

# Measurements of the decay branching ratios of $H \rightarrow b\bar{b}/c\bar{c}/gg$ from higgs production in association with lepton pair at CEPC

To Be Added

*To Be Added*

*Elsevier Inc<sup>a,b</sup>, Global Customer Service<sup>b,\*</sup>*

<sup>a</sup>*1600 John F Kennedy Boulevard, Philadelphia*

<sup>b</sup>*360 Park Avenue South, New York*

---

## Abstract

The measurement on Higgs boson parameters with high precision is one of the primary goals of the high energy Circular Electron Positron Collider(CEPC). A study on the measurements of  $H \rightarrow b\bar{b}/c\bar{c}/gg$  decay branching ratios in CEPC experiment is presented, in the scenario of an integrated luminosity of 5000 fb<sup>-1</sup> is assumed in the analysis, corresponding to 10 years of data taking with nominal luminosity of  $2 \times 10^{34}$  cm<sup>-2</sup>s<sup>-1</sup> at the CEPC. In the study the Higgs boson is produced in association with a pair of leptons, dominantly following the Higgsstrahlung process. The statistic uncertainty is estimated around 1% while the systematic uncertainty are also discussed. This study demonstrate the potential of measuring Higgs boson hadronic decay in CEPC, and will provide key information to understand the Yukawa coupling between Higgs and quarks.

*Keywords:* `elsarticle.cls`, L<sup>A</sup>T<sub>E</sub>X, Elsevier, template

*2010 MSC:* 00-01, 99-00

---

---

<sup>☆</sup>Fully documented templates are available in the elsarticle package on CTAN.

<sup>\*</sup>Corresponding author

*Email address:* `support@elsevier.com` (Global Customer Service)

*URL:* `www.elsevier.com` (Elsevier Inc)

<sup>1</sup>Since 1880.

## 1. Introduction

The discovery of a scalar boson with mass around 125 GeV at LHC [1, 2] completed the final piece of the standard model. This particle, interpreted as the Higgs boson, plays a crucial role in the electroweak spontaneous symmetry broken (EWSB), known as the BEH (Brout-Englert-Higgs) mechanism [3, 4, 5], often referred as the Higgs mechanism. The Higgs mechanism allows the  $W$ ,  $Z$ , quarks and charged leptons to be massive while keeping the  $SU(2)_L \times U(1)_Y$  gauge invariance. The masses of the SM fermions ( $m_{f_i}$ ) in the SM are proportional to their Yukawa couplings ( $h_i$ ) to the Higgs field:  $m_{f_i} = \frac{vh_i}{\sqrt{2}}$ , where  $v \approx 246$  GeV is the vacuum expectation value of the Higgs field (VEV).

Measuring the Yukawa couplings between higgs and SM fermions is essential to understand the origin of the fermions' masses and the detail of EWSB. The dominant direct higgs decays into fermionic final states are  $H \rightarrow b\bar{b}$ ,  $H \rightarrow c\bar{c}$  and  $H \rightarrow \tau^+\tau^-$ , the decay branching ratios of which are predicted to be 57 %, 2.7 % and 6 %. In addition, the Higgs boson can decay to gluon pairs via a heavy quark loops. The large coupling between higgs and top quark leads to considerably large branching ratio of  $H \rightarrow gg$  that is estimated to be about 9 %.

Until now, the LHC is the only collider to directly study the Higgs mechanism. The leading higgs fermionic decay,  $H \rightarrow b\bar{b}$  was studied in both ATLAS and CMS experiment in VH[6, 7], ttH[8] and VBF[9, 10] process, with the LHC Run-I data. The combination of ATLAS and CMS gives  $b\bar{b} \sigma \times Br$  signal strength for  $0.70 \pm 0.29$  in run-I data[11]. A study on  $H \rightarrow b\bar{b}$  in which Higgs are produced in association with  $W$  or  $Z$  boson, using 36.1 fb $^{-1}$  Run-II data with center of mass 13 TeV in ATLAS, constrained the signal strength as  $1.20^{+0.24}_{-0.23}(stat.)^{+0.34}_{-0.28}(sys.)$  [12] with 3.5  $\sigma$  signal significance. The large uncertainty is due to huge QCD or vector boson production with multi-jets backgrounds, which is inevitable in hadron colliders.

The lepton collider is the proper option for high precision measurement as the result of it's clean background and fixed center-of-mass frame. In addition, modern techniques of particle detection and reconstruction improves the preci-

sion of Several future lepton colliders have been proposed with the capability of precise measurement of Higgs boson parameters, like International Linear Collider(ILC),  $e^+e^-$  Future Circular Collider (FCC-ee or TLEP) and Circular Electron Positron Collider(CEPC). The CEPC can operate with center of mass  
35 energy at 240 - 250 GeV and instantaneous luminosity of  $2 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ . 5000  $\text{fb}^{-1}$  collision data will be accumulated, including one million Higgs events, after 10 years of running of CEPC. The Higgs boson are produced mainly via Higgsstrahlung process(96.6%).  $WW$ -fusion and  $ZZ$ -fusion has the fraction of 3.06% and 0.29% in Higgs production.

