PHYSICAL REVIEW D **85,** 034029 (2012)

How to improve top-quark tagging

Tilman Plehn,[1] Michael Spannowsky,[2] and Michihisa Takeuchi[1]
[1]*Institut für Theoretische Physik, Universität Heidelberg, Germany*
[2]*Institute for Particle Physics Phenomenology, Department of Physics, Durham University, DH1 3LE, United Kingdom*
(Received 2 December 2011; published 16 February 2012)

In time for the first tests on LHC data we introduce a set of improvements and tests of purely kinematic top tagging algorithms. First, we show how different jet algorithms can be used for different transverse momentum regimes. Combining pruning and filtering in the reconstruction can enhance the signal over background ratio significantly, while larger jet radii only give minor improvements. Finally, bottom tagging can be added to the top tagger, but at least for the HEPTOPTAGGER does not improve the kinematic selection algorithm.

DOI: 10.1103/PhysRevD.85.034029 PACS numbers: 14.65.Ha

I. INTRODUCTION

The top quark, found in 1995 [1], is the heaviest and so far only observed fermion with a weak-scale mass. Therefore, it is expected to have strong ties to the mechanism which triggers electroweak symmetry breaking. Searches for new physics in the top sector are of high priority because they can shed additional light on the structure of the standard model (SM) at and above the weak scale. Many extensions of the standard model, like supersymmetry or little Higgs models [2], predict top partners to ameliorate the loop-induced effect of the top quark on the Higgs boson's mass. Typical signatures for such extended top sectors include top partners decaying to a top quark and missing energy [3–6] or heavy resonances decaying to two often strongly boosted top quarks [7,8].

During the Tevatron's final years D0 and CDF have measured several anomalies directly related to the top sector: in CDF's single top analysis the ratio between the s- and t-channel production rates deviates by 2.5σ from the SM prediction [9], and both experiments measure an enhanced $t\bar{t}$ forward-backward asymmetry compared to SM predictions [10]. Also including the excess in the dijet invariant mass spectrum of the Wjj final state, measured (only) by CDF [11], all these anomalies show that we need an improved understanding and simulation of top production processes [12].

The top pair production rate at the LHC ranges around 1×10^6 tops per inverse femtobarn of integrated luminosity. On the one hand, this means that top pair production is a very challenging background for searches relying on high multiplicity final states of jets, leptons, and missing energy [3–5,13]. On the other hand, this means that already now we can test top pair events in many different kinematic regimes. In this paper we will focus on moderately boosted top quarks in the semileptonic decay channel.

While the idea of studying the substructure of jets is already a classic [14], the potential for searches of massive Higgs [15–17] and gauge bosons [18] or new physics [19] has only been appreciated recently [20]. In this paper we will focus on tagging boosted top quarks [5,21–29]. Aside from being sensitive probes of new physics they are also the prime candidates to generally establish that fat-jet or subjet methods work at the LHC. Some very promising ATLAS results on the HEPTOPTAGGER [5,22] performance on data can be found in Ref. [30]. CMS has already released first search results using a top tagger [31]. Of the Tevatron anomalies listed above the top forward-backward or charge asymmetry is particularly interesting in the light of boosted top quarks; the ratio of initial state quarks vs gluons increases in the boosted regime, thereby enhancing the otherwise small asymmetry at the LHC [32].

Starting from the default purely kinematic setup of the HEPTOPTAGGER we investigate several avenues on how to improve its performance: In Sec. II we discuss how well the momentum of subjets matches the decay partons in the default setup and which strategies for an improvement should be promising. In Sec. III we investigate the performance of different jet algorithms for the filtering and subjet reconstruction. In Sec. IV we then study the tagging performance if we include pruning in combination with filtering. The pruned top mass we use as an additional kinematic variable according to Ref. [33]. In Sec. V we investigate the possibility to enlarge the size of the fat jet to $R = 1.8$, focusing on the currently most relevant low-p_T tops. Finally, in Sec. VI we augment the kinematic top tagger by a b tag inside the fat jet [34]. Two possible strategies are simply adding the b tag at the end of the top tagging algorithm or including it in a modified extraction of the relevant subjets.

II. SUBJET-PARTON RECONSTRUCTION

Before we can suggest and test improvements to our top tagger it is crucial that we study measures for the quality of the top reconstruction. The geometrical distance between the reconstructed and the true top momenta is simply

$$\Delta R^2_{\text{top}} = \Delta R^2(p^{\text{tagged}}_{\text{top}}, p^{\text{parton}}_{\text{top}}). \tag{1}$$

For a more detailed study we also compute the geometric separation of the top decay products, which requires a proper definition of the parton and jet level constituents. Jet combinatorics is the main challenge, in particular, for hadronic top pair production at the LHC. For example, including up to two additional hard QCD jets the $t\bar{t}$ sample consists of 6 to 8 partons which we label as (1,4) for the bottom quarks, (2,5) for the harder W decay partons, (3,6) for the softer W decay partons, and (7,8,...) for additionally radiated partons. The partons 1–3 and 4–6 come from one top decay each. After hadronization and jet reconstruction the corresponding b, W_1, and W_2 subjets are defined such that W_1 (harder) and W_2 (softer) reconstruct m_W best. Note that we do not apply any b tagging, an issue we will look at in Sec. VI. We can then define

$$\Delta R^2_{\text{sum}} = \min_{\text{mappings}} \sum_{i=1}^{3} \Delta R^2(p_i^{\text{sub}}, p_{j_i}^{\text{parton}}), \tag{2}$$

where the label j_i denotes the ith hardest parton in the tagged top. The best parton-subjet mapping $\{j_i\} = \{j_1, j_2, j_3\}$ is defined by the minimum $\sum_{i=1}^{3} \Delta R^2(p_i^{\text{sub}}, p_{j_i}^{\text{parton}})$ value. For example, for $\{j_i\} = \{1, 5, 7\}$ the hardest subjet corresponds to a b quark, the second hardest to a W decay from the other top, and the softest an additional parton from jet radiations. This way we can categorize all tagged tops into three types:

(i) Type 1: $\{j_1, j_2, j_3\}$ come from one top decay, i.e., $\{1, 2, 3\}$ or $\{4, 5, 6\}$.
(ii) Type 2: only the two hardest $\{j_1, j_2\}$ come from one top decay; j_3 has a different origin.
(iii) Type 3: other.

For semileptonic top pairs the combinatorics is simplified significantly, but we can still categorize all tagged tops along the same lines. Our analysis is based on ALPGEN-PYTHIA [35,36] samples with MLM merging [37] ($R^{\text{MLM}} = 0.4$, $p_T^{\text{MLM}} > 30$ GeV) and we cluster the visible final state using FASTJET [38]. We take into account two hard jets in association with $t\bar{t}$ production and three to five hard jets for W + jets and QCD jets. The $t\bar{t}$ sample we reweight to 918 pb [39]. The left panel of Fig. 1 shows ΔR_{sum} for each type. The quality of the reconstructed subjets direction is the same for all types, with the exception of long type 2 and type 3 tails. In the central and right panels of Fig. 1 we test the actual top momentum reconstruction in terms of ΔR_{top} and $\Delta p_T^{\text{top}}/p_T^{\text{top}} = (p_{T,t}^{\text{tagged}} - p_{T,t}^{\text{parton}})/p_{T,t}^{\text{tagged}}$. Its quality depends on the different types of parton identification and the most poorly reconstructed candidate tops are of type 2 and type 3. Unlike for type 1 tops their distributions do not follow a largely Gaussian shape centered at zero but show a significant shift.

Figure 2 shows the transverse momentum difference between each subjet and the corresponding parton $\Delta p_T/p_T = (p_T^{\text{subjet}} - p_T^{\text{parton}})/p_T^{\text{subjet}}$. We see that the subjet momentum reconstruction is essentially of the same quality for all subjets and for all types, with the exception of p_T^{W1} which is better reconstructed than p_T^{W2} because of its larger value.

In Table I we give the fraction of tagged tops for each of the three types after different requirements on the quality of the reconstruction. For type 1 tags the momentum reconstruction both for subjets and tops is almost perfect. For type 2 and type 3 tags, a large fraction of tagged tops satisfies $\Delta R_{\text{sum}} < 0.4$, which means the individual subjets are reconstructed well but the set of the partons is wrongly picked. Consequently, the top momentum reconstruction for type 3 tags becomes worse. Thanks to a correct assignment for the hardest two subjets in type 2 tags the top momentum reconstruction is not too bad because the wrong third subjet does not contribute much to the top momentum.

From the discussion above we can conclude that the individual subjet-parton momentum reconstruction works well for all types. The limitation to the top momentum reconstruction arises from events where some of the identified subjets do not correspond to a top decay product. From Table I we estimate that $\mathcal{O}(20\%)$ of type 3 tops and $\mathcal{O}(50\%)$ of type 2 tops still give the correct top momentum within 15%. The fraction of tagged tops with good momentum reconstruction within each type does not depend much on $p_{T,t}$; however, the fraction of type 2 and type 3

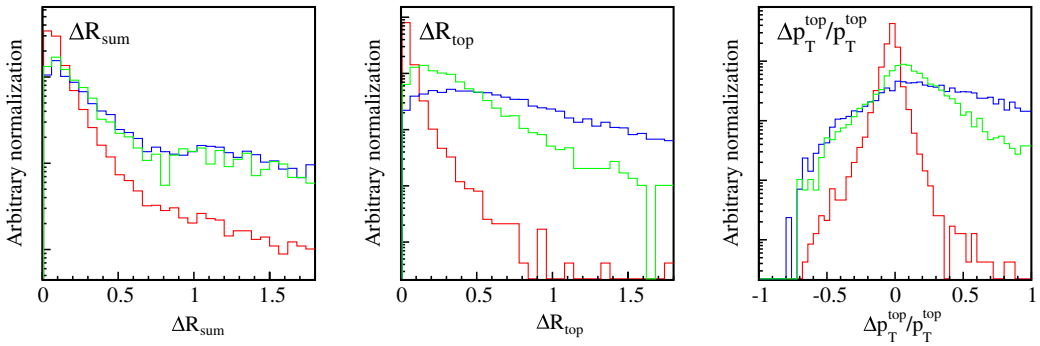

FIG. 1 (color online). Left: ΔR_{sum} as defined in Eq. (2). Center: ΔR_{top} as defined in Eq. (1). Right: $\Delta p_T^{\text{top}}/p_T$. The lines represent type 1 [medium (red)], type 2 [light (green)], type 3 [dark (blue)] for semileptonic top pairs at a collider energy of 14 TeV.

