Jet Origin Identification at the CEPC

PhysRevLett.132.221802

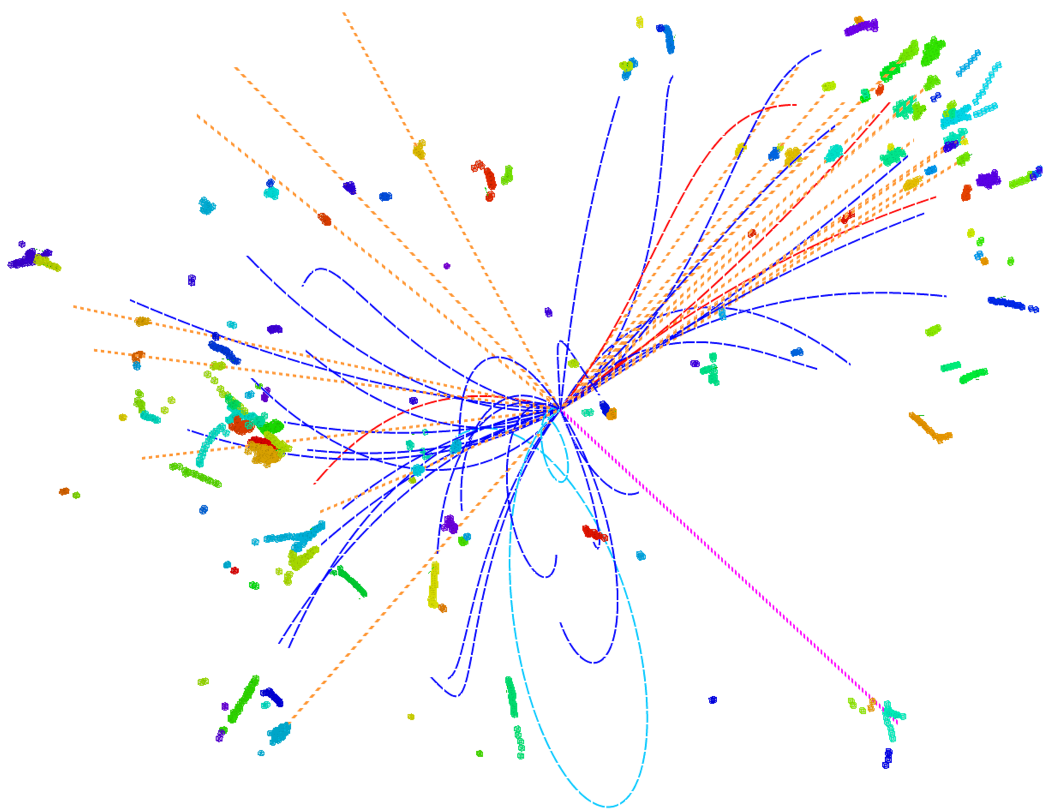
Hao Liang (IHEP), Yongfeng Zhu (PKU), Yuexin Wang (IHEP), Yuzhi Che (IHEP), Manqi Ruan (IHEP), Chen Zhou (PKU), Huilin Qu (CERN)

10,26,2024

2024 CEPC Workshop @ Hangzhou

Definition of Jet Origin Identification

Quarks and gluons carry color charge, and they can not travel freely.
Once generated in high-energy collisions, quarks, and gluon would fragment into numerous particles, which are called jet.



$\nu\bar{\nu}H, H \rightarrow gg$ at $\sqrt{s} = 240\ GeV$

Jet Origin Identification: categorizes jets into 5 quarks (b, c, s, u, d), 5 anti-quark ($\bar{b}, \bar{c}, \bar{s}, \bar{u}, \bar{d}$), and gluon.
It comprises the concepts of jet flavor tagging, jet charge measurement, s-quark tagging, and gluon finding.

motivation

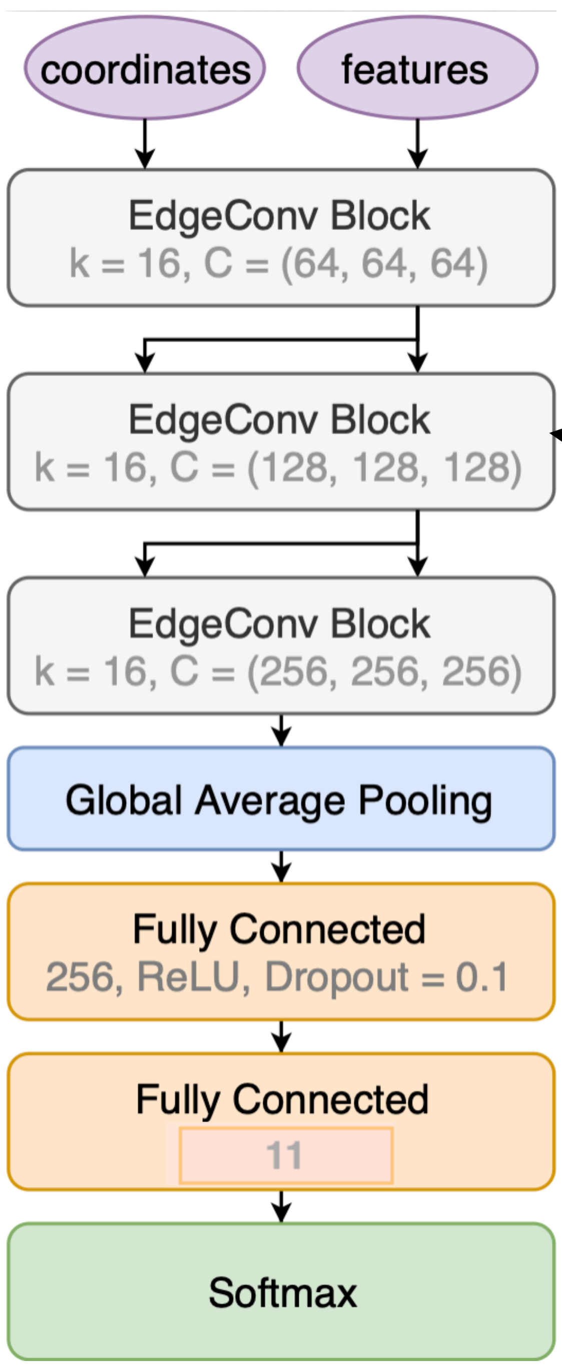
- A successful jet origin identification is critical for experimental particle physics at the energy frontier, especially for studying Higgs properties and understanding the QCD.

