Probing the CP-Violation effects in the $h\tau\tau$ coupling at the LHC

Xin Chen$^{1,2,3,*}$ and Yongcheng Wu$^{3}$

$^1$Department of Physics, Tsinghua University, Beijing 100084, China
$^2$Collaborative Innovation Center of Quantum Matter, Beijing 100084, China
$^3$Center for High Energy Physics, Peking University, Beijing 100084, China

The prospects of measuring the Higgs CP mixing angle with $h \rightarrow \tau \tau$ decays at the LHC is presented. The analysis is based on a new method used to reconstruct the neutrino from tau decay with high precision, and a matrix element based method used to extract the best sensitivity to the Higgs CP. All major hadronic tau decay modes are included. It is predicted, based on a detailed detector simulation, that with $3 \text{ ab}^{-1}$ of data at $\sqrt{s} = 13 \text{ TeV}$, a significant improvement of the measurement of the CP mixing angle to a precision of $6.9^\circ$ can be achieved at the LHC.

To account for the large asymmetry between the matter and anti-matter in our Universe, enough CP violation effects should be presented in the theory. However, in the Standard Model, the CP phase in the CKM matrix is not sufficient for this purpose. New physics which can provide new source of CP violation is therefore needed to introduce more CP violation sources. Possible candidates are Supersymmetry, Left-Right Symmetric model, etc.

On the other hand, the new discovered Higgs boson also opens a window towards the new physics. The precision measurement of Higgs properties will be one of the most important targets of the LHC in the next running periods. Among them, the CP property is an important topic. The pure CP eigenstate assumption has already been investigated at the LHC experiments [18] in the diboson decays, and the pure CP-odd situation is excluded better that 99.9% CL. The $h \rightarrow ZZ^* \rightarrow 4l$ is the golden channel for this measurement, subject to the scale suppression due to dim-6 operators, whereas for the Yukawa coupling, $h \rightarrow \tau \tau$ is the best channel we could use and is widely investigated in the literatures [4–17]. However, the missing neutrino from the decay of each tau makes it difficult to achieve a better precision on measuring the CP mixing angle ($\phi$) of the $h\tau\tau$ interaction which we assume to have the following effective form:

$$\mathcal{L} = -\frac{y_{\tau}}{\sqrt{2}} \tau (\cos \phi + i\gamma^5 \sin \phi) \tau h,$$  \hspace{1cm} (1)

In this work, utilizing the mass constraints and impact parameters, a new method which can be used to calculate the missing neutrino per event with the best precision one can achieve is proposed, which is not tried for LHC before. With the momentum of the neutrino from the tau decay reconstructed, an observable based on the matrix element to retrieve the CP information in the $h\tau\tau$ interaction is calculated for the first time. Unlike in the previous work where only specific tau decay modes are studied, all major tau hadronic decay modes are included and combined in this work, which gives the important prediction on the best we can do with the Higgs CP in its fermionic coupling at the LHC. A detailed simulation study shows that a significant improvement of the measurement of the CP mixing angle can be achieved.

We stress that the detector response effect is important, because if the signal yield is overestimated or the background is underestimated, it can lead to unrealistic CP measurement accuracy.

As already been investigated in [4], the most promising Higgs production channel for our purpose is the VBF channel. Although the gluon-gluon fusion is the dominant production channel for the Higgs at the LHC, it has larger background and lower signal purity, which will impact a lot the measurement of the CP property [18]. Thus, we will use the VBF channel for the Higgs production which possesses a particular topology. Only hadronic decay modes for two taus from Higgs decay are used, since the leptonic decay mode contains two missing neutrinos. Not only the neutrino pair mass induces a new unknown parameter that is hard to construct, but also it dilutes the CP sensitivity by having to integrate out the relative degrees of freedom between the two neutrinos. On the other hand, the hadronic modes don’t have this extra parameter, and also have the best statistics among all the modes [19]. The main backgrounds for this signal come from the Z production associated with additional jets. The processes we consider are listed in following:

