The $Higgs \to b\bar{b}, c\bar{c}, gg$ measuremnet at CEPC & corresponding optimization studies

Yongfeng Zhu

advisor : Manqi Ruan

IHEP

2022 年 5 月 23 日

Motivation

- The g(Hcc), the second generation fermion Yukawa coupling, is one of the most important benchmark at CEPC, and its measurement strongly depend on the Higgs branching ratio measurement accuracy.
- It is very important to estimate the measurement accuracy of Higgs branching ratio to verify the Higgs mechanism.

Contents

- The relative accuracy of signal strength measurement of $\nu\nu H(H \rightarrow b\bar{b}, c\bar{c}, gg)$.
  - key performance : flavor tagging
- The relative accuracy of signal strength measurement of $qqH(H \rightarrow b\bar{b}, c\bar{c}, gg)$.
  - key performance : flavor tagging
  - key performance : color singlet identification
- summary

Sample

- all SM processes at CEPC ($\sqrt{s} = 240GeV$) with integrated luminosity of $5600fb^{-1}$

$llH(H \to b\bar{b}, H \to c\bar{c}, H \to gg)$ *accuracy measurement* at CEPC ($\sqrt{s} = 250\ GeV$ with integrated luminosity $5000fb^{-1}$)

Chinese Physics C Vol. 44, No. 1 (2020) 013001

Measurements of decay branching fractions of $H \to b\bar{b}/c\bar{c}/gg$ in associated $(e^+e^-/\mu^+\mu^-)H$ production at the CEPC*

Yu Bai(白羽)[1;1)] Chun-Hui Chen(陈春晖)[2;2)] Ya-Quan Fang(方亚泉)[3;3)] Gang Li(李刚)[3;4)]
Man-Qi Ruan(阮曼奇)[3;5)] Jing-Yuan Shi(史静远)[1,4] Bo Wang(王博)[1,5] Pan-Yu Kong(孔攀宇)[1,6]
Bo-Yang Lan(兰博扬)[1] Zhan-Feng Liu(刘站峰)[1]

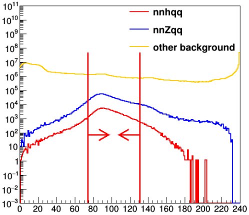
[1]School of Physics, Southeast University, Nanjing 210096, China
[2]Department of Physics and Astronomy, Iowa State University, Ames 50011-3160, USA
[3]Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049, China
[4]School of physics and astronomy, the University of Manchester, Oxford Rd, M13 9PL, UK
[5]Department of Physics, College of Sciences, Northeastern University, Shenyang 110004, China
[6]School of Cyber Science and Engineering, Southeast University, Nanjing 210096, China

	$\mu^+\mu^-H$			e^+e^-H		
	$H \to b\bar{b}$	$H \to c\bar{c}$	$H \to gg$	$H \to b\bar{b}$	$H \to c\bar{c}$	$H \to gg$
accuracy	1.1%	10.5%	5.4%	1.6%	14.7%	10.5%

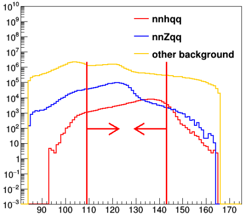
for the detail : `https://arxiv.org/abs/1905.12903v2`

$\nu\nu H(H \rightarrow b\bar{b}, H \rightarrow c\bar{c}, H \rightarrow gg)$ *accuracy measurement*

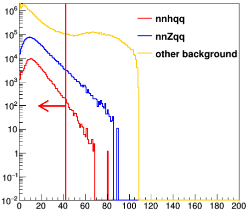
The following cuts were used to identify $\nu\nu Hq\bar{q}/gg$.



