Observation of the Decay $B^- \to D^{(*)+} K^- \ell^+ \nu_\ell$

P. del Amo Sanchez,1 J. P. Lees,1 V. Poireau,1 E. Precipice,1 V. Tisserand,1 J. Garra Tico,2 E. Grauges,2 M. Martinelliab,3 A. Palanoab,3 M. Pappagalloab,3 G. Eigen,4 B. Stugu,4 L. Sun,4 M. Battaglia,5 D. N. Brown,5 B. Hooberman,5 L. T. Kerth,5 Yu. G. Kolomensky,5 G. Lynch,5 I. L. Osipenkov,5 T. Tanabe,5 C. M. Hawkes,6 A. T. Watson,6 H. Koch,7 T. Schroeder,7 D. J. Asgeirsson,8 C. Hearty,7 T. S. Mattison,8 J. A. McKenna,8 A. Khan,9 A. Randle-Conde,9 V. E. Blinov,10 A. R. Buzykaev,10 V. P. Druzhinin,10 V. B. Golubev,10 A. P. Onuchin,10 S. I. Serezhakov,10 Yu. I. Skopnen,10 E. P. Solodov,10 K. Yu. Todyshev,10 A. N. Yushkov,10 M. Bondioli,11 S. Curry,11 D. Kirkby,11 A. J. Lankford,11 M. Mandelkern,11 E. C. Martin,11 D. P. Stoker,11 H. Atmacan,12 J. W. Gary,12 F. Liu,12 O. Long,12 G. M. Vitug,12 C. Campagnari,13 T. M. Hong,13 D. Kovalski,13 J. D. Richman,13 A. M. Eisner,14 C. A. Heusch,14 J. Kroseberg,14 W. S. Lockman,14 A. J. Martinez,14 T. Schalk,14 B. A. Schumm,14 A. Seiden,14 L. O. Winstron,14 C. H. Cheng,15 D. A. Doll,15 B. Echenard,15 D. G. Hiltin,15 P. Ongmonkolkul,15 F. C. Porter,15 A. Y. Rakitin,15 R. Andreassen,16 M. S. Dubrovin,16 G. Mancinelli,16 B. T. Meadows,16 M. D. Sokoloff,16 P. C. Bloom,17 W. T. Ford,17 A. Gaz,17 M. Nagel,17 U. Nauenberg,17 J. G. Smith,17 S. R. Wagner,17 R. Ayad,18, * W. H. Toki,18 H. Jasper,19 T. M. Karbach,19 J. Merkel,19 A. Petzold,19 B. Spaan,19 K. Wacker,19 M. J. Kobel,20 K. R. Schubert,20 R. Schwierz,20 D. Bernard,21 M. Verderi,21 P. J. Clark,22 S. Playfer,22 J. E. Watson,22 M. Andreottiab,23 D. Bettoni,23 A. Bozzi,23 R. Calabreseab,23 A. Cecchib,23 G. Cibinettoab,23 E. Fioravantiab,23 P. Franchininab,23 E. Luppiba,23 M. Mumeratoab,23 M. Negrinia,23 A. Petrella,23 L. Piemontese,23 R. Baldini-Ferroli,24 A. Calcaterra,24 R. de Sangro,24 G. Finocchiaro,24 M. Nicolaci,24 S. Pacetti,24 P. Patteri,24 I. M. Peruzzi,24 M. Piccolo,24 M. Rama,24 A. Zollo,24 R. Contrib,25 E. Guidob,25 M. Lo Vetereab,25 M. R. Mongeab,25 S. Passaggioab,25 C. Patrignanib,25 E. Robuttic,25 S. Tosib,25 B. Bhuyan,26 V. Prasad,26 C. L. Lee,27 M. Morii27 A. Adametz28 J. Marks28 S. Schenk,28 U. Uwer,28 F. U. Bernlochner,29 M. Ebert,29 H. M. Lackner,29 T. Lueck,29 A. Volk,29 P. D. Dauneck,30 M. Tibbetts,30 P. K. Behera,31 U. Mallik,31 C. Chen,32 J. Cochran,32 H. B. Crawley,32 L. Dong,32 W. T. Meyer,32 S. Prell,32 E. I. Rosenberg,32 A. E. Rubin,32 Y. Y. Gao,33 A. V. Gritsan,33 Z. J. Guo,33 N. Arnaud,34 M. Davier,34 D. Derkach,34 J. Firmino da Costa,34 G. Grosdidier,34 F. Le Diberder,34 A. M. Lutz,34 B. Malaescu,34 A. Perez,34 P. Roudoune,34 M. H. Schune,34 J. Serrano,34 V. Sordini,34 A. Stocchi,34 L. Wang,34 G. Wormser,34 D. J. Lange,35 D. M. Wright,35 I. Bingham,36 C. A. Chavez,36 J. P. Coleman,36 J. R. Fry,36 E. Gabathuler,36 R. Gamez,36 D. E. Hutchcroft,36 D. J. Payne,36 C. Touramanis,36 A. J. Bevan,37 F. Di Lodovico,37 R. Saccob,37 M. Sigamani,37 G. Cowan,38 S. Parmesanvar,38 A. C. Wren,38 D. N. Brown,39 C. L. Davis,39 A. G. Denig,40 M. Fritsch,40 W. Gradl,40 A. Hafner,40 K. E. Alwyn,41 D. Bailey,41 R. J. Barlow,41 G. Jackson,41 G. D. Lafferty,41 T. J. West,41 J. Anderson,42 R. Cenci,42 A. Jawahery,42 D. A. Roberts,42 G. S. G. Simi,42 J. M. Tuggle,42 