## 40 2. Detector and MC Sample

### 2.1. Detector

ILD-like detector is designed as the CEPC detector(CEPC-v1) with additional considerations[13]. A vertex detector (VTX) with high pixel resolution is located in the inner most part of the detector. It provides the inner tracks measurement with high spatial resolution, which is key information for track impact parameter (IP) measurements and vertex reconstruction. The heavy flavor jet tagging significantly relies on the capability on IP measurements and vertex reconstruction with high precision. The 6 layers of sensors are laid 16 to 60 mm in radius with 97% to 90% in azimuth angular coverage. The signal layer spatial resolution is  $2.8 \mu\text{m}$  in the 2 inner layers and  $4 \mu\text{m}$  in the 4 outer layers. The overall IP resolution can be estimated as

$$\sigma(r\phi) = a \oplus \frac{b}{p(\text{GeV}) \sin^{3/2} \theta} \mu\text{m} \quad (1)$$

, in which  $a = 5$  and  $b = 10$ .

A Time Project Chamber(TPC) is laid out of VTX to take the major task of track measurement. It covers the solid angle up to  $\cos \theta = 0.98$ . When being  
45 operated in 3.5 T field, the momentum resolution is  $\sigma(1/p_T) = 10^{-4}$  GeV.

A Particle Flow Algorithm-oriented[14] calorimeter system, combined by the electromagnetic calorimeter(ECAL) and hadronic calorimeter(HCAL), was de-

signed with high granularity and precise energy measurement of electrons, photons, taus and hadronic jets. The resolution of ECAL and HCAL are about  
50  $16\%/\sqrt{E(\text{ GeV})}$  and  $50\%/\sqrt{E(\text{ GeV})}$ . The energy resolution of jets from Higgs or  $W^\pm/Z$  decay is estimated about  $\sigma_E/E = 3 - 4\%$ . With granularity better than  $1\text{ cm} \times 1\text{ cm}$  of each cell, the hadrons in jets can be well separated.

The muon system is mounted as the outermost part in the detector. The baseline design of muon detector requires 94% reconstruction efficiency of muons  
55 with energy higher than 6 GeV. The longitudinal and transverse position resolution are required to be  $\sigma_z = 1.5\text{ cm}$  and  $\sigma_{r\phi} = 2.0\text{ cm}$ . The rate of pions mis-identified as muons at energy 30 GeV is required to be less than 1%.

## 2.2. MC Samples and Jets Reconstruction

In this analysis, the signal events are  $l^+l^-H \rightarrow l^+l^- + b\bar{b}/c\bar{c}/gg$ , thus the final  
60 states contains a pair of charged leptons with opposite flavor and two jets. The standard model background includes di-quark events, di-lepton events, vector boson pair production and higgs production with final states different from the signal. Both background and signal events are generated using Whizard[15] configured as no-polarization electron-positron collision at center of mass energy  
65 of 250 GeV. PHYTHIA [16] was implemented to simulate the fragmentation and hadronization. The higgs mass was assumed to be 125 GeV and the coupling was set as that predicted by standard model.

The generated events undergo the detector simulation by Mokka[17], a GEANT4[18]  
based detector simulator. The simulated hits were digitized and reconstructed  
70 with ArborPFA[19].

Jets reconstruction and flavor tagging are essential to this analysis. These tasks are done with the LCFIPLUS [20] toolkit, integrating the functionality of doing vertex finding, jet reconstruction and jet flavor tagging. Flavor tagging will be discussed in section 4.1. Jets are reconstructed by Durham algorithm[21].  
75 This algorithm begins with jet cluster candidates, which are either single reconstructed particles, or compound objects like reconstructed secondary vertices. The procedure iteratively pairs the clusters by selecting the minimum distance

measure, defined as  $y_{i,j} = \min\{E_i^2, E_j^2\}(1 - \cos(\theta_{ij}))/E_{vis}^2$ , where  $E_i$  and  $E_j$  are the energy of  $i$ -th and  $j$ -th cluster, and  $\theta_{ij}$  refers to the angle between them.  $E_{vis}$  are the sum of energy of all the clusters. Clusters with minimum  $y_{ij}$  are merged, reducing the cluster number by 1 until the remaining clusters number equals to the required jet multiplicity. This algorithm gives a series of  $y_{ij}$  value:  $y_{12}, y_{23}, y_{34}$  etc. When  $i$  is larger than the real cluster multiplicity,  $y_{ij}$  reflect the distance of two clusters which are actually from the same cluster, resulting a small  $y_{ij}$  value. Thus  $y_{ij}$  is an indicator of the jet cluster multiplicity.

