Searching for Rare Flavour-Changing Higgs Boson Decays at the FCC-ee

Darren Duan, Department of Physics, Brown University, Providence, RI 02912
Ang Li, Department of Physics, Brookhaven National Laboratory, Upton, NY 11973

September 15, 2026

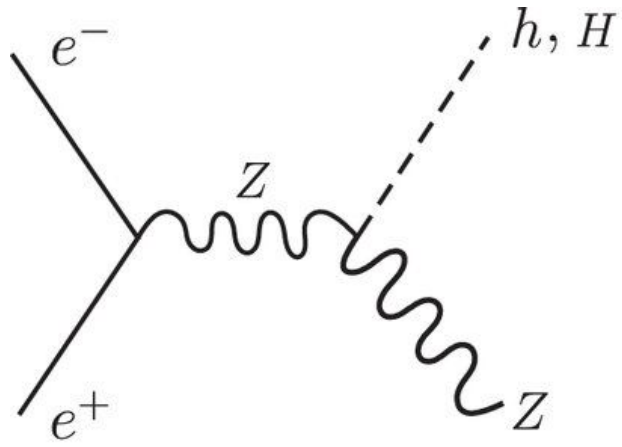
@BrookhavenLab

Motivation and Setup

In the standard model, the Higgs Boson has no tree-level flavor changing neutral currents (FCNC).

Therefore decays like H→sb has low branching ratio. But if BSM theories like Two Higgs Doublet Models (2HDM) or the MSSM are true, observing such signals would be possible!

What limit can we set on the branching ratio of these decays from the FCC?



A previous study on this was done by Michele Selvaggi, which can be found here:

https://repository.cern/records/7xjde-ebj45

Also in Alexis's talk, different Z decays have been considered, https://indico.cern.ch/event/1410813/contributions/5986018/attachments/2870427/5025303/04_06_24_FCCperformanceMeeting.pdf

We're studying the decay of $e^+e^- \to ZH \to \mu\mu q\bar{q}$, where the final state has different flavor quarks.

The channel $e^+e^- \to ZH \to \nu\nu H$ has higher statistics, but the $\mu\mu$ channel boasts more precise momentum tracking.

The MC is produced by WHIZARD3 and showered by PYTHIA6 with a CM energy of 240 GeV.

Cuts:

86 GeV < $m_{\mu\mu}$ < 96 GeV (mass of Z boson = 91.2 GeV)

20 GeV < $p_{\mu\mu}$ < 70 GeV

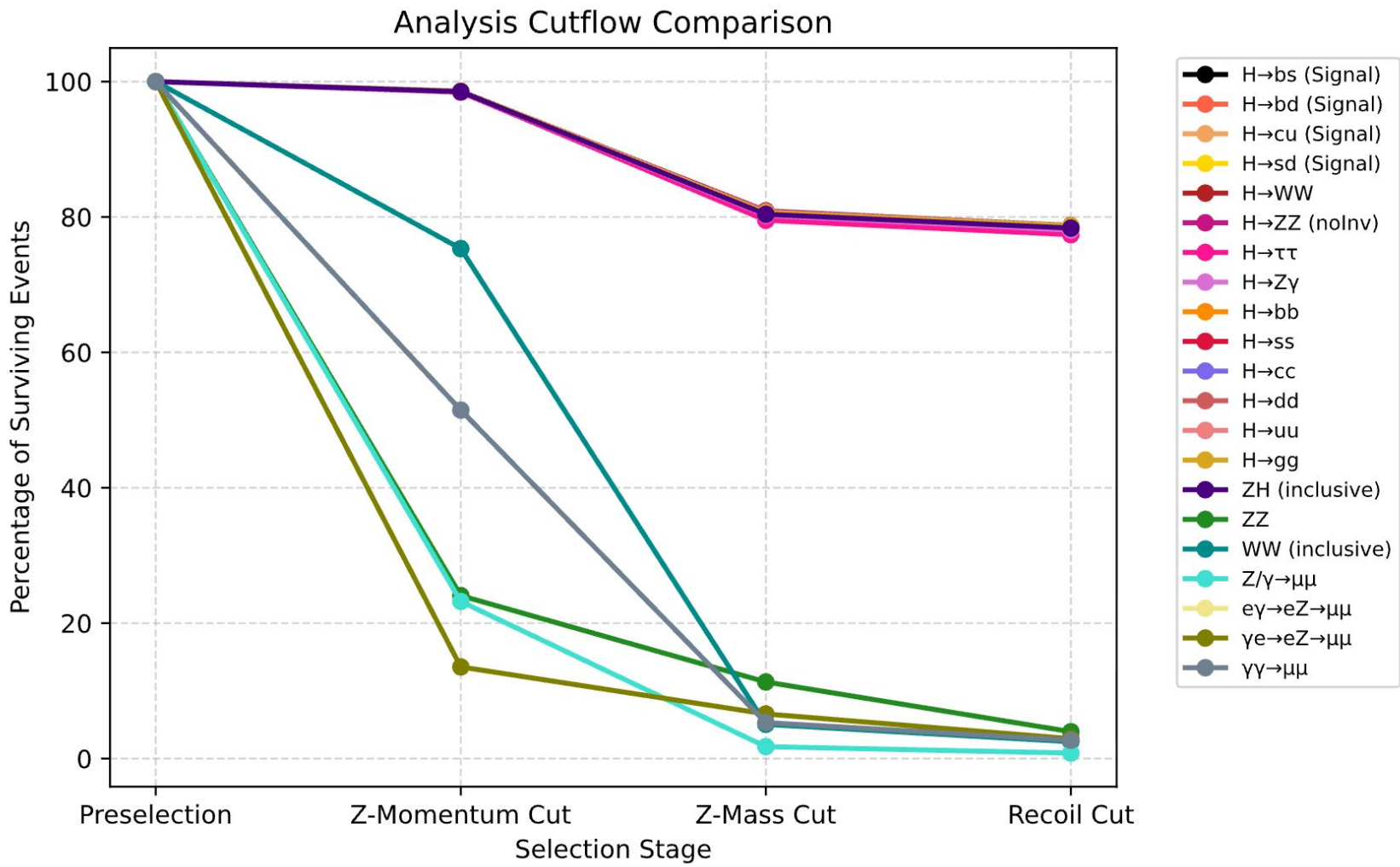
120 GeV < m_{recoil} < 140 GeV

Boosted Decision Tree Setup

Decided on a multi-classifier BDT using XGBoost to differentiate between signal and background samples.

