Prospect for measuring the CP phase in the $h\tau\tau$ coupling at the LHC

Andrew Askew, Prerit Jaiswal, Takemichi Okui, Harrison B. Prosper, and Nobuo Sato

Phys. Rev. D **91**, 075014 — Published 21 April 2015
DOI: 10.1103/PhysRevD.91.075014

Prospect for measuring the CP phase in the $h\tau\tau$ coupling at the LHC

Andrew Askew,[1] Prerit Jaiswal,[2,1] Takemichi Okui,[1] Harrison B. Prosper,[1] and Nobuo Sato[3,1]

[1] *Department of Physics, Florida State University, Tallahassee, FL 32306, USA*
[2] *Department of Physics, Syracuse University, Syracuse, NY 13244, USA*
[3] *Jefferson Lab, Newport News, Virginia 23606, USA*

The search for a new source of CP violation is one of the most important endeavors in particle physics. A particularly interesting way to perform this search is to probe the CP phase in the $h\tau\tau$ coupling, as the phase is currently completely unconstrained by all existing data. Recently, a novel variable Θ was proposed for measuring the CP phase in the $h\tau\tau$ coupling through the $\tau^\pm \to \pi^\pm\pi^0\nu$ decay mode. We examine two crucial questions that the real LHC detectors must face, namely, the issue of neutrino reconstruction and the effects of finite detector resolution. For the former, we find strong evidence that the collinear approximation is the best for the Θ variable. For the latter, we find that the angular resolution is actually not an issue even though the reconstruction of Θ requires resolving the highly collimated $\pi^\pm$'s and π^0's from the τ decays. Instead, we find that it is the missing transverse energy resolution that significantly limits the LHC reach for measuring the CP phase via Θ. With the current missing energy resolution, we find that with $\sim 1000\,\mathrm{fb}^{-1}$ the CP phase hypotheses $\Delta = 0^\circ$ (the standard model value) and $\Delta = 90^\circ$ can be distinguished, at most, at the 95% confidence level.

I. INTRODUCTION

The Standard Model (SM) of particle physics has been extremely successful in explaining all microscopic phenomena observed down to distances of order $\sim 1\,\mathrm{TeV}^{-1}$. However, the model has two notable under-explored corners: the neutrino and Higgs sectors. Having discovered the Higgs boson the next major goal of the Large Hadron Collider (LHC) is to scrutinize the properties of the Higgs boson and search for deviations from the SM predictions.

A particularly interesting question is whether CP is a good symmetry of the Higgs sector. Since all Higgs boson couplings are predicted to be CP even in the SM, any observation of a CP-odd or CP-violating component will constitute unambiguous evidence of physics beyond the SM. CP-violating Higgs interactions may be motivated by the fact that baryogenesis requires the existence of a new source of CP violation and the possibility that such a primordial source of new CP violation may leave traces in the couplings of the Higgs boson.

Recently, a novel variable Θ for measuring the CP phase in the h-τ-τ coupling was proposed in Ref. [1]. The distribution of Θ was shown to have the simple form $c - a\cos(\Theta - 2\Delta)$, where Δ is the phase of the h-τ-τ Yukawa coupling, with $\Delta = 0^\circ$ and 90° corresponding to the CP-even and -odd limits, respectively, while a and c are independent of Δ and Θ. What distinguishes the h-τ-τ Yukawa coupling from other couplings is that its *phase* (i.e. the CP angle Δ) is currently completely unconstrained by data [2], even though its *magnitude* (i.e. the rate of τ^+-τ^- production) is roughly consistent with the SM prediction [3, 4]. Field theoretically, as discussed in Ref. [1], an $\mathcal{O}(1)$ Δ in the h-τ-τ coupling can easily coexist with other Higgs boson couplings whose CP phases are known to be small, such as h-t-$\bar{t}$ and h-b-$\bar{b}$ ($\Delta_t \lesssim 0.01$ and $\Delta_b \lesssim 0.1$, respectively [2]), and h-Z-Z (for which $\Delta_Z = 90^\circ$ is disfavored [5–7]). In Ref. [1] using parton-level analyses it is estimated that the h-τ-τ CP phase, Δ, may be measured from the Θ distribution to an accuracy of $\sim 11^\circ$ at the 14-TeV LHC with $\sim 3000\,\mathrm{fb}^{-1}$ of data, or the $\Delta = 0^\circ$ and $\Delta = 90^\circ$ cases may be distinguished at 5σ level with $\sim 1000\,\mathrm{fb}^{-1}$. However, these conclusions are for an ideal detector. In this paper, we examine the impact of realistic detector effects on the measurement of Δ via Θ.

There are various ways that detector effects can affect those estimates. Most importantly, the reconstruction of the Θ variable requires resolving the charged and neutral pions in the $\tau^\pm \to \rho^\pm\nu \to \pi^\pm\pi^0\nu$ decay mode, but these $\tau^\pm$ are highly boosted as they come from higgs decays. One therefore expects the two photons from each π^0 decay to hit the same tower of the electromagnetic calorimeter (ECAL), thereby making the standard identification of a π^0 using photons useless. Moreover, one expects the same ECAL tower to be hit by the $\pi^\pm$, raising the question of how much of the ECAL tower's activity should be attributed to the π^0.

