and the MC no longer overestimates the data. C_{yield} is thus considered the scale factor that corrects the overestimate of the B^+B^- MC. To quantify the level of agreement, Fig.7.7 shows the ratio of the data and MC yields over the full E_{extra} distribution. The ratio has a value close to 1, except in the region above 4 GeV where the statistics are low. If this region is excluded, a linear fit to the ratio yields a slope of 0.96 ± 0.2 . This gives the size of the residual discrepancy between data and MC, which is $< 5\%$ and will be accounted for in the B_{tag} yield systematic uncertainty.

The m_{ES} sideband substitution will be used to obtain the background estimate at the end of the selection. R_{comb} and C_{yield} vary as a function of cut in the signal selection. At the final stage in the signal selection, data is blinded and thus C_{yield} cannot be determined.

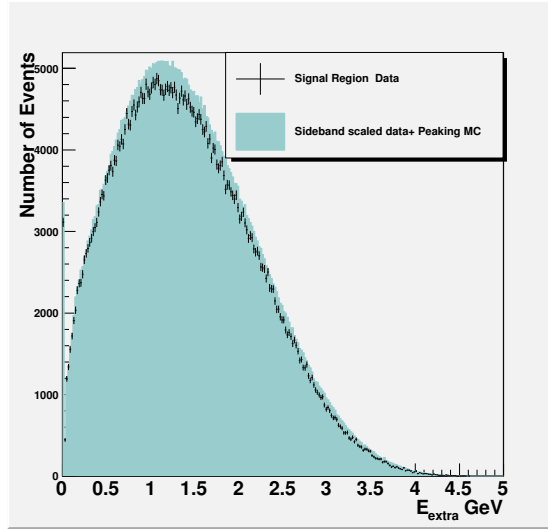


Figure 7.5: E_{extra} distribution of data (black) vs. background MC (blue), where the background MC consists of the combinatorial contribution, determined using sideband data, and a peaking B^+B^- component without C_{yield} correction.

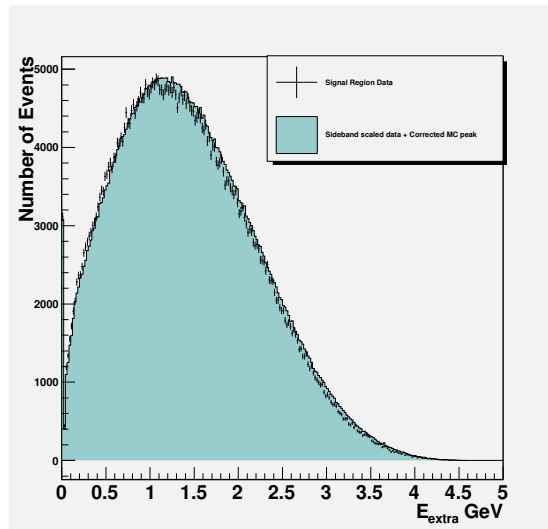


Figure 7.6: E_{extra} distribution after the m_{ES} sideband substitution and peaking background correction. The agreement between data (black) and background MC (blue) is much improved.

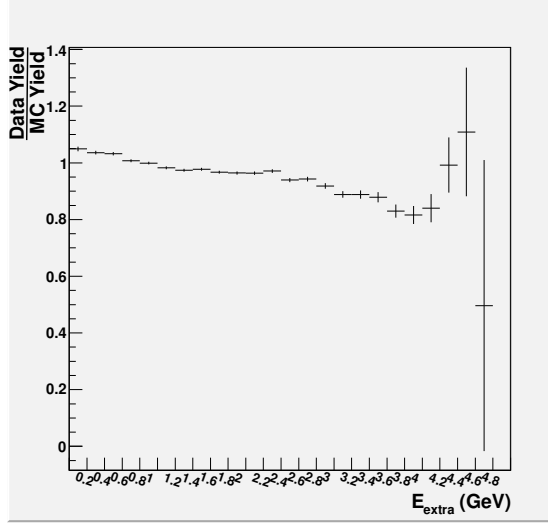


Figure 7.7: Ratio of data and MC yield in the E_{extra} distribution after the m_{ES} sideband substitution.

The peaking data component, at the end of the selection, may consist of $B^+ \rightarrow K^+ \tau^+ \tau^-$ events. It is thus important to apply the C_{yield} correction from an earlier stage, after the s_B cut, when the data sample is still dominated by background events. This ensures that the m_{ES} sideband substitution does not affect the ability to observe signal events in data at the end of the signal selection, while simultaneously correcting for known data-MC differences. For the combinatorial background, R_{comb} is determined at the end of the signal selection, since the signal is not combinatoric. After the MLP cut, data in the m_{ES} sideband region consists of all events that survive the signal selection, except for the m_{ES} cut, and have a m_{ES} ranging between 5.20 and 5.26 GeV/ c^2 . If M denotes the number of surviving events after the full signal selection, including the MLP cut, then

$$\begin{aligned}
 M_{comb} &= M_{data}^{side} \times R_{comb}, \\
 M_{pkg} &= M_{B^+B^-}^{pkg} \times C_{yield} = (M_{B^+B^-}^{sig} - (M_{B^+B^-}^{side} \times R_{B^0\bar{B}^0})) \times C_{yield}, \\
 M_{bkg} &= M_{comb} + M_{pkg},
 \end{aligned} \tag{7.7}$$

where here M_{comb} is the final combinatorial background estimate, M_{pkg} is the final peaking background estimate and M_{bkg} is the total background estimate at the end of the selection. It should be noted that the m_{ES} cut has been applied here according to the *sig* and *side* superscripts. As previously mentioned, C_{yield} is determined after the s_B cut, and is thus only included as a pre-determined scale factor at the end of the selection. The equation for C_{yield} is exactly as given in equation 7.6, if N is the number of events surviving the s_B cut. On the other hand, R_{comb} is determined at the end of the signal selection.

Because signal MC is also classified as peaking B^+B^- MC, the same level of discrepancy between signal MC and data should be applicable. This should be corrected for to obtain an accurate estimate of the signal efficiency. Thus, the final signal efficiency is also scaled by C_{yield} :

$$\begin{aligned}\epsilon_{sig} &= \frac{M_{sig}}{N_{generated}} \times C_{yield}, \\ \epsilon_{sig} &= \frac{M_{sig}}{N_{generated}} \times \frac{N_{data}^{pkg}}{N_{B^+B^-}^{pkg}},\end{aligned}\tag{7.8}$$

where here M_{sig} is the number of events in the generic signal MC that survive the final MLP neural network cut and $N_{generated}$ is the total number of generated events in the generic signal MC.

7.2 Combinatorial background estimate

The value of R_{comb} varies as a function of cut. Fig. 7.8 shows a plot of R_{comb} after each cut in the analysis. The variation in R_{comb} reflects the change in the shape and composition of the combinatorial background as the signal selection cuts are applied.

After applying the lepton PID cut, the value of the ratio almost doubles, implying that the number of MC events in the sideband region is considerably decreased relative to the number of events in the signal region. This can be seen in Table 6.2, where the number of continuum background events is reduced by at least 3 orders of magnitude. The dominant

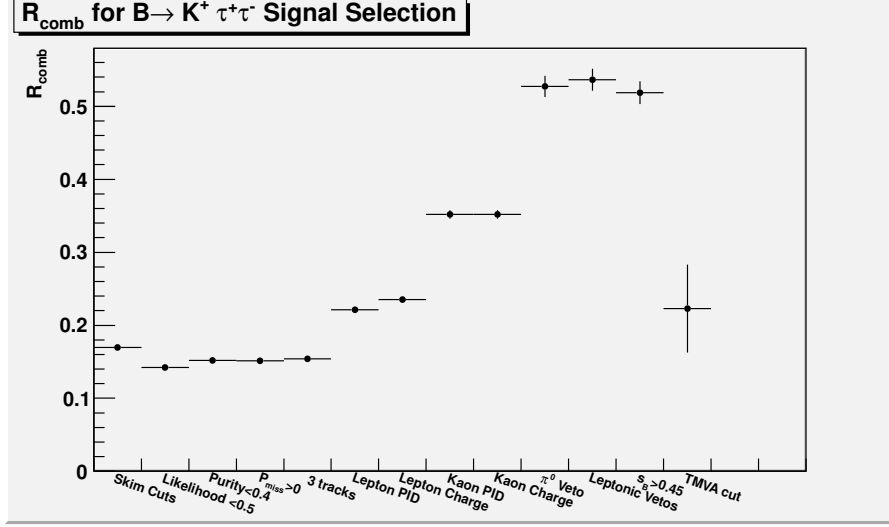
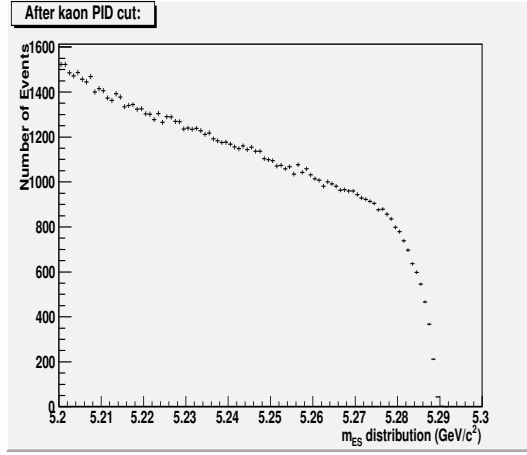


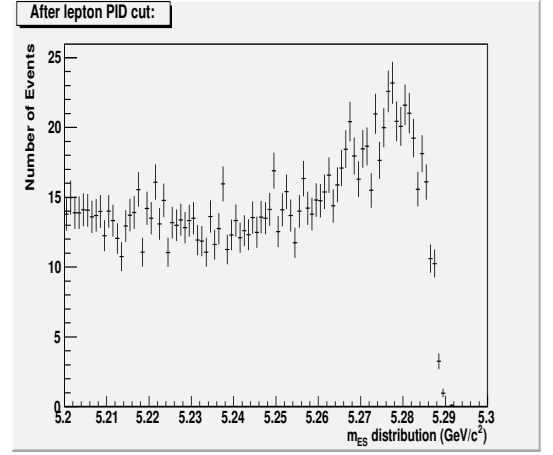
Figure 7.8: R_{comb} as a function of cut in the signal selection.

combinatorial background at this stage is $B^0\bar{B}^0$, which has also been reduced by at least 2 orders of magnitude. Fig. 7.9 and Fig. 7.10 show the *unweighted* m_{ES} distribution of the $B^0\bar{B}^0$ and $c\bar{c}$ backgrounds after each of the kaon PID, lepton PID, π^0 veto, s_B and MLP cuts. The change in the shape of the $B^0\bar{B}^0$ background is evident after applying the lepton PID cut. A broad peaking shape in the $B^0\bar{B}^0$ forms after this cut and becomes more prominent as the remaining cuts are applied, except for the final MLP cut. This peaking shape causes the shift in the value of the $B^0\bar{B}^0$ combinatorial ratio as there are more events in the m_{ES} signal region than in the sideband region. On the other hand, the $c\bar{c}$ background retains its shape after the lepton PID cut but considerably decreases in size. The same is true after the remaining cuts in the signal selection. After the MLP cut, the shape of the $B^0\bar{B}^0$ distribution becomes more or less flat and the number of surviving events is of order 1. The MLP cut suppresses this peaking behaviour in $B^0\bar{B}^0$ and removes a large amount of the surviving combinatorial background.

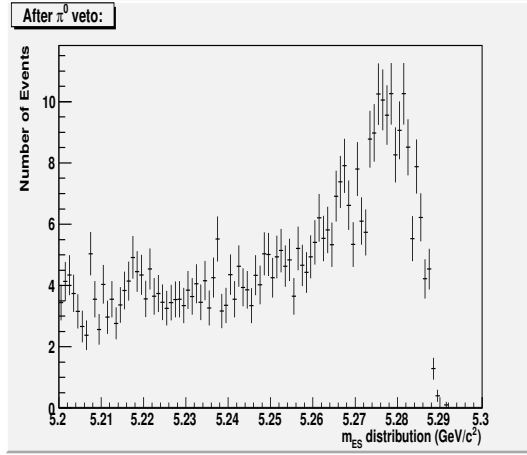
To estimate the combinatorial background at any given stage of the selection, the value of R_{comb} must be determined at that stage. Consequently, for the final background estimate, the value is taken after the MLP cut. This ensures that the selected value of R_{comb} , 0.213 ± 0.018 where the uncertainty is statistical only, accounts for the change in the shape of the different



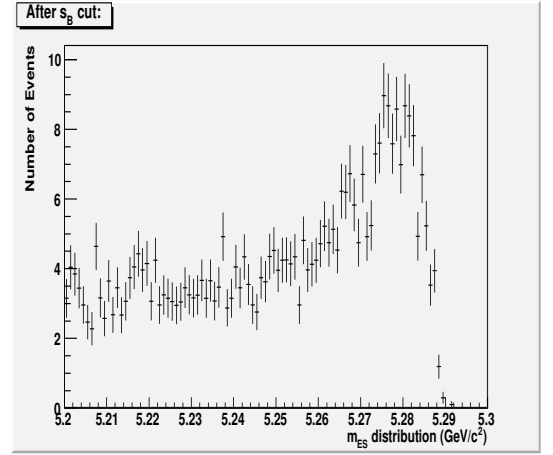
(a) After kaon PID cut



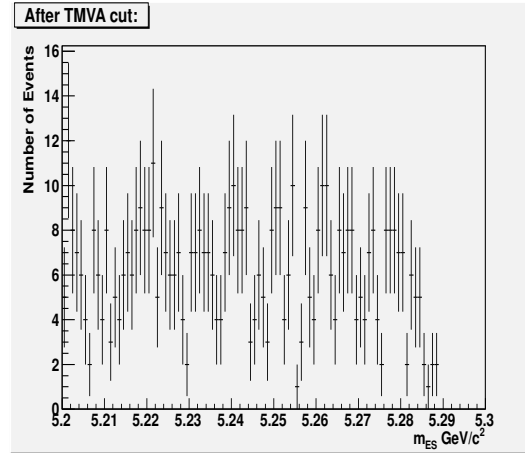
(b) After lepton PID cut



(c) After π^0 veto

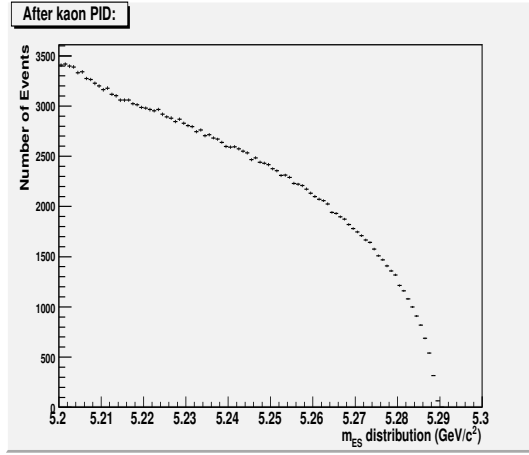


(d) After s_B cut

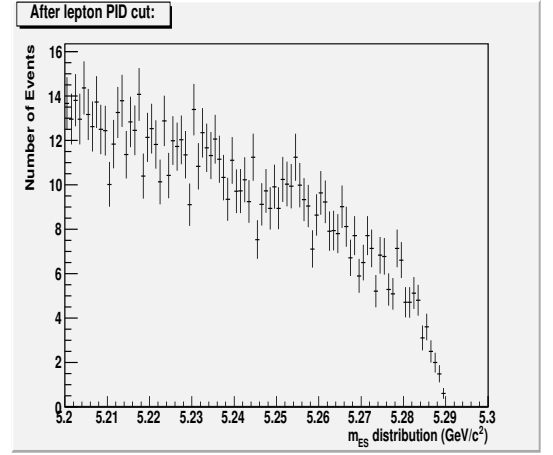


(e) After MLP cut

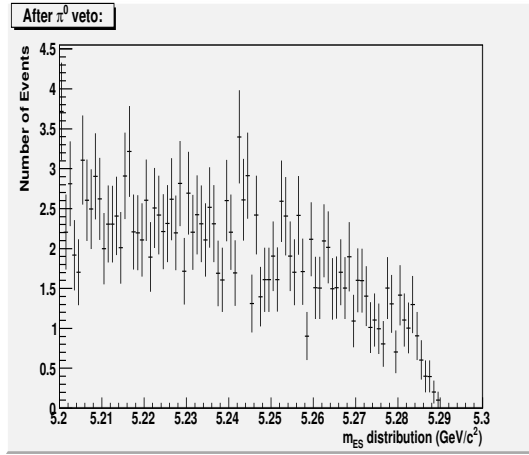
Figure 7.9: m_{ES} distribution of $B^0\bar{B}^0$ background after various cuts in the signal selection.



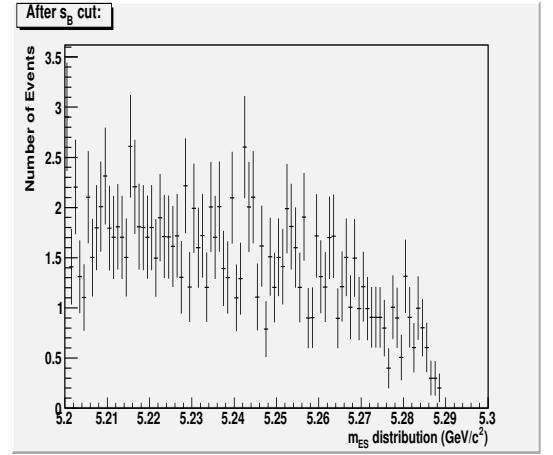
(a) After kaon PID cut



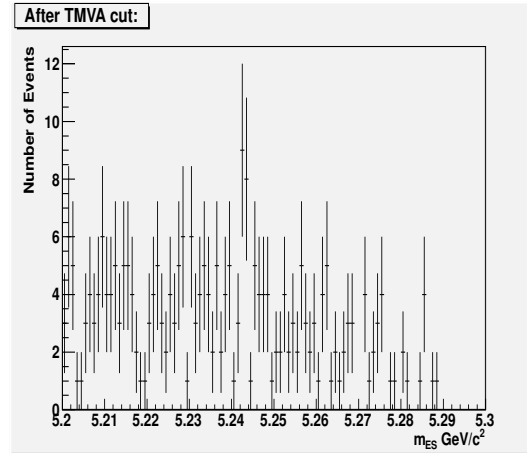
(b) After lepton PID cut



(c) After π^0 veto



(d) After s_B cut



(e) After MLP cut

Figure 7.10: m_{ES} distribution of $c\bar{c}$ background after various cuts in the signal selection.

combinatorial background components.

7.3 Peaking background estimate

In contrast with R_{comb} , C_{yield} is relatively stable throughout the cut flow, after applying the skim-level, purity and continuum likelihood cuts. These initial cuts remove most of the mis-reconstructed B_{tag} candidates, with decay modes that usually have a low purity value. The stability of C_{yield} throughout the rest of the signal selection is expected behaviour. This is true since the C_{yield} discrepancy is attributed primarily to modelling of B_{tag} mode branching fractions, and hence its value is independent of signal-side cuts. Fig. 7.11 shows a plot of C_{yield} after each cut in the signal selection. The B^+B^- background retains its shape throughout the cut flow, as shown in Fig. 7.12, and is only significantly reduced by the final MLP cut. This peaking B^+B^- background overestimates the data throughout the signal selection, excluding the MLP cut where data is still blinded.

The value of C_{yield} chosen for the peaking background estimate is that after the s_B cut, 0.914 ± 0.020 , where the uncertainty is statistical only. This ensures maximum possible agreement between data and MC before the final MLP cut. At the same time, because a large number of B^+B^- background events survives the s_B cut, choosing C_{yield} at this stage does not impact the ability to observe signal events in data.

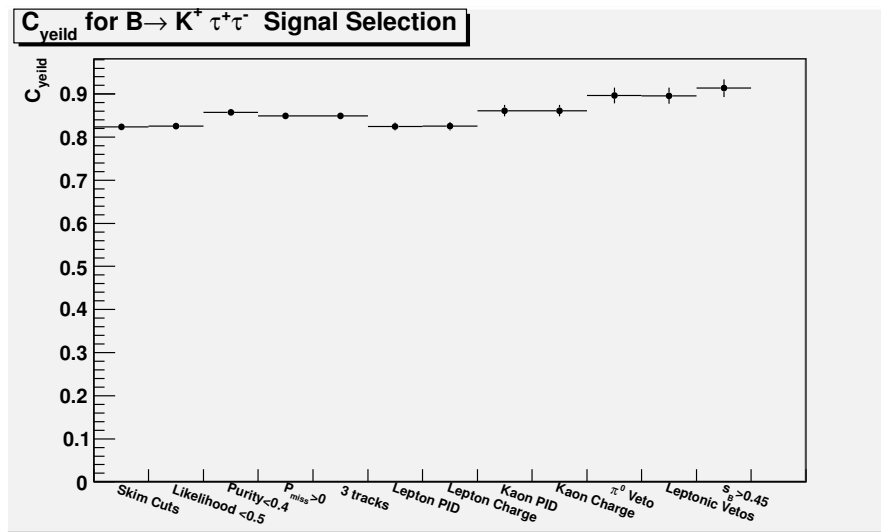
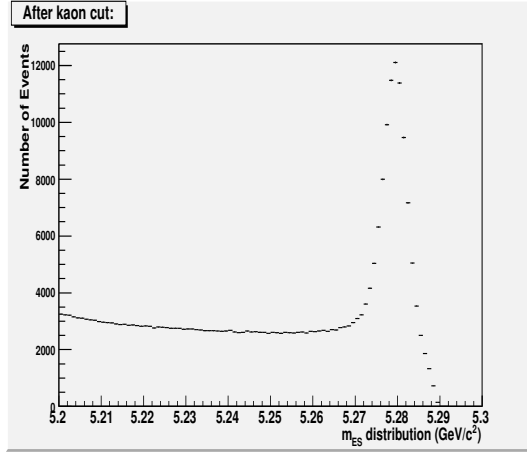
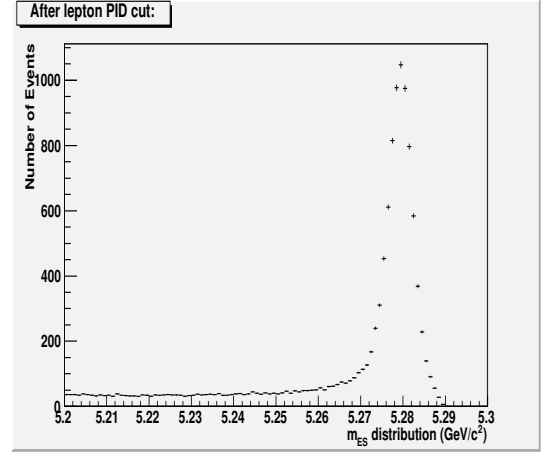


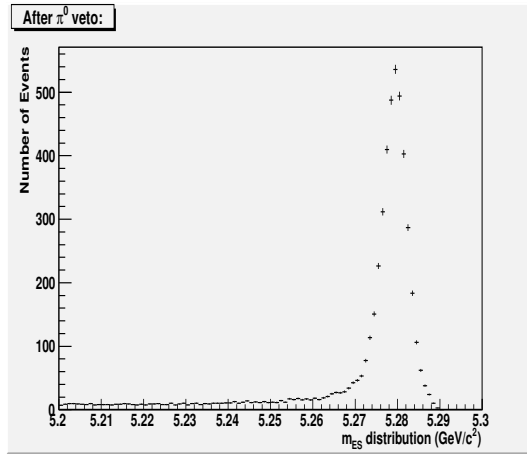
Figure 7.11: C_{yield} as a function of cut.



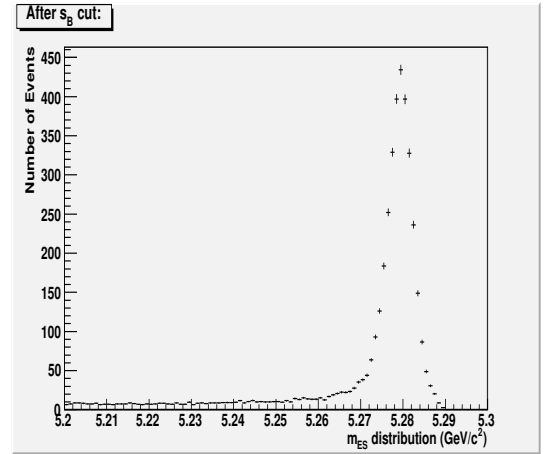
(a) After kaon PID cut



(b) After lepton PID cut



(c) After π^0 veto

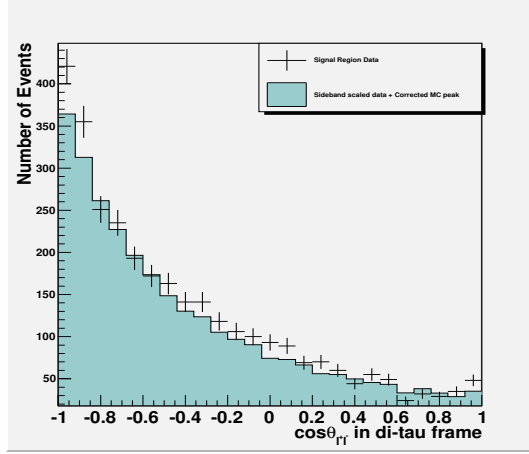


(d) After s_B cut

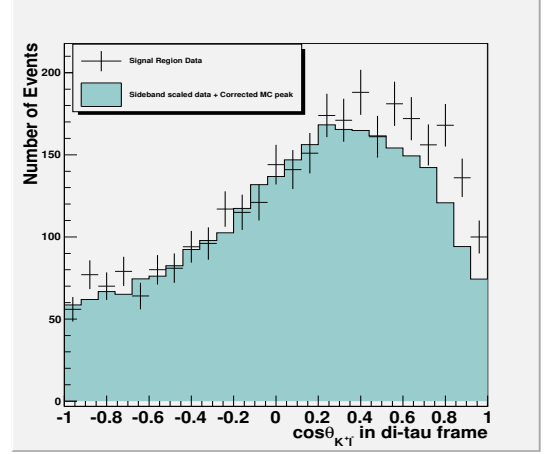
Figure 7.12: m_{ES} distribution of B^+B^- background after various cuts in the signal selection.

7.4 Discriminating variables after m_{ES} sideband substitution

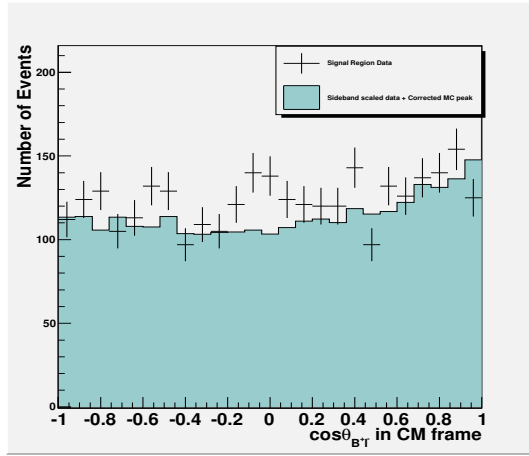
The discriminating variables that enter into the MLP neural network are shown below after the s_B cut and after the m_{ES} sideband substitution.



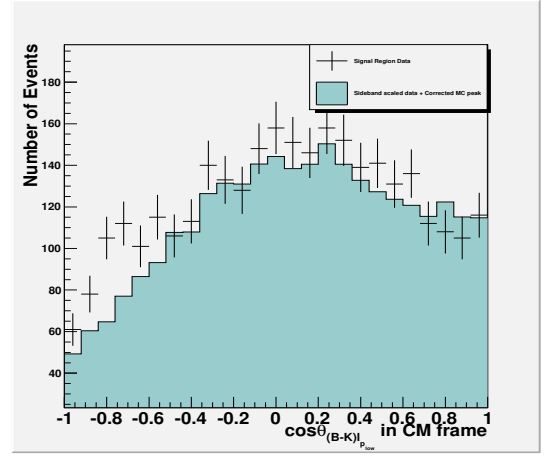
(a) $\cos \theta_{l+l-}$ in di- τ frame.



(b) $\cos \theta_{K+l-}$ in di- τ frame.

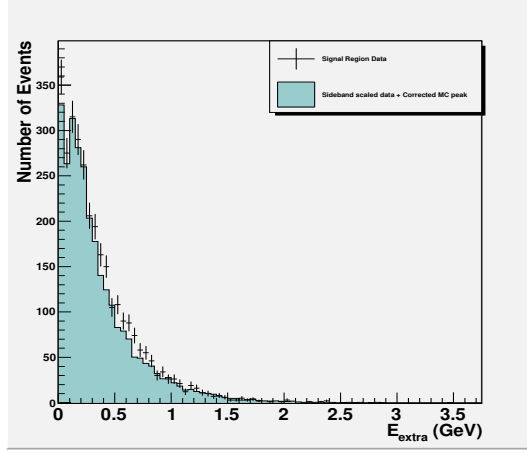


(c) $\cos \theta_{B+l-}$ in CM frame.

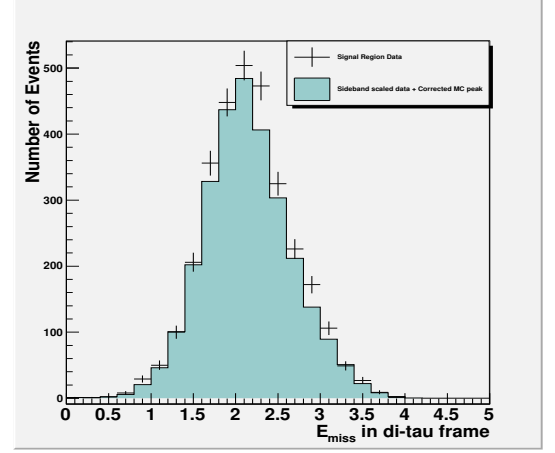


(d) $\cos \theta_{(B-K)l_{low}}$ in di- τ frame.

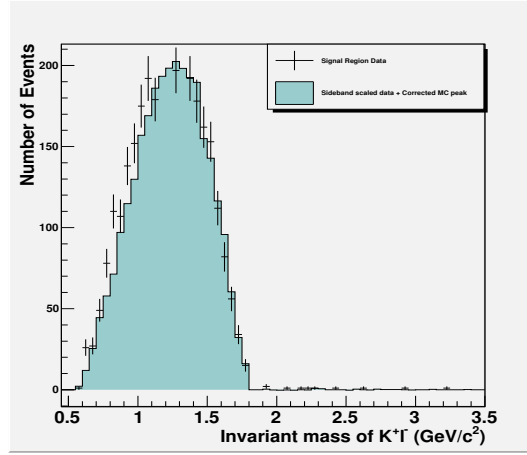
Figure 7.13: Angular variables used in the MLP neural network, after applying the s_B cut and m_{ES} sideband substitution.



(a) E_{extra}



(b) E_{miss}



(c) Invariant mass of $K^+ l^-$ in CM frame.

Figure 7.14: Calorimeter variables used in the MLP neural network, after applying the s_B cut and m_{ES} sideband substitution.

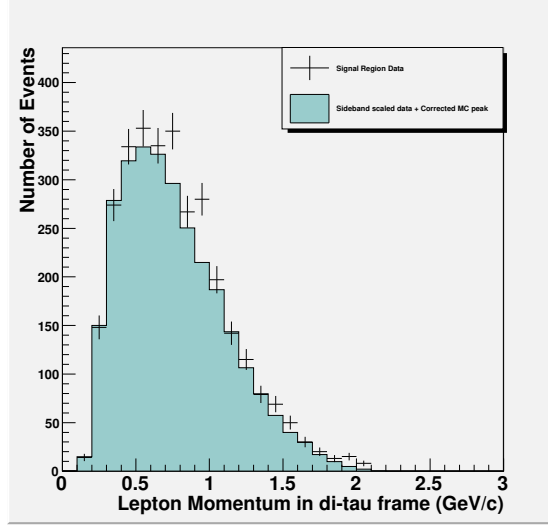


Figure 7.15: Lepton Momentum in the di-tau frame, used in the MLP neural network, after applying the s_B cut and m_{ES} sideband substitution.

7.5 $B^+ \rightarrow D^0 \ell^+ \bar{\nu}_\ell$ study: MLP neural network validation

The MLP cut is the most important cut in the analysis because it suppresses the dominant source of background. It is also the last step in the signal selection, at which point the signal efficiency and background estimate are determined. It is thus important to verify the level of data-MC agreement at this stage selection. As previously mentioned, the data is blinded and therefore a validation within the signal region cannot be done. Instead, a control sample is used to perform a check on the TMVA technique used in this analysis. The control sample chosen for this check is a $B^+ \rightarrow D^0 \ell^+ \bar{\nu}_\ell$, $D^0 \rightarrow K^- \pi^+$ sample. It is obtained from data after applying the full signal selection cuts discussed above, *except* the MLP cut, with the following variation:

- The lepton PID for one of the tracks is reversed to a pion PID, as discussed in section 5.2.2.
- The D^0 veto, discussed in section 5.2.4, is reversed: $1.80 < M_{K^\pm \pi^\mp} < 1.90 \text{ GeV}/c^2$.

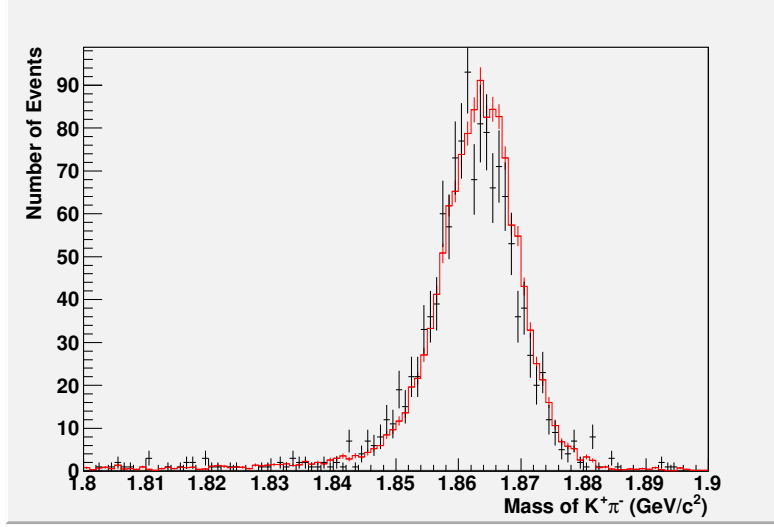


Figure 7.16: B^+B^- MC (red) vs. data (points with error bars) after applying the control sample selection.