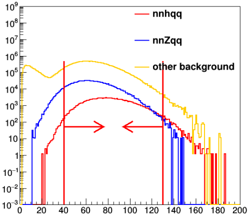
74 < recoil mass < 131



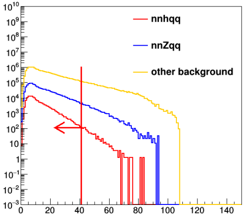
109 < visible energy < 143



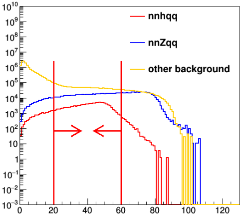
leading lepton energy < 42



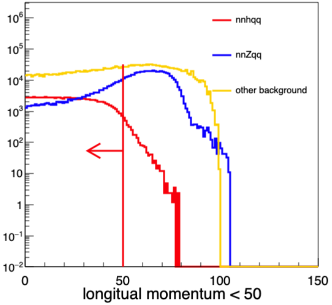
40 < multiplicity < 130

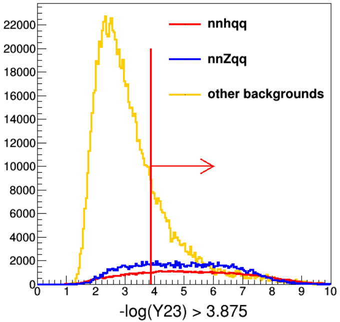


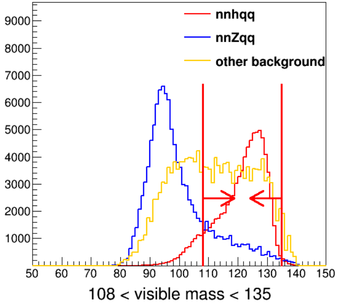
leading neutral energy < 41



20 < transverse momentum < 60


nnhqq
nnZqq
other background
longitual momentum < 50

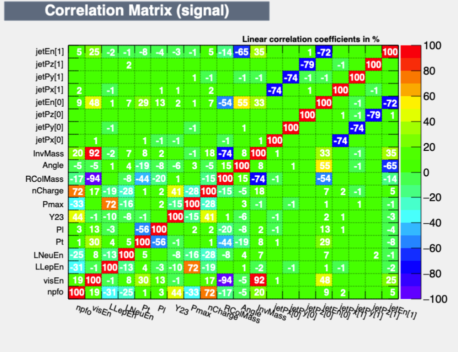

nnhqq
nnZqq
other backgrounds
-log(Y23) > 3.875

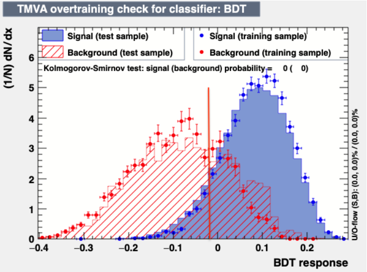

nnhqq
nnZqq
other background
108 < visible mass < 135

BDT for selecting $\nu\nu Hq\bar{q}$

input variables :

- recoilMass : the recoil mass of final state particles
- visEn : visible energy of final state particles
- Npfo : number of final state particles
- leadLepEn : leading lepton energy,
- leadNeuEn : leading neutral energy
- Pt : the transverse momentum of all final state particles
- Pl : the longitudinal momentum of all final state particles
- Pmax : the maximum transverse momentum among final state particles
- Y23
- InvMass : visible mass
- num_charge : number of charge particles
- Angle : the angle between two jets
- 4 momentum of two jets





The full cut chain shown in the following table.

	$\nu\nu Hq\bar{q}$	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

cut based flavor tagging

introduce the flavor tagging performance matrix :

eff \ to / true	b	c	udsg
b	b to b	b to c	b to udsg
c	c to b	c to c	c to udsg
udsg	udsg to b	udsg to c	udsg to udsg

to : identified as

c to c : the ratio of c jet identified as c

	c	b	udsg
udsg	0	0	1
b	0	1	0
c	1	0	0

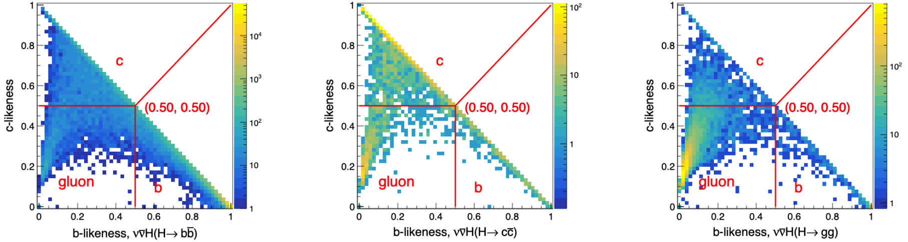
perfect flavor tagging

	c	b	udsg
udsg	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
b	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
c	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$

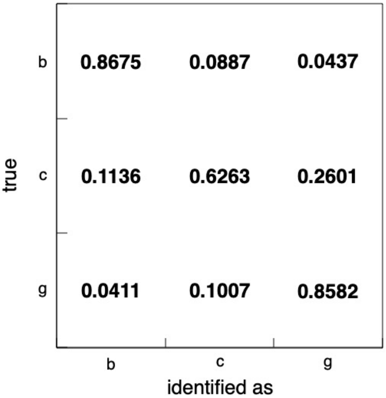
non flavor tagging

Optimized matrix

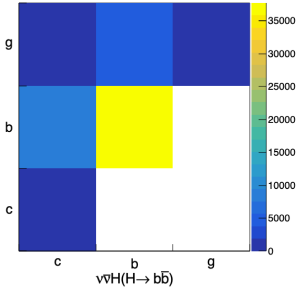
1. The b-likeness and c-likeness of two jets can be displaced in 2D graph.
2. The cut on b-likeness and c-likeness can be find to maximize the value of $eff(b \to b) + eff(c \to c) + eff(udsg \to udsg)$, the trace of flavor tagging matrix.

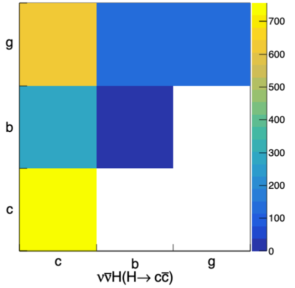


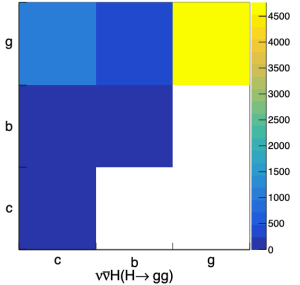
Optimized matrix :

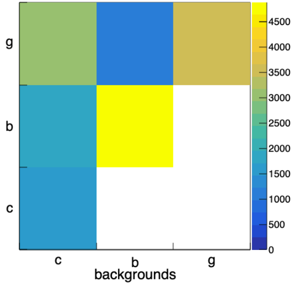


events distribution based on optimized matrix :









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

- 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, tt is 0.49%/5.75%/1.82% for $\nu\nu Hb\bar{b}/c\bar{c}/gg$.

key performance : flavor tagging

changing flavor tagging performance :

$$non\ flavor\ tagging \rightarrow perfect\ flavor\ tagging$$

the changing procedure of flavor tagging performance matrix :

