Precision Measurements of Lifetime and CP Violation Phase in the B_s^0 Meson with the DØ Detector

Jorge Martínez Ortega

Department of Physics

Centro de Investigación y de Estudios Avanzados del IPN

A thesis submitted for the degree of

Doctor of Science
Speciality in Physics

Supervised by Dr. Alberto Sánchez Hernández

Mexico City October 2012

Resumen

El Modelo Estándar (SM) de Física de Partículas es probablemente la teoría más exitosa, en cuanto a la medición de sus predicciones.Sin embargo, la predicción del SM para violación de simetría CP no es suficiente para explicar la abrumadora asimetría que existe en la abundancia de materia y antimateria. Medir algún proceso donde la violación de CP sea mayor a la predicha por el SM sería un indicio claro de Física más allá del Modelo Estándar. La predicción del SM para la fase de violación de CP, ϕ_s , en el mesón B_s^0 es prácticamente cero para la presición de los experimentos actuales. Es decir, medir una desviación de cero en ϕ_s podría ser un indicio de Física más allá del SM.

Por otro lado, la aproximación basada en la Simetría del quark pesado permite hacer cálculos aproximados de las cantidades fundamentales de hadrones que contienen un quark pesado, c, b, t. Dichos cálculos se realizan como expansiones en términos de potencias inversas del quark más pesado del hadrón. Este formalismo se llama "Teoróa Efectiva del Quark Pesado" (Heavy Quark Effective Theory, HQET) y ha sido exitoso en la predicción de algunas cantidades en los hadrones pesados. La predicción de HQET para el cociente de los tiempos de vida de los mesones B_d^0 y B_s^0 es para términos prácticos igual a la unidad. Por lo tanto, medir este con una buena precisión es otra forma de poner a prueba una aproximación basada en el SM.

En está tesis se presenta en detalle el procedimiento para la medición de ϕ_s y el cociente de los tiempos de vida de los mesones B_d^0 y B_s^0, entre otras cantidades, con el detector DØ, ubicado en el Fermi National Accelerator Laboratory, en los Estados Unidos.

Abstract

The Standard Model of Partice Physics (SM) is probably the most successful theory, regarding to his predictions. The SM prediction for CP violation is not enough to explain the overwhelming asymmetry among the matter and anti-matter abundance. Measuring some process where CP violation is different to the one predicted by the SM would be a clear signal for Physics Beyond the Standard Model. The SM prediction for the CP violation phase, ϕ_s, in the B_s^0 meson is practically equal to zero for the current experiments. This means that measuring a deviation from zero in ϕ_s could be an indication for Physics Beyond the SM.

On the other hand, the approximation based on the "heavy quark symmetry" let approximated calculations of the fundamental quantities of those hadrons containing a heavy quark, c, b, t. These calculations are expressed as expansions on inverse powers of the heavy quark mass in such hadron. This formalism is called " Heavy Quark Effective Theory" (HQET), and has been successful predicting some properties of the heavy hadrons. The HQET prediction for the lifetime ratio the B_d^0 and B_s^0 is practically equal to one. So, measuring with good precision the B_s^0 lifetime is also a way to test an approximation based on the SM.

In this thesis it is detailed presented the method to measure the ϕ_s and the lifetime ratio of the B_d^0 and B_s^0, among other quantities, with the DØ located in the Fermi National Accelerator Laboratory, in the United States.

To my beloved son, Francisco Ernesto

Acknowledgements

There are many people and institutions that I would want to thank, hope that I do not forget any.

First I want to thank Alberto Sánchez Hernándes, for his patience help along my PhD studies.

I am thankful to Heriberto Castilla Valdez, who was always ready to help me, not only academically.

To Eduard de la Cruz and Pedro Podesta Lerma, they were always also in a good mood to teach me any things.

To Rosemary Ovando, Elizabeth Mote, Flor Ibáñez, Patricia Villar, Sonya Wright, Julie Saviano, Jackie Cyko and Cheryl Bentham, for their patience help with all the paperwork at Cinvestav and Fermilab.

To CONACyT, that funded my education.

To the Fermi Research Alliance, LLC, and the Universities Research Association, Inc., that made possible my visit to Fermilab.

To Alejandro Garca, Jesus Hernández, Enrique Camacho, Estela Garcés, Ricardo Magaña, Andrés García, from whom I learned many different things, not only physics or programming.

To all my colleagues at the DØ experiment, specially to George Ginther, Dmitri Denisov, Gregorio Bernardi, Stefan Soldner, Darien Wood, William Lee, Geoff Savage, Brendan Casey, Maiko Takahashi, Sung Woo Youn, Andreas Jung, Nirmalya Parua, Marco Verzocchi, Aurelio Juste, Mark Williams, Penelope Kasper, Marjorie Corcoran, Dmitry Tsybychev, Avdhesh Chandra and Daria Zieminska.

To all my friends at Cinvestav, specially to Alonso Contreras, David Bermúdez, Aldrin Cervantes and Isaac Martínez.

To Laura Robles, Liliana Hernández, Roger Hernández, Federico Sforza, Ritoban Basu Thakur, Vito Lombardo, Juan Pablo Velásquez, Oscar Gonzalez, José Palomino, Ruth Ponce, Cesar Castromonte, Luisa Dos Anjos, Angelo Di Canto, Viviana Cavaliere, Michelle Prewitt, Emmanuel Munyangabe, Jesse Green, Pierfrancesco Butti, Minerba Betabcourt and all the good friends with whom I shared happy moments at Fermilab.

To Liliana Toscano, things not always happen as we wish, but I will never forget your invaluable help during many years.

To my parents, Carmen Ortega and Jorge Martinez, I have not enough words to thank them for everything.

To my sister, Carmen Beatriz, and my friends, Yoloxóchitl Corona, Ulises Moya, Goretti Rodríguez, Alma García, María Teresa Gutiérred, María Elena Pacheco, Orietta Sánchez, Martín Ortiz, Paulino Monroy, Pablo López, Renato Arroyo, Rodrigo Castro, Eduardo Razón that helped me always I needed a word of support.

To Gypsy Torres, that in few months had helped me more than she might think.

Contents

List of Figures

LIST OF FIGURES

List of Tables

LIST OF TABLES

1. Introduction

1.1 Short History of Particle Physics

The High Energy Physics (HEP) or Elementary Particle Physics is one of the most active fields in science. The field started in its modern form with the discovery of the electron in 1897 by J. J. Thomson and his team [3]. This gave the first evidence that the atom was not indivisible. During the next two decades some experiments carried out by Ernest Rutherford and his team demonstrated the existence of atomic nucleus in 1909, and in 1919 found the first evidence of the proton [4–7].

By 1930 there existed a discrepancy in β decays, apparently the energy and the momentum were not conserved in these processes. In 1930 the neutrino, ν, was postulated first by Wolfgang Pauli to solve the issue [8]. Pauli hypothesized an undetected particle that he called a "neutron" in keeping with convention employed for naming both the proton and the electron, which in 1930 were known to be respective products for α and β decay. James Chadwick discovered a much more massive nuclear particle in 1932 and also named it a neutron, leaving two kinds of particles with the same name. Enrico Fermi, who developed the theory of beta decay, coined the term neutrino (the Italian equivalent of "little neutral one") in 1933 as a way to resolve the confusion. The detection of a neutrino had to wait until 1956. The proton - neutron model of nucleus was proposed in 1932. In this model almost all of the mass of an atom is located in the nucleus, with a very small contribution from the electrons cloud.

However, a new kind of interaction had to be introduced to explain the mechanism that is responsible for preventing that the nucleus tears itself apart under the action of the electrostatic force. The Strong interaction was then postulated. In 1934 Hideki Yukawa predicted the existence and the approximate mass of the "meson" as the carrier

of the nuclear force that holds atomic nuclei together [9]. The first meson candidate, then known as the"μ meson" (today known as muon) was discovered 1936 by Carl David Anderson and Seth Henry Neddermeyer in the decay products of cosmic ray interactions [10].

The muon had about the right mass to be Yukawa candidate for a carrier of the strong nuclear force, but over the course of the next decade, it became evident that it was not the right particle. It was eventually found that the muon did not participate in the strong nuclear interaction at all, but rather behaved like a heavy version of the electron, and is in fact a lepton rather than a meson. The first true meson to be discovered was the "π meson" (or pion) in 1947 [11, 12], as a product of cosmic rays. It also had about the right mass predicted by Yukawa [13], and over the next few years, more experiments showed that the pion was indeed involved in strong interactions, however, the Yukawa theory was intended to explain the β decay, that not not associated with the strong interaction.

The cosmic rays also helped to discover an essential ingredient of the modern theory, the antimatter. In 1930, Paul Dirac postulated the existence of this new kind on matter. He realized that the Klein-Gordon equation had negative "unphysical" solutions. In order to solve this "inconsistency" Dirac derived a linear equation to ascribe correctly the Quantum Mechanics at the relativistic domain. The Dirac equation, however, also has "negative energy solutions"[14]. Dirac interpreted the negative energy solutions as holes in the "sea" (now commonly referred as Dirac sea) that must behave as opposite charge particles, the antimatter. In 1932 the first antiparticle was discovered, the positron, the antiparticle of the electron [15].

Soon after the discovery of the pion, in 1947, a new kind of "strange particle" was discovered, the kaon, or K meson. In 1947, G. D. Rochester and Clifford Charles Butler of the University of Manchester published two cloud chamber photographs of cosmic ray-induced events, one showing what appeared to be a neutral particle decaying into two charged pions, and one which appeared to be a charged particle decaying into a charged pion and something neutral [16]. The mass of the new particles was about half the mass of the proton. More examples of these "V-particles" were slow in coming.

The first breakthrough was obtained at Caltech, where a cloud chamber was set on the top of the Mount Wilson, to increase cosmic ray exposure. During the decade

of 1950, plenty charged and neutral V-particles were reported [17]. Inspired by this, numerous observations were made over the next several years. The terminology used was, "K meson" for a particle intermediate in mass between the pion and nucleon and "Hyperon" meant any particle heavier than a nucleon. Typical, the decays were extremely slow compared to production reaction, lifetimes are of the order of 10^{-10} s, while production in pion-proton reactions proceeds with a time scale of 10^{-23} s. The problem of this mismatch was solved indepently by Abraham Pais, Murray Gell-Mann and Kazuhiko Nishijima. They postulated the new quantum number called "strangeness" which is conserved in strong interactions but violated by the weak interactions. Strange particles appear copiously due to "associated production" of a strange and an antistrange particle together. It was soon shown that this could not be a multiplicative quantum number, because that would allow reactions which were never seen in the new synchrotrons which were commissioned in Brookhaven National Laboratory in 1953 and in the Lawrence Berkeley Laboratory in 1955 [18].

The development of these new particle accelerators and particle detectors in the 1950s led to the discovery of a huge variety of hadrons, prompting Wolfgang Pauli's remark: "Had I foreseen this, I would have gone into botany". The term particle zoo was used colloquially to describe the relatively extensive list of the known "elementary" particles that almost look like hundreds of species in the zoo. The situation was particularly confusing in the late 1960s, before the discovery of quarks, when hundreds of strongly interacting particles (hadrons) were known.

1.2 Standard Model

The Standard Model (SM) is the theoretical framework capable to explain the experimental facts mentioned in previous section. The evidence shows that everything in the Universe is made from twelve basic building blocks, called fundamental particles, governed by four fundamental interactions. Our best understanding of how these twelve particles and three of the interactions are related to each other is encapsulated in the Standard Model of particles and forces. Developed during the 1960s by Sheldon Glashow, Steven Wienberg, and Abdus Salam, it became accepted after the discovery at the European Organization for Nuclear Research (CERN) in 1973 of the neutral weak

currents caused by Z boson boson exchange [19–21]. The SM has successfully explained a host of experimental results and precisely predicted a wide variety of phenomena.

All the material objects around us are made of matter particles. These matter particles occur in two basic types called quarks and leptons. Each group consists of six particles, which are related in pairs, or generations. The lightest and most stable particles make up the first generation, whereas the heavier and less stable particles belong to the second and third generations. All stable matter in the Universe is made from particles that belong to the first generation, any heavier particles quickly decay to the next most stable level.

The six quarks are paired in the three generations the 'up quark' and the 'down quark' form the first generation, followed by the 'charm quark' and 'strange quark', then the 'top quark' and 'bottom quark'. The six leptons are similarly arranged in three generations the 'electron' and the 'electron-neutrino', the 'muon' and the 'muon-neutrino', and the 'tau' and the 'tau-neutrino'. The electron, the muon and the tau all have a negative electric charge and a mass, whereas the neutrinos are electrically neutral with very small mass.

There are four fundamental interactions at work in the Universe: the strong interaction, the weak interaction, the electromagnetic force, and the gravitational force. They work over different ranges and have different intensity. Gravity is the weakest but it has an infinite range. The electromagnetic force also has infinite range but it is many orders of magnitude stronger than gravity. The weak and strong forces are effective only over a very short range and dominate only at the level of subatomic particles. Despite its name, the weak interaction is much stronger than gravity but it is indeed the weakest of the other three. The strong interaction is, as the name says, the strongest among all the four fundamental interactions.

We know that three of the fundamental interactions result from the exchange of force carrier particles, or 'gauge bosons'. Matter particles transfer discrete amounts of energy by exchanging bosons with each other. Each fundamental force has its own corresponding boson particle the strong interaction is carried by the gluons, the electromagnetic force is carried by the photon, and the W and Z bosons are responsible for the weak force.

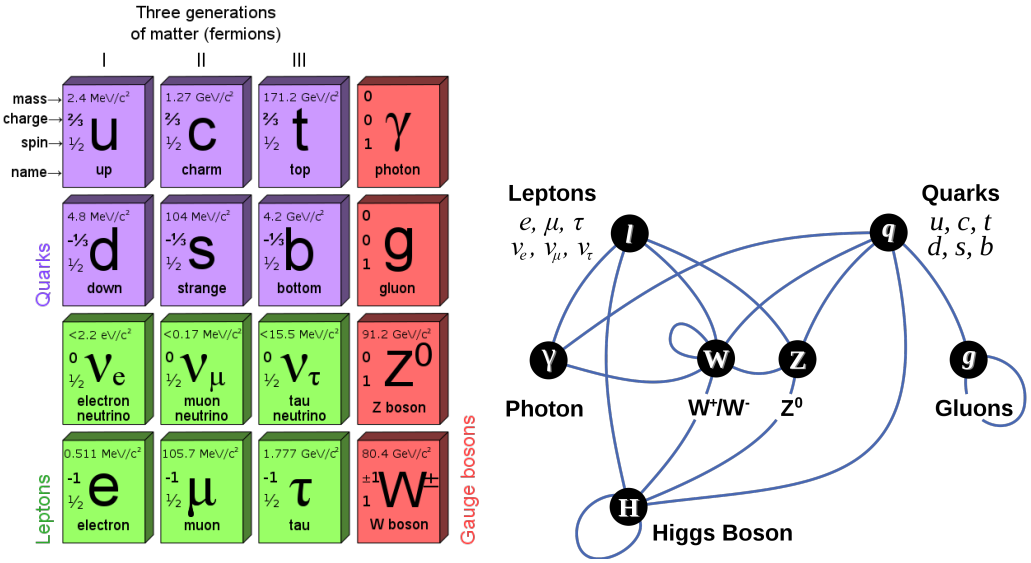


Figure 1.1: Elementary particles and interactions with the gauge bosons.

So despite the effectiveness of the Standard Model at describing the phenomena within its domain, it is nevertheless incomplete. Probably the most shocking contradiction of the Standard Model with reality is that we live in a Universe made from matter. There is no experimental evidence that there are any significant concentrations of antimatter in the observable Universe. There are two main interpretations for this disparity: either the universe began with a small preference for matter (total baryonic number of the universe different from zero), or the universe was originally perfectly symmetric, but somehow a set of phenomena contributed to a small imbalance in favor of matter over time. Assuming the second point of view, the are three necessary conditions, known as Sakharov conditions, to obtain that imbalance, which one of these is CP violation. The Standard Model allows CP violation trough Cabibbo-Kobayashi-Maskawa (CKM) and Pontecorvo-Maki-Nakagawa-Sakata matrices, but these mechanisms seem to be not enough to explain matter - antimatter imbalance observed in the Universe.

1.3 B_s^0 Meson, CP violation and lifetime

In 1964, James Cronin, Val Fitch with coworkers provided clear evidence that CP-symmetry could be broken [22]. This discovery showed that weak interactions violate not only the charge-conjugation symmetry C between particles and antiparticles and the P or parity, but also their combination. The discovery shocked particle physics and opened the door to questions still at the core of particle physics and of cosmology today. The lack of an exact CP-symmetry, but also the fact that it is so nearly a symmetry, created a great puzzle.

In 2001, a new generation of experiments, including the BaBar Experiment at the Stanford Linear Accelerator Center (SLAC) and the Belle Experiment at the High Energy Accelerator Research Organisation (KEK) in Japan, observed direct CP violation in a different system, namely in decays of the B mesons. By now a large number of CP violation processes in B meson decays have been discovered. Before these "B-factory" experiments, there was a logical possibility that all CP violation was confined to kaon physics.

As mentioned in the previous section, the universe is made practically only of matter, rather than consisting of equal parts of matter and antimatter as might be expected. It can be demonstrated that, to create an imbalance in matter and antimatter from an initial condition of balance, the Sakharov conditions must be satisfied:

- Baryonic number violation,
- CP violation,
- Interactions out of thermal equilibrium.

Explanations which do not involve CP violation are less plausible, since they rely on the assumption that the matterantimatter imbalance was present at the beginning, or on other admittedly exotic assumptions. The Big Bang should have produced equal amounts of matter and antimatter if CP-symmetry was preserved. If that was the case, there should have been total cancellation of both, protons should have cancelled with antiprotons, electrons with positrons, neutrons with antineutrons, and so on. This would have resulted in a sea of radiation in the universe with no matter. Since this is

not the case, after the Big Bang, physical laws must have acted differently for matter and antimatter, *i.e.* violating CP-symmetry.

The B_s^0 and $\overline{B}_s^0$ mesons are in fact a superpositions of two different states with well defined masses and decay rates, and then different lifetimes. They are generally produced in strong interactions, but they decay weakly. The $B_s^0 - \overline{B}_s^0$ mesons can be decomposed also as a combination of states invariant[1] under a CP transformation. It can happen that CP-symmetry is broken in the $B_s^0 - \overline{B}_s^0$ system.

The Standard Model contains only two ways to break CP-symmetry. The first of these is in the QCD Lagrangian, and has not been found experimentally. But in the scenario, it would be expected this to lead to either no CP violation or a CP violation that is many, many orders of magnitude too large. The second of these, involving the weak interaction, has been experimentally verified, but can account for only a small portion of CP violation. It is predicted to be sufficient for a net mass of normal matter equivalent to only a single galaxy in the known universe. In this context, finding a process that present CP violation larger that the SM prediction would be an undoubtedly evidence of "New Physics". The CP violation in the B_s^0 system is expected to be rather small. Measuring a CP violation effect in this system would be a clear evidence of New Physics.

The Heavy Quark Effective Theory (HQET), that is in fact a perturbative approximation of the exact theory, has been very successful on the prediction of mass differences and lifetime ratios of the heavy mesons (those containing charm or bottom quarks). The HQET predicts that the B_s^0 lifetime is practically equal to the B_d^0 lifetime. The B_d^0 lifetime has been measured with a precision of 1% [23]. Then, the HQET can be tested measuring the B_s^0 lifetime with a good precision.

1.4 Tevatron, D0 and LHC

The Tevatron, the worlds highest-energy proton-antiproton collider, shut down on Sept. 30, 2011. It is located in the Fermi National Accelerator Laboratory, Fermilab, located 40 miles west to Chicago, Illinois. Since 1983, the most powerful atom smasher of the United States has created particle collisions and provided particle beams to fixed

[1] Invariant up to a phase factor.

target experiments and test beam areas. The Tevatron has informed some of the most important fundamental discoveries of our time, such as the existence of the top quark and five baryons, which helped to test and refine the Standard Model of particle physics and shape our understanding of matter, energy, space and time. The Tevatron research program also yielded countless achievements in detector, accelerator and computing technology.

During almost three decades the Tevatron provided collisions for the general propose detectors CDF and DØ, among other experiments. In particular the DØ experiment has contributed with many physics results during recent years, and it will continue analyzing the collected data during the next few years. One of the most exciting results from DØ, that suggest a CP violation effect not explained by the SM is the "Anomalous like-sign dimuon charge asymmetry" [24, 25].

The measurements in this work were developed in the DØ experiment. The DØ experiment is located at one of the interaction regions, where proton and antiproton beams intersect, on the Tevatron synchrotron ring, labelled 'DØ'. DØ is an international collaboration of several hundred of physicists from around 90 universities and national laboratories from 18 countries.

It is relevant to mention the Large Hadron Collider (LHC), now days the most energetic collider, that is located near Geneva, where it spans the border between Switzerland and France about 100 m underground. It is expected that the LHC will help answer some of the fundamental open questions in physics. Many theorists expect New Physics beyond the Standard Model to emerge at the TeV energy scale.

1.5 Overview of the Thesis

The thesis is about the experimental measurement of the B_s^0 system properties, in particular CP violation parameters and lifetime. In chapter two there is a short overview of the theoretical framework to understand how the Standard Model constrains these quantities. In chapter three and four I describe the experimental apparatus necessary to develop these measurement, *i.e.* the Tevatron Collider and the D0 detector. The method to measure the B_s^0 lifetime is detailed written in chapter five. The measure-

ment of CP violation phase ϕ_s is also detailed described in chapter six. Finally, the conclusions of this work are summarized in the chapter seven.

2. Theoretical Framework

2.1 Introduction

The theoretical Framework used to understand the fundamental interactions in nature is known as The Standard Model (SM). The SM is a Quantum Theory of Gauge Fields, this means that both particles and interactions are understood as quantum fields and that at certain scale (energy) the theory has manifested some gauge symmetries. The gauge symmetries in the SM are described with the group $SU(3)\times SU_L(2)\times U_Y.(1)$. The $SU_L(2)\times U_Y(1)$ symmetry of the theory is broken at low energies. This broke up is an essential part of the SM, the responsible ingredient of this effect is the Higgs boson [26]. After the spontaneous symmetry breaking of $SU_L(2)\times U_Y(1)$, only a $U_{EM}(1)$ symmetry remains.

2.2 The Electroweak Model

The SM is a gauge theory based on the symmetry group $SU(3)\times SU_L(2)\times U_Y(1)$ [27]. The $SU(3)$ group describes the strong color interactions among the quarks, and the $SU_L(2)\times U_Y(1)$ group describes the electroweak interactions, which are carried out via the corresponding spin-1 gauge fields: eight massless gluons and one massless photon for the strong and the electromagnetic interactions, respectively, and three massive bosons, $W^{\pm}$ and Z^0, for the weak interaction. At the present time, three generations of quarks and leptons have been observed. The measured width of the Z boson does not permit a fourth generation with an active light neutrino [28]. The matter fields in the minimal SM are three families of quarks and leptons having spin $s = 1/2$, and a spin-zero Higgs boson, shown in Table 2.1. The index $i = 1, 2, 3$, on the fermion fields is

Field	Dim($SU(3)$)	Dim($SU(2)$)	$Y(U(1))$	Lorentz
$Q_L^i = \begin{pmatrix} u_L^i \\ d_L'^i \end{pmatrix}$	**3**	**2**	$1/6$	$(1/2, 0)$
u_R^i	**3**	**1**	$2/3$	$(0, 1/2)$
d_R^i	**3**	**1**	$-1/3$	$(0, 1/2)$
$L_L^i = \begin{pmatrix} \nu_L^i \\ \ell_L^i \end{pmatrix}$	**1**	**2**	$-1/2$	$(1/2, 0)$
ℓ_R^i	**1**	**1**	-1	$(0, 1/2)$
$H_L^i = \begin{pmatrix} H^+ \\ H^0 \end{pmatrix}$	**1**	**2**	$1/2$	$(0, 0)$

Table 2.1: Matter fields in the standard model.

a family or generation index, and the subscripts L and R denote left- and right-handed fields, respectively,

$$\psi_L = P_L\psi, \qquad \psi_R = P_R\psi, \tag{2.1}$$

where P_L and P_R are the projection operators

$$P_L = \frac{1-\gamma^5}{2}, \qquad P_R = \frac{1+\gamma^5}{2}. \tag{2.2}$$

Q_L^i, u_R^i, d_L^i are the quark fields and L_L^i, e_R^i are the lepton fields[1]. The particles associated with the matter fields are summarized in Table 2.1. The $SU(2) \times U(1)$ symmetry is spontaneously broken by the vacuum expectation value (v.a.v.) of the Higgs doublet H. The spontaneous breakdown of $SU_L(2) \times U_Y(1)$ gives mass to the $W^\pm$ and Z^0 gauge bosons. A single Higgs doublet is the simplest way to achieve the observed pattern of

[1] The γ^5 matrix is defined as $\gamma^5 = \frac{i}{4!}\varepsilon_{\alpha\beta\mu\nu}\gamma^\alpha\gamma^\beta\gamma^\mu\gamma^\nu$, where $\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}I$

spontaneous symmetry breaking, but a more complicated scalar sector, such as two doublets, is possible.

The terms in the SM Lagrangian density that involve only the Higgs doublet,

$$H = \begin{pmatrix} H^+ \\ H^0 \end{pmatrix}, \tag{2.3}$$

are

$$\mathscr{L} = (D_\mu H)^\dagger (D_\mu) - V(H), \tag{2.4}$$

where D_μ is the covariant derivative and $V(H)$ is the Higgs potential,

$$V(H) = \frac{\lambda}{4}(H^\dagger H - \frac{v^2}{2})^2. \tag{2.5}$$

The Higgs potential is minimized when $H^\dagger H = v^2/2$. The $SU(2) \times U(1)$ symmetry can be used to rotate a general v.a.v. into the standard form

$$\langle H \rangle = \begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix}. \tag{2.6}$$

Before the spontaneous symmetry breaking the gauge bosons are $W^i_\mu, i = 1, 2, 3$, and B_μ for $SU(2)$ and $U(1)$ respectively, and the corresponding gauge coupling constants g and g'. The left-handed fermion fields, as summarized in Table 2.1, transform as doublets:

$$\Psi^i = L^i_L = \begin{pmatrix} \nu^i_L \\ \ell^i_L \end{pmatrix} \text{ and } Q^i_L = \begin{pmatrix} u^i_L \\ d'^i_L \end{pmatrix}, \quad d'^i_L \equiv \sum_{j=1}^{3} V_{ij} d^j_L, \quad i = 1, 2, 3, \tag{2.7}$$

where V is de Cabibbo-Kobayashi-Maskawa mixing matrix.

After the symmetry breaking, the Lagrangian for the fermion fields, $\psi^i = u^i_\alpha, d^i_\alpha, \nu^i, \ell^i_\alpha$, $\alpha = L, R$, is

$$\begin{aligned} \mathscr{L}_F &= \sum_k \overline{\psi}^k \left(i \not{\partial} - m_k - \frac{g m_k H}{2 M_W} \right) \psi^k \\ &\quad - \frac{g}{2\sqrt{2}} \sum_k \overline{\Psi}^k \gamma^\mu (1 - \gamma^5)(T^+ W^+_\mu + T^- W^-_\mu) \Psi^k \\ &\quad - e \sum_k q_k \overline{\psi}^k \gamma^\mu \psi^k A_\mu \\ &\quad - \frac{g}{2 \cos\theta_W} \sum_k \overline{\psi}^k \gamma^\mu (g^k_V - g^k_A \gamma^5) \psi^k Z_\mu. \end{aligned} \tag{2.8}$$

m_k s the mass of the fermion field ψ^k, M_W is the W boson mass, $T^{\pm}$ are the isospin raising and lowering operators, $W^{\pm} \equiv (W^1 \pm iW^2)/\sqrt{2}$ are the charged weak boson fields, $\theta_W = \arctan(g'/g)$ is the weak angle, $e = g\sin\theta_W$ is the positron electric charge, q_k is the charge of ψ^k in units of e, $\gamma = A \equiv B\cos\theta_W + W^3\sin\theta_W$ is the photon field, $Z \equiv -B\sin\theta_W + W^3\cos\theta_W$ is the neutral weak boson field, and $g_V^k \equiv t_{3L}(k) - 2q^k sin^2\theta_W$, $g_A^k \equiv t_{3L}$ are the vector and axial-vector couplings, where $t_{3L}(k)$ is the weak isospin of fermion ψ^k ($+1/2$ for u^k and ν^k; $-1/2$ for d^k and ℓ^k).

The boson masses are (at tree level, *i.e.*, to lowest order in perturbation theory),

$$M_H = \lambda v, \tag{2.9a}$$

$$M_W = \frac{1}{2}gv = \frac{ev}{2\sin\theta_W}, \tag{2.9b}$$

$$M_Z = \frac{1}{2}\sqrt{g^2 + g'^2}v = \frac{M_W}{cos\theta_W}, \tag{2.9c}$$

$$M_\gamma = 0. \tag{2.9d}$$

The first term in $\mathscr{L}$ contains the Yukawa couplings an gives rise to the fermion masses. The second term represents the charged-current weak interactions. The third term describes the electromagnetic interaction (QED) and the last term is the weak neutral-current interaction.

2.3 The CKM Matrix

Within the Standard Model, CP symmetry is broken by complex phases in the Yukawa couplings (that is, the couplings of the Higgs scalar to quarks). When all manipulations to remove unphysical phases in this model are exhausted, one finds that there is a single CP-violating parameter [29]. In the basis of mass eigenstates, this single phase appears in the 3×3 unitary matrix that gives the W -boson couplings to an up-type antiquark and a down-type quark. (If the Standard Model is supplemented with Majorana mass terms for the neutrinos, the analogous mixing matrix for leptons has three CP-violating phases.) The beautifully consistent and economical Standard-Model description of CP violation in terms of Yukawa couplings, known as the Kobayashi-Maskawa (KM) mechanism [29], agrees with all measurements to date. Furthermore, one can fit the data allowing new physics contributions to loop processes to compete with, or even

dominate over, the Standard Model ones [30, 31]. Such an analysis provides a model-independent proof that the KM phase is different from zero, and that the matrix of three-generation quark mixing is the dominant source of CP violation in meson decays.

The masses and mixing of quarks have a common origin in SM. They arise from the Yukawa interactions of the Higgs condensate with the fermions,

$$\mathscr{L}_Y = -Y^d_{ij}\overline{Q'}^i_L \phi d'^j_R - Y^u_{ij}\overline{Q'}^i_L \epsilon\phi^* u'^j_R + \text{h.c.}, \tag{2.10}$$

where $Y^{u,d}$ are 3×3 complex matrices, ϕ is the Higgs field, i, j are the generation labels, and ϵ is the 2×2 unitary antisymmetric tensor. Q_L are the left-handed quark doublets and d'_R and u'_R are the right-handed down- and up-type quark singlets respectively, in the weak interaction eigenstate basis. When ϕ acquires a vacuum expectation value, Equation 2.3 yields mass terms for the quarks. The physical states are obtained by diagonalizing $Y^{u,d}$ by four unitary matrices, $V^{u,d}_{L,R}$ as $M_{diag} = V^f_L Y^f V^{f\dagger}_R (v/\sqrt{2})$, $f = u, d$. As a result, the charged-current interactions couple to the physical u^i_L and d^j_L with couplings given by

$$\frac{-g}{\sqrt{2}}(\overline{u}_L, \overline{c}_L, \overline{t}_L)\gamma^\mu W^+_\mu V_{CKM} \begin{pmatrix} d_L \\ s_L \\ b_L \end{pmatrix} + \text{h.c.}, \quad V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \tag{2.11}$$

This Cabibbo-Kobayashi-Maskawa matrix (CKM) is a 3×3 complex unitary matrix. It can be parametrized by three mixing angles and a CP violating phase. The standard convention to parametrize it is

$$V_{CKM} = \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix}, \tag{2.12}$$

where $c_{ij} = \cos\theta_{ij}$, $s_{ij} = \sin\theta_{ij}$, and δ is the phase responsible of all CP-violating phenomena in flavor-changing processes in the SM. All angles can be chosen to lie in the first quadrant, so $s_{ij}, c_{ij} \geq 0$

It is known experimentally that $s_{13} \ll s_{23} \ll s_{12} \ll 1$. It is convenient to exhibit

this hierarchy using the Wolfenstein parametrization [1, 32, 33],

$$s_{12} = \lambda = \frac{|V_{us}|}{\sqrt{|V_{ud}|^2 + |V_{us}|^2}}, \tag{2.13a}$$

$$s_{23} = A\lambda^2 = \lambda \left| \frac{V_{cb}}{V_{us}} \right| \tag{2.13b}$$

$$s_{13}e^{i\delta} = A\lambda^3(\rho + i\eta) = \frac{A\lambda^3(\overline{\rho} + i\overline{\eta})\sqrt{1 - A^2\lambda^4}}{\sqrt{1-\lambda^2}[1 - A^2\lambda^4(\overline{\rho} + i\overline{\eta})]} = V_{ub}^*. \tag{2.13c}$$

It is useful to see V_{CKM} written up to sixth order in λ,

$$V_{CKM} = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 - \frac{1}{8}\lambda^4 & \lambda & A\lambda^3\overline{z} \\ -\lambda + \frac{1}{2}A^2\lambda^5[1 - 2z] & 1 - \frac{1}{2}\lambda^2 - \frac{1}{8}\lambda^4(1 + 4A^2) & A\lambda^2 \\ A\lambda^3[(1 - (1 - \frac{1}{2}\lambda^2)z] & -A\lambda^2 + \frac{1}{2}A\lambda^4[1 - 2z] & 1 - \frac{1}{2}A^2\lambda^4 \end{pmatrix} + \mathcal{O}(\lambda^6), \tag{2.14}$$

where $z = \rho + i\eta$.

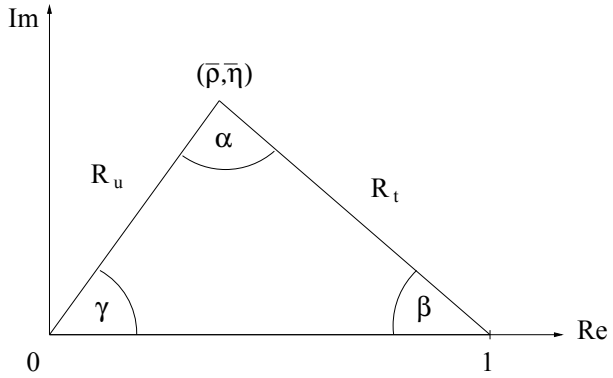


Figure 2.1: Sketch of the unitary triangle, R_u and R_t can be replaced by any $\left|\frac{V_{ui}V_{uj}^*}{V_{ci}V_{cj}^*}\right|$ and $\left|\frac{V_{ti}V_{tj}^*}{V_{ci}V_{cj}^*}\right|$ $(i, j = d, s, t;\ i \neq j)$, respectively.

The CKM matrix is unitary, this imposes

$$\sum_i V_{ij}V_{ik}^* = \sum_i V_{ji}V_{ki}^* = \delta_{jk}. \tag{2.15}$$

The six vanishing combinations can be represented as triangles in a complex plane, as shown in Figure 2.1. The area of all triangles is the same, half of the Jarlskog invariant

Matrix Element	Experimental World Average
$\|V_{ud}\|$	0.97425 ± 0.00022
$\|V_{us}\|$	0.2252 ± 0.0009
$\|V_{cd}\|$	0.230 ± 0.011
$\|V_{cs}\|$	1.006 ± 0.023
$\|V_{cb}\|$	0.0409 ± 0.0011
$\|V_{ub}\|$	0.00415 ± 0.00049
$\|V_{td}\|$	0.0084 ± 0.0006
$\|V_{ts}\|$	0.0429 ± 0.0026
$\|V_{tb}\|$	0.89 ± 0.07

Table 2.2: World average absolute values of the CKM matrix elements, taken from [1].

J, which is a measure of CP violation independent of the phase convention. Jarlskog invariant is defined by

$$J \sum_{m,n} \varepsilon_{ikm} \varepsilon_{jln} = \mathrm{Im}[V_{ij} V_{kl} V_{il}^* V_{kj}^*]. \tag{2.16}$$

When CP is violated J is different from zero and can be related to the single CP-violating parameter. In the Wolfenstein parametrization, $J \simeq \lambda^6 A^2 \eta$.

In the $B_s^0 - \overline{B}_s^0$ system, the relevant unitary triangle is

$$V_{us} V_{ub}^* + V_{cs} V_{cb}^* + V_{ts} V_{tb}^* = 0. \tag{2.17}$$

Experimental measurements of the absolute values of the CKM matrix elements are listed in Table 2.2

For each unitary triangle represented in Figure 2.1 the angles can be defined as:

$$\alpha_{ij} = \arg\left(\frac{V_{ti} V_{tj}^*}{V_{ui} V_{uj}^*}\right), \tag{2.18a}$$

$$\beta_{ij} = \arg\left(\frac{V_{ci} V_{cj}^*}{V_{ti} V_{tj}^*}\right), \tag{2.18b}$$

$$\gamma_{ij} = \arg\left(\frac{V_{ui} V_{uj}^*}{V_{ci} V_{cj}^*}\right). \tag{2.18c}$$

Since CP violation involves phases of CKM elements, many measurements of CP-violating observables can be used to constrain the angles and the $\overline{\rho}, \overline{\eta}$ parameters. In fact, there is an over constrain given by all these measurements, and then can be used to test the CKM mechanism in SM, and improve the determination of the CKM matrix elements.

In particular, measurements of CP violation in B^0 meson decays provide direct information of the unitary triangle defined in Equation 2.3 with angles:

$$\alpha = \phi_2 = \arg\left(\frac{V_{td}V_{tb}^*}{V_{ud}V_{ub}^*}\right), \tag{2.19a}$$

$$\beta = \phi_1 = \arg\left(\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right), \tag{2.19b}$$

$$\gamma = \phi_3 = \arg\left(\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}\right). \tag{2.19c}$$

The experimental world average of these three angles are

$$\sin 2\beta = 0.679 \pm 0.020, \tag{2.20a}$$

$$\alpha = (89.0^{+4.4}_{-4.2})^\circ, \tag{2.20b}$$

$$\gamma = (68^{+10}_{-11})^\circ. \tag{2.20c}$$

The CKM matrix elements can be most precisely determined by a global fit that uses all amiable measurementts and imposes the three generation unitarity. The fit must also use theoretical predictions for hadronic matrix elements, which sometimes have significant uncertainties. There are several approaches to combine experimental data. CKMfitter and References [36] use frequentist statistics, while UTfit uses a Bayesian approach. Both approaches provide consistent results. The constraints implied by the unitarity of the three generation CKM matrix significantly reduce the allowed range of some of the CKM matrix elements. The fit for the Wolfenstein parameters defined in Equations 2.13 gives [1]

$$\lambda = 0.22535 \pm 0.00065, \tag{2.21a}$$

$$A = 0.811^{+0.022}_{-0.012}, \tag{2.21b}$$

$$\overline{\rho} = 0.131^{+0.026}_{-0.013}, \tag{2.21c}$$

$$\overline{\eta} = 0.345^{0.013}_{-0.014}. \tag{2.21d}$$

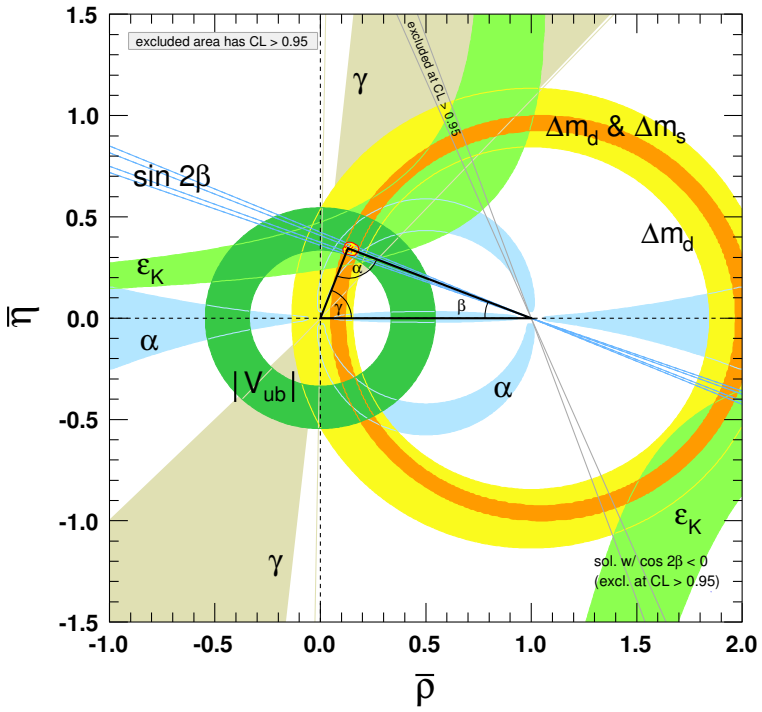


Figure 2.2: Constraints on the $\overline{\rho}, \overline{\eta}$ plane [1]. The shaded areas have 95% C.L.

The fit results [1] for the magnitudes of all nine CKM matrix elements are

$$V_{CKM} = \begin{pmatrix} 0.97427 \pm 0.00015 & 0.22534 \pm 0.00065 & 0.00351^{+0.00015}_{-0.00014} \\ 0.22520 \pm 0.00065 & 0.97344 \pm 0.00016 & 0.0412^{+0.0011}_{-0.0005} \\ 0.00867^{+0.00029}_{-0.00031} & 0.0404^{+0.0011}_{-0.0005} & 0.999146^{+0.000021}_{-0.000046} \end{pmatrix}, \quad (2.22)$$

and the Jarlskog invariant is $J = (2.96^{+0.20}_{-0.16}) \times 10^{-5}$. Constrains on the $\overline{\rho}, \overline{\eta}$ plane from various measurements and the global fit result are shown in Figure 2.2. The latest fit results for the parameter in Equations 2.21 can be found at References [37, 38].

2.4 CP violation in Meson Decays

The CP transformation is a combination of two discrete transformations, charge conjugation, C, and parity transformation P. Under charge conjugation fields representing particles and antiparticles are interchanged, by conjugation of all internal quantum

numbers,

$$C(a|\psi(\vec{x})\rangle) = a^*|\bar{\psi}(\vec{x})\rangle. \tag{2.23}$$

Under parity transformation, the handedness of the space is reversed, $\vec{x} \to -\vec{x}$,

$$P|\psi(\vec{x})\rangle = |\bar{\psi}(-\vec{x})\rangle. \tag{2.24}$$

If CP were an exact symmetry, the interactions of nature would be the same for matter and antimatter when space is reversed. It is observed in many experiments that C and P are symmetric in a wide variety of phenomena, and therefore they are also CP-symmetric. In particular, these symmetries are respected by gravity, electromagnetism and strong interactions. However, weak interactions violate C and P very drastically. The charged W bosons couple to left-handed electrons and right-handed positrons, but neither to right-handed electrons nor left-handed positrons. While C and P symmetries are violated separately by weak interactions, CP symmetry is preserved by most weak interaction processes.

The CP symmetry is, however, violated in some not so usual processes. In 1964 it was discovered that in neutral K decays there is an asymmetry among matter and antimatter [22]. The first evident of direct CP violation appeared on 1988 [39]. Since then, a number of CP-violating phenomena have appeared in different pseudoscalar meson systems. CP violation has a larger effect on B mesons than in K mesons. Recently CP violations has been observed in D meson [40–42]. Until now, there is no evidence of CP violation in the lepton sector.

The phenomenology of CP violation is superficially different in K, D, B, and B_s^0 decays. This is primarily because each of these systems is governed by a different balance between decay rates, oscillations, and lifetime splitting. However, the underlying mechanisms of CP violation are identical for all pseudoscalar mesons.

Here it is described the general formalism for, and classification of, CP violation meson M that might be a charged or neutral K, D, B or B_s^0 meson.

Quantity	Type of CP Violation	Decay
$\|\epsilon\| = (2.228 \pm 0.011) \times 10^{-3}$	Indirect	$K \to \pi\pi, K \to \pi\ell\nu$, $K_L \to \pi^+\pi^-e^+e^-$
$\mathrm{Re}(\epsilon'/\epsilon) = (1.65 \pm 0.26) \times 10^{-3}$	Direct	$K \to \pi\pi$
$S_{\psi K^0} = +0.679 \pm 0.020$	Interference	$B^0 \to \psi K^0$
$S_{\eta' K^0} = +0.59 \pm 0.07$ $S_{\phi K^0} == +0.74^{+0.11}_{-0.13}$ $S_{f_0 K^0} = +0.69^{+0.10}_{-0.12}$ $S_{K^+K^-K^0} = +0.68^{+0.09}_{-0.10}$	Interference	modes related to $b \to q\bar{q}s$
$S_{\pi^+\pi^-} = -0.65 \pm 0.07$	Interference	$B^0 \to \pi^+\pi^-$
$C_{\pi^+\pi^-} = -0.36 \pm 0.06$	Direct	$B^0 \to \pi^+\pi^-$
$S_{\psi\pi^0} = -0.93 \pm 0.15$ $S_{D^+D^-} = -0.98 \pm 0.17$ $S_{D^{*+}D^{*-}} = -0.77 \pm 0.10$	Interference	modes related to $b \to c\bar{c}d$
$A_{K^\pm\pi^\pm} = -0.087 \pm 0.008$	Direct	$\overline{B}^0 \to K^-\pi^+$
$A_{D_+K^\pm} = +0.19 \pm 0.03$	Direct	$B^\pm \to D_+K^\pm$

Table 2.3: Observables where it has been measured CP violation at level above 5σ.

2.4.1 Charged- and neutral-meson decays

The decay amplitudes of M and its CP conjugate $\overline{M}$ to a multiparticle final state f and its CP conjugate $\overline{f}$ are defined as

$$A_f = \langle f|e^{-i\mathscr{H}t}|M\rangle, \tag{2.25a}$$

$$A_{\overline{f}} = \langle \overline{f}|e^{-i\mathscr{H}t}|M\rangle, \tag{2.25b}$$

$$\overline{A}_f = \langle f|e^{-i\mathscr{H}t}|\overline{M}\rangle, \tag{2.25c}$$

$$\overline{A}_{\overline{f}} = \langle \overline{f}|e^{-i\mathscr{H}t}|\overline{M}\rangle, \tag{2.25d}$$

where $\mathscr{H}$ is the hamiltonian governing the time evolution. It can be shown [43] that, at sufficiently short times, Equations 2.25 can be approximated as

$$A_f = \langle f|\mathscr{H}|M\rangle, \tag{2.26a}$$
$$A_{\overline{f}} = \langle \overline{f}|\mathscr{H}|M\rangle, \tag{2.26b}$$
$$\overline{A}_f = \langle f|\mathscr{H}|\overline{M}\rangle, \tag{2.26c}$$
$$\overline{A}_{\overline{f}} = \langle \overline{f}|\mathscr{H}|\overline{M}\rangle, \tag{2.26d}$$

The action of the CP operator on these states introduces phases ξ_M and ξ_f that depend on their flavor content, according to

$$CP|M\rangle = e^{+\xi_M}|\overline{M}\rangle, \tag{2.27a}$$
$$CP|\overline{M}\rangle = e^{-\xi_M}|M\rangle, \tag{2.27b}$$
$$CP|f\rangle = e^{+\xi_f}|\overline{f}\rangle, \tag{2.27c}$$
$$CP|\overline{f}\rangle = e^{-\xi_f}|f\rangle, \tag{2.27d}$$

so that $(CP)^2 = 1$. The phases ξ_M and ξ_f are unphysical because of the flavor symmetry of the strong interaction. If the CP symmetry is conserved by the dynamics, *i.e.*,

$$[CP, \mathscr{H}] = 0, \tag{2.28}$$

then A_f and $\overline{A}_{\overline{f}}$ have the same magnitude and an arbitrary unphysical phase

$$\overline{A}_{\overline{f}} = e^{i(\xi_f - \xi_M)} A_f. \tag{2.29}$$

2.4.2 Neutral-meson mixing

If we have a initial state is a superposition of M^0 and $\overline{M}^0$, say

$$|M(0)\rangle = a(0)|M^0\rangle + b(0)|\overline{M}^0\rangle, \tag{2.30}$$

this state will evolve in time acquiring components that describe all possible decay final states $\{f_1, f_2, \dots\}$, that is,

$$|M(t)\rangle = a(t)|M^0\rangle + b(t)|\overline{M}^0\rangle + c_1(t)|f_1\rangle + c_2(t)|f_2\rangle + \dots. \tag{2.31}$$

In the case that we are only interested in the mixing phenomena, i.e., in computing only the values of $a(t)$ and $b(t)$ but not the terms related to decay final states, $c_i(t) = \langle f_i|e^{-i\mathscr{H}t}|M^0\rangle$; and if the times t in which we are interested are much larger than the typical strong interaction time scale, then we can use the simplified Wigner-Weisskopf formalism [44, 45]. The simplified time evolution is determined by a 2×2 effective hamiltonian, $\mathbf{H}$, that is not Hermitian, since otherwise the mesons would only oscillate and not decay. Any complex matrix, such as $\mathbf{H}$, can be written in terms of hermitian matrices $\mathbf{M}$ and $\mathbf{\Gamma}$ as

$$\mathbf{H} = \mathbf{M} - \frac{i}{2}\mathbf{\Gamma}. \tag{2.32}$$

The matrices $\mathbf{M}$ and $\mathbf{\Gamma}$ are associated with $(M^0, \overline{M}^0) \leftrightarrow (M^0, \overline{M}^0)$ transitions via off-shell (dispersive) and on-shell (absorptive) intermediate states, respectively. Diagonal elements of $\mathbf{H}$ are associated with flavor-conservating transitions $M^0 \to M^0$ and $\overline{M}^0 \to \overline{M}^0$, while off-diagonal elements are associated with flavor-changing transitions $M^0 \longleftrightarrow \overline{M}^0$.

The eigenvectors of $\mathbf{H}$ have well defined masses and decay widths. To specify the components of the strong interaction eigenstates, M^0 and $\overline{M}^0$, in the light (M_L) and heavy (M_H) mass eigenstates three parameters are introduced, p, q, z. The z parameter is introduced for the case that CP and CPT symmetries are both violated in mixing. The mass eigenstates are then defined as

$$|M_L\rangle = \frac{1}{Z}\left(p\sqrt{1-z})|M^0\rangle + q\sqrt{1+z})|\overline{M}^0\rangle\right), \tag{2.33a}$$

$$|M_H\rangle = \frac{1}{Z}\left(p\sqrt{1+z})|M^0\rangle - q\sqrt{1-z})|\overline{M}^0\rangle\right), \tag{2.33b}$$

where $Z = |p|^2\sqrt{(1-z)(1-\overline{z})} + |q|^2\sqrt{(1+z)(1+\overline{z})}$.

The real and imaginary parts of the eigenvalues $\omega_{L,H}$ corresponding to $|M_{L,H}\rangle$ represent their masses and decay rates, respectively. The masses and width differences are

$$\Delta m \equiv m_H - m_L = \mathrm{Re}(\omega_H - \omega_L), \tag{2.34a}$$

$$\Delta\Gamma \equiv \Gamma_H - \Gamma_L = \mathrm{Im}(\omega_H - \omega_L). \tag{2.34b}$$

In this way, Δm is defined to be positive, while the sign of $\Delta\Gamma$ has to be experimentally determined. The sign of $\Delta\Gamma$ has not yet been determined for B^0 mesons, while $\Delta\Gamma < 0$

for K and B_s^0 mesons [46]and $\Delta\Gamma > 0$ is established for D mesons [23]. The SM predicts $\Delta\Gamma < 0$ also for B^0 mesons.

Solving the eigenvalue problem for $\mathbf{H}$ yields

$$\omega_H = \mathbf{M}_{11} + \mathbf{M}_{22} - \frac{i}{2}(\mathbf{\Gamma}_{11} + \mathbf{\Gamma}_{22}) + \sqrt{\left(\mathbf{M}_{12} - \frac{i}{2}\mathbf{\Gamma}_{12}\right)\left(\mathbf{M}_{12}^* - \frac{i}{2}\mathbf{\Gamma}_{12}^*\right) + \frac{1}{4}(\delta m - \frac{i}{2}\delta\Gamma)}, \tag{2.35a}$$

$$\omega_L = \mathbf{M}_{11} + \mathbf{M}_{22} - \frac{i}{2}(\mathbf{\Gamma}_{11} + \mathbf{\Gamma}_{22}) + \sqrt{\left(\mathbf{M}_{12} - \frac{i}{2}\mathbf{\Gamma}_{12}\right)\left(\mathbf{M}_{12}^* - \frac{i}{2}\mathbf{\Gamma}_{12}^*\right) + \frac{1}{4}(\delta m - \frac{i}{2}\delta\Gamma)}, \tag{2.35b}$$

$$\left(\frac{q}{p}\right)^2 = \frac{\mathbf{M}_{12}^* - (i/2)\mathbf{\Gamma}_{12}^*}{\mathbf{M}_{12} - (i/2)\mathbf{\Gamma}_{12}}, \tag{2.36}$$

and

$$z = \frac{\delta m - (i/2)\delta\Gamma}{\Delta m - (i/2)\Delta\Gamma}, \tag{2.37}$$

where

$$\delta m \equiv \mathbf{M}_{11} - \mathbf{M}_{22}, \tag{2.38a}$$

$$\delta\Gamma \equiv \mathbf{\Gamma}_{11} - \mathbf{\Gamma}_{22}, \tag{2.38b}$$

are the difference in effective mass and decay-rate expectation values for the strong interaction states M^0 and $\overline{M}^0$.

