Impact of flavor tagging & CSI on $H \to b\bar{b}/c\bar{c}/gg$ signal strength accuracy measurement

Yongfeng Zhu

advisor : Manqi Ruan

CEPC Workshop, 2022.10.28

Motivation :

- The measurement of $H \to b\bar{b}/c\bar{c}/gg$ signal strength is improtant for Higgs coupling studies.
- In the measurement of the relative accuracy of $H \to b\bar{b}/c\bar{c}/gg$ signal strength, the flavor tagging performance and color singlet identification have significant impact on the final results.
- The vertex detector design has impact on flavor tagging performance.

The relative accuracy of $H \to b\bar{b}/c\bar{c}/gg$ signal strength.

CEPC baseline detector,
all SM samples corresponding to 240 GeV, 5.6 ab^{-1},
full simulation,
$\ell^+\ell^- Higgs$,
$\nu\bar{\nu}Higgs$,
$q\bar{q}Higgs$,

The analysis procedure can be divided into two steps :
first step : select signal with cut flow
second step : calculate the signal strength accuracy

arXiv : 2203.01469

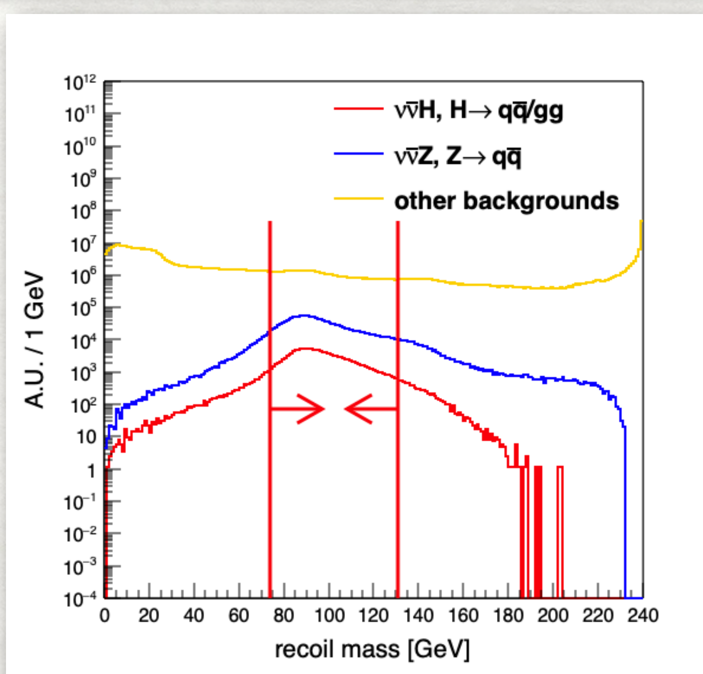
$\nu\bar{\nu}Higgs$ selection

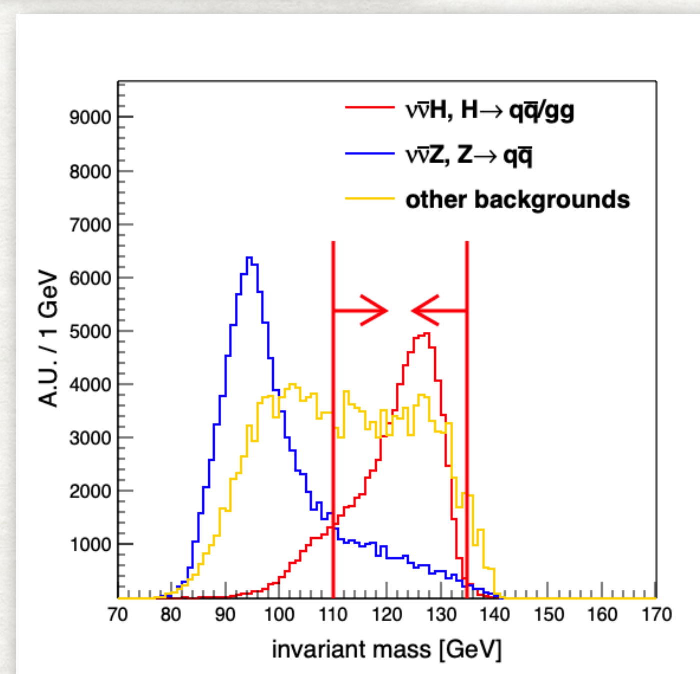
	$\nu\bar{\nu}Hq\bar{q}/gg$	2f	SW	SZ	WW	ZZ	Mixed	ZH	$\frac{\sqrt{S+B}}{S}(\%)$
total	178890	$8.01E8$	$1.95E7$	$9.07E6$	$5.08E7$	$6.39E6$	$2.18E7$	961606	16.86
recoilMass (GeV) $\in (74, 131)$	157822	$5.11E7$	$2.17E6$	$1.38E6$	$4.78E6$	$1.30E6$	$1.08E6$	74991	4.99
$visEn$ (GeV) $\in (109, 143)$	142918	$2.37E7$	$1.35E6$	$8.81E5$	$3.60E6$	$1.03E6$	$6.29E5$	50989	3.92
$leadLepEn$ (GeV) $\in (0, 42)$	141926	$2.08E7$	$3.65E5$	$7.24E5$	$2.81E6$	$9.72E5$	$1.34E5$	46963	3.59
$multiplicity$ $\in (40, 130)$	139545	$1.66E7$	$2.36E5$	$5.24E5$	$2.62E6$	$9.07E5$	4977	42751	3.29
$leadNeuEn$ (GeV) $\in (0, 41)$	138653	$1.46E7$	$2.24E5$	$4.72E5$	$2.49E6$	$8.69E5$	4552	42303	3.12
Pt (GeV) $\in (20, 60)$	121212	248715	$1.56E5$	$2.48E5$	$1.51E6$	$4.31E5$	999	35453	1.37
Pl (GeV) $\in (0, 50)$	118109	52784	$1.05E5$	74936	$7.30E5$	$1.13E5$	847	34279	0.94
-log10(Y23) $\in (3.375, +\infty)$	96156	40861	26088	60349	$2.25E5$	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

Table 1. The event selection of $\nu\bar{\nu}H(H \to q\bar{q}/gg)$.

