Probing the symmetry of nature.
(CP Violation in $B \to J/\psi\ K_L$ Decays)

Kirill Korotushenko

A DISSERTATION

PRESENTED TO THE FACULTY

OF PRINCETON UNIVERSITY

IN CANDIDACY FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE

BY THE DEPARTMENT OF PHYSICS

November 2004

Abstract

In this work we present a measurement of the CP-violation parameter $\sin 2\phi_1$ determined through the difference between the time-dependent decay rates for the $B^0 \to J/\psi\ K_L$ and $\bar{B}^0 \to J/\psi\ K_L$ transitions.

We also show improved selection criteria, which allow us to increase the reconstruction efficiency by 40% and decrease the statistical error in $\sin 2\phi_1$ by 25%. In addition, we will present a detailed overview of the K_L and J/ψ detection and reconstruction.

Using 78 fb^{-1} of e^+e^- collision data, collected with the Belle detector at the KEKB accelerator, we reconstruct 2054 $J/\psi\ K_L$ candidates, with an estimated signal purity of $57 \pm 3\%$. Using 1714 of these candidates, we measure $\sin 2\phi_1 = 0.51 \pm 0.12(\text{statistical}) \pm 0.07(\text{systematic})$

The measurements presented take advantage of the excellent design of the Belle KLM detector. They are in good agreement with the results from other experiments.

Acknowledgments

I would like to say many thanks to my advisor, Daniel Marlow, for his support, and his patience. He allowed me to do what I wanted to do and provided me with an intellectually stimulating environment during the course of my study. It was my pleasure to conduct research under his supervision. I am also grateful for his numerous comments, which helped to improve the quality of this document.

This thesis was a result of a close collaboration with my fellow graduate student, Sven Einar Vahsen, whose work helped shape many improvements and innovations presented here. I would also like to thank him for taking the heat associated with releases of official results. The work of Kazu Hanagaki on K_L detection and particle identification has provided the foundation for several parts of the analysis.

The guidance from Robert Wixted helped me tremendously in the design, implementation and debugging of readout electronics, while useful tips and help from Stanley Chidzik saved me a lot of time during the production cycle. I am grateful to Daniel Marlow for providing computing resources and Christopher Tully, Jeremiah Mans and Wade Fisher for managing it. I am grateful to Alexei Tchouvikov for testing my software and helping me during my visits to Princeton.

I would like to thank Tipanee Pipatanagul for giving me inspiration and moral support throughout my research work. Without her loving support and understanding I might have never completed my Ph.D. journey.

There are many other people who made my stay at Princeton enjoyable. I would like to thank Timur Shutenko, Ilya Kurland and Andrei Litvak for being good friends and for the quality time we spent together. I should not forget Yan Vtorov, Michael Ashikhmin, Alex Buchel, Serguei Komarov, Igor Bilinsky and many others who helped

me to get into graduate school and to go through it. I would also like to thank Oliver Stier, Jagdeep Pawa and Piti Disyatat for their insights into life beyond academia and to Yermolai Solzhenitsyn for showing me the countryside I would not have seen otherwise. I want to thank Alexander Schekochihin, Olexiy Motrunich, Felix Vardy, Akakiy Melikidze, Roman Rafikov, Serguei Gukov, Ilya Nemenman, Ilya Iofin and Daniel Erlihman for occasionally dragging me out of the lab to see the sunlight. Finally, I would like to thank everyone whose direct and indirect efforts helped me to complete my thesis.

I would like to thank Stew Smith for devoting his time to read this document and for his comments and suggestions.

Contents

List of Figures

List of Tables

Chapter 1

Motivation

This work describes the measurement of CP violating parameters in B^0 ($b\bar{d}$) meson decays into J/ψ (1^3S_1 $c\bar{c}$) and K_L ($K_L \approx 1/\sqrt{2}(s\bar{d} + \bar{s}d)$). This decay lies in the shadow of the "gold-plated mode"(GPM) ($B^0 \to J/\psi\ K_S$), which is believed to be the reaction that provides the most powerful tool to probe CP symmetry within the Standard Model (SM). Theoretical contributions are identical for both decays, but experimental complications in $B^0 \to J/\psi\ K_L$ measurements were to blame for lower precision. In this thesis we will show that despite these problems, the K_L mode can provide a precision comparable to the $B^0 \to J/\psi\ K_S$ decay. Adding results from both channels improves the overall statistical sensitivity, and the asymmetry in the $B^0 \to J/\psi\ K_L$ decay should have the sign opposite to the $B^0 \to J/\psi\ K_S$ decay, thus giving an another test of the theoretical approach.

CP violation is one of the greatest physics mysteries of the 20^{th} century. Even though it is a well established phenomenon and the SM describes it in great detail, there is no satisfactory explanation of the origin of CP violation. At least five Nobel prizes have been awarded for work related to C, P and T transformations[1], showing how rich and counterintuitive discrete symmetries are. Despite major advances in the

[1] 1933 Dirac, prediction of antimatter (C); 1933 C.Anderson discovery of positron (C); 1955 O.Chamberlain and E.Segre discovery of antiproton (C); 1956 T.D. Lee and C.N.Yang, parity test in a weak process (P); 1964 J.Christensen, J.Cronin, V.Fitch and R.Turlay, CP violation

understanding of symmetries, many physicists still believe that discoveries related to discrete symmetries are not over.

Another great puzzle possibly related to CP violation comes from cosmology. We still do not understand the prevalence of matter over anti-matter in the Universe. If the Big Bang scenario is true, the early stages of the universe's evolution would have equal numbers of baryons and anti-baryons. However, current observations show that we live in a matter-dominated universe and experiments do not reveal any baryon-number violation. This could indicate that the evolution of the universe is neither T nor CP invariant, given that combined CPT invariance holds. Moreover, the large observed baryon-number-to-photon ratio ($\eta = 6.1^{+0.3}_{-0.2} \times 10^{-10}$[1]) hints at sizeable asymmetries in matter-antimatter behavior. This is one of the reasons that studying CP violation may be a good place to observe physics beyond the Standard Model.

Before the 20^{th} century, physics mainly described *how* matter interacts and propagates, but failed to address questions like: *what* the matter is and *why* it has the properties we observe. Even though over the past century, there has been significant progress toward addressing these issues, many questions remain. For example, in the SM, we do not know *why* we have three generations of quarks or leptons. Even though we understand better *what* mesons are made of and *how* they interact, we do not know *how* to calculate the KM parameters and *what* type of physics is hidden behind them.

Most experimental evidence helps us in the *how* questions, while logic is used to fill in the rest. On the other hand, our improved knowledge of *what* and *why* can dramatically alter our understanding of *how*. In other words, to figure out why quarks and leptons are the way they are, we might have to revisit the laws of motion in exactly the same way as progress in electrodynamics, rather than Newtonian mechanics, gave birth to Lorentz transformations.

With that in mind, we move forward and describe a measurement of CP violation in the neutral B meson system, adding one more piece of the puzzle.

Chapter 2

CP violation

2.1 Standard model.

Weak interactions in the hadron sector are successfully described by the so-called Standard Model (SM). At this point, we know that there are six leptons and six quarks plus their corresponding antiparticles. Leptons and quarks, despite their differences, can be arranged into three families (table 2.1) with transitions within the families being preferred (e.g. $t \to b$ or $e \to \nu_e$). Transitions between adjacent hadron families give birth to a mixing matrix in the quark sector.

In the SM, the weak interaction is carried by the weak bosons: the $W^{\pm}$ and the Z^0. These bosons couple to both leptons and quarks. Given that the charge of W bosons is ± 1, weak transitions are only allowed for $\Delta Q = -1, 0, +1$, thus disallowing any non-integer lepton-hadron weak current. The result of this rule is an entirely detached mixing matrix for quarks and leptons. Another interesting consequence is that only quark-quark transitions are allowed, because quark-antiquark as well as lepton-quark transition will require transfer of non integer charge (like 1/3, 2/3, etc ...).

The charged-current weak interaction responsible for flavor changes is often written in the form $gW^{\mu}J^{cc}_{\mu}$. The structure of the weak current J^{cc}_{μ} encompasses all prior experimental knowledge. The allowed transitions are governed by detached lepton

Charge	1	2	3
-1	$\bar{e}$	$\bar{\mu}$	$\bar{\tau}$
$-\frac{2}{3}$	$\bar{u}$	$\bar{c}$	$\bar{t}$
$-\frac{1}{3}$	d	s	b
0	ν_e, $\bar{\nu_e}$	ν_μ, $\bar{\nu_\mu}$	ν_τ, $\bar{\nu_\tau}$
$\frac{1}{3}$	$\bar{d}$	$\bar{s}$	$\bar{b}$
$\frac{2}{3}$	u	c	t
1	e	μ	τ

Table 2.1: Three generations of leptons and quarks

and hadron contributions with unitary mixing matrixes V_l in the lepton sector and a V_q in the hadron sector. C and P violation are incorporated through the algebra of γ matrixes. CP violation is introduced via a complex phase in the mixing matrix. The summation in the hadron sector comes from the color structure of the hadron interaction. In short, the W boson is coupled to the following current

$$J_\mu^{cc} = \begin{pmatrix} \bar{\nu}_e \\ \bar{\nu}_\mu \\ \bar{\nu}_\tau \end{pmatrix} \gamma_\mu \frac{1-\gamma_5}{2} V_l \begin{pmatrix} e \\ \mu \\ \tau \end{pmatrix} + \sum_{r,g,b} \begin{pmatrix} \bar{u} \\ \bar{c} \\ \bar{t} \end{pmatrix} \gamma_\mu \frac{1-\gamma_5}{2} V_q \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{2.1}$$

The essence of the SM is that the quark mass eigenstates are not the same as weak eigenstates, thus allowing transitions between different families of the quarks or flavor violation (weak interaction does not conserve strangeness, charm, etc ...). For example, in the decay $K \to \pi\pi$, the strange quark originally present in the kaon evaporates without a trace. By convention, the mixing is often expressed in terms of a 3×3 unitary matrix V operating on the charge $-e/3$ quark mass eigenstates (d, s

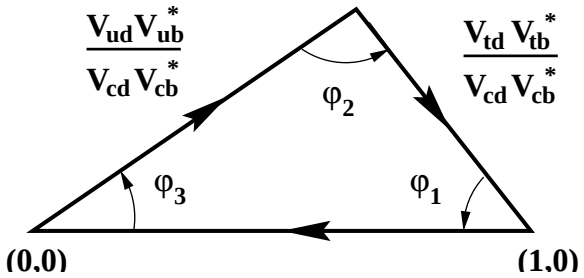

Figure 2.1: Sketch of the unitarity triangle

and b):

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{2.2}$$

This mixing matrix is called the Cabbibo-Kobayashi-Maskawa (CKM) [2, 3] matrix. The CKM matrix should be unitary, because the transition probability of, say, a u quark to decay into all possible final states—c, s and d—should be unity. This requirement can be visualized through the so called unitary triangles, where three complex phases φ_1, φ_2 and φ_3 sum to zero. Fig. 2.1 shows an example of the $b-d$ unitary triangle with respect to the first and the third column of the mixing matrix. That triangle has all sides with comparable length, while the other two unitary triangles are flat. That property makes the $b-d$ triangle the most attractive for phase measurement.

Let us start by determining the number of independent parameters needed to describe an $N \times N$ mixing matrix. Before imposing any constraints, this matrix can have $2N^2$ independent real parameters, which is enough to specify N^2 complex numbers. The unitarity requirement gives rise to N constraints. Another consequence of unitarity is the requirement that all columns be orthogonal. This introduces $N(N-1)/2$ additional equations for both the imaginary and real parts, so the number of independent parameters goes down to $2N^2 - N - N(N-1) = N^2$. But this is not the end, since not all phases contained in the CKM matrix will be meaningful. For

example, each element of the matrix is an amplitude for the corresponding transition between the quark states $V_{\alpha\beta} = \langle\alpha|\mathcal{H}_W|\beta\rangle$. A simple quark-phase transformation $\langle\alpha| \to e^{i\theta}\langle\alpha|$ will lead to a redefinition of $V_{\alpha\beta}$, but should not change the physical outcome. As a result, by redefining quark states, we can remove the relative phases from the CKM matrix. In the case of N quark families, we will have $2N$ quarks and $2N-1$ relative phases. Thus we can freely redefine the relative phases and get the final number of independent parameters $N^2-(2N-1) = (N-1)^2$. The number of independent orthogonal rotations in an N dimensional space is given by the number of Euler angles: $N(N-1)/2$, which result in the $(N-1)(N-2)/2$ complex phases for $N \times N$ matrix. Note that for two generations, we will have a real CKM matrix, while for three generations of quarks, we will get one imaginary phase. The existence of a non-trivial phase in the 3×3 scheme was Kobayashi and Maskawa's (K-M's) main insight in the explanation of the CP violation [3]. That was remarkable prediction, given that the charm, top and bottom quarks had not been discovered yet. This complex phase is a cornerstone, which allows CP violation in the SM. Later we will show in detail how this complex parameter gives rise to CP violating asymmetries. The decays of neutral kaons are dominated by transitions involving only two generations of quarks, which makes the mixing matrix close to real, resulting in small CP violating effects [4, 5, 6]. It is widely expected that only decays with comparable contribution from all three generations of quarks will exhibit large CP violation in the standard model framework.

There are several parametrizations for the above matrix. A standard parametrization of V utilizes angles θ_{12}, θ_{23}, θ_{13}, and a phase, δ_{13}

$$V = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & c_{13}e^{-i\delta 13} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta_{13}} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta 13} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta_{13}} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta 13} & c_{23}c_{13} \end{pmatrix} \tag{2.3}$$

with $c_{ij} = \cos\theta_{ij}$ and $s_{ij} = \sin\theta_{ij}$ for the generation labels $i, j = 1, 2, 3$. This has a distinct advantage of interpretation, for the rotation angles are defined and labelled

in a way that relates to the mixing of two specific generations and if one of those angles vanishes, so does the mixing between the two generations.

A popular approximation that emphasizes the hierarchy in the size of the angles, $s_{12} \gg s_{23} \gg s_{13}$, is due to Wolfenstein [7], where one sets $\lambda \equiv s_{12}$, the sine of the Cabbibo angle, and then writes the other elements in terms of powers of λ:

$$V \approx \begin{pmatrix} 1-\lambda^2/2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\lambda^2/2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2/2 & 1 \end{pmatrix} \tag{2.4}$$

With A, ρ and η real numbers that turned out to be of order unity. This parametrization gives an intuitive grasp of the quark transition strengths and provides a simple tool for dominant decay estimation. The triangle presented on Fig. 2.1 is the only unitary triangle with all sides having comparable length. This makes it especially attractive for determining the weak phase in the SM.

2.2 Discrete symmetries.

Symmetries play an important role in modern physics. Not long ago, people realized that any symmetry in a system leads to a conserved quantity. The most prominent and well known example is a connection between continuous symmetries—translational, rotational and time invariance—and conserved quantities—momentum, angular momentum and energy.

Classical physics does not provide a reason for introducing discrete symmetries, mostly because it only deals with continuous variables. The advent of quantum mechanics changed this situation dramatically. Continuous variables have been replaced by variables allowed to have only a discrete set of values. As a result, new areas of physics, dealing with this type of symmetry emerged. Incomprehensible from a classical point of view, the understanding of fields like atomic spectroscopy, crystallography and solid state physics were boosted immeasurably. Simple rules for allowed

transitions followed from the application of symmetry principles. Oddly enough, this revolution did not touch the state of the vacuum nor "space as a whole," which has preserved a continuous spectrum and therefore did not yield any globally conserved quantity.

The introduction of Lorentz-invariant field theory changed this peaceful landscape. We will not go into the derivation of C,P and T operator properties in either relativistic or non-relativistic quantum mechanics, but rather will provide some rationale why these transformations are favored by relativistic dynamics (for an excellent review of C, P and T operators see [8]).

The Lorentz transformation is a simple rotation in 4-space, but with one distinct feature. Any physical rotation should leave the time axis within the light cone. This innocent property has far reaching consequences. It creates three distinct regions, which can not be connected by any physical process (one unavailable region and two separated physical regions). The existence of two separate, otherwise, identical regions gave birth to a new quantum number which happened to be the same as electromagnetic charge. But while applying charge conjugation (C) we are supposed to end up in the second cone, which effectively means that time will flip the sign $t \rightarrow -t$ (T). However, the T transformation will change the state of the system—i.e., moments will also change sign. In order to get the same state of motion, we have to apply space inversion: $x \rightarrow -x$ (P). Thus, combined CPT transformations should return the system back into an identical kinematic state. This is a rather crude proof of the CPT theorem, but it does illustrate the reasoning behind it. The combined PT transformation can also be viewed as a Lorentz rotation (with $det = 1$), thus any Lorentz invariant quantity should stay invariant under this transformation as well.

In other words, CPT formalism was born only after Lorentz transformations were well digested by the physics community. Let us define these transformations. In the late 1920's, Paul Dirac predicted that every particle has a corresponding antiparticle [9, 10, 11]. The fundamental difference between particles and antiparticles is the

opposite sign of all internal quantum numbers. For example, an *electron* with charge $-e$ will have an antiparticle *positron* with opposite charge $+e$ and the same mass and spin. Some particles, like the photon, are their own antiparticle. Others, like the neutral B mesons, have a distinct antiparticle ($B_d^0 = \bar{b}d$, $\bar{B}_d^0 = b\bar{d}$). In the case of the B meson, particle and antiparticle are distinguished by a quantum number called "beauty", which is equal to the number of beauty quarks in the system. This quantum number is conserved in all interactions except for the weak. In order to make concepts more readable we give definitions from the point of view of classical physics.

- **C** stands for charge conjugation, but it means more than that. This operator is designed to change particle into antiparticle, thus flipping the sign of all internal quantum numbers (electric charge among them).

- **P** stands for parity transformation, or simply flipping the sign of all space coordinates.

- **T** stands for time inversion, or simply flipping the sign of time. It can be viewed as reversal of motion. Quantum mechanics adds complex conjugation for matrix elements.

Modern physics deals with four major forces: strong, electromagnetic, weak and gravity. At our mass and energy scale, gravity forces are negligibly small, so we will omit gravity, which is C, P and T invariant, from further discussion. The electromagnetic and strong forces play an important role in the formation and decay of elementary particles, but despite that we will not pay much attention to the structure of these forces in consequent discussions. The main reason behind this is an assumption of their C, P and T invariance, with all the experimental evidence standing firmly for it.

In the hadron sector, strong and weak forces were originally separated, as their names suggest, by the strength or typical time of interaction. But as time progressed,

it was noted that weak interaction in addition to being weak also violates most quantum number conservation laws (strangeness, bottomness etc ...), while strong and electromagnetic interactions seem to preserve them. The same can be said about C, P and CP invariance. As a result, when we talk about weak and strong interactions, we imply CP-invariant and non-invariant forces and the fact that conserving forces have stronger couplings can be seen as a mere coincidence.

When we write transition amplitudes for two charge conjugate processes $i \to f$ and $\bar{i} \to \bar{f}$, governed by weak and strong decay amplitudes, we can state that

$$\begin{aligned} V_{if} &= \textstyle\sum_j A_j e^{+i\delta^j_{\text{Weak}}} e^{i\delta^j_{\text{Strong}}} \\ V_{\bar{i}\bar{f}} &= \textstyle\sum_j \bar{A}_j e^{-i\delta^j_{\text{Weak}}} e^{i\delta^j_{\text{Strong}}} \end{aligned} \tag{2.5}$$

Interference between the amplitudes which differ in both strong and weak phase can result in different rates for $i \to f$ and $\bar{i} \to \bar{f}$ reaction and herald so called "direct" CP violation. Although there are some hints, experiments to date have not provided strong evidence for the existence of direct CP violation. A second large class of CP violating effects is exploring possible interference between two competing weak phases and is called "indirect" or "mixing induced" CP violation.

Strong forces are invariant under C conjugation and dominant in the formation of mesons, leading to the identical "strong" mass for both neutral particle and antiparticle. This, in turn, opens a possibility for "weak" forces to split the mass eigenstate and to arrange transitions between them. Transitions between neutral particles and their antiparticle counterparts are called "mixing". Mixing plays a crucial role in modern experiments for probing CP, because it provides a transition that does not affect the strong phase. Indirect CP violation is closely associated with mixing and often divided further into violation in mixing (different rates for $(i \to \bar{i}) \to \bar{f}$ and $(\bar{i} \to i) \to f$) and violation in the phase mismatch between the mixing and the decay amplitude to the CP eigenstate (different rates for $[(i \to \bar{i}) \to f] + [(i \to f]$ and $[(\bar{i} \to i) \to f] + [\bar{i} \to f])$. The former will signal either direct CP violation in $i \to \bar{i}$ transition or indirect interference between more than one contributing process. The

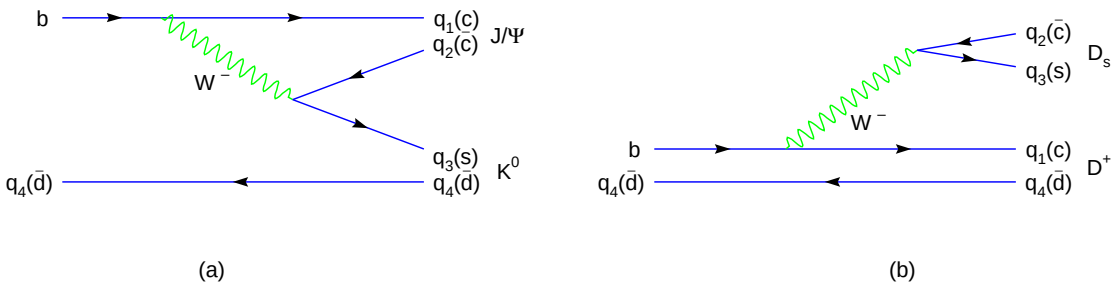

Figure 2.2: Tree level diagram for $b \to c\bar{c}s$ transition.

later will not be able to establish whether the CP violation happens in the mixing part or in the interference between mixing and decay, but nevertheless will establish the phenomena.

CP violation induced by a complex phase can be observed and quantitatively measured in the cases where neutral mesons decay into CP eigenstates. B_d^0 mesons are particularly attractive since the characteristic times for decay and oscillation are of the same order and since b-quark decays can involve all three quark generations.

2.3 Feynman rules for B meson decay.

Decays of the B meson are governed by weak interactions with flavor changing currents standing behind its short lifetime. In the spectator approximation, the simplest diagram for b-quark decay involves one $W^{\pm}$ boson (see Fig. 2.2).[1] The CKM matrix is such that the $b \to c$ transition dominates b-quark decay (the Cabibbo favored $b \to t$ transition is forbidden by energy conservation). That transition involves emission of a W boson, which in turn decays into a $l\nu$, a ud, or a cs pair. The two possible hadronization schemes will result in the following possible final states.

[1]Most of the examples in this short review are based on the $B^0 \to J/\psi\ K_S$ decay, however presented arguments can be used for a variety of other transitions.

$$\bar{B}^0_d \to (c\bar{c}) + \left(s\bar{d}\right) \quad \bar{B}^0_d \to (s\bar{c}) + \left(c\bar{d}\right),$$
$$\bar{B}^0_d \to (c\bar{u}) + \left(d\bar{d}\right) \quad \bar{B}^0_d \to (d\bar{u}) + \left(c\bar{d}\right),$$
$$\bar{B}^0_d \to (l\bar{\nu}) + \left(c\bar{d}\right),$$

where two quark combinations can be replaced with any of the stable mesons.

$$\begin{array}{lcllcl} (c\bar{c}) & : & J/\psi, \chi_c, \eta_c, \ldots & (s\bar{c}) & : & D_s^-, D_s^{*-}, \ldots \\ (c\bar{u}) & : & D^0, D^{*0}, \ldots & \left(c\bar{d}\right) & : & D^+, D^{*+}, \ldots \\ \left(s\bar{d}\right) & : & K^0, K^{*0}, \ldots & \left(u\bar{d}\right) & : & \pi^+, \rho^+, \ldots \end{array}$$

All of the favored transitions listed above will have a charm quark among the decay products. Among the c based mesons, only the $c\bar{c}$ family and the D^0 system are invariant under C. Of the various possibilities $B_d \to J/\psi\ K^0$ decays prove to be highly attractive both theoretically and experimentally. All detectable variations of the K^0 system are eigenstates of CP (K_S: $CP = +1$, K_L: $CP = -1$). J/ψ being the $c\bar{c}$ state in the $1S$ spherical configuration has negative intrinsic charge conjugation number and parity ($C = (-1)^{L+S} = -1$, $P = (-1)^{L+1} = -1$, $CP = +1$). The parity assignment for the $J/\psi K_L$ system is less straightforward due to the J/ψ's vector nature. However, both B_d and K^0 are spinless particles. The conservation of angular momentum will force the $J/\psi K_L$ to be produced in a p-wave thus forcing it to be in $CP = +1$ state. Other transitions of the $b\bar{d} \to c\bar{c} + s\bar{d}$ type either face considerable difficulties in the reconstruction stage or represent $CP = -1$ and $CP = +1$ addmixtures, thus requiring a nontrivial angular analysis.

The final state produced in the tree B meson decay will consist of four quarks, which can hadronize in two fundamentally different ways. The quarks coming from the W decay have the same color and as a result can form a separate colorless meson without any color restrictions (Fig. 2.2(b)), while in the process depicted in Fig. 2.2(a) quarks from the W decay must match the color of the spectator quark. Given that there are three color states and any stable particle must be colorless, one would naively expect the latter channel to be three times less likely than the former one.

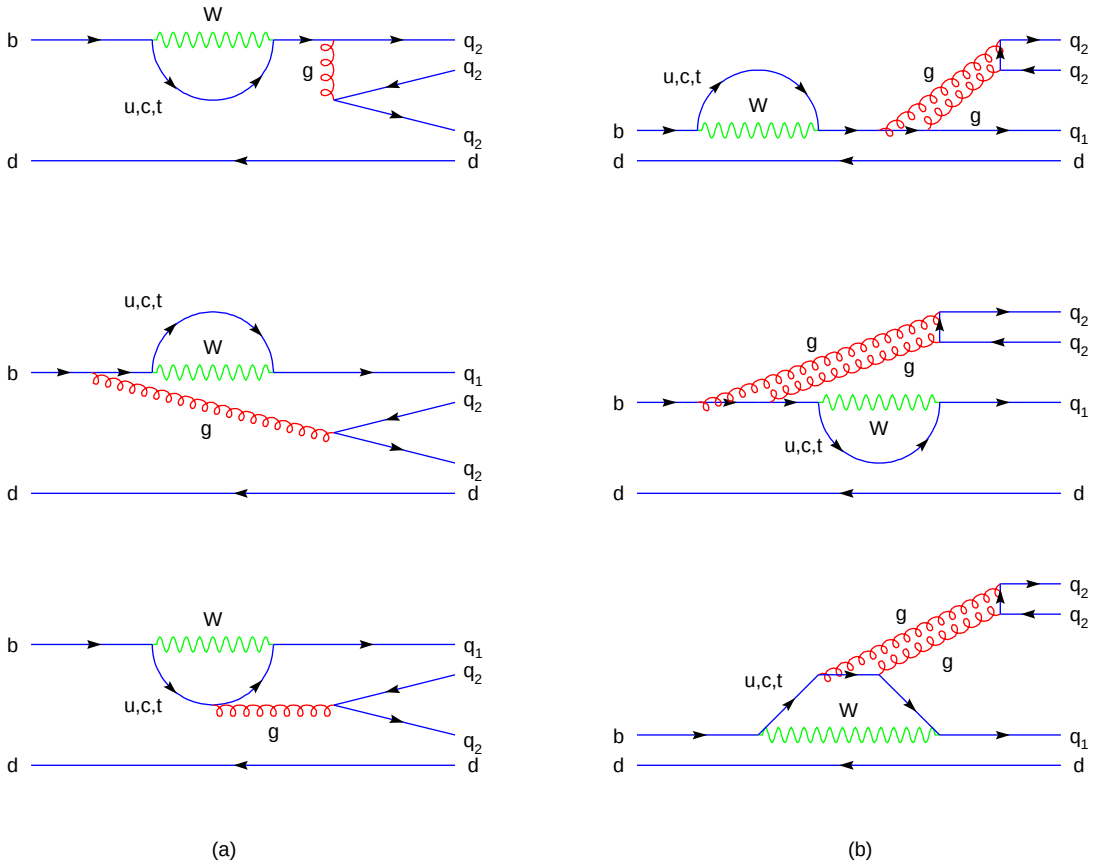

Figure 2.3: Gluonic penguin contamination

More detailed calculations [12] predict an even higher suppression factor of five. This is rather unfortunate, because it leads to suppression of the neutral-B-meson decays into pairs of neutral particles. Recent experimental results show an even greater asymmetry, $Br(B_d \to J/\psi\ K^0) = (8.9 \pm 1.2) \times 10^{-4}$ is almost ten times less likely than $Br(B_d \to D^-\ D_S^+) = (8 \pm 3) \times 10^{-3}$ [13]. To be more accurate, one has to sum the branching ratios for all $c\bar{c}$, $s\bar{d}$, $s\bar{c}$, $c\bar{d}$ states, and indeed, the total sum for all known decays gives a similar result. The example above shows how profoundly the strong interaction can affect otherwise identical weak processes.

If two or more weak transitions of comparable amplitude contribute, the overall phase for the decay will depend on both strong and weak phases. Given that our knowledge of strong phases is far from perfect, we will have to focus on decays involving one dominant diagram, for which the CP violating phase will be well defined

by SM predictions. In order to make sure that we will have no contamination in the $B^0 \to J/\psi\ K_S$ decay, we have to look into the second-order corrections—the diagrams with two weak interaction vertexes (see Fig. 2.3). The number of gluon lines for such diagrams is arbitrary, with the only limitation coming from the right side diagrams, where one-gluon transitions are forbidden because a single gluon cannot turn into a colorless meson. These sorts of two-gluon penguin diagrams have been nicknamed "vacuum-cleaner diagrams"[14]. The lowest order penguin diagram involves one $W^{\pm}$ and one gluon (as the name gluonic penguin suggests). The penguin amplitude for such a process can be written in the form:

$$P = V_{tb}V_{ts}^{*}P_t + V_{cb}V_{cs}^{*}P_c + V_{ub}V_{us}^{*}P_u \tag{2.6}$$

The P_i values describes the gluon amplitude for the corresponding quark i. The values for the P_i diverge, however, and the standard solution is to use the SM unitarity condition $V_{tb}V_{ts}^{*} + V_{cb}V_{cs}^{*} + V_{ub}V_{us}^{*} = 0$ to regroup them in the following way

$$P = V_{cb}V_{cs}^{*}(P_c - P_t) + V_{ub}V_{us}^{*}(P_u - P_t) \tag{2.7}$$

The $B^0 \to J/\psi\ K_S$ decay weak coupling for the first type of penguin is of order λ^2, while the second is of order λ^4. The leading penguin has the same weak phase as the main tree diagram and, as such, will not affect the total phase. The second term is second order (in λ) with respect to the leading tree diagram. It is usually assumed that the gluonic penguin diagrams are further suppressed [15, 16, 12]. The reason behind this suppression is the smallness of the B-meson wavefunction at the origin and the fact that gluon exchange at larger distances requires a rather energetic gluon. Rough estimates of the ratio between the gluonic penguin and the tree diagram (independent of the hadronic matrix elements and the CKM factors) can be found in [17]. It equals

$$\frac{\alpha_s(m_b)}{6\pi} \ln \frac{m_t}{m_b} \approx 0.04 \tag{2.8}$$

The estimate above might be a bit optimistic because the experimental evidence hints at larger-than-expected branching ratios for decays dominated by gluonic penguins.

However, the combination of the λ^2 suppression with the gluonic amplitude guarantees that the weak phase in the $B^0 \to J/\psi\ K_S$ decay will be dominated by the tree-transition phase.

Fortunately, contributions arising from both electroweak penguins and decays involving spectator quarks are small and can safely be neglected when calculating the $B^0 \to J/\psi\ K_S$ decay phase.

In summary, the $B^0 \to J/\psi\ K_S$ decay is believed to be extremely clean, since the leading penguin contamination has the same phase as the main tree diagram, and the second-order penguin contribution is small (see estimate in Eq. 2.8). As a result, possible contamination is believed not to exceed 1%.

2.4 B mixing and $B^0 \to J/\psi\ K_S$ decay.

Despite its clean quark transition signature, the $B^0 \to J/\psi\ K_S$ decay is affected by quark-antiquark interference. The detectable kaon states are not the flavor eigenstates. As a result, observation of $B^0 \to J/\psi K_{S(L)}$ decay corresponds to the superposition of the following four decays:

$$\begin{aligned} \bar{B}^0_d \to J/\psi \bar{K}^0, \quad \bar{B}^0_d \to B^0_d \to J/\psi \bar{K}^0, \\ \bar{B}^0_d \to J/\psi K^0, \quad \bar{B}^0_d \to B^0_d \to J/\psi K^0, \end{aligned}$$

The time evolution of both B^0 and K^0 systems can be described by the two component Schrödinger equation and will result in B^0-$\bar{B}^0$ and K^0-$\bar{K}^0$ oscillations. The observable mass eigenstates B_L, B_H, K_S, K_L can be written:

$$\begin{aligned} |B_L\rangle = p_B|B_d\rangle + q_B|\bar{B}_d\rangle, \quad |K_L\rangle = p_K|K^0\rangle + q_K|\bar{K}^0\rangle, \\ |B_H\rangle = p_B|B_d\rangle - q_B|\bar{B}_d\rangle, \quad |K_S\rangle = p_K|K^0\rangle - q_K|\bar{K}^0\rangle \end{aligned} \tag{2.9}$$

The decay amplitudes for any final state common to both B^0 and $\bar{B}^0$ will be a superposition of direct transition and the transition through oscillations. The calculations presented in Appendix A result in the following asymmetry for the decays of the B_d

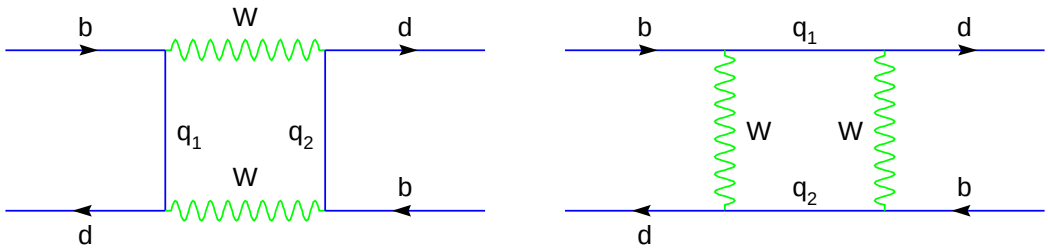

Figure 2.4: Neutral B meson mixing

meson into the CP eigenstate f.

$$\frac{\Gamma(B_d(t) \to f) - \Gamma(\bar{B}_d(t) \to f)}{\Gamma(B_d(t) \to f) + \Gamma(\bar{B}_d(t) \to f)} = \frac{-2e^{0.5\Delta\Gamma t} \sin 2\phi \ \sin \Delta M t}{1 + e^{\Delta\Gamma t} + [1 - e^{\Delta\Gamma t}] \cos 2\phi} \tag{2.10}$$

where 2ϕ is the total weak phase corresponding to both mixing and final state decay, where the factor of two comes in because one sees the difference between particle and antiparticle. ΔM and $\Delta\Gamma$ are respectively the mass splitting and the width difference between the heavy and light mass eigenstates. The full weak phase $2\phi_{B^0 \to J/\psi \ K_S}$ for the $B^0 \to J/\psi \ K_S$ decay can be described as a product for B^0, K^0 mixing and $B^0 \to J/\psi K^0$ transition.

$$e^{i2\phi_{B^0 \to J/\psi \ K_S \text{ decay}}} = \frac{q_B}{p_B} \bar{\rho}(f) \frac{q_K}{p_K} = \frac{q_B}{p_B} \frac{\bar{A}}{A} \frac{q_K}{p_K} \tag{2.11}$$

The factor q/p is defined by the transition amplitudes for the corresponding particle-antiparticle mixing in B and K mesons and $\bar{A}/A$ is the ratio of the particle and antiparticle decay amplitudes for the $B \to J/\psi + K^0$ transition. Below we present estimates for these three factors.

Mixing between the B^0 meson and its antiparticle occurs through box diagrams, where two W bosons are exchanged between a pair of quark lines. The mixing amplitude is proportional to

$$\langle \bar{B}_d^0 | \mathcal{H} | B_d^0 \rangle \sim \sum_{i,j} V_{id} V_{ib} V_{jd}^* V_{jb}^* S(r_i, r_j), \tag{2.12}$$

where $S(r_i, r_j)$ is a loop function which depends on the masses ($r_i \equiv m_i^2/M_W^2$) of the up type quarks running along the internal fermionic lines.

Owing to the unitarity of the CKM matrix, the mixing amplitude vanishes for equal (up-type) quark masses (GIM mechanism [18]); thus the effect is proportional to the fourth power of the mass splitting between the u, c and t quarks. Since the different CKM factors are of similar size, $V_{ud}V_{ub}^* \sim V_{cd}V_{cb}^* \sim V_{td}V_{tb}^* \sim A\lambda^3$, the final amplitude is dominated by the top-quark contribution:

$$\langle \bar{B}_d^0 | \mathcal{H} | B_d^0 \rangle \sim V_{td}V_{tb}V_{td}^*V_{tb}^* S(r_t, r_t), \tag{2.13}$$

Similar arguments are applicable for neutral K mixing and give rise to the following equation:

$$\langle \bar{K}^0 | \mathcal{H} | K^0 \rangle \sim \sum_{i,j} V_{id}V_{is}V_{jd}^*V_{js}^* S(r_i, r_j), \tag{2.14}$$

Unlike the B meson system, mixing in the kaon sector is dominated by the c-quark contribution. The top-quark contribution is suppressed by the smallness of the CKM matrix element ($V_{st}V_{td} \sim \lambda^5$, $V_{sc}V_{cd} \sim \lambda, V_{su}V_{ud} \sim \lambda$), while the up-quark contribution is suppressed by the loop function $S(r_i, r_j)$, which depends on the quark masses ($r_i \equiv m_i^2/M_W^2$).

Taking both B and K mixing phases into account, the final phase shift for the $B^0 \to J/\psi\ K_S$ decay can be written:

$$e^{i2\phi} = \frac{q_B}{p_B}\bar{\rho}(f)\frac{q_K}{p_K} = \frac{q_B}{p_B}\frac{\bar{A}}{A}\frac{q_K}{p_K} = \frac{V_{td}V_{tb}^*}{V_{td}^*V_{tb}}\frac{V_{cb}V_{cs}^*}{V_{cb}^*V_{cs}}\frac{V_{cs}V_{cd}^*}{V_{cs}^*V_{cd}} = \frac{V_{td}V_{tb}^*}{V_{td}^*V_{tb}}\frac{V_{cb}V_{cd}^*}{V_{cb}^*V_{cd}} \tag{2.15}$$

The phase ϕ from Eq. 2.15 is equal to ϕ_1 from the b-d unitarity triangle [2] (see Fig. 2.1). This allows estimation of the CP violating phase in the SM from the observed time-dependent asymmetry in decays of neutral B mesons.

[2] ϕ_1 is the angle between $V_{cd}V_{cb}^*$ and $V_{td}V_{tb}^*$, ϕ_2 is the angle between $V_{td}V_{tb}^*$ and $V_{ud}V_{ub}^*$ and ϕ_3 is the angle between $V_{ud}V_{ub}^*$ and $V_{cd}V_{cb}^*$). An alternative naming convention is $\alpha = \phi_2$, $\beta = \phi_1$ and $\gamma = \phi_3$

2.5 Experimental procedure.

In order to measure $\sin 2\phi_1$, one must observe the time-dependent asymmetry between $B^0 \to J/\psi\ K_S(K_L)$ and $\bar{B}^0 \to J/\psi\ K_S(K_L)$ decays. The proper use of Eq. 2.10 requires knowledge of the B^0 meson production time when it is in a pure flavor eigenstate, B_d or $\bar{B}_d$, and the decay time to the CP eigenstate. While the charged-track decay products of the B^0 can be used to determine the decay vertex of the CP eigenstate B, there is no simple way to determine its production point in an e^+e^--collider experiment, where B^0-$\bar{B}^0$ pairs are produced via the reaction $e^+e^- \to \gamma^* \to \Upsilon(4S) \to B_d\bar{B}_d$.

However, the decay vertex of the second B meson produced in the e^+e^- collision can provide us with the missing information. To a good approximation, all products of the e^+e^- collision will be created through one-photon exchange, and as such will have the quantum numbers of the photon, namely $J^{PC} = 1^{--}$ [3]. Since the B^0 has $J = 0$, it follows that the $B\bar{B}$ pair will have to be in a p-wave. The B mesons, which are bosons, must form only symmetric configurations. Even though B mesons can mix ($B^0 \leftrightarrow \bar{B}^0$), they can never instantaneously be the same. This imposes an interesting restriction on how the B pair can evolve, namely if we detect that one of the mesons decayed as B_d, it means that the other meson will, at the same time, be pure $\bar{B}_d$.

Once we have established that one of the B mesons has decayed into a CP eigenstate, the rest of the event is used to determine the flavor and decay time of the second B meson, thus providing us with information needed to estimate the asymmetry. The determination of the flavor of a neutral meson is called "tagging." Flavor tagging exploits the correlation between the beauty flavor of the parent b-hadron and the charge of the its decay products. However, low efficiency for full reconstruction of B meson forces us to use partial reconstruction to increase tagging efficiency. Certain

[3] $\Upsilon(4S)$ has the same quantum numbers

reconstructed particles from B meson decay, like high momentum lepton have high degree of correlation with the beauty flavor of the parent B meson. Such particles we denote as a "tagging particles" and use to assign flavor probability for parent B meson.

Chapter 3

Experiment overview

3.1 Physics requirements

The need to measure the time evolution of the B mesons puts some major constraints on the accelerator and detector parameters. The short B meson lifetime favors a large boost for the measurement of B meson evolution, which can be achieved by building an asymmetric e^+e^- collider. Increased boost values will, however, decrease the detection efficiency, because more decay products will be thrown forward along the beam line, where the detector acceptance is low. Even for a large boost, one still needs a high-precision silicon vertex detector to extract the position of the B meson decay. The desire to produce as many B_d mesons as possible forces us to operate at the $\Upsilon(4S)$ resonance, which lies right above the $B_d\bar{B}_d$ decay threshold and predominantly decays into $B\bar{B}$. The relatively high mass of the b quark ensures an abundance of possible final states and, as a result, any given decay will have a relatively low branching ratio, thus requiring a high luminosity collider. In the next several sections, we will discuss how these goals are met at the KEK laboratory.

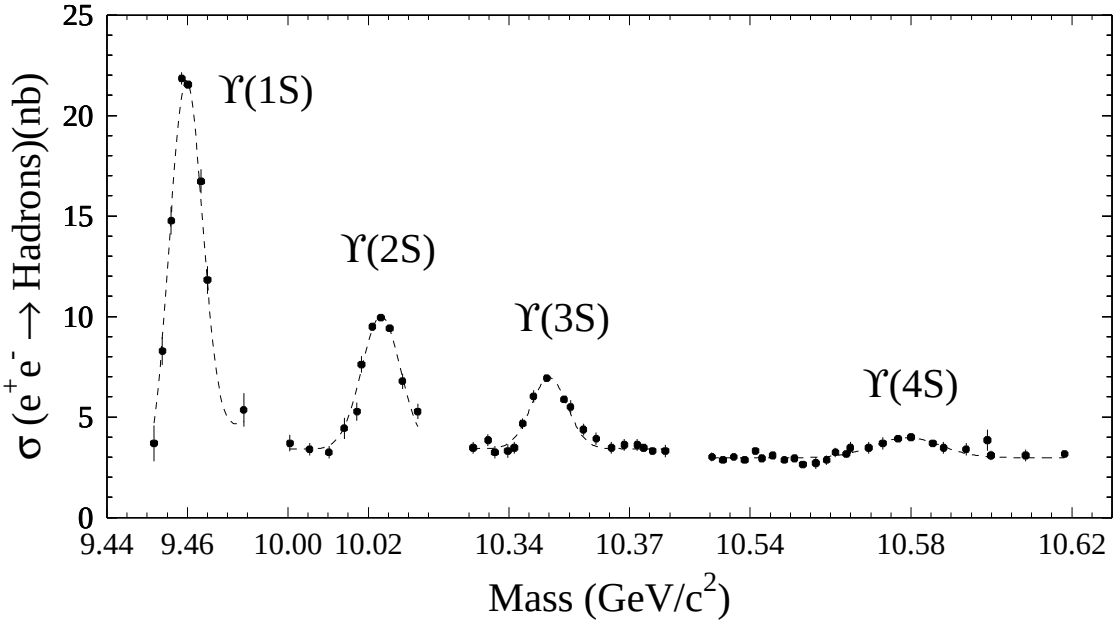

Figure 3.1: The hadron cross section for e^+e^- annihilation around 10 GeV/c^2. The picture represents the measurements by the CLEO collaboration [19].

3.2 Accelerator

3.2.1 Overview

The KEKB accelerator is an advanced electron-positron collider. It has two features separating it from earlier designs — very high luminosity and asymmetric beam energies. The accelerator operates with electron and positron beam energies of 7.996 and 3.500 GeV/c^2, respectively, and a luminosity exceeding 1×10^{34} cm^{-2} s^{-1}. This luminosity is large enough to allow KEKB to claim the highest luminosity achieved by any colliding-beam machine. The required energy asymmetry led to a double-ring collider design. The layout of KEKB is depicted in Fig. 3.2. KEKB has only one interaction point (IP), located in the Tsukuba experimental hall, where the electron and positron beams are collided at a finite (horizontal) angle of ± 11 mrad.

A detailed overview of the design and accelerator properties can be found in [20]. Here we will present a short description of the main parameters and their

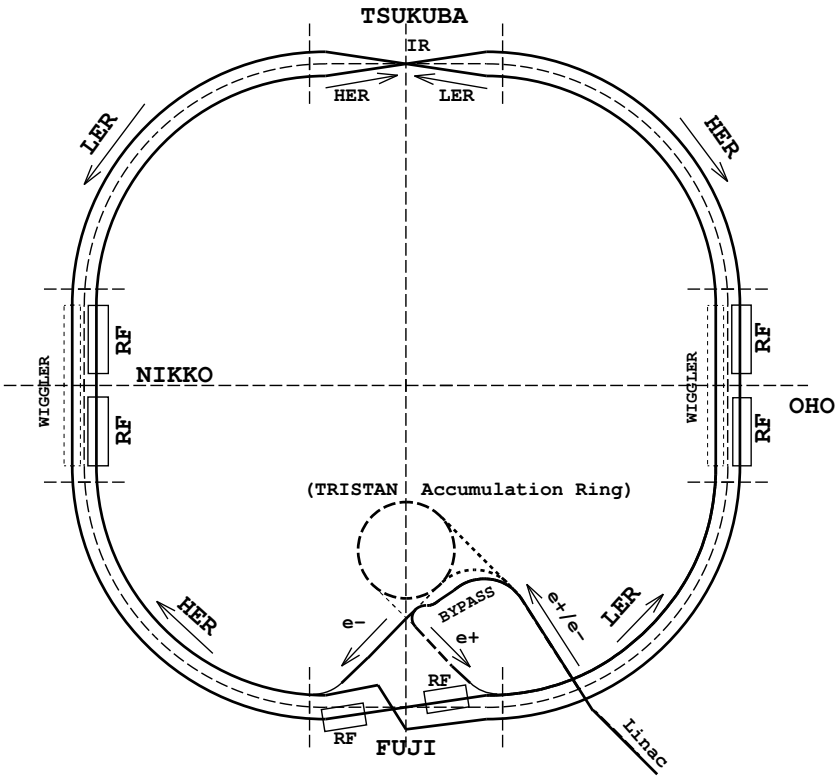

Figure 3.2: KEKB schematic view. Plot is taken from [21].

corresponding effect on the physics we are studying. KEKB was designed to replace the TRISTAN accelerator, and as a result some of the parameters were chosen to reuse the previous equipment so as to minimize cost.

3.2.2 Parameters

Luminosity

One of the most important characteristics of an accelerator is the event rate R. The event rate is proportional to the interaction cross section σ_{int} and a factor called the luminosity.

$R = \sigma_{\text{int}}\mathcal{L}$

If two bunches containing n_1 and n_2 particles collide with frequency f, the luminosity

will be

$$\mathcal{L} = f \frac{n_1 n_2}{4\pi\sigma_x\sigma_y}$$

where σ_x and σ_y characterize the gaussian transverse beam profile in the horizontal and vertical directions.

Crossing angle

Near the IP a rapid two-beam separation is necessary in order to maintain optimized focusing of both the low energy ring (LER) and high-energy-ring (HER) beams and to avoid parasitic collisions. After examining several beam separating options based on bending schemes, finite beam-crossing-angle schemes, and their combination, KEKB settled on a moderately large crossing angle. The absence of separation bend magnets for this solution offers the advantage of reduced synchrotron radiation and simplified IR design. That scheme also allows final focusing with superconducting quadrupole magnets with reasonable inner aperture sizes and offers a flexible configuration that permits a wide range of bunch intensity *vs.* bunch spacing for various center-of-mass energies.

A crossing angle of 22 mrad is close to the minimum that allows elimination of the IP separation bend magnets. It is also nearly the maximum crossing angle that still permits the final focusing of both beams at the IP with common quadrupole magnets.

Bunch parameters

The electron and positron currents in the accelerator are not a continuous, but rather consist of "bunches," which are packs of electrons (positrons) separated by empty space. The bunch spacing is determined by the RF frequency of $f = 508$ MHz. The choice of this frequency was dictated by the desire to reuse RF resources from TRISTAN (a previous accelerator at the KEKB site), so as to reduce cost. Therefore, the allowed bunch spacing is an integer multiple of $c/f = 0.59$ m. The high total luminosity is achieved by using a multibunch operation mode. This reduces the

requirements on single bunch current and the single bunch luminosity, making them easier to realize with established technology.

One of the factors limiting beam stability arises from ions generated in beam-gas collisions. These ions can later be trapped by the negative electric field of the electron beam over multiple beam passes. The standard procedure for reducing this effect is to leave a gap in the bunch train, which causes the trapped ions be driven away by mutual repulsion. This does not remove the ions completely, but it does spread them enough to create the desired beam stability.

Synchrotron radiation from positron bunches creates photoelectrons at the beam walls. These photoelectrons are pulled toward the positively charged positron bunch. When they hit the opposite wall, they generate secondary electrons which can in turn be accelerated by the next bunch. Depending on surface reflectivity, photoelectron and secondary electron yield, this mechanism can lead to the fast build up of an electron cloud and adversely affect the beam stability .

Original design of KEKB underestimated electron-cloud instability and main consideration was given to beam-gas interaction. The LER has higher current and as such will create more ionization. In order to reduce the ion trapping positrons were chosen for the LER. In the HER at least 100 RF buckets have to be left vacant in order to reduce ion trapping.

Energy asymmetry

KEKB was designed to measure CP violation in the decay of neutral B mesons. Doing so requires measuring the decay vertices of the B's. In a symmetric machine, electrons and positrons having equal and opposite momenta are brought into collision at an energy that is tuned to the $\Upsilon(4S)$ resonance. The $\Upsilon(4S)$ subsequently decays to a B-$\bar{B}$ pair. The B mesons from the decays of $\Upsilon(4S)$ have momenta below 0.45 GeV/c, thus the mean flight distance for the B meson in the $\Upsilon(4S)$ rest frame will be $\beta\gamma \cdot c\tau \sim 0.06 \cdot 480\ \mu\text{m} = 30\ \mu\text{m}$. Such a short travel distance is smaller than the typical detector

resolution ($\sim 100\ \mu$m), making a meaningful measurement a practical impossibility. For the B's to have a measurable decay length, they must be boosted in the lab.

An e^+e^- collider with an asymmetric energy collision will do the job. The larger the energy asymmetry, the better the precision of the time-evolution measurement will be. Unfortunately, high values of the asymmetry result in lower reconstruction efficiency, because of the decreased detector acceptance (more products will be boosted in the direction of the beam pipeline). MC simulations have shown that for $\beta\gamma > 0.4$ the sensitivity to the CP violating parameters reaches a broad plateau, and thus the value of 0.42 was chosen. PEP-II uses 0.55, which gives about the same performance.

Energy range

KEKB was designed for B meson production around the $\Upsilon(4S)$ resonance. However, some data has to be collected at other center-of-mass energies. An energy scan is required to identify the peak position of the $\Upsilon(4S)$ resonance. In addition, knowledge of the background composition from non $b\bar{b}$ final states is crucial for many measurements. At 10.5 GeV/c^2 the cross section for $B\bar{B}$ production is negligibly small so this energy was chosen to be a lower bound for the accelerator design. Decays of B_s mesons can provide additional information on b-quark decays. The desire to preserve this option forces the upper limit of the accessible energy range to be above the $\Upsilon(5S)(10.87\ \text{GeV}/c^2)$ resonance, which lies above the $B_s\bar{B}_s$ pair production threshold. It was decided that KEKB should be able to cover the energy between 10.4 and 11.0 GeV/c^2, while keeping the energy asymmetry equal to $\beta\gamma = 0.425$.