C. Dallapiccola,43 E. Salvati,43 R. Cowan,44 D. Dujmic,44 P. H. Fisher,44 G. Sciolla,44 M. Zhao,44 D. Lindemann,45 P. M. Patel,45 S. H. Robertson,45 M. Schram,45 P. Biassoniab,46 A. Lazzaroab,46 V. Lombardoab,46 F. Palomboab,46 S. Strackab,46 L. Cremaldb,46 R. Godang,47, 5 R. Krokeger,47 P. Sonneck,47 D. J. Summers,47 X. Nguyen,48 M. Simard,48 P. Taras,48 G. De Nardoab,49 D. Monorchioab,49 G. Onoratoab,49 C. Sciacciab,49 G. Raven,50 H. L. Snoek,50 C. P. Jessop,50 K. J. Knoepfel,51 J. M. LoSecco,51 F. W. Wang,51 L. A. Corwin,52 K. Kouscheid,52 R. Kass,52 J. P. Morris,52 A. M. Rahimi,52 N. L. Blount,53 J. Braun,53 R. Frey,53 O. Igkonkina,53 J. A. Kolb,53 R. Rahmat,53 N. B. Sinev,53 D. Strom,53 J. Strube,53 E. Torrence,53 G. Castellia,54 E. Feltessei,54 N. Gagliardiab,54 M. Margoniab,54 M. Morandinab,54 M. Posoccoab,54 M. Rotondoab,54 F. Simonettia,54 R. Stroilib,54 E. Ben-Haimb,55 G. R. Bonneauab,55 H. Briand,55 G. Calderini,55 J. Chauveau,55 O. Hamon,55 Ph. Lerusette,55 G. Marchiorib,55 J. Ocarizb,55 J. Prendkib,55 S. Sitt,55 M. Biasinib,55 E. Manonib,56 A. Rossib,56 C. Angelinib,57 G. Batignaniab,57 S. Bettarinib,57 M. Carpinellib,57 G. Casarosab,57 A. Cervellib,57 F. Fortib,57 M. A. Giorgiab,57 A. Lusianiac,57 N. Neriab,57 E. Paoloniab,57 G. Rizzob,57 J. J. Walsh,57 D. Lopes Pegnab,58 C. Luan,58 J. Olsen,58 A. J. S. Smith,58 A. V. Telnob,58 F. Anullib,58 E. Baracchiniab,59 G. Cavob,59 R. Facchinib,59 F. Ferrarotab,59 F. Ferronib,59 M. Gasperob,59 L. Li Giao,59 M. A. Mazzoniab,59 G. Pireddaa,59 F. Renub,59 T. Hartmannb,60 T. Ledig,60 H. Schroeder,60 R. Waldib,60 T. Adyeb,61 F. Franke,61 E. O. Olaiyab,61 F. F. Wilson,61 S. Emery,62 G. Hamel de Monchenault,62 G. Vasseur,62 Ch. Yèche,62 M. Zito,62 M. T. Allen,63 D. Astone,63 D. J. Bard,63 R. Bartoldus,63 J. F. Benitez,63 C. Cartaro,63 M. R. Convery,63
J. Dorfan, D. P. Dubois-Felsmann, W. Dunwoodie, R. C. Field, M. Franco Sevilla, B. G. Fulsom, A. M. Gabareen, T. M. Graham, P. Grenier, C. Hast, W. R. Innes, M. H. Kelsey, H. Kim, P. Kim, M. L. Kocian, D. W. G. S. Leith, S. Li, B. Lindquist, S. Luitz, V. Luth, H. L. Lynch, D. B. MacFarlane, H. Marisiske, D. R. Muller, H. Neal, S. Nelson, C. P. O’Grady, I. Ofte, M. Perl, T. Pulliam, B. N. Ratchiff, A. Roodman, A. A. Sahinkov, V. Santoro, R. H. Schindler, J. Schiewing, A. Snyder, D. Su, M. K. Sullivan, S. Sun, K. Suzuki, J. M. Thompson, J. Va’vra, A. P. Wagner, M. Weaver, C. A. West, W. J. Wisniewski, M. Wittgen, D. H. Wright, H. W. Wulsin, A. K. Yarritu, C. C. Young, V. Ziegler, X. R. Chen, W. Park, M. V. Purohit, R. M. White, J. R. Wilson, S. J. Sekula, M. Bellis, P. R. Burchat, A. J. Edwards, T. S. Miyashita, S. Ahmed, M. S. Alam, J. A. Ernst, B. Pan, M. A. Saeed, S. B. Zain, N. Guttman, A. Soffer, P. Lund, S. M. Spanier, R. E. Eckmann, J. L. Ritchie, A. M. Ruland, J. C. Schilling, R. F. Schwitters, B. C. Wray, J. M. Izen, X. C. Lou, F. Bianchi, D. Gambino, M. Pelliccioni, M. Bomben, L. Lanceri, L. Vitale, N. Lopez-March, F. Martinez-Vidal, D. A. Milanes, A. Oyanguren, J. Albert, Sw. Banerjee, H. H. F. Choi, K. Hamano, G. J. King, R. Kowalewski, M. J. Lewczuk, I. M. Nugent, R. J. Sobie, T. J. Gershon, P. F. Harrison, T. E. Latham, E. M. T. Pucio, H. R. Band, S. Dasu, K. T. Flood, Y. Pan, R. Prepost, C. O. Vuosalo, and S. L. Wu