Basically, the signal selection is reversed such that peaking background events, $B^+ \rightarrow D^0 \ell^+ \bar{\nu}_\ell$, $D^0 \rightarrow K^- \pi^+$ are chosen. These are the events that the MLP neural network is trained to classify as background.

The cuts above are also applied to the generic MC samples to allow for a direct comparison with control sample data. The number of surviving $B^0\bar{B}^0$ and continuum events is very small (~ 32 normalized events). The result, shown in Fig. 7.16, is a clear D^0 peak in both the control sample and B^+B^- MC.

The control sample data and MC are then run through the previously trained MLP neural network. The discriminating variables used in the neural network for this validation test are shown in Appendix E, before and after the m_{ES} sideband substitution. The neural network is optimized such that it can distinguish between $B^+ \rightarrow K^+ \tau^+ \tau^-$ events and the events that dominate the current B^+B^- and data samples. Consequently, events in the control sample tend to populate lower values of the MLP output, as shown in Fig. 7.17. Furthermore, Fig. 7.18a shows the MLP output for both data and the total MC sample, where the latter includes the combinatorial backgrounds ($B^0\bar{B}^0$ and continuum). The level of agreement is satisfactory and is improved using the m_{ES} sideband substitution, as shown in Fig. 7.18b.

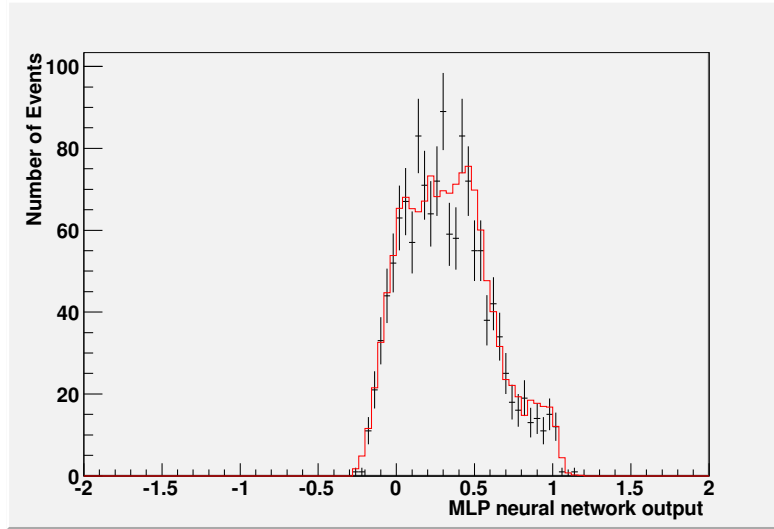
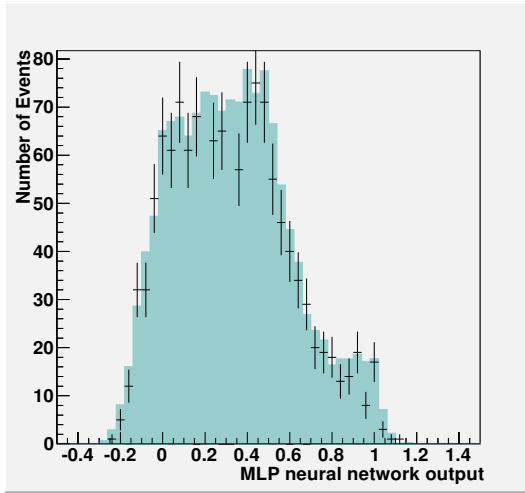
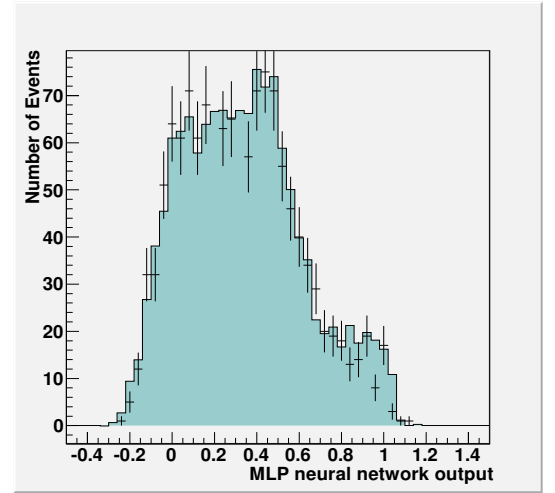


Figure 7.17: MLP neural network output for both the B^+B^- (red) and the data control sample (points with error bars).

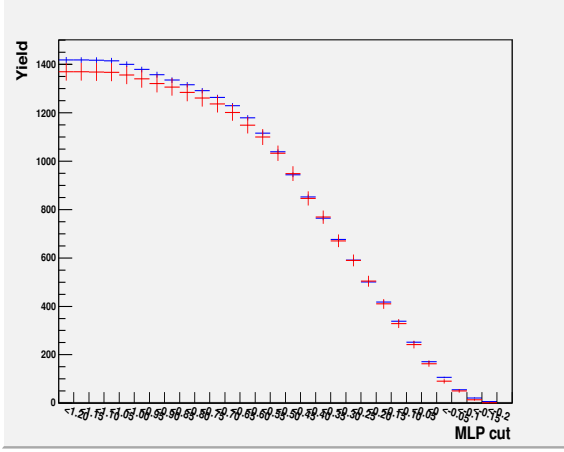


(a) Before m_{ES} sideband substitution.

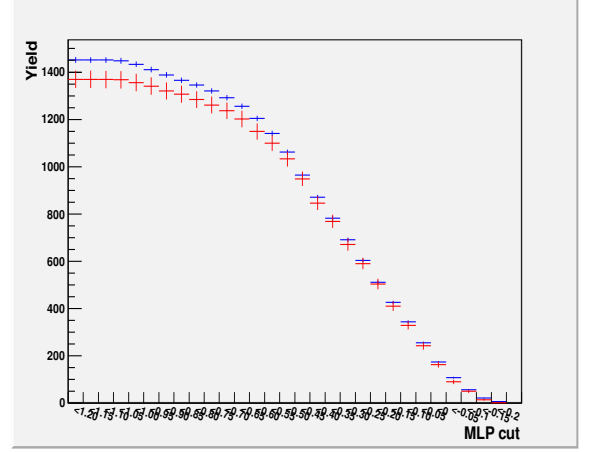


(b) After m_{ES} sideband substitution.

Figure 7.18: MLP neural network output for total MC (blue) and data (points with error bars).



(a) Data (red) and B^+B^- sample (blue).

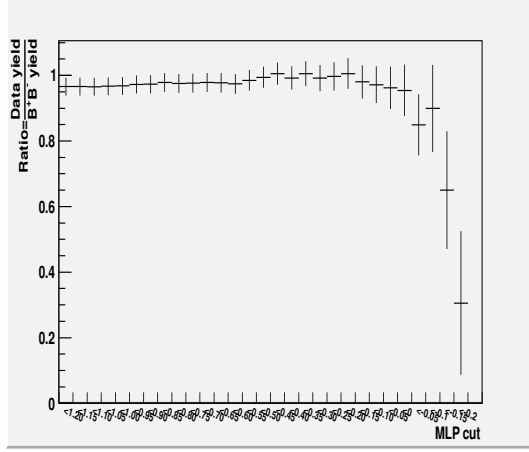


(b) Data (red) and total MC sample (blue).

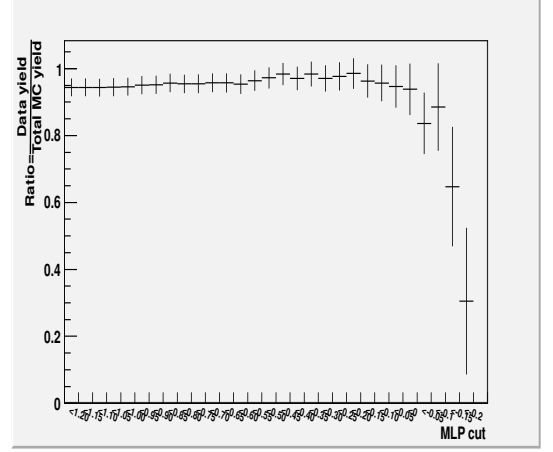
Figure 7.19: Event yields as a function of MLP cut.

To further verify the level of data-MC agreement, the MLP neural network cut is reversed and applied on both the data and B^+B^- MC samples, in increments of 0.05. Fig. 7.19a shows the result and a good agreement between data and the B^+B^- sample is clearly visible. The same is done for total MC and the agreement with data is also satisfactory, as shown in Fig. 7.19b.

To quantify the level of agreement, a ratio is calculated and is given by $R = N_{data}/N_{MC}$, where N_{data} and N_{MC} are the event yields for the data and MC samples respectively. Fig. 7.20a shows a plot of the ratio as a function of each MLP cut for the data and B^+B^- samples. Except near the end of the spectrum, where the MLP cut removes almost all of the available statistics, the value of this ratio is close to unity. Fig. 7.20b also shows the ratio of the data yield to the total MC sample, which is also close to unity for almost the full spectrum. This implies that the level of agreement between data and MC is satisfactory. The validation test shows that the MLP neural network does not bias the data-MC agreement. The level of disagreement is $< 5\%$, which is expected with the MC sample, as a result of the hadronic B_{tag} reconstruction, and will be accounted for in the systematics section.



(a) Data yield/ B^+B^- yield

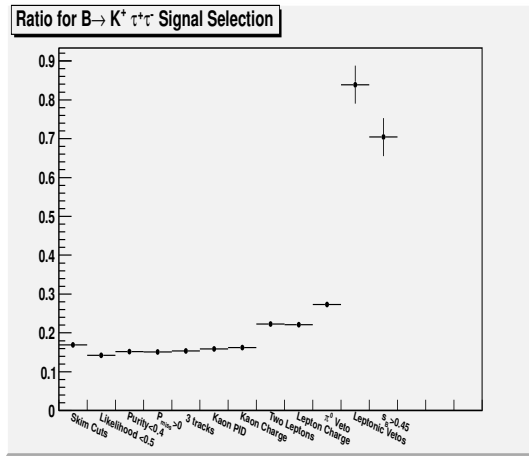


(b) Data yield/Total MC yield

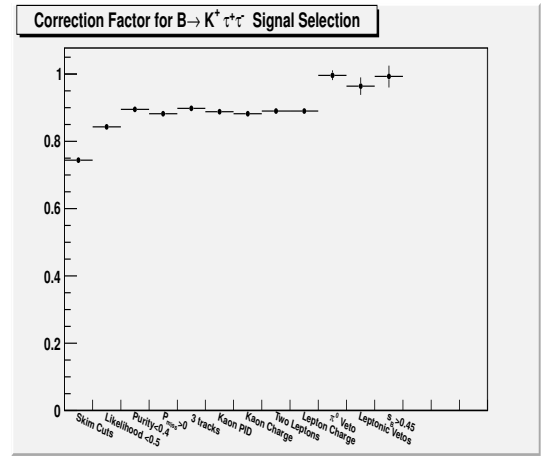
Figure 7.20: Ratio of the event yields as a function of MLP cut.

Fig. 7.21a and 7.21b show the plots of C_{yield} and R_{comb} as a function of cut flow in this validation study. The values of R_{comb} and C_{yield} are important, because they can be used to estimate the final background. As can be seen, the distribution of R_{comb} does not show the same behaviour after the PID cuts as that seen in the R_{comb} distribution of the actual signal selection. This implies that the shape of the background distributions after applying the control sample selection is not the same as that after applying the $B^+ \rightarrow K^+ \tau^+ \tau^-$ signal selection.

Furthermore, the C_{yield} distribution shows a good agreement between data and Monte Carlo with the value of C_{yield} reaching 0.993 after applying the s_B cut. This implies that the decay mode $B^+ \rightarrow D^0 \ell^+ \bar{\nu}_\ell$, $D^0 \rightarrow K^+ \pi^-$ is well modeled in data and in MC. However, the background discussed in section 5.3 consists of a combination of modes and thus the same level of agreement between data and MC for the other modes is not necessarily expected. This is why the values of R_{comb} and C_{yield} from the validation test are not used in the final background estimate.



(a) R_{comb}



(b) C_{yield}

Figure 7.21: R_{comb} and C_{yield} as a function of cut flow in control sample selection.

Chapter 8

Branching Fraction Calculation

The branching fraction for $B^+ \rightarrow K^+ \tau^+ \tau^-$ is calculated using the following equation:

$$\mathcal{B}(B^+ \rightarrow K^+ \tau^+ \tau^-) = \frac{N_{obs} - M_{bkg}}{N_{B\bar{B}} \times \epsilon_{sig}} \quad (8.1)$$
$$\mathcal{B}(B^+ \rightarrow K^+ \tau^+ \tau^-) = \frac{N_{obs} - ((M_{data}^{side} \times R_{comb}) + (M_{B^+B^-}^{sig} \times C_{yield}))}{N_{B\bar{B}} \times (\frac{M_{sigMC}^{sig}}{N_{sigMC}^{gen}} \times C_{yield})}$$

where N_{obs} is the number of signal events observed after unblinding, M_{bkg} is the expected background estimate, $N_{B\bar{B}}$ is the number of $B\bar{B}$ pairs in the data samples, and ϵ_{sig} is the signal efficiency. M_{bkg} is the sum of two components, a combinatorial component calculated by scaling the sideband data, M_{data}^{side} , by R_{comb} , and a peaking component obtained by scaling the peaking B^+B^- MC in the m_{ES} signal region by C_{yield} . The signal efficiency is calculated as the ratio of events in the generic signal MC surviving the MLP cut, M_{sigMC}^{sig} , to the total number of events generated for the generic signal MC sample, N_{sigMC}^{gen} . The signal MC is considered peaking B^+B^- MC and thus the efficiency is scaled by the same C_{yield} to correct for the overestimate of MC to data. The value of $N_{B\bar{B}}$, corresponding to the dataset used for this analysis is determined to be $(470.97 \pm 0.02) \times 10^6$ [77].

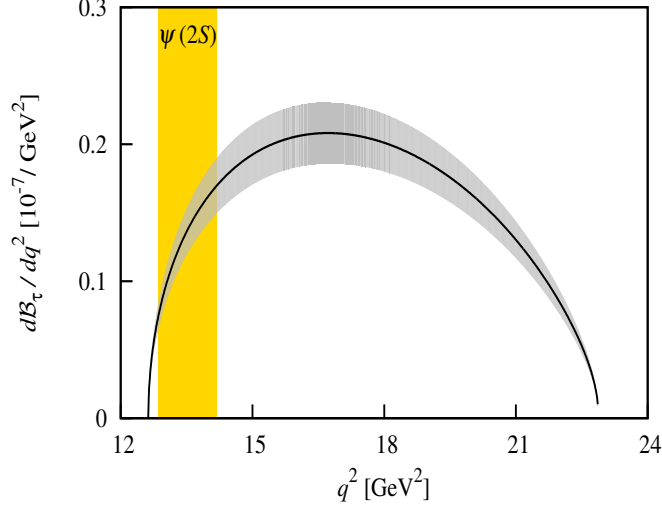


Figure 8.1: Predicted Standard Model contribution to $d\mathcal{B}_\tau/dq^2$ calculated using unquenched lattice QCD approach [57].

8.1 Signal reweighting

Signal MC in this analysis is generated according to the BTOSLLBALL model, where the form factors that describe the long-distance (non-perturbative) effects on the effective Hamiltonian of $B \rightarrow K \ell^+ \ell^-$ are calculated using a LCSR approach [55]. This approach is described in detail in Ref. [84] and Ref. [85]. Recently, the form factors for $B \rightarrow K \ell^+ \ell^-$ have been determined based on unquenched lattice calculations and are extrapolated over the entire q^2 spectrum in a model independent way [57]. The resulting Standard Model predictions were found to be consistent with previous theoretical and experimental results. The predicted q^2 distribution of $B \rightarrow K \tau^+ \tau^-$, using the unquenched lattice QCD approach, is shown in Fig. 8.1. Because these calculations are considered to be the most accurate to date, the final signal efficiency in this analysis is quoted after reweighting the signal s_B distribution to that predicted by the unquenched lattice QCD model. The reweighted signal MC distribution is shown in Fig. 8.2. The effect of the reweighting is found to be very small, less than 2% for all 3 modes combined.

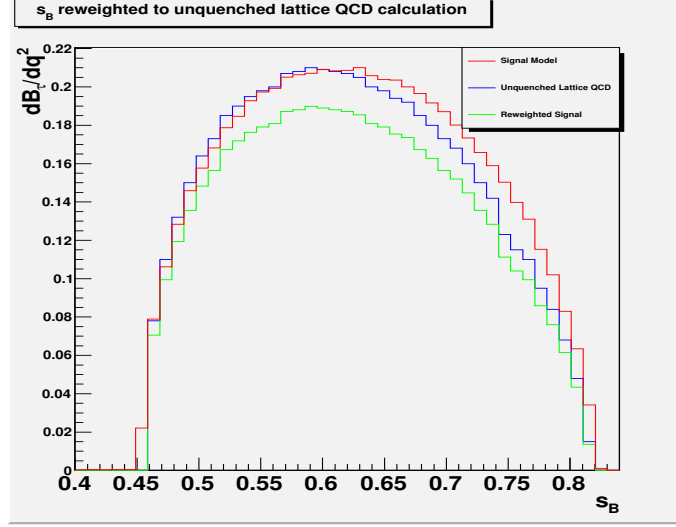


Figure 8.2: Reweighted generic signal MC distribution to the unquenched lattice QCD result. The reweighted signal MC is scaled down by $\sim 10\%$ to distinguish the different curves.

8.2 Final ϵ_{sig} and N_{bkg}

The final signal efficiencies are quoted in Table 8.1, along with the final background estimate and including statistical and systematic uncertainties (discussed in section 9). As can be readily seen, the signal efficiency for the $e^+\mu^-$ mode is almost twice that of the e^+e^- or $\mu^+\mu^-$ mode, since it includes two different final state topologies $e^+\mu^-$ and $e^-\mu^+$. On the other hand, the background estimate for the $e^+\mu^-$ mode is only slightly large than the e^+e^- and $\mu^+\mu^-$ modes. This is a result of optimizing the MLP cut to the best upper limit separately for each mode.

Mode	Efficiency ($\times 10^{-5}$)	N_{bkg}	$N_{B^+B^-}$	N_{pkgBkg}	$N_{sideData}$	$N_{nonPkgBkg}$
e^+e^-	1.1 ± 0.2	49.4 ± 3.8	47.1 ± 2.3	43 ± 3.6	30 ± 5.5	6.4 ± 1.2
$\mu^+\mu^-$	1.3 ± 0.2	45.8 ± 4.0	40.1 ± 2.2	36.7 ± 3.7	43.0 ± 6.6	9.2 ± 1.6
$e^+\mu^-$	2.1 ± 0.3	59.2 ± 4.4	55.2 ± 2.5	50.4 ± 4.2	41 ± 6.4	8.8 ± 1.5

Table 8.1: The final values for $B \rightarrow K^+ \tau^+ \tau^-$ with statistical and systematic uncertainties.

8.3 Branching fraction results

The final signal efficiencies and background estimates, listed in Table 8.1, can be inserted into equation (8.1), with the number of observed events, to determine the central value of $\mathcal{B}(B^+ \rightarrow K^+ \tau^+ \tau^-)$. A range of values of N_{obs} is used to estimate the final branching fraction for each of the e^+e^- , $\mu^+\mu^-$ and $e^+ \mu^-$ modes in Table 8.2, 8.3, and 8.4 respectively. In this analysis, the goal is to combine all 3 signal modes and calculate one final central branching fraction for $B^+ \rightarrow K^+ \tau^+ \tau^-$. The central branching fraction, as well as the 90% confidence level upper and lower limits, is calculated using two different frequentist methods: Barlow [93] and Feldmann-Cousins [94]. The central value of the branching fraction, R , for a single mode can be written as:

$$R = \frac{n - b}{S}, \quad (8.2)$$

where here n is the number of observed events, b is the number of background events, and S is the sensitivity which is given by $\epsilon_{sig} \times N_{B\bar{B}}$.

The central value of the branching fraction for the combination of modes is determined using a maximum likelihood approach. The likelihood is defined as the product of the Poisson distributions of the different signal channels, given by

$$\mathcal{L} = \prod_i \frac{\mu_i^{n_i} e^{-\mu_i}}{n_i!}, \quad (8.3)$$

where here μ_i is the Poisson mean of the number of observed events for each mode calculated using an initial estimate of R : $\mu_i = R \times S_i + b_i$, and n_i is the actual number of observed events from data. The likelihood is thus the product of the probabilities of observing n_i , assuming a Poisson distribution with mean μ_i , where i is a sum over all modes. Maximizing, $\frac{\partial \ln \mathcal{L}}{\partial R} = 0$, gives the following:

$$\sum_i \left(\frac{n_i S_i}{R S_i + b_i} - S_i \right) = 0. \quad (8.4)$$

The resulting central value of the branching fraction for the combined modes, R_{data} , is then the value that maximizes \mathcal{L} given the actual number of observed events n_i .

An upper (lower) limit at the 90% confidence level for a given number of observed events n is the value of the branching fraction for which 10% of all measurements would yield a result which is less (greater) than or equal to n . With the Barlow method, the approach is to perform a set of “toy” Monte Carlo experiments, where the trial values of S and b are generated by Gaussian distributions with a standard deviation equal to σ_S and σ_b [93]. (Here, the uncertainty of $N_{B\bar{B}}$ is not included because it is already accounted for in the B_{tag} yield systematic error [77]). Furthermore, in each trial, the total number of observed events is generated from a Poisson distribution with a mean μ , where μ is calculated using the generated values of S and b and the central value of the branching fraction R_{data} . The trial value of R that gives 10% of toy experiments with a value of n less (greater) than or equal to the observed number of events is the upper (lower) limit [93]. Furthermore, for the combined limit, the same procedure is followed, separately for each mode, to generate the trial values of S_i , b_i and the total number of observed events. The combined upper (lower) limit is the value of R that maximizes the likelihood function, \mathcal{L} , and for which 10% of toy experiments give a value of R that is less (greater) than or equal to R_{data} [93].

In addition, the error on the central value of the branching fraction is determined as the 1σ interval, containing 68% of the distribution, calculated also using the Barlow method. Thus, the upper (lower) bound error on the branching fraction is the trial value of the branching fraction that gives 16% of toy experiments with a value of n less (greater) than the observed number of events.

Using the procedure described above, the central value of the branching fraction is independent of the systematic and statistical uncertainties of both the signal efficiency and the background estimate. However, the uncertainties on the signal efficiency and background estimate are important when determining the error on the central value of the branching fraction, as well as the upper and lower limits. The limits are determined using MC

experiments, where the signal efficiency and background estimate are generated according to Gaussian distributions with σ equal to their uncertainties (systematic and statistical). To avoid double-counting the systematic errors when calculating the combined limit, the common systematic errors between the different modes are only included once, for one of the three modes. This is done to ensure that, upon combining the different modes, the σ of the Gaussian distribution used to generate trial values of the signal efficiency and the background estimate takes into account the systematic errors only once.

The Barlow method can output negative upper or lower limits, which are difficult to interpret and understand. To avoid this, the Feldmann-Cousins method can also be used to calculate the limits of $\mathcal{B}(B^+ \rightarrow K^+ \tau^+ \tau^-)$. Here, a likelihood ratio is employed to rank the possible outcomes of an experiment when determining the boundaries of the acceptance region. This ratio is given by

$$R \equiv P(n|\mathcal{B})/P(n|\mathcal{B}_{best}), \quad (8.5)$$

where P is the probability for observing n events given a branching fraction \mathcal{B} and \mathcal{B}_{best} is the value of \mathcal{B} that maximizes $P(n|\mathcal{B})$ [94]. Thus, R is a ratio of two likelihoods: the first is the likelihood of getting n events given a branching fraction \mathcal{B} and the second is the likelihood of getting n events given the best physically possible branching fraction, \mathcal{B}_{best} . Every possible outcome (number of events) n is assigned a likelihood ratio. The limits are calculated by adding n values to the acceptance region until the sum of the probabilities is 90%. At the end of an experiment, the upper and lower limits for an observed number of events n_o are the maximum and minimum values of \mathcal{B} that have n_o in the acceptance region. To account for the uncertainties, a convolution of the Poisson distribution of P with a Gaussian in b and S is performed. The procedure is similar to that used in the Barlow method and is explained in greater detail in Ref. [95]. Furthermore, to calculate a combined limit for all 3 modes, a product of the Poisson probabilities for each signal mode is used as P and the same logic is applied. Fig. 8.3 shows the central value and the 90% confidence limit interval of $\mathcal{B}(B^+ \rightarrow K^+ \tau^+ \tau^-)$ for each of the leptonic modes as a function of N_{obs} .

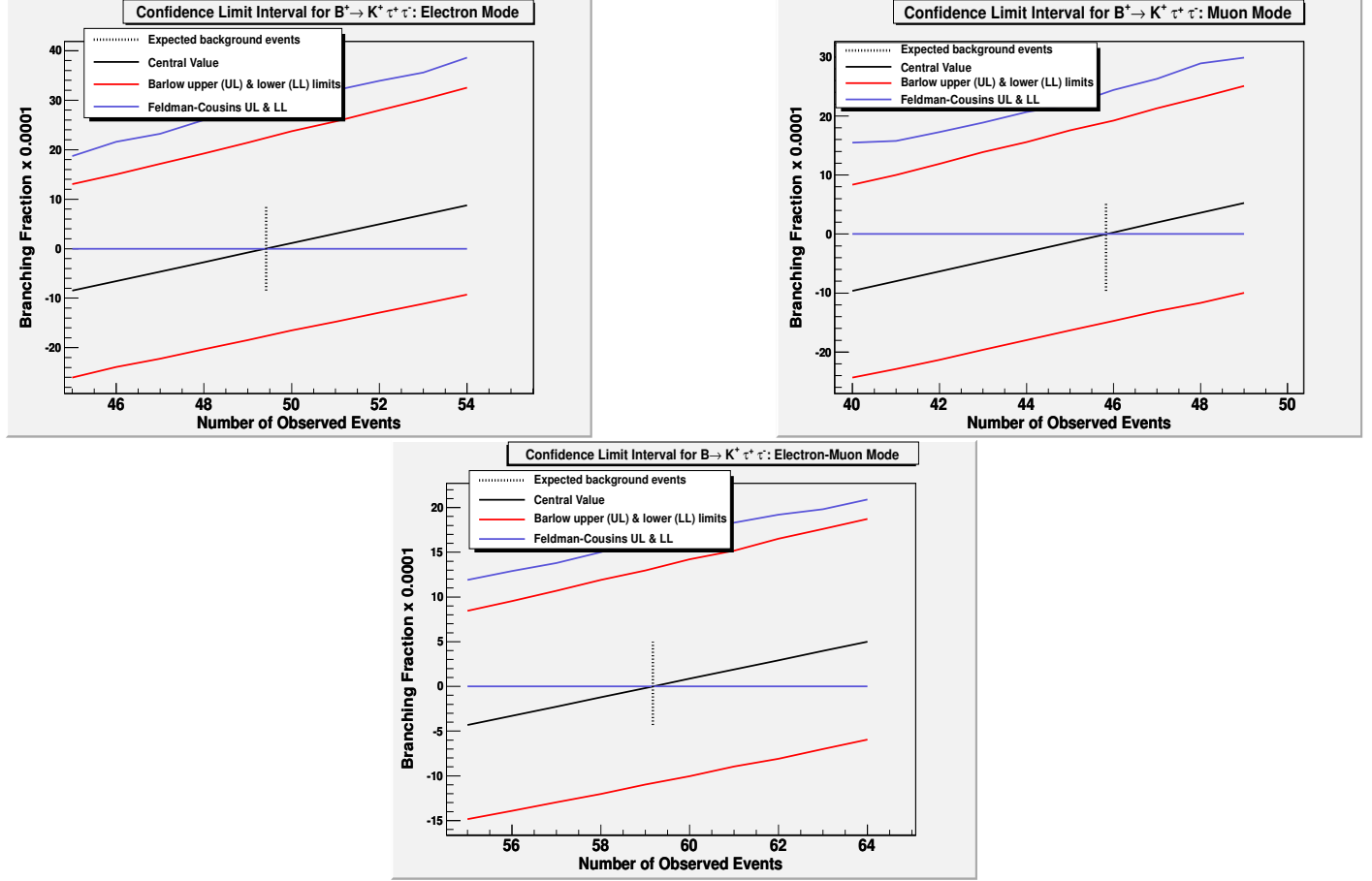


Figure 8.3: Branching fraction as a function of N_{obs} for each of the electron (top-left), muon (top-right), and electron-muon (bottom) modes.

$N_{observed}$	$\mathcal{B}(B^+ \rightarrow K^+ \tau^+ \tau^-, \tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau, \tau^- \rightarrow e^- \bar{\nu}_e \nu_\tau)$				
	Central Value	Barlow		Feldmann-Cousins	
	$(\times 10^{-4})$	Lower Limit $(\times 10^{-4})$	Upper Limit $(\times 10^{-4})$	Lower Limit $(\times 10^{-4})$	Upper Limit $(\times 10^{-4})$
45	-8.46	-25.31	12.25	0.00	19.00
46	-6.55	-23.22	14.27	0.00	20.20
47	-4.63	-21.42	16.39	0.00	21.80
48	-2.72	-19.49	18.60	0.00	24.70
49	-0.81	-17.85	20.78	0.00	26.50
50	1.11	-15.89	23.18	0.00	29.10
51	3.02	-14.11	25.00	0.00	31.30
52	4.94	-12.28	27.49	0.00	33.00
53	6.85	-10.22	29.53	0.00	35.40
54	8.77	-8.66	31.98	0.00	38.10

Table 8.2: $\mathcal{B}(B^+ \rightarrow K^+ \tau^+ \tau^-, \tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau, \tau^- \rightarrow e^- \bar{\nu}_e \nu_\tau)$ values with upper and lower Barlow and Feldmann-Cousins limits.

$N_{observed}$	$\mathcal{B}(B^+ \rightarrow K^+ \tau^+ \tau^-, \tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau, \tau^- \rightarrow \mu^- \bar{\nu}_\mu \nu_\tau)$				
	Central Value	Barlow		Feldmann-Cousins	
	$(\times 10^{-4})$	Lower Limit $(\times 10^{-4})$	Upper Limit $(\times 10^{-4})$	Lower Limit $(\times 10^{-4})$	Upper Limit $(\times 10^{-4})$
40	-9.62	-23.75	7.58	0.00	13.10
41	-7.97	-22.04	9.31	0.00	14.80
42	-6.32	-20.45	11.23	0.00	16.00
43	-4.67	-18.85	13.18	0.00	17.50
44	-3.02	-17.24	14.82	0.00	19.80
45	-1.37	-15.59	16.96	0.00	21.20
46	0.28	-14.07	18.54	0.00	23.80
47	1.93	-12.57	20.58	0.00	25.20
48	3.58	-10.83	22.42	0.00	27.60
49	5.24	-9.26	24.51	0.00	29.40

Table 8.3: $\mathcal{B}(B^+ \rightarrow K^+ \tau^+ \tau^-, \tau^+ \rightarrow \mu^+ \nu_\mu \bar{\nu}_\tau, \tau^- \rightarrow \mu^- \bar{\nu}_\mu \nu_\tau)$ values with upper and lower Barlow and Feldmann-Cousins limits.

$N_{observed}$	$\mathcal{B}(B^+ \rightarrow K^+ \tau^+ \tau^-, \tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau, \tau^- \rightarrow \mu^- \bar{\nu}_\mu \nu_\tau)$				
	Central Value	Barlow		Feldmann-Cousins	
	($\times 10^{-4}$)	Lower Limit ($\times 10^{-4}$)	Upper Limit ($\times 10^{-4}$)	Lower Limit ($\times 10^{-4}$)	Upper Limit ($\times 10^{-4}$)
55	-4.32	-14.37	7.97	0.00	11.70
56	-3.29	-13.30	9.00	0.00	12.60
57	-2.25	-12.48	10.23	0.00	13.00
58	-1.22	-11.39	11.47	0.00	14.50
59	-0.18	-10.37	12.48	0.00	15.40
60	0.86	-9.41	13.82	0.00	16.70
61	1.89	-8.56	14.75	0.00	17.90
62	2.93	-7.35	16.08	0.00	18.80
63	3.97	-6.47	17.09	0.00	19.80
64	5.00	-5.47	18.36	0.00	20.90

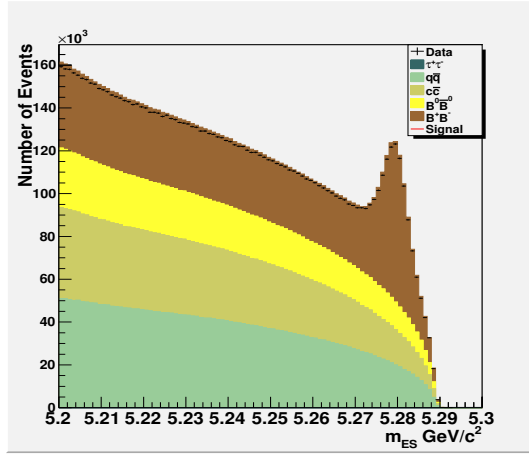
Table 8.4: $\mathcal{B}(B^+ \rightarrow K^+ \tau^+ \tau^-, \tau^+ \rightarrow e^+ \nu_e \bar{\nu}_\tau, \tau^- \rightarrow \mu^- \bar{\nu}_\mu \nu_\tau)$ values with upper and lower Barlow and Feldmann-Cousins limits.