Another obvious issue is the reconstruction of the neutrinos. In Ref. [1], the so-called collinear approximation [8] is adopted, where the neutrino from each τ decay is taken to be exactly collinear with its parent τ. As shown in Ref. [1], the collinear approximation results in a 60–70% reduction of the amplitude of the $\cos(\Theta - 2\Delta)$ curve. So, one might expect that here lies a significant opportunity for improvement. In this paper, however, we present a theoretical calculation that strongly suggests that it is extremely difficult, if not impossible, to outperform the collinear approximation for the purpose of Θ reconstruction.

Adopting the collinear approximation, we then focus on the aforementioned issues of pion identifications in the presence of realistic detector effects, and propose a realistic algorithm to identify "τ candidates" specifically suited for constructing Θ. Surprisingly, we find that the amplitude of the $\cos(\Theta - 2\Delta)$ curve of the signal is virtually unchanged by the finite angular resolution of the

LHC detectors. Instead, we find that the measurement of Δ through the Θ distribution is significantly hampered by the contamination by the $Z \to \tau\tau$ background. This degradation arises due to resolution effects in reconstructing the $\tau\bar{\tau}$ invariant mass as a consequence of large uncertainties in the measurement of the missing transverse energy. Fortunately, this is an area where future improvement can be hoped for, unlike the angular resolution. We leave the important question of pileup effects for future work.

While we focus on the measurement of Δ through the Θ variable, similar but different acoplanarity angles have been studied both for e^+e^- colliders [9–14] as well as for the LHC [15–20]. For the $\tau^\pm \to \pi^\pm\pi^0\nu$ decay mode, which is the most dominant decay mode of the τ, the acoplanarity angles used in those studies are not as effective as the Θ variable at e^+e^- colliders in which the neutrino momenta can be reconstructed up to a two-fold ambiguity [1]. An advantage of the LHC studies mentioned above is that they can be applied to all 1-prong τ decays including leptonic decays. However, realistic detector effects are not considered in those studies. Ref. [21] performs a realistic study for the LHC but assuming universal CP violation for all SM fermions, unlike our study in which we assume Δ is non-zero only for the τ. Finally, the use of the τ impact parameter associated with a displaced τ vertex was studied in Ref. [22] in the e^+e^- context. Our Θ variable does not depend on the measurement of the τ impact parameter.

The paper is organized as follows. In Section II, we review the Θ variable and rewrite it in a simple form by expanding it in terms of a small parameter λ that characterizes the degree of collinearity of the decay products of the taus, with $\lambda \to 0$ in the exactly collinear limit. We use this expansion to argue that the collinear approximation appears the best we can do for the purpose of reconstructing Θ. Section III describes the parton level simulation of the signal and background processes. Section IV describes the detector simulation and selection of the parametrized detector quantities corresponding to the Higgs boson decay products. Section V evaluates the results of the simulation and sensitivity as a function of integrated luminosity, and we conclude in Section VI.

II. THEORY

A. The CP phase Δ and the variable Θ

Following Ref. [1], we define the CP phase Δ by

$$\mathcal{L}_{\tau\text{-Yukawa}} = -\frac{y_\tau}{\sqrt{2}} h\bar{\tau}(\cos\Delta + \mathrm{i}\gamma_5 \sin\Delta)\tau\,, \tag{1}$$

where the scalar field h and Dirac spinor τ describe the Higgs boson and τ lepton, respectively. In order to maintain the SM decay rate for $h \to \tau\tau$, we fix the h-τ-τ Yukawa coupling, y_τ, to its SM value $y_\tau^{\rm SM} = m_\tau/v$ with $v \simeq 174\,\mathrm{GeV}$, but keep the CP phase Δ as a free parameter to be measured. The SM prediction is $\Delta = 0$.[1]

We are interested in the decay chain

$$\begin{aligned} h &\to \tau^+\tau^- \\ &\to \rho^+\bar{\nu}\,\rho^-\nu \\ &\to \pi^+\pi^0\bar{\nu}\,\pi^-\pi^0\nu\,. \end{aligned} \tag{2}$$

As calculated in Ref. [1], the squared matrix element for this decay has the form

$$|\mathcal{M}_{h\to\pi\pi\bar{\nu}\,\pi\pi\nu}|^2 \propto c - a\cos(\Theta - 2\Delta)\,, \tag{3}$$

where a and c are both positive and depend neither on Θ nor Δ. The angular variable Θ is defined through

$$\begin{aligned} &\cos\Theta \propto (k_1\!\cdot\! p_{\tau 2})\,(k_2\!\cdot\! p_{\tau 1}) - (p_{\tau 1}\!\cdot\! p_{\tau 2})\,(k_1\!\cdot\! k_2)\,, \\ &\sin\Theta \propto \epsilon_{\mu\nu\rho\sigma}\, k_1^\mu\, p_{\tau 1}^\nu\, k_2^\rho\, p_{\tau 2}^\sigma\,, \end{aligned} \tag{4}$$

with a common proportionality factor. The subscripts 1 and 2 refer to the τ^+ and τ^- branches of the decay, respectively (e.g., $p_{\tau 1}^\mu = p_{\tau^+}^\mu$ and $p_{\tau 2}^\mu = p_{\tau^-}^\mu$). The 4-momenta k_i^μ $(i = 1, 2)$ are defined by

$$k_i^\mu \equiv y_i q_i^\mu + r p_{\nu i}^\mu, \tag{5}$$

with

$$q_i^\mu \equiv p_{\pi^\pm i}^\mu - p_{\pi^0 i}^\mu\,, \tag{6}$$

and

$$y_i \equiv \frac{2q_i\!\cdot\! p_{\tau i}}{m_\tau^2 + m_\rho^2}\,, \quad r \equiv \frac{m_\rho^2 - 4m_\pi^2}{m_\tau^2 + m_\rho^2} \approx 0.14\,. \tag{7}$$

