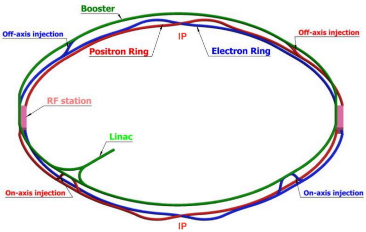
The measurement of Higgs decay (bb/cc/gg, rare, and FCNC) at the CEPC

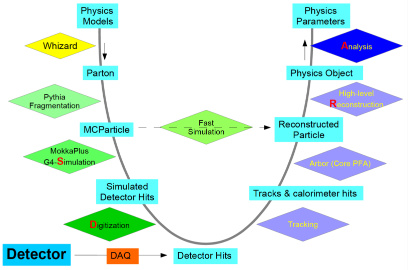
Yongfeng Zhu (PKU), Hao Liang (IHEP), Chen Zhou (PKU), Huilin Qu (CERN), Manqi Ruan (IHEP), etc

Higgs 2023

12.01.2023

CEPC





Operation mode	$\sqrt{s}$ GeV	Running time year	$\int L$ ab^{-1}	Event yields
H	240	7/10	5.6/20	$1 \times 10^6 / 4 \times 10^6$
Z	91.2	2	16/100	$7 \times 10^{11} / 3 \times 10^{12}$
W^-W^+	160	1	2.6/6	$2 \times 10^7 / 1 \times 10^8$
$t\bar{t}$	360	5	1	5×10^5

Motivation

- Measuring the branching fractions of the $Higgs \to b\bar{b}/c\bar{c}/gg$ decays is one of the core CEPC physics objectives.
- The rare and FCNC hadronic decays of the Higgs boson are of great interest to NP.

Contents

- The relative accuracy of signal strength measurement of $\nu\nu H(H \to b\bar{b}, c\bar{c}, gg)$ and $qqH(H \to b\bar{b}, c\bar{c}, gg)$. ($\sqrt{s} = 240$ GeV, $\int L = 5.6ab^{-1}$)
- key performance: flavor tagging performance and color singlet identification
- The upper limits on branching ratio of Higgs rare and FCNC hadronic decays at the CEPC. ($\sqrt{s} = 240$ GeV, $\int L = 20ab^{-1}$)
- summary

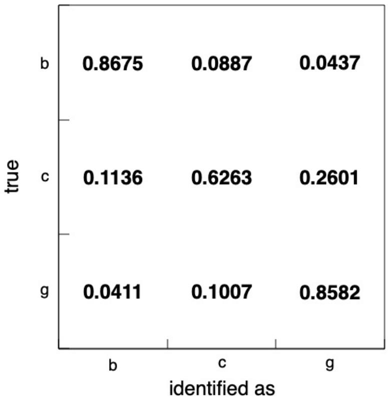
The analysis of $Higgs \to b\bar{b}/c\bar{c}/gg$ in $\nu\nu H$ channel.

