First measurement of ϕ_{3} with a model-independent Dalitz plot analysis of B^{±}→DK^{±}, D→K_{S}^{0}π^{+}π^{-} decay

H. Aihara *et al.* (Belle Collaboration)

Phys. Rev. D **85**, 112014 — Published 26 June 2012

DOI: 10.1103/PhysRevD.85.112014

Belle Preprint 2012-12
KEK Preprint 2012-4

First Measurement of ϕ_3 with a Model-independent Dalitz Plot Analysis of $B^\pm \to DK^\pm$, $D \to K_S^0\pi^+\pi^-$ Decay

H. Aihara,[50] K. Arinstein,[2] D. M. Asner,[38] V. Aulchenko,[2] T. Aushev,[15] A. M. Bakich,[44] K. Belous,[14] B. Bhuyan,[10] M. Bischofberger,[30] A. Bondar,[2] G. Bonvicini,[55] A. Bozek,[33] M. Bračko,[25,16] T. E. Browder,[7] M.-C. Chang,[5] P. Chang,[32] B. G. Cheon,[6] K. Chilikin,[15] R. Chistov,[15] K. Cho,[19] Y. Choi,[43] J. Dalseno,[26,46] Z. Doležal,[3] A. Drutskoy,[15] S. Eidelman,[2] D. Epifanov,[2] J. E. Fast,[38] M. Feindt,[18] V. Gaur,[45] N. Gabyshev,[2] A. Garmash,[2] Y. M. Goh,[6] B. Golob,[23,16] J. Haba,[8] H. Hayashii,[30] Y. Horii,[29] Y. Hoshi,[48] W.-S. Hou,[32] Y. B. Hsiung,[32] H. J. Hyun,[21] T. Iijima,[29,28] A. Ishikawa,[49] R. Itoh,[8] M. Iwabuchi,[57] T. Julius,[27] J. H. Kang,[57] T. Kawasaki,[35] C. Kiesling,[26] H. J. Kim,[21] H. O. Kim,[21] J. B. Kim,[20] J. H. Kim,[19] K. T. Kim,[20] M. J. Kim,[21] Y. J. Kim,[19] K. Kinoshita,[4] B. R. Ko,[20] S. Koblitz,[26] P. Kodyš,[3] S. Korpar,[25,16] P. Križan,[23,16] P. Krokovny,[2] B. Kronenbitter,[18] T. Kuhr,[18] T. Kumita,[52] A. Kuzmin,[2] Y.-J. Kwon,[57] S.-H. Lee,[20] J. Li,[42] Y. Li,[54] J. Libby,[11] C. Liu,[41] Y. Liu,[4] Z. Q. Liu,[12] D. Liventsev,[15] R. Louvot,[22] D. Matvienko,[2] K. Miyabayashi,[30] H. Miyata,[35] R. Mizuk,[15] G. B. Mohanty,[45] A. Moll,[26,46] T. Mori,[28] N. Muramatsu,[40] Y. Nagasaka,[9] E. Nakano,[37] M. Nakao,[8] Z. Natkaniec,[33] S. Nishida,[8] O. Nitoh,[53] S. Ogawa,[47] T. Ohshima,[28] S. Okuno,[17] S. L. Olsen,[42,7] G. Pakhlova,[15] C. W. Park,[43] H. Park,[21] H. K. Park,[21] K. S. Park,[43] T. K. Pedlar,[24] R. Pestotnik,[16] M. Petrič,[16] L. E. Piilonen,[54] A. Poluektov,[2] K. Prothmann,[26,46] M. Ritter,[26] M. Röhrken,[18] M. Rozanska,[33] S. Ryu,[42] H. Sahoo,[7] Y. Sakai,[8] T. Sanuki,[49] Y. Sato,[49] O. Schneider,[22] C. Schwanda,[13] A. J. Schwartz,[4] K. Senyo,[56] O. Seon,[28] M. E. Sevior,[27] M. Shapkin,[14] T.-A. Shibata,[51] J.-G. Shiu,[32] B. Shwartz,[2] A. Sibidanov,[44] F. Simon,[26,46] J. B. Singh,[39] P. Smerkol,[16] Y.-S. Sohn,[57] A. Sokolov,[14] E. Solovieva,[15] S. Stanič,[36] M. Starič,[16] K. Sumisawa,[8] T. Sumiyoshi,[52] G. Tatishvili,[38] K. Trabelsi,[8] M. Uchida,[51] S. Uehara,[8] Y. Unno,[6] S. Uno,[8] P. Urquijo,[1] P. Vanhoefer,[26] G. Varner,[7] K. E. Varvell,[44] A. Vinokurova,[2] V. Vorobyev,[2] C. H. Wang,[31] M.-Z. Wang,[32] P. Wang,[12] Y. Watanabe,[17] K. M. Williams,[54] E. Won,[20] B. D. Yabsley,[44] H. Yamamoto,[49] J. Yamaoka,[7] Y. Yamashita,[34] C. Z. Yuan,[12] Z. P. Zhang,[41] V. Zhilich,[2] V. Zhulanov,[2] and A. Zupanc[18]

(The Belle Collaboration)

[1] *University of Bonn, Bonn*
[2] *Budker Institute of Nuclear Physics SB RAS and Novosibirsk State University, Novosibirsk 630090*
[3] *Faculty of Mathematics and Physics, Charles University, Prague*
[4] *University of Cincinnati, Cincinnati, Ohio 45221*
[5] *Department of Physics, Fu Jen Catholic University, Taipei*
[6] *Hanyang University, Seoul*
[7] *University of Hawaii, Honolulu, Hawaii 96822*
[8] *High Energy Accelerator Research Organization (KEK), Tsukuba*
[9] *Hiroshima Institute of Technology, Hiroshima*
[10] *Indian Institute of Technology Guwahati, Guwahati*
[11] *Indian Institute of Technology Madras, Madras*
[12] *Institute of High Energy Physics, Chinese Academy of Sciences, Beijing*
[13] *Institute of High Energy Physics, Vienna*
[14] *Institute of High Energy Physics, Protvino*
[15] *Institute for Theoretical and Experimental Physics, Moscow*
[16] *J. Stefan Institute, Ljubljana*
[17] *Kanagawa University, Yokohama*
[18] *Institut für Experimentelle Kernphysik, Karlsruher Institut für Technologie, Karlsruhe*
[19] *Korea Institute of Science and Technology Information, Daejeon*
[20] *Korea University, Seoul*
[21] *Kyungpook National University, Taegu*
[22] *École Polytechnique Fédérale de Lausanne (EPFL), Lausanne*
[23] *Faculty of Mathematics and Physics, University of Ljubljana, Ljubljana*
[24] *Luther College, Decorah, Iowa 52101*
[25] *University of Maribor, Maribor*
[26] *Max-Planck-Institut für Physik, München*
[27] *University of Melbourne, School of Physics, Victoria 3010*
[28] *Graduate School of Science, Nagoya University, Nagoya*
[29] *Kobayashi-Maskawa Institute, Nagoya University, Nagoya*
[30] *Nara Women's University, Nara*
[31] *National United University, Miao Li*
[32] *Department of Physics, National Taiwan University, Taipei*
[33] *H. Niewodniczanski Institute of Nuclear Physics, Krakow*

[34] Nippon Dental University, Niigata
[35] Niigata University, Niigata
[36] University of Nova Gorica, Nova Gorica
[37] Osaka City University, Osaka
[38] Pacific Northwest National Laboratory, Richland, Washington 99352
[39] Panjab University, Chandigarh
[40] Research Center for Nuclear Physics, Osaka University, Osaka
[41] University of Science and Technology of China, Hefei
[42] Seoul National University, Seoul
[43] Sungkyunkwan University, Suwon
[44] School of Physics, University of Sydney, NSW 2006
[45] Tata Institute of Fundamental Research, Mumbai
[46] Excellence Cluster Universe, Technische Universität München, Garching
[47] Toho University, Funabashi
[48] Tohoku Gakuin University, Tagajo
[49] Tohoku University, Sendai
[50] Department of Physics, University of Tokyo, Tokyo
[51] Tokyo Institute of Technology, Tokyo
[52] Tokyo Metropolitan University, Tokyo
[53] Tokyo University of Agriculture and Technology, Tokyo
[54] CNP, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061
[55] Wayne State University, Detroit, Michigan 48202
[56] Yamagata University, Yamagata
[57] Yonsei University, Seoul

We present the first measurement of the angle ϕ_3 of the Unitarity Triangle using a model-independent Dalitz plot analysis of $B^\pm \to DK^\pm$, $D \to K^0_S\pi^+\pi^-$ decays. The method uses, as input, measurements of the strong phase of the $D \to K^0_S\pi^+\pi^-$ amplitude from the CLEO collaboration. The result is based on the full data set of 772×10^6 $B\overline{B}$ pairs collected by the Belle experiment at the $\Upsilon(4S)$ resonance. We obtain $\phi_3 = (77.3^{+15.1}_{-14.9} \pm 4.1 \pm 4.3)^\circ$ and the suppressed amplitude ratio $r_B = 0.145 \pm 0.030 \pm 0.010 \pm 0.011$. Here the first error is statistical, the second is the experimental systematic uncertainty, and the third is the error due to the precision of the strong-phase parameters obtained by CLEO.

PACS numbers: 12.15.Hh, 13.25.Hw, 14.40.Nd

I. INTRODUCTION

The angle ϕ_3 (also denoted as γ) is one of the least well-constrained parameters of the Unitarity Triangle. The measurement that currently dominates sensitivity to ϕ_3 uses $B^\pm \to DK^\pm$ decays with the neutral D meson decaying to a three-body final state such as $K^0_S\pi^+\pi^-$ [1, 2]. The weak phase ϕ_3 appears in the interference between $b \to c\bar{u}s$ and $b \to u\bar{c}s$ transitions. The value of ϕ_3 is determined by exploiting differences between the $K^0_S\pi^+\pi^-$ Dalitz plots for D mesons from B^+ and B^- decay. Theoretical uncertainties in the ϕ_3 determination in $B^\pm \to DK^\pm$ decays are expected to be negligible [3], and the main difficulty in its measurement is the very low probability of the decays that are involved. However, the method based on Dalitz plot analysis requires the knowledge of the amplitude of the $D^0 \to K^0_S\pi^+\pi^-$ decay, including its complex phase. The amplitude can be obtained from a model that involves isobar and K-matrix [4] descriptions of the decay dynamics, and thus results in a model uncertainty for the ϕ_3 measurement. In the latest model-dependent Dalitz plot analyses performed by BaBar and Belle, this uncertainty ranges from 3° to 9° [5–10].

A method to eliminate the model uncertainty using a binned Dalitz plot analysis has been proposed by Giri *et al.* [1]. Information about the strong phase in the $D^0 \to K^0_S\pi^+\pi^-$ decay is obtained from the decays of quantum-correlated D^0 pairs produced in the $\psi(3770) \to D^0\overline{D}^0$ process. As a result, the model uncertainty is replaced by a statistical error related to the precision of the strong-phase parameters. This method has been further developed in Refs. [11, 12], where its experimental feasibility has been shown along with a proposed analysis procedure to optimally use the available B decays and correlated D^0 pairs. In this paper, we report the first measurement of ϕ_3 using a model-independent Dalitz plot analysis of the $D \to K^0_S\pi^+\pi^-$ decay from the mode $B^\pm \to DK^\pm$, based on a 711 fb^{-1} data sample (corresponding to 772×10^6 $B\overline{B}$ pairs) collected by the Belle detector at the KEKB asymmetric-energy e^+e^- collider. This analysis uses the recent measurement of the strong phase in $D^0 \to K^0_S\pi^+\pi^-$ and $D^0 \to K^0_S K^+K^-$ decays performed by the CLEO collaboration [13, 14].