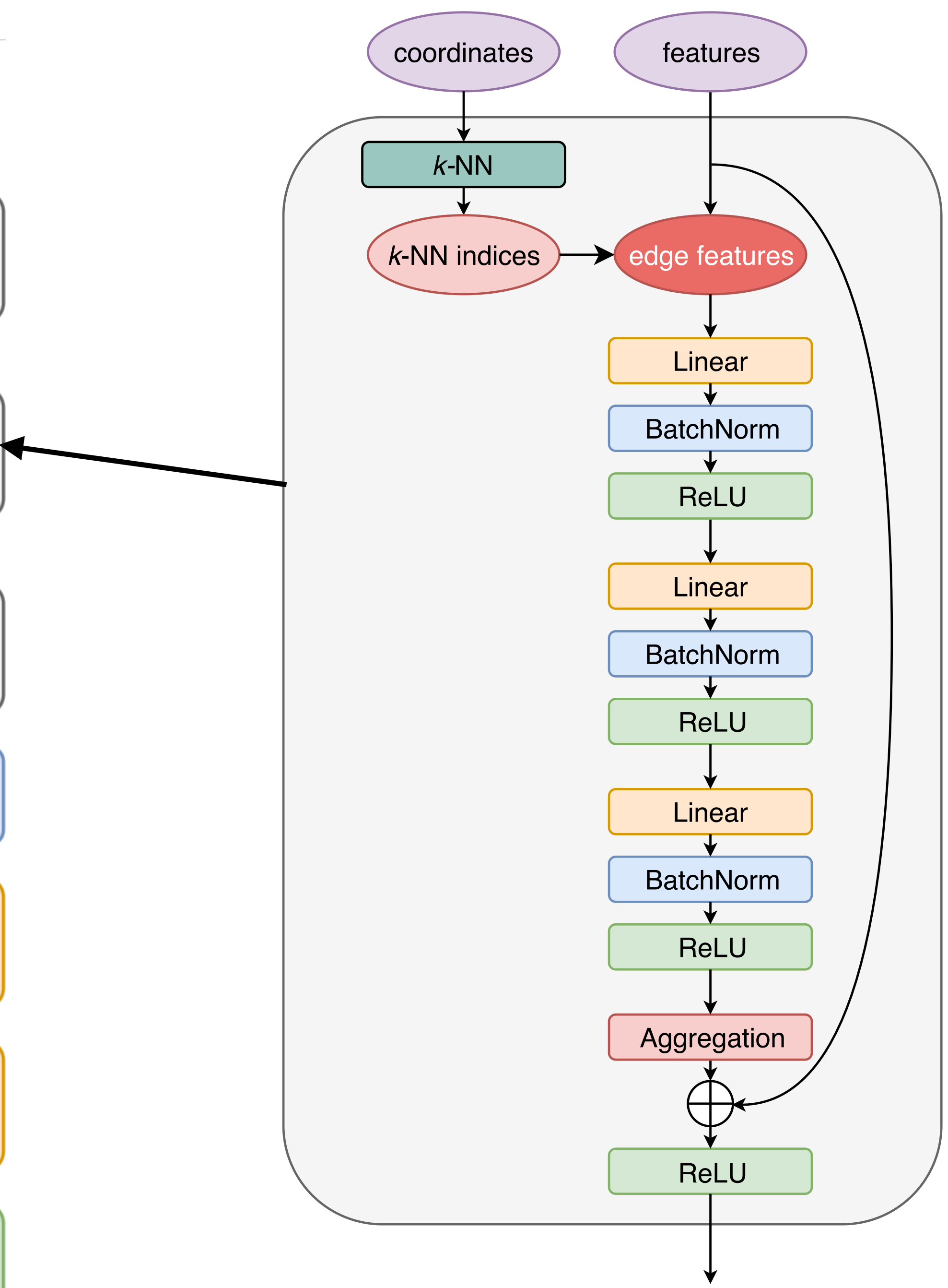
Contents

- samples and ParticleNet
- the performance of jet origin identification
- Benchmark physics analyses
- the dependent factors of JOI performance

Samples and ParticleNet

samples: Based on the CEPC baseline detector, full simulation of $e^+e^- \rightarrow \nu\bar{\nu}H, H \rightarrow b\bar{b}/c\bar{c}/s\bar{s}/u\bar{u}/d\bar{d}/gg$ at $\sqrt{s} = 240\ GeV$ generated with Whizard1.95 and Pythia-6.4.





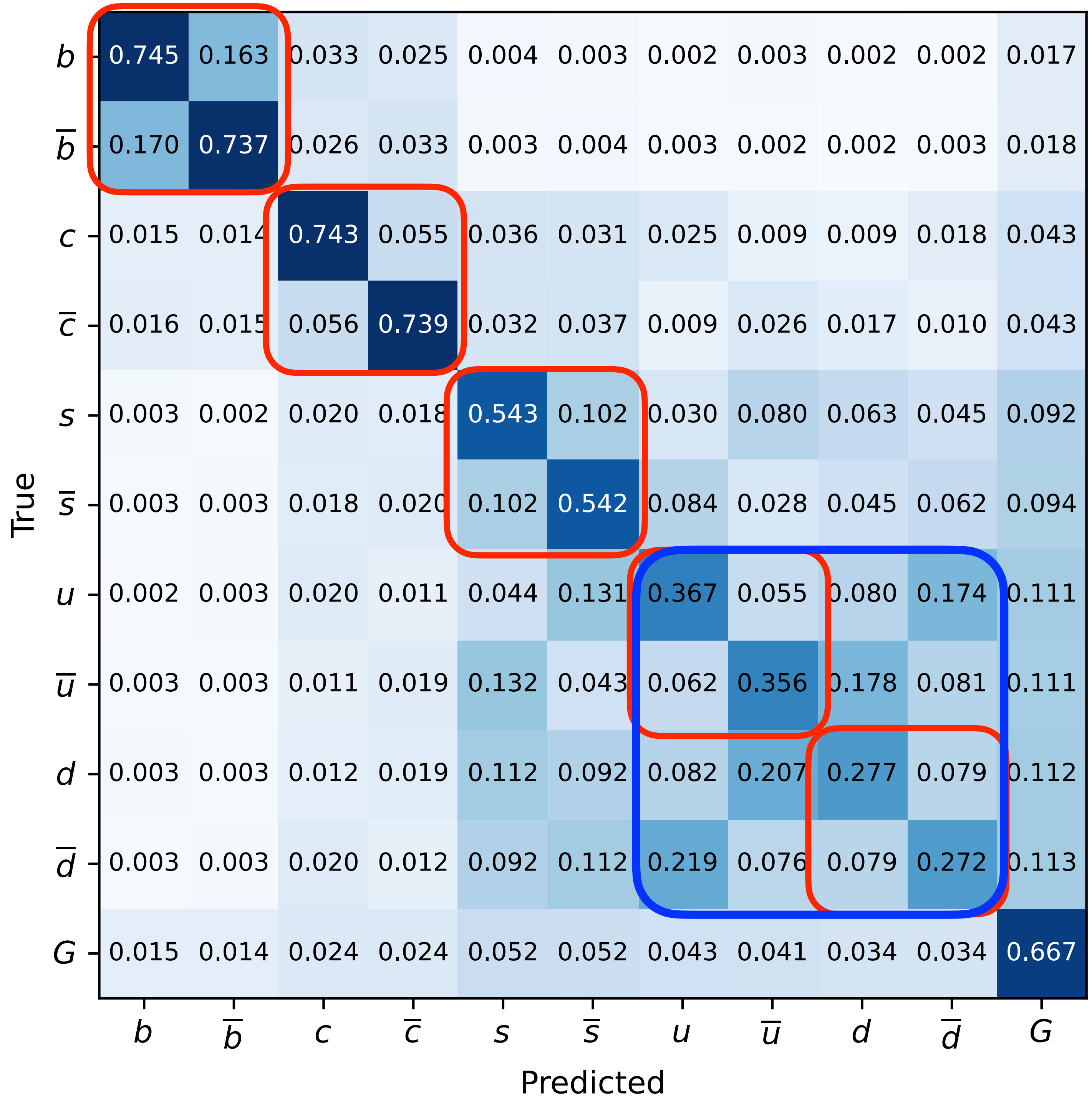
input features:

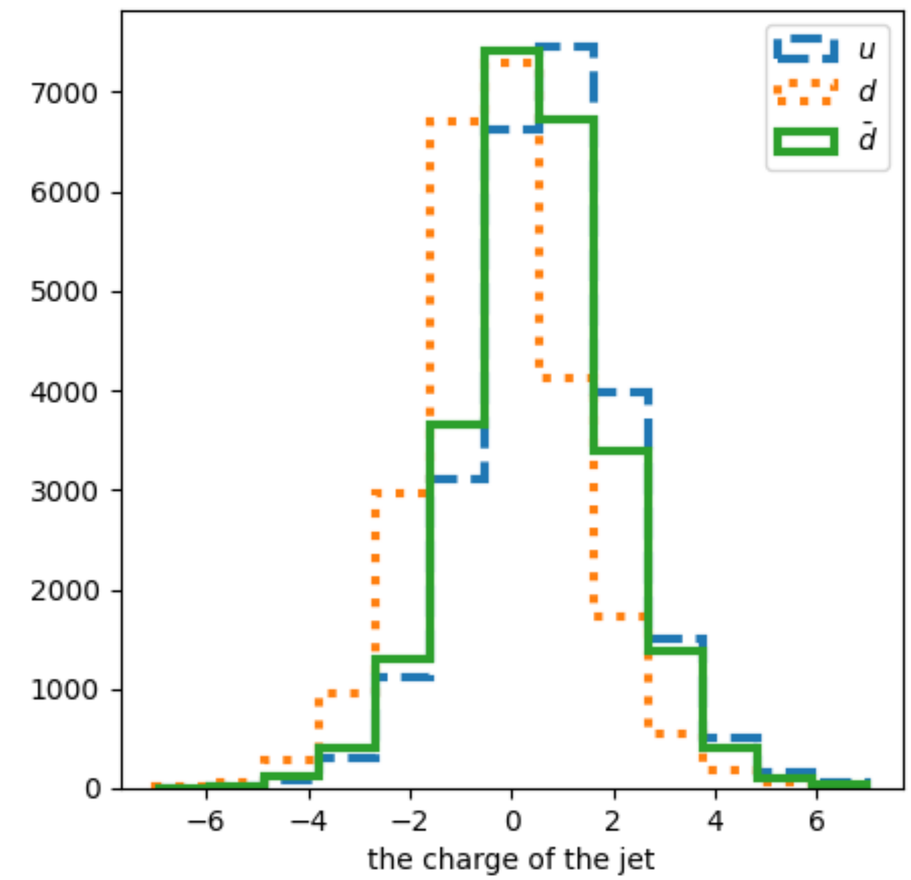
impact parameters, kinematic variables, and PID information of particles in a jet

https://doi.org/10.1103/PhysRevD.101.056019

the performance of jet origin identification

ParticleNet algorithm attaches each jet with 11 scores corresponding to 11 types of jets. Then the jet type is determined according to the maximum score.