$\sqrt{s} = 240$ GeV, $\int L = 5.6ab^{-1}$

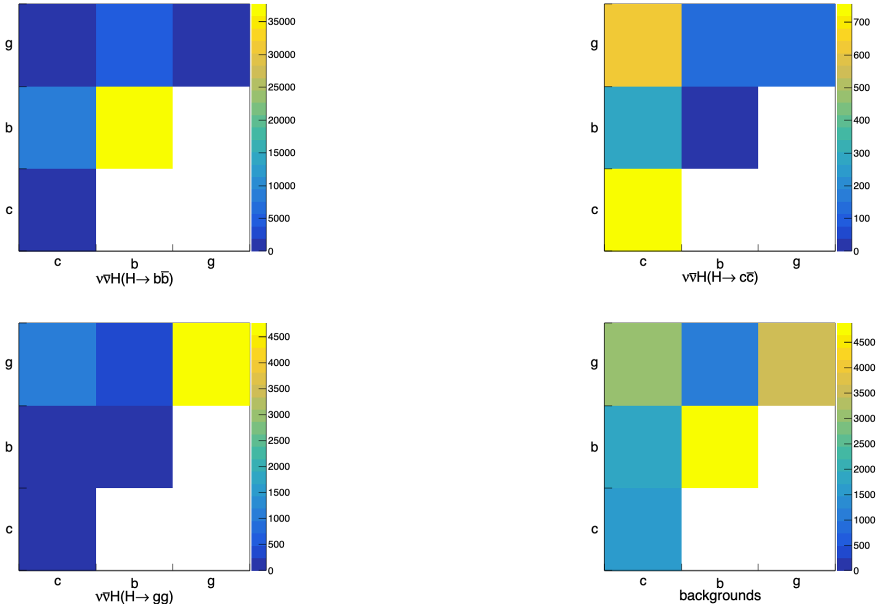
	$\nu\nu Hq\bar{q}/gg$	2f	SW	SZ	WW	ZZ	Mixed	ZH	$\frac{\sqrt{S+B}}{S}$ (%)
total	178890	8.01*E*8	1.95*E*7	9.07*E*6	5.08*E*7	6.39*E*6	2.18*E*7	961606	16.86
recoilMass (GeV) $\in (74, 131)$	157822	5.11*E*7	2.17*E*6	1.38*E*6	4.78*E*6	1.30*E*6	1.08*E*6	74991	4.99
visEn (GeV) $\in (109, 143)$	142918	2.37*E*7	1.35*E*6	8.81*E*5	3.60*E*6	1.03*E*6	6.29*E*5	50989	3.92
leadLepEn (GeV) $\in (0, 42)$	141926	2.08*E*7	3.65*E*5	7.24*E*5	2.81*E*6	9.72*E*5	1.34*E*5	46963	3.59
multiplicity $\in (40, 130)$	139545	1.66*E*7	2.36*E*5	5.24*E*5	2.62*E*6	9.07*E*5	4977	42751	3.29
leadNeuEn (GeV) $\in (0, 41)$	138653	1.46*E*7	2.24*E*5	4.72*E*5	2.49*E*6	8.69*E*5	4552	42303	3.12
Pt (GeV) $\in (20, 60)$	121212	248715	1.56*E*5	2.48*E*5	1.51*E*6	4.31*E*5	999	35453	1.37
Pl (GeV) $\in (0, 50)$	118109	52784	1.05*E*5	74936	7.30*E*5	1.13*E*5	847	34279	0.94
-log10(Y23) $\in (3.375, +\infty)$	96156	40861	26088	60349	2.25*E*5	82560	640	10691	0.76
InvMass (GeV) $\in (116, 134)$	71758	22200	11059	6308	77912	13680	248	6915	0.64
BDT $\in (-0.02, 1)$	60887	9140	266	2521	3761	3916	58	1897	0.47

1. The final state particles of remaining events are clustered into two jets.
2. The b-likeness and c-likeness of two jets were calculated with the LCFIPlus package.
3. The cut on b-likeness and c-likeness can be found to maximize the value of $eff(b \to b) + eff(c \to c) + eff(udsg \to udsg)$.

migration matrix: the elements represent the identification efficiency and misidentification rate of each quark species.



events distribution based on migration matrix :



get the signal strength accuracy by the Log-likelihood function:

$$-2 \cdot log(\ell) = \sum_{i=1}^{i=6} \frac{[S_b \cdot N_{b,i} + S_c \cdot N_{c,i} + S_{light} \cdot N_{light,i} + N_{bkg,i} - N_i]^2}{N_i}$$

- S_b: the signal strength of $\nu\nu Hb\bar{b}$
- $N_{b,i}$: the event number of $\nu\nu Hb\bar{b}$ in *ith* bin
- N_i: the total event number in *i'th* bin of $\nu\nu Hb\bar{b}$, $\nu\nu H/c\bar{c}$, $\nu\nu Hgg$ and backgrounds
- $N_{bkg,i}$ is the expected event number in *ith* bin of backgrounds,
- similar for S_c, S_{light}, $N_{c,i}$, and $N_{light,i}$

$$hessian\ matrix = \begin{bmatrix} \frac{\partial^2 log(\ell)}{\partial S_g \partial S_c} & \frac{\partial^2 log(\ell)}{\partial S_g \partial S_b} & \frac{\partial^2 log(\ell)}{\partial S_g \partial S_g} \\ \frac{\partial^2 log(\ell)}{\partial S_b \partial S_c} & \frac{\partial^2 log(\ell)}{\partial S_b \partial S_b} & \frac{\partial^2 log(\ell)}{\partial S_b \partial S_g} \\ \frac{\partial^2 log(\ell)}{\partial S_c \partial S_c} & \frac{\partial^2 log(\ell)}{\partial S_c \partial S_b} & \frac{\partial^2 log(\ell)}{\partial S_c \partial S_g} \end{bmatrix}$$

