PROCEEDINGS OF SCIENCE

Search for $B^+ \to K^+ \nu \bar{\nu}$ at Belle II

Roberta Volpe on behalf of the Belle II Collaboration[a,*]

[a] *Perugia University and INFN,*
Via Pascoli, 06123 Perugia, Italy

E-mail: roberta.volpe@pg.infn.it

We use a 362 fb^{-1} sample of e^+e^- collisions at $\Upsilon(4S)$ resonance collected with Belle II detector at SuperKEKB collider to search for the rare decay $B^+ \to K^+ \nu \bar{\nu}$. The main strategy consists in exploiting the inclusive properties of the other B to suppress the background. Another analysis is based on a conventional hadronic reconstruction of the accompanying B meson and is used to corroborate the first strategy. Both the procedures have been validated with several control samples. A maximum likelihood fit is used to extract the branching ratio, which results in $[2.7 \pm 0.5(\text{stat}) \pm 0.5(\text{syst})] \times 10^{-5}$ for the main analysis and $[1.1^{+0.9}_{-0.8}(\text{stat})^{+0.8}_{-0.5}(\text{syst})] \times 10^{-5}$ for the support analysis. The combination of the two analyses gives a branching fraction of $[2.3 \pm 0.5(\text{stat})^{+0.5}_{-0.4}(\text{syst})] \times 10^{-5}$ which corresponds to the first evidence of the decay with 3.5 standard deviations and 2.7 standard deviations above the standard model expectation.

16th International Conference on Heavy Quarks and Leptons (HQL2023)
28 November-2 December 2023
TIFR, Mumbai, Maharashtra, India

*Speaker

PoS(HQL2023)049

1. Introduction

The decay $B^+ \to K^+\nu\bar{\nu}$ occurs through the flavor-changing neutral current transition $b \to s\nu\bar{\nu}$. That is suppressed in the Standard Model (SM) because of the Glashow–Iliopoulos–Maiani mechanism, which makes it a rare process. The dominating contributions to the decay are the Feynman diagrams shown in figure 1

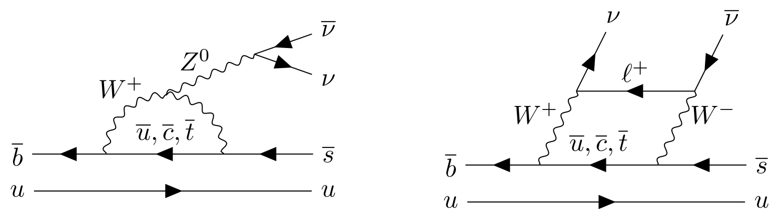

Figure 1: Dominating Feynman diagrams contributing to the $b \to s\nu\bar{\nu}$ transition.

The SM prediction for the branching fraction is $\mathcal{B}(B^+ \to K^+\nu\bar{\nu}) = (5.58 \pm 0.37) \times 10^{-6}$ [1]. The good theoretical precision, due to small hadronic uncertainties, makes this process an ideal environment to search for new physics. Indeed the branching fraction can be enhanced in models that predict high mass non-SM particles, as Leptoquarks [2]. Furthermore, new low-mass undetectable exotic particles (dark matter candidates or mediators of a dark sector) could be produced together with the kaon giving rise to a two-body or three-body decay with missing energy [3] [4]. Before the analysis described in this document, no evidence for a signal has been found and the experimental upper limit on the branching fraction was 1.6×10^{-5} at 90% of confidence level (CL) [5]. The main challenge for this search is the presence of two neutrinos, which precludes the full reconstruction of the event. In this work the signal B meson is produced in the $e^+e^- \to \Upsilon(4S) \to B^+B^-$ process and the accompanying B is used to obtain information on the event kinematics. Two analysis strategies have been exploited, the most performant one is the inclusive tagging analysis method (ITA), exploiting inclusive properties from the B-meson pair-produced along with the signal B and representing the main analysis. A support analysis, employing the well-established hadronic tagging analysis method (HTA), has been used as well. Both the strategies have been applied to the full dataset of e^+e^- collisions produced from 2019 to 2022 by the SuperKEKB collider [6] and collected by the Belle II experiment [7]. Data produced with a center-of-mass (c.m.) energy equal to the mass of the $Y(4S)$, on-resonance data, correspond to an integrated luminosity of 362 fb^{-1}. The Belle II detector is made of several sub-detectors arranged in a cylindrical structure, surrounded by a superconducting solenoid providing a 1.5 T magnetic field parallel to the cylindrical main axis. Starting from the inside, the tracking system is composed of: a silicon pixel detector, a double sided silicon strip detector, a central drift chamber (CDC). A time-of-propagation counter and an aerogel ring-imaging Cherenkov counter provide the identification of charged particles (PID) and an electromagnetic calorimeter (ECL) reconstruct photons and other neutral particles. In the flux return of the solenoid a system to identify muons and K_L mesons is installed. A detailed description of the $B^+ \to K^+\nu\bar{\nu}$ analysis can be found in [8], this document summarizes the main features. A brief description of the reconstruction of the event is given in section 2, the background suppression approach is described in section 3 and the validation of efficiency and background estimation is summarized in 4. The signal extraction settings and results are given in sections 5 and 6 respectively.