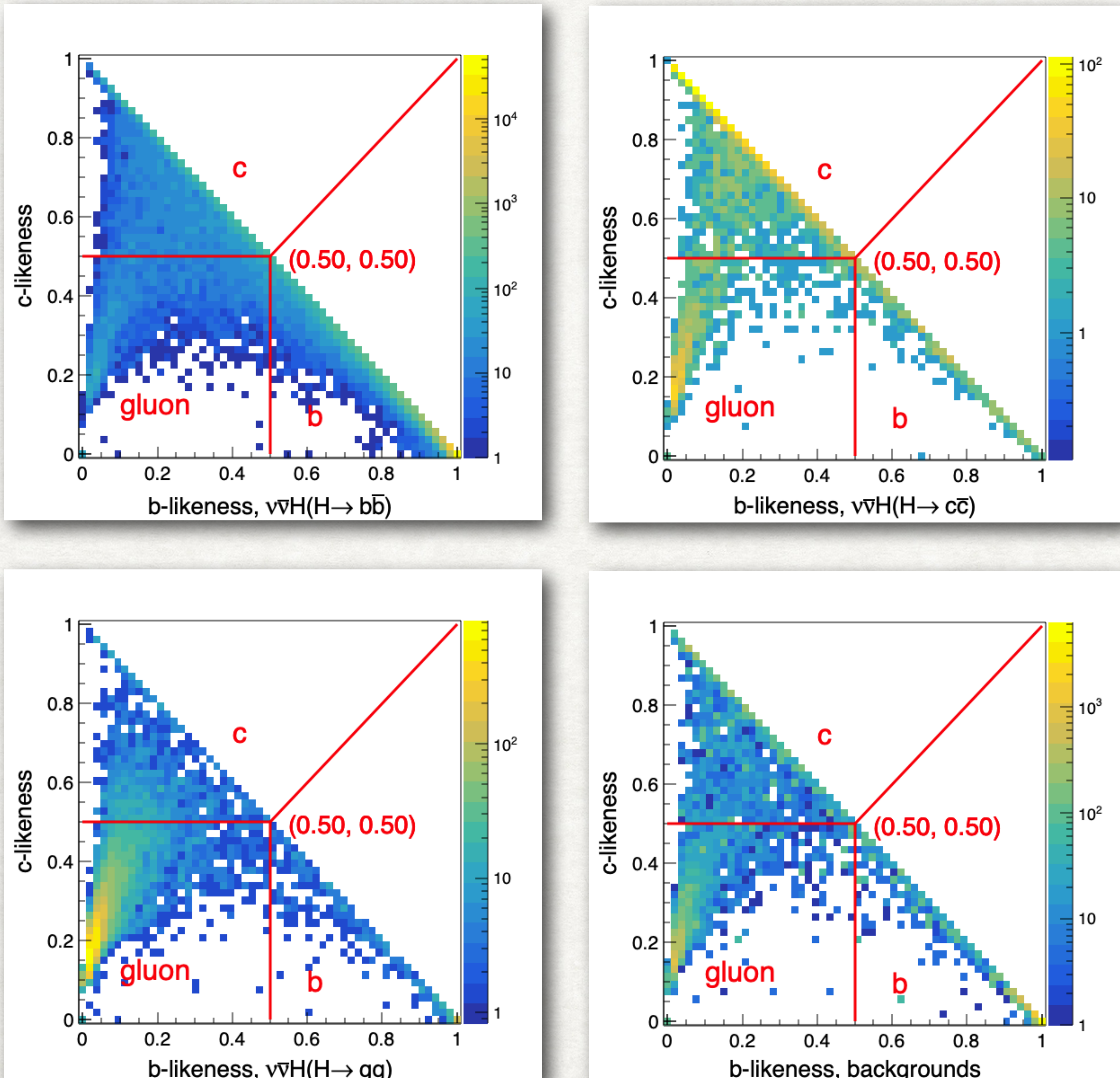
signal-to-background ratio

0.22
0.52
2.91



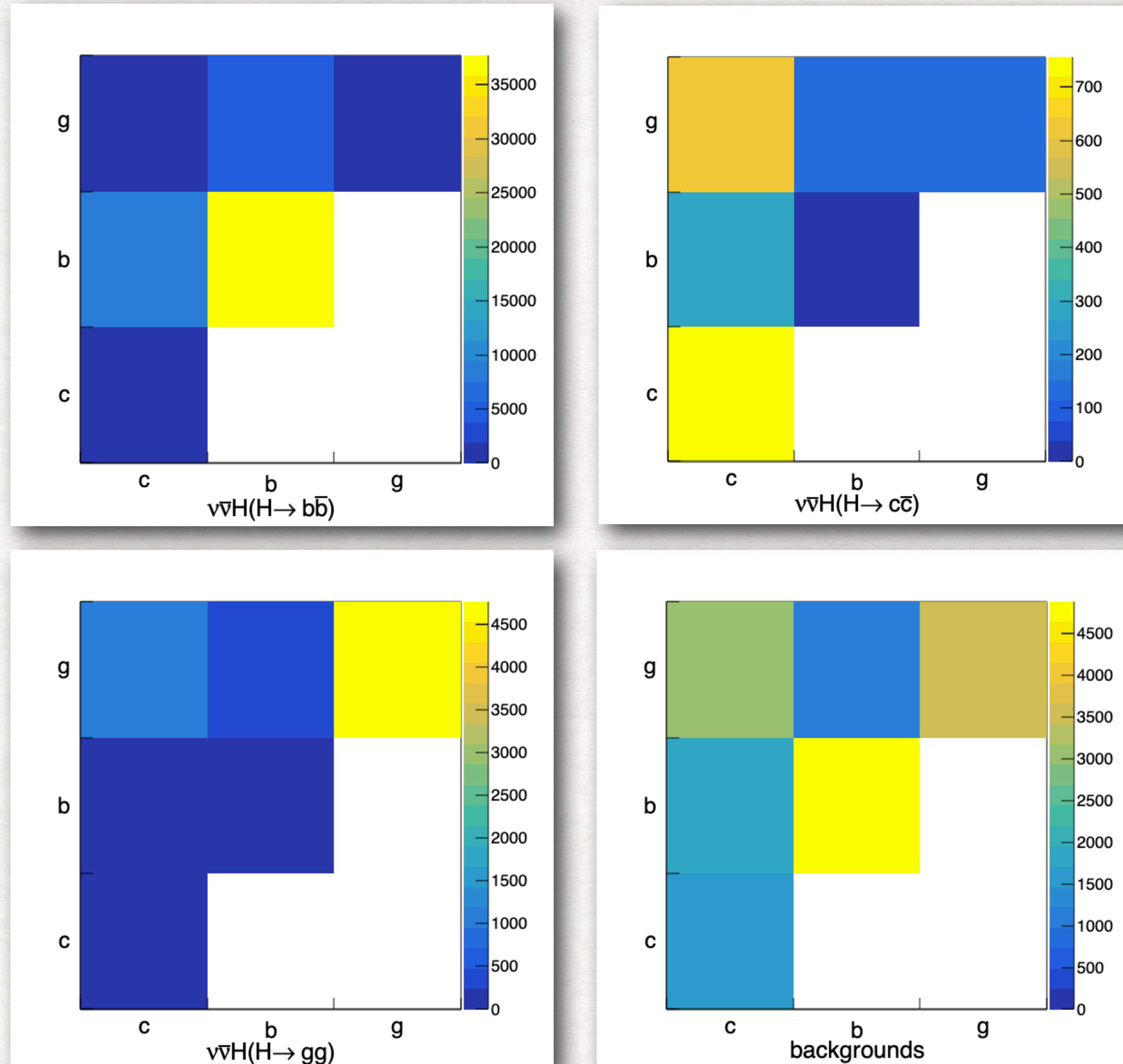


the second step



true \ identified as	b	c	g
b	0.8675	0.0887	0.0437
c	0.1136	0.6263	0.2601
g	0.0411	0.1007	0.8582

The identified flavor combinations based on flavor tagging performance matrix.



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

The signal strength accuracy of $\nu\bar{\nu}H \rightarrow b\bar{b}/c\bar{c}/gg$ be calculated as 0.49%/5.75%/1.82%.

results :

