First Flavor-Tagged Determination of Bounds on Mixing-Induced CP Violation in B^0 \rightarrow J/\psi \phi Decays

This Letter describes the first determination of bounds on the CP-violation parameter $2\beta_s$ using $B^0_s$ decays in which the flavor of the bottom meson at production is identified. The result is based on approximately 2,000 $B^0_s \to J/\psi \phi$ decays reconstructed in a 1.35 fb$^{-1}$ data sample collected with the CDF II detector using $p\bar{p}$ collisions produced at the Fermilab Tevatron. We report confidence regions in the two-dimensional space of $2\beta_s$ and the decay-width difference $\Delta \Gamma$. Assuming the standard model predictions of $2\beta_s$ and $\Delta \Gamma$, the probability of a deviation as large as the level of the observed data is 15%, corresponding to 1.5 Gaussian standard deviations.

Dedicated to the memory of our dear friend and colleague, Michael P. Schmidt.

PACS numbers: 13.25.Hw, 14.40.Nd, 14.65.Fy

---

*With visitors from *University of Athens, 15784 Athens, Greece, 6Universidad Autónoma de Madrid, E-28049 Madrid, Spain,
The accurate determination of charge-conjugation-parity (CP) violation in meson systems has been one of the goals of particle physics since the effect was first discovered in neutral kaon decays in 1964 [1]. Standard model CP-violating effects are described through the Cabibbo-Kobayashi-Maskawa (CKM) mechanism [2], which has proved to be extremely successful in describing the phenomenology of CP violation in $B^0$ and $B^+$ decays in the past decade [3]. However, comparable experimental knowledge of $B^0_s$ decays has been lacking.

In the $B^0_s$ system, the mass eigenstates $B^0_{sL}$ and $B^0_{sH}$ are admixtures of the flavor eigenstates $B^0_s$ and $B^0_{s\bar{s}}$. This causes oscillations between the $B^0_s$ and $B^0_{s\bar{s}}$ states with a frequency proportional to the mass difference of the mass eigenstates, $\Delta m_s \equiv m_{H} - m_{L}$. In the standard model this effect is explained in terms of second-order weak processes involving virtual massive particles that provide a transition amplitude between the $B^0_s$ and $B^0_{s\bar{s}}$ states. The magnitude of this mixing amplitude is proportional to the oscillation frequency, while its phase, responsible for CP violation in $B^0_s \rightarrow J/\psi \phi$ decays, is $-2\beta^{SM}_s = -2 \arg \left( -\frac{V_{ts}V_{ts}^*}{V_{tb}V_{tb}^*} \right)$ [4], where $V_{ij}$ are elements of the CKM quark mixing matrix. The presence of physics beyond the standard model could contribute additional processes and modify the magnitude or the phase of the mixing amplitude. The recent precise determination of the oscillation frequency [5] indicates that contributions of new physics to the magnitude, if any, are extremely small [6]. Global fits of experimental data tightly constrain the CP phase to small values in the context of the standard model, $2\beta^{SM}_s \approx 0.04$ [7]. However, new physics may contribute significantly larger values [6, 8]. The observed CP phase can be expressed as $2\beta_s = 2\beta^{SM}_s - \phi^{NP}$, where $\phi^{NP}$ is due to the additional processes. The decay-width difference between the mass eigenstates, $\Delta \Gamma \equiv \Gamma_L - \Gamma_R$, is also sensitive to the same new physics phase. If $\phi^{NP} \gg 2\beta^{SM}_s$, we expect $\Delta \Gamma = 2|\Gamma_{12}| \cos(2\beta_s)$ [8], where $|\Gamma_{12}|$ is the off-diagonal element of the $B^0_s$-$\bar{B}^0_s$ decay matrix from the Schrödinger equation describing the time evolution of $B^0_s$ mesons [9]. Recent studies of $B^0_s \rightarrow J/\psi \phi$ decays without identification of the initial flavor of the $B^0_s$ meson [9, 10] have provided information on $\Delta \Gamma$ and have some limited sensitivity to the CP phase.

