View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Scientific Open-access Literature Archive and Repository

IL NUOVO CIMENTO Vol. 35 C, N. 6 Novembre-Dicembre 2012
DOI 10.1393/ncc/i2012-11373-7

Colloquia: LaThuile12

Studies of CP violation in the decay $B_s \to J/\psi\phi$ at DØ

R. Magana-Villalba
Centro de Investigación y Estudios Avanzados del IPN - México D.F., Mexico

ricevuto il 7 Settembre 2012

Summary. — We report a measurement of the lifetime, decay rate difference and CP-violating phase for B_s^0 mesons obtained from a flavour-tagged, time-dependent, angular analysis in a sample of 5598 ± 113 $B_s^0 \to J/\psi\phi$ decays selected in a $8\,\text{fb}^{-1}$ data sample of $p\bar{p}$ collisions at $\sqrt{s} = 1.96\,\text{TeV}$ recorded with the DØ experiment at Fermilab.

PACS 13.25.Hw – Decays of bottom mesons.
PACS 11.30.Er – Charge conjugation, parity, time reversal, and other discrete symmetries.

1. – Introduction

In the standard model (SM), the light (L) and heavy (H) mass eigenstates of the mixed 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, due to the interference of the decay with and without mixing, is predicted [1] to be $\phi_s^{J/\psi\phi} = -2\beta_s = 2\arg[-V_{tb}V_{ts}^*/V_{cb}V_{cs}^*] = -0.038 \pm 0.002$, where V_{ij} are elements of the Cabibbo-Kobayashi-Maskawa quark-mixing matrix [2]. New phenomena may alter the observed phase [3] to $\phi_s^{J/\psi\phi} \equiv -2\beta_s + \phi_s^{\Delta}$.

Here 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 $8.0\,\text{fb}^{-1}$ collected with the D0 detector [4] 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 $\phi_s^{J/\psi\phi}$.

2. – Data sample and event reconstruction

The analysis presented here is based on data accumulated between February 2002 and June 2010 and corresponds to $8\,\text{fb}^{-1}$ of integrated luminosity.

© Società Italiana di Fisica

We reconstruct the decay chain $B_s^0 \to J/\psi\phi$, $J/\psi \to \mu^+\mu^-$, $\phi \to K^+K^-$ from candidate $(J/\psi, \phi)$ pairs consistent with coming from a common vertex and having an invariant mass in the range $5.37 \pm 0.20\,\text{GeV}$. Events are collected with a mixture of single and dimuon triggers.

B_s^0 candidate events are required to include two opposite-sign muons accompanied by two opposite-sign tracks. The invariant mass range for muon pairs is $3.096 \pm 0.350\,\text{GeV}$, consistent with J/ψ decay. J/ψ candidates 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 $1.019 \pm 0.030\,\text{GeV}$. Each of the four final-state tracks is required to have at least one SMT hit.

A kinematic fit under the B_s^0 decay hypothesis constrains the dimuon invariant mass to the world-average J/ψ mass [5] and constrains the four-track system to a common vertex.

The primary vertex (PV) is reconstructed using 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 [5], and $\vec{p}$ is the particle momentum. Approximately 5 million events are accepted after the selection described in this section.

3. – Background suppression

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

Three Monte Carlo (MC) samples are used to study background suppression: signal, prompt background, and non-prompt background. All three are generated with PYTHIA [6]. Hadronization is also done in PYTHIA, but all hadrons carrying heavy flavors are passed on to EvtGen [7] 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. For the signal sample, events with a $B_s^0 \to J/\psi\phi$ are selected. There are approximately 10^6 events in each background and the signal MC samples. All events are passed through a full standard chain of GEANT-based [8] detector software of DØ simulation.

4. – Multivariate event selection

To discriminate the signal from background events, we use the TMVA package [9]. 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.

To choose the best set of criteria for the two BDT discriminants, we start with 14 data samples with signal yields ranging from 4000 to 7000 events. For each sample we 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. 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%.

The choice of the final cut on the BDT output is based on an ensemble study. 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 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. Based on these results, we choose the sample that contains about 6500 signal events.

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 ref. [10]. We refer to this second sample as the "Square-cuts" sample.

5. – 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 $p\bar{p}$ interaction.

The OST algorithm 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 $\bar{b}$ quark. Larger $|d|$ values correspond to higher tagging confidence. The OST-discriminating variables and algorithm are described in detail in ref. [11].

The tagging dilution $\mathcal{D}$ is defined as $\mathcal{D} = N_{\rm cor} - N_{\rm wr}/(N_{\rm cor} + N_{\rm wr})$ where $N_{\rm cor}$ ($N_{\rm wr}$) is the number of events with correctly (wrongly) identified initial B-meson flavor.

