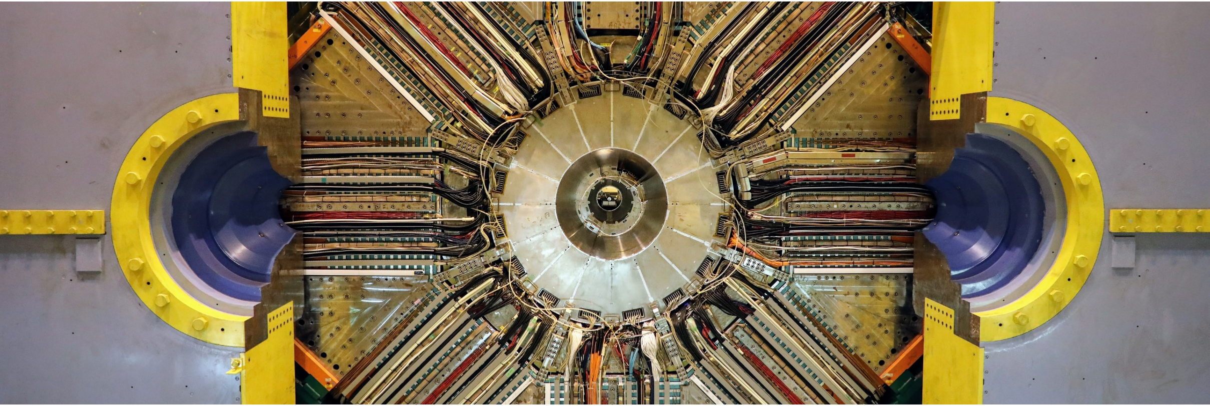
Rare Decays from Belle II



Caspar Schmitt
for the Belle II Collaboration

LMU
LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN
Max-Planck-Institut für Physik (Werner-Heisenberg-Institut)
Belle II

Belle II at today's B factory SuperKEKB

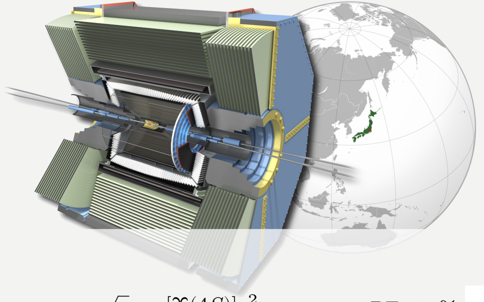
Belle II focuses on (but is not limited to)
flavor physics with B mesons
in precision (indirect) BSM searches.

Advantages wrt. hadron colliders

- kinematically well-known initial state
- very low backgrounds, partly accessible at $\sqrt{s} < m[\Upsilon(4S)]c^2$
- low multiplicities; strong with neutral particles

Disadvantages wrt. hadron colliders

- low cms energy and B boost
- low cross section; need high luminosity!

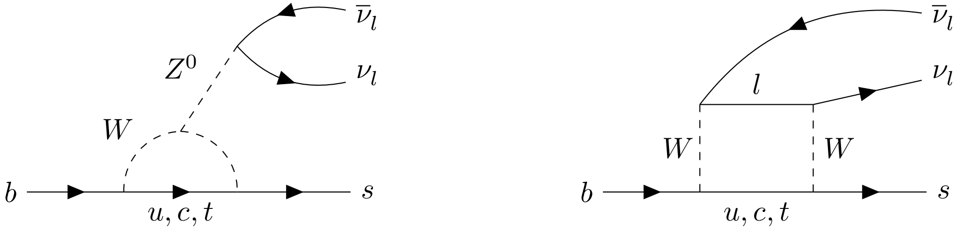


$$e^+e^- \xrightarrow{\sqrt{s}=m[\Upsilon(4S)]c^2} \Upsilon(4S) \xrightarrow{\text{BF}>96\%} B\overline{B}$$

Nano-Beam Scheme:
$\mathcal{L}_{\text{peak}} = 5.1 \cdot 10^{34}\,\text{cm}^{-2}\text{s}^{-1}$

Inclusive $B \to X_s\nu\bar{\nu}$ with Hadronic B Tagging

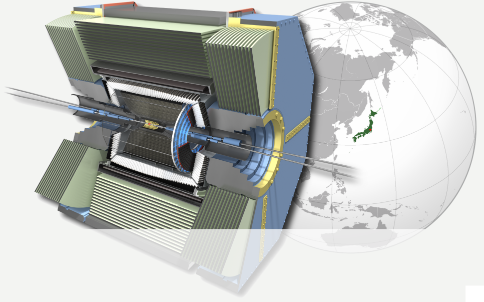
- Probe flavor changing neutral currents (FCNC) in $b \to s\nu\bar{\nu}$.
- Suppression in standard model allows for precise tests on alternate theories.
- Limits enhanced by precise theory prediction, since no γ exchange.