If either CP or CPT is a symmetry of $\mathbf{H}$ (independently of whether T symmetry is violated or not), then the values of δm and $\delta\Gamma$ are both zero, and hence $z = 0$ Also, the difference of the eigenvalues becomes

$$\omega_H - \omega_L = 2\sqrt{\left(\mathbf{M}_{12} - \frac{i}{2}\mathbf{\Gamma}_{12}\right)\left(\mathbf{M}_{12}^* - \frac{i}{2}\mathbf{\Gamma}_{12}^*\right)}. \tag{2.39}$$

If either CP or T is a symmetry of $\mathbf{H}$ (independently of whether CPT symmetry is violated or not), then $\mathbf{\Gamma}_{12}/\mathbf{M}_{12}$ is real, leading to

$$\left(\frac{q}{p}\right)^2 = e^{2i\xi_M} \quad \Rightarrow \quad \left|\frac{q}{p}\right| = 1, \tag{2.40}$$

If, and only if, CP is a symmetry of $\mathbf{H}$ (independently of CPT and T), then both of the above conditions are hold, with the result that the mass eigenstates are orthogonal

$$\langle M_H|M_L\rangle = |p|^2 - |q|^2 = 0. \tag{2.41}$$

2.4.3 CP-violating observables

All CP-violating observables in M and $\overline{M}$ decays to final states f and $\overline{f}$ can be expressed in terms of phase-convention-independent combinations of $A_f, \overline{A}_f, A_{\overline{f}}$ and $\overline{A}_{\overline{f}}$, together with, for neutral mesons only, q/p. CP violation in charged-meson decays depends only in the combination $|\overline{A}_{\overline{f}}/A_f|$, while CP violation in neutral-meson decays is complicated by $M \longleftrightarrow \overline{M}^0$ oscillations, and depends, additionally, on $|q/p|$ and on

$$\lambda_f \equiv \frac{q}{p}\frac{\overline{A}_f}{A_f}. \tag{2.42}$$

The decay rates of the two neutral K mass eigenstates, K_S and K_L, are different enough ($\Gamma_S/\Gamma_L \sim 500$) that the decays can, in most cases, actually be studied independently. However, in the case of the B_s^0 system, as well as in the D and B^0 mesons, it makes sense to define an average decay width

$$\Gamma = \frac{\Gamma_H + \Gamma_L}{2}, \tag{2.43}$$

and it is expected that $\Delta\Gamma/\Gamma$ is having relatively small values, so both eigenstates must be considered in their evolution. Let us denote the state of an initially pure $|M^0\rangle$ or $|\overline{M}^0\rangle$ after an elapsed proper time t as $|M^0_{phys}(t)\rangle$ or $|\overline{M}^0_{phys}(t)\rangle$, respectively. Using the effective Hamiltonian approximation, but not assuming CPT is a good symmetry, the evaluated states are

$$|M^0_{phys}(t)\rangle = (g_+(t) + zg_-(t))|M^0\rangle - \sqrt{1-z^2}\frac{q}{p}g_-(t)|\overline{M}^0\rangle, \tag{2.44a}$$

$$|M^0_{phys}(t)\rangle = (g_+(t) - zg_-(t))|M^0\rangle - \sqrt{1-z^2}\frac{p}{q}g_-(t)|\overline{M}^0\rangle, \tag{2.44b}$$

where

$$g_\pm(t) \equiv \frac{1}{2}\left(e^{-(im_H+\Gamma_H/2)t} \pm e^{-(im_L+\Gamma_L/2)t}\right). \tag{2.45}$$

The obtained decay rates are:

$$\begin{aligned}\frac{e^{\Gamma t}}{N_f}\,\frac{\mathrm{d}\Gamma[M^0_{phys}(t)\to f]}{\mathrm{d}t} = \\ &\left(|A_f|^2|1+z|^2+\left|\frac{q}{p}\right|^2|A_f|^2|1-z^2|\right)\cosh\left(\Delta\Gamma t/2\right)+\\ &\left(|A_f|^2|1+z|^2-\left|\frac{q}{p}\right|^2|A_f|^2|1-z^2|\right)\cos\left(\Delta m t\right)+\\ &2\mathrm{Re}\left[\frac{q}{p}A_f^*\overline{A}_f(1+z^*)\sqrt{1-z^2}\right]\sinh\left(\Delta\Gamma t/2\right)-\\ &2\mathrm{Im}\left[\frac{q}{p}A_f^*\overline{A}_f(1+z^*)\sqrt{1-z^2}\right]\sin\left(\Delta m t\right),\end{aligned} \tag{2.46a}$$

$$\begin{aligned}\frac{e^{\Gamma t}}{N_f}\,\frac{\mathrm{d}\Gamma[\overline{M}^0_{phys}(t)\to f]}{\mathrm{d}t} = \\ &\left(\left|\frac{q}{p}\right|^2|A_f|^2|1-z|^2+|A_f|^2|1-z^2|\right)\cosh\left(\Delta\Gamma t/2\right)-\\ &\left(\left|\frac{p}{q}\right|^2|A_f|^2|1-z|^2-|A_f|^2|1-z^2|\right)\cos\left(\Delta m t\right)+\\ &2\mathrm{Re}\left[\frac{p}{q}A_f\overline{A}_f^*(1-z^*)\sqrt{1-z^2}\right]\sinh\left(\Delta\Gamma t/2\right)-\\ &2\mathrm{Im}\left[\frac{p}{q}A_f\overline{A}_f^*(1-z^*)\sqrt{1-z^2}\right]\sin\left(\Delta m t\right)\end{aligned} \tag{2.46b}$$

If CPT invariance is assumed, $z=0$, then the expressions are reduced to

$$\begin{aligned}\frac{e^{\Gamma t}}{N_f}\,\frac{\mathrm{d}\Gamma[M^0_{phys}(t)\to f]}{\mathrm{d}t} = \\ &\left(|A_f|^2+\left|\frac{q}{p}\right|^2|A_f|^2\right)\cosh\left(\Delta\Gamma t/2\right)+\\ &\left(|A_f|^2-\left|\frac{q}{p}\right|^2|A_f|^2\right)\cos\left(\Delta m t\right)+\\ &2\mathrm{Re}\left[\frac{q}{p}A_f^*\overline{A}_f\right]\sinh\left(\Delta\Gamma t/2\right)-2\mathrm{Im}\left[\frac{q}{p}A_f^*\overline{A}_f\right]\sin\left(\Delta m t\right),\end{aligned} \tag{2.47a}$$

and

$$\frac{e^{\Gamma t}}{N_f}\,\frac{\mathrm{d}\Gamma[\overline{M}^0_{phys}(t)\to f]}{\mathrm{d}t} = \left(\left|\frac{q}{p}\right|^2|A_f|^2+|A_f|^2\right)\cosh\left(\Delta\Gamma t/2\right) - \left(\left|\frac{p}{q}\right|^2|A_f|^2-|A_f|^2\right)\cos\left(\Delta m t\right) + 2\mathrm{Re}\left[\frac{p}{q}A_f\overline{A}^*_f\right]\sinh\left(\Delta\Gamma t/2\right) - 2\mathrm{Im}\left[\frac{p}{q}A_f\overline{A}^*_f\right]\sin\left(\Delta m t\right) \tag{2.47b}$$

N_f is a common, time-independent, normalization factor. Decay rates for the CP conjugate final state $\overline{f}$ are obtained analogously with substitutions

$$N_f\to N_{\overline{f}},\quad A_f\to A_{\overline{f}},\quad \overline{A}_f\to\overline{A}_{\overline{f}},$$

Terms proportional to $|A_f|^2$ or $|\overline{A}_f|^2$ are associates with direct decays with no net $M^0\longleftrightarrow\overline{M}^0$ oscillations. Terms proportional to $|(q/p)A_f|^2$ or $|(p/q)\overline{A}_f|^2$ are associated to decays following a net oscillation. Terms proportional to the real or imaginary part of $(q/p)A^*_f\overline{A}_f$ or $(p/q)A_f\overline{A}^*_f$ are associated to the interference of these two cases.

2.4.4 Classification of CP-violating effects

Assuming that CPT is a symmetry, as has to be for any theory that preserves Lorenz invariance, three types of CP-violating effects can be distinguished in meson decays:

i. Direct CP violation in charged-meson decays, where mixing effects are absent. The only possible source of CP violations is that the decay amplitudes are different in magnitude

$$\left|\frac{\overline{A}_{\overline{f}}}{A_f}\right|\neq 1, \tag{2.48}$$

leading to the asymmetries

$$\mathcal{A}_{f^\pm}\equiv\frac{\Gamma(M^-\to f^-)-\Gamma(M^+\to f^+)}{\Gamma(M^-\to f^-)+\Gamma(M^+\to f^+)}=\frac{|\overline{A}_{\overline{f}}|^2-|A_f|^2}{|\overline{A}_{\overline{f}}|^2+|A_f|^2}. \tag{2.49}$$

ii. CP violation in mixing of neutral mesons. If p and q do not have the same absolute value,

$$\left|\frac{q}{p}\right| \neq 1, \tag{2.50}$$

then the charged-current neutral meson decays $M \to \ell^+ X$ and $\overline{M} \to \ell^- X$[1] do not have the same decay amplitude. In these decays, this is the only source of CP violation, and can be measured via the asymmetry of the "wrong-sign" decays induced by oscillations,

$$\begin{aligned} \mathcal{A}_{SL} &\equiv \frac{\mathrm{d}[\overline{M}^0(t) \to \ell^+ X]/\mathrm{d}t - \mathrm{d}[M^0(t) \to \ell^- X]/\mathrm{d}t}{\mathrm{d}[\overline{M}^0(t) \to \ell^+ X]/\mathrm{d}t + \mathrm{d}[M^0(t) \to \ell^- X]/\mathrm{d}t} \\ &= \frac{|p|^4 - |q|^4}{|p|^4 + |q|^4}. \end{aligned} \tag{2.51}$$

It is worth to note that the semi-leptonic asymmetry of time dependent decay rates is actually time-independent.

iii. CP violation in interference between a decay without mixing, $M^0 \to f$, and a decay with mixing, $M^0 \to \overline{M}^0 \to f$. This effect only occurs in decays to a final states that are common to M^0 and $\overline{M}^0$, includingfi all CP eigenstates. Such interference can lead to a CP violation effect observed, for example, using the asymmetry of neutral meson decays into final CP eigenstates f_{CP},

$$\begin{aligned} \mathcal{A}_{f_{CP}}(t) &\equiv \frac{\mathrm{d}[\overline{M}^0(t) \to f_{CP}]/\mathrm{d}t - \mathrm{d}[M^0(t) \to f_{CP}]/\mathrm{d}t}{\mathrm{d}[\overline{M}^0(t) \to f_{CP}]/\mathrm{d}t + \mathrm{d}[M^0(t) \to f_{CP}]/\mathrm{d}t} \\ &\propto \left[\left(\left|\frac{q}{p}\right|^2 + 1\right)\cos(\Delta m t) - \left(\left|\frac{q}{p}\right|^2 - 1\right)\cosh\left(\frac{\Delta\Gamma t}{2}\right)\right]\left(1 - |\lambda_f|^2\right) \\ &\quad + 2\sinh\left(\frac{\Delta\Gamma t}{2}\right)\mathrm{Re}\left[\left|\frac{q}{p}\right|^2 \lambda_f - \lambda_f^*\right] \\ &\quad - 2\sin(\Delta m t)\mathrm{Im}\left[\left|\frac{q}{p}\right|^2 \lambda_f - \lambda_f^*\right], \end{aligned} \tag{2.52}$$

where,

$$\lambda_f \equiv \frac{q}{p}\frac{\overline{A}_{f_{CP}}}{A_{f_{CP}}}. \tag{2.53}$$

[1] This is the case in the Standard Model, and in most of the extensions, to the lowest order in G_F,

This expression only vanishes if $|q/p| = 1$, $|A_{f_{CP}}| = |\overline{A}_{f_{CP}}|$ and

$$\mathrm{Im}[\lambda_f] = 0. \tag{2.54}$$

This is the kind of CP violation expected in the $B_s^0 - \overline{B}_s^0$ system.

2.5 Decay Rates and Effective Field Theories

The decay rate or *decay width*[1] of a b hadron H_b with momentum $\mathbf{P}$ into some final state f of n particles, or the probability of transition per unit of time from the hadron H_b to the final satte f is computed as

$$\Gamma(H_b \to f) = S_f \frac{(2\pi)^4}{2E(\mathbf{P})} \int \delta^4 \left(P - \sum\nolimits_i k_i\right) |\mathscr{M}_f(P; k_i)|^2 \prod_{j=1}^{n} \frac{\mathrm{d}^3 k_j}{(2\pi)^3 2E_j(\mathbf{k}_j)}, \tag{2.55}$$

where $E_i(\mathbf{k}_i)^2 = k_i^2 + m_i^2$, m_i is the mass of the i-th particle, $\mathscr{M}_f(P; k_i)$ is the invariant amplitude of the decay, and S_f is a combinatorial factor to account for identical configurations in the final state f.

The confinement of quarks and gluons occur on a length scale $R_{had} \sim 1/\Lambda_{QCD} \sim$ 1 fm, which determines the size of hadrons. The parameter Λ_{QCD} ($\approx 200 - 400$ MeV) is a fundamental quantity in Quantum Chromodynamics, QCD, that sets the scale for the running coupling constant. For energies lower than Λ_{QCD} the strong coupling constant is large, and a perturbative development of the physical quantities is not possible. On the other hand, when two quarks are very close ($\ll 1/\Lambda_{QCD}$), the coupling constant is small (asymptotic freedom), and a perturbative approach is feasible. The asymptotic freedom allows us to write the decay amplitude as an *operator product expansion* (OPE),

$$\mathscr{M}_f = -\frac{4G_F}{\sqrt{2}} \sum_j |KM|_j^2 C_j(\mu) \langle f | O_j(\mu) | B \rangle \left[1 + \mathcal{O}\left(\frac{m_b^2}{M_W^2}\right)\right], \tag{2.56}$$

where μ is a renormalization scale, m_b the mass of the b quark, m_W the mass of the W boson, and $|KM|_j^2$ are appropriate products of the corresponding CKM elements (for example, $V_{cb}V_{cs}^*$ in $b \to \bar{c}cs$ decays). Physics from distances shorter than μ^{-1} is

[1] The decay rate, or width, to a final state f is the probability per unit time that the particle will decay to the state f.

contained in the Wilson coefficients C_j (dependence on the top mass, M_W, etc.), that are universal since they do not depend on the final state f. On the other hand, the hadronic matrix elements $\langle f|O_j(\mu)|B\rangle$ account for phenomena occurring at distances longer than μ^{-1} (e.g. dependence on Λ_{QCD}), and are usually evaluated using non-perturbative methods such as Lattice QCD (LQCD) or QCD Sum Rules (QCDSR). Therefore, at a scale of order m_b and up to corrections of order m_b^2/M_W^2, an effective Hamiltonian can be written, where the WIlson coefficients, C_j, are interpreted as effective coupling constants and the operators, O_j, are taken as the corresponding vertices:

$$H_{eff} = \frac{4G_F}{\sqrt{2}} \sum_j |KM|_j^2 C_j O_j + \text{h.c.} \tag{2.57}$$

The coupling constants, C_j, can be calculated perturbatively to include hard QCD effects. A list of O_j operators in the SM and extensions of the SM as well as numerical values of C_j, including their renormalization group evolution, can be found in Refs. [47–49].

2.6 Heavy quark symmetry

Once the short-distance physics has been separated from the long-distance physics, the latter can be simplified taking advantage of approximate symmetries, which imply non-trivial relations between observables. A *heavy quark* is by definition a quark with a mass much larger than Λ_{QCD} ($m_Q \gg \Lambda_{QCD}$): in the SM u, d, s are light quarks, whereas c, b and t are heavy quarks. For heavy quarks, the theory allows a perturbative description of the strong interactions. On the other hand, systems with a heavy quark and other light quarks are more complicated. Fortunately, the typical momenta exchanged between heavy and light quarks is of the order Λ_{QCD}, that is not enough to resolve the quantum numbers of the heavy quark, such as flavor (mass) and spin orientation of the heavy quark [50]. Therefore, light quarks only experience the heavy quark color field, and in the rest frame of the heavy quark, the spin interaction decouples (vanish in the limit $m_Q \to \infty$). Also, as $m_Q \to \infty$, the heavy quark and the hadron have the same velocity (0 in the rest frame), and in this limit the solution of the field

equations of QCD and the configuration of the light constituents are independent of m_Q.

In addition, in the limit $m_Q \to \infty$, hadrons which differ only in the flavor or spin of the heavy quark have the same configuration of the light degrees of freedom. This provides relations between the properties of hadrons containing heavy quarks, plus corrections due to finite heavy quark masses. These relations are encoded in a $SU(2N_h)$ spin-flavor symmetry group (N_h the number of heavy-quarks flavors), called heavy quark symmetry (HQS) under which the *effective* strong interactions are invariant.[1]

2.7 Heavy quark effective theory

It is useful to write a theory where the effects of the heavy quarks become irrelevant at low energies (similarly to the Fermi theory, where the effects of the W and Z are disregarded), where the HQS breaking corrections are developed in a systematic and consistent way in powers of $1/m_Q$. Short distance effects (hard gluons) can be incorporated in this effective theory using renormalization group techniques in the form of Wilson coefficients (see section 2.5). Such simplified description is provided by the heavy quark effective theory (HQET), or Heavy Quark Expanssion (HQE), where a heavy quark interacts with light quarks by the exchange of soft gluons (virtual momenta are small, of the order of Λ_{QCD}). In this framework, heavy particle fields are "integrated out", leaving a non-local effective action that can be expanded in an OPE in powers of $1/m_Q$.

The are many good reviews of HQET [51–54]. Here is mainly followed [50, 55]. Firstly, we observe that the propagator of a heavy quark can be expanded as follows:

$$\frac{i}{\not{p} - m_Q} = \frac{i}{v \cdot k} \frac{1 + \not{v}}{2} + \dots , \tag{2.58}$$

where $p_Q^\mu = m_Q v^\mu + k^\mu$, v is the quark velocity (very close to the hadron's velocity) and $k \sim \Lambda_{QCD} \ll m_Q$ is the residual momentum. As $m_Q \to \infty$, heavy quark flavor

[1] An analogy of the flavor symmetry is the fact that different isotopes have similar chemistry, since the wave function of the electrons are almost independent of the heavy nucleus. Similarly, the spin symmetry resembles the degeneration of the hyperfine levels in atoms.

symmetry emerges. Also, the quark-gluon vertex appears between two propagators of the form

$$P_{\pm} = \frac{1 \pm \not{v}}{2}. \tag{2.59}$$

Spin symmetry emerges from the fact that $P_+ \gamma^\mu P_+ = v^\mu P_+$ and $P_+^2 = P_+$ (just one γ matrix in the end). Therefore, it is convenient to write the heavy quark field $Q(x)$ in terms of velocity dependent fields and factorize m_Q:

$$Q = e^{-im_Q v\cdot x} \left[h_v(x) + H_v(x)\right], \tag{2.60}$$

where

$$\begin{pmatrix} h_v(x) \\ H_v(x) \end{pmatrix} = e^{im_Q v\cdot x} P_{\pm} Q(x). \tag{2.61}$$

After some calculations, the QCD Lagrangian for a heavy quark becomes

$$\mathcal{L}_Q = \bar{Q}(i \not{D} - m_Q)Q = \bar{h}_v iv\cdot D h_v - \bar{H}_v(iv\cdot D + 2m_Q)H_v + \text{interaction terms}, \tag{2.62}$$

where $D^\mu = \partial^\mu - ig_s T_a A_a^\mu$, g_s is the strong coupling constant and T_a are generators of the $SU(3)$ group. From this equation h_v describes apparently a massless field, whereas H_v "receives" twice the heavy quark mass. The HQET Lagrangian is constructed from h_v, eliminating H_v by using the equation of motion of Q, $(i \not{D} - m_Q)Q = 0$, and then looking for an effective Lagrangian that recovers the equation of motion for h_v:

$$\mathcal{L}_{\text{eff}} = \bar{h}_v iv\cdot D h_v + \bar{h}_v i \not{D}_\perp \frac{1}{2m_Q + iv\cdot D} i \not{D}_\perp h_v, \tag{2.63}$$

where $D_\perp^\mu = D^\mu - v^\mu v\cdot D$. Each derivative in the second term of Eq. 2.63 produces powers of k ($\ll m_Q$). Therefore, this term can be expanded in powers of iD/m_Q and use the properties of the P_+ operator to show that:

$$\mathcal{L}_{\text{eff}} = \bar{h}_v iv\cdot D h_v + \frac{1}{2m_Q}\bar{h}_v (i \not{D}_\perp)^2 h_v + C_{\text{mag}}(\mu)\frac{g_s}{4m_Q}\bar{h_v}\sigma_{\mu\nu}G^{\mu\nu}h_v + \mathcal{O}(1/m_Q^2), \tag{2.64}$$

where $[iD^\mu, iD^\nu] = ig_s G^{\mu\nu}$ is the gluon field strength tensor and $\sigma_{\mu\nu} = i\left[\gamma_\mu, \gamma_\nu\right]/2$. Eq. 2.64 is the HQET effective Lagrangian. As expected, in the limit $m_Q \to \infty$, where only the first term remains, the heavy quark symmetry is recovered . The second term describes the kinetic energy resulting from the residual motion of the heavy quark, whereas the third term represents the chromomagnetic coupling of the heavy quark

spin to the gluon field. They violate flavor and spin symmetry, respectively. The coefficient of the chromomagnetic operator, $C_{\rm mag}(\mu) = 1+\mathcal{O}(\alpha_s)$, receives short distance corrections which are treated perturbatively in powers of the strong coupling constant, while the kinetic term is protected from quantum corrections by the Lorentz symmetry.

Finally, in the operators of the EW Lagrangian, the QCD field Q can be replaced in terms of h_v (using the equations of motion for h_v and H_v, and Eq. 2.60) to develop a series of higher-dimension operators that describe $1/m_Q$ effects.

2.8 Heavy quark expansion and b hadron lifetimes

If the heavy hadron H_b has multiple decay modes (or branches) with different final states, the total or *inclusive* decay rate, or decay width, of the particle is obtained summing the decay rate of every branch,

$$\Gamma(H_b) = \sum_f \Gamma(H_b \to f). \tag{2.65}$$

The *branching ratio* for each mode is given by

$$\mathcal{B}(H_b \to f) = \frac{\Gamma(H_b \to f)}{\Gamma(H_b)}, \tag{2.66}$$

and the inverse of $\Gamma(H_b)$ measures the mean lifetime of the H_b particle[1],

$$\tau(H_b) = \frac{1}{\Gamma(H_b)}. \tag{2.67}$$

To calculate the inclusive decay rate, the HQET technology developed in previous sections can be used. This calculation relies on a hypothesis known as quark-hadron duality: the sum over all possible exclusive final states, which is necessary to determine the total lifetime, is equal to the sum over all possible final state quarks [56]. This assumption eliminates bound-state effects related to the individual properties of individual hadrons [50]. We use the optical theorem to relate the inclusive decay rate of a heavy hadron H_b to the imaginary part of the forward matrix element of the transition operator [57]:

$$\Gamma(H_b) = \frac{1}{2m_{H_b}} \langle H_b | \,\mathrm{Im}\; i \int d^4x\, T\{\mathcal{L}_{\rm eff}(x), \mathcal{L}_{\rm eff}(0)\}\, |H_b\rangle, \tag{2.68}$$

[1] The lifetime depends on the Lorentz frame in which it is measured. In this work, we always measure lifetimes in the H_b rest frame.

where $\mathcal{L}_{\rm eff}$ is the effective $\Delta B = 1$ Lagrangian in the Fermi theory,

$$\mathcal{L}_{\rm eff} = -2\sqrt{2}G_F V_{cb} \sum_{d'=d,s;u'=u,c} V^*_{u'd'} \left[C_1(\mu) Q_1^{u'd'}(\mu) + C_2(\mu) Q_2^{u'd'}(\mu) \right] + \text{h.c.}, \quad (2.69)$$

and Q_1 and Q_2 are four-quark operators given by

$$Q_1^{u'd'}(\mu) = \bar{d}'_L \gamma_\mu u'_L \bar{c}_L \gamma^\mu b_L, \qquad Q_2^{u'd'}(\mu) = \bar{c}_L \gamma_\mu u'_L \bar{d}'_L \gamma^\mu b_L. \quad (2.70)$$

If we construct an OPE for these operators, the decay rate is given by [47]

$$\begin{aligned} \Gamma(H_b) &= \sum_k \frac{C_k(\mu)}{2m_b^k} \langle O_k^{\Delta B=0}(\mu) \rangle \\ &= \frac{G_F^2 m_b^5}{192\pi^3} \left\{ C_0 \langle \bar{b}b \rangle + C_2 \frac{\langle \bar{b} g_s \sigma^{\mu\nu} G_{\mu\nu} b \rangle}{m_b^2} \right. \\ &\quad \left. + 16\pi^2 \frac{\sum_i C_3^i \langle (\bar{q}_i \Gamma_i q_i)(\bar{b} \Gamma'_i b) \rangle}{m_b^3} + ... \right\}, \end{aligned} \quad (2.71)$$

where $\langle O \rangle \equiv \langle H_b | O | H_b \rangle / 2m_{H_b}$, q_i stands for light quarks (u,d,s), and Γ_i, Γ'_i denote spin and color structures of the four-quark operators. In the limit $m_Q \to \infty$, we recover the parton decay ($\langle \bar{b}b \rangle \to 1$, see Eq. 2.72), where the b quark decays weakly without feeling the light (spectator) quarks or gluons inside the hadron. The second term in the last expression of Eq. 2.71 describes the effect of the gluon field on the heavy quark through the chromomagnetic part of the gluon field. Hence, it depends on the spin of the heavy quark but not in the flavor of the light quarks. The third term finally gives an explicit dependence on the light quarks, and so on.

To evaluate the first (leading operator) and second (first correction) terms in Eq. 2.71, the matrix elements are expanded in powers of $1/m_b$ using HQET techniques (b_v fields, etc.). Then:

$$\langle \bar{b}b \rangle = 1 - \frac{\mu_\pi^2(H_b) - \mu_G^2(H_b)}{2m_b^2} + \mathcal{O}(1/m_b^3), \quad (2.72)$$

$$\langle \bar{b} g_s \sigma^{\mu\nu} G_{\mu\nu} b \rangle = 2\mu_G^2(H_b) + \mathcal{O}(1/m_b), \quad (2.73)$$

where

$$\mu_\pi^2(H_b) = \frac{1}{2m_{H_b}} \langle H_v(v) | \bar{b}_v (i\vec{D})^2 b_v | H_b(v) \rangle, \quad (2.74)$$

$$\mu_G^2(H_b) = \frac{1}{2m_{H_b}} \langle H_v(v) | \bar{b}_v \frac{g_s}{2} \sigma^{\mu\nu} G_{\mu\nu} b_v) | H_b(v) \rangle. \quad (2.75)$$

To evaluate these HQET matrix elements, we use the spectroscopic formula for heavy hadrons [50]:

$$m_H = m_Q + \bar{\Lambda} + \frac{\mu_\pi^2(H_Q) - \mu_G^2(H_Q)}{2m_Q} + \ldots, \tag{2.76}$$

where $\bar{\Lambda}$ is a parameter independent of the heavy quark mass and spin. Then, the splitting between the ground-state pseudoscalar ($J = 0$) and vector ($J = 1$) b mesons is:

$$\mu_G^2(B) = \frac{3}{4}(M_{B^*}^2 - M_B^2) \approx 0.36 \text{ GeV}^2. \tag{2.77}$$

The value $\mu_\pi^2(B)$ has to be calculated in a non-perturbative framework, still constrained to $\mu_\pi^2(B) > \mu_G^2(B)$ [47]. It is useful to observe that the first and second matrix elements are equivalent to the Lorentz factor $\gamma^{-1} \approx 1 - \vec{p}^{\,2}/2m_b^2$, which decreases (increases) the decay width (lifetime) for a moving particle.

So, with HQET technology the decay rate of the b-mesons is calculated as an expansion of the heavy b-quark mass:

$$\frac{1}{\tau} = \Gamma = \Gamma_0 + \frac{\Lambda^2}{m_b^2}\Gamma_2 + \frac{\Lambda^3}{m_b^3}\Gamma_3 + \ldots, \tag{2.78}$$

where Γ_0 is interpreted as the decay rate of a free b-quark. The lifetime ratio of two given heavy mesons is read as:

$$\frac{\tau_1}{\tau_2} = 1 + \frac{\Lambda^3}{m_b^3}(\Gamma_3^0 + \Gamma_3^{(1)} + \ldots) + \frac{\Lambda^4}{m_b^4}(\Gamma_4 + \ldots) + \ldots. \tag{2.79}$$

2.9 B_0^s Lifetime

In the case of the B_d^0 and B_s^0 mesons, it is expected [58, 59]

$$\left| \frac{\overline{\tau}_{B_d^0}}{\overline{\tau}_{B_s^0}} - 1 \right| < 0.01, \tag{2.80}$$

where $\overline{\tau}$ denotes the the average lifetime of the two mass eigenstates for each system.

The B_s^0 system is a combination of two mass eigenstates with two different masses, M_L and M_H, and two different decay widths, Γ_L and Γ_H. For a decay into a final state f it is set

$$\Gamma_f = \begin{pmatrix} \Gamma_{f,11} & \Gamma_{f,12} \\ \Gamma_{f,12}^* & \Gamma_{f,22} \end{pmatrix}. \tag{2.81}$$

If the initial state is not tagged and neglecting CP-violation, the decay rate to a final mode f can be written as

$$\begin{aligned} P_{B_s^0\bar{B}_s^0\to f}(t) &= \frac{1}{2}(e^{-\Gamma_L t}+e^{-\Gamma_H t})(\Gamma_{f,11}+\Gamma_{f,22}) - (e^{-\Gamma_L t}-e^{-\Gamma_H t})\mathrm{Re}\ \Gamma_{\mathrm{f,12}} \\ &= \left(\frac{1}{2}(\Gamma_{f,11}+\Gamma_{f,22}) - \mathrm{Re}\Gamma_{\mathrm{f,12}}\right) e^{-\Gamma_L t} + \\ &\quad \left(\frac{1}{2}(\Gamma_{f,11}+\Gamma_{f,22}) + \mathrm{Re}\Gamma_{\mathrm{f,12}}\right) e^{-\Gamma_H t} \\ &= A_L e^{-\Gamma_L t} + A_H e^{-\Gamma_H t}, \end{aligned} \tag{2.82}$$

where A_L and A_H are the amplitudes of the states B_L and B_H respectively. Theses states are almost CP eigenstates. So, if the lifetime is measured experimentally with only one exponential, then there is a dependence of the lifetime on the composition of the final mode f [60],

$$\tau_{B_s^0\to f} = \frac{A_L\frac{1}{\Gamma_L^2} + A_H\frac{1}{\Gamma_H^2}}{A_L\frac{1}{\Gamma_L} + A_H\frac{1}{\Gamma_H}}. \tag{2.83}$$

In the case of semi-leptonic modes $\Gamma_{f,11} = \Gamma_{f,22} = \Gamma_{sl}$ and $\Gamma_{f,12} = 0$, so $A_L = A_H$, so the result simplifies to

$$\tau_{B_s^0\to sl} = \frac{1}{\Gamma}\frac{1+(\frac{\Delta\Gamma}{2\Gamma})^2}{1-(\frac{\Delta\Gamma}{2\Gamma})^2}, \tag{2.84}$$

where $\Gamma = (\Gamma_L+\Gamma_H)/2$ and $\Delta\Gamma = \Gamma_L - \Gamma_H$. So, should be noted that $\tau_{B_s^0\to sl}$ is always larger than $\tau_{B_s^0} = 1/\Gamma$. In the case of $\tau_{B_s^0}$ and $\tau_{B_d^0}$, according to the results of the quenched lattice-QCD calculation [61, 62], the ratio is restricted to

$$-0.0004 < \frac{\tau_{B_s^0}}{\tau_{B_d^0}} - 1 < 0. \tag{2.85}$$

If we use equation 2.84 and inequality 2.85 we can conclude that

$$\tau_{B_s^0\to sl} \simeq \tau_{B_d^0}\frac{1+(\frac{\Delta\Gamma}{2\Gamma})^2}{1-(\frac{\Delta\Gamma}{2\Gamma})^2}. \tag{2.86}$$

So, a high precision measurement of the B_s^0 lifetime, can be used to test these models.

The lifetime of the meson B_d^0 is known experimentally with a precision of better than 1% [63]. The world average reported by the Particle Data Group [1, 2] is $\tau_{B_d^0} = 1.519 \pm 0.007$ps. On the experimental side, the B_s^0 lifetime is reported to be, using the

flavor specific measurement, $\tau_{B_s^0} = 1.417 \pm 0.042$ps. With this two values, the lifetimes ratio gives

$$\frac{\tau_{B_s^0}}{\tau_{B_d^0}} = 0.933 \pm 0.028, \tag{2.87}$$

so, the experimentally measured lifetime ratio is smaller that expected when we compare with eq. 2.85.

3. 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

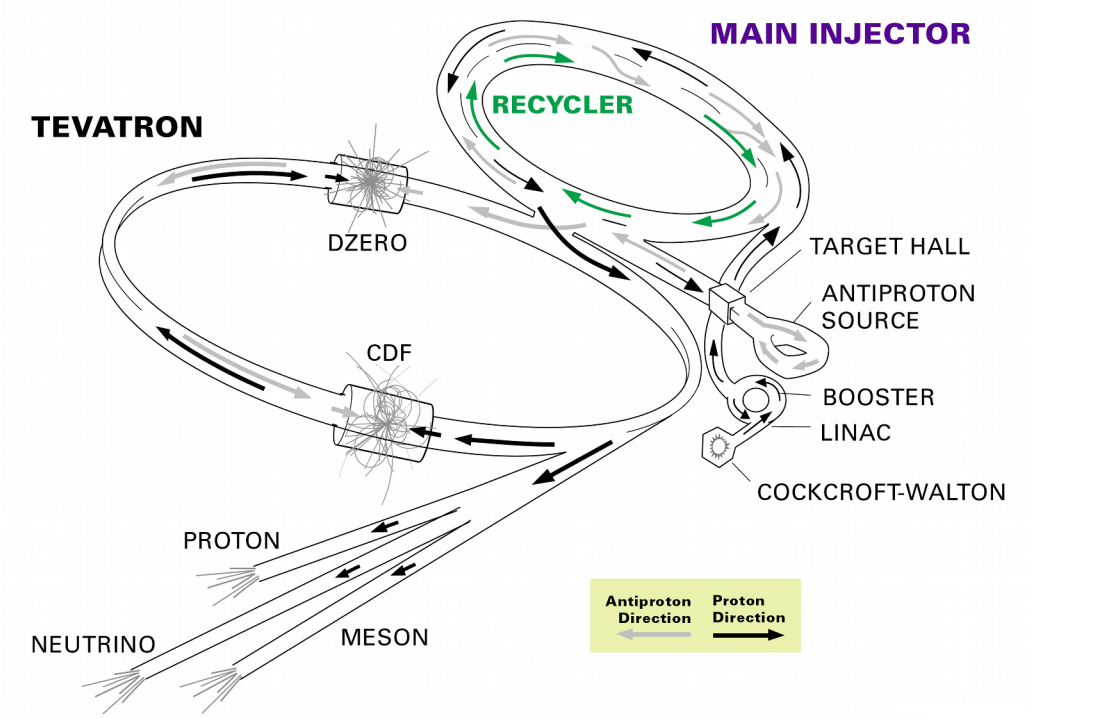


Figure 3.1: Fermilab accelerator chain

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.

3.1 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.

3.2 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 $= 4 \times 201$ 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.

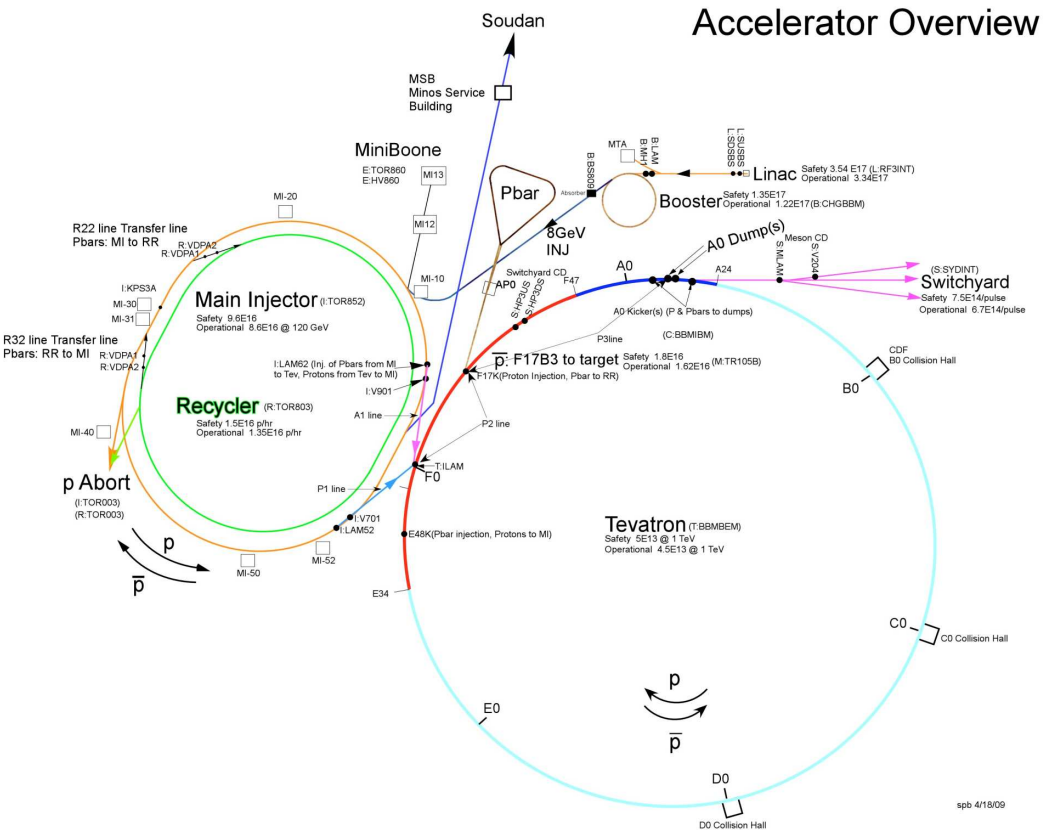


Figure 3.2: Detailed schematic of the accelerators in the Fermilab complex.

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

3.4 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.

3.5 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.

3.6 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 systems in 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.

3.7 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.

3.8 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.

4. The DØ Detector

4.1 Introduction

The DØ detector 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 detector performed very well during Run I of the Tevatron, from 1992 to 1996, leading to the discovery of the top quark and measurement of its mass, a precision measurement of the mass of W boson, detailed analysis of gauge boson couplings, among many other accomplishments.

During Run I, the Tevatron operated using six bounties of protons and antiprotons, with 3500 ns between each crossing and a center-of-mass energy of 1.8 TeV. The peak luminosity was typically $1 - 2 \times 10^{31}\text{cm}^{-2}\text{s}^{-1}$ and approximately 120 pb^{-1} of data were recorded. In Run II, which began in March 2001 and finished in September of 2011, the Tevatron was operated with 36 bounties of protons and antiprotons with a bunch spacing of 396 ns and at an increased center of mass of 1.96 TeV. The peak of instantaneous luminosity increased to be typically around $3.5 \times 10^{32}\text{cm}^{-2}\text{s}^{-1}$.

To take advantage of these improvements in the Tevatron, the DØ detector was upgraded for the Run II. The detector consists of three major subsystems: central tracking detectors, uranium/liquid argon calorimeters, and a muon spectrometer.

4.2 Central Tracking

Excellent tracking in the central region is necessary for studies in b physics, top quark, electroweak, etc. The central tracking system consists of the silicon microstrip tracker (SMT) and the central fiber tracker (CFT) surrounded by a solenoidal magnet of 2T. It surrounds the DØ beryllium beam pipe, which has a wall thickness of 0.508 mm and

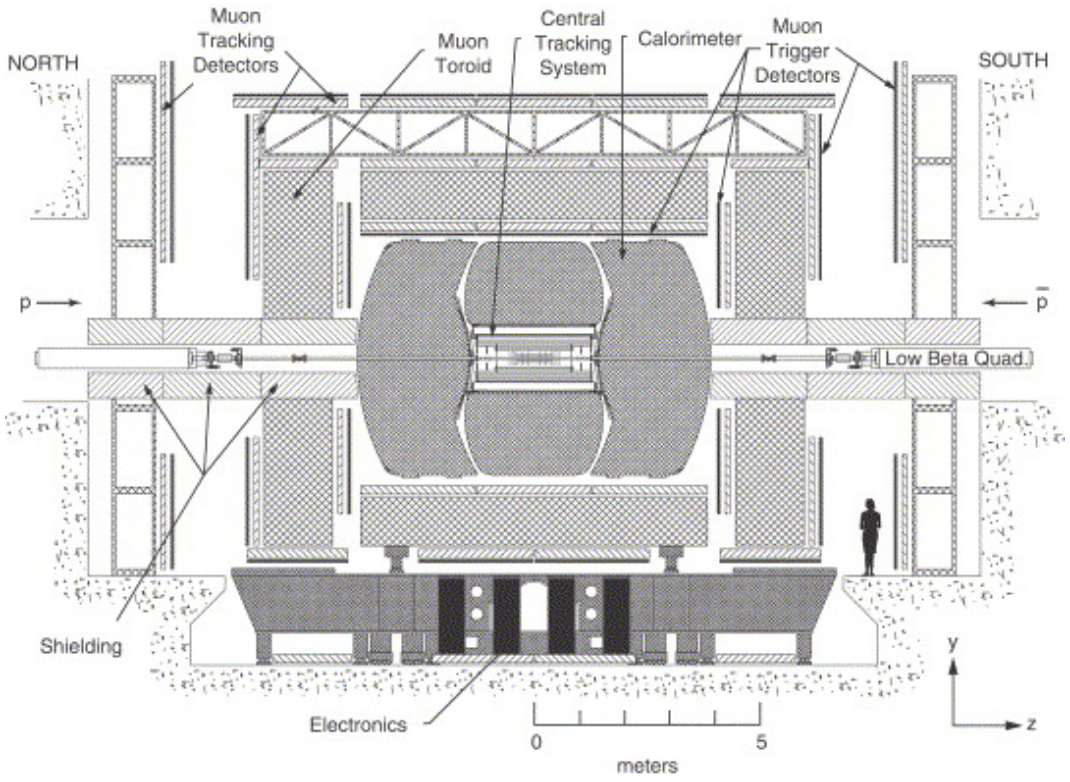


Figure 4.1: Diagram of the Upgraded DØ detector, as installed in the collision hall and viewed from inside the Tevatron ring. The forward proton detector is not shown. The detectors in the central region are shown in Fig. 4.2

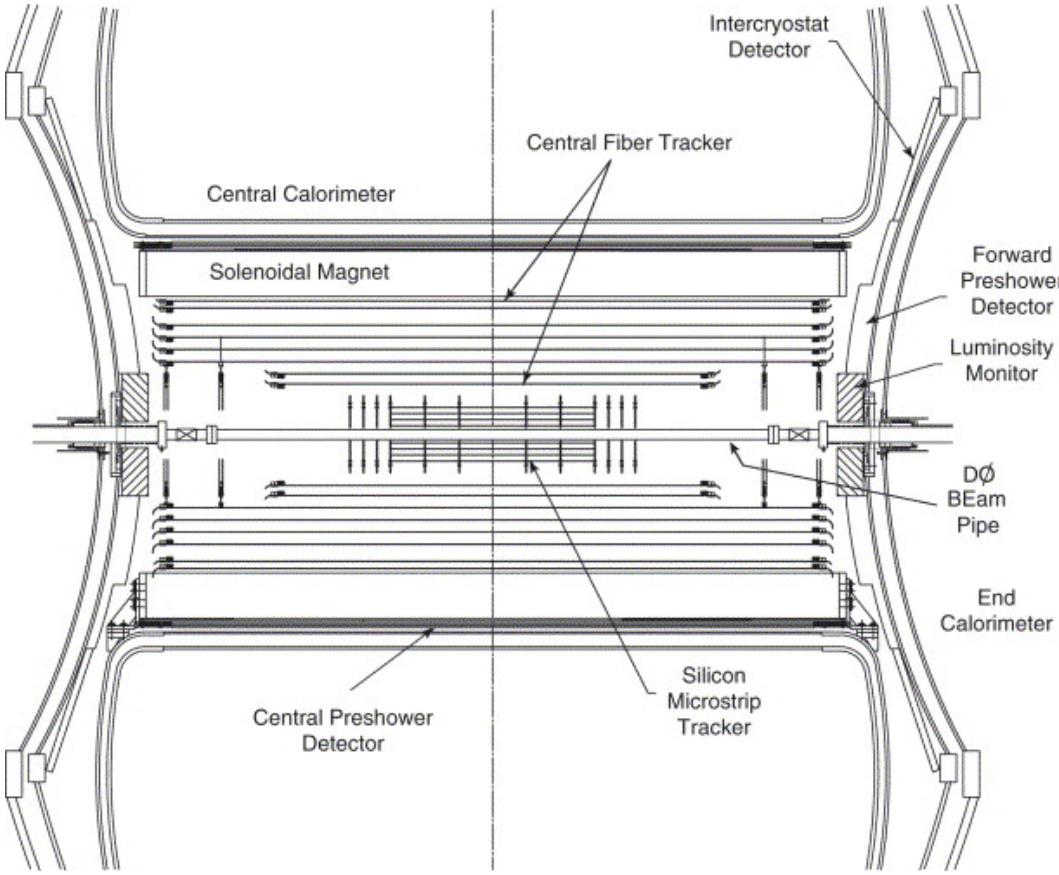


Figure 4.2: Cross-sectional view of the central tracking system for Run IIA in the $x - z$ plane. Also shown are the locations of the solenoid, the preshower detectors, luminosity monitors, and calorimeters.

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. the high resolution of the vertex position allows good measurement of lepton p_T.

Both SMT and CFT provide tracking information to the trigger. The SMT provides signal to the Level 2 and 3 trigger systems. 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 digitized analog signals of the CFT, in addition to the discriminator information available at Level 1 and 2. A schematic view of the central

tracking system is shown in Figure 4.2. The expected traverse momentum resolution (as $\delta p_T/p_T$) of the central tracking is shown in Figure 4.3 for tracks with $p_T = 1, 10$ and 100 GeV

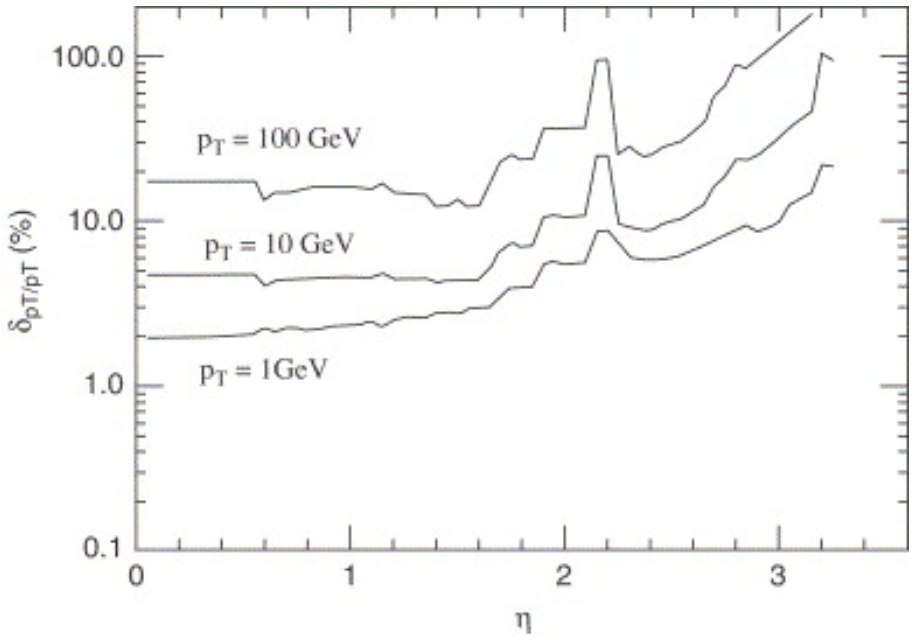


Figure 4.3: Expected relative transverse momentum resolution of the central tracking system as a function of pseudo rapidity for tracks with $p_T = 1, 10$, and 100 GeV. The relevant resolution for b-physics analysis is the 1 GeV.

4.2.1 Silicon microstrip tracker

The SMT provides both tracking and vertexing over nearly full pseudo rapidity, $\eta = -\ln[\tan(\theta/2)]$, coverage of the calorimeter and muon systems. Design of the detector, electronics, and cooling were, 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. Ideally the design of the SMT should be in such a way that most of the particles pass through it perpendicularly for all η, this conditions led to a design of barrel modules interspersed with disks in the center and assemblies of disks in the forward regions. The barrel detector primarily measure the $r - \phi$ coordinate and the disk detectors

measure $r - z$ as well as $r - \phi$. Thus, vertices for particles 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 CFT. An isometric view of the SMT is shown in Figure 4.2. 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 12 ladders each; layers 3 and 4 have 24 ladders each, for a total of 432 ladders. Each barrel is capped at high $|z|$ with a disk of 12 double-sided (DS) wedge detectors, called an"F-disk". Forward of the three disk/barrel assemblies on each side is a unit consisting three F-disk. In the far forward regions two "H-disks" provide tracking at high $|\eta|$. The H-disks are composed of single-sided (SS) wedges. In total, there are $792,576$ channels in 912 readout modules. The SMT is read out by custom-made 128-channel SVXIIe readout chips. These signals are used by the Level 2 and Level 3 trigger systems.

4.2.2 Central Fiber Tracker

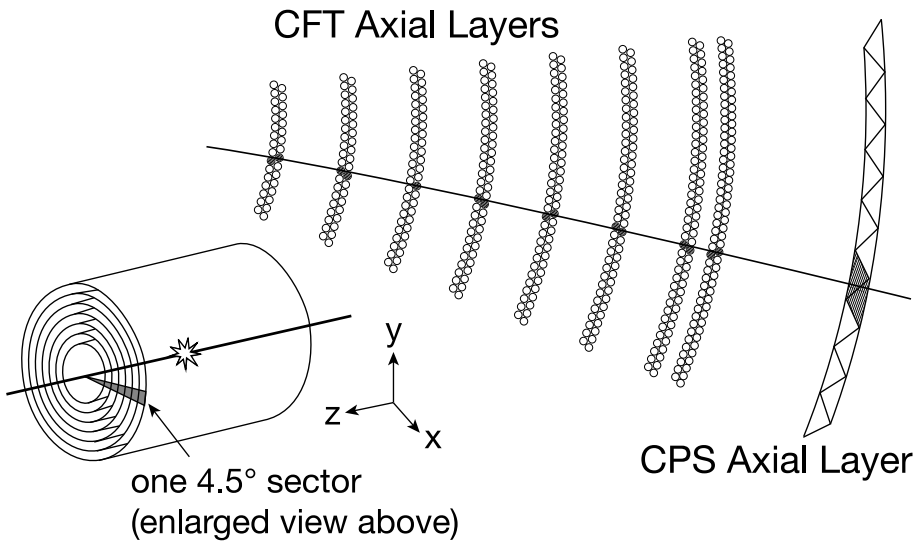


Figure 4.4: Transverse schematic view of the CFT detector [64].

The CFT consists of eight concentric cylinders, 1 to 8 counting from the innermost and outward. Their radii go from 20 cm to 52 cm. Cylinders 1 and 2 are 1.66m and

Layer	Radius (cm)	Fibers× layer	Fiber Separation (μm)	Active length (m)
A	20.04	1280×2	982.4	1.66
Au	20.22	1280×2	990.3	1.66
B	24.93	1600×2	978.3	1.66
Bv	24.93	1600×2	985.1	1.66
C	29.87	1920×2	976.1	2.52
Cu	30.05	1920×2	980.9	2.52
D	34.77	2240×2	974.9	2.52
Dv	34.95	2240×2	979.3	2.52
E	39.66	2560×2	971.7	2.52
Eu	39.86	2560×2	976.3	2.52
F	44.56	2880×2	970.0	2.52
Fv	44.74	2880×2	974.3	2.52
G	49.49	3200×2	969.8	2.52
Gu	49.67	3200×2	973.3	2.52
H	51.97	3520×2	926.1	2.52
Hv	52.15	3520×2	927.8	2.52

Table 4.1: Design parameters of the eight axial and stereo Layers comprising the CFT; indexes refer to the stereo angle of the corresponding layer according to $u_{angle} = +3°$ and $v_{angle} = -3°$.

the remaining are 2.52m long. The reason of this difference is that the SMT H-Disks are supported by cylinder 3. The CFT cylinder covers $|\eta \lesssim 1.7|$.

Each CFT cylinder have two layers of scintillating fibers mounted on its surface. The first layer is oriented along the beam direction and is referred as Axial Layer. The second layer is referred as Stereo Layer because its fiber are oriented with an stereo angle. Odd numbers Stereo Layers are the $u-$Layers and have an stereo angle of $+3°$. Even number Stereo Layers are $v-$Layers and have an stereo angle of $-3°$.

The scintillating fibers are $835\mu m$ in diameter and match the length of the cylinder they are mounted in. The core of the fiber is made from Polystyrene whose refractive index is $n = 1.59$ and is surrounded by two coats of approximately $25\mu m$ each. The

inner is al layer of Polymethyl-Methacrylate with $n = 1.49$ and outer layer of fluoro-acrylic with $n = 1.42$.

The scintillating fibers are optically connected to clear fiber waveguides of identical diameter which are $7.8m$ to $11.9m$ long and carry the scintillation light to Visible Light Photon Counters, VLPC's for readout. The scintillating fibers are structural and chemically similar to the clear fibers, contains fluorescent dyers.

The light is observed by only one end of each scintillating fiber. The opposite end of each of the scintillating fibers is mirrored with a sputtered aluminum coating that provides a reflectivity of about 90%.

The total $76,800$ readout channels of the CFT use around 184 km of scintillating fiber and 800 km of clear fiber. Table 4.1 shows the summary of the design parameter of the CFT.

4.2.3 Solenoidal Magnet

The superconducting Solenoidal Magnet helps to optimize the momentum resolution, $\delta p_{\mathrm{T}}/p_{\mathrm{T}}$, and tracking pattern recognition. The dimensions of the solenoid are: 2.73 m in length and 1.42 in diameter. It provides a central field of 2 T. The polarity of the magnet was reversed regularly in order to reduce detector asymmetry effects. The major parameters of the solenoid design are listed Table 4.2. A perspective view of the solenoid inside the Cap Calorimeter with its chimney and control dewar is shown in Figure 4.5. The $y - z$ view of the magnetic field with both the toroid and solenoid magnets at full current is shown in Figure 4.6.

4.3 Muon Detector

For muon triggering and measurement, the upgraded detector uses the Run I central muon system proportional drift tubes (PDTs) and toroidal magnets [65], 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.

Central filed	2.0 T
Operating current	4749 A
Cryostat warm bore diameter	1.067 m
Cryostat length	2.729 m
Stored Energy	5.3 MJ
Inductance	0.47 H
Cold mass	1460 kg
Thickness	0.87 X_0

Table 4.2: Major parameters of the DØ solenoid.

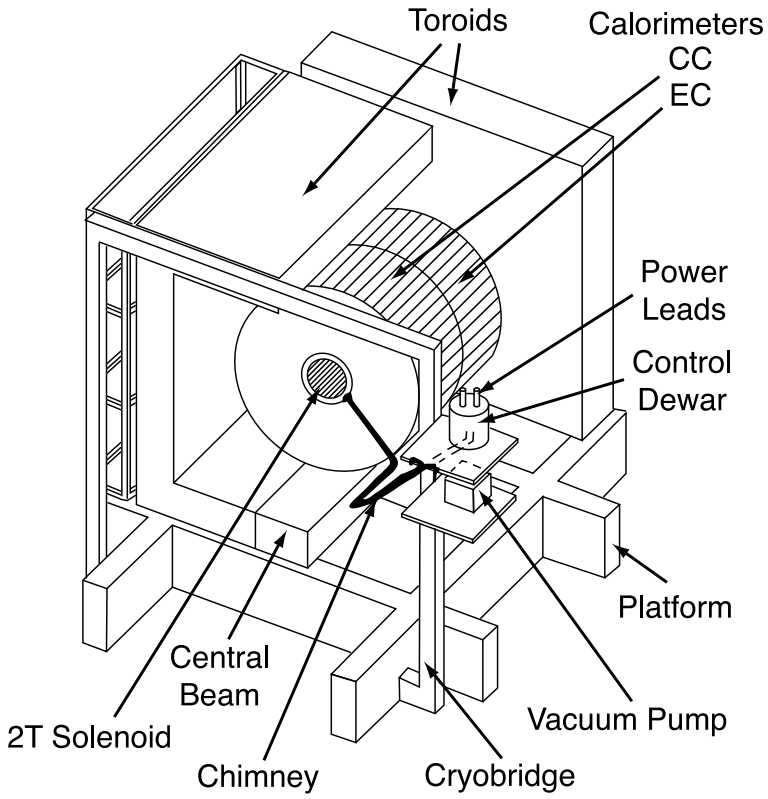


Figure 4.5: Perspective view of the solenoid inside the central calorimeter. One end calorimeter, several muon chambers, and parts of the toroids have been omitted for clarity. Also shown are the service chimney and control dewar [64].

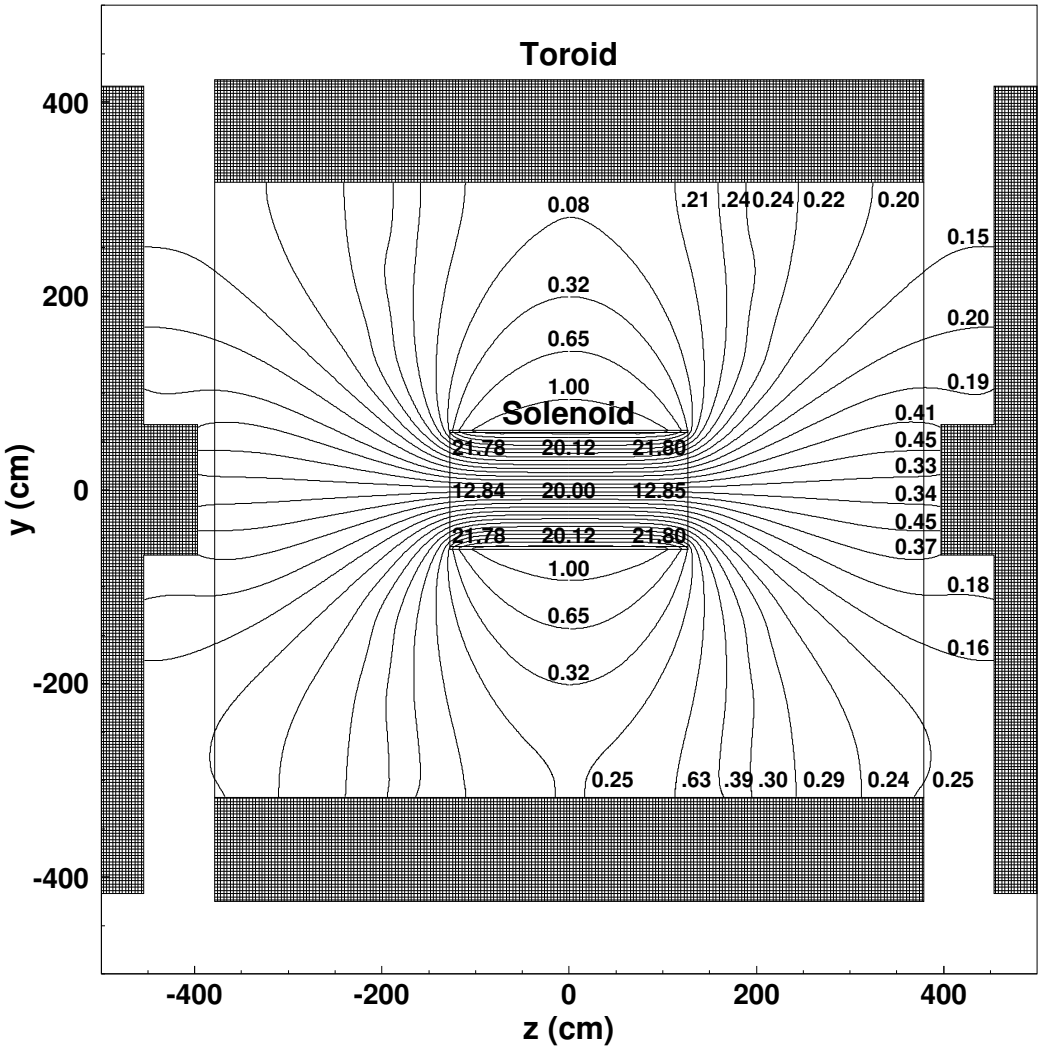


Figure 4.6: The $y-z$ view of the DØ magnetic field (in kG) with both the toroidal and solenoidal magnets at full current (1500 A and 4749 A, respectively). The field in the central toroid is approximately 1.8 T; that in the end toroids is about 1.9 T. The field lines are projections onto the $y-z$ plane; the left and right line ends differ by up to 2.5 m in x.

