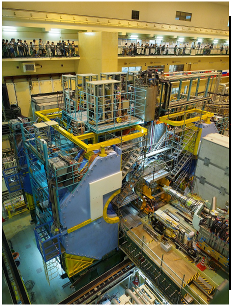
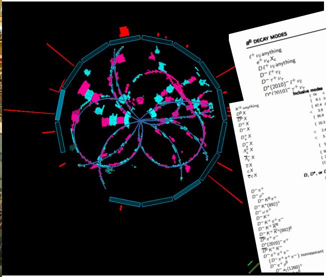
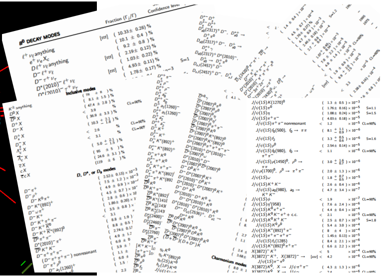
Full Event Interpretation at Belle II







Alexander von Humboldt Stiftung / Foundation

UNIVERSITÄT BONN

Overview

Overview

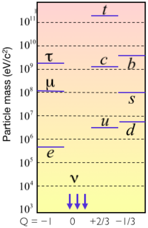
The flavour sector

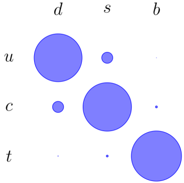
- Nature chose three copies of matter: $\psi \rightarrow \psi_i$
- Identical except for..

  $\mathcal{L} = \psi_i \lambda_{ij} \psi_j h + h.c.$

- This leads to a wide range of flavour phenomenology and puzzles:
  - CP violation
  - quark mixing
  - quark and lepton masses

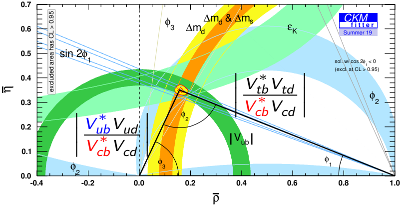
$$\underbrace{\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix}}_{\text{Weak Eigenstates}} = \underbrace{\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}}_{\text{CKM Matrix}} \underbrace{\begin{pmatrix} d \\ s \\ b \end{pmatrix}}_{\text{Mass Eigenstates}}$$

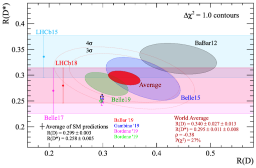


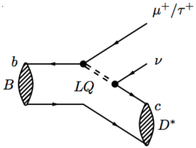


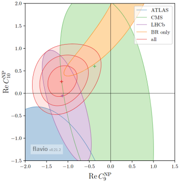
Why semileptonic / missing energy decays?

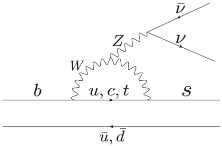
- Precision measurements of the SM:
  - Semileptonic decays are used to determine CKM matrix elements which are essential in global fits for the CKM parameters.
- Excellent probe of new physics:
  - Potential NP in $B \to D^* \tau \nu_\tau$.
  - NP hints in $b \to sll$ should be seen in $b \to s\nu\bar{\nu}$

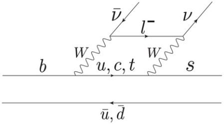


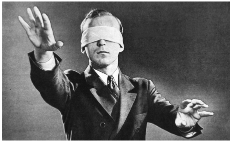








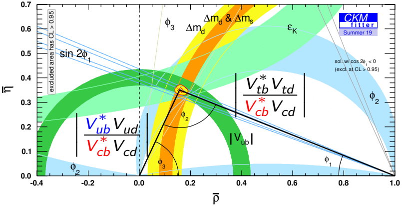


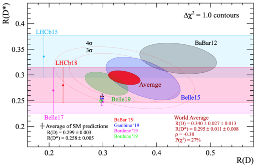


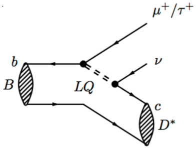
$$R\left(D^{(*)}\right) = \frac{\mathcal{B}(B \to D^{(*)}\tau^- \bar{\nu}_\tau)}{\mathcal{B}(B \to D^{(*)}\mu^- \bar{\nu}_\mu)}$$

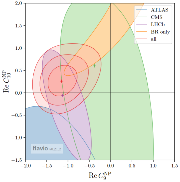
Why semileptonic / missing energy decays?

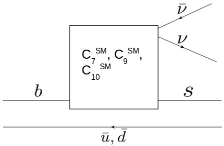
- Precision measurements of the SM:
  - Semileptonic decays are used to determine CKM matrix elements which are essential in global fits for the CKM parameters.
- Excellent probe of new physics:
  - Potential NP in $B \to D^* \tau \nu_\tau$.
  - NP hints in $b \to sll$ should be seen in $b \to s\nu\bar{\nu}$

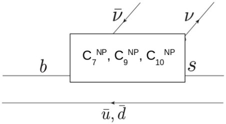


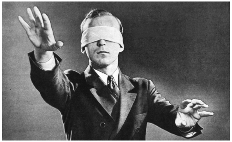








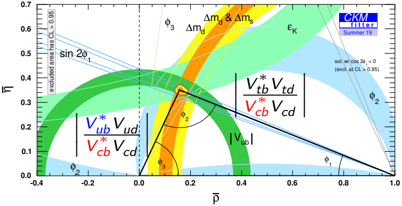


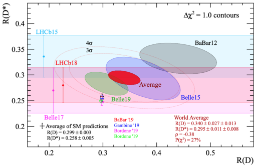


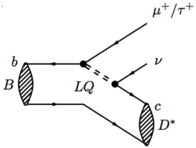
$$R\left(D^{(*)}\right) = \frac{\mathcal{B}(B \to D^{(*)}\tau^- \bar{\nu}_\tau)}{\mathcal{B}(B \to D^{(*)}\mu^- \bar{\nu}_\mu)}$$

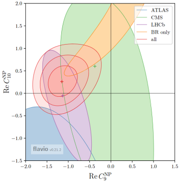
Why semileptonic / missing energy decays?

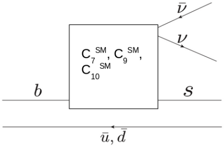
- Precision measurements of the SM:
  - Semileptonic decays are used to determine CKM matrix elements which are essential in global fits for the CKM parameters.
- Excellent probe of new physics:
  - Potential NP in $B \to D^* \tau \nu_\tau$.
  - NP hints in $b \to sll$ should be seen in $b \to s\nu\bar{\nu}$



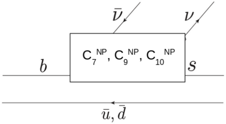






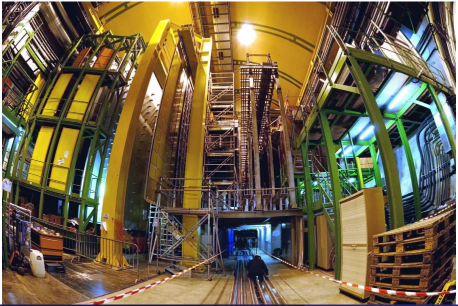


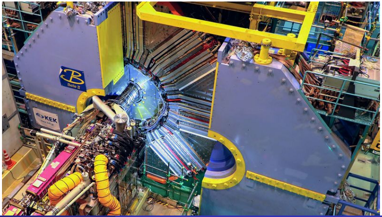
$$R\left(D^{(*)}\right) = \frac{\mathcal{B}(B \to D^{(*)}\tau^-\bar{\nu}_\tau)}{\mathcal{B}(B \to D^{(*)}\mu^-\bar{\nu}_\mu)}$$



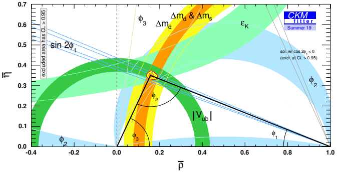
Overview

The B experiments of today

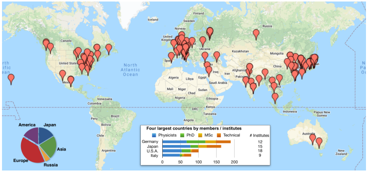




Belle II experiment



- Aim to collect 50 ab^{-1} of e^+e^- collisions at $\sqrt{s} = m_{\Upsilon(4S)}$.
- Wide range of physics: precision CKM measurements, CP violation to new physics searches.