- **Signal:** $pp \rightarrow hjj, h \rightarrow \tau\tau$.
- **Background:** $pp \rightarrow Z + 0, 1, 2, 3j, Z \rightarrow \tau\tau$.
- **Tau Decay Modes Used:**
  - $\tau^\pm \rightarrow \pi^\pm \nu$,
  - $\tau^\pm \rightarrow \rho^\pm \nu \rightarrow \pi^\pm \pi^0 \nu$,
  - $\tau^\pm \rightarrow a^\pm \nu \rightarrow \pi^\pm \pi^\pm \pi^\mp \nu$,

and all six combinations of the tau decay modes are used in our simulation. The other backgrounds, mainly dominated by the QCD, is also important [19]. However, QCD fake background is beyond the scope of this work. It is usually estimated from real data by tau charge or ID reverse. We will just assume the same cross section after all selection cuts as the QCD Z background for simplicity, which in our opinion is a conservative choice.

...
The VBF \( h \to \tau\tau \) signal is generated with Powheg [20] at NLO accuracy in QCD and PDF set NNPDF30NLO [21], and interfaced to Pythia8 [22] for resonance decays, parton shower and hadronization. The QCD and EW \( Z+jets \) background, with \( Z \to \tau\tau \), is generated at LO with MadGraph5 [23] and PDF set NNPDF23LO [24], with up to three extra partons. Samples with different parton multiplicities are merged according to the CKKW-L method [25], and showered by Pythia8. A k-factor of 1.23 is applied to the QCD \( Z \) momentum is smeared with a resolution of 0.06 (0.2) GeV in the central region (0.5 < |\( \eta \) | < 2.5), the coefficients of the two terms increase to 0.10 (0.25) and 1.7–3 (3.1–3), respectively. The calorimeter energy of particles is modelled with a resolution of \( A E \pm B\sqrt{E} \pm C \), whose coefficients are listed in Table 1.

| \( |\eta| \) | 1.7 < \( |\eta| \) < 3.2 | 3.2 < \( |\eta| \) < 4.5 |
|---|---|---|
| EM | 0.0017 | 0.0350 |
| HAD | 0.5205 | 0.706 | 1.00 |

The tracking efficiency for charged tracks is 95% (85%) in the |\( \eta \)| < 1.5 (1.5 < |\( \eta \) | < 2.5) region. The track momentum is smeared with a resolution of 0.06 \( \oplus 1.3^{-3}p_T \) (\( p_T \) in GeV) in the central region (|\( \eta \)| < 0.5) in a magnetic field of 2 Teslas. In the region of 0.5 < |\( \eta \)| < 1.5 (1.5 < |\( \eta \) | < 2.5), the coefficients of the two terms increase to 0.10 (0.25) and 1.7–3 (3.1–3), respectively. The calorimeter energy of particles is modelled with a resolution of \( A E \pm B\sqrt{E} \pm C \), whose coefficients are listed in Table 1.

The jets are formed from the clustered energy deposition of particles in the calorimeter based on the Anti-
\( k_T \) algorithm [28] with a cone parameter of 0.4. The hadronic tau tagging is performed on these jets with an efficiency of 70% (60%) and fake rate of 2% (1%) for 1-prong (3-prong) real and fake tau objects, respectively. To measure the CP with \( h \to \tau\tau \), it is essential to identify the different tau decay modes efficiently. The development of tau substructure algorithms in the ATLAS and CMS experiments has recently made this possible using a particle flow method [29] [30]. In this work, it is assumed that different tau decay modes can be efficiently classified with no crosstalk, and the neutral pion energy can be resolved with a 15% uncertainty. The impact parameters of the tracks are used to better constrain the neutrino momenta from tau decays, as used in [5]. A simple resolution of the form \( a \oplus b/p_T \) (\( p_T \) in GeV) is applied on the impact parameters, where \( a = 8.5 (13.5) \) \( \mu m \) and \( b = 110 (200) \) \( \mu m \) for \( d_0 (z_0) \) based on [31]. Although in HL-LHC, it is expected that the tracking range will be extended to |\( \eta \)| < 4.0, in this work, the tracking is still within |\( \eta \)| < 2.5 consistent with the current ATLAS detector. The \( \chi^2 \) for a single track from the tau decay is Eq. 7 or 8 of [5]. For the 3-prong tau decay, it is assumed that the decay proceeds through the \( a_1 \) channel [32], and the combined \( \chi^2_{a_1} = \sum_i \chi^2_i \), where \( i \) is the track index. The tau flight direction is obtained by minimizing \( \chi^2_{a_1} \). Fig. 1 shows the difference in \( \eta \) and \( \phi \) between their fitted (by minimizing \( \chi^2_{a_1} \)) and true values for the taus which decay via \( a_1\nu \). The momenta of the neutrinos from the hadronic tau decays can be obtained by minimizing the \( \chi^2 \) of

