

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Chilikin, K;Adachi, I;Asner, DM;Aulchenko, V;Aushev, T;Ayad, R;Babu, V;Badhrees, I;Bansal, V;Behera, P;Beleño, C;Berger, M;Bhardwaj, V;Biłka, T;Biswal, J;Bobrov, A;Bondar, A;Bozek, A;Bračko, M;Browder, TE;Campajola, M;Cao, L;Červenkov, D;Chekelian, V;Chen, A;Cheon, BG;Cho, K;Choi, SK;Choi, Y;Cinabro, D;Cunliffe, S;Di Carlo, S;Doležal, Z;Dong, TV;Eidelman, S;Epifanov, D;Fast, JE;Ferber, T;Fulsom, BG;Garg, R;Gaur, V;Gabyshev, N;Garmash, A;Gelb, M;Giri, A;Goldenzweig, P;Grzymkowska, O;Haba, J;Hara, T;Hayasaka, K;Hayashii, H;Hou, WS;Hsu, CL;Inami, K;Ishikawa, A;Itoh, R;Iwasaki, M;Iwasaki, Y;Jacobs, WW;Jia, S;Jin, Y;Joffe, D;Joo, KK;Julius, T;Kaliyar, AB;Karyan, G;Kato, Y;Kiesling, C;Kim, CH;Kim, DY;Kim, SH;Kinoshita, K;Kodyš, P;Korpar, S;Kotchetkov, D;Kroeger, R;Krokovny, P;Kulasiri, R;Kumar, R;Kwon, YJ;Lalwani, K;Lange, JS;Lee, JK;Lee, JY;Lee, SC;Li, LK;Li, YB;Li Gioi, L;Libby, J;Liventsev, D;Lu, PC;Luo, T;Macnaughton, J;Macqueen, C;Masuda, M;Matsuda, T;Matvienko, D;Merola, M;Miyabayashi, K;Mizuk, R

Title:

Evidence for  $B^+ \rightarrow hcK^+$  and observation of  $\eta_c(2S) \rightarrow pp^- \pi^+ \pi^-$

Date:

2019-07-10

Citation:

Chilikin, K., Adachi, I., Asner, D. M., Aulchenko, V., Aushev, T., Ayad, R., Babu, V., Badhrees, I., Bansal, V., Behera, P., Beleño, C., Berger, M., Bhardwaj, V., Biłka, T., Biswal, J., Bobrov, A., Bondar, A., Bozek, A., Bračko, M., ... Mizuk, R. (2019). Evidence for  $B^+ \rightarrow hcK^+$  and observation of  $\eta_c(2S) \rightarrow pp^- \pi^+ \pi^-$ . *Physical Review D*, 100 (1), <https://doi.org/10.1103/PhysRevD.100.012001>.

Persistent Link:

<https://hdl.handle.net/11343/290341>

License:

CC BY

**Evidence for  $B^+ \rightarrow h_c K^+$  and observation of  $\eta_c(2S) \rightarrow p\bar{p}\pi^+\pi^-$** 

K. Chilikin,<sup>41</sup> I. Adachi,<sup>17,13</sup> D. M. Asner,<sup>3</sup> V. Aulchenko,<sup>4,62</sup> T. Aushev,<sup>52</sup> R. Ayad,<sup>77</sup> V. Babu,<sup>78</sup> I. Badhrees,<sup>77,34</sup> V. Bansal,<sup>64</sup> P. Behera,<sup>23</sup> C. Beleño,<sup>12</sup> M. Berger,<sup>74</sup> V. Bhardwaj,<sup>20</sup> T. Bilka,<sup>5</sup> J. Biswal,<sup>31</sup> A. Bobrov,<sup>4,62</sup> A. Bondar,<sup>4,62</sup> A. Bozek,<sup>59</sup> M. Bračko,<sup>46,31</sup> T. E. Browder,<sup>16</sup> M. Campajola,<sup>28,54</sup> L. Cao,<sup>32</sup> D. Červenkov,<sup>5</sup> V. Chekelian,<sup>47</sup> A. Chen,<sup>56</sup> B. G. Cheon,<sup>15</sup> K. Cho,<sup>35</sup> S.-K. Choi,<sup>14</sup> Y. Choi,<sup>75</sup> D. Cinabro,<sup>86</sup> S. Cunliffe,<sup>8</sup> S. Di Carlo,<sup>39</sup> Z. Doležal,<sup>5</sup> T. V. Dong,<sup>17,13</sup> S. Eidelman,<sup>4,62,41</sup> D. Epifanov,<sup>4,62</sup> J. E. Fast,<sup>64</sup> T. Ferber,<sup>8</sup> B. G. Fulsom,<sup>64</sup> R. Garg,<sup>65</sup> V. Gaur,<sup>85</sup> N. Gabyshev,<sup>4,62</sup> A. Garmash,<sup>4,62</sup> M. Gelb,<sup>32</sup> A. Giri,<sup>22</sup> P. Goldenzweig,<sup>32</sup> O. Grzymkowska,<sup>59</sup> J. Haba,<sup>17,13</sup> T. Hara,<sup>17,13</sup> K. Hayasaka,<sup>61</sup> H. Hayashii,<sup>55</sup> W.-S. Hou,<sup>58</sup> C.-L. Hsu,<sup>76</sup> K. Inami,<sup>53</sup> A. Ishikawa,<sup>80</sup> R. Itoh,<sup>17,13</sup> M. Iwasaki,<sup>63</sup> Y. Iwasaki,<sup>17</sup> W. W. Jacobs,<sup>24</sup> S. Jia,<sup>2</sup> Y. Jin,<sup>82</sup> D. Joffe,<sup>33</sup> K. K. Joo,<sup>6</sup> T. Julius,<sup>48</sup> A. B. Kaliyar,<sup>23</sup> G. Karyan,<sup>8</sup> Y. Kato,<sup>53</sup> C. Kiesling,<sup>47</sup> C. H. Kim,<sup>15</sup> D. Y. Kim,<sup>73</sup> S. H. Kim,<sup>15</sup> K. Kinoshita,<sup>7</sup> P. Kodyš,<sup>5</sup> S. Korpar,<sup>46,31</sup> D. Kotchetkov,<sup>16</sup> R. Kroeger,<sup>49</sup> P. Krokovny,<sup>4,62</sup> R. Kulasiri,<sup>33</sup> R. Kumar,<sup>68</sup> Y.-J. Kwon,<sup>88</sup> K. Lalwani,<sup>44</sup> J. S. Lange,<sup>11</sup> J. K. Lee,<sup>71</sup> J. Y. Lee,<sup>71</sup> S. C. Lee,<sup>38</sup> L. K. Li,<sup>25</sup> Y. B. Li,<sup>66</sup> L. Li Gioi,<sup>47</sup> J. Libby,<sup>23</sup> D. Liventsev,<sup>85,17</sup> P.-C. Lu,<sup>58</sup> T. Luo,<sup>10</sup> J. MacNaughton,<sup>50</sup> C. MacQueen,<sup>48</sup> M. Masuda,<sup>81</sup> T. Matsuda,<sup>50</sup> D. Matvienko,<sup>4,62,41</sup> M. Merola,<sup>28,54</sup> K. Miyabayashi,<sup>55</sup> R. Mizuk,<sup>41,51,52</sup> G. B. Mohanty,<sup>78</sup> T. Mori,<sup>53</sup> M. Nakao,<sup>17,13</sup> K. J. Nath,<sup>21</sup> M. Nayak,<sup>86,17</sup> M. Niiyama,<sup>37</sup> N. K. Nisar,<sup>67</sup> S. Nishida,<sup>17,13</sup> K. Nishimura,<sup>16</sup> H. Ono,<sup>60,61</sup> Y. Onuki,<sup>82</sup> P. Pakhlov,<sup>41,51</sup> G. Pakhlova,<sup>41,52</sup> B. Pal,<sup>3</sup> S. Pardi,<sup>28</sup> H. Park,<sup>38</sup> S. Patra,<sup>20</sup> S. Paul,<sup>79</sup> T. K. Pedlar,<sup>43</sup> R. Pestotnik,<sup>31</sup> L. E. Piilonen,<sup>85</sup> V. Popov,<sup>41,52</sup> M. Ritter,<sup>42</sup> A. Rostomyan,<sup>8</sup> G. Russo,<sup>28</sup> Y. Sakai,<sup>17,13</sup> M. Salehi,<sup>45,42</sup> S. Sandilya,<sup>7</sup> T. Sanuki,<sup>80</sup> V. Savinov,<sup>67</sup> O. Schneider,<sup>40</sup> G. Schnell,<sup>1,19</sup> C. Schwanda,<sup>26</sup> Y. Seino,<sup>61</sup> K. Senyo,<sup>87</sup> M. E. Sevier,<sup>48</sup> C. P. Shen,<sup>2</sup> J.-G. Shiu,<sup>58</sup> B. Shwartz,<sup>4,62</sup> F. Simon,<sup>47</sup> A. Sokolov,<sup>27</sup> E. Solovieva,<sup>41</sup> M. Starič,<sup>31</sup> Z. S. Stottler,<sup>85</sup> J. F. Strube,<sup>64</sup> T. Sumiyoshi,<sup>84</sup> W. Sutcliffe,<sup>32</sup> M. Takizawa,<sup>72,18,69</sup> U. Tamponi,<sup>29</sup> K. Tanida,<sup>30</sup> F. Tenchini,<sup>8</sup> K. Trabelsi,<sup>39</sup> M. Uchida,<sup>83</sup> S. Uno,<sup>17,13</sup> P. Urquijo,<sup>48</sup> Y. Usov,<sup>4,62</sup> R. Van Tonder,<sup>32</sup> G. Varner,<sup>16</sup> A. Vinokurova,<sup>4,62</sup> B. Wang,<sup>47</sup> C. H. Wang,<sup>57</sup> M.-Z. Wang,<sup>58</sup> P. Wang,<sup>25</sup> M. Watanabe,<sup>61</sup> S. Watanuki,<sup>80</sup> E. Won,<sup>36</sup> S. B. Yang,<sup>36</sup> H. Ye,<sup>8</sup> J. Yelton,<sup>9</sup> J. H. Yin,<sup>25</sup> C. Z. Yuan,<sup>25</sup> J. Zhang,<sup>25</sup> Z. P. Zhang,<sup>70</sup> V. Zhilich,<sup>4,62</sup> and V. Zhukova<sup>41</sup>

(Belle Collaboration)

<sup>1</sup>University of the Basque Country UPV/EHU, 48080 Bilbao<sup>2</sup>Beihang University, Beijing 100191<sup>3</sup>Brookhaven National Laboratory, Upton, New York 11973<sup>4</sup>Budker Institute of Nuclear Physics SB RAS, Novosibirsk 630090<sup>5</sup>Faculty of Mathematics and Physics, Charles University, 121 16 Prague<sup>6</sup>Chonnam National University, Kwangju 660-701<sup>7</sup>University of Cincinnati, Cincinnati, Ohio 45221<sup>8</sup>Deutsches Elektronen-Synchrotron, 22607 Hamburg<sup>9</sup>University of Florida, Gainesville, Florida 32611<sup>10</sup>Key Laboratory of Nuclear Physics and Ion-beam Application (MOE) and Institute of Modern Physics, Fudan University, Shanghai 200443<sup>11</sup>Justus-Liebig-Universität Gießen, 35392 Gießen<sup>12</sup>II. Physikalisches Institut, Georg-August-Universität Göttingen, 37073 Göttingen<sup>13</sup>SOKENDAI (The Graduate University for Advanced Studies), Hayama 240-0193<sup>14</sup>Gyeongsang National University, Chinju 660-701<sup>15</sup>Hanyang University, Seoul 133-791<sup>16</sup>University of Hawaii, Honolulu, Hawaii 96822<sup>17</sup>High Energy Accelerator Research Organization (KEK), Tsukuba 305-0801<sup>18</sup>J-PARC Branch, KEK Theory Center, High Energy Accelerator Research Organization (KEK), Tsukuba 305-0801<sup>19</sup>IKERBASQUE, Basque Foundation for Science, 48013 Bilbao<sup>20</sup>Indian Institute of Science Education and Research Mohali, SAS Nagar 140306<sup>21</sup>Indian Institute of Technology Guwahati, Assam 781039<sup>22</sup>Indian Institute of Technology Hyderabad, Telangana 502285<sup>23</sup>Indian Institute of Technology Madras, Chennai 600036<sup>24</sup>Indiana University, Bloomington, Indiana 47408<sup>25</sup>Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049

