PHYSICAL REVIEW D **89,** 014007 (2014)

Tagging highly boosted top quarks

S. Schätzel[1] and M. Spannowsky[2]

[1]*Physikalisches Institut, Ruprecht-Karls-Universität Heidelberg, Germany*

[2]*Institute for Particle Physics Phenomenology, University of Durham, Durham DH1 3LE, United Kingdom*

(Received 21 August 2013; published 10 January 2014)

For highly energetic top quarks, the products of the decay $t \to bqq'$ are collimated. The three-prong decay structure can no longer be resolved using calorimeter information alone if the particle jet separation approaches the calorimeter granularity. We propose a new method, the HPTTopTagger, that uses tracks of charged particles inside a fat jet to find top quarks with transverse momentum $p_T > 1$ TeV. The tracking information is complemented by the calorimeter measurement of the fat jet energy to eliminate the sensitivity to jet-to-jet fluctuations in the charged-to-neutral particle ratio. We show that with the HPTTopTagger, a leptophobic narrow-mass Z' boson of mass 3 TeV could be found using 300 fb^{-1} of 14 TeV LHC data.

DOI: 10.1103/PhysRevD.89.014007 PACS numbers: 12.38.-t, 14.65.Ha

I. INTRODUCTION

After the recent discovery of the Higgs boson [1,2], the LHC's next foremost goal is to find evidence for physics beyond the Standard Model, i.e., new particles or forces. With a center-of-mass energy of 13 TeV, starting in spring 2015, the LHC experiments can access an unprecedented energy regime, allowing for the production of very heavy resonances. When heavy TeV-scale resonances decay to electroweak-scale particles (e.g., top quarks, W, Z, and Higgs bosons), these particles are boosted, and for central production, the particle's transverse momentum p_T exceeds its mass m. The decay products of these particles are then collimated in the laboratory frame. Due to the large branching ratio of electroweak-scale resonances into jets, it is beneficial, in many measurements and searches, to use jet substructure methods to disentangle the signal from large QCD backgrounds [3–8]. The reconstruction of intermediately ($2m \leq p_{\mathrm{T}} < 5m$) and highly ($P_{\mathrm{T}} \geq 5m$) boosted top quarks was one of the first motivations to study jet substructure techniques [9,10]. For an overview of these so-called "top-taggers," see [11].

Some of the most successful taggers are either based on jet-shape observables [12–15] or on sophisticated ways of selecting subjets inside a fat jet and comparing their energy sharing [16–20]. The separate identification of the decay products of highly boosted top quarks becomes experimentally challenging when the detector granularity does not allow us to resolve the individual particle jets.[1] This is particularly an issue if jets are reconstructed using calorimeter information alone, as it is currently done with ATLAS data. The cell size of the ATLAS barrel hadronic calorimeter is 0.1×0.1 in (η, ϕ), and topological cell clusters are formed around seed cells with an energy $|E_{\mathrm{cell}}| > 4\sigma_{\mathrm{noise}}$ by adding the neighboring cells with $|E_{\mathrm{cell}}| > 2\sigma_{\mathrm{noise}}$, and then all surrounding cells [22]. The minimal transverse size for a cluster of hadronic calorimeter cells is therefore 0.3×0.3 and is reached if all significant activity is concentrated in one cell. Two particle jets leave distinguishable clusters if each jet hits only a single cell and the jet axes are separated by at least $\Delta R = 0.2$, so that there is one empty cell between the two seed cells.[2] If two top quark decay jets are so close that they do not leave separate clusters, then top taggers based on identifying the three-prong decay structure will fail. Figure 1 shows the angular separation $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ of the two closest final state quarks in hadronic top quark decay $t \to bqq'$ as a function of the top quark p_T. For $p_T = 1.12\,\mathrm{TeV}$ the separation is 0.2, and calorimeter resolution issues should become apparent around that p_T or even earlier if the particle level jets that correspond to the quarks are not collimated enough to hit only a single cell. A precise determination of the threshold top quark p_T requires the use of the ATLAS simulation framework and hence has to be carried out by the ATLAS Collaboration.

Tracking detectors that measure the trajectories of charged particles can remedy the problem because of the much finer spatial resolution. In this article, we therefore propose a novel method of reconstructing highly boosted top quarks using a combination of tracks and calorimeter information. We compare the tagging efficiency of this high-p_T top tagger (HPTTopTagger) with the HEPTopTagger [19] and show that, with the HPTTopTagger, the LHC has a discovery reach for heavy resonances, which decay exclusively into top quarks, up to a resonance mass of 3 TeV with 300 fb^{-1} of data taken at $\sqrt{s} = 14$ TeV.

The article is arranged as follows: in Sec. II we introduce the top-tagging algorithm of the HPTTopTagger. In Sec. III

[1]For similar problems in decays of electroweak gauge bosons, see [21].

[2]A splitting algorithm has to be used in this case to divide this big cluster into two.

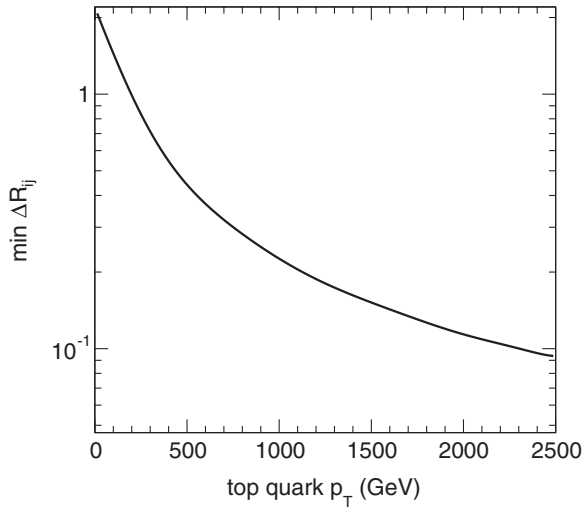


FIG. 1. Angular separation $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ of the two closest quarks in the top quark decay $t \to bqq'$ as a function of the top quark p_T.

we compare the tagging performance of the HPTTopTagger with that of the HEPTopTagger for highly boosted top quarks. We present the reach of the LHC in discovering very heavy resonances in Sec. IV and summarize our findings in Sec. V.

