CEPC Draft

Paper title

Version: x.y

To be submitted to: Journal name

Corresponding editor(s)

Comments are due by: Comments deadline

document created on July 11, 2017 from file `CEPCNote.tex`
cover page automatically created with `cepccover.sty`

CEPC NOTE

CEPC_ANA_HIG_2016_XXX

July 11, 2017

Measurement of $H \to b\bar{b}/c\bar{c}/gg$ Branch Ration from ZH Production with $\nu\bar{\nu}H$, l^+l^-H and $q\bar{q}H$ Final States in CEPC Experiment

CEPC Workgroup

Abstract

A couple of remarks about the paper front page:

- **Title:** Measurement of $Hboson \to b\bar{b}/c\bar{c}/gg$ Branch Ration from ZH Production with $\nu\bar{\nu}H$, $l^+l^-H$and $q\bar{q}H$Final States in CEPC Experiment
- **Author list:** it will be provided by the CEPC Collaboration, and will be made available on their website. On the front page, you should name "The CEPC Collaboration" as author.
- **Abstract:** A study on the measurement on $Hboson \to b\bar{b}/c\bar{c}/gg$ branch ratio, one of the benchmark measurements in CEPC experiment, is presented here. In the scenario concerned, the Higgs boson are produced associated with a Z boson, and subsequently undergo hadronic decay, while the $Zboson$ decays to neutrino pair, charged lepton pair or quark pair.

  An integral luminosity of 5000 fb^{-1} is assumed to estimate the signal and background yielding, corresponding to 10 years running of CEPC with nominal luminosity at 2×10^{34} $\text{cm}^{-2}\text{s}^{-1}$. A cut based analysis is applied to each analysis, combined with multi-variable method. The flavor information in final states are extracted by fitting each flavor components according to their templates. A Toy Monte-Carlo test was done to evaluate the statistic uncertainty. They systematic uncertainty was also disccussed and estimated. We conclude the measurement of $H \to b\bar{b}/c\bar{c}/gg$ reach the precision 0.2 %, 2 % and 3 %.

E-mail address: baiy@seu.edu.cn

Contents

1 Introduction

The discovery of a scalar boson with mass around 125 GeVat LHC [1, 2] completed the fundamental particles list in the standard model. This particle, interpreted as the higgs boson, plays a lead role in the electroweak spontaneous symmetry broken(EWSB), known as the BEH(Brout-Englert-Higgs) mechanism [3, 4]. The higgs meachism guarantees that the W, Z, as well as the fermions like quarks and charged leptons can be massive in $SU(2)_L \times U(1)_Y$ gauge invariant way. The mass of the fermions m_{f_i} in standard model and their couplings to the higgs field h_i, so called Yukawa coupling, are related proportionally: $m_{f_i} = \frac{vh_i}{\sqrt{2}}$, in which v stands for the higgs field vaccume expectation(VEV), evalued around 246 GeV. Therefore measuring the Yukawa coupling between higgs and SM fermions is essential to undertand the origin of the fermions' masses and the detail of EWSB. The dominant higgs fermionic decay are expected to be $H \to b\bar{b}$(Br.57 %), $H \to c\bar{c}$(4 %) and $H \to \tau^+\tau^-$(3 %). In addition, the higgs can also decay to gluon pairs dominantly via a top loop diagram. The large coupling between higgs and top quark lead to considerable branch ratio of $H \to gg$(Br. 9 %).

Until now, the LHC is the only place to directly study the higgs experimentally. The leading higgs fermionic decay, $H \to b\bar{b}$ was studied in both ATLAS and CMS experiment in VH[5, 6], ttH[7] and VBF[8, 9] 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 ± 0.29 in run-I data[**?**]. The large uncertainty is due to huge QCD or vector boson production with muliti-jets backgrounds, which is inevitable in hadron colliders.

The Circular Electron Positron Collider(CEPC) [10] program is proposed with the goal to better understand the EWSB by precisely measuring on these higgs parameters as well as other EW parameters of interest. The CEPC has the advantages in precision measurement:

- Clean backgrounds
- Well defined frame of center momentum
- High luminosity

The works presented in this note demostrate the capability of the $H \to b\bar{b}/c\bar{c}/gg$ measurements in CEPC. The higgs productions associate with charged lepton(electrons or muons) pair, neutrino pair or quark pair are studied. In Section II, a brief introduction of CEPC experiment and the MC sample will be presented. In section III the event selection and the analysis strategy will be described. In section IV the results are listed and discussed. Detail information, auxiliary figures, tables and numbers, as well as analysis method in study can be found in appendix.

2 CEPC Experiment and MC Sample

2.1 CEPC experiment

The CEPC is a future ciruclar electron-positron collider project. Two detectors will be installed at two interaction points in the stoarge ring, 50-100 kilometers in circumference. Electrons and positrons collide at each interaction point with center of mass energy 240 - 250 GeV. The luminosity is design to be 2×10^{34} $cm^{-2}s^{-1}$. The higgs boson are produced mainly via associated production with Z boson(96.6%) as well as much WW fusion (with $\nu_e\bar{\nu}_eH$ in final states,3.06%) and ZZ fusion(0.29%, with e^+e^-H in final state). [10].

2.2 CEPC detector

A ILD-like detector is designed as the CEPC detector(CEPC-v1) with additional considerations[10]. Detail description of ILD model can be found in [11]. All changes need to be implemented into simulation, and iterate with physics analysis and cost estimation.

2.3 MC Samples

In this analysis, the signal events are $e^+e^- \rightarrow ZH \rightarrow q\bar{q} + b\bar{b}/c\bar{c}/gg$. 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[12] with assumption of collision by no-polarization electron-positron, with center of energy of 250 GeV. PHYTHIA 6.4 [13] was used to model 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 signal events and backgrounds events with higgs production undergo the GEANT4[14] based detector simulator Mokka[15] with CEPC-v1. The simulated hits were digitized and reconstructed by the MarlinReco package. The jets are reconstructed with Durham-like algorithm[16] by implementing the toolkit lcfiplus[17]. This toolkit is also capable to reconstruct the primary and secondary vertex, as well as jet flavor tagging.

Background events without higgs production undergo fast simulation, which includes:

- Four momentum of jet(b,c quarks and gluon) is smeared according to a Gaussian function, with jet energy resolution σ set to be 4%.
- Each leptonic track (e/μ) is corrected by momentum resolution and tracking efficiency, whose parameter are obtained from the study of the full simulation.
- The four momentum of neutrino decaying from the final hadron subtracted from the four momentum of jet.

The fast simulation reconstruct jets using a simplified algorithm. It is faster but with disadvantage that the vertex information is missing from reconstruct. Full simulation sample of the backgrounds are also in production and results will be updated using those samples. Detailed information of MC samples can be found in [18].

3 Analysis Strategy and Event Selection

Two steps of analysis are taken. In the first step, a series of object and event selection was applied to maximize the sensitivity to detect the signal; in the second step, a template fit on the 2-Dimension flavor weight distribution, which is got from TMVA based flavor tagging algorithm, is implemented to further distiguish the flavor components in final states. The details of template fit are described in 4.

The final states with 2 jets, 2 jets + 2 leptons(either electron or muon pair), 4 jets are required for $\nu\bar{\nu}H$,$l^+l^-H$and $q\bar{q}H$channel respectively.