Description	**Notation**	**Ring**		**Units**
		LER	**HER**	
Energy	E	3.5	8.0	GeV $/c^2$
Energy spread	σ_ϵ	7.1×10^{-4}	6.7×10^{-4}	
Circumference	C	3016.26		m
Luminosity (Belle CsI)	$\mathcal{L}$	7.25×10^{33}		$cm^{-2}s^{-1}$
Crossing angle	θ_x	± 11		mrad
Beam-beam parameters	ξ_x/ξ_y	0.078/0.049	0.074/0.043	
Beta function at IP	β_x^*/β_y^*	59/0.62	63/0.7	cm
Emittance	ϵ_x/ϵ_y	$1.8 \times 10^{-8}/3.6 \times 10^{-10}$		m
Beam Current	I	1.393	0.869	A
Beam Lifetime		98@1393	276@869	min.@mA
Natural bunch length	σ_z	0.4		cm
Number of bunches		1223		
Bunch spacing	s_b	2.4		m
Time between collisions	τ_{coll}	0.008		μs
Particle/Bunch	N	3.3×10^{10}	1.4×10^{10}	
Synchrotron tune	ν_s	-0.0225	-0.0199	
Betatron tune	ν_x/ν_y	45.513/43.556	45.514/41.580	
Momentum compaction factor	α_p	$1 \times 10^{-4} \sim 2 \times 10^{-4}$		
Energy loss/turn	U_0	0.81/1.5	3.5	MeV/c^2
RF voltage	V_c	6.6	12	MV
RF frequency	f_{RF}	508.887		MHz
Harmonic number	h	5120		
Longitudinal dumping time	τ_ϵ	43/23	23	ms
Total beam power	P_b	2.7/4.5	4.0	MW
Radiation power	P_{SR}	2.1/4.0	3.8	MW
HOM power	P_{HOM}	0.57	0.15	MW
Bending radius	ρ	16.3	104.5	m
Length of bending magnet	l_B	0.915	5.86	m

Table 3.1: Current parameters of KEKB

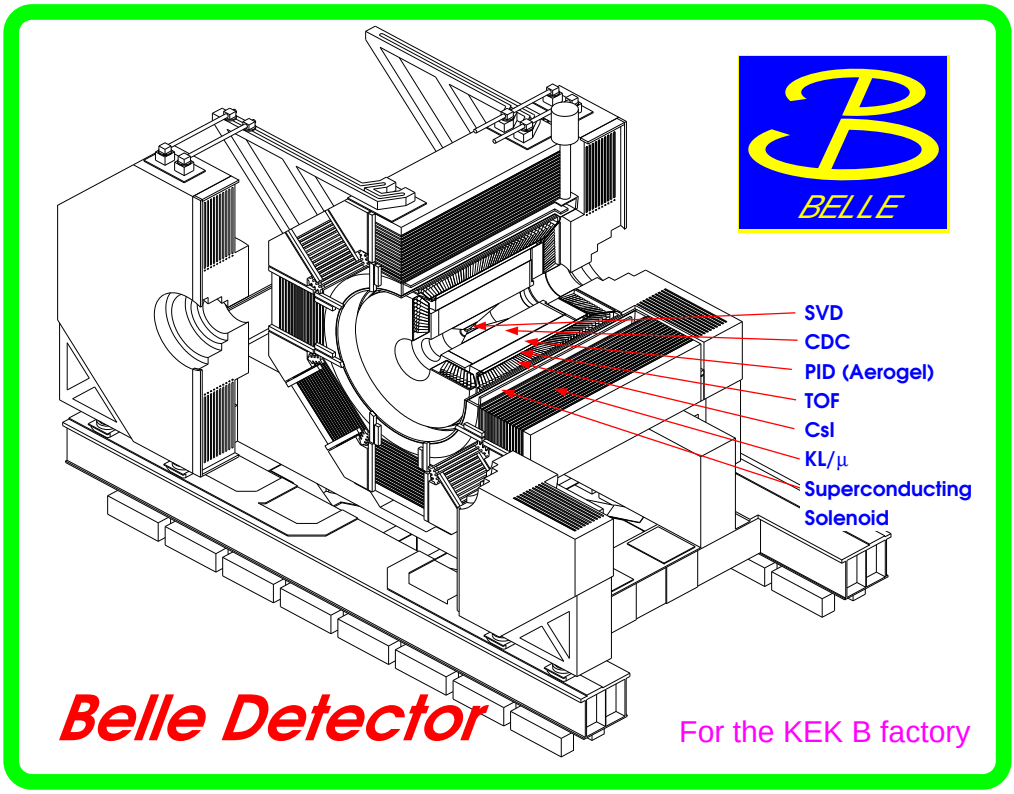

Figure 3.3: Detector view. Plot is taken from [21].

3.3 Detector description

The Belle detector is designed to measure CP violation parameters in neutral B meson decays. That goal requires an accurate determination of the B decay vertex position, and a precision measurement of the momenta and energies of the final state particles. In many cases, particle identification information is also important. The detector is designed to have a large acceptance so as to facilitate reconstruction of modes with many particles in the final state. The decay products have an approximately isotropic angular distribution in the CM frame. The CM boost tends to pitch particles forward in the lab. As a result, the detector has to be asymmetric. The detector tracks charged particles in the range 60 MeV/c $< p_t <$ 4 GeV/c, and provides a positive

identification of e, μ, π, K and p's over a wide kinematic range. The detector (see Fig. 3.3) has an onion like composition with the following major parts:

- Silicon Vertex Detector (SVD) The vertex detector is designed for precision decay vertex position measurement.
- Central Drift Chamber (CDC) The drift chamber is used for charged particle momentum measurements and contributes to the particle identification through dE/dx measurements.
- Aerogel Cherenkov Counter (ACC) The main purpose of the ACC is $K - \pi$ separation in the 1.2 GeV/$c < p <$ 3.5 GeV/c momentum range.
- Time of Flight Counters (TOF) This subdetector is designed to improve $K - \pi$ separation in the $p <$ 1.5 GeV/c momentum range.
- Electromagnetic Calorimeter (ECL) The calorimeter is used for photon energy measurement as well as for electron particle identification.
- Superconducting Solenoid (SS) Produces an approximately uniform magnetic field of 1.5 T in the drift-chamber tracking volume.
- $K_L - \mu$ system (KLM) An iron structure surrounding the solenoid is used to return the flux of the superconducting solenoid and to identify muons and K_L mesons.

The beam pipe is the innermost part of the Belle detector. Its primary purpose is maintaining the vacuum around IP. The beam pipe must dissipate the heat induced by the beam in the form of higher-order-mode losses. The beam pipe must thus satisfy several conflicting criteria. On one hand, it should be strong enough to withstand the pressure difference and the high heat load, while on the other, the amount of material should be kept to a minimum in order to reduce the scattering of the B decay products, since multiple scattering degrades the performance of the innermost

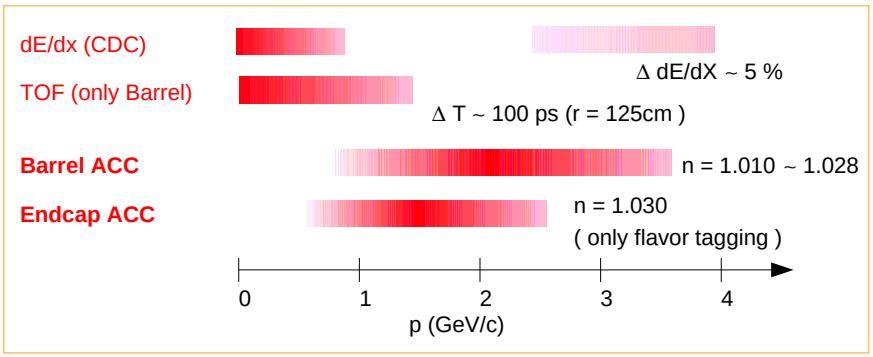

Figure 3.4: PID momentum coverage. Plot is taken from [21].

tracking devices, thus decreasing the vertex resolution. The amount of multiple scattering is proportional to the atomic number Z and square root of the thickness, so it was decided to use a low Z-very strong double-wall 0.5 mm thick beryllium ($Z = 4$) pipe with cooling provided by flow of helium through a 2 mm gap. The total material thickness corresponds to 2.8×10^{-3} radiation lengths.

Particle identification

On average the B decay produces 3.5 charged pions, 0.79 charged kaons, 0.1 protons and 0.1 leptons. In order to distinguish these final states we need a particle identification system.

Electron identification relies on the information from the CDC, ECL and ACC [22]. The threshold for Cherenkov light in the ACC is several MeV/c^2 for electrons and 0.5-1 GeV/c for pions, thus providing good identification for electrons with momentum below 1 GeV/c. The ionizing energy loss in the drift chamber, which exhibits the well known β^{-2} behavior, provides another tool for electron identification. However, the most significant discriminants are provided by the ECL system, which exploits the major differences in the properties of showers induced by hadrons, muons and electrons. Electrons with momentum p deposit all of their energy, $E \approx p$, in the calorimeter, while other particles have higher penetration power and are not fully

contained by the ECL. The energy loss from the CDC, the light yield in the ACC, the shower shape, the value of E/p and the distance between the shower center and the track projection from the drift chamber to the calorimeter are combined to form an electron likelihood, which allows us to select electrons with an efficiency of $92.4 \pm 0.4\%$ with a π fake rate $0.25 \pm 0.02\%$.

Muon identification is based on the high penetration power of muons and relies on the CDC, the ECL, and the KLM [23]. Good muon identification is possible in the momentum range above 1 GeV/c. The penetration depth in the KLM detector in combination with the energy deposition in the CsI and the goodness of fit for tracks both in both the CDC and the KLM allows us to select muons with 93% efficiency and with less than 3% K/π contamination.

Information from the CDC, the ECL, and the KLM is sufficient to identify most particles. However, information from these systems can not distinguish kaons from pions with momenta greater than 0.7 GeV/c. Time of flight and Cherenkov counters are used to cover this region (Fig. 3.4).

The identification of kaons is based on information coming from the ACC, TOF and the dE/dx measurement from the CDC. These measurements are nearly independent, with good separation over a wide range of momenta. Energy loss in the central drift chamber provides $\geqslant 3\sigma$ π/K separation for tracks with momentum less than 0.8 GeV/c and 2σ separation for momenta greater than 2.0 GeV/c. Time of flight measurements provide π/K separation of better than 3σ for momentum less than 1.2 GeV/c.

3.3.1 Magnet

A magnetic field of 1.5 T is created by the main superconducting solenoid. This field covers the volume of the drift chamber and can affect the beam dynamics and result in the reduced beam lifetime (see [21]). To avoid this, the integrated field along

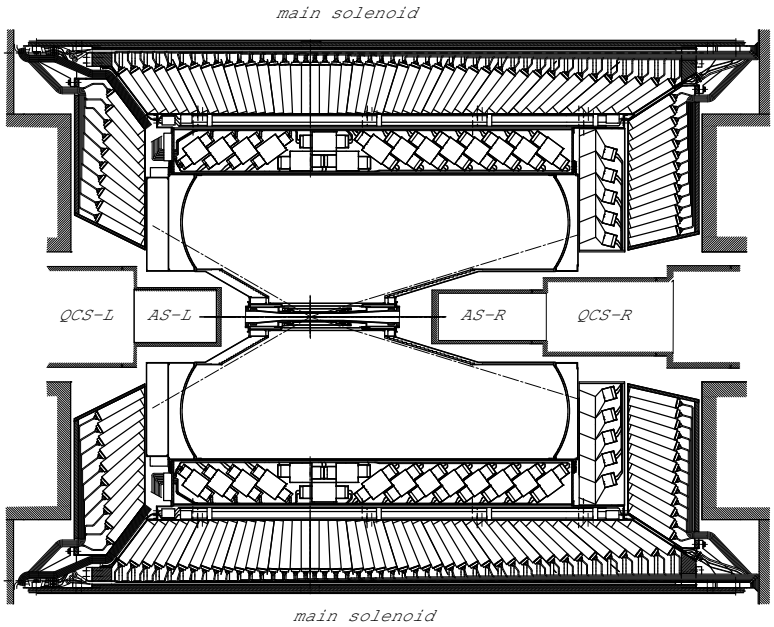

Figure 3.5: Magnets around interaction point. The tracking field is produced by the main solenoid, while AS-L(R) magnets added for beam stability and QCS-L(R) for final focusing. Plot is taken from [21].

the beam path is locally cancelled by a pair of additional superconducting solenoids (S-R and S-L) with a smaller radius placed right outside the silicon vertex detector. Another two quadrupole superconducting magnets (QS-L and QS-R) are used for the final vertical beam focusing.

3.3.2 SVD

The silicon vertex detector (SVD) plays an essential role in the Belle experiment. The primary physics goal of Belle is to observe CP violation in the proper-time distributions, Δt, of neutral B-meson decays. The proper time is reconstructed from the distance between decay vertices of the two B mesons. Because the two B's move slowly in the $\Upsilon(4S)$ center-of-mass (CM) frame ($\beta\gamma = 0.06$), their lab velocities are defined by the beam boost and are nearly equal. As a result, the displacement between their decay vertices in the z direction is proportional to Δt. The KEKB

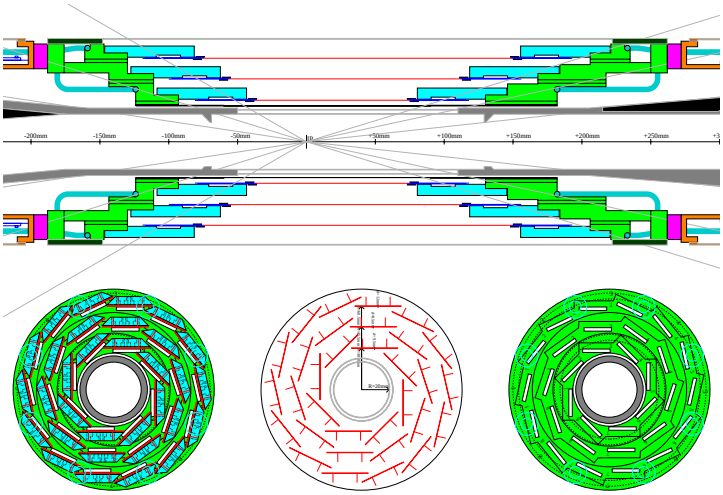

Figure 3.6: Silicon Vertex detector, side view. Plot is taken from [21].

Lorentz factor, $\beta\gamma$, is 0.425, corresponding to an average B meason travel distance of ≈ 200 μm. This value determines the resolution requirement for the vertex detector. [24, 25].

The SVD consists of three concentric approximately cylindrical layers of double sided silicon strip detectors (DSSD). Fig. 3.6 shows the configuration of the SVD. Each layer consists of DSSD ladders that are made of two half ladders and boron-nitride support ribs that reinforce the structure. The radii of the three layers are 30, 45.5 and 60.5 mm. The angular coverage is $23° < \theta < 140°$ which corresponds to 86% of the full solid angle in the CM frame, where θ is the angle from the beam axis. The total material of the SVD at $\theta = 90°$ corresponds to 2.6% of a radiation length, including the beryllium beampipe.

3.3.3 CDC

The efficient reconstruction of charged particle tracks and the precise determination of their momenta is essential to virtually all of the measurements planned for the Belle

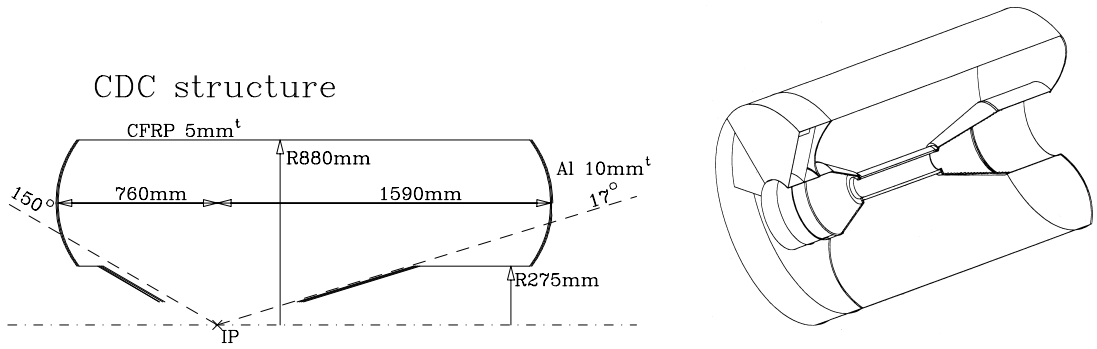

Figure 3.7: Central Drift Chamber. Pictures are taken from [21].

experiment. Specifically, the physics goals of the experiment require a momentum resolution of $\sigma_{p_t}/p_t \sim 0.5\%\sqrt{1+p_t^2}$ (p_t in GeV/c) for all charged particles with $p_t > 100$ MeV/c in the wide polar angle region. In addition, the charged particle tracking system is expected to provide important information for the trigger system and particle identification information in the form of rates of ionization energy loss, dE/dx [21, 26, 27].

The structure of the drift chamber is shown in Fig. 3.7. It has an asymmetric shape, reflecting the CM boost. The angular acceptance is $17° \leq \theta \leq 150°$. The inner radius is 103.5 mm and the outer radius is 874 mm. The small inner radius of the drift chamber allows good tracking efficiency for low p_t tracks, which is important for reconstruction of slow pions in $D^* \to D\pi$ transitions.

The low-Z gas is chosen to reduce multiple coulomb scattering contribution to the momentum resolution. Low-Z gases also have smaller photo-electric cross section, which results in reduced background from synchrotron radiation and spent particles. CDC uses 50% He-50%C_2H_6 mixture. That mixture has a long radiation length ($\sim$ 640 m), its drift velocity saturates at $\sim$ 4cm/μsec with $\sim 2kV$/cmelectric field. That makes drift time less sensitive to the HV and simplify calibration. Good dE/dx resolution provided by large ethane component.

Expected resolution is determined by two components, one due to the measurement error and a second due to multiple coulomb scattering [27].

$$\frac{\sigma_{P_t}}{P_t} = \sqrt{\left(\frac{\sigma_{P_t}}{P_t}\right)^2_{\text{meas}} + \left(\frac{\sigma_{P_t}}{P_t}\right)^2_{\text{MS}}} \tag{3.1}$$

where P_t is transverse momentum in GeV/c with respect to the direction of magnetic field. With components presented below

$$\left(\frac{\sigma_{P_t}}{P_t}\right)_{\text{meas}} = \frac{P_t \sigma_{r\phi}}{0.3 L^2 B}\sqrt{\frac{720}{N+4}} \tag{3.2}$$

$$\left(\frac{\sigma_{P_t}}{P_t}\right)_{\text{MS}} = \frac{0.05}{LB\beta}\sqrt{1.43\frac{L}{X_0}}\left[1 + 0.038 \ln \frac{L}{X_0}\right] \tag{3.3}$$

where β is velocity of the charged particle, B is the magnetic field in Tesla, L is CDC radius in meters, $\sigma_{r\phi}$ is the spatial resolution in meters for each measured point, X_0 is the radiation length in meters and N is the number of measurement points. The Belle CDC P_t resolution for muons is expected to be

$$\sqrt{(0.17\% P_t)^2 + \left(\frac{0.22\%}{\beta}\right)^2} \tag{3.4}$$

Energy loss in the drift chamber is used for charged particle identification and is given by Bethe-Bloch formula[28, 29, 30, 31]

$$-\frac{dE}{dx} = 4\pi N_A r_e^2 m_c z^2 \frac{Z}{A\beta^2}\left[\frac{1}{2}\ln\left(\frac{2m_e\gamma^2\beta^2 T_{\max}}{I}\right) - \beta^2 - \frac{\delta}{2}\right] \tag{3.5}$$

where N_A is Avogadro's number, r_e classical electron radius $e^2/4\pi\epsilon_0 c^2$, M is incident particles mass, m_e is electron mass, ze is charge of incident particle, I is mean excitation energy and δ is density effect correction to ionization energy loss. $T_{\max}$ is maximum kinetic energy which can be imparted to a free electron in a single collision. For a point like charged particle with mass M and momentum $M\beta\gamma$, $T_{\max}$ is given by

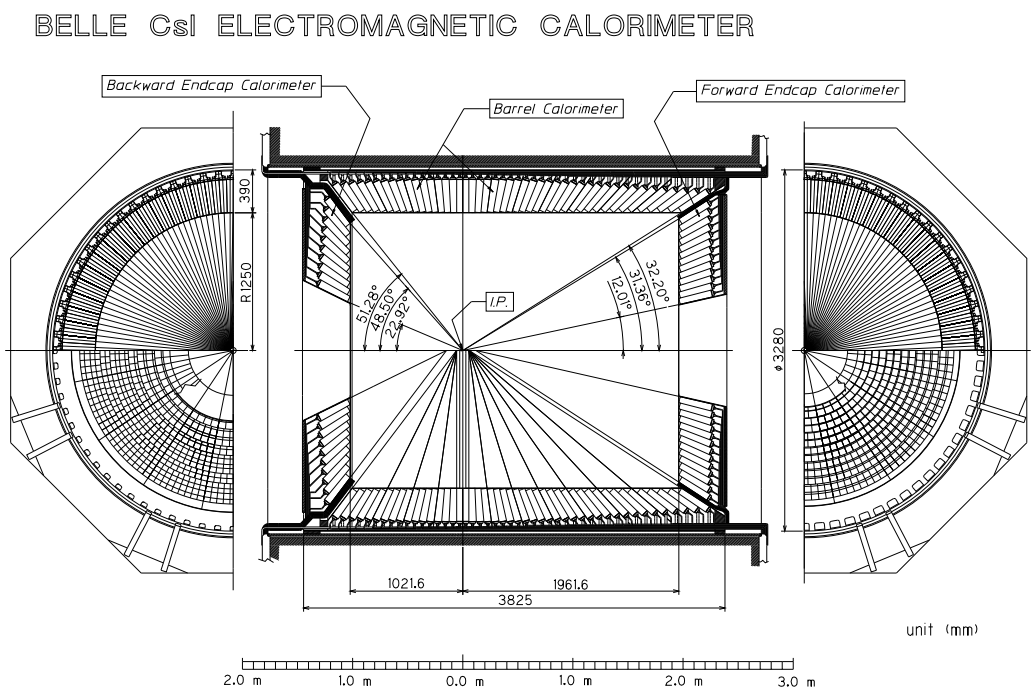

Figure 3.8: ECL dimensions. Plot is taken from [21].

$$T_{\max} = \frac{2m_e\beta^2\gamma^2}{1 + 2\gamma\frac{m_e}{M} + \left(\frac{m_e}{M}\right)^2} \tag{3.6}$$

3.3.4 ECL

The main purpose of the electromagnetic calorimeter is the detection of photons from B-meson decays. This must be done with high efficiency and good resolution in energy and position. Since most photons are the end products of cascade decays, they have relatively low energies, making good performance below 500 MeV/c^2 especially important. On average, $\Upsilon(4S)$ decays yield 10 photons, roughly half having energies below 200 MeV/c^2. High-momentum π^0 detection requires the separation of two nearby photons and a precise determination of their opening angle. This requires a fine-grained segmentation in the calorimeter. In order to satisfy the above requirements, the Belle collaboration has decided to use a highly segmented array of CsI(Tl) crystals with silicon photodiode readout installed in a magnetic field of 1.5 T inside

a superconducting solenoid magnet [32, 33, 21, 34, 35]. Low-energy π^0's reconstruction is especially challenging, as reconstruction efficiency falls fast with minimum detectable energy for photons. Thus a 15 MeV/c^2 threshold for photons cuts around 10% of the π^0. Important two-body decay modes such as $B \to K^*\gamma$ and $B \to \pi^0\pi^0$ produce photon energies up to 4 GeV/c^2 and good resolution at high energy is needed to reduce backgrounds for these modes.

Electron identification in Belle relies primarily on a comparison of the charged particle momentum and the energy deposits in the electromagnetic calorimeter. The good electromagnetic energy resolution results in better hadron rejection. Proper cluster position is also important to electron identification as it is crucial for good track-cluster matching efficiency.

The overall layout of the calorimeter is shown in Fig. 3.8. The calorimeter is comprised of three parts, a barrel and two endcaps, covering azimuthal angles from 17° to 150°. The barrel has an inner radius of 1.25 m and a length of 3 m. The endcaps are placed at $z = +2$m and $z = -1$m from the interaction point. A total of 8,736 crystals is assembled in the structure, each pointing towards the IP with a small nonprojective tilt ($\sim 1°$) to prevent photons from escaping through the gaps between them. The ECL detector covers 91% of 4π in the laboratory frame.

3.3.5 EFC

In order to improve the experiment's sensitivity to two-photon physics processes, we need to further extend the polar angle coverage of the ECL. The extreme forward calorimeter, EFC, covers the angular range from 6.2° to 11.6° in the forward direction and 163.1° to 171.5° in the backward direction. Each side consists of 160 crystals arranged in 5 rows in θ and 32 segments in the ϕ direction. The typical crystal size is 2 cm × 2 cm in cross section and 12.0 X_0 (10.5 X_0) in length (X_0 refers to radiation length) in the forward (backward) direction. The choice of Bismuth Germanate

($Bi_4Ge_3O_{12}$) crystal was dictated by the harsh radiation conditions around the IP. In addition to radiation hardness BGO crystals have excellent e/γ energy resolution, high density and fast decay of scintillation light ($\sim$ 300 ns).

The EFC detector is attached to the front faces of the cryostats of the compensation solenoid magnets of the KEKB accelerator. The location of the extreme forward calorimeter makes it usable as a beam mask to reduce backgrounds for the CDC

In addition, the EFC serves as a beam monitor for KEKB and a luminosity monitor for the Belle experiment. It can also be used as a tagging device for two-photon physics.

Radiation tests of the crystal at Inst. of Inorganic Chemistry (Novosibirsk, Russia) showed high yields even after a 10 MRad radiation dose.

3.3.6 ACC

Particle identification, specifically the ability to distinguish charged kaons from pions, plays a key role in the elucidation of CP violation. An array of silica aerogel threshold Cherenkov counters has been selected as part of the Belle particle identification (PID) system to extend the PID momentum coverage beyond the reach of CDC energy loss measurements (dE/dx) and time-of-flight measurements. The ACC, being a threshold device, generates Cherenkov light for pions, while kaons and protons pass through undetected [36].

Cherenkov light is generated when a charged particle traverses a medium at speed higher than the speed of light in that medium (i.e., $v > c/n$). In analogy with a supersonic shock wave, the Cherenkov light is emitted in a cone around the particle's flight direction having angle θ_c given by

$$\cos\theta_c = \frac{1}{n\beta} \tag{3.7}$$

The threshold momentum is given by

$$p_{\text{thr}} = \frac{m}{\sqrt{n^2-1}} \tag{3.8}$$

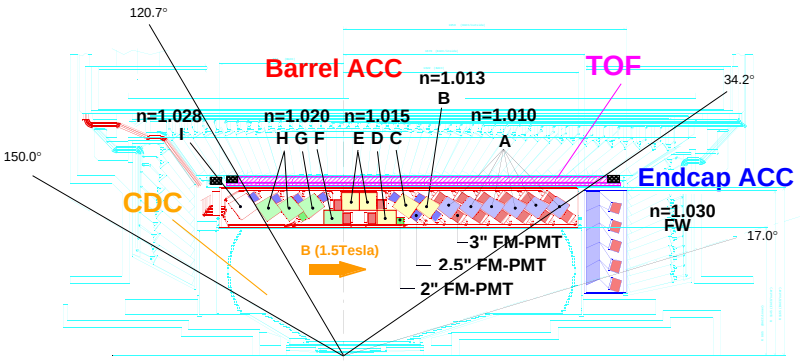

Figure 3.9: ACC and TOF. Plot is taken from [21].

The refractive index of the aerogel was selected to achieve the best K-π separation in the 1.2-3.5 GeV/c momentum range. Five different refraction indices were used, n=1.01, 1.013, 1.015, 1.020 and 1.028, in order to correct for the higher momentum of B meson products in the lab frame for smaller polar angles θ that arises as a consequence of the CM boost.

$$N = 2\pi\alpha L \int_{\lambda_1}^{\lambda_2} \frac{\sin^2\theta_c}{\lambda^2} d\lambda \tag{3.9}$$

Integration over the range from 7.3 eV to 1.5 eV results in the following photon density [37]

$$\frac{dN}{dEdx} = 370 L \sin^2\theta_c \text{ eV}^{-1} \text{ cm}^{-1} \tag{3.10}$$

Photon yield is is proportional to $\sim \sin^2\theta_c \sim 1 - 1/n^2 \ \sim 0.04$ and results in only tens of photons per cm for BELLE configuration. At such small yields we have to pay special attention to the transparency of the aerogel. Transmission length (Λ) is defined by the following equation

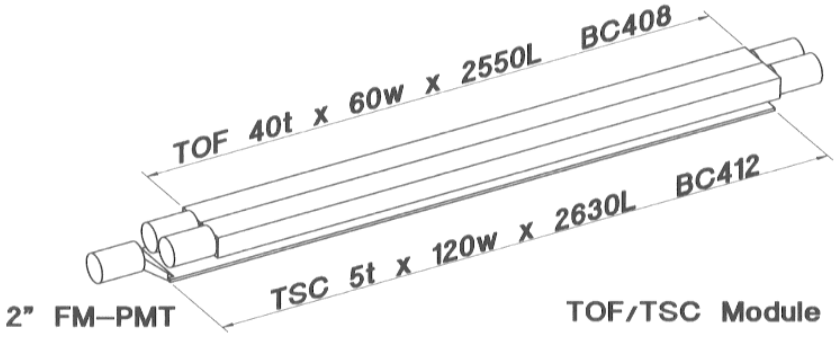

Figure 3.10: Time of flight counters. Plot is taken from [21].

$$T = T_0 exp\left(-\frac{d}{\Lambda}\right) \tag{3.11}$$

where T and d are the light transmission and the thickness of aerogel respectively. Transmission length of aerogel used in ACC is monotonously increasing function of wavelength with $T \approx 0.6$cm for $\lambda = 300$ nm, $T \approx 3.0$cm for $\lambda = 440$ nm and $T \approx 7.5$cm for $\lambda = 550$ nm. Aerogel blocks have side dimensions of 12 cm and provide enough material to detect between 10 and 30 photons per particle above threshold, which is enough for desirable K/π separation in required momentum range.

3.3.7 TOF

The time-of-flight (TOF) system is based on plastic scintillation counters located behind the aerogel system [38]. Precise time-of-flight measurement allows particle-mass estimation from the following equation

$$M^2 = \left(\frac{1}{\beta^2} - 1\right) p^2 = \left[\left(\frac{cT}{L}\right)^2 - 1\right] p^2 \tag{3.12}$$

where T is the measured time of flight, p particle momentum measured in the drift chamber and L is the path length of the particle determined from the CDC track fit. For a 1.2 m flight path, a TOF system with 100 ps time resolution has 2σ separation

for $p - K - \pi$ particles with momenta up to 1.25 GeV/c, which covers 90% of the particles produced in the $\Upsilon(4S)$ decays.

In addition to particle identification, the TOF counters generate fast timing signals for the trigger system.

3.3.8 KLM

The K_L/μ system was designed to identify K_L's and muons with high efficiency over a broad momentum range, 0.6-5 GeV/c. The barrel-shaped region around the interaction point covers an angular range from 45° to 125° in polar angle and the endcaps in the forward and backward directions extend this range to 20° and 155°.

The K_L/μ subdetector consists of alternating layers of 3.7-cm thick glass-electrode resistive plate chamber (RPC) modules and 4.7-cm-thick iron plates. There are 15 detector layers and 14 iron layers in the octagonal barrel region and 14 detector layers and 14 iron layers in each of the endcaps. The iron plates provide a total of 3.9 interaction lengths of material for a particle travelling normal to the detector planes. In addition the electromagnetic calorimeter (CsI) provides another 0.8 interaction length of material to convert K_L's.

Each RPC module consists of two independent RPCs arranged back to back, sandwiched between orthogonal readout strip panels. This redundant superlayer design provides a three-dimensional measurement of the coordinates of a thoroughgoing charged track with better than 98% efficiency, since a streamer that develops in either RPC will induce an image charge on both readout strip planes.

Each RPC is composed of two glass plates separated by 1.9 mm thick, 1.9 mm wide Noryl spacers, which are epoxied to the glass plates and positioned so as to form a 12-cm-wide gas channel that guides the gas through the chambers in a "mouse-maze" pattern. The high voltage is supplied and distributed on the outer surface of the glass plates which have being coated with Kohi-noor 3080 F india ink (or

graphite impregnated plastic film for endcap). The surface resistance of the carbon coating is $0.5 - 5M\Omega/m^2$ $(0.2 - 0.7M\Omega/m^2)$. RPCs are operated in the streamer mode. Applied high voltage is fixed to +4.7(+4.5) kV for the anode plane of the barrel (endcap) and -3.5 kV for the cathode planes for barrel and endcap RPCs. The gas mixture is chosen to be a 30% argon, 62% HFC134a and 8% isobutane, which is environmentally friendly, nonflammable and provides high detection efficiency and stable RPC operation. System parameters were chosen to achieve time resolution in the order of a few nanoseconds with operation rate of $\sim 100\frac{Hz}{m^2}$. The KLM system has hit spatial resolution of 1.2 cm.

For more information on the KLM system see [39, 40] and [41].

History of RPCs

Resistive plate chambers (counters) were first developed in the early 1970s, with the primary design goals of good spatial and time resolution. Work of a group from Novosibirsk [42] showed the possibility of creating detectors with spatial resolution 0.2 mm and time resolution of 30 ps. But this was not the cheap option. Pestov's chambers were operated at a high pressure of 12 atm, had a small gap of 0.1 mm and used a semiconducting glass with low resistivity ($10^9 \sim 10^{10}\ \Omega \cdot \mathrm{cm}$).

At the time, science in the Soviet Union was well funded, so price was not an issue. But still, technological difficulties of producing large surfaces for long operation under high pressure and stringent mechanical requirements associated with $100\mu m$ uniform gaps limited the chamber size to be $30 \times 30\ cm^2$.

In the 1980s the desire to build bigger detectors and less favorable funding prompted an Italian group to scale down the ambitions in terms of resolution. At first, they opted for RPCs made of bakelite ($\rho \cong 10^{11}\ \Omega \cdot \mathrm{cm}$) operated under atmospheric pressure and with a gap about 2 mm. They have reported [43] time resolutions of order of 1 ns, pulses of about 100 ns duration and rise times of about 5 ns. Later the same

group reported the possibility of using RPCs made of glass for low-energy electromagnetic calorimetry [44] and for charged-particle counting in high rate environments ($10 Hz/cm^2$) [45].

In the 1990s the B-factories were born, with need of good muon detection in a low background and event-rate environment. In order to achieve these goals large-area detectors were needed with time accuracy of the order of 10 ns and spatial resolution around 1 cm. Although initial expectations were that satisfying these requirements would be straightforward ("a piece of cake"), unanticipated technological problems were encountered along the way. The biggest obstacle encountered in both Belle and BaBar experiment was surface erosion caused by the water vapor and subsequent efficiency degradation. Another problem was caused by "avalanche" discharge of entire chamber. Fast discharge $\sim$ 10 ns of entire chamber created huge currents flowing into readout electronics, with rather devastating effects. The latest problem encountered by Belle is efficiency loss of the outer endcap layers, due to the unexpected high rate background caused by the neutrons coming from the beam. After overcoming all these obstacles Belle KLM system grew into stable and reliable detector, that met all the design goals.

Chapter 4

Event selection.

4.1 Introduction

This chapter describes the selection of $B \to J/\psi\ K_L$ decays. We use a reconstruction approach similar to previous work [46, 47, 48, 49, 50, 51]: we start from inclusive J/ψ selection, continue with identification of the J/ψ and K_L pair and finish by vetoing the most prominent backgrounds. The decay modes considered as background are listed in Table 4.1 (Hereafter the inclusion of the charge conjugate states is implied). The main difference from previously reported results lies in the extensive use of a probabilistic (likelihood) approach for signal quality assessment. That allows us to either increase the reconstruction efficiency or to improve the signal-to-background ratio. We have developed special reconstruction techniques for all particles based on the likelihood approach and must therefore describe in detail the likelihood construction for all basic particles. In order to maintain consistency, likelihood functions for all the particles mentioned in Table 4.1 were normalized to the signal over background ratio S/B, predicted by MC.

Assigning a likelihood to every process under consideration also allows us to create an efficient structure for background suppression. For events with multiple candidates, we can simply select the one with the highest likelihood value. In addition, for all

particles participating in complex cascade decays we store the value for the highest parent likelihood (HPL). As an example of HPL construction, we assume the hypothetical situation where particle C is reconstructed to be a part of two overlapping decay chains $A_1 \to X_1 B_1$ with $B_1 \to Y_1 C$ and $A_2 \to X_2 B_2$ with $B_2 \to Y_2 C$. The HPL for particle C will be the highest likelihood among $L(A_1), L(B_1), L(A_2), L(B_2)$. In order to illustrate the usage of the HPL we describe the possible veto of photons coming from the $\pi^0 \to \gamma\gamma$ transition in $\chi_c \to J/\psi\ \gamma$ selection. The simplest filter rejects photons with an HPL greater than the χ_c likelihood. In the subsequent analysis we include the HPL in the χ_c likelihood calculation for better efficiency. We also use HPL in selecting the $B \to J/\psi\ K_L$ sample, which has rather high background contamination from various decay channels. This forces us to pay special attention to the selection of the inclusive J/ψ sample and to veto as many J/ψ exclusive modes as possible. The full list of exclusive modes is presented in Table 4.1.

In order to manage combinatorial backgrounds, we perform event reconstruction in two steps. In the first, we estimate the probability of all fully and partially reconstructed modes and store the highest likelihood of the B meson decay through any daughter particles. For example, in some events we might have three decay candidates D_1, D_2 and D_3, where $D_1 = B \to J/\psi\ K^+$, $D_2 = B \to J/\psi\ K^{*0}(K^+\pi^-)$ and $D_3 = B \to J/\psi\ \pi^-$ with estimated likelihoods $L(D_1) > L(D_2) > L(D_3)$. In this configuration we store the following numbers as a HPL L^P: $L^P_{J/\psi} = L(D_1)$, $L^P_{K^+} = L(D_1)$, $L^P_{\pi^-} = L(D_2)$. The HPL approach allows us to suppress backgrounds caused by "clean" decay channels and to manage multiple decay channels in a uniform way. In our analysis, we use the HPL as an additional parameter in the likelihood function for the $B \to J/\psi\ K_L$ decay.

Previously reported results [46, 47] and estimates presented later in this text show that backgrounds related to non J/ψ modes are small for J/ψ exclusive decays and can be ignored at the event selection stage. Special treatment is, however, reserved for such contamination at the final stages of analysis. The small branching fraction

Primary mode	Secondary mode(s)	type
$B \to \psi(2S)X$	$\psi(2S) \to J/\psi\ \pi^+\pi$, $\psi(2S) \to l^+l^-$	inclusive
$B \to \chi_{c1,\ c2}(1P)X$	$\chi_{c1,\ c2}(1P) \to J/\psi\ \gamma$	inclusive
$B^- \to J/\psi\ \pi^-$		full
$B^0 \to J/\psi\ \pi^0$		full
$B^- \to J/\psi\ K^-$		full
$B^0 \to J/\psi\ K_S$	$K_S \to \pi^+\pi^-$, $K_S \to \pi^0\pi^0$	full
$B^0 \to J/\psi\ K^{*0}$	$K^{*0} \to K^+\pi^-$, $K^{*0} \to K_L\pi^0$, $K^{*0} \to K_S\pi^0$	full
$B^- \to J/\psi\ K^{*-}$	$K^{*-} \to K^-\pi^0$, $K^{*-} \to K_L\pi^-$, $K^{*0} \to K_S\pi^0$	full
$B^0 \to J/\psi\ K_S$	$J/\psi\ \pi^-$, $J/\psi\ \pi^0$	partial
$B \to J/\psi\ K^*$	$J/\psi\ K^-$, $J/\psi\ K_S$, $J/\psi\ \pi^-$, $J/\psi\ \pi^0$	partial

Table 4.1: Decay chains used as veto in $B^0 \to J/\psi\ K_L$ selection. Type describes the type of the veto decay. It could be inclusive or exclusive with a partially or fully reconstructed B meson. For partially reconstructed modes, the secondary mode column provides a list of particles used in partial reconstruction. Unless otherwise noted all the decays in the primary and secondary modes proceed through the following transitions $J/\psi \to l^+l^-$, $K_S \to \pi^+\ \pi^-$ and $\pi^0 \to \gamma\ \gamma$

for J/ψ production from $B\bar{B}$ pairs ($BR(B \to J/\psi_{(l^+l^-)}\ X) = (1.357 \pm 0.009) \times 10^{-3}$) makes the use of generic MC prohibitively expensive. However, the reasonably small contamination from non J/ψ modes allows us to tune selection criteria on the inclusive J/ψ MC sample ($B \to J/\psi_{(l^+l^-)}\ X$). In order to create a robust test bed for J/ψ backgrounds, the J/ψ likelihood function is tuned on the data sample. That makes the construction of the J/ψ likelihood quite different from that for the rest of the particles involved. We therefore describe it right after the introduction to the likelihood. The likelihoods for all the other particles are based on the J/ψ inclusive MC and are described right after the $c\bar{c}$ section.

4.2 Monte Carlo generation.

The bulk of this analysis is tuned on the J/ψ inclusive MC sample. Particles are generated using the EvtGen software package developed by the BaBar collaboration [52], while the detector response is simulated via the GEANT3 software package

from CERN [53]. The decay tables were modified to match the up-to-date branching fractions for B meson decays. We did, however, carry out the generation with $\mathrm{BR}(\Upsilon(4S) \rightarrow B^0\bar{B}^0) = \mathrm{BR}(\Upsilon(4S) \rightarrow B^+B^-)$. This creates an excess of neutral B meson decays, but does not affect the presented results, since we estimate signal and background yields from the corresponding shapes and do not rely on branching fractions provided by MC.

In order to create inclusive J/ψ sample, we generate generic $e^+e^- \rightarrow \Upsilon(4S) \rightarrow B\bar{B}$ decays, but only decays containing J/ψ mesons are saved. Furthermore, we force the J/ψ to decay into two leptons. In order to facilitate the CP analysis we have generated six independent samples with $\sin 2\phi_1$ values $\{0.1, 0.3, 0.5, 0.7, 0.9, 1.0\}$. The number of events in the entire sample is 5.5 million.

4.3 Likelihood

In cases where there are many parameters, each providing some information, an efficient selection requires a probability-based approach. The probability or likelihood approach can be viewed as a step forward in the evolution of signal-background separation techniques. In a conventional cut-based analysis "cut values" or simply "cuts," which provide only a binary answer, are used to separate signal from background. A cut is defined as an allowed range for a certain parameter of a physical process. A common example is the mass cut, $M_{\mathrm{min}} < M < M_{\mathrm{max}}$, which is routinely used for resonance selection.

This technique will work well if there is a complete or nearly complete separation between signal and background for the parameters to which the cuts are applied. In many cases, however, the distributions for signal and background events, although significantly different, will have a substantial overlap. As a result the binary answers provided by hard cuts will cost both efficiency and overall performance. In a likelihood approach, the signal is assigned a probability with respect to each underlying

parameter, so that an event falling in the tail of one distribution can be saved if it falls in the preferred region for signal in the other parameters.

We will use the following nomenclature in the subsequent text.

- The *likelihood* is a function of the event parameters, which is designed to separate the signal from the background and to provide a robust estimate of the signal quality.

- A *likelihood parameter* is any variable used in the likelihood construction. In the case of correlated parameters, we will try to create likelihood discriminants based on several likelihood parameters.

- A *likelihood discriminant* is a likelihood related to any number of likelihood parameters that is used as an atom for the construction of the next level likelihood. We treat the discriminants as uncorrelated variables.

For multi-parameter likelihood constructions, special care must be taken to account for possible parameter correlations. The conventional solution uses Fisher's linear discriminant, which is a classification method that projects multi-dimensional data onto a line and performs classification in a one-dimensional space [54, 55, 56]. Here, however, we implement a different technique. We identify the most closely correlated variables and pair them into two-dimensional likelihood discriminants. Proper pair selection is crucial, because leaving correlated variables unpaired reduces the final performance.

In order to illustrate the above definitions and to describe the idea behind the likelihood construction, we will consider a likelihood based on five parameters: P_1, P_2, P_3, P_4, P_5. Furthermore, we assume that four parameters form two pairs of correlated values (P_1, P_2) and (P_3, P_4). For the correlated parameters we fill two-dimensional frequency distributions for both signal, $S(P_i, P_j)$, and background, $B(P_i, P_j)$, for P_5 we fill a one-dimensional distribution. Here we assume that the original sample can be divided

into signal and background parts. In what follows we refer to $S(P_i), S(P_i, P_j), \ldots$ / $B(P_i), B(P_i, P_j), \ldots$ as the signal/background probability density functions (PDF). The likelihood discriminants corresponding to the five parameters above can be written in the form:

$$\begin{aligned} L_1 &= S(P_1, P_2)/(S(P_1, P_2) + B(P_1, P_2)) \\ L_2 &= S(P_3, P_4)/(S(P_3, P_4) + B(P_3, P_4)) \\ L_3 &= S(P_5)/(S(P_5) + B(P_5)) \end{aligned} \tag{4.1}$$

In principle, one could employ likelihoods of even higher dimensionality, n, but the likelihood functions are based on table lookups and the number of entries scales like N^n, thus increasing n quickly becomes impractical.

In order to combine two likelihood discriminants into the next level likelihood we use the standard formula for probability multiplication.

$$L = L_A \otimes L_B = \frac{L_A L_B}{L_A L_B + (1 - L_A)(1 - L_B)} \tag{4.2}$$

The formula above can be extended to any number of discriminants. In the three-discriminant case one obtains.

$$L = L_1 \otimes L_2 \otimes L_3 = \frac{L_1 L_2 L_3}{L_1 L_2 L_3 + (1 - L_1)(1 - L_2)(1 - L_3)} \tag{4.3}$$

Both equations above are valid for uncorrelated discriminants and should be used with caution for real-world applications.

4.3.1 Likelihood performance estimation

The cuts or allowed bands for parameters are often tuned to maximize the figure of merit function, FOM $= S/\sqrt{S + B}$. The FOM can be viewed as a measure of the statistical significance, i.e. the ratio of the signal, S, over the possible fluctuation in the background, $\sim \sqrt{S + B}$. We use the same function for finding the boundary between meaningful signal and the background. At each step, after creating a likelihood

discriminant, we search for the maximum of the FOM as a function of the likelihood discriminant. For each value of the likelihood discriminant, L, $S(L)$ is a sum of all the signal events with likelihoods greater than L. Using an analogous definitions for the background, we can rewrite the formula for FOM:

$$\mathrm{FOM}(L) = \frac{S(L)}{\sqrt{S(L) + B(L)}} = \frac{\sqrt{S(L)}}{\sqrt{1 + B(L)/S(L)}} \tag{4.4}$$

That definition has one unfortunate property: it scales with the size of the data set. It makes it difficult to compare the FOM performance on different data sets. In order to make this function standard for a variety of data sets we will scale the FOM function with the total signal available and express it as percentage.

$$\mathrm{FOM}(L)\% = \frac{1}{\sqrt{1 + B(L)/S(L)}} \cdot \sqrt{\frac{S(L)}{S_{tot}}} \cdot 100\% \tag{4.5}$$

Thus the defined function has all the analytical properties of the FOM function, but is bound between 100% and 0%. The 100% value corresponds to ideal signal-background separation.

4.3.2 Likelihood construction

The construction of the likelihood discriminant is done in several steps. First we settle on the number of parameters going into the discriminant. In the second step, we ensure that the discriminant covers most of the signal region, while some of the background can spill into the overflow bins. In the third step, we define the number of the bins for histogram creation. In the current analysis we aim at 10% precision for each bin, thus the required mean occupancy should be above 100 entries per bin. The likelihood corresponding to each bin is defined as a simple ratio of the number of signal events to the sum of the signal and the background events. This ratio can be changed either by changing the signal or background distributions, here we will adopt the strategy of modifying the background distribution. The bin size for the histograms

in each likelihood discriminant affects the overall normalization as well. There are several ways to compensate for it. The first is by applying the constructed likelihood to the event sample and mapping the likelihood values to the $S/(S+B)$ value for each subregion of the sample. An alternative approach is to control the amount of background going into the discriminant creation so that the likelihood always behaves the same way as $S/(S+B)$. In this work we adhere to the latter approach. Thus for each likelihood discriminant, the signal distribution will always be filled by the entire signal sample, while the bin size and the background distribution will be chosen to ensure proper likelihood normalization and to minimize correlations.

During likelihood construction an overall normalization is performed to make the likelihood function as close as possible to the true "signal probability". That allows us to treat multiple decay channels in the same way. For example, if we have properly normalized likelihood functions for both $J/\psi \to e^+e^-$ and $J/\psi \to \mu^+\mu^-$, in $B \to J/\psi \;\; X$ we do not need separate treatments for the J/ψ sub-decay channels.

The background distribution for the signal-free areas is of no interest to us. Therefore, we do not treat properly regions with high background levels and negligible signal fractions. Neglecting these areas results in distortions for low likelihood values, where the lack of attention to the background distributions is most prominent.

Extracting the PDF distributions for the signal and background from the data presents additional challenges. Unlike the MC case, we can not select signal with 100% probability, thus the distributions that are created will always have background contamination in the signal and signal contamination in the background. The inability to remove background events that closely mimic signal events results in an offset for the final likelihood. If we define the true likelihood as $L_{\rm true} = S/(S+B)$, the one extracted from the data sample will be as follows: $L_{\rm data} = (S + B_{\rm mimic})/(S+B)$. Another fundamental challenge arises because a proper PDF estimation requires separation of our sample into signal and background using the most powerful discriminant,

thus making it unavailable for further inclusion in the likelihood. In addition, special care must be taken with respect to variables that might be correlated with the discriminant used for signal-background separation. Despite these challenges, we will demonstrate that proper likelihood for J/ψ selection can be constructed from the data.

Process	$b\bar{b}$	$q\bar{q}$ (non $b\bar{b}$)	$\tau\bar{\tau}$	QED	$\gamma\bar{\gamma}$	Beam Gas
σ(nb)	1.1	3.3	0.93	37.8	11.1	-
ϵ HadronA (%)	0.994	0.838	0.240	0.002	0.008	-
σ HadronA (nb)	1.09	2.77	0.22	0.06	0.09	0.34

Table 4.2: e^+e^- cross section. We did not include cross sections for e^+e^- scattering and $\mu^+\mu^-$ production, because the clean signature for these reactions allows their heavy suppression. QED includes radiative electron interactions, integrated for detector configuration. This table was taken from [57].

4.4 J/ψ selection

4.4.1 HadronA selection criteria.

B mesons are not the only possible products of an e^+e^- collision. In fact non $b\bar{b}$ final states, from continuum, are by far more likely at the 10 GeV/c^2 collision energy (see Table 4.2). The first step selection, named HadronA ([58]), rejects most of the background induced by e^+e^- scattering, $\mu^+\mu^-$ production and radiative electron interactions. Its performance is summarized in last two lines of the Table 4.2. The second step is a selection of a pure J/ψ sample. This is a crucial point for the $B \to J/\psi\ K_L$ study, because a large combinatorial contamination and the difficulty of K_L identification makes this channel prone to the numerous backgrounds. It will be shown later that requiring the presence of both a J/ψ and a K_L in an event will effectively remove all backgrounds but J/ψ inclusive ones. This allows us to use the J/ψ inclusive MC sample for tuning the K_L selection criteria. The final step is the $B \to J/\psi\ K_L$ selection and suppression of relevant backgrounds.

Throughout this chapter, we will make extensive use of the beam center of mass frame (CM). In this frame, both B mesons have relatively small momentum, $\approx$ 330 MeV/c, and this frame can be used as a good approximation to the B meson rest frame.

Hadronic events are selected based on charged track information and energy deposited in the electromagnetic calorimeter. We require at least three tracks that satisfy $\sqrt{x^2+y^2} < 2.0$ cm, $|z| < 4.0$ cm, and $p_t > 0.1$ Gev/c, where x, y and z represent the distance of closest approach of the track to the beam axis, and p_t is the momentum of the track projected onto the x-y plane. We also require that more than one neutral cluster is observed in the barrel region of the detector and has energy greater than 0.1 GeV/c^2. This cluster need not be matched with any charged track. The sum of all cluster energies, boosted back to the CM assuming each cluster is generated by a massless particle, is required to be between 10% and 80% of the total CM energy. The total visible energy in the CM, which is an energy sum of good charged and unmatched neutral tracks $E_{\rm vis}^{\rm cms}$, is greater than 20% of the total CM energy. The absolute value of the z component of the CM momentum is required to be less than 50% of the CM energy. The event vertex reconstructed from the selected tracks must be within 1.5 cm and 3.5 cm of the interaction region in the directions perpendicular and parallel to the z axis, respectively. MC simulation shows that the selection criteria described above retain more than 99% of $B\bar{B}$ events and J/ψ inclusive events.

4.4.2 Preselection for the inclusive charmonium skimming.

In the present analysis, J/ψ event candidates are reconstructed from lepton pairs only. The lepton modes constitute only 11.8% of all possible J/ψ decays, but they are the only ones sufficiently free of background for inclusive selection. The mass distributions for dilepton pairs in the inclusive charmonium sample are presented in Fig. 4.1, while the CM momentum is shown in Fig. 4.2. The momentum distribution sheds some light on the background composition for the inclusive sample. The sharp signal cut off around 2 GeV/c^2 corresponds to the kinematic end point for J/ψ's produced in B-meson decay. The difference between the solid and the shaded histograms is

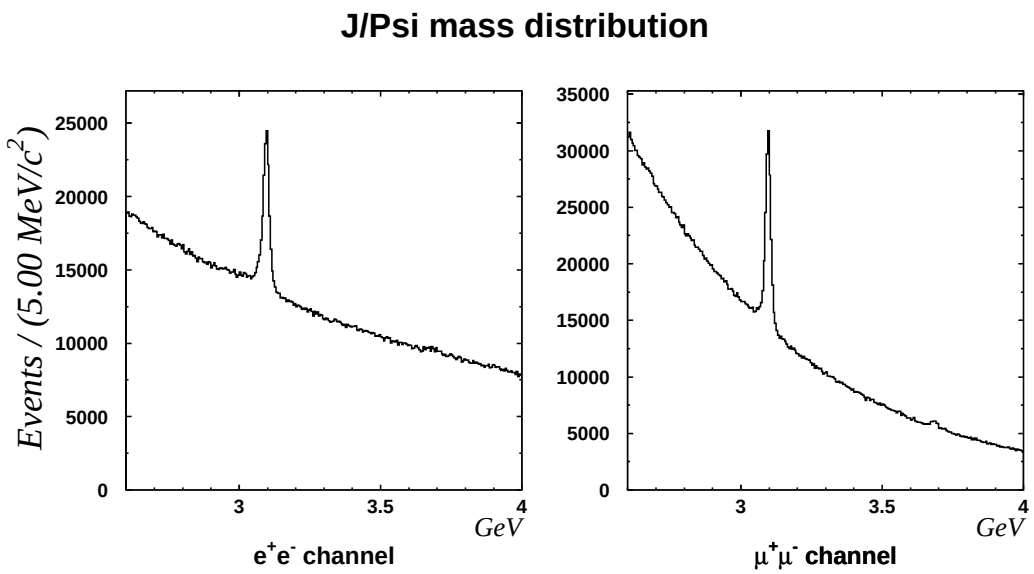

Figure 4.1: The dilepton mass distributions for the inclusive charmonium skim. The plot is based on a 78.13 fb^{-1} data sample.

due to the contribution from inclusive B meson decays. The background for the di-electron decays is dominated by beam-gas-interaction and QED backgrounds, while the di-muon channel is prone to misidentified pions and kaons from the $q\bar{q}$ continuum. The inclusive J/ψ sample serves as a pool for testing and reconstructing dilepton decays of J/ψ and $\psi(2S)$ mesons. Candidates from it satisfy the requirements of the HadronA filter, and should have positive identification for one lepton, that is electron probability should be greater than 0.01. The second lepton should either have positive identification or energy deposited in the electromagnetic calorimeter that is consistent with a lepton hypothesis. The showers associated with electrons should have energy greater than half of the associated track momentum, while for muons we require the deposited energy to be in the $100 - 300$ MeV/c range. In order to account partially for final-state radiation and bremsstrahlung we add photons found within 50 mrad of the e^+ or e^- direction. The mass of the dilepton pair should satisfy 2.5 GeV/c^2 $< M_{l^+l^-} <$ 4.0 GeV/c^2 and the CM momentum should lie in the range 0 GeV/c $< p_{\text{cms}} <$ 5 GeV/c.

