$b \to s\tau\tau$ at Belle II

Gaetano de Marino*,
on behalf of the Belle II Collaboration

THE τ-RRIFIC PENGUINS WORKSHOP

Imperial college, London

*gaetano.demarino@ijs.si

terrific adjective

ter·rif·ic tə-ˈri-fik

OED | Oxford English Dictionary

1. Causing terror, terrifying; terrible, frightful; stirring, awe-inspiring; sublime. Now *rare*.
   Show quotations
   Cite Historical thesaurus ▾

2.a. Of great size or intensity; excessive; very severe. Cf. **tremendous** *adj.* 2a.
   Show quotations
   Cite Historical thesaurus ▾

2.b. As an enthusiastic term of commendation: amazing, impressive; excellent, exceedingly good, splendid.
   Show quotations
   Cite Historical thesaurus ▾

Merriam Webster

1 : unusually fine : **MAGNIFICENT**
  terrific weather

2 : **EXTRAORDINARY**
  terrific speed

3 a : exciting or fit to excite fear or awe
  a *terrific* thunderstorm

  b : very bad : **FRIGHTFUL**

The double nature of $b \to s\tau\tau$ decays

- Intriguing (B)SM tests
- Experimentally very challenging

What do flavor experiments still can say in the pre-FCC era?

Here to give a perspective on *B*-factories

BELLE II AND B-FACTORIES

1. Threshold $B\bar{B}$ production at $\Upsilon(4S)$

→ Two B's and nothing else

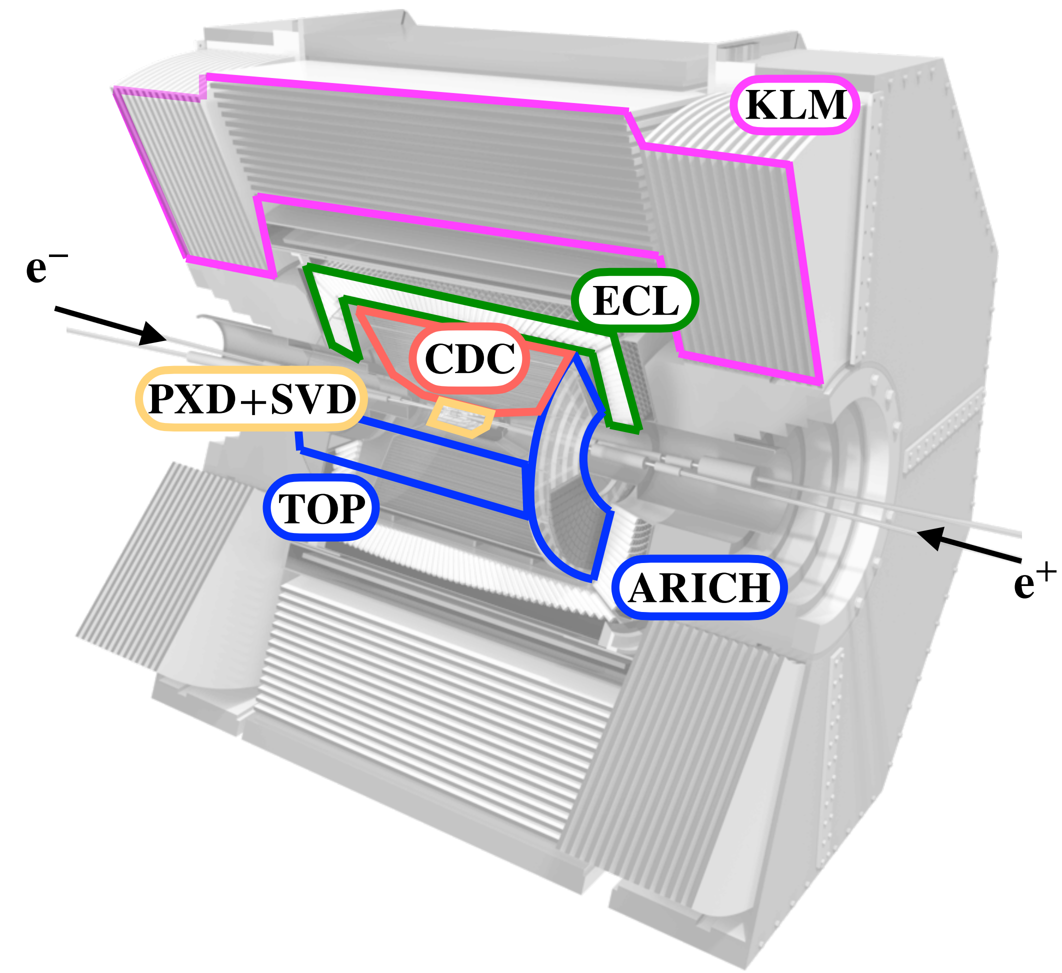
2. Relatively low $q\bar{q}$-background

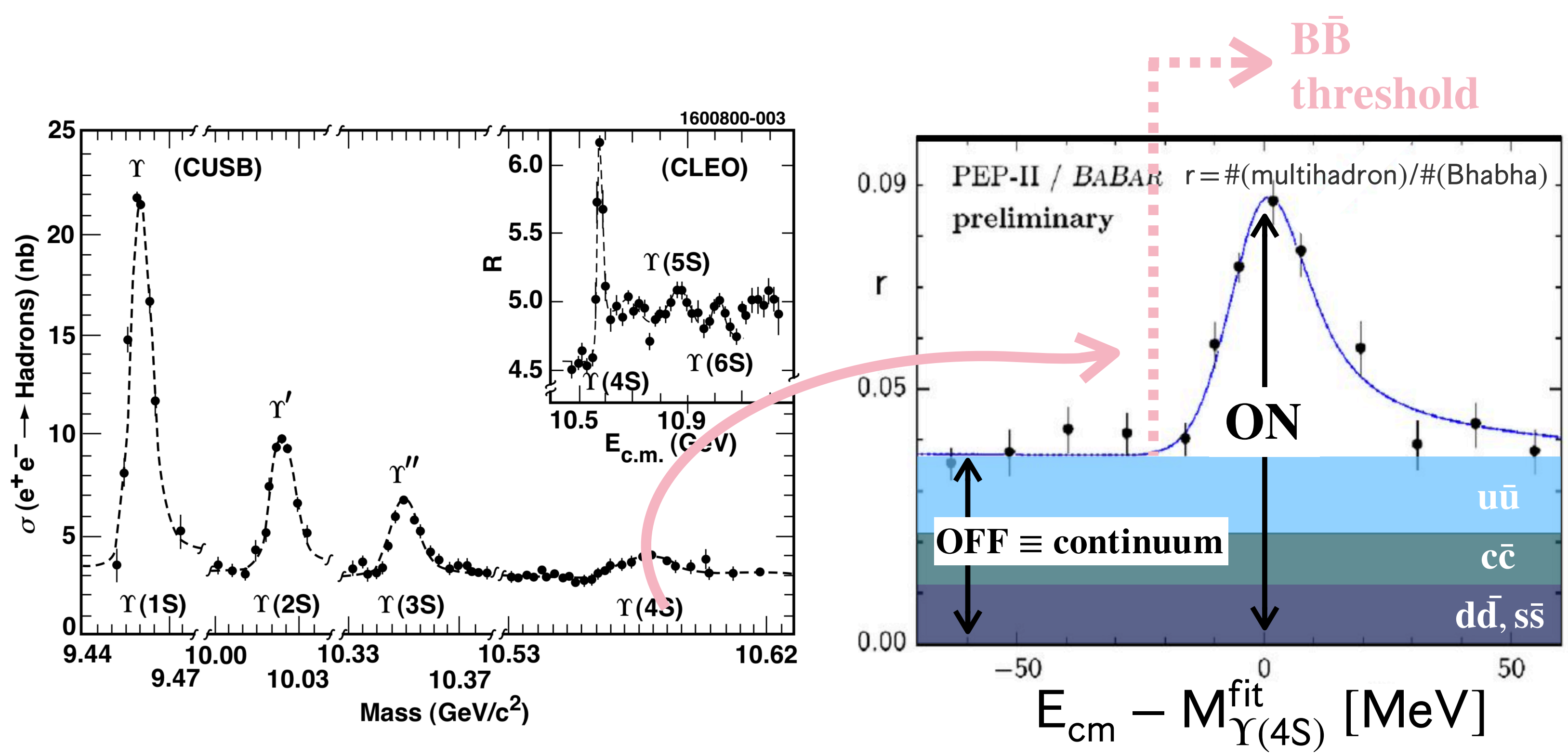
→ Can be calibrated with OFF-resonance data

3. Known initial kinematics

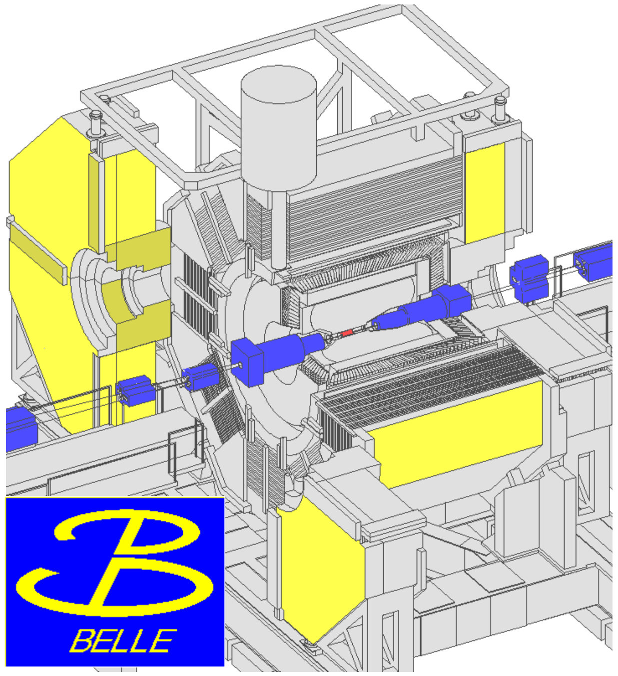
+ almost-4π detector coverage

→ reconstruct final states with ν's

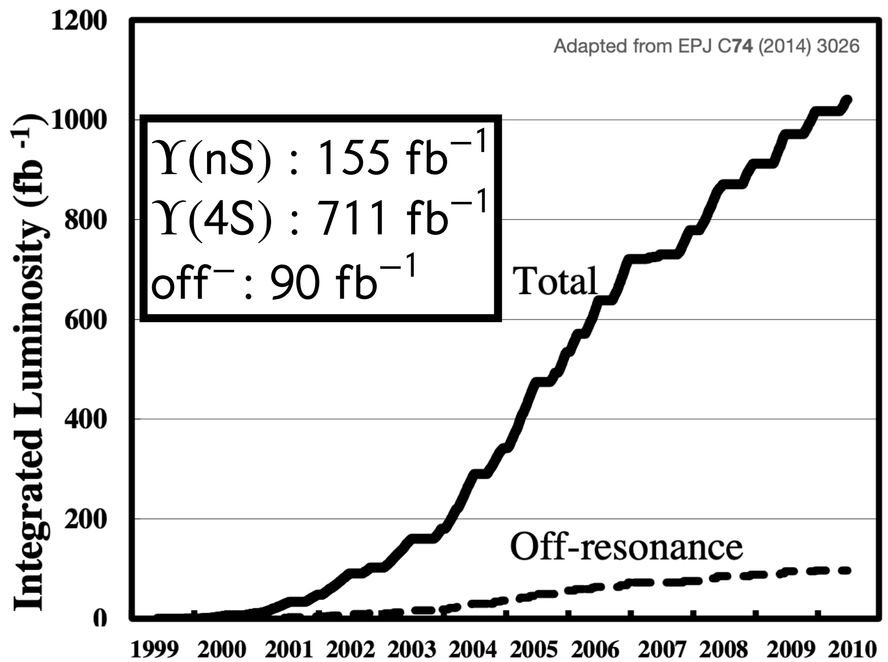




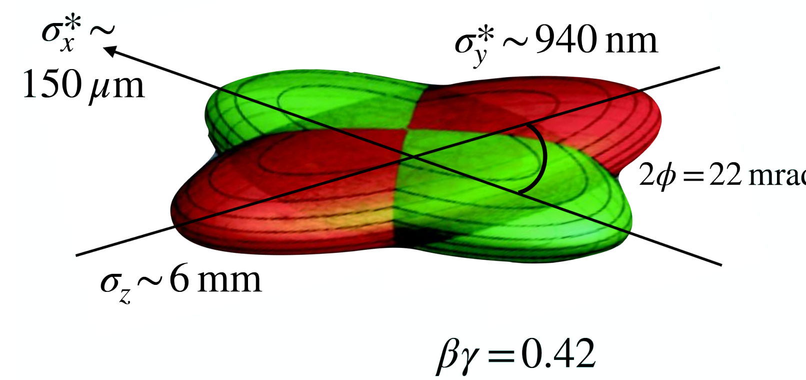
DATASETS

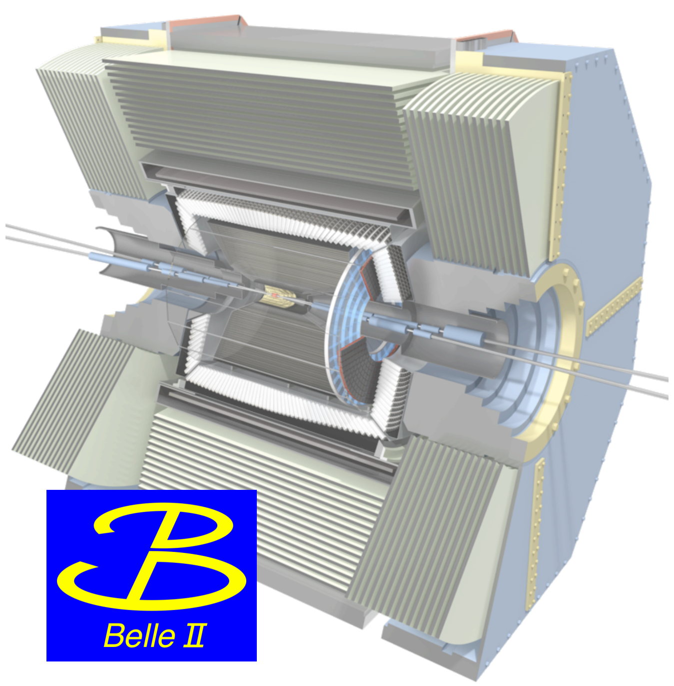


N. of $B_{u/d}/B_s$: 772 M/17 M

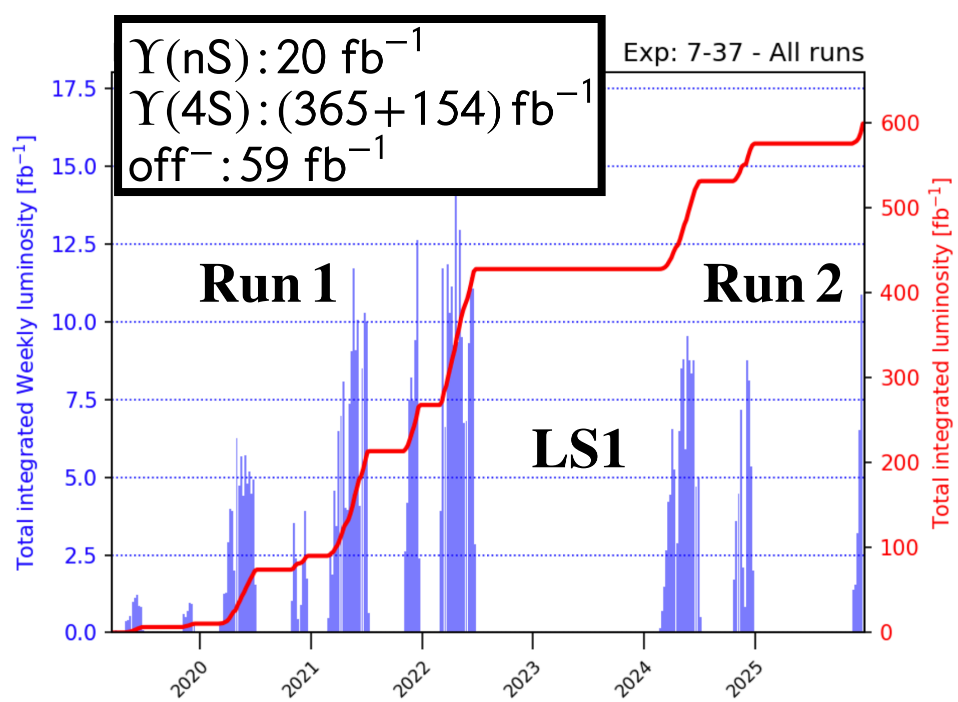


Max luminosity of $2.1{\times}10^{34}$ cm^{-2} s^{-1} (Currents 1.2/1.6 A) (June 2009)

