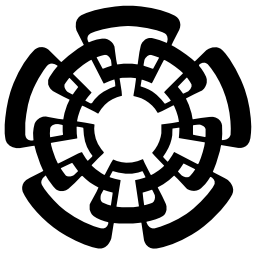


Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional

DEPARTAMENTO DE FÍSICA

CP Violation in $B_s^0 \rightarrow J/\psi\,\phi$ using 8 fb^{-1} of $p\bar{p}$ collisions

Tesis que presenta

Ricardo Magaña Villalba

para obtener el grado de

Doctor en Ciencias

en la especialidad de

Física

Director de tesis: **Dr. Heriberto Castilla Valdez**

México, Distrito Federal. Abril, 2012

Esta tesis está dedicada a

Laura y Miranda.

AGRADECIMIENTOS

Quiero agradecer a mi asesor Heriberto Castilla quien nos guió y apoyó durante estos años. A Eduard de La Cruz y Alberto Sanchez que estuvieron siempre disponibles para resolver dudas y discutir ideas.

También quiero agradecer a mis compañeros de Cinvestav en DØ –Enrique Camacho, Alejandro Garcia, Ivan Heredia, Jorge Martinez y Jesus Orduna–, quienes fueron mi familia lejos de casa.

Al sub grupo dedicado al jpsi phi –Avdhesh Chandra, Marj Corcoran, Aurelio Juste, Penny Kasper, Dmitri Tsybychev, Rick Van Kooten, Mark Williams y Daria Zieminska–, por las innumerables discusiones y el incansable esfuerzo que llevaron a buen termino este trabajo. A los coordinadores de física –Marco Verzocchi y Gregorio Bernardi–, cuyas múltiples revisiones y sugerencias mantienen el alto estándar de calidad en las publicaciones.

Al grupo de operaciones y del CTT, en especial a Marc Buehler, George Ginther, Stefan Grüenendahl, Norik Khalatyan, Bill Lee, Carrie McGivern, Geoff Savage y Jadwiga Warchol por sus enseñanzas en esta otra cara del trabajo experimental: el diseño, operación y mantenimiento continuo del detector.

A mis padres que siempre me han dado su su cariño y apoyo incondicional, A mi esposa Laura quien compartió conmigo la aventura en Fermilab, Y a Areli Robles por su apoyo en el diseño editorial de este trabajo.

Finalmente, agradezco al Cinvestav y el CONACYT cuya colaboración genera el ambiente necesario para producir ciencia.

RESUMEN

La teoría del Big-Bang, la teoría mas aceptada que describe el origen y la evolución del universo, predice un equilibrio entre materia y antimateria. Pero el universo esta dominado por materia. Un ingrediente necesario para explicar esta asimetría entre materia y antimateria es la violación de CP. La violación de CP esta incluida en el Modelo Estándar y describe correctamente la violación de CP presente en los mesones B^+ y B_d^0, pero es muy pequeña para describir la asimetría entre materia y antimateria.

El Modelo Estándar predice un valor pequeño de violación de CP en el mesón B_s^0. La medición de un valor grande de violación de CP en este mesón seria una clara indicación de nueva física mas allá del Modelo Estándar. Esta tesis describe la medición de de la fase de violación de CP, $\phi_s^{J/\psi\phi}$, y la diferencia en el ancho de decaimiento de los eigenestados de masa, $\Delta\Gamma_s$, en el decaimiento $B_s^0 \to J/\psi\phi$.

Los datos usados corresponden a una luminosidad integrada de 8.0 fb^{-1} acumulados por el detector DØ de colisiones protón-antiprotón a una energía en el centro de masa de $\sqrt{s} = 1.96$ TeV producidas por el colisionador Tevatron en Fermilab.

Las regiones de 68% de confianza, incluyendo las incertidumbres sistemáticas, son $\Delta\Gamma_s = 0.163^{+0.065}_{-0.064}$ ps^{-1} y $\phi_s^{J/\psi\phi} = -0.55^{+0.38}_{-0.36}$. El valor p para la predicción del Modelo Estándar es 29.8%.

ABSTRACT

The Big Bang theory, the most accepted theory that describes the origin and evolution of the universe predicts an equilibrium between matter and antimatter, but our universe is dominated by matter. An ingredient required to explain this matter-antimatter asymmetry is CP violation. CP violation is included into the Standard Model, and successfully describes CP violation in the B^+ and B^0_d systems, but is far to small to to describe the matter-antimatter asymmetry.

The Standard Model predicts a small value of CP Violation in the B^0_s system. A measurement of large, anomalous CP violation in the this system would be a clear indication of new physics beyond the Standard Model. This dissertation describes the measurement of the CP-violating phase, $\phi_s^{J/\psi\phi}$, and the decay-width difference for the two mass eigenstates, $\Delta\Gamma_s$, from the flavor-tagged decay $B^0_s \to J/\psi\phi$.

The data sample corresponds to an integrated luminosity of 8.0 fb^{-1} accumulated with the DØ detector using proton-antiproton collisions at $\sqrt{s} = 1.96$ TeV produced at the Fermilab Tevatron Collider.

The 68% Bayesian confidence regions, including systematic uncertainties, are $\Delta\Gamma_s = 0.163^{+0.065}_{-0.064}$ ps^{-1} and $\phi_s^{J/\psi\phi} = -0.55^{+0.38}_{-0.36}$. The p-value for the Standard Model point is 29.8%.

CONTENTS

1 Theory **1**

1.1 The Standard Model 2

1.1.1 Building Blocks of Matter 2

1.1.2 Fundamental Interactions 4

1.2 CP Violation in the B System 9

1.2.1 The SM Description of CP Violation 9

1.2.2 The Unitary Triangle(s) of the CKM Matrix 10

1.2.3 B–$\bar{B}$ Mixing 11

1.2.4 The B_s System 12

1.3 Phenomenology of the $B_s^0 \to J/\psi\phi$ Decay 13

1.3.1 Differential Rates 14

1.3.2 Detector Efficiency and Normalization 18

1.3.3 Time Development 20

1.3.4 Sensitivity to Δm 22

1.3.5 Incorporating Direct CP Violation 23

1.3.6 Incorporating a Contribution from $B_s^0 \to J/\psi K^+K^-$ (Kaons in an S-Wave State) 26

1.3.7 Symmetries 30

2 Accelerator Machine **35**

2.1 Pre-accelerator 37

2.2 Linac 37

2.3 Booster 38
2.4 Main Injector 39
2.5 Tevatron 40
2.6 Antiproton Source 42
2.7 Recycler 44
2.8 Switchyard 120 GeV 45

3 DØ Detector **47**
3.1 Central tracking 49
3.1.1 Silicon microstrip tracker 50
3.1.2 Central fiber tracker 51
3.2 Muon system 52

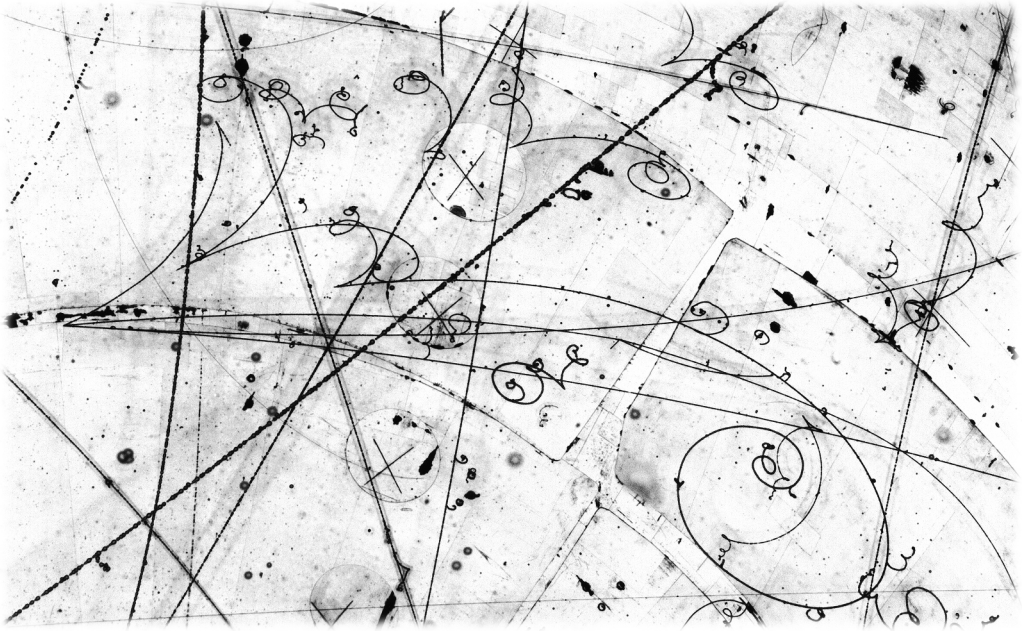
4 Measurement of CP Violation in $B_s \to J/\psi\phi$ **55**
4.1 Data Sample and Event Reconstruction 56
4.2 Background Suppression 57
4.2.1 Signal and background simulation 57
4.2.2 Multivariate event selection 58
4.2.3 BDT Discriminants 59
4.2.4 Selection Criteria 64
4.2.5 Simple Selection 66
4.3 Detector acceptance 67
4.4 Flavor Tagging 67
4.4.1 OST calibration 70
4.5 Maximum Likelihood Fit 71
4.5.1 Signal model 71
4.5.2 Background model 73
4.5.3 Fit results 74
4.5.4 Systematic uncertainties 79
4.5.5 Independent estimate of F_S 81
4.6 Bayesian credibility intervals from MCMC studies 82

4.6.1 The method 83
4.6.2 General properties of MCMC chains for the BDT-selection and Square-cuts samples 84
4.6.3 Results 86
4.7 Summary and Discussion 88

CHAPTER

ONE

THEORY



1.1 The Standard Model

The standard model (SM) of elementary particle physics [34, 75, 62, 36, 37, 35, 68, 43, 42, 59] is a quantum field theory that combines quantum mechanics and the concepts of special relativity. It provides a very elegant theoretical framework to describe the fundamental particles and their interactions. Since its formulation in the 1960s and 1970s it has undergone a large number of experimental tests and passed all of them successfully. Although there remain some open questions that cannot be answered by the SM, it is nevertheless the theory most successfully tested with highest precision up to now.

1.1.1 Building Blocks of Matter

Elementary particles of spin $s = 1/2$, called fermions, are the building blocks of matter in the SM. They obey Fermi-Dirac statistics and the Pauli exclusion principle, which implies that two fermions cannot share the same quantum state.

The SM contains twelve of these elementary fermions, six leptons and six quarks, which can be arranged in three generations, as shown in figure 1.1. Each generation consists of an up-type quark (up (u), charm (c), top (t)) and a down-type quark (down (d), strange (s), bottom (b)), a charged lepton (electron (e), muon (μ), tau (τ)) and the corresponding neutrino (electron-neutrino (ν_e), muon-neutrino (ν_μ), tau-neutrino (ν_τ)). To each elementary fermion exist in addition a corresponding antipartner, which has the same properties as the fermion but opposite charges.

All ordinary matter of our universe is composed of particles of the first generation. The higher-generation fermions appear solely in high-energy interactions as for example in collisions of cosmic rays with molecules of the atmosphere or in the laboratory environment of collider experiments. Once produced, these massive fermions decay subsequently into the lighter fermions of the first generation.

Charged leptons (e, μ, τ) carry electric charge of one elementary charge, while the corresponding neutrinos are electrically neutral. Neutrinos are assumed to be strictly massless in the SM. Various observations [18, 16, 19, 17] however indicate

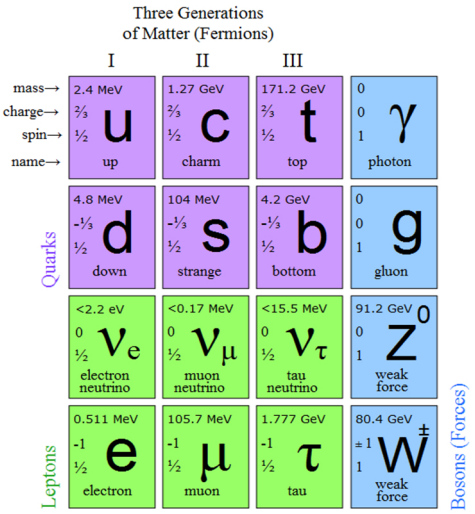


Figure 1.1: The three generation of fermions and their electric charges and masses. The electric charge is given in units of the elementary charge $e = 1.602176487(40) \times 10^{-19}$C.

that neutrinos have non-vanishing masses. This requires an extension [51, 60] of the SM accommodating these results.

The members of the second group of fermions, the quarks, carry fractional electric charge of either 2/3 (up-type quarks) or $-1/3$ (down-type quarks) of elementary charge. Quarks cannot exist as free particles, they are forced to form bound states, called hadrons, of either three quarks (*baryons*) or quark-antiquark pairs (*mesons*). This feature of the interaction between quarks is named *quark confinement*. The top quarks due to its large mass decays before it can form bound states. In order to satisfy the Pauli exclusion principle in the formation of bound states of quarks it becomes necessary to introduce an additional quantum number for quarks. This quantum number is called color [39, 44, 32] and can be three types (red, blue, and green). Color is a non-observable quantum number, thus only colorless quark combinations are realized in nature. This implies that quarks do not occur as free particles and is the theoretical explanation for quark confinement. Colorless mesons consist of a quark of a certain color and of an antiquark carrying the corresponding anticolor. Combinations of three differently colored quarks result in colorless baryons.

	Effect	Strength	Range (m)	Partcile
Strong	Binds quarks and gluons	1	10^{-15}	gluon
Electromagnetic	Iteraction between electric charged particles	1/137	Infinite	photon
Weak	Radioactive β-decay	10^{-6}	10^{-18}	W^+, W^-, Z_0
Gravitation	Attraciton of masses	6×10^{-39}	Infinite	graviton?

Table 1.1: Fundamental forces of nature.

1.1.2 Fundamental Interactions

In our present understanding of the world, all known interactions between particles can be ascribed to four fundamental forces, and gravitation, see table 1.1. Three of them can be formulated as quantum field theories and have therefore been successfully incorporated in the SM, while gravity is described by the theory of general relativity [24, 25]. The fact that gravity and the quantum-field framework of the SM seem to be mathematically incompatible is an indication that the SM in its contemporary form cannot be the ultimate theory of everything. Nevertheless, due to the extreme separation of scales on which gravity and the three forces of the SM are relevant [1], we can make use of both theories to explain the physics in our experimental reach.

The dynamics of a physical system are formulated in functions called *Lagrangians*. In gauge theories the Lagrangians are invariant under a certain group of local transformations. In quantized field theories the quanta of the gauge fields represent particles transmitting the forces. These *gauge bosons* are particles with integer spin that obey Bose-Einstein statistics. This allows several bosons to oc-

[1] Gravity plays an important role on large scales, e.g. for the foundation of galaxies, but can be neglected on subnuclear scales. The electromagnetic, weak and strong force are the dominating forces in elementary particle reactions but can be neglected on large scales.

cupy the same quantum state as opposed to the half-integer spin fermions (which follow Fermi-Dirac statistics). The gauge bosons of the three SM forces have all spin $s = 1$. A compilation of these fore mediators and some of their properties can be found in figure 1.1.

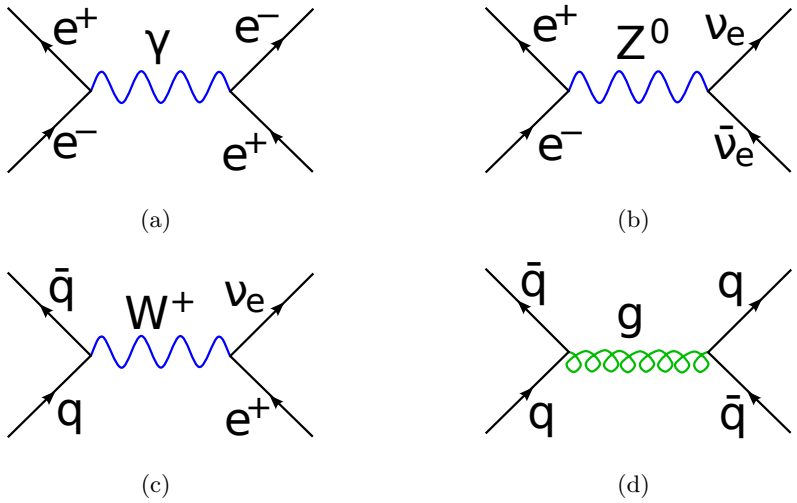


Figure 1.2: Example Feynman diagrams for typical fundamental interactions. In each diagram two incoming particles annihilate and produce a virtual gauge boson that decays into the two outgoing particles: Electron-positron annihilation via (a) the electromagnetic force into a photon or via (b) the weak force into a Z boson, (c) quark-antiquark annihilation into a W boson, and (d) quark-antiquark annihilation via the strong interaction into a gluon. In this description times evolves from left to right, while the spatial dimensions expand the vertical direction.

Figure 1.2 shows how one can imagine such interactions between elementary particles via gauge bosons exchange. So called Feynman diagrams are visualizations of the mathematical expressions describing fundamental interactions. The propagation in space-time is here represented by lines, whereas the couplings are represented by vertices. Applying the Feynman rules one can translate each diagram into a formula and calculate the transition amplitude of a given process by summing over all possible Feynman graphs contributing to this process. From this, the cross section of the process, a measure for the probability for an interaction to occur, can be computed.

Interactions between electrically charged particles are described by the theory of Quantum Electrodynamics (QED) [27, 29, 28, 63, 64, 65, 73]. The gauge boson mediating the electromagnetic force is the photon, which is the excitation of the massless photon field. The infinite range of the electromagnetic force is a direct consequence of the zero-mass of the photon.

The strong force between color charged particles is described by a similar

theory, the theory of Quantum Chromodynamics (QCD) [43, 42, 59, 32]. Massless gluons couplings to the color charge are the gauge bosons of the strong interaction. Unlike the electrically neutral photons, gluons themselves carry the charge they couple to, and gluons therefore do not only mediate the strong interaction but also participate in it. The theory requires gluons to carry one unit of color and one unit of anticolor, or more precisely superpositions of these. Due to the eight possible independent combinations of color and anticolor, which lead to a non-zero net color charge, there exist eight different gluons. The decrease in the strength of the strong interaction at short distances, named asymptotic freedom, and its increase at large distances that forces quarks to create colorless bound states are direct consequences of the color charge of gluons.

The weak force differs in several aspects from the two other forces in the SM. In contrast to the massless photon and gluons, the mediators of the weak force are massive gauge bosons, the electrically neutral Z^0 and the charged $W^{\pm}$ bosons. The large mass of the weak gauge bosons limits the range of the weak force to subnuclear scales. Weak gauge bosons couple to the weak charge, which is called isospin. Thus the weak force is the only force in the SM that affects neutrinos, which possess isospin, but carry neither electric nor color charge.

Another uniqueness of the weak interaction is the capability of flavor change through the exchange of charged W bosons. The eigenstates of the weak interaction are not identical with the mass eigenstates. With the Cabibbo-Kobayashi-Maskawa (CKM) matrix [7, 52] one can transform the different eigenstates into each other. By convention the CKM matrix is a 3×3 unitary matrix that operates on the mass eigenstate of the downtype quarks (d, s, b) resulting in the corresponding weak eigenstates (d', s',b'):

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ d \end{pmatrix} \tag{1.1}$$

For example the weak partner of the top quark is b', a linear combination of the mass eigenstates of the three downtype quarks d, s, and b. In general, the

coupling of two quarks q_1 and q_2 to a W boson is proportional to the corresponding matrix element $V_{q_1 q_2}$. The diagonal elements have by far the largest values, so that transitions of quarks within a family are favored over transitions into quarks of other families. For the given example this implies that the top quarks can in principle decay in any of the three downtype quarks, but the decay into the b quark is the most probable one.

The law of parity conservation (P) says that the physics remain the same under the reversal of all spatial axes. While the electromagnetic and the strong force respect this principle, it is violated in weak interactions mediated by charged W and Z bosons. The parity violation of the weak force is closely linked to the concept of chirality of particles. Something is chiral, if cannot be mapped to its mirror image by simple rotations and translations alone. An example of chiral objects are human hands: the left hand is a non-superposable mirror image of the right hand and vice versa. The two possible chiral states are named left-handed and right-handed after this every day's life example. Particles can be linear combinations of left-handed and right-handed components. In the SM the W boson couples only to the left-handed part of the wave function of particles and to the right-handed part of the wave function of antiparticles. Since the parity transformed object of a left-handed particle is a right-handed particle this behavior of the W bosons means a maximal violation of parity conservation. In contrast, the neutral mediator of the weak interactions, the Z boson, couples to both chiral components, but with different strengths depending on the particular quark or lepton involved.

Although at first sight electromagnetic effects and phenomena of the weak interaction seem very different, above a certain energy scale both can be explained within the electroweak theory [75, 62, 36, 35] as two aspects of one single interaction.

A consequence of local gauge invariance is that all particles described by the theory have to be massless. Introducing mass terms for particles into the Lagrangian would spoil the local gauge invariance. The gauge theory of electroweak interactions contains four massless gauge bosons. But at low energies the electro-

magnetic force with its massless photon and the weak force with its massive gauge bosons appear very different. Therefore the symmetry of the electroweak theory has to be broken, such that three of the four massless gauge bosons acquire mass and one remains massless.

The most promising candidate mechanism providing such a breakdown of the electroweak symmetry is the Higgs mechanism [48, 47, 49] in which the symmetry is spontaneously broken. Spontaneous symmetry breaking occurs, if the Lagrangian possesses symmetries which do not hold for the vacuum state of the sytem. In the Higgs mechanism a complex isodoublet scalar field, the Higgs field, with a non-vanishing vacuum expectation value ($v \approx 246$ GeV) is introduced into the Lagrangian, leading to symmetry breaking terms. According to Goldstone's theorem [56, 57, 38] for each broken generator of a symmetry group one would expect the existence of a massless Goldstone boson vanishes and replaces the missing longitudinal degree of freedom of the former massless gauge boson. In other words the Goldstone boson gets "eaten up" by the gauge boson, giving it its mass. In the Higgs mechanism the symmetry of the electroweak theory is broken in a way that the symmetry of the electromagnetic force remains as a symmetry of the vacuum. As a consequence the gauge bosons of the weak interaction become massive, while the photon remains massless. As a remnant of this mechanism, the theory predicts a massive scalar Higgs particle. This particle is the last particle predicted by the SM yet to be observed; its discovery would be a major success of the SM itself.

But not only the weak gauge bosons acquire their masses through the interaction with the Higgs field. The coupling of the lepton fields to the Higgs field, the so-called Yukawa coupling, leads to mass terms for the charged leptons it the Lagrangian. In a similar way mass terms for the quarks can be found by Yukawa coupling of the quark fields to the Higgs field. The Higgs mechanism can provide an explanation of the masses of the fermions, but since the Yukawa couplings are free parameters of the theory, the SM cannot predict the mass values of any of its fermions.

1.2 CP Violation in the B System

Among the most interesting aspects and unsolved mysteries of modern particle physics is the violation of CP symmetry. Studies of CP violation are particularly exciting, as they may open a window to the physics beyond the SM. There are many interesting ways to explore CP violation, for instance through certain rare K- or D-meson decays.

However, for testing the SM description of CP violation in a quantitative way, the B system appears to be most promising [58, 40, 31, 5, 45].

1.2.1 The SM Description of CP Violation

Within the framework of the SM, CP violation is closely related to the Cabibbo–Kobayashi–Maskawa (CKM) matrix [7, 52], connecting the electroweak eigenstates (d', s', b') of the down, strange and bottom quarks with their mass eigenstates (d, s, b) through the following unitary transformation:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix} \equiv \hat{V}_{\rm CKM} \cdot \begin{pmatrix} d \\ s \\ b \end{pmatrix} . \tag{1.2}$$

The elements of the CKM matrix describe charged-current couplings.

In the case of three generations, three generalized Cabibbo-type angles [7] and a single *complex phase* [52] are needed in order to parametrize the CKM matrix. This complex phase allows one to accommodate CP violation in the SM, as was pointed out by Kobayashi and Maskawa in 1973 [52]. A closer look shows that CP-violating observables are proportional to the following combination of CKM matrix elements [50]:

$$J_{\rm CP} = \pm \,{\rm Im}\left(V_{ik} V_{jl} V_{il}^* V_{jk}^*\right) \quad (i \neq j,\, l \neq k)\,, \tag{1.3}$$

which represents a measure of the "strength" of CP violation in the SM. Since $J_{\rm CP} = \mathcal{O}(10^{-5})$, CP violation is a small effect. However, in scenarios of new physics, typically several additional complex couplings are present, leading to new sources of CP violation.

As far as phenomenological applications are concerned, the following parametrization of the CKM matrix, the "Wolfenstein parametrization" [76], which corresponds to a phenomenological expansion in powers of the small quantity $\lambda \equiv |V_{us}| = \sin\theta_{\rm C} \approx 0.22$, turns out to be very useful:

$$\hat{V}_{\rm CKM} = \begin{pmatrix} 1-\frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \tag{1.4}$$

The terms of $\mathcal{O}(\lambda^4)$ can be taken into account systematically [6], and will play an important role.

1.2.2 The Unitary Triangle(s) of the CKM Matrix

Concerning tests of the CKM picture of CP violation, the central targets are the *unitary triangles* of the CKM matrix. The unitarity of the CKM matrix, which is described by

$$\hat{V}_{\rm CKM}^{\dagger} \cdot \hat{V}_{\rm CKM} = \hat{1} = \hat{V}_{\rm CKM} \cdot \hat{V}_{\rm CKM}^{\dagger}, \tag{1.5}$$

leads to a set of 12 equations, consisting of 6 normalization and 6 orthogonality relations. The latter can be represented as 6 triangles in the complex plane, which all have the same area [3]. However, in only two of them, all three sides are of comparable magnitude $\mathcal{O}(\lambda^3)$, while in the remaining ones, one side is suppressed relative to the others by $\mathcal{O}(\lambda^2)$ or $\mathcal{O}(\lambda^4)$. The orthogonality relations describing the non-squashed triangles are given as follows:

$$V_{ud}\,V_{ub}^* + V_{cd}\,V_{cb}^* + V_{td}\,V_{tb}^* = 0 \tag{1.6}$$

$$V_{ud}^*\,V_{td} + V_{us}^*\,V_{ts} + V_{ub}^*\,V_{tb} = 0. \tag{1.7}$$

The two non-squashed triangles agree at leading order in the Wolfenstein expansion, i.e. at $\mathcal{O}(\lambda^3)$, so that we actually have to deal with a single triangle at this order, which is usually referred to as "the" unitary triangle of the CKM

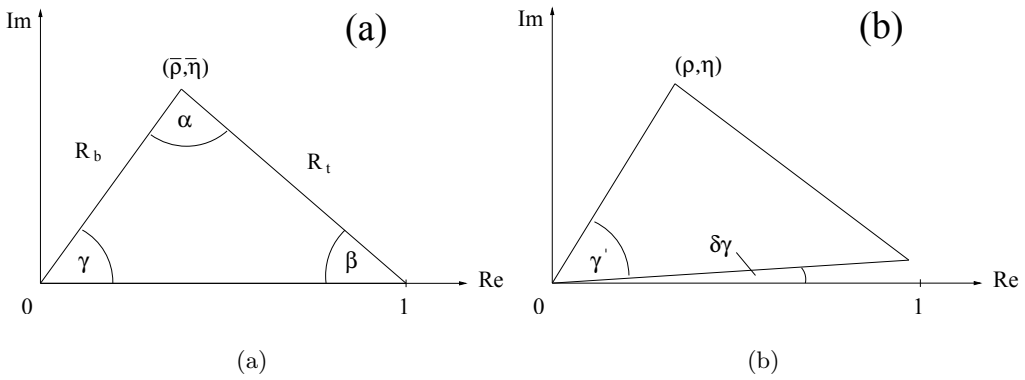


Figure 1.3: The two non-squashed unitary triangles of the CKM matrix: (a) and (b) correspond to the orthogonality relations (1.6) and (1.7), respectively.

matrix [9, 8]. However the experimental accuracy is getting to the point that we will have to take into account the next-to-leading order terms of the Wolfenstein expansion, and distinguish between the unitary triangles described by (1.6) and (1.7), which are illustrated in Figure 1.3. Here, $\overline{\rho}$ and $\overline{\eta}$ are related to the Wolfenstein parameters ρ and η through [6]

$$\overline{\rho} \equiv \left(1 - \lambda^2/2\right)\rho, \quad \overline{\eta} \equiv \left(1 - \lambda^2/2\right)\eta. \tag{1.8}$$

Note the angles of the triangles, in particular those designated by α, β, γ and $\delta\gamma$. And the sides R_b and R_t of the unitary triangle shown in Figure 1.3(a) are given as follows:

$$R_b = \left(1 - \frac{\lambda^2}{2}\right)\frac{1}{\lambda}\left|\frac{V_{ub}}{V_{cb}}\right| = \sqrt{\overline{\rho}^2 + \overline{\eta}^2} = 0.41 \pm 0.07, \tag{1.9}$$

$$R_t = \frac{1}{\lambda}\left|\frac{V_{td}}{V_{cb}}\right| = \sqrt{(1-\overline{\rho})^2 + \overline{\eta}^2} = \mathcal{O}(1), \tag{1.10}$$

1.2.3 B–$\bar{B}$ Mixing

The eigenstates of flavor, $B_q = (\bar{b}q)$ and $\bar{B}_q = (b\bar{q})$ ($q = d,\ s$), degenerate in pure QCD, mix on account of weak interactions. The quantum mechanics of the two-state system, with basis $\{|1\rangle, |2\rangle\} \equiv \{|B_q\rangle, |\bar{B}_q\rangle\}$, is described by a complex, 2×2 Hamiltonian matrix

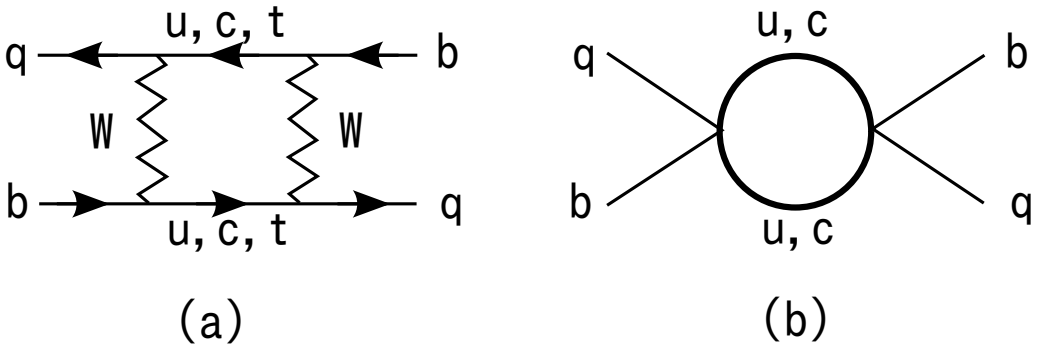


Figure 1.4: (a): General box diagram describing B–$\bar{B}$ mixing. (b): Special case of diagram (a) with internal u and c, whose absorptive part determines Γ_{12}.

$$\mathbf{H} = \mathbf{M} - \frac{i}{2}\mathbf{\Gamma} = \begin{pmatrix} M & M_{12} \\ M_{12}^* & M \end{pmatrix} - \frac{i}{2}\begin{pmatrix} \Gamma & \Gamma_{12} \\ \Gamma_{12}^* & \Gamma \end{pmatrix} \tag{1.11}$$

with Hermitian matrices $\mathbf{M}$ and $\mathbf{\Gamma}$. The off-diagonal elements in (1.11) arise from $\Delta B = 2$ flavor-changing transitions with virtual (M_{12}) or real intermediate states (Γ_{12}), in the latter case corresponding to decay channels common to B and $\bar{B}$.

