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

[arxiv:2104.12624]

Slavomira Stefkova

LHCb FSP meeting, 08.10.2021

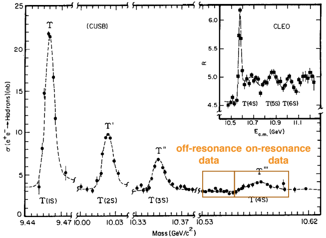
HELMHOLTZ RESEARCH FOR GRAND CHALLENGES

Belle II

SuperKEKB

SuperKEKB is an energy asymmetric e^+e^- collider @ $\sqrt{s} = 10.58$ GeV:

- ▷ $\sqrt{s} = 10.58$ GeV $\leftrightarrow \Upsilon(4S)$ resonance $\rightarrow$
  $\Upsilon(4S) \rightarrow B\bar{B}$ + nothing else with $\mathcal{B} > 96\%$
  $\rightarrow$ clean B sample (on-resonance)
- ▷ @ 60 MeV below $\Upsilon(4S)$ resonance
  $\rightarrow$ control sample to constrain continuum backgrounds ($e^+e^- \rightarrow q\bar{q}$)
  (off-resonance)



SuperKEKB is not only *B*-factory:

- ▷ τ and c pairs have similar cross-sections @ $\sqrt{s} = 10.58$ GeV

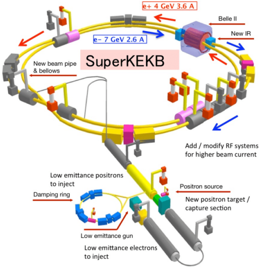
With nanobeam scheme and upgraded rings SuperKEKB aims to reach $30\times$ higher $\mathcal{L}_{\text{inst}}$ than KEKB at cost of $\mathcal{O}(15)$ higher backgrounds:

- ▷ x 1.5 currents
- ▷ x 1/20 β_y^*

$$L = \frac{\gamma_\pm}{2 e r_e}\left(1 + \frac{\sigma_y^*}{\sigma_x^*}\right)\frac{I_\pm \xi_{y\pm}}{\beta_{y\pm}}\frac{R_L}{R_{\xi_y}}$$

beam current

vertical beta function at IP

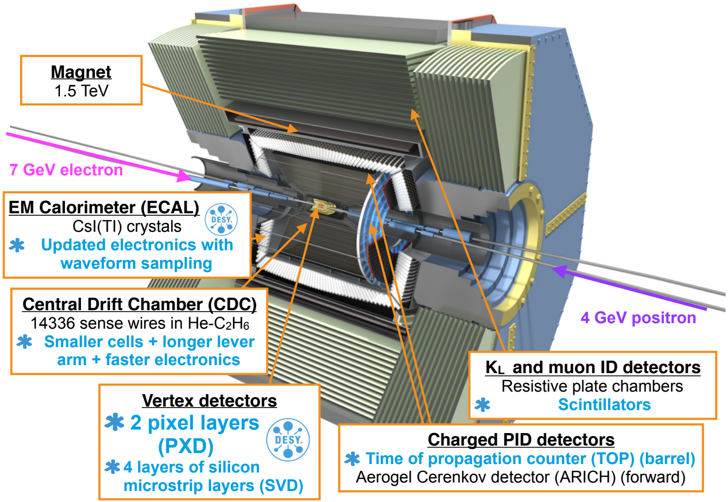


LHC vs SuperKEKB

Very nice talk by by Diego Tonelli for many more details!

LHC	**SuperKEKB**
pp-collisions *b*-quarks produced by gluon fusion all *b*-hadrons species (B_d, B_s, B_c, *b*-baryon) highly boosted topology $\sigma_{bb} = 100\ \mu$b different backgrounds (N/S = 1000)	e^+e^- energy asymmetric collisions $B\bar{B}$ produced from $\Upsilon(4S)$ exclusive $B\bar{B}$ production asymmetric beam energy $\rightarrow$ boost $\sigma_{bb} = 1.1$ nb *B*-backgrounds, continuum backgrounds + QED (N/S=4)
1 fb^{-1}	1 ab^{-1}

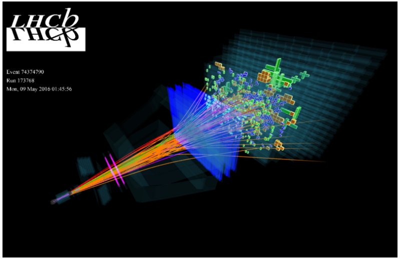
Belle II Detector

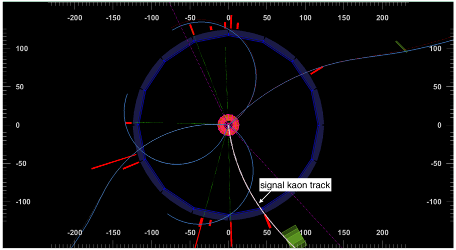


- Belle II detector was designed to give **similar or better performance** than Belle even under $\mathcal{O}(15)$ higher backgrounds
- **DAQ and trigger** systems were also upgraded!

