

Ditau-Jet Tagging and Boosted Higgses from a Multi-TeV Resonance

Andrey Katz,¹ Minho Son,² and Brock Tweedie^{3,4}

¹*Department of Physics, University of Maryland, College Park, MD 20742*

²*Department of Physics, Yale University, New Haven, CT 06511*

³*Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218*

⁴*Physics Department, Boston University, Boston, MA 02215*

New TeV-scale physics processes at the LHC can produce Higgs bosons with substantive transverse Lorentz boost, such that the Higgs's decay products are nominally contained in a single jet. In the case of a light Higgs decaying predominantly to $b\bar{b}$, previous studies have shown that these Higgs-jets can be identified by capitalizing on jet substructure techniques. In this work, we explore the possibility of also utilizing the subdominant but very distinctive decay $h \rightarrow \tau^+\tau^-$. To this end, we introduce the concept of a “ditau-jet,” or a jet consisting of two semi-collinear taus where one or both decay hadronically. We perform simple modifications to ordinary tau tagging methods to account for this configuration, and estimate tag rates of $O(50\%)$ and QCD mistag rates of $O(10^{-4} - 10^{-3})$ for $p_T \sim \text{TeV}$, even in the presence of pileup. We further demonstrate the feasibility of reconstructing the ditau invariant mass by using traditional \cancel{E}_T projection techniques. Given these tools, we estimate the sensitivity of the LHC for discovery of a multi-TeV Z' decaying to Zh , utilizing both leptonic and hadronic Z decay channels. The leptonic Z channel is limited due to low statistics, but the hadronic Z channel is potentially competitive with other searches.

I. INTRODUCTION

As the first collider capable of probing multi-TeV energy scales, the LHC will face a unique set of challenges in reconstructing some of its most extreme events. Particles which have traditionally been considered heavy can be produced with energies far in excess of their mass, leading us to consider novel classifications such as “top-jets,” “ W -jets,” “ Z -jets,” and “Higgs-jets.” While these objects can look uncomfortably similar to ordinary QCD jets, in the past several years a variety of techniques have been developed to robustly discriminate them [1–16].

In a previous paper [17], we made a detailed study of hadronic W -, Z -, and Higgs-jets in the context of the search for a multi-TeV Z' , which commonly arises in composite Higgs and 5D models as well as in simple extended gauge sectors (see, e.g., [18–21] and others in [17]). Our previous study was restricted to the case where the boson-jet consists of quark-antiquark, which is a shared option in the case that the Higgs is light and its decays dominated by $b\bar{b}$. However, since the mass and branching fractions of the Higgs are as yet unknown, it is important to consider a broader set of options.

In this paper, we explore the possibility of utilizing Higgs-jets generated by the decay $h \rightarrow \tau^+\tau^-$. We call these “ditau-jets.” For a standard Higgs near 120 GeV, which is favored by electroweak precision measurements, this mode represents about 7% of the total decay rate and quickly drops as the mass is raised. While this means that the $\tau^+\tau^-$ mode is always subleading, several considerations compel us to take a closer look at it. First, from the perspective of discovering multi-TeV physics that generates Higgs-jets, establishing additional search channels increases our chances of spotting something new, and increases our confidence if we do. Second, from the perspective of testing Higgs physics, obtaining a measurement of this rate with respect to $b\bar{b}$ can give us a measure of the Higgs couplings, and test whether the “Higgs” observed in a new physics process is indeed the one observed at lower energies. Finally, from the perspective of signal discrimination, a high- p_T ditau-jet is in principle much easier to spot than a high- p_T $b\bar{b}$ -jet, as we will see.

Of course, these observations are all strictly theoretical if the ditau-jet production rate is genuinely too small to see. In order to make this Higgs decay mode truly competitive without running the LHC ten times longer, there needs to be a tradeoff in branching fraction with another part of the event. The process $Z' \rightarrow Zh$ provides a ready example. Here, one

of the main discovery channels with a light Higgs is $Zh \rightarrow (l^+l^-)(b\bar{b})$ [17, 22, 23], with branching fraction penalties of roughly 7% on the Z and 70% on the Higgs. We can instead consider the channel $Zh \rightarrow (q\bar{q})(\tau^+\tau^-)$, with respective penalties of about 70% and 7%. In order for this to be a worthwhile tradeoff, we must demonstrate that a ditau-jet can be very cleanly identified, as was the case for the dileptonic Z in the first channel. If this can be accomplished with high efficiency and high purity, then the now-hadronic Z on the other side of the event can also be tagged using jet substructure (analogous to the $b\bar{b}$ Higgs-jet), and we in principle obtain Z' sensitivity at least as good as what was obtained in the untagged analysis of $Zh \rightarrow (l^+l^-)(b\bar{b})$ in [17]. The key question is: How well can we discriminate ditau-jets from QCD?

There are two basic facts about taus that should allow this discrimination to be accomplished quite well, at least in principle. The first is that almost 60% of ditau decays contain at least one lepton. The second is that hadronic tau decays consist of extremely tight cones of activity (mostly charged and neutral pions), with opening angles of $O(m_\tau/p_{T\tau})$. At progressively higher energies, these cones become progressively smaller. QCD jets, by contrast, are color-connected to the rest of the event and tend to radiate at all angles. Highly energetic hadronic taus therefore allow us an extreme form of the kind of color-discrimination explored in [24, 25]. This improves as the energy is raised, suggesting that the hadronic taus produced in TeV-scale processes should be particularly clean.

Nonetheless, the case of two nearby high- p_T taus has not been studied in detail before. The decays that proceed dileptonically should be straightforward to find, but cannot uniquely be ascribed to taus, and are rather rare (12% BR). The remaining decays contain at least one hadronic tau in close proximity to either a lepton or another hadronic tau. These configurations can fail normal tau and/or lepton identification criteria. To deal with this possibility, we propose simple extensions of tau and lepton isolation criteria such that two objects can be “mutually isolated,” i.e. individually isolated from tracker and ECAL activity not caused by the other object. This forms the basis of a ditau-tag, which can be applied in both semi-leptonic and fully hadronic ditau decays, and extends to TeV-scale energies and very small tau separations.

We first study mutual isolation in the context of 2 and 3 TeV $Z' \rightarrow Zh \rightarrow (l^+l^-)(\tau^+\tau^-)$, finding ditau-tagging efficiencies of order 50% with QCD rejection rates at the $10^{-4} - 10^{-3}$ level. We also demonstrate that reconstruction of the Higgs mass peak should be possible,

even given the presence of two neutrinos and nearly-collinear geometry. We proceed to estimate the Z' discovery reach in this channel and in the $(q\bar{q})(\tau^+\tau^-)$ channel, demonstrating that the latter is a possible competitor to the diboson final-state channels that we studied in [17].

The paper is organized as follows. In the next section we review tracker-based hadronic tau-tag techniques and introduce our ditau-tag. In the third section we categorize the performance of the tag and ditau mass reconstruction, and estimate Z' discovery potential. We conclude in section IV.

II. TAGGING BOOSTED TAUS AND DITAUS

In this section, we outline techniques for tagging hadronic taus at high p_T . We start by reviewing some of the techniques that will be applied to single hadronic taus at the LHC. We then move on the case relevant for boosted Higgs decays, namely tagging two taus in close proximity. This includes cases where one or both of the taus decays hadronically. In addition, 12% of ditaus will decay dileptonically, but for these we do not pursue any special techniques.