Diagonalizing (1.11), one obtains the physical eigenstates B_H ('heavy'), B_L ('light') and the corresponding eigenvalues $M_{H,L} - \frac{i}{2}\Gamma_{H,L}$. The mass and width differences read

$$\Delta M_q \equiv M_H^{(q)} - M_L^{(q)} = 2|M_{12}^{(q)}|, \qquad \Delta\Gamma_q \equiv \Gamma_H^{(q)} - \Gamma_L^{(q)} = \frac{2\mathrm{Re}(M_{12}^{(q)*}\Gamma_{12}^{(q)})}{|M_{12}^{(q)}|}\,. \tag{1.12}$$

ΔM is positive by definition, $\Delta\Gamma$ is defined in such a way that a negative value[2] is obtained in the SM for the case of B_s, where a sizable width difference is expected. In the SM, the off-diagonal elements M_{12} and Γ_{12} inducing B mixing are described by the box diagrams in Figure 1.4.

1.2.4 The B_s System

The e^+–e^- B-factories operating at the $\Upsilon(4S)$ resonance will not be in a position to explore the B_s system. Since it is, moreover, very desirable to have large data samples available to study B_s decays, they are of particular interest for hadron machines. There are important differences between the B_d and B_s systems:

[2]Note that also the opposite sign convention for $\Delta\Gamma$ is used in the literature.

- Within the SM, a large B_s^0–$\overline{B_s^0}$ mixing parameter $x_s \equiv \Delta M_s/\Gamma_s = \mathcal{O}(20)$ is expected, whereas the mixing phase $\phi_s = -2\lambda^2\eta$ is expected to be very small.

- There may be a sizable width difference $\Delta\Gamma_s/\Gamma_s = \mathcal{O}(15\%)$, whereas $\Delta\Gamma_d$ is negligible.

The mass difference ΔM_s plays an important role to constrain the apex of the unitarity triangle shown in Figure 1.3(a), and the non-vanishing width difference $\Delta\Gamma_s$ may allow studies of CP-violating effects in "untagged" B_s rates, [22]–[30].

1.3 Phenomenology of the $B_s^0 \to J/\psi\phi$ Decay

Two flavor eigenstates, $|B_s^0\rangle$ and $|\bar{B}_s^0\rangle$, mix via the weak interaction. The two mass eigenstates

$$|B_s^H\rangle = p\,|B_s^0\rangle - q\,|\bar{B}_s^0\rangle, \qquad |B_s^L\rangle = p\,|B_s^0\rangle + q\,|\bar{B}_s^0\rangle$$

are labeled "heavy" and "light". The mass and lifetime differences between the B_s^H and B_s^L states can be defined as

$$\Delta m \equiv m_H - m_L, \quad \Delta\Gamma \equiv \Gamma_L - \Gamma_H, \quad \Gamma = (\Gamma_H + \Gamma_L)/2\,,$$

where $m_{H,L}$ and $\Gamma_{H,L}$ denote the mass and decay width of B_s^H and B_s^L (with this definition both Δm and $\Delta\Gamma$ are expected to be positive quantities). The heavy state decays with a longer lifetime, $\tau_H = 1/\Gamma_H$, while the light state decays with the shorter lifetime $\tau_L = 1/\Gamma_L$, in analogy to the neutral kaon system. The mean lifetime is defined to be $\tau = 1/\Gamma$. Theoretical estimates predict $\Delta\Gamma/\Gamma$ to be on the order of $\sim 15\%$ [54]. Linear polarization eigenstates of the J/ψ and ϕ provide a convenient basis for the analysis of the decay [21]. The two vector mesons can have their spins transversely polarized with respect to their momentum and be either parallel or perpendicular to each other. Alternatively, they can both be longitudinally polarized. We denote these states as $|\mathcal{P}_{||}\rangle$, $|\mathcal{P}_{\perp}\rangle$, and $|\mathcal{P}_0\rangle$.

In the standard model, CP violation occurs through complex phases in the CKM matrix [52]. Large phases occur in the matrix elements V_{ub} and V_{td}. While these matrix elements generate large CP violation in the B^0 system, they do not appear in leading order diagrams contributing to either $B_s^0 \leftrightarrow \bar{B}_s^0$ mixing or to the decay $B_s^0 \to J/\psi\phi$. For this reason the standard model expectation of CP violation in $B_s^0 \to J/\psi\phi$ is small. In the limit of vanishing CP violation, the heavy, long-lived mass eigenstate B_s^H is CP odd and decays to the CP-odd, L=1 orbital angular momentum state $|\mathcal{P}_\perp\rangle$. The light, short-lived mass eigenstate B_s^L is CP even and decays to both CP-even L=0 and L=2 orbital angular momentum states, which are linear combinations of $|\mathcal{P}_0\rangle$ and $|\mathcal{P}_{||}\rangle$.

The small CP violation in $B_s^0 \to J/\psi\phi$ can be quantified in the following way: we define A_i as the decay amplitude $\langle B_s|H|\mathcal{P}_i\rangle$ and $\bar{A}_i$ as the decay amplitude $\langle \bar{B}_s|H|\mathcal{P}_i\rangle$ where i is one of $\{||, \perp, 0\}$. All CP observables in the system are characterized by three quantities $\lambda_i = \frac{q}{p}\frac{\bar{A}_i}{A_i}$. In the standard model the λ_i are given as $\lambda_i = \pm\exp(i2\beta_s)$ where the positive and negative sign applies to the CP even and odd final state, and

$$\beta_s \equiv \arg\left(-\frac{V_{ts}V_{tb}^*}{V_{cs}V_{cb}^*}\right).$$

The standard model expectation [71] is $2\beta_s = 0.037 \pm 0.002$, a very small phase which does not lead to appreciable levels of CP violation. New physics can alter the mixing phase, while leaving λ very nearly unimodular. In this paper we consider, however, also the case in which $|\lambda| \neq 1$.

1.3.1 Differential Rates

The state of an initially pure B_s^0 or $\bar{B}_s^0$ meson after a proper time t has elapsed is denoted as $|B_{s,phys}^0(t)\rangle$ and $|\bar{B}_{s,phys}^0(t)\rangle$. Transitions of these states to the detectable $\mu^+\mu^-K^+K^-$ can be written as:

$$
\begin{aligned}
\langle \mu^+\mu^- \quad K \quad {}^+K^-|H|B^0_{s,phys}(t)\rangle & \\
&= \sum_i \langle \mu^+\mu^- K^+K^-|H|\mathcal{P}_i\rangle\langle\mathcal{P}_i|H|B^0_s\rangle\langle B^0_s|B^0_{s,phys}(t)\rangle \\
&+ \sum_i \langle \mu^+\mu^- K^+K^-|H|\mathcal{P}_i\rangle\langle\mathcal{P}_i|H|\bar{B}^0_s\rangle\langle \bar{B}^0_s|B^0_{s,phys}(t)\rangle, \\
\langle \mu^+\mu^- \quad K \quad {}^+K^-|H|\bar{B}^0_{s,phys}(t)\rangle & \\
&= \sum_i \langle \mu^+\mu^- K^+K^-|H|\mathcal{P}_i\rangle\langle\mathcal{P}_i|H|B^0_s\rangle\langle B^0_s|\bar{B}^0_{s,phys}(t)\rangle \\
&+ \sum_i \langle \mu^+\mu^- K^+K^-|H|\mathcal{P}_i\rangle\langle\mathcal{P}_i|H|\bar{B}^0_s\rangle\langle \bar{B}^0_s|\bar{B}^0_{s,phys}(t)\rangle.
\end{aligned}
\tag{1.13}
$$

where H is the weak interaction Hamiltonian. The expression can be written much more simply, by defining time-dependent amplitudes for $|B^0_s\rangle$ and $|\bar{B}^0_s\rangle$ to reach the states $|\mathcal{P}_i\rangle$ either with or without mixing:

$$
\begin{aligned}
\mathcal{A}_i(t) &\equiv \langle\mathcal{P}_i|H|B^0_s\rangle\langle B^0_s|B^0_{s,phys}(t)\rangle + \langle\mathcal{P}_i|H|\bar{B}^0_s\rangle\langle \bar{B}^0_s|B^0_{s,phys}(t)\rangle, \\
\bar{\mathcal{A}}_i(t) &\equiv \langle\mathcal{P}_i|H|B^0_s\rangle\langle B^0_s|\bar{B}^0_{s,phys}(t)\rangle + \langle\mathcal{P}_i|H|\bar{B}^0_s\rangle\langle \bar{B}^0_s|\bar{B}^0_{s,phys}(t)\rangle.
\end{aligned}
$$

Then:

$$
\begin{aligned}
\langle \mu^+\mu^- K^+K^-|H|B^0_{s,phys}(t)\rangle &= \sum_i \mathcal{A}_i(t)e^{-imt}\langle \mu^+\mu^- K^+K^-|H|\mathcal{P}_i\rangle, \\
\langle \mu^+\mu^- K^+K^-|H|\bar{B}^0_{s,phys}(t)\rangle &= \sum_i \bar{\mathcal{A}}_i(t)e^{-imt}\langle \mu^+\mu^- K^+K^-|H|\mathcal{P}_i\rangle\,,
\end{aligned}
\tag{1.14}
$$

where the time dependence of $\mathcal{A}_i(t)$ and $\bar{\mathcal{A}}_i(t)$ is:

$$
\begin{aligned}
\mathcal{A}_i(t) &= \frac{e^{-\Gamma t/2}}{\sqrt{\tau_H+\tau_L \pm \cos 2\beta_s\,(\tau_L-\tau_H)}}\left[E_+(t) \pm e^{2i\beta_s}E_-(t)\right]a_i\,, \\
\bar{\mathcal{A}}_i(t) &= \frac{e^{-\Gamma t/2}}{\sqrt{\tau_H+\tau_L \pm \cos 2\beta_s\,(\tau_L-\tau_H)}}\left[\pm E_+(t) + e^{-2i\beta_s}E_-(t)\right]a_i\,,
\end{aligned}
\tag{1.15}
$$

and where the upper sign indicates a CP even final state, the lower sign indicates a CP odd final state,

$$E_{\pm}(t) \equiv \frac{1}{2}\left[e^{+(\frac{-\Delta\Gamma}{4}+i\frac{\Delta m}{2})t} \pm e^{-(\frac{-\Delta\Gamma}{4}+i\frac{\Delta m}{2})t}\right], \tag{1.16}$$

and the a_i are complex amplitude parameters satisfying:

$$\sum_i |a_i|^2 = 1\,. \tag{1.17}$$

The final state $\mu^+\mu^-K^+K^-$ is characterized by three decay angles, described in a coordinate system[3] called the transversity basis [20]. In the J/ψ rest frame, the x-axis is taken to lie along the momentum of the ϕ and the z-axis perpendicular to the decay plane of the ϕ. The variables (θ, φ) are the polar and azimuthal angles of the μ^+ momentum in this basis. We also define the angle ψ to be the "helicity" angle in the ϕ decay, i.e. the angle between the K^+ direction and the x-axis in the ϕ rest frame. With these definitions, the muon momentum direction in the J/ψ rest frame is given by the unit vector

$$\hat{n} = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)\,. \tag{1.18}$$

Let $\mathbf{A}(t)$ and $\bar{\mathbf{A}}(t)$ be complex vector functions of time defined as

$$\begin{aligned}
\mathbf{A}(t) &= \left(\mathcal{A}_0(t)\cos\psi, -\frac{\mathcal{A}_{\|}(t)\sin\psi}{\sqrt{2}}, i\frac{\mathcal{A}_{\perp}(t)\sin\psi}{\sqrt{2}}\right), \\
\bar{\mathbf{A}}(t) &= \left(\bar{\mathcal{A}}_0(t)\cos\psi, -\frac{\bar{\mathcal{A}}_{\|}(t)\sin\psi}{\sqrt{2}}, i\frac{\bar{\mathcal{A}}_{\perp}(t)\sin\psi}{\sqrt{2}}\right),
\end{aligned} \tag{1.19}$$

where $\mathcal{A}_i(t)$ have now been normalized.

For experimental measurements we are concerned with normalized probability density functions P_B and $P_{\bar{B}}$ for B and $\bar{B}$ mesons in the variables t, $\cos\psi$, $\cos\theta$, and φ, which can be obtained by squaring Eq. (1.14). The formulae of Ref. [23] are then equivalent to:

[3] An alternate basis called the helicity basis is discussed further in Section 1.3.7.

$$P_B(\theta,\varphi,\psi,t) = \frac{9}{16\pi}|\mathbf{A}(t)\times\hat{n}|^2$$
$$P_{\bar{B}}(\theta,\varphi,\psi,t) = \frac{9}{16\pi}|\bar{\mathbf{A}}(\mathbf{t})\times\hat{n}|^2 \tag{1.20}$$

which give a picture of a time-dependent polarization analyzed in the decay[4]. The factors of $9/16\pi$ are normalization constants, and are present in order that

$$\int \sum_{j=B,\bar{B}} P_j(\psi,\theta,\varphi,t)d(\cos\psi)d(\cos\theta)d\varphi dt = 1\,. \tag{1.21}$$

The quantities $|a_i|^2$ give the time-integrated rate to each of the polarization states. The values of $\mathcal{A}_i(t)$ at $t=0$ will be denoted as A_i. To translate between the a's and the A's one can use the following two sets of transformations:

$$\begin{aligned}
|A_\perp|^2 &= \frac{|a_\perp|^2 y}{1+(y-1)|a_\perp|^2} \qquad & |a_\perp|^2 &= \frac{|A_\perp|^2}{y+(1-y)|A_\perp|^2}\\
|A_{||}|^2 &= \frac{|a_{||}|^2}{1+(y-1)|a_\perp|^2} \qquad & |a_{||}|^2 &= \frac{|A_{||}|^2 y}{y+(1-y)|A_\perp|^2}\\
|A_0|^2 &= \frac{|a_0|^2}{1+(y-1)|a_\perp|^2} \qquad & |a_0|^2 &= \frac{|A_0|^2 y}{y+(1-y)|A_\perp|^2}
\end{aligned} \tag{1.22}$$

where $y \equiv (1-z)/(1+z)$ and $z \equiv \cos 2\beta_s \Delta\Gamma/(2\Gamma)$. The relation (1.17) insures that

$$\sum_i |A_i|^2 = 1 \tag{1.23}$$

Eq. (1.20), together with the definitions in Eqs. (1.15), (1.16), and (1.18) can be used as a decay model for an event generator, and is suitable for use as a fitting function in the absence of detector effects.

[4]Throughout this chapter, when writing the dot product of two complex vectors, we always imply complex conjugation on the second operand.

1.3.2 Detector Efficiency and Normalization

The detector efficiency $\varepsilon(\psi,\theta,\varphi)$, when introduced into the above expression, disturbs the normalization of Eq. (1.21). We restore it by dividing by a normalization factor N,

$$\begin{aligned} P' &(\psi,\theta,\varphi,t) = \frac{1}{N} P(\psi,\theta,\varphi,t)\varepsilon(\psi,\theta,\varphi)\,, \\ N &= \int \sum_{i=B,\bar{B}} P_i(\psi,\theta,\varphi,t)\varepsilon(\psi,\theta,\varphi) d(\cos\psi) d(\cos\theta) d\varphi dt\,. \end{aligned} \tag{1.24}$$

Suppose that the efficiency $\varepsilon(\psi,\theta,\varphi)$ can be parametrized as

$$\varepsilon(\psi,\theta,\varphi) = c_{lm}^k P_k(\cos\psi) Y_{lm}(\theta,\varphi), \tag{1.25}$$

where c_{lm}^k are expansion coefficients, $P_k(\cos\psi)$ are Legendre polynomials, and $Y_{lm}(\theta,\varphi)$ are real harmonics related to the spherical harmonics through the following relations:

$$\begin{aligned} Y_{lm} &= Y_l^m && (m=0)\,, \\ Y_{lm} &= \tfrac{1}{\sqrt{2}}(Y_l^m + (-1)^m Y_l^{-m}) && (m>0)\,, \\ Y_{lm} &= \tfrac{1}{i\sqrt{2}}(Y_l^{|m|} - (-1)^{|m|} Y_l^{-|m|}) && (m<0)\,. \end{aligned} \tag{1.26}$$

The products $P_k(\cos\psi)Y_{lm}(\theta,\varphi)$ constitute an orthonormal basis for functions of the three angles. The detector efficiency (obtained, for example, from Monte Carlo simulation) can be fit to the first few of these polynomials. A straightforward calculation shows that:

$$
\begin{aligned}
N = & \\
& \frac{3}{8\sqrt{\pi}}\left[\frac{4c_{00}^{0}}{3}(|a_0|^2+|a_\parallel|^2+|a_\perp|^2)\right. \\
+ & \left.\frac{4c_{00}^{2}}{15}(2|a_0|^2-|a_\parallel|^2-|a_\perp|^2)\right] \\
+ & \frac{3}{8\sqrt{5\pi}}\left[\frac{2c_{20}^{0}}{3}(|a_0|^2+|a_\parallel|^2-2|a_\perp|^2)\right. \\
+ & \left.\frac{4c_{20}^{2}}{15}(|a_0|^2-\frac{1}{2}|a_\parallel|^2+|a_\perp|^2)\right] \\
- & \frac{9}{16\sqrt{15\pi}}\frac{\sin 2\beta_s(\tau_L-\tau_H)}{\sqrt{((\tau_L-\tau_H)\sin 2\beta_s)^2+4\tau_L\tau_H}} \\
& \times\left[(a_\parallel^* a_\perp + a_\parallel a_\perp^*)(\frac{4}{3}c_{2-1}^{0}-\frac{4}{15}c_{2-1}^{2})\right] \\
+ & \frac{9}{16}\frac{\sqrt{2}}{\sqrt{15\pi}}\frac{\sin 2\beta_s(\tau_L-\tau_H)}{\sqrt{((\tau_L-\tau_H)\sin 2\beta_s)^2+4\tau_L\tau_H}} \\
& \times\left[(a_0^* a_\perp + a_0 a_\perp^*)(\frac{\pi c_{21}^{1}}{8}-\frac{\pi c_{21}^{3}}{32}+...)\right] \\
+ & \frac{9}{8\sqrt{15\pi}}\left[\frac{2c_{22}^{0}}{3}(-|a_0|^2+|a_\parallel|^2)-\frac{4c_{22}^{2}}{15}(|a_0|^2+\frac{1}{2}|a_\parallel|^2)\right] \\
+ & \frac{9}{16}\frac{\sqrt{2}}{\sqrt{15\pi}}\left[(a_0^* a_\parallel + a_0 a_\parallel^*)(\frac{\pi c_{2-2}^{1}}{8}-\frac{\pi c_{2-2}^{3}}{32}+...)\right]. \qquad (1.27)
\end{aligned}
$$

The numerical factors $+\pi/8$ and $-\pi/32$, appearing together with $c_{2,1}^{k}$ and $c_{2,-2}^{k}$ in the infinite series, are the integrals

$$\int P_k(\cos\psi)\cos(\psi)\sin\psi d(\cos\psi)\,. \qquad (1.28)$$

While this series is infinite, the number of basis functions needed to fit detector efficiencies in a particular analysis is finite and determined chiefly by the size of the data sample. With the factors in Eq. (1.28) the normalizing factor can be adapted to account for all terms used in the expansion of the efficiency. Eq. (1.27) represents an analytic normalization of the fitting function and provides an efficient way to compute the likelihood during a maximum log likelihood fit. The orthonormality of the basis functions has been used to reduce the expression to its final form.

1.3.3 Time Development

The short oscillation length of the B_s^0 meson [10], requires us to account for resolution effects when fitting the rates of flavor-tagged decays, even using the best silicon vertex detectors, which have proper decay length resolutions on the order of 25 μm. Certain time-dependent functions arising from particle-antiparticle oscillations, particularly those expressed as the product of exponential decays and harmonic functions with frequency Δm, must be convolved with one or more Gaussian components describing detector resolution. This convolution can be carried out analytically, using the method described in Ref. [33] for the evaluation of certain integrals which are equivalent to complex error functions. In this step one requires that various components of the time dependence first be separated from Eq. (1.20). The time development of $\mathcal{A}_0(t)$ and $\mathcal{A}_\parallel(t)$ amplitudes are identical, but differs from that of $\mathcal{A}_\perp(t)$. We begin by decomposing

$$\mathbf{A}(t) = \mathbf{A}_+(t) + \mathbf{A}_-(t), \quad \bar{\mathbf{A}}(t) = \bar{\mathbf{A}}_+(t) + \bar{\mathbf{A}}_-(t) \tag{1.29}$$

where

$$\begin{aligned} \mathbf{A}_+(t) = \mathbf{A}_+ f_+(t) &= (a_0 \cos\psi, -\frac{a_\parallel \sin\psi}{\sqrt{2}}, 0) \cdot f_+(t)\,, \\ \bar{\mathbf{A}}_+(t) = \bar{\mathbf{A}}_+ \bar{f}_+(t) &= (a_0 \cos\psi, -\frac{a_\parallel \sin\psi}{\sqrt{2}}, 0) \cdot \bar{f}_+(t)\,, \end{aligned} \tag{1.30}$$

and

$$\begin{aligned} \mathbf{A}_-(t) = \mathbf{A}_- f_-(t) &= (0, 0, i\frac{a_\perp \sin\psi}{\sqrt{2}}) \cdot f_-(t)\,, \\ \bar{\mathbf{A}}_-(t) = \bar{\mathbf{A}}_- \bar{f}_-(t) &= (0, 0, i\frac{a_\perp \sin\psi}{\sqrt{2}}) \cdot \bar{f}_-(t)\,, \end{aligned} \tag{1.31}$$

and we define

$$
\begin{aligned}
f_{\pm}(t) &= \frac{e^{-\Gamma t/2}}{\sqrt{\tau_H + \tau_L \pm \cos 2\beta_s \, (\tau_L - \tau_H)}} \left[E_+(t) \pm e^{2i\beta_s} E_-(t) \right] , \\
\bar{f}_{\pm}(t) &= \frac{e^{-\Gamma t/2}}{\sqrt{\tau_H + \tau_L \pm \cos 2\beta_s \, (\tau_L - \tau_H)}} \left[\pm E_+(t) + e^{-2i\beta_s} E_-(t) \right] .
\end{aligned}
\tag{1.32}
$$

We then have in place of Eq. (1.20)

$$
\begin{aligned}
& P_B \quad (\theta, \psi, \varphi, t) \\
&= \frac{9}{16\pi} \left\{ |\mathbf{A}_+(t) \times \hat{n}|^2 + |\mathbf{A}_-(t) \times \hat{n}|^2 + 2Re((\mathbf{A}_+(t) \times \hat{n}) \cdot (\mathbf{A}_-^*(t) \times \hat{n})) \right\} \\
&= \frac{9}{16\pi} \left\{ |\mathbf{A}_+ \times \hat{n}|^2 |f_+(t)|^2 + |\mathbf{A}_- \times \hat{n}|^2 |f_-(t)|^2 \right. \\
&\quad \left. + 2Re((\mathbf{A}_+ \times \hat{n}) \cdot (\mathbf{A}_-^* \times \hat{n}) \cdot f_+(t) \cdot f_-^*(t)) \right\}
\end{aligned}
\tag{1.33}
$$

and

$$
\begin{aligned}
& P_{\bar{B}} \quad (\theta, \psi, \varphi, t) \\
&= \frac{9}{16\pi} \left\{ |\bar{\mathbf{A}}_+(t) \times \hat{n}|^2 + |\bar{\mathbf{A}}_-(t) \times \hat{n}|^2 + 2Re(\bar{\mathbf{A}}_+(t) \times \hat{n}) \cdot (\bar{\mathbf{A}}_-^*(t) \times \hat{n})) \right\} \\
&= \frac{9}{16\pi} \left\{ |\mathbf{A}_+ \times \hat{n}|^2 |\bar{f}_+(t)|^2 + |\mathbf{A}_- \times \hat{n}|^2 |\bar{f}_-(t)|^2 \right. \\
&\quad \left. + 2Re((\mathbf{A}_+ \times \hat{n}) \cdot (\mathbf{A}_-^* \times \hat{n}) \cdot \bar{f}_+(t) \cdot \bar{f}_-^*(t) \right\}
\end{aligned}
\tag{1.34}
$$

where (for $\bar{B}$) the diagonal term in Eq. (1.34) is

$$
|\bar{f}_{\pm}(t)|^2 = \frac{1}{2} \frac{(1 \pm \cos 2\beta_s) e^{-\Gamma_L t} + (1 \mp \cos 2\beta_s) e^{-\Gamma_H t} \pm 2 \sin 2\beta_s e^{-\Gamma t} \sin \Delta m t}{\tau_L (1 \pm \cos 2\beta_s) + \tau_H (1 \mp \cos 2\beta_s)} ,
\tag{1.35}
$$

while (for B) the diagonal term in Eq. (1.33) is

$$
|f_{\pm}(t)|^2 = \frac{1}{2} \frac{(1 \pm \cos 2\beta_s) e^{-\Gamma_L t} + (1 \mp \cos 2\beta_s) e^{-\Gamma_H t} \mp 2 \sin 2\beta_s e^{-\Gamma t} \sin \Delta m t}{\tau_L (1 \pm \cos 2\beta_s) + \tau_H (1 \mp \cos 2\beta_s)}
\tag{1.36}
$$

and (for $\bar{B}$) the cross-term, or interference term in Eq. (1.34) is

$$\bar{f_+}(t)\bar{f_-}^*(t) = \frac{-e^{-\Gamma t}\cos\Delta m t - i\cos 2\beta_s e^{-\Gamma t}\sin\Delta m t + i\sin 2\beta_s(e^{-\Gamma_L t} - e^{-\Gamma_H t})/2}{\sqrt{\left[(\tau_L - \tau_H)\sin 2\beta_s\right]^2 + 4\tau_L\tau_H}}\,, \tag{1.37}$$

while (for B) the interference term in Eq. (1.33) is

$$f_+(t)f_-^*(t) = \frac{e^{-\Gamma t}\cos\Delta m t + i\cos 2\beta_s e^{-\Gamma t}\sin\Delta m t + i\sin 2\beta_s(e^{-\Gamma_L t} - e^{-\Gamma_H t})/2}{\sqrt{\left[(\tau_L - \tau_H)\sin 2\beta_s\right]^2 + 4\tau_L\tau_H}}\,. \tag{1.38}$$

This accomplishes the desired separation. In the fitting function, to accommodate the proper time resolution, one has only to replace all time-dependent functions with their smeared equivalents.