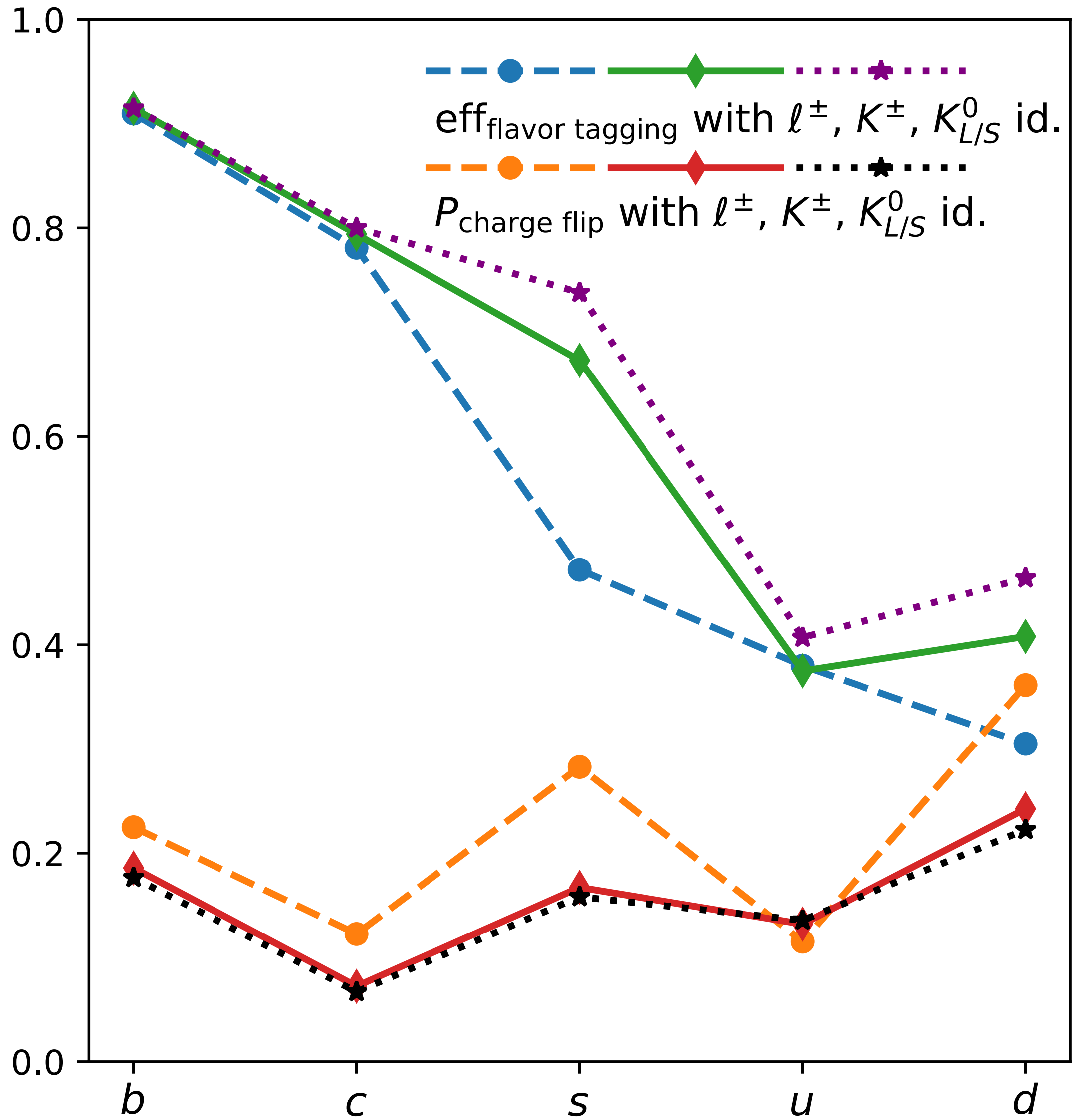




the performance of jet origin identification

jet flavor is defined as $max(b+\bar{b},\ c+\bar{c},\ s+\bar{s},\ u+\bar{u},\ d+\bar{d},\ g)$

jet charge is assigned by comparing the quark and anti-quark likelihoods of the corresponding flavor



To understand the impact of PID, three scenarios are compared.

1, assumes perfect identification of charged leptons ($\ell^{\pm}$)

2, further assumes perfect identification of the charged hadrons ($K^{\pm}$)

3, on top of the second scenario, assumes perfect identification of K_L and K_S.

default scenario: 2 scenario, based on:

Eur. Phys. J. C 80, 7 (2020)

Journal of Instrumentation 16, P06013 (2021)

Eur. Phys. J. C 78, 464 (2018)

Eur. Phys. J. C 83, 93 (2023)

Nucl. Instrum. Meth. A 1047, 167835 (2023)

Benchmark physics analyses

Begin with the existing analyses of $\nu\bar{\nu}H, H \to b\bar{b}/c\bar{c}/gg$, (arXiv:2203.01469) and combining the jet origin identification, we obtain the upper limits on branching ratios of seven Higgs rare and FCNC hadronic decay modes.

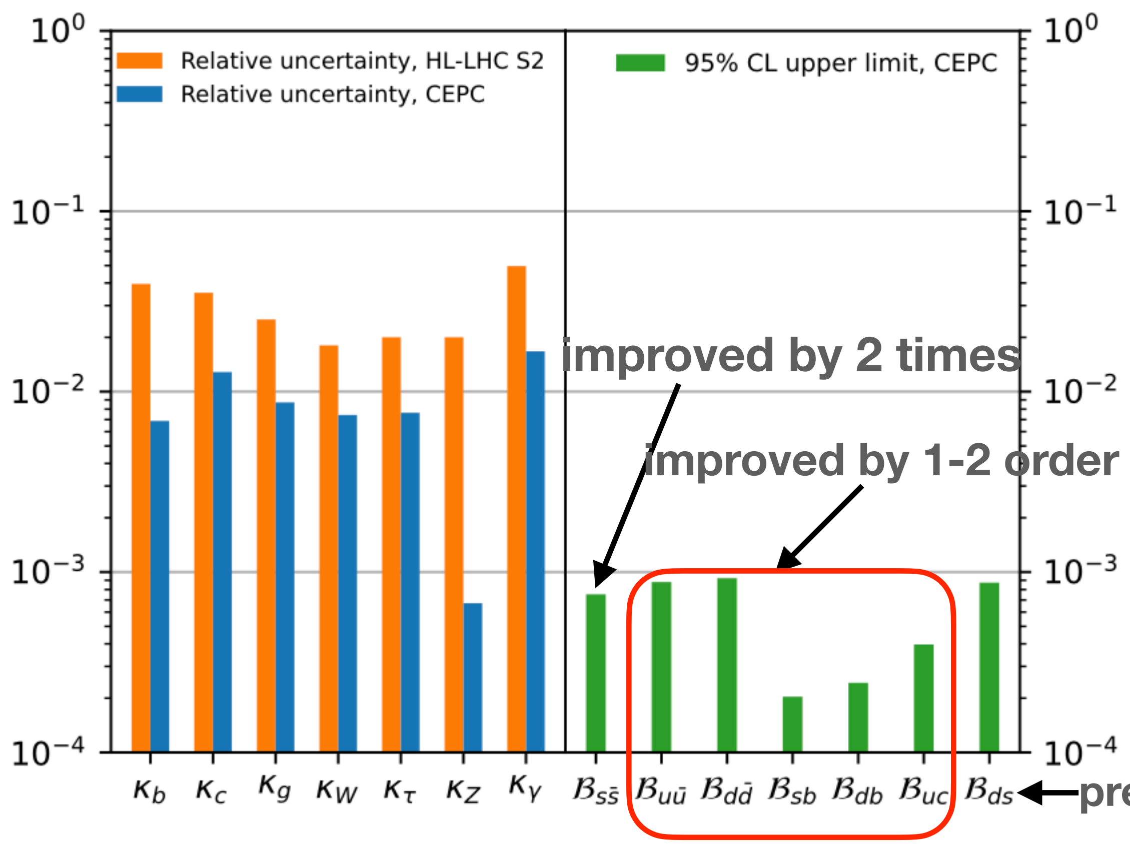


FIG. 5. Expected upper limits on the branching ratios of rare Higgs boson decays from this Letter (green) and the relative uncertainties of Higgs couplings anticipated at CEPC [19] (blue) and HL-LHC [43] (orange) under the kappa-0 fit scenario [54] and scenario S2 of systematics [55], as cited in Ref. [19]. The limit on $B_{s\bar{s}}$ corresponds to an upper limit of 1.7 on the Higgs-strange coupling modifier κ_s (not shown).