LHCb vs Belle II

LHCb	Belle II
single-arm detector longitudinal momentum of *B* not known	hermetic detector known initial state kinematics pro @ neutral object reconstruction (photon, K_L)



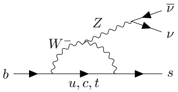


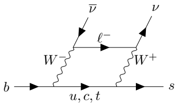
- $B^+ \to K^+\nu\bar{\nu}$ **is a golden channel @ Belle II:** clean environment and well defined initial state but still **challenging** as two neutrinos in the final state leave no signature

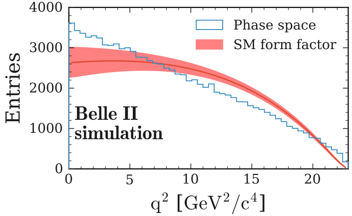
Theory Introduction

SM Theory:

- Flavour-changing neutral current transition (FCNC) occurring at loop level only $\rightarrow$ highly suppressed
- **Clean SM computation** $\rightarrow$ does not suffer from charm-loop contributions
- $\mathcal{B}(B^+ \rightarrow K^+ \nu\bar{\nu}) = (4.6 \pm 0.5) \times 10^{-6}$ [arxiv:1606.00916]
- Leading uncertainty: $B^+ \rightarrow K^+$ form factor
- SM $q^2(\nu\bar{\nu})$ distribution [arXiv:1409.4557]

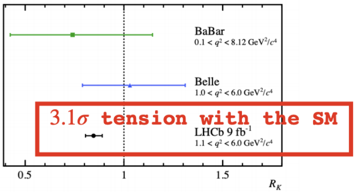






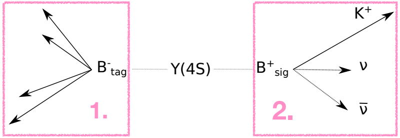
BSM Interest:

- Complementary channel to $b \rightarrow sll$ transitions where tensions with the SM have been observed
- NP scenarios:
  Leptoquarks [PRD 98, 055003 (2018)],
  Axions [PRD 102, 015023 (2020)],
  and Dark Matter candidates [PRD 101, 095006 (2020)]



Previous Measurements

- **Traditionally** searched for with explicit B_{tag} reconstruction
  - 1. Reconstruct B_{tag} using MVA algorithm either in semileptonic or hadronic channels
  - 2. Reconstruct B_{sig}
  - Advantage: Flavour constraint ($B^-_{tag} \rightarrow B^+_{sig}$), kinematically constrained system, high purity
  - Disadvantage: Low efficiency = $\epsilon_{B_{tag}} \times \epsilon_{B_{sig}}$
    hadronic: ~ 0.04 %, semileptonic: ~ 0.2 %



- Summary of **upper limits** from previous *B*-factories:

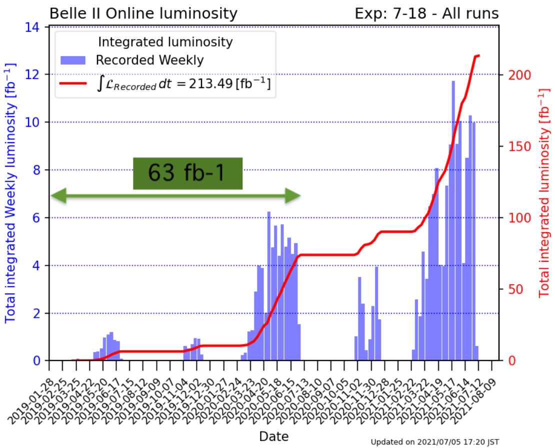
Experiment	Year	Observed limit on $\mathrm{BR}(B^+ \rightarrow K^+ \nu \bar{\nu})$	Approach	Data [fb^{-1}]
BABAR	**2013**	$< 1.6 \times 10^{-5}$ **[Phys.Rev.D87,112005]**	**SL + Had tagging**	429
Belle	**2013**	$< 5.5 \times 10^{-5}$ **[Phys.Rev.D87,111103(R)]**	**Had tagging**	711
Belle	**2017**	$< 1.9 \times 10^{-5}$ **[Phys.Rev.D96,091101(R)]**	**SL tagging**	711

