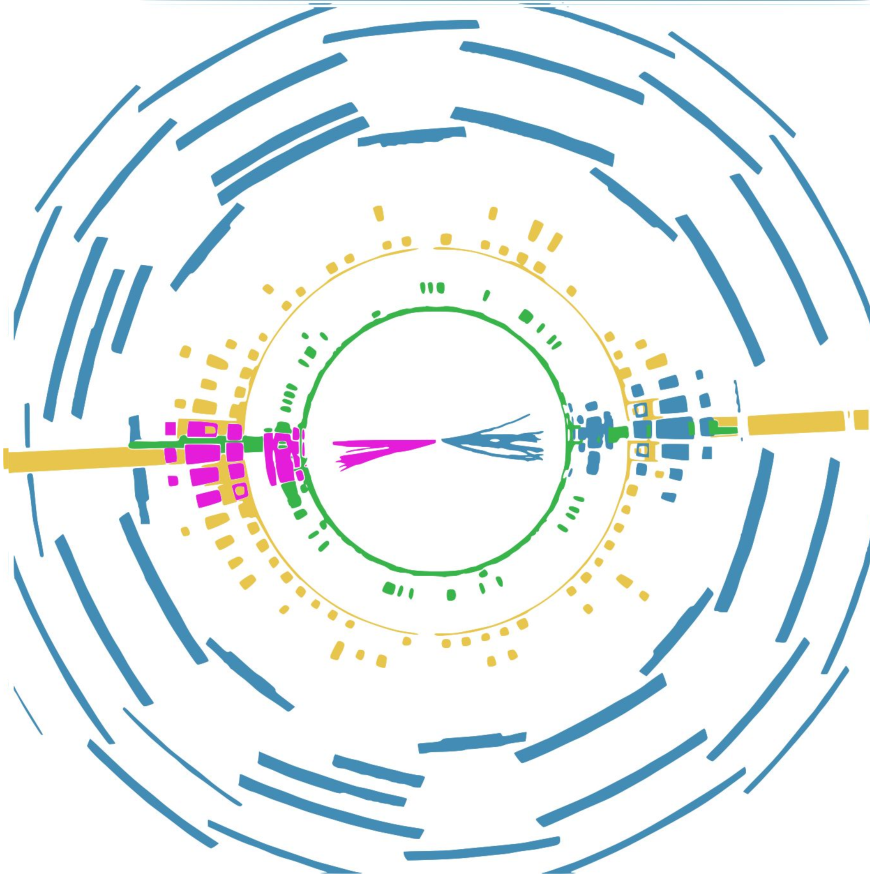
ZH->jjjj

Iza Veliscek

Contributions from: Haider Abidi, Viviana Cavaliere, Jan Eysermans, George Iakovidis, Loukas Gouskos, Andrea Sciandra, Michele Selvaggi

25th March 2024, 2nd US FCC Workshop

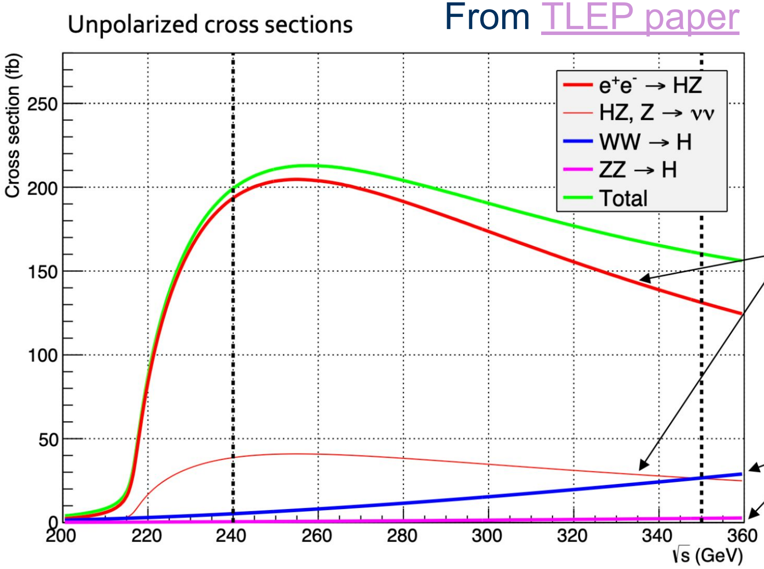


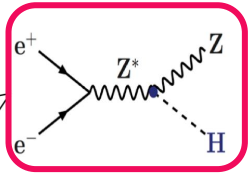
Brookhaven National Laboratory

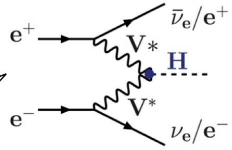
FUTURE CIRCULAR COLLIDER

Introduction

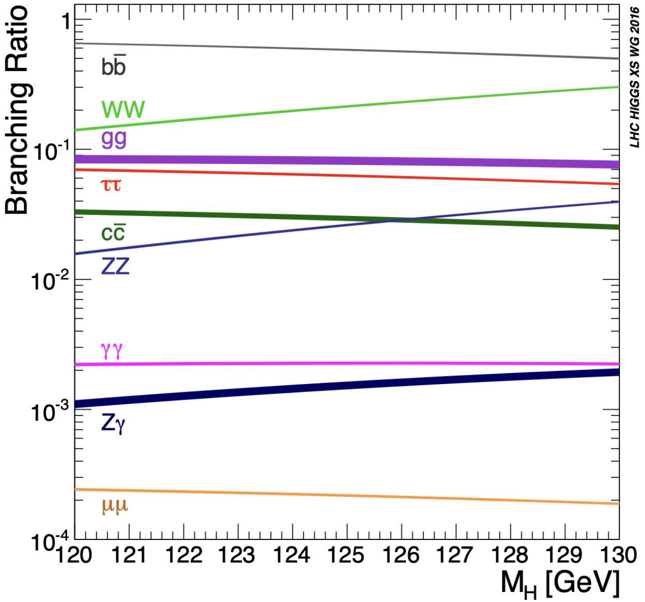
From TLEP paper







From Handbook of LHC Higgs cross sections



- ZH leading Higgs production mode
  - + All hadronic decay has the largest branching fraction
  - - Jet combinatorics, flavour identification
- Abundance of Higgs produced @ √s = 240 GeV
  - Focus on IDEA Detector

Samples Considered

- **IDEA Detector**
  - Delphes fast sim
- **Winter2023 Samples**
  - /eos/experiment/fcc/ee/jet_flavour_tagging/winter2023/wc_pt_7classes_12_04_2023
- **Jet Clustering**
  - N = 4 Durham k_T exclusive algorithm
- **ParticleNet jet tagger**
  - fccee_flavtagging_edm4hep_wc
- Build on ZH(full hadronic) analysis presented in Annecy by George [slides]

Background:

- **WW**
- **ZZ**
- **Zqq**
- **Z(bb/cc/ss/qq/)H(tautau)**
- **Z(bb/cc/ss/qq/)H(WW)**
- **Z(bb/cc/ss/qq/)H(ZZ)**
- **Z(bb/cc/ss/qq/)H(Z/ɣ*)**
- **nunuH(jj)**
- **Missing Z(bb/cc/ss/qq/)H(qq)**

Signals:

- **Z(bb/cc/ss/qq/)H(bb)**
- **Z(bb/cc/ss/qq/)H(cc)**
- **Z(bb/cc/ss/qq/)H(ss)**
- **Z(bb/cc/ss/qq/)H(gg)**

Analysis setup

Preselection

- Exactly 4 jet!

