First model-independent Dalitz analysis of $B^{0} \rightarrow \bar{D}K^{*0}$, $D \rightarrow K_{S}^{0} \pi^{+}\pi^{-}$ decay


1University of the Basque Country UPV/EHU, 48080 Bilbao
2Beihang University, Beijing 100191
3University of Bonn, 53115 Bonn
4Budker Institute of Nuclear Physics SB RAS and Novosibirsk State University, Novosibirsk 630090
5Faculty of Mathematics and Physics, Charles University, Prague
6Chornam National University, Kiev 660-701
7University of Cincinnati, Cincinnati, Ohio 45221
8Deutsches Elektronen-Synchrotron, 22607 Hamburg
9University of Florida, Gainesville, Florida 32611
10University of Hawaii, Honolulu, Hawaii 96822
11University of Hawai‘i, Honolulu, Hawaii 96822
12University of Hawaii, Honolulu, Hawaii 96822
13University of Hawaii, Honolulu, Hawaii 96822
14University of Hawaii, Honolulu, Hawaii 96822
15University of Hawaii, Honolulu, Hawaii 96822
16High Energy Accelerator Research Organization (KEK), Tsukuba 305-0801
17IKERBASQUE, Foundation for Science, 48013 Bilbao
18Indian Institute of Technology Bhubaneswar, Satya Nagar 751007
19Indian Institute of Technology Madras, Chennai 600036
20Indiana University, Bloomington, Indiana 47408
21Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049
22Institute of High Energy Physics, Vienna 1050
23INFN - Sezione di Torino, 10125 Turin
24Institute for Theoretical and Experimental Physics, Moscow 117218
25J. Stefan Institute, 1000 Ljubljana
26Kanagawa University, Yokohama 221-8686

We report a measurement of the amplitude ratio $r_S$ of $B^0 \to D^0 K^{*0}$ and $B^0 \to D^0 K^{*0}$ decays with a Dalitz analysis of $D \to K_S^0\pi^+\pi^-$ decays, for the first time using a model-independent method. We set an upper limit $r_S < 0.87$ at the 68% confidence level, using the full data sample of $772 \times 10^6 \ BB$ pairs collected at the $\Upsilon(4S)$ resonance with the Belle detector at the KEKB $e^+e^-$ collider. This result is obtained from observables $x_+ = +0.4_{-0.6}^{+0.1} \pm 0.0$, $y_1 = -0.6_{-1.0}^{+0.8} \pm 0.1$, $x_+ = +0.1_{-0.4}^{+0.1} \pm 0.1$ and $y_1 = +0.3_{-0.8}^{+0.1} \pm 0.1$, where $x_\pm = r_S \cos(\delta_S \pm \phi_3)$, $y_\pm = r_S \sin(\delta_S \pm \phi_3)$ and $\phi_3$ ($\delta_S$) is the weak (strong) phase difference between $B^0 \to D^0 K^{*0}$ and $B^0 \to D^0 K^{*0}$.

PACS numbers:

INTRODUCTION

Determination of parameters of the standard model (SM) plays an important role in the search for new physics. In the SM, the Cabibbo-Kobayashi-Maskawa (CKM) matrix [1] gives a successful description of current all measurements...
of $CP$ violation. The $CP$-violating parameters $\phi_1$, $\phi_2$ and $\phi_3$ are the three angles of the most equilateral of the CKM unitarity triangles, of which $\phi_3 \equiv \arg \left( -V_{ud}V_{ub}^*/V_{cd}V_{cb}^* \right)$ is the least accurately determined. In the usual quark-phase convention, where the complex phase is negligible in the CKM matrix elements other than $V_{ub}$ and $V_{td}$ [2], the measurement of $\phi_3$ is equivalent to the extraction of the phase of $V_{ub}$. To date, $\phi_3$ measurements have been performed mainly with $B$ meson decays into $D^{(*)}K^{(*)}$ final states [3–12], all of which exploit the interference between the $D^{(*)0}$ and $D^{(*)0}$ decaying into a common final state. In particular, Dalitz analyses of $B^{\pm} \rightarrow D^{(*)}K^{(*)} \pm$, $D \rightarrow K_S^0\pi^-\pi^+$ provide the most precise determination of $\phi_3$. The Dalitz analysis technique for the measurement of $\phi_3$ was proposed in Ref. [13]. Belle reported the first $\phi_3$ measurement with the model-independent Dalitz analysis technique in Ref. [14], which exploits a set of measured strong phases instead of relying on a $D$ decay model into a three-body final state.

In this paper, we present the first measurement of the amplitude ratio of $B^0 \rightarrow D^0 K^{(*)0}$ and $B^0 \rightarrow D^0 K^{(*)0}$ decays with a model-independent Dalitz analysis. We reconstruct $B^0 \rightarrow DK^{(*)0}$, with $K^{(*)0} \rightarrow K^+\pi^-$ (Throughout the paper, charge-conjugate processes are implied; $K^{(*)0}$ refers to $K^*(892)^0$ and $D$ refers to either $D^0$ or $\bar{D}^0$ when the $D^0$ flavor is untagged.) Here, the flavor of the $B$ meson is identified by the kaon charge. Neutral $D$ mesons are reconstructed in the $K_S^0\pi^+\pi^-$ decay mode. The reconstructed final states are accessible through $b \rightarrow c$ and $b \rightarrow u$ processes via the diagrams shown in Fig. 1.