- <sup>26</sup>*Institute of High Energy Physics, Vienna 1050*
- <sup>27</sup>*Institute for High Energy Physics, Protvino 142281*
- <sup>28</sup>*INFN—Sezione di Napoli, 80126 Napoli*
- <sup>29</sup>*INFN—Sezione di Torino, 10125 Torino*
- <sup>30</sup>*Advanced Science Research Center, Japan Atomic Energy Agency, Naka 319-1195*
- <sup>31</sup>*J. Stefan Institute, 1000 Ljubljana, Slovenia*
- <sup>32</sup>*Institut für Experimentelle Teilchenphysik, Karlsruher Institut für Technologie, 76131 Karlsruhe*
- <sup>33</sup>*Kennesaw State University, Kennesaw, Georgia 30144*
- <sup>34</sup>*King Abdulaziz City for Science and Technology, Riyadh 11442, Saudi Arabia*
- <sup>35</sup>*Korea Institute of Science and Technology Information, Daejeon 305-806*
- <sup>36</sup>*Korea University, Seoul 136-713*
- <sup>37</sup>*Kyoto University, Kyoto 606-8502*
- <sup>38</sup>*Kyungpook National University, Daegu 702-701*
- <sup>39</sup>*LAL, Univ. Paris-Sud, CNRS/IN2P3, Université Paris-Saclay, Orsay*
- <sup>40</sup>*École Polytechnique Fédérale de Lausanne (EPFL), Lausanne 1015*
- <sup>41</sup>*P.N. Lebedev Physical Institute of the Russian Academy of Sciences, Moscow 119991*
- <sup>42</sup>*Ludwig Maximilians University, 80539 Munich*
- <sup>43</sup>*Luther College, Decorah, Iowa 52101*
- <sup>44</sup>*Malaviya National Institute of Technology Jaipur, Jaipur 302017*
- <sup>45</sup>*University of Malaya, 50603 Kuala Lumpur*
- <sup>46</sup>*University of Maribor, 2000 Maribor*
- <sup>47</sup>*Max-Planck-Institut für Physik, 80805 München*
- <sup>48</sup>*School of Physics, University of Melbourne, Victoria 3010*
- <sup>49</sup>*University of Mississippi, University, Mississippi 38677*
- <sup>50</sup>*University of Miyazaki, Miyazaki 889-2192*
- <sup>51</sup>*Moscow Physical Engineering Institute, Moscow 115409*
- <sup>52</sup>*Moscow Institute of Physics and Technology, Moscow Region 141700*
- <sup>53</sup>*Graduate School of Science, Nagoya University, Nagoya 464-8602*
- <sup>54</sup>*Università di Napoli Federico II, 80055 Napoli*
- <sup>55</sup>*Nara Women's University, Nara 630-8506*
- <sup>56</sup>*National Central University, Chung-li 32054*
- <sup>57</sup>*National United University, Miao Li 36003*
- <sup>58</sup>*Department of Physics, National Taiwan University, Taipei 10617*
- <sup>59</sup>*H. Niewodniczanski Institute of Nuclear Physics, Krakow 31-342*
- <sup>60</sup>*Nippon Dental University, Niigata 951-8580*
- <sup>61</sup>*Niigata University, Niigata 950-2181*
- <sup>62</sup>*Novosibirsk State University, Novosibirsk 630090*
- <sup>63</sup>*Osaka City University, Osaka 558-8585*
- <sup>64</sup>*Pacific Northwest National Laboratory, Richland, Washington 99352*
- <sup>65</sup>*Panjab University, Chandigarh 160014*
- <sup>66</sup>*Peking University, Beijing 100871*
- <sup>67</sup>*University of Pittsburgh, Pittsburgh, Pennsylvania 15260*
- <sup>68</sup>*Punjab Agricultural University, Ludhiana 141004*
- <sup>69</sup>*Theoretical Research Division, Nishina Center, RIKEN, Saitama 351-0198*
- <sup>70</sup>*University of Science and Technology of China, Hefei 230026*
- <sup>71</sup>*Seoul National University, Seoul 151-742*
- <sup>72</sup>*Showa Pharmaceutical University, Tokyo 194-8543*
- <sup>73</sup>*Soongsil University, Seoul 156-743*
- <sup>74</sup>*Stefan Meyer Institute for Subatomic Physics, Vienna 1090*
- <sup>75</sup>*Sungkyunkwan University, Suwon 440-746*
- <sup>76</sup>*School of Physics, University of Sydney, New South Wales 2006*
- <sup>77</sup>*Department of Physics, Faculty of Science, University of Tabuk, Tabuk 71451*
- <sup>78</sup>*Tata Institute of Fundamental Research, Mumbai 400005*
- <sup>79</sup>*Department of Physics, Technische Universität München, 85748 Garching*
- <sup>80</sup>*Department of Physics, Tohoku University, Sendai 980-8578*
- <sup>81</sup>*Earthquake Research Institute, University of Tokyo, Tokyo 113-0032*
- <sup>82</sup>*Department of Physics, University of Tokyo, Tokyo 113-0033*
- <sup>83</sup>*Tokyo Institute of Technology, Tokyo 152-8550*
- <sup>84</sup>*Tokyo Metropolitan University, Tokyo 192-0397*
- <sup>85</sup>*Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061*

<sup>86</sup>Wayne State University, Detroit, Michigan 48202  
<sup>87</sup>Yamagata University, Yamagata 990-8560  
<sup>88</sup>Yonsei University, Seoul 120-749



(Received 15 March 2019; published 10 July 2019)

A search for the decays  $B^+ \rightarrow h_c K^+$  and  $B^0 \rightarrow h_c K_S^0$  is performed. Evidence for the decay  $B^+ \rightarrow h_c K^+$  is found; its significance is  $4.8\sigma$ . No evidence is found for  $B^0 \rightarrow h_c K_S^0$ . The branching fraction for  $B^+ \rightarrow h_c K^+$  is measured to be  $(3.7^{+1.0}_{-0.9} {}^{+0.8}_{-0.8}) \times 10^{-5}$ ; the upper limit for the  $B^0 \rightarrow h_c K_S^0$  branching fraction is  $1.4 \times 10^{-5}$  at 90% C.L. In addition, a study of the  $p\bar{p}\pi^+\pi^-$  invariant mass distribution in the channel  $B^+ \rightarrow (p\bar{p}\pi^+\pi^-)K^+$  results in the first observation of the decay  $\eta_c(2S) \rightarrow p\bar{p}\pi^+\pi^-$  with  $12.1\sigma$  significance. The analysis is based on the  $711 \text{ fb}^{-1}$  data sample collected by the Belle detector at the asymmetric-energy  $e^+e^-$  collider KEKB at the  $\Upsilon(4S)$  resonance.

DOI: [10.1103/PhysRevD.100.012001](https://doi.org/10.1103/PhysRevD.100.012001)

## I. INTRODUCTION

The decays  $B^+ \rightarrow \chi_{c0} K^+$ ,  $B^+ \rightarrow \chi_{c2} K^+$ , and  $B^+ \rightarrow h_c K^+$  are suppressed by factorization [1,2]. The decays  $B^+ \rightarrow \chi_{cJ} K^+$  have been observed; the current world-average branching fractions are  $\mathcal{B}(B^+ \rightarrow \chi_{c0} K^+) = (1.49^{+0.15}_{-0.14}) \times 10^{-4}$  and  $\mathcal{B}(B^+ \rightarrow \chi_{c2} K^+) = (1.1 \pm 0.4) \times 10^{-5}$  [3]. While  $\mathcal{B}(B^+ \rightarrow \chi_{c0} K^+)$  is smaller than the branching fraction of the factorization-allowed process  $\mathcal{B}(B^+ \rightarrow \chi_{c1} K^+) = (4.84 \pm 0.23) \times 10^{-4}$ , it is not strongly suppressed. Before the first experimental searches, this resulted in an assumption that the process  $B^+ \rightarrow h_c K^+$  may also have a large branching fraction  $\mathcal{B}(B^+ \rightarrow h_c K^+) \approx \mathcal{B}(B^+ \rightarrow \chi_{c0} K^+)$  [2].

However, the decay  $B^+ \rightarrow h_c K^+$  has not been observed experimentally yet. Neither the Belle [4] nor the BABAR [5] collaboration has found a statistically significant signal of  $B^+ \rightarrow h_c K^+$  using the decay mode  $h_c \rightarrow \eta_c \gamma$ . The current branching-fraction upper limit of  $\mathcal{B}(B^+ \rightarrow h_c K^+) < 3.8 \times 10^{-5}$  at 90% C.L. [3] was obtained in the  $h_c$  search by Belle [4].

Also, the LHCb Collaboration searched for the process  $B^+ \rightarrow h_c(\rightarrow p\bar{p})K^+$  [6] and set the upper limit on the branching fraction product  $\mathcal{B}(B^+ \rightarrow h_c K^+) \times \mathcal{B}(h_c \rightarrow p\bar{p}) < 6.4 \times 10^{-8}$  (95% C.L.). However, this measurement does not result in a stronger restriction on  $\mathcal{B}(B^+ \rightarrow h_c K^+)$  because the decay  $h_c \rightarrow p\bar{p}$  has never been observed and the upper limit on its branching fraction is  $\mathcal{B}(h_c \rightarrow p\bar{p}) < 1.5 \times 10^{-4}$  (90% C.L.) [3]. Note that a newer LHCb analysis of the same channel performed in Ref. [7] does not update the upper limit on  $\mathcal{B}(B^+ \rightarrow h_c K^+) \times \mathcal{B}(h_c \rightarrow p\bar{p})$ .

Several new theoretical predictions of  $\mathcal{B}(B^+ \rightarrow h_c K^+)$  were made after the experimental upper limit was set. The branching fraction has been calculated in the QCD factorization approach to be  $2.7 \times 10^{-5}$  [8]. A calculation using perturbative QCD was performed in Ref. [9]; the result is  $\mathcal{B}(B^+ \rightarrow h_c K^+) = 3.6 \times 10^{-5}$ . Another calculation performed in Ref. [10] results in  $\mathcal{B}(B^+ \rightarrow h_c K^+)$  [ $\mathcal{B}(B^0 \rightarrow h_c K^0)$ ] in the interval from  $3.1 \times 10^{-5}$  to  $5.7 \times 10^{-5}$  (from  $2.9 \times 10^{-5}$  to  $5.3 \times 10^{-5}$ ), depending on the assumed value of the  $c$  quark mass. All the results mentioned above are close to each other, and the theoretical values of  $\mathcal{B}(B^+ \rightarrow h_c K^+)$  are slightly below the current experimental upper limit. This motivates an updated study of the decays  $B^+ \rightarrow h_c K^+$ , which may be able to find them.

Here, we present such an updated search for the decays  $B^+ \rightarrow h_c K^+$  and also include a search for the decays  $B^0 \rightarrow h_c K_S^0$ . The analysis is performed using the  $711 \text{ fb}^{-1}$  data sample collected by the Belle detector at the asymmetric-energy  $e^+e^-$  collider KEKB [11]. The data sample was collected at the  $\Upsilon(4S)$  resonance and contains  $772 \times 10^6$   $B\bar{B}$  pairs. The integrated luminosity is 2.8 times greater than the luminosity used in the previous analysis [4]. For further improvement of the sensitivity, the new analysis uses ten  $\eta_c$  decay channels to reconstruct the decay  $h_c \rightarrow \eta_c \gamma$ ; only two channels were used in the old one. The new  $h_c$  decay channel  $h_c \rightarrow p\bar{p}\pi^+\pi^-$  observed recently by the BESIII Collaboration [12] is also used for its reconstruction; in addition, we study the decays of other charmonium states to  $p\bar{p}\pi^+\pi^-$ . The discrimination of the signal and background events is improved by performing a multivariate analysis.

## II. BELLE DETECTOR

The Belle detector is a large-solid-angle magnetic spectrometer that consists of a silicon vertex detector (SVD), a 50-layer central drift chamber (CDC), an array of aerogel threshold Cherenkov counters (ACC),

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP<sup>3</sup>.

a barrel-like arrangement of time-of-flight scintillation counters (TOF), and an electromagnetic calorimeter (ECL) composed of CsI(Tl) crystals located inside a superconducting solenoid coil that provides a 1.5 T magnetic field. An iron flux return located outside of the coil is instrumented to detect  $K_L^0$  mesons and to identify muons. The detector is described in detail elsewhere [13]. Two inner detector configurations were used. A 2.0 cm radius beam pipe and a three-layer silicon vertex detector were used for the first sample of  $140 \text{ fb}^{-1}$ , while a 1.5 cm radius beam pipe, a four-layer silicon detector, and a small-cell inner drift chamber were used to record the remaining data [14].

We use a GEANT-based Monte Carlo (MC) simulation [15] to model the response of the detector, identify potential backgrounds, and determine the acceptance. The MC simulation includes run-dependent detector performance variations and background conditions. Signal MC events are generated with EVTGEN [16] in proportion to the relative luminosities of the different running periods.

### III. EVENT SELECTION

We select events of the type  $B^+ \rightarrow h_c K^+$  and  $B^0 \rightarrow h_c K_S^0$ . Inclusion of charge-conjugate modes is implied hereinafter. The reconstruction is performed with a conversion from the Belle to the Belle II data format [17].

All tracks are required to originate from the interaction point region: we require  $dr < 0.2 \text{ cm}$  and  $|dz| < 2 \text{ cm}$ , where  $dr$  and  $dz$  are the cylindrical coordinates of the point of the closest approach of the track to the beam axis (the  $z$  axis of the laboratory reference frame coincides with the positron-beam axis).

Charged  $\pi$ ,  $K$  mesons and protons are identified using likelihood ratios  $R_{h_1/h_2} = \mathcal{L}_{h_1}/(\mathcal{L}_{h_1} + \mathcal{L}_{h_2})$ , where  $h_1$  and  $h_2$  are the particle-identification (PID) hypotheses ( $\pi$ ,  $K$ , or  $p$ ) and  $\mathcal{L}_{h_i}$  are their corresponding likelihoods. The likelihoods are calculated from the combined time-of-flight information from the TOF, the number of photoelectrons from the ACC, and  $dE/dx$  measurements in the CDC. We require  $R_{K/\pi} > 0.6$  for  $K$  candidates,  $R_{\pi/K} > 0.6$  for  $\pi$  candidates, and  $R_{p/\pi} > 0.6$ ,  $R_{p/K} > 0.6$  for  $p$  candidates. The identification efficiency of the above requirements varies in the ranges (94–99)%, (84–93)%, and (90–98)% for  $\pi$ ,  $K$ , and  $p$ , respectively, depending on the  $h_c$  or  $\eta_c$  decay channel. The misidentification probability for the background particles that are not  $\pi$ ,  $K$ , and  $p$ , varies in the ranges (25–49)%, (4.9–11.3)%, and (0.5–1.9)%, respectively. Without the electron background, which is rejected as described below, the  $\pi$  fake rate drops to (20–35)%.

Electron candidates are identified as CDC charged tracks that are matched to electromagnetic showers in the ECL. The track and ECL cluster matching quality, the ratio of the electromagnetic shower energy to the track momentum, the transverse shape of the shower, the ACC light yield, and

the track  $dE/dx$  ionization are used in our electron-identification criteria. A similar likelihood ratio is constructed:  $R_e = \mathcal{L}_e/(\mathcal{L}_e + \mathcal{L}_h)$ , where  $\mathcal{L}_e$  and  $\mathcal{L}_h$  are the likelihoods for electrons and charged hadrons ( $\pi$ ,  $K$  and  $p$ ), respectively [18]. An electron veto ( $R_e < 0.9$ ) is imposed on  $\pi$ ,  $K$ , and  $p$  candidates. It is not applied for the  $K_S^0$  and  $\Lambda$  daughter tracks because they have independent selection criteria. For the  $h_c$  or  $\eta_c$  decay channels other than  $\eta_c \rightarrow K_S^0 K_S^0 \pi^0$  and  $\eta_c \rightarrow \Lambda \bar{\Lambda}$ , the electron veto rejects from 3.5% to 15% of the background events, while its signal efficiency is no less than 97.5%.

Photons are identified as ECL electromagnetic showers that have no associated charged tracks detected in the CDC. The shower shape is required to be consistent with that of a photon.