II. THE MODEL-INDEPENDENT DALITZ PLOT ANALYSIS TECHNIQUE

The amplitude of the $B^+ \to DK^+$, $D \to K^0_S\pi^+\pi^-$ decay is a superposition of the $B^+ \to \overline{D}{}^0K^+$ and $B^+ \to D^0K^+$ amplitudes

$$A_B(m^2_+, m^2_-) = \overline{A} + r_B e^{i(\delta_B+\phi_3)} A\,, \tag{1}$$

where m^2_+ and m^2_- are the Dalitz plot variables — the squared invariant masses of $K^0_S\pi^+$ and $K^0_S\pi^-$ combinations, respectively, $\overline{A} = \overline{A}(m^2_+, m^2_-)$ is the amplitude of the $\overline{D}{}^0 \to K^0_S\pi^+\pi^-$ decay, $A = A(m^2_+, m^2_-)$ is the amplitude of the $D^0 \to K^0_S\pi^+\pi^-$ decay, r_B is the ratio of the absolute values of the $B^+ \to \overline{D}{}^0K^+$ and $B^+ \to D^0K^+$ amplitudes, and δ_B is the strong-phase difference between them. In the case of CP conservation in the D decay $A(m^2_+, m^2_-) = \overline{A}(m^2_-, m^2_+)$. The Dalitz plot density of the D decay from $B^+ \to DK^+$ is given by

$$\begin{aligned} P_B = |A_B|^2 =& |\overline{A} + r_B e^{i(\delta_B+\phi_3)} A|^2 = \\ & \overline{P} + r_B^2 P + 2\sqrt{P\overline{P}}(x_+C + y_+S)\,, \end{aligned} \tag{2}$$

where $P(m^2_+, m^2_-) = |A|^2$, $\overline{P}(m^2_+, m^2_-) = |\overline{A}|^2$; while

$$x_+ = r_B\cos(\delta_B + \phi_3)\,, \quad y_+ = r_B\sin(\delta_B + \phi_3)\,; \tag{3}$$

and the functions $C = C(m^2_+, m^2_-)$ and $S = S(m^2_+, m^2_-)$ are the cosine and sine of the strong-phase difference $\delta_D(m^2_+, m^2_-) = \arg\overline{A} - \arg A$ between the $\overline{D}{}^0 \to K^0_S\pi^+\pi^-$ and $D^0 \to K^0_S\pi^+\pi^-$ amplitudes.[1] The equations for the charge-conjugate mode $B^- \to DK^-$ are obtained with the substitution $\phi_3 \to -\phi_3$ and $A \leftrightarrow \overline{A}$; the corresponding parameters that depend on the B^- decay amplitude are:

$$x_- = r_B\cos(\delta_B - \phi_3)\,, \quad y_- = r_B\sin(\delta_B - \phi_3)\,. \tag{4}$$

Using both B charges, one can obtain ϕ_3 and δ_B 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 plot density, and the functions C and S are obtained from model assumptions in the fit to the $D^0 \to K^0_S\pi^+\pi^-$ amplitude. In the model-independent approach, the Dalitz plot is divided into $2\mathcal{N}$ bins symmetric under the exchange $m^2_- \leftrightarrow m^2_+$. The bin index i ranges from $-\mathcal{N}$ to $\mathcal{N}$ (excluding 0); the exchange $m^2_+ \leftrightarrow m^2_-$ corresponds to the exchange $i \leftrightarrow -i$. The expected number of events in bin i of the Dalitz plot of the D meson from $B^\pm \to DK^\pm$ is

$$N^\pm_i = h_B\left[K_{\pm i} + r_B^2 K_{\mp i} + 2\sqrt{K_iK_{-i}}(x_\pm c_i \pm y_\pm s_i)\right]\,, \tag{5}$$

where h_B is a normalization constant and K_i is the number of events in the i^{th} bin of the $K^0_S\pi^+\pi^-$ Dalitz plot of the D meson in a flavor eigenstate. A sample of flavor-tagged D^0 mesons is obtained by reconstructing $D^{*\pm} \to D\pi^\pm$ decays (note that charge conjugation is assumed throughout this paper unless otherwise stated). The terms c_i and s_i include information about the functions C and S averaged over the bin region:

$$c_i = \frac{\int\limits_{\mathcal{D}_i} |A||\overline{A}|C\,d\mathcal{D}}{\sqrt{\int\limits_{\mathcal{D}_i} |A|^2 d\mathcal{D} \int\limits_{\mathcal{D}_i} |\overline{A}|^2 d\mathcal{D}}}\,. \tag{6}$$

Here $\mathcal{D}$ represents the Dalitz plot phase space and $\mathcal{D}_i$ is the bin region 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 the c_i and s_i terms are measured in the quantum correlations of D pairs by charm-factory experiments operating at the threshold of $D\overline{D}$ pair production [13, 14]. The measurement involves studies of the four-dimensional density of two correlated $D \to K^0_S\pi^+\pi^-$ Dalitz plots, as well as decays of a D meson tagged in a CP-eigenstate decaying to $K^0_S\pi^+\pi^-$. The wave function of the two mesons is antisymmetric, thus the four-dimensional density of two correlated $D \to K^0_S\pi^+\pi^-$ Dalitz plots is

$$\begin{aligned} |A_{\text{corr}}(m^2_+, m^2_-, m'^2_+, m'^2_-)|^2 = |A\overline{A}' - \overline{A}A'|^2 = \\ P\overline{P}' + \overline{P}P' - 2\sqrt{P\overline{P}'\overline{P}P'}(CC' + SS')\,, \end{aligned} \tag{7}$$

where the primed and unprimed quantities correspond to the two decaying D mesons. Similarly, the density of the decay $D \to K^0_S\pi^+\pi^-$, where the D meson is in a CP-eigenstate, is

$$|A_{\text{CP}}(m^2_+, m^2_-)|^2 = |A \pm \overline{A}|^2 = P + \overline{P} \pm 2\sqrt{P\overline{P}}C\,. \tag{8}$$

CLEO uses these relations to obtain c_i and s_i values. Once they are measured, the system of equations (5) contains only three free parameters (x, y, and h_B) for each B charge, and can be solved using a maximum likelihood method to extract the value of ϕ_3.

We have neglected charm-mixing effects in D decays from both the $B^\pm \to DK^\pm$ process and in quantum-correlated $D\overline{D}$ production. It has been shown [15] that although the charm mixing correction is of first order in the mixing parameters x_D, y_D, it is numerically small (of the order 0.2° for $x_D, y_D \sim 0.01$) and can be neglected at the current level of precision. Future precision measurements of ϕ_3 can account for charm mixing and CP violation (both in mixing and decay) using the measurement of the corresponding parameters.

In principle, the set of relations defined by Eq. (5) can be solved without external constraints on c_i and s_i for $\mathcal{N} \geq 2$. However, due to the small value of r_B, there

[1] This paper follows the convention for strong phases in D decay amplitudes introduced in Ref. [12].

is very little sensitivity to the c_i and s_i parameters in $B^\pm \to DK^\pm$ decays, which results in a reduction in the precision on ϕ_3 that can be obtained [11].

III. CLEO INPUT

The procedure for a binned Dalitz plot analysis should give the correct results for any binning. However, the statistical accuracy depends strongly on the amplitude behavior across the bins. Large variations of the amplitude within a bin result in loss of coherence in the interference term. This effect becomes especially significant with limited statistics when a small number of bins must be used to ensure a stable fit. Greater statistical precision is obtained for the binning in which the phase difference between the D^0 and $\overline{D}^0$ amplitudes varies as little as possible within a bin [12]. For optimal precision, one also has to take the variations of the absolute value of the amplitude into account, along with contributions from background events. The procedure to optimize the binning for the maximal statistical precision of ϕ_3 has been proposed in Ref. [12] and generalized to the case with background in Ref. [14]. It has been shown that as few as 16 bins are enough to reach a statistical precision that is only 10–20% worse than in the unbinned case.

The optimization of binning sensitivity uses the amplitude of the $D \to K^0_S\pi^+\pi^-$ decay. It should be noted, however, that although the choice of binning is model-dependent, a poor choice of model results only in a loss of precision, not bias, of the measured parameters [12]. CLEO measured c_i and s_i parameters for four different binnings with $\mathcal{N} = 8$:

1. Bins equally distributed in the phase difference $\Delta\delta_D$ between the D^0 and $\overline{D}^0$ decay amplitudes, with the amplitude from the BaBar measurement [6].

2. Same as option 1, but with the amplitude from the Belle analysis [10].

3. Optimized for statistical precision according to the procedure from [12] (see Fig. 1). The effect of the background in B data is not taken into account in the optimization. The amplitude is taken from the BaBar measurement [6].

4. Same as option 3, but optimized for an analysis with high background in B data (*e. g.,* at LHCb).

Our analysis uses the optimal binning shown in Fig. 1 (option 3) as the baseline since it offers better statistical accuracy. In addition, we use the equal phase difference binning ($\Delta\delta_D$-binning, option 1) as a cross-check.

The results of the CLEO measurement of c_i and s_i for the optimal binning are presented in Table I. The same results in graphical form are shown in Fig. 2. The values of c_i and s_i calculated from the Belle model [10] are compared to the measurements and are found to be in reasonable agreement with $\chi^2 = 18.6$ for the number of degrees of freedom ndf $= 16$ (the corresponding p-value is $p = 29\%$).

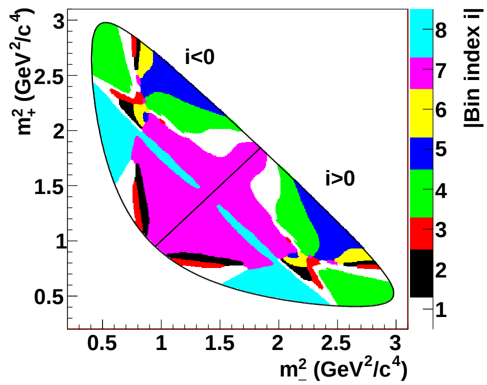

FIG. 1. Optimal binning of the $D \to K^0_S\pi^+\pi^-$ Dalitz plot. The color scale indicated corresponds to the absolute value of the bin index, $|i|$.

TABLE I. Values of c_i and s_i for the optimal binning measured by CLEO [14], and calculated from the Belle $D \to K^0_S\pi^+\pi^-$ amplitude model.

