First studies of $B_c^+ \to \tau^+ \nu_\tau$ with EDM4hep and FCCAnalyses

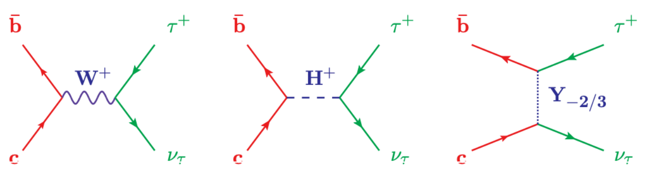
Donal Hill, Clement Helsens, Yasmine Amhis

14/12/20

FCC-ee Physics Performance meeting

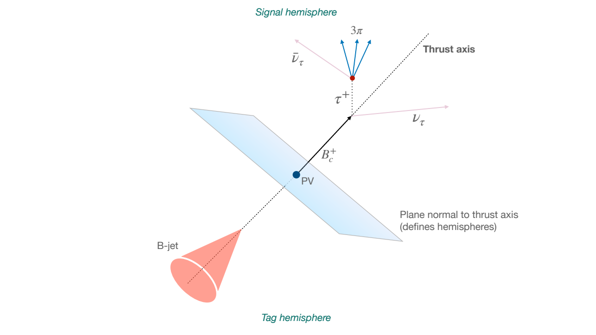
Introduction

- $B_c^+ \to \tau^+ \nu_\tau$ is a unique flavour opportunity at FCC-ee
  - Not possible at LHCb due to missing energy - lack of constraints and reconstructed information
  - No B_c^+ mesons produced at Belle II
- Can be used to measure CKM matrix element $|V_{cb}|$, but is also highly sensitive to New Physics amplitudes at tree level (e.g. charged Higgs, leptoquark)
- Provides strong tests of NP that are complementary to $b \to c\ell\nu$ deviations observed in LHCb and B-factory measurements



Event topology for $B_c^+ \tau^+ \nu_\tau$

- Can reconstruct the thrust axis for $Z^0 \to q\bar{q}$ and use this to define which hemisphere the particles fall in
  - Due to high missing energy in signal decay, the two hemispheres have rather different energy distributions
  - Use thrust calculated in FCCAnalyses to study hemisphere energy distributions in $B_c^+ \to \tau^+ \nu_\tau$ exclusive signal MC and inclusive $Z \to q\bar{q}, c\bar{c}, b\bar{b}$



Using $\tau^+ \to \pi^+\pi^+\pi^-\bar{\nu}_\tau$ and other hadronic modes

- Existing feasibility study for CEPC used electron and muon τ decays
- Idea to use hadronic modes for this study - currently using $\tau^+ \to \pi^+\pi^+\pi^-\bar{\nu}_\tau$ signal MC
  - Multi-track modes like $3\pi\bar{\nu}_\tau$ and $3\pi\pi^0\bar{\nu}_\tau$ provide τ decay vertex - combined measure of B_c^+ + τ^+ flight (lifetime)
  - Modes like $\pi\pi^0\bar{\nu}_\tau$ and $3\pi\pi^0\bar{\nu}_\tau$ can also be used to understand and benefit from calorimeter reconstruction
  - High combined branching fraction across π, $\pi\pi^0$, 3π, $3\pi\pi^0$ modes of 51% ($e + \mu$ = 35%)
- Work shown today does not involve any explicit reconstruction of the decay chain - looking only at **event-level** information in signal and background