8.4 Branching fraction closure test

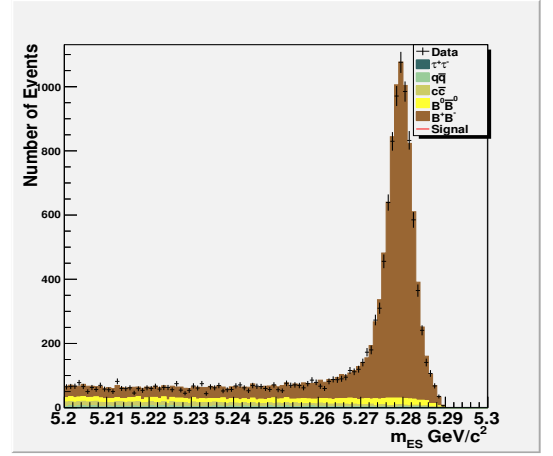
To check the validity of the procedure used to calculate the branching fraction, as given in equation (8.1), a closure test is done by injecting signal MC in to a *mock* data sample. The *mock* data sample is created from background MC. This is done by randomly selecting events from each background MC type: B^+B^- , $B^0\bar{B}^0$, $c\bar{c}$, uds and $\tau^+\tau^-$, while matching the data luminosity for each run. Table 4.2 shows the weight for each MC type for each run. This weight is approximately 0.1 for B^+B^- , $B^0\bar{B}^0$, and $c\bar{c}$ events reflecting the 10:1 ratio of the number of generated events for these MC types when compared to data. The uds and $\tau^+\tau^-$ MC samples have a weight of approximately 0.25, displaying a MC sample that is almost 4 times the data. Thus, for each of the B^+B^- , $B^0\bar{B}^0$, and $c\bar{c}$ MC, approximately 1 event should be randomly chosen from every set of 10 events, whereas for uds and $\tau^+\tau^-$ 1 out of every 4 events should be selected to form the *mock* data sample. To do this accurately, a reproducible random variable is used. The result is a sample of background MC events that has a luminosity equivalent to that of the data, which can be used to test the branching fraction calculation if signal is appended to it.

The signal selection is then applied on the *mock* data sample. The resulting m_{ES} distributions of the sample are shown in Fig. 8.4, after the E_{miss} , lepton PID, and s_B cuts. As can be readily seen, the agreement between the *mock* data and the background MC is almost exact, within statistical error. This is expected since the *mock* data sample is a subset of background MC with the same luminosity as data. Fig. 8.5 shows a plot of R_{comb} and C_{yield} as a function of signal selection, using the *mock* data sample. As expected, the R_{comb} plots are unchanged. The C_{yield} plot reflects a C_{yield} consistent with one.

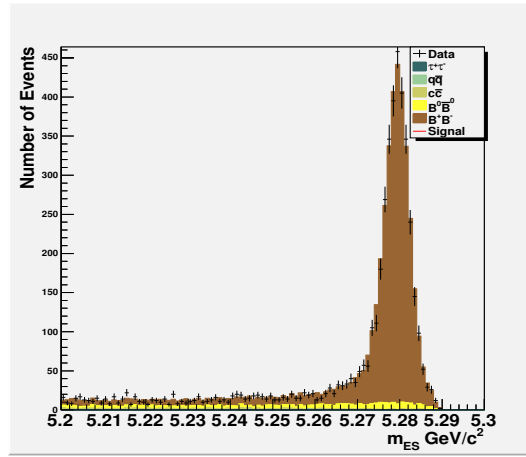
After validating the *mock* data sample, the generic signal MC sample is appended to it such that the branching fraction extraction used for this analysis can be checked. The generic signal MC sample contains 3.308×10^6 $B^+ \rightarrow K^+ \tau^+\tau^-$ events, which results in a



(a) After E_{miss} cut.



(b) After lepton PID cuts.



(c) After s_B cut.

Figure 8.4: m_{ES} distribution of *mock* data sample. The *mock* data sample (black points) is overlaid on the background MC distributions (color-filled). $q\bar{q}$ refers to $q = u, d, s$.

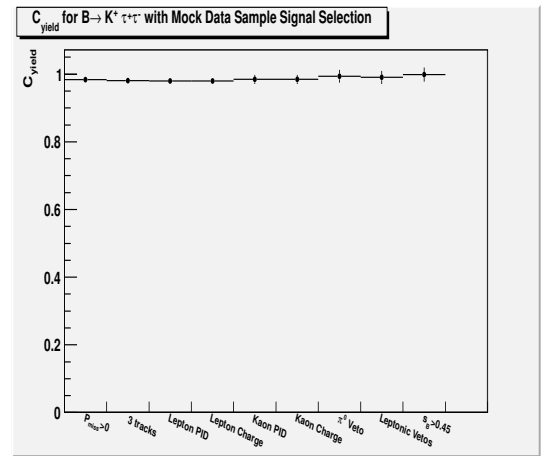
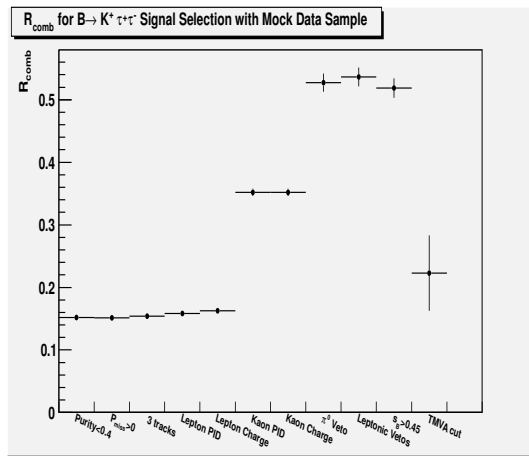


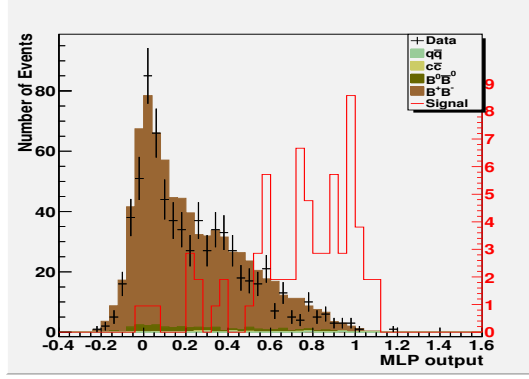
Figure 8.5: R_{comb} (left) and C_{yield} (right) distributions using the *mock* data sample.

branching fraction of 7.1×10^{-3} . This result is given by

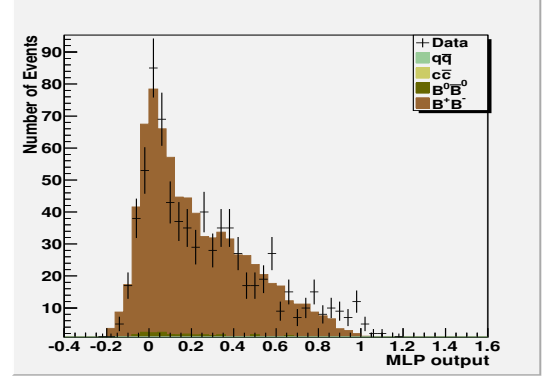
$$\mathcal{B} = \frac{N_{sigMC}^{gen}}{N_{B^\pm}} = \frac{N_{sigMC}^{gen}}{\mathcal{L} \times \sigma_{B^+B^-} \times 2} = 7.1 \times 10^{-3}, \quad (8.6)$$

where N_{sigMC}^{gen} is the number of generated events in the $B^+ \rightarrow K^+ \tau^+ \tau^-$ generic signal MC and $\sigma_{B^+B^-}$ is the B^+B^- cross-section, given approximately 0.55 nb. The factor of two is inserted to account for the total number of B^+ and B^- candidates in the B^+B^- MC. (The branching fraction is the number of events B mesons decaying via $B^+ \rightarrow K^+ \tau^+ \tau^-$ divided by the total number of B mesons, not the total number of B^+B^- pairs).

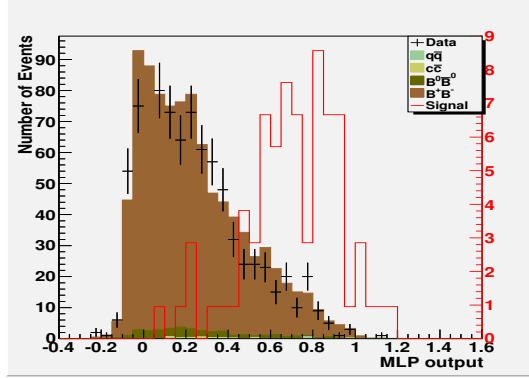
Signal selection cuts are then applied to the *mock* data sample with the added signal MC, and the sample is then run through the MLP neural network. The result is shown in Fig. 8.6 for each of the electron, muon, and electron-muon modes. The neural network outputs of the *mock* data sample are also shown for each mode before adding the signal MC. The signal efficiency and background estimate are then calculated as described in the previous sections. For the background estimate, R_{comb} is unchanged. However, the value of C_{yield} in this sample is almost unity, 0.998 ± 0.020 . The branching fraction is then calculated using equation 8.1, and the results are shown in Table 8.5. As can be seen, the resulting branching fraction for each mode agrees with the expected branching fraction, within statistical errors. Statistical errors are calculated here as the difference between the central value of the branching fraction and the 1σ upper and lower limits. The statistical errors are large for the individual modes but are improved for the combined limit. The combined limit also shows agreement with the expected branching fraction. The final result that will be quoted in this analysis is a combined limit and this closure test shows that the methodology used here will reflect a signal if it is existent.



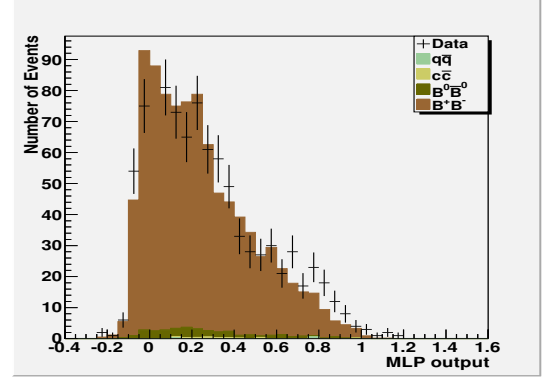
(a) *Mock* data sample and overlaid signal for muon mode.



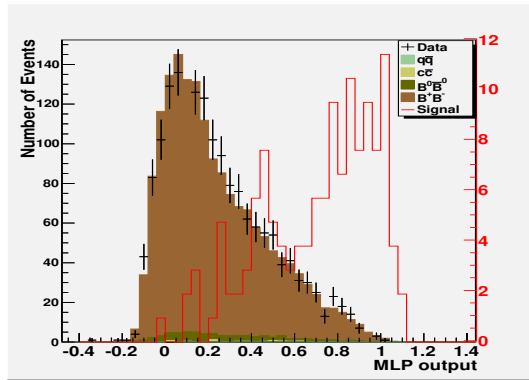
(b) *Mock* data sample and appended signal for muon mode.



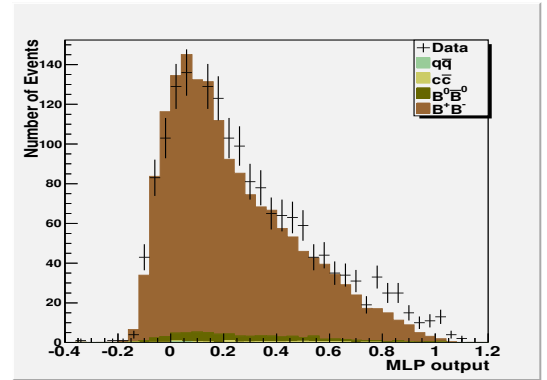
(c) *Mock* data sample and overlaid signal for electron mode.



(d) *Mock* data sample and appended signal for electron mode.



(e) *Mock* data sample for electron-muon mode and overlaid signal.



(f) *Mock* data sample and appended signal for electron-muon mode.

Figure 8.6: MLP output of *mock* data sample. Signal MC (*generic*) distribution is shown in red (left) and is appended to the *mock* data sample (right). The *mock* data sample (black points) is overlaid on the background MC distributions (color-filled). $q\bar{q}$ refers to $q = u, d, s$.

Mode	Calculated \mathcal{B} ($\times 10^{-3}$)	Expected \mathcal{B} ($\times 10^{-3}$)
Electron	$6.10^{+2.34}_{-1.68}$	7.10
Muon	$7.20^{+2.19}_{-1.60}$	7.10
Electron-Muon	$6.76^{+1.60}_{-1.23}$	7.10
Combined	$6.75^{+1.04}_{-0.90}$	7.10

Table 8.5: Result of branching fraction closure test with *mock* data + signal sample.

Chapter 9

Systematic Uncertainties

The signal selection for $B^+ \rightarrow K^+ \tau^+ \tau^-$ events introduces systematic uncertainties on the final signal efficiency and background estimate. As discussed in section 7.1, the combinatorial background is estimated using sideband data. The peaking background and the signal efficiency are calculated using background and signal MC. Therefore, the level of agreement between data and MC is crucial in this analysis. The m_{ES} sideband substitution is used to account for the well-known discrepancies between data and MC. Because it is done at a later stage in the selection, many of the uncertainties associated with data/MC disagreement have already been accounted for in the uncertainty of the sideband substitution method. However, any residual data/MC disagreement after applying the m_{ES} sideband substitution must be addressed. Furthermore, other sources of systematic uncertainties, such as the choice of the theoretical model for the signal MC, are also examined. To do so, it is necessary to investigate each cut in the signal selection and evaluate the level of systematic uncertainty it introduces. The main sources of systematic uncertainties are found to be the following:

- B_{tag} yield correction: systematic uncertainty on the combinatorial and peaking background estimate, as well as the signal efficiency.
- Assumed theoretical model: systematic uncertainty on the signal efficiency.
- Data/MC agreement in PID: systematic uncertainty on the signal efficiency, peaking

and combinatorial background.

- Data/MC agreement in π^0 veto: systematic uncertainty on the signal efficiency, peaking and combinatorial background.
- Data/MC agreement in TMVA approach: systematic uncertainty on the signal efficiency, peaking and combinatorial background.

Except for the systematic uncertainty due to the assumed theoretical model, the systematic uncertainties contribute to both the signal efficiency and background estimate. Given equation 8.1, this correlation must be taken into account when evaluating the final branching fraction. However, as shown in Appendix F, correlations due to common systematic uncertainties between the signal efficiency and the background estimate turn out to have a negligible effect on the final branching fraction.

9.1 B_{tag} yield correction

The final signal efficiency and background estimate are dependent on what stage of the signal selection these values are chosen. In addition, the measurements of C_{yield} and R_{comb} are also highly correlated, since the total number of events in the m_{ES} signal region is constant for any given set of cuts. Thus, a larger R_{comb} implies a larger combinatorial contribution and thus more events will be subtracted from the peaking background estimate (and vice versa). This is why only one systematic uncertainty needs to be evaluated, taking into account the anti-correlation between C_{yield} and R_{comb} . C_{yield} and R_{comb} are related via the equation

$$C_{\text{yield}} = \frac{N_{\text{Data}}^{\text{pkg}}}{N_{B^+B^-}^{\text{pkg}}} = \frac{N_{\text{data}}^{\text{sig}} - (N_{\text{data}}^{\text{side}} \times R_{\text{comb}})}{N_{B^+B^-}^{\text{pkg}}}. \quad (9.1)$$

To evaluate the systematic uncertainty, a set of toy MC experiments is performed to determine the effect of fluctuating R_{comb} , and thus C_{yield} , on the signal efficiency and background estimate. In these toy MC experiments, the value R_{comb} is fluctuated, using 5000 trials,

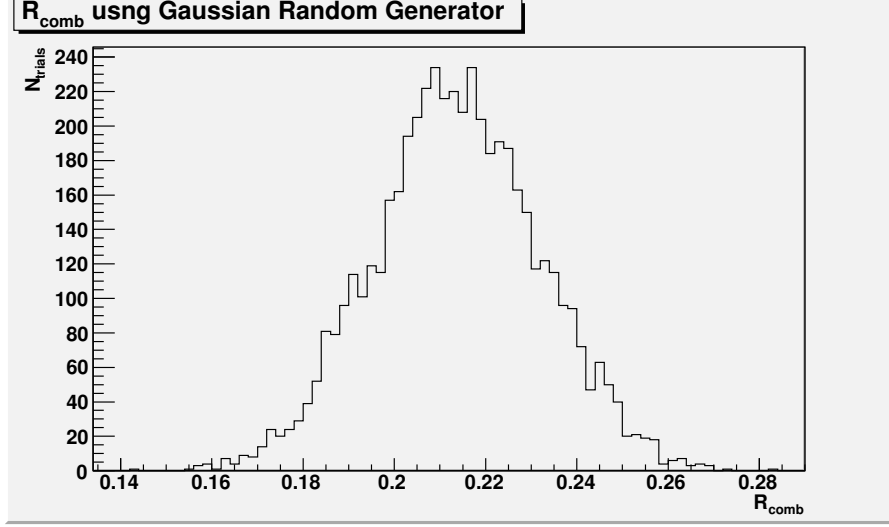


Figure 9.1: Gaussian distribution resulting from the random fluctuation of R_{comb} with a Gaussian generator of $\mu = 0.21$ and $\sigma = 0.018$.

according to a Gaussian distribution with a mean of the chosen R_{comb} (0.21) and a σ equal to the statistical error, as shown in Fig. 9.1. Because R_{comb} and C_{yield} are anti-correlated, fluctuating R_{comb} will cause a resulting variation in the value of C_{yield} . To quantify this variation, the *mock* data sample is used. In contrast to actual data, with the *mock* data sample, the number of signal events, N_{data}^{sig} , can be extracted at the final stage in the signal selection, after the MLP cut. Furthermore, given equation 9.1, the values of both N_{data}^{side} and $N_{pkg}^{B^+B^-}$ can be readily determined from the *mock* data sample and B^+B^- MC, respectively. Thus, the exact anti-correlation between C_{yield} and R_{comb} can be determined at the final stage in the analysis, after applying the MLP cut. This is important since R_{comb} is determined after the MLP neural network cut, and C_{yield} is assumed to be consistent throughout the cut flow, even though it cannot be extracted after the final MLP cut. Thus, for each value of R_{comb} , the anti-correlated value of C_{yield} can be determined with equation 9.1 and is shown in Fig. 9.2. The anti-correlation is shown explicitly in Fig. 9.3.

Fluctuating R_{comb} and C_{yield} will affect both the signal efficiency and the background estimate. To examine the effect on the signal efficiency, the anti-correlated values of C_{yield} , shown in Fig. 9.2, are used to determine the final signal efficiency in each toy MC experiment.

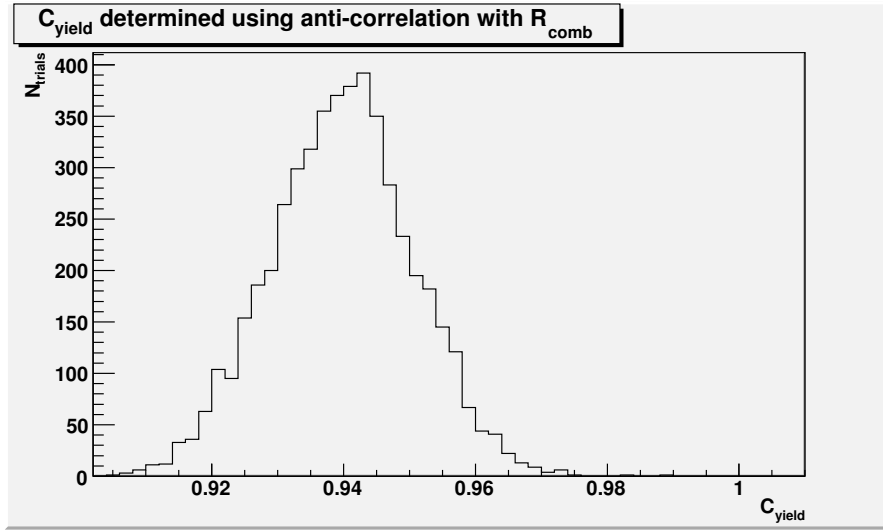


Figure 9.2: Resulting C_{yield} determined using the anti-correlation with R_{comb} .

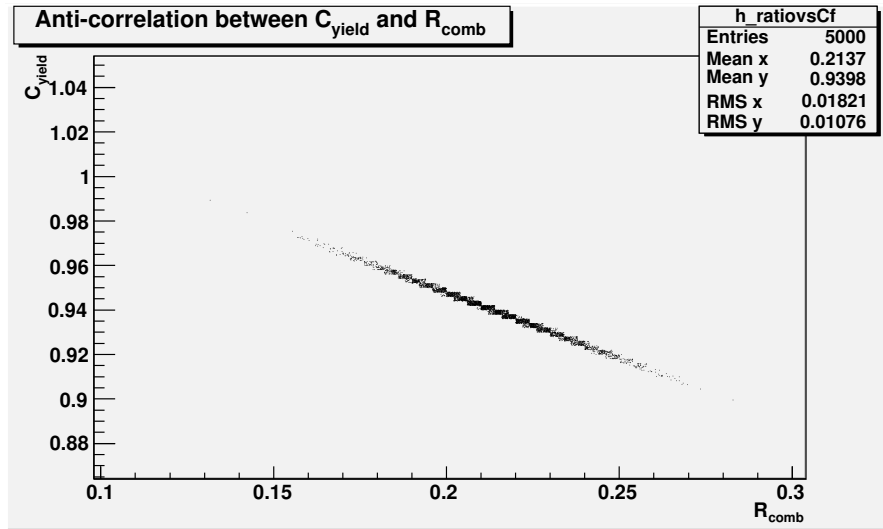


Figure 9.3: 2D plot of R_{comb} and C_{yield} displaying a linear anti-correlation.

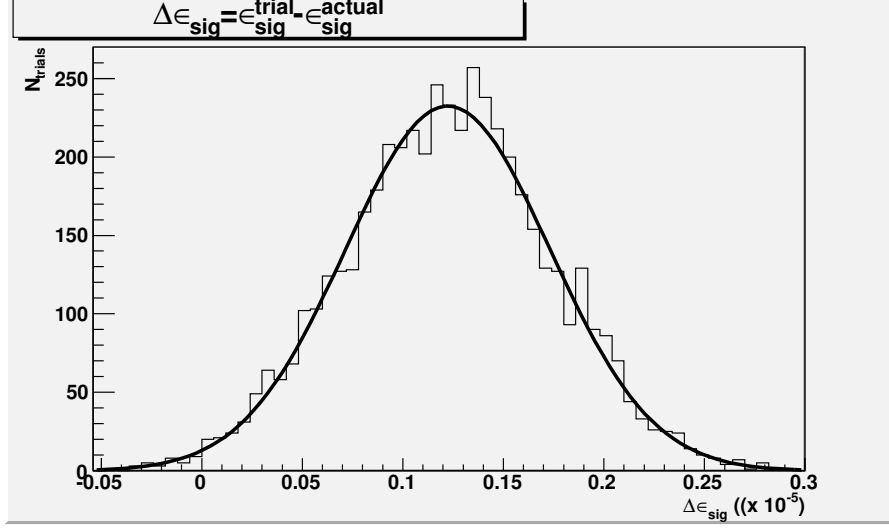


Figure 9.4: Fit to $\Delta\epsilon_{sig}$ after fluctuating R_{comb} with a Gaussian random generator and determining the anti-correlated C_{yield} .

The resulting ϵ_{sig} distribution is determined, and the difference $\Delta\epsilon_{sig}$ is calculated as the difference: $\epsilon_{sig}^{trial} - \epsilon_{sig}^{actual}$. Here, ϵ_{sig}^{trial} is the value of the signal efficiency determined using the C_{yield} resulting from each toy MC trial and ϵ_{sig}^{actual} is the actual signal efficiency used in the branching fraction calculation. A plot of $\Delta\epsilon_{sig}$ is shown in Fig. 9.4. Because the value of C_{yield} is mostly below 1, the value of ϵ_{sig}^{trial} is almost always greater than ϵ_{sig}^{actual} and thus the distribution of $\Delta\epsilon_{sig}$ is offset from zero. The systematic uncertainty due to the B_{tag} yield is determined as the σ of the Gaussian fit to the $\Delta\epsilon_{sig}$ distribution, after fluctuating R_{comb} using a Gaussian random number generator and determining the anti-correlated C_{yield} using equation 9.1 and the *mock* data sample. σ is found to be 5.1×10^{-7} , which translates into a systematic uncertainty of 1.2%.

The methodology used for the signal efficiency cannot be applied on the background estimate. This is true since the final background estimate will be a constant number for each of these toy MC experiments. To understand this, recall that, for these toy MC experiments, the value of C_{yield} is determined with a constant N_{data}^{sig} from the *mock* data sample at the final stage in the signal selection. Doing so, N_{bkg} is being scaled to match the data exactly. The only thing varying between each trial is the ratio of combinatorial background to peaking

background.

However, because the main source of background in this analysis is peaking background, the systematic uncertainty due to the B_{tag} yield correction can be determined by fluctuating C_{yield} only. At the end of the selection, there is a very small fraction of combinatorial background and thus fluctuations in R_{comb} do not contribute to the overall uncertainty. The source of uncertainty on the background estimate is thus related to how accurately the overestimate of B^+B^- MC to the data is determined. The plot of C_{yield} as a function of cut in Fig. 7.11, shows that, except for the initial skim cuts and continuum likelihood, the value of C_{yield} is relatively uniform and close to 0.90. The statistical error on C_{yield} increases along the signal selection because the number of surviving events decreases. After the s_B cut, the statistical error of C_{yield} is largest and covers a range of values that almost spans the C_{yield} spectrum as a function of cut. Therefore, to estimate the systematic uncertainty, the value of C_{yield} is fluctuated by plus or minus this statistical error. The difference in the final background estimate is calculated according to:

$$\begin{aligned}\delta_+ &= \frac{N_+ - N}{N} \\ \delta_- &= \frac{N_- - N}{N} \\ \delta_{\text{total}} &= \frac{|\delta_+| + |\delta_-|}{2}\end{aligned}\tag{9.2}$$

where here N is the number of background events without any fluctuation and N_+ (N_-) is the number of background events when C_{yield} is fluctuated by plus (minus) its statistical uncertainty. The values of N , N_{\pm} , and δ_{\pm} are shown in Table 9.1. Fluctuating C_{yield} by its statistical error results in $\sim 2\%$ systematic uncertainty on the background estimate.

The systematic uncertainty on the B_{tag} yield simultaneously accounts for any data-MC discrepancy due to the B_{tag} cuts: continuum likelihood cut (discussed in section 9.2) and purity cut. These are applied at an early stage in the signal selection, where the main sources of data-MC discrepancy are corrected for by applying the m_{ES} sideband substitution at a later

N	154.5
N_+	157.7
N_-	151.5
δ_+	2.1%
δ_-	1.94%

Table 9.1: Values of N , N_{\pm} , and δ_{\pm} used to calculate the B_{tag} yield systematic uncertainty on the background estimate.

stage. As previously mentioned, the purity cut is applied to get rid of “dirty” hadronic modes that result in poorly reconstructed B_{tag} candidates. This analysis is not sensitive to the specific combination of hadronic modes used in the hadronic B_{tag} reconstruction. Furthermore, B_{tag} candidates that are mis-reconstructed via “dirty” modes form the combinatorial component of the B^+B^- MC. With the m_{ES} sideband substitution, this component is estimated directly from sideband data, which removes the need for a systematic uncertainty to account for any data-MC discrepancy in the “dirty” B_{tag} modes. Any residual discrepancy in the purity distribution of properly reconstructed events is accounted for with the systematic on the B_{tag} yield.

9.2 Continuum Likelihood Cut

As previously mentioned in section 5.1.3, the continuum likelihood cut is used to get rid of non- $B\bar{B}$ events or events where a B_{tag} candidate is poorly reconstructed. The discrepancy observed prior to the continuum likelihood cut is mainly the result of unmodeled background sources, such as radiative Bhabha events, which appear in the data and not the MC. Consequently, it is difficult to extract a meaningful measure of data-MC agreement at such an early stage in the analysis.

However, with the m_{ES} sideband substitution, continuum backgrounds are estimated directly from data. This eliminates the need to evaluate data-MC discrepancies due to these unmodeled backgrounds. Furthermore, other sources of discrepancy in the continuum

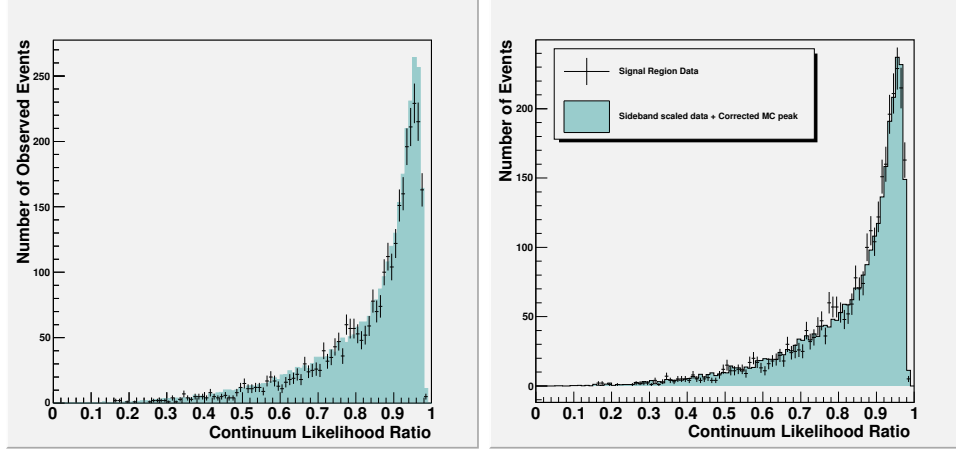


Figure 9.5: Continuum likelihood ratio, after applying the s_B cut and before the MLP cut, before (left) and after (right) the m_{ES} sideband substitution.

likelihood ratio distribution, such as from the B_{tag} reconstruction, are also corrected for in the B^+B^- MC, with the C_{yield} scaling. This can be seen in Fig. 9.5, which shows the continuum likelihood ratio before and after the m_{ES} sideband substitution, after applying the s_B cut but before the MLP cut. The data-MC agreement is satisfactory with a discrepancy less than 5%. This discrepancy has already been accounted for with the systematic uncertainty due to the B_{tag} yield correction, as discussed in section 9.1.

9.3 Theoretical Uncertainty

Because the final signal efficiency is dependent on the choice of theoretical model, it is important to evaluate a systematic uncertainty associated with quoting the final signal efficiency using the unquenched lattice QCD approach. To do so, the final signal efficiency is compared between 3 theoretical models: light cone sum rules (LCSR) [54], unquenched lattice QCD [57], and the lattice-constrained dispersion quark model calculation[90]. Because of its larger statistics, the cocktail signal MC, which has been generated according to the LCSR model, is used here. First, truth information is used to determine the s_B distribution of the cocktail signal MC before any cuts. This distribution, shown in Fig. 9.6, reflects the LCSR

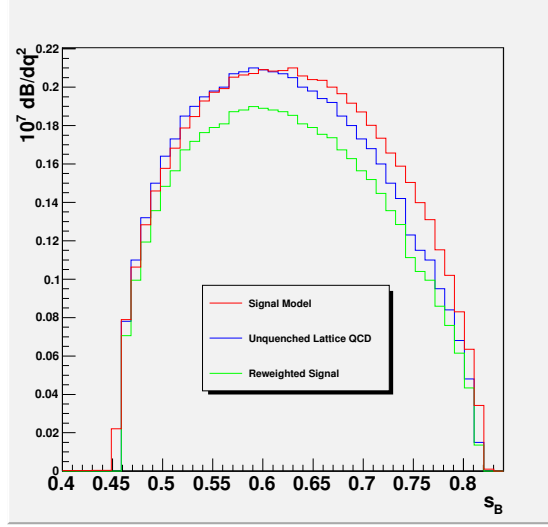


Figure 9.6: s_B distribution of cocktail signal MC after reweighting to the unquenched lattice QCD calculation. The signal MC before and after reweighting is shown in red and blue respectively.

model. The s_B distribution is then reweighted to match that of the unquenched lattice QCD model, as shown in Fig. 9.6. The reweighted sample is then run through the entire signal selection, including the MLP cut and the difference in the final signal efficiency is evaluated by comparing the final signal yields of each sample. When comparing the LCSR model to the unquenched lattice QCD, the difference in signal efficiency is found to be 2.89%. Furthermore, the s_B distribution is then reweighted according to the dispersion quark model [90] and the full signal selection is applied to the resulting sample. The s_B distribution determined using the $B \rightarrow K$ transition form factors of the dispersion quark model is shown in Fig. 9.7, along with the reweighted cocktail signal MC distribution in Fig. 9.8. The difference in signal efficiencies is also then determined between the unquenched lattice QCD and the dispersion quark model, and is found to be 0.8%. The systematic uncertainty due to the assumed theoretical model is then calculated to be the quadrature sum of the difference in signal efficiency when the LCSR and dispersion quark models are compared with the unquenched lattice QCD model, yielding a 3% systematic uncertainty.

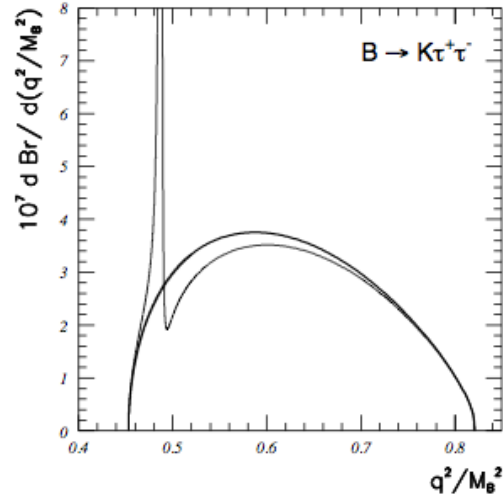


Figure 9.7: s_B distribution resulting from the dispersion quark model, with and without long-distance (peaking) contributions from the $\psi(2S)$ resonance [90].