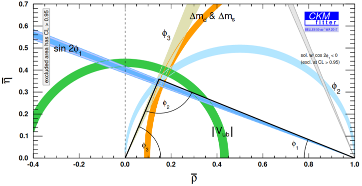


The Belle II Physics Book [arxiv1808.10567]

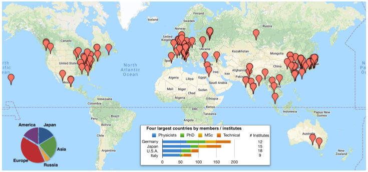
Observables	Expected the. accuracy	Expected exp. uncertainty	Facility (2025)
UT angles & sides			
ϕ_1 [°]	***	0.4	Belle II
ϕ_2 [°]	**	1.0	Belle II
ϕ_3 [°]	***	1.0	LHCb/Belle II
$\|V_{cb}\|$ incl.	***	1%	Belle II
$\|V_{cb}\|$ excl.	***	1.5%	Belle II
$\|V_{ub}\|$ incl.	**	3%	Belle II
$\|V_{ub}\|$ excl.	**	2%	Belle II/LHCb
CP Violation			
$S(B \to \phi K^0)$	***	0.02	Belle II
$S(B \to \eta' K^0)$	***	0.01	Belle II
$\mathcal{A}(B \to K^0\pi^0)[10^{-2}]$	***	4	Belle II
$\mathcal{A}(B \to K^+\pi^-)$ $[10^{-2}]$	***	0.20	LHCb/Belle II
(Semi-)leptonic			
$\mathcal{B}(B \to \tau\nu)$ $[10^{-6}]$	**	3%	Belle II
$\mathcal{B}(B \to \mu\nu)$ $[10^{-6}]$	**	7%	Belle II
$R(B \to D\tau\nu)$	***	3%	Belle II
$R(B \to D^*\tau\nu)$	***	2%	Belle II/LHCb
Radiative & EW Penguins			
$\mathcal{B}(B \to X_s\gamma)$	**	4%	Belle II
$A_{CP}(B \to X_{s,d}\gamma)$ $[10^{-2}]$	***	0.005	Belle II
$S(B \to K_S^0\pi^0\gamma)$	***	0.03	Belle II
$S(B \to \rho\gamma)$	**	0.07	Belle II
$\mathcal{B}(B_s \to \gamma\gamma)$ $[10^{-6}]$	**	0.3	Belle II
$\mathcal{B}(B \to K^*\nu\bar{\nu})$ $[10^{-6}]$	***	15%	Belle II
$R(B \to K^*\ell\ell)$	***	0.03	Belle II/LHCb
Charm			
$\mathcal{B}(D_s \to \mu\nu)$	***	0.9%	Belle II
$\mathcal{B}(D_s \to \tau\nu)$	***	2%	Belle II
$A_{CP}(D^0 \to K_S^0\pi^0)$ $[10^{-2}]$	**	0.03	Belle II
$\|q/p\|(D^0 \to K_S^0\pi^+\pi^-)$	***	0.03	Belle II
$A_{CP}(D^+ \to \pi^+\pi^0)$ $[10^{-2}]$	**	0.17	Belle II
Tau			
$\tau \to \mu\gamma$ $[10^{-10}]$	***	< 50	Belle II
$\tau \to e\gamma$ $[10^{-10}]$	***	< 100	Belle II
$\tau \to \mu\mu\mu$ $[10^{-10}]$	***	< 3	Belle II/LHCb

- Belle II Collaboration: 1050 members, 120 institutes, 26 countries

Belle II experiment



- Aim to collect 50 ab^{-1} of e^+e^- collisions at $\sqrt{s} = m_{\Upsilon(4S)}$.
- Wide range of physics: precision CKM measurements, CP violation to new physics searches.



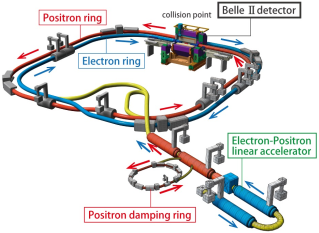
The Belle II Physics Book [arxiv1808.10567]

Observables	Expected the. accuracy	Expected exp. uncertainty	Facility (2025)
UT angles & sides			
ϕ_1 [°]	***	0.4	Belle II
ϕ_2 [°]	**	1.0	Belle II
ϕ_3 [°]	***	1.0	LHCb/Belle II
$\|V_{cb}\|$ incl.	***	1%	Belle II
$\|V_{cb}\|$ excl.	***	1.5%	Belle II
$\|V_{ub}\|$ incl.	**	3%	Belle II
$\|V_{ub}\|$ excl.	**	2%	Belle II/LHCb
CP Violation			
$S(B \to \phi K^0)$	***	0.02	Belle II
$S(B \to \eta' K^0)$	***	0.01	Belle II
$\mathcal{A}(B \to K^0\pi^0)[10^{-2}]$	***	4	Belle II
$\mathcal{A}(B \to K^+\pi^-)$ $[10^{-2}]$	***	0.20	LHCb/Belle II
(Semi-)leptonic			
$\mathcal{B}(B \to \tau\nu)$ $[10^{-6}]$	**	3%	Belle II
$\mathcal{B}(B \to \mu\nu)$ $[10^{-6}]$	**	7%	Belle II
$R(B \to D\tau\nu)$	***	3%	Belle II
$R(B \to D^*\tau\nu)$	***	2%	Belle II/LHCb
Radiative & EW Penguins			
$\mathcal{B}(B \to X_s\gamma)$	**	4%	Belle II
$A_{CP}(B \to X_{s,d}\gamma)$ $[10^{-2}]$	***	0.005	Belle II
$S(B \to K_S^0\pi^0\gamma)$	***	0.03	Belle II
$S(B \to \rho\gamma)$	**	0.07	Belle II
$\mathcal{B}(B_s \to \gamma\gamma)$ $[10^{-6}]$	**	0.3	Belle II
$\mathcal{B}(B \to K^*\nu\bar{\nu})$ $[10^{-6}]$	***	15%	Belle II
$R(B \to K^*\ell\ell)$	***	0.03	Belle II/LHCb
Charm			
$\mathcal{B}(D_s \to \mu\nu)$	***	0.9%	Belle II
$\mathcal{B}(D_s \to \tau\nu)$	***	2%	Belle II
$A_{CP}(D^0 \to K_S^0\pi^0)$ $[10^{-2}]$	**	0.03	Belle II
$\|q/p\|(D^0 \to K_S^0\pi^+\pi^-)$	***	0.03	Belle II
$A_{CP}(D^+ \to \pi^+\pi^0)$ $[10^{-2}]$	**	0.17	Belle II
Tau			
$\tau \to \mu\gamma$ $[10^{-10}]$	***	< 50	Belle II
$\tau \to e\gamma$ $[10^{-10}]$	***	< 100	Belle II
$\tau \to \mu\mu\mu$ $[10^{-10}]$	***	< 3	Belle II/LHCb

- Belle II Collaboration: 1050 members, 120 institutes, 26 countries

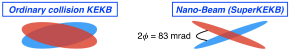
SuperKEKB

- Upgrade of KEKB with original aim $x40\mathcal{L}$

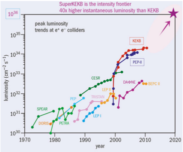


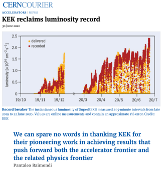
$$L = \frac{\gamma_{e\pm}}{2er_e}\left(1+\frac{\sigma_y^*}{\sigma_x^*}\right)\left(\frac{I_{e\pm}\xi_y^{e\pm}}{\beta_y^*}\right)\left(\frac{R_L}{R_{\xi_y}}\right)$$

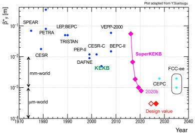
- beam current, $I_{e+/-}$ $\times 1.5$
- Reduction in beam size, β_y, by factor 20