(The BaBar Collaboration)
We report the observation of the decay $B^{-} \to D_{s}^{(*)+} K^{-} \ell^{+} \bar{\nu}_{\ell}$ based on 342 fb$^{-1}$ of data collected at the $T(4S)$ resonance with the BABAR detector at the PEP-II $e^{+}e^{-}$ storage rings at SLAC. A simultaneous fit to three $D_{s}$ decay chains is performed to extract the signal yield from measurements of the squared missing mass in the $B$ meson decay. We observe the decay $B^{-} \to D_{s}^{(*)+} K^{-} \ell^{+} \bar{\nu}_{\ell}$ with a significance greater than five standard deviations (including systematic uncertainties) and measure its branching fraction to be $B(B^{-} \to D_{s}^{(*)+} K^{-} \ell^{+} \bar{\nu}_{\ell}) = [6.13^{+1.04}_{-1.05}(\text{stat.}) \pm 0.43(\text{syst.}) \pm 0.51(B(D_{s}))] \times 10^{-4}$, where the last error reflects the limited knowledge of the $D_{s}$ branching fractions.

PACS numbers: 13.20.He, 12.15.Ji, 12.38.Qk

The study of charmed inclusive semileptonic $B$ meson decays enables the measurement of the CKM matrix element $|V_{cb}|$. This measurement relies on a precise knowledge of all semileptonic $B$ meson decays. Decays of orbitally excited $D$ mesons, from the process $B \to D^{(*)+}\ell\nu$, constitute a significant fraction of these decays [1], and may help explain the discrepancy between the inclusive $B \to X_{c}\ell\nu$ rate, where $X_{c}$ is a charmed hadronic final state, and the sum of the measured exclusive decay rates [1, 2]. So far, analyses of these decays have focused on the reconstruction of $D \to D^{(*)}\pi\ell\nu$ states [3–5]. In such analyses, experimental data are interpreted as a sum of the four $D^{(*)}$ resonances. The results show the dominance of $B$ decays to broad resonances, while QCD sum rules imply the opposite [6]. Conversely, a small contribution from broad $D^{(*)}$ states implies the presence
of a non-resonant $B \to D^{(*)}\pi\ell\nu$ component, which has not yet been observed. Measurement of the branching fraction for the as-yet-unobserved $B^- \to D_s^{(*)-}K^-\ell^-\overline{\nu}_\ell$ decay [7] would provide additional information relevant to this issue, by exploring the hadronic mass distribution above 2.46 GeV/$c^2$ where resonant and non-resonant components are present. In addition, the measurement of $B^- \to D_s^{(*)-}K^-\ell^-\overline{\nu}_\ell$ will provide a better estimate of background in future studies of semileptonic $B_s \to D_s^+ \pi^-\ell^-\overline{\nu}_\ell$ decays.

