

Higgs boson decays into narrow di-photon jets and their search strategies at the Large Hadron Collider

Benjamin Sheff, Noah Steinberg, James D. Wells

*Leinweber Center for Theoretical Physics
Physics Department, University of Michigan
Ann Arbor, MI 48109-1040 USA*

Abstract:

In many extensions of the Standard Model the Higgs boson can decay into two light scalars each of which then subsequently decay into two photons. The underlying event is $h \rightarrow 4\gamma$, but the kinematics from boosted light scalar decays combined with realistic detector resolutions may fail to register the events in straightforward categories and thus may be lost. In this article we investigate the phase space for highly boosted di-photon events from these exotic Higgs decays and discuss search strategies that aim to capture and label events in this difficult region. In the process we develop a new category, ξ -jets, which identifies with high selectivity highly collimated di-photon decay modes of the Higgs boson.

Contents

1	Introduction	2
2	Theory description	4
3	Photon ξ-jets	5
4	Benchmark model points	6
5	Experimental search strategies	6
5.1	Multi-photon final states	7
5.2	ξ -jet final states	7
5.3	ξ -candidates	8
5.4	Reconstructing ξ -jets and Higgs decays	10
6	Conclusion	14
7	Acknowledgments	14

1 Introduction

Nearly a decade after the discovery of the Higgs boson it remains to be decided whether the discovered particle interacts with other known elementary particles in precisely the way the Standard Model dictates[1, 2, 3, 4, 5, 6]. Deviations from SM expectations can arise by virtue of the Higgs boson being composite, part of a larger Higgs sector, coupled through its portal interactions to hidden sector states, or embedded in extra dimensions to name just a few examples. Alternatives remain viable because the SM Higgs boson couplings to other SM states are known only to at best 10% for some, and only to within $\mathcal{O}(1)$ factors for others, including muon, electrons, charm, and Higgs self interactions[7, 8]. The possibility of the Higgs boson decaying into final states that are not allowed by the SM is also not constrained well in many cases.

In this article we take up the case of the Higgs boson (h) decaying into other very light scalars (ϕ_1 and ϕ_2) where each subsequently decays into photons,

$$h \rightarrow \phi_1 \phi_2 \rightarrow (\gamma\gamma)(\gamma\gamma) \quad (\text{target observable}). \quad (1)$$

In several different limits this process has been studied already[9, 10]. In the case of $\phi_{1,2}$ both having mass above about 10 GeV one finds that the events register as unambiguous 4γ events

in the detector that can be searched for well. Within this regime, current studies limit this process to $B(h \rightarrow 4\gamma) \lesssim 3 \times 10^{-4}$ [11, 12].

On the other extreme, if $\phi_{1,2}$ both have mass less than a few hundred MeV, the photons from $\phi_i \rightarrow \gamma\gamma$ are so collimated coming from the highly boosted ϕ_i resultant from their parent Higgs decay, that each ϕ_i decay appears to go to a single photon. In that case, $h \rightarrow \phi_1\phi_2$ is simply combined with the standard $h \rightarrow \gamma\gamma$ analysis, and it becomes a statistical question to determine what overabundance of such a signal would be consistent with data. At 95% CL the answer to this question is that the branching fraction of non-SM contributions to $B(h \rightarrow \gamma\gamma)$ cannot exceed 2.2×10^{-4} [13]. Such light scalars may also be disentangled from the SM $h \rightarrow \gamma\gamma$ process with sophisticated substructure techniques [14, 15].

Combining both extremes leads to an apparent detection $h \rightarrow 3\gamma$. This arises when one of the ϕ_i has mass less than a few hundred MeV and the other more than about 10 GeV. This process is forbidden in the SM, and the branching ratio is currently limited to $B(h \rightarrow 3\gamma) \lesssim \times 10^{-3}$, as can be gleaned from [16].

In between these two extremes, from the point of view of observables, is a murky region where the mass of one or both ϕ_i states is between ~ 0.1 GeV and 10 GeV. In that case, the two photons coming out of the ϕ_i decays are not highly collimated nor are they cleanly separated. Roughly speaking, the ATLAS and CMS detectors see something distinct from a standard photon but that also does not register as two photons when the photon separation is between $0.04 < \Delta R < 0.4$ [17, 18]. It is this difficult middle ground region that we wish to address in this letter.

It should be stated that extending the scalar sector of the SM by one (or multiple) singlets is a mature and well studied subfield [19, 20, 21]. Much of the parameter space for exotic heavy and light scalars (relative to the Higgs boson mass) is well constrained by direct searches and by precision electroweak measurements [22]. Our simplified model highlights a region of parameter space in a class of singlet extended models that has been less explored by previous studies.

The value in exploring such a regime lies in its ability to utilize the available experimental power from the LHC to investigate one of the most interesting loose ends in the Standard Model. Many models exist coupling new light scalars to the Standard Model in ways that are highly susceptible to the search strategy we advocate here [23]. The nature of the Higgs boson makes such couplings to new physics generic and apparent in a broad swath of theory parameter space. Furthermore, the rough knowledge we have of the Higgs boson to date deserves significant tightening in every reasonable direction. Our goal here is to consider this particular case in detail, highlight the experimental challenges for discovery, proffer some suggestions, and suggest a benchmark theory with points that may be useful for serious further study by experimental groups within the ATLAS and CMS collaborations.

2 Theory description

The phenomenon we are after is $h \rightarrow \phi_1 \phi_2$ with subsequent decay of $\phi_i \rightarrow \gamma\gamma$. Such decays arise generically in a broad class of BSM theories, many of which give rise to additional exotic phenomena. Most commonly these are other, similar gauge interactions, such as $Z \rightarrow \phi\gamma$, but the possibilities are wide and varied. Many BSM theories of this type are not yet constrained by experiment and have their most accessible phenomenon as $h \rightarrow \phi_1 \phi_2 \rightarrow 4\gamma$, if there are dedicated searches for it. Our focus lies in this last type of theory.