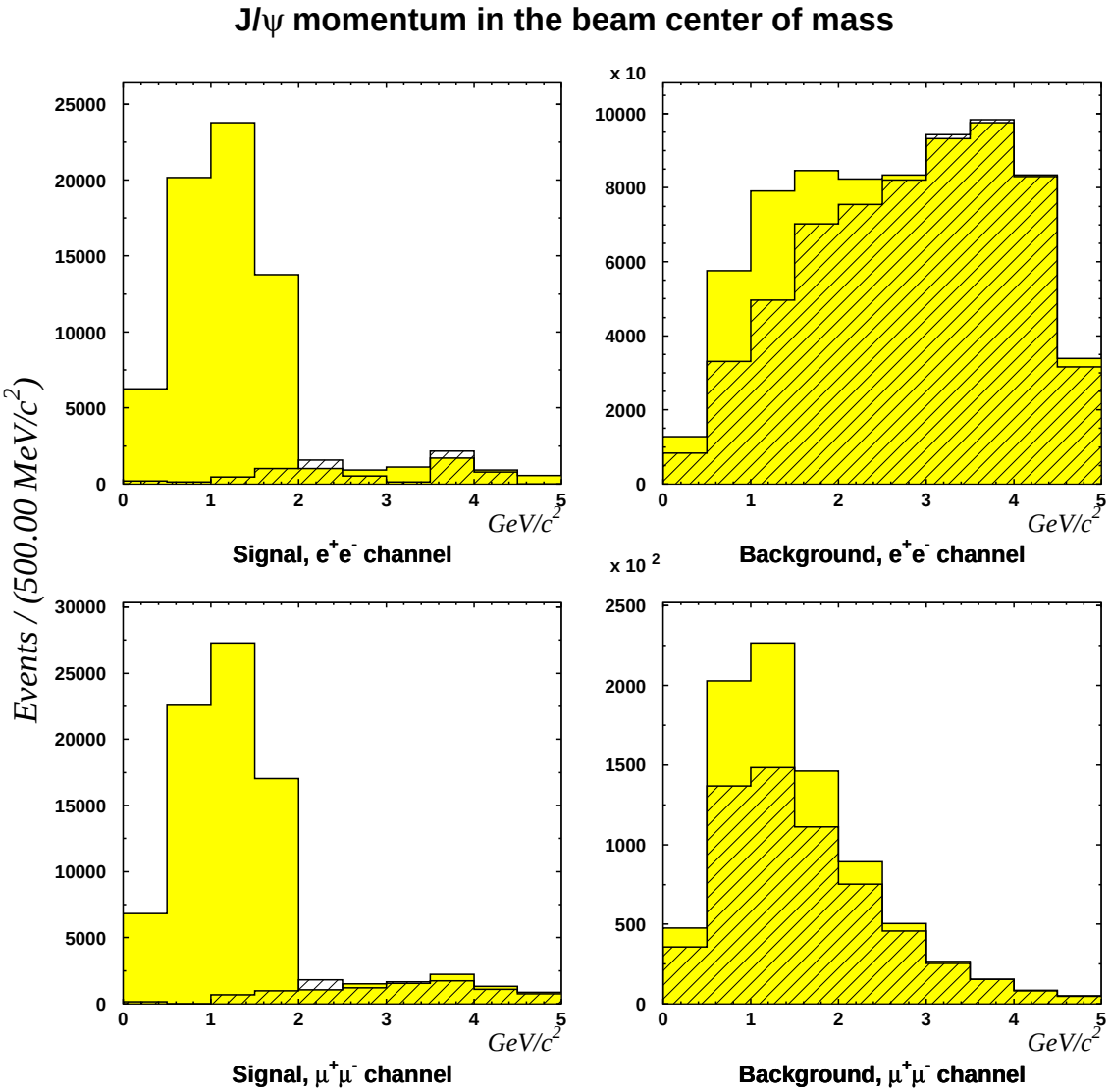

Figure 4.2: The solid histogram represents the on-resonance data, while the hatched area corresponds to the continuum data scaled to the on-resonance luminosity. The momentum spectra were obtained by fitting the dilepton mass distributions for each bin of the momentum spectrum. The number of events for both signal and background were estimated in the $2.9 < M_{ll} < 3.15$ GeV/c^2 mass window. The plot is based on a 78.13 fb^{-1} on-resonance and a 8.83 fb^{-1} continuum data sample. Continuum was taken at an energy 60 MeV/c^2 below the center value for the $\Upsilon(4S)$ resonance.

As can be seen from Fig. 4.2, the inclusive J/ψ skim has a large amount of background in the pure continuum region, $p^*_{J/\psi} > 2$ GeV/c. This region is of no interest for our study and we will select only J/ψ candidates produced in $B\bar{B}$ decays: $p^*_{J/\psi} < 2$ GeV/c (Eq. B.13).

4.4.3 Likelihood construction

The final J/ψ likelihood is based on three discriminants. $L^E_{J/\psi}$ depends on the event shape variables, $L^C_{J/\psi}$ takes into account track properties, and $L^M_{J/\psi}$ is based on the shape of the mass spectrum, . The final analysis uses a combined likelihood,$L_{J/\psi} = L^E_{J/\psi} \otimes L^C_{J/\psi} \otimes L^M_{J/\psi}$. In certain circumstances, however, it is advantageous to use a single discriminant.

The lepton and muon channels have quite different sources of background. Electrons are so well identified that the J/ψ sample consists of virtually only dielectrons, with most of the background coming from beam-gas interactions and QED processes. By contrast, muon identification has a higher fake rate resulting in larger contamination from fast pions and kaons in the dimuon channel. The abundance of pions among the $q\bar{q}$ and B meson decay products make them the largest background. As a result, the event shape likelihood is more efficient for suppressing background in the electron-positron channel, while the muon related backgrounds are more heavily suppressed by the candidate likelihood. The combined event plus candidate likelihood, $L^E_{J/\psi} \otimes L^C_{J/\psi}$, exhibits similar performance for both channels.

Event shape Likelihood, $L^E_{J/\psi}$

Inspired by the work of Steve Olsen [59] on using the event shape variables for inclusive J/ψ skimming we decided to go a bit further and include all the variables he discussed in [59] into one likelihood function. Before proceeding we remind the reader of the definitions for the standard event classification variables used in Belle experiment.

$P_{\mathbf{MM}}$ Missing momentum value. Calculated as the difference between the net beam three-momentum and the sum of all charged tracks and identified photons. Clusters with EFC energy deposition greater than 100 MeV/c^2 are also added as photons.

$E_{\mathbf{vis}}$ This is the total energy of all charged tracks and unmatched neutral showers calculated in the CM frame. $E_{\rm vis} = \sum_{\rm tracks} \sqrt{p_{\rm cm}^2 + m_\pi^2} + \sum_{\gamma' s} E_{\rm cm}$

P_z This is the z component of the total three-momentum of all charged and unmatched neutral tracks in the CM frame.

$N_{\mathbf{tracks}}$ This is the number of charged tracks in the event.

$N_{\mathbf{clusters}}$ This is the number of clusters in the electromagnetic calorimeter.

$E_{\mathbf{sum}}$ This is the sum of the energy deposited in the calorimeter. All clusters are treated as a massless particles, the combined four-momentum is then boosted to CM frame and the resulting energy is called the energy sum, $E_{\rm sum} = \sum_{\rm clusters} E_{\rm cm}$

$M_{\mathbf{heavy\ jet}}$ In the CM frame, the event is divided into two hemispheres by a plane perpendicular to the thrust axis. The invariant mass of all particles in each hemisphere is calculated. The *heavy jet mass* is the larger of the two mass values.

R_2 This is the ratio of the second- and zeroth-order Fox-Wolfram moments of the events H_2/H_0 [60]. $H_l = \sum_{i,j} |\overrightarrow{p}_i||\overrightarrow{p}_j| P_l(\cos\theta_{i,j})$, where the indices i and j run over all final state particles. p is the particle momentum, P_l is the l-th Legendre polynomial, and $\theta_{i,j}$ is the angle between particle directions.

There are many sources of background, among them QED dominated processes, continuum quark production, beam-gas interactions and particle misidentification. Each of these backgrounds has a different signature in the event shape variables.

For example, the visible energy of generic B decays sits around 8 GeV with a mean transverse momentum around zero. Bhabba events, on the other hand, will tend to have an energy not far from either beam energy or half the beam energy. Moreover, this energy will be correlated with the value of the transverse momenta. A high value of p_z can also be attributed to high-energy particles escaping detection because of the poor detector coverage along the beam line. In the case of QED background or light jets from the continuum, these particles will have a total energy comparable to half of the beam energy in the CM frame.

In addition to continuum suppression, it is possible to use the event-shape variables to discriminate against other B meson decays. The inclusive charmonium sample will differ from the rest of the B decays because of the high mass of the J/ψ meson ($M_{J/\psi} = 3.096$ GeV/c^2). This will limit the amount of energy available for the kinematic motion for the rest of the decay, and thus result in a shift for most event classification variables.

Another distinct feature is due to the leptonic nature of J/ψ decay. Muons are characterized by high penetration power, and, as a result will deposit only a small energy in the calorimeter. Electrons on the other hand lose all of their energy in the calorimeter, thus prompting a calorimeter energy deposition shift between the muon and electron channels of approximately 3 GeV.

The event shape variables are not fully independent. In order to exploit the correlation between them, we arrange the variables into pairs and construct a set of two-dimensional probability density functions for likelihood construction. After careful consideration, we decided to use the following pairs of variables for likelihood estimation:

- Visible energy vs. p_z.
- $N_{\rm trk}$ vs. $N_{\rm cls}$
- Thrust vs. $E_{\rm sum}$.

- Heavy jet mass vs. $R2$.
- $P_{\rm sum}$ vs. missing momentum ($P^{\rm MM}$).
- Missing momentum θ vs. missing momentum ϕ.

Candidate likelihood, $L^C_{J/\psi}$

The candidate likelihood deals with the properties of the tracks forming J/ψ's as well as the kinematic parameters of the J/ψ meson. Let us introduce the following definitions:

$L^{\mathbf{ID}}$ Lepton ID.

$P^{\mathbf{IP}}$ Probability of track coming from the Interaction Point (IP). Calculated using the gaussian overlap in projections of the IP profile and track errors into dr and dz track helix parameters. It is used for vetoing beam-gas backgrounds.

$P^{\mathbf{Bhabba}}$ Momentum of the Bhabba partner. Using the direction of one lepton we can calculate the momentum of the second lepton under the assumption that both leptons come from Bhabbas.

The candidate likelihood is based on the following two-dimensional discriminants:

- Azimuthal angle of the J/ψ candidate in the CM frame versus the J/ψ decay angle(angle between J/ψ flight direction and direction of one of the daughter leptons in the J/ψ rest frame).
- $L_1^{\rm ID}$ vs. $L_2^{\rm ID}$.
- $P_1^{\rm IP}$ vs. $P_2^{\rm IP}$.
- $P_1^{\rm Bhabba}$ vs. $P_2^{\rm Bhabba}$.

Mass likelihood, $L^M_{J/\psi}$ (J/ψ mass line shape).

The natural width of the J/ψ is rather narrow, 87 ± 5 keV, and is not observable with the current detector resolution of ~ 10 MeV/c^2. Thus the observable line shape will be dominated by the detector response function. In addition to the detector resolution we have to take into account final state radiation (FSR, [61, 62]) and bremsstrahlung. These contributions are equally large for electrons, and the former is dominant for muons (more detailed discussion can be found in Ph.D. thesis of A.Esrhov [63]).

The combination of bremsstrahlung, final state radiation, and detector resolution creates a nontrivial line shape. There are many ways to approximate it, but we use a double-sided Crystal Ball function [64].

$$\begin{aligned} x - \mu \leqslant -\sigma\alpha_L : \quad & \frac{A \exp\left(-0.5\alpha_L^2\right)}{\left(1 - \frac{\alpha_L}{N_L}\frac{x-\mu}{\sigma} - \frac{\alpha_L^2}{N_L}\right)^{N_L}} \\ -\sigma\alpha_L < x - \mu < \sigma\alpha_R : \quad & A \exp\left(\frac{(x-\mu)^2}{2\sigma^2}\right) \\ x - \mu \geqslant \sigma\alpha_R : \quad & \frac{A \exp\left(-0.5\alpha_R^2\right)}{\left(1 - \frac{\alpha_R}{N_R}\frac{x-\mu}{\sigma} - \frac{\alpha_R^2}{N_R}\right)^{N_R}} \end{aligned} \tag{4.6}$$

This function behaves as a gaussian near its mean, but has tails that decay not exponentially, but according to a power-law function. The coefficients are chosen to make the function smooth. The primary reason for the left tail is FSR in the muon channel and the combination of FSR and bremsstrahlung for di-electron decay. The right-side tail is due to non-gaussian tails in the detector response. In the case of electrons, there is contamination from good J/ψ's, fake photons, or photons from another decay channel. The mean (μ) and standard deviation (σ) serve as estimates for the J/ψ peak position and line width, although one should not use them as the true ones. The central part is approximated with a symmetric function while FSR

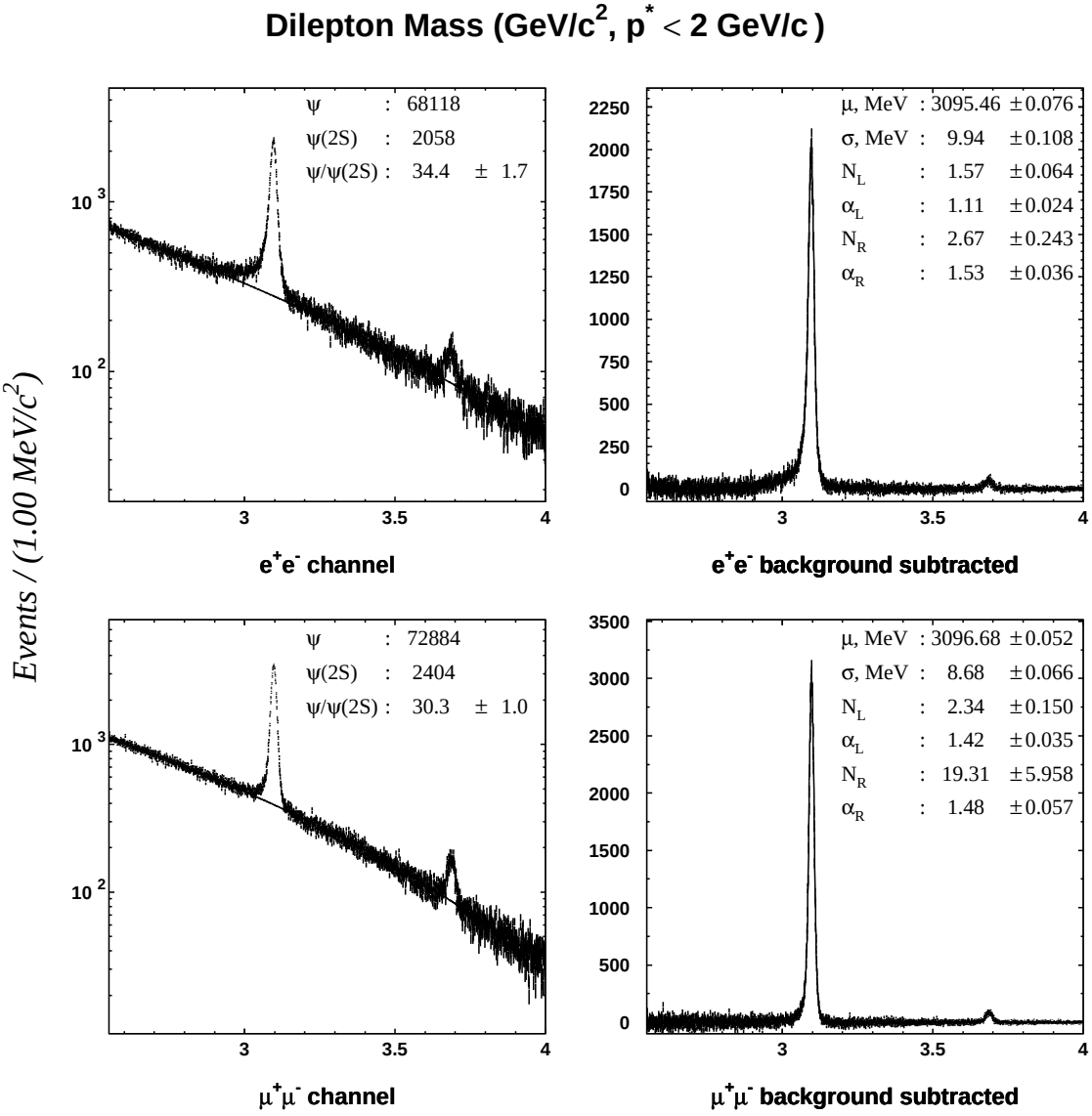

Figure 4.3: The mass spectra for dilepton pairs. The plots are based on a sample satisfying $L^E_{J/\psi} \otimes L^C_{J/\psi} > 0.2$ and $p^*_{J/\psi} < 2$ GeV/c. The number of events for both J/ψ and $\psi(2S)$ is an integral of the corresponding shapes (Eq. 4.6). In the fit we have fixed the position of the $\psi(2S)$ peak and all the parameters of the $\psi(2S)$ shape are equal to those of the J/ψ. The amplitude is the only floating parameter for the $\psi(2S)$. Plot is based on a 78.13 fb^{-1} on-resonance data sample.

and bremsstrahlung skew the line shape to the left. Nevertheless, this is good enough for signal yield estimation and likelihood construction. The line shape parameters can be estimated either from the signal MC or from the data. Here we present parameter estimation from data sample corresponding to 78.13 fb^{-1}. In order to use the line shape in a likelihood framework we will select a sample based on the maximum value of the FOM as a function of combined likelihood, $L^{E}_{J/\psi} \otimes L^{C}_{J/\psi}$.

4.4.4 Likelihood tuning

A two-dimensional PDF construction requires higher statistics than does the one-dimensional case. As a result, the amount of the available data becomes a crucial factor. With 79 fb^{-1} we have collected approximately 70,000 J/ψ candidates in each lepton decay channel. To have a meaningful distribution, we need at least 20 bins for each parameter, thus each 2-D PDF will have at least 400 bins. Given that we are using ten different likelihood discriminators, we will get $\sim$ 4000 bins and the average occupancy will be $\sim$ 20. Such a fine segmentation raises a major concern regarding possible bias for the data used in the likelihood construction.

In order to test the effect of self bias, we have divided the data sample into two equal sub samples, "A" and "B". We have used sample "A" for likelihood construction, while "B" was used for a likelihood performance estimation. In a second approach, we have used the entire data sample for both the likelihood construction and a performance check. After thorough comparison, we have not observed any bias related to the use of the same sample for both construction and performance measurement, and, as a result, we decided to use the full data sample for likelihood construction. This result is not surprising, since for most distributions, 80% of the signal typically occupies no more than 30% of the corresponding PDF, and as a result, even a partial sample will have enough statistics to cover the full phase space.

In our quest for a J/ψ likelihood, we have used an iterative approach. In the beginning we have selected the mass of the lepton pair as the main discriminator between signal and background. We have used wide mass region for J/ψ likelihood construction ($2.5 < M_{l^+l^-} < 5.0$ GeV/c^2), with subregions $3.06 < M_{l^+l^-} < 3.12$ GeV/c^2 and $3.66 < M_{l^+l^-} < 3.71$ GeV/c^2, declared to be signal while the complementary region is taken as background. In order to improve the separation of signal and background, we have performed a mass fit for each of the regions and subtracted sidebands scaled according to the fit. In the case of the heavy-jet mass variable, this approach does not work because it is correlated with the dilepton pair mass.

At the later stages, however, we have used the "full", $L_{J/\psi} = L^E_{J/\psi} \otimes L^C_{J/\psi} \otimes L^M_{J/\psi}$, likelihood from the previous iteration to select the signal. In other words, we declare any candidate to be a signal if its likelihood is greater than a certain cut off value.

4.4.5 Likelihood performance, data check.

After creating the likelihood, we constructed several tests in order to verify its performance and consistency. The bottom left plots of Fig. 4.4 and Fig. 4.5 show a linear correlation between likelihood and $S/(S+B)$. Signal and background fractions were estimated by mass fits in each bin of the likelihood distribution. The top left plots show that most of the J/ψ mesons produced in B meson decays occupy bins with large likelihood values and bottom right plot tells us that for $L > 0.44(0.42)$ we select 87.4(90.1)% of all available J/ψ candidates with purity 82.4(81.9)% in di-electron(di-muon) channel, thus allowing us to select a high statistics J/ψ sample with high purity. The bottom right plots shows the FOM(L), for the events with likelihood greater than parameter L. Both dilepton channels have FOM maximum around likelihood value of 0.4. This value is good for the inclusive J/ψ study, however it corresponds to a 10% signal loss, which is unacceptable for exclusive-mode selection. In order to limit the effect on the inclusive mode selection, we have decided to

use a likelihood cut value which would guarantee a signal loss below 5%. The dilepton mass distributions presented in Fig. 4.6 and Fig. 4.7 shows the performance of the event likelihood discriminant alone (top plots) and combined with the candidate discriminant (bottom plots).

In order to understand the background composition of the events passing the likelihood cut, we have replicated Fig. 4.2 for the events with $L > 0.2$. The result is presented in Fig. 4.8, which clearly shows that continuum contamination is still present in our sample, however significantly reduced. These plots also show that our likelihood does not suppress backgrounds with CM momentum above 2 GeV/c, which is a direct consequence of our selection criteria in the likelihood construction.

4.4.6 Likelihood performance, MC check.

Several simple tests were performed on an 0.5-M-event inclusive J/ψ MC sample. These tests estimate the efficiency for decays relevant to our study. For the J/ψ K decay study, we are interested in J/ψ's with CM momentum in the range 1.3 GeV/c < $p^*_{J/\psi} < 2.0$ GeV/c. This J/ψ momentum range covers the tails induced by the detector resolution for J/ψ K decays and is predominantly filled with J/ψ K and J/ψ K^* modes.

One possible drawback of a likelihood tuned on the inclusive data sample is the possibility of losing some exotic final states, despite keeping good overall efficiency. In order to make sure that this is not the case for the decays of interest, we have checked the detection efficiency for J/ψ detection in all J/ψ K and J/ψ K^* modes. These decay channels encompass a variety of decay topologies and present a good test of the constructed likelihood. As can be seen from Table 4.3, the efficiencies for different decay topologies vary, but not by much more than statistical fluctuations, which are of order 1% for this MC sample. The largest contribution to the efficiency loss for J/ψ K_L decay are the large missing momentum and the smaller track and

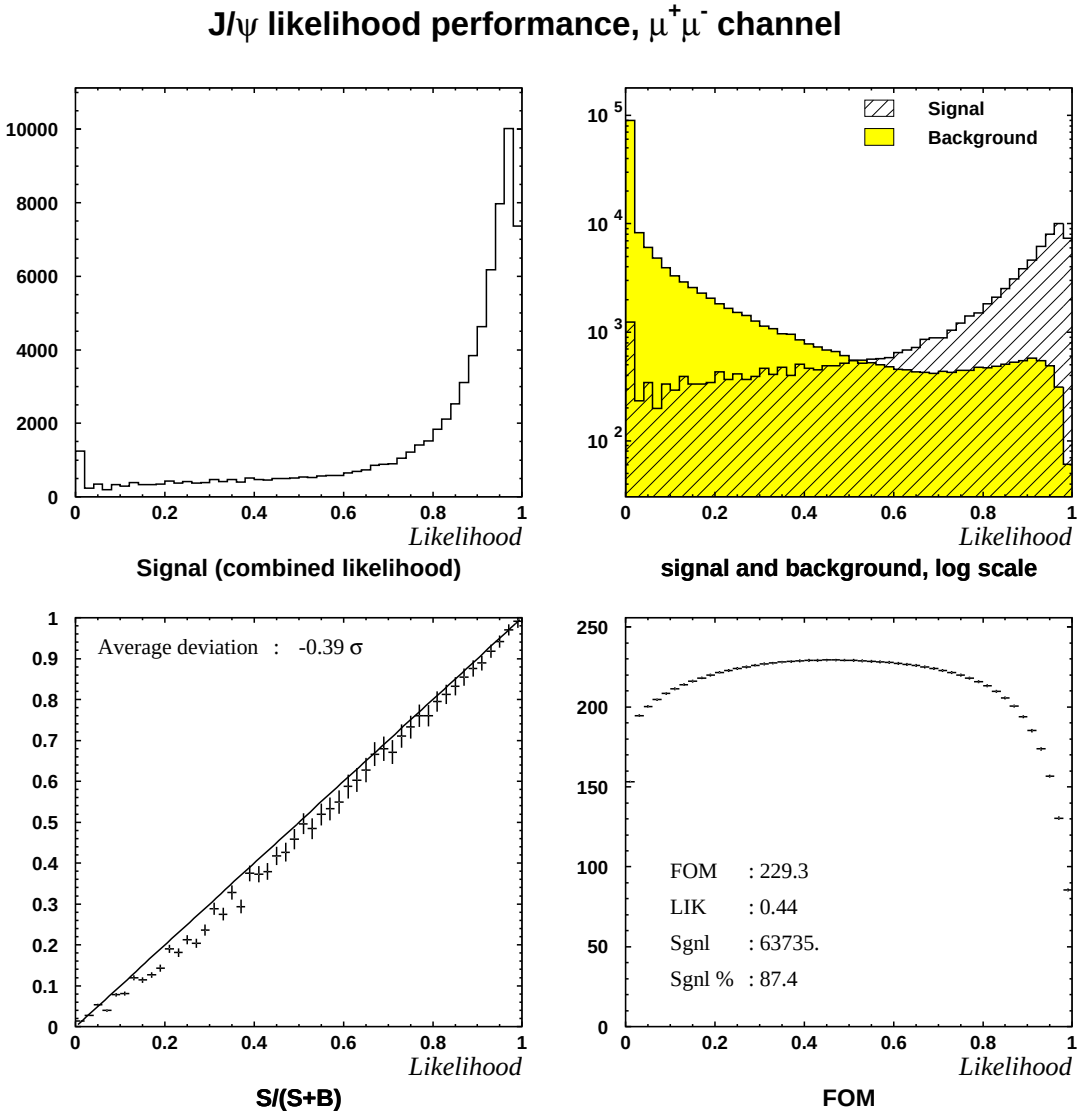

Figure 4.4: The S/(S+B) is obtained using a mass fit for the signal and the background in each bin. The FOM = $S/\sqrt{S+B}$ is an integral value for all events with a likelihood greater than some value. The plot is based on a 78.13 fb^{-1} on-resonance data sample.

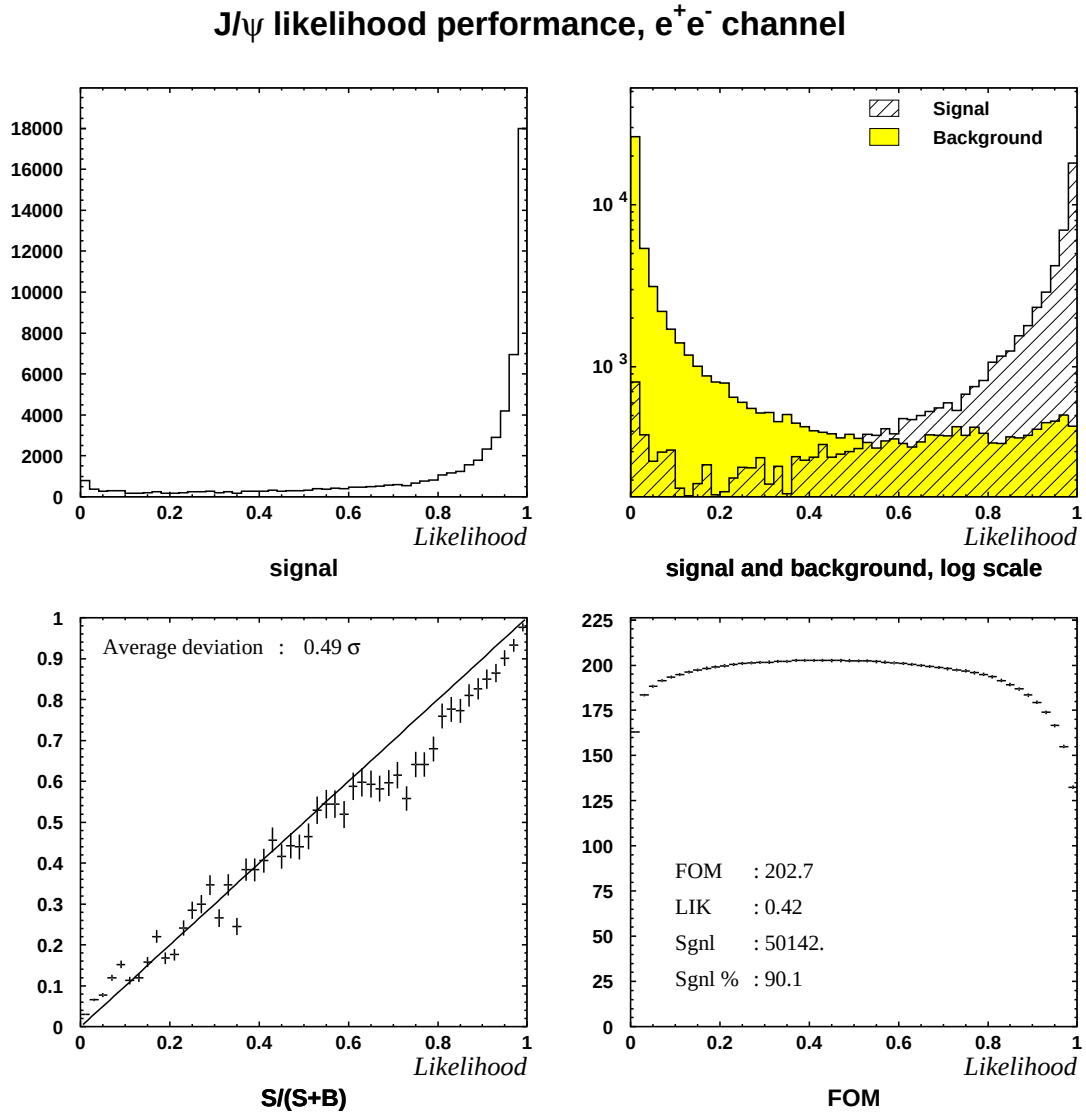

Figure 4.5: The S/(S+B) is obtained using a mass fit for the signal and the background in each bin. The FOM = $S/\sqrt{S+B}$ is an integral value for all events with a likelihood greater than some value. The plot is based on a 78.13 fb^{-1} on-resonance data sample.

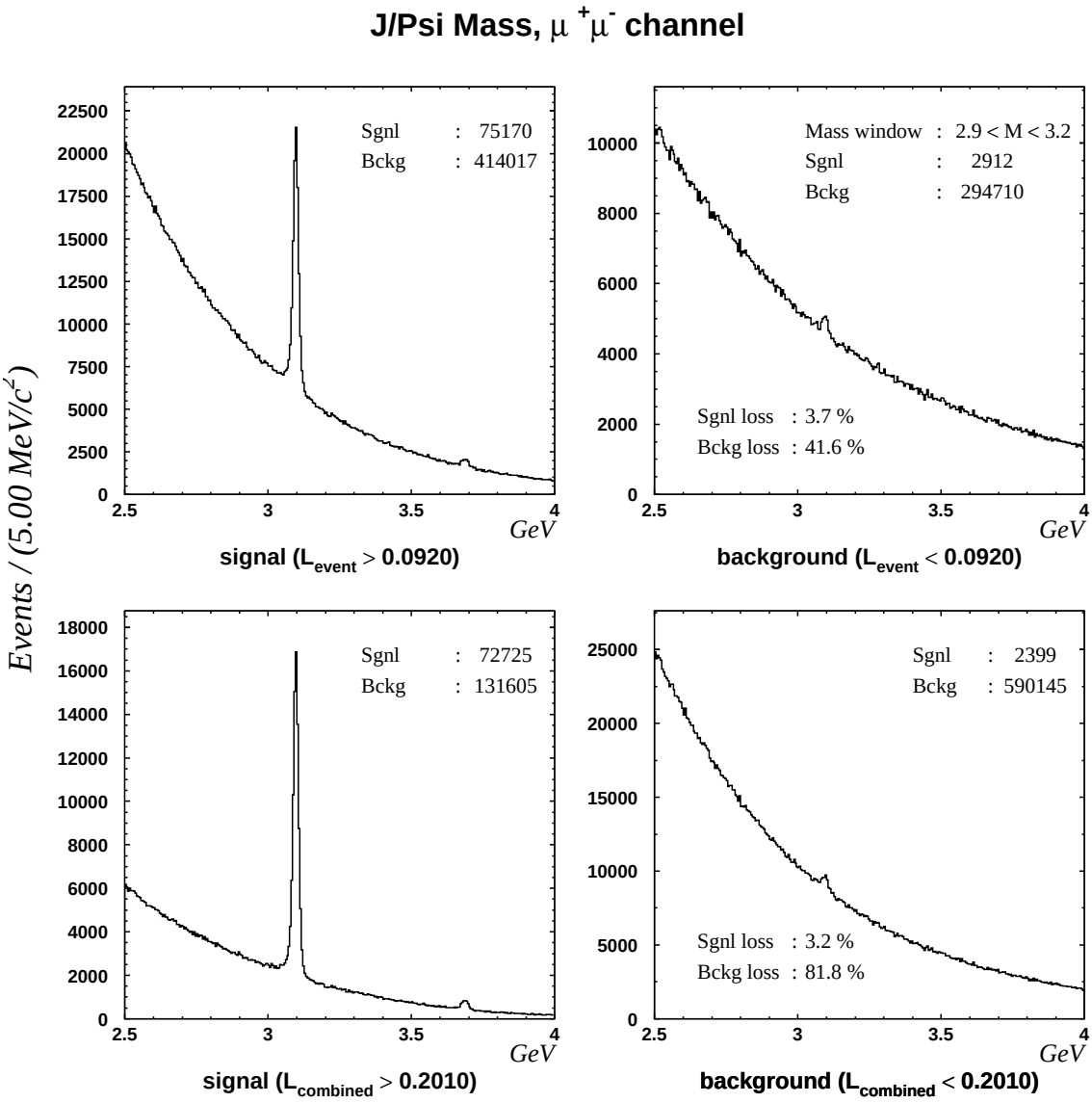

Figure 4.6: The mass spectrum for the di-muon pairs. The number of events for both the J/ψ signal and the background are estimated in the $2.9 < M < 3.2$ GeV/c^2 mass window. We assume the signal shape from Fig. 4.3. The signal-background separation is based on a likelihood cut, which maximize FOM. The plot is based on a 78.13 fb^{-1} on-resonance data sample.

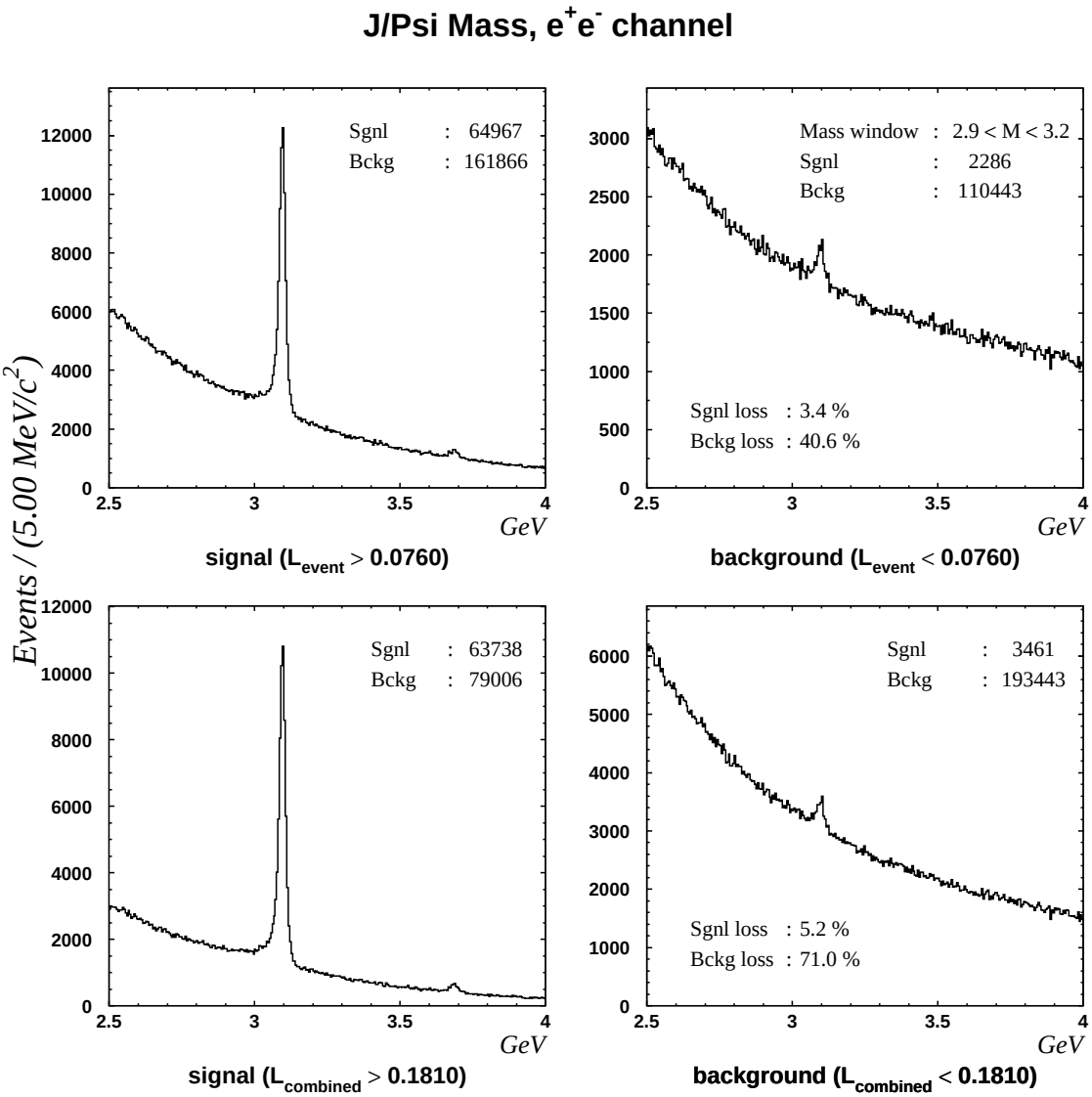

Figure 4.7: The mass spectrum for the di-electron pairs. The number of events for both the J/ψ signal and the background are estimated in the $2.9 < M < 3.2$ GeV/c^2 mass window. We assume the signal shape from Fig. 4.3. The signal-background separation is based on the likelihood cut, which maximize FOM. The plot is based on a 78.13 fb^{-1} on-resonance data sample.

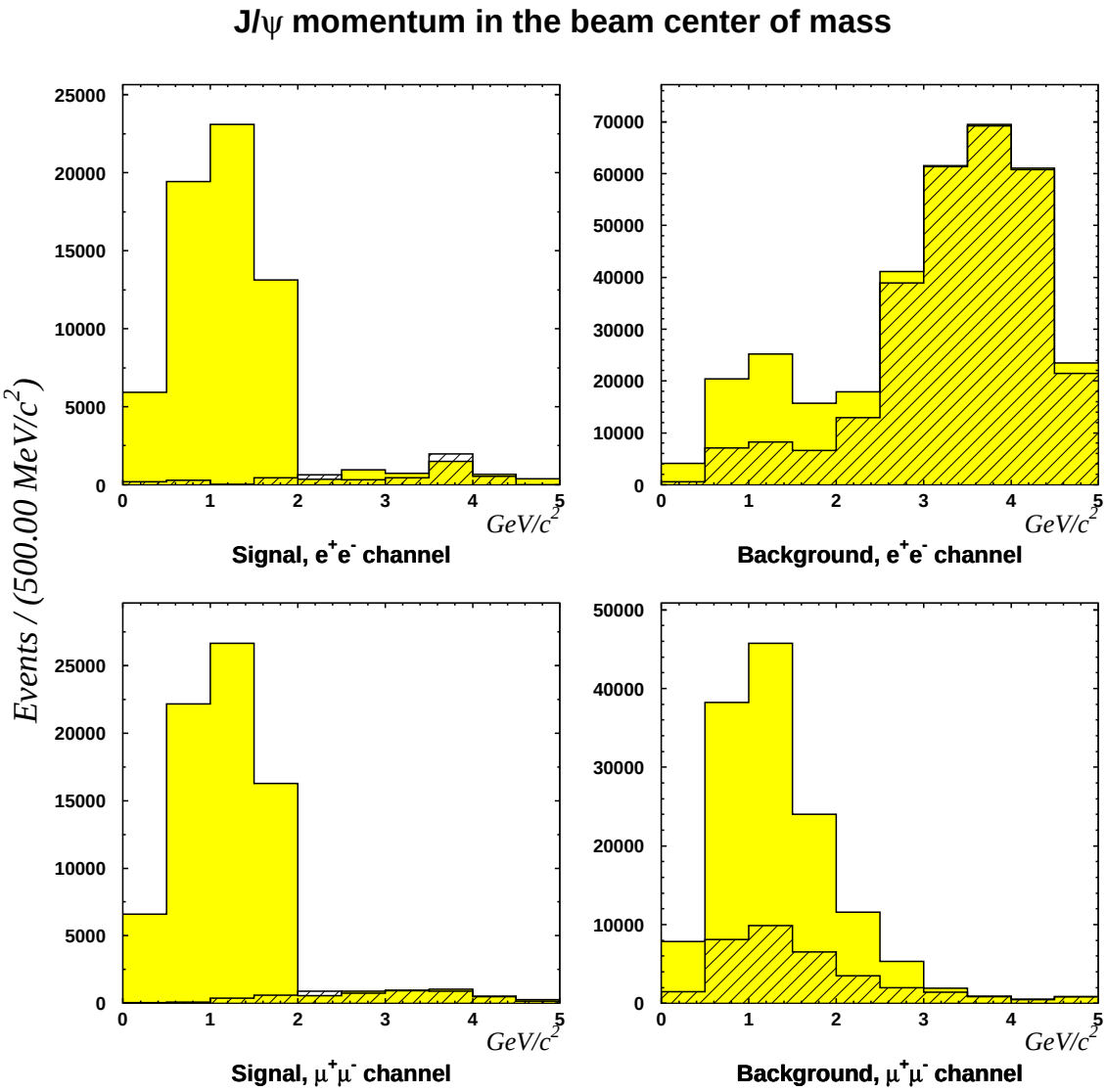

Figure 4.8: The plots are based on a sample satisfying $L^E_{J/\psi} \otimes L^C_{J/\psi} > 0.2$. The solid histogram represents the on-resonance data, while the shaded one corresponds to the continuum data scaled to the on-resonance luminosity. The momentum spectra were obtained by fitting the dilepton mass distributions for each bin of the momentum spectrum. Both the number of signal events and the background events were estimated in the $2.9 < M_{ll} < 3.15$ GeV/c^2 mass window. The plot is based on a 78.13 fb^{-1} on-resonance and a 8.83 fb^{-1} continuum data sample. Continuum was taken at an energy 60 MeV below the peak of the $\Upsilon(4S)$ resonance.

mode X, $B \to J/\psi + X$	J/ψ generated	J/ψ detected	efficiency		
			ε, %	ε_{ee}, %	$\varepsilon_{\mu\mu}$, %
K_L	10148	6709	66.1 ± 0.8	62.1 ± 1.1	70.1 ± 1.2
$K^{*+}((K_S(\pi^+\pi^-)\pi^+)$	7706	5208	67.6 ± 0.9	62.3 ± 1.3	72.9 ± 1.4
$K_S(\pi^+\pi^-)$	6942	4725	68.1 ± 1.0	64.2 ± 1.4	71.9 ± 1.4
$K^{*0}(K^+\pi^-)$	22208	15245	68.6 ± 0.6	66.2 ± 0.8	71.1 ± 0.8
$K^{*0}(K_L\pi^0)$	5471	3752	68.6 ± 1.1	63.6 ± 1.5	73.6 ± 1.6
$K^{*+}(K^+\pi^0)$	11124	7630	68.6 ± 0.8	63.7 ± 1.1	73.5 ± 1.1
K^+	19854	13648	68.7 ± 0.6	65.1 ± 0.8	72.4 ± 0.8
$K^{*+}(K_L\pi^+)$	11182	7727	69.1 ± 0.8	65.6 ± 1.1	72.6 ± 1.1
$K^{*0}((K_S(\pi^+\pi^-)\pi^0)$	3818	2653	69.5 ± 1.3	64.9 ± 1.8	74 ± 2.0
$K^{*0}((K_S(\pi^0\pi^0)\pi^0)$	1712	1203	70 ± 2	65 ± 3	76 ± 3
$K^{*+}((K_S(\pi^0\pi^0)\pi^+)$	3521	2467	70.1 ± 1.4	65.4 ± 1.9	75 ± 2
$K_S(\pi^0\pi^0)$	3041	2150	70.7 ± 1.5	67 ± 2	74 ± 2

Table 4.3: J/ψ efficiency for $B \to J/\psi\ K^{(*)}$ modes. Event selection is based on the 0.5-M-event inclusive J/ψ MC sample. We consider a J/ψ as a detected particle if we observe it through dilepton decay with $2.8 < M_{l^+l^-} < 3.2$ GeV/c^2, $1.2 < p^*_{l^+l^-} < 2.0$ GeV/c and $L_{J/\psi} > 0.03$. Modes are arranged in order of increased J/ψ efficiency.

neutral cluster multiplicity. Events with a large number of tracks or large energy deposition have the highest efficiency. The difference in efficiency between kaon and pion modes is probably due to the larger energy deposition in the calorimeter for the particles with an s quark (K^- and $\bar{K}^0$).

Another test comes from the strictly selected $J/\psi\ K^\pm$ and $J/\psi\ K_S$ events, for which we can obtain clean samples (background contamination less than 3%). We have applied the same tight selection criteria for both data and MC samples and plotted the likelihood distributions for the surviving candidates, results are shown in Fig. 4.9. The J/ψ inclusive MC has a slightly higher mean value than the one obtained from data. The difference is probably due to a higher number of fake J/ψ's in the data. Otherwise agreement between MC and data is pretty good.

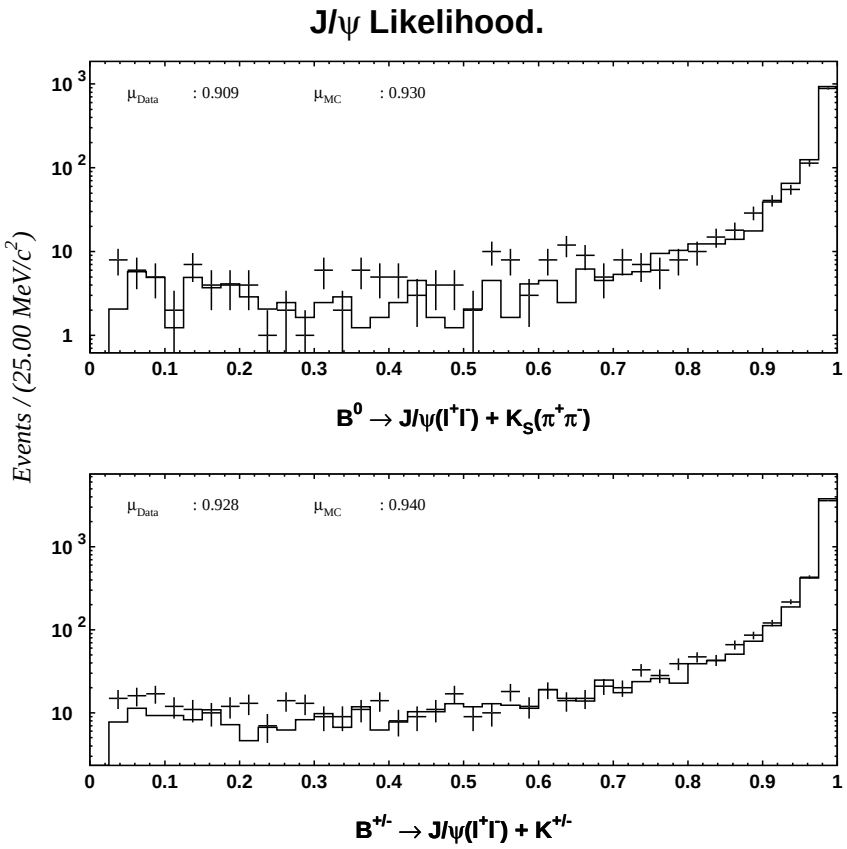

Figure 4.9: The full J/ψ likelihood for the data and the MC. The MC is scaled with respect to the number of entries in the data sample. The plot is based on a 78.13 fb^{-1} on-resonance data sample.

4.5 $c\bar{c}$ resonances decaying through J/ψ.

4.5.1 Introduction.

B^0 can decay to several unstable "charmonium" states, with subsequent transitions resulting in J/ψ produced via radiative or hadronic decays. The most prominent examples are the $\psi(2S)$ and χ_c resonances, which have branching ratios comparable to direct J/ψ or $\psi(1S)$ production. According to recent measurements [65] about 30% of J/ψ produced in inclusive B meson decays are results of the cascade transitions $B \to XY$, with $X \to J/\psi\ Z$. Exclusive selection and subsequent veto of such decays can improve the selection of B meson decays that are plagued by large backgrounds. It also provides an independent quality test for the J/ψ selection and the validity of the likelihood approach for signal-yield estimates of the states decaying through the J/ψ resonance. The dominant decays ending in the production of J/ψ's include the following $c\bar{c}$ resonances $\chi_{c1}(1P)$ ($M = 3510.51 \pm 0.12$ GeV/c^2), $\chi_{c2}(1P)$ ($M = 3556.18 \pm 0.13$ GeV/c^2) and $\psi(2S)$ ($M = 3685.96 \pm 0.09$ GeV/c^2). The inclusive branching fraction for J/ψ production through $B \to \chi_{c1,c2}, \psi(2S)$ transitions is given by the following expressions:

$$\begin{aligned}
&\mathrm{BR}(B \to J/\psi)\{\mathrm{total}\} = (11.5 \pm 0.6) \times 10^{-3} \\
&\mathrm{BR}(B \to J/\psi)\{\mathrm{direct}\} = (8.0 \pm 0.8) \times 10^{-3} \\
&\mathrm{BR}(B \to J/\psi)\{\chi_{c1}\} = \mathrm{BR}(B \to \chi_{c1})\mathrm{BR}(\chi_{c1} \to J/\psi) = (1.14 \pm 0.27) \times 10^{-3} \\
&\mathrm{BR}(B \to J/\psi)\{\chi_{c2}\} = \mathrm{BR}(B \to \chi_{c2})\mathrm{BR}(\chi_{c2} \to J/\psi) = (1.3 \pm 0.7) \times 10^{-4} \\
&\mathrm{BR}(B \to J/\psi)\{\psi(2S)\} = \mathrm{BR}(B \to \psi(2S))\mathrm{BR}(\psi(2S) \to J/\psi) = (1.92 \pm 0.37) \times 10^{-3}
\end{aligned} \tag{4.7}$$

4.5.2 χ_{c1} and χ_{c2} selection.

Both χ_{c1} and χ_{c2} decays into J/ψ predominantly through the radiative process $\chi_c \to \gamma J/\psi$. The almost identical decay signatures allows us to treat both decay channels in the same way and disregard effects of overlapping mass spectrums for χ_{c1} and χ_{c2}

resonances. In what follows χ_c reconstruction will mean reconstruction through both χ_{c1} and χ_{c2} channels.

In order to suppress the most obvious backgrounds we set the likelihood threshold for J/ψ $L_{J/\psi} > 0.03$ and for photons $L_\gamma > 0.03$. The kinematic end-point for χ_c production in B meson decays (Section B.3) allows us to require the momentum of the $J/\psi\ \gamma$ pair to be less than 1.75 GeV/c and to impose a mass window 3.42 GeV/c^2 < $M(J/\psi\ \gamma) - M(J/\psi) + 3.096$ GeV/c^2 < 3.60 GeV/c^2. The likelihood function for the χ_c candidates (L_{χ_c}) is constructed in a regular way, using the following two-dimensional discriminants:

- Photon likelihood vs. photon HPL (to suppress $\pi^0 \to \gamma\gamma$) .
- χ_c decay angle vs. χ_c momentum.
- χ_c mass vs. χ_c momentum in the CM frame.
- J/ψ helicity vs. J/ψ decay angle.

The performance of the likelihood based on a 2.5-M-event inclusive J/ψ MC sample is summarized in Fig. 4.10. The overall separation power is not great, with only 58% of the signal coming in the signal region. This is mostly due to the high level of the photon combinatorial background.

This decay chain also serves as a nice introduction to the challenges we face with K_L selection and yield estimates. The inclusive χ_c selection criteria are tuned on inclusive J/ψ MC, and therefore require a robust estimate for non-inclusive J/ψ component (combinatorial background). In addition, we do not have full confidence in the photon (or the K_L) background representation in the MC. Despite this lack of confidence, we assume that the overall signal/background shape is valid and only the true/fake photon fractions have to be corrected. That forces us to estimate the fake/true photon ratio as well as the combinatorial contamination from the data.