Lepton cuts

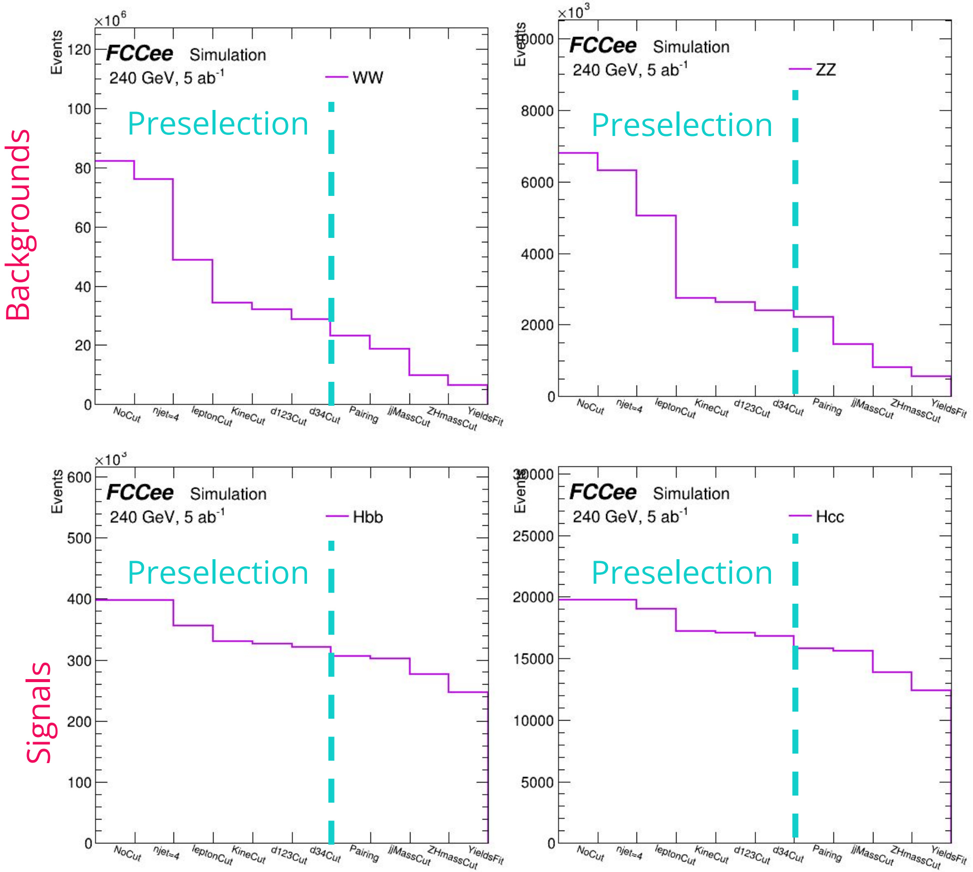
- <= 2 muons and electrons
- Leading muon and electron pT < 20 GeV

Visible Energy

- Visible m > 150 GeV
- Visible E > 150 GeV
- 0.15 <Visible θ < 3.0

d_{ij} Cuts

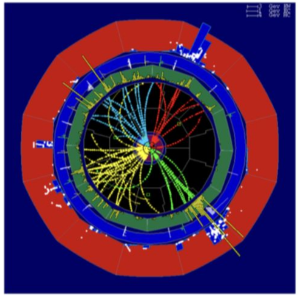
- $15000 < d_{12} < 58000$
- $400 < d_{23} < 18000$
- $100 < d_{34} < 6000$



Jet energy correction

Precision with e^+e^- colliders (4)

- Why are e^+e^- colliders the tool of choice for precision anyway ? (cont'd)
  - Electrons are leptons, i.e., elementary particles: no underlying event
    - Corollary: Final state has known energy and momentum: ($\sqrt{s}$, 0, 0, 0)
  - Example: an $e^+e^- \to W^+W^- \to q\bar{q}q\bar{q}$ candidate
    - Four jets in the event and nothing else
    - Total energy and momentum are conserved
      - ➨ $E_1 + E_2 + E_3 + E_4 = \sqrt{s}$
      - ➨ $P_1^{x,y,z} + p_2^{x,y,z} + p_3^{x,y,z} + p_4^{x,y,z} = 0$
    - Jet directions ($\beta_i = p_i/E_i$) are very well measured



$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ \beta_1^x & \beta_2^x & \beta_3^x & \beta_4^x \\ \beta_1^y & \beta_2^y & \beta_3^y & \beta_4^y \\ \beta_1^z & \beta_2^z & \beta_3^z & \beta_4^z \end{bmatrix} \begin{bmatrix} E_1 \\ E_2 \\ E_3 \\ E_4 \end{bmatrix} = \begin{bmatrix} \sqrt{s} \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

    - Jet energies (or di-jet masses: m_W) determined analytically by inverting the matrix
      - ➨ No systematic uncertainty related to jet energy calibration

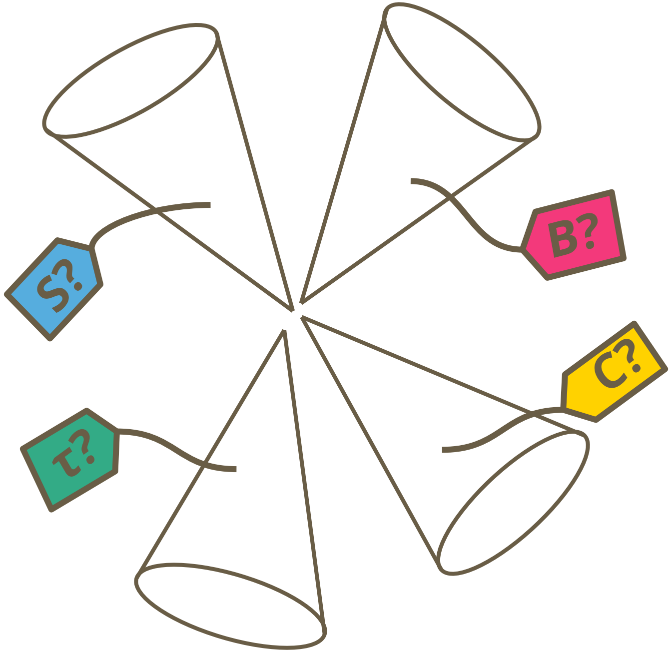
A lot of Z are available anyway to calibrate and align everything

Patrick Janot | Physics at Future Colliders 28-29 July 2016 | 7

- If any jet in event **E<0** OR **E>240** GeV [only a few percent of events]

