Search for the rare decay $B\to K\nu\bar{\nu}$
(The BABAR Collaboration)

1Laboratoire d’Annecy-le-Vieux de Physique des Particules (LAPP), Université de Savoie, CNRS/IN2P3, F-74941 Annecy-Le-Vieux, France
2Universitat de Barcelona, Facultat de Fisica, Departament ECM, E-08028 Barcelona, Spain
3INFN Sezione di Bari, I-70126 Bari, Italy
4Dipartimento di Fisica, Università di Bari, I-70126 Bari, Italy
5University of Bergen, Institute of Physics, N-5007 Bergen, Norway
6Lawrence Berkeley National Laboratory and University of California, Berkeley, California 94720, USA
7University of Birmingham, Birmingham, B15 2TT, United Kingdom
8Ruhr Universität Bochum, Institut für Experimentalphysik 1, D-44780 Bochum, Germany
9University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z1
10Brunel University, Uxbridge, Middlesex UB8 3PH, United Kingdom
11Budker Institute of Nuclear Physics, Novosibirsk 630090, Russia
12University of California at Irvine, Irvine, California 92697, USA
13University of California at Riverside, Riverside, California 92521, USA
14University of California at Santa Barbara, Santa Barbara, California 93106, USA
15University of California at Santa Cruz, Institute for Particle Physics, Santa Cruz, California 95064, USA
16California Institute of Technology, Pasadena, California 91125, USA
17University of Cincinnati, Cincinnati, Ohio 45221, USA
18University of Colorado, Boulder, Colorado 80309, USA
19Colorado State University, Fort Collins, Colorado 80523, USA
20Technische Universität Dortmund, Fakultät Physik, D-44221 Dortmund, Germany
21Technische Universität Dresden, Institut für Kern- und Teilchenphysik, D-01062 Dresden, Germany
22Laboratoire Leprince-Ringuet, CNRS/IN2P3, Ecole Polytechnique, F-91128 Palaiseau, France
23University of Edinburgh, Edinburgh EH9 3JZ, United Kingdom
24INFN Sezione di Ferrara, I-44100 Ferrara, Italy
25Dipartimento di Fisica, Università di Ferrara, I-44100 Ferrara, Italy
26INFN Laboratori Nazionali di Frascati, I-00044 Frascati, Italy
27INFN Sezione di Genova, I-16146 Genova, Italy
28Dipartimento di Fisica, Università di Genova, I-16146 Genova, Italy
29Indian Institute of Technology Guwahati, Guwahati, Assam, 781 039, India
30Harvard University, Cambridge, Massachusetts 02138, USA
31Universität Heidelberg, Physikalisches Institut, Philosophenweg 12, D-69120 Heidelberg, Germany
32Humboldt-Universität zu Berlin, Institut für Physik, Newtonstr. 15, D-12489 Berlin, Germany
33Imperial College London, London, SW7 2AZ, United Kingdom
34University of Iowa, Iowa City, Iowa 52242, USA
35Iowa State University, Ames, Iowa 50011-3160, USA
36Johns Hopkins University, Baltimore, Maryland 21218, USA
37Laboratoire de l’Accélérateur Linéaire, IN2P3/CNRS et Université Paris-Sud 11, Centre Scientifique d’Orsay, B. P. 34, F-91898 Orsay Cedex, France
38Lawrence Livermore National Laboratory, Livermore, California 94550, USA
39University of Liverpool, Liverpool L69 7ZE, United Kingdom
40Queen Mary, University of London, London, E1 4NS, United Kingdom
41University of London, Royal Holloway and Bedford New College, Egham, Surrey TW20 0EX, United Kingdom
42University of Louisville, Louisville, Kentucky 40292, USA
43Johannes Gutenberg-Universität Mainz, Institut für Kernphysik, D-55099 Mainz, Germany
44University of Manchester, Manchester M13 9PL, United Kingdom