A. Review of hadronic tau tag

In some sense, hadronically decaying taus can be treated as “fat” leptons. For electrons and muons, a track and/or small cluster of electromagnetic activity is first identified, and then demanded to be isolated from any additional activity out to a certain ΔR range and up to a certain energy threshold. Hadronic taus are not so well-localized in the detectors, since they decay before reaching the inner tracker. However, the multiplicity of decay products is always fairly small, and their ΔR separation is characterized by the ratio $m_\tau/p_{T\tau}$, which becomes quite small at even modestly large energies. The practical way to identify a tau, then, is to define a small region of hadronic and electromagnetic activity in the detector, and insist that it is isolated from additional activity.

The first step is to define the region of interest. This is done using either a hard track or a hard calorimeter cell as a seed, around which further activity is sought out (see, e.g. [26, 27]). For typical analysis with taus, the choice of track seed or calorimeter seed is not so

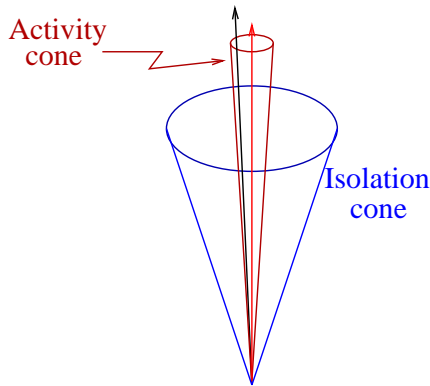


FIG. 1: *Illustration of tracker-seeded hadronic tau tag.*

crucial, and algorithms built on each of these have similar performance. However, as we will ultimately be concentrating on the TeV-scale momentum regime with *two* nearby taus, we choose to focus on a track-seeded approach, which should work more robustly when dealing with small angular scales. It would be interesting to further pursue the extension of calorimeter-seeded approaches to ditau-jets, but we do not do so here.

Let us now review the CMS particle-flow algorithm for tau-tagging, as described in [28]. After reconstructing jets with a $R = 0.5$ cone algorithm, an individual jet is investigated for constituent hadronic tracks. A track will qualify as a seed for hadronic tau-tagging if

- it is the hardest track in the jet
- it has $p_T > 5$ GeV
- it deviates by no more than $\Delta R = 0.1$ from the jet vector.

If one finds such a seed, at the next step an *inner cone* is defined around the seeding track, where all of the products of the hadronic tau decay are expected to be found. For sufficiently high p_T , the radius of this cone is taken to be 0.07.¹

The decay products of the tau should be inside this inner cone, and almost nothing should be outside. To enforce isolation, one further defines an isolation annulus, namely an outer cone of $R_{\text{outer}} = 0.45$ around the leading prong which omits the contents of the inner

¹ The CMS particle-flow tau-tag actually takes advantage of an inner cone radius that shrinks with p_T , tracing the expected physical shrinking of tau decay angles. However, the minimum cone size allowed is 0.07, which is applied to taus with $p_T > 70$ GeV.

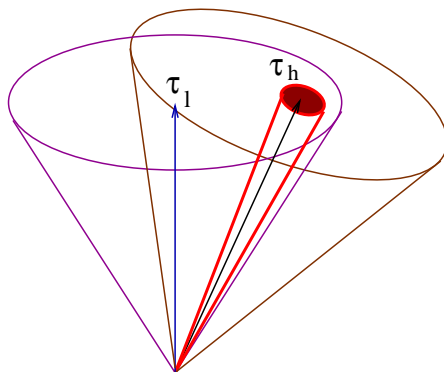


FIG. 2: *Illustration of mutual isolation cones and inner activity cone for a semileptonic ditau.*

cone (Fig. 1). Within the annulus, there should be no photons with $p_T > 1.5$ GeV and no tracks with $p_T > 1$ GeV. (The hadronic calorimeter, which is the least spatially-precise subdetector, is not utilized for this isolation.) Inner cones of activity which are constructed and isolated in this manner are considered tau-tagged. As shown in [28], the efficiency of this algorithm is about 80% in the central region of the detector, and the QCD fake rate is at the few percent scale. The fake rate declines as p_T is increased, while the tag rate plateaus. This makes sense physically, since the tau decay products become more collimated whereas QCD jets continue to radiate at all angles.

B. Ditau tag

Let us now imagine a ditau-jet, such as would be produced in the decay of a highly energetic light (120 GeV) Higgs from a TeV-scale process. We restrict ourselves to cases where at least one of the two taus decays hadronically. If the p_T of the Higgs is 1 TeV, then the ΔR separation between the two taus will be roughly 0.3, which would lead to failure of the single tau-tagging described above. However, the pattern of activity will still look highly unusual compared to an ordinary QCD jet, especially at the energies under consideration.

The simplest and most robust case is semileptonic, which accounts for about 45% of ditau decays. The lepton can nominally fail isolation because of the nearby hadronic tau, and vice versa. As a remedy, we propose the concept of *mutual isolation*. We call the lepton and τ -jet mutually isolated if the jet passes the criteria for a hadronic tau when the lepton is removed from consideration, and at the same time the lepton passes some form of isolation criterion with the tau inner cone removed. Step-by-step, this runs as follows:

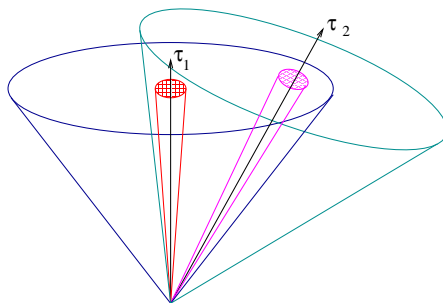


FIG. 3: *Illustration of mutual isolation cones and inner activity cones for an all-hadronic ditau.*

- find a lepton with $p_T > 30$ GeV which fails simple isolation criteria using an $R = 0.4$ cone
- find the hardest hadronic track inside this cone with $p_T > 20$ GeV, and bail out if none is found
- draw a small tau-candidate cone of radius $R_{\text{inner}} = 0.07$ around this seed track
- check whether the lepton passes a weakened isolation criterion using a surrounding $R = 0.4$ cone with this small cone deleted, accounting only for tracker and electromagnetic energy: $\frac{p_T(l)}{p_T(l)+p_T(\text{cone})} > 0.9$
- if the lepton passes, independently cluster all additional activity in the vicinity (we use $R_{\text{jet}} = 0.4$ Cambridge/Aachen(C/A) [29, 30])
- apply the single tau-tag the resulting jets, ignoring the lepton for purposes of annulus isolation

Schematically this algorithm is illustrated in Fig. 2. We will categorize its performance in the context of a resonance search the next section.

It is not difficult to extend this mutual isolation technique to the fully hadronic case, which accounts for about 42% of ditau decays. Now, we start with a candidate jet of $R_{\text{jet}} = 0.4$ (clustered with C/A, though the choice of algorithm is not crucial) and search inside of it for *two* tracks which can seed independent taus:

- from amongst the jet constituents, use the hardest charged hadron as the first seed, and then look for the hardest track $\Delta R > 0.08$ away from this for the second seed (again, each should be above 20 GeV in p_T)

- draw a small cone around each seed, picking the inner radius depending on their separation: $R_{\text{inner}} = \Delta R_{\text{seeds}}/2$ if the distance between the seeds is smaller than 0.14, or $R_{\text{inner}} = 0.07$ if the distance between the seeds is bigger²
- draw an isolation annulus ($R_{\text{outer}} = 0.4$) around each inner cone, with the other inner cone deleted, and check that each tau candidate passes tau isolation criteria
- if the candidates are mutually isolated, then use the original jet to define the ditau four-vector, and the seed directions to define the individual tau trajectories.

