Higgs → \( \tau \tau \) Branching Ratio Measurement at CEPC

To cite this article: YU Dan 2018 J. Phys.: Conf. Ser. 1024 012008

View the article online for updates and enhancements.

Related content
- Investigation of CMOS pixel sensor with 0.18 m CMOS technology for high-precision tracking detector
  L. Zhang, M. Fu, Y. Zhang et al.
- Interference effects on Higgs mass measurement in \( e^+ e^- \) at CEPC
  Guang-Zhi Xu, Gang Li, Yi-Jie Li et al.
- Feasibility study of TPC at electron positron colliders at Z pole operation
  M. Zhao, M. Ruan, H. Qi et al.
Higgs $\rightarrow \tau\tau$ Branching Ratio Measurement at CEPC

Dan YU

E-mail: yudan@ihep.ac.cn

Abstract.

The CEPC (Circular Electron Positron Collider) and ILC (International Linear Collider) detector design provides a high efficiency vertex and Arbor Particle Flow Algorithm\cite{1} uses a new charged particle identification with efficiency higher than 98%. Taking these advantages, the Higgs decaying to $\tau\tau$ channels have been studied in full simulation samples, and the precision of the production cross section times the branching ratio have been improved to 1% level.

1. Introduction

Taus are extremely intriguing physics objects. As the heaviest lepton in the Standard Model (SM), a significant fraction of the SM Higgs boson decays into di-\tau final states. As a result, $g(H\rightarrow\tau\tau)$ could be measured to an accuracy better than 1% at the CEPC. Measuring the polarization of Tau at the Z pole leads to a precise determination of forward-backward asymmetry $A_{fb}(\tau)$ and therefore $\sin^2\theta$ and could be used to check the lepton universality. The reconstruction of $\tau$ functional spectral is of key importance to the CEPC Electro-Weak program, as it grants the access to a precise reconstruction of $\alpha_S$ and EW parameters. In this note, the branching ratio of $H\rightarrow\tau\tau$ is measured in full simulated ZH samples and SM backgrounds.

Most of the tau decaying modes\cite{2} are with one or three charged particles in the final state and an odd number of photons, which is the main idea of the $\tau$ tagging we plan to perform. From the decay modes, the topology of tau’s is simpler than jets, allowing a clear disentanglement of tau events from the others.

2. Monte-Carlo Samples

The CEPC luminosity is planed to be 5000 fb$^{-1}$, the cross section for different Z decay modes are summarized in [3]. All the samples mentioned in this note are generated by the MC generator Whizard, version 1.95\cite{4}. The detector used in the simulation is the CEPC_v1 detector, containing a tracking system with excellent momentum resolution and a high granular calorimeter system (ECAL and HCAL)\cite{5}.

3. Event categories

A successful reconstruction of the $\tau$ lepton is not a trivial task, since it could be generated with various event topologies, and it decays in different final states. In the CEPC collision environment, the $\tau$ events are divided into two categories according to their topology. The reconstruction algorithm and the spectrometer performances have been optimized following this selection criterium.
Due to the limited computing resource, the inclusive $ZH$ events, and SM categories background events are filtered by some preselection using MC truth information to simplify the samples. The excellent performance of PFA ensures that this preselection would not lose information.

### 3.1. $llH$ channel

The first category is the leptonic one, where no physics objects, or only lepton / photon / missing energy is generated together with the $\tau$ candidates$^1$. These events include, for example:

- $ZH, Z \rightarrow l^+ l^- / \nu \nu, H \rightarrow \tau\tau$ events; golden channel for $g(H\tau\tau)$ measurements
- $ZZ, l^+ l^- / \nu \nu / \tau\tau$ events
- $WW$ events with $l\nu\mu\nu$ final states.
- $Z \rightarrow \tau\tau$ events at $Z$ pole operation.

In these events, the global multiplicity is limited while the additional physics objects, if they exist, are easy to identify$^6$. A successful identification of these events relies highly on the reconstruction of photons and charged hadrons. In the following section, the physics performances of $\tau$ reconstruction at $\mu\muH$ and $\nu\nuH$ channel are shown as well as their $Br(H \rightarrow \tau\tau)$ measurement.