1.3.4 Sensitivity to Δm

It can be noticed that the time development of the interference term, expressions 1.37 and 1.38, contain undiluted mixing asymmetries *even in the case of no CP violation*, i.e., when $\beta_s = 0$. Let us try to better understand the mechanism by which the flavor of the B_s^0 meson is tagged at *decay* time, by first rewriting Eq. (1.13) using the B_s^H and B_S^L states in the expansion rather than the B_s^0 and $\bar{B}_s^0$ states:

$$\begin{aligned}\langle\ \mu^+\mu^-\quad K^+K^-|H|B^0_{s,phys}(t)\rangle = \\ \sum_i \langle\mu^+\mu^-K^+K^-|H|\mathcal{P}_i\rangle\langle\mathcal{P}_i|H|B_s^H\rangle\langle B_s^H|B^0_{s,phys}(t)\rangle \\ +\quad \sum_i \langle\mu^+\mu^-K^+K^-|H|\mathcal{P}_i\rangle\langle\mathcal{P}_i|H|B_s^L\rangle\langle B_s^L|B^0_{s,phys}(t)\rangle. \end{aligned} \tag{1.39}$$

Now, we take the limit of zero CP violation in the B_s^0 system, such that $\langle\mathcal{P}_{||}|H|B_s^H\rangle = \langle\mathcal{P}_0|H|B_s^H\rangle = \langle\mathcal{P}_\perp|H|B_s^L\rangle = 0$, and only three of the six terms in Eq. (1.39) remain:

$$
\begin{aligned}
\langle \quad \mu^+\mu^- \quad & K^+K^-|H|B^0_{s,phys}(t)\rangle = \\
& \langle\mu^+\mu^-K^+K^-|H|\mathcal{P}_\perp\rangle\langle\mathcal{P}_\perp|H|B^H_s\rangle\langle B^H_s|B^0_{s,phys}(t)\rangle \\
+ \quad & \langle\mu^+\mu^-K^+K^-|H|\mathcal{P}_0\rangle\langle\mathcal{P}_0|H|B^L_s\rangle\langle B^L_s|B^0_{s,phys}(t)\rangle \\
+ \quad & \langle\mu^+\mu^-K^+K^-|H|\mathcal{P}_{||}\rangle\langle\mathcal{P}_{||}|H|B^L_s\rangle\langle B^L_s|B^0_{s,phys}(t)\rangle. \qquad (1.40)
\end{aligned}
$$

When the expression is squared, the interference terms are the cross terms involving both the product of a CP-even and a CP-odd amplitudes. The time dependence of these terms is contained in the factor:

$$
\begin{aligned}
\langle \quad B^H_s \quad & |B^0_{s,phys}(t)\rangle\langle B^L_s|B^0_{s,phys}(t)\rangle = \\
& \frac{1}{4}\left[\left(\langle B^H_s|B^0_{s,phys}(t)\rangle + \langle B^L_s|B^0_{s,phys}(t)\rangle\right)^2\right. \\
& \left. - \left(\langle B^H_s|B^0_{s,phys}(t)\rangle - \langle B^L_s|B^0_{s,phys}(t)\rangle\right)^2\right] \\
= \quad & \frac{1}{2}\left[\left(\frac{\langle B^H_s| + \langle B^L_s|}{\sqrt{2}}|B^0_{s,phys}(t)\rangle\right)^2 - \left(\frac{\langle B^H_s| - \langle B^L_s|}{\sqrt{2}}|B^0_{s,phys}(t)\rangle\right)^2\right] \\
= \quad & \frac{1}{2}\left[\langle B^0_s|B^0_{s,phys}(t)\rangle^2 - \langle\bar{B}^0_s|B^0_{s,phys}(t)\rangle^2\right].
\end{aligned}
\qquad (1.41)
$$

This factor takes the value +1/2 when the meson is pure B^0_s, and -1/2 when the meson is pure $\bar{B}^0_s$, and in general oscillates between these two values. Thus the interference term effectively tags the flavor of the B^0_s at decay. This provides a way to observe $B^0_s \to \bar{B}^0_s$ flavor oscillations using a sample of flavor-tagged $B^0_s \to J/\psi\phi$ decays which can be collected with a simple dimuon trigger.

1.3.5 Incorporating Direct CP Violation

An asymmetry either in the decay rate ($|\bar{A}_i/A_i| \neq 1$) or in the mixing ($|q/p| \neq 1$) such that $|\lambda| \neq 1$ is direct CP violation. In the case of direct CP violation λ does not lie on the unit circle in the complex plane, and we need two parameters to describe it which we will take to be $\mathcal{C} \equiv Re(\lambda)$ and $\mathcal{S} \equiv Im(\lambda)$. Experimentally, even if one sets out to extract β_s assuming the constraint $|\lambda| = 1$, it is nonetheless

of interest to test that constraint, since sensitivity to $\mathcal{C}$ and $\mathcal{S}$ arise from very different features of the detector. In that case we must revisit not only the functional form of the differential decay rates, but also the normalization. The amplitudes in Eq. (1.15) must now be written as:

$$\begin{aligned} \mathcal{A}_i &= \mathcal{N}_{\pm} e^{-\Gamma t/2} \left[E_+(t) \pm \lambda E_-(t) \right] a_i \,, \\ \bar{\mathcal{A}}_i &= \mathcal{N}_{\pm} e^{-\Gamma t/2} \left[\pm E_+(t) + E_-(t)/\lambda \right] a_i \,, \end{aligned} \tag{1.42}$$

where

$$\begin{aligned} \mathcal{N}_{\pm} &= \left\{ \frac{1}{4|\lambda|^2} \left[\left[(\tau_H + \tau_L)(1+|\lambda|^2)^2 \pm 2\mathcal{C} \cdot (\tau_L - \tau_H)(1+|\lambda|^2) \right] \right. \right. \\ &\quad \left. \left. + \frac{\tau}{1+\Delta m^2 \tau^2} \cdot \left[\pm 4\mathcal{S} \cdot \left(1-|\lambda|^2\right) \Delta m \tau - 2\left(1-|\lambda|^2\right)^2 \right] \right] \right\}^{-\frac{1}{2}} . \end{aligned}$$

These amplitudes can readily be seen to reduce to those of Eq. (1.15) in the limit of $\mathcal{C}^2 + \mathcal{S}^2 \equiv |\lambda|^2 \to 1$. The normalization of detector efficiency, Eq. (1.27), becomes:

$$
\begin{aligned}
N = \quad & \frac{3}{8\sqrt{\pi}}\left[\frac{4c_{00}^0}{3}(|a_0|^2+|a_\parallel|^2+|a_\perp|^2)\right. \\
+ \quad & \left.\frac{4c_{00}^2}{15}(2|a_0|^2-|a_\parallel|^2-|a_\perp|^2)\right] \\
+ \quad & \frac{3}{8\sqrt{5\pi}}\left[\frac{2c_{20}^0}{3}(|a_0|^2+|a_\parallel|^2-2|a_\perp|^2)\right. \\
+ \quad & \left.\frac{4c_{20}^2}{15}(|a_0|^2-\frac{1}{2}|a_\parallel|^2+|a_\perp|^2)\right] \\
- \quad & \frac{9}{16\sqrt{15\pi}}\mathcal{N}_+\mathcal{N}_-\mathcal{S}\cdot(\tau_L-\tau_H) \\
& \times\left[(a_\parallel^* a_\perp + a_\parallel a_\perp^*)(\frac{4}{3}c_{2-1}^0-\frac{4}{15}c_{2-1}^2)\right] \\
+ \quad & \frac{9}{16}\frac{\sqrt{2}}{\sqrt{15\pi}}\mathcal{N}_+\mathcal{N}_-\mathcal{S}\cdot(\tau_L-\tau_H) \\
& \times\left[(a_0^* a_\perp + a_0 a_\perp^*)(\frac{\pi c_{21}^1}{8}-\frac{\pi c_{21}^3}{32}+...)\right] \\
+ \quad & \frac{9}{8\sqrt{15\pi}}\left[\frac{2c_{22}^0}{3}(-|a_0|^2+|a_\parallel|^2)-\frac{4c_{22}^2}{15}(|a_0|^2+\frac{1}{2}|a_\parallel|^2)\right] \\
+ \quad & \frac{9}{16}\frac{\sqrt{2}}{\sqrt{15\pi}}\left[(a_0^* a_\parallel + a_0 a_\parallel^*)(\frac{\pi c_{2-2}^1}{8}-\frac{\pi c_{2-2}^3}{32}+...)\right].
\end{aligned}
$$

Finally, the explicit time development, Eqs. (1.35), (1.36), (1.37) and (1.38), must be replaced with the more general forms:

$$
\begin{aligned}
|\bar{f}_\pm(t)|^2 &= \frac{\mathcal{N}_\pm^2}{4|\lambda|^2}\left[((1+|\lambda|^2)\pm 2\mathcal{C})e^{-\Gamma_L t}+((1+|\lambda|^2)\mp 2\mathcal{C})e^{-\Gamma_H t}\right. \\
&\quad \left.+\left(\pm 4\mathcal{S}\sin\Delta m t-2(1-|\lambda|^2)\cos\Delta m t\right)e^{-\Gamma t}\right], \\
|f_\pm(t)|^2 &= \frac{\mathcal{N}_\pm^2}{4}\left[((1+|\lambda|^2)\pm 2\mathcal{C})e^{-\Gamma_L t}+((1+|\lambda|^2)\mp 2\mathcal{C})e^{-\Gamma_H t}\right. \\
&\quad \left.-\left(\pm 4\mathcal{S}\sin\Delta m t-2(1-|\lambda|^2)\cos\Delta m t\right)e^{-\Gamma t}\right], \\
\bar{f}_+(t)\bar{f}_-{}^*(t) &= \frac{\mathcal{N}_+\mathcal{N}_-}{4|\lambda|^2}\left[-e^{-\Gamma t}\left(2(1+|\lambda|^2)\cos\Delta m t+4i\mathcal{C}\sin\Delta m t\right)\right. \\
&\quad \left.+e^{-\Gamma_L t}\left((1-|\lambda|^2)+2i\mathcal{S}\right)+e^{-\Gamma_H t}\left((1-|\lambda|^2)-2i\mathcal{S}\right)\right], \\
f_+(t)f_-^*(t) &= \frac{\mathcal{N}_+\mathcal{N}_-}{4}\left[e^{-\Gamma t}\left(2(1+|\lambda|^2)\cos\Delta m t+4i\mathcal{C}\sin\Delta m t\right)\right. \\
&\quad \left.+e^{-\Gamma_L t}\left((1-|\lambda|^2)+2i\mathcal{S}\right)+e^{-\Gamma_H t}\left((1-|\lambda|^2)-2i\mathcal{S}\right)\right],
\end{aligned}
$$

which can be seen to reduce to expression 1.35, 1.36 and 1.37, 1.38 as $|\lambda|^2 \to 1$.

1.3.6 Incorporating a Contribution from $B_s^0 \to J/\psi K^+K^-$ (Kaons in an S-Wave State)

It has been suggested [67] that a contribution from S-wave K^+K^- under the ϕ peak in $B_s^0 \to J/\psi\phi$ decay may contribute up to 5-10% of the total rate. A normalized probability density for the decay $B_s^0 \to J/\psi K^+K^-$ (kaons in an S-wave state) can be worked out by considering the polarization vector of the J/ψ in the decay and proceeding as in [21]. The resulting expressions

$$\begin{aligned} Q_B(\theta,\varphi,\psi,t) &= \frac{3}{16\pi}|\mathbf{B}(t)\times\hat{n}|^2\,, \\ Q_{\bar{B}}(\theta,\varphi,\psi,t) &= \frac{3}{16\pi}|\bar{\mathbf{B}}(\mathbf{t})\times\hat{n}|^2 \end{aligned} \tag{1.43}$$

do not depend at all on the angle ψ (which is the helicity angle in the ϕ decay). In the previous expression

$$\begin{aligned} \mathbf{B}(t) &= (\mathcal{B}(t),0,0)\,, \\ \bar{\mathbf{B}}(t) &= \left(\bar{\mathcal{B}}(t),0,0\right) \end{aligned} \tag{1.44}$$

where the time-dependent amplitudes,

$$\begin{aligned} \mathcal{B}(t) &= \frac{e^{-\Gamma t/2}}{\sqrt{\tau_H+\tau_L-\cos 2\beta_s\,(\tau_L-\tau_H)}}\left[E_+(t)-e^{2i\beta_s}E_-(t)\right], \\ \bar{\mathcal{B}}(t) &= \frac{e^{-\Gamma t/2}}{\sqrt{\tau_H+\tau_L-\cos 2\beta_s\,(\tau_L-\tau_H)}}\left[-E_+(t)+e^{-2i\beta_s}E_-(t)\right] \end{aligned} \tag{1.45}$$

reflect the CP-odd nature of the $J/\psi KK$ final state.

When both P-wave and S-wave are present, the amplitudes must be summed and then squared. The P wave has a resonant structure due to the ϕ-propagator, while the S-wave amplitude is flat (but can have any phase with respect the P-wave). Suppose that in our experiment we accept events for which the reconstructed mass $m(K^+K^-) \equiv \mu$ lies within a window $\mu_{lo} < \mu < \mu_{hi}$. The normalized probability in this case is

$$\begin{aligned}\rho_B(\theta,\varphi,\psi,t,\mu) &= \frac{9}{16\pi}\left|\left[\sqrt{1-F_s}g(\mu)\mathbf{A}(t)+e^{i\delta_s}\sqrt{F_s}\frac{h(\mu)}{\sqrt{3}}\mathbf{B}(t)\right]\times\hat{n}\right|^2,\\ \rho_{\bar{B}}(\theta,\varphi,\psi,t,\mu) &= \frac{9}{16\pi}\left|\left[\sqrt{1-F_s}g(\mu)\bar{\mathbf{A}}(t)+e^{i\delta_s}\sqrt{F_s}\frac{h(\mu)}{\sqrt{3}}\bar{\mathbf{B}}(\mathbf{t})\right]\times\hat{n}\right|^2,\end{aligned} \tag{1.46}$$

where we use a nonrelativistic Breit-Wigner to model the ϕ resonance[5]

$$g(\mu)=\sqrt{\frac{\Gamma_\phi/2}{\Delta\omega}}\cdot\frac{1}{\mu-\mu_\phi+i\Gamma_\phi/2} \tag{1.47}$$

a flat model for the S-wave mass distribution

$$h(\mu)=\frac{1}{\sqrt{\Delta\mu}} \tag{1.48}$$

and define

$$\omega_{hi}=\tan^{-1}\frac{2(\mu_{hi}-\mu_\phi)}{\Gamma_\phi}\quad \omega_{lo}=\tan^{-1}\frac{2(\mu_{lo}-\mu_\phi)}{\Gamma_\phi} \tag{1.49}$$

and

$$\Delta\mu=\mu_{hi}-\mu_{lo}\qquad \Delta\omega=\omega_{hi}-\omega_{lo}\,. \tag{1.50}$$

In these equations, F_s is the S-wave fraction; μ_ϕ is the ϕ mass (1019 MeV/c^2); Γ_ϕ is the ϕ width (4.26 MeV/c^2), and δ_s is the phase of the S-wave component relative to the P-wave component.

In the presence of an S-wave contribution, the normalization of Eq. (1.27) must be generalized; in order to do this we first define the quantities

$$\mathcal{F}(\mu)\equiv\sqrt{\frac{F_s(1-F_s)\Gamma_\phi}{2\Delta\mu\Delta\omega}}\cdot\frac{e^{-i\delta_s}}{\mu-\mu_\phi+i\Gamma_\phi/2} \tag{1.51}$$

and

[5]We shall have more to say about that, later.

$$\mathcal{I}_\mu \equiv \int \mathcal{F}(\mu) d\mu = \sqrt{\frac{F_s(1-F_s)\Gamma_\phi}{2\Delta\mu\Delta\omega}} \cdot e^{-i\delta_s} \cdot \log \frac{\mu_{hi} - \mu_\phi + i\Gamma_\phi/2}{\mu_{lo} - \mu_\phi + i\Gamma_\phi/2} \,. \tag{1.52}$$

Then the normalizing factor appropriate for Eq. (1.46) is

$$\mathcal{N} = (1 - F_s) \cdot N + 2Re\left[\mathcal{I}_\mu \cdot N'\right] + F_s \cdot N'' \tag{1.53}$$

where N is given in Eq. (1.27), and

$$\begin{aligned} N' = \quad & \sqrt{3} * a_0^*\left(\frac{1}{6\sqrt{\pi}} c_{00}^1 + \frac{1}{12\sqrt{5\pi}} c_{20}^1 - \frac{1}{4\sqrt{15\pi}} c_{22}^1\right) \\ + \quad & \frac{3}{16}\sqrt{\frac{2}{5\pi}} a_\parallel^* \left(\frac{\pi}{2} c_{2-2}^0 - \frac{\pi}{8} c_{2-2}^2 + ...\right) \\ + \quad & \frac{3}{16}\sqrt{\frac{2}{5\pi}} a_\perp^* \frac{\sin 2\beta_s(\tau_L - \tau_H)}{\sqrt{((\tau_L - \tau_H)\sin 2\beta_s)^2 + 4\tau_L\tau_H}} \left(\frac{\pi}{2} c_{21}^0 - \frac{\pi}{8} c_{21}^2 + ...\right) \end{aligned} \tag{1.54}$$

and

$$N'' = \frac{1}{2\sqrt{\pi}} c_{00}^0 + \frac{1}{4\sqrt{5\pi}} c_{20}^0 - \frac{3}{4\sqrt{15\pi}} c_{22}^0 \,. \tag{1.55}$$

The numerical factors $+\pi/2$ and $-\pi/8$ appearing together with $c_{2,1}^k$ and $c_{2,-2}^k$ in the infinite series are the integrals

$$\int P_k(\cos\psi) \sin\psi d(\cos\psi) \,. \tag{1.56}$$

We now work out the explicit time and mass dependence of the differential rates. We will use Eq. (1.33) together with the analogous equation for the pure S-wave differential rate:

$$\begin{aligned} Q_B(\theta, \psi, \varphi, t) &= \frac{3}{16\pi} |\mathbf{B}(\mathbf{t}) \times \hat{n}|^2 \\ &= \frac{3}{16\pi} |\mathbf{B} \times \hat{n}|^2 |f_-(t)|^2 \end{aligned} \tag{1.57}$$

and

$$\begin{aligned} Q_{\bar{B}}(\theta,\psi,\varphi,t) &= \frac{3}{16\pi}|\bar{\mathbf{B}}(t)\times\hat{n}|^2 \\ &= \frac{3}{16\pi}|\mathbf{B}\times\hat{n}|^2|\bar{f}_-(t)|^2\,. \end{aligned} \tag{1.58}$$

where the vector $\mathbf{B} = \hat{x} = (1,0,0)$. The full probability densities, which can be used in a time-, angle-, and ϕ mass-dependent fit, are obtained by expanding Eq. (1.46). We get

$$\begin{aligned} \rho_B \quad & (\theta,\psi,\varphi,t,\mu) = \\ & (1-F_s)\frac{\Gamma_\phi/2}{\Delta\omega}\cdot\frac{1}{(\mu-\mu_\phi)^2+\Gamma_\phi^2/4}\cdot P_B(\theta,\psi,\varphi,t) \\ & +F_s\frac{1}{\Delta\mu}Q_B(\theta,\psi,\varphi,t) \\ & +2\frac{\sqrt{27}}{16\pi}Re\left[\mathcal{F}(\mu)\left((\mathbf{A}_-\times\hat{n})\cdot(\mathbf{B}\times\hat{n})\cdot|f_-(t)|^2\right.\right. \\ & \left.\left.+(\mathbf{A}_+\times\hat{n})\cdot(\mathbf{B}\times\hat{n})\cdot f_+(t)\cdot f_-^*(t)\right)\right] \end{aligned} \tag{1.59}$$

and

$$\begin{aligned} \rho_{\bar{B}} \quad & (\theta,\psi,\varphi,t,\mu) = \\ & (1-F_s)\frac{\Gamma_\phi/2}{\Delta\omega}\cdot\frac{1}{(\mu-\mu_\phi)^2+\frac{1}{1+F_S}\Gamma_\phi^2/4}\cdot P_{\bar{B}}(\theta,\psi,\varphi,t) \\ & +F_s\frac{1}{\Delta\mu}Q_{\bar{B}}(\theta,\psi,\varphi,t) \\ & +2\frac{\sqrt{27}}{16\pi}Re\left[\mathcal{F}(\mu)\left((\mathbf{A}_-\times\hat{n})\cdot(\mathbf{B}\times\hat{n})\cdot|\bar{f}_-(t)|^2\right.\right. \\ & \left.\left.+(\mathbf{A}_+\times\hat{n})\cdot(\mathbf{B}\times\hat{n})\cdot\bar{f}_+(t)\cdot\bar{f}_-^*(t)\right)\right]\,. \end{aligned} \tag{1.60}$$

In case one does not want to observe the ϕ-mass variable μ, one can integrate it out. Then one obtains

$$\begin{aligned}\rho_B \quad & (\theta,\psi,\varphi,t) = \\ & (1-F_s)\cdot P_B(\theta,\psi,\varphi,t) + F_s Q_B(\theta,\psi,\varphi,t) \\ & +2\frac{\sqrt{27}}{16\pi} Re\left[\mathcal{I}_\mu\left((\mathbf{A}_-\times\hat{n})\cdot(\mathbf{B}\times\hat{n})\cdot|f_-(t)|^2\right.\right. \\ & \left.\left.+(\mathbf{A}_+\times\hat{n})\cdot(\mathbf{B}\times\hat{n})\cdot f_+(t)\cdot f_-^*(t)\right)\right],\end{aligned} \tag{1.61}$$

$$\begin{aligned}\rho_{\bar{B}} \quad & (\theta,\psi,\varphi,t) = \\ & (1-F_S)\cdot P_{\bar{B}}(\theta,\psi,\varphi,t) + F_s Q_{\bar{B}}(\theta,\psi,\varphi,t) \\ & +2\frac{\sqrt{27}}{16\pi} Re\left[\mathcal{I}_\mu\left((\mathbf{A}_-\times\hat{n})\cdot(\mathbf{B}\times\hat{n})\cdot|\bar{f}_-(t)|^2\right.\right. \\ & \left.\left.+(\mathbf{A}_+\times\hat{n})\cdot(\mathbf{B}\times\hat{n})\cdot\bar{f}_+(t)\cdot\bar{f}_-^*(t)\right)\right].\end{aligned} \tag{1.62}$$

1.3.7 Symmetries

In this section we examine the symmetries of our differential rate formulae, starting from the simplest case, K^+K^- in a P-wave, Eq. (1.20), but considering also the case where both P and S waves are included, Eq. (1.46). In the case of pure P-wave, one can readily spot that the probability densities in Eq. (1.20) are invariant to the following transformations:

- A simultaneous rotation of the vectors $\mathbf{A}(t)$ and $\hat{n}$
- An inversion of the vector $\mathbf{A}(t)$
- Complex-conjugation of the vector $\mathbf{A}(t)$

The symmetry to simultaneous rotation of the vectors $\mathbf{A(t)}$ and $\hat{n}$ corresponds to the well-known freedom to choose a convenient basis in which to work.

An example of an alternative basis is the *helicity basis*, which derives from the *transversity basis* by a cyclic permutation of the coordinate axis: $\hat{x}_T = \hat{z}_H$, etc. One can take the angles in Eq. (1.18) to be the polar and azimuthal angles in the helicity basis, but then one must transform $\mathbf{A}(\mathbf{t})$ accordingly, i.e, by permuting the elements of $\mathbf{A}(\mathbf{t})$ in the defining equation, Eq. (1.19). Then, Eq. (1.20) remains valid in the helicity basis. This rotational invariance implies that the choice of basis is irrelevant to the final result since the likelihood is invariant to the choice (though we do not rule out the possibility that the quality of the efficiency expansion, Eq. (1.25), may depend on the choice of basis, as pointed out in [41]).

A more interesting symmetry is the symmetry that results from transforming $\mathbf{A}(t)$ to its complex conjugate. If we take, by convention, a_0 to be real and let $\delta_\parallel = \arg(a_\parallel)$, and $\delta_\perp = \arg(a_\perp)$, then as we will demonstrate below, this conjugation transformation is equivalent to the simultaneous transformation:

$$\begin{aligned} \beta_s &\rightarrow \pi/2 - \beta_s \\ \Delta\Gamma &\rightarrow -\Delta\Gamma \\ \delta_\perp &\rightarrow \pi - \delta_\perp \\ \delta_\parallel &\rightarrow 2\pi - \delta_\parallel \,. \end{aligned} \tag{1.63}$$

That is to say that the simultaneous transformation of these four variables is a symmetry of the likelihood *because* it transforms $\mathbf{A}(\mathbf{t})$ into its complex conjugate. Since for pure P wave state the combined transformation is a well-known symmetry, this observation may appear as a curiosity; however when both P and S wave states are included, we shall see that complex conjugation teaches us how to properly extend the symmetry. First, we show how the combined transformation accomplishes the claimed complex conjugation.

1. Note from Eq. (1.16) that the combined transformation transforms $E_\pm(t) \rightarrow \pm E^*_\pm(t)$.

2. Note also that the combined transformation transforms $e^{-2i\beta_s} \rightarrow -e^{+2i\beta_s}$ and $e^{+2i\beta_s} \rightarrow -e^{-2i\beta_s}$

3. Therefore, in Eq. (1.15), the terms in square brackets are transformed into their complex conjugates.

4. Note that both $\cos 2\beta_s$ and $\tau_L - \tau_H$ change sign under the transformation, so also the piece of Eq. (1.15) in the denominator, under the square root sign, is invariant under the combined transformation; since that piece is real we can say that it is anyway equal to its complex conjugate.

5. The real quantity a_0 does not change under the combined transformation, but since it is real, it is anyway equal to a_0^*.

6. The combined transformation transforms $a_\parallel \to a_\parallel^*$.

7. The combined transformation transforms $ia_\perp \to -ia_\perp^*$.

8. Then looking at Eq. (1.19), one sees finally that the net effect of the combined transformation has been the complex conjugation of the vector $\mathbf{A(t)}$.

Returning now to the full likelihood including both P and S wave states, Eq. (1.46), we can see that, here again, complex conjugation of the term

$$\sqrt{1-F_s}g(\mu)\mathbf{A}(t) + e^{i\delta_s}\sqrt{F_s}\frac{h(\mu)}{\sqrt{3}}\mathbf{B}(t) \tag{1.64}$$

leaves the probability density invariant (in a parameter space now enlarged to include μ_ϕ and Γ_ϕ); however now, complex conjugation of the term $g(\mu)$, Eq. (1.47), implies that the transformation $\Gamma_\phi \to -\Gamma_\phi$ should also be carried out, in addition to the transformation of β_s, $\Delta\Gamma$, $\delta_\parallel$, and $\delta_\perp$. Since negative values of Γ_ϕ are physically meaningless, this transformation is not an admissible symmetry.

However we can find a symmetry transformation that carries one set of physically meaningful parameters into another. Such a symmetry is the transformation of the terms in Eq. (1.64) into their *negative complex conjugate*. This transformation is equivalent to the combined transformation already described, in addition to:

$$\begin{aligned} \delta_s &\to \pi - \delta_s \\ (\mu - \mu_\phi) &\to -(\mu - \mu_\phi)\,. \end{aligned} \tag{1.65}$$

The latter transformation carries us from a point on one side of the ϕ mass peak to another point located symmetrically on the other side. This symmetry is useful when considering likelihood functions in which the dependence on μ is integrated out. If we integrate symmetrically about the ϕ mass peak, we can consider the contribution to the integral coming from a slice in ϕ mass on one hand and the a symmetrically-located slice in ϕ mass on the other hand. While the contribution of either slice is not invariant to the transformation above, the contribution of both slices certainly is, and the combined transformation:

$$\begin{aligned} \beta_s &\rightarrow \pi/2 - \beta_s \\ \Delta\Gamma &\rightarrow -\Delta\Gamma \\ \delta_\perp &\rightarrow \pi - \delta_\perp \\ \delta_\parallel &\rightarrow 2\pi - \delta_\parallel \\ \delta_s &\rightarrow \pi - \delta_s \end{aligned} \tag{1.66}$$

is again a symmetry of the integrated likelihood. We note, however, that this symmetry requires the symmetry of the nonrelativistic ϕ-propagator, Eq. (1.47), and applies to *the likelihood integrated over a finite symmetric interval of integration.*

Symmetries of the likelihood function for $B_s^0 \rightarrow J/\psi\phi$, in the presence of S-wave contribution for a fixed value of $\mu = m(K^+K^-)$ were discussed in a recent publication [77]. These formula can also be used to fit for data falling within a narrow window in μ. Under those assumptions we can drop the ϕ propagator from the expression in Eq. (1.64), absorb the Breit-Wigner terms into the amplitudes $\mathbf{A}(t)$, and write our model for the rates as

$$\sqrt{1-F_s}\mathbf{A}(t) + e^{i\delta_s}\sqrt{\frac{F_s}{3}}\mathbf{B}(t)\,. \tag{1.67}$$

Then one can see that the transformation in which $\delta_s \rightarrow -\delta_s$ replaces $\delta_s \rightarrow \pi - \delta_s$ in Eq. 1.66 accomplishes a complex conjugation of the terms in Eq. 1.67 and is a symmetry *of the likelihood at fixed* μ.

In the more general case one can notice from Eqs. 1.61 and 1.62 that the probability densities integrated over μ are invariant to complex conjugation of

both $\mathbf{A}(t)$ *and* the overlap integral $\mathcal{I}_{mu}$ of Eq. 1.52. This can be accomplished with a more complicated adjustment of δ_s. With a nonrelativistic Breit Wigner the required transformation is

$$\delta_s \to 2\delta_{BW} - \delta_s$$

where $\delta_{BW} \equiv \arg\left(\log\left((\mu_{hi} - \mu_\phi + i\Gamma_\phi/2)/(\mu_{lo} - \mu_\phi + i\Gamma_\phi/2)\right)\right)$. The phase δ_{BW} reduces to $\delta_{BW} = 0$ in the limit of an infinitesimally thin interval in μ, and to $\delta_{BW} = -\pi/2$ in the limit of a finite symmetric interval. This demonstrates real differences in the two formulations, and underscores the need for caution when applying the formulae of Ref. [77] to a finite interval in $\mu = m(K^+K^-)$.

CHAPTER

TWO

ACCELERATOR MACHINE

The **Tevatron** is a circular particle accelerator in the United States, at the Fermi National Accelerator Laboratory (Fermilab), just east of Batavia, Illinois, and is the second highest energy particle collider in the world after the Large Hadron Collider (LHC). The Tevatron is a synchrotron that accelerates protons and antiprotons in a 6.28 km ring to energies of up to 1 TeV, hence its name.

The acceleration occurs in a number of stages. The first stage is the 750 keV Cockcroft-Walton pre-accelerator, which ionizes hydrogen gas and accelerates the negative ions created using a positive voltage. The ions then pass into the 150 m long linear accelerator (*linac*) which uses radio frequency cavities to accelerate the ions to 400 MeV. The ions then pass through a carbon foil, to remove the electrons, and the charged protons then move into the Booster.