	CLEO measurement	Belle model
c_1	$-0.009 \pm 0.088 \pm 0.094$	-0.039
c_2	$+0.900 \pm 0.106 \pm 0.082$	$+0.771$
c_3	$+0.292 \pm 0.168 \pm 0.139$	$+0.242$
c_4	$-0.890 \pm 0.041 \pm 0.044$	-0.867
c_5	$-0.208 \pm 0.085 \pm 0.080$	-0.246
c_6	$+0.258 \pm 0.155 \pm 0.108$	$+0.023$
c_7	$+0.869 \pm 0.034 \pm 0.033$	$+0.851$
c_8	$+0.798 \pm 0.070 \pm 0.047$	$+0.662$
s_1	$-0.438 \pm 0.184 \pm 0.045$	-0.706
s_2	$-0.490 \pm 0.295 \pm 0.261$	$+0.124$
s_3	$-1.243 \pm 0.341 \pm 0.123$	-0.687
s_4	$-0.119 \pm 0.141 \pm 0.038$	-0.108
s_5	$+0.853 \pm 0.123 \pm 0.035$	$+0.851$
s_6	$+0.984 \pm 0.357 \pm 0.165$	$+0.930$
s_7	$-0.041 \pm 0.132 \pm 0.034$	$+0.169$
s_8	$-0.107 \pm 0.240 \pm 0.080$	-0.596

As is apparent from Fig. 2, the chosen binning contains bins where the strong phase difference between D^0 and $\overline{D}^0$ amplitudes is close to zero (bins with $|i| = 2, 7, 8$) and 180° (bin with $|i| = 4$) which provide sensitivity to $x_\pm$, as well as bins with the strong phase difference close to 90° and 270° (bins with $|i| = 1, 3, 5, 6$), more sensitive to $y_\pm$. This ensures that the method is sensitive to ϕ_3 for

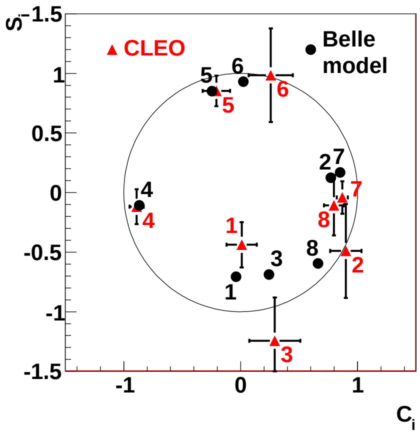

FIG. 2. Comparison of phase terms (c_i, s_i) for the optimal binning measured by CLEO, and calculated from the Belle $D \to K_S^0\pi^+\pi^-$ amplitude model.

any combination of ϕ_3 and δ_B values.

IV. ANALYSIS PROCEDURE

Equation (5) is the key relation used in the analysis, but it only holds if there is no background, a uniform Dalitz plot acceptance and no cross-feed between bins. (Cross-feed is due to invariant-mass resolution and radiative corrections.) In this section we outline the procedures that account for these experimental effects.

A. Efficiency profile

We note that the Eqs. (2), (7) and (8) do not change after the transformation $P \to \epsilon P$ when the efficiency profile $\epsilon(m_+^2, m_-^2)$ is symmetric: $\epsilon(m_+^2, m_-^2) = \epsilon(m_-^2, m_+^2)$. This implies that if the efficiency profile is the same in all of the three modes involved in the measurement (flavor D, correlated $\psi(3770) \to D\overline{D}$, and D from $B \to DK$), the result will be unbiased even if no efficiency correction is applied.

The effect of non-uniform efficiency over the Dalitz plot cancels out when using a flavor-tagged D sample with kinematic properties that are similar to the sample for the signal B decay. This approach allows for the removal of systematic error associated with the possible inaccuracy in the description of the detector acceptance in the Monte Carlo (MC) simulation. The center-of-mass (CM) D momentum distribution for $B \to DK$ decays is practically uniform in the narrow range 2.10 GeV$/c < p_D <$ 2.45 GeV$/c$. We assume that the efficiency profile depends mostly on the D momentum and take the flavor-tagged sample with an average momentum of $p_D = 2.3$ GeV$/c$ (we use a wider range of D momenta than in $B \to DK$ to increase the statistics). The assumption that the efficiency profile depends only on the D momentum is tested using MC simulation, and the remaining difference is treated as a systematic uncertainty.

While calculating c_i and s_i, CLEO applies an efficiency correction, therefore the values reported in their analysis correspond to a flat efficiency profile. To use the c_i and s_i values in the ϕ_3 analysis, they have to be corrected for the Belle efficiency profile. This correction cannot be performed in a completely model-independent way, since the correction terms include the phase variation inside the bin. Fortunately, the calculations using the Belle $D \to K_S^0\pi^+\pi^-$ model show that this correction is negligible even for very large non-uniformity of the efficiency profile. The difference between the uncorrected c_i and s_i terms and those corrected for the efficiency, calculated using the efficiency profile parameterization used in the 605 fb^{-1} analysis [10], does not exceed 0.01, which is negligible compared to the statistical error.

B. Momentum resolution

Momentum resolution leads to migration of events between 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 bins:

$$N_i' = \sum \alpha_{ik} N_k, \tag{9}$$

where N_k is the number of events that bin k would contain without the cross-feed, and N_i' is the reconstructed number of events in bin i. The cross-feed matrix α_{ik} is nearly a unit matrix: $\alpha_{ik} \ll 1$ for $i \neq k$. It is obtained from a signal MC simulation generated with the amplitude model reported in Ref [10]. In the case of a $D \to K_S^0\pi^+\pi^-$ decay from a B, the cross-feed depends on the parameters x and y. However, this is a minor correction to an already small effect due to cross-feed; therefore it is neglected.

Migration of events between the bins also occurs due to final state radiation (FSR). The c_i and s_i terms in the CLEO measurement are not corrected for FSR; we therefore do not simulate FSR to obtain the cross-feed matrix to minimize the bias due to this effect. Comparison of the cross-feed with and without FSR shows that this effect is negligible.

C. Fit procedure

The background contribution has to be accounted for in the calculation of the values N_i and K_i. Statistically the most effective way of calculating the number of signal events (especially in the case of N_i, where the statistics is a limiting factor) 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.

Two different approaches are used in this analysis to obtain the CP violating parameters from the data: separate fits in bins, and a combined fit.

In the first, we fit the data distribution in each bin separately, with the number of events for signal and backgrounds as free parameters. Once the numbers of events in bins N_i are found, we use them in Eq. 5 to obtain the parameters $(x_\pm, y_\pm)$. This is accomplished by minimizing a negative logarithmic likelihood of the form

$$-2\log\mathcal{L}(x,y) = -2\sum_i \log p(\langle N_i\rangle(x,y), N_i, \sigma_{N_i}), \quad (10)$$

where $\langle N_i\rangle(x,y)$ is the expected number of events in the bin i obtained from Eq. 5. Here, N_i and σ_{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 χ^2 fit; however, the assumption of the Gaussian distribution may introduce a bias in the case of low statistics in certain bins.

The procedure described above does not make any assumptions on 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 second approach is to use the combined fit with a common likelihood for all bins. The relative numbers of background events in bins in such a fit can be constrained externally from MC and control samples. In addition, for the case of the combined fit, the two-step procedure of first extracting the numbers of signal events, and then using them to obtain (x, y) is not needed — the expected numbers of events $\langle N_i\rangle$ as functions of (x, y) can be included in the likelihood. Thus the variables (x, y) become free parameters of the combined likelihood fit, and the assumption that the number of signal events has a Gaussian distribution is not needed.

Both approaches are tested with the control samples and MC simulation. We choose the combined fit approach as the baseline, but the procedure with separate fits in bins is also used: it allows a clear demonstration of the CP asymmetry in each bin.

V. EVENT SELECTION

We use a data sample of 772×10^6 $B\overline{B}$ pairs collected by the Belle detector. The decays $B^\pm\to DK^\pm$ and $B^\pm\to D\pi^\pm$ are selected for the analysis. The neutral D meson is reconstructed in the $K^0_S\pi^+\pi^-$ final state in all cases. We also select $D^{*\pm}\to D\pi^\pm$ decays produced via the $e^+e^-\to c\bar{c}$ continuum process as a high-statistics sample to determine the K_i parameters related to the flavor-tagged $D^0\to K^0_S\pi^+\pi^-$ decay.

The Belle detector is described in detail elsewhere [16, 17]. It is a large-solid-angle magnetic spectrometer consisting of a silicon vertex detector (SVD), a 50-layer central drift chamber (CDC) for charged particle tracking and specific ionization measurement (dE/dx), an array of aerogel threshold Cherenkov counters (ACC), time-of-flight scintillation counters (TOF), and an array of CsI(Tl) crystals for electromagnetic calorimetry (ECL) located inside a superconducting solenoid coil that provides a 1.5 T magnetic field. An iron flux return located outside the coil is instrumented to detect K_L mesons and identify muons (KLM).

Charged tracks are required to satisfy criteria based on the quality of the track fit and the distance from the interaction point of the beams (IP). We require each track to have a transverse momentum greater than 100 MeV/c, and the impact parameter relative to the IP to be less than 2 mm in the transverse and less than 10 mm in longitudinal projections. Separation of kaons and pions is accomplished by combining the responses of the ACC and the TOF with the dE/dx measurement from the CDC. Neutral kaons are reconstructed from pairs of oppositely charged tracks with an invariant mass $M_{\pi\pi}$ within 7 MeV/c^2 of the nominal K^0_S mass, flight distance from the IP in the plane transverse to the beam axis greater than 0.1 mm, and the cosine of the angle between the projections of K^0_S flight direction and its momentum greater than 0.95.

The flavor of the neutral D mesons used for K_i determination is tagged by the charge of the slow pion in the decay $D^{*\pm}\to D\pi^\pm$. The slow pion track is required to originate from the D^0 decay vertex to improve the momentum and angular resolution. The selection of signal candidates is based on two variables, the invariant mass of the neutral D candidates $M_D = M_{K^0_S\pi^+\pi^-}$ and the difference of the invariant masses of the $D^{*\pm}$ and the neutral D candidates $\Delta M = M_{(K^0_S\pi^+\pi^-)_D\pi} - M_{K^0_S\pi^+\pi^-}$. We retain the events satisfying the following criteria: $1800\ \mathrm{MeV}/c^2 < M_D < 1920\ \mathrm{MeV}/c^2$ and $\Delta M < 150$ MeV/c^2. We also require the momentum of the D^0 candidate in the CM frame p_D to be in the range $1.8\ \mathrm{GeV}/c < p_D < 2.8\ \mathrm{GeV}/c$ to reduce the effect of the efficiency profile on the ϕ_3 measurement (see Sec. IV A). About 15% of selected events contain more than one $D^{*\pm}$ candidate that satisfies the requirements above; in this case we keep only one randomly selected candidate.

Selection of $B^\pm\to DK^\pm$ and $B^\pm\to D\pi^\pm$ samples is based on the CM-energy difference $\Delta E = \sum E_i - E_{\rm beam}$ and the beam-constrained B meson mass $M_{\rm bc} = \sqrt{E^2_{\rm beam} - (\sum \vec{p}_i)^2}$, where $E_{\rm beam}$ is the CM beam energy, and E_i and $\vec{p}_i$ are the CM energies and momenta of the B candidate decay products. We select events with $M_{\rm bc} > 5.2\ \mathrm{GeV}/c^2$ and $|\Delta E| < 0.18$ GeV for further analysis. We also impose a requirement on the invariant mass of the neutral D candidate $|M_{K^0_S\pi^+\pi^-} - M_{D^0}| < 11\ \mathrm{MeV}/c^2$.