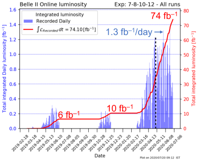


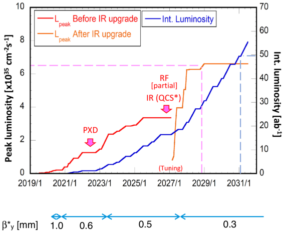
- New aim $x30\mathcal{L}$



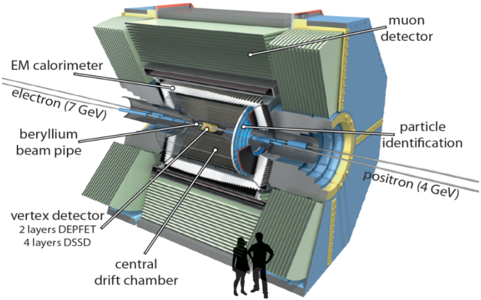


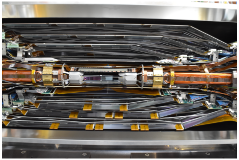


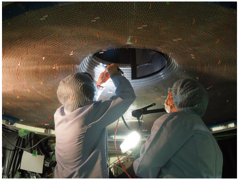




The Belle II Detector



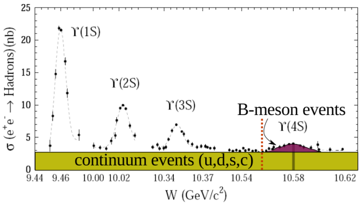


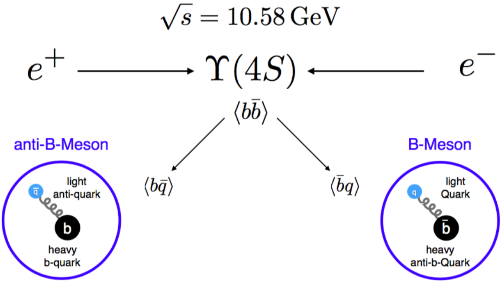


- Inner vertex detector:
  - PXD: 2 layers of DEPFET pixels
  - SVD: 4 layers of DSSD
- Central Drift Chamber for tracking.
- 1.5 T Superconducting solenoid
- Excellent tracking and vertexing down to $p_{\mathrm{T}}{\sim}100$ MeV
- Impact parameter resolution in z $\sim$20 μm
- PID provided by Time of propagation (TOP) counter and a aerogel RICH
- Outer muon and K_{L} detector

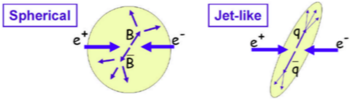
Unique B factory topology

- Collide e^+ and e^- at the energy to make $\Upsilon(4S)$ particles
- $\Upsilon(4S)$ decays to B^+B^- and $B^0\bar{B}^0$ 96% of the time.
- Background from $e^+\ e^- \to q\bar{q}$, $q = u, d, c, s$





Different event topologies

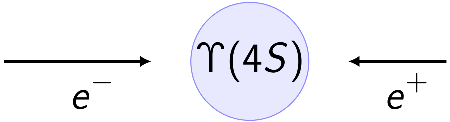


B-mesons Continuum

Overview

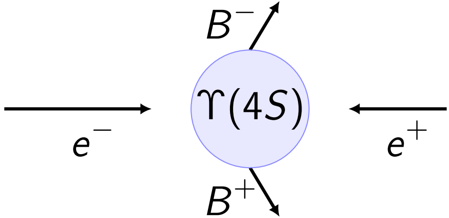
Tag-side B reconstruction

- Collide e^+ and e^- at the energy to make $\Upsilon(4S)$ particles.



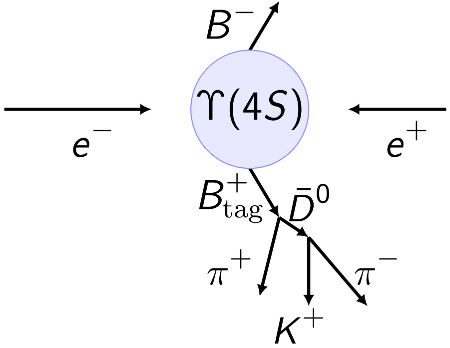
Tag-side B reconstruction

- Collide e^+ and e^- at the energy to make $\Upsilon(4S)$ particles.
- $\Upsilon(4S)$ decays to B^+B^- and $B^0\bar{B}^0$ over 96% of the time.



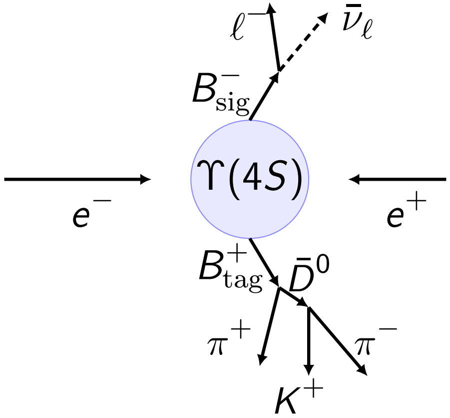
Tag-side B reconstruction

- Collide e^+ and e^- at the energy to make $\Upsilon(4S)$ particles.
- $\Upsilon(4S)$ decays to B^+B^- and $B^0\bar{B}^0$ over 96% of the time.
- Reconstruct one B meson as tag-side (B_{tag}) hadronic or SL.



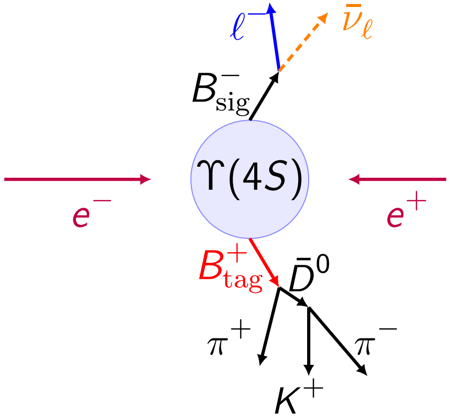
Tag-side B reconstruction

- Collide e^+ and e^- at the energy to make $\Upsilon(4S)$ particles.
- $\Upsilon(4S)$ decays to B^+B^- and $B^0\bar{B}^0$ over 96% of the time.
- Reconstruct one B meson as tag-side (B_{tag}) hadronic or SL.
- Study remaining B meson as signal (B_{sig}).



Tag-side B reconstruction

- Collide e^+ and e^- at the energy to make $\Upsilon(4S)$ particles.
- $\Upsilon(4S)$ decays to B^+B^- and $B^0\bar{B}^0$ over 96% of the time.
- Reconstruct one B meson as tag-side (B_{tag}) hadronic or SL.
- Study remaining B meson as signal (B_{sig}).
- Flavour constraints:
  $B^+_{\text{tag}} \implies B^-_{\text{sig}}$
  Kinematic constraints:
  $p_\nu = p_{e^+e^-} - p_{\ell^-} - p_{B^+}$



Which tag-side reconstruction?