II. HIGHLY BOOSTED TOP QUARK RECONSTRUCTION

The HEPTopTagger uses a mass-drop criterion and adaptive filtering to obtain three subjets that are tested for kinematic compatibility with hadronic top quark decay. These conditions are formulated in the form of ratios of invariant mass combinations of the subjets. For example, the mass m_{23} is defined as the invariant mass of the subleading p_T and the sub-subleading $\boldsymbol{p_T}$ subjet. For most hadronic top quark decays, the ratio m_{23}/m_{123} corresponds to m_W/m_t with the b-jet having the largest p_T. The invariant masses are determined from the 4-momenta of the subjets which have to be reconstructed precisely. The ATLAS Collaboration has calculated calibration constants for the HEPTopTagger subjets and uncertainties that quantify to which precision the subjet energy scale and energy resolution can be modeled in simulation [23]. These uncertainties are crucial for comparing the data to simulated model predictions and for setting exclusion limits, as has been done, for example, in [24]. Calibrations of ATLAS Cambridge/Aachen (C/A) jets [25] are available for radius parameters R from 1.2 down to 0.2 [23]. Jets with a smaller radius parameter approach the minimal hadronic cluster size, as discussed in Sec. I.

A tracking detector can reconstruct the trajectory of a charged particle and can specify the direction of the particle at any point of the trajectory with a precision much better than the granularity of the calorimeter. For example, the angular resolution of the ATLAS inner tracking detector for charged particles with $p_T = 10$ GeV and $\eta = 0.25$ is $\approx 10^{-3}$ in η and ≈ 0.3 mrad in ϕ [26] with a reconstruction efficiency of $> 78\%$ for tracks of charged particles with $p_T > 500$ MeV [27]. The momentum resolution for charged pions is 4% for momenta $p < 10$ GeV, rising to 18% at $p = 100$ GeV [26].

Prong-based tagging algorithms usually require the reconstruction of the top quark and W boson masses to identify a fat jet as being initiated by a top quark decay. In a typical proton-proton collision, about 65% of the jet energy is carried by charged hadrons, 25% by photons, produced mainly from π^0 decays, and only 10% by neutral hadrons (mostly neutrons and K_L^0) [28]. However, these fractions can vary significantly from event to event. Thus, reconstructing the correct resonance mass is a challenging task for a tagging algorithm which is based exclusively on tracks. Fortunately, while no calibrations exist for subjets with $R < 0.2$, the energy of the fat jet can be calibrated to good precision [23] and the inverse of the energy fraction carried by charged tracks

$$\alpha_j = \frac{E_{\text{jet}}}{E_{\text{tracks}}} \tag{1}$$

can be measured for each jet individually, thereby eliminating the sensitivity to fluctuations to a large extent.

Our tagger for highly boosted top quarks uses elements of the HEPTopTagger which do not introduce artificial mass scales in background events; i.e., we do not consider all possible three-subject combinations until we find a top-like structure. Such drastic measures might be necessary when the small boost of the top quark requires us to use a very large jet cone to capture all decay products. In the case of a highly boosted top quark, the decay products are confined to a small area of the detector, and the amount of additional radiation inside a C/A jet with $R = 0.8$ is usually not excessive.[3]

To reconstruct highly boosted top quarks, we propose the following procedure, labeled for later reference as the HPTTopTagger algorithm.

(1) Define a jet j using the C/A algorithm with $R = 0.8$ from calorimeter clusters.
(2) Take the tracks with $p_T > 500$ MeV that are associated with j and recombine them into a track-based jet j_c.
(3) Calculate α_j of Eq. (1) using j and j_c.
(4) Apply the mass drop procedure introduced in [19]: undo the last clustering of the track-based jet j_c into two subjets j_{c1}, j_{c2} with $m_{j_{c1}} > m_{j_{c2}}$. We require $m_{j_{c1}} < 0.8 m_{j_c}$ to keep j_{c1} and j_{c2}. If this condition

[3]The amount of additional radiation in a fat jet strongly depends on the cone size [29], but also on the overall hadronic activity of the event and on the color flow of the underlying hard interaction [30].

does not hold, we keep only j_{c1}. Each subjet j_{ci} we further decompose unless $m_{j_{ci}} < 20$ GeV. The remaining subjets we add to the list of relevant substructures.

(5) If we find fewer than two remaining subjets, we consider the tag to have failed. Else, we take the constituents of all subjets surviving the mass drop procedure and multiply their momenta by α_j each.

(6) We take all the rescaled constituents and filter them with resolution $R_{\text{filt}} = \max(0.05, \min(\Delta R_{ij}/2))$, in which i and j run over all remaining subjets after the mass drop procedure. We recombine all constituents of the four hardest filtered subjets and require the resulting invariant mass to be in a mass window around the top quark mass. We call this object our top candidate.

(7) Again we follow the HEPTopTagger and construct exactly three p_T ordered subjets j_1, j_2, j_3 from the top candidate's constituents. If the masses (m_{12}, m_{13}, m_{23}) satisfy the so-called A-cut of [19], we consider the top tag to be successful.

One guiding principle of the outlined algorithm is not to bias the mass distribution for the top and W candidates. Particularly at high p_T, splittings of massless quarks and gluons can geometrically induce a large jet mass $m_j^2 \sim p_{T,j}^2 \Delta R_{j_1,j_2}^2$ [31]. Depending on the jet p_T cut in the event, the average jet mass can be much bigger than the top quark mass. Therefore, looking explicitly for structures in very hard jets which fulfill simple mass requirements can result in a large fake rate.