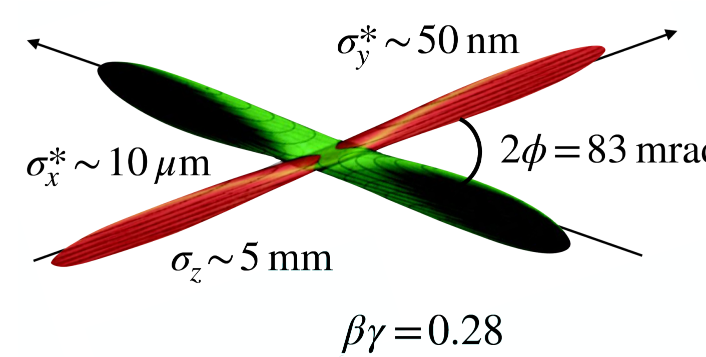




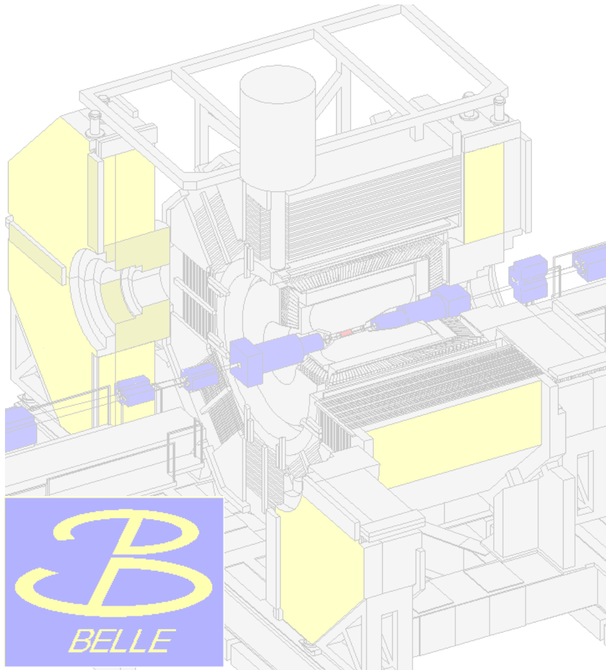
N. of $B_{u/d}/B_s$: 570 M/0 M



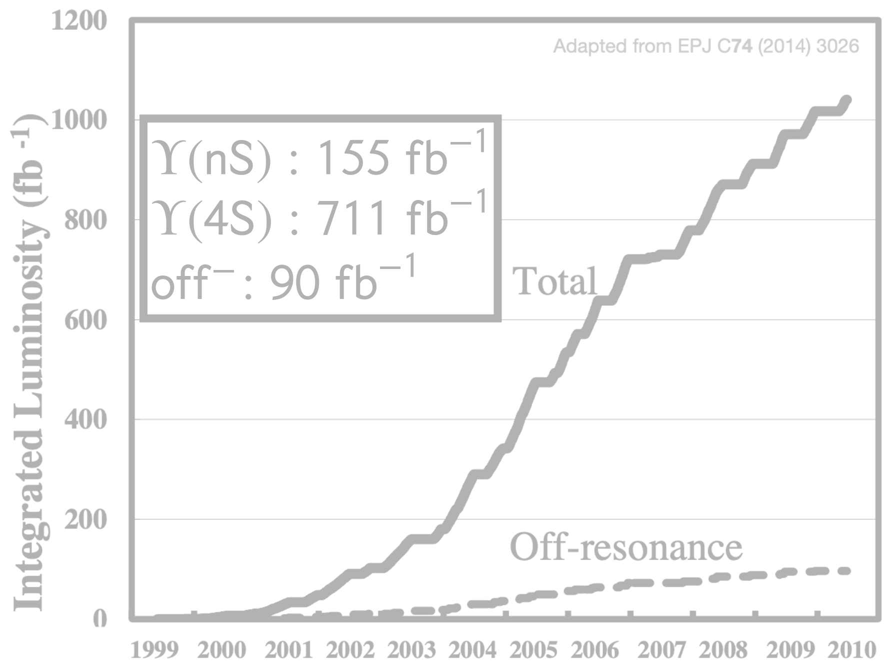
WR luminosity of $5.1{\times}10^{34}$ cm^{-2} s^{-1} (Currents 1.3/1.7 A) (Dec 2024)



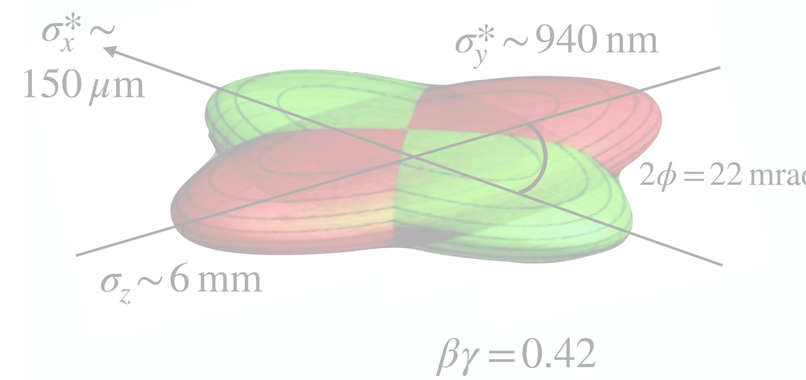
DATASETS

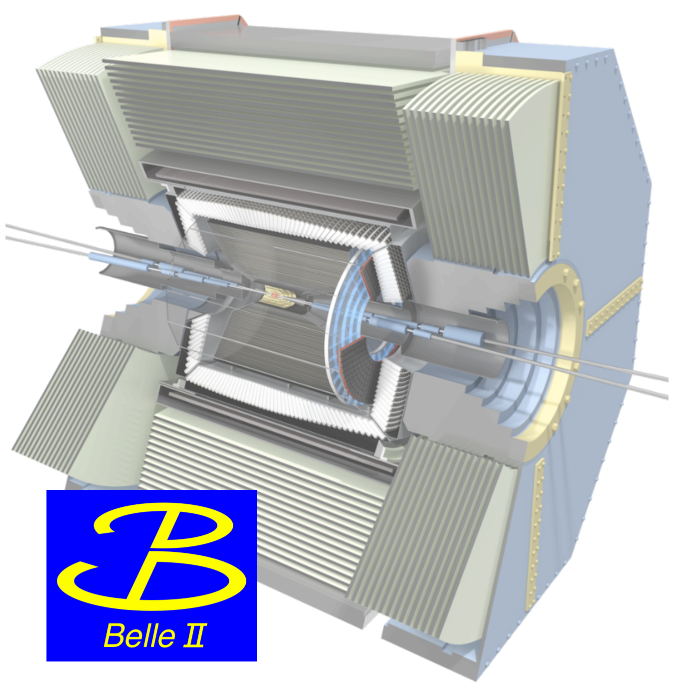


N. of $B_{u/d}/B_s$: 772 M/17 M

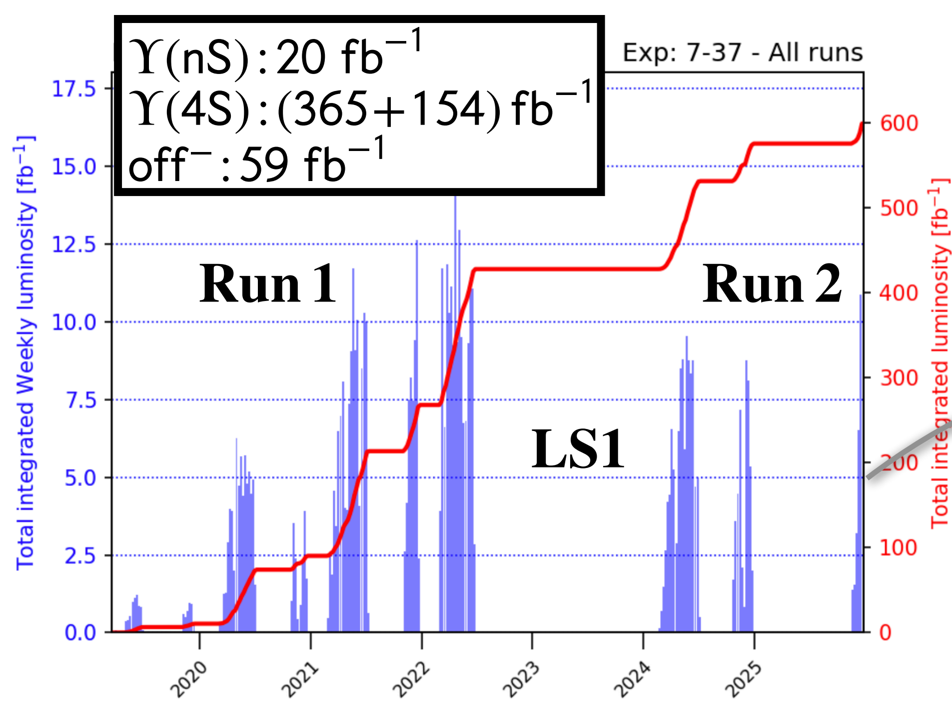


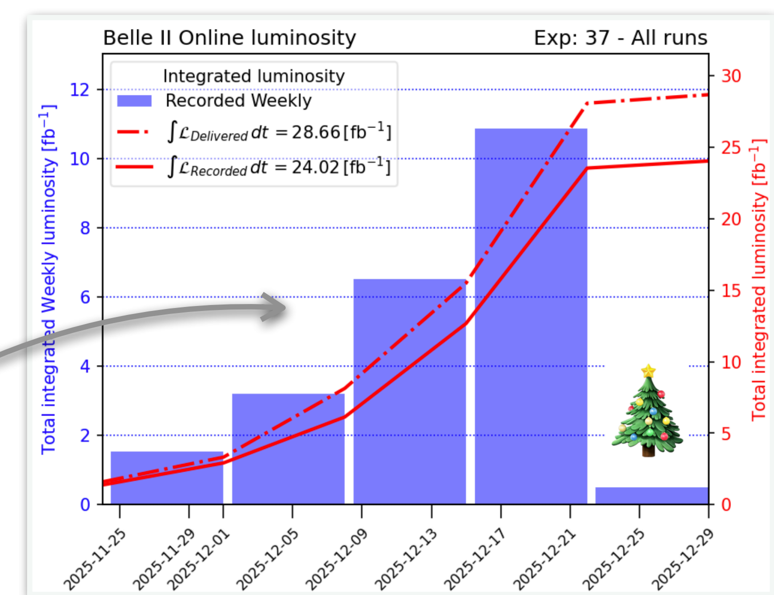
Max luminosity of 2.1×10^{34} cm^{-2} s^{-1}
(Currents 1.2/1.6 A) (June 2009)





N. of $B_{u/d}/B_s$: 570 M/0 M

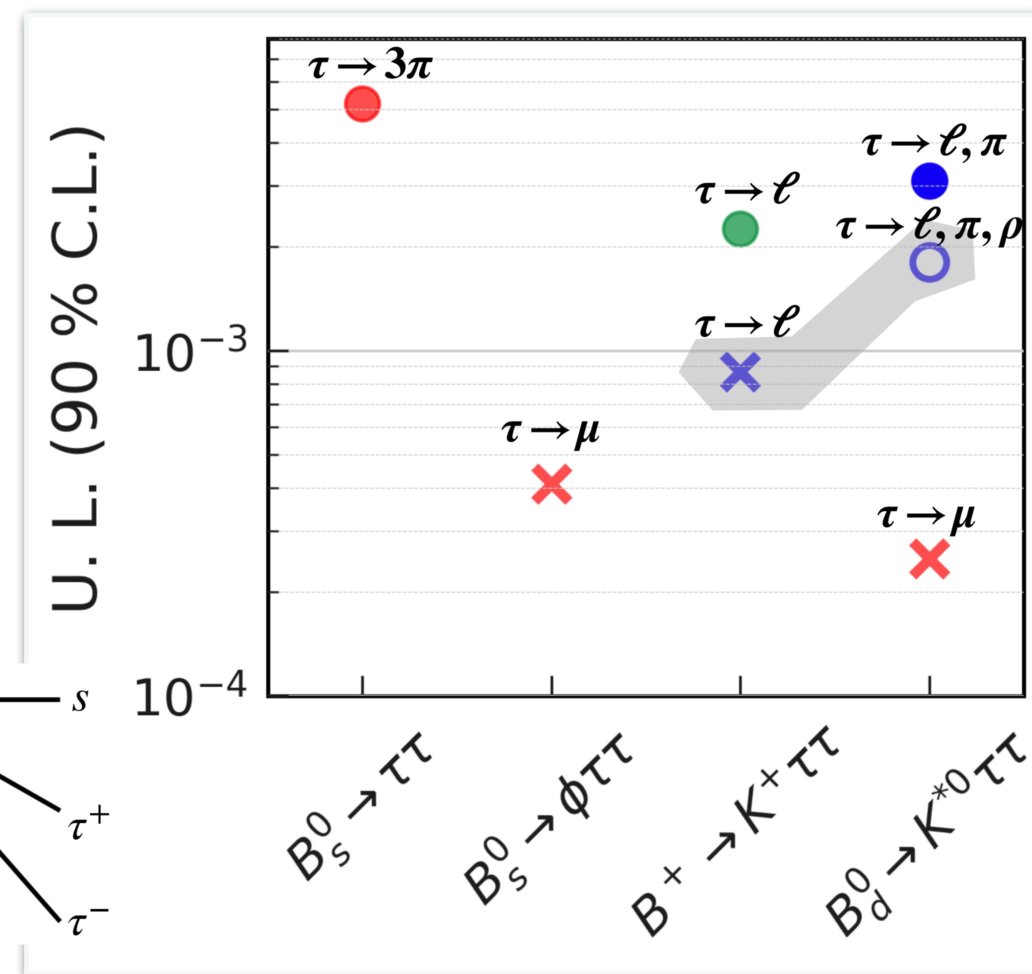


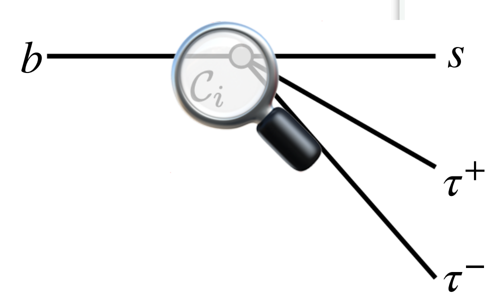


EXPERIMENTAL STATUS

- **Focus on the recent searches for $B^+ \to K^+\tau\tau$ and $B^0 \to K^{*0}\tau\tau$ decays**
  - Results obtained with Run1 (+ Belle) datasets
  - Similar methods were used in previous searches at BaBar and Belle