![FIG. 1: Diagrams for the $\bar{B}^0 \rightarrow D\bar{K}^{*0}$ decay.](image)

Since the $K^*$ has a large natural width of 50 MeV, we consider interference between the signal $B \rightarrow DK^*$ and $B \rightarrow DK\pi$, where $K\pi$ arises from a non-resonant decay or through other kaonic resonances. In this analysis, we use the variables $r_S$, $k$, and $\delta_S$ to parameterize the strong dynamics of the decay. These parameters are defined as [15]

$$r_S^2 \equiv \frac{\Gamma(B^0 \rightarrow D^0 K^+\pi^-)}{\Gamma(B^0 \rightarrow D^0 K^+\pi^-)} = \frac{\int dp A_{b\rightarrow u}^2(p)}{\int dp A_{b\rightarrow c}^2(p)},$$

$$k e^{i\delta_S} \equiv \frac{\int dp A_{b\rightarrow c}(p)A_{b\rightarrow u}(p)e^{i\delta(p)}}{\sqrt{\int dp A_{b\rightarrow c}^2(p)\int dp A_{b\rightarrow u}^2(p)}},$$

where the integration is over the $B^0 \rightarrow DK^{+}\pi^-$ Dalitz distribution region corresponding to the $K^{(*)0}$ resonance. Here, $A_{b\rightarrow c}(A_{b\rightarrow u})(p)$ is the magnitude of the amplitude for the $b \rightarrow c$ ($u$) transition and $\delta(p)$ is the relative strong phase, where the variable $p$ indicates the position within the $DK^{+}\pi^-$ Dalitz distribution. If the $B^0$ decay can be considered as a $DK^{(*)0}$ two-body decay, $r_S$ becomes the ratio of the amplitudes for $b \rightarrow u$ and $b \rightarrow c$ and $k$ becomes 1. According to a simulation study using a Dalitz model based on the measurements in Ref. [16], the value of $k$ is $0.95 \pm 0.03$ within the phase space of the $DK^{(*)0}$ resonance. The value of $r_S$ is expected to be around 0.4, which corresponds naively to $|V_{ub}V_{cs}|/|V_{cb}V_{us}|$ but also depends on strong interaction effects.

**THE MODEL-INDEPENDENT DALITZ ANALYSIS TECHNIQUE**

The amplitude of the $B^0 \rightarrow DK^{(*)0}$, $D \rightarrow K_S^0\pi^+\pi^-$ decay is a superposition of the $B^0 \rightarrow D^0 K^{(*)0}$ and $B^0 \rightarrow D^0 K^{(*)0}$ amplitudes

$$A_B(m_+^2, m_-^2) = \bar{A} + r_S e^{i(\delta_S + \phi_3)} A,$$

where $m_+^2$ and $m_-^2$ are the squared invariant masses of the $K^0_0\pi^+$ and $K^0_0\pi^-$ combinations, respectively, $\bar{A} = \bar{A}(m_+^2, m_-^2)$ is the amplitude of the $B^0 \rightarrow D^0 K^{(*)0}$, $D^0 \rightarrow K_S^0\pi^+\pi^-$ decay and $A = A(m_+^2, m_-^2)$ is the amplitude of the $B^0 \rightarrow D^0 K^{(*)0}$, $D^0 \rightarrow K_S^0\pi^+\pi^-$ decay. In the case of $CP$ conservation in the $D$ decay, we have $A(m_+^2, m_-^2) = \bar{A}(m_+^2, m_-^2)$. The Dalitz distribution density of the $D$ decay from $B^0 \rightarrow DK^{(*)0}$ is given by

$$P_B = |A_B|^2 = |\bar{A} + r_S e^{i(\delta_S + \phi_3)} A|^2 = \bar{P} + r_S^2 P + 2k \sqrt{\bar{PP}}(x + C + y + S),$$

where $C$ and $S$ are the constants.
where \( P(m_+^2, m_-^2) = |A|^2, \tilde{P}(m_+^2, m_-^2) = |\tilde{A}|^2 \), and
\[
    x_+ = r_S \cos(\delta_S + \phi_3), \quad y_+ = r_S \sin(\delta_S + \phi_3),
\]
\[
    x_- = r_S \cos(\delta_S - \phi_3), \quad y_- = r_S \sin(\delta_S - \phi_3).
\]
The functions \( C(m_+^2, m_-^2) \) and \( S(m_+^2, m_-^2) \) are the cosine and sine of the strong-phase difference \( \delta_D(m_+^2, m_-^2) = \arg \tilde{A} - \arg A \) between the \( D^0 \rightarrow K_S^0 \pi^+ \pi^- \) and \( D^0 \rightarrow K_S^0 \pi^+ \pi^- \) amplitudes. Here, we have used the definition of \( k \) given in Eq. (2). The equations for the charge-conjugate mode \( B^0 \rightarrow D K^{*0} \) are obtained with the substitution \( \phi_3 \leftrightarrow -\phi_3 \) and \( A \leftrightarrow \tilde{A} \); the corresponding parameters that depend on the \( B^0 \) decay amplitude are

Using \( B^0 \) and \( \bar{B}^0 \) decays, one can obtain \( r_S, \phi_3 \) and \( \delta_S \) separately.