Dataset

This measurement of $\mathcal{B}(B^+ \to K^+\nu\bar{\nu})$ is based on Moriond 2021 dataset:

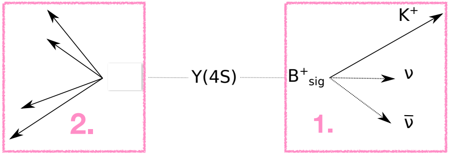
- **63 fb**$^{-1}$ on-resonance data
- **9 fb**$^{-1}$ off-resonance data
- Background samples corresponding to 600 fb^{-1} (mixed B, charged B, $c\bar{c}$, $u\bar{u}$, $s\bar{s}$, $d\bar{d}$, $\tau^+\tau^-$)
- Signal sample corresponding to 1 million events

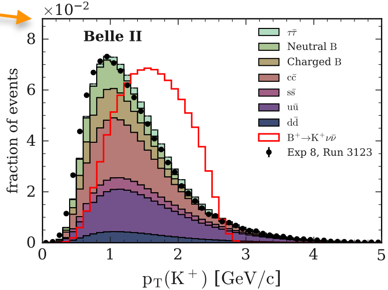
10×smaller $\mathcal{L}$ wrt previous measurements → new approach of inclusive tagging



The Inclusive Tag

- Inclusive tag approach exploits **very distinct signal kinematics**!
- 1. Reconstruct B_{sig}: select the highest-p_T clean track in the event with at least 1 PXD hit as the K^+ candidate (78% true candidate)
- 2. Basic reconstruction of the rest-of-the-event (ROE) object : remaining tracks and clusters
- **Higher signal efficiency but also higher background contamination**
- Background contamination suppressed with several cuts and BDTs (see next slides)

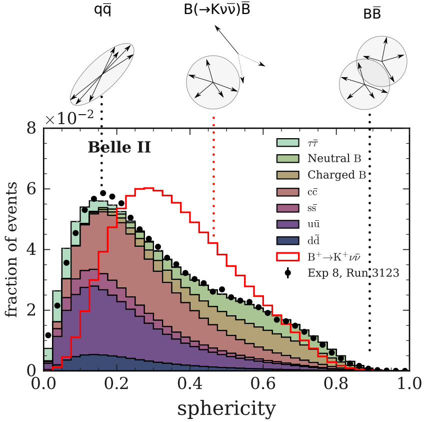


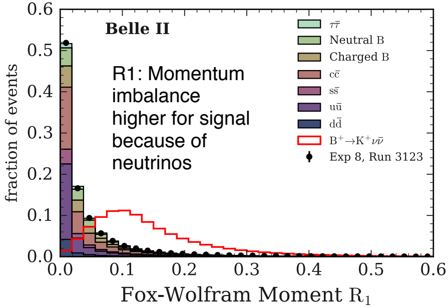


BDT parameters I

To suppress the backgrounds list of potential features (>100) such as:

- variables related to event-shape,



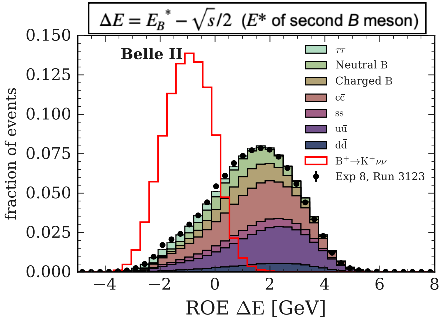


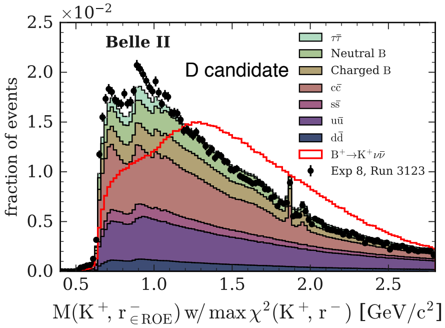
BDT parameters II

To suppress the backgrounds list of potential features (>100) such as:

- variables related to event-shape, ROE-related variables, variables related to the distance wrt to beam spot and tag-vertex, variables related to 2/3-track vertex fits, missing mass ...