This procedure is sketched in Fig. 3. We also categorize its performance in the next section.

There are potentially several basic limitations to these procedures. An important one is the ability of the ATLAS and CMS detectors to cleanly identify tracks and photons at such high energies and small angular separations. In particular, tracking may fail and nearby ECAL cells will share energy to some extent. However, we do not expect that these will pose insurmountable problems. Ideally, a genuine hadronic tau consists of only one or three (or very rarely five) tracks, implying a very modest density of hits in the tracker. This low track (or hit) multiplicity could even be used to strengthen the discrimination against QCD. The caveat is that these tracks are almost always accompanied by photons from π^0 s, which can convert in the tracker (especially at high energy), and increase the activity. We have no way to model this effect in our study, but simply assume that nearby photon conversions do not dramatically degrade tracking quality as we move from 100 GeV taus to 1000 GeV taus. In the ECAL, energy will spread laterally, and we may worry that some fraction will enter the isolation annulus. But the characteristic distance scale of this leakage is fairly energy-independent and is of the same order as the cell size. Our inner cones are intentionally taken to be 3-4 cells in radius, and should contain the vast majority of the electromagnetic activity, especially since the photons are becoming more collinear at higher energies. In the studies below, we will nonetheless incorporate some of the detector effects by using a primitive detector model consisting of separate tracker, ECAL, and HCAL elements. This is mainly relevant for obtaining quasi-realistic estimates of energy and mass resolutions.

² The occurrence of $\Delta R_{\text{seeds}} < 0.14$ is not very common in our analysis below. In particular, accepting these configurations allows us to recapture about 5% of the dihadronic $h \rightarrow \tau^+\tau^-$ from 3 TeV resonance decay. However, a shrinking inner cone can be important for even higher-energy searches, or for ditau produced from boosted particles lighter than 120 GeV.

The other major worry is pileup, which will introduce a spray of uncorrelated particles into the isolation annuli. We find that if we simply accept these additional particles and run the ditau tags using the isolation criteria of subsection II A, the signal rates drop by a factor of order 3 due to isolation failures. However, pileup consists of min-bias events produced at distinct vertices. This means that pileup tracks can in principle be removed by tracing them back to the “wrong” vertex. Assuming that this can be done with high efficiency, we are still left with irreducible pileup neutrals, but these do not tend to be very energetic. We will see below that raising the annulus photon threshold from 1.5 GeV to 2.5 GeV is quite adequate. The tag rates end up only 10% lower, relatively speaking, than what is obtained without pileup using the original threshold, and the fake rates remain acceptable. It is also worth noting that even if pileup tracks cannot be efficiently removed by vertexing, using a slightly higher track p_T threshold would likely have similar impact.

III. HEAVY RESONANCE SEARCHES UTILIZING DITAU-JET TAGGING

In this section, we apply the ditau-tagging of section II to the case of a multi-TeV Z' decaying to Zh with a 120 GeV Standard Model-like Higgs boson. We begin with the case of $Zh \rightarrow (l^+l^-)(\tau^+\tau^-)$. Given that the branching fraction $h \rightarrow \tau^+\tau^-$ is only 7%, we do not expect this mode to be competitive with the much more prevalent $Zh \rightarrow (l^+l^-)(b\bar{b})$. Nonetheless, this serves as a clean context in which to check the performance of the ditau tag and to explore the possibility of Higgs mass reconstruction. Next, we turn to the more challenging mode $Zh \rightarrow (q\bar{q})(\tau^+\tau^-)$. The backgrounds here are potentially severe, but the starting rate now competes with $(l^+l^-)(b\bar{b})$. We will see that ditau-tagging and Higgs mass reconstruction can make the mode viable.³

The Z' samples for these studies are generated with $2 \rightarrow 4$ matrix elements in MadGraph/MadEvent v4.4.32 [31] and showered/hadronized in PYTHIA (plugin version 2.1.3, taus

³ There are other signal topologies which we might consider. For example, we could also gain some rate above $Z \rightarrow l^+l^-$ by using $Z \rightarrow \nu\bar{\nu}$. This can in principle be fairly clean, but the final state always contains sets of neutrinos which are approximately back-to-back, potentially seriously complicating Z' and Higgs reconstruction. Also, the mode $Zh \rightarrow (\tau^+\tau^-)(b\bar{b})$ is superficially nearly equivalent to $(q\bar{q})(\tau^+\tau^-)$, and could be mined for signal. However, the starting rate is about a factor of two smaller, and will be cut down by our Z and Higgs mass cuts. Combining in a dedicated analysis of this mode might also be useful to explore.

decayed with TAUOLA). The Z' width is taken to be 3%, which is smaller than the instrumental width and the signal windows that we use below. (Our estimates should work reasonably well for resonances up to about 15% natural width.) Our background samples are mainly generated with PYTHIA v6.4.11 [32] with default settings. This automatically includes prompt and radiative heavy flavor contributions, which can contaminate our semileptonic ditau analysis. The W -strahlung background is generated with MadGraph. All processes are evaluated at leading-order for a 14 TeV LHC. In addition to the hard collision, each event is superimposed with a pileup of min-bias collisions. The number of extra collisions is drawn from a Poisson distribution with a mean of 20. Charged pileup particles are subsequently removed, under the assumption that these could mostly be traced back to distinct vertices.

The particle-level events are further processed through a primitive detector model, corresponding to a kind of idealized particle-flow reconstruction. Charged particle four-vectors are left untouched. Photons are passed into a perfect energy sampling ECAL with granularity $\Delta\eta \times \Delta\phi = 0.02 \times 0.02$. Neutral hadrons are passed into a perfect energy sampling HCAL with granularity $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$. The calorimeter cells are then mapped into massless four-vectors. We cluster these detector-level particles into jets using FastJet 2.4.1 [33].

Leptons and jets (or subjects) are subsequently smeared. Electrons are smeared by 2%, and muons by $(5\%)\sqrt{E/\text{TeV}}$. Jets and subjects are smeared by the resolution quoted in the CMS TDR for $R = 0.5$ cone jets [34]: $\Delta E/E = 5.6/E \oplus 1.25/\sqrt{E} \oplus 0.033$, with \oplus indicating quadrature sum and E measured in GeV.

We estimate discovery reach for 2 and 3 TeV Z' by using simple counting. For each sample, we construct a box in the $(m_h^{\text{reco}}, m_{Z'}^{\text{reco}})$ plane, centered on the signal. The box is not optimized, other than coarsely by eye. We claim that discovery is possible if two criteria are met: 5σ significance according to Poisson statistics,⁴ and $N_S > 10$.