In this Letter we present the first study of the $B^0_s \rightarrow J/\psi \phi$ decay [11] in which the initial state of the $B^0_s$ meson (i.e., whether it is produced as $B^0_s$ or its anti-particle $\bar{B}^0_s$) is identified in a process known as “flavor tagging”. Such information is necessary to separate the time evolution of mesons produced as $B^0_s$ or $\bar{B}^0_s$. By relating this time development with the CP eigenvalue of the final states, which is accessible through the angular distributions of the $J/\psi$ and $\phi$ mesons, we obtain direct sensitivity to the CP-violating phase. This phase enters the time-development with terms proportional to both $|\cos(2\beta_s)|$ and $|\sin(2\beta_s)|$. Analyses of $B^0_s \rightarrow J/\psi \phi$ decays that do not use flavor tagging are primarily sensitive to $|\cos(2\beta_s)|$ and $|\sin(2\beta_s)|$, leading to a four-fold ambiguity in the determination of $2\beta_s$ [9, 10].

This measurement uses 1.35 fb$^{-1}$ of data collected by the CDF experiment at the Fermilab Tevatron between February 2002 and September 2006. The CDF II detector is described in detail in Ref. [12]. Detector sub-systems relevant for this analysis are described briefly here. The tracking system is composed of silicon micro-strip detectors surrounded by a multi-wire drift chamber. The drift chamber provides tracking information and charged particle identification through the measurement of specific ionization energy loss ($dE/dx$). A time-of-flight (TOF) detector provides additional particle identification. These detectors are immersed within a 1.4 T axial magnetic field. Electromagnetic and hadronic calorimeters surround the solenoid. At the outermost radial extent of the detector, muons are detected in planes of multi-wire drift chambers and scintillators. The data used were collected with a di-muon trigger which preferentially selects events containing $J/\psi \rightarrow \mu^+ \mu^-$ decays [12].

We reconstruct the $B^0_s \rightarrow J/\psi \phi$ decay from the decays $J/\psi \rightarrow \mu^+ \mu^-$ and $\phi \rightarrow K^+ K^-$ and require these final state particles to originate from a common point. We use an artificial neural network (ANN) [13] to separate $B^0_s \rightarrow J/\psi \phi$ signal from background. In the ANN training, we consider the following variables: particle identification of kaons using the TOF and $dE/dx$, the component of momenta of the $B^0_s$ and $\phi$ candidates transverse to the proton beam direction, the invariant mass of the $\phi$ candidate, and the quality of a kinematic fit to the trajectories of the final state particles. We have trained the ANN with signal events from simulated data that are passed through the standard GEANT-based [14] simulation of the CDF II detector [15] and are reconstructed as in real data. We use $B^0_s \rightarrow J/\psi \phi$ mass sideband candidates, defined as those having $m(J/\psi\phi) \in [5.1820, 5.2142]$. 

---

$^c$Chinese Academy of Sciences, Beijing 100864, China, $^d$University of Bristol, Bristol BS8 1TL, United Kingdom, $^e$University Libre de Bruxelles, B-1050 Brussels, Belgium, $^f$University of California Irvine, Irvine, CA 92697, $^g$University of California Santa Cruz, Santa Cruz, CA 95064, $^h$Cornell University, Ithaca, NY 14853, $^i$University of Cyprus, Nicosia CY-1678, Cyprus, $^j$University College Dublin, Dublin 4, Ireland, $^k$University of Edinburgh, Edinburgh EH9 3JZ, United Kingdom, $^l$University of Heidelberg, D-69120 Heidelberg, Germany, $^m$Universidad Iberoamericana, Mexico D.F., Mexico, $^n$University of Manchester, Manchester M13 9PL, England, $^o$Nagasaki Institute of Applied Science, Nagasaki, Japan, $^p$University de Oviedo, E-33007 Oviedo, Spain, $^q$Queen Mary, University of London, London, E1 4NS, England, $^r$Texas Tech University, Lubbock, TX 79409, $^s$IFIC(CSIC-Universitat de Valencia), 46071 Valencia, Spain,
[5.3430, 5.3752] GeV/c^2, as the background sample in the ANN training. Applying the selection on the output variable of the ANN, we observe 2,019 \pm 73 B^0 \rightarrow J/\psi \phi signal events with a signal to background ratio of approximately one. The invariant J/\psi \phi mass distribution is shown in Fig. 1. An event-specific primary interaction point is used in the calculation of the proper decay time, \( t = m(B^0_{s})/p_T(B^0_{s}) \), where \( L_{xy}(B^0_{s}) \) is the distance from the primary vertex to the \( B^0_{s} \rightarrow J/\psi \phi \) decay vertex projected onto the momentum of the \( B^0_s \) in the plane transverse to the proton beam direction, \( m(B^0_s) \) is the mass of the \( B^0_s \) meson [3], and \( p_T(B^0_s) \) is its measured transverse momentum.