The  $\pi^0$  candidates are reconstructed via their decay to two photons. The photon energies in the laboratory frame are required to be greater than 30 MeV. The  $\pi^0$  invariant mass is required to satisfy  $|M_{\pi^0} - m_{\pi^0}| < 15 \text{ MeV}/c^2$ . Here and elsewhere,  $M_{\text{particle}}$  denotes the reconstructed invariant mass of the specified particle, and  $m_{\text{particle}}$  stands for the nominal mass of this particle [3]. This requirement corresponds approximately to a  $3\sigma$  mass window around the nominal mass.

The  $V^0$ -particle ( $K_S^0$  and  $\Lambda$ ) candidates are reconstructed from pairs of oppositely charged tracks that are assumed to be  $\pi^+ \pi^-$  and  $p \pi^-$  for  $K_S^0$  and  $\Lambda$ , respectively. We require  $|M_{K_S^0} - m_{K_S^0}| < 20 \text{ MeV}/c^2$  and  $|M_{\Lambda} - m_{\Lambda}| < 10 \text{ MeV}/c^2$ , corresponding approximately to  $5.5\sigma$  mass windows in both cases. The  $V^0$  candidates are selected by a neural network using the following input variables: the  $V^0$  candidate momentum, decay angle, flight distance in the  $xy$  plane, the angle between the  $V^0$  momentum and the direction from the interaction point to the  $V^0$  vertex, the shortest  $z$  distance between the two daughter tracks, their radial impact parameters, and numbers of hits in the SVD and CDC. The separation of the  $K_S^0$  and  $\Lambda$  candidates is performed by another neural network. The input variables of this network are the momenta and polar angles of the daughter tracks in the laboratory frame, their likelihood ratios  $R_{\pi/p}$ , and the  $V^0$  candidate invariant mass for the  $\Lambda$  hypothesis.

The  $\eta$  candidates are reconstructed in  $\gamma\gamma$  and  $\pi^+ \pi^- \pi^0$  channels. The reconstructed  $\eta$  candidates are denoted by the  $\eta$  decay channel as  $\eta_{2\gamma}$  and  $\eta_{3\pi}$ . The  $\eta$  invariant mass is required to satisfy  $|M_{\eta_{2\gamma}} - m_{\eta}| < 30 \text{ MeV}/c^2$  and  $|M_{\eta_{3\pi}} - m_{\eta}| < 15 \text{ MeV}/c^2$ ; these requirements correspond to  $2.5\sigma$  and  $4\sigma$  mass windows, respectively.

The  $\eta'$  candidates are reconstructed in the  $\eta\pi^+ \pi^-$  decay mode. The invariant mass window is  $|M_{\eta'} - m_{\eta'}| < 15 \text{ MeV}/c^2$ , corresponding to a  $4\sigma$  mass window.

The  $\eta_c$  candidates are reconstructed in ten decay channels:  $K^+ K_S^0 \pi^-$ ,  $K^+ K^- \pi^0$ ,  $K_S^0 K_S^0 \pi^0$ ,  $K^+ K^- \eta$ ,  $K^+ K^- K^+ K^-$ ,

$\eta'(\rightarrow \eta\pi^+\pi^-)\pi^+\pi^-$ ,  $p\bar{p}$ ,  $p\bar{p}\pi^0$ ,  $p\bar{p}\pi^+\pi^-$ , and  $\Lambda\bar{\Lambda}$ . The selected  $\eta_c$  candidates are required to satisfy  $|M_{\eta_c} - m_{\eta_c}| < 50 \text{ MeV}/c^2$ ; the mass-window width is about 1.6 widths of the  $\eta_c$ .

The  $h_c$  candidates are reconstructed in the  $h_c \rightarrow \eta_c\gamma$  and  $h_c \rightarrow p\bar{p}\pi^+\pi^-$  decay channels. The invariant mass of the  $h_c$  candidates is not restricted for the channel  $\eta_c\gamma$ ; for the channel  $p\bar{p}\pi^+\pi^-$ , it is required to be greater than  $2.7 \text{ GeV}/c^2$ . The lower mass limit is selected to be very low to study other charmonium states decaying to the same final state.

The  $B$ -meson candidates are reconstructed via the decay modes  $B^+ \rightarrow h_c K^+$  and  $B^0 \rightarrow h_c K_S^0$ . The  $B$  candidates are selected by their energy and the beam-energy-constrained mass. The difference between the  $B$ -meson and beam energies is defined as  $\Delta E = \sum_i E_i - E_{\text{beam}}$ , where  $E_i$  are the energies of the  $B$  decay products in the center-of-mass frame and  $E_{\text{beam}}$  is the beam energy in the same frame. The beam-energy-constrained mass is defined as  $M_{\text{bc}} = \sqrt{E_{\text{beam}}^2 - (\sum_i \vec{p}_i)^2}$ , where  $\vec{p}_i$  are the momenta of the  $B$  decay products in the center-of-mass frame. We retain  $B$  candidates satisfying the conditions  $5.2 < M_{\text{bc}} < 5.3 \text{ GeV}/c^2$  and  $|\Delta E| < 0.2 \text{ GeV}$ . A mass-constrained fit is applied to the selected  $B$ -meson candidates.

In addition, for the channel  $h_c \rightarrow \eta_c\gamma$ , the  $h_c$  daughter  $\gamma$  energy is required to be greater than  $200 \text{ MeV}$  in the  $B$  rest frame. This requirement removes the background from low-energy photons, including the peaking backgrounds from  $B$  decays to the same final state without the photon. The signal efficiency of this requirement is 100% because the  $\eta_c\gamma$  invariant mass of all excluded events is smaller than the  $h_c$  mass.

Also, the ratio of the Fox-Wolfram moments [19]  $F_2/F_0$  is required to be less than 0.3. This requirement reduces the continuum background, rejecting from 18% to 53% of background events, depending on the  $h_c$  or  $\eta_c$  decay channel. Its signal efficiency is from 93.3% to 96.3%.

#### IV. MULTIVARIATE ANALYSIS AND OPTIMIZATION OF THE SELECTION REQUIREMENTS

##### A. General analysis strategy and data samples

To improve the separation of the signal and background events, we perform a multivariate analysis followed by an optimization of selection requirements. The first stages of the analysis are performed individually for  $h_c \rightarrow p\bar{p}\pi^+\pi^-$ , and each  $\eta_c$  decay channel for the  $h_c$  candidates reconstructed in the  $\eta_c\gamma$  mode [the channels  $\eta_c \rightarrow K^+K^-\eta$  and  $\eta_c \rightarrow \eta'(\rightarrow \eta\pi^+\pi^-)\pi^+\pi^-$  are optimized separately for  $\eta_{2\gamma}$  and  $\eta_{3\pi}$ ]. They include the determination of two-dimensional  $(\Delta E, M_{\text{bc}})$  resolution and the distribution of the background in  $(\Delta E, M_{\text{bc}})$  and the multivariate-analysis stage. The optimization of the selection requirements uses the results of all initial stages as its input. The resolution is

used to determine the expected number of the signal events, and the distribution of the background in  $(\Delta E, M_{\text{bc}})$  is used to determine the expected number of the background events in the signal region. The optimization is performed individually for the channel  $h_c \rightarrow p\bar{p}\pi^+\pi^-$  and globally for all  $\eta_c$  decay channels for the channel  $h_c \rightarrow \eta_c\gamma$ . The data selected using the resulting channel-dependent criteria are merged into a single sample for the  $h_c \rightarrow \eta_c\gamma$  channel. The final fit is performed simultaneously to the  $h_c \rightarrow \eta_c\gamma$  and  $h_c \rightarrow p\bar{p}\pi^+\pi^-$  samples.

The experimental data are used for the determination of the  $(\Delta E, M_{\text{bc}})$  distribution, selection of the background samples for the neural network, and final fit to the selected events. During the development of the analysis procedure, the  $h_c$  signal region was excluded to avoid bias of the  $h_c$  significance. The final fit described in Sec. V was performed on MC pseudoexperiments generated in accordance with the fit result without the  $h_c$  mixed with the  $B^+ \rightarrow h_c K^+$  ( $B^0 \rightarrow h_c K_S^0$ ) signal MC. The  $h_c$  signal region is defined by

$$\sqrt{\left(\frac{\Delta E}{\sigma_{\Delta E}}\right)^2 + \left(\frac{M_{\text{bc}} - m_B}{\sigma_{M_{\text{bc}}}}\right)^2} < 3, \quad (1)$$

where  $\sigma_{\Delta E} = 18 \text{ MeV}$  and  $\sigma_{M_{\text{bc}}} = 2.5 \text{ MeV}/c^2$  are the approximate resolutions in  $\Delta E$  and  $M_{\text{bc}}$ , respectively, and

$$\begin{aligned} 3.50 < M_{h_c} < 3.55 \text{ GeV}/c^2 & \text{ for } h_c \rightarrow \eta_c\gamma, \\ 3.515 < M_{h_c} < 3.535 \text{ GeV}/c^2 & \text{ for } h_c \rightarrow p\bar{p}\pi^+\pi^-. \end{aligned} \quad (2)$$

After completion of the analysis procedure development, this requirement is no longer used.

The signal MC is used for the determination of the resolution and the selection of the signal samples for the neural network. The signal MC is generated using the known information about the angular or invariant-mass distributions of the decay products if it is possible; otherwise, uniform distributions are assumed. The angular distribution is known for the channel  $h_c \rightarrow \eta_c\gamma$ . It does not have any free parameters and is proportional to  $\sin^2\theta_{h_c}$ , where  $\theta_{h_c}$  is the  $h_c$  helicity angle that is defined as the angle between  $-\vec{p}_B$  and  $\vec{p}_{\eta_c}$ , where  $\vec{p}_B$  and  $\vec{p}_{\eta_c}$  are the momenta of the  $B$  and  $\eta_c$  in the  $h_c$  rest frame, respectively. In addition, the  $\eta_c$  decay resonant structure is taken into account if it is known. The distributions for the channels  $K^+K_S^0\pi^-$ ,  $K^+K^-\pi^0$ ,  $K_S^0K_S^0\pi^0$ ,  $K^+K^-\eta_{2\gamma}$ , and  $K^+K^-\eta_{3\pi}$  are based on the results of a Dalitz plot analysis performed in Ref. [20]. The contributions of intermediate  $\phi$  resonances are taken into account for the channel  $K^+K^-K^+K^-$  based on the world-average branching fractions from Ref. [3].

### B. Resolution

The resolution is parametrized by the function

$$S(\Delta E, M_{bc}) = N_{CB} F_{CB}(x_1) G_a^{(12)}(y_1) + N_{G1} G_a^{(21)}(x_2) G_a^{(22)}(y_2) + N_{G2} G_a^{(31)}(x_3) G_a^{(32)}(y_3), \quad (3)$$

where  $F_{CB}$  is an asymmetric Crystal Ball function [21];  $G_a^{(ij)}$  are asymmetric Gaussian functions;  $N_{CB}$ ,  $N_{G1}$ , and  $N_{G2}$  are normalizations; and  $x_i$  and  $y_i$  ( $i = 1, 2, 3$ ) are rotated variables that are given by

$$\begin{pmatrix} x_i \\ y_i \end{pmatrix} = \begin{pmatrix} \cos \alpha_i & \sin \alpha_i \\ -\sin \alpha_i & \cos \alpha_i \end{pmatrix} \begin{pmatrix} \Delta E - (\Delta E)_0 \\ M_{bc} - (M_{bc})_0 \end{pmatrix}. \quad (4)$$

Here,  $((\Delta E)_0, (M_{bc})_0)$  is the central point, and  $\alpha_i$  is the rotation angle. The central point is the same for all three components. The resolution is determined from a binned maximum likelihood fit to signal MC events. Example resolution fit results [for the channel  $B^+ \rightarrow h_c K^+$  with  $h_c \rightarrow \eta_c (\rightarrow K^+ K^- \pi^0) \gamma$ ] are shown in Fig. 1.

### C. Fit to the $(\Delta E, M_{bc})$ distribution

The  $(\Delta E, M_{bc})$  distribution is fitted in order to estimate the expected number of the background events in the signal region. The distribution is fitted to the function

$$N_S S(\Delta E, M_{bc}) + B(\Delta E, M_{bc}), \quad (5)$$

where  $N_S$  is the number of signal events and  $B$  is the background density function that is given by

$$B(\Delta E, M_{bc}) = \sqrt{m_0 - M_{bc}} \exp[-a(m_0 - M_{bc})] \times P_3(\Delta E, M_{bc}), \quad (6)$$

where  $m_0$  is the threshold mass,  $a$  is a rate parameter, and  $P_3$  is a two-dimensional third-order polynomial. The region with  $\Delta E < -0.12$  GeV is excluded for the channel  $h_c \rightarrow p \bar{p} \pi^+ \pi^-$  because of the presence of peaking backgrounds from partially reconstructed B decays with an additional  $\pi$  meson.

Example  $(\Delta E, M_{bc})$  fit results [for the channel  $B^+ \rightarrow h_c K^+$  with  $h_c \rightarrow \eta_c (\rightarrow K^+ K^- \pi^0) \gamma$ ] are shown in Fig. 2.