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

	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

[23] J. Duarte-Campderros, G. Perez, M. Schlaffer, and A. Soffer, Phys. Rev. D **101**, 115005 (2020), arXiv:1811.09636 [hep-ph].

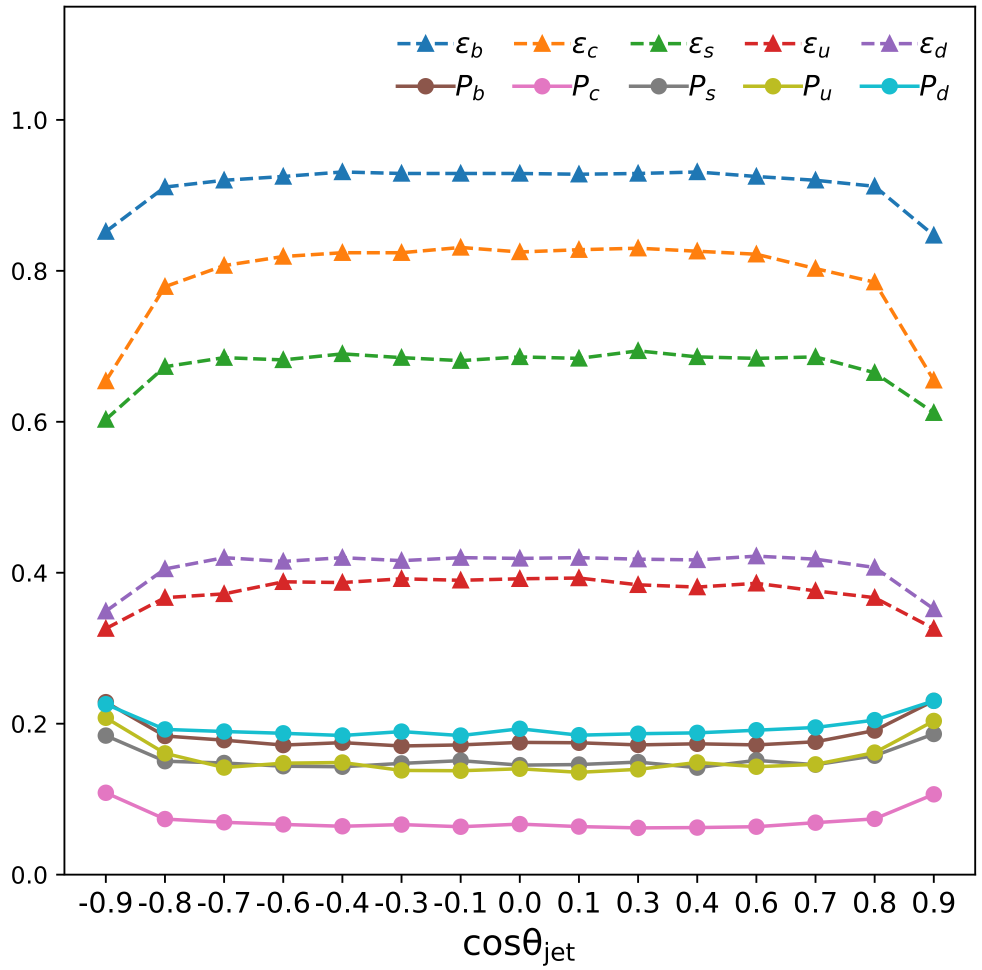
[44] A. Albert *et al.*, "Strange quark as a probe for new physics in the higgs sector," (2022), arXiv:2203.07535 [hep-ex].

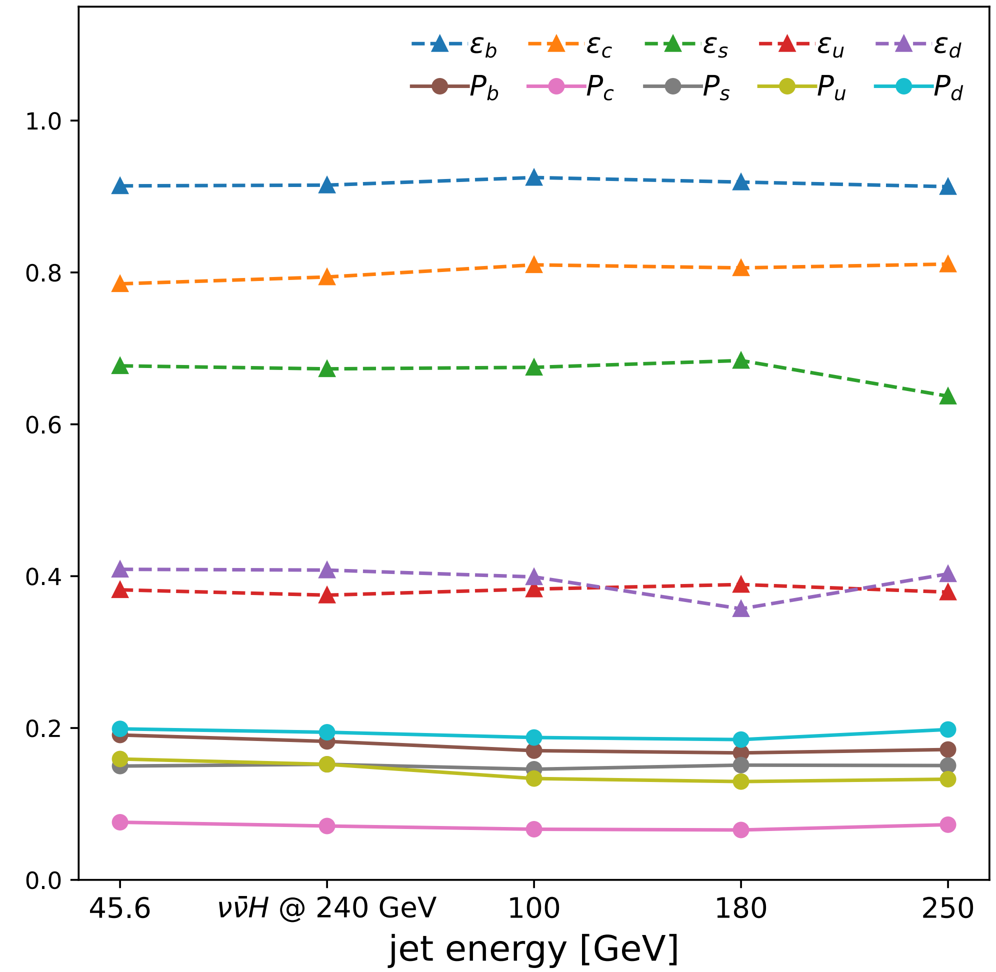
[53] J. de Blas *et al.*, JHEP **01**, 139 (2020), arXiv:1905.03764 [hep-ph].