Up to this point, the description of the model-dependent and model-independent techniques is the same. The model-dependent analysis deals directly with the Dalitz distribution density and the functions \( C \) and \( S \) are obtained from a model based upon a fit to the \( D^0 \rightarrow K_S^0 \pi^+ \pi^- \) amplitude. In the model-independent approach, the Dalitz plot is divided into \( 2N \) bins symmetric under the exchange \( m_+^2 \leftrightarrow m_-^2 \). The bin index \( i \) ranges from \(-N \) to \( N \) (excluding 0); the exchange \( m_+^2 \leftrightarrow m_-^2 \) corresponds to the exchange \( i \leftrightarrow -i \). The expected number of signal events in bin \( i \) of the Dalitz distribution of the \( D \) mesons from \( B^0 \rightarrow DK^{*0} \) is

\[
    N_i^\pm = h_B^{\pm} \left[ K_{\pm i} + r_S^2 K_{\mp i} + 2k \sqrt{K_{\pm i} K_{\mp i}} (x_+ c_i \pm y_+ s_i) \right],
\]
where \( N_i^{\pm(-)} \) stands for the number of \( B^0(\bar{B}^0) \) meson decays, \( h_B^{\pm(-)} \) is the normalization constant and \( K_{\pm i} \) is the number of events in the \( i \)th bin of a flavor-tagged \( D^0 \rightarrow K_S^0 \pi^+ \pi^- \) decays measured with a sample of inclusively reconstructed \( D^0 \rightarrow D^0 \pi^+ \pi^- \) decays. The values of \( K_i \) are measured from a sample of flavor-tagged \( D^0 \) mesons obtained by reconstructing \( D^{\pm} \rightarrow D \pi^{\pm} \) decays. The terms \( c_i \) and \( s_i \) are the amplitude-weighted averages of the functions \( C \) and \( S \) over the bin:

\[
    c_i = \frac{\int_{D_i} |A| \|\tilde{A}\| C dD}{\sqrt{\int_{D_i} |A|^2 dD \int_{D_i} \|\tilde{A}\|^2 dD}}.
\]

Here, \( D \) represents the Dalitz plane and \( D_i \) is the bin over which the integration is performed. The terms \( s_i \) are defined similarly with \( C \) substituted by \( S \). The absence of \( CP \) violation in the \( D \) decay implies \( c_i = c_{-i} \) and \( s_i = -s_{-i} \).

The values of \( c_i \) and \( s_i \) measurements are included, the set of relations defined by Eq. (7) contains only three free parameters \( (x, y, \) and \( h_B) \) for each \( B^0 \) and \( \bar{B}^0 \) and can be solved using a maximum likelihood method to extract the values of \( \phi_3, \delta_S \) and \( r_S \). We have neglected charm-mixing effects in \( D \) decays from both the \( B^0 \rightarrow DK^{*0} \) process and in the quantum-correlated \( DD \) production [19]. In principle, the set of relations defined by Eq. (7) can be solved without external constraints on \( c_i \) and \( s_i \) for \( N \geq 2 \). However, due to the small value of \( r_S \), there is very little sensitivity to the \( c_i \) and \( s_i \) parameters in \( B^0 \rightarrow DK^{*0} \) decays, which results in a reduction in the precision of the other parameters [20].

**ANALYSIS PROCEDURE**

Equation (7) is the key relation used in the analysis, but it holds only if there is no background, no non-uniformity in the Dalitz acceptance and no crossfeed between bin. (Crossfeed is due to invariant-mass resolution and radiative corrections.) In this section, we outline the procedures to account for the acceptance and crossfeed.

First, we discuss the effect of the variation of the efficiency profile over the Dalitz plane. We note that Eq. (4) does not change under the transformation \( P \rightarrow eP \) when the efficiency profile \( e(m_+^2, m_-^2) \) is symmetric: \( e(m_+^2, m_-^2) = e(m_-^2, m_+^2) \). The effect of non-uniform efficiency over the Dalitz plane cancels when using a flavor-tagged \( \tilde{D} \) sample with kinematic properties that are similar to the sample from the signal \( B \) decay. This approach allows for the removal of the systematic uncertainty associated with the possible inaccuracy of the detector acceptance description in the Monte Carlo (MC) simulation. With the efficiency taken into account (that is, in general non-uniform across the bin region), the number of events detected is

\[
    N' = \int p(D) e(D) dD.
\]
Clearly, the efficiency does not factorize. One can use an efficiency averaged over the bin, then correct for it in the analysis:

\[
\bar{\epsilon}_i = \frac{N'_i}{N_i} = \frac{\int p(D)\epsilon(D)dD}{\int p(D)dD}.
\]

(10)

The averaged efficiency \(\bar{\epsilon}_i\) can be determined from MC. The assumption that the efficiency profile depends only on the \(D\) momentum is tested using MC simulation and the residual difference is treated as a systematic uncertainty. This correction cannot be calculated in a completely model-independent way, since the correction terms include the amplitude variation inside the bin. Calculations using the Belle \(D \rightarrow K_S^0\pi^+\pi^-\) model [5] show that this correction is negligible even for very large non-uniformity of the efficiency profile.