The orbital angular momenta of the vector (spin 1) mesons, \( J/\psi \) and \( \phi \), produced in the decay of the pseudo-scalar (spin 0) \( B^0_s \) meson, are used to distinguish the \( CP \)-even S- and D-wave final states from the \( CP \)-odd P-wave final state. We measure the decay angles \( \theta_T, \phi_T \), and \( \psi_T \), defined in Ref. [9], in the transversity basis [16]. The transverse linear polarization amplitudes at \( t = 0 \), \( A_\parallel \) and \( A_\perp \), correspond to \( CP \) even and \( CP \) odd final states, respectively. The longitudinal polarization amplitude \( A_0 \) corresponds to a \( CP \) even final state. The polarization amplitudes are required to satisfy the condition \( |A_0|^2 + |A_\parallel|^2 + |A_\perp|^2 = 1 \).

In order to separate the time development of the \( B^0_s \) meson from that of the \( B^0_s \) meson, we identify the flavor of the \( B^0_s \) or \( B^0_s \) meson at the time of production by means of flavor tagging. Two independent types of flavor tags are used, each exploiting specific features of the production of \( b \) quarks at the Tevatron, where they are mostly produced as \( b\bar{b} \) pairs. The first type of flavor tag infers the production flavor of the \( B^0_s \) or \( B^0_s \) meson from the decay products of the \( b \) hadron produced by the other \( b \) quark in the event. This is known as an opposite-side flavor tag (OST). The OST decisions are based on the charge of muons or electrons from semileptonic \( B \) decays [17, 18] or the net charge of the opposite-side jet [19]. If multiple tags are available for an event, the decision from the highest dilution flavor tag is chosen [20]. The tag dilution \( D \), defined by the probability to correctly tag a candidate \( P_{tag} \equiv (1 + D)/2 \), is estimated for each event. The calibration of the OST dilution is determined from \( B^+ \rightarrow J/\psi K^+ \) and \( B^0 \rightarrow J/\psi K^{0*} \) decays. The second type of flavor tag identifies the flavor of the reconstructed \( B^0_s \) or \( B^0_s \) meson at production by correlating it with the charge of an associated kaon arising from fragmentation processes [21], referred to as a same-side kaon tag (SSKT). The SSKT algorithm and its dilution calibration on simulated data are described in Ref. [22]. The average dilution is \((11 \pm 2)\% \) for the SSKT and \((27 \pm 4)\% \) for the SSNT, where the uncertainties contain both statistical and systematic effects. The measured efficiencies for a candidate to be tagged are \((96 \pm 1)\% \) for the SSNT and \((50 \pm 1)\% \) for the SSKT.

