$\mathcal{B}(B_c^+ \to \tau^+\nu_\tau)$ prospects at FCC-ee

Donal Hill, Clement Helsens, Yasmine Amhis, Olcyr Sumensari

17/5/21

FCC-ee P&P meeting

Introduction

- Physics motivation and more details on workflow given in previous talk by Clement (see here)
- Study feasibility of measurement of $\mathcal{B}(B_c^+ \to \tau^+\nu_\tau)$ at FCC-ee
- Use $\tau^+ \to 3\pi\nu_\tau$ mode to provide the τ^+ decay vertex, which can help to isolate signal
- Show latest performance of multiple MVAs to reject inclusive $Z \to b\bar{b}/c\bar{c}/q\bar{q}$ background, selection optimisation, toy fit studies, and $\mathcal{B}$ precision estimates

Updates since last meeting

- Moved to official `spring2021` production of events (see here)
    - Include ISR/FSR, Beam Energy Spread, and vertex smearing
    - Produced 28 different exclusive modes (200M events per mode) based on scrutinizing the decays passing tight MVA cuts from inclusive sample
    - Produced inclusive $Z \to b\bar{b}$ events using EvtGen for the decay (500M events, inclusive Pythia is 1B)
- Added the Primary Vertex (PV) constraint
- Fixed some issues with PV being considered in some variables used for MVA discrimination
- Dedicated production of 1B events with orthogonal seeds for training (see here)

Pre-selection before first-stage MVA

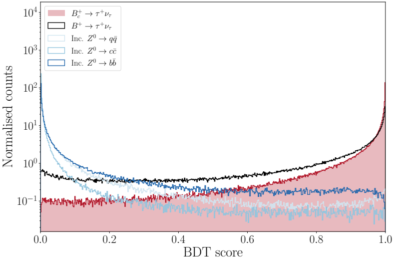
- The following cuts are applied prior to first-stage MVA training:
  - A PV is reconstructed
  - At least one reconstructed 3π candidate in the event
  - One of the 3π must be in the lower energy hemisphere
- Pre-selection cuts applied to signal and inclusive $Z \to b\bar{b}/c\bar{c}/q\bar{q}$ background samples

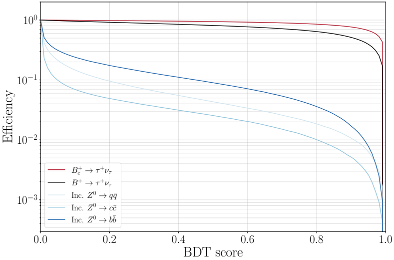
First-stage MVA

- Train on general event-level properties, harnessing difference in hemisphere energies between $B_c^+ \to \tau^+\nu_\tau$ and $Z \to b\bar{b}/c\bar{c}/q\bar{q}$
- Also use information on PV multiplicity, number of vertices reconstructed in event, number of 3π candidates, and distance between decay vertices and PV
- Background training sample is a mix of inclusive $Z \to b\bar{b}/c\bar{c}/q\bar{q}$, combined according to Z branching ratios and efficiency of pre-selection requirements

First-stage MVA performance

- Strong rejection of all three backgrounds, with ROC = **0.984**
    - $Z \to b\bar{b}$ rejected least since it produces most missing energy
- $B^+ \to \tau^+\nu_\tau$ also rejected to an extent (different event-level properties compared to B_c^+)
    - Important background, which is enhanced due to B^+ vs. B_c^+ production rate but lower due to CKM suppression ($|V_{ub}|^2$ vs. $|V_{cb}|^2$)
    - Expect 5 times more events than signal before any cuts





Selection cuts before second-stage MVA

- Apply MVA1 > 0.6 cut to remove “easy” stuff
- Select the reconstructed 3π candidate with the smallest displaced vertex χ^2
- Selected vertex must be in the hemisphere with less energy
- Apply $m(\pi\pi)$ cuts to 3π candidates to select $\rho \to \pi^+\pi^-$ region [0.6, 1.0] GeV
    - Signal decays purely via $a_1 \to \rho\pi$ resonance
- Require $m(3\pi) < m_\tau$, and that the difference in hemisphere energies exceeds 10 GeV

Second-stage MVA

- Trained on signal and inclusive background passing pre-selection (previous slide)

- Use properties of the 3π candidate (mass, momentum, IP, distance from PV, angle of momentum to thrust axis)