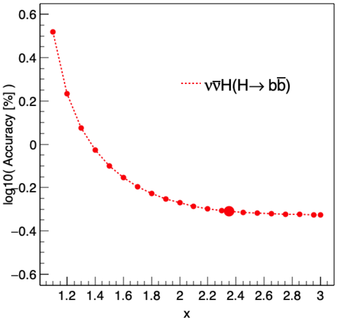
$$temp\ matrix = \frac{x - trace_I}{trace_T - trace_I} \cdot (T - I) + I \qquad (trace_I \leq x \leq trace_T)$$

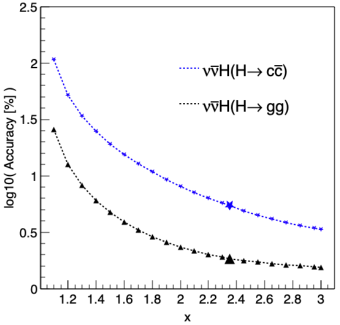
T : matrix with perfect flavor tagging

I : matrix with non flavor tagging

$trace_I, trace_T$: the trace of matrix I and T

The x value and flavor tagging performance matrix have a one to one relation.



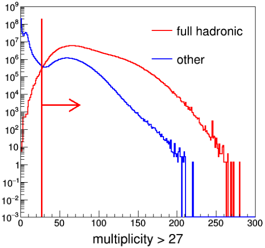


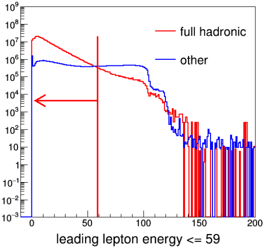
relative accuracy (%)	$\nu\nu Hb\bar{b}$	$\nu\nu Hc\bar{c}$	$\nu\nu Hgg$
cut based flavor tagging	0.49	5.75	1.82
perfect flavor tagging	0.48	3.53	1.61

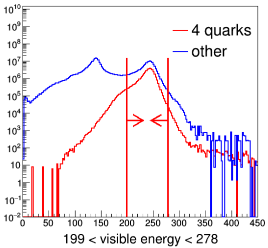
$qqH(H \rightarrow b\bar{b}, H \rightarrow c\bar{c}, H \rightarrow gg)$ *accuracy measurement*

1. Finding the full hadronic samples from all samples.
2. Finding 4-quark samples from the full hadronic samples.
3. Finding $ZH(Z \to q\bar{q}, H \to q\bar{q})$ from 4-quark samples.

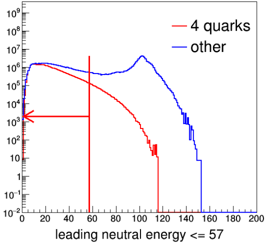
Firstly, finding the full hadronic samples from all samples.

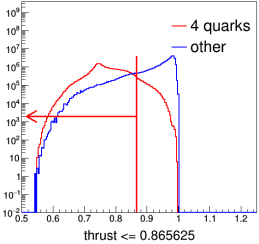


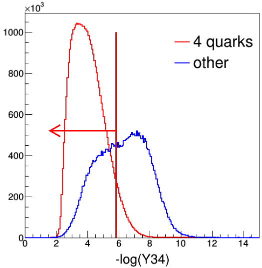




Secondly, finding 4-quark samples from the full hadronic samples.

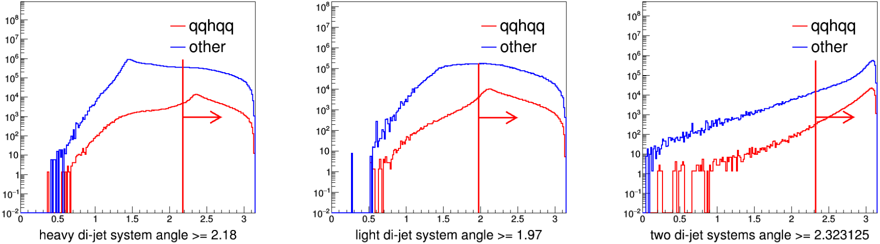




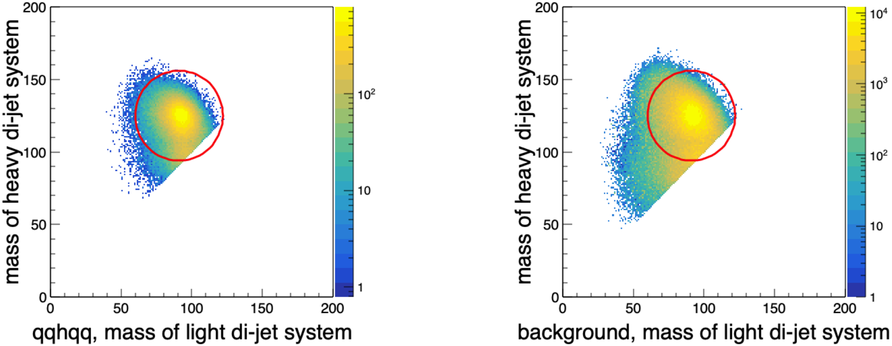


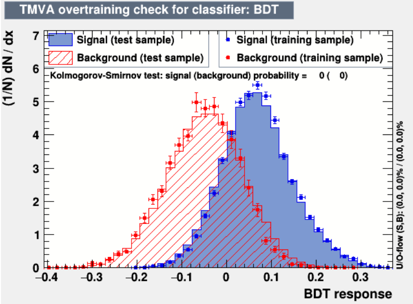
Thirdly, finding $ZH(Z \to q\bar{q}, H \to q\bar{q})$ from 4-quark samples.