### D. Multivariate analysis

To improve the separation of signal and background events, we perform a multivariate analysis for each individual channel. The algorithm used for the multivariate analysis is the multilayer perceptron (MLP) neural network implemented in the TMVA library [22]. The following variables are always included in the neural network: the angle between the thrust axes of the  $B$  candidate and the remaining particles in

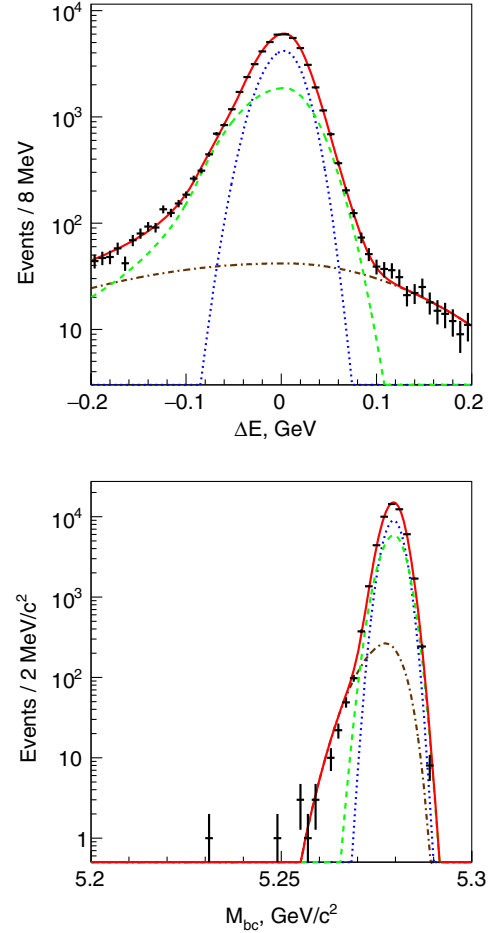


FIG. 1. Projections of the resolution fit results onto  $\Delta E$  and  $M_{bc}$  for the channel  $B^+ \rightarrow h_c K^+$  with  $h_c \rightarrow \eta_c (\rightarrow K^+ K^- \pi^0) \gamma$ . The red solid line is the fit result, the green dashed line is the Crystal Ball component, the blue dotted line is the first Gaussian component, and the brown dash-dotted line is the second Gaussian component.

the event, the angle between the thrust axes of all tracks and all photons in the event, the ratio of the Fox-Wolfram moments  $F_2/F_0$ , the  $B$  production angle, and the vertex fit quality. For the  $h_c$  candidates reconstructed in the  $\eta_c \gamma$  channel, the MLP also includes the  $h_c$  helicity angle, the  $\eta_c$  mass, and the number of  $\pi^0$  candidates that include the  $h_c$  daughter photon as one of their daughters (separately for two groups of  $\pi^0$  candidates with the energy of another photon less and greater than 100 MeV).

For the channels  $\eta_c \rightarrow K^+ K_S^0 \pi^-$ ,  $\eta_c \rightarrow K^+ K^- \pi^0$ , and  $\eta_c \rightarrow K_S^0 K_S^0 \pi^0$ , two invariant masses of the  $\eta_c$  daughter particle pairs (both  $K\pi$  combinations) are added to the neural network.

The following PID variables are included in the neural network if there are corresponding charged particles in the final state: the minimum likelihood ratio  $R_{K/\pi}$  of the  $h_c$  daughter kaons, the minimum of the two likelihood ratios  $R_{p/K}$ ,  $R_{p/\pi}$  of the  $h_c$  daughter protons, and  $R_{K/\pi}$  for the  $B$

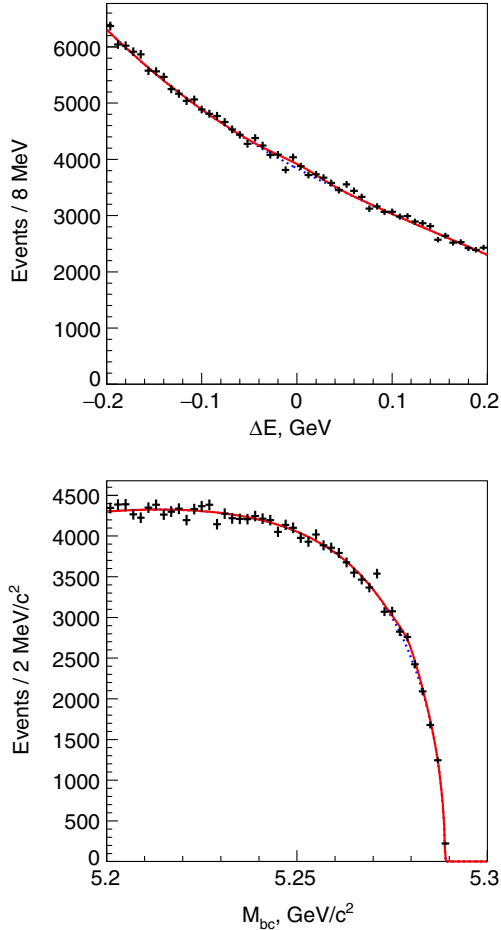


FIG. 2. Projections of the results of the fit to the  $(\Delta E, M_{bc})$  distribution onto  $\Delta E$  (with  $M_{bc} > 5.272 \text{ GeV}/c^2$ ) and  $M_{bc}$  (with  $|\Delta E| < 20 \text{ MeV}$ ) for the channel  $B^+ \rightarrow h_c K^+$  with  $h_c \rightarrow \eta_c(\rightarrow K^+ K^- \pi^0)\gamma$ . The red solid line is the fit result, and the blue dotted line is the background. Since there is no significant signal before the optimization of the selection requirements and for the entire  $\eta_c\gamma$  mass range, the two lines almost coincide.

daughter  $K^+$  (for the channel  $B^+ \rightarrow h_c K^+$ ). Here, the  $h_c$  daughters may be either direct (from the decay  $h_c \rightarrow p\bar{p}\pi^+\pi^-$ ) or indirect (the  $\eta_c$  daughters for the  $h_c$  candidates reconstructed in the  $\eta_c\gamma$  mode).

If there is a  $\pi^0$  or  $\eta$  decaying to  $\gamma\gamma$  in the final state, four additional variables are added: the  $\pi^0$  ( $\eta$ ) mass, the minimal energy of the  $\pi^0$  ( $\eta$ ) daughter photons in the laboratory frame, and the number of  $\pi^0$  candidates that include a  $\pi^0$  ( $\eta$ ) daughter photon as one of their daughters [for each of the  $\pi^0$  ( $\eta$ ) daughter photons]. If there is an  $\eta$  reconstructed in the  $\pi^+\pi^-\pi^0$  decay mode, then only its mass is added to the MLP. If the  $\eta_c$  has a daughter  $\eta'$ , then the mass of the  $\eta'$  candidate is also included to the neural network.

The training and testing signal samples are taken from the signal MC. The background sample is taken from a two-dimensional  $(\Delta E, M_{bc})$  sideband. For the channel  $B^+ \rightarrow h_c K^+$ , the sideband is defined as

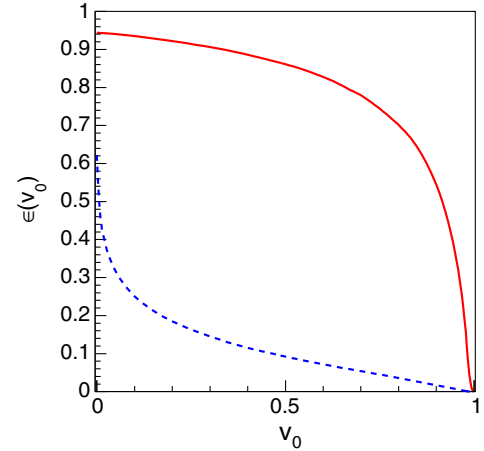


FIG. 3. Efficiency of the MLP output requirement ( $v > v_0$ ) for the channel  $B^+ \rightarrow h_c K^+$  with  $h_c \rightarrow \eta_c(\rightarrow K^+ K^- \pi^0)\gamma$ . The red solid line is the signal efficiency, and the blue dashed line is the background efficiency.

$$3 < \sqrt{\left(\frac{\Delta E}{\sigma_{\Delta E}}\right)^2 + \left(\frac{M_{bc} - m_B}{\sigma_{M_{bc}}}\right)^2} < 8. \quad (7)$$

The background sample is divided into training and testing samples of equal size.

The channel  $B^0 \rightarrow h_c K_S^0$  has a small number of background events. In order to avoid overtraining, the background region for this channel is redefined. It includes all selected events except the central region defined by Eq. (1). In addition, the MLP internal architecture is changed. Instead of the default TMVA neural network with two hidden layers, only one hidden layer is used.

The resulting efficiency of the requirement ( $v > v_0$ ) on the MLP output variable  $v$  for the training sample is shown in Fig. 3 for the channel  $B^+ \rightarrow h_c K^+$  with  $h_c \rightarrow \eta_c(\rightarrow K^+ K^- \pi^0)\gamma$ . Note that the efficiency is given by

$$\epsilon(v_0) = \epsilon_{\text{MLP}}(v_0) \times \epsilon_{\text{multiple}}(v_0), \quad (8)$$

where  $\epsilon$  is the full efficiency,  $\epsilon_{\text{MLP}}$  is the raw MLP output requirement efficiency and  $\epsilon_{\text{multiple}}$  is the best-candidate selection efficiency. Because of the correction by  $\epsilon_{\text{multiple}}$ , the efficiency at the minimal MLP output value  $v_{\text{min}}$  is not 1 but rather  $\epsilon_{\text{multiple}}(v_{\text{min}})$ .

The best-candidate selection is performed for each of the multivariate-analysis channels separately in the following way. The selected  $(\Delta E, M_{bc})$  region is subdivided into three bins in both  $\Delta E$  and  $M_{bc}$ . The selection is performed for each of the bins separately. The candidate with the largest MLP output is selected. One of the bins ( $-67 < \Delta E < 67 \text{ MeV}$ ,  $5.267 < M_{bc} < 5.3 \text{ GeV}/c^2$ ) always contains the entire signal region selected by the optimization procedure as described in Sec. IV E. Thus, the signal region of the final data sample does not contain multiple

candidates that originate from the same  $h_c$  channel. However, multiple candidates from different  $h_c$  channels are possible.

The best-candidate selection efficiency increases for larger values of the MLP output cutoff value  $v_0$ . For the  $v_0$  values obtained as the result of the optimization of the selection requirements as described in Sec. IV E, the selection procedure removes from 3% to 15% of data events, depending on the multivariate-analysis channel.

### E. Optimization of the selection requirements

Optimization of the selection requirements is performed by maximizing the value

$$F_{\text{opt}} = \frac{\sum_i N_{\text{sig}}^{(i)}}{\frac{a}{2} + \sqrt{\sum_i N_{\text{bg}}^{(i)}}}, \quad (9)$$

where  $i$  is the channel index,  $N_{\text{sig}}^{(i)}$  is the expected number of the signal events for the  $i$ th channel,  $N_{\text{bg}}^{(i)}$  is the expected number of the background events in the signal region, and  $a = 3$  is the target significance. This optimization method is based on Ref. [23].

The signal region is defined as

$$\left(\frac{\Delta E}{R_{\Delta E}^{(i)}}\right)^2 + \left(\frac{M_{\text{bc}} - m_B}{R_{M_{\text{bc}}}^{(i)}}\right)^2 < 1. \quad (10)$$

where  $R_{\Delta E}^{(i)}$  and  $R_{M_{\text{bc}}}^{(i)}$  are the half-axes of the signal region ellipse. The parameters determined by the optimization are  $R_{\Delta E}^{(i)}$ ,  $R_{M_{\text{bc}}}^{(i)}$ , and the minimal value of the MLP output ( $v_0^{(i)}$ ) for each channel.

The expected number of signal events for  $B^+ \rightarrow h_c K^+$  is calculated as

$$N_{\text{sig}}^{(i)} = 2N_{\Upsilon(4S)} \mathcal{B}(\Upsilon(4S) \rightarrow B^+ B^-) \mathcal{B}(B^+ \rightarrow h_c K^+) \times \mathcal{B}(h_c \rightarrow i) \epsilon_{\text{SR}}^{(i)} \epsilon_S^{(i)}(v_0^{(i)}), \quad (11)$$

where  $N_{\Upsilon(4S)}$  is the number of  $\Upsilon(4S)$  events,  $\mathcal{B}(h_c \rightarrow i)$  is the branching fraction of the  $h_c$  to its  $i$ th decay channel,  $\epsilon_{\text{SR}}^{(i)}$  is the reconstruction efficiency for the specific signal region SR, and  $\epsilon_S^{(i)}(v_0^{(i)})$  is the efficiency of the requirement ( $v > v_0$ ) on the MLP output variable  $v$  for the signal events. The number of  $\Upsilon(4S)$  events is assumed to be equal to the number of  $B\bar{B}$  pairs; the branching fraction  $\mathcal{B}(\Upsilon(4S) \rightarrow B^+ B^-)$  is calculated under the same assumption [3]. The signal-region-dependent reconstruction efficiency is calculated as

$$\epsilon_{\text{SR}}^{(i)} = \epsilon_R^{(i)} \int_{\text{SR}} S_i(\Delta E, M_{\text{bc}}) d\Delta E dM_{\text{bc}}, \quad (12)$$

where  $\epsilon_R^{(i)}$  is the reconstruction efficiency and  $S_i$  is the signal probability density function for  $i$ th  $h_c$  decay channel (the integral of  $S_i$  over the signal region is the efficiency of the signal region selection). The unknown branching fraction  $\mathcal{B}(B^+ \rightarrow h_c K^+)$  can be set to an arbitrary value because the maximum of  $F_{\text{opt}}$  does not depend on it. The expected number of signal events for  $B^0 \rightarrow h_c K_S^0$  is calculated similarly.

The expected number of background events is calculated as

$$N_{\text{bg}}^{(i)} = \epsilon_B^{(i)}(v_0^{(i)}) \frac{N_{h_c \text{ region}}}{N_{\text{full}}} \int_{\text{SR}} B_i(\Delta E, M_{\text{bc}}) d\Delta E dM_{\text{bc}}, \quad (13)$$

where  $\epsilon_B^{(i)}(v_0^{(i)})$  is the efficiency of the MLP output requirement for the background events,  $N_{h_c \text{ region}}$  is the number of background events in the  $h_c$  region defined by Eq. (2),  $N_{\text{full}}$  is the full number of the background events, and  $B_i$  is the background density function defined in Eq. (6) for the  $i$ th  $h_c$  decay channel.

The optimization is performed separately for two channel groups. The first group includes the multivariate-analysis channels corresponding to the decay  $h_c \rightarrow \eta_c \gamma$ ; the index  $i$  runs over all  $\eta_c$  decay channels. The second group consists of the single channel  $h_c \rightarrow p\bar{p}\pi^+\pi^-$ . The separate optimization is required by the difference of further data processing: the data from the first group of channels are combined into a single  $h_c \rightarrow \eta_c \gamma$  data sample, while the  $h_c \rightarrow p\bar{p}\pi^+\pi^-$  data are fitted with another function, as described below in Sec. V. The optimization results are shown in Table I. We also check the improvement achieved by MLP usage by changing the selection method to rectangular cuts. The values of  $F_{\text{opt}}$  are found to be about 30% and 10% smaller for the channels  $h_c \rightarrow \eta_c \gamma$  and  $h_c \rightarrow p\bar{p}\pi^+\pi^-$ , respectively.