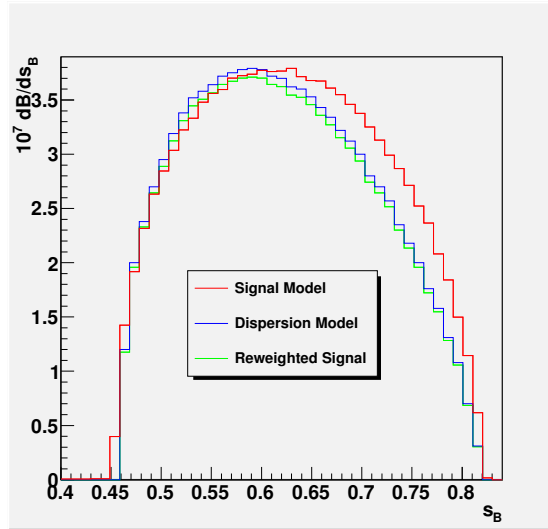


Figure 9.8: Cocktail signal s_B distribution after reweighting to the dispersion quark model. The signal MC before and after reweighting is shown in red and blue, respectively.

9.4 Particle Identification cuts

PID plays an important role in the signal selection and it is thus essential to evaluate any additional discrepancies between data and MC resulting from this cut. The choice of PID selector and tightness level has a large effect on the number and category of surviving background events. It is thus necessary to investigate the level of data-MC agreement for each individual PID selector used in this analysis, listed in Table 5.1. Because a different combination of the electron and muon selectors is used for each of the three signal modes, a lepton PID systematic uncertainty will be evaluated separately for each one.

The performance plots for each of the kaon BDT tight selector, electron KM tight selector and muon BDT loose selector are shown in Fig. 9.9-9.11 [22]. As can be readily seen, the performance of a selector depends on the momentum of the particle in consideration, both in terms of the overall efficiency and misidentification rates and in terms of the level of data-MC agreement. In this analysis, the momentum range of interest is that of the kaon and leptons in the generic signal MC. The momentum distribution of the kaon is shown in Fig. 9.12, after applying the kaon PID cuts. The electron and muon momentum distributions are shown in Figs. 9.13 - 9.16 for each of the electron, muon, and electron-muon modes. To evaluate the systematic uncertainty due to the PID cuts, the level of agreement between data and MC as a function of the particle's momentum should be determined. To do so, the ratio of the data to MC efficiency, $\epsilon_{data}/\epsilon_{MC}$, is used. A weighted average of this ratio is calculated using the momentum distribution shown in Figs. 9.13 - 9.16 for each particle type. The results are shown in Table 9.2, where a weighted average is evaluated for each charge of each particle type in a given signal channel.

The resulting kaon systematic uncertainty is applicable to all three signal channels, whereas the lepton PID uncertainty is determined, for each mode, based on the number of electrons and muons in the mode. For the kaon PID, the uncertainty due to the positive and negative tracks passing the tight BDT selector is averaged, yielding a total systematic uncertainty of 2.05% for all 3 modes. Furthermore, because there are two leptons in every

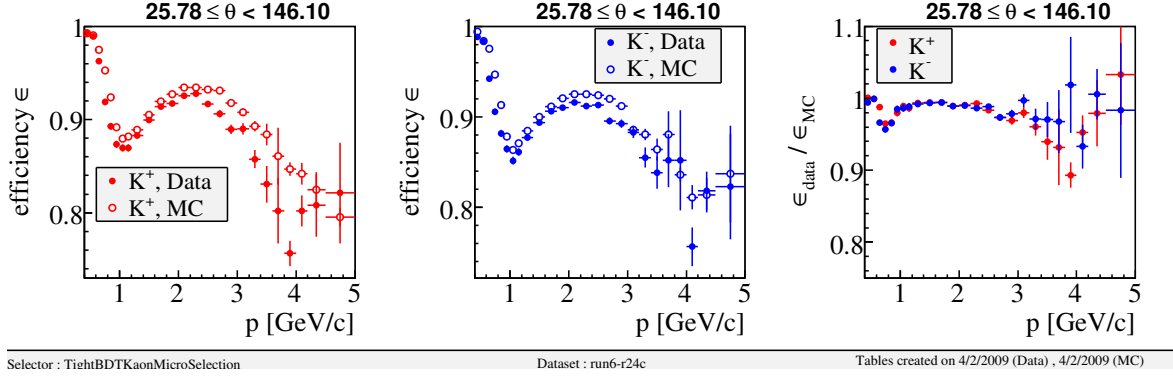


Figure 9.9: Performance plot for the kaon BDT tight selector as a function of the lab frame momentum . Data and MC efficiency for positive (left) and negative (center) kaons. The ratio of the data to MC efficiency $\epsilon_{data}/\epsilon_{MC}$ is shown in the right hand plot[22].

Mode	PID Uncertainty (%)					
	K^+	K^-	e^+	e^-	μ^+	μ^-
Electron	2.08	2.02	2.39	2.37	0.00	0.00
Muon	2.08	2.02	0.00	0.00	3.79	3.17
Electron-Muon	2.08	2.02	2.23	1.56	3.04	3.11

Table 9.2: List of PID systematic uncertainties for each charge of each particle type, calculated as a function of the particle’s momentum in a specific signal mode.

mode, the PID uncertainty associated with one track is added linearly with that for the other track within a signal mode. For the electron mode, the PID uncertainty of e^+ (2.39%) is added to that of e^- (2.37%), resulting in a total systematic uncertainty of 4.76%. The muon mode has a total PID uncertainty of 6.95%, which results from adding the uncertainty of μ^- (3.16%) with that of μ^+ (3.79%). Finally, the total PID uncertainty of the electron-muon mode is determined to be 4.97%. This is calculated by adding separately, the PID uncertainty of e^+ (2.23%) with that of μ^- (3.11%), and the PID uncertainty of e^- (1.55%) with μ^+ (3.04%). These two uncertainties are then averaged to yield the total uncertainty on the electron-muon mode.

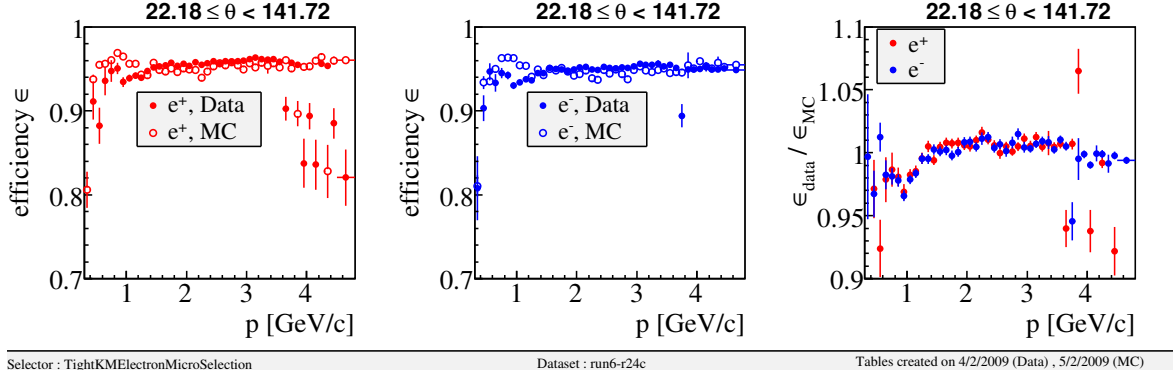


Figure 9.10: Performance plot for the electron KM tight selector as a function of the lab frame momentum. Data and MC efficiency for positive (left) and negative (center) electrons. The ratio of the data to MC efficiency $\epsilon_{\text{data}}/\epsilon_{\text{MC}}$ is shown in the right hand plot [22].

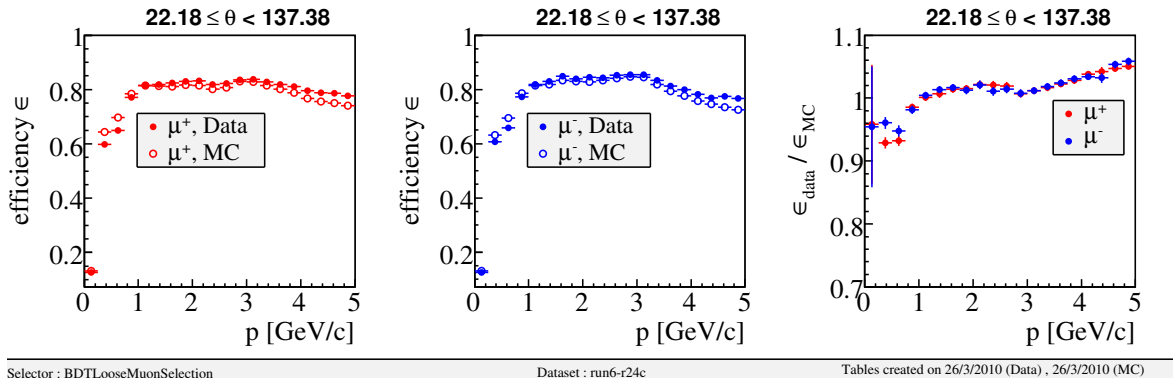


Figure 9.11: Performance plot for the muon BDT loose selector as a function of the lab frame momentum. Data and MC efficiency for positive (left) and negative (center) muons. The ratio of the data to MC efficiency $\epsilon_{\text{data}}/\epsilon_{\text{MC}}$ is shown in the right hand plot [22].

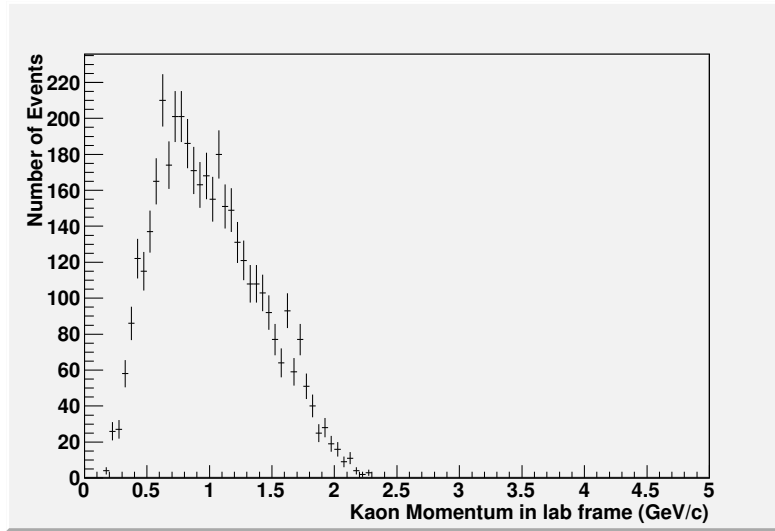


Figure 9.12: Lab frame momentum of kaon in all 3 signal modes, after applying kaon PID.

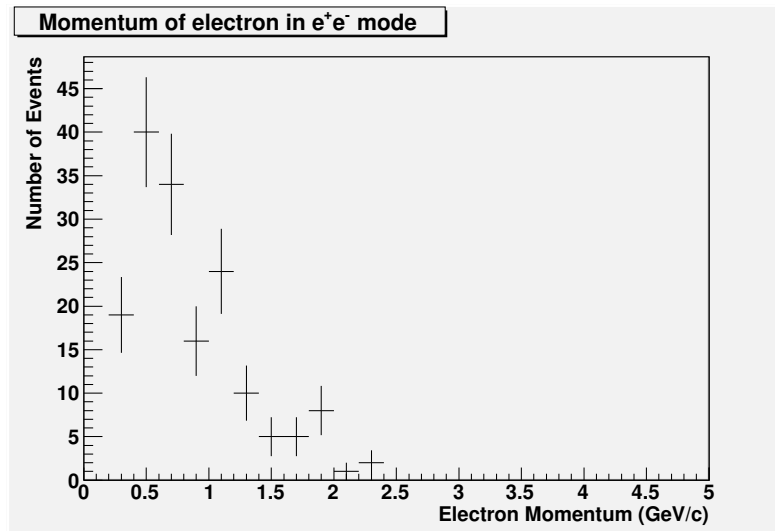


Figure 9.13: Lab frame momentum of electron in the Electron mode, after applying the electron PID cut.

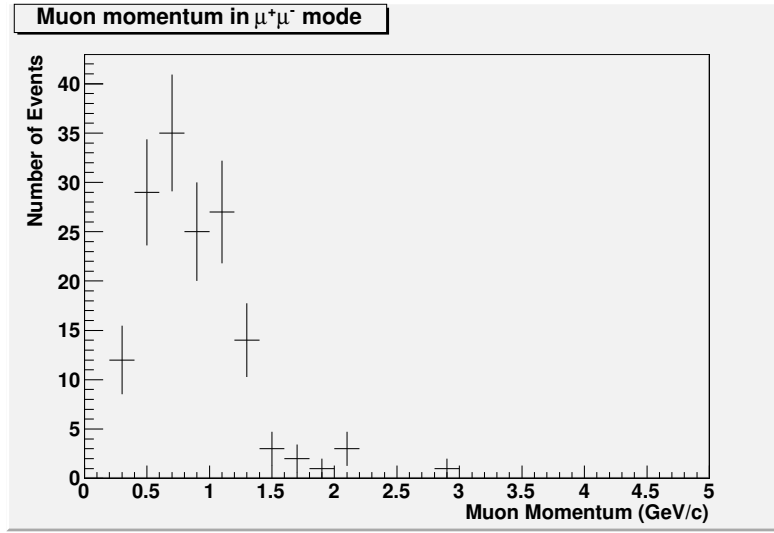


Figure 9.14: Lab frame momentum of muon in the Muon mode, after applying the muon PID cut.

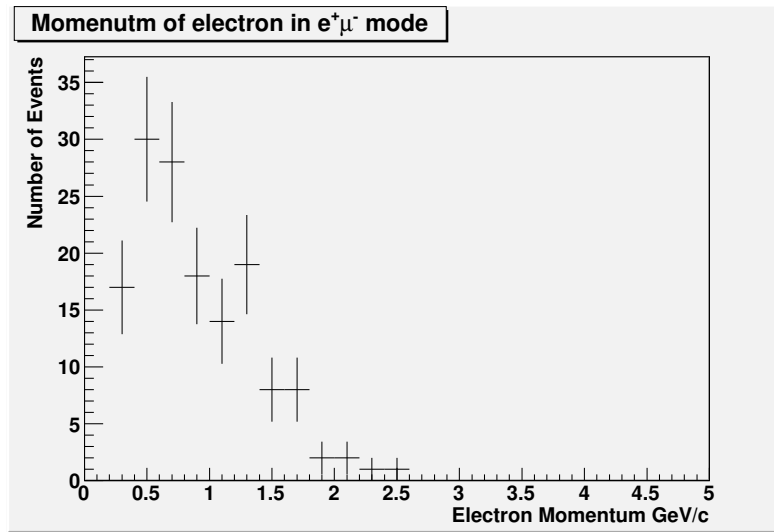


Figure 9.15: Lab frame momentum of electron in the Electron-Muon mode, after applying PID cuts.

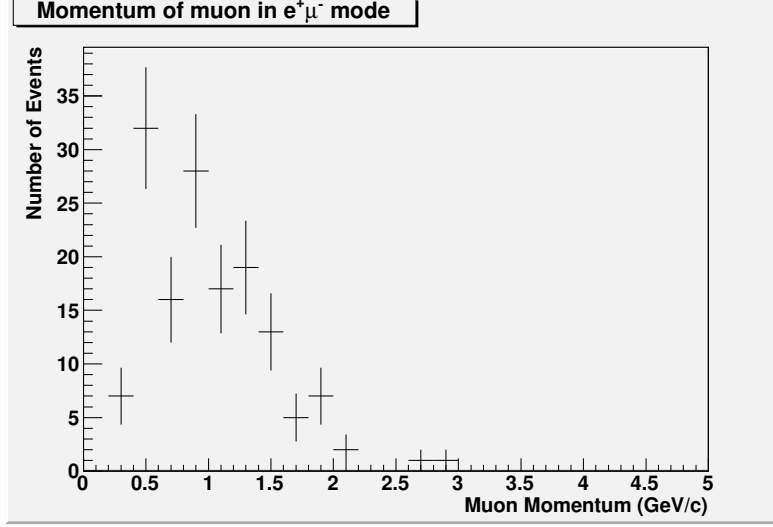


Figure 9.16: Lab frame momentum of muon in the Electron-Muon mode, after applying PID cuts.

9.5 π^0 veto

Systematic uncertainties associated with π^0 reconstruction and veto are important to investigate, even though the m_{ES} sideband substitution is applied after the π^0 veto. In this analysis, π^0 's are reconstructed in the same way as those in the *BABAR* pi0Loose list. The pi0Loose list is basically a standard recipe for π^0 reconstruction, as determined by the Neutrals Analysis Working Group [92] (AWG) of the *BABAR* collaboration. To quantify the systematic uncertainty, the level of data/MC agreement is initially evaluated. The mass window, used in the π^0 reconstruction, is fluctuated by $\pm\delta$ and the difference in the final signal efficiency and background estimate is determined.

To determine δ , the π^0 mass distribution in data and MC is first examined. Fig. 9.17 shows the MC distribution of all π^0 candidates that pass the reconstruction discussed in section 5.2.3, after all cuts that precede the π^0 veto in the signal selection. This distribution is then fit with a combination of a Crystal Ball and a linear function. The same fit is repeated for the data distribution, shown in Fig. 9.18, with all parameters fixed except for σ , the standard deviation of the Gaussian peak, and N , the total yield.

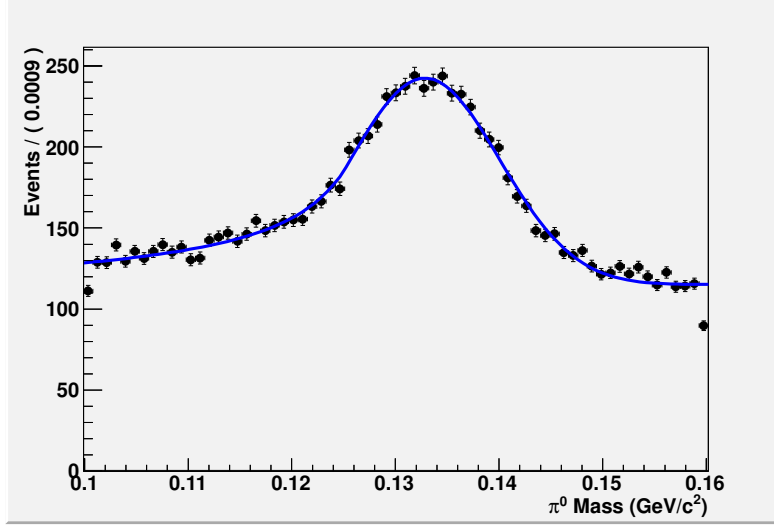


Figure 9.17: Background MC π^0 mass fit.

δ is then determined as the difference in the σ value of the data and MC π^0 mass fits. The MC fit yields a σ of 7.27056 MeV/ c^2 and that of data is 7.0997 MeV/ c^2 , which results in $\delta = 0.17086$ MeV/ c^2 . The mass window of the π^0 veto is then fluctuated by $\pm\delta$ on each side. Thus, the loose π^0 veto consists of a mass cut of $99.830 < M_{\pi^0} < 160.171$ MeV/ c^2 , while the tight π^0 veto has a mass cut of $100.171 < M_{\pi^0} < 15.983$ MeV/ c^2 . The difference in signal efficiency and background estimate is then evaluated and the results are shown in Table 9.3 for each mode. These results are consistent with zero, within the available statistics, and thus a measurable effect cannot be determined within the context of this analysis.

Nevertheless, the systematic uncertainty due to the π^0 veto is not only related to the data-MC discrepancy in the final mass distributions, but also it is associated with the efficiency of reconstructing a single photon candidate. Furthermore, hadronic interactions in the EMC and the photon background are not perfectly modeled in the MC [92]. Because this analysis reconstructs π^0 candidates in the same way as the *BABAR* pi0Loose list, the systematic uncertainty associated with the π^0 veto will be the same as that determined according to the standard *BABAR* recipe: 3% [77].

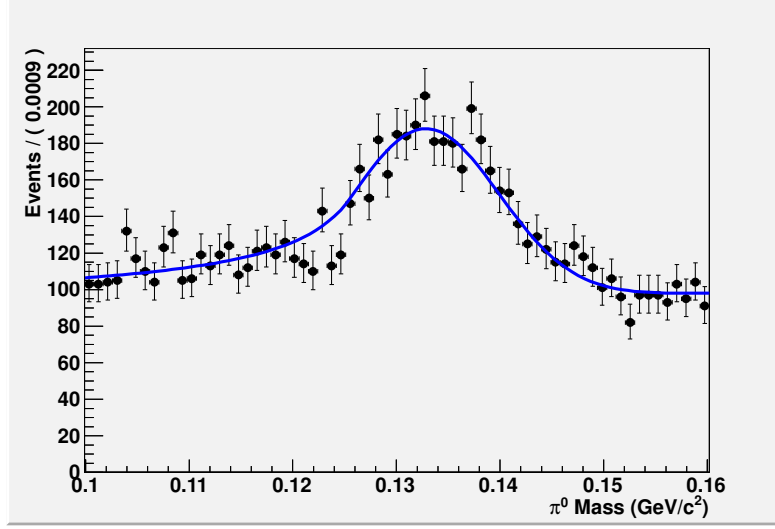


Figure 9.18: Fit of π^0 mass distribution in data.

π^0 mass window	N_{events} signal	Δ_{sig}	N_{events} bkg	Δ_{bkg}
Electron Mode				
$100 < M_{\pi^0} < 160$ MeV	4196		55.5425	
Loose π^0 veto	4195	0.024%	55.5425	0.000%
Tight π^0 veto	4198	0.048%	55.5425	0.000%
Muon Mode				
$100 < M_{\pi^0} < 160$ MeV	3873		58.9501	
Loose π^0 veto	3872	0.026%	58.9501	0.000%
Tight π^0 veto	3875	0.052%	59.0521	0.170%
Electron-Muon Mode				
$100 < M_{\pi^0} < 160$ MeV	6595		64.883	
Loose π^0 veto	6592	0.045%	64.7819	0.160%
Tight π^0 veto	6595	0.000%	64.883	0.000%

Table 9.3: Systematic uncertainty due to the π^0 veto, evaluated by fluctuating the π^0 mass window by $\pm\delta$.

9.6 s_B cut

The s_B cut rejects all background events that have a kaon momentum that is higher than what is kinematically allowed for $B^+ \rightarrow K^+ \tau^+ \tau^-$. As can be seen in Table 6.3, this cut has a negligible effect on the signal efficiency. In addition, the m_{ES} sideband substitution is done after applying the s_B cut, and therefore any disagreement between data and MC on the background estimate has already been accounted for.

9.7 MLP cut

It is important to quantify the systematic uncertainty associated with the cut on the MLP output. The validation test, discussed in section 7.5, shows a good agreement between the data control sample and background MC. The ratio of data-to-MC yield as a function of the MLP cut, shown in Fig. 7.20a and Fig. 7.20b, ranges between 0.965 to 1.005. This value gives the maximum range of the data-MC disagreement due to this multivariate technique. In addition, the discriminating variables that enter into the MLP neural network are shown in section 5.4 and section 7.4, before and after the m_{ES} sideband substitution respectively. These show a good agreement between data and MC. Furthermore, prior to unblinding the signal region, the data distribution in the MLP output sideband region is examined and shown in Fig. 9.19. The MLP output sideband region is defined as the region where the MLP neural network output is less than 0.5. As can be readily seen, the data-MC agreement is satisfactory and is improved after applying the m_{ES} sideband substitution, also shown in Fig. 9.19. The m_{ES} sideband substitution done here yields a C_{yield} of 0.9178, which is close to the nominal C_{yield} used in this analysis. This provides an additional sanity check on the TMVA approach and shows that this multivariate technique does not have a negative impact on the level of data-MC agreement.

Nevertheless, it is important to quantify any residual data-MC discrepancy introduced by this technique and thus determine the systematic uncertainty associated with it. To

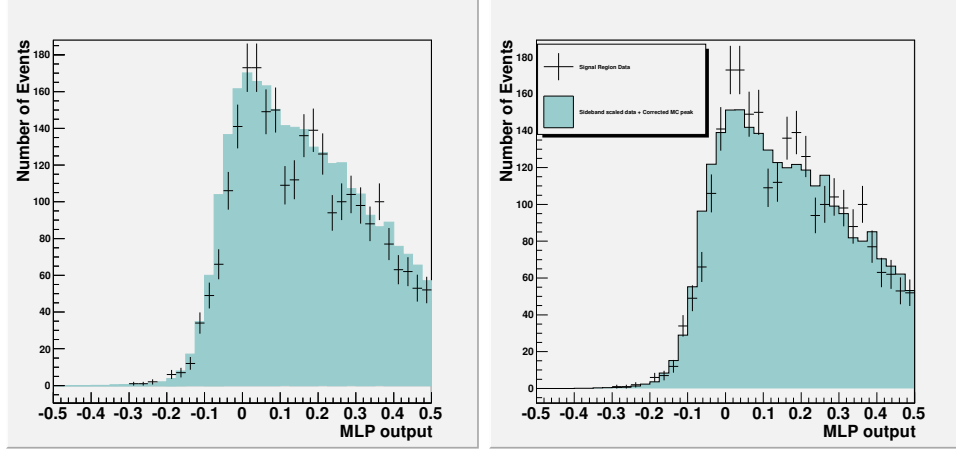


Figure 9.19: MLP output distribution in sideband region before (left) and after (right) m_{ES} sideband substitution.

do so, the results of the validation test, discussed in section 7.5, are used. Fig. 9.20 shows the MLP output distribution of the data and total MC sample, after selecting for the $B^+ \rightarrow D^0 \ell^+ \nu_\ell$, $D^0 \rightarrow K^- \pi^+$ sample and running it through the MLP neural network. To quantify the data-MC discrepancy, the ratio of the data-and-MC yield is calculated for each bin in the distribution. The results are shown in Fig. 9.21 and show a consistent value of about 0.9. A linear fit is then made to this distribution, which gives the average value of this ratio, as shown in 9.22. The result is 0.916 ± 0.026 , which is very close to the nominal C_{yield} used in this analysis. The deviation of this value from the nominal C_{yield} provides the level of discrepancy the MLP cut introduces. Thus, the statistical error on this linear fit, 2.6%, is assigned as the systematic uncertainty associated with the MLP cut.

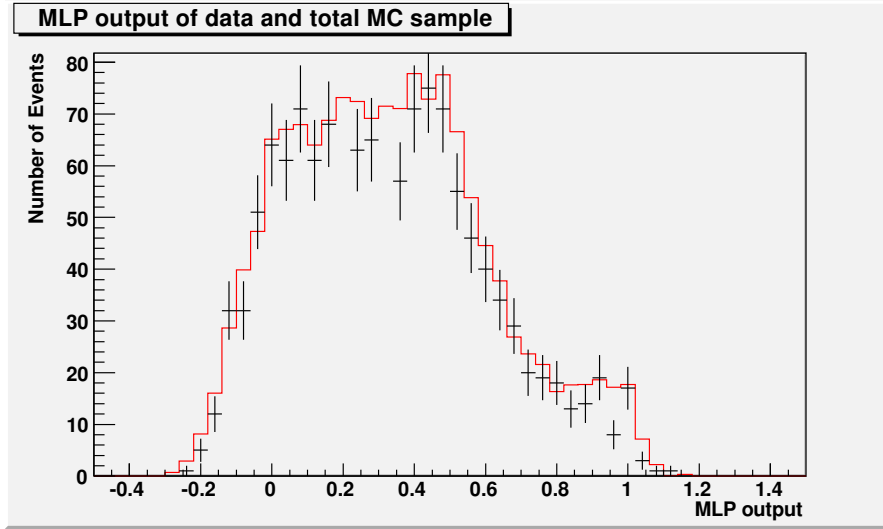


Figure 9.20: MLP output distribution for data (points) and total MC (red) after applying the $B^+ \rightarrow D^0 \nu_\ell$, $D^0 \rightarrow K^+ \pi^-$ selection.

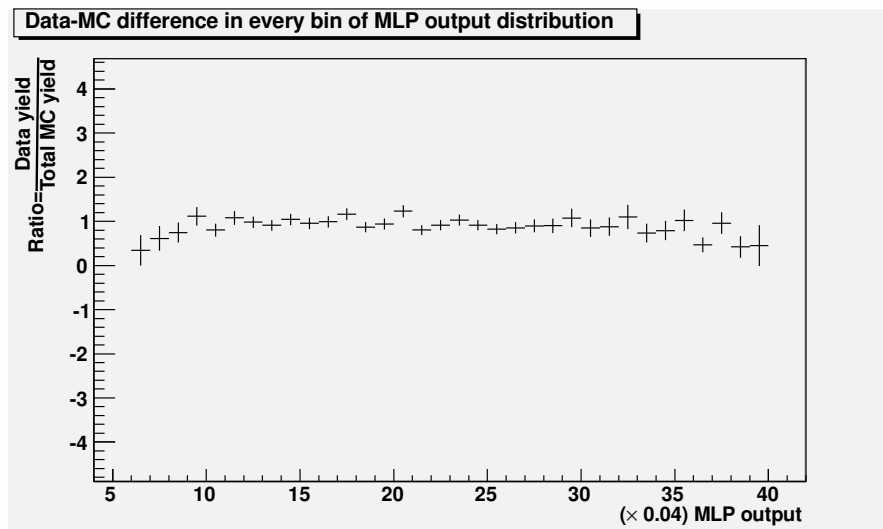


Figure 9.21: Ratio of data to total MC yields, in each bin of the MLP output distribution.

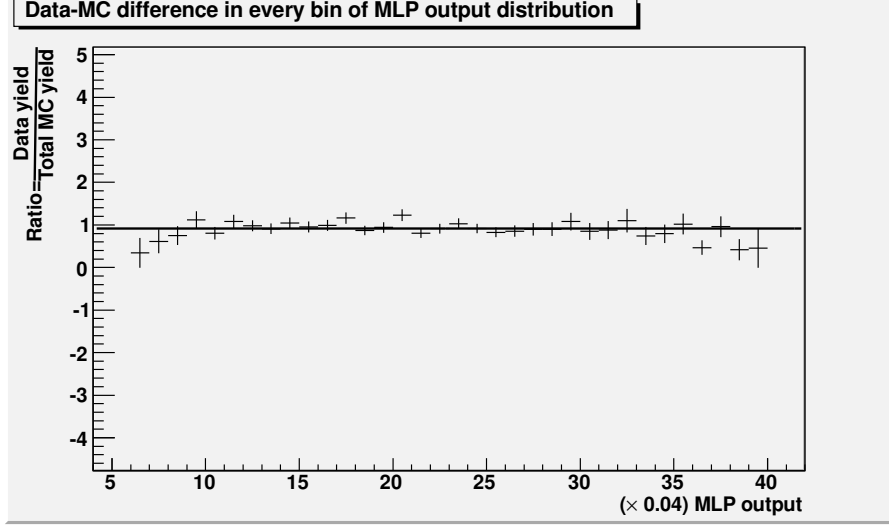


Figure 9.22: Fit to the ratio of data yield to total MC yield in each bin of the MLP output distribution.

9.8 Summary of systematics

Table 9.4 includes all the systematic uncertainties evaluated in this analysis, except for the lepton PID systematic uncertainties, which are evaluated separately for every mode. The latter are shown in Table 9.5. Furthermore, the systematic uncertainties calculated here are applied to both the signal efficiency and the background estimate, except for the uncertainty due to the chosen theoretical model which is applicable to the signal efficiency only. These uncertainties are then added in quadrature to the total statistical error on the signal efficiency and background estimate, before calculating the branching fraction.

Source	Systematic Error
B_{tag} yield	$\epsilon_{sig} : 1.2\%$ $N_{bkg} : 1.60\%$
Theoretical Model	3.0%
Kaon PID	2.05%
π^0 Veto	3%
s_B cut	-
MLP cut	2.6%

Table 9.4: Summary of systematics uncertainties on the signal efficiency and background estimate (except for the uncertainty due to theoretical model) common for all 3 signal channels.

Mode	Lepton PID systematic
Electron	4.76%
Muon	6.95%
Electron-Muon	4.97%

Table 9.5: Summary of lepton PID systematics for each of the 3 signal channels. These uncertainties are applied on the signal efficiency and background estimate.

Chapter 10

Results

After finalizing the signal selection and determining the signal efficiency and background estimate, with the associated systematic uncertainties, the signal region data is unblinded, giving access to the final signal yields in data. The final numbers are shown in Table 10.1 for each of the electron, muon, and electron-muon modes, along with the combined result for all three modes. As previously mentioned, the signal region is defined as the region in the MLP output distribution > 0.70 for the e^+e^- and $\mu^+\mu^-$ modes and > 0.75 for the $e^+ \mu^-$ mode. For both the electron and muon modes, the number of observed events is consistent with the background, within statistical error. The electron-muon mode shows an excess in data of 3.7σ . The combined mode shows a less significant excess, $< 2\sigma$. To examine the excess observed in the electron-muon mode, the distribution of the discriminating variables, used in the MLP neural network, are examined in section 10.1.

The unblinded MLP distributions are shown in Fig. 10.1 for each mode, with the associated signal distributions. The m_{ES} sideband substituted plots are shown in Fig. 10.2 for each mode. As can be seen, the data does not show a significant peak in the signal region of each mode. With the observed number of events, the combined central branching fraction and the upper and lower limits are calculated using the Barlow method. The branching fraction

Mode	$\epsilon_{sig}(\times 10^{-5})$	N_{bkg}	N_{obs}
Electron	1.11 ± 0.12	49.4 ± 3.8	45.0 ± 6.7
Muon	1.29 ± 0.21	45.8 ± 54.0	39.0 ± 6.2
Electron-Muon	2.05 ± 0.26	59.2 ± 4.4	92.0 ± 9.6
Combined	4.77 ± 0.42	154.4 ± 9.6	176.0 ± 13.2

Table 10.1: Efficiency (ϵ_{sig}), background estimate (N_{bkg}), and number of observed events (N_{obs}) for each of the electron, muon, and electron-muon modes. ϵ_{sig} and N_{bkg} include statistical and systematic errors, while N_{obs} has a statistical error only.