51 most discriminating variables w/o loss of performace are chosen as an input to BDTs

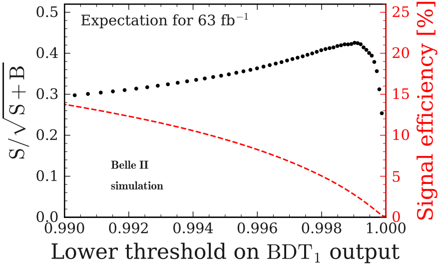


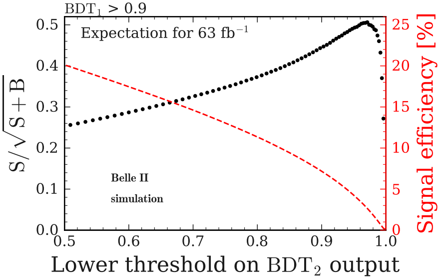


Two-stage BDT

Two consecutive BDTs (BDT_1 and BDT_2) have been trained on simulated subsamples to suppress the backgrounds:

- ▷ BDT_1 trained on the chosen 51 variables on $\sim 10^6$ events for all types of backgrounds and signal
- ▷ BDT_2 is trained with the same set of variables but only on events with $BDT_1 > 0.9$ ($\sim 28\%\ \epsilon_{sig}$)
- ▷ **Boosting of statistics in signal region $\rightarrow$ improvement of signal purity of 35% @ 4% ϵ_{sig}**
- ▷ No overtraining is observed





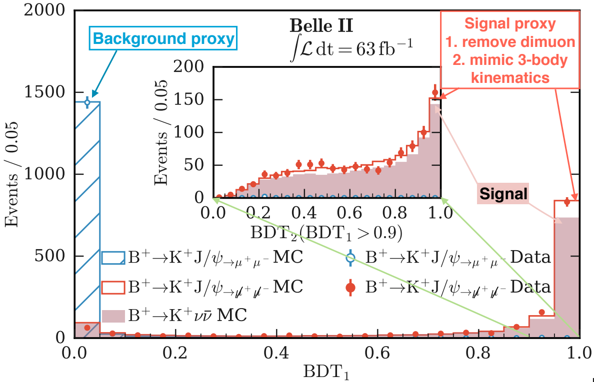
Validation I: $B^+ \to J/\psi(\to \mu^+\mu^-)K^+$

BDT_1 and BDT_2 validated with data/MC comparison using $B^+ \to J/\psi(\to \mu^+\mu^-)K^+$

- ▷ Used because of high BF and clean signature
- ▷ Validation for both signal and B-backgrounds !
- ▷ **Excellent agreement** $\to$ for $\mathrm{BDT}_2 > 0.95$, data/MC = 1.06 ± 0.10

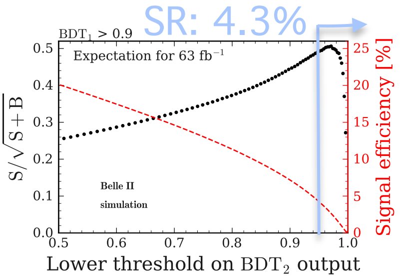
Signal-like $B^+ \to J/\psi(\to \mu^+\mu^-)K^+$

- ▷ 0. Reconstruct $B^+ \to J/\psi(\to \mu^+\mu^-)K^+$
- ▷ 1. Ignore dimuon from J/ψ to mimic missing energy
- ▷ 2. Replace four-momenta of K^+ by that of the signal to mimic 3-body kinematics

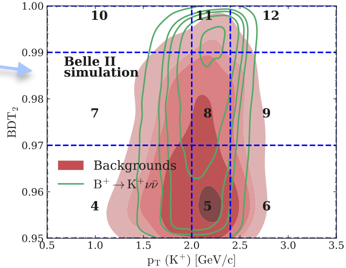


Fit Region Definition

- Signal region: maximum sensitivity $\rightarrow$ $BDT_2 > 0.95$ $\rightarrow$ **4.3%** ϵ_{sig}
- In SR, kaon PID > 0.9 $\rightarrow$ keep 62% kaons, remove 97% pions
- **24 bins in $p_T \times BDT_2$ space**
  (12 bins on-resonance + 12 bins off-resonance)
- Bin boundaries determined from 2D grid optimisation



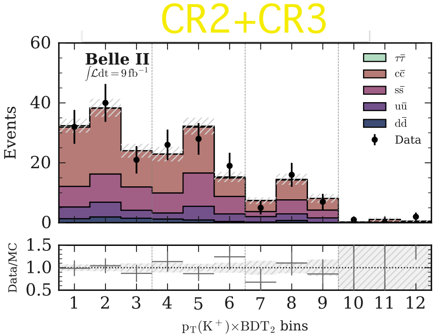
Region	2D Bin Boundary Definition	Physics Processes	$\sqrt{s}$
Signal Region (SR)	$p_T(K^+) \in [0.5, 2.0, 2.4, 3.5]$ GeV/c $BDT_2 \in [0.95, 0.97, 0.99, 1.0]$	signal + all backgrounds	$\Upsilon(4S)$
Control Region 1 (CR1)	$p_T(K^+) \in [0.5, 2.0, 2.4, 3.5]$ GeV/c $BDT_2 \in [0.93, 0.95]$	signal + all backgrounds	$\Upsilon(4S)$
Control Region 2 (CR2)	$p_T(K^+) \in [0.5, 2.0, 2.4, 3.5]$ GeV/c $BDT_2 \in [0.95, 0.97, 0.99, 1.0]$	continuum backgrounds	off-resonance (-60 MeV/c^2)
Control Region 3 (CR3)	$p_T(K^+) \in [0.5, 2.0, 2.4, 3.5]$ GeV/c $BDT_2 \in [0.93, 0.95]$	continuum backgrounds	off-resonance (-60 MeV/c^2)