[54] J. De Blas, G. Durieux, C. Grojean, J. Gu, and A. Paul, JHEP **12**, 117 (2019), arXiv:1907.04311 [hep-ph].

the dependent factors of JOI performance

Dependence on the jet energy, and jet polar angle



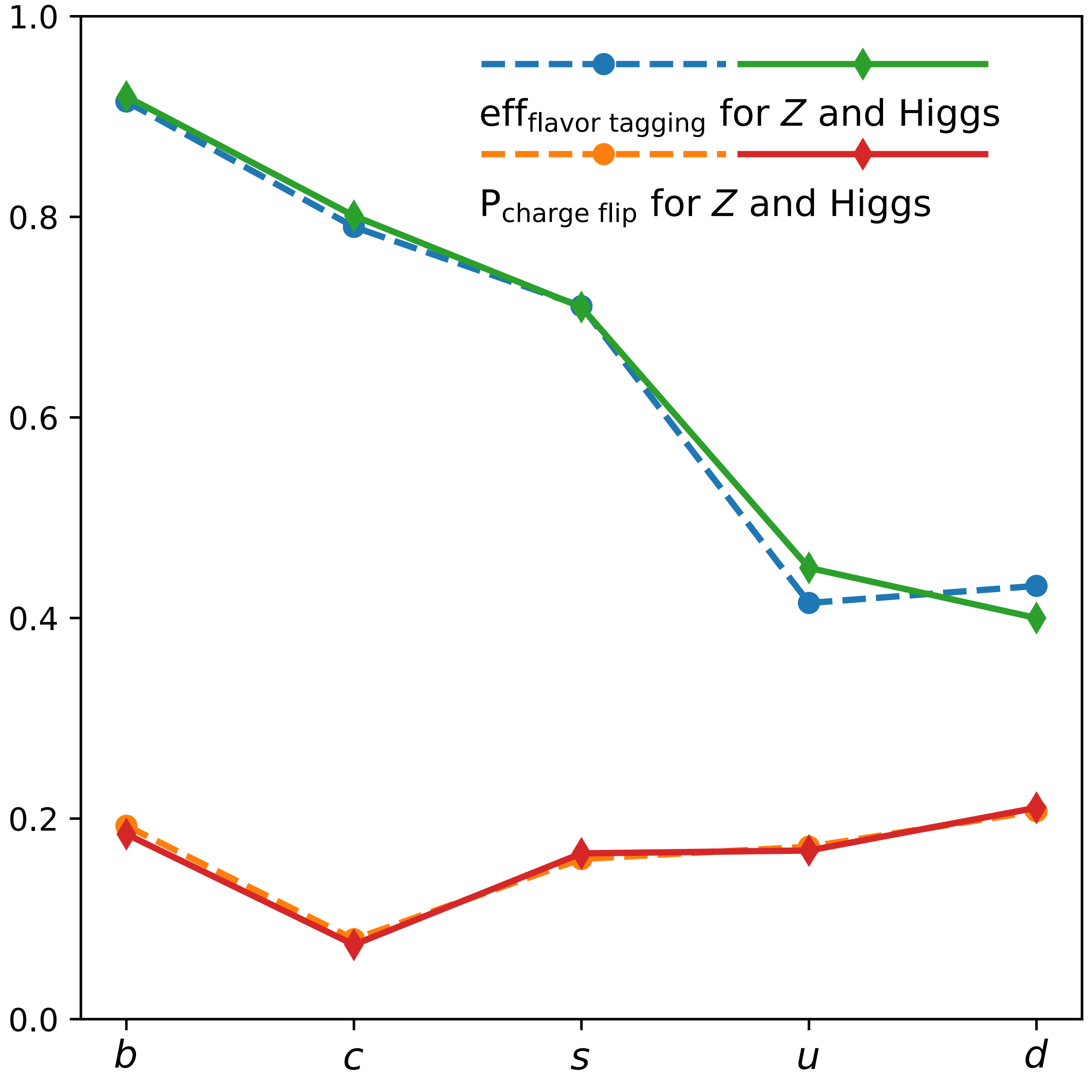


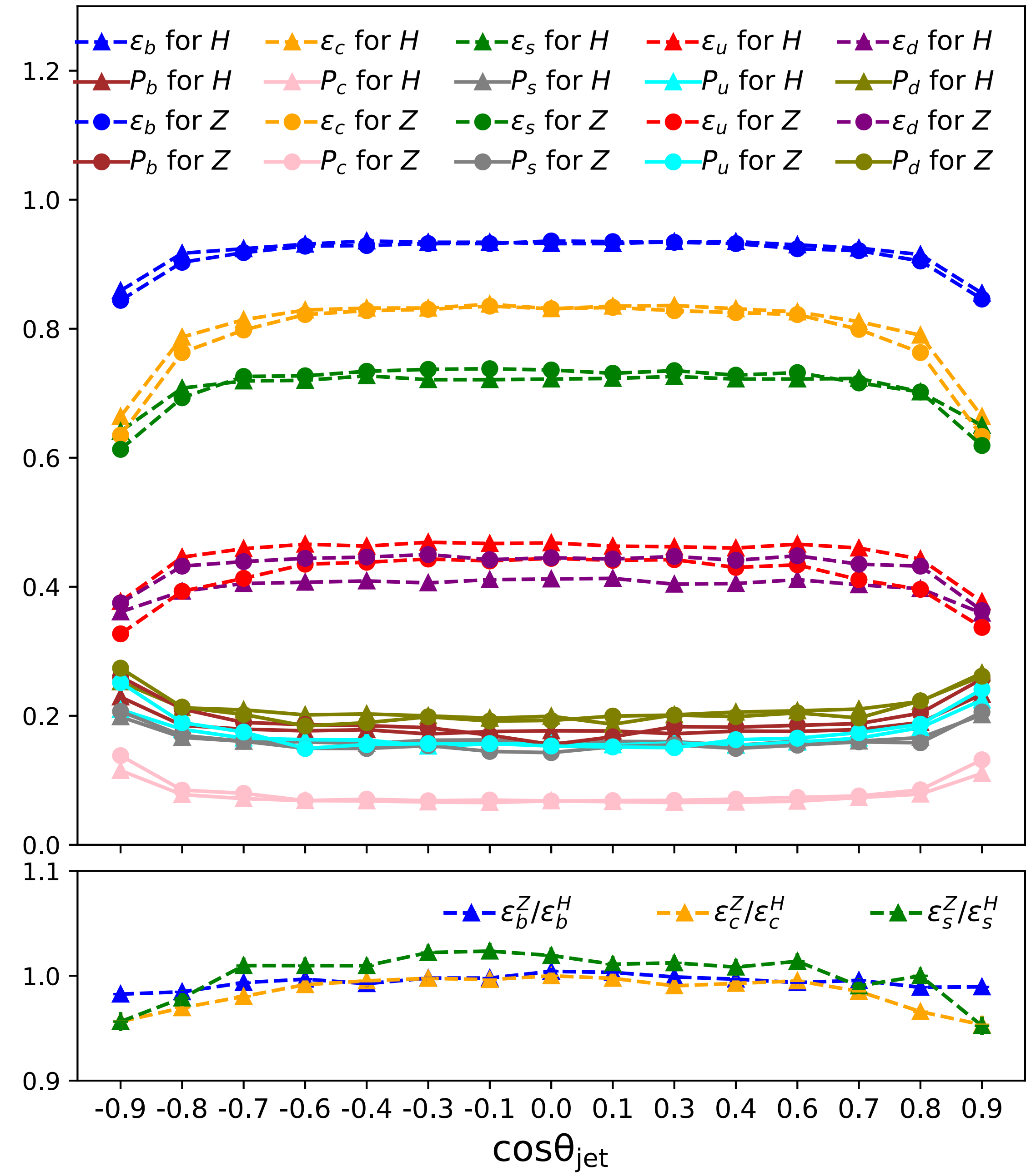
- The jet tagging efficiencies and charge flip rates are better in the barrel region of the detector and exhibit slight degrading in the endcap region.
- The jet tagging efficiencies and charge flip rates are rather stable with various jet energy.