Tagging techniques

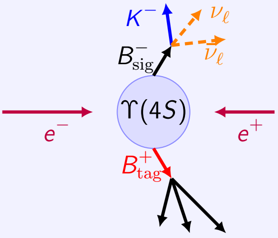
Purity →

← Efficiency

Inclusive

$B \to \text{anything}$

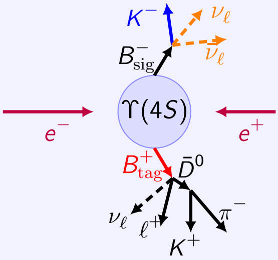
$\epsilon \approx \mathcal{O}(100\%)$



Semileptonic

$B \to D^{(*)}\ell\nu_\ell$

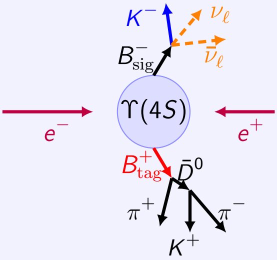
$\epsilon \approx \mathcal{O}(1\%)$



Hadronic

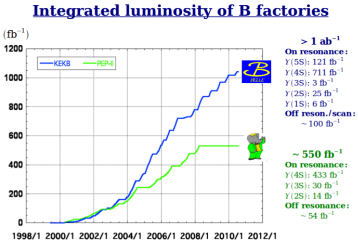
$B \to \text{hadrons}$

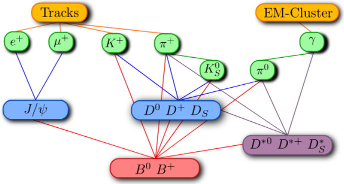
$\epsilon \approx \mathcal{O}(0.1\%)$



Previous tagging alogrithms

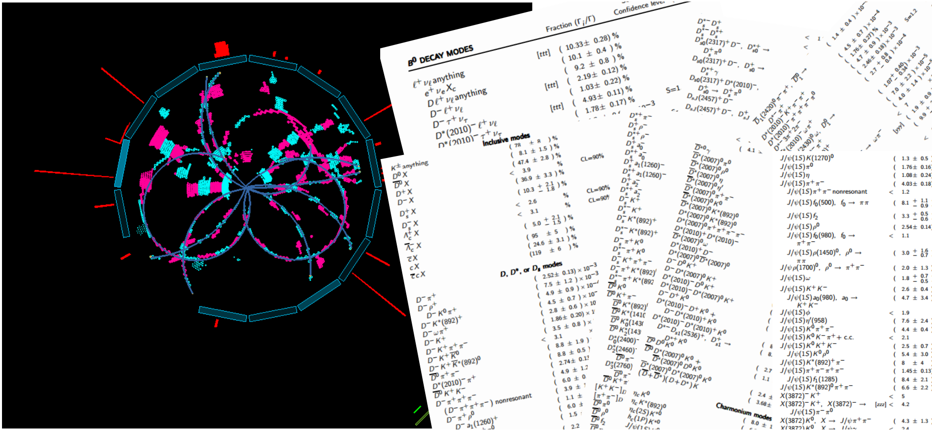
- Full Reconstruction
  - The Belle tagging algorithm and predecessor of the FEI.
  - Hierarchal approach.
  - Neurobayes Neural Network used for classifiers.
- Semi-exclusive-reconstruction
  - The BaBar tagging algorithm
  - Uses D and D^* mesons as a seed.
  - Combines these with up to 5 charmless mesons.



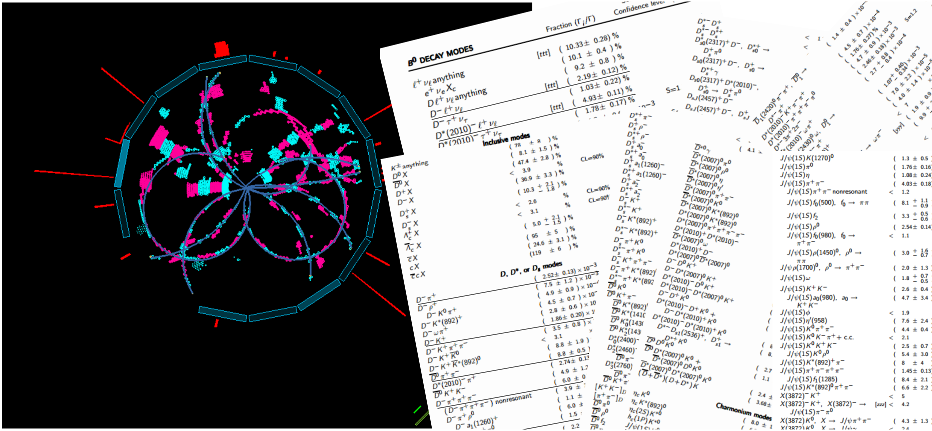


Overview

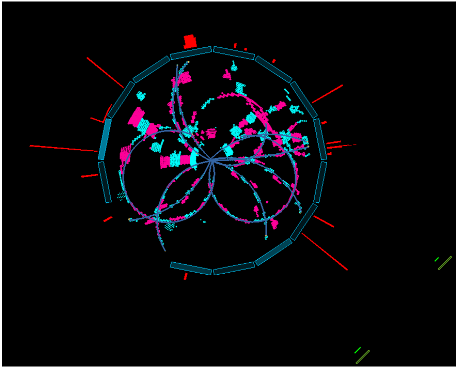
The Task


B0 DECAY MODES
Fraction (Γi/Γ)
Confidence level
Inclusive modes
D, D*, or Ds modes
Charmonium modes

The Task


B0 DECAY MODES
Fraction (Γi/Γ)
Confidence level
Inclusive modes
D, D*, or Ds modes
Charmonium modes

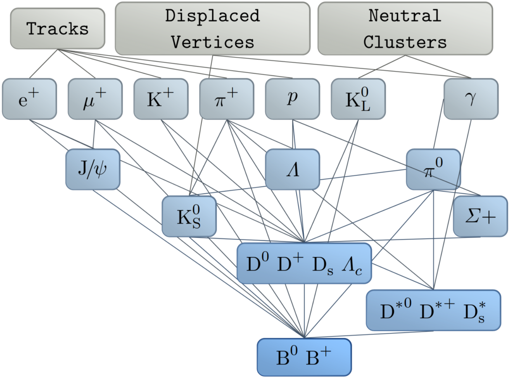
Combinatorics



- $\sim$10 tracks in this event
- Let's assume 5 positively charged and 5 negatively charged.
- Now lets reconstruct $D^0 \to K^-\pi^+\pi^+\pi^-$
- $\binom{5}{2}^2 = 100$ possible combinations
- Reconstructing $B^+ \to (D^0 \to K^-\pi^+\pi^+\pi^-)\pi^+$ introduces $\binom{3}{1} \times 100 = 300$ combinations.

The Full Event Interpretation

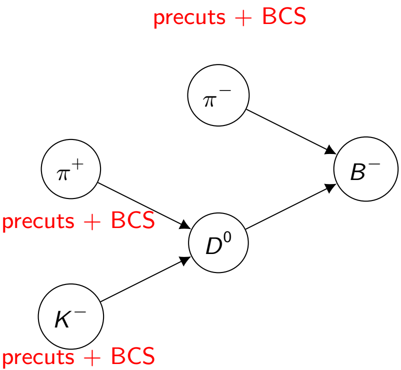
- Utilises O(200) decay channels with a classifiers trained for each.
- Reconstructs O(10000) unique decays chains in six stages.
- Baryonic decays recently added.



Keck, T. et al. Comput Softw Big Sci (2019) 3: 6.

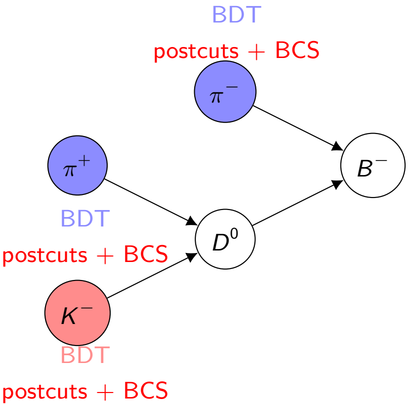
The Algorithm

- Particle candidates assigned from tracks and clusters after a precuts + Best Candidate Selection (BCS).



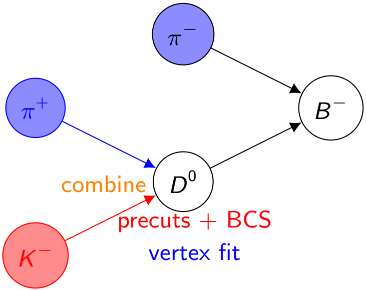
The Algorithm

- Particle candidates assigned from tracks and clusters after a precuts + Best Candidate Selection (BCS).
- For each particle a pre-trained BDT is applied and post cuts + BCS are made.



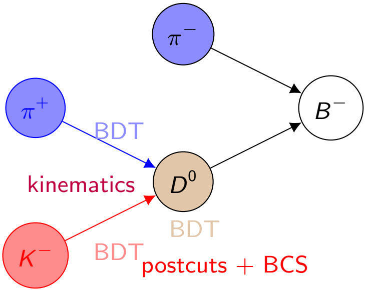
The Algorithm

- Particle candidates assigned from tracks and clusters after a precuts + Best Candidate Selection (BCS).
- For each particle a pre-trained BDT is applied and post cuts + BCS are made.
- Stable particles are combined to reconstruct decays of intermediate particles. After precuts + BCS a vertex fit is performed.



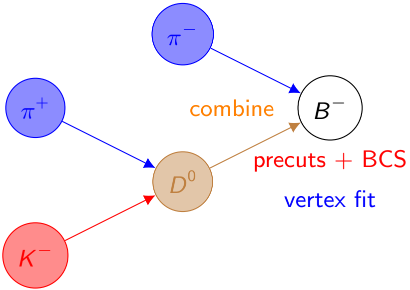
The Algorithm

- Particle candidates assigned from tracks and clusters after a precuts + Best Candidate Selection (BCS).
- For each particle a pre-trained BDT is applied and post cuts + BCS are made.
- Stable particles are combined to reconstruct decays of intermediate particles. After precuts + BCS a vertex fit is performed.
- Intermediate classifiers use daughter kinematics and classifiers.