arXiv:1905.12903

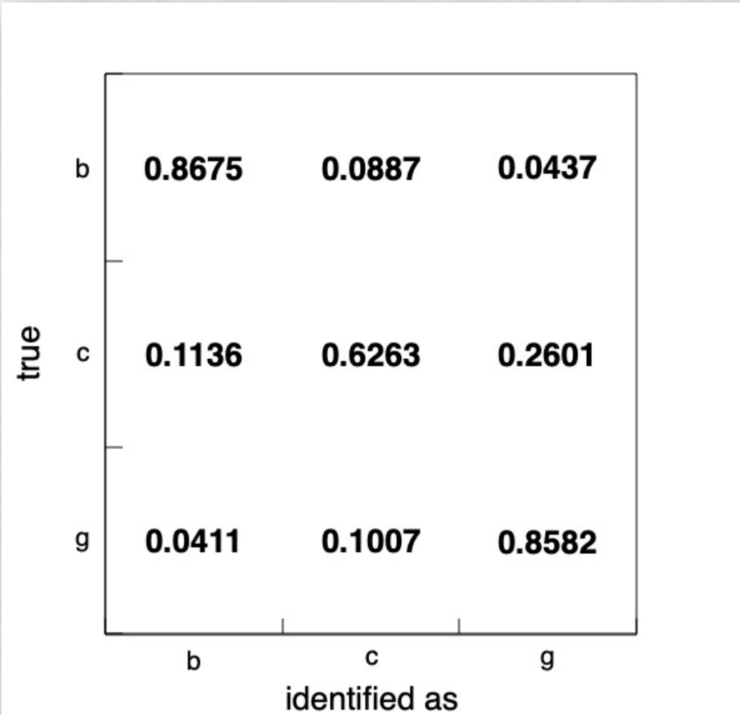
Z decay mode	$H \to b\bar{b}$	$H \to c\bar{c}$	$H \to gg$
$Z \to e^+e^-$	1.57%	14.43%	10.31%
$Z \to \mu^+\mu^-$	1.06%	10.16%	5.23%
$Z \to q\bar{q}$	0.35%	7.74%	3.96%
$Z \to \nu\bar{\nu}$	0.49%	5.75%	1.82%
combination	0.27%	4.03%	1.56%

Table 3. The signal strength accuracies for different channels.

the dependency of $H \to b\bar{b}/c\bar{c}/gg$ accuracy

on

flavor tagging performance



$\nu\bar{\nu}H$

Obtain different flavor tagging performance matrices.

