$B \to K^{(*)}\nu\nu$ with inclusive tag @Belle II

Yubo Han (Univ. of Hawaii)

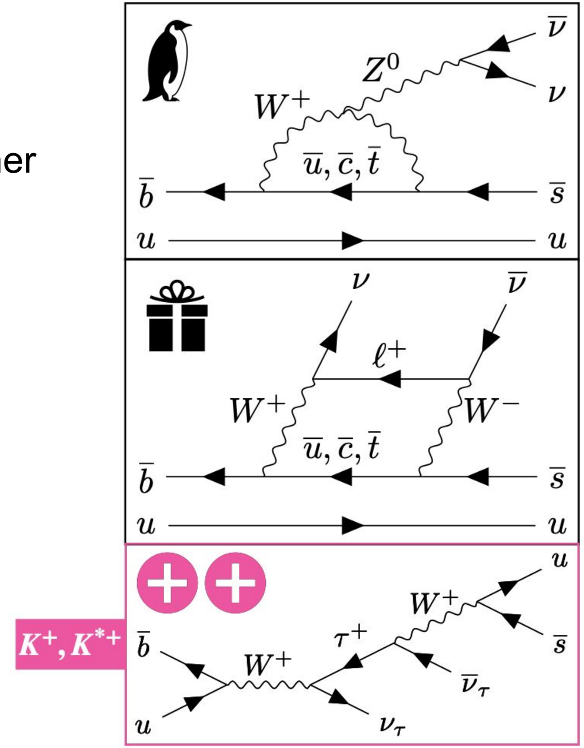
yubohan@hawaii.edu

2025 US Belle II summer workshop, Virginia Tech

Why B→K$^{(*)}$vv?

- B→K$^{(*)}$vv in SM: Flavour-changing neutral current transition
  - Highly suppressed in SM (Br(b→s)<10^{-5}), only occur at higher orders in SM
  - Precise predicted in SM, leading theoretical uncertainty from hadronic form factors

  → Highly sensitive to new physics



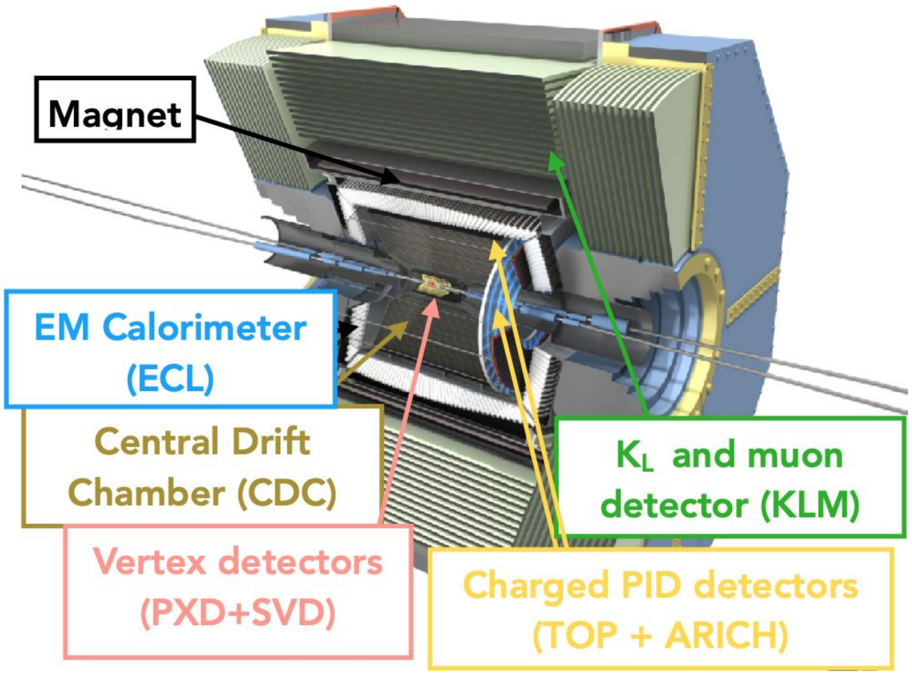
SM prediction [Eur.Phys.J.C 83 (2023) 3, 252] [x10^{-6}]

Decay	SM total	LD contribution	SD contribution
$B^+ \to K^+\nu\bar{\nu}$	5.22 ± 0.32	0.63 ± 0.06	4.59 ± 0.32
$B^0 \to K^0_S\nu\bar{\nu}$	2.12 ± 0.15	—	2.12 ± 0.15
$B^+ \to K^{*+}\nu\bar{\nu}$	11.27 ± 1.51	1.07 ± 0.10	10.20 ± 1.51
$B^0 \to K^{*0}\nu\bar{\nu}$	9.47 ± 1.40	—	9.47 ± 1.40

background signal

Why Belle II: the challenge

- $B \to K^{(*)} \nu\nu$ is experimentally **challenging:**
    - 3-body decay with two neutrinos in the final states
        - Full reconstruction of the kinematic properties of the signal B not possible
        - Usually rely on the reconstruction of another B
    - Very low branching fraction, large background contamination