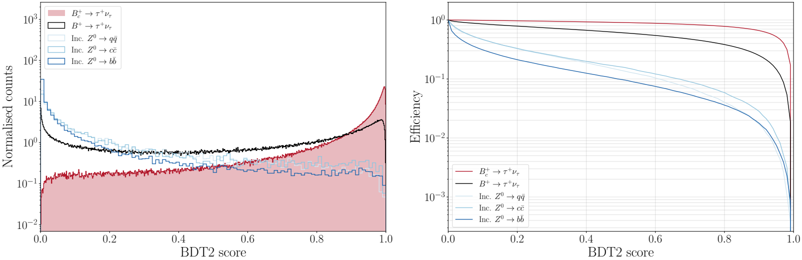
- Use IP to the PV for other decay vertices in the event
    - Associated charm hadron produced with $B_c^+ \to \tau^+\nu_\tau$ originates from PV, unlike the charm hadron in $B^0 \to D^{(*)-}\tau^+\nu_\tau$ for example

- Mass of PV - this is larger for background because the associated charm hadron in $B_c^+ \to \tau^+\nu_\tau$ carries energy away from the PV

- Nominal B energy = $m(Z)$ - sum of all reco. event energy apart from $E(3\pi)$

Second-stage MVA performance

- All backgrounds rejected very well, with ROC = **0.966**
- Excellent rejection of $B^+ \to \tau^+\nu_\tau$ mode as well, even though it is not used in training
  - This reduces the B^+ mode to a level where it must be constrained from independent measurements
  - Described more in toy fit setup later



2D MVA cut optimisation

- Tune the two MVA cuts to maximise purity of the signal $S/(S+B)$, to enable a precise $\mathcal{B}(B_c^+ \to \tau^+\nu_\tau)$ measurement
  - Estimate S and B expected from 5×10^{12} Z at FCC-ee

- Inclusive $Z \to c\bar{c}$ and $Z \to q\bar{q}$ samples are rejected at 10^9 level with sufficiently tight MVA cuts
  - All generated events are rejected
  - Do not consider these in subsequent studies

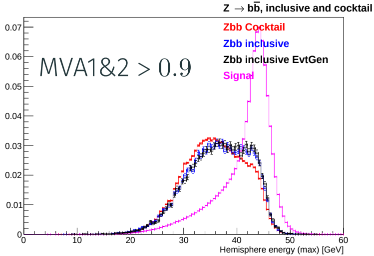
- Inclusive $Z \to b\bar{b}$ statistics are insufficient to assess background rejection to a high enough level or create fit templates
  - Generate a set of exclusive B^0, B^+, B_s^0, and Λ_b^0 modes where all B-hadron decay products are decayed inclusively

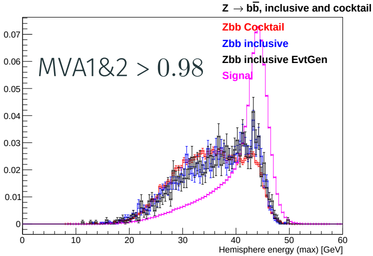
Exclusive background

- Focus on modes where a τ or charm hadron produces a displaced 3π, along with another charm hadron (this topology is most like the signal)
  - $B \to D^{(*)} D_s^{(*)}$
  - $B \to D^{(*)} 3\pi$
  - $B \to D^{(*)} \tau^+ \nu_\tau$

- The sum of all these exclusive modes matches the distributions in inclusive $Z \to b\bar{b}$ rather well
  - An overall rate difference of **2.5** is observed, since we are only generating about 10% of possible B hadron decay widths
  - **Multiply the exclusive sample yields by 2.5 in the cut optimisation to account for this**

Comparing inclusive, exclusive, and signal (normalised)



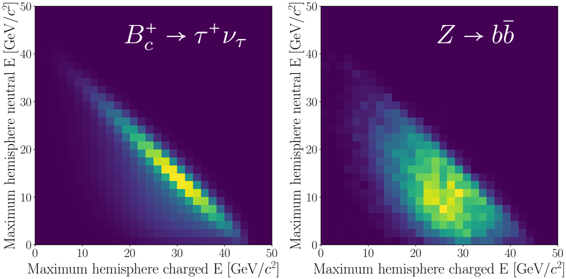


- Comparing shape of inclusive Pythia8 $Z \to b\bar{b}$, inclusive EvtGen $Z \to b\bar{b}$, exclusive sum of B hadron modes, and signal
- Maximum hemisphere energy considered as fit variable
- All plots can be found here

Why use the maximum hemisphere energy?