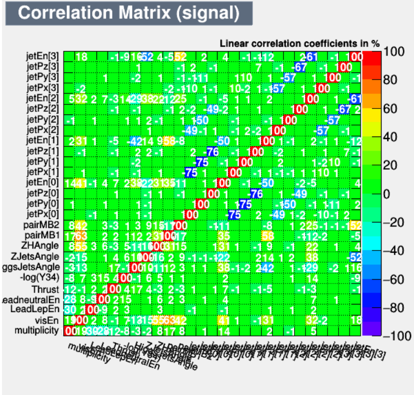
After finding 4-quark samples, the method of maximize $\chi^2 = \frac{(M_{12}-M_{B1})^2}{\sigma_{B1}^2} + \frac{(M_{34}-M_{B2})^2}{\sigma_{B2}^2}$ can be used to pair four jets into two di-jet systems.



Then a circle can be used to find ZH events.



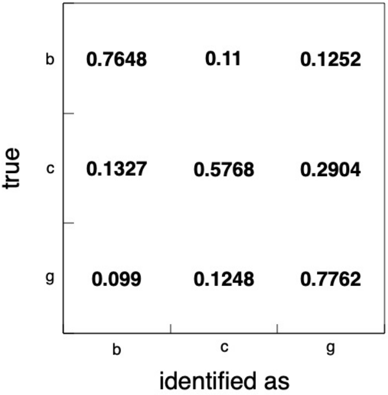

TMVA overtraining check for classifier: BDT
Signal (test sample)
Background (test sample)
Signal (training sample)
Background (training sample)
Kolmogorov-Smirnov test: signal (background) probability = 0 (0)
(1/N) dN / dx
BDT response
U/O-flow (S,B): (0.0, 0.0)% / (0.0, 0.0)%

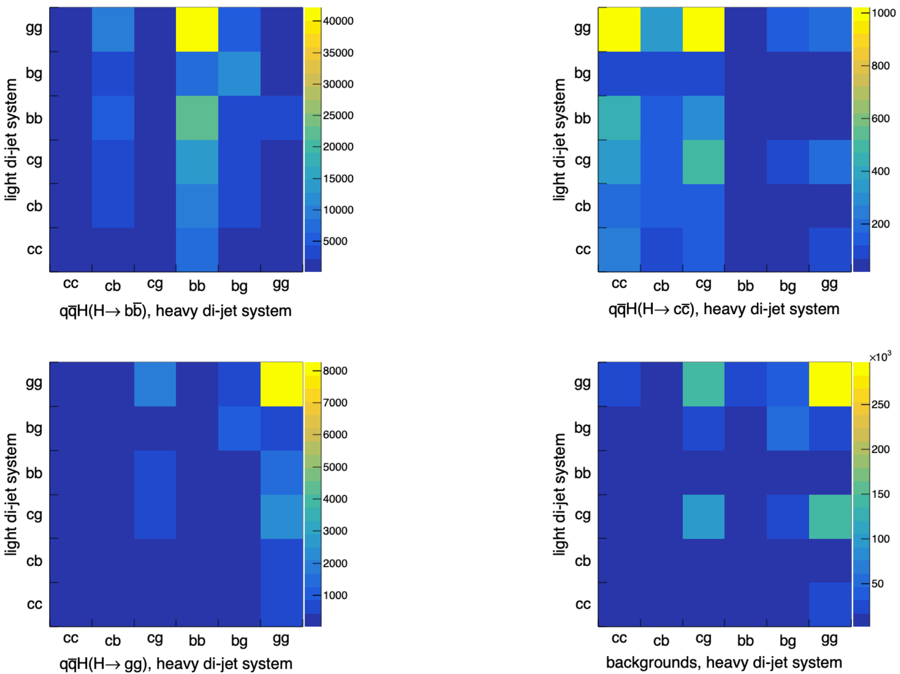

Correlation Matrix (signal)
Linear correlation coefficients in %
jetEn[3]
jetPz[3]
jetPy[3]
jetPx[3]
jetEn[2]
jetPz[2]
jetPy[2]
jetPx[2]
jetEn[1]
jetPz[1]
jetPy[1]
jetPx[1]
jetEn[0]
jetPz[0]
jetPy[0]
jetPx[0]
pairMB2
pairMB1
ZHAngle
ZJetsAngle
-log(Y34)
Thrust
LeadLepEn
visEn
multiplicity

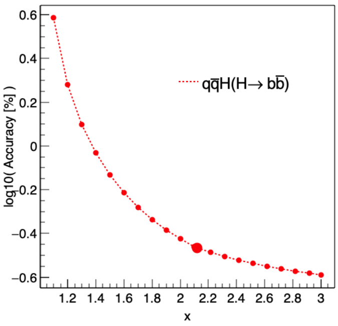
the cut chain of first three stages shown in the following table :