B. The boosted $\tau^\pm$ limit

The fact that the taus from Higgs boson decay are highly boosted can be exploited to obtain an even simpler expression for Θ. We define two lightlike 4-vectors n_1^μ and n_2^μ whose spatial components, $\vec{n}_1$ and $\vec{n}_2$, are unit vectors taken along the 3-momenta of the $\rho^\pm$ from the τ decays:

$$n_1 \equiv (1, \hat{n}_{\rho 1})\,, \quad n_2 \equiv (1, \hat{n}_{\rho 2})\,. \tag{8}$$

We also define

$$\bar{n}_1 \equiv (1, -\hat{n}_{\rho 1})\,, \quad \bar{n}_2 \equiv (1, -\hat{n}_{\rho 2})\,. \tag{9}$$

Then, an arbitrary 4-momentum a^μ can be expanded in terms of n_1 and $\bar{n}_1$ as

$$\begin{aligned} a^\mu &= (a\cdot\bar{n}_1)\,\frac{n_1^\mu}{2} + (a\cdot n_1)\,\frac{\bar{n}_1^\mu}{2} + a^{\perp\mu} \\ &\equiv a^+\frac{n_1^\mu}{2} + a^-\frac{\bar{n}_1^\mu}{2} + a^{\perp\mu}\,, \end{aligned} \tag{10}$$

[1] For a fully SU(2) × U(1) gauge-invariant form of the interaction (1) as well as an example of renormalizable realizations of it, see Ref. [1].

where $a^\perp$ by definition satisfies $a^\perp \cdot n_1 = a^\perp \cdot \bar{n}_1 = 0$. Alternatively, a^μ can be expanded in a similar manner in terms of n_2 and $\bar{n}_2$. To avoid notational clutter, we do not make explicit whether $a^\pm$ and $a^\perp$ are defined in the n_1-$\bar{n}_1$ basis or the n_2-$\bar{n}_2$ basis, but it should be clear that the momenta of all particles originating from the τ^+ will be expanded in the n_1-$\bar{n}_1$ basis, while those from τ^- will be in n_2-$\bar{n}_2$.

Next, we introduce a small expansion parameter $\lambda \ll 1$ to quantify the degree of collinearity between the $\tau^\pm$ and the respective $\rho^\pm$. First, by construction, "collinear" means that p^μ is dominated by the $+$ component, p^+, while the remaining components $p^\perp$ and p^- are small. Therefore, by definition, we assign the λ scaling law $p^\perp \sim \lambda p^+$ to the $\perp$ component. This implies that numerically we have $\lambda \sim m_\tau/m_h \sim 10^{-2}$. Then, in the ultra-relativistic limit, we have $0 \simeq p \cdot p = p^+ p^- + p^\perp \cdot p^\perp \sim p^+ p^- + \mathcal{O}(\lambda^2 p^+ p^+)$, so we see that $p^- \sim \lambda^2 p^+$. To summarize, we characterize the collinearity of p as

$$(p^+, p^-, p^\perp) \sim (1, \lambda^2, \lambda)\, p^+ . \tag{11}$$

C. The reconstruction of neutrinos

Since neutrinos are invisible at the LHC, we must devise a method to assign momenta to the neutrinos from the $\tau^\pm$ decays. However, the choice of method is a delicate issue for the reconstruction of the Θ variable. Performing the λ expansion in the expressions (4), we see that all $\mathcal{O}(\lambda^0)$ and $\mathcal{O}(\lambda^1)$ terms exactly cancel, so we are left with

$$\begin{aligned} (k_1 \cdot p_{\tau 2})\,(k_2 \cdot p_{\tau 1}) - (p_{\tau 1} \cdot p_{\tau 2})\,(k_1 \cdot k_2) \sim \mathcal{O}(\lambda^2)\,, \\ \epsilon_{\mu\nu\rho\sigma}\, k_1^\mu\, p_{\tau 1}^\nu\, k_2^\rho\, p_{\tau 2}^\sigma \sim \mathcal{O}(\lambda^2)\,. \end{aligned} \tag{12}$$

Therefore, for the purpose of reconstructing Θ, values must be carefully assigned to the invisible neutrino momenta so as not to disturb these delicate cancellations. In particular, the λ scaling law (11) must be respected.

A simple method with this property is provided by the *collinear approximation* [8], where we set

$$p_{\tau i}^{\perp\mu} = p_{\tau i}^- = 0\,, \quad p_{\nu i}^{\perp\mu} = p_{\nu i}^- = 0\,, \tag{13}$$

for each i. In other words, the collinear approximation applies the $\lambda \to 0$ limit to the τ and ν momenta. Then, to $\mathcal{O}(\lambda^2)$, the expressions (4) become

$$\begin{aligned} &\cos\Theta \propto (k_1^\perp \cdot n_2)\,(k_2^\perp \cdot n_1) - (n_1 \cdot n_2)\,(k_1^\perp \cdot k_2^\perp)\,, \\ &\sin\Theta \propto \epsilon_{\mu\nu\rho\sigma}\, k_1^{\perp\mu} n_1^\nu\, k_2^{\perp\rho} n_2^\sigma\,, \end{aligned} \tag{14}$$

again with a common proportionality factor. Finally, noticing that the definitions of k_i^μ in (5) give $k_i^{\perp\mu} = y_i q_i^{\perp\mu}$ in the collinear approximation, we arrive at the following very simple expressions that determine Θ:

$$\begin{aligned} &\cos\Theta \propto (q_1^\perp \cdot n_2)\,(q_2^\perp \cdot n_1) - (n_1 \cdot n_2)\,(q_1^\perp \cdot q_2^\perp)\,, \\ &\sin\Theta \propto \epsilon_{\mu\nu\rho\sigma}\, q_1^{\perp\mu} n_1^\nu\, q_2^{\perp\rho} n_2^\sigma\,, \end{aligned} \tag{15}$$

with a common proportionality factor. The 4-vectors q_i^μ, n_1^μ and n_2^μ are defined in (6) and (8), respectively. In the rest of the paper, we use the expressions (15) to compute Θ unless noted otherwise.