To devise an experimental strategy and analysis to discover this class of targeted theories, we must begin by constructing a representative theory within the class and finding ways to find evidence for it. Ideally the representative theory should be maximally simple without losing the key features under consideration for our exotic Higgs decays. In this case, there is such a simple theory, and its lagrangian is

$$\begin{aligned} \mathcal{L} = & \mathcal{L}_{\text{SM}} + \frac{1}{2}(\partial_\mu \phi_1)(\partial^\mu \phi_1) + \frac{1}{2}(\partial_\mu \phi_2)(\partial^\mu \phi_2) - \frac{1}{2}m_1^2 \phi_1^2 - \frac{1}{2}m_2^2 \phi_2^2 \\ & + \lambda_\phi |H|^2 \phi_1 \phi_2 + \frac{1}{\Lambda_1} \phi_1 F_{\mu\nu} F^{\mu\nu} + \frac{1}{\Lambda_2} \phi_2 F_{\mu\nu} F^{\mu\nu} \quad (\text{representative theory}) \end{aligned} \quad (2)$$

where $F^{\mu\nu}$ is the photon field strength tensor. Of course, one could write down non-trivial $|H|^2 \phi_1^2$ and $|H|^2 \phi_2^2$ terms among others, but that would add complexity without contributing significantly to the final phenomenology. One might also object that $\phi_i F_{\mu\nu} F^{\mu\nu}$ should be traded in for gauge-invariant couplings of ϕ_i to hypercharge field strength tensor $\phi_i B_{\mu\nu} B^{\mu\nu}$ and $SU(2)$ field strength tensor $\phi_i W_{\mu\nu}^a W^{a,\mu\nu}$. That would be fine, except that upon diagonalizing these interactions to those of the mass eigenstates one finds nevertheless $\phi_i F^2$ terms, which will completely dominate in the decays of ϕ_i over $\phi_i Z^{\mu\nu} F_{\mu\nu}$ and $\phi_i Z^2$ terms due to the Z boson being much heavier than the ϕ_i that we will consider below.¹ The $\phi_i Z F$ interaction can give rise to $Z \rightarrow \phi_i \gamma$ decays, constrained by searches at the Tevatron and the LHC[24, 11], but as the scale of Λ_i becomes higher, this constraint goes away while $B(\phi_i \rightarrow \gamma\gamma)$ remains 100%². For that reason we drop these extra consideration and extraneous interactions from the theory description and retain only the lagrangian of Eq. 2.

From the point of view of devising experimental search strategies to find evidence for the Higgs boson decaying into a single light scalar, say σ such that $h \rightarrow \sigma\sigma \rightarrow 4\gamma$, the benchmark theory above is adequate. It merely corresponds to the case of $m_1 = m_2$. That is not to say the two theories are exactly the same, only that the subsequent search strategies are the same. That is why we propose to work with only one theory – the representative theory of Eq. 2 – which we believe to form a basis upon which benchmark points can be established and strategies devised.

¹Barring any tuned cancellations, typical branching ratios to $\gamma\gamma$ are $10^7(10^{15})$ times larger than the branching ratios to $\gamma Z^*(Z^* Z^*)$.

²As the Λ_i increase, so does the decay length of the scalar. We have checked that the scalar decay length can be under 1 mm even for large (PeV) values of Λ_i which evade the $Z \rightarrow \phi_i \gamma$ constraint.

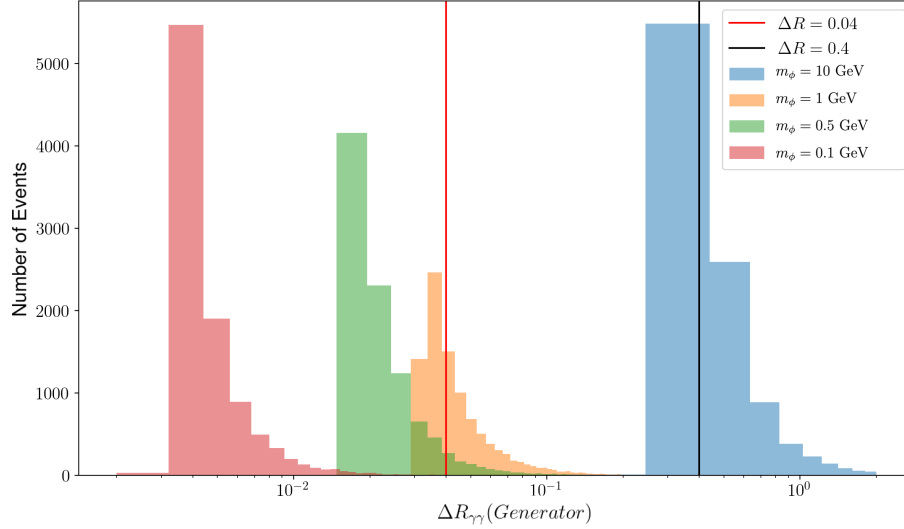


Figure 1: ΔR separation between photon pairs from ϕ_i decays, sampled over 10,000 events at varying masses of the BSM scalars.

3 Photon ξ -jets

As we mentioned in the introduction, the target observable of Eq. 1 implies photon separation from ϕ_i decays that is sensitive to the ϕ_i masses. This separation is given by

$$\Delta R_{\gamma\gamma} \sim 2m_{\phi_i}/E_{\phi_i} \sim 4m_{\phi_i}/m_h. \quad (3)$$

This is illustrated in fig. 1, which shows that $m_\phi = 10$ GeV gives well separated photons ($\Delta R > 0.4$) and $m_\phi = 0.1$ GeV gives very collimated photons ($\Delta R < 0.04$), and mass of 1 GeV gives intermediate separation. Recall that $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$, and $\Delta\phi$ is the azimuthal angle separation and $\Delta\eta$ is the pseudo-rapidity separation of the two photons in $\phi_i \rightarrow \gamma\gamma$ decay.

We turn to the ambiguous case in which the ϕ_i masses fall within the “intermediate mass” range of $1 \text{ GeV} < m_{\phi_i} < 10 \text{ GeV}$. Within the LHC environment, the production of Higgs bosons and their subsequent decay into such scalars yields photon pairs separated by

$$0.04 < \Delta R_{\gamma\gamma} < 0.4 \quad (\text{intermediate separation}). \quad (4)$$

It is well known that photon pairs that fall within the intermediate separation range of Eq. 4 are extremely difficult to separate or identify. We will speak much more on that below, but here we wish to pay respect to that difficulty by giving it a name. We call two photons that are within the range specified by Eq. 4 a “ ξ -jet”. The ξ -jet is a purely theoretical object, and it is defined by underlying “truth data” and not with respect to any detector performance.

If a photon has another photon within the intermediate separation annulus of Eq. 4, and nothing else is within the outer ring of that annulus, then it ceases to be a photon and the two together form a ξ -jet. Such a concept can be generalized to more than two photons but it is of not much importance here to do that. We also specify as a theoretical object that a photon is defined to be either a single photon or two photons within $\Delta R < 0.04$ of each other.

With these theory definitions of photon and ξ -jet, our target observable is broken into several distinct and non-overlapping final states, depending on the masses of the ϕ_i intermediate states in the decay chain:

$$h \rightarrow \phi_1 \phi_2 \rightarrow 4\gamma \Rightarrow 4\gamma, 2\gamma, 3\gamma, \gamma\xi, \gamma\gamma\xi, 2\xi \quad (\text{observable partitions}) \quad (5)$$

The first three of these observables we have already discussed. The remaining observables have not been fully explored in the literature, and we wish to consider them in more detail below.

4 Benchmark model points

We are interested in exploring three observables: $\gamma\xi$, $\gamma\gamma\xi$, and 2ξ . To do so we need benchmark points that give rise to each of these types of observables. They can be obtained rather straightforwardly from our representative theory of Eq. 2 where the masses of ϕ_1 and ϕ_2 are chosen to be various permutations of the masses 0.1 GeV, 1 GeV and 10 GeV. In particular, $m_\phi = 0.1$ GeV generally always gives $\phi \rightarrow \gamma$ decays, $m_\phi = 1$ GeV typically gives $\phi \rightarrow \xi$ decays, and $m_\phi = 10$ GeV generally gives $\phi \rightarrow \gamma\gamma$ decays according to our definitions in the previous section. These are so far entirely defined theoretically. In the next section we will pursue more carefully how a theoretical ξ -jet registers in an experimental analysis.

From these considerations we can construct the following four benchmark points A, B, C, and D specified in Table 1. Fig. 2 shows the relative fraction of each observable for each benchmark point. The dominant and subdominant modes of decay for each benchmark point are listed in Table 1 and can be gleaned from the fraction data given in Fig. 2. Table 1 shows that several combinations of light scalar masses give interesting decay signatures involving combinations of ξ -jets and photons which (to the authors' best knowledge) are not being searched for in current LHC analyses.

5 Experimental search strategies

So far our discussion has been mainly theoretical. We have identified a rare Higgs decay whose cascade we claim may be difficult to detect by experiment. In this section, we discuss how our theoretical objects translate into experimental manifestations. We have suggested that some mass ranges of ϕ_i are problematic for experiment. We will discuss some details on why they are challenging and some strategies by which to possibly overcome those challenges.

Point	m_1 (GeV)	m_2 (GeV)	Dominant mode	Subdominant mode
A	1	10	$\gamma\gamma\xi$	$\gamma\xi \simeq 2\gamma \simeq 3\gamma$
B	0.1	1	$\gamma\xi$	2γ
C	1	1	2ξ	$\gamma\xi \simeq 2\gamma$
D	0.1	10	3γ	$\gamma\xi$

Table 1: Benchmark points for $h \rightarrow \phi_1\phi_2 \rightarrow 4\gamma$ which then partition into various theory-object observables (modes) according to our definitions of ξ (photon pairs with $0.04 < \Delta R < 0.4$) and γ (an isolated photon with $\Delta R > 0.4$ or two photons within $\Delta R < 0.04$).

5.1 Multi-photon final states

Isolated photons or extremely highly collimated photons both get identified simply as photons, and analysis based on those standard objects (photons) proceed without much subtlety regarding how to process the data into well-defined final states of 2γ , 3γ and 4γ .

5.2 ξ -jet final states

Some of the final states from the decays of Eq. 5 yield ξ -jets. Underneath, a ξ -jet is merely two photons with intermediate ΔR separation (see Eq. 4). But a key question is, how does a ξ -jet, defined as a theoretical object, get processed into various experimental categories? A perfect detector would register it as merely two photons, a bad detector as a single photon or nothing, and a realistic good detector, such as ATLAS or CMS, registers it as something altogether different within several possible categories of varying sensitivity and selectivity³.

To address this question of how a ξ registers in a detector it is useful to describe the various categories into which a single photon can fall. As an example we take the standard categories which ATLAS uses for photon identification. There are eight possible standard categories, six are the permutations among three isolation possibilities (non-isolated, loose isolation, and tight isolation) and two ID possibilities (loose ID and tight ID). The other two categories are jet and “lost.” Jet is the standard QCD jet from fragmentation of quarks or gluons, and “lost” refers to the possibility that the data does not conform to any other category and is not registered in any higher abstracted category except for mere energy depositions in the detector.

A ξ -jet will register with some probability into one or more of the standard photon categories. The probability to do so depends on the underlying event kinematics. Under typical assumptions, the ξ -jet will often register as “lost” due to the inability to resolve the two photons yet the event covers more than one cell in the electromagnetic calorimeter which a single photon would not do. As no category becomes applicable, it has no option but to be relegated

³By perfect (non-existent) “sensitivity” we mean a category test that passes with 100% (0%) rate if the underlying event is a ξ -jet, and by perfect (non-existent) “selectivity” we mean a category test that passes with 0% (100%) rate if the underlying event is not a ξ -jet. Good sensitivity means low false negative rate, and good selectivity means low false positive rate.

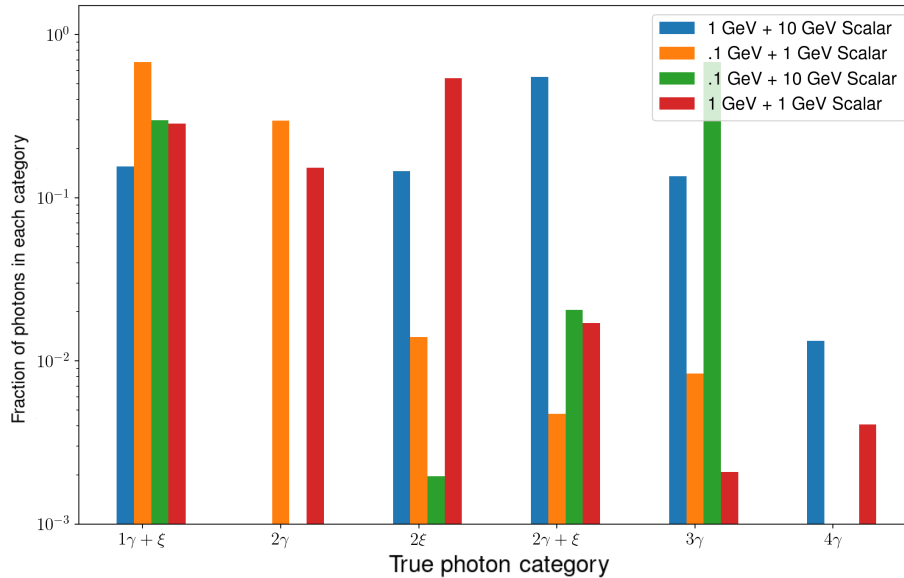


Figure 2: Branching fraction into each final state theory observable for the benchmark points A (blue), B (orange), C (red) and D (green) given in Table 1.

to “lost.”

The implication of a ξ -jet arising from a Higgs decay being categorized as “lost” is that an analysis that requires reconstructing the invariant mass of the Higgs boson from well-defined decay products can no longer register the events. It is therefore necessary to build a ninth category “ ξ -candidate” under which ξ events can fall. ξ -candidates must be defined entirely through detector response, with the goal of producing high sensitivity to underlying ξ -jets with reasonably good selectivity (i.e., mostly only ξ -jets register as ξ -candidates).

5.3 ξ -candidates

A detailed definition of the “ ξ -candidates” category satisfying the demands stated above is best constructed by a team of experimental experts within the ATLAS and CMS collaborations deeply familiar with their detectors. However, it is likely that such a definition meeting the demands of sensitivity and selectivity will have several key characteristics which we would like to discuss here. We will then make illustrative estimates of the utility of a ξ -candidates definitions based on these characteristics.

We make use of MadGraph_aMC@NLO[25] simulations to produce our signal events at leading order with the lagrangian of Eq. 2, which are then hadronized via PYTHIA8[26, 27]. For our detector studies we utilize Delphes[28] fast detector simulation framework with the default CMS card and FastJet[29] for jet clustering algorithms.

Variable	Definition	Cut	Reasoning
$\log\theta_J$	hadronic energy fraction	< -0.8	exclude QCD and τ
N_T	Number of tracks	$= 0$	excludes single converted photons and jet activity
τ_2/τ_1	Ratio of 2- to 1-subjettiness	< 0.3	Selects events with 2 subjets

Table 2: ξ definition meant to capture underlying events with $0.025 < \Delta R < 0.25$. These objects are defined as a cone of radius $\Delta R = 0.25$ about a central cluster in the EM calorimeter, centered on the highest energy pixel. Unless otherwise stated, the region is within the ξ region. Here θ_J is the hadronic energy fraction.

To begin one must have a cluster, established by standard techniques. One useful criteria to impose on the pre- ξ -candidate cluster is a strong isolation requirement against QCD activity within a small cone around the ξ -candidate system, reducing QCD backgrounds from decaying pions. Additional criteria for the definition must also appeal to the stoutness of the photon jet — there are two photons separated enough to not look like one photon and that separation shows up as a larger-than-normal spatial spread among cells within the electromagnetic detector. Furthermore, vetoing on charged tracks eliminates electron-induced showers. Finally, recently established jet n-subjettiness algorithms[30] can be employed to select clusters that have discernible sub-jet structure compatible with 2 collimated photons. Refs.[14, 31] go into detail on the ability to use these and other, similar variables to separate ξ -candidates (called photon-jets in these papers) from photons and QCD jets, but all of these considerations will be in play in the definitions below.

Our ξ -jet theory definition was for underlying two-photon clusters with ΔR separation in the range of 0.04 to 0.4. In addition, within the range of $0.025 < \Delta R < 0.04$ there is a possibility of using electromagnetic shape variables to discern that the underlying event was likely not a single photon, but certainly not clear enough to indicate the possibility of two photons. Nevertheless, our ξ -candidate list of criteria will be applicable for two-photon jets separations down to about $\Delta R \gtrsim 0.025$ and up to about $\Delta R \lesssim 0.25$. We will not discuss the range $0.25 \lesssim \Delta R \lesssim 0.4$ here, because our understanding is that more traditional photon identification tools may be applicable to separate the photons just well enough to help discern signal from photon backgrounds.

Let us now turn to a more precise definition of ξ -candidates (underlying two-photon separation $0.025 \lesssim \Delta R \lesssim 0.25$). This regime targets events that have two photons in sufficiently close proximity that their cores overlap, thereby interfering with one another's identification procedure. This should appear as a cluster of energy in the EM calorimeter, with no tracks or corresponding energy in the hadronic calorimeter, and high 2-subjettiness. We provide an example definition of ξ -candidate criteria in Table 2. Below in Fig. 3 we also show distributions of signal and background for QCD jets and ξ -jets. These distributions reproduce those of [31, 14] and show that ξ -jets can be separated from QCD backgrounds with high efficiency.

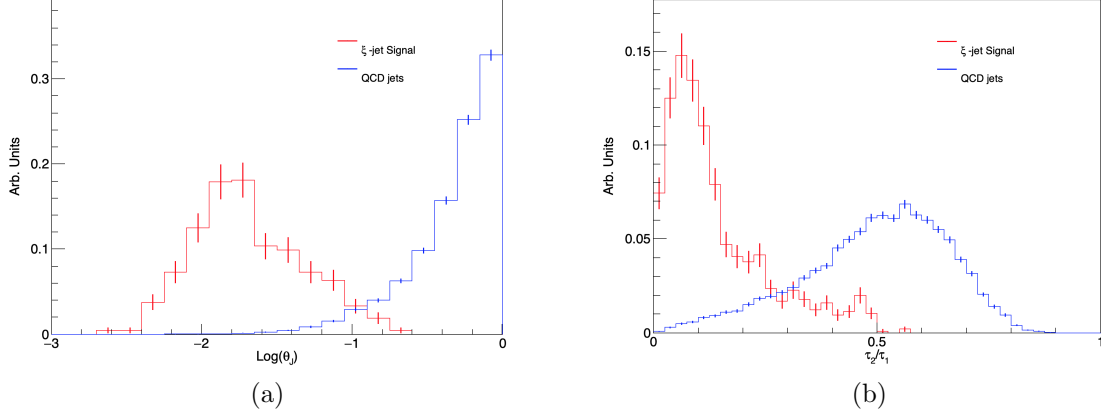


Figure 3: Subset of kinematic variables useful for discriminating of ξ -jets (green) and QCD jets (blue) which are a major background. Here θ_J is the hadronic energy fraction for a jet, and τ_2/τ_1 is the ratio of 2-jettiness to 1-jettiness which is useful for picking out events with 2 subjets.

5.4 Reconstructing ξ -jets and Higgs decays

Now that we have precise definitions of photons and ξ -candidates we can ask how well the Higgs boson signal can be reconstructed, especially in the case of its decay into one or more ξ -jets. Fig. 4 shows the analysis flow of our reconstruction of ξ -jets using Delphes fast detector simulation. Additional photons not covered by that flow, as well as electrons, muons, jets, etc. are identified and labeled by other analysis flows.

First, one must reconstruct the ξ -jets which we attempt to do by following a strategy similar to Ref. [14]. The method is as follows. First energy flow (eflow) objects[32] (composed of deposits in calorimeter cells) are clustered into jets using the anti-kt algorithm with $R = 0.25$. Then we re-cluster those energy deposits that were found in each jet using the kt algorithm, which determines a recombination tree for the jets. This tree specifies the subjets at each level of recombination N from $N = 1$ (the full jet) to $N =$ the number of constituent eflow objects in the jet (no recombination). From here we can compute the N -subjettiness variable for the jet for each N . This variable becomes small when the parameter N is large enough to describe all of the relevant substructure of the jet. It is defined to be

$$\tau_N = \frac{\sum_k p_{T_k} \times \min[\Delta R_{1,k}, \Delta R_{2,k}, \dots, \Delta R_{N,k}]}{\sum_k p_{T_k} \times R}, \quad (6)$$

where k runs over all the constituents of the jet, p_{T_k} is the transverse momentum for the k -th constituent, and R is the characteristic jet radius used in the original jet clustering algorithm.

After jet clustering is completed we then check if a reconstructed ξ -candidate already contains a reconstructed photon. Reconstructed photons are composed of eflow objects originating from the ECAL which must pass isolation requirements (cuts on electromagnetic and

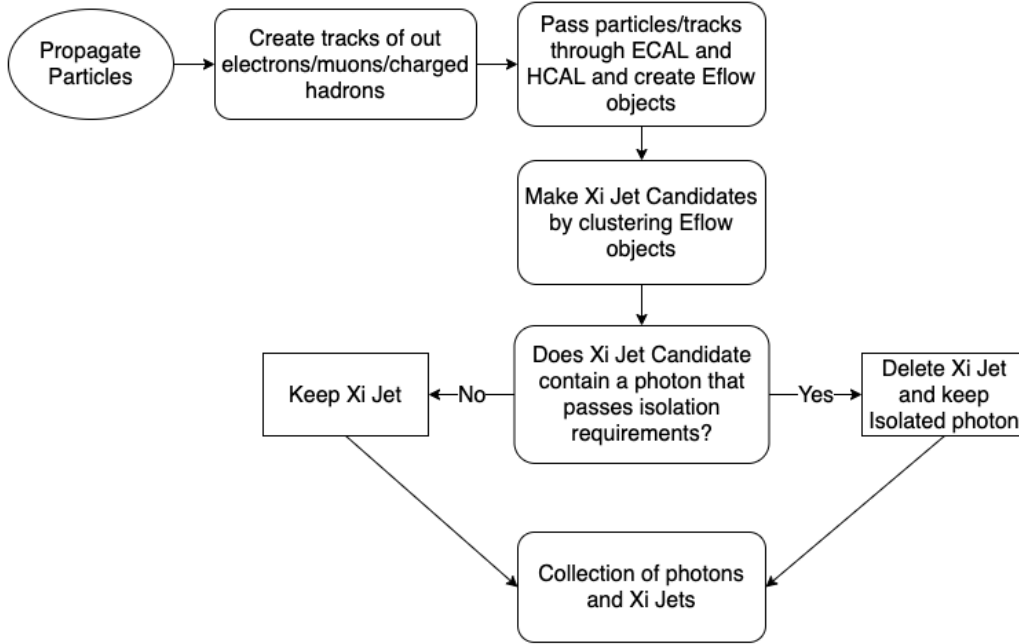


Figure 4: the analysis flow of our reconstruction of ξ -jets using Delphes fast detector simulation. Additional photons not covered by that flow, as well as electrons, muons, jets, etc. are identified and labeled by other analysis flows.

hadronic activity within a cone around the photon). If a ξ -candidate contains an already reconstructed, isolated photon then this ξ -candidate is deleted.

Before applying additional cuts, we would like to characterize the efficiency at which we reconstruct ξ -jets. To do this we utilize Delphes GenJet objects. GenJets are jets that are clustered, not with calorimeter cells or towers or eflow objects, but with the actual generator level particles. By utilizing GenJets we can define “generated ξ -jets” and see at what rate we correctly reconstruct these.

GenJets are clustered with the same strategy as above, first with the anti-kt algorithm with $R = 0.25$, and then reclustered with the kt algorithm. A GenJet is selected as a generated ξ -jet if it has: 1) At most two photons with $p_T > 0.5$ GeV, 2) no non-photons with $p_T > 0.5$ GeV. Since our theoretical ξ -jets were defined as pairs of photons with ΔR between 0.04 and 0.4, we throw out ξ -jets with $\Delta R < 0.025$ as these will most likely be reconstructed as one photon.

Once a generated ξ -jet is identified, we loop over all reconstructed ξ -candidates and attempt to find a match. Matching is done by comparing the ΔR between the momentum of the generated and reconstructed jets. If $\Delta R_{gen/reco} < 0.05$ we consider this jet as matched. We also require that the reconstructed ξ -jets pass a cut on the required hadronic energy fraction. This cut is that $\log(E_{had}/E_{jet}) < -0.8$. Below in Fig. 5 we show $\Delta R_{gen/reco}$, which shows the level of matching between generated and reconstructed ξ -jets. It also serves as a check that this is independent of our model parameters.

Now we would like to understand how often we can reconstruct the Higgs mass using our reconstructed photons and ξ -candidates. To simplify matters we will choose $m_{\phi_1} = m_{\phi_2}$, which is equivalent to having only one light scalar in addition to the observed h_{125} . We scan over light scalar masses from 100 MeV to 14 GeV. This range ensures we see a smooth transition between photon dominated decays and ξ -jet dominated decays. The following discussion can be generalized by choosing different masses for the light scalars. After reconstruction, we first collect all of our reconstructed objects, which for now are photons and ξ -candidates. We only require our reconstructed ξ -candidates to pass our hadronic energy fraction cut, otherwise no cuts (besides minimum p_T cuts which are used for clustering). We then form all the possible subsets of this collection, which have between 1 and 4 objects (as at most the Higgs decayed into 4 separable photons). If one combination of ξ -candidates and photons yields an invariant mass within a 3 GeV window around 125 GeV ($122 \text{ GeV} < M_{inv} < 128 \text{ GeV}$) then we consider this a match. Virtually no events contain multiple combinations of photons and ξ -jets which satisfy this requirement. We split each match into the following categories based on what number and type of objects make up the matching set.

1. Photons only: Matches with 2, 3, or 4 photons
2. Photons + ξ -jets: Matches with 1 photon + 1 ξ , 2 photons + 1 ξ
3. ξ -jets only: Events with 2 and only 2 ξ -jets
4. ξ -jets inclusive: Includes the Photons + ξ -jets category as well as the ξ -jets only category

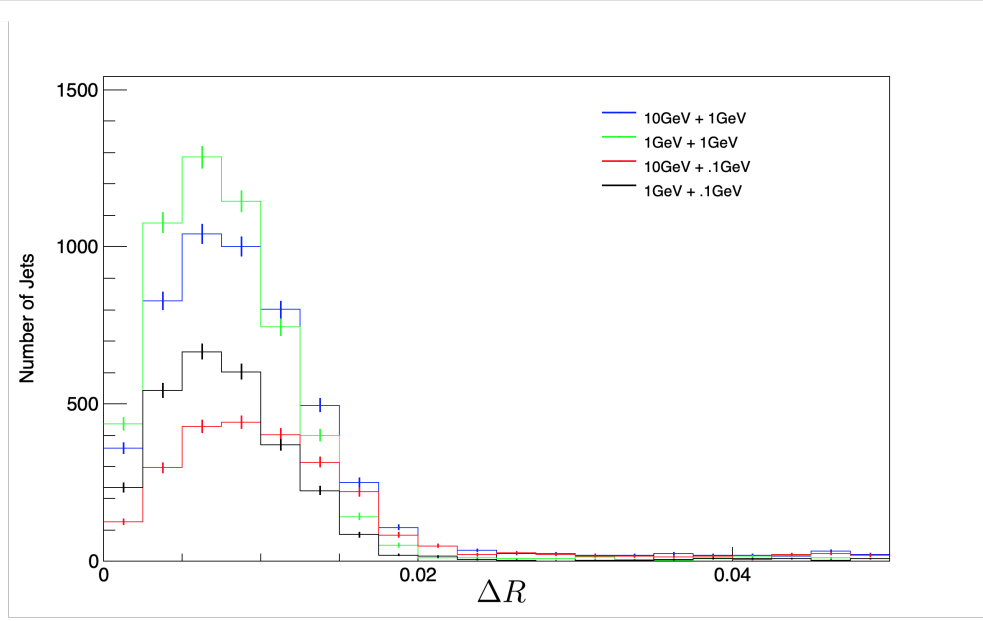


Figure 5: ΔR between reconstructed and generated ξ -jets. Distribution is independent of our model parameters showing good matching between the two.

5. Other combinations: Any combination not included above
6. All: Any match in the accepted mass range

Fig. 6 then shows the efficiency of reconstructing the Higgs mass as a function of the light scalar mass. Several key observations can be made here. At very low scalar masses, photons only makes up the dominant signal channel as the pairs of photons from ϕ decay are extremely collimated. From 100-300 MeV the signal from photons + ξ jets becomes the most efficient channel as one of the pairs of photons is collimated enough to form a ξ -jet. Immediately above 300 MeV the signal from pairs of ξ -jets (ξ -jets only) becomes an order of magnitude more efficient than the photon only channel and remains so until 6 GeV. Overall, searches including ξ -jets are more than an order of magnitude more efficient at reconstructing the Higgs from masses between 100 MeV and 10 GeV.

Fig. 6 shows that searches including ξ -jets would be invaluable if a light scalar connected to the gauge and Higgs sector as in Eq. 2 exists in nature. We would like to stress that even though our analysis and definitions are quite simple, our results should be robust even after the introduction of more strict experimental search strategies and analysis cuts. It is interesting to compute how many such ξ -jet events one could expect for a given luminosity at the LHC. This is of course a function of the ϕ mass and the efficiency for reconstructing the h_{125} . To give an estimate, we can take the $m_\phi = 2$ GeV point as an example. This has an

efficiency for reconstruction of about 50%. If we take $\text{Br}(h \rightarrow \phi\phi) = 10^{-4}$ and an integrated luminosity of 300 fb^{-1} , this leaves us with about 7500 reconstructed events.

While a comprehensive study of standard model backgrounds is necessary for an experimental search, we can still make qualitative statements about discriminating ξ -jets from objects which fake ξ -jets. The largest backgrounds will be from jj , and γj production where a QCD jet fakes a ξ -jet. References [31, 14] use a Boosted Decision Trees based on energy and substructure variables to discriminate between QCD jets and photon jets. They quote a fake rate from QCD jets of $10^{-4} - 10^{-5}$, though this fake rate is dependent on the rate at which one accidentally rejects ξ -jets. Additionally, the requirement that the invariant mass of the two ξ -candidates needs to fall within a 3 GeV window of the h_{125} mass lowers the background as well, as the rate for QCD jet production tends to fall at high invariant masses. Combined, these factors should allow for a bump hunt search for ξ -jets with high sensitivity.

Lastly we would like to comment on our results in comparison to other results in the literature concerning similar rare Higgs decays to scalars. It is our understanding that these articles are principally concerned with only the separation of photon jets and non photon jets, e.g. analyzing purity and efficiency for photon jets vs. single photons and other QCD related jets. In contrast, we have tried to paint a picture which quantifies how such rare Higgs decays to ϕ_i can be missed if only traditional techniques involving photons are used. In providing a clear theoretical and experimental definition of a photon and a ξ -jet we have been able to provide concrete branching fractions to states involving combinations of ξ -jets and photons only in terms of observables, that which a detector actually measures. The extremely high ratio of events involving ξ -jets to events involving only photons serves as a testament to the need for such an analysis. We hope that this can guide future experimental inquiries into models such as 2.

6 Conclusion

The discovery of the Higgs boson has lent strong support to the Standard Model, but also has allowed us to search for new avenues along which to extend it. In this work we have investigated exotic decays of the 125 GeV Higgs boson into light scalars which as of yet may be missed via current analysis techniques. We have discussed, first theoretically and then experimentally, a new object dubbed a ξ -jet which could play a pivotal role in the discovery of any light scalars minimally coupled to the standard model Higgs and to photons as in Eq. 2. If experimentalists are able to identify and reconstruct ξ -jets these new objects could be strong evidence for an extended Higgs sector and Beyond the Standard Model physics.

7 Acknowledgments

We thank Advanced Research Computing at the University of Michigan, Ann Arbor for their computational resources, as well as D. Amidei, C. Hayes, R. Hyneman and T. Schwarz for

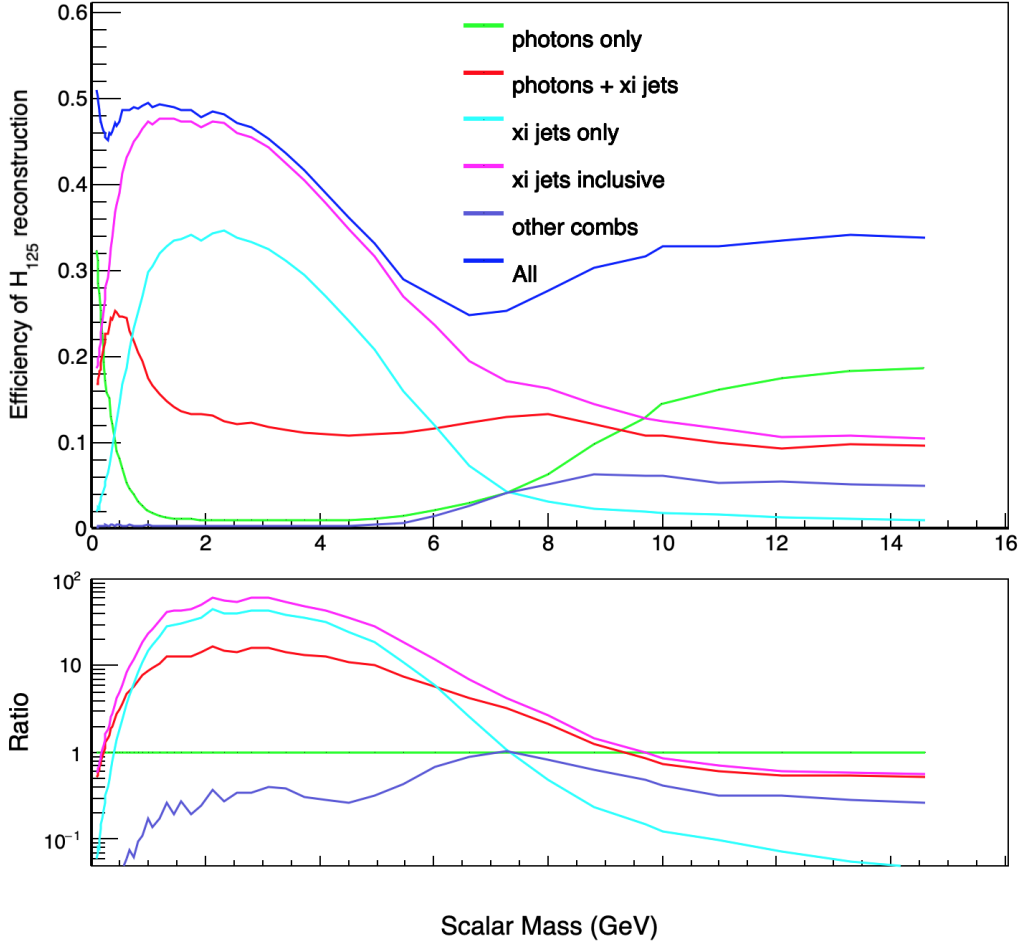


Figure 6: Top: Efficiency of h_{125} reconstruction as a function of scalar mass split into different categories based on number of ξ -jets and photons. Bottom: Ratio to only using photons. The reconstruction efficiency is more than an order of magnitude better when including ξ -jets (pink) over a wide range of masses from 100 MeV to 10 GeV. More specifically, the 2 ξ -jet channel dominates from the range of 300 MeV to 6 GeV.

helpful conversations on these issues. This work was supported in part by the DOE under grant DE-SC0007859. B. Sheff is supported by the NSF GRFP program, and N. Steinberg is supported by a fellowship from the Leinweber Center for Theoretical Physics.

References

- [1] CMS Collaboration, A. M. Sirunyan *et al.*, “Search for invisible decays of a Higgs boson produced through vector boson fusion in proton-proton collisions at $\sqrt{s} = 13$

- TeV,” *Phys. Lett. B* **793** (2019) 520–551, [arXiv:1809.05937 \[hep-ex\]](#).
- [2] **ATLAS** Collaboration, M. Aaboud *et al.*, “Measurements of Higgs boson properties in the diphoton decay channel with 36 fb^{-1} of pp collision data at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS detector,” *Phys. Rev. D* **98** (2018) 052005, [arXiv:1802.04146 \[hep-ex\]](#).
 - [3] G. Belanger, B. Dumont, U. Ellwanger, J. Gunion, and S. Kraml, “Global fit to Higgs signal strengths and couplings and implications for extended Higgs sectors,” *Phys. Rev. D* **88** (2013) 075008, [arXiv:1306.2941 \[hep-ph\]](#).
 - [4] **ATLAS** Collaboration, G. Aad *et al.*, “Combined measurements of Higgs boson production and decay using up to 80 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13 \text{ TeV}$ collected with the ATLAS experiment,” *Phys. Rev. D* **101** no. 1, (2020) 012002, [arXiv:1909.02845 \[hep-ex\]](#).
 - [5] **ATLAS, CMS** Collaboration, G. Aad *et al.*, “Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV ,” *JHEP* **08** (2016) 045, [arXiv:1606.02266 \[hep-ex\]](#).
 - [6] D. Curtin *et al.*, “Exotic decays of the 125 GeV Higgs boson,” *Phys. Rev. D* **90** no. 7, (2014) 075004, [arXiv:1312.4992 \[hep-ph\]](#).
 - [7] **ATLAS** Collaboration, G. Aad *et al.*, “A search for the dimuon decay of the Standard Model Higgs boson with the ATLAS detector,” [arXiv:2007.07830 \[hep-ex\]](#).
 - [8] **ATLAS** Collaboration, G. Aad *et al.*, “Combination of searches for Higgs boson pairs in pp collisions at $\sqrt{s} = 13 \text{ TeV}$ with the ATLAS detector,” *Phys. Lett. B* **800** (2020) 135103, [arXiv:1906.02025 \[hep-ex\]](#).
 - [9] B. A. Dobrescu, G. L. Landsberg, and K. T. Matchev, “Higgs boson decays to CP odd scalars at the Tevatron and beyond,” *Phys. Rev. D* **63** (2001) 075003, [arXiv:hep-ph/0005308](#).
 - [10] P. Draper and D. McKeen, “Diphotons from Tetraphotons in the Decay of a 125 GeV Higgs at the LHC,” *Phys. Rev. D* **85** (2012) 115023, [arXiv:1204.1061 \[hep-ph\]](#).
 - [11] A. Collaboration, “Search for new phenomena in events with at least three photons collected in pp collisions at $\sqrt{s} = 8 \text{ TeV}$ with the ATLAS detector,” *The European Physical Journal C* **76** no. 4, (Apr., 2016) 210. <http://arxiv.org/abs/1509.05051>. [arXiv: 1509.05051](#).
 - [12] S. Chang, P. J. Fox, and N. Weiner, “Visible Cascade Higgs Decays to Four Photons at Hadron Colliders,” *Phys. Rev. Lett.* **98** (2007) 111802, [arXiv:hep-ph/0608310](#).
 - [13] M. T. et al. (Particle Data Group) *Phys. Rev. D* **98**, 030001 (2018) .

- [14] S. D. Ellis, T. S. Roy, and J. Scholtz, “Phenomenology of Photon-Jets,” *Phys. Rev.* **D87** no. 1, (2013) 014015, [arXiv:1210.3657 \[hep-ph\]](#).
- [15] **ATLAS** Collaboration, “Search for a Higgs boson decaying to four photons through light CP-odd scalar coupling using 7 TeV pp collision data taken with ATLAS detector at the LHC,”.
- [16] H. A. Chawdhry, M. Czakon, A. Mitov, and R. Poncelet, “NNLO QCD corrections to three-photon production at the LHC,” *Journal of High Energy Physics* **2020** no. 2, (Feb., 2020) 57. <http://arxiv.org/abs/1911.00479>. [arXiv: 1911.00479](#).
- [17] **ATLAS** Collaboration, M. Aaboud *et al.*, “A search for pairs of highly collimated photon-jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector,” *Phys. Rev.* **D99** no. 1, (2019) 012008, [arXiv:1808.10515 \[hep-ex\]](#).
- [18] A. Collaboration, “Electron and photon performance measurements with the ATLAS detector using the 2015-2017 LHC proton-proton collision data,” *Journal of Instrumentation* **14** no. 12, (Dec., 2019) P12006–P12006. <http://arxiv.org/abs/1908.00005>. [arXiv: 1908.00005](#).
- [19] T. Robens and T. Stefaniak, “Status of the Higgs Singlet Extension of the Standard Model after LHC Run 1,” *Eur. Phys. J. C* **75** (2015) 104, [arXiv:1501.02234 \[hep-ph\]](#).
- [20] V. Barger, P. Langacker, M. McCaskey, M. J. Ramsey-Musolf, and G. Shaughnessy, “LHC Phenomenology of an Extended Standard Model with a Real Scalar Singlet,” *Phys. Rev. D* **77** (2008) 035005, [arXiv:0706.4311 \[hep-ph\]](#).
- [21] T. Robens, T. Stefaniak, and J. Wittbrodt, “Two-real-scalar-singlet extension of the SM: LHC phenomenology and benchmark scenarios,” *Eur. Phys. J. C* **80** no. 2, (2020) 151, [arXiv:1908.08554 \[hep-ph\]](#).
- [22] D. López-Val and T. Robens, “ Δr and the W-boson mass in the singlet extension of the standard model,” *Phys. Rev. D* **90** (2014) 114018, [arXiv:1406.1043 \[hep-ph\]](#).
- [23] C. Csaki, R. T. D’Agnolo, M. Geller, and A. Ismail, “Crunching Dilaton, Hidden Naturalness,” [arXiv:2007.14396 \[hep-ph\]](#).
- [24] **CDF** Collaboration, T. A. Aaltonen *et al.*, “First Search for Exotic Z Boson Decays into Photons and Neutral Pions in Hadron Collisions,” *Phys. Rev. Lett.* **112** (2014) 111803, [arXiv:1311.3282 \[hep-ex\]](#).
- [25] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations,” *JHEP* **07** (2014) 079, [arXiv:1405.0301 \[hep-ph\]](#).

- [26] T. Sjostrand, S. Mrenna, and P. Z. Skands, “PYTHIA 6.4 Physics and Manual,” *JHEP* **05** (2006) 026, [arXiv:hep-ph/0603175](#).
- [27] T. Sjostrand, S. Mrenna, and P. Z. Skands, “A Brief Introduction to PYTHIA 8.1” *Comput. Phys. Commun.* **178** (2008) 852–867, [arXiv:0710.3820 \[hep-ph\]](#).
- [28] **DELPHES 3** Collaboration, J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens, and M. Selvaggi, “DELPHES 3, A modular framework for fast simulation of a generic collider experiment,” *JHEP* **02** (2014) 057, [arXiv:1307.6346 \[hep-ex\]](#).
- [29] M. Cacciari, G. Salam, and G. Soyez, “FastJet user manual,” *Eur. Phys. J. C* **72** no. 1896, (2012) , [arXiv:1111.6097 \[hep-ph\]](#).
- [30] J. Thaler and K. Van Tilburg, “Identifying Boosted Objects with N-subjettiness,” *JHEP* **03** (2011) 015, [arXiv:1011.2268 \[hep-ph\]](#).
- [31] A. Chakraborty, A. M. Iyer, and T. S. Roy, “A Framework for Finding Anomalous Objects at the LHC,” *Nuclear Physics B* **932** (July, 2018) 439–470. <http://arxiv.org/abs/1707.07084>. [arXiv: 1707.07084](#).
- [32] **CMS** Collaboration, “Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET,”.