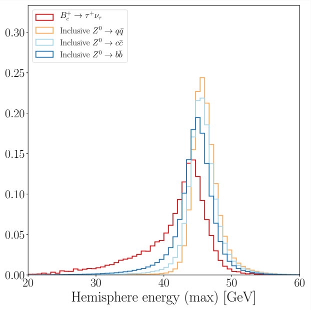
MC samples used

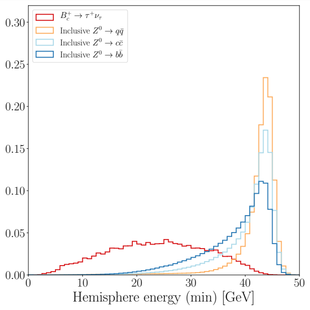
- ROOT files at:
  /eos/experiment/fcc/ee/tmp/flatntuples/Z_Zbb_Bc2TauNu/
- Samples produced with Pythia, EvtGen and Delphes in EDM4hep, with post-processing in FCCAnalyses to calculate thrust and hemisphere energy information
- 12,000 $B_c^+ \to \tau^+ \nu_\tau$ after filtering (filter keeps events with a B_c^+ produced in hadronisation)
  - $\tau^+ \to 3\pi\bar{\nu}_\tau$ generated via TAUHADNU model
- 1 million inclusive $Z^0 \to q\bar{q}, c\bar{c}, b\bar{b}$ each
  - MVA studies (see later) combine these into a single 1 million event sample using Z^0 branching fractions

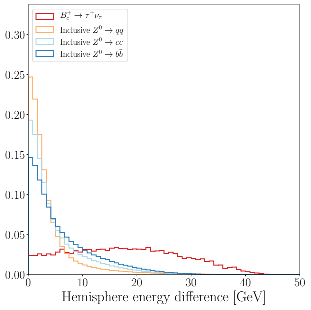
Event-level variables studied (using reco. particles)

Variable	Description
E^{diff} [GeV]	Max. – Min. hemisphere energy
E^{max} [GeV]	Max. hemisphere energy (higher of the two)
E^{min} [GeV]	Min. hemisphere energy (lower of the two)
$E_c^{\text{max(min)}}$ [GeV]	Charged energy in max. (min.) hemisphere
$E_n^{\text{max(min)}}$ [GeV]	Neutral energy in max. (min.) hemisphere
$M_c^{\text{max(min)}}$	Charged multiplicity in max. (min.) hemisphere
$M_n^{\text{max(min)}}$	Neutral multiplicity in max. (min.) hemisphere

$E^{\text{max/min/diff}}$ - clear separation for $B_c^+ \to \tau^+ \nu_\tau$

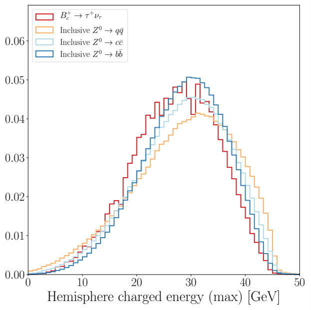


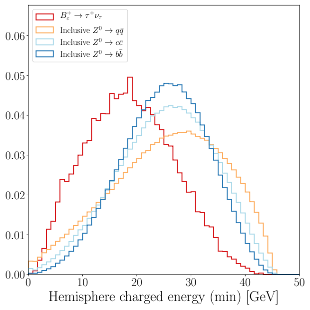




$E_c^{\text{max/min}}$

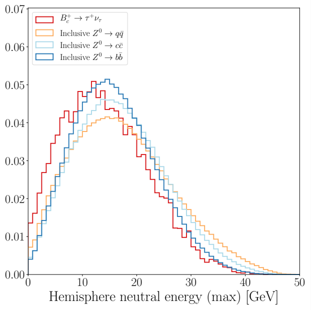
- More separation power in the minimum energy hemisphere
- This side is predominantly signal due to missing neutrinos
- In inclusive background, hemispheres have similar energy on average

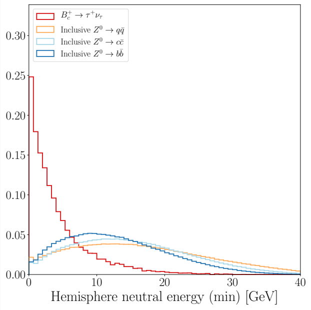




$E_n^{\text{max/min}}$ - more power in min. E hemisphere (mostly signal side)