P. DEL AMO SANCHEZ et al. PHYSICAL REVIEW D 82, 112002 (2010)
We present a search for the rare decays $B^+ \rightarrow K^+ \nu \bar{\nu}$ and $B^0 \rightarrow K^0 \nu \bar{\nu}$ using $459 \times 10^6$ $B \bar{B}$ pairs collected with the BABAR detector at the SLAC National Accelerator Laboratory. Flavor-changing neutral-current decays such as these are forbidden at tree level but can occur through one-loop diagrams in the standard model (SM), with possible contributions from new physics at the same order. The presence of two neutrinos in the final state makes identification of signal events challenging, so reconstruction in the semileptonic decay channels $B \rightarrow D^{(*)} \ell \nu$ of the $B$ meson recoiling from the signal $B$ is used to

---

*Now at Temple University, Philadelphia, PA 19122, USA.
†Also with Università di Perugia, Dipartimento di Fisica, Perugia, Italy.
‡Also with Università di Roma La Sapienza, I-00185 Roma, Italy.
*Now at University of South AL, Mobile, AL 36688, USA.
k Also with Università di Sassari, Sassari, Italy.
The decays $B \to K \nu \bar{\nu}$ arise from flavor-changing neutral currents (FCNC), which are forbidden at tree level in the SM. The lowest-order SM processes contributing to these decays are the $W$ box and the $Z$ penguin diagrams shown in Fig. 1. New physics contributions may enter at the same order as the SM. These contributions, some of which could increase the branching fraction by up to 10 times relative to the SM, include: unparticle models [1], minimal supersymmetric extension of the SM at large $\tan\beta$ [2], models with a single universal extra dimension [3], scalar weakly interacting massive particle (WIMP) dark matter [4] and WIMP-less dark matter [5]. A recent SM prediction (ABSW model [6]) for the total $B \to K \nu \bar{\nu}$ branching fraction is $(4.5 \pm 0.7) \times 10^{-6}$, while an earlier prediction (BHI model [7]), based on a different form factor model, is $(3.8_{-0.5}^{+1.2}) \times 10^{-6}$. The BHI model was used by previous analyses [8,9] and provides a baseline for comparison between results. The current experimental upper limit (UL) on the total branching fraction for $B^+ \to K^+ \nu \bar{\nu}$ (charge conjugation is implied throughout) is $1.4 \times 10^{-5}$ at the 90% CL from the Belle Collaboration [8], while an earlier BABAR analysis set an UL of $5.2 \times 10^{-5}$ (90% CL) [9]. The only existing UL on the total branching fraction for $B^0 \to K^0 \nu \bar{\nu}$ is $1.6 \times 10^{-4}$ (90% CL) from Belle [8].

We report results of a search for $B^+ \to K^+ \nu \bar{\nu}$ and $B^0 \to K^0 \nu \bar{\nu}$, with branching fractions for both decays as well as for the combination $B \to K \nu \bar{\nu}$. We also report on partial branching fractions for $B^+ \to K^+ \nu \bar{\nu}$ in two regions of dineutrino invariant mass squared ($q^2$). The low-$q^2$ region ($q^2 < 0.4 m_B^2$) is selected by requiring $p_{K^+} > 1.5$ GeV/c and the high-$q^2$ region ($q^2 > 0.4 m_B^2$) by $p_{K^+} < 1.5$ GeV/c in the $Y(4S)$-center-of-mass system (CMS) [10], where $m_B$ is the mass of the $B$ meson and $p_{K^+}$ is the magnitude of the CMS 3-momentum of the signal $K^+$ candidate. The high-$q^2$ region is of theoretical interest because the partial branching fraction in this region could be enhanced under some new physics models [6].

![lower-order Feynman diagrams for $B \to K \nu \bar{\nu}$](image)

FIG. 1. Lowest-order Feynman diagrams for $B \to K \nu \bar{\nu}$, with the $W$ box on the left and $Z$ penguin on the right.