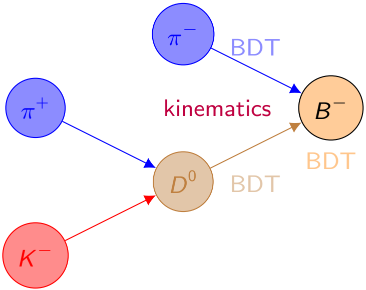
The Algorithm

- Particle candidates assigned from tracks and clusters after a precuts + Best Candidate Selection (BCS).
- For each particle a pre-trained BDT is applied and post cuts + BCS are made.
- Stable particles are combined to reconstruct decays of intermediate particles. After precuts + BCS a vertex fit is performed.
- Intermediate classifiers use daughter kinematics and classifiers.
- Intermediates and stable particles are combined into a B candidate.



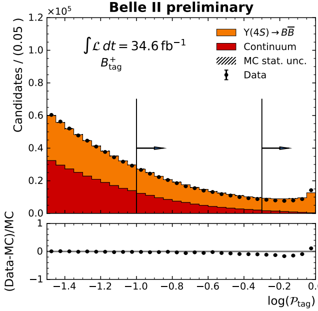
The Algorithm

- Particle candidates assigned from tracks and clusters after a precuts + Best Candidate Selection (BCS).
- For each particle a pre-trained BDT is applied and post cuts + BCS are made.
- Stable particles are combined to reconstruct decays of intermediate particles. After precuts + BCS a vertex fit is performed.
- Intermediate classifiers use daughter kinematics and classifiers.
- Intermediates and stable particles are combined into a B candidate.
- B classifier takes daughter classifiers and kinematics as inputs.

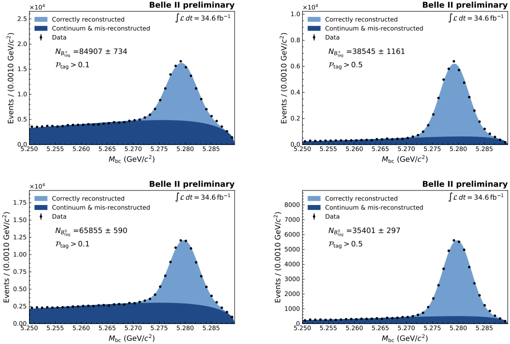


Role of the tag-side B classifier.

- B classifier value, $\mathcal{P}_{\text{tag}}$, discriminates correctly reconstructed tag-sides from background.



- Determine the correctly reconstructed tag-side yield by fitting M_{bc}.

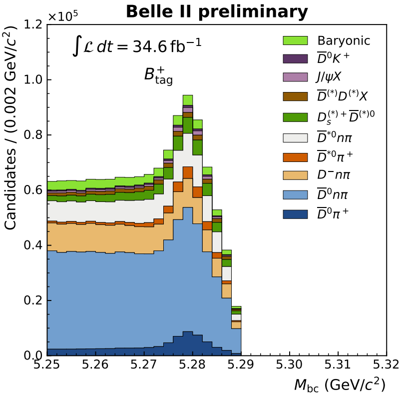


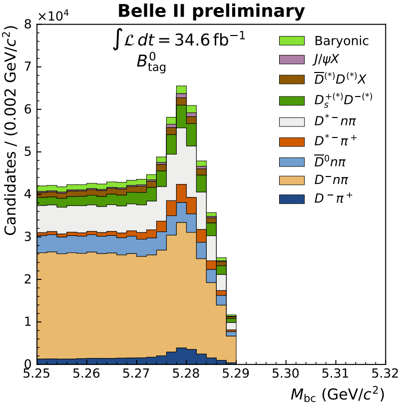
- Select a high purity sample by cutting on $\mathcal{P}_{\text{tag}}$.

$$M_{bc} = \sqrt{E^2_{beam}/4 - (p^{\text{cm}}_{B_{\text{tag}}})^2}$$

Hadronic tag-sides by decay mode

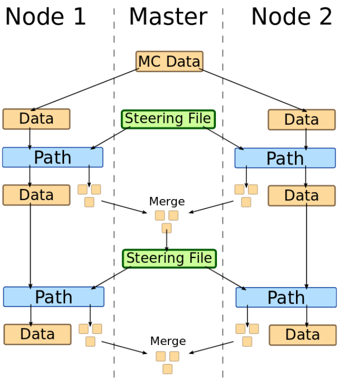
- 29 and 26 hadronic B^+ and B^0 tag-side decay modes are reconstructed.
- Contribution of different categories of modes are shown for data below.





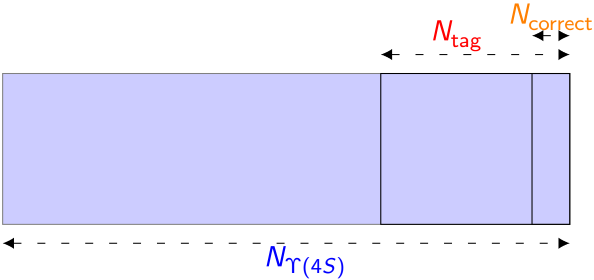
Training the FEI

- Both training and application phases can be distributed via a map reduce approach.
- For training:
  - O(100M) simulated $\Upsilon(4S) \to B\bar{B}$ events
  - Monte carlo / data is partitioned and processed at different nodes.
  - A prereconstruction stage aggregates statistics on MC particles present.
  - At each of the reconstruction phases training data is generated.
  - Training data of each stage is subsquently merged and classifiers trained.



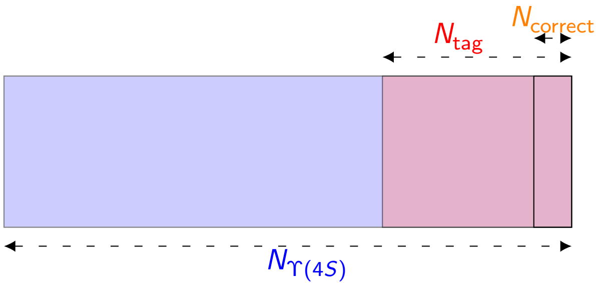
Overview

How does one quantify tagging performance?



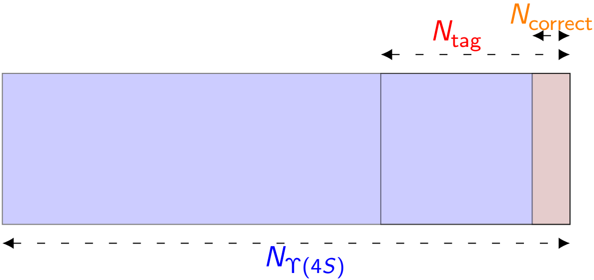
- **tagging efficiency** $= N_{\text{tag}} / N_{\Upsilon(4S)}$

How does one quantify tagging performance?



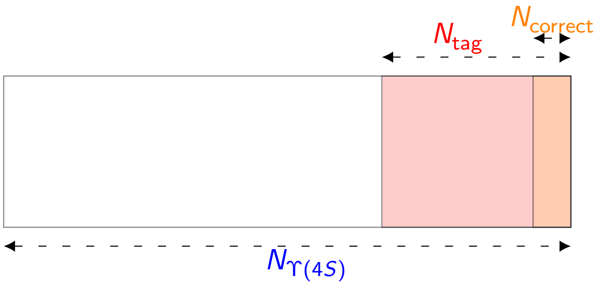
- **tagging efficiency** $=$ $N_{tag}/N_{\Upsilon(4S)}$

How does one quantify tagging performance?



- **tagging efficiency** $= N_{\text{tag}}/N_{\Upsilon(4S)}$
- **tag-side efficiency** $= N_{\text{correct}}/N_{\Upsilon(4S)}$

How does one quantify tagging performance?



- **tagging efficiency** $= N_{\text{tag}} / N_{\Upsilon(4S)}$
- **tag-side efficiency** $= N_{\text{correct}} / N_{\Upsilon(4S)}$
- **purity** $= N_{\text{correct}} / N_{\text{tag}}$

Tagging performance in Monte Carlo

- The table below summarises maximum tag-side efficiency.
- The FEI has over a factor two higher hadronic maximum tag-side efficiency then previous methods (FR and SER).