Leading-order SM diagrams. Long-distance contribution via intermediate τ is sub-dominant.

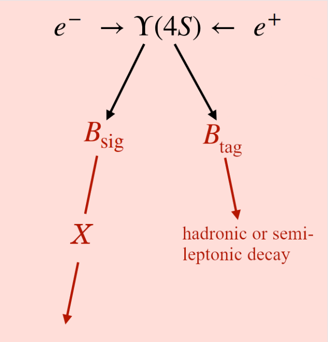
Belle II at B Factory SuperKEKB

Search for $B \to X_s \nu\bar{\nu}$ only possible at Belle II due to neutrinos in final state.



$$e^+e^- \xrightarrow{\sqrt{s}=m[\Upsilon(4S)]c^2} \Upsilon(4S) \xrightarrow{\text{BF}>96\%} B\overline{B}$$

Hadronic B Tagging at Belle II (HTA)



allows kinematic constraints on B_{sig}

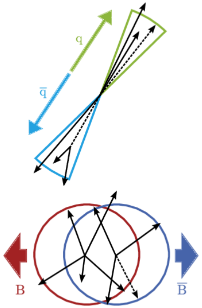
$$|\Delta E^{\text{tag}}| \equiv |E^*_{B_{\text{tag}}} - \sqrt{s}/2|$$

$$M_{bc}^{\text{tag}} \equiv \sqrt{s/4 - |p^*_{B_{\text{tag}}}|^2}$$

Inclusive X_s Reconstruction by Summing 30 Exclusive Decay Modes

- Kinematic constraints $|\Delta E^{\text{tag}}| < 0.2\,\text{GeV}$ and $M_{bc}^{\text{tag}} > 5.27\,\text{GeV/c}^2$.
- Reconstructed total of 30 exclusive final states:
  - $K + i\pi$ with $0 \leq i \leq 4$
  - $3K + i\pi$ with $0 \leq i \leq 1$

  with $K = K^+, K_S^0$ and each at most one π^0.
- Exclusive modes amount to 93% of $\mathcal{B}(B \to X_s \nu\overline{\nu})$.
- Require $E_{\text{extra}}^{\gamma} < 1.3\,\text{GeV}$ without extra tracks or π^0 or K_S^0 and with $\theta(\vec{p}_{\text{miss}})$ in detector acceptance.
- Train BDT η on kinematic and topological features of B_{sig}, B_{tag}, $\vec{p}_{\text{miss}}$ and reconstructed charm vetos $D \to K_{\text{sig}} + i\pi$.



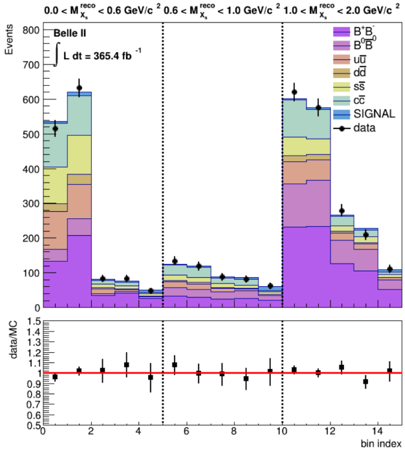
Event-shapes encode topology.

Signal Extraction for $B \to X_s \nu\bar{\nu}$

- Signal region maximizes significance while excluding charm backgrounds. $[M_{X_s} < 1.84 \cup 1.89 < M_{X_s} < 2.0\,\text{GeV}] \times [\eta > 0.86]$
- Maximum likelihood fit in $M_{X_s} \times \eta$ with 3×5 bins, maximizing sensitivity.