Using the shape of the hadronic mass spectrum in $B$ semileptonic decays, a rough estimate on the branching fraction $\mathcal{B}(B^- \to D_s^{(*)-}K^-\ell^-\overline{\nu}_\ell)$ is of the order of $10^{-3}$ [8, 9], which is consistent with the limit set by the ARGUS Collaboration, $\mathcal{B}(B^- \to D_s^{(*)-}K^-\ell^-\overline{\nu}_\ell) < 5 \times 10^{-3}$ at 90% confidence level [10]. A comparison between this expectation and the actual measurement can confirm or refute the expected rapid decrease of the hadronic mass distribution at high values.

In this paper, we present the observation of $B^- \to D_s^{(*)-}K^-\ell^-\overline{\nu}_\ell$ decays, where $\ell = e, \mu$. This analysis does not differentiate between final states with $D_s^{(*)}$ and $D_s^{(')}$, where $D_s^{(*)}$ decays via emission of neutral decay products that are not reconstructed. The results are based on a data sample of $N_{\ell\nu\overline{\nu}} = (376.9 \pm 4.1) \times 10^6 B\overline{B}$ pairs recorded at the $\Upsilon(4S)$ resonance with the BaBar detector [11] at the PEP-II asymmetric energy $e^+e^-$ storage rings at the SLAC National Accelerator Laboratory. This corresponds to an integrated luminosity of $342 \, fb^{-1}$. In addition, $37 \, fb^{-1}$ of data collected about $40 \, MeV$ below the resonance are used for background studies. A GEANT-based Monte Carlo (MC) simulation [12] of $B\overline{B}$ and continuum events ($e^+e^- \to q\overline{q}$ with $q = u, d, s, c$) is used to study the detector response and acceptance, validate the analysis technique, and evaluate signal efficiencies. The sample of simulated $B\overline{B}$ events is equivalent to approximately 3 times the data sample. The signal MC events are generated by adapting the decay model of Goity and Roberts [13] to describe $D_s^{(*)}K^-$ final states. Two alternative signal MC samples are used to estimate systematic uncertainties: a sample based on the ISGW2 model [14], in which $B^-\pi$ mesons decay to $D_0^{(*)}\ell\overline{\nu}$ with $D_0^{(*)}\to D_s^{(*)}K^-$, and a sample based on a simple phase space model. The signal MC samples are equivalent to approximately 10 times the expected signal yield.

We reconstruct $D_s^+$ candidates in three decay chains: $D_s^+ \to \phi\pi^+\ell\nu$ with $\phi \to K^+K^-$, $D_s^+ \to \overline{K}^0K^+$ with $\overline{K}^0 \to K^-\pi^+$, and $D_s^+ \to K^0K^+$ with $K^0 \to \pi^+\pi^-$. The $\phi$, $\overline{K}^0$, and $K^0$ candidates are formed by combining oppositely charged tracks. To suppress combinatorial background from the $D_s^+$ reconstruction in the first two decay chains, we employ a feed-forward neural network (multilayer perceptron, MLP [15]) with three input variables and four hidden layers. The input variables are the absolute value of the difference between the reconstructed and the nominal mass values of the $\phi/K^0\overline{\nu}$ candidate [1], the absolute value of the cosine of the helicity angle of the $\phi/K^0\overline{\nu}$, and the $\chi^2$ probability of the fit to the $D_s^+$ candidate. The helicity angle is defined as the angle between the $D_s^+$ candidate and one kaon originating from the $\phi/K^0\overline{\nu}$ in the $\phi/K^0\overline{\nu}$ rest frame. To suppress combinatorial background in the $D_s^+ \to K^0K^+$ decay chain, we require the invariant mass of the charged pions forming the $K^0$ candidate to satisfy $0.490 \, GeV/c^2 < m(\pi\pi) < 0.506 \, GeV/c^2$, the flight length of the $K^0$ to be larger than 1 mm, the cosine of the laboratory angle between the $K^0$ momentum and the line connecting the $K^0$ decay vertex and the primary vertex of the event to be positive, and the probability of the $D_s^+$ candidate’s vertex fit to be larger than 0.001. The selection criteria are optimized to maximize the statistical significance of the signal. No requirement on the mass of the $D_s^+$ candidates is applied, since this distribution is used to extract the signal yield.