Used 20 total samples split into 8 classes:

- Decays of H $\rightarrow$ uu, dd, bs, bd, cu, sd each get their own class. These are our signal samples.
- Decays of $H \rightarrow \mathrm{bb}, \mathrm{cc}, \mathrm{ss}, \mathrm{gg}, \mathrm{WW}, \mathrm{ZZ}, Z\gamma, \tau\tau$ are grouped in one class, representing Standard Model Higgs decays.
- Decays of $e^+e^- \rightarrow ZZ, WW, \mu\mu$ and $e\gamma \rightarrow eZ \rightarrow e\mu\mu, \gamma e \rightarrow eZ \rightarrow e\mu\mu, \gamma\gamma \rightarrow \mu\mu$ were grouped in a single class for Standard Model Backgrounds.

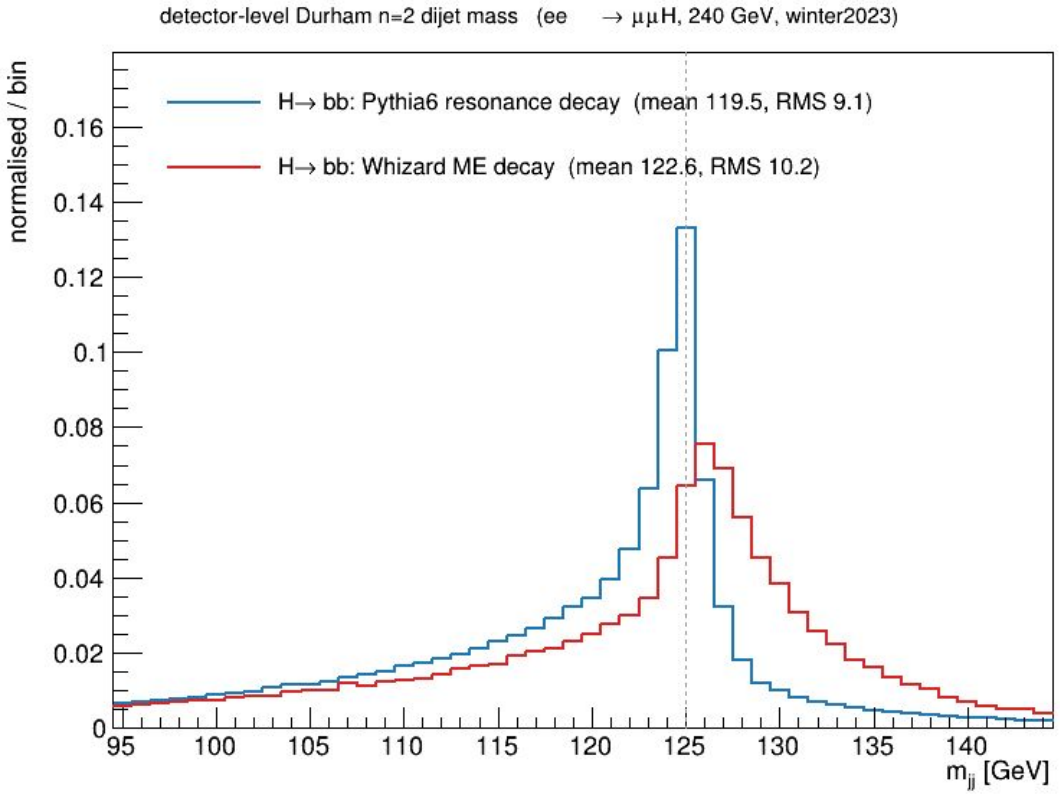


Cutflow Graph for each selection

Other than the 8 rare decay classes, the other samples were from the Winter 2023 production campaign.

Rare decay samples were produced with the same procedure, using WHIZARD3 to calculate the matrix element and PYTHIA6 for showering.

Process / Channel	Sample Name
$e^+e^- \to \mu^+\mu^- H \to \mu^+\mu^- bs$	wzp6_ee_mumuH_Hbs_W4p1MeV_ecm240
$e^+e^- \to \mu^+\mu^- H \to \mu^+\mu^- bd$	wzp6_ee_mumuH_Hbd_W4p1MeV_ecm240
$e^+e^- \to \mu^+\mu^- H \to \mu^+\mu^- cu$	wzp6_ee_mumuH_Hcu_W4p1MeV_ecm240
$e^+e^- \to \mu^+\mu^- H \to \mu^+\mu^- sd$	wzp6_ee_mumuH_Hsd_W4p1MeV_ecm240
$e^+e^- \to \mu^+\mu^- H \to \mu^+\mu^- WW$	wzp6_ee_mumuH_HWW_ecm240
$e^+e^- \to \mu^+\mu^- H \to \mu^+\mu^- ZZ$	wzp6_ee_mumuH_HZZ_noInv_ecm240
$e^+e^- \to \mu^+\mu^- H \to \mu^+\mu^- \tau^+\tau^-$	wzp6_ee_mumuH_Htautau_ecm240
$e^+e^- \to \mu^+\mu^- H \to \mu^+\mu^- b\bar{b}$	wzp6_ee_mumuH_Hbb_ecm240
$e^+e^- \to \mu^+\mu^- H \to \mu^+\mu^- s\bar{s}$	wzp6_ee_mumuH_Hss_ecm240
$e^+e^- \to \mu^+\mu^- H \to \mu^+\mu^- c\bar{c}$	wzp6_ee_mumuH_Hcc_ecm240
$e^+e^- \to \mu^+\mu^- H \to \mu^+\mu^- d\bar{d}$	wzp6_ee_mumuH_Hdd_ecm240
$e^+e^- \to \mu^+\mu^- H \to \mu^+\mu^- u\bar{u}$	wzp6_ee_mumuH_Huu_ecm240
$e^+e^- \to \mu^+\mu^- H \to \mu^+\mu^- gg$	wzp6_ee_mumuH_Hgg_ecm240
$e^+e^- \to \mu^+\mu^- H \to \mu^+\mu^- Z\gamma$	wzp6_ee_mumuH_HZa_ecm240
$e^+e^- \to ZZ$	p8_ee_ZZ_ecm240
$e^+e^- \to WW$	p8_ee_WW_ecm240
$e^+e^- \to Z\ell^+\ell^-$	wzp6_ee_mumu_ecm240
$e\gamma \to eZ \to e\mu^+\mu^-$	wzp6_egamma_eZ_Zmumu_ecm240
$\gamma e \to eZ \to e\mu^+\mu^-$	wzp6_gammae_eZ_Zmumu_ecm240
$\gamma\gamma \to \mu^+\mu^-$	wzp6_gaga_mumu_60_ecm240