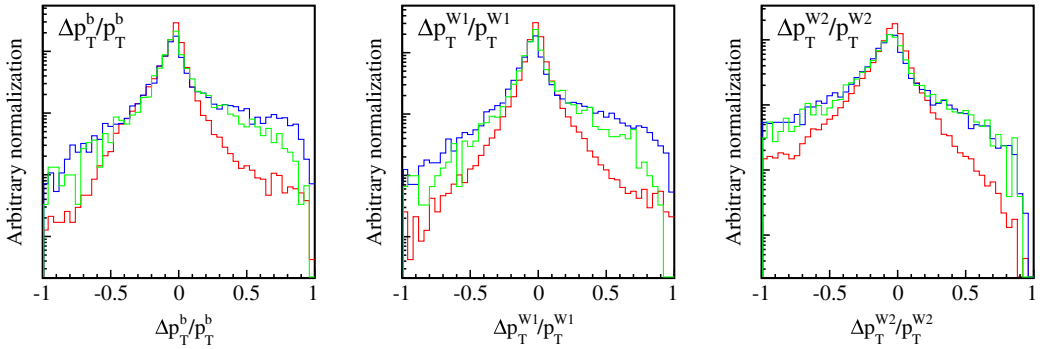

FIG. 2 (color online). Relative p_T differences between subjets and the corresponding partons $\Delta p_T/p_T$ for the b subjet (left), the W_1 subjet (center), and the W_2 subjet (right). Again, the lines show type 1 [medium (red)], type 2 [light (green)], and type 3 [dark (blue)].

tags in all tagged tops decreases for higher p_T and effectively leads to a better momentum reconstruction. In total, 85% of all tagged tops and up to 90% of all tags with $p_T^{\text{tagged}} > 250$ GeV reproduce the true momentum within a 15% error bar.

An assigned top tagging efficiency should describe which fraction of hadronic tops in any event sample is tagged. Such an efficiency we define step by step:

(1) All decay products satisfy $R_{\text{C/A}} < 1.5$.
(2) All decay products appear in a fat jet; i.e., there exist unfiltered subjets with $\Delta R(p^{\text{parton}}, p^{\text{subjet}}) < 0.4$.
(3) All decay products appear in a fat jet with a top candidate fulfilling $150 < m_{jjj}^{\text{filter}} < 200$ GeV.
(4) All decay products appear in a fat jet with a tagged top, i.e., after the mass plane cut.

For the first step we use $R_{\text{C/A}} = \max\{R_1, R_2\}$ based on the two Cambridge-Aachen (C/A) distance measures [40] for the necessary clustering steps $R_{1,2}$. For the third step we apply a mass drop criterion and extract filtered jet masses for the relevant subjet combinations [5,15] For the last step we rely on the jet mass plane cuts which the HEPTOPTAGGER algorithm is based on and which are described in detail in Ref. [5]. All these conditions are defined for signal events, so we can show them as a function of the true $p_{T,t}$ at parton level.

The left panel of Fig. 3 shows the event fractions corresponding to all four categories normalized by the number of hadronic tops as functions of p_T. First, the dotted entries show the fraction of tops with $R_{\text{C/A}} < 1.5$, while the solid entries show the fraction of top decay products inside a fat jet. For any fat-jet analysis they fix an upper bound on all tagging efficiencies.

The lighter (red) dotted and solid entries show the fraction of candidate and tagged top events, all constrained to type 1 tags. The difference between the two is simply given by the mass plane cuts. After these cuts roughly 30% of all hadronic tops are tagged above $p_T \sim 250$ GeV. Note that there exist type 2 and type 3 candidates and tags, i.e., hadronic tops whose decay products are included in a fat jet but the extracted subjets do not fall into type 1.

The right panel of Fig. 3 first shows the fraction of type 1 candidates and type 1 tags relative to the number of fat jets including all three top decay products, corresponding to the second category in the above list. Types 2 and type 3 are shown as the lightest (green) and darkest (blue) entries. The difference between all top decay products inside a fat jet and type 1 candidates is about 25% and almost constant for $200 < p_T < 500$ GeV. It is partly (at most 10%) due to the existence of type 2 or 3 candidates, which means the tagger wrongly selects subjets even though the fat jet does include all decay products. Alternatively, there might be overlapping of subjets such that any three subjets are inconsistent with the top mass constraint. The difference between candidates and tags corresponds to the mass plane cuts, where some signal loss is inevitable for rejecting QCD and W + jets backgrounds. The numbers of type 2 and type 3 tags are negligible; most of the type 2 and type 3

TABLE I. Tagged top rates (in fb) in the semileptonic $t\bar{t}$ sample for 14 TeV collider energy. The percentages for the additional cuts are relative to the numbers of tagged tops in each row.

	All p_T^{tagged}				$p_T^{\text{tagged}} > 250$ GeV			
	Tagged	$\Delta R_{\text{sum}} < 0.4$	$\Delta R_{\text{top}} < 0.2$	$\lvert\frac{\Delta p_T^{\text{top}}}{p_T}\rvert < 0.15$	Tagged	$\Delta R_{\text{sum}} < 0.4$	$\Delta R_{\text{top}} < 0.2$	$\lvert\frac{\Delta p_T^{\text{top}}}{p_T}\rvert < 0.15$
Total	14 156	11 904 (84%)	10 841 (77%)	11 037 (78%)	6029	5279 (88%)	5191 (86%)	5170 (86%)
Type 1	10 318	9531 (92%)	10 102 (98%)	9897 (96%)	4919	4624 (94%)	4858 (99%)	4774 (97%)
Type 2	1336	896 (67%)	477 (36%)	623 (47%)	412	273 (66%)	218 (53%)	244 (59%)
Type 3	2503	1478 (59%)	263 (11%)	517 (21%)	698	381 (55%)	115 (16%)	152 (22%)

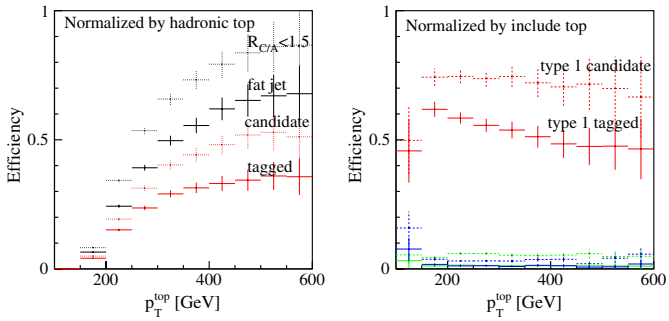

FIG. 3 (color online). Left: tagging efficiencies normalized to the number of hadronic tops as a function of p_T. Right: tagging efficiencies relative to the number of hadronic tops included in a fat jet. The different curves are discussed in the text.

tags shown in Table I come from fat jets not including a top. The mass plane cuts efficiently remove such contributions. Hence, we see that once the top decay products are captured in a fat jet, 50% to 60% of the tops are tagged.

The rapid drop of efficiency below $p_T \sim 250$ GeV in the left panel of Fig. 3 is simply due to the rapid drop of the fraction of decay products within $R_{C/A} < 1.5$ and inside the fat jet. The ratio of type 1 tags to decay products within a fat jet is not small for $150 < p_T < 250$ GeV. Therefore, we will test an increased R value in Sec. V.

An obvious question from the right panel of Fig. 3 is why the tagging efficiency degrades for larger p_T, where in principle the tagging performance should tend to increase. This effect of the mass plane cut is caused by a misreclustering after filtering and is more pronounced for the C/A algorithm. The C/A algorithm relies on the R distance exclusively; therefore the softest two of the five filtered subjets are not necessarily combined according to their shower history. As a result, we find unbalanced invariant masses from the three reclustered subjets. This happens more frequently in the high-p_T regime where all five subjets are not well separated, so the rejection probability by the mass plane cut increases with p_T. This tendency and the possibility of changing the underlying jet algorithm and its effect on mass plane cuts we will discuss in Sec. III.

To include backgrounds and mistagging efficiencies we need to define a slightly different set of scenarios, namely, as a function of the fat jet p_T. We define four scenarios similar to the ones before, but now in terms of fat jets:

(1) Fat jets with three subjets or more after the mass drop criteria.
(2) Fat jets where all distances between top decay products and their closest subjets are less than 0.4.
(3) Fat jets with a top candidate.
(4) Fat jets with a top tag.

The left panel of Fig. 4 shows all corresponding fractions relative to the number of fat jets as a function of p_T. The black dotted and solid symbols show the fraction of events with at least three subjets and of all top decay products included. More than half of the fat-jet events with at least three subjets do not include the top decay products, even for the semileptonic $t\bar{t}$ sample. The medium (red) entries show the fraction of candidates (dotted) and tags (solid). Because all efficiencies are shown as a function of the fat jet p_T, we can also show the backgrounds in dark (blue) and light (green) symbols. On the plateau we find tagged tops in roughly 20% of the fat jets for semileptonic top pairs and 2%–4% for W + jets and QCD jets.