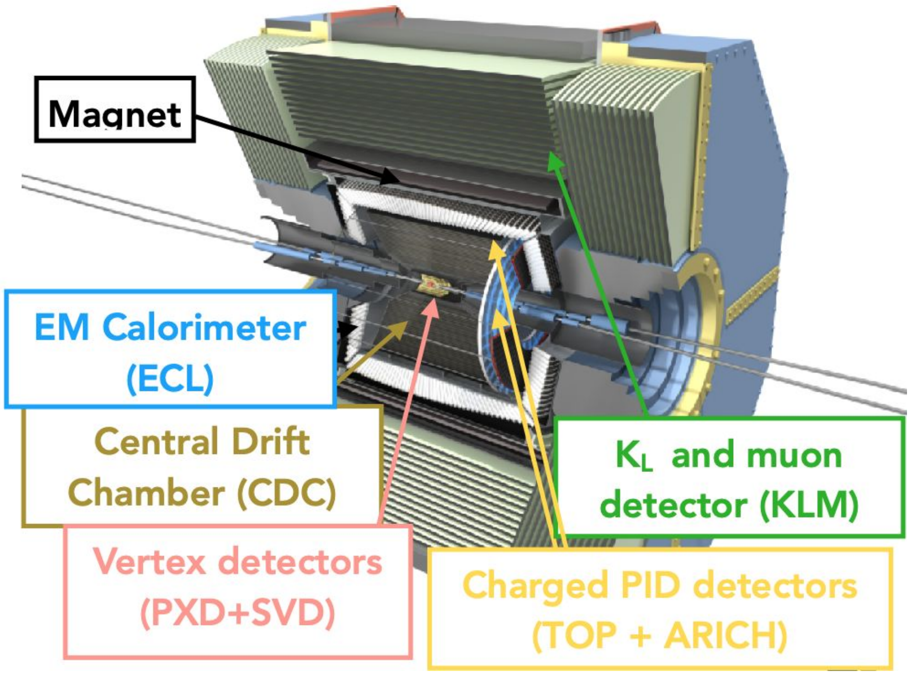


Why Belle II: the opportunity

- **SuperKEKB** is an asymmetric e^+e^- collider at Υ(4S) energy:
  - Υ(4S) → BBbar in 96%
- **Belle II** is a **hermetic detector**:
  - Well-known initial-state kinematics
  - Rather clean environment
  - Excellent particle identification and tracking performance
  - Good neutral particle reconstruction

  → Excellent for decays with neutral or invisible final states

- Belle II has a Unique capability to study this decay

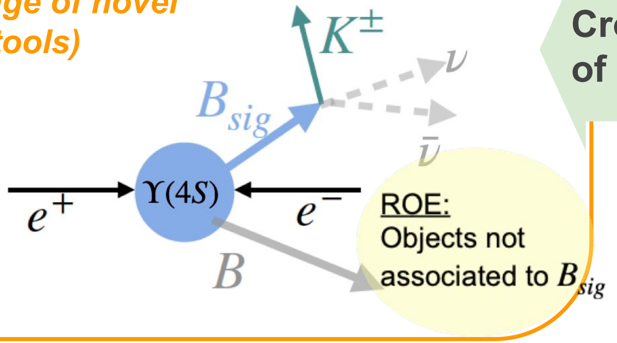


Why $B \to K^{(*)}\nu\nu$ Inclusive @Belle II?

Inclusive tagging (ITA)

New wrt HTA

- Reconstruct B_{sig} through signal K
- Use properties of the rest of event (ROE) *inclusively*
- ***High efficiency***
- *(usage of novel ML tools)*

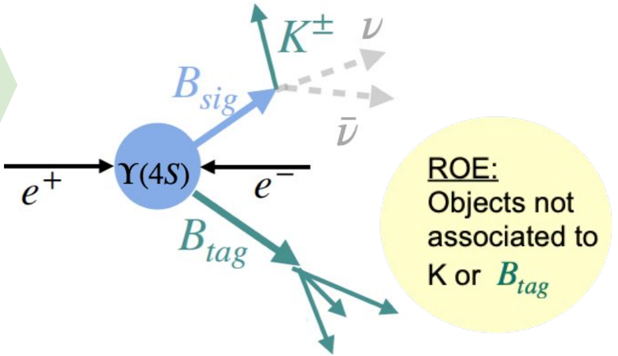


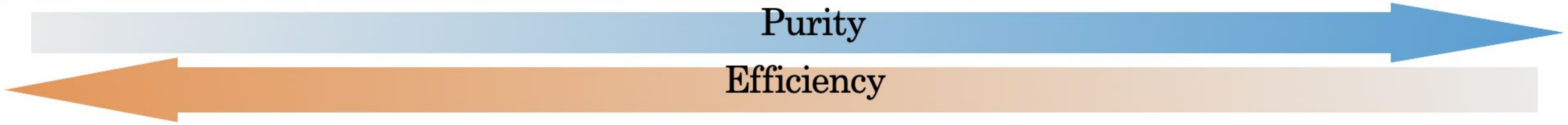
Cross check of each other

Hadronic tagging (HTA)

Well-established tradition method

- Reconstruct B_{tag} through hadronic decay mode. B_{sig} through signal K
- Different definition of ROE
- Eff ~1%





Overview of B→K$^{(*)}$vv inclusive analysis @Belle II

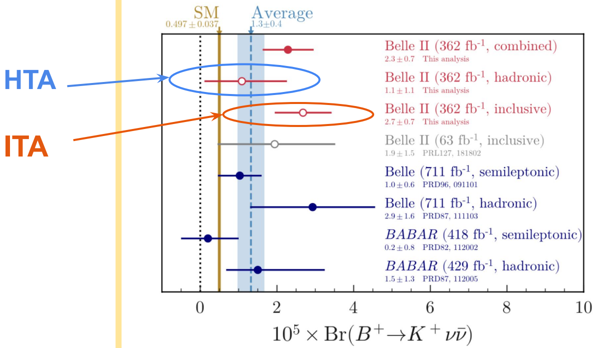
2024 2025 **Target: 2025 summer**

B→K$^{(*)}$vv for Moriond 2022 with 189.2 fb^{-1} data

[BELLE2-NOTE-PH-2021-047]

★ B^{+}→K^{+}vv Belle II Run 1 data
- 3.5 σ wrt null
- 2.7 σ wrt SM
- First evidence of B^{+}->K^{+}vv decays

PhysRevD.109.112006

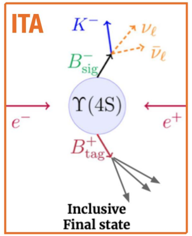


★ B→K$^{(*)}$vv with inclusive tag using Run I data
- Started ~ Feb. 2024
- Similar analysis flow with potential improvement
- Combined analysis of 4 channels