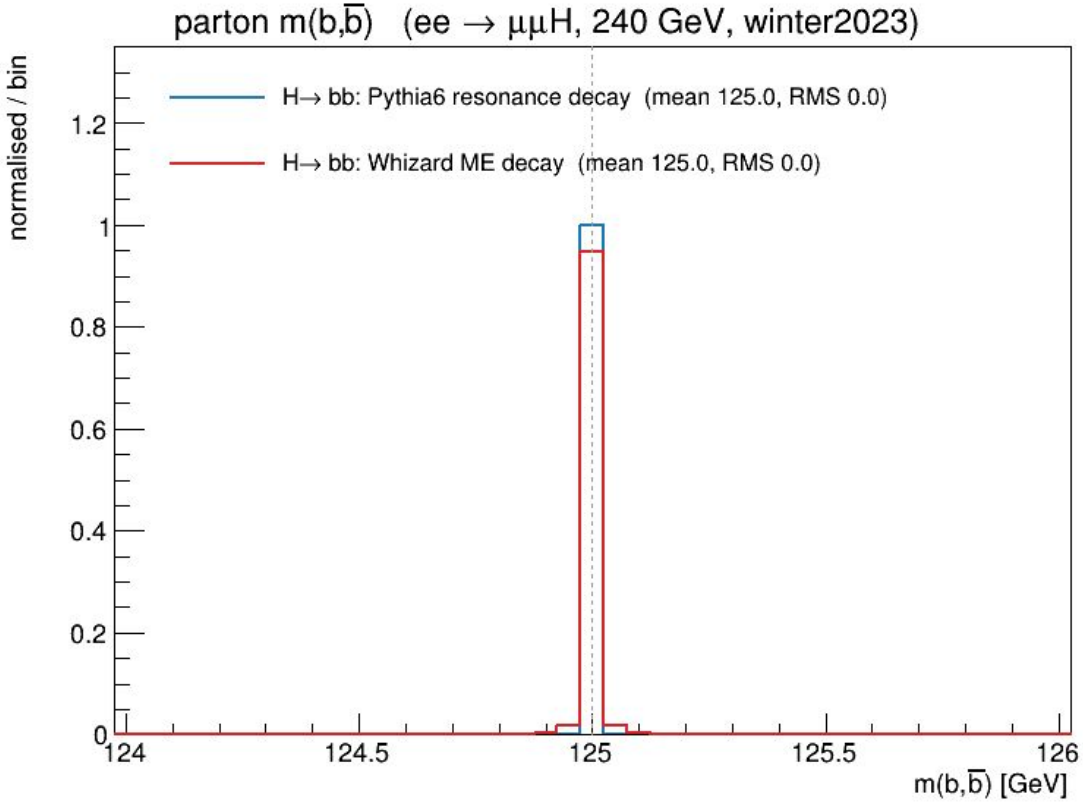
Noticed a strange 1 GeV shift between two of the centrally produced samples.

Compared 2 cases:

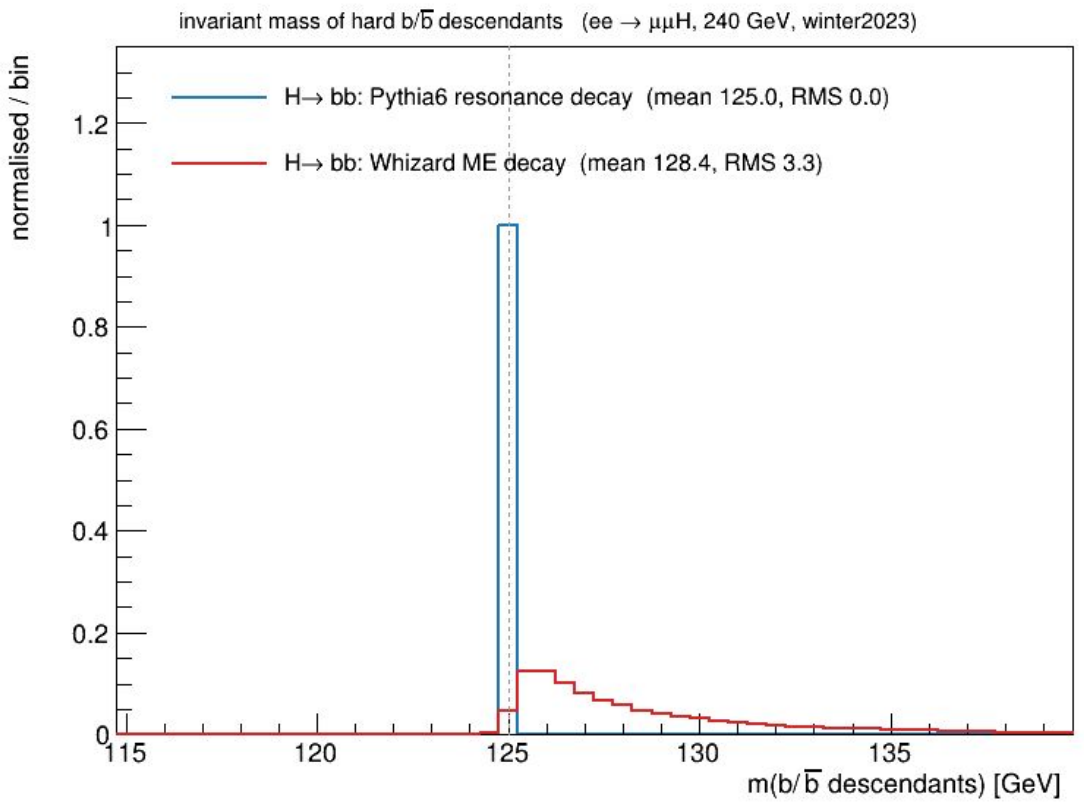
1.) WHIZARD3 calculates $e^+e^- \rightarrow \mu\mu H$, PYTHIA6 decays $H \rightarrow b\bar{b}$