A. Leptonic Z and ditau tag efficiency

In this analysis, we first identify isolated leptons. Within a cone of size 0.4, the lepton should account for 90% of the energy, neglecting photons and any other leptons. We then seek out an opposite-sign same-flavor pair with $m_{ll} = [75, 105]$ GeV to serve as the Z

⁴ More specifically, we allow N_s for which the integrated probability for background to fluctuate to N_s events or larger is less than 3×10^{-7} .

candidate. We further seek out leptons which fail basic isolation but which are mutually isolated from hadronic taus. We then cluster the remaining particles in the event into $R = 0.4$ C/A jets and check them for either single hadronic tau tag or hadronic ditau tag as in section II. (For the single tag, we use an outer annulus radius of 0.4, in contrast to the CMS choice of 0.45.) We require events to have a good Z candidate and *exactly* two tau candidates (hadronic taus and/or leptons).⁵ The vector sum of the visible tau candidate activity should have $|\eta| < 1.5$, placing the taus in the region where the ECAL granularity is highest in the real detectors.

Next, we reconstruct the ditau invariant mass and the Z' mass. A clear obstacle here is the presence of *at least* two neutrinos in the final state from the tau decays. Nonetheless, we know that these neutrinos are very well-aligned with the visible tau decay products. Using this fact, and given the measured components of missing transverse energy (\cancel{E}_T), it is possible to determine the neutrino three-vector associated to each tau and to fully reconstruct the Higgs. This is a standard method for light Higgs searches in vector boson fusion. (Details can be found, for example, in [35–38].) Reconstruction of the Z' is even more straightforward, since all of the neutrinos are fairly collinear from the more global perspective of the entire event. This reconstruction is not very sensitive to the detailed energy sharing between neutrinos from the different taus.

The quality of the Higgs mass reconstruction will depend sensitively on the quality of the missing energy measurement, especially its direction. While we do not have the tools to model this accurately, we will make the case that this reconstruction should be possible. We start by defining the \cancel{E}_T vector to be the transverse vector that balances the sum of the reconstructed Z and the visible products of the two taus. We re-emphasize that these have been nominally smeared according to expectations at the LHC, and note that this construction of \cancel{E}_T somewhat pessimistically subjects us to fluctuations from ISR and other uncorrelated event activity.

The reconstructed Higgs mass is shown in Fig. 4, including a comparison to a reconstruc-

⁵ In the fairly rare cases with two sets of opposite-sign same-flavor lepton pairs in the Z mass window, we take the one that comes closest to the Z pole mass as the Z candidate, and the other two as the ditau candidate. This subjects us to some ZZ background which we do not account for, but could easily be removed with a double- Z veto.

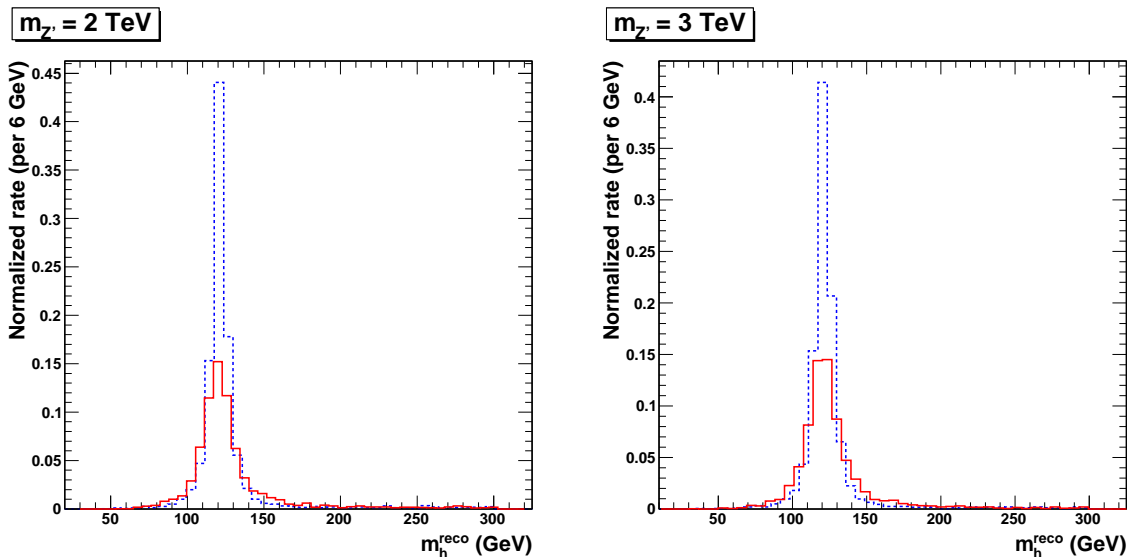


FIG. 4: Reconstructed Higgs invariant mass from 2 TeV (left) and 3 TeV (right) Z' decays. The solid red line shows Higgs reconstruction using our default modeling of the \cancel{E}_T measurement, and the dashed blue line shows Higgs reconstruction if perfect \cancel{E}_T is used.

tion using a perfect \cancel{E}_T measurement (tracing the true sum of unsmeared neutrinos).⁶ The signal has a well-defined peak in both cases. However, in going from a perfect \cancel{E}_T measurement to a smeared measurement, the peak broadens and develops a substantive high-mass tail. The tail in fact extends well beyond the range of the plots. It is fed mainly by cases where the direction of the \cancel{E}_T vector falls outside of the wedge defined by the two taus in the transverse plane. In such cases, one of the neutrinos is reconstructed *anti*-parallel to one of the taus, and the mass broadly overshoots.⁷ We take as a nominal Higgs mass window $m_h^{\text{reco}} = [100, 300]$ GeV, which encompasses the bulk of the peak while staying somewhat higher than the Z mass and recapturing some of the lost high-mass events. Reconstruction

⁶ The plot sums over all ditau decay channels, but the individual distributions are not very different. While this is perhaps the first time such an estimate has been made at high Higgs p_T , we point out for comparison the result for Higgs production from vector boson fusion in [38]. This achieves 10% mass resolution in both dileptonic and semileptonic decay modes.

⁷ It would likely be useful to consider more advanced reconstruction methods that correct for \cancel{E}_T direction mismeasurements, for example re-aligning \cancel{E}_T with the closest tau in pathological cases. We also emphasize that aside from this “out-of-wedge” case, we find that the mass reconstruction for nearly-aligned taus is quite stable. This is in contrast to the other extreme where the taus are moving in nearly opposite directions, and the sum of neutrino energies can become ambiguous.