2. Event reconstruction and basic selection

The trigger selection, based on number of tracks in the CDC or energy deposits in the ECL, has an efficiency close to 100%. The two analysis strategies differ for the tagging method, while the particle reconstruction is kept as similar as possible, the event reconstruction for the two is summarized in the following.
For the ITA, the event reconstruction starts with the reconstruction of charged and neutral particles. The charged particles are required to have a transverse momentum $p_T > 0.1$ GeV, to be within the CDC acceptance and (for the ones not coming from a K_S candidate) to be close to the interaction point, by requiring minimum longitudinal and transverse distances (impact parameters) from the average interaction point of $|d_z| < 3.0$ cm and $d_r < 0.5$ cm, respectively. The K_S candidates are reconstructed starting from two opposite sign tracks compatible to be pions originating from a common vertex. The ECL deposits with $E > 0.1$ GeV, in the CDC acceptance and not matched with tracks are considered as photons. In order to reject misreconstructed particles and cosmic muons, each particle is required to have an energy $E < 5.5$ GeV. Particle identification likelihoods, based on PID detectors and other detector information, are employed to identify the charged kaons. The chosen requirement gives a 68% of efficiency for signal kaons and 1.2% of probability to identify a pion as a kaon. Conditions are imposed on the event as follows. The total momentum of all reconstructed particles is used to compute the missing momentum as its complement and the polar angle of the missing momentum, θ, must be $17° < \theta < 160°$. The number of tracks in the event, N_{trk}, is required to be $4 < N_{trk} < 10$, to reduce high multiplicity and low multiplicity background contributions and the total energy of the event is required to be $E > 4$ GeV. One of the most important quantity to select the signal kaon in an event is the mass squared of the neutrino pair, which, in the ITA is computed as: $q^2_{rec} = s^2 + M^2_K - \sqrt{s}E^*_K$ where M_K is the known mass of K^+ mesons and E^*_K is the reconstructed energy of the kaon in the c.m. system, assuming the signal B at rest in the c.m. frame. The candidate with the lowest q^2_{rec} is chosen. The Rest of the Event (ROE) is composed of all the charged particles, photons and K_S not associated to the signal kaon.
The HTA starts with the reconstruction of a B meson (B_{tag}), through the Full Event Interpretation (FEI) [9]. Requirements on the output of the FEI are used to reduce the background. In addition the B_{tag} and signal kaon are required to have opposite charges and events with $N_{trk} > 12$ are rejected. The kaon identification and the restrictions on missing momentum are the same as in the ITA. The number of tracks, coming form the impact point and with at least 20 hits in the CDC, not associated with the B_{tag} nor with the signal kaon, is required to be zero, all the other tracks are named *extra tracks*. The photons not associated with B_{tag} nor with the signal kaon are named extra photons. Moreover events are rejected if a K^0_S-meson, π^0-meson, or Λ-baryon candidate is reconstructed from the extra tracks and photons.
For both the strategies, control samples from data are used to test the simulation of the detector response and, when a difference with respect to data is found, correction factors are introduced with corresponding systematic uncertainties. Here only the most important ones are mentioned. The photon energy is corrected, moreover an additional correction is needed due to a contribution of clusters mimicking photons but arising from neutral hadrons, charged hadrons and beam background particles. For the ITA a multiplicative hadronic energy correction is inferred empirically using data, while for the HTA a correction to the number of the selected extra photons is applied. The probability

to have incorrect identification of charged particles is different in data and MC. Correction factors and their uncertainties are applied to the simulation as functions of the particle's charge, momentum, and polar angle. The K_L^0 are reconstructed using only the ECL and the modeling of its response is studied by using $e^+e^- \to \phi(\to K_L^0 K_S^0)\gamma$ events. The outcome is that the simulation overestimates the efficiency by 17% in the ITA. Corrections are applied to the ITA and a corresponding systematic uncertainty of 100% is applied both for the ITA and the HTA.