After the optimization, the resulting selection criteria are applied. The selected events for the channel  $h_c \rightarrow \eta_c \gamma$  are merged. The resolution and distribution in  $(\Delta E, M_{\text{bc}})$  are determined again for the  $h_c \rightarrow p\bar{p}\pi^+\pi^-$  sample, since the knowledge of the background distribution in  $(\Delta E, M_{\text{bc}})$  is necessary for the final fit described in Sec. V. The  $(\Delta E, M_{\text{bc}})$  fit results are shown in Fig. 4.

### F. Resolution in $M_{h_c}$

The resolution in  $M_{h_c}$  is determined from a fit to the combined  $h_c \rightarrow \eta_c \gamma$  or  $h_c \rightarrow p\bar{p}\pi^+\pi^-$  signal MC samples with  $\eta_c$  decaying to the reconstructed channels only. All final selection criteria are applied. The distribution of the difference of the reconstructed and true masses is fitted to the sum of an asymmetric Gaussian and asymmetric double-sided Crystal Ball functions,

$$R_{h_c}(\Delta M) = N[F_{\text{CB}}(\Delta M)f_{\text{CB}} + G_a(\Delta M)(1 - f_{\text{CB}})], \quad (14)$$

where  $\Delta M$  is the difference of the reconstructed and true  $h_c$  masses,  $N$  is the common normalization, and  $f_{\text{CB}}$  is the

TABLE I. Results of the optimization of the selection requirements. The signal-region half-axes  $R_{\Delta E}^{(i)}$  ( $R_{M_{bc}}^{(i)}$ ) are in MeV (MeV/ $c^2$ ); all other values are dimensionless.

Channel	$B^+ \rightarrow h_c K^+$						$B^0 \rightarrow h_c K_S^0$					
	Parameters			Efficiency			Parameters			Efficiency		
	$R_{\Delta E}^{(i)}$	$R_{M_{bc}}^{(i)}$	$v_0^{(i)}$	$\epsilon_{SR}^{(i)}$	$\epsilon_S^{(i)}(v_0^{(i)})$	$\epsilon_B^{(i)}(v_0^{(i)})$	$R_{\Delta E}^{(i)}$	$R_{M_{bc}}^{(i)}$	$v_0^{(i)}$	$\epsilon_{SR}^{(i)}$	$\epsilon_S^{(i)}(v_0^{(i)})$	$\epsilon_B^{(i)}(v_0^{(i)})$
Channel group 1: $h_c \rightarrow \eta_c \gamma$												
$\eta_c(\rightarrow K^+ K_S^0 \pi^-) \gamma$	32.7	4.82	0.804	6.27%	59.5%	5.08%	34.2	4.97	0.702	4.31%	69.8%	8.55%
$\eta_c(\rightarrow K^+ K^- \pi^0) \gamma$	36.2	3.90	0.958	4.27%	31.7%	0.56%	43.5	4.54	0.942	3.38%	39.3%	0.85%
$\eta_c(\rightarrow K_S^0 K_S^0 \pi^0) \gamma$	42.3	4.49	0.976	1.79%	17.8%	0.18%	35.9	4.07	0.954	1.05%	35.9%	0.64%
$\eta_c(\rightarrow K^+ K^- \eta_{2\gamma}) \gamma$	34.4	4.16	0.977	4.21%	20.2%	0.22%	37.8	4.49	0.967	3.10%	28.0%	0.41%
$\eta_c(\rightarrow K^+ K^- \eta_{3\pi}) \gamma$	24.9	3.59	0.978	1.75%	29.7%	0.23%	33.2	4.55	0.986	1.50%	23.7%	0.17%
$\eta_c(\rightarrow K^+ K^- K^+ K^-) \gamma$	25.3	4.13	0.770	4.89%	53.2%	6.69%	29.9	4.80	0.734	3.71%	56.9%	8.49%
$\eta_c(\rightarrow \eta'(\rightarrow \eta_{2\gamma} \pi^+ \pi^-) \pi^+ \pi^-) \gamma$	30.5	4.21	0.958	2.69%	40.5%	0.63%	32.0	4.50	0.946	1.87%	45.6%	0.96%
$\eta_c(\rightarrow \eta'(\rightarrow \eta_{3\pi} \pi^+ \pi^-) \pi^+ \pi^-) \gamma$	26.8	4.16	0.990	1.01%	29.2%	0.13%	24.6	3.87	0.986	0.59%	32.3%	0.26%
$\eta_c(\rightarrow p\bar{p}) \gamma$	38.9	5.48	0.654	17.70%	75.7%	10.66%	42.5	5.85	0.513	12.47%	82.8%	15.43%
$\eta_c(\rightarrow p\bar{p}\pi^0) \gamma$	30.2	3.75	0.954	4.65%	30.3%	0.50%	31.8	4.01	0.934	3.33%	39.4%	0.93%
$\eta_c(\rightarrow p\bar{p}\pi^+\pi^-) \gamma$	24.1	4.03	0.912	6.31%	30.0%	1.53%	24.3	4.09	0.860	4.23%	41.8%	3.26%
$\eta_c(\rightarrow \Lambda\bar{\Lambda}) \gamma$	40.4	5.66	0.727	4.04%	70.6%	6.79%	41.5	5.19	0.586	2.65%	76.3%	11.24%
Channel group 2: $h_c \rightarrow p\bar{p}\pi^+\pi^-$												
$p\bar{p}\pi^+\pi^-$	13.5	4.36	0.598	14.81%	64.6%	18.40%	13.8	4.56	0.519	10.30%	71.2%	24.20%

Crystal Ball fraction. Example resolution fit results (for the channel  $B^+ \rightarrow h_c K^+$  with  $h_c \rightarrow \eta_c \gamma$ ) are shown in Fig. 5.

## V. FIT TO THE DATA

### A. Default model

For the  $h_c \rightarrow \eta_c \gamma$  final sample, the distribution in the  $h_c$  mass in the  $(\Delta E, M_{bc})$  sideband cannot be used to constrain the background level in the signal region because of the presence of peaking backgrounds, such as the background from  $B$  decays to a similar final state with a  $\pi^0$  instead of the  $h_c$  daughter  $\gamma$ . If the second photon from this  $\pi^0$  has a small energy, then the  $\Delta E$  and  $M_{bc}$  values are close to 0 and the  $B$  mass, respectively. Thus, the fit is based on the signal distribution only for the channel  $h_c \rightarrow \eta_c \gamma$ .

For the channel  $h_c \rightarrow p\bar{p}\pi^+\pi^-$ , there is a signal from  $B$  decays to the same final state ( $B^+ \rightarrow p\bar{p}\pi^+\pi^- K^+$  or  $B^0 \rightarrow p\bar{p}\pi^+\pi^- K_S^0$ ) that do not proceed via any charmonium state, called the noncharmonium signal hereinafter. Because of the possible interference of the charmonium and noncharmonium signals, the distribution of the noncharmonium signal in the  $p\bar{p}\pi^+\pi^-$  invariant mass needs to be determined by the fit. Thus, both signal and background

distributions are included in the fit for the channel  $h_c \rightarrow p\bar{p}\pi^+\pi^-$ .

We perform a simultaneous extended unbinned maximum likelihood fit to the  $h_c \rightarrow \eta_c \gamma$  signal,  $h_c \rightarrow p\bar{p}\pi^+\pi^-$  background, and  $h_c \rightarrow p\bar{p}\pi^+\pi^-$  signal distributions. The charmonium states are represented by the Breit-Wigner amplitude

$$A_R(M_R) = \frac{1}{M_R^2 - m_R^2 + iM_R\Gamma_R}, \quad (15)$$

where  $M_R$  is the invariant mass,  $m_R$  is the nominal mass, and  $\Gamma_R$  is the width of the resonance  $R$ . The signal-region density function for the channel  $h_c \rightarrow \eta_c \gamma$  is given by

$$S_{\eta_c\gamma}(M) = (N_{h_c} |A_{h_c}(M)|^2) \otimes R_{h_c}^{(\eta_c\gamma)}(\Delta M) + P_2(M), \quad (16)$$

where  $N_{h_c}$  is the number of signal events,  $R_{h_c}^{(\eta_c\gamma)}$  is the  $h_c$  mass resolution for the channel  $\eta_c \gamma$ , and  $P_2$  is a second-order polynomial. The background density function  $B_{p\bar{p}\pi^+\pi^-}(M)$  for the channel  $h_c \rightarrow p\bar{p}\pi^+\pi^-$  is a third-order polynomial. The signal density function for the channel  $h_c \rightarrow p\bar{p}\pi^+\pi^-$  is given by

$$S_{p\bar{p}\pi^+\pi^-}(M) = \left( |P_3(M) + \sum_{R=\eta_c, \chi_{c0}, \eta_c(2S)} \sqrt{N_R} e^{i\varphi_R} A_R(M)|^2 + \sum_{R=J/\psi, \chi_{c1}, h_c, \chi_{c2}, \psi(2S)} N_R |A_R(M)|^2 \right) \otimes R_{h_c}^{(p\bar{p}\pi^+\pi^-)}(\Delta M), \quad (17)$$

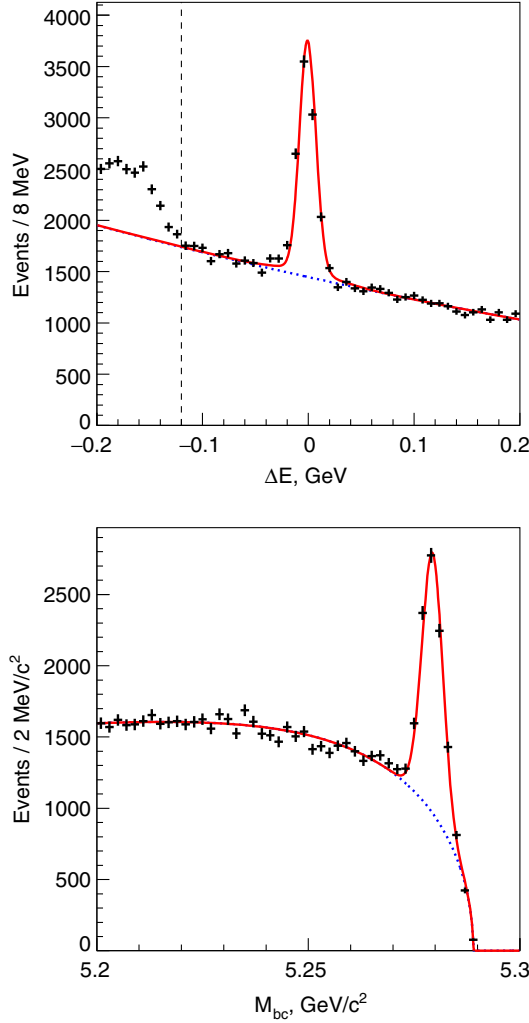


FIG. 4. Projections of the results of the fit to the  $(\Delta E, M_{bc})$  distribution onto  $\Delta E$  (with  $M_{bc} > 5.272 \text{ GeV}/c^2$ ) and  $M_{bc}$  (with  $|\Delta E| < 20 \text{ MeV}$ ) for the channel  $B^+ \rightarrow h_c K^+$  with  $h_c \rightarrow p\bar{p}\pi^+\pi^-$  after the application of the final MLP output selection criterion. The red solid line is the fit result, and the blue dotted line is the background. The region with  $\Delta E < -0.12 \text{ GeV}$  is excluded from the fit because of the presence of peaking backgrounds from partially reconstructed B decays with an additional  $\pi$  meson. The cutoff value is marked by a vertical dashed line.

where  $P_3$  is a third-order polynomial representing the noncharmonium signal. The wide states are added coherently to the signal density function, while the states that are narrower than the resolution are added incoherently. The amplitudes are normalized in such a way that all the parameters  $N_R$  represent the yields of the corresponding states. The signal distribution is fitted to the function

$$S_{p\bar{p}\pi^+\pi^-}(M) + wB_{p\bar{p}\pi^+\pi^-}(M), \quad (18)$$

where  $w$  is the weight of the background events in the signal region that is calculated as the ratio of integrals of the

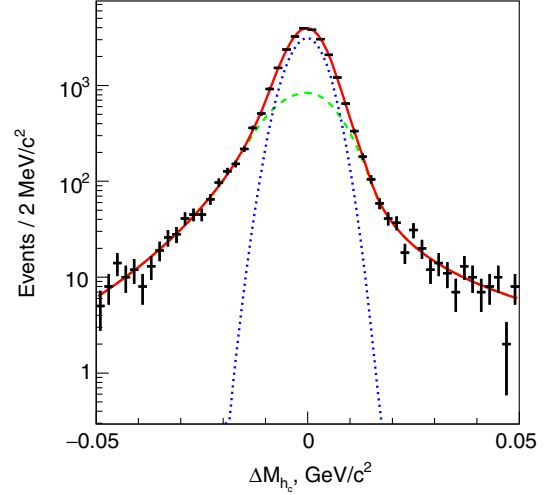


FIG. 5. Resolution in  $M_{h_c}$  for the channel  $B^+ \rightarrow h_c K^+$  with  $h_c \rightarrow \eta_c \gamma$ . The red solid line is the fit result, the green dashed line is the Crystal Ball component, and the blue dotted line is the Gaussian component.

background distribution in  $(\Delta E, M_{bc})$  over the signal and background regions. The model described above is the default one; additional models are considered to study systematic uncertainties. In the default model, the masses and widths of all resonances are fixed to their world-average values [3]; all other parameters are free.

The best-candidate selection procedure described in Sec. IV D guarantees that there are no multiple candidates in the  $h_c \rightarrow p\bar{p}\pi^+\pi^-$  signal sample, but multiple candidates in the  $h_c \rightarrow \eta_c \gamma$  signal sample are possible if they originate from different  $\eta_c$  decay channels. However, the fraction of the events with multiple candidates is found to be negligibly small. No events with multiple candidates (for the  $h_c$  masses within the default fitting regions) are observed for both  $B^+ \rightarrow h_c K^+$  and  $B^0 \rightarrow h_c K_S^0$  channels.