Mode	Branching Fraction ($\times 10^{-3}$)	Lower Limit ($\times 10^{-3}$)	Upper Limit ($\times 10^{-3}$)
Electron	$-0.85^{+1.49}_{-1.11}(stat)^{+0.77}_{-0.62}(sys)$	-2.61	1.31
Muon	$-1.13^{+1.20}_{-0.88}(stat)^{+0.68}_{-0.70}(sys)$	-2.62	0.65
Electron-Muon	$3.40^{+1.10}_{-0.88}(stat)^{+0.85}_{-0.56}(sys)$	2.06	5.19

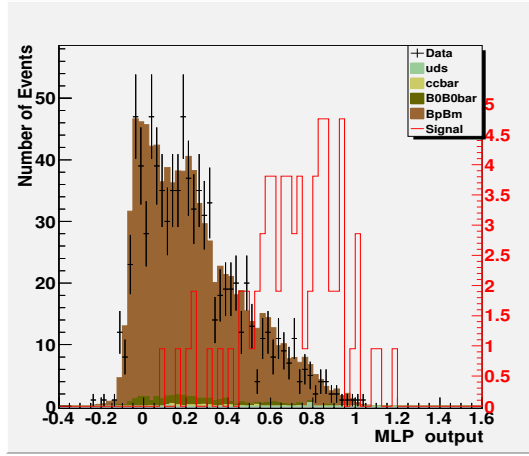
Table 10.2: $\mathcal{B}(B^+ \rightarrow K^+\tau^+\tau^-)$ central value, 90% upper and lower limits for each mode separately determined using the Barlow method.

for all 3 modes combined is:

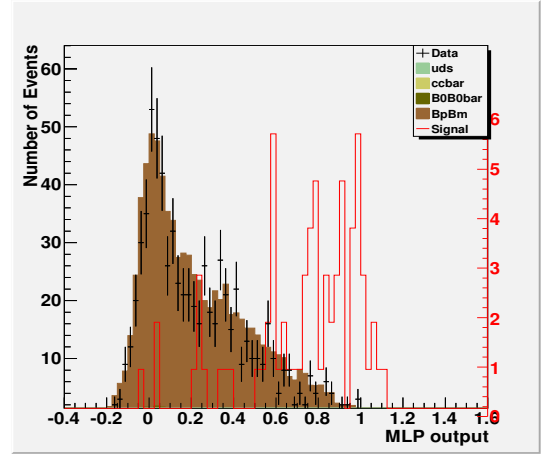
$$\mathcal{B}(B^+ \rightarrow K^+\tau^+\tau^-) = 1.311^{+0.66}_{-0.61}(stat)^{+0.35}_{-0.25}(sys) \times 10^{-3} \quad (10.1)$$

The combined upper limits, at the 90% confidence level, is $\mathcal{B}(B^+ \rightarrow K^+\tau^+\tau^-) < 2.25 \times 10^{-3}$.

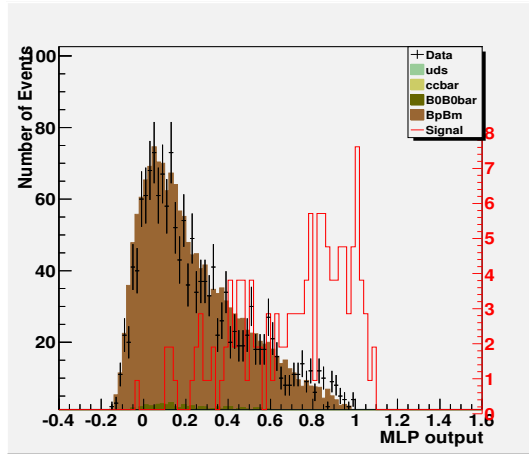
For each mode separately, Table 10.2 shows the central value, upper and lower limits calculated also using the Barlow [93] method.



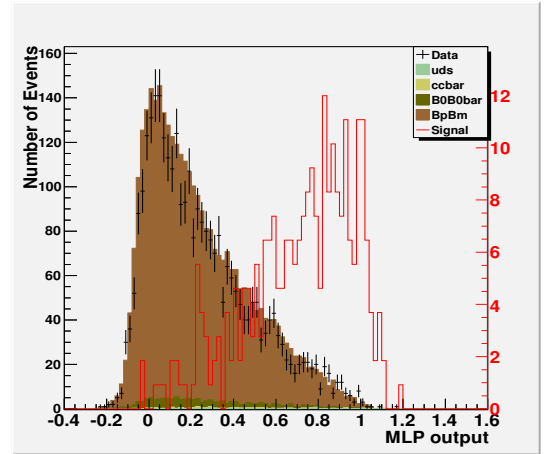
(a) Electron Mode



(b) Muon mode

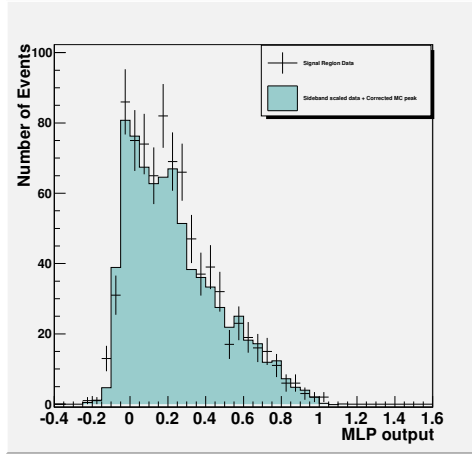


(c) Electron-muon mode

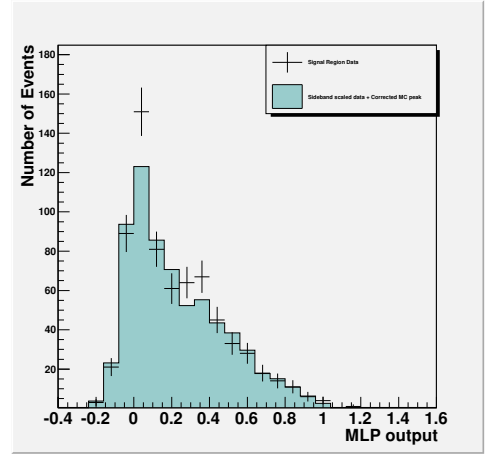


(d) Combined

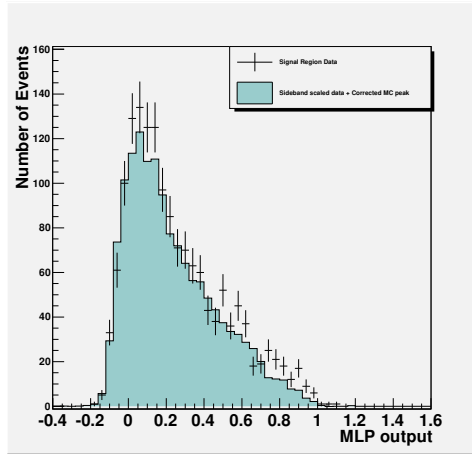
Figure 10.1: MLP output distributions with unblinded data and signal MC.



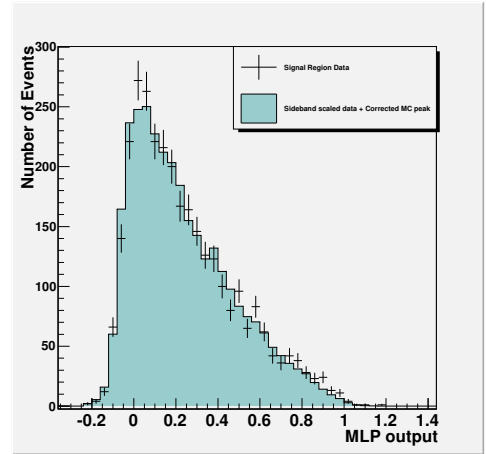
(a) Electron Mode



(b) Muon Mode



(c) Electron-Muon Mode



(d) Combined

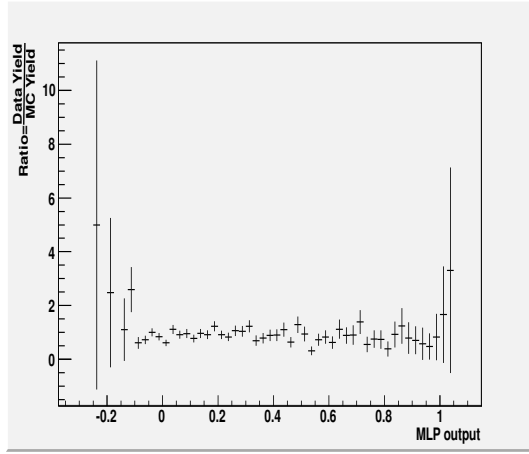
Figure 10.2: MLP output distributions with unblinded data after m_{ES} sideband substitution.

10.1 Discussion

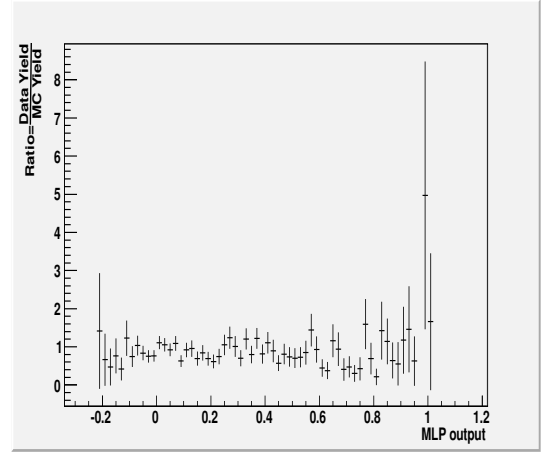
After unblinding the signal region, the data distribution is further examined. The plots, shown in Fig. 10.1, reveal a generally good agreement between data and MC in the signal region and thus the expected background. To quantify this, Fig. 10.3 shows the ratio of data to MC yield, N_{data}/N_{MC} , for each bin of the MLP output distribution. As can be readily seen, the value of the ratio fluctuates about 1 for all 3 modes, which shows a general consistency between data and MC. For the electron and muon modes, there is a slight deviation in various regions of the MLP output distribution to values below 1. This implies that the MC overestimates the data and this is exactly what the m_{ES} sideband substitution corrects for. However, for the electron-muon mode, the data actually exceeds the MC especially in the signal region. This effect can be easily seen in the m_{ES} plots for all 3 modes after the MLP neural network cut in Fig. 10.4. By applying the m_{ES} sideband substitution to the signal region of the MLP output, the discrepancy between data and MC improves for the electron and muon modes, but yet increases for the electron-muon mode.

It is evident that there is a negative deviation from the expected background estimate for the electron and muon modes, but a positive one for the electron-muon mode. Fig. 10.5 shows the C_{yield} and R_{comb} , as a function of cut, determined separately for each mode. As can be readily seen, R_{comb} is consistent with the chosen value and among all 3 modes. However, after the MLP cut where the signal region data is unblinded, C_{yield} reflects the negative deviation in the electron and muon modes, and the positive one in the electron-muon mode, consistent with the observed data yields in the three modes. This can be verified using Fig. 10.6, which shows the peaking data vs. peaking MC for each mode separately. Furthermore, as can be readily seen, the fluctuation in C_{yield} only occurs at the final stage of the analysis. Before the MLP cut, the MC consistently overestimates the data for all cuts. This implies that such a result could not have been predicted by studies performed prior to unblinding the signal region.

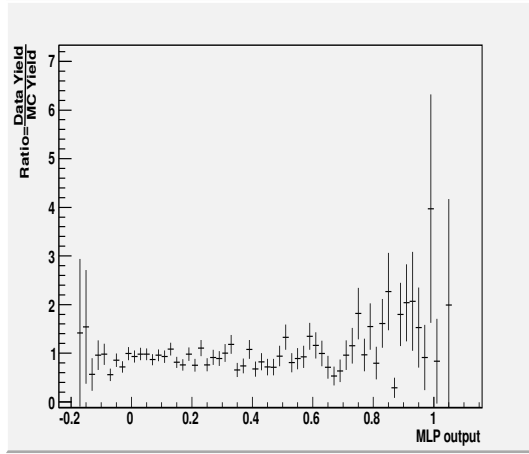
Although the excess observed in the $e^+ \mu^-$ mode is not statistically significant, we can



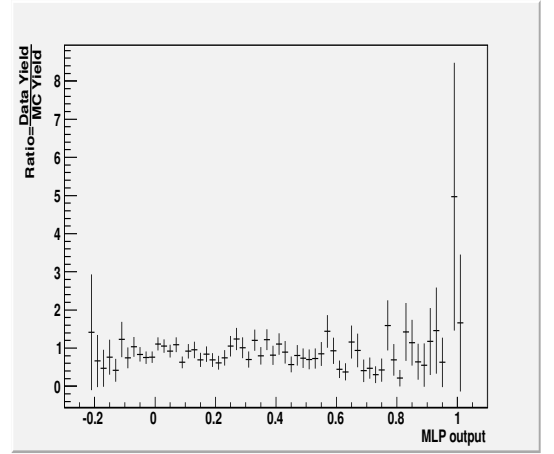
(a) Electron Mode



(b) Muon mode

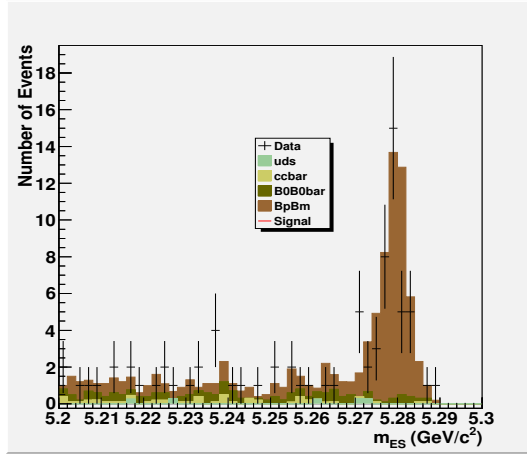


(c) Electron-muon mode

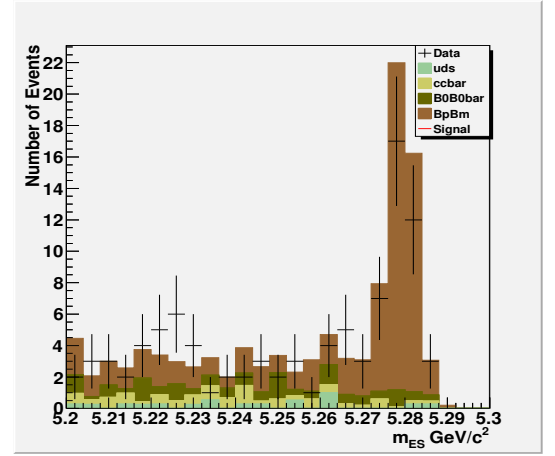


(d) Combined

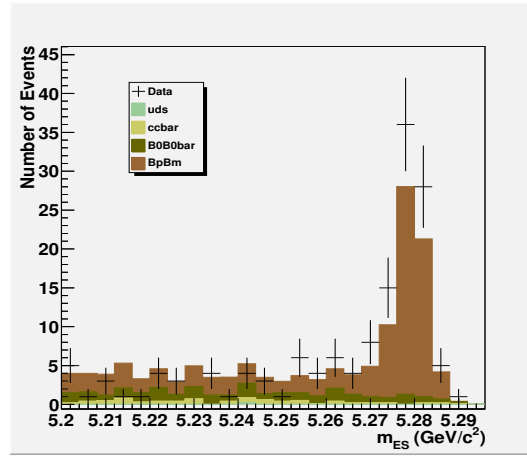
Figure 10.3: Ratio of data to MC yield for each bin of the MLP output distribution.



(a) Electron Mode

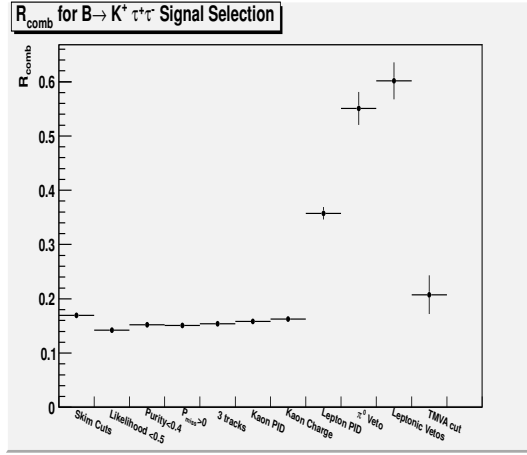


(b) Muon mode

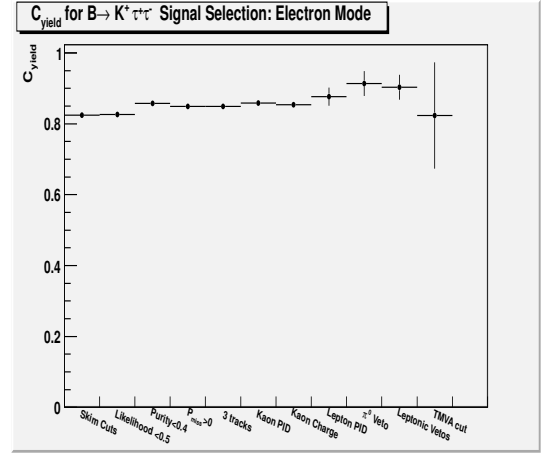


(c) Electron-muon mode

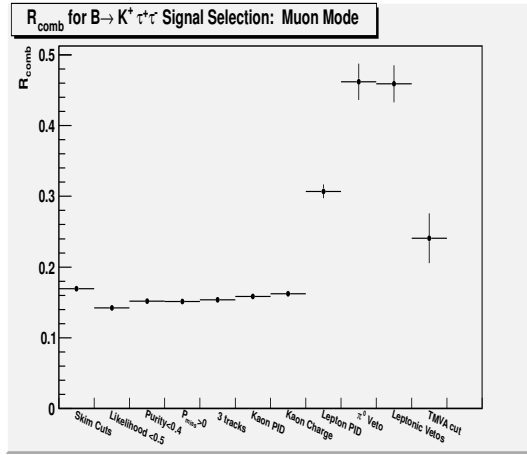
Figure 10.4: m_{ES} distribution of each mode after the MLP neural network cut.



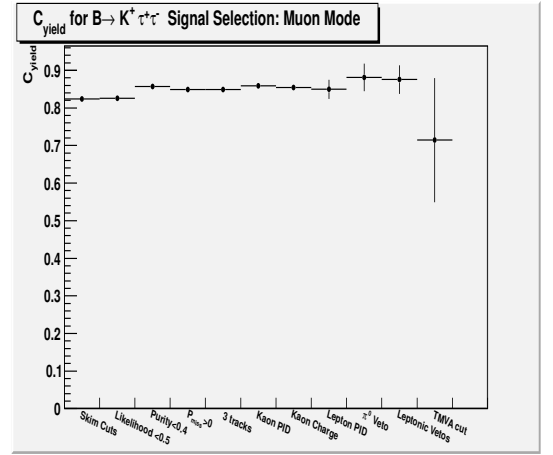
(a) R_{comb} for electron mode



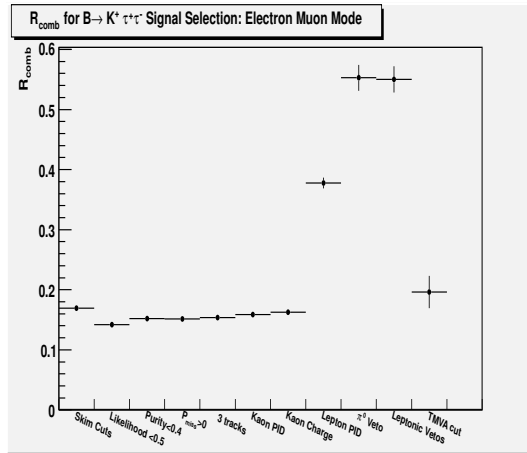
(b) C_{yield} for electron mode.



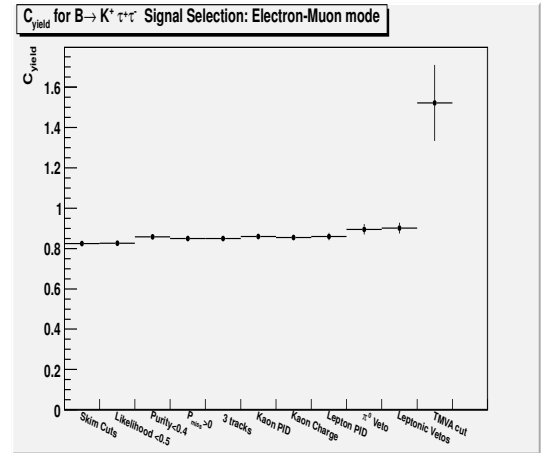
(c) R_{comb} for muon mode.



(d) C_{yield} for muon mode.

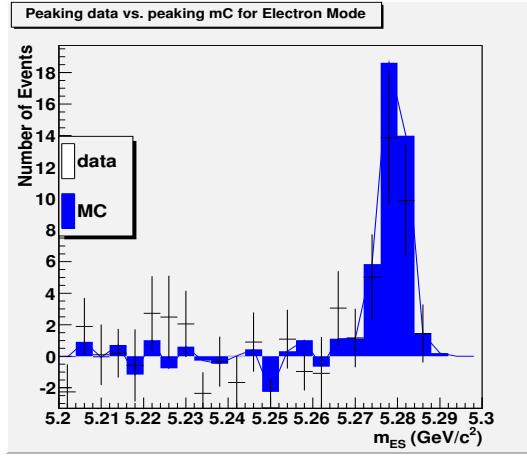


(e) R_{comb} for electron-muon mode.

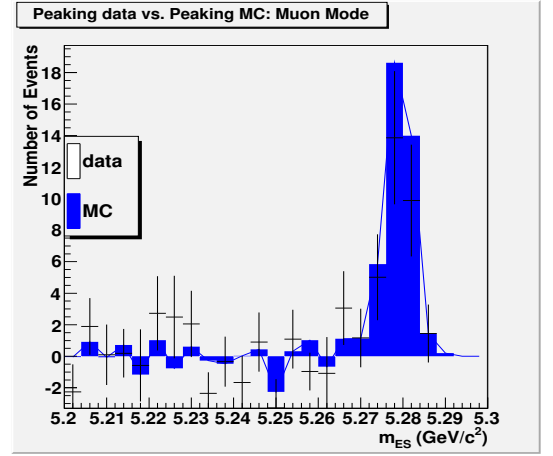


(f) C_{yield} for electron-muon mode.

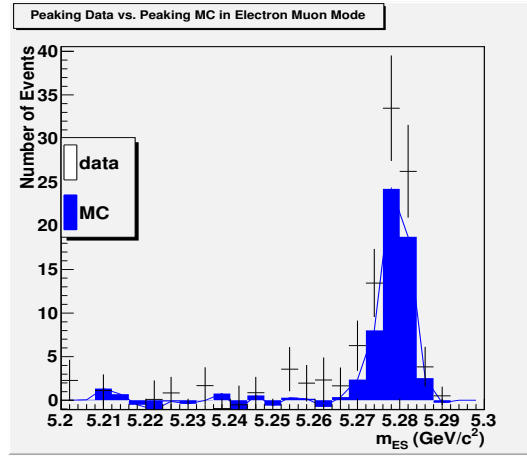
Figure 10.5: R_{comb} (left) and C_{yield} (right) for each mode as a function of cut.



(a) Electron Mode



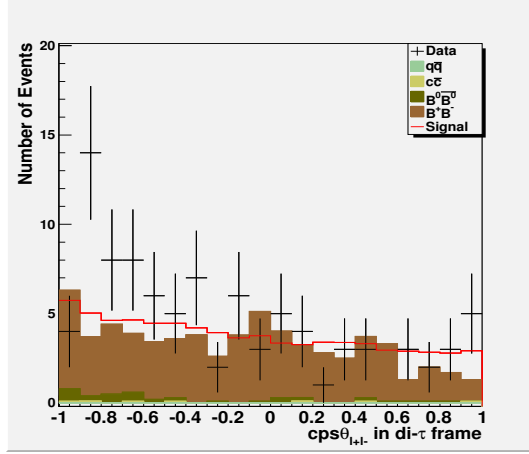
(b) Muon mode



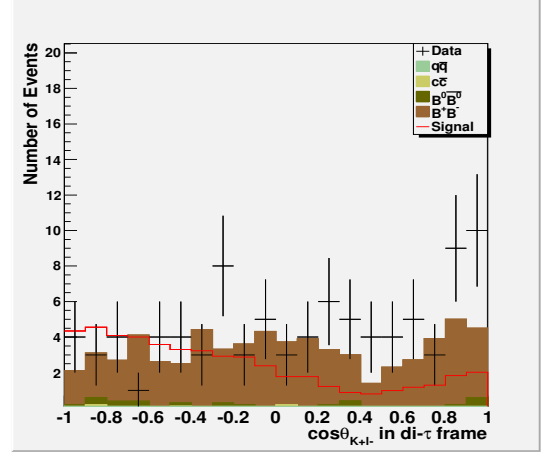
(c) Electron-muon mode

Figure 10.6: Peaking data vs. peaking MC after MLP neural network cut.

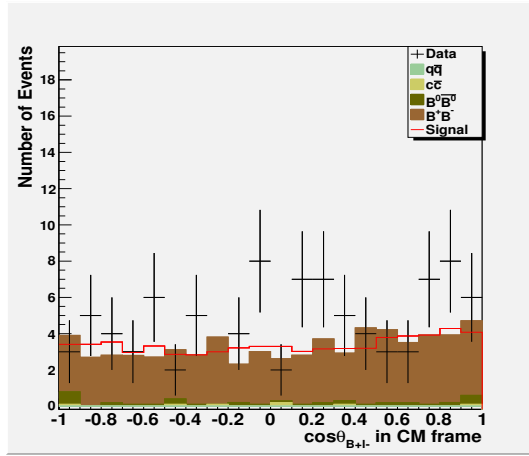
still examine the data in this mode for consistency with a possible $B^+ \rightarrow K^+ \tau^+ \tau^-$ signal. Therefore, plots of the discriminating variables used in the MLP neural network are shown for the electron-muon mode after applying the final cut in Figs. 10.7, 10.8, and 10.9. As can be readily seen, there is no clear indication that the excess observed is signal-like. There is also no significant indication of systematic data-MC disagreement in the plots. Furthermore, plots of the discriminating variables, separately for this mode, are examined before and after the MLP cut, with and without the m_{ES} sideband substitution. These are shown in Appendix G and the agreement between data and MC is also satisfactory. Finally, the absence of signal in the electron and muon modes further suggests that the excess in the electron-muon mode is more likely a fluctuation than a $B^+ \rightarrow K^+ \tau^+ \tau^-$ signal.



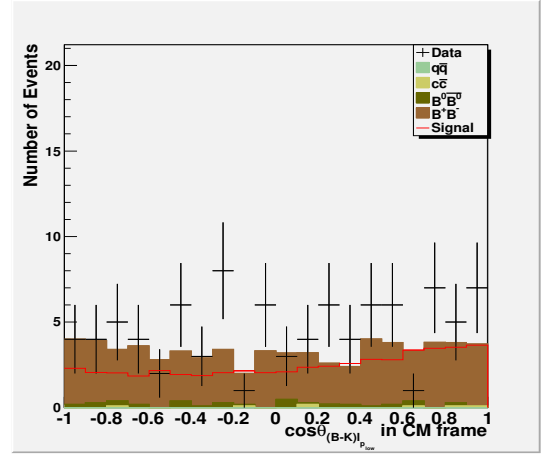
(a) $\cos \theta_{l+l-}$ in di- τ frame.



(b) $\cos \theta_{K+l-}$ in di- τ frame.

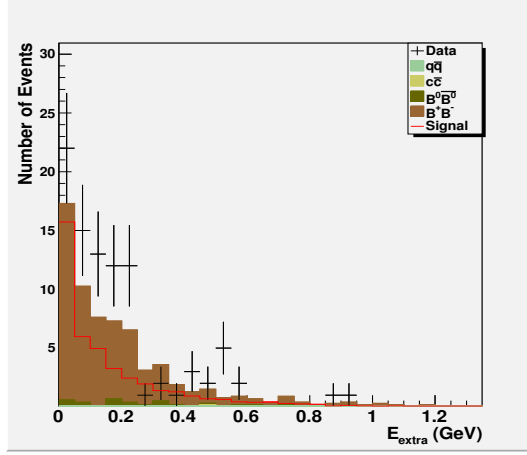


(c) $\cos \theta_{B+l-}$ in CM frame.

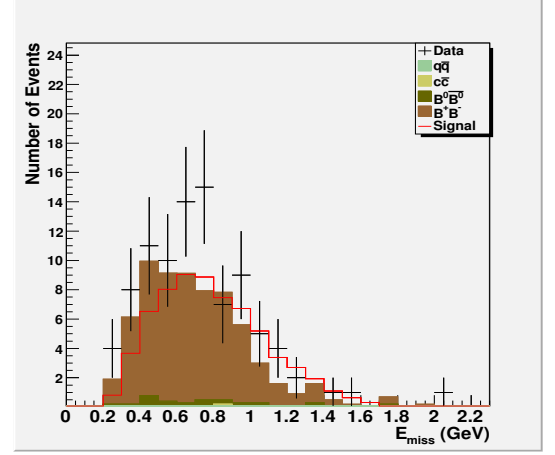


(d) $\cos \theta_{(B-K)l_{low}}$ in di- τ frame.

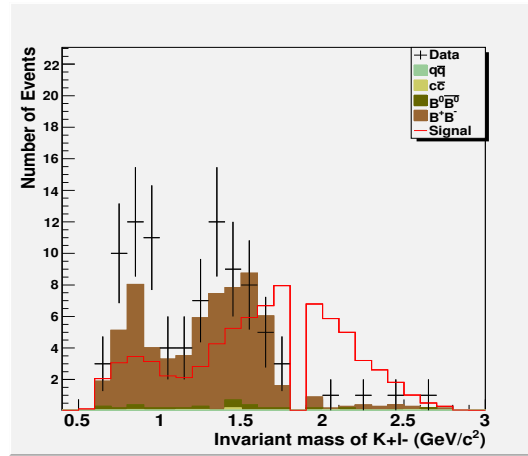
Figure 10.7: Angular variables of the electron-muon mode, used in the MLP neural network, after applying the final cut in the analysis.



(a) E_{extra}

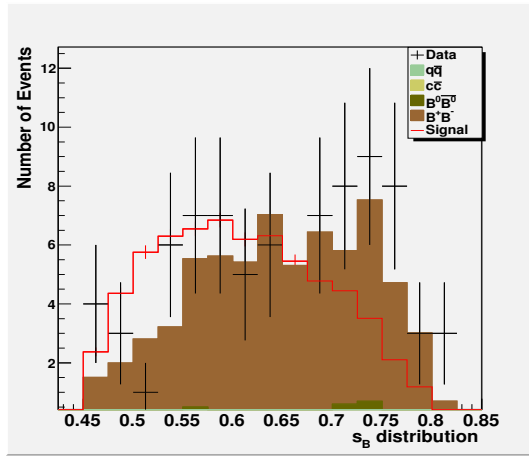


(b) E_{miss}

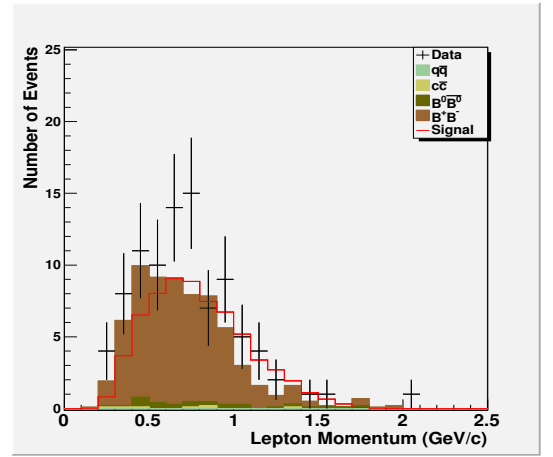


(c) Invariant mass of $K+l^-$ in CM frame.

Figure 10.8: Calorimeter variables of the electron-muon mode, used in the MLP neural network, after applying the final cut in the analysis.



(a) s_B



(b) Lepton Momentum

Figure 10.9: s_B distribution (left) and kinematic variable of the electron-muon mode (right), after applying the final cut in the analysis. The latter is used in the MLP neural network.

Chapter 11

Conclusion

In conclusion, a search for the FCNC process $B^+ \rightarrow K^+ \tau^+ \tau^-$ has been presented, where the $\tau^+ \tau^-$ pair decays leptonically into either $e^+ e^- \nu_e \bar{\nu}_e$ (electron mode), $\mu^+ \mu^- \nu_\mu \bar{\nu}_\mu$ (muon mode), or $e^+ \mu^- \nu_e \bar{\nu}_\mu$ (electron-muon mode) in the final state. Hadronic B_{tag} reconstruction is employed in this analysis. Signal and background MC samples are used to develop the signal selection and study the main backgrounds. Peaking background events of the form $B^+ \rightarrow D^{(*)} \ell^+ \nu_\ell$, $D^{(*)} \rightarrow K \ell \bar{\nu}_\ell$ are suppressed using a MLP neural network, trained with 8 angular, calorimeter and kinematic discriminating variables. Furthermore, using one of the peaking background modes, $B^+ \rightarrow D^0 \ell \bar{\nu}_\ell$, $D^0 \rightarrow K^+ \pi^-$, the signal selection is verified and the level of data-MC agreement is tested. In addition, the background estimate is determined using the m_{ES} sideband substitution, where combinatorial background is calculated using sideband data and peaking background is scaled to correct for data-MC discrepancies. Finally, the branching fraction is determined after unblinding the signal region data. No significant signal is observed. In the absence of signal, the combined branching fraction for all 3 modes is determined, using the Barlow method, to be $\mathcal{B}(B^+ \rightarrow K^+ \tau^+ \tau^-) = 1.31_{-0.61}^{+0.66}(\text{stat})_{-0.25}^{+0.35}(\text{sys}) \times 10^{-3}$. The upper limit at the 90% confidence level is $\mathcal{B}(B^+ \rightarrow K^+ \tau^+ \tau^-) < 2.25 \times 10^{-3}$.