The candidate histogram in the left panel of Fig. 4 exceeds the numbers for top decay products in the fat jet because there exist fat jets whose three main subjets are not from a top decay and accidentally give m_t. The central panel of Fig. 4 shows the composition of each type separately for candidates (dotted) and tags (solid). Indeed, a considerable fraction of candidates are of type 3. On the other hand, type 3 and type 2 tags are effectively rejected by the mass plane cut, so most of the tagged tops are of type 1. The fraction of type 2 and type 3 tags ranges around 2%–4%, similar to QCD and W + jets backgrounds.

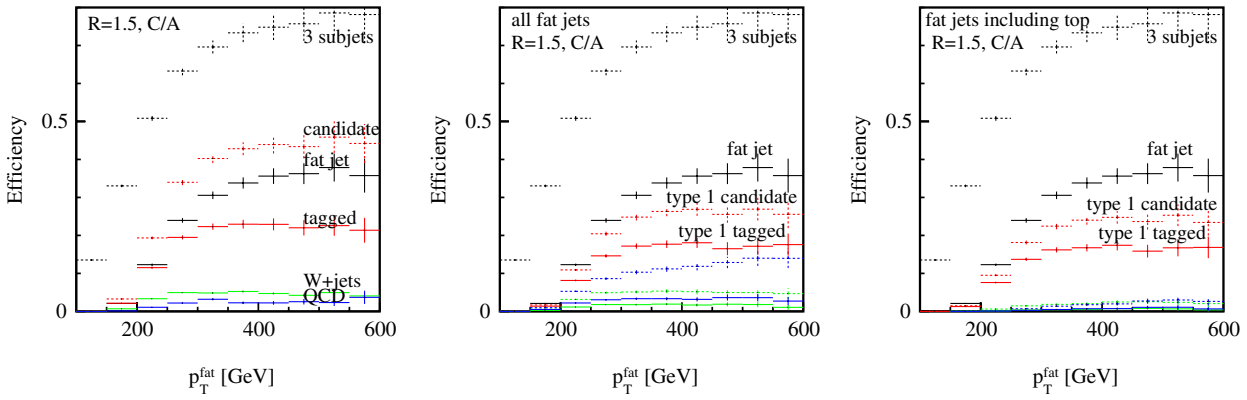

FIG. 4 (color online). Left: efficiencies $\epsilon_{t\bar{t}}$, $\epsilon_{W+\text{jets}}$, ϵ_{QCD} as functions of the fat jet p_T. Center: fraction of tagged tops for type 1 [medium (red) crosses], type 2 [light (green)], and type 3 [dark (blue)]. The dotted lines show the corresponding candidate fractions. Right: fraction of type 1, type 2, and type 3 only for fat jets including a top.

The right panel shows the fractions of candidate and tagged fat jets including a hadronic top, so they are also constrained to belong to the second category. Compared with the central panel we are now less likely to encounter type 2 or type 3 candidates. Most tagged tops are now of type 1 and most type 2 and type 3 tags correspond to fat jets which do not include a hadronic top. Consequently, we find that there is not much room to improve our algorithm in selecting subjets after applying the mass drop criterion.

III. ALTERNATIVE JET ALGORITHMS

The combination of the C/A clustering algorithm [40] with a mass drop criterion [15] is a core feature of the HEPTOPTAGGER and hence not negotiable. However, in the mass reconstruction after filtering the C/A algorithm should be compared to alternatives, like the k_T [38,41] or anti-k_T [42] jet algorithms. After identifying three subjets based on the mass drop criteria there are two steps left to extract the b, W_1, and W_2 subjets: filtering [15] and reclustering. We find that the choice of jet algorithm for the filtering does not have a visible effect on the efficiency at particle level, while for the reclustering step it does. In our explanations we therefore focus on effects of this reclustering which combines five filtered subjets into three top decay subjets while we always use the same jet algorithm for filtering and for reclustering.

Our first result is that the anti-k_T algorithm fails to reliably identify the three hard top decay products. It tends to first recombine a pairing with large transverse momentum, such that of the three reclustered subjets one is very hard and two are very soft. Applying our W and top mass cuts will typically reject such unbalanced combinations.

Using the C/A algorithm a similar problem arises, but only for very large $p_{T,t}$, where the five filtered subjets are close. According to QCD the two softest filtered subjets should then be merged into the main three subjets, which the C/A algorithm achieves as long as the three main subjets are geometrically well separated. Once the geometric distances become small the probability to correctly reconstruct the three main subjets decreases. Such signal events appear in the lower left corner of the ($\arctan m_{13}/m_{12}$ vs m_{23}/m_{123}) distributions [5] shown in Fig. 5. Events close to the x axis have $m_{23} \sim 0$ which usually means $p_3 \sim 0$ for the third-hardest decay subjet.

The k_T algorithm recombines soft filtered subjets most reliably, so it can resolve the three main top decay products best. We can see this in the lower panels of Fig. 5 where hardly any signal events migrate to $m_{23} \sim 0$. As a result, the signal efficiency of passing the mass plane cut increases. This tagging efficiency we show in Fig. 6. For both cases the filtering and reclustering algorithms are the same. In the low-p_T regime the difference between the two algorithms is indeed small. In the central panel we see that unlike the C/A results in Fig. 3 the k_T efficiency hardly

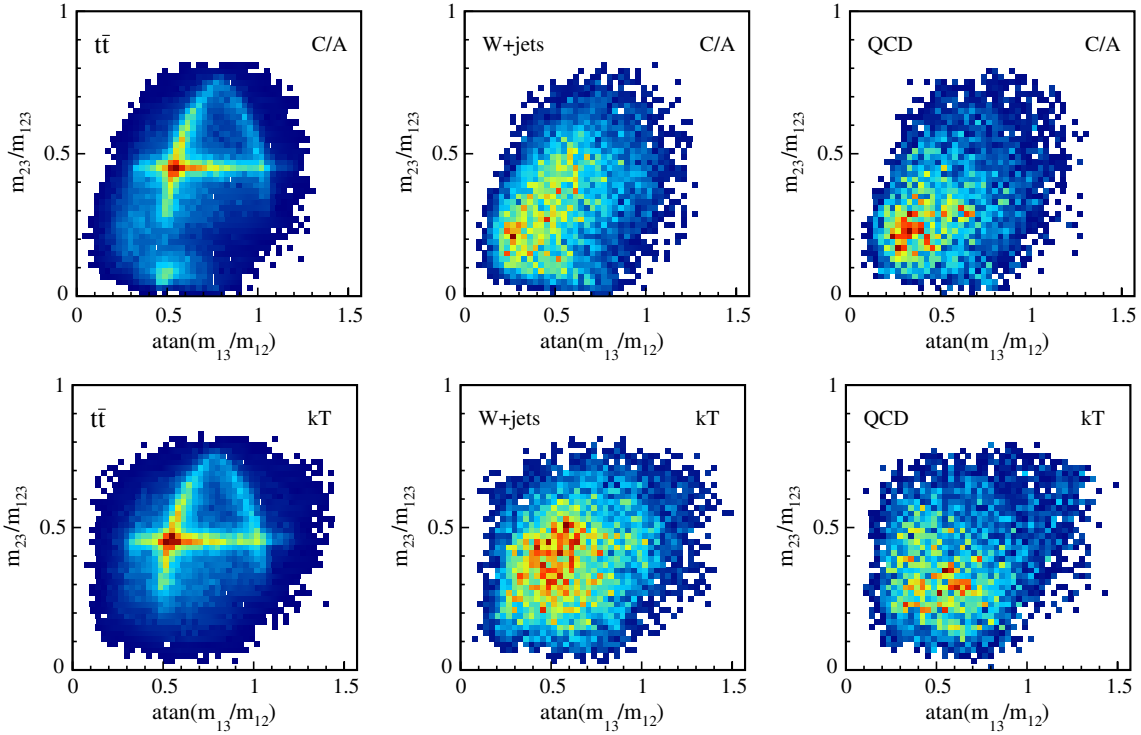

FIG. 5 (color online). Mass plane for top candidates in $t\bar{t}$, W + jets, and QCD events (left to right). The upper three panels use C/A reclustering while the lower three panels use k_T reclustering.

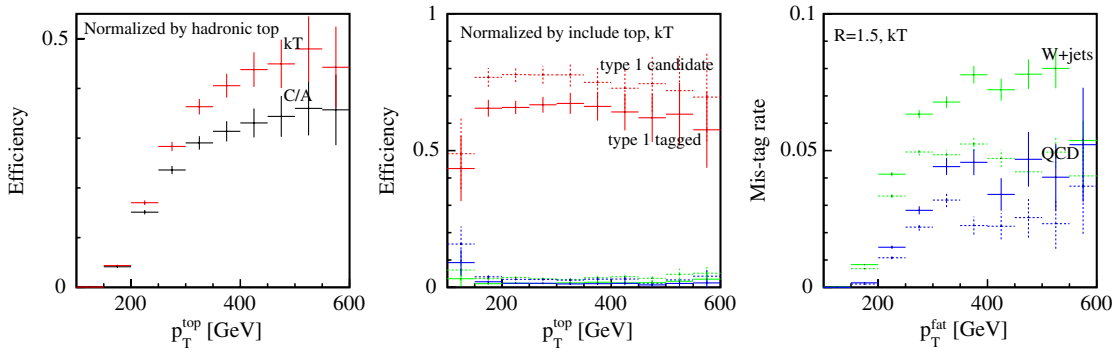

FIG. 6 (color online). Left: signal efficiencies relative to the number of hadronic tops for the k_T [light (red)], and C/A (black) algorithms. Center: k_T tagging efficiency relative to the number of tops included in a fat jet for type 1 [medium (red)], type 2 [light (green)], and type 3 [dark (blue)] tags. Right: mistag rates for QCD [dark (blue)] and W + jets [light (green)] for the C/A (dotted) and k_T algorithm (solid).

decreases towards high $p_{T,t}$; i.e., the efficiency of the mass plane cut is essentially constant. The fractions of type 2 and type 3 tags which include a top also decrease because the subjet momentum reconstruction improves. If the subjets are better matched to the hard top decay products more tagged tops are categorized as type 1.