The sensitivity to fluctuations in the fraction of charged particles is reduced by scaling the subjet momenta by α_j. This procedure relies on the assumption that the energy fraction carried by charged particles is similar in all subjets. Figure 2 compares the particle level subjet α, i.e., the inverse of the energy fraction carried by charged particles with $p_T > 500$ MeV and $|\eta| < 2.5$ inside a subjet, to the particle level α_j calculated from the fat jet. The events used for this figure contain decays of the Z' boson of Sec. IV to $t\bar{t}$ with $m_{Z'} = 3$ TeV. The distributions are very similar for $m_{Z'} = 5$ TeV. For most of the leading p_T subjets (subjets 1), the ratio is ≈ 0.95, whereas for most of the subjets 2 (subleading p_T subjets) and subjets 3, the ratios are ≈ 0.9 and ≈ 0.8, respectively. The fat jet quantity α_j is therefore a good approximation to the subjet α. We note that the subjet α cannot be calculated at the detector level because no calibrations exist for the small calorimeter subjets ($R < 0.2$) we are interested in.

In step 2 we use only tracks for j_c. By including photons measured in a finely grained electromagnetic calorimeter in addition to the tracks, it is possible to obtain a richer jet substructure and a better energy resolution of the rescaled track jet, i.e., a better mass reconstruction of the top candidate. This enhancement is currently not implemented because the fast detector simulation we are using

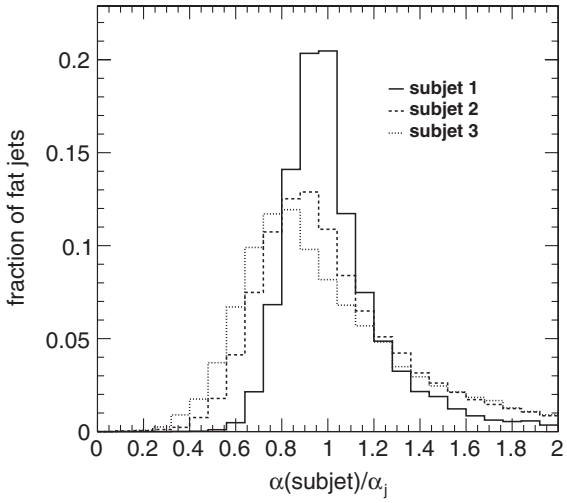


FIG. 2. The ratio $\alpha(\text{subjet})/\alpha_j$ for the three subjets found by the HPTTopTagger in fat jets in events with $Z' \to t\bar{t}$ decays where $m_{Z'} = 3$ TeV. Subjet 1 is the leading p_T subjet and subjet 2 the subleading p_T subjet.

applies the same segmentation to the electromagnetic and hadronic parts of the calorimeter. In Fig. 3 we illustrate the impact of adding photon information. Shown in panel (a) is the fat jet energy fraction carried by charged particles (with $p_T > 500$ MeV) and photons and by the charged particles alone. The distribution is wider in the latter case because of fluctuations in the photon fraction. The effect on the reconstructed top mass is shown in panel (b). Here the distribution obtained at the particle level when applying the above prescription and using only charged particles in step 2 is compared to the two cases in which charged particles and photons or all particles are used. The impact of the fluctuations in the charged-to-neutral particle ratio are already significantly reduced when adding photon information, leading to a narrower mass distribution. This better top quark momentum reconstruction will also improve the $t\bar{t}$ resonance mass resolution. The sensitivity to the charged-to-neutral fluctuations can be reduced by choosing a large enough top quark mass window. With a window from 140 to 210 GeV, we see only a small efficiency increase of a couple of percent when adding photons.

It is well known that track-based observables are not infrared safe [32]. Therefore, nonperturbative contributions have to be taken into account to obtain finite and well-defined results. In full event generators like PYTHIA, hadronization models, including fragmentation functions, are used. These functions are nonperturbative objects following perturbative evolution equations and are usually fitted to LEP data. The fraction of charged particles is unknown within limits imposed by measurements of the jet fragmentation function [33]. In Sec. III we compare the top quark reconstruction efficiency for the HPTTopTagger between HERWIG++ and PYTHIA 8.

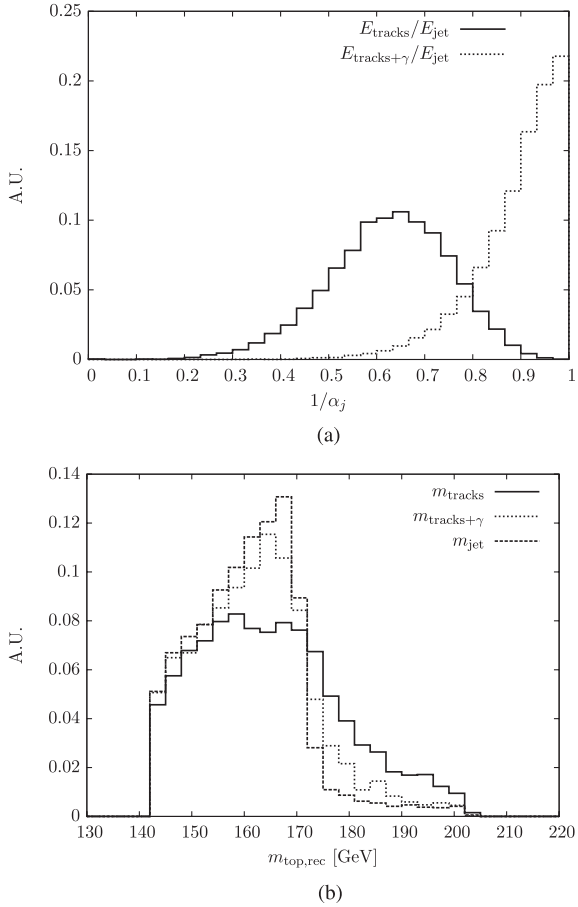


FIG. 3. (a) The fat jet energy fraction carried by charged particles with $p_T > 500$ MeV ($E_{\text{tracks}}/E_{\text{jet}}$) and by charged particles with $p_T > 500$ MeV plus photons. (b) The top quark candidate mass reconstructed at the particle level using the HPTTopTagger as defined in Sec. II (m_{tracks}) and when adding photons ($m_{\text{tracks}+\gamma}$) or all particles (m_{jet}) in step 2.