This analysis is based on a data sample of $(459.0 \pm 5.1) \times 10^6$ $B\bar{B}$ pairs, corresponding to an integrated luminosity of $\sim 418$ fb$^{-1}$ of $e^+e^-$ colliding-beam data and recorded at the $Y(4S)$ resonance with the BABAR detector [11] at the SLAC PEP-II asymmetric-energy $B$ factory. Charged particle tracking is provided by a five-layer silicon vertex tracker and a 40-layer drift chamber in a 1.5 T magnetic field. A CsI(Tl) electromagnetic calorimeter (EMC) is used to measure photon energies and directions and to identify electrons. All quantities in this paper that are measured by the EMC are required to exceed a minimum 20 MeV cluster energy, unless a higher threshold is explicitly noted. The magnetic flux return from the solenoid, instrumented with resistive plate chambers and limited streamer tubes (IFR), provides muon identification.

We identify $K^+$ candidates by using a detector of internally reflected Cherenkov light (DIRC) as well as ionization energy loss information from the tracking system.

Because of the presence of two neutrinos in the $B \to K \nu \bar{\nu}$ final state, it is not possible to exploit the kinematic constraints on the $B$ mass and energy that are typically used to distinguish signal and background events in $B$ meson decays at the $Y(4S)$. Instead, before looking for the signal decay, we first reconstruct a $B$ decay ($B_{\text{rec}}$) in one of several exclusive $D^{(*)}\ell\nu$ semileptonic final states. We then search for the signal $B \to K \nu \bar{\nu}$ among the remaining charged and neutral particles in the detector that are not part of the $B_{\text{rec}}$. We collectively refer to these remaining particles as $B_{\text{rec}}$ for rest of the event. This strategy is common to several BABAR analyses [12,13] and has the advantage of higher efficiency compared with reconstruction of the $B_{\text{rec}}$ in hadronic decay modes [9].

We reconstruct the $D$ candidates in the following decay modes: $K^{-}\pi^{+}$, $K^{-}\pi^{+}\pi^{+}$, $K^{-}\pi^{+}\pi^{+}\pi^{-}$, $K^{-}\pi^{+}\pi^{0}$, $K^{-}\pi^{+}\pi^{0}$, and $K^{0}_{S}\pi^{+}\pi^{-}$. The $K^{0}_{S}$ candidates, reconstructed in the $K^{0}_{S}\to\pi^{+}\pi^{-}$ mode, are required to have a $\pi^{+}\pi^{-}$ invariant mass within 25 MeV/c$^2$ of the nominal $K^{0}_{S}$ mass. $D$ candidates are similarly required to have a reconstructed invariant mass within 60 MeV/c$^2$ of the nominal value [14], except for the $K^{-}\pi^{+}\pi^{0}$ mode where the range is 100 MeV/c$^2$. We form $D^{*0}\to D^{0}\pi^{0}$, $D^{*+}\to D^{0}\pi^{+}$, and $D^{*+}\to D^{+}\pi^{0}$ candidates with a required mass difference ($m(D^{*}) - m(D)$) in the range 130–170 MeV/c$^2$. In addition, we combine $D$ and $\gamma$ candidates to form $D^{\gamma}$ candidates with a required mass difference in the range 120–170 MeV/c$^2$. A $D^{(\gamma)}$ candidate is combined with an identified electron or muon with momentum above 0.8 GeV/c in the CMS to form a $B_{\text{rec}}$ candidate. In events with multiple reconstructed $B_{\text{rec}}$ candidates, we select the...
candidate with the highest probability that the daughter tracks originate from a common vertex. After a $B_{\text{rec}}$ candidate has been identified, the remaining charged and neutral decay products are used to classify the $B_{\text{rec}}$ as either a background event or a possible signal candidate.

As a first step in refining the selection of $B_{\text{rec}}$ candidates, we veto $K$ candidates, which, when combined with a remaining charged or neutral pion candidate, have a $K\pi$ invariant mass within $75 \text{ MeV}/c^2$ of the nominal $K^+$(892) mass. We also veto events where a remaining charged track can be combined with a $\pi^0$ candidate to yield a $\rho^+$ candidate, with a mass window $0.45 < m(\rho^+) < 1.10 \text{ GeV}/c^2$. Similarly vetoed are events where three remaining charged tracks can be combined to yield an $a_1^+$ candidate, with a mass window $0.6 < m(a_1^+) < 2.0 \text{ GeV}/c^2$. These vetoes eliminate, with little loss of signal efficiency, sizable backgrounds that consist mostly of random track combinations. After the vetoes, $B^+$ ($B^0$) signal candidate events are required to possess $K^+$ ($K^0_S$) signal events that have passed this selection, $99\%$ ($92\%$) of events have a correctly identified signal $K^+$ ($K^0_S$). However, a large background still remains.