### 3. Event selection and inclusive higgs decay measurement

Three steps of analysis are taken. In the first step, an inclusive study of  $ZH$  production in  $l^+l^-H$  channel was done. A series of cuts was applied to the lepton pairs. The  $ZH$  events number are extracted by fitting the lepton pair recoil mass spectrum. In the second step, the cuts on jets information are applied, to further suppress the backgrounds. In the last stage, a fit was applied on both the flavor tagging information and lepton pair recoil mass, to extract the events number of  $ZH \rightarrow l\bar{l} + b\bar{b}/c\bar{c}/gg$ . The higgs decay branch ratio was calculated by dividing the individual  $ZH \rightarrow l\bar{l} + b\bar{b}/c\bar{c}/gg$  number by the inclusive  $ZH \rightarrow l^+l^-H$  number.

#### 3.1. $l^+l^-H$ Event Selection

The  $l^+l^-H$  channel is the composite of two sub-channels of  $e^+e^-H$  and  $\mu^+\mu^-H$  process. The dominant backgrounds are semi-leptonic  $ZZ$  process and other  $ZH$  production followed by other types of higgs decay (mainly by higgs decay to off-shell  $W$  or  $Z$  pair, and either of the two intermediate bosons undergo hadronic decay).

Two isolated tracks with opposite charge, reconstructed as electrons or muons, are required in addition to a pair of jets. If there are additional isolation leptons, a probability was calculated according to a 2-dimension PDF of  $l\bar{l}$  invariant mass and recoil mass. The lepton pair with highest PDF in  $l\bar{l}$  was assumed to be from  $Z$  boson decay.

The azimuth  $\mu^+\mu^-$  system are required to in the range of  $|\cos\theta_{\mu^+\mu^-}| < 0.81$ . The angle between the two isolation tracks  $\psi$  is required to satisfy  $\cos\psi > -0.93$ . The invariant mass of  $\mu^+\mu^-$  are required in  $Z$ -mass window, which is between  
110 77.5 GeV and 104.5 GeV. The recoil mass of  $\mu^+\mu^-$  is required to be between  
124 GeV and 131.5 GeV. At this stage, the inclusive  $ZH \rightarrow l^+l^-H$  can be extracted by fitting the recoil mass spectrum.

Then there cuts applied on objects other than the lepton pair. Events with extra isolation leptons are rejected. Two jets are required, each's azimuth angle  
115  $\theta_{jet}$  is required in the range of  $|\cos\theta_{jet}| < 0.96$ . The two jets combined energetic particles<sup>2</sup> multiplicity are required to be larger than 20. The jet pair invariant mass are required in the range between 75 GeV and 150 GeV. To reject the background events from other higgs decay, y-value cut was set to suppress events with jet multiplicity other than 2. The signal and dominant backgrounds events  
120 yields after applying cuts are summarized in table 3.1 for  $\mu^+\mu^-H$  and  $e^+e^-H$  analysis respectively.

#### 4. Measurement of $H \rightarrow b\bar{b}/c\bar{c}/gg$

After applying cuts on jets, the  $H \rightarrow b\bar{b}/c\bar{c}/gg$  are mixed. Flavor tagging is necessary to distinguish different flavor final states. Based on the high per-  
125 formance flavor tagging algorithm, a flavor tagging template fit method is used to simultaneously extract the event number of each flavor final state.[22] These template fits are cooperated with lepton pair recoil mass fit, making the dominant background estimation model independent. In the following text, the flavor tagging techniques, flavor tagging template fit and the 3D flavor tagging-recoil  
130 mass fit will be described.

##### 4.1. Flavor tagging

A TMVA based algorithm are implemented to distinguish the jets' flavor. The reconstructed jets are categorized according to the secondary vertex multi-

---

<sup>2</sup>Defined as particles with energy larger than 0.4 GeV.

$\mu^+\mu^-H \rightarrow \mu^+\mu^- + b\bar{b}/c\bar{c}/gg$ Channel					
	Signal	$\mu^+\mu^-H$ Background	Other Higgs Background	Irreducible SM Background	other SM Background
Original	$2.45 \times 10^4$	$1.10 \times 10^4$	$1.01 \times 10^6$	$1.05 \times 10^6$	$4.96 \times 10^8$
Lepton Cut	$1.51 \times 10^4$	$6.56 \times 10^3$	227	$1.09 \times 10^4$	$2.79 \times 10^4$
Jets Cuts	$1.34 \times 10^4$	$1.80 \times 10^3$	108	$7.75 \times 10^3$	29.4
$e^+e^-H \rightarrow e^+e^- + b\bar{b}/c\bar{c}/gg$ Channel					
	Signal	$e^+e^-H$ Background	Other Higgs Background	Irreducible SM Background	other SM Background
Original	$2.63 \times 10^4$	$1.17 \times 10^4$	$1.01 \times 10^6$	$1.62 \times 10^6$	$4.95 \times 10^8$
Lepton Cut	$9.17 \times 10^3$	$3.53 \times 10^3$	128	$9.00 \times 10^3$	$7.11 \times 10^4$
Jets Cuts	$7.97 \times 10^3$	986	58.0	$6.46 \times 10^3$	67.4