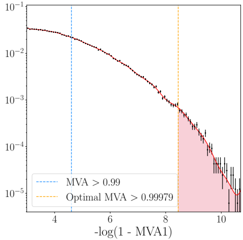
- In an inclusive $Z \to b\bar{b}$ event, either side could have max. E
    - However, in a signal decay, the max. E hemisphere will most likely be the **non-signal side**

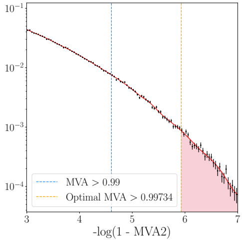
- **Result:** in signal events, max. E hemisphere looks a lot like an inclusive decay (peak near $m(Z)/2$ with tail)
    - In background, our selection requirements bias this hemisphere **down in energy**, giving us discrimination



Exclusive background - efficiency determination

- Full selection leaves too few events in exclusive samples to determine an accurate background efficiency
- Measure efficiency of [MVA1 > 0.95 & MVA2 > 0.95] cut for each exclusive exclusive background mode
- Parameterise shape of remaining MVA distributions using summed sample of exclusive modes
  - Fit $-\log(1-$ MVA$)$ with **cubic spline**
  - Integrate splines above certain cuts to determine efficiencies (and multiply by ϵ[MVA1 > 0.95 & MVA2 > 0.95])





Decay mode	N(expected)	N(generated)	Expected / Generated	Final ϵ
$B^+ \to \bar{D}^0 \tau^+ \nu_\tau$	5.01×10^9	2×10^8	25.0	1.46×10^{-9}
$B^+ \to \bar{D}^{*0} \tau^+ \nu_\tau$	1.22×10^{10}	2×10^8	61.1	1.1×10^{-9}
$B^+ \to \bar{D}^0 3\pi$	3.64×10^9	1.9×10^8	19.2	1.56×10^{-9}
$B^+ \to \bar{D}^{*0} 3\pi$	6.7×10^9	2×10^8	33.5	1.04×10^{-9}
$B^+ \to \bar{D}^0 D_s^+$	5.85×10^9	2×10^8	29.3	2.52×10^{-10}
$B^+ \to \bar{D}^{*0} D_s^+$	4.94×10^9	1.75×10^8	28.2	2.72×10^{-10}
$B^+ \to \bar{D}^{*0} D_s^{*+}$	1.11×10^{10}	2×10^8	55.6	2.42×10^{-10}
$B^0 \to D^- \tau^+ \nu_\tau$	7.02×10^9	2×10^8	35.1	2.69×10^{-9}
$B^0 \to D^{*-} \tau^+ \nu_\tau$	1.02×10^{10}	2×10^8	51.0	1.25×10^{-9}
$B^0 \to D^- 3\pi$	3.9×10^9	2×10^8	19.5	3.4×10^{-9}
$B^0 \to D^{*-} 3\pi$	4.69×10^9	2×10^8	23.4	9.84×10^{-10}
$B^0 \to D^- D_s^+$	4.68×10^9	2×10^8	23.4	3.23×10^{-10}
$B^0 \to D^{*-} D_s^+$	5.2×10^9	2×10^8	26.0	2.32×10^{-10}
$B^0 \to D^{*-} D_s^{*+}$	1.15×10^{10}	2×10^8	57.5	2.35×10^{-10}
$B_s^0 \to D_s^- \tau^+ \nu_\tau$	3.53×10^9	2×10^8	17.6	3.71×10^{-9}
$B_s^0 \to D_s^{*-} \tau^+ \nu_\tau$	2.35×10^9	2×10^8	11.8	2.27×10^{-9}
$B_s^0 \to D_s^- 3\pi$	8.85×10^8	2×10^8	4.4	5.53×10^{-9}
$B_s^0 \to D_s^{*-} 3\pi$	1.05×10^9	2×10^8	5.2	3.38×10^{-9}
$B_s^0 \to D_s^- D_s^+$	6.39×10^8	2×10^8	3.2	4.09×10^{-10}
$B_s^0 \to D_s^{*-} D_s^+$	2.02×10^9	2×10^8	10.1	3.17×10^{-10}
$B_s^0 \to D_s^{*-} D_s^{*+}$	2.09×10^9	2×10^8	10.5	2.56×10^{-10}
$\Lambda_b^0 \to \Lambda_c^- \tau^+ \nu_\tau$	1.83×10^9	2×10^8	9.1	1.36×10^{-9}
$\Lambda_b^0 \to \Lambda_c^{*-} \tau^+ \nu_\tau$	1.83×10^9	2×10^8	9.1	9.44×10^{-10}
$\Lambda_b^0 \to \Lambda_c^- 3\pi$	4.31×10^8	2×10^8	2.2	5.58×10^{-9}
$\Lambda_b^0 \to \Lambda_c^{*-} 3\pi$	4.31×10^8	2×10^8	2.2	9.21×10^{-10}
$\Lambda_b^0 \to \Lambda_c^- D_s^+$	6.15×10^8	2×10^8	3.1	3.46×10^{-10}
$\Lambda_b^0 \to \Lambda_c^{*-} D_s^+$	6.15×10^8	2×10^8	3.1	2.72×10^{-10}
$\Lambda_b^0 \to \Lambda_c^{*-} D_s^{*+}$	6.15×10^8	2×10^8	3.1	2.5×10^{-10}