The Booster is a small circular synchrotron, around which the protons pass up to 20,000 times to attain an energy of around 8 GeV. From the Booster the particles pass into the Main Injector, which perform a number of tasks. It can accelerate protons up to 150 GeV; it can produce 120 GeV protons for antiproton creation; it can increase antiproton energy to 120 GeV and it can inject protons or antiprotons into the Tevatron. The antiprotons are created by the Antiproton Source. 120 GeV protons are collided with a nickel target producing a range of particles including antiprotons which can be collected and stored in the accumulator ring. The ring can then pass the antiprotons to the Main Injector.

The Tevatron can accelerate particles from the Main Injector up to 980 GeV. The protons and antiprotons are accelerated in opposite directions, crossing paths in the CDF and DØ detectors to collide at 1.96 TeV. To hold the particles on track the Tevatron uses 774 niobium-titanium superconducting dipole magnets cooled in liquid helium producing 4.2 T field. The field ramps over about 20 s as the particles are accelerated. Another 240 NbTi quadrupole magnets are used to focus the beam. The initial design luminosity of the Tevatron was 10^{30} cm^{-2} s^{-1}, however the accelerator has following upgrades been able to deliver luminosities up to 4×10^{32} cm^{-2} s^{-1}.

A detailed description of the accelerators found in the Fermilab complex is given in the following sections.

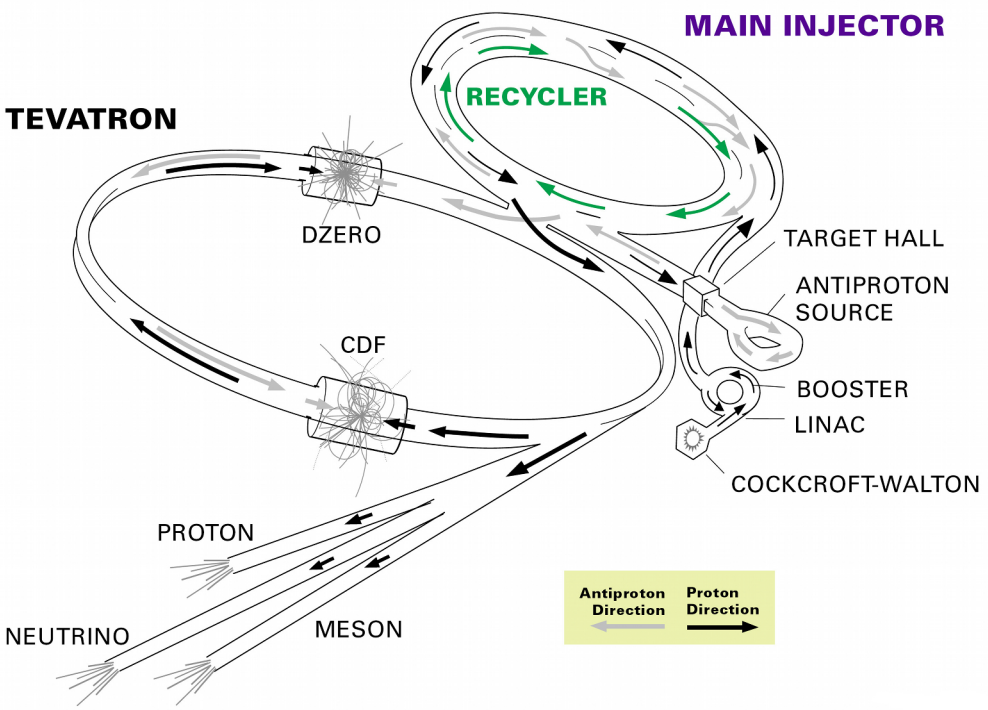


Figure 2.1: Fermilab's accelerator chain

2.1 Pre-accelerator

The Pre-accelerator, or *Preacc*, is really the first accelerator. It is the source of negatively charged hydrogen ions accelerated by the linear accelerator. The Preacc consists of the source housed in an electrically charged dome. The source converts hydrogen gas to ionized hydrogen gas (H^-). The dome is charged to a potential of -750 kV. The ionized gas is allowed to accelerate though a column from the charged dome to the grounded wall to an energy of 750 KeV. The Preacc accelerates beam every 66 milliseconds (a 15 Hz repetition rate) whether beam is being requested or not. After beam exits the accelerating column, it travels trough a transfer line called the 750 KeV line (referring tho the transported beam's kinetic energy) and then enters the Linac.

2.2 Linac

The Linear accelerator or *Linac* is the next level of acceleration for the negatively charged hydrogen ions. It takes the ions with an energy of 750 KeV and accelerates

them to an energy of 400 MeV. The Linac consists of two main sections, the low energy drift tube Linac and the high-energy side coupled cavity Linac. The drift tube Linac (DTL) makes up the first five radio frequency (RF) stations. A large power amplifier tube (PA) powers each drift tube. These tubes amplify the 201 MHz RF signal used to drive the low energy cavities and accelerate the beam. The last 7 RF stations use Klystron amplifiers instead of the outdated tube technology of the low energy end. The Klystron amplify an 850 MHz RF signal that is then fed into a series of side coupled cavity Linac (SCL) modules. The difference in these two resonating frequencies is a result of the differences in geometry between the DTL cavities and the SCL modules. In the DTL, every RF cycle is used to accelerate beam; in the SCL, only every fourth cycle is used (805 MHz = 4×201 MHz). Between the low energy DTL and the high energy SCL there is a "transition section", made up of the bunchier and the vernier. These two additional Klystron stations are used to ease the change in accelerating structure and RF frequency and improve the efficiency of the transfer. The Linac can accelerate beam once every 66 milliseconds (a 15Hz repetition rate). Beam in the DTL is focused by means of quadrupole magnets located inside the drift tubes, which in turn are located inside the RF cavities. The beam traveling through the SCL is focused by quadrupoles placed between the accelerating modules (outside of the accelerating cavities). After beam is accelerated in the Linac, the 400 MeV H^- ions are sent to the 400 MeV line, a transfer line which connects the Linac to the Booster.

2.3 Booster

Booster is the next level of acceleration. It takes the 400 MeV negative hydrogen ions from the Linac and strips the electrons off, which leaves only the proton, and accelerating the protons to 8 GeV. The Booster is the first circular accelerator, or synchrotron, in the chain of accelerators. It consists of a series of magnets arranged around a 75 meter radius circle, with 19 RF cavities interspersed. The accelerated proton beam in the Booster is then sent to the MI-8 line (a transfer

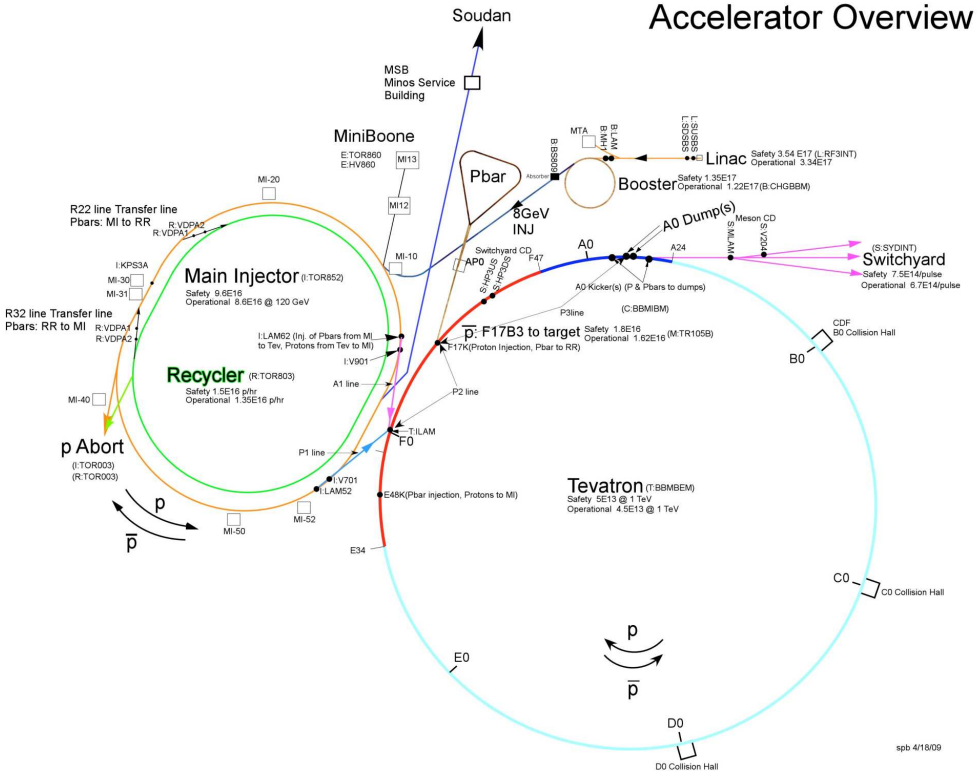


Figure 2.2: Detailed schematic of the accelerators in the Fermilab complex.

line from the Booster to the Main Injector). Booster can accelerate beam once every 66 milliseconds (15 Hz).

2.4 Main Injector

The Main Injector (MI) is a circular synchrotron seven times the circumference of the Booster and slightly more than half the circumference of the Tevatron. The ring is divided up into 6 sections, or sectors, labeled MI-10 trough MI-60. MI-60 is the region adjacent to the Tevatron. Main Injector has 18 accelerating cavities. It can accelerate 8 GeV protons from the Booster to either 120 GeV or 150 GeV, depending on their destination. When used to stack antiprotons the final energy is 120 GeV. When used to inject into the Tevatron, the final beam energy is 150 GeV. As well as accepting protons from Booster, the Main Injector can accept antiprotons from the Antiproton Source. The Main Injector can accelerate beam as fast as every 2.2 seconds. The Main Injector can operate

in different modes, the most important for colliding beams at Tevatron are:

- **Antiproton Production**

  The most commonly used method is where two Booster batches are injected into the Main Injector and then "slipped" together. This method is called "slip-stacking" as the first injected batch is slowed down using the RF, and the second batch is injected and as it slips by the first batch they are merged together into one.

- **Shot Setup**

  This mode relates to the act of extracting antiprotons from the Recycler, commonly referred as a shot. Shot setup is the time before actually transferring the antiprotons when the various transfer lines are tuned up with protons to ensure efficient antiproton transmission. When loading the Tevatron with protons, 7 bunches are injected from booster and accelerated to 150 GeV. A process called coalescing makes one bunch out of the 7 originals, and this coalesced bunch is extracted at MI-52 and travels down the P1 line and into the Tevatron. By repeating this process 36 times in a row you load the protons necessary for a 36x36 store. When loading antiprotons, 4 bunches are extracted from the Recycler, accelerated to 150 GeV in the MI, and extracted at MI-62, sent down the A1 line and into the Tevatron. This process is repeated 9 times to give a total of 36 antiproton bunches.

2.5 Tevatron

The Tevatron is the largest of the Fermilab accelerators, with a circumference of approximately 4 miles. It is a circular synchrotron with eight accelerating cavities. The Tevatron can accept both protons and antiprotons from Main Injector and accelerate them from 150 GeV to 980 GeV. In Collider Mode, the Tevatron can store beam for hours at a time. Because the Tevatron is a primarily storage ring, the length of time between acceleration cycles is widely variable.

The Tevatron is the only cryogenically cooled accelerator at Fermilab. The

magnets used in the Tevatron are made up of a superconducting niobium-titanium alloy that needs to be kept extremely cold ($\tilde{4}$K) to remain a superconductor. The benefit of having superconducting magnets is the increased magnetic fields possible when high currents can be run through thin wires without fear of damage related to excessive resistive heating. This low operating temperature is responsible for the Tevatron's extensive cryogenic "plumbing" and unique magnet protection systems.

The Tevatron is not a perfect circle either. The ring is divided into six sectors labeled A through F. Each sector has five service buildings, a "0" building and "1" through "4" buildings. Each "0" location contains a large straight section, and each such straight section has a special function. The A0 straight section is where the Tevatron tunnel connects to the Switchyard. It is also the location of the beam abort for the Tevatron (the colliding beams abort). The CDF collision hall is located at the B0 straight section, while C0 is unused. The D0 experiment is named for the location it occupies in the tunnel, while E0's only claim to fame is that it was the site of the transfer line from the old Main Ring to the Tevatron. Perhaps the busiest section of tunnel at this laboratory is located at F0. This is where Tevatron RF cavities are located, as well as the connection points of both the P1 and A1 transfer lines from Main Injector. The P2 transfer line also passes trough, carrying beam to and from the Antiproton Source and protons on their way out to the Switchyard.

As mentioned above, the primary purpose of the Tevatron is to act as a storage ring where protons and antiprotons can collide with each other and produce interesting secondary particles. When operating in Collider mode, protons and antiprotons are injected at 150 GeV and then accelerated to 980 GeV. Once the final energy is reached, the two counter-rotating particle beams pass through each other for hours at time (or until some component failure causes the beam to be lost). This stable situation of 980 GeV proton and antiproton collisions is called a Store. After the number of collisions per second (described by the luminosity of the store) drops to low to be useful for the experimenters, the store is ended and the Tevatron prepared for a new store.

2.6 Antiproton Source

- **Target**

  The antiproton target station is not an accelerator. It is added here to maintain some continuity in the discussion of the antiproton source. The target station is found at the end of the AP1 transfer line, and is located beneath the AP0 service building. When we are collecting antiprotons (stacking), 120 GeV protons coming from the MI trough P1→P2→AP1 lines strike a nickel alloy target. These high-energy protons striking the target produce a spray of all sorts of secondary particles. Using magnets to choose wich momentum and charge we can collect 8 GeV antiprotons from this spray. These antiprotons are directed down the AP2 transfer line and into the Debuncher.

- **Debuncher**

  The Debuncher is one of the two synchrotrons that make up the Antiproton Source (commonly refereed to as the Pbar Source). The Debuncher is rounded triangular-shaped synchrotron with a mean radius of 90 meters. It can accept 8 GeV protons from Main Injector for beam studies, and 8 GeV antiprotons from the target station.

  Its primary purpose is to efficiently capture the high momentum spread antiprotons coming off the target, using RF manipulation called bunch rotation. There are also beam-cooling systems that act to make the beam more manageable. This so-called *stochastic cooling* is accomplished by picking up a signal from the circulating antiprotons on one side of the ring, amplifying the signal, and then applying that signal to the antiproton beam at another part of the ring. There are three cooling system sin the Debuncher: a momentum system, and two transverse systems (horizontal and vertical).

  The Debuncher does not 'accelerate' beam in the same sense as the other accelerators, but maintain the beam at a constant energy of 8 GeV. The antiproton beam can be transferred to the Accumulator via the D/A transfer

line, located beneath the AP10 service building.

- **Accumulator**

  The accumulator is the second synchrotron of the antiproton source, It is also a triangular-shaped synchrotron of radius 75 meters and is housed in the same tunnel as the Debuncher. It is the storage ring for the antiprotons; all of the antiprotons made are stored here at 8 GeV and cooled until transferred to the Recycler.

  The Accumulator has a number of different cooling systems: stacktail momentum, core momentum, and core transverse (horizontal and vertical). An 8 GeV antiproton beam can be extracted from the Accumulator and sent down the AP3 transfer line (which eventually mets up the AP1 line) towards the MI.

The Antiproton Source can operate in three different modes:

- **Antiproton Production (Stacking)**

  120 GeV beam extracted from MI at MI-52 travels down the P1→P2→AP1 lines until striking the nickel target beneath the AP0 service building. Out of the spray of random secondary particles, 8 GeV antiprotons are taken down the AP2 line and into the Debuncher.

  The Debuncher cools the Pbar and then transfers them down the D/A line and into the Accumulator. The Accumulator further cools the beam and stores it until it's needed.

- **Shots**

  During shot setup, 4 bunches of antiprotons are extracted from the Accumulator and sent down the AP3→AP1→P2→P1 lines into MI, and into the Recycler.

- **Reverse Protons**

  This is a mode used during shot setup to tune up the transfer lines. This mode is also used during studies period.

2.7 Recycler

The Recycler is an antiproton storage ring located along the ceiling of the Main Injector tunnel. The proposed purpose of the Recycler was to "Recycle" the antiprotons from a Tevatron store, cooling them and storing them alongside those sent from the Antiproton Source. This was abandoned after early problems in Collider Run II.

The Recycler now accepts transfers only from the Antiproton source and cools them further than the Pbar Accumulator is capable. The Recycler uses both a stochastic cooling system (like the Antiproton Source) and electron cooling system. Stochastic cooling is used to cool the beam in Recycler, but loses it effectiveness with higher intensities. Once above 200×10^{30} antiprotons in the Recycler, Electron Cooling is Required.

Electron cooling works on the principle of momentum transfer between electrons and antiprotons. A highly concentrated, cool beam of electrons is driven at the same energy as the antiprotons, and laid on the top of the antiprotons. The resulting glancing collisions between electrons and antiprotons transfer some of the momentum from the "hot" antiprotons to the "cool" electrons. With enough electrons, a substantial longitudinal cooling force is produced by absorbing momenta from the antiprotons, allowing for more compact, brighter bunches to send to the Tevatron. These electrons are produced in a 5 MeV Pelletron, and guided through beamlines to a section of Recycler beam pipe in the RR-30 section. Once the electron beam has made its pass through the antiprotons, it is returned to the Pelletron to recover the charge.

- **Stashing**

  The process of accepting Pbars from the Pbar source and cooling them to prepare for more transfers, and eventually once the stash is large enough to begin HEP shot setup.

- **Studies**

8 GeV protons from the Booster are sent into the Main Injector first (but not accelerated) and then transferred to the Recycler. This is done as protons are easier to make, and less costly to lose for studies. This is typically done as a *check out* of the Recycler systems after a Main Injector enclosure access.

- **Shot setup**

  After cooling the *stash*, the antiprotons are "mined" into 9 "parcels". Each of the parcels is split into four bunches, which are extracted to the Tevatron, after acceleration in MI.

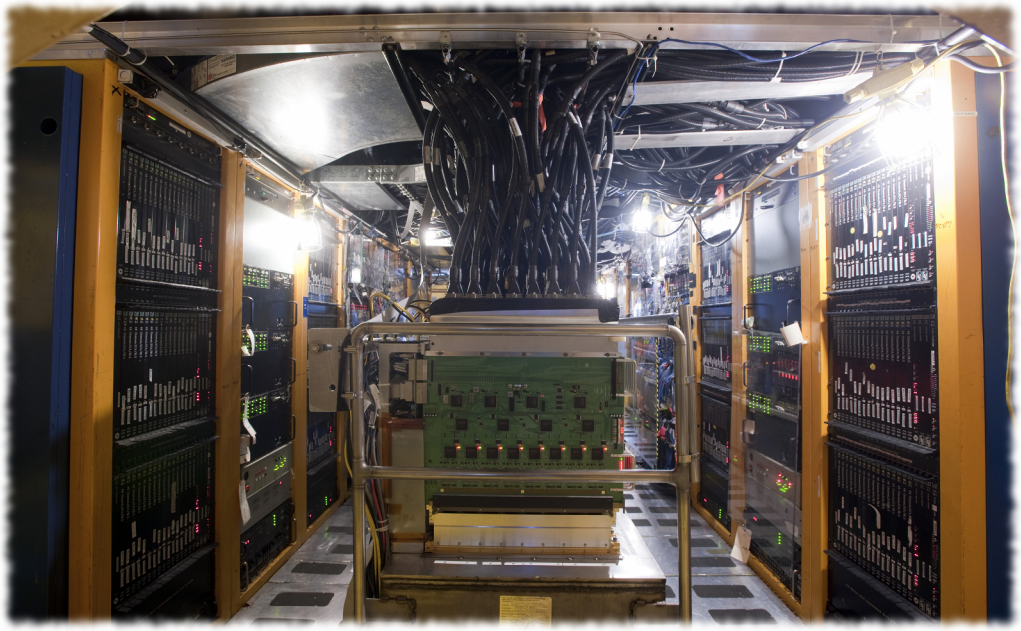
2.8 Switchyard 120 GeV

The Switchyard is not really an accelerator, but rather a complex crossroads where 120 GeV beam coming from the MI could have been directed to a number of final destinations in the main fixed target beamlines: the Proton line, the Meson line, and the Neutrino line.

CHAPTER

THREE

DØ DETECTOR



The DØ experiment was proposed in 1983 to study proton-antiproton collisions at a center-of-mass energy of 1.8 TeV at the Fermilab Tevatron collider. The focus of the experiment was the study of high mass states and large p_T phenomena. The detector performed very well during Run I and Run II of the Tevatron, 1992–1996 and 2001–2011, leading to the discovery of the top quark and measurement of its mass, a precision measurement of the mass of the W boson, detailed analysis of gauge boson couplings, studies of jet production, and greatly improved limits on the production of new phenomena such as leptoquarks, and supersymmetric particles, among many other accomplishments [1].

During Run I, the Tevatron operated using six bunches each of protons and antiprotons, with 3500 ns between bunch crossings and a center-of-mass energy of 1.8 TeV. The peak luminosity was typically 1–2$\times 10^{31}$ cm^{-2}s^{-1} and approximately 120 pb^{-1} of data were recorded by DØ. Following the completion of the new Main Injector and associated Tevatron upgrades [66, 2], the collider began running again in 2001. In Run II, which began in March 2001 and ended in September 2011, the Tevatron was operated with 36 bunches of protons and antiprotons with a bunch spacing of 396 ns and at an increased center-of-mass energy of 1.96 TeV. The instantaneous luminosity increased by more than a factor of one hundred to greater than 10^{33} cm^{-2}s^{-1} and 10.7 fb^{-1} of data were recorded. To take advantage of these improvements in the Tevatron and to enhance the physics reach of the experiment, the DØ detector was significantly upgraded after the Run I.

The Run II detector consisted of three major subsystems: central tracking detectors, uranium/liquid-argon calorimeters, and a muon spectrometer. A side view of the Run II DØ detector is shown in Figure 3.1.

The following sections describe the design and performance of the Run II DØ detector subsystems used in this analysis. For a full description of the detector see [13].

In the detector description and data analysis, we use a right-handed coordinate system in which the z-axis is along the proton direction and the y-axis is upward (Figure 3.1). The angles ϕ and θ are the azimuthal and polar angles, respectively. The r coordinate denotes the perpendicular distance from the z

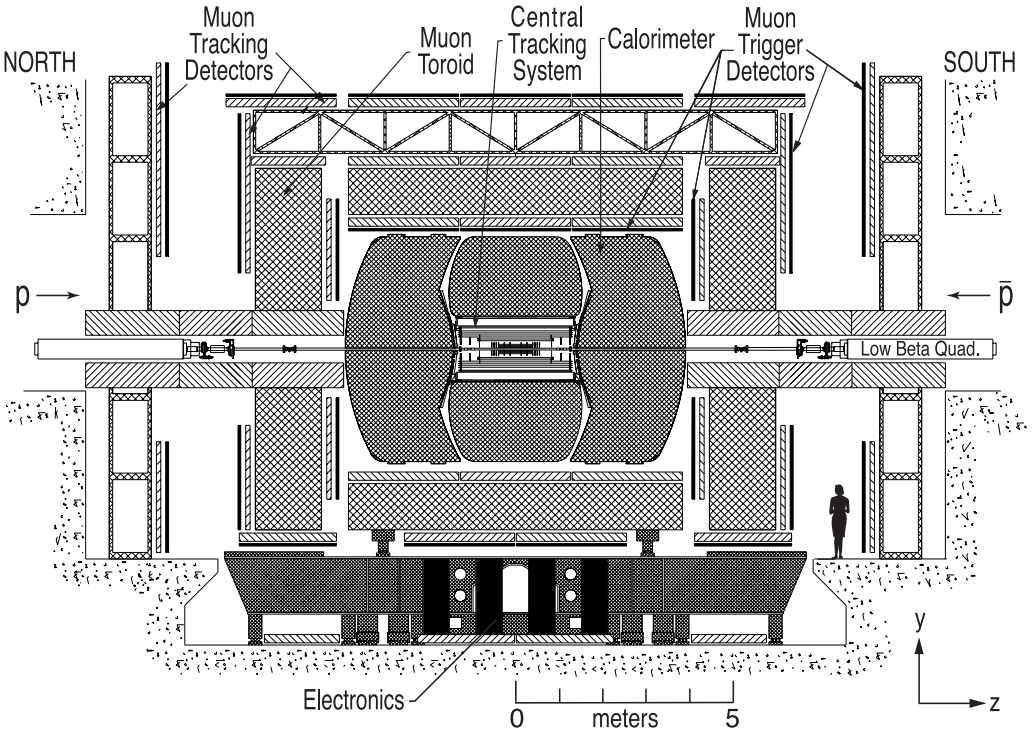


Figure 3.1: Diagram of the Run II DØ detector, as installed in the collision hall and viewed from inside the Tevatron ring. The detectors in the central region of the detector are shown in Figure 3.2.

axis. The pseudorapidity, $\eta = -\ln[\tan(\theta/2)]$, approximates the true rapidity, $y = 1/2\ln[(E + p_z c)/(E - p_z c)]$, for finite angles in the limit that $(mc^2/E) \to 0$. We use the term "forward" to describe the regions at large $|\eta|$.

3.1 Central tracking

Excellent tracking in the central region is necessary for studies of top quark, electroweak, and b physics and to search for new phenomena, including the Higgs boson. The central tracking system consists of the silicon microstrip tracker (SMT) and the central fiber tracker (CFT) surrounded by a solenoidal magnet. It surrounds the DØ beryllium beam pipe, which has a wall thickness of 0.508 mm and an outer diameter of 38.1 mm, and is 2.37 m long. The two tracking detectors locate the primary interaction vertex with a resolution of about 35 μm along the beamline. They can tag b-quark jets with an impact parameter resolution of better than 15 μm in $r - \phi$ for particles with transverse momentum $p_T >$ 10 GeV/c at $|\eta| = 0$. The high resolution of the vertex position allows good measurement of lepton p_T, jet transverse energy (E_T), and missing transverse

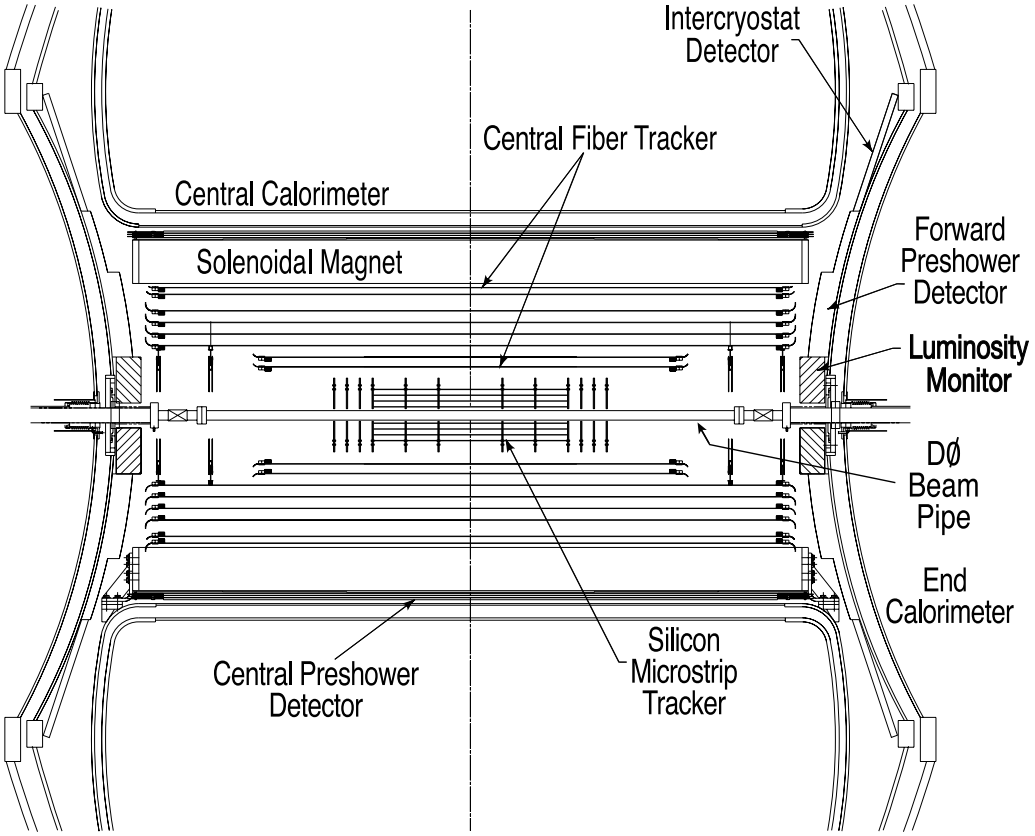


Figure 3.2: Cross-sectional view of the new central tracking system in the $x - z$ plane. Also shown are the locations of the solenoid, the preshower detectors, luminosity monitor, and the calorimeters.

energy $\not\!\!E_T$. Calibration of the electromagnetic calorimeter using E/p for electrons is also possible.

Both the SMT and CFT provide tracking information to the trigger. The SMT provides signals to the Level 2 and 3 trigger systems and is used to trigger on displaced vertices from b-quark decay. The CFT provides a fast and continuous readout of discriminator signals to the Level 1 trigger system; upon a Level 1 trigger accept, track information based on these signals is sent to Level 2. The Level 3 trigger receives a slower readout of the CFT's digitized analog signals, in addition to the discriminator information available at Level 1 and Level 2.

A schematic view of the central tracking system is shown in Figure 3.2.