3. Background suppression

Boosted Decision Tree (BDT) algorithms are built with several input variables: general event-shape variables, variables characterizing the kaon candidate, the kinematic properties of the ROE (for the ITA) and extra tracks and extra photons (for the HTA), the B_{tag} variables for the HTA. Furthermore for kaons that are identified as coming from D_0 and D^+ meson decays, variables describing the fit quality and kinematic properties of the resulting candidates are also included. The ITA uses a first BDT (BDT_1) as an event filter and a second classifier (BDT_2) for the final event selection. The most discriminant input variable of BDT_2 is the cosine of the angle between the momentum of the signal-kaon candidate and the thrust axis of the ROE computed in the c.m. frame. The HTA uses a single classifier, $BDTh$, and for it the most discriminant variable is the sum of extra-photon energy deposits in ECL, named E_{extra}. The multivariate classifiers are trained with simulated samples for signal and background. The output of the BDT_2 for the ITA and $BDTh$ for the HTA, are mapped into variables whose distributions are uniform for simulated signal events: $\eta(BDT_2)$ and $\eta(BDTh)$ respectively. For the ITA the selections $BDT_1 > 0.9$ and $\eta(BDT_2) > 0.92$ define the signal region, which is further split in 3×4 bins of $\eta(BDT_2)$ and q^2_{rec}. For the HTA the signal region is defined as $\eta(BDTh) > 0.4$ and is divided in 6 bins in $\eta(BDTh)$. After the full selection, for the ITA the signal efficiency is 8% with an expected purity of 0.8%, while for the HTA the signal efficiency is 0.4% with an expected purity of 3.7%.

4. Validation of the analysis

The optimization of the strategy and the training of the multivariate classifiers have been performed using simulated samples of signal and background. The modeling of the signal efficiency and the background estimation have been thoroughly validated using control samples data. When needed, corrections are applied and further validated. In this document only a few examples of validation approaches are given and only for the ITA. Similar methods have been used for the HTA validation.

4.1 Signal efficiency validation in the ITA

The agreement of the signal efficiency in data and simulation is validated with a sample selected as $B^+ \to K^+ J/\psi(\to \mu^+\mu^-)$. For each event the muon pair is disregarded and the kaon is replaced by the kaon simulated in the signal events, to reflect the three-body topology of the signal signature. This signal-embedding procedure is performed for both data and $B^+ \to K^+ J/\psi$ simulation. Figure 2 summarizes the results in the distributions of BDT_1 and BDT_2. Good agreement is observed both before and after the signal embedding, resulting in a ratio of the efficiencies in data and simulation of 1.00 ± 0.03. The validation of the kaon identification is computed separately and is described in [8].

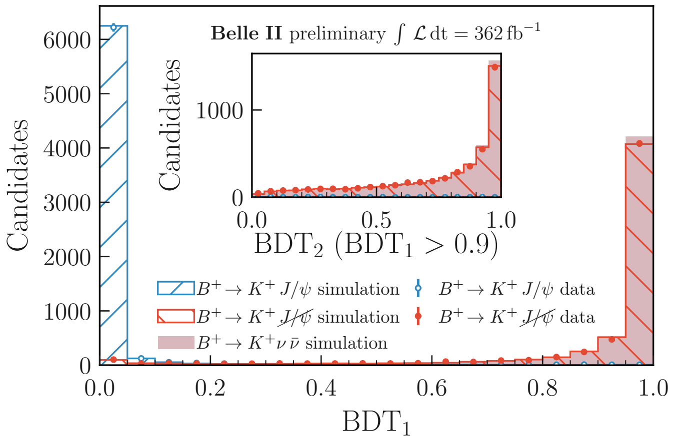

Figure 2: Distribution of the classifier output BDT_1 and BDT_2 for $BDT_1 > 0.9$. The simulation histograms are scaled to the total number of $B^+ \to K^+ J/\psi$ events selected in the data.