Going beyond the signal, the same feature of fewer events with $m_{23} \sim 0$ also appears for the backgrounds. The mass plane cut then leads to a less efficient rejection, in particular, for soft masses. This increase of mistag probabilities we also show in Fig. 6. The solid crosses show the results for the k_T algorithm while the dotted crosses show those for the C/A algorithm. Quantitatively, the mistag efficiencies increase by a slightly larger factor than the tagging efficiency, so switching to the k_T algorithm does not improve S/B but can improve $S/\sqrt{B}$ slightly. In addition, these results are obtained without detector simulation and pileup. Switching to the k_T algorithm might have additional implications from both of them, so only a detailed experimental analysis can determine which jet algorithm to use in the HEPTOPTAGGER framework, and its result might well be process dependent.

IV. PRUNING

From the Higgs tagger based on the C/A algorithm and a mass drop criterion [15] we know that it can be advantageous to combine filtering and pruning in the tagging procedure [33]. Pruning removes soft radiation while clustering the fat jet [23,24]: first, a sequential jet algorithm combines unfiltered subjets until no pair of constituents is geometrically closer than $R_{\rm cut}$, representing an effective subjet cone size usually associated with an intrinsic fat-jet scale. We choose $R_{\rm cut} = m_j/p_{T,j}$. This cutoff can act differently for different underlying jet algorithms. After this recombination the unfiltered subjet merging continues, but with the additional restriction that each combined pair of subjets has to be sufficiently hard,

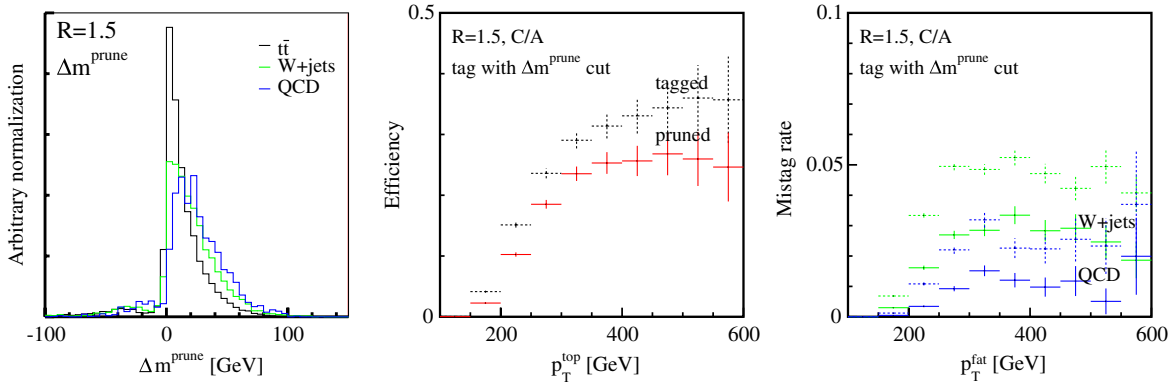

FIG. 7 (color online). Left: $\Delta m^{\rm prune}$ distribution for type 1 tags in $t\bar{t}$ (black line), W + jets [light (green)], and QCD [dark (blue)] events. Center: signal efficiencies as a function of $p_{T,t}$ without (black crosses) and with [light (red)] the pruning cut $\Delta m^{\rm prune} < 20$ GeV defined in Eq. (5). Right: mistag probabilities for QCD [dark (blue) crosses] and W + jets [light (green)] as functions $p_T^{\rm fat}$ with (solid) and without (dotted) the pruning cut.

TABLE II. Tagged top rates (in fb for $t\bar{t}$ and W + jets and pb for QCD jets) after cuts on Δm^{prune} or $\Delta m^{\text{unfilter}}$. The percentages are relative to the numbers without pruned or unfiltered mass cut in each category. $\epsilon_{S/B}$, $\epsilon_{S/\sqrt{B}}$ denote improvement factors relative to no cuts.

	Tagged	$\Delta m^{\text{prune}} < 15$	20	30	$\Delta m^{\text{unfilter}} < 15$	20	30
$t\bar{t}$ [fb]	14 156	7773 (55%)	9072 (64%)	10 875 (77%)	6237 (44%)	7926 (56%)	10 505 (74%)
Type 1	10 318	6255 (61%)	7152 (69%)	8316 (81%)	5141 (50%)	6377 (62%)	8129 (79%)
Type 2	1336	551 (41%)	693 (52%)	893 (67%)	403 (30%)	570 (43%)	847 (63%)
Type 3	2503	967 (39%)	1227 (49%)	1666 (67%)	693 (28%)	979 (39%)	1529 (61%)
W + jet [fb]	6590	2716 (41%)	3373 (51%)	4459 (68%)	2052 (31%)	2797 (42%)	4162 (63%)
$\epsilon_{S/B}$	1	1.33	1.25	1.14	1.41	1.32	1.18
$\epsilon_{S/\sqrt{B}}$	1	0.86	0.9	0.93	0.79	0.86	0.93
QCD [pb]	1229	359 (29%)	474 (39%)	719 (59%)	207 (17%)	331 (27%)	609 (50%)
$\epsilon_{S/B}$	1	1.88	1.66	1.31	2.62	2.08	1.5
$\epsilon_{S/\sqrt{B}}$	1	1.02	1.03	1	1.07	1.08	1.05

$$z = \frac{\min(p_{T,i}, p_{T,j})}{|\vec{p}_{T,i} + \vec{p}_{T,j}|} > z_{\text{cut}}, \tag{3}$$

where we choose $z_{\text{cut}} = 0.1$. Otherwise, the constituents i and j are not combined and the one with smaller transverse momentum is discarded. This algorithm continues until all constituents have been combined or eliminated.

There are two ways we can include pruning in our top tagger. First, we can prune the fat jet before we run the C/A algorithm extracting the relevant splittings using the mass drop criterion. Alternatively, we can use the pruning procedure in parallel to the filtering procedure and combine the two pieces of information. This approach ensures that the additional pruning step does not affect the performance of the rest of the tagging algorithm, so we investigate it in this section.

Since pruning is originally targeted at removing soft radiation, its impact is similar to that of filtering [15]. To quantify the difference we apply pruning to the constituents of the three subjets which we obtain after the usual mass drop criterion. These three subjets are selected such that they give the best filtered top mass among all combinations. The difference between the two algorithms is illustrated by the variable

$$\Delta m^{\text{prune}} = m^{\text{prune}} - m^{\text{filter}}, \tag{4}$$

where m^{filter} is the filtered mass for the selected three subjets, and m^{prune} is the jet mass of the pruned jet. The left panel in Fig. 7 shows Δm^{prune} for tagged tops in $t\bar{t}$ (type 1), QCD jets, and W + jets events. We find that Δm^{prune} is larger for background events than for signal events. Pruning generally collects more constituents than filtering, which discards some of the filtered subjets, so the pruned mass increases in a busy jet environment. Background events rely on such busy events to mimic the generically hard top decay products. Therefore, selecting tags with small Δm^{prune} effectively rejects the backgrounds. Even though we do not show them, type 2 and type 3 tags behave similarly to the backgrounds samples, because QCD jet radiation partly contributes to these tags. Thus, pruning also purifies the type 1 fraction for all tagged tops.

The event numbers after imposing

$$-10 \text{ GeV} < \Delta m^{\text{prune}} < \{15, 20, 30\} \text{ GeV} \tag{5}$$

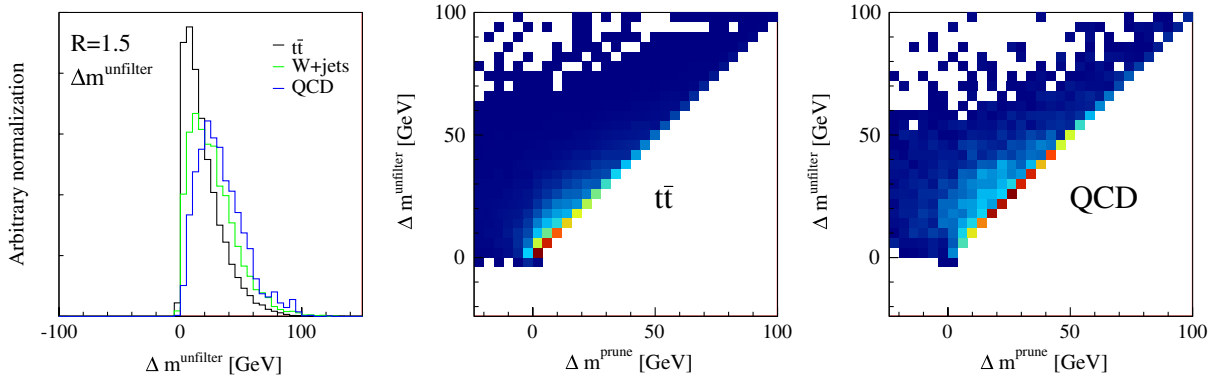

FIG. 8 (color online). Left: $\Delta m^{\text{unfilter}}$ for type 1 $t\bar{t}$ (black line), W + jets [light (green)], and QCD [medium (blue)] events. Center and right: two-dimensional correlation of the two mass differences.

we show in the left half of Table II. The percentages are relative to the number of tagged tops in each row. The different efficiencies we can translate into improvement factors for S/B and for $S/\sqrt{B}$. For example, we can improve S/B by roughly a factor two without loss of $S/\sqrt{B}$. In addition, the quality of the momentum reconstruction improves with the increased fraction of type 1 tags.