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

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

M_I : perfect flavor tagging matrix

Tr_I : the trace of perfect flavor tagging matrix

M_{opt} : optimized flavor tagging matrix

Tr_{opt} : the trace of optimized flavor tagging matrix

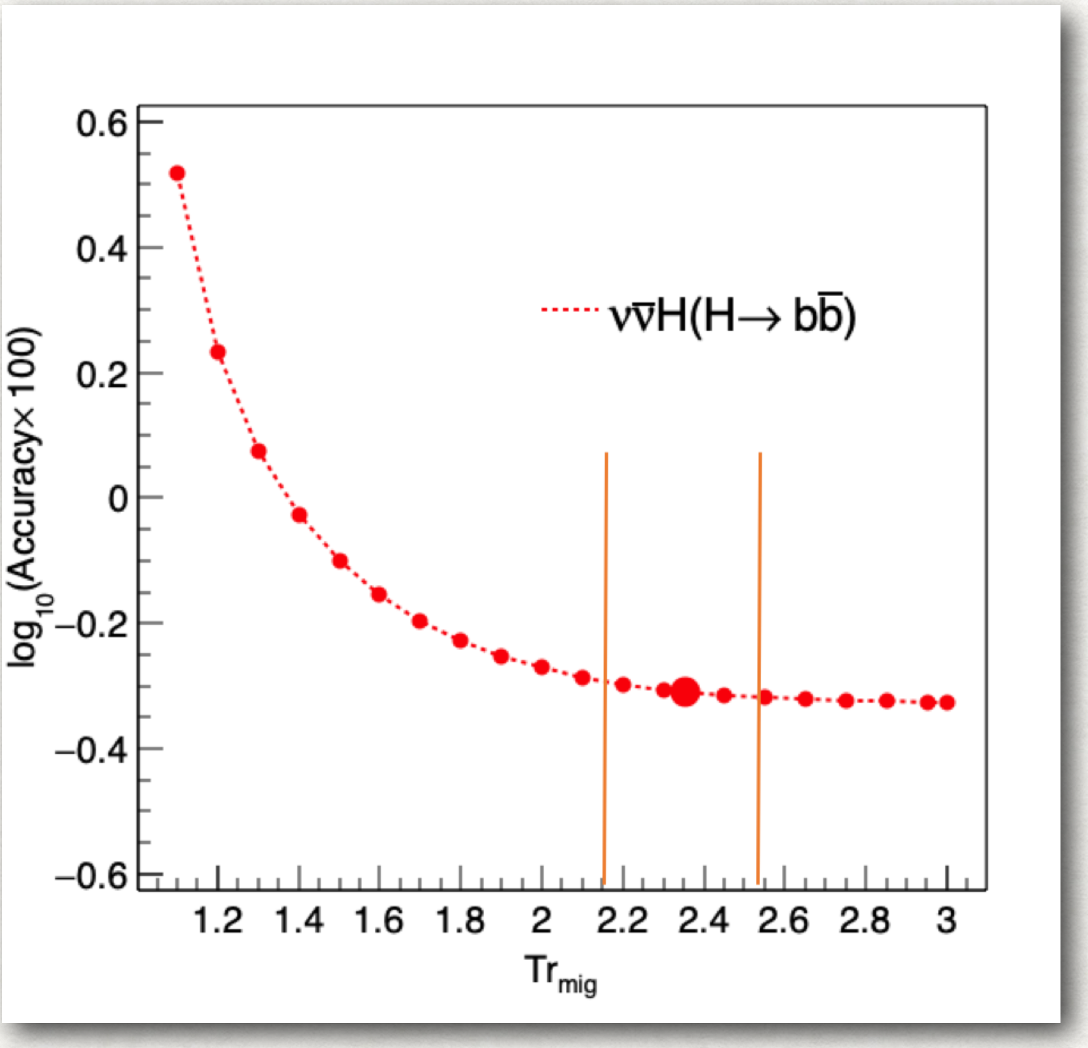
true \ identified as	b	c	g
b	1	0	0
c	0	1	0
g	0	0	1