Comparison between different physics processes

$$Z \to q\bar{q} \text{ at } \sqrt{s} = 91.2\ GeV$$

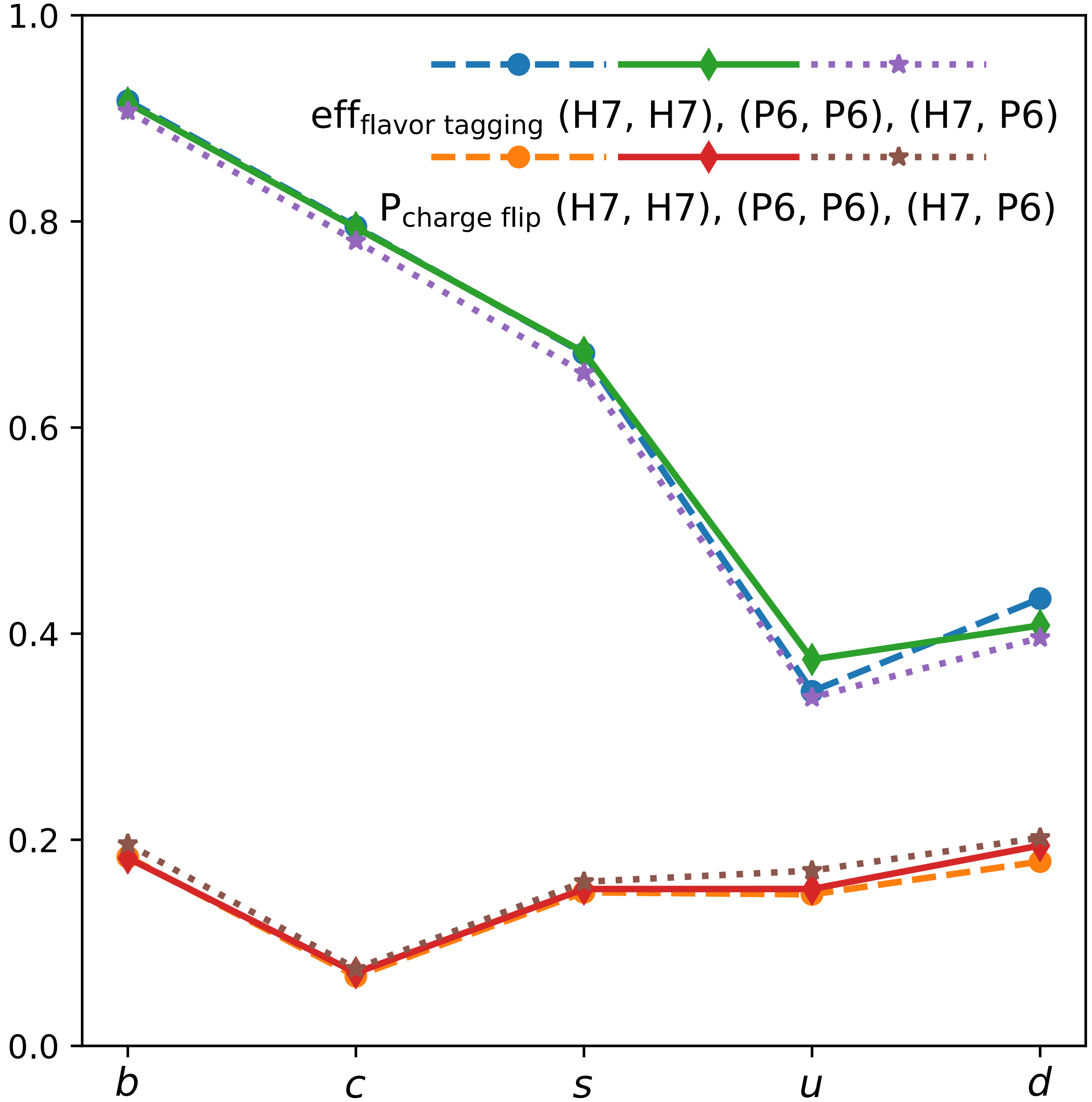
$$\nu\bar{\nu}H, H \to q\bar{q} \text{ at } \sqrt{s} = 240\ GeV$$





The jet origin identification performance agrees with each other, especially in the fiducial barrel region of the detector for the flavor tagging performance of b, c, and s.

Comparison between different hadronization models



Pythia-6.4
Herwig-7.2.2

$\nu\bar{\nu}H, H \to jj$ at $\sqrt{s} = 240\ GeV$

(A, B) means: training on A, test on B

The jet origin identificaion performanc agrees with each other, especially for b, c, and s jets, while exhibits small but visible differences for u and d jets.

Summary

- We propose and realize the concept of jet origin identification that distinguishes b, c, s, u, d, $\bar{b}, \bar{c}, \bar{s}, \bar{u}, \bar{d}$, and gluon-jet.
- Using jet origin identification, we estimate the upper limits on the branching ratios of seven Higgs rare and FCNC hadronic decay modes and achieve significant improvement compared to previous studies.
- We found that the jet origin identification performance, especially for b/c/s-jet, is rather stable versus the jet energy, jet polar angle, different physics processes, and even different hadronization models. The observed stability is vital for applying jet origin identification in real experiments.

Many thanks !

Backup

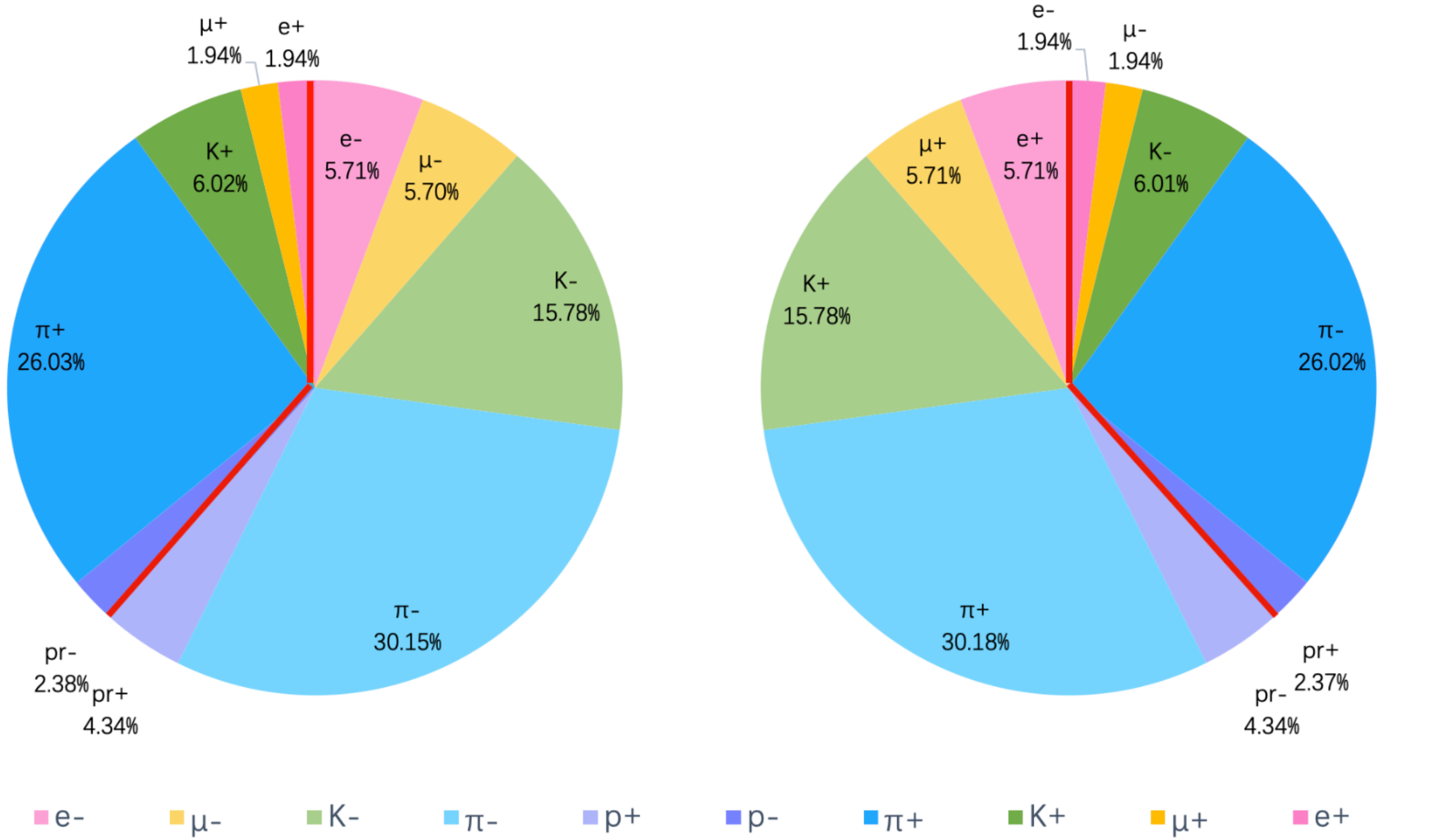
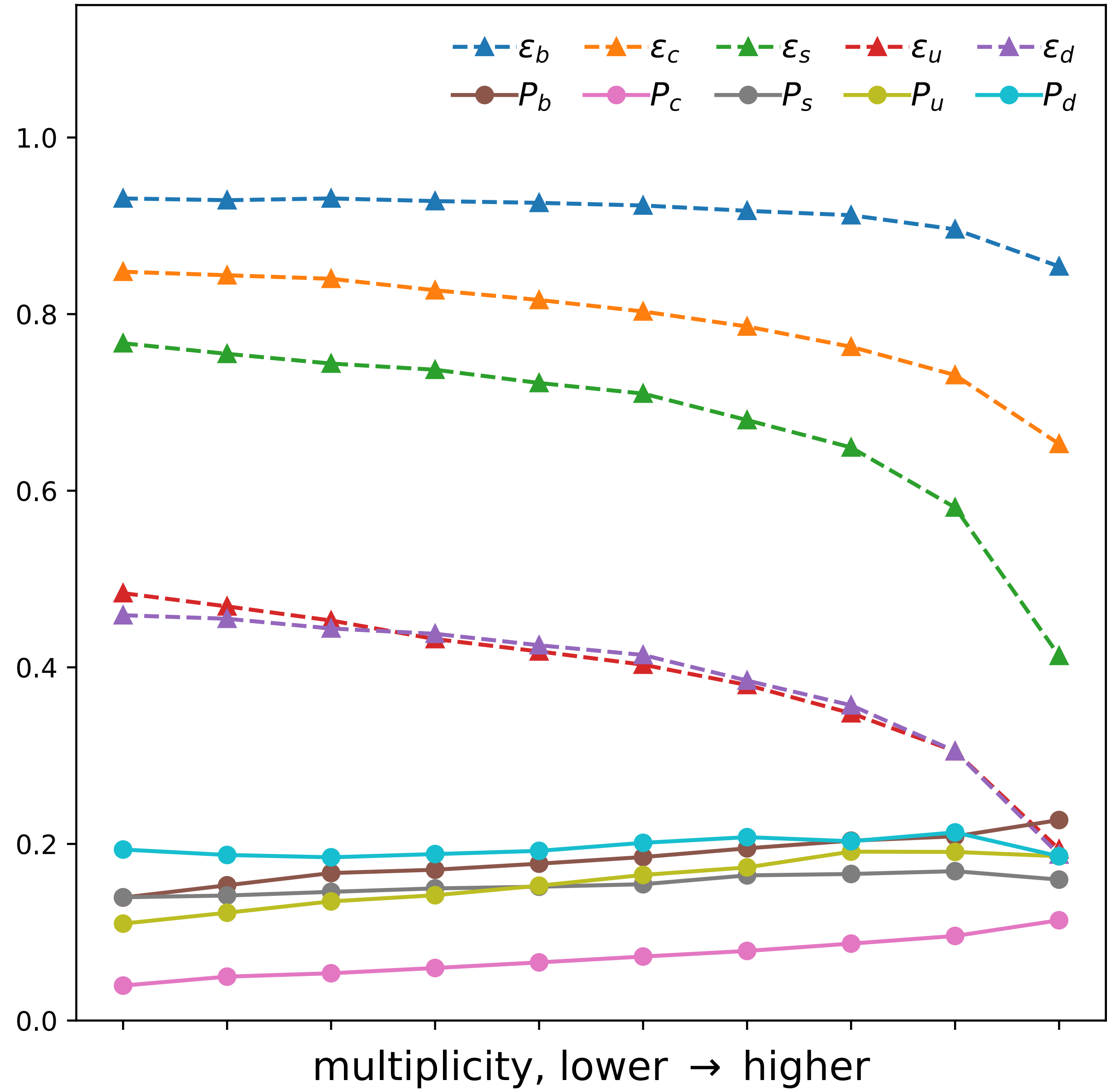