Although the cone size for highly boosted top quarks does not need to be big, we find that using filtering [5,34] in step 6 improves the performance of the tagger in separating top jets from QCD jets. Our goal is to achieve a flat tagging efficiency independent of the top quark's transverse momentum. Thus, to decide if a tag was successful we use invariant masses in step 7.

III. PERFORMANCE OF TOP-TAGGING ALGORITHMS

We use PYTHIA 8.175 [35] to obtain fully showered and hadronized final states of Standard Model $t\bar{t}$ and dijet production, as well as events with hypothetical leptophobic Z' bosons. The Delphes program [36] is used in version 2.0.3 to obtain a fast simulation of the response of a LHC detector. We use the Delphes ATLAS detector card with a tracking efficiency of 78%. Ultimately, at very high top quark p_T, the track reconstruction will struggle to resolve the tracks left by nearby particles. The limit is reached when the hits are so close that they are part of the same reconstructed track, and the track reconstruction efficiency suffers as a consequence. The ATLAS tracking efficiency is 80% for $p_T = 500$ MeV and rises to 86% for $p_T > 10$ GeV [27,37]. To take the close-by effect into account, we have used a reduced tracking efficiency of 78% which corresponds to a 10% relative loss of efficiency at high p_T. We assume that this efficiency is a conservative lower limit and treat it as constant in p_T. A careful study of the tracking efficiency as a function of p_T is needed but can only be done with access to the full detector simulation. This is beyond the scope of this article.

The smallest simulated calorimeter entities are calorimeter cells, which for $|\eta| \leq 2.5$ have a size 0.1×0.17 in (η, ϕ) (and double this ϕ size for $|\eta| > 2.5$). No clustering of these cells is performed, leading to smaller calorimeter structures than those that would be available with the real ATLAS calorimeter. The resolution power of the calorimeter is therefore overestimated, and the impact of tracking information will be larger in reality. For jet finding we use the FastJet [38] program.

C/A $R = 0.8$ jets with $p_T > 800$ GeV are built from calorimeter cells and are required to lie within $|\eta| < 2.5$. Tracks with $p_T > 500$ GeV and $|\eta| < 2.5$ are matched to these jets using ghost association [39,40] as follows. A ghost of every track is created by setting the p_T to a small value (10 eV) and using the tracks η and ϕ at the calorimeter surface. The energy of the ghost is set to 1.001 times its momentum to ensure a positive ghost mass. The ghost tracks are added to the calorimeter jet clustering but do not change the jet because their energy is negligible. If the ghost track ends up in the jet then the original track is taken to be associated with the jet. We then cluster all associated tracks into a C/A jet. The calorimeter jet and the track jet serve as inputs j and j_c to the HPTTopTagger procedure defined in Sec. II.

The HEPTopTagger as proposed in [19] was designed to work for mildly boosted top quarks, which required a large radius parameter of $R = 1.5$ to geometrically catch the decay products. For the reconstruction of highly boosted top quarks, we use as inputs to the HEPTopTagger the calorimeter $R = 0.8$ jets to compare directly to the HPTTopTagger but note that the HEPTopTagger is optimized to achieve a high rejection of background that is picked up by using the large radius. We also modify the original algorithm by stopping the mass drop procedure already if the subjet mass is below 50 GeV (originally 30 GeV) because this was the preferred value in [24].

ATLAS subjet calibrations and subjet simulation uncertainties exist for radius parameters down to $R = 0.2$ [23]. To demonstrate the performance of the HEPTopTagger if only those jets were to be used, we implement the following changes to the original algorithm and refer to the modified tagger as HEPTopTagger' in the following:

(i) The minimal filter radius is set to 0.2.

(ii) Each exclusively clustered subjet is required to have $R' > 0.2$ in which R' is given by the distance in (η, ϕ) space between the jet axis and the jet constituent farthest away from the axis.

(iii) The filtered subjets and the exclusive subjets must have $p_T > 20$ GeV.

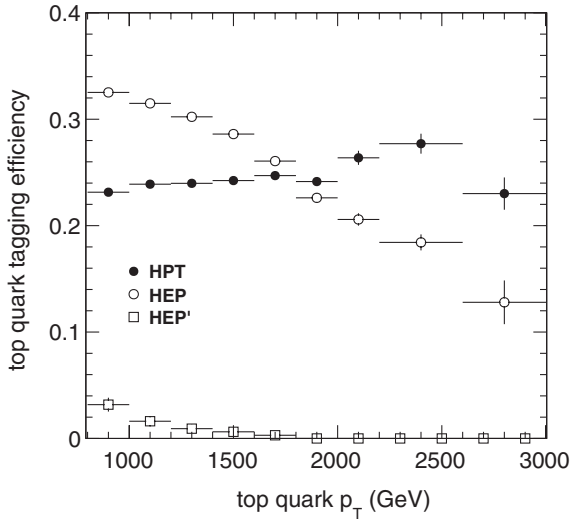
(iv) We stop the mass drop procedure if the two parent jets are closer than $\Delta R = 0.2$.

The same calorimeter fat jets are fed to the three different taggers for $t\bar{t}$ (signal) and light quark or gluon dijet events (background). The top quark tagging efficiency is shown in Fig. 4(a) as a function of the top quark p_T. A top quark is taken to be tagged if a reconstructed top quark candidate is found within $\Delta R = 0.6$ of the top quark. The HPTTopTagger efficiency is stable at $\approx 24\%$ up to 3 TeV in p_T. If the tracking efficiency is artificially set to 100%, then the efficiency rises to 28%. This rather small sensitivity to the tracking inefficiency is explained by the fact that the energy in lost tracks is recovered by the α_j scaling.