A set of scintillation counters, cover the top and upper sides of the outer layer of central muon PDTs, and 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

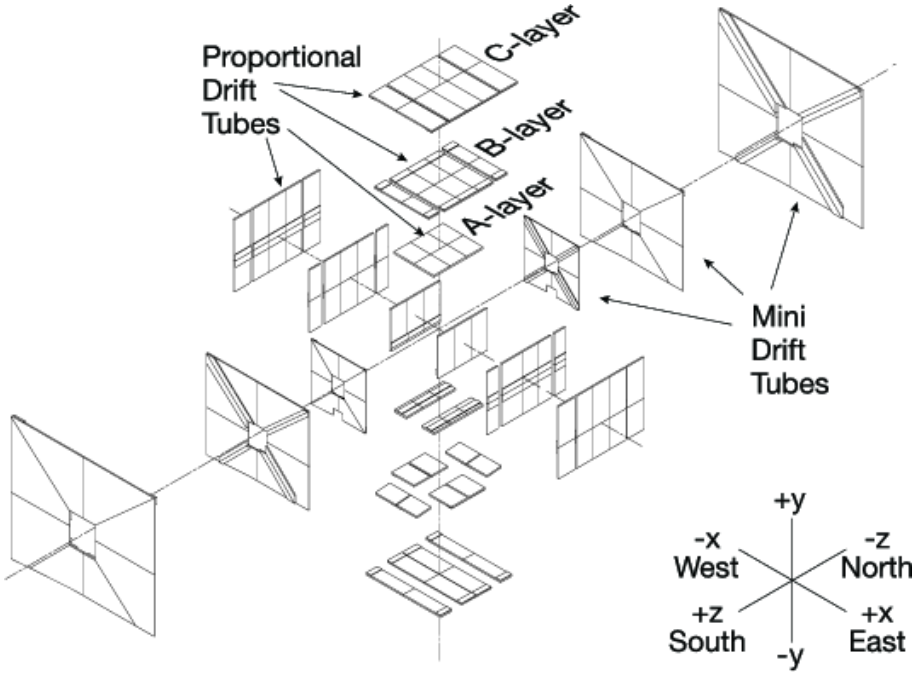


Figure 4.7: Exploded view of the muon wire chambers.

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 4.7 and 4.8.

4.3.1 Toroidal magnets

The toroidal magnets are described in detail in Ref. [65] and visible in Figures 4.1, 4.6, and 4.6. Having a stand-alone muon-system momentum measurement *i*) enables a low-p_T cutoff in the Level 1 muon trigger, *ii*) allows for cleaner matching with central detector tracks, *iii*) rejects π/K decays, and *iv*) improves the momentum resolution for high momentum muons.

The central toroid is a square annulus 109 cm thick whose inner surface is about

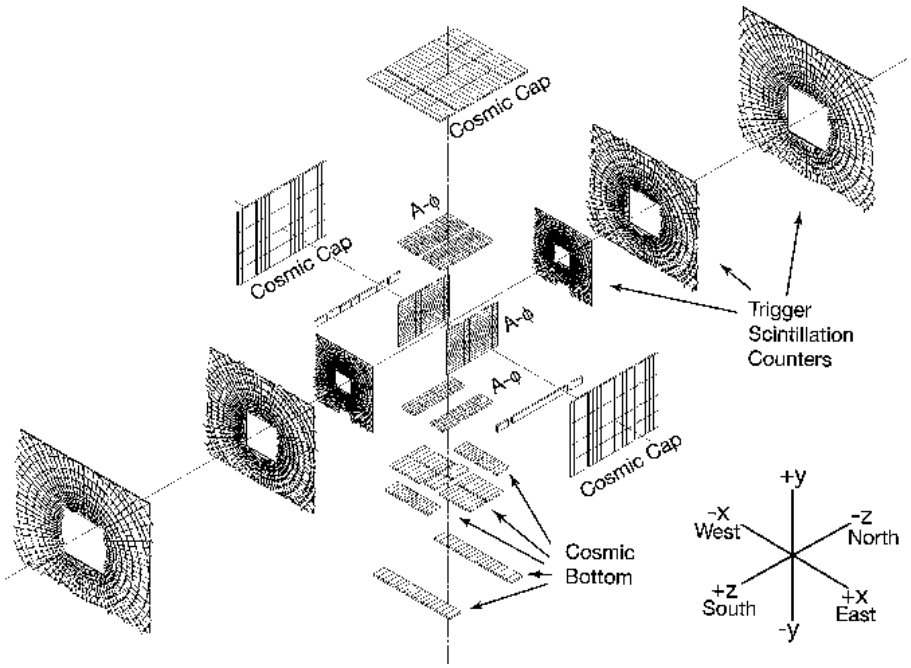


Figure 4.8: Exploded view of the muon scintillation detectors.

318 cm from the Tevatron beamline, it covers the region $|\eta| \lesssim 1$. To allow access to the inner parts of the detector, it was constructed in three sections. The center-bottom section is a 150-cm-wide beam, fixed to the detector platform, which provides a base for the calorimeters and central tracking detectors. Two C-shaped sections, which can be moved perpendicularly to the center beam, complete the central toroid. The magnet is wound using twenty coils of ten turns each. The two end toroids are located at $454 \leq |z| \leq 610$ cm. In the center of each end toroid there is a 183 cm square hole centered on the beamline, in x and y the magnets extend 426 cm from the beamline. The end toroid windings are eight coils of eight turns each.

The magnets were operated in series, with at a current of 1500 A. the internal fields in the central toroid were approximately 1.9 T and those in the end toroids were approximately 2.0 T. The polarity of the magnets during data collection is regularly reversed.

4.3.1.1 Central muon system

The central muon system is composed of the central toroid (Subsection 4.3.1), PDT chambers, the cosmic cap and bottom scintillation counters, and the $A\phi$ scintillation counters.

There are three layers of PDT chambers. The inner-most, called A-layer, is located inside the central toroid. The other two layers, B- and C-layers, are located outside. A PDT is a rectangular extruded aluminum tube, called drift-cell, filled with a mixture of 84% argon, 8% methane, and 8% CF_4. Each drift-cell has an anode wire at the center. Vernier cathode pads above and below the wire determine the hit position along the wire.[1]

A set of scintillation counters, called the cosmic cap and cosmic bottom, is installed on the top, sides and bottom of the C-layer. Scintillation counters are fast enough to allow the association of a muon in a PDT with the correct bunch crossing and to discriminate against the cosmic ray background.

The $A\phi$ scintillation counters cover the A-layer. They provide fast triggering, muon identification and out-of-time backscatter rejection from the forward direction. Together with the CFT, it is possible to trigger at Level 1 (section 4.3.2) for high-p_T single muon and low-p_T dimuon triggers.

4.3.1.2 Forward muon system

The forward muon system is composed of the end toroids (Subsection 4.3.1), MDT chambers, scintillation counters and shielding around the beam pipe.

The function of the MDT chambers is essentially the same as the PDT chambers in the central region. MDTs are chosen because they provide better resolution and radiation hardness than the PDTs. There are three layers of MDTs (A, B and C), with the A-layer inside the end toroids. MDTs have 8 drift cells, covered with a stainless steel foil and inserted into PVC sleeves, and filled with a mixture of 90% CF_4 and 10% methane.

[1] Vernier pads were only fully instrumented on the A-layer, and about 10% on B/C-layers for aging studies, since they were expensive and did not improve the resolution for tracks passing all three layers.

The muon trigger scintillation counters are installed over the three MDT layers. They provide good time resolution, background rejection and high muon detection.

Finally, the shielding to reduce non-muon background in the muon system consists of layers of iron (hadronic and electromagnetic absorber), polyethylene (neutron absorber) and lead (gamma rays absorber) in a steel structure around the beam pipe and low-beta quadrupoles.

4.3.2 Trigger

More than 2.5 million proton-antiproton bunch crossings occurred at DØ every second during a normal Store. Given the total SM cross section of the Tevatron collider, $\sigma_{tot} \approx 10^8$ nb [66], we expected on the order of 1 to 10 collision in each bunch crossing (pile-up). Since in general not every high energy collision leads to interesting physics[1] and it is also impossible to save all that information due to limited disk space (something similar happens at the LHC), the DØ experiment implements three trigger levels (Level 1 to Level 3, or L1, L2 and L3). Each succeeding level examines fewer events but in greater detail and complexity. Events rejected by any of the trigger levels are permanently lost. Accepted events are buffered at each level to prevent further lost of data.

During Tevatron collisions, the available trigger bandwidth had to be divided into different kind of physics and interests, without excluding the possibility of totally new and unexpected physics. For some physics analyses, such as b physics studies, things got more complicated and inefficient (e.g. pattern recognition) at higher luminosities due to the busy environment.

The L1 consists of hardware elements that process signals from the luminosity monitors, calorimeter, muon system and fiber tracker. The requirements at L1 are very basic, such as energy depositions or hit patterns associated with tracks above a preset momentum threshold (see also footnote 1 in page 61). The set of decisions (also called triggers) taken at L1 reduces the event rate from 2.5 MHz to about 2 kHz.

The L2 is composed by serial or parallelized CPU-based cards or programmable logic based cards for each detector subsystem, and a global processor for integration of the

[1]For calibrations purposes, 1 out of 800,000 beam crossings were randomly recorded (zero-bias trigger). Similarly, 1 out of 700,000 collisions, identified as signal coincidences in the luminosity monitors at the expected time, were saved to disk (minimum-bias trigger).

data. The L2 is able to identify physics objects such as jets, electrons, photons, missing energy in the calorimeter, muons and charged tracks (with certain p_T, ϕ, isolation or impact parameter). The global processor analyze all objects received from the detector subsystems and their correlations, to reduce the event rate to about 1 kHz.

Finally, these data are received by the L3 to be analyzed in much more detail. This process is performed in parallel by nearly 500 computers, programmed with complex algorithms. Decisions at L3 are based on complete physics objects and on the relationships between them (vertices, angular separations, etc.).

During Tevatron collisions, to further reduce the rate to an acceptable level, about 100 Hz, L3 triggers had to be "pre-scaled"; for example, 9 out of 10 events with two good muons and $p_T > 2$ GeV/c were randomly rejected between 260 and 310 $\mu b^{-1}s^{-1}$.

5. B_s^0 lifetime

5.1 Selection and Reconstruction

The update of this measurement is done using 10.7 fb^{-1} of data collected with the DØ detector from 2002 to 2010. At the end of the section results with the complete sample of 10.7 fb^{-1} are also shown. We take reconstructed events consistent with the decay mode $B_s^0 \to D_s^+ \mu^- X$. To process them we use the BANA package [67]. To select the events, we first look for a muon, then two tracks of opposite charge, that will be assigned as kaons coming from a ϕ, and finally a charged track that will be assigned the pion mass, and will be combined with the ϕ candidate to form a D_s^- candidate. The muon candidate is required to satisfy the following requirements:

- be detected in the layer inside the toroid and at least in one layer outside the toroid, *i. e.* `nseg` $= 3$;
- transverse momentum $p_T > 2.0$ GeV/c,
- total momentum $p > 3.0$ GeV/c,
- at least two hits in the CFT, N(CFT) > 1,
- at least two hits in the SMT, N(SMT) > 1,
- at least one matched central track,
- $\chi^2 < 25$ in the global muon fit.

Each of the two oppositely charged tracks consistent with coming from the ϕ, identifying them as kaons, must pass the following cuts:

- the invariant mass of the combined tracks must be in the range 1.008 GeV/$c^2 < M(K^+K^-) < 1.032$ GeV/c^2,

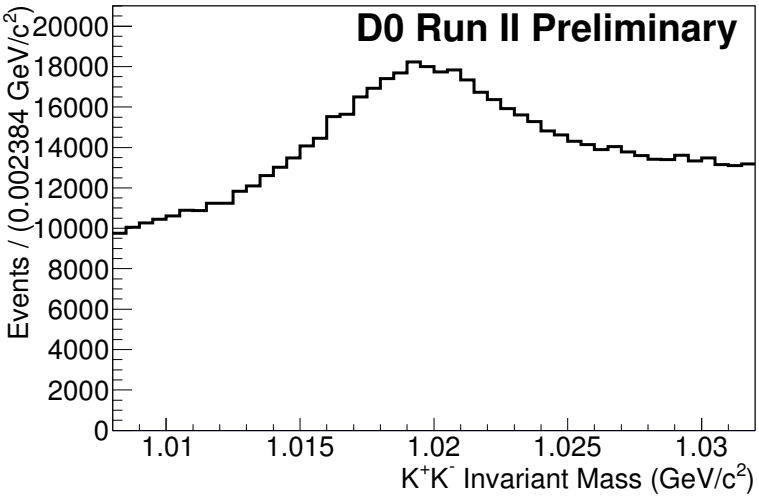


Figure 5.1: Mass Distribution of the ϕ candidates.

- transverse momentum $p_T > 1.0$ GeV/c,
- at least two hits in the CFT, N(CFT) > 1,
- at least two hits in the SMT, N(SMT) > 1,

The additional charged track, the π candidate, must pass:

- transverse momentum $p_T > 0.7$ GeV/c,
- at least two hits in the CFT, N(CFT) > 1,
- at least two hits in the SMT, N(SMT) > 1,
- the combination of this track with the ϕ candidate must fulfill 1.6 GeV/c^2 $< M(\phi\pi) <$ 2.3 GeV/c^2,
- the χ^2 of the vertex fit has to be smaller than 16.26, corresponding to a $p-$value larger than 0.1%,
- all tracks have to be in the same jet, and be associated with the same primary vertex,

in addition

- the invariant mass of the combination of the muon candidate with the D_s^- candidate must be in the range of 3 GeV/c^2 $< M(D_s^-\mu) <$ 5 GeV/c^2,

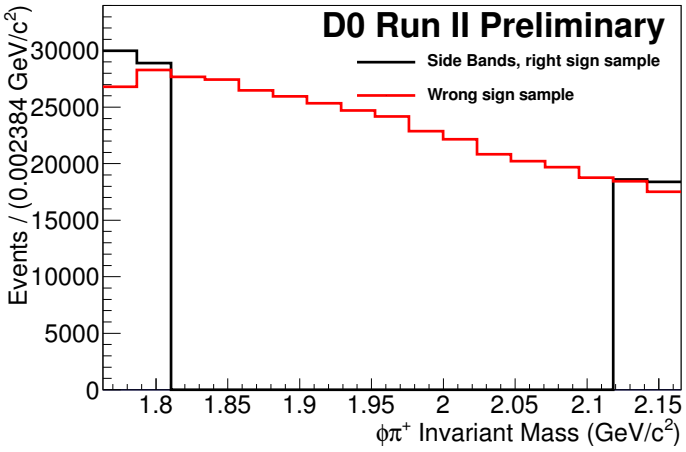


Figure 5.2: Mass distributions of the D_s^- candidates Side Band sample (**SB**, black) and Wrong Sign sample (**WS**, red)

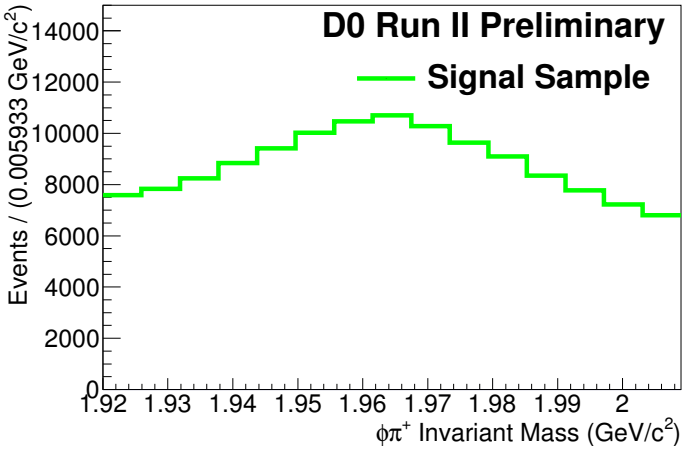


Figure 5.3: Mass distribution of the D_s^- candidates Signal Sample, the sample is the 2σ region around the D_s^+ mass peak.

- the χ^2 of the vertex fit has to be smaller than 15.13 , corresponding to a $p-$value larger than 0.01%.

At this point we have not made any restriction on the charge of the μ and π, however for the reconstructed B_s^0 meson signal candidates these two tracks have to have opposite charge, but we also keep the wrong combination sample to help modeling the background. In the right-sign sample, the reconstructed D_s^- is required to be displaced from the primary vertex in the same direction of its momentum in order to reduce

combinatoric background. In addition to the previous requirements,

We define the signal sample, **SS**, as the set of events that lay in the D_s^- mass window of $1.958 \pm 2 \times 0.0219$. The events with the right-sign combination but un the D_s^- mass regions of $[1.958-9\times0.0219, 1.958-7\times0.0219]$ and $[1.958+7\times0.0219, 1.958+9\times0.0219]$ form the side band sample, **SBS**. With the events with the wrong sign combination $\mu^\pm\pi^\pm$ we define the wrong sign sample, **WSS**. The union of the side band sample and the wrong sign sample define the background sample, **BS**.

5.2 Pseudo Proper Decay Length and K-factor

The B_s^0 lifetime, τ, is given in terms of the decay length, L, by the relation,

$$L = c\tau\beta\gamma = c\tau\frac{p}{m}, \tag{5.1}$$

where $c\tau$ is the proper decay length, m the invariant mass, and p the total linear momentum. Projecting onto the transverse plane, this relation gives:

$$L_{xy} = c\tau\frac{p_T}{m}, \tag{5.2}$$

where p_T is the transverse momentum of the B_s^0, and L_{xy} is the transverse decay length. The transverse decay length is defined as:

$$L_{xy} = \frac{\vec{X}\cdot\vec{p}_T}{|\vec{p}_T|}, \tag{5.3}$$

where $\vec{X}$ is the displacement vector from the primary to the secondary vertex in the transverse plane. In this case, as we use the semileptonic channel, we can not fully reconstruct the B_s^0 momentum. We use the combined momentum of the muon and D_s^- meson, $\vec{p}_T(D_s^\pm\mu^\mp)$ as our best approximation of the B_s^0 momentum. To model this effect in the fit, a correction factor, K, is introduced. The definition of this K factor is:

$$K = \frac{p_T(D_s^-\mu^+)}{p_T(B_s^0)}, \tag{5.4}$$

and is found using simulated Monte Carlo events. Taking this into account, we can find the B_s^0 lifetime using:

$$c\tau = K\lambda = K\frac{\vec{X}\cdot\vec{p}_T(D_s^-\mu)}{|\vec{p}_T(D_s^-\mu)|}, \tag{5.5}$$

where λ is defined as pseudo-proper decay length. The K-factor correction is applied statistically by smearing the exponential decay distribution as described below when extracting the $c\tau(B_s^0)$ from the pseudo-proper decay length in the lifetime fit.

We use the MC samples described in Sec. 5.3, where the B_s^0 has been forced to decay to semileptonic modes that include a strange charmed meson, D_s^-, D_s^*, D_{s0}^*, or $D_{s1}^{\prime *}$.

5.3 Monte Carlo samples

Simulated events are generated to study the different contributions of the all the possible decay channels that can contribute to our data sample and to obtain the K-factor distributions that are used to correct the pseudo proper decay length computed in our selection program. We have generated MC samples using PYTHIA [68] for the production and hadronization phase, and EvtGen [69] for decaying the b and c hadrons. We generated B_s^0 meson samples with $c\tau = 439$ microns, and no mixing. The signal sample includes contributions from $D_s^-\mu^+\nu$, $D_s^{*-}\mu^+\nu$, $D_{s0}^{*-}\mu^+\nu$, $D_{s1}^{\prime -}\mu^+\nu$ and $D_s^{(*)-}\tau^+\nu$.

To evaluate non-combinatorial backgrounds, some additional processes were generated including $\bar{B}^0 \to D_s^{(*)-}D^{(*)+}$, and $\bar{B}^- \to D_s^{(*)-}D^{(*)0}X$, where the "right-sign" $D_s^-\mu^+$ combination can be obtained by allowing $D^{(*)+/0}$ to decay through semileptonic modes. The $B_s \to D_s^{(*)-}D_s^{(*)+}X$ and $B_s \to D_s^{(*)-}D_s^{(*)+/0}X$ processes were also generated.

To be able to fully simulate these samples some kinematic cuts were applied prior to D0gstar, (using d0_mess filter): muons had to have $p_T > 1.9$ GeV/c and $|\eta| < 2.1$, kaons (and pions) from ϕ (D_s^-) had to have $p_T > 0.6$GeV/c and $|\eta| < 3.0$, and the p_T of the D_s^- had to be larger than 1.0 GeV/c. The samples were then processed using the standard full chain procedure D0gstar-D0simD0reco that are a GEANT-based [70] detector simulation. All samples were generated as private production but following the official production criteria.

The produced thumbnails were then filtered using the same procedure as in data, described in section 5.1.

5.3.1 Weighting process

The particles simulated are softer in momentum than in real data and there is no full agreement in some distributions between MC and real data. In order to fix this

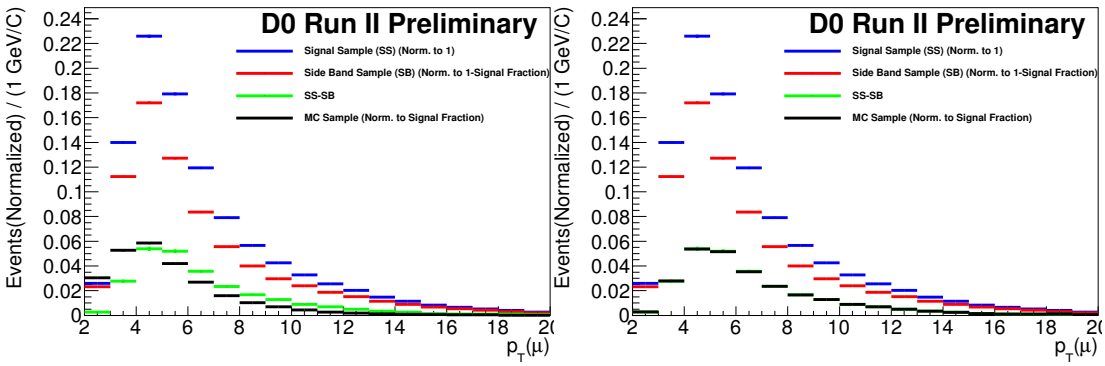


Figure 5.4: Blue and red show $p_T(\mu)$ distributions of the signal sample and sideband sample respectively. The green distribution shows the signal $p_T(\mu)$ distribution obtained subtracting the red distribution to the blue distribution. The black distribution show the MC $p_T(\mu)$ distribution, at the left plot with no weight applied, and weighted to match the signal distribution at the right plot.

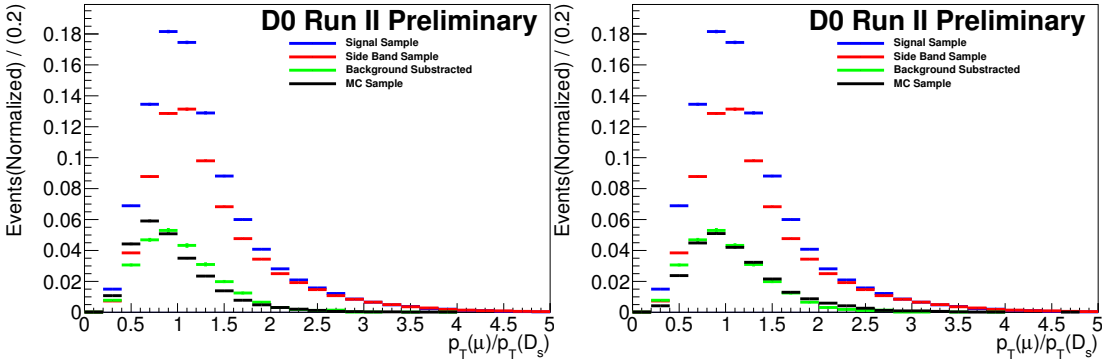


Figure 5.5: Blue and red show $p_T(\mu)/p_T(D_s)$ distributions of the signal sample and sideband sample respectively. The green distribution shows the signal $p_T(\mu)/p_T(D_s)$ distribution obtained subtracting the red distribution to the blue distribution. The black distribution show the MC $p_T(\mu)/p_T(D_s)$ distribution, at the left plot with no weight applied, and weighted at the right plot. The agreement between green and black distributions at the right plot shows that no trigger correction is needed.

discrepancy a weight is applied to each event in such a way that the MC transversal momentum distribution of the μ, $p_T(\mu)$, particle is identical to the same distribution in real data. To obtain the $p_T(\mu)$ distribution for signal in real data, the p_T distribution of the μ candidate in the side bands sample is subtracted from de same distribution in the signal sample.

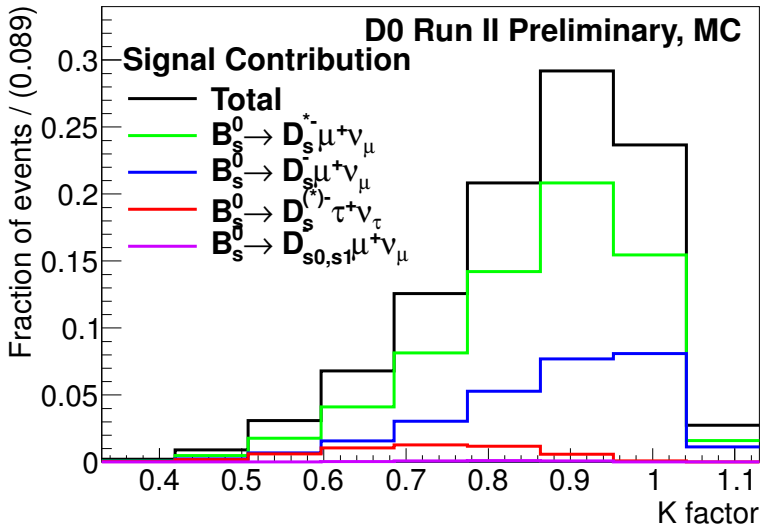


Figure 5.6: K-factor distributions of the decays contributing to the signal. All the histograms are normalized considering the contributions from Table 5.1. The black histogram shows the total K-factor of the signal.

In figure 5.5 there are plotted the distribution of the ratio of momentum among μ and D_s before and after applying the weights. As the main triggers used in this analysis are the single muon triggers, it could be possible that the correlation of μ and D_s transversal momenta could be altered. In the right plot at Fig. 5.5, after the weights were applied, there is a good agreement on MC and real data, we use this fact to justify that no trigger correction is necessary after the selection.

5.4 K-factors

The decay modes that are considered in our signal, include at least one particle that we cannot detect. As a result is necessary to use a correction factor, the so-called K-factor, to take into account the deficit of momentum corresponding to the undetected neutrino. The K factor is computed using MC samples for those decays that form the D_s^- meson mass peak. The signal decays that contribute to the D_s^- meson peak are listed in Table 5.1.

The reconstruction efficiency times the MC production efficiency of the signal is $\epsilon_{prod}\epsilon_{reco} = 12 \times 10^{-5}$, and we estimate that the signal contribution to the sample is

Decay Mode	BR	Contribution to Signal
$D_s^- \mu^+ \nu_\mu$	$(2.20 \pm 0.17)\%$	$(27.5 \pm 2.4)\%$
$D_s^{*-} \mu^+ \nu_\mu \times (D_s^{*-} \to D_s^- \gamma / D_s^- \pi^0)$	$(5.27 \pm 0.23)\%$	$(66.1 \pm 4.4)\%$
$D_{s(J)}^{*-} \mu^+ \nu_\mu \times (D_{s(J)}^{*-} \to D_s^{*-} \pi^0 / D_s^- \gamma)$	$(0.03 \pm 0.42)\%$	$(0.4 \pm 5.3)\%$
$D_s^{(*)-} \tau^+ \nu_\tau \times (\tau^+ \to \mu^+ \bar{\nu}_\mu \nu_\tau)$	$(2.65 \pm 1.28)\% \times (17.36 \pm 0.5)\% = (0.40 \pm 0.22)\%$	$(5.9 \pm 2.7)\%$

Table 5.1: Branching fractions of the decays that contribute to the signal. The total branching ratio of all these modes is $(7.90 \pm 0.55)\%$ and they contribute to the $(80.51 \pm 2.11)\%$ of the events in the D_s^+ mass peak.

Decay	$\sum \mathrm{Br}(\mathrm{B} \to \mathrm{D_s^0 D} \times \mathrm{D} \to \mu)$	$X_{section} \epsilon_{prod} \epsilon_{reco} (\mu\mathrm{b})$	Contribution to D_s^+
$B^+ \to D_s^+ DX$	$(0.11 \pm 0.02)\%$	0.71 ± 0.03	$(3.81 \pm 0.75)\%$
$B^0 \to D_s^+ DX$	$(0.13 \pm 0.02)\%$	0.64 ± 0.03	$(4.13 \pm 0.70)\%$
$B_s^0 \to D_s^+ D_s^{(*)} X$	$(0.06 \pm 0.01)\%$	0.37 ± 0.02	$(1.11 \pm 0.36)\%$
$B_s^0 \to D_s^+ DX$	$(0.13 \pm 0.06)\%$	0.14 ± 0.01	$(0.92 \pm 0.44)\%$
$c\bar{c} \to D_s^+ \mu^-$	$(0.12 \pm 0.02)\%$	1.56 ± 0.1	$(9.53 \pm 1.65)\%$

Table 5.2: Branching fractions of the background decays that contribute to the D_s^- meson mass peak, The B meson decays contribute to 9.98% of the D_s^- mass peak.

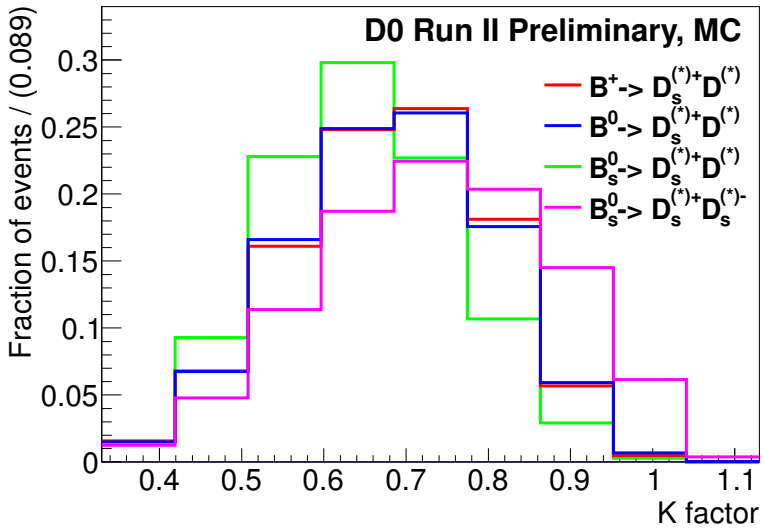


Figure 5.7: K-factor distributions of the decays contributing to the B-meson decays background. There are two B_s^0 decays decays considered in this background, the reason to take them as background is that they are not semileptonic

$(80.5 \pm 2.1)\%$. Also we need to compute a K factor distribution for each background contribution similar to our signal that comes from a B-meson and decays into a D_s^- meson. These decays are listed in Table 5.2. This table also shows the contribution of each decay to the signal sample. The total contribution of the B-meson decay background is 9.98%.

There is another important background contribution coming from events having a D-meson decaying to a μ and a D_s fully reconstructed. These events come from $c\bar{c}$ pairs created at the moment of the collision. The $c\bar{c}$ background have no K-factor associated because there is real B meson decay and they have lifetime equal to zero.

K-factor distributions for signal and background samples are shown in Figures 5.6 and 5.7 respectively.

5.4.1 $p_T(\mu)$ dependence

There is found that the signal K-factor have a slight dependence on the $p_T(\mu)$. When we plot a profile of the K-factor with respect to the $p_T(\mu)$ it is observed that the events with lower $p_T(\mu)$ have, in average, a lower K-factor. Taking this effect into account, the signal sample is divided in three $p_T(\mu)$ bins:

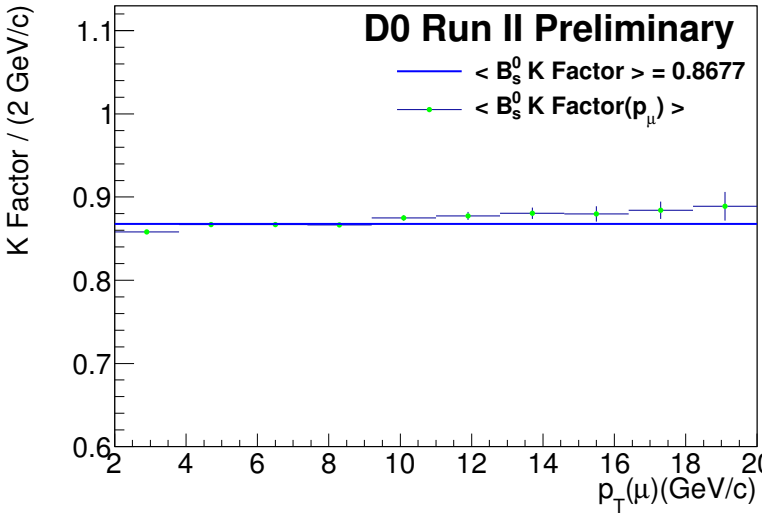


Figure 5.8: Dependence of K-factor on $p_T(\mu)$. The continuous line show the average K-factor in our MC signal sample. The distribution shows the average K-factor per each $p_T(\mu)$ bin.

- from 2 GeV/c to 4 GeV/c
- from 4 GeV/c to 9 GeV/c,
- larger than 9 GeV/c.

A K-factor distribution is computed for each of these subsamples.

5.5 Signal Fraction

To estimate the number of candidates of D_s^- mesons in the signal sample, a fit is made over the D_s^- mass distribution. The mass window taken is $1.958 \pm 10 \times 0.0219$ GeV/c^2 or 1.738 GeV/c^2 $< M(\phi\pi) <$ 2.178 GeV/c^2. We refer these events as the "D_s^- Mass Sample", **DMS**. The peak of the D^- meson also lays in this mass range. The model that is used to fit this distributions is comprised of a Gaussian distribution for the D_s^- mass, a Gaussian distribution for D^- mass, and a second-order polynomial for the

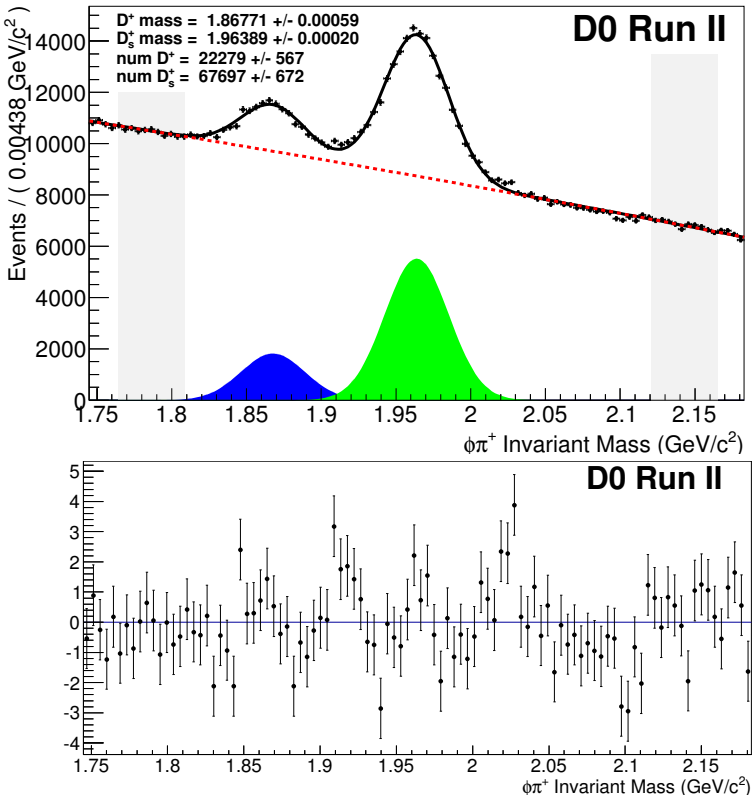


Figure 5.9: [Top figure] Invariant mass distribution of the D_s^- candidates. The curves are projections of the Unbinned maximum likelihood fit for the D_s^- mass distribution. The green area is the D_s^- mass peak, the blue one is the D^- mass peak and the red dashed curve is the combinatorial background distribution. The black curve is the sum of the three contributions. [Bottom figure] Residual plot of the mass unbinned maximum likelihood fit and the D_s^- distribution.

background, both Gaussian distribution have the same width σ,

$$\mathcal{L}_{Mass} = \prod_{i \in \mathbf{DMS}} \frac{1}{N_{\mathbf{DMS}}} \Big[N(D_s^-) e^{-\frac{(M_i - m(D_s^-))^2}{2\sigma^2}} + N(D^-) e^{-\frac{(M_i - m(D^-))^2}{2\sigma^2}} + N(Bkg)(1 + A_1 M_i + A_2 M_i^2) \Big]. \quad (5.6)$$

5.6 Lifetime Fit Model

To perform the lifetime fit we use two samples, the "Signal Sample", **SS**, and the "Background Sample", **BS**. The signal sample is defined as those events in the D_s^-

Parameter	Central Value	Statistical Uncertainty	
A_1	-0.136 (GeV/c^2)$^{-1}$	$+0.882$ (GeV/c^2)$^{-1}$	-0.904 (GeV/c^2)$^{-1}$
A_2	-0.631 (GeV/c^2)$^{-2}$	$+0.230$ (GeV/c^2)$^{-2}$	-0.225 (GeV/c^2)$^{-2}$
$m(D^-)$	1.8677 GeV/c^2	± 0.0006 GeV/c^2	
$m(D_s^-)$	1.9639 GeV/c^2	± 0.0002 GeV/c^2	
σ	0.0215 MeV/c^2	± 0.0002 GeV/c^2	
$N(Bkg)$	902705	$+1625$	-1655
$N(D^-)$	22279	$+663$	-649
$N(D_s^-)$	67697	$+898$	-876

Table 5.3: Parameter list of the mass distribution obtained from the maximum likelihood fit.

mass distribution which lay in the rank of $\pm 2 \times 21.9$ MeV from the fitted mean mass. This region is defined from 1920 MeV to 2007.6 MeV. The number of candidates in this sample is 64909 ± 644 over both periods, runs IIa and IIb . The Background Sample is obtained from sidebands of D_s^- mass distribution,$(-9 \times 21.9$ MeV$, -7 \times 21.9$ MeV$) \cup$ $(7 \times 21.9$ MeV$, 9 \times 21.9$ MeV$)$, and candidates with the "wrong-sign" combination in the interval $(-9 \times 21.9$ MeV$, 9 \times 21.9$ MeV$)$. We assume that the combinatoric background is due to random track combinations, so that, the side-band sample events can be used to model the background in the signal sample.

The pseudo-proper decay length distribution obtained from the signal sample is fit using an unbinned maximum log-likelihood method. The B_s^0 lifetime distribution, and background shape are determined in a simultaneous fit. As a crosscheck, we perform a simultaneous fit of the B_d^0 lifetime using the suppressed decay $B_d^0 \to D^- \mu X$. This measurement also helps for determining the ratio $\tau(B_s^0/\tau(B_d^0)$. The sample used for the B_d^0 is composed by those events in the rank $\pm 2 \times 21.4$ MeV around the D^- mass, 1867.75 MeV.

To model the B_s^0 lifetime distribution we use the approximation described at the end of section 2.9, where, instead of two exponential decays, we use only one exponential that measures the specific flavor lifetime. During the rest of this chapter, we use $\tau_{B_s^0}$ to actually mean $\tau_{B_s^0 \to sl}$ as given at Equation 2.84. The likelihood function $\mathcal{L}$ is defined

as follows:

$$\begin{aligned}\mathcal{L} &= \mathcal{C}_{signal} \prod_{i\in\mathbf{SS}} [f_{signal}\mathcal{F}^i_{signal} + (1-f_{signal})\mathcal{F}^i_{bkg}] \prod_{j\in\mathbf{BS}} \mathcal{F}^j_{bkg} \times \\ &\prod_{k\in\mathbf{B0}} [f_{B0}\mathcal{F}^k_{B0} + (1-f_{B0})\mathcal{F}^k_{bkg}], \end{aligned} \tag{5.7}$$

where the products run over the events on signal sample, **SS**, and background sample, **BS**; f_{signal} is the fraction of signal events in signal sample, estimated by the fit of the D_s^- mass distribution; and $\mathcal{F}^i_{signal/background}$ is the signal/background probability density evaluated for the i-th event. $\mathcal{C}_{signal}$ is a gaussian constriction on the signal fraction.

5.6.1 Signal Probability Density

The signal probability distribution $\mathcal{F}_{signal}$ composed of a weighted sum of several probability distributions, $\mathcal{E}_{decay}$, each of these distributions is defined as an exponential decay convoluted with a resolution function $\mathcal{R}$ and smeared with a K-factor distribution $\mathcal{H}_{decay}(K)$,

$$\begin{aligned}\mathcal{E}^j_{decay}(\lambda_j, \sigma(\lambda_j); s, \tau(B_{decay})) = \int dK \mathcal{H}_{decay}(K) \Bigg[\frac{K}{c\tau(B_{decay})} e^{-K\lambda_j / c\tau(B_{decay})} \\ \otimes \mathcal{R}_{tot}(\lambda_j, \sigma(\lambda_j), s) \Bigg], \end{aligned} \tag{5.8}$$

where λ_j is the pesudo-proper decay length, PPDL, and $\sigma(\lambda_j)$ is the error for the PPDL measurement for the j-Th. event. For each possible contribution from B-meson decay, a $\mathcal{E}$ function is introduced, with the correct PPDL and K-factor. The resolution function is given by,

$$\mathcal{R}(\lambda_j, \sigma(\lambda_j); s) = \left(\frac{1}{\sqrt{2\pi}\ s\sigma(\lambda_j)} \right) e^{\frac{-(\lambda_j - \lambda_0)^2}{2(s\sigma(\lambda_j))^2}}. \tag{5.9}$$

Since a priori we do not know the overall scale of the decay length uncertainty, which is estimated on a event-by-event basis, the scale factor, s, is introduced as a free parameter in the fit. As, in principle, events that lay in the central region of the tracking system have a better spatial resolution than those events collected in the forward/backward

region we use a resolution composed of two gaussian distributions, with different scale factors, given by,

$$\mathcal{R}_{tot}(\lambda_j, \sigma(\lambda_j); s_1, s_2) = f_{fine}\mathcal{R}(\lambda_j, \sigma(\lambda_j); s_1) + (1 - f_{fine})\mathcal{R}(\lambda_j, \sigma(\lambda_j); s_2) \tag{5.10}$$

In the fit, the integration over the K factor distribution is approximated by a sum as follows:

$$\int dK\ \mathcal{H}(K) \to \sum_{k-bin} \Delta K \mathcal{H}(K_{k-bin}), \tag{5.11}$$

where the sum is taken over the bin of the histogram $\mathcal{H}(K_{k-bin})$, with bin size ΔK.

Explicitly, the probability density $\mathcal{F}^i_{signal}$ takes the form:

$$\begin{aligned}\mathcal{F}^i_{signal} = f_{\bar{c}c}F^i_{\bar{c}c} + (1-f_{\bar{c}c})\Big[& f_{B_s\to D_s^- D}\mathcal{E}^i_{B_s\to D_s^- D} + f_{B^+\to D_s^- D}\mathcal{E}^i_{B^+\to D_s^- D} + \\ & f_{B_s\to D_s* D_s*}\mathcal{E}^i_{B_s\to D_s* D_s*} + f_{B_0\to D_s^- D}\mathcal{E}^i_{B_0\to D_s^- D} + \\ & \Big(1 - f_{B_s\to D_s^- D} - f_{B_s\to D_s* D_s*} - f_{B_0\to D_s^- D} - f_{B^+\to D_s^- D}\Big)\mathcal{E}^i_{B_s\to D_s^+\mu^- X}\Big],\end{aligned} \tag{5.12}$$

where $f_{\bar{c}c}$ is the expected fraction of $\bar{c}c$ events in the signal sample, $F_{\bar{c}c}$ is the PDF lifetime for the $\bar{c}c$ events (a gaussian distribution). $f_{B_a\to D_b D_c}$ is the fraction of events for the decay $f_{B_a\to D_b D_c}$, and $\mathcal{E}_{B_a\to D_b D_c}$ is the PPDL PDF associated to each decay. The last term of the sum in eq. 5.12 is the term associated to our signal events $B_s \to D_s^+\mu^- X$

5.6.2 Background Probability Density

The background probability density $\mathcal{F}_{bkg}$ is defined as follows:

$$\begin{aligned}\mathcal{F}^j_{bkg}(\lambda_j, \sigma(\lambda_j)) \;=\; & f_{CSL}\mathcal{E}(\lambda_j, \sigma(\lambda_j); s_1, s_2, \tau_{CSL}) + \\ & f_{CLL}\mathcal{E}(\lambda_j, \sigma(\lambda_j); s_1, s_2, \tau_{CLL}) + \\ & \Big[(1 - f_{SSL} - f_{SLL} - f_{CSL} - f_{CLL}) \times \\ & \mathcal{R}(\lambda_j, \sigma(\lambda_j); s_1, s_2)\Big] + \\ & \begin{cases} \frac{f_{SSL}}{\lambda_{SSL}} e^{\lambda_j/\lambda_{SSL}} + \frac{f_{SLL}}{\lambda_{SLL}} e^{\lambda_j/\lambda_{SLL}} & (\lambda < 0) \\ \frac{f_{SSL}}{\lambda_{SSL}} e^{-\lambda_j/\lambda_{SSL}} + \frac{f_{SLL}}{\lambda_{SLL}} e^{-\lambda_j/\lambda_{SLL}} & (\lambda \geq 0) \end{cases}\end{aligned} \tag{5.13}$$

where the details of the quantities in eq. 5.13 are shown in table 5.4

f_{CSL}	fraction of events in the convoluted exponential with Short PPDL
f_{CLL}	fraction of events in the convoluted exponential with Long PPDL
f_{SSL}	fraction of events in the symmetric exponential with Short PPDL
f_{SLL}	fraction of events in the symmetric exponential with Long PPDL
τ_{CSL}	slope of events in the convoluted exponential with Short PPDL
τ_{CLL}	slope of events in the convoluted exponential with Long PPDL
λ_{SSL}	slope of events in the symmetric exponential with Short PPDL
λ_{SLL}	slope of events in the symmetric exponential with Long PPDL

Table 5.4: List of parameters used in the Lifetime Unbinned maximum likelihood fit to model the combinatorial background.

5.6.3 B_d^0 Probability Density

The probability density for the B_d^0 sample is analogous to the Signal Probability Density, but changing the decays for those that have $D^-\mu X$ as final state, so the B_d^0 Probability Density is defined as:

$$\begin{aligned}\mathcal{F}^i_{B_d^0} &= f_{\bar{c}c}F^i_{\bar{c}c} + (1-f_{\bar{c}c})\Big[f_{B^+\to D^-\mu X}\mathcal{E}^i_{B^+\to D^-\mu X} + \\ &\quad f_{B_d^0\to D^-DX}\mathcal{E}^+_{B_d^0\to D^-DX} f_{B_s^0\to D^-\mu X}\mathcal{E}^i_{B_s^0\to D^-\mu X} \\ &\quad +(1-f_{B_d^0\to D^-DX} - f_{B_0\to D_s^- D} - f_{B_s^0\to D^-\mu X})\mathcal{E}^i_{B_d^0\to D^-\mu^+X}\Big],\end{aligned} \tag{5.14}$$

In analogy with the Signal Probability Density, the last term in eq. 5.14 corresponds to the events $B_d^0 \to D^-\mu^+X$.

5.7 Fit result

Table 5.5 shows the fit results for our nominal model. Where we have used the MINUIT [71] (MIGRAD, HESSE and MINOS) algorithms included in the RooFit [72, 73] package to perform the fit. Figure 5.10 shows the pseudo-proper decay length of the B_s^0 candidates in signal sample and the projections of $\mathcal{F}_{signal}, \mathcal{F}_{bkg}$ and $\mathcal{F}_{signal} + \mathcal{F}_{bkg}$. Figures 5.12 and 5.13 show the pseudo-proper decay length distribution for the side-band, and wrong-sign sample respectively. The Background PDF F_{bkg} is also projected on each distribution.

5.8 Crosschecks

As a crosscheck of the fitting procedure, the B_d^0 lifetime is measured. The channel $B_d^0 \to D^{(*)+}$ is used as the signal. To obtain the signal sample a 2σ mass window is taken around the D^+ mass from the same sample used for the B_s lifetime analysis. It is assumed also that there are some other B meson decays that contribute to the D^+ mass

Parameter	Central Value	Statistical Uncertainty	
f_{SSL}	0.1050	± 0.0020	
f_{SLL}	0.0346	± 0.0012	
f_{CLL}	0.0958	± 0.0031	
f_{CSL}	$0.1571e$	$+0.0028$	-0.0029
λ_{SSL}	114.3 μm	$+2.8$ μm	-2.7 μm
λ_{SLL}	481.9 μm	$+9.2$ μm	-8.8 μm
τ_{CLL}	616.6 μm	$+7.6$ μm	-7.5μm
τ_{CSL}	232.3 μm	$+4.0$ μm	-4.1 μm
f_{fine}	0.7997	$+0.0033$	-0.0034
$s1$	1.2983	$+0.0048$	-0.0049
$s2$	4.1734	$+0.0628$	-0.0618
$c\tau(B_d^0)$	458.9 μm	$\pm 6.4\mu$m	
$\tau(B_s)/\tau(B^0)$	0.9218	$+0.0140$	-0.0138
$c\tau(B_s)$	423.0 μm	$\pm 2.9\mu$m	

Table 5.5: Lifetime Maximum-Likelihood fit results. Two fits are performed, one for obtaining the ratio $\tau(B_s)/\tau(B^0)$ and the other for obtaining $c\tau(B_s)$, both give the same results for the other parameters.

Decay	Contribution
$c\bar{c}$	0.093
$B^0 \to D^- \mu^+ X$	0.808
$B^+ \to D^- \mu^+ X$	0.084
$B^0 \to D^- DX$	0.008
$B_s^0 \to D^- \mu^+ X$	0.005

Table 5.6: Contributions from different decays to the D^+ mass peak.

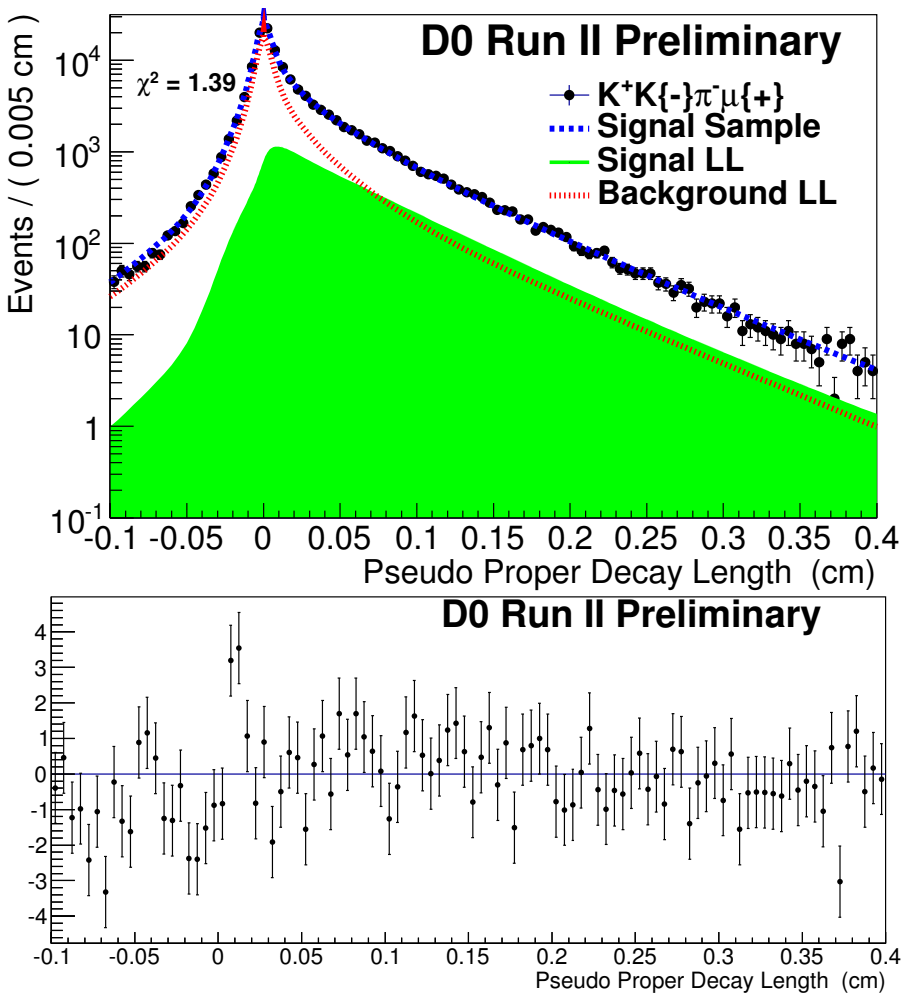


Figure 5.10: [Top figure] Distribution of pseudo proper decay length of the signal sample with projections of the Maximum-Likelihood fit. The green area is the projection of $\mathcal{F}_{signal}$, the red one is the combinatorial background and the blue curve is the sum of the two. [Bottom figure] residual of the lifetime unbinned maximum likelihood fit and the pseudo proper decay length in the region $[-0.1, 0.1]$.

peak and that the $c\bar{c}$ contribution is the same as in the D_s^- mass peak, $f_{c\bar{c}} = 0.093$. From MC samples we compute the contributions from different decays, they are shown in Table 5.6.