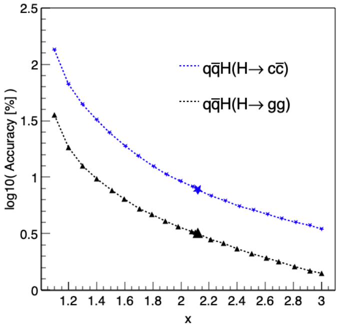
	qqHqq	2f	SW	SZ	WW	ZZ	Mixed	ZH	$\frac{\sqrt{S+B}}{S}$ (%)
total	527488	$8.01E8$	$1.95E7$	$9.07E6$	$5.08E7$	$6.39E6$	$2.18E7$	613008	5.71
multiplicity $\in (27, +\infty)$	527488	$3.04E8$	$1.46E7$	$3.37E6$	$4.85E7$	$6.00E6$	$1.81E7$	577930	3.77
leadLepEn $\in (0, 59)$	527036	$2,98E8$	$6.76E6$	$2.44E6$	$3.93E7$	$5.40E6$	$1.79E7$	531411	3.65
visEn $\in (199, 278)$	510731	$1.21E8$	$1.29E6$	551105	$2.14E7$	$3.06E6$	$1.71E7$	180571	2.52
leadNeuEn $\in (0, 57)$	509623	$5.68E7$	716161	168030	$2.04E7$	$2.93E6$	$1.65E7$	176387	1.94
thrust $\in (0, 0.86)$	460535	$7.81E6$	473732	132126	$1.88E7$	$2.60E6$	$1.54E7$	167863	1.47
$-log(Y_{34})$ $\in (0, 5.8875)$	451468	$4.90E6$	181432	119836	$1.74E7$	$2.40E6$	$1.45E7$	165961	1.40
HiggsJetsA $\in (2.18, 2\pi)$	326207	$2.83E6$	110156	58613	$4.54E6$	870276	$3.74E6$	96560	1.08
ZJetsA $\in (1.97, 2\pi)$	279030	$1.37E6$	33491	37101	$2.39E6$	496611	$2.00E6$	74005	0.93
ZHiggsA $\in (2.32, 2\pi)$	274530	$1.32E6$	17026	33847	$2.28E6$	468340	$1.91E6$	69620	0.92
circle	268271	1.20E6	10193	31567	$2.13E6$	424514	1.79E6	65434	0.90
BDT $\in (0.02, 1)$	192278	378300	40	307	271436	141446	244126	30022	0.57

optimized matrix




light di-jet system
gg
bg
bb
cg
cb
cc
40000
35000
30000
25000
20000
15000
10000
5000
cc cb cg bb bg gg
$q\bar{q}H(H\rightarrow b\bar{b})$, heavy di-jet system
light di-jet system
1000
800
600
400
200
$q\bar{q}H(H\rightarrow c\bar{c})$, heavy di-jet system
light di-jet system
8000
7000
6000
5000
4000
3000
2000
1000
$q\bar{q}H(H\rightarrow gg)$, heavy di-jet system
light di-jet system
$\times 10^3$
250
200
150
100
50
backgrounds, heavy di-jet system





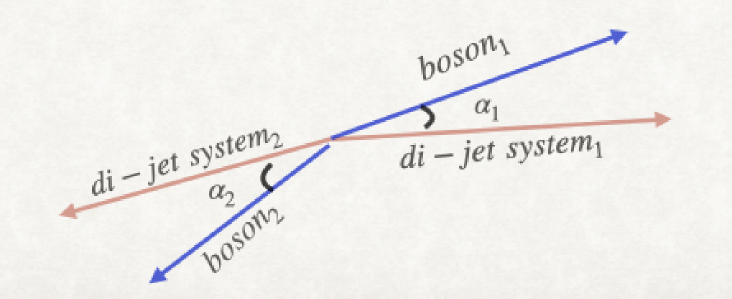
relative accuracy (%)	$qqHb\bar{b}$	$qqHc\bar{c}$	$qqHgg$
cut based flavor tagging	0.35	7.22	3.79
perfect flavor tagging	0.26	3.48	1.41

key performance : color singlet identification

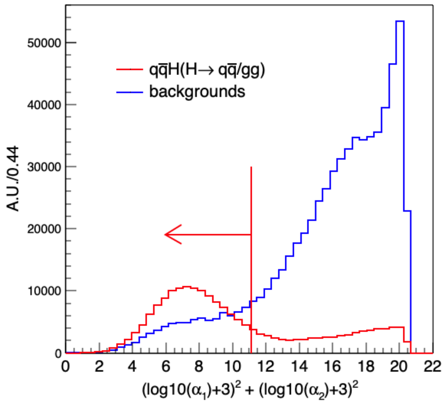
(i.e. jet clustering and jet matching)

color singlet identification

For di-boson event, there are two MC truth bosons and two di-jet systems, the variable $\alpha_i = angle(di-jet\ system_i, truth\ boson_i), (i = 1, 2)$ is used to characterize the performance of jet clustering and jet matching.

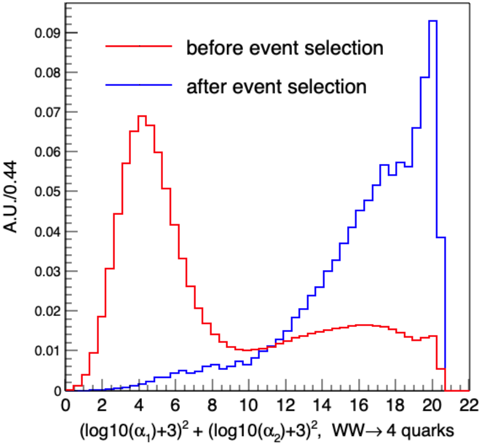


the α variable is just a cheated variable used to characterize the performance of color singlet identification



	$qqHb\bar{b}$	$qqHc\bar{c}$	$qqHgg$
w.o. alpha cut	0.35%	7.22%	3.79%
w.i. alpha cut	0.33%	4.37%	2.08%

The distribution of $(log(\alpha_1) + 3)^2 + (log(\alpha_2) + 3)^2$ for the samples of $e^+e^- \rightarrow W^+W^- \rightarrow 4\ quarks$, the red line corresponds to all samples, and the blue line corresponds to that after the event selection process.