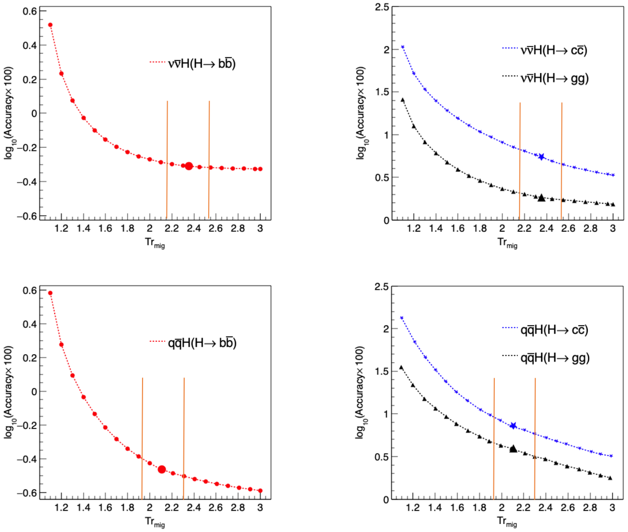
- The error covariance is obtained from the Hessian matrix.
- The relative accuracy of signal strength is the square roots of the diagonal elements of the covariance matrix, it is 0.49%/5.75%/1.82% for $\nu\nu Hb\bar{b}/c\bar{c}/gg$.

The similar method can be applied to $q\bar{q}Hb\bar{b}/c\bar{c}/gg$.

	qqHqq	2f	SW	SZ	WW	ZZ	Mixed	ZH	$\frac{\sqrt{S+B}}{S}$ (%)
total	527488	8.01*E*8	1.95*E*7	9.07*E*6	5.08*E*7	6.39*E*6	2.18*E*7	613008	5.71
multiplicity $\in (27, +\infty)$	527488	3.04*E*8	1.46*E*7	3.37*E*6	4.85*E*7	6.00*E*6	1.81*E*7	577930	3.77
leadLepEn $\in (0, 59)$	527036	2, 98*E*8	6.76*E*6	2.44*E*6	3.93*E*7	5.40*E*6	1.79*E*7	531411	3.65
visEn $\in (199, 278)$	510731	1.21*E*8	1.29*E*6	551105	2.14*E*7	3.06*E*6	1.71*E*7	180571	2.52
leadNeuEn $\in (0, 57)$	509623	5.68*E*7	716161	168030	2.04*E*7	2.93*E*6	1.65*E*7	176387	1.94
thrust $\in (0, 0.86)$	460535	7.81*E*6	473732	132126	1.88*E*7	2.60*E*6	1.54*E*7	167863	1.47
$-log(Y_{34})$ $\in (0, 5.8875)$	451468	4.90*E*6	181432	119836	1.74*E*7	2.40*E*6	1.45*E*7	165961	1.40
HiggsJetsA $\in (2.18, 2\pi)$	326207	2.83*E*6	110156	58613	4.54*E*6	870276	3.74*E*6	96560	1.08
ZJetsA $\in (1.97, 2\pi)$	279030	1.37*E*6	33491	37101	2.39*E*6	496611	2.00*E*6	74005	0.93
ZHiggsA $\in (2.32, 2\pi)$	274530	1.32*E*6	17026	33847	2.28*E*6	468340	1.91*E*6	69620	0.92
circle	268271	1.20E6	10193	31567	2.13*E*6	424514	1.79E6	65434	0.90
BDT $\in (0.02, 1)$	192278	378300	40	307	271436	141446	244126	30022	0.57

With log-likelihood function,
$-2 \cdot log(\ell) = \sum_{i=1}^{i=36} \frac{[S_b \cdot N_{b,i} + S_c \cdot N_{c,i} + S_{light} \cdot N_{light,i} + N_{bkg,i} - N_i]^2}{N_i}$, the signal strength accuracy is 0.35%/7.74%/3.96% for $q\bar{q}Hb\bar{b}/c\bar{c}/gg$.

key performance: flavor tagging



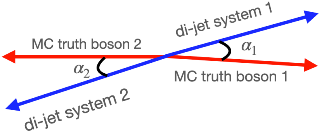
The perfect flavor tagging performance can improve the $\nu\bar{\nu}H(H \to b\bar{b}/c\bar{c}/gg)$ signal strength accuracy by 2%/63%/13%

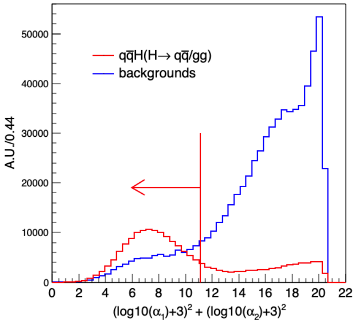
and

$q\bar{q}H(H \to b\bar{b}/c\bar{c}/gg)$ signal strength accuracy by 35%/122%/181%.