TOSS EVENT

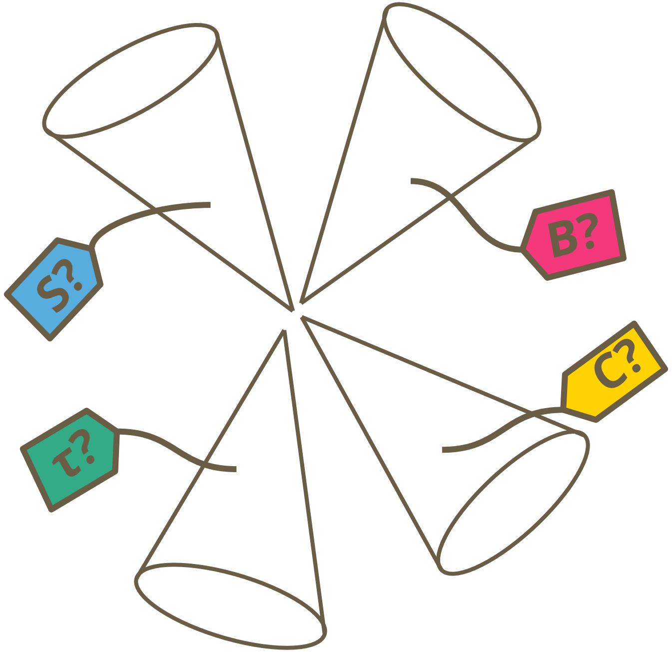
Jet “tagging”



ParticleNet jet tagger

- Scores provided for the “flavours”:
  - B, C, S, g, τ, U, D
    - q: U,D
- Scores ~ probability jet is of flavour X
- NOT traditional flavour tagging
  - Maximum flavour score ~ flavor of jet
  - Sums of same flavour scores for jet pairs ~ flavour of jet pair

Jet pairing

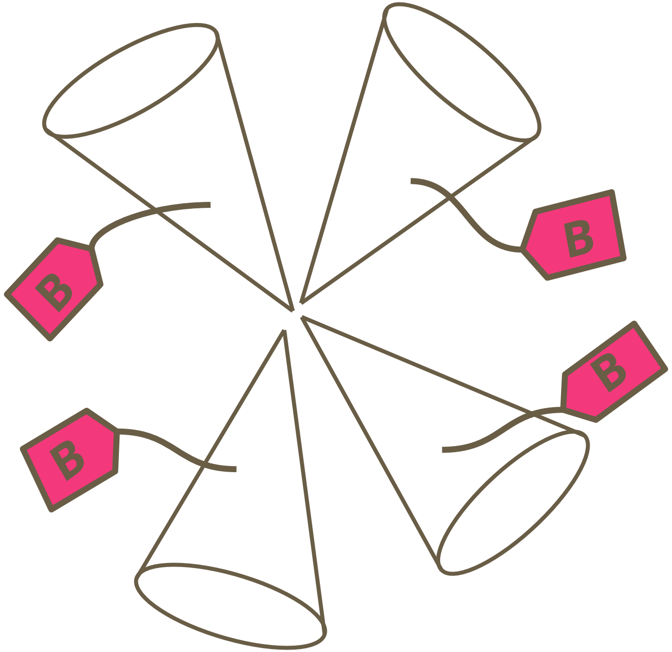


Each jet has a maximum tagger score from a different flavour

-

TOSS EVENT

Jet pairing



CASE 1: All jets have the maximum score from the same flavour

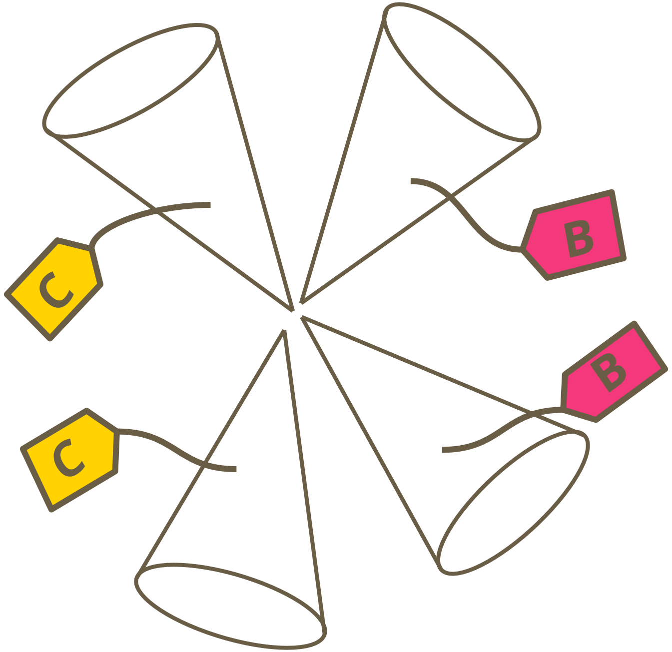
Finding the H&Z candidates

Consider all possible jet pairs

- $\chi_H = (m_{ij} - m_{H,true})^2$
- $\chi_Z = (m_{lk} - m_{Z,\ true})^2$
- $\chi_{comb} = \chi_H + \chi_Z$

The jet paring that gives the **minimum** $\boldsymbol{\chi_{comb}}$ is chosen!

Jet pairing

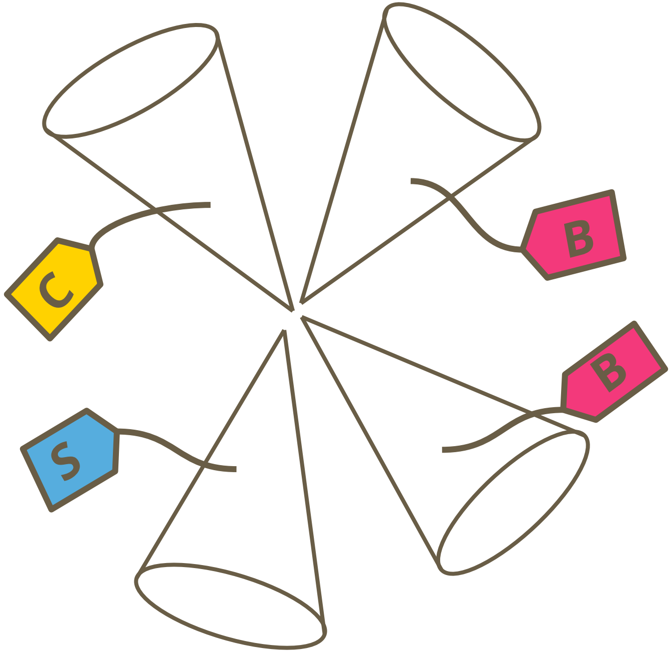


CASE 2: Two jet pairs with same maximum score from the same flavour, but different flavour of the pairs

Finding the H&Z candidates

- Jet paired, if they have the same flavour maximum score
- Z candidate: Pair with minimum $\chi_Z = (m_{lk} - m_{Z,\,true})^2$