Summary :

- The total signal strength of $H \to b\bar{b}, c\bar{c}, gg$ can be measured to a relative accuracy of 0.27%/3.82%/1.52%, combining all four different channels of $\mu\mu H$, $eeH, \nu\nu H$ and qqH.

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%	7.22%	3.79%
$Z \to \nu\bar{\nu}$	0.49%	5.35%	1.77%
combination	0.27%	3.82%	1.52%

- The flavor tagging and color singlet identification (CSI) are the critical performances for these benchmarks. Their impact on the anticipated physics reach is evaluated.
  - for $\nu\nu H$ channel
    - The flavor tagging is critical for the $\nu\bar{\nu}H(H \to b\bar{b}/c\bar{c}/gg)$ measurement. Using an ideal flavor tagging, the anticipated accuracy could be improved by 2%/52%/10%.
  - for qqH channel
    - With perfect flavor tagging, the anticipated accuracy of $q\bar{q}H(H \to b\bar{b}/c\bar{c}/gg)$ could be improved by 35%/107%/169%.
    - If we can quantify the CSI performance and select the events with good CSI performance, the $q\bar{q}H(H \to b\bar{b}/c\bar{c}/gg)$ accuracy can be improved by 6%/65%/82%.
- A good color singlet identification, or even a reliable color singlet identification performance evaluator at reconstruction level, is highly appreciated.

https://arxiv.org/abs/2203.01469

Thanks !

Backup

Z decay mode	$H \to b\bar{b}$	$H \to c\bar{c}$	$H \to gg$	scenario
CEPC	0.27%	3.91%	1.37%	240GeV, 5600fb^{-1}
CEPC CDR	0.3%	3.1%	1.2%	240GeV, 5600fb^{-1}
ILC	1.0%	6.02%	8.5%	250 GeV, 250fb^{-1}
FCC	0.3%	2.2%	1.9%	240 GeV, 5000fb^{-1}

ILC result

Eur. Phys. J. C (2013) 73:2343
DOI 10.1140/epjc/s10052-013-2343-8

THE EUROPEAN PHYSICAL JOURNAL C

Special Article - Tools for Experiment and Theory

A study of measurement precision of the Higgs boson branching ratios at the International Linear Collider

Hiroaki Ono[1,a], **Akiya Miyamoto**[2]

[1]School of Life Dentistry at Niigata, Nippon Dental University, 1-8 Hamaura-cho, Chuo-ku, Niigata 951-8580, Japan
[2]High Energy Accelerator Research Organization, 1-1 Oho, Tsukuba 305-0801, Japan

Table 1 Summary of the $\nu\bar{\nu}H$ channel background reduction assuming $\mathscr{L} = 250$ fb^{-1} with $P(e^-, e^+) = (-0.8, +0.3)$

CM energy	250 GeV			350 GeV
Cut names	Condition	Sig.	Bkg.	Condition
Generated		19360	44827100	
Missing mass	$80 < M_{miss} < 140$ GeV	15466	6214050	$50 < M_{miss} < 240$ GeV
Transverse visible momentum	$20 < P_T < 70$ GeV	13727	549340	$10 < P_T < 140$ GeV
Longitudinal visible momentum	$P_L < 60$ GeV	13342	392401	$P_L < 130$ GeV
# of charged tracks	$N_{chd} > 10$	12936	374877	$N_{chd} > 10$
Maximum track momentum	$P_{max} < 30$ GeV	11743	205038	$P_{max} < 60$ GeV
Y_{23} value	$Y_{23} < 0.02$	7775	74439	$Y_{23} < 0.02$
Y_{12} value	$0.2 < Y_{12} < 0.8$	7438	62584	$0.2 < Y_{12} < 0.8$
Di-jet mass	$100 < M_{jj} < 130$ GeV	6691	19061	$100 < M_{jj} < 130$ GeV
Likelihood ratio	$LR > 0.165$	6293	10940	$LR > 0.395$
Significance (Efficiency)	$S/\sqrt{S+B}$	47.9 (32.5 %)		$S/\sqrt{S+B}$

Table 2 Summary of $q\bar{q}H$ channel background reduction assuming $\mathscr{L} = 250$ fb^{-1} with $P(e^-, e^+) = (-0.8, +0.3)$

CM energy	250 GeV			350 GeV
Cut names	Condition	Sig.	Bkg.	Condition
Generated		52507	45904900	
χ^2	$\chi^2 < 10$	32447	2608980	$\chi^2 < 10$
# of charged tracks	$N_{chd} > 4$	25281	1120950	$N_{chd} > 4$
Y_{34} value	$-\log(Y_{34}) > 2.7$	25065	1002125	$-\log(Y_{34}) > 2.7$
thrust	thrust < 0.9	24688	935950	thrust < 0.85
thrust angle	$\lvert\cos\theta_{thrust}\rvert < 0.9$	21892	696201	$\lvert\cos\theta_{thrust}\rvert < 0.9$
Higgs jets angle	$105° < \theta_H < 160°$	20062	622143	$70° < \theta_H < 120°$
Z di-jet mass	$80 < M_Z < 100$ GeV	16359	411863	$80 < M_Z < 100$ GeV
H di-jet mass	$105 < M_H < 130$ GeV	16359	411863	$105 < M_H < 130$ GeV
Likelihood ratio	$LR > 0.375$	13726	166807	$LR > 0.15$
Significance (Efficiency)	$S/\sqrt{S+B}$	32.3 (26.1 %)		$S/\sqrt{S+B}$