Further separation of the background from $e^+e^-\to q\bar{q}$ $(q = u, d, s, c)$ continuum events is done by calculating two variables that characterize the event shape. One is the cosine of the thrust angle $\cos\theta_{\rm thr}$, where $\theta_{\rm thr}$ is the angle between the thrust axis of the B candidate daughters and that of the rest of the event, calculated in the

CM frame. The other is a Fisher discriminant $\mathcal{F}$ composed of 11 parameters [18]: the production angle of the B candidate, the angle of the B thrust axis relative to the beam axis, and nine parameters representing the momentum flow in the event relative to the B thrust axis in the CM frame. We use the ΔE, $M_{\rm bc}$, $\cos\theta_{\rm thr}$, and $\mathcal{F}$ variables in the maximum likelihood fit.

In both flavor D^0 and $B^\pm \to DK^\pm$ ($B^\pm \to D\pi^\pm$) samples, the momenta of the tracks forming a D^0 candidate are constrained to give the nominal D^0 mass in the calculation of the Dalitz plot variables.

VI. FLAVOR-TAGGED SAMPLE $D^{*\pm} \to D\pi^\pm$, $D \to K^0_S\pi^+\pi^-$

The number of events K_i in bin i of the flavor-tagged $D \to K^0_S\pi^+\pi^-$ decay is obtained from a two-dimensional unbinned fit to the distribution of M_D and ΔM variables. The fits in each Dalitz plot bin are performed independently. The fit uses a signal PDF and two background components: purely random combinatorial background and background with a real D^0 and random slow pion track. The signal distribution is a product of the PDFs for M_D (triple Gaussian) and ΔM (sum of bifurcated Student's t-distribution and bifurcated Gaussian distribution). The combinatorial background is parameterized by a linear function in M_D and by a function with a kinematic threshold at the π^+ mass in ΔM:

$$p_{\rm comb}(\Delta M) = \sqrt{y}(1 + Ay[1 + B(M_D - m_{D^0})])e^{-yC} \tag{11}$$

where $y = \Delta M - m_{\pi^+}$, m_{π^+} and m_{D^0} are the nominal masses of π^+ and D^0, respectively, and A, B, and C are free parameters. A small correlation between the M_D and ΔM distributions is introduced that is controlled by the parameter B. The random slow pion background is parameterized as a product of the signal M_D distribution and combinatorial ΔM background shape.

The parameters of the signal and background distributions are obtained from the fit to data. The parameters of the signal PDF are constrained to be the same in all bins. The free parameters in each bin are the number of signal events K_i, the parameters of the background distribution, and fractions of the background components.

The fit results from the flavor-tagged D sample integrated over the whole Dalitz plot are shown in Fig. 3. The number of signal events calculated from the integral of the signal distribution is 426938 ± 825, the background fraction in the signal region $|M_D - m_{D^0}| < 11$ MeV/c^2, 144.5 MeV/c^2 $< \Delta M <$ 146.5 MeV/c^2 is $10.1 \pm 0.1\%$. The signal yield in bins is shown in Table II.

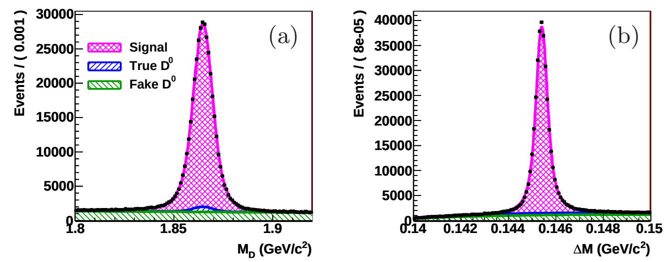

FIG. 3. Projections of the flavor-tagged $D^{*\pm} \to D\pi^\pm$, $D \to K^0_S\pi^+\pi^-$ data with 1.8 GeV/c $< p_D <$ 2.8 GeV/c. (a) M_D distribution for 144.5 MeV/c^2 $< \Delta M <$ 146.5 MeV/c^2. (b) ΔM distribution for 1854 MeV/c^2 $< M_D <$ 1876 MeV/c^2. Histograms show the fitted signal and background contributions, points with the error bars are the data. The full $D \to K^0_S\pi^+\pi^-$ Dalitz plot is used.

TABLE II. Signal yields in Dalitz plot bins for the flavor-tagged $D^{*\pm} \to D\pi^\pm$, $D \to K^0_S\pi^+\pi^-$ sample with 1.8 GeV/c $< p_D <$ 2.8 GeV/c.

Bin i	K_i	K_{-i}
1	43261 ± 255	8770 ± 124
2	58005 ± 268	1827 ± 63
3	62808 ± 274	1601 ± 58
4	44513 ± 253	26482 ± 202
5	21886 ± 177	13146 ± 143
6	28876 ± 197	1765 ± 68
7	48001 ± 265	22476 ± 196
8	9279 ± 125	26450 ± 181
Total	426938 ± 825	

VII. SELECTION OF $B^\pm \to D\pi^\pm$ AND $B^\pm \to DK^\pm$ SAMPLES

The decays $B^\pm \to DK^\pm$ and $B^\pm \to D\pi^\pm$ have similar topology and background sources and their selection is performed in a similar way. The mode $B^\pm \to D\pi^\pm$ has an order of magnitude larger branching ratio and a smaller amplitude ratio $r_B \sim 0.01$ due to the ratio of weak coefficients $|V_{ub}V^*_{cd}|/|V_{cb}V^*_{ud}| \sim 0.02$ and the color suppression factor. This results in the small CP violation in this mode, therefore it is used as a control sample to test the procedures of the background extraction and Dalitz plot fit. In addition, signal resolutions in ΔE and $M_{\rm bc}$ and the Dalitz plot structure of some background components are constrained from the control sample and used in the signal fit.

The number of signal events is obtained by fitting the 4D distribution of variables $M_{\rm bc}$, ΔE, $\cos\theta_{\rm thr}$ and $\mathcal{F}$. The fits to the $B^\pm \to D\pi^\pm$ and $B^\pm \to DK^\pm$ samples use the following three background components in addition to the signal PDF:

- Combinatorial background from the process

$e^+e^- \to q\overline{q}$, where $q = (u, d, s, c)$.

- Random $B\overline{B}$ background, in which the tracks forming the $B^\pm \to D\pi^\pm$ 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 $(M_{\rm bc}, \Delta E)$ distribution are quite smooth.

- Peaking $B\overline{B}$ background, in which all tracks forming the $B^\pm \to D\pi^\pm$ candidate come from the same B meson. This kind of background is dominated by $B \to D^*\pi$ decays reconstructed without the π or γ from the D^* decay.

In addition, the $B^\pm \to DK^\pm$ fit includes a fourth component that models $B^\pm \to D\pi^\pm$ decays in which the pion is misidentified as a kaon.

The PDF for the signal parameterization (as well as for each of the background components) is a product of the $(M_{\rm bc}, \Delta E)$ and $(\cos\theta_{\rm thr}, \mathcal{F})$ PDFs. The $(M_{\rm bc}, \Delta E)$ PDF is a 2D double-Gaussian function, which has a correlation between $M_{\rm bc}$ and ΔE. The double-Gaussian function models both the core and tails of the distribution. The $(\cos\theta_{\rm thr}, \mathcal{F})$ distribution is parameterized by the sum of two functions (with different coefficients) of the form

$$p(x, \mathcal{F}) = \exp(C_1 x + C_2 x^2 + C_3 x^3)\times G(\mathcal{F}, F_0(x), \sigma_{FL}(x), \sigma_{FR}(x)), \quad (12)$$

where $x = \cos\theta_{\rm thr}$, $G(\mathcal{F}, F, \sigma_L, \sigma_R)$ is the bifurcated Gaussian distribution with the mean F and the widths σ_L and σ_R, and functions F_0, σ_{FL} and σ_{FR} are polynomials that contain only even powers of x. The parameters of the signal PDF are obtained from the signal MC simulation. However, to account for the possible imperfection of the simulation, we allow all the width parameters to scale by a common factor, which is obtained from the $B^\pm \to D\pi^\pm$ sample.

The combinatorial background from continuum $e^+e^- \to q\overline{q}$ production is obtained from the experimental sample collected at a CM energy below the $\Upsilon(4S)$ resonance (off-resonance data). The parameterization in variables $(\cos\theta_{\rm thr}, \mathcal{F})$ follows Eq. (12). The parameterization in $(M_{\rm bc}, \Delta E)$ is the product of an exponential distribution in ΔE and the empirical shape proposed by the ARGUS collaboration [19] in $M_{\rm bc}$:

$$p_{\rm comb}(M_{\rm bc}, \Delta E) = \exp(-\alpha\Delta E) M_{\rm bc}\sqrt{y}\exp(-cy), \quad (13)$$

where $y = 1 - M_{\rm bc}/E_{\rm beam}$, $E_{\rm beam}$ is the CM beam energy, and α and c are empirical parameters.

The parameters for random and peaking $B\overline{B}$ backgrounds are obtained from a generic MC sample. Generator information is used to distinguish between the two: the latter contains only the events in which the candidate is formed of tracks coming from both B mesons. The $(M_{\rm bc}, \Delta E)$ distributions for each of these backgrounds are parameterized by the sum of three components:

- the product of an exponential (in ΔE) and Argus (in $M_{\rm bc}$) functions, as for continuum background (as expected, this component dominates the random $B\overline{B}$ background);

- the product of an exponential in the ΔE and bifurcated Gaussian distribution in $M_{\rm bc}$, where the mean of the Gaussian distribution is linear as a function of ΔE; and

- a two-dimensional Gaussian distribution in ΔE and $M_{\rm bc}$, which includes a correlation and is asymmetric in $M_{\rm bc}$. This component is small compared to the random $B\overline{B}$ contribution, but dominates the peaking $B\overline{B}$ background, which mostly consists of partially reconstructed B decays.

The peaking background coming from B^+B^- and $B^0\overline{B}^0$ decays is treated separately in $(M_{\rm bc}, \Delta E)$ variables, while a common $(\cos\theta_{\rm thr}, \mathcal{F})$ distribution is used. In the case of the $B^\pm \to DK^\pm$ fit, $B^\pm \to D\pi^\pm$ events with the pion misidentified as a kaon are treated as a separate background category. The distributions of $M_{\rm bc}, \Delta E$ and $\cos\theta_{\rm thr}, \mathcal{F}$ variables are parameterized in the same way as for the signal events and are obtained from MC simulation.

The Dalitz plot distributions of the background components are discussed in the next section. Note that the Dalitz distribution is described by the relative number of events in each bin. The numbers of events in bins can be free parameters in the fit, thus there will be 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 $B\overline{B}$ background in the case of $B^\pm \to D\pi^\pm$), or is constrained by a much larger number of events than the signal (such as the continuum background).