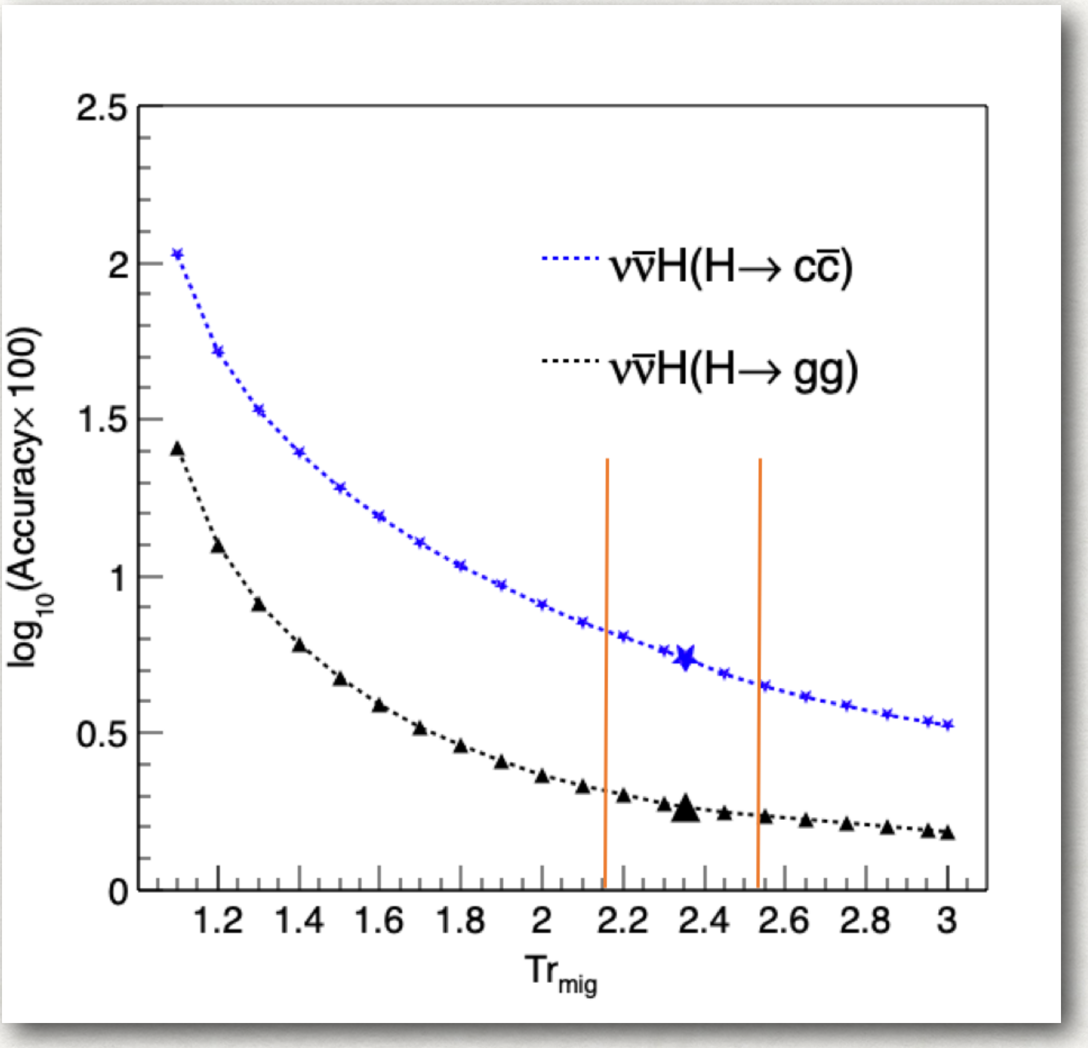
identified as
perfect

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

identified as
none

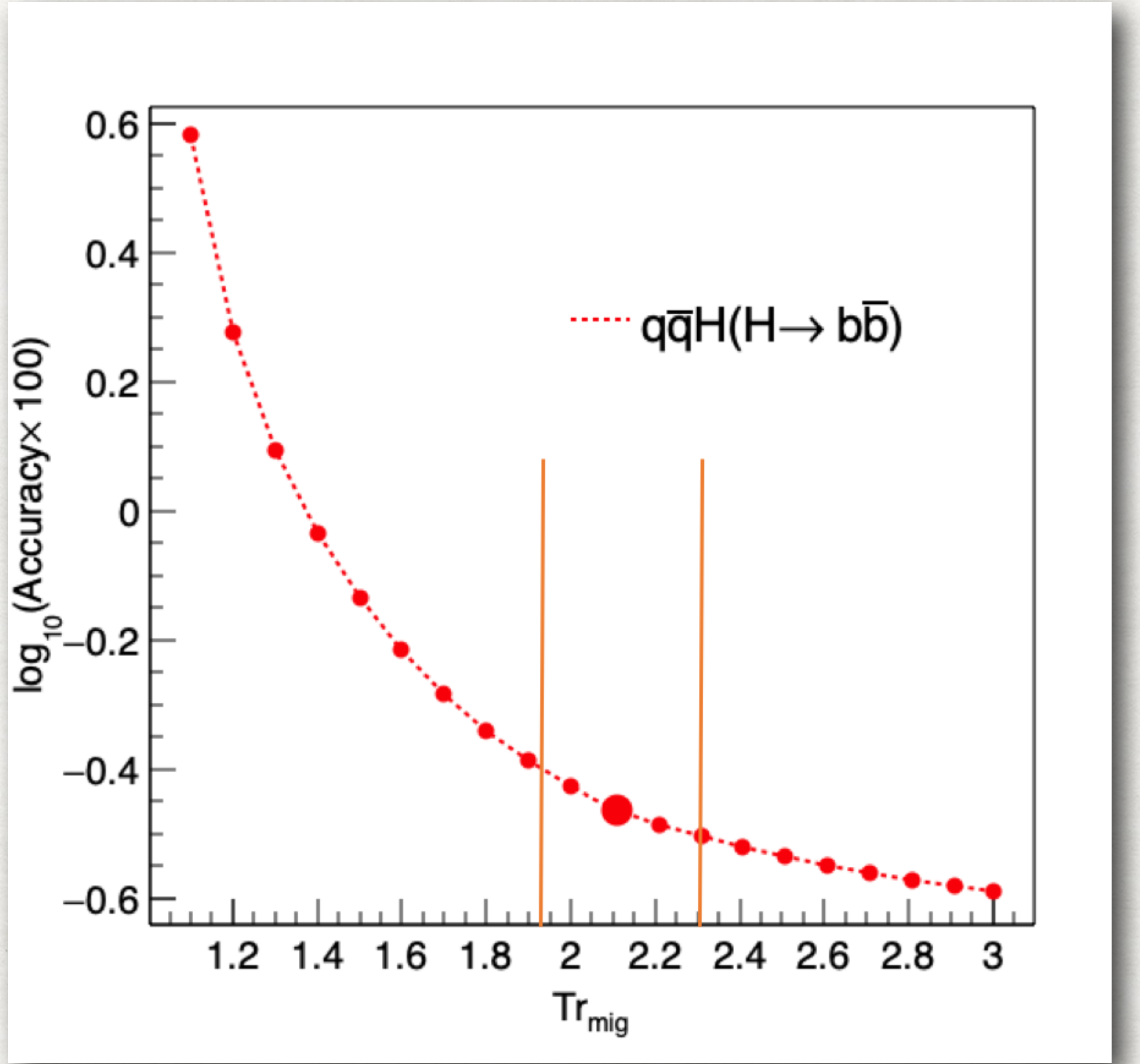
relative accuracy at different migration matrices