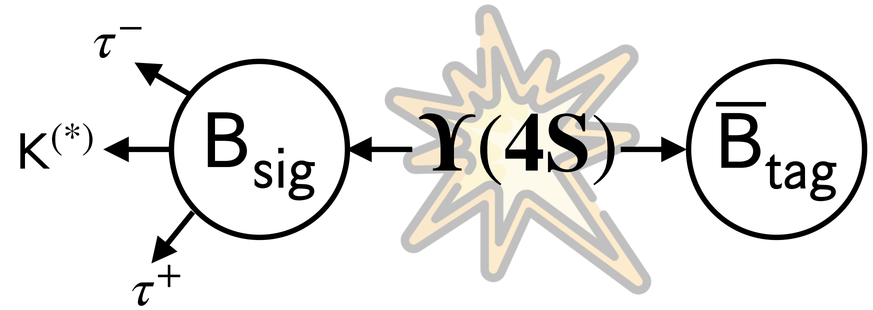
PRL 118 (2017) 25, 251802
PRL 118 (2017) 3, 031802
PRD 108 (2023) 1, L011102
PRL 135, 151801 (2025) HEPData
CKM 2025
2510.13716

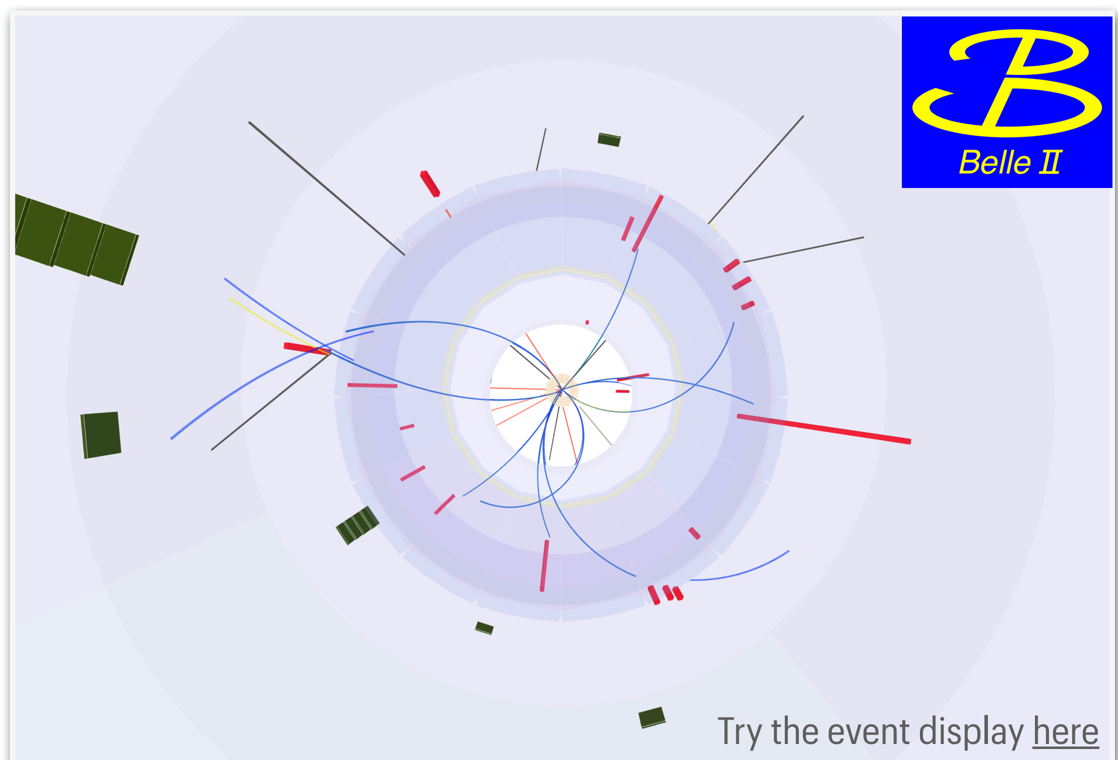




reconstruction

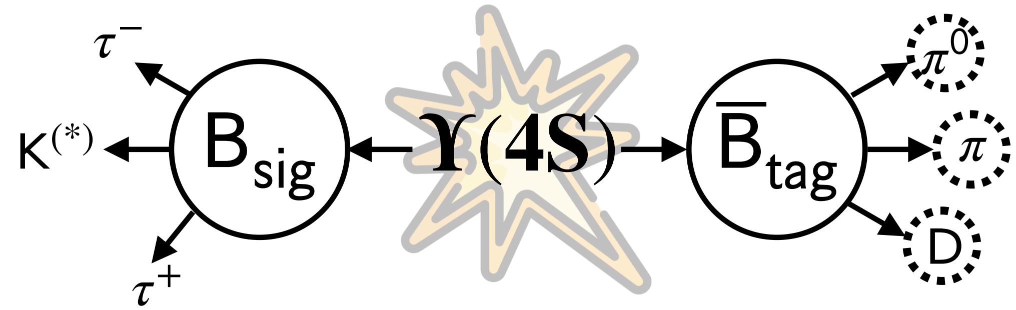
OVERVIEW





Clean environment but low *b*-cross section
Need high performance from all sub-detectors and optimised analysis strategies

BTAG



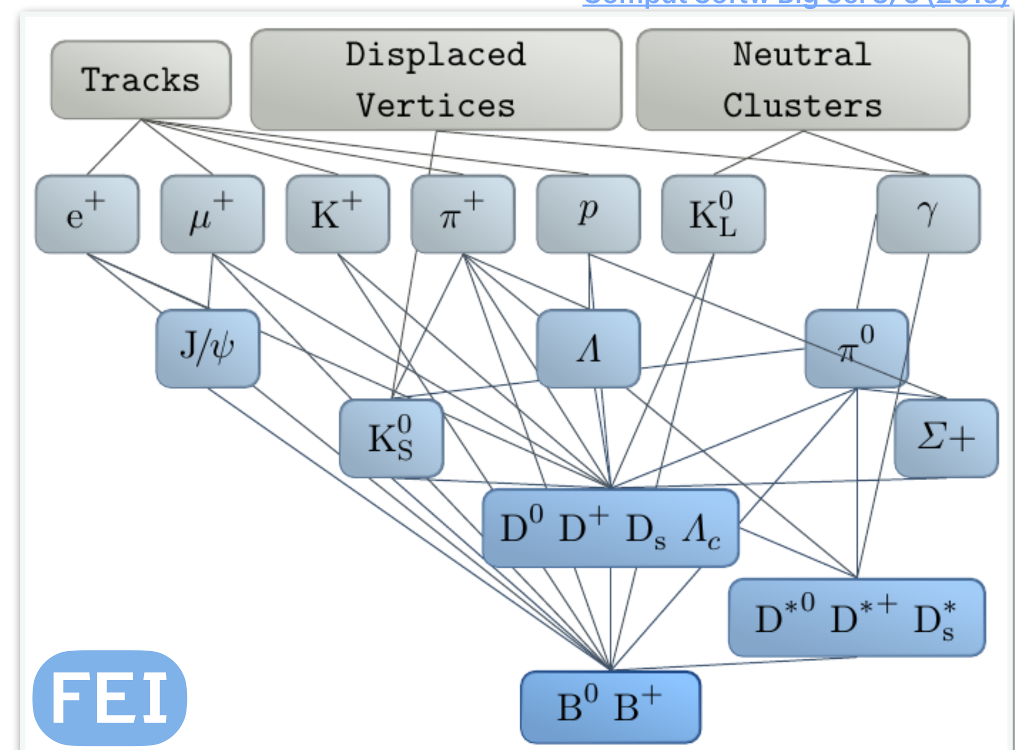
- **Exclusive hadronic B-tagging**
  More effective for modes lacking clear signatures
  → **employed for all $B \to K\tau\tau$ searches at B-factories so far**

  Used for:
  - background filtering
  - flavor information
  - full kinematic info

  At Belle II we use a multi-variate based algorithm for B_{tag} reconstruction to increase the efficiency $\varepsilon_{tag} \sim 0.5\ \%$ for purity $\pi_{tag} \sim 50\ \%$

  The quality of the B-candidate tuned with the BDT output called $\mathscr{P}_{FEI}$

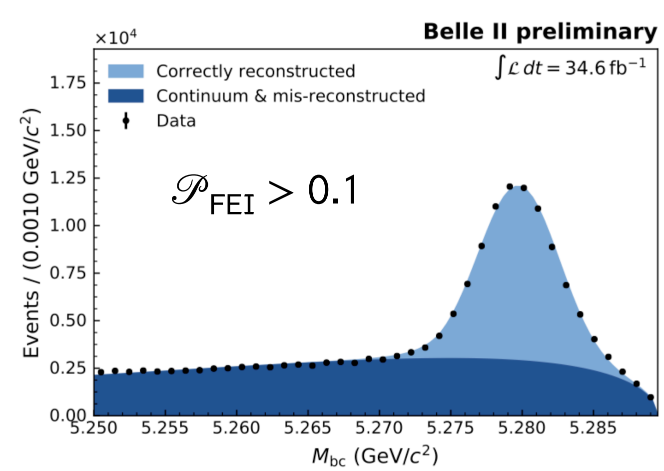
Comput Softw Big Sci 3, 6 (2019)

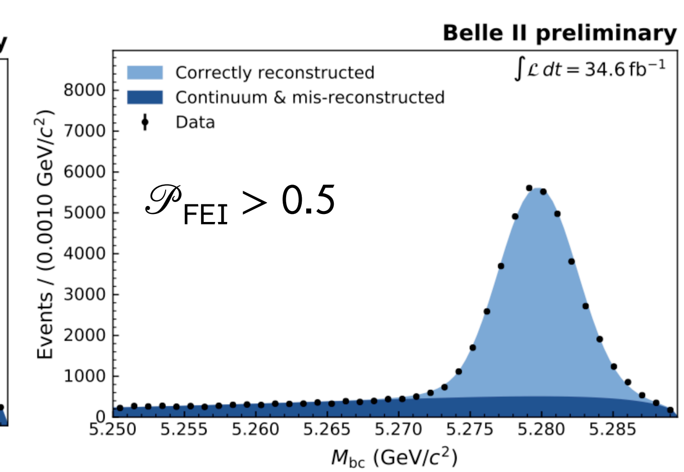


*Also configurable for semileptonic decays

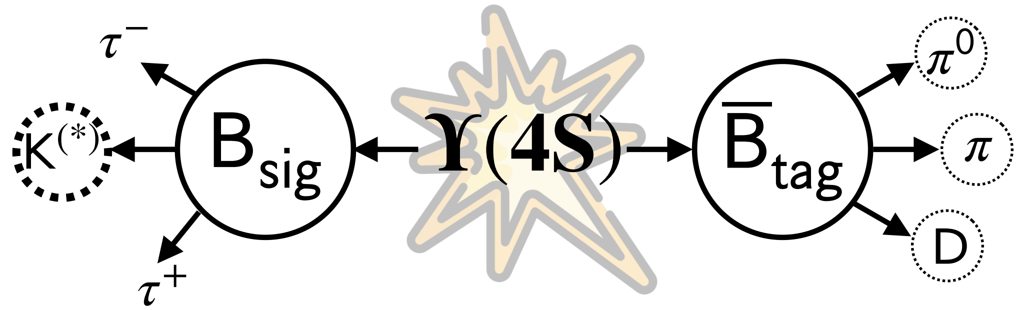
$$M_{bc} = \sqrt{(E^*_{beam}/c^2)^2 - (p^*_B/c)^2}$$

Usually work with less pure working points

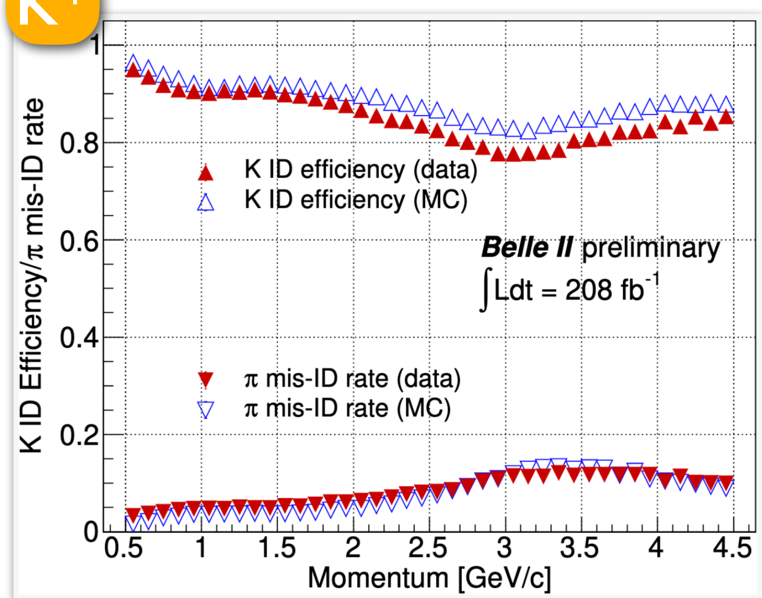




BSIG I

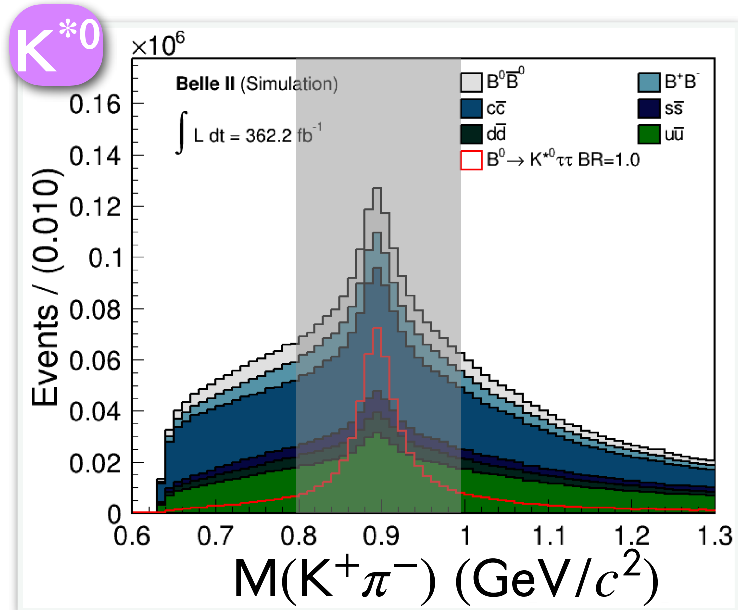


K^+



$\varepsilon \sim 90\,\%$ @ $\pi \to K \sim 6\,\%$
with likelihood based selector
(NN-based in future)