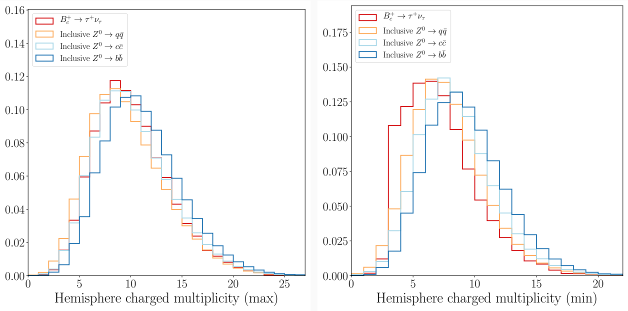
- More separation power in the minimum energy hemisphere
- This side is predominantly signal due to missing neutrinos
- In inclusive background, hemispheres have similar energy on average





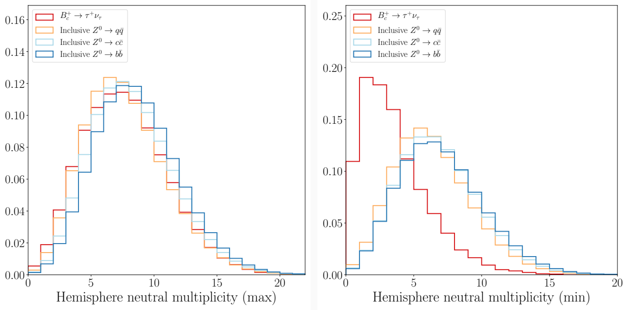
$M_c^{\mathrm{max/min}}$

- Non-signal sides are similar in terms of charged particle content
- Signal side slightly lower in multiplicity, since we only have three charged tracks in signal decay



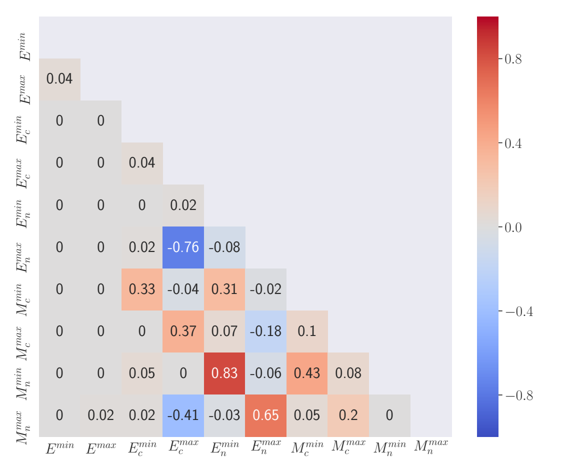
$M_n^{\text{max/min}}$

- Non-signal sides are similar in terms of neutral particle content
  - Neutral particles are charge-zero objects reconstructed in PFlow
- Signal side quite a bit more quiet



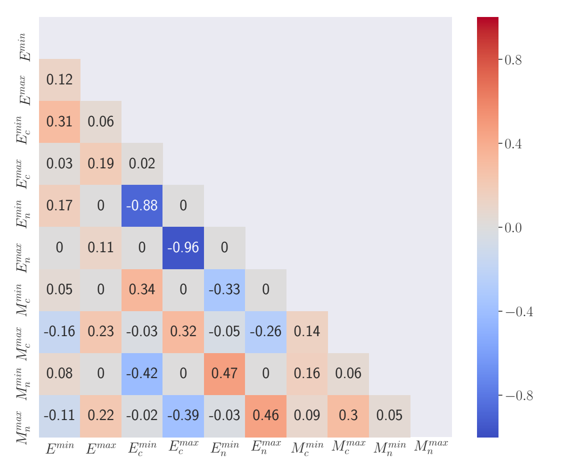
Variable correlations - $B_c^+ \to \tau^+ \nu_\tau$