			$\mathcal{B}\ [10^{-5}]$		
$M_{X_s}\ [\text{GeV}/c^2]$	ϵ	N_{sig}	Central value	UL_{obs}	UL_{exp}
$[0, 0.6]$	0.25%	10^{+18+18}_{-17-16}	$0.5^{+0.9+0.9}_{-0.8-0.8}$	2.5	2.4
$[0.6, 1.0]$	0.11%	36^{+27+31}_{-25-26}	$3.8^{+2.8+3.2}_{-2.6-2.7}$	10.0	7.2
$[1.0, M^{\text{max}}_{X_s})$	0.06%	33^{+44+64}_{-42-53}	$7.2^{+9.6+13.9}_{-9.2-11.6}$	35.3	28.3
Full range	0.11%	80^{+61+93}_{-59-79}	$11.5^{+8.9+13.5}_{-8.5-11.4}$	35.6	27.9

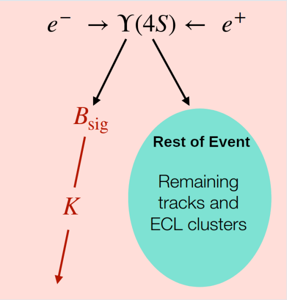
No signal at $\mathcal{B} = \left[1.2^{+0.9}_{-0.9}(\text{stat.})^{+1.4}_{-1.1}(\text{syst.})\right]$, yielding best upper limit $3.6 \cdot 10^{-4}$ at 90% CL.



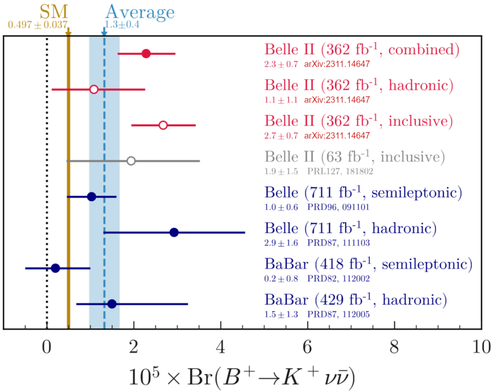
$M(X_s)$ regions enriched in K, K^* and $(K n\pi)_{\text{non-res}}$.

Reinterpreting $B^+ \to K^+\nu\bar{\nu}$ in BSM Scenarios [arXiv:2507.12393,2402.08417]

Inclusive B tagging at Belle II (ITA)



allows higher efficiencies than HTA at lower purities



ITA+HTA exceeds SM prediction by 2.7σ.

What BSM models does this favor?

Reweighting the SM-dependent likelihood

- $B^+ \to K^+\nu\bar{\nu}$ was measured using SM signal templates.
- Reweighting these to some BSM model must take into account experimental acceptance and efficiencies.
- Re-simulating event-by-event is expensive and non-reproducible by community.
- Re-weight the number density $n(x)$ of expected events after selection from one onto another theory prediction for the cross section $\sigma(q^2)$.

$$n_{\text{BSM}}(x) = L \int dq^2\, \epsilon(x|q^2)\sigma_{\text{BSM}}(q^2) = L \int dq^2\, \underbrace{\epsilon(x|q^2)\sigma_{\text{SM}}(q^2)}_{n_{\text{SM}}(x,q^2)} \frac{\sigma_{\text{BSM}}(q^2)}{\sigma_{\text{SM}}(q^2)}$$

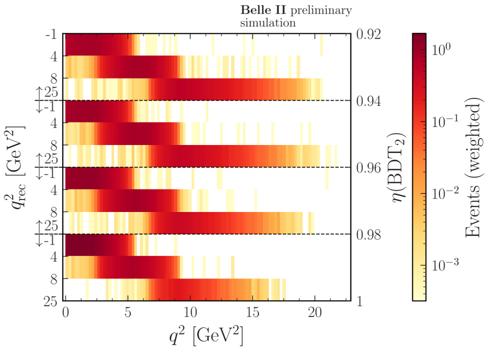
Knowledge of joint number density $n_{\text{SM}}(x, q^2)$
and kinematic theory predictions for $\sigma(q^2)$ suffices for reinterpretation!

Reweighting $B^+ \to K^+\nu\bar{\nu}$

- In practice, we evaluate the relations only in a finite number of bins.

$$n_{\mathrm{BSM}}(x, q^2) \to n_{\mathrm{BSM},x,q^2} = \sum_{q^2\ \mathrm{bins}} n_{\mathrm{SM},x,q^2} \frac{\sigma_{\mathrm{BSM},q^2}}{\sigma_{\mathrm{SM},q^2}}$$

- We publish n_{SM,x,q^2} for $B^+ \to K^+\nu\bar{\nu}$ in sufficiently fine binning.