The steps for di-$\tau$ events tagging are:

- For $\mu\muH$, veto the $\mu$s decayed from $Z$ by choosing the $\mu$ pair with invariant mass closest to $Z$ mass
- Find the leading track among the remaining particles and collect the tracks and photons close to this track ($\approx 1$ rad, to be grouped in region A), and their numbers are noted as $NTrkA$ and $NPhA$.
- Collect the rest tracks and photons and group them in region B with their numbers noted as $NTrkB$ and $NPhB$.
- Get the angle between the leading tracks in region A or B and the furthest track in this region, noted as $Cone_{T-T}(A/B)$.
- $Cone_{T-P}(A/B)$ is the angle between the leading tracks in region A or B and the furthest photon in this region.
- $Cone_{P-P}(A/B)$, the angle between the leading photon in a region and the furthest photon in this region.

The inclusive SM background could be efficiently subtracted by requesting the proper number of leptons, if exist, and limit the number of photon and charged hadrons. For $\mu\muH$ events, the event selection could be further enhanced by request on the recoil mass against the di-lepton system.

The pre-selections applied to select $\mu\muH$ and $\nu\nuH$ suppress the SM background by 4 orders of magnitude and the remaining backgrounds are dominated by $ZH$ events. The cut chain of the event selection for $\mu\muH$ and $\nu\nuH$ are shown in Table 1 and Table 2.

### 3.2. Hadronic

The second category is the hadronic one, where the $\tau$ lepton(s) are generated with jets. For instance, we have:

- $ZH, Z \rightarrow qq, H \rightarrow \tau\tau$

$^1$ The charge is ignored for event classifications.
Table 1. $\mu\muH$ cut flow

|                | $\mu\muH$ cut flow | $\mu\muH$ cut flow | $\mu\muH$ cut flow | $\mu\muH$ cut flow | $\mu\muH$ cut flow | $\mu\muH$ cut flow | $\mu\muH$ cut flow |
|----------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| total generated| 2292                | 33557               | 571144              | 44180832            | 15361538            | 7809747             | 418595861           |
| after preselection | 2246              | 32894               | 122674              | 223691              | 0                   | 86568               | 1075886             |
| $N_{Trk}(A/B) < 6$ & $N_{Ph}(A/B) < 7$ | 2219              | 1039                | 2559                | 352                 | 0                   | 9397                | 25583               |
| BDT $>0.78$ | 2135               | 885                 | 484                 | 24                  | 0                   | 157                 | 161                 |
| efficiency | 93.15%              | 2.63%               | <0.01%              | <0.01%              | <0.01%              | <0.01%              | <0.01%              |

Table 2. $\nu\nuH$ cut flow

|                | $\nu\nuH$ cut flow | $\nu\nuH$ cut flow | $\nu\nuH$ cut flow | $\nu\nuH$ cut flow | $\nu\nuH$ cut flow | $\nu\nuH$ cut flow |
|----------------|---------------------|---------------------|---------------------|---------------------|---------------------|---------------------|
| total generated| 15497               | 231670              | 571144              | 44180832            | 17361538            | 7809747             | 418595861           |
| after preselection | 9434              | 214830              | 1239457             | 7463105             | 3327803             | 956694              | 12826280            |
| $N_{Trk}(A/B) < 6$ & $N_{Ph}(A/B) < 7$ | 9260              | 8858                | 24760               | 1354852             | 17389               | 676185              | 1535029             |
| BDT $>0.78$ | 8836               | 6587                | 15450               | 89729               | 1355                | 10739               | 11243               |
| efficiency | 57.02%              | 2.84%               | 0.27%               | 0.20%               | <0.01%              | 0.14%               | <0.01%              |

- $ZZ \rightarrow qq\tau\ell$
- $WW \rightarrow qql\tau$
- $ZH, Z \rightarrow qq, H \rightarrow WW \rightarrow l\nu\tau\nu$

To find the $\tau$ lepton in the hadronic categories is much more challenging than that in the leptonic environments. The identification algorithm would always be a compromise between the signal efficiency and purity. Taking the $\text{Br}(H \rightarrow \tau\tau)$ measurement at $qqH$ events for example, the $\tau$ candidate finding steps are:

- Find tracks with energy higher than a defined $E_{min}$ as the seed
- Collect tracks and photons within a small angle
- Calculate invariant mass with these particles
- Calculate the energy in a larger cone ConeB around the seed.

