

Measurement of Branching Fractions
and Mass Spectra in $B \rightarrow K\pi\pi\gamma$ Decays

Thesis by

Alexander Samuel

in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2006

(Defended June 13, 2005)

Copyright © 2006

Alexander Samuel

All rights reserved.

Acknowledgments

Thank you to the people who taught me particle physics: Melissa, Peter, Fotis, Hitoshi, Frank, David, and Anders. Also thanks to the *BABAR* Radiative Penguin Analysis Working Group, especially Wouter and Al, for politely informing me when I am doing something wrong.

Abstract

We present a measurement of the branching fractions of the exclusive radiative penguin processes $B \rightarrow K\pi\pi\gamma$ in a sample of 232 million $e^+e^- \rightarrow B\bar{B}$ decays recorded by the *BABAR* detector at the PEP-II asymmetric-energy storage ring. We reconstruct four final states: $K^+\pi^-\pi^+\gamma$, $K^+\pi^-\pi^0\gamma$, $K_S^0\pi^-\pi^+\gamma$, and $K_S^0\pi^+\pi^0\gamma$, where $K_S^0 \rightarrow \pi^+\pi^-$, in the range $m_{K\pi\pi} < 1.8 \text{ GeV}/c^2$. We measure the branching fractions $\mathcal{B}(B^+ \rightarrow K^+\pi^-\pi^+\gamma) = (2.95 \pm 0.13(\text{stat.}) \pm 0.19(\text{syst.})) \times 10^{-5}$, $\mathcal{B}(B^0 \rightarrow K^+\pi^-\pi^0\gamma) = (4.07 \pm 0.22(\text{stat.}) \pm 0.31(\text{syst.})) \times 10^{-5}$, $\mathcal{B}(B^0 \rightarrow K^0\pi^+\pi^-\gamma) = (1.85 \pm 0.21(\text{stat.}) \pm 0.12(\text{syst.})) \times 10^{-5}$, and $\mathcal{B}(B^+ \rightarrow K^0\pi^+\pi^0\gamma) = (4.56 \pm 0.42(\text{stat.}) \pm 0.30(\text{syst.})) \times 10^{-5}$. We also measure the distribution of $m_{K\pi\pi}$.

Contents

Acknowledgments	iii
Abstract	iv
Contents	v
List of Figures	viii
List of Tables	xii
1 Introduction	1
2 Background	5
2.1 Theoretical motivation	5
2.2 Previous measurements	8
3 PEP-II and the <i>BABAR</i> Detector	10
3.1 PEP-II	10
3.2 The <i>BABAR</i> detector	12
3.3 The silicon vertex tracker	15
3.4 The drift chamber	17
3.5 The Cherenkov detector	18
3.6 The calorimeter	20
3.7 Triggers and data acquisition	21
3.8 Particle selection criteria	23

4	Samples	27
4.1	Data samples	27
4.2	Monte Carlo samples	27
5	Candidate Selection	33
5.1	Skim	34
5.2	Sample selection	34
5.3	Cuts	35
5.4	Cut optimization	48
5.5	Fit region	53
5.6	Multiple candidate selection	55
5.7	Efficiency	59
5.8	Resolution	63
6	Background Processes	64
6.1	$b \rightarrow s\gamma$ background	65
6.2	Backgrounds from $B \rightarrow K\pi\pi\pi^0$ and $B \rightarrow K\pi\pi\eta$	71
6.3	Generic backgrounds	77
6.4	Total backgrounds	81
7	Control Samples	83
7.1	Off-resonance data	83
7.2	ΔE sidebands	86
7.3	$D\pi^\pm$ control samples	89
8	Fits	101

	vii	
8.1	Fit functions	102
8.2	Combined fit	116
8.3	$m_{K\pi\pi}$ distribution	121
9	Results	127
9.1	Branching fraction fit	127
9.2	$m_{K\pi\pi}$ distributions	131
10	Systematic Uncertainties	132
10.1	Branching fraction	132
10.2	Fit bias	135
10.3	Signal and crossfeed shape systematics	136
10.4	$b \rightarrow s\gamma$ model uncertainty	137
10.5	Peaking backgrounds	152
10.6	Combined systematics	153
11	Conclusions	154
	Bibliography	156

List of Figures

1	PEP-II schematic	10
2	PEP-II integrated luminosity	12
3	Side elevation of the <i>BABAR</i> detector	13
4	End elevation of the <i>BABAR</i> detector	14
5	SVT end elevation	15
6	SVT side elevation	16
7	DIRC schematic	19
8	DIRC K^\pm efficiency and π^\pm misidentification rate	20
9	Arrangement of crystals in the EMC	20
10	<i>GoodTracksLoose</i> efficiency distributions	24
11	<i>TightKaonMicroSelection</i> efficiency	25
12	<i>TightKaonMicroSelection</i> pion mistag rate	25
13	<i>LooseLHPionMicroSelection</i> efficiency	26
14	Distribution and cut on $m_{K\pi\pi}$	37
15	Distribution of best $m_{\gamma\gamma}$ for π^0 veto	38
16	Distribution of best $m_{\gamma\gamma}$ for η veto	39
17	Distribution of K_S^0 mass	40
18	Distribution of $K_S^0 \cos \theta_{\text{flight}}$	41
19	Distribution of K_S^0 decay length significance	41
20	Distribution of K_S^0 vertex probability	41
21	Distribution of $m_{\gamma\gamma}$ for π^0 candidates	42

22	Photon energy distribution from π^0 candidates	42
23	Distribution of CM energy of π^0 candidates	43
24	Distribution of vertex probability	44
25	Distribution of L_2/L_0	46
26	Distribution of $\cos \theta_B^*$	47
27	Distribution of $\cos \theta_T$	47
28	Distribution of Fisher discriminants	48
29	Efficiency in signal MC vs. $m_{K\pi\pi}$	62
30	$m_{K\pi\pi}$ resolution	63
31	m_{ES} and ΔE distributions in the $B \rightarrow K\pi\gamma$ cocktail MC	68
32	m_{ES} and ΔE distributions in the signal cocktail MC	69
33	m_{ES} and ΔE distributions in inclusive $b \rightarrow s\gamma$ MC	70
34	Background distributions in mode $K^+\pi^-\pi^0\gamma$ from $B^0 \rightarrow \bar{D}^0\pi^0$	72
35	Background distributions in mode $K^+\pi^-\pi^0\gamma$ from $B^0 \rightarrow \bar{D}^0\eta$	72
36	Background distributions in mode $K_S^0\pi^-\pi^+\gamma$ from $B^0 \rightarrow \bar{D}^0\pi^0$	73
37	Background distributions in mode $K_S^0\pi^-\pi^+\gamma$ from $B^0 \rightarrow \bar{D}^0\eta$	73
38	Background distributions in mode $K_S^0\pi^+\pi^0\gamma$ from $B \rightarrow D\rho$	74
39	Background distributions from $B \rightarrow K^*\rho$	76
40	m_{ES} and ΔE distribution of background candidates from $udsc$ MC	78
41	m_{ES} and ΔE distribution of background candidates from $B\bar{B}$ MC	79
42	$m_{K\pi\pi}$ distribution of background candidates from $udsc$ MC	80
43	$m_{K\pi\pi}$ distribution of background candidates from $B\bar{B}$ MC	80
44	Cumulative backgrounds from B decays	82
45	m_{ES} and ΔE distributions in off-resonance data and $udsc$ MC	84

46	$K\pi\pi$ distributions in off-resonance data and $udsc$ MC	85
47	Distributions in the low ΔE sideband for mode $K^+\pi^-\pi^+\gamma$	88
48	m_{ES} distributions in $D\pi^\pm$ control samples	91
49	ΔE distributions in $D\pi^\pm$ control samples	91
50	$m_{K\pi\pi}$ distributions in $D\pi^\pm$ control samples	92
51	Distributions in the $K^-\pi^+\pi^+\pi^-$ control sample	93
52	Distributions in the $K^+\pi^-\pi^0\pi^+$ control sample	94
53	Distributions in the $K_s^0\pi^-\pi^+\pi^+$ control sample	95
54	Distributions in the $K_s^0\pi^-\pi^0\pi^+$ control sample	96
55	Signal shape fits	105
56	Category 1 crossfeed shape fits	108
57	Category 2 crossfeed shape fits	109
58	Category 3 crossfeed shape fits	110
59	Category 4 crossfeed shape fits	111
60	Continuum shape fits	115
61	Projections of combined toy MC fit	120
62	$m_{K\pi\pi}$ sPlot distributions for a toy MC fit	123
63	Mean signal $m_{K\pi\pi}$ distributions from toy MC samples	124
64	Mean $udsc + B\bar{B}$ $m_{K\pi\pi}$ distributions from toy MC samples	125
65	Mean feed-down $m_{K\pi\pi}$ distributions from toy MC samples	125
66	Bias of sPlot $m_{K\pi\pi}$ distributions from toy MC samples	126
67	Likelihood ratio distribution in toy MC studies	128
68	m_{ES} and ΔE projections of on-resonance data	129
69	$m_{K\pi\pi}$ sPlot distributions in on-resonance data	131

70	Bias systematic uncertainty to $m_{K\pi\pi}$ distribution	136
71	Generated $m_{K\pi\pi}$ distributions in signal and $b \rightarrow s\gamma$ MC	139
72	m_{ES} distributions of crossfeed background	141
73	ΔE distributions of crossfeed background	142
74	$m_{K\pi\pi}$ distributions of crossfeed background	143
75	generated vs. reconstructed $m_{K\pi\pi}$ distributions in crossfeed background from signal MC	145
76	generated vs. reconstructed $m_{K\pi\pi}$ distributions in crossfeed background from $b \rightarrow s\gamma$	146
77	m_{ES} and ΔE distributions of feed-down background candidates	147
78	$m_{K\pi\pi}$ distributions of feed-down background candidates	148
79	$b \rightarrow s\gamma$ model systematic uncertainty to $m_{K\pi\pi}$ distribution	151
80	$m_{K\pi\pi}$ sPlot distributions in on-resonance data	155

List of Tables

1	Lowest-lying $J > 0$ kaon resonances	6
2	Selected predictions of $\mathcal{B}(B \rightarrow K_X \gamma)$	6
3	PEP-II parameters	11
4	Assumed resonance branching fractions	28
5	Assumed branching fractions of K_X resonances	29
6	MC exclusive background samples	32
7	exclusive $B \rightarrow D\pi^+$, $D \rightarrow K\pi\pi$ MC samples	32
8	Comparison of K^\pm PID selectors	52
9	Comparison of π^\pm PID selectors	52
10	Candidate multiplicity distribution in signal MC	56
11	Candidate multiplicity distribution in truth-matched signal MC	56
12	Descriptions of alternative MCS techniques.	57
13	Efficiency comparison of MCS techniques	58
14	Candidate selection cut efficiencies in signal MC	60
15	Marginal efficiency in signal MC	61
16	Cumulative efficiency in signal MC	61
17	Efficiency in signal MC by generated resonance	61
18	Efficiency in signal MC by run	61
19	Yields in $b \rightarrow s\gamma$ MC with $B \rightarrow K\pi\pi\gamma$ removed	66
20	Crossfeed yields by MC generation mode	67
21	$B \rightarrow D\rho$ background processes	74

22	$K^* \rho$ background processes	75
23	Background yields from $B \rightarrow K \pi \pi \pi^0$ and $B \rightarrow K \pi \pi \eta$	76
24	Peaking background yields from $B \rightarrow K \pi \pi \pi^0$ and $B \rightarrow K \pi \pi \eta$	77
25	$udsc$ and $B\bar{B}$ background yields	77
26	Expected signal and background yields	81
27	Yields in $udsc$ MC compared to off-resonance data	83
28	Yields in the low ΔE sideband	87
29	Yields in the high ΔE sideband	87
30	$B \rightarrow D \pi^\pm$, $D \rightarrow K \pi \pi$ branching fractions	89
31	Last-cut efficiencies in the $D \pi^\pm$ control sample	98
32	Shift in m_{ES} measured from $D \pi^\pm$ control samples	100
33	Signal shape parameters	105
34	Crossfeed background categories	107
35	Category 1 crossfeed shape parameters	108
36	Category 2 crossfeed shape parameters	109
37	Category 3 crossfeed shape parameters	110
38	Category 4 crossfeed shape parameters	111
39	Toy MC results with floating $B\bar{B}$ background component	114
40	Toy MC results with no $B\bar{B}$ background component	114
41	Continuum shape parameters	115
42	Fit parameter statistics from toy fits	121
43	Branching fraction bias estimates from toy fits	121
44	Parameters from fit to on-resonance data	127
45	Fit yields	130

46	Fit parameters in toy MC samples	135
47	Parameters from signal shape toy fits	136
48	Parameters from fit to data with shifted m_{ES}	137
49	Crossfeed backgrounds estimated from signal MC and inclusive $b \rightarrow s\gamma$ MC	140
50	Crossfeed backgrounds estimated from signal MC and normalized inclusive $b \rightarrow$ $s\gamma$ MC	140
51	Fit parameters in $K\pi\gamma$ toy MC studies	149
52	Fit parameters in signal shape toy MC studies	149
53	Fit parameters in crossfeed toy MC studies	150
54	Fit parameters in feed-down shape toy MC studies	150
55	Fit parameters in feed-down yield toy MC studies	150
56	Summary of branching fraction systematic uncertainties	153

1 Introduction

In the standard model (SM), the radiative penguin transition $b \rightarrow s\gamma$ proceeds via a weak amplitude. This process underlies decays $B \rightarrow X_s\gamma$, where X_s is a hadronic final state with strangeness $S = +1$ for B^+ or B^0 decays, $S = -1$ for B^- or \bar{B}^0 decays. Measurements of the $b \rightarrow s\gamma$ process can differ from the predictions of the SM weak interaction in the presence of beyond-SM physics, such as supersymmetry.

Radiative penguin decays of B mesons have previously been observed in reconstructed $K\pi\gamma$ and $K\pi\pi\gamma$ exclusive states, as well as in inclusive analyses in which the hadronic state is not reconstructed or is partially reconstructed. Among the $K\pi\pi\gamma$ final states, two of the six possible charge combinations, $K^+\pi^-\pi^+\gamma$ and $K^0\pi^+\pi^-\gamma$ (in the channel $K^0 \rightarrow K_s^0 \rightarrow \pi^+\pi^-$), have previously been observed.

In this analysis, we present new, more precise measurements of the branching fractions of the two previously observed decays. We also present a first observation of the decays $B^0 \rightarrow K^+\pi^-\pi^0\gamma$, and $B^+ \rightarrow K^0\pi^+\pi^0\gamma$ and measure these branching fractions, in the channel $K^0 \rightarrow K_s^0 \rightarrow \pi^+\pi^-$. In all four charge modes, we measure the invariant mass distributions $m_{K\pi\pi}$ of the hadronic system. We use in our measurements B mesons produced in $e^+e^- \rightarrow \Upsilon(4S) \rightarrow B\bar{B}$ reactions by the PEP-II collider and reconstructed by the *BABAR* detector.

In this analysis, we reconstruct four charge modes of the $B \rightarrow K\pi\pi\gamma$ process. These are,

- $K^+\pi^-\pi^+\gamma$
- $K^+\pi^-\pi^0\gamma$
- $K^0\pi^-\pi^+\gamma, K^0 \rightarrow K_S^0 \rightarrow \pi^+\pi^-$
- $K^0\pi^+\pi^z\gamma, K^0 \rightarrow K_S^0 \rightarrow \pi^+\pi^-$.

Throughout this analysis charge conjugate modes are implied. We do not reconstruct the isospin-related final states with two π^0 particles, $K^+\pi^0\pi^0\gamma$ and $K_S^0\pi^0\pi^0\gamma$. We also do not reconstruct K_L^0 final states or decays of K_S^0 other than to $\pi^+\pi^-$.

We reconstruct $B \rightarrow K\pi\pi\gamma$ candidate samples by combining charged tracks and neutral clusters detected in data events or Monte Carlo (MC) simulated events. We choose tracks and clusters that satisfy the electric charge and kinematic properties of the final states we seek to reconstruct. For each reconstructed B candidate, we compute the kinematic variables

$$\begin{aligned}\Delta E &= E_B^* - E_{\text{beam}}^* \\ m_{\text{ES}} &= \sqrt{E_{\text{beam}}^{*2} - \mathbf{p}_B^{*2}},\end{aligned}$$

where E_B^* and \mathbf{p}_B^* are the energy and momentum of the B candidate in the center of mass (CM) frame of the e^+e^- system, and E_{beam}^* is the CM energy of each beam. We expect correctly reconstructed candidates to satisfy $\Delta E \approx 0$ within measurement precision. We use m_{ES} instead of the invariant mass of the reconstructed B since the beam energy is better measured than the energy of the B , and because m_{ES} and ΔE are nearly uncorrelated. For correctly reconstructed candidates, we expect $m_{\text{ES}} \approx m_B$, with the precision dominated by the uncertainty in the beam energy.

Our sample of reconstructed B candidates will contain backgrounds from various sources. The largest of these consists of combinations of tracks from light-quark continuum events. The largest

contribution from $B\bar{B}$ events is from $b \rightarrow s\gamma$ processes, including a significant component of misreconstructed $B \rightarrow K\pi\pi\gamma$ decays. We impose selection criteria on the sample of reconstructed B candidates designed to remove misreconstructed and misidentified particle candidates and to reject events inconsistent with $B\bar{B}$ production processes. We also restrict our consideration to the range $m_{K\pi\pi} < 1.8 \text{ GeV}/c^2$, since we expect background processes to dominate our sample at higher masses. We optimize the selection criteria to maximize the expected precision of our branching fraction measurements, based on MC models and prior expectations of signal and background yields.

We measure the branching fraction by fitting the candidates' distribution in m_{ES} and ΔE to a probability density function that includes signal and background components. The shapes of signal components are determined from MC samples. The background components consist of contributions from continuum background, $b \rightarrow s\gamma$ processes, and other B decays. We fit all four reconstructed modes simultaneously, as this allows a consistent handling of correctly reconstructed and misreconstructed candidates from $B \rightarrow K\pi\pi\gamma$ decays.

We extract the $m_{K\pi\pi}$ distribution of correctly reconstructed $B \rightarrow K\pi\pi\gamma$ candidates in our data sample using a statistical technique for disentangling the distribution of one variable in a single signal component of a maximum-likelihood fit with signal and background components to another set of variables.

In this analysis, we choose to present measurements of the $m_{K\pi\pi}$ spectrum for all decays we reconstruct in each charge mode, which is largely free of model uncertainty (though we do rely on models of inclusive and exclusive $b \rightarrow s\gamma$ decays to model backgrounds), instead of attempting to determine branching fractions $B \rightarrow K_X\gamma$ for specific K_X resonances. Disentangling the resonance structure would require careful modeling of amplitudes and relative phases of multiple interfering processes, which is beyond the statistical power of our current dataset. A correct model must also include proper treatment the decays of the coherent K_X resonances, not all of which are well-

measured. We expect that larger B factory datasets available in the future will enable a Dalitz analysis of the resonance structure as well as measurement of the photon polarization.

We used a blind methodology, in which we formulate the selection and fit procedures using MC simulations of our signal and background processes, and data control samples distinct from the data that contains our signal candidates. Only when we have finalized and validated our analysis procedure (except, in a few instances, the procedures for estimating systematic uncertainties) do we examine candidates passing our signal selection criteria and produce our results. The blind methodology is intended to reduce inadvertent bias in the choice of selection criteria and fit procedure.

2 Background

2.1 Theoretical motivation

In the SM, couplings of quarks of different families are mediated by the weak interaction. The SM weak interaction does not predict tree-level flavor changing neutral current processes, such as decays of b quarks to s quarks. However, such reactions can occur in one-loop amplitudes, such as *radiative penguin* decays of b quarks, $b \rightarrow s\gamma$, in which a W boson is emitted and reabsorbed by the quark line. Inside the loop, u , c , and t quarks can contribute; the latter is the dominant term. These radiative penguin decays can provide sensitive tests of the SM (see, for example, [1]), as new particles can contribute loop diagrams at the same order of perturbation theory as the lowest-order SM process.

In the decay of a B meson, the s quark produced in this process, combined with the spectator quark, produces a hadronic system of one or more particles. The decay may proceed through resonant or non-resonant amplitudes are possible. The hadronic system recoiling against the real photon must satisfy $J > 0$, which excludes the decay $B \rightarrow K\gamma$. Decays through higher kaon resonances are possible. The six lowest-lying $J > 0$ kaon resonances and their principle decays are listed in the table below.

Resonance	J^P	Mass (MeV/ c^2)	Width (MeV/ c^2)	Decay	Branching frac.
$K^*(892)$	1^-	$\begin{cases} 891.66 \pm 0.26 (K^{*+}) \\ 896.10 \pm 0.27 (K^{*0}) \end{cases}$	$\begin{cases} 50.8 \pm 0.9 (K^{*+}) \\ 50.7 \pm 0.6 (K^{*0}) \end{cases}$	$K\pi$	$\sim 100\%$
$K_1(1270)$	1^+	1273 ± 7	90 ± 20	$K\rho$ $K_0^*(1430)\pi$ $K^*\pi$ $K\omega$	$(42 \pm 6)\%$ $(28 \pm 4)\%$ $(16 \pm 5)\%$ $(11 \pm 2)\%$
$K_1(1400)$	1^+	1402 ± 7	174 ± 13	$K^*\pi$ $K\rho$	$(94 \pm 6)\%$ $(3 \pm 3)\%$
$K^*(1410)$	1^-	1414 ± 15	232 ± 21	$K^*\pi$ $K\pi$ $K\rho$	$> 40\%$ $(6.6 \pm 1.3)\%$ $< 7\%$
$K_2^*(1430)$	2^+	$\begin{cases} 1425.6 \pm 1.5 (K_2^{*+}) \\ 1432.4 \pm 1.3 (K_2^{*0}) \end{cases}$	$\begin{cases} 98.5 \pm 2.7 (K_2^{*+}) \\ 109 \pm 5 (K_2^{*0}) \end{cases}$	$K\pi$ $K^*\pi$ $K^*\pi\pi$ $K\rho$	$(49.9 \pm 1.2)\%$ $(24.7 \pm 1.5)\%$ $(13.4 \pm 2.2)\%$ $(8.7 \pm 0.8)\%$
$K^*(1680)$	1^-	1717 ± 27	332 ± 110	$K\pi$ $K\rho$ $K^*\pi$	$(38.7 \pm 2.5)\%$ $(31.4_{-2.1}^{+4.7})\%$ $(29.9_{-4.7}^{+2.2})\%$

TABLE 1. Properties and principle decays of the six lowest-lying $J > 0$ resonances [2]. Limits are at 90% confidence level.

The branching fractions of decay of B mesons to $K_X\gamma$ exclusive final states are not well predicted, and are the subject of ongoing theoretical investigation. A selection of theoretical predictions of $B \rightarrow K_X\gamma$ branching fractions is shown in the table below.

Source	$B \rightarrow K^*\gamma$	$B \rightarrow K_1(1270)\gamma$	$B \rightarrow K_1(1400)\gamma$	$B \rightarrow K_2^*(1430)\gamma$
Cheng and Chua (2004) [3]	3.27 ± 0.74	0.02 to 0.84	0.003 to 0.80	1.48 ± 0.30
Ebert et al. [4]	4.5 ± 1.5	0.45 ± 0.15	0.78 ± 0.18	1.7 ± 0.6
Safir [5]	5.81 ± 2.27	0.67 ± 0.27	0.30 ± 0.13	1.67 ± 0.67
Veseli and Olsson [6]	9.99 ± 3.81	1.44 ± 0.53	0.70 ± 0.30	2.07 ± 0.97

TABLE 2. Selected predictions of $\mathcal{B}(B \rightarrow K_X\gamma)$ in units of 10^{-5} , drawn from [3]. The predictions of Cheng and Chua depend on the choice of the $K_1(1270) - K_1(1400)$ mixing angle.

In SM $b \rightarrow s\gamma$ decays, the parity-violating coupling of the W produces an s quark that is approximately left-handed, up to $\mathcal{O}(m_s/m_b)$, so the recoiling photon is also approximately left-handed. In $\bar{b} \rightarrow \bar{s}\gamma$, the photon is approximately right-handed. Measurement of the photon polarization would be a strong test of the SM, since non-SM processes can introduce diagrams with different polarization to the decay.

The polarization of the photon cannot be measured directly, nor can it be inferred from a recoiling $K\pi$ hadronic system. Gronau et al. have shown that the photon polarization can, however, be measured in $B \rightarrow K^+\pi^-\pi^0\gamma$ and $B \rightarrow K_s^0\pi^+\pi^0\gamma$ decays [7]. In these decays, interference between $K^{*0}\pi^0$ and $K^{*+}\pi^-$ processes or between $K^{*+}\pi^0$ and $K^{*0}\pi^+$ processes can produce decay distributions sensitive to the photon polarization. The overall decay rate does not depend on the photon polarization, but the decay rate variation with θ_d , the angle between the normal to the $K\pi\pi$ decay plane and the photon direction in the $K\pi\pi$ center of mass frame, is related to the polarization.

In this analysis, we undertake to observe the previously unobserved decays $B \rightarrow K^+\pi^-\pi^0\gamma$ and $B \rightarrow K_s^0\pi^+\pi^0\gamma$, as well as to produce improved measurements of the branching fractions of $B \rightarrow K^+\pi^-\pi^+\gamma$ and $B \rightarrow K_s^0\pi^-\pi^+\gamma$. Observation of the two $\pi^\pm\pi^0$ modes is the first step to measuring the photon polarization in $K\pi\pi\gamma$ decays. In addition, we measure the $m_{K\pi\pi}$ distributions in these decays, which provide information about the resonance structure.

2.2 Previous measurements

Radiative penguin B decays have been observed in exclusive two-body final states, where X_s is $K^+\pi^-$ [8, 9] or $K_s^0\pi^0$ [10], in decays $B \rightarrow K^*\gamma$. The rates and kinematics of these decays are governed not only by the weak interaction, but also by the QCD physics by which the s quark and the spectator quark hadronize to form the X_s system. In the case of decays to $K\pi\gamma$ final states, the hadronic part can decay through resonances. In addition to $B \rightarrow K^*\gamma$ decays, the $K\pi\gamma$ final state has been observed in decays of the $K_2^*(1430)$ resonance as well [11, 12]. The most recent *BABAR* measurements of these branching fractions are,

$$\mathcal{B}(B^+ \rightarrow K^*(892)^+\gamma) = (3.87 \pm 0.28 \pm 0.26) \times 10^{-5}$$

$$\mathcal{B}(B^0 \rightarrow K^*(892)^0\gamma) = (3.92 \pm 0.20 \pm 0.24) \times 10^{-5}$$

$$\mathcal{B}(B^+ \rightarrow K_2^*(1430)^+\gamma) = (1.45 \pm 0.40 \pm 0.15) \times 10^{-5}$$

$$\mathcal{B}(B^0 \rightarrow K_2^*(1430)^0\gamma) = (1.22 \pm 0.25 \pm 0.10) \times 10^{-5},$$

where the first uncertainty in each measurement is statistical and the second is systematic.

Decays to two exclusive three-body final states, $B \rightarrow K^+\pi^-\pi^+\gamma$ and $B \rightarrow K_s^0\pi^-\pi^+\gamma$ have also been observed [13], by the Belle Collaboration. The measured branching fractions are,

$$\mathcal{B}(B^+ \rightarrow K^+\pi^+\pi^-\gamma) = (2.50 \pm 0.18 \pm 0.22) \times 10^{-5}$$

$$\mathcal{B}(B^0 \rightarrow K^0\pi^+\pi^-\gamma) = (2.4 \pm 0.4 \pm 0.3) \times 10^{-5},$$

where the first uncertainty is statistical and the second is systematic. The second process was measured using $K^0 \rightarrow K_s^0 \rightarrow \pi^+\pi^-$ decays. Decays to $K\pi\pi\gamma$ can also display interesting hadronic structure. There are five kaon resonances with spin of at least one that decay to $K\pi\pi$ and contribute in the mass range below $1.8 \text{ GeV}/c^2$. The decays of these resonances themselves exhibit resonance structure, in $K^*\pi$, $K\rho$, and $K_0^*(1430)\pi$ combinations. By selecting specific secondary

resonance decays, Belle claims measurements or upper limits of branching fractions to specific K_X resonances:

$$\mathcal{B}(B^+ \rightarrow K_1(1270)^+ \gamma) = (4.3 \pm 0.9 \pm 0.9) \times 10^{-5}$$

$$\mathcal{B}(B^0 \rightarrow K_1(1270)^0 \gamma) < 5.8 \times 10^{-5}$$

$$\mathcal{B}(B^+ \rightarrow K_1(1400)^+ \gamma) < 1.5 \times 10^{-5}$$

$$\mathcal{B}(B^0 \rightarrow K_1(1400)^0 \gamma) < 1.2 \times 10^{-5},$$

where the upper limits are at a 90% confidence level. These measurements are based on the model assumption about the K_X decays, in particular that interference among processes decaying to the same final state can be neglected.

The radiative penguin process has also been observed by inclusive measurements, in which the hadronic part X_s is not reconstructed or is partially constructed. The latter technique is known as the *semi-inclusive* analysis, in which as many exclusive X_s final states as possible are reconstructed and combined to approximate an inclusive measurement. The current world average [2] branching fraction of the inclusive process is ,

$$\mathcal{B}(b \rightarrow s \gamma) = (3.3 \pm 0.4) \times 10^{-4}.$$

Interpretation of the results of a semi-inclusive measurement depends on understanding the exclusive decays $B \rightarrow X_s \gamma$, which this analysis aims to improve.

3 PEP-II and the *BABAR* Detector

In this chapter, we present a brief description of the PEP-II collider and of the construction and performance of the *BABAR* detector components used in this analysis.

3.1 PEP-II

PEP-II [14] is an asymmetric-energy e^+e^- collider at the Stanford Linear Accelerator Center designed for high-luminosity production of $B\bar{B}$ pairs in a moving center of mass frame. It consists of two storage rings, the *high-energy ring* (HER) storing 9 GeV electrons, and the *low energy ring* (LER) storing 3.1 GeV positrons, each with a circumference of 2200 m. The storage rings are hexagonal, with a single interaction region occupied by the *BABAR* detector. Particles are injected into both from the preexisting 3 km linear accelerator. The figure below shows a schematic of PEP-II.

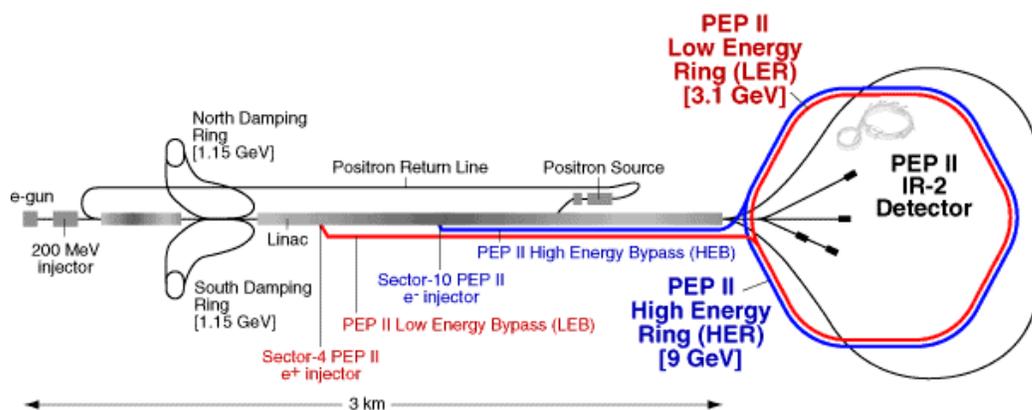


FIGURE 1. Schematic illustration of the PEP-II storage ring and the linear accelerator that injects it. The HER and its injection line are shown in blue. The LER and its injection line are shown in red.

The two beam energies are chosen to produce a center of mass (CM) energy of $\sqrt{s} = 10.58$ GeV,

the nominal mass of the $\Upsilon(4S)$ resonance. At this energy, the $b\bar{b}$ cross section is approximately 1.05 nb, while the cross section for light-quark continuum production ($u\bar{u}$, $d\bar{d}$, $s\bar{s}$, and $c\bar{c}$, collectively known as $udsc$) is about 3.4 nb; 0.94 nb for $\tau^+\tau^-$ production; 1.16 nb for $\mu^+\mu^-$ production; and about 40 nb for e^+e^- elastic (Bhabha) scattering [15]. The collider is run about 20 MeV below the $\Upsilon(4S)$ resonance for a fraction of data taking, to produce data samples without $b\bar{b}$ events for background studies.

The $\Upsilon(4S)$ decays to a $B\bar{B}$ pair, with available momentum $p = 335 \text{ MeV}/c$ in the CM frame. Due to the asymmetry of the beam energies in PEP-II, the e^+e^- system is boosted with $\beta\gamma = 0.56$ in the lab frame. The boost is chosen to increase the typical distance between the B decay vertices to $\beta\gamma c\tau \sim 250 \text{ }\mu\text{m}$, which can be measured by *BABAR*'s silicon tracking detector.

The following table lists PEP-II parameters as of June 2004 [16], the end of the period in which data used in this analysis were obtained.

Parameter	LER	HER
energy	3.1 GeV	9.0 GeV
number of bunches	1588	1588
horizontal beam size σ_x	170 μm	170 μm
vertical beam size σ_y	7.2 μm	7.2 μm
bunch length σ_z	13 mm	13 mm
horizontal beta at IP β_x^*	32 cm	32 cm
vertical beta at IP β_y^*	10.55 mm	10.5 mm
tune shift ξ_x/ξ_y	0.053/0.064	0.055/0.046
current	2.45 A	1.55 A
crossing angle	0 mrad	
luminosity	$9.21 \times 10^{33}/\text{cm}^2/\text{s}$	

TABLE 3. Typical operating parameters of the PEP-II storage ring as of June 2004.

In the period up to summer of 2004, in which the data used in this analysis were recorded, the record PEP-II luminosity was $9.21 \times 10^{33}/\text{cm}^2/\text{sec}$. The record daily integrated luminosity collected by the *BABAR* experiment (incorporating data taking efficiency) was 681.08 pb^{-1} . The following figure shows the cumulative integrated luminosity delivered by PEP-II and recorded by *BABAR* over this data-taking period.

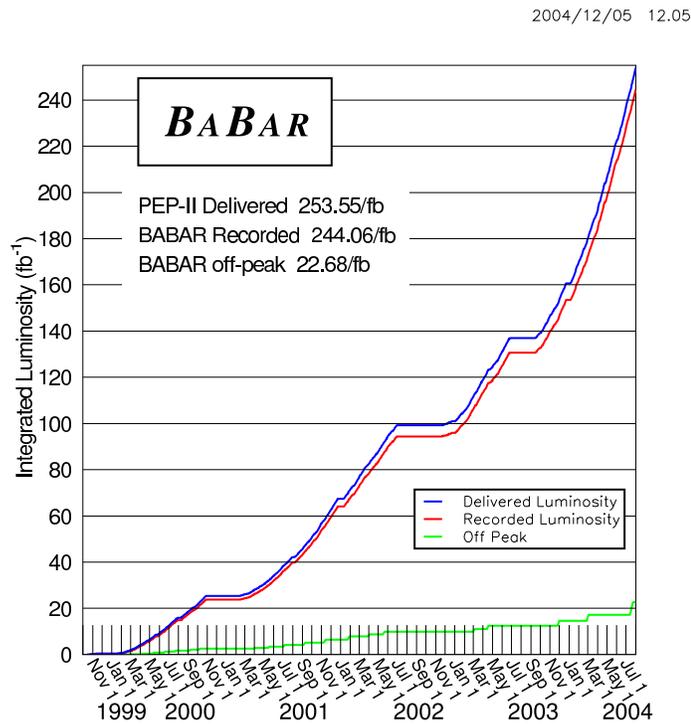


FIGURE 2. Cumulative integrated luminosity delivered by PEP-II (blue) and recorded by *BABAR* (red) through June 2004. The off peak integrated luminosity is shown in green.

3.2 The *BABAR* detector

The *BABAR* detector [17] is a general purpose particle physics detector installed at the PEP-II. It was designed primarily to measure CP violation in the decays of B mesons and perform high-precision tests of the Standard Model weak mixing matrix, but is capable of a wide range of measurements in B , charm, and τ physics.

The *BABAR* detector, from the inside out, consists of a silicon vertex tracker (SVT) for measurement of track angles and precise location of decay vertices; a drift chamber (DCH) for measurement of charged track momenta, a detector of internally-reflected Cherenkov light (DIRC) for identifying charged particles; a CsI(Tl) crystal calorimeter (EMC) for measuring the energy of photons and other neutral particles; a 1.5 T superconducting solenoid; and an instrumented steel flux return interspersed with resistive plate chambers for detecting muons and other weakly interacting long-lived

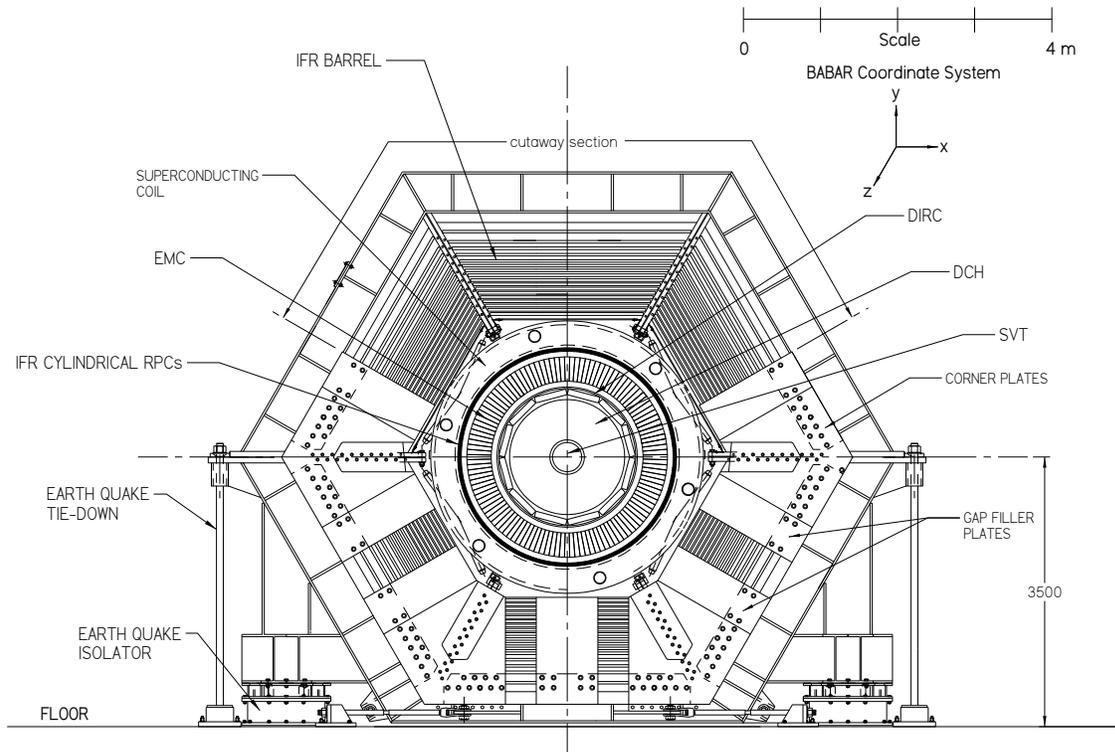


FIGURE 4. Elevation of the *BABAR* detector in section perpendicular to the beam line.

3.3 The silicon vertex tracker

The SVT is constructed of five layers of double-sided silicon microstrip sensors. On one side of each sensor, strips oriented parallel to the beam measure ϕ , while on the other side, strips oriented transversely to the beam measure z . The sensors are $300\ \mu\text{m}$ thick; strip pitch varies from $50\ \mu\text{m}$ to $210\ \mu\text{m}$. Position resolution is improved by interpolating among energy deposits on adjacent strips.

The innermost three layers, at radii of 32 mm, 40 mm, and 54 mm from the beam, are composed of six circumferential segments. The segments are pitched slightly to provide overlapping coverage at their ends. The outermost two layers are arranged each as a staggered pair of layers with slightly different radii. The radius of the outermost layer is 144 mm. The following figure illustrates the configuration of the five SVT layers.

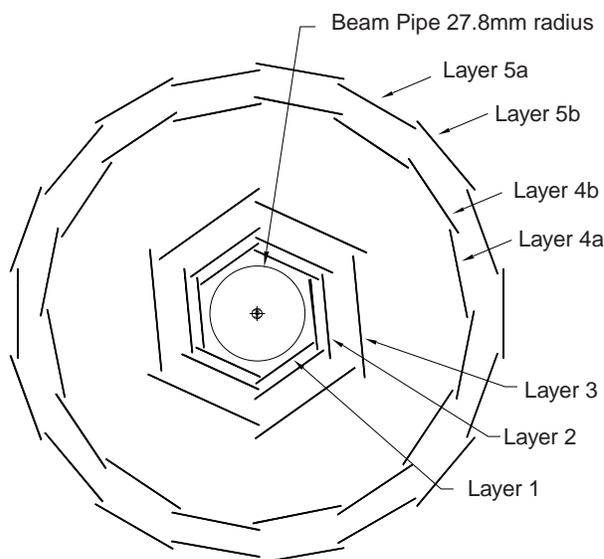


FIGURE 5.
Schematic end
view diagram of
the layers of the
SVT.

The inner layers are barrel-shaped, while the outer layers are tilted in at the ends to produce an arch shape in z . Each layer provides polar angle coverage down to $350\ \text{mrad}$ in the forward direction and $520\ \text{mrad}$ in the backward direction; smaller angles are obstructed by permanent dipole magnets mounted around the interaction point. A side cross-section of the SVT is shown in the figure below.

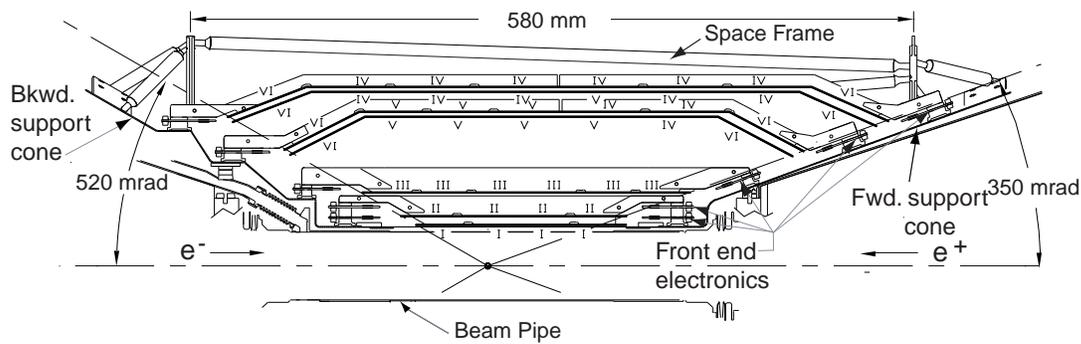


FIGURE 6. Side elevation diagram of the top half-section of the SVT.

The inner strips provide precise measurements of a track's angle and impact parameter. The outer layers improve pattern recognition and provide additional charge deposition measurements for determining dE/dx .

The sensors are read out from both ends by front-end electronics mounted just outside the active region. These electronics sample the charge collected in the strips at 30 MHz into a circular buffer. When a trigger arrives, hits in the appropriate time window are extracted from the buffer, multiplexed, and transferred to data acquisition electronics outside the detector.

Helical tracks reconstructed in the two tracking chambers are parameterized by five values: impact parameter in the x - y plane, d_0 ; impact parameter along the beam line, z_0 ; azimuthal angle at the point of closest approach to the interaction point, ϕ_0 ; $\tan \lambda$, where λ is the pitch angle of the helix; and κ , the curvature of the track. For most tracks, the SVT dominates the measurement of the first four of these parameters, with average precision,

- $\sigma_{d_0} = 23 \mu\text{m}$
- $\sigma_{z_0} = 29 \mu\text{m}$
- $\sigma_{\phi_0} = 0.43 \text{ mrad}$
- $\sigma_{\tan \lambda} = 0.53 \times 10^{-3}$.

3.4 The drift chamber

The DCH is a 40-layer, 7104-cell drift chamber with axial and stereo layers. It measures the helical trajectory of a charged particle traversing a magnetic field, and also provides energy loss and precise timing information.

The inner radius of the DCH is 23.6 cm and the outer radius is 80.9 cm. In the 1.5 T magnetic field, charged particles with transverse momentum $p_T > 180 \text{ MeV}/c$ reach the outer radius. Motivated by the asymmetry of the beam energy, the DCH is positioned asymmetrically in z , extending 174.9 cm forward and 101.5 cm backward from the nominal interaction point.

The DCH is arranged in 10 superlayers, each composed of 4 layers of hexagonal drift cells. The innermost and every third superlayer is axial, with wires parallel to the z axis. The remaining stereo superlayers are arranged at small angles to the z axis, to provide z coordinate measurements of tracks. The stereo angles vary between 45 mrad and 76 mrad, and alternate in sign between stereo superlayers.

The DCH is filled with a gas mixture of 80% helium and 20% isobutane. Charged particles passing through the DCH ionize the gas; ionization electrons are accelerated toward high-voltage sense wires, producing an avalanche of secondary ionizations along the way. The time of arrival of the ionization electrons at the sense wire determines the distance of closest approach of the track to the wire; a pattern matching algorithm uses this and the position of each wire to determine the track's trajectory. The integrated charge of the ionization electrons deposited in successive cells is used to measure dE/dx . The DCH was operated with voltage on the sense wires of 1900 V and 1960 V during Run 1, and of 1930 V subsequently.

The sense wires are read out by front-end electronics mounted on the rear endplate of the drift chamber, which digitize the arrival time and integrated charge of wire hits. This data is transferred to readout and trigger electronics. The DCH provides tracking and event timing information used

to trigger the other detector subsystems.

The DCH dominates the precision of the momentum measurement for most tracks. The precision of the transverse momentum is parameterized by,

$$\sigma_{p_T}/p_T = (0.13 \pm 0.01)\% \cdot p_T + (0.45 \pm 0.03)\% .$$

In addition, the average resolution of dE/dx measured in the DCH is 7%.

3.5 The Cherenkov detector

The DIRC is a particle identification system consisting of quartz bars surrounding the DCH. Charged particles traversing the active elements produce Cherenkov radiation, which is captured by internal reflection and exits through the back of the detector into an imaging region instrumented with photomultiplier tubes (PMTs). Using geometric and timing information, the detected radiation is associated with charged tracks. The particle's velocity and thus its mass and species is inferred from the angle of the radiation relative to the particle's trajectory,

The active elements of the DIRC consist of 144 synthetic quartz bars, 17×35 mm in cross section and 4.9 m long, arranged around the DCH and running parallel to the beam axis. Radiation produced inside the bars, which have an index of refraction $n = 1.473$, is reflected by the finished surfaces and bounces forward or backward to the ends of the bars. Mirrors affixed to the forward ends of the bars reflect radiation back toward the rear of the detector. The backward ends of the quartz bars open into a large, water-filled standoff box, through which radiation is projected onto an array of PMTs.

The conical radiation pattern of Cherenkov light emitted by charged particles in the DIRC is focused onto one or several rings or segments of rings on the PMT array. Reconstruction algorithms use the geometric and timing information to associate PMT hits with charged tracks reconstructed

in the tracking detectors and projected into the DIRC volume, and determine the Cherenkov angle θ_C . A particle with $\beta = 1$ at normal incidence in the center of the bar produces approximately 23 photoelectrons in the PMTs.

A schematic of one azimuthal segment of the DIRC is shown in the figure below.

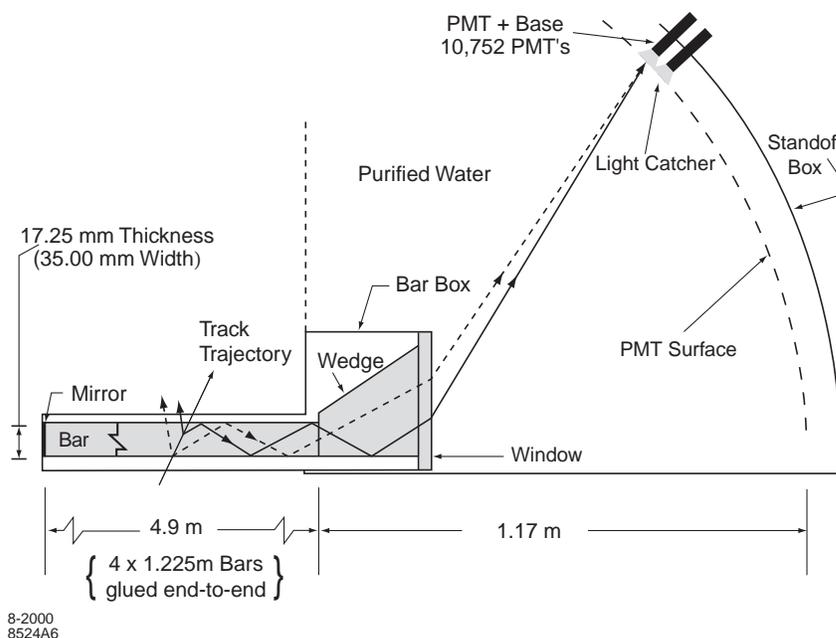


FIGURE 7. Schematic cross-section of one azimuthal segment of the DIRC, showing the trajectory of radiation emitted by a hypothetical particle.