All K factors for these decays were computed and the fit performed with the proper correct definitions in the background samples as well. It was found that the B^0 lifetime

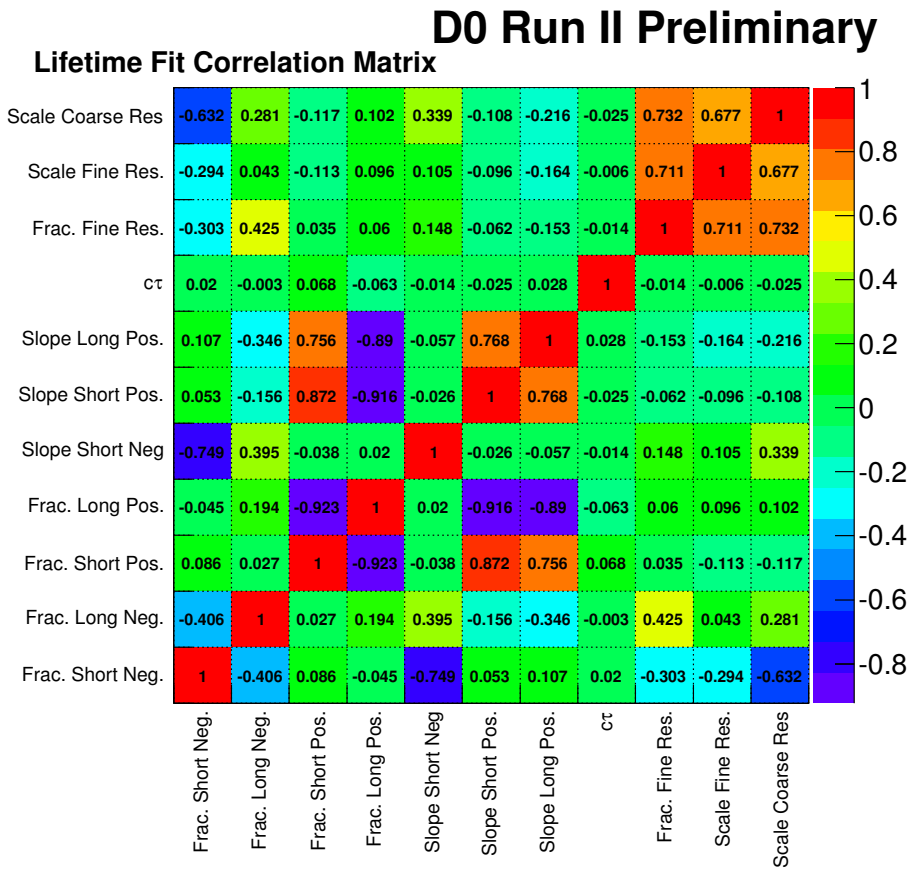


Figure 5.11: Correlation matrix of the lifetime fit parameters.

(times c) obtained is 458.9 ± 6.4 μm, this is in good agreement with the world average 455.7 ± 2.1 μm. The Figure 5.14 shows the distribution and the projection of the fit.

5.8.1 MC pseudo experiments generation and fit

To check that the fit is able to obtain the measurement of the lifetime, there were generated Monte Carlo pseudo experiments with lifetime equal to 432 μm. The distribution of the fit results is shown in Figure 5.15. Even when the distribution is not following perfectly a Gaussian distribution and the mean is slightly shifted, the result is consistent with the generation value.

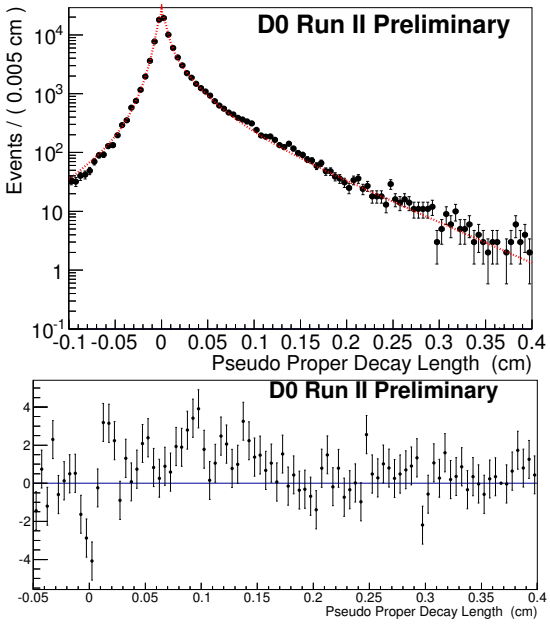


Figure 5.12: Background PDF component projected over the side band sample.

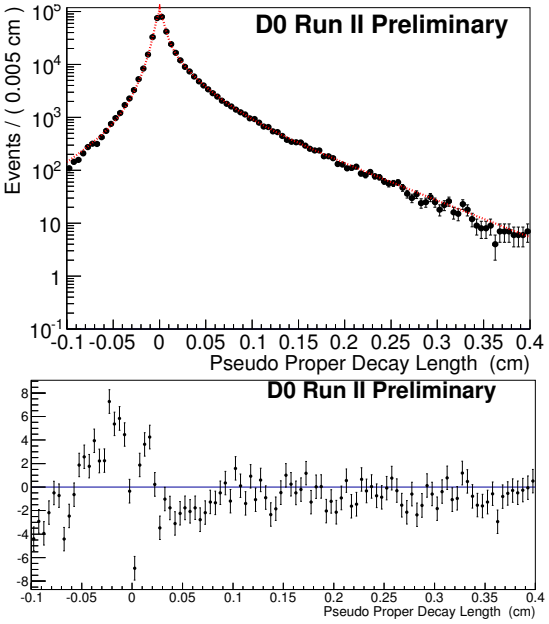


Figure 5.13: Background PDF component projected over the wrong sign sample.

Generated value (μm)	Fit value (μm)
420	423 ± 3
430	431 ± 3
440	440 ± 4
450	456 ± 4
460	468 ± 5

Table 5.7: Fit results over samples with different lifetime. The fitted lifetime in all samples are consistent with the input lifetime, with no evidence of bias in the fits.

5.8.2 Fit bias test

Another way to check that the fitting procedure is actually able to determine the lifetime, we fit several Monte Carlo samples generated with different lifetime. In the Table 5.7 are shown the results of the fits, all lifetime obtained by the fit are consistent with the generated values.

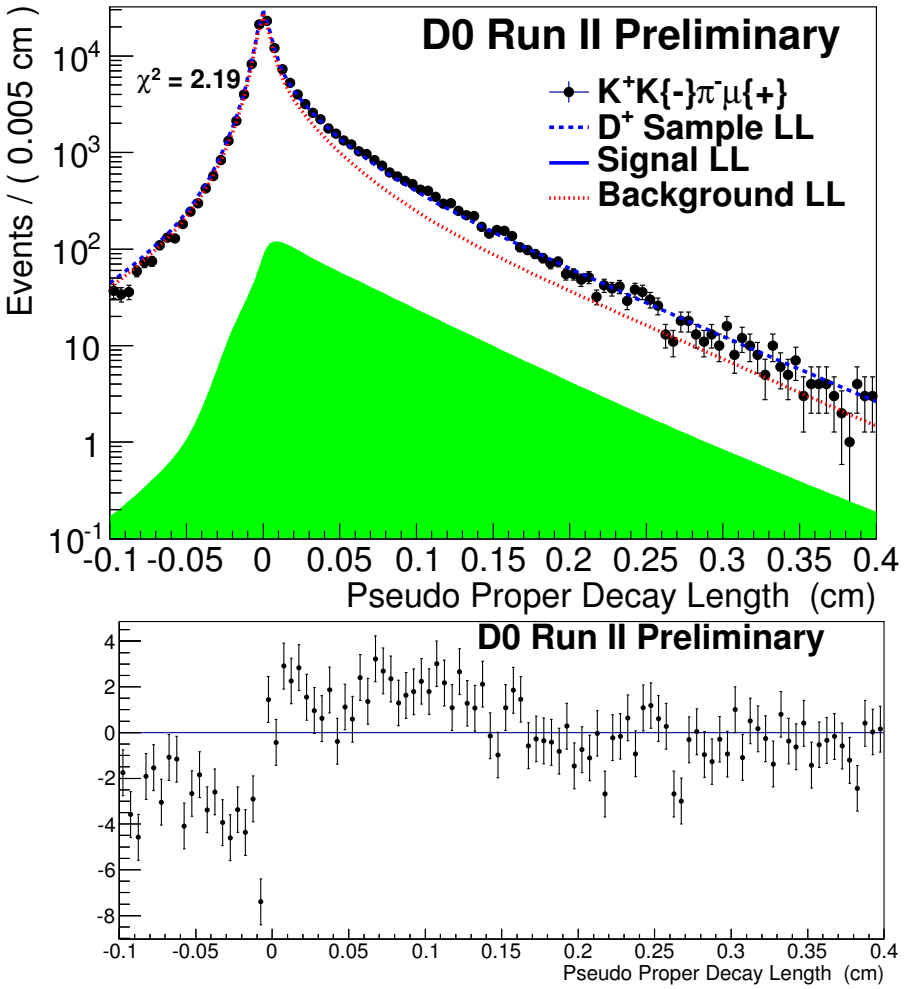


Figure 5.14: Distribution of pseudo proper decay length of the D^+ signal sample with projections of the Maximum-Likelihood fit. The blue area is the projection of $\mathcal{F}^{D^+}_{signal}$, the red one is the combinatorial background and the green curve is the sum of the two

5.9 Systematic uncertainties

5.9.1 Decay Length Resolution

In the nominal fit a double-Gaussian function is used to model the resolution. The fit gives a fraction of 78.26% for the contribution of the narrower gaussian. However, a single gaussian resolution can also be a valid model for the resolution. The systematic uncertainty due to this change in the resolution model is 6.7 μm

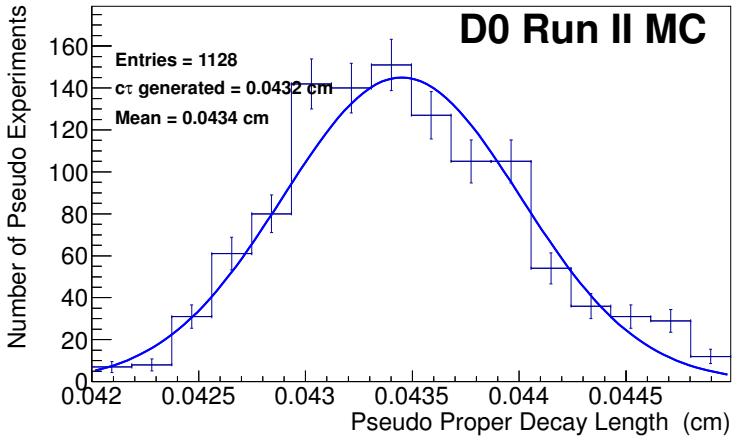


Figure 5.15: Fit over the pseudo-experiments output distribution. For each pseudo-experiment a MC sample of the same number as the real data sample is generated following the lifetime PDF with the parameters fixed to the nominal result.

Model	Variation
no Side Bands	$+3.7$ μm
no Low Side Bands	$+5.7$ μm
no High Side Bands	-2.3 μm

Table 5.8: List of systematic uncertainties associated with the combinatorial background modeling.

5.9.2 Combinatorial Background Evaluation

To evaluate the shape of the combinatorial background in the Signal Sample, side-band events and wrong-sign sample are used to help the fit. To associate a systematic uncertainty of this method we use a sample with only wrong sign combination, a sample using only the high-mass and wrong-sign and a sample with low mass and wrong-sign. Table 5.8 shows the difference associated with each sample. The maximum difference is 5.7 μm with no side-bands, and is the taken as uncertainty for this evaluation. This uncertainty is only positive but it is taken symmetric in the total effect.

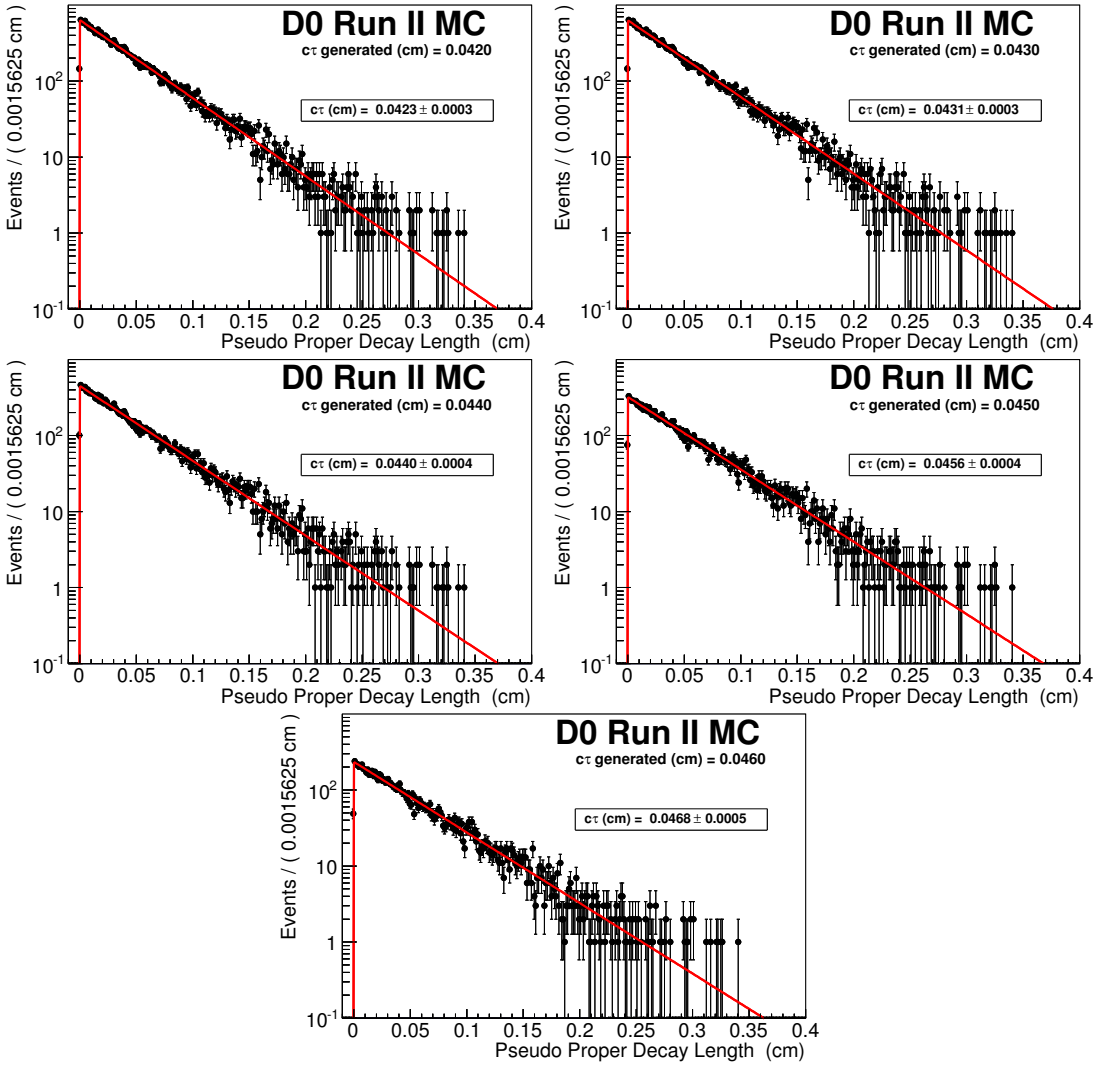


Figure 5.16: Lifetime fit projection over samples with different lifetime. The fitted lifetime in all samples are consistent with the input lifetime, with no evidence of bias in the fits.

5.9.3 K-Factor Determination

To determine the K-factor of the signal events, we assume that the signal is composed according to the fractions given in the table 5.1. To associate a systematic uncertainty to the determination of the K-factor we change this assumption considering that the signal is composed only of the $B_s \to \mu\nu D_s^-$ and $B_s \to \mu\nu D_s^*$ decay modes. The change

Decay	-1σ	$+1\sigma$
$c\bar{c} \to D_s^+ \mu X$	-4.1 μm	$+4.2$ μm
$B_s^0 \to D_s^+ DX$	$+0.8$ μm	-0.8 μm
$B_s^0 \to D_s^{+(*)} D_s^{-(*)}$	$+0.4$ μm	-0.4 μm
$B^+ \to D_s^+ DX$	$+1.7$ μm	-1.8 μm
$B^0 \to D_s^+ DX$	$+1.2$ μm	-1.2 μm

Table 5.9: Systematic uncertainties associated to the variation of the contributions of the different decays contributing to the non combinatoric background. A total uncertainty of $^{+4.8}_{-4.7}$ μm is obtained when each is added in quadrature.

in the lifetime measurement using this assumption is -2.5 μm

5.9.4 Non-Combinatorial Background Composition

In the case of the non-combinatorial background, *i.e.*, the background due to a incomplete reconstruction of a B-meson decay, there is a systematic effect that is due to the uncertainty in their contribution of D_s^- mesons. There is also a contribution from $c\bar{c} \to D_s^- D$ events that it is necessary to take into account. Table 5.9 summarizes the effect of varying each decay fraction. The total effect of these variations, summed in quadrature is $^{+4.8}_{-4.5}$ μm.

5.9.5 B-mesons lifetime

In the likelihood fit, the B^+ and B^0 lifetime are set to the world average, *i.e.*, 492 ± 0.2 μm and 456 ± 0.2 μm respectively. If we vary the lifetimes in one σ, a change of ± 0.1 μm for the B^+ and no change for the B^0.

5.9.6 Summary

Table 5.10 shows the contributions to the systematic uncertainty from different sources. The most significant effect comes from the combinatorial background evaluation. The total effect, computed adding all of them in quadrature, is $^{+10.0}_{-5.4}$ μm

Uncertainty source	Estimate
Resolution Modeling	$+6.7$ μm
Combinatorial Background modeling	$^{+5.7}_{-2.3}$ μm
K-factor determination	-2.5 μm
Non Combinatorial Background	$^{+4.8}_{-4.5}$ μm
B-mesons lifetime	± 0.1 μm
Total	$\mathbf{^{+10.0}_{-5.8}}$ μm

Table 5.10: Summary of systematic uncertainty contributions.

5.10 Results and conclusions

Using 8.1 fb^{-1} of data collected with the DØ detector we measured the B_s^0 lifetime in the inclusive semileptonic channel $B_s^0 \to D_s^- \mu^+ + X$. We found:

$$\tau_{B_s} = 423.0 \pm 2.9 \ \mu\text{m (stat.)}^{+10.0}_{-5.6} \ \mu\text{m (syst.)}, \tag{5.15}$$

that is consistent with the world average, as is shown in figure 7.1. With respect to the DØ previous measurement, we reduced the statistical uncertainty by a factor of 3. But the systematic uncertainty was not reduced. The uncertainty in this measurement is dominated by the systematic effects. Taking the world average of B^0 lifetime and this measurement, one can compute the ratio $\tau_{B_s^0}/\tau_{B^0}$,

$$\frac{\tau_{B_s^0}}{\tau_{B^0}} = 0.930^{+0.022}_{-0.014}. \tag{5.16}$$

If we take our measurement of the B^0 lifetime, the ratio becomes

$$\frac{\tau_{B_s^0}}{\tau_{B^0}} = 0.922^{+0.032}_{-0.026}. \tag{5.17}$$

This result is consistent, but with some tension, with the prediction of inequality 2.85.

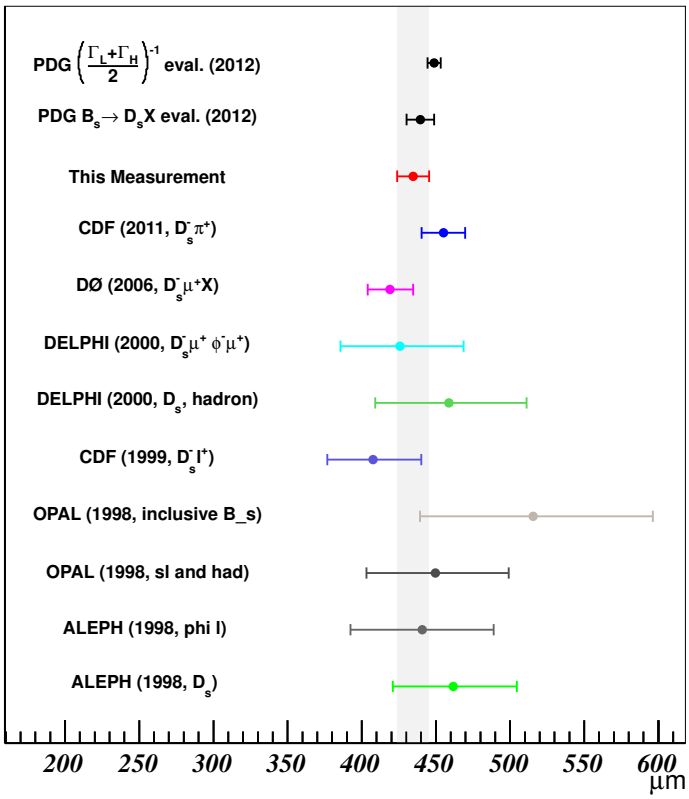


Figure 5.17: B_s lifetime (times c) measurements [2].

6. CP violation phase in the decay $B_s^0 \to J/\Psi$

6.1 Introduction

The meson-antimeson mixing and the phenomenon of charge-conjugation-parity *(CP)* violation in neutral mesons systems are key problems of particle physics. In the standard model (SM), the light (L) and heavy (H) mass eigenstates of the B_s^0 system are expected to have sizeable mass and decay width differences: $\Delta M_s \equiv M_H - M_L$ and $\Delta\Gamma_s \equiv \Gamma_L - \Gamma_H$. The two mass eigenstates are expected to be almost pure *CP* eigenstates. The *CP*-violating phase that appears in $b \to c\bar{c}s$ decays is due to the interference of the decay with and without mixing, and it is predicted [74, 75] to be $\phi_s^{J/\psi\phi} = -2\beta_s^{SM} = 2\arg[-V_{tb}V_{ts}^*/V_{cb}V_{cs}^*] = -0.038 \pm 0.002$, where V_{ij} are elements of the Cabibbo-Kobayashi-Maskawa quark-mixing matrix [29]. New phenomena [76–86] may alter the observed phase to $\phi_s^{J/\psi\phi} \equiv -2\beta_s \equiv -2\beta_s^{SM} + \phi_s^{\Delta}$. A significant deviation from $\phi_s^{J/\psi\phi}$ from its small SM value would indicate the presence of processes beyond SM.

The analysis of the decay chain $B_s^0 \to J/\psi\phi$, $J/\psi \to \mu^+\mu^-$, $\phi \to K^+K^-$ separates the *CP*-even and *CP*-odd states using the angular distributions of the decay products as a function of proper decay time. The first direct constraint on $\phi_s^{J/\psi\phi}$ [87] was derived by analysing $B_s^0 \to J/\psi\phi$ decays where the flavor (i.e., B_s^0 or $\overline{B}_s^0$) at the time of production was not determined ("tagged"). It was followed by an improved analysis [88], based on 2.8 fb^{-1} of integrated luminosity, that included the information on the B_s^0 flavor at production. The CDF collaboration has performed a measurement [89] of $\phi_s^{J/\psi\phi}$ using 1.35 fb^{-1} of data.

In this Article, we present new results from the time-dependent amplitude analysis of the decay $B_s^0 \to J/\psi\phi$ using a data sample corresponding to an integrated luminosity

of 8.0 fb^{-1} collected with the D0 detector [64] at the Fermilab Tevatron Collider. We measure $\Delta\Gamma_s$; the average lifetime of the B_s^0 system, $\overline{\tau}_s = 1/\overline{\Gamma}_s$, where $\overline{\Gamma}_s \equiv (\Gamma_H + \Gamma_L)/2$; and the CP-violating phase $\phi_s^{J/\psi\phi}$. Chapter 4 describes the D0 detector. Section 6.2 presents the event reconstruction and the data set. Sections 6.3 and 6.4 describe the event selection requirements and the procedure of determining the flavor of the initial state of the B_s^0 candidate. In Sec. 6.5 we describe the analysis formalism and the fitting method, present fit results, and discuss systematic uncertainties in the results. We obtain the confidence level (C.L.) intervals for physics parameters using a procedure based on the Markov Chain Monte Carlo (MCMC) technique, presented in Sec. 6.6. We summarize and discuss the results in Sec. 6.7.

6.2 Data Sample and Event Reconstruction

The analysis presented here is based on data accumulated between February 2002 and June 2010. Events are collected with a mixture of single- and dimuon triggers. They are rejected if they only satisfy triggers that impose a requirement on the track impact parameter with respect to the $p\overline{p}$ interaction vertex. This is done to avoid introducing a bias in the B_s^0 lifetime distribution.

Candidate $B_s^0 \to J/\psi\phi$, $J/\psi \to \mu^+\mu^-$, $\phi \to K^+K^-$ events are required to include two opposite-sign muons accompanied by two opposite-sign 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 oppositely charged 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 [2] 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, 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

$$t = \frac{M_{B_s} \vec{L}_{xy}^B \cdot \vec{p}}{p_T^2} \tag{6.1}$$

where M_{B_s} is the world-average B_s^0 mass [2], 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.

6.3 Background Suppression

The selection criteria are designed to optimimize 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.

6.3.1 Signal and background simulation

Three Monte Carlo (MC) samples are used to study background suppression: signal, prompt background, and non-prompt background. All three are generated with PYTHIA [68]. Hadronization is also done in PYTHIA, but all hadrons carrying heavy flavors are passed on to EvtGen [69] 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

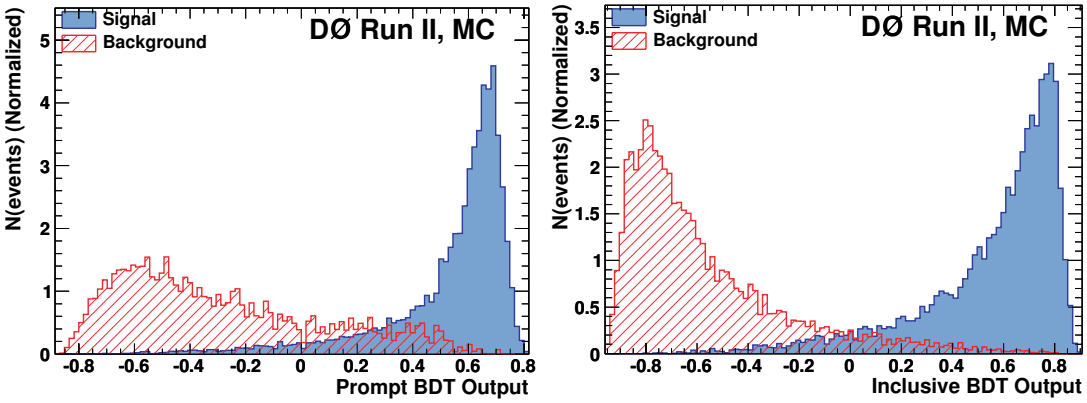


Figure 6.1: (color online). BDT discriminant output for the prompt (top) and non-prompt (bottom) classifiers. The signal and background events are taken from simulation. Events used for BDT training are excluded from these samples.

approximately 10^6 events in each background and the signal MC samples. All events are passed through a full GEANT-based [70] 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 Appendix B).

6.3.2 Multivariate event selection

To discriminate the signal from background events, we use the TMVA package [90]. 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 Appendix A.

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 6.1 shows the BDT output discriminant for the prompt and non-prompt cases.

6.3.3 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 6.1. Figure 6.2 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 using only the variables whose importance (see Appendix A) exceeds 0.01, giving 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.

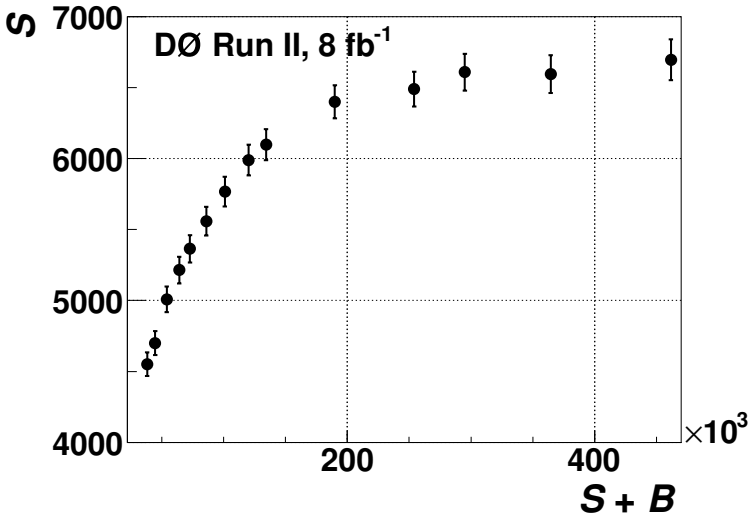


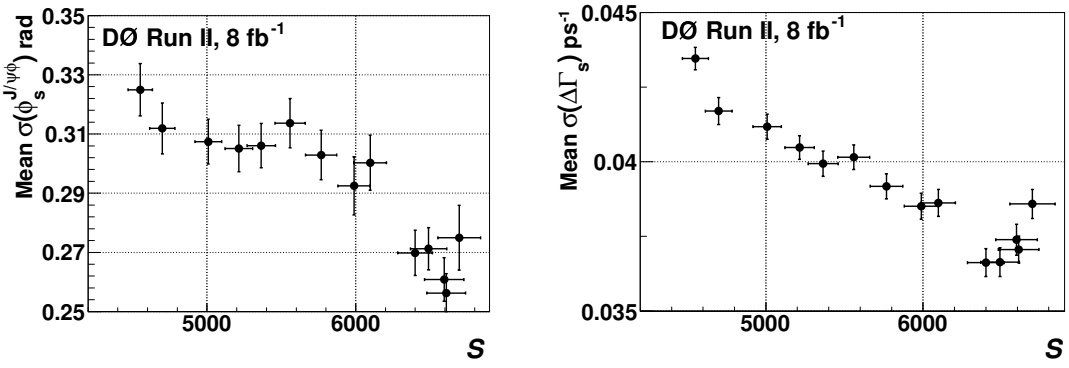
Figure 6.2: Number of $B_s^0 \to J/\psi\phi$ signal events as a function of the total number of events for the 14 criteria sets considered.

The choice of the final cut on the BDT output is based on an ensemble study. For

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

Table 6.1: 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.

each point in Table 6.1, 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 candidate 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 $\phi_s^{J/\psi\phi}$ ($\sigma(\phi_s^{J/\psi\phi})$) 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 Figs. 6.3(a) and 6.3(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



(a) (b)

Figure 6.3: 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.

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 6.1) as a final selection and refer to it as the "BDT selection". Figure A.3 in Appendix A shows the ratios of the normalized distributions of the three angles (see Section 6.5) 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.

6.3.4 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. [87] and [88]. 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.

6.4 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 are based primarily on the presence of a muon or an electron from the semi-leptonic decay of the other b hadron produced in the $p\overline{p}$ interaction. If a charged lepton is not found, the algorithm attempts to reconstruct the decay vertex of the opposite-side b hadron and determine the net charge of particles forming the vertex.

The OST algorithm, based on the Likelihood Ratio method, assigns to each event a value of the predicted tagging parameter d, in the range [−1,1], with $d > 0$ tagged as an initial b quark and $d < 0$ tagged as an initial $\overline{b}$ 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. [91].

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

$$\mathcal{D} = \frac{N_{\text{cor}} - N_{\text{wr}}}{N_{\text{cor}} + N_{\text{wr}}}, \tag{6.2}$$

where N_{cor} (N_{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.

6.4.1 OST calibration

The dilution calibration is based on four independent $B_d^0 \to \mu\nu D^{*\pm}$ 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. [91].

For each sample we perform an analysis of the $B_d^0 - \overline{B}_d^0$ oscillations described in Ref. [92]. 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 Fig. 6.4. The dependence of the dilution on $|d|$ is parametrized as

$$|\mathcal{D}| = \frac{p_0}{(1+\exp((p_1-|d|)/p_2))} - \frac{p_0}{(1+\exp(p_1/p_2))}. \tag{6.3}$$

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

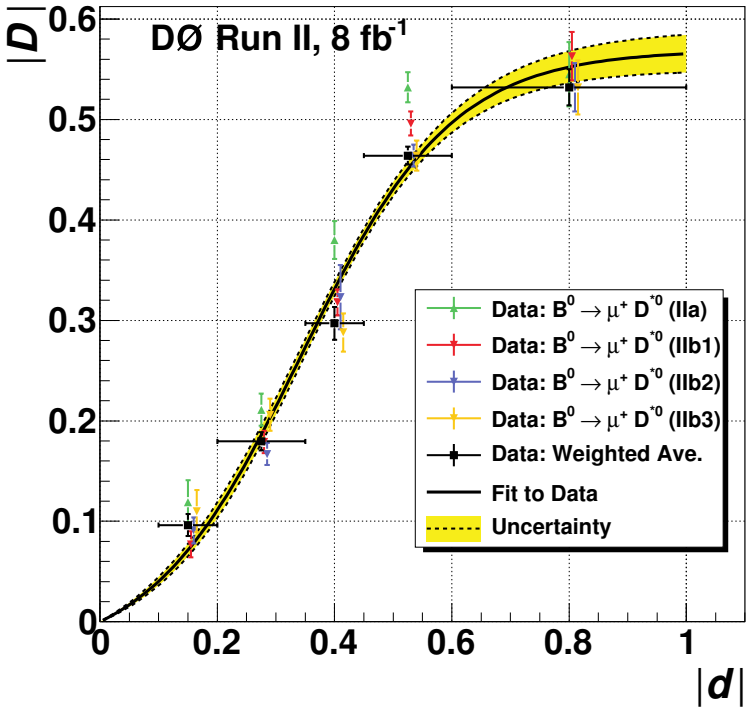


Figure 6.4: (color online). Parametrization of the dilution $|\mathcal{D}|$ as a function of the tagging parameter $|d|$ for the combined opposite-side tagger. The curve is the result of the weighted fit to four self-tagging control data samples (see text).

6.5 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 adopt the formulae and notation of Ref. [93]. 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 [73].

6.5.1 Signal model

The angular distribution of the signal is expressed in the transversity basis. 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 Fig. 6.5.

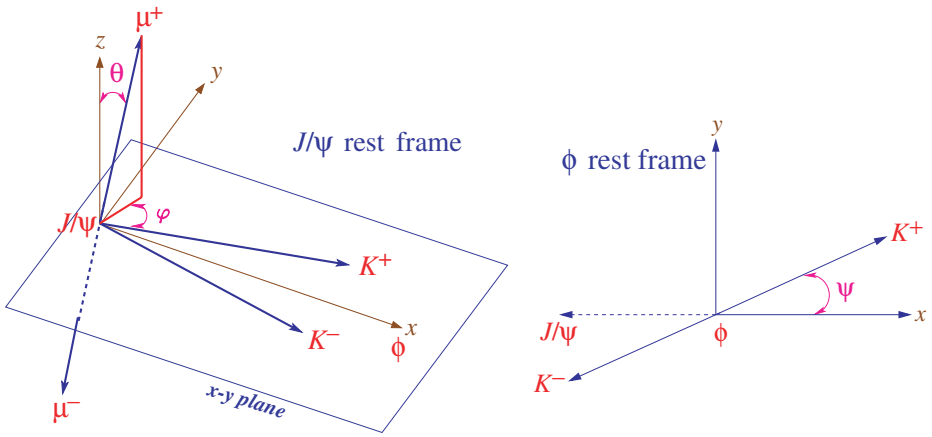


Figure 6.5: (color online). Definition of the transversity polar and azimuthal angles θ and φ and the angle ψ.

In this basis, the decay amplitude of the B_s^0 and $\overline{B}_s^0$ mesons is decomposed into

three independent components corresponding to linear polarization states of the vector mesons J/ψ and ϕ, which are polarized either longitudinally (0) or transversely to their direction of motion, and parallel ($\|$) or perpendicular ($\perp$) to each other.

The time dependence of amplitudes $\mathcal{A}_i(t)$ and $\bar{\mathcal{A}}_i(t)$ (i denotes one of $\{||, \perp, 0\}$), for B_s^0 and $\overline{B}_s^0$ states to reach the final state $J/\psi\ \phi$ is:

$$\begin{aligned} \mathcal{A}_i(t) &= F(t)\left[E_+(t) \pm e^{2i\beta_s} E_-(t)\right] a_i\,, \\ \bar{\mathcal{A}}_i(t) &= F(t)\left[\pm E_+(t) + e^{-2i\beta_s} E_-(t)\right] a_i\,, \end{aligned} \tag{6.4}$$

where

$$F(t) = \frac{e^{-\Gamma_s t/2}}{\sqrt{\tau_H + \tau_L \pm \cos 2\beta_s\,(\tau_L - \tau_H)}}\,, \tag{6.5}$$

and τ_H and τ_L are the lifetimes of the heavy and light B_s^0 eigenstates.

In the above equations the upper sign indicates a *CP*-even final state, the lower sign indicates a *CP*-odd final state,

$$E_\pm(t) \equiv \frac{1}{2}\left[e^{(\frac{-\Delta\Gamma_s}{4} + i\frac{\Delta M_s}{2})t} \pm e^{-(\frac{-\Delta\Gamma_s}{4} + i\frac{\Delta M_s}{2})t}\right], \tag{6.6}$$

and the amplitude parameters a_i give the time-integrated decay rate to each of the polarization states, $|a_i|^2$, satisfying:

$$\sum_i |a_i|^2 = 1\,. \tag{6.7}$$

The 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, are

$$\begin{aligned} P_B(\theta, \varphi, \psi, t) &= \frac{9}{16\pi}|\mathbf{A}(t) \times \hat{n}|^2, \\ P_{\bar{B}}(\theta, \varphi, \psi, t) &= \frac{9}{16\pi}|\bar{\mathbf{A}}(\mathbf{t}) \times \hat{n}|^2, \end{aligned} \tag{6.8}$$

where $\hat{n}$ is the muon momentum direction in the J/ψ rest frame,

$$\hat{n} = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)\,, \tag{6.9}$$

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

$$\begin{aligned} \mathbf{A}(t) &= \left(\mathcal{A}_0(t)\cos\psi, -\frac{\mathcal{A}_\parallel(t)\sin\psi}{\sqrt{2}}, i\frac{\mathcal{A}_\perp(t)\sin\psi}{\sqrt{2}}\right), \\ \bar{\mathbf{A}}(t) &= \left(\bar{\mathcal{A}}_0(t)\cos\psi, -\frac{\bar{\mathcal{A}}_\parallel(t)\sin\psi}{\sqrt{2}}, i\frac{\bar{\mathcal{A}}_\perp(t)\sin\psi}{\sqrt{2}}\right). \end{aligned} \tag{6.10}$$

The values of $\mathcal{A}_i(t)$ at $t = 0$ are denoted as A_i. They are related to the parameters a by

$$\begin{aligned} |A_\perp|^2 &= \frac{|a_\perp|^2 y}{1 + (y-1)|a_\perp|^2}, \\ |A_{||}|^2 &= \frac{|a_{||}|^2}{1 + (y-1)|a_\perp|^2}, \\ |A_0|^2 &= \frac{|a_0|^2}{1 + (y-1)|a_\perp|^2}, \end{aligned} \tag{6.11}$$

where $y \equiv (1-z)/(1+z)$ and $z \equiv \cos 2\beta_s \Delta\Gamma_s/(2\overline{\Gamma}_s)$. By convention, the phase of A_0 is set to zero and the phases of the other two amplitudes are denoted by $\delta_{||}$ and $\delta_\perp$.

For a given event, the decay rate is the sum of the functions P_B and $P_{\bar{B}}$ weighted by the flavor tagging dilution factors $(1 + \mathcal{D})/2$ and $(1 - \mathcal{D})/2$, respectively.

The contribution from the decay to $J/\psi K^+K^-$ with the kaons in an $\mathcal{S}$ wave is expressed in terms of the $\mathcal{S}$-wave fraction F_S and a phase δ_s. The squared sum of the $\mathcal{P}$ and $\mathcal{S}$ waves is integrated over the KK mass. For the $\mathcal{P}$ wave, we assume the non-relativistic Breit-Wigner model

$$g(M(KK)) = \sqrt{\frac{\Gamma_\phi/2}{\Delta M(KK)}} \cdot \frac{1}{M(KK) - M_\phi + i\Gamma_\phi/2} \tag{6.12}$$

with the ϕ meson mass $M_\phi = 1.019$ GeV and width $\Gamma_\phi = 4.26$ MeV [1], and with $\Delta M(KK) = 1.03 - 1.01 = 0.02$ GeV.

For the $\mathcal{S}$-wave component, we assume a uniform distribution in the range $1.01 < M(KK) < 1.03$ GeV. In the case of the BDT selection, it is modified by a KK-mass dependent factor corresponding to the BDT selection efficiency. We constrain the oscillation frequency to $\Delta M_s = 17.77 \pm 0.12$ ps^{-1}, as measured in Ref. [94]. Table 6.2 lists all physics parameters used in the fit.

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

$$\beta_s \to \pi/2 - \beta_s, \tag{6.13a}$$
$$\Delta\Gamma_s \to -\Delta\Gamma_s, \tag{6.13b}$$
$$\delta_\parallel \to 2\pi - \delta_\parallel, \tag{6.13c}$$
$$\delta_\perp \to \pi - \delta_\perp, \tag{6.13d}$$
$$\delta_s \to \pi - \delta_s. \tag{6.13e}$$

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 simulation, as described in Appendix B.

The signal time resolution in the MC simulation is modeled by a superposition of five Gaussian functions. This function is used in the analysis. The background-subtracted signal distribution agrees well with the MC model, as seen in Fig. 6.6. Two random variations of the function, also shown in the figure, are used in alternative fits, to estimate the systematic effect due to time resolution.

Parameter	Definition
$\lvert A_0\rvert^2$	$\mathcal{P}$-wave longitudinal amplitude squared, at $t = 0$
$A1$	$\lvert A_\parallel\rvert^2/(1 - \lvert A_0\rvert^2)$
$\overline{\tau}_s$ (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 6.2: Definition of nine real measurables for the decay $B_s^0 \to J/\psi\phi$ used in the Maximum Likelihood fitting.

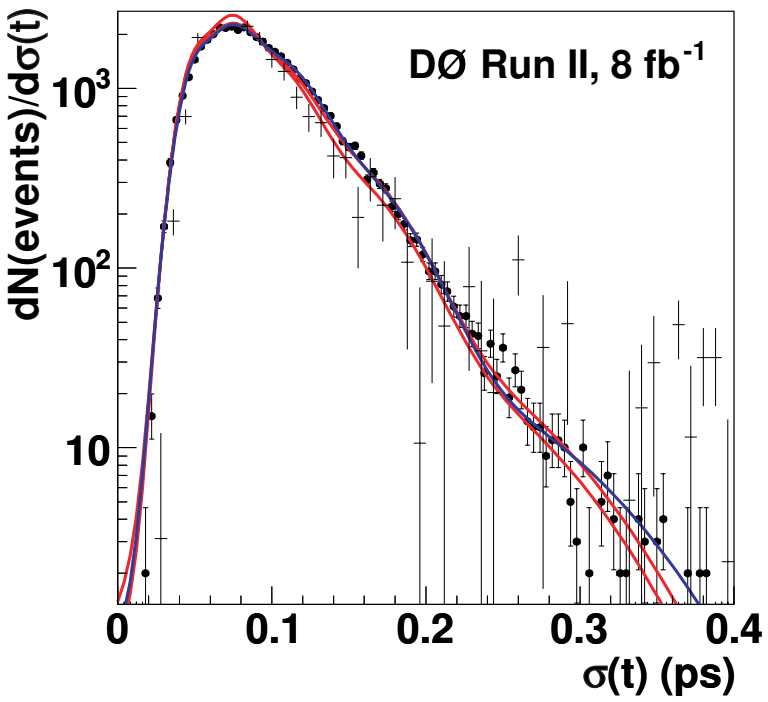


Figure 6.6: (color online). The distribution of the time resolution for the signal, MC (squares) and background-subtracted data (crosses). The blue curve is the sum of five Gaussian functions fitted to the MC distribution. The two red lines are variations of the default function used in the studies of systematic effects.

6.5.2 Background model

The proper decay time distribution of the background is described by a sum of a prompt component, modeled as a Gaussian function centered at zero, 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_{lm}^{k} and c_{lm}^{k}, 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 17×2 parameters 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_{1-1}^{0}, c_{20}^{0}, c_{22}^{0}, c_{00}^{2}, and c_{22}^{2}, and four for prompt background: c_{1-1}^{0}, c_{20}^{0}, c_{22}^{0}, and c_{2-1}^{2}. 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 6.2, 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.

6.5.3 Fit results

The fit results for the BDT sample and for the Square-cuts sample are shown in Table 6.3 and Table 6.4. 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 6.7 – 6.10 illustrate the quality of the fit for the background, for all data, and for the signal-enhanced sub-samples.

An independent measurement of the S-wave fraction is described in Appendix C and the result is in agreement with F_S determined from the maximum likelihood fit.

6.5.4 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 $B_d^0 \to \mu\nu D^{(*)\pm}$. 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

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 6.3: Maximum likelihood fit results for the BDT selection for the nominal fit (Default), for two alternative time resolution functions, $\sigma_A(t)$ and $\sigma_B(t)$, 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. The uncertainties are statistical.

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

Table 6.4: Maximum likelihood fit results for the 'Square-cuts' sample.

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 Appendix B) 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. (6.12) 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 with a width artificially increased by a factor of two. An MC-based estimate of the smearing 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. (6.12). We repeat the fits using this altered $\phi(1020)$ propagator as a measure of the sensitivity to the $M(KK)$ resolution.

Tables 6.3 and 6.4 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 C.L. ranges for all the measured physics quantities, we conduct MCMC studies described in the next section.

6.6 Confidence 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.

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 proper integrating 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.

6.6.1 The method

The MCMC technique uses the Metropolis-Hastings algorithm [95] to generate a random sample proportional 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 maximum likelihood fit result, we start from the best-fit point x. We randomly generate a point x' according to the multivariate normal distribution

$$PDF_{MC}(x') = e^{(-(x'-x)\cdot\Sigma\cdot(x'-x)/2)}, \tag{6.14}$$

where Σ is the covariance matrix. 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.

6.6.2 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 6.3 and 6.4.

Figures 6.11 and 6.12 illustrate the dependence of $\phi_s^{J/\psi\phi}$ on other physics parameters, in particular on $\cos\delta_\perp$ and $\cos\delta_s$. 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 (see also Fig. D.3 in Appendix D).

b) A correlation between $|\phi_s^{J/\psi\phi}|$ and $\overline{\tau}_s$, with the highest $\overline{\tau}_s$ occuring 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 $+0.8$. 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 $B_d^0 \to J/\psi K^*$ [2], up to the sign ambiguities. Ref. [96] 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 $\cos\delta_\perp < 0$.

To obtain the C.L. ranges 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.

6.6.3 Results

Figure 6.13 shows 68%, 90% and 95% C.L. contours 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 C.L. interval for each parameter, while the most probable value defines the central value.

The correlation between the two phases, $\delta_\perp$ and δ_s, prevents us from making separate point estimates. For the BDT selection, the measured $\mathcal{S}$-wave fraction F_S(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 results for the BDT-based sample, listed in Table 6.5: The one-dimensional estimates of physics parameters for the Square-cuts sample are in Table 6.6

To obtain the final C.L. ranges 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 6.14 shows 68%, 90% and 95% C.L. contours in the ($\phi_s^{J/\psi\phi}$,$\Delta\Gamma_s$) plane. The p-value for the SM point [62] ($\phi_s^{J/\psi\phi}, \Delta\Gamma_s$) =

Parameter	Value (±Stat. Error)
$\overline{\tau}_s$	$1.426^{+0.035}_{-0.032}$ ps
$\Delta\Gamma_s$	$0.129^{+0.076}_{-0.053}$ ps^{-1}
$\phi_s^{J/\psi\phi}$	$-0.49^{+0.48}_{-0.40}$
$\lvert A_0\rvert^2$	$0.552^{+0.016}_{-0.017}$
$\lvert A_\parallel\rvert^2$	$0.219^{+0.020}_{-0.021}$
$\delta_\parallel$	3.15 ± 0.27
$\cos(\delta_\perp - \delta_s)$	-0.06 ± 0.24
F_S(eff)	0.146 ± 0.035

Table 6.5: BDT fit results.

Parameter	Value (±Stat. Error)
$\overline{\tau}_s$	$1.444^{+0.041}_{-0.033}$ ps
$\Delta\Gamma_s$	$0.179^{+0.059}_{-0.060}$ ps^{-1}
$\phi_s^{J/\psi\phi}$	$-0.56^{+0.36}_{-0.32}$
$\lvert A_0\rvert^2$	0.565 ± 0.017
$\lvert A_\parallel\rvert^2$	$0.249^{+0.021}_{-0.022}$
$\delta_\parallel$	3.15 ± 0.19
$\cos(\delta_\perp - \delta_s)$	$-0.20^{+0.26}_{-0.27}$
F_S	0.173 ± 0.036

Table 6.6: Square-cuts fit results.

$(-0.038, 0.087$ ps$^{-1})$ is 29.8%. The one-dimensional 68% C.L. ranges are listed in Section 6.7 below.

6.7 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% C.L. intervals, including systematic uncertainties, with the oscillation frequency constrained to $\Delta M_s = 17.77 \pm 0.12$ ps^{-1}, are shown in Table 6.7

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

In the previous publication [88], which was based on a subset of this data sample, we 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. The result of this study (Appendix C) is consistent with the result of the

Maximum Likelihood fit shown above.

It is a unique feature of the decay $B_s^0 \to J/\psi\phi$ that thanks to the sizeable lifetime difference between the two mass eigenstates, there is a sensitivity to $\phi_s^{J/\psi\phi}$ even in the absence of the flavor tagging information. The interference terms $A_\parallel - A_\perp$ and $A_0 - A_\perp$ are proportional to $(e^{-\Gamma_H t} - e^{-\Gamma_L t}) \sin\phi_s^{J/\psi\phi}$. Also, if $\cos\phi_s^{J/\psi\phi}$ is significantly different from unity, the decay rates of the *CP*-even and *CP*-odd components have two slopes each. We confirm (see Appendix D) that the independent inputs to the measurement of $\phi_s^{J/\psi\phi}$ (the oscillatory behavior, relying on the flavor tagging information, and the exponential time evolution of the angular distribution) give consistent results.

$\overline{\tau}_s$	$1.443^{+0.038}_{-0.035}$ ps
$\Delta\Gamma_s$	$0.163^{+0.065}_{-0.064}$ ps^{-1}
$\phi_s^{J/\psi\phi}$	$-0.55^{+0.38}_{-0.36}$
$\lvert A_0\rvert^2$	$0.558^{+0.017}_{-0.019}$
$\lvert A_\parallel\rvert^2$	$0.231^{+0.024}_{-0.030}$
$\delta_\parallel$	3.15 ± 0.22
$\cos(\delta_\perp - \delta_s)$	$-0.11^{+0.27}_{-0.25}$
F_S	0.173 ± 0.036

Table 6.7: Measured values and the 68% C.L. intervals, including systematic uncertainties

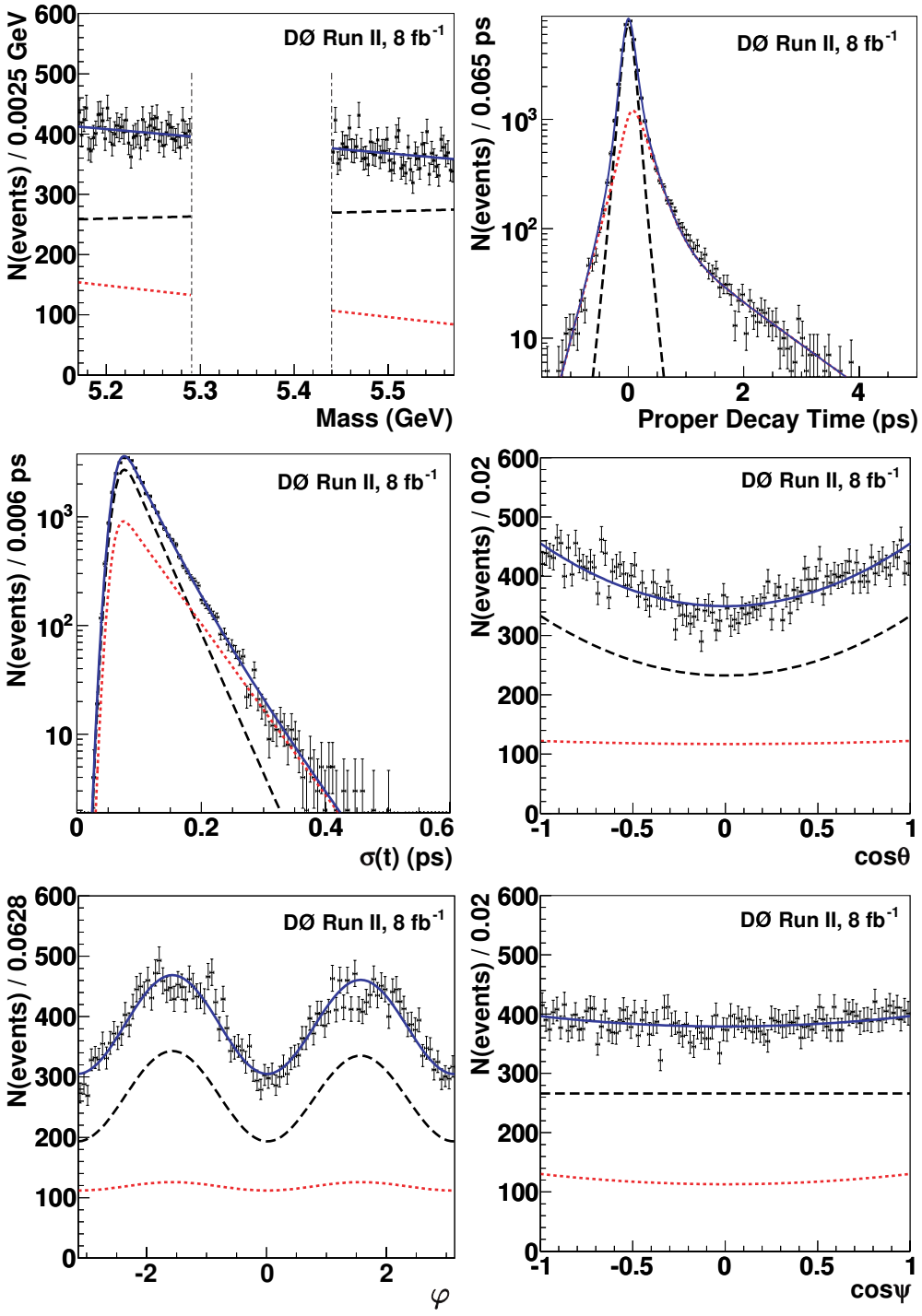


Figure 6.7: (color online). The distributions in the background (B_s^0 mass sidebands) region of candidate mass, proper decay time, decay time uncertainty, transversity polar and azimuthal angles, and $\cos\psi$ for the BDT sample. The curves show the prompt (black dashed) and non-prompt (red dotted) components, and their sum (blue solid).

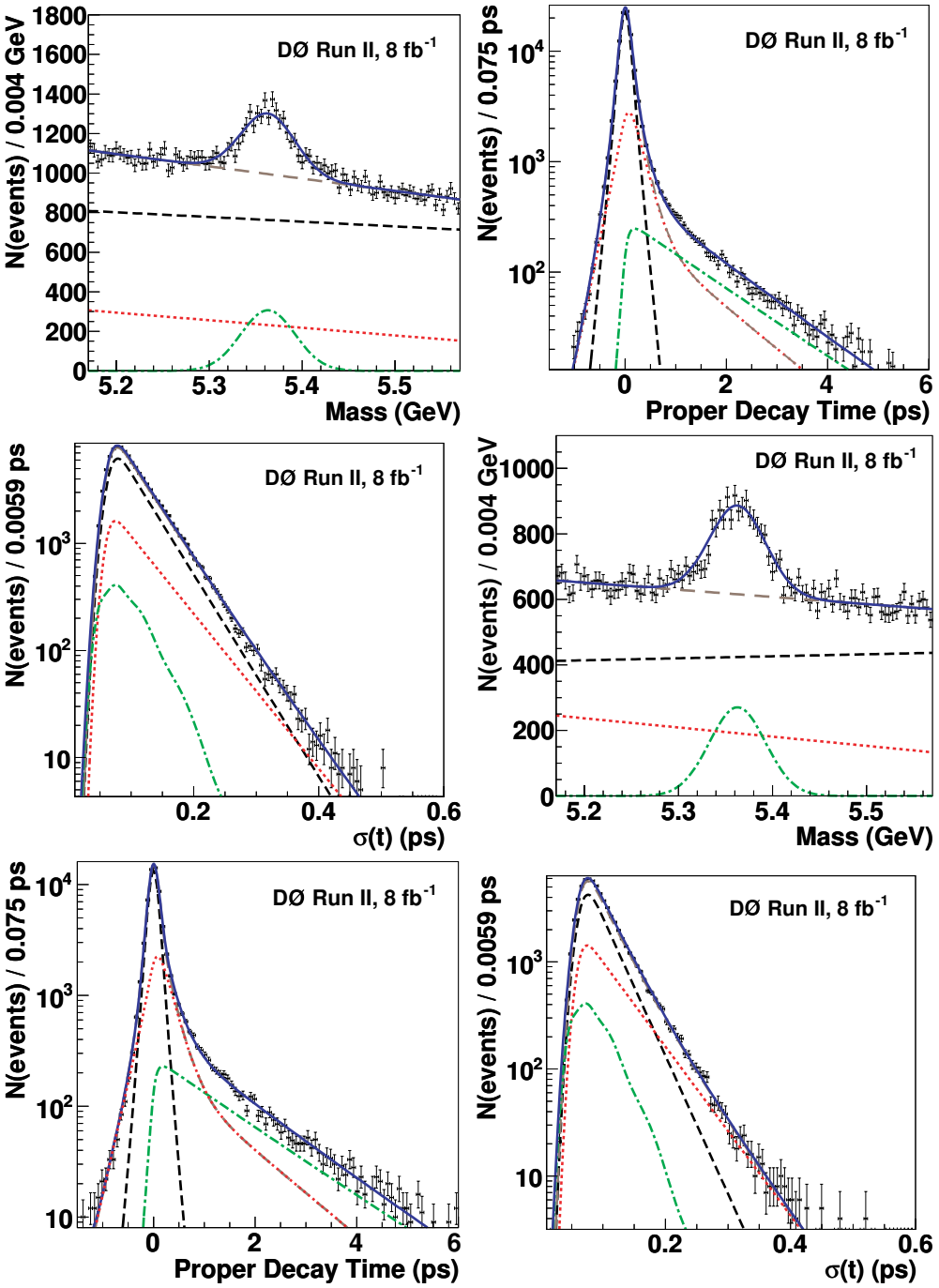


Figure 6.8: (color online). Invariant mass, proper decay time, and proper decay time uncertainty distributions for B_s^0 candidates in the (top) BDT sample and (bottom) Square-cuts sample. The curves are projections of the maximum likelihood fit. Shown are the signal (green dashed-dotted curve), prompt 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).