2.) WHIZARD3 calculates $e^+e^- \rightarrow \mu\mu b\bar{b}$, PYTHIA6 just showers the event

Parton Level Decay



Particle Level



Difference arose from whether PYTHIA6 received an explicit H in the record

Parton level sees no difference.
After showering the blue curve conserves 4-momentum, while the red has global recoil

We decided that for our purposes, the PYTHIA6 Resonance Decay was more accurate. All rare decay samples were generated with WHIZARD3 calculating $e^+e^- \rightarrow \mu\mu H$

Used 41 training variables:

- A Graph Neural Network Tagger produced jet tagging scores, probabilities that a jet originated from various particles

- m_{jj}, the reconstructed invariant dijet mass from the Higgs decay

- The recoil mass $M^2_{recoil} = \left(\sqrt{s} - E_{\mu\mu}\right)^2 - \vec{p}^2_{\mu\mu}$, describing the invariant mass of the system excluding the muons

- Jet Substructure variables $d_{12}, d_{23}, d_{34}, \text{and } d_{45}$, general distance parameters describing the jets in time

For completeness:

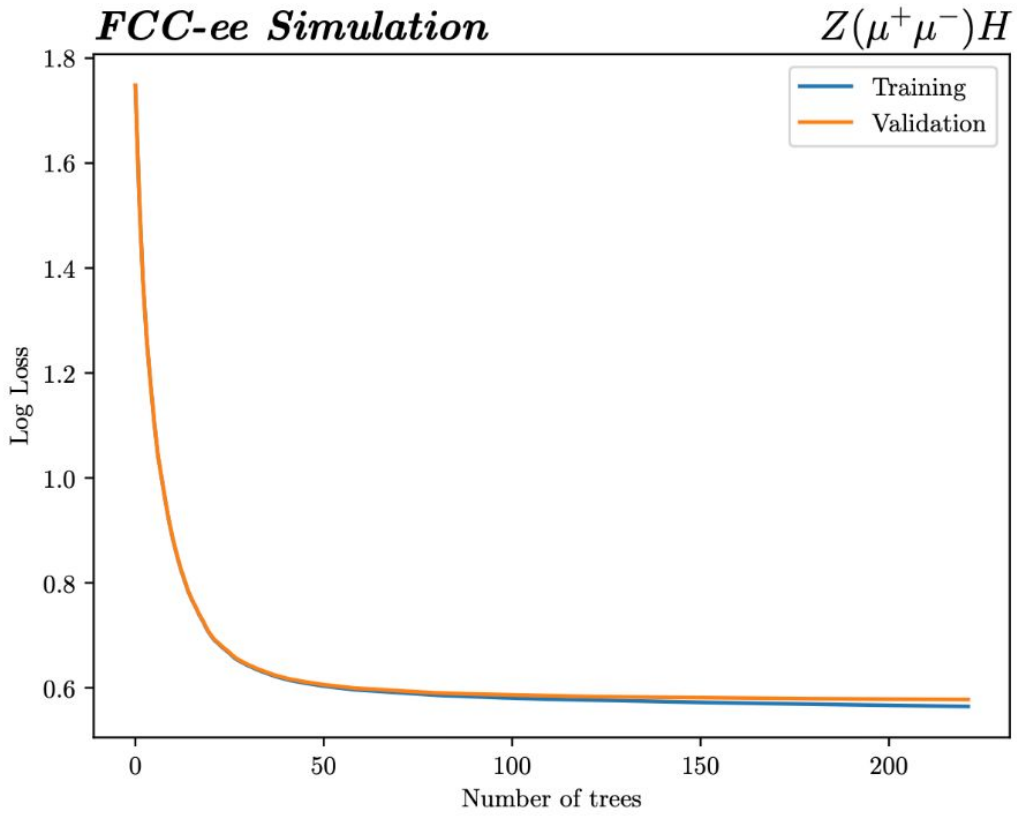
Category	Variable / Symbol	Description
$Z \to \ell\ell$ **System**	$m_{\ell\ell}, p_{\ell\ell}, \theta_{\ell\ell}$	Invariant mass, momentum, and polar angle of Z candidate
	m_{recoil}	Recoil mass against $Z \to \ell\ell$ system
	$p_{\ell,1}, \theta_{\ell,1}, p_{\ell,2}, \theta_{\ell,2}$	Momentum and polar angle of leading and subleading leptons
	Acollinearity, Acoplanarity	Spatial angles between the two leptons
Higgs System	m_{jj} (m_H)	Invariant mass of dijet Higgs candidate
Jet Kinematics	$p_{j1}, \theta_{j1}, \phi_{j1}, m_{j1}, E_{j1}$	Kinematic variables and energy of leading jet
	$p_{j2}, \theta_{j2}, \phi_{j2}, m_{j2}, E_{j2}$	Kinematic variables and energy of subleading jet
	$N_{\text{const},1}, N_{\text{const},2}$	Constituent multiplicity for jets 1 and 2
	$d_{12}, d_{23}, d_{34}, d_{45}$	Durham jet clustering distance measures
Flavor Tags	Flavour scores (j_1, j_2)	Tagger score outputs (b, c, s, u, d, g, τ)

