Eur. Phys. J. C (2014) 74:2792
DOI 10.1140/epjc/s10052-014-2792-8

THE EUROPEAN PHYSICAL JOURNAL C

Special Article - Tools for Experiment and Theory

Boosted objects and jet substructure at the LHC. Report of BOOST2012, held at IFIC Valencia, 23rd–27th of July 2012

A. Altheimer[1], A. Arce[2], L. Asquith[3], J. Backus Mayes[4], E. Bergeaas Kuutmann[5,49], J. Berger[6], D. Bjergaard[2], L. Bryngemark[7], A. Buckley[8], J. Butterworth[9], M. Cacciari[10], M. Campanelli[9], T. Carli[11], M. Chala[12], B. Chapleau[13], C. Chen[14], J. P. Chou[15], Th. Cornelissen[16], D. Curtin[17], M. Dasgupta[18], A. Davison[9], F. de Almeida Dias[19], A. de Cosa[20], A. de Roeck[11], C. Debenedetti[8], C. Doglioni[21], S. D. Ellis[22], F. Fassi[23], J. Ferrando[24], S. Fleischmann[16], M. Freytsis[25,47], M. L. Gonzalez Silva[26], S. Gonzalez de la Hoz[23], F. Guescini[21], Z. Han[27], A. Hook[4], A. Hornig[22], E. Izaguirre[4], M. Jankowiak[4], J. Juknevich[28], M. Kaci[23], D. Kar[24], G. Kasieczka[30], R. Kogler[29], A. Larkoski[4], P. Loch[31], D. Lopez Mateos[27], S. Marzani[32], L. Masetti[33], V. Mateu[34], D. W. Miller[35], K. Mishra[36], P. Nef[4], K. Nordstrom[24], E. Oliver Garcia[23], J. Penwell[37], J. Pilot[38], T. Plehn[30], S. Rappoccio[39,48], A. Rizzi[40], G. Rodrigo[23], A. Safonov[41], G. P. Salam[10,11], J. Salt[23], S. Schaetzel[30], M. Schioppa[42], A. Schmidt[29,30], J. Scholtz[22], A. Schwartzman[4], M. D. Schwartz[27], M. Segala[43], M. Son[44], G. Soyez[45], M. Spannowsky[32], I. Stewart[34], D. Strom[46], M. Swiatlowski[4], V. Sanchez Martinez[23], M. Takeuchi[30], J. Thaler[34], E. N. Thompson[1], N. V. Tran[36], C. Vermilion[25], M. Villaplana[23], M. Vos[23,a], J. Wacker[4], J. Walsh[47]