The resolution of the reconstructed Cherenkov angle is measured in $\mu^+\mu^-$ events to be 2.5 mrad.

The kaon efficiency and pion misidentification rates determined from $D^0 \rightarrow K^-\pi^+$ reconstructed in D^* decays are shown in the figure below.

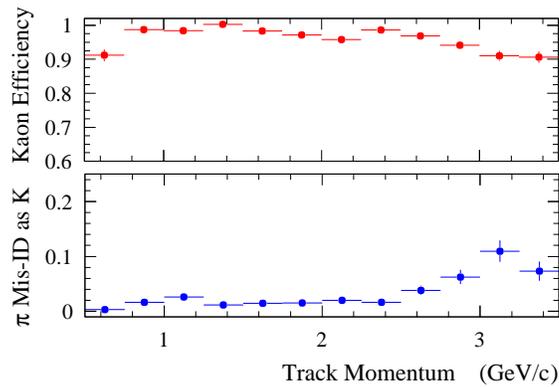


FIGURE 8. *top* Efficiency for K^\pm identification and *bottom* misidentification rate for π^\pm in the DIRC as a function of track momentum, determined from $D^0 \rightarrow K^- \pi^-$ decays.

The DIRC provides poor identification efficiency for particles with momenta below about 0.7 GeV/c; for these, dE/dx measured in the tracking detectors is used for identification.

3.6 The calorimeter

The EMC consists of 6580 thallium-doped CsI crystals, divided into a barrel-shaped central region surrounding the DIRC, and a forward endcap extending down to 15.8° from the beam line. The individual crystals, which are angled to point toward the interaction point, have a typical front area of 4.7×4.7 cm, while the depth varies from 29.6 to 32.4 cm (16.0 to 17.5 radiation lengths). The figure below illustrates the arrangements of the crystals in the EMC.

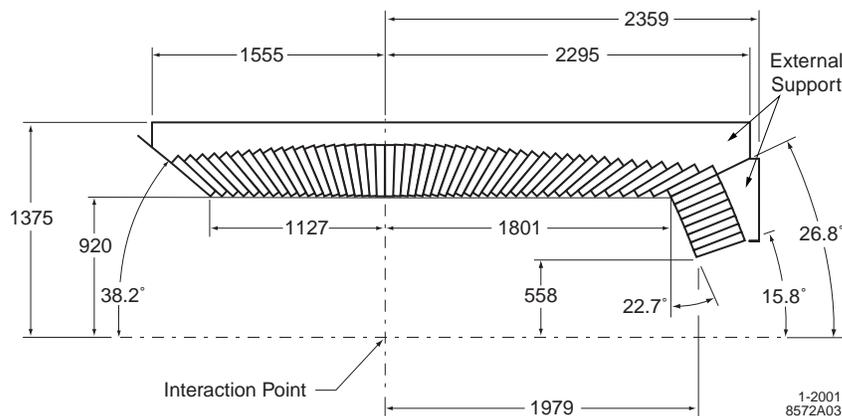


FIGURE 9. Half cross-sectional schematic of the EMC illustrating the arrangement of crystals.

Each crystal is wrapped on the front and sides with a thin reflective coating, and instrumented on the back face with two silicon photodiodes. The photodiodes are connected to preamplifiers mounted directly behind them, which are wired in turn to electronics mounted at the ends of the EMC support structure. These electronics further amplify the signals from the individual crystals, convert them to digital signals, and multiplex them for transfer to data acquisition hardware.

The crystals are calibrated with a radioactive source and using Bhabha scattering events. In calibration runs, the EMC is irradiated with 6.13 MeV gamma rays produced by an activated liquid (flourinert) circulated over the EMC face. Bhabha scattering $e^+e^- \rightarrow e^+e^-$ events collected during normal data taking produce 3 to 9 GeV clusters and are used for high-energy calibration. In addition, we apply corrections to the energies of reconstructed photon candidates to compensate for shower leakage into cracks between detector segments and from the sides and back of the detector.

Neutral clusters are reconstructed from sets of contiguous crystals in which energy deposits are measured in an event. The reconstruction algorithm searches for local minima, or *bumps*, in the cluster and attempts to distinguish merged clusters from single particle showers. Tracks reconstructed in the tracking chamber are matched to EMC bumps to distinguish neutral particles.

The energy resolution achieved by the EMC is parameterized as

$$\frac{\sigma_E}{E} = \frac{(2.32 \pm 0.30)\%}{\sqrt[4]{E(\text{GeV})}} \oplus (1.85 \pm 0.12)\% .$$

The angular resolution of reconstructed clusters is parameterized as

$$\sigma_\theta = \sigma_\phi = \left(\frac{3.87 \pm 0.07}{\sqrt{E(\text{GeV})}} + 0.00 \pm 0.04 \right) \text{ mrad} .$$

3.7 Triggers and data acquisition

The trigger system is divided into two stages, a hardware *Level 1* trigger and a software *Level 3* trigger.

The Level 1 trigger is implemented in hardware and receives inputs from the readout electronics of the DCH and EMC. The DCH trigger (DCT) constructs track segments from hits in individual layers, and assembles track segments within azimuthal segments to form 2D tracks. The EMC trigger (EMT) groups crystals into sectors and searches for activity in these sectors consistent with a minimum ionizing particle or particle shower. The Level 1 trigger combines objects constructed in the EMT and DCT to produce a trigger decision as the disjunction of several preset criteria. A Level 1 accept can be triggered by DCT objects only, EMT objects only, or a combination of the two. The typical output rate of the Level 1 trigger is 1 kHz.

A Level 1 accept decision is propagated to the readout electronics of all the detector components, and triggers readout of detector channels into the data acquisition system. The data acquisition system is implemented as a combination of special-purpose electronics and a farm of UNIX workstations. One workstation node is assigned to each Level 1 accept, and collects the contributions from the detector subsystems into a complete event.

The Level 3 trigger is implemented in software and runs on the same workstation farm as the software component of the data acquisition system. The trigger performs more detailed processing of event data, reconstructing 3D tracks and localized EMC clusters using look-up tables. As with the Level 1 trigger, objects reconstructed in the DCH or in the EMC, or a combination of the two produce a Level 3 trigger. The Level 3 output rate is limited to 120 Hz, which includes physics triggers, prescaled Bhabha events, random triggers, and diagnostic triggers.

The combined Level 1 and Level 3 trigger efficiency is greater than 99.9% for $B\bar{B}$ events and greater than 95% for light-quark continuum events.

The Level 3 trigger is hosted on the workstation farm nodes by the Online Event Processing (OEP) system, which manages the storage and forwarding of completed events. OEP also hosts a real-time data quality monitoring system, which accumulates statistical distributions of measured

quantities both for real-time monitoring of graphical displays by physicists operating the detector, and for automated comparison against reference distributions. Completed events are forwarded from the workstation nodes to a staging area, where they are assembled into complete runs and sent to a tape storage system for archiving. The data are subsequently processed by full event reconstruction programs in preparation for physics analysis; these programs also perform additional data quality checks and offline calibrations.

3.8 Particle selection criteria

The *BABAR* experiment has defined standard selection criteria for tracks and particle candidates. The performance of these criteria has been well-studied, and efficiency is understood in data and MC. We describe here the standard selection criteria used in this analysis. Criteria are provided for “very loose,” “loose,” “tight,” and “very tight” selection.

3.8.1 CHARGED TRACKS: *GoodTracksLoose*

We use the standard *GoodTracksLoose* selection for charged tracks. These tracks are required to satisfy,

- at least 12 hits in the DCH,
- impact parameter in the $x - y$ plane $d_0 < 1.5$ cm,
- distance of closest approach in z to the nominal interaction point $z_0 < 10$ cm,
- momentum $p < 10$ GeV/ c , and
- transverse momentum $p_T > 100$ MeV/ c .

The figure below shows the efficiency of the *GoodTracksLoose* selection as a function of transverse momentum, angles, and the track multiplicity in the event. Efficiencies are shown for a segment of

the data sample in which the DCH was operated at 1930 V, and for a MC sample simulated with the same conditions.

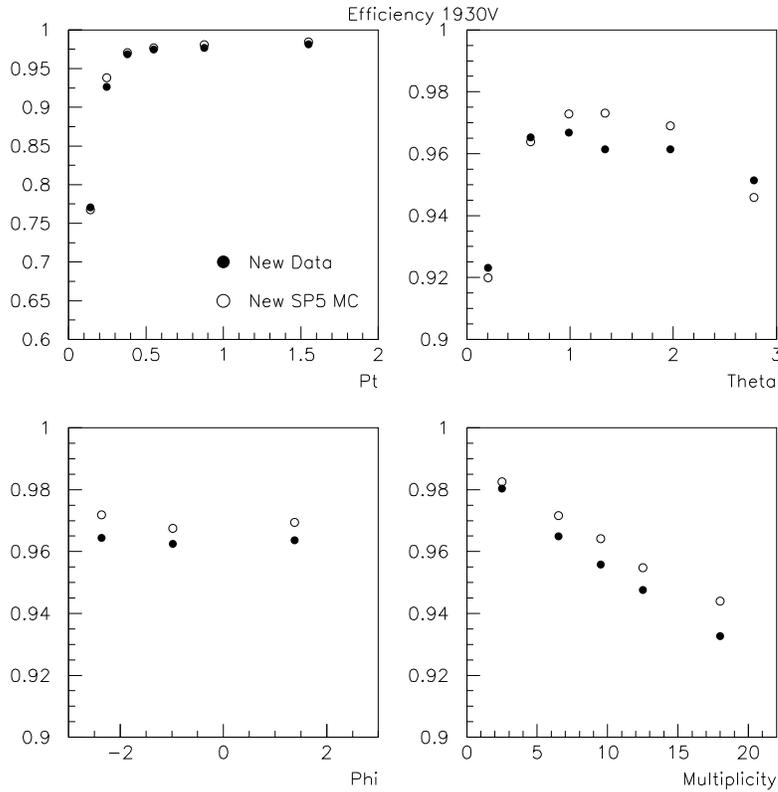


FIGURE 10. Efficiency distributions for *GoodTracksLoose* as a function of transverse momentum p_T , polar angle θ , azimuthal angle ϕ , and multiplicity. Solid points show data collected with DCH at 1930 V; empty points show simulation assuming the same DCH conditions.

Standard efficiency corrections have been measured to compensate for the difference between data and MC, and we weight the candidates in MC samples accordingly. The average correction is 0.992.

3.8.2 PHOTONS: *GoodPhotonLoose*

We use the standard *GoodPhotonLoose* selection for photons. This selection applies to EMC clusters that are not matched to a charged track, and requires

- energy in the lab frame of at least 100 MeV,
- at least 4 EMC crystals in the cluster, and
- a lateral moment less than 0.8.

3.8.3 CHARGED KAON AND PION IDENTIFICATION

A charged track is identified as a kaon or pion candidate using dE/dx information from the SVT and DCH for low-momentum ($p < 0.7 \text{ GeV}/c$) tracks, and DIRC information for high-momentum ($p > 0.6 \text{ GeV}/c$) tracks. For each particle hypothesis (pion, kaon, electron, muon, or proton), a likelihood is constructed using the expected distributions of dE/dx in the SVT and DCH, and number of detected photons and reconstructed Cherenkov angle in the DIRC.

The PID selection criteria are cuts on the likelihood ratios between pairs of particle hypothesis. In this analysis, we use the *TightKaonMicroSelection* for charged kaons. The efficiency for this selection is shown in the following figure, followed by the mistag rate for charged pions.

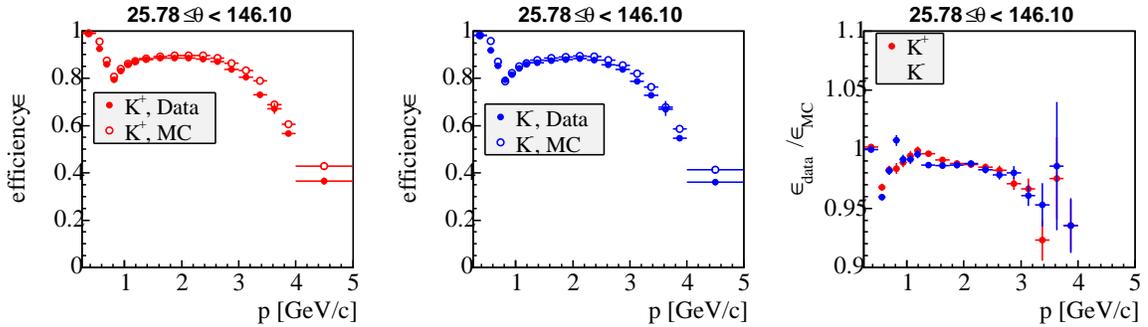


FIGURE 11. Efficiency of *TightKaonMicroSelection* PID selection criteria for K^+ (red) and K^- (blue) as a function of momentum. Filled points are Run 4 data and empty points are MC; the right plot shows the ratio of data to MC efficiency.

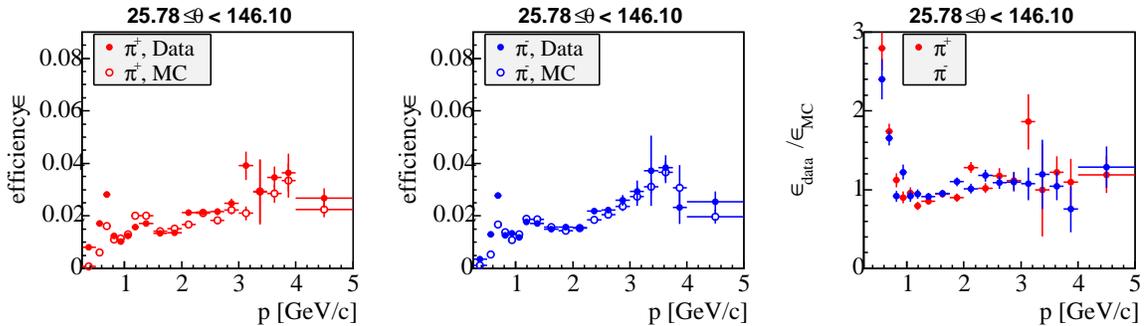


FIGURE 12. Charged pion mistag of *TightKaonMicroSelection* PID selection criteria for π^+ (red) and π^- (blue) as a function of momentum. Filled points are Run 4 data and empty points are MC; the right plot shows the ratio of the mistag rate in data to MC.

We use the *LooseLHPionMicroSelection* for charged pions. The efficiency for this selection is shown in the following figure.

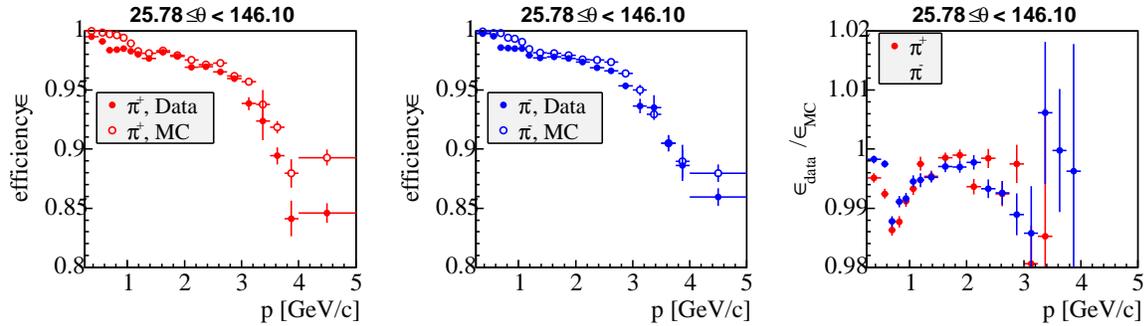


FIGURE 13. Efficiency of *LooseLHPionMicroSelection* PID selection criteria for π^+ (red) and π^- (blue) as a function of momentum. Filled points are Run 4 data and empty points are MC; the right plot shows the ratio of data to MC efficiency.

Standard efficiency corrections have been measured to compensate for the difference between data and MC in charged PID selection efficiency; we weight the candidates in MC samples accordingly.

The average correction is 0.999 for pions and 0.987 for kaons.

4 Samples

In this chapter, we describe the data and MC samples used in this analysis

4.1 Data samples

We use the *BABAR* dataset current as of early 2005, which consists of 210.6 fb^{-1} of data taken on the $\Upsilon(4S)$ resonance (*on-resonance*), and 21.6 fb^{-1} of data taken approximately 20 MeV below the $\Upsilon(4S)$ resonance (*off-resonance*). Off-resonance data samples contain light-quark continuum events comparable to those in the on-resonance data, but do not include $B\bar{B}$ events. The data are divided into four runs: Run 1 taken in 1999–2000, Run 2 taken in 2001 and early 2002, Run 3 taken in late 2002 and 2003, and Run 4 taken in 2004. The number of $\Upsilon(4S) \rightarrow B\bar{B}$ decays in the on-resonance data sample is $(231.8 \pm 1.5) \times 10^6$.

4.2 Monte Carlo samples

We also use Monte Carlo (MC) simulated data samples in the development of this analysis and to estimate efficiencies and background rates. The production and decay processes are simulated with the *EvtGen* [18] event generator, which incorporates the current understanding of the decays of B mesons and their decay products. Decays of light quarks are simulated with the *JETSET* [19] fragmentation model.

These simulated decays are processed in a detailed model of the *BABAR* detector, implemented using *Geant4* [20]. The detector simulation produces output similar to that of the detector’s data acquisition system, and the simulated data is processed using the same event reconstruction code as real data. Information about the simulated physics process, the MC “truth” information, is retained

in simulated events, and objects reconstructed in the detector are associated with information about the underlying generated particles. This allows us to determine on an event-by-event basis whether our hypothesis about candidates reconstructed in MC samples are correct, which is known as *MC truth matching*.

MC samples are generated using historical detector configurations, conditions, and background data. As with the data sample, MC samples are divided into four runs, and we scale the effective luminosity of each MC sample to the on-resonance data luminosity of the corresponding run. Throughout this analysis, distributions and yields derived from MC samples are normalized to the equivalent integrated luminosity of the on-resonance data sample using world-average branching fractions [2], except where noted. Efficiency studies have uncovered residual discrepancies between the simulated and actual detector response; we apply corrections, which are standardized for most *BABAR* analyses, to the simulated data to reduce the discrepancies. These corrections are described in Chapter 5.

4.2.1 SIGNAL MODEL

We model our signal processes with a cocktail of exclusive MC samples listed in the table below. Each sample is scaled according to the listed branching fractions. We use separate simulated samples for $B^+ B^-$ decays and for $B^0 \bar{B}^0$ decays.

Mode	Assumed \mathcal{B}	Source
$B \rightarrow K_1(1270) \gamma$	4.28×10^{-5}	Belle measurement
$B \rightarrow K_1(1400) \gamma$	0.80×10^{-5}	ansatz
$B \rightarrow K^*(1410) \gamma$	0.80×10^{-5}	ansatz
$B \rightarrow K_2^*(1430) \gamma$	1.34×10^{-5}	average of <i>BABAR</i> measurements
$B \rightarrow K^*(1680) \gamma$	0.20×10^{-5}	ansatz

TABLE 4.
Branching fractions assumed for radiative decays to K_X resonances.

The branching fraction for $B \rightarrow K_1(1270) \gamma$ is measured by the Belle Collaboration [13], while the $B \rightarrow K_2^*(1430) \gamma$ branching fraction is measured by *BABAR* in $K_2^*(1430) \rightarrow K\pi$ [11]. The other branching fractions have not been measured; we choose ansatz values. Belle has published

a limit $\mathcal{B}(B \rightarrow K_1(1400) \gamma) < 1.5 \times 10^{-5}$; for this mode we use 0.8×10^{-5} , a value typical of theoretical estimates [6, 3], (excluding those estimates that predict a larger rate of $K_1(1400) \gamma$ than $K_1(1270) \gamma$). Lacking better information, we use the same rate for $\mathcal{B}(B \rightarrow K^*(1410) \gamma)$. For $\mathcal{B}(B \rightarrow K^*(1680) \gamma)$, theoretical estimates suggest a smaller branching fraction than for the lower-lying resonances, so we use 0.2×10^{-5} .

The following table lists the branching fractions we use [2] of K_X resonances to $K\pi\pi$ via $K\rho$ and $K^*\pi$.

resonance	assumed $\mathcal{B}(\rightarrow K\rho)$	assumed $\mathcal{B}(\rightarrow K^*\pi)$
$K_1(1270)$	42%	44%
$K_1(1400)$	3%	94%
$K^*(1410)$	7%	86%
$K_2^*(1430)$	9%	25%
$K^*(1680)$	31%	30%

TABLE 5.
Branching fractions
assumed for decays
of K_X resonances.

For the purposes of normalizing our signal MC samples, we have included the $K_1(1270) \rightarrow K_0^*(1430) \pi$ fraction in the $K_1(1270) \rightarrow K^*\pi$ component. For $K^*(1410)$, the measured branching fractions are $\mathcal{B}(K^*(1410) \rightarrow K\pi) = 7\%$ and $\mathcal{B}(K^*(1410) \rightarrow K\rho) < 7\%$; we have taken the upper limit as the branching fraction for $K\rho$, and assumed the fraction not accounted for by $K\pi$ and $K\rho$ decays always to $K^*\pi$.

These simulated samples are implemented in the *EvtGen* event generator as sequential incoherent decays, in which each intermediate resonance is on-shell and there is no interference among channels that produce identical final states. Helicity amplitudes are carried for each particle, but amplitudes and phases are not computed for the entire decay tree; instead, the rate of a decay tree is computed from the branching fractions of the individual decay processes. This is known to produce results that are incorrect for our signal model in several respects:

1. Interference among various $B \rightarrow K_X \gamma$ decays is not simulated. Note that the relative phases for these processes are not known.

2. Interference among decays of K_X resonances are not simulated, for instance between $K_X^+ \rightarrow K^{*+}\pi^0$ and $K^{*+} \rightarrow K^0\pi^+$, or between $K_X^+ \rightarrow K^{*0}\pi^+$ and $K^{*0} \rightarrow K^0\pi^0$.
3. Breit-Wigner line shapes, irrespective of the production process, are used for intermediate resonances, both the primary K_X resonance and secondary resonances, such as K^* and ρ . This is known to be inaccurate, especially in the case of $K_1(1270) \rightarrow \rho K$, which is close to threshold.
4. The simulation cannot accommodate the decay $K_1(1270) \rightarrow K_0^*(1430)\pi$, which is below threshold for the nominal value of the $K_0^*(1430)$ mass.

In addition, our signal model does not include a non-resonant $B \rightarrow K\pi\pi\gamma$ component. It has not been established whether there is a non-resonant component in these decays.

Our analysis, however, does not depend strongly on details of the signal model or on the signal branching fractions we assume in the model. Our dependence on the model is as follows:

- We use the model to optimize our selection procedure. If the model is incorrect, our selection may be suboptimal.
- We use the model to estimate efficiencies for reconstructing signal events. We have established that these efficiencies do not depend on the distribution of $m_{K\pi\pi}$ in our signal model.
- We use the model to estimate backgrounds from misreconstructed $B \rightarrow K\pi\pi\gamma$ decays. We have performed studies to estimate the uncertainty of our results due to modeling of these backgrounds.

4.2.2 $B \rightarrow K\pi\gamma$

The inclusive $b \rightarrow s\gamma$ event generator used in the *BABAR* MC simulation does not reproduce the correct distribution of $B \rightarrow K\pi\gamma$ events for low $m_{K\pi}$, so we use exclusive MC samples to model

these decays. In the kinematic region we consider, $B \rightarrow K\pi\gamma$ is dominated by the $K^*(892)$ resonance, with a smaller contribution from $K_2^*(1430)$. We use a cocktail of exclusive $K^*(892)$ and $K_2^*(1430)$ MC samples to estimate backgrounds from $B \rightarrow K\pi\gamma$.

4.2.3 INCLUSIVE $b \rightarrow s\gamma$

To evaluate backgrounds from $b \rightarrow s\gamma$ processes other than $B \rightarrow K\pi\gamma$ and $B \rightarrow K\pi\pi\gamma$, we use inclusive $b \rightarrow s\gamma$ MC samples. These simulate $B \rightarrow X_{su}\gamma$ and $B \rightarrow X_{sd}\gamma$ decays using the model of Kagan and Neubert [21] with the b quark mass set to $m_b = 4.80 \text{ GeV}/c^2$. The X_{su} and X_{sd} , diquark states with strangeness $S = -1$, are decayed by *JETSET* using a generic fragmentation model. To assess the model dependence of the MC predictions of backgrounds from $b \rightarrow s\gamma$ processes, we also study $b \rightarrow s\gamma$ MC samples generated with the same Kagan and Neubert model with $m_b = 4.65 \text{ GeV}/c^2$.

BABAR has measured [22] the inclusive $b \rightarrow s\gamma$ branching fraction by a method in which many exclusive final states are reconstructed and combined, the “semi-inclusive” technique. In this analysis, it was found that the fragmentation model used in the inclusive $b \rightarrow s\gamma$ simulation does not accurately reproduce the multiplicity distribution in data. The semi-inclusive analysis measured ratios of data to MC yields for most low-multiplicity (five or fewer particles) final states. We reweight the inclusive $b \rightarrow s\gamma$ MC sample by these ratios.

4.2.4 EXCLUSIVE BACKGROUND PROCESSES

We study specific B decays which can potentially produce background candidates with kinematic properties similar to those of our signal. We use exclusive MC samples shown in the table below. The middle column shows the signal mode in which we expect to reconstruct these processes as “peaking” background.

Process	Signal mode
$B^0 \rightarrow D^- \rho^+, D^- \rightarrow K_S^0 \pi^-$	$K_S^0 \pi^- \pi^+ \gamma$
$B^- \rightarrow D^0 \rho^-, D^0 \rightarrow K^- \pi^+$	$K^+ \pi^- \pi^+ \gamma$
$B^- \rightarrow D^0 \rho^-, D^0 \rightarrow K_S^0 \pi^0$	$K_S^0 \pi^+ \pi^0 \gamma$
$B^0 \rightarrow D^0 \pi^0, D^0 \rightarrow K^- \pi^+ \pi^0$	$K^+ \pi^- \pi^0 \gamma$
$B^0 \rightarrow D^0 \pi^0, D^0 \rightarrow K_S^0 \pi^+ \pi^-$	$K_S^0 \pi^- \pi^+ \gamma$
$B^0 \rightarrow D^0 \eta, D^0 \rightarrow K^- \pi^+ \pi^0$	$K^+ \pi^- \pi^0 \gamma$
$B^0 \rightarrow D^0 \eta, D^0 \rightarrow K_S^0 \pi^+ \pi^-$	$K_S^0 \pi^- \pi^+ \gamma$
$B^+ \rightarrow K^{*0} \rho^+, K^{*0} \rightarrow K^+ \pi^-$	$K^+ \pi^- \pi^+ \gamma$
$B^+ \rightarrow K^{*0} \rho^+, K^{*0} \rightarrow K_S^0 \pi^0$	$K_S^0 \pi^+ \pi^0 \gamma$
$B^+ \rightarrow K^{*+} \rho^0, K^{*+} \rightarrow K^+ \pi^0$	$K^+ \pi^- \pi^+ \gamma$
$B^0 \rightarrow K^{*+} \rho^-, K^{*+} \rightarrow K^+ \pi^0$	$K^+ \pi^- \pi^0 \gamma$
$B^0 \rightarrow K^{*+} \rho^-, K^{*+} \rightarrow K_S^0 \pi^+$	$K_S^0 \pi^- \pi^+ \gamma$
$B^0 \rightarrow K^{*0} \rho^0, K^{*0} \rightarrow K_S^0 \pi^0$	$K_S^0 \pi^+ \pi^0 \gamma$

TABLE 6. MC samples used for studies of peaking B decays.

4.2.5 OTHER GENERIC PROCESSES

The largest background processes in this analysis are continuum production of $u\bar{u}$, $d\bar{d}$, $s\bar{s}$, and $c\bar{c}$ quark pairs. We study these backgrounds using generic uds and $c\bar{c}$ MC samples, as well as in off-resonance data.

We also use generic $B^0 \bar{B}^0$ and $B^+ B^-$ samples to evaluate backgrounds from B decays. We remove $b \rightarrow s\gamma$ events from these samples using MC truth information.

4.2.6 $D\pi^\pm$ CONTROL SAMPLE

We study the accuracy of the MC simulation of our hadronic selection and event shape variables using a control sample of $B \rightarrow D\pi^+$, $D \rightarrow K\pi\pi$ decays. We compare candidates reconstructed in these modes in on-resonance data to MC samples. The samples are listed below.

Process	Signal mode
$B^0 \rightarrow D^- \pi^+, D^- \rightarrow K^+ \pi^- \pi^-$	$K^+ \pi^- \pi^+ \gamma$
$B^+ \rightarrow \bar{D}^0 \pi^+, \bar{D}^0 \rightarrow K^+ \pi^- \pi^0$	$K^+ \pi^- \pi^0 \gamma$
$B^+ \rightarrow \bar{D}^0 \pi^+, \bar{D}^0 \rightarrow K^0 \pi^+ \pi^-$	$K_S^0 \pi^- \pi^+ \gamma$
$B^0 \rightarrow D^- \pi^+, D^- \rightarrow K^0 \pi^- \pi^0$	$K_S^0 \pi^+ \pi^0 \gamma$

TABLE 7. MC samples used for comparison to the $B \rightarrow D\pi^+$, $D \rightarrow K\pi\pi$ control sample.

5 Candidate Selection

Event selection proceeds in two steps. First, we process the data and MC samples, construct $K\pi\pi\gamma$ candidates, and apply a simple set of loose *sample selection* requirements that are close to 100% efficient for reconstructed signal candidates. Processing the data and MC samples is time- and computationally-intensive; creating a preselected sample of candidates allows us to develop our selection procedure more efficiently. Our candidate construction and sample selection process operates on a standard skimmed subset of the entire data and MC event samples; this is described in section 5.1. We reconstruct candidates, and preselect them for inclusion in our candidate sample according to sample selection cuts described in section 5.2.

We then apply an optimized set of *candidate selection* cuts to the candidates in these samples. The cuts are described in section 5.3. We determine optimal cut values to maximize the figure of merit $S^2/(S + B)$, where S is the yield in truth-matched signal MC and B is the yield in the $udsc$ MC, both evaluated in a small region around the signal peak in m_{ES} and ΔE . The optimization is described in section 5.4. We also restrict candidates to a rectangular region in m_{ES} and ΔE , the *fit region*. Optimization of the fit region is deferred to section 5.5. After imposing the optimized candidate selection, we find that many signal events contain multiple candidates. Section 5.6 describes our choice from among alternative techniques for *multiple candidate selection* (MCS).

The final set of *all cuts* consists of sample selection, candidate selection, and the fit region cut, followed by MCS.

5.1 Skim

We use the standard *BABAR BtoXGamma* skim when running over data and MC samples. This skim requires that an event is accepted by either the *BGFMultiHadron* or *BGFNeutralHadron* background filters, and that the CM energy of the highest-energy photon candidate in the event falls between 1.5 and 3.5 GeV. The *BGFMultiHadron* filter requires that the event contains three or more charged tracks, and that $R_2^{\text{ch}} < 0.98$, where R_2^{ch} is the ratio of second to zeroth Fox-Wolfram moments computed from the momenta of charged tracks. This second requirement reduces the rate of Bhabha events accepted by the filter. The *BGFNeutralHadron* filter accepts events with fewer than three tracks, but only neutral particles are detected in the event. The filters are designed to accept *B* physics events while reducing the rate of Bhabha events, continuum events, and beam backgrounds.

5.2 Sample selection

We reconstruct *B* candidates by the following procedure.

- High-energy photons are selected from neutral clusters satisfying the *GoodPhotonLoose* criteria (section 3.8.2). We further require that the photon candidate has a CM energy of at least 1 GeV.
- K^\pm candidates are selected from tracks satisfying the *KLHVeryLoose* PID selection (section 3.8.3).
- $K_S^0 \rightarrow \pi^+\pi^-$ candidates are selected from pairs of oppositely-charged tracks with invariant mass within $25 \text{ MeV}/c^2$ of $m_{K_S^0}$. The mass is calculated from track momenta evaluated at the K_S^0 vertex, which is determined from a geometric fit of the two tracks.
- π^\pm candidates are selected from charged tracks satisfying the *GoodTracksLoose* criteria (sec-

tion 3.8.1).

- $\pi^0 \rightarrow \gamma \gamma$ candidates are selected from pairs of photons with invariant mass between 100 and 160 MeV/c^2 and with total energy of at least 200 MeV.

We reconstruct B candidates in all four modes from right-sign combinations of π^\pm , π^0 , K^\pm , and K_s^0 candidates. We exclude candidates in which a charged track or neutral cluster is used in more than one particle of the final state, but we do allow multiple overlapping candidates in the same event.

Our sample selection consists of these B candidates satisfying, in addition,

- $m_{ES} > 5.2 \text{ GeV}/c^2$,
- $|\Delta E| < 0.5 \text{ GeV}$, and
- $m_{K\pi\pi} < 2.2 \text{ GeV}/c^2$.

5.3 Cuts

This section describes the selection requirements we impose on B candidates. The cut values we use are presented with each cut, but the description of the cut optimization procedure is deferred to the end.

5.3.1 HIGH-ENERGY PHOTON SELECTION

We require the high-energy photon to pass the following requirements.

- We require that none of the crystals making up the photon cluster, or associated electronics, is marked as dead or hot in the detector's running conditions at the time the event was collected.
- We require $-0.74 < \cos \theta_\gamma < 0.93$, where θ_γ is the angle between the z axis and the direction of EMC cluster centroid in the lab frame. This selects photons falling in the fiducial region of the EMC.

- We require $M_2 < 0.002$, where M_2 is the geometric second moment of the crystals in θ - ϕ coordinates, weighted by energy. This removes clusters with an oblong shape in the transverse plane, which are associated with merged decays of high-energy π^0 and η mesons.
- We require that the three-dimensional distance between the centroid of the EMC cluster and the centroid of the nearest other bump be greater than 25 cm. This isolation requirement removes many photons from π^0 and η decays.

These cuts are identical to the high-energy photon selection used in most *BABAR* radiative penguin analyses, and have been validated thoroughly. We therefore have not further optimized the values of these cuts for the present analysis.

5.3.2 $K\pi\pi$ MASS RANGE

Given prior expectations of the $K\pi\pi$ resonance structure, we did not expect the region $m_{K\pi\pi} > 1.8\text{ GeV}/c^2$ to contain large number of signal events. Continuum backgrounds, however, increase as a function of $m_{K\pi\pi}$, as shown in the figures below. We therefore restrict our consideration to the region $m_{K\pi\pi} < 1.8\text{ GeV}/c^2$. This cut also removes background candidates from B decays in which the hadronic part of the final state is produced by the decays $D \rightarrow K\pi\pi$, as well as a great deal of other backgrounds from B decays.

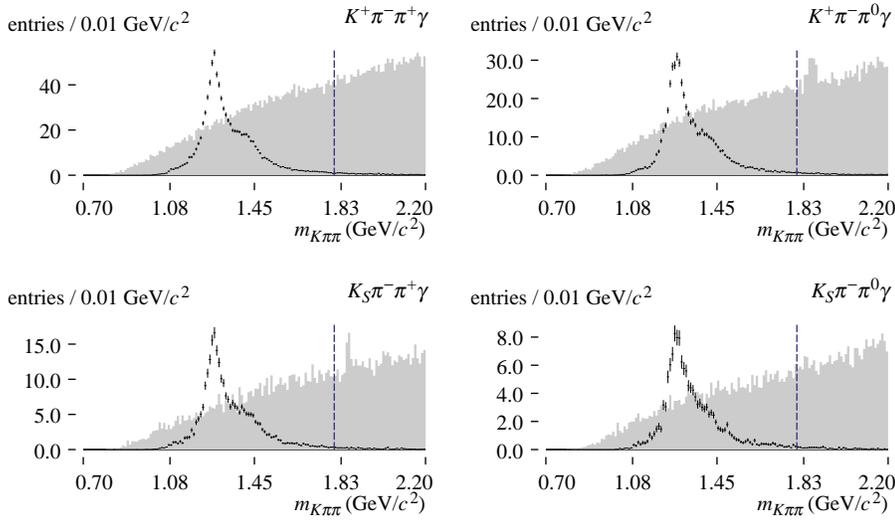


FIGURE 14. Distribution of $m_{K\pi\pi}$ in truth-matched signal MC (points) and $udsc$ MC (shaded; scaled arbitrarily). The dashed blue lines indicate cuts. All other candidate selection cuts have been applied.

5.3.3 π^0 AND η VETOES

Asymmetric π^0 and η decays are a large source of high-energy photon candidates in the continuum. We reduce this background by vetoing any photon candidate which, in combination with another photon in the event, is consistent with the decay of a π^0 or η .

To veto photons from π^0 decays, we combine our photon candidate with other photons of at least 50 MeV in the event, and select combinations with at least 200 MeV total energy. We reject our photon candidate if we find a combination with invariant mass within 25 MeV/ c^2 of the π^0 mass. The distribution of the invariant mass of the combination closest to the π^0 mass is shown below.

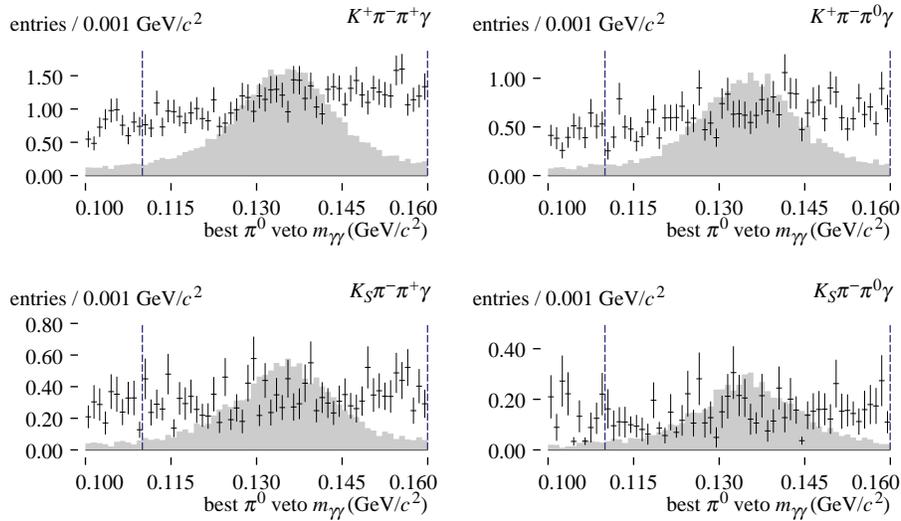


FIGURE 15. Distribution of $m_{\gamma\gamma}$ for the $\gamma\gamma$ veto pair closest to the π^0 mass in truth-matched signal MC (points) and $udsc$ MC (shaded; scaled arbitrarily). The dashed blue lines indicate cuts. The second photon in the veto is required to have $E_\gamma > 50 \text{ MeV}/c^2$. All other candidate selection cuts have been applied.

To veto photons from $\eta \rightarrow \gamma\gamma$ decays, we combine our photon candidate with other photons of at least 250 MeV. We reject our photon candidate if we find a combination with invariant mass within 40 MeV of the η mass. The distribution of the invariant mass of the combination closest to the η mass is shown below.

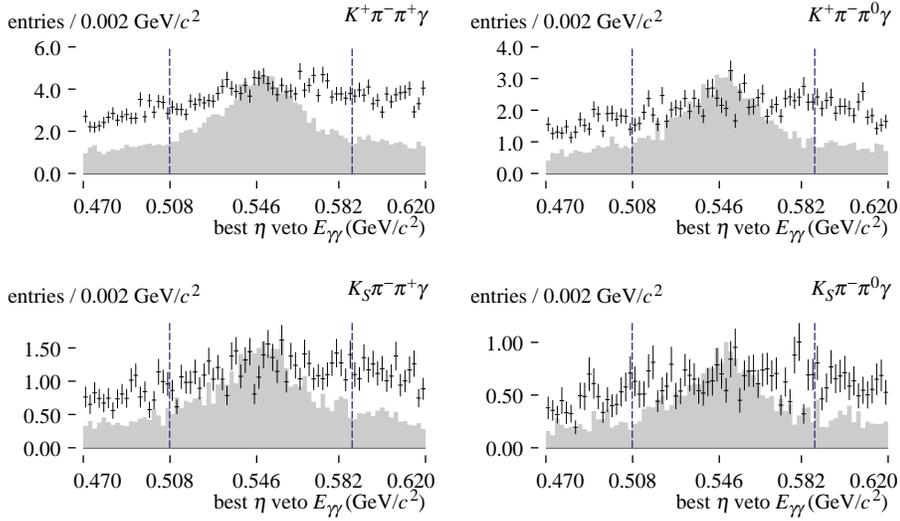


FIGURE 16. Distribution of $m_{\gamma\gamma}$ for the $\gamma\gamma$ veto pair closest to the η mass in truth-matched signal MC (points) and $udsc$ MC (shaded; scaled arbitrarily). The dashed blue lines indicate cuts. The second photon in the veto is required to have $E_\gamma > 250 \text{ MeV}/c^2$, and all other candidate selection cuts have been applied.

5.3.4 K^\pm PARTICLE ID

We require K^\pm candidates to satisfy the *TightKaonMicroSelection* PID selector (section 3.8.3) and the *GoodTracksLoose* track selection (section 3.8.1).

5.3.5 π^\pm PARTICLE ID

We require π^\pm candidates to satisfy the *LooseLHPionMicroSelection* PID selector (section 3.8.3). Our sample selection also requires that π^\pm tracks satisfy the *GoodTracksLoose* selection (section 3.8.1).

5.3.6 K_S^0 SELECTION

We use the *TreeFitter* [23] fitting algorithm to perform a geometric fit using the π^+ and π^- tracks from a K_S^0 candidate. The fit determines the K_S^0 decay vertex, with which we can evaluate the K_S^0 four-momentum and trajectory. We evaluate the K_S^0 mass using π^\pm track momenta projected from

this vertex. To determine the trajectory, we also require a measurement of the K_S^0 production vertex, i.e., the B decay vertex. We obtain this using a global fit to the entire reconstructed final state, which is described in section 5.3.8.

The displaced decay of the K_S^0 is useful for rejecting background. We compute the decay length significance $d_{K_S^0}/\sigma(d_{K_S^0})$, the ratio of the three-dimensional length of the K_S^0 trajectory, and the error on that quantity obtained from the vertex fit. We also compute θ_{flight} , the angle between the K_S^0 trajectory and its momentum vector.

We impose the following cuts on K_S^0 candidates:

- $|m_{\pi^+\pi^-} - m_{K_S^0}| < 11 \text{ MeV}/c^2$
- $\cos \theta_{\text{flight}} > 0.995$
- $d_{K_S^0} > 5 \sigma(d_{K_S^0})$

We do not explicitly cut on the goodness-of-fit of the $\pi^+\pi^-$ vertex fit, but the mass and decay length significance cuts implicitly remove candidates with failed or poor fits. The distributions of the three variables used for K_S^0 selection are shown below.

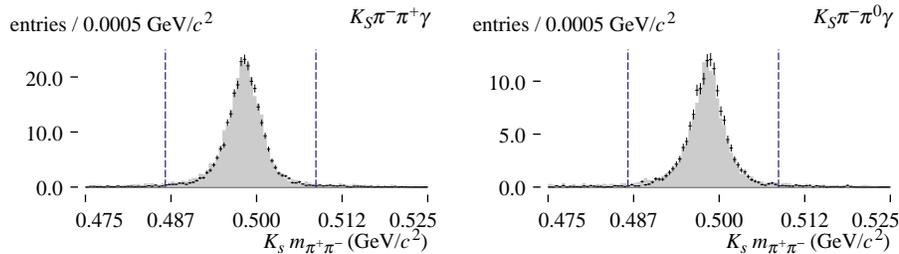


FIGURE 17. Distribution of reconstructed K_S^0 mass $m_{\pi^+\pi^-}$ in truth-matched signal MC (points) and $udsc$ MC (shaded; scaled arbitrarily). The dashed blue lines indicate cuts. All other candidate selection cuts have been applied.

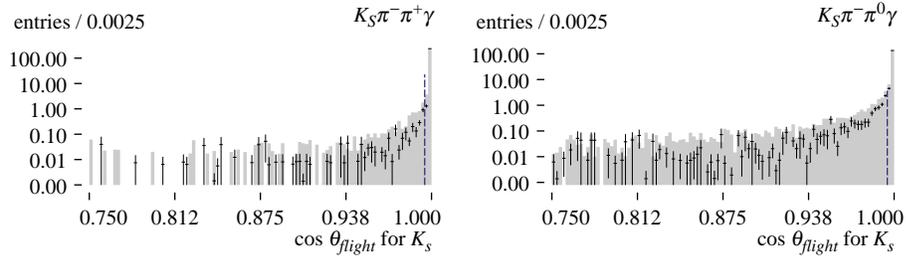


FIGURE 18. Distribution of $K_S^0 \cos \theta_{\text{flight}}$ in truth-matched signal MC (points) and $udsc$ MC (shaded; scaled arbitrarily). The dashed blue lines indicate cuts. All other candidate selection cuts have been applied.

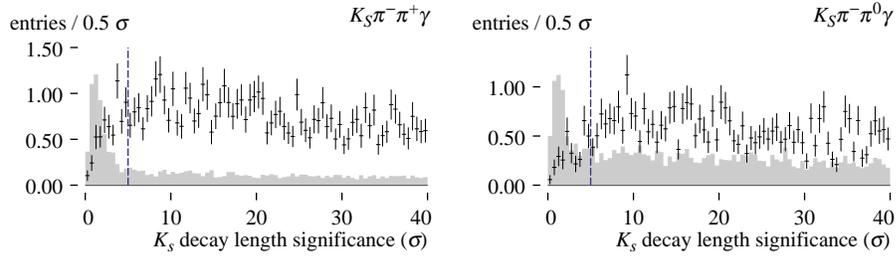


FIGURE 19. Distribution of K_S^0 flight distance divided by uncertainty on that quantity in truth-matched signal MC (points) and $udsc$ MC (shaded; scaled arbitrarily). The dashed blue lines indicate cuts. All other candidate selection cuts have been applied.

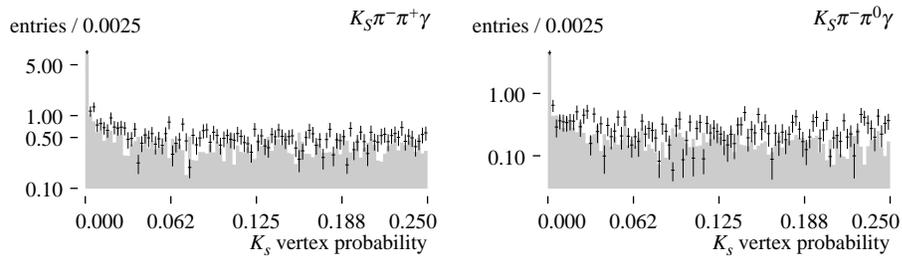


FIGURE 20. Distribution of K_S^0 vertex probability in truth-matched signal MC (points) and $udsc$ MC (shaded; scaled arbitrarily). All candidate selection cuts have been applied.

Standard corrections that tune the efficiency of these K_S^0 selection cuts in MC to match the efficiency in data have been calculated. We re-weight MC samples with these efficiency corrections.

5.3.7 π^0 SELECTION

We require that the invariant mass of π^0 candidates fall within $16 \text{ MeV}/c^2$ of the nominal π^0 mass.

The invariant mass distribution is shown below.

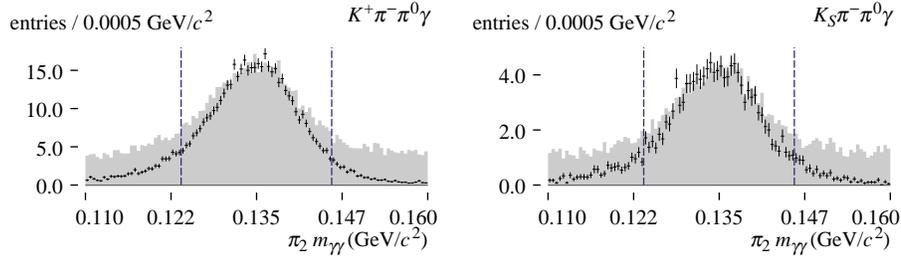


FIGURE 21. Distribution of $m_{\gamma\gamma}$ for π^0 candidates in truth-matched signal MC (points) and $udsc$ MC (shaded; scaled arbitrarily). The dashed blue lines indicate cuts. All other candidate selection cuts have been applied.

In addition, we place cuts on the energy of the π^0 and its constituent photons. We require that the energies of the two photons are at least 50 MeV in the lab frame, and that the π^0 candidate energy is at least 450 MeV in the CM frame. Distributions of the photon and π^0 energies are shown below.