Jet pairing



CASE 3: Two jets with maximum score from the same flavour form a pair

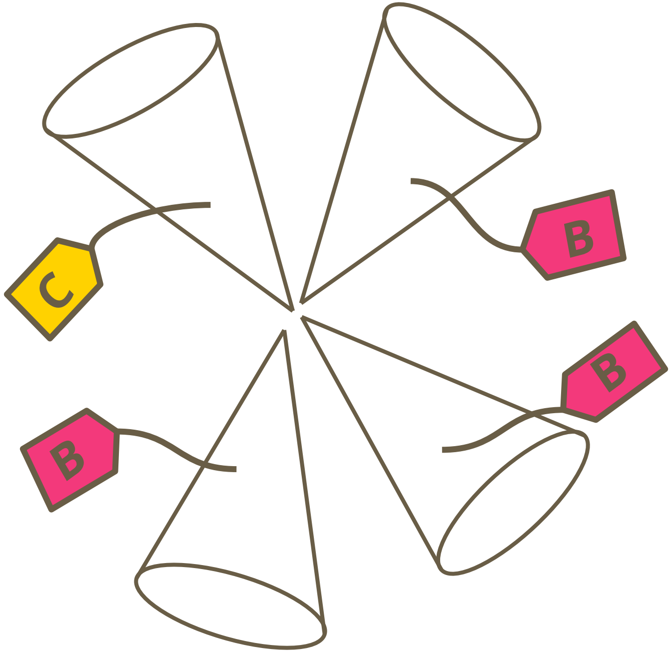
Recover second pair:

- Consider all sums of tagger scores
  - **Max**($\sum_{ij}$Bscore, $\sum_{ij}$Cscore, $\sum_{ij}$Sscore, …)
    - Determines the flavour of the pair

Finding the H&Z candidates

- Same flavour pairs (Case 1)
  - Min($\chi_{comb}=\chi_H+\chi_Z$)
- Different flavour pairs (Case 2)
  - Min($\chi_Z=(m_{lk} - m_{Z,\ true})^2$

Jet pairing



CASE 4: Three jets with maximum score from the same flavour

Recover first pair: [check code]

- Maximum tagger score sum
  - **Max**($\sum_{ij}$Bscore, $\sum_{ik}$Bscore, $\sum_{jk}$Bscore, …)
    - Determines the flavour of the 1st pair

Recover second pair:

- Consider all sums of tagger scores
  - **Max**($\sum_{ij}$Bscore, $\sum_{ij}$Cscore, $\sum_{ij}$Sscore, …)
    - Determines the flavour of the pair

Finding the H&Z candidates

- Same as for Case 3

A few more cuts

WW & ZZ rejection

$$\sqrt{(m_{z_{jj}} - m_W)^2 + (m_{H_{jj}} - m_W)^2} > 10$$

$$\sqrt{(m_{z_{jj}} - m_Z)^2 + (m_{H_{jj}} - m_Z)^2} > 10$$

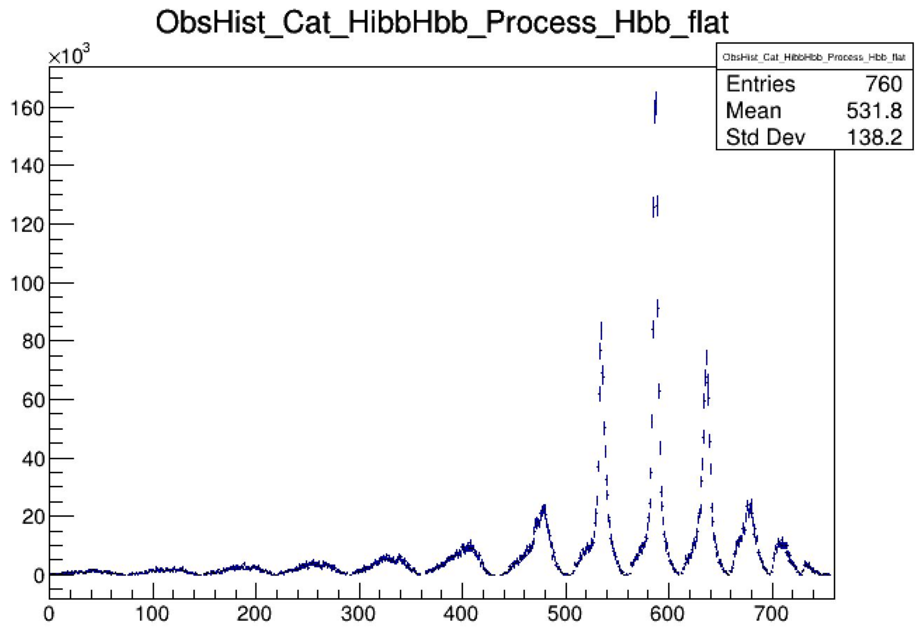
Mass window

$$50 < m_{Z_{jj}} < 125\,\text{GeV}, m_{H_{jj}} > 90\,\text{GeV}$$

Reject events identified/contain as:

- H->ττ
- H->qq, q=u,d
- Z->ττ
- Z->gg

Categorization



Unfolded 2D $m_{jj,H}$ - $m_{jj,Z}$ mass plane

Hbb signal in the high purity Z(bb)H(bb) category

- Categorize by H->j_1j_2 decay
  - Categorize by Z->j_3j_4 decay
  - Additionally by H flavour score
    - **Purity category :**
      - High (>1.8 (1.4 for Hss))
      - Mid(1.1 (0.8) < score < 1.6 (1.4) (Hss cut in ())
      - Low (<1.1 (0.8 for Hss))
- 36 Categorised in total!
- + 1 GeV binning in $m_{jj,H}$
- + 5 GeV binning in $m_{jj,Z}$

Likelihood scan

- **Asimov** (expected) **data = SM = background estimation + SM signal**
    - How compatible are different μ_{xx} to the asimov data set, i.e. how sensitive are we?
    - Compare the **test statistic (λ)** of the different μ_{xx} on this dataset.

$$\lambda(\mu_{xx}) = -2\ ln\left(\frac{L(\mu_{xx}, \hat{\hat{\theta}})}{L(\bar{\mu}_{xx}, \hat{\theta})}\right)$$

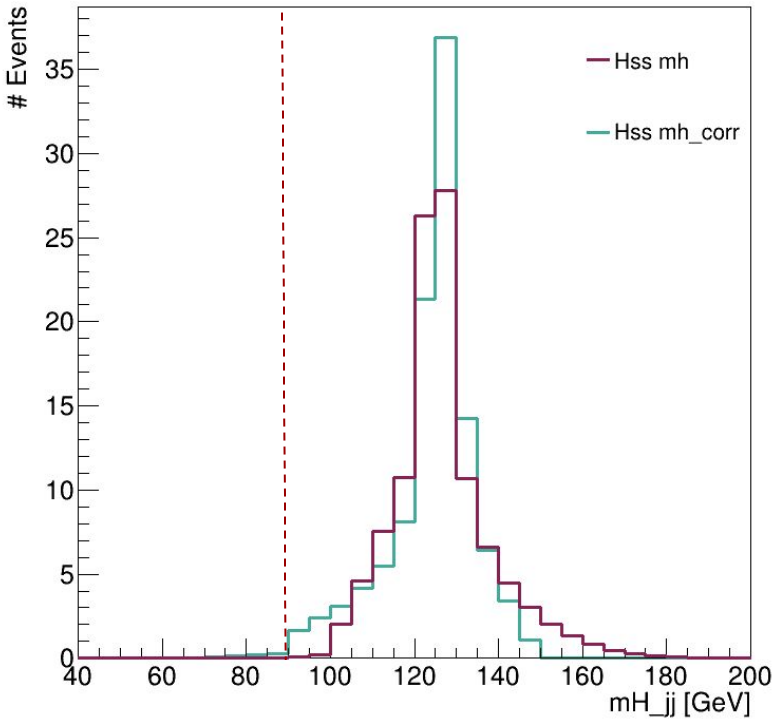
Maximize $L(\kappa_\lambda, \theta)$ by holding μ_{xx} fixed and fitting the model.