Figure 5. The percentages of species of final state leading charged particles within the b jet (left) and the $\bar{b}$ jet (right) by WHIZARD 1.95.

arXiv:2306.14089

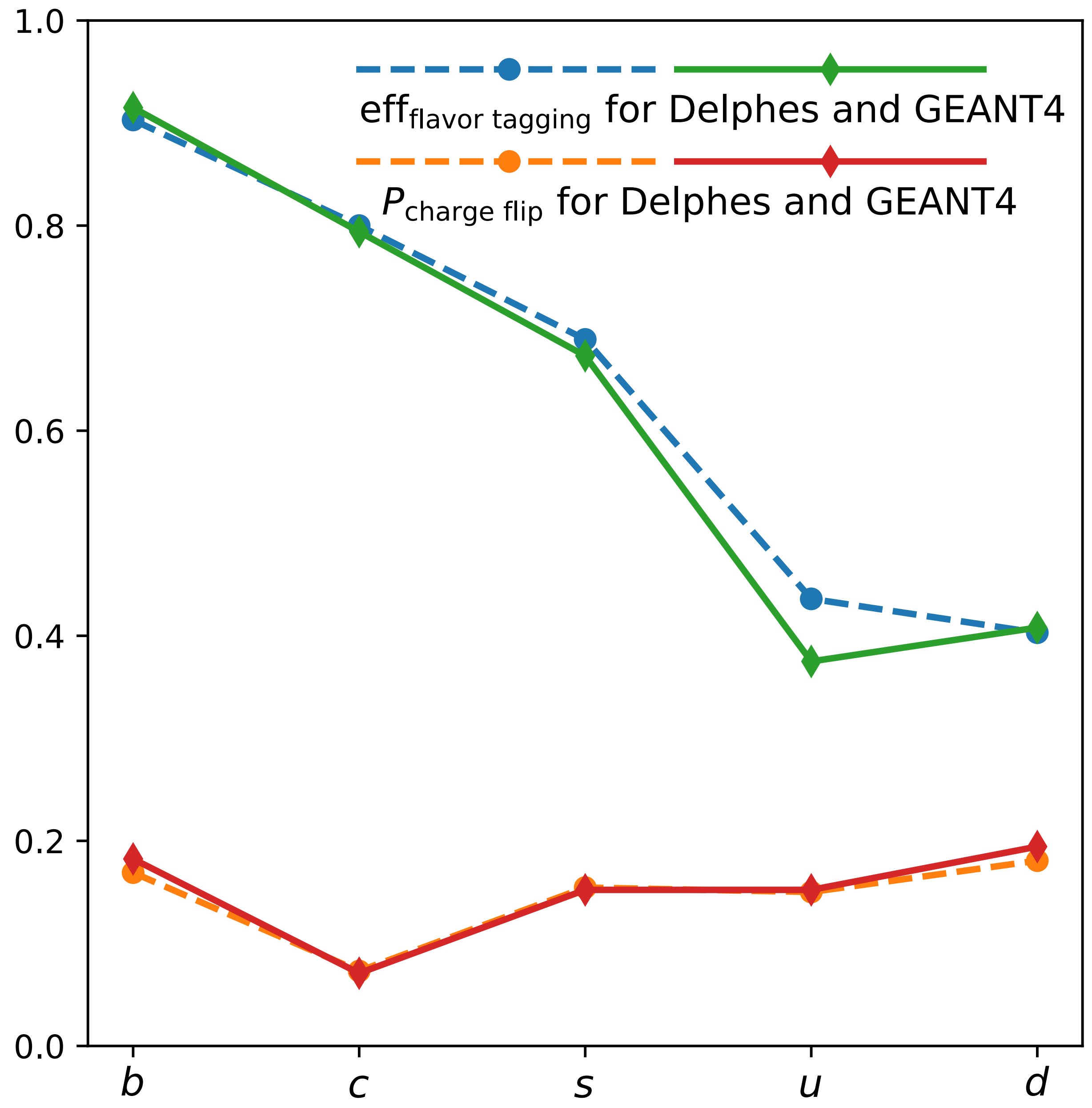
Dependence on the particle multiplicity of the jet