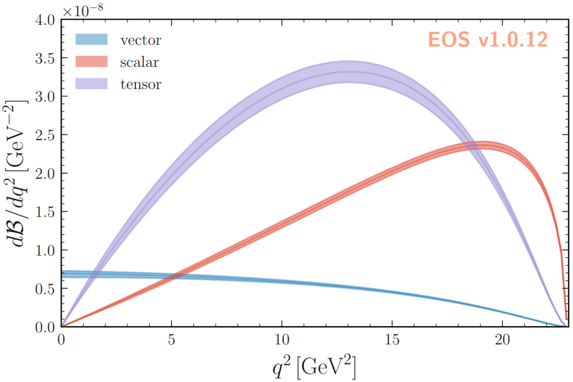
redist

Redist Python Package

Reinterpreting $B^+ \to K^+\nu\bar{\nu}$ in Weak Effective Theory

- Reweight n_{SM,x,q^2} to WET prediction using 6 dimensional operators.

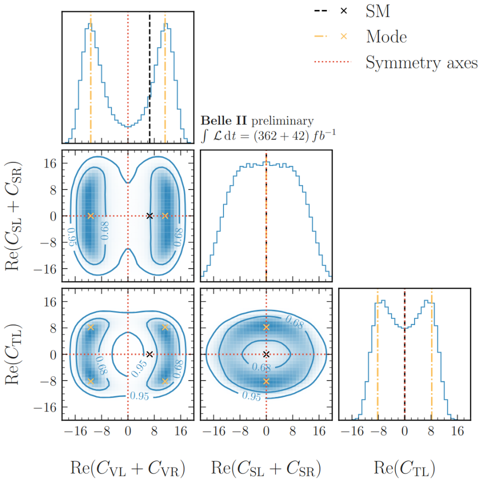
$$\frac{d\mathcal{B}}{dq^2} \propto \frac{\lambda_{BK}}{24q^2}|f_+(q^2)|^2|C_{VL}+C_{VR}|^2 + \frac{(M_B^2-M_K^2)^2}{8(m_b-m_s)^2}|f_0(q^2)|^2|C_{SL}+C_{SR}|^2 + \frac{2\lambda_{BK}|f_T(q^2)|^2}{3(M_B+M_K)^2}|C_{TL}|^2$$

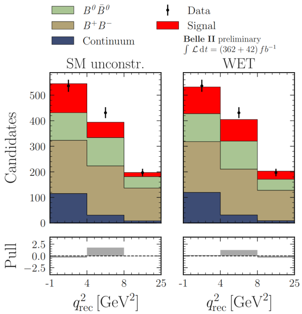


Reinterpreting in WET yields POI $|C_{VL}+C_{VR}|, |C_{SL}+C_{SR}|, |C_{TL}|$, with $C_i^{\mathrm{SM}}=0$ except

$$C_{VL}^{\mathrm{SM}} = 6.6 \pm 0.1$$

Reinterpreting $B^+ \to K^+\nu\bar{\nu}$ in Weak Effective Theory





Parameters	Mode	68% HDI	95% HDI
$\lvert C_{\mathrm{VL}} + C_{\mathrm{VR}}\rvert$	11.3	[7.82, 14.6]	[1.86, 16.2]
$\lvert C_{\mathrm{SL}} + C_{\mathrm{SR}}\rvert$	0.00	[0.00, 9.58]	[0.00, 15.4]
$\lvert C_{\mathrm{TL}}\rvert$	8.21	[2.29, 9.62]	[0.00, 11.2]

WET is favored over SM with large $|C_{VL} + C_{VR}|$ and non-zero $|C_{TL}|$.

$B^0 \to K^{*0}\tau^+\tau^-$ with Hadronic Tag

- FCNC suppression and $b \to s$ analogous to $B^+ \to K^+\nu\bar{\nu}$, but with intermediate γ.
- Constrain BSM for $B^+ \to K^+\nu\bar{\nu}$ and $b \to c\tau\nu_\tau$ anomalies.
- HTA approach with B_{sig} reconstructed in combinations of

$$B^0_{\text{sig}} \to K^{*0}\left[\to K^+\pi^-\right]\tau^+\tau^- \quad \begin{cases} \tau \to \mu\bar{\nu}_\mu\nu_\tau \\ \tau \to e\bar{\nu}_e\nu_\tau \\ \tau \to \pi\nu_\tau \\ \tau \to \rho\left[\to \pi\pi^0\right]\nu_\tau. \end{cases}$$