Hyperparameter	Value
Objective	`multi:softprob`
Number of classes	8
$N_{\text{estimators}}$	350
Learning rate (η)	0.20
Maximum tree depth	3
Subsample ratio	0.5
Feature subsample (`colsample_bytree`)	0.5
Minimum child weight	10
Gamma (γ)	3
Max delta step	0
Early stopping rounds	25
Evaluation metrics	`mlogloss`, `merror`

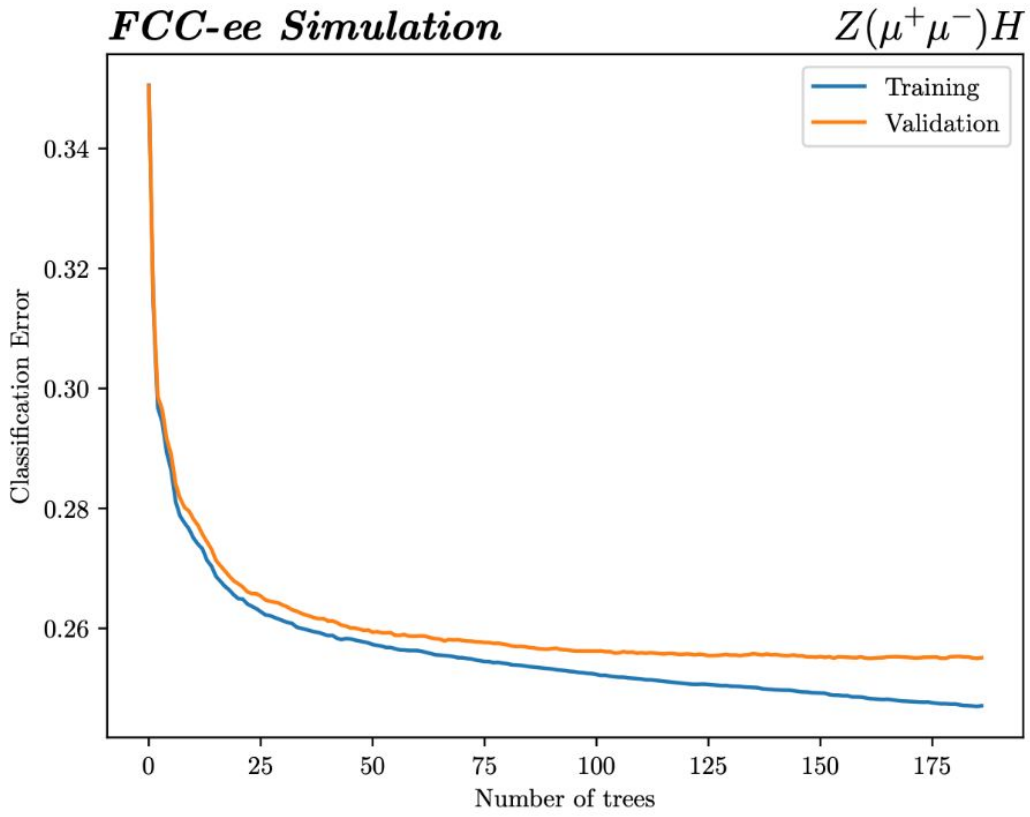
Important Hyperparameters for Training

Boosted Decision Tree Output and Performance

The quality of the BDT may be observed from plots of its performance metrics:



(a) Log loss evaluation.