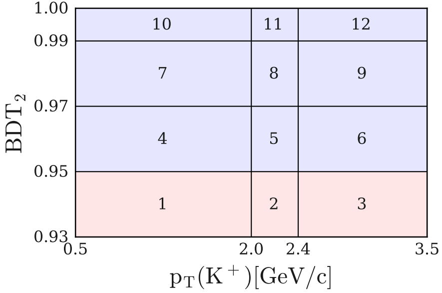


Validation II: Continuum

Data/MC comparison between off-resonance data and continuum simulation in $p_T \times \mathrm{BDT}_2$ bins

- **Very good agreement in shape**
- **Discrepancy in scale**: Data/MC factor = 1.40 $\pm$0.12
- Introduction of normalisation uncertainty of 50% to all the backgrounds (conservative)





Statistical Model

- Likelihood implemented within `pyhf` package
- Cross-check with `sghf`: simplified Gaussian model
- **Inclusion of systematics in the model via nuisance parameters:** background normalisation uncertainty, tracking inefficiency, neutral energy miscalibration for photons, neutral energy miscalibration for unmatched photons, uncertainty on PID correction due to limited statistics, uncertainty on branching fractions of leading bkg processes, uncertainty on SM form factor
- All 7 background samples considered separately: mixed *B*, charged *B*, $c\bar{c}$, $u\bar{u}$, $s\bar{s}$, $d\bar{d}$, $\tau^+\tau^-$
- Total number of fit parameters: 175 nuisance parameters ($\vec{\phi}$) and 1 parameter of interest (signal strength=μ)
- **1 μ = SM BF = $(4.6 \pm 0.5) \times 10^{-6}$**

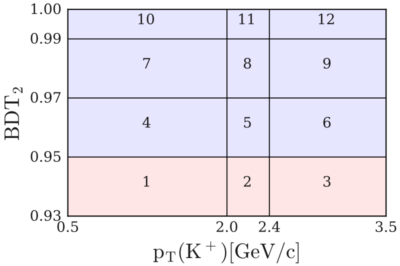
$$f(n, a \mid \eta, \chi) = \prod_{r \in \text{regions}} \prod_{b \in \text{bins}} \text{Pois}(n_{rb} \mid \nu_{rb}(\eta, \chi)) \prod_{\chi} c_\chi(a_\chi \mid \chi)$$

Simultaneous measurements of multiple regions

Constraints

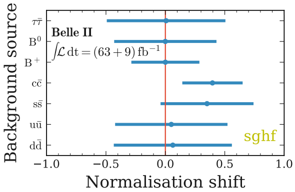
η = parameter of interest
χ = nuisance parameters

on-resonance off-resonance
Regions = {SR, CR1, CR2, CR3}

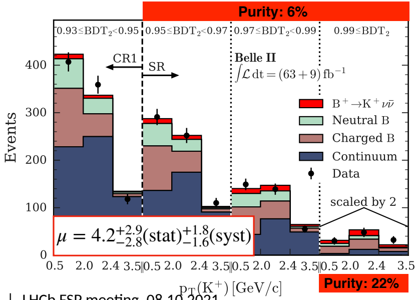


Fit To Data

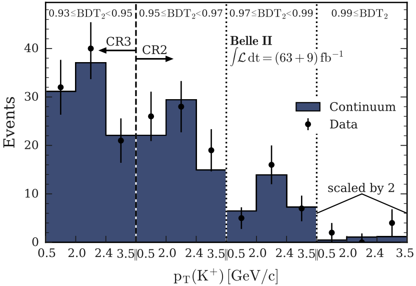
- Binned simultaneous ML fit to data to extract signal strength μ
- **Result:** $\mu = 4.2^{+2.9}_{-2.8}(\text{stat})^{+1.8}_{-1.6}(\text{syst}) = 4.2^{+3.4}_{-3.2}$
- Continuum bkgs pulled up by up to 40%, *B*-bkgs stay the same