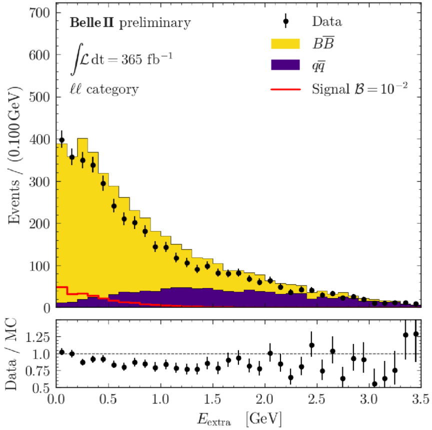
Challenges

- Low $\mathcal{B}_{\text{SM}} = (0.98 \pm 0.10)\cdot 10^{-7}$ with best limit $< 3.1 \cdot 10^{-3}$ at 90% CL by Belle.
- Final states with up to 4 neutrinos and low-momentum K^{*0} provide **no peaking observables**.

Selections in $B^0 \to K^{*0}\tau^+\tau^-$

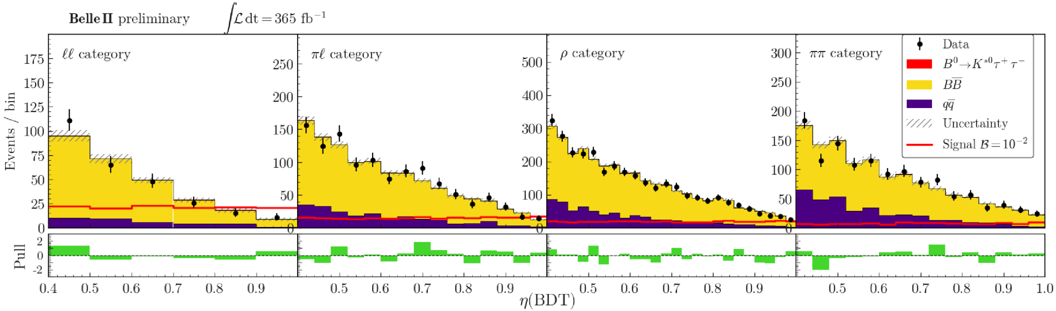
- Require $B_{\text{sig}}\, B_{\text{tag}} = B^0\overline{B^0}$ of opposite flavor.
- Require no extra tracks and $\theta\,(p_{\text{extra}} = p_{\text{collider}} - \sum p_i)$ in detector acceptance.
- Single-candidate selection with best $M(K^{*0})$, preferring first ρ then e, μ modes.

Split in 4 signal categories $ll, \pi l, \pi\pi, \rho$ **which differ in number of neutrinos and hence kinematics**.



Signal Extraction in $B^0 \to K^{*0}\tau^+\tau^-$

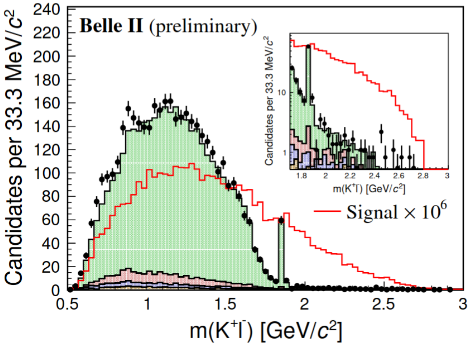
- Binary BDT classifier η trained separately for 4 signal categories, notably on E_{extra} and $q^2 = (p_{\tau^+} + p_{\tau^-})^2$.
- Binned maximum likelihood fit of η in 4 signal categories.

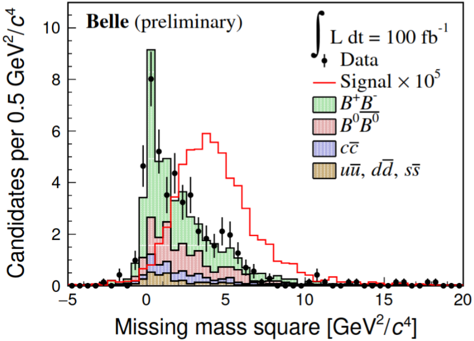


No evidence for signal at $\mathcal{B} = [-0.15 \pm 0.86(\text{stat.}) \pm 0.52(\text{syst.})] \cdot 10^{-3}$, yielding best upper limit of $1.8 \cdot 10^{-3}$ at 90% CL.