An unbinned maximum likelihood fit is performed to extract the parameters of interest, \( 2\beta_s, \Delta \Gamma \), plus nuisance parameters to the measurement, which include the signal fraction \( f_s \), the mean \( B^0_s \) width \( \Gamma = (\Gamma_s + \Gamma_{\bar{s}})/2 \), the mixing frequency \( \Delta m_s \), the magnitudes of the polarization amplitudes \( |A_0|^2, |A_\parallel|^2 \), and \( |A_\perp|^2 \), and the strong phases \( \delta_0 \equiv \arg(A^*_0A_0) \) and \( \delta_\perp \equiv \arg(A^*_1A_0) \). The fit uses information on the reconstructed \( B^0_s \) candidate mass \( m \) and its uncertainty \( \sigma_m \), the \( B^0_s \) candidate proper decay time \( t \) and its uncertainty \( \sigma_t \), the transversity angles \( \vec{\rho} = \{\cos \theta_T, \phi_T, \cos \psi_T\} \), and tag information \( D \) and \( \xi \), where \( D \) is the event-specific dilution and \( \xi = \{-1, 0, +1\} \) is the tag decision, in which +1 corresponds to a candidate tagged as \( B^0_s \), -1 to a \( B^0_s \) and 0 to an untagged candidate. The single-event likelihood is described in terms of signal (\( P_s \)) and background (\( P_b \)) probability distribution functions (PDFs) as

\[
\begin{align*}
&f_s P_s(m|\sigma_m)P_s(t, \vec{p}, \xi|D, \sigma_t)P_s(\sigma_t)P_s(D) \\
&+ (1 - f_s) P_b(m|\sigma_m)P_b(t|\sigma_t)P_b(\vec{p})P_b(\sigma_t)P_b(D).
\end{align*}
\]

The signal mass PDF \( P_s(m|\sigma_m) \) is parameterized as a single Gaussian with a standard deviation determined separately for each candidate, while the background mass PDF, \( P_b(m) \), is parameterized as a first order polynomial. The distributions of the decay time uncertainty and the event-specific dilution are observed to be different in signal and background, so we include their PDFs explicitly in the likelihood. The signal PDFs \( P_s(\sigma_t) \) and \( P_s(D) \) are determined from sideband-subtracted data distributions, while the background PDFs \( P_b(\sigma_t) \) and \( P_b(D) \) are determined from the \( J/\psi \phi \) invariant mass sidebands. The PDFs of the decay time uncertainties, \( P_s(\sigma_t) \) and \( P_b(\sigma_t) \), are described with a sum of Gamma function distributions, while the dilution PDFs \( P_s(D) \) and \( P_b(D) \) are included as histograms that have been extracted from data.

The time and angular dependence of the signal PDF
for a single flavor tag can be written in terms of two PDFs, $P$ for $B^0_s$, and $P$ for $B^0$, as

$$P_s(t, \bar{\rho}, \xi, |D, \sigma) = \frac{1 + \xi D}{2} P(t, \bar{\rho} \varepsilon) + \frac{1 - \xi D}{2} P(t, \bar{\rho} \varepsilon),$$

which is trivially extended in the case of two independent flavor tags (OST and SSKT). The detector acceptance effects on the transversity angle distributions, $\varepsilon(\bar{\rho})$, are modeled with $B^0_s \rightarrow J/\psi \phi$ simulated data. Three-dimensional joint distributions of the transversity angles are used to determine $\varepsilon(\bar{\rho})$, in order to correctly account for any dependencies among the angles. The time and angular probabilities for $B^0_s$ can be expressed as

$$d^4 P(t, \bar{\rho}) = |A_0|^2 T_{+} f_1(\bar{\rho}) + |A_1|^2 T_{+} f_2(\bar{\rho})$$

+ $|A_1|^2 T_{-} f_3(\bar{\rho}) + |A_2|^2 T_{-} f_4(\bar{\rho})$

+ $|A_0|^2 |A_2| \cos(\delta) T_{+} f_5(\bar{\rho})$

+ $|A_0|^2 |A_1| V_+ f_6(\bar{\rho}),$

where the functions $f_1(\bar{\rho}), \ldots, f_6(\bar{\rho})$ are defined in Ref. [9]. The probability $P$ for $B^0_s$ is obtained by substituting $U_+ \rightarrow U_-$ and $V_+ \rightarrow V_-$. The time-dependent term $T_{\pm}$ is defined as

$$T_{\pm} = e^{-\Gamma t} \times [\cosh(\Delta \Gamma t/2) \mp \cos(2\beta_s) \sinh(\Delta \Gamma t/2)$$