The dependence of the tagging dilution on the tagging parameter d is calibrated with data for which the flavor (B or $\overline{B}$) is known. The dilution calibration is based on independent $B_d^0 \to \mu\nu D^{*\pm}$ data samples. We perform an analysis of the B_d^0-$\overline{B}_d^0$ oscillations described in ref. [12]. We divide the samples into 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 running

TABLE I. – *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 $t = 0$
$A1$	$\|A_\parallel\|^2/(1-\|A_0\|^2)$
$\overline{\tau}_s$ (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	$\arg(A_s/A_0)$

period of time is parametrized by function

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

and the function is fitted to the data.

6. – 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 66455 events in the Square-cuts sample. We adopt the formulae and notation of ref. [13]. 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.

6·1. *Signal model.* – The angular distribution of the signal is expressed in the transversity basis.

The P wave is decomposed into three independent components corresponding to linear polarization states of the vector mesons J/ψ and ϕ, which are either longitudinal (0) or transverse to their direction of motion, and parallel ($\|$) or perpendicular ($\perp$) to each other, The time evolution of the angular distribution of the decay products, expressed in terms of the magnitudes $|A_0|$, $|A_{||}|$, and $|A_\perp|$, and two phases, $\delta_{||}$ and $\delta_\perp$. By convention, the phase of A_0 is set to zero.

The contribution from the decay to $J/\psi K^+K^-$ with the kaons in the S wave is expressed in terms of the S-wave fraction F_S and a phase δ_s. The squared sum of the P and S waves is integrated over the KK mass. For the P wave, we assume the nonrelativistic Breit-Wigner model with the ϕ meson mass 1.019 GeV and width 4.26 MeV. 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 $\Delta M_s = 17.77 \pm 0.12\,\mathrm{ps}^{-1}$, as measured in ref. [14]. Table I 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 simulation.

6˙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.

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 17×2 parameters allowed to vary. In subsequent fits those that converge at values within two standard deviations of zero are set to zero. All background parameters described above are varied simultaneously with physics parameters. In total, there are 36 parameters used in the fit.

6˙3. *Systematic uncertainties.* – There are several possible sources of systematic uncertainty in the measurements. These uncertainties are estimated for:

- *Flavor tagging*: 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 alter the nominal parameters by their uncertainties.

- *Proper decay time resolution*: To assess the effect, we have used two alternative parameterizations 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.

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

The differences between the best-fit values and the alternative fit values provide a measure of systematic effects. For the best estimate of the CL ranges for all the measured physics quantities, we conduct Markov Chain Monte Carlo (MCMC).

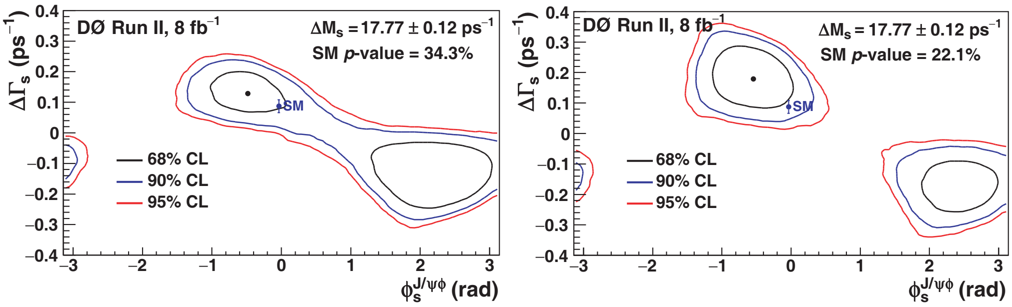

Fig. 1. – Two-dimensional 68%, 90% and 95% CL contour for BDT and square cuts selection. The standard model expectation is indicated as a point with an error.

7. – Confidence intervals

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.

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 $B_d^0 \to J/\psi K^*$ [5], up to the sign ambiguities. Reference [15] 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 $\cos\delta_\perp < 0$.

7·1. *Results*. – The fit assigns 5598 ± 113 (5050 ± 105) events to the signal for the BDT (Square-cuts) sample.

Figure 1 shows 68%, 90% and 95% CL 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 CL interval for each parameter, while the most probable value defines the central value.

The one-dimensional estimates of physics parameters for the BDT-cuts and Square-cuts sample are shown in table II.