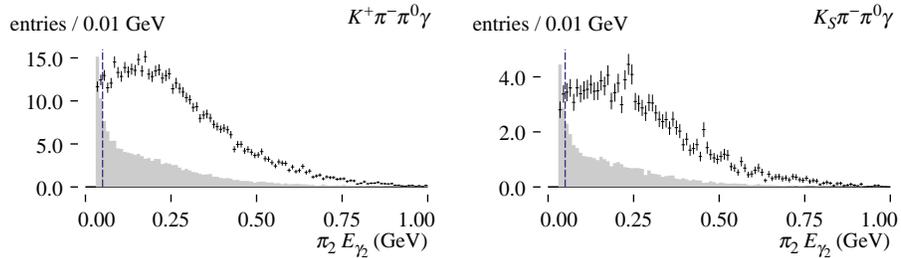


FIGURE 22. Distribution of the energy in the lab frame of the lower energy photon in π^0 candidates, in truth-matched signal MC (points) and $udsc$ MC (shaded; scaled arbitrarily). The dashed blue lines indicate cuts. All other candidate selection cuts have been applied.

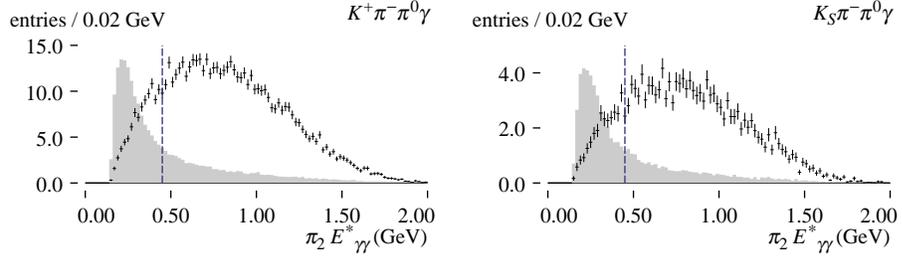


FIGURE 23. Distribution of the energy in the CM frame of the energy of π^0 candidates, in truth-matched signal MC (points) and $udsc$ MC (shaded; scaled arbitrarily). The dashed blue lines indicate cuts. All other candidate selection cuts have been applied.

Standard corrections have been calculated that tune the efficiency of the π^0 reconstruction in MC to match the efficiency in data. We re-weight MC samples with these efficiency corrections.

5.3.8 VERTEX SELECTION

We perform a geometric fit to the final state particles in our reconstructed B candidates, both charged and neutral, using the *TreeFitter* [23] fitting algorithm. This is a global fit to the entire B decay tree. For final states including a K_S^0 candidate, the fit uses the previously-fitted K_S^0 decay vertex. We cut on the fit probability of the B decay vertex P_{vtx} to reduce combinatoric backgrounds. We have found in our optimization procedure that selecting candidates with a vertex probability strictly greater than zero was optimal. Since this is sensitive to floating-point precision, we select candidates with a probability $P_{\text{vtx}} > 10^{-4}$. Vertex probability distributions are shown below; note the logarithmic vertical scale.

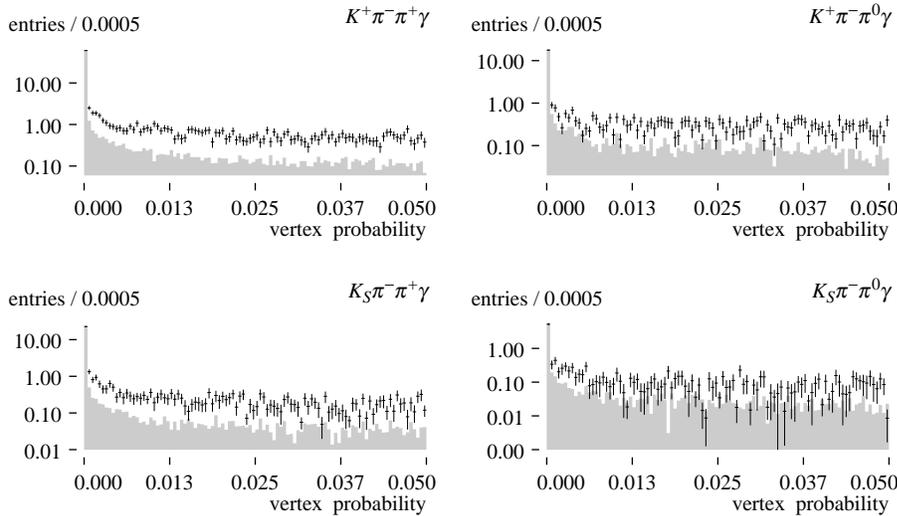


FIGURE 24. Distribution of the vertex probability for tracks from the B decay in truth-matched signal MC (points) and $udsc$ MC (shaded; scaled arbitrarily). All candidate selection criteria except for the cut on this quantity have been applied.

5.3.9 FISHER DISCRIMINANT

We construct a Fisher discriminant[24] to distinguish between $B\bar{B}$ and $udsc$ events. We compute separately for each mode a Fisher discriminant trained to distinguish candidates in the signal MC sample from candidates in the $udsc$ MC. The component variables of the Fisher discriminant are:

- $|\cos\theta_B^*|$, where θ_B^* is the polar angle in the CM frame of the B , i.e., the angle between the direction of the B candidate and the z axis;

- $|\cos \theta_{\text{thrust}}|$, where θ_{thrust} is the angle between the thrust axis of the B candidate and the thrust axis of the rest of the event;
- L_2/L_0 , where L_n is the n th Legendre moment around the B thrust axis of the rest of the event.

The Legendre moment is defined as,

$$L_n = \sum_i |\vec{p}_i^*| \cdot |\cos \theta_i^*|^n,$$

where the sum is over all charged and neutral particles in the event excluding those that comprise the reconstructed B candidate, \vec{p}^* is the CM momentum, and $\cos \theta^*$ is the CM polar angle.

In $udsc$ events, the momenta of the two lighter recoiling quarks tend to be large, and thus particles from the fragmentation of these quark pairs tend to produce two back-to-back jets, while in $B\bar{B}$ events, both B mesons are approximately at rest in the CM frame, and their decay products tend to produce spherical energy distributions. In jet-like $udsc$ events, the thrust axes of the B candidate and the rest of the event tend to be antiparallel, while in $B\bar{B}$ events, the directions of the thrust axes are uncorrelated. The value of L_2/L_0 is larger in $udsc$ events, which deviate from a spherical energy distribution.

We use a definition of the Fisher discriminant slightly different from that conventionally used. The Fisher discriminant is given by,

$$\mathcal{F}(\vec{x}) = \vec{x}^T \Sigma^{-1} (\vec{\mu}_s - \vec{\mu}_b)$$

where Σ is a dispersion matrix and $\vec{\mu}_s$ and $\vec{\mu}_b$ are the positions of the centroids of the two classes (i.e. signal and background). The Fisher discriminant assumes the dispersion matrices for the two classes are equal, $\Sigma_s = \Sigma_b \equiv \Sigma$, even though this is not generally the case. Conventionally, the weighted sum of Σ_s and Σ_b is used as the combined dispersion matrix Σ , or the two classes

are combined into one for computation of Σ . When calculating our Fisher discriminants, we have instead used the unweighted sum $\Sigma = \Sigma_s + \Sigma_b$. Since the MC statistics of our signal MC sample are much larger than the statistics of our *udsc* MC sample, this in effect weights the dispersion of the signal sample more heavily than in the conventional computation. We find that this method of constructing the Fisher discriminant produces a substantially better value of $S^2/(S + B)$ when we impose the optimal cut. We believe this is due to the fact that the optimal cut is in tail of the Fisher discriminant distribution for background events but in the heart of the distribution for signal events. We scale and shift the Fisher discriminant so that its distribution in the signal MC events has a centroid of zero and a root-mean-squared of one. The distributions of the quantities used to compute the Fisher discriminants are shown below, followed by distributions of the Fisher discriminants themselves.

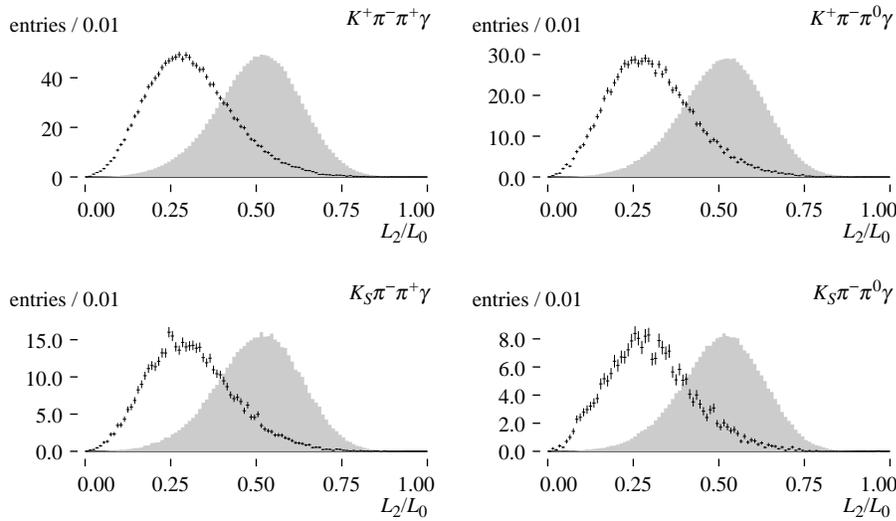


FIGURE 25. Distribution of the normalized second Legendre polynomial computed around the thrust axis in truth-matched signal MC (points) and *udsc* MC (shaded; scaled arbitrarily). All other selection cuts except for the Fisher discriminant cut have been applied.

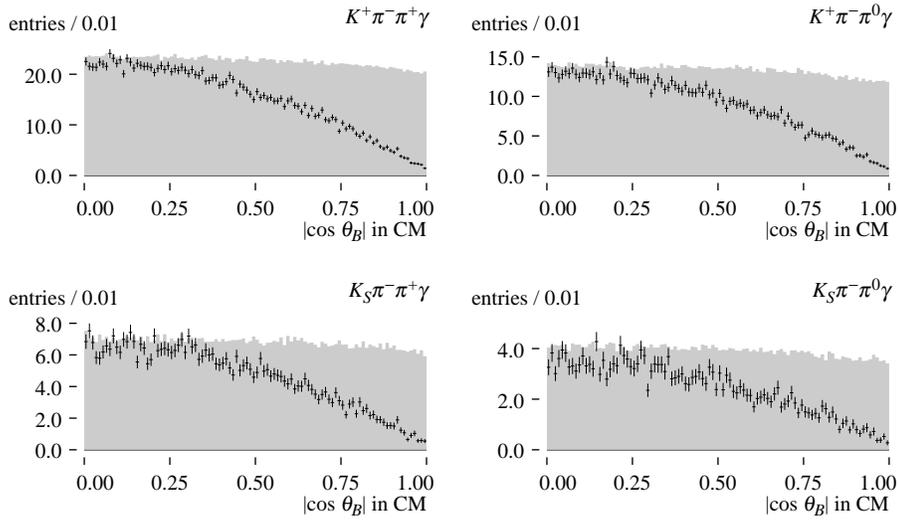


FIGURE 26. Distribution of the cosine of the polar angle of the reconstructed B in the CM frame in truth-matched signal MC (points) and $udsc$ MC (shaded; scaled arbitrarily). All other selection cuts except for the Fisher discriminant cut have been applied.

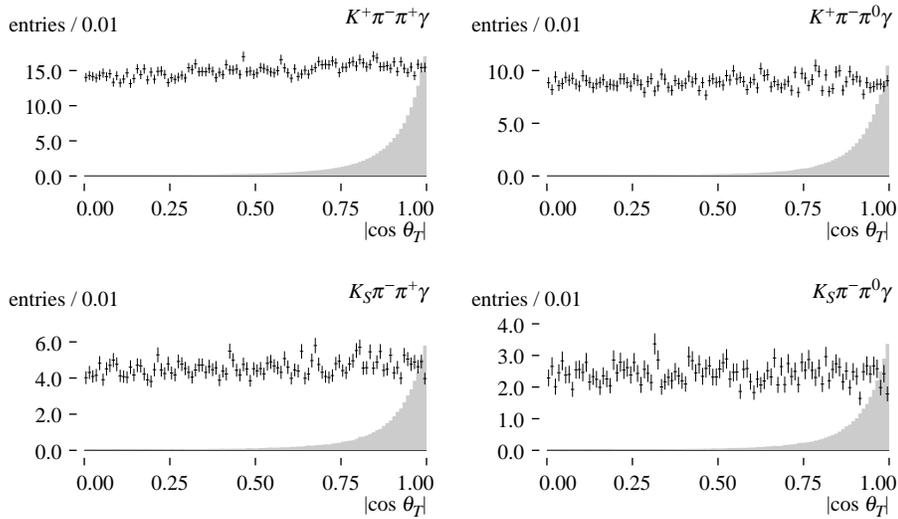


FIGURE 27. Distribution of the cosine of the thrust angle in truth-matched signal MC (points) and $udsc$ MC (shaded; scaled arbitrarily). All other selection cuts except for the Fisher discriminant cut have been applied.

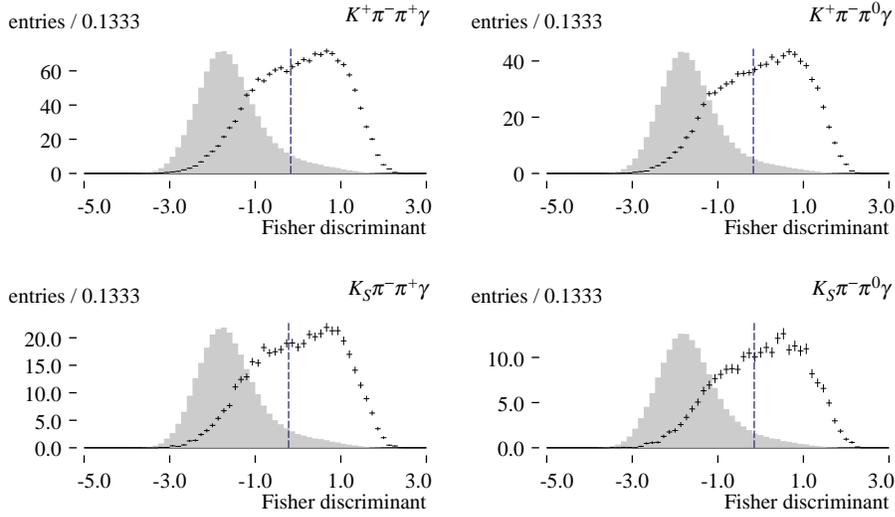


FIGURE 28. Distribution of the Fisher discriminants in truth-matched signal MC (points) and $udsc$ MC (shaded; scaled arbitrarily). The dashed blue lines indicate cuts. All other candidate selection cuts have been applied.

5.4 Cut optimization

We optimize the cuts listed in the previous subsection by maximizing the figure of merit $S^2/(S+B)$, where S is yield in truth-matched signal MC, and B is the yield in $udsc$ MC. We use $udsc$ MC because these processes are expected to be the major source of background.

We are primarily interested in optimizing the cuts to maximize the signal relative to the background in the region of m_{ES} and ΔE close to the signal peak; the branching fraction measurements we obtain from our fit procedure are not very sensitive to the number of background candidates outside this region. Therefore, we count S and B only inside an elliptical region given by,

$$\sqrt{\left(\frac{m_{ES} - m_B}{\sigma_{m_{ES}}}\right)^2 + \left(\frac{\Delta E}{\sigma_{\Delta E}}\right)^2} \leq 3$$

where $\sigma_{m_{ES}}$ and $\sigma_{\Delta E}$ are the widths of the signal distributions in the respective variables determined in fits to truth-matched signal MC events (see section 8.1). The region contains approximately 87% of signal MC events in all modes.

We optimize between six and thirteen continuous parameters (not including coefficients of the

Fisher discriminant), depending on the mode. These are,

- The size of the $m_{\gamma\gamma}$ window for the π^0 veto, and the minimum energy cut on the second photon used in the veto.
- The size of the $m_{\gamma\gamma}$ window for the η veto, and the minimum energy cut on the second photon used in the veto.
- The size of the $m_{\pi^+\pi^-}$ window for K_S^0 candidates.
- Minimum $K_S^0 \cos \theta_{\text{flight}}$.
- Minimum K_S^0 flight distance.
- Minimum K_S^0 vertex probability.
- The size of the symmetric $m_{\gamma\gamma}$ window for π^0 candidates.
- The minimum energy in the lab frame of the photons composing π^0 candidates.
- The minimum CM energy of π^0 candidates.
- Minimum B vertex probability.
- Minimum value of the Fisher discriminant.

Simultaneous optimization of these cuts requires us to minimize the (negative of the) figure of merit in a very high-dimensional space, a task that is challenging for minimization algorithms. Note that we did not optimize quality cuts on the high energy photon candidate.

In our experiments with the widely-used *Minuit* minimizer, we found that it was not able reliably to find a global maximum of the figure of merit, and would converge on a local maximum or wander slowly among several local maxima. Experiments with iterative sequential optimization of single cut parameters also produced discouraging results.

We instead used a simple implementation of minimization (of the negative figure of merit) by *simulated annealing* [25]. We give an outline of the algorithm to find the point \vec{x} that minimizes a function f here.

1. Set initial values for a small step size s and “temperature” T .
2. Initialize a starting point \vec{x} in the parameter space.
3. Choose a step $\delta\vec{x}$ of size s in a randomly-chosen direction, and evaluate $\delta f = f(\vec{x} + \delta\vec{x}) - f(\vec{x})$.
4. If $\delta f \leq 0$, update $\vec{x} \leftarrow \vec{x} + \delta\vec{x}$.
5. If $\delta f > 0$, update $\vec{x} \leftarrow \vec{x} + \delta\vec{x}$ with probability $P = e^{-\delta f/T}$. Otherwise, leave \vec{x} unchanged.
6. Decrease T slightly according to a preset *annealing schedule*.
7. If no step has been accepted in the last N tries, decrease s .
8. If s has achieved the target minimization scale, stop. Otherwise, return to step 3.

We choose as our annealing schedule $T \leftarrow 0.999 \times T$ at each step, and $N = 20$ as the threshold for reducing the step size. The initial value of T is chosen to be about 10% of the maximum variation of f by step size s around the initial point.

We ran five tries of the simulated annealing minimizer, with the same parameters and starting values but different random number sequences, for each optimization, and verified that all obtained the same minimum.

Simulated annealing produced more stable minimization runs than the other techniques we tried, but we still had difficulty verifying that the results represented global minima in the full multi-dimensional cut space. We therefore optimized the parameters several at a time, in this sequence:

1. We computed the Fisher discriminant with only the photon selection criteria applied (but still restricted to the elliptical region in $m_{ES}-\Delta E$ space).
2. We fixed the values of the minimum photon energy for the second leg of the π^0 and η veto cuts at reasonable values determined in previous attempts at cut optimization.
3. We performed simulated annealing minimization of each mode, varying the locations of all cuts. We minimized the figure of merit computed with the photon selection criteria and the charged PID selection applied (see below), and restricted to the elliptical region in $m_{ES}-\Delta E$ space.
4. We determined that the optimized mass windows for the π^0 and η veto were nearly the same in all modes, by examining the differential distributions of the figure of merit in these parameters. We similarly determined that the minimum photon energy cuts on the second leg were close to optimal.
5. We fixed the π^0 and η veto mass windows and the minimum distance to the closest cluster, and re-optimized, varying the remaining parameters. From the results, we determined the optimal cuts on π^0 candidates.
6. We fixed the cuts on the π^0 candidates as well, and minimized once again to determine the remaining parameters.

In most cases, we were able to adjust cuts to be uniform across modes without substantial loss in the figure of merit. Where possible, we also rounded cut values to numbers typically used in *BABAR* analyses.

We chose PID selection criteria for K^\pm and π^\pm candidates by computing the figures of merit for several choices. (The figures of merit were computed with a previous version of the other

candidate selection cuts.) Below are the figures of merit per mode for different choices of standard *BABAR* K^\pm PID selectors. The *KMicro* selectors impose progressively tighter cuts on dE/dx and the DIRC angle; the *KNNTight* selector combines these quantities in a neural net; the *KLHTight* selector uses likelihood ratios of these quantities. All candidate selection cuts except π^\pm PID selection are applied.

K^\pm PID selector	$K^+\pi^-\pi^+\gamma$	$K^+\pi^-\pi^0\gamma$
<i>KMicroLoose</i>	329.74	179.73
<i>KMicroTight</i>	346.23	194.72
<i>KMicroVeryTight</i>	339.61	189.68
<i>KNNTight</i>	337.04	183.80
<i>KLHTight</i>	337.52	186.38

TABLE 8.
 $S^2/(S+B)$ for choices of K^\pm particle ID, computed from signal and *udsc* MC. All other candidate selection cuts have been applied.

Likewise, below are the figures of merit per mode for different choices of π^\pm PID selectors. In π^+ π^- modes, the selection was applied to both charged pions. Again, a previous optimization of the other candidate selection cuts was used, and S and B were computed over the entire range of m_{ES} and ΔE in our n-tuples.

π^\pm PID selector	$K^+\pi^-\pi^+\gamma$	$K^+\pi^-\pi^0\gamma$	$K_S^0\pi^-\pi^+\gamma$	$K^+\pi^-\pi^0\gamma$
<i>piLHVeryLoose</i>	342.40	190.36	99.99	24.00
<i>piLHLoose</i>	346.23	194.72	102.18	24.54
<i>piLHTight</i>	337.82	192.68	98.90	24.33

TABLE 9.
 $S^2/(S+B)$ for choices of π^\pm particle ID, computed from signal and *udsc* MC. All other candidate selection cuts have been applied.

We choose the *KMicroTight* selector for charged kaons and the *piLHLoose* selector for charged pions.

5.5 Fit region

The fit region is the rectangular region in m_{ES} and ΔE in which we perform the fit to extract the signal yield. We impose the fit region selection after the candidate selection cuts, but before performing MCS.

The choice of the fit region affects the fit sensitivity not only by determining the domain of the fit distributions, but also by changing the efficiency of MCS, and thus the signal and background yields. These effects are potentially competing—by tightening the fit region, we may either improve or worsen the fit, increase the signal yield by improving MCS, or decrease the signal yield by cutting out signal events.

We determine the optimal choice of the fit region with a toy MC study¹. We consider a three-dimensional parameter space: the fit region is determined by the minimum cut on m_{ES} , and the minimum and maximum cuts on ΔE . The toy MC runs sample this parameter space on a grid, using all possible combinations of

- minimum m_{ES} cut of 5.20, 5.22, and 5.24 GeV/ c^2 ,
- minimum ΔE cut of -0.50 to -0.10 GeV in increments of 0.05 GeV, and
- maximum ΔE cut of 0.15 to 0.50 GeV in increments of 0.05 GeV.

The procedure for the toy MC study is as follows:

- We perform an unbinned maximum likelihood fit of a parameterized analytic function to the truth-matched signal MC over the full m_{ES} and ΔE range, with no MCS applied, to determine signal shape parameters. The fit function is described below.
- We fit the *udsc* MC to determine continuum shape parameters (see section 8.1.3).

¹This study was performed with a previous, slightly different version of the candidate selection cuts.

- We measure the distribution of feed-up, crossfeed, and feed-down backgrounds (see section 6.1) over the m_{ES} and ΔE range as a two-dimensional histogram.
- For each fit region, we perform 1,000 toy MC experiments:
 1. We compute the yields including analysis cuts, restriction to the fit region, and MCS for truth-matched signal and the various background processes.
 2. For each of the signal and background processes, we generate a sample of events according to the analytic fit or binned distributions, and restricted to the fit region in m_{ES} and ΔE . The number of events in each sample is generated from a Gaussian distribution with a mean at the yield measured in step 1, and standard deviation of the square root of the yield (to simulate counting statistics).
 3. We fit the combined sample as we would fit the data. The floating parameters are the signal and $udsc$ yields and the two $udsc$ shape parameters. The signal shape parameters are fixed. We use the binned distributions for the $B \rightarrow K\pi\pi\gamma$ and remaining $b \rightarrow s\gamma$ components and fixed the normalizations.
- For each fit region, we compute the mean and standard deviation over the ensemble of toy MC fits to estimate the fit bias and precision.

We parameterize the shape of the signal MC sample using a product of two Crystal Ball [26] functions, one in m_{ES} and one in ΔE . The parameterization in ΔE is augmented with a double Gaussian core along with the usual power-law tail, given by,

$$\text{CB}_2(x; \beta, \mu_1, \sigma_1, \mu_2, \sigma_2, \alpha, n) = \begin{cases} \frac{1}{a} \left(\frac{n}{\alpha}\right)^n \frac{\beta \exp\left\{\frac{\alpha^2}{2}\right\} + (1-\beta) \exp\left\{-\frac{(\mu_1 - \mu_2 - \alpha\sigma_1)^2}{2\sigma_2^2}\right\}}{((\mu_1 - x)/\sigma_1 + n/\alpha - \alpha)^n} & x < \mu_1 - \alpha\sigma_1 \\ \frac{1}{a} \left[\beta e^{\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2} + (1-\beta) e^{\frac{1}{2}\left(\frac{x-\mu_2}{\sigma_2}\right)^2} \right] & x > \mu_1 - \alpha\sigma_1, \end{cases}$$

where β is the fraction of the peak ascribed to the Gaussian with parameters μ_1, σ_1 , and $(1 - \beta)$ is the fraction of the peak ascribed to the Gaussian with parameters μ_2, σ_2 . The tail is a power-law function with exponent n , joined to the core at $\alpha\sigma_1$ below the peak. We fix μ_2 , as allowing it to float produces unstable fits without improving the likelihood of our fits. For the m_{ES} factor, we use an ordinary Crystal Ball shape with a single Gaussian core, equivalent to setting $\beta = 1$.

Our studies show that the fit precision does not depend strongly on the choice of the fit region. We choose tight cuts on ΔE for our fit region, in order to suppress background from the $b \rightarrow s\gamma$ processes. Our fit region is,

- $m_{\text{ES}} > 5.20 \text{ GeV}/c^2$
- $-0.15 < \Delta E < 0.15 \text{ GeV}$

We blinded the fit region in on-resonance data until the analysis methodology had been finalized. We define two ΔE sideband regions for control sample studies. These are given by $-0.50 < \Delta E < -0.15 \text{ GeV}$ and $0.15 < \Delta E < 0.50 \text{ GeV}$.

5.6 Multiple candidate selection

In this section, we describe our choice of techniques for selecting a single candidate from events in which more than one candidate is reconstructed and passes selection criteria. We only consider candidates that have passed the candidate selection cuts described earlier in this chapter, and that fall inside the fit region. We consider each mode independently, and count candidate multiplicity only within each mode. We make no requirement among multiple candidates reconstructed in different modes from the same event.

The following tables show the distribution of candidate multiplicity (the number of candidates reconstructed in an event) in the signal MC sample. The first table includes all events in which

one or more candidates are reconstructed; the second only includes events in which truth-matched candidates have been reconstructed.

Multiplicity	$K^+\pi^-\pi^+\gamma$	$K^+\pi^-\pi^0\gamma$	$K_S^0\pi^-\pi^+\gamma$	$K_S^0\pi^+\pi^0\gamma$
1	987	617	314	186
2	183	126	57	49
3	28	22	9	12
4	7	7	2	4
≥ 5	3	3	1	3

TABLE 10. Candidate multiplicity distribution (the number of events with each candidate multiplicity) scaled to analysis luminosity, for signal MC events in which at least one or more candidates were reconstructed. All selection cuts and the fit region cut have been applied.

Multiplicity	$K^+\pi^-\pi^+\gamma$	$K^+\pi^-\pi^0\gamma$	$K_S^0\pi^-\pi^+\gamma$	$K_S^0\pi^+\pi^0\gamma$
1	623	381	192	93
2	124	79	38	27
3	18	14	6	6
4	4	4	1	2
≥ 5	2	2	0	1

TABLE 11. Candidate multiplicity distribution (the number of events with each candidate multiplicity) scaled to analysis luminosity, for those signal MC events with in which truth-matched candidates were reconstructed. All selection cuts and the fit region cut have been applied.

Here and below, we consider each mode independently. Only candidates reconstructed in the same mode are counted in the multiplicity. We do not seek to eliminate the possibility that candidates are reconstructed in two or more different modes in the same event.

The following table summarizes the strategies we considered for selecting a single candidate from each event.

Name	Modes	Criterion	Tie-breaker
<i>random</i>	all	random	
<i>delta_e</i>	all	min $ \Delta E $	
<i>vtx_prob</i>	all but $K_S^0 \pi^+ \pi^0 \gamma$	max P_{vtx}	random
<i>vtx_prob+pi0_mass</i>	π^0 modes	max P_{vtx}	min $ m_{\gamma\gamma} - m_{\pi^0} $
<i>vtx_prob+ks_mass</i>	K_S^0 modes	max P_{vtx}	min $ m_{\pi^+ \pi^-} - m_{K_S^0} $
<i>vtx_prob+ks_vtx_prob</i>	K_S^0 modes	max P_{vtx}	max $K_S^0 P_{\text{vtx}}$
<i>pi0_mass</i>	$K_S^0 \pi^+ \pi^0 \gamma$	min $ m_{\gamma\gamma} - m_{\pi^0} $	random
<i>pi0_mass+ks_vtx_prob</i>	$K_S^0 \pi^+ \pi^0 \gamma$	min $ m_{\gamma\gamma} - m_{\pi^0} $	max $K_S^0 P_{\text{vtx}}$

TABLE 12. Descriptions of alternative MCS techniques.

We expect MCS using K_S^0 or π^0 invariant mass or K_S^0 vertex probability to work well if misreconstructed candidates tend to include fake K_S^0 or π^0 candidates. MCS using ΔE or B vertex probability would fare better than these at removing wrong combinations of correctly-identified particles, for instance a $K_S^0 \pi^- \pi^+ \gamma$ candidate which includes a real K_S^0 from the other B in the event. Inspection of MC tables truth for signal events indicates that most wrong combinations involve correctly-identified candidates from the other B in the event.

In evaluating MCS techniques, the important figure of merit is efficiency on signal events—how well the technique picks the true candidate over the wrong alternatives. The MCS efficiency is the yield of truth-matched signal MC candidates after all other candidate selection cuts, the fit region cut, and MCS, divided by the number of signal MC *events* in which the truth-matched candidate (along with zero or more other candidates) passes all other cuts. Signal events in which the correct $K\pi\pi\gamma$ combination is not reconstructed do not enter into the efficiency at all, nor do other background processes, since for these we do not care which candidate is chosen.

The efficiency of any MCS technique depends on the candidate multiplicity; the more incorrect combinations are removed before MCS, the better it will perform. We therefore first impose all other candidate selection cuts and also restrict candidates to the fit region in m_{ES} and ΔE . We evaluate

MCS techniques for various choices of the m_{ES} and ΔE region by computing the efficiency for signal MC events as a function of the minimum m_{ES} cut and minimum and maximum ΔE cuts. The table below shows the MCS efficiency for each technique for the full m_{ES} - ΔE region and illustrative loose and tight fit regions.

Mode	MCS Technique	MCS efficiency for:		
		$m_{ES} > 5.20$ $ \Delta E < 0.5$	$m_{ES} > 5.20$ $ \Delta E < 0.3$	$m_{ES} > 5.20$ $ \Delta E < 0.15$
$K^+\pi^-\pi^+\gamma$	<i>random</i>	78.8%	84.0%	90.4%
	<i>delta_e</i>	93.1%	93.2%	93.7%
	<i>vtx_prob</i>	82.9%	87.4%	92.4%
$K^+\pi^-\pi^0\gamma$	<i>random</i>	80.2%	84.5%	90.1%
	<i>delta_e</i>	92.4%	92.5%	92.9%
	<i>vtx_prob</i>	82.1%	86.3%	91.0%
	<i>vtx_prob+pi0_mass</i>	84.5%	88.2%	92.4%
$K_S^0\pi^-\pi^+\gamma$	<i>random</i>	76.9%	82.6%	89.1%
	<i>delta_e</i>	92.3%	92.4%	92.9%
	<i>vtx_prob</i>	81.7%	86.3%	91.6%
	<i>vtx_prob+ks_mass</i>	81.7%	86.3%	91.6%
	<i>vtx_prob+ks_vtx_prob</i>	81.7%	86.4%	91.6%
$K_S^0\pi^+\pi^0\gamma$	<i>random</i>	70.7%	76.2%	84.1%
	<i>delta_e</i>	87.7%	87.8%	88.4%
	<i>pi0_mass</i>	72.5%	77.8%	85.5%
	<i>pi0_mass+ks_vtx_prob</i>	72.3%	78.3%	85.9%

TABLE 13. Efficiencies of MCS techniques for full, loose, and tight regions in m_{ES} and ΔE . MCS efficiency is the fraction of signal MC events containing a truth-matched candidate, after selections cuts and the fit region cut have been applied, in which the MCS procedure chooses the true candidate. Note that these results were generated with a previous and slightly different version of the candidate selection cuts imposed.

We conclude that

- MCS using ΔE is most effective,
- random MCS is least effective,
- MCS based on B vertex probability performs acceptably,
- in π^0 modes, $m_{\gamma\gamma}$ improves MCS, and
- in K_S^0 modes, neither $m_{\pi^+\pi^-}$ nor the K_S^0 vertex probability improves MCS.

Because we fit the ΔE distribution to extract the yield of signal events, however, we are sensitive to any distortion created by our selection procedure. Selecting the best candidate based on $|\Delta E|$ produces such a distortion, so we exclude this MCS method. We choose *vtx_prob* in the $K^+\pi^-\pi^+\gamma$ and $K_S^0\pi^-\pi^+\gamma$ modes, *vtx_prob+pi0_mass* in the $K^+\pi^-\pi^0\gamma$ mode, and *pi0_mass* in the $K_S^0\pi^+\pi^0\gamma$ mode. The efficiency of MCS depends on the choice of the fit region, which is imposed before MCS is performed.

5.7 Efficiency

This table lists the last-cut efficiency—the efficiency of each cut after all the other listed cuts have been applied—of the candidate selection cuts in truth-matched signal MC. The efficiencies are computed for events passing sample selection in the region MC for $m_{ES} > 5.20 \text{ GeV}/c^2$, $|\Delta E| < 0.5 \text{ GeV}$, and $m_{K\pi\pi} < 1.8 \text{ GeV}/c^2$.

Cut	$K^+\pi^-\pi^+\gamma$	$K^+\pi^-\pi^0\gamma$	$K_S\pi^-\pi^+\gamma$	$K_S\pi^-\pi^0\gamma$
Fisher	58.0%	58.3%	59.0%	58.3%
B vertex probability	94.0%	97.0%	93.0%	96.8%
second moment	99.2%	99.3%	99.2%	99.2%
$\cos\theta_\gamma$	98.3%	98.7%	98.6%	99.0%
bump distance	97.6%	98.0%	97.6%	98.3%
π^0 veto	94.0%	94.2%	94.2%	94.5%
η veto	96.7%	96.9%	96.7%	96.5%
K^\pm PID	85.0%	85.2%		
K_S^0 mass			97.5%	97.4%
K_S^0 decay length			97.6%	97.9%
$K_S^0 \cos\theta_{\text{flight}}$			98.9%	92.4%
π^\pm PID for π_1	98.9%	99.0%	98.8%	98.9%
π^\pm PID for π_2	98.7%		98.8%	
π^0 mass		92.9%		92.7%
$\pi^0 E_\gamma$		95.6%		95.6%
$\pi^0 E_\gamma^*$		83.3%		82.6%

TABLE 14.
Last-cut efficiency of candidate selection cuts in truth-matched signal MC. Uncertainties are due to MC statistics.

The selection criteria of the *BtoXGamma* skim are more than 99.9% efficient for truth-matched candidates reconstructed in signal MC samples in all modes. We therefore neglect the skim selection in our efficiency calculations.

Efficiencies in signal MC for the fit region requirement, as well as for MCS (which is applied after the fit region only), are summarized below. The first line in each table is the fraction of signal events generated in each mode in which we reconstruct a candidate that passes our sample selection and is truth-matched; this includes the aggregate reconstruction efficiency for all final state particles, the implicit geometrical acceptance, and the loose sample selection cuts described in section 5.2. The truth match requirement is always imposed after all other selection requirements. The first table lists the efficiency of each cut relative to the previous. The second lists cumulative efficiencies of cuts as applied in the stated order; the bottom line shows our overall selection efficiency in truth-matched signal MC.

Cut	$K^+\pi^-\pi^+\gamma$	$K^+\pi^-\pi^0\gamma$	$K_S\pi^-\pi^+\gamma$	$K_S\pi^-\pi^0\gamma$
reconstruction	$(46.25 \pm 0.08)\%$	$(29.34 \pm 0.07)\%$	$(39.34 \pm 0.14)\%$	$(24.53 \pm 0.11)\%$
candidate selection	$(33.98 \pm 0.12)\%$	$(25.90 \pm 0.12)\%$	$(36.31 \pm 0.22)\%$	$(24.24 \pm 0.23)\%$
fit region	$(90.22 \pm 0.12)\%$	$(88.69 \pm 0.18)\%$	$(90.59 \pm 0.22)\%$	$(88.51 \pm 0.34)\%$
MCS	$(92.51 \pm 0.12)\%$	$(90.85 \pm 0.17)\%$	$(91.53 \pm 0.22)\%$	$(86.38 \pm 0.39)\%$

TABLE 15. Efficiency in truth-matched signal MC. The efficiency of each cut is relative to the sample accepted by the previous. Uncertainties are due to MC statistics.

Cut	$K^+\pi^-\pi^+\gamma$	$K^+\pi^-\pi^0\gamma$	$K_S\pi^-\pi^+\gamma$	$K_S\pi^-\pi^0\gamma$
reconstruction	$(46.25 \pm 0.08)\%$	$(29.34 \pm 0.07)\%$	$(39.34 \pm 0.14)\%$	$(24.53 \pm 0.11)\%$
candidate selection	$(15.71 \pm 0.06)\%$	$(7.60 \pm 0.04)\%$	$(14.29 \pm 0.10)\%$	$(5.95 \pm 0.06)\%$
fit region	$(14.18 \pm 0.06)\%$	$(6.74 \pm 0.04)\%$	$(12.94 \pm 0.10)\%$	$(5.26 \pm 0.06)\%$
MCS	$(13.12 \pm 0.06)\%$	$(6.12 \pm 0.04)\%$	$(11.85 \pm 0.09)\%$	$(4.55 \pm 0.05)\%$

TABLE 16. Cumulative efficiencies in truth-matched signal MC. Uncertainties are due to MC statistics.

The following tables present the efficiency after all selection criteria, broken down by the generated resonance, and by the run conditions of the signal MC sample.

resonance	$K^+\pi^-\pi^+\gamma$	$K^+\pi^-\pi^0\gamma$	$K_S\pi^-\pi^+\gamma$	$K_S\pi^-\pi^0\gamma$
$K_1(1270)$	$(13.23 \pm 0.13)\%$	$(6.13 \pm 0.08)\%$	$(11.76 \pm 0.21)\%$	$(4.66 \pm 0.11)\%$
$K_1(1400)$	$(13.16 \pm 0.12)\%$	$(6.05 \pm 0.08)\%$	$(12.07 \pm 0.20)\%$	$(4.45 \pm 0.12)\%$
$K^*(1410)$	$(13.28 \pm 0.12)\%$	$(6.10 \pm 0.07)\%$	$(11.80 \pm 0.19)\%$	$(4.53 \pm 0.11)\%$
$K_2^*(1430)$	$(12.71 \pm 0.25)\%$	$(5.87 \pm 0.15)\%$	$(11.64 \pm 0.40)\%$	$(4.31 \pm 0.22)\%$
$K^*(1680)$	$(12.94 \pm 0.17)\%$	$(6.37 \pm 0.10)\%$	$(12.01 \pm 0.29)\%$	$(4.75 \pm 0.14)\%$

TABLE 17. Overall efficiency by generated resonance in signal MC. Uncertainties are due to MC statistics.

run	$K^+\pi^-\pi^+\gamma$	$K^+\pi^-\pi^0\gamma$	$K_S\pi^-\pi^+\gamma$	$K_S\pi^-\pi^0\gamma$
Run 1	$(12.75 \pm 0.27)\%$	$(6.20 \pm 0.17)\%$	$(11.53 \pm 0.43)\%$	$(4.49 \pm 0.24)\%$
Run 2	$(13.64 \pm 0.18)\%$	$(6.25 \pm 0.11)\%$	$(11.85 \pm 0.28)\%$	$(4.75 \pm 0.16)\%$
Run 3	$(13.84 \pm 0.25)\%$	$(6.33 \pm 0.15)\%$	$(12.63 \pm 0.41)\%$	$(4.59 \pm 0.22)\%$
Run 4	$(12.77 \pm 0.12)\%$	$(5.91 \pm 0.08)\%$	$(11.57 \pm 0.20)\%$	$(4.53 \pm 0.11)\%$

TABLE 18. Overall efficiency by run in signal MC. Uncertainties are due to MC statistics.

The following figures show the overall signal efficiency binned in $m_{K\pi\pi}$.

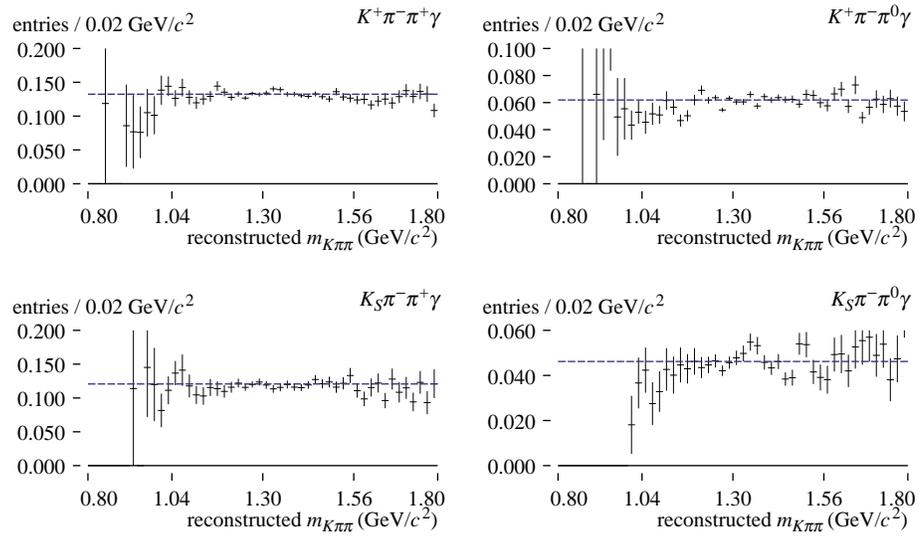


FIGURE 29. Overall efficiency in signal MC as a function of $m_{K\pi\pi}$. The dashed blue line shows the average efficiency.

5.8 Resolution

We estimate the resolution of $m_{K\pi\pi}$ by comparing the reconstructed value to the value generated by the MC generator. The figure below shows the distribution of the difference between the reconstructed and generated value for truth matched signal MC candidates.

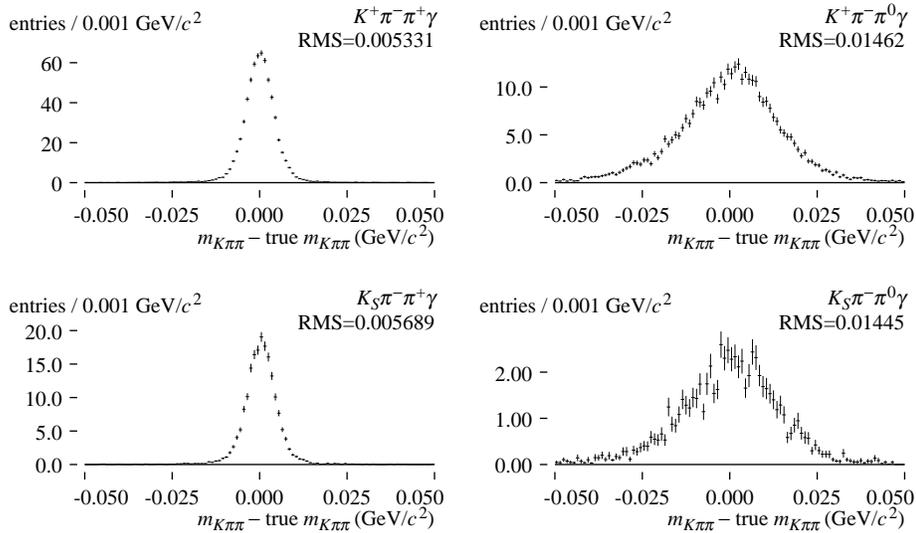


FIGURE 30. Distribution of the difference between the reconstructed and generated values of $m_{K\pi\pi}$ in truth-matched signal MC. Candidate selection and MCS have been applied.

We take the RMS values of these distributions as the $m_{K\pi\pi}$ resolutions. For each of the four modes, the mean of the distribution is much smaller than the RMS.

6 Background Processes

We consider three categories of background processes:

1. The $b \rightarrow s\gamma$ processes that include our signal modes are also a major source of background candidates, due to the presence of a high-energy photon. This photon combines with tracks from the fragmentation of the hadronic part of the decay products and/or tracks from the decay of the other B in the event to form background candidates.
2. While we do not expect the total yield of background events from B decays to be large, potentially peaking background processes are of particular concern. These are processes in which the decay products of a single B can be misreconstructed as a signal candidate, without the addition of tracks from the other B in the event. These are of concern, since their distributions in m_{ES} and ΔE are similar to those of signal events, and thus may contribute spuriously to signal yields in the fit.
3. Combinatoric backgrounds are candidates formed from random combinations of tracks and photons from $udsc$ pair production or $B\bar{B}$ decays. These processes contribute the largest number of background candidates to the analysis, primarily from $udsc$. However, these candidates can be separated in the fit because their distribution in m_{ES} and ΔE is not shaped like the distribution of signal candidates.

Studies of these background processes are presented in the following sections.

6.1 $b \rightarrow s\gamma$ background

Radiative penguin $b \rightarrow s\gamma$ processes not only include the signal modes in this analysis, but are a source of background as well. The rarest component of the signature of our signal is the high-energy photon; kaons and pions are plentiful in B decays. Radiative penguin processes are a dangerous background because they are a source of high-energy photons.

We can divide these background processes into three groups. First, our analysis can misreconstruct events produced in $B \rightarrow K\pi\pi\gamma$ decays, by replacing one or more kaon or pion in the final state with candidates from the other side of the event. These candidates are called *crossfeed* candidates. A crossfeed candidate may be reconstructed in the same mode in which it was produced, or a different mode. In the former case, it is called *self-crossfeed*. Second, $B \rightarrow K\pi\gamma$ processes, specifically $B \rightarrow K^*(892)\gamma$ and $B \rightarrow K_2^*(1430)\gamma$, can produce background candidates in which an additional track from the decay of the other B is included; this is called *feed-up*. Finally, other $b \rightarrow s\gamma$ processes with higher-multiplicity final states can produce background candidates; these are called *feed-down*.

The effects of MCS are different on background candidates from these three processes. In a $B \rightarrow K\pi\pi\gamma$ event in which our analysis reconstructs multiple candidates of which one is the true $B \rightarrow K\pi\pi\gamma$ candidate, it is MCS that determines whether this event is a “signal event” or “background event,” depending on whether it selects the true candidate or another candidate. Thus, the choice of MCS affects the $B \rightarrow K\pi\pi\gamma$ self-crossfeed background rate. In $b \rightarrow s\gamma$ events other than $B \rightarrow K\pi\pi\gamma$ containing background candidates, however, all candidates are background, and it is not a concern which one MCS chooses—we are always left with one background candidate per event.

6.1.1 YIELDS

We estimate the background yields from the three different $b \rightarrow s\gamma$ components separately:

- We estimate the background from $B \rightarrow K\pi\gamma$ using a cocktail of $B \rightarrow K^*(892)\gamma$ and $B \rightarrow K_2^*(1430)\gamma$ MC samples. The distributions are shown in Figure 31.
- We estimate the background from $B \rightarrow K\pi\pi\gamma$ from the exclusive signal MC cocktail, with signal candidates vetoed by MC truth. The distributions are shown in Figure 32.
- We use the inclusive $b \rightarrow s\gamma$ MC sample to estimate rates of remaining $b \rightarrow s\gamma$ backgrounds. We have removed from this sample all $B \rightarrow K\pi\gamma$ and $B \rightarrow K\pi\pi\gamma$ events, based on MC truth information. The distributions are shown in Figure 33.

Systematic uncertainties in the choice of models for $b \rightarrow s\gamma$ backgrounds are discussed in section 10.4.

The table below lists total background yields from $b \rightarrow s\gamma$ processes after all cuts, including the fit region cut and MCS.

Mode	$B \rightarrow K\pi\gamma$	$B \rightarrow K\pi\pi\gamma$	other $b \rightarrow s\gamma$
$K^+ \pi^- \pi^+ \gamma$	213.7 ± 2.3	383.6 ± 3.4	308.7 ± 11.2
$K^+ \pi^- \pi^0 \gamma$	83.0 ± 1.4	282.7 ± 2.9	219.0 ± 8.5
$K_S^0 \pi^- \pi^+ \gamma$	64.5 ± 1.2	133.4 ± 2.0	115.2 ± 6.8
$K_S^0 \pi^+ \pi^0 \gamma$	39.2 ± 1.0	121.6 ± 1.9	107.9 ± 6.1

TABLE 19. Yields in the fit region scaled to data luminosity from the three categories of $b \rightarrow s\gamma$ background processes. Candidate selection, the fit region cut, MCS, and MC truth veto have been applied.