3.1 l^+l^-H Event Selection

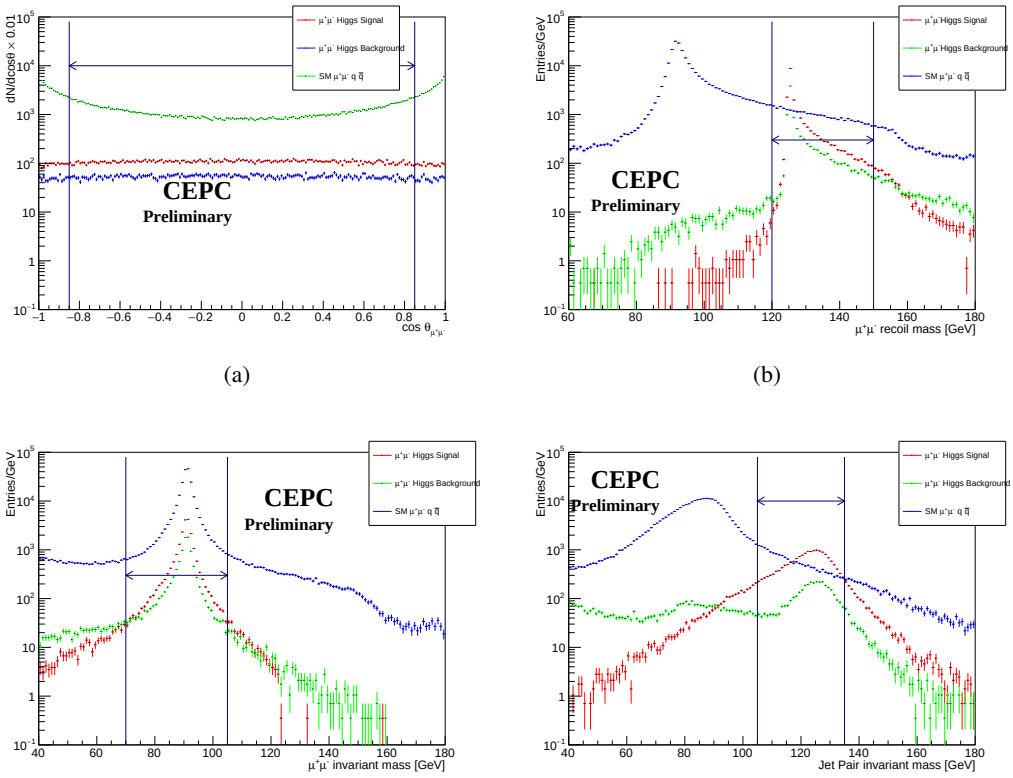
The l^+l^-H channel is the composite of two sub-channels of e^+e^-H and $\mu^+\mu^-H$ process. The dominant backgrounds are semi-leptonic ZZ process and other ZH production followed by other types of higgs decay(mainly by higgs decay to off-shell W or Z pair, and both intermediate boson undergo hadronic decay). Two isolated tracks with opposite charge, reconstructed as electrons or muons, are required in addition to a pair of jets. The recoil mass provide a very clear signature of the signal events, which have been delicately studied in [19]. The cut criteria of lepton pair recoil mass, together with the invariant mass of jet pair and lepton pair, are opptimized to maximum the signal significance. The polar angluar of the lepton pair recoil system are more concentrated in central region than the irreducible backgrounds from $t-$channel ZZ production. The distribution of recoil and invariant masses and lepton pair polar angluar distribution are presented in figure 3.1 and 3.1 for and e^+e^-H analysis. 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 yields after applying cuts are summarized in table 3.1 and 3.1 for $\mu^+\mu^-H$ and e^+e^-H analysis respectively.

Event Yields	Signal	$\mu^+\mu^-H$ background	SM $\mu^+\mu^-q\bar{q}$ process
$\sigma\times$Lumi	24532.3	10967.6	1051700
Object Selection	17563.7	9203.5	296779.5
0.85< $\cos\theta_{\mu^+\mu^-}$ <0.85	18580.8	8213.7	193043.4
120 GeV< $M_{\mu^+\mu^- recoil}$ <150 GeV	17953.2	7849.6	19711.2
70 GeV< $M_{\mu^+\mu^-}$ <105	17557.1	7255.4	17485.2
105 GeV< M_{JJ} <135 GeV	14392.6	2927.0	5769.2
y_{23} & y_{34} cut	13707.8	1575.2	5152.8

Table 1: Event Yields of signal and dominant backgrounds with cuts of $\mu^+\mu^-H$ channel, normalized to 5000 fb^{-1}.

Event Yields	Signal	e^+e^-H background	SM $e^+e^-q\bar{q}$ process
$\sigma\times$Lumi	26438.4	11918.6	1639129
Object Selection	21245.7	9192.0	296779.5
0.78< $\cos\theta_{e^+e^-}$ <0.78	14002.0	7200.8	372942.9
120 GeV< $M_{e^+e^- recoil}$ <160 GeV	13773.9	6581.0	167607.0
70 GeV< $M_{e^+e^-}$ <105	13143.0	6051.4	27027.4
105 GeV< M_{JJ} <135 GeV	9637.7	1935.3	6941.2
y_{23} & y_{34} cut	9148.4	1101.9	6081.4

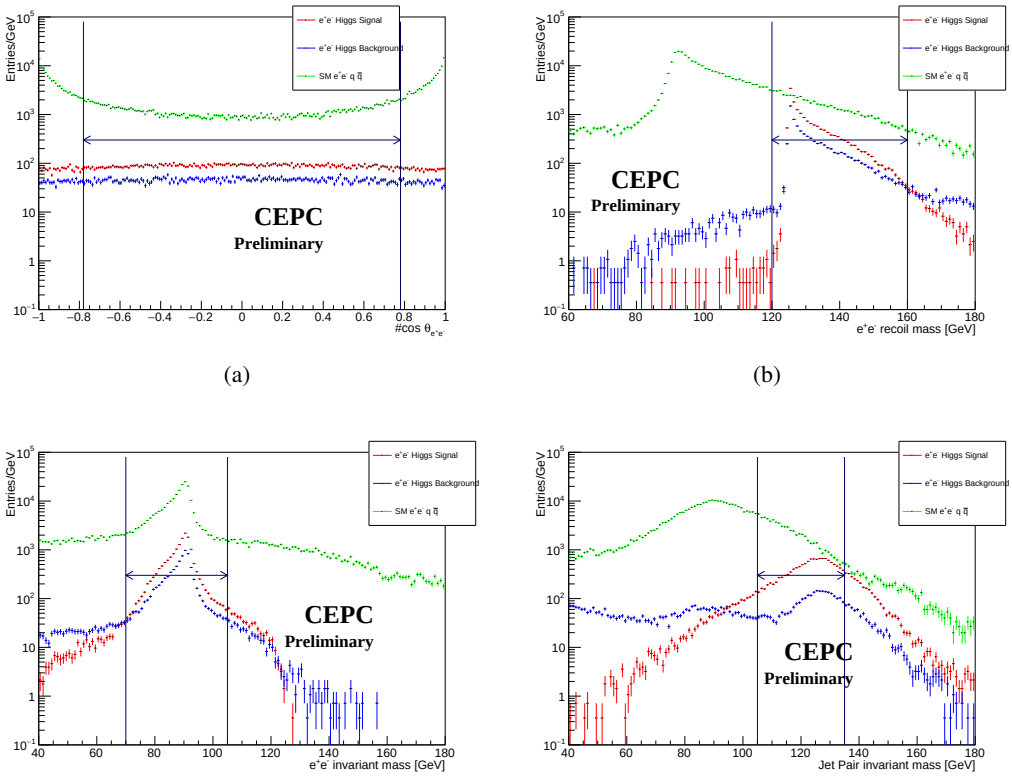
Table 2: Event Yields of signal and dominant backgrounds with cuts of e^+e^-H channel, normalized to 5000 fb^{-1}.



(a) (b)

(c) (d)

Figure 1: Distribution of $\mu^+\mu^-$ system polar angle(top left), $\mu^+\mu^-$ recoil mass(top right), $\mu^+\mu^-$ invariant mass(bottom left) and jet pair invariant mass(bottom right) for signal and dominant backgrounds in $\mu^+\mu^- H$ analysis.