To obtain the final CL 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 2 shows 68%, 90% and 95% CL

TABLE II. – *The one-dimensional estimates of physics parameters for the BDT-cuts sample, Square-cuts sample and final result values with systematic error.*

Parameter	BDT-cut sample	Square-cut sample	Final result
$\overline{\tau}_s$ (ps)	$1.426^{+0.035}_{-0.032}$	$1.444^{+0.041}_{-0.033}$	$1.443^{+0.038}_{-0.035}$
$\Delta\Gamma_s$ (ps^{-1})	$0.129^{+0.076}_{-0.053}$	$0.179^{+0.059}_{-0.060}$	$0.163^{+0.065}_{-0.064}$
$\phi_s^{J/\psi\phi}$	$-0.49^{+0.48}_{-0.40}$	$-0.56^{+0.36}_{-0.32}$	$-0.55^{+0.38}_{-0.36}$
$\lvert A_0\rvert^2$	$0.552^{+0.016}_{-0.017}$	0.565 ± 0.017	$0.558^{+0.017}_{-0.019}$
$\lvert A_\parallel\rvert^2$	$0.219^{+0.020}_{-0.021}$	$0.249^{+0.021}_{-0.022}$	$0.231^{+0.024}_{-0.030}$
$\delta_\parallel$	3.15 ± 0.27	3.15 ± 0.19	3.15 ± 0.22
$\cos(\delta_\perp - \delta_s)$	-0.06 ± 0.24	$-0.20^{+0.26}_{-0.27}$	$-0.11^{+0.27}_{-0.25}$
F_S	0.146 ± 0.035	0.173 ± 0.036	0.173 ± 0.036

contours in the ($\phi_s^{J/\psi\phi}$,$\Delta\Gamma_s$) plane. The p-value for the SM point [16] $(\phi_s^{J/\psi\phi}, \Delta\Gamma_s) = (-0.038, 0.087\,\text{ps}^{-1})$ is 29.8%.

8. – 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 final result values for the 68% CL intervals, including systematic uncertainties, with the oscillation frequency constrained to $\Delta M_s = 17.77 \pm 0.12\,\text{ps}^{-1}$, are shown in the last column of table II. 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%.

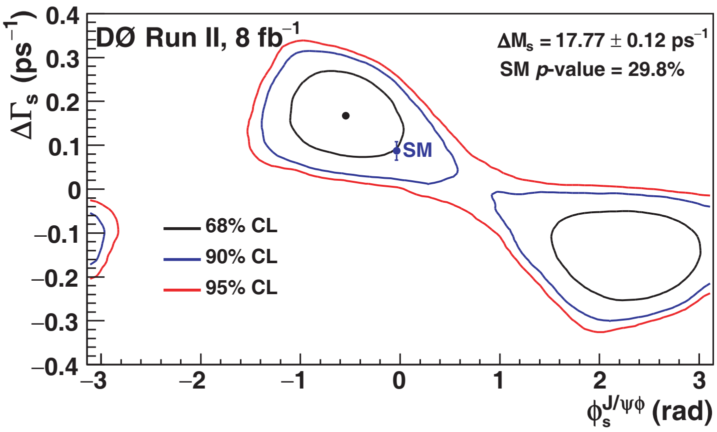

Fig. 2. – Two-dimensional 68%, 90% and 95% CL contours including systematic uncertainties. The standard model expectation is indicated as a point with an error.

* * *

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).

REFERENCES

[1] Bona M. *et al.*, *JHEP*, **10** (2006) 081.
[2] Kobayashi M. and Maskawa T., *Prog. Theor. Phys.*, **49** (1973) 652.
[3] Bona M. *et al.*, *JHEP*, **10** (2006) 081.
[4] Abazov V. M. *et al.* (D0 Collaboration), *Nucl. Instrum. Methods Phys. Res. A*, **565** (2006) 463.
[5] Nakamura K. *et al.* (Particle Data Group), *J. Phys. G*, **37** (2010) 075021.
[6] Bengtsson H. U. and Sjöstrand T., *JHEP*, **05** (2006) 026.
[7] Lange D. J., *Nucl. Instrum. Methods A*, **462** (2001) 152.
[8] Brun R. and Carminati F., CERN Program Library Long Writeup No. W5013, 1993 (unpublished).
[9] `http://root.cern.ch`.
[10] Abazov V. M. *et al.* (D0 Collaboration), *Phys. Rev. Lett.*, **101** (2008) 241801.
[11] Abazov V. M. *et al.* (D0 Collaboration), *Phys. Rev. D*, **74** (2006) 112002.
[12] Abazov V. M. *et al.* (D0 Collaboration), *Phys. Rev. Lett.*, **97** (2006) 021802.
[13] Azfar F. *et al.*, *JHEP*, **11** (2010) 158.
[14] Abulencia A. *et al.* (CDF Collaboration), *Phys. Rev. Lett.*, **97** (2006) 242003.
[15] Gronau M. and Rosner J. L., *Phys. Lett. B*, **669** (2008) 321.
[16] Lenz A. and Nierste U., arXiv:1102.4274 [hep-ph].