K^{*0}

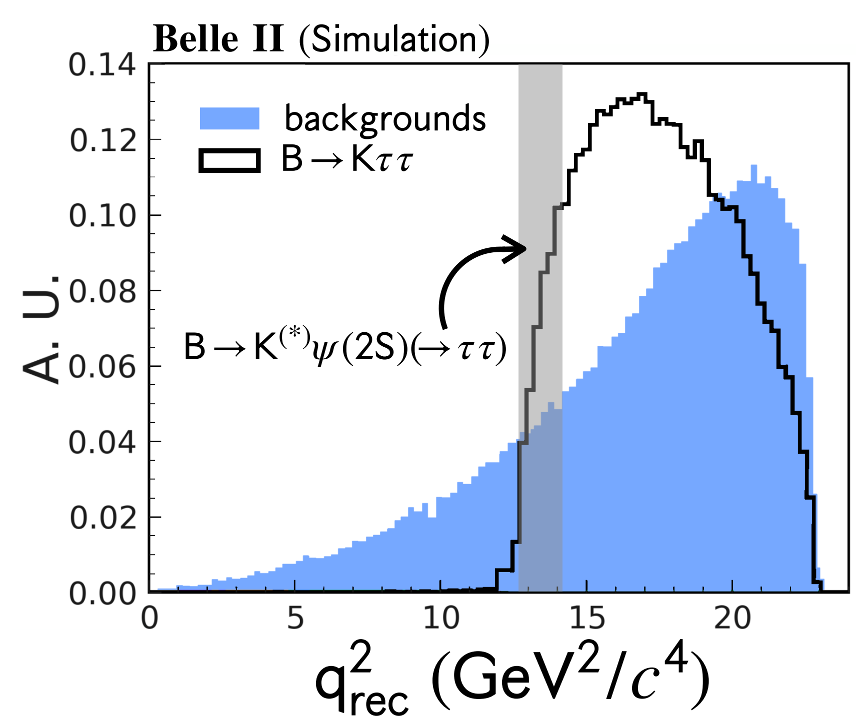


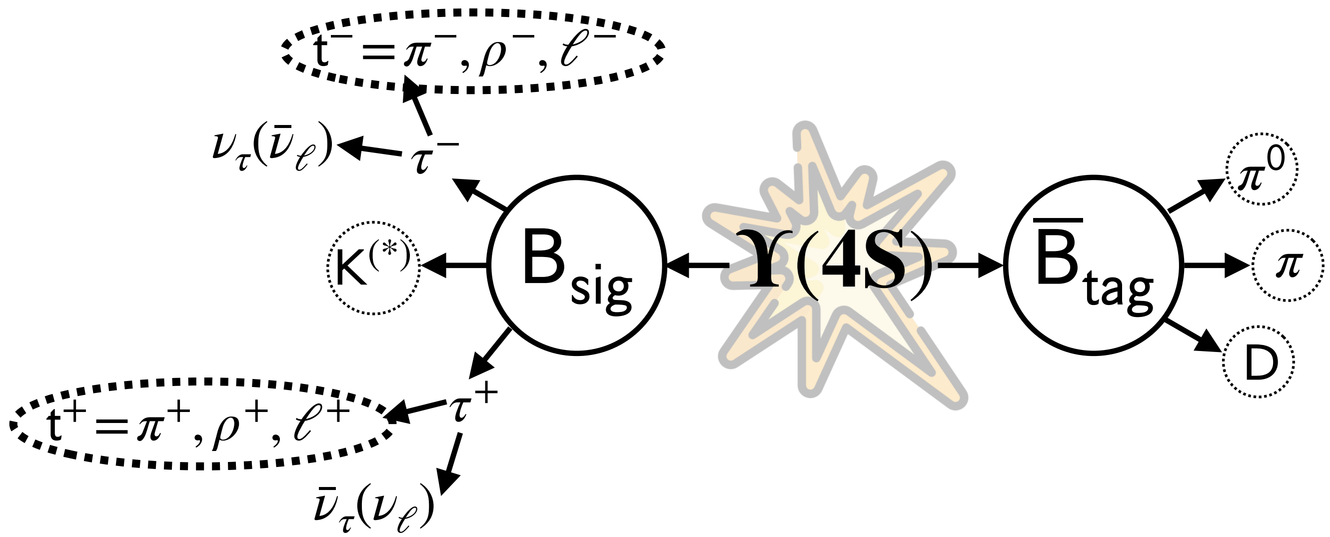
Focus on resonant $K\pi$ only

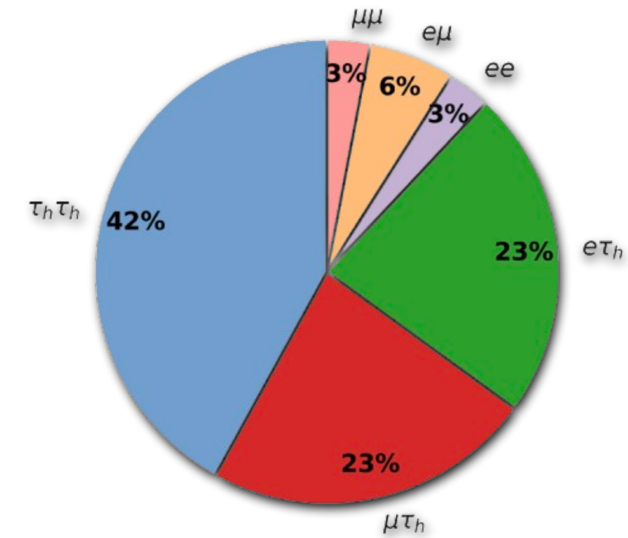
$$q^2_{rec} = m^2_B + m^2_K - 2\left(E^*_B E^*_K + p^*_B p^*_K \cos\theta\right)$$

excellent resolution~ 0.07 GeV² for good tags

- q^2_{rec} **used to**
  - suppress background via BDT K^{*0}
  - remove out-of-phase-space events + resonant contribution
    $q^2_{rec} > 14.18\ \mathrm{GeV}^2/c^4$ (K^+) K^+







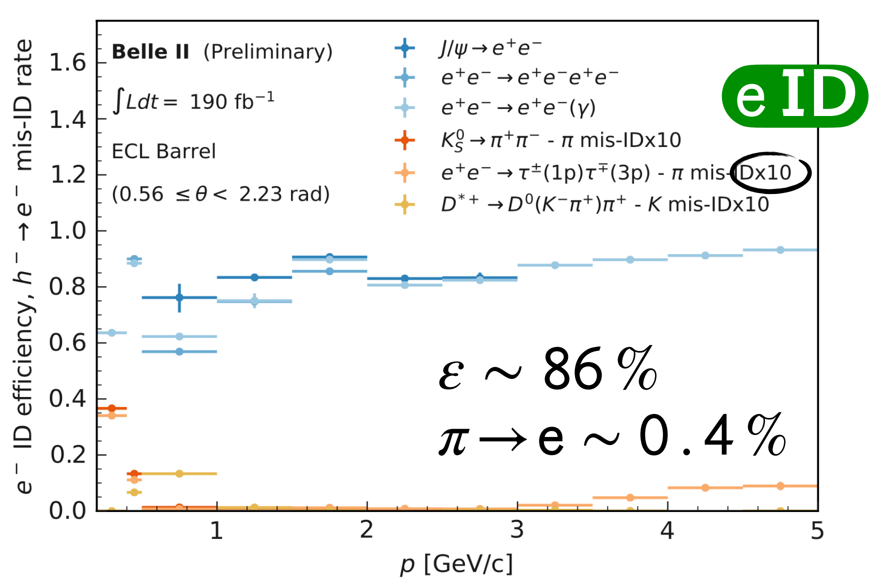
- $\ell\ell$ 0.12
- $\ell\pi$ 0.08
- $\ell\rho$ 0.18
- $\pi\pi$ 0.01
- $\rho\pi$ 0.06
- $\rho\rho$ 0.06

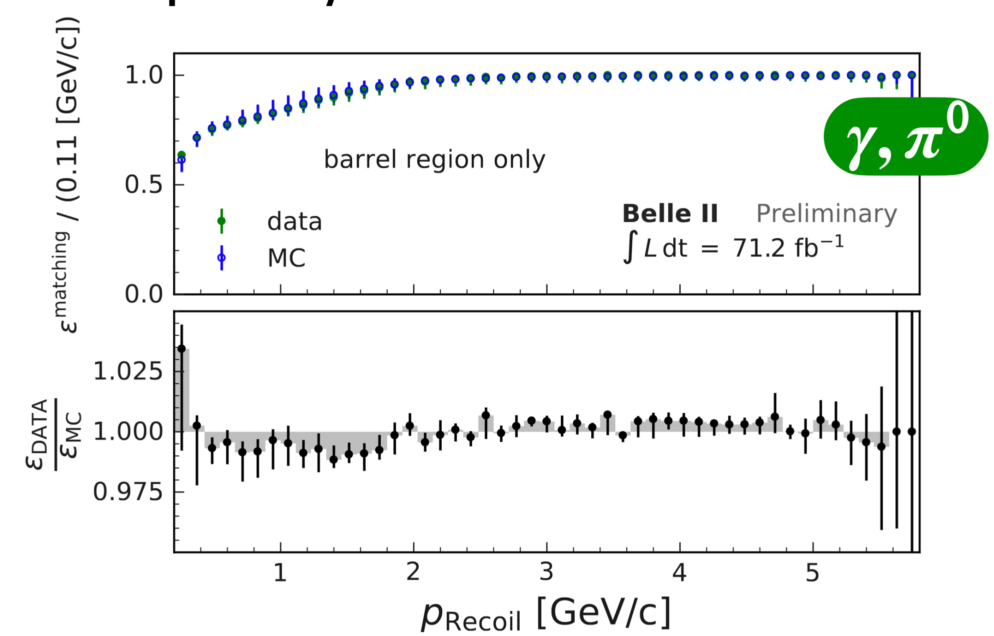
} 0.51

- **Remove events compatible with $\tau \rightarrow$ 3-prong decays**
  The larger multiplicity brings more combinatorial without good vertex information, contrarily to LHCb
- **Focus on**

K^+ leptonic decays

K^{*0} 1-prong decays → leptonic and hadronic final states are treated separately



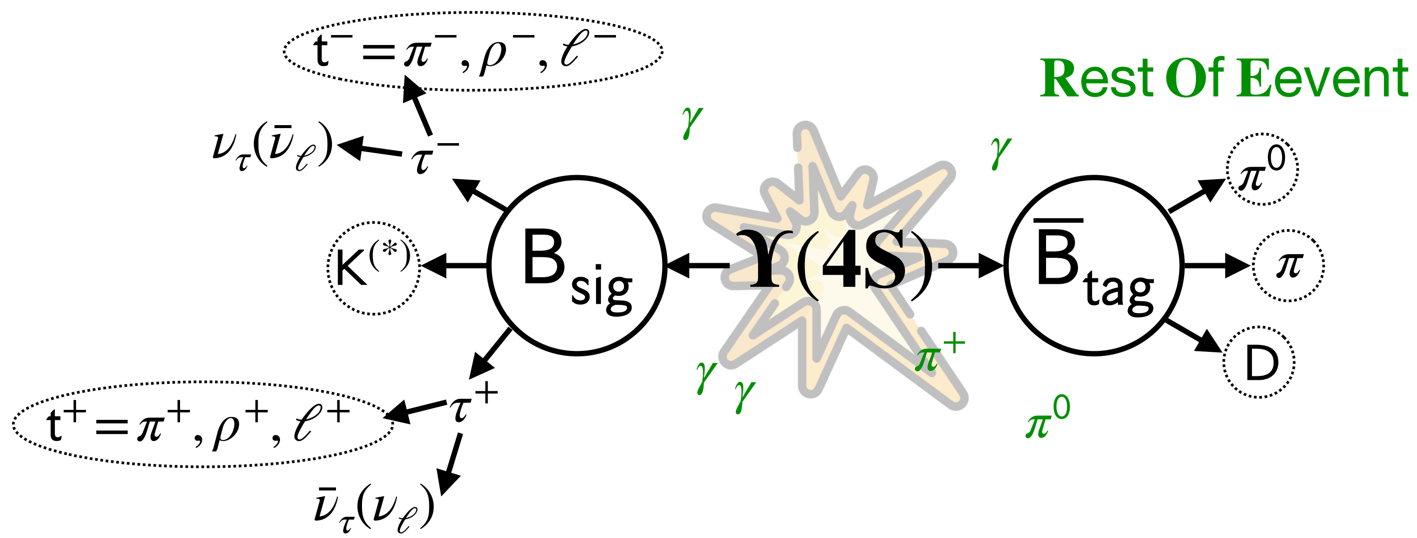


Good lepton ID and similar e-μ efficiency
At ~90% eff., μID has larger pion fake rate

High photon efficiency > 90% (p>1.5 GeV/c)
Resolutions of 7.7% (2.2%) at 100 MeV (1 GeV)