One might think we should be able to do better than the collinear approximation by using statistical correlations of the neutrino momenta with other, visible, momenta in the process. However, the delicate cancellation shown in (12) implies that the correlations must get the neutrino momenta correctly at the level of $\mathcal{O}(\lambda^2) \sim 10^{-4}$ in order to outperform the collinear approximation. Although we could not prove that this is impossible, it certainly suggests that it is extremely difficult.

III. PARTON-LEVEL SIMULATION

In this section, we outline our generation of the parton level samples used in the detector simulation. We generated tree level event samples with a combination of `MadGraph5` [23] and `MCFM` as explained below for the background ($Z + j$) and the signal ($h + j$) with $p_{\mathrm{T\,min}} = 70$ GeV for the jet. The associated jet facilitates triggering and provides sufficient $\tau^+\tau^-$ transverse momentum for the collinear approximation to hold. Other associated production modes could be used, notably $h + Z$ and $h + W$, but they have lower cross sections than $h + j$, though potentially better signal to noise that could be exploited in a further study. Multijet events are a potential background; however, this background is found to be small once tau selection criteria are imposed. Therefore, consideration of $Z + j$ alone is sufficient for our purposes.

The jet-p_{T} cut was imposed at the generator level in anticipation of the high-p_{T} jet requirement we impose in our analysis to ensure that the event will pass the trigger as well as to avoid kinematic configurations where the $\tau^+\tau^-$ is produced back-to-back in the transverse plane and the collinear approximation for the neutrino momenta fails.

The model files were generated using `FeynRules` package [24] by modifying the SM template provided by `FeynRules` to include the CP violating coupling in the h-τ-τ vertex as well as the effective vertices for the decay of $\tau \to \rho + \nu$ and $\rho \to \pi + \pi^0$ given by

$$\mathcal{L}_{\text{eff}} = -\rho_\mu^*\, \bar{\nu}\gamma^\mu P_{\mathrm{L}} \tau - \rho^\mu \big(\pi^0 \partial_\mu \pi^* - \pi^* \partial_\mu \pi^0\big) + \text{c.c.}\,, \tag{16}$$

where $P_{\mathrm{L}} \equiv (1-\gamma_5)/2$ is the projection operator onto left-handed chirality. We implement the Higgs boson production from gg fusion using the effective field theory (EFT) (in which the top quark is integrated out) available from the `FeynRules` repository. In principle, such an effective vertex is applicable only when the energy scale of the process lies well below the mass of the top quark. However, using the `MCFM` program [25], we have verified that the Higgs boson p_{T} distribution from the EFT in the region of interest ($70 < p_{\mathrm{T}} < 200$ GeV) agrees at the percent level with the full 1-loop QCD calculation. Finally, note that all the coupling constants in Eq. (16) are set to unity. To get the physical cross section, we use the cross sections from `MCFM` re-weighted by $\tau \to \rho + \nu$ and $\rho \to \pi + \pi^0$ branching fractions. The CTEQ6l1 PDF set was used throughout.

Parton showering and hadronization effects are included by processing the `Madgraph5` samples using the `Pythia8` program [26]. In principle, higher order corrections can be implemented by including events with higher jet multiplicity from tree level matrix elements and performing the corresponding jet merging with `Pythia`'s parton showering. However, we have verified that the inclusion of these corrections do not affect the Higgs boson p_T distribution and they differ at most by a few percent in the region of interest.

IV. DETECTOR SIMULATION

In order to incorporate detector effects into the measurement of the Θ angle, we use the Snowmass detector model [27], implemented in `Delphes 3` [28]. In `Delphes`, energy from particles deposited within the calorimeter is partitioned between the electromagnetic and hadronic sections according to a fixed ratio. This ratio is unity for the fraction of energy from a charged pion deposited in the hadronic calorimeter, which fails to account for energy depositions from hadronic showers that begin before the hadronic calorimeter. We account for this effect, including potential systematic effects due to overlap between the energy deposition of the neutral pion and the charged pion, by distributing the energy of charged pions in both calorimeter sections according to a probability distribution that models the one given in Ref. [29] for 30 GeV pions. 85% percent of the time the particle is presumed to deposit energy via ionization only, and the energy is entirely deposited within the hadronic calorimeter. For the remaining 15% of events, the fraction of the particle energy deposited in the electromagnetic calorimeter is drawn from a Gaussian distribution of mean 0.4 and width 0.3 (selecting only physical values). The rest of the energy is assigned to the hadronic calorimeter. The Gaussian models the probability with which the pion interacts early, in which case a significant fraction of its energy will be deposited in the ECAL.