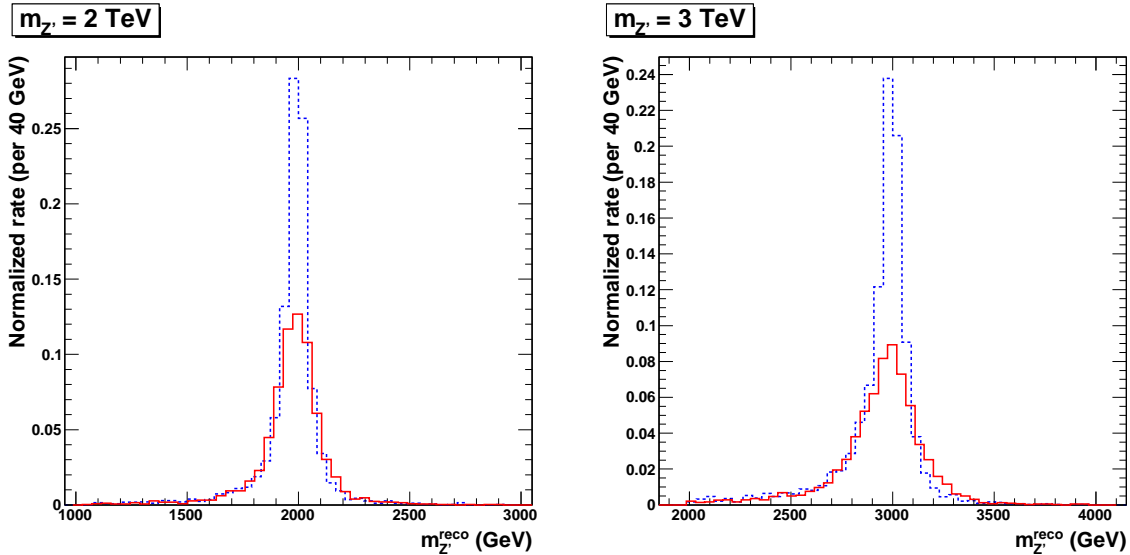


FIG. 5: Reconstructed Z' invariant mass for 2 TeV (left) and 3 TeV (right), with the Higgs mass window cut applied. Plots are normalized to the signal rate before the Higgs mass window cut. The solid red line uses our nominal \cancel{E}_T reconstruction, and the dashed blue line uses perfect \cancel{E}_T .

of the Z' after application of this window is shown in Fig. 5.

One of the major backgrounds for this analysis is Z +jets, with the jets faking a ditau pair.⁸ We use this background to obtain an estimate for ditau-tag fake rates from QCD jets. We restrict to events with a reconstructed leptonic Z and where the recoiling hard parton is in the central region $|\eta| < 1$. We run separate analyses for 2(3) TeV by taking $\sqrt{\hat{s}} = [1600, 2400]([2600, 3400])$ GeV and $\hat{p}_T = [800, 1200]([1300, 1700])$ GeV. The ditau tag rate is defined as the fraction of these events that pass our reconstruction. For the equivalent analyses with the signal, we simply start with all events that pass basic event reconstruction. We show the tag and mistag rates for the individual ditau decay modes in table I. The samples are heavily dominated by merged ditau configurations. We see that the signal tag rates are at the 50 – 70% level in the semileptonic and all-hadronic channels,

⁸ The diboson background $ZZ \rightarrow (l^+l^-)(\tau^+\tau^-)$ is small, and is ultimately brought down to a negligible level with our Higgs mass window. The irreducible Zh background is similarly negligible. There will also be a background where the Z is produced recoiling against a quark, which subsequently emits a W boson (W -strahlung) that decays leptonically or into a tau. The fraction of Z +jets events in which this occurs is sub-percent [9], before demanding that the associated jet is tau-tagged. This background ends up subleading to Z +jets with mistagged ditau.

	2 TeV			3 TeV		
	dilep	semilep	all-had	dilep	semilep	all-had
signal ditau tag	86%	50%	49%	98%	51%	70%
+ Higgs window	48%	30%	30%	47%	29%	41%
ditau mistag gZ (qZ)	- (-)	0.02%(0.03%)	0.02%(0.07%)	- (-)	0.03%(0.03%)	< 0.01%(0.07%)

TABLE I: *Efficiencies (in percent) for reconstruction of ditaus. For the signal, the efficiencies are determined channel-by-channel. For the background, which does not have physical taus, the efficiencies represent the percentage of all events passing into each channel.*

and the mistag rates are at the $10^{-4} - 10^{-3}$ level (corresponding roughly to the square of individual hadronic tau tag rates expected at lower energies [28]). Note that the background simulations run out of statistics after application of the Higgs mass window, corresponding to final efficiencies of less than about 10^{-5} . The final signal efficiencies after application of the mass window are about 30 – 40%.

We now turn to discovery reach. For our signal box, we use the Higgs mass window above, and define Z' mass windows of [1650, 2350] GeV and [2500, 3500] GeV. For 2 and 3 TeV resonances, we find negligible background rate, amounting to less than 1 event even after 300 fb^{-1} . (The final $Z' \rightarrow Zh$ reconstruction efficiency, after all cuts, is about 25 – 30% times the individual Z and Higgs BR s.) Discovery is therefore purely determined by our requirement that at least 10 events are observed. We combine all three ditau decay channels, and display the reach in table II. Resonances with $\sigma \times BR(Zh)$ of a few 10's of fb can be discovered with reasonable LHC luminosities, roughly independent of mass. These are fairly high cross sections for multi-TeV resonances. As anticipated, the search sensitivity is limited by the low branching fraction of $Zh \rightarrow (l^+l^-)(\tau^+\tau^-)$. However, this may still be an interesting mode to pursue for nonstandard Higgses or other scalars. Extending the search to lower Z' masses, where the cross section is expected to be larger, should also be possible.

B. Hadronic Z

As in the previous analysis, we start by looking for isolated leptons, and then try to find leptons mutually isolated from hadronic taus. The event should have no good dileptonic

	$M_{Z'} = 2 \text{ TeV}$	$M_{Z'} = 3 \text{ TeV}$
$\mathcal{L} = 100 \text{ fb}^{-1}$	77 fb	70 fb
$\mathcal{L} = 300 \text{ fb}^{-1}$	26 fb	23 fb

TABLE II: *Discovery reach for $Z' \rightarrow Zh \rightarrow (l^+l^-)(\tau^+\tau^-)$. The numbers denote the minimum discoverable $\sigma(q\bar{q} \rightarrow Z') \times BR(Z' \rightarrow Zh)$ at the given integrated luminosity. We assume $BR(h \rightarrow \tau^+\tau^-) = 0.068$, as for a standard Higgs at 120 GeV.*

Z candidate. The remaining particles are then clustered with $R = 0.4$ C/A in search for hadronic tau-jets and ditau-jets. Of the identified taus (or leptons), two should be on the same side of the event, and their vector sum (neglecting neutrinos) should again have $|\eta| < 1.5$.

Subsequently, we look for the hadronic Z using jet substructure methods. All particles not associated with taus are reclustered into $R = 1.4$ quasi-hemispheric fat-jets. The hardest of these should have $|\eta| < 1.5$, $p_T > 600$ GeV, and be greater than $\pi/2$ away from the taus in ϕ . This is declustered into two subjets as in [3]. The mass of the summed subjets is then required to be in the window [75, 105] GeV. The final “ Z -tag” rate is about 70%, with mistag rates at the 10% level or smaller. For more details on this procedure and its performance (studied for the nearly identical case of W -jets), see [17].

To perform the resonance search, we use the [100, 300] GeV mass window for the Higgs, as above, and for the 2(3) TeV Z' mass window we again use [1650, 2350]([2500, 3500]) GeV. The reconstructed Z' mass peaks are shown in Fig. 6.

The backgrounds for this analysis can be formidable, especially in the all-hadronic ditau mode, where we subject ourselves to dijet-induced fakes.⁹ However, even in this mode we find that we may achieve enough discriminating power, reducing the dijet contribution by a factor of $O(10^{-6})$ with the combination of cuts on the hadronic Z and ditau Higgs.

Another important background is Z +jets, where this time the Z decays to taus and the jets fake a hadronic Z . This is the only major background with large production cross section and a genuine ditau pair. We display the reconstructed ditau invariant mass for this

⁹ Triggering on all-hadronic signal events may not be guaranteed, even with a dedicated hadronic tau trigger. However, these events are extremely energetic, with leading jets usually well in excess of 500 GeV. The total QCD rate above this p_T is sub-Hertz at $10^{-34} \text{ cm}^{-1} \text{ s}^{-1}$ luminosity.

