Jet origin identification and its application at future e+e- Higgs factory

Hao Liang[1,2], Yongfeng Zhu[3], Yuexin Wang[1,4], Yuzhi Che[1,2], Manqi Ruan[1,2], Chen Zhou[3], and Huilin Qu[5]
[1] Institute of High Energy Physics, Chinese Academy of Sciences, Beijing, China
[2] University of Chinese Academy of Sciences, Beijing, China
[3] State Key Laboratory of Nuclear Physics and Technology of Peking University, Beijing, China
[4] China Center of Advanced Science and Technology, Beijing, China, [5] CERN, Geneva, Switzerland

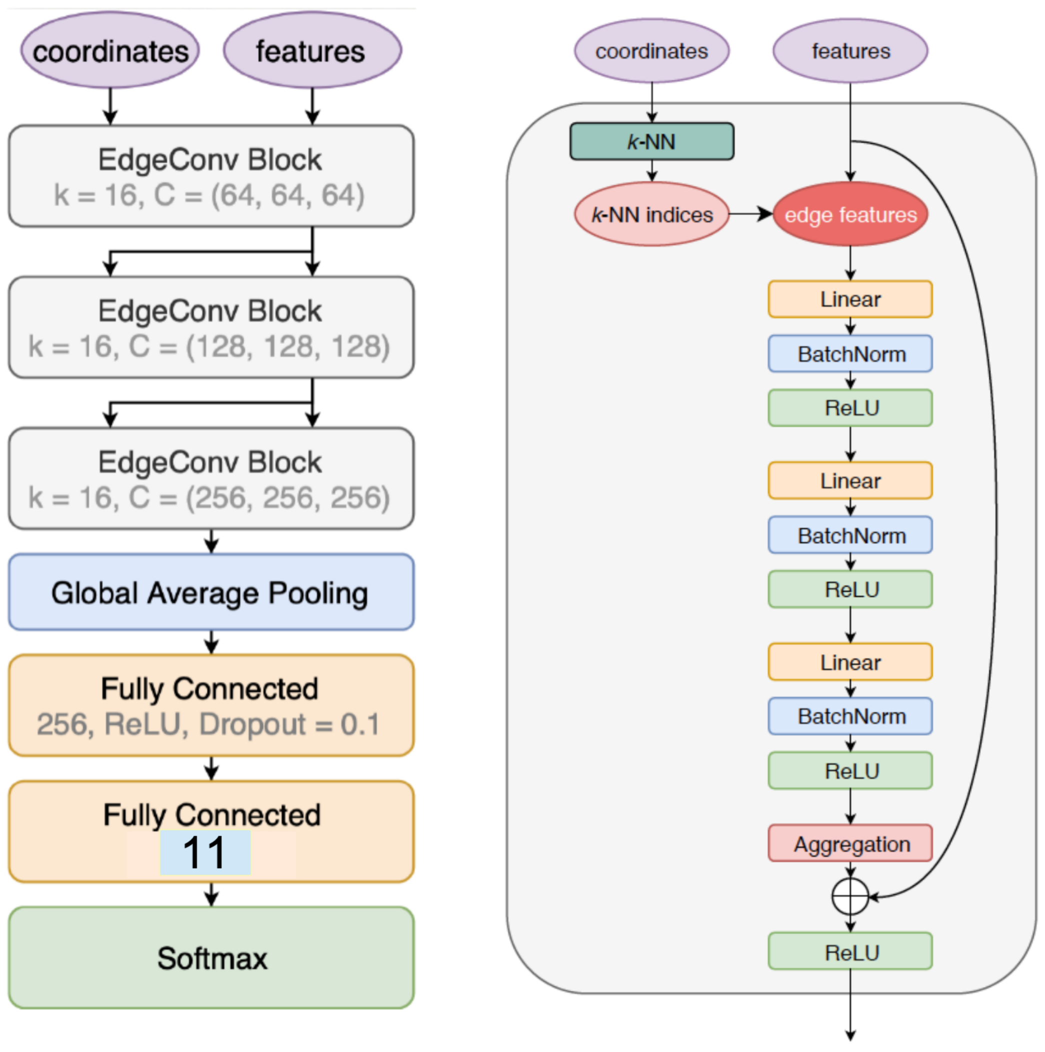
Accepted by Physics Review Letter on May 1, 2024