The fit results are shown in Fig. 6 for the channel  $B^+ \rightarrow h_c K^+$  and in Fig. 7 for the channel  $B^0 \rightarrow h_c K_S^0$ . The signal yields and phases are listed in Table II. The statistical significance of the decays  $B^+ \rightarrow h_c K^+$  and  $B^0 \rightarrow h_c K_S^0$ , as well as the significances of other charmonium states in the channel  $p\bar{p}\pi^+\pi^-$ , are calculated from the difference of  $(-2 \ln L)$ , where  $L$  is the maximum likelihood, between the models with and without these states taking the number of degrees of freedom into account. The significances of the decays  $B^+ \rightarrow h_c K^+$  and  $B^0 \rightarrow h_c K_S^0$  in the default model are found to be  $5.0\sigma$  and  $0.8\sigma$ , respectively. The significances of the decays  $B^+ \rightarrow h_c K^+$  and  $B^0 \rightarrow h_c K_S^0$  with the systematic error taken into account are  $4.8\sigma$  and  $0.7\sigma$ , respectively; the procedure of the calculation of the systematic uncertainty is described in Sec. V B. Thus, we find evidence for the decay  $B^+ \rightarrow h_c K^+$  but do not find evidence for  $B^0 \rightarrow h_c K_S^0$ . The significances of charmonium states in the channel  $p\bar{p}\pi^+\pi^-$  (except the  $h_c$ , which is

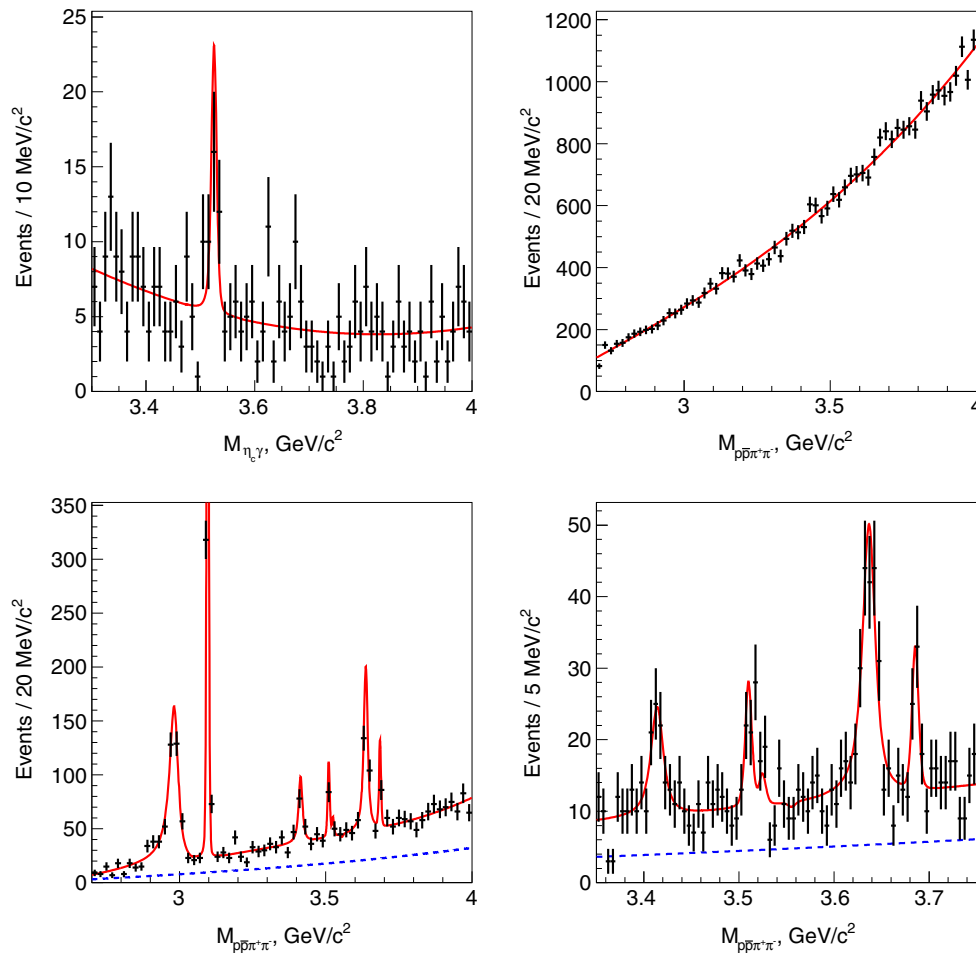


FIG. 6. Fit results in the  $B^+ \rightarrow h_c K^+$  channel:  $h_c \rightarrow \eta_c \gamma$  signal (top left),  $h_c \rightarrow p\bar{p}\pi^+\pi^-$  background (top right), and  $h_c \rightarrow p\bar{p}\pi^+\pi^-$  signal in the entire fitting region (bottom left) and in the  $\chi_{cJ}$  region (bottom right). The red solid line is the fit result, and the blue dashed line is the background. The maximum of the  $p\bar{p}\pi^+\pi^-$  signal fit result at the  $J/\psi$  peak is more than two times greater than the number of data events in the corresponding bin because the  $J/\psi$  peak is narrower than the bin size. Thus, a part of the fit result at the  $J/\psi$  peak is not shown.

reconstructed in two decay channels) are shown in Table III. The significances of the  $\eta_c(2S)$  in the default model are  $12.3\sigma$  and  $5.9\sigma$  for the processes  $B^+ \rightarrow (c\bar{c})(\rightarrow p\bar{p}\pi^+\pi^-)K^+$  and  $B^0 \rightarrow (c\bar{c})(\rightarrow p\bar{p}\pi^+\pi^-)K_S^0$ , respectively. The significances including the systematic error are  $12.1\sigma$  and  $5.8\sigma$ , respectively. Consequently, the decay  $\eta_c(2S) \rightarrow p\bar{p}\pi^+\pi^-$  is observed for the first time in both  $B^+ \rightarrow (c\bar{c})(\rightarrow p\bar{p}\pi^+\pi^-)K^+$  and  $B^0 \rightarrow (c\bar{c})(\rightarrow p\bar{p}\pi^+\pi^-)K_S^0$  processes.

### B. Systematic uncertainty: Model dependence

For a systematic-uncertainty study, we consider additional models. They include the models with free masses and widths of the  $h_c$  and all other charmonium states (with Gaussian constraints in accordance with the errors of their current world-average values), with increased order (3) of the background probability density function polynomial, different fitting ranges, scaled resolution, and variation of

the relative fraction of the channels  $h_c \rightarrow \eta_c \gamma$  and  $h_c \rightarrow p\bar{p}\pi^+\pi^-$ . For the model with scaled resolution, the resolution function  $R_{h_c}(\Delta M)$  is changed to

$$R_{h_c}(\Delta M) \rightarrow \frac{1}{S} R_{h_c}\left(\frac{\Delta M}{S}\right), \quad (19)$$

where  $S$  is the resolution scaling parameter. The variation of the relative fraction of the channels  $h_c \rightarrow \eta_c \gamma$  and  $h_c \rightarrow p\bar{p}\pi^+\pi^-$  is performed by changing the expected yields in these channels by  $\pm 1\sigma$ , where the error is due to the error of the corresponding branching fractions. The results are listed in Table IV.

### C. Branching fraction

Using the number of reconstructed events, we calculate the branching fractions of the decays  $B^+ \rightarrow h_c K^+$  and  $B^0 \rightarrow h_c K_S^0$ , as well as the branching fraction products

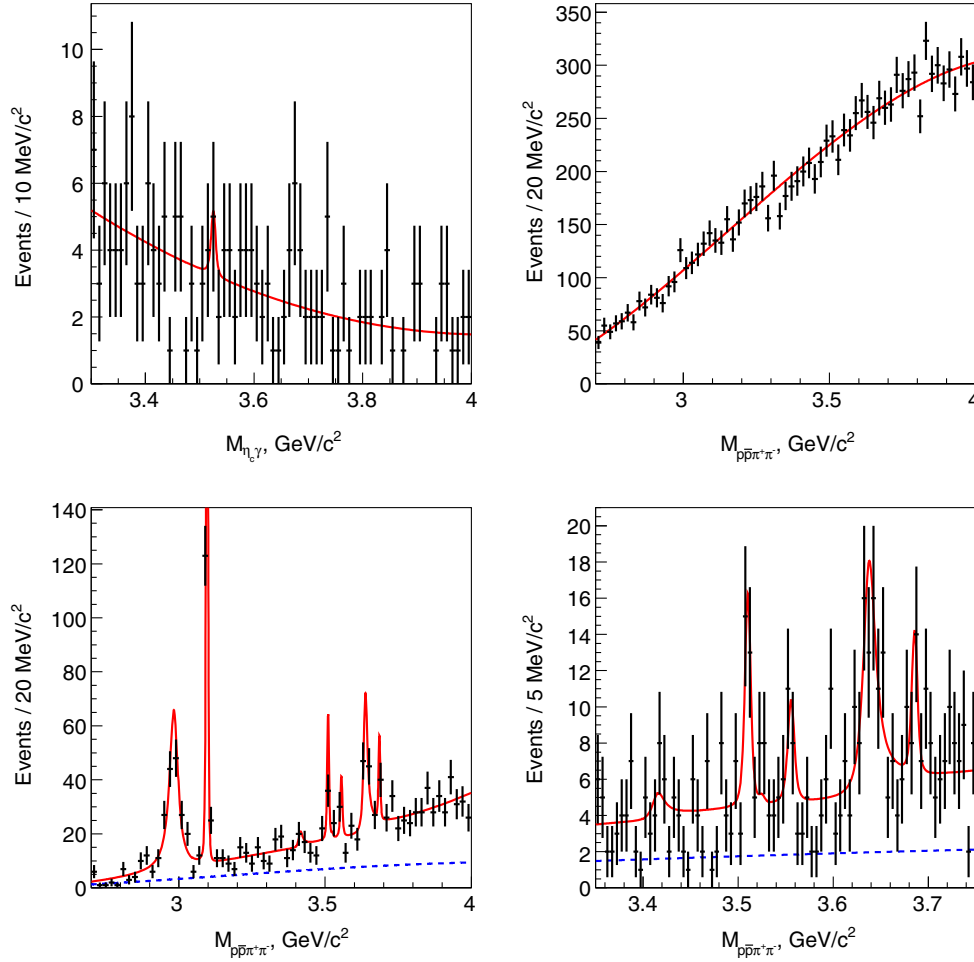


FIG. 7. Fit results in the  $B^0 \rightarrow h_c K_S^0$  channel:  $h_c \rightarrow \eta_c \gamma$  signal (top left),  $h_c \rightarrow p \bar{p} \pi^+ \pi^-$  background (top right), and  $h_c \rightarrow p \bar{p} \pi^+ \pi^-$  signal in the entire fitting region (bottom left) and in the  $\chi_{cJ}$  region (bottom right). The red solid line is the fit result, and the blue dashed line is the background. The maximum of the  $p \bar{p} \pi^+ \pi^-$  signal fit result at the  $J/\psi$  peak is more than two times greater than the number of data events in the corresponding bin because the  $J/\psi$  peak is narrower than the bin size. Thus, a part of the fit result at the  $J/\psi$  peak is not shown.

$\mathcal{B}(B^+ \rightarrow (c\bar{c})K^+) \times \mathcal{B}((c\bar{c}) \rightarrow p \bar{p} \pi^+ \pi^-)$  and  $\mathcal{B}(B^0 \rightarrow (c\bar{c})K_S^0) \times \mathcal{B}((c\bar{c}) \rightarrow p \bar{p} \pi^+ \pi^-)$ . The decay  $\psi(2S) \rightarrow p \bar{p} \pi^+ \pi^-$  can proceed via the  $J/\psi$ :  $\psi(2S) \rightarrow J/\psi(\rightarrow p \bar{p}) \pi^+ \pi^-$ .

TABLE II. The resulting signal yields and phases in the default model. The errors are statistical only.

Parameter	$B^+ \rightarrow h_c K^+$	$B^0 \rightarrow h_c K_S^0$
$N_{\eta_c}$	$229 \pm 18$	$96 \pm 11$
$\varphi_{\eta_c}$	$-1.12 \pm 0.09$	$-1.47 \pm 0.14$
$N_{J/\psi}$	$345 \pm 19$	$128 \pm 12$
$N_{\chi_{c0}}$	$25.5 \pm 7.1$	$0.9 \pm 1.6$
$\varphi_{\chi_{c0}}$	$-1.51 \pm 0.23$	$-2.05 \pm 1.20$
$N_{\chi_{c1}}$	$34.5 \pm 8.8$	$21.2 \pm 6.1$
$N_{h_c}$	$32.6 \pm 8.0$	$3.1 \pm 3.8$
$N_{\chi_{c2}}$	$-1.6 \pm 6.3$	$11.6 \pm 5.5$
$N_{\eta_c(2S)}$	$86.1 \pm 11.9$	$24.0 \pm 6.8$
$\varphi_{\eta_c(2S)}$	$-1.41 \pm 0.14$	$-1.90 \pm 0.23$
$N_{\psi(2S)}$	$36.9 \pm 8.8$	$13.0 \pm 5.6$

To remove the events with a  $J/\psi$ , the  $\psi(2S)$  yield is taken from an alternative fit with an additional  $J/\psi$  veto defined as  $|M_{p\bar{p}} - m_{J/\psi}| > 50 \text{ MeV}/c^2$ .

The sources of the systematic uncertainty of the branching fractions include the model dependence (the same set of alternative fit models is used as in Sec. VB), overtraining

TABLE III. Significances of the charmonium states decaying to  $p \bar{p} \pi^+ \pi^-$  except the  $h_c$  in the default model.

State	$B^+ \rightarrow (c\bar{c})K^+$	$B^0 \rightarrow (c\bar{c})K_S^0$
$\eta_c$	$20.1\sigma$	$12.5\sigma$
$J/\psi$	$33.9\sigma$	$20.8\sigma$
$\chi_{c0}$	$6.0\sigma$	$0.6\sigma$
$\chi_{c1}$	$4.9\sigma$	$4.5\sigma$
$\chi_{c2}$	$0.3\sigma$	$2.5\sigma$
$\eta_c(2S)$	$12.3\sigma$	$5.9\sigma$
$\psi(2S)$	$5.0\sigma$	$2.8\sigma$

TABLE IV. Model dependence of the  $h_c$  and  $\eta_c(2S) \rightarrow p\bar{p}\pi^+\pi^-$  significance.