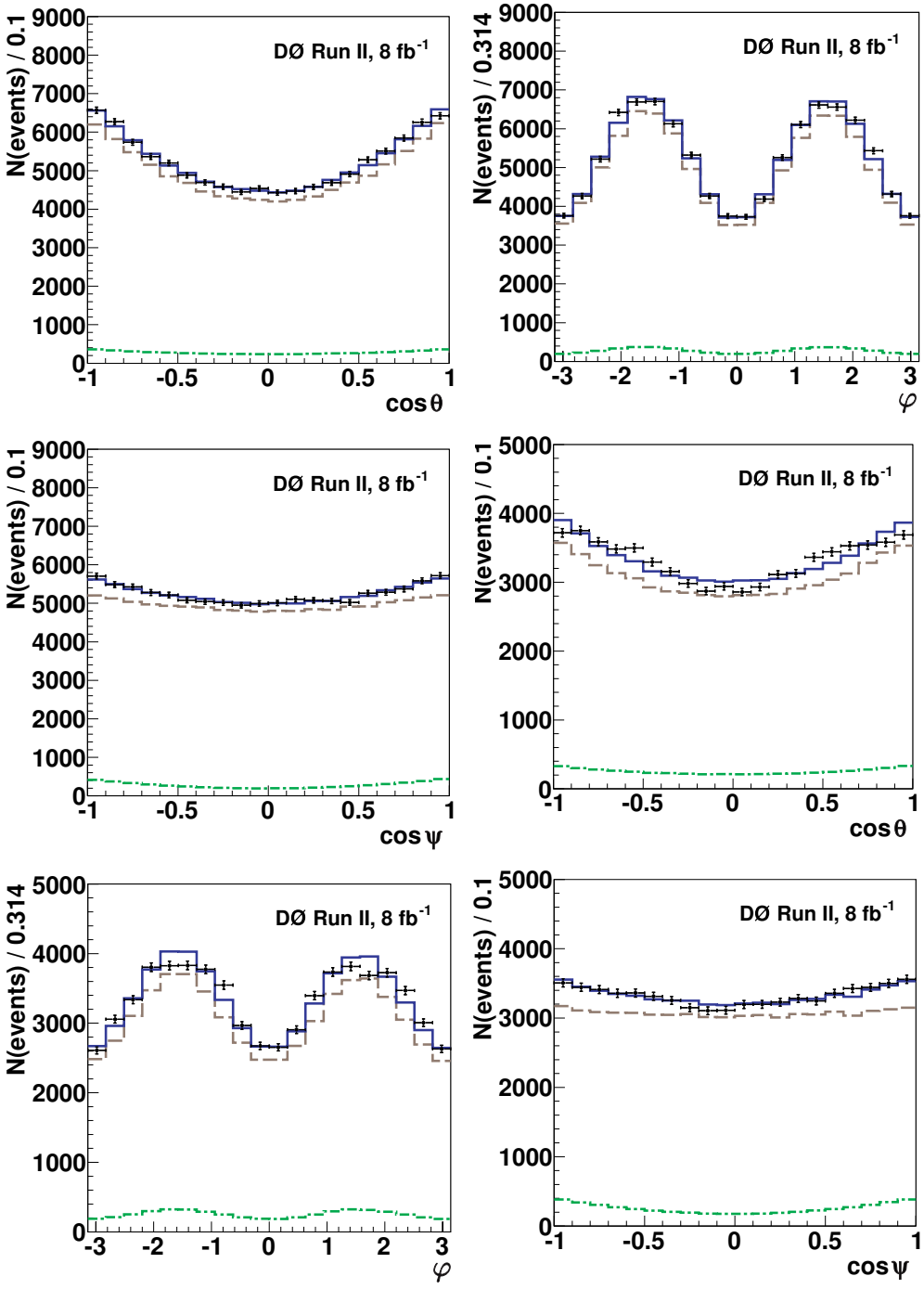


Figure 6.9: (color online). Distributions of transversity polar and azimuthal angles and $\cos\psi$ for B_s^0 candidates in the BDT sample (top) and Square-cuts sample (bottom). The curves are projections of the maximum likelihood fit. Shown are the signal (green dashed-dotted curve), total background (brown long-dashed curve) and the sum of signal and total background (blue solid curve).

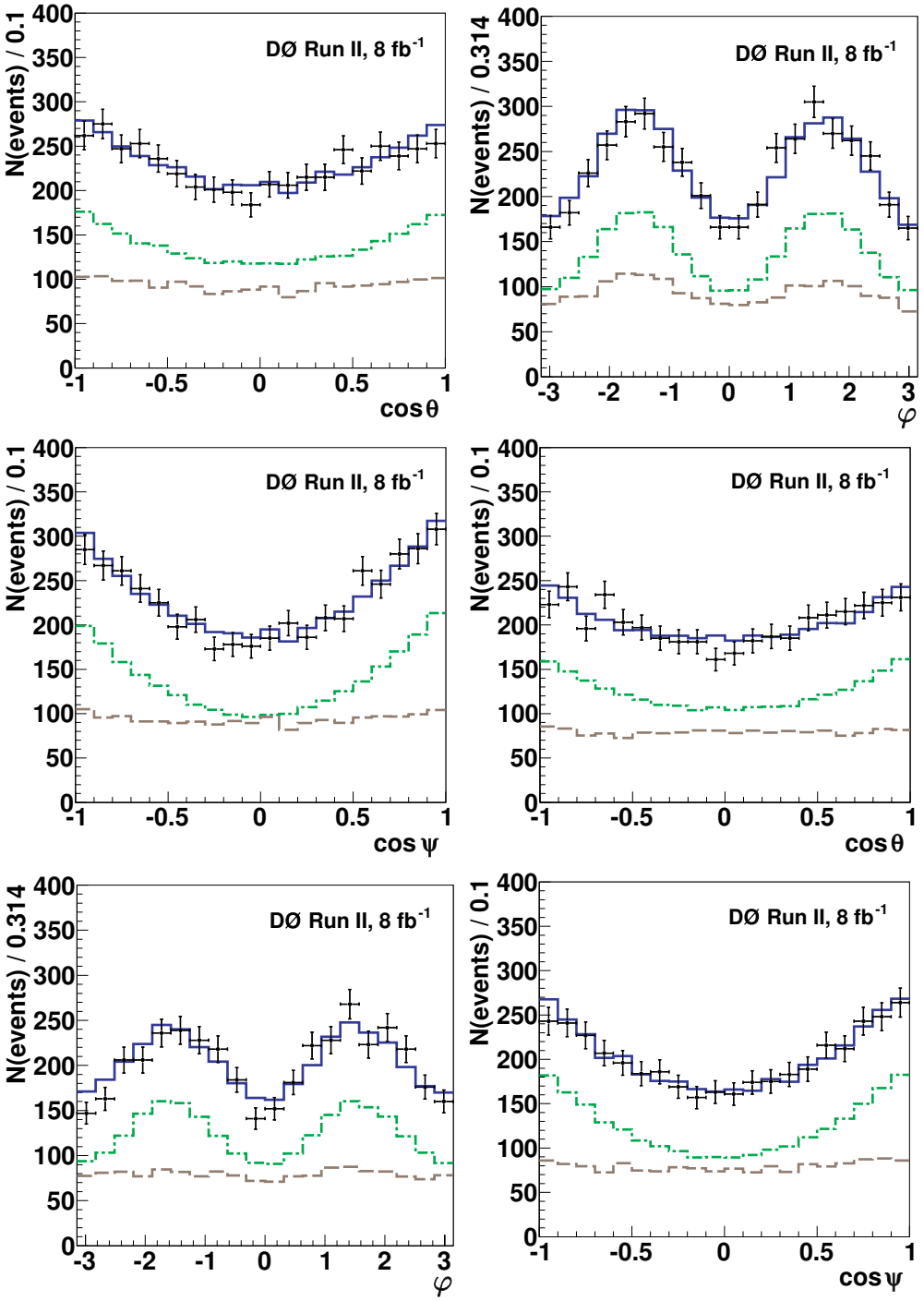


Figure 6.10: (color online). Distributions of transversity polar and azimuthal angles and $\cos\psi$ for B_s^0 candidates in the BDT sample (top) and Square-cuts sample (bottom). in the signal mass region ($5.31 < M(B_s) < 5.43$ GeV) and with an additional signal-enhancing requirement $t > 1.0$ ps. The 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).

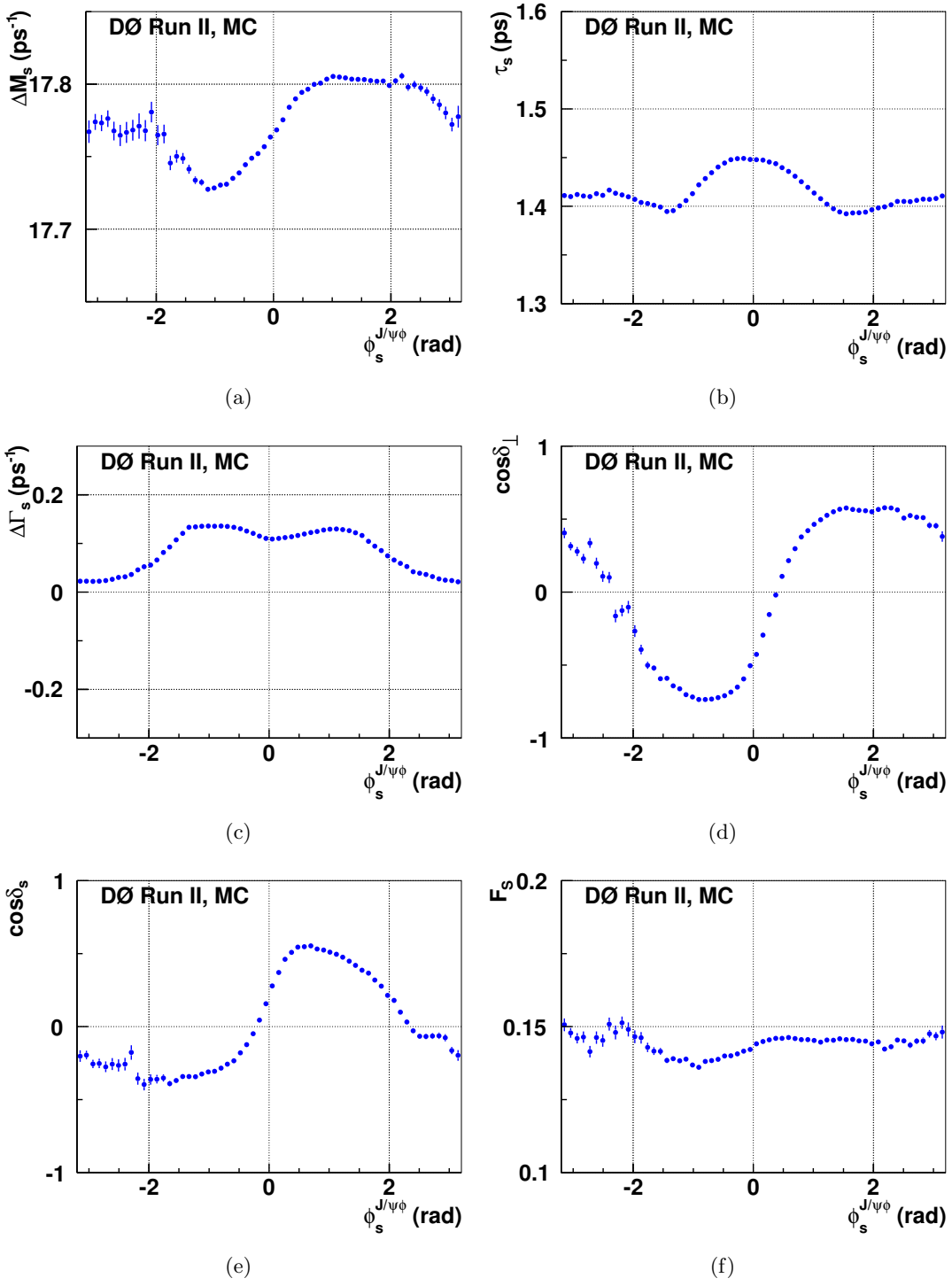


Figure 6.11: Profiles of ΔM_s, $\overline{\tau}_s$, $\Delta\Gamma_s$, $\cos\delta_\perp$, $\cos\delta_s$, and F_S, for $\Delta\Gamma_s > 0$, versus $\phi_s^{J/\psi\phi}$ from the MCMC simulation for the BDT selection data sample.

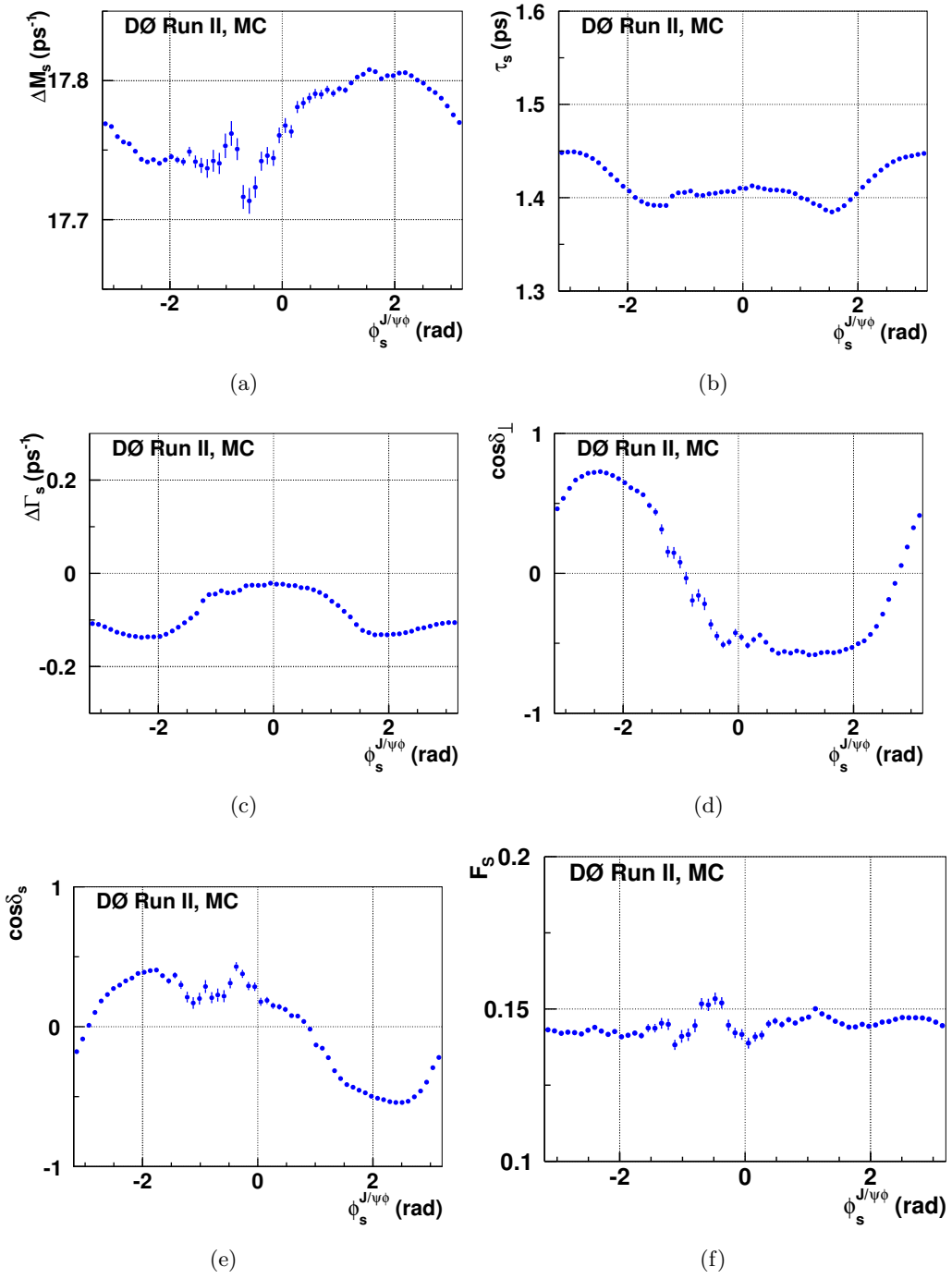
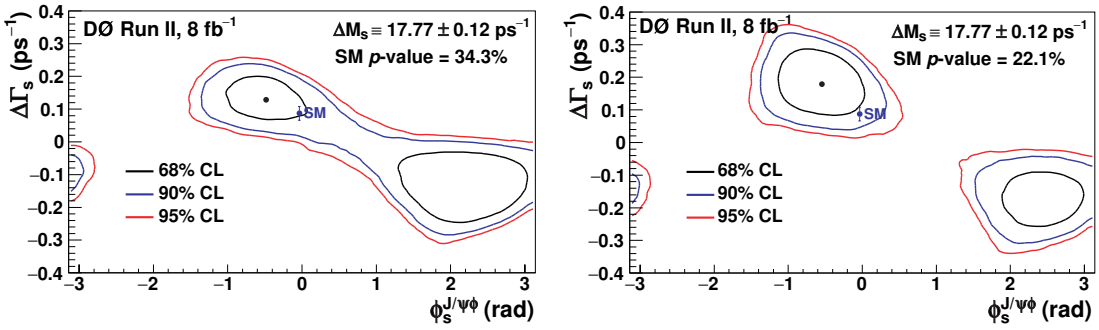


Figure 6.12: Profiles of ΔM_s, $\overline{\tau}_s$, $\Delta\Gamma_s$, $\cos\delta_\perp$, $\cos\delta_s$, and F_S, for $\Delta\Gamma_s < 0$, versus $\phi_s^{J/\psi\phi}$ from the MCMC simulation for the BDT selection data sample.



(a) BDT selection (b) Square-cuts selection

Figure 6.13: (color online). Two-dimensional 68%, 90% and and 95% C.L. contours for (a) the BDT selection and (b) the Square-cuts sample. The standard model expectation is indicated as a point with an error.

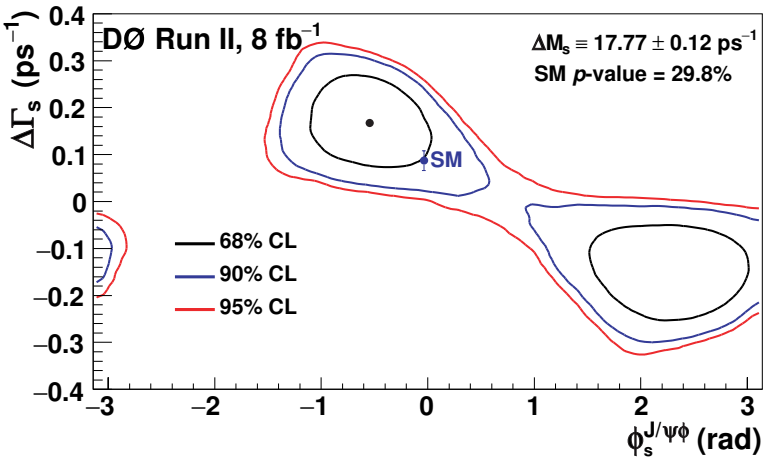


Figure 6.14: (color online). Two-dimensional 68%, 90% and 95% C.L. contours including systematic uncertainties. The standard model expectation is indicated as a point with an error.

6. CP VIOLATION PHASE IN THE DECAY $B_S^0 \to J/\Psi$

7. Conclusions

7.0.1 B_s^0 Lifetime

Using 8.1 fb^{-1} of data collected with the DØ detector we measured the B_s^0 lifetime in the inclusive semileptonic channel $B_s^0 \to D_s^- \mu^+ + X$. We found:

$$c\tau_{B_s \to SL} = 423.0 \pm 2.9 \ \mu\text{m (stat.)}^{+10.0}_{-5.6} \ \mu\text{m (syst.)}, \tag{7.1}$$

$$\tau_{B_s \to SL} = 1.41 \pm 0.01 \text{ ps (stat.)}^{+0.03}_{-0.02} \text{ ps (syst.)}, \tag{7.2}$$

that is consistent with the world average, as is shown in figure 7.1. With respect to the DØ previous measurement, we reduced the statistical uncertainty by a factor of 3. But the systematic uncertainty was not reduced. The uncertainty in this measurement is dominated by the systematic effects. Taking the world average of B^0 lifetime and this measurement, one can compute the ratio $\tau_{B_s^0}/\tau_{B^0}$,

$$\frac{\tau_{B_s^0}}{\tau_{B^0}} = 0.930^{+0.022}_{-0.014}. \tag{7.3}$$

If we take our measurement of the B^0 lifetime, the ratio becomes

$$\frac{\tau_{B_s^0}}{\tau_{B^0}} = 0.922^{+0.032}_{-0.026}. \tag{7.4}$$

This result is consistent, but with some tension, with the prediction of inequality 2.85.

7.1 CP Violation in $B_s^0 - \overline{B}_s^0$ meson

In summary, we report a new messurement of the B_s^0 mixing parameters and average lifetime in the decay process $B_s^0 \to J/\psi\phi$. In the fits, we constrain the oscillation frequency to $\Delta M_s = 17.77 \pm 0.12$ ps^{-1}, as measured in Ref. [94]. We also measure the amplitudes and phases of the polarization amplitudes. For the first time we allow for an admixture of the KK S wave. Our results are:

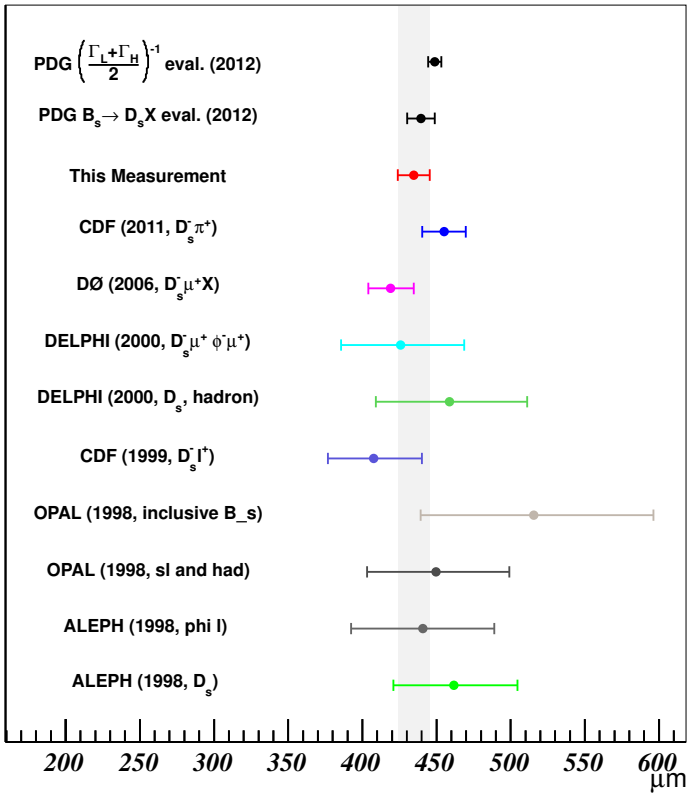


Figure 7.1: B_s lifetime (times c) measurements [2].

$$
\begin{aligned}
\overline{\tau}_s &= 1.44 \pm 0.04 \text{ ps} \\
\Delta\Gamma_s &= 0.16 \pm 0.07 \text{ ps}^{-1} \\
\phi_s^{J/\psi\phi} &= -0.57 \pm 0.43 \\
|A_0|^2 &= 0.56 \pm 0.02 \\
|A_\parallel|^2 &= 0.24 \pm 0.03 \\
F_S &= 0.175 \pm 0.035 \\
\delta_\parallel &= 3.14 \pm 0.21 \\
\cos(\delta_\perp - \delta_s) &= -0.14 \pm 0.28
\end{aligned}
$$

The p-value for the SM point $(\phi_s^{J/\psi\phi}, \Delta\Gamma_s) = (-0.038, 0.087 \text{ ps}^{-1})$ is 30.0%. The

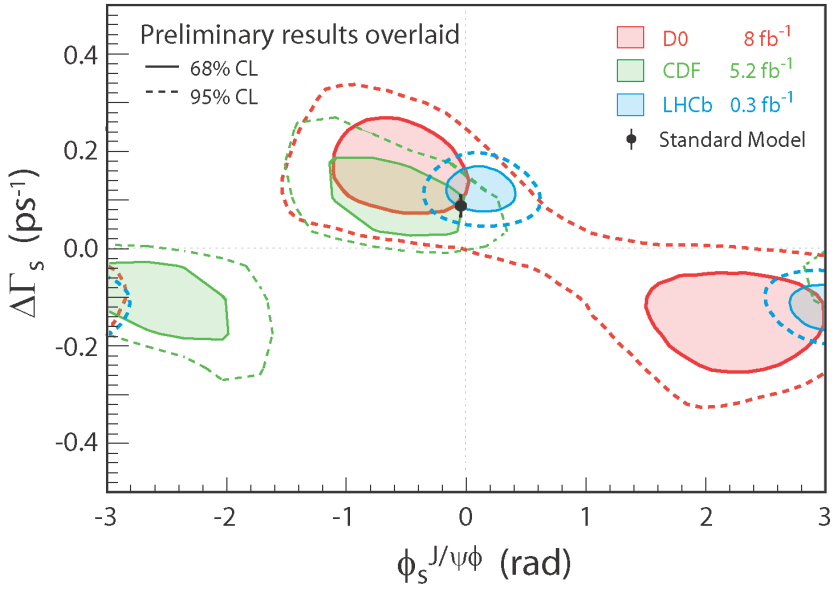


Figure 7.2: Different experimental measuremets of phi_s and $\Delta\Gamma$.

results supersede our previous measurements [88] that were based on a smaller data sample.

7.2 General Conclusions

The measurements of the B_s^0 life time and CP violation parameters are consistent with the SM model. For the CP violation parameters they are also in agreement with the results from other experiments, as shown in Figure 7.2, and it seems that a big CP violation effect will not be detected in oscillations of the $B_s^0 - \overline{B}_s^0$ system. However, for the lifetime measurement, even when there is still in agreement with the HQET, there is some tension that would indicate that some review to the theory has to be done, or the experimental method should be further reviewed.

A. 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 A.1. The inclusive BDT uses 35 variables, listed in Table A.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

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

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 Fig. A.1. The distributions for the six most important variables in the training on inclusive $B \to J/\psi X$ decays are shown in Fig. A.2.

Figure A.3 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.

A. BDT DISCRIMINANTS

Rank	Variable	Importance	Separation
1	KK invariant mass	0.3655	0.3540
2	Maximum ΔR between either K meson and the B_s^0 candidate	0.1346	0.4863
3	Isolation using the maximum ΔR between either K and the B_s^0	0.0390	0.1784
4	Uncorrected p_T of the B_s^0	0.0346	0.3626
5	Minimum ΔR between either K and the B_s^0	0.0335	0.4278
6	p_T of the trailing K meson	0.0331	0.4854
7	p_T of the ϕ meson	0.0314	0.4998
8	p_T of the leading K meson	0.0283	0.4884
9	Trailing muon momentum	0.0252	0.0809
10	p_T of the leading muon	0.0240	0.1601
11	Maximum ΔR between either muon and the B_s^0	0.0223	0.1109
12	Maximum χ^2 of either K meson with the J/ψ vertex	0.0217	0.0162
13	Dimuon invariant mass	0.0215	0.0145
14	Maximum χ^2 of either of the K candidate track	0.0213	0.021
15	B_s^0 isolation using the larger K/B_s ΔR and tracks from the PV	0.0207	0.1739
16	p_T of the J/ψ meson	0.0205	0.1809
17	Minimum ΔR between either muon and the B_s^0 candidate	0.0188	0.1023
18	Trailing K momentum	0.0105	0.3159
19	χ^2 of the B_s^0 candidate vertex	0.0093	0.0119
20	B_s^0 isolation using $\Delta R < 0.75$	0.0084	0.0241
21	Minimum χ^2 of the J/ψ vertex with either K	0.0081	0.0069
22	p_T of the trailing muon	0.0079	0.0922
23	Minimum of the χ^2 of the J/ψ and ϕ vertices	0.0073	0.0057
24	Isolation using $\Delta R < 0.5$	0.0070	0.0405
25	Uncorrected B_s^0 total momentum	0.0068	0.2103
26	Minimum χ^2 of either K track fit	0.0065	0.0266
27	Isolation using $\Delta R < 0.5$ and particles from the PV	0.0057	0.0401
28	Leading K meson momentum	0.0051	0.3217
29	Leading muon momentum	0.0048	0.0908
30	ϕ meson momentum	0.0048	0.3233
31	Maximum χ^2 of the J/ψ or ϕ vertices	0.0044	0.0061
32	Isolation using $\Delta R < 0.75$ and particles from the PV	0.0037	0.0259
33	J/ψ meson momentum	0.0037	0.1004

Table A.1: Variables used to train the prompt BDT, ranked by their importance in the training.

Rank	Variable	Importance	Separation
1	KK invariant mass	0.2863	0.3603
2	B_s^0 isolation using the larger K/B_s ΔR and tracks from the PV	0.1742	0.4511
3	Minimum dE/dx of either K	0.0778	0.1076
4	χ^2 of B_s^0	0.0757	0.2123
5	p_T of the ϕ meson	0.0559	0.4856
6	p_T of the leading K meson	0.0504	0.4745
7	Isolation using the maximum ΔR between either K and the B_s^0	0.0429	0.4468
8	p_T of the trailing K meson	0.0350	0.4774
9	Maximum χ^2 of either K meson with the J/ψ vertex	0.0260	0.2051
10	Isolation using $\Delta R < 0.5$ and particles from the PV	0.0229	0.1703
11	Isolation using $\Delta R < 0.75$ and tracks from the PV	0.0154	0.2238
12	Minimum χ^2 of of either K with the J/ψ vertex	0.0151	0.1308
13	Minimum ΔR between either K meson and the B_s^0 candidate	0.0115	0.3104
14	Dimuon invariant mass	0.0099	0.0190
15	Total momentum of the ϕ meson	0.0091	0.3307
16	p_T of the J/ψ meson	0.0089	0.1198
17	Trailing muon momentum	0.0082	0.0594
18	Isolation using $\Delta R < 0.5$	0.0073	0.1695
19	Maximum ΔR between either K meson and the B_s^0 candidate	0.0070	0.3794
20	Maximum dE/dx of either K meson	0.0069	0.0528
21	Trailing K meson momentum	0.0068	0.3253
22	J/ψ vertex χ^2	0.0063	0.0057
23	Leading K meson momentum	0.0058	0.3277
24	Maximum χ^2 of either K candidate track	0.0054	0.0267
25	Isolation using $\Delta R < 0.75$	0.0046	0.2203
26	Minimum ΔR between either muon and the B_s^0 candidate	0.0041	0.0729
27	Minimum χ^2 of either K candidate track	0.0039	0.0284
28	uncorrected p_T of B_s^0 candidate	0.0036	0.2485
29	p_T of the trailing muon	0.0029	0.0702
30	J/ψ momentum	0.0027	0.0645
31	Maximum ΔR between either muon and the B_s^0 candidate	0.0026	0.0872
32	Vertex χ^2 of the ϕ meson	0.0017	0.0098
33	Uncorrected B_s^0 momentum	0.0014	0.1675
34	p_T of the leading muon	0.0011	0.1008
35	Leading muon momentum	0.0009	0.0547

Table A.2: Variables used to train the non-prompt BDT, ranked by their importance in the training.

A. BDT DISCRIMINANTS

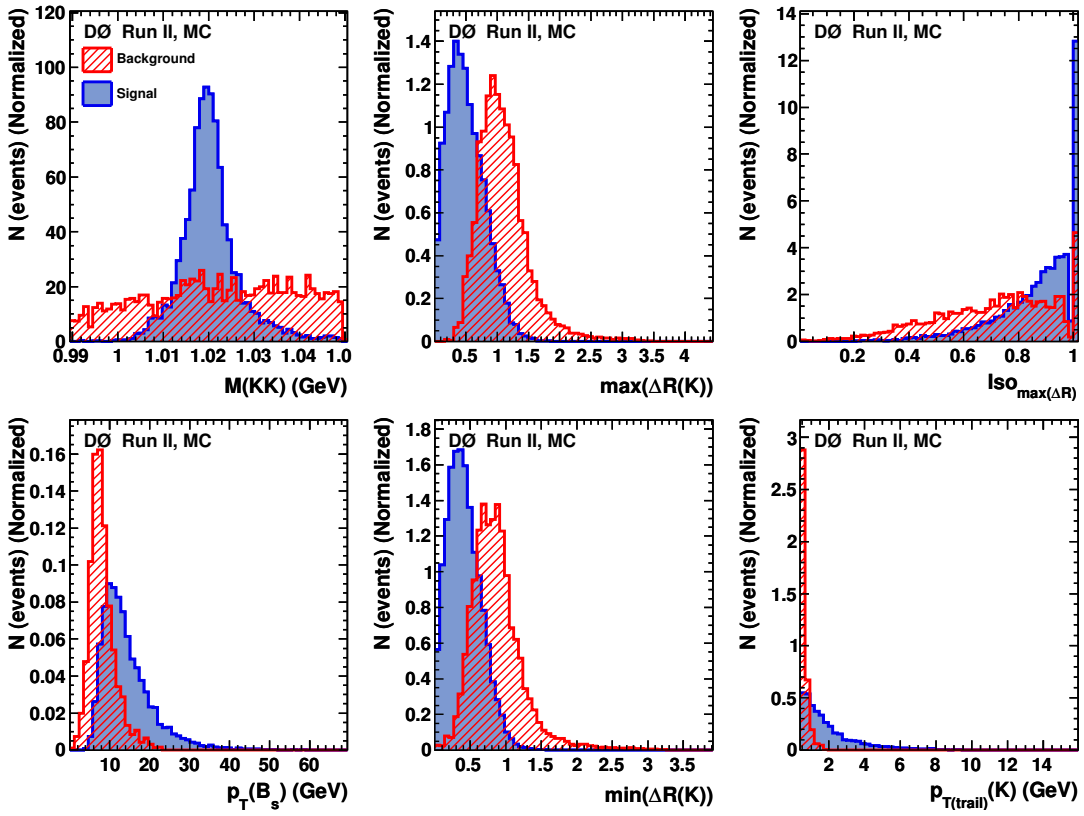


Figure A.1: (color online) The distributions of the six most important variables used in the BDT trained on prompt J/ψ production for the $B_s^0 \to J/\psi\phi$ signal (solid blue) and prompt J/ψ events (red dashed) histograms.

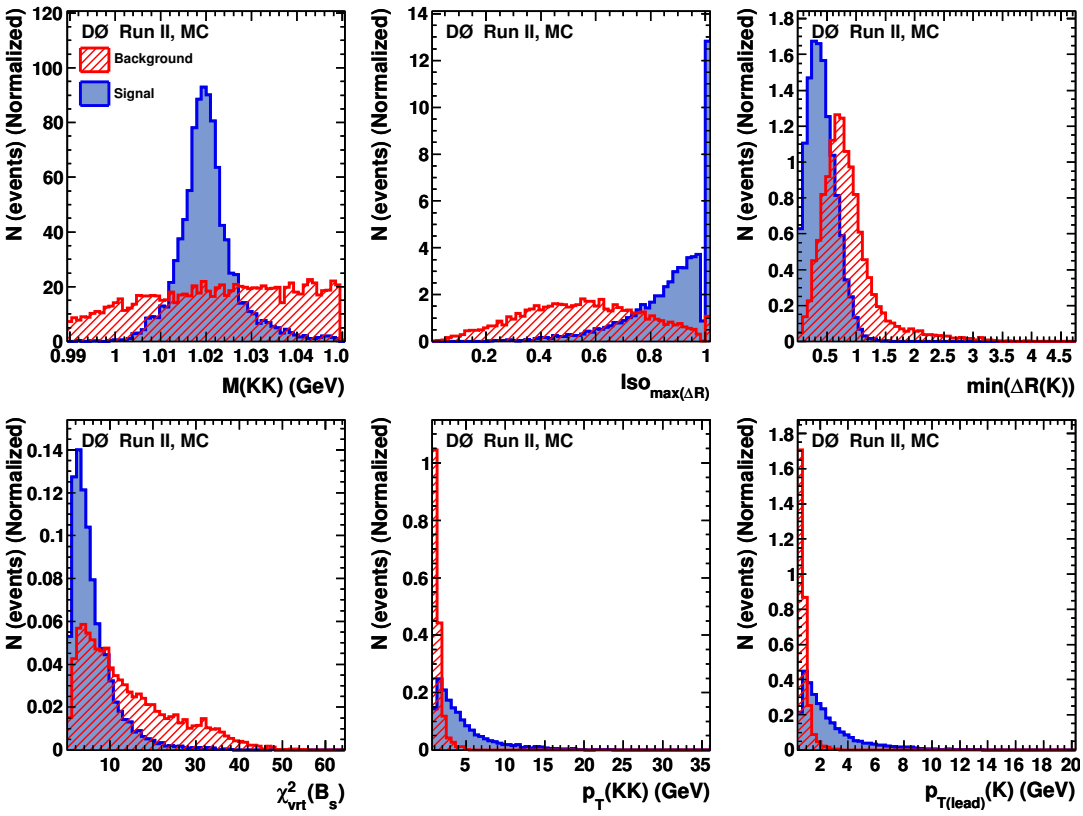
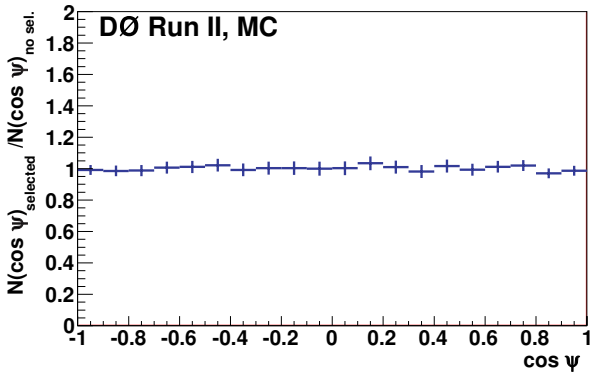
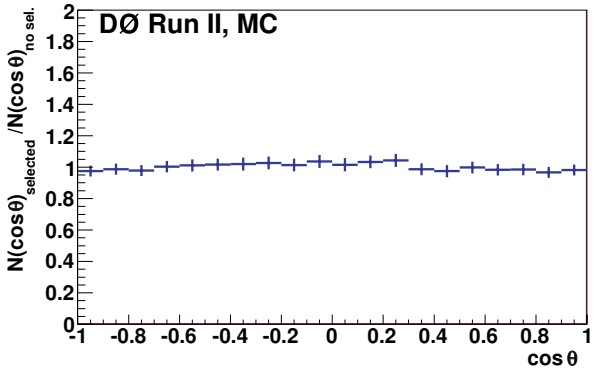


Figure A.2: (color online) The distributions of the six most important variables used in the BDT trained on inclusive $B \to J/\psi X$ decays for the $B_s^0 \to J/\psi\phi$ signal (solid blue) and inclusive $B \to J/\psi X$ decays (red dashed) histograms.

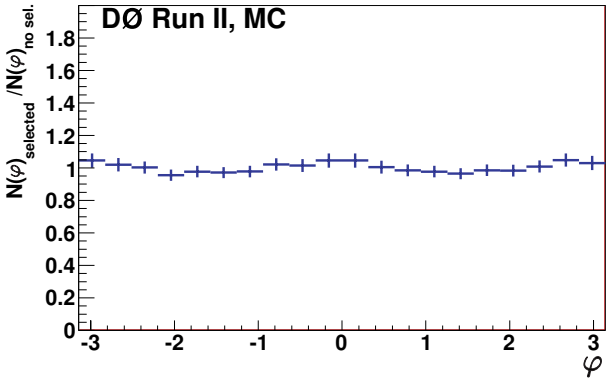
A. BDT DISCRIMINANTS



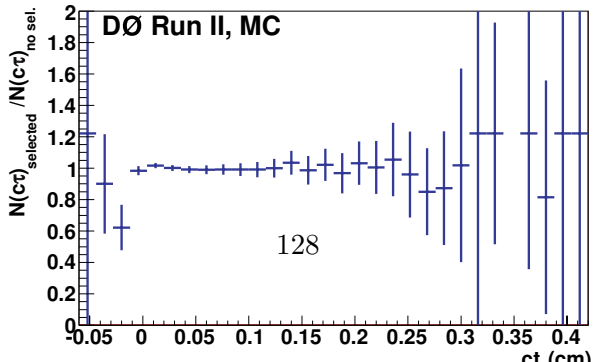
(a)



(b)



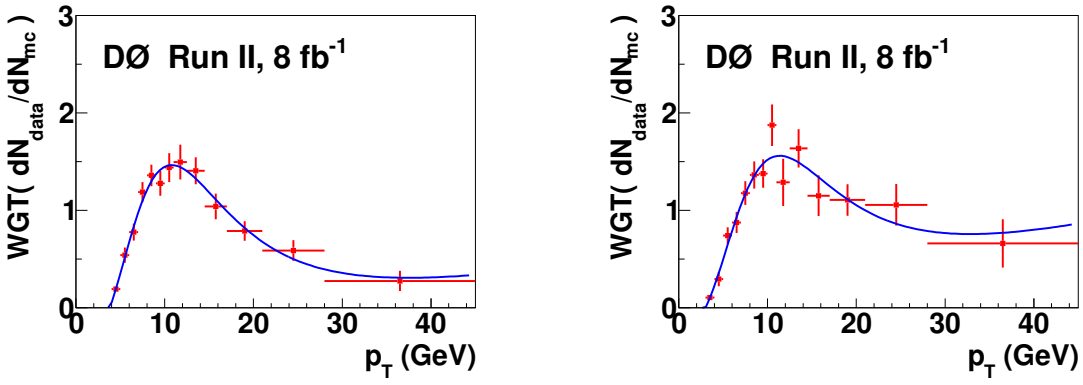
(c)



B. Detector acceptance

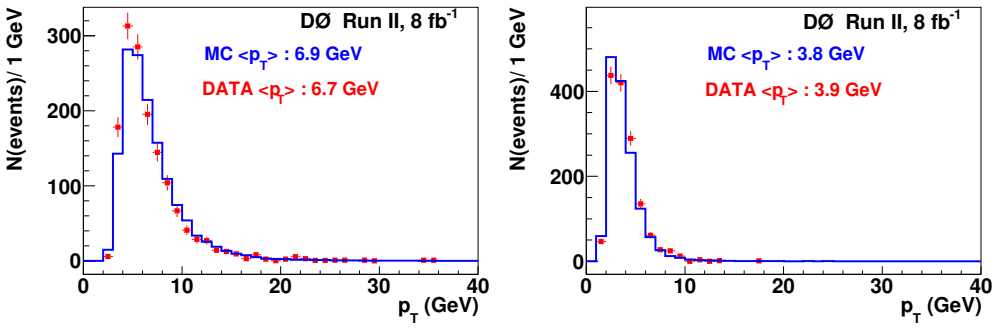
We take into account the shaping of the signal distribution by the detector acceptance and kinematic selection by introducing acceptance functions in the three angles of the transversity basis. The acceptance functions are derived from Monte Carlo simulation. Due to the event triggering effects, the 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 Fig. B.1.

Figure B.2 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 is shown in Fig. B.3. The acceptance in ψ is shown in Fig. B.4.



(a) (b)

Figure B.1: 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 B.2: Transverse momentum distributions of the four final-state particles in data (points) and weighted MC (solid histogram), for the BDT-based event selection.

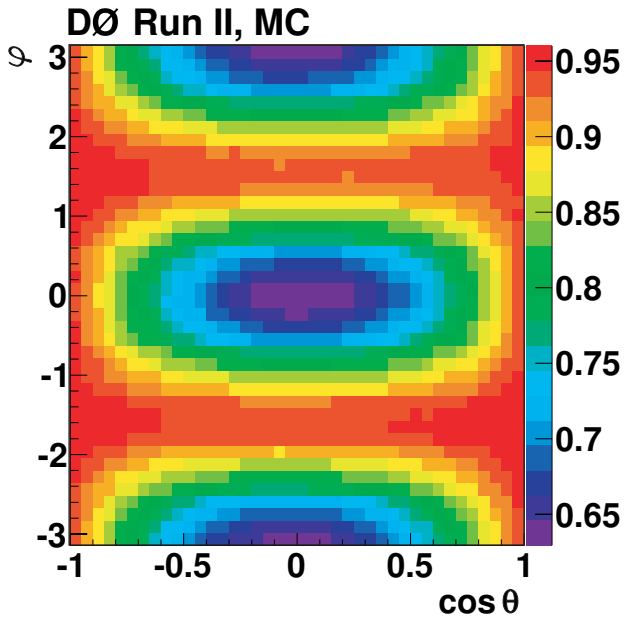


Figure B.3: Map of the detector acceptance on the plane φ – $\cos\theta$.

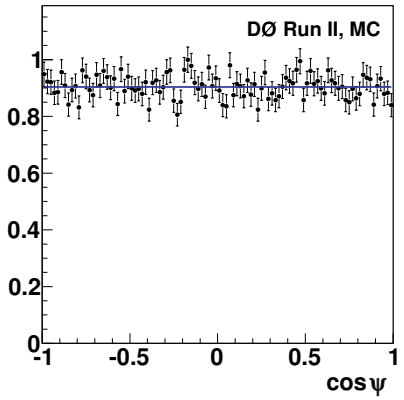


Figure B.4: Detector acceptance as a function of $\cos\psi$. The acceptance is uniform in $\cos\psi$.

C. Independent estimate of F_S

In the Maximum Likelihood fit, the information on 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. Also, by studying the dependence of the $\phi(1020)$ helicity angle distribution on its mass, we can resolve the sign ambiguity in the physics measurements.

For this study, we use the "Square-cuts" sample, for which the event selection is not biased in $M(K^+K^-)$.

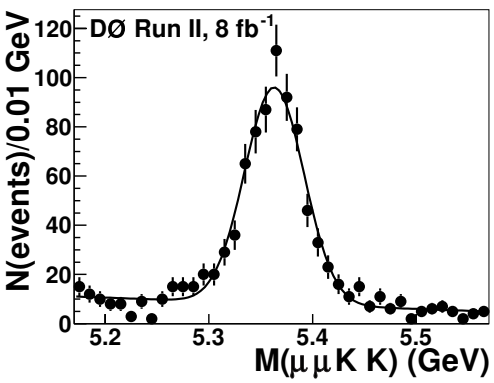
Using events with $ct > 0.02$ cm to suppress background, we divide the data into 2-MeV slices in $M(K^+K^-)$ and extract the B_s^0 signal in each slice by fitting the B_s^0 candidate mass distribution to a Gaussian signal and a linear background. For $M(K^+K^-)$ slices above 1.03 GeV, we fix the signal width to 0.03 GeV to suppress the effect of the $B_d^0 \to J/\psi K^*$ background. Our MC studies show that the $B_d^0 \to J/\psi K^*$ background peaks at about 5.40 GeV, with a width of about 0.06 GeV. Two examples of such fits are shown in Fig. C.1.

The resulting $M(K^+K^-)$ distribution of the pure B_s^0 signal is shown in Fig. C.2. The B_s^0 yield is extracted by performing a fit to the sum of a $\mathcal{P}$ wave and a constant term, representing a three-body decay $B_s^0 \to J/\psi K^+K^-$. (or a quasi-two-body decay involving a wide resonance, such as f_0). There is a clear non-$\phi(1020)$ component of the B_s^0 decay. The $M(K^+K^-)$ distribution from the MC simulation of the $B_s^0 \to J/\psi\phi$ decay is shown in Fig. C.3.

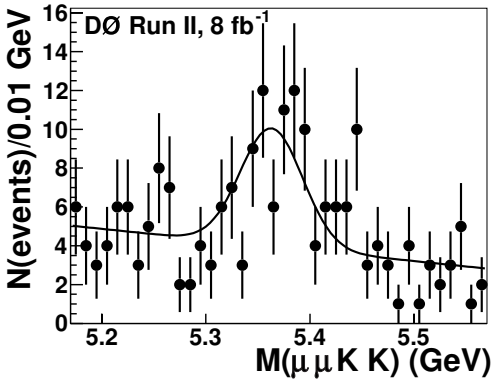
This constant term may be due to an $\mathcal{S}$ wave, or a non-resonant $\mathcal{P}$ wave, or a combination of both. If we assign it entirely to the $\mathcal{S}$ wave, we interpret its fraction as F_S. From the fits in Fig. C.1, the measured $\mathcal{S}$-wave fraction in the range $1.01 < M(K^+K^-) < 1.03$ GeV is $F_S = 0.14 \pm 0.02$.

In an effort to resolve the $\cos\delta_s$ sign ambiguity, in Fig. C.4 we plot the distribution of the ϕ helicity angle $\cos\psi$ for $M(K^+K^-) < 1.019$ GeV and for $M(K^+K^-) > 1.019$ GeV.

The $\mathcal{P}$ wave has a characteristic symmetric distribution in $\cos\psi$, as a sum of the A_0 term, proportional to $\cos^2\psi$, and $A_|$ and $A_\perp$, proportional to $\sin^2\psi$. The $\mathcal{S}$ wave has no dependence on ψ. Thus, in the presence of both $\mathcal{P}$ and $\mathcal{S}$ waves, in addition to the quadratic term, there is a linear term that changes sign from positive to negative, indicative of a $\mathcal{P}$ – $\mathcal{S}$ interference. Assuming a Breit-Wigner propagator for the $\mathcal{P}$ wave, whose real part changes sign from negative to positive around the $\phi(1020)$ meson mass, and a constant $\mathcal{S}$-wave amplitude, the asymmetry of the $\cos\psi$ distribution is proportional to $-\cos\delta_s$. There is a small asymmetry that is changing sign from positive to negative, indicating that the assignment $\cos\delta_s < 0$ is marginally favored by the data.



(a) $1.018 < M(K^+K^-) < 1.020$ GeV



(b) $1.048 < M(K^+K^-) < 1.050$ GeV

Figure C.1: The invariant mass distribution of B_s^0 candidates with $ct > 0.02$ cm in two slices of $M(K^+K^-)$. Fits to a sum of a Gaussian function and a polynomial are used to extract the B_s^0 yield in each slice.

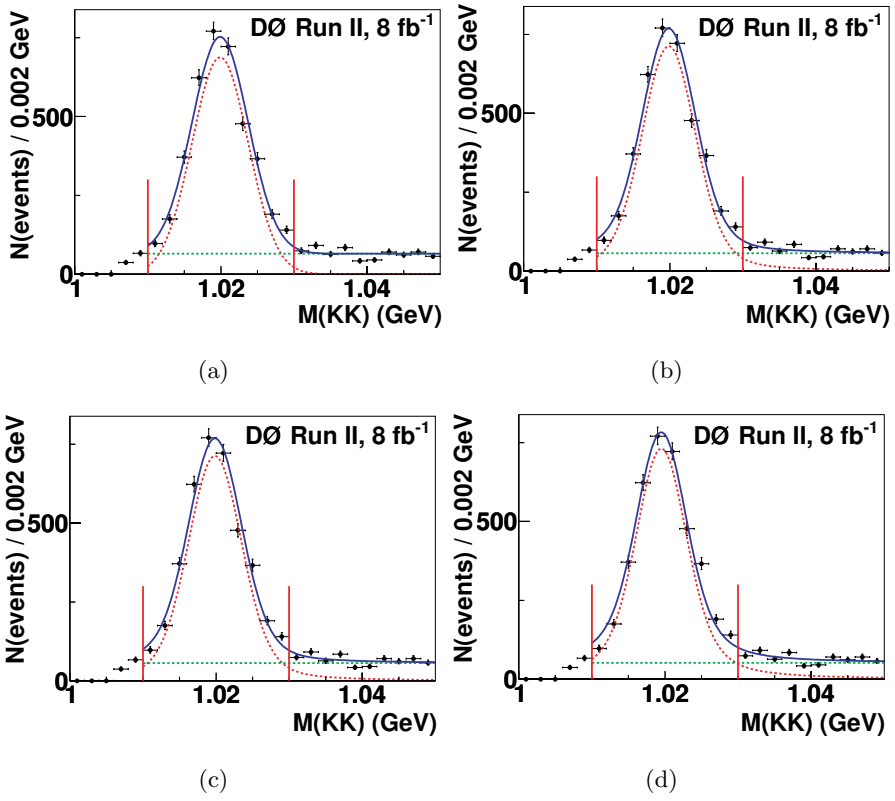


Figure C.2: The K^+K^- mass distribution of pure B_s^0 signal with four fits to a sum of the $\mathcal{P}$-wave $\phi(1020)$ meson decay and a constant term (presumed to be due to $\mathcal{S}$ wave), assuming different shapes of the $\mathcal{P}$ wave. The mass shapes, and the resulting $\mathcal{S}$-wave fraction in the 1.01 – 1.03 GeV window are: (a) a Gaussian function with an unconstrained width, $F_S = 0.17 \pm 0.01$; (b) a smeared Breit-Wigner function with free mass, width, and smearing, $F_S = 0.16 \pm 0.01$; (c) a Breit-Wigner function with the mass and width parameters taken from PDG [2] and with a free smearing, $F_S = 0.14 \pm 0.01$; and (d) a Breit-Wigner function with the mass and width parameters taken from PDG and the smearing from the MC simulation, $F_S = 0.12 \pm 0.01$.

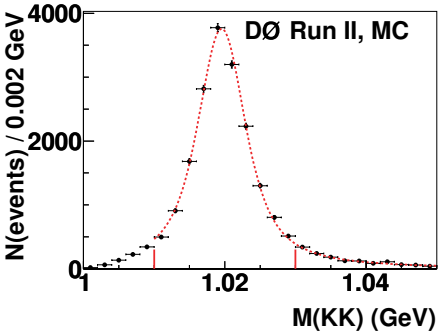
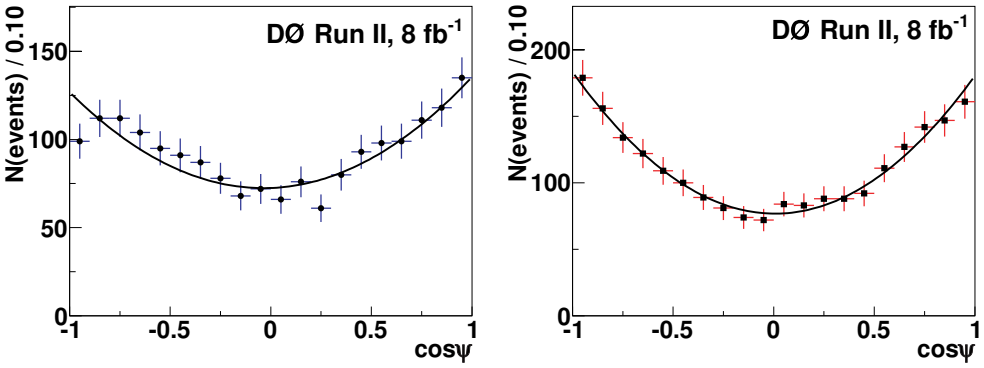


Figure C.3: The K^+K^- mass distribution in MC simulation. The fit is the Breit-Wigner function convoluted with a Gaussian resolution.