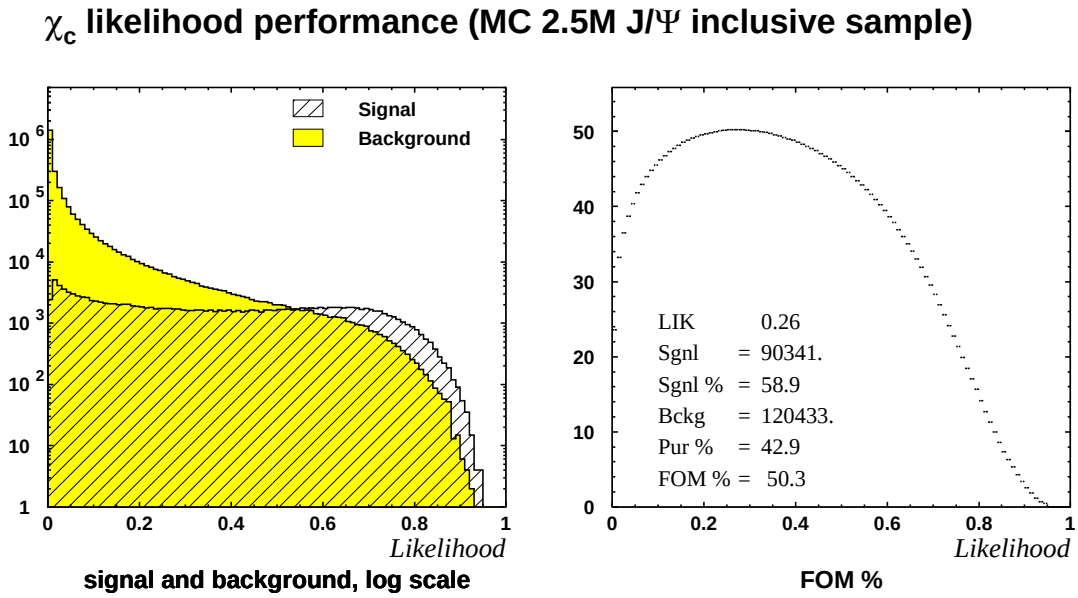

Figure 4.10: The χ_c likelihood performance ($J/\psi\ \gamma$). The plots are based on a 2.5-M-event J/ψ inclusive MC sample. The left plot shows the distribution for the signal and the background events on a logarithmic scale. The right plot shows the FOM for events with likelihood greater than the abscissa value. The values presented on the plot show the maximum value for the FOM, the likelihood at which it was achieved $L_{\max}$, the number of signal events for $L > L_{\max}$ and the percentages of signal and background events satisfying $L > L_{\max}$.

For inclusive decays we have only one free parameter to estimate the signal yield, and mass is the variable of choice. Our approach is based on likelihood, which incorporates mass in its construction, as a more powerful variable for signal/background yield estimates. While switching to an analysis solely based on likelihood, we are presenting yield estimates based on the mass variable alongside a number based on likelihood. All subsequent signal/background yield estimates will be presented in a form similar to that of Fig. 4.11, which represents the estimates of the signal and background fractions yield in both mass and likelihood variables for the inclusive χ_c sample from the data. The top left picture shows shapes and estimates for the mass variable, while the top right utilizes the likelihood. The bottom left picture uses simultaneous approximation in both mass and likelihood variables. The bottom right picture uses yield results from the bottom left picture and shapes corresponding to the full likelihood $L = L_{\chi_c} \otimes L_{J/\psi}$.

We have to discuss this plot in detail because it will be a cornerstone for the signal/background estimation in the $B \to J/\psi\ K_L$ decay reconstruction. We start from combinatorial background.

Estimating the shape of the spectrum for combinatorial J/ψ events is a multi-step process. As a first step, we extract two sub-samples from the χ_c candidate sample. The first sample represents the combinatorial background and has $L_{J/\psi} < 0.20$, while the second one corresponds to the inclusive J/ψ sample with $L_{J/\psi} > 0.90$. For both sub-samples and the overall sample we fill the histograms with the number of bins, $N_{\rm bins}$, and with respect to the variable of interest X. We refer to normalized histograms as shapes. The shapes for overall, inclusive and combinatorial samples are denoted by $S_{\rm tot}$, $S_{\rm incl}$, $S_{\rm comb}$. The corresponding yields are $N_{\rm tot}$, $N_{\rm incl}$, $N_{\rm comb}$. Fraction size can be estimated by finding the solution for the following equation.

$$N_{\rm tot} S_{\rm tot}(X) = N_{\rm incl} S_{\rm incl}(X) + N_{\rm comb} S_{\rm comb}(X) \tag{4.8}$$

Eq. 4.8 is effectively a set of $N_{\rm bins}$ independent linear equations with two unknown

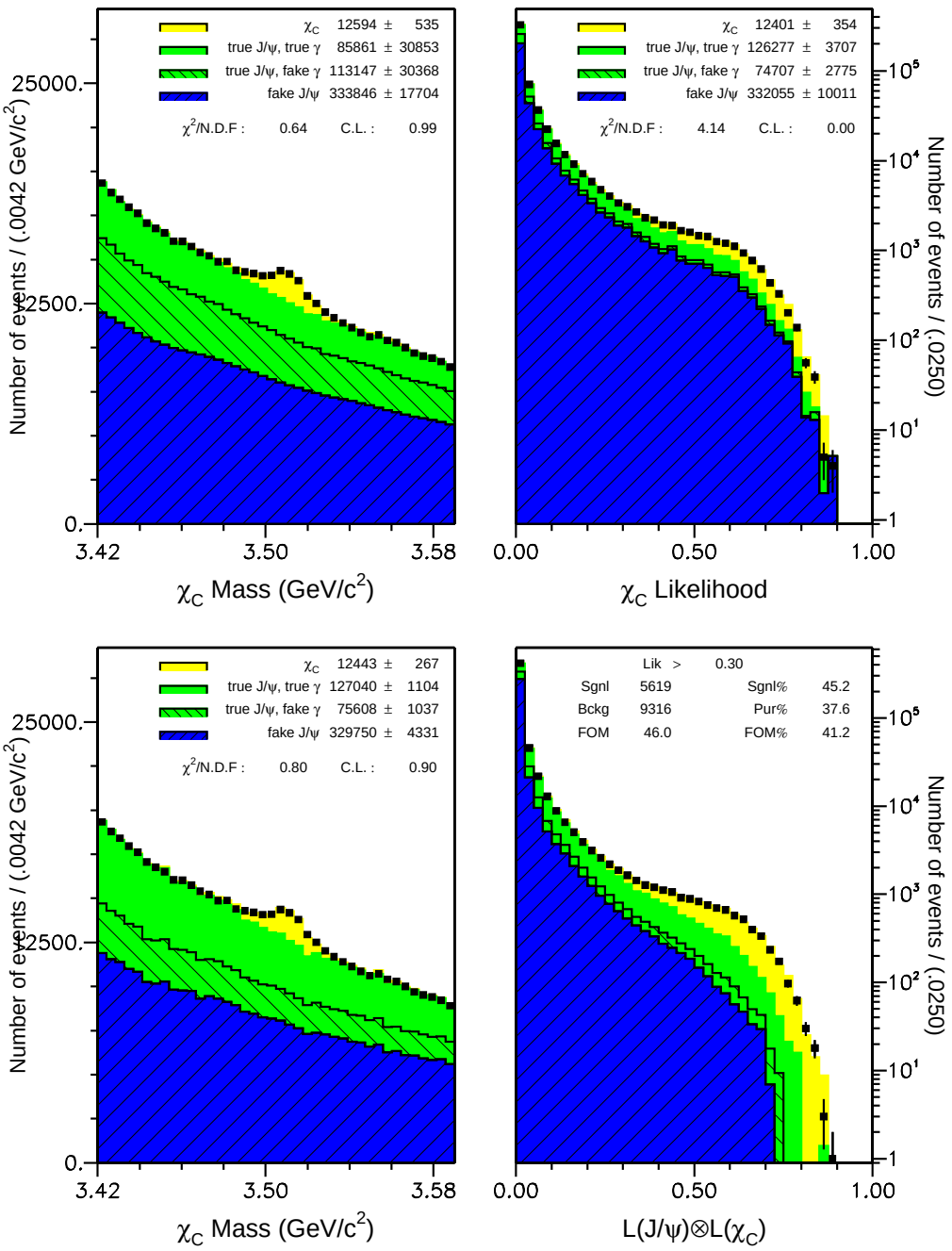

Figure 4.11: Estimates for the $\chi_c \to J/\psi\ \gamma$ signal, inclusive J/ψ background and combinatorial J/ψ background yield. All plots are based on a 78.13 fb^{-1} on-resonance data sample with $L_{J/\psi} > 0.03$ and $L_\gamma > 0.03$. The shape of the combinatorial background is based on a sample with $L_{J/\psi} < 0.25$ and inclusive J/ψ shape from $L_{J/\psi} > 0.90$. Top left plot shows yield estimates using mass variable, while top right gives estimates using likelihood. Results on the bottom left plot are based on both mass and likelihood variables. Bottom right plot shows shapes for full likelihood along with likelihood cutoff value and yields, which maximize the FOM.

parameters $N_{\rm incl}$ and $N_{\rm comb}$, which can be estimated by minimizing either χ^2 or the likelihood. Such a procedure can be performed in any variable, but in the subsequent analysis we use both mass (the traditional approach) and likelihood variables. However, such an estimate for $N_{\rm comb}$ includes contamination from the inclusive sample with $L_{J/\psi} < 0.20$. In order to account for it we also fill MC shapes corresponding to the inclusive and combinatorial shape criteria $S^{MC}_{\rm incl}$, $S^{MC}_{\rm comb}$ with yield $N^{MC}_{\rm incl}$ and $N^{MC}_{\rm comb}$. The corrected combinatorial shape then takes the form of:

$$S'_{\rm comb}(X) = S_{\rm comb}(X) - \frac{N^{MC}_{\rm comb}}{N_{\rm comb}} \frac{N_{\rm incl}}{N^{MC}_{\rm incl}} S^{MC}_{\rm comb}(X) \tag{4.9}$$

After obtaining the shape for the combinatorial background we are ready for the yield estimate of all background and signal components. The shapes for the signal and the backgrounds with fake and true photons are based on MC samples. We denote shapes for the signal and the inclusive backgrounds with fake and true photons as $S_{\rm s}$, $S_{\rm bif}$, and $S_{\rm bit}$, respectively. The corresponding yields are denoted as $N_{\rm s}$, $N_{\rm bif}$, and $N_{\rm bit}$. In order to enforce the proper combinatorial yield, we minimize simultaneously the two equations presented in Eq. 4.10.

$$\begin{aligned} N_{\rm tot}S_{\rm tot}(X) &= N_{\rm incl}S_{\rm incl}(X) + N_{\rm comb}S'_{\rm comb}(X) \\ N_{\rm tot}S_{\rm tot}(X) &= N_{\rm s}S_{\rm s}(X) + N_{\rm bif}S_{\rm bif}(X) + N_{\rm bit}S_{\rm bit}(X) + N_{\rm comb}S'_{\rm comb}(X) \end{aligned} \tag{4.10}$$

The results of the minimization for both mass and likelihood variables are presented in the top two plots in Fig. 4.11. However, one can see that shapes for the combinatorial and for the inclusive background with true and fake photons all have similar shapes in the mass variable. The combinatorial background is restricted by the first equation in Eq. 4.10, but inclusive backgrounds are free to float, and as a result produce large errors in the corresponding yields. On the other hand, the χ_c likelihood has the photon likelihood incorporated in it. That makes the shapes for fake and true photons quite distinct and as a result the overall errors are much smaller for the likelihood based yields. The agreement between the signal and combinatorial yield for both mass and likelihood variables is excellent. Despite all the advantages of the likelihood

Sample	true γ	fake γ	inclusive	combinatorial	total
MC	8.85	7.88	16.73	-	-
Data	10.21 ± 0.24	6.08 ± 0.15	16.29 ± 0.28	26.5 ± 0.7	42.8 ± 0.7

Table 4.4: Ratio B/S for different χ_c background fractions. The inclusive background comprises true J/ψ with both true and fake γ's, while the total background is a sum of inclusive and combinatorial backgrounds.

approach, we decided to be conservative in the yield estimate and base our final numbers on simultaneous minimization for both likelihood and mass variables. That combines the better likelihood-based estimate for fake/true photon backgrounds while retaining the dependability of the mass based approach to signal and combinatorial background yield. The results of mass-likelihood minimization are presented in the bottom left plot of Fig. 4.11. In this analysis we use the full likelihood $L = L_{\chi_c} \otimes L_{J/\psi}$ for final event selection. In order to estimate final performance we are using the shapes predicted by MC and yields produced by mass-likelihood optimization. The combinatorial shape is obtained by subtracting yield normalized MC shapes from the overall data.

The results of signal and background yield estimates are presented in Fig. 4.11. The corresponding background to signal ratio is shown in Table 4.4. Despite the large amount of background in the original sample (97.5%), the likelihood treatment makes it possible to select 45% of the inclusive χ_c sample with a purity of 37.8%. The differences between MC and data for background composition are in the range of 20%. This number, though large, does not affect the signal yield estimates.

4.5.3 ψ(2S) selection.

$\psi(2S)$ decays into J/ψ predominantly through a two-pion transition $\psi(2S) \to \pi\,\pi\,J/\psi$. In the current analysis we do not consider the neutral pion transition, because $\psi(2S) \to J/\psi\ \pi^0\ \pi^0$ has a smaller branching fraction, a low reconstruction efficiency

Sample	true $\pi^+\pi^-$	fake $\pi^+\pi^-$	inclusive	combinatorial	total
MC	2.62	7.99	10.51	-	-
Data	4.25 ± 0.15	11.09 ± 0.21	15.34 ± 0.26	20.9 ± 0.5	36.3 ± 0.6

Table 4.5: Ratio B/S for different $\psi(2S)$ background fractions. Inclusive backgrounds comprise true J/ψ with both true and fake $\pi^+\pi^-$ pairs, while the total background is a sum of inclusive and combinatorial backgrounds.

($\approx 30\%$) and high backgrounds.

Here, $\psi(2S)$ candidates are reconstructed by combining a J/ψ meson with a pair of charged tracks with opposite charge. In order to suppress the most obvious background, we set the likelihood threshold for J/ψ $L_{J/\psi} > 0.03$. The kinematic end-point for $\psi(2S)$ production in B meson decays (Section B.3) allows us to require the momentum of the J/ψ γ pair to be less than 1.50 GeV/c. The reconstructed mass is required to be in the mass window 3.665 GeV/c^2 $< M(J/\psi\,\pi^+\pi^-) - M(J/\psi) + 3.096$ GeV/c^2 $<$ 3.700 GeV/c^2. The likelihood function for $\psi(2S)$ candidates ($L_{\psi(2S)}$) is constructed in the standard way and is based on the following two dimensional discriminants:

- Di-pion mass vs. $\psi(2S)$ momentum.
- π^+ likelihood vs. π^- likelihood.
- IP probability for π^+ vs. IP probability for π^-.
- $\psi(2S)$ decay angle with respect to J/ψ vs. $\psi(2S)$ decay angle with respect to high momentum π
- $\psi(2S)$ mass vs. $\psi(2S)$ momentum in the CM frame.

The performance of the likelihood on a 5.5-M-event inclusive J/ψ MC sample is summarized in Fig. 4.12. The overall separation power is good, with 76% of the signal coming in the signal region.

The results of signal and background yield estimates are presented in Fig. 4.13 with the corresponding background to signal ratio shown in Table 4.5. Despite the large

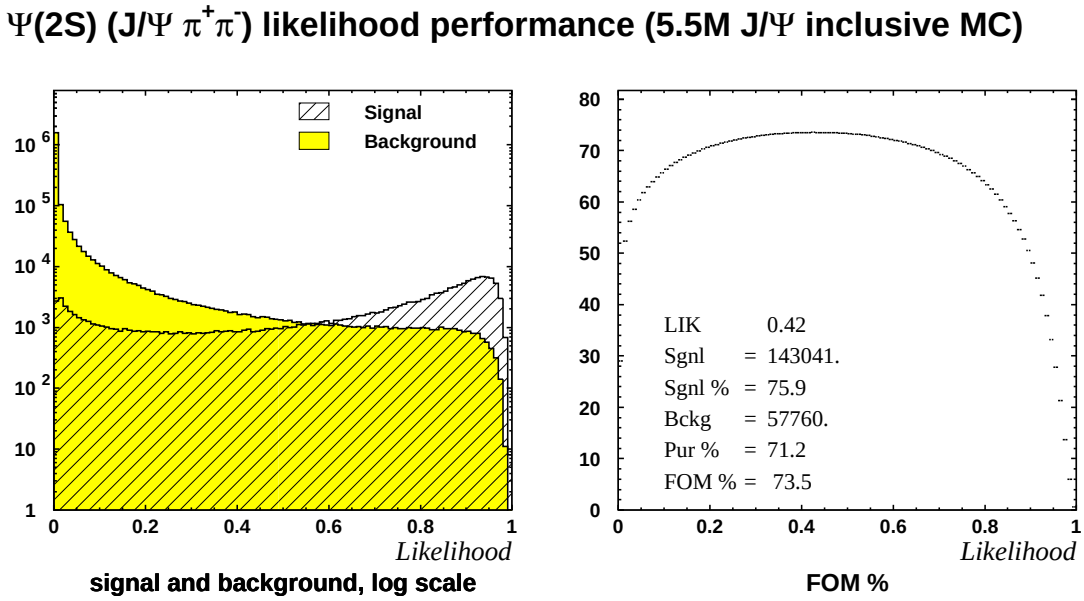

Figure 4.12: The $\psi(2S)$ likelihood performance ($J/\psi\ \pi^+\pi^-$). The plots are based on a 5.5-M-event J/ψ inclusive MC sample. The left plot shows the distribution for signal and background events on a logarithmic scale. The right plot shows the FOM for events with likelihood greater than the abscissa value. The values presented on the plot show the maximum value for the FOM, the likelihood at which it was achieved $L_{\max}$, the number of signal events for $L > L_{\max}$ and the percentages of signal and background events satisfying $L > L_{\max}$.

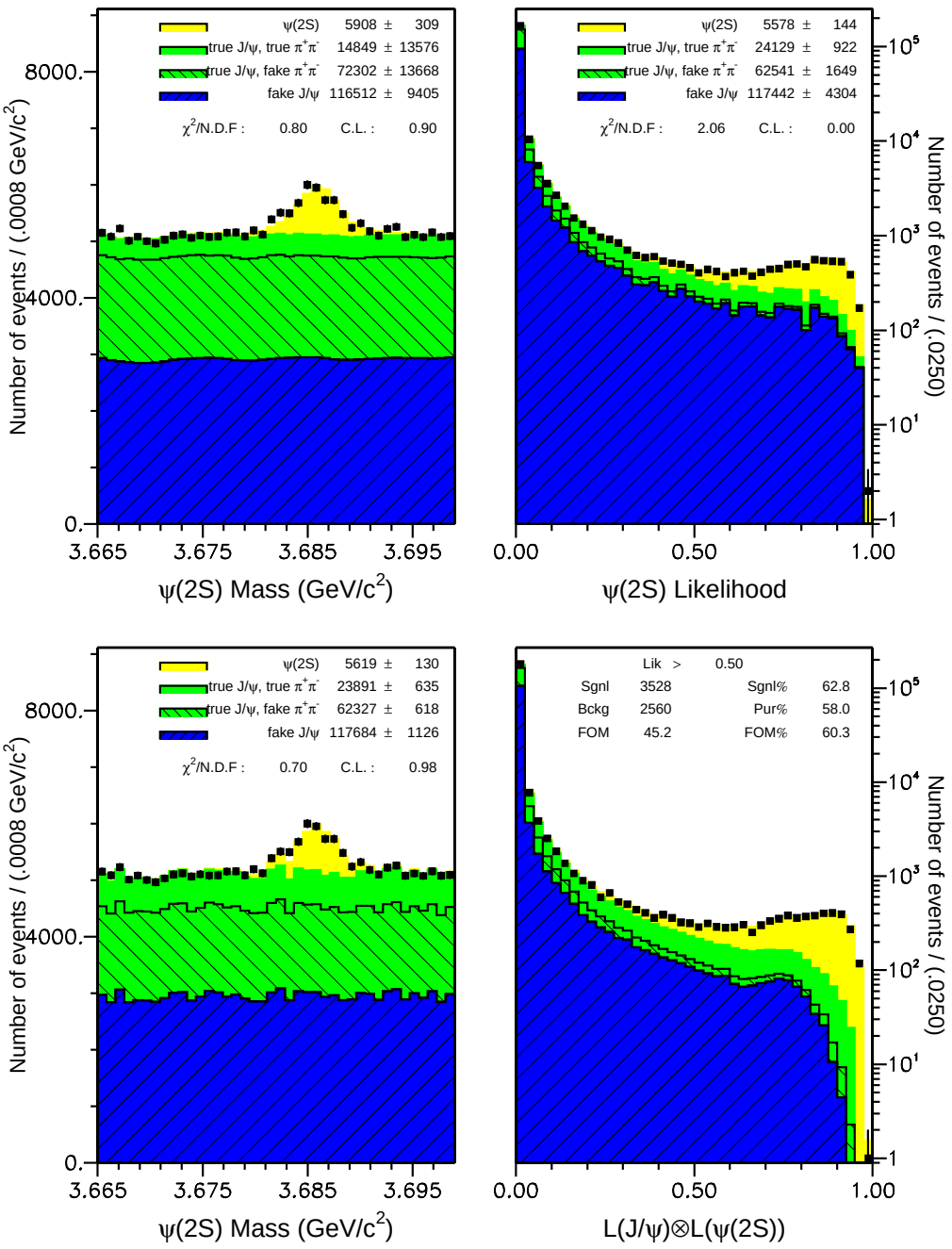

Figure 4.13: The $\psi(2S)$ yield estimates ($J/\psi\ \pi^+\pi^-$). Estimates for the signal, inclusive J/ψ background and combinatorial J/ψ background yield. All plots are based on a 78.13 fb^{-1} on-resonance data sample with $L_{J/\psi} > 0.03$. The combinatorial background shape is based on a sample with $L_{J/\psi} < 0.25$ and the inclusive J/ψ shape from $L_{J/\psi} > 0.90$. Top left plot shows yield estimates using mass variable, while top right gives estimates using likelihood. Results on the bottom left plot are based on both mass and likelihood variables. Bottom right plot shows shapes for full likelihood along with likelihood cutoff value and yields, which maximize the FOM.

amount of background in the original sample (97%), the likelihood treatment makes it possible to select 62.8% of the inclusive $\psi(2S)$ sample with a purity of 58%. The $\psi(2S)$ sample, like the one for χ_c, exhibits a discrepancy between the background fraction predicted by the MC and the one measured in the data. However, the fake/true ratio in the $\psi(2S)$ sample is much closer to the value predicted by MC. This could be attributed to better representation of charged tracks compared to neutral tracks in the MC. The larger-than-expected true inclusive backgrounds for both χ_c and $\psi(2S)$ are consistent with a lack of knowledge of higher level-resonances in $B_d \to (c\bar{c})(s\bar{d})$ transitions. Our MC sample has a known deficit of low momentum J/ψ mesons, which should have higher multiplicity for pions and photons.

4.6 Charged particle identification

In this analysis we use standard Belle particle identification software for assessing the signal probability for π, K, μ and e. However, we scale the kaon and pion probability functions to account for the large number of leptons in the J/ψ inclusive sample. An overall normalization is performed to match the signal-to-background ratio for the inclusive J/ψ MC sample. Hadron identification for charged tracks is based on the response from the ACC, the TOF and ionization measurements in the CDC. Electron identification relies heavily on the energy resolution of the ECL and the momentum measurements from the CDC with information from the TOF, the ACC and specific ionization also taken into account. The major device for muon identification is the KLM. Charged particle identification is described in detail elsewhere [22, 23, 66].

4.7 Photon selection

4.7.1 Introduction

Reliable identification of particles interacting within the electromagnetic calorimeter is crucial to the K_L analysis, because more than 50% of K_L's coming from B meson decays interact in the calorimeter. The largest contamination is caused by photons and charged tracks. MC simulation shows that the average $B\bar{B} \to J/\psi + X$ decay produces ~ 7 detectable photons and ~ 8 charged tracks with an energy deposition in the calorimeter. Over 90% of the photons come from π^0 decays, thus forcing us to pay special attention to π^0 selection. In the following section we start by selecting clusters induced by charged tracks and follow with an overview of photon and π^0 selection. We will also discuss the selection criteria for the conversion photons that produce an e^+e^- pair. But we first describe the clustering algorithm and introduce cluster properties.

4.7.2 Cluster properties.

For every event the ECL subdetector provides an energy measurement in each hit crystal. The low-level reconstruction software combines crystal hits into clusters and these clusters are available for further study. The shower (cluster) is a group of N crystal hits. The reconstruction software treats each hit crystal as a photon, i.e. the shower is formed by combining the four vectors of photons with energy defined by the energy deposited in each corresponding crystal and direction constrained by the position of the corresponding crystal center and the beam interaction point. Crystal hits with low energy depositions (below 0.1 MeV) are rejected in order to minimize electronic noise, the seed crystal for the shower required to exceed 5 MeV. The large amount of low-energy background forces us to reject all clusters with an energy below 20 MeV. That cuts away the crystal hits with low energy deposition at the cluster

edge and results in lower energy resolution for the photons. That effect is called shower leakage and also affects hadronic clusters. On average, hadronic showers are bigger and for the same momentum are less energetic, thus increasing losses from rejected crystal hits with low energy deposition.

In order to quantify cluster properties we must introduce several definitions. Many cluster properties coincide or correlate with definitions for standard properties of distributions, like mean and variance. The cluster energy is defined as the first moment of the energy distribution for all ECL crystal hits combined into a cluster.

$$E = \sum_{i=1}^{N} E_i \tag{4.11}$$

where E_i represents the energy deposited in each of the crystals. The cluster position is the energy weighted first moment:

$$\left\langle \vec{X} \right\rangle_E = \frac{\sum\limits_{i=1}^{N} E_i \vec{X}_i}{\sum\limits_{i=1}^{N} E_i} = \frac{\sum\limits_{i=1}^{N} E_i \vec{X}_i}{E} \tag{4.12}$$

where $\vec{X}_i$ is the position vector for each crystal hit (x_i, y_i, z_i). Throughout this document we use $< \cdots >$ for simple averaging and $< \cdots >_E$ for averaging using the measured E. The cluster width is the second moment of the energy distribution.

$$w_E^2 = \frac{\sum\limits_{i=1}^{N} (E_i \vec{X}_i)^2 \Big) - \Big(\sum\limits_{i=1}^{N} E_i \vec{X}_i \Big)^2}{\Big(\sum\limits_{i=1}^{N} E_i \Big)^2} = \left\langle \vec{X}^2 \right\rangle_E - \left\langle \vec{X} \right\rangle_E^2 \equiv \left\langle \left(\vec{X} - \left\langle \vec{X} \right\rangle_E \right)^2 \right\rangle_E \tag{4.13}$$

The shower mass is the invariant mass, assuming that the N hit crystals represents N photons. That variable includes both simple geometrical effects of the shower development in the calorimeter as well as a physics behind the shower formation.

$$m^2 = E^2 - p^2 = \left(\sum_{i=1}^{N} E_i \right)^2 - \left(\sum_{i=1}^{N} E_i \frac{\vec{X}_i}{r_i} \right)^2 = E^2 - E^2 \left\langle \frac{\vec{X}}{r} \right\rangle_E^2 \tag{4.14}$$

where $r_i = \sqrt{x_i^2 + y_i^2 + z_i^2}$. Noting that $\left(\frac{\vec{X}}{r}\right)^2 = 1$, we can rewrite the above equation in the following form

$$E^2 - E^2 \left\langle \frac{\vec{X}}{r} \right\rangle_E^2 = E^2 \left[\left\langle \left(\frac{\vec{X}}{r}\right)^2 \right\rangle_E - \left\langle \frac{\vec{X}}{r} \right\rangle_E^2 \right] = E^2 \left\langle \left(\frac{\vec{X}}{r} - \left\langle \frac{\vec{X}}{r} \right\rangle \right)^2 \right\rangle_E \tag{4.15}$$

Most of the good ECL clusters are well localized and as a result r_i does not vary much. In this approximation one can replace it with $\left\langle \vec{X} \right\rangle_E$ and pull it outside of the averaging brackets:

$$E^2 \left\langle \left(\frac{\vec{X}}{r} - \left\langle \frac{\vec{X}}{r} \right\rangle \right)^2 \right\rangle_E \approx \frac{E^2}{\left\langle \vec{X} \right\rangle_E} \left\langle \left(\vec{X} - \left\langle \vec{X} \right\rangle_E \right)^2 \right\rangle_E = \frac{E^2}{\left\langle \vec{X} \right\rangle_E} w_E^2 \tag{4.16}$$

The shower width (w_E^2) and the shower mass (m^2) are strongly correlated, but for historic reasons we use the shower mass for selection of the fast neutral pions. Decay products of ultra-relativistic pions often form one cluster. However, such clusters still tend to have a mass approximately equal to the mass of the π^0, thus allowing good particle identification.

4.7.3 Selection criteria for the clusters induced by charged tracks.

Charged tracks always deposit energy in the calorimeter while traversing CsI crystals. This energy deposition forms three types of clusters: minimum ionizing (μ, π, K), hadronic shower (π, K) and electromagnetic shower (e). The strong correlation between charged-track momentum and cluster energy in combination with the abundance of leptons in our sample forces us to dedicate special treatment to the electron-induced clusters. The minimum ionizing and hadronic clusters are treated together, because they share the same set of discriminating variables.

In order to find ECL clusters associated with charged tracks we extrapolate the trajectory for each charged track from the central drift chamber boundary to the

electromagnetic calorimeter. This extrapolation takes into account the highly non-uniform magnetic field in the vicinity of the main super-conducting magnet and the flux return structure and interactions with the inner detector and the calorimeter. In the track extrapolation algorithm we use a π hypothesis and assume minimum-ionizing energy losses. We also assume that the cluster induced by a minimum ionizing particle will have a center position corresponding to the half distance between its entry and exit point from electromagnetic calorimeter.

In the first step, we create a likelihood for showers induced by either the μ, π or K. This likelihood ($L^{ECL}_{\mu,\pi,K}$) is based on the following two-dimensional discriminants:

- Distance between cluster position and extrapolation versus track momentum.
- Cluster energy versus E9/E25.
- Number of crystal hits versus the width of the cluster.
- Angle to closest KLM cluster versus azimuthal angle.
- Cluster width versus difference between the cluster energy deposition and energy loss predicted by the track extrapolation.

The good tracking precision of the CDC allows us to extrapolate the charged-particle trajectory into the calorimeter and test it against the ECL cluster position. In addition, minimum ionizing particles tend to deposit between 100 and 200 MeV/c^2 of energy in well contained compact clusters. Minimum ionizing particles have a high probability of penetrating into the KLM detector and forming traces there, thus ECL cluster alignment with KLM hits can provide additional separation power. The large beam background, however, tends to populate the regions of the KLM with low polar angle. All of the above allows us to select clusters induced by minimum ionizing particles with good purity. The performance of $L^{ECL}_{\mu,\pi,K}$ is summarized in Fig. 4.14.

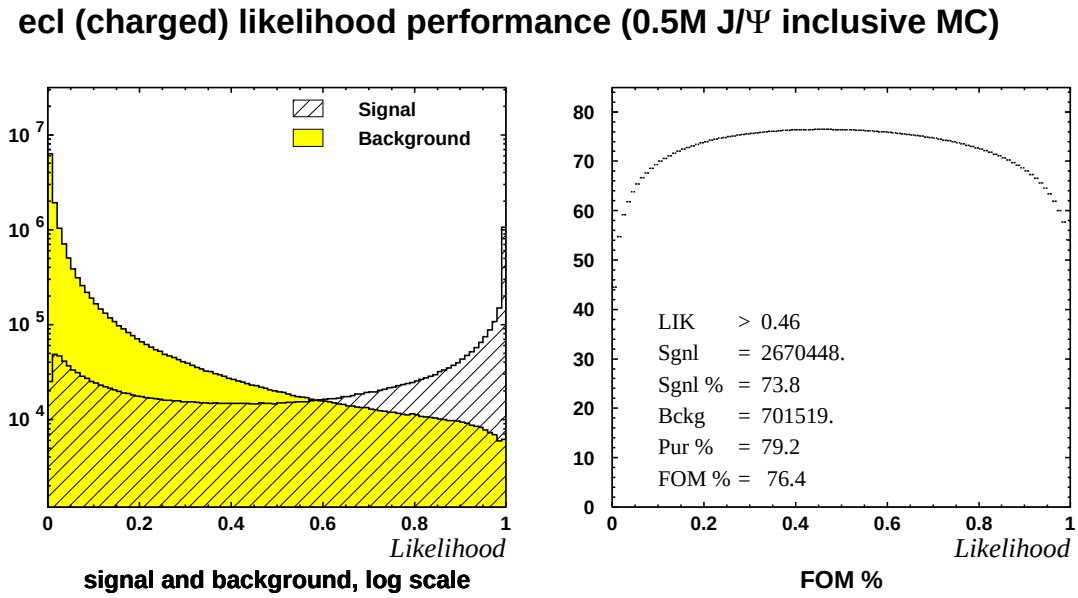

Figure 4.14: The charged track likelihood performance (ECL). The plots are based on a 0.5-M-event J/ψ inclusive MC sample. The left plot shows the distribution for signal and background events on a the logarithmic scale. The right plot shows the FOM for events with likelihood greater than the abscissa value. The values presented on the plot show the maximum value for the FOM, the likelihood at which it was achieved $L_{\max}$, the number of signal events for $L > L_{\max}$ and the percentage of the ECL signal and background events satisfying $L > L_{\max}$.

The performance of the likelihood algorithm based on a 0.5-M-event sample of J/ψ MC events is summarized in Fig. 4.14. The overall separation power is good, with 74% of the signal coming in the signal region.

For electromagnetic showers we select electron-induced clusters that have $L^{ECL}_{\mu,\pi,K} < 0.95$ and create a likelihood function (L^{ECL}_{e}) based on the following two-dimensional discriminants:

- Distance between cluster position and extrapolation versus azimuthal angle.
- Cluster energy versus track momentum.
- Number of crystal hit versus width of the cluster.
- Angle with closest KLM cluster versus E9/E25.

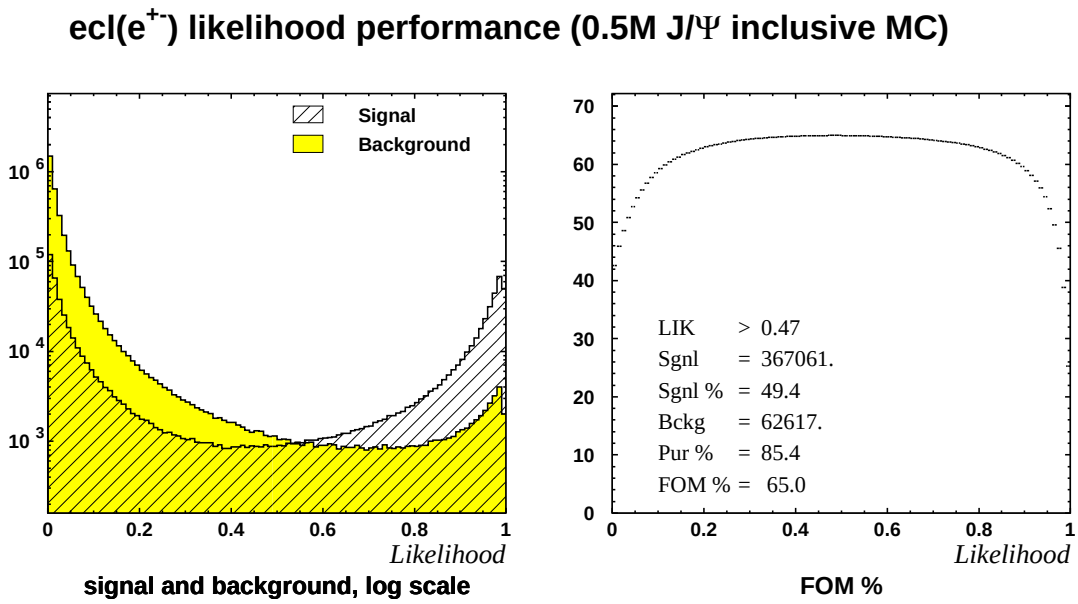

Figure 4.15: The likelihood performance for electron-induced showers (ECL). The plots are based on a 0.5-M-event J/ψ inclusive MC sample. The left plot shows the distribution for signal and background events on a logarithmic scale. The right plot shows the FOM for events with likelihood greater than the abscissa value. The values presented on the plot show the maximum value for the FOM, the likelihood at which it was achieved $L_{\max}$, the number of signal events for $L > L_{\max}$ and the percentage of the ECL signal and background events satisfying $L > L_{\max}$.

- Cluster width versus difference between the cluster energy deposition and energy loss predicted by the track extrapolation.

The performance of the likelihood based on a 0.5-M-event inclusive J/ψ MC sample is summarized in Fig. 4.15. The overall separation power is satisfactory, with 50% of the signal coming in the signal region. The reason for the low separation power lies in the large background from low-momentum pions, which tend to deposit all of their energy in the calorimeter.

4.7.4 Selection criteria for the clusters induced by photons.

For the photon-induced showers we select clusters that have $L^{ECL}_{\mu,\pi,K,e} < 0.95$ and create a likelihood function (L^{ECL}_{γ}) based on the following two-dimensional discriminants:

- Distance between cluster position and closest charged track extrapolation vs. track momentum.
- Cluster energy vs. E9/E25.
- Number of crystals hit vs. width of the cluster.
- Angle with closest KLM cluster vs. cluster azimuthal angle.

Despite vetoing clusters with close association to charged tracks, we still need to include information about the closest track momentum and extrapolated position. This allows us to suppress split-off clusters from hadronic interactions and backgrounds related to misreconstructed tracks. The angle with respect to the closest KLM cluster is used to suppress neutral-hadron backgrounds. The cluster properties work to distinguish photon clusters from all types of backgrounds.

The performance of the likelihood based on a 0.5-M-event inclusive J/ψ MC sample is summarized in Fig. 4.16. The overall separation power is excellent, with 85% of the signal coming in the signal region.

4.7.5 Selection criteria for photons producing electron pairs.

Conversion photons contribute less than 1% of all photons produced in B meson decays. Nevertheless even this small increase in the photon reconstruction efficiency pays well in multi-photon final states, such as $K_S \to \pi^0\pi^0$ with $\pi^0 \to \gamma\gamma$. It also adds another testing facility for electron identification and vertex resolution. The $\gamma \to e^+e^-$ identification takes advantage of the excellent vertex resolution of the Belle detector. Photon conversion requires the presence of material and this moves the decay vertex for a large number of γ's away from the beam-interaction point. The vertex displacement is determined by the SVD detector geometry and could be as large as 15 cm. Such a distance is enough to cleanly separate tracks belonging

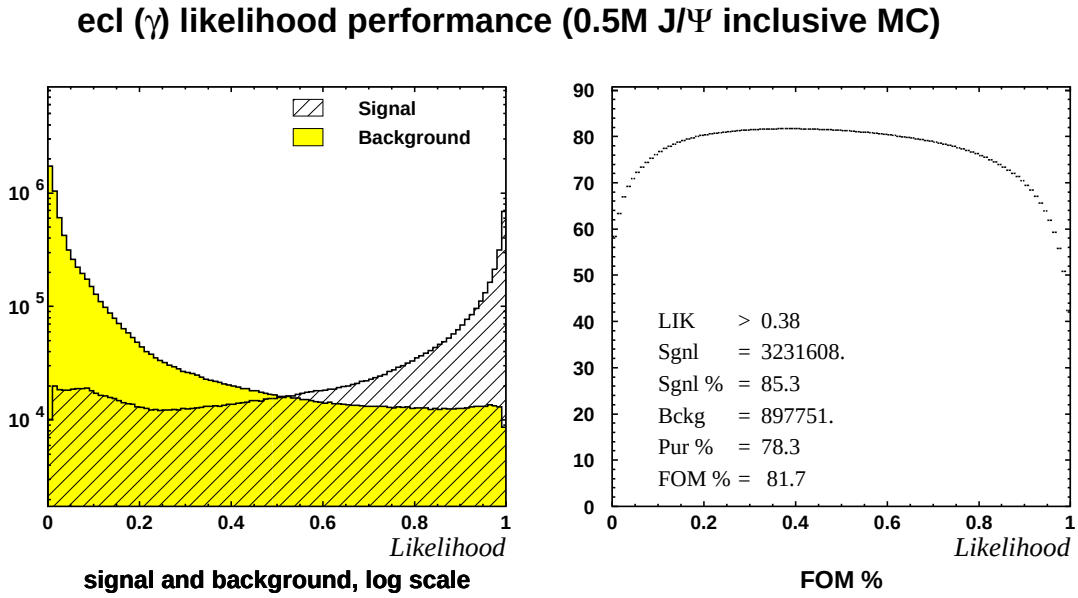

Figure 4.16: The photon likelihood performance (ECL). The plots are based on a 0.5-M-event J/ψ inclusive MC sample. The left plot shows the distribution for signal and background events on a logarithmic scale. The right plot shows the FOM for events with likelihood greater than the abscissa value. The values presented on the plot show the maximum value for the FOM, the likelihood at which it was achieved $L_{\max}$, the number of signal events for $L > L_{\max}$ and the percentage of the ECL signal and background events satisfying $L > L_{\max}$.

conversion particles from the majority of tracks coming from the IP. The γ likelihood is based on the following two dimensional discriminants:

- χ^2 of the vertex fit vs. z distance of the vertex from the IP.
- IP probability of the first track vs. IP probability of the second track.
- γ decay angle vs. decay time, assuming that the γ was produced at the beam-interaction point.
- γ mass vs. γ CM momentum.
- e probability of the first track vs. e probability of the second track.

The performance of the likelihood based on a 0.5-M-event inclusive J/ψ MC sample is summarized in Fig. 4.17. The overall separation power is excellent with 91% of the signal coming in the signal region.

4.7.6 Selection criteria for the π^0.

The π^0 decays into a two-photon final state with 98.8% probability, the remaining 1.2% is represented by three-body decay $\pi^0 \to e^+e^-\gamma$. That three body decay was overlooked during reconstruction stage and as a result is not included in this analysis. This omission results in a lower reconstruction efficiency for particle with π^0 among the decay products but does not affect the asymmetry measurements. The selection performance for the π^0 is heavily dependent on the photon-selection performance. All of the photon properties are embedded in the photon-likelihood function. The most common kinematic parameters for two-body decays are the mass of the particle, its momentum and decay angle. The abundance of low-momentum photons and the high background levels associated with them makes low momentum π^0 reconstruction quite difficult. In addition, the average $B\bar{B}$ decay produces about 5 π^0s which introduce

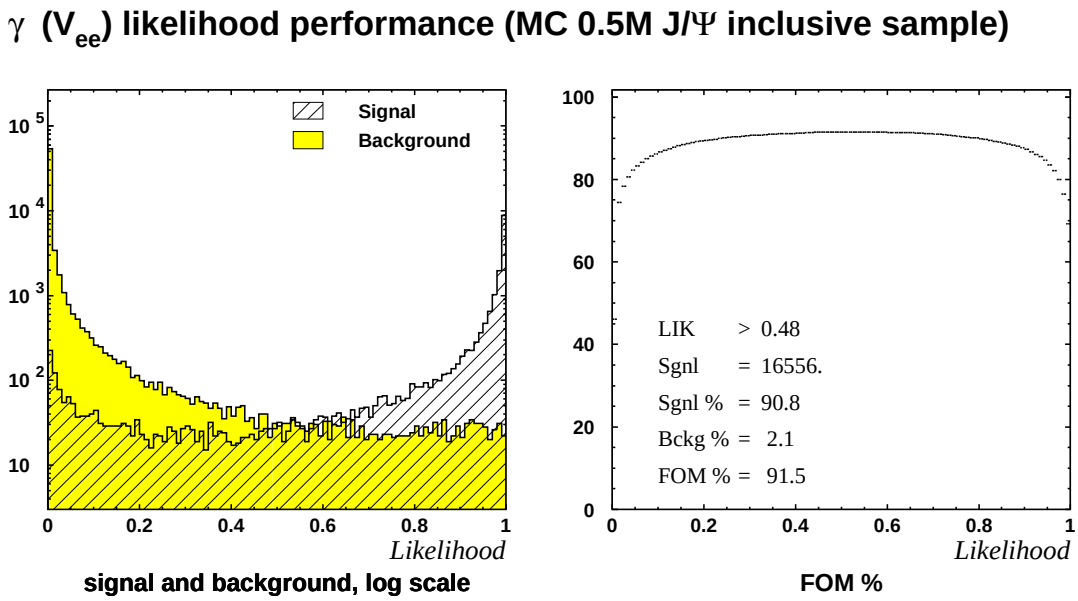

Figure 4.17: The photon likelihood performance (e^+e^-). The plots are based on a 0.5-M-event J/ψ inclusive MC sample. The left plot shows the distribution for signal and background events on a logarithmic scale. The right plot shows the FOM for events with likelihood greater than the abscissa value. The values presented on the plot show the maximum value for the FOM, the likelihood at which it was achieved $L_{\max}$, the number of signal events for $L > L_{\max}$ and the percentage of the ECL signal and background events satisfying $L > L_{\max}$.

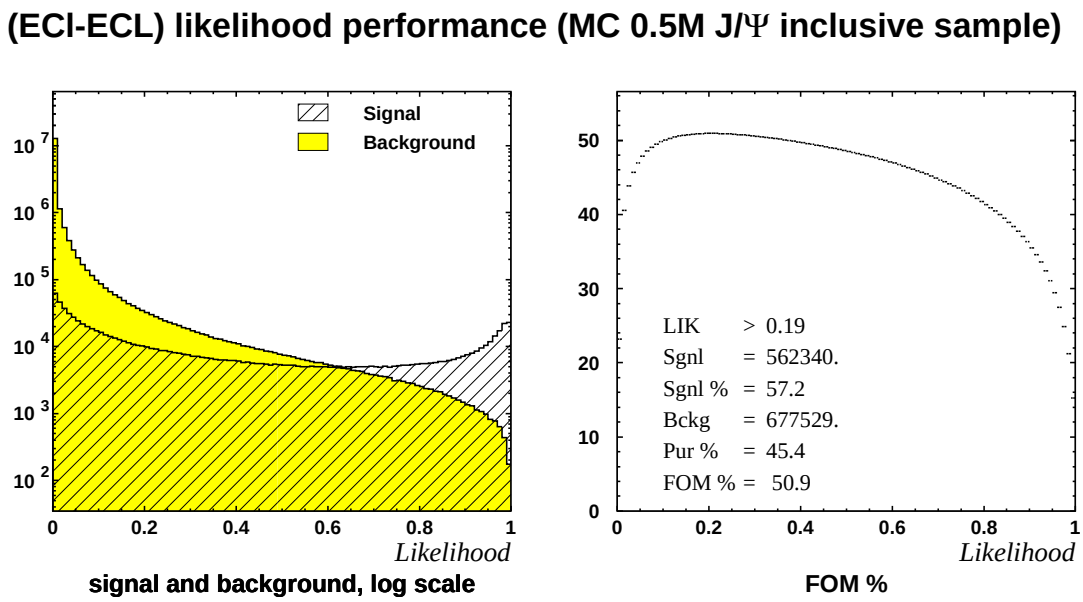

Figure 4.18: The π^0 likelihood performance. The plots are based on a 0.5-M-event J/ψ inclusive MC sample. The left plot shows the distribution for signal and background events in the logarithmic scale. The right plot shows the FOM for events with likelihood greater than the abscissa value. The values presented on the plot show the maximum value for the FOM, the likelihood at which it was achieved $L_{\max}$, the number of signal events for $L > L_{\max}$ and the percentage of ECL signal and background events satisfying $L > L_{\max}$.

high levels of combinatorial background. Despite these challenges it is possible to select 57% of the inclusive π^0 sample with 45% purity.

The π^0 likelihood is based on the following two dimensional discriminants:

- π^0 decay angle vs. π^0 momentum.
- π^0 mass vs. π^0 CM momentum.
- γ likelihood of the first photon vs. γ likelihood of the second photon.

The performance of the likelihood based on a 0.5-M-event inclusive J/ψ MC sample is summarized in Fig. 4.18. The overall separation power is satisfactory with 57% of the signal coming in the signal region.

4.8 K_S selection.

In identifying K_S's we take advantage of the excellent vertex resolution of the Belle detector. The relatively large K_S lifetime moves the decay vertex for the K_S away from the beam interaction point. The vertex displacement corresponding to the lifetime of a 1 GeV/c K_S is $\beta\gamma c\tau = 5.3$ cm. Such a distance is enough to cleanly separate tracks belonging to long-lived particles from the majority of tracks coming from the IP. The K_S likelihood is based on the following two-dimensional discriminants:

- χ^2 of the vertex fit versus z distance of the vertex from the interaction point.
- Closest radial distance between track and interaction point versus distance in the z direction.
- K_S decay angle (angle between the pion and the K_S flight direction in the K_S rest frame) versus proper decay time, assuming that K_S was produced at the beam interaction point.
- K_S mass versus K_S CM momentum.
- π probability of one track versus π probability of the second track.

The overall performance for the K_S likelihood is summarized in Fig. 4.19.

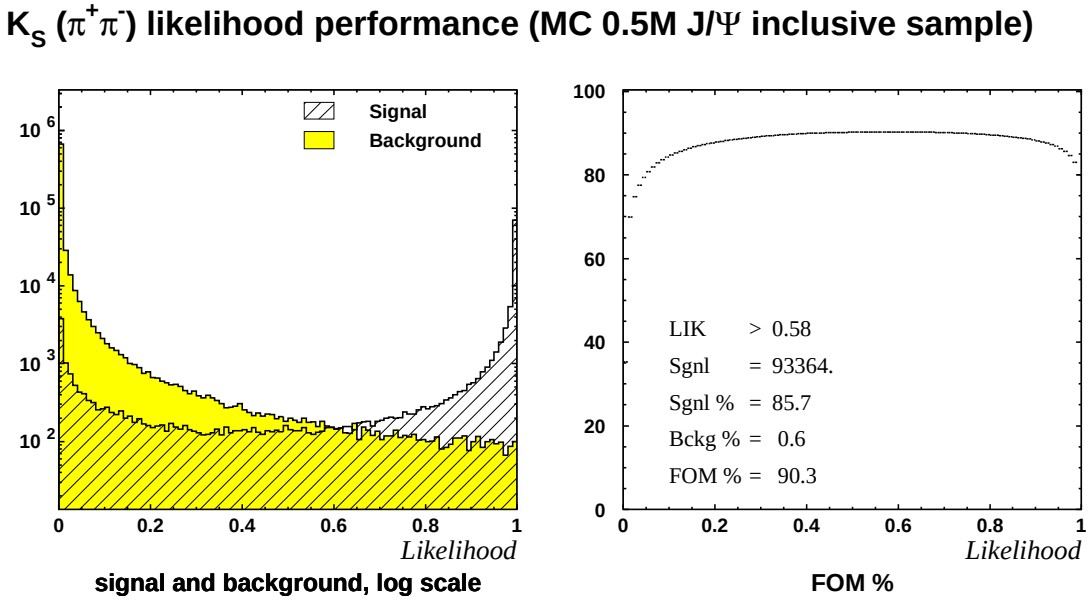

Figure 4.19: The $K_S \to \pi^+\pi^-$ likelihood performance. The plots are based on a 0.5-M-event J/ψ inclusive MC sample. The left plot shows the distribution for both signal and background events on a logarithmic scale. The right plot shows the FOM for events with likelihood greater than the abscissa value. The values presented on the plot show the maximum value for the FOM, the likelihood at which it was achieved $L_{\max}$, the number of signal events for $L > L_{\max}$ and the percentage of the ECL signal and background events satisfying $L > L_{\max}$.

4.9 K_{L} selection.

4.9.1 Introduction

This study concentrates on the $B \rightarrow J/\psi\ K_L$ decay chain. The kinematics of the Belle experiment will force K_L's from the decay above to have CM momenta greater than 1.5 GeV/c which corresponds to a lab momentum greater than 1 GeV/c. Thus we optimize K_L detection for the momentum range greater than 1 GeV/c.

K_L's can be detected through either their hadronic interactions with the material of the detector or through their decay products. A 1 GeV/c threshold limits number of K_L decays within detector boundaries to 2%. Such a small fraction of in-flight decays renders K_L reconstruction from decay products irrelevant to our study. The energy released in hadronic interactions of the K_L allows us to determine the point along the K_L trajectory. This information with the assumption that all K_L's are produced at the beam interaction point defines the K_L trajectory, but does not provide any direct information regarding the K_L's energy. The lack of energy information creates a major challenge when it comes to event reconstruction and background suppression.

There are two sub detectors capable of seeing products of K_L interactions with detector material: the electromagnetic calorimeter and the KLM. Depending on where the interaction occurs, we classify all K_L candidates into three distinct groups: ECL, ECL-KLM and KLM (Fig. 4.20). ECL candidates have products of K_L interactions with traces in only the electromagnetic calorimeter. For the ECL-KLM events the K_L starts interacting in the ECL with energy spilling into the KLM detector. And finally the KLM group embraces the K_Ls which interact only in the KLM detector. The above classifications stresses not only the separate decay signatures, but also the vast differences in the background composition. The ECL candidates are overwhelmed by photon backgrounds, while the KLM candidates are subject to a large contamination from neutrons produced along the beam line. The typical hadronic interaction produces a number of neutral and charged particles, resulting in large clusters with

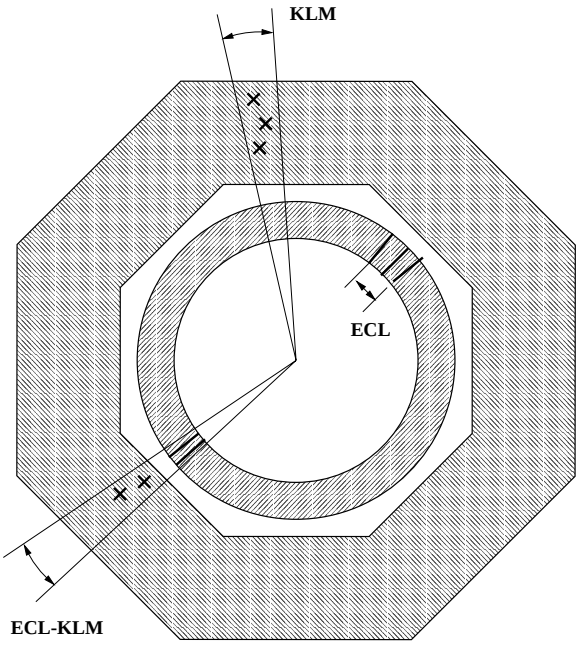

Figure 4.20: Interaction based K_L classification. Depending on where the interaction occurs, we classify all K_L candidates into three distinct groups: ECL, ECL-KLM and KLM.

irregular shape. The current analysis is based on the Belle standard cluster finding algorithm in both the ECL and KLM detectors. We have tried to compensate for the known limitations of the Belle standard codes without changes in the underlying algorithms.