Gitlab repository
EWP analysis tracker
BELLE2-NOTE-PH-2024-031

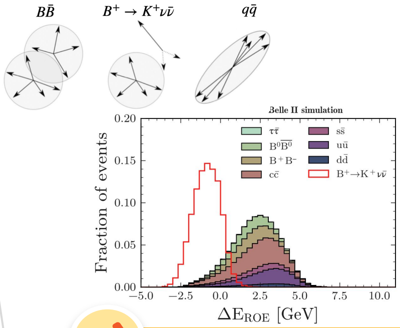
- Run 1 data:
  - data taken from 03.2019 - 06.2022
  - *365 fb^{-1} @Y4S, 42 fb^{-1} offres*

Analysis workflow in a nutshell *(K^+ channel as an example)*



1. Basic selection and reconstruction

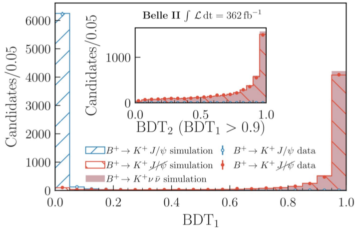
- Basic reconstruction of objects to suppress beam-related background)
- Best candidate selection applied here



2. Background suppression

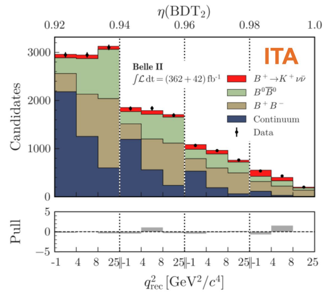
- Two subsequent BDTs involved: BDT1 + BDT2
  - BDT1: first-level filter
  - BDT2: final selection

***BDT:** ML method combines many input features to make predictions*



3. Validation

- Validate the selection and modeling with control samples
- Correction introduced if needed
- Systematic uncertainty determination



4. Statistical interpretation

- Maximum likelihood fit with pyhf

***Pyhf:** python histFactory for building and fitting statistical models*

Analysis workflow: improvements

1. Basic selection and reconstruction
2. Background suppression
3. Validation
4. Statistical interpretation

New for KS0, K^{*0}, K^{*+}

Optimized candidate selection
>=1 best candidates at early stage

Update the variables list for BDT2 training and exploration of other ML tools

Validation with more control samples (fakes)

Optimization of the SR definition

*Optimize the treatment of D**, leading B decays*

Will not cover all of them today

Event selections

Common for all 4 channels:

- Charged tracks: P_t>0.1 GeV/c, 0.1<E<5.5 GeV, within central region of the detector, close to IP, nPXDHits>0
- Neutrals: 0.06<E<5.5 GeV, within central region of the detector and not matched to tracks
- low-multiplicity event veto: Visible energy>4 GeV 0.3<theta of missing momentum < 2.8
- *(mass squared of the neutrino pair)* $q^2 > -1$ GeV/c^2 ($q^2 = s/(4c^4) + M^2_{K(*)} - \sqrt{s} E_K^*/c^4$), total charge (Q_net^2) <=4

B->K_S^0νν → Best Candidate selection

K_S^0:
- K_S^0:merged, 0.485<M<0.51GeV/c^2, cos(p,vertex)>0.98
- B^0: 2<nROETracks<9

B->K^{*0}νν → MVA to select the best 2 candidates

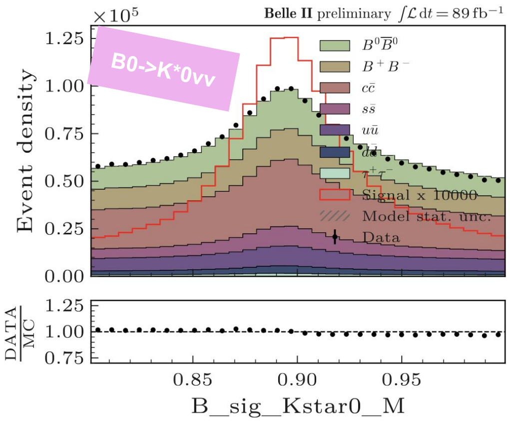
K^{*0} (-> $K^+\pi^-$) :
- K^{*0} : KID>0.75, πID>0.05, (+Treefit) K and π 0.8<M(K^{*0})<1.0 GeV/c^2
- B^0: 2 <nROETracks < 9

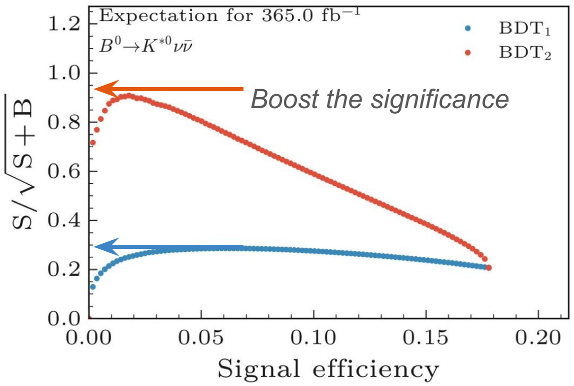
B->K^{*+}νν → 5 Best Candidates

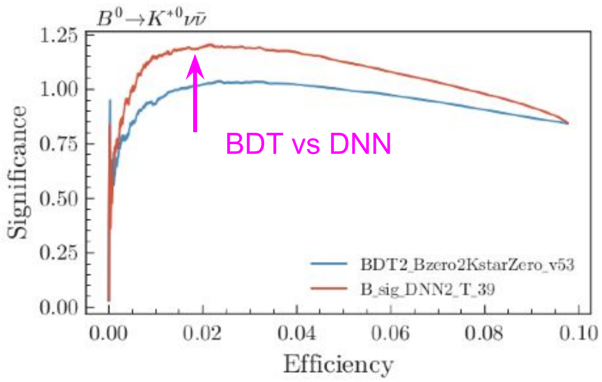
K^{*+}(-> $K^+\pi^0$, KS0 π^+)
- K*+: KID>0.9,nCDCHits>10,abs(dz)<0.3,dr<0.05,E<3.0, π^0: stdPi0s(photonlist=eff50_May2020), mass-constrained kFit 0.8<M<1.0GeV/c^2, no vertex fit. K_S^0: same as K_S^0 channel
- B^+: 3<=nROETracks<=9