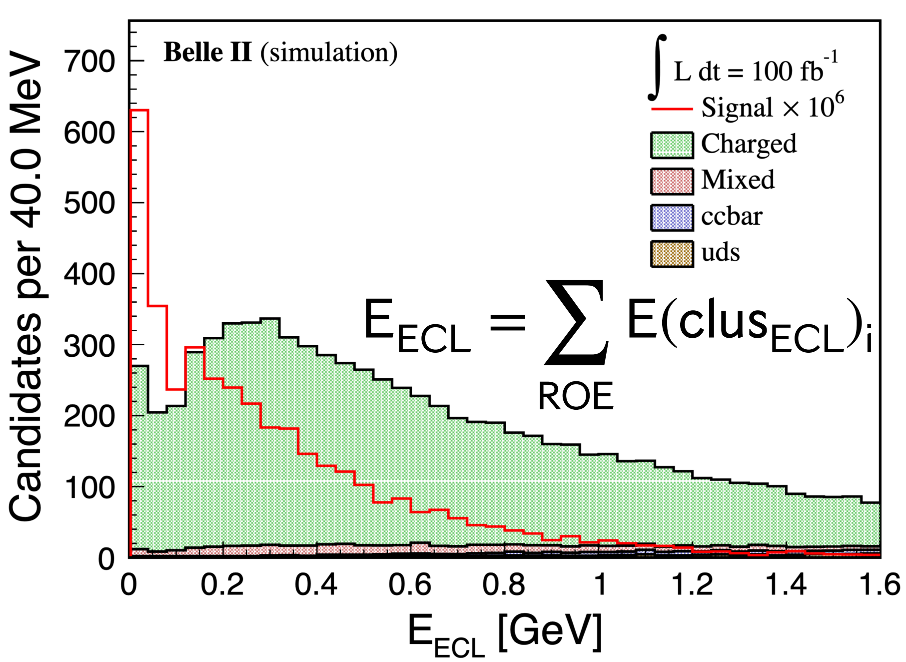
JINST 15 C10016

ROE



In tagged analyses, events with neutral extra-mesons (π^0, K_S) are usually rejected

The total energy of the extra-photons (ECL clusters) is one of the most discriminating variables



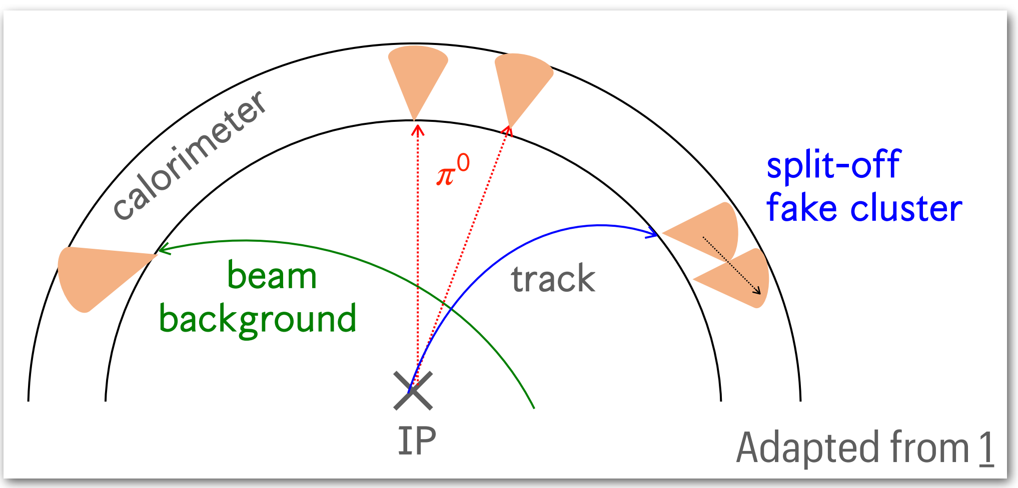
- **No extra activity in ECL expected for signal events**
  Tails due to mis-reconstruction and ECL energy resolution
- **Background has larger ECL deposits**
  Events at ~0 due to energy threshold adopted in the extra-photon selection
- **Used for**

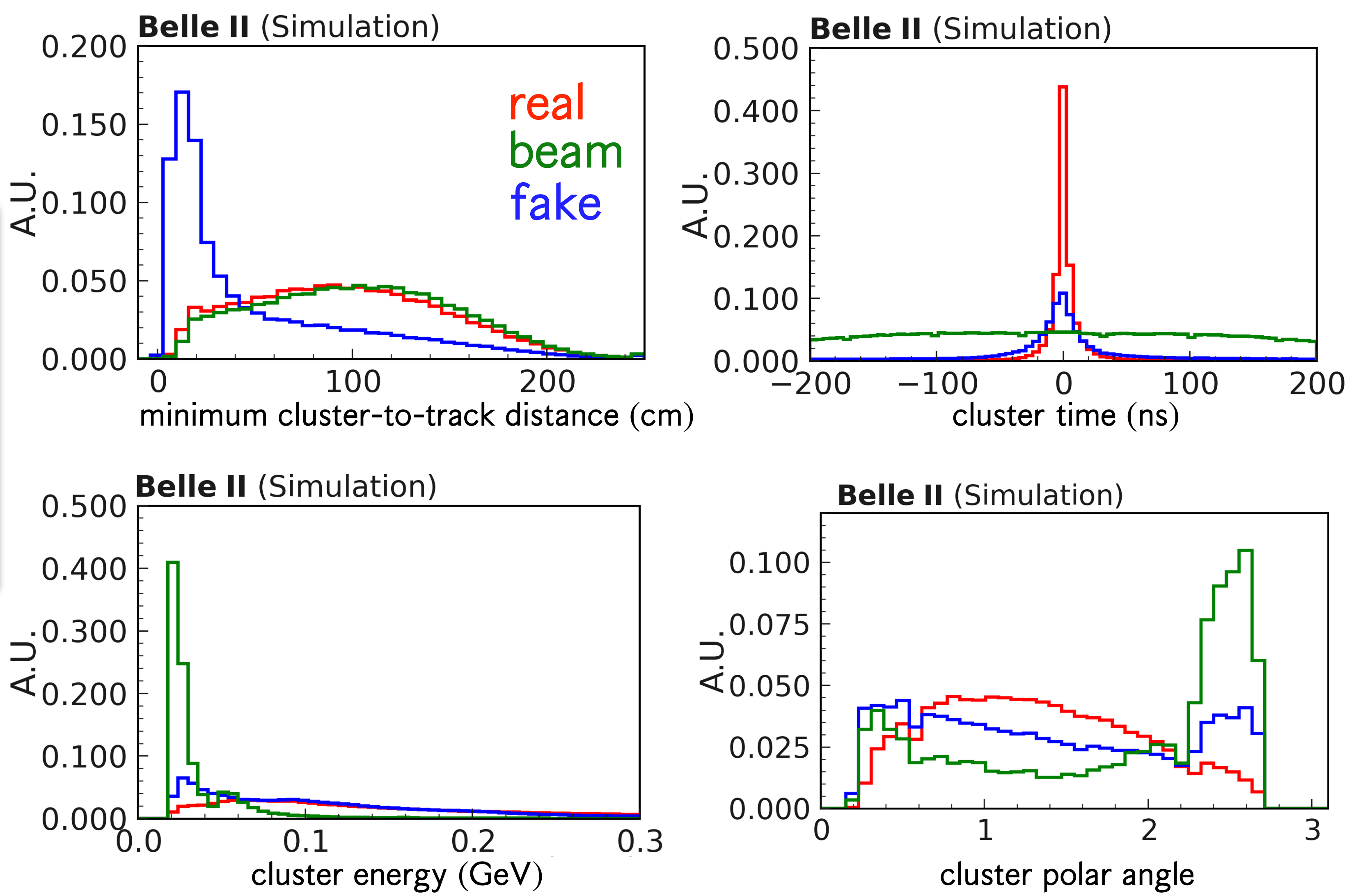
  background suppression in BDT K^{*0}

  signal extraction K^+

E_{ECL} VALIDATION : INTRO

- **E_{ECL} is prone to mismodeling due to its inclusive nature.** Extra-photons can originate from
  real photons
  beam backgrounds
  fake photons — spurious clusters from hadron deposits
- **Data/MC agreement in E_{ECL} depends on the photon selection and the MC campaign** (ECL clustering algorithm)
  - Checked in analysis specific sidebands or background-enhanced samples



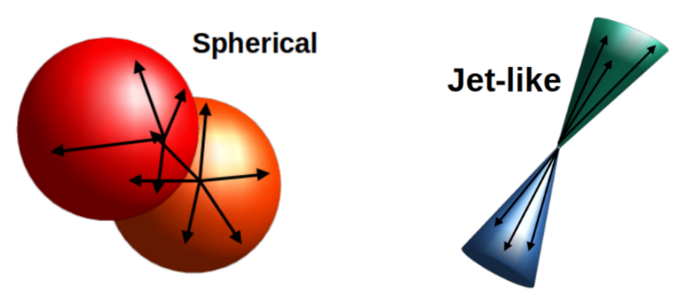


background suppression

KEY-POINTS - $q\bar{q}$

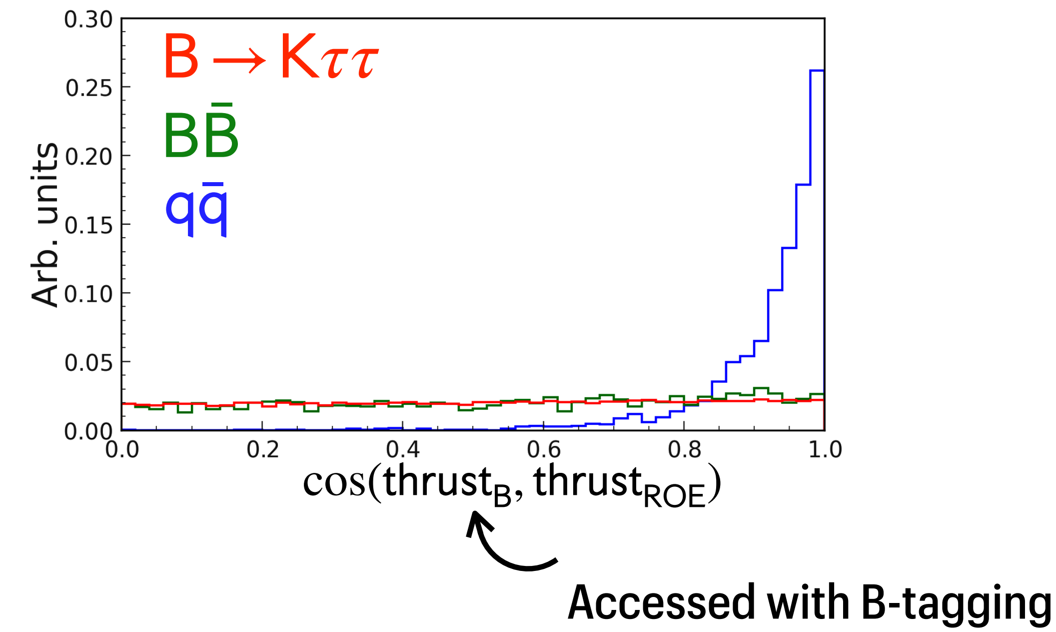
- **With hadronic B-tagging the $q\bar{q}$ background is generally a minor background**

1. Reduced further with event shape information (pre-cut or BDT)

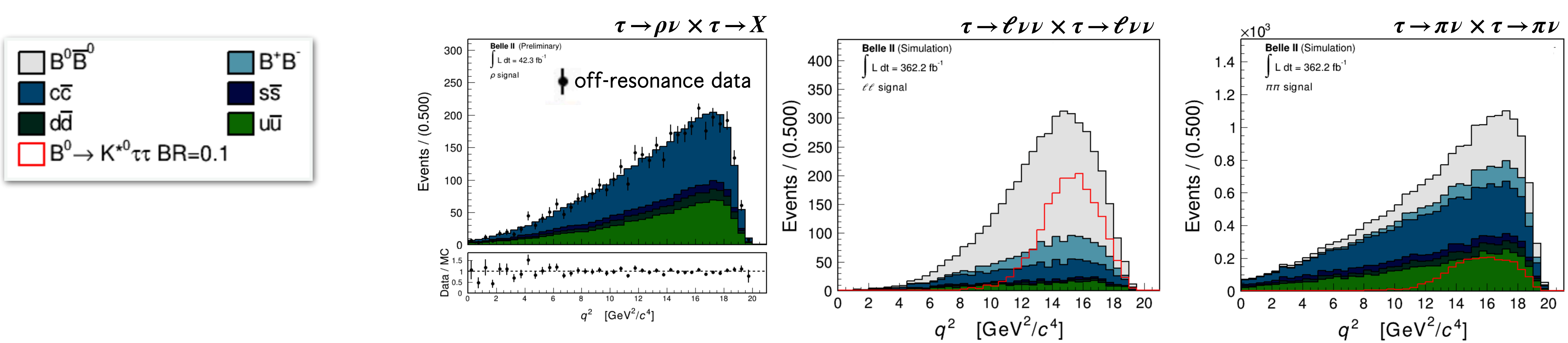


$e^-e^+ \to B\bar{B}$ $\quad$ $e^-e^+ \to q\bar{q}$

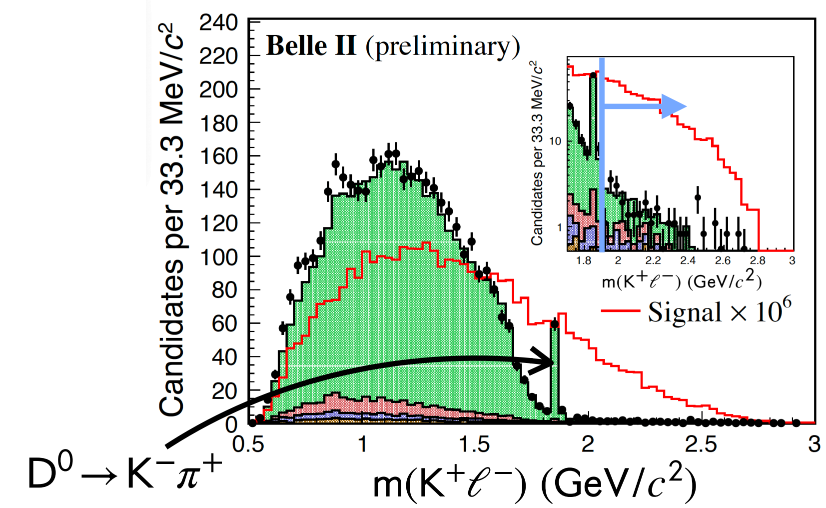
$\mathrm{p(B)} \approx 0.3\,\mathrm{GeV/c}$ $\quad$ $\mathrm{p(q)} \approx 5\,\mathrm{GeV/c}$



2. Separation variables shapes are checked with off-resonance data K^{*0}
   Low statistics for the leptonic categories
   For the hadronic categories, larger contributions



KEY-POINTS - $B\bar{B}$



B$\bar{\text{B}}$ backgrounds are more challenging

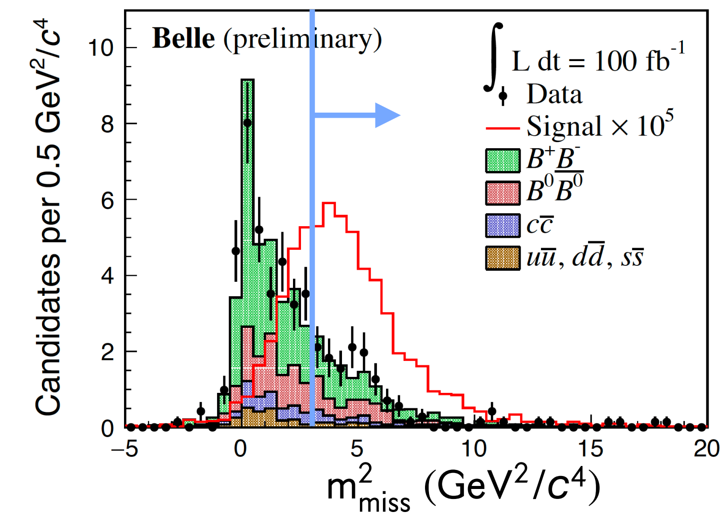
- For leptonic categories, $B \to D(KX)\ell X\nu$
- For hadronic categories, also $B \to Dn\pi$