Although this analysis uses the full *BABAR* data sample, the resulting limit is of $\mathcal{O}(10^{-3})$.

This is about four orders of magnitude above the SM expectation. Selecting for the leptonic decay modes of the τ limits the final signal efficiency and thus the sensitivity, but allows for a clean selection of signal events. The measured upper limit is approximately three orders of magnitude over the sensitivity threshold to contributions of the neutral Higgs boson in 2HDM [45]. The same is true for other extensions of the SM discussed in Refs. [47]-[54]. Nevertheless, this analysis lays the ground work for the search for such a rare decay. The hadronic B_{tag} reconstruction and the trained neural network, with the discriminating variables used, have the potential to be a successful strategy for extracting a branching fraction measurement with a higher statistics data sample.

The Belle-2 experiment is currently under construction and will be the next generation “super” B -factory [96]. It is the extension of experiments like *BABAR* and Belle to higher luminosities. This is an exciting prospect for $B^+ \rightarrow K^+ \tau^+ \tau^-$. With a data sample that is of order 100 times larger than the *BABAR* data, sensitivity to new physics scenarios that might be affecting $B^+ \rightarrow K^+ \tau^+ \tau^-$ observables may come within our experimental reach. Data taking by the Belle-2 experiment is scheduled to start in 2018 and the target integrated luminosity is 50 ab^{-1} . The resulting new data sample will allow for a more accurate measurement of $B^+ \rightarrow K^+ \tau^+ \tau^-$, which may potentially lead to the observation of physics beyond the Standard Model.

Bibliography

- [1] C. Huang and Q. Yan, Phys. Lett. B **442**, 209-216 (1998).
- [2] G. Aad *et al.* [ATLAS Collaboration], Phys. Lett. B **716**, 1-29 (2012).
- [3] N. Arkani-Hamed, S. Dimopoulos, G. Dvali, Phys. Lett. B **429**, 263-272 (1998).
- [4] J. P. Lees *et al.* [BABAR Collaboration], Phys. Rev. D **86**, 03 (2012).
- [5] J. T. Wei *et al.* [Belle Collaboration], Phys. Rev. Lett. **103**, 171801 (2009).
- [6] T. Aaltonen *et al.* [CDF Collaboration], Phys. Rev. Lett. **107**, 201802 (2011).
- [7] R. Aaij *et al.* [LHCb Collaboration], JHEP **08**, 131 (2013).
- [8] R. Barbieri, G. Isidori and A. Pattori, Eur. Phys. J. C **76** 2, 67 (2016)
- [9] L. Calibbi, A. Crivellin and T. Ota, Phys. Rev. Lett. **115**, 181801 (2015).
- [10] G. C. Branco *et al.*, Phys .Rept. **516**, 1-102 (2012).
- [11] J. Hewitt, Phys. Rev. D **53**, 4964-4969 (1996).
- [12] S. Choudhry, N. Guar, and A. Gupta, Phys. Rev. D **60**, 115004 (1999).
- [13] J. P. Lees *et al.* [BABAR Collaboration], Phys. Rev. Lett. **109**, 101802 (2012).
- [14] A. Bozek *et al.* [Belle Collaboration], Phys. Rev. D **82**, 072005 (2010).
- [15] R. Aaij *et al.* [LHCb Collaboration], Phys. Rev. Lett. **115**, 111803 (2015).

- [16] C. Quigg, “Gauge Theories of the Strong, Weak and Electromagnetic Interactions”, The Benjamin/Cummings Publishing Company (1983).
- [17] B. R. Martin and G. Shaw, “Particle Physics” , John Wiley & Sons (2007).
- [18] K. A. Olive *et al.* (Particle Data Group), Chin. Phys. C **38**, 090001 (2014).
- [19] W. Wang, International Journal of Modern Physics A **29**, 1430040 (2014)
- [20] R. Aaij *et al.* [LHCb Collaboration], Phys. Rev. Lett. **115**, 072001 (2015).
- [21] A. Djouadi, Phys. Rept. **457**, 1-216 (2008).
- [22] B. Aubert *et al.* [BABAR Collaboration], Nucl. Instrum. Meth. A **729**, 615-701 (2013).
- [23] A. Abashian *et al.* [Belle Collaboration], Nuc. Instrum. Meth. A **479**, 117-232 (2002).
- [24] See for example: B. Aubert *et al.* [BABAR Collaboration], Phys. Rev. Lett. **94**, 161803 (2005); B. Aubert *et al.* [BABAR Collaboration], Phys. Rev. D **71**, 032005 (2005).
- [25] M. Beyer, “CP Violation in Particle, Nuclear and Astrophysics”, Springer-Verlag Berlin Heidelberg, (2002).
- [26] J. H. Christenson *et al.*, Phys. Rev. Lett. **13**, 138 (1964).
- [27] “The 2008 Nobel Prize in Physics - Press Release”, Nobelprize.org, [http : //www.nobelprize.org/nobel – prizes/physics/laureates/2008/press.html](http://www.nobelprize.org/nobel-prizes/physics/laureates/2008/press.html).
- [28] T. Mannel, Nuclear Physics B (Proc. Suppl.) **167**, 170-174 (2007).
- [29] S. Robertson, “Searches for New Physics in Rare Leptonic and Electroweak Decays”, Presented at Cornell University, Ithaca, New York (2007).
- [30] J. Goldstone, A. Salam and S. Weinberg, Phys. Rev. **127**, 965 (1962).
- [31] See for example: G. Aad *et al.* [ATLAS collaboration], Eur. Phys. J. C **75**, .231 (2015); S. Chatrchyan *et al.* [CMS collaboration], Phys. Rev. D **89**, 092007 (2014).

- [32] T. Kajita, Rep. Prog. Phys. **69**, 1607-1635 (2006).
- [33] G. Aad *et al.* [ATLAS Collaboration], J. Instrum. **3**, S08003 (2008).
- [34] S. Martin, arXiv:hep-ph/9709356 (1997).
- [35] O. Deschamps, S. Descotes-Genon, S. Monteil, V. Niess, S. T’Jampens, V. Tisserand, Phys. Rev. D **82**, 073012 (2010); A. K. Grant, Phys. Rev. D **51**, 207 (1995)
- [36] J. Polchinski, arXiv:hep-th/9210046 (1999).
- [37] M. Neubert, arXiv:hep-ph/0512222 (2005).
- [38] K. Wilson and W. Zimmermann, Commun. Math. Phys. **24**, 87-106 (1972).
- [39] A. J. Bevan *et al.* [BABAR and Belle Collaborations], “The Physics of the B -factories”, Eur. Phys. J. C **74**, 3026 (2014).
- [40] B. Grinstein, M. J. Savage, and M. Wise, Nucl. Phys. B **319**, 271 (1989).
- [41] T. Inami and C.S.Lim, Prog. Theor. Phys. **65**, 297 (1981).
- [42] N. Cabibbo and L. Maiani, Phys. Lett. B **79**, 109 (1978).
- [43] D. Du, C. Liu, and D. Zhang, Phys. Lett. B **317**, 179 (1993).
- [44] R. Casalbuoni, A. Deandrea, N. Di Bartolomeo, R. Gatto and G. Nardulli, Phys. Lett. B **312**, 315 (1993).
- [45] T. M. Aliev, M. Savci, A. Ozpineci, J. Phys. G **24**, 49-65 (1998).
- [46] Y. Dai, C. Huang and H. Huang, Phys. Lett. B **390**, 257 (1997).
- [47] C. Huang and Y. Qi-Shu, Phys. Lett. B **442**, 209-216 (1998).
- [48] J. Hewett and J. D. Wells, Phys. Rev. D **55**, 5549-5560 (1997).
- [49] Y. Dai, C. Huang, and H. Huang, Phys. Lett. B **390**, 257-262 (1997).

- [50] Q. Yan *et al.*, Phys. Rev. D **62**, 094023 (2000).
- [51] D. Guetta and E. Nardi, Phys. Rev. D **58**, 012001 (1998).
- [52] S. Choudhury, N. Guar, and A. Gupta, Phys. Rev. D **60**, 115004 (1999).
- [53] Y. Kim, P. Ko, J. Lee, Nucl. Phys. B **544**, 64-88 (1999).
- [54] A. Ali, P. Ball, L.T. Handoko, and G. Hiller, Phys. Rev. D **61**, 074024 (2000).
- [55] A. Ali, W. Lunghi, C. Greub, and G. Hiller, Phys. Rev. D. **66**, 034002 (2002).
- [56] M. Zhong, Y. L. Wu, and W. Y. Wang, Int. J. Mod. Phys. A **18**, 1959 (2003).
- [57] C. Bouchard *et al.*, Phys. Rev. Lett. **111**, 162002 (2013).
- [58] R. Aaij *et al.* [LHCb Collaboration], JHEP **07**, 133 (2012).
- [59] R. Aaij *et al.* [LHCb Collaboration], JHEP **02**, 105 (2013).
- [60] R. Aaij *et al.* [LHCb Collaboration], Phys. Rev. Lett. **113**, 151601 (2014).
- [61] R. Aaij *et al.* [LHCb Collaboration], arXiv:1403.8044v3 (2014).
- [62] M. Schram, “Heavy flavor meson decays with tau leptons in the final state at BaBar”,
Presented at 2010-11th International Workshop on Tau Lepton Physics, Manchester,
UK (2010).
- [63] B. Aubert *et. al* [*BABAR* Collaboration], Nucl. Instrum. Meth. A **479**, 1-116 (2002).
- [64] S. Robertson, “Observation of CP violation in B meson decays”, Presented at University
of Calgary, Alberta (2001).
- [65] C. Bozzi *et. al*, Nucl. Instrum. Meth. A **453**, 78-83 (2000).
- [66] G. Sciolla *et. al*, Nucl. Instrum. Meth. A **419**, 310-314 (1998).
- [67] B. Aubert *et. al*, Nucl. Instrum. Meth. A **479**, 1-116 (2002).

- [68] J. Schwiening, “Performance of the *BABAR* DIRC”, Presented at the Fifth Workshop on Rich Detectors at Playa del Carmen, Mexico, SLAC-PUB-11017 (2005).
- [69] E. D. Frank, R. G. Jacobsen, and E. Sexton-Kennedy, “Architecture of the *BABAR* reconstruction system”, Presented at Computing in High Energy Physics, Chicago, IL, USA (1997).
- [70] D. J. Lange, Nucl. Instrum. Meth. A **462**, 152 (2001).
- [71] T. Sjostrand, “PYTHIA 5.7 and JETSET 7.4 Physics and Manual”, CERN-TH 7112/93 (1993).
- [72] S. Agostinelli *et al.* [GEANT4 Collaboration], Nucl. Instrum. Meth. A **506**, 250 (2003).
- [73] F. Rademakers and R. Brun, Linux Journal, **51** (1998).
- [74] J. P. Lees *et al.* [*BABAR* Collaboration], Phys. Rev. Lett. **109**, 101802 (2012).
- [75] J. P. Lees *et al.* [*BABAR* Collaboration], arXiv:1401.2703 (2013).
- [76] D. Boutigny *et al.* [*BABAR* Collaboration], “The *BABAR* Physics Book: Physics at an Asymmetric Factory”, SLAC Stanford, CA (1998).
- [77] J. P. Lees *et al.* [*BABAR* Collaboration], Phys. Rev. D **89**, 112005 (2013).
- [78] W. Verkerke and D. Kirkby, “RooFit users manual V2.07”, <http://roofit.sourceforge.net/docs/> (2006).
- [79] G. Fox and S. Wolfram, Nucl. Phys. B **149**, 413-496 (1979).
- [80] S. R. Safavian and D. Landgrebe, IEEE transactions on systems, man, and cybernetics **21**, 660-674 (1991).
- [81] K. Machová *et al.*, Acta Polytechnica Hungarica **3.2**, 121-132 (2006).

- [82] T. G. Dietterich and G. Bakiri, Journal of Artificial Intelligence Research, **2**, 263-286 (1995).
- [83] B. Denby, Neural Computation **5**, 505-549 (1993).
- [84] I. I. Balitsky, V. M. Braun, and A. V. Kolesnichenko, Nucl. Phys. B **312**, 509 (1989).
- [85] V. L. Chernyak and I. R. Zhitnisky, Nucl. Phys. **345**, 137 (1990).
- [86] W. Jaus and D. Wyler, Phys. Rev. D **41**, 3405 (1990); B. Stech, Phys. Lett. **B 354**, 447 (1995); Z. Phys. C **75**, 245 (1997); C. Q. Geng and C. P. Kao, Phys. Rev. D **54**, 5636 (1996); D. Melikhov, N. Nikitin, S. Simula, Phys. Lett. B **410**, 290 (1997).
- [87] A. Abada *et al.* [APE Collaboration], Phys. Lett. B **365**, 275 (1996); J. M. Flynn and C. T. Sachrajda [UKQCD Collaboration] arXiv:hep-lat/9710057; J. M. Flynn *et al.*, Nucl. Phys. B **461**, 327 (1996); L. Del Debbio *et al.* [UKQCD Collaboration]m arXiv: hep-lat/9708008 (1997).
- [88] G. Burdman, Phys. Rev. D **52**, 6400 (1995); W. Roberts, Phys. Rev. D **54**, 863 (1996).
- [89] L. Lellouch, Nucl. Phys. B **479**, 353 (1996).
- [90] D. Melikov, N. Nikiten, and S. Simula, Phys. Rev. D **57**, 6814-6829 (1998).
- [91] D. Melikhov, Phys. Rev. D **53**, 2460 (1996); Phys. Lett. B **380**, 363 (1996); Phys. Lett. B **394**, 385 (1997).
- [92] B. Aubert *et al.* [BABAR Collaboration], arXiv:hep-ex/0207063v1 (2002).
- [93] R. Barlow, Comput. Phys. Commun. **149**, 97 [arXiv:hep-ex/0203002] (2002).
- [94] G. Feldmann and R. Cousins, Phys. Rev. D **57**, 3873 (1998).
- [95] B. Aubert *et al.* [BABAR collaboration], Phys. Rev. Lett. **99**, 201801 (2007).
- [96] T. Abe *et al.*, arXiv:1011.0352 (2010).

[97] R. Aaij *et al.* [LHCb Collaboration], Int. J. Mod. Phys. A **30**, 1530022 (2015).

Appendices

Appendix A

Discriminating variables

The discriminating variables that enter into the MLP neural network are shown here for both the generic and cocktail signal MC samples. These are done after all signal cuts, except the MLP cut. As can be readily seen, the distributions show no major discrepancy between the two signal samples, except that the cocktail has larger statistics. The cocktail signal MC is scaled arbitrarily to show any differences in the shape of the distribution.

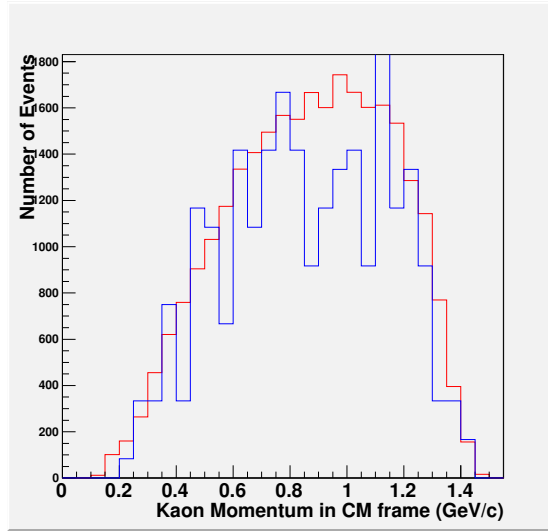


Figure A.1: Kaon Momentum in CM frame for cocktail (red) and generic (blue) signal MC samples.

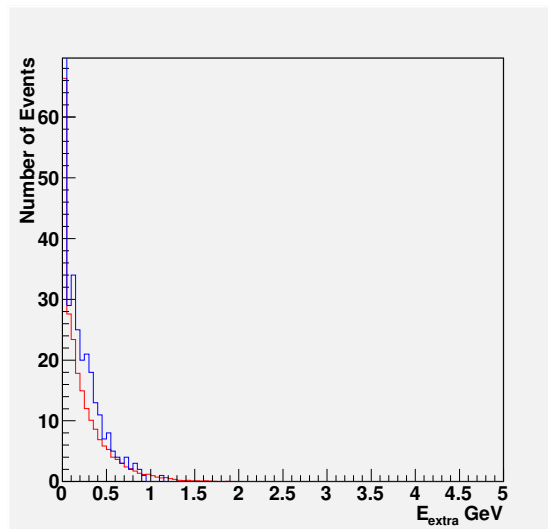


Figure A.2: E_{extra} for cocktail (red) and generic (blue) signal MC samples.

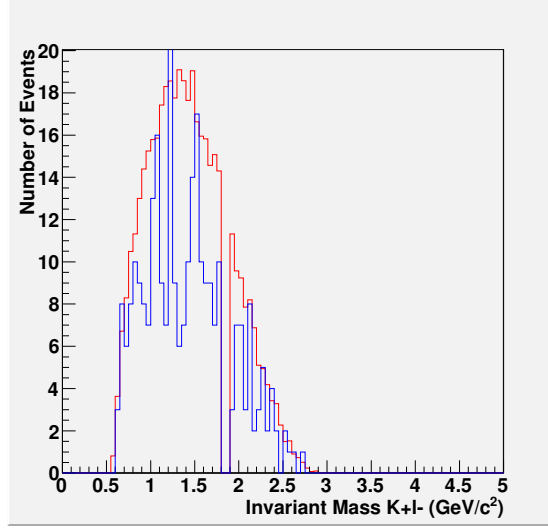


Figure A.3: Invariant mass of kaon with oppositely charge lepton in lab frame for cocktail (red) and generic (blue) signal MC samples.

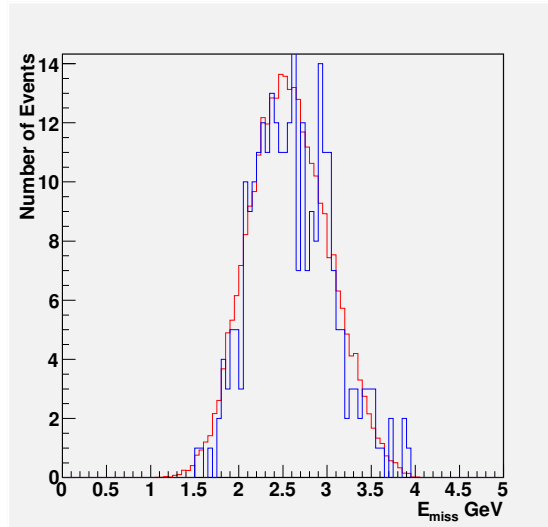


Figure A.4: E_{miss} in di-tau frame for cocktail (red) and generic (blue) signal MC samples.

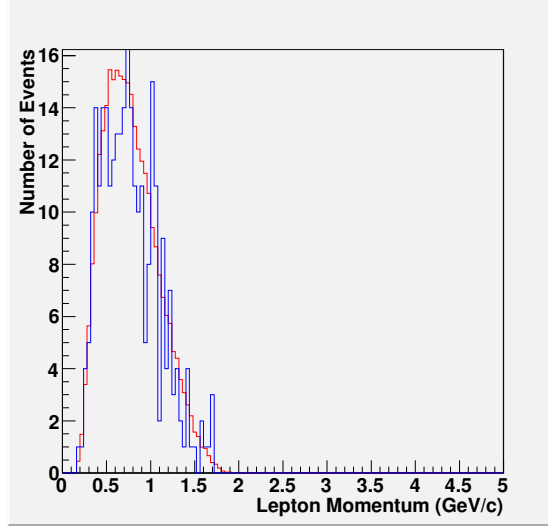


Figure A.5: Momentum of lepton with charge opposite to kaon for cocktail (red) and generic (blue) signal MC samples.

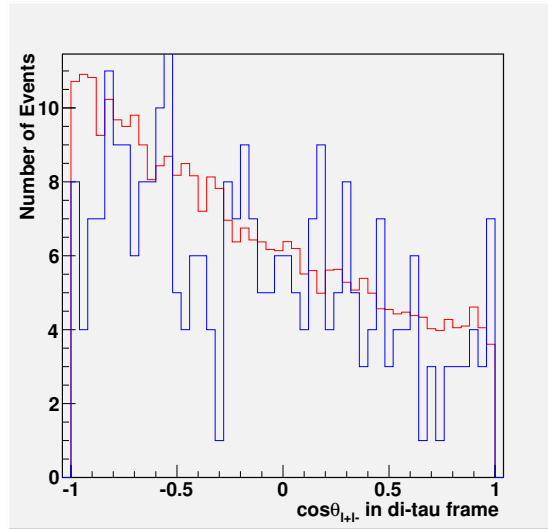


Figure A.6: $\cos \theta_{\ell+\ell-}$ in di-tau frame for cocktail (red) and generic (blue) signal MC samples.

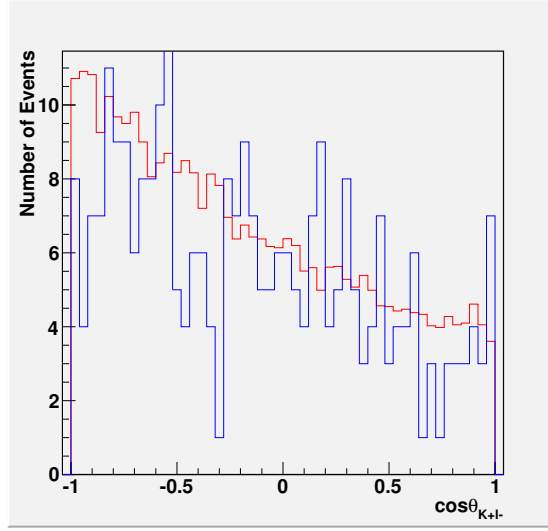


Figure A.7: $\cos \theta_{K+l-}$ in di-tau frame for cocktail (red) and generic (blue) signal MC samples.

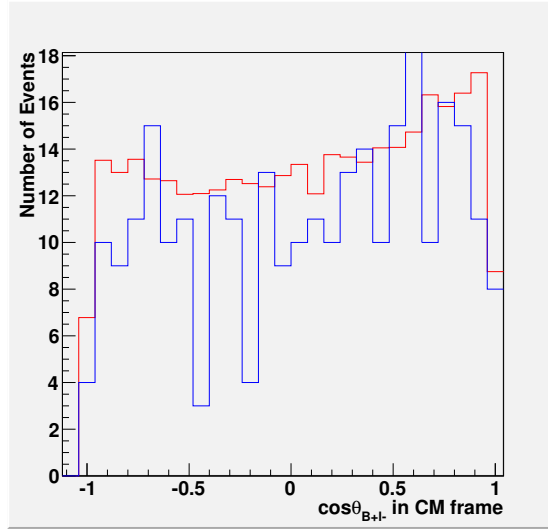


Figure A.8: $\cos \theta_{B+l-}$ in CM frame for cocktail (red) and generic (blue) signal MC samples.

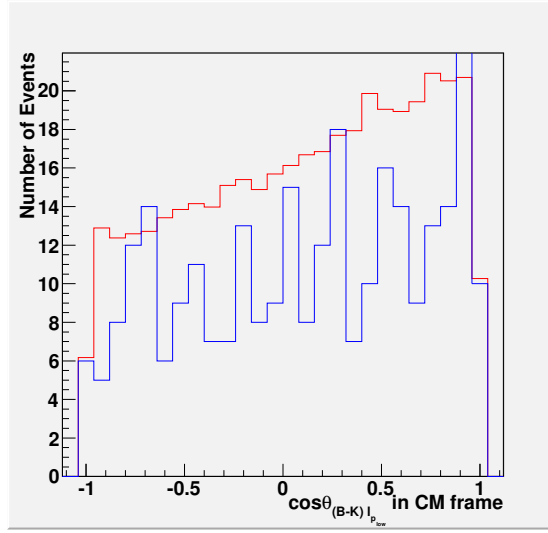


Figure A.9: $\cos\theta_{(B_{\text{sig}}-K)l_{p_{low}}}$ in CM frame for cocktail (red) and generic (blue) signal MC samples.

Appendix B

Region above D meson in M_{K+l-} distribution

Looking at plot 5.13c , the region above $1.8 \text{ GeV}/c^2$ in the invariant mass distribution of the combination of the K with the oppositely charged lepton, referred to as M_{K+l-} , is almost background free and thus it is interesting to treat it separately. One suggestion is to train and test the MLP neural network on all events with M_{K+l-} below $1.8 \text{ GeV}/c^2$, which corresponds to the D^0 mass, and apply the MLP cut. Signal and background events in the region $M_{K+l-} > 1.8 \text{ GeV}/c^2$ are then added to the final distribution, in attempt to increase the signal efficiency and improve the sensitivity of this analysis. Fig. B.1 shows the MLP output of the signal and background events for each of the electron, muon, and electron-muon modes, when the cut $M_{K+l-} < 1.8 \text{ GeV}/c^2$ is applied before training and testing the MLP neural network. As can be seen for the electron mode, the double-peak structure in the MLP output is no longer apparent. This implies that there is a set of background events in the region $M_{K+l-} > 1.8 \text{ GeV}/c^2$ that has a distinct structure, causing a small second peak in the MLP output distribution. The MLP cut is then applied to the events below the D^0 mass, while events above $1.8 \text{ GeV}/c^2$ are just added to the final distribution. The results are shown in Table B.1. The signal efficiency is not much increased, implying that almost all

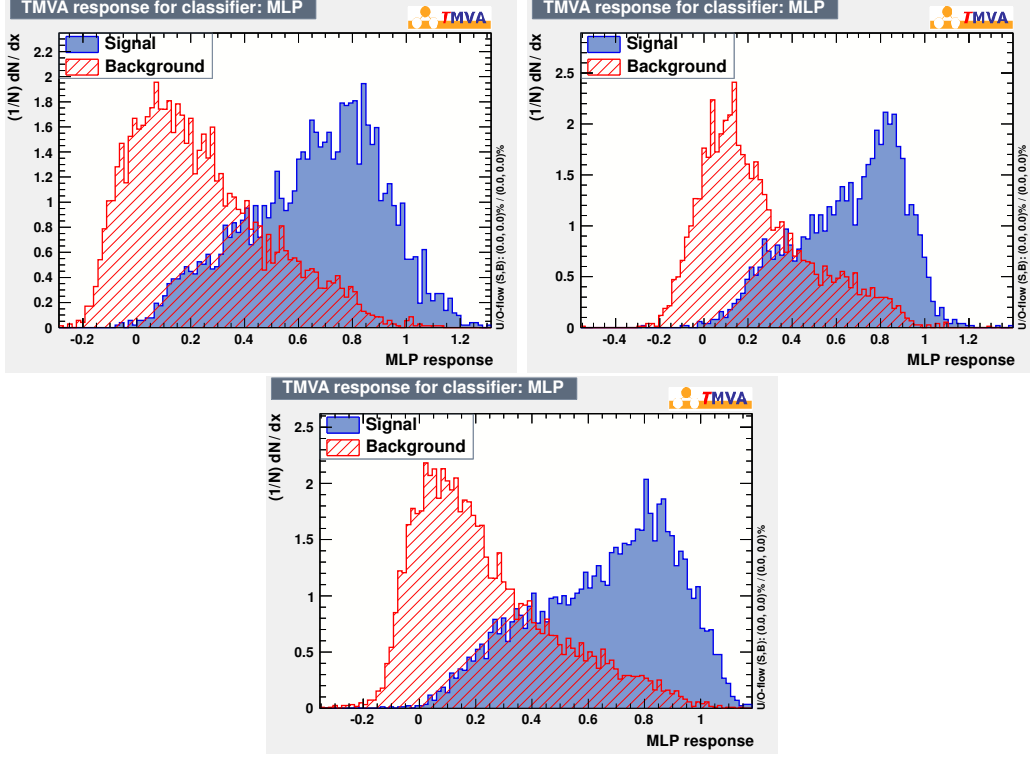


Figure B.1: MLP output for each of the electron (top-left), muon (top-right) and electron muon (bottom) modes, when the MLP neural network is trained and tested in the region $M_{K+\ell^-} < 1.8 \text{ GeV}/c^2$.

events with $M_{K+\ell^-} > 1.8 \text{ GeV}/c^2$ in the signal MC already have a MLP output value greater than the MLP cut. Thus, simply appending these events does not improve the final signal efficiency. The background estimate increases because a small subset of background events is present above the D^0 mass. From the results in Table B.1, it is concluded that treating the region above the D^0 mass in the $M_{K+\ell^-}$ distribution does not improve the sensitivity of this analysis.

Mode	Efficiency ($\times 10^{-5}$)	N_{bkg}	N_{B+B^-}	N_{pkgBkg}	$N_{sideData}$	$N_{nonPkgBkg}$
Electron Mode	1.19 ± 0.233	58.5 ± 5.95	56.2 ± 2.48	51.4 ± 5.77	32 ± 5.66	7.13 ± 1.46
Muon Mode	1.35 ± 0.265	52.8 ± 5.66	45.8 ± 2.41	41.8 ± 5.29	49 ± 7	10.9 ± 2
Electron Muon	2.02 ± 0.344	60.4 ± 5.93	53.4 ± 2.51	48.8 ± 5.58	52 ± 7.21	11.6 ± 2.01

Table B.1: The final values for $B \rightarrow K^+ \tau^+ \tau^-$.

Appendix C

Kaon Momentum after MLP cut

As previously mentioned, the main motivation behind this analysis is the search for new physics contributions. Therefore, it is essential that none of the discriminating variables used in the MLP neural network could affect our ability to detect new physics. Now, the presence of any new physics contribution will impact the kaon momentum distribution, which must reside in the low region due to the heavy mass of the $\tau^+\tau^-$ pair. It is thus important to verify that the neural network used does not hinder our sensitivity to the full kinematically allowed region of kaon momentum. To ensure this is the case, a plot of the kaon momentum and the s_B distribution is shown below for all 3 modes combined. This plot is made after applying a MLP cut of > 0.7 . Data is still blinded at this stage and is therefore not shown in this distribution. As can be readily seen, the kaon momentum distribution is still fully accessible after the MLP cut.

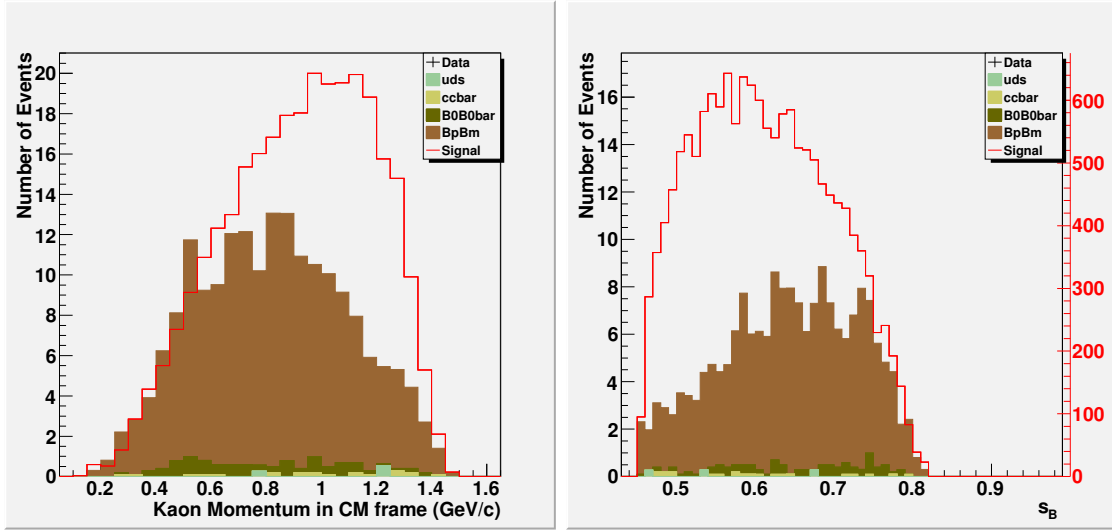


Figure C.1: Kaon Momentum in CM frame (left) and s_B distribution (right) after MLP cut for both signal and background MC.

Appendix D

Combinatorial B^+B^- and $B^0\bar{B}^0$ components in m_{ES} distribution

The assumption made, using the m_{ES} sideband substitution, is that the combinatorial component of B^+B^- MC in the signal region has the same shape of that of the $B^0\bar{B}^0$ MC. This assumption is made to calculate the combinatoric ratio, R_{comb} , and to isolate the peaking component of B^+B^- MC, as discussed in section 7.1. To verify this assumption, the combinatorial components of B^+B^- and $B^0\bar{B}^0$ are examined in the continuum likelihood sideband region, which is here defined as the region where the continuum ratio is less than 0.5. Furthermore, the combinatorial component is further isolated by applying the wrong cut on the charge of the B_{tag} . Thus, for $B^0\bar{B}^0$ combinatorial component, the applied cut on the B_{tag} charge is ± 1 , and for the B^+B^- MC, the charge must be zero. Fig. D.1 shows the m_{ES} distribution of the B^+B^- MC in the continuum likelihood sideband region, after requiring a neutral B_{tag} candidate. At this point, the purity cut and E_{miss} cut are also applied as well. The corresponding $B^0\bar{B}^0$ distribution, after requiring a charged B_{tag} candidate, is also shown in Fig. D.1. Both distributions have a very similar shape in the m_{ES} signal region. This verifies the assumption made in the m_{ES} sideband substitution method that the combinatorial B^+B^- component has the same shape as the combinatorial $B^0\bar{B}^0$.