Exploration of ML tools

- 2 subsequent filters(BDTs) involved:
  - BDT1: event-shape based variable for continuum suppression
  - BDT2: final selection exploring more variable (boost the performance in BDT1>0.9 region)
- Optimization regarding BDT2:
  - Review the variable list for all the channels
  - Optimization of the BDT configuration
  - Explore DNN method -> might switch to DNN in the future







(b) $B^0 \to K^{*0}\nu\bar{\nu}$ channel

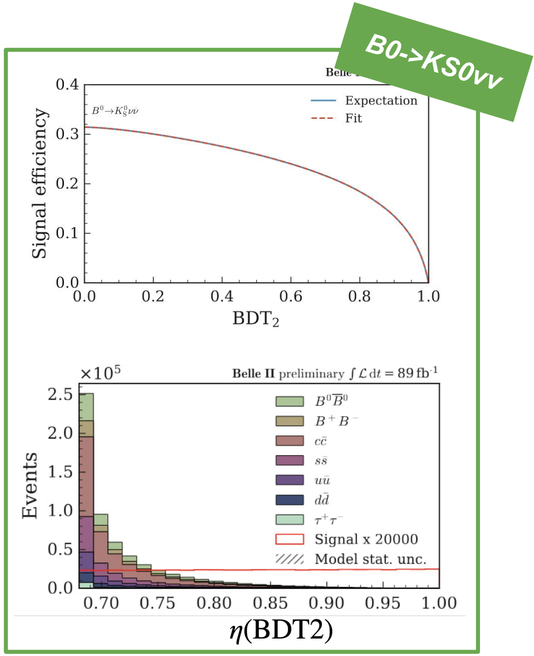
Optimization of Signal region definition

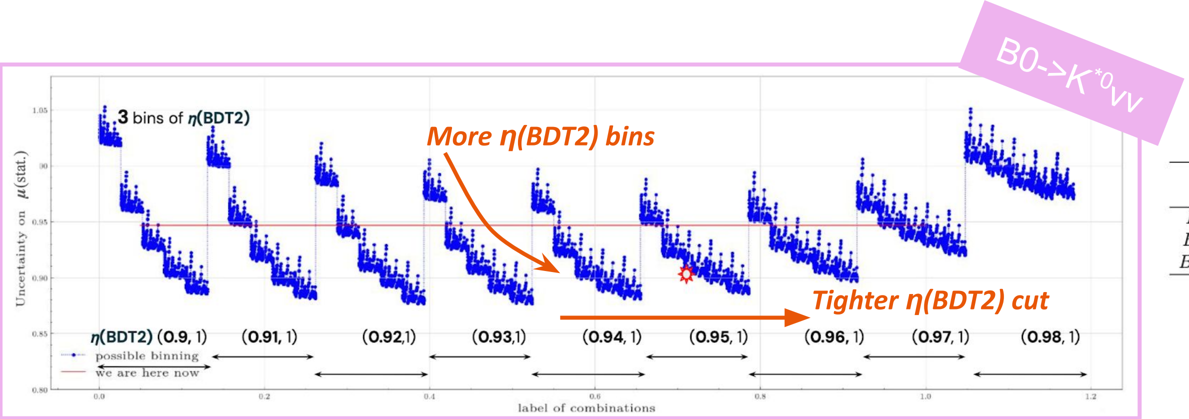
- 2D binning using *η(BDT2)* and q^2:

$$\boxed{\eta(\mathrm{BDT}_2) \equiv 1 - \int_{\mathrm{BDT}_2}^{1} \epsilon(b)\mathrm{d}b}$$

η(BDT2): More physical variable than BDT2 output

  - K_S^0 : keep the same as of K+
  - K^{*0}, K^{*+}: optimize based on Asimov fit to minimize the expected uncertainty on μ
  - Avoid too many η(BDT2) bins to have enough MC events per bin





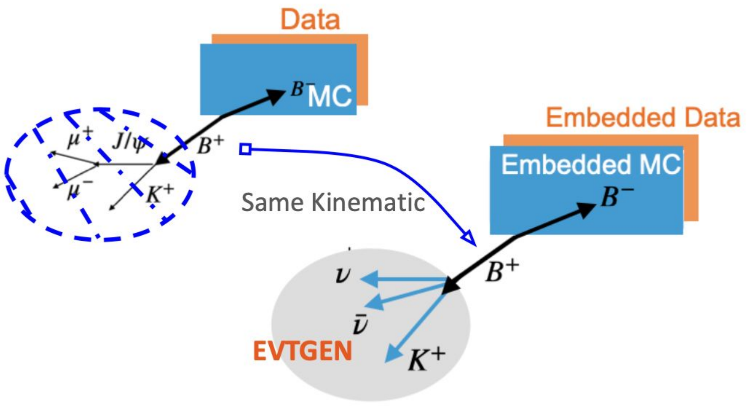
Decay	Bin boundaries ($\eta(\mathrm{BDT}_2) \times q^2$)
$B^0 \to K_S^0 \nu\bar{\nu}$	[[0.92, 0.94, 0.96, 0.98, 1.00] × [-1, 4, 8, 25]]
$B^0 \to K^{*0} \nu\bar{\nu}$	[[0.95, 0.96, 0.97, 0.98, 0.99, 1.00] × [-1, 4, 8, 25]]
$B^+ \to K^{*+} \nu\bar{\nu}$	[[0.97 , 0.975, 0.98 , 0.985, 0.99 , 0.995, 1.00] × [-1, 4, 8, 25]]