Model	$h_c$ significance		$\eta_c(2S) \rightarrow p\bar{p}\pi^+\pi^-$ significance	
	$B^+ \rightarrow h_c K^+$	$B^0 \rightarrow h_c K_S^0$	$B^+ \rightarrow h_c K^+$	$B^0 \rightarrow h_c K_S^0$
Default	$5.0\sigma$	$0.8\sigma$	$12.3\sigma$	$5.9\sigma$
Free masses and widths	$5.0\sigma$	$0.8\sigma$	$12.3\sigma$	$6.0\sigma$
Polynomial order ( $h_c \rightarrow \eta_c\gamma$ )	$4.8\sigma$	$0.8\sigma$	$12.3\sigma$	$5.9\sigma$
Polynomial order ( $h_c \rightarrow p\bar{p}\pi^+\pi^-$ background)	$5.0\sigma$	$0.8\sigma$	$12.2\sigma$	$5.9\sigma$
Polynomial order ( $h_c \rightarrow p\bar{p}\pi^+\pi^-$ signal)	$5.0\sigma$	$0.9\sigma$	$12.2\sigma$	$5.9\sigma$
Fitting range variation ( $h_c \rightarrow \eta_c\gamma$ )	$5.0\sigma$	$0.9\sigma$	$12.3\sigma$	$5.9\sigma$
Fitting range variation ( $h_c \rightarrow p\bar{p}\pi^+\pi^-$ )	$5.0\sigma$	$0.8\sigma$	$12.1\sigma$	$5.8\sigma$
Scaled resolution	$5.0\sigma$	$0.8\sigma$	$12.3\sigma$	$6.0\sigma$
Fraction of $h_c \rightarrow \eta_c\gamma$ and $h_c \rightarrow p\bar{p}\pi^+\pi^-$	$4.9\sigma$	$0.7\sigma$	$12.2\sigma$	$5.9\sigma$

(the difference between the efficiency in the training and testing samples), the error of the difference of the PID requirements efficiency between the data and MC, the difference of the MLP efficiency between the data and MC, tracking efficiency, number of  $\Upsilon(4S)$  events, the  $\eta_c \rightarrow K^+ K_S^0 \pi^-$  and  $\Upsilon(4S) \rightarrow B^+ B^-$  or  $\Upsilon(4S) \rightarrow B^0 \bar{B}^0$  branching fractions. All systematic error sources are listed in Table V for the channel  $B^+ \rightarrow (c\bar{c})K^+$  [ $B^+ \rightarrow h_c K^+$  for the  $h_c$  and  $B^+ \rightarrow (c\bar{c})(\rightarrow p\bar{p}\pi^+\pi^-)K^+$  for all other charmonium states] and in Table VI for the channel  $B^0 \rightarrow (c\bar{c})K_S^0$  [ $B^0 \rightarrow h_c K_S^0$  for the  $h_c$  and  $B^0 \rightarrow (c\bar{c})(\rightarrow p\bar{p}\pi^+\pi^-)K_S^0$  for all other charmonium states]. The errors of the tracking efficiency and the difference of the efficiency of the PID requirements depend on the multivariate-analysis channel in the case of the calculation of  $\mathcal{B}(B^+ \rightarrow h_c K^+)$  and  $\mathcal{B}(B^0 \rightarrow h_c K_S^0)$ ; the errors related to the MLP efficiency and overtraining are estimated separately for  $h_c \rightarrow \eta_c\gamma$  and  $h_c \rightarrow p\bar{p}\pi^+\pi^-$ . The values presented in Tables V and VI are weighted averages.

The difference of the PID requirements efficiency between the data and MC is estimated from several control samples, such as  $D^{*+} \rightarrow D^0(\rightarrow K^-\pi^+)\pi^+$  for  $K$  and  $\pi$ ,  $\Lambda \rightarrow p\pi^+$  for  $p$ ,  $\Lambda_c^+ \rightarrow \Lambda\pi^+$  for  $\Lambda$ ,

$D^0 \rightarrow K_S^0 \pi^+ \pi^-$  for  $K_S^0$ , and  $\tau^- \rightarrow \pi^-\pi^0\nu_\tau$  for  $\pi^0$ . The resulting overall efficiency ratio depends on the final state and the momenta of the decay products; thus, it is different for all branching fractions. For example, for the  $h_c$ , it is found to be  $(95.0 \pm 3.8)\%$  for the channel  $B^+ \rightarrow h_c K^+$  and  $(93.0 \pm 4.5)\%$  for the channel  $B^0 \rightarrow h_c K_S^0$ .

The error caused by the difference of the MLP efficiency between the data and MC is estimated for the channel  $h_c \rightarrow \eta_c\gamma$  using the decay mode  $B^0 \rightarrow \eta_c\pi^-K^+$ . This decay is reconstructed using selection criteria that are as similar as possible to the signal mode  $B^+ \rightarrow h_c K^+$ . The same MLP optimized for  $B^+ \rightarrow h_c K^+$  is applied to the control channel. Some MLP input variables used for the signal channel are undefined for  $B^0 \rightarrow \eta_c\pi^-K^+$ , for example, the number of  $\pi^0$  candidates that include the  $h_c$  daughter  $\gamma$  as one of their daughters. Such variables are held constant. The ratio of the number of signal candidates before and after the application of the MLP selection requirements is used to measure the difference of the MLP efficiency between the data and MC:

$$r_{\text{MLP}} = \frac{N(\text{all } \eta_c \text{ channels, with MLP cut})}{N(\eta_c \rightarrow K^+ K_S^0 \pi^-, \text{ no MLP cut})}. \quad (20)$$

TABLE V. Relative systematic uncertainties of the branching fractions for the channel  $B^+ \rightarrow (c\bar{c})K^+$  [ $B^+ \rightarrow h_c K^+$  for the  $h_c$  and  $B^+ \rightarrow (c\bar{c})(\rightarrow p\bar{p}\pi^+\pi^-)K^+$  for all other charmonium states].

Error source	$h_c$	$\eta_c$	$J/\psi$	$\chi_{c0}$	$\chi_{c1}$	$\chi_{c2}$	$\eta_c(2S)$	$\psi(2S)$
Model dependence	$(^{+9.09}_{-9.16})\%$	$(^{+3.45}_{-1.47})\%$	$(^{+1.96}_{-0.04})\%$	$(^{+4.59}_{-5.32})\%$	$(^{+7.10}_{-2.98})\%$	$(^{+21.13}_{-65.93})\%$	$(^{+1.75}_{-4.62})\%$	$(^{+1.74}_{-7.07})\%$
PID	3.99%	3.64%	3.62%	3.50%	3.50%	3.51%	3.52%	3.53%
Overtraining	0.41%	0.14%	0.14%	0.14%	0.14%	0.14%	0.14%	0.14%
Tracking	1.60%	1.75%	1.75%	1.75%	1.75%	1.75%	1.75%	1.75%
MLP efficiency	12.73%	0.25%	0.25%	0.25%	0.25%	0.25%	0.25%	0.25%
Number of $\pi^0$ candidates	11.60%	...	...	...	...	...	...	...
$\eta_c$ mass and width	0.99%	...	...	...	...	...	...	...
$h_c$ branching fraction	10.22%	...	...	...	...	...	...	...
$\mathcal{B}(\Upsilon(4S) \rightarrow B^+ B^-)$	1.17%	1.17%	1.17%	1.17%	1.17%	1.17%	1.17%	1.17%
Number of $\Upsilon(4S)$ events	1.37%	1.37%	1.37%	1.37%	1.37%	1.37%	1.37%	1.37%
Total	$(^{+22.51}_{-22.54})\%$	$(^{+5.62}_{-4.67})\%$	$(^{+4.83}_{-4.41})\%$	$(^{+6.30}_{-6.86})\%$	$(^{+8.31}_{-5.25})\%$	$(^{+21.57}_{-66.07})\%$	$(^{+4.68}_{-6.34})\%$	$(^{+4.68}_{-8.29})\%$

TABLE VI. Relative systematic uncertainties of the branching fractions for the channel  $B^0 \rightarrow (c\bar{c})K_S^0$  [ $B^0 \rightarrow h_c K_S^0$  for the  $h_c$  and  $B^0 \rightarrow (c\bar{c})(\rightarrow p\bar{p}\pi^+\pi^-)K_S^0$  for all other charmonium states].

Error source	$h_c$	$\eta_c$	$J/\psi$	$\chi_{c0}$	$\chi_{c1}$	$\chi_{c2}$	$\eta_c(2S)$	$\psi(2S)$
Model dependence	$^{+30.94}_{-24.05}\%$	$^{+4.91}_{-24.37}\%$	$^{+1.45}_{-1.80}\%$	$^{+134.01}_{-24.56}\%$	$^{+6.79}_{-2.28}\%$	$^{+8.63}_{-4.78}\%$	$^{+4.51}_{-5.36}\%$	$^{+10.89}_{-9.01}\%$
PID	4.86%	3.93%	3.93%	3.87%	3.87%	3.86%	3.83%	3.81%
Overtraining	0.15%	0.19%	0.19%	0.19%	0.19%	0.19%	0.19%	0.19%
Tracking	1.95%	2.10%	2.10%	2.10%	2.10%	2.10%	2.10%	2.10%
MLP efficiency	12.79%	0.25%	0.25%	0.25%	0.25%	0.25%	0.25%	0.25%
Number of $\pi^0$ candidates	11.66%	...	...	...	...	...	...	...
$\eta_c$ mass and width	0.96%	...	...	...	...	...	...	...
$h_c$ branching fraction	10.27%	...	...	...	...	...	...	...
$\mathcal{B}(\Upsilon(4S) \rightarrow B^0 \bar{B}^0)$	1.23%	1.23%	1.23%	1.23%	1.23%	1.23%	1.23%	1.23%
Number of $\Upsilon(4S)$ events	1.37%	1.37%	1.37%	1.37%	1.37%	1.37%	1.37%	1.37%
Total	$^{+37.33}_{-31.86}\%$	$^{+6.89}_{-24.85}\%$	$^{+5.04}_{-5.16}\%$	$^{+134.10}_{-25.02}\%$	$^{+8.31}_{-5.30}\%$	$^{+9.87}_{-6.76}\%$	$^{+6.56}_{-7.17}\%$	$^{+11.88}_{-10.18}\%$

Only the channel  $\eta_c \rightarrow K^+ K_S^0 \pi^-$  is used before the MLP selection because only this channel is sufficiently clean for the determination of the number of the signal events without the MLP selection. The ratio  $r_{\text{MLP}}$  is extracted from a simultaneous fit to the  $\eta_c$  mass distribution before and after the application of the MLP selection. The relative difference between the values of  $r_{\text{MLP}}$  in data and MC is found to be  $(14.4 \pm 9.0)\%$ . For conservative treatment, the statistical error is added in quadrature to the central value of the difference. The resulting systematic uncertainty caused by the MLP selection efficiency difference in data and MC is 17.0%.

The estimation of the MLP efficiency error for the hadronic channel  $p\bar{p}\pi^+\pi^-$  is done by performing the fit using the  $p\bar{p}\pi^+\pi^-$  data without the MLP selection and comparing the resulting branching fraction products  $\mathcal{B}(B^+ \rightarrow J/\psi K^+) \times \mathcal{B}(J/\psi \rightarrow p\bar{p}\pi^+\pi^-)$  with the results of the default procedure. Their relative difference is found to be 0.2%.

The same estimates of the MLP efficiency uncertainty for the channels  $h_c \rightarrow \eta_c \gamma$  and  $h_c \rightarrow p\bar{p}\pi^+\pi^-$  are used for both  $B^+ \rightarrow h_c K^+$  and  $B^0 \rightarrow h_c K_S^0$ . The final value of the MLP efficiency uncertainty is calculated as a weighted average of the errors for the two  $h_c$  decay channels. The result is slightly different for the channels  $B^+ \rightarrow h_c K^+$  and  $B^0 \rightarrow h_c K_S^0$  because of the difference in the relative number of the expected  $h_c \rightarrow \eta_c \gamma$  and  $h_c \rightarrow p\bar{p}\pi^+\pi^-$  events. The MLP efficiency error for the branching fraction products for the channels  $B^+ \rightarrow (c\bar{c})(\rightarrow p\bar{p}\pi^+\pi^-)K^+$  and  $B^0 \rightarrow (c\bar{c})(\rightarrow p\bar{p}\pi^+\pi^-)K_S^0$  is equal to the error for the channel  $p\bar{p}\pi^+\pi^-$ , since all charmonium states other than the  $h_c$  are reconstructed in this channel only.

The MLP efficiency uncertainty for the channel  $h_c \rightarrow \eta_c \gamma$  does not include the uncertainty caused by the difference between the data and MC in the distributions of the variables that are not defined for the channel  $B^0 \rightarrow \eta_c \pi^- K^+$ . There are four such variables: the  $\eta_c$  mass, the  $h_c$  helicity angle, and two numbers of  $\pi^0$  candidates that include the  $h_c$

daughter photon as one of their daughters. The distribution of the  $h_c$  helicity angle for the signal events is known precisely; thus, there is no additional uncertainty caused by the difference of its distribution in data and MC. The difference of the numbers of  $\pi^0$  candidates is taken into account by removing these variables from the neural network for the channel  $B^+ \rightarrow h_c K^+$ , performing an alternative optimization, and comparing the resulting  $h_c$  branching fractions in the channel  $h_c \rightarrow \eta_c \gamma$ . The relative difference is found to be 15.6%. The error due to the  $\eta_c$  mass distribution uncertainty is estimated by varying the  $\eta_c$  mass and width by  $\pm 1\sigma$  and reweighting the selected MC events in accordance with the relative difference between the modified and default  $\eta_c$  mass distributions. The largest resulting efficiency difference is considered as the systematic uncertainty related to the  $\eta_c$  mass distribution. This uncertainty is estimated to be 1.3% for both the  $B^+ \rightarrow h_c K^+$  and  $B^0 \rightarrow h_c K_S^0$  channels.

The ratio  $r_{\text{MLP}}$  used for determination of the MLP efficiency uncertainty includes the number of reconstructed events relatively to the number of events in  $h_c \rightarrow \eta_c (\rightarrow K^+ K_S^0 \pi^-) \gamma$ . The total expected number of events can be calculated as

$$2N_{\Upsilon(4S)} \left( \sum_i \epsilon_i \mathcal{B}_i \right) = 2N_{\Upsilon(4S)} \mathcal{B}_{h_c \rightarrow \eta_c \gamma} \left( \sum_i \epsilon_i \frac{\mathcal{B}_i}{\mathcal{B}_{h_c \rightarrow \eta_c \gamma}} \right), \quad (21)$$

where  $\epsilon_i$  and  $\mathcal{B}_i$  are the efficiency and branching fraction for  $i$ th channel, respectively. The last term in Eq. (21) is proportional to  $r_{\text{MLP}}$ . Consequently, for the channel  $h_c \rightarrow \eta_c \gamma$ , one needs to take into account only the error of  $\mathcal{B}(h_c \rightarrow \eta_c (\rightarrow K^+ K_S^0 \pi^-) \gamma)$ . Errors of the branching fractions of all other channels relative to  $h_c \rightarrow \eta_c (\rightarrow K^+ K_S^0 \pi^-) \gamma$  enter the MLP efficiency error.