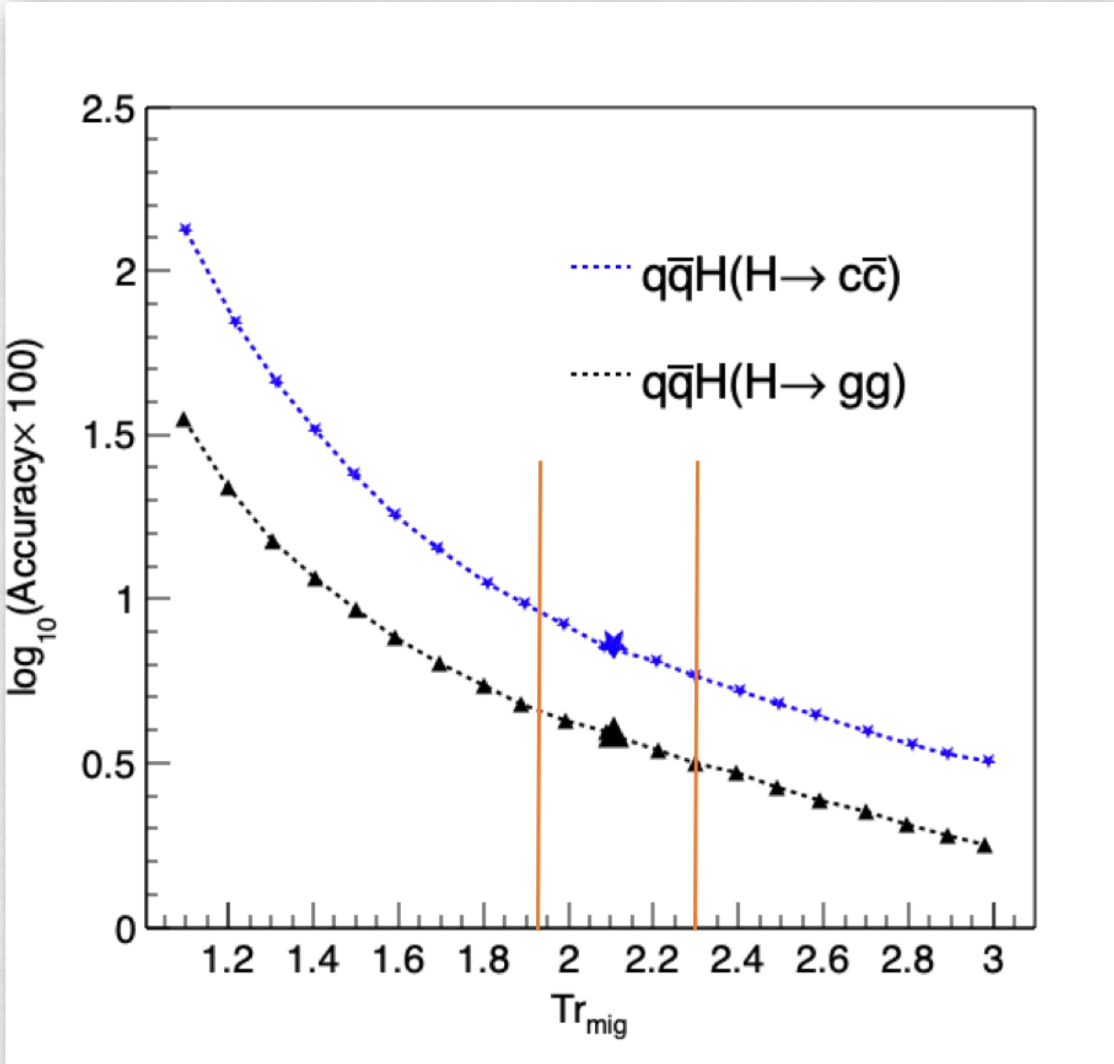




perfect flavor tagging performance can improve $H \to b\bar{b}/c\bar{c}/gg$ signal strength accuracy by 2%/63%/13%

$q\bar{q}H$





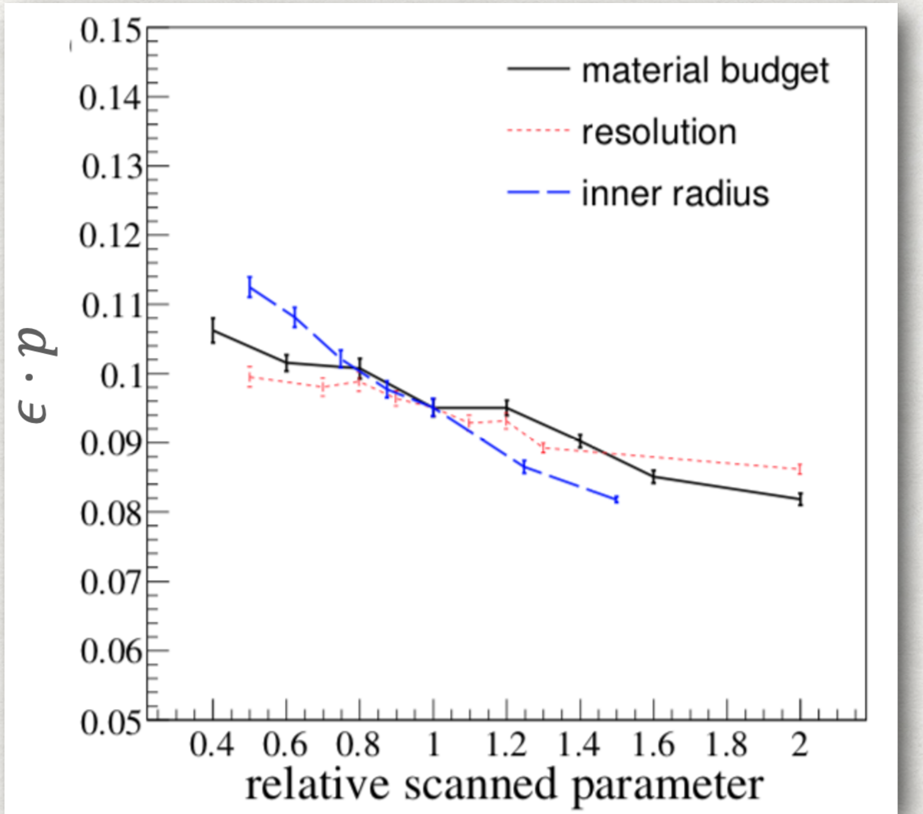
perfect flavor tagging performance can improve $H \to b\bar{b}/c\bar{c}/gg$ signal strength accuracy by 35%/122%/181%

The dependency of flavor tagging performance on vertex detector parameters

Z. Wu *et al* 2018 *JINST* **13** T09002

Table 1. The baseline design parameters of the CEPC vertex system.