Introduction

- The jet origin identification (encompassing flavor and charge tagging) plays a vital role in Higgs property measurement. Additionally, the jet charge is crucial for weak mixing angle and time-dependent CP measurements.
- Using AI technology, the concept of jet origin identification of **11** different origins, 5 quarks (b, c, s, u, d), 5 anti-quarks $(\bar{b}, \bar{c}, \bar{s}, \bar{u}, \bar{d})$, and the gluon is realized.
- The identification is applied to the measurement of **rare** $(H \to s\bar{s}, u\bar{u}, d\bar{d})$, **FCNC** $(H \to sb, db, us, ds)$, and $|V_{cb}|$ at the Circular Electron Positron Collider (CEPC).

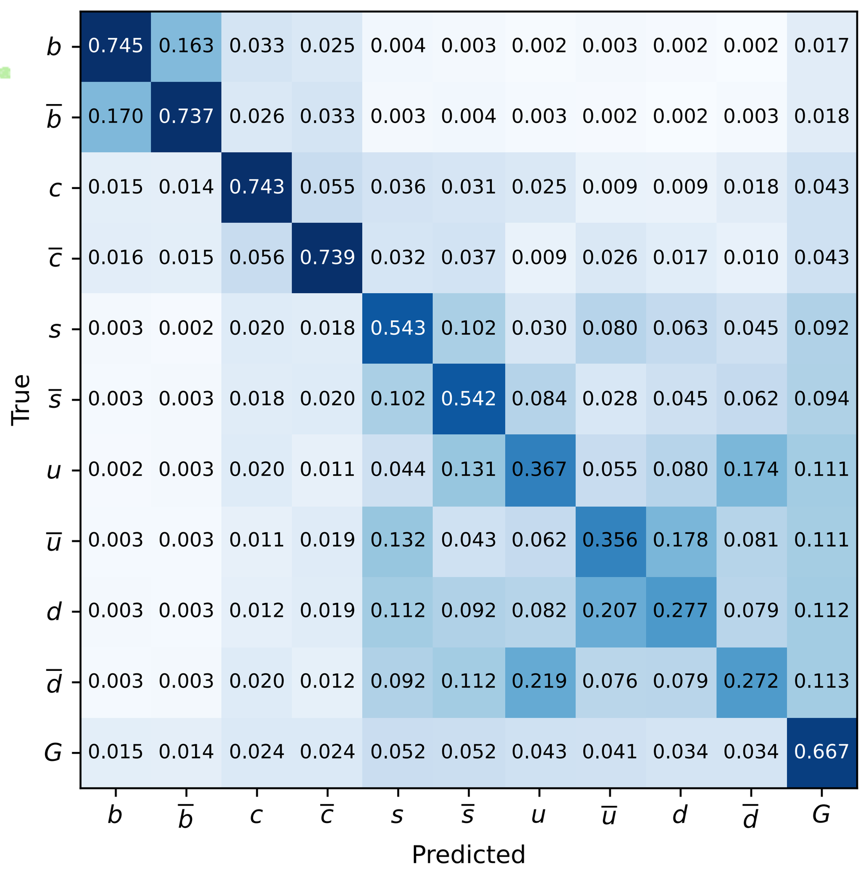
Sample and Methodology

- **Sample**: Di-jet events at $\sqrt{s}$ = 91.2, 240 GeV simulated with CEPC baseline
- **Software**: Pythia-6.4 (generator), MokkaPlus (Geant-4 based simulation), Arbor (particle flow reconstruction), e+e−-k$_t$ algorithm (jet clustering), ParticleNet (graph neural network)