3.1.1 Silicon microstrip tracker

The SMT provides both tracking and vertexing over nearly the full η coverage of the calorimeter and muon systems. Design of the detector, electronics, and

cooling are, in large part, dictated by the accelerator environment. The length of the interaction region ($\sigma \approx 25$ cm) sets the length scale of the device. With a long interaction region, it is a challenge to deploy detectors such that the tracks are generally perpendicular to detector surfaces for all η. This led us to a design of barrel modules interspersed with disks in the center and assemblies of disks in the forward regions. The barrel detectors primarily measure the $r-\phi$ coordinate and the disk detectors measure $r-z$ as well as $r-\phi$. Thus vertices for particles at high η are reconstructed in three dimensions by the disks, and vertices of particles at small values of η are measured in the barrels and central fiber tracker.

An isometric view of the SMT is shown in Figure 3.3. The detector has six barrels in the central region. Each barrel has four silicon readout layers. The silicon modules installed in the barrels are called "ladders." Layers 1 and 2 have twelve ladders each; layers 3 and 4 have twenty-four ladders each, for a total of 432 ladders. Each barrel is capped at high $|z|$ with a disk of twelve double-sided wedge detectors, called an "F-disk." Forward of the three disk/barrel assemblies on each side is a unit consisting of three F-disks. In the far forward regions, two large-diameter disks, "H-disks," provide tracking at high $|\eta|$. Twenty-four full wedges, each consisting of two back-to-back single-sided "half" wedges, are mounted on each H-disk. There are 144 F-wedges and 96 full H-wedges in the tracker; each side of a wedge (upstream and downstream) is read out independently. There is a grand total of 912 readout modules, with 792,576 channels. The centers of the H-disks are located at $|z| = 100.4,\ 121.0$ cm; the F-disks are at $|z| = 12.5,\ 25.3,\ 38.2,\ 43.1,\ 48.1$, and 53.1 cm. The centers of the barrels are at $|z| = 6.2,\ 19.0,\ 31.8$ cm. The SMT is read out by custom-made 128-channel SVXIIe readout chips.

3.1.2 Central fiber tracker

The CFT consists of scintillating fibers mounted on eight concentric support cylinders and occupies the radial space from 20 to 52 cm from the center of the beampipe. To accommodate the forward SMT H-disks, the two innermost

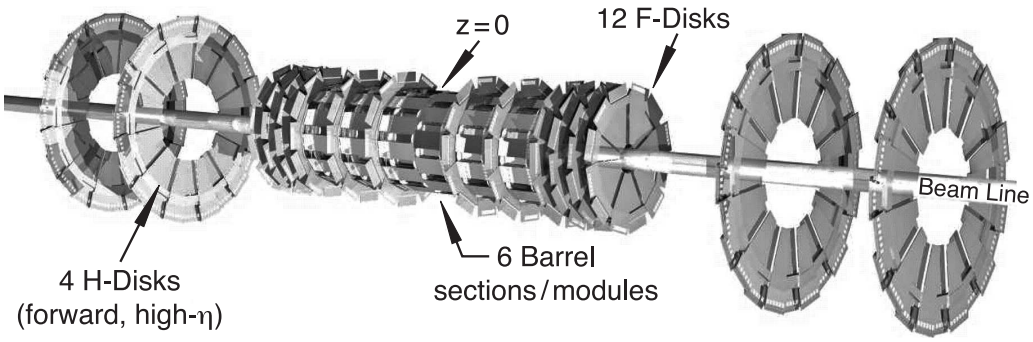


Figure 3.3: The disk/barrel design of the silicon microstrip tracker.

cylinders are 1.66 m long; the outer six cylinders are 2.52 m long. The outer cylinder provides coverage for $|\eta| \lesssim 1.7$. Each cylinder supports one doublet layer of fibers oriented along the beam direction (z) and a second doublet layer at a stereo angle in ϕ of $+3^\circ$ (u) or -3° (v). Doublet layers with fibers oriented along the beam axis are referred to as axial layers, while the doublet layers oriented at small angles are referred to as stereo layers. From the smallest cylinder outward, the fiber doublet orientation is $zu{-}zv{-}zu{-}zv{-}zu{-}zv{-}zu{-}zv$. The scintillating fibers are coupled to clear fiber waveguides which carry the scintillation light to visible light photon counters (VLPCs) for read out. The small fiber diameter (835 μm) gives the CFT an inherent doublet layer resolution of about 100 μm as long as the location of the individual fibers is known to better than 50 μm.

Discriminator signals from the axial doublet layers are used to form a fast Level 1 hardware trigger based upon the number of track candidates above specified p_T thresholds (with a minimum threshold of 1.5 GeV/c). Level 1 track candidates are used by the Level 2 trigger, while the Level 3 trigger uses the full CFT readout information.

3.2 Muon system

For muon triggering and measurement, the upgraded detector uses the original central muon system proportional drift tubes (PDTs) and toroidal magnets [72], central scintillation counters (some new and some installed during Run I), and a completely new forward muon system. The central muon system provides coverage for $|\eta| \lesssim 1.0$. The new forward muon system extends muon detection to

$|\eta| \approx 2.0$, uses mini drift tubes (MDTs) rather than PDTs, and includes trigger scintillation counters and beam pipe shielding. The small angle muon system [72] of the original detector, including its associated magnets, has been removed.

During Run I, a set of scintillation counters, the cosmic cap [70], was installed on the top and upper sides of the outer layer of central muon PDTs. This coverage has been extended to the lower sides and bottom of the detector, to form the cosmic bottom. These trigger scintillation counters are fast enough to allow us to associate a muon in a PDT with the appropriate bunch crossing and to reduce the cosmic ray background. Additional scintillation counters, the $A\phi$ counters, have been installed on the PDTs mounted between the calorimeter and the toroidal magnet. The $A\phi$ counters provide a fast detector for triggering and identifying muons and for rejecting out-of-time background events.

The scintillation counters are used for triggering; the wire chambers are used for precise coordinate measurements as well as for triggering. Both types of detectors contribute to background rejection: the scintillator with timing information and the wire chambers with track segments.

Exploded views of the muon system are shown in Figures 3.4 and 3.5.

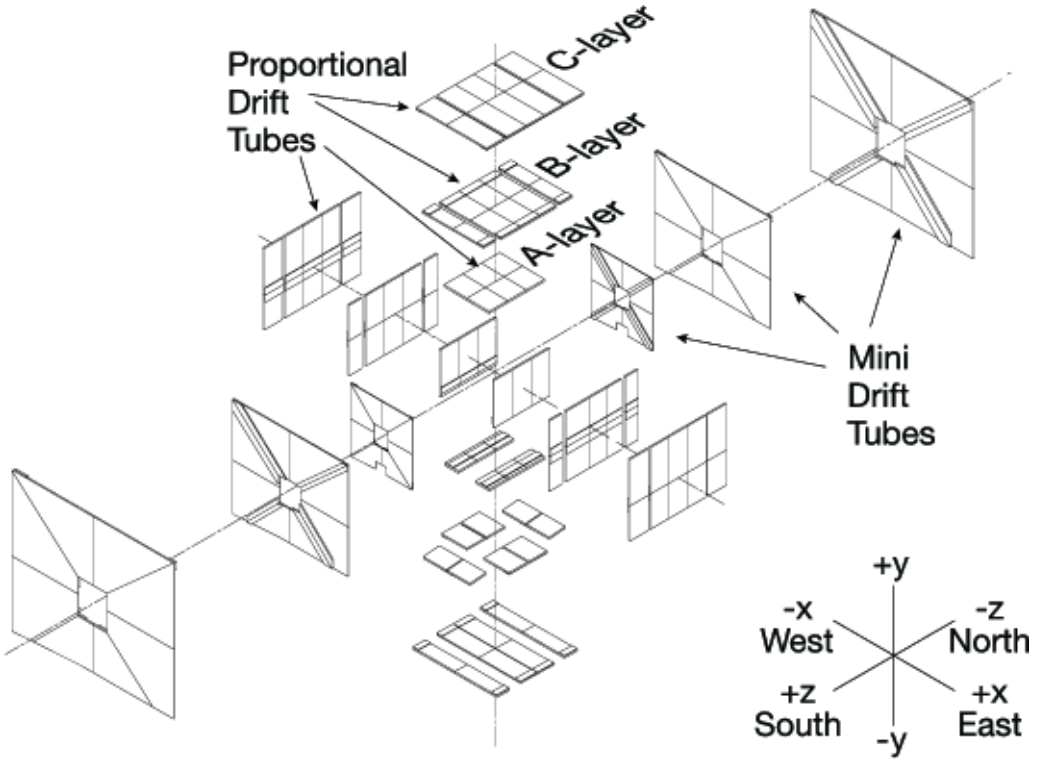


Figure 3.4: Exploded view of the muon wire chambers.

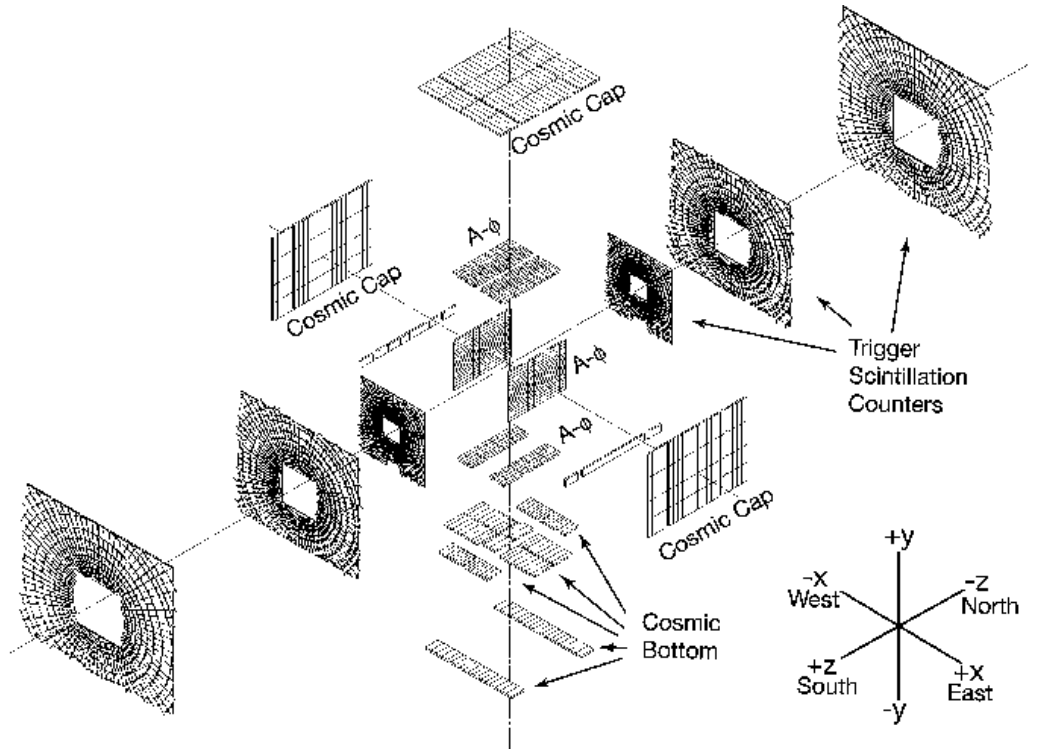
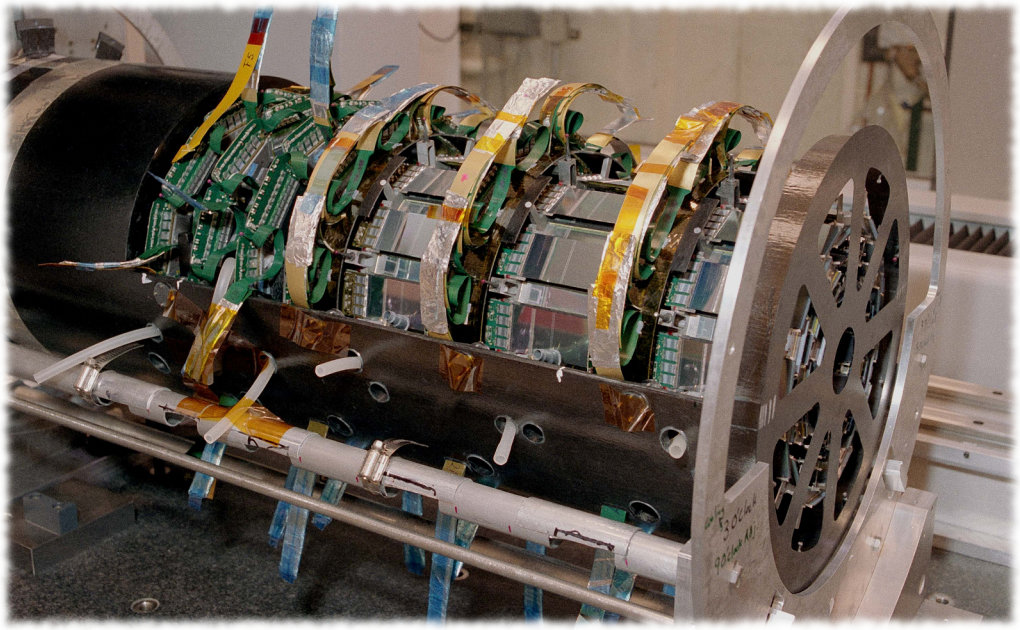


Figure 3.5: Exploded view of the muon scintillation detectors.

CHAPTER

FOUR

MEASUREMENT OF CP VIOLATION IN $B_S \to J/\psi\phi$



4.1 Data Sample and Event Reconstruction

The analysis presented here is based on 8 fb^{-1} proton-antiproton collision data accumulated during RunII between February 2002 and June 2010. Events are collected with a mixture of single- and dimuon triggers. Some triggers require track displacement with respect to the primary vertex (large track impact parameter). Since this condition biases the B_s^0 lifetime measurement, the events selected exclusively by these triggers are removed from our sample.

Candidate $B_s^0 \to J/\psi\phi$, $J/\psi \to \mu^+\mu^-$, $\phi \to K^+K^-$ events are required to include two opposite charge muons accompanied by two opposite charge tracks. Both muons are required to be detected in the muon chambers inside the toroid magnet, and at least one of the muons is required to be also detected outside the toroid. Each of the four final-state tracks is required to have at least one SMT hit.

To form B_s^0 candidates, muon pairs in the invariant mass range 3.096 ± 0.350 GeV, consistent with J/ψ decay, are combined with pairs of opposite charge tracks (assigned the kaon mass) consistent with production at a common vertex, and with an invariant mass in the range 1.019 ± 0.030 GeV. A kinematic fit under the B_s^0 decay hypothesis constrains the dimuon invariant mass to the world-average J/ψ mass [26] and constrains the four-track system to a common vertex.

Trajectories of the four B_s^0 decay products are adjusted according to the decay-vertex kinematic fit. The re-adjusted track parameters are used in the calculation of the B_s^0 candidate mass and decay time, and of the three angular variables characterising the decay as defined later. B_s^0 candidates are required to have an invariant mass in the range 5.37 ± 0.20 GeV. In events where multiple candidates satisfy these requirements, we select the candidate with the best decay vertex fit probability.

To reconstruct the primary vertex (PV), we select tracks that do not originate from the candidate B_s^0 decay, and apply a constraint to the average beam-spot position in the transverse plane. We define the signed decay length of a B_s^0 meson,

L_{xy}^B, as the vector pointing from the PV to the decay vertex, projected on the B_s^0 transverse momentum p_T. The proper decay time of a B_s^0 candidate is given by $t = M_{B_s}\vec{L}_{xy}^B \cdot \vec{p}/(p_T^2)$ where M_{B_s} is the world-average B_s^0 mass [26], and $\vec{p}$ is the particle momentum. The distance in the beam direction between the PV and the B_s^0 vertex is required to be less than 5 cm. Approximately 5 million events are accepted after the selection described in this section.

4.2 Background Suppression

The selection criteria are designed to optimize the measurement of $\phi_s^{J/\psi\phi}$ and $\Delta\Gamma_s$. Most of the background is due to directly produced J/ψ mesons accompanied by tracks arising from hadronization. This "prompt" background is distinguished from the "non-prompt", or "inclusive $B \to J/\psi + X$" background, where the J/ψ meson is a product of a b-hadron decay while the tracks forming the ϕ candidate emanate from a multi-body decay of a b hadron or from hadronization. Two different event selection approaches are used, one based on a multi-variate technique, and one based on simple limits on kinematic and event quality parameters.

4.2.1 Signal and background simulation

Three Monte Carlo simulated samples are used to study background suppression: signal, prompt background, and non-prompt background. All three are generated with PYTHIA [61]. Hadronization is also done in PYTHIA, but all hadrons carrying heavy flavors are passed on to EvtGen [53] to model their decays. The prompt background MC sample consists of $J/\psi \to \mu^+\mu^-$ decays produced in $gg \to J/\psi g$, $gg \to J/\psi\gamma$, and $g\gamma \to J/\psi g$ processes. The signal and non-prompt background samples are generated from primary $b\bar{b}$ pair production with all b hadrons being produced inclusively and the J/ψ mesons forced into $\mu^+\mu^-$ decays. For the signal sample, events with a B_s^0 are selected, their decays to $J/\psi\phi$ are implemented without mixing and with uniform angular distributions, and the B_s^0 mean lifetime is set to $\overline{\tau}_s = 1.464$ ps. There are approximately 10^6 events in each background and the signal MC samples. All events are passed through a full GEANT-based [4]

detector simulation. To take into account the effects of multiple interactions at high luminosity, hits from randomly triggered $p\bar{p}$ collisions are overlayed on the digitized hits from MC. These events are reconstructed with the same program as used for data. The three samples are corrected so that the p_T distributions of the final state particles in $B_s^0 \to J/\psi\phi$ decays match those in data (see Section 4.3).

4.2.2 Multivariate event selection

To discriminate the signal from background events, we use the TMVA package [69]. In preliminary studies using MC simulation, the Boosted Decision Tree (BDT) algorithm was found to demonstrate the best performance. Since prompt and non-prompt backgrounds have different kinematic behavior, we train two discriminants, one for each type of background. We use a set of 33 variables for the prompt background and 35 variables for the non-prompt background. The variables and more details of the BDT method are given in Section 4.2.3.

The BDT training is performed using a subset of the MC samples, and the remaining events are used to test the training. The signal MC sample has about 84k events, the prompt background has 29k events, and the non-prompt background has 39k events. Figure 4.1 shows the BDT output discriminant for the prompt and non-prompt cases.

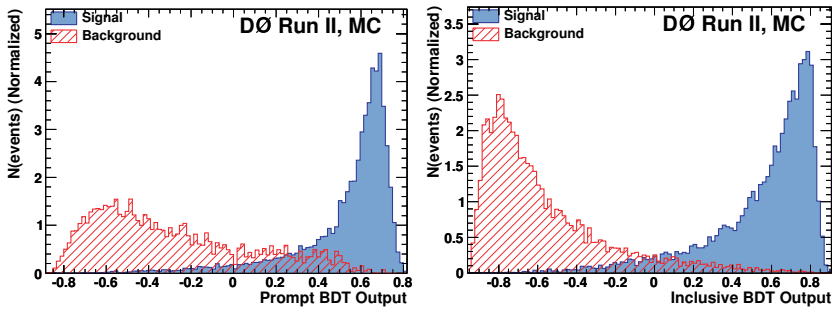


Figure 4.1: BDT discriminant output for the prompt (top) and non-prompt (bottom) classifiers. The signal and background events are taken from simulation. Events used for BDT training are excluded from these samples.

4.2.3 BDT Discriminants

Two BDT discriminants are used to reject background. One is trained to remove the prompt background (the "prompt BDT"), and the other is trained to remove inclusive B decays (the "inclusive BDT"). The prompt BDT uses 33 variables, listed in Table 4.1. The inclusive BDT uses 35 variables, listed in Table 4.2. In these tables, ΔR is defined as $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$, where η is the pseudorapidity and ϕ is the azimuthal angle. The term "uncorrected" refers to the correction due to the J/ψ mass constraint. "Leading" ("trailing") muon or kaon refers to the particle with larger (smaller) p_T, and dE/dx is the energy loss per unit path length of a charged particle as it traverses the silicon detector. Isolation is defined as $p(B)/\sum_{<\Delta R} p$ where $p(B)$ is the sum of the momenta of the four daughter particles of the B_s^0 candidate, and the sum is over all particles within a cone defined by ΔR, including the decay products of the B_s^0 candidate. The tables also show the importance and separation for each variable. The separation $\langle S^2 \rangle$ of a classifier y is defined as

$$\langle S^2 \rangle = \frac{1}{2} \int \frac{(\hat{y}_S(y) - \hat{y}_B(y))^2}{\hat{y}_S(y) + \hat{y}_B(y)} dy, \tag{4.1}$$

where y_S is the output of the discriminant function for signal events and y_B is the discriminant function for background. The importance of each BDT input variable is derived by counting in the training how often the variable is used to split decision tree nodes and by weighting each split occurrence by its separation gain squared and by the number of events in the node.

The distributions for the six most important variables in training on prompt J/ψ decays are shown in Figure 4.2. The distributions for the six most important variables in the training on inclusive $B \to J/\psi X$ decays are shown in Figure 4.3.

Figure 4.4 compares the shapes of the distributions of the three angular variables and the lifetime, before and after the BDT requirements. The figures show that the BDT requirements do not affect these differential distributions significantly.

Rank	Variable
1	KK invariant mass
2	Maximum ΔR between either K meson and the B_s^0 candidate
3	Isolation using the maximum ΔR between either K and the B_s^0
4	Uncorrected p_T of the B_s^0
5	Minimum ΔR between either K and the B_s^0
6	p_T of the trailing K meson
7	p_T of the ϕ meson
8	p_T of the leading K meson
9	Trailing muon momentum
10	p_T of the leading muon
11	Maximum ΔR between either muon and the B_s^0
12	Maximum χ^2 of either K meson with the J/ψ vertex
13	Dimuon invariant mass
14	Maximum χ^2 of either of the K candidate track
15	B_s^0 isolation using the larger K/B_s ΔR and tracks from the PV
16	p_T of the J/ψ meson
17	Minimum ΔR between either muon and the B_s^0 candidate
18	Trailing K momentum
19	χ^2 of the B_s^0 candidate vertex
20	B_s^0 isolation using $\Delta R < 0.75$
21	Minimum χ^2 of the J/ψ vertex with either K
22	p_T of the trailing muon
23	Minimum of the χ^2 of the J/ψ and ϕ vertices
24	Isolation using $\Delta R < 0.5$
25	Uncorrected B_s^0 total momentum
26	Minimum χ^2 of either K track fit
27	Isolation using $\Delta R < 0.5$ and particles from the PV
28	Leading K meson momentum
29	Leading muon momentum
30	ϕ meson momentum
31	Maximum χ^2 of the J/ψ or ϕ vertices
32	Isolation using $\Delta R < 0.75$ and particles from the PV
33	J/ψ meson momentum

Table 4.1: Variables used to train the prompt BDT, ranked by their importance in the training.

Table 4.2: Variables used to train the non-prompt BDT, ranked by their importance in the training.

Rank	Variable
1	KK invariant mass
2	B_s^0 isolation using the larger K/B_s ΔR and tracks from the PV
3	Minimum dE/dx of either K
4	χ^2 of B_s^0
5	p_T of the ϕ meson
6	p_T of the leading K meson
7	Isolation using the maximum ΔR between either K and the B_s^0
8	p_T of the trailing K meson
9	Maximum χ^2 of either K meson with the J/ψ vertex
10	Isolation using $\Delta R < 0.5$ and particles from the PV
11	Isolation using $\Delta R < 0.75$ and tracks from the PV
12	Minimum χ^2 of of either K with the J/ψ vertex
13	Minimum ΔR between either K meson and the B_s^0 candidate
14	Dimuon invariant mass
15	Total momentum of the ϕ meson
16	p_T of the J/ψ meson
17	Trailing muon momentum
18	Isolation using $\Delta R < 0.5$
19	Maximum ΔR between either K meson and the B_s^0 candidate
20	Maximum dE/dx of either K meson
21	Trailing K meson momentum
22	J/ψ vertex χ^2
23	Leading K meson momentum
24	Maximum χ^2 of either K candidate track
25	Isolation using $\Delta R < 0.75$
26	Minimum ΔR between either muon and the B_s^0 candidate
27	Minimum χ^2 of either K candidate track
28	uncorrected p_T of B_s^0 candidate
29	p_T of the trailing muon
30	J/ψ momentum
31	Maximum ΔR between either muon and the B_s^0 candidate
32	Vertex χ^2 of the ϕ meson
33	Uncorrected B_s^0 momentum
34	p_T of the leading muon
35	Leading muon momentum

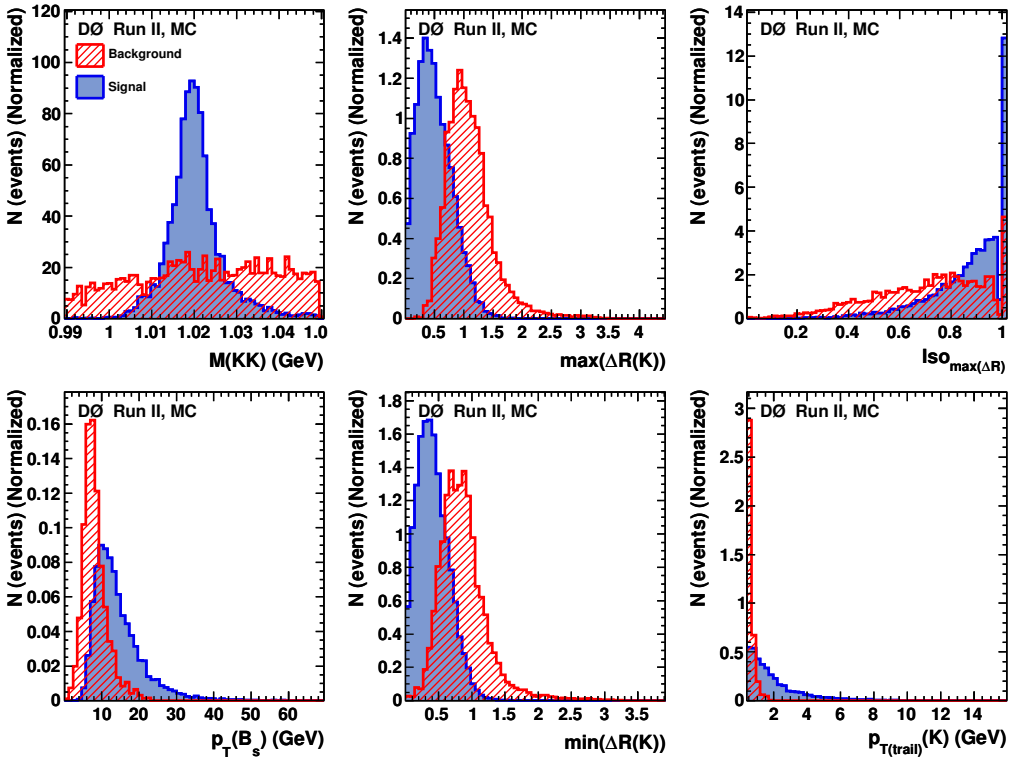


Figure 4.2: The distributions of the six most important variables used in the BDT trained on prompt J/ψ production for the $B_s^0 \to J/\psi\phi$ signal (solid blue) and prompt J/ψ events (red dashed) histograms.

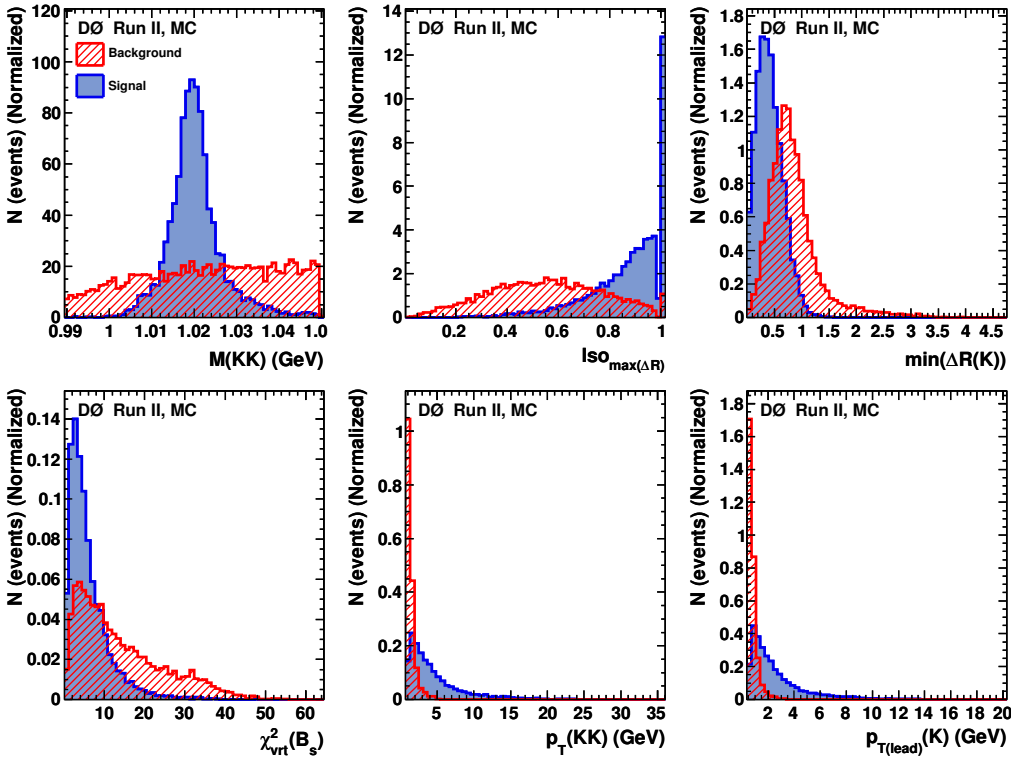


Figure 4.3: The distributions of the six most important variables used in the BDT trained on inclusive $B \to J/\psi X$ decays for the $B_s^0 \to J/\psi\phi$ signal (solid blue) and inclusive $B \to J/\psi X$ decays (red dashed) histograms.