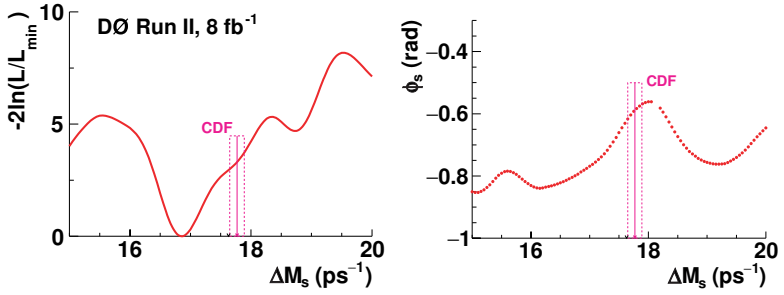
(a) (b)

Figure C.4: Projections onto the variable $\cos\psi$ for events with $M(K^+K^-)$ (a) below and (b) above the $\phi(1020)$ meson mass. The curves are results of fits to a second-degree polynomial. The linear term is (a) 4.3 ± 4.0 and (b) -1.4 ± 4.4.

D. Sensitivity to ΔM_s

D.0.1 Likelihood scan as a function of ΔM_s

While it is not possible to perform a precise measurement of the $B_s^0 - \overline{B}_s^0$ oscillation frequency with the present data, we can study all aspects of the sensitivity to its value. The inspection of MCMC chains reveals a positive correlation between the oscillation frequency ΔM_s and $\phi_s^{J/\psi\phi}$. For a more detailed insight into the sensitivity to ΔM_s, we repeat the Maximum Likelihood fits, varying ΔM_s at values between 15 and 20 ps^{-1}. The results are shown in Fig. D.1. The best fit corresponds to ΔM_s near 17 ps^{-1}, and $\phi_s^{J/\psi\phi}$ around -0.8. It agrees with the CDF value of the ΔM_s [94] to within about 1.5 standard deviations.



(a) (b)

Figure D.1: (a) The likelihood variation as a function of ΔM_s and (b) the best-fit value of $\phi_s^{J/\psi\phi}$ versus ΔM_s for the Square-cuts sample. No external constraints are applied in the fits, the only condition being $\cos\delta_\perp < 0$ and $\cos\delta_s < 0$ which selects the fit with $\Delta\Gamma_s > 0$, $\phi_s^{J/\psi\phi} < 0$. Also shown is the 68% range from CDF [94] - used as a constraint in this analysis.

D.0.2 $B_s^0 - \overline{B}_s^0$ oscillation

Under the hypothesis of CP conservation in the B_s^0 decay, and a possible mixing-induced CP violation, the non-vanishing CP-violating mixing angle should manifest itself as a $B_s^0 - \overline{B}_s^0$ oscillation with the amplitude proportional to $\sin(\phi_s^{J/\psi\phi})$. The observed time-dependent asymmetry $\Delta N \equiv N(B_s^0) - N(\overline{B}_s^0) = N_S \cdot C \cdot \sin(\phi_s^{J/\psi\phi})$, is diluted by a product C of several factors: (i) a factor of $(1-2|A_\perp|^2)\cdot(1-2F_s) \approx 0.6\cdot 0.7$ due to the presence of the CP-odd decay, (ii) a factor of $\epsilon \cdot \mathcal{D}^2 \approx 0.03$ due to the flavor tagging efficiency and accuracy, and (iii) a factor of $\exp(-(\Delta M_s \sigma)^2/2) \approx 0.2$ due to the limited time resolution. Thus, with $N_S \approx 6000$ events, and $C \approx 0.0025$, we expect $N_S \cdot C \approx 15$.

In Fig. D.2 we show the proper decay length evolution of ΔN in the first 90 μm, corresponding to approximately twice the mean B_s^0 lifetime. The curve represents a fit to the function $N_0 \cdot \sin(\Delta M_s t) \cdot \exp(-t/\tau_s)$, with N_0 unconstrained and with $\Delta M_s \equiv 17.77$ ps^{-1}. The fit gives $N_0 = -6$ for the BDT-based sample and -8 for the Square-cuts sample, with a statistical uncertainty of ± 4, corresponding to $\sin(\phi_s^{J/\psi\phi}) = N_0/N_S \cdot C \approx -0.4 \pm 0.3$. This one-dimensional analysis gives a result for $\phi_s^{J/\psi\phi}$ that is consistent with the result of the full analysis.

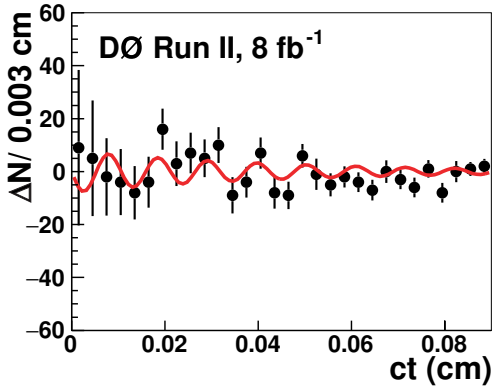
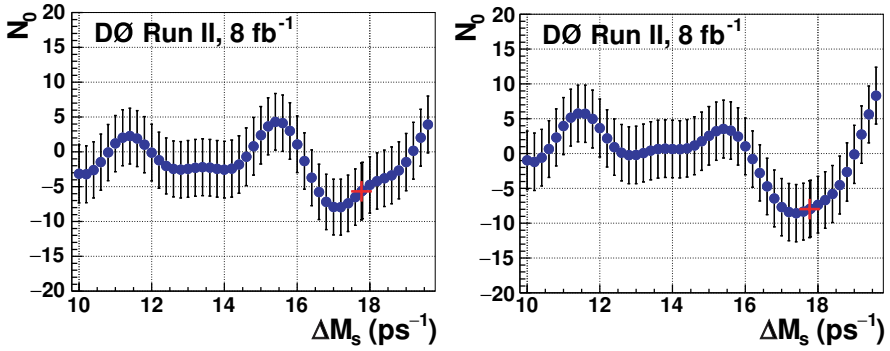


Figure D.2: Proper decay length evolution of the difference $\Delta N = N(B_s^0) - N(\overline{B}_s^0)$ in the first 0.09 cm (3 ps) for the Square-cuts sample. The curve represents the best fit to the oscillation with the frequency of $\Delta M_s = 17.77$ ps^{-1}.

Following the Amplitude Method described in Ref. [97], we fit the above distributions at discrete values of ΔM_s, and plot the fitted value of N_0 as a function of the probe frequency. The results are shown in Fig. D.3. There is an undulating structure, with no significantly large deviations from zero. At ΔM_s near 17.77 ps^{-1} the data

prefer a negative oscillation amplitude (and hence a negative value of $\sin \phi_s^{J/\psi\phi}$).



(a) (b)

Figure D.3: The fitted magnitude of the $B_s^0 - \overline{B}_s^0$ oscillation as a function of ΔM_s for (a) BDT selection and (b) Square cuts. The red crosses correspond to $\Delta M_s = 17.77$ ps^{-1}.

D. SENSITIVITY TO ΔM_S

E. Published paper

Measurement of the CP-violating phase $\phi_s^{J/\psi\phi}$ using the flavor-tagged decay $B_s^0 \to J/\psi\phi$ in 8 fb^{-1} of $p\bar{p}$ collisions

V. M. Abazov,[34] B. Abbott,[72] B. S. Acharya,[28] M. Adams,[48] T. Adams,[46] G. D. Alexeev,[34] G. Alkhazov,[38] A. Alton,[60,*] G. Alverson,[59] G. A. Alves,[2] M. Aoki,[47] M. Arov,[57] A. Askew,[46] B. Åsman,[40] S. Atkins,[57] O. Atramentov,[64] K. Augsten,[9] C. Avila,[7] J. BackusMayes,[79] F. Badaud,[12] L. Bagby,[47] B. Baldin,[47] D. V. Bandurin,[46] S. Banerjee,[28] E. Barberis,[59] P. Baringer,[55] J. Barreto,[3] J. F. Bartlett,[47] U. Bassler,[17] V. Bazterra,[48] A. Bean,[55] M. Begalli,[3] M. Begel,[70] C. Belanger-Champagne,[40] L. Bellantoni,[47] S. B. Beri,[26] G. Bernardi,[16] R. Bernhard,[21] I. Bertram,[41] M. Besançon,[17] R. Beuselinck,[42] V. A. Bezzubov,[37] P. C. Bhat,[47] V. Bhatnagar,[26] G. Blazey,[49] S. Blessing,[46] K. Bloom,[63] A. Boehnlein,[47] D. Boline,[69] E. E. Boos,[36] G. Borissov,[41] T. Bose,[58] A. Brandt,[75] O. Brandt,[22] R. Brock,[61] G. Brooijmans,[67] A. Bross,[47] D. Brown,[16] J. Brown,[16] X. B. Bu,[47] M. Buehler,[47] V. Buescher,[23] V. Bunichev,[36] S. Burdin,[41,†] T. H. Burnett,[79] C. P. Buszello,[40] B. Calpas,[14] E. Camacho-Pérez,[31] M. A. Carrasco-Lizarraga,[55] B. C. K. Casey,[47] H. Castilla-Valdez,[31] S. Chakrabarti,[69] D. Chakraborty,[49] K. M. Chan,[53] A. Chandra,[77] E. Chapon,[17] G. Chen,[55] S. Chevalier-Théry,[17] D. K. Cho,[74] S. W. Cho,[30] S. Choi,[30] B. Choudhary,[27] S. Cihangir,[47] D. Claes,[63] J. Clutter,[55] M. Cooke,[47] W. E. Cooper,[47] M. Corcoran,[77] F. Couderc,[17] M.-C. Cousinou,[14] A. Croc,[17] D. Cutts,[74] A. Das,[44] G. Davies,[42] K. De,[75] S. J. de Jong,[33] E. De La Cruz-Burelo,[31] F. Déliot,[17] M. Demarteau,[47] R. Demina,[68] D. Denisov,[47] S. P. Denisov,[37] S. Desai,[47] C. Deterre,[17] K. DeVaughan,[63] H. T. Diehl,[47] M. Diesburg,[47] P. F. Ding,[43] A. Dominguez,[63] T. Dorland,[79] A. Dubey,[27] L. V. Dudko,[36] D. Duggan,[64] A. Duperrin,[14] S. Dutt,[26] A. Dyshkant,[49] M. Eads,[63] D. Edmunds,[61] J. Ellison,[45] V. D. Elvira,[47] Y. Enari,[16] H. Evans,[51] A. Evdokimov,[70] V. N. Evdokimov,[37] G. Facini,[59] T. Ferbel,[68] F. Fiedler,[23] F. Filthaut,[33] W. Fisher,[61] H. E. Fisk,[47] M. Fortner,[49] H. Fox,[41] S. Fuess,[47] A. Garcia-Bellido,[68] G. A. García-Guerra,[31,‡] V. Gavrilov,[35] P. Gay,[12] W. Geng,[14,61] D. Gerbaudo,[65] C. E. Gerber,[48] Y. Gershtein,[64] G. Ginther,[47,68] G. Golovanov,[34] A. Goussiou,[79] P. D. Grannis,[69] S. Greder,[18] H. Greenlee,[47] Z. D. Greenwood,[57] E. M. Gregores,[4] G. Grenier,[19] Ph. Gris,[12] J.-F. Grivaz,[15] A. Grohsjean,[17] S. Grünendahl,[47] M. W. Grünewald,[29] T. Guillemin,[15] G. Gutierrez,[47] P. Gutierrez,[72] A. Haas,[67,§] S. Hagopian,[46] J. Haley,[59] L. Han,[6] K. Harder,[43] A. Harel,[68] J. M. Hauptman,[54] J. Hays,[42] T. Head,[43] T. Hebbeker,[20] D. Hedin,[49] H. Hegab,[73] A. P. Heinson,[45] U. Heintz,[74] C. Hensel,[22] I. Heredia-De La Cruz,[31] K. Herner,[60] G. Hesketh,[43,‖] M. D. Hildreth,[53] R. Hirosky,[78] T. Hoang,[46] J. D. Hobbs,[69] B. Hoeneisen,[11] M. Hohlfeld,[23] Z. Hubacek,[9,17] N. Huske,[16] V. Hynek,[9] I. Iashvili,[66] Y. Ilchenko,[76] R. Illingworth,[47] A. S. Ito,[47] S. Jabeen,[74] M. Jaffré,[15] D. Jamin,[14] A. Jayasinghe,[72] R. Jesik,[42] K. Johns,[44] M. Johnson,[47] A. Jonckheere,[47] P. Jonsson,[42] J. Joshi,[26] A. W. Jung,[47] A. Juste,[39] K. Kaadze,[56] E. Kajfasz,[14] D. Karmanov,[36] P. A. Kasper,[47] I. Katsanos,[63] R. Kehoe,[76] S. Kermiche,[14] N. Khalatyan,[47] A. Khanov,[73] A. Kharchilava,[66] Y. N. Kharzheev,[34] J. M. Kohli,[26] A. V. Kozelov,[37] J. Kraus,[61] S. Kulikov,[37] A. Kumar,[66] A. Kupco,[10] T. Kurča,[19] V. A. Kuzmin,[36] J. Kvita,[8] S. Lammers,[51] G. Landsberg,[74] P. Lebrun,[19] H. S. Lee,[30] S. W. Lee,[54] W. M. Lee,[47] J. Lellouch,[16] L. Li,[45] Q. Z. Li,[47] S. M. Lietti,[5] J. K. Lim,[30] D. Lincoln,[47] J. Linnemann,[61] V. V. Lipaev,[37] R. Lipton,[47] Y. Liu,[6] A. Lobodenko,[38] M. Lokajicek,[10] R. Lopes de Sa,[69] H. J. Lubatti,[79] R. Luna-Garcia,[31,¶] A. L. Lyon,[47] A. K. A. Maciel,[2] D. Mackin,[77] R. Madar,[17] R. Magaña-Villalba,[31] S. Malik,[63] V. L. Malyshev,[34] Y. Maravin,[56] J. Martínez-Ortega,[31] R. McCarthy,[69] C. L. McGivern,[55] M. M. Meijer,[33] A. Melnitchouk,[62] D. Menezes,[49] P. G. Mercadante,[4] M. Merkin,[36] A. Meyer,[20] J. Meyer,[22] F. Miconi,[18] N. K. Mondal,[28] G. S. Muanza,[14] M. Mulhearn,[78] E. Nagy,[14] M. Naimuddin,[27] M. Narain,[74] R. Nayyar,[27] H. A. Neal,[60] J. P. Negret,[7] P. Neustroev,[38] S. F. Novaes,[5] T. Nunnemann,[24] G. Obrant,[38,‡‡] J. Orduna,[77] N. Osman,[14] J. Osta,[53] G. J. Otero y Garzón,[1] M. Padilla,[45] A. Pal,[75] N. Parashar,[52] V. Parihar,[74] S. K. Park,[30] J. Parsons,[67] R. Partridge,[74,§] N. Parua,[51] A. Patwa,[70] B. Penning,[47] M. Perfilov,[36] K. Peters,[43] Y. Peters,[43] K. Petridis,[43] G. Petrillo,[68] P. Pétroff,[15] R. Piegaia,[1] M.-A. Pleier,[70] P. L. M. Podesta-Lerma,[31,**] V. M. Podstavkov,[47] P. Polozov,[35] A. V. Popov,[37] M. Prewitt,[77] D. Price,[51] N. Prokopenko,[37] S. Protopopescu,[70] J. Qian,[60] A. Quadt,[22] B. Quinn,[62] M. S. Rangel,[2] K. Ranjan,[27] P. N. Ratoff,[41] I. Razumov,[37] P. Renkel,[76] M. Rijssenbeek,[69] I. Ripp-Baudot,[18] F. Rizatdinova,[73] M. Rominsky,[47] A. Ross,[41] C. Royon,[17] P. Rubinov,[47] R. Ruchti,[53] G. Safronov,[35] G. Sajot,[13] P. Salcido,[49] A. Sánchez-Hernández,[31] M. P. Sanders,[24] B. Sanghi,[47] A. S. Santos,[5] G. Savage,[47] L. Sawyer,[57] T. Scanlon,[42] R. D. Schamberger,[69] Y. Scheglov,[38] H. Schellman,[50] T. Schliephake,[25] S. Schlobohm,[79] C. Schwanenberger,[43] R. Schwienhorst,[61] J. Sekaric,[55] H. Severini,[72] E. Shabalina,[22] V. Shary,[17] A. A. Shchukin,[37] R. K. Shivpuri,[27] V. Simak,[9] V. Sirotenko,[47] P. Skubic,[72] P. Slattery,[68] D. Smirnov,[53] K. J. Smith,[66] G. R. Snow,[63] J. Snow,[71] S. Snyder,[70] S. Söldner-Rembold,[43] L. Sonnenschein,[20] K. Soustruznik,[8] J. Stark,[13] V. Stolin,[35] D. A. Stoyanova,[37] M. Strauss,[72] D. Strom,[48] L. Stutte,[47] L. Suter,[43] P. Svoisky,[72] M. Takahashi,[43] A. Tanasijczuk,[1] M. Titov,[17] V. V. Tokmenin,[34] Y.-T. Tsai,[68] K. Tschann-Grimm,[69] D. Tsybychev,[69] B. Tuchming,[17] C. Tully,[65] L. Uvarov,[38] S. Uvarov,[38] S. Uzunyan,[49] R. Van Kooten,[51] W. M. van Leeuwen,[32] N. Varelas,[48] E. W. Varnes,[44] I. A. Vasilyev,[37]

 © 2012 American Physical Society

P. Verdier,[19] L. S. Vertogradov,[34] M. Verzocchi,[47] M. Vesterinen,[43] D. Vilanova,[17] P. Vokac,[9] H. D. Wahl,[46] M. H. L. S. Wang,[47] J. Warchol,[53] G. Watts,[79] M. Wayne,[53] M. Weber,[47,††] L. Welty-Rieger,[50] A. White,[75] D. Wicke,[25] M. R. J. Williams,[41] G. W. Wilson,[55] M. Wobisch,[57] D. R. Wood,[59] T. R. Wyatt,[43] Y. Xie,[47] C. Xu,[60] S. Yacoob,[50] R. Yamada,[47] W.-C. Yang,[43] T. Yasuda,[47] Y. A. Yatsunenko,[34] Z. Ye,[47] H. Yin,[47] K. Yip,[70] S. W. Youn,[47] J. Yu,[75] S. Zelitch,[78] T. Zhao,[79] B. Zhou,[60] J. Zhu,[60] M. Zielinski,[68] D. Zieminska,[51] and L. Zivkovic[74]

(The D0 Collaboration)

[1]*Universidad de Buenos Aires, Buenos Aires, Argentina*
[2]*LAFEX, Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro, Brazil*
[3]*Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil*
[4]*Universidade Federal do ABC, Santo André, Brazil*
[5]*Instituto de Física Teórica, Universidade Estadual Paulista, São Paulo, Brazil*
[6]*University of Science and Technology of China, Hefei, People's Republic of China*
[7]*Universidad de los Andes, Bogotá, Colombia*
[8]*Charles University, Faculty of Mathematics and Physics, Center for Particle Physics, Prague, Czech Republic*
[9]*Czech Technical University in Prague, Prague, Czech Republic*
[10]*Center for Particle Physics, Institute of Physics, Academy of Sciences of the Czech Republic, Prague, Czech Republic*
[11]*Universidad San Francisco de Quito, Quito, Ecuador*
[12]*LPC, Université Blaise Pascal, CNRS/IN2P3, Clermont, France*
[13]*LPSC, Université Joseph Fourier Grenoble 1, CNRS/IN2P3, Institut National Polytechnique de Grenoble, Grenoble, France*
[14]*CPPM, Aix-Marseille Université, CNRS/IN2P3, Marseille, France*
[15]*LAL, Université Paris-Sud, CNRS/IN2P3, Orsay, France*
[16]*LPNHE, Universités Paris VI and VII, CNRS/IN2P3, Paris, France*
[17]*CEA, Irfu, SPP, Saclay, France*
[18]*IPHC, Université de Strasbourg, CNRS/IN2P3, Strasbourg, France*
[19]*IPNL, Université Lyon 1, CNRS/IN2P3, Villeurbanne, France and Université de Lyon, Lyon, France*
[20]*III. Physikalisches Institut A, RWTH Aachen University, Aachen, Germany*
[21]*Physikalisches Institut, Universität Freiburg, Freiburg, Germany*
[22]*II. Physikalisches Institut, Georg-August-Universität Göttingen, Göttingen, Germany*
[23]*Institut für Physik, Universität Mainz, Mainz, Germany*
[24]*Ludwig-Maximilians-Universität München, München, Germany*
[25]*Fachbereich Physik, Bergische Universität Wuppertal, Wuppertal, Germany*
[26]*Panjab University, Chandigarh, India*
[27]*Delhi University, Delhi, India*
[28]*Tata Institute of Fundamental Research, Mumbai, India*
[29]*University College Dublin, Dublin, Ireland*
[30]*Korea Detector Laboratory, Korea University, Seoul, Korea*
[31]*CINVESTAV, Mexico City, Mexico*
[32]*Nikhef, Science Park, Amsterdam, the Netherlands*
[33]*Radboud University Nijmegen, Nijmegen, the Netherlands and Nikhef, Science Park, Amsterdam, the Netherlands*
[34]*Joint Institute for Nuclear Research, Dubna, Russia*
[35]*Institute for Theoretical and Experimental Physics, Moscow, Russia*
[36]*Moscow State University, Moscow, Russia*
[37]*Institute for High Energy Physics, Protvino, Russia*
[38]*Petersburg Nuclear Physics Institute, St. Petersburg, Russia*
[39]*Institució Catalana de Recerca i Estudis Avançats (ICREA) and Institut de Física d'Altes Energies (IFAE), Barcelona, Spain*
[40]*Stockholm University, Stockholm and Uppsala University, Uppsala, Sweden*
[41]*Lancaster University, Lancaster LA1 4YB, United Kingdom*
[42]*Imperial College London, London SW7 2AZ, United Kingdom*
[43]*The University of Manchester, Manchester M13 9PL, United Kingdom*
[44]*University of Arizona, Tucson, Arizona 85721, USA*
[45]*University of California Riverside, Riverside, California 92521, USA*
[46]*Florida State University, Tallahassee, Florida 32306, USA*
[47]*Fermi National Accelerator Laboratory, Batavia, Illinois 60510, USA*
[48]*University of Illinois at Chicago, Chicago, Illinois 60607, USA*
[49]*Northern Illinois University, DeKalb, Illinois 60115, USA*
[50]*Northwestern University, Evanston, Illinois 60208, USA*
[51]*Indiana University, Bloomington, Indiana 47405, USA*

[52]*Purdue University Calumet, Hammond, Indiana 46323, USA*
[53]*University of Notre Dame, Notre Dame, Indiana 46556, USA*
[54]*Iowa State University, Ames, Iowa 50011, USA*
[55]*University of Kansas, Lawrence, Kansas 66045, USA*
[56]*Kansas State University, Manhattan, Kansas 66506, USA*
[57]*Louisiana Tech University, Ruston, Louisiana 71272, USA*
[58]*Boston University, Boston, Massachusetts 02215, USA*
[59]*Northeastern University, Boston, Massachusetts 02115, USA*
[60]*University of Michigan, Ann Arbor, Michigan 48109, USA*
[61]*Michigan State University, East Lansing, Michigan 48824, USA*
[62]*University of Mississippi, University, Mississippi 38677, USA*
[63]*University of Nebraska, Lincoln, Nebraska 68588, USA*
[64]*Rutgers University, Piscataway, New Jersey 08855, USA*
[65]*Princeton University, Princeton, New Jersey 08544, USA*
[66]*State University of New York, Buffalo, New York 14260, USA*
[67]*Columbia University, New York, New York 10027, USA*
[68]*University of Rochester, Rochester, New York 14627, USA*
[69]*State University of New York, Stony Brook, New York 11794, USA*
[70]*Brookhaven National Laboratory, Upton, New York 11973, USA*
[71]*Langston University, Langston, Oklahoma 73050, USA*
[72]*University of Oklahoma, Norman, Oklahoma 73019, USA*
[73]*Oklahoma State University, Stillwater, Oklahoma 74078, USA*
[74]*Brown University, Providence, Rhode Island 02912, USA*
[75]*University of Texas, Arlington, Texas 76019, USA*
[76]*Southern Methodist University, Dallas, Texas 75275, USA*
[77]*Rice University, Houston, Texas 77005, USA*
[78]*University of Virginia, Charlottesville, Virginia 22901, USA*
[79]*University of Washington, Seattle, Washington 98195, USA*

(Received 15 September 2011; published 22 February 2012)

We report an updated 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 $B_s^0 \to J/\psi\phi$. The data sample corresponds to an integrated luminosity of 8.0 fb^{-1} accumulated with the D0 detector using $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV produced at the Fermilab Tevatron collider. The 68% Bayesian credibility intervals, 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%.

DOI: 10.1103/PhysRevD.85.032006 PACS numbers: 13.25.Hw, 11.30.Er

I. INTRODUCTION

The meson-antimeson mixing and the phenomenon of charge-conjugation-parity (*CP*) violation in neutral mesons systems are key problems of particle physics. In the standard model (SM), the light (L) and heavy (H) mass eigenstates of the B_s^0 system are expected to have sizeable mass and decay-width differences: $\Delta M_s \equiv M_H - M_L$ and $\Delta\Gamma_s \equiv \Gamma_L - \Gamma_H$. The two mass eigenstates are expected to be almost pure *CP* eigenstates. The *CP*-violating phase that appears in $b \to c\bar{c}s$ decays is due to the interference of the decay with and without mixing, and it is predicted [1] to be $\phi_s^{J/\psi\phi} = -2\beta_s^{\rm SM} = 2\arg[-V_{tb}V_{ts}^*/V_{cb}V_{cs}^*] = -0.038 \pm 0.002$, where V_{ij} are elements of the Cabibbo-Kobayashi-Maskawa quark-mixing matrix [2]. New phenomena [3–23] may alter the observed phase to $\phi_s^{J/\psi\phi} \equiv -2\beta_s \equiv -2\beta_s^{\rm SM} + \phi_s^{\Delta}$. A significant deviation of $\phi_s^{J/\psi\phi}$ from its small SM value would indicate the presence of processes beyond SM.

The analysis of the decay chain $B_s^0 \to J/\psi\phi$, $J/\psi \to \mu^+\mu^-$, $\phi \to K^+K^-$ separates the *CP*-even and *CP*-odd states using the angular distributions of the decay products. It is a unique feature of the decay $B_s^0 \to J/\psi\phi$ that because of the sizeable lifetime difference between the two mass eigenstates, there is a sensitivity to $\phi_s^{J/\psi\phi}$ even in the absence of the flavor-tagging information. The first direct constraint on $\phi_s^{J/\psi\phi}$ [24,25] was derived by analyzing $B_s^0 \to J/\psi\phi$ decays where the flavor (i.e., B_s^0 or $\bar{B}_s^0$) at

*with visitor from Augustana College, Sioux Falls, SD, USA
†with visitor from The University of Liverpool, Liverpool, UK
‡with visitor from UPIITA-IPN, Mexico City, Mexico
§with visitor from SLAC, Menlo Park, CA, USA
‖with visitor from University College London, London, UK
¶with visitor from Centro de Investigacion en Computacion - IPN, Mexico City, Mexico
**with visitor from ECFM, Universidad Autonoma de Sinaloa, Culiacán, Mexico
††with visitor from Universität Bern, Bern, Switzerland
‡‡Deceased

the time of production was not determined (''tagged''). It was followed by an improved analysis [26], based on 2.8 fb^{-1} of integrated luminosity, that included the information on the B_s^0 flavor at production. In addition, the CDF collaboration has performed a measurement [27] of $\phi_s^{J/\psi\phi}$ using 1.35 fb^{-1} of data. After the submission of this article, new measurements of the *CP* violation parameters in the $B_s^0 \to J/\psi\phi$ decay have been published by the CDF [28] and the LHCb [29] Collaborations.

In this article, we present new results from the time-dependent amplitude analysis of the decay $B_s^0 \to J/\psi\phi$ using a data sample corresponding to an integrated luminosity of 8.0 fb^{-1} collected with the D0 detector [30] at the Fermilab Tevatron Collider. In addition to the increase in the size of the data sample used in the analysis, we also take into account the $\mathcal{S}$-wave K^+K^- under the ϕ peak that has been suggested [31] to contribute between 5% and 10%. We measure $\Delta\Gamma_s$; the average lifetime of the B_s^0 system, $\bar{\tau}_s = 1/\bar{\Gamma}_s$, where $\bar{\Gamma}_s \equiv (\Gamma_H + \Gamma_L)/2$; and the *CP*-violating phase $\phi_s^{J/\psi\phi}$. Section II briefly describes the D0 detector. Section III presents the event reconstruction and the data set. Sections IV and V describe the event selection requirements and the procedure of determining the flavor of the initial state of the B_s^0 candidate. In Sec. VI, we describe the analysis formalism and the fitting method, present fit results, and discuss systematic uncertainties in the results. We obtain the Bayesian credibility intervals for physics parameters using a procedure based on the Markov Chain Monte Carlo (MCMC) technique, presented in Sec. VII. We summarize and discuss the results in Sec. VIII.

II. DETECTOR

The D0 detector consists of a central tracking system, calorimetry system, and muon detectors, as detailed in Refs. [30,32,33]. The central tracking system comprises a silicon microstrip tracker (SMT) and a central fiber tracker (CFT), both located inside a 1.9 T superconducting solenoidal magnet. The tracking system is designed to optimize tracking and vertexing for pseudorapidities $|\eta| < 3$, where $\eta = -\ln[\tan(\Theta/2)]$, and Θ is the polar angle with respect to the proton beam direction.

The SMT can reconstruct the $p\bar{p}$ interaction vertex (PV) for interactions with at least three tracks with a precision of 40 μm in the plane transverse to the beam direction and determine the impact parameter of any track relative to the PV with a precision between 20 and 50 μm, depending on the number of hits in the SMT.

The muon detector is positioned outside the calorimeter. It consists of a central muon system covering the pseudorapidity region $|\eta| < 1$ and a forward muon system covering the pseudorapidity region $1 < |\eta| < 2$. Both central and forward systems consist of a layer of drift tubes and scintillators inside 1.8 T toroidal magnets and two similar layers outside the toroids.

The trigger and data acquisition systems are designed to accommodate the high instantaneous luminosities of Tevatron Run II.

III. DATA SAMPLE AND EVENT RECONSTRUCTION

The analysis presented here is based on data accumulated between February 2002 and June 2010. Events are collected with a mixture of single-muon 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 [34] 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 characterizing 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 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 [34], 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 5×10^6 events are accepted after the selection described in this section.

IV. 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 multibody decay of a b hadron or from hadronization. Two different event selection approaches are used, one based on a multivariate technique, and one based on simple limits on kinematic and event quality parameters.

A. 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 [35]. Hadronization is also done in PYTHIA, but all hadrons carrying heavy flavors are passed on to EVTGEN [36] 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 $\bar{\tau}_s = 1.464$ ps. There are approximately 10^6 events in each background and the signal MC samples. All events are passed through a full GEANT-based [37] detector simulation. To take into account the effects of multiple interactions at high luminosity, hits from randomly triggered $p\bar{p}$ collisions are overlaid 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 Appendix B).

B. Multivariate event selection

To discriminate the signal from background events, we use the TMVA package [38]. 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 Appendix A.

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 84 k events, the prompt background has 29 k events, and the non-prompt background has 39 k events. Figure 1 shows the BDT output discriminant for the prompt and non-prompt cases.

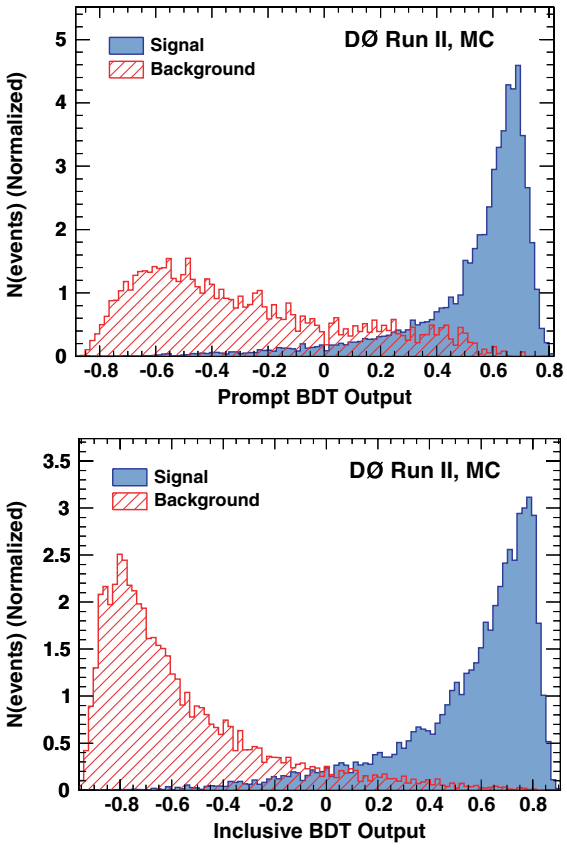


FIG. 1 (color online). BDT discriminant output for the prompt (top) and non-prompt (bottom) classifiers. The signal and background events are taken from simulation. Events used for BDT training are excluded from these samples.

C. 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 I. Figure 2 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 using only the variables whose importance (see Appendix A) exceeds 0.01, giving 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

TABLE I. 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	38 130	0.45	0.42
1	4699	44 535	0.45	0.29
2	5008	53 942	0.39	0.35
3	5213	64 044	0.36	0.30
4	5364	72 602	0.33	0.28
5	5558	85 848	0.13	0.41
6	5767	100 986	0.21	0.29
7	5988	120 206	0.13	0.29
8	6097	134 255	0.07	0.29
9	6399	189 865	0.04	0.10
10	6489	254 022	−0.05	−0.01
11	6608	294 949	−0.13	0.00
12	6594	364 563	−0.18	−0.14
13	6695	461 744	−0.35	−0.08

about 10%. Therefore, we proceed with the original number of variables.

The choice of the final cut on the BDT output is based on an ensemble study. For each point in Table I, 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 pseudoexperiments in the 5D space of B_s^0 candidate 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 $\phi_s^{J/\psi\phi}$ ($\sigma(\phi_s^{J/\psi\phi})$) 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 Figs. 3(a) and 3(b). The mean statistical

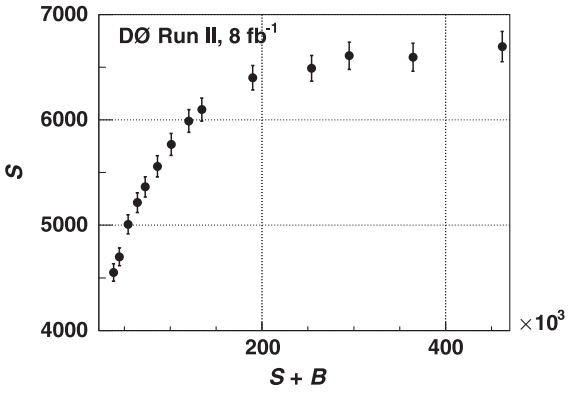


FIG. 2. Number of $B_s^0 \to J/\psi\phi$ signal events as a function of the total number of events for the 14 criteria sets considered.

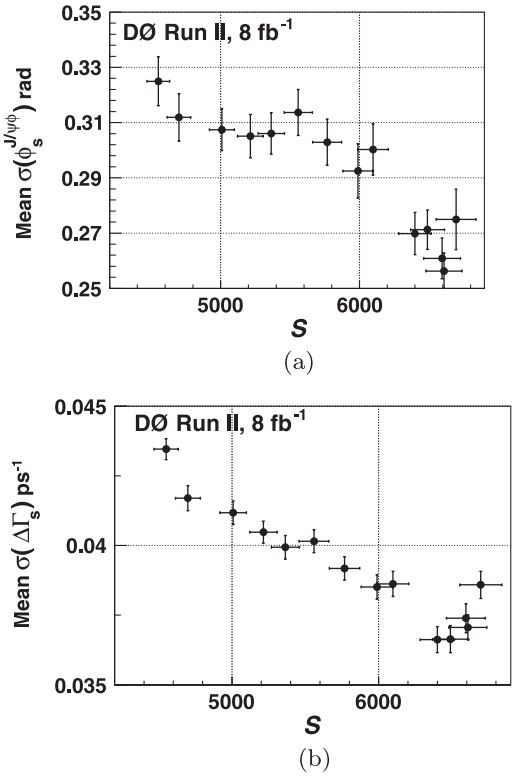


FIG. 3. Ensemble study results for (a) mean value of $\sigma(\phi_s)$ as a function of the number of signal events and (b) mean value of $\sigma(\Delta\Gamma_s)$ as a function of the number of signal events.

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 I) as a final selection and refer to it as the "BDT selection." Figure 17 in Appendix A shows the ratios of the normalized distributions of the three angles (see Sec. VI) 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.

D. 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. [24,26]. 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 ± 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.

V. 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" (OST)]. The OST-discriminating variables x_i are based primarily on the presence of a muon or an electron from the semileptonic decay or the decay-vertex charge of the other b hadron produced in the $p\bar{p}$ interaction.

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:

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

A given variable x_i can be undefined for some events. For example, there are events that do not 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 a fraction of the number of candidates with $d \neq 0$, is 18%. The OST-discriminating variables and algorithm are described in detail in Ref. [39].

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

$$\mathcal{D} = \frac{N_{\text{cor}} - N_{\text{wr}}}{N_{\text{cor}} + N_{\text{wr}}}, \tag{2}$$

where N_{cor} (N_{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 $\bar{B}$) is known.

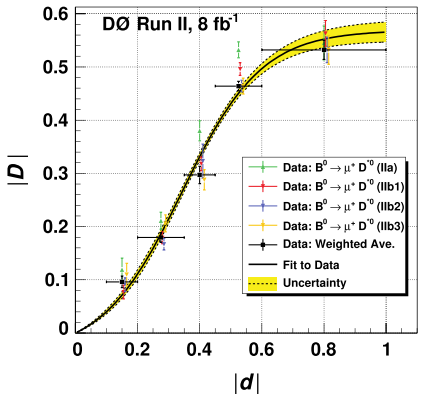


FIG. 4 (color online). Parametrization of the dilution $|\mathcal{D}|$ as a function of the tagging parameter $|d|$ for the combined opposite-side tagger. The curve is the result of the weighted fit to four self-tagging control data samples (see text).

OST calibration

The dilution calibration is based on four independent $B_d^0 \to \mu\nu D^{*\pm}$ 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. [39].

For each sample we perform an analysis of the $B_d^0 - \bar{B}_d^0$ oscillations described in Ref. [40]. 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 Fig. 4. The dependence of the dilution on $|d|$ is parametrized as

$$|\mathcal{D}| = \frac{p_0}{(1+\exp((p_1-|d|)/p_2))} - \frac{p_0}{(1+\exp(p_1/p_2))}. \tag{3}$$

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

VI. 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 104 683 events in the BDT-based sample and 66 455 events in the Square-cuts

sample. We adopt the formulas and notation of Ref. [41]. 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 [42].

A. Signal model

The angular distribution of the signal is expressed in the transversity basis [43]. 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 Fig. 5.

In this basis, the decay amplitude of the B_s^0 and $\bar{B}_s^0$ mesons is decomposed into three independent components corresponding to linear polarization states of the vector mesons J/ψ and ϕ, which are polarized either longitudinally (0) or transversely to their direction of motion, and parallel ($\parallel$) or perpendicular ($\perp$) to each other.

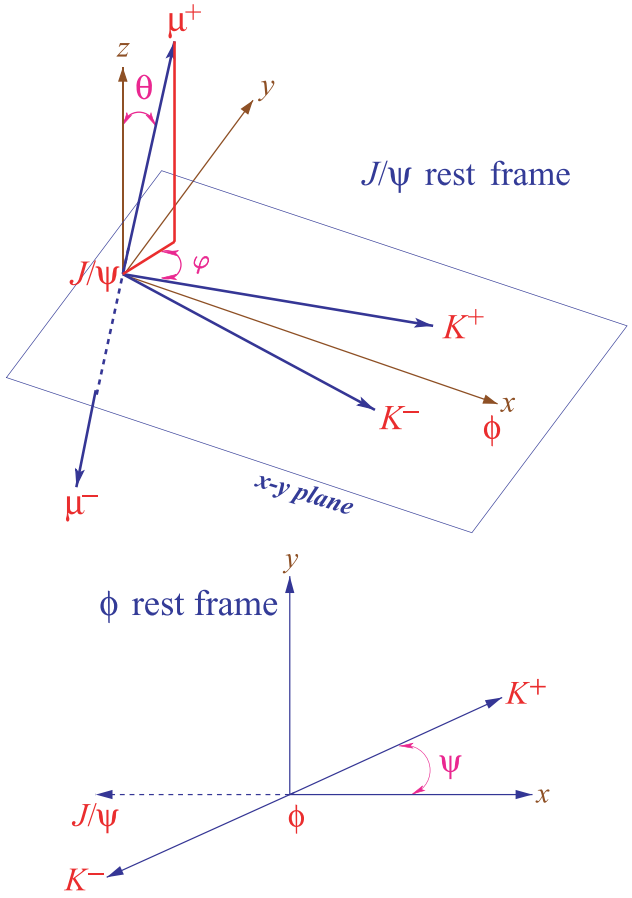


FIG. 5 (color online). Definition of the angle ψ, and the transversity angles θ and φ.

The time dependence of amplitudes $\mathcal{A}_i(t)$ and $\bar{\mathcal{A}}_i(t)$ (i denotes one of $\{\parallel, \perp, 0\}$), for B_s^0 and $\bar{B}_s^0$ states to reach the final state $J/\psi\ \phi$ is

$$\begin{aligned} \mathcal{A}_i(t) &= F(t)[E_+(t) \pm e^{2i\beta_s}E_-(t)]a_i, \\ \bar{\mathcal{A}}_i(t) &= F(t)[\pm E_+(t) + e^{-2i\beta_s}E_-(t)]a_i, \end{aligned} \tag{4}$$

where

$$F(t) = \frac{e^{-\Gamma_s t/2}}{\sqrt{\tau_H + \tau_L \pm \cos 2\beta_s(\tau_L - \tau_H)}}, \tag{5}$$

and τ_H and τ_L are the lifetimes of the heavy and light B_s^0 eigenstates.

In the above equations the upper sign indicates a CP-even final state, the lower sign indicates a CP-odd final state,

$$E_\pm(t) \equiv \frac{1}{2}\left[e^{((-\Delta\Gamma_s)/(4)+i(\Delta M_s)/(2))t} \pm e^{-((-\Delta\Gamma_s)/(4)+i(\Delta M_s)/(2))t}\right], \tag{6}$$

and the amplitude parameters a_i give the time-integrated decay rate to each of the polarization states, $|a_i|^2$, satisfying:

$$\sum_i |a_i|^2 = 1. \tag{7}$$

The interference terms $A_\parallel - A_\perp$ and $A_0 - A_\perp$ are proportional to $(e^{-\Gamma_H t} - e^{-\Gamma_L t})\sin\phi_s^{J/\psi\phi}$. Also, if $\cos\phi_s^{J/\psi\phi}$ is significantly different from unity, the decay rates of the CP-even and CP-odd components have two slopes each.

The 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, are

$$\begin{aligned} P_B(\theta, \varphi, \psi, t) &= \frac{9}{16\pi}|\mathbf{A}(t) \times \hat{n}|^2, \\ P_{\bar{B}}(\theta, \varphi, \psi, t) &= \frac{9}{16\pi}|\bar{\mathbf{A}}(\mathbf{t}) \times \hat{n}|^2, \end{aligned} \tag{8}$$

where $\hat{n}$ is the muon momentum direction in the J/ψ rest frame,

$$\hat{n} = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta), \tag{9}$$

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

$$\begin{aligned} \mathbf{A}(t) &= \left(\mathcal{A}_0(t)\cos\psi, -\frac{\mathcal{A}_\parallel(t)\sin\psi}{\sqrt{2}}, i\frac{\mathcal{A}_\perp(t)\sin\psi}{\sqrt{2}}\right), \\ \bar{\mathbf{A}}(t) &= \left(\bar{\mathcal{A}}_0(t)\cos\psi, -\frac{\bar{\mathcal{A}}_\parallel(t)\sin\psi}{\sqrt{2}}, i\frac{\bar{\mathcal{A}}_\perp(t)\sin\psi}{\sqrt{2}}\right). \end{aligned} \tag{10}$$

The values of $\mathcal{A}_i(t)$ at $t = 0$ are denoted as A_i. They are related to the parameters a_i by

$$|A_\perp|^2 = \frac{|a_\perp|^2 y}{1+(y-1)|a_\perp|^2}, \quad |A_\parallel|^2 = \frac{|a_\parallel|^2}{1+(y-1)|a_\perp|^2}, \quad |A_0|^2 = \frac{|a_0|^2}{1+(y-1)|a_\perp|^2}, \tag{11}$$

where $y \equiv (1-z)/(1+z)$ and $z \equiv \cos 2\beta_s \Delta\Gamma_s/(2\bar{\Gamma}_s)$. By convention, the phase of A_0 is set to zero and the phases of the other two amplitudes are denoted by $\delta_\parallel$ and $\delta_\perp$.

For a given event, the decay rate is the sum of the functions P_B and $P_{\bar{B}}$ weighted by the flavor-tagging dilution factors $(1+\mathcal{D})/2$ and $(1-\mathcal{D})/2$, respectively.

The contribution from the decay to $J/\psi K^+K^-$ with the kaons in an $\mathcal{S}$-wave is expressed in terms of the $\mathcal{S}$-wave fraction F_S and a phase δ_s. The squared sum of the $\mathcal{P}$ and $\mathcal{S}$ waves is integrated over the KK mass. For the $\mathcal{P}$-wave, we assume the nonrelativistic Breit-Wigner model

$$g(M(KK)) = \sqrt{\frac{\Gamma_\phi/2}{\Delta M(KK)}} \cdot \frac{1}{M(KK) - M_\phi + i\Gamma_\phi/2} \tag{12}$$

with the ϕ meson mass $M_\phi = 1.019$ GeV and width $\Gamma_\phi = 4.26$ MeV [34], and with $\Delta M(KK) = 1.03 - 1.01 = 0.02$ GeV.

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. [44]. Table II lists all physics parameters used in the fit.

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 Appendix B.

TABLE II. Definition of nine real measurables for the decay $B_s^0 \to J/\psi\phi$ used in the maximum-likelihood fitting.

Parameter	Definition
$\|A_0\|^2$	$\mathcal{P}$-wave longitudinal amplitude squared, at $t = 0$
$A1$	$\|A_\parallel\|^2/(1-\|A_0\|^2)$
$\bar{\tau}_s$ (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	$\arg(A_s/A_0)$

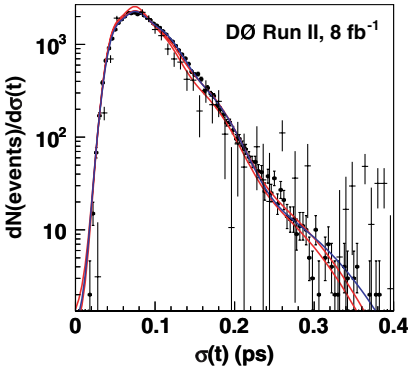


FIG. 6 (color online). The distribution of the uncertainty in the decay time for the signal, MC (squares) and background-subtracted data (crosses). The (blue) curve is the sum of five Gaussian functions fitted to the MC distribution. The two (red) lines are variations of the default function used in the studies of systematic effects.

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 Fig. 6. 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.

B. 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 II, 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.

C. 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 Fig. 6, 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 Tables III and IV. 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 multimodal 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 7–10 illustrate the quality of the fit for the background, for all data, and for the signal-enhanced subsamples.

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

D. Systematic uncertainties

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

(i) *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.

(ii) *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

TABLE III. Maximum-likelihood fit results for the BDT selection. The uncertainties are statistical.

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
$\bar{\tau}_s$ (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 IV. 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
$\bar{\tau}_s$ (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

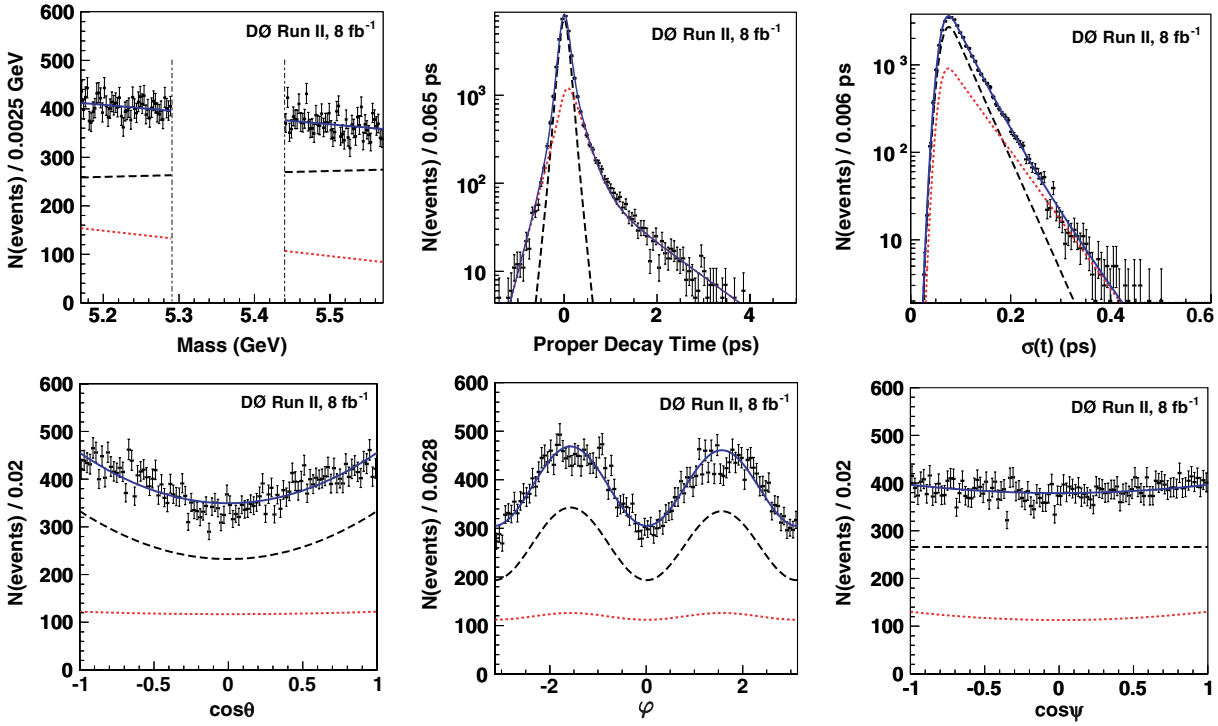


FIG. 7 (color online). The distributions in the background (B_s^0 mass sidebands) region of candidate mass, proper decay time, decay-time uncertainty, transversity polar and azimuthal angles, and $\cos\psi$ for the BDT sample. The curves show the prompt (black dashed) and non-prompt (red dotted) components, and their sum (blue solid).

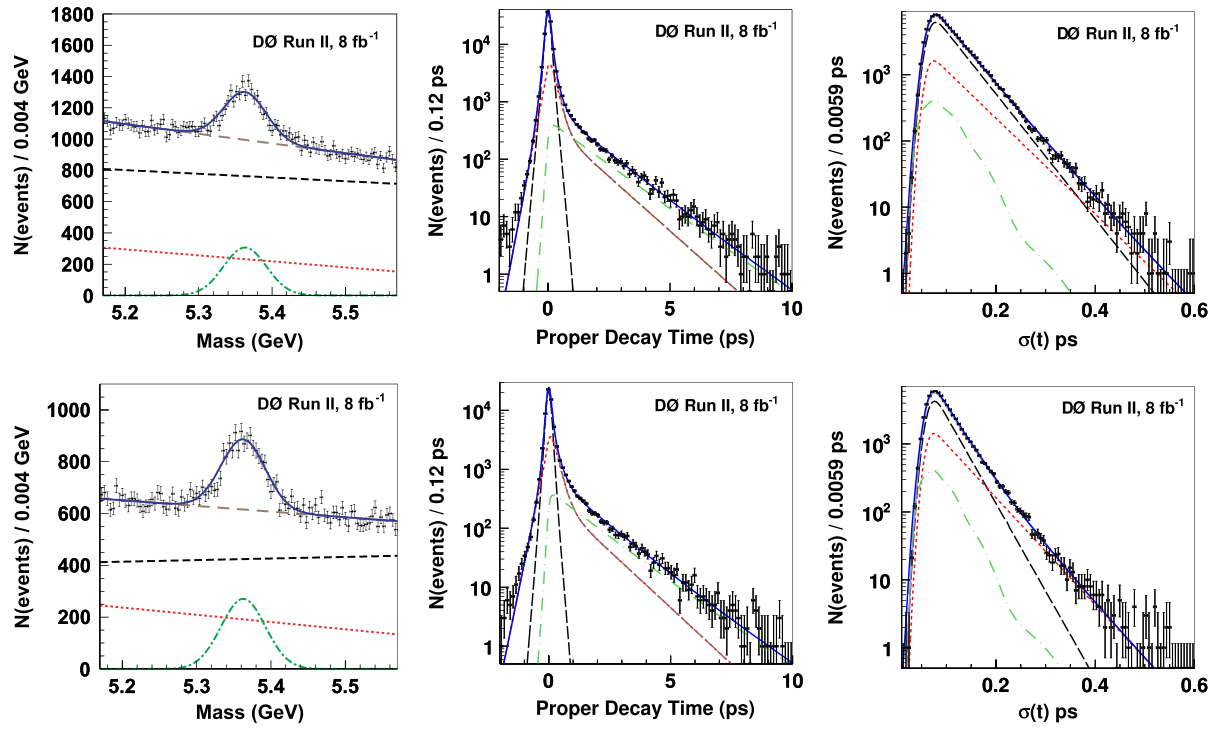


FIG. 8 (color online). Invariant mass, proper decay time, and proper decay-time uncertainty distributions for B_s^0 candidates in the (top) BDT sample and (bottom) Square-cuts sample. The curves are projections of the maximum-likelihood fit. Shown are the signal (green dashed-dotted curve), prompt background (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).

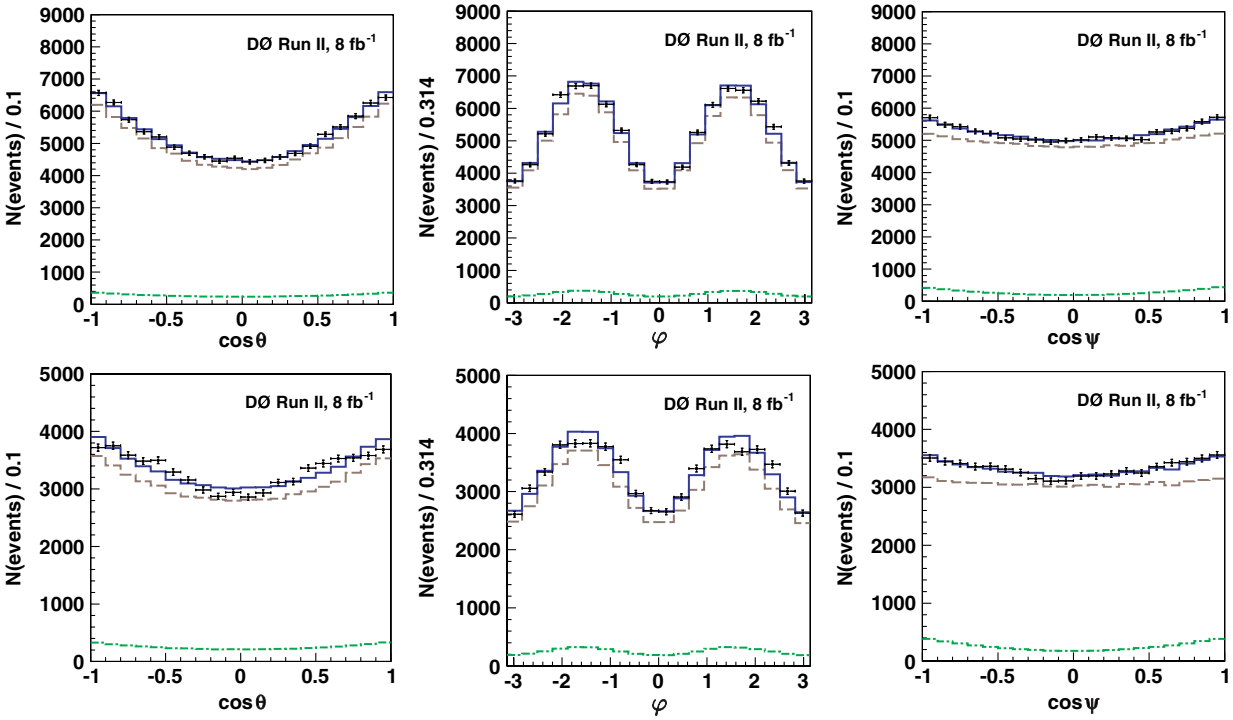


FIG. 9 (color online). Distributions of transversity polar and azimuthal angles and $\cos\psi$ for B_s^0 candidates in the BDT sample (top) and Square-cuts sample (bottom). The curves are projections of the maximum-likelihood fit. Shown are the signal (green dashed-dotted curve), total background (brown long-dashed curve) and the sum of signal and total background (blue solid curve).