Further background suppression is achieved using a multivariate event selection algorithm, a bagged decision tree (BDT) [15,16], that can leverage many weak discriminating variables to achieve high background rejection. Such an algorithm needs to be trained with simulated signal and background events, henceforth referred to as Monte Carlo (MC) events. We use a GEANT4 [17] detector simulation to obtain large samples of simulated signal events generated with a pure phase-space model (which are later rescaled to the BHI signal model), as well as samples of nonresonant $e^+ e^- \rightarrow q\bar{q}$ ($q = u, d, s, c$), $BB$, and $\tau^+ \tau^-$ background events, whose sizes are one ($uds$), two ($c\bar{c}$), three ($BB$), and one ($\tau^+ \tau^-$) times luminosity. These background events are augmented with a separate sample, with a size 13 times luminosity, of simulated $BB$ doubly semileptonic events, the largest source of background.

We construct two ensembles of BDTs, one for the $K^+$ signal mode and one for the $K^0_S$. To create an ensemble, we repeatedly divide the total signal and background datasets in half randomly, creating 20 distinct BDT training and validation datasets, where each dataset has a 50% correlation with any other because approximately 50% of the events are shared. This procedure makes optimal use of the limited statistics of MC events that pass the initial event selection and results in a more statistically precise unbiased estimate of background contributions. Use of the ensemble of 20 BDTs created for each final state also averages out the variations in BDT response compared to a single BDT trained and validated with a single division of the simulated signal and background datasets [18,19]. The choice of 20 divisions, instead of a lower or higher number, represents a balance between minimizing the variation versus minimizing the overhead of multiple BDTs.

Each BDT of the $K^+$ ($K^0_S$) ensemble uses 26 (38) discriminating variables, described in the Appendix. These variables fall into four general categories: quantities related to the missing energy in the event, to the overall event properties, to the signal kinematics, and to the overall reconstruction quality of the $B_{\text{rec}}$.

Some quantities are given in two different frames and thus allow the BDTs to extract from them additional discriminating power. Several additional variables were initially considered but were pruned during the BDT optimization process because they were found to add little additional sensitivity.

“Missing Energy” quantities relate to the fact that signal events are expected to possess significant missing energy and momentum because the signal decay includes two neutrinos. In contrast, the dominant background events usually acquire missing energy and momentum as a result of particles passing outside of the detector fiducial acceptance, with the result that distributions of quantities related to missing energy differ between signal and background.

After the $B_{\text{rec}}$ and $K$ or $K^0$ signal candidate have been identified, signal events are expected to have little or no additional activity in the detector, other than a few low-energy clusters in the calorimeter resulting from hadronic shower remnants, beam backgrounds, or similar sources. In contrast, background events arising from higher-multiplicity $B$ decays typically possess additional charged or neutral particles within the detector. Variables which characterize this additional detector activity can provide discriminating power between signal and background, and are indicated by the term “extra” in the following.

The strongest discriminant for both $K^+$ and $K^0$ ensembles is $E_{\text{extra}}$, the sum of all detector activity not explicitly associated with either the $B_{\text{rec}}$ or $K$ signal candidate, followed by $p_T^{K^+}$ for the $K^+$ ensemble and by the lab energy of the signal $K^0_S$ for the $K^0$ ensemble. The reconstructed mass of the $D$ from the $B_{\text{rec}}$ is the third ranking variable for both channels.

Figure 2 shows signal, background, and data distributions from the validation set of $K^+$ and $K^0$ BDT output for a BDT randomly selected from the 20 BDT in the ensemble. The other 19 BDTs are similar to that shown.