4.2 Background estimation and its validation in the ITA

For the ITA the background is composed of continuum events $q\bar{q}$ for the 40% and B-meson decay events for the 60%. The modeling of $q\bar{q}$ contribution is validated using the off-resonance data, which is a sample obtained with e^+e^- collisions at a c.m. energy 60 MeV below the mass of the $\Upsilon(4S)$. The moderate disagreement in shape is corrected with the procedure described in [10]. Among the B-meson decay contributions, the charged B^+B^- are the most important ones and can be separated in: (*i*) B-mesons hadronic decays involving D mesons and a kaon (38%); (*ii*) other hadronic B decays (14%); (*iii*) semi-leptonic B decays to charm-mesons that decay in turn to kaons (47%); (*iv*) leptonic decays (1%). Processes involving K_L mesons are particularly relevant because they are poorly known, in addition the detector response can be mis-modeled and K_L can fake missing energy. The decays of the kind $B \to D \to K^0_L X$ are evaluated by using a control sample selected with a pion identification instead of the kaon identification. An excess of data over simulation is found and in order to evaluate it, the sample is separated in three contributions: the B decays involving $B \to D \to K^0_L X$ decays, all the other B decays and the $q\bar{q}$. A fit to the q^2_{rec} distribution is performed with the three fractions of the contributions as parameters and the estimated normalization factor for $B \to D \to K^0_L X$ is found to be an increase in rate of (30 ± 2)%. Figure 3 (left) shows the post-fit data/simulation comparison for the q^2_{rec} distribution. Applying the same normalization factor to other variables, a good agreement is found, in particular for the other main variable of the analysis $\eta(BDT_2)$, shown in 3(right).

The charmless hadronic B decays with K^0_L mesons are scrutinised as well. Three-body $B^+ \to K^+K^0_LK^0_L$ decays are modeled using Dalitz spectra of $B^+ \to K^+K^0_SK^0_S$ decays measured by BaBar [11] and assuming equal probabilities for the two decays. The sPlot technique is applied to determine the distribution of the invariant $K^0_SK^0_S$ mass after the background subtraction. The good data/simulation agreement is visible in figure 4 (left). Similar strategies are used to estimate

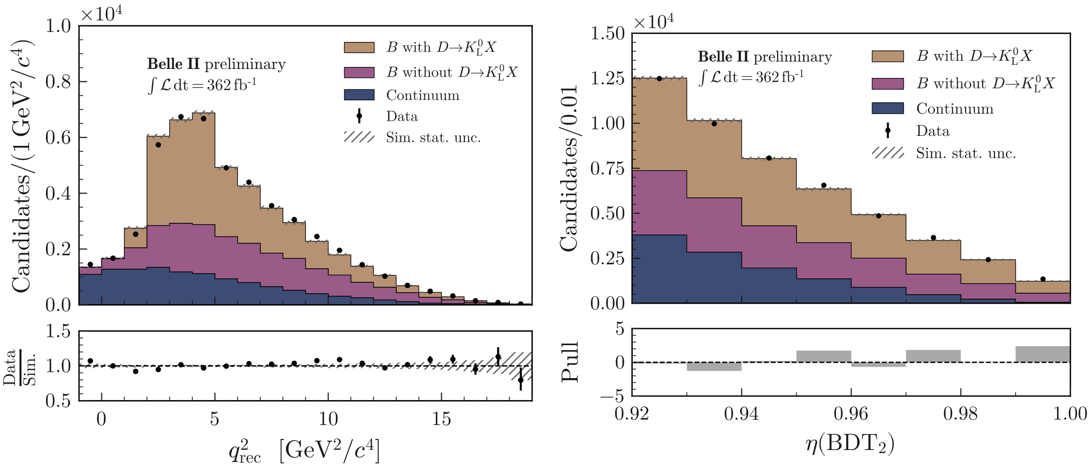

Figure 3: Post-fit data and simulation distributions of the pion enriched sample for q^2_{rec} (left) and $\eta(BDT_2)$(right).