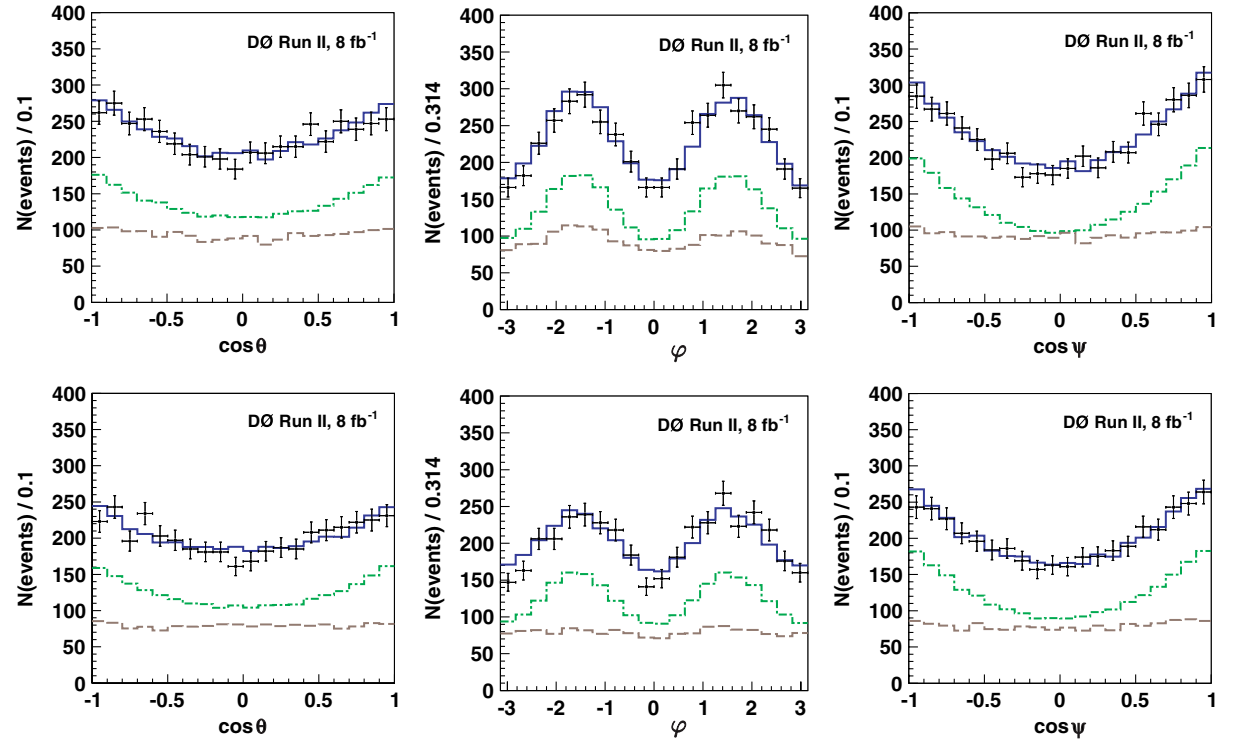


FIG. 10 (color online). Distributions of transversity polar and azimuthal angles and $\cos\psi$ for B_s^0 candidates in the BDT sample (top) and Square-cuts sample (bottom). The signal contribution is enhanced, relative to the distributions shown in Fig. 9, by additional requirements on the reconstructed 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).

obtained by random sampling of the resolution function.

(iii) *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 Appendix B) are different for the two samples. Imperfections in the modeling 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.

(iv) $M(KK)$ *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. (12) 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. (12) 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. (12). We repeat the fits using this altered $\phi(1020)$ propagator as a measure of the sensitivity to the $M(KK)$ resolution.

Tables III and IV 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.

VII. 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.

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 multi-dimensional space, taking into account non-Gaussian parameter distributions and correlations, we use the MCMC technique.

A. The method

The MCMC technique uses the Metropolis-Hastings algorithm [45] 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

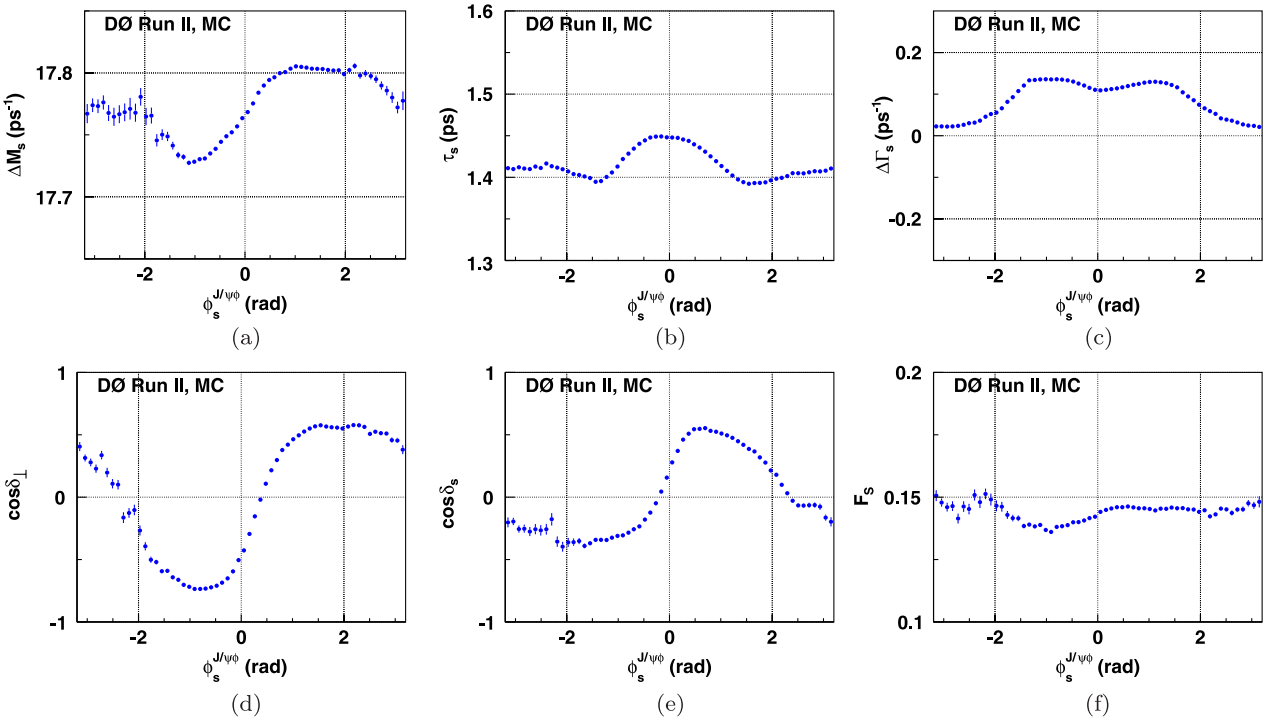


FIG. 11 (color online). Profiles of ΔM_s, $\bar{\tau}_s$, $\Delta\Gamma_s$, $\cos\delta_\perp$, $\cos\delta_s$, and F_S, for $\Delta\Gamma_s > 0$, versus $\phi_s^{J/\psi\phi}$ from the MCMC simulation for the BDT selection data sample.

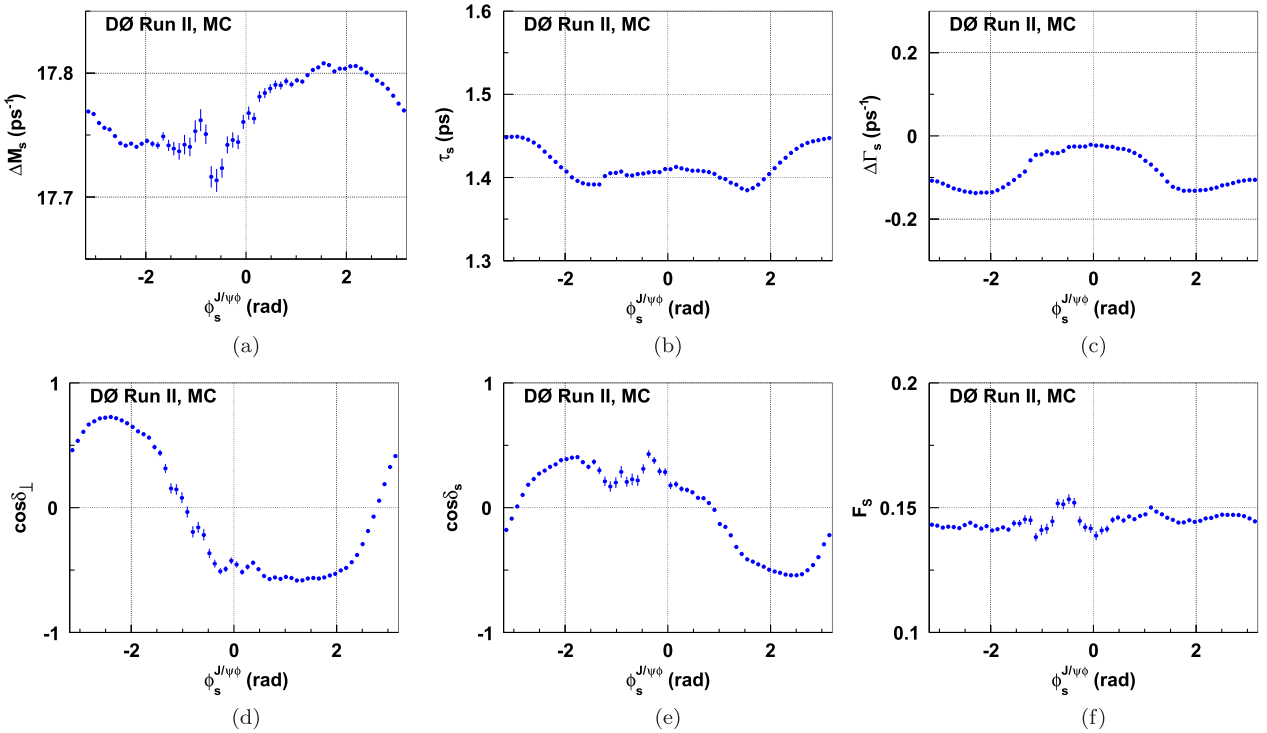


FIG. 12 (color online). Profiles of ΔM_s, $\bar{\tau}_s$, $\Delta\Gamma_s$, $\cos\delta_\perp$, $\cos\delta_s$, and F_S, for $\Delta\Gamma_s < 0$, versus $\phi_s^{J/\psi\phi}$ from the MCMC simulation for the BDT selection data sample.

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.

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

We generate 8 MCMC chains, each containing 1×10^6 states: a nominal and three alternative chains each for the BDT-selection and Square-cuts samples, according to the fit results presented in Tables III and IV.

Figures 11 and 12 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 (see also Fig. 26 in Appendix D)

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

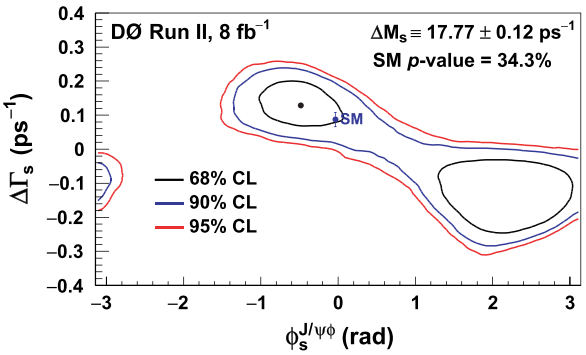
(c) For $\phi_s^{J/\psi\phi}$ near zero, $|\Delta\Gamma_s|$ increases with $|\phi_s^{J/\psi\phi}|$.

(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 $+0.8$. 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.

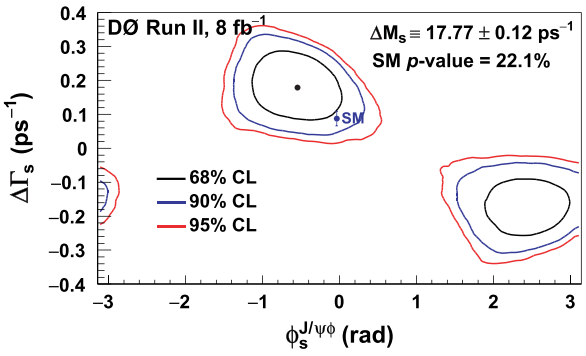
(e) A strong positive correlation between $\phi_s^{J/\psi\phi}$ and $\cos\delta_s$ near $\phi_s^{J/\psi\phi} = 0$, with $\phi_s^{J/\psi\phi}$ changing sign as the average $\cos\delta_s$ increases between -0.4 and $+0.4$.

(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 $B_d^0 \to J/\psi K^*$ [34], up to the sign ambiguities. Reference [46] 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 $\cos\delta_\perp < 0$.



(a)BDT selection



(b)Square-cuts selection

FIG. 13 (color online). Two-dimensional 68%, 90%, 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.

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

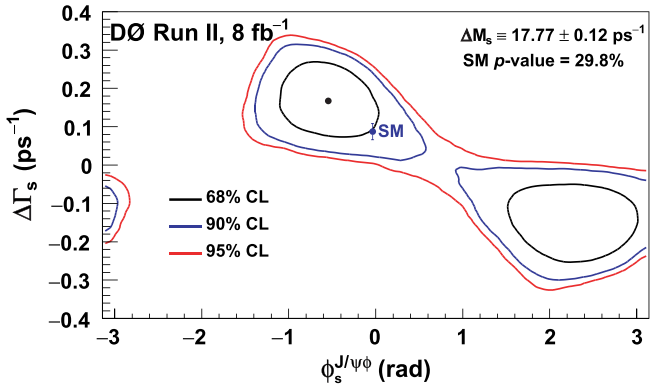


FIG. 14 (color online). Two-dimensional 68%, 90%, and 95% credible regions including systematic uncertainties. The standard model expectation is indicated as a point with an error.

sample. We then combine all eight chains, to produce the final result.

C. Results

Figure 13 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.

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(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:

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

F_S(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 nonresonant $B_s^0 \to J/\psi K^+K^-$.

TABLE V. Variables used to train the prompt BDT, ranked by their importance in the training.

Rank	Variable	Importance	Separation
1	KK invariant mass	0.3655	0.3540
2	Maximum ΔR between either K meson and the B_s^0 candidate	0.1346	0.4863
3	Isolation using the maximum ΔR between either K and the B_s^0	0.0390	0.1784
4	Uncorrected p_T of the B_s^0	0.0346	0.3626
5	Minimum ΔR between either K and the B_s^0	0.0335	0.4278
6	p_T of the trailing K meson	0.0331	0.4854
7	p_T of the ϕ meson	0.0314	0.4998
8	p_T of the leading K meson	0.0283	0.4884
9	Trailing muon momentum	0.0252	0.0809
10	p_T of the leading muon	0.0240	0.1601
11	Maximum ΔR between either muon and the B_s^0	0.0223	0.1109
12	Maximum χ^2 of either K meson with the J/ψ vertex	0.0217	0.0162
13	Dimuon invariant mass	0.0215	0.0145
14	Maximum χ^2 of either of the K candidate track	0.0213	0.021
15	B_s^0 isolation using the larger K/B_s ΔR and tracks from the PV	0.0207	0.1739
16	p_T of the J/ψ meson	0.0205	0.1809
17	Minimum ΔR between either muon and the B_s^0 candidate	0.0188	0.1023
18	Trailing K momentum	0.0105	0.3159
19	χ^2 of the B_s^0 candidate vertex	0.0093	0.0119
20	B_s^0 isolation using $\Delta R < 0.75$	0.0084	0.0241
21	Minimum χ^2 of the J/ψ vertex with either K	0.0081	0.0069
22	p_T of the trailing muon	0.0079	0.0922
23	Minimum of the χ^2 of the J/ψ and ϕ vertices	0.0073	0.0057
24	Isolation using $\Delta R < 0.5$	0.0070	0.0405
25	Uncorrected B_s^0 total momentum	0.0068	0.2103
26	Minimum χ^2 of either K track fit	0.0065	0.0266
27	Isolation using $\Delta R < 0.5$ and particles from the PV	0.0057	0.0401
28	Leading K meson momentum	0.0051	0.3217
29	Leading muon momentum	0.0048	0.0908
30	ϕ meson momentum	0.0048	0.3233
31	Maximum χ^2 of the J/ψ or ϕ vertices	0.0044	0.0061
32	Isolation using $\Delta R < 0.75$ and particles from the PV	0.0037	0.0259
33	J/ψ meson momentum	0.0037	0.1004

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

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

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 14 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 [47] ($\phi_s^{J/\psi\phi}$, $\Delta\Gamma_s$) = (-0.038, $0.087\ \text{ps}^{-1}$) is 29.8%. The one-dimensional 68% credible intervals are listed in Sec. VIII below.

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

TABLE VI. Variables used to train the non-prompt BDT, ranked by their importance in the training.

Rank	Variable	Importance	Separation
1	KK invariant mass	0.2863	0.3603
2	B_s^0 isolation using the larger K/B_s ΔR and tracks from the PV	0.1742	0.4511
3	Minimum dE/dx of either K	0.0778	0.1076
4	χ^2 of B_s^0	0.0757	0.2123
5	p_T of the ϕ meson	0.0559	0.4856
6	p_T of the leading K meson	0.0504	0.4745
7	Isolation using the maximum ΔR between either K and the B_s^0	0.0429	0.4468
8	p_T of the trailing K meson	0.0350	0.4774
9	Maximum χ^2 of either K meson with the J/ψ vertex	0.0260	0.2051
10	Isolation using $\Delta R < 0.5$ and particles from the PV	0.0229	0.1703
11	Isolation using $\Delta R < 0.75$ and tracks from the PV	0.0154	0.2238
12	Minimum χ^2 of of either K with the J/ψ vertex	0.0151	0.1308
13	Minimum ΔR between either K meson and the B_s^0 candidate	0.0115	0.3104
14	Dimuon invariant mass	0.0099	0.0190
15	Total momentum of the ϕ meson	0.0091	0.3307
16	p_T of the J/ψ meson	0.0089	0.1198
17	Trailing muon momentum	0.0082	0.0594
18	Isolation using $\Delta R < 0.5$	0.0073	0.1695
19	Maximum ΔR between either K meson and the B_s^0 candidate	0.0070	0.3794
20	Maximum dE/dx of either K meson	0.0069	0.0528
21	Trailing K meson momentum	0.0068	0.3253
22	J/ψ vertex χ^2	0.0063	0.0057
23	Leading K meson momentum	0.0058	0.3277
24	Maximum χ^2 of either K candidate track	0.0054	0.0267
25	Isolation using $\Delta R < 0.75$	0.0046	0.2203
26	Minimum ΔR between either muon and the B_s^0 candidate	0.0041	0.0729
27	Minimum χ^2 of either K candidate track	0.0039	0.0284
28	uncorrected p_T of B_s^0 candidate	0.0036	0.2485
29	p_T of the trailing muon	0.0029	0.0702
30	J/ψ momentum	0.0027	0.0645
31	Maximum ΔR between either muon and the B_s^0 candidate	0.0026	0.0872
32	Vertex χ^2 of the ϕ meson	0.0017	0.0098
33	Uncorrected B_s^0 momentum	0.0014	0.1675
34	p_T of the leading muon	0.0011	0.1008
35	Leading muon momentum	0.0009	0.0547

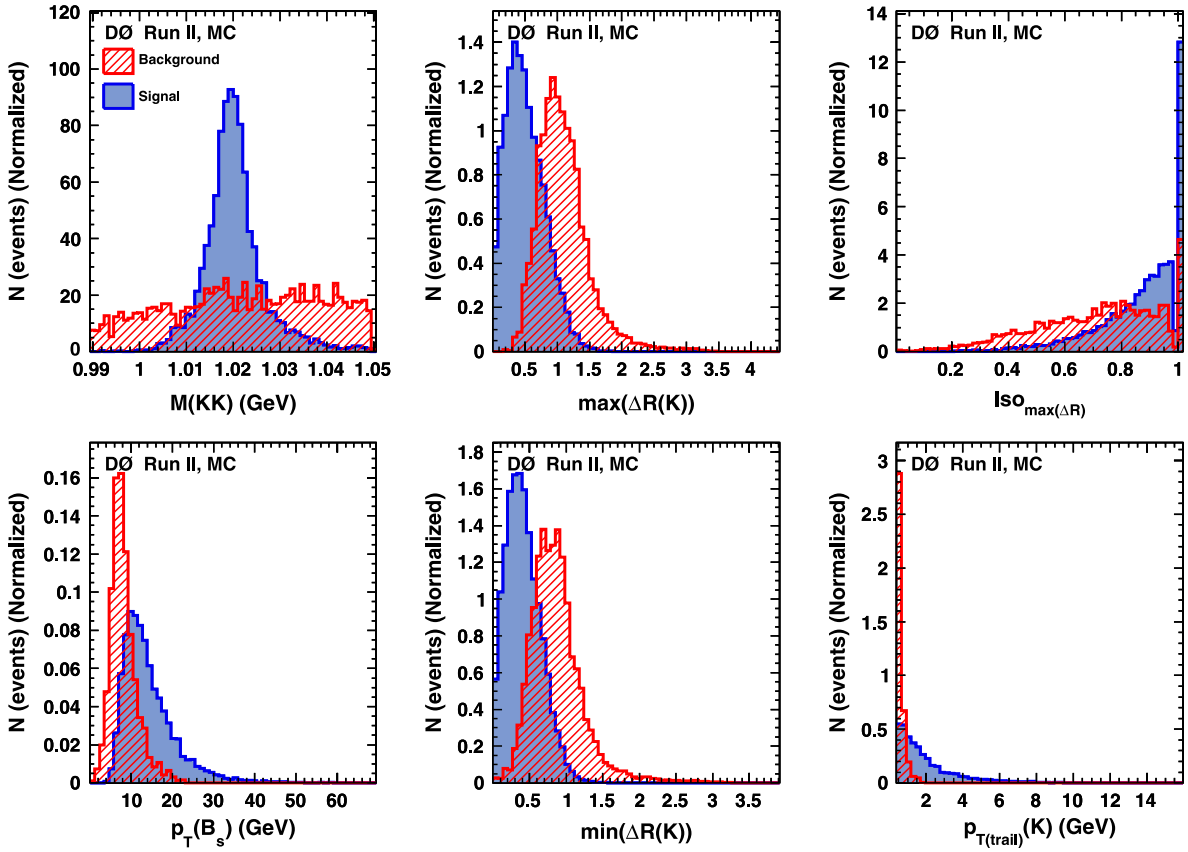


FIG. 15 (color online). The distributions of the six most important variables used in the BDT trained on prompt J/ψ production for the $B_s^0 \to J/\psi\phi$ signal (solid blue) and prompt J/ψ events (red dashed) histograms.

$$\begin{aligned} \bar{\tau}_s &= 1.443^{+0.038}_{-0.035}\ \text{ps}, \quad \Delta\Gamma_s = 0.163^{+0.065}_{-0.064}\ \text{ps}^{-1}, \\ \phi_s^{J/\psi\phi} &= -0.55^{+0.38}_{-0.36}, \quad |A_0|^2 = 0.558^{+0.017}_{-0.019}, \\ |A_\parallel|^2 &= 0.231^{+0.024}_{-0.030}, \quad \delta_\parallel = 3.15 \pm 0.22, \\ \cos(\delta_\perp - \delta_s) &= -0.11^{+0.27}_{-0.25}, \quad F_S = 0.173 \pm 0.036, \end{aligned} \tag{13}$$

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 the previous publication [26], which was based on a subset of this data sample, we 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 nonresonant K^+K^- in the final state based on the $M(KK)$ distribution of the B_s^0 signal yield. The result of this study (Appendix C) is consistent with the result of the maximum-likelihood fit shown above.

ACKNOWLEDGMENTS

We thank the staffs at Fermilab and collaborating institutions, and acknowledge support from the DOE and NSF (USA); CEA and CNRS/IN2P3 (France); FASI, Rosatom and RFBR (Russia); CNPq, FAPERJ, FAPESP and FUNDUNESP (Brazil); DAE and DST (India); Colciencias (Colombia); CONACyT (Mexico); KRF and KOSEF (Korea); CONICET and UBACyT (Argentina); FOM (The Netherlands); STFC and the Royal Society (United Kingdom); MSMT and GACR (Czech Republic); CRC Program and NSERC (Canada); BMBF and DFG (Germany); SFI (Ireland); The Swedish Research Council (Sweden); and CAS and CNSF (China). We thank J. Boudreau who has suggested and developed the use of the MCMC method for this study.

APPENDIX A: BDT DISCRIMINANTS

Two BDT discriminants are used to reject background. One is trained to remove the prompt background (the

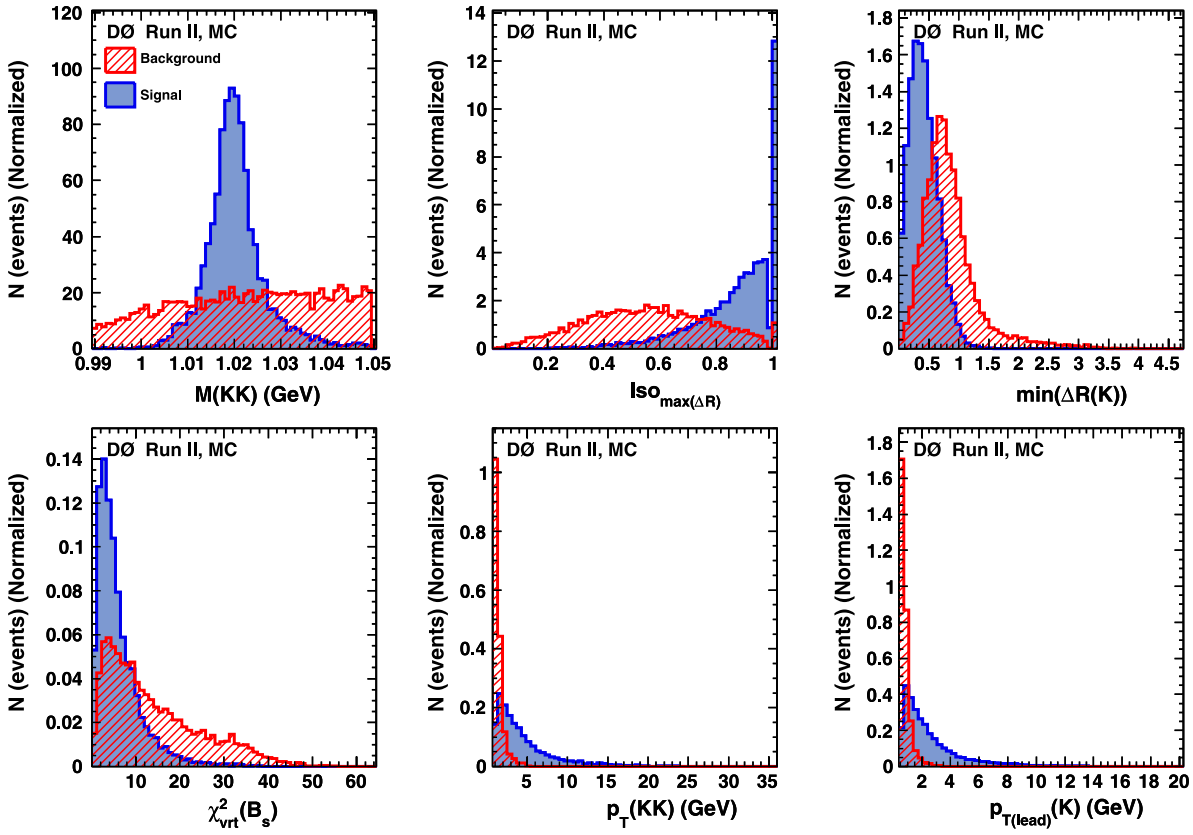


FIG. 16 (color online). The distributions of the six most important variables used in the BDT trained on inclusive $B \to J/\psi X$ decays for the $B_s^0 \to J/\psi\phi$ signal (solid blue) and inclusive $B \to J/\psi X$ decays (red dashed) histograms.

"prompt BDT"), and the other is trained to remove inclusive B decays (the "inclusive BDT"). The prompt BDT uses 33 variables, listed in Table V. The inclusive BDT uses 35 variables, listed in Table VI. 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

$$\langle S^2 \rangle = \frac{1}{2} \int \frac{(\hat{y}_S(y) - \hat{y}_B(y))^2}{\hat{y}_S(y) + \hat{y}_B(y)} dy, \qquad \text{(A1)}$$

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 Fig. 15. The distributions for the six most important variables in the training on inclusive $B \to J/\psi X$ decays are shown in Fig. 16.

Figure 17 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.

APPENDIX B: DETECTOR ACCEPTANCE

We take into account the shaping of the signal distribution by the detector acceptance and kinematic selection by

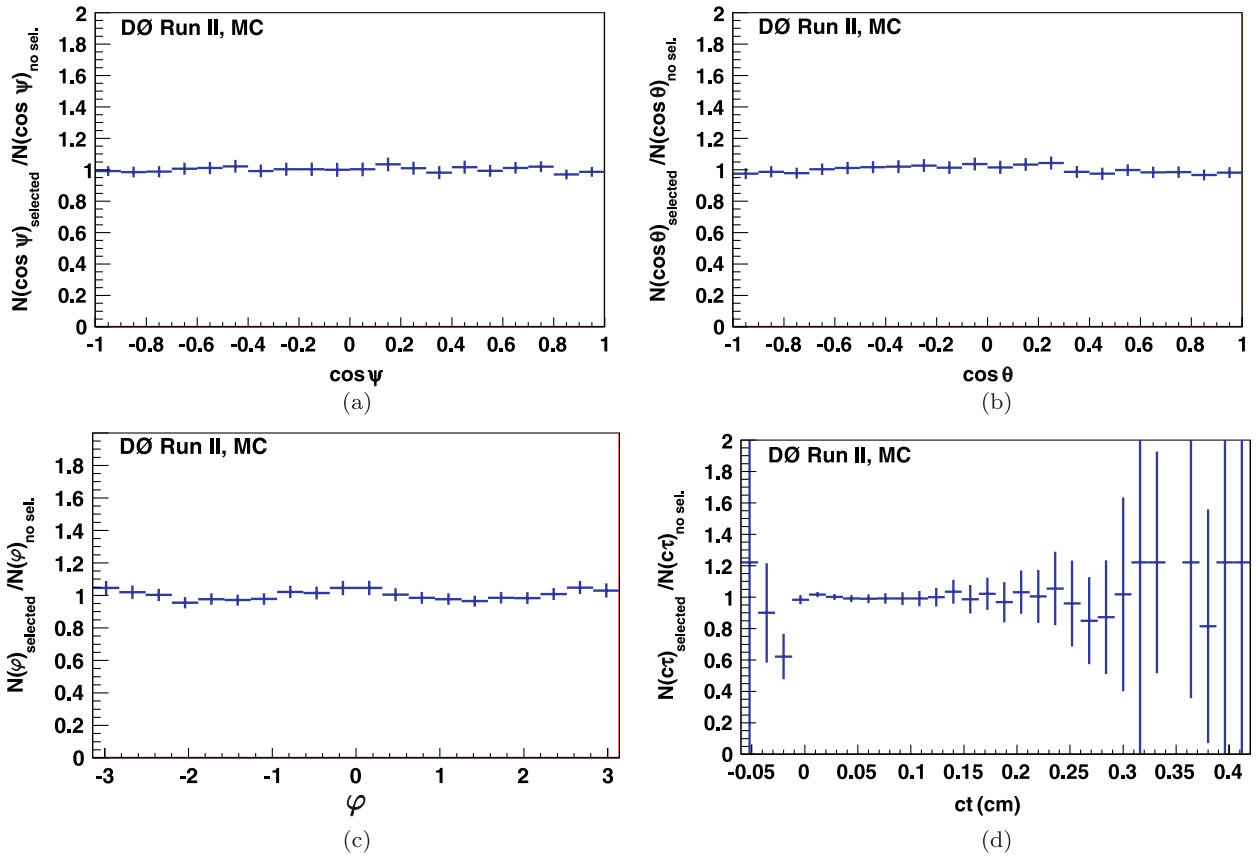


FIG. 17 (color online). Test of uniformity of the efficiencies of the BDT selection using a MC sample with $\phi_s = -0.5$. The figure shows the ratios of the normalized distributions of (a–c) the three angles and (d) the proper decay length, before and after the BDT selection.

introducing acceptance functions in the three angles of the transversity basis. The acceptance functions are derived from Monte Carlo simulation. Because of 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 Fig. 18.

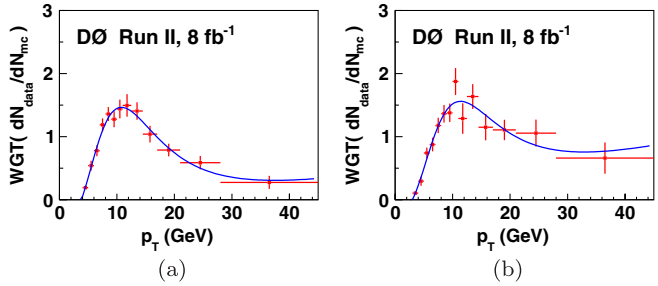
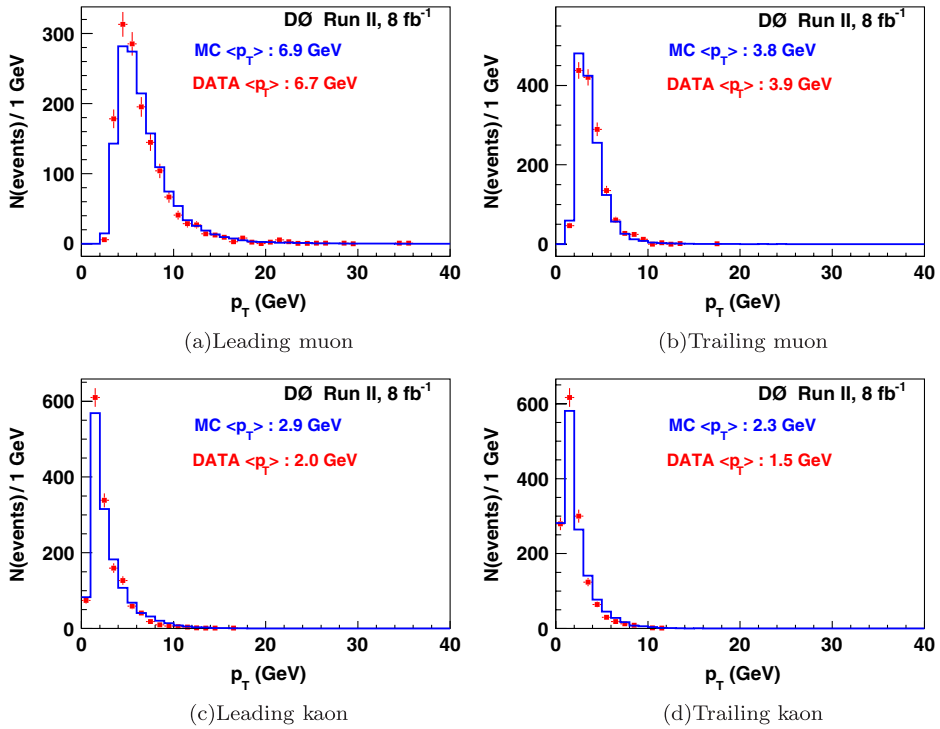


FIG. 18 (color online). Weight factor as a function of $p_T(J/\psi)$ used to correct MC p_T distribution of B_s^0 and B_d^0 decay objects for (a) central region, and (b) forward region. The curves are empirical fits to a sum of a Landau function and a polynomial.

Figure 19 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 is shown in Fig. 20. The acceptance in ψ is shown in Fig. 21.

APPENDIX C: 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



(a)Leading muon (b)Trailing muon

(c)Leading kaon (d)Trailing kaon

FIG. 19 (color online). Transverse momentum distributions of the four final-state particles in data (points) and weighted MC (solid histogram), for the BDT-based event selection.

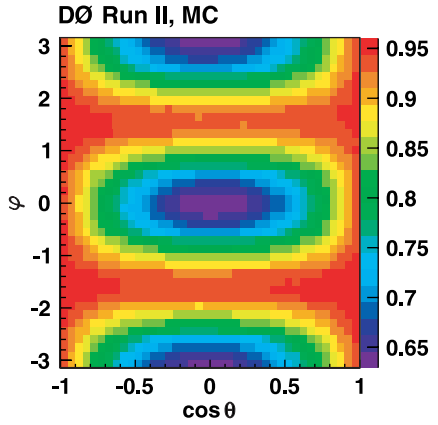


FIG. 20 (color online). Map of the detector acceptance on the plane $\varphi - \cos\theta$.

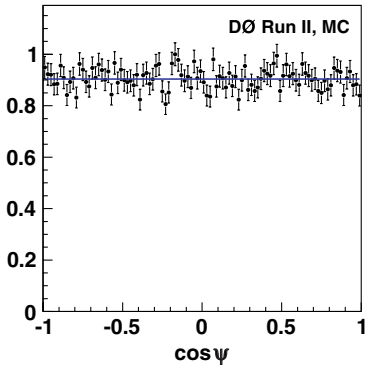


FIG. 21 (color online). Detector acceptance as a function of $\cos\psi$. The acceptance is uniform in $\cos\psi$.

the $M(K^+K^-)$ mass information to make an independent estimate of the nonresonant K^+K^- contribution in the final state.

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 Fig. 22, 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.

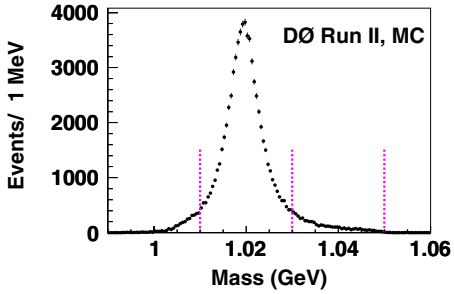


FIG. 22 (color online). Invariant mass distribution of kaon pairs from the full simulation of the decay $\phi \to K^+K^-$. Vertical dashed lines delineate the two $M(KK)$ invariant mass bins considered.

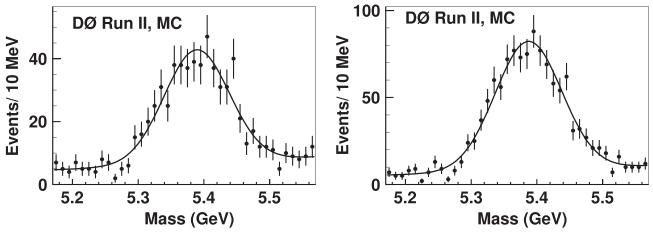


FIG. 23. The simulated distributions of the invariant mass of the $B_d^0 \to J/\psi K^*$ decay products reconstructed under the $B_s^0 \to J/\psi\phi$ hypothesis for $1.01 < M(KK) < 1.03$ GeV (left) and $1.03 < M(KK) < 1.05$ GeV (right). The curves are results of fits assuming a sum of two Gaussian functions.

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 Fig. 23. The mass distributions, with fits using the above templates, are shown in Fig. 24. 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 nonresonant 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 nonresonant $\mathcal{P}$-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 $F_S = 0.12 \pm 0.03$.

APPENDIX D: $B_s^0 - \bar{B}_s^0$ OSCILLATION

Under the hypothesis of CP conservation in the B_s^0 decay, and a possible mixing-induced CP violation, the

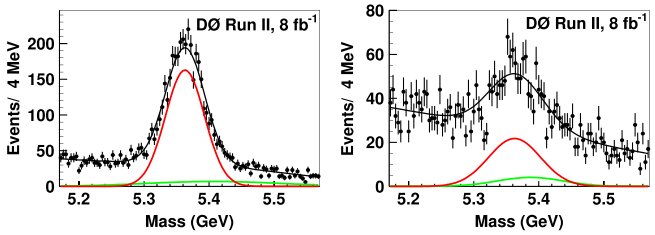


FIG. 24 (color online). Invariant mass distributions of B_s^0 candidates with decay length $ct > 0.02$ cm for $1.01 < M(KK) < 1.03$ GeV (left) and $1.03 < M(KK) < 1.05$ GeV (right). Fits to a sum (black line) of a Gaussian function representing the signal (red), an MC simulation-based template for the $B^0 \to J/\psi K^*$ reflection (green line), and a linear function representing the background are used to extract the B_s^0 yield.

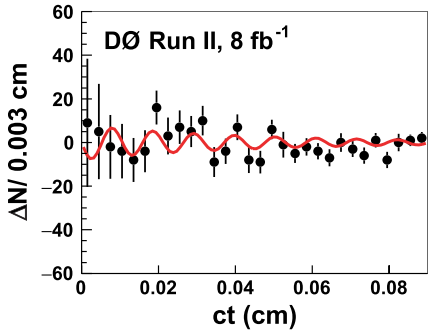


FIG. 25 (color online). Proper decay length evolution of the difference $\Delta N = N(B_s^0) - N(\bar{B}_s^0)$ in the first 0.09 cm (3 ps) for the Square-cuts sample. The curve represents the best fit to the oscillation with the frequency of $\Delta M_s = 17.77$ ps^{-1}.

nonvanishing CP-violating mixing angle should manifest itself as a $B_s^0 - \bar{B}_s^0$ oscillation with the amplitude proportional to $\sin(\phi_s^{J/\psi\phi})$. The observed time-dependent asymmetry $\Delta N \equiv N(B_s^0) - N(\bar{B}_s^0) = N_S \cdot C \cdot \sin(\phi_s^{J/\psi\phi})$, is diluted by a product C of several factors: (i) a factor of $(1 - 2|A_\perp|^2) \cdot (1 - 2F_s) \approx 0.6 \cdot 0.7$ due to the presence of the CP-odd decay, (ii) a factor of $\epsilon \cdot \mathcal{D}^2 \approx 0.03$ due to the flavor-tagging efficiency and accuracy, and (iii) a factor of $\exp(-(\Delta M_s \sigma)^2/2) \approx 0.2$ due to the limited time resolution. Thus, with $N_S \approx 6000$ events, and $C \approx 0.0025$, we expect $N_S \cdot C \approx 15$.

In Fig. 25 we show the proper decay length evolution of ΔN in the first 90 μm, corresponding to approximately twice the mean B_s^0 lifetime. The curve represents a fit to the function $N_0 \cdot \sin(\Delta M_s t) \cdot \exp(-t/\tau_s)$, with N_0 unconstrained and with $\Delta M_s \equiv 17.77$ ps^{-1}. The fit gives $N_0 = -6$ for the BDT-based sample and -8 for the Square-cuts sample, with a statistical uncertainty of ± 4, corresponding to $\sin(\phi_s^{J/\psi\phi}) = N_0/N_S \cdot C \approx -0.4 \pm 0.3$. This one-dimensional analysis gives a result for $\phi_s^{J/\psi\phi}$ that is consistent with the result of the full analysis.

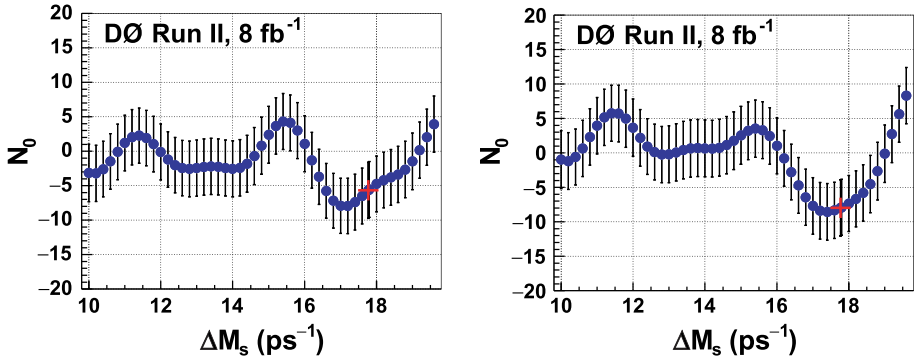


FIG. 26 (color online). The fitted magnitude of the $B_s^0 - \bar{B}_s^0$ oscillation as a function of ΔM_s for (a) BDT selection and (b) Square-cuts. The (red) crosses correspond to $\Delta M_s = 17.77$ ps^{-1}.

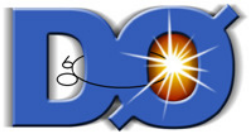
Following the Amplitude Method described in Ref. [48], we fit the above distributions at discrete values of ΔM_s, and plot the fitted value of N_0 as a function of the probe frequency. The results are shown in Fig. 26. There is an undulating structure, with no significantly large deviations from zero. At ΔM_s near 17.77 ps^{-1} the data prefer a negative oscillation amplitude (and hence a negative value of $\sin\phi_s^{J/\psi\phi}$). The statistical uncertainty of the result of this simple approach does not take into account uncertainties of the dilution factors, related to the time resolution, *CP*-odd fraction, and the $\mathcal{S}$-wave fraction.

[1] M. Bona *et al.*, J. High Energy Phys. 10 (2006) 081.
[2] M. Kobayashi and T. Maskawa, Prog. Theor. Phys. **49**, 652 (1973).
[3] M. Kreps *et al.*, arXiv:1103.4962.
[4] J. Drobnak *et al.*, Phys. Lett. B **701**, 234 (2011).
[5] R. M. Wang *et al.*, Phys. Rev. D **83**, 095010 (2011).
[6] A. K. Alok *et al.*, J. High Energy Phys. 11 (2011) 122.
[7] J. Shelton and K. M. Zurek, Phys. Rev. D **83**, 091701 (2011).
[8] Z. J. Ajaltouni and E. Di Salvo, J. Phys. G **37**, 125001 (2010).
[9] S. Nandi and A. Soni, Phys. Rev. D **83**, 114510 (2011).
[10] A. Datta *et al.*, Phys. Rev. D **83**, 094501 (2011).
[11] J. Girrbach *et al.*, J. High Energy Phys. 06 (2011) 044.
[12] A. J. Buras *et al.*, J. High Energy Phys. 05 (2011) 005.
[13] Z. Ligeti *et al.*, Phys. Rev. Lett. **105**, 131601 (2010).
[14] A. J. Buras *et al.*, Phys. Lett. B **694**, 402 (2011).
[15] B. A. Dobrescu, P. J. Fox, and A. Martin, Phys. Rev. Lett. **105**, 041801 (2010).
[16] J. P. Saha, B. Misra, and A. Kundu, Phys. Rev. D **81**, 095011 (2010).
[17] A. J. Buras *et al.*, J. High Energy Phys. 10 (2010) 106.
[18] Y. Bai and A. E. Nelson, Phys. Rev. D **82**, 114027 (2010).
[19] G. Isidori, Y. Nir, and G. Perez, Annu. Rev. Nucl. Part. Sci. **60**, 355 (2010).
[20] A. Soni *et al.*, Phys. Rev. D **82**, 033009 (2010).
[21] L. L. Everett *et al.*, Phys. Rev. D **82**, 094924 (2010).
[22] F. J. Botella, G. C. Branco, and M. N. Rebelo, Phys. Lett. B **687**, 194 (2010).
[23] C. W. Chiang *et al.*, J. High Energy Phys. 04 (2010) 031.
[24] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **98**, 121801 (2007).
[25] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **100**, 121803 (2008).
[26] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **101**, 241801 (2008).
[27] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **100**, 161802 (2008).
[28] T. Aaltonen *et al.* (CDF Collaboration), arXiv:1112.1726 [Phys. Rev. D (to be published)].
[29] R. Aaji *et al.* (LHCb Collaboration), arXiv:1112.3183 [Phys. Rev. Lett. (to be published)].
[30] V. M. Abazov *et al.* (D0 Collaboration), Nucl. Instrum. Methods Phys. Res. Sect. A **565**, 463 (2006).
[31] S. Stone and L. Zhang, Phys. Rev. D **79**, 074024 (2009).
[32] R. Angstadt *et al.*, Nucl. Instrum. Methods Phys. Res. Sect. A **622**, 298 (2010).
[33] V. M. Abazov *et al.*, Nucl. Instrum. Methods Phys. Res. Sect. A **552**, 372 (2005).
[34] K. Nakamura *et al.* (Particle Data Group), J. Phys. G **37**, 075021 (2010).
[35] H. U. Bengtsson and T. Sjöstrand, J. High Energy Phys. 05 (2006) 026.
[36] D. J. Lange, Nucl. Instrum. Meth. Sect. A **462**, 152 (2001).
[37] R. Brun and F. Carminati, CERN Program Library Long Writeup No. W5013, 1993.

[38] A. Hoecker *et al.*, arXiv:physics/0703039, http://pos.sissa.it/archive/conferences/050/040/ACAT_040.pdf.
[39] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. D **74**, 112002 (2006).
[40] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **97**, 021802 (2006).
[41] F. Azfar *et al.*, J. High Energy Phys. 11 (2010) 158.
[42] W. Verkerke and D. Kirkby, *Statistical Problems in Particle Physics, Astrophysics and Cosmology, Proceedings of PHYSTAT05*, edited by L. Lyons and M. K. Unel (Imperial College Press, London, 2006) p. 186, http://www.icpress.co.uk/physics/p446.html.
[43] A. S. Dighe, I. Dunietz, and R. Fleischer, Eur. Phys. J. C **6**, 647 (1999).
[44] A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. Lett. **97**, 242003 (2006).
[45] W. K. Hastings, Biometrika **57**, 97 (1970).
[46] M. Gronau and J. L. Rosner, Phys. Lett. B **669**, 321 (2008).
[47] A. Lenz and U. Nierste, arXiv:1102.4274.
[48] H. G. Moser and A. Roussarie, Nucl. Instrum. Methods Phys. Res. Sect. A **384**, 491 (1997).

E. PUBLISHED PAPER

F. DØ internal note



DØ Note 6306
Version 1.0

Measurement of B_s lifetime using the semileptonic channel $B_s \to D_s^+ \mu^- X$ with 8.0 fb^{-1} of data collected with the DØ detector

Jorge Martínez-Ortega and Alberto Sánchez-Hernández
CINVESTAV-IPN México
(Dated: February 29, 2012)

An update measurement of the B_s^0 lifetime is presented using 8.0 fb^{-1} of data collected by the DØ detector during Run II of the Tevatron collider at the Fermi National Accelerator Laboratory from 2002 to 2010. The Tevatron collides protons with anti-protons at energy of center of mass of 1.96 TeV. It is used a sample of semi-leptonic decays $B_s^0 \to D_s^+ \mu^- X$, where the B_s^0 can not be completely reconstructed. A K factor is introduced to take account of this issue. A Maximum Likelihood fit is performed to measure the flavor-specific lifetime. The obtained central value of the flavor-specific lifetime is or $\tau_{B_s} = 1.448$ ps $\pm$ 0.014 ps(stat.) $\pm$ 0.023 ps(syst.), this measurement is dominated by the systematic uncertainty.

I. INTRODUCTION

This note describes a measurement of the B_s^0 lifetime in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$TeV with the DØ detector at Fermilab. The sample is composed by several partial semi-leptonic decays of the form $B_s^0 \to D_s^+ \mu X$. To estimate the number of D_s^+ candidates a Maximum Likelihood fit is performed. The sample yields more than 35000 D_s^+ candidates.

In the semi-leptonic decays there is an amount of momentum carried by neutral particles that can not be measured with the DØ detector. To properly account for this momentum imbalance a K factor, computed from Monte Carlo simulation samples ,is introduced. In addition to the B_s^0 decays that form the signal, there are other meson decays that contribute to the D_s^+ mass peak. The Monte Carlo samples are also used to estimate the fractions of the different decays that contribute to the D_s^+ candidates.

FInally, a Maximum Likelihood fit over the pseudo-proper decay length distribution is performed to estimate the B_s^0 lifetime. The fit is tested measuring the B_d^0 lifetime. The systematic uncertainties are due to the sample composition, the combinatoric background modeling, the resolution modeling and the K factor determination.

II. MOTIVATION

The B_s system is a combination of two states with two different masses, M_L and M_H, and two different decay widths, Γ_L and Γ_H. If the initial state is not tagged and neglecting CP-violation, the decay rate to a final mode f can be written as

$$P_{B_s^0\bar{B}_s^0\to f}(t) = Ae^{-\Gamma_L t} + Be^{-\Gamma_H t}, \tag{1}$$

where A and B are the amplitudes of the states B_L and B_H respectively. Theses states are almost CP eigenstates. So, if the lifetime is measured experimentally with only one exponential, then there is a dependence of the lifetime on the composition of the final mode f [1],

$$\tau_{B_s^0\to f} = \frac{A\frac{1}{\Gamma_L^2} + B\frac{1}{\Gamma_H^2}}{A\frac{1}{\Gamma_L} + B\frac{1}{\Gamma_H}}. \tag{2}$$

In the case of semi-leptonic modes $A = B$, so the result simplifies to

$$\tau_{B_s^0\to SL} = \frac{1}{\Gamma}\frac{1+(\frac{\Delta\Gamma}{2\Gamma})^2}{1-(\frac{\Delta\Gamma}{2\Gamma})^2}, \tag{3}$$

where $\Gamma = (\Gamma_L + \Gamma_H)/2$ and $\Delta\Gamma = \Gamma_L - \Gamma_H$. So, should be noted that $\tau_{B_s^0\to SL}$ is always larger than $\tau_{B_s^0} = 1/\Gamma$.

On the other hand, in the framework of the Heavy Quark Expansion Theory (HQET) the decay rate of the b-mesons is calculated in a expansion of the heavy b-quark mass:

$$\frac{1}{\tau} = \Gamma = \Gamma_0 + \frac{\Lambda^2}{m_b^2}\Gamma_2 + \frac{\Lambda^3}{m_b^3}\Gamma_3 + \ldots, \tag{4}$$

where Γ_0 is interpreted as the decay rate of a free b-quark. The lifetime ratio of two given heavy mesons is read as:

$$\frac{\tau_1}{\tau_2} = 1 + \frac{\lambda^3}{m_b^3}(\Gamma_3^0 + \Gamma_3^{(1)} + \ldots) + \frac{\Lambda^4}{m_b^4}(\Gamma_4 + \ldots) + \ldots. \tag{5}$$

In the case of $\tau_{B_s^0}$ and $\tau_{B_d^0}$, according to the results of the quenched lattice-QCD calculation [2, 3], the ratio is restricted to

$$-0.0004 < \frac{\tau_{B_s^0}}{\tau_{B_d^0}} - 1 < 0. \tag{6}$$

So, if we use equation 3 and inequality 6 we can conclude that $\tau_{B_s^0 SL}$ should always be almost equal or larger than $\tau_{B_d^0}$. So, a high precision measurement of the B_s^0 lifetime, can be used to test these models. From this point, otherwise is explicitly stated, $\tau_{B_s}^0$ will refer to the flavor-specific lifetime measured with semi-leptonic decays.

The lifetime of the meson B_d^0 is known experimentally with a precision of better than 1% [4]. The world average reported by the Particle Data Group [5] is $\tau_{B_d^0} = 1.519 \pm 0.007$ps .On the experimental part of B_s^0 lifetime it is reported that the flavor specific measurement lifetime is $\tau_{B_s^0} = 1.417 \pm 0.042$ps. With this two values, the lifetimes ratio gives

$$\frac{\tau_{B_s^0}}{\tau_{B_d^0}} = 0.933 \pm 0.028. \tag{7}$$

III. THE D0 DETECTOR

The D0 detector consists of a magnetic central-tracking system, two pre shower detectors, three liquid-argon/uranium calorimeters, a muon system and a luminosity detector.

The central-tracking system is comprised of a silicon micro-strip tracker and a central fiber tracker, both located within a 1.9 T superconducting solenoidal magnet. The SMT has approximately 800000 individual strips, with typical pitch of 50 to 80 μm, and designed to optimize the capabilities for track and vertex reconstruction at absolute $\eta < 2.5$, where $\eta = \ln[\tan(\theta/2)]$, and θ is the polar angle with respect to the proton beam direction. The system has a six-barrel

longitudinal structure, each with a set of four layers arranged axially around the beam pipe, and interspersed with 14 radial disks. The innermost silicon detector layer is known as layer 0, was introduced in 2006 and surrounds the beryllium beam pipe with single-sided sensors staggered at radii 1.6 and 1.76 cm.

The CFT has eight thin coaxial barrels, each supporting two doublets of overlapping scintillating fibers of 835μm in diameter, one doublet being parallel to the collision axis, and the other alternating by $\pm 3ø$ relative to the axis. Light signals are transferred via clear fibers to solid-state photon counters that have approximately 80% quantum efficiency.