We choose as the target signal efficiency the one that maximizes expected signal significance averaged over the
20 BDTs, under the assumption of a branching fraction of $3.8 \times 10^{-6}$. This signal significance is $s/\sqrt{s + b}$, where $s$ is the number of signal events, and $b$ is the number of background events. Optimization using a figure of merit based upon signal efficiency and independent of assumed branching fraction yields similar results. For each BDT, a BDT output value that yields the target signal efficiency is calculated. For example, the BDT output cuts for the BDTs shown in Fig. 2 are 0.976 for the $K^+$ BDT and 0.955 for the $K^0$ BDT. The mean background for target signal efficiency is obtained by averaging the individual background estimates from each of the 20 BDTs. Thus, we treat each ensemble of 20 BDTs as a set of correlated estimators for the numbers of signal and background events in a signal region defined by the target signal efficiency.

The low-$q^2$ (high-$q^2$) measurement uses the $K^+$ ensemble but only includes events with $p_K^* > 1.5$ GeV/c ($p_K^* < 1.5$ GeV/c), which means that only those events are used to calculate the signal efficiency and the background prediction. The low-$q^2$ measurement has the same BDT output cuts and background prediction as the primary $K^+$ measurement, with only the signal efficiency changed by the restriction on $p_K^*$. On the other hand, the high-$q^2$ measurement has its own set of BDT output cuts based upon its own optimized signal efficiency, along with its own background prediction.

The total optimized signal efficiency for the $K^+$ ($K^0$) mode is 0.16% (0.06%), while the efficiency for the $K^+$ low-$q^2$ (high-$q^2$) region is 0.24% (0.28%). The uncertainty in the signal efficiency is discussed below. Figure 3 shows the BDT selection efficiency versus $p_K^*$ for the $K^+$, $K^0$, and high-$q^2$ measurements, where the BDT selection efficiency considers only the effect of the BDT output cut.

To measure the branching fractions, we use the value obtained from simulated events of the predicted background in the signal region, the number of observed data events, and the signal efficiency, as shown by the following equation: $B = (N_{\text{obs}} - N_{\text{bkg}}) / \epsilon N_B$, where $B$ is the branching fraction, $N_{\text{obs}}$ is the number of observed data events, $N_{\text{bkg}}$ is the number of predicted background events, $\epsilon$ is the total signal efficiency, and $N_B$ is the number of $B$ mesons, either charged or neutral [20], that are relevant to the branching fraction. We account for the 50% correlation between each of the datasets when computing the statistical uncertainty of the estimated background contribution by using a standard method for combining correlated uncertainties [19].

Data control samples are used to ensure that both signal-like and background-like events in actual data are classified similarly to simulated events. The vetoed $a_1^+$ events offer a high-statistics control sample, which can be used to compare the $K^+$ and $K^0$ BDT output distributions.
for background events in both simulated and actual data. We find good agreement between data and MC events in the BDT output distribution for both final states, with only a \((+5 \pm 2\%)\) data-MC discrepancy. For the \(K^+\) mode we make a \(+5\%\) adjustment to the expected number of background events, based upon a weighting technique that corrects data-MC discrepancy in the sideband \(K^+\) BDT output next to the signal region, and we assign the full adjustment as a systematic uncertainty. Likewise, for the high-\(q^2\) \(K^+\) measurement, we make a \(+25\%\) correction to the expected number of background events and assign the full correction as a systematic uncertainty. In the \(K_S^0\) final state, we find a \((+10 \pm 3\%)\) data-MC discrepancy in the sideband BDT output next to the signal region, and we make a \(+10\%\) correction and assign the full correction as a systematic uncertainty.