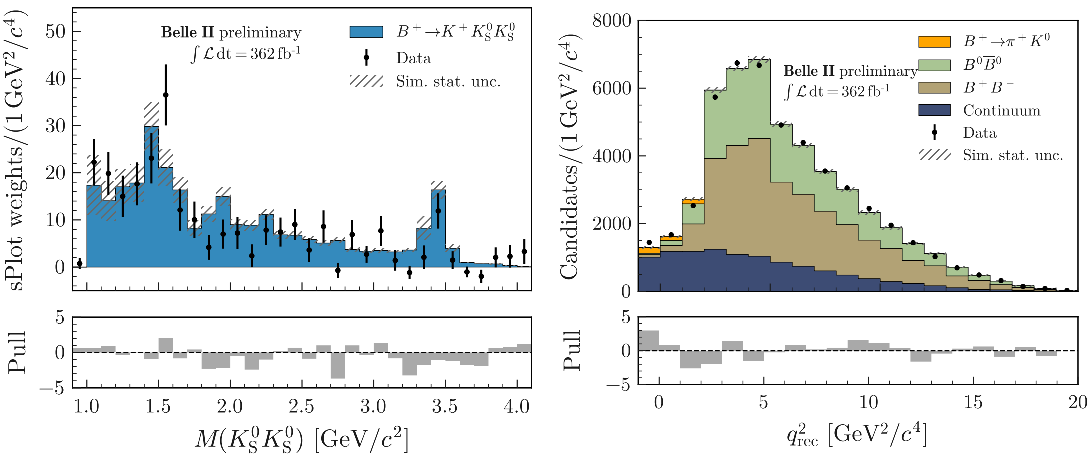

Figure 4: Left: Distribution of the invariant $K^0_S K^0_S$ mass in background subtracted data. The simulated distribution is normalized to the number of BB events. The pull distribution is shown in the bottom panel. Right: q^2 distribution for data and MC obtained with the fit results to determine the branching fraction of the $B^+ \to \pi^+ K^0$ decay.

and validate the background contributions from $B^+ \to K^+ K^0_L K^0_S$ and $B^+ \to K^+ n\bar{n}$. Validation procedures applied to other background contributions both for the ITA and the HTA are described in [8].

5. Signal extraction

Binned maximum likelihood fits are performed on data counts in the signal regions to extract the signal yield, both for the ITA and the HTA. For the ITA both on-resonance and off-resonance data are used, each one divided in 3×4 bins of $\eta(BDT_2)$ and q^2_{rec} and the yields of the seven individual background categories (B^+B^-, $B^0\bar{B}^0$, $c\bar{c}$, $s\bar{s}$, $u\bar{u}$, $d\bar{d}$, $\tau^+\tau^-$). The HTA uses only on-resonance data and the signal region is divided in six bins of $\eta(BDTh)$ and the background categories considered are BB (B^+B^-, $B^0\bar{B}^0$), $c\bar{c}$ and light quark pairs, while $\tau^+\tau^-$ can be neglected. The parameter of interest is μ, the signal branching fraction relative to its SM expectation, which is taken as the value 4.97×10^{-6}, excluding the long distance contribution from τ decays [1]. The systematic uncertainties are included in the likelihood as nuisance parameters. The most important ones are: the normalization of the $B\bar{B}$ background, the limited size of the simulated samples both for the ITA and the HTA. For the ITA another important contribution comes from the poor knowledge of some background contributions: $B^+ \to K^+K^0_L K^0_L$, $B \to D^{**}$. For the HTA another main contribution comes from the modeling of the extra photon multiplicity. Before extracting the result, an additional check for the ITA method has been performed by measuring the branching fraction of the $B^+ \to \pi^+K^0$ decay. Similar signal extraction settings to the nominal analysis are used, the main differences are: the pion identification is used instead of kaon identification, the only on-resonance data are used, not all the systematic sources are considered. The measured value is $\mathcal{B}(B^+ \to \pi^+K^0) = (2.5 \pm 0.5) \times 10^{-5}$, consistent with the PDG value. The post fit distribution of the q^2_{rec} is shown in figure 4(right).

6. Results

The results for ITA of the the simultaneous fit to off-resonance and on-resonance data, together with the observed yields are illustrated in Figure 5. The signal strength is determined to be