Table 1: Cut flow of  $\mu^+\mu^-H$  and  $e^+e^-H$  channel.

plicity. In each category, a b-tagging and a c-tagging training with BDT method  
135 was implemented over  $Z$ -pole di-jet events, employing variables including jets  
kinematic variables, tracks' impact parameters and secondary vertex information.  
The training output gives a b-jet likeness weight and a c-jet likeness weight for  
each jet, representing the resemblance of the jet to a b-jet or a c-jet.

#### 4.2. Template fit

The b/c-weight likeness of the two individual jets from higgs decay, say  
 $L_{b1/c1}$  and  $L_{b2/c2}$ , can be used to construct the combined B/C likeness, defined  
as:

$$X_{B/C} = \frac{L_{b1/c1}L_{b2/c2}}{L_{b1/c1}L_{b2/c2} + (1 - L_{b1/c1})(1 - L_{b2/c2})}.$$

140 The conservation of quark flavor in higgs decay guarantee that  $X_B(X_C)$  is close  
to 1 if higgs decay to  $b\bar{b}(c\bar{c})$  while close to 0 in other case. The distribution of  
various process in  $X_B - X_C$  space are taken as the flavor PDFs templates.

#### 4.3. Three dimension fit on flavor templates and recoil mass

A set of 3-dimension PDFs are defined by the production of the flavor PDF and lepton pair recoil mass PDFs:

$$PDF^{3D}(X_B, X_C, M_{recoil}) = PDF^{flavor}(X_B, X_C) \times PDF^{recoil\_mass}(M_{recoil})$$

The recoil mass PDFs of  $l^+l^-H$  events was described by a crystal ball function plus a double sided exponential function, while the recoil mass of dominant background events are described by a first order Chebychev polynomial function;  $PDF^{flavor}$  is the flavor PDFs defined in section 4.2. There are 4 parameters for crystal-ball shape and 2 parameters for polynomial, in addition, there are 9 parameters for signal and background normalizations. The shape parameters of recoil mass PDFs are free in the fit. The normalization parameters for signals,  $ZZ \rightarrow \mu^+\mu^- + q\bar{q}(\mu^+\mu^-H$  channel) and  $e^+e^-Z \rightarrow e^+e^- + q\bar{q}(e^+e^-H$  channel) also float, while normalization parameters for  $H \rightarrow WW^*$  and  $H \rightarrow ZZ^*$  are fixed. Both shape and normalization parameters for other backgrounds are fixed. The shape of those minor backgrounds are taken from MC simulation in the form of a 3-dimension histogram in  $X_B-X_C-M_{recoil}$  space. In figure 4.3, the model from fit are compared to pseudo-data, project to lepton pair recoil mass,  $X_B$  and  $X_C$ .

### 5. Uncertainties of Measurements

The statistic uncertainty was estimated by applying the toyMC method. These procedure includes 2000 to 5000 iterations. In each iteration, the 'data' is filled in a 3D histogram in  $X_B-X_C-recoilmass$  space. Then, in each bin of the histogram, the event yields fluctuated according to a Poisson distribution. The 3D fit is applied to the fluctuated 'data'. The dispersion of fitted signal event yields in signal region represents the statistic uncertainty. The results of toyMC test for  $H \rightarrow b\bar{b}$ ,  $H \rightarrow c\bar{c}$  and  $H \rightarrow gg$  are represented in Fig 5, in terms of standard deviation of fitted signal strength. The statistic uncertainty of fitted signal strength for  $H \rightarrow b\bar{b}$ ,  $H \rightarrow c\bar{c}$  and  $H \rightarrow gg$  are estimated to be 1.11%,