Cut optimisation - results

- Estimate yields with $N_Z = 5 \times 10^{12}$, $\mathcal{B}(Z \to b\bar{b}) = 0.1512$, and B hadron production fractions from Pythia
- Estimate signal yield using $\mathcal{B}(B_c^+ \to \tau^+ \nu_\tau) = 0.0236$ (SM prediction), $f_c = 0.0004$ from Pythia, and selection efficiency from our signal MC
- Estimate $B^+ \to \tau^+ \nu_\tau$ and exclusive B-hadron background yields using PDG $\mathcal{B}$'s, production fractions, and MC ϵ's
- Best signal yield of **4147** events at **89% purity** ($\epsilon = 0.31\%$)
- **207** events expected for $B^+ \to \tau^+ \nu_\tau$ mode ($\epsilon = 3.1 \times 10^{-5}$)
- Background level of **325** events estimated from a sum of all exclusive modes, scaled by 2.5 (inclusive vs. exclusive ratio)

Toy fit to measure signal yield

- **Energy of the maximum energy hemisphere** used as the fit variable

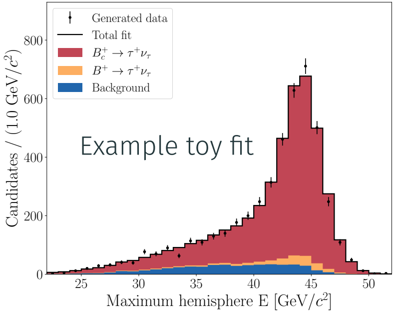
- Create histogram templates in $B_c^+ \to \tau^+\nu_\tau$, $B^+ \to \tau^+\nu_\tau$, and exclusive (B^0, B^+, B_s^0, Λ_b^0) background
  - Use MC with MVA1&2 > 0.99 for templates, to have sufficient stats
  - Sum exclusive backgrounds according to their expected yields from the MVA cut optimisation

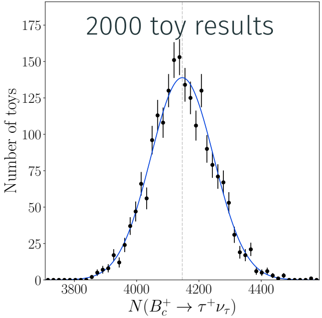
- $B^+ \to \tau^+\nu_\tau$ yield constrained in fit to the value found in cut optimisation, with **5% relative uncertainty**
  - Projected uncertainty on $\mathcal{B}(B^+ \to \tau^+\nu_\tau)$ from Belle II physics book
  - This mode contributes a **yield of 5% relative to signal** after all cuts, so the uncertainty on this does not impact us so much

Toy fit to measure signal yield

- Toy datasets generated from a summed histogram PDF of signal, $B^+ \to \tau^+\nu_\tau$, and total background, with yields as mentioned
  - Each bin is Poisson varied independently to create toys

- Signal and background yields vary freely in the toy fit - signal is measured with **2.4% relative precision**
  - Fit with $10\times$ higher background level measures 2.9% uncertainty