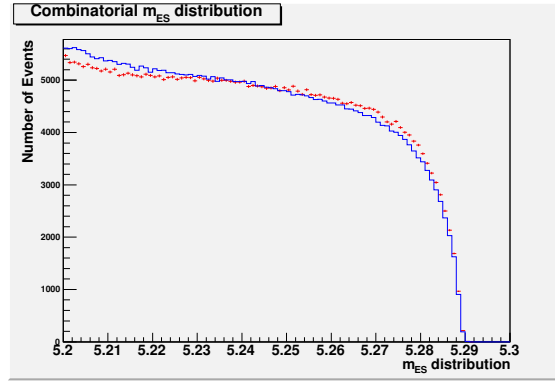


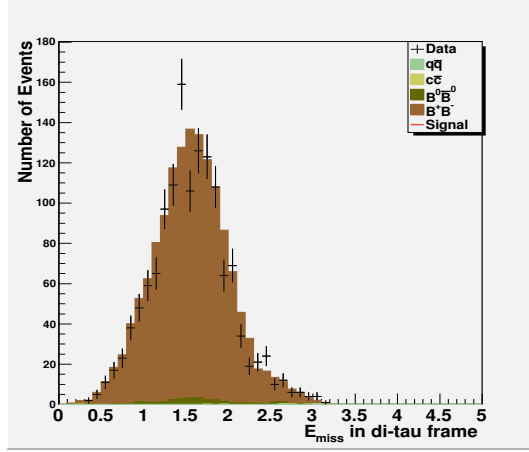
Figure D.1: Combinatorial B^+B^- (blue) and $B^0\bar{B}^0$ (red) distribution in the continuum likelihood sideband region after requiring a neutral and charged B_{tag} candidate, respectively.

Appendix E

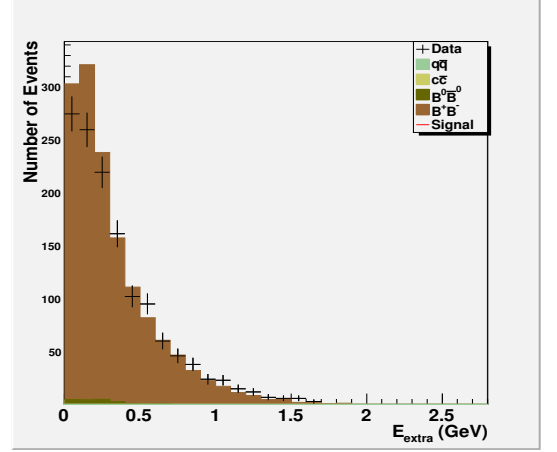
Discriminating Variables of Validation Test

Plots of the discriminating variables of the $B^+ \rightarrow D^0 \ell \nu_e$, $D^0 \rightarrow K^+ \pi^-$ control sample and the corresponding background MC, used in the MLP neural network validation test, are shown below, before and after the m_{ES} sideband substitution. The agreement between data and MC is satisfactory.

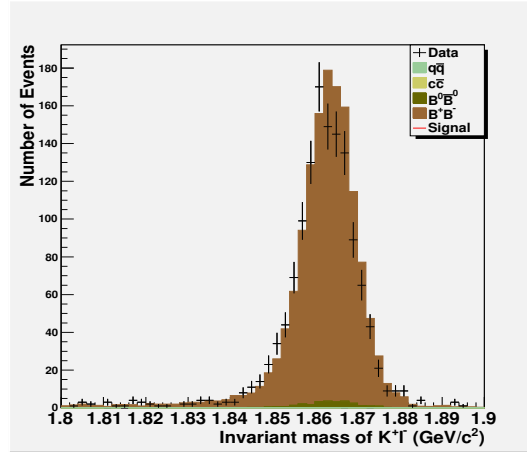
E.1 Discriminating variables before m_{ES} sideband substitution



(a) Missing energy in the di-tau frame.



(b) E_{extra} distribution in the CM frame.



(c) Invariant mass of the kaon and oppositely charged lepton sum.

Figure E.1: Calorimeter variables used in the MLP neural network validation test. The data distribution (black points) is overlaid on the background MC distributions (color-filled). $q\bar{q}$ refers to $q = u, d, s$.

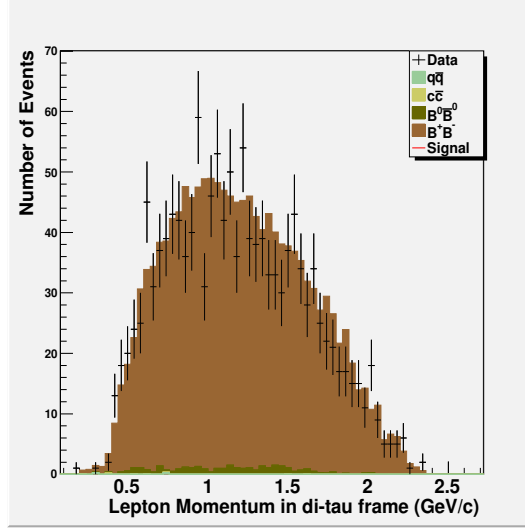
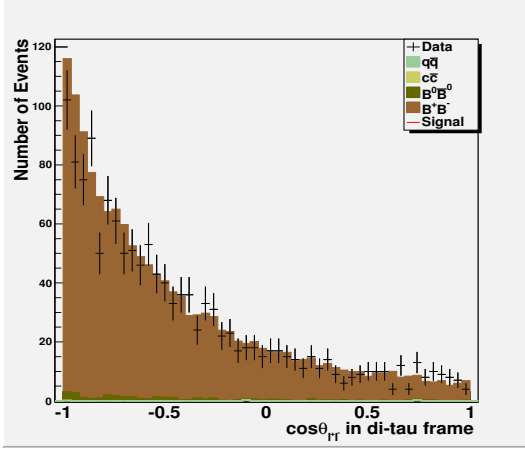
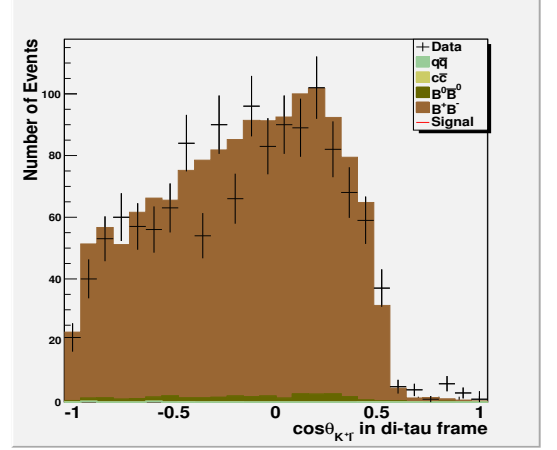


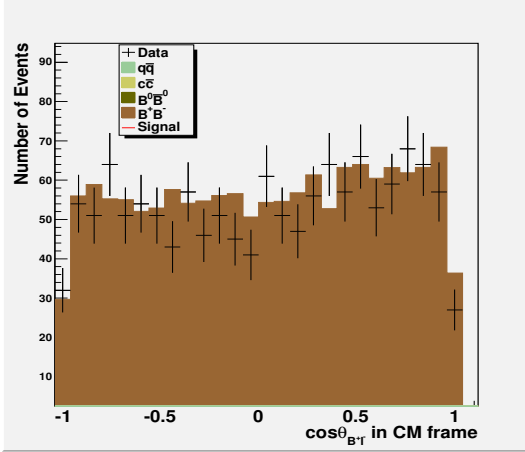
Figure E.2: Momentum of lepton, with charge opposite to that of the kaon, in the di-tau frame used in the MLP neural network validation test. The data distribution (black points) is overlaid on the background MC distributions (color-filled). $q\bar{q}$ refers to $q = u, d, s$.



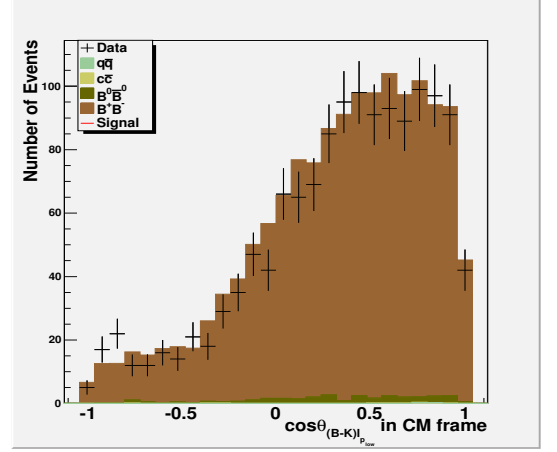
(a) $\cos \theta_{l+l-}$ in di-tau frame.



(b) $\cos \theta_{K+l-}$



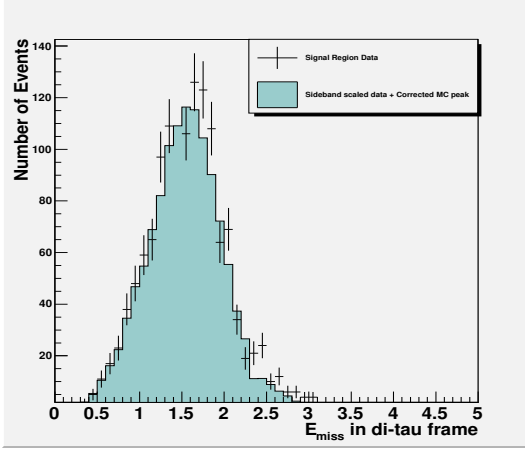
(c) $\cos \theta_{B+l-}$ in center-of-mass frame.



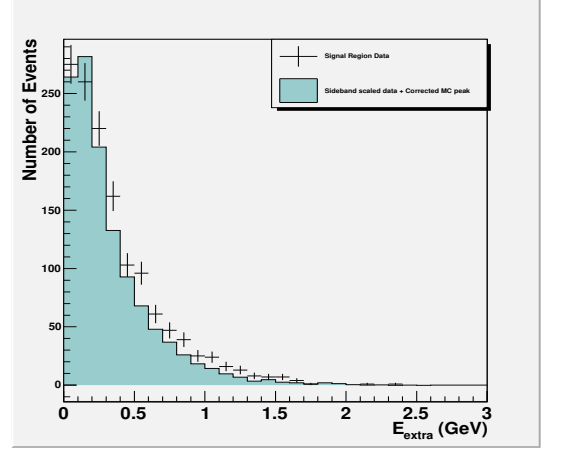
(d) $\cos \theta_{(B_{\text{sig}}-K)l_{\text{low}}}$ in CM frame.

Figure E.3: Angular variables used in the MLP neural network validation test: angle between two leptons in di-tau frame (top-left), angle between Kaon and oppositely charged lepton in di-tau frame (top-right), angle between B_{sig} and oppositely charged lepton in CM frame (bottom-left), and angle between the Kaon recoil vector and the lepton with low momentum in CM frame (bottom-right). The data distribution (black points) is overlaid on the background MC distributions (color-filled). $q\bar{q}$ refers to $q = u, d, s$.

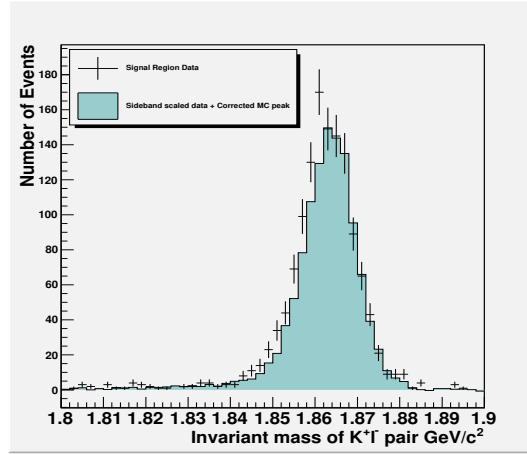
E.2 Discriminating variables after m_{ES} sideband substitution



(a) Missing energy in the di-tau frame.



(b) E_{extra} distribution in the CM frame.



(c) Invariant mass of the kaon and oppositely charged lepton sum.

Figure E.4: Calorimeter variables used in the MLP neural network validation test. The data distribution (black points) is overlaid on the corrected B^+B^- MC + sideband data (color-filled). $q\bar{q}$ refers to $q = u, d, s$.

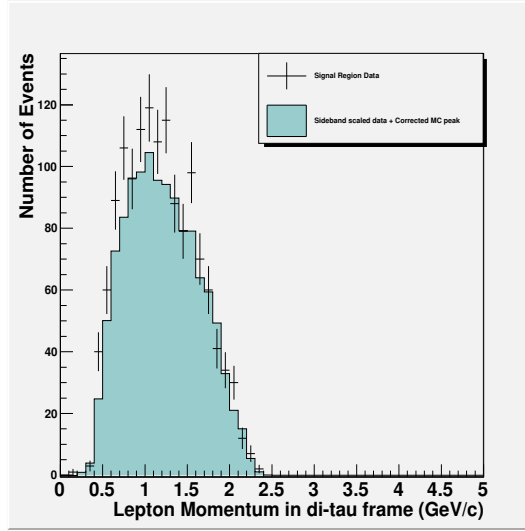
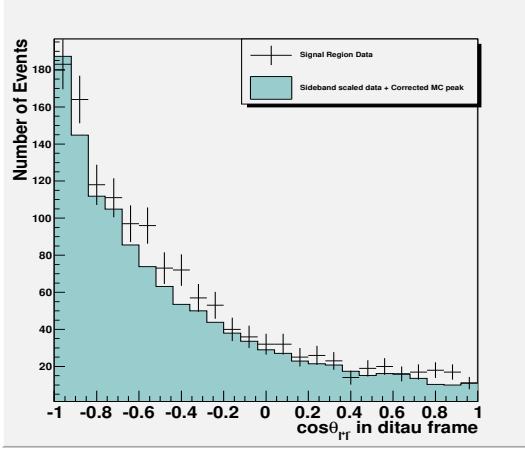
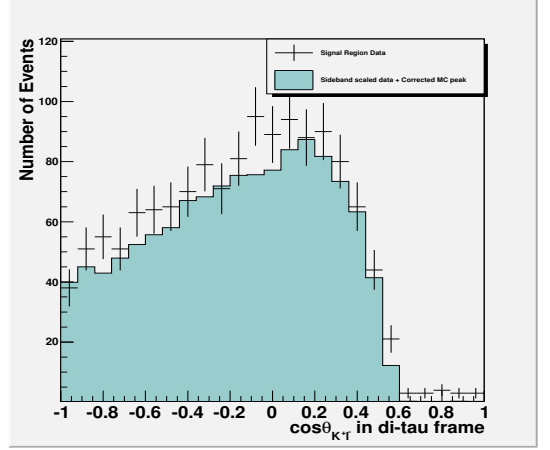


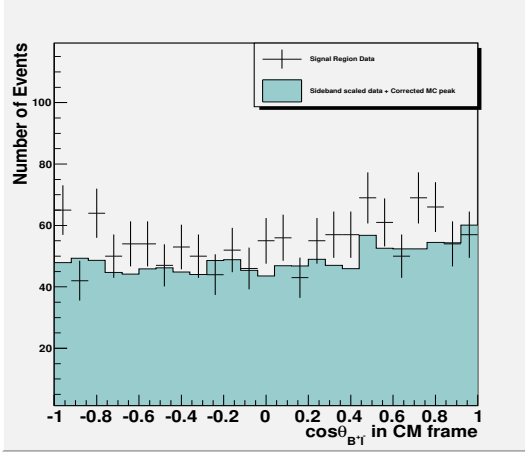
Figure E.5: Momentum of lepton, with charge opposite to that of the kaon, in the di-tau frame used in the MLP neural network validation test. The data distribution (black points) is overlaid on the corrected B^+B^- background + sideband data (color-filled). $q\bar{q}$ refers to $q = u, d, s$.



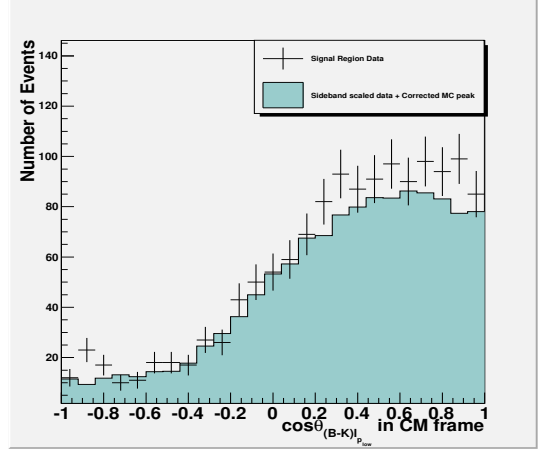
(a) $\cos \theta_{l+l-}$ in di-tau frame.



(b) $\cos \theta_{K+l-}$



(c) $\cos \theta_{B+l-}$ in center-of-mass frame.



(d) $\cos \theta_{(B_{\text{sig}}-K)l_{p_{\text{low}}}}$ in CM frame.

Figure E.6: Angular variables used in the MLP neural network validation test: angle between two leptons in di-tau frame (top-left), angle between Kaon and oppositely charged lepton in di-tau frame (top-right), angle between B_{sig} and oppositely charged lepton in CM frame (bottom-left), and angle between the Kaon recoil vector and the lepton with low momentum in CM frame (bottom-right). The data distribution (black points) is overlaid on the corrected B^+B^- background + sideband data (color-filled). $q\bar{q}$ refers to $q = u, d, s$.

Appendix F

Comment on correlations

As can be seen in section 9, the evaluated systematic errors are common between ϵ_{sig} and N_{bkg} . Given (8.1) used to calculate the central branching fraction, it is thus important to evaluate any correlation between ϵ_{sig} and N_{bkg} . This can be done using toy MC experiments, analogous to the procedure followed for evaluating the B_{tag} yield uncertainty. The approach here is to select a specific systematic uncertainty, $8.26\%^1$, and fluctuate the signal efficiency using a Gaussian random generator with $\mu = \epsilon_{sig}$ and a σ equal to the systematic uncertainty, $8\% \times \epsilon_{sig}$. The same approach is used with the background estimate, N_{bkg} , and a set of toy MC experiments are done. The resulting smear in the branching fraction is then evaluated with the fluctuation in ϵ_{sig} , N_{bkg} or both. The central value of the branching fraction is evaluated according to equation 8.1. To evaluate any correlation, 3 sets of toy MC experiments are done, one where ϵ_{sig} is fluctuated and N_{bkg} is kept fixed, another where N_{bkg} is fluctuated and ϵ_{sig} is kept fixed, and finally one where both ϵ_{sig} and N_{bkg} are fluctuated using the Gaussian random generator. These sets of toy MC are done separately for a range of N_{sig} , from -15 to +15, where here $N_{sig} = N_{observed} - N_{bkg}$. For each value of N_{sig} , a 1000 toy MC experiments are made, where ϵ_{sig} and N_{bkg} are fluctuated using the Gaussian random generators. Furthermore, for each trial, the difference in the branching fraction is calculated

¹This value is the MLP cut systematic uncertainty which was initially determined to be around 8.26%. After further study, the value of 8.26% was found to be an overestimation.

as $\Delta BF = BF_{trial} - BF_{actual}$ where BF_{trial} is the central value of the branching fraction with a fluctuation in N_{bkg} , ϵ_{sig} , or both, and BF_{actual} is the central value of the branching fraction given a specific N_{sig} and the actual values of N_{bkg} and ϵ_{sig} without any fluctuation. A sample result, for a given $N_{sig} = 4$, is shown below. Fig. F.1 shows the smear in ΔBF when N_{bkg} is fluctuated, whereas Fig. F.2 displays ΔBF for the case where ϵ_{sig} is varied. As can be seen, the σ of the Gaussian fit to the branching fraction is larger when the background estimate is fluctuated. Furthermore, Fig. F.3 shows the smear in the branching fraction when both ϵ_{sig} and N_{bkg} are fluctuated, again for $N_{sig} = 4$. It can be readily seen that the value of the σ of the Gaussian fit to the branching fraction, the smear, when both are fluctuated has a value that is very close to that resulting from the distribution when only N_{bkg} is varied.

To quantify the correlation, Fig. F.4, F.5, F.6 show a plot of σ as a function of N_{sig} when N_{bkg} , ϵ_{sig} and both are fluctuated, respectively. The average value when N_{bkg} is fluctuated alone is 0.7794×10^{-3} . When only the ϵ_{sig} is fluctuated, the average value of σ is of order 10^{-5} . Finally, , when both ϵ_{sig} and N_{bkg} are fluctuated, the average smear in the branching fraction is 0.7525×10^{-3} . Thus, fluctuating N_{bkg} by its systematic uncertainty affects the branching fraction substantially more than a fluctuation in ϵ_{sig} . The effect of the correlation due to common systematic errors between the signal efficiency and the background estimate can be quantified using the difference in the branching fraction smear when both N_{bkg} and ϵ_{sig} are fluctuated to when N_{bkg} is fluctuated alone. The correlation between ϵ_{sig} and N_{bkg} causes an absolute smear in the branching fraction of 2.69×10^{-5} . This translates into a 0.28% systematic uncertainty on the background estimate and thus the effect of the correlation is negligible. Correlations between N_{bkg} and ϵ_{sig} , given equation 8.1, can be ignored for all systematic uncertainties evaluated in this analysis.

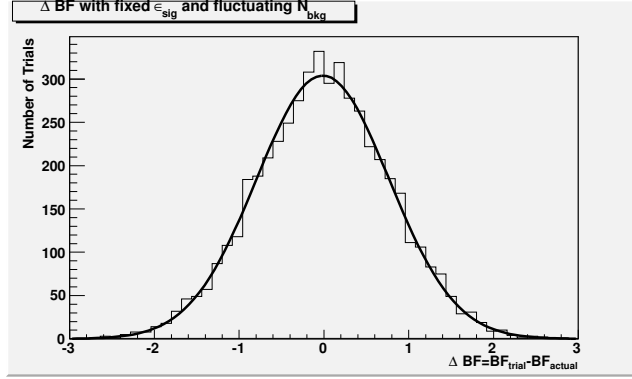


Figure F.1: Given $N_{sig} = 4$, the smear in ΔBF , when N_{bkg} is fluctuated using a Gaussian random generator with $\mu = N_{bkg}$ and $\sigma = 8.26\% \times N_{bkg}$. Here, $\Delta BF = BF_{trial} - BF_{actual}$, where BF_{trial} is the trial central value given a fluctuation in N_{bkg} and BF_{actual} is the actual value of the branching fraction for $N_{sig} = 4$.

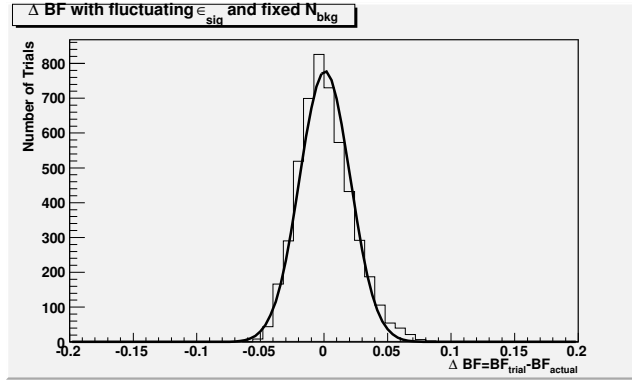


Figure F.2: Given $N_{sig} = 4$, the smear in ΔBF , when ϵ_{sig} is fluctuated using a Gaussian random generator with $\mu = \epsilon_{sig}$ and $\sigma = 8.26\% \times \epsilon_{sig}$.

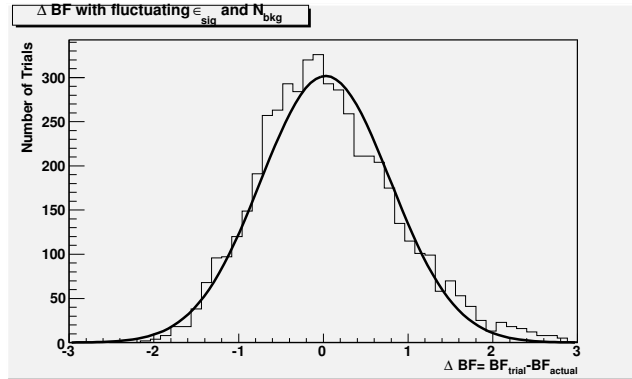


Figure F.3: Given $N_{sig} = 4$, the smear in ΔBF , when N_{bkg} and ϵ_{sig} are fluctuated.

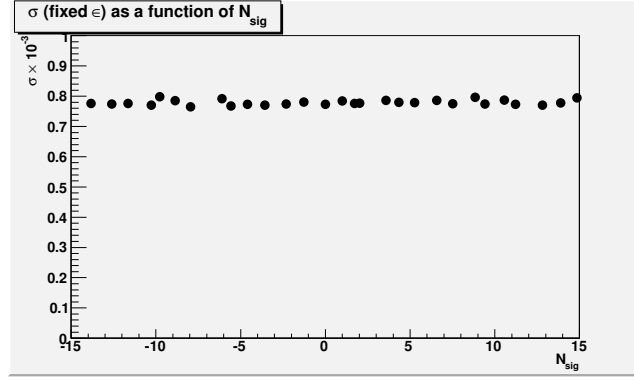


Figure F.4: σ as a function N_{sig} when N_{bkg} is fluctuated.

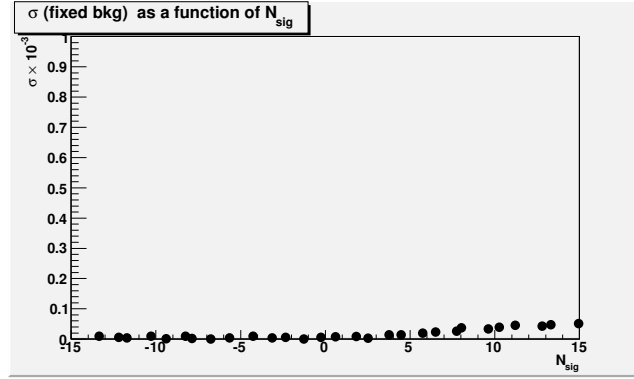


Figure F.5: σ as a function N_{sig} when ϵ_{sig} is fluctuated.

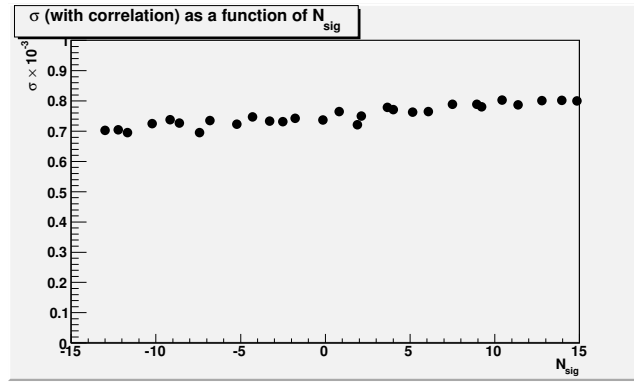


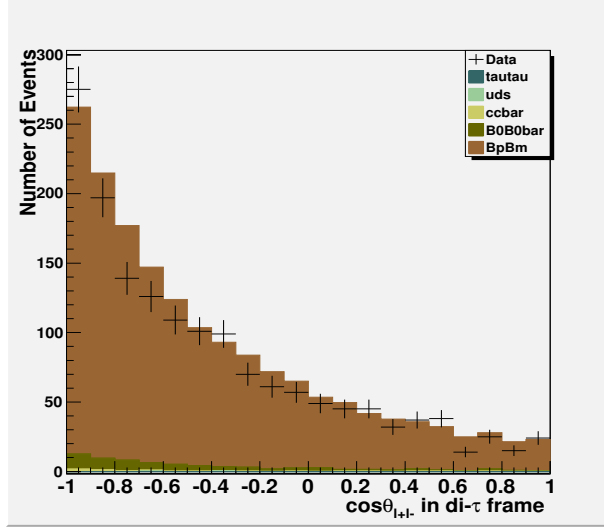
Figure F.6: σ as a function N_{sig} when both N_{bkg} and ϵ_{sig} are fluctuated.

Appendix G

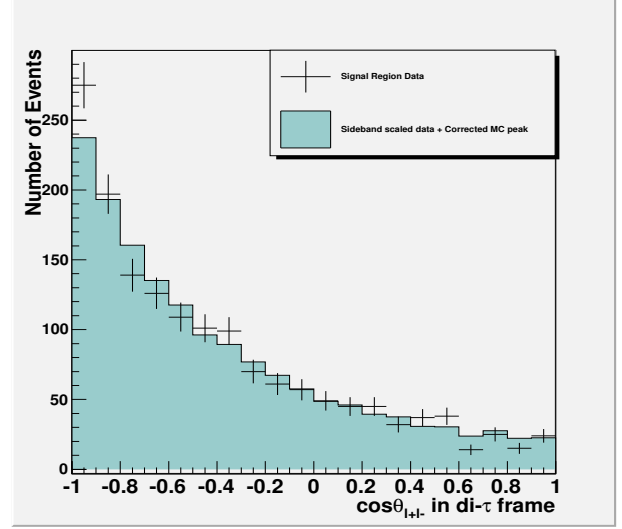
Discriminating Variables before and after MLP cut

In the following appendix, each discriminating variable entering into the MLP neural network is examined separately for the electron-muon mode. The goal is to ensure that there is no significant disagreement between data and MC for these variables, leading to the excess observed in that channel. The figures below show the distribution of these variables before and after the final MLP cut, with and without the m_{ES} sideband substitution. After looking closely, the conclusion is that there are no significant discrepancies between data and MC. While the agreement is not exact, the m_{ES} sideband substitution provides a significant improvement, specifically before applying the final MLP cut. However, after the MLP cut, low statistics in the data distribution make it hard to see a much improved agreement.

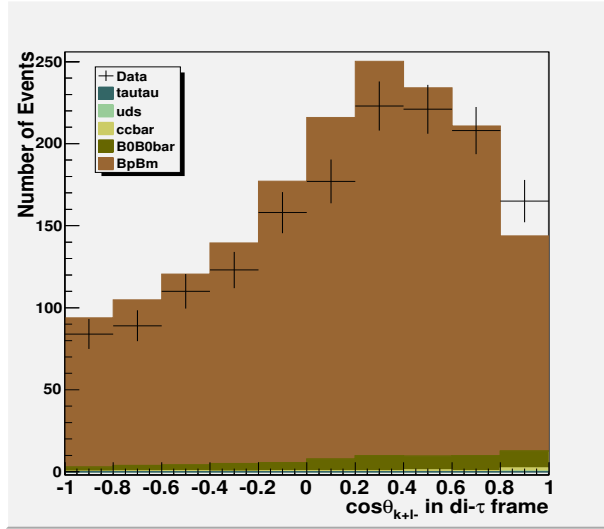
G.1 Discriminating variables before the final MLP cut



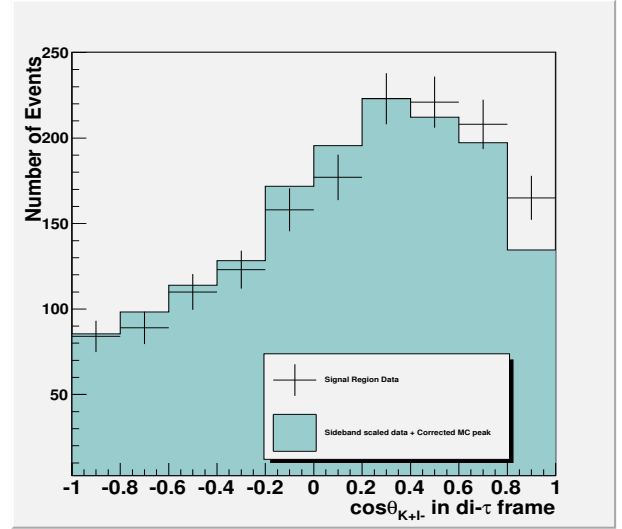
(a) $\cos \theta_{l+l-}$ in di- τ frame.



(b) $\cos \theta_{l+l-}$ in di- τ frame.

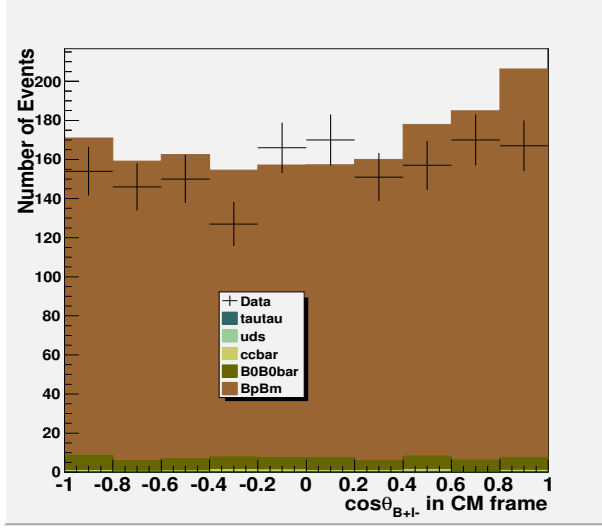


(c) $\cos \theta_{K+l-}$ in di- τ frame.

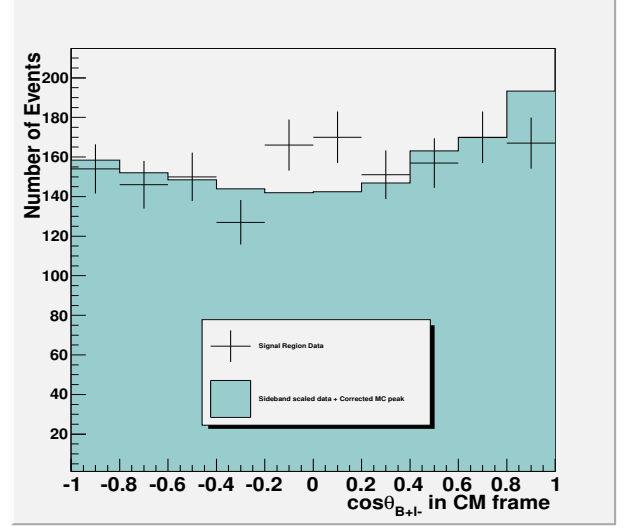


(d) $\cos \theta_{K+l-}$ in di- τ frame.