\[
\chi^2 = \left( \frac{m_{\text{fit}} - m_h}{\sigma_h} \right)^2 + \left( \frac{m_{\text{fit}} - m_\tau}{\sigma_\tau} \right)^2 + \left( \frac{m_{\text{fit}} - m_\nu}{\sigma_\nu} \right)^2 + \left( \frac{E_{\text{fit}} - E_y}{\sigma_{\text{fit}}} \right)^2 + \left( \frac{E_{\text{fit}} - E_x}{\sigma_{\text{fit}}} \right)^2 + \chi^2_{\text{IP}},
\]

where \( m_h = 125 \) GeV, \( m_\tau = 1.777 \) GeV, \( \sigma_h = 10 \) GeV, \( \sigma_\tau = 0.1 (0.2) \) GeV for taus decaying to \( a_1\nu \) or \( \pi\nu \) (\( \nu \)), \( E_y \) is the missing transverse energy, \( \sigma_{\text{fit}} = 0.67 \sqrt{\Sigma E_T} \) (\( \Sigma E_T \) in GeV) is its resolution. The \( \chi^2_{\text{IP}} \) is the impact parameter contribution. For 1-prong taus, it is just the sum of contributions from each tau, whereas for 3-prong taus, it is \( (\eta_{-1}^{(i)} - \eta_\tau)^2/0.007^2 + (\phi_{-1}^{(i)} - \phi_\tau)^2/0.007^2 \), where \( \eta_\tau \) and \( \phi_\tau \) are the tau direction obtained by minimizing \( \chi^2_{a_1} \) in the previous step. For the states with an intermediate \( \rho \) meson, extra terms of \( (m_\rho - 0.775)^2/0.2^2 + (f_{\pi}\rho - 1)^2/0.15^2 \) are added to Eq. 2 where \( f_{\pi}\rho \) is the energy scale factor multiplied to the \( \sigma_0 \).

In the per-event minimization of Eq. 2 the \( \eta \) and \( \phi \) of one neutrino are firstly scanned over, from which the magnitudes of the neutrinos’ momenta and the direction of the other neutrino can be obtained via the tau mass.
distribution of unconstrained $m_{\tau\tau}$ is done without the mass constraint term in Eq. 2. The Higgs mass at 125 GeV. Thus in the first step, the fit is done without the mass constraint term in Eq. 2. Conversely, the scan is repeated starting from the parameters of the other neutrino. Finally, a fit using MINUIT [33] is performed around the minimal point found by the scans for a better estimation. The $\Delta R$ and momentum difference between the fitted and true neutrinos in the $a_1 + \pi$ channel are shown in Fig. 2.

With the Higgs mass constraint term in Eq. 2, the background ditau mass is also biased to the nominal Higgs mass at 125 GeV. Thus in the first step, the fit is done without the mass constraint term in Eq. 2. The distribution of unconstrained $m_{\tau\tau}$ is shown in Fig. 3. In the second step, the mass constraint term is put back in Eq. 2 and the fit is done to extract the CP information.