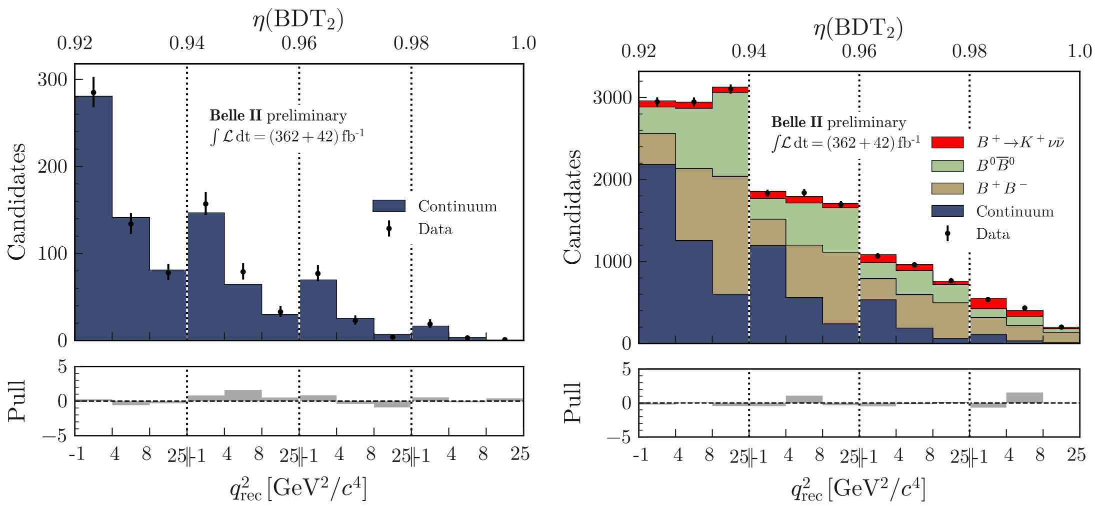

Figure 5: Observed yields and fit results in bins of the $\eta(BDT_2) \times q^2_{rec}$ space for the off-resonance (left) and the on-resonance (right) samples.

$\mu = 5.4 \pm 1.0(\text{stat}) \pm 1.1(\text{syst}) = 5.4 \pm 1.5$, corresponding to $\mathcal{B}(B^+ \to K^+\nu\bar{\nu}) = [2.7 \pm 0.5\ (\text{stat}) \pm 0.5\ (\text{syst})] \times 10^{-5}$. By evaluating the profile likelihood for several μ values, we found the significance of the observed excess with respect to the background-only hypothesis, which is 3.5 standard deviations (σ) and the significance of the observed signal with respect to the SM expectation, which is 2.9σ. Figure 6 shows the post-fit distributions for $\eta(BDT_2)$ and q^2_{rec} with a different binning with respect to the one used for the fit. The post-fit distributions are checked also considering only events on the most signal-rich region, $\eta(BDT_2) > 0.98$. The distributions of $\eta(BDT_2) > 0.98$ and q^2_{rec} for these events are shown in Figure 7.

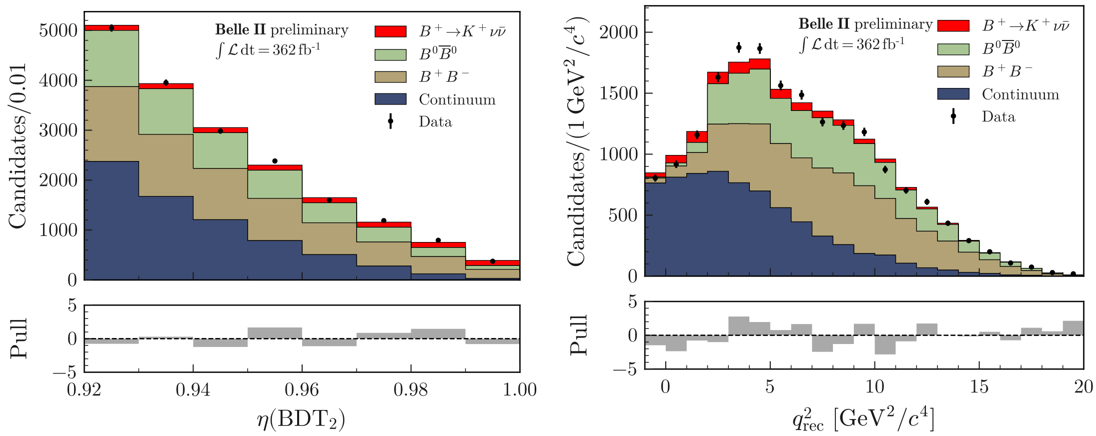

Figure 6: Observed yields and post-fit simulation data for the ITA, for $\eta(BDT_2)$ (left) and q^2_{rec} (right).

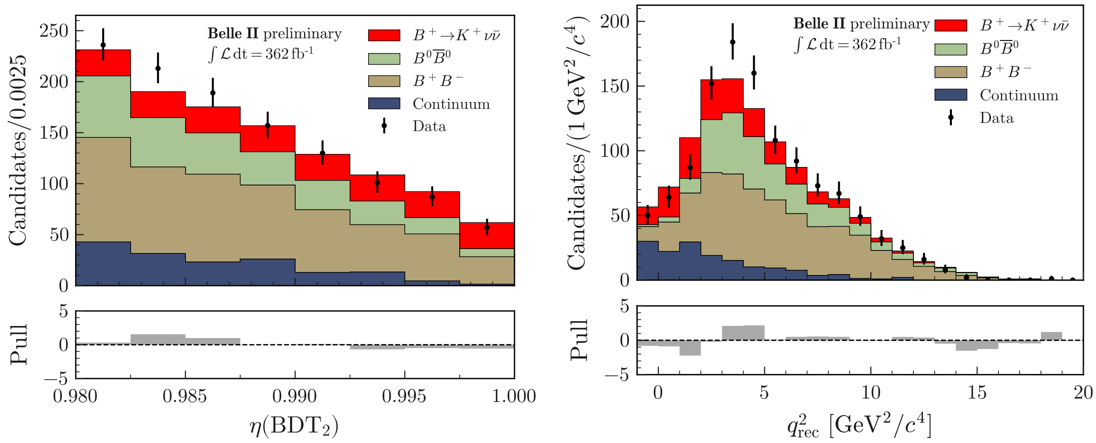

Figure 7: Observed yields and post-fit simulation data for the ITA, after requiring $\eta(BDT_2) > 2$, for $\eta(BDT_2)$ (left) and q^2_{rec} (right).