	R(mm)	Z(mm)	single-point resolution(μm)	material budget
Layer 1	16	62.5	2.8	$0.15\%/X_0$
Layer 2	18	62.5	6	$0.15\%/X_0$
Layer 3	37	125.0	4	$0.15\%/X_0$
Layer 4	39	125.0	4	$0.15\%/X_0$
Layer 5	58	125.0	4	$0.15\%/X_0$
Layer 6	60	125.0	4	$0.15\%/X_0$



$$\epsilon \cdot p = 0.095\left(1 - 0.14\frac{\Delta x_{\text{material}}}{x_{\text{material}}}\right)\left(1 - 0.09\frac{\Delta x_{\text{resolution}}}{x_{\text{resolution}}}\right)\left(1 - 0.23\frac{\Delta x_{\text{radius}}}{x_{\text{radius}}}\right)$$

In $q\bar{q}H$ channel :

$$Tr_{mig} = 2.12 + 0.05 \cdot log_2\frac{R^0_{material}}{R_{material}} + 0.04 \cdot log_2\frac{R^0_{resolution}}{R_{resolution}} + 0.10 \cdot log_2\frac{R^0_{radius}}{R_{radius}}$$

$R^0_{material}$: the default material budget of vertex detector

$R_{material}$: the changed material budget

In $\nu\bar{\nu}H$ channel :

$$Tr_{mig} = 2.35 + 0.05 \cdot log_2\frac{R^0_{material}}{R_{material}} + 0.04 \cdot log_2\frac{R^0_{resolution}}{R_{resolution}} + 0.10 \cdot log_2\frac{R^0_{radius}}{R_{radius}}$$

Table 2. Reference geometries.

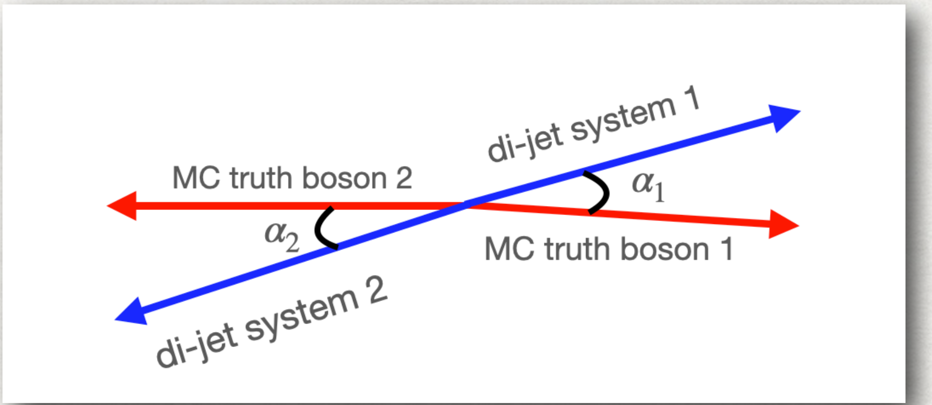
	Scenario A (Aggressive)	Scenario B (Baseline)	Scenario C (Conservative)
Material per layer/X_0	0.075	0.15	0.3
Spatial resolution/μm	1.4 - 3	2.8 - 6	5 - 10.7
R_{in}/mm	8	16	23
trace in $q\bar{q}H$ channel	**2.31**	**2.12**	**1.93**
trace in $\nu\bar{\nu}H$ channel	**2.54**	**2.35**	**2.16**

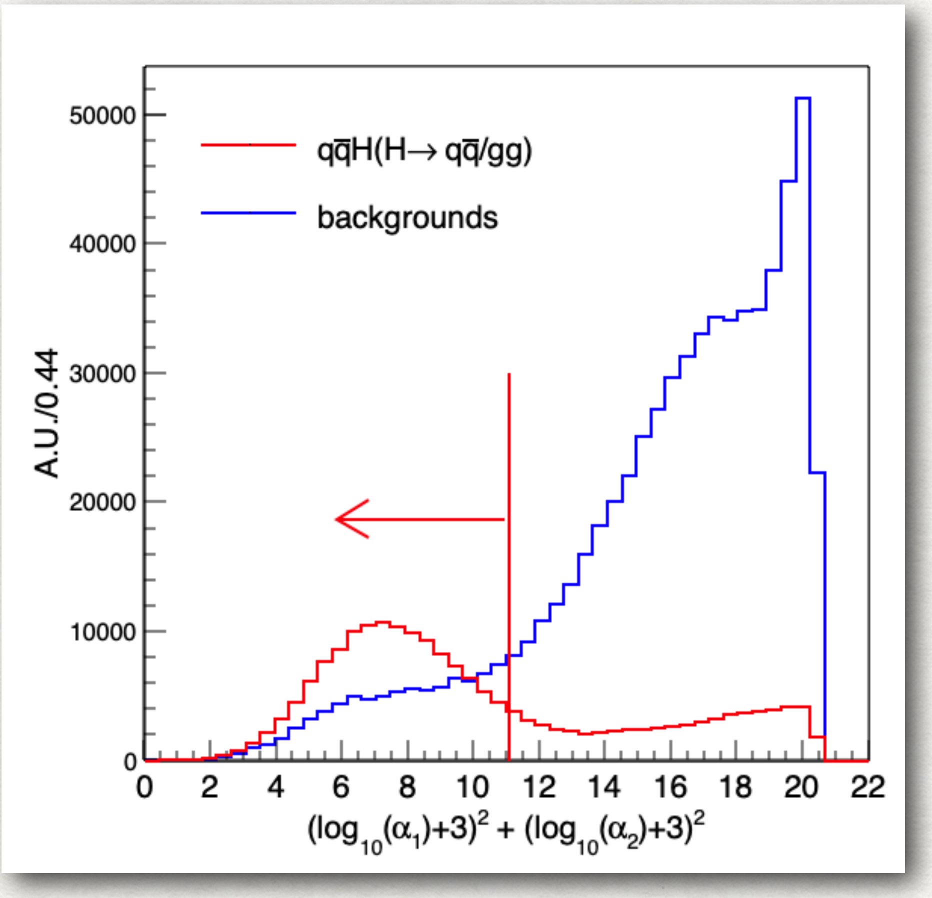
the dependency of $qqH \rightarrow b\bar{b}/c\bar{c}/gg$ accuracy

on

color singlet identification (CSI)