(a) (b)

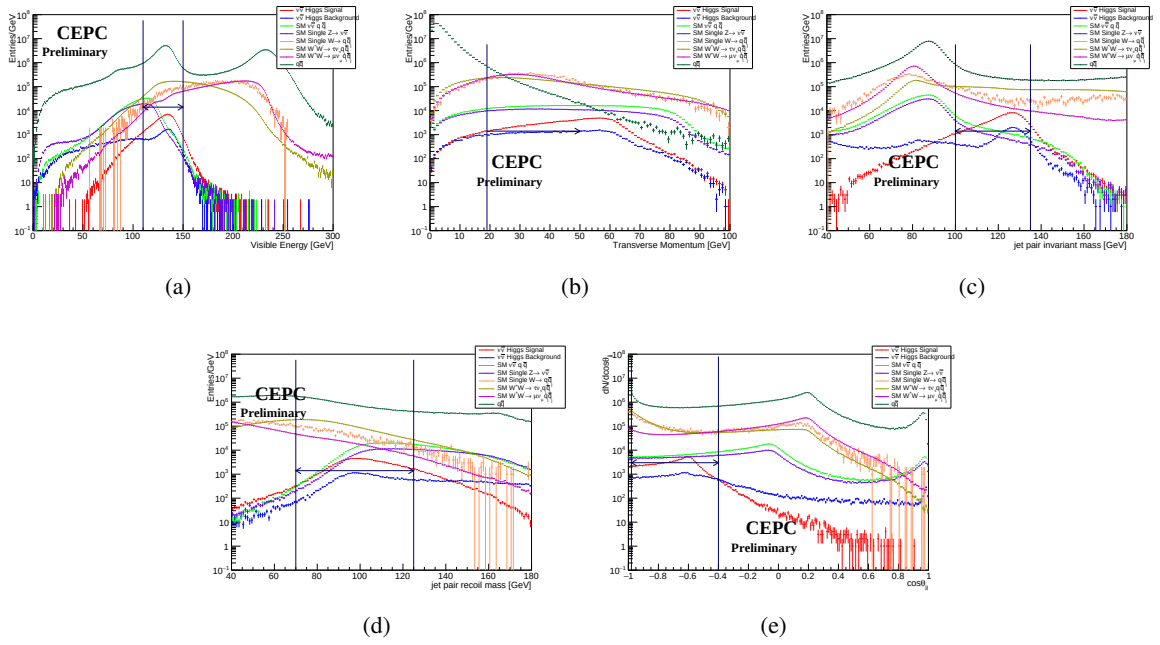
(c) (d)

Figure 2: Distribution of e^+e^- system polar angle(top left), e^+e^- recoil mass(top right), e^+e^- invariant mass(bottom left) and jet pair invariant mass(bottom right) for signal and dominant backgrounds in e^+e^-H analysis.

3.2 $\nu\bar{\nu}H$ Event Selection

The $\nu\bar{\nu}H$ channel includes ZH production followed by invisible $Z-$decay, or $t-$channel $W-$fusion process. The dominant backgrounds are quark pair production and gauge boson pair production, followed by hadronic and invisible decay of each boson. The observable particles in the signal form two energetic jets, initiated from two quarks of higgs decay. Thus isolation leptons are rejected, and a minimum number of PFOs are required the quark pair. The signal events are featured in the kinematic distribution of invisible section. The visible energy of the signal events are significantly lower than reducible SM backgrounds like semi-leptonic WW events. The visible transverse momentum are required to larger than 19 GeVto reject events, which tend to have low visible tranverse momentum due to high fraction of radiation return events. The invariant mass of the jets characterize the higgs production which is obverse discriminator against the non-higgs production SM events [1]. The angle between two jets and y-th values are also useful to distinguish signal events from SM backgrounds. Meanwhile the recoil system of these two jets, which is not directly observable, has the charaterstic of the Z boson associate with higgs in final states. The distribution of the above varibles and cut value can be found in figure 3.2 and 3.2 for signal and background events.

The Boost decision tree [20] method is implemented to the survived events from cut chain. The variables used in BDT are mentioned aboved: visible energy, transverse momentum, $y-th$ value, jet pair recoil and invariant mass. The events yields of signal and background in cutflow and BDT selection can be found in table 3.2.

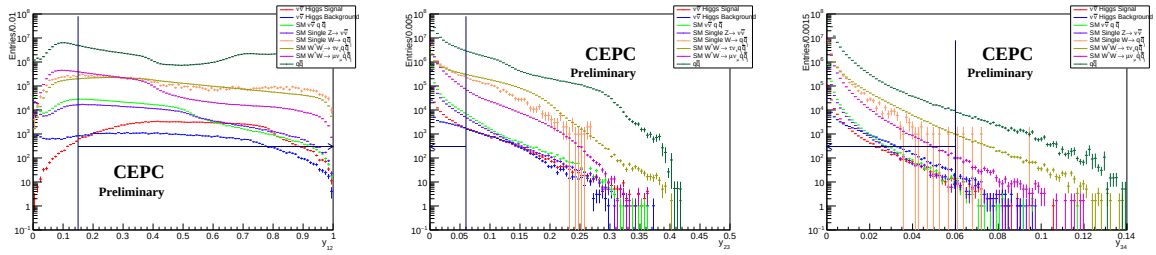


(a) (b) (c)

(d) (e)

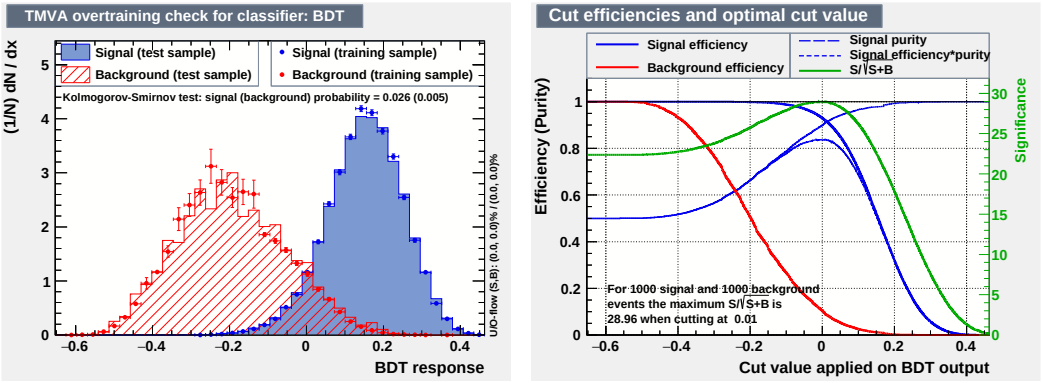
Figure 3: Visible energy(top left), visible transverse momentum(top middle), jet pair system invariant mass(top right), jet system recoil mass(bottom left) and jet pair system polar angluar(bottom right) distribution in $\nu\bar{\nu}H$ analysis.

[1] The peak of in signal region of events is due to radiation events.



(a) (b) (c)

Figure 4: The y_{12}(left), y_{23}(middle) and y_{34} right distribution in $\nu\bar{\nu}H$ analysis.