[1] Nevis Laboratory, Columbia University, Irvington, NY 10533, USA
[2] Duke University, Durham, NC 27708, USA
[3] Argonne National Laboratory, Lemont, IL 60439, USA
[4] SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA
[5] Deutsches Elektronen-Synchrotron, DESY, 15738 Zeuthen, Germany
[6] Cornell University, Ithaca, NY 14853, USA
[7] Lund University, 22100 Lund, Sweden
[8] University of Edinburgh, Edinburgh EH9 3JZ, UK
[9] University College London, London WC1E 6BT, UK
[10] CERN, 1211 Geneva 23, Switzerland
[11] LPTHE, UPMC Univ. Paris 6 and CNRS UMR, 7589 Paris, France
[12] CAFPE and U. of Granada, 18071 Granada, Spain
[13] McGill University, Montreal, QC H3A 2T8, Canada
[14] Iowa State University, Ames, IA 50011, USA
[15] Rutgers University, Piscataway, NJ 08854, USA
[16] Bergische Universitaet Wuppertal, 42097 Wuppertal, Germany
[17] YITP, Stony Brook University, Stony Brook, NY 11794-3840, USA
[18] University of Manchester, Manchester M13 9PL, UK
[19] UNESP-Universidade Estadual Paulista, São Paulo 01140-070, Brazil
[20] INFN and University of Naples, 80216 Naples, Italy
[21] University of Geneva, 1211Geneva 4 , Switzerland
[22] University of Washington, Seattle, WA 98195, USA
[23] Instituto de Física Corpuscular, IFIC/CSIC-UVEG, 46071 Valencia, Spain
[24] University of Glasgow, Glasgow G12 8QQ, UK
[25] Lawrence Berkeley National Laboratory, University of California, Berkeley, CA 94720, USA
[26] Universidad de Buenos Aires, 1428 Buenos Aires, Argentina
[27] Harvard University, Cambridge, MA 02138, USA
[28] Weizmann Institute, 76100 Rehovot, Israel
[29] Universitaet Hamburg, 22761 Hamburg, Germany
[30] Universitaet Heidelberg, 69117 Heidelberg, Germany
[31] University of Arizona, Tucson, AZ 85719, USA
[32] IPPP, University of Durham, Durham DH1 3LE, UK
[33] Universitaet Mainz, 55099 Mainz, Germany
[34] MIT, Cambridge, MA 02139, USA
[35] University of Chicago, Chicago, IL 60637, USA
[36] Fermi National Accelerator Laboratory, Batavia, IL 60510, USA
[37] Indiana University, Bloomington, IN 47405, USA
[38] University of California, Davis, CA 95616, USA
[39] Johns Hopkins University, Baltimore, MD 21218, USA

[40] INFN and University of Pisa, 56127 Pisa, Italy
[41] Texas A & M University, College Station, TX 77843, USA
[42] INFN and University of Calabria, 87036 Rende, Italy
[43] Brown University, Richmond, RI 02912, USA
[44] Yale University, New Haven, CT 06511, USA
[45] CEA Saclay, 91191 Gif-sur-Yvette, France
[46] University of Illinois, Chicago, IL 60607, USA
[47] University of California, Berkeley, CA 94720, USA
[48] *Present address:* University at Buffalo, State University of New York, Buffalo, NY 14214, USA
[49] *Present address:* Humboldt University, 10099 Berlin, Germany

Received: 4 December 2013 / Accepted: 17 February 2014 / Published online: 21 March 2014

Abstract This report of the BOOST2012 workshop presents the results of four working groups that studied key aspects of jet substructure. We discuss the potential of first-principle QCD calculations to yield a precise description of the substructure of jets and study the accuracy of state-of-the-art Monte Carlo tools. Limitations of the experiments' ability to resolve substructure are evaluated, with a focus on the impact of additional (pile-up) proton proton collisions on jet substructure performance in future LHC operating scenarios. A final section summarizes the lessons learnt from jet substructure analyses in searches for new physics in the production of boosted top quarks.

1 Introduction

With a centre-of-mass energy of 7 TeV in 2010 and 2011 and of 8 TeV in 2012 the LHC has pushed the energy frontier well into the TeV regime. Another leap in energy is expected with the start of the second phase of operation in 2014, when the centre-of-mass energy is to be increased to 13–14 TeV. For the first time experiments produce large samples of W and Z bosons and top quarks with a transverse momentum p_T that considerably exceeds their rest mass m ($p_T \gg m$). The same is true also for the Higgs boson and, possibly, for as yet unknown particles with masses near the electroweak scale. In this new kinematic regime, well-known particles are observed in unfamiliar ways. *Classical* reconstruction algorithms that rely on a one-to-one jet-to-parton assignment are often inadequate, in particular for hadronic decays of such boosted objects.

A suite of techniques has been developed to fully exploit the opportunities offered by boosted objects at the LHC. Jets are reconstructed with a much larger radius parameter to capture the energy of the complete (hadronic) decay in a single jet. The internal structure of these *fat* jets is a key signature to identify boosted objects among the abundant jet production at the LHC. Many searches use a variety of recently proposed substructure observables. Jet grooming techniques[1] improve the resolution of jet substructure measurements, help to reject background, and increase the resilience to the impact of multiple proton-proton interactions.