Nuisance parameters

Best-fit coupling

Find μ_{xx} that maximizes ***L*** for the data, i.e. let μ_{xx} vary in a global fit.

Yet another correction to $m_{jj,H}$



- Besides the energy correction to the jets based on COM
- After all selection:
  - $mH\_jj\_corr = mH\_jj + mZ\_jj - mZ_{truth}$
- As before fit mH_jj_corr against mZ_jj

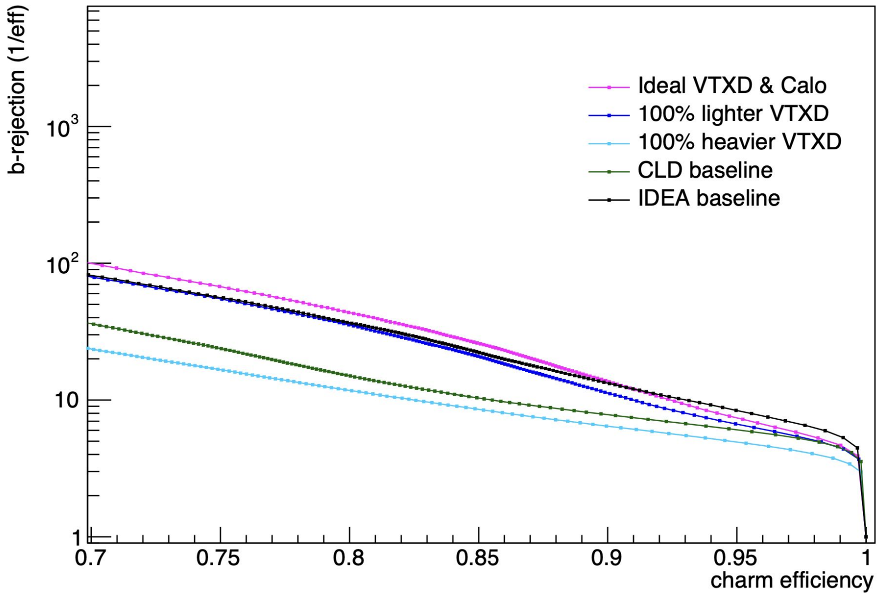
68% CL precision / variation	μ_{Hbb}	μ_{Hcc}
BASE	±0.3	±3.9
Base (fit Mh_jj_corr_Mz_jj	±0.3	±3.9

IDEA tracker variations: Approximating the impact of tagging performance on the analysis

Andrean re-trained tagger for different detectors [see Andrea' presentation]:

- **Baseline:** IDEA baseline
- **idealVXDCalo:**
    - Best material budget, hit resolution and calorimeter granularity
- **lighterVXD_100pc:**
    - ~ No material interaction (X_0>>1m)
- **heavierVXD_100pc:**
    - Super small radiation length (X_0<<1m)
- **CLD**
    - Fast sim of the CLD **o1_v01**

Plot from Andrea

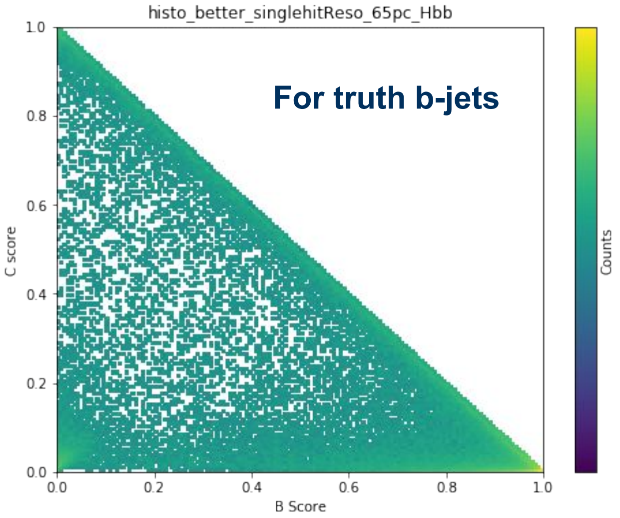


FUTURE CIRCULAR COLLIDER

Approximating the impact on tagging

Propagating the impact of retraining the tagger:

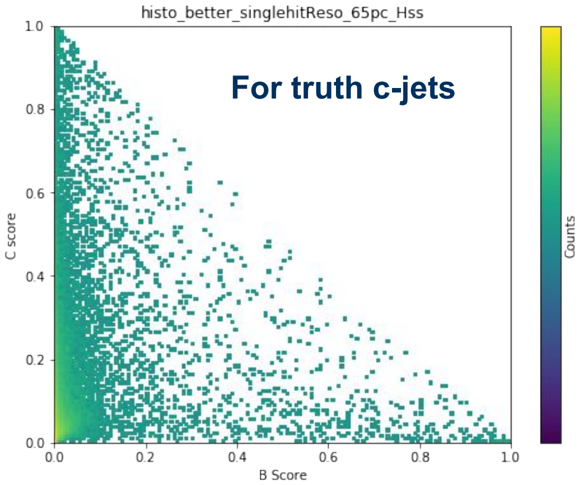
- Account only for impact on **b-,c- and s-score**
- Histo per jet flavour (4x) per detector variation **[Thanks Andrea!]**
  - Sample from histogram to update the b- c- and s-score score
    - Depends on the jet truth label!