Central and forward pre-shower detectors located just outside of the superconducting coil (in front of the calorimetry) are constructed of several layers of extruded triangular scintillator strips that are read out using wavelength-shifting fibers and VLPCs.

The next layer of detection involves three liquid-argon/uranium calorimeters: a central section covering absolute η up to approximately 1.1, and two end calorimeters that extend the coverage up to 4.2, all housed in separate cryostats. In addition to the pre-shower detectors, scintillators between the central calorimeter and end calorimeter cryostats provide sampling of developing showers at absolute $1.1 < |\eta| < 1.4$.

A muon system resides beyond the calorimetry, and consists of a layer of tracking detectors and scintillation trigger counters before 1.8 T toroids, followed by two similar layers after the toroids. Tracking at absolute $\eta < 1$ relies on 10 cm wide drift tubes, while 1 cm mini-drift tubes are used beyond $\eta = 1$ and up to 2.

Luminosity is measured using plastic scintillator arrays located in front of the end calorimeter cryostats, covering $2.7 < |\eta| < 4.4$.

Trigger and data acquisition systems are designed to accommodate the high luminosities of Run II. Based on preliminary information from tracking, calorimetry, and muon systems, the output of the first level of the trigger is used to limit the rate for accepted events to approximately 2 kHz.

At the next trigger stage, with more refined information, the rate is reduced further to approximately 1 kHz. These first two levels of triggering rely mainly on hardware and firmware. The third and final level of the trigger, with access to all the event information, uses software algorithms and a computing farm, and reduces the output rate to approximately 100 to 150 Hz, which is written to tape.

IV. DATA SELECTION AND RECONSTRUCTION

The update of this measurement is done using 8.0fb^{-1} of data collected with the DØ detector from 2002 to 2010. We take reconstructed events consistent with the decay mode $B_s \to D_s^+ \mu^- X$. We take the muphipi_lifetime-tag B-group skims in `http://www-d0.fnal.gov/~markw/d0_private/bgroup_webpage/SAM_access/skim_defs.html`. To process them we use the AATrack package. This skim was prepared removing all possible requirements, which could induce a bias the lifetime measurement. To select the events, first we look for a muon, then two tracks of opposite charge, that will be assigned as kaons coming from a ϕ, and finally a charged track that will be assigned the pion mass, and will be combined with the ϕ candidate to form a D_s^- candidate. The muon candidate is selected to pass the following cutoffs:

- required to be detected in the layer inside the toroid and at least in one layer outside the toroid,`nseg` $= 3$;
- carrying transverse momentum $p_T > 2.0$ GeV/c,
- carrying total momentum $p > 3.0$ GeV/c,
- having at least 2 hits in the CFT, N(CFT) > 1,
- having at least 2 hits in the SMT, N(SMT) > 1,
- at least one central track matched,
- having $\chi^2 < 25$ in the global muon fit.

Each of the two oppositely charged tracks consistent with coming from the ϕ, identifying them as kaons, must pass the following cuts:

- the invariant mass of the combined tracks must lay between, 1.008 GeV/c^2 $< M(K^+K^-) <$ 1.032 GeV/c^2,
- carrying transverse momentum $p_T > 1.0$ GeV/c,
- having at least 2 hits in the CFT, N(CFT) > 1,

- having at least 2 hits in the SMT, N(SMT) > 1,

The additional charged track, the *pi* candidate, must pass:

- carrying transverse momentum $p_T > 0.7$ GeV/c,
- having at least 2 hits in the CFT, N(CFT) > 1,
- having at least 2 hits in the SMT, N(SMT) > 1,
- the combination of this track with the ϕ candidate must fulfill 1.6 GeV/c^2 < $M(\phi\pi)$ < 2.3 GeV/c^2,
- the χ^2 of the vertex fit has to be smaller than 16.26,

in addition

- the invariant mass of the combination of the muon candidate with the D_s^- candidate has to be 3 GeV/c^2 < $M(D_s^-\mu)$ < 5 GeV/c^2,
- the χ^2 of the vertex fit hast to be smaller than 15.13

All tracks have to be in the same jet, and be associated with the same primary vertex. At this point we have not made any restriction on the charge of the μ and π, however for the B_s^0 these two tracks have to have opposite charge, but we also keep the wrong combination sample to help modeling the background. In the right sign sample the reconstructed D_s^- is required to be displaced from the primary vertex in the same direction of its momentum in order to get some suppression of combinatoric background.

V. MONTE CARLO SAMPLES

We need to generate simulated events in order to study the different contributions of the all the possible decay channels that can contribute to our data sample and to get the k-factor distributions the are used to correct the pseudo proper decay length computed in our selection program. We have generated MC samples using PYTHIA [6] for the production and hadronization phase, and EvtGen for decaying the b and c hadrons. We generated B_s^0 meson samples with $c\tau = 439$ microns, and no mixing. The signal sample includes contributions from $D_s^-\mu^+\nu$, $D_s^{*-}\mu^+\nu$, $D_{s0}^{*-}\mu^+\nu$, $D_{s1}^{'-}\mu^+\nu$ and $D_s^{(*)-}\tau^+\nu$.

To be able to evaluate non-combinatorial backgrounds, additional processes were generated including $\bar{B}^0 \to D_s^{(*)-}D^{(*)+}$, and $\bar{B}^- \to D_s^{(*)-}D^{(*)0}X$, where the "right-sign" $D_s^-\mu^+$ combination can be obtained by allowing $D^{(*)+/0}$ to decay semileptonically. The $B_s \to D_s^{(*)-}D_s^{(*)+}X$ and $B_s \to D_s^{(*)-}D_s^{(*)+/0}X$ processes were also generated.

To be able to fully simulate these samples some kinematic cuts were applied prior to D0gstar, (using d0_mess filter): muons had to have $p_T > 1.9$ GeV/c and $|\eta| < 2.1$, the kaons (and pions) from $\phi(D_s^-)$ had to have $p_T > 0.6$GeV/c and $|\eta| < 3.0$, and the p_T of the D_s^- had to be greater than 1.0GeV/c. The samples were then processed using the standard full chain procedure D0gstar-D0simD0reco. All samples were generated as private production but following the official production criteria.

The produced thumbnails were then filtered using the same procedure as in data, described in previous section, above.

VI. PSEUDO PROPER DECAY LENGTH AND K-FACTOR

The B_s lifetime, τ, is given in terms of the decay length, L, by the relation,

$$L = c\tau\beta\gamma = c\tau\frac{p}{m}, \tag{8}$$

where $c\tau$ is the proper decay length, m the invariant mass, and p the total linear momentum. Using the projection in the transverse plane, this relation is changed to

$$L_{xy} = c\tau\frac{p_T}{m}, \tag{9}$$

where p_T is the transverse momentum of the B_s, and L_{xy} is the so-called transverse decay length. The transverse decay length is, in general, given by the relation

$$L_{xy} = \frac{\vec{X} \cdot \vec{p}_T}{|\vec{p}_T|}, \tag{10}$$

where $\vec{X}$ is the displacement vector from the primary to the secondary vertex in that transverse plane. In this case, as we use the semileptonic channel, we can not fully reconstruct the B_s^0 momentum, due to this we use the combined momentum of the muon and D_s^- meson, $\vec{p}_T(D_s^{\pm}\mu^{\mp})$ as our best approximation of B_s^0 momentum. In order to model this effect in the fit, a correction factor, K, has to be introduced. The definition of this K factor is given by:

$$K = \frac{p_T(D_s^-\mu)}{p_T(B_s^0)}, \tag{11}$$

and is computed using Monte Carlo events. Taking this into account, we can find the B_s^0 lifetime with.

$$c\tau = K\lambda = \frac{\vec{X} \cdot \vec{p}_T(D_s^-\mu)}{|\vec{p}_T(D_s^-\mu)|}, \tag{12}$$

where λ is called pseudo-proper decay length. The K-factor correction is applied statistically by smearing the exponential decay distribution as described below when extracting the $c\tau(B_s^0)$ from the pseudo-proper decay length in the lifetime fit.

We use the MC samples described in Sec. V, where the B_s^0 has been decayed semileptonically through D_s^-, D_s^*, D_{s0}^*, and $D_{s1}^{'*}$.

VII. K-FACTORS

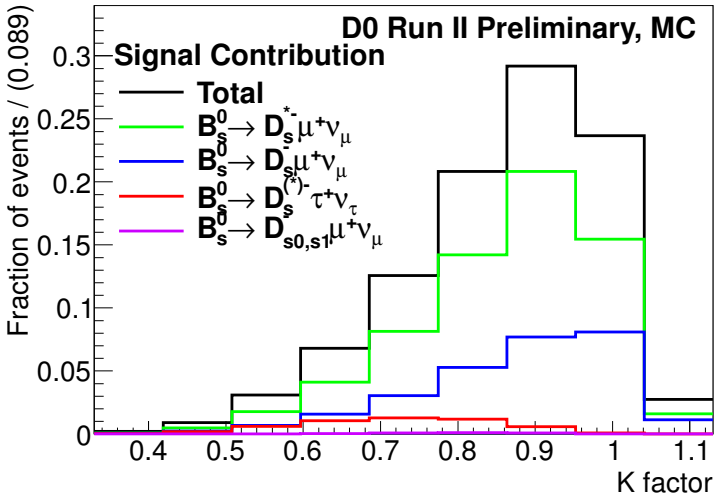


FIG. 1: K-factor distributions of the decays contributing to the signal. All the histograms are normalized considering the contributions from table I. The black histogram shows the total K-factor of the signal.

The decay modes that are considered in our signal, include at least one particle that we cannot detect. As a result is necessary to use a correction factor, the so-called K-factor, to take into account the deficit of momentum corresponding to the undetected neutrino. The K factor is computed using MC samples for those decays that form the D_s^- meson mass peak. The signal decays that contribute to the D_s^- meson peak are listed in table I.

Decay Mode	BR	Contribution to Signal
$D_s^- \mu^+ \nu_\mu$	$(2.20 \pm 0.17)\%$	$(27.5 \pm 2.4)\%$
$D_s^{*-} \mu^+ \nu_\mu$	$(5.27 \pm 0.23)\%$	$(66.1 \pm 4.4)\%$
$D_{s(J)}^{(*)-} \mu^+ \nu_\mu$	$(0.03 \pm 0.42)\%$	$(0.4 \pm 5.3)\%$
$D_s^{(*)-} \tau^+ \nu_\tau \times (\tau^+ \to \mu^+ \bar{\nu}_\mu \nu_\tau)$	$(2.65 \pm 1.28)\% \times (17.36 \pm 0.5)\% = (0.40 \pm 0.22)\%$	$(5.9 \pm 2.7)\%$

TABLE I: Branching fractions of the decays that contribute to the signal. The total branching ratio of all these modes is $(7.90 \pm 0.55)\%$ and they contribute to the $(80.51 \pm 2.11)\%$ of the events in the D_s^+ mass peak.

The reconstruction efficiency times the MC production efficiency of the signal is $\epsilon_{prod}\epsilon_{reco} = 11.9983 \times 10^{-5}$, and we estimate that the signal contribution to the sample is $(80.51 \pm 2.11)\%$. Also we need to compute a K factor distribution for each background contribution similar to our signal that comes from a B-meson and decay into a D_s^- meson. These decays are listed in table II. This table also shows the contribution of each decay to the signal sample. The total contribution of the B-meson decays background is 9.98%.

K factor distributions for signal and background samples are shown in figures 1 and 2 respectively.

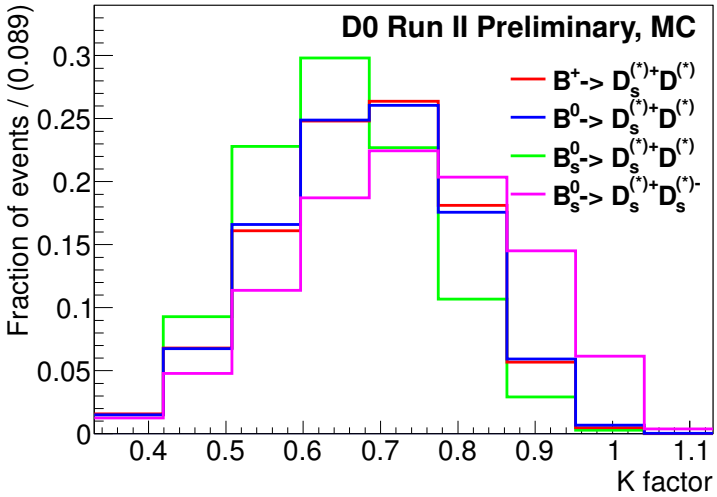


FIG. 2: K-factor distributions of the decays contributing to the B mesons background.

Decay	$\sum \mathrm{Br}(\mathrm{B} \to \mathrm{D_s^0 D} \times \mathrm{D} \to \mu)$	$X_{section}\epsilon_{prod}\epsilon_{reco}(\mu\mathrm{b})$	Contribution to D_s^+
$B^+ \to D_s^+ DX$	$(0.11 \pm 0.02)\%$	0.71 ± 0.03	$(3.81 \pm 0.75)\%$
$B^0 \to D_s^+ DX$	$(0.13 \pm 0.02)\%$	0.64 ± 0.03	$(4.13 \pm 0.70)\%$
$B_s^0 \to D_s^+ D_s^{(*)} X$	$(0.06 \pm 0.01)\%$	0.37 ± 0.02	$(1.11 \pm 0.36)\%$
$B_s^0 \to D_s^+ DX$	$(0.13 \pm 0.06)\%$	0.14 ± 0.01	$(0.92 \pm 0.44)\%$
$c\bar{c} \to D_s^+ \mu^-$	$(0.12 \pm 0.02)\%$	1.56 ± 0.1	$(9.53 \pm 1.65)\%$

TABLE II: Branching fractions of the background decays that contribute to the D_s^- meson mass peak, The B meson decays contribute to the 9.98% to the D_s^+ mass peak.

VIII. SIGNAL FRACTION

In order to estimate the number of candidates of D_s^--mesons in the Signal Sample, a fit over the D_s^- mass distribution. The mass window taken is $(1.958 \pm 10 \times 0.022)\mathrm{GeV/c^2}$ or $1.738\ \mathrm{GeV/c^2} < M(\phi\pi) < 2.178\ \mathrm{GeV/c^2}$, we refer these events as the "D_s^- Mass Sample" **DMS**, the peak of the D^+ meson also lays in this mass range. The model that is used to fit this distributions is comprised of a gaussian distribution for D_s^- mass, a gaussian distribution for

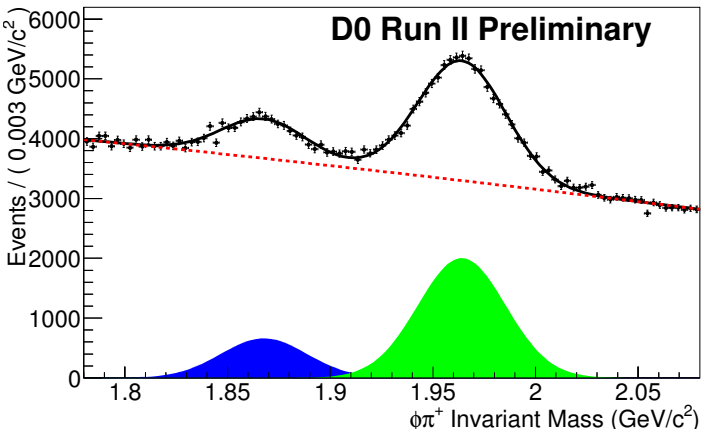


FIG. 3: Invariant mass distribution of the D_s^+ candidates. The curves are projections of the Maximum Likelihood fit for the D_s^+ mass distribution. The green area is the D_s^+ mass peak, the blue one is the D^+ mass peak and the red dashed curve is the combinatorial background distribution. The black curve is the sum of the three contributions.

D^+ mass, and a second order polynomial for the background,

$$\mathcal{L}_{Mass} = \prod_{i \in \mathbf{DMS}} \frac{1}{N_{\mathbf{DMS}}} \left[N(D_s^-) e^{-\frac{(M_i - m(D_s^-))^2}{2\sigma^2(D_s^-)}} + N(D^+) e^{-\frac{(M_i - m(D^+))^2}{2\sigma^2(D^+)}} + N(Bkg)(1 + A_1 M_i + A_2 M_i^2) \right] \quad (13)$$

Parameter	Central Value	Uncertainty ($\pm$)
$N(Bkg)$	497040	104000
A_1	$0.9006(\mathrm{GeV/c^2})^{-1}$	$0.0905(\mathrm{GeV/c^2})^{-1}$
A_2	$0.9089(\mathrm{GeV/c^2})^{-2}$	$0.0234(^{e}\mathrm{V/c^2})^{-2}$
$N(D^+)$	11593	541
$m(D^+)$	$1.8677\mathrm{MeV/c^2}$	$0.0008\mathrm{MeV/c^2}$
$\sigma(D^+)$	$0.021\mathrm{MeV/c^2}$	$0.001\mathrm{MeV/c^2}$
$N(D_s^-)$	35781	476
$m(D_s^-)$	$1.9641\mathrm{MeV/c^2}$	$0.0003\mathrm{MeV/c^2}$
$\sigma(D_s^-)$	$0.0214\mathrm{MeV/c^2}$	$0.0003\mathrm{MeV/c^2}$

TABLE III: Parameter of the mass distribution obtained from Maximum-likelihood fit.

IX. LIFETIME FIT MODEL

To perform the lifetime fit we use to samples, the "Signal Sample" **SS**, and the "Background Sample" **BS**. The signal sample is defined as those events in the D_s^- mass distribution which lay in the rank of $\pm 2\sigma$ from fitted mean mass, this region is defined from 1914MeV to 2002MeV. The number of candidates in this sample is 35781 over both periods, runs IIa and IIb . The Background sample is obtained from sidebands of D_s^- mass distribution,$(-9\sigma, -7\sigma) \cup (7\sigma, 9\sigma)$, and candidates with the "wrong-sign" combination from the interval $(-9\sigma, 9\sigma)$. We assume that the combinatoric background is due to random track combinations, and then, the side band sample events can be used to model the background in the signal sample.

The pseudo-proper decay length distribution obtained from the signal sample is fit using an not binned maximum log-likelihood method. Both B_s^0 lifetime and background shape are determined in a simultaneous fit using the signal and background samples for each period of data. The likelihood function $\mathcal{L}$ for each period is given by:

$$\mathcal{L} = \mathcal{C}_{signal} \prod_{i \in \mathbf{SS}} [f_{signal} \mathcal{F}^i_{signal} + (1 - f_{signal}) \mathcal{F}^i_{bkg}] \prod_{j \in \mathbf{BS}} \mathcal{F}^j_{bkg}, \quad (14)$$

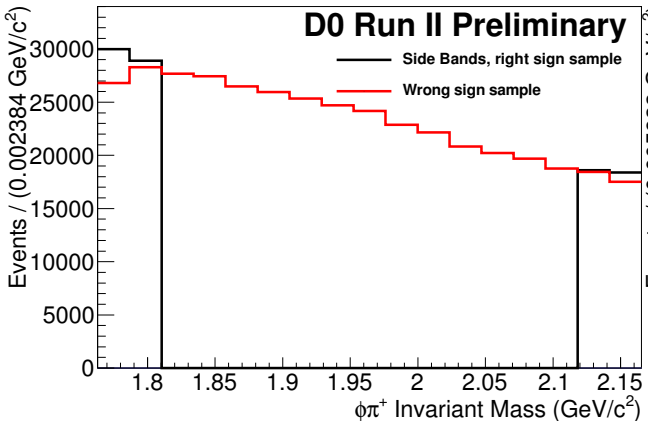


FIG. 4: Mass distributions of Side Band sample (**SB**, black) and Wrong Sign sample (**WS**, red)

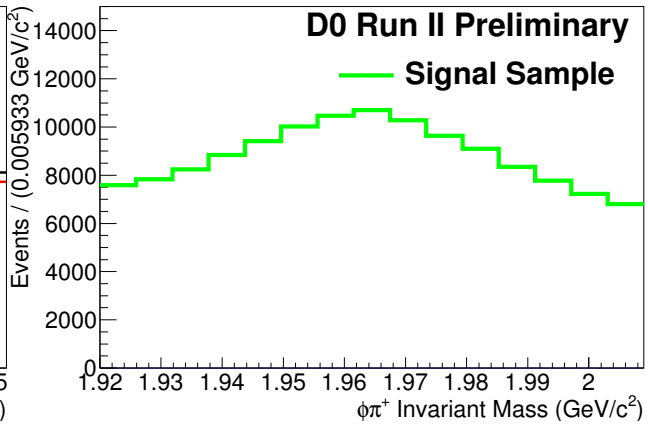


FIG. 5: Mass distribution of Signal Sample, the sample is the 2σ region around the D_s^+ mass peak.

where the products run over the events on signal sample, **SS**, and background sample, **BS**; f_{signal} is the fraction of signal events in signal sample, estimated by the fit of the D_s^- mass distribution; and $\mathcal{F}^i_{signal/background}$ is the signal/background probability density evaluated for the i-th event. $\mathcal{C}_{signal}$ is a gaussian constriction on the signal fraction.

A. Signal Probability Density

The signal probability distribution $\mathcal{F}_{signal}$ composed of a weighted sum of several probability distributions, $\mathcal{E}_{decay}$, each of these distributions is defined as a exponential decay convoluted with a resolution function $\mathcal{R}$ and smeared with a K-factor distribution $\mathcal{H}_{decay}(K)$,

$$\mathcal{E}^j_{decay}(\lambda_j, \sigma(\lambda_j), s) = \int dK \mathcal{H}_{decay}(K) \left[\frac{K}{c\tau(B_{decay})} e^{-K\lambda_j / c\tau(B_{decay})} \otimes \mathcal{R}(\lambda_j, \sigma(\lambda_j), s) \right], \tag{15}$$

where λ_j is the pesudo-proper decay length, PPDL, and $\sigma(\lambda_j)$ is the error for the PPDL measurement for the j-Th. event. For each possible contribution from B-meson decay, a $\mathcal{E}$ function is introduced, with the correct PPDL and K-factor. The resolution function is given by,

$$\mathcal{R}(\lambda_j, \sigma(\lambda_j), s) = \left(\frac{1}{\sqrt{2\pi}\ s\sigma(\lambda_j)} \right) e^{\frac{-(\lambda_j - \lambda_0)^2}{2(s\sigma(\lambda_j))^2}}. \tag{16}$$

Since a priori we do not know the overall scale of the decay length uncertainty, which is estimated on a event-by-event basis, the scale factor, s, is introduced as a free parameter in the fit. As, in principle, events that lay in the central region of the tracking system have a better spatial resolution than those events collected in the forward/backward region we use a resolution composed of two gaussian distributions, with different scale factors, given by,

$$\mathcal{R}_{tot}(\lambda_j, \sigma(\lambda_j), s_1, s_2) = f_{fine}\mathcal{R}(\lambda_j, \sigma(\lambda_j), s_1) + (1 - f_{fine})\mathcal{R}(\lambda_j, \sigma(\lambda_j), s_2) \tag{17}$$

In the fit, the integration over the K factor distribution is approximated by a sum as follows:

$$\int dK\ \mathcal{H}(K) \rightarrow \sum_{k-bin} \Delta K \mathcal{H}(K_{k-bin}), \tag{18}$$

where the sum is taken over the bin of the histogram $\mathcal{H}(K_{k-bin})$, with bin size ΔK.

Explicitly, the probability density $\mathcal{F}^i_{signal}$ takes the form:

$$\mathcal{F}^i_{signal} = f_{\bar{c}c}F^i_{\bar{c}c} + (1-f_{\bar{c}c})[f_{B_s\to D_s^- D}\mathcal{E}^i_{B_s\to D_s^- D} + f_{B_s\to D_s*D_s*}\mathcal{E}^i_{B_s\to D_s*D_s*} + f_{B_0\to D_s^- D}\mathcal{E}^i_{B_0\to D_s^- D} + f_{B^+\to D_s^- D}\mathcal{E}^i_{B^+\to D_s^- D} + (1-f_{B_s\to D_s^- D} - f_{B_s\to D_s*D_s*} - f_{B_0\to D_s^- D} - f_{B^+\to D_s^- D})\mathcal{E}^i_{B_s\to D_s^+\mu^- X}], \quad (19)$$

where $f_{\bar{c}c}$ is the expected fraction of $\bar{c}c$ events in the signal sample, $F_{\bar{c}c}$ is the PDF lifetime for the $\bar{c}c$ events (a gaussian distribution). $f_{B_a\to D_b D_c}$ is the fraction of events for the decay $f_{B_a\to D_b D_c}$, and $\mathcal{E}_{B_a\to D_b D_c}$ is the PPDL PDF associated to each decay. The last term of the sum in eq. 19 is the term associated to our signal events $B_s \to D_s^+\mu^- X$

B. Background Probability Density

The background probability density $\mathcal{F}_{bkg}$ is defined as follows:

$$\mathcal{F}^j_{bkg}(\lambda_j,\sigma(\lambda_j)) = \begin{cases} (1-f_{LSL}-f_{LLL}-f_{RSL}-f_{RLL})\mathcal{R}(\lambda_j,\sigma(\lambda_j),s) + \frac{f_{LSL}}{\lambda_{LSL}}e^{\lambda_j/\lambda_{LSL}} + \frac{f_{LLL}}{\lambda_{LLL}}e^{\lambda_j/\lambda_{LLL}} & (\lambda<0) \\ (1-f_{LSL}-f_{LLL}-f_{RSL}-f_{RLL})\mathcal{R}(\lambda_j,\sigma(\lambda_j),s) + \frac{f_{RSL}}{\lambda_{RSL}}e^{-\lambda_j/\lambda_{RSL}} + \frac{f_{RLL}}{\lambda_{RLL}}e^{-\lambda_j/\lambda_{RLL}} & (\lambda\geq 0) \end{cases} \quad (20)$$

where the details of the quantities in eq. 20 are shown in table IV

X. FIT RESULT

Table V shows the fit results for our nominal model. Where we have used the MINUIT (HESSE and MINOS) algorithms included in the RooFit package to perform the fit. The Fig. 6 shows the pseudo-proper decay length of the B_s^0 candidates in signal sample and the projections of $\mathcal{F}_{signal}, \mathcal{F}_{bkg}$ and $\mathcal{F}_{signal} + \mathcal{F}_{bkg}$. Figs 7 and 8 show the pseudo-proper decay length distribution for the side band and wrong sign sample respectively. The Background PDF F_{bkg} is also projected on each distribution.

XI. CROSSCHECKS

A. B^0 lifetime

As a crosscheck of the fitting procedure, the B_d^0 lifetime is measured. The channel $B_d^0 \to D^{(*)+}$ is used as the signal. To obtain the signal sample a 2σ mass window is taken around the D^+ mass from the same sample used for the B_s lifetime analysis. It is assumed also that there are some other B meson decays that contribute to the D^+ mass peak and that the $c\bar{c}$ contribution is the same as in the D_s^- mass peak, $f_{c\bar{c}} = 0.093$. From MC samples we compute the contributions from different decays, they are shown in Table VI.

All K factors for these decays were computed and the fit performed with the proper correct definitions in the background samples as well. It was found that the B^0 lifetime obtained is $447 \pm 10\mu m$, this is en good agreement with the world average $455.7 \pm 2.1\mu m$.

B. Split Sample Tests

To test the stability of the measurement, we have performed some consistency checks based on sample splitting procedure. No systematic uncertainty will be quoted from these tests since observed differences were statistically

f_{LSL}	fraction of events in the Left exponential with Short PPDL
f_{LLL}	fraction of events in the Left exponential with Long PPDL
f_{RSL}	fraction of events in the Right exponential with Short PPDL
f_{RLL}	fraction of events in the Right exponential with Long PPDL
λ_{LSL}	slope of events in the Left exponential with Short PPDL
λ_{LLL}	slope of events in the Left exponential with Long PPDL
λ_{RSL}	slope of events in the Right exponential with Short PPDL
λ_{RLL}	slope of events in the Right exponential with Long PPDL

TABLE IV: List of parameters used in the Lifetime Maximum Likelihood fit to model the combinatorial background.

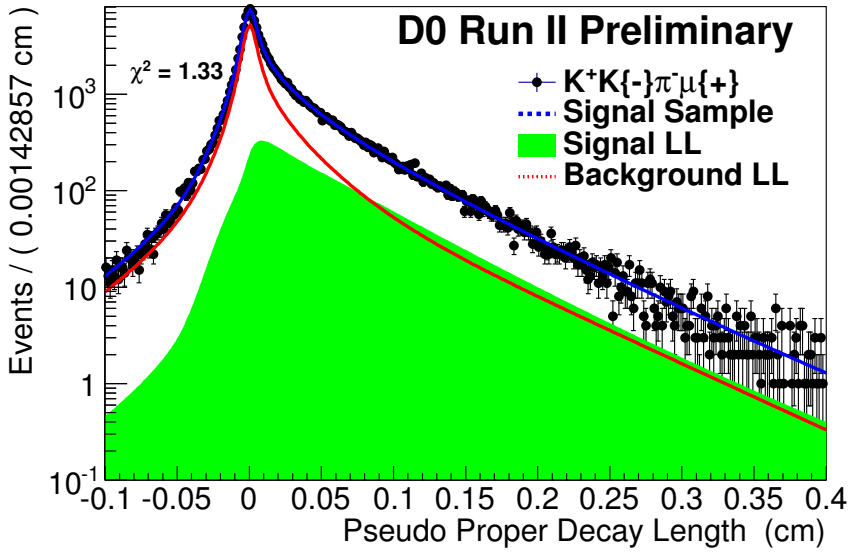


FIG. 6: Distribution of pseudo proper decay length of the signal sample with projections of the Maximum-Likelihood fit. The green area is the projection of $\mathcal{F}_{signal}$, the red one is the combinatorial background and the blue curve is the sum of the two.

Parameter	Central Value	Statistical Uncertainty
f_{LSL}	0.05212	$+0.00135$ -0.00137
f_{LLL}	0.02263	$+0.00083$ -0.00079
f_{RSL}	0.19452	$+0.00339$ -0.00337
f_{RLL}	0.10553	± 0.00035
λ_{LSL}	$77.7\mu m$	$\pm 2.7\mu m$
λ_{LLL}	$420.4\mu m$	$\pm 8.4\mu m$
λ_{RSL}	$236.9\mu m$	$\pm 3.6\mu m$
λ_{RLL}	$632.7\mu m$	$+8.5\mu m$ $-8.1\mu m$
f_{fine}	0.78257	0.00349 0.00355
$s1$	1.340	± 0.006
$s2$	4.383	± 0.061
$c\tau(B_s)$	$434.5\mu m$	$\pm 4.3\mu m$

TABLE V: Lifetime Maximum-Likelihood fit results.

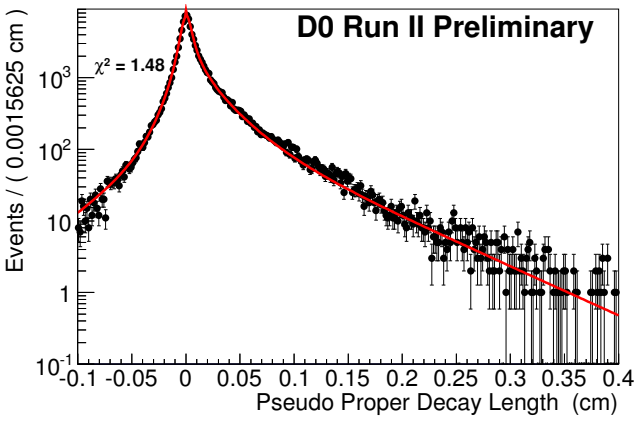


FIG. 7: Background PDF component projected over the side band sample.

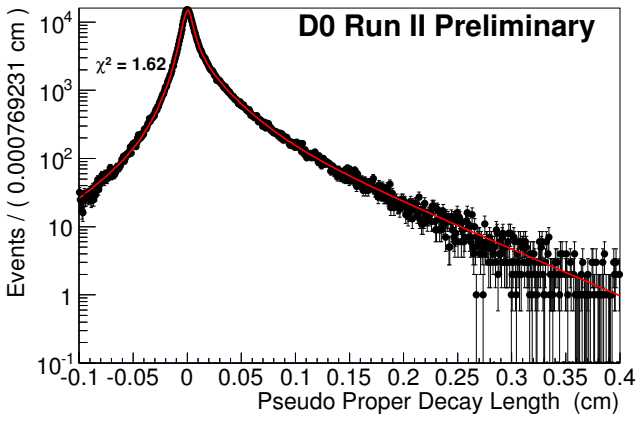


FIG. 8: Background PDF component projected over the wrong sign sample.

Decay	Contribution
$c\bar{c}$	0.093
$B^0 \to D^- \mu^+ X$	0.808
$B^+ \to D^- \mu^+ X$	0.084
$B^0 \to D^- D X$	0.008
$B_s^0 \to D^- \mu^+ X$	0.005

TABLE VI: Contributions from different decays to the D^+ mass peak.

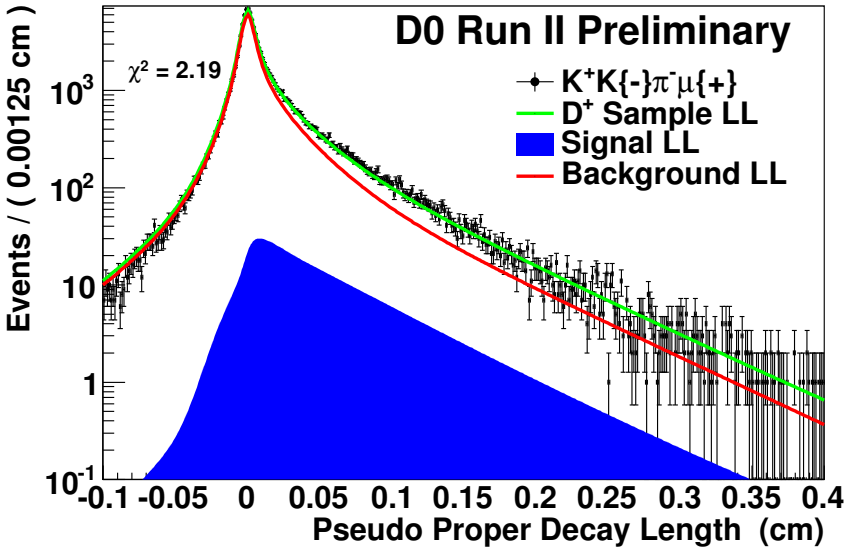


FIG. 9: Distribution of pseudo proper decay length of the D^+ signal sample with projections of the Maximum-Likelihood fit. The blue area is the projection of $\mathcal{F}^{D^+}_{signal}$, the red one is the combinatorial background and the green curve is the sum of the two

consistent with zero. Table VII shows the fitted lifetimes for different period of data taking, the measurements are consistent with the nominal value..

Sub-sample	Fit result
Run IIA	$(432.1^{+6.9}_{-6.8})\mu$m
Run IIB	$(435.3 \pm 5.3)\mu$m

TABLE VII: Fit results for the $c\tau_{B_s^0}$ for different sub-samples. The uncertainties are statistical. The measured lifetime in the sub-samples are consistent with the fit for the complete sample where $c\tau_{B_s^0} = 434.5\mu$m.

XII. SYSTEMATIC UNCERTAINTIES

A. Decay Length Resolution

In the nominal fit,a double-Gaussian function is used to model the resolution. The fit gives a fraction of 78.26% for the contribution of the narrower gaussian. However, a single gaussian resolution can be also be a valid model for the resolution. The systematic uncertainty due to this change in the resolution model is $0.9\mu m$

B. Combinatorial Background Evaluation

To evaluate the shape of the combinatorial background in the Signal Sample, side band events and wrong sign combination are used to help the fit. To associate a systematic uncertainty of this method we use a sample with only wrong sign combination, a sample with only high mass and wrong sign and a sample with low mass and wrong sign. The Table VIII shows the uncertainty associated with each sample. The maximum difference is 5μm with no Side Bands, and is the quoted uncertainty for this evaluation. This uncertainty is only positive but it is taken symmetric in the total effect.

Model	Variation
no Side Bands	$+5.0\mu m$
no Low Side Bands	$+1.5\mu m$
no High Side Bands	$+2.2\mu m$

TABLE VIII: List of systematic uncertainties associated with the combinatorial background modeling.

C. K Factor Determination

To determine the K factor of the signal events, we assume that the signal is composed according to the fractions given in the table I. To associate a systematic uncertainty to the determination of the K factor we change this assumption considering that the signal is composed only with the $B_s \to \mu\nu D_s^-$ and $B_s \to \mu\nu D_s^*$. The change in the lifetime measurement using this assumption is $1.5\mu m$

D. Non-Combinatorial Background Composition

Decay	-1σ	$+1\sigma$
$c\bar{c} \to D_s^+ \mu X$	$-3.8\mu m$	$+3.8\mu$m
$B_s^0 \to D_s^+ DX$	$+1.2\mu m$	-0.5μm
$B_s^0 \to D_s^{+(*)} D_s^{-(*)}$	$+0.2\mu m$	-0.6μm
$B^+ \to D_s^+ DX$	$+1.8\mu m$	-1.8μm
$B^0 \to D_s^+ DX$	$+1.3\mu m$	-1.3μm

TABLE IX: Systematic uncertainties associated to the variation of the contributions of the different decays contributing to the non combinatoric background. When each contribution is added in quadratures it is obtained a total uncertainty of $^{+4.6}_{-4.5}$

In the case of the non-combinatorial background, *i.e.*, the background due to a incomplete reconstruction of a B-meson decay, there is a systematic effect that is due to the uncertainty in their contribution of D_s^- mesons. There is also a contribution from $c\bar{c} \to D_s^- D$ events that it is necessary to take into account. In the Table IX is summarized the effect of the variations on each decay fraction. The total effect of these variations, summed on quadratures is $^{+4.6}_{-4.5}$.

E. Summary

The table X shows the contributions to the systematic uncertainty from different sources. The most significative effect comes from the resolution modeling. The total effect, computed adding al of them in quadratures, is ± 10 μm

XIII. RESULTS AND CONCLUSIONS

The likelihood fit gives the central value and statistic uncertainty of

$$\tau_{B_s} = 434.5^{+4.2}_{-4,1}\ \mu\text{m}. \tag{21}$$

Uncertainty source	Variation
Resolution Modeling	0.9 μm
Combinatorial Background modeling	5.0 μm
K factor determination	1.5 μm
Non Combinatorial Background	4.6 μm
Total	**7.0** μm

TABLE X: Summary of systematic uncertainty contributions.

When the statistic and systematic uncertainties are added the final result can be written as,

$$\tau_{B_s} = 434.5^{+4.2}_{-4.1}\ \mu\text{m (stat.)} \pm 7.0\ \mu\text{m (syst.)}. \tag{22}$$

The uncertainty in this measurement is dominated by the systematic effects. Taking the world average of B^0 lifetime and this measurement, one can compute the ratio $\tau_{B^0_s}/\tau_{B^0}$,

$$\frac{\tau_{B^0_s}}{\tau_{B^0}} = 0.953 \pm 0.019. \tag{23}$$

The p-value for the restrictions from lattice-QCD, eq. 6 is 1.4%.

[1] K. Hartkorn and H.-G. Moser, (1999).
[2] D. Becirevic, (2001), hep-ph/0110124.
[3] U. Nierste and A. Lenz, (2011), 1102.4274.
[4] Heavy Flavor Averaging Group, D. Asner *et al.*, (2010), 1010.1589.
[5] Particle Data Group, K. Nakamura *et al.*, (2010), J. Phys.**G 37**, 075021.

Bibliography

[1] J. Beringer et al. Review of particle physics. *Phys. Rev. D*, 86:010001, Jul 2012. xiii, 16, 17, 18, 19, 36, 100

[2] K. Nakamura et al. The Review of Particle Physics. 2010. xiv, xv, 36, 87, 90, 91, 108, 120, 135

[3] J.J. Thomson. Cathode rays. *Phil.Mag.*, 44:293–316, 1897. 1

[4] H. Geiger and E. Marsden. On a diffuse reflection of the α-particles. *Proceedings of the Royal Society of London. Series A*, 82(557):495–500, 1909. 1

[5] H. Geiger and E. Marsden. The Laws of Deflexion of α-Particles. *Phil.Mag.*, 27:604–623, 1913.

[6] E. Rutherford. The scattering of alpha and beta particles by matter and the structure of the atom. *Phil.Mag.*, 21:669–688, 1911.

[7] E. Rutherford. The structure of the atom. *Phil.Mag.*, 27:489–498, 1914. 1

[8] Wikipedia. Neutrino — Wikipedia, the free encyclopedia, 2012. [Online; accessed 17-August-2012]. 1

[9] Wikipedia. Meson — Wikipedia, the free encyclopedia, 2012. [Online; accessed 17-August-2012]. 2

[10] S.H. Neddermeyer and C.D. Anderson. NOTE ON THE NATURE OF COSMIC RAY PARTICLES. *Phys.Rev.*, 51:884–886, 1937. 2

[11] C.M.G. Lattes, G.P.S. Occhialini, and C.F. Powell. OBSERVATIONS ON THE TRACKS OF SLOW MESONS IN PHOTOGRAPHIC EMULSIONS. 1. *Nature*, 160:453–456, 1947. 2

[12] C.M.G. Lattes, G.P.S. Occhialini, and C.F. Powell. OBSERVATIONS ON THE TRACKS OF SLOW MESONS IN PHOTOGRAPHIC EMULSIONS. 2. *Nature*, 160:486–492, 1947. 2

[13] Hideki Yukawa. Meson theory in its developments. *AAPPS Bull.*, 17N1:4–8, 2007. 2

[14] Paul A.M. Dirac. The Quantum theory of electron. *Proc.Roy.Soc.Lond.*, A117:610–624, 1928. 2

[15] Carl D. Anderson. The positive electron. *Phys. Rev.*, 43:491–494, Mar 1933. 2

[16] G.D. Rochester and C.C. Butler. EVIDENCE FOR THE EXISTENCE OF NEW UNSTABLE ELEMENTARY PARTICLES. *Nature*, 160:855–857, 1947. 2

[17] G. D. Rochester. The discovery of the v-particles. *AIP Conference Proceedings*, 300(1):17–38, 1994. 3

[18] Wikipedia. Kaon — Wikipedia, the free encyclopedia, 2012. [Online; accessed 17-August-2012]. 3

[19] Steven Weinberg. Conceptual foundations of the unified theory of weak and electromagnetic interactions. *Rev. Mod. Phys.*, 52:515–523, Jul 1980. 4

[20] Abdus Salam. Gauge unification of fundamental forces. *Rev. Mod. Phys.*, 52:525–538, Jul 1980.

[21] Sheldon Lee Glashow. Towards a unified theory: Threads in a tapestry. *Rev. Mod. Phys.*, 52:539–543, Jul 1980. 4

[22] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay. Evidence for the 2π decay of the k_2^0 meson. *Phys. Rev. Lett.*, 13:138–140, Jul 1964. 6, 20

[23] Y. Amhis et al. Averages of b-hadron, c-hadron, and tau-lepton properties as of early 2012. 2012. 7, 24

[24] Victor Mukhamedovich Abazov et al. Evidence for an anomalous like-sign dimuon charge asymmetry. *Phys.Rev.Lett.*, 105:081801, 2010. 8

[25] Victor Mukhamedovich Abazov et al. Measurement of the anomalous like-sign dimuon charge asymmetry with $9fb^{-}1$ of p pbar collisions. *Phys.Rev.*, D84:052007, 2011. 8

[26] Peter W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964. 11

[27] C. Grojean and M. Spiropulu. High-energy Physics. Proceedings, 18th European School, ESHEP 2010, Raseborg, Finland, 20 Jun. - 3 Jul., 2010. 2012. 11

[28] D. Decamp et al. Determination of the Number of Light Neutrino Species. *Phys.Lett.*, B231:519, 1989. 11

[29] Makoto Kobayashi and Toshihide Maskawa. *cp*-violation in the renormalizable theory of weak interaction. *Progress of Theoretical Physics*, 49(2):652–657, 1973. 14, 89

[30] The UTfit collaboration. `http://www.utfit.org/UTfit/Results`. Fit results: Summer 2012 (pre-ICHEP12). 15

[31] The CKM fitter collaboration. `http://ckmfitter.in2p3.fr/www/html/ckm_results.html`. Preliminary results as of Summer 2012 (ICHEP 12 conference). 15

[32] Lincoln Wolfenstein. Parametrization of the kobayashi-maskawa matrix. *Phys. Rev. Lett.*, 51:1945–1947, Nov 1983. 16

[33] Andrzej J. Buras, Markus E. Lautenbacher, and Gaby Ostermaier. Waiting for the top quark mass, $k^+ \to \pi^+\nu\nu\bar{b}^0_s - B\bar{s}^0$ mixing, and *CP* asymmetries in *B* decays. *Phys. Rev. D*, 50:3433–3446, Sep 1994. 16

[34] P.F. Harrison and Helen R. Quinn, editors. *The BABAR physics book: Physics at an asymmetric B factory*. SLAC-R-504, 1998.

[35] Stephane Plaszczynski and Marie-Helene Schune. Overall determination of the CKM matrix. page hf8/019, 1999.

[36] G.P. Dubois-Felsmann, D.G. Hitlin, F.C. Porter, and G. Eigen. Sensitivity of CKM fits to theoretical uncertainties and their representation. 2003. 18

[37] The UTfit collaboration. `http://www.utfit.org/UTfit/Results`. 19

[38] The CKM fitter collaboration. `http://ckmfitter.in2p3.fr/www/html/ckm_results.html`. 19

[39] H. Burkhardt et al. First Evidence for Direct CP Violation. *Phys.Lett.*, B206:169, 1988. 20

[40] R. Aaij et al. Evidence for CP violation in time-integrated $D^0 \to h^- h^+$ decay rates. *Phys.Rev.Lett.*, 108:111602, 2012. 20

[41] B.R. Ko et al. Evidence for CP Violation in the Decay $D^+ \to K^0_S \pi^+$. *Phys.Rev.Lett.*, 109:021601, 2012.

[42] T. Aaltonen et al. Measurement of the difference of CP–violating asymmetries in $D^0 \to K^+K^-$ and $D^0 \to \pi^+\pi^-$ decays at CDF. *Phys.Rev.Lett.*, 109:111801, 2012. 20

[43] P. K. Kabir and A. Pilaftsis. Unitarity and the time evolution of quantum mechanical states. *Phys. Rev. A*, 53:66–69, Jan 1996. 22

[44] V. Weisskopf and E. Wigner. ber die natrliche linienbreite in der strahlung des harmonischen oszillators. *Zeitschrift fr Physik A Hadrons and Nuclei*, 65:18–29, 1930. 10.1007/BF01397406. 23

[45] V. Weisskopf and E. Wigner. Berechnung der natrlichen linienbreite auf grund der diracschen lichttheorie. *Zeitschrift fr Physik A Hadrons and Nuclei*, 63:54–73, 1930. 10.1007/BF01336768. 23

[46] R. Aaij et al. Determination of the sign of the decay width difference in the B^0_s system. *Phys. Rev. Lett.*, 108:241801, Jun 2012. 24

[47] K. Anikeev, D. Atwood, F. Azfar, S. Bailey, C.W. Bauer, et al. B physics at the Tevatron: Run II and beyond. 2001. 30, 34, 35

[48] (ed.) Buras, A.J. and (ed.) Lindner, M. Heavy flavors. 1992.

[49] Andrzej J. Buras, Mikolaj Misiak, and Jorg Urban. Two loop QCD anomalous dimensions of flavor changing four quark operators within and beyond the standard model. *Nucl.Phys.*, B586:397–426, 2000. 30

[50] Matthias Neubert. Effective field theory and heavy quark physics. pages 149–194, 2005. 30, 31, 33, 35

[51] Ikaros I.Y. Bigi. The QCD perspective on lifetimes of heavy flavor hadrons. 1995. 31

[52] Matthias Neubert. Heavy quark effective theory and weak matrix elements. 1998.

[53] Matthias Neubert. B decays and the heavy quark expansion. *Adv.Ser.Direct.High Energy Phys.*, 15:239–293, 1998.

[54] Nikolai Uraltsev. Topics in the heavy quark expansion. 2000. 31

[55] Iván Heredia de la Cruz. *Measurement of the Λ_b^0 lifetime in the exclusive decay $\Lambda_b^0 \to J/\psi\Lambda^0$ with the DØ detector.* PhD thesis, Centro de Investigación y Estudios Avanzados del IPN, México, Jun 2012. 31

[56] Alexander Lenz. Theoretical update of B-Mixing and Lifetimes. 2012. 33

[57] Fabrizio Gabbiani, Andrei I. Onishchenko, and Alexey A. Petrov. Lambda(b) lifetime puzzle in heavy quark expansion. *Phys.Rev.*, D68:114006, 2003. 33

[58] Ikaros I.Y. Bigi, B. Blok, Mikhail A. Shifman, N. Uraltsev, and Arkady I. Vainshtein. Nonleptonic decays of beauty hadrons: From phenomenology to theory. 1994. 35

[59] M. Beneke, G. Buchalla, and I. Dunietz. Width difference in the $B_s - \overline{b}_s$ system. *Phys. Rev. D*, 54:4419–4431, Oct 1996. 35

[60] K. Hartkorn and H.-G. Moser. A new method of measuring $\Delta\Gamma/\Gamma$" in the $B_s^0 - \bar{B}_s^0$" system. *The European Physical Journal C - Particles and Fields*, 8:381–383, 1999. 10.1007/s100529901079. 36

[61] Damir Becirevic. Theoretical progress in describing the B meson lifetimes. 2001. 36

[62] Ulrich Nierste and Alexander Lenz. Numerical updates of lifetimes and mixing parameters of B mesons. 2011. 36, 108

[63] D. Asner et al. Averages of b-hadron, c-hadron, and tau-lepton Properties. 2010. 36

[64] V.M. Abazov et al. The upgraded dø detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 565(2):463 – 537, 2006. 53, 56, 90

[65] S. Abachi et al. The dø detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 338(23):185 – 253, 1994. 55, 58

[66] John M. Campbell, J.W. Huston, and W.J. Stirling. Hard Interactions of Quarks and Gluons: A Primer for LHC Physics. *Rept.Prog.Phys.*, 70:89, 2007. 61

[67] Guennadi Borissov. The b ana packge. 63

[68] Torbjrn Sjstrand, Patrik Edn, Christer Friberg, Leif Lnnblad, Gabriela Miu, Stephen Mrenna, and Emanuel Norrbin. High-energy-physics event generation with pythia6.1. *Computer Physics Communications*, 135(2):238 – 259, 2001. 67, 91

[69] David J. Lange. The evtgen particle decay simulation package. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 462(12):152 – 155, 2001. BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines. 67, 91

[70] Rene Brun, Federico Carminati, and Simone Giani. GEANT Detector Description and Simulation Tool. 1994. 67, 92

[71] F. James and M. Roos. Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations. *Comput.Phys.Commun.*, 10:343–367, 1975. 77

[72] Rene Brun and Fons Rademakers. Root an object oriented data analysis framework. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 389(12):81 – 86, 1997. New Computing Techniques in Physics Research V. 77

[73] W. Verkerke and D. Kirkby. The RooFit toolkit for data modeling. *ArXiv Physics e-prints*, June 2003. 77, 98

[74] Alexander Lenz and Ulrich Nierste. Theoretical update of $B_s - \bar{B}_s$ mixing. *JHEP*, 0706:072, 2007. 89

[75] 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. 89

[76] Andrzej J. Buras. Messages on Flavour Physics Beyond the Standard Model. *Nucl.Phys.Proc.Suppl.*, 185:157–160, 2008. 89

[77] Zoltan Ligeti, Michele Papucci, Gilad Perez, and Jure Zupan. Implication s of the dimuon CP asymmetry in $B_{d,s}$ decays. *Phys.Rev.Lett.*, 105:131601, 2010.

[78] Andrzej J. Buras, Maria Valentina Carlucci, Stefania Gori, and Gino Isidori. Higgs-mediated FCNCs: Natural Flavour Conservation vs. Minimal Flavour Violation. *JHEP*, 1010:009, 2010.

[79] Bogdan A. Dobrescu, Patrick J. Fox, and Adam Martin. CP violation in B_s mixing from heavy Higgs exchange. *Phys.Rev.Lett.*, 105:041801, 2010.

[80] Jyoti Prasad Saha, Basudha Misra, and Anirban Kundu. Constraining Scalar Leptoquarks from the K and B Sectors. *Phys.Rev.*, D81:095011, 2010.

[81] Andrzej J. Buras, Bjorn Duling, Thorsten Feldmann, Tillmann Heidsieck, Christoph Promberger, et al. Patterns of Flavour Violation in the Presence of a Fourth Generation of Quarks and Leptons. *JHEP*, 1009:106, 2010.

[82] Gino Isidori, Yosef Nir, and Gilad Perez. Flavor Physics Constraints for Physics Beyond the Standard Model. *Ann.Rev.Nucl.Part.Sci.*, 60:355, 2010.

[83] Amarjit Soni, Ashutosh Kumar Alok, Anjan Giri, Rukmani Mohanta, and Soumitra Nandi. SM with four generations: Selected implications for rare B and K decays. *Phys.Rev.*, D82:033009, 2010.

[84] Lisa L. Everett, Jing Jiang, Paul G. Langacker, and Tao Liu. Phenomenological Implications of Supersymmetric Family Non-universal $U(1)$-prime Models. *Phys.Rev.*, D82:094024, 2010.

[85] F.J. Botella, G.C. Branco, and M.N. Rebelo. Minimal Flavour Violation and Multi-Higgs Models. *Phys.Lett.*, B687:194–200, 2010.

[86] Cheng-Wei Chiang, Alakabha Datta, Murugeswaran Duraisamy, David London, Makiko Nagashima, et al. New Physics in $B_s^0 \to J/\psi\phi$: A General Analysis. *JHEP*, 1004:031, 2010. 89

[87] V.M. Abazov et al. Lifetime difference and CP-violating phase in the B_s^0 system. *Phys.Rev.Lett.*, 98:121801, 2007. 89, 95

[88] V.M. Abazov et al. Measurement of B_s^0 mixing parameters from the flavor-tagged decay $B_s^0 \to J/\psi\phi$. *Phys.Rev.Lett.*, 101:241801, 2008. 89, 95, 109, 121

[89] T. Aaltonen et al. Measurement of lifetime and decay-width difference in $B_s^0 \to J/\psi\phi$ decays. *Phys.Rev.Lett.*, 100:121803, 2008. 89

[90] Andreas Hocker, J. Stelzer, F. Tegenfeldt, H. Voss, K. Voss, et al. TMVA - Toolkit for Multivariate Data Analysis. *PoS*, ACAT:040, 2007. 92

[91] V.M. Abazov et al. Measurement of B_d mixing using opposite-side flavor tagging. *Phys.Rev.*, D74:112002, 2006. 96, 97

[92] V.M. Abazov et al. First direct two-sided bound on the B_s^0 oscillation frequency. *Phys.Rev.Lett.*, 97:021802, 2006. 97

[93] F. Azfar, J. Boudreau, N. Bousson, J.P. Fernandez, K. Gibson, et al. Formulae for the Analysis of the Flavor-Tagged Decay $B_s^0 \to J\psi\phi$. *JHEP*, 1011:158, 2010. 98

[94] A. Abulencia et al. Observation of B0(s) - anti-B0(s) Oscillations. *Phys.Rev.Lett.*, 97:242003, 2006. 100, 119, 137

[95] W. K. HASTINGS. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109, 1970. 106

[96] Michael Gronau and Jonathan L. Rosner. Flavor symmetry for strong phases and determination of β_s, $\Delta\Gamma$ in $B_s \to J/\psi\phi$. *Phys.Lett.*, B669:321–326, 2008. 108

[97] H.G. Moser and A. Roussarie. Mathematical methods for B0 anti-B0 oscillation analyses. *Nucl.Instrum.Meth.*, A384:491–505, 1997. 138

[98] Ricardo Magańa Villalba. *CP Violation in $B_s \to J/\psi\phi$ using 8 fb^{-1} of $p\overline{p}$ of collisions.* PhD thesis, Centro de Investigación y Estudios Avanzados del IPN, México, Apr 2012.

[99] Marco Antonio Carrasco Lizárraga. *Measurement of the B_s^0 lifetime using the semileptonic decay channel $B_s^0 \to D_s^- \mu + \nu X$.* PhD thesis, Centro de Investigación y Estudios Avanzados del IPN, México, 2009.

[100] Jose Hernández Orduna, de Jesus. *Observation of the doubly strange b-Baryon* Ω_b^-. PhD thesis, Centro de Investigación y Estudios Avanzados del IPN, México, 2011.

[101] Pedro L.M. Podesta Lerma. *Measurement of the* B_s *Lifetime*. PhD thesis, Centro de Investigación y Estudios Avanzados del IPN, México, 2005.