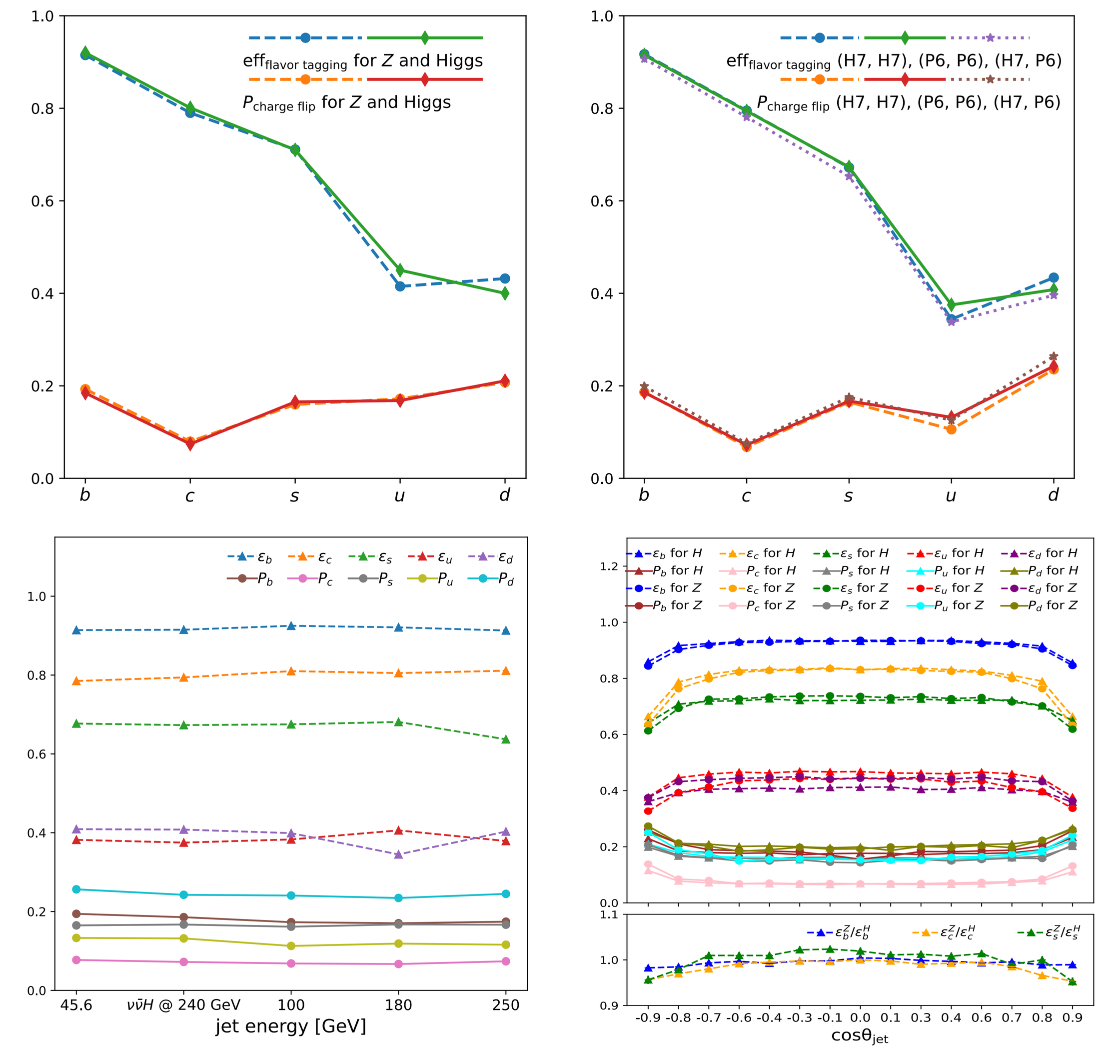


The Architecture of ParticleNet

Performance of Jet Origin Identification



The confusion matrix M$_{11}$ for $\nu\bar{\nu}H, H \to jj$ events at $\sqrt{s}$ =240 GeV.



- The comparisons of jet origin identification (jet flavor tagging efficiencies [ε] and jet charge flip rates [P]) between different physics processes ($Z \to q\bar{q}, \sqrt{s}$ = 91.2 GeV and $\nu\nu H, H \to q\bar{q}, \sqrt{s}$ = 240 GeV), hadronization generators (Herwig-7.2.2[H7], Pythia-6.4 [P6]), jet energies (M$_H$=91.2, 200, 360, 500 GeV), and jet polar angles.
- All results show that the jet origin identification stays stable, with respect to considered $\sqrt{s}$, generators, jet energies, and jet polar angles.