In July 2012 IFIC Valencia organized the 2012 edition [4] of the BOOST series of workshops, the main forum for the physics of boosted objects and jet substructure[2]. Working groups formed during the 2010 and 2011 workshops prepared reports [9,10] that provide an overview of the state of the field and an entry point to the now quite extensive literature and present new material prepared by participants. In this paper we present the report of the working groups set up during BOOST2012. Each contribution addresses an important aspect of jet substructure analysis as a tool for the study of boosted objects at the LHC.

A good understanding of jet substructure is a prerequisite to further progress. Predictions of jet substructure based on first-principle, analytical calculations may provide a more precise description of jet substructure and allow deeper insight. However, resummation of the leading logarithms in this case is notoriously difficult and the predictions may be subject to considerable uncertainties. In fact, one might ask:

– Can jet substructure be predicted by first-principle QCD calculations and compared to data in a meaningful way?

The findings of the working group that was set up to evaluate the limitations and potential of the most popular approaches are presented in Sect. 2.

While progress toward analytical predictions continues, searches for boosted objects that employ jet substructure rely

[1] We refer to three related techniques as jet grooming: filtering [1], trimming [2] and pruning [3]. Unless stated otherwise, all studies in this paper of these techniques apply a common set of parameters that is widely used in the community.

[2] Previous BOOST workshops took place at SLAC (2009, [5]), Oxford University (2010, [6]) and Princeton University University (2011, [7]). BOOST2013 [8] was organized by the University of Arizona from August $12th$ to $16th$.

[a] e-mail: marcel.vos@ific.uv.es

on the predictions of mainstream Monte Carlo models. It is therefore vital to answer this question:

– How accurately is jet substructure described by state-of-the-art Monte Carlo tools?

The BOOST2010 report [9] provided a partial answer, based on pre-LHC tunes of several popular leading-order generators. After the valuable experience gained in the first three years of operation of the LHC, it seems appropriate to revisit this question in Sect. 3.

A further potential limitation to the performance of jet substructure analyses is the level to which the detector response can be understood and modelled. Again, the first years of LHC operation have provided valuable experience on how well different techniques work in a realistic experimental environment. In particular, the impact of multiple proton-proton interactions (pile-up) on substructure measurement has been evaluated exhaustively and mitigation schemes have been developed. Anticipating a sharp increase in the pile-up activity in future operating scenarios of the LHC, one might worry that in the future the detector performance might be degraded considerably for the sensitive substructure analyses. A third working group was therefore given the following charge:

– How does the impact of additional proton proton collisions limit the performance of jet substructure analysis at the LHC, now and in future operating scenarios?

Section 4 presents the contributions regarding jet reconstruction performance under extreme contributions, with up to 200 additional proton-proton collisions in each bunch crossing. We present the expectations for fake jet rates and the impact of pile-up on jet mass measurements under these conditions.

In the first years of operation of the LHC several groups in ATLAS and CMS have deployed techniques specifically developed for the study of boosted objects in several analyses. Jet substructure analysis has become an important tool in many searches for evidence for new physics. In Sect. 5 we present the lessons learnt in several studies of boosted top quark production that have been the first to apply these techniques and answer the following question:

– How powerful a tool is jet substructure analysis in studies of boosted top production, and how can it be made even more powerful?

We hope that the answers to the above questions prepared by the working groups may shed some light on this rapidly evolving field.