A lepton and a kaon, both with negative charge, are combined with the $D_s^+$ candidate to form a $B^-$ candidate. Leptons are required to have momentum $|p_\ell| > 0.8 \, GeV/c$ [16] to reject those not directly originating from $B$ mesons. The probability of the vertex fit of the $B$ candidate is required to be larger than 0.01.

Three event-shape variables that are sensitive to the topological differences between jet-like continuum events and more spherical $B\overline{B}$ events are used as input to a neural network to suppress background from continuum events. These variables are the normalized second Fox-Wolfram moment $R_2$ [17], the monomial $L_2$ [18], and the cosine of the angle between the flight direction of the reconstructed $B$ candidate and the rest of the event. A neural network whose input variables are the $B$ candidate mass, the $B$ candidate sphericity, and the thrust value of the rest of the event, is used to reduce the background from other $B$ decays, providing a slight, but not negligible improvement in the sensitivity of the measurement.

After applying these selection criteria, the remaining background events are divided into two classes, depending on whether or not they contain a correctly reconstructed $D_s^+$ meson. The first class is the more important of the two. We refer to it as $D_s^+$ background events in the following. Most of these events contain a $D_s$ originating from decays such as $B \to D_sD_s$, where the kaon and lepton tracks used to form a $B$ candidate are taken from the other $B$ meson in the event. The angular correlation between the flight directions of the $D_s$ and the $D$ is used to suppress the $D_s$ background candidates. The direction of the $D$ meson is estimated from the direction of a previously unused charged or neutral kaon candidate that is assumed to be from $D \to K^\pm\pi^\mp\ell\nu\overline{\nu}$ decays. Requiring the cosine of the angle between the flight direction of the $D_s$ candidate and the additional kaon to be larger than $-0.5$, about 30% of the $D_s$-background events are rejected, as shown in Fig. 1. About 8% of the remaining
background events are modeled using a Fermi function:

\[ f(M^2_m) = \frac{1}{\epsilon |(M^2_{m} - M^2_0)/E_C| + 1}, \]  

(2)

where \( M^2_0 \) represents the \( M^2_m \) drop-off value and \( E_C \) the smearing of the Fermi-edge. The values for \( M^2_0 \) and \( E_C \), \( M^2_0 = (0.303 \pm 0.034) \) GeV\(^2/c^4 \) and \( E_C = (0.333 \pm 0.018) \) GeV\(^2/c^4 \) for the electron channel, and \( M^2_0 = (0.247 \pm 0.041) \) GeV\(^2/c^4 \) and \( E_C = (0.346 \pm 0.022) \) GeV\(^2/c^4 \) for the muon channel, are fixed to the values derived from fits to MC distributions and are the same for all \( D^+_s \) decay chains. Signal events are modeled by a Gaussian distribution, with the same mean and width for all reconstruction channels. The width is fixed to the value determined from the simulation. The mean of the distribution is determined in the fit, allowing for contributions from events with a \( D^+_s \) in the final state.

The total number of events with a \( D^+_s \) and the number of combinatorial background events in the signal region have been determined from fits to the \( m(D^+_s) \) distributions. The number of signal \( B^- \rightarrow D_s^{(*)+} K^- \ell^- \nu_\ell \) and \( D^+_s \) background events are extracted from the fits to the \( M^2_m \) distributions, separately for the electron and muon samples. For these fits the three \( D^+_s \) decay channels are combined, taking into account their detection branching fractions \( \epsilon_{BR} = B(D_s^+ \rightarrow D^0 d) \times B(D^+ \rightarrow d\bar{s}d4) \) and individual reconstruction efficiencies \( \epsilon_{reco} \). For illustration, these efficiencies and the branching ratios are listed in table I, together with the total fitted number of signal events and the estimated contributions from each of the three channels.

The fit is performed in the range \(|M^2_m| < 1.5 \) GeV\(^2/c^4 \) has 10 free parameters: the mean value of \( M^2_m \), the total number of fitted signal events \( N^{\text{signal}} \), five parameters that describe the shape of the combinatorial background, and three sideband normalization parameters. The number of signal and \( D^+_s \) background events are free in the fit, only the sum of both values is constrained to the result of the fits to the \( m(D^+_s) \) distributions.

The likelihood function is

\[ \mathcal{L} = \frac{e^{-N^{\text{signal}}}}{n!} (N^{\text{signal}})^n \prod_{j}^{N_j} \prod_{i}^{N_{j,i}} \mathcal{P}(M^2_{m,i}, \alpha_j), \]  

(3)

with \( N_j \) the number of events and \( \mathcal{P}(M^2_{m,i}, \alpha_j) \) the probability density function (PDF) for a given fit slice \( j \) (signal region or sideband of each \( D^+_s \) decay chain) with the fit parameters \( \alpha_i \) and \( n = \sum_i N_{j,i} \) the total number of events.