The results of the fit to $B^\pm \to D\pi^\pm$ and $B^\pm \to DK^\pm$ data with the full Dalitz plot taken are shown in Figs. 4 and 5, respectively. We obtain a total of 19106 ± 147 signal $B^\pm \to D\pi^\pm$ events and 1176 ± 43 signal $B^\pm \to DK^\pm$ events — 55% more than in the 605 fb^{-1} model-dependent analysis [10]. The improvement partially comes from the larger integrated luminosity of the sample, and partially from the larger selection efficiency due to improved track reconstruction.

VIII. DATA FITS IN BINS

The data fits in bins for both $B^\pm \to D\pi^\pm$ and $B^\pm \to DK^\pm$ samples are performed with two different procedures: separate fits for the number of events in bins and the combined fit with the free parameters (x, y) as discussed in Sec. IV C. The combined fit is used to obtain the final values for (x, y), while the separate fits provide a crosscheck of the fit procedure and a way to visualize the extent of CP violation within the sample. A study

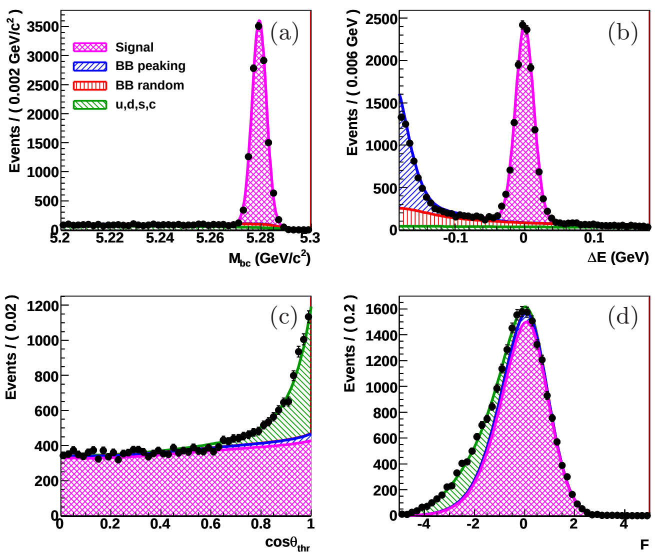

FIG. 4. Projections of the $B^\pm \to D\pi^\pm$ data. (a) $M_{\rm bc}$ distribution with $|\Delta E| < 30$ MeV and $\cos\theta_{\rm thr} < 0.8$ requirements. (b) ΔE distribution with $M_{\rm bc} > 5270$ MeV/c^2 and $\cos\theta_{\rm thr} < 0.8$ requirements. (c) $\cos\theta_{\rm thr}$ and (d) $\mathcal{F}$ distributions with $|\Delta E| < 30$ MeV and $M_{\rm bc} > 5270$ MeV/c^2 requirements. Histograms show the fitted signal and background contributions, points with error bars are the data. The entire $D \to K^0_S\pi^+\pi^-$ Dalitz plot is used.

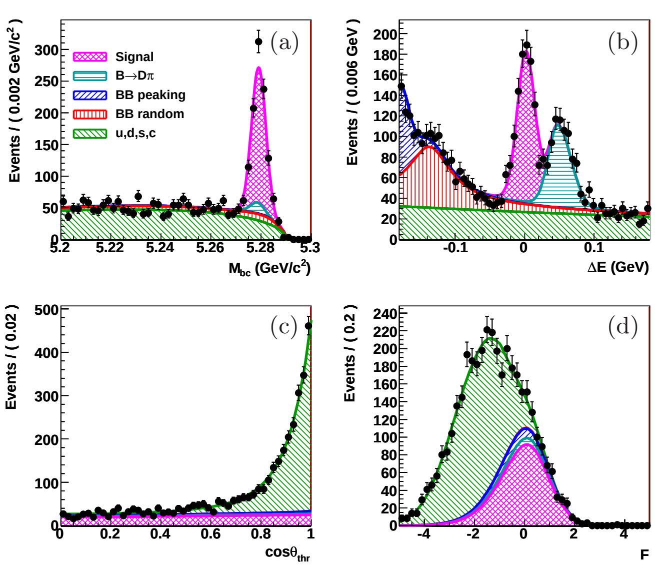

FIG. 5. Projections of the $B^\pm \to DK^\pm$ data. (a) $M_{\rm bc}$ distribution with $|\Delta E| < 30$ MeV and $\cos\theta_{\rm thr} < 0.8$ requirements. (b) ΔE distribution with $M_{\rm bc} > 5270$ MeV/c^2 and $\cos\theta_{\rm thr} < 0.8$ requirements. (c) $\cos\theta_{\rm thr}$ and (d) $\mathcal{F}$ distributions with $|\Delta E| < 30$ MeV and $M_{\rm bc} > 5270$ MeV/c^2 requirements. Histograms show the fitted signal and background contributions, points with error bars are the data. The entire $D \to K^0_S\pi^+\pi^-$ Dalitz plot is used.

with MC pseudo-experiments is performed to check that the observed difference in the fit results between the two approaches agrees with expectation.

In the case of separate fits in bins, we first perform the fit to all events in the Dalitz plot. The fit uses background shapes fixed to those obtained from fits to the generic MC samples of continuum and $B\overline{B}$ decays. The signal shape parameters are fixed to those obtained from a fit to the signal MC sample except for the mean value and width scale factors of ΔE and $M_{\rm bc}$ PDFs. As a next step, we fit the 4D $(M_{\rm bc}, \Delta E, \cos\theta_{\rm thr}, \mathcal{F})$ distributions in each bin separately, with the signal peak positions and width scale factors fixed to the values obtained from the fit to all events. The free parameters of each fit are the number of signal events, and the number of events in each background category.

The numbers of signal events in bins for the $B^\pm \to D\pi^\pm$ sample extracted from the fits are given in Table III. These numbers are used in the fit to extract (x, y) using Eq. (5) after the cross-feed and efficiency correction for both N_i and K_i. Figure 6 illustrates the results of this fit. The numbers of signal events in each bin for B^+ and B^- are shown in Fig. 6(a) together with the numbers of events in the flavor-tagged D^0 sample (appropriately scaled). The difference in the number of signal events shown in Fig. 6(b) does not reveal CP violation. Figures 6(c) [(d)] show the difference between the numbers of signal events for B^+ [B^-] data and scaled flavor-tagged D^0 sample, both for the data and after the (x, y) fit. The χ^2/ndf is reasonable for both the (x, y) fit and the comparison with the flavor-specific CP conserving amplitude.

Unlike $B^\pm \to D\pi^\pm$, the $B^\pm \to DK^\pm$ sample has significantly different signal yields in bins of B^+ and B^- data (see Fig. 7(b) and Table IV). The probability to obtain this difference as a result of a statistical fluctuation is 0.42%. This value can be taken as the model-independent measure of the CP violation significance. The significance of ϕ_3 being nonzero is in general smaller since $\phi_3 \neq 0$ results in a specific pattern of charge asymmetry. The fit of the signal yields to the expected pattern described by the parameters (x, y) is of good quality [Figs. 7(c,d)], which is consistent with the hypothesis that the observed CP violation is solely explained by the mechanism involving nonzero ϕ_3.

The default combined fit constrains the random $B\overline{B}$ background in bins from the generic MC, and takes the $(x_\pm, y_\pm)$ variables as free parameters. Fits to B^+ and B^- data are performed separately. The plots illustrating the combined fit results are given in the Appendix. Additional free parameters are the yields of continuum and peaking $B\overline{B}$ backgrounds in each bin, the fraction of the random $B\overline{B}$ background, and the means and scale factors of the signal $M_{\rm bc}$ and ΔE distributions. The values of (x, y) are then corrected for the fit bias obtained from MC pseudo-experiments. The value of the bias depends on the initial x and y values and is of the order 5×10^{-3} for the $B^\pm \to DK^\pm$ sample and less than 10^{-3} for $B^\pm \to D\pi^\pm$ sample.

TABLE III. Signal yields in Dalitz plot bins for the $B^\pm \to D\pi^\pm$, $D \to K^0_S\pi^+\pi^-$ sample with the optimal binning.

Bin i	N_i^-	N_i^+
-8	564.2 ± 25.3	587.0 ± 25.7
-7	462.3 ± 23.8	462.8 ± 23.9
-6	47.9 ± 7.7	39.2 ± 7.2
-5	314.1 ± 19.0	286.2 ± 18.2
-4	592.6 ± 26.5	645.7 ± 27.8
-3	22.2 ± 6.2	27.2 ± 6.3
-2	42.7 ± 7.6	54.0 ± 8.7
-1	190.8 ± 15.4	210.8 ± 16.3
1	959.2 ± 32.6	980.2 ± 33.1
2	1288.7 ± 37.0	1295.9 ± 37.1
3	1395.8 ± 38.4	1352.2 ± 37.9
4	1045.5 ± 34.7	1065.1 ± 34.9
5	479.3 ± 23.3	532.2 ± 24.5
6	623.7 ± 26.0	663.5 ± 26.7
7	1081.0 ± 35.3	1049.2 ± 34.8
8	210.0 ± 16.1	212.1 ± 16.3
Total	9467.1 ± 103.6	9639.1 ± 104.7

TABLE IV. Signal yields in Dalitz plot bins for the $B^\pm \to DK^\pm$, $D \to K^0_S\pi^+\pi^-$ sample with the optimal binning.

Bin i	N_i^-	N_i^+
-8	49.8 ± 8.2	37.8 ± 7.5
-7	42.2 ± 8.6	24.9 ± 7.2
-6	0.0 ± 1.9	3.4 ± 2.9
-5	9.6 ± 4.5	23.6 ± 6.2
-4	32.9 ± 7.5	42.1 ± 8.3
-3	3.5 ± 2.8	0.7 ± 2.5
-2	11.3 ± 4.1	0.0 ± 1.3
-1	16.6 ± 5.4	7.7 ± 4.4
1	37.6 ± 8.0	65.1 ± 9.9
2	68.6 ± 9.6	75.5 ± 9.8
3	83.4 ± 10.1	82.4 ± 10.2
4	49.3 ± 9.1	86.5 ± 11.4
5	34.0 ± 7.3	38.3 ± 7.6
6	34.8 ± 6.8	41.9 ± 7.5
7	70.8 ± 10.6	46.4 ± 9.0
8	9.4 ± 4.3	14.2 ± 5.1
Total	574.9 ± 29.9	601.6 ± 30.8

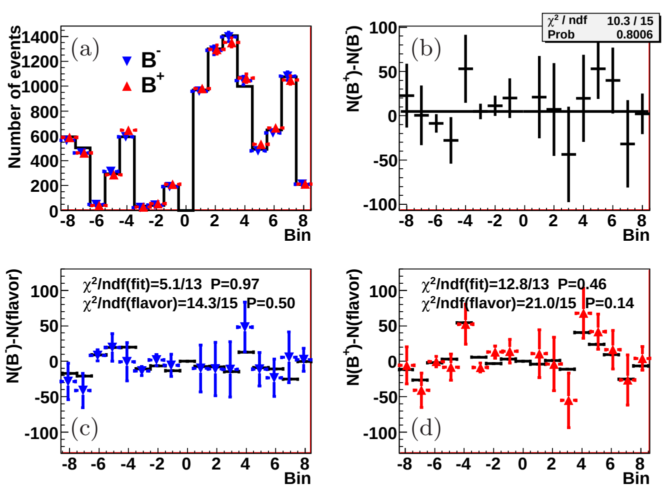

FIG. 6. Results of the fit to the $B^\pm \to D\pi^\pm$ sample. (a) Signal yield in bins of the $D \to K^0_S\pi^+\pi^-$ Dalitz plot: from $B^- \to D\pi^-$ (red), $B^+ \to D\pi^+$ (blue) and flavor sample (histogram). (b) Difference of signal yields between the $B^+ \to D\pi^+$ and $B^- \to D\pi^-$ decays. (c) Difference of signal yields between the $B^- \to D\pi^-$ and flavor samples (normalized to the total $B^- \to D\pi^-$ yield): yield from the separate fits (points with error bars), and as a result of the combined (x, y) fit (horizontal bars). (d) Same as (c) for $B^+ \to D\pi^+$ data.