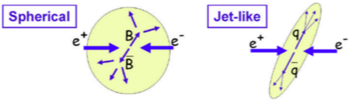
Tag	FR	SER	FEI Belle MC	FEI Belle II MC
Hadronic B^+	0.28%	0.4%	0.76%	0.66%
SL B^+	0.31%	0.3%	1.80%	1.45%
Hadronic B^0	0.18%	0.2%	0.46%	0.38%
SL B^0	0.34%	0.6%	2.04%	1.94%

Tagging performance in Belle data

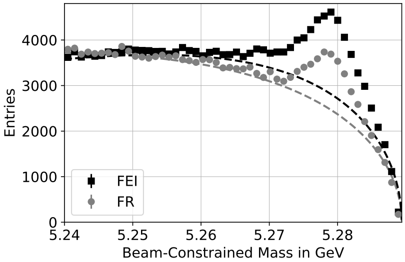
$$m_{bc} = \sqrt{E_{\mathrm{B}}^2 - p_B^2}$$

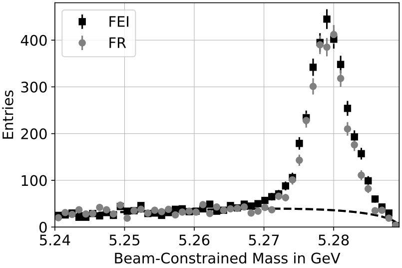
$$E_B = \sqrt{s}/2$$

Different event topologies



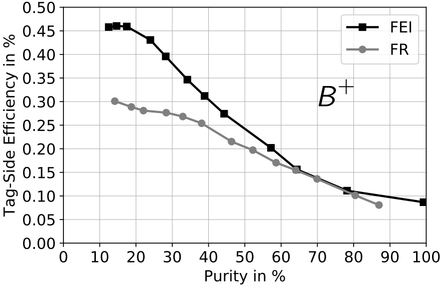
B-mesons Continuum

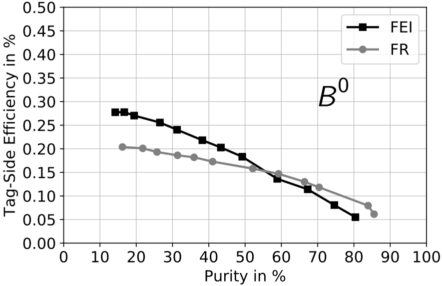




Tagging performance in Belle data

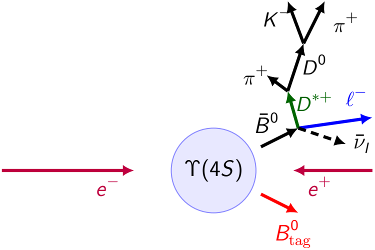
- Compare the tag-side efficiency vs purity for charged and neutral tags between the FEI and FR.

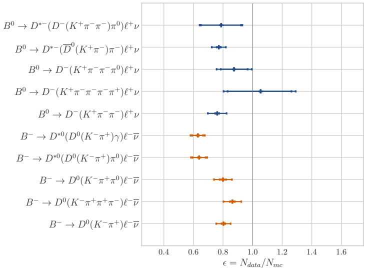




Calibration of FEI with Belle data

- A calibration is required due to signifcant differences in the efficiency in MC and Data.
- Use the FEI on Belle data to reconstruct several well known $B \to D^{(*)}\ell\nu$ semileptonic decays.
- $\epsilon = N_{DATA}/N_{MC}$

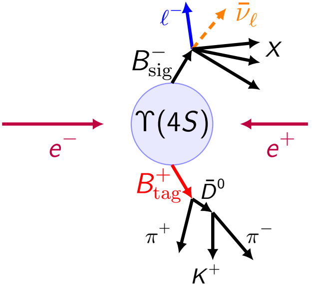




$\epsilon_{\text{charged}} = 0.74 \pm 0.05$ $\epsilon_{\text{neutral}} = 0.86 \pm 0.07$

Calibrating the FEI at Belle II data

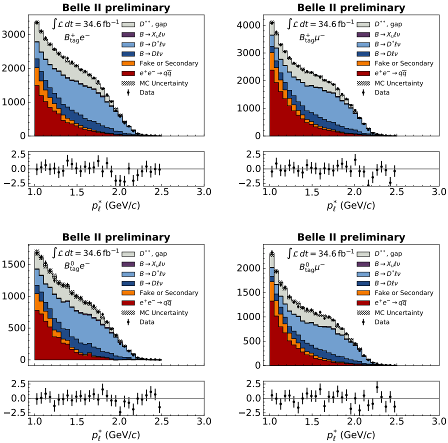
- Can calibrate the FEI by measuring a signal-side (arxiv2008.06096)
- Use $B \to Xl\nu$ given the large branching fraction ($\sim$20%).



- $M_{bc} > 5.27$ GeV$/c^2$, $\mathcal{P}_{\rm tag} > 0.001$, 0.01, 0.1, Lepton ID, $p^*_\ell > 1$ GeV$/c$

$*$ $\Longrightarrow$ B Rest Frame

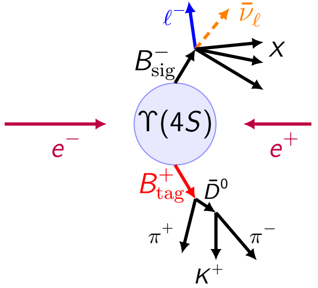
- Calibration factor, $\epsilon_{cal} = N^{X\ell\nu}_{Data}/N^{X\ell\nu}_{MC}$



Here $\mathcal{P}_{\rm tag} > 0.001$

Calibrating the FEI at Belle II data

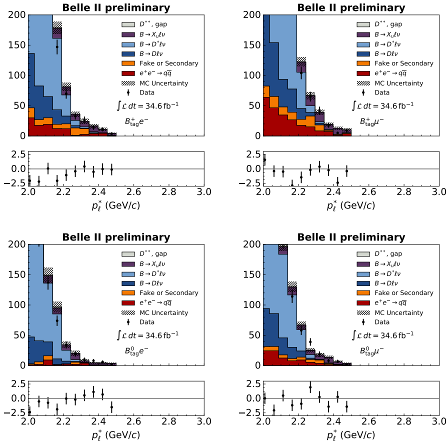
- Can calibrate the FEI by measuring a signal-side (arxiv2008.06096)
- Use $B \to Xl\nu$ given the large branching fraction ($\sim$20%).



- $M_{bc} > 5.27$ GeV$/c^2$, $\mathcal{P}_{\text{tag}} > 0.001$, 0.01, 0.1, Lepton ID, $p^*_\ell > 1$ GeV$/c$

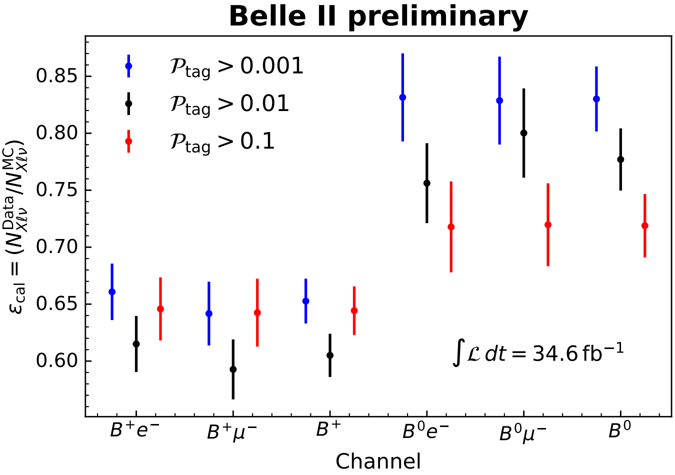
* $\implies$ B Rest Frame

- Calibration factor, $\epsilon_{cal} = N^{X\ell\nu}_{Data} / N^{X\ell\nu}_{MC}$



Here $\mathcal{P}_{\text{tag}} > 0.001$

Calibration results



Calibration results

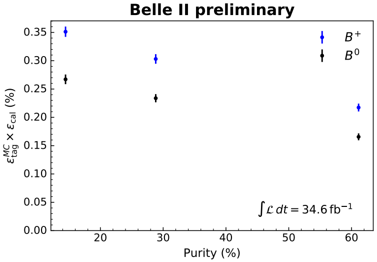
$\mathcal{P}_{B^+} >$	ϵ	% uncertainty
0.001	0.653 ± 0.020	3.02
0.01	0.605 ± 0.019	3.13
0.1	0.644 ± 0.021	3.30

$\mathcal{P}_{B^0} >$	ϵ	% uncertainty
0.001	0.830 ± 0.029	3.44
0.01	0.777 ± 0.027	3.51
0.1	0.719 ± 0.028	3.87