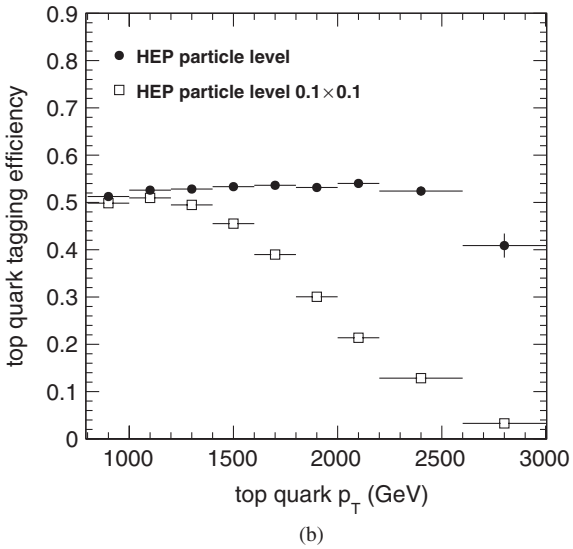
For the HEPTopTagger, the efficiency drops from $\approx 32\%$ for $800 < p_T < 1000$ GeV to $\approx 13\%$ for $2600 < p_T < 3000$ GeV due to the segmentation of the calorimeter which prevents all three top quark decay particle jets from being reconstructed. To prove this claim, we show in Fig. 4(b) the top quark finding efficiency for the HEPTopTagger at the particle level. When the constituents of the C/A $R = 0.8$ jets are stable particles, the HEPTopTagger efficiency is stable at 53%. If we granularize the particles into (η, ϕ) cells of size 0.1×0.1, the efficiency starts to drop at a top quark p_T of 1.2 TeV.

The efficiency of the HEPTopTagger' for finding top quarks using calorimeter cells is less than 4% for $p_T > 800$ GeV [Fig. 4(a)]. From this we conclude that with the present available ATLAS jet calibrations and uncertainties, it is not possible to find top quarks at hig p_T. To obtain calibrations and uncertainties also for jets with $R < 0.2$ we suggest the use of the reconstructed top mass peak in $t\bar{t}$ events. The position of the peak can be used for calibration, and the difference between simulation and data can serve to estimate the simulation uncertainty. We note that at higher top quark p_T, the fraction of subjets with small R will be higher. This effect can be studied by binning the mass distribution in p_T of the top candidate.

The efficiency for tagging fat jets constructed from calorimeter cells is shown in Fig. 5 as a function of the fat jet p_T. For $t\bar{t}$ events, the numerical values are similar to the top quark tagging efficiencies. The fake rate, defined as the probability to tag fat jets originating from light quarks or gluons, is stable at 1.6% for the HPTTopTagger, while it increases for the HEPTopTagger from $\approx 2\%$ for



(a)



(b)

FIG. 4. Efficiencies for tagging top quarks using (a) calorimeter cells and (b) stable particles.

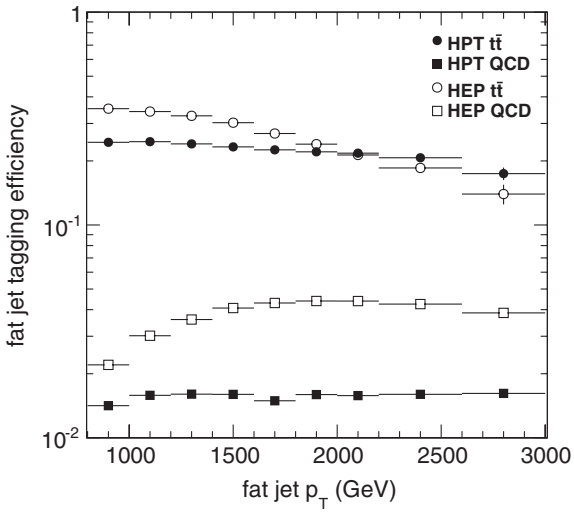
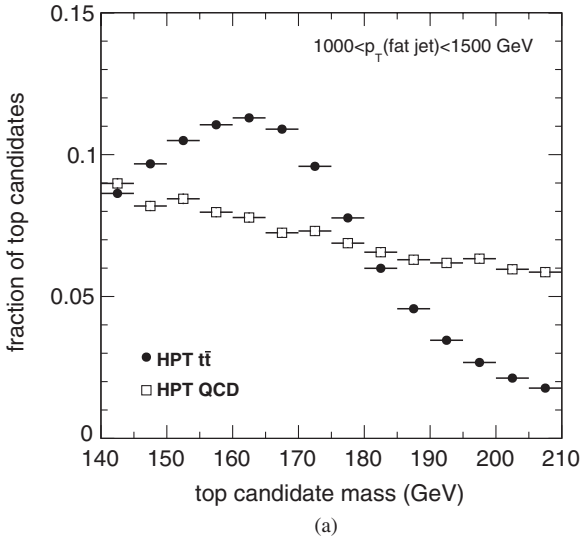


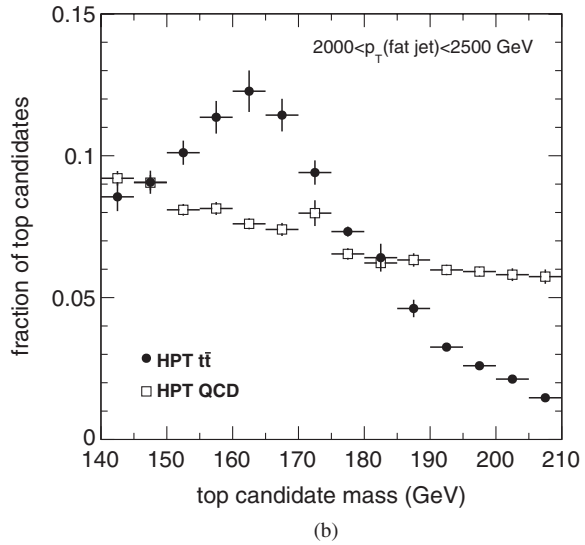
FIG. 5. Efficiencies for tagging C/A $R = 0.8$ calorimeter fat jets.

$p_T = 800$ GeV to 4.5% for $p_T = 2$ TeV. Because of the comparable signal efficiency and the much lower fake rate, the HPTTopTagger outperforms the HEPTopTagger when the resonance search strategy requires an improvement on the signal-to-background ratio. The fake rate is smaller with the HPTTopTagger because not all possible three-subject combinations are tried when looking for a toplike structure. This is different in the HEPTopTagger where all triplets of substructure objects (the objects after the mass drop) are tested for compatibility with top quark decay, which increases the efficiency but also the fake rate.