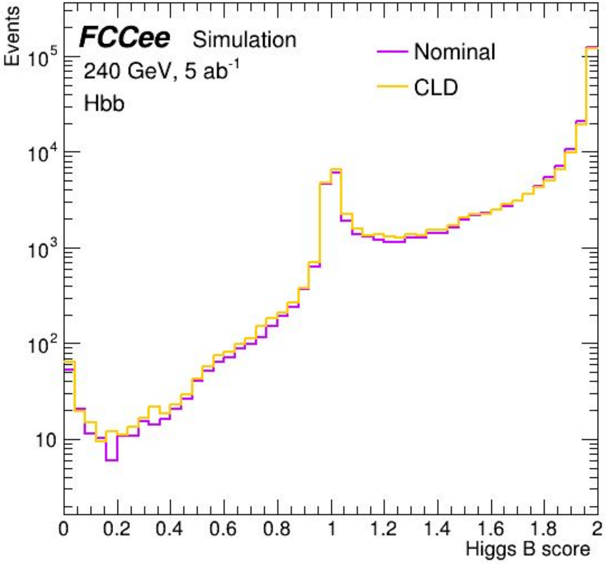
Drawbacks of the strategy

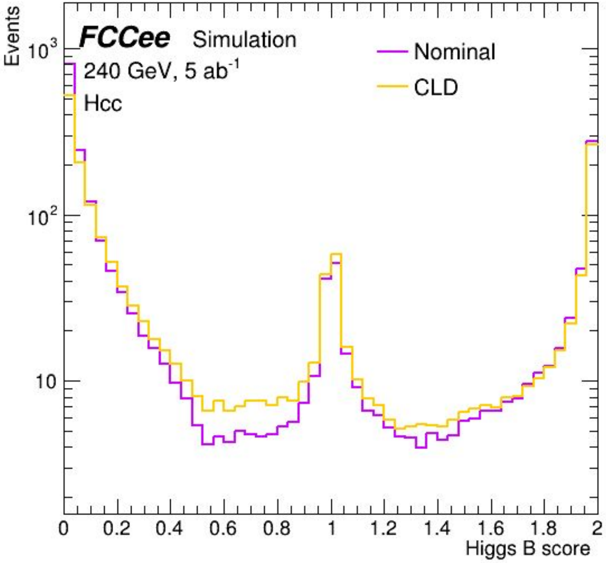
- Jet truth labelling not optimal
  - 88% accuracy in Z(qq)H(bb) samples [**Thanks Jan E.!**]
  - Does not tag gluon jets
- Ignoring some correlations
  - Correlation of the b-,c-, s- score to u/d, gluon score neglected

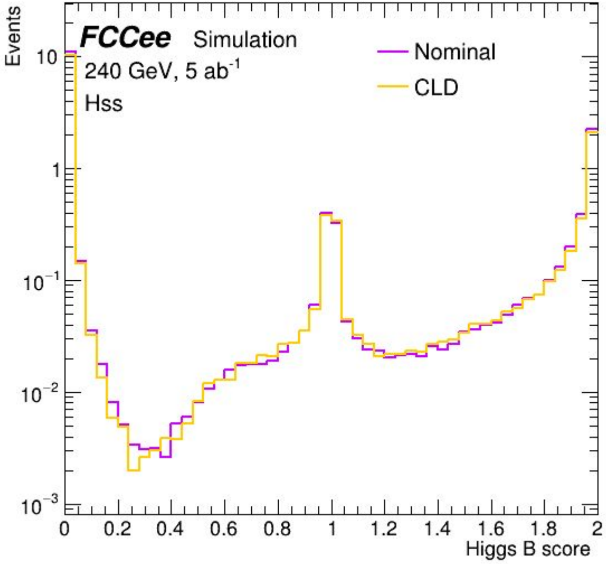
* Older tagger training, tau's not included