We also separate the $B \rightarrow K\pi\pi\gamma$ background candidates by the process with which the background event was generated. Those for which the production process is the same as the reconstructed mode—but not correctly reconstructed, since an MC truth veto is applied—are self-crossfeed candidates. The table below breaks down yields from $B \rightarrow K\pi\pi\gamma$ based on the generation process.

Generation mode	$K^+\pi^-\pi^+\gamma$	$K^+\pi^-\pi^0\gamma$	$K_S^0\pi^-\pi^+\gamma$	$K_S^0\pi^+\pi^0\gamma$
$K^+\pi^-\pi^+\gamma$	180.1 ± 2.3	43.7 ± 1.1	5.7 ± 0.4	0.3 ± 0.1
$K^+\pi^-\pi^0\gamma$	177.8 ± 2.3	187.7 ± 2.4	2.6 ± 0.3	3.8 ± 0.3
$K_S\pi^-\pi^+\gamma$	5.5 ± 0.4	0.4 ± 0.1	61.2 ± 1.3	23.8 ± 0.8
$K_S\pi^-\pi^0\gamma$	1.0 ± 0.2	4.6 ± 0.4	57.6 ± 1.3	72.9 ± 1.5
$K^+\pi^0\pi^0\gamma$	4.4 ± 0.4	33.3 ± 0.9	0.0 ± 0.0	0.4 ± 0.1
$K_S\pi^0\pi^0\gamma$	0.1 ± 0.0	0.5 ± 0.1	1.7 ± 0.2	16.2 ± 0.7
$K_L\pi^-\pi^+\gamma$	10.7 ± 0.6	1.1 ± 0.2	3.4 ± 0.3	0.4 ± 0.1
$K_L\pi^-\pi^0\gamma$	1.1 ± 0.2	8.5 ± 0.5	0.3 ± 0.1	2.1 ± 0.2
$K_L\pi^0\pi^0\gamma$	0.0 ± 0.0	0.3 ± 0.1	0.0 ± 0.0	0.1 ± 0.0
$K_S\pi^-\pi^+\gamma, K_S \rightarrow \pi^0\pi^0$	2.6 ± 0.3	3.0 ± 0.3	0.8 ± 0.1	1.2 ± 0.2
$K_S\pi^-\pi^0\gamma, K_S \rightarrow \pi^0\pi^0$	0.2 ± 0.1	4.4 ± 0.4	0.0 ± 0.0	1.2 ± 0.2
$K_S\pi^0\pi^0\gamma, K_S \rightarrow \pi^0\pi^0$	0.0 ± 0.0	0.1 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
$K_S\pi^-\pi^+\gamma, K_S \rightarrow \text{other}$	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
$K_S\pi^-\pi^0\gamma, K_S \rightarrow \text{other}$	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
$K_S\pi^0\pi^0\gamma, K_S \rightarrow \text{other}$	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Total	383.6 ± 3.4	287.7 ± 2.9	133.3 ± 2.0	122.4 ± 1.9

TABLE 20. Yields in truth-vetoed signal MC scaled to luminosity broken down by MC generation mode. Candidate selection cuts, the fit region cut, and MCS have been applied.

The following figures show the m_{ES} and ΔE distributions of the three $b \rightarrow s\gamma$ background components.

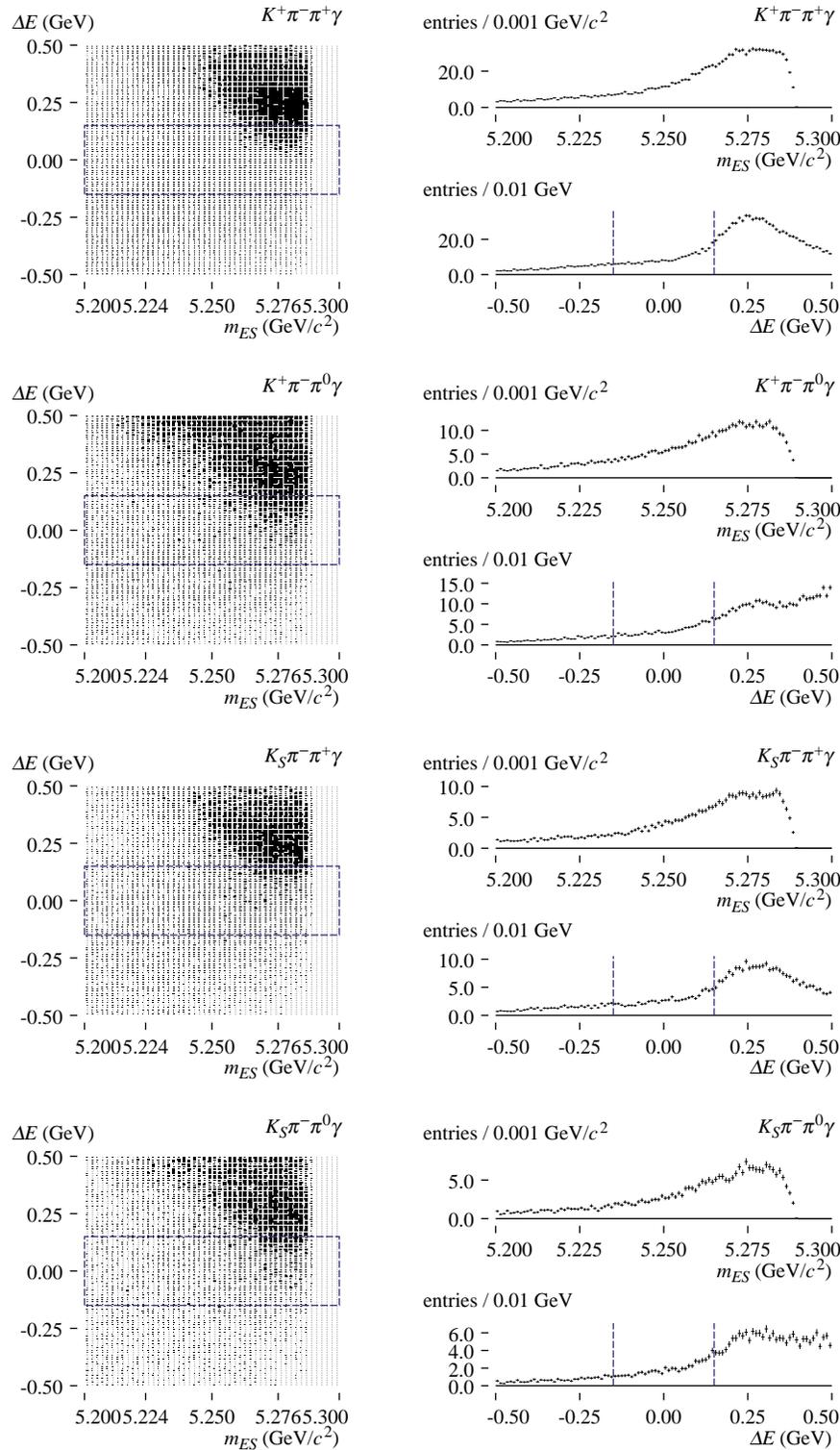


FIGURE 31. Joint and marginal distributions of m_{ES} and ΔE in the $B \rightarrow K\pi\gamma$ cocktail MC sample. Candidate selection has been applied, but not the fit region cut or MCS. The fit region is indicated in blue.

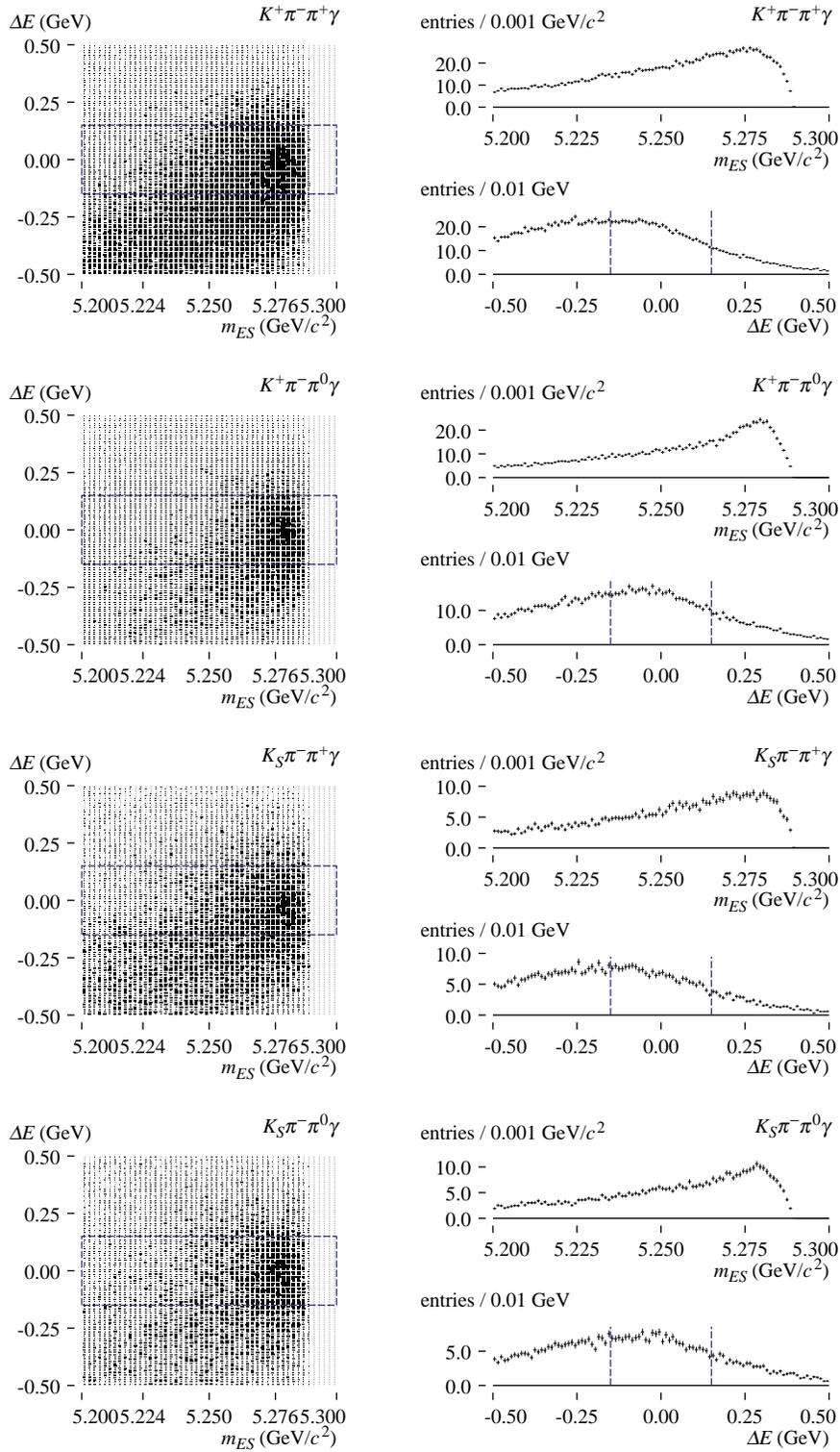


FIGURE 32. Joint and marginal distributions of m_{ES} and ΔE in the signal cocktail MC sample with MC truth veto applied. Candidate selection has been applied, but not the fit region cut or MCS. The fit region is indicated in blue.

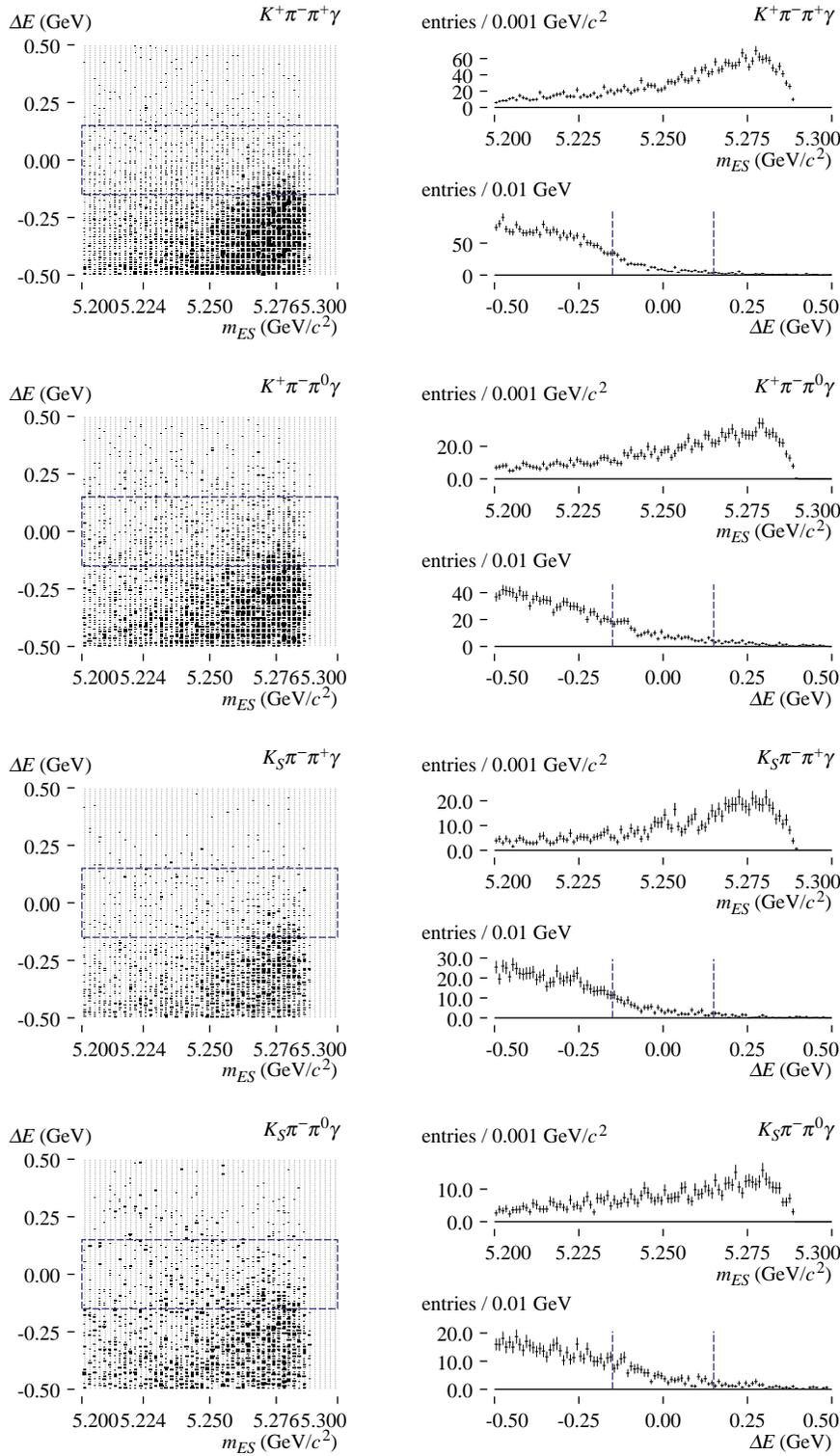


FIGURE 33. Joint and marginal distributions of m_{ES} and ΔE in inclusive $b \rightarrow s\gamma$ MC sample with $B \rightarrow K\pi\gamma$ and $B \rightarrow K\pi\pi\gamma$ events removed. Candidate selection has been applied, but not the fit region cut or MCS. The fit region is indicated in blue.

6.2 Backgrounds from $B \rightarrow K\pi\pi\pi^0$ and $B \rightarrow K\pi\pi\eta$

We have studied several additional exclusive B decays that are potentially sources of peaking backgrounds. These are B decays with $K\pi\pi\pi^0$ or $K\pi\pi\eta$ final states, in which the π^0 or η is misreconstructed as a photon. While we explicitly veto photons from π^0 and η decays, some of these will nevertheless slip past the vetoes. This can occur if the two EMC clusters from $\pi^0 \rightarrow \gamma\gamma$ or $\eta \rightarrow \gamma\gamma$ merge, or if one photon is not detected. However, because of our kinematic selection requirements, a high-energy π^0 or η must be produced.

When studying these modes, we pay special attention to candidates in which the kaon, both pions, and photon all were produced from the same B , as determined from MC truth. We present distributions and yields for these ‘‘peaking’’ candidates, as well as distributions for all candidates from these samples. Only the former are of concern, since we assume the latter can be included with other generic B decays.

In the distributions shown below for these processes, we show the total background contributions in gray, with backgrounds from candidates reconstructed from a single B drawn as black points with error bars. Candidate selection and the fit region cut are applied, except for cuts (indicated in blue) on the quantity shown in each plot.

6.2.1 $B^0 \rightarrow \bar{D}^0\pi^0$ AND $B^0 \rightarrow \bar{D}^0\eta$

The most dangerous processes in this category are $B^0 \rightarrow \bar{D}^0\pi^0$ and $B^0 \rightarrow \bar{D}^0\eta$, because of their relatively large branching fractions, respectively 2.91×10^{-4} and 2.2×10^{-4} . The π^0 or η from the two-body decay of the B can easily produce a photon candidate in the energy range required by this analysis. The D undergoes a three-body decay, either $\bar{D}^0 \rightarrow K_S^0\pi^+\pi^-$, reconstructed in the $K_S^0\pi^-\pi^+\gamma$ signal mode, or $\bar{D}^0 \rightarrow K^+\pi^-\pi^0$, reconstructed as $K^+\pi^-\pi^0\gamma$.

In the $K^+\pi^-\pi^0\gamma$ signal mode, we find 5.6 peaking background candidates in the fit region,

after candidate selection and MCS, contributed by $B^0 \rightarrow \bar{D}^0\pi^0$ and 2.9 from $B^0 \rightarrow \bar{D}^0\eta$. The distributions peak slightly below zero in ΔE , as shown below.

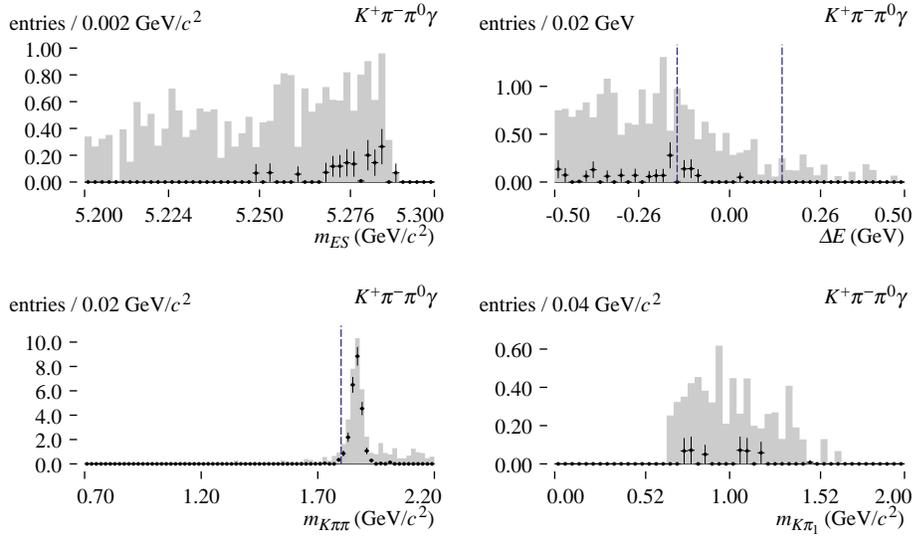


FIGURE 34. Distributions of m_{ES} , ΔE , and invariant masses in peaking background candidates generated as $B^0 \rightarrow \bar{D}^0\pi^0$ and reconstructed in mode $K^+\pi^-\pi^0\gamma$. All candidates passing candidate selection cuts are in gray; those reconstructed from a single B are in black.

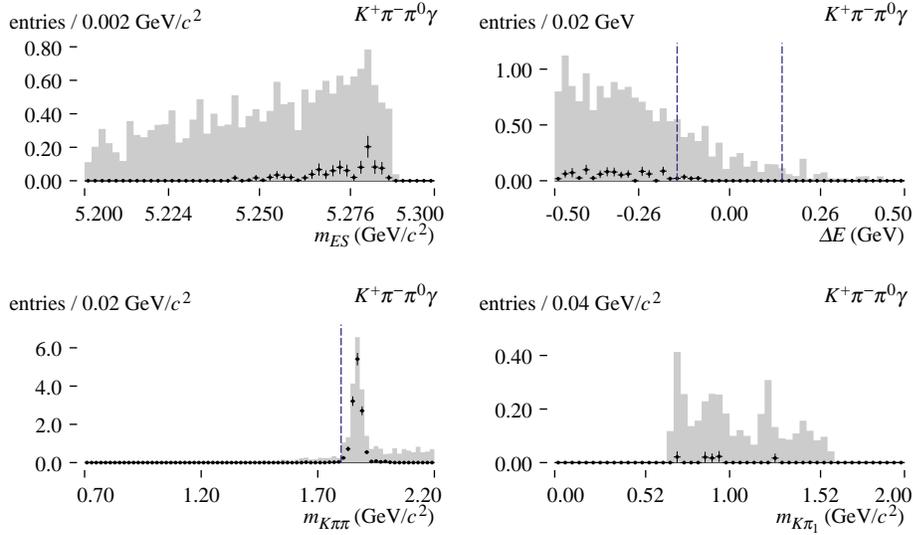


FIGURE 35. Distributions of m_{ES} , ΔE , and invariant masses in peaking background candidates generated as $B^0 \rightarrow \bar{D}^0\eta$ and reconstructed in mode $K^+\pi^-\pi^0\gamma$. All candidates passing candidate selection cuts are in gray; those reconstructed from a single B are in black.

In the $K_S^0 \pi^- \pi^+ \gamma$ signal mode, we find 1.6 peaking background events in the fit region, after analysis and MCS, contributed by $B^0 \rightarrow \bar{D}^0 \pi^0$ and 1.1 from $B^0 \rightarrow \bar{D}^0 \eta$. The distributions peak slightly below zero in ΔE , as shown below.

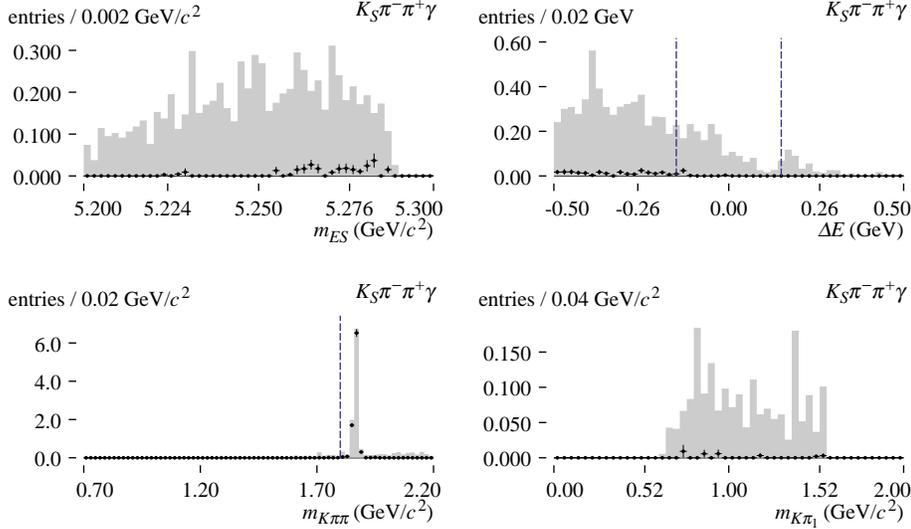


FIGURE 36. Distributions of m_{ES} , ΔE , and invariant masses in background candidates generated as $B^0 \rightarrow \bar{D}^0 \pi^0$ and reconstructed in mode $K_S^0 \pi^- \pi^+ \gamma$. All candidates passing candidate selection cuts are in gray; those reconstructed from a single B are in black.

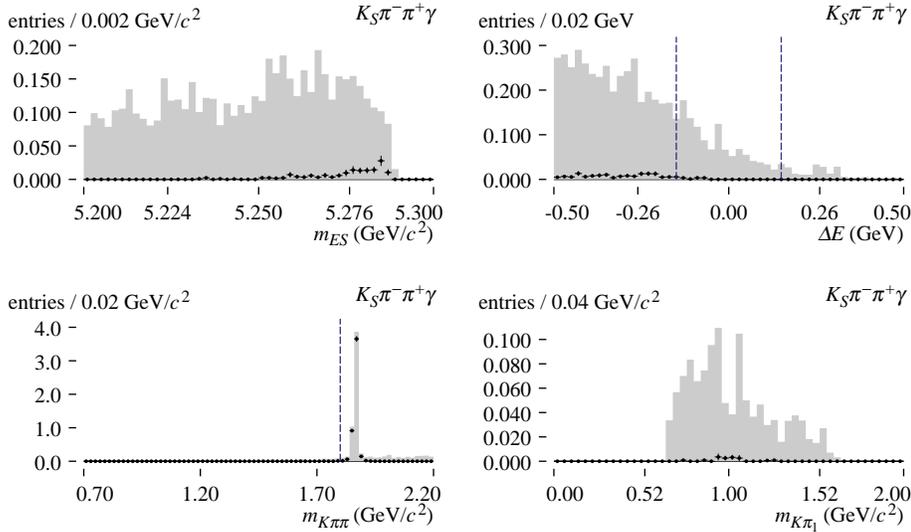


FIGURE 37. Distributions of m_{ES} , ΔE , and invariant masses in background candidates generated as $B^0 \rightarrow \bar{D}^0 \eta$ and reconstructed in mode $K_S^0 \pi^- \pi^+ \gamma$. All candidates passing candidate selection cuts are in gray; those reconstructed from a single B are in black.

6.2.2 $B \rightarrow D \rho, D \rightarrow K \pi$

We consider also the decays $B \rightarrow D \rho$, which have large branching fractions. For these processes to produce backgrounds in our analysis, however, a π^0 from the secondary decay of either the D or ρ must fake our high-energy photon. The $B \rightarrow D \rho$ processes and their branching fractions are listed below.

Process	\mathcal{B}	Secondary decay	Secondary \mathcal{B}	Signal mode
$B^+ \rightarrow \bar{D}^0 \rho^+$	1.34%	$\bar{D}^0 \rightarrow K^+ \pi^-$	3.80%	$K^+ \pi^- \pi^+ \gamma$
		$\bar{D}^0 \rightarrow K^0 \pi^0$	2.30%	$K_S^0 \pi^+ \pi^0 \gamma$
$B^0 \rightarrow D^- \rho^+$	7.7×10^{-3}	$D^- \rightarrow K^0 \pi^-$	2.82%	$K_S^0 \pi^- \pi^+ \gamma$
$B^0 \rightarrow \bar{D}^0 \rho^0$	2.9×10^{-4}	$\bar{D}^0 \rightarrow K^0 \pi^0$	2.30%	$K_S^0 \pi^- \pi^+ \gamma$

TABLE 21. Branching fractions \mathcal{B} of $B \rightarrow D \rho$ modes. For each, the secondary decays and the signal modes to which they contribute backgrounds are listed.

We do not study the last of these, $B^0 \rightarrow \bar{D}^0 \rho^0$, because of its small branching fraction.

We find that the process $B^+ \rightarrow \bar{D}^0 \rho^+$ contributes 3.1 peaking background events in the fit region, after candidate selection and MCS, in the $K_S^0 \pi^+ \pi^0 \gamma$ mode. Distributions of m_{ES} , ΔE , and invariant masses are shown below. We find fewer than one background event in the other modes.

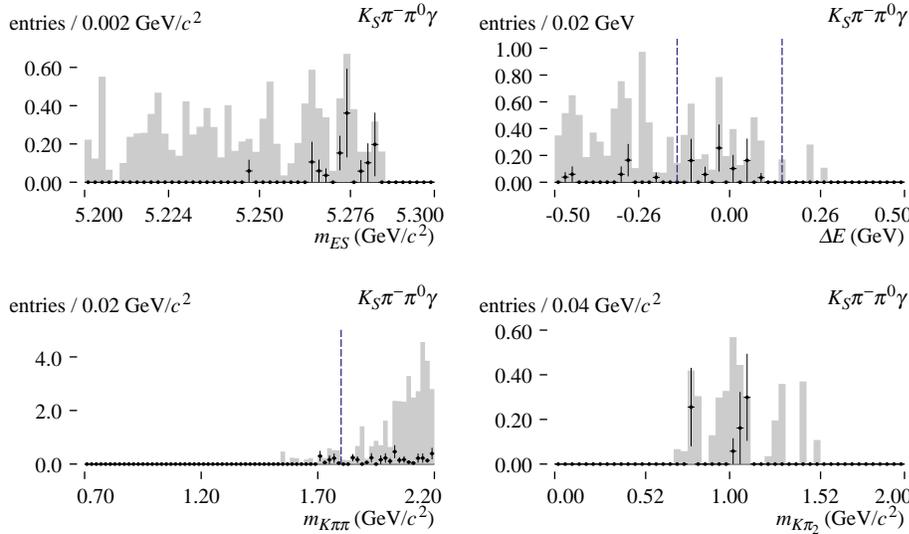


FIGURE 38. Distributions of m_{ES} , ΔE , and invariant masses in background candidates generated as $B^+ \rightarrow \bar{D}^0 \rho^+$, where $\bar{D}^0 \rightarrow K^+ \pi^0$, and reconstructed in mode $K_S^0 \pi^+ \pi^0 \gamma$. All candidates passing candidate selection cuts are in gray; those reconstructed from a single B are in black.

6.2.3 $B \rightarrow K^* \rho$

The rare decays $B \rightarrow K^* \rho$ have recently been measured by *BABAR* in three charge combinations. The branching fractions and longitudinal polarization fractions for these measurements are listed below. The decay $B^0 \rightarrow K^{*0} \rho^0$ has not been observed; the current upper limit on the branching fraction is 34×10^{-6} at 90% C.L. We use an ansatz branching fraction of 15×10^{-6} and assume 50% longitudinal polarization for this mode.

Process	\mathcal{B}	f_L	Secondary decay	Signal mode
$B^+ \rightarrow K^{*0} \rho^+$	17.0×10^{-6}	0.79	$K^{*0} \rightarrow K^+ \pi^-$ $K^{*0} \rightarrow K^0 \pi^0$	$K^+ \pi^- \pi^+ \gamma$ $K_S^0 \pi^+ \pi^0 \gamma$
$B^+ \rightarrow K^{*+} \rho^0$	10.6×10^{-6}	0.96	$K^{*+} \rightarrow K^+ \pi^0$	$K^+ \pi^- \pi^+ \gamma$
$B^0 \rightarrow K^{*+} \rho^-$	11.8×10^{-6}	0.27	$K^{*+} \rightarrow K^+ \pi^0$ $K^{*+} \rightarrow K^0 \pi^+$	$K^+ \pi^- \pi^0 \gamma$ $K_S^0 \pi^- \pi^+ \gamma$
$B^0 \rightarrow K^{*0} \rho^0$	(15×10^{-6})	(0.50)	$K^{*0} \rightarrow K^0 \pi^0$	$K_S^0 \pi^- \pi^+ \gamma$

TABLE 22. Branching fraction \mathcal{B} and longitudinal polarization fraction f_L of $B \rightarrow K^* \rho$ modes. For each, the secondary decays and the signal modes to which they contribute backgrounds are listed. The branching fraction and polarization fraction for $K^{*0} \rho^0$ are ansatz values.

In the $K^+\pi^-\pi^+\gamma$ signal mode, we find a background estimate of 1.7 events in the fit region, after candidate selection and MCS, from $B \rightarrow K^*\rho$. Kinematic distributions are shown below. We expect fewer than one background event from this process in the other three signal modes.

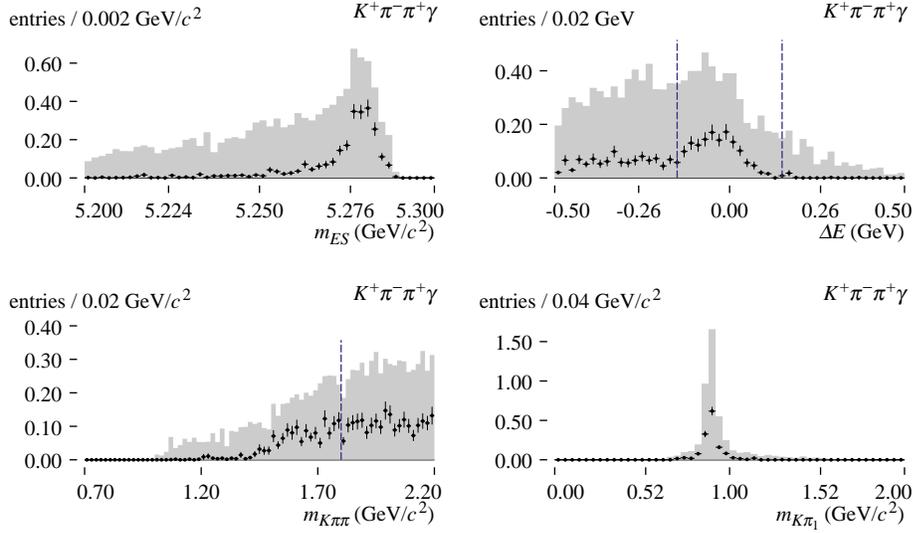


FIGURE 39. Distributions of m_{ES} , ΔE , and invariant masses in background candidates generated as $B \rightarrow K^*\rho$ and reconstructed in mode $K^+\pi^-\pi^+\gamma$. All candidates passing candidate selection cuts are in gray; those reconstructed from a single B are in black.

6.2.4 COMBINED YIELDS

Yields per signal mode for the exclusive B decay processes described above are summarized in the table below. The first table summarizes background yields for the samples described above. The second table breaks out the yields from candidates reconstructed from a single B .

Process	$K^+\pi^-\pi^+\gamma$	$K^+\pi^-\pi^0\gamma$	$K_s^0\pi^-\pi^+\gamma$	$K_s^0\pi^+\pi^0\gamma$
$B \rightarrow D\pi^0$	5.4 ± 0.6	5.6 ± 0.6	1.6 ± 0.2	0.9 ± 0.2
$B \rightarrow D\eta$	4.7 ± 0.3	3.2 ± 0.3	1.1 ± 0.1	0.7 ± 0.1
$B \rightarrow D\rho$	5.8 ± 1.2	5.4 ± 0.9	3.5 ± 0.7	3.2 ± 0.6
$B \rightarrow K^*\rho$	3.5 ± 0.1	1.0 ± 0.1	0.5 ± 0.0	0.5 ± 0.0
total	19.3 ± 1.4	15.2 ± 1.1	6.7 ± 0.8	5.2 ± 0.6

TABLE 23. Background yields from $B \rightarrow K\pi\pi\pi^0$ and $B \rightarrow K\pi\pi\eta$ scaled to data luminosity. Uncertainties are from MC statistics only.

Process	$K^+\pi^-\pi^+\gamma$	$K^+\pi^-\pi^0\gamma$	$K_s^0\pi^-\pi^+\gamma$	$K_s^0\pi^+\pi^0\gamma$
$B \rightarrow D\pi^0$	0.1 ± 0.1	0.4 ± 0.2	0.0 ± 0.0	0.0 ± 0.0
$B \rightarrow D\eta$	0.0 ± 0.0	0.1 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
$B \rightarrow D\rho$	0.2 ± 0.2	0.0 ± 0.0	0.0 ± 0.0	0.8 ± 0.3
$B \rightarrow K^*\rho$	1.2 ± 0.1	0.1 ± 0.0	0.1 ± 0.0	0.0 ± 0.0
total	1.4 ± 0.2	0.5 ± 0.2	0.2 ± 0.0	0.8 ± 0.3

TABLE 24. Background yields from $B \rightarrow K\pi\pi\pi^0$ and $B \rightarrow K\pi\pi\eta$ scaled to data luminosity. Only candidates decaying from a single B are included. Uncertainties are from MC statistics only.

6.3 Generic backgrounds

We consider two sources of combinatoric backgrounds: $udsc$ continuum processes and generic $B\bar{B}$ decays. To study combinatoric backgrounds, we remove $b \rightarrow s\gamma$ processes from the generic $B\bar{B}$ MC sample, based on MC truth information. These are all events generated as $B \rightarrow X_s\gamma$, where X_s is either an X_{su} or X_{sd} (a MC pseudoparticle specifying generic fragmentation including an s quark), or a kaon resonance. We refer to backgrounds from all $B\bar{B}$ processes other than $b \rightarrow s\gamma$ as generic $B\bar{B}$ backgrounds.

Yields from $udsc$ and from generic $B\bar{B}$ processes are summarized below. For the latter, the yields of candidates with all tracks coming from the same B are indicated.

Process	$K^+\pi^-\pi^+\gamma$	$K^+\pi^-\pi^0\gamma$	$K_s^0\pi^-\pi^+\gamma$	$K_s^0\pi^+\pi^0\gamma$
$udsc$	6971.8 ± 70.4	4381.5 ± 55.2	2054.9 ± 38.5	1693.9 ± 34.3
generic B	340.2 ± 8.9	235.7 ± 7.3	101.5 ± 4.8	87.6 ± 4.5
(from same B)	11.7 ± 1.7	11.0 ± 1.6	6.1 ± 1.3	1.6 ± 0.6

TABLE 25. Background yields from $udsc$ and $B\bar{B}$ (with $b \rightarrow s\gamma$ decays removed) MC samples, scaled to the analysis luminosity. Candidate selection, the fit region cut, and MCS have been applied.

The following two figures show the m_{ES} and ΔE distributions for these two MC samples.

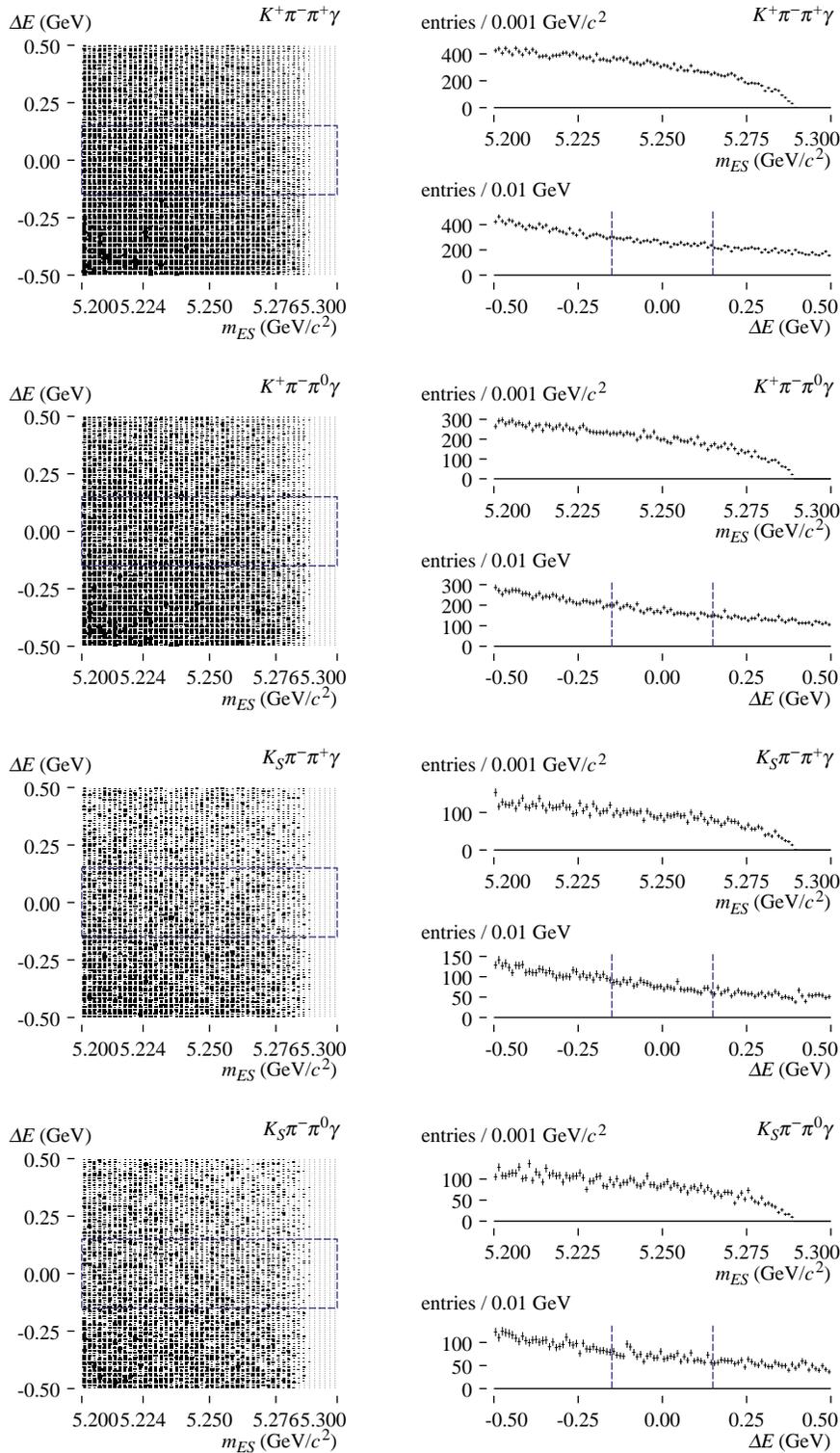


FIGURE 40. Joint and marginal distributions of m_{ES} and ΔE in background candidates from $udsc$ MC. Candidate selection excluding MCS have been applied. The fit region is shown in blue.

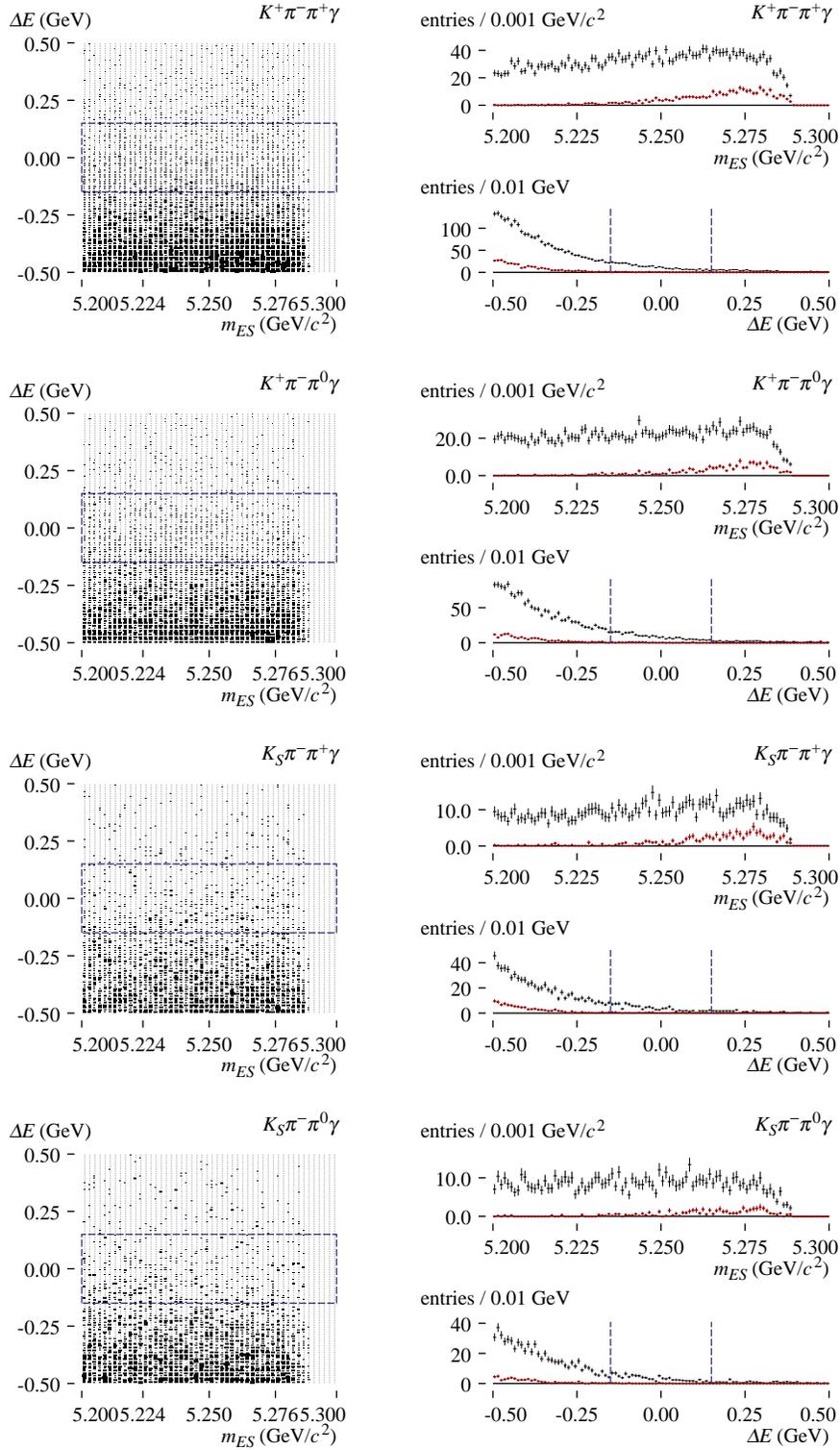


FIGURE 41. Joint and marginal distributions of m_{ES} and ΔE in background candidates from $B\bar{B}$ MC, with $b \rightarrow s\gamma$ decays removed. Candidate selection excluding MCS have been applied. Backgrounds from candidates reconstructed from a single B are shown in red. The fit region is shown in blue.

These plots show the distributions of $m_{K\pi\pi}$ in the $udsc$ and generic $B\bar{B}$ MC samples.

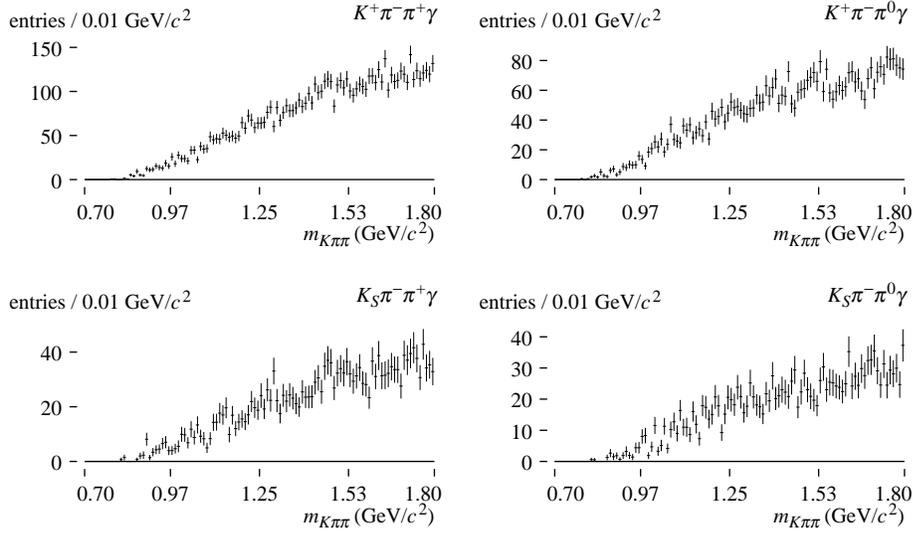


FIGURE 42. $m_{K\pi\pi}$ distributions of background candidates in $udsc$ MC. All candidate selection criteria have been applied.

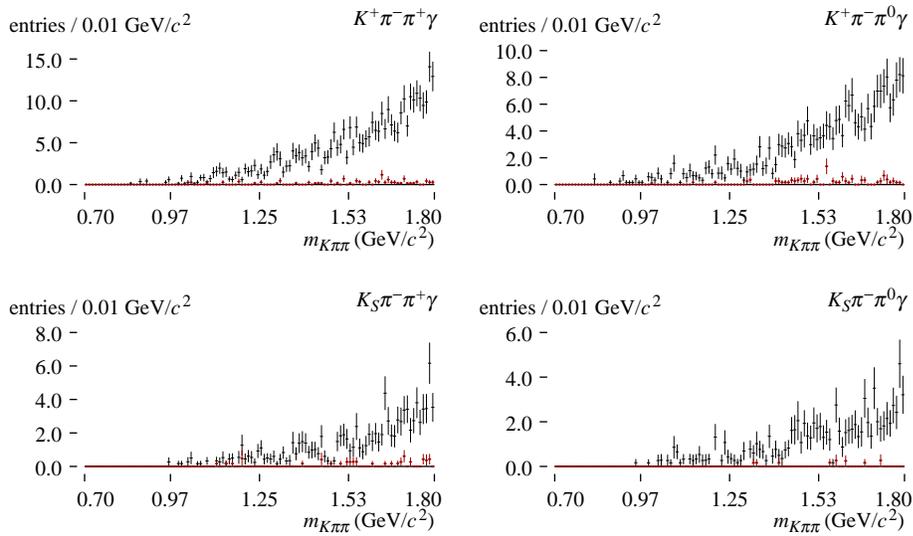


FIGURE 43. $m_{K\pi\pi}$ distributions of background candidates in $B\bar{B}$ MC, with $b \rightarrow s\gamma$ decays removed. All candidate selection criteria have been applied. Backgrounds from candidates reconstructed from a single B are shown in red.