Asimov fit: fit to the expectations

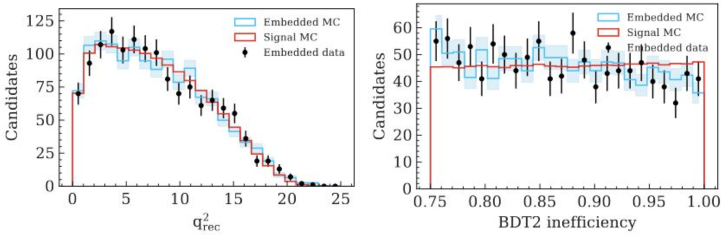
Validations with control samples

Validating the modeling of the background is critical for this analysi

- ➢ Off-resonance data
  - ○ Normalization of continuum background
  - ○ Continuum shape corrections via BDTc
- ➢ Signal-embedded $B \rightarrow K^{(*)} J/\psi$
  - ○ New embedding procedure
  - ○ Signal selection efficiency
- ➢ η(BDT2) sideband
  - ○ Check data-mc agreement with all corrections applied
- ➢ Kaon mass sideband
  - ○ Check modeling of fake $K^{(*)}$
- ➢ $D \rightarrow K^{(*)} X$ sample
  - ○ Check modeling of $D \rightarrow K^{(*)} X$
  - ○ Dominant background and could be used to constrain normalization systematic uncertainty
- ➢ $B \rightarrow K^{(*)} K^0 K^0$ modeling



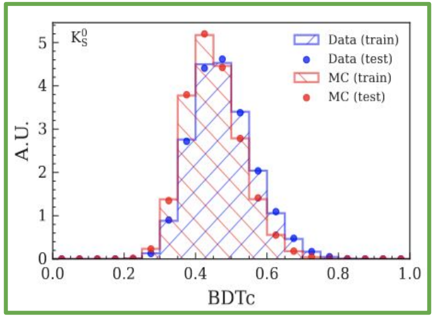
K_S^0



Validation

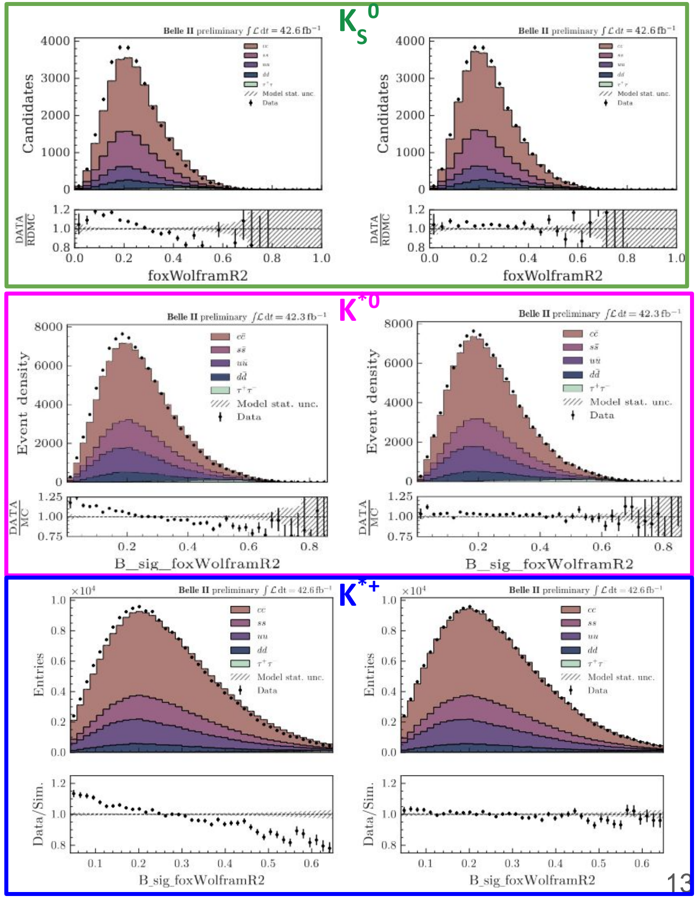
Correction example: BDTc

- Correct the discrepancies in shapes of the continuum background
- Data-driven events re-weighting
  - Training of BDT to separate data and simulation event
  - Assign weight to simulation events to suppress events with large separation



- Improved data-mc agreement after BDTc correction

w/o BDTc
w/ BDTc

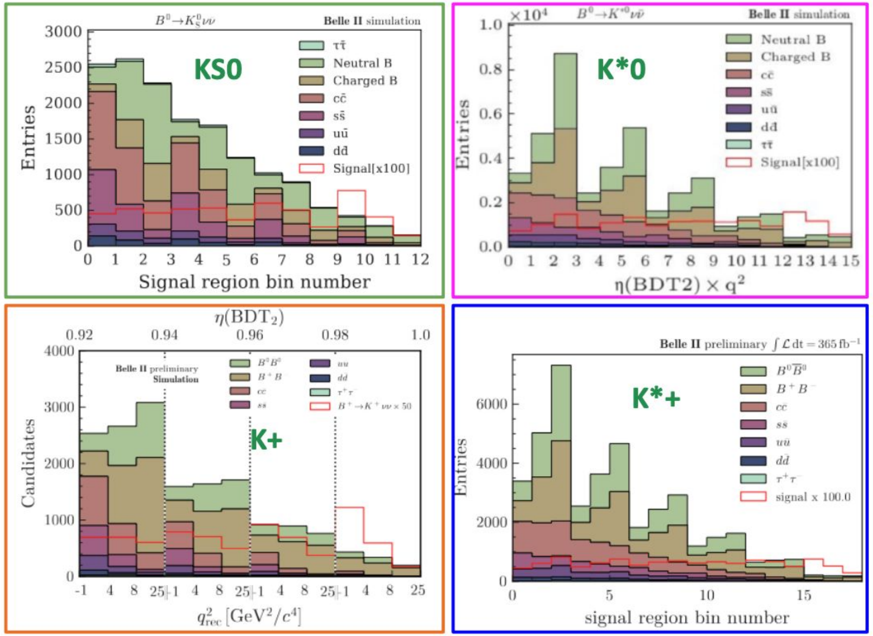


Signal extraction

- Sample compositions in Signal Region
- Fitting strategy:
  - 2D Binned maximum likelihood to extract parameter of interest: signal strength μ