Tau identification in `Delphes` is abstracted to an efficiency, which is insufficient for our purposes. Therefore, we develop an algorithm to identify taus explicitly that assigns the charged pion and the neutral pion to a track and a calorimeter tower, respectively. We select jets reconstructed with the anti-k_T algorithm [30] with a distance parameter of 0.5, which are reconstructed in `Delphes` using `FastJet` [31]. At least three jets are required in each event with $p_T > 30$ GeV within $|\eta| < 2.5$ (the tracking limit of the detector), of which the lead jet must have $p_T > 100$ GeV (well above the parton level requirement of 70 GeV). All jets with $p_T >$30 GeV are examined as potential τ-candidates. The lead track within $\Delta R < 0.2$ of the jet axis is selected, and the sum of the transverse momenta of the remaining tracks, within $\Delta R < 0.4$, is required to be less than 7 GeV in order to exclude jets that are unlikely to be taus. These selection requirements emulate the more sophisticated selection of taus in the LHC experiments, and are treated as equivalent to the `Delphes` efficiency of 65%.

We make further selections on the individual elements within the tau candidate in order to ensure that the decay products have been measured with sufficient precision for the calculation of Θ. The lead track is required to have $p_T > 5$ GeV, or this candidate is rejected. We then proceed to examine calorimeter towers within $\Delta R < 0.2$ of the jet axis. We select the tower with the largest transverse energy in the electromagnetic portion of the tower. The tower in question must also have an electromagnetic fraction (EMF) of 0.9, that is, at least 90% of its total energy within the ECAL. If this tower has $p_T > 5$ GeV, is within $\Delta R < 0.2$ of the selected track, and the sum of the selected tower transverse energy and the track momentum is greater than 20 GeV, then this jet becomes a τ-candidate suitable for Θ reconstruction, and the charged pion is assumed to be the selected track with the neutral pion the selected tower.

Finally, the electromagnetic portion of the actual ATLAS and CMS detectors have finer granularity than a single tower, which is not properly accounted for by the Snowmass detector. To approximate a position resolution of one detector element of the LHC detectors, we introduce the resolution for the position in η-ϕ of the neutral pion as a Gaussian smearing with standard deviation $0.025/\sqrt{12}$. A resolution no better than a tower (as in the default Snowmass detector) was found to degrade significantly the sensitivity to the modulation in the Θ distribution. Finally, events are required to have at least two τ-candidates with opposite charges. Should there be more than two τ-candidates, the highest p_T pair with opposite charge is selected.

V. RESULTS

We now come to the critical question of whether the signal modulation survives detector effects. In order to answer that question, consider the analysis cut-flow shown in Table I. The signal acceptance – the ratio of the second number to the first in column two of Table I – is 0.430, while that of the background is 0.254. The products of the efficiencies of all the remaining cuts including the 65% τ reconstruction efficiency are $0.65^2 \times 143979/1078279 = 0.0564$ and $0.65^2 \times 134628/1069747 = 0.0531$ for the signal and background, respectively. Therefore, overall, the signal and background efficiencies are $0.430 \times 0.0564 = 2.42 \times 10^{-2}$ and $0.254 \times 0.0531 = 1.35 \times 10^{-2}$, respectively.

To see the effects of finite detector resolution on our ability to reconstruct the signal Θ distribution (3), we

Selection criterion	Z+j→ $\tau\tau$+j	h+j→ $\tau\tau$+j
None	9618732	4698651
At least three jets $p_\mathrm{T} > 30\,\mathrm{GeV}$, lead jet $p_\mathrm{T} > 100\,\mathrm{GeV}$	2445860	2019316
Two τ-candidates only ΔR and isolation requirements	1069747	1078279
τ-track $p_\mathrm{T} > 5\,\mathrm{GeV}$	903429	941368
Tower $E_\mathrm{T} > 5\,\mathrm{GeV}$	804566	856299
Tower EMF > 0.9	156140	160312
Sum $p_\mathrm{T} > 20\,\mathrm{GeV}$	143508	149591
Opposite Charge	134628	143979

TABLE I: Cut-flows for Z→ $\tau\tau$ and $h \to \tau\tau$. The generator and reconstruction level cuts on the jet are $p_\mathrm{T} > 70\,\mathrm{GeV}$, and $100\,\mathrm{GeV}$, respectively. See text for more details.

parameters	MC Truth	Delphes Sim.
c	0.0397 ± 0.0001	0.0397 ± 0.0002
a	0.0089 ± 0.0001	0.0085 ± 0.0002
γ	1.03 ± 0.01	1.03 ± 0.02
δ	-0.003 ± 0.003	0.01 ± 0.01
$\alpha \equiv a/c$	0.225 ± 0.002	0.214 ± 0.004

TABLE II: Fits to the signal Θ distribution before and after the inclusion of detector effects using the functional form $c - a\cos(\gamma\Theta - 2\delta)$.

present in Fig. 1 the reconstructed Θ distribution from the signal samples generated with $\Delta = 0$ before and after the inclusion of detector effects, labelled as *Monte Carlo Truth* and *Delphes Simulation*, respectively. We perform fits to these distributions using the functional form $c - a\cos(\gamma\Theta - 2\delta)$ with free parameters a, c, γ, and δ in order to extract the modulation amplitude $\alpha \equiv a/c$. The fit results are presented in Table II and show that $\gamma \simeq 1$ and $\delta \simeq 0$ as expected from the analytical result (3) and the assumption $\Delta = 0$, even after detector effects. Most crucially, the dilution of the modulation amplitude α due to the detector effects is only about $\sim 4\%$, going down from 0.225 to 0.214.