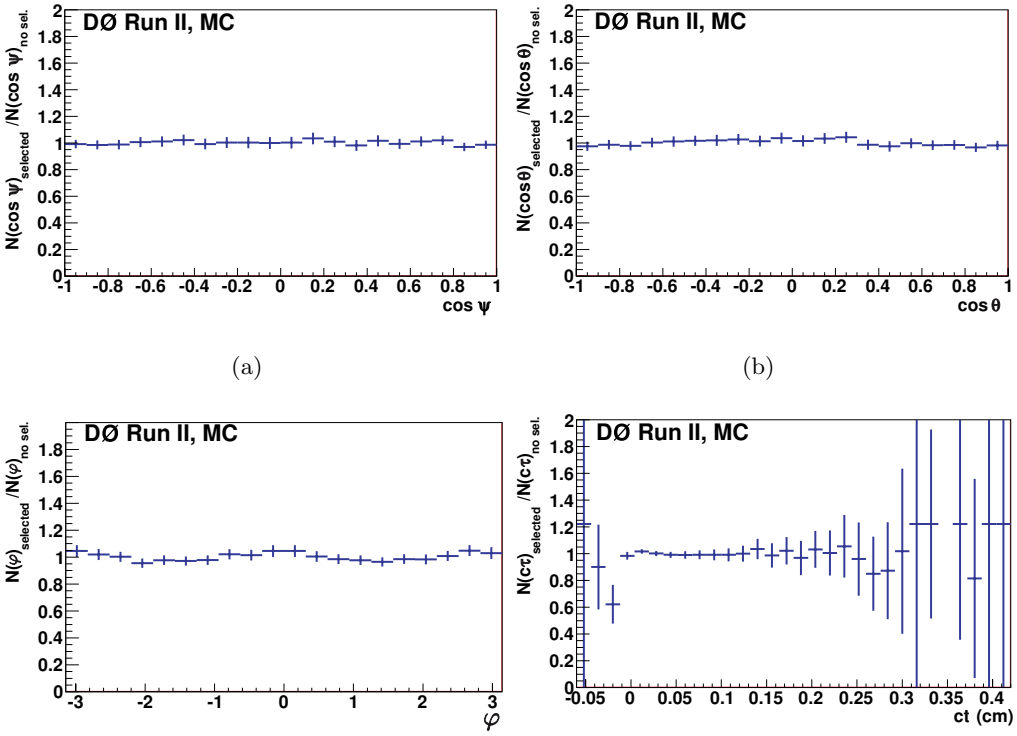


Figure 4.4: Test of uniformity of the efficiencies of the BDT selection using a MC sample with $\phi_s = -0.5$. The figure shows the ratios of the normalized distributions of (a – c) the three angles and (d) the proper decay length, before and after the BDT selection.

4.2.4 Selection Criteria

To choose the best set of criteria for the two BDT discriminants, we first step through the values of both BDT discriminants from -0.4 to 0.8 in increments of 0.01 and measure the B_s^0 signal yield for each choice of cuts. Next, we define 14 signal yield regions between 4000 and 7000 events, and for each region choose the pair of BDT cuts which gives the highest significance $S/\sqrt{S+B}$, where S (B) is the number of signal (background) events in the data sample. The 14 points, in increasing order of the signal size S, are shown in Table 4.3. Figure 4.5 shows the number of signal events as a function of the total number of events for the 14 points. As the BDT criteria are loosened, the total number of events increases by a factor of ten, while the number of signal events increases by about 50%.

As a test of possible detrimental effects of training on variables with low separation power, we have repeated the above procedure a reduced number of variables, 18 variables for the prompt background and 13 variables for the non-prompt background. The resulting number of background events for a given number of signal events is larger by about 10%. Therefore, we proceed with the original number of variables.

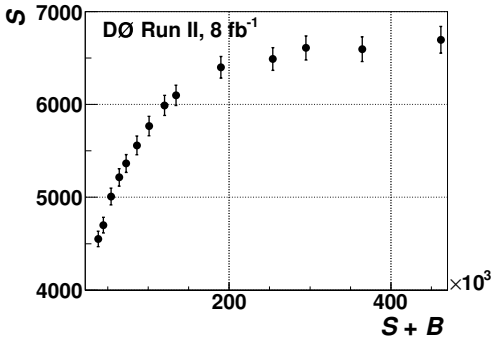
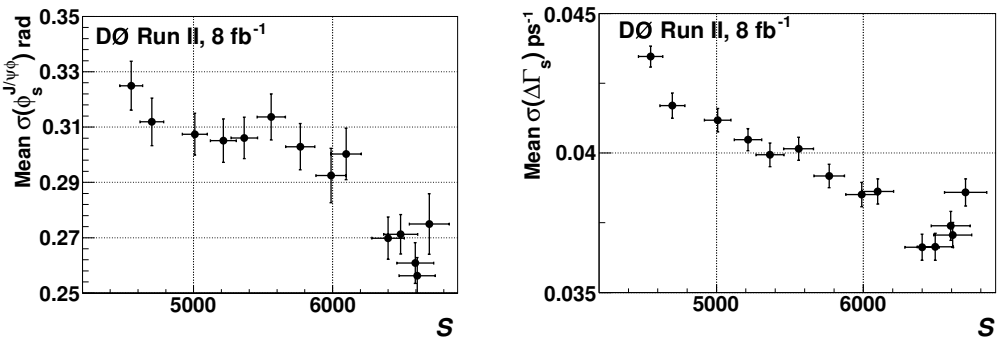


Figure 4.5: Number of $B_s^0 \to J/\psi\phi$ signal events as a function of the total number of events for the 14 criteria sets considered.

The choice of the final cut on the BDT output is based on an ensemble study. For each point in Table 4.3, we perform a maximum-likelihood fit to the event distribution in the 2-dimensional (2D) space of B_s^0 candidate mass and proper time. This 2D fit provides a parametrization of the background mass and proper time distribution. We then generate pseudo-experiments in the 5D space of B_s^0 candi-



(a) (b)

Figure 4.6: Ensemble study results for (a) mean value of $\sigma(\phi_s)$ as a function of the number of signal events and (b) mean value of $\sigma(\Delta\Gamma_s)$ as a function of the number of signal events.

date mass, proper time, and three independent angles of decay products, using as input the parameters as obtained in a preliminary study, and the background from the 2D fit. We perform a 5D maximum likelihood fit on the ensembles and compare the distributions of the statistical uncertainties of $\phi_s^{J/\psi\phi}$ ($\sigma(\phi_s^{J/\psi\phi})$)

Table 4.3: Numbers of signal and signal-plus-background events for different sets of BDT criteria, shown in the last two columns, that give the largest value of $S/\sqrt{S+B}$ for a given S.

Criteria Set	S	$S+B$	Non-prompt BDT	Prompt BDT
0	4550	38130	0.45	0.42
1	4699	44535	0.45	0.29
2	5008	53942	0.39	0.35
3	5213	64044	0.36	0.30
4	5364	72602	0.33	0.28
5	5558	85848	0.13	0.41
6	5767	100986	0.21	0.29
7	5988	120206	0.13	0.29
8	6097	134255	0.07	0.29
9	6399	189865	0.04	0.10
10	6489	254022	−0.05	−0.01
11	6608	294949	−0.13	0.00
12	6594	364563	−0.18	−0.14
13	6695	461744	−0.35	−0.08

and $\Delta\Gamma_s$ ($\sigma(\Delta\Gamma_s)$) for the different sets of criteria. The dependence of the mean values of $\sigma(\phi_s^{J/\psi\phi})$ and $\sigma(\Delta\Gamma_s)$ on the number of signal events is shown in Figures 4.6(a) and 4.6(b). The mean statistical uncertainties of both $\phi_s^{J/\psi\phi}$ and $\Delta\Gamma_s$ systematically decrease with increasing signal, favoring looser cuts. The gain in the parameter resolution is slower for the three loosest criteria, while the total number of events doubles from about 0.25×10^6 to 0.5×10^6. The fits used for these ensemble tests were simplified, therefore the magnitude of the predicted uncertainty is expected to underestimate the final measured precision. However, the general trends should be valid.

Based on these results, we choose the sample that contains about 6500 signal events, (labeled "Set 10" in Table 4.3) as a final selection and refer to it as the "BDT selection". Figure 4.4 in Section 4.2.3 shows the ratios of the normalized distributions of the three angles (see Section 4.5) and the lifetime before and after the BDT selection. The ratios are consistent with unity, which means that the BDT requirements do not significantly alter these distributions.

4.2.5 Simple Selection

We select a second event sample by applying criteria on event quality and kinematic quantities. We use the consistency of the results obtained for the BDT and for this sample as a measure of systematic effects related to imperfect modeling of the detector acceptance and of the selection requirements.

The criteria are the same as in Refs. [14] and [15]. Each of the four tracks is required to have at least two SMT hits and at least eight hits in SMT or CFT. We require minimum momentum in the transverse plane p_T for B_s^0, ϕ, and K meson candidates of 6.0 GeV, 1.5 GeV, and 0.7 GeV, respectively. Muons are required to have p_T above 1.5 GeV. For events in the central rapidity region (an event is considered to be central if the higher p_T muon has $|\eta(\mu_{\text{leading}})| < 1$), we require the transverse momentum of the J/ψ meson to exceed 4 GeV. In addition, J/ψ candidates are accepted if the invariant mass of the muon pair is in the range 3.1 $\pm$ 0.2 GeV. Events are required to satisfy the condition $\sigma(t) < 0.2$ ps where $\sigma(t)$

is the uncertainty on the decay proper time obtained from the propagation of the uncertainties in the decay-vertex kinematic fit, the primary vertex position, and the B_s^0 candidate transverse momentum. We refer to this second sample as the "Square-cuts" sample.

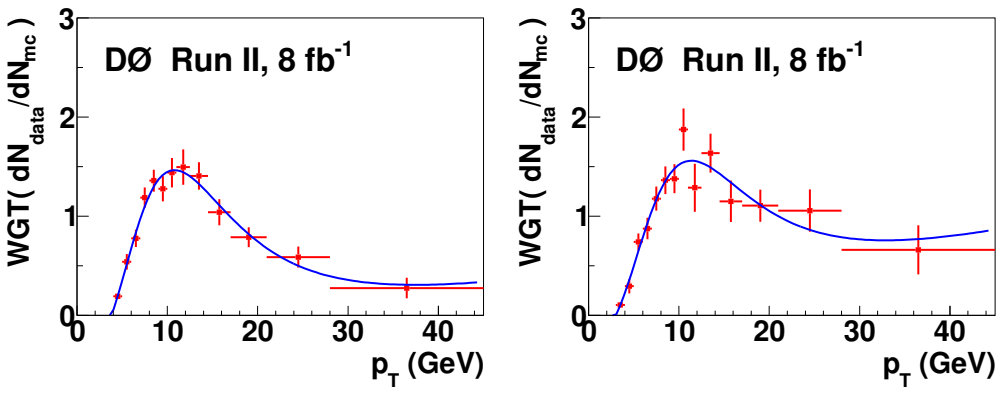
4.3 Detector acceptance

We take into account the shaping of the signal distribution by the detector acceptance and kinematic selection by introducing acceptance functions in the three angles of the transversity basis . The acceptance functions are derived from Monte Carlo simulation. Due to the event triggering effects, the momentum spectra of final-state objects in data are harder than in MC. We take into account the difference in the p_T distribution of the final-state objects in data and MC by introducing a weight factor as a function of $p_T(J/\psi)$, separately for the central ($|\eta(\mu_{\text{leading}})| < 1$) and forward regions. The weight factor is derived by forcing an agreement between the J/ψ transverse momentum spectra in data and MC. The behavior of the weight factor as a function of $p_T(J/\psi)$ for the BDT-based selection, for the central and forward regions, is shown in Figure 4.7.

Figure 4.8 shows the background-subtracted p_T distributions of the leading and trailing muon and leading and trailing kaon, in the central region. There is a good agreement between data and MC for all final-state particles after applying the weight factor. The acceptance in φ and θ and ψ are shown in Figure 4.9.

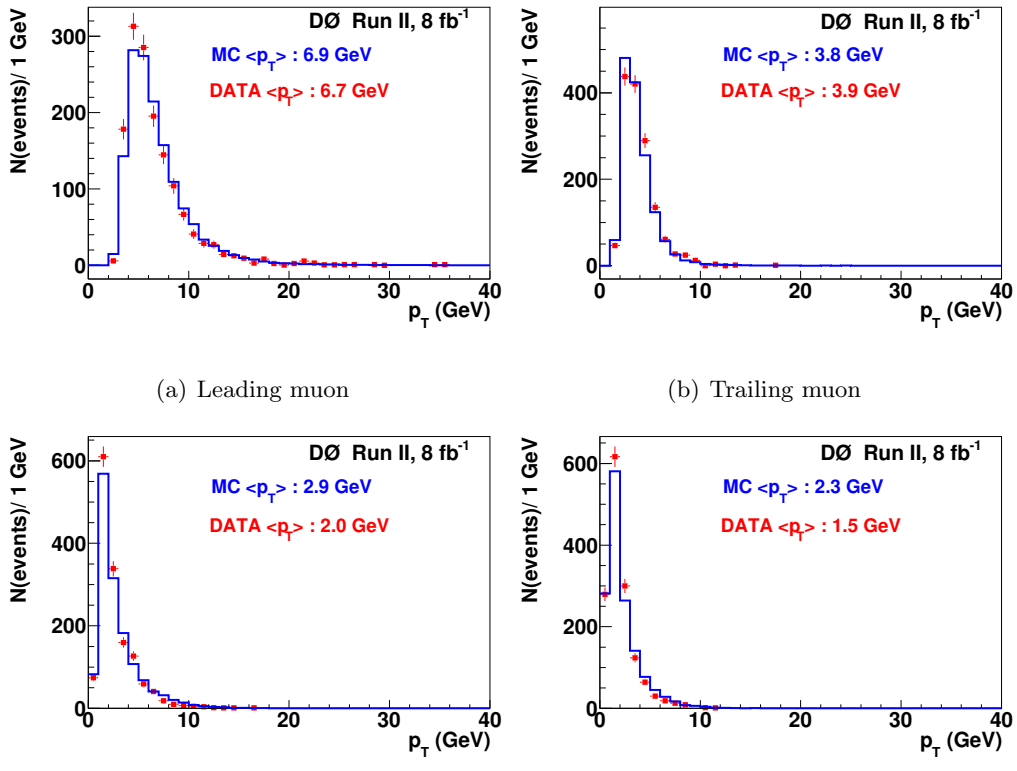
4.4 Flavor Tagging

At the Tevatron, b quarks are mostly produced in $b\bar{b}$ pairs. The flavor of the initial state of the B_s^0 candidate is determined by exploiting properties of particles produced by the other b hadron ("opposite-side tagging", or OST). The OST-discriminating variables x_i are based primarily on the presence of a muon or an electron from the semi-leptonic decay or the decay vertex charge of the other b hadron produced in the $p\overline{p}$ interaction.



(a) (b)

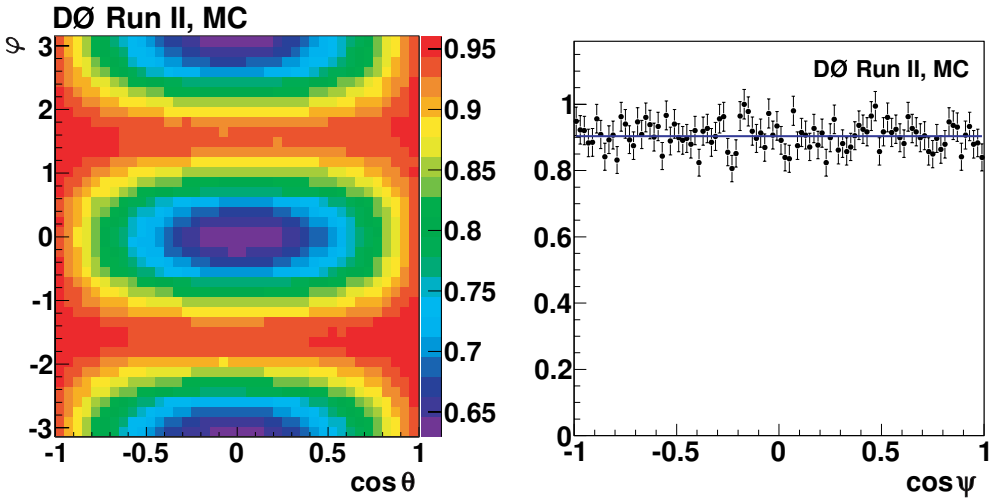
Figure 4.7: Weight factor as a function of $p_T(J/\psi)$ used to correct MC p_T distribution of B_s^0 and B_d^0 decay objects for (a) central region, and (b) forward region. The curves are empirical fits to a sum of a Landau function and a polynomial.



(a) Leading muon (b) Trailing muon

(c) Leading kaon (d) Trailing kaon

Figure 4.8: Transverse momentum distributions of the four final-state particles in data (points) and weighted MC (solid histogram), for the BDT-based event selection.



(a) (b)

Figure 4.9: (a) Map of the detector acceptance on the plane φ–$\cos\theta$. (b) Detector acceptance as a function of $\cos\psi$. The acceptance is uniform in $\cos\psi$.

For the initial b quark, the probability density function (PDF) for a given variable x_i is denoted as $f_i^b(x_i)$, while for the initial $\bar{b}$ quark it is denoted as $f_i^{\bar{b}}(x_i)$. The combined tagging variable y is defined as:

$$y = \prod_{i=1}^{n} y_i; \;\; y_i = \frac{f_i^{\bar{b}}(x_i)}{f_i^b(x_i)}. \tag{4.2}$$

A given variable x_i can be undefined for some events. For example, there are events that don't contain an identified muon from the opposite side. In this case, the corresponding variable y_i is set to 1.

In this way the OST algorithm assigns to each event a value of the predicted tagging parameter $d = (1-y)/(1+y)$ in the range $[-1,1]$, with $d > 0$ tagged as an initial b quark and $d < 0$ tagged as an initial $\bar{b}$ quark. Larger $|d|$ values correspond to higher tagging confidence. In events where no tagging information is available d is set to zero. The efficiency ϵ of the OST, defined as fraction of the number of candidates with $d \neq 0$, is 18%. The OST-discriminating variables and algorithm are described in detail in Ref. [12].

The tagging dilution $\mathcal{D}$ is defined as

$$\mathcal{D} = \frac{N_{\rm cor} - N_{\rm wr}}{N_{\rm cor} + N_{\rm wr}}, \tag{4.3}$$

where $N_{\rm cor}$ ($N_{\rm wr}$) is the number of events with correctly (wrongly) identified initial B-meson flavor. The dependence of the tagging dilution on the tagging parameter d is calibrated with data for which the flavor (B or $\overline{B}$) is known.

4.4.1 OST calibration

The dilution calibration is based on four independent $B^0_d \to \mu\nu D^{*\pm}$ data samples corresponding to different time periods, denoted IIa, IIb1, IIb2, and IIb3, with different detector configurations and different distributions of instantaneous luminosity. The Run IIa sample was used in Ref. [12].

For each sample we perform an analysis of the $B^0_d - \overline{B}^0_d$ oscillations described in Ref. [11]. We divide the samples in five ranges of the tagging parameter $|d|$, and for each range we obtain a mean value of the dilution $|\mathcal{D}|$. The mixing frequency ΔM_d is fitted simultaneously and is found to be stable and consistent with the world average value. The measured values of the tagging dilution $|\mathcal{D}|$ for the four data samples above, in different ranges of $|d|$, are shown in Figure 4.10. The dependence of the dilution on $|d|$ is parametrized as

$$|\mathcal{D}| = \frac{p_0}{(1+\exp((p_1-|d|)/p_2))} - \frac{p_0}{(1+\exp(p_1/p_2))}. \qquad (4.4)$$

and the function is fitted to the data. All four measurements are in good agreement and hence a weighted average is taken.

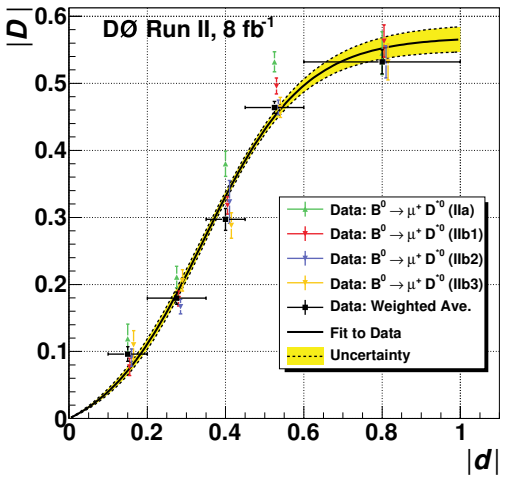


Figure 4.10: Parametrization of the dilution $|\mathcal{D}|$ as a function of the tagging parameter $|d|$ for the combined opposite-side tagger. The curve is the result of the weighted fit to four self-tagging control data samples (see text).

4.5 Maximum Likelihood Fit

We perform a six-dimensional (6D) unbinned maximum likelihood fit to the proper decay time and its uncertainty, three decay angles characterizing the final state, and the mass of the B_s^0 candidate. We use events for which the invariant mass of the K^+K^- pair is within the range 1.01 – 1.03 GeV. There are 104683 events in the BDT-based sample and 66455 events in the Square-cuts sample. We use the formulae described in Section 1.3 The normalized functional form of the differential decay rate includes an $\mathcal{S}$-wave KK contribution in addition to the dominant $\mathcal{P}$-wave $\phi \to K^+K^-$ decay. To model the distributions of the signal and background we use the software library RooFit [74].

4.5.1 Signal model

The angular distribution of the signal is expressed in the transversity basis [20]. In the coordinate system of the J/ψ rest frame, where the ϕ meson moves in the x direction, the z axis is perpendicular to the decay plane of $\phi \to K^+K^-$, and $p_y(K^+) \geq 0$. The transversity polar and azimuthal angles θ and φ describe the direction of the positively-charged muon, while ψ is the angle between $\vec{p}(K^+)$ and $-\vec{p}(J/\psi)$ in the ϕ rest frame. The angles are shown in Figure 4.11.

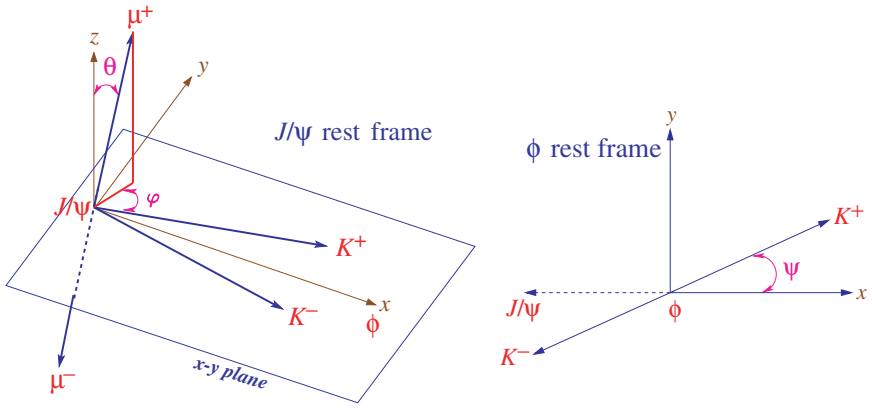


Figure 4.11: Definition of the angle ψ, and the transversity angles θ and φ.

For the $\mathcal{S}$-wave component, we assume a uniform distribution in the range $1.01 < M(KK) < 1.03$ GeV. We constrain the oscillation frequency to $\Delta M_s = 17.77 \pm 0.12$ ps^{-1}, as measured in Ref. [10]. Table 4.4 lists all physics parameters

used in the fit.

Parameter	Definition
$\|A_0\|^2$	$\mathcal{P}$-wave longitudinal amplitude squared, at $t = 0$
$A1$	$\|A_\parallel\|^2/(1 - \|A_0\|^2)$
$\overline{\tau}_s$ (ps)	B_s^0 mean lifetime
$\Delta\Gamma_s$ (ps^{-1})	Heavy-light decay width difference
F_S	K^+K^- $\mathcal{S}$-wave fraction
β_s	CP-violating phase ($\equiv -\phi_s^{J/\psi\phi}/2$)
$\delta_\parallel$	$\arg(A_\parallel/A_0)$
$\delta_\perp$	$\arg(A_\perp/A_0)$
δ_s	$\arg(A_s/A_0)$

Table 4.4: Definition of nine real measurables for the decay $B_s^0 \to J/\psi\phi$ used in the Maximum Likelihood fitting.

For the signal mass distribution we use a Gaussian function with a free mean value, width, and normalization. The function describing the signal rate in the 6D space is invariant under the combined transformation

$\beta_s \to \pi/2 - \beta_s$, $\Delta\Gamma_s \to -\Delta\Gamma_s$, $\delta_\parallel \to 2\pi - \delta_\parallel$, $\delta_\perp \to \pi - \delta_\perp$, and $\delta_s \to \pi - \delta_s$.

In addition, with a limited flavor-tagging power, there is an approximate symmetry around $\beta_s = 0$ for a given sign of $\Delta\Gamma_s$.

We correct the signal decay rate by a detector acceptance factor $\epsilon(\psi, \theta, \varphi)$ parametrized by coefficients of expansion in Legendre polynomials $P_k(\psi)$ and real harmonics $Y_{lm}(\theta, \varphi)$. The coefficients are obtained from Monte Carlo simulated samples, as described in Section 4.3.

The signal decay time resolution is given by a Gaussian centered at zero and width given by the product of a global scale factor and the event-by-event uncertainty in the decay time measurement. The distribution of the uncertainty in the decay time measurement in the MC simulation is modeled by a superposition of five Gaussian functions. The background-subtracted signal distribution agrees well with the MC model, as seen in Figure 4.12. Variations of the parameters within one sigma of the best fit are used to define two additional functions, also shown in the figure, that are used in alternative fits to estimate the systematic

effect due to time resolution.

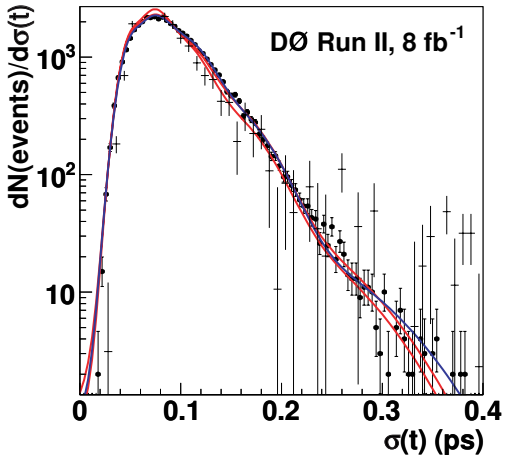


Figure 4.12: The distribution of the uncertainty in the decay time for the signal, MC (squares) and background-subtracted data (crosses). The blue curve is the sum of five Gaussian functions fitted to the MC distribution. The two red lines are variations of the default function used in the studies of systematic effects.

4.5.2 Background model

The proper decay time distribution of the background is described by a sum of a prompt component, modeled as the same resolution function used in the signal decay time, and a non-prompt component. The non-prompt component is modeled as a superposition of one exponential decay for $t < 0$ and two exponential decays for $t > 0$, with free slopes and normalizations. The lifetime resolution is modeled by an exponential convoluted with a Gaussian function, with two separate parameters for prompt and non-prompt background. To allow for the possibility of the lifetime uncertainty to be systematically underestimated, we introduce a free scale factor.

The mass distributions of the two components of background are parametrized by low-order polynomials: a linear function for the prompt background and a quadratic function for the non-prompt background. The angular distribution of background is parametrized by Legendre and real harmonics expansion coefficients. A separate set of expansion coefficients c^k_{lm} and c^k_{lm}, with $k = 0$ or 2 and $l = 0, 1, 2$, is used for the prompt and non-prompt background. A preliminary fit is first performed with all 17 parameters ${c^{(P)}}^k_{lm}$ for prompt and 17 parameters ${c^{(NP)}}^k_{lm}$ for non-prompt allowed to vary. In subsequent fits those that converge at values within two standard deviations of zero are set to zero. Nine free parameters remain, five for non-prompt background: ${c^{(NP)}}^0_{1-1}$, ${c^{(NP)}}^0_{20}$, ${c^{(NP)}}^0_{22}$, ${c^{(NP)}}^2_{00}$, and

$c^{(NP)}{}^2_{22}$, and four for prompt background: $c^{(P)}{}^0_{1-1}$, $c^{(P)}{}^0_{20}$, $c^{(P)}{}^0_{22}$, and $c^{(P)}{}^2_{2-1}$. All background parameters described above are varied simultaneously with physics parameters. In total, there are 36 parameters used in the fit. In addition to the nine physics parameters defined in Table 4.4, they are: signal yield, mean mass and width, non-prompt background contribution, six non-prompt background lifetime parameters, four background time resolution parameters, one time resolution scale factor, three background mass distribution parameters, and nine parameters describing background angular distributions.

4.5.3 Fit results

The maximum likelihood fit results for the nominal fit (Default), for two alternative time resolution functions, $\sigma_A(t)$ and $\sigma_B(t)$ shown in Figure 4.12, and for an alternative $M(KK)$ dependence of the $\phi(1020) \to K^+K^-$ decay with the decay width increased by a factor of two are shown in Table 4.5 and Table 4.6. These alternative fits are used to estimate the systematic uncertainties. The fit assigns 5598 ± 113 (5050 ± 105) events to the signal for the BDT (Square-cuts) sample. Only the parameters whose values do not suffer from multi-modal effects are shown. A single fit does not provide meaningful point estimates and uncertainties for the four phase parameters. Their estimates are obtained using the MCMC technique. Figures 4.13 – 4.16 illustrate the quality of the fit for the background, for all data, and for the signal-enhanced sub-samples.