$\nu\bar{\nu}H, H \to jj$ at $\sqrt{s} = 240\ GeV$

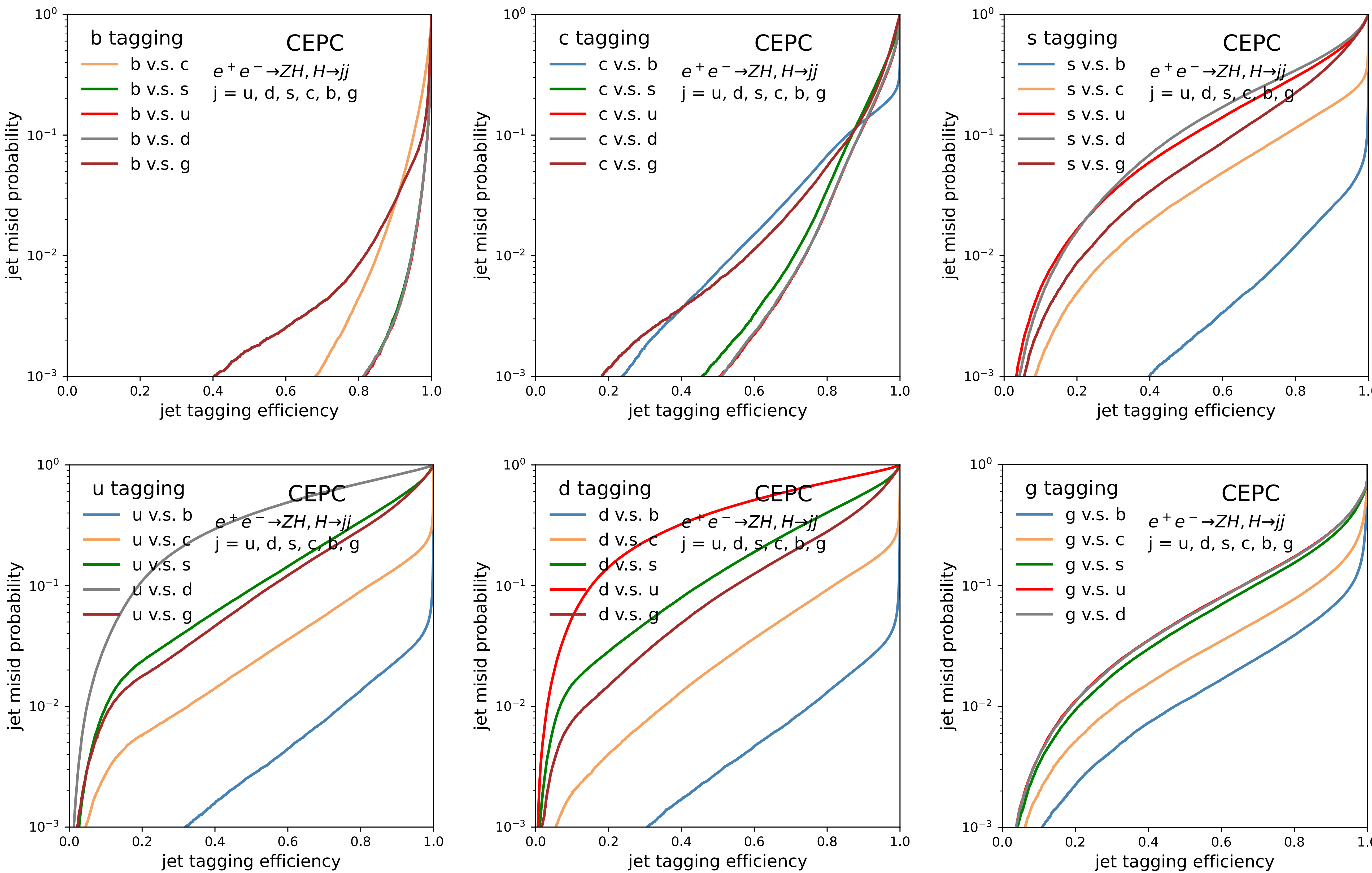
The JOI performance will degrade with increasing particle multiplicity of the jet.

Comparison between full simulation and Delphes

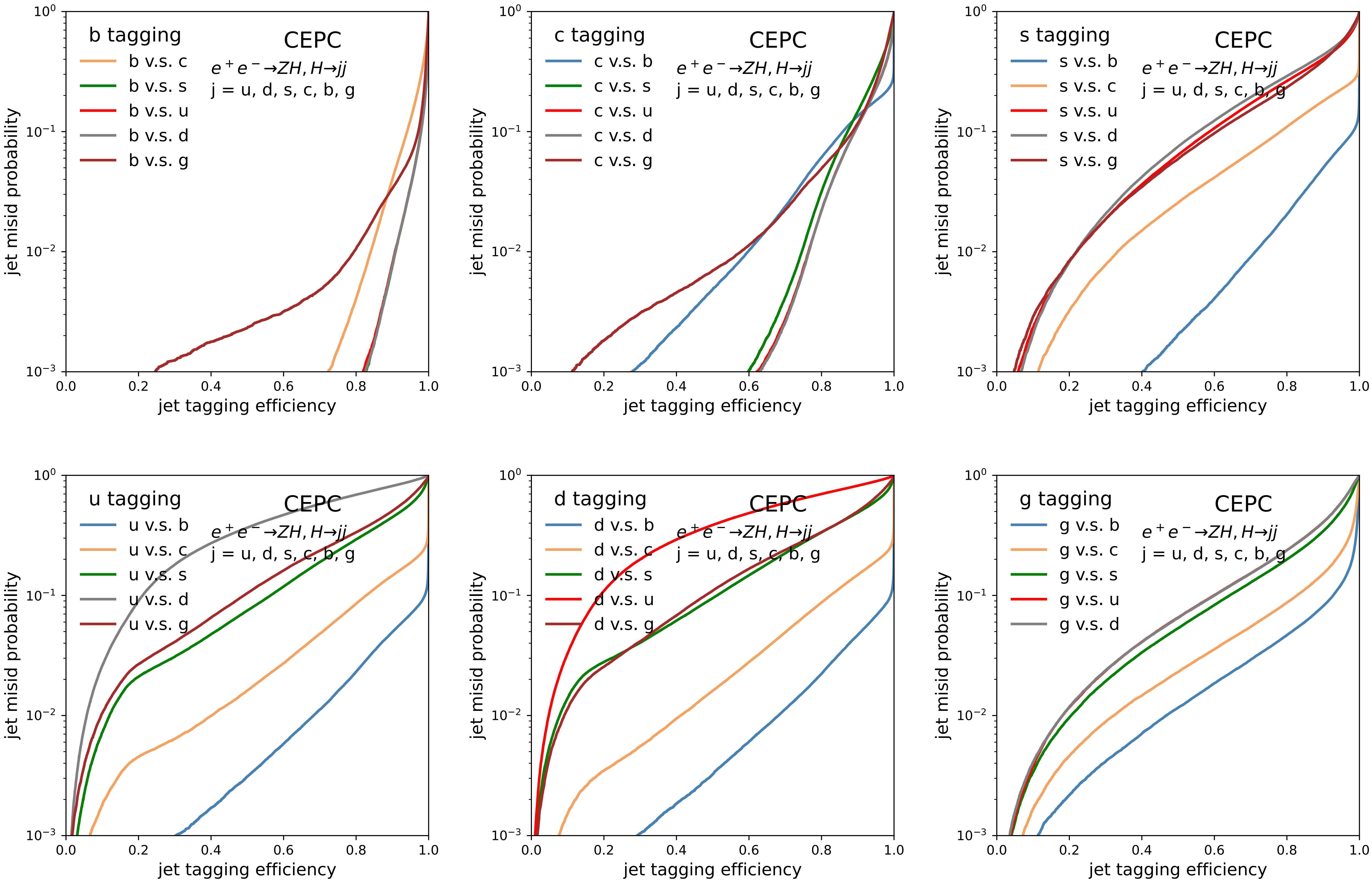


The Delphes fast simulation performance agrees with the full simulation one.

full simulation



Delphes



b tagging
CEPC
$e^+e^- \to ZH, H \to jj$
j = u, d, s, c, b, g
b v.s. c
b v.s. s
b v.s. u
b v.s. d
b v.s. g
c tagging
CEPC
$e^+e^- \to ZH, H \to jj$
j = u, d, s, c, b, g
c v.s. b
c v.s. s
c v.s. u
c v.s. d
c v.s. g
s tagging
CEPC
$e^+e^- \to ZH, H \to jj$
j = u, d, s, c, b, g
s v.s. b
s v.s. c
s v.s. u
s v.s. d
s v.s. g
u tagging
CEPC
$e^+e^- \to ZH, H \to jj$
j = u, d, s, c, b, g
u v.s. b
u v.s. c
u v.s. s
u v.s. d
u v.s. g
d tagging
CEPC
$e^+e^- \to ZH, H \to jj$
j = u, d, s, c, b, g
d v.s. b
d v.s. c
d v.s. s
d v.s. u
d v.s. g
g tagging
CEPC
$e^+e^- \to ZH, H \to jj$
j = u, d, s, c, b, g
g v.s. b
g v.s. c
g v.s. s
g v.s. u
g v.s. d
jet misid probability
jet tagging efficiency
10^0
10^{-1}
10^{-2}
10^{-3}
0.0
0.2
0.4
0.6
0.8
1.0