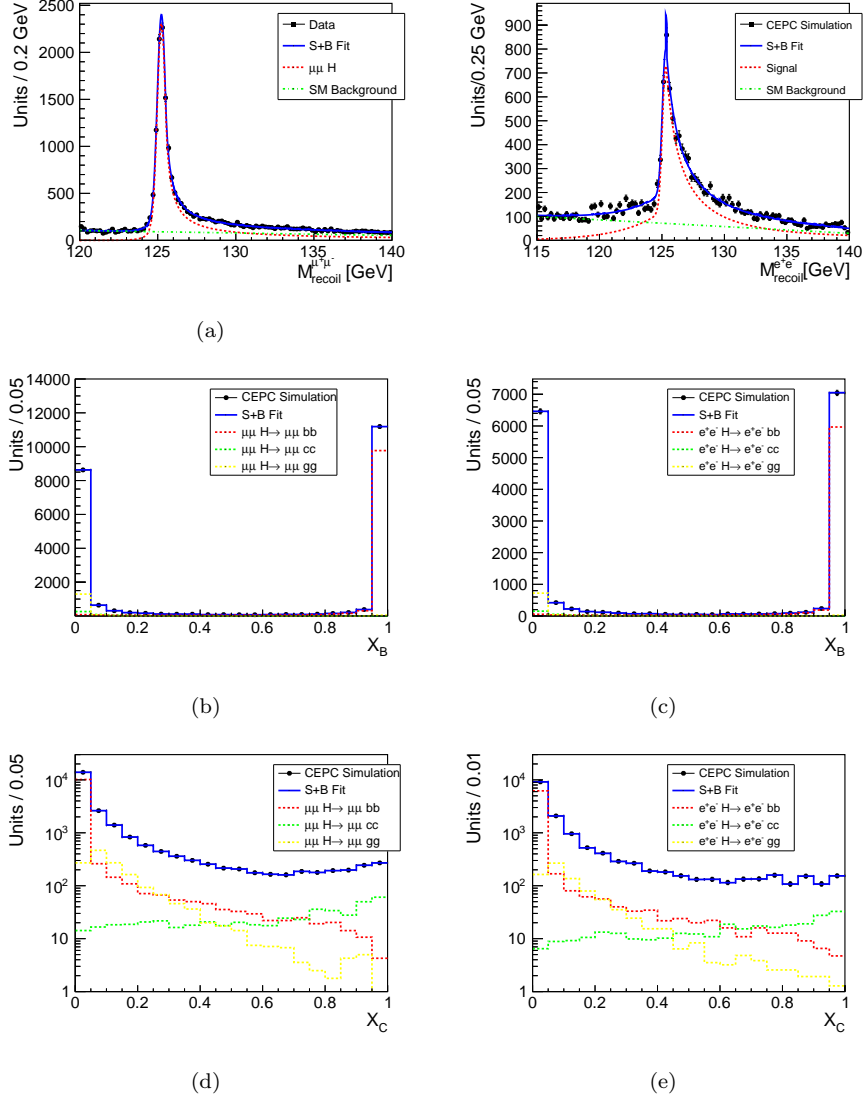


Figure 1: 3D-fit result project in three dimension: (a) fit result project on recoil mass distribution in  $\mu^+\mu^-H$  channel, (b) fit result project on recoil mass distribution in  $e^+e^-H$  channel, (c) fit result project on B-likeness distribution in  $\mu^+\mu^-H$  channel, (d) fit result project on B-likeness distribution in  $e^+e^-H$  channel, (e) fit result project on C-likeness distribution in  $\mu^+\mu^-H$  channel, (f) fit result project on C-likeness distribution in  $e^+e^-H$  channel.

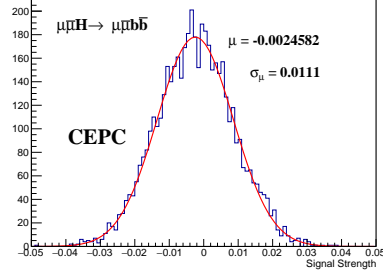
10.5% and 5.44% respectively. The statistic uncertainty of inclusive higgs decay is estimated in similar way, except for that the toyMC sample was generated by  
170 fluctuate lepton one-dimension recoil mass spectrum, instead of three-dimension flavor-recoil mass distribution.

The uncertainties luminosity, lepton cut efficiencies,  $Z \rightarrow \mu^+\mu^-/e^+e^-$  modeling and ISR correction are also included in  $\sigma_{e^+e^-H \rightarrow e^+e^-+b\bar{b}/c\bar{c}/gg}$  and  $\sigma_{\mu^+\mu^-H \rightarrow \mu^+\mu^-+b\bar{b}/c\bar{c}/gg}$ . However, when measure the branch ratio of  $H \rightarrow$   
175  $b\bar{b}/c\bar{c}/gg$ , these uncertainties are also included in inclusive  $H$  decay so they will be canceled. The remaining systematic sources include uncertainty from fit method, uncertainty of jet selection efficiencies, uncertainty from flavor templates and uncertainty due to non-uniformity in isolation lepton selection efficiency.