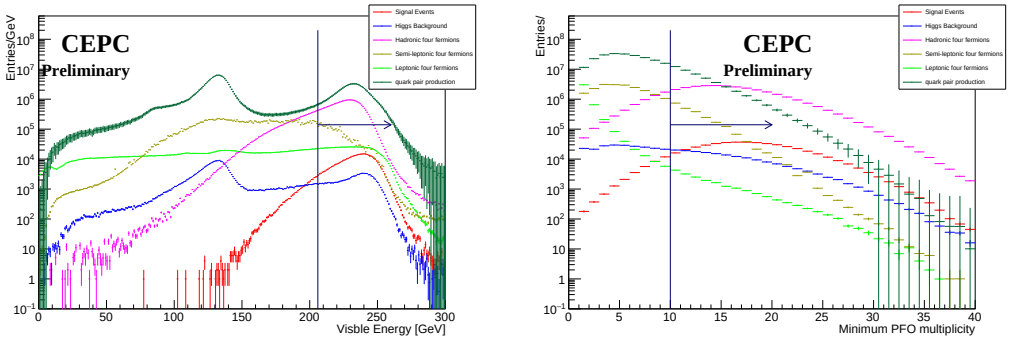


(a) (b)

Figure 5: Over training check of BDT distribution(left) and optimization of the BDT cut performance.

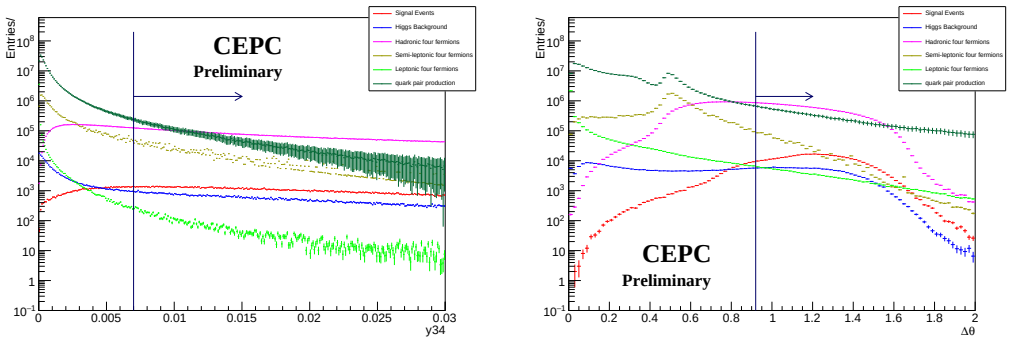
cutflow	signal	$\nu\bar{\nu}H$ bkg	zzsl	sznuqq	wwsltauq	wwslmuq	swslqq	qq
No Cut	170633	76604.1	1.08992×10^6	744218	1.19114×10^7	1.19114×10^7	1.30255×10^7	2.46847×10^8
2 jets in final state	170512	73227.4	1.08987×10^6	744201	1.19112×10^7	1.19112×10^7	1.30255×10^7	2.46842×10^8
NPFO	170349	42734.8	980878	656657	1.17604×10^7	1.16597×10^7	1.22262×10^7	2.39672×10^8
E_{total}	152374	33867.2	451233	250618	5.06253×10^6	1.27372×10^6	2.07027×10^6	1.01743×10^8
pT	142048	31579.8	413994	229568	4.31686×10^6	1.19619×10^6	1.93607×10^6	297012
IsoLep Veto	141112	27966	410719	227762	3.73815×10^6	365116	682854	294929
M_{inv}	134583	26165.2	41340.5	23255.3	2.20577×10^6	66320	336493	111687
M_{recoil}	125958	24817.4	37889.9	20720.1	1.75479×10^6	29908	237815	85653.4
y12	125228	24164.8	37138.4	20306.8	1.61702×10^6	27807.4	228934	83451.1
y23	126365	13478.1	29136.3	15976.7	1.07172×10^6	18577.8	126308	71353
y34	107347	5708.84	26728	14616.3	889531	16016	110520	69372.9
costheta	104023	5169.31	21169.6	11891.4	506063	9354.92	85850.1	48209.6
BDT Cut	83852.1	1961.65	2704.18	1566.07	11116.3	476.269	986.783	6170.45

Table 3: Signal and background yields in the cutflow of $\nu\bar{\nu}H$ analysis, normalized to 5000 fb^{-1}



(a) (b)

Figure 6: Distribution of visible energy(left) and minimum jets' pfo multiplicity(right) for the signal and SM backgrounds.



(a) (b)

Figure 7: Distribution of y_{34}(left) and $\Delta\theta$ (right) for the signal and SM backgrounds.

3.3 $q\bar{q}H$ Event Seleciton

The $q\bar{q}H$ channel refers to ZH production in which both Z and Higgs bosons decays hadronically. We require exclusively 4 jets reconstructed by Durham-like algorithm[16] in final states, corresponding to the leading logrithm approximation of 4-partners final states. The domiant SM backgrounds are consist of diboson production followed by hadronic decays of both bosons, and quark pair production. Events with loosened isolation lepton candidates were rejected. All of the 4 jets are required to contain at least 10 PFOs to remove the events with fake jets. The visible energy of each event are required to be larger than 206 GeVto reject the events with energetic nuetrinos. In figure 3.3 visible energy and minimum PFO multicplicity of the jets distributions of signal and backgrounds are presented.

To suppress the background from 2 jets events, the y_{34} are used as discriminator, which can effectively select the 4 jets events from those with lower jet multiplicity. A cut on the energetic-weighted angular dispersion was defined as $\Delta\theta$ variable, which was used in ALEPH 4-bjets channel analysis[21], is also used to furuther suppress the $q\bar{q}$ and semi-leptonic 4 fermions backgrounds. . The distribution of y_{34} and $\Delta\theta$ can be found in figure 3.3.

The 4 jets in the final states are paired in order to get minimum χ^2 defined as:

$$\chi^2_{HZ} = \min\{(m_{ij} - m_H)/\sigma^2_H + (m_{kl} - m_Z)/\sigma^2_Z\} \quad (1)$$

in which i,j,k,l are the jet index, running from 1-4; m_{ij} is the invariant mass of one jet pair and m_{kl} is that

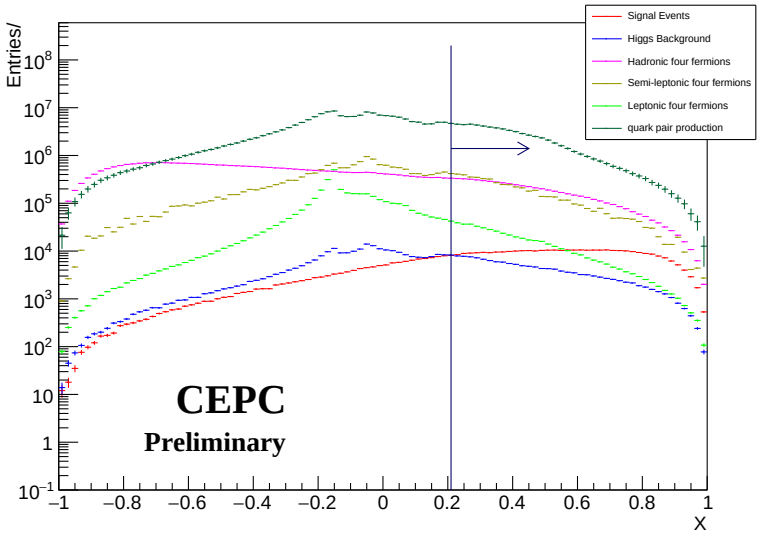


Figure 8: The distribution of X defined in formula 2.