	Higgs Mass	$H \to b\bar{b}$	$H \to c\bar{c}$	$H \to gg$
CEPC	125 GeV	57.7%	2.91%	8.57%
ILC	120 GeV	65.7%	3.6%	5.5%

	ILC	CEPC
luminosity	$250fb^{-1}$	$5600fb^{-1}$
polarization	ep(-80%+30%)	
cross-section		
total events		
Higgs Mass	120 GeV	125
$H \to b\bar{b}$	65.36%	57.9%
$H \to c\bar{c}$	3.244%	2.87%
$H \to gg$	8.409%	8.17%

$\nu\nu H$	ILC	CEPC
total events	$250fb^{-1}$	$5600fb^{-1}$
before BDT	ep(-80%+30%)	
after BDT		

FCCee result

Table 1.1. Relative statistical uncertainty on $\sigma_{\mathrm{HZ}} \times \mathrm{BR}(\mathrm{H} \to \mathrm{XX})$ and $\sigma_{\nu\bar{\nu}\mathrm{H}} \times \mathrm{BR}(\mathrm{H} \to \mathrm{XX})$, as expected from the FCC-ee data, obtained from a fast simulation of the CLD detector and consolidated with extrapolations from full simulations of similar linear-collider detectors (SiD and CLIC).

$\sqrt{s}$ (GeV)	240		365	
Luminosity (ab^{-1})	5		1.5	
$\delta(\sigma\mathrm{BR})/\sigma\mathrm{BR}$ (%)	HZ	$\nu\bar{\nu}$ H	HZ	$\nu\bar{\nu}$ H
$\mathrm{H} \to$ any	± 0.5		± 0.9	
$\mathrm{H} \to \mathrm{b\bar{b}}$	± 0.3	± 3.1	± 0.5	± 0.9
$\mathrm{H} \to \mathrm{c\bar{c}}$	± 2.2		± 6.5	± 10
$\mathrm{H} \to \mathrm{gg}$	± 1.9		± 3.5	± 4.5
$\mathrm{H} \to \mathrm{W^+W^-}$	± 1.2		± 2.6	± 3.0
$\mathrm{H} \to \mathrm{ZZ}$	± 4.4		± 12	± 10
$\mathrm{H} \to \tau\tau$	± 0.9		± 1.8	± 8
$\mathrm{H} \to \gamma\gamma$	± 9.0		± 18	± 22
$\mathrm{H} \to \mu^+\mu^-$	± 19		± 40	
$\mathrm{H} \to$ invisible	<0.3		<0.6	

Notes. All numbers indicate 68% CL intervals, except for the 95% CL sensitivity in the last line. The accuracies expected with $5\,\mathrm{ab}^{-1}$ at 240 GeV are given in the middle column, and those expected with $1.5\,\mathrm{ab}^{-1}$ at $\sqrt{s} = 365$ GeV are displayed in the last column.

CEPC CDR result

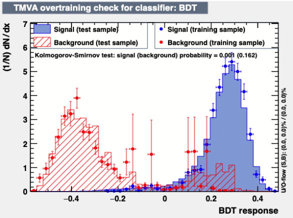
Table 6. Expected relative precision on $\sigma(ZH)\times\mathrm{BR}$ for the $H\to b\bar{b}$, $c\bar{c}$ and gg decays from a CEPC dataset of $5.6\,\mathrm{ab}^{-1}$.

Z decay mode	$H\to b\bar{b}$	$H\to c\bar{c}$	$H\to gg$
$Z\to e^+e^-$	1.3%	12.8%	6.8%
$Z\to \mu^+\mu^-$	1.0%	9.4%	4.9%
$Z\to q\bar{q}$	0.5%	10.6%	3.5%
$Z\to \nu\bar{\nu}$	0.4%	3.7%	1.4%
combination	0.3%	3.1%	1.2%

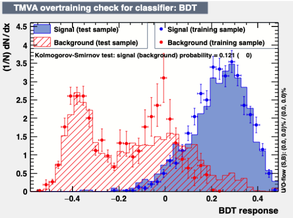
Z decay mode	$H\to b\bar{b}$	$H\to c\bar{c}$	$H\to gg$
$Z\to e^+e^-$	1.6%	14.7%	10.5%
$Z\to \mu^+\mu^-$	1.1%	10.5%	5.4%
$Z\to q\bar{q}$	0.34%	5.58%	1.91%
$Z\to \nu\bar{\nu}$	0.48%	5.92%	1.95%
combination	0.26%	3.67%	1.31%

The BDT was used to identify $\nu\nu H(H \to b\bar{b}, H \to c\bar{c}, H \to gg)$.