To validate our signal efficiency estimates and assess their systematic uncertainties, we use high-purity samples of \(B^+ \to K^+ J/\psi (\to \ell^+ \ell^-)\) decays (where \(\ell^+ \ell^- = e^+ e^-\), \(\mu^+ \mu^-\)). The two leptons from the \(J/\psi\) are discarded in order to model the unseen neutrinos of the signal decay, and then the events are subjected to the same selection requirements as other signal candidates. Classifying \(J/\psi K\) data and MC events, we find only a \((-10 \pm 10\%)\) data-MC discrepancy in the BDT output distribution. Although we do not make any correction, we assign a \(10\%\) systematic uncertainty to the estimated signal efficiency for all four measurements (\(K^+, K_S^0\), low-\(q^2\) \(K^+\), high-\(q^2\) \(K^+\)) based on these results. We also assign a signal efficiency systematic uncertainty of \(10\%\) to account for the theoretical uncertainties of the signal models. Adding these in quadrature, we assign a total uncertainty of \(14\%\) in the estimation of signal efficiency for both final states. Table I summarizes all of the systematic uncertainties.

### Table I. Systematic uncertainties

<table>
<thead>
<tr>
<th>Category</th>
<th>Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Signal efficiency</td>
<td>14%</td>
</tr>
<tr>
<td>(K^+) background prediction</td>
<td>5%</td>
</tr>
<tr>
<td>High-(q^2) (K^+) background prediction</td>
<td>25%</td>
</tr>
<tr>
<td>(K_S^0) background prediction</td>
<td>10%</td>
</tr>
</tbody>
</table>

### Table II. Total signal efficiencies and MC expectations of the number of data events. The uncertainties shown are systematic for \(N_{sngl}\), with statistical negligible, and statistical followed by systematic for \(N_{bgld}\).

<table>
<thead>
<tr>
<th>Mode</th>
<th>(\epsilon) (in %)</th>
<th>(N_{sngl})</th>
<th>(N_{bgld})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(K^+)</td>
<td>0.16</td>
<td>2.9 (\pm) 0.4</td>
<td>17.6 (\pm) 2.6 (\pm) 0.9</td>
</tr>
<tr>
<td>(K_S^0)</td>
<td>0.06</td>
<td>0.5 (\pm) 0.1</td>
<td>3.9 (\pm) 1.3 (\pm) 0.4</td>
</tr>
<tr>
<td>low-(q^2) (K^+)</td>
<td>0.24</td>
<td>2.9 (\pm) 0.4</td>
<td>17.6 (\pm) 2.6 (\pm) 0.9</td>
</tr>
<tr>
<td>high-(q^2) (K^+)</td>
<td>0.28</td>
<td>2.1 (\pm) 0.3</td>
<td>187 (\pm) 10 (\pm) 46</td>
</tr>
</tbody>
</table>

Table II shows the total signal efficiencies and the expected number of signal and background events in the data. We performed a blind analysis where data events with BDT outputs above the optimized values were not counted or plotted until the analysis methodology and sources of systematic uncertainty were fixed as described above.

Table III shows our results. The noninteger number of observed events results from averaging the integer yields from the 20 BDTs of each type. We calculate two-sided 68\% confidence intervals for the number of excess events based on the statistical and systematic uncertainties in the background estimates and the statistical uncertainty on the number of events observed in the data. Figure 4 shows the averaged BDT outputs in the signal region for \(K^+, K^0\), and high-\(q^2\) \(K^+\) data overlaid with the background and signal contributions, while Fig. 5 shows similar plots for the \(p_T\) distribution in the signal region. Figure 6 shows the integrated numbers of events (observed, predicted background, and excess over background) in the signal region for \(K^+\), \(K^0\), and high-\(q^2\) \(K^+\) data for each of the 20 BDTs of each type. Table IV gives the branching fraction central values, along with corresponding 90\% and 95\% CL upper limits, assuming the BHI signal model (the ABSW model gives similar results). The upper limits are calculated using a frequentist method [21]. The quoted uncertainties include all statistical and systematic uncertainties. Our results constrain the \(B \to K\nu\nu\) branching fraction at the 90\% CL to a few times the SM expectation, with limits of \(1.3 \times 10^{-5}\) for \(B^+ \to K^+ \nu\nu\) and \(5.6 \times 10^{-5}\) for \(B^0 \to K^0 \nu\nu\).