$$\mu = \frac{\mathscr{B}(B^+ \to K^+\nu\bar{\nu})}{\mathscr{B}_{SM}(B^+ \to K^+\nu\bar{\nu})}$$

  - Simultaneously extract signal strength for each channel
  - Correlations between channels were taken into account (PID, Kshort eff)
  - Crossfeed between channels also included



Systematics

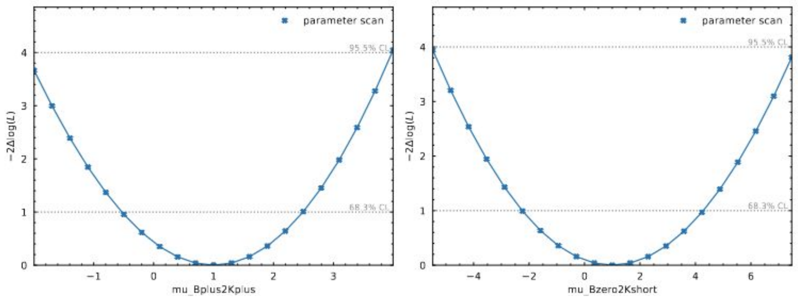
- Impact of each systematic uncertainty estimated using Asimov fit:
  - removing them one-by-one and check the uncertainty change of the nominal fit
  - New for this iteration
  - Not the final version →to be updated

Source	Uncertainty on μ $B^+ \to K^+\nu\bar{\nu}$	$B^0 \to K^0_S\nu\bar{\nu}$	$B^+ \to K^{*+}\nu\bar{\nu}$	$B^0 \to K^{*0}\nu\bar{\nu}$
Global normalizations	0.88	2.28	1.39	0.92
MC stats	0.55	1.38	1.07	0.65
K^0_L efficiency	0.34	0.01	0.31	0.47
Kaon ID	0.06	0.02	0.03	0.05
Photon energy	0.09	0.05	0.36	0.12
Hadronic energy	0.36	0.14	0.80	0.16
Tracking	0.09	0.12	0.07	0.07
B counting	0.02	0.00	0.00	0.00
Luminosity	0.00	0.00	0.00	0.00
Leading BF	0.26	0.10	0.20	0.15
D^{**}	0.44	0.26	0.49	0.46
$B \to K^{(*)}n\bar{n}$	0.24	0.02	0.05	0.00
$B^+ \to K^+K^0_LK^0_L$	0.58	0.02	0.10	0.02
$B^+ \to K^+K^0_SK^0_L$	0.02	0.00	0.00	0.00
$B^0 \to K^0K^0K^0$	0.00	0.35	0.01	0.00
$B^0 \to K^{*0}K^0_LK^0_L$	0.02	0.03	0.04	0.35
$D \to K^0_L$	0.21	0.01	0.08	0.08
Offresonance luminosity	0.01	0.00	0.00	0.00
BDTc	0.05	0.98	0.51	0.10
π^0 efficiency (total)	0.02	0.02	0.05	0.03
K^0_S efficiency (stat)	0.01	0.08	0.02	0.00
K^0_S efficiency (syst)	0.03	0.40	0.02	0.00
Fake K^*	0.13	0.00	0.30	0.14
Signal form factors	0.06	0.06	0.12	0.13

Tests of the fitting template

- Asimov fit: (365 fb^{-1})

Channel	σ_μ symmetric	$\pm\sigma_\mu$ profiled
$B^+ \to K^+\nu\bar{\nu}$	1.50	+ 1.49 - 1.53
$B^0 \to K^0_S\nu\bar{\nu}$	3.24	+ 3.29 - 3.25
$B^+ \to K^{*+}\nu\bar{\nu}$	2.29	+ 2.36 - 2.30
$B^0 \to K^{*0}\nu\bar{\nu}$	1.63	+ 1.67 - 1.62



- Isospin average fit:

  Single μ is used for K^+ and K_s^0 channel and a second μ used for K^{*+} and K^{*0} channel
  10% and 16% improvement observed (${\sigma_{\mu,\,Ks0}}$ =1.14 with 500 fb^{-1})

Channel	σ_μ symmetric	$\pm\sigma_\mu$ profiled
$B^+ \to K^+\nu\bar{\nu}$	1.34	+ 1.35 - 1.37
$B^0 \to K^0_S\nu\bar{\nu}$	1.34	+ 1.35 - 1.37
$B^+ \to K^{*+}\nu\bar{\nu}$	1.37	+ 1.41 - 1.37
$B^0 \to K^{*0}\nu\bar{\nu}$	1.37	+ 1.41 - 1.37

Summary

★ **B →K$^{(*)}$νν with inclusive tag using Run I data with improvements across multiple stages:**

- Optimized and channel-wised event selection strategy
- Channel-wise BDTs and SR definition
- Extensive studies and validation of background modeling
- Improved control of key systematic uncertainties
- Combined fit with systematic uncertainties correlated across channels

★ Current stage: Finalizing analysis under **Working Group review**

Thanks!