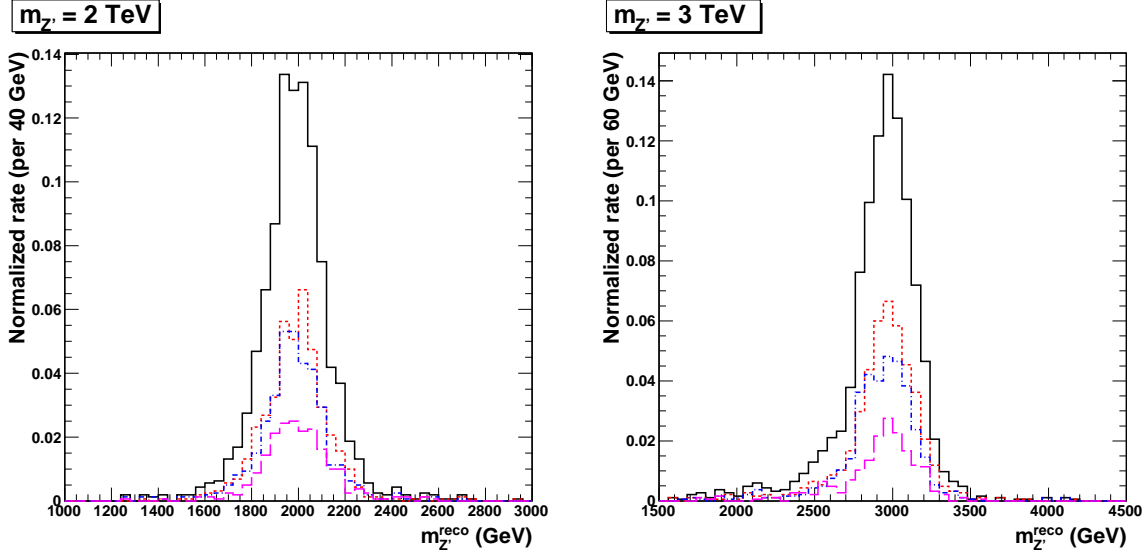


FIG. 6: Reconstructed Z' invariant mass in the hadronic Z channel (black solid), and broken down into contributions from dileptonic ditau (pink long-dashed), semileptonic ditau (blue dot-dashed), and all-hadronic ditau (red short-dashed).

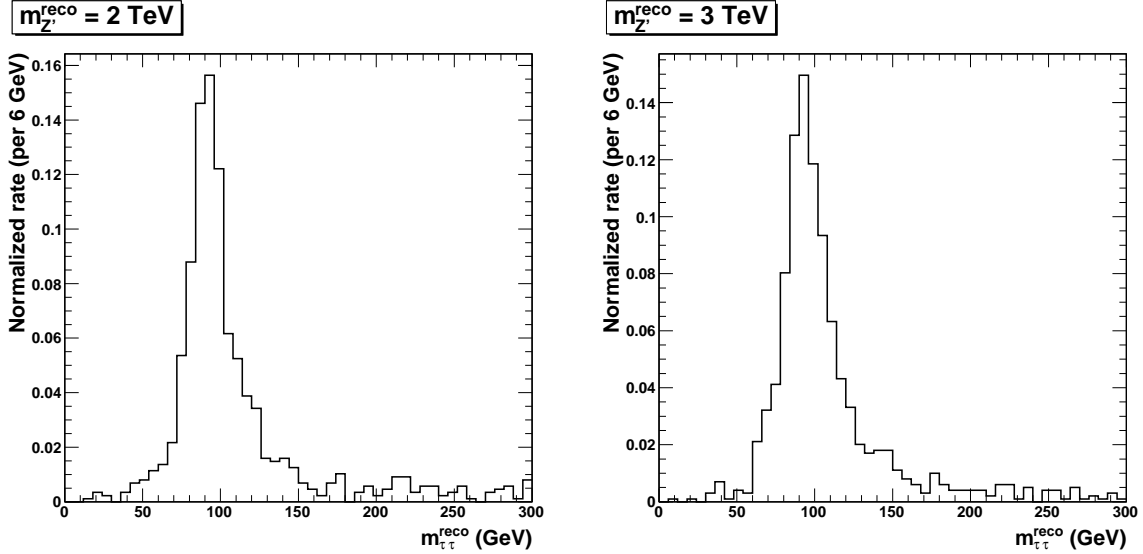


FIG. 7: Fully reconstructed ditau invariant mass from the $(Z \rightarrow \tau^+\tau^-) + \text{jets}$ background.

background in figure 7. Our Higgs mass window cut removes about 3/4 of the events.

We also consider $t\bar{t}$ and $W + \text{jets}$ (W -strahlung) backgrounds. The former can be dangerous, for example, when one of the tops decays hadronically and the other leptonically or

	2 TeV			3 TeV		
	dilep	semilep	all-had	dilep	semilep	all-had
Signal Eff.	25% (32%)	12% (16%)	16% (22%)	25% (36%)	14% (20%)	19% (29%)
$\sigma(\text{dijet QCD})$	0 ab (0 ab)	0 (39)	140 (1560)	0 (0)	0 (37)	7 (465)
$\sigma(Z+\text{jets})$	3 (13)	6 (42)	10 (52)	0 (1)	1 (2)	1 (2)
$\sigma(t\bar{t})$	0 (8)	2 (26)	6 (34)	0 (0)	1 (3)	0 (4)
$\sigma(W\text{-strahlung})$	0 (0)	10 (144)	0 (0)	0 (0)	0 (84)	0 (0)
$\sigma(\text{SM diboson})$	1 (1)	2 (3)	2 (3)	0 (0)	0 (0)	0 (0)

TABLE III: *Final signal efficiencies (channel-by-channel, without incorporating branching fractions) and background cross sections, in ab, after(before) the Higgs mass cut. Here, “0” indicates cross sections which we estimate to be less than 1 ab, or for which we have otherwise run out of statistics in our simulations (see text).*

into a tau. Such a $t\bar{t}$ event can closely mimic the signal when the b -jet on the hadronic side is soft (and thrown away by the hadronic Z -tagger), and the b -jet on the leptonic side is misidentified as a tau. Final-state W -strahlung can occur in any process that produces a hard final state quark, in this case simple $2 \rightarrow 2$ parton scattering. This kind of emission greatly enhances the chance that a generic dijet event passes our analysis cuts if the emitted W decays leptonically or into a tau. However, this competes with the fact that the emission is fairly rare.

Finally, we include Standard Model diboson backgrounds, including the irreducible Zh . These are all subleading.