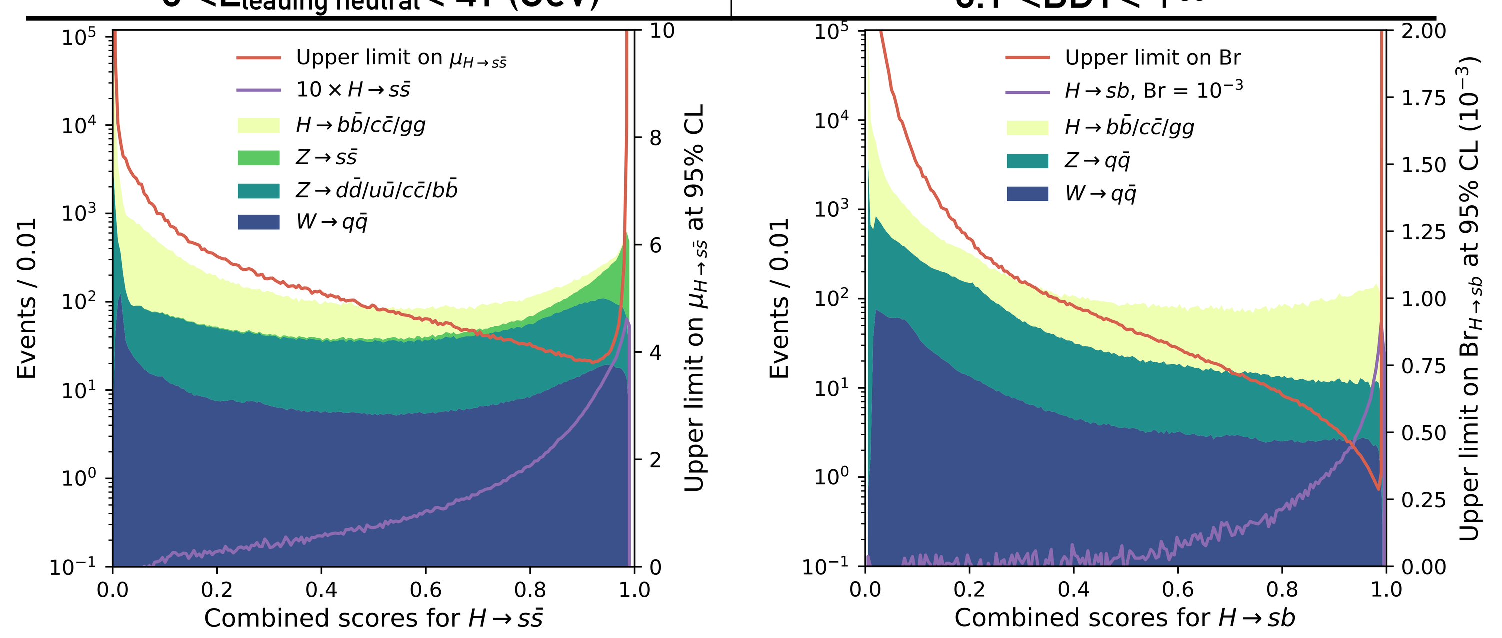
Benchmark Physics Analyses

Rare and exotic Higgs decay

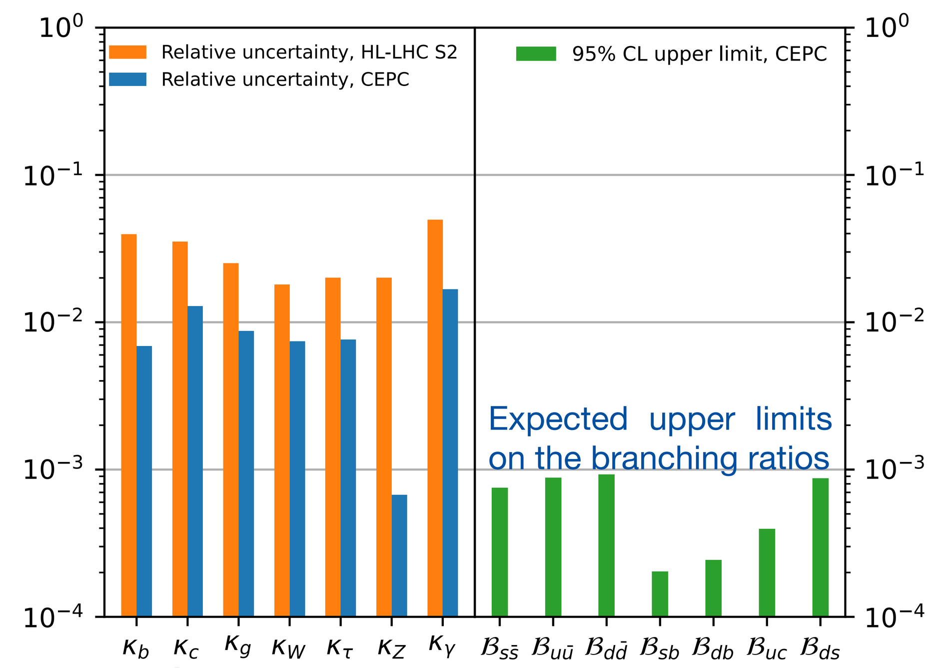
- Combine all jet flavor-tagging scores of the two jets into a BDT; the optimal cut on the combined scores leads to the upper limit of 3.8 on the signal strength $H \to s\bar{s}$ at 95% CL.
- By combining all three channels $(\mu^+\mu^-H, e^+e^-H, \nu\bar{\nu}H)$, the branching ratios of $H \to u\bar{u}$ and $d\bar{d}$ can be constrained to 0.091% and 0.095% at 95% CL, respectively.

Event Selections

74 < M_{recoil} < 131 (GeV)	20 < P_T < 60 (GeV)
109 < E_{vis} < 143 (GeV)	0 < P_l < 50 (GeV)
0 < $E_{leading\ lepton}$ < 42 (GeV)	3.375 < $-\log_{10}(y_{23})$ < +∞
40 < Multiplicity < 130	110 < $M_{invariant}$ < 134 (GeV)
0 < $E_{leading\ neutral}$ < 41 (GeV)	0.1 < BDT < +∞



The distributions of combined scores for signal and SM backgrounds, where the signals are (left panel) $H \to s\bar{s}$ and (right panel) $H \to sb$, respectively, in the $\nu\bar{\nu}H$ process, with CEPC nominal parameters.