4.9.2 $K_{\mathbf{L}}$ detection.

We will start with a description of the detector material, since it is there where the interactions necessary for the detection of neutral kaons take place. The central part of the detector is carefully designed to minimize the amount of material. This is a major requirement for high precision momentum measurements in the drift chamber and accurate energy measurements in the calorimeter. The total amount of material in front of the electromagnetic calorimeter is equivalent to less than 5 cm of Al. Most of it comes from the drift chamber structure, the cables and the photomultiplier tubes in the aerogel Cherenkov counters. The first significant absorption takes place in the

calorimeter, which consist of 30 cm of CsI crystals. The coil of superconducting solenoid provides less material than one layer of KLM system, and can be viewed as a first absorption layer outside ECL detector. Most neutral-kaon interactions take place in the iron of the magnetic flux return, which consists of 15 layers of 4.7 cm iron plates interleaved with 14 resistive plate chambers (RPCs). The main material inside the RPC is four glass plates with a total thickness of 1 cm per layer. Thus in total we will have 5 cm of Al followed by 30 cm of CsI, and 65.8 cm of Fe interlaced with 14 cm of glass. Neutral hadrons do not interact electromagnetically so the main contribution to their total cross section comes from nuclear interaction.

An exact description of the hadronic interaction between kaons and nuclei is rather complex and well beyond the scope of this document, so we will try to concentrate on the qualitative features of the main processes that are involved.

The total hadronic cross section is often divided into three components: elastic, quasi-elastic and absorption. As its name suggests the elastic cross section is responsible for scattering off the entire nucleus without exciting either the nucleus or the incident particle. The quasi-elastic part corresponds to the scattering from individual nucleons in the nucleus. This type of scattering does not leave the nucleus intact but preserves the incident particle. The last type of interaction is absorption, wherein the incident particle is lost and new particles are created.

Both the mean free path of a particle, λ, in a medium and its cross section, σ, can be seen as a probability measure for the particle to interact with the medium. The relation between them can be expressed in the following way: $\sigma\lambda = V/N = A/(N_A\rho)$, where V is an interaction volume with N particles. A medium is characterized by its atomic weight, A, its density, ρ, and Avogadro's number, N_{A}.

The (nuclear) interaction length is the mean free path of a particle in a given medium for undergoing an interaction that is neither elastic nor quasi-elastic (diffractive) and is designated by λ_a with relevant cross-section — σ_a . The collision length

(also known as the nuclear collision length), λ, follows from the total nuclear cross-section, σ.

This division is important for the K_L detector, because for our momentum range the energy transfer to the nucleus due to the elastic scattering is rather small and will not be observable. Quasi-elastic scattering could possibly contribute to the visible energy in the calorimeter, but will not be a factor in the KLM. As a result, the estimate of the detection efficiency uses only absorption contributions. However, we should not forget about K_S regeneration. Regeneration phenomena arise because the different total cross sections for s and $\bar{s}$ quarks destroy coherence of the K_L system, thus giving birth to the K_S component. All regenerated K_S will decay without regeneration within several cm because their flight length is much smaller than the mean free path of K^0 in media. Most of the K_S decay products will have either photons or charged tracks, thus making them visible in both the calorimeter and the KLM sub-detector.

There are no comprehensive measurements of neutral kaon interactions with material. However, well measured cross sections for charged kaons provide a good approximation to those for neutral ones. Charged kaon interactions with material were studied extensively for a Kp, Kn and Kd. The most precise data, however, are available for the high momentum range. Cross sections in the momentum accessible in the Belle experiment, $p_{K_L} < 5$ GeV/c, suffer from limited statistics, and also exhibit a complicated energy dependence due to the large contributions from resonance production (see Fig. 4.21 and Fig. 4.22). In addition, the mass eigenstates of the neutral kaon system are linear superpositions of the flavor eigenstates and the K_L cross section is approximately as an average of the cross sections for K^+ and K^-. Another complication arises from regeneration phenomena and the absence of Coulomb scattering for the neutral kaons. All of these phenomena makes straightforward estimates difficult and contribute to the discrepancies between MC and data. The K_L selection criteria have been tuned on the J/ψ inclusive MC sample with detector simulation

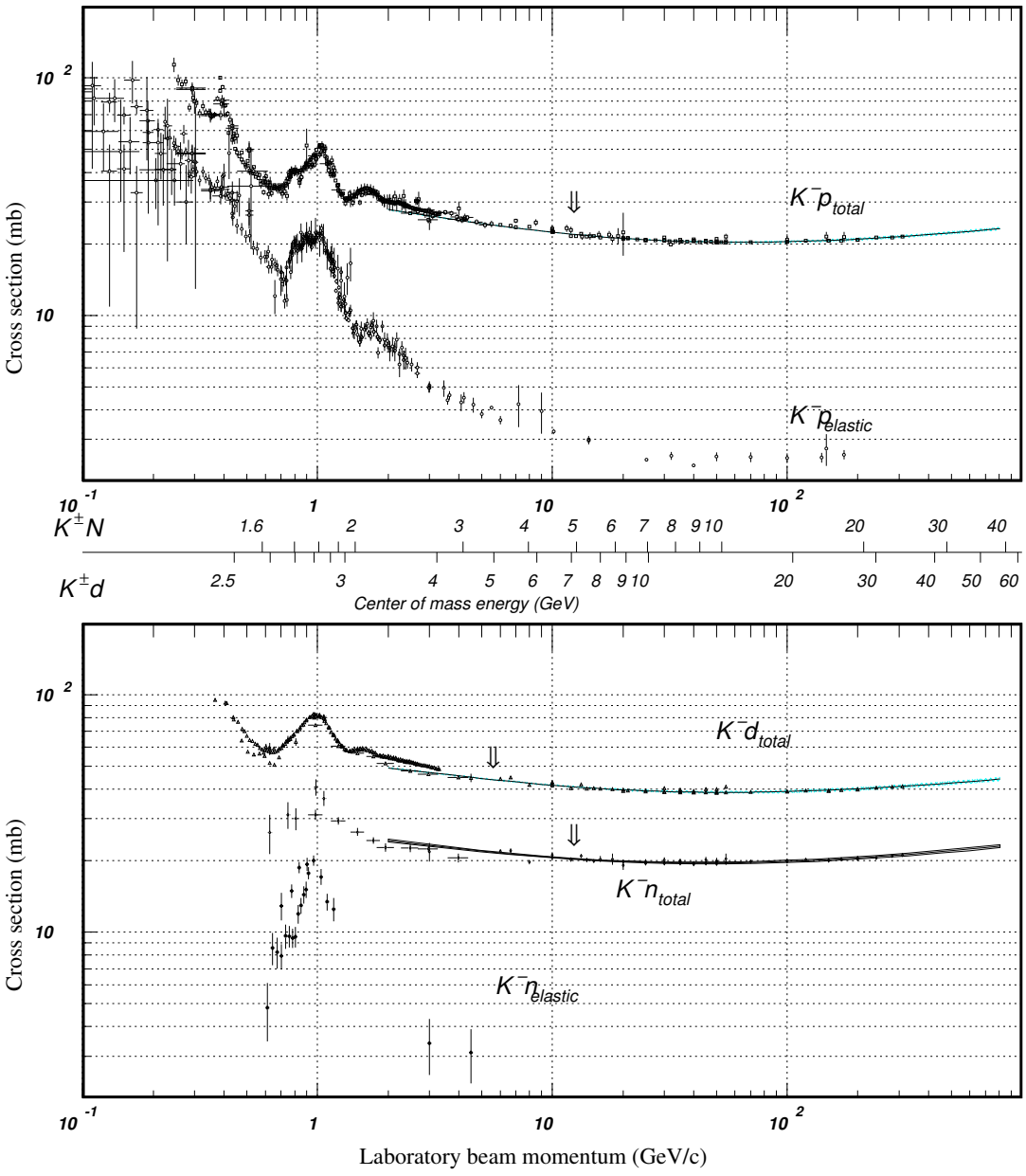

Figure 4.21: Total and elastic cross sections for $K^- p$ and total cross sections for $K^- d$ (total only), and $K^- n$ collisions as a function of laboratory beam momentum and total center-of-mass energy. Plots are taken from [65]. Corresponding computer-readable data files may be found at http://pdg.lbl.gov/xsect/contents.html (Courtesy of the COMPAS Group, IHEP, Protvino, Russia, August) 1999

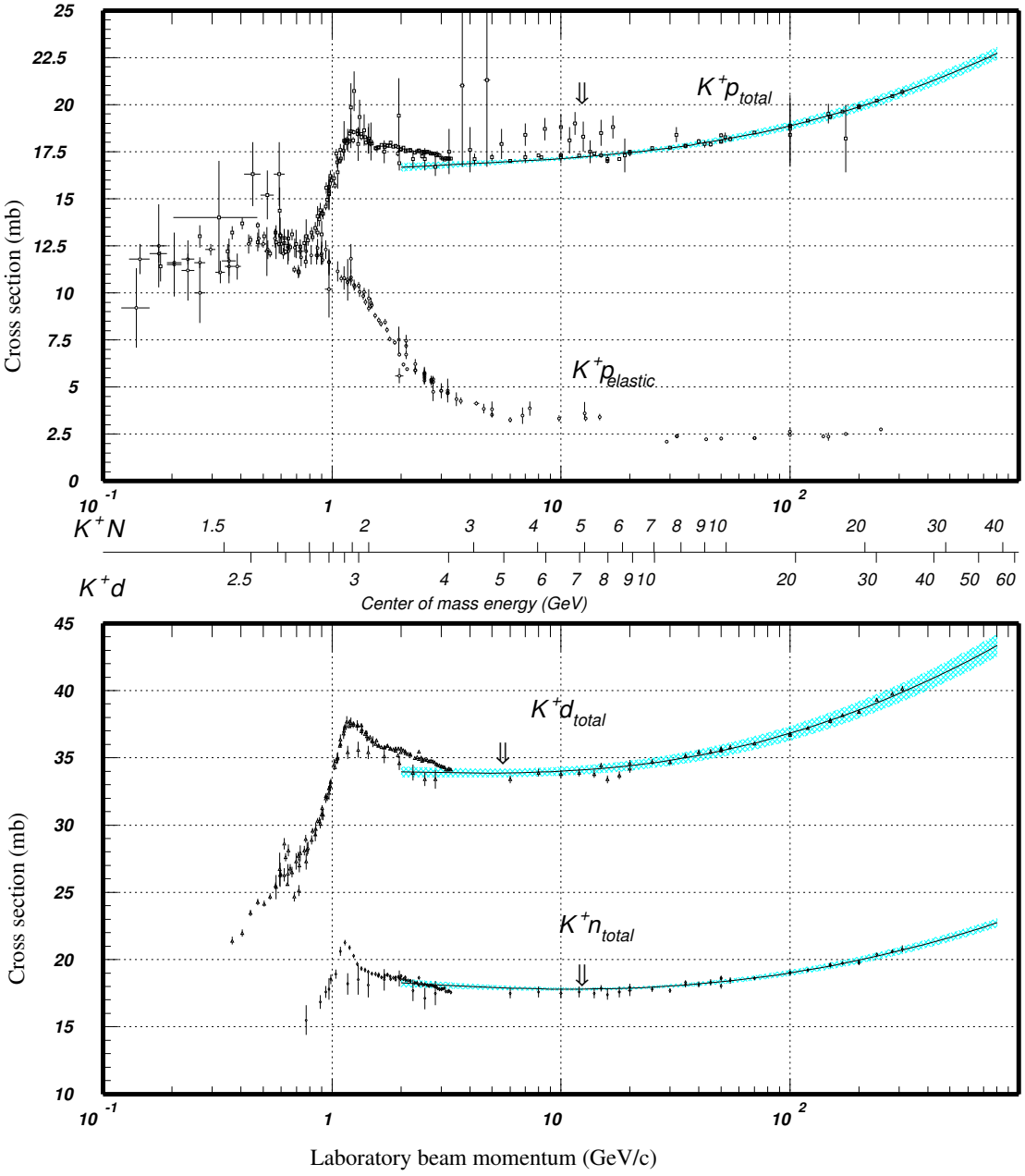

Figure 4.22: Total and elastic cross sections for K^+p and total cross sections for K^+d and K^+n collisions as a function of laboratory beam momentum and total center-of-mass energy. Plots are taken from [65]. Corresponding computer-readable data files may be found at http://pdg.lbl.gov/xsect/contents.html (Courtesy of the COMPAS Group, IHEP, Protvino, Russia, August) 1999

performed by the GEANT3 software from CERN, which is the only available option for the detector simulation at Belle. It is known that GEANT3 does not accurately predict interactions of strange particles with the detector material, with discrepancies in interaction cross section for the charged kaons of order of 10% for our energy range. This deficiency does not impact the asymmetry measurements, but it does affect the fractional weight for K_L interacting in the ECL, ECL-KLM and KLM. The fractional weight correction is extracted from the data using MC shapes with floating weights assigned to each class of events. This procedure will be described in detail in the Section 5.4.

4.9.3 Missing momentum

The large momentum of the K_L ($p > 1$ GeV/c) in the lab frame allows us to use missing momentum for event selection. The large solid angle coverage and excellent performance of the Belle detector facilitates the momentum measurement for most of the charged and neutral particles produced in the e^+e^- collision. This knowledge, combined with predefined beam momentum, allows us to estimate the momentum for the particles with undetectable momentum, like neutrinos and K_Ls. In order to improve the sensitivity for the K_L modes we also test the correlation between the directions of the K_L clusters and the missing momentum.

4.9.4 ECL candidates

The Belle software for the electromagnetic calorimeter is optimized for photon detection. As a result, the larger hadronic showers often split into several clusters with photon like signatures. In order to correct for this deficiency, we attempt to reconstruct proper hadron showers by combining close showers into larger hadronic ones. This process we call reclustering. In the first step, any two or more neutral clusters within a predefined distance are grouped together. For the group with the highest

number of clusters we calculate the center of gravity and use it as a seed direction. After finding the seed we move to the second step in which we loop through all the clusters and choose only those which have distance from the seed less than a predefined value. That group is declared to be a new super-cluster and clusters from the super-cluster get removed from the original list. After that we return to the first step and continue iteration until reclustering is finished.

The electromagnetic calorimeter sees the products of K_L decay in flight, the energy deposition from the hadronic interaction of the K_L with the CsI material, and the energy deposition from decay products of the regenerated K_S's. The energy deposition could be anything from 0 to $\sqrt{p_K^2 + M_K^2}$. However, the average hadronic collision results in pionic final states, where 33% of the energy appears in the form of π^0's, which immediately decay to photon pairs. This crude estimate hints at an average energy deposition greater than $0.33\sqrt{p_K^2 + M_K^2} \sim 400$ MeV/c^2 (for a 1 GeV/c kaon). The large low-energy photon contamination and the relatively large energy depositions for hadronic events lead us to consider only clusters with total energy deposition greater than 75 MeV/c^2 for the ECL-KLM candidates and 300 MeV/c^2 for ECL candidates.

The major backgrounds for neutral hadrons are charged tracks and photons. The clusters produced by charged particles are removed by requiring a minimum distance between the clusters and the projected impact position on the ECL to be less than 15 cm. The photon background is much harder to suppress, but the cluster properties for the photons are somewhat different from hadronic ones. On average photon clusters are more compact and exhibit no correlation with missing momentum.

The likelihood construction for the ECL candidates follows the same procedure as the J/ψ likelihood. The following parameters are used to form the two-dimensional likelihood discriminants:

- Missing momentum (P^{MM}) vs. transverse missing momentum (P_t^{MM}).

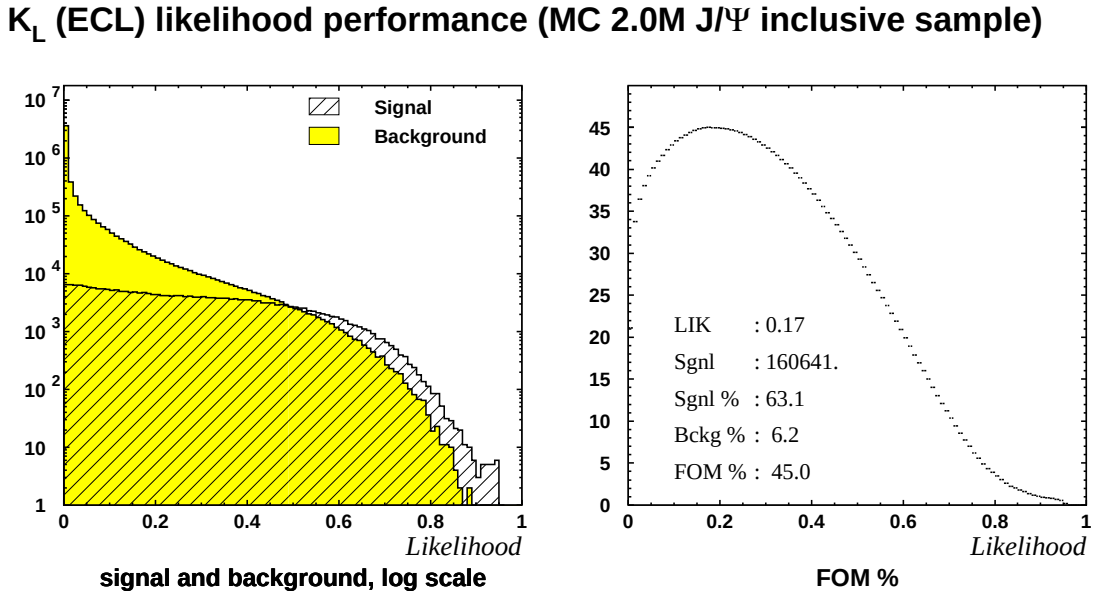

Figure 4.23: The K_L likelihood performance (ECL). The plots are based on a 2.0-M-event J/ψ inclusive MC sample. The signal is defined as any true K_L reconstructed as an ECL candidate. The background is any non-K_L candidate satisfying the selection criteria. The left plot shows the distribution for both signal and background events on a logarithmic scale as a function of the likelihood. The right plot shows the FOM as a function of likelihood for the events with likelihood greater than the abscissa value. The values presented on the plot show the maximum value for the FOM, the likelihood at which it was achieved $L_{\max}$, the number of signal events for $L > L_{\max}$ and the percentage of the ECL signal events satisfying $L > L_{\max}$.

- Missing momentum ϕ vs. K_L ϕ.
- Missing momentum θ vs. K_L θ.
- Angle with closest charged track vs. charged track momentum.
- Cluster energy vs. missing momentum (P^{MM}).
- Cluster width vs. E9/E25.

The performance of the likelihood based on the above parameters is summarized in Fig. 4.23

4.9.5 KLM candidates

The original cluster finder software was optimized for well defined clusters with low background. KLM hits were assigned to the same cluster only if they fell within a 5° cone with respect to each other. This constraint is rather tight and quite often results in splitting the original cluster into two sub-clusters. In order to test the cluster properties we perform reclustering for the KLM candidates the same way as we did for the ECL. Unfortunately, the stored base KLM cluster information is rather limited, and did not allow for thorough testing of the cluster properties. The only possible reclustering is based on the angular distance and does not have the information about the number of RPC hits which should be used as a weight. The optimized clustering angle was found to be 14°. This value was tuned to maximize the number of detected K_L within a 5° cone of the actual K_L direction. The increased angle not only improves the angular resolution, but also decreases the number of the single-layer-hit candidates, which are background prone.

Clusters that lie within a 15° cone of the charged-track projection into the KLM are vetoed to suppress backgrounds. No other restrictions are imposed on the KLM clusters.

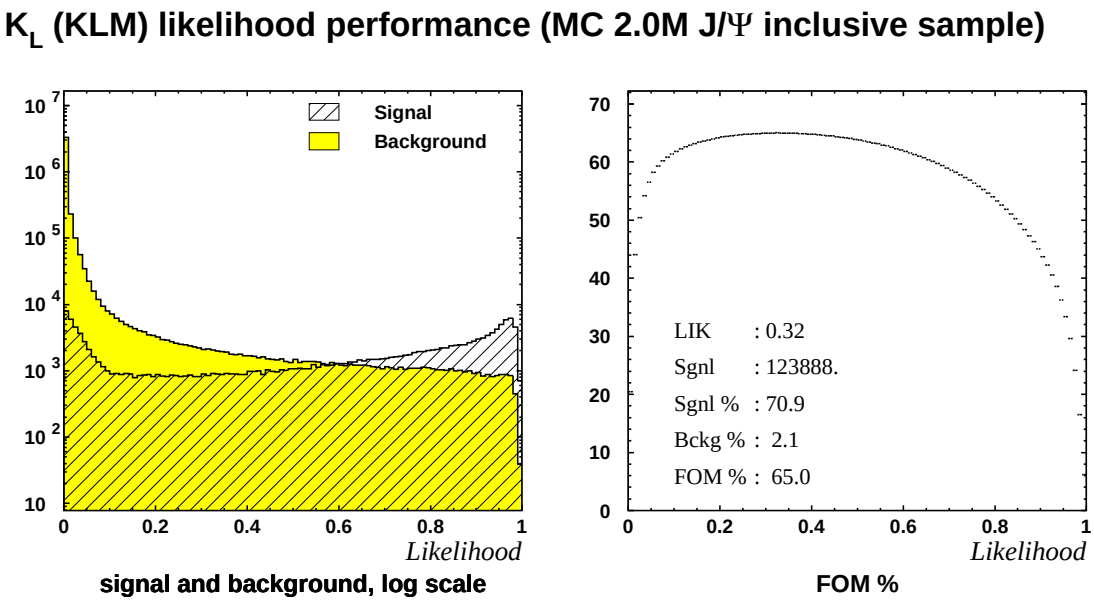

Figure 4.24: The K_L likelihood performance (KLM). The plots are based on a 2.0-M-event J/ψ inclusive MC sample. The signal is defined as any true K_L reconstructed as a KLM candidate. The background is any non-K_L candidate satisfying the selection criteria. The left plot shows the distribution for both signal and background events on a logarithmic scale as a function of the likelihood. The right plot shows the FOM as a function of likelihood for the events with likelihood greater than the abscissa value. The values presented on the plot show the maximum value for the FOM, the likelihood at which it was achieved $L_{\max}$, the number of signal events for $L > L_{\max}$ and the percentage of the KLM signal events satisfying $L > L_{\max}$.

The likelihood construction for the KLM candidates follows the standard likelihood procedure with the following parameters forming the two dimensional likelihood discriminants:

- Missing momentum (P^{MM}) vs. transverse missing momentum (P_t^{MM}).
- Missing momentum ϕ vs. K_L ϕ.
- Missing momentum θ vs. K_L θ.
- Angle with closest charged track vs. charged track momentum.
- Number of KLM layers hit vs. first layer hit.
- Number of KLM layers hit vs. missing momentum (P^{MM}).

The performance of the likelihood based on the above parameters is summarized in Fig. 4.24. The large number of events with low likelihood is due to the background-prone single-layer KLM clusters.

4.9.6 ECL-KLM candidates

The ECL-KLM candidates are the result of hadronic interactions in the ECL with secondary particles penetrating into the KLM. The secondary particles are much less energetic than the original K_L and thus have low penetration power. In addition, the high efficiency of the RPCs for charged particle detection will ensure that the first layer of the KLM cluster will be close to the ECL. Taking into account that long-lived secondary particles also will not be able to travel far, we can conclude that ECL-KLM candidates will have a well contained KLM cluster with a first layer close to the ECL cluster. The ECL-KLM clusters are selected from preselected ECL and KLM clusters. We select only KLM clusters that have hits in the first 5 layers of the KLM. This significantly reduces the combinatorial background from randomly

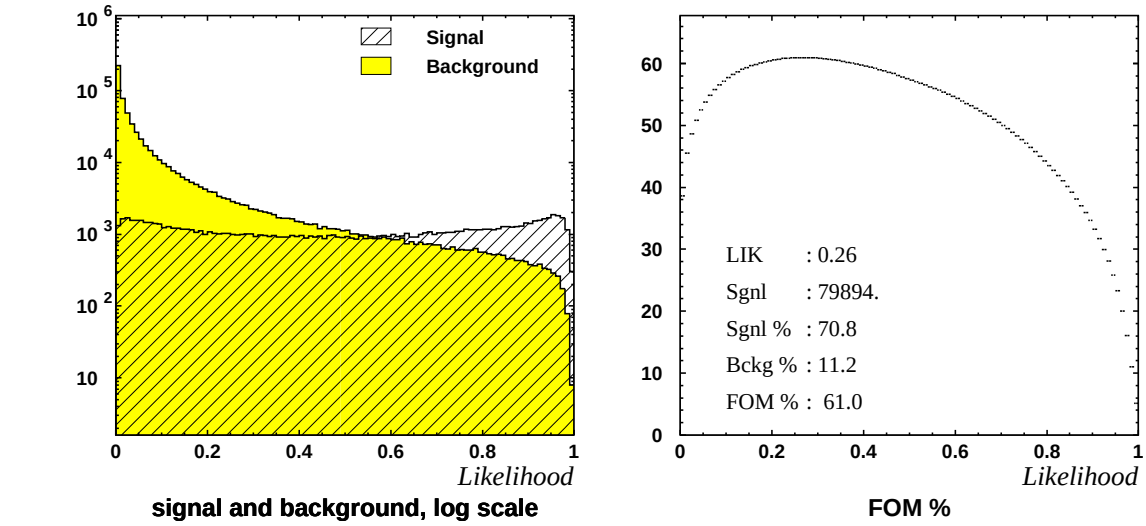

Figure 4.25: The K_L likelihood performance (ECL-KLM). The plots are based on a 2.0-M-event J/ψ inclusive MC sample. The signal is defined as any true K_L reconstructed as an ECL-KLM candidate. The background is any non-K_L candidate satisfying the selection criteria. The left plot shows the distribution for both signal and background events on a logarithmic scale as a function of the likelihood. The right plot shows the FOM as a function of likelihood for the events with likelihood greater than the abscissa value. The values presented on the plot show the maximum value for the FOM, the likelihood at which it was achieved $L_{\max}$, the number of signal events for $L > L_{\max}$ and the percentage of the ECL-KLM signal events satisfying $L > L_{\max}$.

matched KLM and ECL clusters. We combine ECL and KLM clusters if the angle between them is less than 10°.

The likelihood construction for the ECL-KLM candidates follows the standard likelihood procedure with the following parameters forming the two dimensional likelihood discriminants:

- Missing momentum ($P^{\rm MM}$) vs. transverse missing momentum ($P_t^{\rm MM}$).
- Missing momentum ϕ vs. K_L ϕ.
- Missing momentum θ vs. K_L θ.
- Angle with closest charged track vs. charged track. momentum.
- Number of KLM layers hit vs. first layer hit.

- Number of KLM layers hit vs. missing momentum (P^{MM}).
- Cluster energy vs. missing momentum (P^{MM}).
- Cluster width vs. E9/E25.

The performance of the likelihood based on the parameters above is summarized in Fig. 4.25

4.10 $K^*(892)$ selection.

The $K^*(892)$ decays predominantly into $K\pi$ pairs. The large width ($\sim$ 50 MeV/c^2) of the $K^*(892)$ makes it difficult to select a completely pure inclusive sample, but it is nonetheless advantageous to create a likelihood function that can be used to discriminate against B decay backgrounds containing $K^*(892)$'s. Due to the large combinatorial backgrounds involved, we will concentrate only on decays involving $K^{\pm}$ and $K_S \to \pi\pi^-$. The decays of $K^*(892)$ that produces K_L's can not be selected in inclusive mode, due to the lack of the discriminating parameters and reconstructed separately through the exclusive B meson decays (Section 4.14). Despite these omission K^* likelihood along likelihoods for other inclusive particles helps in reducing backgrounds in $B \to J/\psi\ K_L$ decay.

The selection procedure for this mode starts from choosing any K and π combination satisfying the following criteria.

- $L_K \otimes L_\pi > 0.03$
- $p^*_{K^*} < 3.0$ GeV/c
- 0.75 GeV/c^2< $M(K^*) <$ 1.05 GeV/c^2

The performance of this discriminant derived from a 0.5-M-event inclusive J/ψ MC sample is summarized in Figs. E.1, E.2, E.3, E.4.

4.11 Kinematic parameters for exclusive B decays

The B mesons studied by Belle are produced in the reaction $e^+e^- \to \Upsilon(4S) \to B\bar{B}$. To achieve the maximum event rate, the collision energy of the e^+e^- beams is set to the mass of the $\Upsilon(4S)$, which lies just above the threshold for $B\bar{B}$ production. Since both the collision energy and the masses of the final state B mesons are known, both the energies of the outgoing B mesons are determined (Appendix D). Thus the reconstructed mass and momentum of the B mesons are excellent tools for the separation of signal and background and also for the assessment of signal-to-background levels and yields. For various practical reasons, the mass and momentum variables are often replaced by related variables, called $\Delta E = E_B - E_{\rm beam}/2$ and $M_{\rm bc} = \sqrt{E^2_{\rm beam}/4 - p^2_B}$, but the general idea is the same. Here we use CM momentum of the B meson, p^*_B, and the mass of the B meson, $M_B = \sqrt{E^2_B - p^2_B}$. This choice of variables is a relic of the original $B \to J/\psi K_L$ event selection and now stays for backward compatibility.

Another series of kinematic parameters exploits angular correlations between decay products. The $B \to J/\psi\ K$ and $B \to J/\psi\ \pi$ families of decays have three common helicity angles, which we define below using the convention presented in Appendix B.4.

$\Upsilon(4S)$ **helicity** Defining transition is $e^+e^- \to \Upsilon(4S) \to B\ \bar{B}$. The $\Upsilon(4S)$ helicity is the cosine of the angle between the B meson and the beam momentum in the $\Upsilon(4S)$ rest frame. This is the same as the θ angle of the B meson flight direction in the CM frame.

B **helicity** Defining transitions are $\Upsilon(4S) \to B\ \bar{B}$ and $B \to J/\psi\ K$. The B helicity is the cosine of the angle between the J/ψ and the $\Upsilon(4S)$ in the B meson rest frame.

J/ψ **helicity** Defining transitions are $B \to J/\psi\ K$ and $J/\psi \to l^+\ l^-$. The J/ψ helicity is the cosine of the angle between between the lepton and the B in the

J/ψ rest frame.

In the $B \to J/\psi\ K_L$ decay reconstruction we have to use the B mass in order to calculate the K_L 4-momentum. As a result we lose one of the two main discriminants against the background and the two-dimensional M_B vs. p_B^* plane gets reduced to a one-dimensional distribution for the CM momentum of the B. This reduction results in increased background contamination. In the K_L analysis, all candidates are assigned momenta consistent with the decay mode, thus moving the kinematic parameters of the background into the physical region.

4.12 $B^{\pm} \to J/\psi\ K^{\pm}$.

Charged B meson decays into $J/\psi\ K^{\pm}$ final states have both high branching fractions and clean signatures. Even though they are of little interest for CP-violation studies, these modes exhibit the same kinematics as $B^0 \to J/\psi\ K_L$ making them an excellent demonstration tool for signal properties. Here we will not only describe the selection criteria, but also illustrate the parameters used in $B^0 \to J/\psi\ K_L$ selection. For illustration we select J/ψ with $L_{J/\psi} > 0.03$, 2.5 GeV/c^2 $< M_{J/\psi} <$ 3.2 GeV/c^2 and charged tracks coming from the IP region. We require the invariant mass and the CM momentum of the $J/\psi K$ pair to be correspondingly 4.9 GeV/c^2 $< M_B <$ 5.6 GeV/c^2 and 0.0 GeV/c $< p_B^* <$ 2.0 GeV/c. Despite such loose selection criteria, the selected sample has a rather high signal-to-background ratio. Its performance is summarized in Fig. 4.26. There are two well defined regions with a high density of events, denoted "A" and "B". Events falling into the "A" box with $J/\psi K$ mass 5.2 GeV/c^2 $< M_B <$ 5.37 GeV/c^2 and CM momentum 0.15 GeV/c $< p_B^* <$ 0.6 GeV/c have more than 95% $B \to J/\psi\ K$ signal purity and can be used to illustrate signal properties. The "B" box is dominated by events coming from $B \to J/\psi\ K^*(K^{\pm}\pi)$ decays, where the slow pion has been omitted. The relatively high density of events in the "B" box will allow us to do partial reconstruction of $B \to J/\psi\ K^*$ decays.

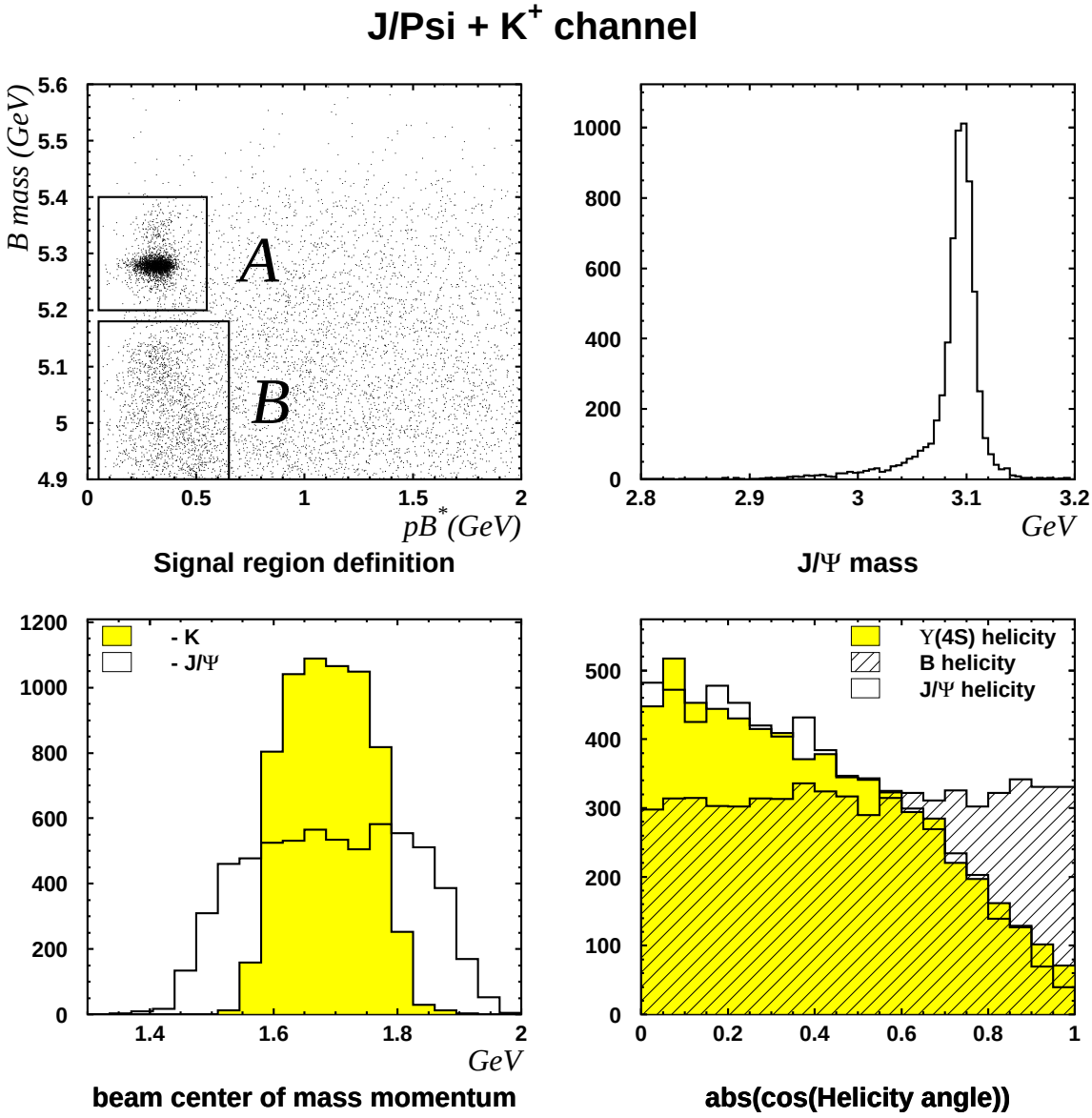

Figure 4.26: The kinematic variables related to $B^{\pm} \to J/\psi\ K^{\pm}$ decay. The upper left plot shows the distribution of B-meson mass versus CM momentum. The box "A" defines the signal region for $B^{\pm} \to J/\psi\ K^{\pm}$ decay, while box "B" outlines the region for $B^{\pm} \to J/\psi\ K^{\pm}X$ decays. All remaining plots are for events from the signal region "A" only. The upper right plot shows the J/ψ mass distribution. The bottom left plot depicts the CM momentum distributions for both J/ψ and $K^{\pm}$. The bottom right plot shows the helicity distributions for three helicity angles. The plot is based on a 78.13 fb^{-1} data sample.

		Signal region				
decay	S_R	S	B	S/S_R,%	$S/(S+B)$	FOM, %
$J/\psi K^{\pm}$	123927	121325	3101	97.9	97.5	97.7

Table 4.6: Summary for the $B \to J/\psi K^{\pm}$ likelihood performance based on a a 5.5-M-event J/ψ inclusive MC sample. S_R denotes the total number of reconstructed signal events, while $S(B)$ is the number of signal(background) events within the signal region.

Using events from the "A" box we plot the distribution for the CM momentum of both J/ψ and K. The bottom left plot from Fig. 4.26 shows that the decay products of the B meson also occupy a well defined region in the CM momentum space. The momentum for the both J/ψ and K are determined by the masses of all the particles involved and the beam CM momentum spread. The top right plot from Fig. 4.26 shows the mass distribution for J/ψ mesons and confirms the high purity of that sample.

The helicity behavior of two body B meson decays involving the J/ψ is illustrated in the bottom right plot of Fig. 4.26. The helicity angle distribution for both $\Upsilon(4S)$ and J/ψ exhibits the expected $\sin^2\theta_h$ dependence and the B helicity angle distribution remains flat.

For $B^{\pm} \to J/\psi\ K^{\pm}$ decay we select any J/ψ and K combination satisfying the following criteria.

- $L_{J/\psi} > 0.03$
- $5.00\ \mathrm{GeV}/c^2 < M_B < 5.40\ \mathrm{GeV}/c^2$
- $p^*_B < 1.0\ \mathrm{GeV}/c$
- $1.20\ \mathrm{GeV}/c < p^*_{J/\psi} < 2.00\ \mathrm{GeV}/c$
- $1.25\ \mathrm{GeV}/c < p^*_K < 1.90\ \mathrm{GeV}/c$

The performance of the likelihood based on a 5.5-M-event inclusive J/ψ MC sample is summarized in Figs. E.5 with the main results presented in Table 4.6.

4.13 $B^0 \to J/\psi\ K_S$.

The decay $B \to J/\psi\ K_S$ is the CP-mirror of the $B \to J/\psi\ K_L$ transition. As a result, $B \to J/\psi\ K_S$ represents not just a potential background to $B \to J/\psi\ K_L$, but one that can directly affect the observed CP asymmetry. Keeping that in mind, we pay special attention to the reconstruction efficiency for veto modes. In order to further increase the efficiency, $B \to J/\psi\ K_S$ events are selected based on the $B \to J/\psi\ \pi\ X$ decay topology. Despite missing one of the pions from the K_S decay, we can still select a high purity sample, because events from $B \to J/\psi\ K_S$ decays occupy a well defined region in the mass-momentum plane. Here we present the selection criteria for four types of K_S decays, decays into charged and neutral pions and corresponding partial reconstruction for single-pion final states.

For $B^\pm \to J/\psi\ K_S$ decay, we select J/ψ ($L_{J/\psi} > 0.03$) and K_S combinations satisfying the following criteria.

- $5.0\ \mathrm{GeV}/c^2 < M_B < 5.4\ \mathrm{GeV}/c^2$
- $p^*_B < 1.0\ \mathrm{GeV}/c$
- $1.20\ \mathrm{GeV}/c < p^*_{J/\psi} < 2.00\ \mathrm{GeV}/c$
- $1.25\ \mathrm{GeV}/c < p^*_{K_S} < 1.90\ \mathrm{GeV}/c$

For partially reconstructed $B^\pm \to J/\psi\ K_S$ decays we select J/ψ ($L_{J/\psi} > 0.03$) and π combinations with following selection criteria.

- $4.35\ \mathrm{GeV}/c^2 < M_B < 5.20\ \mathrm{GeV}/c^2$
- $p^*_B < 1.30\ \mathrm{GeV}/c$
- $1.35\ \mathrm{GeV}/c < p^*_{J/\psi} < 1.90\ \mathrm{GeV}/c$
- $0.75\ \mathrm{GeV}/c < p^*_\pi < 1.80\ \mathrm{GeV}/c$

		Signal region				
decay	S_R	S	B	S/S_R,%	$S/(S+B)$	FOM, %
$J/\psi K_S(\pi^+\pi^-)$	37929	35843	1423	94.5	96.2	95.3
$J/\psi K_S(\pi^0\pi^0)$	9477	7165	1654	75.6	81.2	78.4
$J/\psi K_S(X\pi^\pm)$	37174	29851	32316	80.3	48.0	62.1
$J/\psi K_S(X\pi^0)$	27408	15897	40296	58.0	28.3	40.5

Table 4.7: Summary of the $B \to J/\psi K_S$ likelihood performance based on a 5.5-M-event J/ψ inclusive MC sample. S_R denotes the total number of reconstructed signal events, while $S(B)$ is the number of signal(background) events within the signal region.

The performance of the likelihood based on a 5.5-M-event inclusive J/ψ MC sample is summarized in Figs. E.6, E.7, E.8 and E.9 with the main results presented in Table 4.7. Note that there is a significant overlap between the $J/\psi K_S(\pi^+\pi^-)$ and the $J/\psi K_S(X\pi^\pm)$ samples. Moreover, for some events we might have $J/\psi K_S(\pi^+\pi^-)$ and two candidates for $J/\psi K_S(X\pi^\pm)$, corresponding to both pions. Despite the overlap, partially reconstructed decays increase the efficiency by more than 30% for $J/\psi K_S$ decays.

4.14 $B^0 \to J/\psi\ K^*$.

The $B \to J/\psi\ K^*(892)$ decay sequence is identical to $B \to J/\psi\ K$ decay with respect to quark transitions. The only differences are the higher mass ($\sim$ 890 MeV/c^2) and spin-1 value for the K^* meson. This similarity along with the wide mass width ($\sim$ 50 MeV/c^2) and abundance of kaons among the K^* decay products (BR($K^*(892) \to K\pi$) $\approx$ 100%) make $B \to J/\psi\ K^*(892)$ transitions a major source of background to $B \to J/\psi\ K_L$. The kinematic variables related to $B \to J/\psi\ K^*(892)$ decay are shown in Fig. 4.27. The spin value of the K^* results in a flat distribution for the J/ψ helicity, while the lower efficiency for low-momentum charged tracks suppresses particles with low CM momentum. Charged particle production in $B \to J/\psi\ K^*(892)$ decay exhibits a momentum cutoff around 300 MeV/c for kaons and around 60 MeV/c for pions.

partial decay	actual decays
$B \to J/\psi K_S(\pi^+\pi^-)$	$B^\pm \to J/\psi K^{*\pm}(K_S\pi^\pm)$, $B^0 \to J/\psi K^{*0}(K_S\pi^0)$
$B \to J/\psi K^\pm$	$B^\pm \to J/\psi K^{*\pm}(K^\pm\pi^0)$, $B^0 \to J/\psi K^{*0}(K^\pm\pi^\pm)$
$B \to J/\psi \pi^\pm$	$B^\pm \to J/\psi K^{*\pm}(K^0\pi^\pm)$, $B^0 \to J/\psi K^{*0}(K^\mp\pi^\pm)$
$B \to J/\psi \pi^0(\gamma\gamma)$	$B^\pm \to J/\psi K^{*\pm}(K^\pm\pi^0)$, $B^0 \to J/\psi K^{*0}(K^0\pi^0)$

Table 4.8: Decay chains used for partial reconstruction in $B^0 \to J/\psi\ K^*(892)$ selection.

Low-momentum charged pions are suppressed because their high curvature in the magnetic field prevents them from reaching the drift chamber, while low-momentum kaons tend to decay in flight inside the drift chamber. A similar situation is present for neutral pions, where large combinatorial backgrounds from low-energy photons limit the sensitivity to low-momentum pions. Unfortunately, the signature of $B \to J/\psi\ K^*(892)$ decay with low-momentum pion production is similar to that for $B \to J/\psi\ K$ decay. This is particulary worrisome for decays involving the K_L, where the lack of a K_L energy measurement makes separation of $B \to J/\psi\ K^*(K_L\ \pi)$ from $B \to J/\psi\ K_L$ especially hard. The aforementioned problems forced us to attempt partial reconstruction of $B \to J/\psi\ K^*(892)(K\pi)$ decays based only on the $B \to J/\psi\ K$ or the $B \to J/\psi\ \pi$ topology. This method is analogous to the partial reconstruction of $B \to J/\psi\ K_S(X\pi)$ decays discussed in the previous section.

The family of partially reconstructed decays consists of the four channels presented in Table 4.8. In all such channels we calculate the kinematic parameters of the B meson assuming the corresponding two-body decay $B \to J/\psi\ K$ or $B \to J/\psi\ \pi$. Partially reconstructed decays have good purity because of the good particle identification and the low energy loss for the missing particle. This is especially useful for the suppression of decays with a missing low-momentum particle, since they tend to populate the low-background area, as can be seen from the higher density of events in the top part of region "B" in the top left part of Fig. 4.26. This is welcome news, because it allows us to offset detector limitations for low-momentum particle detection and to effectively recover most $B \to J/\psi\ K^*$ decays with either a kaon or a pion

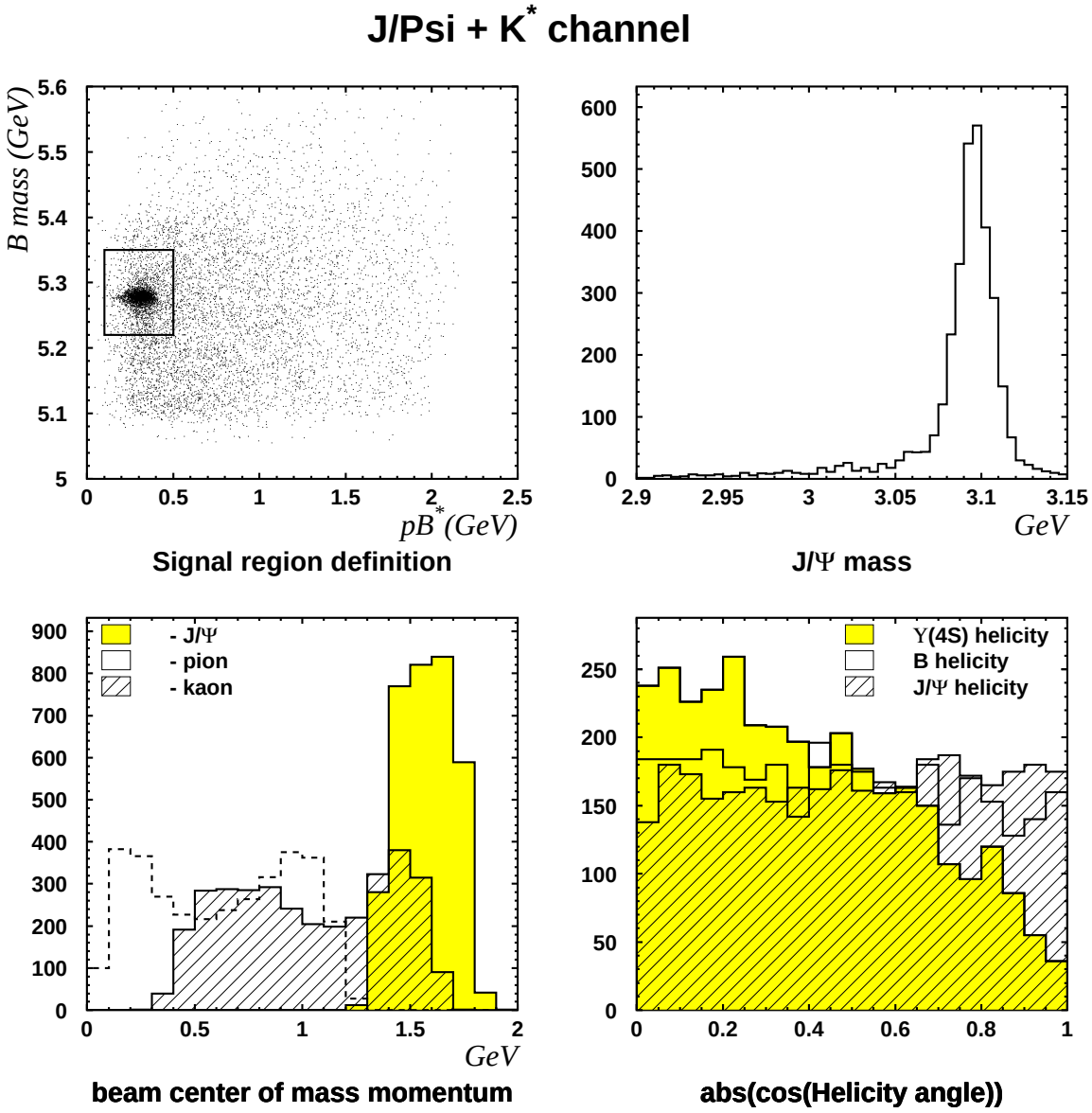

Figure 4.27: The kinematic variables related to $B^{\pm} \to J/\psi\ K^*(892)(K^{\pm}\pi^{\mp})$ decay. The upper left plot shows the distribution of B-meson mass versus CM momentum. The box on the top left plot outlines the signal region for $B^{\pm} \to J/\psi\ K^*$ decay. All remaining plots are for events from the signal region only. The upper right plot shows the J/ψ mass distribution. The bottom left plot depicts the CM momentum distributions for J/ψ and products of K^* decay: K and π. The bottom right plot shows the helicity distributions for three helicity angles. The plot is based on a 78.13 fb^{-1} data sample.

detected.

For $B^{\pm} \to J/\psi\ K^*$ decay, we select J/ψ ($L_{J/\psi} > 0.03$) and K-π combinations satisfying the following criteria.

- $5.00\ \mathrm{GeV}/c^2 < M_B < 5.40\ \mathrm{GeV}/c^2$
- $p^*_B < 1.00\ \mathrm{GeV}/c$
- $1.20\ \mathrm{GeV}/c < p^*_{J/\psi} < 1.90\ \mathrm{GeV}/c$

For partially reconstructed $B^{\pm} \to J/\psi\ K^*(XK)$ decays, we select J/ψ ($L_{J/\psi} > 0.03$) and K combinations with the following selection criteria.

- $4.35\ \mathrm{GeV}/c^2 < M_B < 5.20\ \mathrm{GeV}/c^2$
- $p^*_B < 1.30\ \mathrm{GeV}/c$
- $1.20\ \mathrm{GeV}/c < p^*_{J/\psi} < 1.90\ \mathrm{GeV}/c$
- $0.70\ \mathrm{GeV}/c < p^*_{\pi} < 1.75\ \mathrm{GeV}/c$

For partially reconstructed $B^{\pm} \to J/\psi\ K^*(X\pi)$ decays we select J/ψ ($L_{J/\psi} > 0.03$) and π combinations with the following selection criteria.

- $4.30\ \mathrm{GeV}/c^2 < M_B < 4.85\ \mathrm{GeV}/c^2$
- $p^*_B < 1.15\ \mathrm{GeV}/c$
- $1.30\ \mathrm{GeV}/c < p^*_{J/\psi} < 1.90\ \mathrm{GeV}/c$
- $0.75\ \mathrm{GeV}/c < p^*_{\pi} < 1.30\ \mathrm{GeV}/c$

The performance of the likelihood based on a 5.5-M-event inclusive J/ψ MC sample is summarized in Figs. E.10, E.11, E.12, E.13, E.14, E.15, E.16, E.17, E.18 and E.19 with the main results presented in Table 4.9.

		Signal region				
decay	S_R	S	B	S/S_R,%	$S/(S+B)$	FOM, %
$J/\psi K^*(K^\pm\pi^\mp)$	102023	91107	9336	89.3	90.7	90.0
$J/\psi K^*(K^\pm\pi^0$	60955	42973	23520	70.5	64.6	67.5
$J/\psi K^*(K_S\pi^\pm)$	26171	22376	3856	85.5	85.3	85.4
$J/\psi K^*(K_S\pi^0)$	20039	13827	7113	69.0	66.0	67.5
$J/\psi K^*(K_L\pi^\pm)$	38508	22181	25523	57.6	46.5	51.7
$J/\psi K^*(K_L\pi^0)$	33186	14967	33643	45.1	30.8	37.3
$J/\psi K^*(XK^\pm)$	130900	98758	85552	75.0	53.6	63.4
$J/\psi K^*(X\pi^\pm)$	99754	78706	61085	78.9	56.3	66.7
$J/\psi K^*(XK_S)$	45410	33922	18771	74.7	64.4	69.3
$J/\psi K^*(X\pi^0)$	67110	57178	123207	85.2	31.7	52.0

Table 4.9: Summary of the $B \to J/\psi K^*$ likelihood performance based on a 5.5-M-event J/ψ inclusive MC sample. S_R denotes the total number of reconstructed signal events, while $S(B)$ is the number of signal(background) events within the signal region.

4.15 $B \to J/\psi\ \pi$.

Although the branching fraction for $B \to J/\psi\ \pi$ decay is small (≈ 0.05 BR($B \to J/\psi\ K$), it nonetheless represents a potentially important background to $B \to J/\psi\ K_L$, since contamination from π^0 decay products can mimic a K_L in the ECL.

For $B^\pm \to J/\psi\ \pi$ decay we select J/ψ ($L_{J/\psi} > 0.03$) and π combinations satisfying the following criteria.

- 5.00 GeV/c^2 $< M_B <$ 5.40 GeV/c^2
- $p^*_B <$ 1.00 GeV/c
- 1.40 GeV/c $< p^*_{J/\psi} <$ 2.00 GeV/c
- 1.50 GeV/c $< p^*_\pi <$ 1.90 GeV/c

The performance of the likelihood based on a 5.5-M-event inclusive J/ψ MC sample is summarized in Fig. E.21 and Fig. E.20 with the main results presented in Table 4.10.

		Signal region				
decay	S_R	S	B	S/S_R,%	$S/(S+B)$	FOM, %
$J/\psi\pi^{\pm}$	4341	3551	459	81.8	81.6	85.1
$J/\psi\pi^0$	16563	3277	938	76.4	77.7	77.1

Table 4.10: Summary of the $B \to J/\psi\pi$ likelihood performance based on a 5.5-M-event J/ψ inclusive MC sample. S_R denotes the total number of reconstructed signal events, while $S(B)$ is the number of signal(background) events within the signal region.