We list the signal efficiencies and major background cross sections for the different ditau modes and Z' masses in table III. The backgrounds are all quite small, at a level much less than a fb, even for the all-hadronic mode. Indeed, the dominant backgrounds in reality will likely be instrumental in nature, but determining this will require a more advanced simulation or data-driven study.¹⁰

¹⁰ Accurate estimation of the fully QCD background is further complicated by the large statistics involved. Our final QCD samples (consisting of several 10's of millions of events) represent event weights of about 7 ab, or 0.7 for 100 fb^{-1} luminosity and 2 for 300 fb^{-1} . Consequently, a quoted cross section of 7 ab is more accurately $O(7 \text{ ab})$, and a quoted cross section of 0 ab is more accurately $\lesssim 7 \text{ ab}$. We infer from intermediate stages of the analysis that the dileptonic and semileptonic cross sections stay well below

	$M_{Z'} = 2 \text{ TeV}$	$M_{Z'} = 3 \text{ TeV}$	$M_{Z'} = 2 \text{ TeV} (\geq 1 \text{ lepton})$	$M_{Z'} = 3 \text{ TeV} (\geq 1 \text{ lepton})$
$\mathcal{L} = 30 \text{ fb}^{-1}$	66 fb	35 fb	76 fb	63 fb
$\mathcal{L} = 100 \text{ fb}^{-1}$	32 fb	12 fb	26 fb	21 fb
$\mathcal{L} = 300 \text{ fb}^{-1}$	17 fb	6 fb	13 fb	7 fb
S/B = 1	23 fb	4 fb	6 fb	1 fb

TABLE IV: *Discovery reach for $Z' \rightarrow Zh \rightarrow (q\bar{q})(\tau^+\tau^-)$. The numbers denote the minimum discoverable $\sigma(q\bar{q} \rightarrow Z') \times BR(Z' \rightarrow Zh)$ at the given integrated luminosity, or that which is required to match the background cross section after all cuts. We assume $BR(h \rightarrow \tau^+\tau^-) = 0.068$, as for a standard Higgs at 120 GeV. The last two columns use only the channels with at least one leptonic tau, omitting the all-hadronic channel.*

We list the discovery reach estimates in table IV, broken down into cases without/with the inclusion of the all-hadronic mode. While all-hadronic is likely the most experimentally challenging, we have seen that its backgrounds appear to be controllable, and that including it roughly doubles the statistics. However, in the event that instrumental backgrounds make this mode non-viable, the combined dileptonic and semileptonic analyses can still give good, even superior sensitivity, especially at very high integrated luminosity.

As expected, the overall gain in sensitivity is significant relative to the $(l^+l^-)(\tau^+\tau^-)$ channel. For example, discovery of a 3 TeV Z' with $\sigma \times BR(Zh)$ of about 6-7 fb is possible at 300 fb^{-1} luminosity (versus 23 fb with the leptonic Z). These can be compared to the results in [17], utilizing $h \rightarrow b\bar{b}$ in association with a leptonic or invisible Z , which respectively could probe down to about 2.5 fb and 1.5 fb. For 2 TeV resonances, the difference in sensitivities is more modest, with $(q\bar{q})(\tau^+\tau^-)$ (omitting all-hadronic ditau) only about 50% less sensitive than the non- b -tagged $(l^+l^-)(b\bar{b})$. In particular, the Z' states of the warped 5D models of [18, 23] may be simultaneously visible at 2 TeV, with production cross section of about 15 fb, in the three Zh modes $(l^+l^-)(b\bar{b})$, $(\nu\bar{\nu})(b\bar{b})$, and $(q\bar{q})(\tau^+\tau^-)$.

Finally, we note that all of our results in this subsection can also be trivially applied to $W' \rightarrow Wh \rightarrow (q\bar{q}')(\tau^+\tau^-)$, which has almost identical structure.

the all-hadronic, and take “0 ab” to be reasonable at our level of precision. However, for the 3 TeV all-hadronic cross section, which is represented by a single simulation event, we conservatively increase the number to 15 ab for the purpose of calculating discovery reach.

IV. CONCLUSIONS

When a light Higgs is generated in a multi-TeV process, it can subsequently decay into a pair of taus and manifest itself in the detector as a ditau-jet. While this decay channel is expected to be subdominant, it is also extremely distinctive. Both the Higgs and the tau are color-singlets, so any hadrons produced in the tau decays are highly beamed into one or two very small regions of the detector without any surrounding radiation produced in the final state. This unusual configuration should in principle become progressively easier to spot at higher energies. In this paper, we have explored the possibility of capitalizing on this fact in order to construct a powerful high- p_T Higgs tag, and to determine if $h \rightarrow \tau^+ \tau^-$ can have any role to play in the search for heavy new physics at the LHC.

In order to identify tau pairs with $\Delta R_{\tau\tau} \lesssim 0.4$ and one or two hadronic decays, we extend the usual tau algorithms to incorporate “mutual isolation,” where the inner activity cone of one tau is deleted from the isolation annulus of its partner. We have found that for ditaus at $p_T \sim \text{TeV}$ we can achieve tag and mistag rates that are roughly the square of what has been estimated by the LHC experimentalists for single taus at $p_T \sim 100 \text{ GeV}$. Specifically, we estimate that $O(50\%)$ tag rate and $O(10^{-4} - 10^{-3})$ mistag rates should be possible. Of course, in this study we were not able to faithfully model instrumental effects, which are likely very important. A more detailed detector study is warranted to verify the physics picture which we have laid out here.

We have also demonstrated that the ditau pair mass can be reconstructed, with the Higgs appearing as a broad but well-localized bump. This is perhaps counterintuitive, as it is known that Higgs mass reconstruction becomes badly ambiguous in the *anti*-collinear limit, due to a degeneracy in the apportioning of $\vec{\cancel{E}}_T$ between the two taus. However, taking the collinear limit and assuming that all of the neutrinos are moving in roughly the same direction as the visible tau products, this ambiguity is no longer so severe since we know the approximate sum of neutrino energies. Nevertheless, fluctuations in visible energy measurements can cause the missing energy vector to point outside of the 2D wedge in the transverse plane defined by the visible taus, naively leading to reconstructed neutrinos oriented toward the wrong side of the event. In this case, the reconstructed ditau mass can be highly overestimated. While this population of misreconstructed events does not cause major problems for us, it seems likely that even a slightly modified method could perform

much better.

We applied the ditau-jet tag and invariant mass reconstruction to the simple new physics process of a multi-TeV Z' decaying to Zh . (Some of our results also trivially extend to $W' \rightarrow Wh$.) The low-statistics channel $(l^+l^-)(\tau^+\tau^-)$ mainly served to calibrate our ditau techniques, but is largely background-free and has sensitivity to resonances with $\sigma \times BR(Zh)$ of a few 10's of fb after running the LHC for several hundred fb^{-1} . The channel $(q\bar{q})(\tau^+\tau^-)$ is more promising from the perspective of rate, but faces enormous backgrounds at the outset. Nonetheless, we have shown that the combination of ditau-tagging, Higgs mass reconstruction, and hadronic Z -tagging (as in [17]) can potentially reduce the background to a manageable level while preserving most of the signal. We estimate that this channel has sensitivity at the 10 fb level or lower, competing with the more conventional channels $(l^+l^-)(b\bar{b})$ and $(\nu\bar{\nu})(b\bar{b})$ [17], especially around 2 TeV.

The techniques which we discuss here can also be applied in a broader new physics context. In addition to resonant production of the Higgs, we might consider Higgses produced in SUSY cascades or in the decays of other pair-produced heavy particles. New scalars besides the Higgs are another possible target. For example, a light scalar ($m \ll 100$ GeV) decaying into tau pairs could be identified as a ditau-jet within events of much lower energies.

Acknowledgments

We are grateful to Kaustubh Agashe, Sarah Demers, Petar Maksimovic, Frank Paige, and Raman Sundrum for useful discussions. The authors are also grateful to the organizers of the workshop “BOOST2010” at Oxford. AK was partially supported by NSF grant PHY-0801323. MS was supported in part by the DoE under grant No. DE-FG-02-92ER40704. BT was supported by JHU grant No. 80020033 and by DoE grant No. DE-FG-02-91ER40676.