K^+

1. Focus on the background-depleted region **$m(K^+\ell^-) > 1.9$ GeV/c^2**
2. Optimise selection on **lepton momentum, missing mass and signal window of E_{ECL} to minimise expected UL (with background-only MC)**

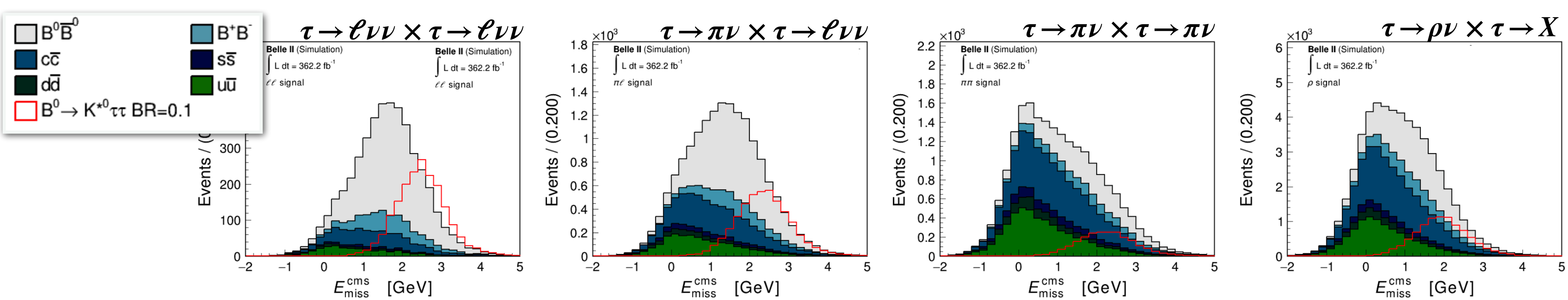
$$m^2_{\text{miss}} = p^2_{\text{miss}} = \begin{pmatrix} E^*_{\text{miss}} \\ \boldsymbol{p}^*_{\text{miss}} \end{pmatrix}^2$$

$$p_{\text{miss}} = p_{e^+e^-} - p_{B_{\text{sig}}} - p_{B_{\text{tag}}}$$



K^{*0}

1. **Combine inputs in BDT**
   Residual energy in calorimeter, q^2, K^{*0} properties, missing momentum, event shape ...
2. **BDT trained for each $\tau\tau$ category given the different background nature/level**



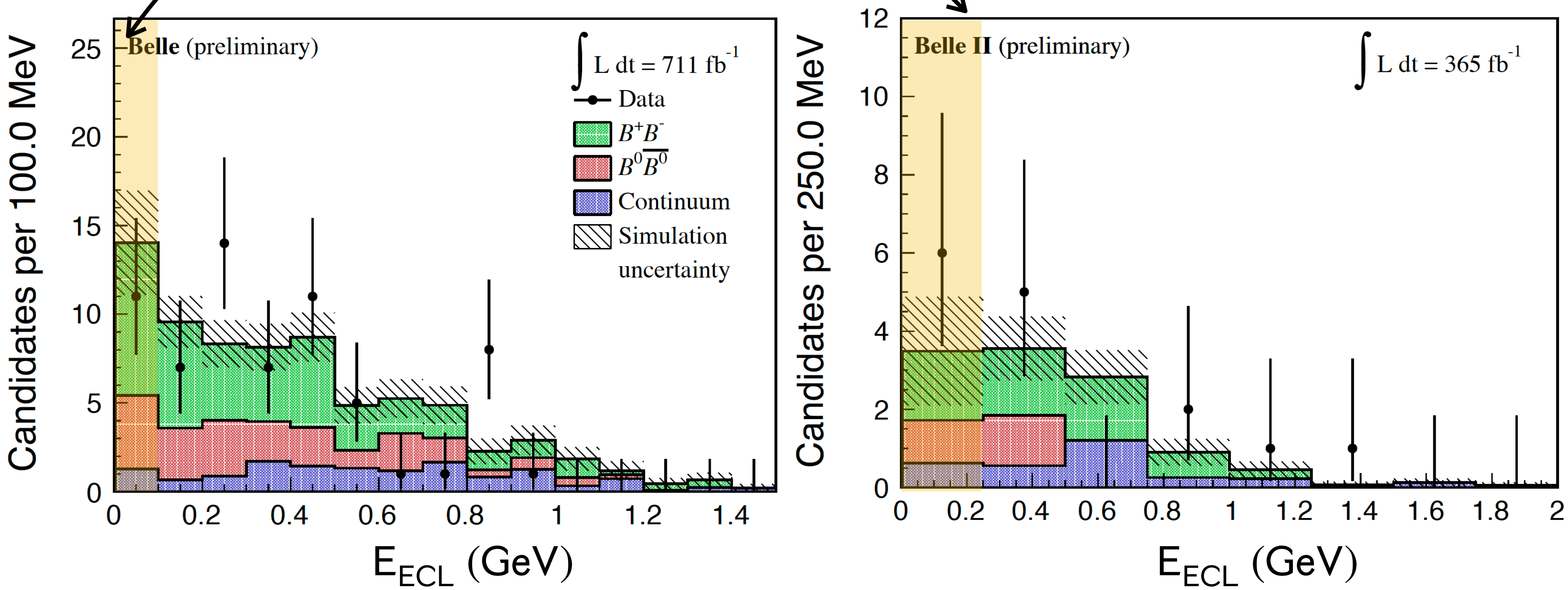
fitting

FITTING - K^+

Combined fit to Belle and Belle II samples

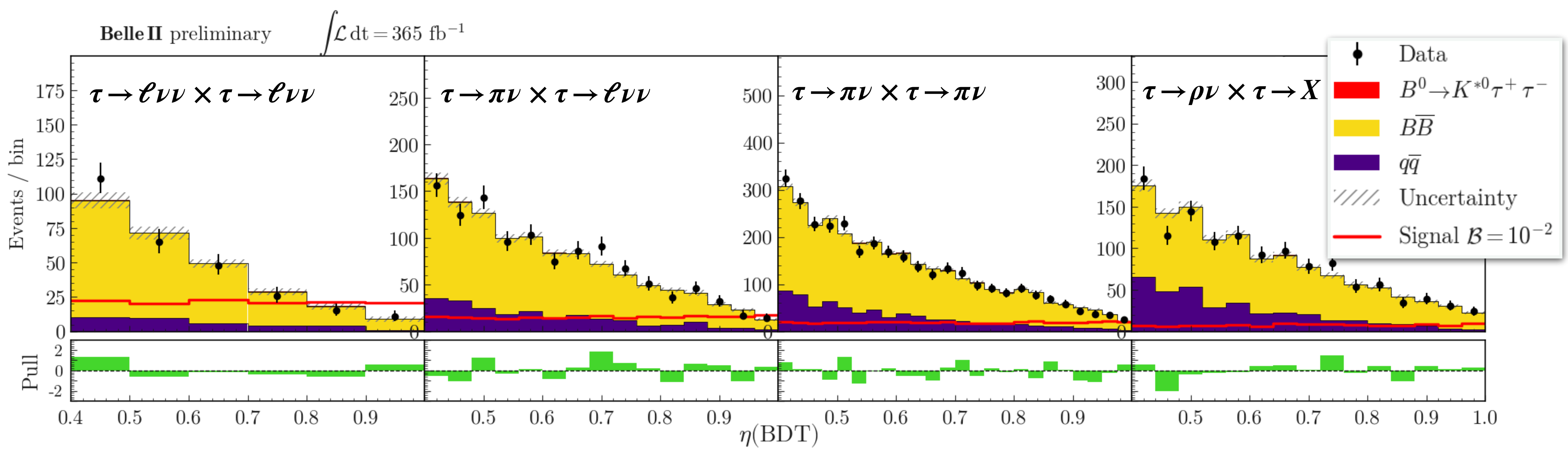
Signal extraction: Poisson-event counts

Optimised signal regions



FITTING - K^{*0}

- **Simultaneous binned ML fit to BDT output to four $\tau\tau$ categories:** $\ell\ell$, $\ell\pi$, $\pi\pi$, ρX
- Signal region determined to maintain **high signal efficiency while limiting the impact of background-related systematic** uncertainties on the expected branching fraction
- **Fit 3 separate components:**
  $q\bar{q}$ (u, d, s, c)
  $B\bar{B}$ (neutral and charged)
  signal $B^0 \to K^{*0}\tau\tau$



systematics

*our results are statistically dominated but we carefully assess the systematic uncertainties

SYSTEMATIC UNCERTAINTIES - K^{*0}

External inputs
Efficiency calibrations (aux measurements)
Dedicated measurements

Source	Impact on $\mathcal{B} \times 10^{-3}$	
$B \to D^{**}\ell/\tau\nu$ branching fractions	0.29	1.
Simulated sample size	0.27	2.
$q\bar{q}$ normalization	0.18	3.
ROE cluster multiplicity	0.17	4.
π and K ID	0.14	4.
B decay branching fraction	0.11	4.
Combinatorial $B\bar{B}$ normalization	0.09	
Signal and peaking $B^0\bar{B}^0$ normalization	0.07	
Lepton ID	0.04	
π^0 efficiency	0.03	
f_{00}	0.01	
$N_{\Upsilon(4S)}$	0.01	
$D \to K_L^0$ decays	0.01	
Signal form factors	0.01	
Luminosity	<0.01	
Total systematics	0.52	
Statistics	0.86	

1. 50% uncertainty on the $B \to D^{**}\ell/\tau\nu$ branching fractions (5-10% of the residual $B\bar{B}$ background)

2. Limited size of simulated samples used for the templates due to low hadronic B-tagging efficiency

3. $q\bar{q}$ normalisation uncertainties obtained from off-resonance (15-70%)

4. Corrections affecting both shapes and normalisations across bins, templates and categories have non-trivial correlations
   - A covariance matrix is obtained with toys and decomposed
   - The main eigenvectors, preserving the correlations of the complete representation, are taken as independent nuisance parameters acting on the shape of the templates

SYSTEMATIC UNCERTAINTIES - K^{*0}

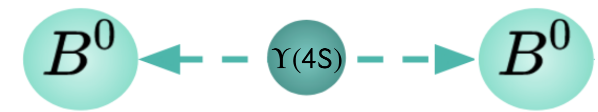
External inputs

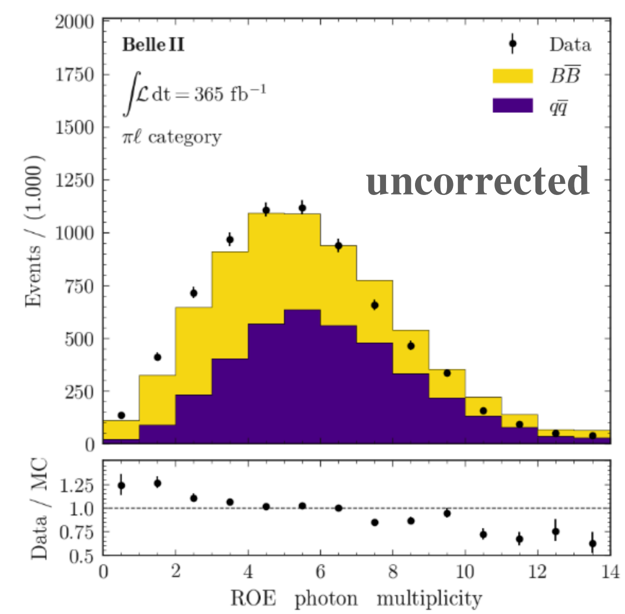
Efficiency calibrations (aux measurements)

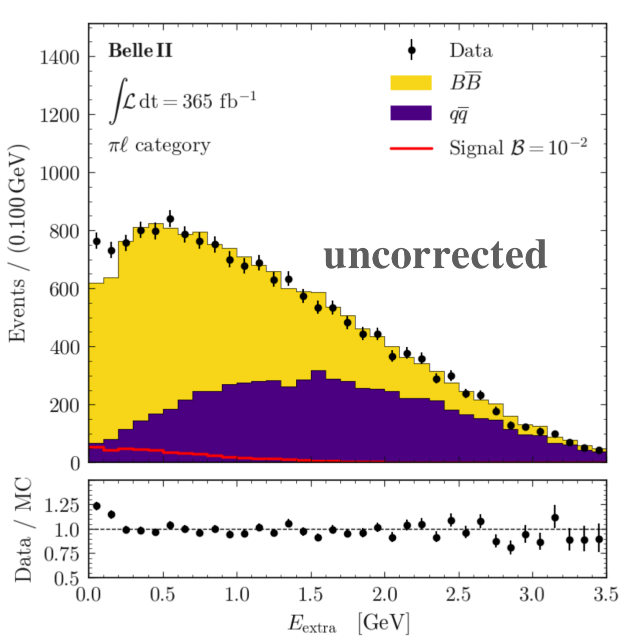
Dedicated measurements

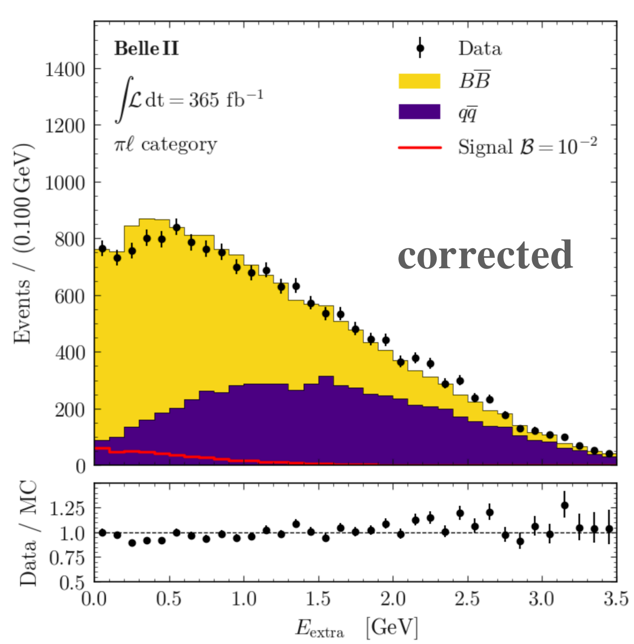
Source	Impact on $\mathcal{B} \times 10^{-3}$
$B \to D^{**}\ell/\tau\nu$ branching fractions	0.29
Simulated sample size	0.27
$q\bar{q}$ normalization	0.18
ROE cluster multiplicity	0.17
π and K ID	0.14
B decay branching fraction	0.11
Combinatorial $B\bar{B}$ normalization	0.09
Signal and peaking $B^0\bar{B}^0$ normalization	0.07
Lepton ID	0.04
π^0 efficiency	0.03
f_{00}	0.01
$N_{\Upsilon(4S)}$	0.01
$D \to K^0_L$ decays	0.01
Signal form factors	0.01
Luminosity	<0.01
Total systematics	0.52
Statistics	0.86