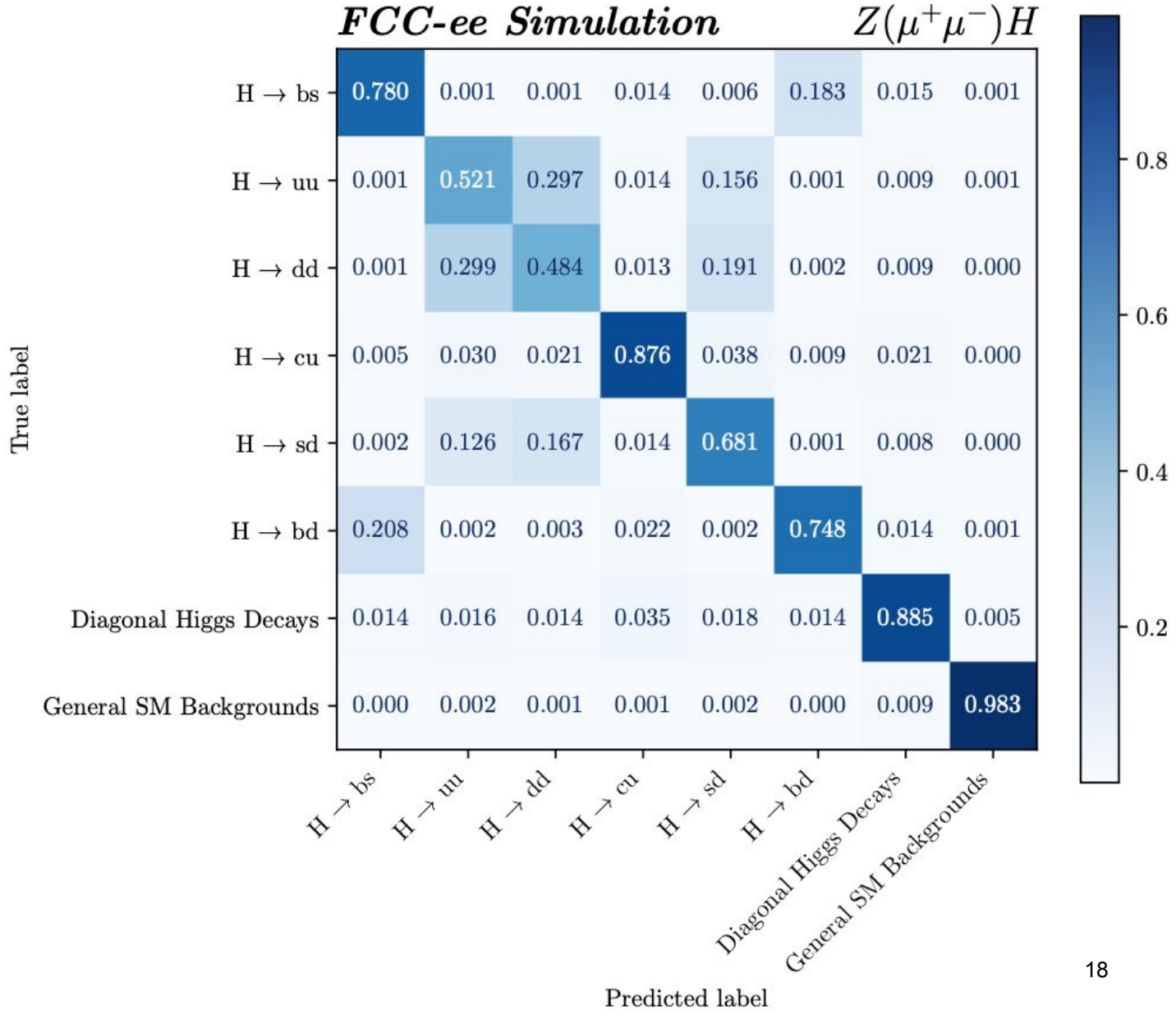


(b) Classification Error

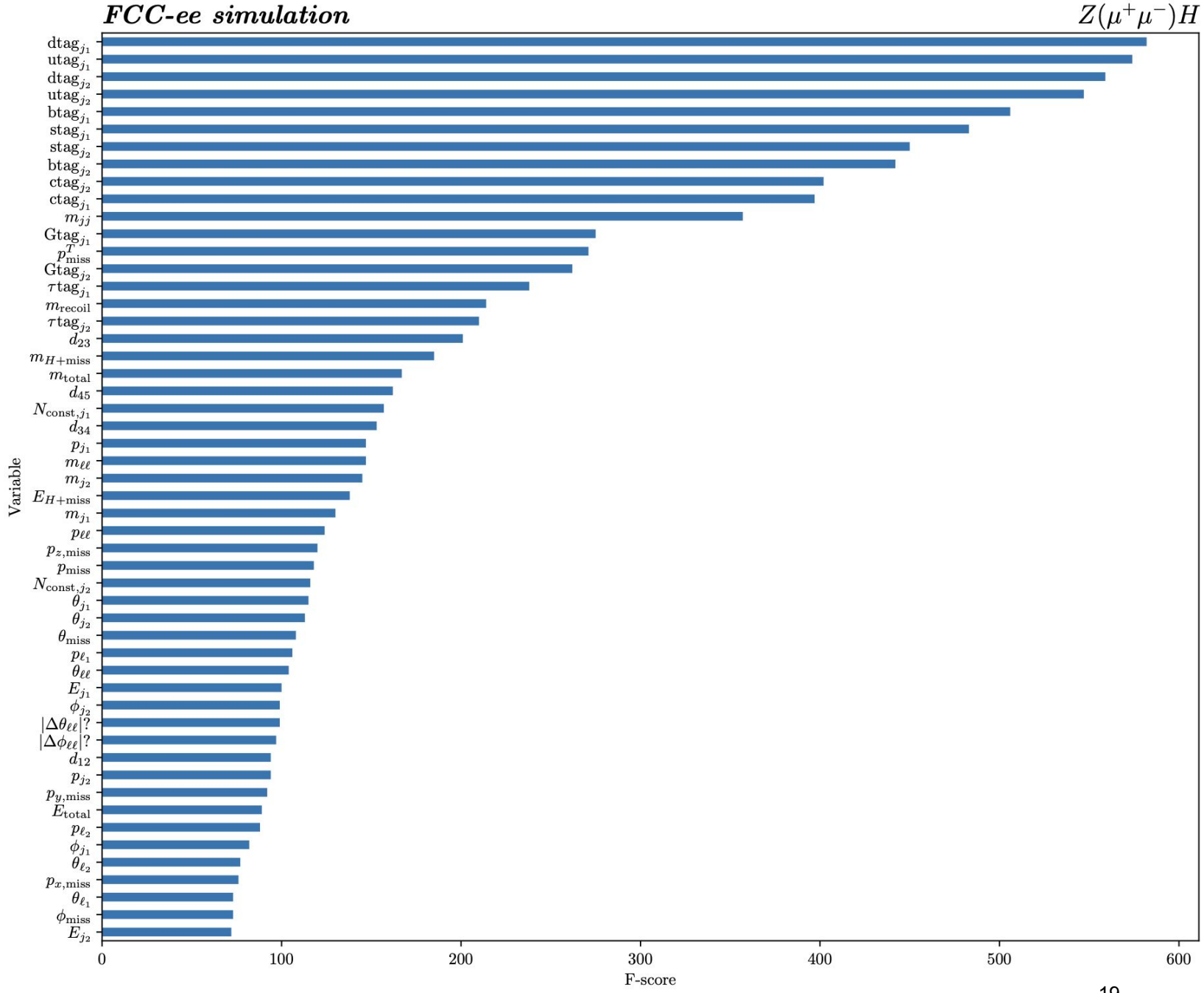
Further optimization should involve training up to 100 trees to prevent overfitting.