Tr_{mig} : the trace of migration matrix

key performance: Color Singlet Identification (CSI)



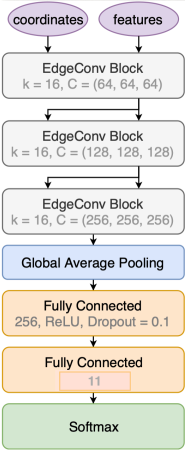


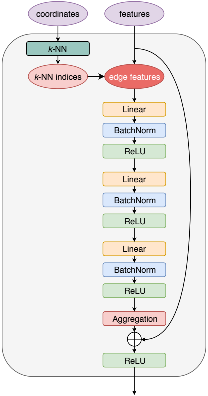
	$qqHb\bar{b}$	$qqHc\bar{c}$	$qqHgg$
w.o. alpha cut	0.35%	7.74%	3.96%
w.i. alpha cut	0.33%	4.37%	2.08%
improved by	6%	77%	90%

note: The CSI in this report is just a demonstrator to illustrate the importance of CSI, we need to construct a CSI evaluator at the reconstruction level in the future.

The rare and FCNC hadronic decays of the Higgs.

ParticleNet (arXiv:1902.08570): Graph Neural Network based flavor tagging algorithm





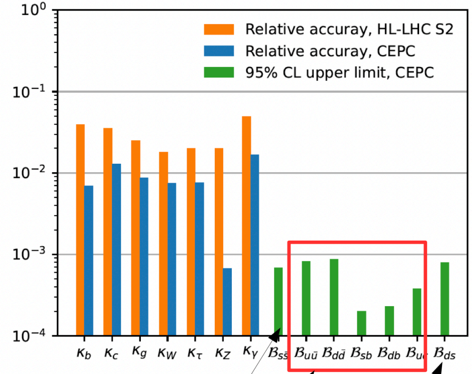
Jet origin identification:

True \ Predicted	b	$\bar{b}$	c	$\bar{c}$	s	$\bar{s}$	u	$\bar{u}$	d	$\bar{d}$	G
b	0.745	0.163	0.033	0.025	0.004	0.003	0.002	0.003	0.002	0.002	0.017
$\bar{b}$	0.170	0.737	0.026	0.033	0.003	0.004	0.003	0.002	0.002	0.003	0.018
c	0.015	0.014	0.743	0.055	0.036	0.031	0.025	0.009	0.009	0.018	0.043
$\bar{c}$	0.016	0.015	0.056	0.739	0.032	0.037	0.009	0.026	0.017	0.010	0.043
s	0.003	0.002	0.020	0.018	0.543	0.102	0.030	0.080	0.063	0.045	0.092
$\bar{s}$	0.003	0.003	0.018	0.020	0.102	0.542	0.084	0.028	0.045	0.062	0.094
u	0.002	0.003	0.020	0.011	0.044	0.131	0.367	0.055	0.080	0.174	0.111
$\bar{u}$	0.003	0.003	0.011	0.019	0.132	0.043	0.062	0.356	0.178	0.081	0.111
d	0.003	0.003	0.012	0.019	0.112	0.092	0.082	0.207	0.277	0.079	0.112
$\bar{d}$	0.003	0.003	0.020	0.012	0.092	0.112	0.219	0.076	0.079	0.272	0.113
G	0.015	0.014	0.024	0.024	0.052	0.052	0.043	0.041	0.034	0.034	0.667

b/c/s/g tagging efficiency can be higher than 91%/80%/64%/67%

arxiv: 2310.03440

$\sqrt{s} = 240$ GeV, $\int L = 20ab^{-1}$



Improved by ~3 times

Improved by 1-2 orders of magnitudes

Presumably... firstly quantified

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

	Bkg. (10^3)			Upper limit (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

[28] J. Duarte-Campderros, G. Perez, M. Schlaffer, and A. Soffer. Probing the Higgs–strange-quark coupling at e^+e^- colliders using light-jet flavor tagging. *Phys. Rev. D*, 101(11):115005, 2020.