2 Measurements and first-principle QCD predictions for jet substructure

Section prepared by the Working Group: 'Predictions and measurements of jet substructure observables', A. Davison, A. Hornig, S. Marzani, D.W. Miller, G. Salam, M. Schwartz, I. Stewart, J. Thaler, N.V. Tran, C. Vermilion, J. Walsh

The internal structure of jets has traditionally been characterized in jet shape measurements. A detailed introduction to the current theoretical understanding and of the calculations needed for observables that probe jet substructure is provided in last year's BOOST report [10]. Here, rather than give a comprehensive review of the literature relevant to the myriad of developments, we focus on the progress made in the last year in calculations of jet substructure at hadron colliders. Like the Tevatron experiments ATLAS and CMS have performed measurements of the energy flow within the jet [11,12]. Both collaborations have moreover performed dedicated jet substructure measurements on jets reconstructed with a large radius parameter ($R = 0.8 - 1.2$, as opposed to the usual $R = 0.4 - 0.7$). These experimental results are briefly reviewed before we introduce analytical calculations and summarize the status of the two main approaches.

2.1 Jet substructure measurements by ATLAS

The first measurement of jet mass for large-radius jets, with radii of 1.0 and 1.2, and several substructure observables was performed by ATLAS on data from the 2010 run of the LHC [13]. These early studies include also a first measurement of the jet mass distribution for filtered [1] Cambridge-Aachen jets. A number of further jet shapes were studied with the same data set in Reference [14]. These early studies were crucial to establish the ability of the experiment to resolve jet substructure and to validate the Monte Carlo description of jet substructure. They are moreover unique, as the impact of pile-up could be trivially avoided by selecting events with a single primary vertex. The results, fully corrected for detector effects, are available for comparison to calculations.

Since then, the ATLAS experiment has performed a direct and systematic comparison of the performance of several grooming algorithms on inclusive jet samples, purified samples of high-p_T W bosons and top quarks, and Monte Carlo simulations of boosted W and top-quark signal samples [15]. The parameters of large-radius ($R = 1.0$) trimmed [2], pruned [3] and mass-drop filtered jet algorithms were optimized in the context of Standard Model measurements and new physics searches using multiple performance measures, including efficiency and jet mass resolution. The impact of pile-up on the jet mass measurement is studied quantitatively. The mitigating effect of trimming and mass-drop filtering is established in events with up to 15 additional proton proton interactions.

For a subset of the jet algorithms tested, dedicated jet energy scale and mass scale calibrations were derived and systematic uncertainties evaluated for a wide range of jet transverse momenta. Relative systematic uncertainties were obtained by comparing ratios of track-based quantities to calorimeter-based quantities in the data and MC simulation. In situ measurements of the mass of jets containing boosted hadronically decaying W bosons further constrain the jet mass scale uncertainties for this particular class of jets to approximately $\pm 1\%$.

2.2 Jet substructure measurements by CMS

The CMS experiment measured jet mass distributions with approximately 5 fb^{-1} of data at a center-of-mass energy of $\sqrt{s} = 7$ TeV [16]. The measurements were performed in several p_T bins and for two processes, inclusive jet production and vector boson production in association with jets. For inclusive jet production, the measurement corresponds to the average jet mass of the highest two p_T jets. In vector boson plus jet ($V+$ jet) production the mass of the jet with the highest p_T was measured. The measurements were performed primarily for jets clustered with the anti-k_t algorithm with distance parameter $R = 0.7$ (AK7). The mass of ungroomed, filtered, trimmed, and pruned jets are presented in bins of pt. Additional measurements were performed for anti-k_t jets with smaller and larger radius parameter ($R = 0.5$, 0.8), after applying pruning [3] and filtering [1] to the jet, and for Cambridge-Aachen jets with $R = 0.8$ and $R = 1.2$.

The jet mass distributions are corrected for detector effects and can be compared directly with theoretical calculations or simulation models. The dominant systematic uncertainties are jet energy resolution effects, pile-up, and parton shower modeling.

The study finds that, for the grooming parameters examined, the pruning algorithm is the most aggressive grooming algorithm, leading to the largest average reduction of the jet mass with respect to the original jet mass. Due to this fact, CMS also finds that the pruning algorithm reduced the pile-up dependence of the jet mass the most of the grooming algorithms.