In order to measure the CP effects which is retrieved from a differential distribution, we need to also improve $S/B$ as much as possible. For this purpose, the following cuts are used to select reconstructed events:

**Tau cuts:** The tau candidate should have one or three tracks with a unit charge. The leading track has $p_T > 5$ GeV. For the 3-prong tau, $p_T > 2$ GeV on the other tracks. The two taus have opposite charge, and are within $|\eta| < 2.5$. To take into account the trigger, $p_T > 40, 30$ GeV are required on the two taus. They should also have $|\Delta \phi| < 2.9$ to avoid the back-to-back topology.

**VBF Cut:** $p_T^{a1} > 50$ GeV, $p_T^{a2} > 40$ GeV, $|\Delta \eta_{jj}| > 3.8$, $m_{jj} > 500$ GeV, $\eta_{j1} \times \eta_{j2} < 0$

![Figure 2](image_url)

**Figure 2:** The $\Delta R$ and momentum difference between the fitted and true neutrinos from the hadronic tau decays in the $a_1\nu + \pi\nu$ channel.

![Figure 3](image_url)

**Figure 3:** The expected unconstrained $m_{\tau\tau}$ distribution with 300 fb$^{-1}$ after the VBF cuts with all channels combined.

**TABLE II:** The events for signal (in total and also in each decay mode) and background processes left after all selection cuts at the LHC with 300 fb$^{-1}$ luminosity.

| Process | Signal $Z \to \tau\tau$ | $Z \to \tau\tau$ (EW) | QCD-fake(assumed) |
|---------|------------------------|----------------------|------------------|
| Events  | 125                    | 91                   | 91               |
| $p + p$ | 41.3                   | 34.3                 | 1.7              |
| $a_1 + p$ | 31.0                 | 21.3                 | 2.2              |
| $\pi + p$ | 30.7                 | 22.5                 | 2.0              |
| $a_1 + \pi$ | 11.5               | 6.2                  | 0.50             |
| $a_1 + a_1$ | 5.6                | 2.8                  | 0                |
| $\pi + \pi$ | 4.9                 | 3.4                  | 0.25             |
| $\eta_{j1}$ | 3.4                  | 3.4                  | 3.4              |

**Tau Centrality:** $\min \{\eta_{j1}, \eta_{j2}\} < \eta_{\tau_{1,2}} < \max \{\eta_{j1}, \eta_{j2}\}$

**Higgs Mass:** 115 GeV $< m_{\tau\tau} < 150$ GeV

**Missing Energy:** $E_{\text{proj}} - E_{\text{fit}}^{\nu_1 + \nu_2} > -6$ GeV

where $j1$ and $j2$ are the leading and subleading jets, $m_{\tau\tau}$ is the unconstrained mass, $E_{\text{proj}}$ is the projection of $E_T$ onto the transverse direction of the vectorial sum of two neutrinos’ fitted momenta, $E_{\text{fit}}^{\nu_1 + \nu_2}$. This variable is useful because for the $Z \to \tau\tau$ events, the fitted neutrino momenta are stretched to comply to the Higgs mass constraint, resulting in a larger $E_{\text{proj}}^{\nu_1 + \nu_2}$ than $E_{\text{proj}}$.

Based on the above reconstruction and selection, the left events at 300 fb$^{-1}$ LHC are listed in Table. [3] for signal and background processes, from which one find that the most important modes are those involving the $\rho$ meson, whereas the $\pi + \pi$ mode as investigated in [4] is not the best one. Using the reconstructed momentum for all final states, we calculate the matrix element event by event which according to our parameterization (Eq. 1) has the following form:

$$|M|^2 \propto A + B \cos 2\phi + C \sin 2\phi$$

(3)

where $A$, $B$ and $C$ are calculated based on Eq. 1 and the effective Lagrangians and form factors for the $\tau$ decay vertices detailed in [41], which depends on the momenta of all final state particles as inputs. From the coefficients
as we have used all final states information (especially the neutrino momentum) incorporated into the matrix element to reconstruct an “angle” and retrieve the CP mixing information, compared with usual construction methods (one example is that used in [4] for Higgs CP study in VBF production channel and will be called as $\phi_{4\pi}$), this can achieve higher sensitivity if the neutrino can be reconstructed with high precision. This can be seen from Fig. 6(a) for $\rho + \rho$ mode using $\phi_{4\pi}$ as an example. In this comparison, the truth-level momenta are used to reconstruct all variables for simplicity, and we stress here that the neutrino reconstruction precision will influence the final sensitivity.