+ $\eta \sin(2\beta_s) \sin(\Delta m_s t)],$$

where $\eta = +1$ for $P$ and $-1$ for $\bar{P}$. The other time-dependent terms are defined as

$$U_\pm = \pm e^{-\Gamma t} \times [\sin(\delta_\perp - \delta_\parallel) \cos(\Delta m_s t)$$

- $\cos(\delta_\perp - \delta_\parallel) \cos(2\beta_s) \sin(\Delta m_s t)$

$\pm \cos(\delta_\perp - \delta_\parallel) \sin(2\beta_s) \sinh(\Delta \Gamma t/2)]$$

+ $V_\pm = \mp e^{-\Gamma t} \times [\sin(\delta_\perp) \cos(\Delta m_s t)$$

- $\cos(\delta_\perp) \cos(2\beta_s) \sin(\Delta m_s t)$

$\pm \cos(\delta_\perp) \sin(2\beta_s) \sinh(\Delta \Gamma t/2)].$

These relations assume that there is no direct CP violation in the system. The time-dependence is convolved with a Gaussian proper time resolution function with standard deviation $\sigma$, which is adjusted by an overall calibration factor determined from the fit using promptly decaying background candidates. The average of the resolution function is 0.1 ps, with a root-mean-square deviation of 0.04 ps.

We model the lifetime PDF for the background, $P_b(t|\sigma)$, with a delta function at $t = 0$, a single negative exponential, and two positive exponentials, all of which are convolved with the Gaussian resolution function. The background angular PDFs are factorized, $P_b(\bar{\rho}) = P_b(\cos \theta) P_b(\phi) P_b(\cos \psi)$, and are obtained using $B^0_s$ mass sidebands events.

Possible asymmetries between the tagging rate and dilution of $B^0_s$ and $B^0$ mesons have been studied with control samples and found to be statistically insignificant. We allow important sources of systematic uncertainty, such as the determination of overall calibration factors associated with the proper decay time resolution and the dilutions, to float in the fit. The mixing frequency $\Delta m_s = 17.77 \pm 0.12$ ps$^{-1}$ is constrained in the fit within the experimental uncertainties [5]. Systematic uncertainties coming from alignment, detector sculpting, background angular distributions, decays from other B mesons, the modeling of signal and background are found to have a negligible effect on the determination of both $\Delta \Gamma$ and $\beta_s$ relative to statistical uncertainties.

An exact symmetry is present in the signal probability distribution, as can be seen in Eq. (3), which is invariant under the simultaneous transformation $(2\beta_s \rightarrow \pi - 2\beta_s, \Delta \Gamma \rightarrow - \Delta \Gamma, \delta_\parallel \rightarrow 2\pi - \delta_\parallel, and \delta_\perp \rightarrow \pi - \delta_\perp$. This causes the likelihood function to have two minima. This symmetry can be removed by restricting any of the above parameters within appropriate ranges. However, even after removal of the exact symmetry, approximate symmetries remain, producing local minima. Since the log-likelihood function is non-parabolic, we cannot meaningfully quote point estimates. Instead we choose to construct a confidence region in the $2\beta_s - \Delta \Gamma$ plane.