We have examined MC truth information for background candidates reconstructed from a single B that remain in the generic $B\bar{B}$ MC sample after $b \rightarrow s\gamma$ and the exclusive B peaking modes have been removed. They are produced by a wide variety of B decay processes; the most prominent of these are,

- $B \rightarrow D\rho$, $D \rightarrow K\pi\pi$ and other $B \rightarrow K\pi\pi\pi\pi^0$ processes,
- $B \rightarrow D^*\rho$, and

- $B \rightarrow K^*\eta$ and $B \rightarrow K^*\pi^0$, where $\eta/\pi^0 \rightarrow \gamma\gamma$, one photon converts, and a conversion electron is mis-identified as a pion.

We find no indication that these peak near zero in ΔE , and do not study them further.

6.4 Total backgrounds

The table below summarizes the total expected signal and background yields. The figures following show the m_{ES} and ΔE distributions from the various MC samples we use to estimate B backgrounds.

Backgrounds from $udsc$ MC, which are much larger, are not shown.

	$K^+\pi^-\pi^+\gamma$	$K^+\pi^-\pi^0\gamma$	$K_S^0\pi^-\pi^+\gamma$	$K_S^0\pi^+\pi^0\gamma$
signal	714.6 ± 4.5	433.1 ± 3.7	220.8 ± 2.5	111.2 ± 1.9
$K\pi\gamma$	213.7 ± 2.3	83.0 ± 1.4	64.5 ± 1.2	39.2 ± 1.0
$K\pi\pi\gamma$	383.6 ± 3.4	282.7 ± 2.9	133.4 ± 2.0	121.6 ± 1.9
other $b \rightarrow s\gamma$	308.7 ± 11.2	219.0 ± 8.5	115.2 ± 6.8	107.9 ± 6.1
other B decays	315.4 ± 8.6	220.6 ± 7.0	96.3 ± 4.7	82.1 ± 4.4
$udsc$	6971.8 ± 70.4	4381.5 ± 55.2	2054.9 ± 38.5	1693.9 ± 34.3

TABLE 26. Expected yields in the data sample for signal and background after all selection cuts, the fit region cut, and MCS. Uncertainties are due to MC statistics.

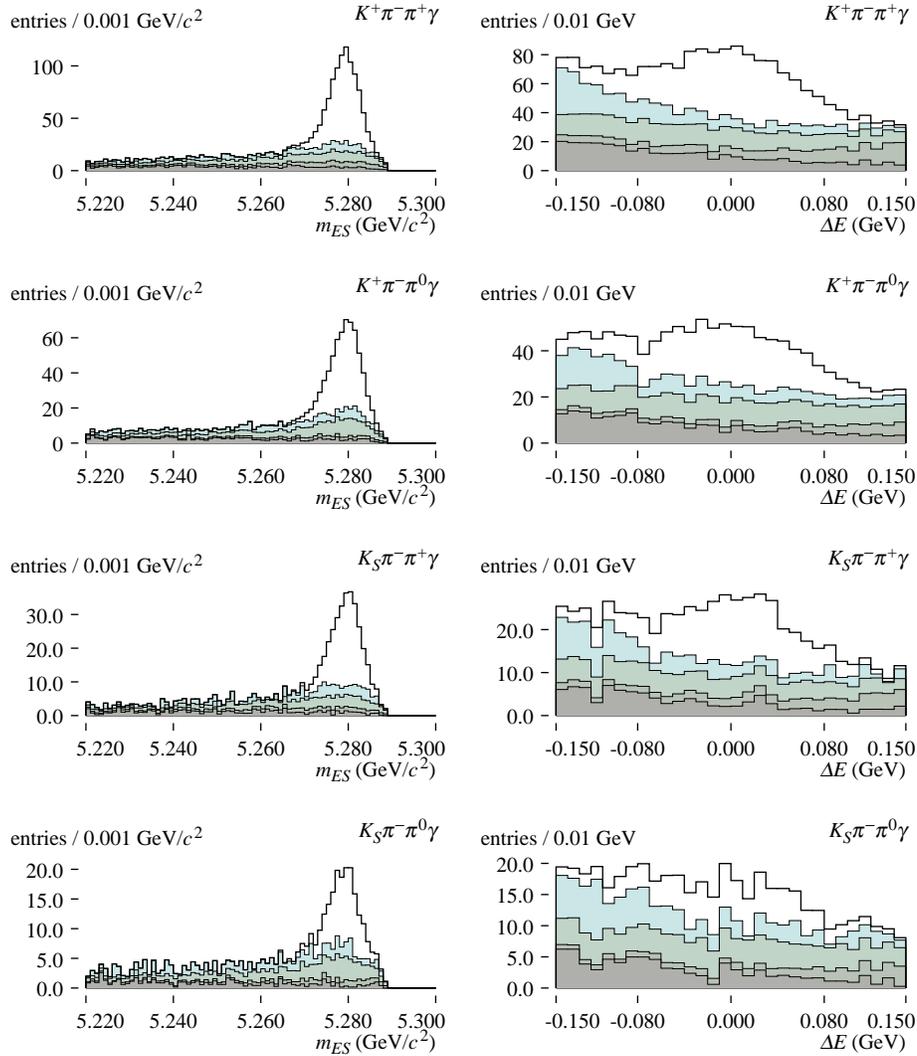


FIGURE 44. Stacked m_{ES} and ΔE distributions of background contributions in MC from B decays, after all selection criteria. Top to bottom:

- truth-matched signal MC, in white.
- $b \rightarrow s\gamma$ feed-down,
- $B \rightarrow K\pi\pi\gamma$ crossfeed,
- $B \rightarrow K\pi\gamma$ feed-up,
- generic B decays,

The background contribution from $udsc$ processes is not shown.

7 Control Samples

In this section we describe three control sample studies performed to assess the accuracy of our MC samples in predicting the shapes and yields of backgrounds in real data. First, we examine off-resonance data, and compare it to the $udsc$ MC cocktail. Second, we examine the ΔE sidebands below and above the fit region in on-resonance data, and compare with a mixture of off-resonance data and B MC samples. Third, we reconstruct $B \rightarrow D \pi^\pm$, $D \rightarrow K \pi \pi$ in on-resonance data and compare with corresponding MC samples, to study the efficiency of particle quality, event shape, and vertex cuts.

7.1 Off-resonance data

We process off-resonance data with the sample selection as our other samples. Candidate selection is identical as well, except we adjust m_{ES} by the difference between the nominal on-resonance beam energy and the actual off-resonance beam energy for each run.

We compare the off-resonance data with $udsc$ MC samples. The MC samples are scaled to the integrated luminosity of the off-resonance data. The yields are compared below.

Mode	$udsc$ MC	off-resonance data	ratio
$K^+ \pi^- \pi^+ \gamma$	718.0 ± 7.2	880.0 ± 29.7	$(81.6 \pm 2.9)\%$
$K^+ \pi^- \pi^0 \gamma$	451.2 ± 5.7	487.0 ± 22.1	$(92.7 \pm 4.4)\%$
$K_S^0 \pi^- \pi^+ \gamma$	211.6 ± 4.0	244.0 ± 15.6	$(86.7 \pm 5.8)\%$
$K_S^0 \pi^+ \pi^0 \gamma$	174.4 ± 3.5	188.0 ± 13.7	$(92.8 \pm 7.0)\%$

TABLE 27. Yields in scaled $udsc$ MC compared to off-resonance data. Candidate selection, the fit region cut, and MCS have been applied. Uncertainties are statistical only.

These plots compare the m_{ES} and ΔE distributions.

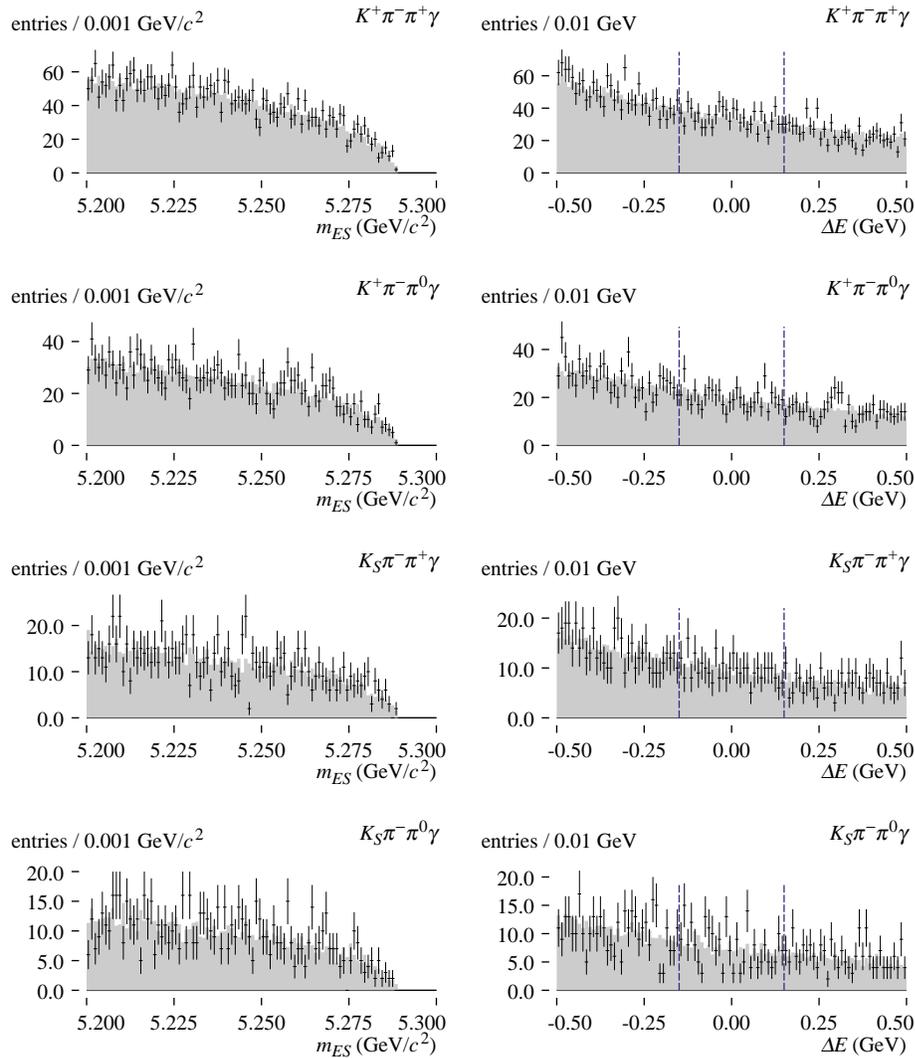


FIGURE 45. Distributions of m_{ES} and ΔE in off-resonance data (points) and $udsc$ MC (shaded). The $udsc$ MC distributions are normalized to the data distributions. Candidate selection cuts excluding the fit region cuts and MCS have been applied. The fit region is indicated in blue.

These plots compare the $m_{K\pi\pi}$ distributions.

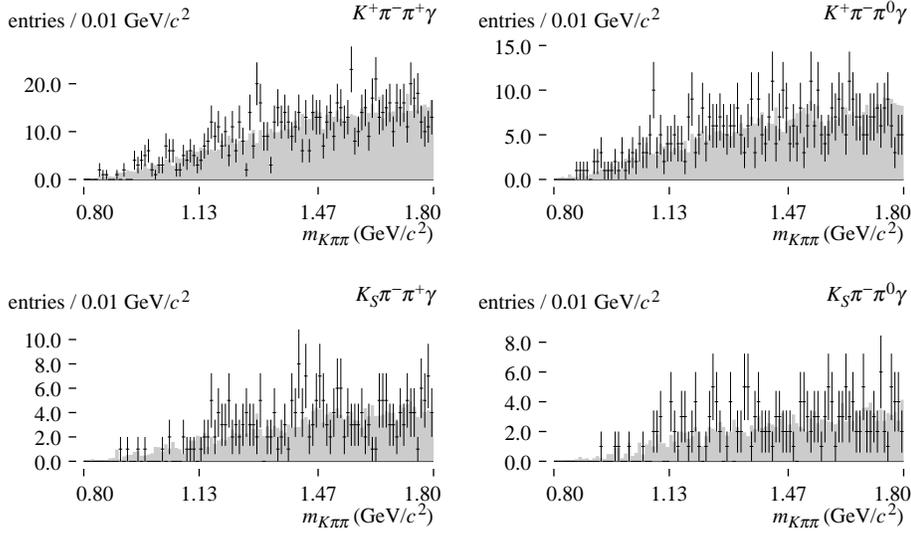


FIGURE 46. Distributions of $m_{K\pi\pi}$ in off-resonance data (points) and $udsc$ MC (shaded). The $udsc$ MC distributions are normalized to the data distributions. Candidate selection, the fit region cuts, and MCS have been applied.

We conclude that the $udsc$ MC is in fair agreement with the off-resonance data regarding yields of background events, and in excellent agreement regarding distributions. The discrepancy in the total background yield from the $udsc$ MC samples is probably due to incorrect modeling of the light-quark fragmentation. In our fit to on-resonance data, we do not rely on MC samples to determine the yield of the background component from $udsc$ processes; the normalization of this component is allowed to float. Further, we parameterize the m_{ES} and ΔE shapes of this component and allow the shape parameters to float in the fit. This control sample study gives us confidence that the shape parameterization, which is validated using $udsc$ MC, is reasonable.

7.2 ΔE sidebands

The ΔE sideband samples are the region in on-resonance data above and below the fit region in ΔE . The *low sideband* consists of candidates in the range $-0.50 < \Delta E < -0.15$ GeV, and the *high sideband* consists of candidates in the range $0.15 < \Delta E < 0.50$ GeV. We expect these regions to contain only a small number of signal events, but to contain sizable contributions from background processes. We studied this sideband before unblinding the fit region to assess our understanding of these backgrounds.

We compare the yields and distributions in these control samples to a cocktail composed of,

- MC estimates of the peaking B decay modes studied in section 6.2,
- MC estimates background components from $B \rightarrow K\pi\gamma$, $B \rightarrow K\pi\pi\gamma$, and other $b \rightarrow s\gamma$ processes,
- background estimates from the generic $B\bar{B}$ MC, with $b \rightarrow s\gamma$ processes removed, plus
- the off-resonance data in the same ΔE region, scaled by the ratio of the luminosities of the on-resonance to off-resonance data samples.

Yields in the ΔE sidebands in on-resonance data, along with yields in the corresponding comparison cocktail samples, are given in the tables below.

Mode	On-res yield	Cocktail yield	Ratio
$K^+ \pi^- \pi^+ \gamma$	21764	22112	$(101.6 \pm 2.0)\%$
$K^+ \pi^- \pi^0 \gamma$	12759	12597	$(98.7 \pm 2.5)\%$
$K_S^0 \pi^- \pi^+ \gamma$	6414	6307	$(98.3 \pm 3.5)\%$
$K_S^0 \pi^+ \pi^0 \gamma$	4917	4820	$(98.0 \pm 4.0)\%$

TABLE 28. Yields in the low ΔE sideband after candidate selection in on-resonance data and comparison cocktail sample.

Mode	On-res yield	Cocktail yield	Ratio
$K^+ \pi^- \pi^+ \gamma$	9454	9446	$(99.9 \pm 3.2)\%$
$K^+ \pi^- \pi^0 \gamma$	5667	5572	$(98.3 \pm 4.1)\%$
$K_S^0 \pi^- \pi^+ \gamma$	2750	2749	$(100.0 \pm 5.9)\%$
$K_S^0 \pi^+ \pi^0 \gamma$	2156	2390	$(110.9 \pm 7.0)\%$

TABLE 29. Yields in the high ΔE sideband after candidate selection in on-resonance data and comparison cocktail sample.

The following figure compares distributions in the low ΔE sideband sample for the $B \rightarrow K^+ \pi^- \pi^+ \gamma$ mode to the off-resonance data and simulation sample. Distributions in the other three modes and in the high ΔE sideband are similar. Low statistics of the off-resonance sample limits the precision of the comparison.

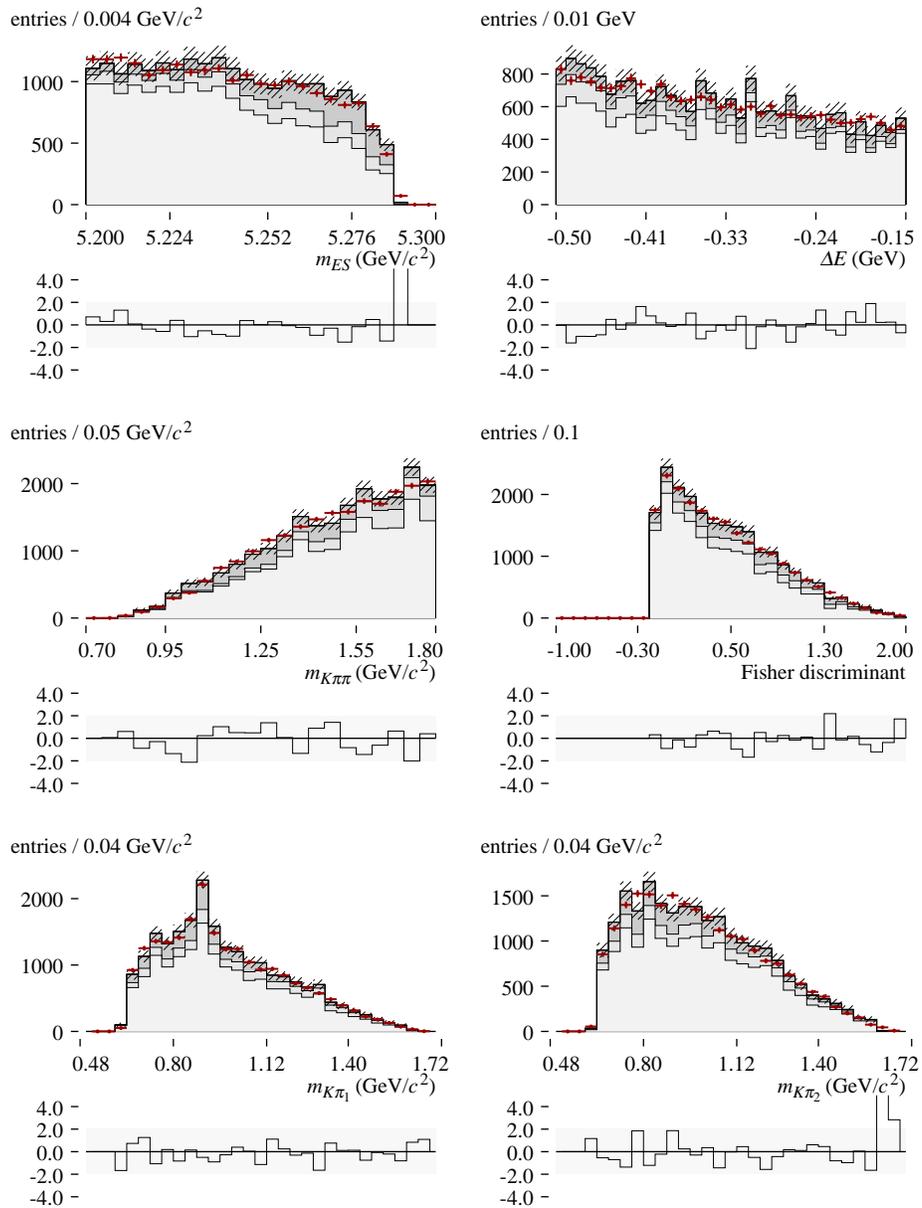


FIGURE 47. Yields in the low ΔE sideband for mode $K^+\pi^-\pi^+\gamma$ after candidate selection in on-resonance data (red points) and comparison cocktail sample (shaded). Cocktail components from bottom are off-resonance data, generic $B\bar{B}$ MC, $b \rightarrow s\gamma$ MC samples, and (tiny) peaking B decay modes. Crosshatching shows the statistical uncertainty of the total cocktail distribution. The plot below each distribution shows the difference normalized to the bin error, with $\pm 2\sigma$ shaded.

7.3 $D\pi^\pm$ control samples

We study the efficiencies of our candidate selection criteria in samples of $B \rightarrow D\pi^\pm$, $D \rightarrow K\pi\pi$ candidates reconstructed in on-resonance data. For each of our signal modes, there is a corresponding $D\pi^\pm$ control sample in which the high-energy photon's analogue is the prompt π^\pm from the B decay. Since the decay $D^+ \rightarrow K^+\pi^-\pi^+$ is doubly Cabbibo-suppressed, we instead use the Cabbibo-allowed decay $D^+ \rightarrow K^-\pi^+\pi^+$ as the analogue of our $K^+\pi^-\pi^+\gamma$ mode. The branching fractions of the four control sample processes are listed below.

Process		Secondary decay		Signal mode
$B^0 \rightarrow D^-\pi^+$	2.76×10^{-3}	$D^- \rightarrow K^+\pi^-\pi^-$	9.2%	$K^+\pi^-\pi^+\gamma$
		$D^- \rightarrow K^0\pi^-\pi^0$	9.7%	$K_S^0\pi^+\pi^0\gamma$
$B^+ \rightarrow \bar{D}^0\pi^+$	4.98×10^{-3}	$\bar{D}^0 \rightarrow K^+\pi^-\pi^0$	13.0%	$K^+\pi^-\pi^0\gamma$
		$\bar{D}^0 \rightarrow K^0\pi^+\pi^-$	6.0%	$K_S^0\pi^-\pi^+\gamma$

TABLE 30. Branching fractions of $B \rightarrow D\pi^\pm$, $D \rightarrow K\pi\pi$ processes reconstructed in the $D\pi^\pm$ control samples.

We measure the momentum of the reconstructed B in these control samples more precisely than in our signal modes, because the EMC cluster from the high-energy photon is replaced with a better-measured charged track from the prompt π^\pm . For correctly reconstructed candidates, the value of $m_{K\pi\pi}$ is equal to the D mass, up to detector resolution. We therefore expect to be able to isolate very pure samples of these decays with tight cuts on kinematic variables. It is important to note, however, that the kinematic distribution in these control samples is not quite the same as the distributions in our signal candidates, since the D mass lies just above the upper limit of $m_{K\pi\pi}$ we reconstruct for signal candidates, $1.8 \text{ GeV}/c^2$.

We reconstruct $D\pi^\pm$ candidates in on-resonance data. We use the same sample selection procedure that we use to reconstruct signal candidates, except that we select charged tracks satisfying the *GoodTracksLoose* track selection instead of high-energy photons. We use the same loose sample selection cuts on m_{ES} and ΔE that we use when reconstructing signal candidates, and also require

that $m_{K\pi\pi}$ fall within $50 \text{ MeV}/c^2$ of the D mass.

We perform a geometric fit to the reconstructed D to produce a vertex probability, on which we place a candidate selection cut. A fit to the reconstructed B reconstructed as $B \rightarrow D\pi^\pm$ would not be analogous to the vertex fit in $B \rightarrow K\pi\pi\gamma$ signal candidates, since the track of the prompt π^\pm introduces additional geometric information not present in signal candidates. Additionally, we expect the D decay vertex in the control sample to be displaced from the B decay vertex, which is not the case in events from signal processes. Therefore, to model the vertex probability in the control samples, we vertex only the kaon and two pions that comprise the D .

We select $D\pi^\pm$ candidates with these cuts.

- hadronic selection cuts on $K^\pm, \pi^\pm, K_s^0, \pi^0$ as for signal modes,
- Fisher discriminant cuts as for signal modes,
- B vertex probability cut applied to D vertex probability,
- $|\Delta E| < 25 \text{ MeV}$, and
- $|m_{K\pi\pi} - m_D| < 10 \text{ MeV}/c^2$ in the $\pi^+\pi^-$ modes
 $|m_{K\pi\pi} - m_D| < 20 \text{ MeV}/c^2$ in the $\pi^+\pi^0$ modes.

We do not cut on m_{ES} ; instead, we fit the m_{ES} distribution to determine the yield of $D\pi^\pm$ candidates. However, when plotting distributions, we impose a cut $5.276 < m_{ES} < 5.284 \text{ GeV}/c^2$. The distributions of m_{ES} , ΔE , and $m_{K\pi\pi}$ in the $D\pi^\pm$ control samples are shown below.

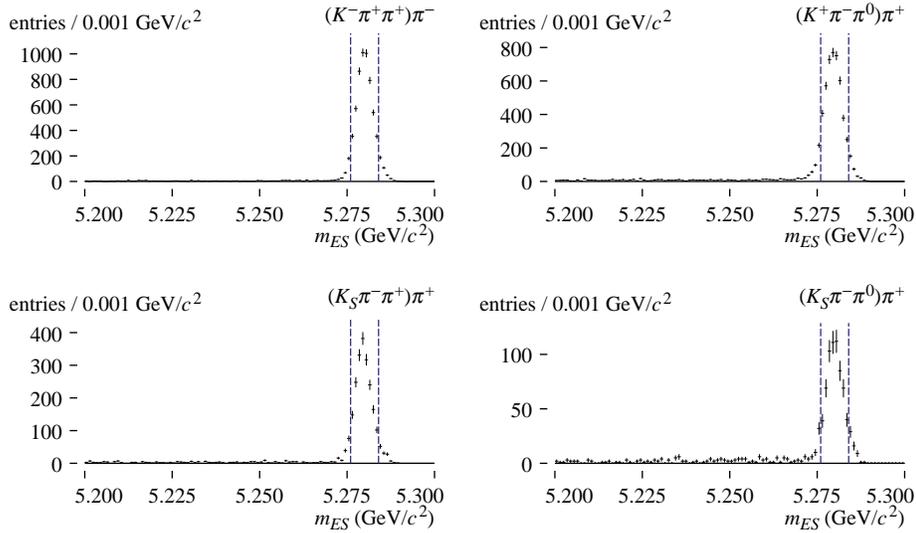


FIGURE 48. Distribution of m_{ES} in $D\pi^\pm$ control samples from on-resonance data. Hadronic quality, Fisher, ΔE , and $m_{K\pi\pi}$ cuts have been applied. The m_{ES} cuts are indicated in blue.

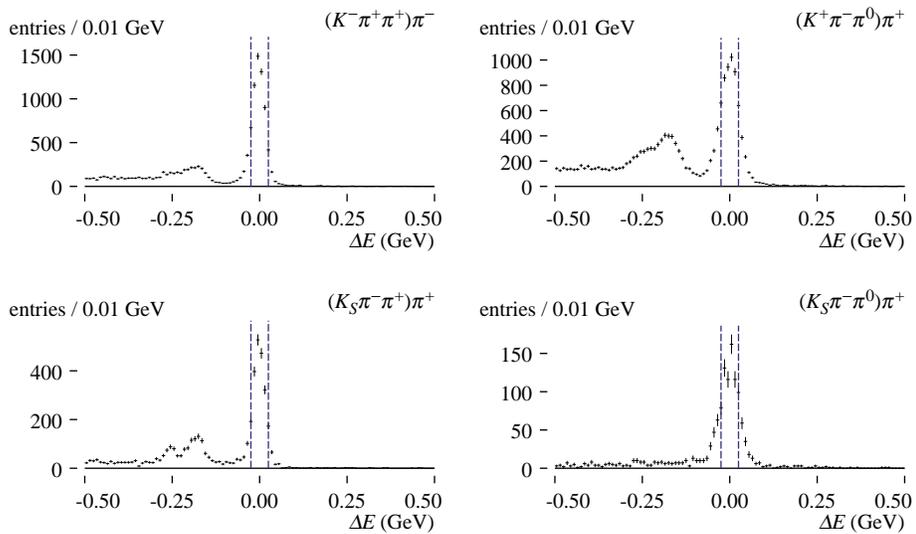


FIGURE 49. Distribution of ΔE in $D\pi^\pm$ control samples from on-resonance data. Hadronic quality, Fisher, m_{ES} , and $m_{K\pi\pi}$ cuts have been applied. The ΔE cuts are indicated in blue.

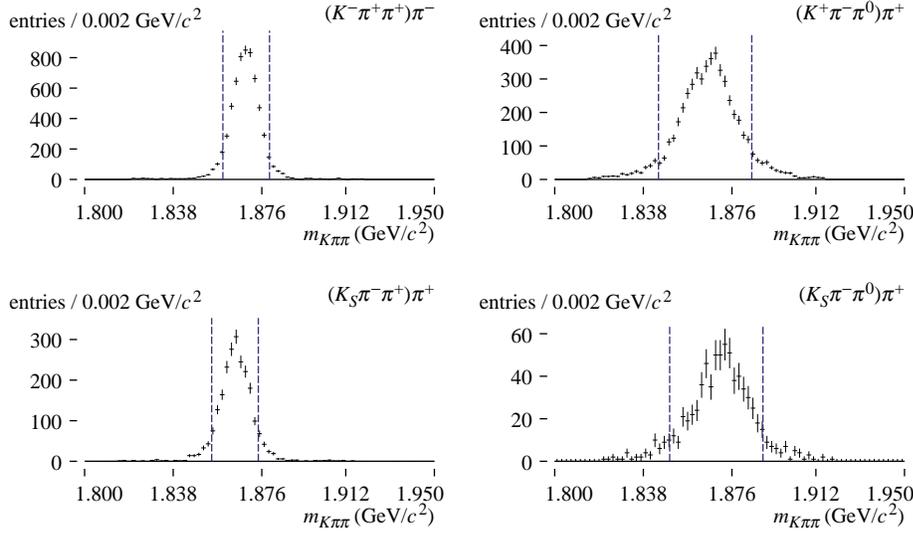


FIGURE 50. Distribution of $m_{K\pi\pi}$ in $D\pi^\pm$ control samples from on-resonance data. Hadronic quality, Fisher, m_{ES} , and ΔE cuts have been applied. The $m_{K\pi\pi}$ cuts are indicated in blue.

We study the efficiency of the hadronic selection cuts, Fisher discriminant cut, and B vertex cut by imposing all of them on the $D\pi^\pm$ control sample and then relaxing one cut at a time. We determine the number of signal candidates in the sample with all cuts imposed, and for each cut divide this by the number of candidates with the cut relaxed. We compare these efficiencies computed in the $D\pi^\pm$ on-resonance control samples to corresponding $D\pi^\pm$ MC samples, and to signal MC samples.

Because these cuts still admit some amount of background in the control sample, we determine the number of signal $D\pi^\pm$ candidates in each sample (with all cuts imposed, or one cut relaxed) from an unbinned maximum likelihood fit to the m_{ES} distribution, after applying the ΔE and $m_{K\pi\pi}$ cuts. We use a Crystal Ball shape for the signal component, the same shape we use for the m_{ES} distribution of our signal candidates. We use an Argus function for the background component. Parameterizations of signal and background shapes are described in section 8.1.

In the $D\pi^\pm$ MC, we simply count truth-matched candidates passing the cuts when we compute efficiencies. In $K\pi\pi\gamma$ signal MC, we count truth-matched candidates passing the cuts; we do not apply the m_{ES} , ΔE , or $m_{K\pi\pi}$ cuts.

The plots on the following pages show distributions of the variables on which we impose cuts,

in the $D\pi^\pm$ on-resonance control sample, $D\pi^\pm$ MC, and signal MC. Distributions of the three variables comprising the Fisher discriminant are also shown. Tight cuts on m_{ES} have been imposed in these distributions, but the control samples in on-resonance data contain background contributions. From the m_{ES} fits, we estimate that the background component is approximately 95% in the worst case: when the Fisher cut is relaxed in the $K_s^0\pi^-\pi^0\pi^+$ control sample.

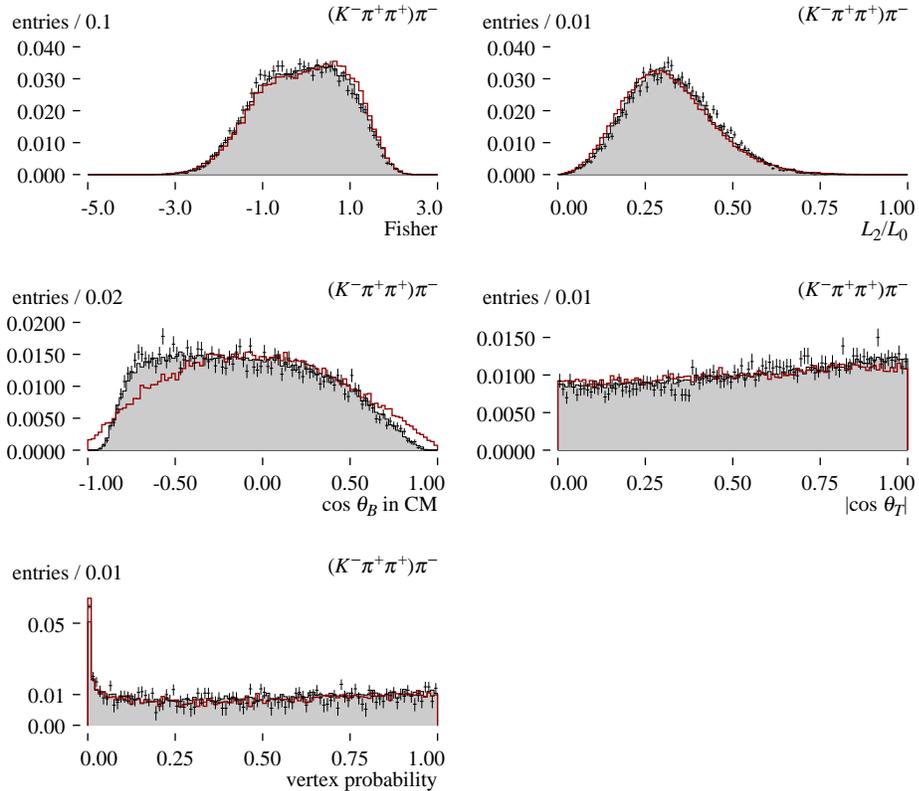


FIGURE 51. Distributions in the $K^-\pi^+\pi^-\pi^-$ on-resonance control sample (points) and corresponding MC (shaded). The cuts described in the text have been applied, except for the cut on the variable shown in the plot (or the Fisher cut, for plots of Fisher input variables). Distributions for truth-matched $K\pi\pi\gamma$ signal MC are shown in red for comparison. All three distributions are independently normalized to unit area.

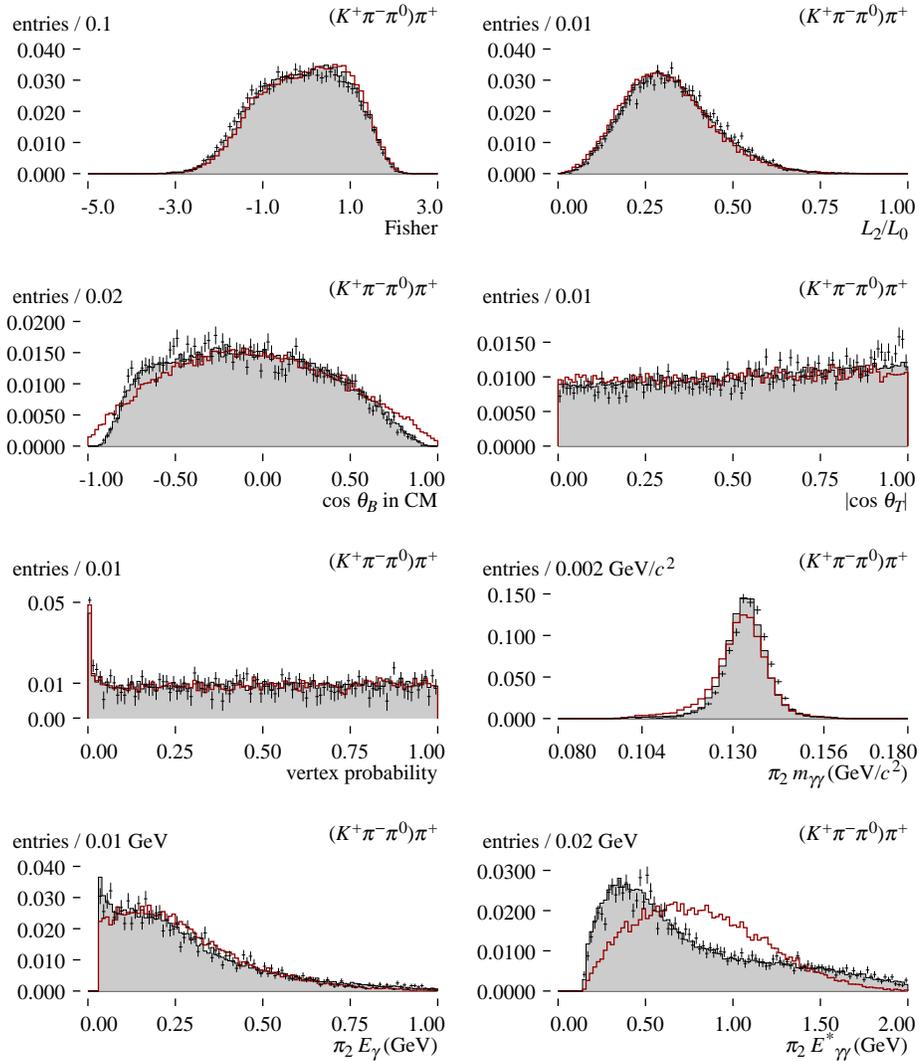


FIGURE 52. Distributions in the $K^+\pi^-\pi^0\pi^+$ on-resonance control sample (points) and corresponding MC (shaded). Quality cuts and m_{ES} , ΔE , and $m_{K\pi\pi}$ cuts have been applied, except for the cut on the variable shown in the plot (or the Fisher cut, for plots of Fisher input variables). Distributions for truth-matched $K\pi\pi\gamma$ signal MC are shown in red for comparison. All three distributions are normalized to unit area.

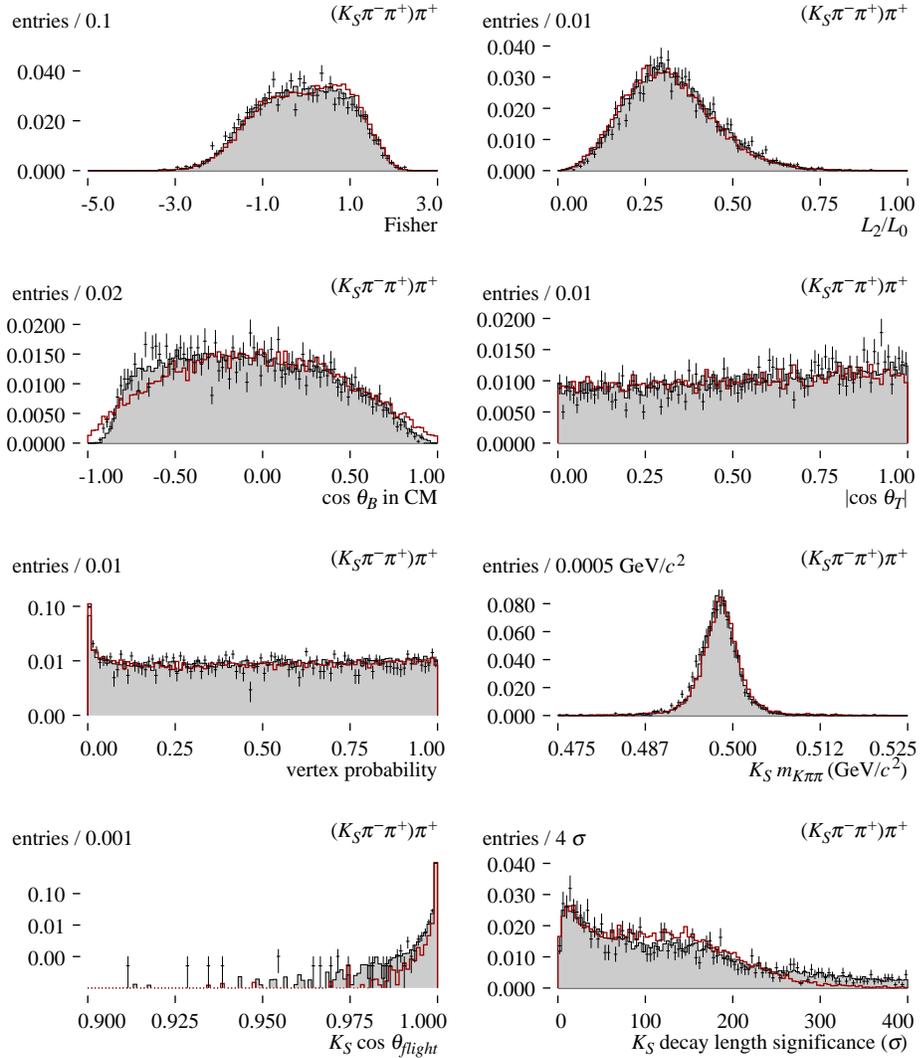


FIGURE 53. Distributions in the $K_S^0 \pi^- \pi^+ \pi^+$ on-resonance control sample (points) and corresponding MC (shaded). Quality cuts and m_{ES} , ΔE , and $m_{K\pi\pi}$ cuts have been applied, except for the cut on the variable shown in the plot (or the Fisher cut, for plots of Fisher input variables). Distributions for truth-matched $K\pi\pi\gamma$ signal MC are shown in red for comparison. All three distributions are normalized to unit area.

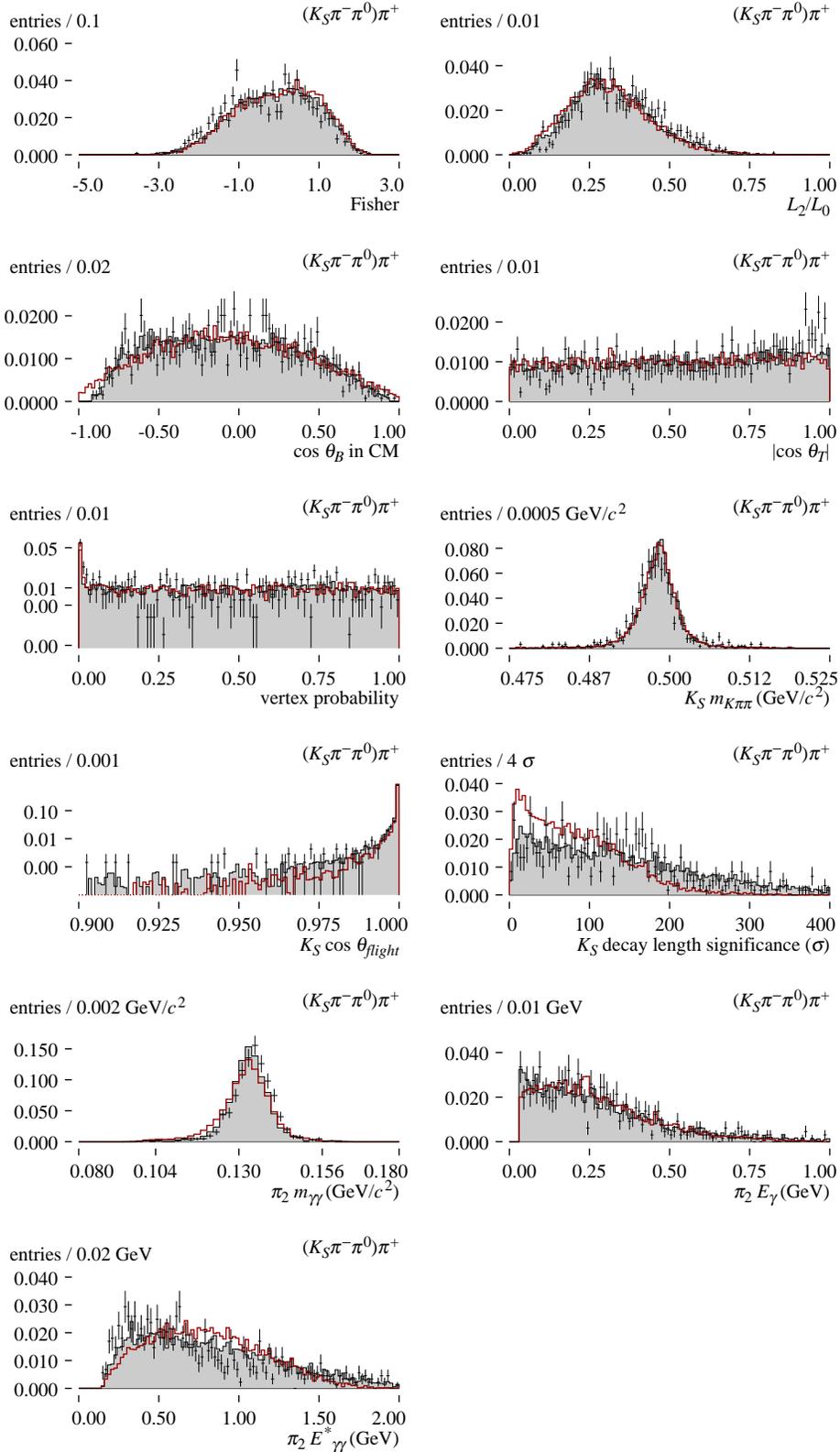


FIGURE 54. Distributions in the $K_S^0 \pi^- \pi^0 \pi^+$ on-resonance control sample (points) and corresponding MC (shaded). Quality cuts and m_{ES} , ΔE , and $m_{K\pi\pi}$ cuts have been applied, except for the cut on the variable shown in the plot (or the Fisher cut, for plots of Fisher input variables). Distributions for truth-matched $K\pi\pi\gamma$ signal MC are shown in red for comparison. All three distributions are normalized to unit area.

These plots show that the kinematic distributions of the $B \rightarrow D\pi^\pm$, $D \rightarrow K\pi\pi$ decays in our $D\pi^\pm$ control sample do not match those of our signal $K\pi\pi\gamma$ processes in all cases. Discrepancies are particularly wide in the distributions of $\cos\theta_B$ and of the π^0 CM energy. This is not surprising, since the underlying physics process is different.

The following table lists last-cut efficiencies in the $D\pi^\pm$ control sample and corresponding MC samples. The three K_s^0 quality cuts are treated individually—the efficiency for each is computed with the other two cuts imposed; likewise for the three π^0 quality cuts. Efficiencies are also presented for K_s^0 cuts and for the π^0 cuts taken pairwise.