$B^+ \to K^+\tau^-\tau^+$ with Hadronic Tag

- HTA approach on **Belle and Belle II** with only leptonic $\tau \to l\nu_l$ i.e. $l = e, \mu$.
- Require no extra tracks and kinematically suppress $B^+ \to K^+ J\psi(\to ll)$ and $\gamma \to e^+e^-$ and $B^+ \to \psi(2S)[\to \tau^+\tau^-]K^+$.
- Charm-depleted signal region in $M(K^+l^-) > 1.9\,\mathrm{GeV}$ with large M^2_{miss}, p_l and small E_{extra}.



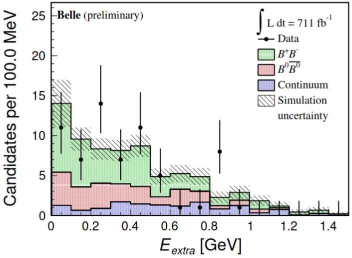


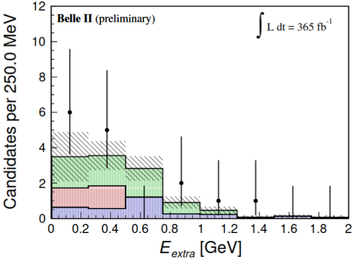
Signal Extraction in $B^+ \to K^+ \tau^- \tau^+$

Compare counts with background-only expectation, extrapolated from sidebands in $E_{\text{extra}}, q^2, M_{bc}$.

	Belle	Belle II
N_{bkg}	$14.05 \pm 1.60 \pm 1.85$	$3.48 \pm 0.73 \pm 0.91$
N_{obs}	11	6
$\epsilon_{\text{sig}} (\times 10^{-5})$	$1.4 \pm 0.05 \pm 0.15$	$1.26 \pm 0.04 \pm 0.17$
$\mathcal{B}(B^+ \to K^+ \tau^+ \tau^-) (\times 10^{-4})$	$-2.76^{+3.31}_{-2.70} \pm 2.24$	$5.05^{+5.62}_{-4.27} \pm 2.46$
Observed (expected) limit	$0.6\ (1.0) \times 10^{-3}$	$2.1\ (1.2) \times 10^{-3}$

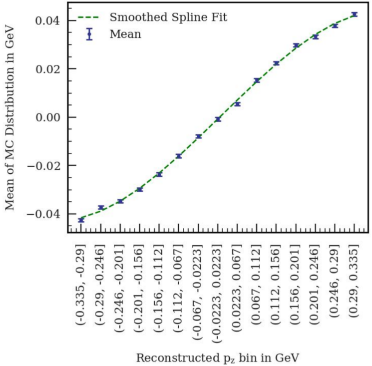
No evidence for signal, yielding combined best upper limit of $\mathcal{B} < 0.87 \cdot 10^{-3}$ at 90% CL.

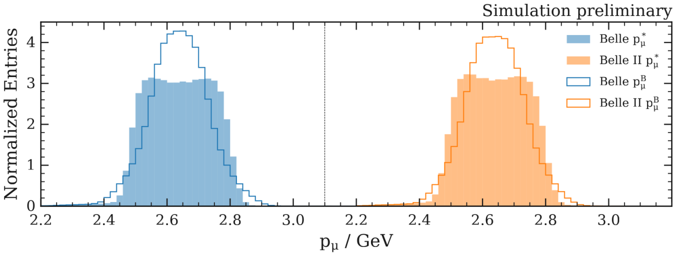




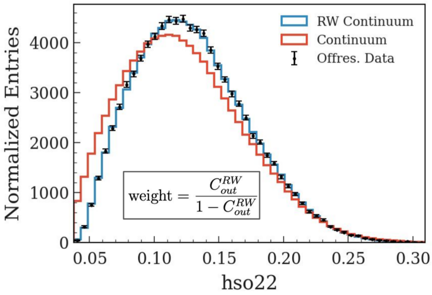
$B^+ \to \mu^+\nu_\mu$ with Inclusive Tag

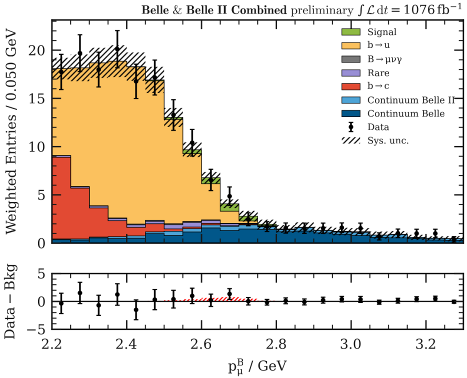
- ITA approach with single high-momentum μ on signal side for **Belle and Belle II**.
- Boost μ to approximate B_{sig} rest frame using p_{ROE} knowledge.
  Signal peaks at $p_\mu^B = 2.64$ GeV in B rest frame with improved resolution.
- Calibration of p_{ROE} in simulation accounts for unreconstructed particles.