After folding in the detector efficiency and resolution effects, and imposing the selection cuts described previously, all decay channels listed above are used to estimate the sensitivity. The result is presented in Fig. 6(b) for 300 fb$^{-1}$ (solid line) and also 3 ab$^{-1}$ (dashed line) luminosity. The 1-$\sigma$ precision at 300 fb$^{-1}$ can reach 27$^\circ$ (0.47) and can further be pushed down to 6.9$^\circ$ (0.12) at 3 ab$^{-1}$.

In conclusion, a new method is described in this paper to measure the CP-violation effect in the h$\tau\tau$ interaction, based on the neutrino momentum reconstruction and matrix element. All major hadronic tau decay modes are included simultaneously. With detailed detector simulation, it is predicted that at 13 TeV LHC with 300 fb$^{-1}$ (3 ab$^{-1}$) integrated luminosity, a precision up to 27$^\circ$ (6.9$^\circ$) can be achieved for the CP-mixing angle ($\phi$) measurement, significantly improving the previous predictions using only particular tau decay modes or partial event reconstructions. Discussion: (1) It is expected that if a MVA method is used [35], the signal purity can be further improved which is not tried in this work. (2) Although the gluon-gluon-fusion Higgs signal is not included in our calculation, its relative contribution at the order of 25% in the VBF signal region can further improve the result to some extent.

The authors would like to thank Tao Han for the helpful discussion. XC is thankful to the support from the National Thousand Young Talents program and the NSFC of China.

References

1. xin.chen@cern.ch
2. wuyongcheng12@mails.tsinghua.edu.cn
3. [1] G. Aad et al. (ATLAS), Phys. Lett. B726, 120 (2013), arXiv:1307.4342 [hep-ex].
4. [2] S. Chatrchyan et al. (CMS), Phys. Rev. Lett. 110, 081803 (2013), arXiv:1212.6639 [hep-ex].
5. [3] S. Chatrchyan et al. (CMS), Phys. Rev. D89, 092007 (2014), arXiv:1312.5353 [hep-ex].
6. [4] T. Han, S. Mukhopadhyay, B. Mukhopadhyaya, and Y. Wu, JHEP 05, 128 (2017), arXiv:1612.00413 [hep-ph].
7. [5] X. Chen and Y. Wu, (2017), arXiv:1703.04855 [hep-ph].
8. [6] K. Hagiwara, K. Ma, and S. Mori, Phys. Rev. Lett. 118, 171802 (2017), arXiv:1609.00943 [hep-ph].

FIG. 4. The distribution of the angle $\phi_{\text{ME}}$ in truth and after the fit for the $a_1\nu + \pi\nu$ channel in the pure CP even $h \to \tau\tau$ signal.

FIG. 5. The expected distribution of $\phi_{\text{ME}}$ for signal process only at the 13 TeV 300 fb$^{-1}$ LHC after all selection cuts for three different choices of CP mixing angle $\phi$: 0.0 (black line), 0.64 (red line) and 1.57 (blue line).
FIG. 6. The ΔNLL as a function of the CP mixing angle φ for 300 fb⁻¹ (solid line) and 3000 fb⁻¹ (dashed line). In Panel (a) only ρ + ρ mode is used and the LL is calculated from the expected zero-CP events in φME distribution (black line) and also in φe distribution (red line) both constructed from truth-level information, while in Panel (b) all tau decay channels are used and the LL is calculated from the expected zero-CP events in φME distribution constructed from reconstructed information.