The parameters for cut of $\tau$ tagging is:

- Number of tracks/photons
- Energy proportion in the smaller cone

Here the cuts of parameters are optimized to the value $\epsilon \cdot p$, where $\epsilon$ is the efficiency of finding an opposite charged $\tau$ pair in $qq\tau\tau$ events and $p$ is the probability of tagging a opposite charged $\tau$ pair in the backgrounds.

After these cuts, the remaining $\tau$s in an event is collected and the two leading energetic ones with opposite charge are chosen to calculate the invariant mass of the $\tau\tau$. The recoil mass and invariant mass of $qq$ system are also applied for the background suppression. The total efficiencies for the signal and the dominating background, $ZZ$ and $WW$, are 49.97%, 0.41% and 0.04%, respectively. The cut chain is summarized in Table 3.
Table 3. Cut Flow of MC sample for $qqH \rightarrow \tau\tau$ selection on signal and inclusive SM backgrounds

| Cut Criteria | $qqH \tau\tau$ total generated (scaled to 5 ab$^{-1}$) | $qqH \tau\tau$ inclusive bkg | $ZH$ inclusive bkg | $ZZ$ | $WW$ | singleW | singleZ | $2f$ |
|--------------|---------------------------------|-----------------------------|-------------------|------|------|--------|--------|------|
| (1st preselection) | 45465 677854 310245 5039286 42425195 1267564 1398362 148401031 | 24674 7342 33721 93955 723989 33887 54386 103642 | 24284 6290 32344 88245 597480 24927 36039 56615 | 22937 2103 4887 65625 21718 738 1893 556 |
| (2nd preselection) | 45145 174650 226059 293306 12452091 125735 117306 547402 | 24674 7342 33721 93955 723989 33887 54386 103642 | 24284 6290 32344 88245 597480 24927 36039 56615 | 22937 2103 4887 65625 21718 738 1893 556 |
| $N_{\tau+} > 0, N_{\tau-} > 0$ | 24674 7342 33721 93955 723989 33887 54386 103642 | 24284 6290 32344 88245 597480 24927 36039 56615 | 22937 2103 4887 65625 21718 738 1893 556 |
| $20 GeV < M_{\tau+\tau-} < 20 GeV$ | 24674 7342 33721 93955 723989 33887 54386 103642 | 24284 6290 32344 88245 597480 24927 36039 56615 | 22937 2103 4887 65625 21718 738 1893 556 |
| $70 GeV < M_{qq} < 110 GeV$ | 24674 7342 33721 93955 723989 33887 54386 103642 | 24284 6290 32344 88245 597480 24927 36039 56615 | 22937 2103 4887 65625 21718 738 1893 556 |
| $100 GeV < M_{qq}^{Rec} < 170 GeV$ | 24674 7342 33721 93955 723989 33887 54386 103642 | 24284 6290 32344 88245 597480 24927 36039 56615 | 22937 2103 4887 65625 21718 738 1893 556 |
| Efficiency | 49.97% 0.31% 1.26% 0.41% 0.04% 0.01% 0.01% 0.01% |

4. Results

To conclude, the $\tau$ reconstruction at the CEPC is currently categorized into leptonic and hadronic events and reconstructed using different strategies and $\tau$ finding algorithms. In the leptonic events, where the $\tau$ lepton is generated only in association with leptons, photons or missing energy, the $\tau$ events identification relies strongly on a successful reconstruction of the photons and charged hadrons.

In the hadronic events, it is more difficult to suppress the background, for further study, the correlation with other channels might be applied. The total signal yields are extracted from a fit to a variable defined by the impact parameters on the leading tracks from the two tau objects, and the fit results are shown in Figure 1. In summary, with 5 ab$^{-1}$ of CEPC data and all three channels combined, the Branching Ratio of $H \rightarrow \tau\tau$ measurement is $6.28 \pm 0.07\%$, and the fractional error on the $\sigma \times BR$ can reach $0.81\%$.