We use the Feldman-Cousins likelihood ratio ordering [23] to determine the confidence level (CL) for a 20 x 40 grid evenly spaced in $2\beta_s \in [-\pi/2, 3\pi/2]$ and $\Delta \Gamma \in [-0.7, 0.7]$. The other parameters in the fit are treated as nuisance parameters (e.g. $B^0_s$ mean width, transversity amplitudes, strong phases) [24]. The coverage against deviations of the nuisance parameters from the measured values is confirmed by randomly sampling the nuisance parameter space within $\pm \sigma$ of the values determined from the fit to data. The 68% and 95% confidence regions obtained are shown in Fig. 2. The solution centered in $0 \leq 2\beta_s \leq \pi/2$ and $\Delta \Gamma > 0$ corresponds to $\cos(\delta_\perp) < 0$ and $\cos(\delta_\perp - \delta_\parallel) > 0$, while the opposite is true for the solution centered in $\pi/2 \leq \beta_s \leq \pi$ and $\Delta \Gamma < 0$. Assuming the standard model predicted values of $2\beta_s = 0.04$ and $\Delta \Gamma = 0.096$ ps$^{-1}$ [8], the probability to observe a likelihood ratio equal to or higher than what is observed in data is 15%. Additionally, we present a Feldman-Cousins confidence interval of $2\beta_s$, where the $\Delta \Gamma$ is treated as a nuisance parameter, and find that $2\beta_s \in [0.32, 2.82]$ at the 68% confidence level. The CP phase $2\beta_s$, $\Delta \Gamma$, $\Gamma$, and the linear polarization amplitudes are consistent with those measured in Ref. [9].

We also exploit current experimental and theoretical information to extract tighter bounds on the CP-violating phase. By applying the constraint $|\Gamma_{12}| = 0.048 \pm 0.018$ [8] in the relation $\Delta \Gamma = 2|\Gamma_{12}| \cos(2\beta_s)$, we obtain $2\beta_s \in [0.24, 1.36] \cup [1.78, 2.90]$ at the 68% CL. If we additionally constrain the strong phases $\delta_\parallel$ and $\delta_\perp$
FIG. 2: Feldman-Cousins confidence region in the $2\beta_s - \Delta \Gamma$ plane, where the standard model favored point is shown with error bars [8]. The intersection of the horizontal and vertical dotted lines indicates the reflection symmetry in the $2\beta_s - \Delta \Gamma$ plane.

to the results from $B^0 \rightarrow J/\psi K^{*0}$ decays [25] and the $B^0_s$ mean width to the world average $B^0_s$ width [3], we find $2\beta_s \in [0.40,1.20]$ at the 68% CL.

In summary we present confidence bounds on the $CP$-violation parameter $2\beta_s$ and the width difference $\Delta \Gamma$ from the first measurement of $B^0_s \rightarrow J/\psi \phi$ decays using flavor tagging. Assuming the standard model predicted values of $2\beta_s = 0.04$ and $\Delta \Gamma = 0.096$ ps$^{-1}$, the probability of a deviation as large as the level of the observed data is 15%, which corresponds to 1.5 Gaussian standard deviations. Treating $\Delta \Gamma$ instead as a nuisance parameter and fitting only for $2\beta_s$, we find that $2\beta_s \in [0.32, 2.82]$ at the 68% confidence level. The presented experimental bounds restrict the knowledge of $2\beta_s$ to two of the four solutions allowed in measurements that do not use flavor tagging [9, 10] and improve the overall knowledge of this parameter.

We would like to thank U. Nierste for several useful suggestions. We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contributions. This work was supported by the U.S. Department of Energy and National Science Foundation; the Italian Istituto Nazionale di Fisica Nucleare; the Ministry of Education, Culture, Sports, Science and Technology of Japan; the Natural Sciences and Engineering Research Council of Canada; the National Science Council of the Republic of China; the Swiss National Science Foundation; the A.P. Sloan Foundation; the Bundesministerium für Bildung und Forschung, Germany; the Korean Science and Engineering Foundation and the Korean Research Foundation; the Science and Technology Facilities Council and the Royal Society, UK; the Institut National de Physique Nucléaire et Physique des Particules/CNRS, France; the Russian Foundation for Basic Research; the Comisión Interministerial de Ciencia y Tecnología, Spain; the European Community’s Human Potential Programme; the Slovak R&D Agency; and the Academy of Finland.

[9] ArXiv:hep-ex/... (CDF note 9100), to be submitted to PRL.
[11] Charge-conjugate states are implied throughout the paper unless otherwise specified.