Cut(s)	$D\pi^\pm$ data	$D\pi^\pm$ MC	Signal MC
	$(K^-\pi^+\pi^+)\pi^-$		$K^+\pi^-\pi^+\gamma$
Fisher	$(52.8 \pm 0.5)\%$	$(53.9 \pm 0.2)\%$	$(56.4 \pm 0.1)\%$
K^\pm PID	$(86.8 \pm 0.4)\%$	$(82.6 \pm 0.1)\%$	$(83.4 \pm 0.1)\%$
π^\pm PID for π_1	$(98.7 \pm 0.1)\%$	$(97.5 \pm 0.1)\%$	$(97.9 \pm 0.1)\%$
π^\pm PID for π_2	$(98.2 \pm 0.2)\%$	$(97.5 \pm 0.1)\%$	$(97.7 \pm 0.1)\%$
B vertex probability	$(96.2 \pm 0.2)\%$	$(97.4 \pm 0.1)\%$	$(94.0 \pm 0.1)\%$
	$(K^+\pi^-\pi^0)\pi^+$		$K^+\pi^-\pi^0\gamma$
Fisher	$(53.7 \pm 0.5)\%$	$(54.3 \pm 0.2)\%$	$(56.6 \pm 0.2)\%$
K^\pm PID	$(86.0 \pm 0.5)\%$	$(81.9 \pm 0.2)\%$	$(83.6 \pm 0.2)\%$
$\pi^0 E_{\gamma\gamma}^*$	$(68.9 \pm 0.5)\%$	$(65.4 \pm 0.2)\%$	$(83.3 \pm 0.2)\%$
$\pi^0 E_\gamma$	$(94.4 \pm 0.3)\%$	$(93.2 \pm 0.1)\%$	$(95.5 \pm 0.1)\%$
π^0 mass	$(96.8 \pm 0.2)\%$	$(95.3 \pm 0.1)\%$	$(93.0 \pm 0.1)\%$
π^\pm PID for π_1	$(99.1 \pm 0.1)\%$	$(97.7 \pm 0.1)\%$	$(98.1 \pm 0.1)\%$
B vertex probability	$(97.8 \pm 0.2)\%$	$(97.9 \pm 0.1)\%$	$(96.9 \pm 0.1)\%$
$\pi^0 E_\gamma$ and $\pi^0 E_{\gamma\gamma}^*$	$(61.6 \pm 0.5)\%$	$(57.1 \pm 0.2)\%$	$(78.1 \pm 0.2)\%$
π^0 mass and $\pi^0 E_{\gamma\gamma}^*$	$(63.9 \pm 0.5)\%$	$(59.2 \pm 0.2)\%$	$(76.6 \pm 0.2)\%$
π^0 mass and $\pi^0 E_\gamma$	$(90.4 \pm 0.4)\%$	$(86.9 \pm 0.2)\%$	$(88.5 \pm 0.2)\%$
	$(K_S\pi^-\pi^+)\pi^+$		$K_S\pi^-\pi^+\gamma$
Fisher	$(55.5 \pm 0.8)\%$	$(55.2 \pm 0.3)\%$	$(57.3 \pm 0.3)\%$
$K_S^0 \cos\theta_{\text{flight}}$	$(97.7 \pm 0.3)\%$	$(97.4 \pm 0.1)\%$	$(98.8 \pm 0.1)\%$
K_S^0 decay length	$(98.3 \pm 0.3)\%$	$(98.3 \pm 0.1)\%$	$(97.6 \pm 0.1)\%$
K_S^0 mass	$(98.8 \pm 0.2)\%$	$(99.1 \pm 0.1)\%$	$(97.5 \pm 0.1)\%$
π^\pm PID for π_1	$(99.4 \pm 0.2)\%$	$(97.5 \pm 0.1)\%$	$(97.6 \pm 0.1)\%$
π^\pm PID for π_2	$(98.4 \pm 0.3)\%$	$(97.9 \pm 0.1)\%$	$(97.9 \pm 0.1)\%$
B vertex probability	$(94.8 \pm 0.5)\%$	$(96.3 \pm 0.1)\%$	$(93.0 \pm 0.2)\%$
$K_S^0 \cos\theta_{\text{flight}}$ and K_S^0 decay length	$(89.3 \pm 0.6)\%$	$(93.2 \pm 0.2)\%$	$(94.4 \pm 0.2)\%$
K_S^0 mass and $K_S^0 \cos\theta_{\text{flight}}$	$(96.2 \pm 0.4)\%$	$(96.5 \pm 0.1)\%$	$(96.3 \pm 0.1)\%$
K_S^0 mass and K_S^0 decay length	$(96.8 \pm 0.4)\%$	$(97.3 \pm 0.1)\%$	$(95.0 \pm 0.1)\%$
	$(K_S\pi^-\pi^0)\pi^+$		$K_S\pi^-\pi^0\gamma$
Fisher	$(51.4 \pm 1.4)\%$	$(53.5 \pm 0.5)\%$	$(56.3 \pm 0.4)\%$
$K_S^0 \cos\theta_{\text{flight}}$	$(88.1 \pm 1.1)\%$	$(87.4 \pm 0.4)\%$	$(92.2 \pm 0.3)\%$
K_S^0 decay length	$(98.6 \pm 0.4)\%$	$(99.2 \pm 0.1)\%$	$(97.9 \pm 0.1)\%$
K_S^0 mass	$(96.8 \pm 0.7)\%$	$(98.1 \pm 0.2)\%$	$(97.3 \pm 0.2)\%$
$\pi^0 E_{\gamma\gamma}^*$	$(71.5 \pm 1.4)\%$	$(76.4 \pm 0.5)\%$	$(82.4 \pm 0.3)\%$
$\pi^0 E_\gamma$	$(94.9 \pm 0.8)\%$	$(93.5 \pm 0.3)\%$	$(95.7 \pm 0.2)\%$
π^0 mass	$(97.3 \pm 0.6)\%$	$(95.3 \pm 0.3)\%$	$(92.9 \pm 0.2)\%$
π^\pm PID for π_1	$(97.8 \pm 0.6)\%$	$(97.5 \pm 0.2)\%$	$(98.0 \pm 0.1)\%$
B vertex probability	$(97.7 \pm 0.6)\%$	$(97.8 \pm 0.2)\%$	$(96.7 \pm 0.2)\%$
$K_S^0 \cos\theta_{\text{flight}}$ and K_S^0 decay length	$(75.7 \pm 1.4)\%$	$(83.1 \pm 0.5)\%$	$(86.1 \pm 0.3)\%$
K_S^0 mass and $K_S^0 \cos\theta_{\text{flight}}$	$(85.3 \pm 1.2)\%$	$(85.5 \pm 0.5)\%$	$(89.5 \pm 0.3)\%$
K_S^0 mass and K_S^0 decay length	$(95.1 \pm 0.8)\%$	$(97.3 \pm 0.2)\%$	$(95.1 \pm 0.2)\%$
$\pi^0 E_\gamma$ and $\pi^0 E_{\gamma\gamma}^*$	$(65.2 \pm 1.5)\%$	$(68.6 \pm 0.5)\%$	$(77.5 \pm 0.4)\%$
π^0 mass and $\pi^0 E_{\gamma\gamma}^*$	$(67.0 \pm 1.5)\%$	$(70.5 \pm 0.5)\%$	$(75.8 \pm 0.4)\%$
π^0 mass and $\pi^0 E_\gamma$	$(91.4 \pm 1.0)\%$	$(87.6 \pm 0.4)\%$	$(88.5 \pm 0.3)\%$

TABLE 31. Efficiencies for each cut with all other cuts applied in the $D\pi^\pm$ on-resonance control samples (left column). Efficiencies in $D\pi^\pm$ MC (middle) and in truth-matched $K\pi\pi\gamma$ signal MC (right) are shown for comparison. Among the three K_S^0 cuts and among the three π^0 cuts, efficiencies are shown for relaxing the cuts pairwise as well. Uncertainties are computed assuming a binomial process with known number of trials.

In the preceding table, we have computed the uncertainty for each measured efficiency assuming a binomial random process with a known number of trials. However, in the $D\pi^\pm$ on-resonance control sample, we have determined both the number of trials (the number of candidates with the cut relaxed) and the number of successes (the number of candidates with the cut imposed) from fits to the m_{ES} distributions. We expect the fit uncertainties in these two to be strongly correlated, and therefore substantially to cancel in the efficiency uncertainty. Based on toy MC studies we have performed with similar fits, we estimate that the uncertainties on the efficiencies measured in the $D\pi^\pm$ data sample are several tenths of a percent larger than stated in the table.

The table also lists efficiencies for relaxing the K_s^0 cuts and π^0 cuts pairwise. In a few cases, this results in a significantly lower efficiency in the $D\pi^\pm$ control sample than in $D\pi^\pm$ MC. We expect this is due to non- $D\pi^\pm$ backgrounds in the control sample. The effect is particularly pronounced when the K_s^0 decay angle and decay length cuts are both relaxed; the remaining K_s^0 mass cut is quite loose, and admits significant combinatoric background.

As illustrated in the preceding plots, kinematic differences between $D\pi^\pm$ decays and signal $K\pi\pi\gamma$ decays prevent direct comparison between efficiencies in the $D\pi^\pm$ control sample and in signal MC. We therefore compare the control sample with $D\pi^\pm$ MC samples, not with signal MC samples, when we estimate systematic uncertainties due to selection efficiency.

We also expect the $D\pi^\pm$ MC simulation to exhibit differences from the data in modes other than $(K^+\pi^-\pi^-)\pi^+$, due to incomplete knowledge of the resonance substructure of these decays. In particular, differences in the π^0 energy distributions, and thus in the efficiencies of the cuts on these variables, are not surprising given the incomplete knowledge incorporated in the event generator's D decay model.

Systematic uncertainties for the K_s^0 and π^0 selection cuts are provided by standard recipes. We

conclude that the discrepancies between $D\pi^\pm$ data and MC shown in Table 31 are acceptable within these uncertainties. For the Fisher discriminant, vertex probability, and charged PID cuts, we assign uncertainties based on the discrepancies measured in this control sample. Systematic uncertainties are described in Chapter 10.

The $D\pi^\pm$ control sample also provides a measurement of the shift in measured m_{ES} due to uncertainty in the beam energy. The table below shows the values of the Crystal Ball mean parameter μ fitted to the $D\pi^\pm$ samples with all cuts imposed. The difference between this value and the B mass for each mode is the inferred shift in m_{ES} .

Mode	μ	$\mu - m_B$
$(K^-\pi^+\pi^+)\pi^-$	5.2801	0.00065
$(K^+\pi^-\pi^0)\pi^+$	5.2796	0.00054
$(K_S\pi^-\pi^+)\pi^+$	5.2797	0.00062
$(K_S\pi^-\pi^0)\pi^+$	5.2800	0.00062

TABLE 32. Values of mean parameter μ of Crystal Ball signal components fitted to m_{ES} distributions of $D\pi^\pm$ control samples after all cuts, and differences between parameter value and B mass.

8 Fits

To extract the signal yield, we perform a bivariate unbinned maximum likelihood fit in m_{ES} and ΔE to determine the branching fractions of our four $B \rightarrow K\pi\pi\gamma$ modes. We fit all four modes simultaneously, since the four signal processes each contribute to crossfeed background for all four final states. We use the *Minuit* minimization program to perform the fit.

The likelihood function we use contains terms for the following components:

1. Correctly-constructed signal candidates.
2. Crossfeed, which consists of misreconstructed signal events.
3. Feed-up backgrounds from $B \rightarrow K\pi\gamma$ events.
4. Feed-down backgrounds from higher-multiplicity $b \rightarrow s\gamma$ events.
5. Combined backgrounds from continuum $udsc$ events and $B\bar{B}$ events other than $b \rightarrow s\gamma$.

Each term is a product of a candidate yield for that component and a two-dimensional probability density function (PDF) describing the joint $m_{ES} - \Delta E$ distribution of these candidates. In each mode, one piece of the crossfeed component consists of candidates misreconstructed from events that are produced in the same mode; these are *self-crossfeed* candidates. For example, a $B^+ \rightarrow K^+\pi^-\pi^+\gamma$ event may be misreconstructed if the wrong π^+ , produced in the decay of the B^- , is chosen. The yields both of correctly reconstructed $B^+ \rightarrow K^+\pi^-\pi^+\gamma$ signal candidates and of such misreconstructed candidates are proportional to the same branching fraction $\mathcal{B}(B^+ \rightarrow K^+\pi^-\pi^+\gamma)$. We write them as separate terms in the fit, but conceptually they are a single contribution. By writing them as separate fit components, we can use different efficiencies and PDFs for the two. This allows us to separate the $m_{K\pi\pi}$ distribution for correctly reconstructed signal events, as we

will discuss below; the $m_{K\pi\pi}$ distribution of self-crossfeed candidates does not accurately reflect the mass spectrum of the underlying physics process. However, the branching fraction effectively is determined in the fit by both the signal and crossfeed yields.

The correctness of this aspect of the fit procedure depends on our ability to model the crossfeed components correctly (as, in fact, it depends on our ability to model the signal as well). However, the other alternative would be to fix the crossfeed background component in the fit and subtract it from the $m_{K\pi\pi}$ spectrum, which also depends on correct modeling.

Section 8.1 describes the fit functions we use for the components of our fit: the signal and crossfeed components, the continuum $udsc$ background component, and the $B \rightarrow K\pi\pi\gamma$ and other $b \rightarrow s\gamma$ components. In the section 8.2, we describe the formulation of the combined fit to the data. In section 8.3, we describe the method for extracting the $m_{K\pi\pi}$ spectrum.

8.1 Fit functions

In this section, we describe the fit functions we use for the components of our fit. We use analytic parameterizations for the signal and crossfeed components, which we fit to MC samples to determine shape parameters. We also use analytic parameterizations for the shape of backgrounds from continuum and generic $B\bar{B}$ decays. We rely on binned distributions measured from MC for the distributions of feed-up backgrounds from $B \rightarrow K\pi\gamma$ and for feed-down backgrounds from higher-multiplicity $b \rightarrow s\gamma$ decays.

We use maximum likelihood fits on MC samples to determine shape parameters. The unbinned fit procedure cannot accommodate samples of candidates with different weights. When we perform a fit to determine shape parameters, we unweight the sample MC using an acceptance-rejection method, where the weight for each candidate is computed to produce the correct efficiency for the

MC sample and the correct luminosity relative to other MC samples (so that MC samples generated with each of Run 1 through Run 4 conditions are represented in correct proportion). Since we only extract shape parameters from these fits, we do not normalize the total MC sample size to the analysis luminosity.

8.1.1 SIGNAL

Signal candidates are $B \rightarrow K\pi\pi\gamma$ events that are correctly reconstructed. Our fit to determine the $B \rightarrow K\pi\pi\gamma$ branching fractions treats signal and crossfeed on equal footing, so we do not have to distinguish the two. However, to extract the $m_{K\pi\pi}$ spectrum, we must be able to separate signal from crossfeed, at least on a statistical basis. We can distinguish signal candidates from crossfeed candidates in MC samples by looking at MC truth information. In data samples, we rely on their differing m_{ES} and ΔE distributions to distinguish them on a statistical basis.

We use the signal MC cocktail to determine the shape in m_{ES} and ΔE for the signal candidates. We apply the full analysis cuts to the MC sample, including the the fit region cut, and perform MCS, and then use a MC truth match to select correctly reconstructed candidates. In the final fit to the data, we fix the signal shape parameters to the values obtained from the fits to MC samples, and float only the sizes of the signal components to extract branching fractions.

We model the signal shape by a product of a Crystal Ball function [26] in m_{ES} and a Crystal Ball function in ΔE . The Crystal Ball shape is a Gaussian shape with an extended tail on the low side. This tail accommodates mismeasured m_{ES} and ΔE values, primarily due to energy leakage in the EMC of the high-energy photon. The expression for the standard Crystal Ball function is given below.

$$CB_1(x; \mu, \sigma, \alpha, n) = \begin{cases} \frac{1}{a} \left(\frac{n}{\alpha}\right)^n \frac{\exp(\alpha^2/2)}{((\mu-x)/\sigma+n/\alpha-\alpha)^n} & x < \mu - \alpha\sigma \\ \frac{1}{a} \exp\left[\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] & x > \mu - \alpha\sigma \end{cases}$$

The parameters μ and σ describe a Gaussian, which is truncated on the low side at $\mu - \alpha\sigma$ and joined continuously to a power function with exponent n ; $1/a$ is a normalization constant.

We find that the value of the mean parameter μ in the m_{ES} Crystal Ball function determined from signal MC is compatible with m_B , within fit errors, in all modes. We fix this parameter to m_B in the fit. We allow the parameter μ in the ΔE shape to float. We also find that large values of n in the ΔE shape describe the distributions adequately, so we fix n to infinity in the ΔE shape, which is equivalent to replacing the power-law tail of the Crystal Ball shape with an exponential tail.

Projection plots and parameters from the shape fits to truth-matched signal MC samples are presented below.

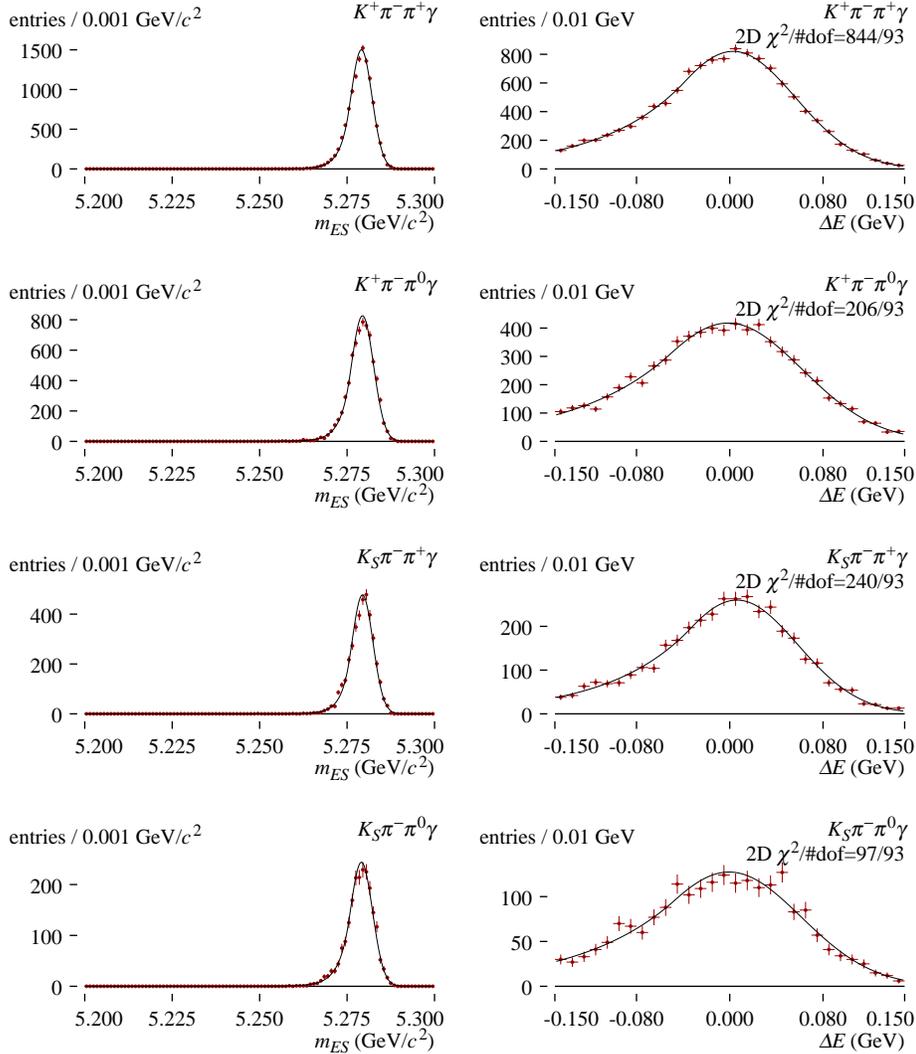


FIGURE 55. Projections of m_{ES} and ΔE fits of signal shape functions to truth-matched signal MC candidates. Candidate selection cuts, the fit region cut, and MCS have been applied to the candidates.

Parameter	$K^+\pi^-\pi^+\gamma$	$K^+\pi^-\pi^0\gamma$	$K_S^0\pi^-\pi^+\gamma$	$K_S^0\pi^-\pi^0\gamma$
$m_{ES} \sigma$	0.0030 ± 0.0000	0.0032 ± 0.0000	0.0030 ± 0.0000	0.0032 ± 0.0001
$m_{ES} \alpha$	1.2222 ± 0.0251	1.1996 ± 0.0309	1.2505 ± 0.0476	1.1101 ± 0.0481
$m_{ES} n$	20.0000 ± 0.7281	20.0000 ± 2.7765	20.0000 ± 2.8312	20.0000 ± 3.2139
$\Delta E \mu$	0.0028 ± 0.0010	-0.0014 ± 0.0016	0.0061 ± 0.0017	-0.0006 ± 0.0029
$\Delta E \sigma$	0.0536 ± 0.0008	0.0637 ± 0.0013	0.0524 ± 0.0013	0.0619 ± 0.0023
$\Delta E \alpha$	0.7616 ± 0.0310	0.7762 ± 0.0520	0.7513 ± 0.0509	0.7543 ± 0.0937

TABLE 33. Shape parameters obtained from $m_{ES}-\Delta E$ fit to truth-matched signal MC candidates. Candidate selection cuts, the fit region cut, and MCS have been applied to the candidates.

8.1.2 CROSSFEED

Crossfeed candidates are backgrounds produced in $B \rightarrow K\pi\pi\gamma$ decays and misreconstructed. The charge mode of the production process is not necessarily the same as the charge mode in which the candidate is reconstructed. A charged pion is often replaced by a neutral pion from the decay of the other B in the event, or vice versa. The replacement of a K^+ for a K_s^0 or vice versa is much less common. Examination of MC truth information of crossfeed candidates in signal MC indicate that the correct high-energy photon is reconstructed.

We include under the heading “crossfeed” all background candidates produced in $B \rightarrow K\pi\pi\gamma$ decays. These include background candidates produced in processes that we do not reconstruct in this analysis. Such processes are the two remaining charge modes, $B \rightarrow K\pi^0\pi^0\gamma$; $B \rightarrow K_L^0\pi\pi\gamma$ decays; and $B \rightarrow K_s^0\pi\pi\gamma$ decays where the K_s^0 decays to $\pi^0\pi^0$ or another final state. For the purposes of our fit, it is convenient to combine all of these processes into four crossfeed background categories, based on the production branching fraction that governs the yield in the category. To do this, we use the following relations:

- We can relate $\mathcal{B}(B \rightarrow K\pi^0\pi^0\gamma)$ to $\mathcal{B}(B \rightarrow K\pi^+\pi^-\gamma)$ or to $\mathcal{B}(B \rightarrow K\pi^+\pi^0\gamma)$, using model-dependent production ratios. We choose $\mathcal{B}(B \rightarrow K\pi^+\pi^-\gamma)$.

The production ratio of $K\pi^0\pi^0\gamma$ to $K\pi^+\pi^-\gamma$ is model-dependent. The ratio is 0.19 in our signal MC cocktail and 0.24 in the inclusive $b \rightarrow s\gamma$ MC. We use the former to model the crossfeed backgrounds, and substitute the inclusive $b \rightarrow s\gamma$ MC in studies of systematic uncertainty due to model dependence.

- We relate $\mathcal{B}(B \rightarrow K_L^0\pi\pi\gamma)$ to $\mathcal{B}(B \rightarrow K_s^0\pi\pi\gamma)$ using known branching fractions.
- We relate $\mathcal{B}(B \rightarrow K_s^0\pi\pi\gamma)$ where K_s^0 decays to other than $\pi^+\pi^-$ to the corresponding $K_s^0 \rightarrow \pi^+\pi^-$ process using known branching fractions.

Note that the rate of crossfeed backgrounds produced in modes that we do not reconstruct is a small part of our total crossfeed background, so our choice of treatment of these does not greatly affect our results.

The processes that we include in the four crossfeed background categories are listed in the table below. When we determine the crossfeed efficiencies and shape parameters for use in the fit, we combine together all the processes in each category using MC truth information, and treat each as a single production mode.

Category 1	Category 2	Category 3	Category 4
$K^+\pi^-\pi^+\gamma$	$K^+\pi^-\pi^0\gamma$	$K_S^0\pi^-\pi^+\gamma$	$K_S^0\pi^+\pi^0\gamma$
$K^+\pi^0\pi^0\gamma$		$K_S^0\pi^0\pi^0\gamma$	
		$K_L^0\pi^+\pi^-\gamma$	$K_L^0\pi^-\pi^0\gamma$
		$K_L^0\pi^0\pi^0\gamma$	
		$K_S^0\pi^+\pi^-\gamma, K_S^0 \rightarrow \pi^0\pi^0$	$K_S^0\pi^-\pi^0\gamma, K_S^0 \rightarrow \pi^0\pi^0$
		$K_S^0\pi^0\pi^0\gamma, K_S^0 \rightarrow \pi^0\pi^0$	
		$K_S^0\pi^+\pi^-\gamma, K_S^0 \rightarrow \text{other}$	$K_S^0\pi^-\pi^0\gamma, K_S^0 \rightarrow \text{other}$
		$K_S^0\pi^0\pi^0\gamma, K_S^0 \rightarrow \text{other}$	

TABLE 34. Production processes included in the four crossfeed background categories.

For crossfeed candidates, we use the same parameterization in m_{ES} that we use for signal candidates.

In ΔE , we find that the crossfeed shape does not peak appreciably, and we use a linear function to parameterize the distribution, written in the form

$$\text{Lin}(x) = \frac{1}{a}(1 + c_1x)$$

where a is a normalization constant.

We determine shape parameters for crossfeed backgrounds from the signal MC sample from which correctly reconstructed candidates have been removed with a veto on MC truth information. The fit parameters and plots of the projections of the crossfeed shape fits to signal MC are shown below. Note that in our implementation of the Crystal Ball function, a value of the parameter n above 100 is used to indicate an infinite value, which produces an exponential tail to the shape

function. Because of the way that the fits are specified, values of n slightly above 100 may appear in the results, but these are equivalent; the fit uncertainties on these values are meaningless.

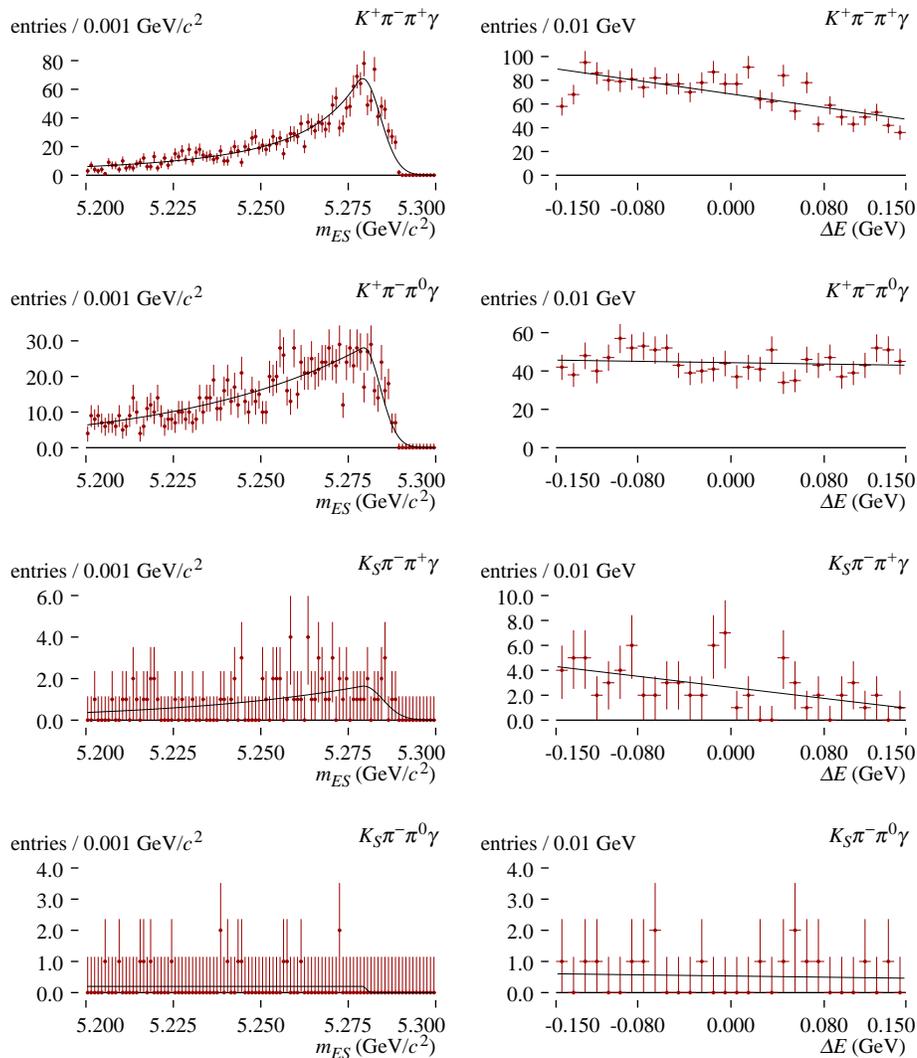


FIGURE 56. Projections of m_{ES} and ΔE of fits to truth-vetoed candidates in signal MC produced in category 1 and reconstructed in each of the four modes.

Param.	$K^+\pi^-\pi^+\gamma$	$K^+\pi^-\pi^0\gamma$	$K_S^0\pi^-\pi^+\gamma$	$K_S^0\pi^-\pi^0\gamma$
$m_{ES} \sigma$	0.0051 ± 0.0002	0.0047 ± 0.0002	0.0055 ± 0.0010	0.0010 ± 0.0020
$m_{ES} \alpha$	0.3023 ± 0.0402	0.0885 ± 0.0063	0.1033 ± 0.0348	0.0066 ± 5.0189
$m_{ES} n$	1.9570 ± 0.4830	100.0321 ± 0.1238	100.2232 ± 73.0874	0.0000 ± 0.1105
$\Delta E c_1$	-2.0636 ± 0.2462	-0.2008 ± 0.3108	-4.2195 ± 1.1346	-0.8861 ± 2.8941

TABLE 35. Shape parameters obtained from $m_{ES}-\Delta E$ fits to truth-vetoed candidates in signal MC produced in category 1 and reconstructed in each of the four modes.

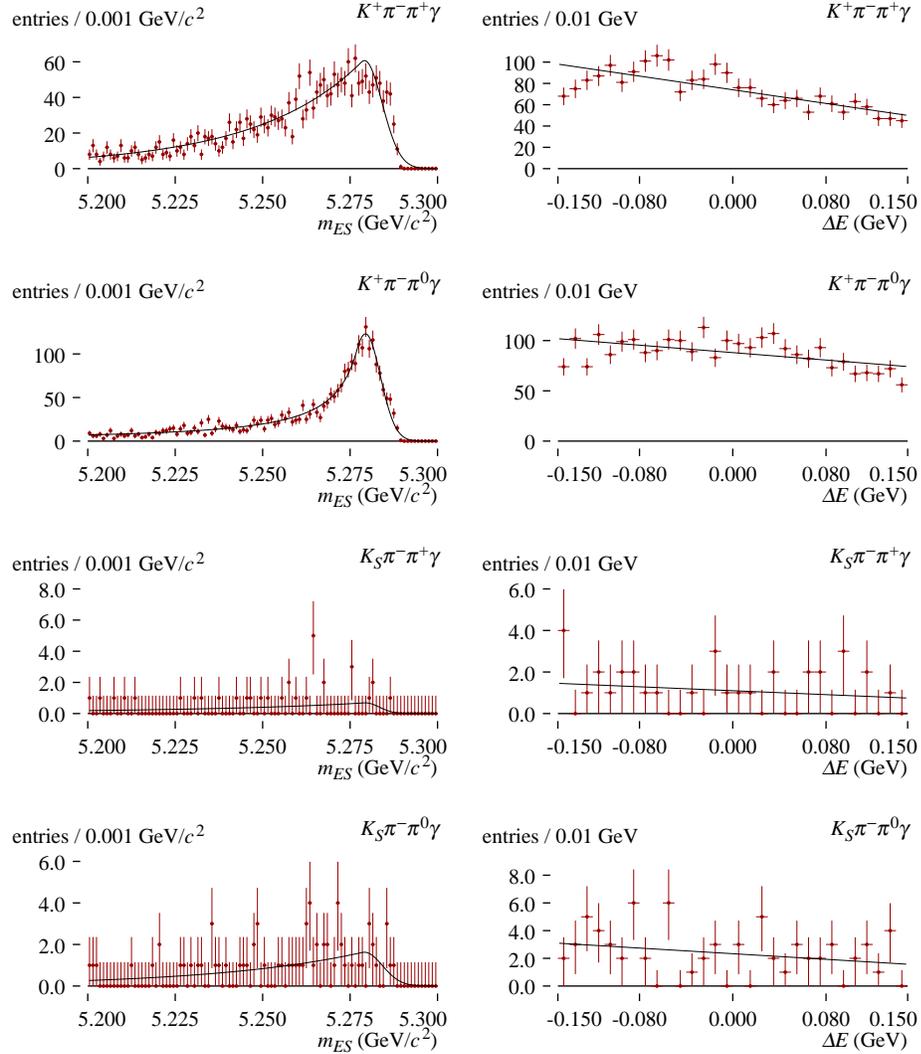


FIGURE 57. Projections of m_{ES} and ΔE of fits to truth-vetoed crossfeed candidates from the signal MC sample produced in category 2 and reconstructed in each of the four modes.

Param.	$K^+\pi^-\pi^+\gamma$	$K^+\pi^-\pi^0\gamma$	$K_S^0\pi^-\pi^+\gamma$	$K_S^0\pi^-\pi^0\gamma$
$m_{ES} \sigma$	0.0051 ± 0.0002	0.0043 ± 0.0001	0.0039 ± 0.0013	0.0051 ± 0.0010
$m_{ES} \alpha$	0.1649 ± 0.0157	0.5876 ± 0.0505	0.0780 ± 0.0834	0.1185 ± 0.0321
$m_{ES} n$	9.4627 ± 8.8236	1.1575 ± 0.1072	3.2347 ± 18.1531	100.0384 ± 0.0727
$\Delta E c_1$	-2.1680 ± 0.2449	-1.0497 ± 0.2326	-2.1477 ± 1.8589	-2.1740 ± 1.2899

TABLE 36. Shape parameters obtained from $m_{ES}-\Delta E$ fits to truth-vetoed candidates in signal MC produced in category 2 and reconstructed in each of the four modes.

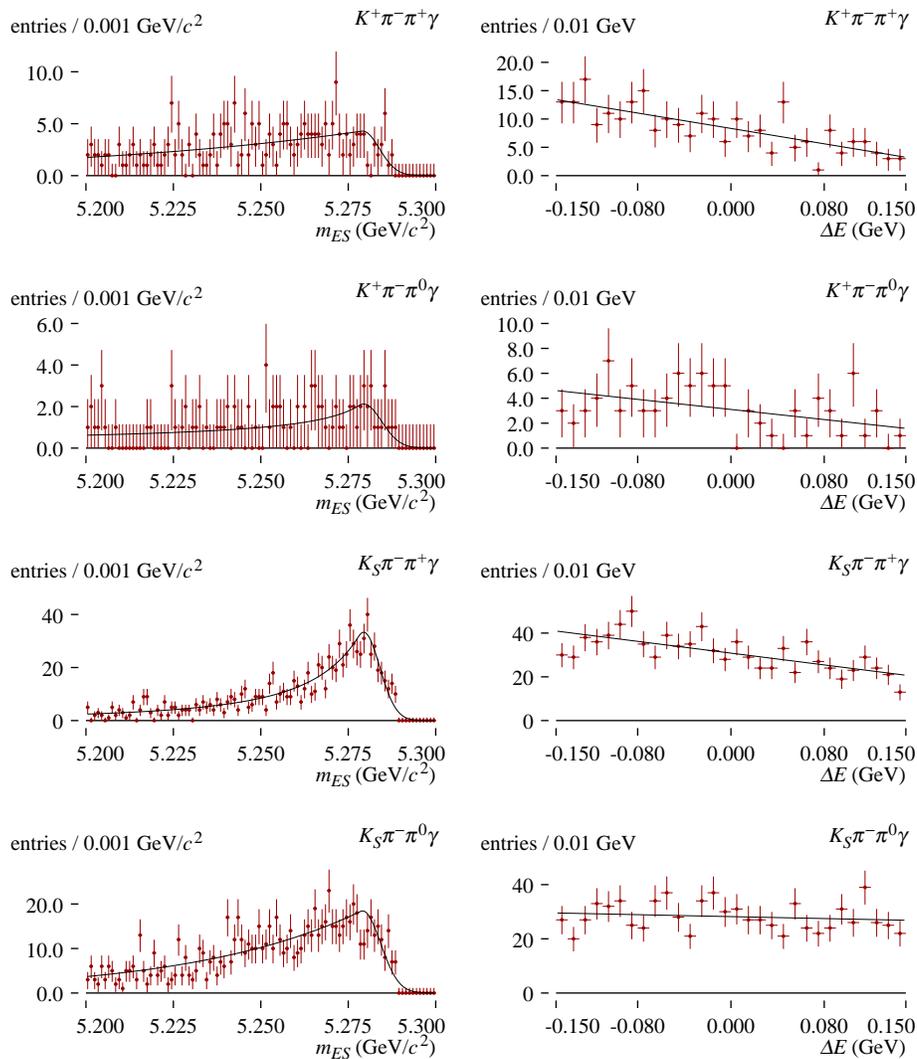


FIGURE 58. Projections of m_{ES} and ΔE of fits to truth-vetoed crossfeed candidates from the signal MC sample produced in category 3 and reconstructed in each of the four modes.

Param.	$K^+\pi^-\pi^+\gamma$	$K^+\pi^-\pi^0\gamma$	$K_S^0\pi^-\pi^+\gamma$	$K_S^0\pi^+\pi^0\gamma$
$m_{ES} \sigma$	0.0048 ± 0.0006	0.0052 ± 0.0009	0.0046 ± 0.0002	0.0051 ± 0.0003
$m_{ES} \alpha$	0.0546 ± 0.0134	0.2622 ± 0.3620	0.3076 ± 0.0500	0.1052 ± 0.0092
$m_{ES} n$	100.1480 ± 2.3005	0.5784 ± 0.7230	2.0151 ± 0.5567	100.5922 ± 71.1335
$\Delta E c_1$	-4.0713 ± 0.6335	-3.2679 ± 1.1756	-2.1848 ± 0.3698	-0.3178 ± 0.4028

TABLE 37. Shape parameters obtained from $m_{ES}-\Delta E$ fits to truth-vetoed candidates in signal MC produced in category 3 and reconstructed in each of the four modes.

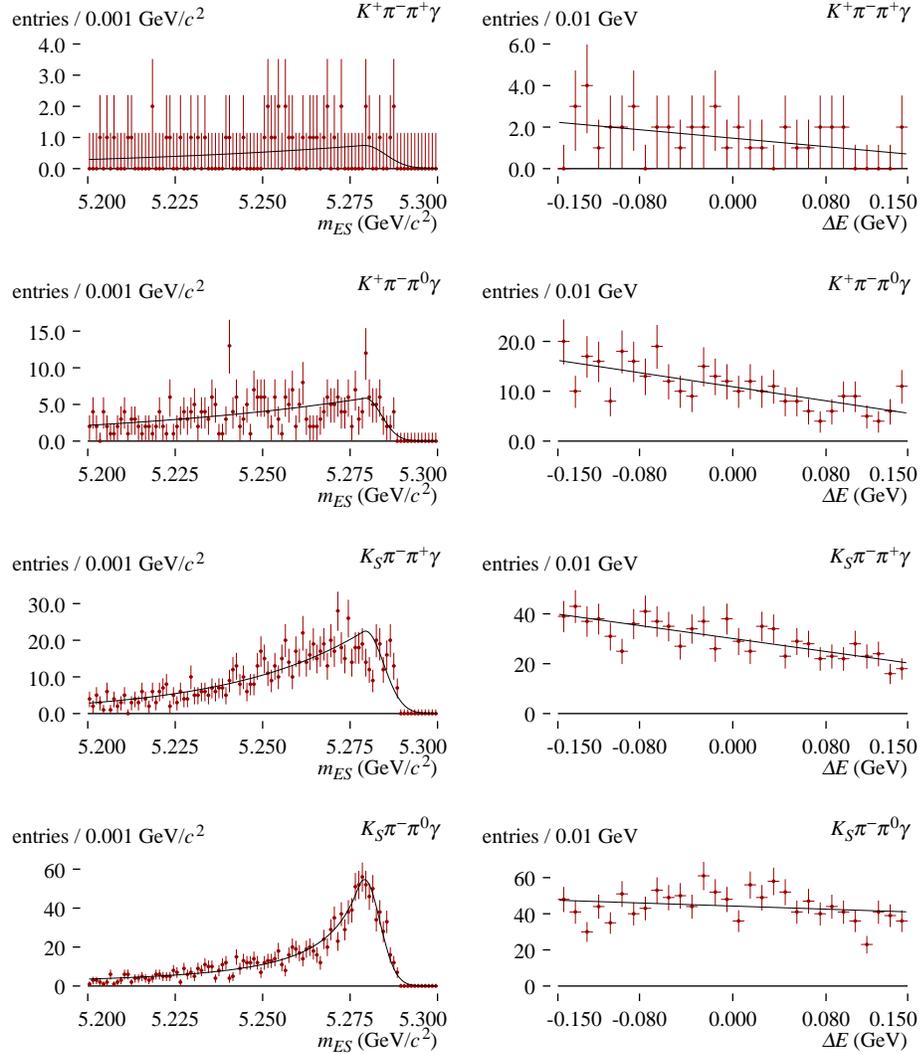


FIGURE 59. Projections of m_{ES} and ΔE of fits to truth-vetoed crossfeed candidates from the signal MC sample produced in category 4 and reconstructed in each of the four modes.

Param.	$K^+\pi^-\pi^+\gamma$	$K^+\pi^-\pi^0\gamma$	$K_S^0\pi^-\pi^+\gamma$	$K_S^0\pi^-\pi^0\gamma$
$m_{ES} \sigma$	0.0061 ± 0.0016	0.0045 ± 0.0005	0.0052 ± 0.0003	0.0044 ± 0.0002
$m_{ES} \alpha$	0.0716 ± 0.0401	0.0574 ± 0.0112	0.1357 ± 0.0095	0.4313 ± 0.0639
$m_{ES} n$	100.8403 ± 50.5633	100.0031 ± 0.0350	100.5768 ± 1.0648	1.4465 ± 0.2659
$\Delta E c_1$	-3.4739 ± 1.6277	-3.2373 ± 0.5697	-2.1660 ± 0.3781	-0.4841 ± 0.3329

TABLE 38. Shape parameters obtained from $m_{ES}-\Delta E$ fits to truth-vetoed candidates in signal MC produced in category 4 and reconstructed in each of the four modes.

8.1.3 CONTINUUM AND GENERIC B

In our data fit, we do not fix from MC the shape of continuum $udsc$ and generic $B\bar{B}$ background. Instead, we allow the shape parameters of our continuum fit function to float, along with the overall normalization. We do, however, fit the MC samples of these processes to verify that our parameterization is adequate.

The shapes of distributions in m_{ES} and ΔE of backgrounds from both continuum $udsc$ and generic $B\bar{B}$ processes are each well-parameterized by the product of Argus and exponential functions, albeit with different parameter values. However, the large background rate from $udsc$ swamps the generic $B\bar{B}$ background, and distinguishing the two components in the fit without fixing both their shapes is not possible.

We model the distribution of these backgrounds as a product of an Argus function [27] in m_{ES} and an exponential function in ΔE . The Argus function, with which we model the m_{ES} distribution, is given by

$$\text{Argus}(x; \zeta, E_b) = \frac{1}{a} x \sqrt{1 - x^2/E_b^2} e^{-\zeta \sqrt{1 - x^2/E_b^2}} .$$

The parameter ζ is the Argus shape parameter, the cutoff E_b is at the nominal beam energy, and $1/a$ is a factor to normalize the integral of the distribution to unity over the fit region. The ΔE parameterization is given simply by

$$\text{Ex}(x; s) = \frac{1}{a} e^{sx} .$$

The parameter s determines the exponential shape, and $1/a$ is a normalization constant.

We studied three methods of including the $B\bar{B}$ background contribution in the fit². For each technique, we generated 500 toy MC samples thrown from the signal and background distributions determined from MC samples, performed fits to each, and examined the mean and RMS over the

²This study was performed with a previous, slightly different version of the candidate selection cuts.

ensemble of toy samples of the resulting fit parameters. We included in the toy MC samples contributions from $b \rightarrow s\gamma$ background processes. The fit functions included signal terms, fixed terms for $b \rightarrow s\gamma$ background processes, and one or two additional terms, as described below.

1. We used separate $udsc$ and $B\bar{B}$ terms in the fit function. The Argus and exponential shape parameters of both were allowed to vary in the fit along with their normalizations.
2. We included a separate $B\bar{B}$ contribution, but fixed the Argus and exponential shape parameters to values obtained from a separate fit to the $B\bar{B}$ MC sample. The normalization of the $B\bar{B}$ component, and the normalization and shape parameters of the $udsc$ contribution were allowed to vary in the fit.
3. We did not include a separate $B\bar{B}$ fit contribution, forcing the $B\bar{B}$ background events to be absorbed into the $udsc$ contribution.

The first method produced wildly unreliable fit results, as there are not sufficient statistics in our samples to determine the $B\bar{B}$ and $udsc$ shape parameters simultaneously in the fit.

The second method overestimated the $B\bar{B}$ background contribution by a factor of two to three, at the expense of the $udsc$ background contribution. The fit, however, determined the sum of the $B\bar{B}$ and $udsc$ background yields correctly. The fitted values of the signal yield showed a negative bias, especially in the K^+ modes, where the $B\bar{B}$ background contributions are larger. The results of these toy MC studies are shown below.

	$K^+\pi^-\pi^+\gamma$	$K^+\pi^-\pi^0\gamma$	$K_S^0\pi^-\pi^+\gamma$	$K_S^0\pi^+\pi^0\gamma$
input signal yield	659.9	365.7	190.2	70.5
input $udsc$ yield	6738.5	3644.8	1873.7	1544.4
input $B\bar{B}$ yield	270.9	165.7	87.3	69.2
fit signal yield	652.4 ± 39.3	357.6 ± 31.9	187.4 ± 21.8	68.3 ± 17.4
fit $udsc$ yield	6305.7 ± 371.1	3513.1 ± 221.2	1680.6 ± 165.5	1397.8 ± 159.8
fit $B\bar{B}$ yield	717.6 ± 374.3	306.5 ± 222.8	228.7 ± 169.2	221.3 ± 159.6
signal yield bias	-7.5 ± 1.2	-8.1 ± 1.0	-2.8 ± 0.7	-2.2 ± 0.6

TABLE 39. Input and resulting fit yields for 1,000 toy MC fit studies. The shape parameters of the $B\bar{B}$ background component were fixed. The shape parameters of the $udsc$ component were allowed to float.

The third method produced reasonable fits to the data, and correctly determined the sum of the $B\bar{B}$ and $udsc$ background yields. The signal yields showed a smaller bias, only in the K^+ modes, at the limit of statistical significance. The results of these toy MC studies are shown below.

	$K^+\pi^-\pi^+\gamma$	$K^+\pi^-\pi^0\gamma$	$K_S^0\pi^-\pi^+\gamma$	$K_S^0\pi^+\pi^0\gamma$
input signal yield	659.9	365.7	190.2	70.5
input $udsc$ yield	6738.5	3644.8	1873.7	1544.4
input $B\bar{B}$ yield	270.9	165.7	87.3	69.2
fit signal yield	656.8 ± 41.2	362.3 ± 31.6	192.0 ± 20.8	70.3 ± 17.2
fit $udsc$ yield	7015.3 ± 88.6	3818.7 ± 67.4	1963.6 ± 47.3	1617.1 ± 43.7
signal yield bias	-3.1 ± 1.3	-3.4 ± 1.0	1.8 ± 0.7	-0.2 ± 0.5

TABLE 40. Input and resulting fit yields for 1,000 toy MC fit studies. No separate $B\bar{B}$ component was included in the fit; $B\bar{B}$ background events were absorbed into the $udsc$ component.

Based on these studies, we choose to combine backgrounds from continuum $udsc$ processes and generic B decays. The generic B background component includes all B decay processes *except* for those from $b \rightarrow s\gamma$ decays, which we handle separately. The projection plots and fit parameters of validation fits to combined $udsc$ and generic $B\bar{B}$ MC samples are below.

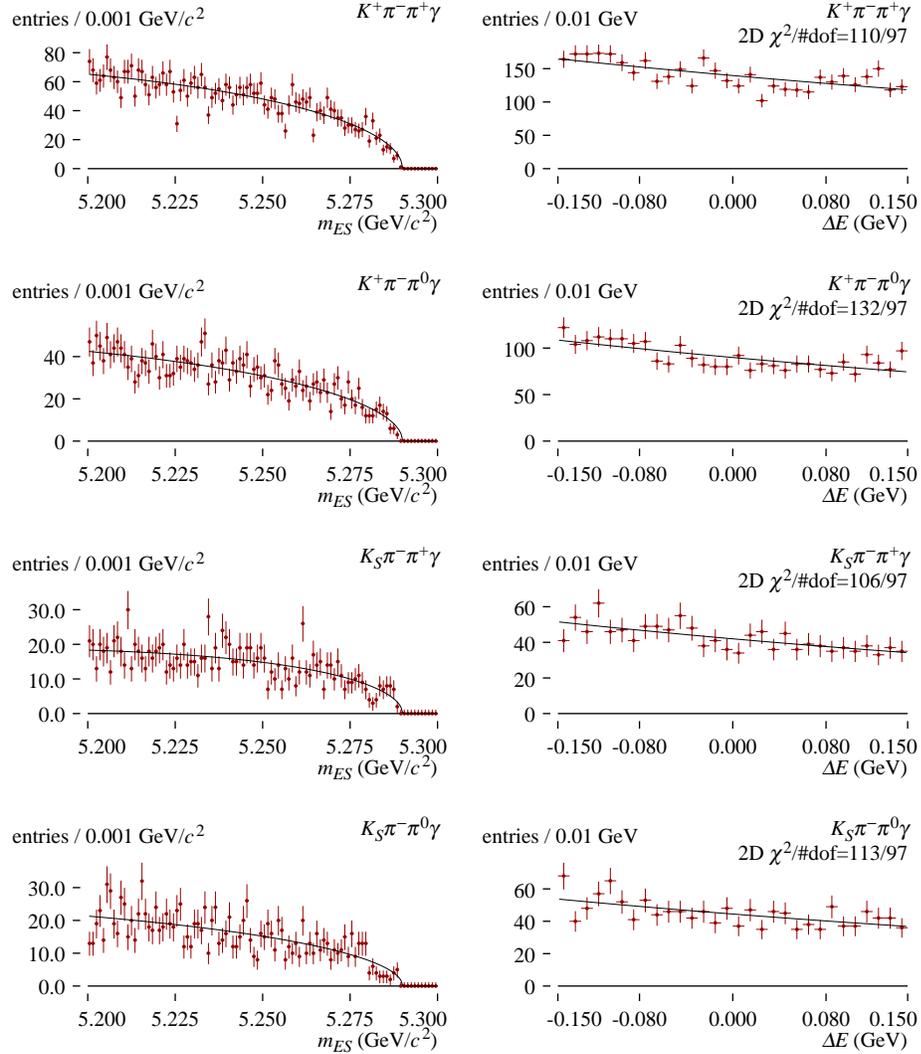


FIGURE 60. Projections of m_{ES} and ΔE of fits to events from $udsc$ MC and generic $B\bar{B}$ MC (with $b \rightarrow s\gamma$ and peaking decays removed).

Parameter	$K^+\pi^-\pi^+\gamma$	$K^+\pi^-\pi^0\gamma$	$K_S^0\pi^-\pi^+\gamma$	$K_S^0\pi^+\pi^0\gamma$
m_{ES} Argus ζ	4.756 ± 1.733	3.859 ± 2.161	9.618 ± 3.139	2.906 ± 3.077
ΔE exponential s	-1.105 ± 0.178	-1.260 ± 0.222	-1.366 ± 0.326	-1.260 ± 0.316

TABLE 41. Shape parameters obtained from $m_{ES}-\Delta E$ fits to events from $udsc$ MC and generic $B\bar{B}$ MC (with $b \rightarrow s\gamma$ and peaking decays removed).