The post-fit distribution of the fit variable $\eta(BDTh)$ for the HTA is shown in figure 8(left) along with the post-fit distribution of q^2. The fit results in a $\mu = 2.2^{+1.8}_{-1.7}$ (stat)$^{+1.6}_{-1.1}$ (syst) corresponding to $\mathcal{B}(B^+ \to K^+ \nu \bar{\nu}) = [1.1^{+0.9}_{-0.8}(\text{stat})^{+0.8}_{-0.5}\ (\text{syst})] \times 10^{-5}$. This result is compatible with the background-only hypothesis at 1.1σ and in agreement with the SM at 0.6σ.

Several consistency checks are performed to scrutiny the validity of the analysis: simulation and data events are divided into approximately same-size statistically independent samples based on different criteria. Quite good compatibility is observed between the split samples for the ITA and the HTA.

The results of the two analyses are compatible, with a difference in the signal strength of 1.2σ. Furthermore the overlap of the data sample is small, only 2% of the full ITA selected sample. Therefore, after the removal of the common events from the ITA sample, a combination of the

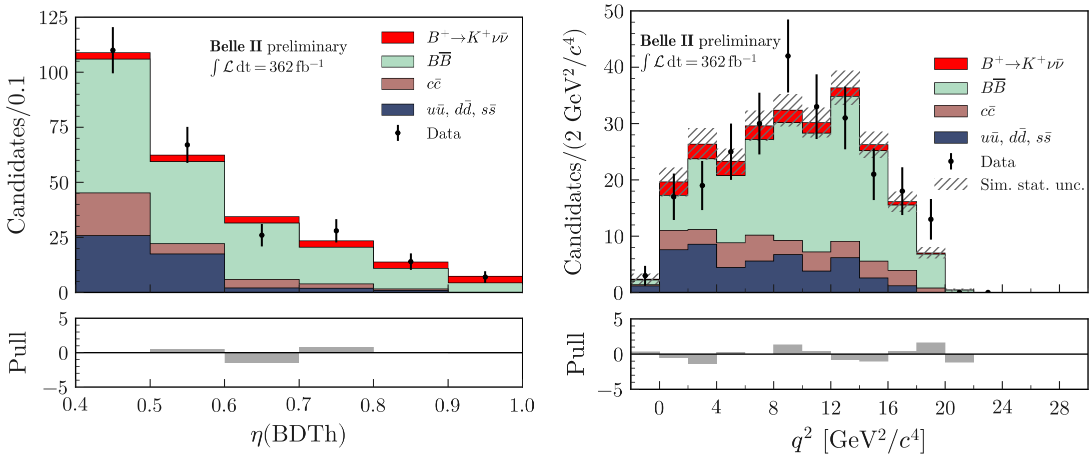

Figure 8: Observed yields and post-fit simulation data for the HTA, for $\eta(BDTh)$ (left) and q^2_{rec} (right)

two analyses is performed with a profile likelihood fit, incorporating correlations between common systematic uncertainties. The combined result for the signal strength is $\mu = 4.6 \pm 1.0(\text{stat}) \pm 0.9(\text{syst}) = 4.6 \pm 1.3$ corresponding to :

$$\mathcal{B}(B^+ \to K^+\nu\bar{\nu}) = [2.3 \pm 0.5(\text{stat})^{+0.5}_{-0.4}\,(\text{syst})] \times 10^{-5} = (2.3 \pm 0.7) \times 10^{-5} \tag{1}$$

This results in a significance with respect to the background-only hypothesis of 3.5σ and in a 2.7σ above the SM expectation.

Figure 9(left) show the values of the quantity $-2 \log L$, with L the likelihood, as a function of μ, for the ITA, the HTA and combined analyses. The value for each scan point is determined by fitting the data, where all parameters but μ are varied.