Confusion matrix between each of the 8 classes of the multiclassifier

c and b quarks have excellent tagging, which explains the behavior exhibited in this graph.



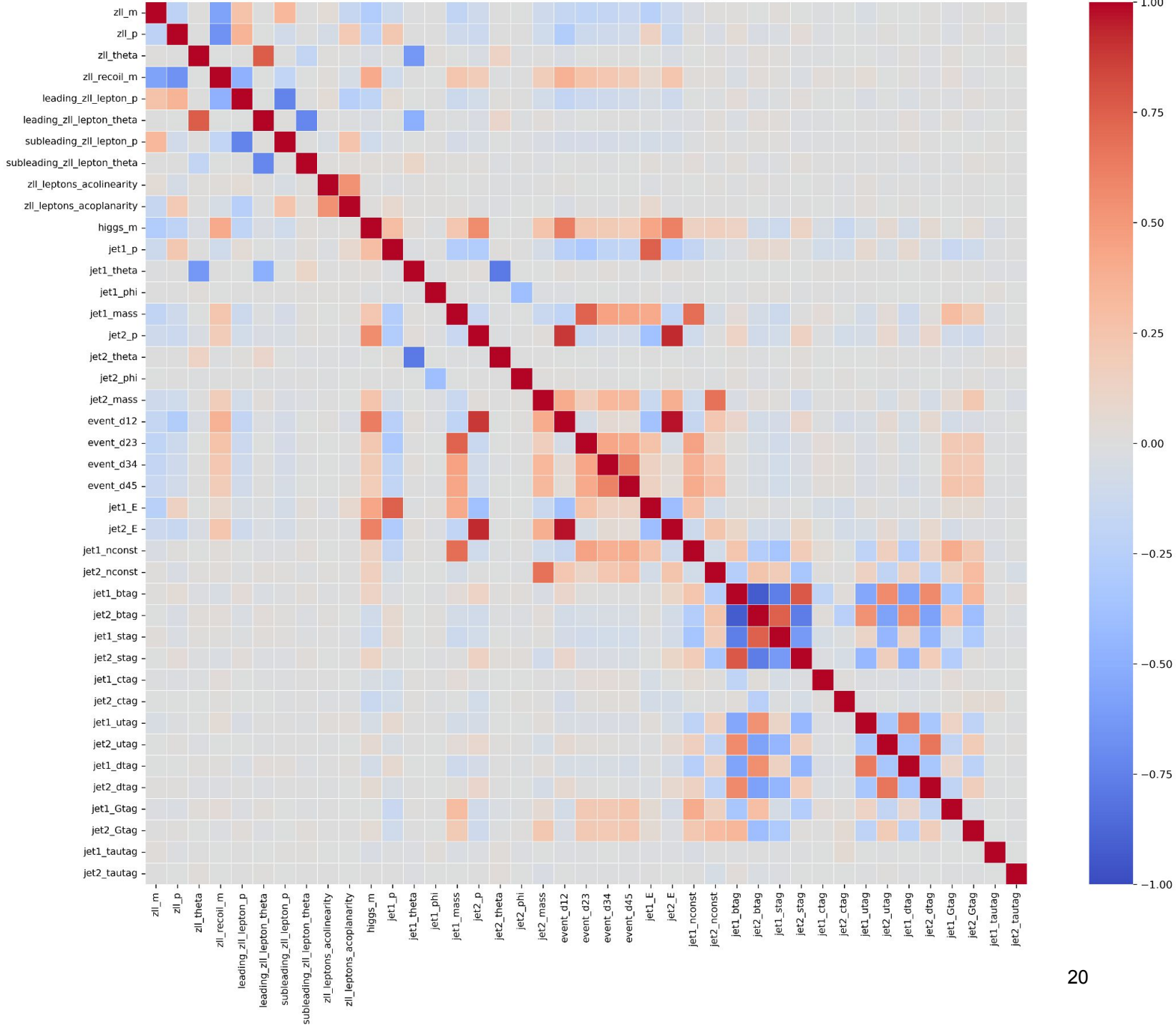
Plot of Importance of 41 input variables