The jet mass distributions are compared against different simulation programs: PYTHIA 6 [17,18] (version 424, tune Z2), HERWIG++ [19,20] (version 2.4.2, tune 23), and PYTHIA 8 (version 145, tune 4C), in the case of inclusive jet production. In general the agreement between simulation and data is reasonable although HERWIG++ appears to have the best agreement with the data for more aggressive grooming algorithms. The $V+$ jet channel appears to have better agreement overall than the inclusive jets production channel which indicates that quark jets are modeled better in simulation. The largest disagreement with data comes from the low jet mass region, which is more affected by pile-up and soft QCD effects.

The jet energy scale and jet mass scale of these algorithms were validated individually. The jet energy scale was investigated in MC simulation, and was found to agree with the ungroomed energy scale within 3%, which is assigned as an additional systematic uncertainty. The jet mass scale was investigated in a sample of boosted W bosons in a semileptonic $t\bar{t}$ sample. The jet mass scale derived from the mass of the boosted W jet agrees with MC simulation within 1%, which is also assigned as a systematic uncertainty.

2.3 Analytical predictions for jet substructure

Next-to-leading order (NLO) calculations in the strong coupling constant have been performed for multi-jet production, even in association with an electro-weak boson. This means that substructure observables, such as the jet mass, can be computed to NLO accuracy using publicly available codes [21,22]. However, whenever multiple scales, e.g. a jet's transverse momentum and its mass, are involved in a measurement, the prediction of the observables will contain logarithms of ratios of these scales at each order in perturbation theory. These logarithms are so important for jet shapes that they qualitatively change the shapes as compared to fixed order. Resummation yields a more efficient organization of the perturbative expansion than traditional fixed-order perturbation theory. Accurate calculations of jet shapes are impossible without resummation. In general one can moreover interpolate between, or *merge*, the resummed and fixed-order result.

In resummation techniques the perturbative expansion of cross-sections for generic observables v is schematically organized in the form[3]

$$\sigma(v) = \int_0^v dv' \frac{d\sigma}{dv'} = \sum_{\substack{\text{partonic} \\ \text{configurations} \\ \delta}} \sigma_0^{(\delta)} g_0^{(\delta)}(\alpha_s) e^{\beta}, \qquad (1)$$

$$\beta = L g_1^{(\delta)}(\alpha_s L) + g_2^{(\delta)}(\alpha_s L) + \alpha_s g_3^{(\delta)}(\alpha_s L) + \ldots \qquad (2)$$

where $\sigma_0 = \sum \sigma_0^{(\delta)}$ is the corresponding Born cross-section, $L = \ln v$ is a logarithm of the observable in question and the

[3] The actual form of Eq. (1) is in general rather complex. For more than three hard partons it involves a non-trivial matrix structure in colour space. Moreover, the actual form of the constant terms $g_0(\alpha_s)$ depends on the flavor of the jet under consideration.

functions $g_i^{(\delta)}$ resum the large logarithmic contributions to all orders in perturbation theory[4].

The notation used in traditional fixed-order perturbation theory refers to the lowest-order calculation as leading order (LO) and higher-order calculations as NLO, next-to-next to leading order (NNLO), and so on (with N^{n}LO referring to the $\mathcal{O}(\alpha_s^n)$ correction to the LO result). When organized instead in resummed perturbation theory as in Eq. (1), the lowest order, in which only the function $g_1^{(\delta)}$ is retained, is referred to as leading-log (LL) approximation. Similarly, the inclusion of all $g_i^{(\delta)}$ with $1 \leq i \leq k+1$ and of g_0 up to order α_s^{k-1} gives the next^k-to-leading log approximation to $\ln\sigma$; this corresponds to the resummation of all the contributions of the form $\alpha_s^n L^m$ with $2(n-k)+1 \leq m \leq 2n$ in the cross section σ. This can be extended to $2(n-k) \leq m \leq 2n$ by also including the order α_s^k contribution to $g_0^{(\delta)}$.