Impact on the analysis - Higgs B score

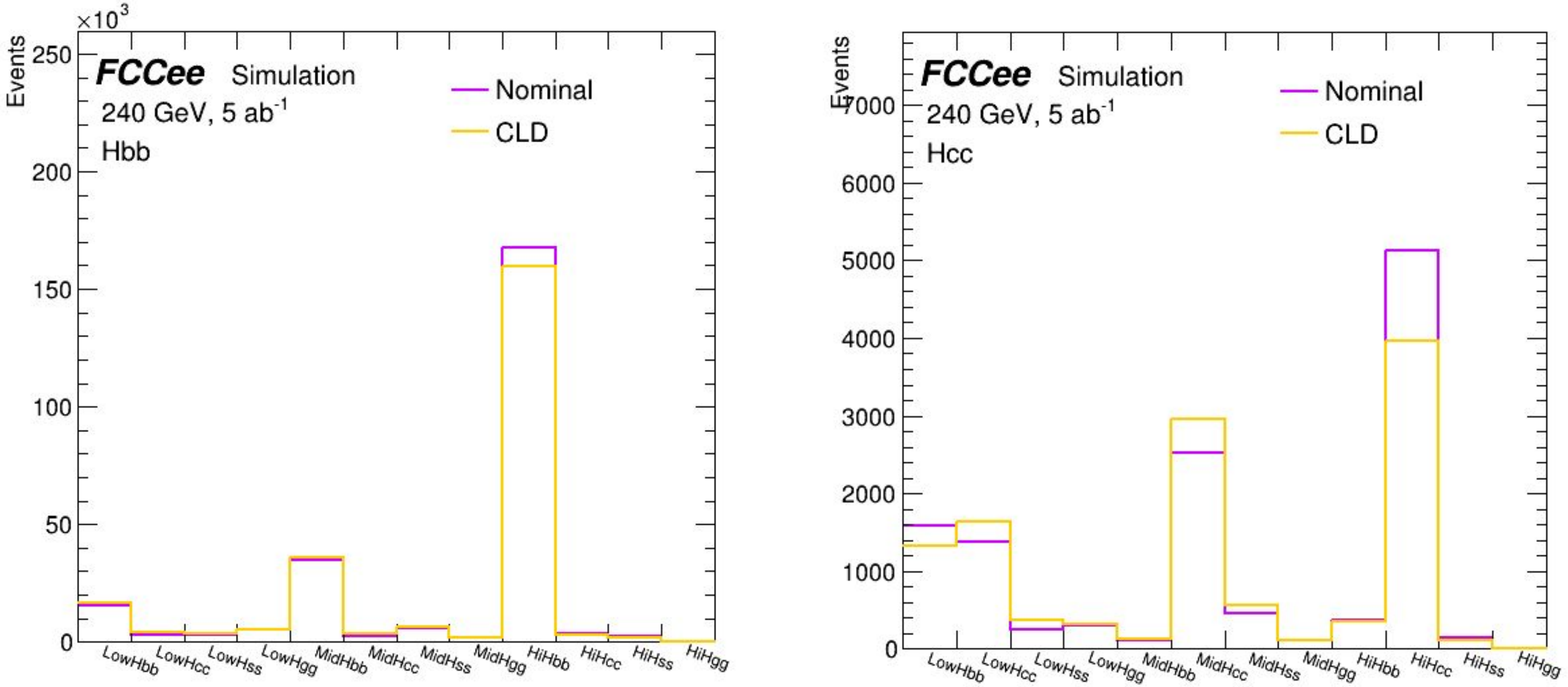




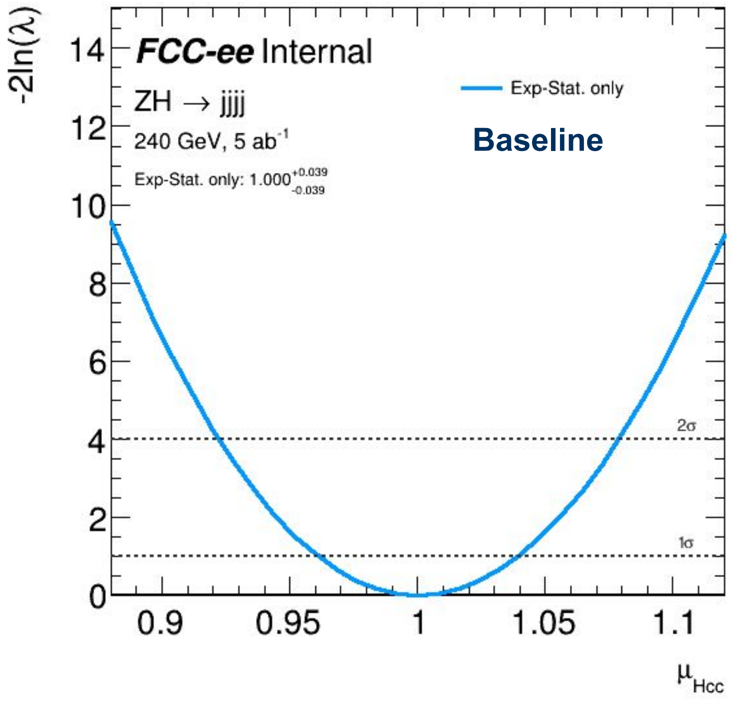


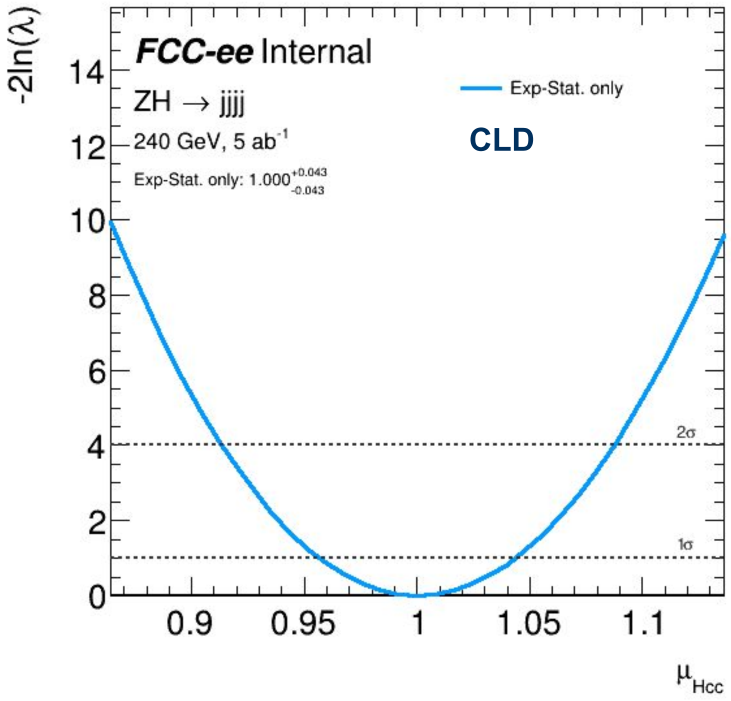
Truth H->bb jets flavour: The hit in performance of the tagger has the largest effect on the Higgs C-score. Smaller c-jet rejection leads to a larger Higgs C score.

Impact on the analysis - Migration between fit categorise



Likelihood scans - μ_{Hcc}





Largest change in expected precision on μ_{Hcc} observed when the tagger is re-trained with the CLD simulation.

Results

- **IDEA baseline very close to ideal vertex & calo detector**
- Robust analysis strategy
  - Small change in event selection
  - Main effect is migrates events between categories, dues to changes in performance
- No change in $\boldsymbol{\mu}_{Hgg}$ as expected
  - G-score not varied nor truth gluon jet score corrected
- Largest impact on $\boldsymbol{\mu}_{Hcc}$ w/ CLD trained tagger
- Caveats remainder!
  - Only approximate propagation of tagging effects
  - Ignored correlations of between b/c/s with g and light scores

variation / 68% CL precision	μ_{Hbb}	μ_{Hcc}
BASE	±0.3%	±3.9%
idealVXDCalo	±0.3%	+3.9% -3.8%
lighterVXD_100pc	±0.3%	±3.9%
heavierVXD_100pc	±0.4%	+4.6% -4.5%
CLD	±0.4%	±4.3%

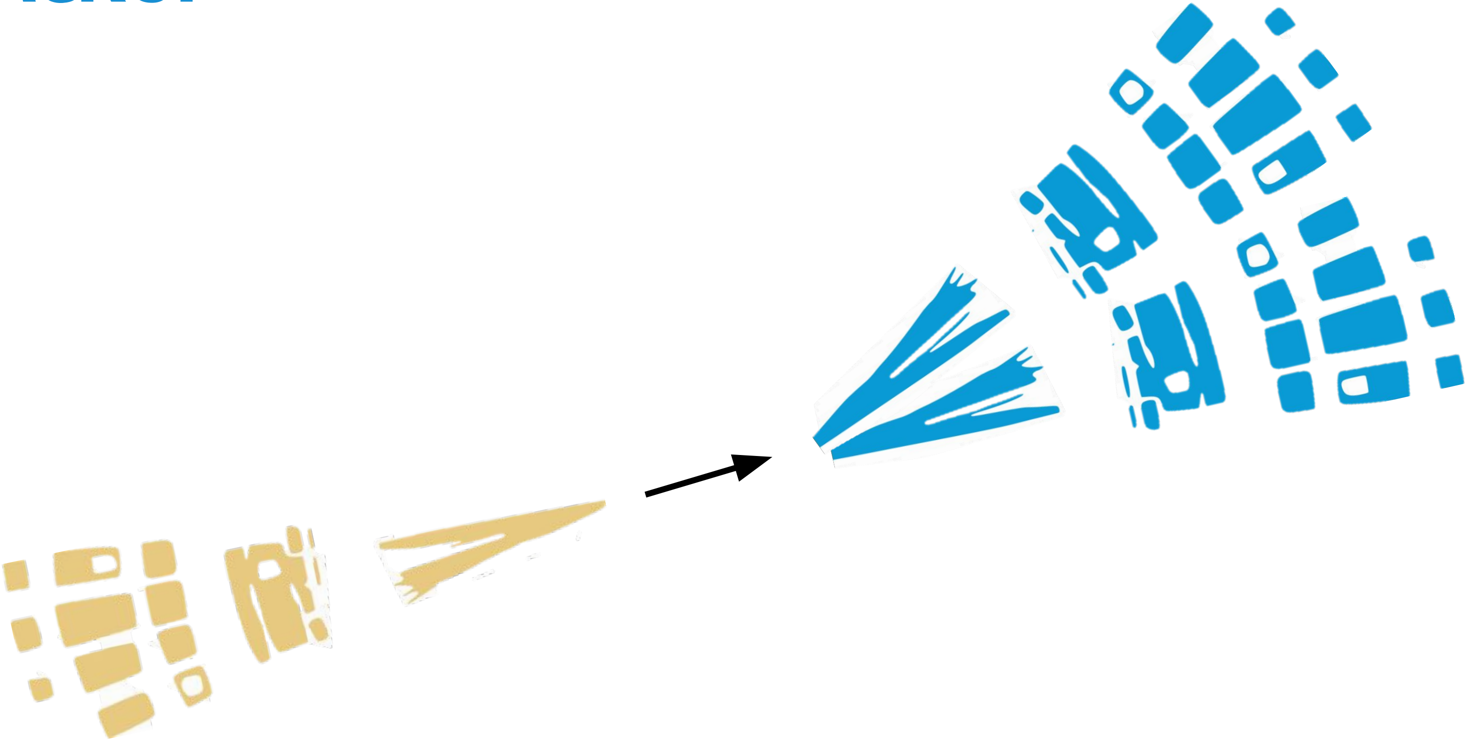
Conclusion

- Correction of the reconstructed Higgs mass does not significantly improve the expected precision on μ_{Hxx}
  - mH_jj_corr = mH_jj + mZ_jj - mZ_{truth}
- First look at the impact of flavour tagging given different detector layouts
  - Partricle net retrained for various detector layouts
  - Changes in tagger performance propagated to the ZH->jjjj analysis
    - Sever approximation taken to have a quick estimation of the impact
      - Determine how big of a change in the tagging performance is worth rerunning the whole analysis chain
  - The analysis roubouts
    - Very small impact on the expected μ_{Hxx} precision measurement