There are two sources of crossfeed: momentum resolution and flavor misidentification. Momentum resolution leads to migration of events among the bins. In the binned approach, this effect can be corrected in a non-parametric way. The migration can be described by a linear transformation of the number of events in each bin

\[
N'_i = \sum \alpha_{ik} N_k,
\]

(11)

where \(N_k\) is the number of events that bin \(k\) would contain without the crossfeed and \(N'_i\) is the reconstructed number of events in bin \(i\). The crossfeed matrix \(\alpha_{ik}\) is nearly the unit matrix; it is obtained from a signal MC simulation generated with the amplitude model reported in Ref. [5]. Most of the off-diagonal elements are null; only a few have values \(\alpha_{ik} \leq 0.04\). In the case of a \(D \rightarrow K_S^0\pi^+\pi^-\) decay from a \(B\), the crossfeed depends on the parameters \(x\) and \(y\). However, this is a minor correction to an already small effect and so is neglected.

The final effect to be considered is due to misidentification of the \(B\) flavor. Double misidentification in \(K^{*0}\) reconstruction from \(K^+\pi^-\) leads to migration of events between \(N_i^+ \leftrightarrow N_i^-\) due to assignment of the wrong flavor to the \(B\) candidate. If the fraction of doubly-misidentified events is \(\beta\), the number of events in each bin can be written as

\[
N'_i = N_i^\pm + \beta N_i^\mp.
\]

(12)
The value of $\beta$ is obtained from MC simulation and is found to be $(0.12 \pm 0.01)\%$. Therefore, the effect of flavor misidentification is neglected in this analysis.

**Fit Procedure**

We fit the data distribution in each bin separately, with the number of signal and background events as free parameters. The values of $N_i$ found in Eq. (7) to obtain the parameters $(x_\pm, y_\pm)$. This is accomplished by minimizing the negative logarithmic likelihood of the form

$$-2\log L(x, y) = -2\sum_i \log p(\langle N_i \rangle (x, y), N_i, \sigma_{N_i}), \quad (13)$$

where $\langle N_i \rangle$ is the expected number of signal events in the bin $i$ obtained from Eq. (7). Here, $N_i$ and $\sigma_{N_i}$ are the observed number of events in data and the uncertainty on $N_i$, respectively. If the probability density function (PDF) $p$ is Gaussian, this procedure is equivalent to a $\chi^2$ fit; however, the assumption of the Gaussian distribution introduces a bias in the case of low yield in some bins.

The procedure described above does not make any assumptions about the Dalitz distribution of the background events, since the fits in each bin are independent. Thus, there is no associated systematic uncertainty. However, in the case of a small number of events and many background components, this can be a limiting factor. Our approach is to use the combined fit with a common likelihood for all bins. The background contribution then must be accounted for in the calculation of the values $N_i$. Statistically, the most effective way of calculating the number of signal events is to perform, in each bin $i$ of the Dalitz plot, an unbinned fit in the variables used to distinguish the signal from the background. In this analysis, we obtain the $CP$ violating parameters $(x_\pm, y_\pm)$ from a combined fit in all bins. The relative numbers of background events in the bins of such a fit are constrained externally from MC samples. In the combined fit, the expected numbers of events $\langle N_i \rangle$ as functions of $(x, y)$ are fixed from a large flavor-tagged sample and external observables. Thus, the variables $(x, y)$ become free parameters of the combined likelihood fit and the assumption that the signal yield obeys a Gaussian distribution is not needed.

**EVENT RECONSTRUCTION AND SELECTION**

This analysis is based on a data sample that contains $772 \times 10^6 \ BB$ pairs, collected with the Belle detector at the KEKB asymmetric-energy $e^+e^-$ (3.5 on 8 GeV) collider [23] operating at the $\Upsilon(4S)$ resonance. The Belle detector is a large-solid-angle magnetic spectrometer that consists of a silicon vertex detector, a 50-layer central drift chamber (CDC), an array of aerogel threshold Cherenkov counters (ACC), a barrel-like arrangement of time-of-flight scintillation counters (TOF), and an electromagnetic calorimeter comprised of CsI(Tl) crystals located inside a superconducting solenoid coil that provides a 1 T magnetic field. An iron flux-return located outside of the coil is instrumented to detect $K^0_L$ mesons and to identify muons. The detector is described in detail elsewhere [24].