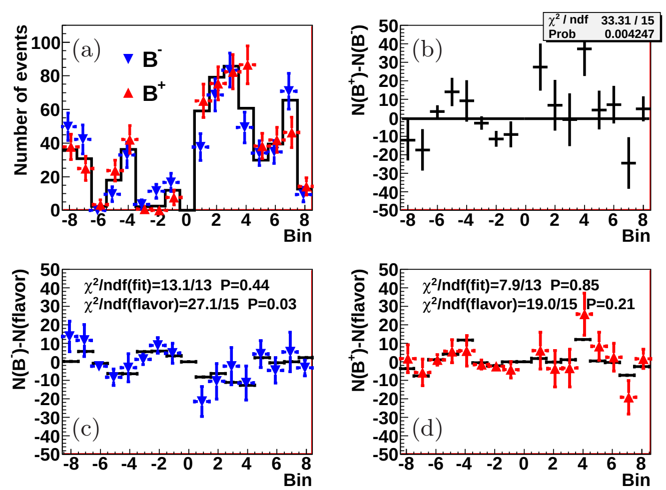

FIG. 7. Results of the fit of the $B^\pm \to DK^\pm$ sample. (a) Signal yield in bins of the $D \to K^0_S\pi^+\pi^-$ Dalitz plot: from $B^- \to DK^-$ (red), $B^+ \to DK^+$ (blue) and flavor sample (histogram). (b) Difference of signal yields between the $B^+ \to DK^+$ and $B^- \to DK^-$ decays. (c) Difference of signal yields between the $B^- \to DK^-$ and flavor samples (normalized to the total $B^- \to DK^-$ yield): yield from the separate fits (points with error bars), and as a result of the combined (x, y) fit (horizontal bars). (d) Same as (c) for $B^+ \to DK^+$ data.

TABLE V. (x, y) parameters and their statistical correlations from the combined fit of the $B^\pm \to D\pi^\pm$ and $B^\pm \to DK^\pm$ samples. The first error is statistical, the second error is systematic, and the third error is due to the uncertainty on the c_i and s_i parameters.

Parameter	$B^\pm \to D\pi^\pm$	$B^\pm \to DK^\pm$
x_-	$-0.0045 \pm 0.0087 \pm 0.0049 \pm 0.0026$	$+0.095 \pm 0.045 \pm 0.014 \pm 0.010$
y_-	$-0.0231 \pm 0.0107 \pm 0.0041 \pm 0.0065$	$+0.137^{+0.053}_{-0.057} \pm 0.015 \pm 0.023$
$\mathrm{corr}(x_-, y_-)$	-0.189	-0.315
x_+	$-0.0172 \pm 0.0089 \pm 0.0060 \pm 0.0026$	$-0.110 \pm 0.043 \pm 0.014 \pm 0.007$
y_+	$+0.0129 \pm 0.0103 \pm 0.0059 \pm 0.0065$	$-0.050^{+0.052}_{-0.055} \pm 0.011 \pm 0.017$
$\mathrm{corr}(x_+, y_+)$	-0.205	$+0.059$

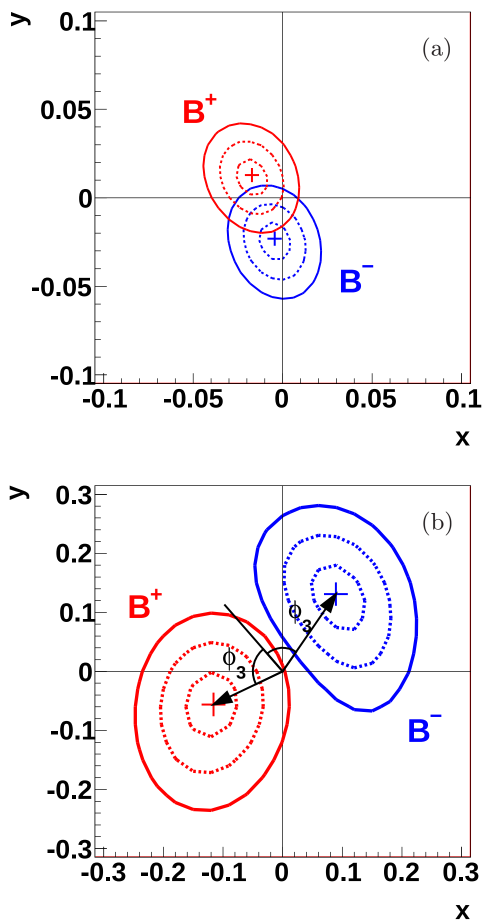

FIG. 8. One, two, and three standard deviation (x, y) confidence levels for (a) $B^\pm \to D\pi^\pm$ and (b) $B^\pm \to DK^\pm$ decays (statistical only). Note different scale for $B^\pm \to D\pi^\pm$ and $B^\pm \to DK^\pm$ modes. The weak phase ϕ_3 appears as half the opening angle between (x_+, y_+) and (x_-, y_-) vectors.

The values of (x, y) parameters and their statistical correlations obtained from the combined fit for $B^\pm \to D\pi^\pm$ control sample and $B^\pm \to DK^\pm$ sample are given in Table V. The measured values of $(x_\pm, y_\pm)$ for both samples with their statistical likelihood contours are shown in Fig. 8.

IX. SYSTEMATIC ERRORS

Systematic errors in the (x, y) fit are obtained for the default procedure of the combined fit with the optimal binning. The systematic errors are summarized in Table VI.

The uncertainty due to the signal shape used in the fit includes the sources listed below:

1. The choice of parameterization used to model the shape. The corresponding uncertainty is estimated by using the non-parametric Keys PDF function [20] instead of the parameterized distribution.

2. Correlation between the $(M_{\rm bc}, \Delta E)$ and $(\cos\theta_{\rm thr}, \mathcal{F})$ distributions. To estimate its effect, we use a 4D binned histogram to describe the distribution.

3. The MC description of the $(\cos\theta_{\rm thr},\ \mathcal{F})$ distribution. Its effect is estimated by floating the parameters of the distribution in the fit to the $B^\pm \to D\pi^\pm$ control sample.

4. The dependence of the signal width on the Dalitz plot bin. The uncertainty due to this effect is estimated by performing the $B^\pm \to D\pi^\pm$ fit with the shape parameters floated separately for each bin, and then using the results in the fit to $B^\pm \to DK^\pm$ data.

We do not assign an uncertainty due to the difference in $(M_{\rm bc}, \Delta E)$ shape between the MC and data since the width of the signal distribution is calibrated using $B^\pm \to D\pi^\pm$ data.

For the uncertainty due to the continuum background shape, we use the same four sources as considered for the signal distribution. The uncertainty due to the choice of parameterization is estimated by using the Keys PDF as an alternative. The effect of correlation between the

$(M_{\rm bc}, \Delta E)$ and $(\cos\theta_{\rm thr}, \mathcal{F})$ distributions is estimated by using a distribution split into the sum of two components (u, d, s and charm contributions) with independent $(M_{\rm bc}, \Delta E)$ and $(\cos\theta_{\rm thr}, \mathcal{F})$ shapes. The uncertainty due to the MC description of the $(M_{\rm bc}, \Delta E)$ and $(\cos\theta_{\rm thr}, \mathcal{F})$ distributions is estimated by floating their parameters in the $B^\pm \to D\pi^\pm$ fit. To estimate the effect of correlation of the shape with the Dalitz plot variables we fit the shapes separately in each Dalitz plot bin.

The uncertainties due to the shapes of random and peaking $B\overline{B}$ backgrounds are estimated differently for the $B^\pm \to D\pi^\pm$ and $B^\pm \to DK^\pm$ samples. In the $B^\pm \to D\pi^\pm$ case, the effect of the background shape uncertainty is estimated by performing a fit with $\Delta E > -0.1$ GeV: this requirement rejects the peaking $B\overline{B}$ background and a large part of the random $B\overline{B}$ background. In the case of the $B^\pm \to DK^\pm$ sample, the uncertainty is estimated by performing an alternative fit with the $(M_{\rm bc}, \Delta E)$ and $(\cos\theta_{\rm thr}, \mathcal{F})$ shapes taken from the $B^\pm \to D\pi^\pm$ sample (shifted by 50 MeV in ΔE to account for difference in pion and kaon masses in the calculation of ΔE).

In the case of the fit to the $B^\pm \to DK^\pm$ sample, the uncertainty due to the $B^\pm \to D\pi^\pm$ background shape in the $(\cos\theta_{\rm thr}, \mathcal{F})$ variables is estimated by taking the $(\cos\theta_{\rm thr}, \mathcal{F})$ shape for signal events. The Dalitz plot distribution uncertainty is estimated by using the number of flavor-tagged events in bins, rather than the number of $B^\pm \to D\pi^\pm$ events used in the default fit. Uncertainties due to correlations are treated in the same way as in the case of the signal distribution.

There is an uncertainty due to the Dalitz plot efficiency shape because of the difference in average efficiency over each bin for the flavor and $B^\pm \to DK^\pm$ samples. A maximum difference of 1.5% is obtained in a MC study. The uncertainty is taken to be the maximum of the two comparable in size quantities:

- the root mean square (RMS) of x and y from smearing the numbers of events in the flavor sample K_i by 1.5% (larger for y parameters);
- the bias of x and y between fits with and without efficiency correction for K_i obtained from signal MC (larger for x parameters).

The uncertainty due to cross-feed of events between bins is estimated by varying the momentum resolution by 20% — the MC resolution scaling factor obtained from the fit to $B^\pm \to D\pi^\pm$ sample — and by taking the bias between the fits with and without final state interactions taken into account.

The uncertainty arising from the finite sample of flavor-tagged $D \to K^0_S\pi^+\pi^-$ decays is evaluated by varying the values of K_i within their statistical errors.

The final results for (x, y) are corrected for the fit bias obtained from fits of MC pseudo-experiments. The uncertainty due to the fit bias is taken from the difference of biases for various input values of x and y.

The uncertainty due to the limited precision of the 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. We have performed a study of this procedure using both MC pseudo-experiments and analytical calculations. We find that the uncertainty obtained in this way is sample-dependent for small B data samples and its average value scales in inverse proportion to the square root of the sample size. It reaches a constant value for large B data samples (in the systematics-dominated case). This explains the somewhat higher uncertainty compared to the CLEO estimate given in [14], which was obtained in the limit of a very large B sample. In addition, the uncertainty in (x, y) is proportional to r_B, and, thus, the uncertainty in the phases ϕ_3 and δ_B is independent of r_B. As a result, the uncertainty of (x, y) in the $B^\pm \to DK^\pm$ sample fit is 3–4 times larger than in the $B^\pm \to D\pi^\pm$ sample.