In order to check that our τ reconstruction algorithm does not introduce artificial modulations in the background Θ distribution, we present in Fig. 2 the Monte Carlo Truth and Delphes Simulation Θ distributions using the background samples. Clearly, the background is consistent with a flat distribution after the detector effects and we may conclude that the latter do not bias the shape of the Θ distribution.

Now, we address the crucial question: how well can we distinguish the Θ distribution with $\Delta = 0$ from that with $\Delta \neq 0$? Given N events, each yielding a measurement of Θ_i, we construct a test statistic t designed to distinguish between the SM hypothesis $\Delta = 0$ and the alternatives

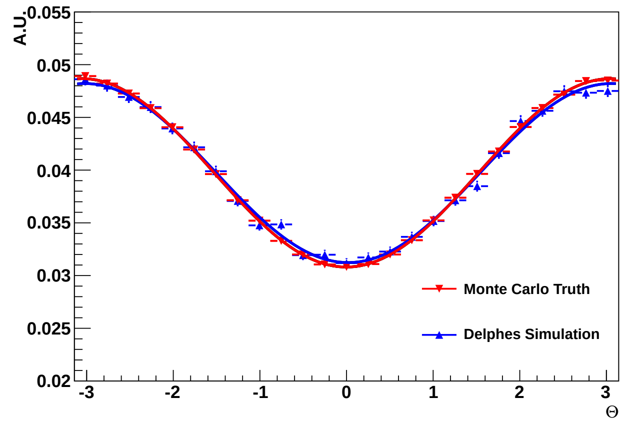

FIG. 1: Comparison of parton level Θ distribution (Monte Carlo Truth) with the distribution after the `Delphes` simulation and event reconstruction for $h \to \tau\tau$.

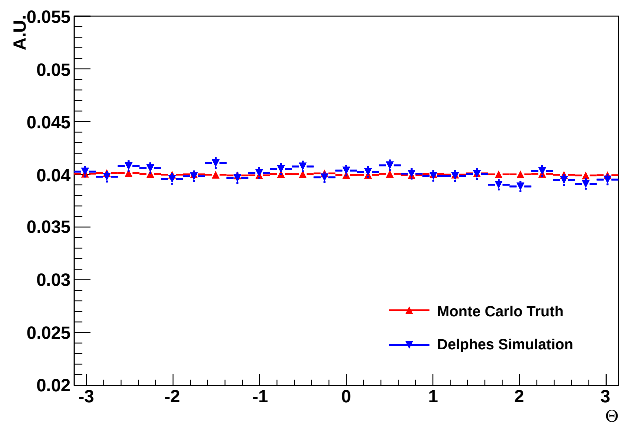

FIG. 2: Comparison of parton level Θ distribution (Monte Carlo Truth) with the distribution after the `Delphes` simulation and event reconstruction for $Z \to \tau\tau$.

$\Delta \neq 0$. A standard choice for t is

$$t = 2\ln\frac{L(\Delta)}{L(0)}\,, \tag{17}$$

where

$$L(\Delta) = \prod_{i=1}^{N} f(\Theta_i|\Delta) \tag{18}$$

is the likelihood function, while $f(\Theta_i|\Delta)$ is the Θ distribution for a given value of Δ.

In order to enhance the signal to background ratio (S/B), we use a boosted decision tree (BDT), trained using the invariant mass of the reconstructed $\tau^\pm$ system, the kinematic variables $(p_\mathrm{T}, \phi, \eta)$ of the final state visible particles, and the missing transverse energy. Therefore, each event can be characterized by the measurement of

Θ_i and the value of the BDT discriminant d_i, which yields the modified likelihood function,

$$L(\Delta) = \prod_{i=1}^{N} f(\Theta_i, d_i|\Delta)\,, \tag{19}$$

where the probability density function of the data $\{(\Theta, d)\}$ is modelled as follows,

$$\begin{aligned} &f(\Theta, d|\Delta) \\ &\propto B\,\rho^{\mathrm{B}}_{\mathrm{BDT}}(d) + S\,\rho^{\mathrm{S}}_{\mathrm{BDT}}(d)\left[1-\alpha\cos(\Theta-2\Delta)\right]\,, \end{aligned} \tag{20}$$

where $\rho^{\mathrm{S}}_{\mathrm{BDT}}(d)$ and $\rho^{\mathrm{B}}_{\mathrm{BDT}}(d)$ are, respectively, the signal and background BDT distributions for the discriminant parameter d, while B and S are, respectively, the total expected background and signal counts given by

$$B = \epsilon_{\mathrm{B}}\sigma_{\mathrm{B}}\mathcal{L}\,, \quad S = \epsilon_{\mathrm{S}}\sigma_{\mathrm{S}}\mathcal{L}\,. \tag{21}$$

Here, σ_{S} and σ_{B} are the associated cross section times branching fractions, the ϵ's are the signal and background efficiencies discussed above and $\mathcal{L}$ is the integrated luminosity.