Samples

Data:

On-res Data (365 fb^{-1}): Proc13 + Prompt

Off-res Data (42 fb^{-1})

MC:

MC15rd:

4*365 fb^{-1} generic MC

4M signal MC for each channel

Off-res MC (4*42 fb^-1)

MC15ri:

400 fb^{-1} generic MC *-> for training/testing of BDTs*

10M/40M signal MC for each channel

basf2 : light-2311-nebelung

The whole picture of the template

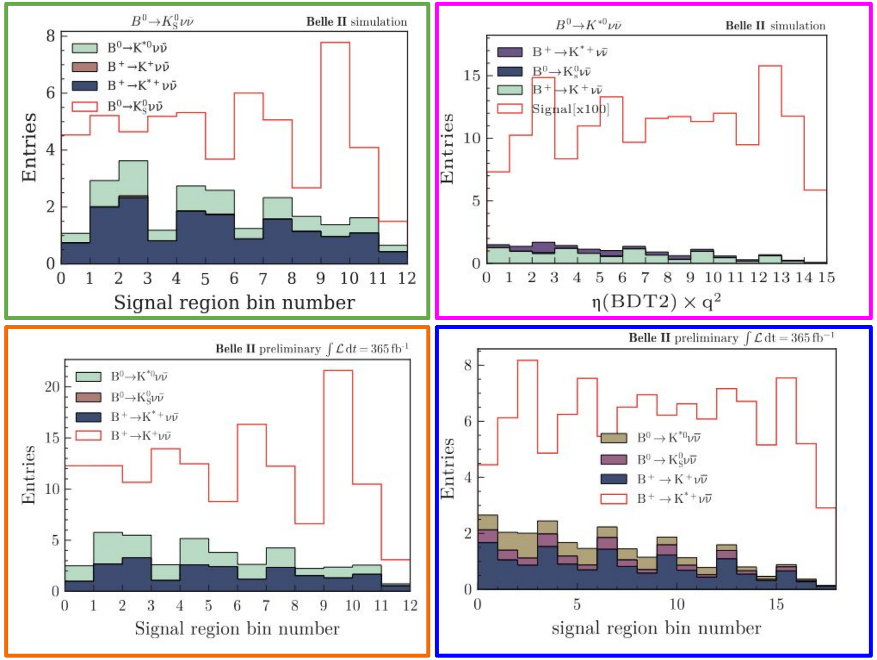
Parameter	Modifier	Process	Channel	N. of pars & Correlation scheme
$\mu_{B^+\to K^+\nu\bar{\nu}}$	Unconstr.Normalization	$B^+ \to K^+\nu\bar{\nu}$	all	1, shared between channels
$\mu_{B^0\to K^0_S\nu\bar{\nu}}$	Unconstr.Normalization	$B^0 \to K^0_S\nu\bar{\nu}$	all	1, shared between channels
$\mu_{B^+\to K^{*+}\nu\bar{\nu}}$	Unconstr.Normalization	$B^+ \to K^{*+}\nu\bar{\nu}$	all	1, shared between channels
$\mu_{B^0\to K^{*0}\nu\bar{\nu}}$	Unconstr.Normalization	$B^0 \to K^{*0}\nu\bar{\nu}$	all	1, shared between channels
$\mu_{B^+B^-}$	Normalization Unc. (50%)	B^+B^-	all	4, uncorrelated, one per channel
$\mu_{B^0\bar{B}^0}$	Normalization Unc. (50%)	$B^0\bar{B}^0$	all	4, uncorrelated, one per channel
$\mu_{u\bar{u}}$	Normalization Unc. (50%)	$u\bar{u}$	all	4, uncorrelated, one per channel
$\mu_{d\bar{d}}$	Normalization Unc. (50%)	$d\bar{d}$	all	4, uncorrelated, one per channel
$\mu_{c\bar{c}}$	Normalization Unc. (50%)	$c\bar{c}$	all	4, uncorrelated, one per channel
$\mu_{s\bar{s}}$	Normalization Unc. (50%)	$s\bar{s}$	all	4, uncorrelated, one per channel
$\mu_{\tau^+\tau^-}$	Normalization Unc. (50%)	$\tau+\tau^-$	all	4, uncorrelated, one per channel
N_{BB}	Normalization Unc.	signals, B^+B^-, $B^0\bar{B}^0$	all	1, shared between processes and channels
$f^{00/\pm}$	Normalization Unc.	signals, B^+B^-, $B^0\bar{B}^0$	all	1, shared between processes and channels
Luminosity	Normalization Unc.	$u\bar{u}, d\bar{d}, c\bar{c}, s\bar{s}, \tau^+\tau^-$	all	1, shared between channels
Leading branching fractions	correlated shape	B^+B^-, $B^0\bar{B}^0$	all	104 nuisance parameter, shared between channels.
$\mathcal{B}(B \to D^{**}X)$	correlated shape	B^+B^-, $B^0\bar{B}^0$	all	4, shared between processes and channels
$\mathcal{B}(B^0 \to Kn\bar{n})$,$\mathcal{B}(KK^0_L K^0_L)$...	correlated shape	B^+B^-, $B^0\bar{B}^0$	all	$4 \times N_{\mathrm{decays}}$, uncorrelated, shared between processes
$\mathcal{B}(D \to K_L X)$	correlated shape	B^+B^-, $B^0\bar{B}^0$	all	1, shared between processes and channels
Continuum modelling (BDT_c)	correlated shape	all	all	4, one per channel channel
Off-resonance normalization	Normalization Unc. (5%)	all	all	1, shared between continuum processes and channels.
Efficiency - tracking	correlated shape	all	all	1, shared between processes and channels
Efficiency - siganl K PID	correlated shape	all	all	11, shared between processes and channels
Efficiency - hadronic energy scale	correlated shape	all	all	1, shared between processes and channels
Efficiency - Photon energy sclale	correlated shape	all	all	1, shared between processes and channels
Efficiency - π^0	correlated shape	all	$B^+ \to K^{*+}\nu\bar{\nu}$	11+1, shared between processes and channels
Efficiency - K^0_L in ECL	correlated shape	all	all	1, shared between processes and channels
Efficiency - K^0_S	correlated shape	all	$B^0 \to K^0_S\nu\bar{\nu}$, $B^+ \to K^{*+}\nu\bar{\nu}$	11+1, shared between processes and channels
Signal efficiency	Normalisation Unc.	signals	all	4, one per channel
Fake K^*	correlated shape	B^+B^-, $B^0\bar{B}^0$, $u\bar{u}, dd, c\bar{c}, s\bar{s}, \tau^+\tau^-$	$B^0 \to K^{*0}\nu\bar{\nu}$, $B^+ \to K^{*+}\nu\bar{\nu}$	2, shared between processes
Signal form factors (pseudo scalar)	correlated shape	signals	$B^+ \to K^+\nu\bar{\nu}$, $B^0 \to K^0_S\nu\bar{\nu}$	3, correlated between processes
Signal form-factors (vector)	correlated shape	signals	$B^+ \to K^+\nu\bar{\nu}$, $B^0 \to K^0_S\nu\bar{\nu}$	9, correlated between processes
MC sample size	MC Stat. Uncertainty	all	all	114, one per SR bin