- Some strong correlations but also quite a lot of mutual information



Variable correlations - inclusive $Z^0 \to q\bar{q}$ ($q = u, d, s$)

- Differences in correlation structure compared to signal (similar in $c\bar{c}$ and $b\bar{b}$, see backup slides)

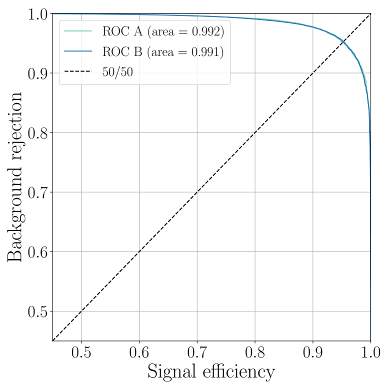


	E^{min}	E^{max}	E_c^{min}	E_c^{max}	E_n^{min}	E_n^{max}	M_c^{min}	M_c^{max}	M_n^{min}	M_n^{max}
E^{min}										
E^{max}	0.12									
E_c^{min}	0.31	0.06								
E_c^{max}	0.03	0.19	0.02							
E_n^{min}	0.17	0	-0.88	0						
E_n^{max}	0	0.11	0	-0.96	0					
M_c^{min}	0.05	0	0.34	0	-0.33	0				
M_c^{max}	-0.16	0.23	-0.03	0.32	-0.05	-0.26	0.14			
M_n^{min}	0.08	0	-0.42	0	0.47	0	0.16	0.06		
M_n^{max}	-0.11	0.22	-0.02	-0.39	-0.03	0.46	0.09	0.3	0.05	

Multivariate analysis

- Use hemisphere energy information to distinguish $B_c^+ \to \tau^+ \nu_\tau$ from $Z^0 \to q\bar{q}, c\bar{c}, b\bar{b}$
- Create combined background sample of 1 million events using Z^0 PDG branching fractions
- Use *GradientBoostingClassifier* in *scikit-learn* with:
  - *n_estimators* = 200
  - *learning_rate* = 0.4
  - All other hyper-parameters set to defaults
- Split samples into A and B, and train two BDTs (A and B)
  - Apply BDT A (B) to sample B (A) to get predictions for full sample

ROC AUC and feature ranking

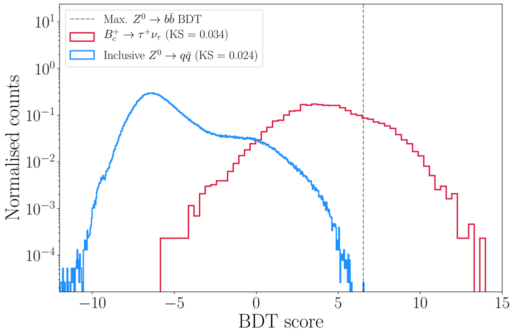


Variable	Feature importance
E^{min}	0.823
E_n^{min}	0.0998
E^{max}	0.0473
E_c^{min}	0.0192
E_c^{max}	0.00419
E_n^{max}	0.00277
M_c^{min}	0.00155
M_n^{min}	0.000881
M_c^{max}	0.000608
M_n^{max}	0.000447

- Lower energy hemisphere dominates, but also contributions from charged and neutral sub-totals and the maximum hemisphere energy

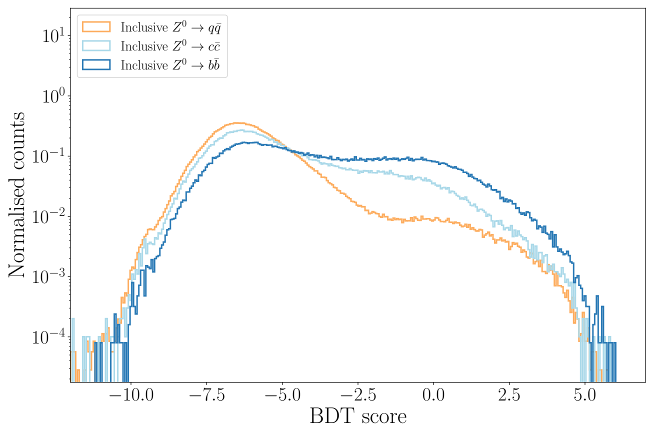
BDT score distributions

- Reject all 10^6 background events with BDT > 6.5 cut
- This cut is 19% efficient on $B_c^+ \to \tau^+ \nu_\tau$ signal
- Need larger background samples to test rejection to higher level
  - More signal will help improve the KS-test scores, which aren't bad but can be improved