The top quark mass reconstructed with the HPTTopTagger is shown in Fig. 6 in two bins of the calorimeter fat jet p_T. While a clear peak is visible for events with top quarks, the background distribution is smoothly falling and shows no shaping into a peak. This holds true for low and high fat jet p_T.



(a)



(b)

FIG. 6. Top quark mass reconstructed with the HPTTopTagger in two bins of the calorimeter fat jet p_T.

TABLE I. The efficiency for tagging fat particle jets with $p_T > 1.2$ TeV for two samples of $t\bar{t}$ events with top quark $p_T > 1$ TeV.

Efficiency for tagging	PYTHIA 8 (%)	HERWIG++ 2.5 (%)
The leading p_T fat jet	27.0(5)	28.2(5)
The subleading p_T fat jet	18.1(4)	18.6(5)
Both fat jets	4.9(1)	5.3(2)

The HPTTopTagger sensitivity to the imperfect knowledge of charged particle production, i.e., the hadronization model, is small. Measurements of the jet fragmentation function and comparisons with different generators have been reported in [33]. The difference between the string model [41] based in PYTHIA 8 and the cluster model [42] based in HERWIG++ 2.5 [43] gives a conservative estimate of the difference in charged particle production. The efficiency for tagging fat particle jets with $p_T > 1.2$ TeV is shown in Table I for two samples of $t\bar{t}$ events with top quark $p_T > 1$ TeV, one generated with PYTHIA 8 and one with HERWIG++ 2.5. The efficiencies are compatible within the relative statistical uncertainty of 3%. The top quark candidate mass, reconstructed at the particle level, is compared in Fig. 7. The average mass from HERWIG++ is larger by only 1.7 GeV.

IV. RECONSTRUCTING HEAVY RESONANCES AT THE LHC

In the following we discuss top taggers in the context of detecting a leptophobic topcolor Z' boson that decays to two top quarks [44]. The width of the resonance is set to $\Gamma_{Z'}/m_{Z'} = 3.2\%$. We choose two mass points, $m_{Z'} = 3$ TeV and $m_{Z'} = 5$ TeV, for which the production

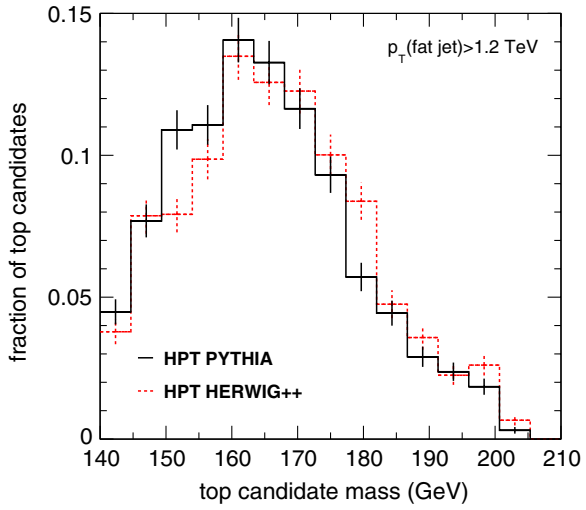


FIG. 7 (color online). Top quark mass reconstructed at the particle level with the HPTTopTagger using $t\bar{t}$ events generated with PYTHIA and HERWIG++ for fat jets with $p_T > 1.2$ TeV.

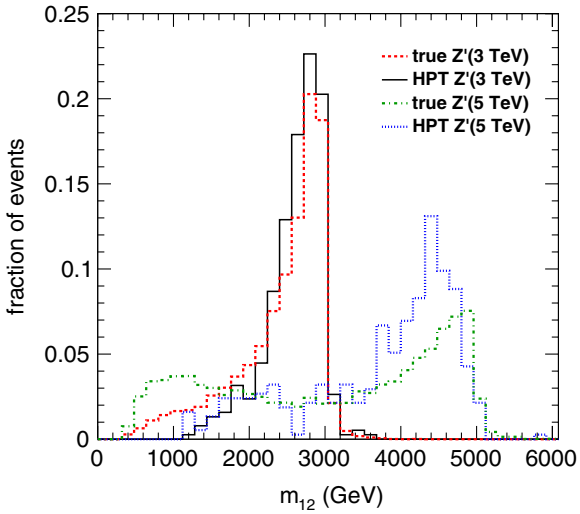


FIG. 8 (color online). Invariant mass of the two top quarks from the decay $Z' \to t\bar{t}$ (" true Z' ") after QCD radiation for two Z' masses and the corresponding reconstructed distributions when using the HPTTopTagger.

cross sections in pp collisions at $\sqrt{s} = 14$ TeV are 3.5 fb and 97 ab, respectively.

Top taggers imposing requirements on the top quark mass and/or the W boson mass reconstruct on-shell top quarks right before they decay. Therefore, radiation off the top quark, while necessary to reconstruct the Z' resonance, is discarded. Particularly for heavy Z' bosons, which result in highly boosted top quarks, gluon radiation is not unlikely, as can be seen from Fig. 8 which shows the invariant mass of the two top quarks after QCD radiation. Events with $m_{12} \leq 4$ TeV amount to 2/3 of the production cross section of a 5 TeV Z' boson. Those events require a refined reconstruction strategy beyond the simple double top tag discussed here.