An independent measurement of the $\mathcal{S}$-wave fraction is described in Section 4.5.5 and the result is in agreement with F_S determined from the maximum likelihood fit.

Parameter	Default	$\sigma_A(t)$	$\sigma_B(t)$	$\Gamma_\phi = 8.52$ MeV
$\lvert A_0\rvert^2$	0.553 ± 0.016	0.553 ± 0.016	0.552 ± 0.016	0.553 ± 0.016
$\lvert A_\parallel\rvert^2/(1-\lvert A_0\rvert^2)$	0.487 ± 0.043	0.483 ± 0.043	0.485 ± 0.043	0.487 ± 0.043
$\overline{\tau}_s$ (ps)	1.417 ± 0.038	1.420 ± 0.037	1.417 ± 0.037	1.408 ± 0.434
$\Delta\Gamma_s$ (ps^{-1})	0.151 ± 0.058	0.136 ± 0.056	0.145 ± 0.057	0.170 ± 0.067
F_S	0.147 ± 0.035	0.149 ± 0.034	0.147 ± 0.035	0.147 ± 0.035

Table 4.5: Maximum likelihood fit results for the BDT selection. The uncertainties are statistical.

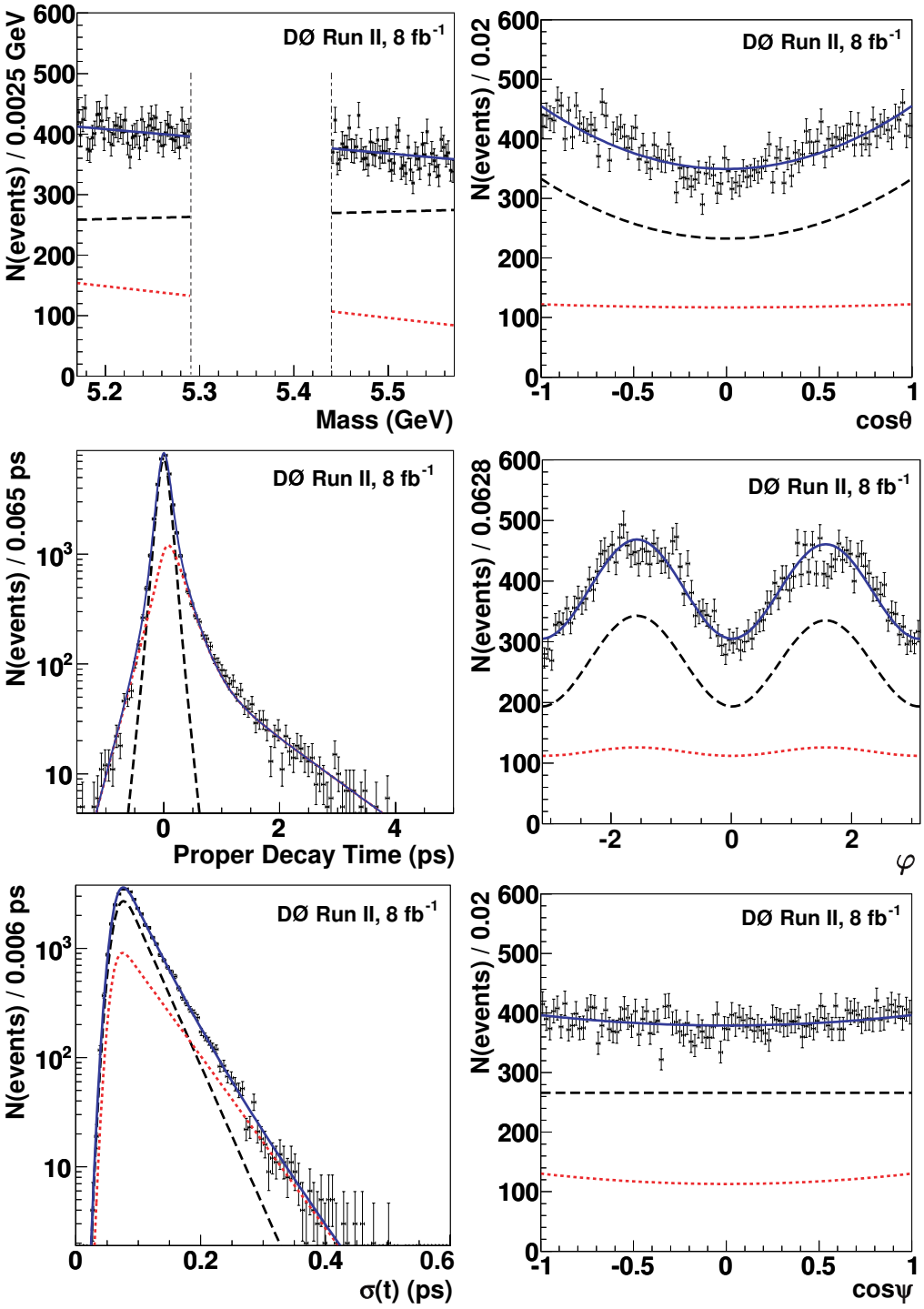


Figure 4.13: The distributions in the background (B_s^0 mass sidebands) region of candidate mass, proper decay time, decay time uncertainty, transversity polar and azimuthal angles, and $\cos\psi$ for the BDT sample. The curves show the prompt (black dashed) and non-prompt (red dotted) components, and their sum (blue solid).

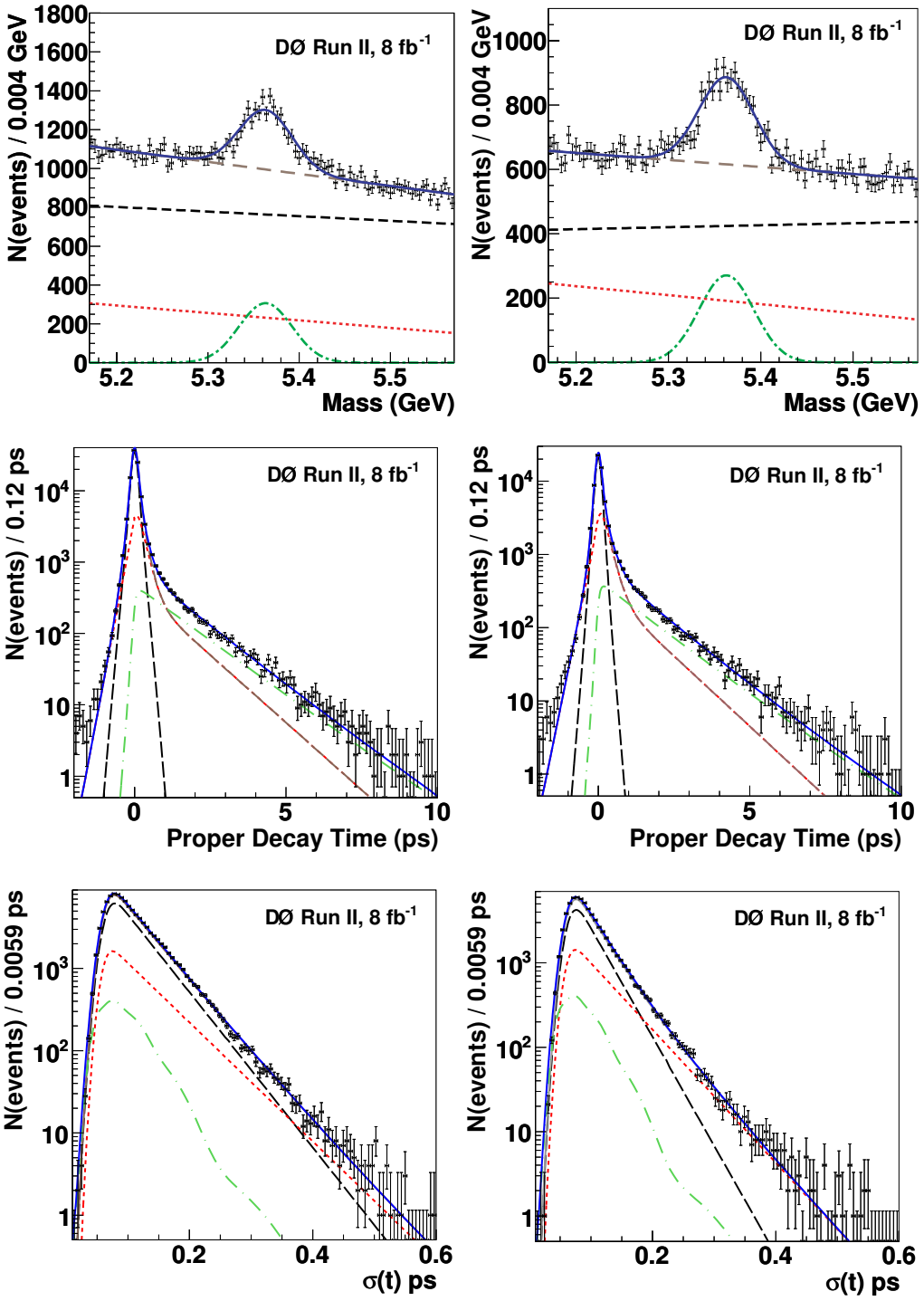


Figure 4.14: Invariant mass, proper decay time, and proper decay time uncertainty distributions for B_s^0 candidates in the BDT sample (left) and Square-cuts sample (right). The curves are projections of the maximum likelihood fit. Shown are the signal (green dashed-dotted curve), prompt background (black dashed curve), non-prompt background (red dotted curve), total background (brown long-dashed curve), and the sum of signal and total background (solid blue curve).

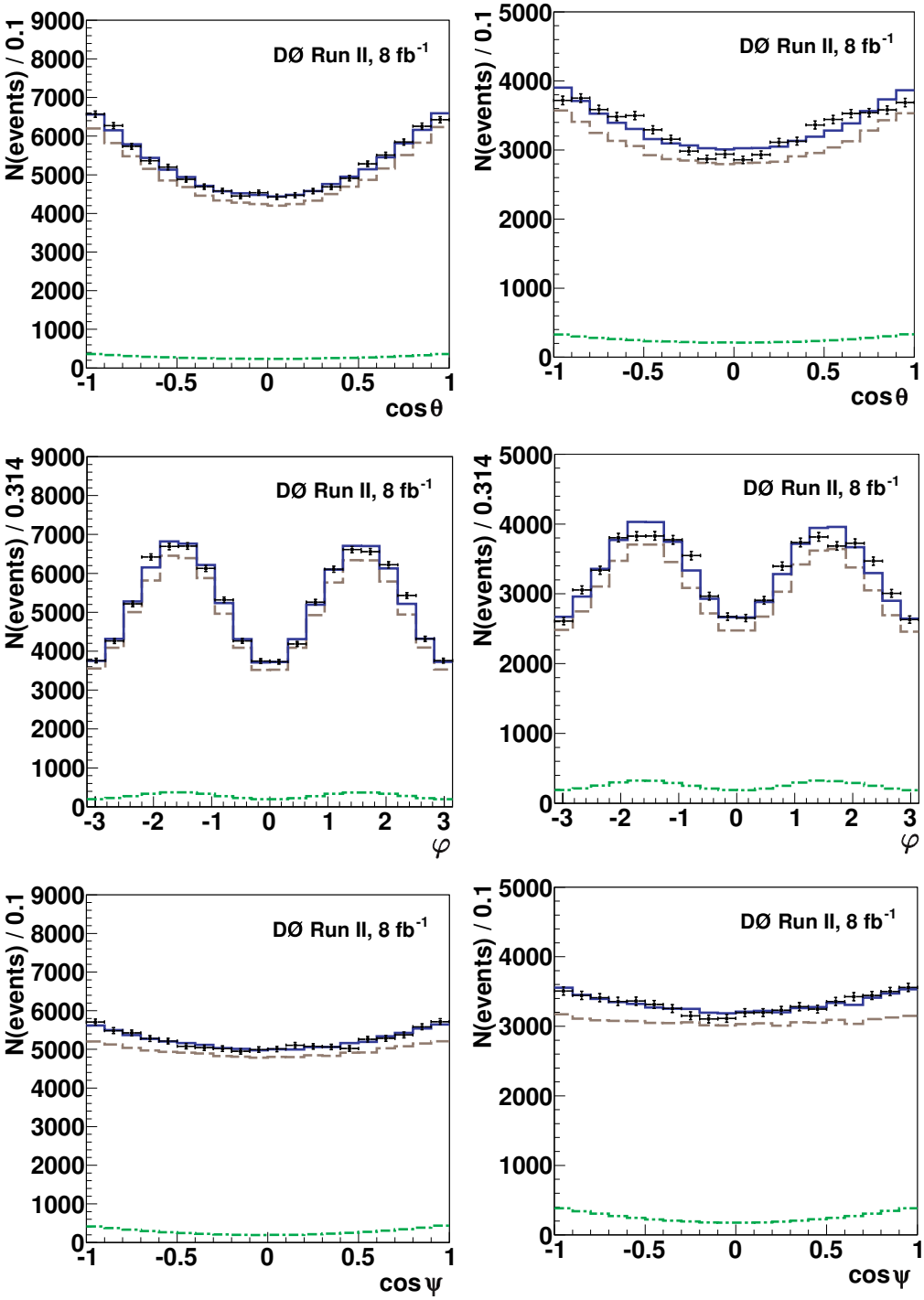


Figure 4.15: Distributions of transversity polar and azimuthal angles and $\cos\psi$ for B_s^0 candidates in the BDT sample (left) and Square-cuts sample (right). The curves are projections of the maximum likelihood fit. Shown are the signal (green dashed-dotted curve), total background (brown long-dashed curve) and the sum of signal and total background (blue solid curve).

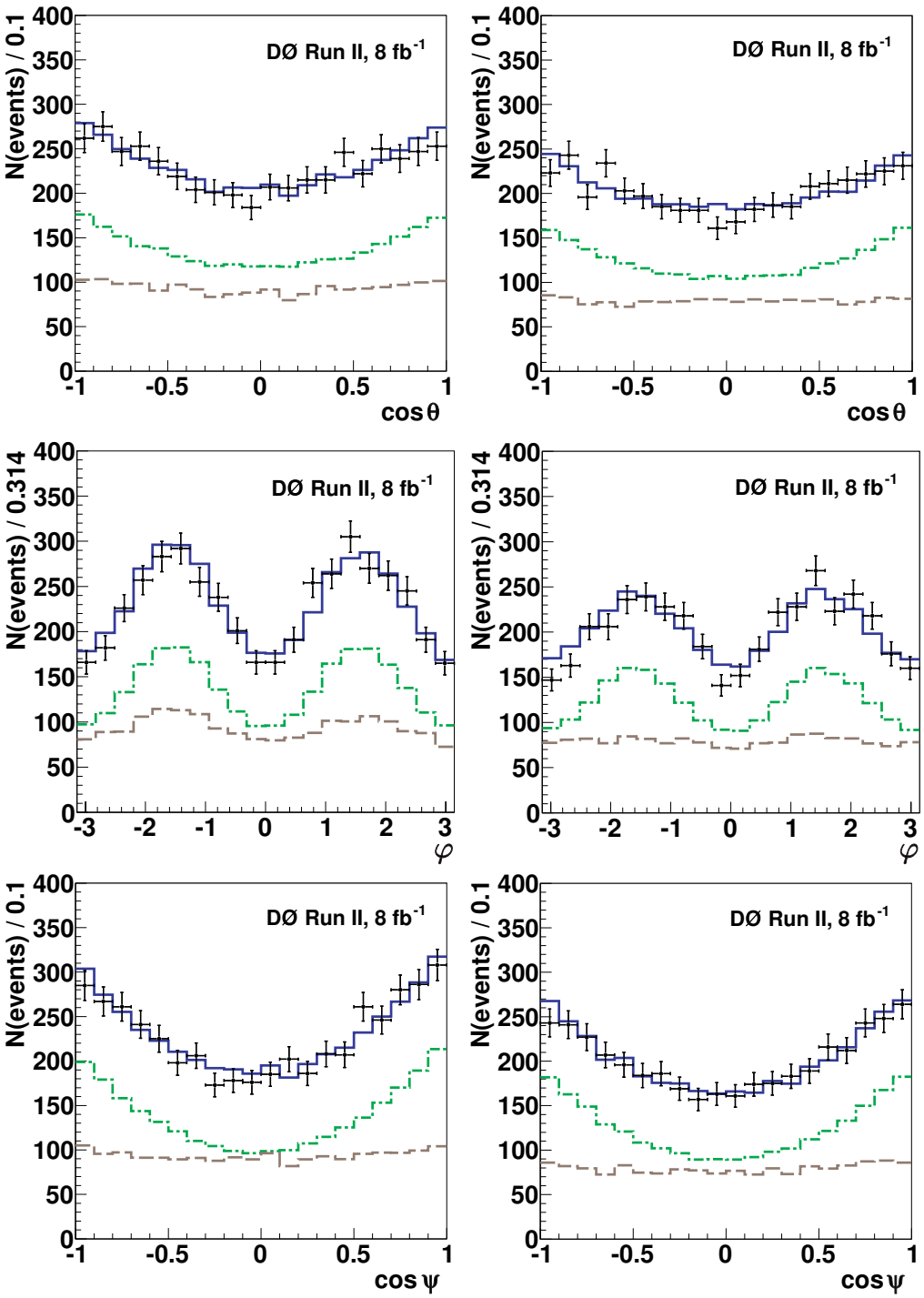


Figure 4.16: Distributions of transversity polar and azimuthal angles and $\cos\psi$ for B_s^0 candidates in the BDT sample (left) and Square-cuts sample (right). The signal contribution is enhanced, relative to the distributions shown in Figure 4.15, by additional requirements on the reconstructed mass of the B_s^0 candidates ($5.31 < M(B_s^0) < 5.43$ GeV) and on the proper time $t > 1.0$ ps. The curves are projections of the maximum likelihood fit. Shown are the signal (green dashed-dotted curve), total background (brown long-dashed curve) and the sum of signal and total background (blue solid curve).

Table 4.6: Maximum likelihood fit results for the 'Square-cuts' sample.

Parameter	Default	$\sigma_A(t)$	$\sigma_B(t)$	$\Gamma_\phi = 8.52$ MeV
$\lvert A_0\rvert^2$	0.566 ± 0.017	0.564 ± 0.017	0.567 ± 0.017	0.566 ± 0.017
$\lvert A_\parallel\rvert^2/(1-\lvert A_0\rvert^2)$	0.579 ± 0.048	0.579 ± 0.048	0.577 ± 0.048	0.579 ± 0.048
$\overline{\tau}_s$ (ps)	1.439 ± 0.039	1.450 ± 0.038	1.457 ± 0.037	1.438 ± 0.042
$\Delta\Gamma_s$ (ps^{-1})	0.199 ± 0.058	0.194 ± 0.057	0.185 ± 0.056	0.202 ± 0.060
F_S	0.175 ± 0.035	0.169 ± 0.035	0.171 ± 0.035	0.175 ± 0.035

4.5.4 Systematic uncertainties

There are several possible sources of systematic uncertainty in the measurements. These uncertainties are estimated as described below.

- **Flavor tagging**: The measured flavor mistag fraction suffers from uncertainties due to the limited number of events in the data samples for the decay $B^0_d \to \mu\nu D^{(*)\pm}$. The nominal calibration of the flavor tagging dilution is determined as a weighted average of four samples separated by the running period. As an alternative, we use two separate calibration parameters, one for the data collected in running periods IIa and IIb1, and one for the IIb2 and IIb3 data. We also alter the nominal parameters by their uncertainties. We find the effects of the changes to the flavor mistag variation negligible.

- **Proper decay time resolution**: Fit results can be affected by the uncertainty of the assumed proper decay time resolution function. To assess the effect, we have used two alternative parametrizations obtained by random sampling of the resolution function.

- **Detector acceptance**: The effects of imperfect modeling of the detector acceptance and of the selection requirements are estimated by investigating the consistency of the fit results for the sample based on the BDT selection and on the Square-cuts selection. Although the overlap between the two samples is 70%, and some statistical differences are expected, we interpret the differences in the results as a measure of systematic effects.

The two event selection approaches have different merits. The BDT-based approach uses more information on each event, and hence it allows a higher signal yield at lower background. However, it accepts signal events of lower quality (large vertex χ^2 or proper decay time uncertainty) that are rejected by the square cuts. Also, the BDT-based approach uses the $M(KK)$ distribution as a discriminant in the event selection, affecting the results for the parameters entering the $\mathcal{S}-\mathcal{P}$ interference term, particularly the $\mathcal{S}$-wave fraction F_S and the phase parameters.

The main difference between the two samples is in the kinematic ranges of final-state kaons, and so the angular acceptance functions and MC weights (see Section 4.3) are different for the two samples. Imperfections in the modelling of the B_s^0 decay kinematics and estimated acceptances, and in the treatment of the MC weighting, are reflected in differences between fit results. The differences are used as an estimate of this class of systematic uncertainty.

- ***M*(*KK*) resolution:** The limited $M(KK)$ resolution may affect the results of the analysis, especially the phases and the $\mathcal{S}$-wave fraction F_S, through the dependence of the $\mathcal{S}-\mathcal{P}$ interference term on the $\mathcal{P}$-wave mass model. In principle, the function of Eq. (1.47) should be replaced by a Breit-Wigner function convoluted with a Gaussian. We avoid this complication by approximating the smeared $\mathcal{P}$-wave amplitude by a Breit-Wigner function where the width Γ_ϕ of Eq. (1.47) is set to twice the world average value to account for the detector resolution effects. A MC simulation-based estimate of the scale factor for the event selection criteria used in this analysis yields a value in the range 1.5 – 1.7. The resulting complex integral of the $\mathcal{S}-\mathcal{P}$ interference has an absolute value behavior closer to the data, but a distorted ratio of the real and imaginary parts compared to Eq. (1.47). We repeat the fits using this altered $\phi(1020)$ propagator as a measure of the sensitivity to the $M(KK)$ resolution.

Tables 4.5 and 4.6 compare results for the default fit and the alternative fits discussed above. The differences between the best-fit values provide a measure of systematic effects. For the best estimate of the credible intervals for all the measured physics quantities, we conduct MCMC studies described in the next section. Other sources of systematic uncertainties like the functional model of the background mass, lifetime and angle distributions were studied and give a negligible contribution.

4.5.5 Independent estimate of F_S

In the Maximum Likelihood fit, the invariant mass of the K^+K^- pair is not used. To do so would require a good model of the $M(K^+K^-)$ dependence of background, including a small $\phi(1020)$ component, as a function of the B_s^0 candidate mass and proper time. However, we can use the $M(K^+K^-)$ mass information to make an independent estimate of the non-resonant K^+K^- contribution in the final state.

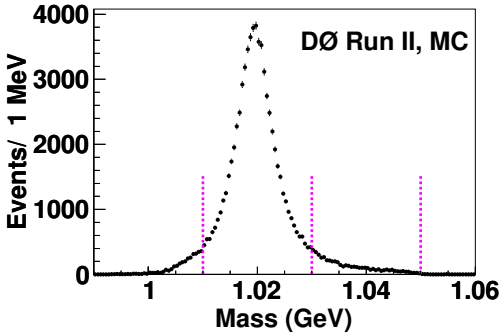


Figure 4.17: Invariant mass distribution of kaon pairs from the full simulation of the decay $\phi \to K^+K^-$. Vertical dashed lines delineate the two $M(KK)$ invariant mass bins considered.

For this study, we use the "Square-cuts" sample, for which the event selection is not biased in $M(K^+K^-)$. Using events with decay length $ct > 0.02$ cm to suppress background, we extract the B_s^0 signal in two ranges of $M(K^+K^-)$: $1.01 < M(KK) < 1.03$ GeV and $1.03 < M(KK) < 1.05$ GeV. The first range is that used by both selections, and contains the bulk of the $\phi \to K^+K^-$ signal. The second range will still contain a small Breit-Wigner tail of $\phi \to K^+K^-$. From the simulated $M(K^+K^-)$ distribution of the $B_s^0 \to J/\psi\phi$ decay, shown in Figure 4.17, we obtain the fraction of the K^+K^- decay products in the upper

mass range to be 0.061 ± 0.001 of the total range $1.01 < M(KK) < 1.05$ GeV. The $\mathcal{S}$-wave component is assumed to be a flat distribution in $M(KK)$ across this range. Given that the widths of the ranges are the same, the number of candidates due to the $\mathcal{S}$-wave contribution should be the same for both.

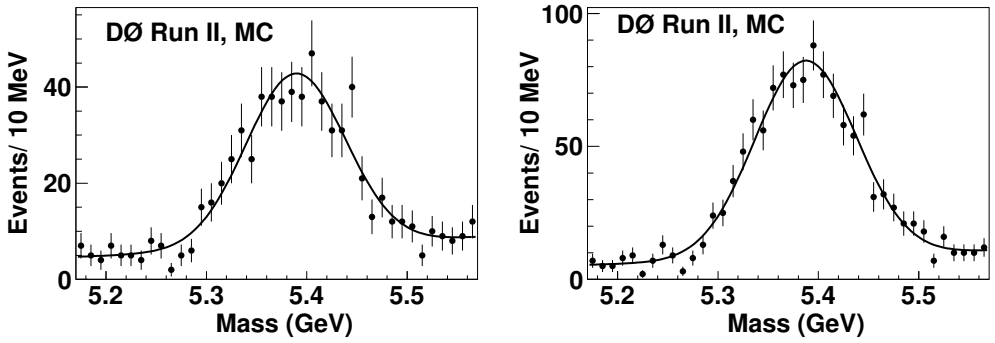


Figure 4.18: The simulated distributions of the invariant mass of the $B_d^0 \to J/\psi K^*$ decay products reconstructed under the $B_s^0 \to J/\psi\phi$ hypothesis for $1.01 < M(KK) < 1.03$ GeV (left) and $1.03 < M(KK) < 1.05$ GeV (right). The curves are results of fits assuming a sum of two Gaussian functions.

The B_s^0 signal in each mass range is extracted by fitting the B_s^0 candidate mass distribution to a Gaussian function representing the signal, a linear function for the background, and MC simulation-based templates for the $B^0 \to J/\psi K^*$ reflection where the pion from the K^* decay is assumed to be a kaon. The two shape templates used, one for each mass range, are shown in Figure 4.18. The mass distributions, with fits using the above templates, are shown in Figure 4.19. The fits result in the B_s^0 yield of 3027 ± 93 events for $1.01 < M(KK) < 1.03$ GeV and 547 ± 94 events for $1.03 < M(KK) < 1.05$ GeV. In the mass range $1.01 < M(KK) < 1.03$ GeV, we extract the fraction of B_s^0 candidates decaying into non-resonant KK to be 0.12 ± 0.03. The error includes the uncertainties in the signal and background modeling. This excess may be due to an $\mathcal{S}$-wave, or a non-resonant $\mathcal{P}$-wave, or a combination of both. If we assign it entirely to the $\mathcal{S}$-wave, and assume it to be independent of $M(KK)$, we obtain the measured $\mathcal{S}$-wave fraction in the range $1.01 < M(K^+K^-) < 1.03$ GeV to be $F_S = 0.12 \pm 0.03$.

4.6 Bayesian credibility intervals from MCMC studies

The maximum likelihood fit provides the best values of all free parameters, including the signal observables and background model parameters, their statistical

uncertainties and their full correlation matrix.

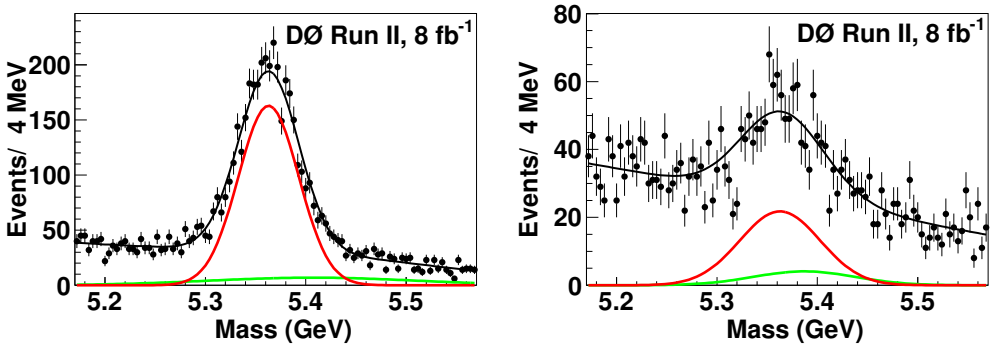


Figure 4.19: Invariant mass distributions of B_s^0 candidates with decay length $ct > 0.02$ cm for $1.01 < M(KK) < 1.03$ GeV (left) and $1.03 < M(KK) < 1.05$ GeV (right). Fits to a sum (black line) of a Gaussian function representing the signal (red), an MC simulation-based template for the $B^0 \to J/\psi K^*$ reflection (green line), and a linear function representing the background are used to extract the B_s^0 yield.

In addition to the free parameters determined in the fit, the model depends on a number of external constants whose inherent uncertainties are not taken into account in a given fit. Ideally, effects of uncertainties of external constants, such as time resolution parameters, flavor tagging dilution calibration, or detector acceptance, should be included in the model by introducing the appropriate parametrized probability density functions and allowing the parameters to vary. Such a procedure of maximizing the likelihood function over the external parameter space would greatly increase the number of free parameters and would be prohibitive. Therefore, as a trade-off, we apply a random sampling of external parameter values within their uncertainties, we perform the analysis for thus created "alternative universes", and we average the results. To do the averaging in the multidimensional space, taking into account non-Gaussian parameter distributions and correlations, we use the MCMC technique.

4.6.1 The method

The MCMC technique uses the Metropolis-Hastings algorithm [46] to generate a sample representative to a given probability distribution. The algorithm generates a sequence of "states", a Markov chain, in which each state depends only on the previous state.