Using MC experiments from a generator, which includes parameterizations of detector performance for signal reconstruction and background expectations, it has been verified that the fit is able to extract signal branching fractions for \( B(B^- \rightarrow D_s^{(*)+} K^- \ell^- \nu_\ell) > 3 \times 10^{-4} \). Values of fit biases are also determined with this procedure and are taken into account in the analysis.
TABLE I: Signal yields, selection efficiencies $\epsilon_{\text{reco}}$, and branching fractions $\epsilon_{BR} = B(D^+_s \to D_1 d_2) \times B(D_1 \to d_3 d_4)$ for the individual and combined decay chains. The signal yields of each decay chain are computed using $N^{\text{signal}}$ and the efficiencies and are given for illustration only. The errors on the signal yields are the fit errors, the uncertainties of $\epsilon_{\text{reco}}$ are the systematic uncertainties and the uncertainties of $\epsilon_{BR}$ represent the limited knowledge of the branching fractions of the $D^+_s$.

| $D^+_s$ decay chain | $N^{\text{signal}}_{\text{electron}}$ | $\epsilon_{\text{reco, electron}}$ [%] | $N^{\text{signal}}_{\text{muon}}$ | $\epsilon_{\text{reco, muon}}$ [%] | $\epsilon_{BR}$ [%] |
|---------------------|--------------------------------------|--------------------------------------|----------------------------------|--------------------------|---------------------|
| all                 | 259.4 $^{+67}_{-67}$               | 209.7 $^{+53}_{-52}$                |                                  |                          |                     |
| $D^+_s \to \phi\pi^+$, $\phi \to K^+K^-$ | 115.7 $^{+30}_{-30}$                | (2.76 $\pm$ 0.08)                 | 92.1 $^{+23}_{-23}$               | (1.62 $\pm$ 0.06)         | (2.18 $\pm$ 0.33)   |
| $D^+_s \to \vec{K}^{\ast 0}K^+$, $\vec{K}^{\ast 0} \to K^+\pi^+$ | 85.2 $^{+22}_{-22}$                | (1.79 $\pm$ 0.06)                 | 70.2 $^{+17}_{-17}$               | (1.09 $\pm$ 0.05)         | (2.60 $\pm$ 0.40)   |
| $D^+_s \to K^{0}_S K^+$, $K^{0}_S \to \pi^+\pi^-$ | 58.5 $^{+15}_{-15}$               | (2.98 $\pm$ 0.08)                 | 47.4 $^{+12}_{-11}$               | (1.78 $\pm$ 0.06)         | (1.02 $\pm$ 0.09)   |

Fit results are given in Table I. Reconstruction efficiencies for the three decay chains are obtained by counting simulated signal events in the range $|M_{m}^2| < 1.2\, \text{GeV}^2/c^4$. As reported in Ref. [19] the reconstruction efficiency of the $D^+_s \to \phi\pi^+$ decay chain depends on the requirement on the $\phi$ mass. The impact of this effect is covered by the systematic uncertainties on $\epsilon_{BR}$. Figure 2 shows the sideband subtracted $M_{m}^2$ distributions summed over the decay channels.

The bias-corrected signal yields are $N^{\text{signal}}_{\text{electron}} = 301^{+68}_{-67}$ and $N^{\text{signal}}_{\text{muon}} = 206^{+53}_{-52}$. The bias correction is $+42 (-4)$ events for the electron (muon) channel. Extended simulations showed that the source of the bias is a fluctuation of the underlying combinatorial background distribution.

FIG. 2: Sideband subtracted $M_{m}^2$ distributions with fitted functions superimposed: (a) for the electron channel and (b) for the muon channel. All $D^+_s$ reconstruction chains have been summed. Solid lines represent the full distribution, dashed lines are the $D^+_s$ background component and dotted lines represent the fitted signal component.

The systematic uncertainties are divided into two categories: additive uncertainties (Table II) are related to the number of extracted signal events, while multiplicative uncertainties (Table III) are related to the calculated branching fraction. The uncertainty due to the $D^+_s$ daughter branching fractions is quoted separately.