On-resonance data

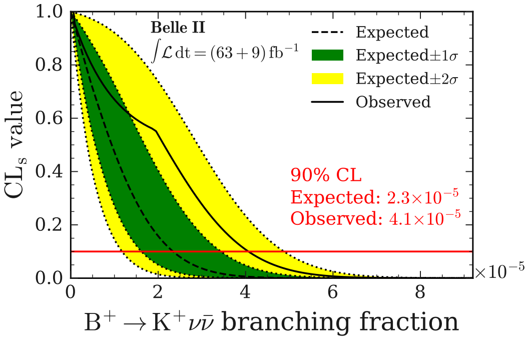


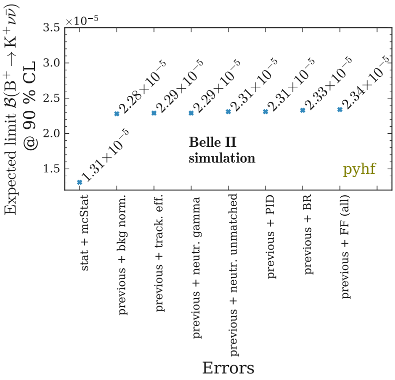
Off-resonance data



Limit Setting

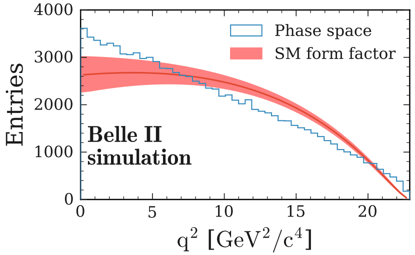
- As no significant signal is observed $\rightarrow$ limit setting
- Use both pyhf and sghf to compute a limit $\rightarrow$ consistent results
- **Result: $\mathcal{B}(B^+ \rightarrow K^+\nu\bar{\nu}) < 4.1 \times 10^{-5}$ @ 90 CL**
- Leading systematic: background normalisation

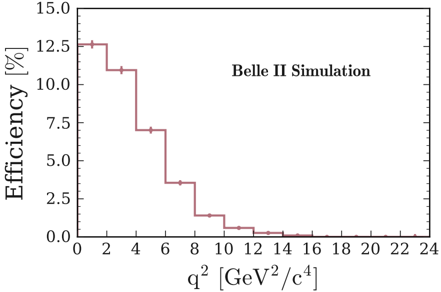




Reinterpretation

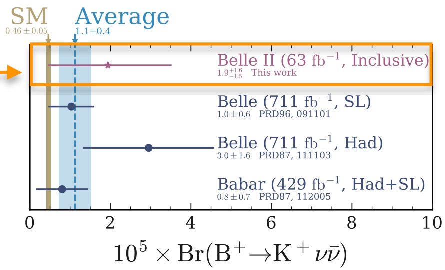
- Publish ϵ_{sig} as a function of $q^2(\nu\bar{\nu})$
- Reminder: default signal model $\rightarrow$ PHSP model with SM form factor reweighting [arXiv:1409.4557]
- **At low q^2 maximum signal efficiency of $\sim$ 13%, but no sensitivity for $q^2 > 16$ GeV^2/c^2**





Comparison with Other Measurements

- **Competitive limit**
- **Comparison with other experiments** via σ_{BR} assuming same luminosity $\rightarrow$ the performance of inclusive tag:
  - 3.5 better than hadronic tag
  - 20% better than semileptonic tag
  - 10% better than combined hadronic and semileptonic tag



Experiment	Year	Observed limit on $BR(B^+ \rightarrow K^+\nu\bar{\nu})$	Approach	Data [fb^{-1}]
BABAR	**2013**	$< 1.6 \times 10^{-5}$ **[Phys.Rev.D87,112005]**	**SL + Had tagging**	429
Belle	**2013**	$< 5.5 \times 10^{-5}$ **[Phys.Rev.D87,111103(R)]**	**Had tagging**	711
Belle	**2017**	$< 1.9 \times 10^{-5}$ **[Phys.Rev.D96,091101(R)]**	**SL tagging**	711
Belle II	**2021**	$< 4.1 \times 10^{-5}$	**Inclusive tagging**	63

Conclusion and Prospects

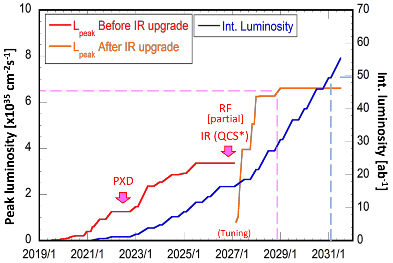
- **Paper got recently accepted by PRL :)** Paper link: arxiv:2104.12624
- **Bigger dataset** (+ possible combination with Belle dataset)
- **Attacking biggest systematic** (background normalisations, e.g continuum modelling)
- **More channels** (K^*, K_S^0, K^{*+}...)
- **Possible improvement in background suppression** (use of NN architecture, discriminating vars)
- **Combined analysis of inclusive and exclusive tagged events**