In order to quantify how well one might be expected to distinguish a $\Delta \neq 0$ hypothesis from the SM hypothesis $\Delta = 0$, we generate three samples of Θ and d values with $\Delta = 0$, $\pi/4$, and $\pi/2$. For each sample, we calculate the values of t as defined in Eq. (17), which yield the t distributions as illustrated in Fig. 3 for the $\Delta = 0$ and $\pi/4$ samples. To quantify the discrimination between the two hypothesis, the p-value of the SM t distribution is computed using the median of the non-SM t distribution. As proxies for systematic uncertainties in the S and α parameters, the p-values are varied in the 10% neighborhood of their nominal values. In Fig. 4 we present the reach as a function of the integrated luminosity for $\Delta = \pi/4$ and $\pi/2$. Because we have chosen the "observed" value of t to be the median of the non-SM hypothesis, the reach in this context is interpreted as follows: if the non-SM hypothesis is true, then there is a 50% chance to reject the SM hypothesis $\Delta = 0$ at the Z-sigma level, where $Z = \Phi^{-1}(p) \equiv \sqrt{2}\,\mathrm{erf}^{-1}(2p-1)$ with p being the p-value shown in Fig. 3.

It is clear from Fig. 4 that the optimistic conclusion obtained for an ideal detector does not survive the smearing effects of a realistic detector and our reconstruction algorithms. At best, at ~ 1500 fb^{-1}, there is a 50:50 chance of rejecting the $\Delta = 0$ hypothesis at 95% confidence level. It is important to note where the poor performance stems from. As mentioned earlier, the reduction of the modulation amplitude α of the signal due to the angular resolution is negligible. Rather, we find that the discrimination power of the BDT dominantly comes from the invariant mass of the reconstructed τ^+-τ^- system, and consequently the S/B ratio is sensitive to the mass resolution, which in turn is sensitive to the resolution of the missing transverse energy (MET). Therefore, our analysis shows that the current MET resolution of the LHC detectors is too low to achieve a sufficient reduction of the total background count B.

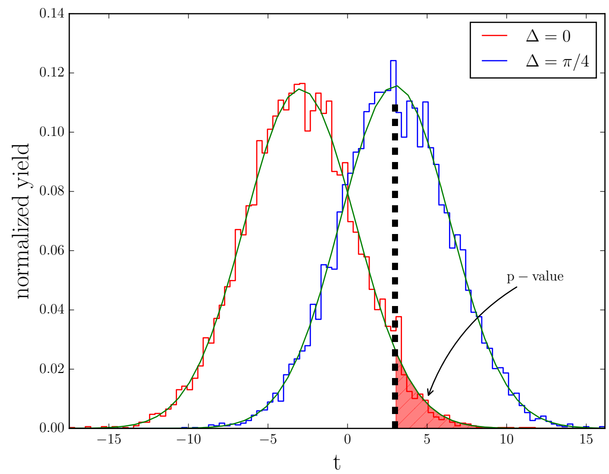

FIG. 3: Distributions of the t statistic for the SM hypothesis ($\Delta = 0$) and a non-SM hypothesis ($\Delta = \pi/4$). The "observed" value of the statistic is taken to be the median of the t distribution of the non-SM hypothesis.

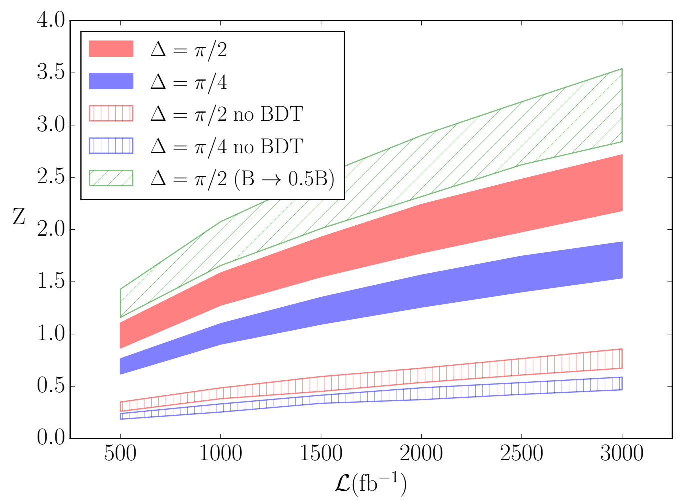

FIG. 4: The reach, quantified on the Z-sigma scale of the normal distribution, versus the integrated luminosity. The uppermost plot shows the effect of reducing the background by a factor of two, while keeping every else fixed.

To illustrate the effects of improved missing transverse energy resolution, we show in the uppermost plot in Fig. 4 the effect of increasing S/B by a factor of two by decreasing B by a factor two, while keeping everything else the same. Increasing the S/B by increasing S by a factor two (while keeping α unchanged) has a similar but more pronounced effect.

VI. CONCLUSIONS

We have examined a recent proposal [1] for measuring a potentially non-zero CP phase Δ in Higgs to $\tau^+\tau^-$ decays using an angular variable Θ that can be reconstructed from the decay chains $\tau^\pm \to \rho^\pm\nu \to \pi^\pm\pi^0\nu$. A non-zero phase of the h-τ-τ Yukawa coupling would be unambiguous evidence of physics beyond the SM. Contrary to previous expectations [1], with $\sim 1000\,\text{fb}^{-1}$ we find that the standard model CP phase, $\Delta = 0°$, can be distinguished from the $\Delta = 90°$ hypothesis at no more than 95% confidence level.

We have addressed two obvious effects that are expected to degrade our ability to measure Δ from the Θ distribution. First, we examined the issue of neutrino momentum reconstruction, for which the original proposal [1] employed the collinear approximation. We have presented a theoretical argument that strongly suggests that this cannot be improved further.