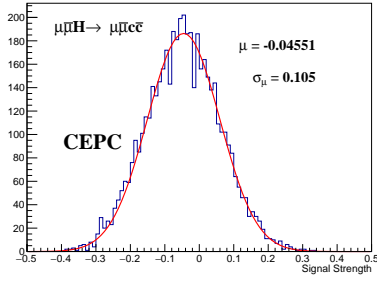
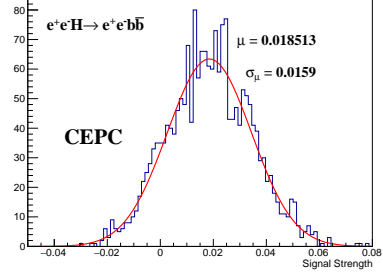
180 The fit method have two types of systematic uncertainty. The first kind of uncertainty is due to imperfect modeling of the PDFs. The PDFs include the recoil mass modeling and flavor template. The latter will be discussed as the systematic of flavor tagging and here we only focus on the recoil mass modeling. The toyMC samples used in statistic uncertainty study are used to estimate  
185 this type of systematic uncertainty. The difference between the central value of fitted signal strength in toyMC test from the MC prediction are taken as the systematic uncertainty due to the modeling.

The other kind of systematic uncertainty in the fit comes from the uncertainty of we fixed parameters. The normalization parameters for  $H \rightarrow WW^*$   
190 and  $H \rightarrow ZZ^*$  backgrounds are fixed in the fit. We set these normalization parameters 5% higher and lower to find their impact on the fitted signal yields. We conservatively vary the yields of backgrounds other than  $l^+l^-H$  and  $ZZ \rightarrow l^+l^- + q\bar{q}$  by 100% to estimate the systematic uncertainty of fixing them.

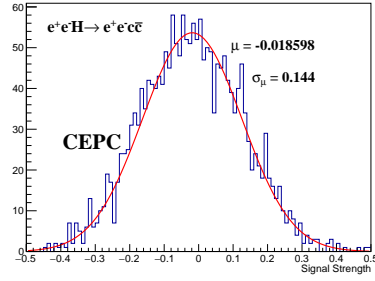
Systematic uncertainty on lepton veto efficiency for  $H \rightarrow b\bar{b}$  can be estimate  
195 in  $Z - pole\ b\bar{b}$  events. With one jet tagged as  $b - jets$ . The efficiency of lepton veto can be studied with the precision of 0.0002% with 2 billion  $b\bar{b}$  events in the  $Z$ -pole sample. . Thus we take 0.002% as the systematic uncertainty on lepton veto. For  $H \rightarrow c\bar{c}$  with similar method the uncertainty of lepton veto efficiency



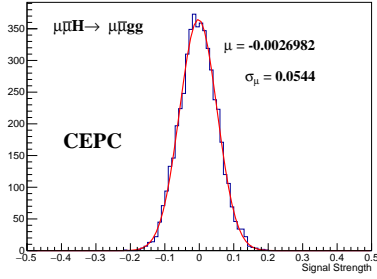
(a)



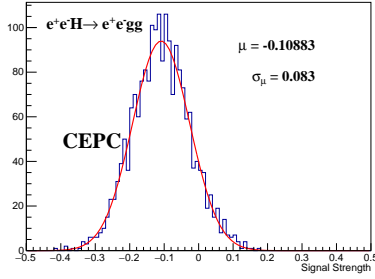
(b)



(c)



(d)



(e)

Figure 2: Toy MC test result in terms of signal strength and uncertainty in signal region from template fit.

is estimated as 0.0001%. For  $H \rightarrow gg$  we assume the gluon jets sample yielding  
200 1% of that of  $b\bar{b}$ , and estimate the uncertainty of lepton veto to be 0.008%. The  
systematic uncertainties of jet PFO multiplicity, jet  $\cos\theta$  cut and  $y$ -th value  
cut was estimated in similar way, such that by assuming these variables can be  
calibrated with the  $Z$ -pole sample.

The systematic uncertainty of the efficiency on jet pair invariant mass cut  
205 can be estimated from the jet energy resolution. We apply a smearing on jet  
pair mass distribution according to a gaussian distribution corresponding to the  
jet energy resolution. We take the value of 4% as the jet energy resolution from  
CEPC pre-CDR[13] and get the uncertainty for  $H \rightarrow b\bar{b}$ ,  $H \rightarrow c\bar{c}$  and  $H \rightarrow gg$   
are  $\begin{smallmatrix} +0.68\% \\ -0.20\% \end{smallmatrix}$ ,  $\begin{smallmatrix} +0.43\% \\ -1.08\% \end{smallmatrix}$  and  $\begin{smallmatrix} +0.71\% \\ -1.68\% \end{smallmatrix}$  respectively.