The reconstructed ditop invariant mass is shown in Fig. 9 for the Z' signal and QCD dijet production, which constitutes the most important background ($t\bar{t}$ production is smaller by a factor of ≈0.1) for 300 fb^{-1} of pp collisions at a center-of-mass energy of 14 TeV. The only imposed requirement is that there be at least two tagged C/A $R = 0.8$ jets in the event. The plotted quantity m_{12} is the invariant mass of the two leading p_T top quark candidates. Based on the expected position of the signal, we have defined mass windows, in which we compare the number of signal (S) and background (B) events. For the 3 TeV Z' boson, we find a signal-to-background ratio $S/B = 0.45(7)$ and a significance $S/\sqrt{B} = 4.1(4)$ in the window $2560 < m_{12} < 3040$ GeV. The discovery of such a Z' boson with the HPTTopTagger is therefore within reach. The uncertainties are statistical and dominated by the finite number of simulated background events. For comparison, with the same generated events, the significance when using the HEPTopTagger is only 3.3(3) and the difference

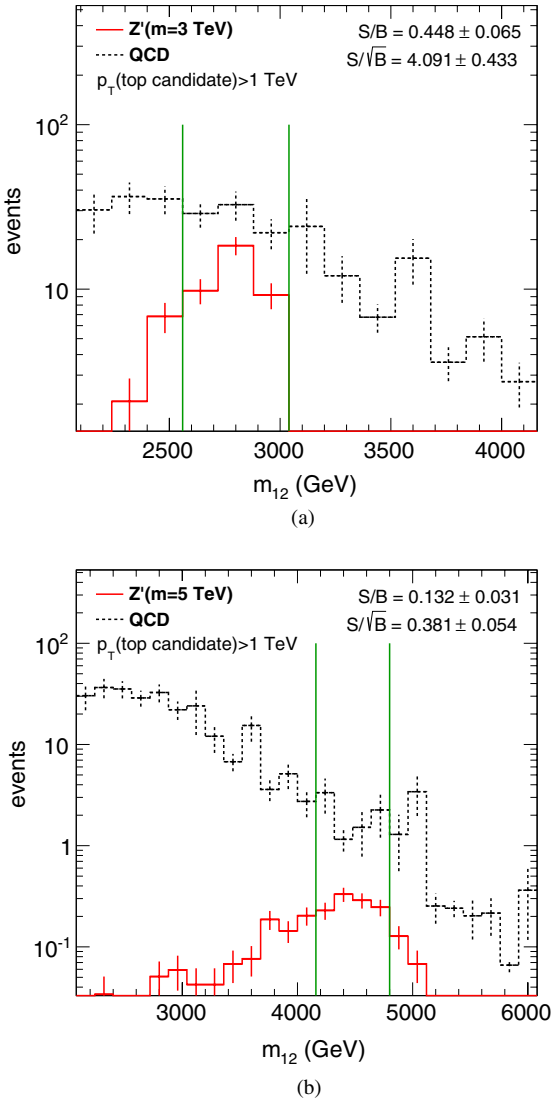


FIG. 9 (color online). Invariant ditop mass, reconstructed with the HPTTopTagger, from 300 fb^{-1} of decays of Z' bosons of mass (a) 3 TeV and (b) 5 TeV, produced in pp collisions at $\sqrt{s} = 14$ TeV. Also shown is the background from QCD dijet production. The signal-to-noise ratio S/B and the significance $S/\sqrt{B}$ are given for the indicated mass window.

to the HPTTopTagger results directly from the different fat jet tagging efficiencies shown in Fig. 5. There is no sensitivity to a 5 TeV Z' boson, with $S/B = 0.13(3)$ and $S/\sqrt{B} = 0.38(5)$ in the window $4160 < m_{12} < 4800$ GeV, because the background level is too high. The sensitivity might be improved by applying b-quark tagging if the related systematic uncertainties are small enough at high p_T.

V. SUMMARY AND OUTLOOK

Traditional top quark finding algorithms, that are based on identifying the three-prong hadronic decay structure, fail when the decay products can no longer be resolved. For

calorimeter granularities of 0.1×0.1, we find this merging of particle jets to start at top quark transverse momenta of ≈ 1.2 TeV.

We propose the HPTTopTagger, a new algorithm to find boosted top quarks with transverse momentum $p_T > 1$ TeV, that combines track and calorimeter information. The finer spatial resolution of tracking detectors allows the separation of close-by particle jets that would merge in the calorimeter. We have shown that with the HPTTopTagger, a Z' boson of mass 3 TeV is within discovery reach when using 300 fb^{-1} of 14 TeV LHC data.

Including photons, measured in a finely grained electromagnetic calorimeter, in the HPTTopTagger algorithm could improve the performance. For heavy resonances, the effect of QCD radiation off the top quarks becomes important and a possible way to reconstruct the resonance mass is to include jets in addition to the two top quark candidates.

While the HPTTopTagger has a smaller tagging efficiency for top quarks with $p_T < 1$ TeV compared to standard tagging approaches using subjets, it performs better for highly boosted top quarks. We point out that it is straightforward to combine the HPTTopTagger approach with any subjet-based top tagger, particularly with the HEPTopTagger, to obtain an improved tagging performance over a large p_T range.

ACKNOWLEDGMENTS

This work was funded in the UK by STFC. We thank the Perimeter Institute and the University of Oregon for hospitality. We thank Simone Marzani for valuable discussions.