Signal Extraction in $B^+ \to \mu^+\nu_\mu$





- **Re-weight continuum simulation** using BDT η_{RW} trained to distinguish simulated and observed continuum.
- Validate efficiencies and p_μ^B resolution in $B \to D^0[K\pi]\pi$ control channel.
- Binned maximum likelihood fit of p_μ^B in 4 bins of BDT η for each Belle & Belle II.

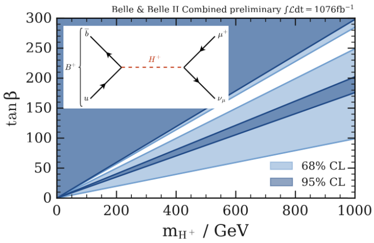
$$\mathcal{B} = [4.36 \pm 1.89(\mathrm{stat.}) \pm 1.03(\mathrm{syst.})]\cdot 10^{-7} \Rightarrow \boldsymbol{\mathcal{B} < [7.13(\mathrm{bay.}), 6.24(\mathrm{freq.})] \cdot 10^{-7}}\ 90\%\ \mathrm{CL}$$

BSM constraints from $B^+ \to \mu^+\nu$

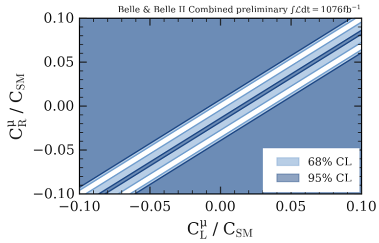
Extracted $|V_{ub}|$ consistent with in- and exclusive measurements.

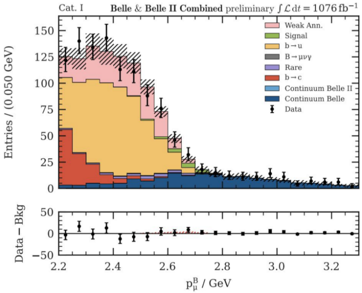
For $p_l^B > 2.2\,\text{GeV}$, $\Delta\mathcal{B}(B \to X_u l\nu) = (0.286 \pm 0.031) \cdot 10^{-3}$.

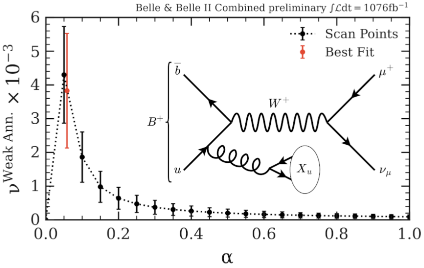
$$\mathcal{B}(B^+ \to \mu^+\nu_\mu) = \mathcal{B}_{SM} \times \left(1 - \frac{m_B^2 \tan^2\beta}{m_{H^+}^2}\right)^2$$



$$\mathcal{B}(B^+ \to \mu^+\nu_\mu) = \mathcal{B}_{SM} \times \left|1 + \frac{m_B^2}{m_b m_\mu}\left(\frac{C_R^\mu}{C_{SM}} - \frac{C_L^\mu}{C_{SM}}\right)\right|^2$$







Summary

We present the latest results by Belle II on

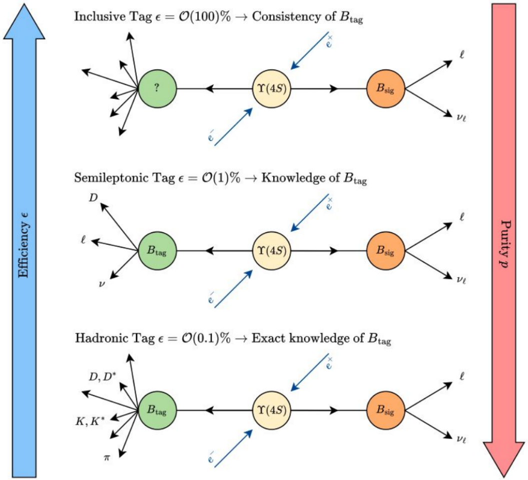
- **Best upper limit on** $B \to X_s \nu\overline{\nu}$ using sum of 30 reconstructed exclusive X_s final states.
- **Published model-agnostic likelihood for anomalous** $B^+ \to K^+\nu\overline{\nu}$. WET reinterpretation favors larger vector and non-zero tensor operators over SM.
- **Best upper limit on** $B^0 \to K^{*0}\tau^+\tau^-$ with total of 16 final states. Twice better than Belle with half the data.
- **Best upper limit for** $B^+ \to K^+\tau^+\tau^-$ **combining Belle and Belle II**, improved by a factor 2.6 relative to previous limits.
- **Best upper limit for** $B^+ \to \mu^+\nu_\mu$ **combining Belle and Belle II**, constraining $|V_{ub}|$, $B \to X_u l\nu$, two-Higgs-doublets and weak annihilation.