Sources of uncertainty in %

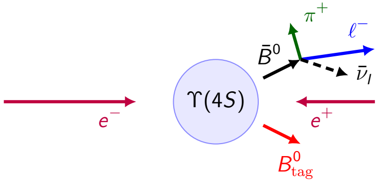
Channel	Fit Model	$\mathcal{B}(B^{0/+} \to X\ell\nu)$	Lepton ID	Fit Stat.	Tracking	MC Stat.	$D^*\ell\nu$ FF	$D\ell\nu$ FF
B^+e^-	2.67	2.09	0.76	0.93	0.91	0.39	0.41	0.06
$B^+\mu^-$	2.93	2.1	2.13	0.86	0.91	0.37	0.38	0.06
B^0e^-	3.72	2.1	0.73	1.22	0.91	0.62	0.43	0.07
$B^0\mu^-$	3.17	2.09	2.13	1.19	0.91	0.6	0.41	0.06

- Tag-side efficiency in simulation against purity corrected by ϵ_{cal}.
- Tag-side efficiency = No. of events with a correctly reconstructed tag-side (N_{corr}) / No. of $\Upsilon(4S) \to B\bar{B}$
- Purity = N_{corr} / No. of events with a tag-side

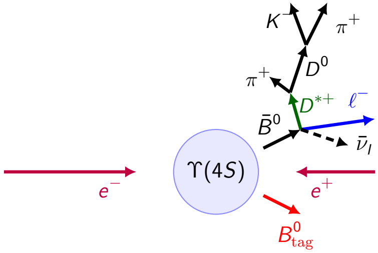


Overview

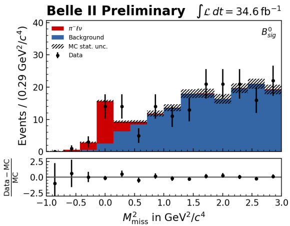
Rediscovering $B \to \pi\ell\nu$ and $B \to D^*\ell\nu$ with tagging

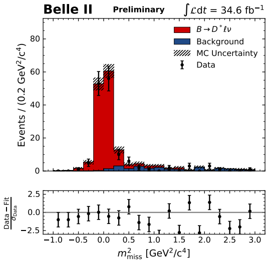


$$m^2_{\text{miss}} = (p_{e^+e^-} - p_{B_{\text{tag}}} - p_\ell - p_{\pi/D^*})^2$$



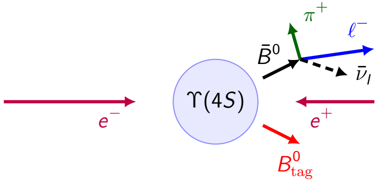
- Data-simulation comparisons with the calibration applied.



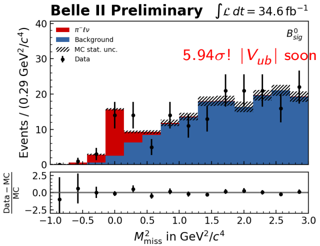


$\pi\ell\nu$: arxiv2008.08819 $D^*\ell\nu$: arxiv2008.10299

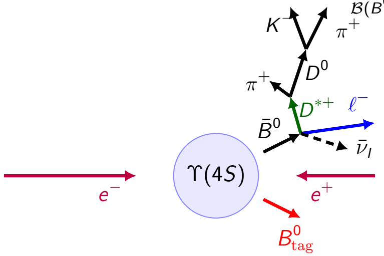
Rediscovering $B \to \pi \ell \nu$ and $B \to D^* \ell \nu$ with tagging



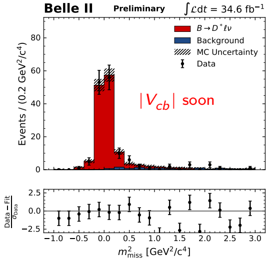
$$\mathcal{B}(B \to \pi \ell \nu) = (1.62 \pm 0.42(\text{stat}) \pm 0.07(\text{sys})) \times 10^{-4}$$



$$m^2_{\rm miss} = (p_{e^+e^-} - p_{B_{\rm tag}} - p_\ell - p_{\pi/D^*})^2$$



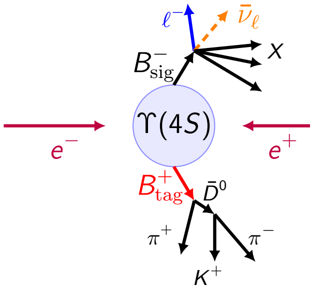
$$\mathcal{B}(B^0 \to D^{*+} \ell \nu) = (4.45 \pm 0.41(\text{stat}) \pm 0.27(\text{sys}) \pm 0.45(\pi_s)) \times 10^{-2}$$



$\pi\ell\nu$: arxiv2008.08819 $D^*\ell\nu$: arxiv2008.10299

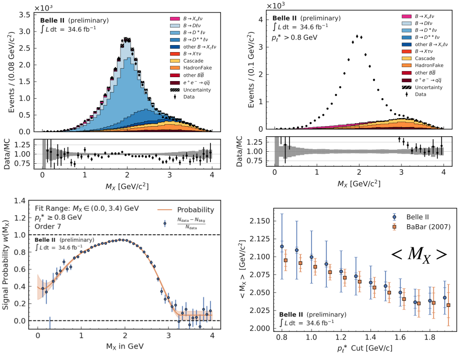
First physics: M_X moments of $B \to X\ell\nu$ decays.

- Measurement of the M_X moments [arxiv2009.04493]



- Plan to perform the first measurement of the q^2 moments in the near future.

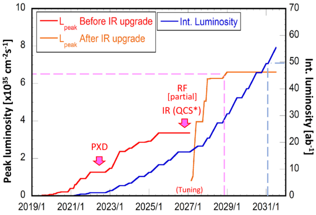
- Fit M_X functional form after a background subtraction.
- Determine M_X moments after correcting for detector and resolution effects.

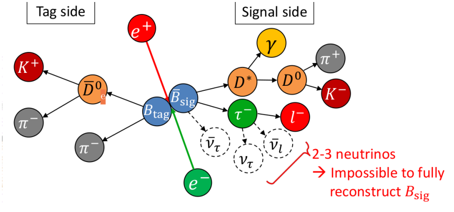


Overview

FEI in Belle 2

- By 2031 2-3$\times10^8$ B^+ and 1-2$\times10^8$ B^0 tags.



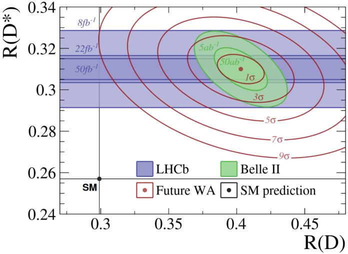


- Eventually systematically limited by tagging calibration

	5 ab^{-1}	50 ab^{-1}
R_D	$(\pm6.0\pm3.9)\%$	$(\pm2.0\pm2.5)\%$
R_{D^*}	$(\pm3.0\pm2.5)\%$	$(\pm1.0\pm2.0)\%$

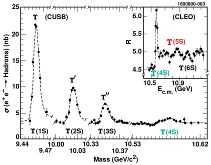
stat. sys.

- LHCb and Belle II will resolve $R(D^*)$ anomaly.



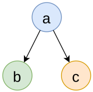
FEI Developments and the Future

- Algorithm has been successfully applied to the $\Upsilon(5S)$ resonance.



- Graph networks naturally suit particle decays.

Input graph



Adjacency matrix

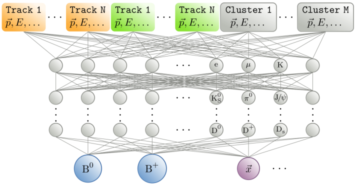
	a	b	c
a	0	1	1
b	1	0	0
c	1	0	0

Feature matrix

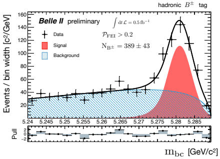
a0	a1
b0	b1
c0	c1

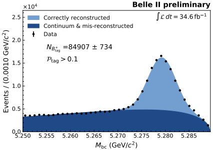
Figure: Example of matrices fully defining a graph.

- Exploring deep extensions of the FEI.



- We can look forward to exciting physics results from the growing number of B tags at Belle II!