This additional cut becomes more important for larger thresholds of the subjet mass in the top tagger. For example, compared with the default choice $m_{\text{subjet}} > 50$ GeV a reduction to 30 GeV makes the cut on Δm^{prune} less efficient. This arises because with a larger m_{subjet} threshold more constituents can contribute in the background case.

According to the above discussion the difference in the pruned mass distribution for signal and backgrounds is due to the subjet multiplicity inside the fat jet. A simpler intuitive measure for this feature is the fat-jet mass before filtering. Algorithmically, we select the relevant three subjets only after filtering, but the original mass of the unfiltered subjets includes additional information:

$$\Delta m^{\text{unfilter}} = m^{\text{unfilter}} - m^{\text{filter}}. \tag{6}$$

In Fig. 8 we show this distribution for signal and backgrounds. Indeed, the left panel is very similar to the Δm^{prune} distribution of Fig. 7. The two-dimensional correlation confirms that almost all events for signal and background lie on the central diagonal of the Δm^{prune} vs $\Delta m^{\text{unfilter}}$ plane.

In the right half of Table II we show the corresponding efficiencies after cutting on $\Delta m^{\text{unfilter}}$. From both variables we can obtain significant improvements on S/B, and it remains an experimental question which of them is more stable once we include detector effects and pileup.

V. FATTER JETS

In the standard model the cross section for top pairs falls very steeply with increasing transverse momentum. The fraction of top pair event above different $p_{T,t}^{\text{min}}$ values at a 14 TeV LHC we calculated using ALPGEN [35]:

$p_{T,t}^{\text{min}}$ [GeV]	0	100	150	200	250	300	400	500
Fraction	100%	53%	28%	14%	6.8%	3.4%	0.96%	0.33%

Extending the top tagging reach by 50 GeV towards smaller $p_{T,t}$ corresponds to doubling the number of accessible top pairs.

The question becomes where the observed limitations of top tagging in this regime really come from and whether these constraints can be removed. From the left panel of Fig. 3 we know that the fraction of hadronic tops which can be included in a fat jet rapidly drops around $p_{T,t} =$ 200–250 GeV. Compared to a well suited data sample with $p_{T,t} > 300$ GeV the tagging probability roughly drops to half its value. On the other hand, in the right panel of Fig. 3 we see how the fraction of tagged tops relative to the number of hadronic tops included in a fat jet increases. This suggests that larger values than $R = 1.5$ should significantly improve the tagging efficiency around 200–250 GeV. In this section we will study an increase to $R = 1.8$ for our standard HEPTOPTAGGER setup to test such an option.

We show in the left panel of Fig. 9 the tagging efficiency for $R = 1.8$ as a function of $p_{T,t}$ as well as the fraction of hadronic tops included in a fat jet. By increasing R from 1.5 to 1.8 we increase the tagging efficiency for tops with $p_{T,t} = 150$–250 GeV by a factor 1.5 to 3. This is mainly an effect of more hadronic tops fully included in the fat jet. In the 300–450 GeV range the effect of an increased R is small, and for $p_{T,t} > 450$ the efficiency even slightly decreases due to combinatorics.

The central panel shows the fractions of type 1 candidates and type 1 tags relative to all hadronic tops included in a fat jet as functions of hadronic top p_T. All the way down to $p_{T,t} = 100$ GeV the efficiencies are flat, which means we have a fair chance to collect very moderately boosted tops. Type 2 and type 3 fractions we show at the

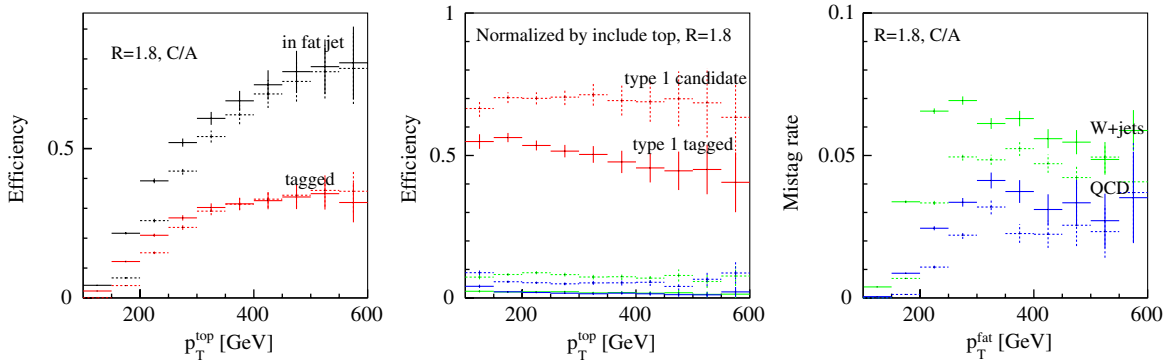

FIG. 9 (color online). Left: tagging efficiency [solid light (red) crosses] and the fraction of tops included in the fat jet for $R = 1.8$ (solid) and for $R = 1.5$ (dotted) as a function of $p_{T,t}$ for semileptonic $t\bar{t}$ events. Center: tagging candidates and tags relative to all tops included in a fat jet. Type 2 and type 3 tags are shown in [lightest (green) and darkest (blue)]. Right: mistag rate as a function of fat jet p_T for QCD [dark (blue)] and W + jets [light (green)].

TABLE III. Numbers of tagged tops with $R = 1.8$ with several cuts on Δm^{prune} or $\Delta m^{\text{unfilter}}$. The percentages are relative to the numbers of tagged tops without pruned or unfiltered mass cut in each category. $\epsilon_{S/B}$ and $\epsilon_{S/\sqrt{B}}$ denote improvement factors relative to the $R = 1.5$ numbers with no cuts as shown in Table II.

$R = 1.8$	Tagged [fb]	$\Delta m^{\text{prune}} < 15$	20	30	$\Delta m^{\text{unfilter}} < 15$	20	30
$t\bar{t}$ [fb]	27 853	10 695(38%)	13 221(47%)	17 453(63%)	8253(30%)	11 131(40%)	16 148(58%)
Type 1	17 502	8114(46%)	9716(56%)	12 195(70%)	6463(37%)	8403(48%)	11 507(66%)
Type 2	3628	934(26%)	1252(35%)	1847(51%)	655(18%)	996(27%)	1666(46%)
Type 3	6723	1647(24%)	2252(34%)	3410(51%)	1135(17%)	1732(26%)	2975(44%)
W + jet [fb]	16 920	4274(25%)	5791(34%)	8551(51%)	3063(18%)	4521(27%)	7620(45%)
$\epsilon_{S/B}$	0.77	1.16	1.06	0.95	1.25	1.15	0.99
$\epsilon_{S/\sqrt{B}}$	1.23	0.94	1	1.08	0.86	0.95	1.06
QCD [pb]	4402	644(15%)	936(21%)	1627(37%)	337(8%)	584(13%)	1279(29%)
$\epsilon_{S/B}$	0.55	1.44	1.23	0.93	2.13	1.65	1.1
$\epsilon_{S/\sqrt{B}}$	1.04	1.04	1.07	1.07	1.11	1.14	1.12

bottom of the figure. While the fraction of type 2 and type 3 tags increases the fraction of type 1 candidates and tags does not drastically change compared to $R = 1.5$ as shown in the right panel of Fig. 3.

Finally, increasing R also increases the mistag rate significantly. The right panel of Fig. 9 shows the mistag efficiency as functions of the fat jet p_T for QCD and W + jets events. We observe a larger increase for these background processes than for the signal, which means we will not improve S/B through larger jet sizes. However, we might improve the statistical significance measure $S/\sqrt{B}$.

In Table III we show the improvements in S/B and $S/\sqrt{B}$ relative to the $R = 1.5$ case. We see that roughly twice the number of tops get tagged, mainly at low transverse momenta. However, S/B still decreases by a factor $1/2$, while $S/\sqrt{B}$ slightly improves as long as QCD is the main background.

To compensate for the increased backgrounds we can also apply pruning for $R = 1.8$. The corresponding efficiencies for different pruned mass cuts we also include in Table III. Adding pruning shifts back the performance to a similar level as our $R = 1.5$ results, and there is no obvious advantage in combining it with larger fat jets, unless there should be a specific reason to target the low-$p_{T,t}$ regime.

VI. BOTTOM TAGGING

A major difference in the background rejection between the C/A based Higgs tagger [15] and corresponding top taggers is additional b tags. Only based on kinematic conditions it appears unlikely to achieve a QCD or W + jets rejection of more than a factor $1/100$. However, we can gain a significant improvement by requiring a b tag for one of the top decay jets. At the same time, for moderately boosted top quarks the kinematic tagging algorithm might benefit from the identification of the b jet, so we can first ensure that it is captured and, second, use this information in the kinematic reconstruction.

Our first attempt to improve top tagging through an additional b tag will leave the kinematic top tagging algorithm unchanged and will instead focus on the selection of the subjet which should be b-tagged. All of the usual top taggers treat three subjets democratically, i.e., without any b-tagging information. We label the b, W_1, and W_2 subjets such that W_1 and W_2 subjets (ordered by hardness) reconstruct m_W best; j_b, j_{W1}, and j_{W2} are the corresponding parton labels from Monte Carlo truth, as defined in Sec. II. An obvious question is for what fraction of all tags the subjet labeled b really points to the bottom quark $b = j_b$. For type 1 tags by definition one of the subjets corresponds to a bottom, so $b = j_b$ implies that all subjets are correctly

TABLE IV. Kinematic identification probabilities (b) for all three top decay partons for different types of top tags and for different numbers of subjet pairings consistent with m_W (n_W).