We reconstruct $B^0 \to \bar{D}K^{*0}$ events with $K^{*0} \to K^+\pi^-$ and $D \to K^0_S\pi^+\pi^-$. The event selection described below is developed from studies of off-resonance data and MC simulated events.

The $K^0_S$ candidates are identified using the output of a neural network. Inputs to the network for a pair of oppositely-charged pions are the invariant mass, 20 kinematic parameters and particle identification (PID) information from the ACC, TOF and the ionization energy loss in the CDC. The $K^0_S$ selection has a simulated purity of 92.2% and an efficiency of 75.1%. Charged kaon and pion candidates are identified using PID information. The efficiency is 80–90% and the probability of misidentification is 6–10%, depending upon the momentum of hadrons and obtained using dedicated data control samples. We reconstruct neutral $D$ mesons by combining a $K^0_S$ candidate with a pair of oppositely-charged pion candidates. We require that the invariant mass be within $\pm 15 \mathrm{MeV}/c^2$ ($\pm 3\sigma$) of the nominal $D^0$ mass. $K^{*0}$ candidates are reconstructed from $K^+\pi^-$ pairs. We require that the invariant mass be within $\pm 50 \mathrm{MeV}/c^2$ of the nominal $K^{*0}$ mass. We combine $D$ and $K^{*0}$ candidates to form $B^0$ mesons. Candidate events are identified by the energy difference $\Delta E \equiv \sum_i E_i - E_b$, and the beam-constrained mass $M_{bc}c^2 \equiv \sqrt{E_b^2 - |c\sum_i \vec{p}_i|^2}$, where $E_b$ is the beam energy and $\vec{p}_i$ and $E_i$ are the momenta and energies, respectively, of the $B^0$ meson decay products in the $e^+e^-$ center-of-mass (CM) frame. We select events with $5.20 \mathrm{GeV}/c^2 < M_{bc} < 5.29 \mathrm{GeV}/c^2$ and $-0.10 \mathrm{GeV} < \Delta E \ < 0.15 \mathrm{GeV}$.

Among other $B$ decays, the most serious background is from $B^0$ decaying to the same final state as $B^0 \to \bar{D}K^{*0}$. To suppress this background, we exclude candidates for which the invariant mass of the $K^{*0}\pi^+$ system is within
Fisher discriminants based on modified Fox-Wolfram moments [25].

The angle in the CM frame between the thrust axes of the B decay and that of remaining particles.

The signed difference of the vertices between the B candidate and the remaining charged tracks.

The distance of closest approach between the trajectories of the $K^*$ and $D$ candidates.

The expected flavor dilution factor described in Ref. [26].

The angle $\theta$ between the $B$ meson momentum direction and the beam axis in the CM frame.

The angle between the $D$ and $\Upsilon(4S)$ directions in the rest frame of the B candidate.

The projection of the sphericity vector with the largest eigenvalue onto the $e^+e^-$ beam direction.

The angle of the sphericity vector with the largest eigenvalue with respect to that of the remaining particles.

The angle of the sphericity vector with the second largest eigenvalue.

The angle of the sphericity vector with the smallest eigenvalue.

The magnitude of the thrust of the particles not used to reconstruct the signal.

<table>
<thead>
<tr>
<th>TABLE I: Variables used for $q\bar{q}$ suppression.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>11</td>
</tr>
<tr>
<td>12</td>
</tr>
</tbody>
</table>

$\pm 4 \text{ MeV}/c^2$ of the nominal $D^+$ mass. This criterion leads to a negligible contamination from this mode and a relative loss of 0.6% in the signal efficiency.

The large combinatorial background of true $D^0$ and random $K^+$ and $\pi^-$ combinations from the $e^+e^- \to c\bar{c}$ process and other $BB$ decays is reduced if $D^0$ candidates that are a decay product of $D^{*+} \to D^0\pi^+$ are eliminated. We use the mass difference $\Delta M$ between the $[K_S^0\pi^+\pi^-]_{D}\pi^+$ and $[K_S^0\pi^+\pi^-]_{D}$ systems for this purpose: if $\Delta M > 0.15 \text{ GeV}/c^2$ for any additional $\pi^+$ candidate not used in the $B$ candidate reconstruction, the event is retained. This requirement removes 19% of $c\bar{c}$ background and 11% of $BB$ background according to MC simulation. The relative loss in signal efficiency is 5.5%.

In the rare case where there are multiple candidates in an event, the candidate with $M_{bc}$ closest to the nominal value is chosen. The relative loss in signal efficiency is 0.8%.