of another pair; m_H and m_Z are Higgs and Z boson mass; σ_H and σ_Z are the width of reconstructed jet pair invariant mass from Higgs and Z boson. A variable χ^2_{WW} is defined in the similar way as formula 1, in which the Z/Higgs masses and widths are replaced by those of W boson. The χ^2_{HZ} and χ^2_{ZZ} are combined into a variable X:

$$X = \frac{\chi^2_{WW} - \chi^2_{HZ}}{chi^2_{WW} + \chi^2_{HZ}} \tag{2}$$

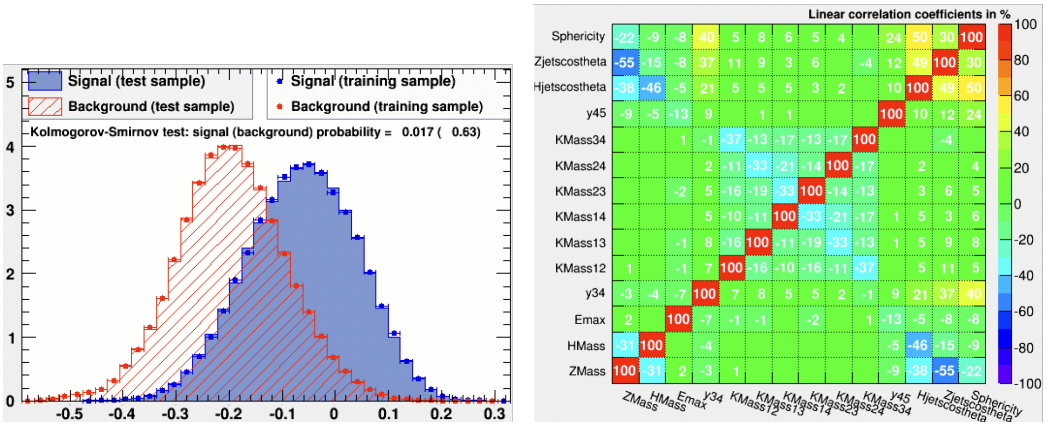
The X value distrutes in the range from -1 to 1. For signal events, it tends to be positive values(close to 1) while for the dominant WW hadronic background, the X value tends to be negative(close to -1), as shown in figure 3.3. Thus X provides discrimination power against WW hadronic events.

The BDT[20] method is applied to further suppress the SM backgrounds. The variables like reconstructed Higgs and Z mass, different combination of jet pair invariant masses, y_{34} and y_{45} and sphericity, reconstructed Z and Higss $\cos\theta$ and highest energy of of jets are used in the training. Figure 3.3 shows the signal and background distribution of linear correlation of these variables and the outcome BDT response of signal and backgrounds.

	signal combined	$H \to b\bar{b}$	$H \to c\bar{c}$	$H \to gg$	background combined	higgs background	4 fermions hadronic	quark pair	4 fermion semi-leptonic
4 jets and iso-lepton veto	493947	413299	19362	61286	75 M	299583	36.83 M	23.86M	14.72 M
E_{vis} >206 GeV	459972	381470	18690	59812	50.6 M	109529	28.19 M	20.37M	1.967M
$y_{34} > 0.007$	393979	325137	15976	52866	26.4 M	100813	21.32 M	5.207 M	218394
$N_{jet,pfo} \geq 9$	371240	305982	14903	50355	21.4 M	82281	18.78 M	2.601 M	27487
$\Delta\theta$ >0.92	318163	261808	12610	43745	13.55 M	71987	12.16 M	1.315 M	4745
$X > 0.21$	236652	197510	9562	29580	3.15 M	38579	2.2 M	907188	3012
BDT >-0.19	211281	177447	8324	25510	1.52 M	32653	1.08 M	405567	580

Table 4: Signal and background yields of cutflow in $q\bar{q}H$analysis, normalized to 5000 fb^{-1}

A yth-value is defined according to the PFOs distribution to suppress the background from di-jet events.



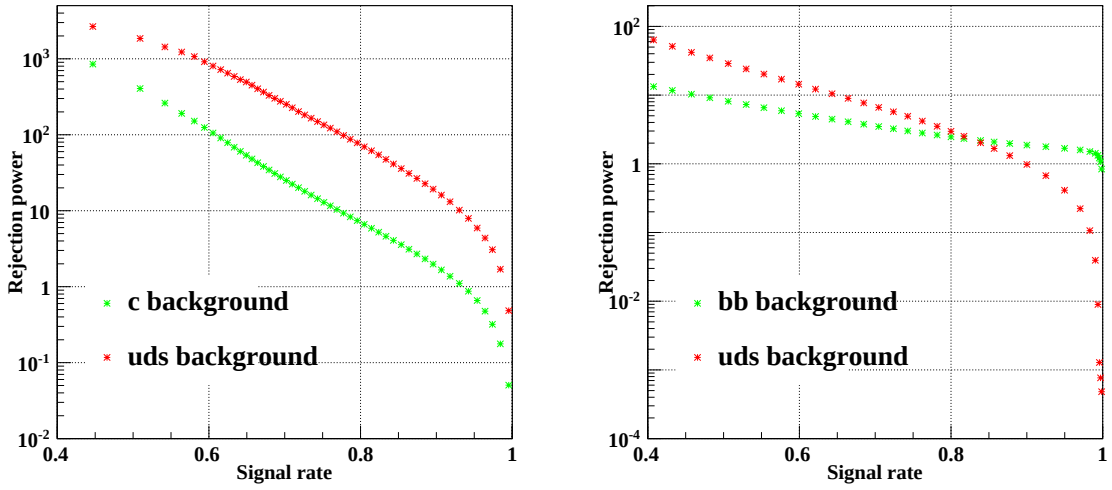
(a) (b)

Figure 9: Linear correlation map of variables used in BDT training(left) and the outcome BDT response of signal and backgrounds.

4 Flavor tagging and template fit

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 mulitiplicity. In each category, a b-tagging and a c-tagging training with BDT method was implemented over $Z-$pole di-jet events, employing variables including jets kinematic varibles, 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. The performance of b-tagging and c-tagging is presented in figure 4.1.



(a) (b)

Figure 10: Flavor tagging performance curve. Left: b-tagging efficiency as function of rejection against *c* and light jet. Right: c-tagging efficiency as function of rejetion against *b* and light jet.

4.2 Template fit

The b-likenesses of the two jets from higgs decay, say L_{b1} and L_{b2}, can be combined to construct a discriminator variable $X_B = \dfrac{L_{b1}L_{b2}}{1 - L_{b1}L_{b2}}$. The conservation of quark flavor in higgs decay guarantee that X_B is close to 1 if higgs decay to $b\bar{b}$ while close to 0 in other case. Similar variable are defined for c-likeness: $X_C = \dfrac{L_{c1}L_{c2}}{1 - L_{c1}L_{c2}}$. A set of template is defined according to the signal and background's $X_B - X_C$ distribution. The $X_B - X_C$ distribution in 'data', which is composed from signal and background simulation events, is shown in figure 4.2.