-
- [1] M. H. Seymour, *Searches for New Particles Using Cone and Cluster Jet Algorithms: A Comparative Study*, *Z. Phys.* **C62**, 127 (1994).
 - [2] J. M. Butterworth, B. E. Cox, and J. R. Forshaw, *WW Scattering at the CERN LHC*, *Phys. Rev.* **D65**, 096014 (2002), hep-ph/0201098.

- [3] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, *Jet Substructure as a New Higgs Search Channel at the LHC*, Phys.Rev.Lett. **100**, 242001 (2008), 0802.2470.
- [4] G. Brooijmans, *High p_T Hadronic Top Quark Identification. Part I: Jet Mass and Ysplitter* (2008), ATL-PHYS-CONF-2008-008.
- [5] J. Thaler and L.-T. Wang, *Strategies to Identify Boosted Tops*, JHEP **07**, 092 (2008), 0806.0023.
- [6] D. E. Kaplan, K. Rehermann, M. D. Schwartz, and B. Tweedie, *Top Tagging: A Method for Identifying Boosted Hadronically Decaying Top Quarks*, Phys. Rev. Lett. **101**, 142001 (2008), 0806.0848.
- [7] T. Plehn, G. P. Salam, and M. Spannowsky, *Fat Jets for a Light Higgs*, Phys. Rev. Lett. **104**, 111801 (2010), 0910.5472.
- [8] T. Plehn, M. Spannowsky, M. Takeuchi, and D. Zerwas, *Stop Reconstruction with Tagged Tops* (2010), 1006.2833.
- [9] K. Rehermann and B. Tweedie, *Efficient Identification of Boosted Semileptonic Top Quarks at the LHC* (2010), 1007.2221.
- [10] J. M. Butterworth, J. R. Ellis, A. R. Raklev, and G. P. Salam, *Discovering Baryon-Number Violating Neutralino Decays at the LHC*, Phys. Rev. Lett. **103**, 241803 (2009), 0906.0728.
- [11] G. D. Kribs, A. Martin, T. S. Roy, and M. Spannowsky, *Discovering the Higgs Boson in New Physics Events using Jet Substructure*, Phys. Rev. **D81**, 111501 (2010), 0912.4731.
- [12] L. G. Almeida et al., *Substructure of High- p_T Jets at the LHC*, Phys. Rev. **D79**, 074017 (2009), 0807.0234.
- [13] L. G. Almeida, S. J. Lee, G. Perez, G. Stermann, and I. Sung, *Template Overlap Method for Massive Jets*, Phys. Rev. **D82**, 054034 (2010), 1006.2035.
- [14] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, *Techniques for Improved Heavy Particle Searches with Jet Substructure*, Phys. Rev. **D80**, 051501 (2009), 0903.5081.
- [15] D. Krohn, J. Thaler, and L.-T. Wang, *Jet Trimming*, JHEP **02**, 084 (2010), 0912.1342.
- [16] J. Thaler and K. Van Tilburg, *Identifying Boosted Objects with N -subjettiness* (2010), 1011.2268.
- [17] A. Katz, M. Son, and B. Tweedie, *Jet Substructure and the Search for Neutral Spin-One Resonances in Electroweak Boson Channels* (2010), 1010.5253.
- [18] K. Agashe, A. Delgado, M. J. May, and R. Sundrum, *RS1, Custodial Isospin and Precision*

- Tests*, JHEP **08**, 050 (2003), hep-ph/0308036.
- [19] N. Arkani-Hamed, A. G. Cohen, and H. Georgi, *Electroweak Symmetry Breaking from Dimensional Deconstruction*, Phys. Lett. **B513**, 232 (2001), hep-ph/0105239.
- [20] M. Schmaltz and D. Tucker-Smith, *Little Higgs Review*, Ann. Rev. Nucl. Part. Sci. **55**, 229 (2005), hep-ph/0502182.
- [21] P. Langacker, *The Physics of Heavy Z' Gauge Bosons*, Rev. Mod. Phys. **81**, 1199 (2009), 0801.1345.
- [22] J. E. Garcia, M. Lechowski, E. Ros, and D. Rousseau (The Atlas), *Search for the Decays $Z_H \rightarrow Zh$ and $W_H \rightarrow Wh$ in the Little Higgs Model Assuming $m(h) = 120$ GeV* (2004), ATL-PHYS-2004-001.
- [23] K. Agashe et al., *LHC Signals for Warped Electroweak Neutral Gauge Bosons*, Phys. Rev. **D76**, 115015 (2007), 0709.0007.
- [24] J. Gallicchio and M. D. Schwartz, *Seeing in Color: Jet Superstructure*, Phys. Rev. Lett. **105**, 022001 (2010), 1001.5027.
- [25] A. Falkowski, D. Krohn, L.-T. Wang, J. Shelton, and A. Thalapillil, *Unburied Higgs* (2010), 1006.1650.
- [26] M. Heldmann and D. Covalli (The Atlas), *An improved tau-identification for the atlas experiment* (2006), ATL-PHYS-PUB-2006-008.
- [27] E. Richter-Was and T. Szymocha (The Atlas), *Hadronic Tau Identification with Track Based Approach: The $Z \rightarrow \tau\tau$, $W \rightarrow \tau\nu$ and Dijet Events from DC1 Data Samples* (2005), ATL-PHYS-PUB-2005-005.
- [28] The CMS Collaboration, *CMS Strategies for Tau Reconstruction and Identification Using Particle Flow Techniques* (2009), CMS PAS PFT-08-001.
- [29] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, *Better Jet Clustering Algorithms*, JHEP **08**, 001 (1997), hep-ph/9707323.
- [30] M. Wobisch and T. Wengler, *Hadronization Corrections to Jet Cross Sections in Deep-Inelastic Scattering* (1998), hep-ph/9907280.
- [31] J. Alwall et al., *MadGraph/MadEvent v4: The New Web Generation*, JHEP **09**, 028 (2007), 0706.2334.
- [32] T. Sjostrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **05**, 026 (2006), hep-ph/0603175.

- [33] M. Cacciari and G. P. Salam, *Dispelling the N^3 Myth for the k_t Jet-Finder*, Phys. Lett. **B641**, 57 (2006), hep-ph/0512210.
- [34] The CMS Collaboration, *CMS Technical Design Report* (2006).
- [35] R. K. Ellis, I. Hinchliffe, M. Soldate, and J. J. van der Bij, *Higgs Decay to $\tau^+\tau^-$: A Possible Signature of Intermediate Mass Higgs Bosons at the SSC*, Nucl. Phys. **B297**, 221 (1988).
- [36] D. L. Rainwater, D. Zeppenfeld, and K. Hagiwara, *Searching for $H \rightarrow \tau\tau$ in Weak Boson Fusion at the LHC*, Phys. Rev. **D59**, 014037 (1999), hep-ph/9808468.
- [37] T. Plehn, D. L. Rainwater, and D. Zeppenfeld, *A Method for Identifying $H \rightarrow \tau\tau \rightarrow e^{+-}\mu^{-+}$ Missing $p(T)$ at the CERN LHC*, Phys. Rev. **D61**, 093005 (2000), hep-ph/9911385.
- [38] G. Aad et al. (The ATLAS), *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics* (2009), 0901.0512.