To discriminate signal events from the large combinatorial background dominated by the two-jet-like $e^+e^- \to q\bar{q}$ continuum process, where $q$ indicates $u, d, s$ or $c$, a multivariate analysis is performed using the 12 variables introduced in Table I. To effectively combine these 12 variables, we employ the NeuroBayes neural network package [27]. The NeuroBayes output is denoted as $C_{NB}$ and lies within the range $[-1, 1]$; events with $C_{NB} \sim 1$ are signal-like and events with $C_{NB} \sim -1$ are $q\bar{q}$-like. Training of the neural network is performed using signal and $q\bar{q}$ MC samples. The $C_{NB}$ distribution of signal events peaks at $C_{NB} \sim 1$ and is therefore difficult to represent with a simple analytic function. However, the transformed variable

$$C'_{NB} = \ln \frac{C_{NB} - C_{NB, \text{low}}}{C_{NB, \text{high}} - C_{NB}},$$

where $C_{NB, \text{low}} = -0.6$ and $C_{NB, \text{high}} = 0.9992$, has a distribution that can be modeled by a Gaussian for signal as well as background. The events with $C_{NB} < C_{NB, \text{low}}$ are rejected; the relative loss in signal efficiency is 7.4%. The fit region is defined as $\Delta E \in [-0.1, 0.3]$ GeV and $M_{bc} > 5.21 \text{ GeV}/c^2$.

The number of signal events is obtained by fitting the three-dimensional distribution of variables $M_{bc}$, $\Delta E$, and $C'_{NB}$ using the extended maximum likelihood method. We form three-dimensional PDFs for each component as the product of one-dimensional PDFs for $\Delta E$, $M_{bc}$ and $C'_{NB}$, since the correlations among the variables are found to be small.

Backgrounds are divided into the following components:

- **Continuum background from $q\bar{q}$ events.**

- **$B\bar{B}$ background**, in which the tracks forming the $B^0 \to DK^{*0}$ candidate come from decays of both $B$ mesons in the event. The number of possible $B$ decay combinations that contribute to this background is large; therefore, both the Dalitz distribution and distribution of the fit parameters are quite smooth. In this analysis, $BB$ backgrounds are further subdivided into two components: events reconstructed with a true $D \to K_S^0\pi^+\pi^-$ decay, referred to as $D_{true} BB$ background, and those reconstructed with a combinatorial $D$ candidate, referred to as $D_{fake} BB$ background.
the additional requirements. Fits to Figs. 4 and 5, respectively. The plots show the projections of the data and the fitting model on the $\Delta$ total yields of $(\text{of the of the Table II. In this study, these (of the ARGUS function for

the frequentist approach with Feldman-Cousins ordering [30], which is described in Sec. .

shape parameters of the PDFs are fixed from MC samples.

for $\bar{B}$ over the Dalitz plot in such an approach. This procedure is justified for background that is either well separated from

background, an exponential function for the $\bar{B}$ signal, a Crystal Ball function [28] for $D_{\text{true}}$ $B\bar{B}$ background, an ARGUS function [29] for $D_{\text{fake}}$ $B\bar{B}$ background, an ARGUS function for $q\bar{q}$ background, a sum of a Gaussian and ARGUS functions for $D^0\rho^0$ background and a Gaussian for $D^0a_1^+$ background. For each component, the $C'_{\text{NB}}$ PDF is the sum of a Gaussian and bifurcated Gaussian. The shape parameters of the PDFs are fixed from MC samples.

The Dalitz distributions of the background components are discussed in the next section. The numbers of events in

each bin are free parameters in the fit; thus, there is no uncertainty due to the modeling of the background distribution

over the Dalitz plot in such an approach. This procedure is justified for background that is either well separated from

the signal (such as peaking $B\bar{B}$ background) or is constrained by a much larger number of events than the signal

(such as $q\bar{q}$ background). The results of the fit over the full Dalitz plot are shown in Fig. 3. We obtain a total of

$44.2_{-12.1}^{+13.3}$ signal events. The statistical significance is $2.8\sigma$ relative to the no-signal hypothesis.

DATA FIT IN BINS

A combined fit is performed to obtain the $B^0 \to D K^{*0}$ yield in each bin. The combined fit constrains the amount

of the $D_{\text{true}}$ $B\bar{B}$ background in bins from the ratio of $E^0$ ($K_i$) and $D^0$ ($K_{-i}$) from the generic MC and the amount

of the $D_{\text{fake}}$ $B\bar{B}$, $q\bar{q}$, $D^0\rho^0$ and $D^0a_1^+$ backgrounds in bins from the MC and takes the $(x, y, \rho)$ variables as free

parameters. Fits to $B^0$ and $B^0$ data are performed separately for each bin of the Dalitz distribution as shown in

Figs. 4 and 5, respectively. The plots show the projections of the data and the fitting model on the $\Delta E$ variable, with

the additional requirements $M_{bc} > 5.27\text{ GeV}/c^2$ and $C'_{\text{NB}} > 2$. In this fit, the additional free parameters are the
total yields of $D_{\text{true}}$ $B\bar{B}$, $D_{\text{fake}}$ $B\bar{B}$, $q\bar{q}$ and peaking $B\bar{B}$ backgrounds over the entire Dalitz plane. The values of the

$(x, y)$ parameters and their statistical correlations, obtained from the combined fit for the signal sample, are given in
Table II. In this study, these $(x, y)$ values from the likelihood distribution of the combined fit are corrected using the
frequentist approach with Feldman-Cousins ordering [30], which is described in Sec.
FIG. 4: Projections of the combined fit of the $B^0 \rightarrow DK^{*0}$ sample on $\Delta E$ for each Dalitz bin, with the $M_{bc} > 5.27$ GeV/$c^2$ and $C'_{NB} > 2$ requirements. The fill styles for the signal and background components are the same as in Fig. 3.