Event-wise extra photons multiplicity correction based on data/MC agreement in the same-flavor sample







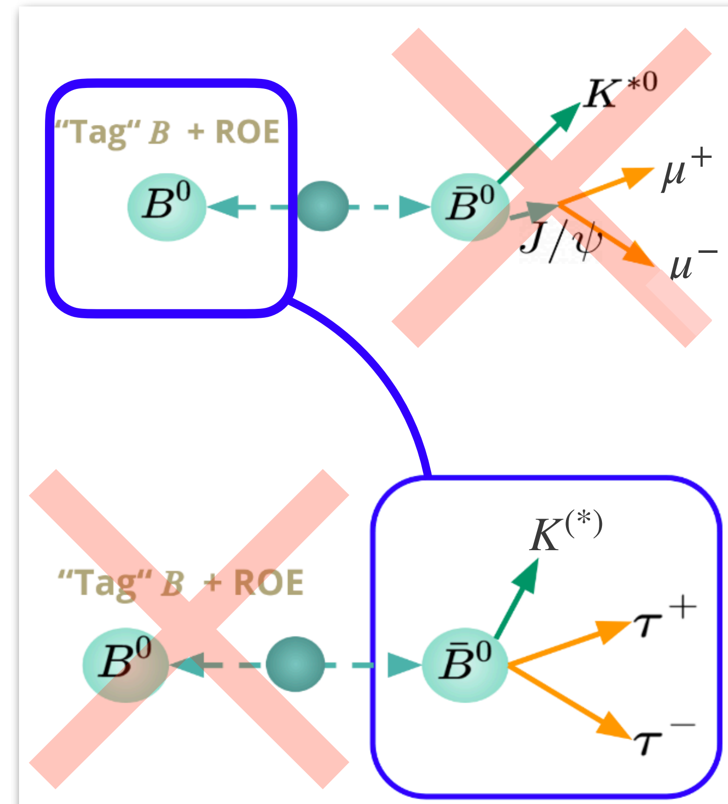


Assign systematic uncertainty from the residual data/MC difference in the η(BDT) < 0.4 sideband

SYSTEMATIC UNCERTAINTIES - K^{*0}

External inputs
Efficiency calibrations (aux measurements)
Dedicated measurements

Source	Impact on $\mathcal{B} \times 10^{-3}$
$B \to D^{**}\ell/\tau\nu$ branching fractions	0.29
Simulated sample size	0.27
$q\bar{q}$ normalization	0.18
ROE cluster multiplicity	0.17
π and K ID	0.14
B decay branching fraction	0.11
Combinatorial $B\bar{B}$ normalization	0.09
Signal and peaking $B^0\bar{B}^0$ normalization	0.07
Lepton ID	0.04
π^0 efficiency	0.03
f_{00}	0.01
$N_{\Upsilon(4S)}$	0.01
$D \to K^0_L$ decays	0.01
Signal form factors	0.01
Luminosity	<0.01
Total systematics	0.52
Statistics	0.86

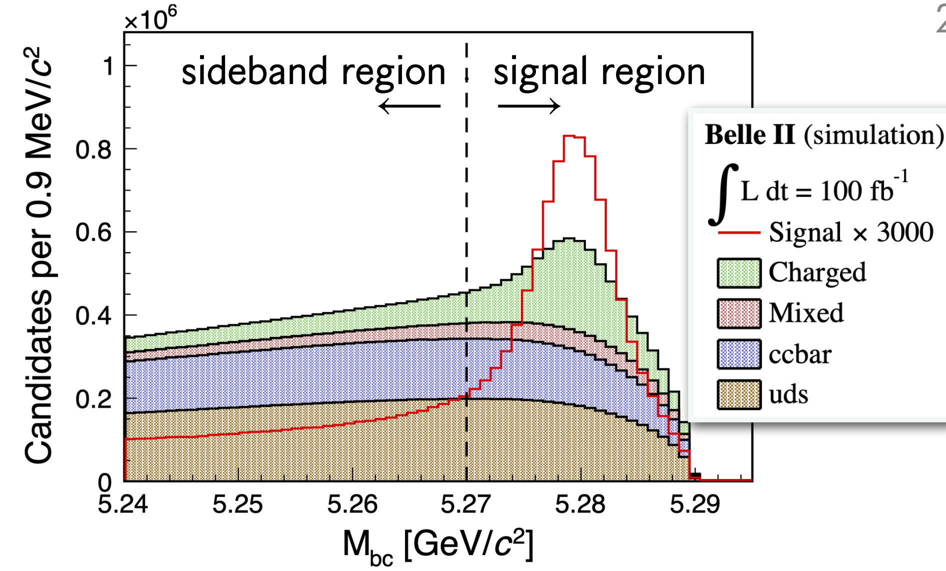


- **Reconstruct** $B^0 \to K^{(*)0} J/\psi(\mu\mu)$ events (clean)
- **Remove** $B \to K^{(*)} J/\psi(\mu\mu)$ from reconstructed objects and **replace** with $B^0 \to K^{*0}\tau\tau$ (both data and MC)
- Data/MC efficiency ratio used as tag efficiency correction
  - The correction is used for signal MC and correctly reconstructed B_{tag} in $B^0\bar{B}^0$ events

Cons: low statistics
Pros: analysis-specific

SYSTEMATIC UNCERTAINTIES - K+

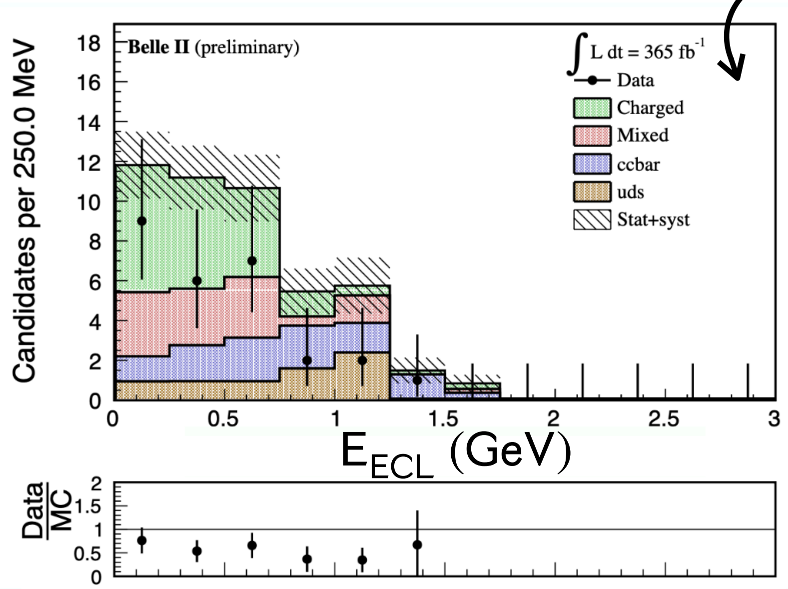
Source	Size in Belle	Size in Belle II
Expected background yield	± 2.5 (absolute)	± 1.2 (absolute)
FEI scale factor	10.0%	13.3%
Tracking efficiency	1.1%	0.8%
Particle identification corrections	5.4%	1.6%
π^0 veto efficiency	2.2%	2.5%
Simulated sample-size	3.1%	3.3%
Decay model	3.8%	4.0%
Total B-meson yield	1.4%	1.5%
$\Upsilon(4S)$ branching fraction	$^{+1.4\%}_{-2.1\%}$	$^{+1.4\%}_{-2.1\%}$
Total	$\pm 2.24 \times 10^{-4}$	$\pm 2.46 \times 10^{-4}$



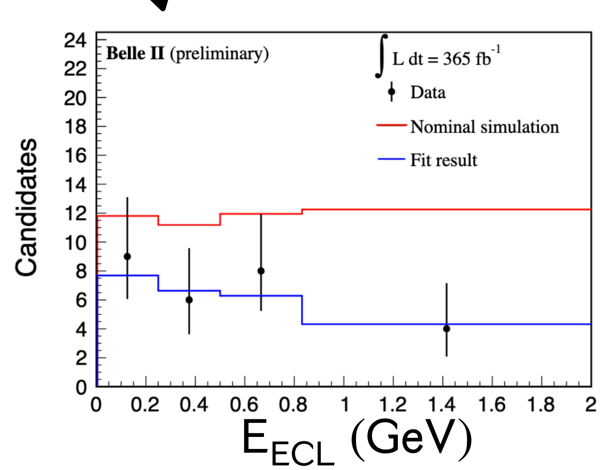
- E_{ECL} shape calibrated by **fitting the simulation to match data in the sidebands and extrapolating it into the signal region**
- Uncertainty combines the stat. uncertainty due to the **sizes of the sideband samples**, and the syst. uncertainty associated with the **shape assumed for the background scaling**

Sidebands

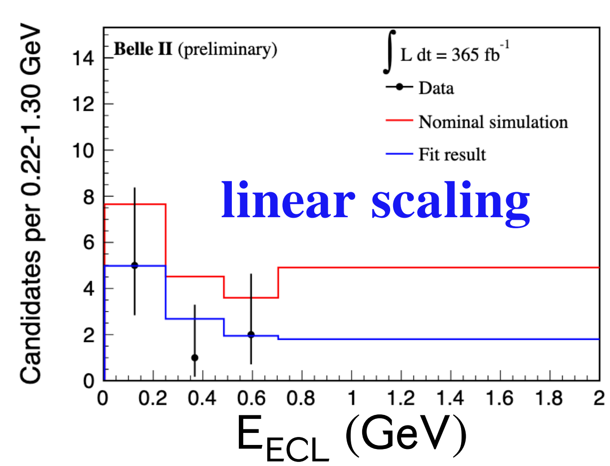
1. $M_{bc} \in [5.20, 5.27]$ GeV$/c^2$
2. $q^2 < 14.18$ GeV$^2/c^4$
3. $E_{ECL} \notin$ SR



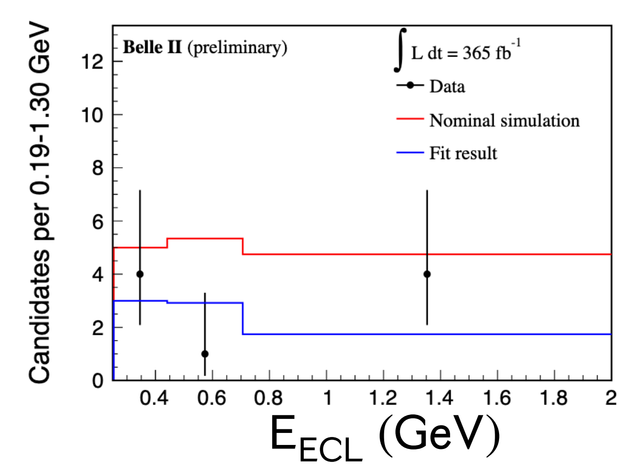
1.



2.



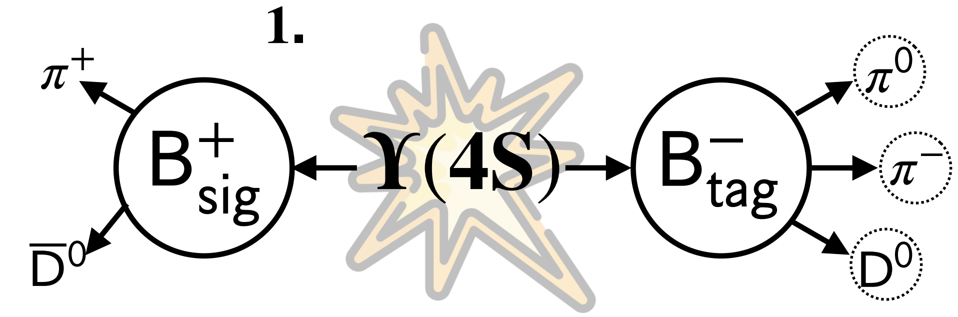
3.



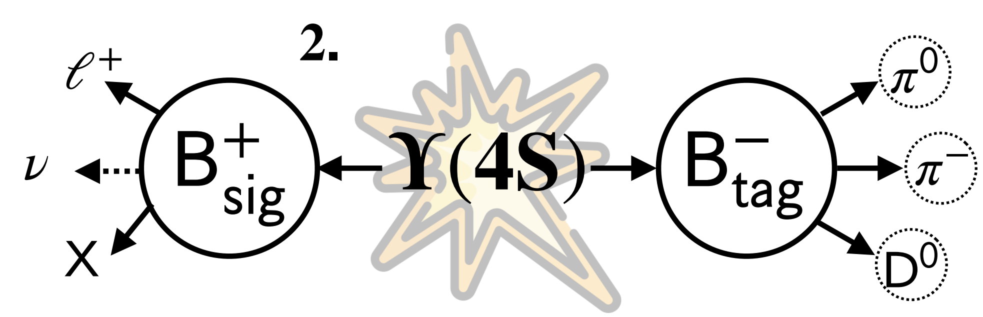
SYSTEMATIC UNCERTAINTIES - K^+

Source	Size in Belle	Size in Belle II
Expected background yield	± 2.5 (absolute)	± 1.2 (absolute)
FEI scale factor	10.0%	13.3%
Tracking efficiency	1.1%	0.8%
Particle identification corrections	5.4%	1.6%
π^0 veto efficiency	2.2%	2.5%
Simulated sample-size	3.1%	3.3%
Decay model	3.8%	4.0%
Total B-meson yield	1.4%	1.5%
$\Upsilon(4S)$ branching fraction	$^{+1.4\%}_{-2.1\%}$	$^{+1.4\%}_{-2.1\%}$
Total	$\pm 2.24 \times 10^{-4}$	$\pm 2.46 \times 10^{-4}$