The systematic uncertainty arising from the choice of the $D^+_s$ background PDF is evaluated using 100 statistically independent MC experiments. Each experiment corresponds to different values for the two parameters that describe the PDF, $M_{m}^2$ and $E_C$, which are distributed according to the error matrix for these parameters. We take the width of a Gaussian fitted to the resulting $N^{\text{signal}}$ distribution as a systematic uncertainty. The impact of shape differences between data and MC have been studied, as well as shape differences due to varying compositions of the $D^+_s$ background, and both found to be negligible. A similar procedure is used to estimate the uncertainty due to using the $D^+_s$ branching fractions $\epsilon_{BR}$ for the combination of the individual channel signal yields. MC samples of $\epsilon_{BR}$ are produced for each decay channel using the information of Ref. [1]. This leads to differences in the total number of extracted signal events. The width of a Gaussian fitted to the resulting distribution of signal yields is taken as systematic uncertainty.

The width of the Gaussian PDF of $M_{m}^2$ for signal, as well as the number of fitted $D^+_s$, are varied by $\pm$1$\sigma$ to evaluate these systematic uncertainties. This approach also takes into account the variation of the width due to a contribution of $D^{*+}_s$ to the signal yield. The systematic uncertainty related to the bias correction is given by the statistical uncertainty of the correction.

We evaluate the uncertainty of the signal MC model by calculating the difference of the efficiencies between the alternative signal models and the Goity-Roberts signal MC model. The impact of the finite statistics of the simulated signal sample is deduced from the uncertainty on the efficiency determination. The uncertainty arising from particle identification, as well as from the $K^{0}_S$ reconstruction, is determined using dedicated high purity control samples for the corresponding particles. Uncertainties arising from track and photon reconstruction, as
well as from radiative corrections, are evaluated by varying their reconstruction efficiencies and the energy radiated by photons in the simulation. The uncertainty on the number of $B$ mesons in the data set is determined as described in Ref. [20] and the $D_s^+$ daughter branching fraction uncertainties are taken from Ref. [1].

TABLE III: Multiplicative systematic uncertainties in percent.

| source | syst. uncert. electron (muon) ch. [%] | $\phi\pi^+$ | $K^\circ K^+$ | $K_s^0 K^+$ |
|--------|-------------------------------------|------------|-------------|-------------|
| Signal MC model | 7.6 (0.2) | 3.1 (6.4) | 5.9 (2.1) |
| N(Signal MC) | 2.7 (3.5) | 3.3 (4.2) | 2.5 (3.3) |
| Particle ID | 0.6 (1.6) | 1.2 (4.9) | 3.6 (7.9) |
| $K_s^0$ eff. | - (-) | - (-) | 2.0 (3.1) |
| Tracking eff. | 0.4 (0.1) | 0.5 (0.2) | 1.8 (2.4) |
| Photon eff. | 0.6 (0.9) | 0.4 (0.9) | 0.5 (0.7) |
| Radiative corr. | 2.0 (2.1) | 2.2 (2.5) | 1.9 (1.9) |
| Total $\Delta\epsilon_{\text{reco}}$ | 8.4 (4.5) | 5.2 (9.5) | 8.1 (9.9) |
| $B$ counting | 1.1 (1.1) | 1.1 (11) | 1.1 (11) |
| $B(D_s^+)$ | 15.1 (15.1) | 15.4 (15.4) | 6.0 (6.0) |

A second fit, imposing an $\Lambda^{\text{signal}} = 0$ hypothesis, is used to estimate the significance of the measurement. Since the mean of the Gaussian is a free parameter in the signal PDF, the difference in the number of free parameters ($\Delta NDF$) of the fits is larger than one. As shown in Ref. [21], the resulting probability distribution cannot be approximated by a chisquare distribution with an integer number of degrees of freedom. Thus, only a significance range, representing the significances for $\Delta NDF = 2$ and $\Delta NDF = 1$, is calculated. Including statistical and systematic uncertainties, the ranges are $[3.3 - 3.7]\sigma$ and $[3.5 - 3.9]\sigma$ for the electron and muon channel, respectively. Combining both lepton channels results in a significance larger than 5.0$\sigma$.

The branching fractions for the individual lepton channels are $B(B^- \rightarrow D_s^{(*)+} K^- e^- \bar{\nu}_e) = [5.81^{+1.30}_{-1.30} \,(\text{stat.}) \pm 0.54 \,(\text{syst.}) \pm 0.49(B(D_s))] \times 10^{-4}$ and $B(B^- \rightarrow D_s^{(*)+} K^- \mu^- \bar{\nu}_\mu) = [6.68^{+1.72}_{-1.69} \,(\text{stat.}) \pm 0.69 \,(\text{syst.}) \pm 0.56(B(D_s))] \times 10^{-4}$, where the last uncertainty reflects the limited knowledge of the $D_s$ branching fractions. The measurements are combined, taking into account the correlations between their systematic uncertainties, yielding $B(B^- \rightarrow D_s^{(*)+} K^- \ell^- \bar{\nu}_\ell) = [6.13^{+1.04}_{-1.03} \,(\text{stat.}) \pm 0.43 \,(\text{syst.}) \pm 0.51(B(D_s))] \times 10^{-4}$.