	All n_W		$n_W = 1$		$n_W = 2$	
$t\bar{t}$ [fb]	14 156		8058		6099	
$b = j_b$	9325	(66%)	5882	(73%)	3442	(56%)
$b = j_{W1}$	2971	(21%)	1242	(15%)	1728	(28%)
$b = j_{W2}$	1666	(12%)	833	(10%)	833	(14%)
Type 1	10 318		5808		4509	
$b = j_b$	7917	(77%)	5044	(87%)	2874	(64%)
$b = j_{W1}$	1695	(16%)	502	(9%)	1193	(26%)
$b = j_{W2}$	706	(7%)	263	(4%)	443	(10%)
Type 2	1336		781		555	
$b = j_b$	565	(42%)	341	(44%)	224	(40%)
$b = j_{W1}$	499	(37%)	294	(38%)	205	(37%)
$b = j_{W2}$	392	(29%)	226	(29%)	166	(30%)
Type 3	2503		1468		1035	
$b = j_b$	842	(34%)	498	(34%)	344	(33%)
$b = j_{W1}$	777	(31%)	447	(30%)	331	(32%)
$b = j_{W2}$	568	(23%)	344	(23%)	224	(22%)
W + jet [fb]	6590		3733		2857	
QCD [pb]	1229		713		516	

assigned. For such tags a W decay angle analysis [26] will work well.

In the first column of Table IV we summarize the probabilities to correctly assign the label b in the kinematics-based top tagging. The percentages are defined relative to the number of tagged tops in each category. For type 1 tags there will be exactly one b parton in the tagged top while for type 2 and type 3 tags there can be any number of b partons. This means that only for type 1 tags all three fractions sum to 100%. Almost 80% of type 1 tags correctly assign $b = j_b$, so they make a good test sample for the question if identifying $b = j_b$ through a b tag leads to an efficient rejection for W + jets and QCD events. From Table IV we also see that the second-most likely subjet to be kinematically identified as a b really is j_{W1}. However, this probability is less than a third of that for $b = j_b$. To rely on an additional b tag for the case $b = j_{W1}$ does not even improve $S/\sqrt{B}$, so the only way to increase S/B is to apply a b tag only to the kinematically identified j_b.

As we show in Fig. 5, a significant fraction of events passing the mass plane cut have two pairs of subjets consistent with the W mass. This happens because the upper bound $m_{bj}^2 < m_t^2 - m_W^2$ is numerically close to m_W^2 [12]. We expect that tagged tops where only one pair of subjets is consistent with m_W should more reliably give $b = j_b$. To test this, we count the number of subjet pairs consistent with m_W as $n_W = 2$ when $\Delta m_2 - \Delta m_1 < 10\% \times m_W$ where $\Delta m = |m_{jj} - m_W|$ for each pairing. Because $n_W = 3$ essentially never appears we find $n_W = 1$ for the other top tags.

The distributions for kinematically identified b subjets for different n_W we also show in Table IV. The fractions of tagged tops with $n_W = 1$ are almost the same ($\sim 57\%$) for $t\bar{t}$ and QCD, W + jets. Type 1 tagged tops with $n_W = 1$ are more likely to give $b = j_b$ than those with $n_W = 2$, while for type 2 and type 3 we do not observe a significant difference.

Consequently, the best strategy to improve S/B in top tagging is to target $n_W = 1$ events and check the kinematically identified b subjet with a b tag. In terms of a tagging efficiency ε_b and a misidentification rate $\varepsilon_b^{\text{mis}}$ for light flavors (we ignore c quarks in this simple estimate) we show the possible improvements in Table V. We expect an enhancement by $0.66\varepsilon_b/\varepsilon_b^{\text{mis}}$ ($0.73\varepsilon_b/\varepsilon_b^{\text{mis}}$ for selecting $n_W = 1$) for S/B, so assuming $\varepsilon_b = 50\%$ and $\varepsilon_b^{\text{mis}} = 2\%$ we find that S/B improves by 16.5 (18 for selecting $n_W = 1$). Similarly, we find an improvement in $S/\sqrt{B}$ around a factor 2.

As an alternative to an external b tag, we can use the b-tagging information during the kinematic top tagging algorithm. We modified in four steps the selection of the three decay subjets after the mass drop criteria:

(1) Group the subjets into either b-tagged or non-b-tagged subjets.
(2) Take all possible triplets of one b-tagged and two non-b-tagged subjets and select the one with the best filtered top mass.
(3) Check if this m^{filter} satisfies our criterion.
(4) Apply modified mass plane cuts.

Since we require exactly one b-tagged subjet in step 2, we adapt the labels b, W_1, W_2 to reflect this. In step 4, we should change the invariant masses in terms of (p_1, p_2, p_3) into (p_b, p_{W1}, p_{W2}). Our two-dimensional mass plane becomes $\arctan(m_{b2}/m_{b1})$ vs m_{12}/m_{123}, where $m_{bi}^2 = (p_b + p_{Wi})^2$ and $m_{12}^2 = (p_{W1} + p_{W2})^2$. Figure 10 shows the two-dimensional distribution of top candidates in the modified mass plane for semileptonic $t\bar{t}$ events, W + jets, and QCD events. Unlike before, the signal events now show a clear W mass peak only for m_{12}. To see how well this new algorithm might do we assume a 100% b-tagging efficiency of perfect purity. Because for the backgrounds all three subjets could equally likely be mistagged, we simply reweight each of the possibilities by $\varepsilon_b^{\text{mis}}$. Following these plots we can apply a stricter mass plane cut on the modified mass plane than before. For our test we use

$$\left| \frac{m_{12}}{m_{123}} - \frac{m_W}{m_t} \right| < 15\% \quad \text{and} \quad 0.2 < \arctan\left(\frac{m_{b2}}{m_{b1}}\right) < 1.3. \tag{7}$$

Based on this modified algorithm the left panel of Fig. 11 shows the signal efficiency as a function of $p_{T,t}$. Again, for the signal we assume a perfect b tag. We can

TABLE V. Efficiencies of a b tag after top tagging and for the modified tagger with b tagging. We assume $\varepsilon_b = 0.5$ and $\varepsilon_b^{\text{mis}} = 0.02$ and quote the improvement factors $\epsilon_{S/B}$, $\epsilon_{S/\sqrt{B}}$ against the QCD background.

	Tagged		$b = j_b$		$b = j_b$ and $n_W = 1$		b tag with cut for m_{12}
$t\bar{t}$	1	0.33	$(0.66\varepsilon_b)$	0.21	$(0.42\varepsilon_b)$	0.41	$(0.82\varepsilon_b)$
Type 1	0.73	0.28	$(0.56\varepsilon_b)$	0.18	$(0.36\varepsilon_b)$	0.36	$(0.72\varepsilon_b)$
W + jet	1	0.02	$(\varepsilon_b^{\text{mis}})$	0.0114	$(0.57\varepsilon_b^{\text{mis}})$	0.0318	$(1.59\varepsilon_b^{\text{mis}})$
QCD	1	0.02	$(\varepsilon_b^{\text{mis}})$	0.0116	$(0.58\varepsilon_b^{\text{mis}})$	0.0332	$(1.66\varepsilon_b^{\text{mis}})$
$\epsilon_{S/B}$	1	16.5	$(0.66\varepsilon_b/\varepsilon_b^{\text{mis}})$	18.42	$(0.72\varepsilon_b/\varepsilon_b^{\text{mis}})$	12.35	$(0.49\varepsilon_b/\varepsilon_b^{\text{mis}})$
$\epsilon_{S/\sqrt{B}}$	1	2.33	$(0.66\varepsilon_b/\sqrt{\varepsilon_b^{\text{mis}}})$	1.95	$(0.55\varepsilon_b/\sqrt{\varepsilon_b^{\text{mis}}})$	2.25	$(0.64\varepsilon_b/\sqrt{\varepsilon_b^{\text{mis}}})$

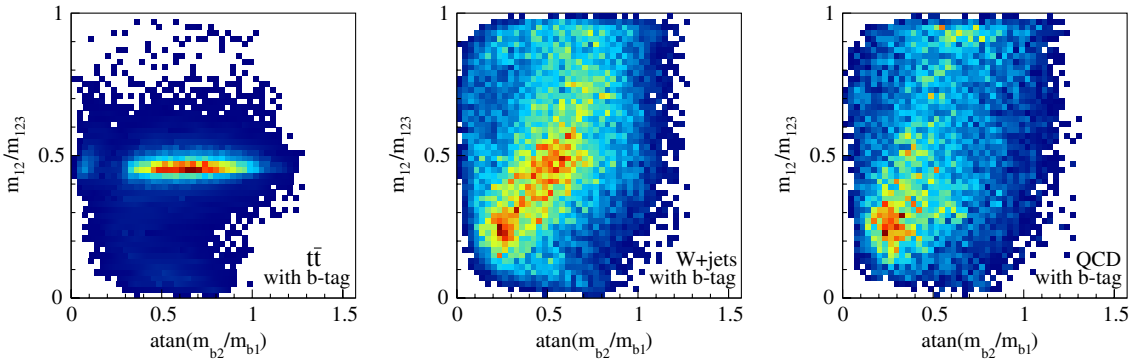

FIG. 10 (color online). Modified mass plane distribution for signal, W + jets, and QCD events. The mass plane is now defined as $\arctan(m_{b2}/m_{b1})$ vs m_{12}/m_{123}.

simply multiply by ε_b to compute the final top tagging efficiency, ignoring the small effect of $\varepsilon_b^{\text{mis}}$. Similarly, we ignore cases with two b subjets in the tagged top as subleading by a factor $\varepsilon_b(1-\varepsilon_b)$ and only appearing for type 2 or type 3.