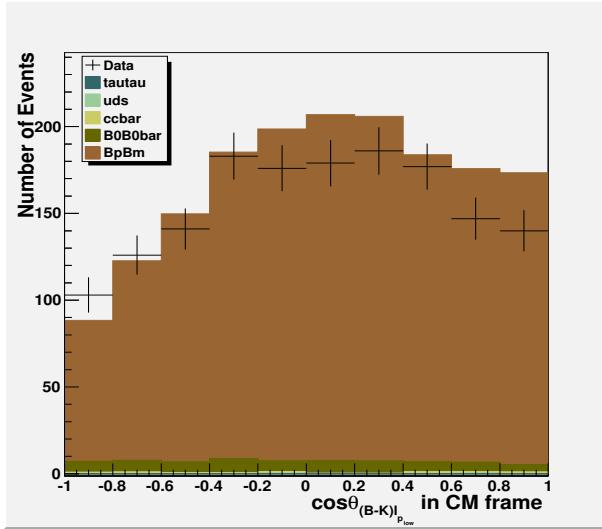
Figure G.1: Angular variables used in the MLP neural network before applying the final cut in the analysis, before (left) and after (right) m_{ES} sideband substitution.



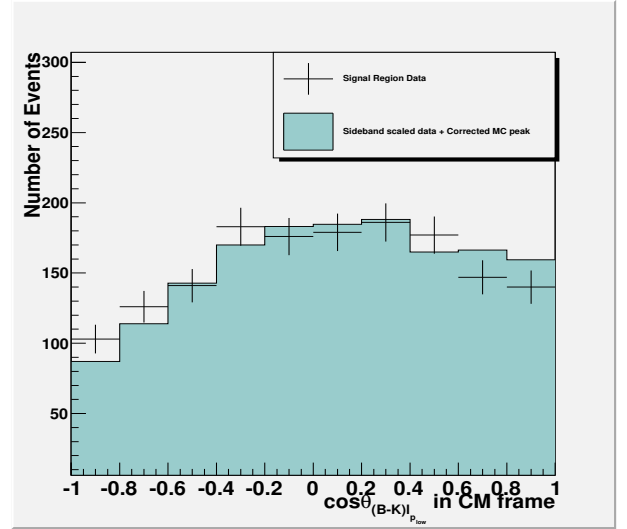
(a) $\cos \theta_{B+l-}$ in CM frame.



(b) $\cos \theta_{B+l-}$ in CM frame.

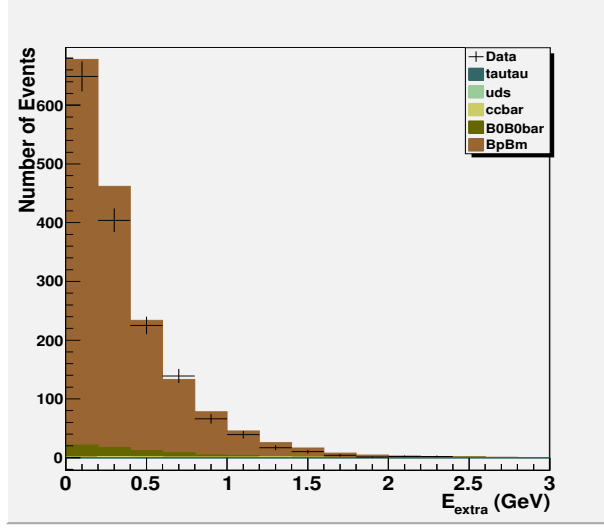


(c) $\cos \theta_{(B-K)l_{p_{low}}}$ in di- τ frame.

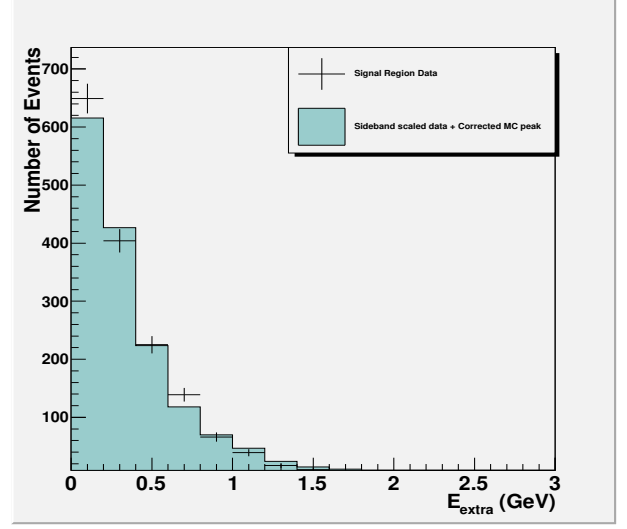


(d) $\cos \theta_{(B-K)l_{p_{low}}}$ in di- τ frame.

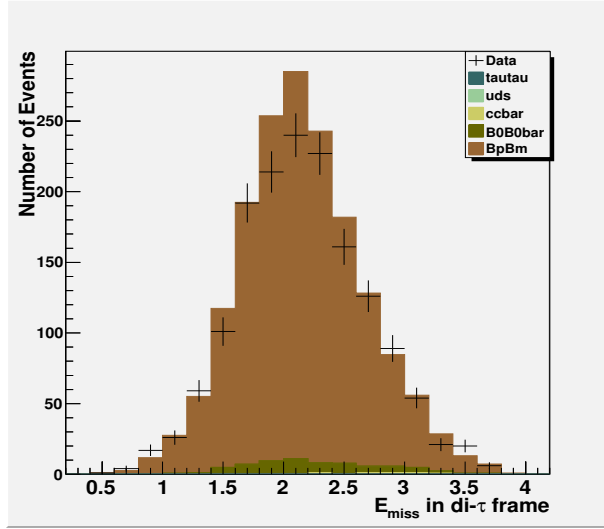
Figure G.2: Angular variables used in the MLP neural network before applying the final cut in the analysis, before(left) and after (right) m_{ES} sideband substitution.



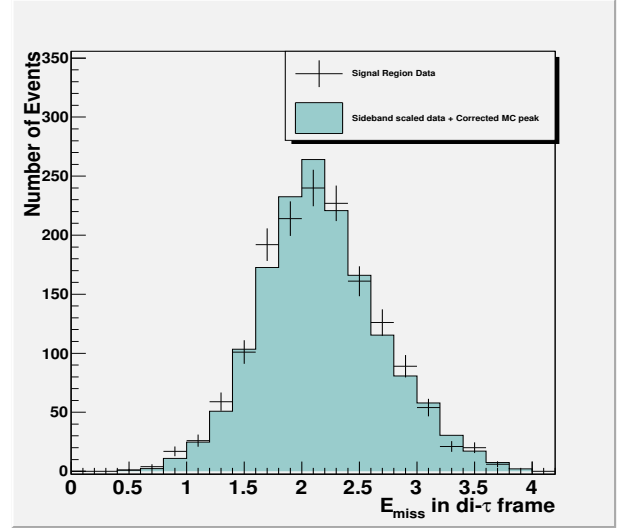
(a) E_{extra}



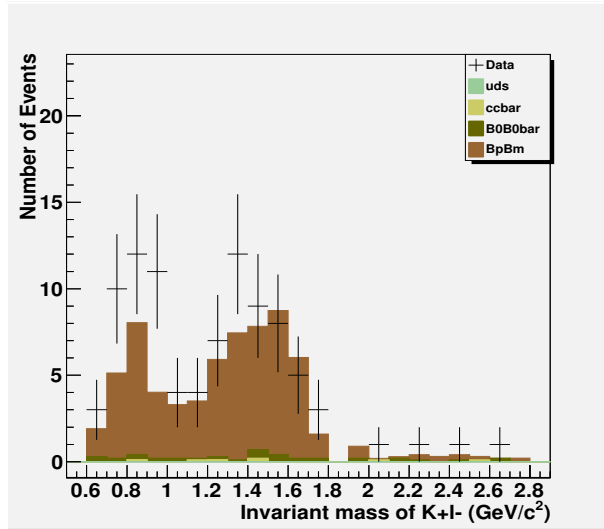
(b) E_{extra}



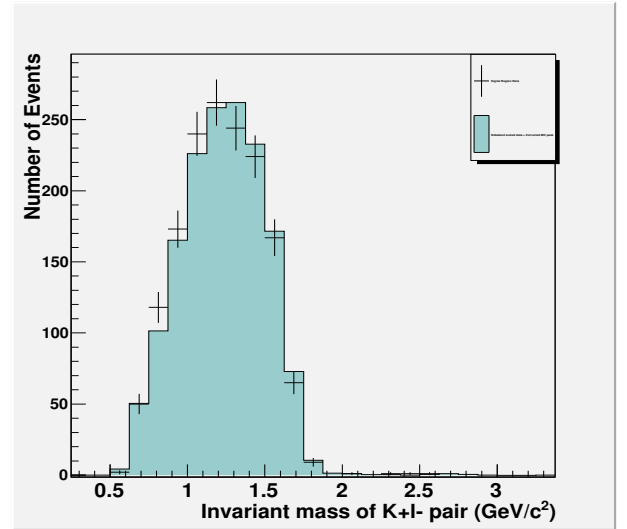
(c) E_{miss}



(d) E_{miss}

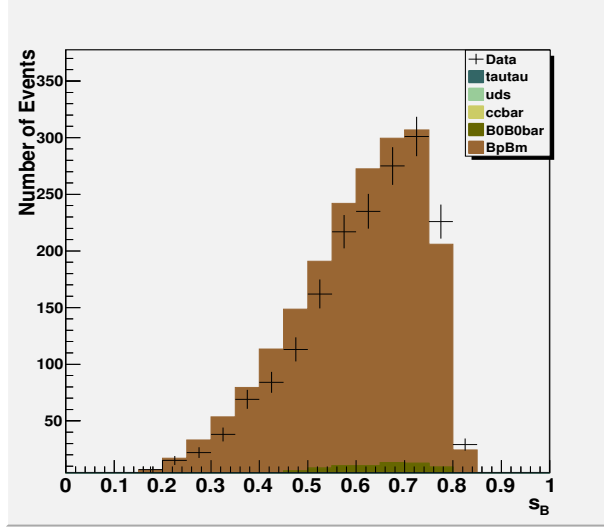


(e) Invariant mass of $K+l^-$ in CM frame.

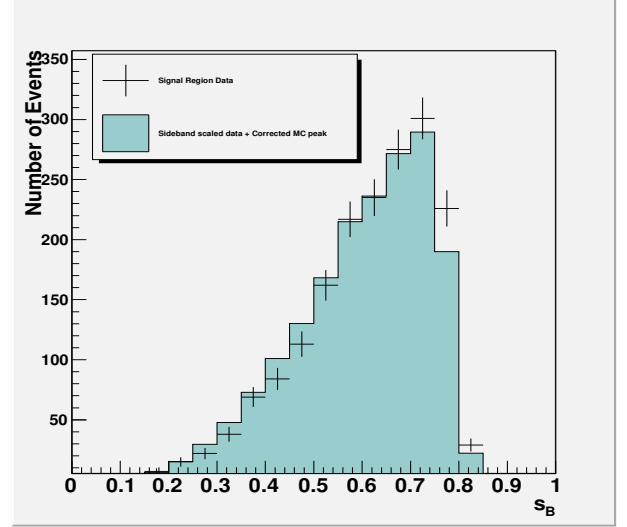


(f) Invariant mass of $K+l^-$ in CM frame.

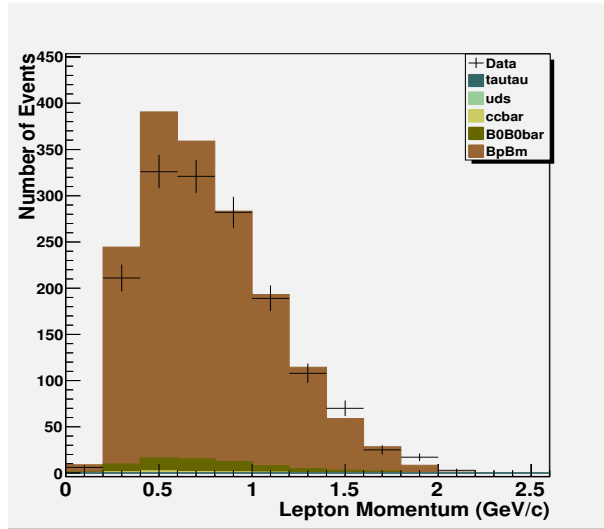
Figure G.3: Calorimeter variables used in the MLP neural network before applying the final cut in the analysis, before (left) and after (right) m_{ES} sideband substitution.



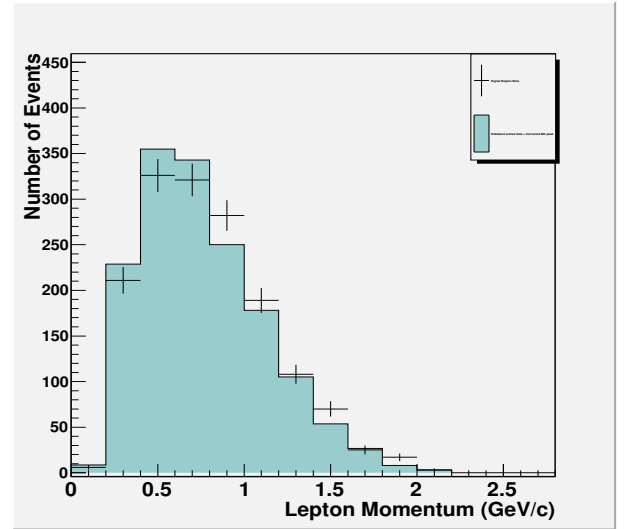
(a) s_B



(b) s_B



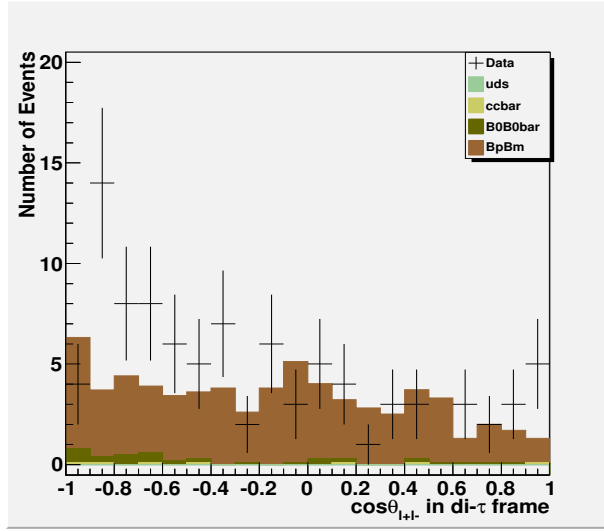
(c) Lepton Momentum



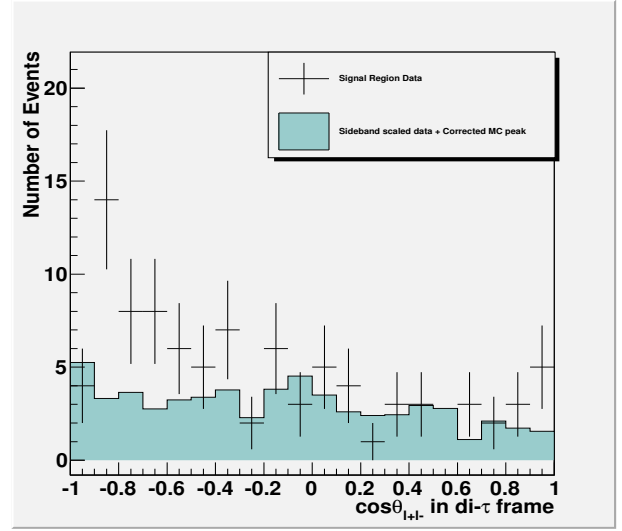
(d) Lepton Momentum

Figure G.4: Kinematic variables used in the MLP neural network before applying the final cut in the analysis, before (left) and after(right) m_{ES} sideband substitution.

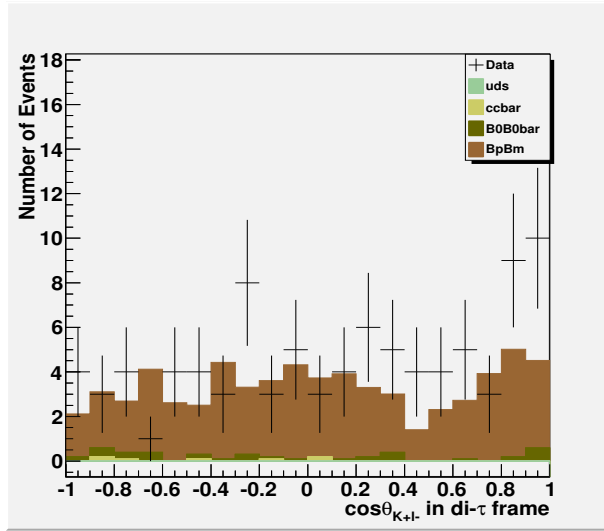
G.2 Discriminating variables after the final MLP cut



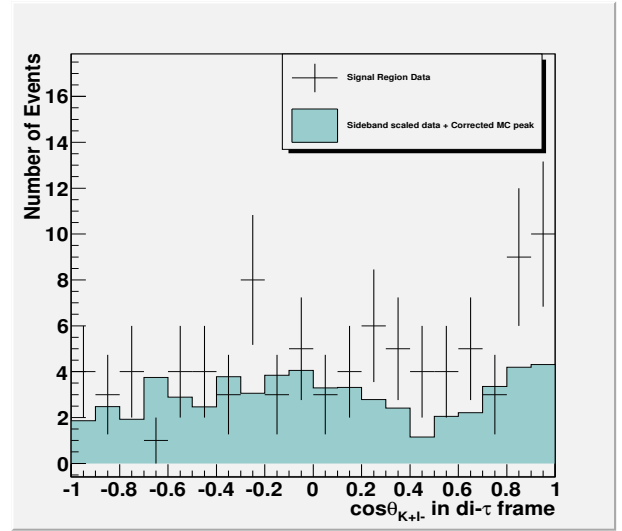
(a) $\cos \theta_{l+l-}$ in di- τ frame.



(b) $\cos \theta_{l+l-}$ in di- τ frame.

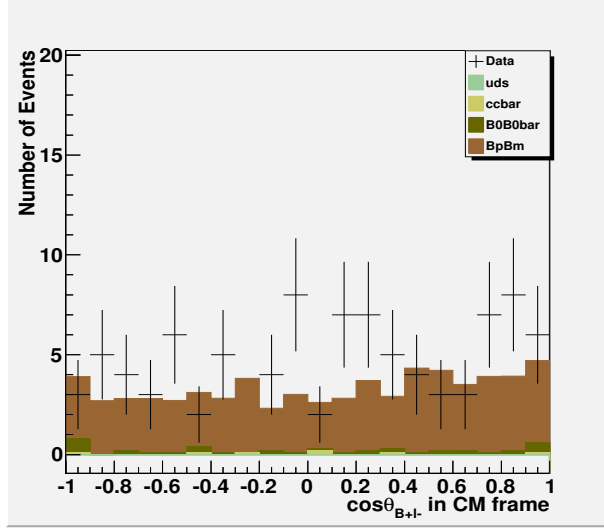


(c) $\cos \theta_{K+l-}$ in di- τ frame.

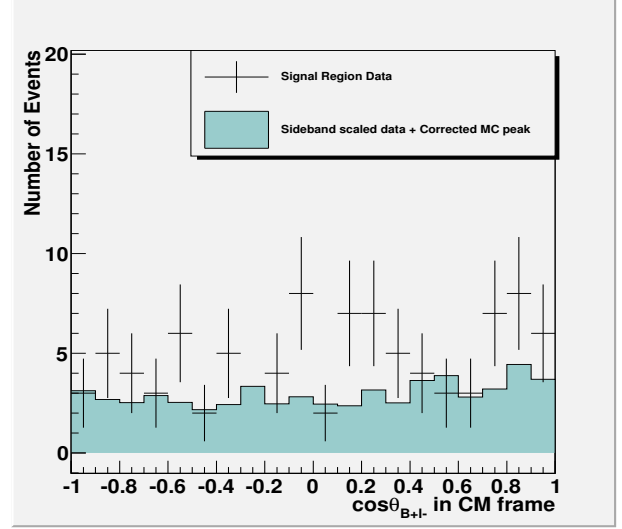


(d) $\cos \theta_{K+l-}$ in di- τ frame.

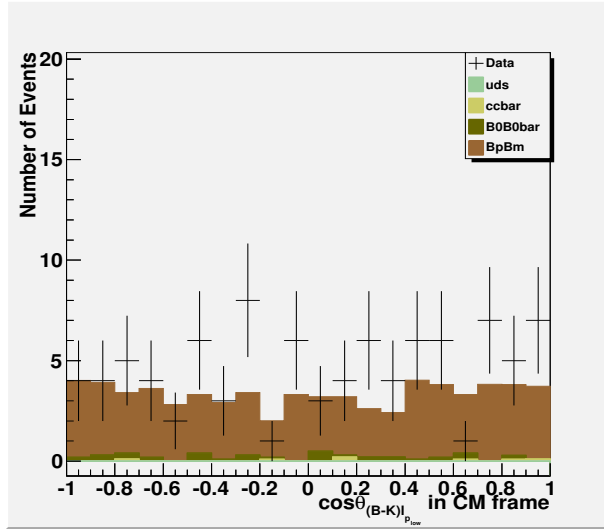
Figure G.5: Angular variables used in the MLP neural network after applying the final cut in the analysis, before (left) and after (right) m_{ES} sideband substitution.



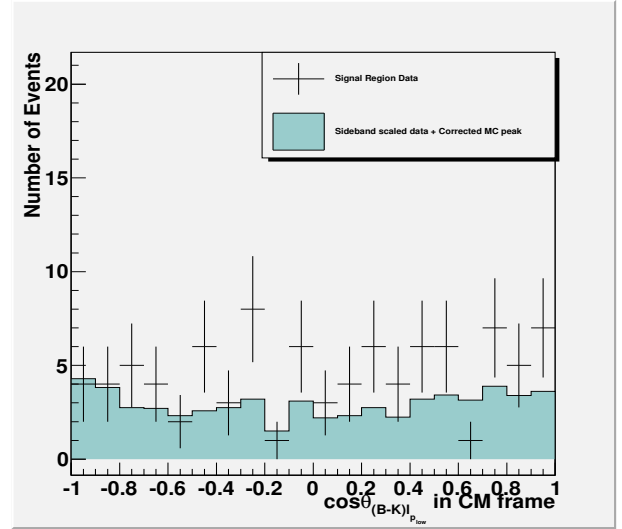
(a) $\cos \theta_{B+l-}$ in CM frame.



(b) $\cos \theta_{B+l-}$ in CM frame.

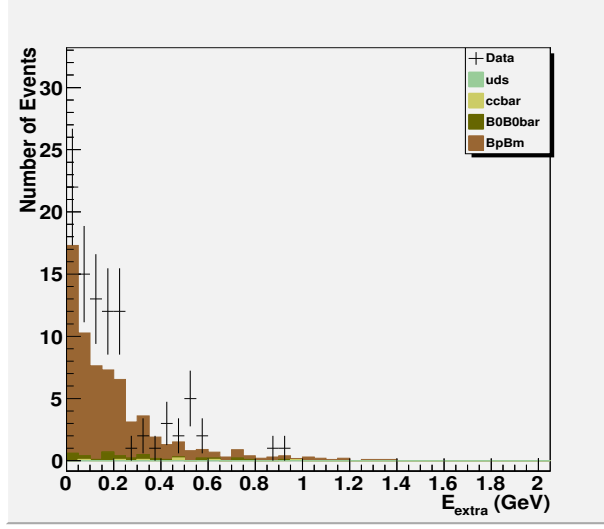


(c) $\cos \theta_{(B-K)l_{p_{low}}}$ in di- τ frame.

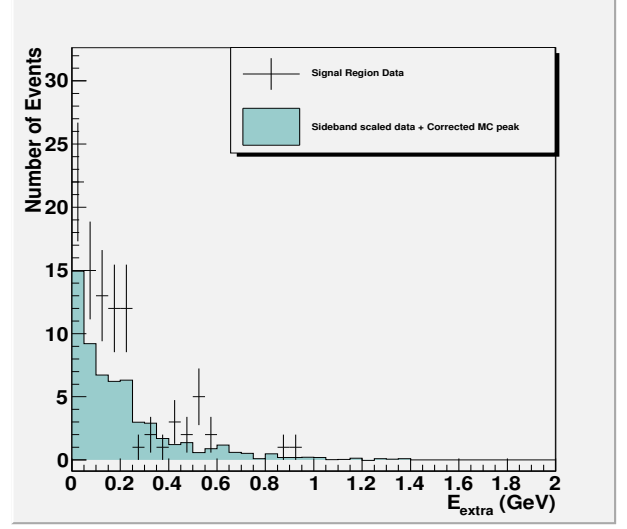


(d) $\cos \theta_{(B-K)l_{p_{low}}}$ in di- τ frame.

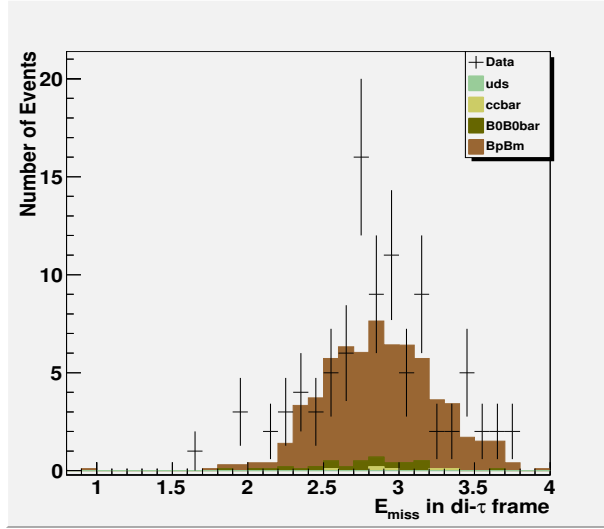
Figure G.6: Angular variables used in the MLP neural network after applying the final cut in the analysis, before(left) and after (right) m_{ES} sideband substitution.



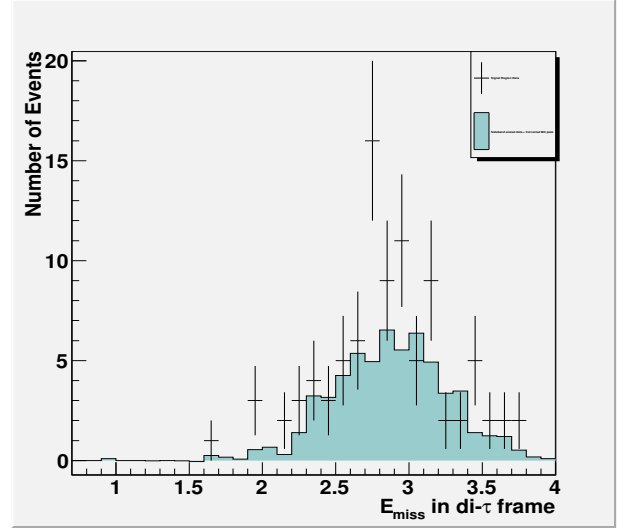
(a) E_{extra}



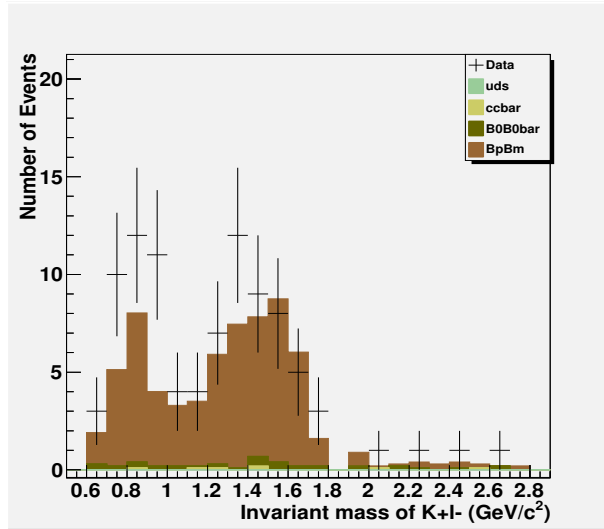
(b) E_{extra}



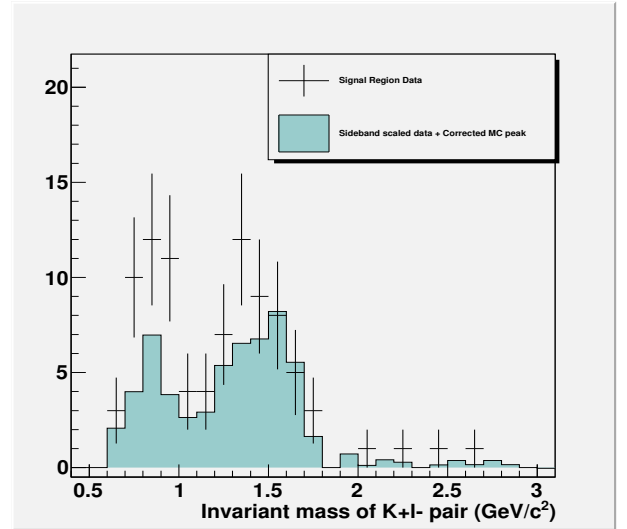
(c) E_{miss}



(d) E_{miss}

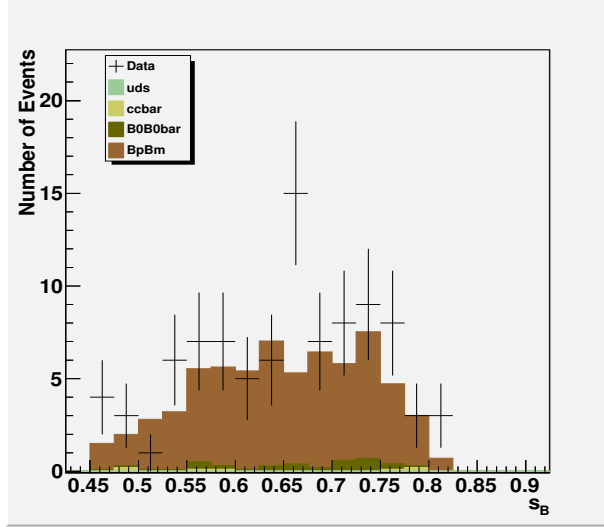


(e) Invariant mass of K+l- in CM frame.

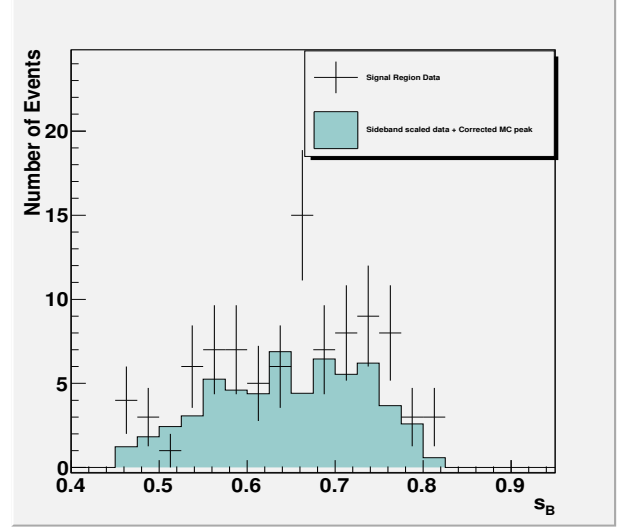


(f) Invariant mass of K+l- in CM frame.

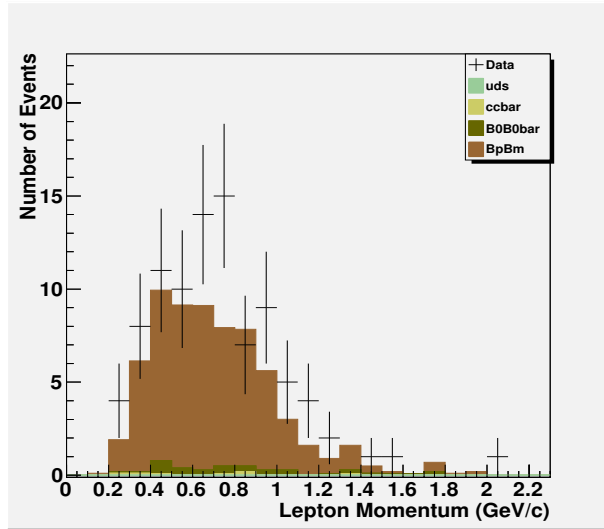
Figure G.7: Calorimeter variables used in the MLP neural network after applying the final cut in the analysis, before (left) and after (right) m_{ES} sideband substitution.



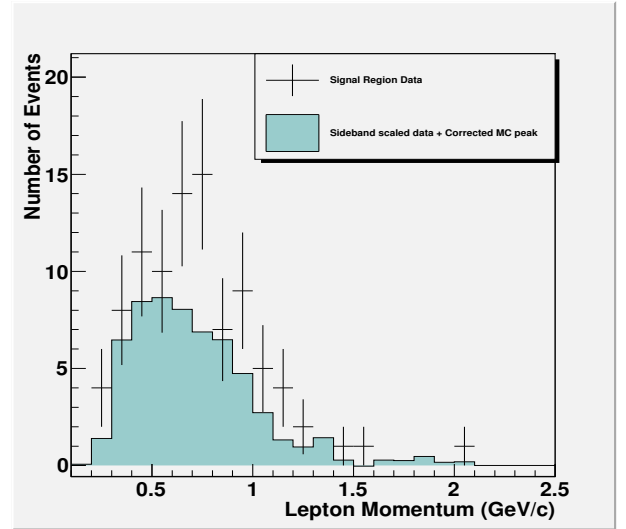
(a) s_B



(b) s_B



(c) Lepton Momentum



(d) Lepton Momentum

Figure G.8: Kinematic variables used in the MLP neural network after applying the final cut in the analysis, before (left) and after (right) m_{ES} sideband substitution.