Signal extraction

- Sample compositions in SR
- Fitting strategy:
  - 2D Binned maximum likelihood to extract parameter of interest: signal strength μ

$$\mu = \frac{\mathcal{B}(B^+ \to K^+ \nu \bar{\nu})}{\mathcal{B}_{SM}(B^+ \to K^+ \nu \bar{\nu})}$$

  - Simultaneously extract signal strength for each channel
  - Correlations between channels were taken into account (PID, Kshort eff)
  - Crossfeed between channels also included
    - **KS0:** 41.4% **K*0:** 8.6%
    - **K+:** 27.7% **K*+:** 28.1%



Systematic Uncertainties: New for KS0, K* channels

Systematic table for K+ channel from the published result

Source	Correction	Uncertainty size	Impact on σ_μ
Normalization of $B\bar{B}$ background	—	50%	0.90
Normalization of continuum background	—	50%	0.10
Leading B-decay branching fractions	—	$O(1\%)$	0.22
Branching fraction for $B^+ \to K^+ K^0_L K^0_L$	q^2 dependent $O(100\%)$	20%	0.49
p-wave component for $B^+ \to K^+ K^0_S K^0_L$	q^2 dependent $O(100\%)$	30%	0.02
Branching fraction for $B \to D^{**}$	—	50%	0.42
Branching fraction for $B^+ \to K^+ n\bar{n}$	q^2 dependent $O(100\%)$	100%	0.20
Branching fraction for $D \to K^0_L X$	+30%	10%	0.14
Continuum-background modeling, $\mathrm{BDT_c}$	Multivariate $O(10\%)$	100% of correction	0.01
Integrated luminosity	—	1%	< 0.01
Number of $B\bar{B}$	—	1.5%	0.02
Off-resonance sample normalization	—	5%	0.05
Track-finding efficiency	—	0.3%	0.20
Signal-kaon PID	p, θ dependent $O(10-100\%)$	$O(1\%)$	0.07
Photon energy	—	0.5%	0.08
Hadronic energy	-10%	10%	0.37
K^0_L efficiency in ECL	-17%	8%	0.22
Signal SM form-factors	q^2 dependent $O(1\%)$	$O(1\%)$	0.02
Global signal efficiency	—	3%	0.03
Simulated-sample size	—	$O(1\%)$	0.52

Extend to 4 channels

Update treatment of LeadingB

*Update treatment of D***

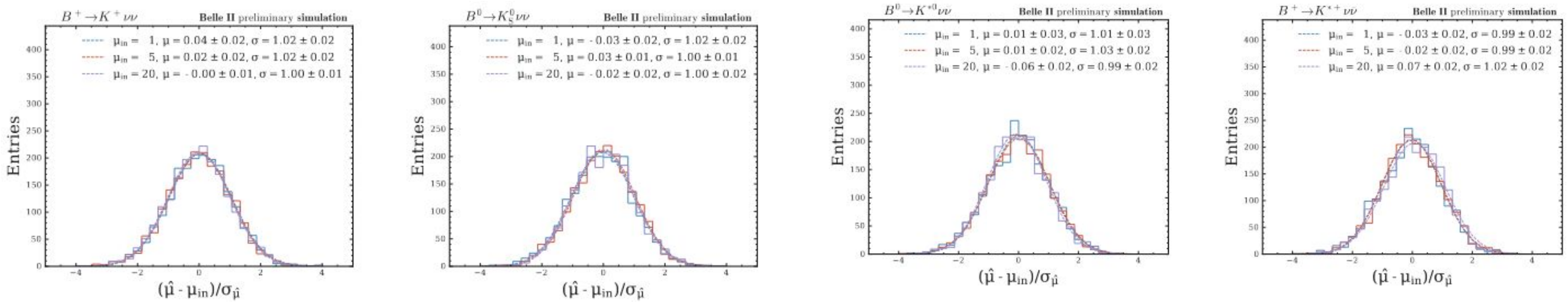
+ Ks efficiency, Pi0 efficiency

+ Signal crossfeed between 4 channels

Tests of the fitting template

- Signal injection study:

  -> No bias observed



- Towards unblinding:
  - Yield stability per unit of integrated luminosity
  - Check the fit quality in SR
  - Bind tests on half-split samples
  - …

Common corrections

(the weights involved in getting the nominal yield)

D->KL0: scale up by 30% +/-10%   *(taken from previous measurement)*

PID corrections: efficiency and fake rate using systematics framework

π0 efficiency: obtained from neutrals performance group

Unmatched photon-candidates energy: -10% in MC

KL0 efficiency: -17% +/- 8% in MC

KS0 efficiency: follow instructions from tracking performance group

Re-evaluate KL0 eff using rd mc