Aim: 50 ab-1

$10^5 \times \sigma_{\rm BR}$ uncertainty for next analyses, assuming 25% improvement + 40% K_S^0

	$63\ \text{fb}^{-1}$ (arXiv:2104.12624)	$197\ \text{fb}^{-1}$ (Summer 2021 – current lumi)	$450\ \text{fb}^{-1}$ (Summer 2022 – expected)	$(450 + 700)\ \text{fb}^{-1}$ (+ Belle I sample)
$\sigma_{BR}(K^+)$	1.55	0.78	0.52	**0.32**
$\sigma_{BR}(K^+ + K_S^0)$	–	0.68	0.45	**0.28**

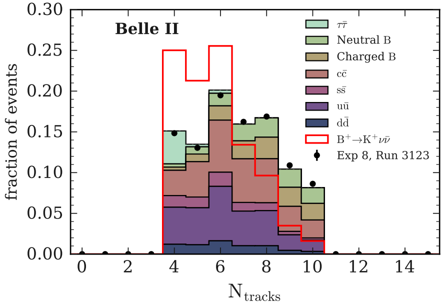
Preliminary



BACKUP

Basic Event Selection

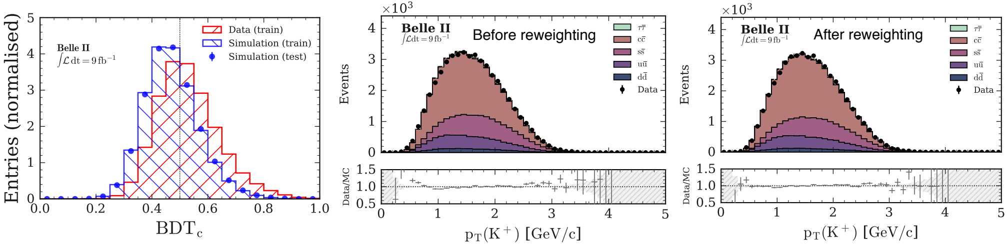
- Track cleanup: $p_T > 0.1$GeV/c, $\theta \in$ CDC, $|dr| < 0.5$cm, $|dz| < 3.0$cm, E < 5.5 GeV
- Photon cleanup: E > 0.1 GeV, $\in$ CDC, E < 5.5 GeV
- Other loose preselection to reject low-multiplicity background:
  - $4 \leq \texttt{nTracksCleaned} \leq 10$
  - $0.3 < \theta(\mathbf{p}_{miss}) < 2.8$ rad
  - Visible E in CMS frame> 4GeV



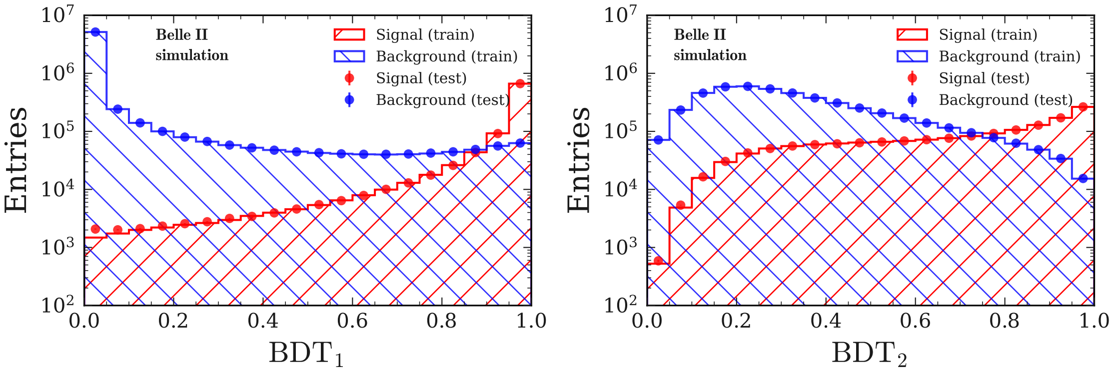
Continuum Modelling Improvement

Additional BDT$_c$ is trained on events with BDT$_1 > 0.9$ in order to correct mismodeling of continuum simulation:

- ▷ Signal = off-resonance data , background = continuum simulation
- ▷ Continuum simulation events are reweighted with $\frac{p}{1-p}$, where p = BDT$_c$ output
- ▷ Method taken from here: J. Phys.: Conf. Ser. 368 012028