4.16 $B^0 \to J/\psi\ K_L$.

The selection procedure for this mode starts from choosing any J/ψ and K_L combination satisfying the following criteria.

- $L_{J/\psi} > 0.03$
- 1.2 GeV/$c < p^*_{J/\psi} <$ 2.0 GeV/c
- $L_{K_L} > 0.01$
- $L_{J/\psi} \otimes L_{K_L} > 0.01$
- $p^*_B <$ 2 GeV/c

J/ψ mesons from $B \to J/\psi\ K$ decay have CM momenta in the range 1.4 GeV/$c <$ $p^*_{J/\psi} <$ 2.0 GeV/c. To achieve better stability of the likelihood construction and yield estimate, however, we decrease the low boundary of the momentum cut. This will not affect the final selection criteria because likelihood values for low CM momenta will be below the likelihood cutoff.

The likelihood construction for this mode differs from the previously discussed decays. In $L(B \to J/\psi\ K_L)$ we are using "parent likelihoods" (for definition see Section 4.11). This makes it conditional on likelihoods for other exclusive and inclusive decays, which must be created first. The same order is preserved on the reconstruction stage, where all of the previously discussed inclusive and exclusive modes have

		Signal region				
decay	S_R	S	B	S/S_R,%	$S/(S+B)$	FOM, %
$J/\psi K_L$(KLM)	23400	18977	10638	81.1	64.1	72.1
$J/\psi K_L$(ECL-KLM)	16563	11992	9001	72.4	57.1	64.3
$J/\psi K_L$(ECL)	14986	12109	5730	80.8	67.9	74.1

Table 4.11: Summary of the $B \to J/\psi K_L$ likelihood performance based on a 5.5-M-event J/ψ inclusive MC sample. S_R denotes the total number of reconstructed signal events, while $S(B)$ is the number of signal (background) events within the signal region.

to be reconstructed first. Effectively we need to veto any J/ψ related decay. This is achieved by including $L^P(J/\psi)$—which stores the highest likelihood among possible J/ψ parents: B, $\psi(2S)$ andχ_c— in the $B \to J/\psi\ K_L$ likelihood. The other major source of background comes from $B \to J/\psi\ K^{*0}(K_L\ \pi)$ related decays, which can be suppressed by using $L^P(K_L)$.

The momentum assignment for the reconstructed K_L distorts the distributions for the helicity angles. The B helicity is fully defined by p^*_B and $p^*_{J/\psi}$, thus making only two out of the three commonly used helicity angles suitable for likelihood construction. The other two parameters are also affected by the momentum assignment and as a result differ from the related signal-background distribution for $B \to J/\psi\ K^{\pm}$ decay. However, they show no strong correlation with other major kinematic variables and can be used for signal selection.

The dramatically different background compositions for K_L's interacting in only the ECL, in both the ECL and the KLM, and the KLM only, force us to create three separate likelihood functions. The performance of the likelihood based on a 5.5-M-event inclusive J/ψ MC sample is summarized in Figs. (E.22, E.23, and E.24) with the main results tabulated in Table 4.11.

A separate treatment for the ECL, ECL-KLM and KLM components allows a yield estimation for each component, which is important for our goal. Previous analyses [46, 47, 48, 49, 50, 51] use the shape of the B meson CM momentum (p^*_B) for signal yield estimates. In this work, however, we switch to an approach based on likelihood. The

		Background types		
	S	$J/\psi^T\ K_L^T$	$J/\psi^T\ K_L^F$	J/ψ^F
KLM p_B^*	758 ± 60	82 ± 150	1679 ± 110	713 ± 60
KLM L	683 ± 50	991 ± 260	815 ± 190	788 ± 220
KLM $L\&p_B^*$	684 ± 40	1019 ± 190	984 ± 160	580 ± 80
ECL-KLM p_B^*	278 ± 35	116 ± 90	638 ± 75	313 ± 30
ECL-KLM L	221 ± 30	500 ± 130	302 ± 100	352 ± 110
ECL-KLM $L\&p_B^*$	224 ± 22	423 ± 100	341 ± 90	383 ± 50
ECL p_B^*	720 ± 60	1227 ± 180	2551 ± 160	1981 ± 110
ECL L	756 ± 60	1913 ± 536	2182 ± 500	1678 ± 500
ECL $L\&p_B^*$	740 ± 40	1353 ± 390	2375 ± 360	2045 ± 190

Table 4.12: Signal and background yields for the $B \to J/\psi K_L$ sample with $p_B^* < 0.8$ GeV/c.

limited experience in using a yield estimation based on the likelihood leads us to be cautious and we present the signal and background yield estimates for both methods.

All of the likelihood functions for B-meson decays were based on an inclusive J/ψ MC sample. As a result, the MC can provide only shapes for signal, background with true K_L's, and background with fake K_L's. We will account for fake J/ψ components using the approach discussed in Section 4.5.2. The results of this procedure for KLM, ECL-KLM, and KLM candidates are presented in the top part of Figs. (4.28, 4.30, and 4.29).

The top left plot shows the results for the fit in the conventional p_B^* space, while the top right plot presents the result for likelihood space. Note that the p_B^* space provides a robust estimate for the fake J/ψ component with low sensitivity to the true/fake K_L ratio, while likelihood space estimates the true/fake K_L ratio well. In order to exploit these complementary properties we present simultaneous fits in both spaces in the bottom left plot. Having obtained yield estimates we are ready for the final step, which is finding the criteria for selecting the $J/\psi\ K_L$ sample. At this stage we start using the $L^F = L(J/\psi\ K_L) \otimes L(J/\psi)$ likelihood. In L^F space we do not know the fake J/ψ distribution. However, we can find it by subtracting the MC-predicted

shape from the data, using the result from the previous step. The result is presented in the bottom right plot. Now that we know all the distributions we can find the likelihood cut value that maximizes the FOM. Overall yield estimates are presented in Table 4.12. The yield estimates are consistent between different fit procedures and confirm that the fractions of K_L interacting in the KLM, the KLM-ECL, and the ECL are different between MC and data. The MC estimates the fractions among detectable K_L from $B \to J/\psi K_L$ decay as 42.5%, 30.2%, and 27.3% for KLM, KLM-ECL, and ECL respectively, while an estimate based on data yields 41.5%, 13.6% and 44.9%. This indicates that K_L cross section for ECL detection is low in the MC.

The final selection criteria for the $B \to J/\psi K_L$ decay are the following

- $L(J/\psi) > 0.03$
- $1.2 \text{ GeV}/c < p^*_{J/\psi} < 2.0 \text{ GeV}/c$
- $L_{K_L} > 0.01$
- $L(J/\psi) \otimes L(K_L) > 0.01$
- $p^*_B < 0.8 \text{ GeV}/c$
- $L(J/\psi) \otimes L(B \to J/\psi K_L) > \begin{array}{l} 0.32\text{(KLM)} \\ 0.25\text{(ECL-KLM)} \\ 0.38\text{(KLM)} \end{array}$

This selection provides a sample with 489 signal and 267 background events detected in the KLM, 522 signal and 480 background events in the ECL, and 163 signal and 133 background events of the ECL-KLM type. Overall we select 2054 events (1174 signal events and 880 events from background) with sample purity of $S/(S+B) = 57\%$ and FOM=25.9. For comparison, the official release of the Belle CP selection [51] for the same data set provides sample of 1330 events with a purity of 63% and a FOM of 22.9.

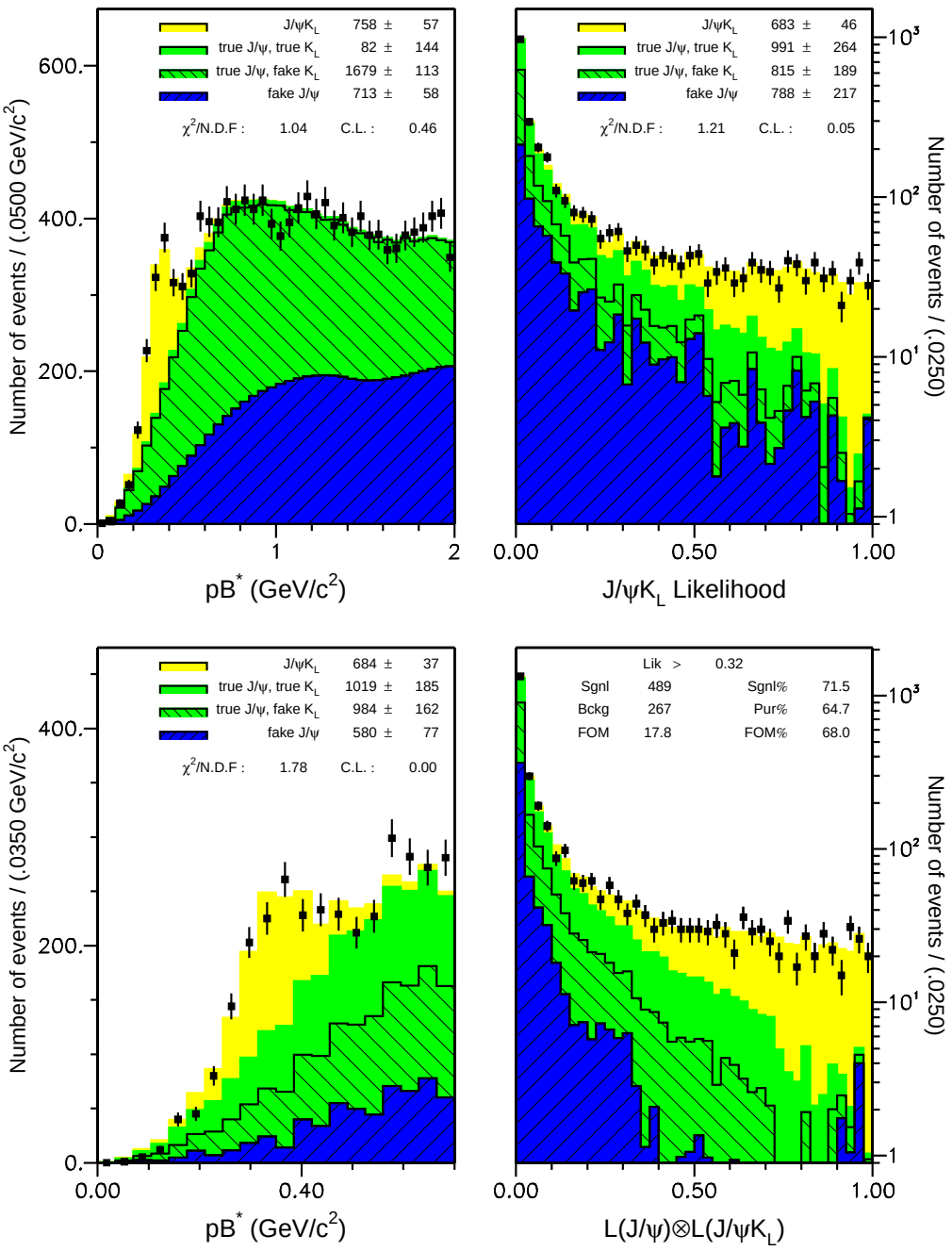

Figure 4.28: Estimates for the $B^0 \to J/\psi\ K_L$(KLM) signal, inclusive J/ψ background and combinatorial J/ψ background yield. All plots are based on a 78.13 fb^{-1} on-resonance data sample with $L_{J/\psi} > 0.03$ and $L_{K_L} > 0.03$. The combinatorial background shape is based on a sample with $L_{J/\psi} < 0.25$ and inclusive J/ψ shape from a sample with $L_{J/\psi} > 0.90$. Top left plot shows yield estimates using mass variable, while top right gives estimates using likelihood. Results on the bottom left plot are based on both mass and likelihood variables. Bottom right plot shows shapes for full likelihood along with likelihood cutoff value and yields, which maximize the FOM.

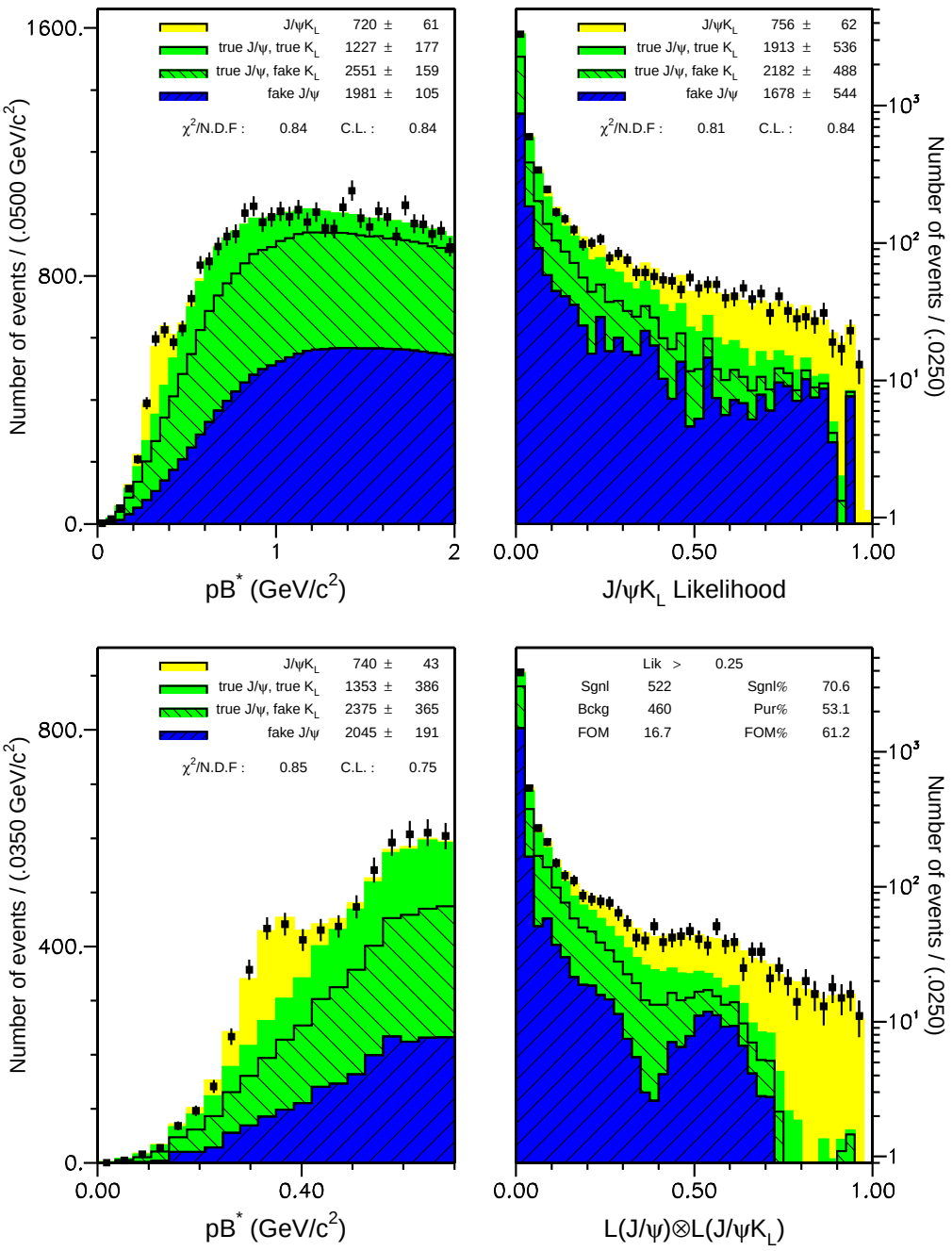

Figure 4.29: Estimates for the $B^0 \to J/\psi\ K_L$(ECL) signal, inclusive J/ψ background and combinatorial J/ψ background yield. All plots are based on a 78.13 fb^{-1} on-resonance data sample with $L_{J/\psi} > 0.03$ and $L_{K_L} > 0.03$. The combinatorial background shape is based on a sample with $L_{J/\psi} < 0.25$ and inclusive J/ψ shape from a sample with $L_{J/\psi} > 0.90$. Top left plot shows yield estimates using mass variable, while top right gives estimates using likelihood. Results on the bottom left plot are based on both mass and likelihood variables. Bottom right plot shows shapes for full likelihood along with likelihood cutoff value and yields, which maximize the FOM.

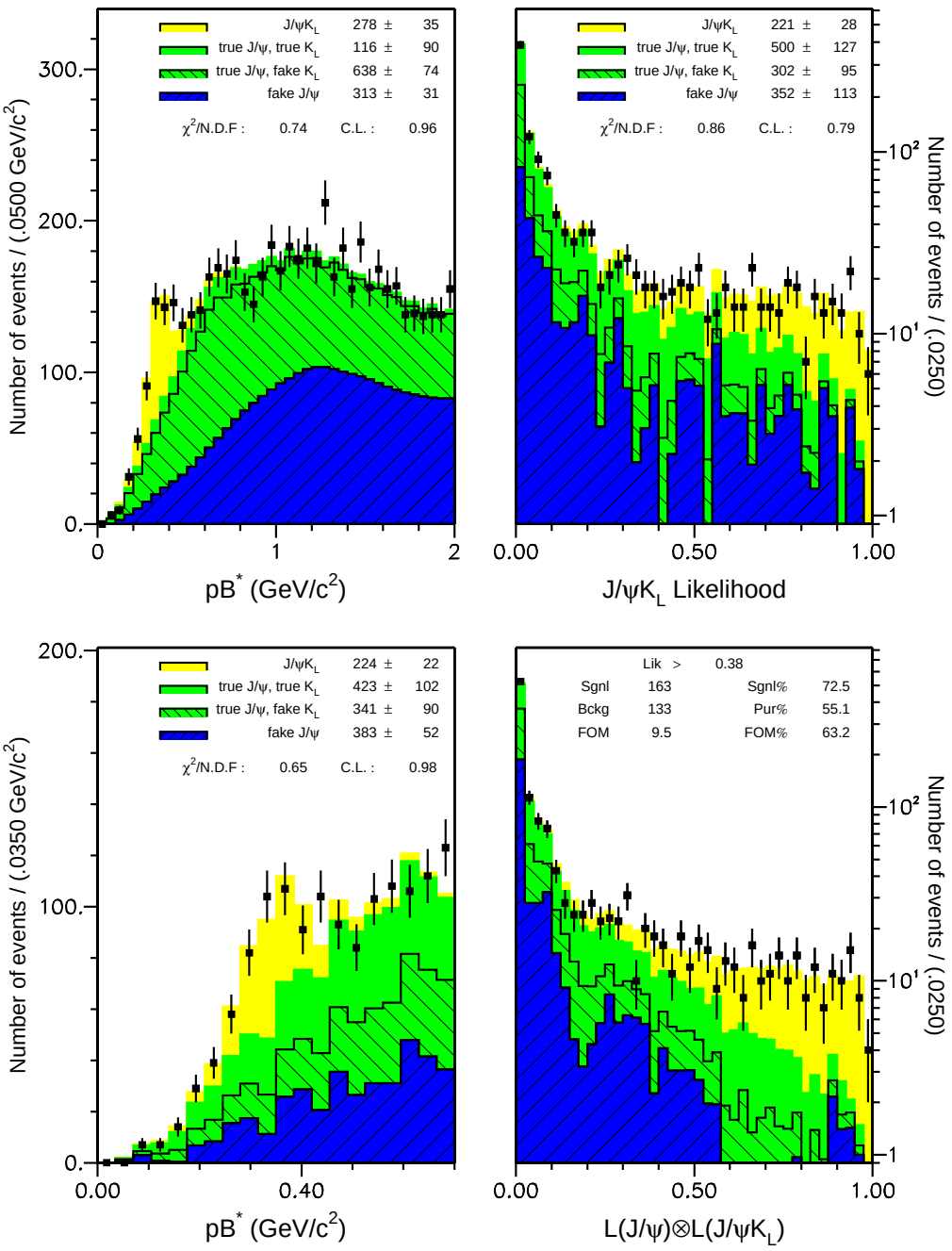

Figure 4.30: Estimates for the $B^0 \to J/\psi\ K_L$(ECL-KLM) signal, inclusive J/ψ background and combinatorial J/ψ background yield. All plots are based on a 78.13 fb^{-1} on-resonance data sample with $L_{J/\psi} > 0.03$ and $L_{K_L} > 0.03$. The combinatorial background shape is based on a sample with $L_{J/\psi} < 0.25$ and inclusive J/ψ shape from a sample with $L_{J/\psi} > 0.90$. Top left plot shows yield estimates using mass variable, while top right gives estimates using likelihood. Results on the bottom left plot are based on both mass and likelihood variables. Bottom right plot shows shapes for full likelihood along with likelihood cutoff value and yields, which maximize the FOM.

Chapter 5

Measurement of the CP Asymmetry.

This chapter outlines the procedure for the $\sin 2\phi_1$ measurement based on the sample we have selected in Chapter 4. In order to measure the CP violating asymmetry, we must know the flavor of the other B meson in the event. The procedure for assigning a specific flavor to the accompanying B meson is called flavor tagging and is described in Section 5.1. After flavor tagging, we must determine the vertex position of both B mesons. That information allows us to calculate the proper time difference,Δt, between the B meson decays (Section 5.2). The final step is performing an unbinned likelihood estimate for the $\sin 2\phi_1$ parameter (referred to as the CP fit) based on the Δt distribution (Section 5.4).

5.1 Flavor Tagging: B^0_d or $\overline{B^0_d}$ meson?

Time-dependent CP violation measurements at Belle rely heavily on the ability to determine the flavor of a B_d meson. A full description of flavor tagging is beyond this document, but can be found elsewhere [67]. Here we present short overview. In Belle, flavor tagging is implemented in a software package called "`Hamlet`". Once one of the

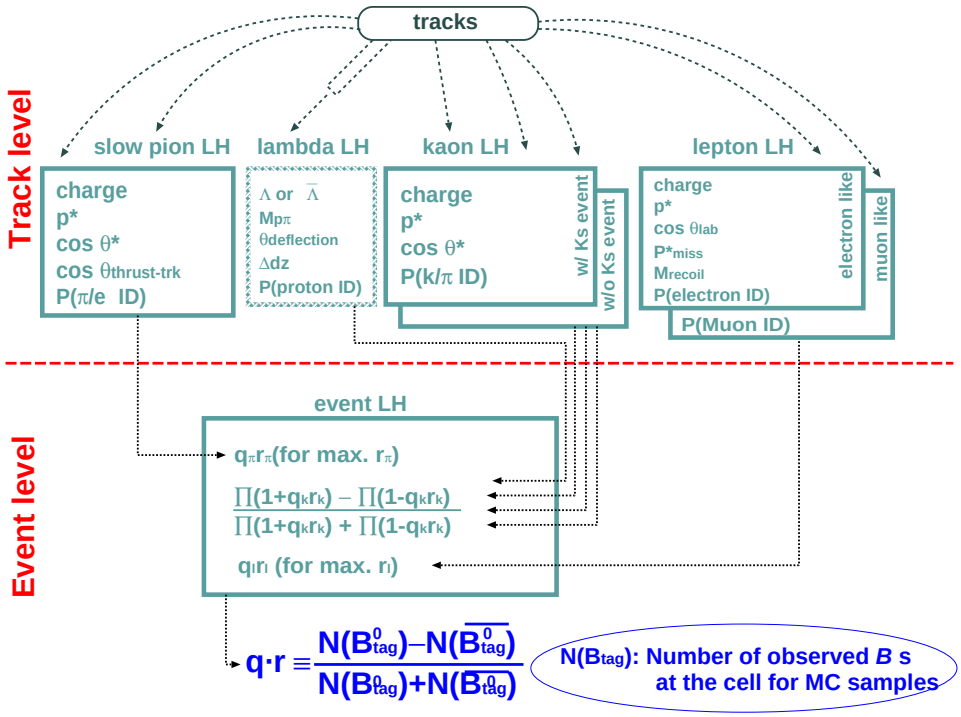

Figure 5.1: Schematic overview of the flavor tagging algorithm [67].

B mesons in the event has been reconstructed as a CP eigenstate, we can assume that all other particles in the event belong to the second B meson. In principle, we could fully reconstruct the second B, which would yield unambiguous information regarding its flavor, but the low reconstruction efficiency for full reconstruction renders this impractical. Instead, we rely on partial reconstruction. Among the prominent flavor-identifying features of B meson decays are slow pions coming from $D^* \to D$ transition, energetic leptons from semileptonic b quark decays and well identified Λ's or kaons from $b \to c \to s$ cascade transitions. Flavor determination is based on a likelihood approach, similar to the one-dimensional likelihood described in Section 4.3. The resulting likelihood functions are constructed using generic $B\bar{B}$ MC. `Hamlet` returns two values: "q" and "r". The value of q is the most likely flavor for the tag particle, $q = 1$ corresponds to B^0_d, while $q = -1$ corresponds to $\bar{B}^0_d$. The parameter r is the likelihood value and characterizes the reliability of the q assignment. If r is properly

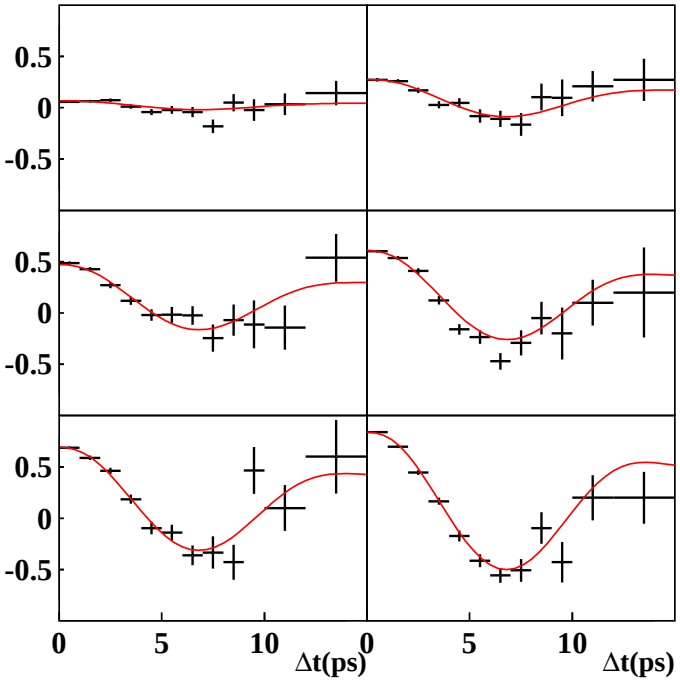

Figure 5.2: Measured time-dependent asymmetries $A(\Delta t)$ (Eq. 5.2) between same-flavor and opposite-flavor events. For each of the six tagging categories ("r-bins"). $B \to D^{*\pm}\, l^{\mp}\, \nu$ decays are used to determine the flavor of one B meson, while the flavor tagging algorithm is used to determine that of the other [67].

normalized, we can write $r = 1 - 2w$, where w is the "wrong-tag fraction." In reality, we obtain r from MC, while w is determined from data. If our MC were perfect, then the r value from the tagging algorithm would equal $1 - 2w$. Any error in the assumed wrong-tag fraction will bias the measured $\sin 2\phi_1$ value. To avoid this bias, the mapping of r into w space is done by reconstructing B^0_d and $\bar{B}^0_d$ decays into self-tagged final states $D^{*\pm}l^{\mp}\nu$, $D^{*\pm}\pi^{\mp}$ and $D^{*\pm}\rho^{\mp}$, where the charge of the decay products defines the flavor of the decaying B meson. After that, we use the flavor-tagging algorithm to assign the tag value for the second B meson. The probabilities of observing same flavor (SF) and opposite flavor (OF) mesons are governed by the following equations:

$$\begin{aligned} P_{\rm OF} &\sim 1 + (1-2w)\cos(\Delta m_d \Delta t) \\ P_{\rm SF} &\sim 1 - (1-2w)\cos(\Delta m_d \Delta t) \end{aligned} \tag{5.1}$$

l	r	w_l
1	0.000-0.250	0.458± 0.006
2	0.250-0.500	0.336± 0.009
3	0.500-0.625	0.228± 0.010
4	0.625-0.750	0.160± 0.009
5	0.750-0.875	0.112± 0.009
6	0.875-1.000	0.020± 0.006

Table 5.1: Wrong-tag fractions (w_l) for the different tagging categories (r bins). The quoted errors on w include both statistical and systematic uncertainties.

which would result in the following time dependent asymmetry.

$$A(\Delta t) = \frac{P_{\rm OF} - P_{\rm SF}}{P_{\rm OF} + P_{\rm SF}} = (1 - 2w)\cos(\Delta m_d \Delta t) \tag{5.2}$$

Measurement of the amplitude of this asymmetry in each bin provides the $(1 - 2w)$ mapping and is presented in Fig. 5.2 and Table 5.1. The errors on the six wrong tag fractions are used when estimating the systematic error on $\sin 2\phi_1$.

5.2 Vertexing.

The time-dependent asymmetry estimation requires precise knowledge of the decay positions for both B mesons. On the CP side, the $B \to J/\psi\ K_L$ vertex position is provided by the two leptons from J/ψ decay. The J/ψ lifetime (about 8×10^{-21} s) is so short that it is effectively zero. The vertex on the tagging side is estimated from a fit to the remaining charged tracks in the event. Tracks from long lived particles, like K_S and Λ, are excluded to prevent a bias in the vertex position. A more detailed discussion of the method is presented in [50].

The RMS resolutions obtained from MC simulation for the B meson vertices are roughly 75 μm for $B_{\rm CP}$ and 140 μm for $B_{\rm tag}$ [50]. The finite resolution for both the CP and tag side, as well as smearing due to the B meson motion in the CM frame can be described by the resolution function $R_{\rm sig}(\Delta t)$. This function provides resolution an event-by-event basis from the z errors given by the two B vertex fits. The resolution

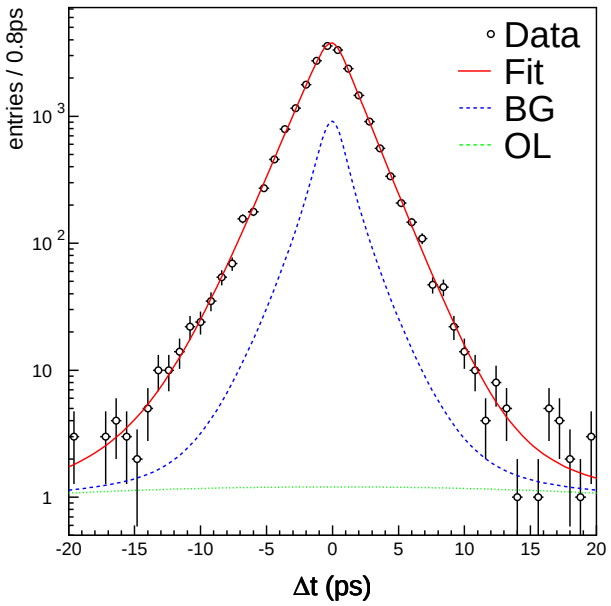

Figure 5.3: Distribution of the proper-time interval Δt, from a B-meson lifetime fit, using 78.1 fb^{-1} of data. One B is fully reconstructed from hadronic decays, while the other B's vertex is determined from the remaining charged tracks in the event. The points with error bars show the experimental data, while the solid line is the fitted PDF. The PDF was obtained by smearing the theoretical expectation with the event-dependent resolution function, $R(\Delta t)$. The yellow line shows a small component of broad outliers, while the blue line shows the sum of all background, including outliers. Excellent agreement between the PDF and experimental data is seen out to ten times the B_d lifetime.

function is measured in fully reconstructed hadronic decays of neutral and charged B decays. A detailed discussion of the function's construction and properties can be found in [68]. A lifetime fit for B meson decays (see Fig.5.3) shows the excellent performance of the resolution function.

5.3 Systematic Uncertainties.

There are a number of parameters that are used as an input to the CP fit without corresponding errors. The uncertainties introduced as a result are taken as part of the systematic error. In order to estimate these errors we vary each experimentally measured parameter by one standard deviation, perform a $\sin 2\phi_1$ estimation, and record the resulting deviation from the fit result. Parameters estimated from MC are

class of parameters	systematic error
signal purity	0.050
BG composition	0.034
vertex reconstruction	0.024
resolution function	0.023
fit bias	0.011
wrong-tag fractions	0.009
physics parameters	0.008
BG shape in fit	0.005
total	0.07

Table 5.2: Contributions to the systematic error on $\sin 2\phi_1$. Each entry shows the cumulative uncertainty arising from a class of cuts and parameters used toward the CP fit

varied by two standard deviations. For cuts, the procedure differs from case to case. We assume no correlations between error distribution. The systematic error estimate is taken from the official release of the $B \to J/\psi K_L$ analysis[48]. That approach is justified, because the CP fit procedure implemented in this analysis is identical to the official release and is based exactly on the same data and MC sample. On top of that signal and background fraction estimates were made with similar techniques, even though in different variable spaces. A detailed discussion of all systematic errors and estimates can be found in [48], here we present only results.

Table 5.2 gives a summary of the observed largest deviations in $\sin 2\phi_1$. By adding these in quadrature, we obtain the total systematic error: $\sigma_{\rm syst} = 0.07$.

The signal purity component incorporates errors from the normalization of four components: signal, background with true K_L's, background with fake K_L's, and background with fake J/ψ's. Background composition errors are due to the uncertainties in the normalization and p_B^* shape of individual background modes, in the particular CP eigenstates, which were determined from the MC. Vertex reconstruction errors describes the presence of the outliers (events with large δt), vertex quality, track-quality and tag-side vertex construction. The fit bias contribution is due to

the possible omission of rare CP eigenstates in the background treatment and effect of a difference between the wrong-tag fractions in $B^0 \to J/\psi K_L$ decays and in the flavor-specific B decays used to measure the wrong tag fractions. The "physics parameters" error are due to the errors in the neutral B meson lifetime, mass, and mixing parameter.

5.4 CP fit

We determine $\sin 2\phi_1$ from an unbinned maximum-likelihood fit to the observed Δt distributions. The probability density function (PDF) expected for the signal is given by

$$P_{\rm sig}(\Delta t, q, w_i) = \frac{e^{-|\Delta t|/\tau_{B^0}}}{4\tau_{B^0}} \left[1 - q(1-2w)\sin 2\phi_1 \sin(\Delta m_d \Delta t)\right] \tag{5.3}$$

where we fix the B^0 lifetime τ_{B^0} and mass difference at their world average values [13]. Each PDF is convolved with the appropriate $R_{\rm sig(\Delta t)}$ to determine the likelihood value for each event as a function of $\sin 2\phi_1$:

$$\begin{aligned} P_i = \quad & (1-f_{\rm ol}) \int [f_{\rm sig} P_{\rm sig}(\Delta t', q, w_i) R_{\rm sig}(\Delta t - \Delta t') \\ & + (1-f_{\rm sig}) P_{\rm bkg}(\Delta t') R_{\rm bkg}(\Delta t - \Delta t')] d(\Delta t') + f_{\rm ol} P_{\rm ol}(\Delta t) \end{aligned} \tag{5.4}$$

where $f_{\rm sig}$ is the signal fraction calculated as a function of $L(J/\psi) \otimes L(B \to J/\psi K_L)$. $P_{\rm bkg}(\Delta t)$ is the PDF for combinatorial background events, which is modeled as a sum of exponential and prompt components. The PDF is convolved with $R_{\rm bkg}$, which is regarded as a resolution function for the background. To account for the small number of events that give large Δt in both signal and background, we introduce a PDF for the outlier component, $P_{\rm ol}$, and its fraction $f_{\rm ol}$. The only free parameter in the final fit is $\sin 2\phi_1$, which is determined by maximizing the likelihood function $L = \prod_i P_i$, where the product is over all events. A test of the event selection and the CP fit procedure on the MC sample is presented in Fig. 5.4. The overall agreement is good. The result of the fit for the data sample is

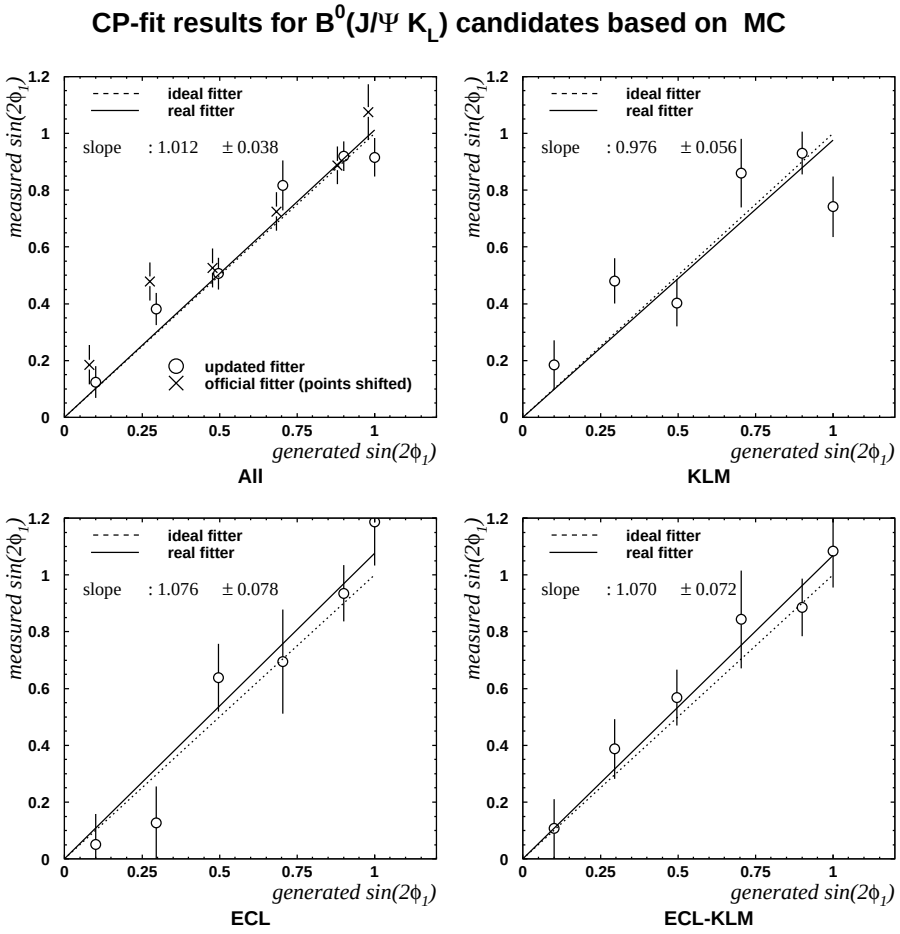

Figure 5.4: MC based test of CP fit.

$$\sin 2\phi_1 = 0.51 \pm 0.12 \text{ (statistical) } \pm 0.07 \text{ (systematic)}$$

This result is based on the 1714 events passed through flavor tagging and vertexing (KLM 635, ECL 815, ECL-KLM 264). The results of the CP fit on a different samples are presented in the Table 5.3. Overall the results are consistent. However, they are on the lower side of the $\sin 2\phi_1$ results from both Belle and BaBar [68, 69][1]. The difference is less then 2σ.

In order to check the consistency of the fit we perform a $\sin 2\phi_1$ estimate for different likelihood ranges (Table 5.4). All of the results are consistent with each other, indicating no evidence of problems with the likelihood method.

5.4.1 Conclusion

The most important result of this thesis is a significant improvement in the statistical sensitivity of the $\sin 2\phi_1$ measurement in $B \to J/\psi K_L$ decay. The statistical error for our method is 25% lower than the official number from Belle, which is to be expected from the larger number of events and lower background level. When I started my Ph.D., people were skeptical about the possibility of measuring $\sin 2\phi_1$ value in the $B \to J/\psi K_L$ decay, however this thesis not only proves them wrong but shows that precision of this measurement approaches the one provided by $B \to J/\psi K_S(\pi^+\pi^-)(\sin 2\phi_1 = 0.73 \pm 0.10 \text{ (statistical)}[68])$[2].

The second result is a successful use of the likelihood approach, which allowed us to build a lean but powerful particle identification and signal quality assessment mechanism. The developed method allows analysis of complex decay channels with large cross channel contamination and partial loss of information.

[1] Belle: 0.719 ± 0.074(statistical) ± 0.035(systematic),
BaBar: 0.741 ± 0.067(statistical) ± 0.034(systematic)

[2] For its distinct signature and low background contamination $B \to J/\psi K_S(\pi^+\pi^-)$ decay even received nickname "gold-plated mode" and until now was providing the bulk of the precision for CP violation measurements in B^0 decays.

mode	$\sin 2\phi_1$		
KLM	0.569	+0.177	-0.184
ECL-KLM	0.763	+0.304	-0.267
ECL	0.365	+0.183	-0.186
total	0.512	+0.118	-0.120
official	0.781	+0.164	-0.171

Table 5.3: $\sin 2\phi_1$ measurement in $B \to J/\psi K_L$ sample with statistical errors only. Belle/BaBar average value for all modes[68, 69] is $\sin 2\phi_1 = 0.730 \pm 0.050$(stat)

	$\sin 2\phi_1$		
$L > 0.4$	0.573	+0.126	-0.129
$L > 0.5$	0.570	+0.136	-0.140
$L > 0.6$	0.561	+0.151	-0.156
$L > 0.7$	0.545	+0.173	-0.181
$L > 0.8$	0.677	+0.214	-0.225
$0.25 < L < 0.50$	0.356	+0.228	-0.229
$0.50 < L < 0.75$	0.623	+0.189	-0.199
$0.75 < L < 1.00$	0.518	+0.191	-0.199

Table 5.4: $\sin 2\phi_1$ measurement in $B \to J/\psi K_L$ sample with statistical errors only.

Appendix A

B meson oscillations

The time evolution of the B^0 system can be written in the following way [8].

$$|\Psi(t)\rangle = a(t)|B_d\rangle + b(t)|\bar{B}_d\rangle + \ decay\ related\ part \tag{A.1}$$

Here we will be interested only in the simple case of pure particle-antiparticle oscillations, where we will discard oscillations that occur through "decay related states." In addition, we will deal with times much greater than the typical strong interaction scale. Given the above, time evolution of the oscillating system, can be described by the two component Schrödinger equation

$$\imath\hbar\frac{\partial}{\partial t}\Psi(t) = H\,\Psi(t) \tag{A.2}$$

the matrix H is given by

$$H\ = M - \frac{\imath}{2}\Gamma = \begin{pmatrix} M_{11} - \frac{\imath}{2}\Gamma_{11} & M_{12} - \frac{\imath}{2}\Gamma_{12} \\ M_{21} - \frac{\imath}{2}\Gamma_{21} & M_{22} - \frac{\imath}{2}\Gamma_{22} \end{pmatrix} \tag{A.3}$$

CPT invariance implies that particles and their antiparticles have the same masses and decay widths [70, 71]. Constraints imposed by CPT invariance on the mass matrix in Eq. A.3 require

$$M_{11} = M_{22},\ \Gamma_{11} = \Gamma_{22} \tag{A.4}$$

The mass and decay rate eigenvalues are given by:

$$\begin{aligned} M_1 - \tfrac{\imath}{2}\Gamma_1 &= M_{11} - \tfrac{\imath}{2}\Gamma_{11} + \tfrac{q}{p}\left(M_{12} - \tfrac{\imath}{2}\Gamma_{12}\right) \\ M_2 - \tfrac{\imath}{2}\Gamma_2 &= M_{11} - \tfrac{\imath}{2}\Gamma_{11} - \tfrac{q}{p}\left(M_{12} - \tfrac{\imath}{2}\Gamma_{12}\right) \end{aligned} \tag{A.5}$$

with

$$\frac{q}{p} = \sqrt{\frac{M_{12}^* - \frac{i}{2}\Gamma_{12}^*}{M_{12} - \frac{i}{2}\Gamma_{12}}} \tag{A.6}$$

The mass and decay width differences ($\Delta M = M_1 - M_2$, $\Delta\Gamma = \Gamma_1 - \Gamma_2$) imply that observable mass states for oscillating systems (like $K_{\text{Short}} - K_{\text{Long}}$ or $B_{\text{Heavy}} - B_{\text{Light}}$), are not the particle-antiparticle combinations. For the case of the B^0 system the observable states can be written as:

$$\begin{aligned} |B_L\rangle &= p|B_d\rangle + q|\bar{B}_d\rangle, \\ |B_H\rangle &= p|B_d\rangle - q|\bar{B}_d\rangle \end{aligned} \tag{A.7}$$

Where B_d and $\bar{B}_d$ are flavor eigenstates containing the $\bar{b}$ and b quarks respectively. The time evolution of a system starting from either a pure B_d or $\bar{B}_d$ state is described by

$$\begin{aligned} |B_d(t)\rangle &= f_+(t)|B_d\rangle + \tfrac{q}{p} f_-(t)|\bar{B}_d\rangle, \\ |\bar{B}_d(t)\rangle &= f_+(t)|\bar{B}_d\rangle + \tfrac{p}{q} f_-(t)|B_d\rangle \end{aligned} \tag{A.8}$$

with

$$f_\pm(t) = \frac{1}{2} e^{-\imath M_1 t} e^{-\frac{1}{2}\Gamma_1 t}\left[1 \pm e^{-\imath \Delta M t} e^{\frac{1}{2}\Delta\Gamma t}\right] \tag{A.9}$$

Denoting the amplitudes for the decay of B_d and $\bar{B}_d$ by $A(f)$ and $\bar{A}(f)$, respectively, into the same final CP eigensstate f, and their ratios by $\rho(f)$ and $\bar{\rho}(f)$, one can write the time evolution for this simplest case, where $A(f) = \bar{A}(f)$ in the form

$$\begin{aligned} &\Gamma(B_d(t) \to f) \propto e^{-\Gamma_1 t}|A(f)|^2 \times \\ &\times \left(1 + e^{\Delta\Gamma t} + Re\left(\tfrac{q}{p}\bar{\rho}(f)\right)\left[1 - e^{\Delta\Gamma t}\right] - 2Im\left(\tfrac{q}{p}\bar{\rho}(f)\right) e^{0.5\Delta\Gamma t}\sin\Delta M t\right) \\ &\Gamma(\bar{B}_d(t) \to f) \propto e^{-\Gamma_1 t}|A(f)|^2 \times \\ &\times \left(1 + e^{\Delta\Gamma t} + Re\left(\tfrac{q}{p}\bar{\rho}(f)\right)\left[1 - e^{\Delta\Gamma t}\right] + 2Im\left(\tfrac{q}{p}\bar{\rho}(f)\right) e^{0.5\Delta\Gamma t}\sin\Delta M t\right) \end{aligned} \tag{A.10}$$

In the above equations we have used $\frac{q}{p}\bar{\rho}(f) = \left(\frac{p}{q}\rho(f)\right)^*$. A close look at Eq. A.10 shows that a CP asymmetry arises if there is a mixing-induced mass splitting ($\Delta M \neq$

0) and if there is a nonzero phase between the transitions involved ($Im\left(\frac{q}{p}\bar{\rho}(f)\right) \neq 0$). Denoting $\frac{q}{p}\bar{\rho}(f) = e^{i2\phi}$, Eq. A.10 becomes

$$\frac{\Gamma(B_d(t) \to f) - \Gamma(\bar{B}_d(t) \to f)}{\Gamma(B_d(t) \to f) + \Gamma(\bar{B}_d(t) \to f)} = \frac{-2e^{0.5\Delta\Gamma t}\sin 2\phi \ \sin \Delta M t}{1 + e^{\Delta\Gamma t} + [1 - e^{\Delta\Gamma t}]\cos 2\phi} \tag{A.11}$$

The mixing parameters, $\Delta\Gamma$ and ΔM, are well measured in semileptonic decays of B mesons, thus allowing a precise determinatioin of $\sin 2\phi$.

The time-dependant analysis of the oscillating system presented above shows that the asymmetry will be present at any given time, but the integral over time will be zero. In the Section 2.2 we defined the strong interaction as CP conserving, while the weak interaction had a CP violating phase. Following the definitions of Eq. 2.5 we can write:

$$\begin{aligned}
V_{if} &= Ae^{+i\delta_W}e^{i\delta_S} & V_{\bar{i}\bar{f}} &= Ae^{-i\delta_W}e^{i\delta_S} \\
V_{i\bar{i}} &= A'e^{+i\delta'_W}e^{i\delta'_S} & V_{\bar{i}i} &= A'e^{-i\delta'_W}e^{i\delta'_S} \\
V'_{if} &= V_{i\bar{i}}V_{\bar{i}\bar{f}} = AA'e^{-i\delta_W+i\delta'_W}e^{i\delta_S+i\delta'_S} & V'_{\bar{i}\bar{f}} &= V_{\bar{i}i}V_{if} = AA'e^{+i\delta_W-i\delta'_W}e^{i\delta_S+i\delta'_S}
\end{aligned}$$

$$\begin{aligned}
\Gamma(i \to f) &\propto A^2 + A^2A'^2 + 2A^2A'\cos\left((2\delta_W - \delta'_W) - (\delta'_S)\right) \\
\Gamma(\bar{i} \to \bar{f}) &\propto A^2 + A^2A'^2 + 2A^2A'\cos\left(-(2\delta_W - \delta'_W) - (\delta'_S)\right)
\end{aligned} \tag{A.12}$$

The corresponding asymmetry can be written in the following way.

$$\frac{\Gamma(i \to f) - \Gamma(\bar{i} \to \bar{f})}{\Gamma(i \to f) + \Gamma(\bar{i} \to \bar{f})} = \frac{2A'\sin(\delta'_S)\sin(2\delta_W - \delta'_W)}{1 + A'^2 + 2A'\cos(\delta'_S)\cos(2\delta_W - \delta'_W)} \tag{A.13}$$

As one would expect, Eq. A.13 has a structure similar to Eq. A.11, with the only difference arising from oscillations. Note that the difference in the decay rates, $\Delta\Gamma$, plays exactly the same role as a nonzero CP conserving phase. The total asymmetry in Eq. A.13 is zero when $\delta'_S = 0$, showing that even though weak forces can mimic a CP non-conserving phase at any given moment, they can not introduce it for the processes averaged over time.

Appendix B

Kinematics

B.1 Notation

We will use bold font, **u**, to denote four-vectors, while a roman font, "p", will be reserved for the magnitude of the three-vector $\vec{p}$. We will use the following definition of four-coordinates:

$$\mathbf{x} = \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix} \quad \text{and} \quad u^i = \frac{dx^i}{ds} = \begin{pmatrix} \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} \\ \frac{\vec{v}}{\sqrt{1-\frac{v^2}{c^2}}} \end{pmatrix} \tag{B.1}$$

where $ds^2 = x^i x_i = c^2dt^2 - dx^2 - dy^2 - dz^2$ and $\vec{v} = (\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt})$. $\vec{v}$ is the usual velocity three-vector, and u^i is a velocity four-vector. In this notation the energy-momentum vector will look like:

$$\mathbf{p} = \begin{pmatrix} p^0 \\ p^1 \\ p^3 \\ p^4 \end{pmatrix} = \begin{pmatrix} E/c \\ p_x \\ p_y \\ p_z \end{pmatrix} = mc\mathbf{u} = mc \begin{pmatrix} u^0 \\ u^1 \\ u^2 \\ u^3 \end{pmatrix} \tag{B.2}$$

and the scalar product of two four-vectors can be written in the following way.

$$\mathbf{p}_1 \ \mathbf{p}_2 = \frac{E_1 E_2}{c^2} - \vec{p}_1 \cdot \vec{p}_2 \tag{B.3}$$

In most of the subsequent formulas we will use a system of units where $c = 1$.

B.2 Particle decay into the massive final states

K_L reconstruction relies only on the measurement of the K_L direction. As a result, any decay channel involving a K_L in the final state requires one extra constraint in order to define the momentum of the K_L. In the case of B meson decays we will use the B meson mass as the constraint. The corresponding kinematics are presented below.

We will describe the decay of a massive initial state, X, with mass M_X into the massive final state Z (mass M_Z) plus a collection of other particles, which are fully reconstructed and thus have a defined four-momentum. For the massive state Z we require only the direction to be known. The equations governing energy-momentum conservation for a such system can be summarized as follows

$$\begin{aligned} \mathbf{p}_X &= \mathbf{p}_Y + \mathbf{p}_Z \\ (\mathbf{p}_X)^2 &= m_X^2 \\ (\mathbf{p}_Z)^2 &= m_Z^2 \end{aligned} \tag{B.4}$$

Squaring the equation for $\mathbf{p}_X$ and taking into account the rest yields

$$m_X^2 - \left(E_Y^2 - (\vec{p}_Y)^2 + m_Z^2\right) = 2 \ \mathbf{p}_Y \ \mathbf{p}_Z \tag{B.5}$$

which can be rewritten in the following way

$$m_X^2 - \left(E_Y^2 - (\vec{p}_Y)^2 + m_Z^2\right) = 2E_Y\sqrt{m_Z^2 + p_Z^2} - 2p_Y p_Z \cos\theta \tag{B.6}$$

In order to simplify the subsequent equations let us introduce an additional coefficients.

$$A = \frac{m_X^2 - \left(E_Y^2 - (\vec{p}_Y)^2 + m_Z^2\right)}{2E_Y}$$
$$B = \frac{p_Y \cos\theta}{E_Y} \tag{B.7}$$

Solving Eq. B.6 and using the definitions of Eq. B.7 leads to the following result

$$p_Z = \frac{AB \pm \sqrt{A^2B^2 + (1-B^2)(A^2 - m_Z^2)}}{1-B^2} \tag{B.8}$$

Eq. B.8 provides two possible solutions. In order to understand which one is meaningful we note that $B < 1$ follows from $p_Y^2 + m_Y^2 = E_Y^2$. The second step relies on the masses of the particles involved in $B \to J/\psi\ K$ decay. Given that the mass of the B meson is much larger than the mass of the kaon, it follows that $A^2 - m_Z^2 > 0$ which implies that the square-root is larger than AB. Since a negative value for p_Z is meaningless, there is no ambiguity for $B \to J/\psi\ K$ decay and the last formula can be written in the compact form with minus sign omitted, i.e.,

$$p_Z = \frac{AB + \sqrt{A^2 - m_Z^2(1-B^2)}}{1-B^2} \tag{B.9}$$

B.3 Kinematic end point in the two body decays

For the inclusive decays $A \to B\ X$, where A and B are particles with well defined mass.

$$\begin{aligned} \mathbf{p}_A &= \mathbf{p}_B + \mathbf{p}_X \\ (\mathbf{p}_A)^2 &= m_A^2 \\ (\mathbf{p}_B)^2 &= m_B^2 \end{aligned} \tag{B.10}$$

In the A rest frame, the momentum of B and X will have the same magnitude and opposite directions, so the above system of equations can be rewritten as

$$m_A^2 = m_B^2 + m_X^2 + 2\sqrt{m_B^2 + |\vec{p}|^2}\sqrt{m_X^2 + |\vec{p}|^2} + 2|\vec{p}|^2 \tag{B.11}$$

particle	$\psi(2S)$	J/ψ	D	K	γ
endpoint, GeV/c	1.35	1.73	2.31	2.62	2.64

Table B.1: Momentum endpoint values for inclusive B meson decays in the B rest frame.

which has the following solution.

$$p = \frac{\sqrt{(m_A^2 - (m_B + m_X)^2)(m_A^2 - (m_B - m_X)^2)}}{2m_A} \tag{B.12}$$

The highest possible momentum for the decay products is achieved in decays where X is massless and is given by the following expression.

$$p_{\max} = \frac{m_A^2 - m_B^2}{2m_A} \tag{B.13}$$

Endpoint values for selected inclusive B meson decays are summarized in Table B.1. In our experiment, the beam center of mass frame serves as a good approximation to the B-meson rest frame. However, one must take into account the smearing induced by the B momentum in the CM. MC simulation shows that CM-frame end-point values should be increased by approximately 400 MeV/c. Fig. 4.2 shows an example of an inclusive momentum spectrum. One can clearly see the endpoint of the inclusive $B \to J/\psi\ X$ transition around 2 GeV/c. The events at the higher values of CM momentum can not come from B decays and are due to other processes. The endpoint momentum can be used as a continuum veto, especially for abundant final states with clean particle ID signatures, such as π^0, ω, K_S, D^*.