8.2 Combined fit

Using the shape functions described in the previous section, we perform a simultaneous fit in all four modes to determine the four signal branching fractions in our data sample. The data sample is divided by run and mode into sixteen subsamples, to account for efficiency differences among runs, but all are fitted simultaneously.

8.2.1 LIKELIHOOD FUNCTION

We write the likelihood function as a sum of five pieces, namely,

1. correctly reconstructed signal,
2. crossfeed background,
3. combined $udsc$ and $B\bar{B}$ background,
4. background from $B \rightarrow K\pi\gamma$, and
5. background from other $b \rightarrow s\gamma$ decays.

The likelihood function for a candidate in run r reconstructed in mode m with measured m_{ES} and ΔE is,

$$\begin{aligned} \mathcal{L}(m, r, m_{\text{ES}}, \Delta E) = N_{B\bar{B}}^r \cdot 2 \cdot \frac{1}{2} & \left(\mathcal{B}^m \mathcal{B}_2^m \epsilon_s^{m,r} f_s^m(m_{\text{ES}}, \Delta E) + \sum_g \mathcal{B}^g \mathcal{B}_2^g \epsilon_x^{g,m,r} f_x^{gm}(m_{\text{ES}}, \Delta E) \right) + \\ & + L^r \sigma_c^{m,r} f_c^m(m_{\text{ES}}, \Delta E; \zeta^m, s^m) + \sum_b n_b^{m,r} f_b^m(m_{\text{ES}}, \Delta E) \end{aligned}$$

where

- $N_{B\bar{B}}^r$ is the $B\bar{B}$ event count for run r , the factor of 2 accounts for two B mesons in each event, and $1/2$ is the branching fraction both of $\Upsilon(4S) \rightarrow B^0\bar{B}^0$ and $\Upsilon(4S) \rightarrow B^+B^-$;
- \mathcal{B}^m is the $B \rightarrow K\pi\pi\gamma$ branching fraction for mode m , which floats in the fit;

- \mathcal{B}_2^m is the secondary branching fraction for mode m to the final states we reconstruct, which includes $\pi^0 \rightarrow \gamma \gamma$ and $K_s^0 \rightarrow \pi^+ \pi^-$ branching fractions, fixed to world average values;
 - $\epsilon_s^{m,r}$ is the efficiency for a signal MC candidate produced in mode m and run r to be reconstructed correctly;
 - f_s^m is the PDF for correctly reconstructed signal candidates in mode m ;
 - the index g is the crossfeed background category, as defined in section 8.1.2;
 - $\epsilon_x^{g,m,r}$, the crossfeed efficiency matrix, is the efficiency for a signal MC event produced in category g with run r conditions to be misreconstructed in mode m ;
 - $f_x^{g,m}$ is the PDF for crossfeed backgrounds from mode g to mode m , whose shape parameters are determined from fits from MC;
 - L^r is the integrated luminosity of run r in fb^{-1} ,
 - $\sigma_c^{m,r}$ is the number of $udsc$ and $B\bar{B}$ background events in mode m and run r per fb^{-1} of integrated luminosity, which floats in the fit;
 - f_c^m is the PDF for $udsc$ and $B\bar{B}$ background events;
 - ζ^m and s^m are the Argus and exponential shape parameters, respectively, for $udsc$ and $B\bar{B}$ background events in mode m , which float in the fit;
 - the index b runs over the remaining background categories: feed-up from $B \rightarrow K\pi\gamma$ decays and feed-down from $b \rightarrow s\gamma$ decays;
 - $n_b^{m,r}$ is the number of background events in category b in mode m and run r , fixed from MC;
- and
- f_b^m is the PDF for background category b in mode m , determined as a binned distribution from MC.

The first term in the likelihood, in parentheses, accounts for signal and crossfeed. The second term accounts for the $udsc$ and $B\bar{B}$ backgrounds. The third term sums over the remaining two background contributions. Note that the signal component (the first term in the parentheses) and the self-crossfeed term (the $g = m$ term in the first sum) are both multiplied by the same floating scale factor \mathcal{B}^m , and can together be regarded as a single term in the likelihood.

Twenty-eight parameters float in the fit: the four branching fractions we seek to measure, \mathcal{B}^m ; 16 scale factors for $udsc$ and $B\bar{B}$ backgrounds, $\sigma_c^{m,r}$; and eight shape parameters of the $udsc$ and $B\bar{B}$ background distributions, ζ^m and s^m .

The combined likelihood is the product of likelihoods over candidates in the sample, times a Poisson factor $e^{-k}k^n/n!$ for each run r and mode m , where n is the number of candidates in the sample and k is the integral of \mathcal{L} . We perform a maximum likelihood fit using *Minuit* to determine the parameter values.

8.2.2 PRECISION AND BIAS

We study the precision and bias of this fit procedure using toy MC samples. Each toy MC sample is a set of $(m_{\text{ES}}, \Delta E, m_{K\pi\pi})$; we do not use the $m_{K\pi\pi}$ values for fit studies, but for studying our procedure for determining the $m_{K\pi\pi}$ distribution described later in this chapter. The candidates included in each toy sample are as follows,

- candidates sampled from signal MC, chosen by acceptance-rejection using luminosity and efficiency weights;
- m_{ES} and ΔE thrown from distributions fitted to $udsc$ MC, and $m_{K\pi\pi}$ thrown from a binned distribution obtained from $udsc$ MC;
- m_{ES} and ΔE thrown from distributions fitted to generic $B\bar{B}$ MC, and $m_{K\pi\pi}$ thrown from a binned distribution obtained from $B\bar{B}$ MC; and

- feed-up and feed-down background candidates sampled from $K\pi\gamma$ and inclusive $b \rightarrow s\gamma$ MC samples, respectively, chosen by acceptance-rejection using luminosity and efficiency weights.

By sampling signal, crossfeed, feed-up, and feed-down candidates from MC samples, we preserve correlations among these variables that may be present. We do not find substantial correlations among them in $udsc$ or generic $B\bar{B}$ samples; in these samples, we do not have large MC statistics from which to draw toy samples, so we generate values from statistical distributions. We normalize the size of each component to the data luminosity, varying the number of candidates within counting uncertainty. The mean total size of each toy MC sample in each mode is the total expected yield for signal and background (see table 26): 8908 candidates in $K^+\pi^-\pi^+\gamma$, 5920 in $K^+\pi^-\pi^0\gamma$, 2685 in $K_s^0\pi^-\pi^+\gamma$, and 2116 in $K_s^0\pi^+\pi^0\gamma$.

We generate 500 such toy samples. The subsets candidates drawn from MC in the toy samples are largely but not entirely independent; the statistically smallest MC samples we use represent about 22 times the analysis luminosity. We fit each toy sample using the fit procedure described above. The following figure illustrates one of the toy fits.

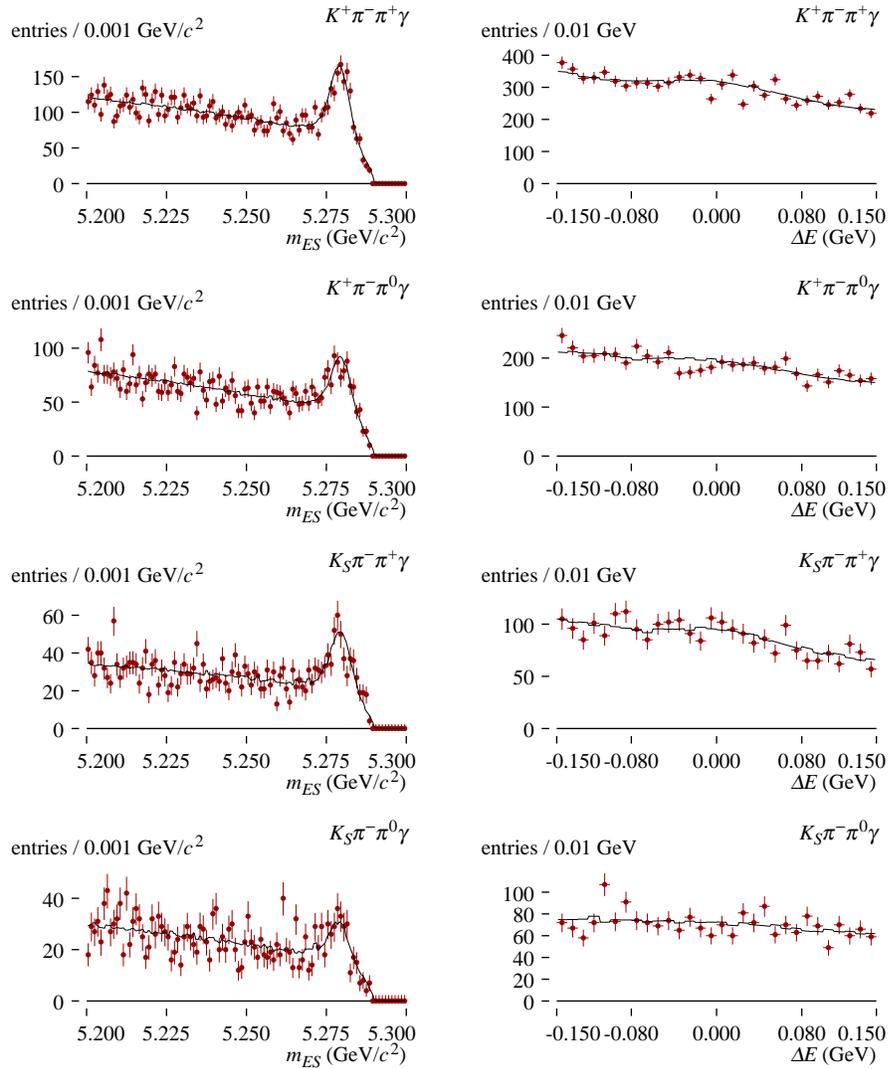


FIGURE 61. Projections of m_{ES} and ΔE of the combined fit to one toy MC sample, summed over runs. Error bars indicate statistical uncertainty on the bin contents.

The table below summarizes the results of these fits. For each parameter, the table gives the mean fit value and the square root of the sample variance over the 500 toy samples.

Parameter	$K^+\pi^-\pi^+\gamma$	$K^+\pi^-\pi^0\gamma$	$K_S^0\pi^-\pi^+\gamma$	$K_S^0\pi^+\pi^0\gamma$
$\mathcal{B} \times 10^6$	24.13 ± 1.11	29.50 ± 2.06	24.47 ± 1.86	28.93 ± 4.63
n_c^1	35.89 ± 1.55	21.10 ± 1.13	10.37 ± 0.81	7.01 ± 0.76
n_c^2	35.39 ± 0.71	22.54 ± 0.80	10.33 ± 0.46	8.59 ± 0.42
n_c^3	34.70 ± 0.99	21.07 ± 0.92	9.73 ± 0.67	8.50 ± 0.63
n_c^4	33.67 ± 0.67	21.83 ± 0.60	9.98 ± 0.38	8.44 ± 0.37
ζ	1.71 ± 1.53	3.51 ± 1.70	3.79 ± 2.40	2.51 ± 2.61
s	-1.08 ± 0.17	-1.25 ± 0.22	-1.31 ± 0.23	-1.13 ± 0.27

TABLE 42. Mean and square RMS of parameters from combined fits to 500 toy MC samples. All parameters in all modes are fit simultaneously. The parameters n_c^r are the number of $udsc$ and $B\bar{B}$ events in run r .

The table below shows estimates of the bias on the fit values of the branching fractions in each mode. The bias is the difference between the mean fit value of the branching fraction over the 500 toy MC samples minus the branching fraction value used when generating the toy MC sample. The stated uncertainty is the precision on the estimate of the mean: the square root of the sample variance divided by the square root of the number of toy samples.

Mode	\mathcal{B} bias $\times 10^6$	Uncertainty $\times 10^6$
$B \rightarrow K^+\pi^-\pi^+\gamma$	0.390	0.056
$B \rightarrow K^+\pi^-\pi^0\gamma$	-0.383	0.093
$B \rightarrow K_S^0\pi^-\pi^+\gamma$	0.297	0.094
$B \rightarrow K_S^0\pi^+\pi^0\gamma$	1.103	0.179

TABLE 43. Estimates of bias on branching fractions values from fits to 20 toy MC samples. The uncertainty in each mode is the statistical uncertainty on the estimate of the mean value over the toy fits.

8.3 $m_{K\pi\pi}$ distribution

We use the sPlot method [28] to measure the distribution of $m_{K\pi\pi}$ in our signal events. The sPlot method assigns to each candidate a weight computed from the distributions of measured variables (m_{ES} and ΔE) in the categories of events in the sample. The categories are signal, the four crossfeed

components, $udsc$ & generic $B\bar{B}$, feed-up, and feed-down. Using these weights, it is possible to extract the distribution of another measured variable ($m_{K\pi\pi}$) for a single category.

This method poses an important advantage in our analysis: unlike a subtraction scheme, it does not require prior knowledge of the background distributions of the variable we seek to measure in the signal. In particular, the $m_{K\pi\pi}$ distribution of crossfeed candidate depends on the signal $m_{K\pi\pi}$ distribution (albeit loosely, since a crossfeed candidate is a signal candidate with one track replaced by another from the other B in the event). To subtract the crossfeed $m_{K\pi\pi}$ distribution from our sample would require us to know the distribution that we seek to measure. In addition, the hadronic mass spectrum in other $b \rightarrow s\gamma$ processes is not precisely known, which limits our understanding the $m_{K\pi\pi}$ distribution in feed-up and feed-down background candidates. Using the sPlot technique, we do not need to know these distributions.

To make an sPlot, we assign a weight to each candidate for each category i given by,

$$w_i(m_{\text{ES}}, \Delta E) = \frac{\sum_j V_{ij} f_j(m_{\text{ES}}, \Delta E)}{\sum_j n_j f_j(m_{\text{ES}}, \Delta E)}$$

where the index over j ranges over the categories, f_j is the joint $m_{\text{ES}}-\Delta E$ PDF for category j , and n_j is the number of events in category j . The covariance matrix V is given by,

$$(V^{-1})_{ij} = \sum_e \frac{f_i(m_{\text{ES}}^e, \Delta E^e) f_j(m_{\text{ES}}^e, \Delta E^e)}{\sum_k n_k f_k(m_{\text{ES}}^e, \Delta E^e)}$$

where i , j , and k are category indices, e indexes the candidates in the sample, and m_{ES}^e and ΔE^e are the measured values for candidate e .

We construct histograms of the $m_{K\pi\pi}$ distributions for the candidates in our sample, using the weight for each candidate computed by the sPlot formula above. The statistical uncertainty on each bin is the RMS of weights of candidates in that bin. The following figure shows the sPlot distributions from one toy MC sample.

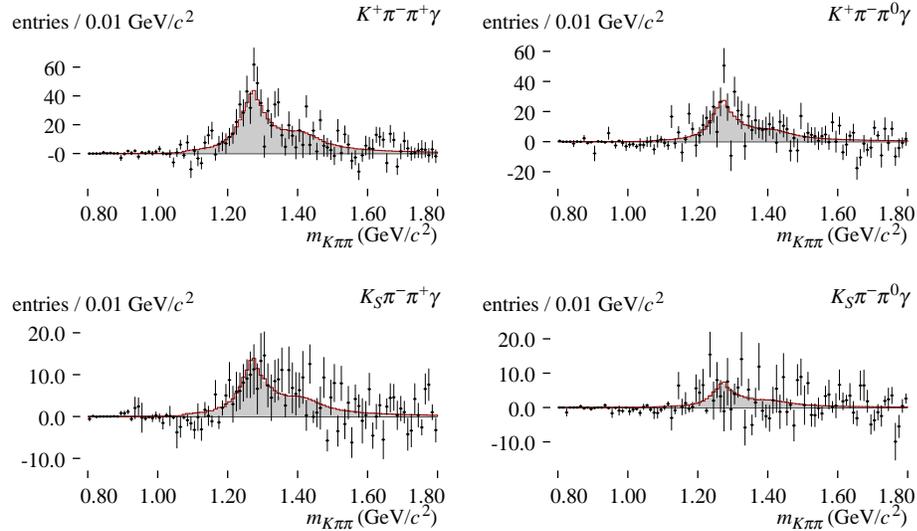


FIGURE 62. Distributions of $m_{K\pi\pi}$ in a toy MC sample (points). Error bars show statistical uncertainty from the sPlot only. The shaded histogram shows the distribution for truth-matched candidates in the signal MC cocktail sample. The red line shows the generator-level distribution in the signal MC sample.

To test the validity of the sPlot procedure, we apply it to the 500 toy MC samples we used to study the fit procedure. We can construct $m_{K\pi\pi}$ distributions not only for the signal component, but for other components as well, and compare them to the MC distributions we used to generate the toy samples. From each sample, we construct the $m_{K\pi\pi}$ sPlot for each component in each mode, adding together the distributions from four runs.

The following figure shows $m_{K\pi\pi}$ plots for the signal component, with the bin values averaged over the 500 toy samples. The error bars show the statistical uncertainty on the mean over the top samples.

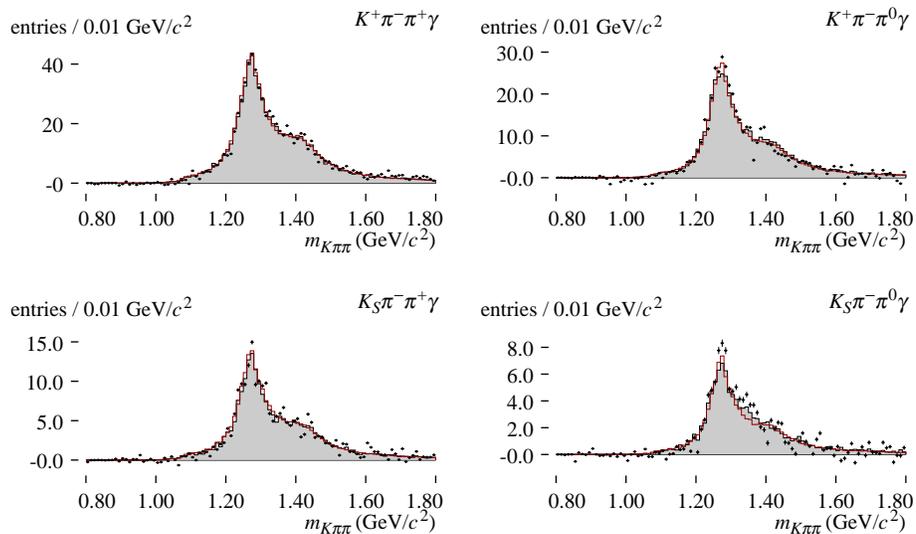


FIGURE 63. Distribution averaged over 500 toy MC samples of $m_{K\pi\pi}$ weighted with sPlot weights of the signal component (points). The error bars show the square root of the sample variance. The shaded histograms show the MC-truth distributions of $m_{K\pi\pi}$ events for truth-matched signal MC candidates passing all cuts. The generator-level distributions for the entire signal MC samples, scaled by signal efficiencies, is shown in red.

The following two figures show the sPlot $m_{K\pi\pi}$ distributions using the weights for the $udsc + B\bar{B}$ background category, and for the $b \rightarrow s\gamma$ feed-down background category. The sPlot distributions are compared to the $m_{K\pi\pi}$ distributions from the corresponding MC samples.

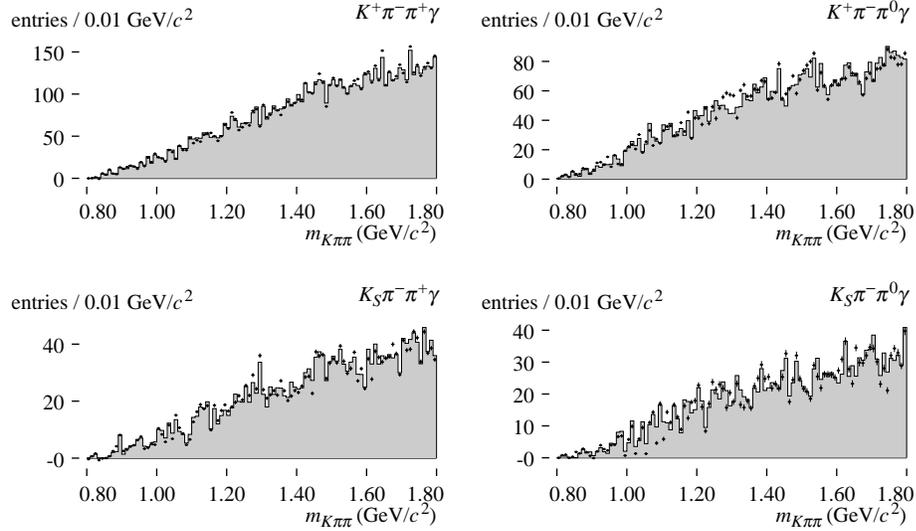


FIGURE 64. Distribution averaged over 500 toy MC samples of $m_{K\pi\pi}$ weighted with sPlot weights of the $udsc + B\bar{B}$ component (points). The error bars show the square root of the sample variance. The shaded histograms show the MC-truth distributions of $m_{K\pi\pi}$ events for candidates $udsc$ and generic $B\bar{B}$ MC passing all cuts.

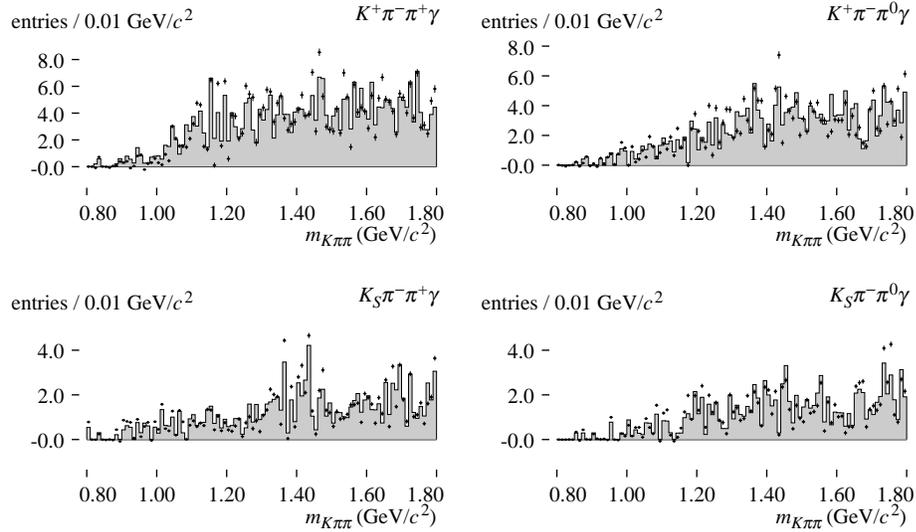


FIGURE 65. Distribution averaged over 500 toy MC samples of $m_{K\pi\pi}$ weighted with sPlot weights of the $b \rightarrow s\gamma$ feed-down component (points). The error bars show the square root of the sample variance. The shaded histograms show the MC-truth distributions of $m_{K\pi\pi}$ events for feed-down $b \rightarrow s\gamma$ MC candidates passing all cuts.

The following plot shows the $m_{K\pi\pi}$ bias: the difference between the mean value in each toy MC sPlot for the signal category, and the $m_{K\pi\pi}$ distribution for truth-matched signal events passing all cuts.

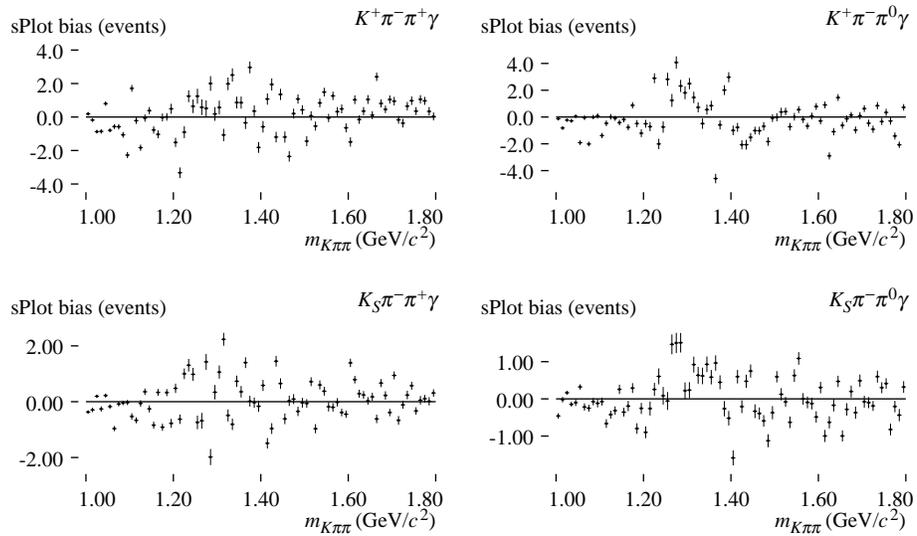


FIGURE 66. Difference between mean of sPlot $m_{K\pi\pi}$ distributions from 500 toy MC samples and generated $m_{K\pi\pi}$ distribution of signal MC samples. Error bars show the uncertainty on the bias, namely the square root of the sample variance divided by the square root of the number of toy MC samples.

9 Results

This chapter presents the results of applying the selection and fit procedures described in previous chapters to the on-resonance data sample. We unblind the signal region only after finalizing the selection and fit procedures.

9.1 Branching fraction fit

The table below lists parameter values obtained from the simultaneous fit to all modes of the on-resonance data sample. Uncertainties are obtained from the fit. (The parabolic errors computed by default in MINUIT do not differ substantially from those computed with the MINOS algorithm.)

Parameter	$K^+\pi^-\pi^+\gamma$	$K^+\pi^-\pi^0\gamma$	$K_S^0\pi^-\pi^+\gamma$	$K_S^0\pi^+\pi^0\gamma$
$\mathcal{B} \times 10^6$	29.487 ± 1.258	40.737 ± 2.201	18.503 ± 2.055	45.600 ± 4.196
n_c^1	41.525 ± 1.554	23.178 ± 1.178	12.296 ± 0.841	7.653 ± 0.673
n_c^2	43.214 ± 0.899	23.481 ± 0.672	12.002 ± 0.476	8.674 ± 0.403
n_c^3	40.653 ± 1.206	24.509 ± 0.950	12.638 ± 0.674	9.151 ± 0.578
n_c^4	40.459 ± 0.687	24.141 ± 0.541	10.986 ± 0.363	8.614 ± 0.322
$m_{\text{ES}} \text{ Argus } \zeta$	0.019 ± 0.001	4.872 ± 1.950	1.807 ± 2.634	0.029 ± 0.002
$\Delta E \text{ exponential } s$	-1.189 ± 0.131	-1.132 ± 0.174	-1.244 ± 0.248	-0.875 ± 0.289

TABLE 44. Parameters obtained from the combined fit to the on-resonance data sample.

We estimate the goodness of fit using a log likelihood ratio statistic on the binned $m_{\text{ES}}-\Delta E$ distributions, assuming bin contents to be Poisson distributed. The statistic is given by,

$$-\log \lambda = -\log \frac{\prod_i e^{-\mu_i} \mu_i^{n_i} / n_i!}{\prod_i e^{-n_i} n_i^{n_i} / n_i!},$$

where the products run over all bins in the $m_{\text{ES}}-\Delta E$ distributions for the four modes and four runs, μ_i is the integral of the likelihood function over bin i , and n_i is the observed number of candidates in bin i . Using 100 bins in m_{ES} and 30 bins in ΔE , we find for our fit $-\log \lambda = 18284$.

For sufficiently large values of n_i , this statistic is χ^2 distributed. Our yields, however, are not

large enough to assure this. We therefore estimate the distribution of $-\log \lambda$ using a toy MC. We generate 5000 toy MC samples distributed according to the fitted likelihood function, in which the number of samples per mode and run is identical to the yield in the data sample. The distribution of $-\log \lambda$ over these toy MC samples is shown below. Based on this distribution, we estimate the fit probability to be $P = 10\%$.

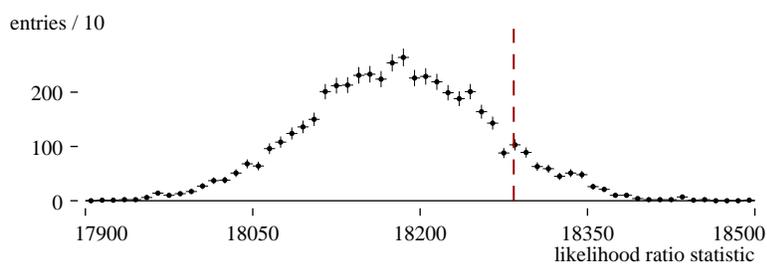


FIGURE 67. Distribution of the likelihood ratio statistic $-\log \lambda$ in 5000 toy MC samples generated according to the likelihood function resulting from the fit. Each toy sample contains the same number of events as the data sample. The value of the statistic obtained for the data fit is indicated by the dotted red line.

Distributions in m_{ES} and ΔE of the data sample and fit projections are shown below.

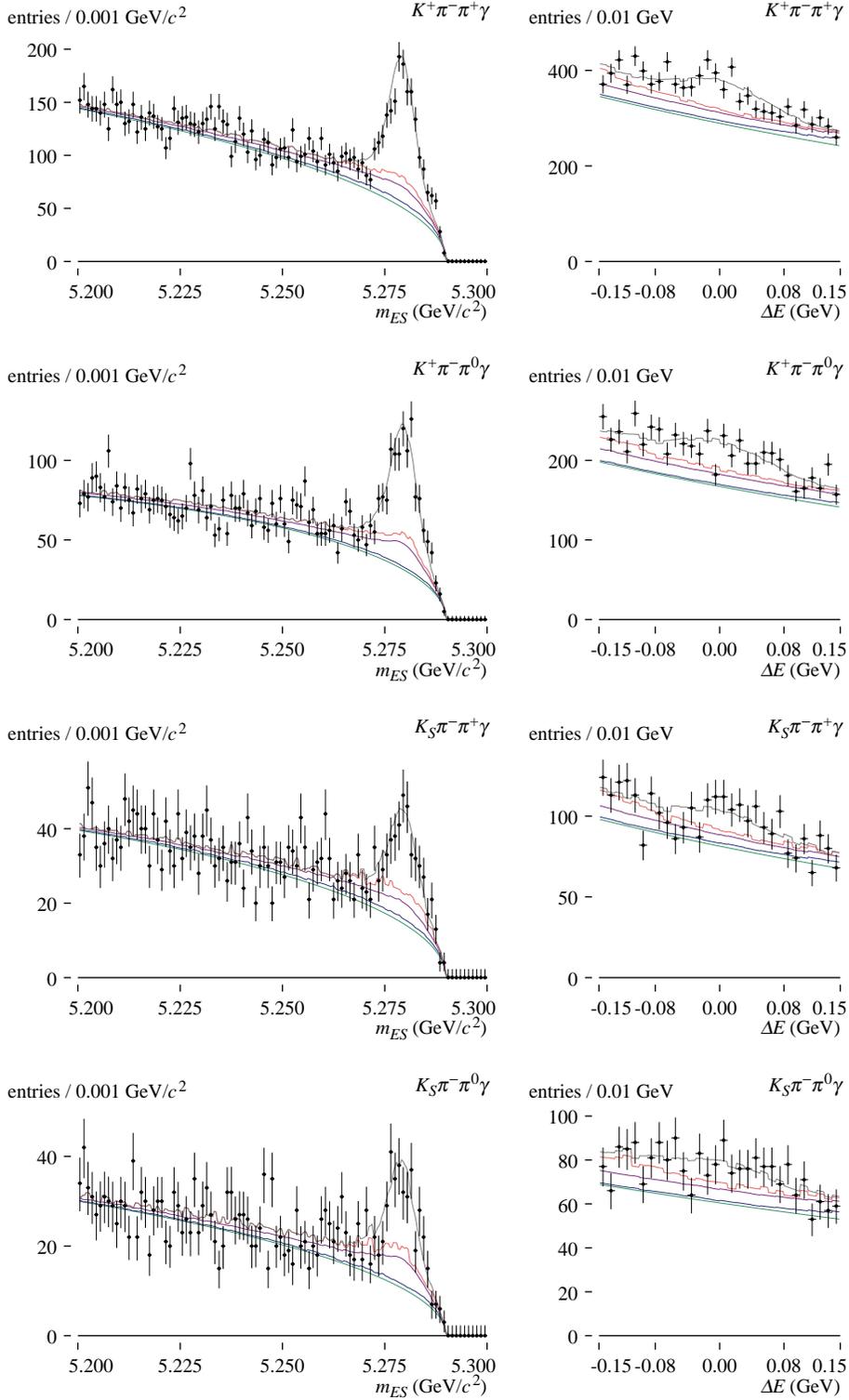


FIGURE 68. Distributions of m_{ES} and ΔE of the on-resonance data sample after full selection (points). Projections of components of the fitted likelihood function are shown as curves; bottom to top: $udsc$ and generic $B\bar{B}$ (green), feed-up (blue), crossfeed (purple), feed-down (red) and total including signal (gray).

The table below shows the normalizations of the components of the fit. These numbers estimate the yields for signal and each background category in the data sample.

Fit component	$K^+\pi^-\pi^+\gamma$	$K^+\pi^-\pi^0\gamma$	$K_S^0\pi^-\pi^+\gamma$	$K_S^0\pi^+\pi^0\gamma$
signal	899.0 ± 38.3	571.9 ± 30.9	175.7 ± 19.5	164.4 ± 15.1
crossfeed category 1	233.6 ± 10.0	100.3 ± 4.3	7.7 ± 0.3	0.9 ± 0.0
crossfeed category 2	245.5 ± 13.3	254.3 ± 13.7	3.6 ± 0.2	5.3 ± 0.3
crossfeed category 3	16.5 ± 1.8	4.0 ± 0.4	54.0 ± 6.0	32.9 ± 3.7
crossfeed category 4	4.1 ± 0.4	28.8 ± 2.6	88.7 ± 8.2	117.5 ± 10.8
feed-up	213.7 ± 0.0	83.0 ± 0.0	64.5 ± 0.0	39.2 ± 0.0
feed-down	308.7 ± 0.0	218.9 ± 0.0	115.2 ± 0.0	107.9 ± 0.0
$udsc$ and $B\bar{B}$	8727.4 ± 190.9	5044.0 ± 147.3	2456.0 ± 102.4	1818.6 ± 87.7
total	10648.5 ± 195.4	6305.2 ± 151.2	2965.3 ± 104.7	2286.6 ± 89.7

TABLE 45. Yields of components of the fit to on-resonance data computed from fitted parameters. The feed-down and feed-up components are fixed in the fit to values determined from MC. Uncertainties are calculated from errors obtained from the fit.

9.2 $m_{K\pi\pi}$ distributions

The figure below shows the $m_{K\pi\pi}$ sPlot distributions in the on-resonance data sample. The content of each bin has been converted to a branching fraction by scaling the fraction of the total yield to the total measured branching fraction in that mode.

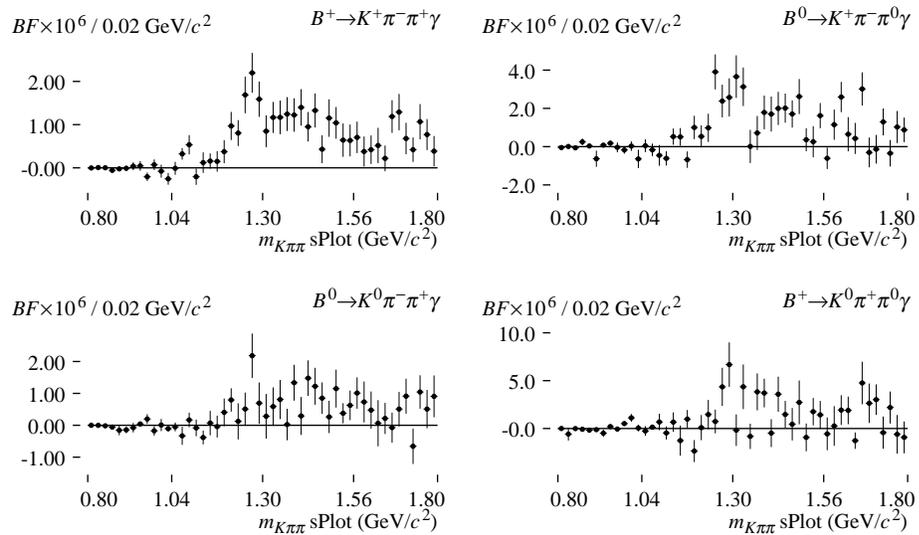


FIGURE 69. Distributions of $m_{K\pi\pi}$ in the on-resonance data sample. Error bars show statistical uncertainty from the sPlot only.

10 Systematic Uncertainties

This chapter describes our estimates of systematic uncertainties on the branching fraction and mass spectrum measurements.

10.1 Branching fraction

In this section, we list the systematic uncertainties that we estimate for our measurements of the $B \rightarrow K\pi\pi\gamma$ branching fractions.

10.1.1 $B\bar{B}$ COUNT UNCERTAINTY

The official $B\bar{B}$ count for the Run 1-4 dataset has an uncertainty of 0.6%. We use this value as the systematic uncertainty on $N_{B\bar{B}}$.

10.1.2 INPUT BRANCHING FRACTION UNCERTAINTY

In our fit, we assume $\mathcal{B}(\Upsilon(4S) \rightarrow B^+B^-) = \mathcal{B}(\Upsilon(4S) \rightarrow B^0\bar{B}^0) = 0.5$. A recent *BABAR* measurement [29] has found $\mathcal{B}(\Upsilon(4S) \rightarrow B^0\bar{B}^0) = 0.487 \pm 0.010$ (stat.) ± 0.008 (sys.). Since the measured value is statistically compatible with one half, we combine the statistical and systematic error on this measurement and convert it to a relative error to obtain a 2.4% uncertainty.

The branching fraction $\mathcal{B}(K^0 \rightarrow K_s^0 \rightarrow \pi^+\pi^-)$ is well measured. We assign no systematic uncertainty due to this.

10.1.3 TRACK RECONSTRUCTION EFFICIENCY

We have applied standard tracking efficiency corrections to MC samples, determined by examining tracks reconstructed by the SVT. These corrections are subject to a 0.8% systematic uncertainty per K^\pm and π^\pm candidate in our reconstructed final states (not including π^\pm used to construct K_s^0

candidates).

10.1.4 PHOTON RECONSTRUCTION EFFICIENCY

The efficiency for π^0 reconstruction in the EMC in MC samples has been studied using the ratio of $\tau \rightarrow \rho\nu$ to $\tau \rightarrow \pi\nu$ yields. This study also validates the single photon reconstruction efficiency in MC. Based on the results of this study, we assign a 1.8% systematic uncertainty to the reconstruction efficiency of the high-energy photon in our reconstructed B candidates.

10.1.5 SKIM

The $BtoXGamma$ skim selection criteria are very close to 100% efficient, so we assume there is no significant systematic uncertainty.

10.1.6 PHOTON SELECTION

Systematic uncertainties in the efficiencies of photon selection cuts have been studied as part of the $B \rightarrow K^*\gamma$ analysis [8] by embedding photon clusters in hadronic events and studying the distributions of selection variables. Our photon selection cuts are identical to those used in this analysis, except that we use a slightly more restrictive π^0 veto: a wider veto window, $25 \text{ MeV}/c^2$ around the π^0 mass. We adopt the systematic uncertainties determined in this analysis for our photon selection criteria: a 2% uncertainty on the efficiency of the bump distance cut, and a 1% efficiency uncertainty due to the π^0 and η vetoes.

10.1.7 K^\pm AND π^\pm PID

We estimate the systematic uncertainties on the efficiency due to the hadronic quality cuts, Fisher cuts, and vertex probability cuts from the efficiency differences between $D\pi^\pm$ control samples and $D\pi^\pm$ MC, listed in table 31. We take the average discrepancy between the π^\pm efficiency in the $D\pi^\pm$ control sample and the efficiency in $D\pi^\pm$ MC as the systematic uncertainty on that cut. Likewise,

we take the average discrepancy between K^\pm selection efficiencies as the systematic uncertainty.

10.1.8 K_s^0 SELECTION

We have applied standard K_s^0 efficiency corrections to MC samples, which are obtained from a study of $B^0 \rightarrow \phi K_s^0$ and $B^0 \rightarrow \pi^+ D^-, D^- \rightarrow K_s^0 \pi^-$ decays. The systematic uncertainty to the corrections determined in this study, which are parameterized by transverse momentum and polar angle, translates to 1.3% in the $K_s^0 \pi^- \pi^+ \gamma$ mode and 1.6% in the $K_s^0 \pi^+ \pi^0 \gamma$ mode. We assign the larger value, 1.6%, as the systematic uncertainty for our K_s^0 selection efficiency in both modes. The efficiencies of the K_s^0 cuts measured in our $D\pi^\pm$ control sample are in agreement with efficiencies in $D\pi^\pm$ MC within this uncertainty.

10.1.9 π^0 SELECTION

The efficiency for π^0 reconstruction using the EMC in MC samples has been studied using the ratio of $\tau \rightarrow \rho \nu$ to $\tau \rightarrow \pi \nu$ yields. We have applied standard π^0 efficiency corrections to MC samples based on this study, for which the corresponding systematic uncertainty is 3%. The efficiencies of the π^0 mass and photon energy cuts measured in our $D\pi^\pm$ control sample are in agreement with efficiencies in $D\pi^\pm$ MC within this uncertainty.

10.1.10 FISHER DISCRIMINANT CUT

The efficiencies of the Fisher discriminant cuts agree well between the $D\pi^\pm$ control samples and $D\pi^\pm$ MC. The absolute value of the difference between them averaged over all four modes is 1.0%, which we assign as the systematic uncertainty on the Fisher cut efficiency.

10.1.11 VERTEX PROBABILITY CUT

For the vertex probability cut, we also see good agreement between the efficiencies in the $D\pi^\pm$ control sample data and in the corresponding MC. As with the Fisher discriminant cut, we assign

the average of the absolute value of the difference, 0.7%, as the systematic uncertainty.

10.2 Fit bias

We assign systematic uncertainties due to fit bias based on the toy MC study presented in section 8.2.

In this study, we generated toy MC samples by choosing signal and background events from MC samples. The table below shows the fitted branching fractions averaged over 500 toy MC samples, and the difference between the average fitted values and the input branching fractions. The signal and background components of the toy MC samples are sampled from the same MC samples with which we model the signal and background components of our fit.

Mode	Average BF $\times 10^6$	BF bias $\times 10^6$
$K^+\pi^-\pi^+\gamma$	23.78 ± 1.25	0.39 ± 0.06
$K^+\pi^-\pi^0\gamma$	29.88 ± 2.09	-0.38 ± 0.09
$K_S^0\pi^-\pi^+\gamma$	23.69 ± 2.10	0.30 ± 0.09
$K_S^0\pi^+\pi^0\gamma$	31.36 ± 4.01	1.10 ± 0.18

TABLE 46. Ensemble average and RMS of fit parameters over 500 fits to independent toy MC samples, and difference from the input signal branching fractions. The toy MC is generated according to our fit model.

In each mode, we assign the bias of the fitted branching fraction divided by the input branching fraction as the systematic uncertainty due to fit bias.

For the $m_{K\pi\pi}$ measurement, we take the difference in each bin between the average reconstructed $m_{K\pi\pi}$ distribution and the generator-level $m_{K\pi\pi}$ distribution in our signal MC sample as the systematic uncertainty. The figures below illustrate this systematic uncertainty; the distributions show the generator-level $m_{K\pi\pi}$ distributions, and the error bars show the bin-by-bin systematic uncertainty due to bias computed in this way.

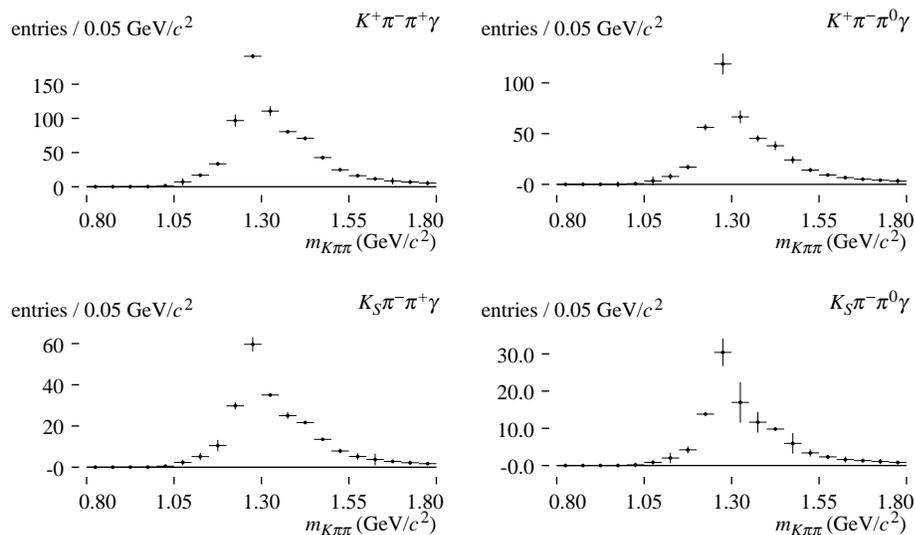


FIGURE 70. Generator-level signal MC distributions of $m_{K\pi\pi}$, with systematic uncertainty due to fit bias indicated by error bars.

10.3 Signal and crossfeed shape systematics

To assess the uncertainty of the measured branching fractions due to uncertainties in the signal and crossfeed shape parameters, we refit the data sample, varying the shape parameters. We perform 250 such toy fits, varying all the shape parameters simultaneously according to normal distributions with the mean and RMS values determined from the shape fits (section 8.1). The following table shows the mean branching fractions fitted in these toy fits, and the difference between these means and the branching fractions obtained in the main fit.

Mode	Average BF $\times 10^6$	δ BF $\times 10^6$
$K^+\pi^-\pi^+\gamma$	29.52 ± 0.20	0.03
$K^+\pi^-\pi^0\gamma$	41.92 ± 1.39	1.19
$K_S^0\pi^-\pi^+\gamma$	18.66 ± 0.23	0.16
$K_S^0\pi^+\pi^0\gamma$	45.69 ± 0.71	0.09

TABLE 47. Mean and RMS values of branching fractions extracted from 250 toy fits to the on-resonance data sample with signal and crossfeed shape parameters varied according to their uncertainties. The right column shows the difference between the mean value and the value from the actual data fit.

We assign a systematic uncertainty in each mode due to signal and crossfeed shape uncertainty equal to the relative difference between the mean value from these toy fits and the value obtained from the on-resonance data fit.

The Crystal Ball mean parameter μ of the m_{ES} shape used for signal and crossfeed distributions is not determined from MC; instead, we fix it to m_B in each mode. We have measured the shift of this parameter in the $D\pi^\pm$ control sample (section 7.3). The shift averaged over all runs and modes is $+0.65 \text{ MeV}/c^2$. To assess the systematic uncertainty due to the shift in m_{ES} , we repeat the fit to the on-resonance data sample with the μ parameter for signal and crossfeed distributions shifted by $0.65 \text{ MeV}/c^2$. The table below shows the results of this fit.

Mode	Fitted BF $\times 10^6$	δ BF $\times 10^6$
$K^+\pi^-\pi^+\gamma$	29.20 ± 612644.44	-0.29
$K^+\pi^-\pi^0\gamma$	40.55 ± 612236.15	-0.19
$K_S^0\pi^-\pi^+\gamma$	18.79 ± 613096.27	0.29
$K_S^0\pi^+\pi^0\gamma$	45.33 ± 612081.59	-0.27

TABLE 48. Branching fractions extracted from fit to on-resonance data sample with Crystal Ball parameters μ for signal and crossfeed shape distributions shifted by $0.65 \text{ MeV}/c^2$. The right column shows the difference between the fitted value and the value from the actual data fit.

We assign a systematic uncertainty in each mode due to the m_{ES} shift equal to the relative difference between the branching fraction extracted from this fit and the branching fraction from the on-resonance data fit.

10.4 $b \rightarrow s\gamma$ model uncertainty

Our fit includes components to account for backgrounds from $b \rightarrow s\gamma$ processes. We divide these into three categories: feed-up from $K\pi\gamma$, crossfeed from misreconstructed $K\pi\pi\gamma$, and feed-down from higher-multiplicity $b \rightarrow s\gamma$ processes. We model the $K\pi\gamma$ background with exclusive $B \rightarrow K^*(892)\gamma$ and $B \rightarrow K_2^*(1430)\gamma$ MC samples. We model the $K\pi\pi\gamma$ background with the same

cocktail of inclusive MC samples that we use as our signal model. We model the remaining $b \rightarrow s\gamma$ processes with inclusive MC.

There is considerable uncertainty in the accuracy of the MC samples we use for these backgrounds. In this section, we assess the impact on our final results of variations in our procedure for modeling $b \rightarrow s\gamma$ backgrounds.

10.4.1 $K\pi\pi\gamma$ MODEL

For $B \rightarrow K\pi\pi\gamma$ processes, we use our signal MC sample to study crossfeed backgrounds by imposing a veto on MC truth, which selects crossfeed candidates only. Our cocktail of exclusive signal MC modes is produced with very different physics than the $K\pi\pi\gamma$ component of the inclusive $b \rightarrow s\gamma$ MC. The former is composed entirely of explicitly-specified primary and secondary resonances. The latter is produced by a fragmentation model, which does not include the full complement of kaon resonances, and also produces events in non-resonant decays.