We are grateful for the extraordinary contributions of our PEP-II colleagues in achieving the excellent luminosity and machine conditions that have made this work possible. The success of this project also relies critically on the expertise and dedication of the computing organizations that support BABAR. The collaborating institutions wish to thank SLAC for its support and the kind hospitality extended to them. This work is supported by the U.S. Department of Energy and National Science Foundation, the Natural Sciences and Engineering Research Council (Canada), the Commissariat à l’Energie Atomique and Institut National de Physique.

### Table III. Observed and excess data events, with statistical uncertainties [21] shown for \(N_{obs}\) and combined statistical and systematic uncertainties shown for \(N_{excess}\). The last column shows the probability that excess events could be due solely to a background fluctuation.

<table>
<thead>
<tr>
<th>Mode</th>
<th>(N_{obs})</th>
<th>(N_{excess})</th>
<th>Prob.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(K^+)</td>
<td>19.4 (\pm) 4</td>
<td>1.8 (\pm) 0.2</td>
<td>38%</td>
</tr>
<tr>
<td>(K^0)</td>
<td>6.1 (\pm) 1</td>
<td>2.7 (\pm) 1.1</td>
<td>23%</td>
</tr>
<tr>
<td>low-(q^2) (K^+)</td>
<td>19.4 (\pm) 4</td>
<td>1.8 (\pm) 0.2</td>
<td>38%</td>
</tr>
<tr>
<td>high-(q^2) (K^+)</td>
<td>164 (\pm) 13</td>
<td>(-23 \pm 49)</td>
<td>33%</td>
</tr>
</tbody>
</table>
FIG. 4 (color online). Averaged BDT signal-region output for (a) $K^+$, (b) $K_S^0$, and (c) high-$q^2$ $K^+$ data, with expected signal and background contributions. The signal estimate assumes a branching fraction of $3.8 \times 10^{-6}$.

FIG. 5 (color online). Averaged $p_K^*$ signal-region output for (a) $K^+$, (b) $K_S^0$, and (c) high-$q^2$ $K^+$ data, with expected signal and background contributions. The signal estimate assumes a branching fraction of $3.8 \times 10^{-6}$.

FIG. 6 (color online). Integrated numbers of observed (red triangles), expected background (black circles), and excess events (blue squares) for data for each BDT: (a) $K^+$, (b) $K_S^0$, and (c) high-$q^2$ $K^+$. The individual uncertainties are purely statistical and assume no correlation between data sets. The horizontal dashed lines show the sum of the statistical and systematic uncertainties on the mean number of excess events.

TABLE IV. Branching fraction (BF) central values and upper limits. The low- and high-$q^2$ values are partial BFs, while the rest are total BFs.

<table>
<thead>
<tr>
<th>Mode</th>
<th>BF</th>
<th>90% CL</th>
<th>95% CL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\times 10^{-5}$</td>
<td>$\times 10^{-5}$</td>
<td>$\times 10^{-5}$</td>
</tr>
<tr>
<td>$K^+$</td>
<td>$0.2^{+0.8}_{-0.7}$</td>
<td>1.3</td>
<td>1.6</td>
</tr>
<tr>
<td>$K^0$</td>
<td>$1.7^{+3.1}_{-2.1}$</td>
<td>5.6</td>
<td>6.7</td>
</tr>
<tr>
<td>Comb. $K^+$, $K^0$</td>
<td>$0.5^{+0.7}_{-0.6}$</td>
<td>1.4</td>
<td>1.7</td>
</tr>
<tr>
<td>Low-$q^2$ $K^+$</td>
<td>$0.2^{+0.6}_{-0.5}$</td>
<td>0.9</td>
<td>1.1</td>
</tr>
<tr>
<td>High-$q^2$ $K^+$</td>
<td>$1.8^{+3.8}_{-3.8}$</td>
<td>3.1</td>
<td>4.6</td>
</tr>
</tbody>
</table>