To generate a Markov chain for a given data sample, we start from the best-fit point $\vec{x}$. We randomly generate a point $\vec{x}'$ in the parameter space according to

the multivariate normal distribution $\exp(-(\vec{x}' - \vec{x}) \cdot \Sigma \cdot (\vec{x}' - \vec{x})/2)$, where Σ is the covariance matrix between the best fit current point $\vec{x}$ in the chain and next random point $\vec{x}'$. The best-fit point and the covariance matrix are obtained from a maximum likelihood fit over the same data sample. The new point is accepted if $\mathcal{L}(x')/\mathcal{L}(x) > 1$, otherwise it is accepted with the probability $\mathcal{L}(x')/\mathcal{L}(x)$. The process is continued until a desired number of states is achieved. To avoid a bias due to the choice of the initial state, we discard the early states which may "remember" the initial state. Our studies show that the initial state is "forgotten" after approximately 50 steps. We discard the first 100 states in each chain.

4.6.2 General properties of MCMC chains for the BDT-selection and Square-cuts samples

We generate 8 MCMC chains, each containing one million states: a nominal and three alternative chains each for the BDT-selection and Square-cuts samples, according to the fit results presented in Tables 4.5 and 4.6.

Figures 4.20 and 4.21 illustrate the dependence of $\phi_s^{J/\psi\phi}$ on other physics parameters, in particular on $\cos\delta_\perp$ and $\cos\delta_s$. Each point shows the Markov Chain representation of the likelihood function integrated over all parameters except the parameter of interest in a slice of $\phi_s^{J/\psi\phi}$. For clarity, the profiles are shown for $\Delta\Gamma_s > 0$ and $\Delta\Gamma_s < 0$ separately. The distributions for the Square-cuts sample are similar. We note the following salient features of these correlations for $\Delta\Gamma_s > 0$:

a) A positive correlation between $\phi_s^{J/\psi\phi}$ and ΔM_s, with the best fit of $\phi_s^{J/\psi\phi}$ changing sign as ΔM_s increases.

b) A correlation between $|\phi_s^{J/\psi\phi}|$ and $\overline{\tau}_s$, with the highest $\overline{\tau}_s$ occurring at $\phi_s^{J/\psi\phi} = 0$.

c) For $\phi_s^{J/\psi\phi}$ near zero, $|\Delta\Gamma_s|$ increases with $|\phi_s^{J/\psi\phi}|$.

d) A strong positive correlation between $\phi_s^{J/\psi\phi}$ and $\cos\delta_\perp$ near $\phi_s^{J/\psi\phi} = 0$, with $\phi_s^{J/\psi\phi}$ changing sign as the average $\cos\delta_\perp$ increases between -0.8 and $+0.8$.

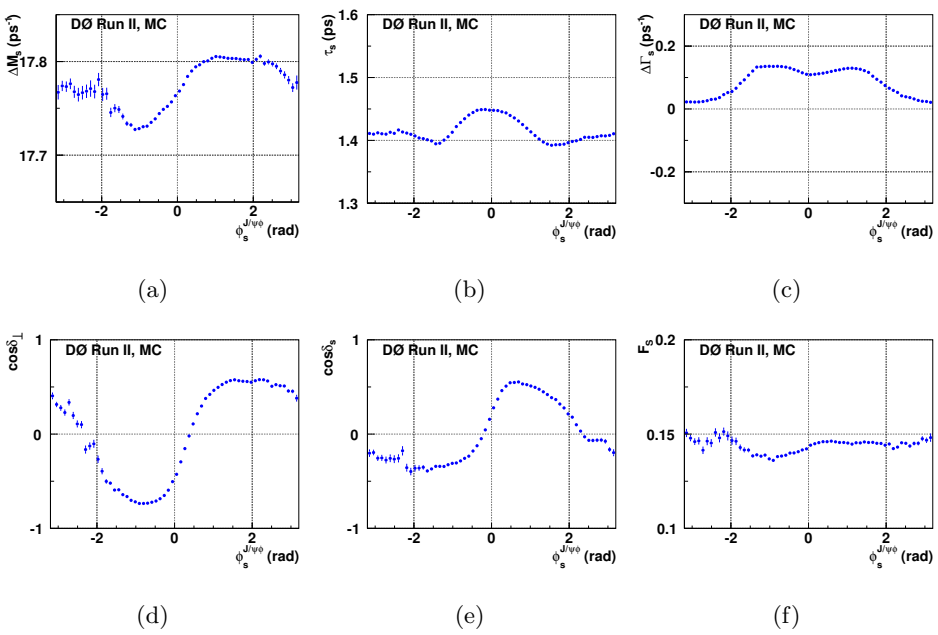


Figure 4.20: Profiles of ΔM_s, $\overline{\tau}_s$, $\Delta\Gamma_s$, $\cos\delta_\perp$, $\cos\delta_s$, and F_S, for $\Delta\Gamma_s > 0$, versus $\phi_s^{J/\psi\phi}$ from the MCMC simulation for the BDT selection data sample.

For the related decay $B_d^0 \to J/\psi K^*$ the measured value is $\cos\delta_\perp = -0.97$. This indicates that a constraint of $\cos\delta_\perp$ to the $B_d^0 \to J/\psi K^*$ value would result in $\phi_s^{J/\psi\phi} < 0$ with a smaller uncertainty.

e) A strong positive correlation between $\phi_s^{J/\psi\phi}$ and $\cos\delta_s$ near $\phi_s^{J/\psi\phi} = 0$, with $\phi_s^{J/\psi\phi}$ changing sign as the average $\cos\delta_s$ increases between -0.4 and $+0.4$.

f) A weak correlation between $\phi_s^{J/\psi\phi}$ and F_S, with F_S a few percent lower for $\phi_s^{J/\psi\phi} < 0$.

While we do not use any external numerical constraints on the polarization amplitudes, we note that the best-fit values of their magnitudes and phases are consistent with those measured in the $U(3)$-flavor related decay $B_d^0 \to J/\psi K^*$ [26], up to the sign ambiguities. Ref. [41] predicts that the phases of the polarization amplitudes in the two decay processes should agree within approximately 0.17 radians. For $\delta_\perp$, our measurement gives equivalent solutions near π and near zero, with only the former being in agreement with the value of 2.91 ± 0.06 measured for $B_d^0 \to J/\psi K^*$ by B factories. Therefore, in the following we limit the range of $\delta_\perp$ to $\cos\delta_\perp < 0$.

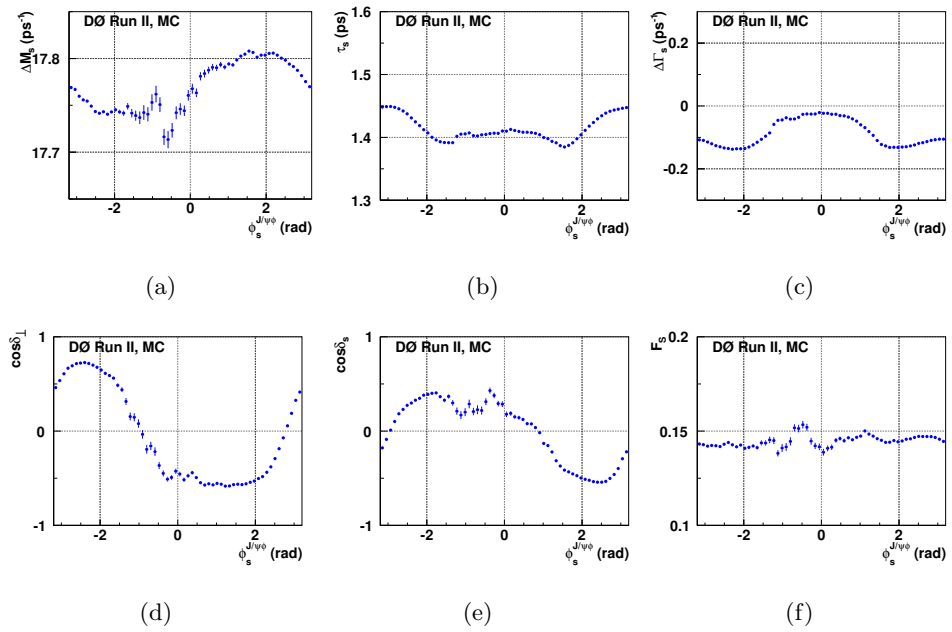
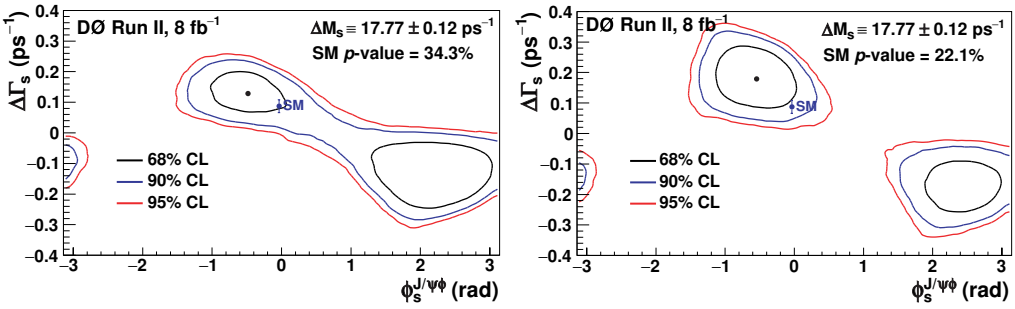


Figure 4.21: Profiles of ΔM_s, $\overline{\tau}_s$, $\Delta\Gamma_s$, $\cos\delta_\perp$, $\cos\delta_s$, and F_S, for $\Delta\Gamma_s < 0$, versus $\phi_s^{J/\psi\phi}$ from the MCMC simulation for the BDT selection data sample.

To obtain the credible intervals for physics parameters, taking into account non-Gaussian tails and systematic effects, we combine the MCMC chains for the nominal and alternative fits. This is equivalent to an effective averaging of the resulting probability density functions from the fits. First, we combine the four MCMC chains for each sample. We then combine all eight chains, to produce the final result. The credible regions are chosen by selecting the highest probability points until the desired confidence level is reached.

4.6.3 Results

Figure 4.22 shows 68%, 90% and 95% credible regions in the ($\phi_s^{J/\psi\phi}$,$\Delta\Gamma_s$) plane for the BDT-based and for the Square-cuts samples. The point estimates of physics parameters are obtained from one-dimensional projections. The minimal range containing 68% of the area of the probability density function defines the one standard deviation credible interval for each parameter, while the most probable value defines the central value.



(a) BDT selection

(b) Square-cuts selection

Figure 4.22: Two-dimensional 68%, 90% and and 95% credible regions for (a) the BDT selection and (b) the Square-cuts sample. The standard model expectation is indicated as a point with an error.

The large correlation coefficient (0.85) between the two phases, $\delta_\perp$ and δ_s, prevents us from making separate point estimates. Their individual errors are much larger than the uncertainty on their difference. For the BDT selection, the measured $\mathcal{S}$-wave fraction $F_S(\text{eff})$ is an effective fraction of the K^+K^- $\mathcal{S}$-wave in the accepted sample, in the mass range $1.01 < M(K^+K^-) < 1.03$ GeV. It includes the effect of the diminished acceptance for the $\mathcal{S}$-wave with respect to the $\mathcal{P}$-wave in the event selection.

This procedure gives the following results for the BDT-based sample:

$$
\begin{aligned}
\overline{\tau}_s &= 1.426^{+0.035}_{-0.032}\ \text{ps}, \\
\Delta\Gamma_s &= 0.129^{+0.076}_{-0.053}\ \text{ps}^{-1}, \\
\phi_s^{J/\psi\phi} &= -0.49^{+0.48}_{-0.40}, \\
|A_0|^2 &= 0.552^{+0.016}_{-0.017}, \\
|A_\parallel|^2 &= 0.219^{+0.020}_{-0.021}, \\
\delta_\parallel &= 3.15 \pm 0.27, \\
\cos(\delta_\perp - \delta_s) &= -0.06 \pm 0.24, \\
F_S(\text{eff}) &= 0.146 \pm 0.035.
\end{aligned}
$$

$F_S(\text{eff})$ in this case refers to the "effective" F_S since it is not a physical parameter: the BDT cut on the phi mass leads to the measurement of F_S in this case

to depend on the efficiency of the selection to non-resonant $B_s^0 \to J/\psi K^+K^-$.

The one-dimensional estimates of physics parameters for the Square-cuts sample are:

$$\begin{aligned}
\overline{\tau}_s &= 1.444^{+0.041}_{-0.033}\ \text{ps}, \\
\Delta\Gamma_s &= 0.179^{+0.059}_{-0.060}\ \text{ps}^{-1}, \\
\phi_s^{J/\psi\phi} &= -0.56^{+0.36}_{-0.32}, \\
|A_0|^2 &= 0.565 \pm 0.017, \\
|A_\parallel|^2 &= 0.249^{+0.021}_{-0.022}, \\
\delta_\parallel &= 3.15 \pm 0.19, \\
\cos(\delta_\perp - \delta_s) &= -0.20^{+0.26}_{-0.27}, \\
F_S &= 0.173 \pm 0.036.
\end{aligned}$$

To obtain the final credible intervals for physics parameters, we combine all eight MCMC chains, effectively averaging the probability density functions of the results of the fits to the BDT- and Square-cuts samples. Figure 4.23 shows 68%, 90% and 95% credible regions in the ($\phi_s^{J/\psi\phi}$,$\Delta\Gamma_s$) plane. The p-value for the SM point [55] $(\phi_s^{J/\psi\phi}, \Delta\Gamma_s) = (-0.038, 0.087\ \text{ps}^{-1})$ is 29.8%. The one-dimensional 68% credible intervals are listed in Section 4.7 below.

4.7 Summary and Discussion

We have presented a time-dependent angular analysis of the decay process $B_s^0 \to J/\psi\phi$. We measure B_s^0 mixing parameters, average lifetime, and decay amplitudes. In addition, we measure the amplitudes and phases of the polarization amplitudes. We also measure the level of the KK $\mathcal{S}$-wave contamination in the mass range (1.01 – 1.03) GeV, F_S. The measured values and the 68% credible intervals,

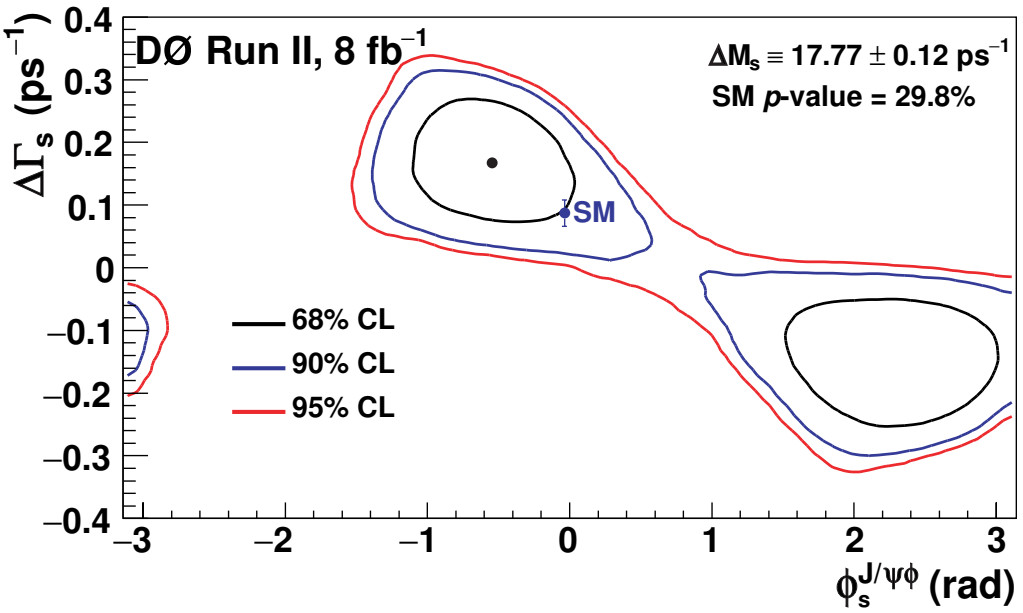


Figure 4.23: Two-dimensional 68%, 90% and 95% credible regions including systematic uncertainties. The standard model expectation is indicated as a point with an error.

including systematic uncertainties, with the oscillation frequency constrained to $\Delta M_s = 17.77 \pm 0.12$ ps^{-1}, are:

$$\begin{aligned}
\overline{\tau}_s &= 1.443^{+0.038}_{-0.035} \text{ ps}, \\
\Delta\Gamma_s &= 0.163^{+0.065}_{-0.064} \text{ ps}^{-1}, \\
\phi_s^{J/\psi\phi} &= -0.55^{+0.38}_{-0.36}, \\
|A_0|^2 &= 0.558^{+0.017}_{-0.019}, \\
|A_\parallel|^2 &= 0.231^{+0.024}_{-0.030}, \\
\delta_\parallel &= 3.15 \pm 0.22, \\
\cos(\delta_\perp - \delta_s) &= -0.11^{+0.27}_{-0.25}. \\
F_S &= 0.173 \pm 0.036,
\end{aligned} \tag{4.5}$$

The p-value for the SM point $(\phi_s^{J/\psi\phi}, \Delta\Gamma_s) = (-0.038, 0.087 \text{ ps}^{-1})$ is 29.8%.

In previous analysis [15] constrained the strong phases to those of $B_d^0 \to J/\psi K^*$ whereas this analysis has a large enough data sample to reliably let them float. Also, the previous publication did not have a large enough data sample to allow

for the measurement of a significant level of KK $\mathcal{S}$-wave, whereas it is measured together with its relative phase in the current analysis. The results supersede our previous measurements.

Independently of the Maximum Likelihood analysis, we make an estimate of the non-resonant K^+K^- in the final state based on the $M(KK)$ distribution of the B_s^0 signal yield.

BIBLIOGRAPHY

[1] DØ Physics Results `http://www-d0.fnal.gov/results/index.html`.

[2] Tevatron Run II Handbook. `http://www-bd/fnal.gov/lug/`.

[3] R. Aleksan, B. Kayser, and D. London. Determining the quark mixing matrix from CP-violating asymmetries. *Phys. Rev. Lett.*, 73:18–20, Jul 1994.

[4] R. Brun and F. Carminati. CERN program library long writeup no. w5013. unpublished, 1993.

[5] A.J. Buras. CP violation and rare decays of K and B mesons. 1999.

[6] A.J. Buras, M.E. Lautenbacher, and G. Ostermaier. Waiting for the top quark mass, $K^+ \to \pi^+ \nu\bar{\nu}$, $b_s^0 \bar{b}_s^0$ mixing, and CP asymmetries in B decays. *Phys. Rev. D*, 50:3433–3446, Sep 1994.

[7] N. Cabibbo. Unitary symmetry and leptonic decays. *Phys. Rev. Lett.*, 10:531–533, Jun 1963.

[8] J. Cecilia and S. Raymond. Unitarity polygons and CP violation areas and phases in the standard electroweak model. *Physics Letters B*, 208(2):268 – 274, 1988.

[9] L. L. Chau and W. Y. Keung. Comments on the parametrization of the kobayashi-maskawa matrix. *Phys. Rev. Lett.*, 53:1802–1805, Nov 1984.

[10] CDF Collaboration. Observation of $B_s^0 - \overline{B}_s^0$ oscillations. *Phys. Rev. Lett.*, 97:242003, Dec 2006.

[11] DØCollaboration. Direct limits on the B_s^0 oscillation frequency. *Phys. Rev. Lett.*, 97:021802, Jul 2006.

[12] DØCollaboration. Measurement of B_d mixing using opposite-side flavor tagging. *Phys. Rev. D*, 74:112002, Dec 2006.

[13] DØCollaboration. The upgraded dø detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 565(2):463 – 537, 2006.

[14] DØCollaboration. Lifetime difference and CP-violating phase in the B_s^0 system. *Phys. Rev. Lett.*, 98:121801, Mar 2007.

[15] DØCollaboration. Measurement of B_s^0 mixing parameters from the flavor-tagged decay $B_s^0 \to J/\psi\phi$. *Phys. Rev. Lett.*, 101:241801, Dec 2008.

[16] SNO Collaboration. Measurement of the rate of $\nu_e + d \to p + p + e^-$ interactions produced by ^8B solar neutrinos at the sudbury neutrino observatory. *Phys. Rev. Lett.*, 87:071301, Jul 2001.

[17] SNO Collaboration. Direct evidence for neutrino flavor transformation from neutral-current interactions in the sudbury neutrino observatory. *Phys. Rev. Lett.*, 89:011301, Jun 2002.

[18] Super-Kamiokande Collaboration. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.*, 81:1562–1567, Aug 1998.

[19] Super-Kamiokande Collaboration. Solar ^8B and hep neutrino measurements from 1258 days of super-kamiokande data. *Phys. Rev. Lett.*, 86:5651–5655, Jun 2001.

[20] A.S. Dighe, I. Dunietz, and R. Fleischer. Extracting CKM phases and B_s–$\overline{B}_s$ mixing parameters from angular distributions of non-leptonic b decays. *The European Physical Journal C - Particles and Fields*, 6:647–662, 1999. 10.1007/s100529800954.

[21] A.S. Dighe, I. Dunietz, H.J. Lipkin, and J.L. Rosner. Angular distributions and lifetime differences in $B_s^0 \to J/\psi\phi$ decays. *Physics Letters B*, 369(2):144 – 150, 1996.

[22] I. Dunietz. $B_s^0\bar{B}_s^0$ mixing, *cp* violation, and extraction of ckm phases from untagged B_s data samples. *Phys. Rev. D*, 52:3048–3064, Sep 1995.

[23] I. Dunietz, R. Fleischer, and U. Nierste. In pursuit of new physics with B_s decays. *Phys. Rev. D*, 63:114015, May 2001.

[24] A. Einstein. Die grundlage der allgemeinen relativitätstheorie. *Ann. Phys.*, 354:769–822, 1916.

[25] A. Einstein. Prinzipielles zur allgemeinen relativitätstheorie. *Ann. Phys.*, 360:241–244, 1918.

[26] K. Nakamura *et al.* (Particle Data Group). Review of particle physics. *Journal of Physics G: Nuclear and Particle Physics*, 37:075021, 2010.

[27] R. P. Feynman. The theory of positrons. *Phys. Rev.*, 76:749–759, Sep 1949.

[28] R. P. Feynman. Mathematical formulation of the quantum theory of electromagnetic interaction. *Phys. Rev.*, 80:440–457, Nov 1950.

[29] R.P. Feynman. Space-time approach to quantum electrodynamics. *Phys. Rev.*, 76:769–789, Sep 1949.

[30] R. Fleischer. Towards extractions of the ckm angle γ from $B_{u,d} \to \pi k$ and untagged $B_s \to k\overline{K}$ decays. *Phys. Rev. D*, 58:093001, Sep 1998.

[31] R. Fleischer. Recent theoretical developments in CP violation in the B system. *Nucl.Instrum.Meth.*, A446:1–17, 2000.

[32] H. Fritzsch, M. Gell-Mann, and H. Leutwyler. Advantages of the color octet gluon picture. *Physics Letters B*, 47(4):365 – 368, 1973.

[33] W. Gautschi. Efficient computation of the complex error function. *SIAM Journal on Numerical Analysis*, 7(1):187–198, 1970.

[34] M. Gell-Mann. A schematic model of baryons and mesons. *Physics Letters*, 8(3):214 – 215, 1964.

[35] H. Georgi and S.L. Glashow. Unified weak and electromagnetic interactions without neutral currents. *Phys. Rev. Lett.*, 28:1494–1497, May 1972.

[36] S.L. Glashow. Partial-symmetries of weak interactions. *Nuclear Physics*, 22(4):579 – 588, 1961.

[37] S.L. Glashow, J. Iliopoulos, and L. Maiani. Weak interactions with lepton-hadron symmetry. *Phys. Rev. D*, 2:1285–1292, Oct 1970.

[38] J. Goldstone. Field theories with superconductor solutions. *Il Nuovo Cimento (1955-1965)*, 19:154–164, 1961. 10.1007/BF02812722.

[39] O.W. Greenberg. Spin and unitary-spin independence in a paraquark model of baryons and mesons. *Phys. Rev. Lett.*, 13:598–602, Nov 1964.

[40] M. Gronau. CP violation and B physics. pages 507–516, 1999.

[41] M. Gronau and J. L. Rosner. Flavor symmetry for strong phases and determination of in. *Physics Letters B*, 669(5):321 – 326, 2008.

[42] D. J. Gross and F. Wilczek. Ultraviolet behavior of non-abelian gauge theories. *Phys. Rev. Lett.*, 30:1343–1346, Jun 1973.

[43] D.J. Gross and F. Wilczek. Asymptotically free gauge theories. I. *Phys. Rev. D*, 8:3633–3652, Nov 1973.

[44] M.Y. Han and Y.Nambu. Three-triplet model with double SU(3) symmetry. *Phys. Rev.*, 139:B1006–B1010, Aug 1965.

[45] P.F. Harrison and H.R. Quinn. The babar physics book. SLAC Report 504 (1998).

[46] W. K. Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57:97, 1970.

[47] P. W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964.

[48] P. W. Higgs. Broken symmetries, massless particles and gauge fields. *Physics Letters*, 12(2):132 – 133, 1964.

[49] P. W. Higgs. Spontaneous symmetry breakdown without massless bosons. *Phys. Rev.*, 145:1156–1163, May 1966.

[50] C. Jarlskog. Commutator of the quark mass matrices in the standard electroweak model and a measure of maximal CP nonconservation. *Phys. Rev. Lett.*, 55:1039–1042, Sep 1985.

[51] S. Khalil and E. Torrente-Lujan. Neutrino mass and oscillation as probes of physics beyond the standard model. *J.Egyptian Math.Soc.*, 9:91–141, 2001.

[52] M. Kobayashi and T. Maskawa. CP-violation in the renormalizable theory of weak interaction. *Progress of Theoretical Physics*, 49(2):652–657, 1973.

[53] D. J. Lange. The EvtGen particle decay simulation package. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 462(1–2):152 – 155, 2001. ¡ce:title¿BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines¡/ce:title¿.

[54] A. Lenz and U. Nierste. Theoretical update of $B_s - \bar{B}_s$ mixing. *JHEP*, 0706:072, 2007.

[55] A. Lenz and U. Nierste. Numerical updates of lifetimes and mixing parameters of B mesons. 2011.

[56] Y. Nambu and G. Jona-Lasinio. Dynamical model of elementary particles based on an analogy with superconductivity. I. *Phys. Rev.*, 122:345–358, Apr 1961.

[57] Y. Nambu and G. Jona-Lasinio. Dynamical model of elementary particles based on an analogy with superconductivity. II. *Phys. Rev.*, 124:246–254, Oct 1961.

[58] Y. Nir. CP violation in and beyond the standard model. 1999.

[59] D. H. Politzer. Reliable perturbative results for strong interactions. *Phys. Rev. Lett.*, 30:1346–1349, Jun 1973.

[60] P. Ramond. Neutrinos: A glimpse beyond the standard model. *Nuclear Physics B - Proceedings Supplements*, 77(1–3):3 – 9, 1999.

[61] M. Stephen S. Torbjörn and S. Peter. Pythia 6.4 physics and manual. *Journal of High Energy Physics*, 05:26, 2006.

[62] A. Salam and J.C. Ward. Electromagnetic and weak interactions. *Physics Letters*, 13(2):168 – 171, 1964.

[63] J. Schwinger. Quantum electrodynamics. I. a covariant formulation. *Phys. Rev.*, 74:1439–1461, Nov 1948.

[64] J. Schwinger. Quantum electrodynamics. II. vacuum polarization and self-energy. *Phys. Rev.*, 75:651–679, Feb 1949.

[65] J. Schwinger. Quantum electrodynamics. III. the electromagnetic properties of the electron¯radiative corrections to scattering. *Phys. Rev.*, 76:790–817, Sep 1949.

[66] V. Shiltsev. Status of tevatron collider run II and novel technologies for luminosity upgrades. Proc. 2004 European Particle Accelerator Conference, Lucern, Switzerland, Vol. 1, 2004, p. 239.

[67] S. Stone and L. Zhang. S-waves and the measurement of CP violating phases in B_s decays. *Phys. Rev. D*, 79:074024, Apr 2009.

[68] G. 't Hooft and M. Veltman. Regularization and renormalization of gauge fields. *Nuclear Physics B*, 44(1):189 – 213, 1972.

[69] The ROOT team. ROOT. `http://root.cern.ch`.

[70] B.S. Acharya *et al.* Scintillation counters for the DØ muon upgrade. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 401(1):45 – 62, 1997.

[71] M. Bona *et. al.* The Unitarity Triangle Fit in the Standard Model and Hadronic Parameters from Lattice QCD: A Reappraisal after the Measurements of Δm_s and $BR(B \to \tau\nu_\tau)$). *JHEP*, 0610:081, 2006.

[72] S. Abachi *et al.* The dØ detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 338(2–3):185 – 253, 1994.

[73] S. Tomonaga. On a relativistically invariant formulation of the quantum theory of wave fields. *Progress of Theoretical Physics*, 1(2):27–42, 1946.

[74] W. Verkerke and D. Kirbky. The roofit toolkit for data modeling. `http://roofit.sourceforge.net`.

[75] S. Weinberg. A model of leptons. *Phys. Rev. Lett.*, 19:1264–1266, Nov 1967.

[76] L. Wolfenstein. Parametrization of the kobayashi-maskawa matrix. *Phys. Rev. Lett.*, 51:1945–1947, Nov 1983.

[77] Y. Xie, P. Clarke, G. Cowan, and F. Muheim. Determination of $2\beta_s$ in $B_s^0 \to J/\psi K^+K^-$ Decays in the Presence of a K^+K^- S-Wave Contribution. *JHEP*, 0909:074, 2009.