[1] G. Aad *et al.* (ATLAS Collaboration), Phys. Lett. B **716**, 1 (2012).
[2] CMS Collaboration, Phys. Lett. B **716**, 30 (2012).
[3] M. H. Seymour, Z. Phys. C **62**, 127 (1994).
[4] J. M. Butterworth, B. E. Cox, and J. R. Forshaw, Phys. Rev. D **65**, 096014 (2002).
[5] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, Phys. Rev. Lett. **100**, 242001 (2008).
[6] A. Abdesselam, E. B. Kuutmann, U. Bitenc, G. Brooijmans, J. Butterworth, P. Bruckman de Renstrom, D. Buarque Franzosi, R. Buckingham *et al.*, Eur. Phys. J. C **71**, 1661 (2011).
[7] L. G. Almeida, R. Alon, and M. Spannowsky, Eur. Phys. J. C **72**, 2113 (2012).
[8] A. Altheimer, S. Arora, L. Asquith, G. Brooijmans, J. Butterworth, M. Campanelli, B. Chapleau, A. E. Cholakian *et al.*, J. Phys. G **39**, 063001 (2012).
[9] M. Gerbush, T. J. Khoo, D. J. Phalen, A. Pierce, and D. Tucker-Smith, Phys. Rev. D **77**, 095003 (2008); U. Baur and L. H. Orr, *ibid.* **77**, 114001 (2008); P. Fileviez Perez, R. Gavin, T. McElmurry, and F. Petriello, *ibid.* **78**, 115017 (2008).
[10] K. Agashe, A. Belyaev, T. Krupovnickas, G. Perez, and J. Virzi, Phys. Rev. D **77**, 015003 (2008); B. Lillie, L. Randall, and L.-T. Wang, J. High Energy Phys. 09 (**2007**) 074; U. Baur and L. H. Orr, Phys. Rev. D **76**, 094012 (2007); V. Barger, T. Han, and D. G. E. Walker, Phys. Rev. Lett. **100**, 031801 (2008).
[11] T. Plehn and M. Spannowsky, J. Phys. G **39**, 083001 (2012).
[12] J. Thaler and L.-T. Wang, J. High Energy Phys. 07 (**2008**) 092.
[13] L. G. Almeida, S. J. Lee, G. Perez, I. Sung, and J. Virzi, Phys. Rev. D **79**, 074012 (2009).
[14] J. Thaler and K. Van Tilburg, J. High Energy Phys. 03 (**2011**) 015; J. High Energy Phys.02 (**2012**) 093.
[15] M. Jankowiak and A. J. Larkoski, J. High Energy Phys. 06 (**2011**) 057.
[16] D. E. Kaplan, K. Rehermann, M. D. Schwartz, and B. Tweedie, Phys. Rev. Lett. **101**, 142001 (2008).
[17] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, Phys. Rev. D **80**, 051501 (2009).
[18] CMS Collaboration, Report No. CMS-PAS-JME-09-001.
[19] T. Plehn, G. P. Salam, and M. Spannowsky, Phys. Rev. Lett. **104**, 111801 (2010); T. Plehn, M. Spannowsky, M. Takeuchi, and D. Zerwas, J. High Energy Phys. 10 (**2010**) 078; http://www.thphys.uni-heidelberg.de/~plehn.
[20] D. E. Soper and M. Spannowsky, Phys. Rev. D **84**, 074002 (2011); Phys. Rev. D**87**, 054012 (2013).
[21] A. Katz, M. Son, and B. Tweedie, J. High Energy Phys. 03 (**2011**) 011.
[22] G. Aad *et al.* (ATLAS Collaboration), Eur. Phys. J. C **73**, 2305 (2013).
[23] G. Aad *et al.* (ATLAS Collaboration), J. High Energy Phys. 09 (**2013**) 076.
[24] G. Aad *et al.* (ATLAS Collaboration), J. High Energy Phys. 01 (**2013**) 116.
[25] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, J. High Energy Phys. 08 (**1997**) 001; M. Wobisch and T. Wengler, arXiv:hep-ph/9907280.
[26] G. Aad *et al.* (ATLAS Collaboration), JINST **3**, S08003 (2008).
[27] G. Aad *et al.* (ATLAS Collaboration), New J. Phys. **13**, 053033 (2011).
[28] CMS Collaboration, Report No. CMS-PAS-PFT-10-001, Report No. CMS-PAS-PFT-10-002.
[29] M. Dasgupta, L. Magnea, and G. P. Salam, J. High Energy Phys. 02 (**2008**) 055; M. Cacciari, G. P. Salam, and S. Sapeta, *ibid.* 04 (**2010**) 065.
[30] S. Ask, J. H. Collins, J. R. Forshaw, K. Joshi, and A. D. Pilkington, J. High Energy Phys. 01 (**2012**) 018; K. Joshi, A. D. Pilkington, and M. Spannowsky, Phys. Rev. D **86**, 114016 (2012).

[31] S. D. Ellis, J. Huston, K. Hatakeyama, P. Loch, and M. Tonnesmann, Prog. Part. Nucl. Phys. **60**, 484 (2008); S. D. Ellis, A. Hornig, C. Lee, C. K. Vermilion, and J. R. Walsh, Phys. Lett. B **689**, 82 (2010); M. Dasgupta, K. Khelifa-Kerfa, S. Marzani, and M. Spannowsky, J. High Energy Phys. 10 (**2012**) 126.

[32] W. J. Waalewijn, Phys. Rev. D **86**, 094030 (2012); D. Krohn, M. D. Schwartz, T. Lin, and W. J. Waalewijn, Phys. Rev. Lett. **110**, 212001 (2013); H.-M. Chang, M. Procura, J. Thaler, and W. J. Waalewijn, *ibid.* **111**, 102002 (2013); Phys. Rev. D **88**, 034030 (2013); A. J. Larkoski and J. Thaler, J. High Energy Phys. 09 (**2013**) 137.

[33] G. Aad *et al.* (ATLAS Collaboration), Eur. Phys. J. C **71**, 1795 (2011).

[34] M. Dasgupta, A. Fregoso, S. Marzani, and G. P. Salam, J. High Energy Phys. 09 (**2013**) 029.

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

[36] S. Ovyn, X. Rouby, and V. Lemaitre, arXiv:0903.2225.

[37] G. Aad *et al.* (ATLAS Collaboration), Phys. Lett. B **688**, 21 (2010).

[38] M. Cacciari and G. P. Salam, Phys. Lett. B **641**, 57 (2006); M. Cacciari, G. P. Salam, and G. Soyez, Eur. Phys. J. C **72**, 1896 (2012); http://fastjet.fr.

[39] M. Cacciari and G. P. Salam, Phys. Lett. B **659**, 119 (2008).

[40] M. Cacciari, G. P. Salam, and G. Soyez, J. High Energy Phys. 04 (**2008**) 005.

[41] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, Phys. Rep. **97**, 31 (1983).

[42] B. R. Webber, Nucl. Phys. **B238**, 492 (1984).

[43] M. Bahr, S. Gieseke, M. A. Gigg, D. Grellscheid, K. Hamilton, O. Latunde-Dada, S. Platzer, P. Richardson *et al.*, Eur. Phys. J. C **58**, 639 (2008).

[44] R. M. Harris, C. T. Hill, and S. J. Parke, arXiv:hep-ph/9911288.