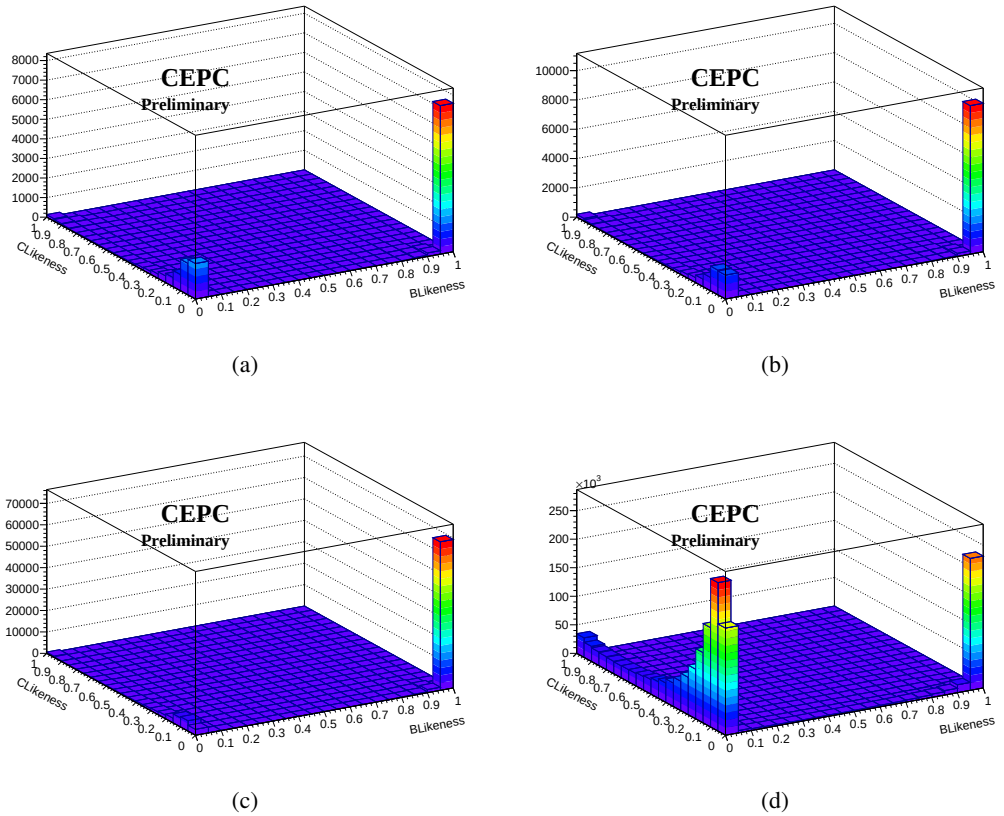


Figure 11: $X_B - X_C$ distribution of pesudo-data in each channel: templates of e^+e^-H(top left),$\mu^+\mu^-H$(top right),$\nu\bar{\nu}H$ (bottom left) and $q\bar{q}H$(bottom right)

A likelihood was constructed according to the template distribution:

$$\log L = \sum_{i=1}^{N} \log Possion(\mu_i, n_i) \tag{3}$$

in which i is the bin index(we have 20 bins for X_B and X_C so in all $N = 400$); $Possion(\mu_i, n_i)$ is the Possion likelihood with for n_i obvserved event and expectation μ_i, which include the parameter of interest:

$$\mu_i = \sum_{j=1}^{s} f_j \times \mu_{i,j} \tag{4}$$

in which j is the index of templates (one template for each components); f_j is the fraction of jth components which need to be worked out; $\mu_{i,j}$ stands for the expected event yields of jth components in ith

bin. Minimize $-\log L$ we can fit the f_j. The distribution of template for $H \to gg/bb/cc$ is shown in figure 4.2. The $H \to b\bar{b}$ and $H \to c\bar{c}$ concentrate in region (1,0) and (0,1) respectively, while $H \to gg$ tends to distribute around (0,0). It can be found a tail at (1,0) show up in H distribution, which are due to contribution of gluon splitting $g \to b\bar{b}$(e^+e^-H,$\mu^+\mu^-H$and $\nu\bar{\nu}H$) and hadronic $Z-$decay($q\bar{q}H$).

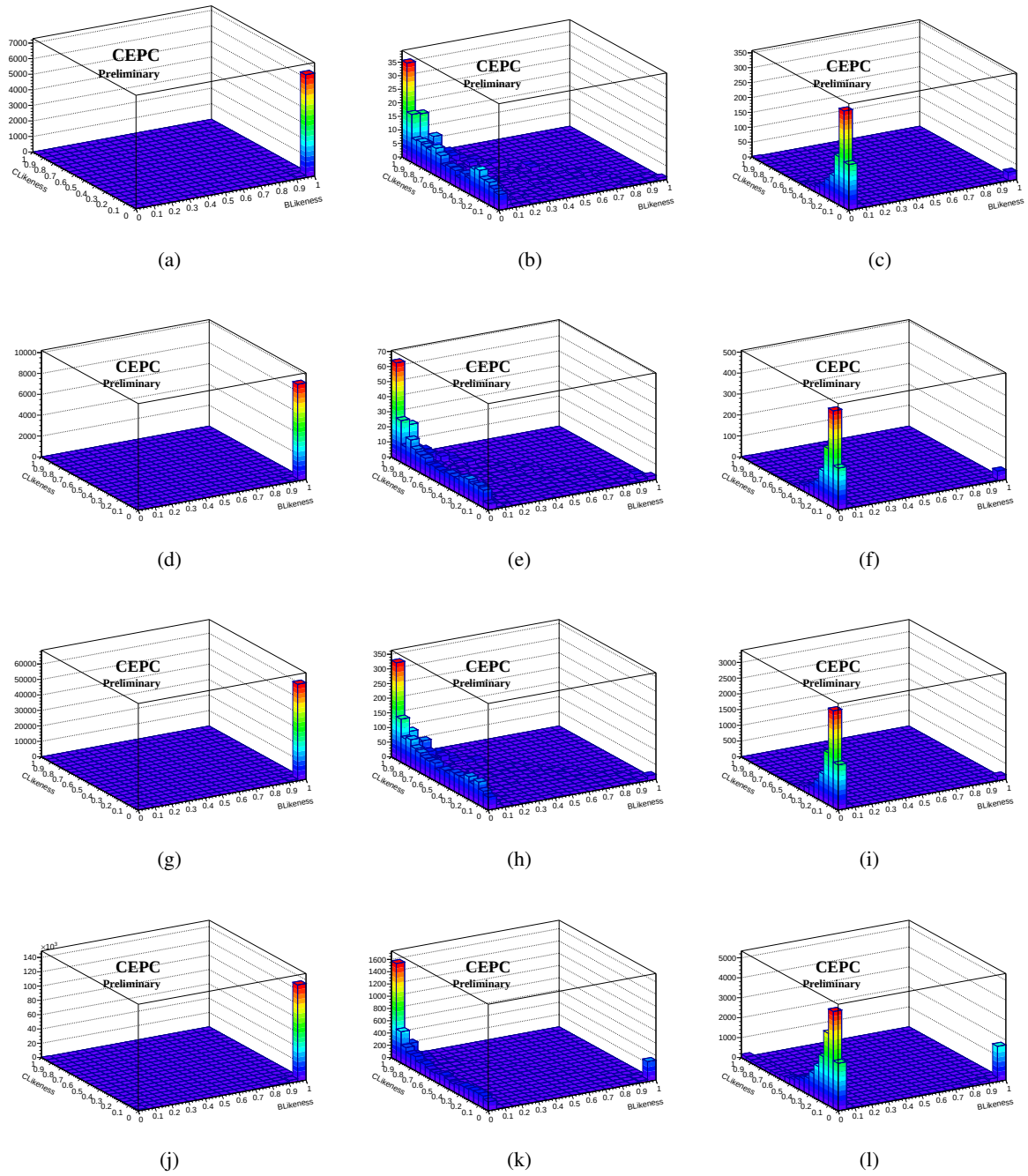


Figure 12: Template of signal events. From top row to bottom row: templates of e^+e^-H,$\mu^+\mu^-H$,$\nu\bar{\nu}H$ and $q\bar{q}H$; from left column to right column: templates of $H \to b\bar{b}$,$H \to c\bar{c}$ and $H \to gg$.

In the fit, the $H \to b\bar{b}$,$H \to c\bar{c}$ and $H \to gg$ combined events yield is set as free parameter, which is equal to $L \times \sigma_{ZH \to 4jets}$. The fraction of each of the 3 flavor finals states, say f_b,f_c and f_g, contribute

2 independent free parameters[2]. The shape of backgrounds are fixed. The background distribution was shown in figure 4.2.