Second, we have closely studied how finite resolutions of a realistic detector affects our ability to distinguish between the SM signal and its deviations induced by a nonzero CP violating phase Δ, using the Snowmass detector model implemented in the `Delphes` detector simulation program. Contrary to the naive expectation that a finite angular resolution should hamper the reconstruction of Θ as it requires resolving the charged and neutral pions from the τ decays, we have found that the finite angular resolution of the LHC detectors reduces the modulation amplitude of the signal Θ distribution by only $\sim 4\%$, compared to that of an ideal detector. Instead, we have found that the experimental ability to distinguish the hypotheses $\Delta = 0$ from $\Delta \neq 0$ is strongly limited by the missing transverse energy resolution. The finite energy resolution significantly degrades the use of the $\tau\tau$ invariant mass distribution as the principal discriminant between the $h \to \tau\tau$ signal and the $Z \to \tau\tau$ background. In contrast, the reconstruction of Θ itself is nearly unaffected by the finite energy resolution, as the Θ variable is basically determined by the angular topology of the Higgs boson decay products. Fortunately, the resolution of missing transverse energy is an area in which improvements can be hoped for, e.g., by using multivariate methods [32].

Acknowledgement

This work was supported in part by the US Department of Energy under grant DE-FG02-13ER41942. The work of Nobuo Sato was partially supported by the US Department of Energy contract No. DE-AC05-06OR23177, under which Jefferson Science Associates, LLC operates Jefferson Lab.

[1] R. Harnik, A. Martin, T. Okui, R. Primulando, and F. Yu, Phys.Rev. **D88**, 076009 (2013), 1308.1094.
[2] J. Brod, U. Haisch, and J. Zupan, JHEP **1311**, 180 (2013), 1310.1385.
[3] S. Chatrchyan et al. (CMS Collaboration), JHEP **1405**, 104 (2014), 1401.5041.
[4] G. Aad et al. (ATLAS Collaboration), ATLAS-CONF-2014-061, ATLAS-COM-CONF-2014-080 (2014).
[5] S. Chatrchyan et al. (CMS Collaboration), Phys.Rev. **D89**, 092007 (2014), 1312.5353.
[6] G. Aad et al. (ATLAS Collaboration) (2014), 1408.5191.
[7] V. M. Abazov et al. (D0 Collaboration), Phys.Rev. **D88**, 032008 (2013), 1304.5422.
[8] R. K. Ellis, I. Hinchliffe, M. Soldate, and J. van der Bij, Nucl.Phys. **B297**, 221 (1988).
[9] J. R. Dell'Aquila and C. A. Nelson, Nucl.Phys. **B320**, 61 (1989).
[10] B. Grzadkowski and J. Gunion, Phys.Lett. **B350**, 218 (1995), hep-ph/9501339.
[11] G. Bower, T. Pierzchala, Z. Was, and M. Worek, Phys.Lett. **B543**, 227 (2002), hep-ph/0204292.
[12] M. Worek, Acta Phys.Polon. **B34**, 4549 (2003), hep-ph/0305082.
[13] K. Desch, A. Imhof, Z. Was, and M. Worek, Phys.Lett. **B579**, 157 (2004), hep-ph/0307331.
[14] S. Berge, W. Bernreuther, and H. Spiesberger, Phys.Lett. **B727**, 488 (2013), 1308.2674.
[15] S. Berge, W. Bernreuther, and J. Ziethe, Phys.Rev.Lett. **100**, 171605 (2008), 0801.2297.
[16] S. Berge and W. Bernreuther, Phys.Lett. **B671**, 470 (2009), 0812.1910.
[17] S. Berge, W. Bernreuther, B. Niepelt, and H. Spiesberger, Phys.Rev. **D84**, 116003 (2011), 1108.0670.
[18] S. Berge, W. Bernreuther, and H. Spiesberger (2012), 1208.1507.
[19] S. Berge, W. Bernreuther, and S. Kirchner, Eur.Phys.J. **C74**, 3164 (2014), 1408.0798.
[20] S. Berge, W. Bernreuther, and S. Kirchner (2014), 1410.6362.
[21] M. J. Dolan, P. Harris, M. Jankowiak, and M. Spannowsky, Phys.Rev. **D90**, 073008 (2014), 1406.3322.
[22] K. Desch, Z. Was, and M. Worek, Eur.Phys.J. **C29**, 491 (2003), hep-ph/0302046.
[23] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, et al., JHEP **1407**, 079 (2014), 1405.0301.
[24] C. Degrande, C. Duhr, B. Fuks, D. Grellscheid, O. Mattelaer, et al., Comput.Phys.Commun. **183**, 1201 (2012), 1108.2040.
[25] J. M. Campbell, R. K. Ellis, and C. Williams, JHEP **1107**, 018 (2011), 1105.0020.
[26] T. Sjostrand, S. Mrenna, and P. Z. Skands, Comput.Phys.Commun. **178**, 852 (2008), 0710.3820.
[27] J. Anderson, A. Avetisyan, R. Brock, S. Chekanov, T. Cohen, et al. (2013), 1309.1057.
[28] J. de Favereau et al. (DELPHES 3), JHEP **1402**, 057 (2014), 1307.6346.
[29] S. Abdullin et al. (USCMS Collaboration, ECAL/HCAL Collaboration), Eur.Phys.J. **C60**, 359 (2009).
[30] M. Cacciari, G. P. Salam, and G. Soyez, JHEP **04**, 063 (2008), 0802.1189.

[31] M. Cacciari, G. P. Salam, and G. Soyez, Eur.Phys.J. **C72**, 1896 (2012), 1111.6097.

[32] V. Khachatryan *et al.* [CMS Collaboration], JINST **10**, no. 02, P02006 (2015) [arXiv:1411.0511 [physics.ins-det]].