Overview

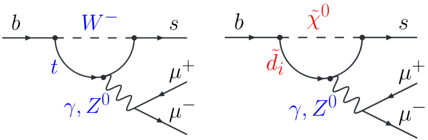
Conclusion

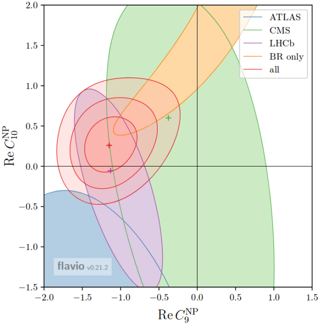
- The Full Event Interpretation (FEI) is an algorithm for tag-side B reconstruction at Belle 2.
- It trains O(200) decay channel classifiers which are used in the reconstruction of O(10000) decay chains.
- The FEI outperforms its predecessors with a higher tag-side efficiency.
- The FEI has been used to measure $B \to X\ell\nu$, $B \to D^{*}\ell\nu$ and $B \to \pi\ell\nu$ decays in early Belle II data.
- The FEI is an essential to the Belle II physics program and resolving the B physics anomalies.

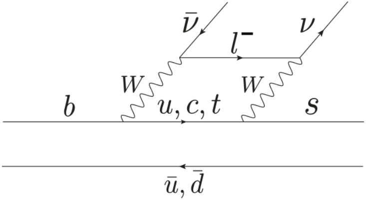
Overview

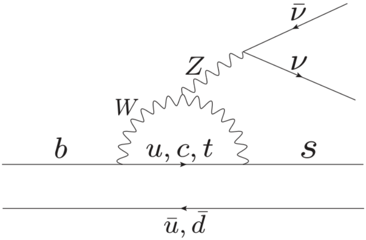
Flavour anomalies

$$\mathcal{H}_{\text{eff}} = -\frac{4\,G_F}{\sqrt{2}} V_{tb} V_{ts}^{*} \frac{e^2}{16\pi^2} \sum_i (C_i O_i + C_i' O_i') + \text{h.c.}$$



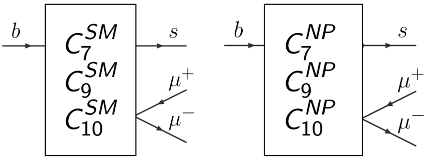


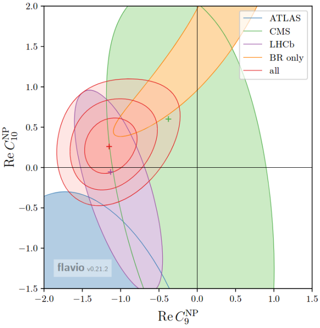


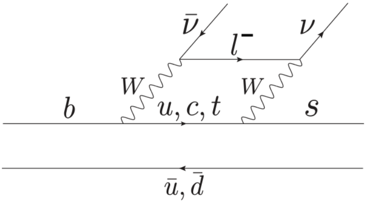


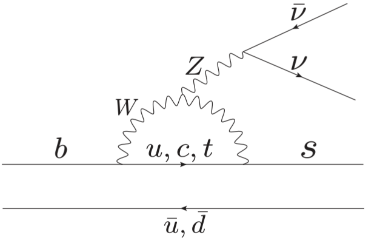
Flavour anomalies

$$\mathcal{H}_{\text{eff}} = -\frac{4\,G_F}{\sqrt{2}} V_{tb} V_{ts}^{*} \frac{e^2}{16\pi^2} \sum_i (C_i O_i + C_i' O_i') + \text{h.c.}$$



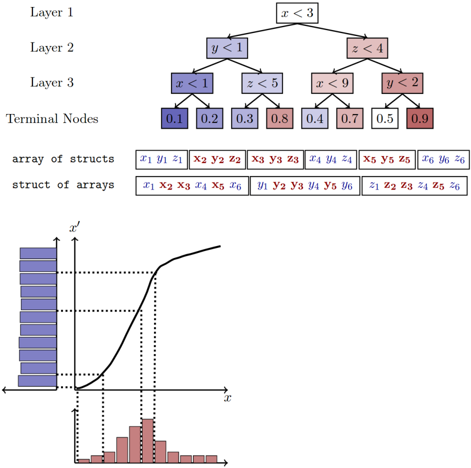






Need for speed

- Utilise FastBDT:
  - Computes cumulative probability histograms (CPH) of nodes in the same level simulataneously.
  - Stores data as an array of structs.
  - BDT cut decisions optimised based on equal frequency bins.
- Utilise FastFit:
  - Uses eigen libraries to gain from vectorisation.
  - Overall factor of 2.7 speed up in the FEI



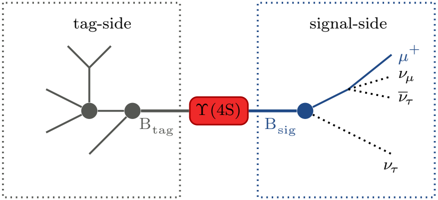
Time break down

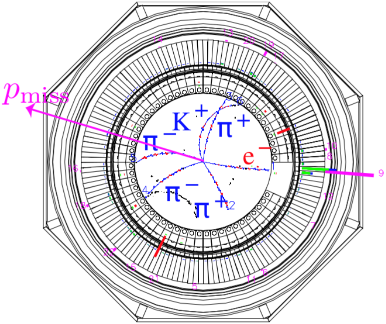
- In application 38% of the time is spent on vertex fitting, 27% on particle combination and 15% on classifier inference.

Task	Training	Application
read/write `DataStore`	30	0
vertex fitting	26	38
particle combination	19	27
classifier inference	11	15
training data & monitoring	6	0
best candidate selection	3	6
other	5	14

Specific vs Generic FEI

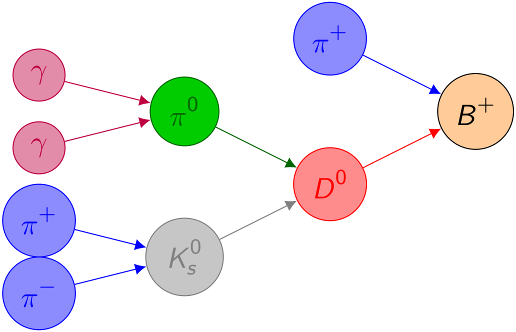
- **Generic FEI** - Reconstruct signal after reconstructing a tag-side B candidate.
- **Specific FEI** - Reconstruct a tag-side B candidate after reconstructing signal





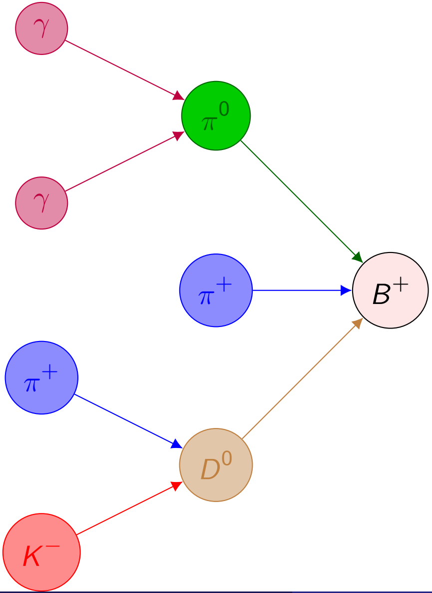
$B^+_{tag} \rightarrow \bar{D}^0\pi^+\pi^+\pi^-$ $\hookrightarrow K^+\pi^-$ $\quad B^-_{sig} \rightarrow \tau^-\bar{\nu}_\tau$ $\hookrightarrow e^-\bar{\nu}_e\nu_\tau$

The Algorithm



- Same $B^+ \to D^0\pi^+$ classifier.
- Different decay chain as $D^0 \to K_s^0\pi^0$.
- $D^0 \to K_s^0\pi^0$ has its own classifier.

The Algorithm



- Different $B^+ \to D^0\pi^+\pi^0$ decay with its own classifier.
- Original D decay chain as $D^0 \to K^-\pi^+$.

References

The Full Event Interpretation – An exclusive tagging algorithm for the Belle II experiment - Thomas Keck et al. https://arxiv.org/abs/1807.08680

Machine learning algorithms for the Belle II experiment and their validation on Belle data - Thomas Keck https://publikationen.bibliothek.kit.edu/1000078149

Analysis Software and Full Event Interpretation for the Belle II Experiment - Christian Pulvermacher http://ekp-invenio.physik.uni-karlsruhe.de/record/48741

FastBDT: A speed-optimized and cache-friendly implementation of stochastic gradient-boosted decision trees for multivariate classification - Thomas Keck https://arxiv.org/abs/1609.06119