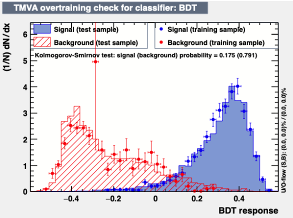
- # of final state particles
- visible energy of final state particles
- leading lepton energy, leading neutral energy
- thrust, Y23
- the transverse momentum of all final state particles
- the longitudinal momentum of all final state particles
- the maximum transverse momentum among final state particles
- the angle of two jets
- # of charge particles
- visible mass
- recoil mass of final state particles
- the b-likeness of each jet
- the b-likeness of each jet
- the energy difference of two hemispheres



$\nu\nu Hb\bar{b}$



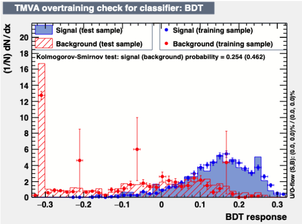
$\nu\nu Hc\bar{c}$



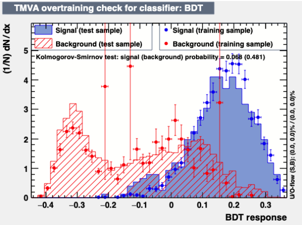
$\nu\nu Hgg$

	signal	2f	SW	SZ	WW	ZZ	Mixed	ZH	total bkg	$\frac{\sqrt{S+B}}{S}$ (%)
$\nu\nu Hc\bar{c}$	1868	5181	120	693	3578	1345	33	1467	12420	6.39
$\nu\nu Hb\bar{b}$	47928	7632	0	940	27	1478	0	305	10385	0.50
$\nu\nu Hg\bar{g}$	5568	726	210	1273	2222	1772	0	789	6995.22	2.01

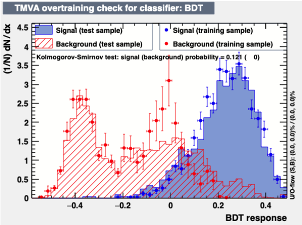
BDT with different input variables for $\nu\nu Hc\bar{c}$



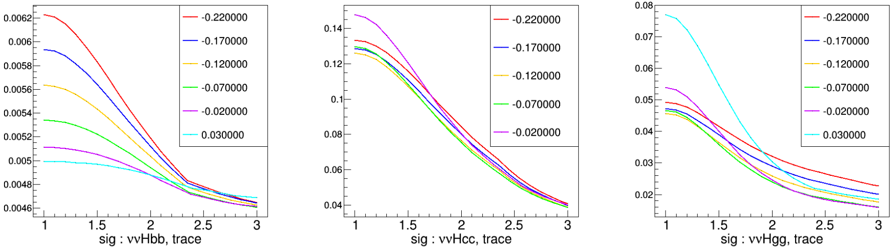
8.64% with only b/c-likeness



7.39% with b/c-likeness & jets 4-momentum



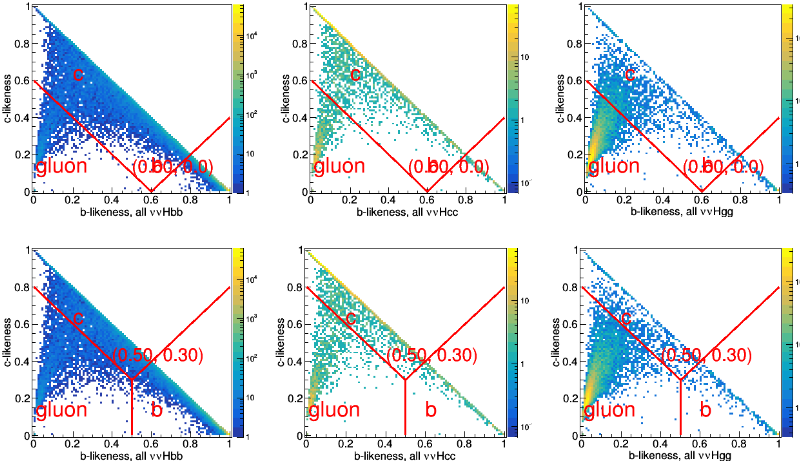
6.39% with b/c-likeness & jets 4-momentum & cut variables with different BDT cut value :



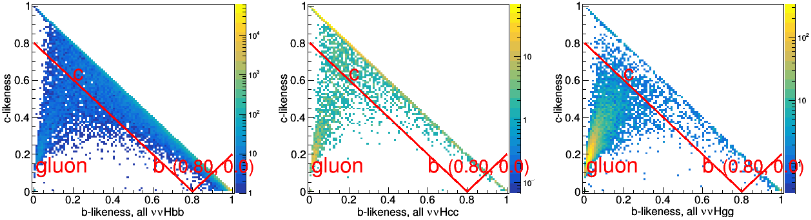
see relative accuracy as optimization target

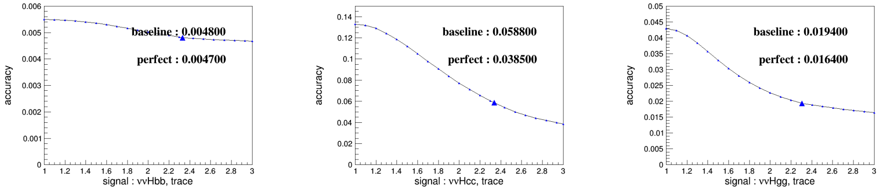
bb

cc



gg





relative accuracy (%)		$\nu\nu Hb\bar{b}$	$\nu\nu Hc\bar{c}$	$\nu\nu Hgg$
trace as optimization target	cut based flavor tagging	0.48	5.92	1.95
	perfect flavor tagging	0.47	3.85	1.64
accuracy as optimization target	cut based flavor tagging	0.48	5.88	1.94
	perfect flavor tagging	0.47	3.85	1.64