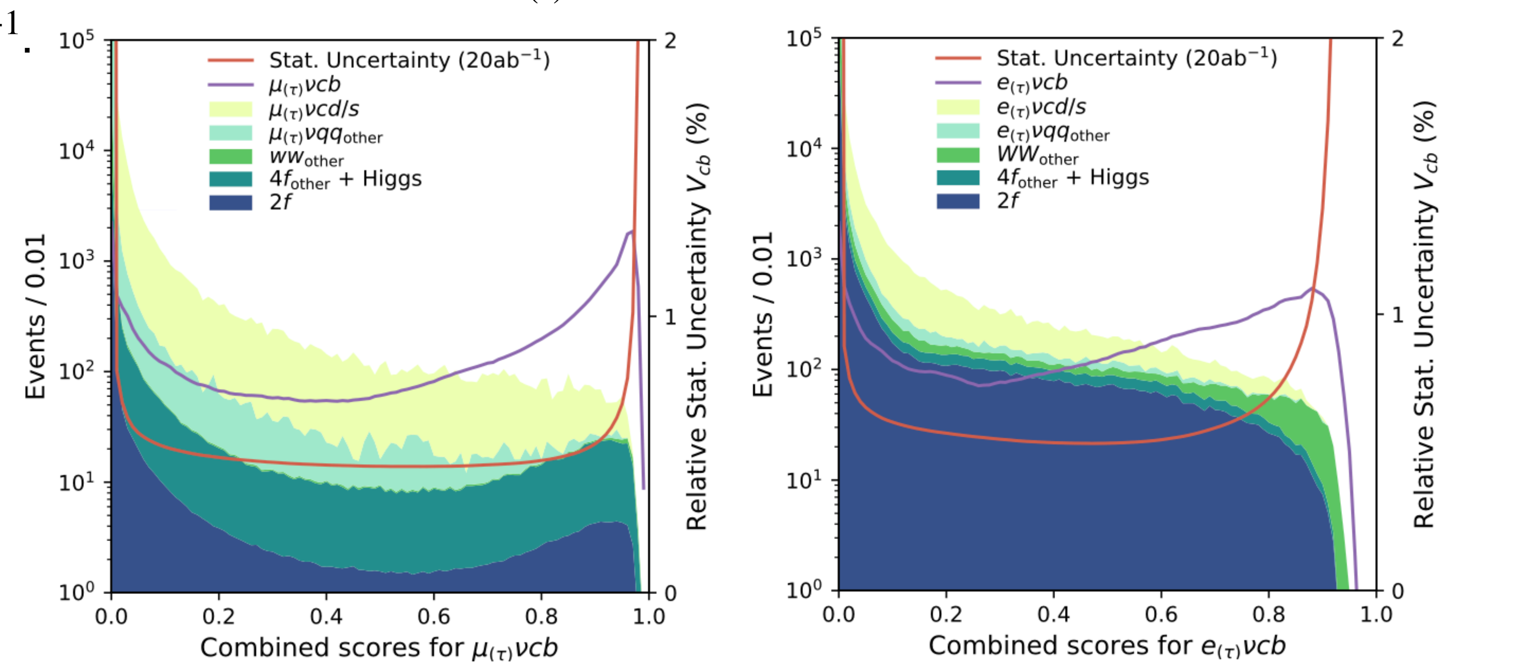
	Bkg. (10^3)			Upper limits on Br. (10^{-3})						
	H	Z	W	$s\bar{s}$	$u\bar{u}$	$d\bar{d}$	sb	db	uc	ds
$\nu\bar{\nu}$H	151	20	2.1	0.81	0.95	0.99	0.26	0.27	0.46	0.93
$\mu^+\mu^-$H	50	25	0	2.6	3.0	3.2	0.5	0.6	1.0	3.0
e^+e^-H	26	16	0	4.1	4.6	4.8	0.7	0.9	1.6	4.3
Comb.	-	-	-	0.75	0.91	0.95	0.22	0.23	0.39	0.86

Summary of background yields from $H \to b\bar{b}, c\bar{c}, gg$, Z, and W prior to the flavor-based event selection, with the expected upper limits on Higgs decay branching ratios at 95% CL under the background-only hypothesis.

$|V_{cb}|$ Measurement from W boson decays at future Higgs factories

https://doi.org/10.1007/JHEP12(2024)071

- Through $WW \to \ell\nu cb$ events at future Higgs factories, after kinematic cuts and jet flavor tagging cuts based on JOI, the relative statistical uncertainties of the CKM matrix elements $|V_{cb}|$ are estimated.
- The relative statistical uncertainties $|V_{cb}|$ are projected to be 0.91% for the $WW \to \mu_{(\tau)}\nu cb$ channel and 1.2% for the $WW \to e_{(\tau)}\nu cb$ channel, assuming a baseline integrated luminosity of $5ab^{-1}$.



The distributions of combined scores for signal and SM backgrounds, where the signals are (left panel) $WW \to \mu_{(\tau)}\nu cb$ and (right panel) $WW \to e_{(\tau)}\nu cb$, respectively, with CEPC nominal parameters.

Conclusion

1. Jet origin identification distinguishing 11 types of color jets has been performed.
2. For b, c, s jets:
   - Jet flavor tagging efficiencies: 92%, 79%, 67%
   - Jet charge flip rates: 19%, 7%, 17%.
3. The upper limits on branching ratios of rare $(H \to s\bar{s}, u\bar{u}, d\bar{d})$ and FCNC $(H \to sb, db, uc, ds)$ decay are 2x10^{-4} to 1x10^{-3} at 95% CL.
4. The resulting upper limit for $H \to s\bar{s}$ decay is approximately 3 times the prediction of the Standard Model.
5. Simultaneously evolving with state-of-the-art detector technology, reconstruction algorithms, and Artificial Intelligence, the jet origin identification algorithm developed here indicates that colored SM particles can potentially be identified with comparable performance to leptons and photons.