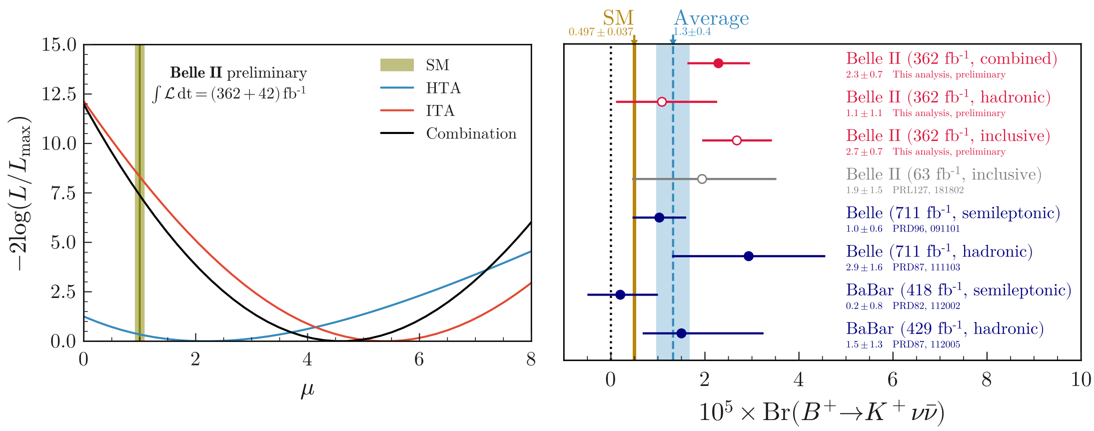

Figure 9: $-2 \log L$ for several values of μ, for the ITA, the HTA and combined analyses (left). Branching ratio measurements obtained in this work, the ones given by previous experiments, and the SM value (right).

Figure 9 (right) shows a comparison of the measurements for $\mathcal{B}(B^+ \to K^+\nu\bar{\nu})$ obtained in this work with the previous results by other experiments and the value predicted by the SM. The weighted average is computed assuming symmetrized and uncorrelated uncertainties, excluding the

PoS(HQL2023)049

superseded measurement of Belle II (63 fb^{-1}, Inclusive) and the uncombined results of Belle II shown as open data points. For the ITA the result is in agreement with the previous measurement obtained with hadronic and inclusive tagging methods. A tension with previous semi-leptonic measurement is observed, 2.3σ with BaBar measurement and 1.8σ with Belle. The HTA result is in agreement with all the previous measurements.

7. Summary

A search for the rare decays $B^+ \to K^+\nu\bar{\nu}$ is carried out with data corresponding to an integrated luminosity of 362 fb^{-1}, collected by the Belle II experiment. Two analysis strategies have been employed, the ITA with high sensitivity and the HTA, which is less performant but consists in a well-established approach. The combination of the two analyses yields a branching fraction of $\mathcal{B}(B^+ \to K^+\nu\bar{\nu}) = (2.3 \pm 0.7) \times 10^{-5}$, providing the first evidence of the decay with a significance of 3.5 standard deviations and giving an excess of 2.7 standard deviations over the SM expectations.

References

[1] W. G. Parrott et al., Phys. Rev. D 107, 014511 (2023), [Erratum: Phys.Rev.D 107, 119903 (2023)], arXiv:2207.13371 [hep-ph].

[2] D. Becirevic et al., Phys. Rev. D 98, 055003 (2018) arXiv:1806.05689 [hep-ph].

[3] J. M. Camalich et al., Phys. Rev. D 102, 015023 (2020), arXiv:2002.04623 [hep-ph].

[4] A. Filimonova et al., Phys. Rev. D 101, 095006 (2020) arXiv:1911.03490 [hep-ph].

[5] R. L. Workman et al., Prog. Theor. Exp. Phys. 2022, 083C01 (2022).

[6] K. Akai et al., Nucl. Instrum. Methods Phys. Res., Sect. A 907, 188 (2018) arXiv:1809.01958 [physics.accph].

[7] T. Abe et al. (Belle II Collaboration), KEK Report 2010-1, arXiv.1011.0352 (2010).

[8] Belle II Collaboration, Belle II Preprint 2023-017, KEK Preprint 2023-35 arXiv:2311.14647.

[9] T. Keck et al., Comput. Softw. Big Sci. 3, 6 (2019), arXiv:1807.08680 [hep-ex].

[10] D. Martschei et al., J. Phys. Conf. Ser. 368, 012028 (2012)

[11] J. P. Lees et al. (BaBar Collaboration), Phys. Rev. D 85, 112010 (2012), arXiv:1201.5897 [hep-ex].