Taking into account the probability of 77% with which the default top tagger correctly assigns $b = j_b$, the modified algorithm slightly increases the number of tagged tops if we assume perfect b tagging. It returns almost the same tagging efficiency as the purely kinematic tagger before adding the factor ε_b. This is because the default top tagger identifies the crucial type 1 tops even without b-tagging information while type 2 or type 3 configurations are comparably rare. The use of b tagging really only helps to identify which subjet of a type 1 tag corresponds to the bottom and to reject type 2 and type 3 tags. To confirm this, in the central panel of Fig. 11 we show the efficiency as a function of the fat jet's p_T. Compared to $p_{T,t}$ this effectively removes some type 2 and type 3 tops.

The right panel of Fig. 11 shows the mistag rate of the modified algorithm. Here we simply assume that one of the three subjets found by the kinematic top tagger is misidentified. These candidate tops have to be multiplied by $3 \times \varepsilon_b^{\text{mis}}$ before imposing the modified mass plane cut. We neglect the effect of misidentifying b subjets other than among the three subjets selected by the kinematic tagger. It should be less than 10% given that more than 90% of the fat jets in which three subjets fulfill the top mass criteria only have one such combination.

The main question is if the restricted mass plane cuts now reduce the backgrounds more efficiently. We find that the cuts from Eq. (7) efficiently drop tagged tops from background samples as compared to democratic mass plane cuts, but they do not compensate for the combinatorial factor 3 in the mistagging probability.

Up to this point we have only considered semileptonic $t\bar{t}$ events. In particular for the possible improvement through b tagging we should consider the more combinatorics prone purely hadronic decays. The left panel of Fig. 12

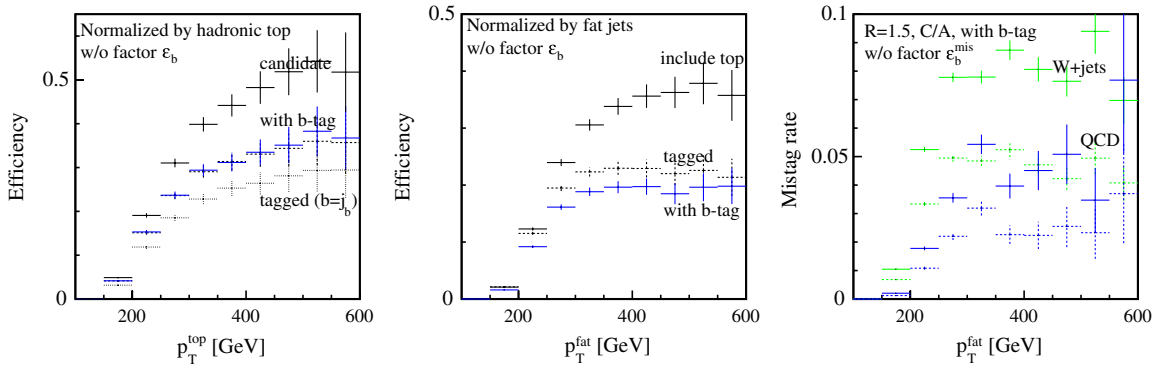

FIG. 11 (color online). Left: tagging efficiency of the modified algorithm as a function of $p_{T,t}$ and assuming $\varepsilon_b = 1$. Center: tagging efficiency of the modified algorithm as a function of the p_T of the fat jet. Right: mistagging rate of the modified algorithm as a function of the p_T of the fat jet for QCD [dark (blue) crosses] and W + jets samples [light (green)]. Again, we omit $\varepsilon_b^{\text{mis}}$. In all panels dotted lines show the HEPTOPTAGGER results as a reference.

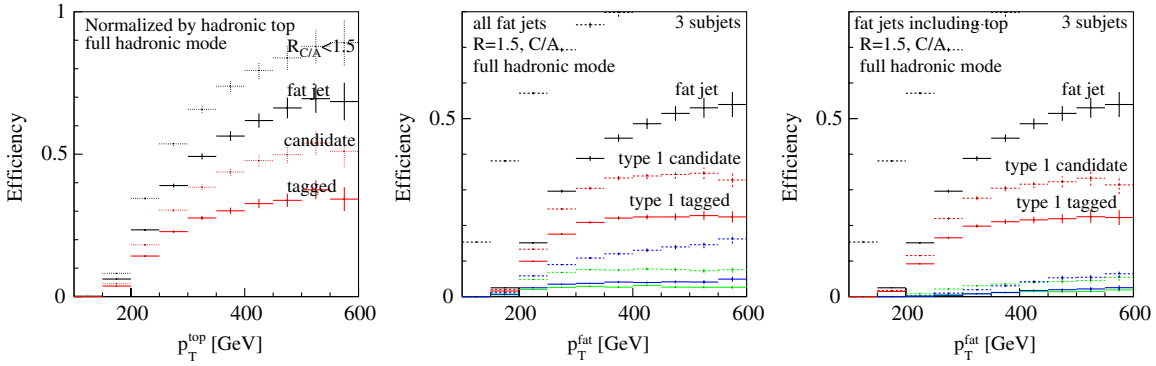

FIG. 12 (color online). Left: tagging efficiencies $\epsilon_{t\bar{t}}$ normalized to the number of hadronic tops as a function of the p_T for the fully hadronic $t\bar{t}$ sample. Center: efficiencies for type 1 [medium (red) crosses], type 2 [light (green)], and type 3 [dark (blue)] as functions of the fat jet p_T for the fully hadronic $t\bar{t}$ sample. The dotted lines show the corresponding candidate fractions. Right: fraction of type 1, type 2, and type 3 only for fat jets including a top.

shows the tagging efficiencies as a function of $p_{T,t}$ for the fully hadronic $t\bar{t}$ sample. Compared to Fig. 3 the tagger works with almost the same efficiency, provided we normalize the number of tagged events to all hadronic tops.

The central panel of Fig. 12 shows the default top tagger efficiencies as a function of the transverse momentum of the fat jet, again for hadronic decays. The tagging efficiency and the fraction of candidate tops we show for type 1, type 2, and type 3 tags. Since the number of fat jets including a hadronic top shown as the black solid line increases compared to semileptonic events, the resulting candidate and tagged efficiency becomes larger. The number of type 2 and type 3 tags increases simply with the jet multiplicities. In the right panel of Fig. 12 we show the same fractions as in the central panel but requiring that all fat jets include a hadronic top. For such fat jets we find type 2 or type 3 almost as rarely as in the semileptonic sample; i.e., type 2 and type 3 candidates contribute at most 30% relative to type 1 candidates. This means that even for the fully hadronic $t\bar{t}$ sample a modified algorithm including b tagging will not provide enough of an improvement to compensate for an expected factor 50% increase in the mistag rate.

In summary, using b tagging when selecting the relevant three subjets in a fat jet does not enhance S/B. The purely kinematic HEPTOPTAGGER selects the correct set of subjets too reliably to gain a significant improvement as long as the hadronic top is fully captured in a fat jet. The b subjet selected based on kinematics is usually identified correctly. Relying on b tagging inside the kinematic algorithm is hurt by the combinatorial mistagging efficiency of $3 \times \epsilon^{\text{mis}}$ while there is no such factor 3 for the signal. This disadvantage is hard to compensate for by improved mass plane cuts. To improve S/B, the best approach is to use the b tag only for the most probable b subjet and simply add it to the kinematic tagging algorithm.

VII. OUTLOOK

In this paper we have proposed and tested several modifications to kinematic top tagging, as implemented in the HEPTOPTAGGER. As a starting point we have shown that provided the top quark is boosted enough to be collected inside a fat jet the usual kinematic criteria, i.e., a search for mass drops in the clustering history and the reconstruction of three independent invariant mass variables from the suspected top decay subjets, do not exhibit obvious shortcomings.

(1) One possible improvement, curing, for example, the decreasing efficiency of C/A based taggers towards larger boost, is a switch to the k_T algorithm once we identified the main subjets using mass drops. It keeps the tagging efficiency relative to the number of tops caught inside a fat jet on a plateau over the entire range $p_{T,t} \sim 150$–600 GeV.

(2) Using pruning in combination with the usual filtering procedure we gain an additional kinematic variable. Cutting on it can almost double S/B relative to pure QCD backgrounds. However, this improvement should be taken with a grain of salt until it can be confirmed by a proper detector simulation in the presence of pileup.

(3) To extend the top tagger to lower boost we can increase the original size of the fat jet from $R = 1.5$ to $R = 1.8$. Indeed, the efficiency for low transverse momenta increases, but the background combinatorics do as well. We find that neither S/B nor $S/\sqrt{B}$ strongly benefit from this modification.

(4) Including a b tag as an additional step in the top tagging procedure can very significantly enhance the background rejection. This is well-known from Higgs tagging. We also find that, for the HEPTOPTAGGER, including such a b tag in a modified algorithm is not promising.

At a time when we can expect the first officially tagged top quarks at the LHC, these studies can guide us towards possible modifications and improvements of top taggers with different analyses in mind. They show that, for example, in the case of the HEPTOPTAGGER there is still room for adjustment but that top taggers in general have within a few years reached an impressive level of maturity and reliability.

ACKNOWLEDGMENTS

We would like to thank our experimental ATLAS colleagues Dirk Zerwas, Sebastian Schätzel, and, in particular, Gregor Kasieczka for their support and advice. Without personal contacts progress in this field of LHC phenomenology would not be possible. Felix Kling we would like to thank for carefully checking our analysis code.