The final  $h_c$  branching fraction error is calculated as a weighted average of the errors for the channels  $h_c \rightarrow \eta_c \gamma$

TABLE VII. Measured branching fractions and branching fraction products and their comparison with the current world-average values [3].

Branching fraction	Value or confidence interval (90% C.L.)	World-average value
$\mathcal{B}(B^+ \rightarrow h_c K^+)$	$(3.7_{-0.9}^{+1.0+0.8}) \times 10^{-5}$	$< 3.8 \times 10^{-5}$
$\mathcal{B}(B^+ \rightarrow \eta_c K^+) \times \mathcal{B}(\eta_c \rightarrow p \bar{p} \pi^+ \pi^-)$	$(39.4_{-3.9}^{+4.1+2.2}) \times 10^{-7}$	$(57.8 \pm 20.2) \times 10^{-7}$
$\mathcal{B}(B^+ \rightarrow J/\psi K^+) \times \mathcal{B}(J/\psi \rightarrow p \bar{p} \pi^+ \pi^-)$	$(56.4_{-3.2}^{+3.3+2.7}) \times 10^{-7}$	$(60.6 \pm 5.3) \times 10^{-7}$
$\mathcal{B}(B^+ \rightarrow \chi_{c0} K^+) \times \mathcal{B}(\chi_{c0} \rightarrow p \bar{p} \pi^+ \pi^-)$	$(3.7_{-1.0}^{+1.2+0.2}) \times 10^{-7}$	$(3.1 \pm 1.1) \times 10^{-7}$
$\mathcal{B}(B^+ \rightarrow \chi_{c1} K^+) \times \mathcal{B}(\chi_{c1} \rightarrow p \bar{p} \pi^+ \pi^-)$	$(4.7_{-1.2}^{+1.3+0.4}) \times 10^{-7}$	$(2.4 \pm 0.9) \times 10^{-7}$
$\mathcal{B}(B^+ \rightarrow \chi_{c2} K^+) \times \mathcal{B}(\chi_{c2} \rightarrow p \bar{p} \pi^+ \pi^-)$	$< 1.9 \times 10^{-7}$	$(0.15 \pm 0.06) \times 10^{-7}$
$\mathcal{B}(B^+ \rightarrow \eta_c(2S) K^+) \times \mathcal{B}(\eta_c(2S) \rightarrow p \bar{p} \pi^+ \pi^-)$	$(11.2_{-1.6}^{+1.8+0.5}) \times 10^{-7}$	Not seen
$\mathcal{B}(B^+ \rightarrow \psi(2S) K^+) \times \mathcal{B}(\psi(2S) \rightarrow p \bar{p} \pi^+ \pi^-)$	$[0.5, 3.5] \times 10^{-7}$	$(3.7 \pm 0.3) \times 10^{-7}$
$\mathcal{B}(B^0 \rightarrow h_c K_S^0)$	$< 1.4 \times 10^{-5}$	Not seen
$\mathcal{B}(B^0 \rightarrow \eta_c K_S^0) \times \mathcal{B}(\eta_c \rightarrow p \bar{p} \pi^+ \pi^-)$	$(19.0_{-2.9}^{+3.2+1.3}) \times 10^{-7}$	$(20.9 \pm 7.8) \times 10^{-7}$
$\mathcal{B}(B^0 \rightarrow J/\psi K_S^0) \times \mathcal{B}(J/\psi \rightarrow p \bar{p} \pi^+ \pi^-)$	$(24.3_{-2.2}^{+2.3+1.2}) \times 10^{-7}$	$(26.2 \pm 2.4) \times 10^{-7}$
$\mathcal{B}(B^0 \rightarrow \chi_{c0} K_S^0) \times \mathcal{B}(\chi_{c0} \rightarrow p \bar{p} \pi^+ \pi^-)$	$< 1.3 \times 10^{-7}$	$(1.5 \pm 0.6) \times 10^{-7}$
$\mathcal{B}(B^0 \rightarrow \chi_{c1} K_S^0) \times \mathcal{B}(\chi_{c1} \rightarrow p \bar{p} \pi^+ \pi^-)$	$(3.7_{-1.0}^{+1.2+0.3}) \times 10^{-7}$	$(1.0 \pm 0.4) \times 10^{-7}$
$\mathcal{B}(B^0 \rightarrow \chi_{c2} K_S^0) \times \mathcal{B}(\chi_{c2} \rightarrow p \bar{p} \pi^+ \pi^-)$	$[0.7, 3.8] \times 10^{-7}$	Not seen
$\mathcal{B}(B^0 \rightarrow \eta_c(2S) K_S^0) \times \mathcal{B}(\eta_c(2S) \rightarrow p \bar{p} \pi^+ \pi^-)$	$(4.2_{-1.2}^{+1.4+0.3}) \times 10^{-7}$	Not seen
$\mathcal{B}(B^0 \rightarrow \psi(2S) K_S^0) \times \mathcal{B}(\psi(2S) \rightarrow p \bar{p} \pi^+ \pi^-)$	$< 1.9 \times 10^{-7}$	$(1.7 \pm 0.2) \times 10^{-7}$

and  $h_c \rightarrow p \bar{p} \pi^+ \pi^-$ . Since branching fraction products are measured for all other charmonium states, they do not have a similar systematic error source.

The resulting branching fractions with both statistical and systematic errors are listed in Table VII. For insignificant decays or decay chains, the confidence intervals are calculated in the frequentist approach [24] using an asymmetric Gaussian as the branching-fraction probability density function and the measured central value and errors as its parameters. This is the first measurement of  $\mathcal{B}(B^0 \rightarrow h_c K_S^0)$ . Also, the branching fraction products  $\mathcal{B}(B^+ \rightarrow (c\bar{c})K^+) \times \mathcal{B}((c\bar{c}) \rightarrow p \bar{p} \pi^+ \pi^-)$  and  $\mathcal{B}(B^0 \rightarrow (c\bar{c})K_S^0) \times \mathcal{B}((c\bar{c}) \rightarrow p \bar{p} \pi^+ \pi^-)$  are measured directly in  $B$  decays for the first time. The current world-average values of the same branching fractions [3] are also presented in Table VII for comparison if they are known. The values of the branching fraction products are calculated by multiplying the individual branching fractions listed in Ref. [3] assuming uncorrelated errors. The measured branching fractions are consistent with the world averages; the largest deviation is observed for the branching-fraction product  $\mathcal{B}(B^0 \rightarrow \chi_{c1} K_S^0) \times \mathcal{B}(\chi_{c1} \rightarrow p \bar{p} \pi^+ \pi^-)$ , which differs from the world-average value by  $2.4\sigma$  taking its error into account. The results for  $\mathcal{B}(B^+ \rightarrow \eta_c K^+) \times \mathcal{B}(\eta_c \rightarrow p \bar{p} \pi^+ \pi^-)$  and  $\mathcal{B}(B^0 \rightarrow \eta_c K_S^0) \times \mathcal{B}(\eta_c \rightarrow p \bar{p} \pi^+ \pi^-)$  have a better precision than the world-average values.

## VI. CONCLUSIONS

A search for the decays  $B^+ \rightarrow h_c K^+$  and  $B^0 \rightarrow h_c K_S^0$  has been performed. Evidence for the decay  $B^+ \rightarrow h_c K^+$  is

found; its significance is  $4.8\sigma$ . No evidence is found for  $B^0 \rightarrow h_c K_S^0$ . The branching fraction of  $B^+ \rightarrow h_c K^+$  is measured to be  $(3.7_{-0.9}^{+1.0+0.8}) \times 10^{-5}$ ; the upper limit for the  $B^0 \rightarrow h_c K_S^0$  branching fraction is  $1.4 \times 10^{-5}$  at 90% C.L. The measured value of  $\mathcal{B}(B^+ \rightarrow h_c K^+)$  is consistent with the existing upper limit of  $3.8 \times 10^{-5}$  (90% C.L.) obtained in the previous Belle analysis [4] and supersedes it. The resulting branching fraction  $\mathcal{B}(B^+ \rightarrow h_c K^+)$  agrees with the existing theoretical predictions [8–10]. In addition, a study of the  $p \bar{p} \pi^+ \pi^-$  invariant mass distribution in the channel  $B^+ \rightarrow (p \bar{p} \pi^+ \pi^-) K^+$  results in the first observation of the decay  $\eta_c(2S) \rightarrow p \bar{p} \pi^+ \pi^-$  with  $12.1\sigma$  significance.

## ACKNOWLEDGMENTS

We thank the KEKB group for the excellent operation of the accelerator; the KEK cryogenics group for the efficient operation of the solenoid; the KEK computer group and the Pacific Northwest National Laboratory (PNNL) Environmental Molecular Sciences Laboratory (EMSL) computing group for strong computing support; and the National Institute of Informatics, and Science Information NETwork 5 (SINET5) for valuable 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 including Grants No. DP180102629, No. DP170102389, No. DP170102204, No. DP150103061, and No. FT130100303; Austrian Science Fund (FWF); the

National Natural Science Foundation of China under Contracts No. 11435013, No. 11475187, No. 11521505, No. 11575017, No. 11675166, and No. 11705209; Key Research Program of Frontier Sciences, Chinese Academy of Sciences (CAS), Grant No. QYZDJ-SSW-SLH011; the CAS Center for Excellence in Particle Physics (CCEPP); the Shanghai Pujiang Program under Grant No. 18PJ1401000; the Ministry of Education, Youth and Sports of the Czech Republic under Contract No. LTT17020; the Carl Zeiss Foundation, the Deutsche Forschungsgemeinschaft, the Excellence Cluster Universe, and the VolkswagenStiftung; the Department of Science and Technology of India; the Istituto Nazionale di Fisica Nucleare of Italy; National Research Foundation (NRF) of Korea Grants No. 2015H1 A2A1033649, No. 2016R1D1A1B01010135, No. 2016K1 A3A7A09005 603, No. 2016R1D1A1B02012900,

No. 2018R1A2B3003 643, No. 2018R1A6A1A06024970, and No. 2018R1D1 A1B07047294; Radiation Science Research Institute, Foreign Large-size Research Facility Application Supporting project, the Global Science Experimental Data Hub Center of the Korea Institute of Science and Technology Information and Korea Research Environment Open Network/Global Ring Network for Advanced Application Development; the Polish Ministry of Science and Higher Education and the National Science Center; the Russian Foundation for Basic Research Grant No. 18-32-00277; the Slovenian Research Agency; Ikerbasque, Basque Foundation for Science, Spain; the Swiss National Science Foundation; the Ministry of Education and the Ministry of Science and Technology of Taiwan; and the United States Department of Energy and the National Science Foundation.

- 
- [1] M. Bauer, B. Stech, and M. Wirbel, *Z. Phys. C* **34**, 103 (1987).
- [2] M. Suzuki, *Phys. Rev. D* **66**, 037503 (2002).
- [3] M. Tanabashi *et al.* (Particle Data Group), *Phys. Rev. D* **98**, 030001 (2018).
- [4] F. Fang *et al.* (Belle Collaboration), *Phys. Rev. D* **74**, 012007 (2006).
- [5] B. Aubert *et al.* (BABAR Collaboration), *Phys. Rev. D* **78**, 012006 (2008).
- [6] R. Aaij *et al.* (LHCb Collaboration), *Eur. Phys. J. C* **73**, 2462 (2013).
- [7] R. Aaij *et al.* (LHCb Collaboration), *Phys. Lett. B* **769**, 305 (2017).
- [8] C. Meng, Y. J. Gao, and K. T. Chao, arXiv:hep-ph/0607221.
- [9] X. Q. Li, X. Liu, and Y. M. Wang, *Phys. Rev. D* **74**, 114029 (2006).
- [10] M. Beneke and L. Vernazza, *Nucl. Phys.* **B811**, 155 (2009).
- [11] S. Kurokawa and E. Kikutani, *Nucl. Instrum. Methods Phys. Res., Sect. A* **499**, 1 (2003), and other papers included in this volume; T. Abe *et al.*, *Prog. Theor. Exp. Phys.* **2013**, 03A001 (2013), and references therein.
- [12] M. Ablikim *et al.* (BESIII Collaboration), *Phys. Rev. D* **99**, 072008 (2019).
- [13] A. Abashian *et al.* (Belle Collaboration), *Nucl. Instrum. Methods Phys. Res., Sect. A* **479**, 117 (2002); also see detector section in also see Sec. 2 in J. Brodzicka *et al.*, *Prog. Theor. Exp. Phys.* **2012**, 04D001 (2012).
- [14] Z. Natkaniec *et al.* (Belle SVD2 Group), *Nucl. Instrum. Methods Phys. Res., Sect. A* **560**, 1 (2006).
- [15] R. Brun *et al.*, GEANT 3.21, CERN Report No. CERN DD/EE/84-1, 1984.
- [16] D. J. Lange, *Nucl. Instrum. Methods Phys. Res., Sect. A* **462**, 152 (2001).
- [17] M. Gelb *et al.*, *Comput. Soft. Big Sci.* **2**, 9 (2018).
- [18] K. Hanagaki, H. Kakuno, H. Ikeda, T. Iijima, and T. Tsukamoto, *Nucl. Instrum. Methods Phys. Res., Sect. A* **485**, 490 (2002).
- [19] G. C. Fox and S. Wolfram, *Phys. Rev. Lett.* **41**, 1581 (1978).
- [20] J. P. Lees *et al.* (BABAR Collaboration), *Phys. Rev. D* **89**, 112004 (2014).
- [21] T. Skwarnicki, Ph.D. thesis, Institute for Nuclear Physics, 1986; DESY Internal Report No. DESY F31-86-02, 1986.
- [22] H. Voss, A. Hocker, J. Stelzer, and F. Tegenfeldt, *Proc. Sci. ACAT2007* (**2007**) 040.
- [23] G. Punzi, arXiv:physics/0308063.
- [24] G. J. Feldman and R. D. Cousins, *Phys. Rev. D* **57**, 3873 (1998).