B.4 Angular correlations

The decays of particles with non-zero spin often exhibit a non-uniform angular distribution of their decay daughters. These distributions can be used either for assigning

the right spin values for newly discovered particles or for suppressing unwanted backgrounds.

The conventional quantum mechanical treatment of an angular momentum works well in non-relativistic cases, but runs into problems when dealing with highly energetic particles, because the angular projection eigenstates (J_z or $\vec{J} \cdot \vec{n}$) are mixed by the boosts.

There are two main approaches to tackle this problem, named the tensor and helicity formalism. The covariant tensor formalism, which is often referred to as the Rarita-Schwinger formalism [72], was actually first fully explored by Zemach [73, 74]. For an overview of tensor properties and a history of the subject, the reader can consult the recent work of V. Filippini [75, 76, 77]. Here we will be using the more conventional helicity approach, which was first developed by Jackob and Wick [78] For later examples see [79]. A proper treatment of the helicity amplitudes can be found in [80]. Both formalisms are often used in noncovariant form, because for the most trivial cases good fits can be obtained without requiring Lorentz invariance. This work is not an exception, because the decays under study have a simple topology. A description of the covariant helicity-coupling amplitudes can be found in the work of Chung [81, 82, 83].

The helicity formalism is more general than the tensor one, because it is valid for any mass and spin of the particles involved in the decay [76]. However, we will not use a covariant formalism here. The complications brought by the Lorentz transformations can be circumvented by using the helicity operator $h = \vec{J} \cdot \vec{p}/|\vec{p}|$, which commutes with both the total angular momentum J^2 and boosts along the helicity axis $\hat{p}$. The helicity operator mimics a good quantum number for certain configurations. The helicity is best used for two-body decays $A \to B\,C$, where in the rest frame of either A, B or C, the remaining two particles are collinear, and their direction is a natural choice for the quantization axis.

The naming convention for the helicity angles comes from the CLEO collaboration

and its definition is presented below together with another common angular variable — the decay angle.

- The helicity angle is defined for two-generation transitions, such as $A \to B\ C$ and $B \to D\ F$. The B helicity angle is the cosine of the angle between the direction of D and A in the rest frame of B.

- The decay angle is defined for one-generation transitions, such as $A \to B\ C$. The B decay angle is the cosine of the angle between the boost direction of A and the direction of B in the rest frame of A.

A well-known case for the helicity angle application is the decay of a spin-zero particle into a spin-zero and a spin-one particle, followed by the decay of the spin one particle into spinless states: $A^{\sigma(0)} \to B^{\sigma(1)}\ C^{\sigma(0)}$ where $B \to D^{\sigma(0)}\ F^{\sigma(0)}$. The classic example is $D_s \to \phi\ \pi$ decay, where $\phi \to KK$. For that decay the helicity angle is called the "ϕ helicity" and the best frame for writing the angular amplitudes is the ϕ rest frame, because there we will have both the spinless particles D_s and π aligned. This ensures that the total spin-angular momentum projection of the pair is 0. Another pair of spinless particles $K\bar{K}$ will form another spin-zero projection axis. The angle between these two axes is called the ϕ helicity angle.

The production reaction $e^+e^- \to \Upsilon(4S) \to B\ \bar{B}$ with subsequent decay $B \to J/\psi\ K$ also presents non-uniform helicity distributions. The $\Upsilon(4S)$ helicity originates from the electron-positron interaction and exploits the spinless nature of the B meson for the reaction $e^+e^- \to \gamma^* \to \Upsilon(4S) \to B\bar{B}$. For both vector and pseudo-vector couplings, the result of the e^+e^- interaction must be in the helicity states $(+1/2, -1/2)$ or $(-1/2, +1/2)$. The relativistic electron can only have a $\pm 1/2$ helicity projection on its direction, while the virtual photon must be longitudinally polarized. The spin one $\Upsilon(4S)$ intermediate state can be viewed as a spectator, which allows one to access all possible helicity states of the photon. Thus the $\Upsilon(4S)$ system will decay from the ± 1 helicity states with respect to the beam direction. But for the decay $\Upsilon(4S) \to B\bar{B}$

both B mesons are spinless and therefore must have zero total helicity along the $B\bar{B}$ flight direction in the $\Upsilon(4S)$ rest frame. The angular distribution is then given by an incoherent sum over the squares of d-functions for rotating $m = \pm 1$ onto $m' = 0$, for $l = 1$. Both of these, squared, give $\sin^2(\theta_h)$.

The angular distribution for the B helicity angle is trivial. Because the B meson is a spinless particle we do not expect any correlations between its decay products and those of its parent particles.

The situation for the J/ψ helicity angle is a carbon copy of the $\Upsilon(4S)$ helicity, with the dilepton decay through $m = \pm 1$ helicity states being the final state instead of the initial state.

Appendix C

The $c\bar{c}$ system.

The charmonium system is a series of heavy-meson states formed from $c\bar{c}$ quark pairs. Charmonium states follow the standard naming convention for $q\bar{q}$ mesons, for which the four possible PC combinations ($-+, +-, --$ and $++$) are denoted as $\eta_c,\ \psi,\ h_c$ and χ_c, respectively.

The resemblance between the quarkonium energy-level structure and the electron-positron system prompted the use of spectroscopic notation for charmonium states. The main energy levels are characterized by the "principle quantum number." The angular momentum is denoted L and the total angular momentum, which takes into account spin, is given by $J = L + S$. Energy levels are denoted as $n\ ^{2S+1}L_J$, where $L = S,\ P,\ D,\ F,\ G,\ H,\ I,\ K,$ ... for $L = 0,\ 1,\ ,2,\ ,$ .... Charge and parity can be calculated using the following expressions $P = (-1)^{L+1},\ C = (-1)^{L+S}$, and for G parity $= (-1)^{L+S+I}$ (the C quantum number is only relevant to neutral mesons). Exotic mesons with J^{PC} quantum numbers that a $q\bar{q}$ system can not have, namely $J^{PC} = 0^{--}, 0^{+-}, 1^{-+}, 2^{+-}, 3^{-+}, \ldots$ will be omitted from further discussion.

In the most favorable decay a "hidden charm" $c\bar{c}$ meson will produce two "open charm" D mesons ($c\bar{u}$ or $c\bar{d}$) through gluon emission. But charmonium levels corresponding to $n = 1$ and $n = 2$ are located below the DD production threshold, thus making DD production energetically forbidden. In addition, a color-neutral $c\bar{c}$

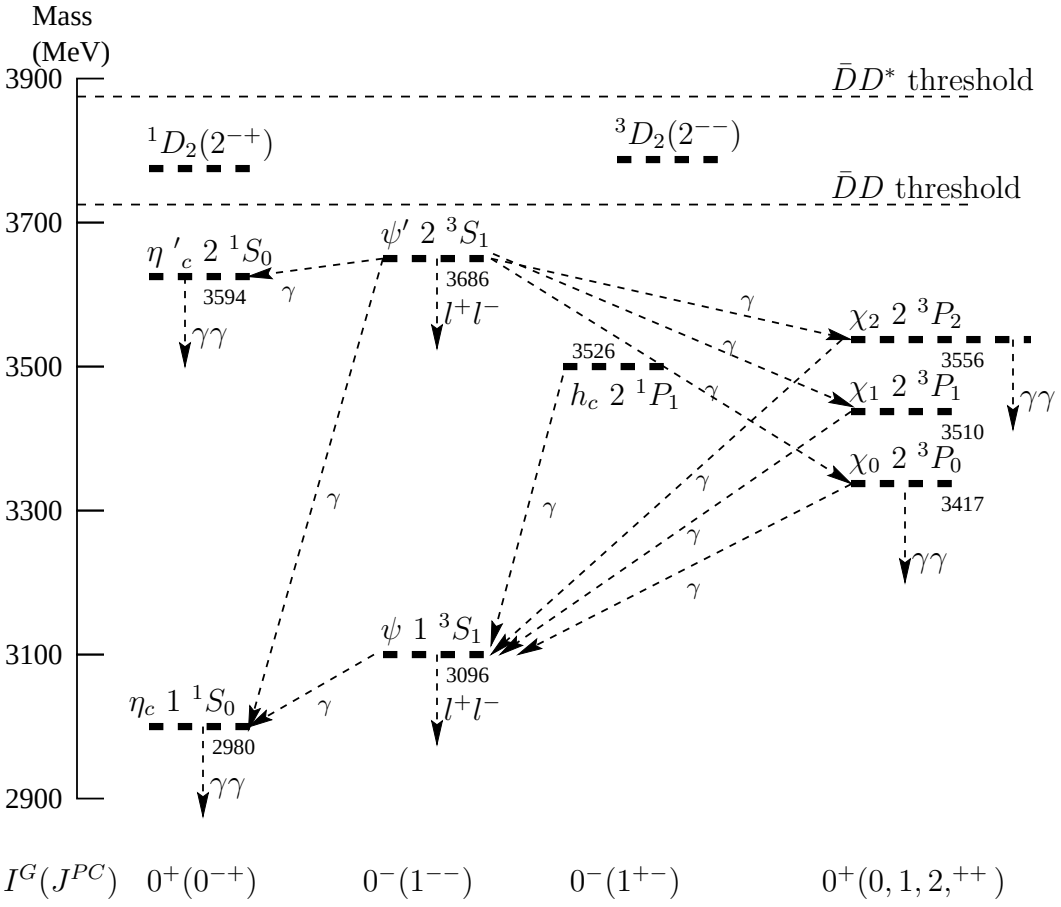

Figure C.1: Charmonium spectra with dominant electromagnetic transitions. Hadronic transitions not shown.
$P = (-1)^{L+1}$, $C = (-1)^{L+S}$, and G parity $= (-1)^{L+S+I}$.
Selection rules for dipole radiation: $\Delta J = 0,\ \pm 1$ and $\Delta L = 1$.

meson cannot decay through only one gluon because the gluon carries color. Two-gluon decays are dominant for $C = +1$ states (η_c, χ_c), but for $C = -1$ states (ψ, h) two-gluon processes are forbidden due to charge-symmetry violation, thus moving this process into the 3-gluon realm. This suppression is often called the Zweig rule for unconnected quark lines. The most interesting consequence of this suppression is that the probability of electromagnetic annihilation becomes comparable with that for a hadronic transition. Given that lepton signatures are much more prominent in most modern detectors, it is not surprising that the first $c\bar{c}$ meson was discovered through leptonic decay ($J/\psi \to l^+l^-$ or $J/\psi \to \gamma^* \to l^+l^-$).

Appendix D

B meson momentum in the beam center of mass frame.

The energies of both beams of the electron-positron collider are known with a high precision. Indeed they are so well defined that for an experiment running on the $\Upsilon(4S)$ resonance, the energies of the $\Upsilon(4S)$ decay products are constrained not only by the natural width of the resonance but rather by the beam energy spread. The $\Upsilon(4S)$ predominantly decays into a $B\bar{B}$ pair and one naively expects a rather sharp CM momentum spectrum of the resulting B mesons. Unfortunately, the kinematics are such that a small uncertainty in the beam energy translates into a large error in the B meson CM momentum, p_B^*. The shape of the detected B meson momentum is affected by the beam energy spread, $\Upsilon(4S)$ mass shape and the detector resolution. In the next several sections we will estimate the contributions from each of the above factors.

We will start with the assumption that the central values for both beam energies are the same as those specified in the technical design report. Later we will discuss the effects of the mean value drift, but for now we will concentrate on the spread due to the beam itself.

Value	e^-	e^+	Units
$\langle E \rangle$	7.9965	3.500	GeV
δE	5.4	2.5	MeV

Table D.1: Technical design report values for the beam energy and its spread.

Beam Energy spread

The KEKB technical design report specifies that the beam energy spread should not exceed the fractional values listed below. For the electron and positron beams, these are $\delta E_+ = 7.1 \times 10^{-4} \langle E_+ \rangle$ and $\delta E_- = 6.7 \times 10^{-4} \langle E_- \rangle$, as presented in Table D.1.

The values for the beam energies were chosen to be on the $\Upsilon(4S)$ resonance ($M_{\Upsilon(4S)} = 10.580$ GeV/c^2). The central values for both the CM beam energy and momentum are given by the following expression.

$$\begin{aligned} \langle E^*_{\rm beam} \rangle &= \sqrt{2}\sqrt{\langle E_+ \rangle \langle E_- \rangle + \sqrt{\langle E_+ \rangle^2 - m_e^2}\sqrt{\langle E_- \rangle^2 - m_e^2} * cos\delta + m_e^2} = 10.580 \text{ GeV} \\ \langle p^*_{\rm beam} \rangle &= 0 \end{aligned} \tag{D.1}$$

$$\gamma = 1.09 \quad \beta = 0.391 \quad \gamma\beta = 0.425 \tag{D.2}$$

The effect of the electron mass is unobservable and we keep it in the formulas above only for completeness. The correction corresponding to the finite crossing angle, $\delta = 22$ mrad, is of order 0.99994, which translates into a 0.7 MeV energy shift, large enough to be observable in the experiment.

The beam energy spread can be estimated with approximate formulas, which neglect the electron mass and the crossing angle.

$$\begin{aligned} \langle E^*_{\rm beam} \rangle &= 2\sqrt{\langle E_- \rangle \langle E_+ \rangle} \\ \langle p^*_{\rm beam} \rangle &= 0 \end{aligned} \tag{D.3}$$

The energy spread can be obtained from Eq. D.3, assuming an uncorrelated gaussian energy distribution for both beams.

$$\delta E^*_{\text{beam}} = \sqrt{\frac{\langle E_- \rangle}{\langle E_+ \rangle} (\delta E_+)^2 + \frac{\langle E_+ \rangle}{\langle E_- \rangle} (\delta E_-)^2} \approx 5.2 \text{ MeV} \tag{D.4}$$

In the same approximation, the momentum for both beams can be written as vectors aligned with the z axis and a momentum equal in magnitude to the corresponding energy. We define the boost parameters to be constant, because we cannot predict the exact energy value for any given event. As such, the beam center of mass frame is defined as a frame with the parameters corresponding to the technical design report boost.

$$\begin{aligned} \left\langle p^*_- \right\rangle &= \gamma \langle p_- \rangle - \beta\gamma \langle E_- \rangle \approx \gamma (1 - \beta) \langle E_- \rangle \\ \left\langle p^*_+ \right\rangle &= \gamma \langle p_+ \rangle + \beta\gamma \langle E_+ \rangle \approx \gamma (1 + \beta) \langle E_+ \rangle \end{aligned} \tag{D.5}$$

The beam momentum spread will be limited in the CM frame as well.

$$\delta p^*_{\text{beam}} = \gamma \sqrt{((1 - \beta)\, \delta E_-)^2 + ((1 + \beta)\, \delta E_+)^2} \approx 4.8 \text{ MeV}/c \tag{D.6}$$

That can be used as a powerful constraint for partial reconstruction of the second B meson in cases where the first one is already fully reconstructed.

The energy distribution for the produced $\Upsilon(4S)$ resonances is affected by both the beam energy spread and the shape of the $\Upsilon(4S)$ cross section. In the extreme case of the narrow $\Upsilon(4S)$ resonance the energy of the products will be fully defined by $\Upsilon(4S)$ width. In the other extreme, where the machine energy is much smaller than the natural width of the $\Upsilon(4S)$, the machine setting determines the final state energy. At KEKB, we have comparable contribution from both components. The natural width of the $\Upsilon(4S)$ resonance is 14 ± 5 MeV ([65, 19, 84]) and a naive comparison with the result of the Eq. D.4 would suggest disregarding the effects induced by the $\Upsilon(4S)$ shape. However, the resonance width is always quoted for the Breit-Wigner distribution, for which the variance and higher moments are infinite. The Breit-Wigner distribution is defined by the position of its maximum (about which the distribution is symmetric), and by the full width at half maximum (FWHM). Thus

when we compare errors generated by the Breit-Wigner shape with Gaussian ones, it is better to use FWHM instead of variance (for gaussian $FWHM = 2.355\sigma = 12.2$ from Eq. D.4). The comparable FWHM for both contributions make the energy distribution of the produced $\Upsilon(4S)$ dependent on both beam energy spread and $\Upsilon(4S)$ cross section. In the current experiment we can not disentangle the effects caused by the energy spread from effects induced by the $\Upsilon(4S)$[1], and can only measure the convolution of the above parameters. That results in reduced energy spread and can be approximated with inverse quadrature:

$$\left(\frac{1}{\delta E^*}\right)^2 \approx \left(\frac{1}{\delta E^*_{beam}}\right)^2 + \left(\frac{2.35}{\delta \Gamma^*_{\Upsilon(4S)}}\right)^2 \quad \Rightarrow \quad \delta E^* \approx 3.9 \text{ MeV} \tag{D.7}$$

$\Upsilon(4S)$ decay into B meson pair.

The main goal of our experiment is to study B meson pairs produced in $\Upsilon(4S)$ decays. Knowledge of the p^*_B spectrum shape is essential to our selection process and as a result we have to understand how the beam energy spread affects it.

$$\begin{aligned} p^*_B &= \sqrt{\tfrac{E^2_{tot}}{4} - m^2_B} \approx \sqrt{2m_B(\tfrac{E_{tot}}{2} - m_B)} \\ \delta p^*_B &= \tfrac{\delta E_{tot}}{2}\sqrt{\tfrac{m_B}{E_{tot}-2m_B}} \end{aligned} \tag{D.8}$$

The above formulas are dependent on the total energy and the mass of the B meson, thus producing different peak positions for charged and neutral B's. In the assumption that we are running on the top of the resonanse, we can use PDG values for the masses [65].

$$\begin{aligned} M_{B^\pm} &= (5279.4 \pm 0.5) \text{ MeV}/c^2 \\ M_{B^0} &= (5279.0 \pm 0.5) \text{ MeV}/c^2 \\ M_{\Upsilon(4S)} &= (10580.0 \pm 3.5) \text{ MeV}/c^2 \end{aligned} \tag{D.9}$$

[1] It is possible to measure both of this values, but such measurement will require a precise energy scan for the $\Upsilon(4s)$ resonance [85].

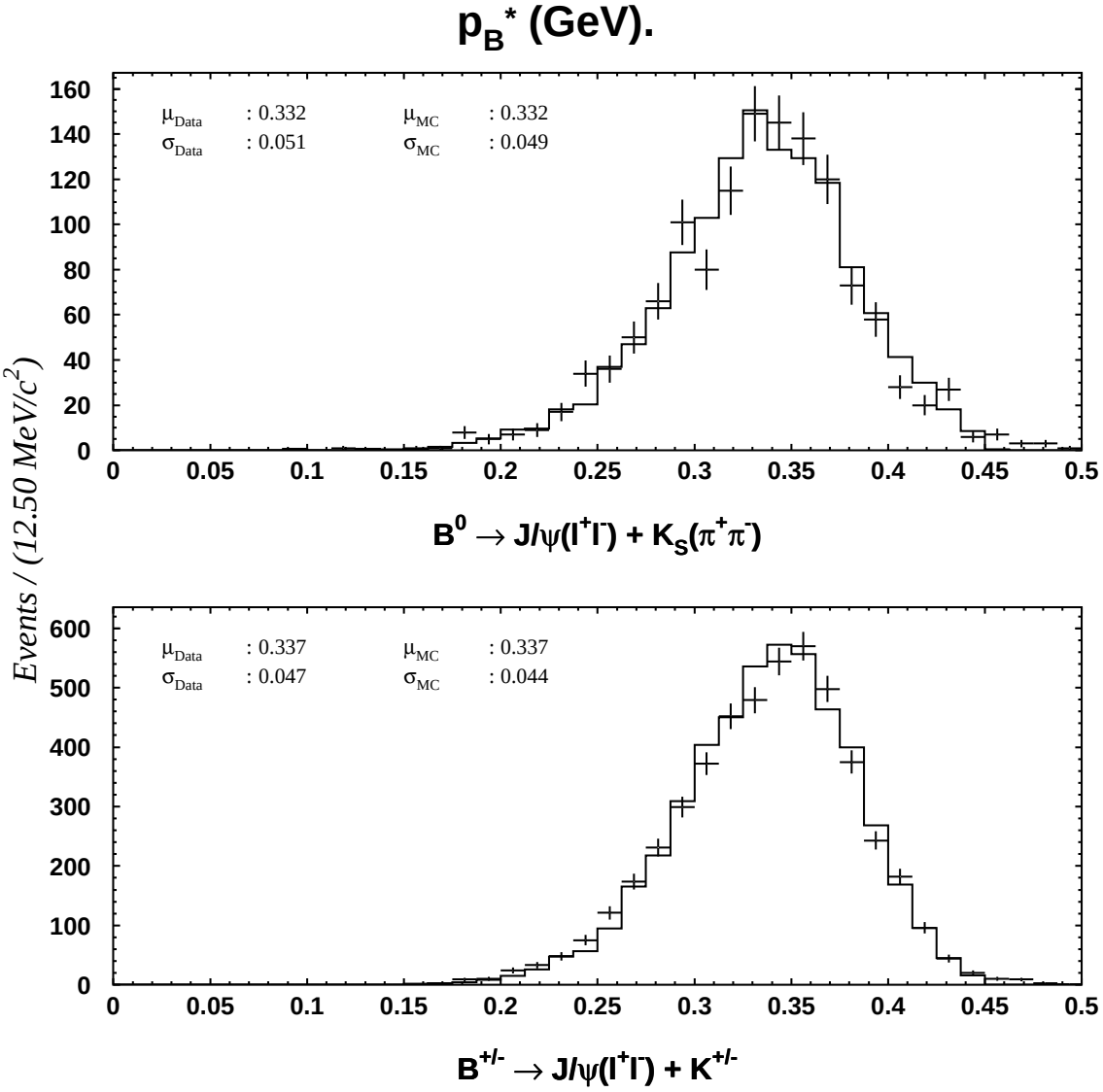

Figure D.1: The B meson momentum distribution in the beam center of mass frame. The solid line corresponds to the events reconstructed from the inclusive J/ψ MC, while crosses represent the data. MC histograms were scaled to the data.

which will result in the following average CM momentum values with corresponding spreads:

$$\begin{aligned} p^*_{B^\pm} &= (334.6 \pm 41)\ \mathrm{MeV}/c \\ p^*_{B^0} &= (340.8 \pm 40)\ \mathrm{MeV}/c \end{aligned} \tag{D.10}$$

The large p^*_B spread is a consequence of the small difference between the CM collision energy and the mass of the B-meson pairs. In the center of mass frame we have only $E_{tot} - 2m_B \approx 20$ MeV available for the kinematic motion of the B meson. The momentum spread will equal the available energy spread magnified by the factor $\frac{1}{2}\sqrt{\frac{m_B}{E_{\mathrm{tot}}-2m_B}} \approx 8.0$. This dependence could be a useful tool for an estimation of the accelerator performance. The momentum resolution, pertinent to the B meson reconstruction, will not significantly affect the width of the p^*_B spectrum, because its contribution is small. The typical resolution is on the order of 10 MeV/c for the decays involving only charged tracks and 20 MeV/c for neutral-particle modes, thus contributing only $1 \sim 2$ MeV to the energy resolution. However, in decays involving multiple neutral (or unmeasured) particles in the final state the detector resolution should be taken into account.

Beam energy drift.

In the experiment, the mean of the beam energy changes with time. The drift is slow enough to leave δp^*_B unaffected over short time intervals. Over the long term, however, the drift is sufficiently large to produce noticeable effects, which require correction.

Fortunately, the accelerator design implementation guarantees that the energy drifts for the electron and positron rings stay highly correlated. As a result, the combined beam momentum will stay constant and we do not have to worry about the changes in the boost vector. The only observable effect of the energy drift will be a shift in the mean value of p^*_B. At the Belle experiment we use fully reconstructed B meson decays in order to extract the run dependant beam energy value. The

relatively abundant decays $B \rightarrow DX$ provide enough statistics to calculate energy corrections for the less likely decays. Fig. D.1 shows the spectrum for the energy corrected $B \rightarrow J/\psi\ K$ decays. The MC was generated assuming PDG [65] values for the masses involved and the beam parameters stipulated in the accelerator design, but no effort was made to account for the shape of the $\Upsilon(4S)$ resonance. In the MC the $\Upsilon(4S)$ energy shape was defined by the convolution of the gaussian beam spreads for both beams. The shape of the p_B^* spectrum for the data agrees well with MC prediction.

Appendix E

Likelihood Plots.

Here we collect the standard likelihood plots. Each page in this appendix is dedicated to some decay channel and consist of two parts: on top one can find a plot describing the likelihood performance, while the bottom part presents the likelihood discriminants.

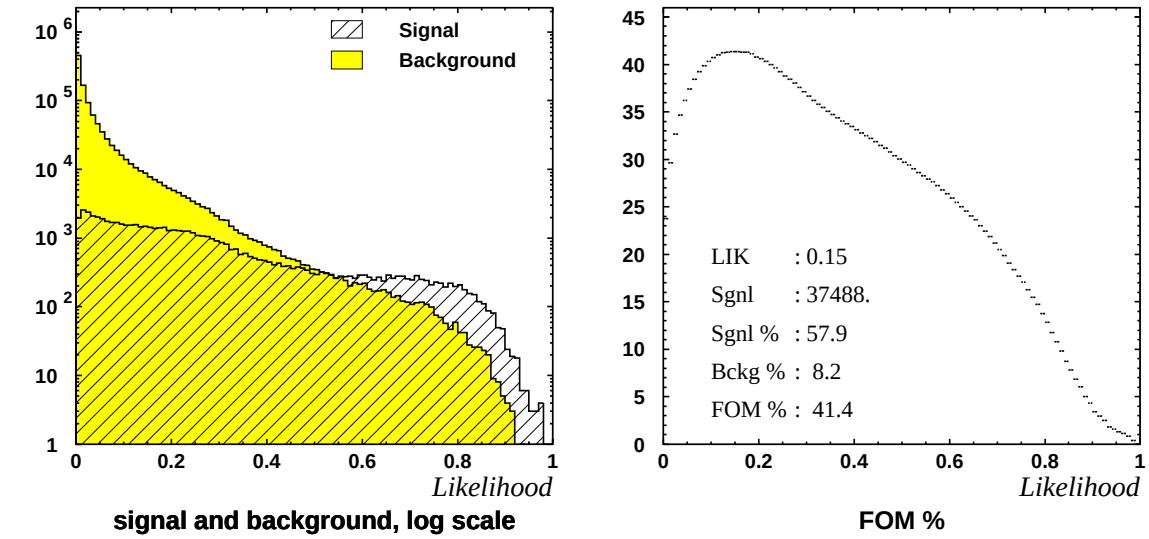

Figure E.1: The $K^{*0}(892) \rightarrow K^{\pm}\ \pi^{\mp}$ likelihood performance. The plots are based on a 0.5 M-event J/ψ inclusive MC sample. For a description of the plot parameters see the discussion of Fig. 4.10.

The $K^{*0}(892) \rightarrow K^{\pm}\ \pi^{\mp}$ likelihood function is based on the following two dimensional discriminants:

- K^* decay angle vs. K^* momentum.
- K^* mass vs. K^* CM momentum.
- K likelihood vs. π likelihood.

The performance of the likelihood based on a 0.5-M-event inclusive J/ψ MC sample is summarized in Fig. E.1.

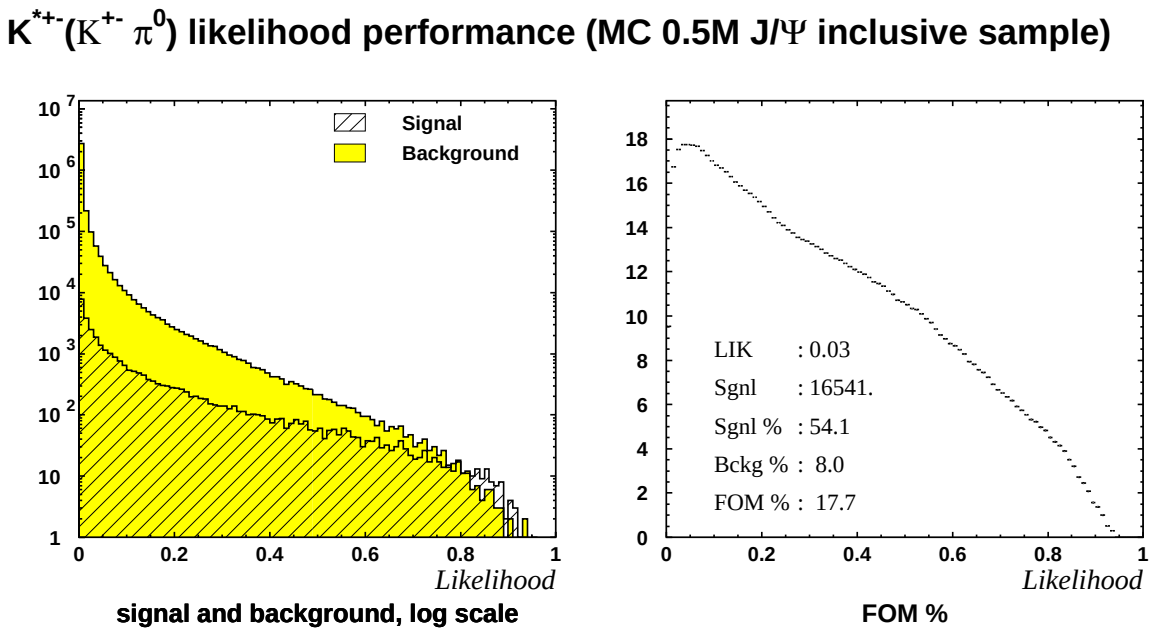

Figure E.2: The $K^{*\pm}(892) \rightarrow K^{\pm}\ \pi^0$ likelihood performance. The plots are based on a 0.5 M-event J/ψ inclusive MC sample. For a description of the plot parameters see the discussion of Fig. 4.10.

The $K^{*\pm}(892) \rightarrow K^{\pm}\ \pi^0$ likelihood function is based on the following two dimensional discriminants:

- K^* decay angle vs. K^* momentum.
- K^* mass vs. K^* CM momentum.
- K likelihood vs. π likelihood.

The performance of the likelihood based on the 0.5-M-event inclusive J/ψ MC sample is summarized in Fig. E.2.

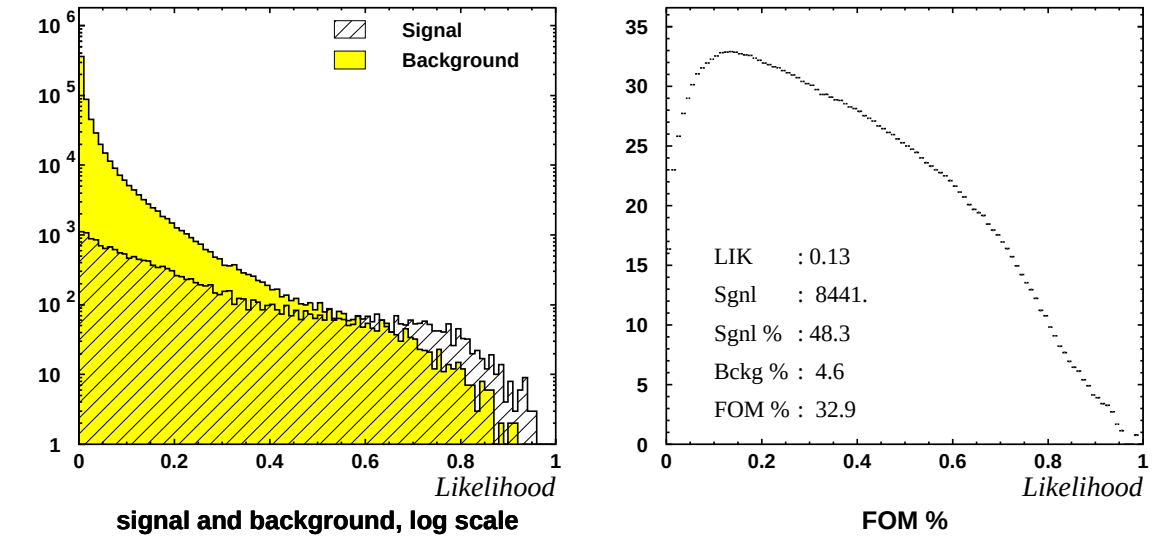

Figure E.3: The $K^{*\pm}(892) \rightarrow K_S\ \pi^{\mp}$ likelihood performance. The plots are based on a 0.5 M-event J/ψ inclusive MC sample. For a description of the plot parameters see the discussion of Fig. 4.10.

The $K^{*\pm}(892) \rightarrow K_S\ \pi^{\mp}$ likelihood function is based on the following two dimensional discriminants:

- K^* decay angle vs. K^* momentum.
- K^* mass vs. K^* CM momentum.
- K likelihood vs. π likelihood.

The performance of the likelihood based on a 0.5-M-event inclusive J/ψ MC sample is summarized in Fig. E.3.

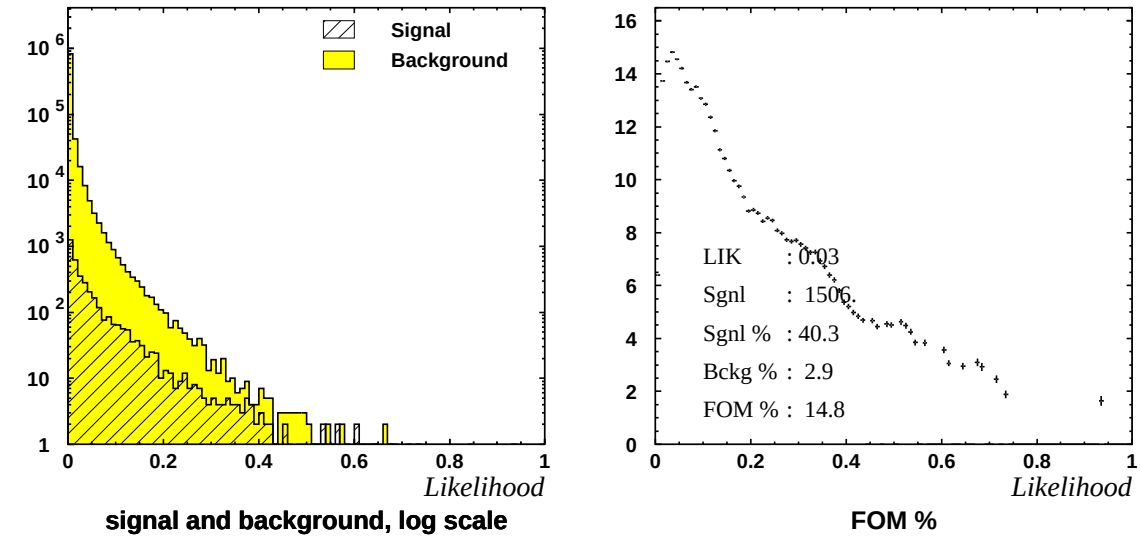

Figure E.4: The $K^{*0}(892) \rightarrow K_S\ \pi^0$ likelihood performance. The plots are based on a 0.5 M-event J/ψ inclusive MC sample. For a description of the plot parameters see the discussion of Fig. 4.10.

The $K^{*0}(892) \rightarrow K_S\ \pi^0$ likelihood function is based on the following two dimensional discriminants:

- K^* decay angle vs. K^* momentum.
- K^* mass vs. K^* CM momentum.
- K likelihood vs. π likelihood.

The performance of the likelihood based on a 0.5-M-event inclusive J/ψ MC sample is summarized in Fig. E.4.

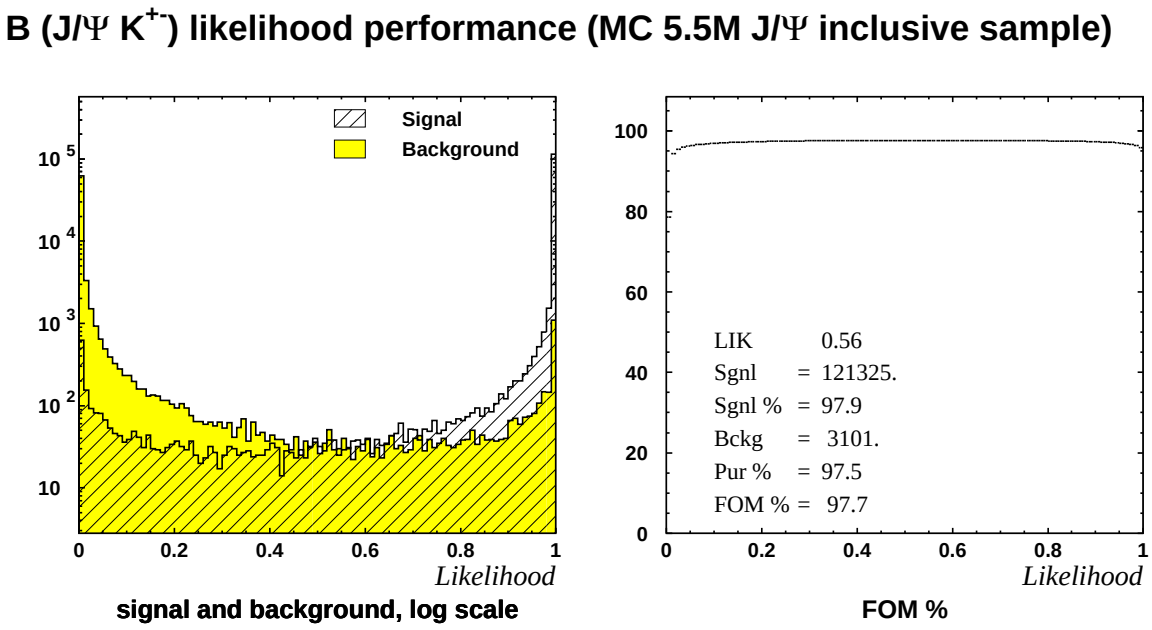

Figure E.5: The $B^{\pm} \to J/\psi\ K^{\pm}$ likelihood performance. The plots are based on a 5.5-M-event J/ψ inclusive MC sample. For a description of the plot parameters see the discussion of Fig. 4.10.

The $B^{\pm} \to J/\psi\ K^{\pm}$ likelihood function is based on the following two dimensional discriminants:

- $M(B) - M(J/\psi) + 3.096$ GeV/c^2 vs. B CM momentum.
- K likelihood vs. B decay angle.
- B vs. J/ψ decay angle.
- $\Upsilon(4S)$ vs. J/ψ helicity.

The performance of the likelihood based on a 5.5-M-event inclusive J/ψ MC sample is summarized in Fig. E.5. The overall separation power is excellent with 97% of the signal coming in the signal region. The largest sources of background are due to $B^{\pm} \to J/\psi\ \pi^{\pm}$ and combinatorial background with true J/ψ and a random kaon.

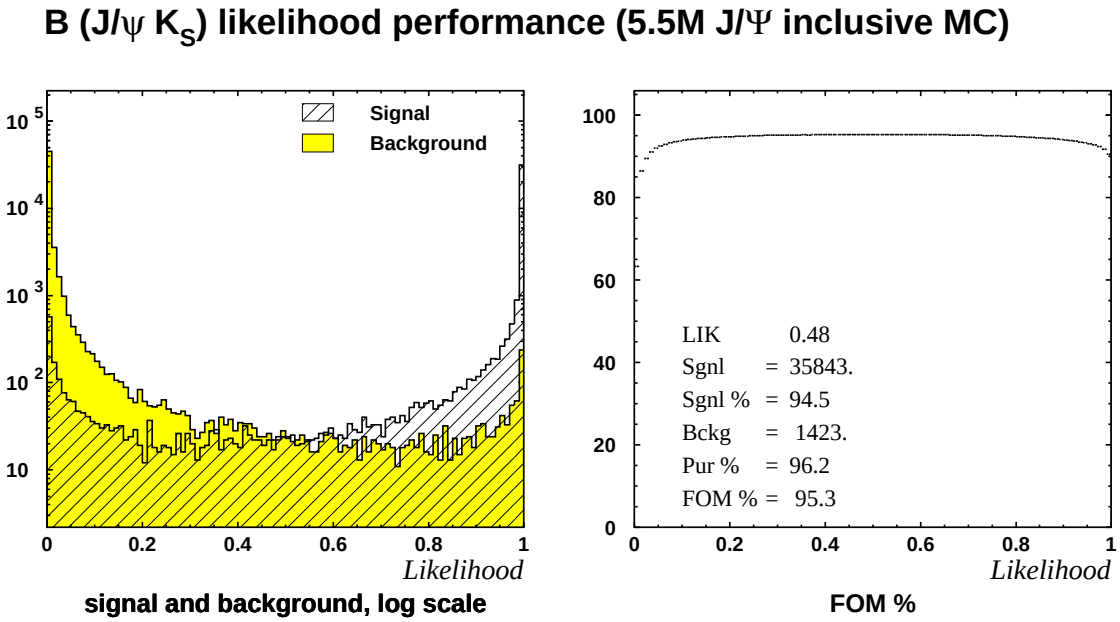

Figure E.6: The $B^0 \to J/\psi\ K_S(\pi^+\pi^-)$ likelihood performance. The plots are based on a 5.5-M-event J/ψ inclusive MC sample. For a description of the plot parameters see the discussion of Fig. 4.10.

We will start from charged two pion final state. This is rather clean channel, with $B^0 \to J/\psi\ K_s(\pi^+\pi^-)$ likelihood based on the following two dimensional discriminants:

- $M(B) - M(J/\psi) + 3.096$ GeV/c^2 vs. B CM momentum.
- K likelihood vs. B decay angle.
- B vs. J/ψ decay angle.
- $\Upsilon(4S)$ vs. J/ψ helicity.

The performance of the likelihood based on a 5.5-M-event inclusive J/ψ MC sample is summarized in Fig. E.6. The overall separation power is excellent with 94% of the signal coming in the signal region.

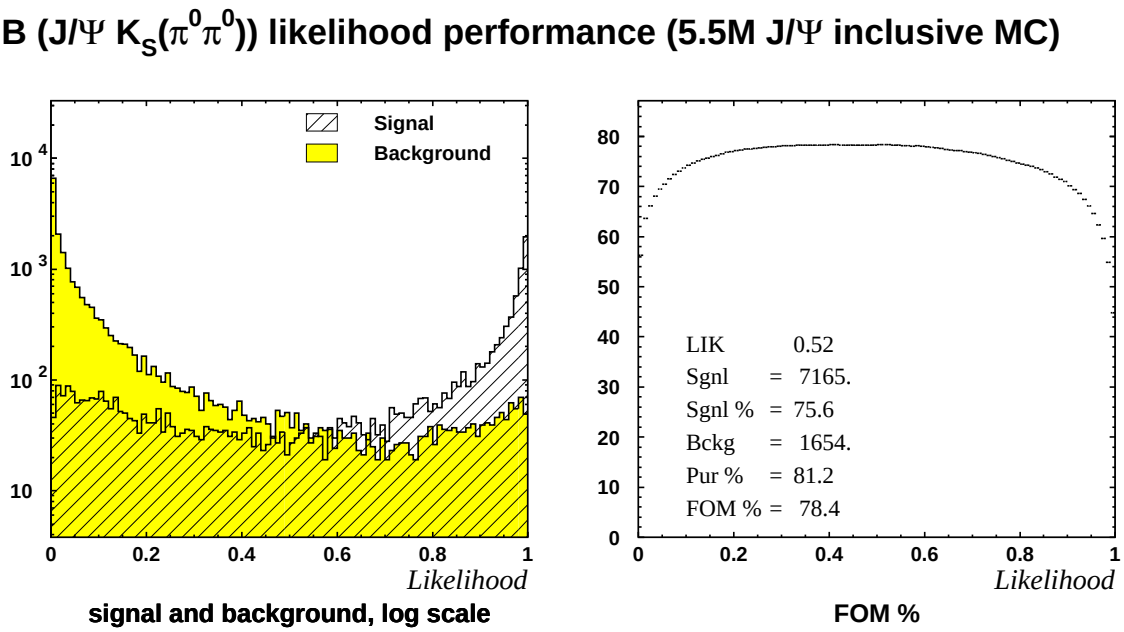

Figure E.7: The $B^0 \to J/\psi\ K_S(\pi^0\pi^0)$ likelihood performance. The plots are based on a 5.5-M-event J/ψ inclusive MC sample. For a description of the plot parameters see the discussion of Fig. 4.10.

$B^0 \to J/\psi\ K_S(\pi^0\pi^0)$ likelihood is based on the following two dimensional discriminants:

- $M(B) - M(J/\psi) + 3.096$ GeV/c^2 vs. B CM momentum.
- J/ψ vs. K CM momentum.
- K likelihood vs. B decay angle.
- B vs. J/ψ decay angle.
- $\Upsilon(4S)$ vs. J/ψ helicity.

The performance of the likelihood based on a 5.5-M-event inclusive J/ψ MC sample is summarized in Fig. E.7. The overall separation power is satisfactory with 75% of the signal coming in the signal region.

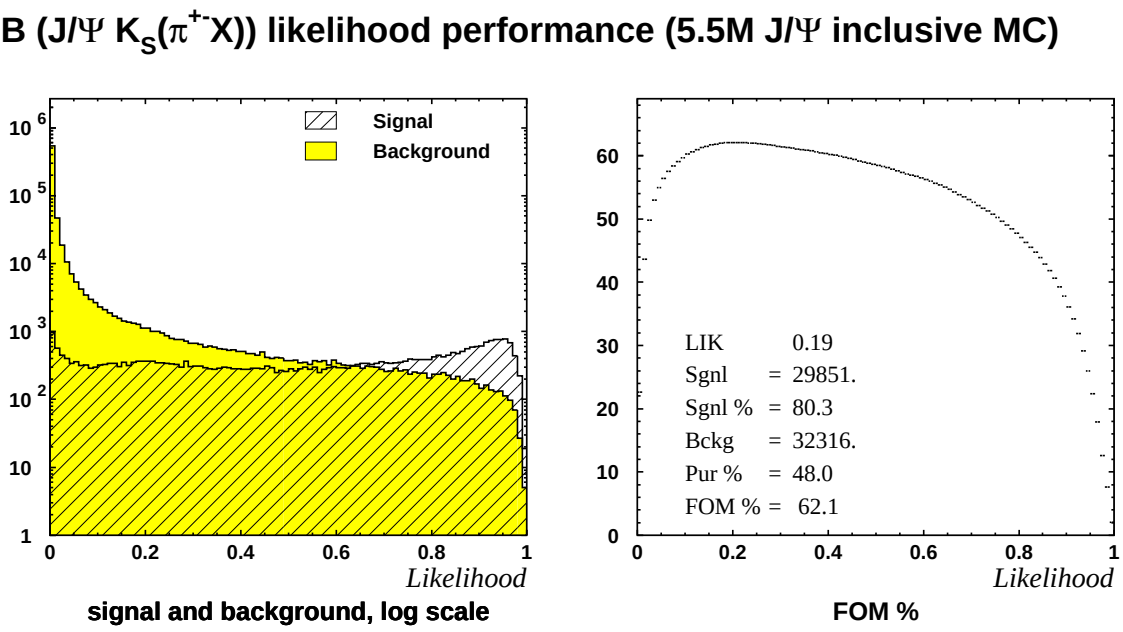

Figure E.8: The likelihood performance for partially reconstructed $B^0 \to J/\psi\ K_S(X\pi^\pm)$ decay. The plots are based on a 5.5-M-event J/ψ inclusive MC sample. For a description of the plot parameters see the discussion of Fig. 4.10.

Partially reconstructed decays $B^0 \to J/\psi\ K_S(X\pi^\pm)$ have likelihood based on the following two dimensional discriminants:

- $M(B) - M(J/\psi) + 3.096$ GeV/c^2 vs. B CM momentum.
- J/ψ vs. π CM momentum.
- π likelihood vs. B decay angle.
- B vs. J/ψ decay angle.
- $\Upsilon(4S)$ vs. J/ψ helicity.
- IP probability of the pion track vs. pion momentum in the lab frame.

The performance of the likelihood based on a 5.5-M-event inclusive J/ψ MC sample is summarized in Fig. E.8. The overall separation power is good with 80% of the signal coming in the signal region.

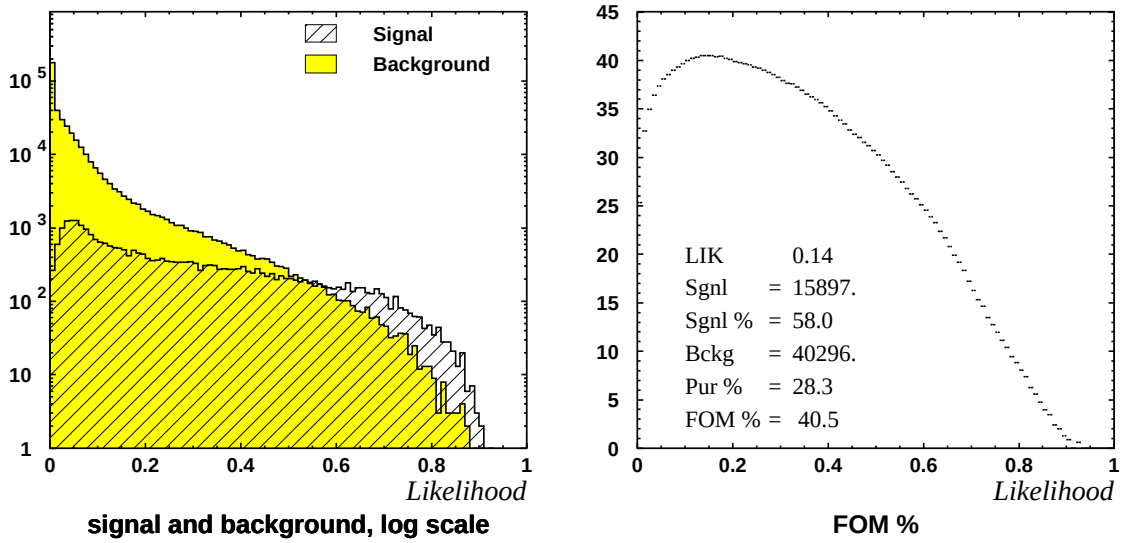

Figure E.9: The likelihood performance for partially reconstructed $B^0 \to J/\psi\ K_S(X\pi^0)$ decay. The plots are based on a 5.5-M-event J/ψ inclusive MC sample. For a description of the plot parameters see the discussion of Fig. 4.10.

Partially reconstructed decays $B^0 \to J/\psi\ K_s(X\pi^0)$ have likelihood based on the following two dimensional discriminants:

- $M(B) - M(J/\psi) + 3.096$ GeV/c^2 vs. B CM momentum (for $B \to J/\psi\ \pi$ decay).
- J/ψ vs. π CM momentum.
- π likelihood vs. B decay angle.
- B vs. J/ψ decay angle.
- $\Upsilon(4S)$ vs. J/ψ helicity.

The performance of the likelihood based on a 5.5-M-event inclusive J/ψ MC sample is summarized in Fig. E.9. The overall separation power is satisfactory with 58% of the signal coming in the signal region.

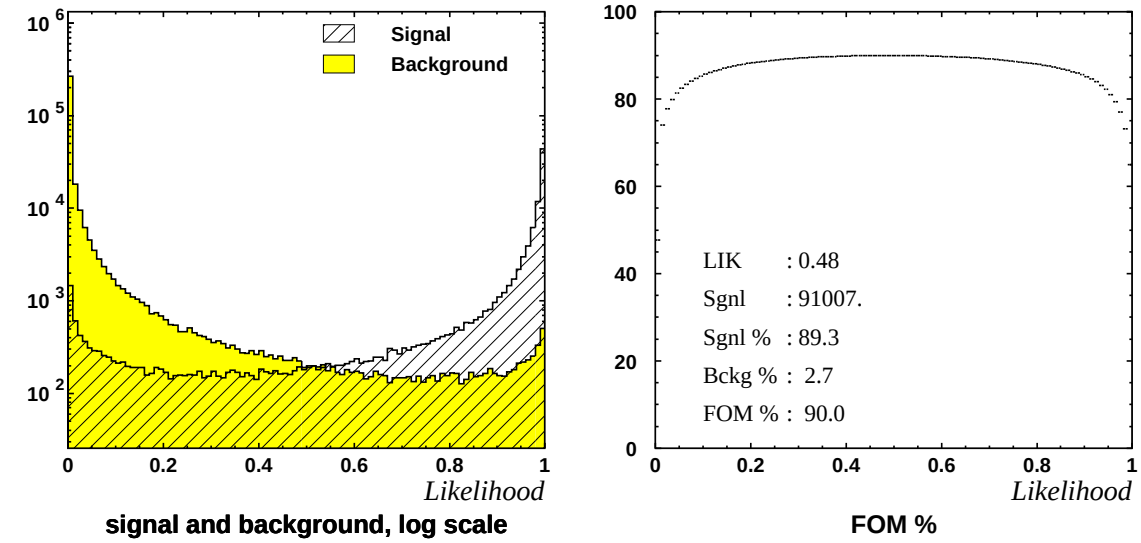

Figure E.10: The $B^0 \to J/\psi\ K^*(K^+\pi^-)$ likelihood performance. The plots are based on a 5.5-M-event J/ψ inclusive MC sample. For a description of the plot parameters see the discussion of Fig. 4.10.

We will start with selection criteria for neutral K^* decaying into charged kaon and charged pion final state. This is rather clean channel, with $B^0 \to J/\psi\ K^0(K^\pm\pi^\mp)$ likelihood based on the following two dimensional discriminants:

- $M(B) - M(J/\psi) + 3.096$ GeV/c^2 vs. B CM momentum.
- K^* likelihood vs. B decay angle.
- B vs. J/ψ decay angle.
- $\Upsilon(4S)$ vs. J/ψ helicity.
- π vs. sum of K and π CM momentum.