[1] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **74**, 2626 (1995); S. Abachi *et al.* (D0 Collaboration), Phys. Rev. Lett. **74**, 2422 (1995).
[2] D. E. Morrissey, T. Plehn, and T. M. P. Tait, arXiv:0912.3259.
[3] P. Meade and M. Reece, Phys. Rev. D **74**, 015010 (2006).
[4] K. Rehermann and B. Tweedie, J. High Energy Phys. 03 (2011) 059; T. Plehn, M. Spannowsky, and M. Takeuchi, J. High Energy Phys. 05 (2011) 135.
[5] T. Plehn, M. Spannowsky, M. Takeuchi, and D. Zerwas, J. High Energy Phys. 10 (2010) 078; http://www.thphys.uni-heidelberg.de/~plehn/.
[6] J. Ellis, F. Moortgat, G. Moortgat-Pick, J. M. Smillie, and J. Tattersall, Eur. Phys. J. C **60**, 633 (2009); K. Rolbiecki, J. Tattersall, and G. Moortgat-Pick, Eur. Phys. J. C **71**, 1571 (2011); M. Perelstein and A. Weiler, J. High Energy Phys. 03 (2009) 141.
[7] See, e.g., V. Barger, T. Han, and D. G. E. Walker, Phys. Rev. Lett. **100**, 031801 (2008); U. Baur and L. H. Orr, Phys. Rev. D **76**, 094012 (2007); T. Han, R. Mahbubani, D. G. E. Walker, and L. T. E. Wang, J. High Energy Phys. 05 (2009) 117.
[8] See, e.g., U. Baur and L. H. Orr, Phys. Rev. D **77**, 114001 (2008); P. Fileviez Perez, R. Gavin, T. McElmurry, and F. Petriello, Phys. Rev. D **78**, 115017 (2008); Y. Bai and Z. Han, J. High Energy Phys. 04 (2009) 056; B. Bhattacherjee, M. Guchait, S. Raychaudhuri, and K. Sridhar, Phys. Rev. D **82**, 055006 (2010).
[9] T. Aaltonen *et al.* (CDF Collaboration, Phys. Rev. D **82**, 112005 (2010).
[10] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **100**, 142002 (2008); T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. D **83**, 112003 (2011).
[11] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **106**, 171801 (2011).
[12] T. Plehn and M. Takeuchi, J. Phys. G **38**, 095 006 (2011); Z. Sullivan and A. Menon, Phys. Rev. D **83**, 091504 (2011).
[13] G. D. Kribs, A. Martin, and T. S. Roy, Phys. Rev. D **84**, 095024 (2011); X.-J. Bi, Q.-S. Yan, and P.-F. Yin, Phys. Rev. D **85**, 035005 (2012).
[14] M. H. Seymour, Z. Phys. C **62**, 127 (1994); J. M. Butterworth, B. E. Cox, and J. R. Forshaw, Phys. Rev. D **65**, 096014 (2002).
[15] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, Phys. Rev. Lett. **100**, 242001 (2008).
[16] ATLAS note, Report No. ATL-PHYS-PUB-2009-088; J. Gallicchio, J. Huth, M. Kagan, M. D. Schwartz, K. Black, and B. Tweedie, J. High Energy Phys. 04 (2011) 069; J.-H. Kim, Phys. Rev. D **83**, 011502 (2011); D. E. Soper and M. Spannowsky, Phys. Rev. D **84**, 074002 (2011).
[17] G. D. Kribs, A. Martin, T. S. Roy, and M. Spannowsky, Phys. Rev. D **81**, 111501 (2010); **82**, 095012 (2010); C. R. Chen, M. M. Nojiri, and W. Sreethawong, J. High Energy Phys. 11 (2010) 012; A. Falkowski, D. Krohn, L.-T. Wang, J. Shelton, and A. Thalapillil, Phys. Rev. D **84**, 074022 (2011); A. Katz, M. Son, and B. Tweedie, Phys. Rev. D **83**, 114033 (2011); B. Bellazzini, C. Csaki, J. Hubisz, and J. Shao, Phys. Rev. D **83**, 095018 (2011); J. Thaler and Z. Thomas, J. High Energy Phys. 07 (2011) 060; C. Englert, T. S. Roy, and M. Spannowsky, Phys. Rev. D **84**, 075026 (2011); S. Yang and Q.-S. Yan, arXiv:1111.4530.
[18] C. Hackstein and M. Spannowsky, Phys. Rev. D **82**, 113012 (2010); C. Englert, C. Hackstein, and M. Spannowsky, Phys. Rev. D **82**, 114024 (2010); A. Katz, M. Son, and B. Tweedie, J. High Energy Phys. 03 (2011) 011; Y. Cui, Z. Han, and M. D. Schwartz, Phys. Rev. D **83**, 074023 (2011).
[19] J. M. Butterworth, J. R. Ellis, and A. R. Raklev, J. High Energy Phys. 05 (2007) 033; J. M. Butterworth, J. R. Ellis, A. R. Raklev, and G. P. Salam, Phys. Rev. Lett. **103**, 241803 (2009).
[20] A. Abdesselam *et al.*, Eur. Phys. J. C **71**, 1661 (2011); L. G. Almeida, R. Alon, and M. Spannowsky, arXiv:1110.3684.
[21] M. Gerbush, T. J. Khoo, D. J. Phalen, A. Pierce, and D. Tucker-Smith, Phys. Rev. D **77**, 095003 (2008); G. Brooijmans, Reports No. ATL-PHYS-CONF-2008-008 and ATL-COM-PHYS-2008-001, 2008; L. G. Almeida, S. J. Lee, G. Perez, I. Sung, and J. Virzi, Phys. Rev. D **79**, 074012 (2009).
[22] T. Plehn, G. P. Salam, and M. Spannowsky, Phys. Rev. Lett. **104**, 111801 (2010).
[23] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, Phys. Rev. D **80**, 051501 (2009); **81**, 094023 (2010); C. K. Vermilion, arXiv:1101.1335.
[24] J. R. Walsh and S. Zuberi, arXiv:1110.5333.
[25] J. Thaler and L.-T. Wang, J. High Energy Phys. 07 (2008) 092; D. Krohn, J. Thaler, and L.-T. Wang, J. High Energy Phys. 06 (2009) 059; 02 (2010) 084.
[26] D. E. Kaplan, K. Rehermann, M. D. Schwartz, and B. Tweedie, Phys. Rev. Lett. **101**, 142001 (2008).

[27] L. G. Almeida, S. J. Lee, G. Perez, G. Sterman, I. Sung, and J. Virzi, Phys. Rev. D **79**, 074017 (2009); L. G. Almeida, S. J. Lee, G. Perez, G. Sterman, and I. Sung, Phys. Rev. D **82**, 054034 (2010).

[28] J. Thaler and K. Van Tilburg, J. High Energy Phys. 03 (2011) 015; J. Thaler and K. Van Tilburg, arXiv:1108.2701.

[29] A. Hook, M. Jankowiak, and J. G. Wacker, arXiv:1102.1012; V. Barger and P. Huang, arXiv:1102.3183; M. Jankowiak and A. J. Larkoski, J. High Energy Phys. 06 (2011) 057.

[30] G. Kasieczka, S. Schatzel, and A. Schoning, Report No. ATL-PHYS-INT-2011-007; G Kasieczka, talk given at DPG 2011 [http://www.dpg-verhandlungen.de/year/2011/conference/karlsruhe/part/t/session/56/contribution/5].

[31] CMS Collaboration, Report No. CMS-PAS-EXO-11 [http://cdsweb.cern.ch/record/1369210/files/EXO-11-020-pas.pdf].

[32] J. L. Hewett, J. Shelton, M. Spannowsky, T. M. P. Tait, and M. Takeuchi, Phys. Rev. D **84**, 054005 (2011).

[33] D. E. Soper and M. Spannowsky, J. High Energy Phys. 08 (2010) 029.

[34] G. Piacquadio (Freiburg U.), Report No. CERN-THESIS-2010-027.

[35] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, J. High Energy Phys. 07 (2003) 001.

[36] T. Sjostrand, S. Mrenna, and P. Z. Skands, J. High Energy Phys. 05 (2006) 026.

[37] M. L. Mangano, M. Moretti, and R. Pittau, Nucl. Phys. **B632**, 343 (2002); S. Hoeche, F. Krauss, N. Lavesson, L. Lonnblad, M. Mangano, A. Schalicke, and S. Schumann, Proceedings of the HERA and the LHC Workshop, CERN/DESY 2004/2005 [arXiv:hep-ph/0602031].

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

[39] P. Nason, S. Dawson, and R. K. Ellis, Nucl. Phys. **B303**, 607 (1988); W. Beenakker, H. Kuijf, W. L. van Neerven, and J. Smith, Phys. Rev. D **40**, 54 (1989); S. Moch and P. Uwer, Phys. Rev. D **78**, 034003 (2008).

[40] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, J. High Energy Phys. 08 (1997) 001; M. Wobisch and T. Wengler, in Monte Carlo Generators for HERA Physics, Hamburg 1998/1999, p. 270 [arXiv:hep-ph/9907280].

[41] S. Catani, Y. L. Dokshitzer, M. Olsson, G. Turnock, and B. R. Webber, Phys. Lett. B **269**, 432 (1991); S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, Nucl. Phys. **B406**, 187 (1993); S. D. Ellis and D. E. Soper, Phys. Rev. D **48**, 3160 (1993); M. Cacciari, arXiv:hep-ph/0607071.

[42] M. Cacciari, G. P. Salam, and G. Soyez, J. High Energy Phys. 04 (2008) 063.