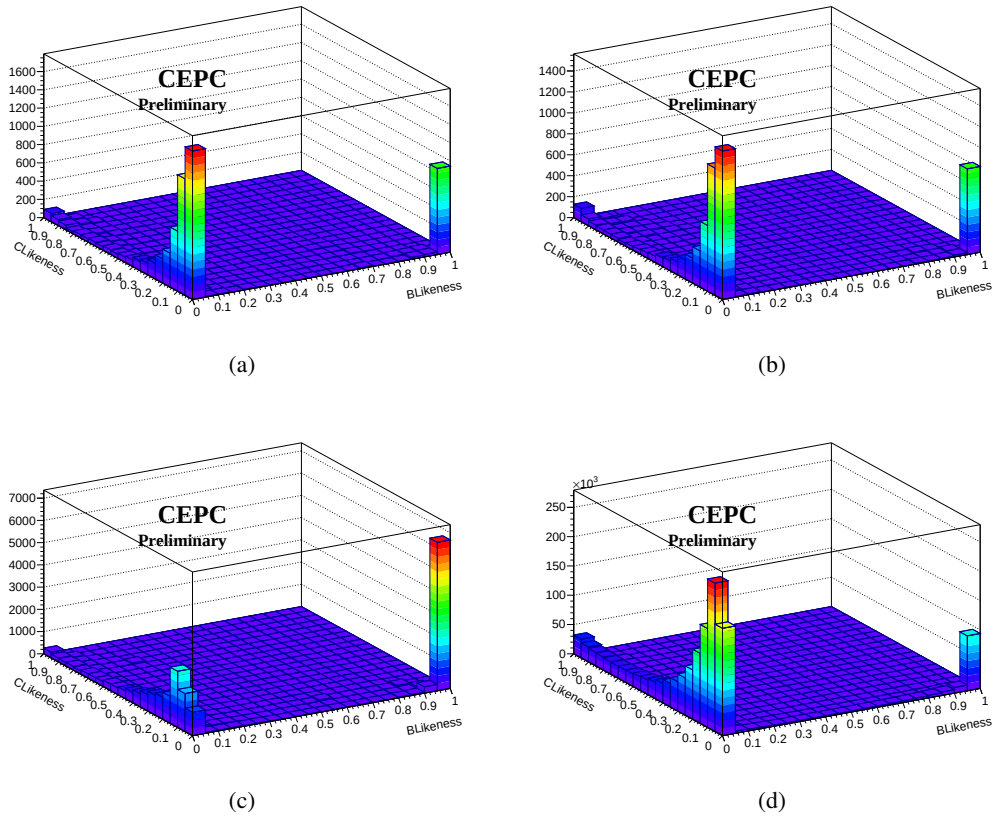


Figure 13: Templates of background events in e^+e^-H(top left),$\mu^+\mu^-H$(top right),$\nu\bar{\nu}H$ (bottom left) and $q\bar{q}H$(bottom right) channel.

4.3 ToyMC test with templatefit

A ToyMC test was applied to evaluate the uncertainty from template fit. In this test, the template fit was done repeatly to the different datasets, in which events yields at each bin on the $X_B - X_C$ distribution are set to be random value according to Possion distribution. This test deomonstrate the unceratiny from the fit due to data statistic fluctuation. The fit results for e^+e^-H, $\mu^+\mu^-H$, $\nu\bar{\nu}H$ and $q\bar{q}H$ are shown in figure 14.

5 Validtion of the Template Fit Method

A delicate flavor tagging commissioning is required to validate the template fit method. The discrepancy between data and monte carlo in template distribution will lead to biased fitting results. Thus it is necessary to apply the tempalte fit to a control sample, and estimate the results in terms of the bias in each flavor components. In this section, we are not trying to directly esimate that bias since we have no collision data, but to demonstrate feasibility of the procedure of the validation.
The semi-leptonic ZZ events can be selected as the control sample. There are several advantage to do so:

[2]The renormalization requires $f_b + f_c + f_g = 1$, reducing the number of free parameter by 1.