X. RESULTS FOR ϕ_3, r_B AND δ_B

We use the frequentist approach with the Feldman-Cousins ordering [21] to obtain the physical parameters $\mu = (\phi_3, r_B, \delta_B)$ from the measured parameters $z = (x_-, y_-, x_+, y_+)$, as was done in previous Belle analyses [9, 10]. In essence, the confidence level α for a set of physical parameters μ is calculated as

$$\alpha(\mu) = \int\limits_{\mathcal{D}(\mu)} p(z|\mu)dz \bigg/ \int\limits_{\infty} p(z|\mu)dz \,, \qquad (14)$$

where $p(z|\mu)$ is the probability density to obtain the measurement result z given the set of physics parameters μ. The integration domain $\mathcal{D}(\mu)$ is given by the likelihood ratio (Feldman-Cousins) ordering:

$$\frac{p(z|\mu)}{p(z|\mu_{\rm best}(z))} > \frac{p(z_0|\mu)}{p(z_0|\mu_{\rm best}(z_0))}, \qquad (15)$$

where $\mu_{\rm best}(z)$ is μ that maximizes $p(z|\mu)$ for the given z, and z_0 is the result of the data fit.

In contrast to previous Belle analyses [9, 10], the probability density $p(z|\mu)$ is a multivariate Gaussian PDF with the errors and correlations between $x_\pm$ and $y_\pm$ taken from the data fit result. In the previous analyses, this PDF was taken from MC pseudo-experiments.

As a result of this procedure, we obtain the confidence levels (CL) for the set of physical parameters ϕ_3, r_B, and δ_B. The confidence levels for one and two standard deviations are taken at 20% and 74% (appropriate for the case of a three-dimensional Gaussian distribution). The projections of the 3D surfaces bounding one, two, and three standard deviations volumes onto (ϕ_3, r_B) and (ϕ_3, δ_B) planes are shown in Fig. 9.

Systematic errors in μ are obtained by varying the measured parameters z within their systematic errors (Gaussian distributions are assumed) and calculating the RMS of $\mu_{\rm best}(z)$. In this calculation we assume that the systematic errors are uncorrelated between the B^+ and B^-

TABLE VI. Systematic uncertainties in the (x, y) measurements for $B^\pm \to D\pi^\pm$ and $B^\pm \to DK^\pm$ samples in units of 10^{-3}.

	$B^\pm \to D\pi^\pm$				$B^\pm \to DK^\pm$			
Source of uncertainty	Δx_-	Δy_-	Δx_+	Δy_+	Δx_-	Δy_-	Δx_+	Δy_+
Signal shape	0.9	1.9	1.1	5.0	7.3	7.4	7.3	5.1
u, d, s, c continuum background	0.9	1.4	0.8	1.3	6.7	5.6	6.6	3.2
$B\overline{B}$ background	3.3	1.6	4.5	1.1	7.7	8.4	7.4	5.4
$B^\pm \to D\pi^\pm$ background	–	–	–	–	1.2	4.2	1.9	1.9
Dalitz plot efficiency	3.0	1.9	3.2	1.6	4.8	2.0	5.6	2.1
Crossfeed between bins	0.3	0.6	0.1	0.7	0.0	3.9	0.1	1.0
Flavor-tagged statistics	1.7	2.0	1.6	2.0	1.5	2.7	1.7	1.9
Fit bias	0.4	0.5	0.4	0.5	3.2	5.8	3.2	5.8
c_i and s_i precision	2.6	6.5	2.6	6.5	10.1	22.5	7.2	17.4
Total without c_i,s_i precision	4.9	4.1	6.0	5.9	14.0	15.3	14.1	10.6
Total	5.6	7.7	6.5	8.8	17.3	27.2	15.9	20.4

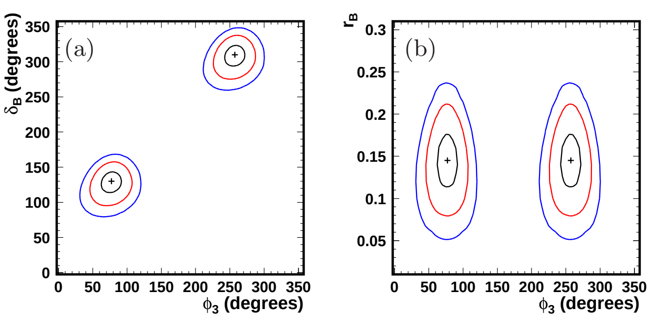

FIG. 9. Two-dimensional projections of confidence region onto (ϕ_3, δ_B) and (ϕ_3, r_B) planes (one, two, and three standard deviations, statistical only).

samples. In the case of c_i and s_i systematics, we test this assumption. When the fluctuation in c_i and s_i is generated, we perform fits to both B^+ and B^- data with the same fluctuated (c_i, s_i). We observe no significant correlation between the resulting x_- and x_+ (y_- and y_+).

The final results are:

$$\begin{aligned} \phi_3 &= (77.3^{+15.1}_{-14.9} \pm 4.1 \pm 4.3)^\circ \\ r_B &= 0.145 \pm 0.030 \pm 0.010 \pm 0.011 \\ \delta_B &= (129.9 \pm 15.0 \pm 3.8 \pm 4.7)^\circ, \end{aligned} \tag{16}$$

where the first error is statistical, the second is systematic error without c_i and s_i uncertainty, and the third error is due to c_i and s_i uncertainty. Extraction of ϕ_3, r_B, and δ_B has a two-fold ambiguity, (ϕ_3, δ) and $(\phi_3+180^\circ, \delta+180^\circ)$, leading to the same values of $x_\pm$ and $y_\pm$. Here we choose the solution that satisfies $0 < \phi_3 < 180^\circ$.

The significance of CP violation (ϕ_3 being non-zero) is calculated as the CL of the point $\phi_3 = 0$. This calculation accounts for a small deviation from Gaussian errors for x and y observed and parameterized using a large number of MC pseudo-experiments. The statistical significance equals 99.64% or 2.9 standard deviations. This value is in good agreement with the χ^2 probability from the difference of the number of events in bins for B^+ and B^- data. With the systematic uncertainties included, the significance decreases to 99.35% or 2.7 standard deviations.

XI. CONCLUSION

We report the results of a measurement of the Unitarity Triangle angle ϕ_3 using a model-independent Dalitz plot analysis of $D \to K^0_S\pi^+\pi^-$ decay in the process $B^\pm \to DK^\pm$. The measurement was performed with the full data sample of 711 fb^{-1} (772×10^6 $B\overline{B}$ pairs) collected by the Belle detector at the $\Upsilon(4S)$ resonance. Model independence is achieved by binning the Dalitz plot of the $D \to K^0_S\pi^+\pi^-$ decay and using the strong-phase coefficients for bins measured by the CLEO experiment [14]. We obtain the value $\phi_3 = (77.3^{+15.1}_{-14.9} \pm 4.1 \pm 4.3)^\circ$; of the two possible solutions we choose the one with $0 < \phi_3 < 180^\circ$. We also obtain the value of the amplitude ratio $r_B = 0.145 \pm 0.030 \pm 0.010 \pm 0.011$. In both results, the first error is statistical, the second is systematic error without c_i and s_i uncertainty, and the third error is due to c_i and s_i uncertainty.

This analysis is the first application of a novel method for measuring ϕ_3. Compared to the result of the model-dependent measurement performed by Belle with the $B^\pm \to DK^\pm$ mode, $\phi_3 = (80.8^{+13.1}_{-14.8} \pm 5.0(\text{syst}) \pm 8.9(\text{model}))^\circ$ [10], this measurement has somewhat poorer statistical precision despite a larger data sample used. There are two factors responsible for lower statistical sensitivity: 1) the statistical error for the same statistics is inversely proportional to the r_B value, and the central value of r_B in this analysis is smaller, and

2) the binned approach is expected to have the statistical precision that is, on average, 10–20% poorer than the unbinned one [12].

More important is that the large model uncertainty of the model-dependent result (8.9°) is replaced by the purely statistical uncertainty of 4.3° due to the limited size of the CLEO $\psi(3770)$ data sample. Although the model-independent approach does not offer significant improvement over the unbinned model-dependent Dalitz plot analysis with the current data sample, it is promising for future measurements at super flavor factories [22, 23] and LHCb [24]. We expect that the statistical error of the ϕ_3 measurement using the statistics of a 50 ab^{-1} data sample that will be available at a super-B factory will reach 1–2°. With the use of BES-III data [25] the error due to the phase terms in the $D \to K^0_S\pi^+\pi^-$ decay will decrease to 1° or less. We also expect that the experimental systematic error can be kept at the level below 1°, since most of its sources are limited by the statistics of the control channels.

ACKNOWLEDGMENTS

We thank the KEKB group for the excellent operation of the accelerator; the KEK cryogenics group for the efficient operation of the solenoid; and the KEK computer group, the National Institute of Informatics, and the PNNL/EMSL computing group for valuable computing and SINET4 network support. We acknowledge support from the Ministry of Education, Culture, Sports, Science, and Technology (MEXT) of Japan, the Japan Society for the Promotion of Science (JSPS), and the Tau-Lepton Physics Research Center of Nagoya University; the Australian Research Council and the Australian Department of Industry, Innovation, Science and Research; the National Natural Science Foundation of China under contract No. 10575109, 10775142, 10875115 and 10825524; the Ministry of Education, Youth and Sports of the Czech Republic under contract No. LA10033 and MSM0021620859; the Department of Science and Technology of India; the Istituto Nazionale di Fisica Nucleare of Italy; the BK21 and WCU program of the Ministry Education Science and Technology, National Research Foundation of Korea, and GSDC of the Korea Institute of Science and Technology Information; the Polish Ministry of Science and Higher Education; the Ministry of Education and Science of the Russian Federation and the Russian Federal Agency for Atomic Energy; the Slovenian Research Agency; the Swiss National Science Foundation; the National Science Council and the Ministry of Education of Taiwan; and the U.S. Department of Energy and the National Science Foundation. This work is supported by a Grant-in-Aid from MEXT for Science Research in a Priority Area ("New Development of Flavor Physics"), and from JSPS for Creative Scientific Research ("Evolution of Tau-lepton Physics").

This research is partially funded by the Russian Presidential Grant for support of young scientists, grant number MK-1403.2011.2.

APPENDIX

The results of the combined fit to $B^+ \to DK^+$ and $B^- \to DK^-$ samples separately for each bin of the Dalitz plot are shown in Figs. 10 and 11, respectively. The plots show the projections of the data and the fitting model on the ΔE variable, with the additional requirements $M_{\rm bc} > 5270$ MeV/c^2 and $\cos\theta_{\rm thr} < 0.8$.