Brookhaven National Laboratory

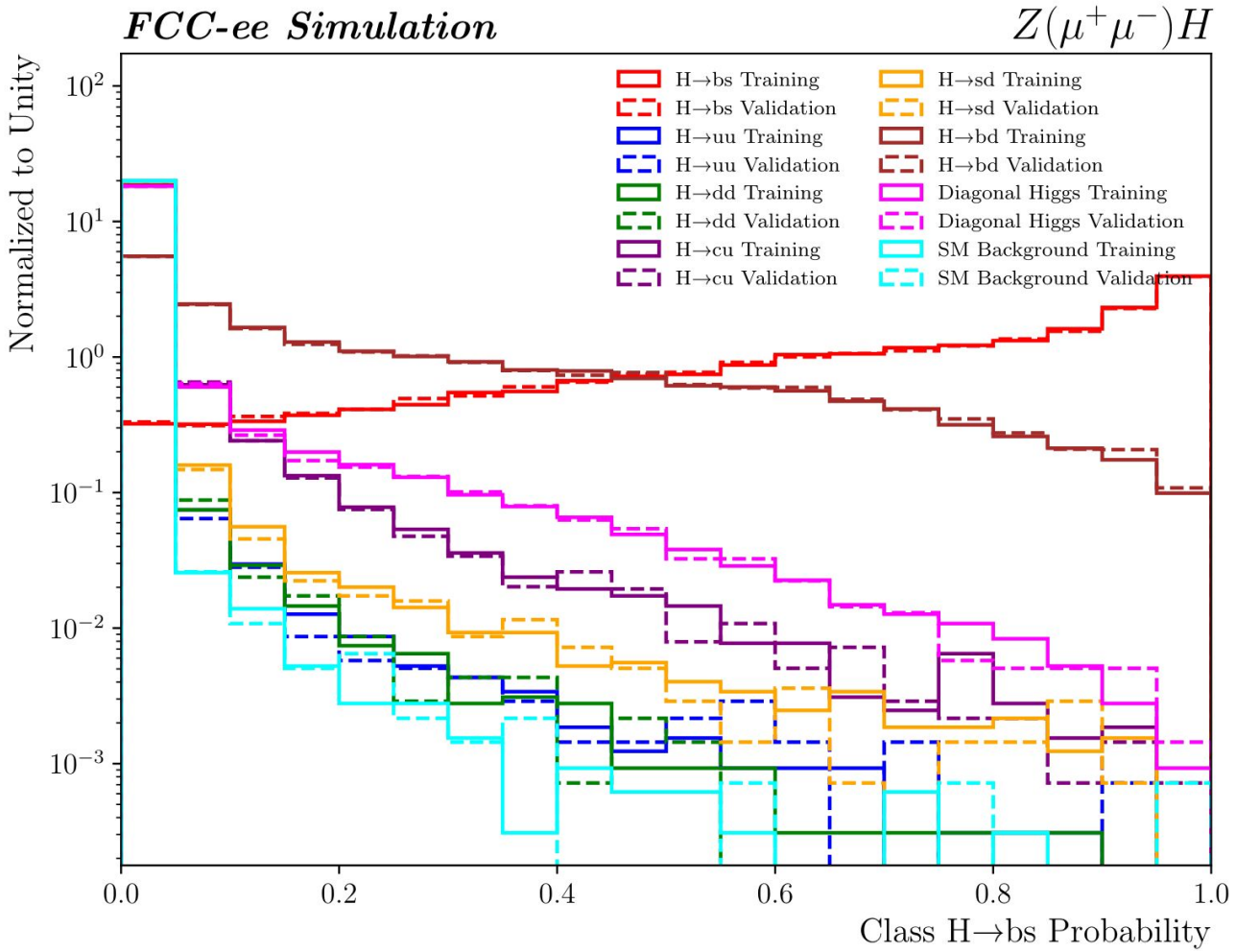
Feature Correlation Matrix between each variable

Highly correlated variables should be trimmed so the BDT learns important information from each split.



Brookhaven National Laboratory

The multiclassifier BDT outputs a probability that an event belongs to each of the 8 classes. In total, these scores sum to 1.



A representative plot of the output BDT score for the rare decay class $H \rightarrow bs$.

Brookhaven National Laboratory

Maximum Likelihood Fit

We can make a maximum likelihood fit to set a limit on what branching ratio they could have. We used the CMS Higgs Combine Tool.

This involves calculating the quantity $N_i(\mu, \boldsymbol{\theta}) = \mu \cdot S_i(\boldsymbol{\theta}) + \sum_{b \in \mathrm{bgr}} B_{b,i}(\boldsymbol{\theta})$ describing the predicted amount of events in each bin of the BDT output.

θ are nuisance parameters to model experimental uncertainty. We used:

- 0.5% luminosity uncertainty
- Dynamically generated statistical uncertainty for each bin

We compare the prediction to an Asimov dataset to derive our final confidence levels.

Results

Decay Channel	SM Prediction	68% CL UL	95% CL UL	3σ Reach	5σ Reach
$H \to bs$	9×10^{-8}	2.707×10^{-4}	5.47×10^{-4}	8.351×10^{-4}	1.43×10^{-3}
$H \to uu$	1.2×10^{-7}	1.621×10^{-3}	2.899×10^{-3}	4.214×10^{-2}	1.516×10^{-1}
$H \to dd$	5.5×10^{-7}	1.665×10^{-3}	2.846×10^{-3}	4.181×10^{-2}	1.507×10^{-1}
$H \to cu$	3×10^{-20}	6.485×10^{-4}	1.289×10^{-3}	1.966×10^{-3}	3.341×10^{-3}
$H \to sd$	2×10^{-15}	1.353×10^{-3}	2.684×10^{-3}	4.111×10^{-3}	6.909×10^{-3}
$H \to bd$	4×10^{-9}	2.763×10^{-4}	5.583×10^{-4}	8.534×10^{-4}	1.46×10^{-3}

Final Confidence Levels for this study

As comparison, here are the CL upper limits set by Michele Selvaggi's study. We have found comparable results with this channel with minimal optimization of the BDT.

Table 2 Summary of 95% CL upper limits on the branching ratios of rare and FCNC Higgs decays, assuming 10.8 ab^{-1} using only the $\nu\nu$H(jj) channel at $\sqrt{s} = 240$ GeV, compared to SM predictions and existing indirect constraints.

Decay	SM Prediction	this work (95% CL)	current constraints
$H \to uu$	1.2×10^{-7}	$< 1.2 \times 10^{-3}$	3×10^{-2}
$H \to dd$	5.5×10^{-7}	$< 1.2 \times 10^{-3}$	7×10^{-3}
$H \to bs$	9×10^{-8}	$< 3.1 \times 10^{-4}$	$< 1.6 \times 10^{-3}$
$H \to bd$	4×10^{-9}	$< 2.2 \times 10^{-4}$	$< 10^{-3}$
$H \to cu$	3×10^{-20}	$< 2.0 \times 10^{-4}$	$< 2 \times 10^{-3}$
$H \to sd$	2×10^{-15}	$< 6.5 \times 10^{-4}$	—

Further Steps

This study could be greatly improved through hyperparameter optimization, as well as more careful selection of input variables.

The previous study also implemented event categorization to isolate high purity events and gain greater enhancement of the signal over the background. This could similarly be implemented!

Other methods other than a BDT may be attempted for future studies, including neural networks or other machine learning strategies.

This study will be continued by Alex Smith in collaboration with Giovanni Marchiori and Dian Yu from the APC group.