The figures below compare the generator-level distributions of $m_{K\pi\pi}$ in our signal MC to those from the $K\pi\pi\gamma$ component of the generic $b \rightarrow s\gamma$ MC.

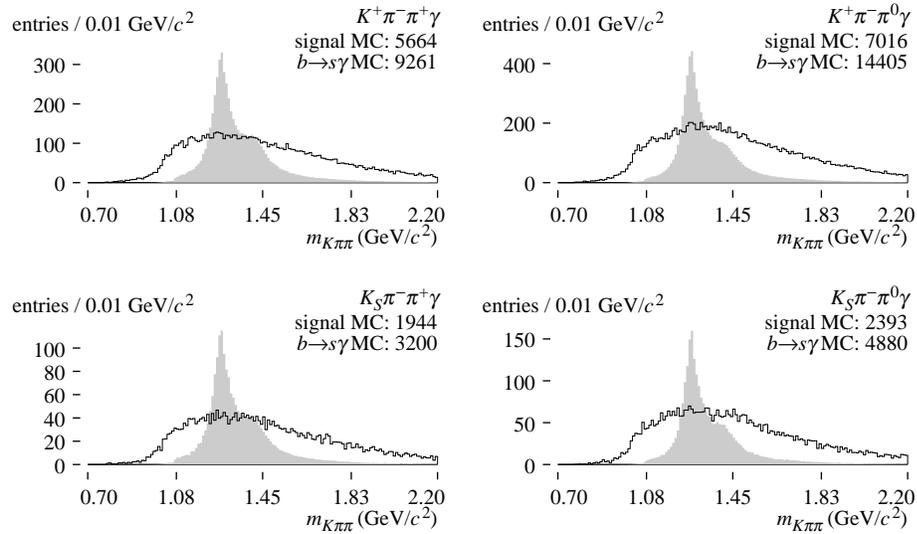


FIGURE 71. Generator-level distributions of $m_{K\pi\pi}$ of events in the signal MC cocktail (shaded) and $K\pi\pi\gamma$ events in the KN480 inclusive $b \rightarrow s\gamma$ MC (curves). Both samples are weighted to the analysis luminosity, and fragmentation corrections have been applied to the latter. Integrals, which include overflow bins (not shown), are noted on the plots.

The following tables show the predicted yield of crossfeed background predicted by the exclusive signal MC cocktail and by the inclusive $b \rightarrow s\gamma$ MC. The first shows the yields of the two samples normalized independently to the integrated luminosity of the data sample. The second shows the same results with the $B \rightarrow K\pi\pi\gamma$ yield in each mode in the inclusive $b \rightarrow s\gamma$ MC normalized to that of the exclusive signal MC cocktail.

Generation mode	$K^+\pi^-\pi^+\gamma$		$K^+\pi^-\pi^0\gamma$		$K_S^0\pi^-\pi^+\gamma$		$K_S^0\pi^+\pi^0\gamma$	
$K^+\pi^-\pi^+\gamma$	180	287	44	68	6	9	0	0
$K^+\pi^-\pi^0\gamma$	178	375	183	301	3	4	4	8
$K_S^0\pi^-\pi^+\gamma$	6	11	0	1	61	104	24	28
$K_S^0\pi^+\pi^0\gamma$	1	3	5	9	58	106	72	124

TABLE 49. Comparison of crossfeed background yields estimated from signal MC and from $K\pi\pi\gamma$ events in inclusive $b \rightarrow s\gamma$ MC. For each generation and reconstruction mode, the table shows the crossfeed yield estimated from signal MC (left) and from the $K\pi\pi\gamma$ component of inclusive $b \rightarrow s\gamma$ MC. Analysis cuts, MCS, and MC truth veto have been applied. Yields are normalized to the analysis luminosity.

Generation mode	$K^+\pi^-\pi^+\gamma$		$K^+\pi^-\pi^0\gamma$		$K_S^0\pi^-\pi^+\gamma$		$K_S^0\pi^+\pi^0\gamma$	
$K^+\pi^-\pi^+\gamma$	180	207	44	49	6	7	0	0
$K^+\pi^-\pi^0\gamma$	178	216	183	173	3	2	4	5
$K_S^0\pi^-\pi^+\gamma$	6	8	0	1	61	75	24	20
$K_S^0\pi^+\pi^0\gamma$	1	2	5	5	58	61	72	71

TABLE 50. Identical to table 49, except that the $b \rightarrow s\gamma$ MC sample has been renormalized to the same number of events in each $K\pi\pi\gamma$ mode with $m_{K\pi\pi} < 1.8 \text{ GeV}/c^2$ at generator level as the exclusive signal MC.

The $K\pi\pi\gamma$ component of the inclusive $b \rightarrow s\gamma$ predicts a substantially larger rate of crossfeed background. Our fit procedure, however, is not sensitive to the overall normalization of $K\pi\pi\gamma$ production in our model. In our fit function, crossfeed background components are scaled by the same floating branching fraction parameters as the signal components. In effect, we measure the $K\pi\pi\gamma$ branching fractions simultaneously in the signal and crossfeed background components. However, model uncertainty in the $m_{K\pi\pi}$ shape can affect the rate of crossfeed background because of the cut we place on this variable. This uncertainty is small; the plots on the following pages show that there is little contribution of crossfeed candidates from $K\pi\pi\gamma$ processes with generated $m_{K\pi\pi} > 1.8 \text{ GeV}/c^2$. We estimate a systematic uncertainty to account for this effect.

The following plots show the m_{ES} , ΔE , and reconstructed $m_{K\pi\pi}$ distributions for the $K\pi\pi\gamma$ crossfeed background, as predicted by the exclusive signal MC cocktail and by inclusive $b \rightarrow s\gamma$ MC. For the latter, the estimates from both from the Kagan and Neubert MC model with $m_B =$

4.65 GeV/c² (KN465) and the from the model with $m_B = 4.80$ GeV/c² (KN480) are shown, with the fragmentation corrections measured by the semi-inclusive $b \rightarrow s\gamma$ analysis applied to both.

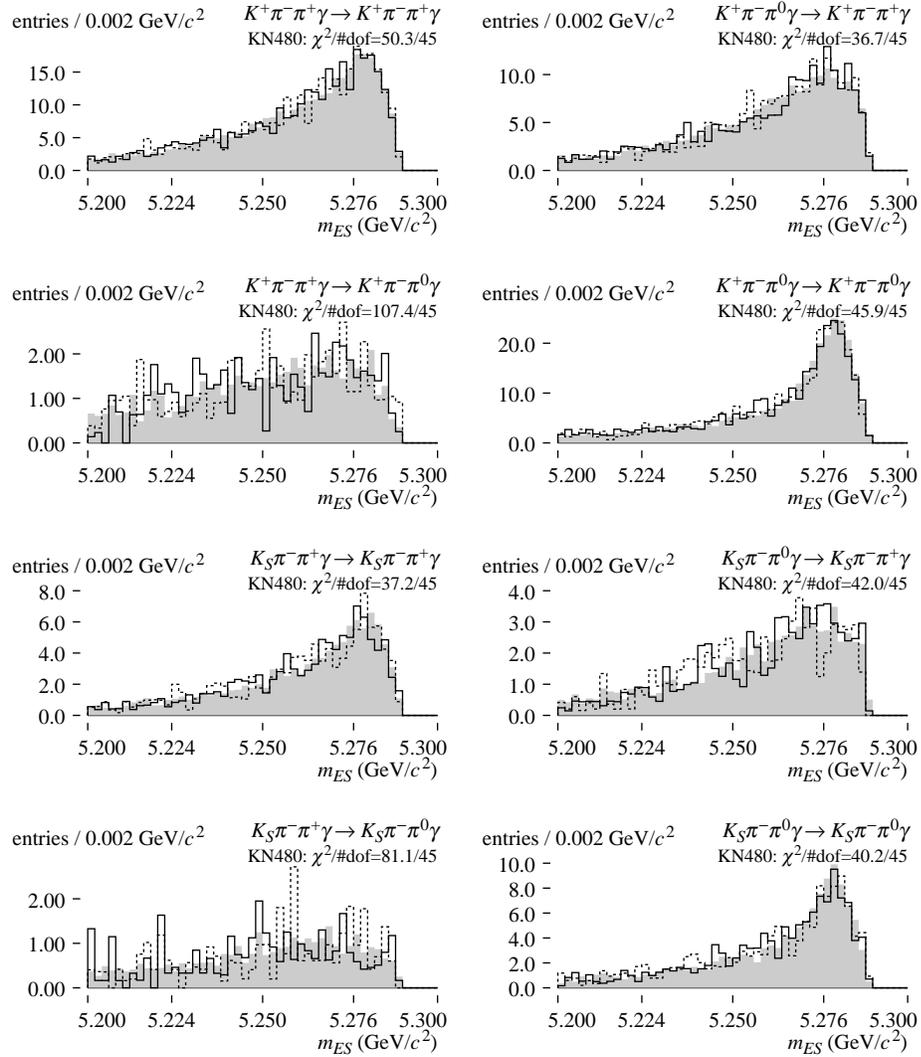


FIGURE 72. Distributions of m_{ES} in crossfeed background estimated from exclusive signal MC (shaded), KN480 $b \rightarrow s\gamma$ MC (solid), and KN465 $b \rightarrow s\gamma$ MC (dotted). Generated and reconstructed modes are indicated in the upper-right of each plot. Crossfeed between K^\pm and K_S^0 modes are small and not shown. The $b \rightarrow s\gamma$ MC samples have been normalized to the exclusive signal MC sample. Candidate selection cuts, the fit region cut, and MC truth veto have been applied. The binned chi-square comparing the KN480 $b \rightarrow s\gamma$ MC and exclusive signal MC distributions are indicated on each plot.

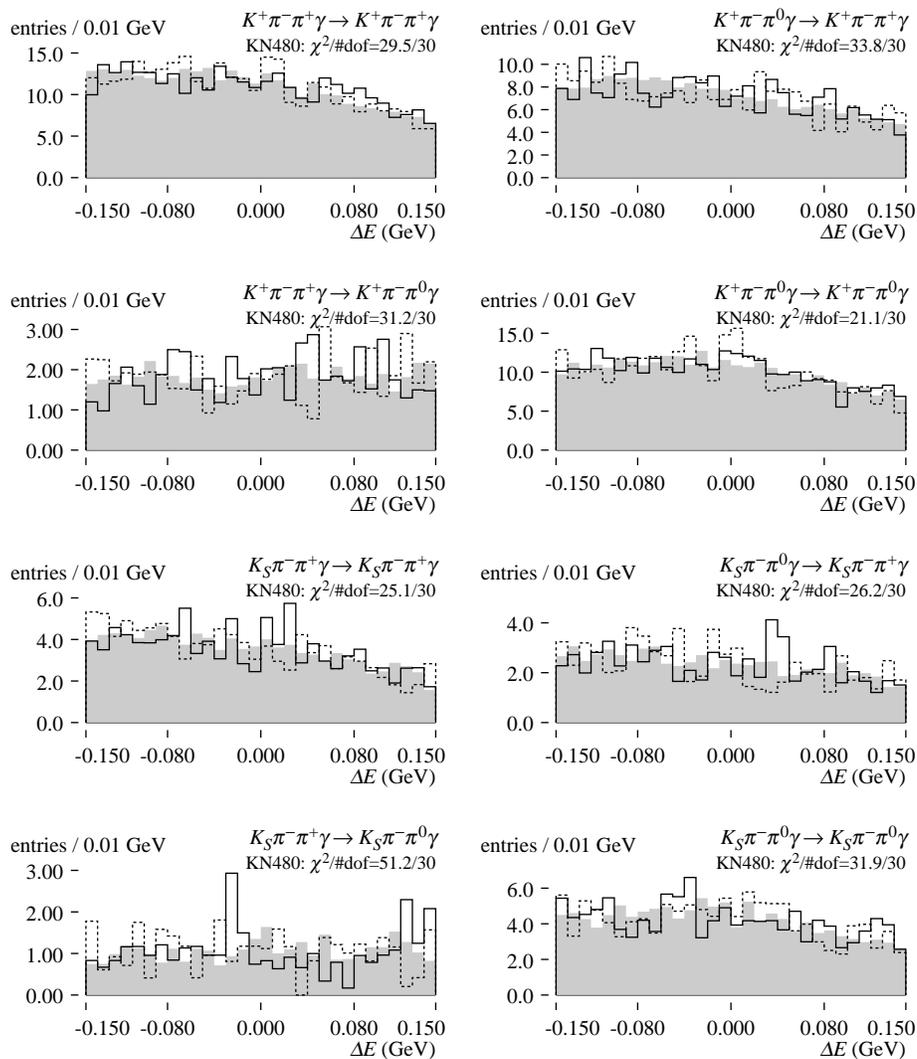


FIGURE 73. Distributions of ΔE in crossfeed background estimated from exclusive signal MC (shaded), KN480 $b \rightarrow s\gamma$ MC (solid), and KN465 $b \rightarrow s\gamma$ MC (dotted). Generated and reconstructed modes are indicated in the upper-right of each plot. Crossfeed between K^\pm and K_s^0 modes are small and not shown. The $b \rightarrow s\gamma$ MC samples have been normalized to the exclusive signal MC sample. Candidate selection cuts, the fit region cut, and MC truth veto have been applied. The binned chi-square comparing the KN480 $b \rightarrow s\gamma$ MC and exclusive signal MC distributions are indicated on each plot.

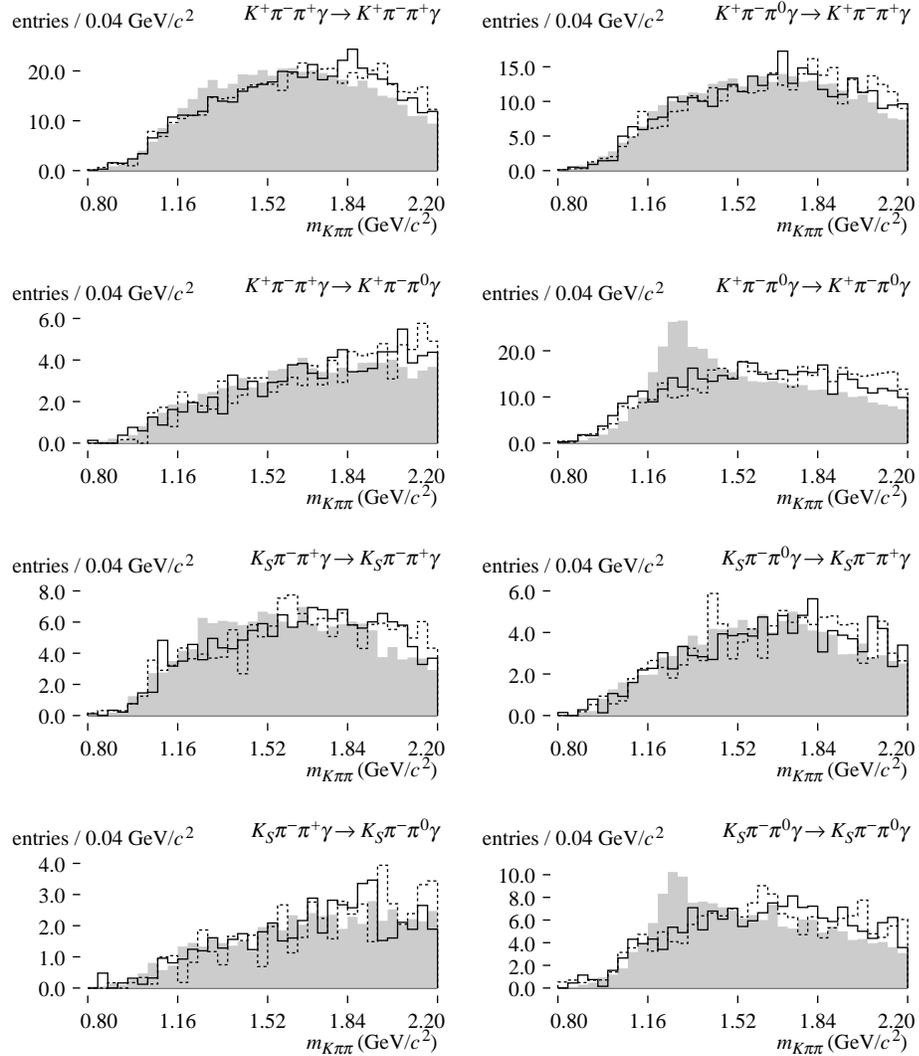


FIGURE 74. Distributions of $m_{K\pi\pi}$ in crossfeed background estimated from exclusive signal MC (shaded), KN480 $b \rightarrow s\gamma$ MC (solid), and KN465 $b \rightarrow s\gamma$ MC (dotted). Generated and reconstructed modes are indicated in the upper-right of each plot. Crossfeed between K^\pm and K_S^0 modes are small and not shown. The $b \rightarrow s\gamma$ MC samples have been normalized to the exclusive signal MC sample. Candidate selection cuts, the fit region cut, except for the $m_{K\pi\pi}$ cut, and MC truth veto have been applied.

The $m_{K\pi\pi}$ self-crossfeed distributions for modes containing a π^0 estimated from the signal MC cocktail show a peak near $1270 \text{ GeV}/c^2$. Based on manual examination of MC truth listings, we attribute the peak to $B \rightarrow K_1(1270)\gamma$ candidates which are reconstructed correctly except for one of the photons from the decay of the π^0 . In some events, the mistaken photon is chosen from elsewhere

in the event. In other events, the photon from the π^0 candidate converts in the detector, and the reconstructed photon is truth-matched to a conversion electron. In both cases, the reconstructed kinematic quantities can be close to the correct values for a correctly reconstructed candidate, but MC truth matching indicates the candidate is misreconstructed, so it is classified as self-crossfeed.

The following plots show the correlation between generated and reconstructed $m_{K\pi\pi}$ for cross-feed backgrounds, in the exclusive signal MC cocktail and in the KN480 $b \rightarrow s\gamma$ MC. These plots indicate that both the inclusive and exclusive models predict little crossfeed from $K\pi\pi\gamma$ events with true $m_{K\pi\pi} > 1.8 \text{ GeV}/c^2$ in our sample. We conclude that the model dependence of the $m_{K\pi\pi} < 1.8 \text{ GeV}/c^2$ selection does not produce a large systematic uncertainty.

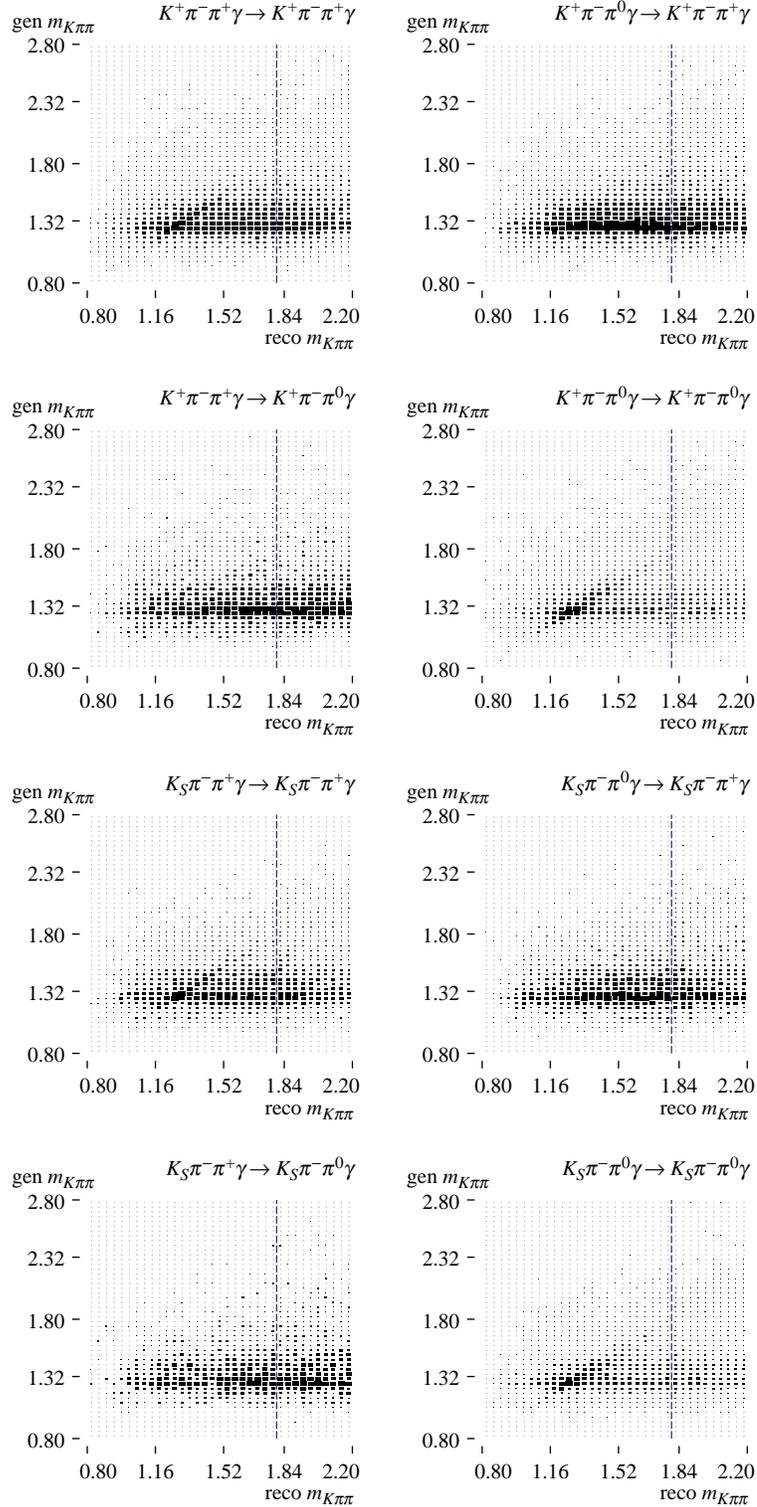


FIGURE 75. Distributions of generated $m_{K\pi\pi}$ vs. reconstructed $m_{K\pi\pi}$ in crossfeed background estimated from exclusive signal MC. Generated and reconstructed modes are indicated in the upper-right of each plot. Crossfeed between K^\pm and K_S^0 modes are small and not shown. Candidate selection cuts the fit region cut, except for the $m_{K\pi\pi}$ cut (indicated in blue), and MC truth veto have been applied.

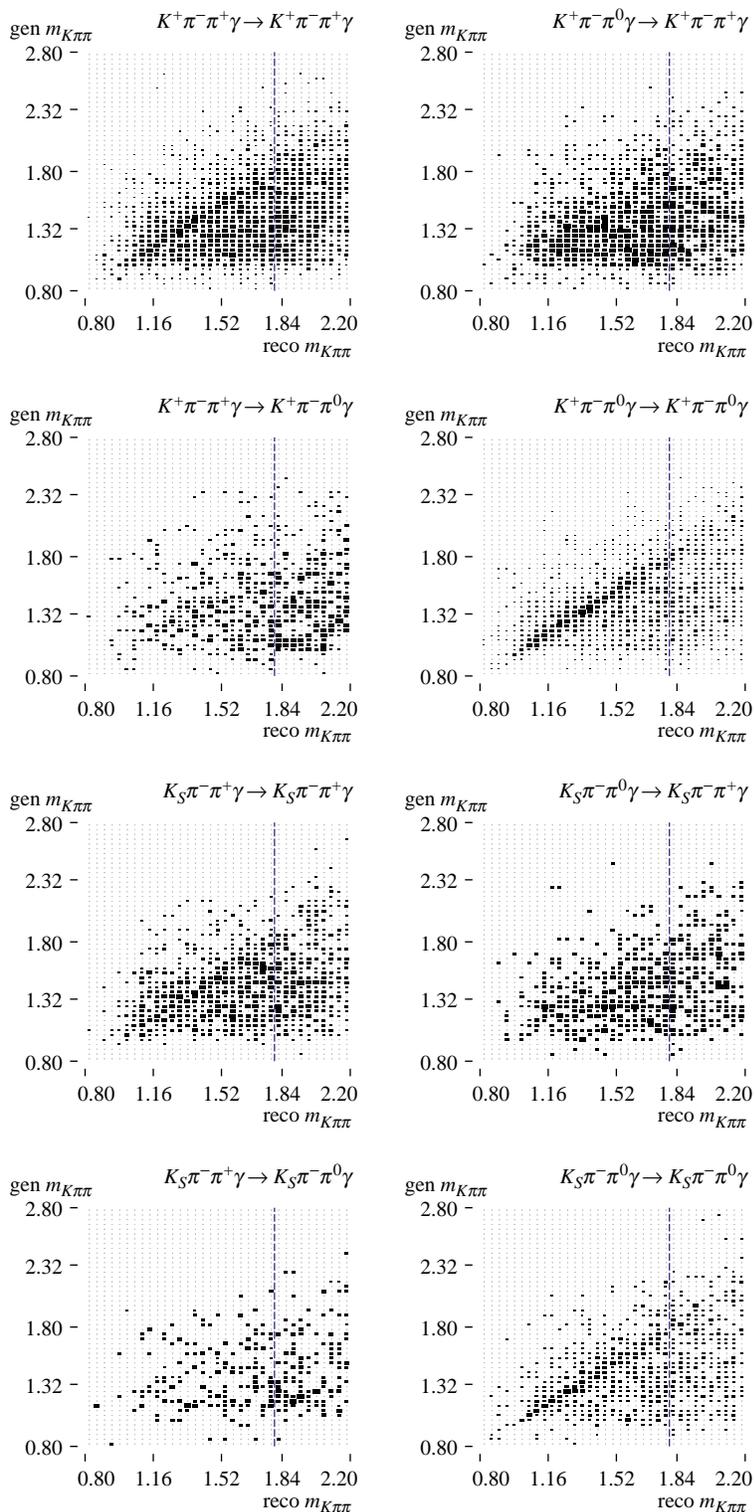


FIGURE 76. Distributions of generated $m_{K\pi\pi}$ vs. reconstructed $m_{K\pi\pi}$ in crossfeed background estimated from $b \rightarrow s\gamma$ MC. Generated and reconstructed modes are indicated in the upper-right of each plot. Crossfeed between K^\pm and K_S^0 modes are small and not shown. Candidate selection cuts and the fit region cut, except for the $m_{K\pi\pi}$ cut (indicated in blue), and MC truth vetox have been applied.

10.4.2 FEED-DOWN MODEL

We use the KN480 inclusive $b \rightarrow s\gamma$ MC sample, with $K\pi\gamma$ and $K\pi\pi\gamma$ events removed, to estimate feed-down backgrounds. The following plots compare this sample with the KN465 inclusive MC sample.

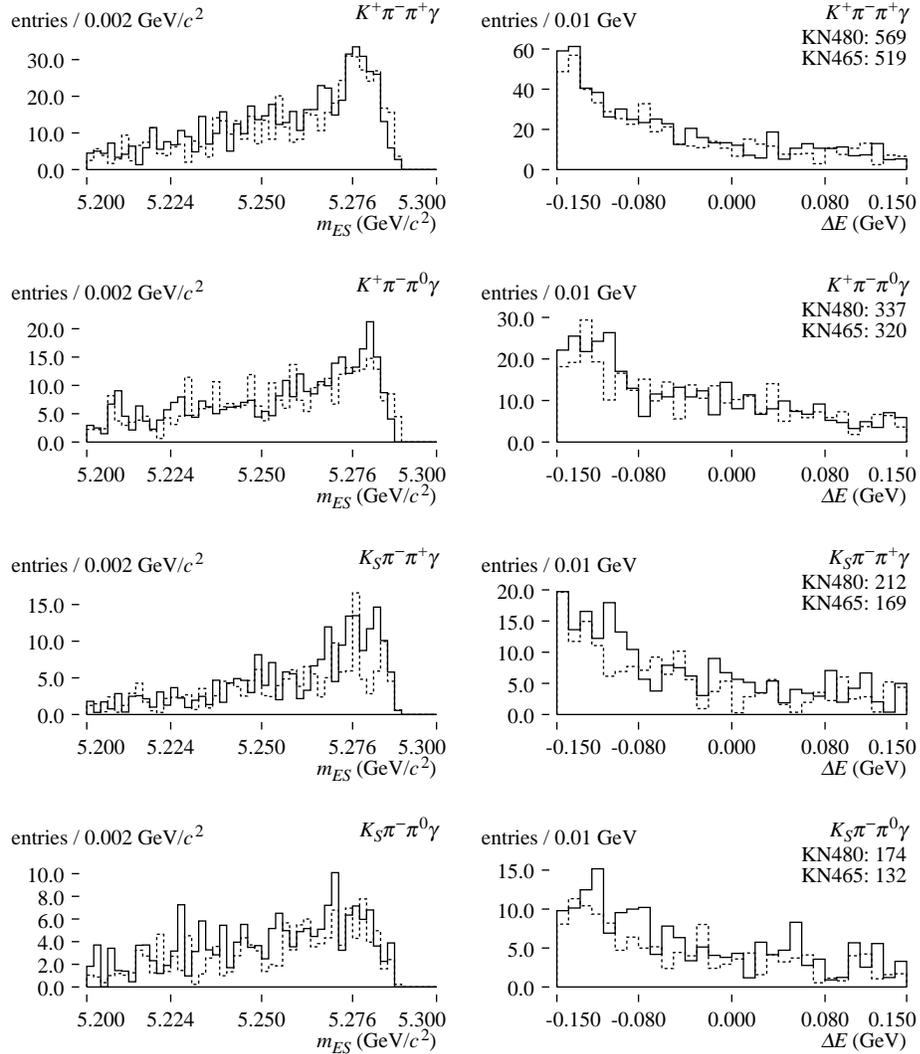


FIGURE 77. Distributions of m_{ES} and ΔE in feed-down background candidates from KN480 inclusive $b \rightarrow s\gamma$ MC (solid) and KN465 inclusive $b \rightarrow s\gamma$ MC (dashed), scaled to analysis luminosity. Candidate selection cuts, the fit region cut, and MCS have been applied. Total yields are noted on the ΔE plots.

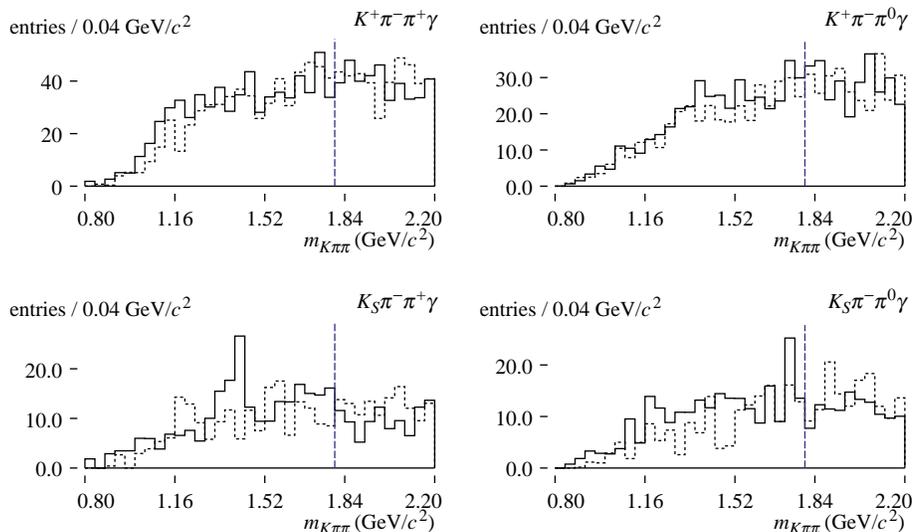


FIGURE 78. Distributions of $m_{K\pi\pi}$ in feed-down background candidates from KN480 inclusive $b \rightarrow s\gamma$ MC (solid) and KN465 inclusive $b \rightarrow s\gamma$ MC (dashed), scaled to analysis luminosity. Candidate selection cuts, except for the $m_{K\pi\pi}$ cut (indicated in blue), have been applied.

10.4.3 MODEL VARIATION

We have performed toy MC studies to assess the impact of varying the $b \rightarrow s\gamma$ background model on our fit parameters. We start with the toy MC study used in section 10.2. We then vary the MC samples from which we sample background processes in the toy MC samples, but keep the fit procedure and fit components the same. The change in the means of the fitted branching fractions indicate how sensitive the fit procedure is to the choice of $b \rightarrow s\gamma$ background models.

The table below shows the result of repeating the original MC study, with the feed-up $K\pi\gamma$ background component for generated $m_{K\pi} > 1.1 \text{ GeV}/c^2$ sampled from the $K\pi\gamma$ component of inclusive $b \rightarrow s\gamma$ MC (instead of from exclusive $B \rightarrow K_2^*(1430)\gamma$). The components of the fit function are unchanged.

Mode	Average BF $\times 10^6$	δ BF $\times 10^6$
$K^+\pi^-\pi^+\gamma$	23.78 ± 1.25	-0.01 ± 0.08
$K^+\pi^-\pi^0\gamma$	29.97 ± 2.08	0.10 ± 0.13
$K_S^0\pi^-\pi^+\gamma$	23.82 ± 2.23	0.13 ± 0.14
$K_S^0\pi^+\pi^0\gamma$	31.56 ± 4.27	0.20 ± 0.26

TABLE 51. Fitted mean and RMS branching fractions for 500 toy MC studies with $K\pi\gamma$ feed-up background for $m_{K\pi} > 1.1 \text{ GeV}/c^2$ sampled from inclusive $b \rightarrow s\gamma$ MC. The last column shows the difference from the branching fractions in table 46.

The table below shows the result of repeating the original MC study, with the signal component sampled from inclusive $b \rightarrow s\gamma$ MC (instead of from the exclusive signal MC cocktail). The $b \rightarrow s\gamma$ MC is normalized in each mode to yield the same number of candidates as the signal MC cocktail. The crossfeed background component is still sampled from the signal MC cocktail. The components of the fit function are unchanged.

Mode	Average BF $\times 10^6$	δ BF $\times 10^6$
$K^+\pi^-\pi^+\gamma$	23.80 ± 1.19	0.02 ± 0.08
$K^+\pi^-\pi^0\gamma$	29.73 ± 1.97	-0.14 ± 0.13
$K_S^0\pi^-\pi^+\gamma$	23.41 ± 2.13	-0.28 ± 0.13
$K_S^0\pi^+\pi^0\gamma$	31.56 ± 3.93	0.20 ± 0.25

TABLE 52. Fitted mean and RMS branching fractions for 500 toy MC studies with the signal component sampled from normalized inclusive $b \rightarrow s\gamma$ MC. The last column shows the difference from the branching fractions in table 46.

The table below shows the result of repeating the original MC study, with the crossfeed $K\pi\pi\gamma$ background component sampled from inclusive $b \rightarrow s\gamma$ MC (instead of from the exclusive signal MC cocktail). The $K\pi\pi\gamma$ component from inclusive $b \rightarrow s\gamma$ MC is normalized to the generator-level $K\pi\pi\gamma$ yield, without cuts, of exclusive signal MC cocktail. The signal component is still sampled from the signal MC cocktail. The components of the fit function are unchanged.

Mode	Average BF $\times 10^6$	δ BF $\times 10^6$
$K^+\pi^-\pi^+\gamma$	24.07 ± 1.17	0.29 ± 0.08
$K^+\pi^-\pi^0\gamma$	29.94 ± 2.05	0.06 ± 0.13
$K_S^0\pi^-\pi^+\gamma$	23.92 ± 2.11	0.23 ± 0.13
$K_S^0\pi^+\pi^0\gamma$	31.28 ± 4.03	-0.08 ± 0.26

TABLE 53. Fitted mean and RMS branching fractions for 500 toy MC studies with $K\pi\pi\gamma$ crossfeed background sampled from renormalized inclusive $b \rightarrow s\gamma$ MC. The last column shows the difference from the branching fractions in table 46.

The table below shows the result of repeating the original MC study, with the feed-down $b \rightarrow s\gamma$ background component sampled from KN465 inclusive $b \rightarrow s\gamma$ MC (instead of from KN480 $b \rightarrow s\gamma$ MC). The fragmentation corrections measured in the same-inclusive $b \rightarrow s\gamma$ analysis are applied to both. The components of the fit function are unchanged.

Mode	Average BF $\times 10^6$	δ BF $\times 10^6$
$K^+\pi^-\pi^+\gamma$	23.92 ± 1.28	0.13 ± 0.08
$K^+\pi^-\pi^0\gamma$	29.87 ± 2.19	-0.00 ± 0.14
$K_S^0\pi^-\pi^+\gamma$	22.87 ± 1.97	-0.82 ± 0.13
$K_S^0\pi^+\pi^0\gamma$	31.60 ± 4.26	0.23 ± 0.26

TABLE 54. Fitted mean and RMS branching fractions for 500 toy MC studies with feed-down background sampled from KN465 inclusive $b \rightarrow s\gamma$ MC. The last column shows the difference from the branching fractions in table 46.

The table below shows the result of repeating the original MC study, with the feed-down $b \rightarrow s\gamma$ background component increased by 12%, the world average uncertainty on the $b \rightarrow s\gamma$ branching fraction. The feed-down component is still sampled from KN480 inclusive MC. The components of the fit function are unchanged.

Mode	Average BF $\times 10^6$	δ BF $\times 10^6$
$K^+\pi^-\pi^+\gamma$	23.81 ± 1.19	0.03 ± 0.08
$K^+\pi^-\pi^0\gamma$	30.11 ± 2.05	0.23 ± 0.13
$K_S^0\pi^-\pi^+\gamma$	23.97 ± 2.16	0.28 ± 0.13
$K_S^0\pi^+\pi^0\gamma$	31.22 ± 4.42	-0.15 ± 0.27

TABLE 55. Fitted mean and RMS branching fractions for 500 toy MC studies with feed-down background sampled with a 12% increased yield. The last column shows the difference from the branching fractions in table 46.

10.4.4 MODEL SYSTEMATIC UNCERTAINTIES

For each of the toy MC studies described in the previous section, we take the difference between the average fitted branching fraction in each mode (tables 51–55), and subtract from it the average fitted branching fraction in the same mode in the original toy MC study (table 46). We add these differences in quadrature to obtain systematic uncertainties due to $b \rightarrow s\gamma$ background models.

We perform a similar procedure for uncertainties in the $m_{K\pi\pi}$ measurements. We take the difference between the average sPlot distributions in the modified toy MC studies, and subtract the average sPlot distributions in the original toy MC study. We add in quadrature the differences in each bin to arrive at a systematic uncertainty in each bin due to $b \rightarrow s\gamma$ background models. We exclude from this calculation the toy MC sample in which we sampled the signal component from $b \rightarrow s\gamma$ MC, since the generated $m_{K\pi\pi}$ spectrum is different.

The figures below illustrate this systematic uncertainty; the distributions show the generator-level $m_{K\pi\pi}$ distributions, and the error bars show the bin-by-bin systematic uncertainty due to bias computed in this way.

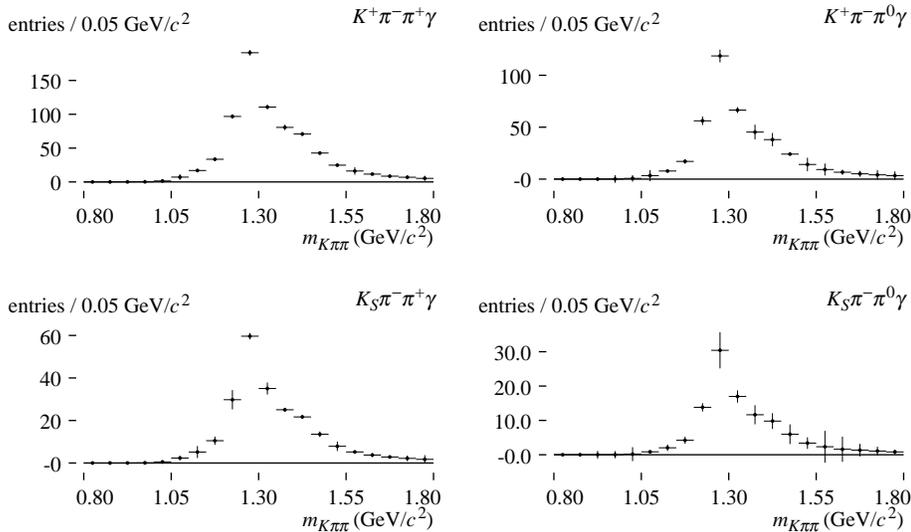


FIGURE 79. Generator-level signal MC distributions of $m_{K\pi\pi}$, with systematic uncertainty due to $b \rightarrow s\gamma$ models indicated by error bars.

The systematic uncertainty in the $m_{K\pi\pi}$ distribution due to $b \rightarrow s\gamma$ model variation is small compared to the statistical uncertainty in our data sample. We therefore neglect this systematic uncer-

tainty when presenting the distribution.

10.5 Peaking backgrounds

Table 24 lists the candidate yields in peaking background channels we have studied in which the entire final state is produced by a single B . The largest contribution is 1.4 candidates expected in the $K^+\pi^-\pi^+\gamma$ mode. We assign the expected number of peaking events as a systematic uncertainty to the yield measured in each mode. We expect that candidates from these decay processes with one or more tracks in the final state taken from the other side of the event are modeled adequately by the generic B MC.

10.6 Combined systematics

The following table is a summary of systematic uncertainties on the measured branching fraction in each mode.

Source	$K^+\pi^-\pi^+\gamma$	$K^+\pi^-\pi^0\gamma$	$K_S^0\pi^-\pi^+\gamma$	$K_S^0\pi^+\pi^0\gamma$
$B\bar{B}$ count	0.6%	0.6%	0.6%	0.6%
$\Upsilon(4S)$ branching fraction	2.4%	2.4%	2.4%	2.4%
photon efficiency	1.8%	1.8%	1.8%	1.8%
photon selection efficiency	2.0%	2.0%	2.0%	2.0%
π^0 and η veto efficiency	1.0%	1.0%	1.0%	1.0%
tracking efficiency	1.4%	1.1%	1.1%	0.8%
π^\pm PID efficiency	1.4%	1.0%	1.4%	1.0%
K^\pm PID efficiency	4.2%	4.2%		
K_S^0 selection efficiency			1.6%	1.6%
π^0 selection efficiency		3.0%		3.0%
Fisher cut efficiency	1.0%	1.0%	1.0%	1.0%
vertex probability cut efficiency	0.7%	0.7%	0.7%	0.7%
fit bias	1.6%	1.3%	1.3%	3.5%
$b \rightarrow s\gamma$ background model	1.4%	1.0%	4.0%	1.3%
peaking backgrounds	0.2%	0.1%	0.0%	0.6%
m_{ES} shift	1.0%	0.5%	1.6%	0.6%
shape parameters	0.1%	2.9%	0.9%	0.2%
total	6.6%	7.5%	6.5%	6.6%

TABLE 56. Summary of systematic uncertainties to the fitted branching fraction in each mode.

11 Conclusions

In this analysis, we have measured the branching fractions of the radiative penguin decay $B \rightarrow K\pi\pi\gamma$ in four charged modes. The results are,

$$\mathcal{B}(B^+ \rightarrow K^+\pi^-\pi^+\gamma) = (2.95 \pm 0.13(\text{stat.}) \pm 0.19(\text{syst.})) \times 10^{-5}$$

$$\mathcal{B}(B^0 \rightarrow K^+\pi^-\pi^0\gamma) = (4.07 \pm 0.22(\text{stat.}) \pm 0.31(\text{syst.})) \times 10^{-5}$$

$$\mathcal{B}(B^0 \rightarrow K^0\pi^+\pi^-\gamma) = (1.85 \pm 0.21(\text{stat.}) \pm 0.12(\text{syst.})) \times 10^{-5}$$

$$\mathcal{B}(B^+ \rightarrow K^0\pi^+\pi^0\gamma) = (4.56 \pm 0.42(\text{stat.}) \pm 0.30(\text{syst.})) \times 10^{-5}.$$

The branching fractions we measure in the $\pi^+\pi^-$ modes are in agreement with previous measurements. The two $\pi^\pm\pi^0$ modes were previously unobserved. The large branching fractions we have established for these and substantial candidate yields we have obtained in our data sample provide encouragement that measurements of the photon polarization of these modes will be possible in larger future B factory data samples.

We have also measured the spectrum of the $K\pi\pi$ invariant mass in these decays, shown below.

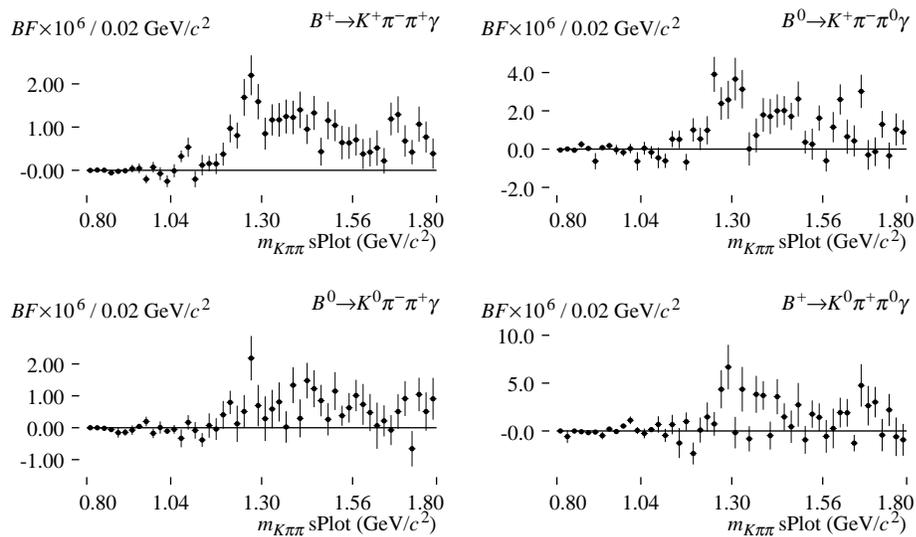


FIGURE 80. Distributions of $m_{K\pi\pi}$ in the on-resonance data sample. Error bars show statistical uncertainty from the sPlot only; these dominate the systematic uncertainties due to $b \rightarrow s\gamma$ models.

References

- [1] Hewett, J. L. and J. D. Wells, *Phys. Rev. D* **55** (1997) 5549.
- [2] Eidelman, S. et al., *Phys. Lett. B* **592** (2004) 1.
- [3] Cheng, H. and C. Chua, hep-ph/0401141 (2004).
- [4] Ebert, D., et al., *Phys. Rev. D* **64** (2001) 054001.
- [5] Safir, A. S., *Eur. Phys. J. direct* **C3** (2001) 15.
- [6] Veseli, S. and M. G. Olsson, *Phys. Lett. B* **367** (1996) 309.
- [7] Gronau, M. et al., *Phys. Rev. Lett.* **88** (2002) 051802.
Gronau, M. and D. Pirjol, *Phys. Rev. D* **66**, (2002) 054008.
- [8] Aubert, B. et al. (BABAR Collaboration), *Phys. Rev. D* **70** (2004) 112006.
- [9] Nakao, M. et al. (Belle Collaboration), *Phys. Rev. D* **69** (2004) 112001,
- [10] Aubert, B. et al. (BABAR Collaboration), *Phys. Rev. Lett.* **93** (2004) 201801.
- [11] Aubert, B. et al. (BABAR Collaboration), *Phys. Rev. D* **70** (2004) 091105.
- [12] Nishida, S. et al. (Belle Collaboration), *Phys. Rev. Lett.* **89** (2002) 231801.
- [13] Yang, H. et al. (Belle Collaboration), hep-ex/0412039 (2004).
- [14] "PEP-II: An Asymmetric B Factory. Conceptual Design Report. June 1993," SLAC-418 (1993).
- [15] Harrison, P. F. and H. R. Quinn, eds., SLAC-504 (1998).

- [16] Seeman, J. et al., EPAC-2004-MOPLT143.
- [17] Aubert, B. et al. (BABAR Collaboration), *Nucl. Instr. and Methods A* **479** (2002) 1.
- [18] Lange, D. J., *Nucl. Instr. and Methods* **462** (2001) 152.
- [19] Sjöstrand, T. et al., *Computer Physics Commun.* **135** (2001) 238.
- [20] Agostinelli, S. et al. (Geant4 Collaboration), *Nucl. Instr. and Methods* **506** (2003) 250.
- [21] Kagan, A. and M. Neubert, *Euro. Phys. J.* **7** (1999) 5.
- [22] Aubert, B. et al. (BABAR Collaboration), *Phys. Rev. Lett.* **93** 021804 (2004).
- [23] Hulsbergen, W. D., physics/0503191 (2005).
- [24] Fisher, R. A., *Ann. Eugenics* **7**, 179 (1936).
- [25] Metropolis, N. et al., *J. Chem. Phys.* **21** (1953) 1087.
- [26] Skwarnicki, T., DESY F31-86-02 (1986).
- [27] Albrect, H. et al. (ARGUS Collaboration), *Phys. Lett.* **B185** (1987) 218.
- [28] Le Diberder, F. R. and M. Pivk, BABAR Analysis Document #509.
- [29] Aubert, B. et al. (BABAR Collaboration), hep-ex/0504001.