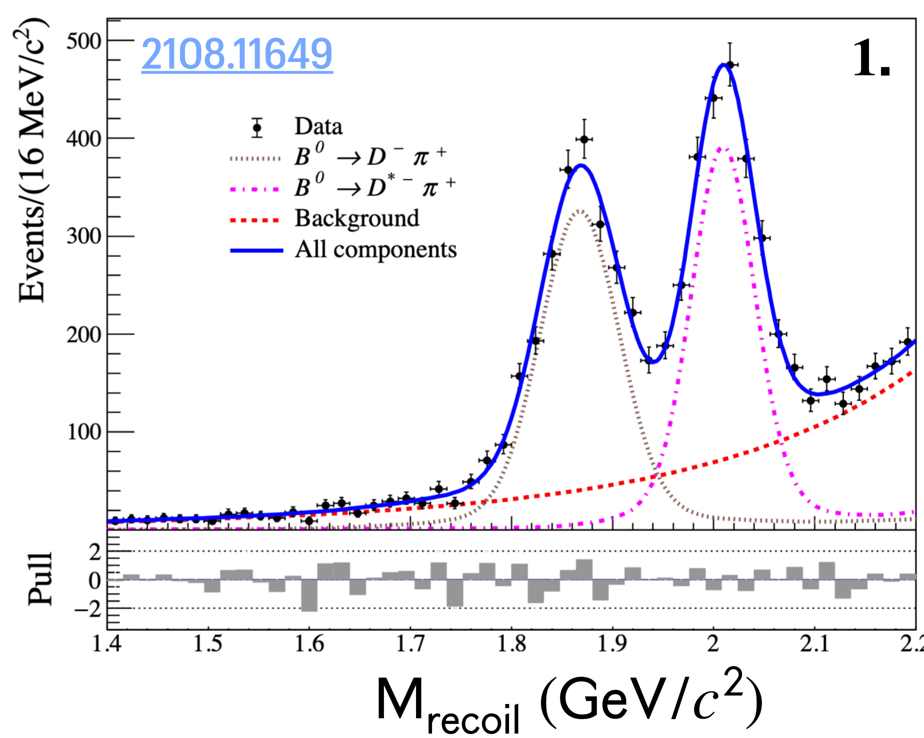
$B^+ \to \overline{D}^0 \pi^+$ for the B-tagging efficiency correction

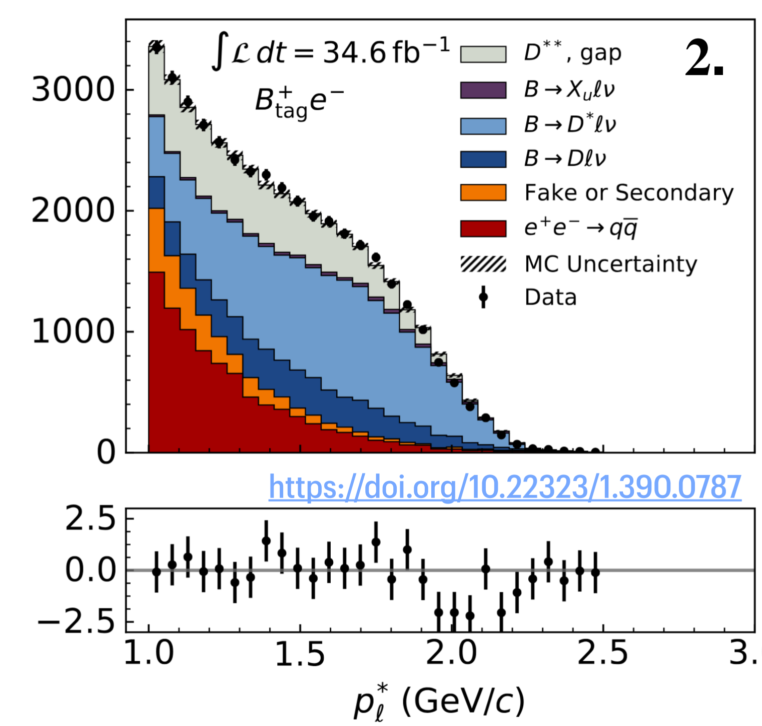


Compared to $B^+ \to X \ell^+ \nu_\ell$ for the associated uncertainty



$$M^2_{recoil} = (p_{e^+e^-} - p_\pi - p_{B_{tag}})^2$$
$$= m^2_\pi + m^2_B - 2(E_\pi \sqrt{s}/2 + \vec{p}_{B_{tag}} \cdot \vec{p}_{B_\pi})$$





results

RESULTS

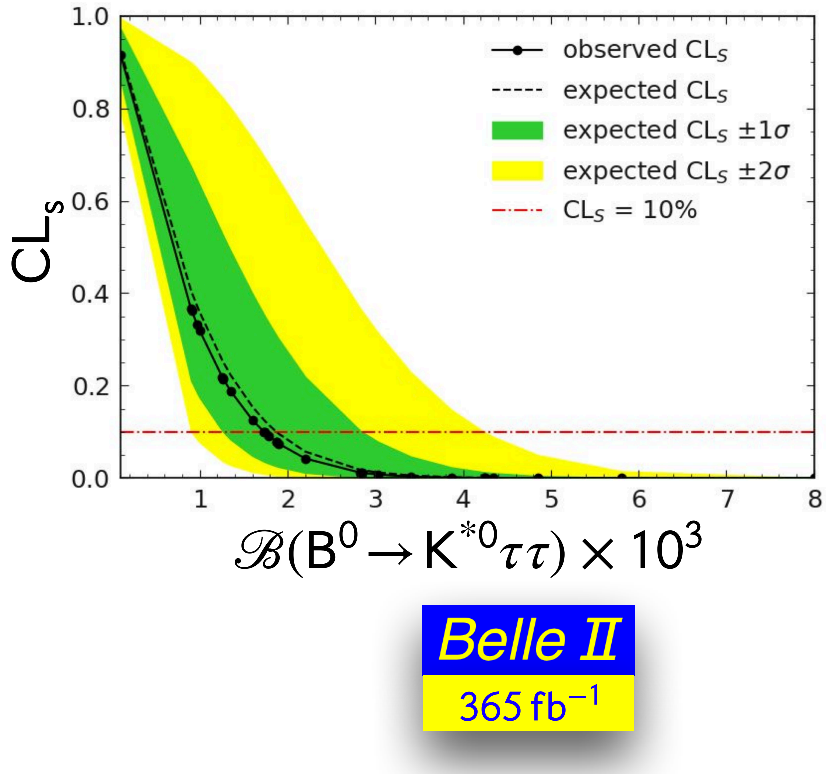
K^{*0}

$\mathcal{B}(B^0 \to K^{*0}\tau\tau) = [-0.15 \pm 0.86\,(\text{stat}) \pm 0.52\,(\text{syst})] \times 10^{-3}$

$\boldsymbol{\mathcal{B}(B^0 \to K^{*0}\tau\tau) < 1.8 \times 10^{-3}\ (90\,\%\ \text{CL})}$

- **Upper limit derived with CLs method**

×2 better limit using half sample size compared to Belle
(better tagging algorithm, multivariate analysis, additional final states)



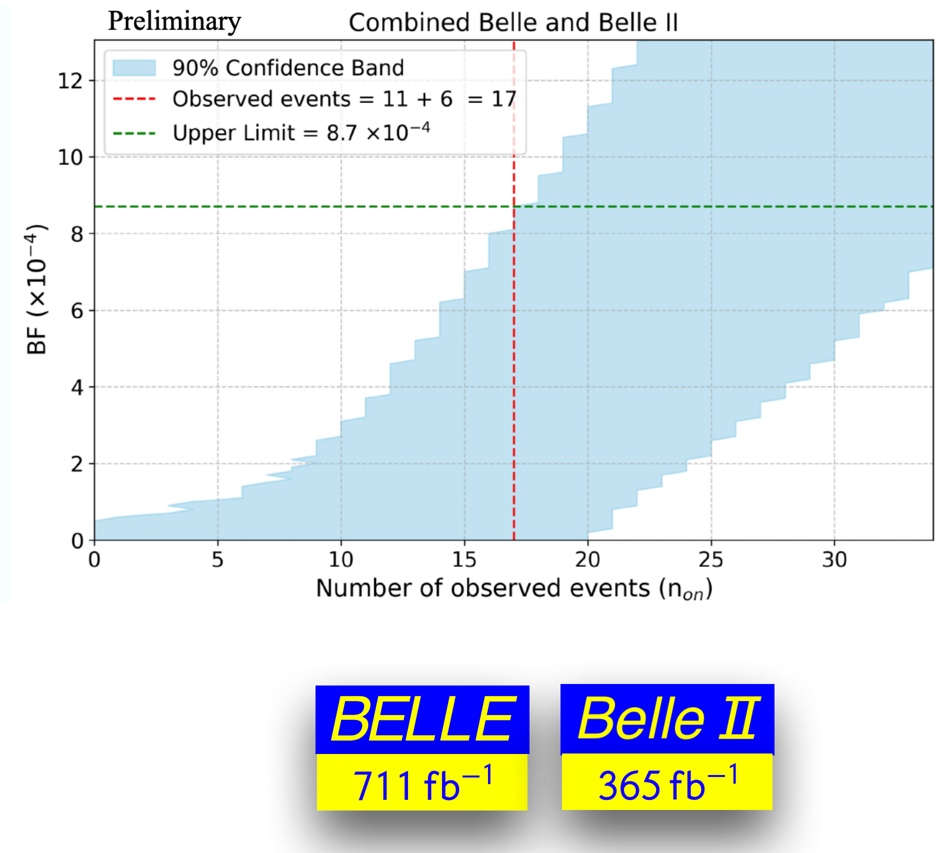
K^+

$\mathcal{B}(B^+ \to K^+\tau\tau) = [3.13^{\pm 3.70}_{\pm 3.30}] \times 10^{-4}$

$\boldsymbol{\mathcal{B}(B^+ \to K^+\tau\tau) < 8.7 \times 10^{-4}\ (90\,\%\ \text{CL})}$

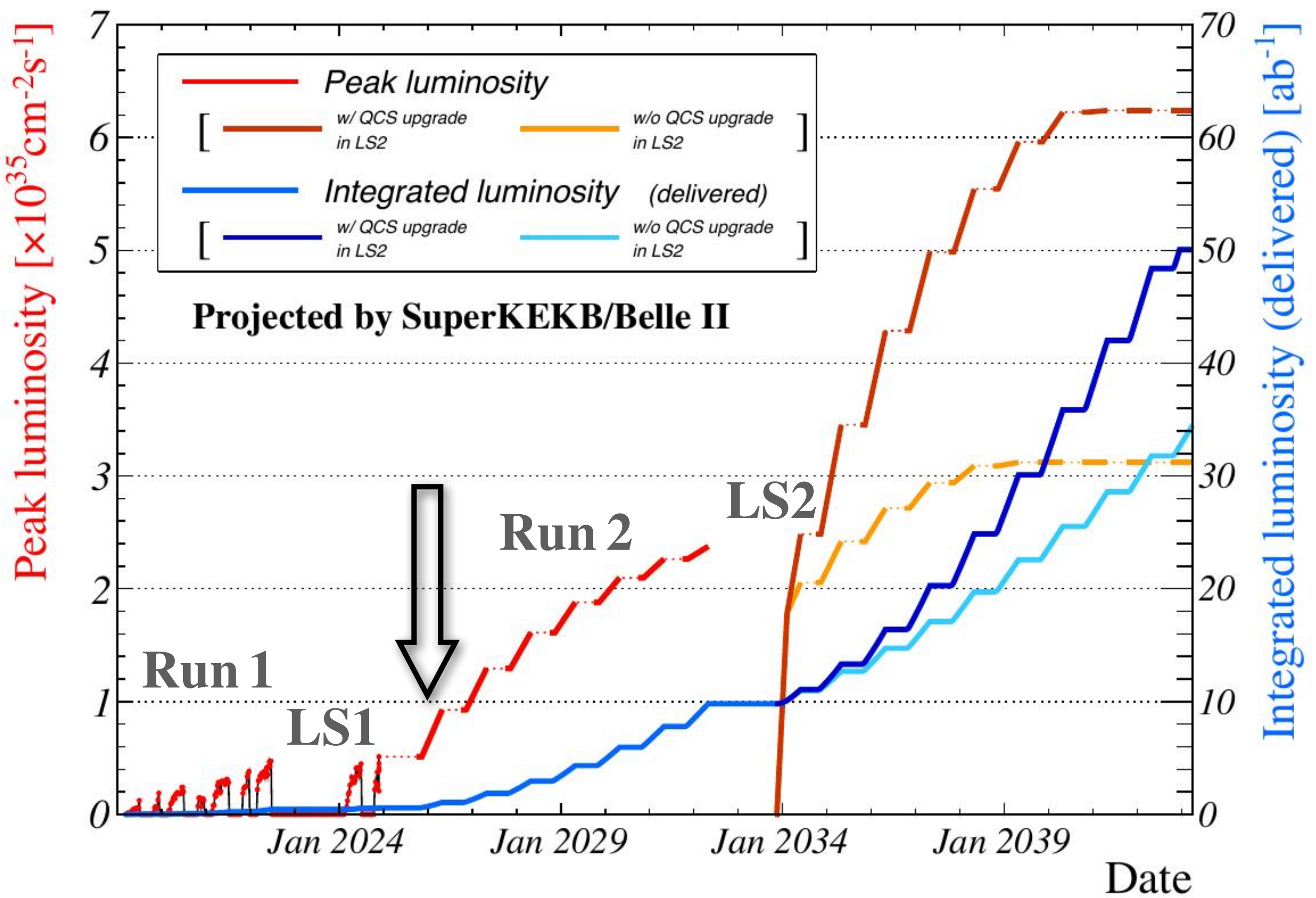
- **Upper limit derivation:** frequentist exclusion limits using LR ordering
Confidence bands constructed assuming the least signal-sensitivity within one-standard deviation range of systematic uncertainties

×2.6 better limit compared to BaBar
(better tagging algorithm, larger dataset, same final states)



CONCLUSION

- Belle II is producing new results which significantly improve over 1st-generation *B*-factories
- Expected to have competitive/leading sensitivity depending on the mode
- Results are statistically limited but have handles to reduce sizeable systematics



CONCLUSION&OUTLOOK

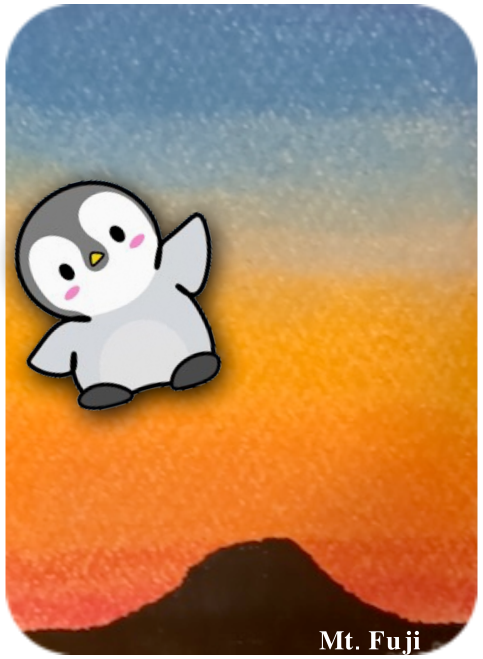
Increase efficiency

- **Add τ channels when not explored already**
  - Further optimisation of analysis strategy when adding less pure modes
    For $B^0 \to K^{*0}\tau\tau$, 10% / 35% BF uncertainty improvement when adding ρ / hadronic modes
- **Alternative *B*-tagging approaches (semileptonic, inclusive — more sensitive for other searches)**

Reduce systematic uncertainties

- Many have **statistical nature**
  - background normalisations, simulated samples, ...
- **Improvements in reconstruction, MC modelling, and analysis techniques**
  - systematics for specific backgrounds/neutral energy corrections

Thank you for your attention! Questions?



SUPPORTING MATERIAL