[50] Alexander Albert et al. Strange quark as a probe for new physics in the Higgs sector. In *Snowmass 2021*, 3 2022.

[59] J. de Blas et al. Higgs Boson Studies at Future Particle Colliders. *JHEP*, 01:139, 2020.

[60] Jorge De Blas, Gauthier Durieux, Christophe Grojean, Jiayin Gu, and Ayan Paul. On the future of Higgs, electroweak and diboson measurements at lepton colliders. *JHEP*, 12:117, 2019.

Summary :

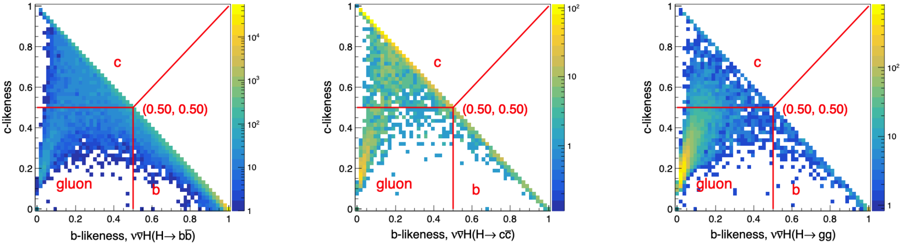
- Combine all four different channels of $\mu\mu H$, eeH (from Chin. Phys. C, 44(1):013001, 2020), $\nu\nu H$ and qqH, the total signal strength of $H \to b\bar{b}, c\bar{c}, gg$ can be measured to a relative accuracy of 0.27%/4.03%/1.56%. JHEP, 11:100, 2022

Z decay mode	$H \to b\bar{b}$	$H \to c\bar{c}$	$H \to gg$
$Z \to e^+e^-$	1.57%	14.43%	10.31%
$Z \to \mu^+\mu^-$	1.06%	10.16%	5.23%
$Z \to q\bar{q}$	0.35%	7.74%	3.96%
$Z \to \nu\bar{\nu}$	0.49%	5.35%	1.77%
combination	0.27%	4.03%	1.56%

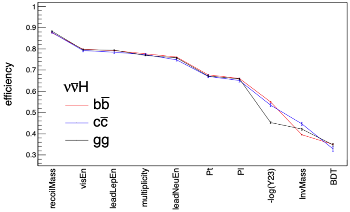
- The flavor tagging and color singlet identification (CSI) are the critical performances for these benchmarks.
- With jet origin identification, the branching ratio upper limit for $H \to s\bar{s}$ ($H \to u\bar{u}/d\bar{d}$) improved by a factor of 3 (∼10) compared to previous studies.

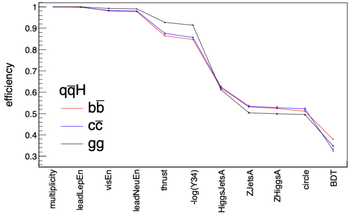
Many thanks !

Backup


c-likeness
b-likeness, $\nu\bar{\nu}H(H\to b\bar{b})$
b-likeness, $\nu\bar{\nu}H(H\to c\bar{c})$
b-likeness, $\nu\bar{\nu}H(H\to gg)$
c
gluon
b
(0.50, 0.50)

systematic uncertainty





systematic uncertainties are categorized into three groups:

- The first group are those that are significantly smaller than the statistical uncertainties, including the reconstructed energy/momentum scale of the physics objects.
- The second group are those comparable to the statistical uncertainty, especially the integrated luminosity.
- The third group are those that can be significantly larger than the statistical uncertainty, including CSI and the jet configuration.

$$M_{mig} = \frac{Tr_{mig} - Tr_{opt}}{Tr_I - Tr_{opt}} \cdot (M_I - M_{opt}) + M_{opt}$$

$$M_{mig} = \frac{Tr_{mig} - Tr_{opt}}{Tr_{1/3} - Tr_{opt}} \cdot (M_{1/3} - M_{opt}) + M_{opt}$$

true \ identified as	b	c	g
b	1	0	0
c	0	1	0
g	0	0	1

perfect

true \ identified as	b	c	g
b	1/3	1/3	1/3
c	1/3	1/3	1/3
g	1/3	1/3	1/3

none