In summary, using a data sample of about 376.9 million $B\overline{B}$ pairs, we find evidence for the decay $B^- \rightarrow D_s^{(*)+} K^- \ell^- \bar{\nu}_\ell$. The signal has a significance larger than 5.0$\sigma$, after taking systematic effects into account. The measured branching fraction, $B(B^- \rightarrow D_s^{(*)+} K^- \ell^- \bar{\nu}_\ell) = [6.13^{+1.04}_{-1.03} \,(\text{stat.}) \pm 0.43 \,(\text{syst.}) \pm 0.51(B(D_s))] \times 10^{-4}$, where the last uncertainty reflects the limited knowledge of the $D_s$ branching fractions, is consistent with the previous upper limit reported by the ARGUS collaboration and with theoretical expectations.

We are grateful for the excellent luminosity and machine conditions provided by our PEP-II colleagues and for the substantial dedicated effort from the computing organizations that support BABAR. The collaborating institutions wish to thank SLAC for its support and kind hospitality. This work is supported by DOE and NSF (USA), NSERC (Canada), CEA and CNRS-IN2P3 (France), BMF and DFG (Germany), INFN (Italy), FOM (The Netherlands), NFR (Norway), MES (Russia), MICIN (Spain), STFC (United Kingdom). Individuals have received support from the Marie Curie EIF (European Union), the A. P. Sloan Foundation (USA) and the Binational Science Foundation (USA-Israel).

* Now at Temple University, Philadelphia, Pennsylvania 19122, USA
† Also with Universit` a di Perugia, Dipartimento di Fisica, Perugia, Italy
‡ Also with Universit` a di Roma La Sapienza, I-00185 Roma, Italy
§ Now at University of South Alabama, Mobile, Alabama 36688, USA
¶ Also with Universit` a di Sassari, Sassari, Italy
[1] C. Amsler et al. (Particle Data Group), Phys. Lett. B667, 1 (2008).
[2] B. Aubert et al. (Babar Collaboration), Phys. Rev. Lett. 100, 151802 (2008).
[3] B. Aubert et al. (Babar Collaboration), Phys. Rev. Lett. 101, 261802 (2008).
[4] D. Liventsev et al. (Belle Collaboration), Phys. Rev. D 77, 091503 (2008).
[5] J. Abdallah et al. (DELPHI Collaboration), Eur. Phys. Jour. C 35, 35 (2006).
[6] N. Uraltsev, Phys. Lett. B 501, 86 (2001).
[7] Throughout this paper, whenever a mode is given, the charge conjugate is implied.
[8] E. Golowich et al., Z. Phys. C 48, 89 (1990).
[9] P. Abreu et al. (DELPHI Collaboration), Phys. Lett. B 289, 199 (1992).
[10] H. Albrecht et al. (ARGUS Collaboration), Z. Phys. C 60, 11 (1993).
[11] B. Aubert et al. (BABAR Collaboration), Nucl. Instrum. Methods Phys. Res., Sect. A 479, 1 (2002).
[12] S. Agostinelli et al. (GEANT4 Collaboration), Nucl. Instr. Methods Phys. Res., Sect. A 506, 250 (2003).
[13] J. L. Goity and W. Roberts, Phys. Rev. D 51, 3459 (1995).
[14] N. Isgur, D. Scora, B. Grinstein, and M. Wise, Phys. Rev. D 39, 799 (1989).
[15] A. Höcker et al., Pos ACAT:040 (2007).
[16] Throughout this paper, all kinematic variables are defined in the center of mass frame.
[17] G. C. Fox and S. Wolfram, Phys. Rev. Lett. 41, 1581 (1978).
\[ L_2 = \sum |\vec{p}_i| \cos^2(\theta_i), \text{ where the sum is over all tracks in} \]
\[ \text{the event not used to form the signal } B \text{ candidate, and} \]
\[ \vec{p}_i \text{ and } \theta_i \text{ are the momenta and angles with respect to the} \]
\[ \text{thrust axis of the signal } B \text{ candidate, respectively.} \]

[19] J. P. Alexander et al. (CLEO Collaboration), Phys. Rev. Lett. 100, 161804 (2008).

[20] B. Aubert et al. (BABAR Collaboration), Phys. Rev. D 67, 032002 (2003).

[21] R. B. Davies, Biometrika 74, 33 (1987).