Nucéaire et de Physique des Particules (France), the Bundesministerium für Bildung und Forschung and Deutsche Forschungsgemeinschaft (Germany), the Istituto Nazionale di Fisica Nucleare (Italy), the Foundation for Fundamental Research on Matter (The Netherlands), the Research Council of Norway, the Ministry of Education and Science of the Russian Federation, Ministerio de Ciencia e Innovación (Spain), and the Science and Technology Facilities Council (United Kingdom). Individuals have received support from the Marie-Curie IEF program (European Union), the A. P. Sloan Foundation (USA) and the Binational Science Foundation (USA-Israel).
APPENDIX: DEFINITIONS OF BDT VARIABLES

In the following the notation \([K^+]\) or \([K^0]\) indicates that a variable is used only by that ensemble; otherwise it is used by both BDT ensembles.

**BDT input variables related to missing 4-momentum**

The event missing 4-momentum is computed from the difference between the 4-momentum of the combined \(e^+e^-\) beams and the 4-momenta of all charged and neutral particles reconstructed in the detector.

(i) Energy component of missing momentum 4-vector
(ii) Energy component of missing momentum 4-vector (CMS)
(iii) Magnitude of the missing momentum 3-vector
(iv) Magnitude of the missing momentum 3-vector (CMS)
(v) Cosine of the angle with respect to the beam axis of the missing momentum 3-vector
(vi) Cosine of the angle with respect to the beam axis of the missing momentum 4-vector
(vii) Uncertainty in the \(x\)-component of the signal \(K^+\) point of closest approach to the \(e^+e^-\) interaction point, as determined from a three dimensional fit, with the \(x\)-axis defined perpendicular to the beam axis in the horizontal plane of the detector \([K^+]\)
(viii) Magnitude of the CMS energy, invariant mass and 3-momentum magnitude of the \(D\)—lepton combination used in the reconstruction of the \(B_{rec}\)
(ix) Reconstructed decay mode of the \(D\) from the \(B_{rec}\)
(x) Reconstructed invariant mass of the \(D\) candidate such that \(100 < (m(D^0, \gamma) - m(D^0)) < 150\) MeV/c²
(xi) Angle with respect to the beam axis of a candidate \(Y(4S)\) 3-momentum vector computed from the \(B_{rec}\) and \(B_{rec}\) 4-momenta \([K^0]\)
(xii) Normalized second Fox-Wolfram moment of the overall event

**BDT input variables related to overall event properties**

(i) \(E_{extra} = \sum E_i\), where \(E_i\) is the energy of an isolated EMC cluster or a charged track and the sum is over all tracks or clusters which are not part of the \(B_{rec}\) or the \(B_{rec}\)
(ii) Total energy of all reconstructed charged and neutral particles in the event
(iii) Minimum invariant mass obtained from the combination of any three charged tracks in the event
(iv) Total charge of all tracks in the event \([K^0]\)
(v) Total charge of all tracks matched to EMC energy deposits \([K^0]\)
(vi) Number of extra EMC clusters
(vii) Number of \(K_L\) candidates in the EMC
(viii) Number of IFR \(K_L\) candidates \([K^+]\)
(ix) Number of extra reconstructed tracks
(x) Magnitude of the 3-momentum of a candidate \(Y(4S)\) computed from the \(B_{rec}\) and \(B_{rec}\) 4-momenta \([K^0]\)
(xi) Angle with respect to the beam axis of a candidate \(Y(4S)\) 3-momentum vector computed from the \(B_{rec}\) and \(B_{rec}\) 4-momenta \([K^0]\)
(xii) Normalized second Fox-Wolfram moment of the overall event

**BDT input variables related to signal kinematics**

(i) Cosine of the angle between the signal \(K\) and the event thrust axis
(ii) Cosine of the angle between the signal \(K\) and the \(Dl\) thrust axis
(iii) Energy of the signal kaon \([K^0]\)


[10] All kinematic quantities in this paper are defined in the lab frame unless marked with an asterisk, in which case they are in the CMS.


[20] Equal branching fractions for $Y(4S) \rightarrow B^0 \bar{B}^0$ and $Y(4S) \rightarrow B^+ B^-$ are assumed in this paper.