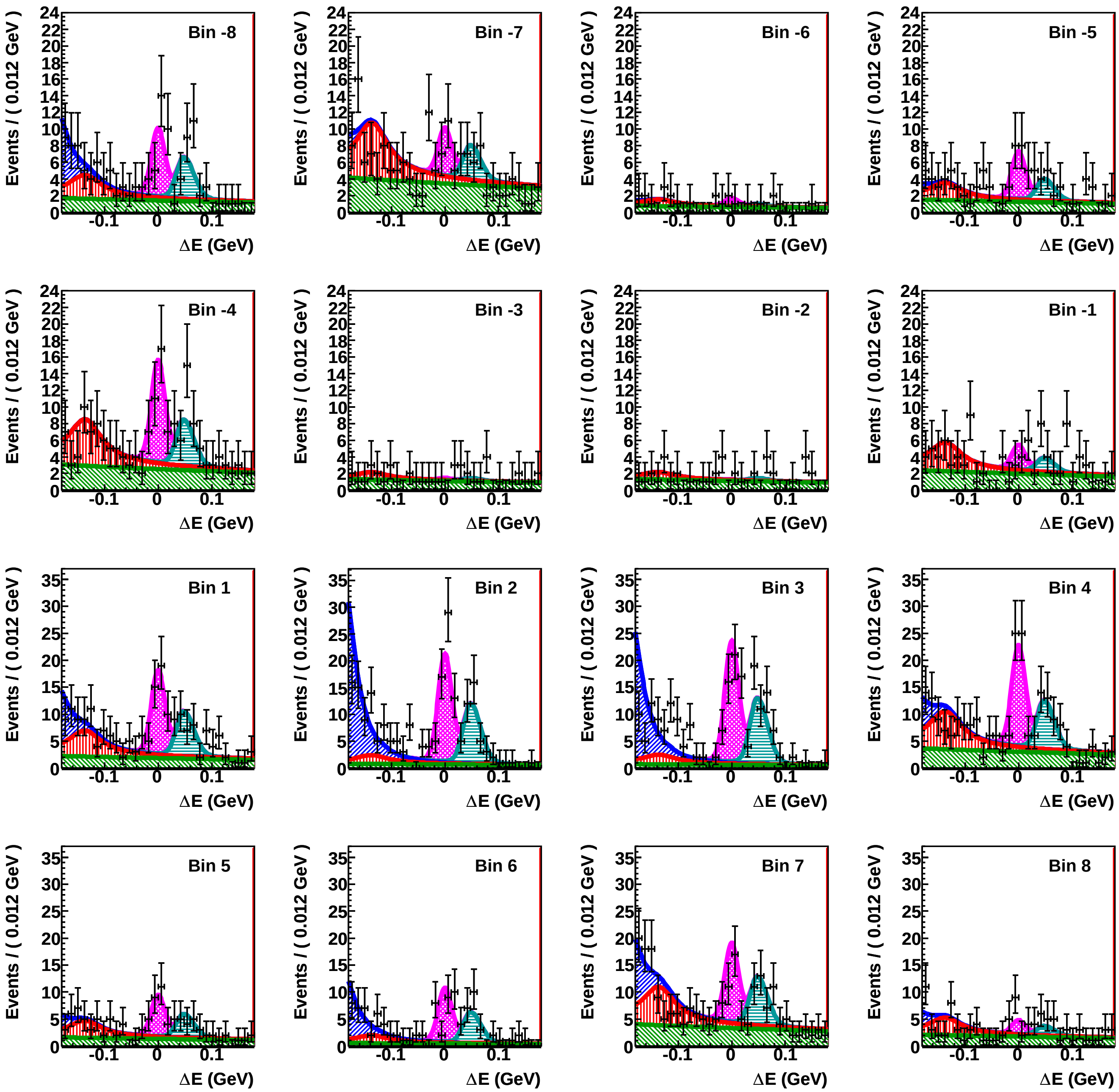

FIG. 10. Projections of the combined fit of the $B^+ \to DK^+$ sample on ΔE for each Dailtz plot bin, with the $M_{\rm bc} > 5270$ MeV$/c^2$ and $\cos\theta_{\rm thr} < 0.8$ requirements. The fill styles for the signal and background components are the same as in Fig. 5.

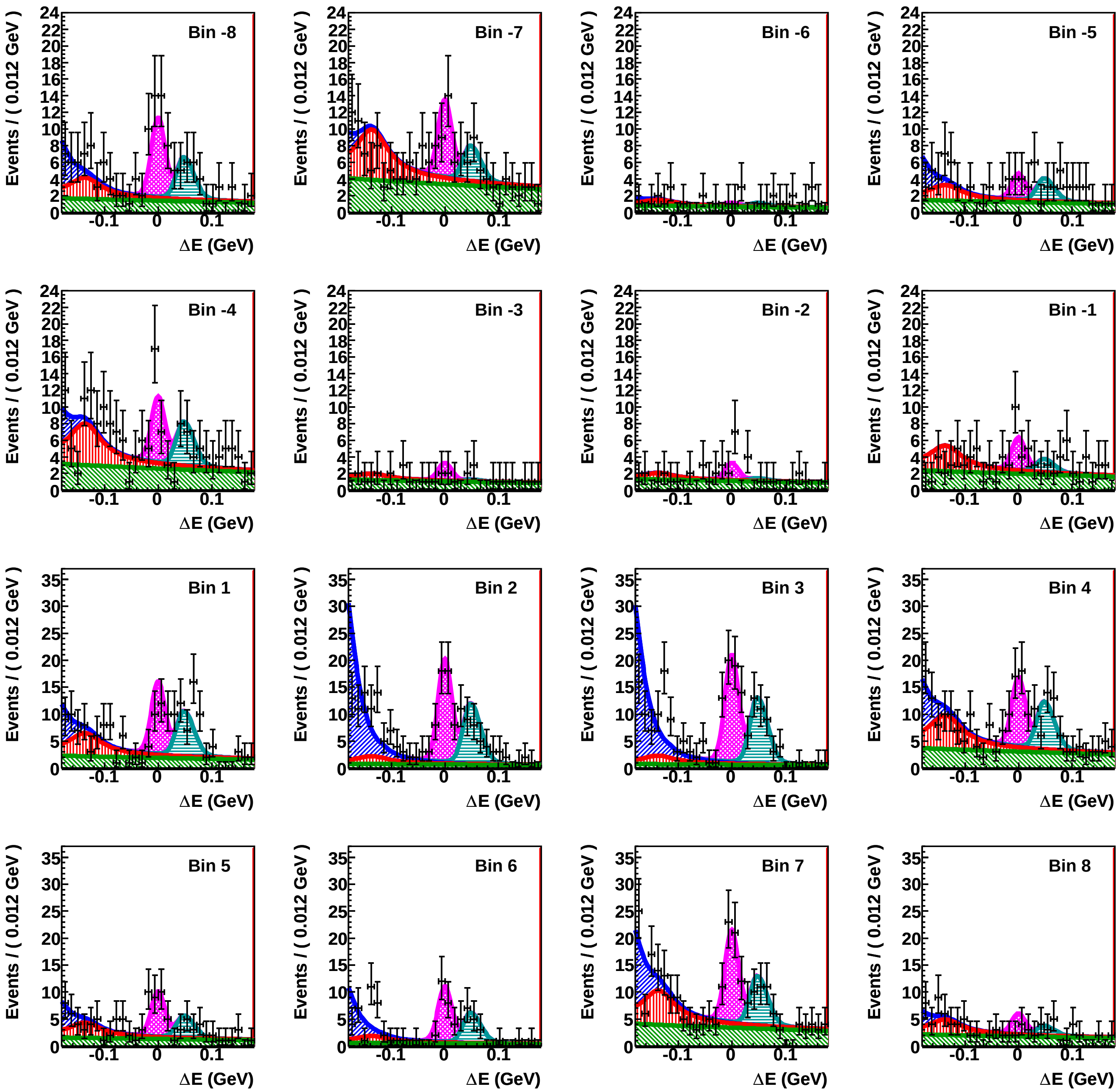

FIG. 11. Projections of the combined fit of the $B^- \to DK^-$ sample on ΔE for each Dailtz plot bin, with the $M_{\rm bc} > 5270$ MeV$/c^2$ and $\cos\theta_{\rm thr} < 0.8$ requirements. The fill styles for the signal and background components are the same as in Fig. 5.

[1] A. Giri, Yu. Grossman, A. Soffer, J. Zupan, Phys. Rev. D **68**, 054018 (2003).
[2] A. Bondar. Proceedings of BINP Special Analysis Meeting on Dalitz Analysis, 24-26 Sep. 2002, unpublished.
[3] J. Zupan, arXiv:1101.0134 (hep-ph).
[4] I.J.R. Aitchison, Nucl. Phys. A **189**, 417 (1972),
[5] BaBar Collaboration, B. Aubert, *et al.*, Phys. Rev. Lett. **95**, 121802 (2005).
[6] BaBar Collaboration, B. Aubert, *et al.*, Phys. Rev. D **78**, 034023 (2008).
[7] BaBar Collaboration, P. del Amo Sanchez, *et al.*, Phys. Rev. Lett **105**, 121801 (2010).
[8] Belle Collaboration, A. Poluektov, *et al.*, Phys. Rev. D **70**, 072003 (2004).
[9] Belle Collaboration, A. Poluektov, *et al.*, Phys. Rev. D **73**, 112009 (2006).
[10] Belle Collaboration, A. Poluektov, A. Bondar, B. Yabsley, *et al.*, Phys. Rev. D **81**, 112002 (2010).
[11] A. Bondar and A. Poluektov, Eur. Phys. J. C **47**, 347 (2006).
[12] A. Bondar and A. Poluektov, Eur. Phys. J. C **55**, 51 (2008).
[13] CLEO Collaboration, R.A. Briere, *et al.*, Phys. Rev. D **80**, 032002 (2009).
[14] CLEO Collaboration, J. Libby, *et al.*, Phys. Rev. D **82**, 112006 (2010).
[15] A. Bondar, A. Poluektov and V. Vorobiev, Phys. Rev. D **82**, 034033 (2010).
[16] Belle Collaboration, A. Abashian, *et al.*, Nucl. Instr. and Meth. A **479**, 117 (2002).
[17] Z. Natkaniec *et al.* (Belle SVD2 Group), Nucl. Instr. and Meth. A **560**, 1 (2006).
[18] CLEO Collaboration, D.M. Asner, *et al.*, Phys. Rev. D **53**, 1039 (1996).
[19] ARGUS Collaboration, H. Albrecht, *et al.*, Phys. Lett. B **241**, 278 (1990).
[20] K.S. Cranmer, Comput. Phys. Commun. **136**, 198 (2001). [hep-ex/0011057].
[21] G.J. Feldman and R.D. Cousins, Phys. Rev. D **57**, 3873 (1998).
[22] T. Abe, *et al.*, Belle II Technical Design Report, KEK Report 2010-1, arXiv:1011.0352 [physics.ins-det].
[23] SuperB collaboration, SuperB Conceptual Design Report, INFN/AE-07/2, SLAC-R-856, LAL 07-15, arXiv:0709.0451v2 [hep-ex].
[24] J. Libby, CERN-LHCb-2007-141.
[25] BES Collaboration, M. Ablikim *et al.*, Nucl. Instrum. Meth. A **614**, 345 (2010).