CSI : the reconstruction of W/Z/Higgs when they decay to two jets





	$qqHb\bar{b}$	$qqHc\bar{c}$	$qqHgg$
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 a demonstrator to illustrate the importance of CSI. Our next step is to construct a CSI evaluator at the reconstruction level.

Summary :

- The relative accuracy of $H \to b\bar{b}/c\bar{c}/gg$ signal strength has been measured at the CEPC with full simulation of all SM samples corresponding to 240 GeV.

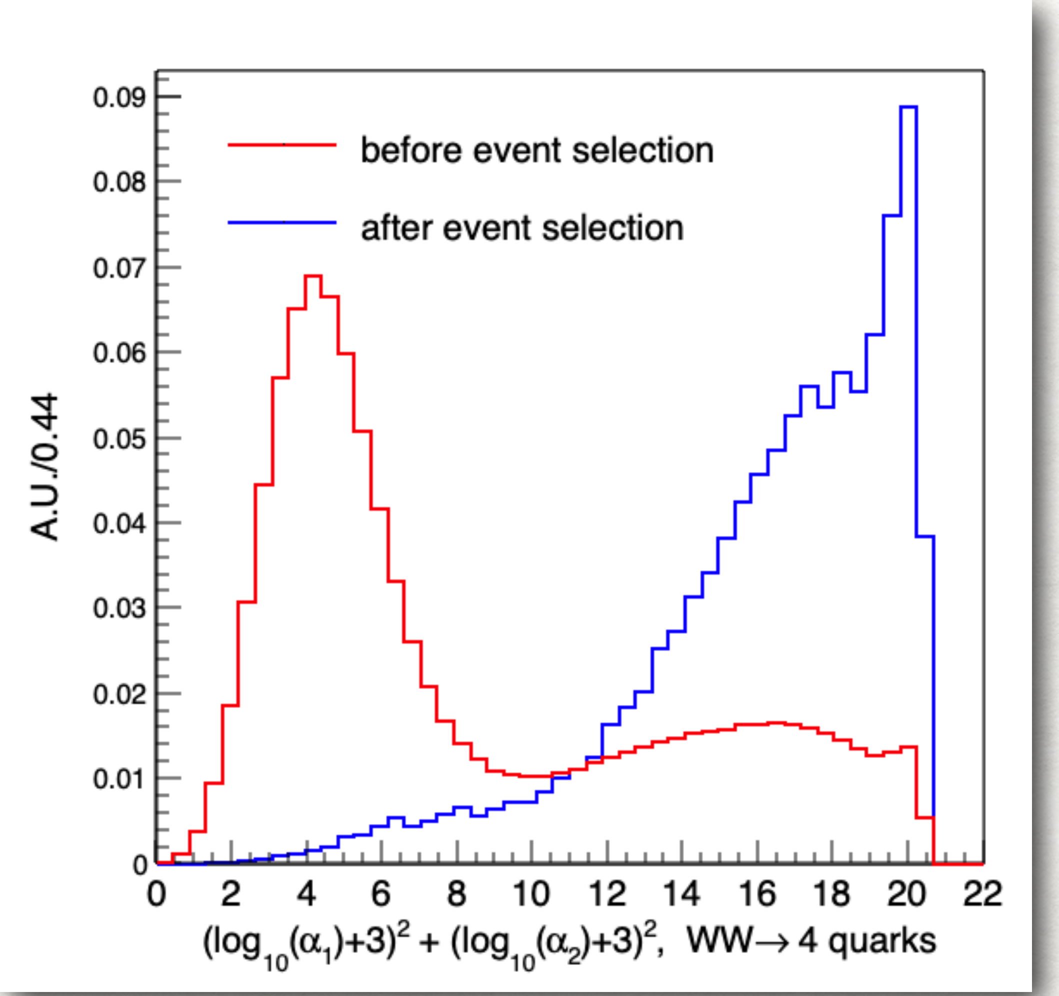
	$H \to b\bar{b}$	$H \to c\bar{c}$	$H \to gg$
$5.6ab^{-1}$	0.27%	4.03%	1.56%
$20ab^{-1}$	0.14%	2.13%	0.82%

- The perfect flavor tagging performance could improve the accuracy of $H \to b\bar{b}/c\bar{c}/gg$ by 2%/63%/13% in the $\nu\bar{\nu}Higgs$ channel and 35%/122%/181% in the $q\bar{q}H$ channel.
- The vertex detector design has great impact on flavor tagging performance.
- The good CSI performance or CSI evaluator at the reconstruction level is essential for the improvement of measurement accuracy.

Backup

The signal strength accuracy of H->bb/cc/gg at CEPC, ILC and FCC-ee.

	bb	cc	gg
CEPC	0.27%	4.03%	1.56%
ILC, 250 GeV, $2ab^{-1}$	0.7%	4%	
FCC-ee, 240 GeV,$5ab^{-1}$	0.3%	2.2%	1.9%



This plot is used to illustrate that the event selection procedure can suppress the backgrounds with good CSI performance.

Systematic uncertainties

We categorize the leading systematic uncertainties into three groups:

- The first group includes the reconstructed energy/momentum scale of the physics objects, which are significantly smaller than the statistical uncertainties.
- 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.

The detailed discussion can be found in https://arxiv.org/abs/2203.01469.