BDT score for each background type

- BDT is trained on a combined sample of $Z^0 \to q\bar{q}, c\bar{c}, b\bar{b}$
- Sub-distributions show that light-quark background is best separated from signal, and $b\bar{b}$ largest in upper tail



Signal purity estimate

- Assume 3×10^{12} Z^0 in FCC-ee operation
- With $\mathcal{B}(Z^0 \to \text{hadrons}) = 69.9\%$, leads to 4.2×10^{12} inclusive background decays
- $N(B_c^+ \to \tau^+ \nu_\tau) = 868,000$ using the following factors

Factor	Value
$N(Z^0)$	3×10^{12}
$\mathcal{B}(Z^0 \to b\bar{b})$	0.1512
B_c^+ production rate	7.9×10^{-5} [CEPC paper]
$\mathcal{B}(B_c^+ \to \tau^+ \nu_\tau)$	0.0236 [CEPC paper]
$\mathcal{B}(\tau^+ \to \{\pi, \pi\pi^0, 3\pi, 3\pi\pi^0\}\bar{\nu}_\tau)$	0.513

- Signal purity before any selection is thus 2×10^{-7}

Signal purity estimate

- Let's target 1000 signal events with 1000 background (50% purity) for a ~ 3% precision $\mathcal{B}$ measurement

- Total background rejection required: 4.2×10^9

- Total signal efficiency required: 1.2×10^{-3}

- BDT achieves 10^6 rejection for 19% signal efficiency:
  - Brings us from 2×10^{-7} to 0.04 purity (**another factor 10 needed**)
  - 4×10^3 further background rejection required
  - Can tolerate an additional signal efficiency of 0.6%

- Selections based on specific signal properties (3π vertex quality, resonant structure, PV separation) must be studied to understand additional background rejection capabilities

Summary and next steps

- Event-level hemisphere energy information provides good discrimination for missing energy mode $B_c^+ \to \tau^+ \nu_\tau$

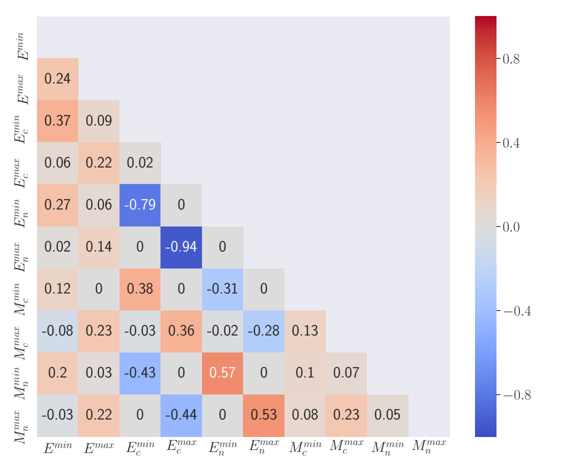
- Still factors of rejection required for a high-purity scenario, but signal-specific properties (e.g. 3π vertex and flight) and b-tagging for non-signal side will help

- Larger background and signal samples to be generated, allowing rejection to be better understood

- Most dangerous physics background is $B^+ \to \tau^+ \nu_\tau$ - will study this vs. signal in dedicated manner

Backups

Variable correlations - inclusive $Z^0 \to c\bar{c}$



Variable correlations - inclusive $Z^0 \to b\bar{b}$