210 Since flavor tagging method are implemented via flavor template fitting,  
the flavor tagging systematic uncertainty is directly caused by the difference  
between templates from MC prediction and templates in data. To evaluate such  
difference, delicate flavor tagging commissioning and calibration are demanded.  
Although there is no such commissioning or calibration have been done yet,  
215 we can briefly estimate the systematic uncertainty by assuming a difference  
between data and MC. This difference are studied in terms of their impact on  
the  $H \rightarrow b\bar{b}/c\bar{c}/gg$  branch ratio measurement.

More than 80% of the  $b\bar{b}$  events concentrated in the region with highest b-  
likeness(b-likeness>0.95) and lowest c-likeness(c-likeness<0.05). A deviation of  
220 the fraction of the contents in this bin was simulated, and the impact of this  
deviation to  $b\bar{b}$ ,  $c\bar{c}$  fraction of any dataset can be evaluated as a function of this  
deviation. This function of a control sample with known  $b\bar{b}$  and  $c\bar{c}$  fraction can  
be used to estimate the deviation, while the function of the dataset in signal  
region can be used to estimate the systematic uncertainty due to this kind of  
225 deviation. We select the  $ZZ \rightarrow q\bar{q} + \mu^+\mu^-$  as the control sample. This control  
sample has the purity upto 99.65%, with statistic around 200000 of  $b\bar{b}$  events  
and  $c\bar{c}$  events, and 600000 of light jet pair events. The kinematic distribution  
of  $ZZ \rightarrow q\bar{q} + \mu^+\mu^-$  is similar to the signal events, so the deviation in flavor  
tagging are assumed to be the same for the  $ZZ \rightarrow q\bar{q} + \mu^+\mu^-$  control sample and

	$\mu^+\mu^-H$			$e^+e^-H$		
	1.11%	10.5%	5.44%	1.59%	14.4%	8.3%
	$H \rightarrow b\bar{b}$	$H \rightarrow c\bar{c}$	$H \rightarrow gg$	$H \rightarrow b\bar{b}$	$H \rightarrow c\bar{c}$	$H \rightarrow gg$
Fixed Background	-0.17% +0.06%	+4.1% -4.2%	7.6%	-0.17% +0.06%	+4.1% -4.2%	7.6%
Event Selection	+0.68% -0.20%	+0.43% -1.1%	+0.71% -1.7%	+0.68% -0.20%	+0.43% -1.1%	+0.71% -1.7%
Flavor Tagging	0.67%	10.4%	1.1%	0.67%	10.4%	1.1%
Non uniformity	0.016%			0.016%		
Combined	+0.96% -0.72%	+11.2% -11.3%	+7.7% -7.9%	+0.96% -0.72%	+11.2% -11.3%	+7.7% -7.9%

Table 2: Uncertainties of  $H \rightarrow b\bar{b}$ ,  $H \rightarrow c\bar{c}$  and  $H \rightarrow gg$ .

the signal events. The systematic uncertainty is determined by the precision of the knowledge on the flavor components in  $ZZ \rightarrow q\bar{q} + \mu^+\mu^-$ . The uncertainty of flavor components fraction are largely determined by  $ZZ$  control sample statistics. Had we use much larger sample (like  $Z$ -pole hadronic events), and have a better understanding on the relationship between the flavor tagging variables and kinematic feature, the uncertainty will be greatly reduced.

The non-uniformity of individual higgs decay channels mainly comes from leptonic or semi-leptonic higgs decay, which happens with diboson as intermediate states. Since there are extra isolation leptons, there are higher chance for these events to have at least two isolation leptons. On the other hand, the final states requires at least two particles as jet candidates, which will reduce the chance of leptonic process with neutrinos in the final states to pass the selection. Since the efficiency difference of the signal events and inclusive  $l^+l^-H$  are mainly due to  $H \rightarrow WW^*$  and  $H \rightarrow ZZ^*$  events according to table ??, the uncertainty was estimated by compare the inclusive higgs efficiency with over estimated and under estimated  $H \rightarrow WW^*/ZZ^*$  fraction by 5%.

The systematic uncertainty estimation are summarized in table 5.

## 6. Conclusion

The  $H \rightarrow b\bar{b}/c\bar{c}/gg$  cross-section was studied in  $\mu^+\mu^-H$  and  $e^+e^-H$  process with simulation data corresponding to  $5000 \text{ fb}^{-1} e^+e^-$  collision at  $\sqrt{s} =$   
250 250 GeV in CEPC experiment. The statistic uncertainty in  $b\bar{b}$ ,  $c\bar{c}$  and  $gg$  measurement are around 1%, 10% and 5% respectively. The systematic uncertainty are also studied which are the same level of statistic uncertainty. This study demonstrate the capability of making precision measurement of Higgs Yukawa coupling with quarks.