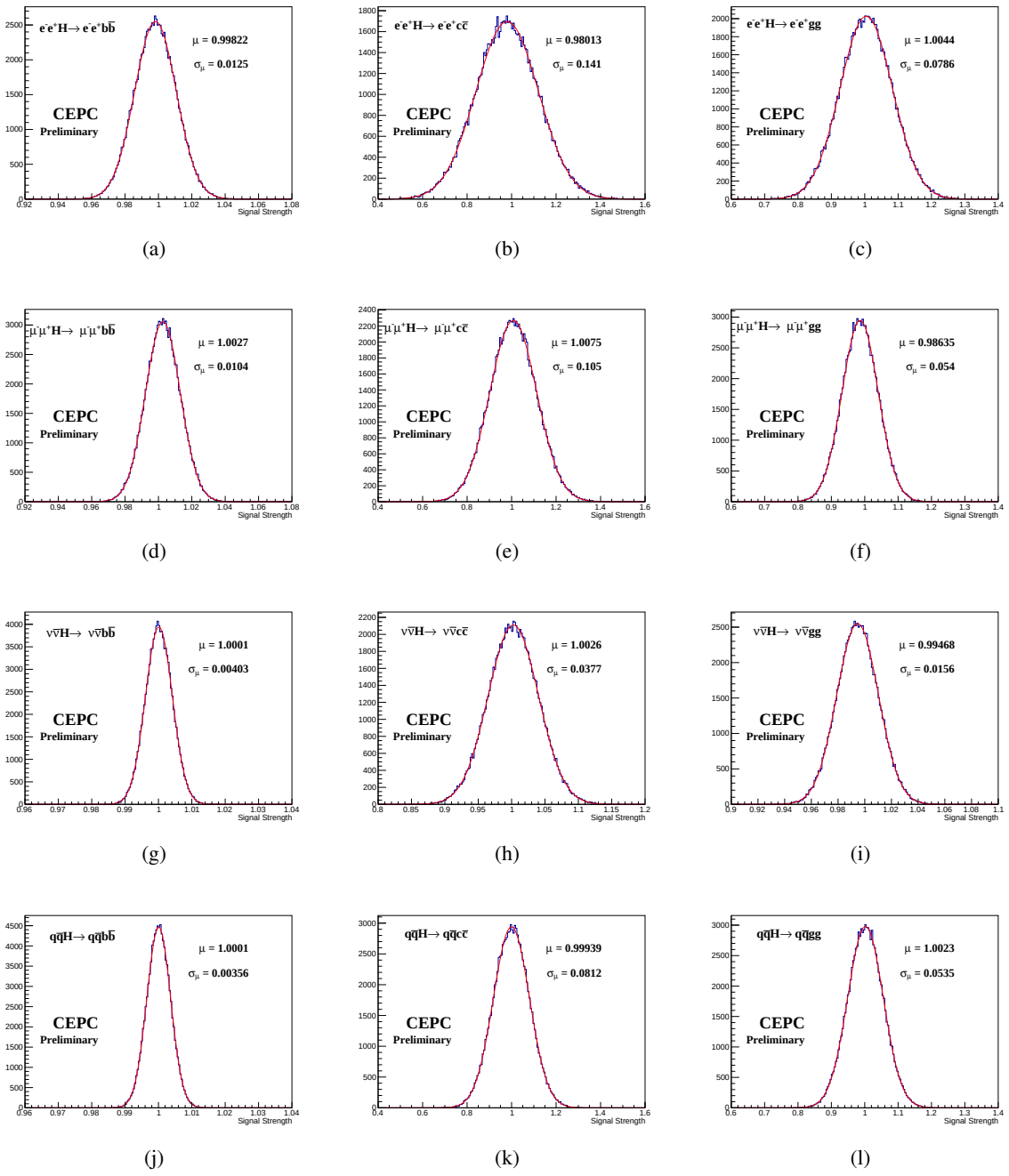


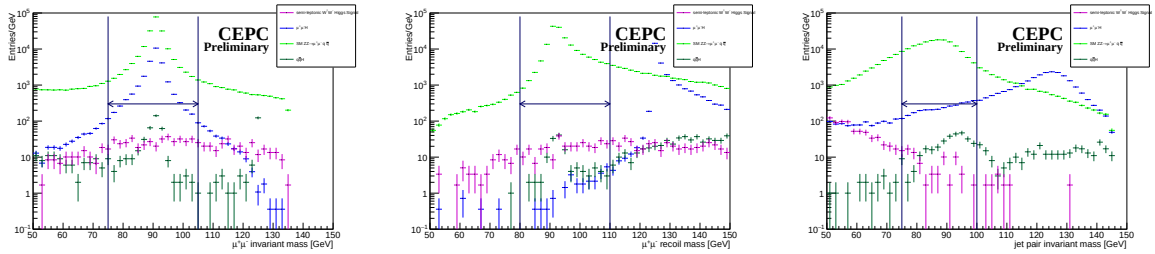
Figure 14: Toy MC test result in terms of signal strength and uncertainty from template fit. The signal strength from template fit in each channel for each higgs hadronic decay is presented. Plots in each row are from the same channel, from top row to bottom row are : e^+e^-H, $\mu^+\mu^-H$, $\nu\bar{\nu}H$ and $q\bar{q}H$; plots in each column are from the same higgs decay mode, from left to right column are $H \to b\bar{b}$, $H \to c\bar{c}$ and $H \to gg$.

	ZZ semi-leptonic	W^+W^- semi-leptonic	$\mu^+\mu^-H$	$q\bar{q}H$
75 GeV< $M_{\mu^+\mu^-}$ <105 GeV	193.3k	412.2	25.96k	366.68
80 GeV< $M_{\mu^+\mu^-recoil}$ <110 GeV	157.4k	132.9	13.99	5.01
75 GeV< M_{jj} <100 GeV	124.1k	8.41	2.51	1.00
Purity		99.99%		

Table 5: Events yields of semi-leptonic ZZ events and other dominant process.

- The cross section of ZZ events is large. There will be about 1.1 million semi-leptonic ZZ channel with $\mu^+\mu^- q\bar{q}$ and $\nu\bar{\nu}q\bar{q}$ each, and over 1.6 million events with $e^+e^- q\bar{q}$ events.
- The hadonic $Z-$decay provide abundant $b\bar{b}$ and $c\bar{c}$ events
- The signiture of ZZ events is very clear, by which purity of the control sample can be guaranteed
- The kinematic feature of jets in the ZZ semi-leptonic jets is similar to that in signal

The ZZ events was selected in $ZZ \to \mu^+\mu - q\bar{q}$ channel. The invariant mass of $\mu^+\mu^-$, jet-pair and the $\mu^+\mu^-$ recoil mass are required in Z-resonance region, which can be seen in figure 5. The event yields of $ZZ \to \mu^+\mu - q\bar{q}$ and other process are shown in tabel 5.



(a) (b) (c)

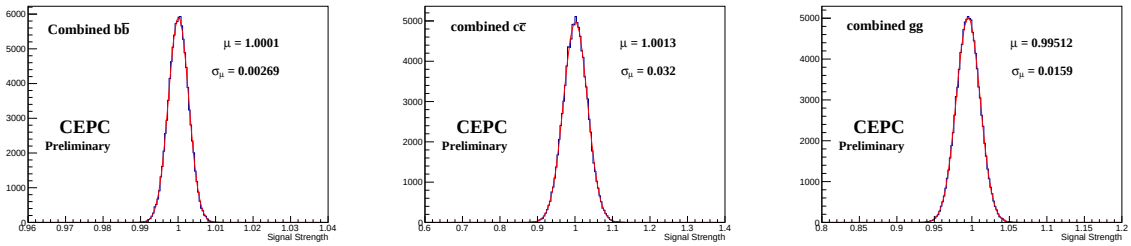
Figure 15: The invariant mass of $\mu^+\mu^-$(left), the recoil mass of $\mu^+\mu^-$ and the invariant mass of jet pair in ZZ control sample

6 Results

To extract the signal strength from all the four sub channels, a combination template fit was applied. This fit concentrate also makes use of $X_B - X_C$ distribution in each channel. The combined log-likelihood is defined as the sum of that in each sub channel. The free parameters are the same as that in the fit of each channel, described in 4.1. The fraction of each flavor are set as common parameterwhile the over all hadronic yields in each channel are set as indepedent parameter. The combined fit results are shown in figure 6.

7 Summary

The combination template fit from e^+e^-H,$\mu^+\mu^-H$,$\nu\bar{\nu}H$ and $q\bar{q}H$evaluate the uncertainty of $H \to b\bar{b}$, $H \to c\bar{c}$ and $H \to gg$ to be 0.27%, 3.2% and 1.6%, reflecting the statistic uncertainty with 5000 fb^{-1}integral luminosity data taken at $\sqrt{s}$ =250 GeV at CEPC. These results are done with all the backgrounds and



(a) (b) (c)

Figure 16: Combined template fit results of signal strength for $H \to b\bar{b}$, $H \to c\bar{c}$ and $H \to gg$ process.

signals from full simulation, and is consistent to the number estimated in pre-CDR. The precision of $H \to b\bar{b}$ is mainly constrained by $q\bar{q}H$channel, while the other two hadronic decay modes are mainly constrained by $\nu\bar{\nu}H$ channel. In $q\bar{q}H$ channel, $H \to gg/c\bar{c}$ suffered from huge backgrounds from hadronic diboson process, and the mis-combination of jet pair degenerate the percision. To solve these problem, it is necessary to optimize the detector and reconstruction performance, which future work should concentrate on.

References

[1] ATLAS Collaboration, Phys. Lett. B **716** (2012) 1–29.

[2] CMS Collaboration, Phys. Lett. B **716** (2012) 30–61.

[3] P.W.Higgs, Phys.Rev.Lett **13** (1964) 508–509.

[4] F.Englert and R.Brout, Phys.Rev.Lett **13** (1964) 321–323.

[5] ATLAS Collaboration, JHEP **01** (2015) 069, 1409.6212.

[6] CMS Collaboration, Phys.Rev.D **89** (2014) 012003, 1310.3687.

[7] CMS Collaboration, Eur.Phys.J.C **75** (2015) 251, 1502.02485.

[8] ATLAS Collaboration, JHEP **11** (2016) 112, 1606.02181.

[9] CMS Collaboration, Phys.Rev.D **92** (2015) 032008, 1506.1010.

[10] M. et.al, Tech. Rep. IHEP-EP-2015-01, The CEPC-SPPC Study Group, 2015.

[11] T. et al., Tech. Rep. KEK-REPORT-2006-6, ILD Concept Group, 2009.

[12] T. W.Kilian and J. Reuter, Eur.Phys.J.C **71** (2011) 1742.

[13] S. T.Sjostrand and P.Skands, JHEP **2006** no. 026, (2006).

[14] GEANT4 Collaboration Collaboration, Nucl.Instrum.Meth. **506** (2003) 250–303.

[15] P. de Freitas and H.Videau,. LC-TOOL-2003-010.

[16] M. G. B. S.Catani, Y.L.Dokshitzer, Phys.Lett.B **269** (1991) 432–438.

[17] T. T. T. Suehara, J.NIMA **808** (2015) 109–116.

[18] X. M. Gang Li,, "Generated sample status for cepc simulation studies." CEPC-TLS-GEN-2015-001, April, 2015.

[19] M.-Q. R. D.-Y. W. G. L. S. J. Y. B. Zhen-Xing Chen, Ying Yang, Chinese Physics C **41** no. 2, (2017).

[20] J. Z. Y. L. I. S. G. M. Byron P. Roe, Hai-Jun Yang, J.NIMA **543** (2005) 577–584.

[21] ALEPH Collaboration Collaboration, Phys.Lett.B **499** (2001) 53–56.