The performance of the likelihood based on a 5.5-M-event inclusive J/ψ MC sample is summarized in Fig. E.10.

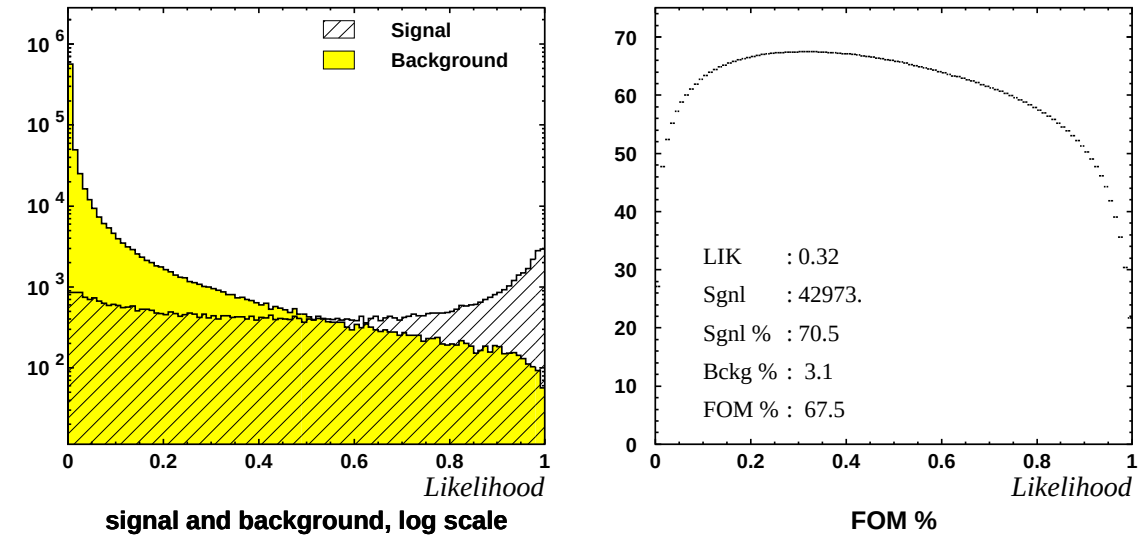

Figure E.11: The $B^{\pm} \to J/\psi\ K^{*}(K^{\pm}\pi^{0})$ likelihood performance. The plots are based on a 0.5 M-event J/ψ inclusive MC sample. For a description of the plot parameters see the discussion of Fig. 4.10.

The $B^{\pm} \to J/\psi\ K^{\pm}(K^{\pm}\pi^{0})$ likelihood is based on the following two dimensional discriminants:

- $M(B) - M(J/\psi) + 3.096$ GeV/c^2 vs. B CM momentum.
- K^* likelihood vs. B decay angle.
- B vs. J/ψ decay angle.
- $\Upsilon(4S)$ vs. J/ψ helicity.
- π vs. sum of K and π CM momentum.
- $M(K^*)$ vs. K^* CM momentum.

The performance of the likelihood based on the 0.5-M-event inclusive J/ψ MC sample is summarized in Fig. E.11.

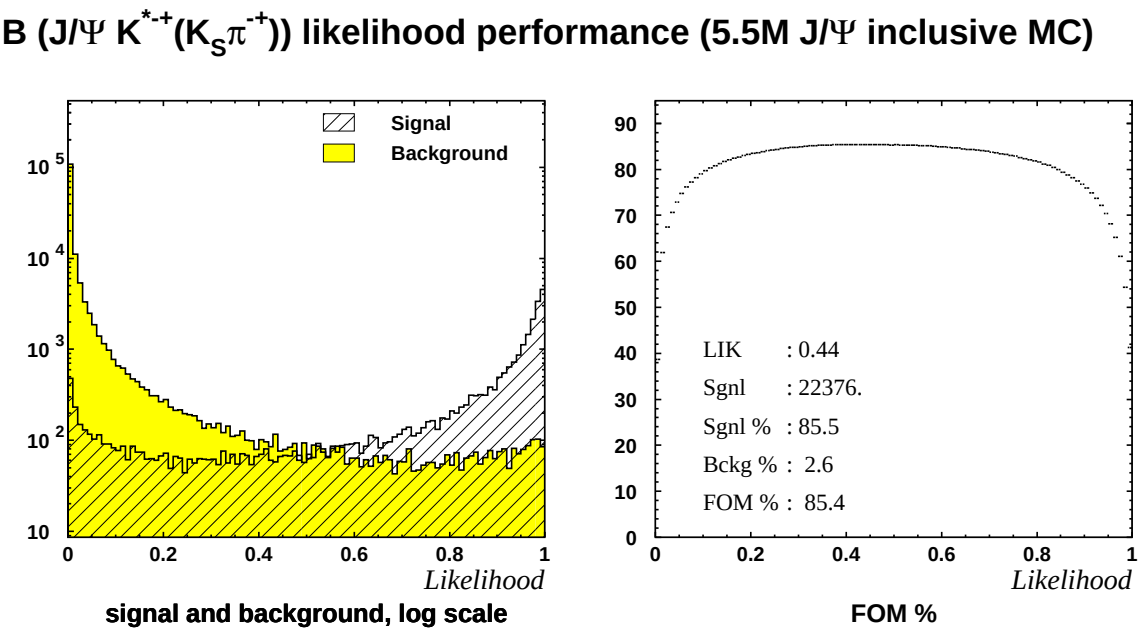

Figure E.12: The $B^{\pm} \to J/\psi\ K^{*}(K_S\pi^{\pm})$ likelihood performance. The plots are based on a 5.5-M-event J/ψ inclusive MC sample. For a description of the plot parameters see the discussion of Fig. 4.10.

Decays of charged K^* involving K_S and charged pions enjoy excellent particle identification for K_S. The $B^{\pm} \to J/\psi\ K^{\pm}(K_S\pi^{\pm})$ likelihood is based on the following two dimensional discriminants:

- $M(B) - M(J/\psi) + 3.096$ GeV/c^2 vs. B CM momentum.
- K^* likelihood vs. B decay angle.
- B vs. J/ψ decay angle.
- $\Upsilon(4S)$ vs. J/ψ helicity.
- π vs. sum of K and π CM momentum.

The performance of the likelihood based on a 5.5-M-event inclusive J/ψ MC sample is summarized in Fig. E.12.

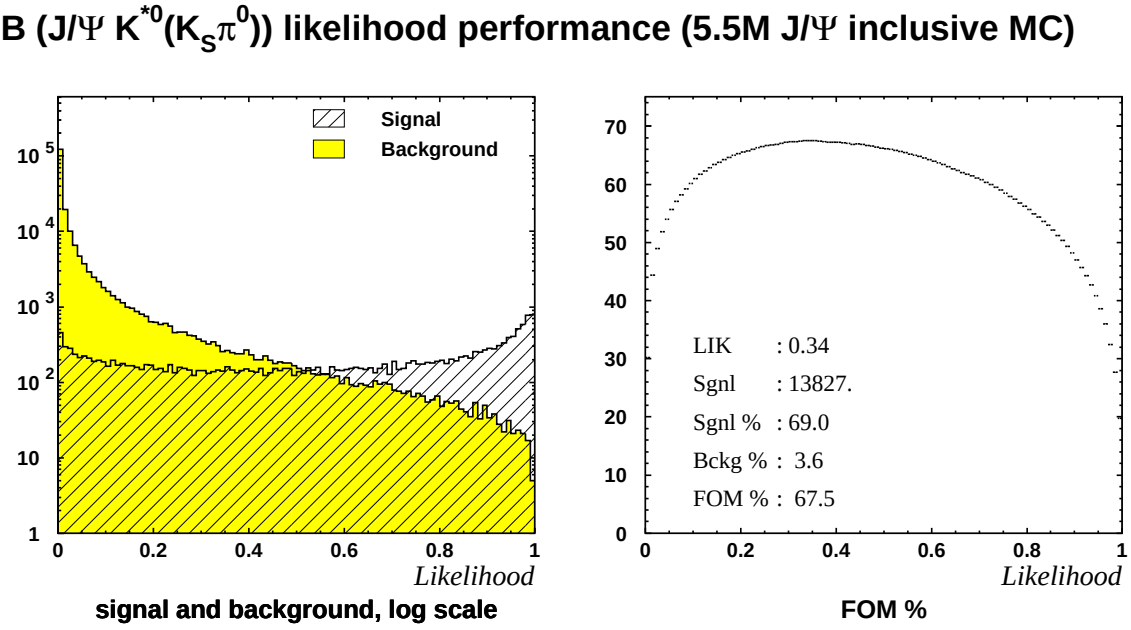

Figure E.13: The $B^0 \to J/\psi\ K^{*0}(K_S\ \pi^0)$ likelihood performance. The plots are based on a 5.5-M-event J/ψ inclusive MC sample. For a description of the plot parameters see the discussion of Fig. 4.10.

The $B^0 \to J/\psi\ K^{*0}(K_S\pi^0)$ likelihood is based on the following two dimensional discriminants:

- $M(B) - M(J/\psi) + 3.096$ GeV/c^2 vs. B CM momentum.
- K^* likelihood vs. B decay angle.
- B vs. J/ψ decay angle.
- $\Upsilon(4S)$ vs. J/ψ helicity.
- π vs. sum of K and π CM momentum.
- $M(K^*)$ vs. K^* CM momentum.

The performance of the likelihood based on a 5.5-M-event inclusive J/ψ MC sample is summarized in Fig. E.13.

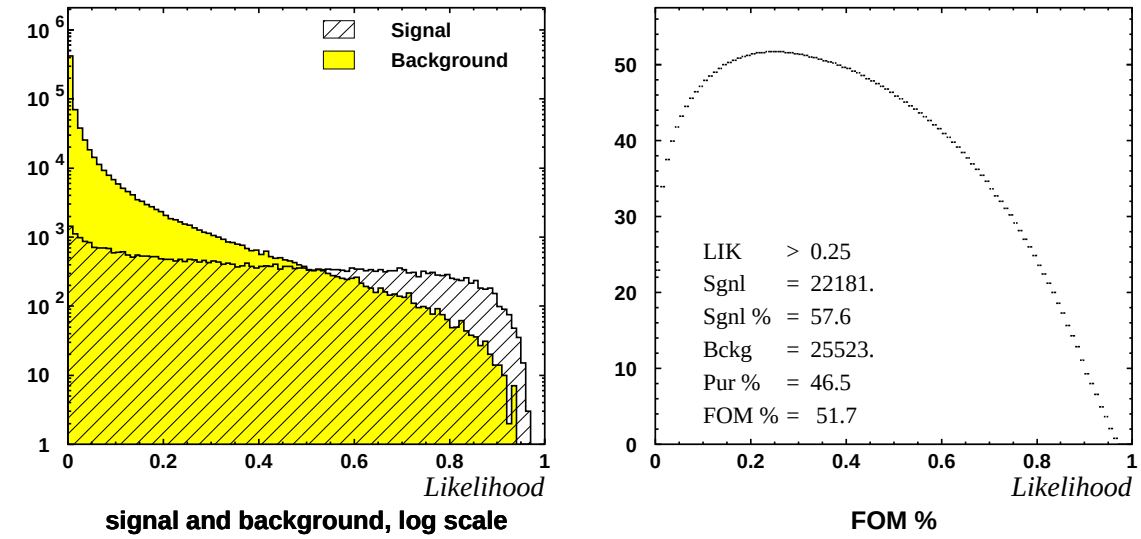

Figure E.14: The $B^{\pm} \to J/\psi\ K^*(K_L\pi^{\pm})$ likelihood performance. The plots are based on a 5.5-M-event J/ψ inclusive MC sample. For a description of the plot parameters see the discussion of Fig. 4.10.

The $B^{\pm} \to J/\psi\ K^{\pm}(K_L\pi^{\pm})$ likelihood is based on the following two dimensional discriminants:

- $\Upsilon(4S)$ vs. J/ψ helicity.
- $L(K) \otimes\ L(\pi)$ vs. K^* decay angle.
- J/ψ vs. B CM momentum.
- $M(K^*)$ vs. $L^P(J/\psi)$.

The performance of the likelihood based on a 5.5-M-event inclusive J/ψ MC sample is summarized in Fig. E.14.

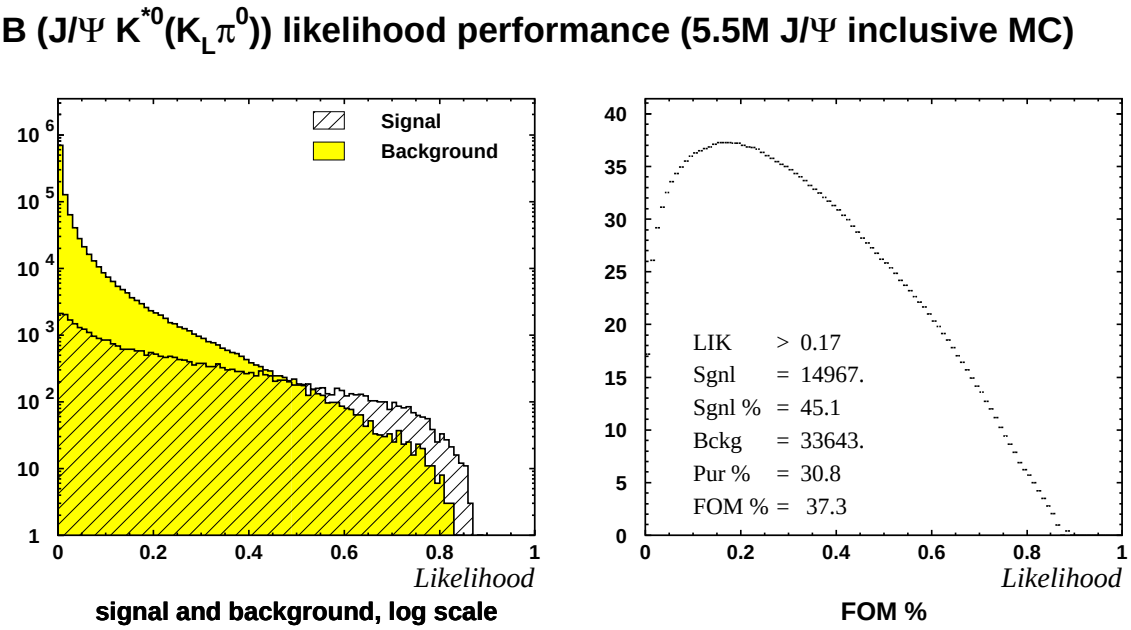

Figure E.15: The $B^0 \to J/\psi\ K^*(K_L\pi^0)$ likelihood performance. The plots are based on a 5.5-M-event J/ψ inclusive MC sample. For a description of the plot parameters see the discussion of Fig. 4.10.

The $B^0 \to J/\psi\ K^0(K_L\pi^0)$ likelihood is based on the following two dimensional discriminants:

- $\Upsilon(4S)$ vs. J/ψ helicity.
- $L(K) \otimes\ L(\pi)$ vs. K^* decay angle.
- J/ψ vs. B CM momentum.
- $M(K^*)$ vs. $L^P(J/\psi)$.

The performance of the likelihood based on a 5.5-M-event inclusive J/ψ MC sample is summarized in Fig. E.15.

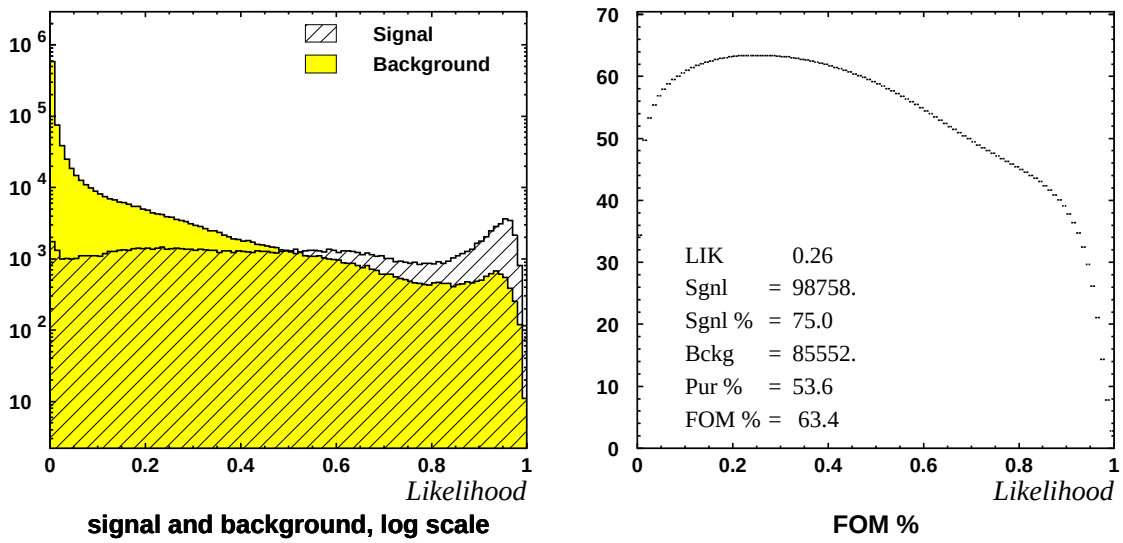

Figure E.16: The likelihood performance for partially reconstructed $B \to J/\psi\ K^*(K^\pm X)$ decay. The plots are based on a 5.5-M-event J/ψ inclusive MC sample. For a description of the plot parameters see the discussion of Fig. 4.10.

The $B \to J/\psi\ K^*(K^\pm\ X)$ likelihood is based on the following two dimensional discriminants:

- $M(B) - M(J/\psi) + 3.096$ GeV/c^2 vs. B CM momentum.
- J/ψ vs. K CM momentum.
- K likelihood vs. B decay angle.
- B vs. J/ψ decay angle.
- $\Upsilon(4S)$ vs. J/ψ helicity.

The performance of the likelihood based on a 5.5-M-event inclusive J/ψ MC sample is summarized in Fig. E.16.

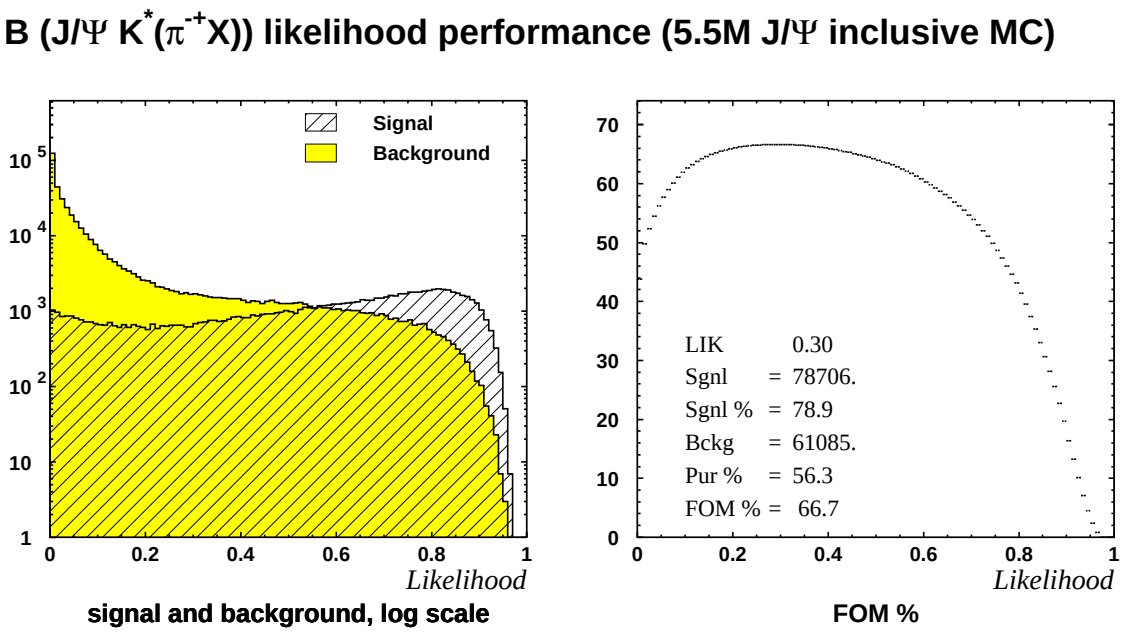

Figure E.17: The likelihood performance for partially reconstructed $B \to J/\psi\ K^*(X\pi^\pm)$ decay. The plots are based on a 5.5-M-event J/ψ inclusive MC sample. For a description of the plot parameters see the discussion of Fig. 4.10.

The $B \to J/\psi\ K^*(\pi^\pm\ X)$ likelihood is based on the following two dimensional discriminants:

- $M(B) - M(J/\psi) + 3.096$ GeV/c^2 vs. B CM momentum.
- J/ψ vs. π CM momentum.
- π likelihood vs. B decay angle.
- B helicity vs. J/ψ decay angle.
- $\Upsilon(4S)$ helicity vs. J/ψ helicity.
- IP probability of the pion track vs. pion momentum in the lab frame.

The performance of the likelihood based on a 5.5-M-event inclusive J/ψ MC sample is summarized in Fig. E.17.

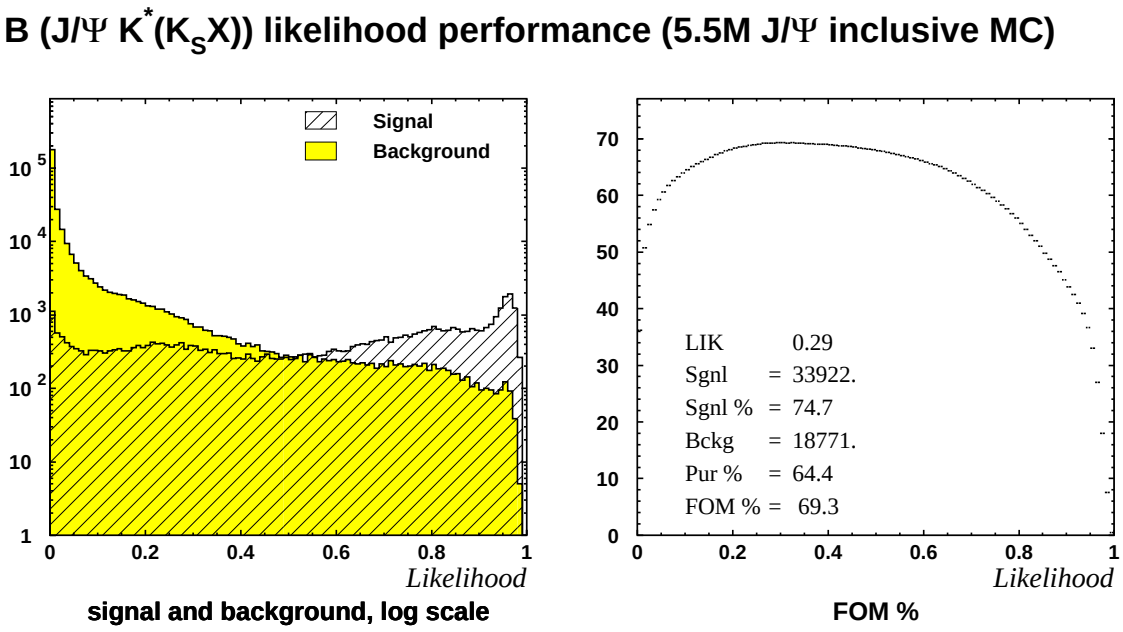

Figure E.18: The likelihood performance for partially reconstructed $B \to J/\psi\ K^*(K_S X)$ decay. The plots are based on a 5.5-M-event J/ψ inclusive MC sample. For a description of the plot parameters see the discussion of Fig. 4.10.

The $B \to J/\psi\ K^*(K_S\ X)$ likelihood is based on the following two dimensional discriminants:

- $M(B) - M(J/\psi) + 3.096$ GeV/c^2 vs. B CM momentum.
- J/ψ vs. K CM momentum.
- K likelihood vs. B decay angle.
- B helicity vs. J/ψ decay angle.
- $\Upsilon(4S)$ helicity vs. J/ψ helicity.

The performance of the likelihood based on a 5.5-M-event inclusive J/ψ MC sample is summarized in Fig. E.18.

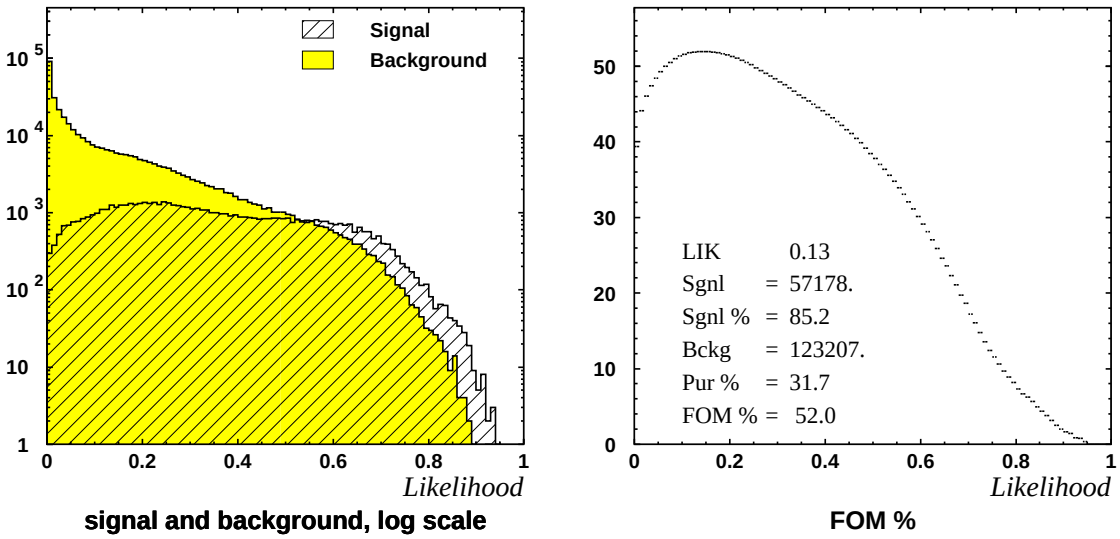

Figure E.19: The likelihood performance for partially reconstructed $B \to J/\psi\ K^*(X\pi^0)$ decay. The plots are based on a 5.5-M-event J/ψ inclusive MC sample. For a description of the plot parameters see the discussion of Fig. 4.10.

The $B \to J/\psi\ K^*(\pi^0\ X)$ likelihood is based on the following two dimensional discriminants:

- $M(B) - M(J/\psi) + 3.096$ GeV/c^2 vs. B CM momentum.
- J/ψ vs. π CM momentum.
- π likelihood vs. B decay angle.
- B helicity vs. J/ψ decay angle.
- $\Upsilon(4S)$ helicity vs. J/ψ helicity.

The performance of the likelihood based on a 5.5-M-event inclusive J/ψ MC sample is summarized in Fig. E.19.

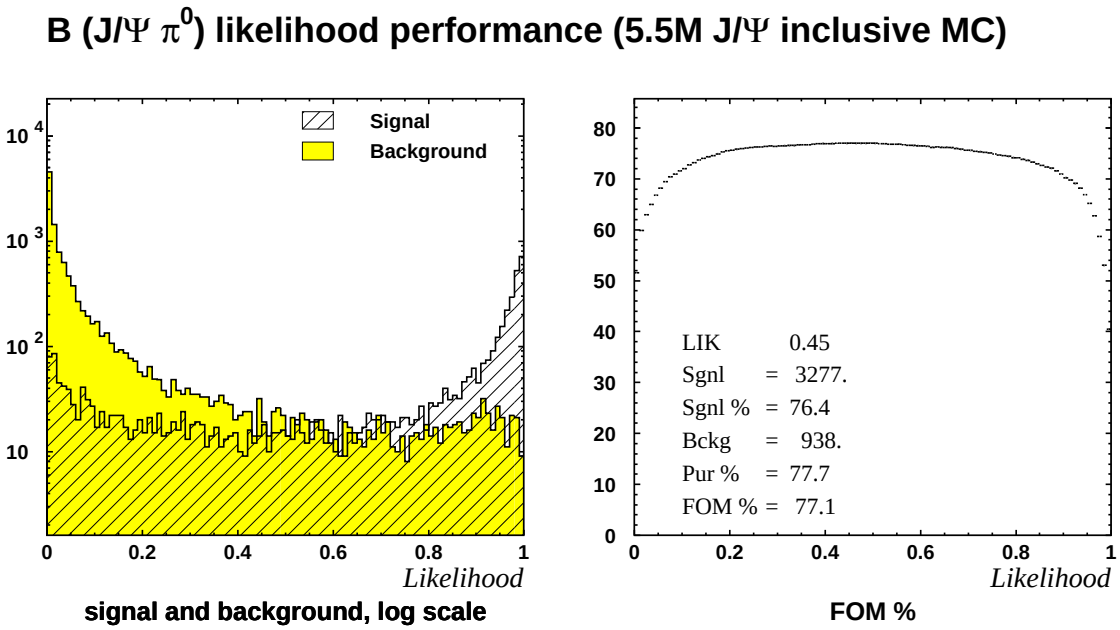

Figure E.20: The likelihood performance for the $B^0 \to J/\psi\ \pi^0$ decay. The plots are based on a 5.5-M-event J/ψ inclusive MC sample. For a description of the plot parameters see the discussion of Fig. 4.10.

The $B^{\pm} \to J/\psi\ \pi^0$ likelihood is based on the following two dimensional discriminants:

- $M(B) - M(J/\psi) + 3.096$ GeV/c^2 vs. B CM momentum.
- J/ψ vs. π CM momentum.
- π likelihood vs. B decay angle.
- B vs. J/ψ decay angle.
- $\Upsilon(4S)$ vs. J/ψ helicity.

The performance of the likelihood based on the 0.5-M-event inclusive J/ψ MC sample is summarized in Fig. E.20. The overall separation power is good with 76% of the signal coming in the signal region.

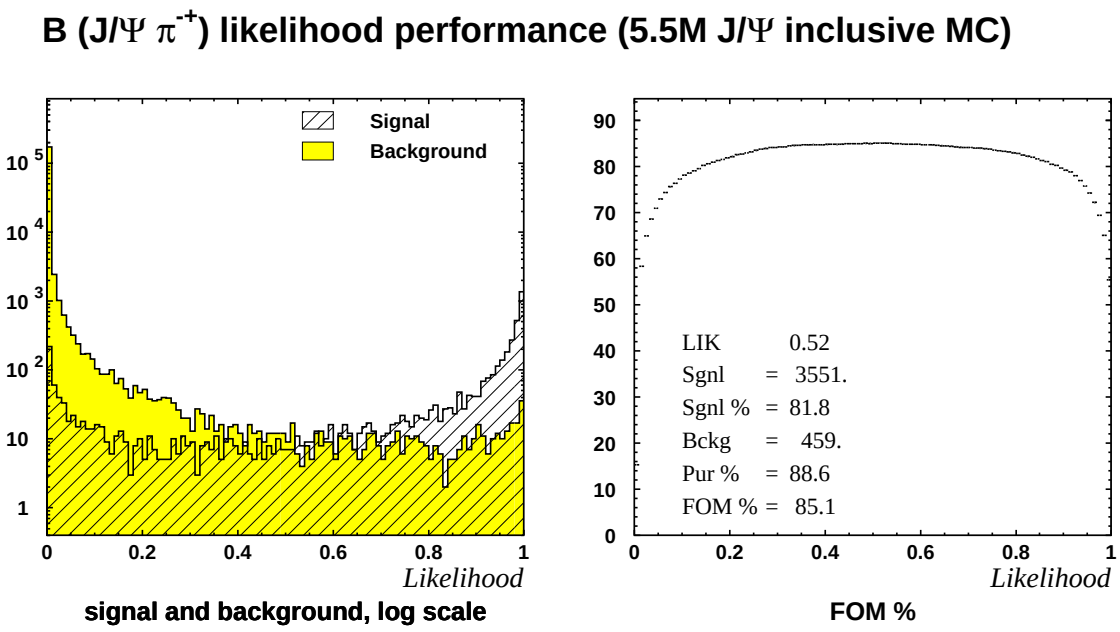

Figure E.21: The likelihood performance for the $B^{\pm} \to J/\psi\ \pi^{\pm}$ decay. The plots are based on a 5.5-M-event J/ψ inclusive MC sample. For a description of the plot parameters see the discussion of Fig. 4.10.

The $B^{\pm} \to J/\psi\ \pi^{\pm}$ likelihood is based on the following two dimensional discriminants:

- $M(B) - M(J/\psi) + 3.096$ GeV/c^2 vs. B CM momentum.
- J/ψ vs. π CM momentum.
- π likelihood vs. B decay angle.
- B vs. J/ψ decay angle.
- $\Upsilon(4S)$ vs. J/ψ helicity.

The performance of the likelihood based on a 5.5-M-event inclusive J/ψ MC sample is summarized in Fig. E.21. The overall separation power is good with 81% of the signal coming in the signal region.

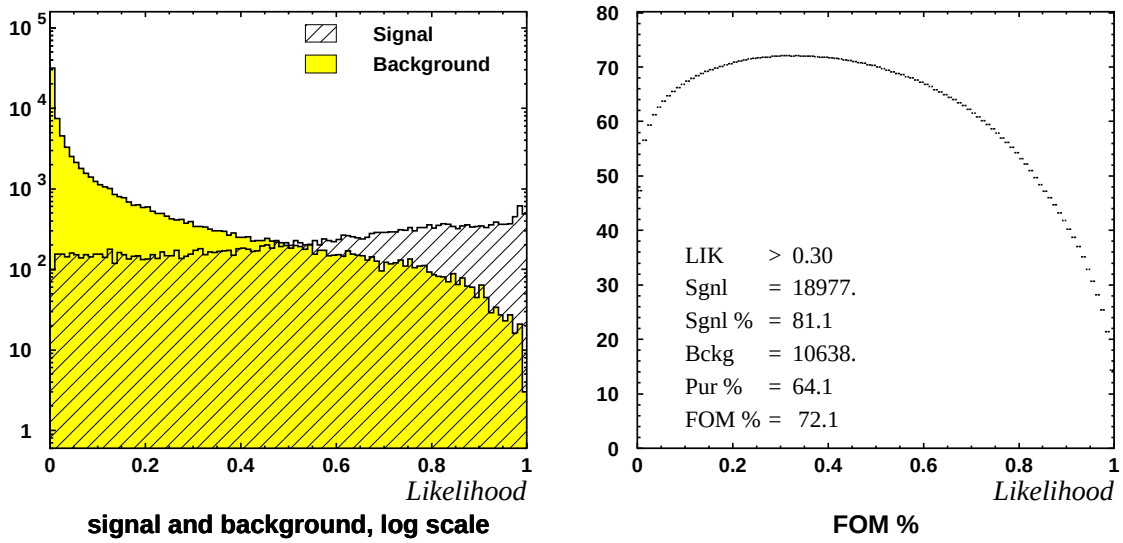

Figure E.22: The likelihood performance for the $B^0 \to J/\psi\ K_L$(KLM) decay. The plots are based on a 5.5-M-event J/ψ inclusive MC sample. For a description of the plot parameters see the discussion of Fig. 4.10.

The $B^0 \to J/\psi\ K_L$(KLM) likelihood based on the following two dimensional discriminants:

- $\Upsilon(4S)$ vs. J/ψ helicity.
- K_L likelihood vs. B decay angle.
- $L^P(K_L)$ vs. $L^P(J/\psi)$.
- J/ψ vs. B CM momentum.
- B vs. J/ψ decay angle.
- $(p_{K_L} - p_{MM})/p_{K_L}$ vs. $(p^T_{K_L} - p^T_{MM})/p^T_{K_L}$[1].

The performance of the likelihood based on a 5.5-M-event inclusive J/ψ MC sample is summarized in Fig. E.22.

[1] P^T means transverse momentum and $P_M M$ denotes projection of the missing momentum on the K_L flight direction

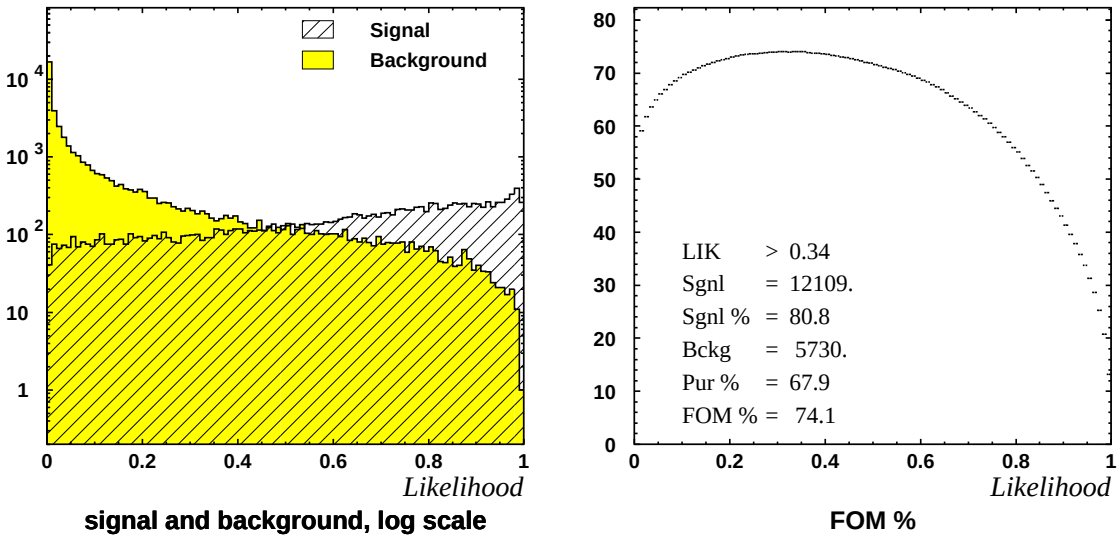

Figure E.23: The likelihood performance for the $B^0 \to J/\psi\ K_L$(ECL-KLM) decay. The plots are based on a 5.5-M-event J/ψ inclusive MC sample. For a description of the plot parameters see the discussion of Fig. 4.10.

The $B^0 \to J/\psi\ K_L$(ECL-KLM) likelihood based on the following two dimensional discriminants:

- $\Upsilon(4S)$ vs. J/ψ helicity.
- K_L likelihood vs. B decay angle.
- $L^P(K_L)$ vs. $L^P(J/\psi)$.
- J/ψ vs. B CM momentum.
- B vs. J/ψ decay angle.
- $(p_{K_L} - p_{MM})/p_{K_L}$ vs. $(p^T_{K_L} - p^T_{MM})/p^T_{K_L}$.

The performance of the likelihood based on a 5.5-M-event inclusive J/ψ MC sample is summarized in Fig. E.23.

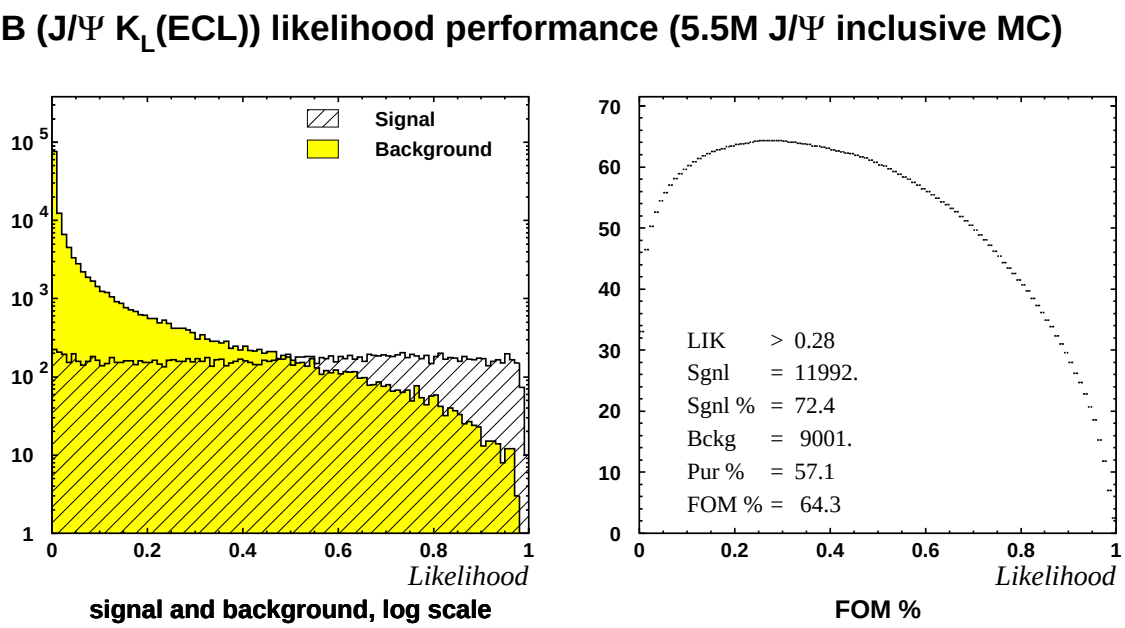

Figure E.24: The likelihood performance for the $B^0 \to J/\psi\ K_L$(ECL) decay. The plots are based on a 5.5-M-event J/ψ inclusive MC sample. For a description of the plot parameters see the discussion of Fig. 4.10.

The $B^0 \to J/\psi\ K_L$(ECL) likelihood based on the following two dimensional discriminants:

- $\Upsilon(4S)$ vs. J/ψ helicity.
- K_L likelihood vs. B decay angle.
- $L^P(K_L)$ vs. $L^P(J/\psi)$.
- J/ψ vs. B CM momentum.
- B vs. J/ψ decay angle.
- $(p_{K_L} - p_{MM})/p_{K_L}$ vs. $(p^T_{K_L} - p^T_{MM})/p^T_{K_L}$.

The performance of the likelihood based on a 5.5-M-event inclusive J/ψ MC sample is summarized in Fig. E.24.

Bibliography

[1] D. N. Spergel *et al.*, "First Year Wilkinson Microwave Anisotropy Probe (WMAP) Observation: Determination of Cosmological Parameters," Astrophys. J. Suppl **148**, 175 (2003).

[2] N. Cabibbo, "Unitary Symmetry and Leptonic Decays," Phys. Rev. Lett. **10**, 531 (1963).

[3] M. Kobayashi and T. Maskawa, "CP violation in the renormalizable theory of weak interaction," Prog. Theor. Phys. **49**, 652 (1973).

[4] B. Kayser, "CP violation and quantum mechanics in the B system," proceedings of 30th Rencontres de Moriond: Electroweak Interactions and Unified Theories (1995).

[5] B. Kayser, "CP violation, mixing, and Quantum mechanics," proceedings of 28th International Conference on High-energy Physics (ICHEP-96) (1996).

[6] B. Kayser, "CP violation in the K and B systems," Trieste HEP Cosmology (1996).

[7] L.Wolfenstein, Phys. Rev. Lett. **51** (1945).

[8] I. I. Y. Bigi and A. I. Sanda, "CP violation," Cambridge Monogr. Part. Phys. Nucl. Phys. Cosmol. (2000).

[9] P. A. M. Dirac, "Quantum theory of emission and absorption of radiation," Proc. Roy. Soc. Lond. **A114**, 243 (1927).

[10] P. A. M. Dirac, "The Quantum theory of electron," Proc. Roy. Soc. Lond. **A117**, 610 (1928).

[11] P. A. M. Dirac, "Quantised singularities in the electromagnetic field," Proc. Roy. Soc. Lond. **A133**, 60 (1931).

[12] M. Gronau, O. F. Hernandez, D. London, and J. L. Rosner, "Electroweak penguins and two body B decays," Phys. Rev. **D52**, 6374 (1995).

[13] D. E. Groom *et al.*, "Review of particle physics," Eur. Phys. J. **C15**, 1 (2000).

[14] "Flavor-changing neutral currents: present and future studies," (1997).

[15] M. Gronau, J. L. Rosner, and D. London, "Weak coupling phase from decays of charged B mesons to πK and $\pi\pi$," Phys. Rev. Lett. **73**, 21 (1994).

[16] M. Gronau, O. F. Hernandez, D. London, and J. L. Rosner, "Decays of B mesons to two light pseudoscalars," Phys. Rev. **D50**, 4529 (1994).

[17] M. Gronau, "Large penguin effects in the CP asymmetry of $B^0_d \to \pi^+\pi^-$," Phys. Lett. **B300**, 163 (1993).

[18] S. L. Glashow, J. Iliopoulos, and L. Maiani, "Weak Interactions with Lepton - Hadron Symmetry," Phys. Rev. **D2**, 1285 (1970).

[19] D. Besson *et al.*, "Observation of new structure in the e^+e^- annihilation cross-section above B anti-B threshold," Phys. Rev. Lett. **54**, 381 (1985).

[20] "KEKB B-factory Design Report," Tech. rep., National Laboratory for High Energy Physics (June 1995).

[21] "Belle Technical Design Report," Tech. rep., National Laboratory for High Energy Physics (June 1995).

[22] K. Hanagaki, H. Kakuno, H. Ikeda, T. Iijima, and T. Tsukamoto, "Electron identification in Belle," Nucl. Instrum. Meth. **A485**, 490 (2002).

[23] A. Abashian *et al.*, "Muon identification in the Belle experiment at KEKB," Nucl. Instrum. Meth. **A491**, 69 (2002).

[24] M. Hazumi *et al.*, "Performance of the Belle silicon vertex detector," Nucl. Instr. and Meth. **A 473**, 1 (2001).

[25] H. Aihara *et al.*, "The Belle silicon vertex detector," Nucl. Instrum. Meth. **A466**, 268 (2001).

[26] S. Uno *et al.*, "The Belle central drift chamber," Nucl. Instrum. Meth. **A379**, 421 (1996).

[27] H. Hirano *et al.*, "A high resolution cylindrical drift chamber for the KEKB-factory," Nucl. Instrum. Meth. **A455**, 294 (2000).

[28] H. Bethe, "Theory of the passage of fast corpuscular rays through matter," Annalen Phys. **5**, 325 (1930).

[29] F. Bloch, "Stopping power of atoms with several electrons," Z. fur Physik **81**, 363 (1933).

[30] E. Fermi, "On Cerenkov radiation," Phys. Rev. **57**, 485 (1940).

[31] U. Fano, "Penetration of protons, alpha particles and mesons," Ann. Rev. Nucl. Sci. **13**, 1 (1963).

[32] H. Ikeda *et al.*, "A detailed test of the CsI(Tl) calorimeter for Belle with photon beams of energy between 20 MeV and 5.4 GeV," Nucl. Instrum. Meth. **A441**, 401 (2000).

[33] H. Ikeda, "Development of the CsI(Tl) Calorimeter for the Measurement of CP violation at KEK B-factory," Ph.D. thesis (1999).

[34] A. Abashian *et al.*, "The Belle Detector," Nucl. Instr. and Meth. **A 479**, 117 (2002).

[35] H. Sagawa *et al.*, "The Belle CsI calorimeter," Nucl. Instrum. Meth. **A453**, 259 (2000).

[36] T. Iijima *et al.*, "Aerogel Cherenkov counter for the Belle detector," Nucl. Instrum. Meth. **A453**, 321 (2000).

[37] e. Nappi, E. and e. Ypsilantis, T., "Experimental techniques of Cerenkov light imaging. Proceedings, 1st Workshop on Ring Imaging Cerenkov Detectors, Bari, Italy, June 2-5, 1993," Nucl. Instrum. Meth **A343**, 1 (1994).

[38] H. Kichimi *et al.*, "The Belle TOF system," Nucl. Instrum. Meth. **A453**, 315 (2000).

[39] A. Abashian *et al.*, "The KLM detector subsystem for the Belle experiment at the KEK B-factory," Nucl. Instrum. Meth. **A449**, 112 (2000).

[40] R. Wixted, S. Chidzik, K.Korotushenko, D.R.Marlow, and C. Mindas, "RPC Readout Version 2.0. Reference Manual." Belle Note 141 (1996).

[41] M. Yamaga *et al.*, "RPC systems for Belle detector at KEKB," Nucl. Instrum. Meth. **A456**, 109 (2000).

[42] Yu.N.Pestov, "Status and future development of sparc counters with a localized discharge," Nucl. Instr. and Meth. **196**, 45 (1982).

[43] M.Anelli, G.Bencivenni, G.Felici, and L.Magro, "Glass electrode sparc counters," Nucl. Instr. and Meth. **A300**, 572 (1991).

[44] G.Bencivenni *et al.*, "Glass electrode sparc counters and their use in low energy e.m. calorimeters." Nucl. Instr. and Meth. **A315**, 507 (1992).

[45] G.Bencivenni *et al.*, "A glass sparc counter for high rate environment." Nucl. Instr. and Meth. **A332**, 368 (1993).

[46] M. Yamaga, "Measurement of CP violation in $B^0 \to J/\psi + K_L$ decay at KEK B-factory," Ph.D. thesis, Tohoku University (2001).

[47] M. Yokoyama, "Observation of Large CP Violation in the Neutral B Meson System Using $B^0 \to J/\psi K_L$ Decay," Ph.D. thesis, University of Tokyo (2002).

[48] S. Vahsen, "Measurement of CP Violation in $J/\psi K_L$ Decays," Ph.D. thesis, Princton University (2003).

[49] A. Abashian *et al.*, "Measurement of the CP violation parameter $\sin(2\phi_1)$ in B^0_d meson decays," Phys. Rev. Lett. **86**, 2509 (2001).

[50] K. Abe *et al.*, "Observation of mixing-induced CP violation in the neutral B meson system," Phys. Rev. **D66**, 032007 (2002).

[51] K. Abe *et al.*, "An improved measurement of mixing-induced CP violation in the neutral B meson system," Phys. Rev. **D66**, 071102 (2002).

[52] D. J. Lange, "The EvtGen particle decay simulation package," Nucl. Instrum. Meth. **A462**, 152 (2001).

[53] CERN, "CERN Program Library Long Writup W5013," (1993).

[54] R. Fisher, "The Use of Multiple Measurements in Taxonomic Problems," Annals of Eugenics **7**, 179 (1936).

[55] R. Fisher, "The Statistical Utilization of Multiple Measurements," Annals of Eugenics **8**, 376 (1938).

[56] E. Parzen, "On Estimation of a Probability Density Function and Mode," Annals of Mathematical Statistics **33**, 1065 (September 1962).

[57] B. Casey *et al.*, "Measurement of the Number of BBbar Events in Experiment 5 Data and the BBbar Cross Section at KEKB," Belle Note 296 (2001).

[58] B. Casey *et al.*, "HadronB," Belle Note 390 (2001).

[59] S. Olsen, "Event shape cuts for inclusive charmonium skimming," Belle Note 304 (2001).

[60] G. C. Fox and S. Wolfram, "Event shapes in e^+e^- annihilation," Nucl. Phys. **B149**, 413 (1979).

[61] H. Taureg *et al.*, "Study of the photon spectrum in the decay $K^0_S \to \pi^+\pi^-\gamma$," Phys. Lett. **B65**, 92 (1976).

[62] T. Armstrong *et al.*, "Observation of the radiative decay $J/\psi \to e^+e^-\gamma$," Phys. Rev. **D54**, 7067 (1996).

[63] A. V. Ershov, "Beauty meson decays to charmonium," Ph.D. thesis, Harvard University (2001).

[64] R. Nernst *et al.*, "Observation of three p states in the radiative decay of $\Upsilon(2S)$," Phys. Rev. Lett. **54**, 2195 (1985).

[65] K. Hagiwara *et al.*, "Review of Particle Physics," Physical Review D **66** (2002).

[66] H. Hamasaki *et al.*, "KID identification in Belle," Belle Note 321 (2001).

[67] H. Kakuno *et al.*, "Flavor tagging by Multidimensional likelihood method." Belle Note 384 .

[68] K. Abe *et al.*, "Precise measurement of B meson lifetimes with hadronic decay final states," Phys. Rev. Lett. **88**, 171801 (2002).

[69] B. Aubert *et al.*, "Measurement of the CP Asymmetry Amplitude $\sin 2\beta$ wit B^0 Mesons," Phys. Rev. Lett. **89**, 201802 (2002).

[70] T. D. Lee, R. Oehme, and C.-N. Yang, "Remarks on Possible Noninvariance under Time Reversal and Charge Conjugation," Phys. Rev. **106**, 340 (1957).

[71] G. Luders and B. Zumino, "Some consequences of TCP-Invariance," Phys. Rev. **106**, 385 (1957).

[72] W. Rarita and J. S. Schwinger, "On a theory of particles with half integral spin," Phys. Rev. **60**, 61 (1941).

[73] C. Zemach, "Three-Pion Decays of Unstable Particles," Phys. Rev. **133**, B1201 (1964).

[74] C. Zemach, "Use of Angular-Momentum Tensors," Phys. Rev. **140**, B97 (1964).

[75] V. Filippini, "Spin parity methods in meson spectroscopy," In *Folgaria 1992, Proceedings, Common problems and trends of modern physics* 469-475.

[76] V. Filippini, A. Fontana, and A. Rotondi, "Covariant spin tensors in meson spectroscopy," Phys. Rev. **D51**, 2247 (1995).

[77] V. Filippini, "Covariant tensor in spin-parity analysis," Nucl. Phys. **A655**, 99 (1999).

[78] M. Jacob and G. C. Wick, "On the general theory of collisions for particles with spin," Ann. Phys. **7**, 404 (1959).

[79] J. D. Hansen, G. T. Jones, G. Otter, and G. Rudolph, "Formalism and Assumptions Involved in Partial Wave Analysis of three - Meson Systems," Nucl. Phys. **B81**, 403 (1974).

[80] C. Amsler and J. C. Bizot, "Simulation of Angular Distributions and Correlations in the Decay of Particles with Spin," Comput. Phys. Commun. **30**, 21 (1983).

[81] S. U. Chung, "Spin Formalisms," Lectures given in Academic Training Program of CERN, 1969- 1970.

[82] S. U. Chung, "Helicity coupling amplitudes in tensor formalism," Phys. Rev. **D48**, 1225 (1993).

[83] S. U. Chung, "A General formulation of covariant helicity coupling amplitudes," Phys. Rev. **D57**, 431 (1998).

[84] H. Albrecht *et al.*, "A Measurement of the electronic widths Γ_{ee} of the $\Upsilon(1s)$, $\Upsilon(2s)$, and $\Upsilon(4s)$ resonances, and of the total decay width Γ of the $\Upsilon(4s)$," Z. Phys. **C65**, 619 (1995).

[85] T.Matsumoto *et al.*, "Measurement of $\sqrt{s}$ from full reconstructed B decays at $\Upsilon(4S)$ resonance." Belle Note 418 (201).