## 255 References

- [1] Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC, Phys. Lett. B 716 (2012) 1–29. doi:10.1016/j.physletb.2012.08.020.
- [2] Observation of a new boson at a mass of 125 GeV with the CMS experiment  
260 at the LHC, Phys. Lett. B 716 (2012) 30–61. doi:10.1016/j.physletb.2012.08.021.
- [3] P.W.Higgs, Broken Symmetries and Masses of Gauge Bosons, Phys.Rev.Lett 13 (1964) 508–509. doi:10.1103/PhysRevLett.13.508.
- [4] F.Englert, R.Brout, Broken Symmetry and Mass of Gauge Vector Mesons,  
265 Phys.Rev.Lett 13 (1964) 321–323. doi:10.1103/PhysRevLett.13.321.
- [5] G.S.Guralnik, C.R.Hagen, T.W.B.Kibble, Global Conservation Laws and Massless Particles, Phys.Rev.Lett 13 (1964) 585–327. doi:10.1103/PhysRevLett.13.585.
- [6] Search for the  $b\bar{b}$  decay of the Standard Model Higgs boson in associated (W/Z)H production with the ATLAS detector, JHEP 01 (2015) 069.  
270 arXiv:1409.6212, doi:10.1007/JHEP01(2015)069.

- [7] Search for the standard model Higgs boson produced in association with a W or a Z boson and decaying to bottom quarks, Phys.Rev.D 89 (2014) 012003. [arXiv:1310.3687](#), [doi:10.1103/PhysRevD.89.012003](#).
- 275 [8] Search for a standard model Higgs boson produced in association with a top-quark pair and decaying to bottom quarks using a matrix element method, Eur.Phys.J.C 75 (2015) 251. [arXiv:1502.02485](#), [doi:10.1140/epjc/s10052-015-3454-1](#).
- 280 [9] Search for the Standard Model Higgs boson produced by vector-boson fusion and decaying to bottom quarks in  $s = 8$  TeV pp collisions with the ATLAS detector, JHEP 11 (2016) 112. [arXiv:1606.02181](#), [doi:10.1007/JHEP11\(2016\)112](#).
- [10] Search for the standard model Higgs boson produced through vector boson fusion and decaying to  $b\bar{b}$ , Phys.Rev.D 92 (2015) 032008. [arXiv:1506.1010](#),  
285 [doi:10.1103/PhysRevD.89.012003](#).
- [11] Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC  $pp$  collision data at  $\sqrt{s} = 7$  and 8 TeV, JHEP 08 (2016) 045. [arXiv:1606.02266](#), [doi:10.1007/JHEP08\(2017\)045](#).
- 290 [12] M. Aaboud, G. Aad, et al., Evidence for the  $H \rightarrow b\bar{b}$  decay with the ATLAS detector, JHEP12 2017 (024). [arXiv:arXiv:1708.03299](#), [doi:https://doi.org/10.1007/JHEP12\(2017\)024](#).
- [13] M.Ahmad, et al., CEPC-SPPC Preliminary Conceptual Design Report, Tech. Rep. IHEP-EP-2015-01, The CEPC-SPPC Study Group (2015).
- 295 [14] J.Brient, Improving the jet reconstruction with the particle flow method: An introduction.
- [15] W.Kilian, T.Ohl, J. Reuter, WHIZARD-simulating multi-particle processes at LHC and ILC, Eur.Phys.J.C 71 (2011) 1742. [doi:https://doi.org/10.1140/epjc/s10052-011-1742-y](#).

- 300 [16] T.Sjöstrand, et al., Pythia 6.4 physics and manuel, JHEP05 2006 (026).  
arXiv:hep-ph/0603175, doi:10.1088/1126-6708/2006/05/026.
- [17] P. de Freitas, H.Videau, Detector simulation with MOKKA/GEANT4:  
Present and future (2003) 623–627LC-TOOL-2003-010.
- [18] GEANT4: A simulation toolkit, Nucl.Instrum.Meth. 506 (2003) 250–303.
- 305 [19] M. Ruan, Arbor, a new approach of the Particle Flow AlgorithmarXiv:  
1502.02485.
- [20] T. Suehara, T. Tanabe, LCFIPlus: A framework for jet analysis in linear  
collider studies, J.NIMA 808 (2015) 109–116. arXiv:1506.08371, doi:  
10.1016/j.nima.2015.11.054.
- 310 [21] S.Catani, Y.L.Dokshitzer, M.Olsson, G.Turnock, B.Webber, New cluster-  
ing algorithm for multi-jet cross-sections in  $e^+e^-$  annihilation, Phys.Lett.B  
269 (1991) 432–438. doi:10.1016/0370-2693(91)90196-W.
- [22] A. M. Hiroaki Ono, A study of measurement precision of the higgs boson  
branching ratios at the international linear collider, Eur. Phys. J. C 73.  
315 doi:10.1140/epjc/s10052-013-2343-8.