Belle II is achieving world-leading results and investigating intriguing anomalies!

Backup

Reconstruction Methods at B factories

https://publikationen.bibliothek.kit.edu/1000078149



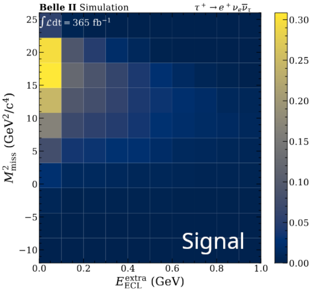
$B^+ \to \tau^+\nu_\tau$ with Hadronic Tag

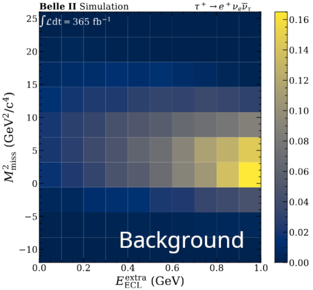
- Reconstruct B_{tag} in HTA and B_{sig} in 4 signal categories, covering 72% of τ decays

$$B^+_{\text{sig}} \to \nu_\tau\tau^+ \left[\to e\overline{\nu}_e,\ \mu\overline{\nu}_\mu,\ \pi,\ \rho(\to \pi\pi^0)\right].$$

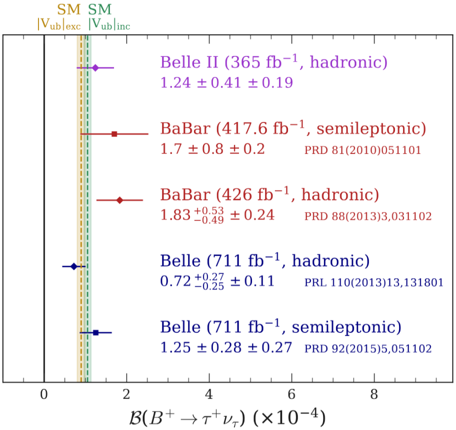
- Require single track with $p > 0.5\,\text{GeV}$.
- Train two BDT classifiers on event topologies for leptonic and hadronic modes.
- Binned maximum likelihood fit for 4 signal categories in $E^{\text{extra}}_{\text{neutral clusters}} \times M^2_{\text{miss}}$.

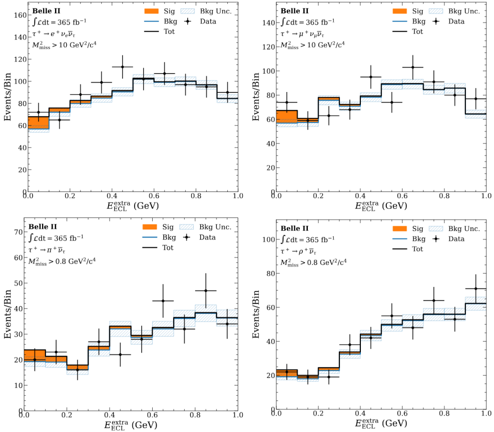
Evidence for signal at
$\mathcal{B} = [\mathbf{1.25 \pm 0.41(stat.) \pm 0.19(syst.)}] \cdot \mathbf{10^{-4}}$.





$|V_{ub}|$ from $B^+ \to \tau^+\nu_\tau$





$$\mathcal{B} = \frac{G_F^2 m_B m_\tau^2}{8\pi}\left[1 - \frac{m_\tau^2}{m_B^2}\right]^2 f_B^2 |V_{ub}|^2 \tau_B$$

$B^+ \to \tau^+\nu_\tau$ provides direct measurement of $|\mathrm{V_{ub}}|$, independent of in- and exclusive semileptonic $\mathrm{B} \to \mathrm{X_u l}$.