SYSTEMATIC UNCERTAINTIES

The systematic uncertainties of $(x, y)$ are obtained by taking deviation from the default procedure under varied assumptions. The systematic uncertainties are summarized in Table III; most are negligible compared to the statistical uncertainty. There is an uncertainty due to the Dalitz efficiency variation because of the difference in average efficiency over each bin for the flavor-tagged and $B^0 \rightarrow DK^{*0}$ samples. A maximum difference of 1.5% is obtained in a MC study. The uncertainty is taken as the maximum of two quantities:

- the root mean square of $x$ and $y$ from smearing the numbers of events in the flavor-tagged sample $K_1$ by 1.5%,
FIG. 5: Projections of the combined fit of the $B^0 \rightarrow D\bar{K}^{*0}$ sample on $\Delta E$ for each Dalitz bin, with the $M_{bc} > 5.27$ GeV/c$^2$ and $C_{NB} > 2$ requirements. The fill styles for the signal and background components are the same as in Fig. 3.

or

- the bias in $x$ and $y$ between the fits with and without efficiency correction for $K_i$ obtained from signal MC.

The uncertainty due to crossfeed of events between bins is estimated by taking the bias between the fits with and without the correction. The uncertainties due to the fixed parameterization of the signal and background PDFs are estimated by varying them by $\pm 1\sigma$. The uncertainty due to the $C'_{NB}$ PDF distributions for $B\bar{B}$ is estimated by replacing them with the signal $C'_{NB}$ PDF. The uncertainty due to the $D_{true}$ and $D_{fake}$ $BB$ fractions is estimated by varying them between 0 and 1. The uncertainty arising from the finite sample of flavor-tagged $D \rightarrow K_S^0 \pi \pi$ decays is evaluated by varying the values of $K_i$ within their statistical uncertainties. The uncertainty due to $k$ in Eq. (2)
is evaluated by varying the value of $k$ within its error $[16]$. The uncertainty due to the limited precision of $c_i$ and $s_i$ parameters is obtained by smearing the $c_i$ and $s_i$ values within their total errors and repeating the fits for the same experimental data. Total systematic uncertainties in the $(x, y)$ are obtained by summing all uncertainties in quadrature and listed in Table III.

### RESULT

We use the frequentist approach with Feldman-Cousins ordering to obtain the physical parameters $\mu = (\phi_3, r_S, \delta_S)$ (or true parameters $\mu = \mu_{\text{true}} = (x_-, y_-, x_+, y_+)$) from the measured parameters $z = z_{\text{meas}} = (x_-, y_-, x_+, y_+)$ taken from the likelihood distribution. In essence, the confidence level $\alpha$ for a set of physical parameters $\mu$ is calculated as

$$
\alpha(\mu) = \frac{\int_{D(\mu)} p(z \mid \mu) dz}{\int_{\infty} p(z \mid \mu) dz},
$$

where $p(z \mid \mu)$ is the probability density to obtain the measurement $z$ given the set of true parameters $\mu$. The integration domain $D(\mu)$ is given by the likelihood ratio (Feldman-Cousins) ordering:

$$
\frac{p(z \mid \mu)}{p(z \mid \mu_{\text{best}}(z))} > \frac{p(z_0 \mid \mu)}{p(z_0 \mid \mu_{\text{best}}(z_0))},
$$

where $\mu_{\text{best}}(z)$ is $\mu$ that maximizes $p(z \mid \mu)$ for the given $z$, and $z_0$ is the result of the data fit. This PDF is taken from MC pseudo-experiments.

Systematic errors in $\mu$ are obtained by varying the measured parameters $z$ within their systematic errors assuming a Gaussian distribution. In this calculation we assume that the correlations of systematic errors between the $B^0$ and $\bar{B}^0$ samples are small.

$B^0$ and $\bar{B}^0$ samples As a result of this procedure, we obtain the confidence levels (C.L.) for $(x, y)$ and the physical parameter $r_S$. The C.L. contours on $(x, y)$ are shown in Fig. 6. The likelihood profile for $r_S$ is shown in Fig. 7. The final results are:

$$
x_- = +0.4^{+1.0+0.0}_{-0.6-0.1} \pm 0.0,
$$

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$x_-$</th>
<th>$y_-$</th>
<th>corr.$(x_-, y_-$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$+0.29 \pm 0.32$</td>
<td>$-0.33 \pm 0.41$</td>
<td>+7.0%</td>
</tr>
</tbody>
</table>

TABLE II: $(x, y)$ parameters and their statistical correlations from the combined fit of the $B^0 \rightarrow DK^{*0}$ sample. The error is statistical. The values and errors are obtained from the likelihood distribution.