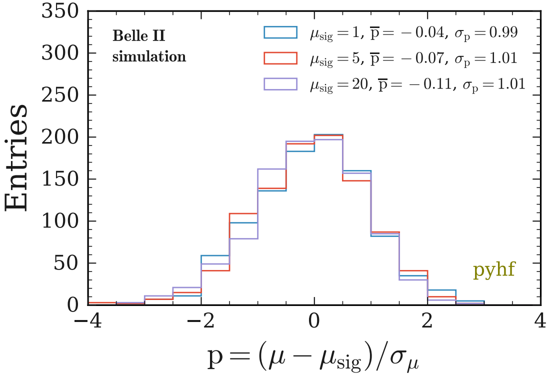


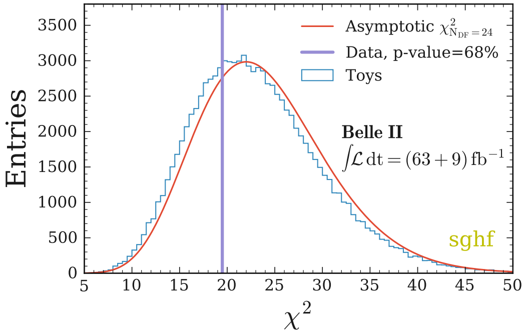
Overtraining



Fit Validation

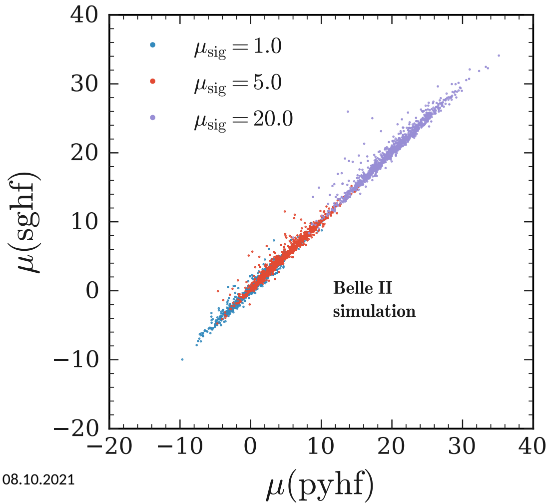
- Test with injected signal $\rightarrow$ check pulls = $\frac{\mu_{fit} - \mu_{inj}}{\sigma_\mu}$ for 1, 5, 20 $\times$ signal
- Test the fit quality $\rightarrow$ high *p*-value, good agreement with χ^2 distribution





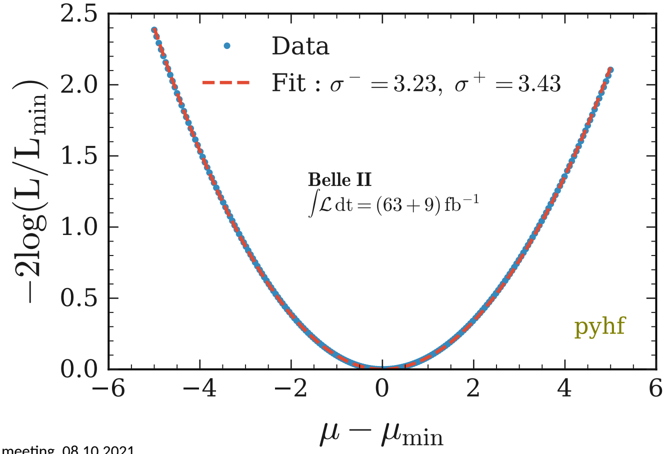
pyhf versus sghf

- Check correlation between pyhf and sghf fitted μ for 1, 5, 20 $\times$ signal $\rightarrow$ very good correlation



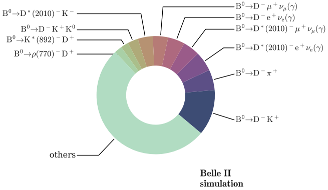
Profile Likelihood Scan

- Assymetric uncertainty on signal strength μ estimated by fitting of parabola of the points from profile likelihood scan

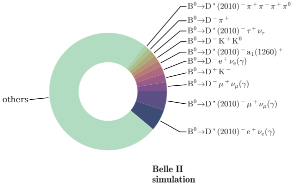


Background Composition in the Fit Region

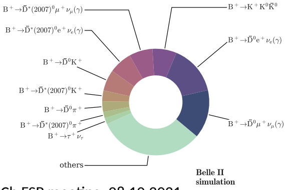
- $B^0\bar{B}^0$ signal side:



- $B^0\bar{B}^0$ tag side:



- B^+B^- signal side:



- $B^+\bar{B}^-$ tag side: