Phenomenology of Photon-Jets

Stephen D. Ellis, Tuhin S. Roy, and Jakub Scholtz
Physics Department, University of Washington,
Seattle, WA 98195-1560, USA

(Dated: May 3, 2014)

One of the challenges of collider physics is to unambiguously associate detector based objects with the corresponding elementary physics objects. A particular example is the association of calorimeter-based objects such as "jets", identified with a standard (IR-safe) jet algorithm, with the underlying physics objects, which may be QCD-jets (arising from a scattered parton), electrons, photons and, as discussed here, photon-jets (a group of collinear photons). This separation is especially interesting in the context of Higgs searches, where the signal includes both di-photon (in the Standard Model) and di-photon-jet decays (in a variety of Beyond the Standard Model scenarios), while QCD provides an ever-present background. Here we describe the implementation of techniques from the rapidly evolving area of jet-substructure studies to not only enhance the more familiar photon-QCD separation, but also separately distinguish photon-jets, i.e., separate usual jets into three categories: single photons, photon-jets and QCD. The efficacy of these techniques for separation is illustrated through studies of simulated data.

I. INTRODUCTION

The Large Hadron Collider (LHC) has clearly exhibited its ability to make discoveries with the observation of a new resonance [1, 2] with even spin that decays to photons and Z bosons as expected of the Standard Model (SM) Higgs particle. Thus precise measurements of the decays of this resonance into various channels (whether standard or not), are of the utmost importance. At the same time, it is essential to verify our understanding of the existing channels, in particular, $h \rightarrow \gamma\gamma$. How well are these photons defined? Can physics objects other than single photons leave signatures in the detector similar to that of a photon? Not surprisingly, the answer is yes [3, 4]. Given the granularity of the calorimeters, an object consisting of (nearly) collinear photons, typically labeled a photon-jet, will generate a signature similar to that of a single photon. The possibility that the Higgs particle decays to multiple collinear photons is not new [3, 5]. Simple models where the Higgs decays to almost massless scalars that each in turn decay to a pair of photons, typically do not give rise to events with four separately identifiable photons, but rather to pairs of photon-jets, each with 2 photons. Slightly more complicated models can produce Higgs decays to photon-jets with 4, 6, ··· photons. We will discuss concrete models where the Higgs decays to photon-jets with 2 and 4 photons per photon-jet. Thus it is essential to develop tools to separate single photons from photon-jets from QCD-jets. Otherwise we are unlikely to understand either the signal or the background.

ATLAS recently made attempts to identify photon-jets from Higgs decays [6]. These analyses rely on relaxing the isolation/shower shape criteria, which use the differing distributions of energy deposition within the calorimeter cells to quite successfully discriminate single photons from QCD-jets. Unfortunately, the parameters of the underlying model can be easily adjusted so that the resultant photon-jets pass the strictest isolation/shower shape criteria just like photons. More importantly, loosening isolation criteria results in a larger fake rate for QCD-jets. Discriminating photon-jets from QCD-jets is more challenging than separating single photons from QCD-jets.

Fortunately jet substructure techniques [7–12] have recently been developed to distinguish QCD-jets from jets containing boosted heavy particle decays, and we can use this work for detection of photon-jets. More broadly, ‘jets’, as defined by an infrared safe jet clustering algorithm, are being proposed as a universal language to describe all calorimeter objects including single photons, photon-jets and QCD-jets. By using the tools developed in jet substructure physics, we do not need to rely on isolation cuts. We supplement the traditional/conventional variables currently used to discriminate photons from QCD-jets with substructure variables that probe in detail the energy distribution within the jet. Note that the photons-jets are composed of energetic photons distributed inside the jet, where the distribution is a result of the kinematic features of the model, e.g., the masses and spins of the intermediate particles. The existence of this structure within photon-jet suggests that substructure variables will be efficient at finding and discriminating photon-jets. We show that our analysis is capable of separating photon-jets from both single photons and QCD-jets at least as efficiently as the traditional discriminators separate photons from QCD-jets.

There is another important advantage to applying jet substructure techniques to purely electro-magnetic calorimeter (ECal) objects. The introduction of ‘grooming’ algorithms (including filtering [10], pruning [15], and trimming [17]) promised to suppress the undesirable contributions to purely hadronic jets from the underlying event (the largely uncorrelated soft interactions surrounding the interesting hard scattering) and from pile-up (the truly uncorrelated proton-proton collisi-
sions that occur in the same time window). Indeed, the recent results from studies at ATLAS [18] and CMS [19] indicate this grooming is effective. We expect that this substructure-based grooming will work as well for all ECal based objects.

It should be noted that in the context of Higgs physics, the decay to photon-jets is not the only example where the collinearity of the decay products adds complexity to the analysis. Collinearity plays a role for traditional decays of the Higgs boson when it is boosted. In Ref. [10], the authors exploited the collinearity of the $b$-quarks in boosted Higgs decays (both quarks in a single jet) to greatly enhance the chances of detecting the $h \rightarrow b\bar{b}$ channel, featuring jet substructure as a mainstream tool (see also Refs. [7–9]). The application of jet substructure in Higgs physics has now become a very active area of research, applied both to the SM Higgs [20–22] as well as beyond the SM Higgs scenarios [23–25]. For reviews, more detailed descriptions, and references see Refs. [22, 29].

The paper is organized as follows: in Sec. II we start with a simplified model for photon-jets. We propose a set of benchmark points, where we take different combinations of masses and parameters in the simplified model to produce photon-jets displaying a variety of distinct kinematics. In Sec. III we define the details of our simulation. We describe, in detail, how we generate samples of photon-jets, one for each of the benchmark points, QCD-jets, and single photons. We present our analysis in Sec. IV. We describe all the variables that we use in this work to discriminate photon-jets from QCD-jets from single photons. Then we combine these variables in a multivariate analysis. We train boosted decision trees (BDTs) using the samples of jets and use these to optimize the discriminating power of our analyses. We also show how these BDTs can be used to simultaneously separate photon-jets, photons, and QCD-jets from each other. Our conclusions are presented in Sec. V.

II. SIMPLE MODEL FOR PHOTON-JETS

By definition, photons-jets refer to calorimeter objects consisting of more than one hard photon. However, such a broad definition presents a challenge since all photon-jets are not the same. They differ in terms of the number of hard constituent photons as well as in the distribution of those photons within the photon-jet. To provide a systematic phenomenological study of photon-jets we classify these objects in more detail in terms of the production mechanism and consider a broad range. We will refer to the various production scenarios as ‘benchmark’ scenarios. We find that a simple model in the spirit of Ref. [31] with two new particles is sufficient to characterize these benchmarks. The model includes a small number of interactions and we can vary the strength of these interaction and the new particle masses in order to generate the benchmark scenarios. In particular, we introduce two scalar fields $n_1$ and $n_2$ of mass $m_1$ and $m_2$ respectively. Without loss of generality, we choose the naming convention such that $m_1 > m_2$. Neither $n_1$ nor $n_2$ carry any SM charges. We use the following interactions to generate photon-jets

$$\frac{1}{2} \mu_h h n_1^2 + \frac{1}{2} \mu_{12} n_1 n_2^2 + \left( \frac{n_1}{m_1} + \frac{n_2}{m_2} \right) F_{\mu\nu} F^{\mu\nu}, \quad (1)$$

where $\mu_h, \mu_{12}$ are mass parameters, $n_1, n_2$ are dimensionless coupling constants, and $F_{\mu\nu}$ is the electromagnetic field strength operator.

This simple model bears a resemblance to a Higgs portal scenario [32, 34] because of the $\mu_h$ coupling. In the Higgs portal language, $n_1$ and $n_2$ constitute a ‘hidden’ sector while the coupling $\mu_h$ provides a tunnel to the corresponding ‘hidden valley’. The electromagnetic couplings (proportional to the $\eta$ parameters) provide ways for the new particles to decay back to SM particles, photons in this case. With respect to Higgs physics, this simple model provides a realistic example where the SM Higgs field decays through the new particles to multiple photons. In the limit $m_1 \ll m_h$, the resultant photons (the decay products of $n_1$) are essentially collinear.

In Table I we list the benchmark scenarios (labeled photon-jet study points or PJSPs) that we investigate in this work. All are generated by varying the parameters in Eq. (1). The symbol X in Table I denotes that a non-zero value is selected for that parameter, which then determines the decay mode. We have chosen the benchmarks in such a way that the parameters denoted by X only change the total width of the decaying particles. As long as the decays are prompt, the exact values of these parameters are irrelevant to the phenomenological properties of the photon-jets. In all these study points we take the Higgs particle to decay to a pair of $n_1$ particles. The small $n_1$ mass ($m_1 \ll m_h$) ensures that the decay products of the $n_1$ are highly collimated. In the Higgs particle rest frame, which is close to the laboratory frame on average, each $n_1$ has momentum $\sim m_h/2$ and the typical angular separation between the $n_1$ decay products is of the order of $4m_1/m_h$. Note that, given

| Study Points | $m_1$ (GeV) | $m_2$ (GeV) | $\mu_{12}$ (GeV) | $\eta_1$ | $\eta_2$ |
|-------------|-------------|-------------|-----------------|---------|---------|
| PJSP 1      | 0.5         | 1.5         | 0               | X       |         |
| PJSP 2      | 1.0         | 10.0        |                 |         |         |
| PJSP 3      | 2.0         | 0.5         |                 | X       | X       |
| PJSP 4      | 5.0         | 0.5         |                 |         |         |
| PJSP 5      | 5.0         | 0.5         |                 |         |         |
| PJSP 6      | 10.0        | 1.0         |                 |         |         |
| PJSP 7      | 10.0        | 0.5         |                 |         |         |
| PJSP 8      | 10.0        | 1.0         |                 |         |         |

TABLE I. The study points used in our analysis. For PJSP 1–3, $n_2$ does not participate in the decay chain since $\mu_{12} = 0$ and the $m_2$ and $n_2$ columns are empty. By X we denote that a non-zero value is chosen for the parameter, which facilitates prompt decays, but the specific value plays no role.
we always consider $m_1 \leq 10$ GeV, we expect the typical angular separation between the $n_1$ decay products to be $\lesssim 1/3$ (we use $m_0 = 120$ GeV). As long as the angular size of photon-jets is larger than $1/3$, we expect to capture all the decay products of the $n_1$ in each photon-jet for all the benchmark points.

For the study points PJSP 1 – 3 the mass parameter $\mu_{12}$ is set to zero and $n_1 \rightarrow \gamma\gamma$ is the only possible $n_1$ decay mode. Hence these scenarios are characterized by photons-jets with typically 2 hard photons per jet, and no other coupling values are included in the table. In these scenarios the Higgs particle cascade decays to four photons. (In this case, a highly boosted decay results in a single photon-jet containing all 4 photons.)

For study points PJSP 4 – 8 we set $\eta_1$ to zero and $\mu_{12}$ to a non-zero value. In these contrasting scenarios the only $n_1$ decay mode involves the chain $n_1 \rightarrow n_2 n_2 \rightarrow \gamma\gamma\gamma\gamma$. Hence the Higgs decays again to two photon-jets, but now each photon-jet typically contains four photons (the $n_1$ decay products). (In this case, a highly boosted Higgs yields a single photon-jet containing 8 photons.)

### III. SIMULATION DETAILS

In order to generate samples of photons-jets, we implement the simple model of Eq. 1 in MadGraph 5 [35]. For each benchmark point we generate matrix elements corresponding to the process $pp \rightarrow h \rightarrow n_1 n_1$ (via gluon fusion) using MadGraph 5 with $m_h = 120$ GeV, which we employ as input to Pythia 8.1 [36, 37] in order to generate the full events and for the subsequent $n_1$ decays. Since the Higgs production is evaluated at lowest order, the produced Higgs particles have zero transverse momentum. We use the QCD dijet events generated by standalone Pythia 8.1 to provide a sample of QCD-jets. In order to define a sample of single photons, we also generate $pp \rightarrow h \rightarrow \gamma\gamma$ events where the photons are well separated. Finally, we include initial state radiation (ISR), final state radiation (FSR) and multiple parton interactions (MI, i.e., the UE) as implemented in Pythia 8.1 to simulate the relevant busy hadronic collider environment.

The Pythia output final states are subjected to our minimal detector simulation. In the following we describe briefly how we treat the final state particles in each event:

- We identify all charged particles with transverse momentum $p_T > 2$ GeV and pseudorapidity $|\eta| < 2.5$ as charged tracks.
- In a real detector, tracks are also generated if photons convert within the pixel part of the tracker. In this work, we simulate this photon conversion process by associating with each photon a probability for it to convert in the tracker.  

The probability is a function of the number of radiation lengths of material the photon has to traverse in order to escape the inner part of the tracker. We use the specifications of the ATLAS detector in order to model this pseudorapidity dependent probability distribution. The details of this procedure are outlined in the Appendix A1.

- In our simulation, all particles (except charged particles with $E < 0.1$ GeV) reach the calorimeters, and all of these (except muons with $p_T > 0.5$ GeV) deposit all of their energy in the calorimeters. The electromagnetic calorimeter (ECal) is modeled as cells of size $0.025 \times 0.025$ in ($\eta$-$\phi$), whereas the hadronic calorimeter (HCal) is taken to have more coarse granularity with $0.1 \times 0.1$ cells. Besides photons and electrons, soft muons and soft hadrons (soft means $E < 0.5$ GeV) are treated as depositing all of their energy in the ECal. More energetic hadrons are absorbed in the HCal, while more energetic muons escape the calorimeter. For a more detailed picture see Appendix A2.

- We attempt to simulate the showering that occurs within the ECal. We distribute the energy of each particle that is absorbed in the ECal into a (3 $\times$ 3) grid of cells (centered on the direction of the original particle) according to a precomputed Molière matrix corresponding to the Molière radius of lead. For details on this transverse smearing see Appendix A3. The structure induced by this shower simulation is observable in our final results.

- We implement calorimeter energy smearing for both the ECal and the HCal. The calorimetric response is parametrized through a Gaussian smearing of the accumulated cell energy $E$ with a variance $\sigma$:

$$\frac{\sigma}{E} = \frac{S}{\sqrt{E}} + C,$$

where $S$ and $C$ are the stochastic and constant terms. For the ECal and the HCal, we use $(S, C)$ to be $(0.1, 0.01)$ and $(0.5, 0.03)$, respectively, in order to approximately match the reported calorimeter response from ATLAS [38].

- Each calorimeter cell that passes an energy threshold becomes an input for our jet clustering algorithm. For the ECal cells we require $E_T > 0.1$ GeV,
while for the HCal cells we use the somewhat harder cut $E_T > 0.5$ GeV. We sum all the energy deposited in a given calorimeter cell and construct a massless 4-vector with the 3-vector direction corresponding to the location of that cell.

- As the final step we cluster the 4-vectors corresponding to the calorimeter cells into jets using Fastjet 3.0.3 [39, 40]. In particular, we use the anti-$k_T$ jet clustering algorithm [41] with $R = 0.4$ and and require $p_T > 50$ GeV for every jet. Only the leading jet from each event is retained for further analysis in order to maintain independence among the jets in the sample.

IV. ANALYSIS

In this section we describe the analysis of 10 samples of jets generated according to the prescription of the previous sections. The first sample contains QCD jets derived from QCD dijet events. The second sample consists of jets from $h \rightarrow \gamma\gamma$ events where each jet typically contains one of the photons from the Higgs decays, plus contributions from the rest of the event (ISR, FSR, UE). We refer to the jets in this sample as single photon jets, or simply single photons. The remaining 8 samples of jets are the photon-jet samples and correspond to the 8 study points in Table I. As noted above, in these events the Higgs particle decays into 4 or 8 photons and the corresponding photon-jets typically contain either 2 or 4 photons. The resulting $p_T$ distributions for QCD-jets (red), photon-jets (blue) (PJSP 8) and single photons (green) are indicated in Fig. 1.

As expected, the $p_T$ distribution for QCD-jets is a falling distribution, while both the single photon and photon-jet distributions exhibit a peak near $m_h/2(=60$ GeV). We understand this last point as arising from the production of Higgs particles with zero transverse momentum followed by 2-body decays (either 2 photons or 2 $n_1$’s). It is the remnants of these two bodies that are typically captured in the jets yielding the indicated peaks near $p_T \sim m_h/2$. For the photon-jet sample we only show the $p_T$ spectrum for the study point PJSP 8, but note that the $p_T$ distributions are almost identical for all other benchmark points. As indicated in Fig. 1, the jets in all of these samples of events have crudely comparable transverse momentum distributions in the range 50–100 GeV, although the QCD sample is more strongly peaked at the low end. Thus studying the jets in these samples should provide a useful laboratory in which to study photon-jets, QCD jets and single photons.

The remainder of this section describes a systematic analysis aimed at distinguishing photon-jets from QCD-jets as well as from single photons. We begin with brief descriptions of the variables that provide the discriminating power. The variables are organized into two groups: (i) conventional variables and (ii) substructure variables. We demonstrate how each of these variables individually discriminates photons-jets from the jet samples. Later in this section, we combine these variables in a multivariate analysis in order to maximize the separation of photon-jets from QCD-jets as well as from single photons.

A. Conventional Variables

The conventional variables we describe below are well known, well understood, and play essential roles in the identification of single photons, i.e., the separation from QCD-jets. We expect these variables to play a similar role in separating photon-jets from QCD-jets, since the probability distributions as functions of these variables are similar for photon-jets and for single photons. On the other hand, they cannot be expected to efficiently discriminate photon-jets from single photons.

1. Hadronic Energy Fraction, $\theta_J$

We define the hadronic energy fraction $\theta_J$ for a jet to be the fraction of its energy deposited in the hadronic calorimeter:

$$\theta_J = \frac{1}{E_j} \sum_{i \in \text{HCal} \in J} E_i$$

where $E_J$ is the total energy of the jet, and $E_i$ is the energy of the $i$-th HCal cell that is a constituent of the jet. This is the most powerful variable for discriminating a single photon or a photon-jet (objects that deposit most of their energy in the ECal) from QCD-jets. Since

\[\text{arb.units}\]

\[0.005 \quad 0.010 \quad 0.015 \quad 0.020 \quad 0.025 \quad 0.030 \quad 0.035 \]

\[p_T \quad h \quad \gamma \gamma \]

\[\text{arb.units}\]

\[0.035 \quad 0.030 \quad 0.025 \quad 0.020 \quad 0.015 \quad 0.010 \quad 0.005 \]

\[50 \quad 60 \quad 70 \quad 80 \quad 90 \quad 100 \]

\[p_T \quad (\text{GeV})\]

\[\text{FIG. 1. The } p_T \text{ distribution of jets for QCD-jets (red), single photons (green), and for photon-jets (blue) from the study point PJSP 8. Jets are constructed as described in the text (the anti-}k_T\text{ algorithm with } R = 0.4).}\]
a QCD-jet typically contains 2/3 charged pions and 1/3 neutral pions, we expect to see a peak at \( \theta_J \sim 2/3 \) (log \( \theta_J \sim -0.2 \)) for QCD-jets. Isolated single photons and photon-jets, on the other hand, should exhibit very small \( \theta_J \) values. However, we start with objects identified by a jet algorithm so there will be contributions from the rest of the event and pile-up, and from leakage from the ECAL into the HCAL. Thus the precise value of \( \theta_J \) for single photons and photon-jets will depend on detailed detector properties and on the contribution from the underlying event and pile-up. Nevertheless, we expect single photons/photon-jets to exhibit very small values of \( \theta_J \).

Any reasonable cut on \( \theta_J (\theta_J \sim 0.1) \) will reduce the QCD-jet contribution by factors of \( 10^{-2} - 10^{-3} \), while barely changing the photon/photon-jet contribution. We impose a preliminary cut by keeping only \( \theta_J \leq 0.25 \) (log \( \theta_J \leq -0.6 \)). About 2% of the original QCD-jets survive this cut, while approximately 94% of the single photons/photon-jets survive. We use the modified jet samples that pass this preliminary \( \theta_J \) cut for the remainder of this paper.

2. Number of Charged Tracks, \( \nu_J \)

In conventional collider phenomenology, the number of charged particles (tracks) associated with an object is often used to distinguish objects from each other. Although photons and electrons generate similar signatures in the ECAL, the latter are typically associated with a track while the former are not. Tracks also play an important role in rejecting QCD-jets since, as mentioned before, a QCD-jet typically contains several charged pions.

In our simulated data we keep all charged particles with \( p_T > 2 \) GeV and assume that all of these correspond to tracks in a real detector. In order to associate these tracks with the jets, which are constructed entirely from calorimeter cells, we perform the following analysis. First replace each track by an arbitrarily soft light-like four vector with the same \((\eta, \phi)\) direction as the track, and then include these soft four-vectors in the jet clustering process along with the calorimeter cells. (We explicitly check that the inclusion of these soft four-vectors does not affect the outcome of the clustering procedure.) A track is associated with a jet if the soft four vector corresponding to that track is clustered into that jet. The resulting total number of tracks associated with a jet yields the value of \( \nu_J \) for that jet. Figure 3 shows the relative probability distribution versus the number of charged tracks associated with a jet. The algorithm for associating tracks with jets is given in the text. For photon-jets we show the distribution for jets from the study points PJSP 1 (dotted) and PJSP 8 (solid).

![FIG. 2. The probability distributions for jets as functions of log \( \theta_J \) for QCD-jets (red), single photons (green) and photon-jets from PJSP 8 (blue). The first bin of the plot (at \( \theta_J = 10^{-3} \)) has an open lower boundary, i.e., it includes all jets with log \( \theta_J \) < -3.0.](image)

![FIG. 3. The relative probability distribution for QCD jets (red), single photons (green) and photon-jets (blue) versus the number of charged tracks associated with a jet. The algorithm for associating tracks with jets is given in the text. For photon-jets we show the distribution for jets from the study points PJSP 1 (dotted) and PJSP 8 (solid).](image)

3 As a check we also consider the more traditional construction where a track is associated with a jet if it is within an angular distance \( R \) or less from the given jet’s direction, where \( R \) is the size-parameter used in the clustering algorithm. For anti-\( k_T \) jets both methods yield identical associations of tracks and jets. For the \( k_T \) or C/A algorithms, where jets are not exactly circular, the method described in the text is a more natural definition of whether a track is associated with a jet or not.
of tracks per jet ($\nu_J$) for QCD-jets (red), single photons (green) and photon-jets (blue). As expected, the number of tracks associated with QCD-jets varies over a broad range and only a tiny fraction of QCD-jets have no associated tracks. The single photon/photon-jet samples, on the other hand, are dominated by jets with no associated tracks. Photons that convert yield tracks associated with the corresponding jets. Since the probability of conversion increases with the number of photons per jet, the probability of obtaining one of more associated tracks increases from single photon jets (single photons) to jets with two photons (typical for PJSP 1, the dotted blue curve) to jets with four photons (typical for PJSP 8, the solid blue curve). As with the variable $\theta_j$, $\nu_j$ offers some separation between QCD-jets and single photons, but much less between single photons and photon-jets (and even less between the different types of photon-jets).

B. Jet Substructure

Next we want to focus on variables that explicitly characterize the internal structure of jets, i.e., characterize the energetic subjet components of the jet. Recall that in this analysis we have identified jets using the the anti-$k_T$ jet algorithm with $R = 0.4$, but we do not expect the general features of our analysis to depend on this specific choice. The next step is to determine a ‘recombination tree’ for the jets we want to study (here leading jet in each event). To this end we apply the $k_T$ algorithm \cite{42,43} to the calorimeter cells identified as constituents of the jet in the first step. (We could as well use the Cambridge/Aachen (C/A) algorithm \cite{44,45}, but not the anti-$k_T$ algorithm in this step as anti-$k_T$ does not tend to produce a physically relevant recombination tree.) This recombination tree specifies the subjets at each level of recombination $N$ from $N = 1$ (the full jet) to $N = \text{the number of constituent calorimeter cells in the jet (no recombination)}$. At the next step the subjet variables we study fall into two classes. In the first class we attempt to count the effective number of relevant subjets without using any properties of the subjets in the tree except their directions in $p_\perp$. In this case the useful variable (defined in detail below) is called $N$-subjettiness. The $N$-subjettiness variable for a given jet becomes numerically small when the parameter $N$ is large enough to describe all of the relevant substructure, i.e., this value of $N$ provides a measure of the number of subjets without explicitly identifying the subjets. $N$-subjettiness involves all components of the original jet for all values of $N$.

The rest of the substructure variables we study more explicitly resolve a jet into a set of subjets. We define both the level in the recombination tree at which we choose to work, i.e., the number of subjets we have split the jet into and how many of these subjets to use in the subsequent analysis. We use $N_{\text{pre-filter}}$ (this notation should become clear shortly) and $N_{\text{hard}}$ to label these two parameters. Thus we start with the 4-vectors corresponding to the (calorimeter cell) constituents of a given jet, and then (re)cluster these constituents using the chosen subjet algorithm (which is not necessarily the algorithm used to originally identify the jet) in exclusive mode, i.e., we continue (re)clustering until there are precisely $N_{\text{pre-filter}}$ 4-vectors left – the $N_{\text{pre-filter}}$ exclusive subjets. Out of these $N_{\text{pre-filter}}$ subjets we pick the $N_{\text{hard}}$ largest $p_T$ subjets and discard the rest. All the substructure variables discussed below (except $N$-subjettiness) are constructed using these $N_{\text{hard}}$ subjets. Note that by choosing $N_{\text{pre-filter}} > N_{\text{hard}}$, we have performed a version of jet ‘grooming’ typically labeled filtering \cite{10,13,14}. This will ensure that our results are relatively insensitive to the effects of the underlying event and pile-up. Ideally, the integers ($N_{\text{hard}}, N_{\text{pre-filter}}$) should be chosen based on the topology of the object we are looking for. However, the naive topology will be influenced by the interaction with the detector and the details of the jet clustering algorithm. For example, a 4 photon photon-jet will often appear in the detector to have fewer than 4 distinct lobes of energy, i.e., one or more photons often merge inside a single lobe of energy. In our simulation, we find that the choice $N_{\text{hard}} = 3$ and $N_{\text{pre-filter}} = 5$ is an acceptable compromise, working reasonably well for single photons and photon-jets from all the study points. Further optimization will be possible in the context of real detectors and searches for specific photon-jet scenarios.

1. $N$-Subjettiness, $\tau_N$

"$N$-subjettiness", introduced in Ref. \cite{47,48}, is a modified version of "$N$-jettiness" from Ref. \cite{49}. It is adapted in a way such that it becomes a property of a jet rather than of an event. $N$-subjettiness provides a simple way to effectively count the number of subjets inside a given jet. It captures whether the energy flow inside a jet deviates from the one-lobe configuration expected to characterize a typical QCD-jet. We use the definition of $N$-subjettiness proposed in Ref. \cite{47}. The starting point is a jet, the full set of 4-vectors corresponding to the (calorimeter cell) constituents of the jet (here found with the anti-$k_T$ algorithm for $R = 0.4$), and the recombination tree found with the $k_T$ algorithm as outlined above. From this tree we know the 4-vectors describing the exclusive subjets for any level $N$, i.e., the level where there are exactly $N$ subjets. With this information we can define $N$-subjettiness to be

$$
\tau_N = \frac{\sum_k p_{T_k} \times \min\{\Delta R_{1,k}, \Delta R_{1,k}, \ldots, \Delta R_{N,k}\}}{\sum_k p_{T_k} \times R},
$$

(4)

where $k$ runs over all the (calorimeter cell) constituents of the jet, $p_{T_k}$ is the transverse momentum for the $k$-th constituent, $\Delta R_{l,k} = \sqrt{(\Delta \eta_{l,k})^2 + (\Delta \phi_{l,k})^2}$ is the angular distance between the $l$-th subjet (at the level when there are $N$ subjets) and the $k$-th constituent of the jet, and $R$ is the characteristic jet radius used in the original jet clustering algorithm.
In the context of single photons, photon-jets and QCD-jets, we use N-subjettiness in two different ways. The first application is to use the ratios $\tau_{N+1}/\tau_N$ in the same way N-subjettiness is used to tag boosted massive particles such as a $W$ boson or a hadronic decaying top [37, 38]. In particular, for a jet with $N_0$ distinct lobes of energy, $\tau_{N_0}$ is expected to be much smaller than $\tau_{N_0-1}$ (of course, we are assuming $N_0 > 1$), whereas for $N > N_0$, $\tau_{N+1}$ is expected to be comparable to $\tau_N$. Thus a two photon photon-jet is expected to be characterized by $\tau_2/\tau_1 \ll 1$. On the other hand, one lobed QCD-jets and single photons should exhibit comparable values for $\tau_2$ and $\tau_1$, and consequently $\tau_2/\tau_1 \sim 1$.

The second way in which we use N-subjettiness consists of using the magnitude of $\tau_1$ itself. Even for a jet with one lobe of energy the exact magnitude of $\tau_1$ represents a measure of how widely the energy is spread. A pencil-like energy profile, like that of a single photon or a narrow photon-jet, should yield a much smaller $\tau_1$ compared to QCD-jets with a much broader profile. In fact, $\tau_1$ is an indicator of jet mass, and, for jets with identical energy, $\tau_1$ is proportional to the square of the jet mass.

Figure 4 shows the probability distributions versus log $\tau_1$ and $\tau_{N+1}/\tau_N$ for $N = 1, 2, 3$ corresponding to single photons, QCD-jets and photon-jets from different study points. Note that for photon-jets, the jet mass is almost always given by the mass parameter $m_1$ in Table 8. Thus for PJSP 8 and PJSP 3, where $m_1$ has the same value, the probability distributions versus log $\tau_1$ are almost identical. For study points PJSP 8, PJSP 4 and PJSP 1 the peak in log $\tau_1$ shifts to the left as the value of $m_1$ decreases (from 10 GeV to 2 GeV to 0.5 GeV). Note also that the PJSP 1 and PJSP 3 distributions exhibit a small $\tau_1$ (small mass) enhancement at essentially the same $\tau_1$ value as the primary peak in the single photon (green curve) distribution. This presumably corresponds to those kinematic configurations where only one of the (two) photons from the $\eta_1$ decay is included in the jet. Thus we expect that a (small) fraction of the time these scenarios will look very single photon-like.

Clearly the ratio $\tau_2/\tau_1$ gives significant separation for the different photons-jet scenarios. The study points PJSP 8 and PJSP 3 are now separated, although both exhibit peaks at small values of the ratio. This suggests an intrinsic 2-lobe structure corresponding to 2 photons for PJSP 3 and 4 photons in two relatively tight pairs ($m_2 \ll m_1$) for PJSP 8. PJSP 4 with presumably a more distinctive 4 photon structure exhibits a broader peak at a larger value of $\tau_2/\tau_1$. Single photons and PJSP 1 exhibit even broader distributions presumably corresponding to an intrinsically 1-lobe structure. The QCD-jet distribution is also broad but with an enhancement around $\tau_2/\tau_1 = 0.8$, presumably arising from a typical 1-lobe structure but some contribution from showers with more structure and from the underlying event. The ratios $\tau_3/\tau_2$ or $\tau_4/\tau_3$ seem to be less effective in discriminating photon-jets from single photons and QCD-jets. This can be understood by noting that quite often the hard photons inside a photon-jet become collinear at the scale of the size of the cell. So even for photon-jets with 4 hard photons, we rarely find jets with 4 distinct centers of energy. In general we expect the ratio $\tau_{N+1}/\tau_N$ becomes less and less useful with increasing $N$.

Note that the distributions for single photons and photon-like photon-jets tend to exhibit a double peak structure in $\tau_3/\tau_2$ or $\tau_4/\tau_3$. We believe that this feature arises from both the contributions due to the underlying event and due to our implementation of transverse smearing in the ECAL (see Appendix A3).

2. Transverse momentum of the Leading Subjet

Now we proceed to discuss the second class of subjet variables constructed from the 3 hardest subjets out of the 5 exclusive subjets. As the first such variable consider the fraction of the jet transverse momentum carried by the leading subjet, which provides significant information about the jet itself. In particular, it indicates the fraction of the jet’s total $p_T$ carried by the leading subjet only. Since photon-jets result from the decay of massive particles into hard and often widely separated photons inside the jet, the subjets are usually of comparable hardness. The leading subjet for single photons and for QCD-jets, on the other hand, typically carry nearly the entire $p_T$ of the jet. So for the majority of these jets, the $p_T$ of the
leading subjet (label it \( p_{T_L} \)) is of the order of the \( p_T \) of the entire jet (\( p_{T_J} \)). Instead of using the ratio \( p_{T_L}/p_{T_J} \), directly we find that it is more instructive to define the variable

\[
\lambda_J = \log \left( 1 - \frac{p_{T_L}}{p_{T_J}} \right). \tag{5}
\]

The advantage of using the definition in Eq. (5) is that it focuses on the behavior near \( p_{T_L} \approx p_{T_J} \).

The discussion above depends crucially on how the subjets are constructed, especially for QCD-jets. QCD partons typically shower into many soft partons/hadrons. After showering and hadronization, single hard partons yield many soft hadrons distributed throughout the jet. The way in which these jets are clustered into subjets dictates the \( p_T \) distribution of subjets. For example, for anti-\( k_T \) subjets, the hardest subjet will always have \( p_{T_L} \approx p_{T_J} \). The \( k_T \) algorithm, on the other hand, clusters the softer elements first and results in more evenly distributed subjets. The C/A jet algorithm clusters taking into consideration only the geometric separations of the elements, and produces qualitatively different results. Single photons, on the other hand, shower very little (no QCD Shower) and deposit energy in only a handful of cells (per hard photon). Therefore we expect that our results for single photons or photon-jets will be less sensitive to the details of the clustering algorithm. To verify this point we use both \( k_T \) and C/A subjets to evaluate \( \lambda_j \) from Eq. (5). The simultaneous use of different clustering algorithms to extract information from the same jet should not come as a surprise. As shown in Ref. [60], substantial further information can be extracted if one employs a broad sampling out of all of the physically sensible clustering histories (trees) for a given jet. In this sense the current analysis is modest in that we only use two specific clustering procedures.

![FIG. 5. Probability distribution for \( \lambda_j \) from Eq. (5). As in Fig. 4 the solid red is for QCD-jets, the solid green for single photons, dotted blue for PJSP 1, dash-dotted blue for PJSP 3, dashed blue for PJSP 4 and solid blue for PJSP 8. The left (right) figure shows the distribution when \( \lambda \) is calculated using C/A (\( k_T \)) subjets.](image)

In Fig. 5 we plot the probability distribution of jets as a function of \( \lambda_j \) for QCD-jets, single photons, and photon-jets. The left (right) panel shows the distribution when we use the C/A (\( k_T \)) algorithm to find the subjets. Note how the distribution for QCD-jets (the red curve) moves more to the right (i.e., the \( p_T \) of the jet gets more evenly distributed among its subjets) as we go from C/A subjets to \( k_T \) subjets. The various photon-jet study points also look more similar when using the \( k_T \) algorithm. In this case the PJSP 1 and PJSP 3 distributions exhibit enhancements suggesting the presence of both single photon-like behavior (\( \lambda_j \sim -1.2 \)) and QCD-like behavior (\( \lambda_j \sim -0.2 \) to \(-0.3 \)). The more complex structure of the PJSP 4 and PJSP 8 jets exhibit a distribution closer to QCD alone. Finally note that the C/A subjets display the jet substructure information differently from the \( k_T \) case with the peak in the QCD-jet distribution at least somewhat separated from the peaks in the photon-jet distributions. Also for C/A all of the photon-jet scenarios exhibit at least a little single photon-like enhancement (for \( k_T \) this is only true for PJSP 1 and PJSP 3). There is clearly some discrimination to be gained from using more than one definition of the subjets. 

3. Energy-Energy Correlation, \( \epsilon_J \)

Another useful variable is the “energy-energy correlation”. We define it as:

\[
\epsilon_J = \frac{1}{E_J^2} \sum_{i>j \in \text{N}_{\text{hard}}} E_i E_j, \tag{6}
\]

where \( E_J \) is the total energy of a given jet, and the indices \( i, j \) run over the (3 hardest) subjets of the jet. From the definition, it should be clear that \( \epsilon_J \) is sensitive to the energy of the subleading jets. In particular, the energy-energy correlation can be expressed as

\[
\epsilon_J = \frac{E_L(ENL+ENN)+ENLENN}{E_J^2} \approx \frac{E_L(E-J)\times ENLENN}{E_J^2}, \tag{7}
\]

where \( E_L, ENL, \) and \( E_{NNL} \) are the energies of the leading subjet, the next-to-leading subjet, and the next-to-next-to-leading subjet.

![FIG. 6. Probability distribution versus \( \epsilon_j \) from Eq. (6). As in Fig. 4 the solid red is for QCD-jets, the solid green for single photons, dotted blue for PJSP 1, dash-dotted blue for PJSP 3, dashed blue for PJSP 4 and solid blue for PJSP 8. The left (right) figure shows the distribution when \( \epsilon_J \) is evaluated using C/A (\( k_T \)) subjets.](image)
We show the probability distribution of jets as a function of $\epsilon_J$ for QCD-jets, single photons and photon-jets in Fig. 6. Note that for single photons (the green curve), $E_{NL}$ and $E_{N,NL}$ are negligible and hence we expect $\epsilon_J$ for single photons to be well approximated by $E_L (E_J - E_L) / E^2_J$. In fact, the sharp peak for single photons in Fig. 5 at $-1.2$ ($k_T$ algorithm) corresponds to the sharp peak at about $0.04$ in Fig. 6. More generally the qualitative features in Fig. 5 are repeated in Fig. 6. For C/A subjets the distributions for all of the photon-jet study points exhibit two peaks, the large $\epsilon_J$ value enhancement presumably corresponding to the energy being shared approximately equally among several final photons, while the small value enhancement arises from the case when one photon dominates (perhaps because some of the photons are not in the jet). For $k_T$ subjets only the PJSP 1 and PJSP 3 distributions exhibit the small $\epsilon_J$ single photon-like enhancement. We also see that again the two algorithms yield distinctly different distributions for QCD-jets.

4. Subjet Spread, $\rho_J$

We define “subjet spread” as a measure of the geometric distribution of the subjets.

$$\rho_J = \frac{1}{R} \sum_{(i,j) \in N_{\text{hard}}} \Delta R_{i,j}, \quad (8)$$

where $\Delta R_{i,j}$ is the angular distance between the $i$-th and $j$-th (hard) subjets, and $R$ is the size parameter of the jet algorithm.

![FIG. 7. Probability distribution for subjet-spread $\rho_J$ from Eq.(8). As in Fig. 4 the solid red is for QCD-jets, the solid green for single photons, dotted blue for PJSP 1, dash-dotted blue for PJSP 3, dashed blue for PJSP 4 and solid blue for PJSP 8. The left (right) figure shows the distribution when $\rho_J$ is calculated using C/A ($k_T$) subjets.]

The left (right) panel of Fig. 7 shows the probability distribution of jets as a function of $\rho_J$ for QCD-jets, single photons and photon-jets when the C/A ($k_T$) subjets are used to evaluate Eq.(8). For this variable only the QCD-jet distribution changes dramatically when changing the choice of subjet algorithm from C/A to $k_T$. By using both algorithms this feature will provide some ability to discriminate between QCD-jets and single photons or photon-jets. For the single photon case the the strong peak at small $\rho_J$ confirms that all of the subjets are close to each other, forming a hard core. Subjet spread is quite sensitive to the mass $m_1$ as can be seen from the different photon-jet distributions. In particular, the position of the peaks for photon-jets with different $m_1$ simply follow the $m_1$ value. The PJSP 3 and PJSP 8 distributions are nearly the same (with the same $m_1$ value), while the PJSP 1 and PJSP 4 distributions are just similar (with somewhat different $m_1$ values), but distinct from PJSP 3 and PJSP 8. The $m_1$ dependence is not surprising since the opening angle between the decay products of the $n_1$ particle depends on $m_1$. Finally we note that the PJSP 3 and PJSP 8 distributions do have an enhancement at small $\rho_J$ values presumably corresponding to configurations where the extra photons are not captured in the jet.

5. Subjet Area of the Jet

As defined in Ref. [51], the “area” associated with a jet is an unambiguous concept that represents quantitatively the amount of surface in the $(\eta,\phi)$ plane included in a jet. In this analysis, we use the “active area” definition for the area of the jet. The active area of a jet is calculated by adding a uniform background of arbitrarily soft ‘ghost’ particles to the event (so that each ghost represents a fixed area) and then counting the number of ghosts clustered into the given jet. The area of a jet is often used to provide a quantitative understanding of the largely uncorrelated contributions to a jet from the underlying event and pile-up. However, it is rarely used in phenomenology for the purpose of discovering new particles or tagging jets. We use “subjet area” as a measure of the ‘cleanliness’ of the jet. We show that it can be a useful tool for distinguishing a single photon or a photon-jet from noisier QCD-jets. We define the subjet area fraction as

$$\delta_J = \frac{1}{A_J} \sum_{i \in N_{\text{hard}}} A_i, \quad (9)$$

where $A_i$ is the area of the $i$-th subjet and $A_J$ is the area of the entire jet. Note that this definition of $\delta_J$ is only useful when the subjets are constructed geometrically by merging the nearest neighbors first (i.e., using the C/A algorithm). In Fig. 8 we show the probability distribution for jets as a function of $\delta_J$ for QCD-jets, single photons, and photon-jets. As expected, the figure shows that single photons (the green curve) are significantly cleaner (exhibit smaller $\delta_J$ values) than QCD-jets (the red curve) and that photon-jets (the blue curves) tend to lie in between. Fixing $m_1$ such that the first splitting is fairly wide, we can investigate the effects of $m_2$. If $m_2$ is small, then the two photons coming from the $n_2$ decays will be very close together, and the subjet that contains them will not collect many ghosts. On the other hand, a large $m_2$ will split the two photons further apart and,
if still contained in the same subset, that subset will collect substantially more ghosts resulting in a subset with a larger active area. QCD-jets contain many soft particles and so the subsets in QCD jets have larger areas. Thus we see that the QCD distribution peaks for $\delta_j$ near 0.5, while the single photon distribution exhibits both a large peak at small ($\sim 10^{-2}$) $\delta_j$ and a smaller peak at larger ($\sim 0.4$) $\delta_j$ values. The photon-jet cases interpolate between these two behaviors and this variable can clearly provide some discriminating power.

C. Multivariate Analysis

We have, so far, introduced a set of well-understood variables. In this subsection, we will employ these variables in a multivariate discriminant, specifically in a Boosted Decision Tree (BDT) \cite{52}. A decision tree is a hierarchical set of one-sided cuts used to discriminate signal versus background. The ‘boosting’ of a decision tree extends this concept from one tree to several trees in a hierarchical set of one-sided cuts used to discriminate signal versus background. The ‘boosting’ of a decision tree extends this concept from one tree to several trees in a hierarchical set of one-sided cuts used to discriminate signal versus background. The ‘boosting’ of a decision tree extends this concept from one tree to several trees in a hierarchical set of one-sided cuts used to discriminate signal versus background.

We construct multiple BDT analyses in order to estimate how well the photon-jets can be separated from single photons and QCD-jets as background. We organize the results of our analysis into three subsections. First, we show the results of BDTs optimized to discriminate QCD-jets from photon-jets, the analogue of the separation of single photons from QCD-jets. In both cases the single photons were treated as the signal. Here we extend that discussion and focus on the photon-jets as the signal. We make use of all the variables discussed earlier in order to minimize the background fake rate ($F$ = the fraction of the background jets that pass the cuts) for a given signal acceptance rate ($A$ = the fraction of the signal jets that pass the cuts). For demonstration purposes we also consider BDTs made with a subset of the full set of variables. To be specific, we consider three different sets of variables:

\begin{align}
D & \equiv \left\{ \log \theta_J, \nu_J, \log \tau_1, \frac{\tau_2}{\tau_1}, \frac{\tau_3}{\tau_2}, \frac{\tau_4}{\tau_3}, (\lambda_j, \epsilon_j, \rho_j, \delta_j) |_{C/A} (\lambda_j, \epsilon_j, \rho_j) \right\}_{k_T} \\
D_C & \equiv \left\{ \log \theta_J, \nu_J \right\} \\
D_S & \equiv \left\{ \log \tau_1, \frac{\tau_2}{\tau_1}, \frac{\tau_3}{\tau_2}, \frac{\tau_4}{\tau_3}, (\lambda_j, \epsilon_j, \rho_j, \delta_j) |_{C/A} (\lambda_j, \epsilon_j, \rho_j) \right\}_{k_T},
\end{align}

where the subscripts $C/A$ or $k_T$ in Eqs. (10) and (12) imply that the observables are calculated using $C/A$ or $k_T$ subjets. The sets $D_C$ and $D_S$ consist of the conventional and the jet substructure variables respectively, whereas $D$ is the set of all variables.

In a previous paper \cite{52} we described the more conventional separation of single photons from QCD-jets along with an initial introduction to the separation of single photons from photon-jets. In both cases the single photons were treated as the signal. Here we extend that discussion and focus on the photon-jets as the signal. We organize the results of our analysis into three subsections. First, we show the results of BDTs optimized to discriminate QCD-jets from photon-jets, the analogue of the separation of single photons from QCD-jets. In the following subsection, we repeat the same study, but optimize it for treating single photons as the background to photon-jets. Finally, we demonstrate how the BDTs might be used for an effective three-way separation of single photons from photon-jets from QCD-jets.

D. QCD-Jets as Background for Photon-jets

We use all of the variables in the set of discriminants $D$ in the BDTs in order to maximize the extraction of signal jets ( photon-jets) from background (QCD-jets). This is similar to the separation of single photons from QCD-jets performed in Ref. \cite{52}. The BDTs are trained individually for each study point. The results for fake rate versus acceptance are shown in Fig. 9 for all of the study points. In this plot the lower right is desirable and the upper left is undesirable. Note that the acceptance rate for photon-jets is bounded above by about 0.94 due to our preselection cut $\theta_J \geq 0.25$ (see Section IV B 4). The same cut eliminates approximately 98% of the QCD-jets yielding a fake rate below $10^{-2}$ except at the largest acceptance.

For 2 photon-jets (study points 1 to 3) the separation becomes easier as $m_1$ increases yielding increas-
It is instructive to quantify the improvements made possible by including the jet substructure variables as discriminants. To achieve this comparison we consider BDTs using only the conventional variables (i.e., we use the set $D_C$ of discriminants to train the BDTs). For a given acceptance of signal we thus obtain two different fake rates — one when we use only the conventional variables (labeled $F_C$), and another when we use conventional+jet substructure variables (labeled $F_{C+S}$). For a given acceptance, the ratio $F_C/F_{C+S}$ quantifies the improvement due to using jet substructure variables in this analysis. The improvement rates for conventional plus jet substructure variables over only conventional variables versus acceptance for discriminating photon-jets from QCD-jets is shown in Fig. 10 for the different study points. While Figs. 9 and 10 indicate that the conventional variables provide some discrimination between photon-jets and QCD-jets, Figs. 11 to 15 indicate that the jet substructure variables provide a substantial number of new distinguishing features. Fig. 10 shows that these new features in the jet substructure variables can provide substantial improvement. Factors of 4 to 50 improvement in the discrimination of photon-jets from QCD-jets are possible at an acceptance of about 70%. As expected more improvement is possible in those physics scenarios where the photon-jets have more structure. Further, our results demonstrate that the use of jet substructure variables provides a tool to distinguish the different physics scenarios, i.e., the different study points, which is not possible with conventional variables alone.

E. Single Photons as Background to Photon-Jets

Now consider the same analysis as in the previous section but with single photons treated as the background. This new sort of separation is essential if we want to consider physics scenarios with photon-jets. Again we use all of the variables in the set of discriminants $D$ in the BDTs in order to maximize the extraction of signal jets (photon-jets) from background (single photons). The BDTs are trained individually for each study point. The results for fake rate versus acceptance are shown in Fig. 11. As in Fig. 9 the lower right is desirable and the upper left is undesirable. Again the acceptance rate for photon-jets is bounded above by about $0.94$ due to our preselection cut $\theta_J > 0.25$ (see Section IV.B.4). For the same reason a similar limit ($0.94$) holds also for the fake rate from single photons (although this is difficult to see on the logarithmic scale).

The results in Fig. 11 teach us several lessons. A photon-jet from PJSP 1 consists of a pair of highly collinear photons. Such a jet is quite photon-like and thus difficult to separate from single photons. Hence the corresponding (solid black) curve is most towards the upper left. One needs to cut away almost half of the signal sample ($A \sim 0.55$) in order to reduce the fake rate to 1 in $10^5$. We also see that it is a challenge to separate the photon-jets for PJSP 3 from single photons (the solid red curve). In this scenario $m_1 = 10$ GeV and the $n_1$ decays directly to two photons. Because of the large $m_1$ value, almost 30% of these ($R = 0, 4$) jets do not contain both of the photons from the $n_1$ decay, i.e., about 30% of this jet sample are actually single photons (in the jet), and not photon-jets. We saw this point earlier in essen-
FIG. 11. The background fake rate versus signal acceptance curves are shown for all study points. Here the photon-jets from the different study points are treated as the signal and single photons are the background. These curves employ all variables in the set of discriminants $D$.

Essentially all of the individual jet substructure variable plots, Figs. 4 to 8 where the PJSP 3 distribution exhibited an enhancement that overlapped with the corresponding peak in the single photon distribution. A larger separation of PJSP 3 from single photons can be obtained at an acceptance just below 0.7, where these single photons configurations are cut away and the fake rate drops below $10^{-3}$. The photon-jets of PJSP 2 represent a ‘sweet’ spot between PJSP1 and PJSP3 where the 2 photons are typically well enough separated to be resolved but close enough to be in the same jet. Thus the PJSP 2 (solid purple) curve is well below and to the right compared to the PJSP 1 (solid black) and PJSP 3 (solid red) curves. Similarly the photon-jets at the other study points can be well separated at even larger acceptance rates using the combination of jet substructure and conventional discriminants. For example, for the study points PJSP4 and PJSP6, even at 85% acceptance, one obtains a fake rate smaller than $10^{-3}$.

Again it is instructive to determine the impact of the jet substructure variables for this analysis. As in the previous subsection we consider BDTs using only the conventional variables (i.e., we use the set $D_C$ of discriminants to train the BDTs) to compare to the results from the full set $D$ of variables. We plot the ratio of fake rates at fixed acceptance for these two analyses in Fig. 12 versus the acceptance. A comparison of Fig. 12 and Fig. 11 indicates that the bulk of the separation of photon-jets from single photons is provided by the jet substructure variables, i.e., the improvement factor typically differs by less than a factor of 10 from one over the fake rate. Further, the improvement factor ranges from 10 to more that $10^3$ even at acceptances as large as 90% for all physics scenarios except PJSP1 and PJSP3. Even in these challenging cases substantial improvement is possible at lower acceptance rates. This is not a surprise since the conventional variables are ineffective at distinguishing between photon-jets and single photons. Recall from Fig. 2 that the hadronic energy fraction distributions are nearly identical for photon-jets and single photons. The distribution of the number of charge tracks associated with a jet, shown in Fig. 3 also indicates only slight differences, arising from the somewhat different conversion rates for photon-jets versus single photons. So it is clear that, if we want to be able to discriminate between photon-jets and single photons (and we do), the jet substructure variables provide the necessary tool.

**F. Three-way Separation**

Finally we come to the really interesting challenge: the simultaneous separation all three samples: single photons, photon-jets, and QCD-jets. In principle, one could perform three BDT training exercises, separating photon-jets from single photons, separating photon-jets from QCD-jets and separating single photons from QCD-jets. In principle, one could perform three BDT training exercises, separating photon-jets from single photons, separating photon-jets from QCD-jets and separating single photons from QCD-jets, using one of the variable sets of Eqs. 10 - 12 in each case. Then the responses from each of these BDTs for each jet could be used to separate the experimentally identified jets in the corresponding 3-dimensional ‘physics object’ space. In order to illustrate these ideas in a fairly simple analysis here we will focus on a two-dimensional analysis employing the two BDTs we have been discussing, separating photon-jets from single photons and separating photon-jets from QCD-jets. There are still the related questions of which set of variables to use for each BDT and, in fact, how to characterize the ‘best separation’ Qualitatively at least, we find good 2-dimensional separation for the following definitions of the

---

4 With three BDTs and the three BDT response numbers for each jet, the ‘best separation’ presumably corresponds to the three distinct physics objects being sent to three diagonally opposite vertices of the BDT response cube (on an equilateral triangle with side of length $\sqrt{2}$ times the length of the edge of the cube).
BDTs. One is trained to separate QCD-jets and photon-jets based only on the conventional discriminants ($D_C$) and is plotted on the vertical axis in the following plots, while the other BDT is trained to separate photon-jets from single photons with the substructure discriminants ($D_S$) alone and is plotted along the horizontal axis. We present the results in terms of two-dimensional contour plots where the numerical values associated with a given contour corresponds to the relative probability to find a calorimeter object of the given kind (indicated by the color) in a cell of size 0.1 × 0.1 in BDT response units. (Note that, by construction, the BDT responses have values in the range −1 to +1, where +1 means ‘signal-like’ and −1 means ‘background-like’.) The color coding in these figures matches the previous choices. Red is for QCD-jets, blue for photon-jets and green is for single photons.

As a first example, Fig. 13 indicates the 2-dimensional distributions resulting from the BDTs for PJSP 2, a scenario with typically two photons in the photon-jet with small angular separation due to the small value of $m_1$. When interpreting the following figures it is important to recall that the jet samples indicated in these figures are constrained to satisfy $\theta_J \leq 0.25$, which means that we are only keeping the approximately 2% of QCD-jets that are most ‘photon-like’. However, Fig. 13 indicates a pretty clear separation between the QCD-jets and the true photon objects (little red above 0.0 in the vertical direction). On the other hand, as we expect from our previous one-dimensional discussions in Subsection IV E, the blue (photon-jet) contours in the upper-left green (single photon) region indicate that it is a challenge to completely separate (PJSP 2) photon-jets from single photons.

In the case of PJSP 3 photon-jets, as indicated in Fig. 14, the photon-jet versus single photon separation challenge is even larger, as we have already discussed. Again we have photon-jets with potentially two photons but, due to the relatively large $m_1$ value, one of those photons is sometimes outside of the identified jet. This explains the small region with a solid blue (probability 0.1) contour inside the green (single photon) region.

The corresponding results for the more complex (and more easily separated) photon-jets of PJSP 4 and PJSP 8, typically with 4 photons in a photon-jet, are displayed in Figs. 15 and 16. In these scenarios the three-way photon-jet versus single photon versus QCD-jet separation is fairly cleanly achieved using just the $D_S$ (horizontal) and $D_C$ (vertical) variable sets. At the 0.005 level there is only a tiny overlap of photon-jets with QCD-jets for PJSP 4 (near the location (0.5,0.0) in Fig. 15) and no overlap for PJSP 8 (Fig. 16).

Before ending this section we should discuss one other point. From our previous discussion, one would expect to improve the photon-jet versus single photon separation by using the full $D$ set of variables (instead of the $D_S$ variables alone), and this expectation raises one of the interesting, and challenging, features of simultaneous separations. Since we are currently training the BDTs so that each BDT separates one type of signal from one type of background, while, at the same time, trying to perform a three-way separation, it can happen that an improvement in one separation corresponds to a degradation in another of the separations. To illustrate this point we first reproduce the results in Fig. 14 but now include a
contour at relative probability 0.001, which we did not include earlier to avoid plots that are too busy. The resulting plot is shown in Fig. 17. Now we perform the same analysis but using the full variable set $D$ in both BDTs. The resulting contour plot is displayed in Fig. 18, which illustrates the relevant points. The 0.001 level boundaries for single photons (green) and photon-jets (blue) are now somewhat better separated, although the effectively one-photon-jets from PJSP 3 (when one of the photons is very soft or is outside of the jet) still lie within the single photon boundary. At the same time, however, the separation between single photons (green) and the (typically more numerous) QCD-jets (red) is somewhat degraded (the green and red regions have moved towards each other). Due to the coupling between the different pairwise separations, optimizing such a three-way separation takes careful work and likely depends on the details of the actual analysis and detector.

These results clearly suggest that a three-way separation is possible, including the ability to distinguish different photon-jet scenarios. Further enhancement will arise from using the full 3-dimensional structure and from using a realistic detector simulation in the training. A thorough optimization in the context of a real detector and actual data may select different, more effective choices of the discriminating variables.

V. CONCLUSION

In this paper we have attempted to link several concepts, some conventional and some less so, with the goal of enhancing the searches for and the analyses of both Standard Model and Beyond the Standard Model physics. We advocate employing general techniques for analyzing and interpreting the detector objects identified by applying standard jet algorithms to the calorimeter cells of typical hadron collider detectors, allowing a universal language for such objects. We have demonstrated the efficacy of employing the recent developments in jet substructure techniques to separate and identify these
detector objects in terms of physics objects. Continuing the efforts begun in Ref. [21], we have focused on identifying three specific physics objects, the familiar single photons, QCD-jets and the Beyond the Standard Model (and LHC) relevant photon-jets. In particular, we have demonstrated that it is possible to achieve significant separation between photon-jets and their dominant backgrounds, i.e., single photons and QCD-jets. We expect that both the ATLAS and CMS groups could enhance their searches for signatures of new physics by adopting the methods described. These methods should allow the separation of photon-jets from single photons from QCD-jets, and also provide some identification of the specific dynamics yielding the photon-jets.

We note that our simulation does not take into account the impact of magnetic fields inside the detectors. On the other hand, one might interpret this absence of a magnetic field as making our results more conservative. When the magnetic field bends the electrons and positrons from converted photons, this serves to generate more structure inside the jet. The substructure variables, as we have described, tend to become more powerful with more structure. A more detailed analysis is, however, beyond the scope of this paper.

Finally, it is worth mentioning that the formalism and techniques developed in this paper for photon-jets should work in a similar way for the case of collinear electrons, often labeled ‘electron-jets’ [25, 26]. An electron-jet is characterized by a large number of charged tracks along with a small hadronic energy fraction. Also, we expect the electrons inside these jets to bend in a magnetic field, creating more substructure. Therefore we anticipate that multivariate analyses similar to those described here will be correspondingly effective at separating electron-jets from QCD-jets (and photon-jets).

ACKNOWLEDGEMENTS

The authors would like to acknowledge stimulating conversations with Henry Lubatti and Gordon Watts regarding the project. We especially thank Henry Lubatti for his careful reading of the manuscript. TSR would like to thank the hospitality of CERN, where part of the work was completed. The work of SDE, TSR and JS was supported, in part, by the US Department of Energy under contract numbers DE-FGO2-96ER40956. JS would also like to acknowledge partial support from a DOE High Energy Physics Graduate Theory Fellowship. Computing resources were provided by the University of Washington supported by the US National Science Foundation contract ARRA-NSF-0959141.

Appendix A: Technical Details

1. Conversions

In the material of the detector photons convert into electron-positron pairs. Our implementation of photon conversion is based on the properties of the ATLAS detector [59]. We associate an $\eta$ dependent probability of conversion with every photon. This probability is a function of the number of radiation lengths a photon passes through (labeled $n(\eta)$) in order to escape the first layer of the pixels. We model the probability to convert using the expression:

$$P(\eta) = 1 - \exp\left(-\frac{7}{9} n(\eta)\right), \quad (A1)$$

where the factor of $7/9$ comes from conversion between radiation length and mean free path. A plot of the extracted radiation length profile of the inner pixel detector in the ATLAS detector that we use to determine $n(\eta)$ is displayed in Fig. 19. The yellow-black dashed line shows...
the extracted $\eta$-dependent number-of-radiation-lengths a photon needs to travel before it exits the pixel layer. If the photon converts before the dashed line, it is treated as a charged track, otherwise no charged track is included in our simulation.

2. Details of the Calorimeters

Our simulation of calorimeters closely resembles the calorimeters implemented in the widely used simulation tool PGS [60]. The electromagnetic calorimeter ECAL covers $|\eta| \leq 2.5$ in the pseudorapidity direction, whereas the hadronic calorimeter HCAL covers the range $|\eta| \leq 5.0$. Both of the calorimeters provide a full $2\pi$ coverage in $\phi$, the azimuthal direction. The ECAL has granularity of $0.025 \times 0.025$ in the $\eta-\phi$ plane. The HCAL, on the other hand, has a coarser granularity with cells of size $0.1 \times 0.1$.

We assume that all particles generated at the interaction point (i.e., the Pythia output), except charged particles with energy less than 0.1 GeV, reach the ECAL. Photons/electrons within $|\eta| \leq 2.5$ deposit 99% of their energy in the ECAL and the rest in the HCAL. The ECAL also fully absorbs hadrons with energy less than 0.5 GeV. For more energetic hadrons, 0.5 GeV is deposited in the ECAL and the rest in the HCAL. Photons/electrons/hadrons that lie outside the pseudorapidity range of the ECAL but within the range of the HCAL (i.e., $5.0 \geq |\eta| > 2.5$) are completely absorbed in the HCAL.

In our simulation, muons also deposit energy in the calorimeters. Within the range $|\eta| \leq 2.5$, muons with $E < 0.5$ GeV deposit all of their energy in the ECAL; muons with $0.5$ GeV $< E < 2.5$ GeV deposit 0.5 GeV in the ECAL and the rest in the HCAL; muons with $E > 2.5$ GeV deposit 0.5 GeV in the ECAL and 2.0 GeV in the HCAL. For muons within the rapidity range $5.0 \geq |\eta| > 2.5$, those with $E < 2.0$ GeV are fully absorbed in the HCAL, while muons with $E > 2.0$ GeV deposit 2.0 GeV in the HCAL.

3. Moliere Matrix

In order to simulate the transverse smearing in the ECAL, we deposit the energy of each particle not just into the specific cell through which it passes, but also into the surrounding cells. The ($\eta-\phi$) coordinates of a particle determine the cell in which it deposits most of its energy (call it the $(i,j)$-th cell in the grid). We distribute its energy in the neighboring cells according to the following table:

| $i$ | $j$ | $E$ |
|-----|-----|-----|
| $i-1$ | 0.01(4) | 0.01(4) | 0.00(4) |
| $i$   | 0.01(4) | 0.92(4) | 0.01(4) |
| $i+1$ | 0.00(4) | 0.01(4) | 0.00(4) |

We estimate these numbers by integrating the energy deposited in an electromagnetic shower in a cell of size $(0.025 \times 0.025)$ and situated 1.5 m from the origin of the shower. We assume that the cell is made of lead ($R_M$(Pb) = 1.6 cm).

[1] G. Aad et al. (ATLAS Collaboration), Phys.Lett.B (2012), 1207.7214.
[2] S. Chatrchyan et al. (CMS Collaboration), Phys.Lett.B (2012), 1207.7235.
[3] B. A. Dobrescu, G. L. Landsberg, and K. T. Matchev, Phys.Rev. D63, 075003 (2001), hep-ph/0005308.
[4] N. Toro and I. Yavin (2012), 1202.6377.
[5] P. Draper and D. McKeen, Phys.Rev. D85, 115023 (2012), 1204.1061.
[6] Tech. Rep. ATLAS-CONF-2012-079, CERN, Geneva (2012).
[7] M. H. Seymour, Z.Phys. C62, 127 (1994).
[8] G. Brooijmans, Tech. Rep. ATIL-PHY-CONF-2008-008, ATL-COM-PHY-2008-001, CERN, Geneva (2008).
[9] J. Butterworth, J. R. Ellis, and A. Raklev, JHEP 0705, 033 (2007), hep-ph/0702150.
[10] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, Phys.Rev.Lett. 100, 242001 (2008), 0802.2470.
[11] J. Thaler and L.-T. Wang, JHEP 0807, 092 (2008), 0806.0023.
[12] D. E. Kaplan, K. Rehermann, M. D. Schwartz, and B. Tweedie, Phys.Rev.Lett. 101, 142001 (2008), 0806.0847.
[13] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, AIP Conf.Proc. 1078, 189 (2009), 0809.2530.
[14] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam (2008), 0810.0409.
[15] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, Phys.Rev. D80, 051501 (2009), 0903.5081.
[16] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, Phys.Rev. D81, 094023 (2010), 0912.0033.
[17] D. Krohn, J. Thaler, and L.-T. Wang, JHEP 1002, 084 (2010), 0912.1342.
[18] Tech. Rep. ATLAS-CONF-2012-065, CERN, Geneva (2012).
[19] Tech. Rep. CMS-PAS-EXO-11-095, CERN, Geneva (2012).
[20] T. Plehn, G. P. Salam, and M. Spannowsky, Phys.Rev.Lett. 104, 111801 (2010), 0910.5472.
[21] J. Gallicchio, J. Huth, M. Kagan, M. D. Schwartz, K. Black, et al., JHEP 113012 (2010), 1010.3698.
[22] C. Hackstein and M. Spannowsky, Phys.Rev. D81, 094023 (2010), 0912.0033.
[23] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, Phys.Rev. D82, 094023 (2010), 0912.0033.
[24] G. D. Kribs, A. Martin, T. S. Roy, and M. Spannowsky, Phys.Rev. D82, 095012 (2010), 1006.1656.
[25] G. D. Kribs, A. Martin, and T. S. Roy, Phys.Rev. D84, 095024 (2011), 1012.2866.
[26] A. Katz, M. Son, and B. Tweedie, Phys.Rev. **D83**, 114033 (2011), 1011.4523.

[27] C. Englert, T. S. Roy, and M. Spannowsky, Phys.Rev. **D84**, 075026 (2011), 1106.4545.

[28] M. Son, C. Spethmann, and B. Tweedie (2012), 1204.0525.

[29] A. Abdesselam, E. B. Kuutmann, U. Bitenc, G. Brooijmans, J. Butterworth, et al., Eur.Phys.J. **C71**, 1661 (2011), 1012.4523.

[30] A. Altheimer, S. Arora, L. Asquith, G. Brooijmans, J. Butterworth, et al., J.Phys.G **G39**, 063001 (2012), 1201.0008.

[31] D. Alves et al. (LHC New Physics Working Group) (2011), 1105.2838.

[32] R. Schabinger and J. D. Wells, Phys.Rev. **D72**, 093007 (2005), hep-ph/0509209.

[33] B. Patt and F. Wilczek (2006), hep-ph/0605188.

[34] M. J. Strassler and K. M. Zurek, Phys.Lett. **B651**, 374 (2007), hep-ph/0604261.

[35] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, JHEP **1106**, 128 (2011), 1106.0522.

[36] T. Sjostrand, S. Mrenna, and P. Z. Skands, JHEP **0605**, 026 (2006), hep-ph/0603175.

[37] T. Sjostrand, S. Mrenna, and P. Z. Skands, Comput.Phys.Commun. **178**, 852 (2008), 0710.3820.

[38] private correspondence with Henry Lubatti.

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

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

[41] M. Cacciari, G. P. Salam, and G. Soyez, JHEP **0804**, 063 (2008), 0802.1189.

[42] S. Catani, Y. L. Dokshitzer, M. Seymour, and B. Webber, Nucl.Phys. **B406**, 187 (1993).

[43] S. D. Ellis and D. E. Soper, Phys.Rev. **D48**, 3160 (1993), hep-ph/9305266.

[44] Y. L. Dokshitzer, G. Leder, S. Moretti, and B. Webber, JHEP **9708**, 001 (1997), hep-ph/9707323.

[45] M. Wobisch and T. Wengler (1998), hep-ph/9907280.

[46] M. Wobisch (2000).

[47] J. Thaler and K. Van Tilburg, JHEP **1103**, 015 (2011), 1011.2268.

[48] J. Thaler and K. Van Tilburg, JHEP **1202**, 093 (2012), 1108.2701.

[49] I. W. Stewart, F. J. Tackmann, and W. J. Waalewijn, Phys.Rev.Lett. **105**, 092002 (2010), 1004.2489.

[50] S. D. Ellis, A. Hornig, T. S. Roy, D. Krohn, and M. D. Schwartz, Phys.Rev.Lett. **108**, 182003 (2012), 1201.1914, URL http://jets.physics.harvard.edu/Jjets/html/Welcome.html.

[51] M. Cacciari, G. P. Salam, and G. Soyez, JHEP **0804**, 005 (2008), 0802.1188.

[52] Y. Freund and R. E. Schapire, pp. 148–156 (1996).

[53] A. Hocker, J. Stelzer, F. Tegenfeldt, H. Voss, K. Voss, et al., PoS **ACAT**, 040 (2007), physics/0703039.

[54] S. D. Ellis, T. S. Roy, and J. Scholz (2012), 1210.1855.

[55] J. T. Ruderman and T. Volansky, JHEP **1002**, 024 (2010), 0908.1570.

[56] C. Cheung, J. T. Ruderman, L.-T. Wang, and I. Yavin, JHEP **1004**, 116 (2010), 0909.0290.

[57] A. Falkowski, J. T. Ruderman, T. Volansky, and J. Zupan, JHEP **1005**, 077 (2010), 1002.2952.

[58] A. Falkowski, J. T. Ruderman, T. Volansky, and J. Zupan, Phys.Rev.Lett. **105**, 241801 (2010), 1007.3496.

[59] G. Aad et al. (ATLAS Collaboration), JINST **3**, S08003 (2008).

[60] J. Conway, R. Culbertson, R. Demina, B. Kilminster, M. Kruse, S. Mrenna, J. Nielsen, M. Roco, A. Pierce, J. Thaler, et al. (2012), URL http://www.physics.ucdavis.edu/~conway/research/software/pgs/pgs4-general.htm.