<table>
<thead>
<tr>
<th>Source of uncertainty</th>
<th>$\Delta x_-$</th>
<th>$\Delta y_-$</th>
<th>$\Delta x_+$</th>
<th>$\Delta y_+$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dalitz efficiency</td>
<td>$\pm 0.00$</td>
<td>$\pm 0.01$</td>
<td>$\pm 0.01$</td>
<td>$\pm 0.00$</td>
</tr>
<tr>
<td>Crossfeed between bins</td>
<td>$\pm 0.00$</td>
<td>$\pm 0.00$</td>
<td>$\pm 0.00$</td>
<td>$\pm 0.00$</td>
</tr>
<tr>
<td>PDF shape</td>
<td>$\pm 0.07$</td>
<td>$\pm 0.01$</td>
<td>$\pm 0.04$</td>
<td>$\pm 0.00$</td>
</tr>
<tr>
<td>Flavor-tag statistics</td>
<td>$\pm 0.00$</td>
<td>$\pm 0.00$</td>
<td>$\pm 0.00$</td>
<td>$\pm 0.00$</td>
</tr>
<tr>
<td>$c_i, s_i$ precision</td>
<td>$\pm 0.03$</td>
<td>$\pm 0.08$</td>
<td>$\pm 0.05$</td>
<td>$\pm 0.08$</td>
</tr>
<tr>
<td>$k$ precision</td>
<td>$\pm 0.00$</td>
<td>$\pm 0.01$</td>
<td>$\pm 0.00$</td>
<td>$\pm 0.00$</td>
</tr>
<tr>
<td>Total without $c_i, s_i$ precision</td>
<td>$\pm 0.01$</td>
<td>$\pm 0.07$</td>
<td>$\pm 0.04$</td>
<td>$\pm 0.04$</td>
</tr>
<tr>
<td>Total</td>
<td>$\pm 0.03$</td>
<td>$\pm 0.12$</td>
<td>$\pm 0.05$</td>
<td>$\pm 0.09$</td>
</tr>
</tbody>
</table>

TABLE III: Systematic uncertainties in the $(x, y)$ measurement for the $B^0 \rightarrow DK^{*0}$ mode. Values are rounded to two significant digits and those less than 0.005 are quoted as 0.00.
FIG. 6: C.L. contours for \((x_-, y_-)\) (blue) and \((x_+, y_+)\) (red). The dots show the most probable \((x, y)\) values; the lines show the 68\% contours. The fluctuations arise from the statistics of the pseudo-experiments and C.L. step used.

FIG. 7: Likelihood profile for \(r_S\). The blue points are for \(\bar{B}^0\) \((x_-, y_-)\), red are for \(B^0\) \((x_+, y_+)\) and black are \(\bar{B}^0\) and \(B^0\) combined. The two horizontal lines show 68\% and 95\% C.L.

\[
y_- = -0.6^{+0.8+0.1}_{-1.0-0.0} \pm 0.1, \tag{18}
\]
\[
x_+ = +0.1^{+0.7+0.0}_{-0.4-0.1} \pm 0.1, \tag{19}
\]
\[
y_+ = +0.3^{+0.5+0.0}_{-0.8-0.3} \pm 0.1, \tag{20}
\]
\[
r_S < 0.87 \quad \text{at 68\% C.L.,} \tag{21}
\]

where the first error is statistical, the second is systematic without uncertainties in \((c_i, s_i)\), and the third is from the \((c_i, s_i)\) precision from CLEO.
CONCLUSION

We report the first measurement of the amplitude ratio $r_S$ using a model-independent Dalitz analysis of $D \to K_S^0 \pi^+ \pi^-$ decays in the process $B^0 \to D K^{*0}$ with the full data sample of 711 fb$^{-1}$ corresponding to $772 \times 10^6$ $B\bar{B}$ pairs collected by the Belle detector at the $Y(4S)$ resonance. Model independence is achieved by binning the Dalitz plot of the $D \to K_S^0 \pi^+ \pi^-$ decay and using the strong-phase coefficients with binning as in the CLEO experiment [18]. We obtain the value $r_S < 0.87$ at 68% C.L. This measurement results in lower statistical precision than the model-dependent measurement from BaBar with the $B^0 \to D K^0$ mode [8] despite the larger data sample due to the smaller $B^0 \to D K^{*0}$ signal observed. The result is consistent with the most precise $r_S$ measurement reported by the LHCb Collaboration [31] of $r_S = 0.240^{+0.055}_{-0.048}$ that uses $B^0 \to [K^+ K^-, K^0 \pi^+ \pi^+, \pi^+ \pi^-] D K^{*0}$ decays.

We have confirmed the feasibility of the model-independent Dalitz analysis method with neutral $B \to D K^*$. The value of $r_S$ indicates the sensitivity of the decay to $\phi_3$ because the statistical uncertainty is proportional to $1/r_S$. This technique will be used to measure $\phi_3$ at future B-factories like Belle II. In future high statistics experiments such as Belle II and the LHCb upgrade, this method will give a precise and model independent determination of $\phi_3$.