Figure 1. Post-fit distributions of the impact parameter variable based on the leading tracks from the two taus in the $Z \rightarrow \ell\ell$ (a) and $Z \rightarrow qq$ (b) channel, expected for 5 ab$^{-1}$ of CEPC $ZH$ data.
With these channels analyzed and fitted, the cross section of Higgs decaying to $\tau\tau$ can be summarized as in Table 4.

### Table 4. Combined cross section

|           | BR ($H \rightarrow \tau\tau$) | $\delta (\sigma \times BR)/(\sigma \times BR)$ |
|-----------|--------------------------------|-----------------------------------------------|
| $\mu\mu H$ | $6.40 \pm 0.18$               | 2.26%                                         |
| eeH (extrapolated) | $6.37 \pm 0.18$ | 2.72%                                         |
| $\nu\nu H$ | $6.19 \pm 0.17$               | 4.29%                                         |
| qqH        | $6.23 \pm 0.04$               | 0.93%                                         |
| combined   | $6.28 \pm 0.07$               | 0.81%                                         |

In both cases, a precise reconstruction of the impact parameter is essential for the $\tau$ events identification, as shown in the figures, the statistics can be fitted only if the position resolution is good enough to distinguish the two peaks for $\tau$s and backgrounds.

### 5. Extrapolating in ILC

Comparing these cross sections for three polarization scenarios in ILC at 250GeV with the cross section at CEPC as shown in previous section, a simple extrapolation can be done as in Table 5. The assumption here is that the efficiency for each signal and background stays the same for ILC and CEPC.

### Table 5. Extrapolated accuracy $\delta (\sigma \times BR)/(\sigma \times BR)$ in ILC 250GeV (2000 fb$^{-1}$)

| Type     | (-80%,-30%) | (80%,-30%) | non-pol |
|----------|-------------|------------|---------|
| $\mu\mu H$ | 2.98        | 3.63       | 3.57    |
| $\nu\nu H$ | 5.48        | 4.34       | 6.78    |
| qqH      | 1.26        | 1.36       | 1.47    |
| combined | 1.13        | 1.22       | 1.33    |

### 6. Discussion

In this paper, different channels with Higgs decaying into $\tau\tau$ at CEPC have been studied and the combined accuracy is reaching 1% level. This result is also extrapolated to ILC and also gives the reasonable accuracy.

This accuracy has still space to be improved. One choice is to use the collinear approximation to recover the momentum of neutrino(s) from $\tau$. This method needs to assume that $\tau$ decay products almost fly back-to-back. The collinear approximation will help to reconstruct the invariant mass of tau pair system and its comparison to the Higgs mass could be a powerful variable to suppress ZZ/WW backgrounds with $\tau$ final states.

Another method is to fully reconstruct hadronically decaying $\tau$ momenta by making use of the interaction point position, the impact parameters of the $\tau$ decay products, and the transverse momentum of the Z boson recoiling against the $\tau\tau$ system. Since more than 60% of $\tau$s decays into hadrons, this method will help to improve the performance of these channels.

Besides, a jet clustering algorithm can be applied in the qqH channel in order to suppress the 2f backgrounds with jets.

However, this study here is based on a perfect vertex detector, the resolution is not taken into account. Since the result was obtained from the impact parameter, the influence of the vertex detector design to the performance should be studied in the future.
7. Bibliography

[1] M. Ruan, *Arbor, a new approach of the Particle Flow Algorithm*. arXiv:1403.4784 (2014).

[2] C. Patrignani et al. Particle Data Group, *Chin. Phys. C*, **40**, 100001 (2016)

[3] X. Mo, G. Li. Generated Sample Staits for CEPC Simulation Studies. *CEPC NOTE 2015*

[4] Kilian W, Ohl T, Reuter J. WHIZARDsimulating multi-particle processes at LHC and ILC[J]. *The European Physical Journal C*, 2011, **71**(9): 1742.

[5] M. Ahmad et al (The CEPC-SPPC Study Group), CEPC-SppC Preliminary Conceptual Design Report: Physics and Detector, http://cepc.ihep.ac.cn/preCDR/main preCDR.pdf, retrieved 4th May 2015

[6] D. YU, M. Ruan, V. Boudry, H. Videau. Lepton identification at particle flow oriented detector for the future $e^+e^-$ Higgs factories. *Eur. Phys. J. C* (2017) **77**: 591.