Example toy fit


2000 toy results

Converting from a signal yield to $\mathcal{B}(B_c^+ \to \tau^+\nu_\tau)$

$$\mathcal{B}(B_c^+ \to \tau^+\nu_\tau) = \frac{N(B_c^+ \to \tau^+\nu_\tau)}{N(B_c^+ \to J/\psi\mu^+\nu_\mu)} \times \frac{\epsilon(B_c^+ \to J/\psi\mu^+\nu_\mu)}{\epsilon(B_c^+ \to \tau^+\nu_\tau)}$$

$$\times \frac{\mathcal{B}(J/\psi \to \mu\mu)}{\mathcal{B}(\tau^+ \to 3\pi\nu_\tau)} \times \mathcal{B}(B_c^+ \to J/\psi\mu^+\nu_\mu)$$

- Signal yield measured in toy fit is based on an assumed $\mathcal{B}(B_c^+ \to \tau^+\nu_\tau) = 0.0236$
- Can convert our toy fit yield back to $\mathcal{B}(B_c^+ \to \tau^+\nu_\tau)$ via correction factors, and assess the precision on branching ratio
- Use some external branching fractions as well as a theory prediction for $\mathcal{B}(B_c^+ \to J/\psi\mu^+\nu_\mu) = 0.013 \pm 0.001$

$\mathcal{B}(B_c^+ \to \tau^+\nu_\tau)$ calculation inputs

- Expected signal yield from cut optimisation procedure
- Yield statistical uncertainty from toy fits, **scaled up by $\sqrt{2}$ to account for potential systematics** (template statistics, background modelling)
- Normalisation yield calculated in similar manner to signal, assuming fairly high efficiency for three-muon mode
  - Main loss due to $m(J/\psi\mu) > 5.3$ GeV cut needed to reject the $B^{+/0}$ background (see arXiv:1407.2126)

Term	Value	Explanation
$\epsilon(B_c^+ \to \tau^+\nu_\tau)$	$(3.12 \pm 0.03) \times 10^{-3}$	From our MC (assume 1% precision)
$\epsilon(B_c^+ \to J/\psi\mu^+\nu_\mu)$	0.150 ± 0.002	Assume with 1% precision
$\mathcal{B}(J/\psi \to \mu\mu)$	$(5.96 \pm 0.03) \times 10^{-2}$	From PDG
$\mathcal{B}(\tau^+ \to 3\pi\nu_\tau)$	$(9.31 \pm 0.05) \times 10^{-2}$	From PDG
$N(B_c^+ \to \tau^+\nu_\tau)$	4147 ± 142	Uses $N_{b\bar{b}}$, f_c, $\mathcal{B}$'s, and ϵ
$N(B_c^+ \to J/\psi\mu^+\nu_\mu)$	70301 ± 265	Uses $N_{b\bar{b}}$, f_c, $\mathcal{B}$'s, and ϵ
$\mathcal{B}(B_c^+ \to J/\psi\mu^+\nu_\mu)$	0.013 ± 0.001	Calculation from Olcyr Sumensari

$\mathcal{B}(B_c^+ \to \tau^+\nu_\tau)$ anticipated precision

$$\mathcal{B}(B_c^+ \to \tau^+\nu_\tau) = (2.360 \pm 0.203) \times 10^{-2}$$

- Recover the input branching ratio central value as expected
- Relative uncertainty is 8.6%, including uncertainties on all of the input terms
- Uncertainty dominated by the current theory uncertainty on $\mathcal{B}(B_c^+ \to J/\psi\mu^+\nu_\mu)$, which is 7.7%
  - Using this and this lattice paper for current estimate along with exclusive $|V_{cb}|$
  - **Uncertainty is reducible if decay form factors are measured** e.g. in an angular analysis (LHCb & FCC-ee can both do this)
- Signal yield uncertainty from fit is 2.4%, which increases to 3.4% assuming the same level of systematics
  - Excellent precision using only the $\tau^+ \to 3\pi\nu_\tau$ mode!

Summary

- Full framework in place using FCC-ee software tools for generation, processing, and analysis of $B_c^+ \to \tau^+ \nu_\tau$ mode and backgrounds

- Selections exploiting missing energy signature of signal and 3π properties can achieve a high purity final dataset
    - Analysis should be feasible even if background level is higher than our estimate (e.g. fit stable with 10 times larger background)

- Signal yield measurement ~ 3% precision, but requires accurate $\mathcal{B}(B_c^+ \to J/\psi \mu^+ \nu_\mu)$ to convert to a precise $\mathcal{B}(B_c^+ \to \tau^+ \nu_\tau)$

- Paper draft currently in preparation detailing the workflow, achievable precision, and relevant phenomenology