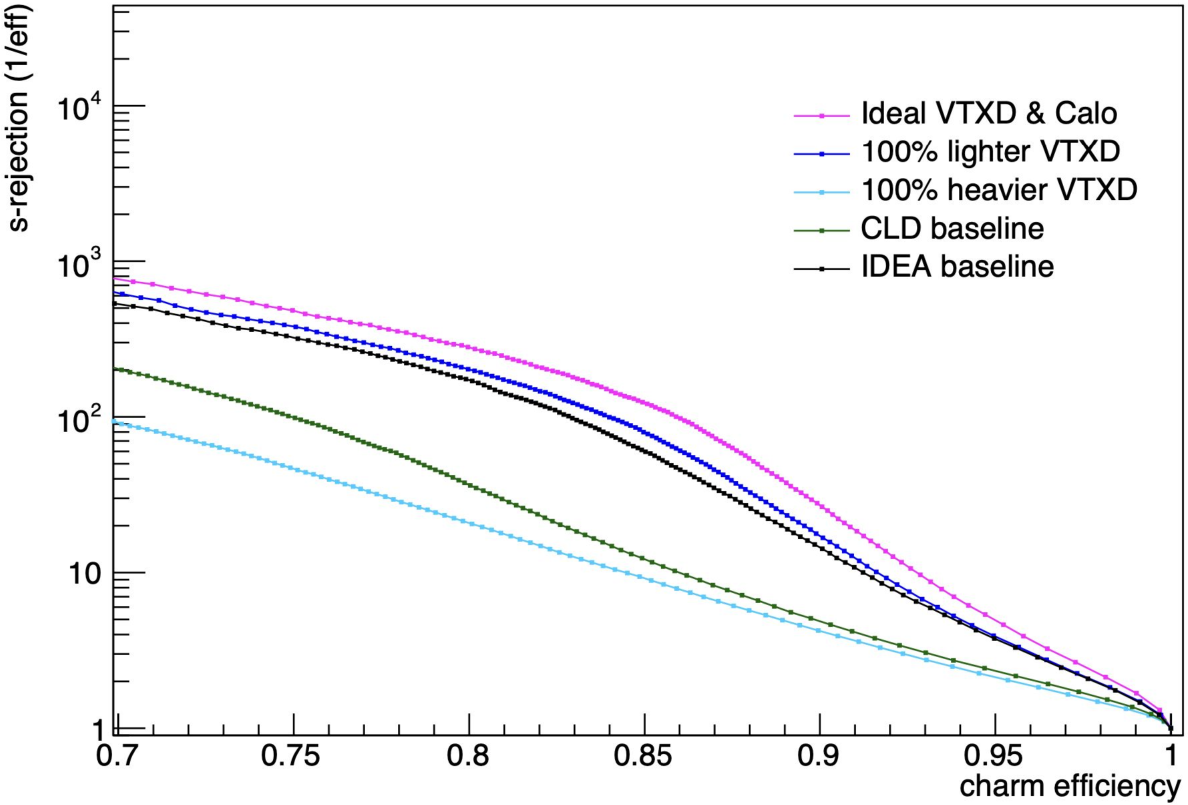
Next Steps

- Consider different jet clustering strategies
  - Will have a summer student working on it for ~ 8 weeks
- Optimize analysis strategy
  - Still some space to push for more precise $\boldsymbol{\mu}_{Hxx}$ measurements
- Future parametrize impact of various detector layout
  - Evaluate re-trained tagger without approximations
  - Jet mass resolution
- Combine with Z(vv)H(jj) and Z(jj)H(jj)

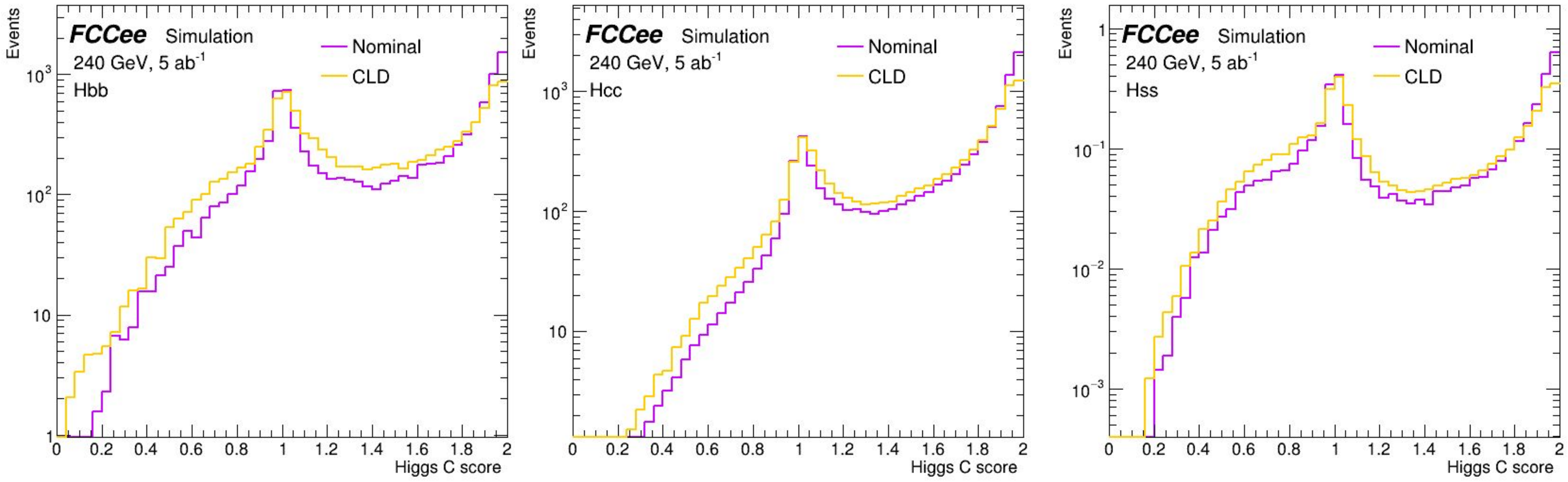
BACKUP



Tagger performance



Impact on the analysis - Higgs C score



Truth H->cc jets flavour: The better rejection of the Nominal tagger is reflected in a higher fraction of truth H->cc events, with a very high Higgs C score. [see next slide]

FUTURE CIRCULAR COLLIDER

Migration of ZZ events