Typical Monte Carlo event generators such as PYTHIA, HERWIG++ and Sherpa [23] are based on calculations to Leading-Log precision. Next-to-Leading-Log (NLL) accuracy has also been achieved for some specific observables, but it is difficult to say whether this can be generally obtained. Analytic calculations provide a way of obtaining precise calculations for jet substructure. Multiple observables have been resummed (most often at least to NLL but not uncommonly to NNLL and as high as NNNLL accuracy for a few cases) and others are actively being studied and calculated in the theory community.

Often for observables of experimental interest, non-global logarithms (NGLs) arise [24], in particular whenever a hard boundary in phase-space is present (such as a rapidity cut or a geometrical jet boundary). These effects enter at NLL level and therefore modify the structure of the function $g_2^{(\delta)}$ in Eq. (1). Until very recently [25], the resummation of NGLs was confined to the limit of large number of colours N_C [24,26,27].

Moreover, we should stress that another class of contributions, usually referred to as clustering logarithms, affects the $g_2^{(\delta)}$ series of Eq. (1) if an algorithm other than anti-k_t is used to define the jets [28,29]. The analytic structure of these clustering effects has been recently explored in Ref. [30,31] for the case of Cambridge-Aachen and k_t algorithms.

Furthermore, recent studies have shown that strict collinear factorization is violated if the observable considered is not sufficiently inclusive [32,33]. As a consequence, coherence-violating (or super-leading) logarithms appear, which further complicate the resummation of certain observables. These contributions affect not only non-global dijets observables, such as the fraction of dijets events with a central gap [34,35], but also some classes of global event shapes [36].

[4] In the following we concentrate on the case of jet masses with a cut on the jet p_T. In this case $L = \ln m_J^2/p_T^2$ and $\sigma(v)$ in Eq. (1) is the integrated (cumulative) distribution for $m_J^2 < v p_T^2$.

Of course, to fully compare to data one needs to incorporate the effects of hadronization and multi-particle interactions (MPI). Progress on this front has also been made, both in purely analytical approaches (especially for hadronization effects [37]) and in interfacing analytical results with parton showers that incorporate these effects.

The two main active approaches to resummation are referred to as traditional perturbative QCD resummation (pQCD) and Soft Collinear Effective Theory (SCET). They describe the same physical effects, which are captured by the Eqs. 1 and 2. However, the techniques employed in pQCD and SCET approaches often differ. Calculations in pQCD exploit factorization and exponentiation properties of QCD matrix elements and of the phase-space associated to the observable at hand, in the soft or collinear limits. The SCET approach is based on factorization at the operator level and exploits the renormalization group to resum the logarithms. The two approaches also adopt different philosophies for the treatment of NGLs. A more detailed description of these differences is given in the next Sections.

2.4 Resummation in pQCD

Jet mass was calculated in pQCD in [38]. A more extensive study can be found in Ref. [39] where the jet mass distribution for Z+jet and inclusive jet production, with jets defined with the anti-k_t algorithm, were calculated at NLL accuracy and matched to LO. In particular, for the Z+jet case, the jet mass distribution of the highest p_T jet was calculated whereas for inclusive jet production, essentially the average of jet mass distributions of the two highest p_T jets was calculated. For the Z+jet case, one has to consider soft-wide angle emissions from a three hard parton ensemble, consisting of the incoming partons and the outgoing hard parton. For three or fewer partons, the colour structure is trivial. Dijet production on the other hand involves an ensemble of four hard partons and the consequent soft wide-angle radiation has a non-trivial colour matrix structure. The rank of these matrices grows quickly with the number of hard partons, making the calculations for multi jet final states a formidable challenge[5].

The jet mass is a non-global observable and NGLs of m_J/p_T for jets with transverse momentum p_T are induced. Their effect was approximated using an analytic formula with coefficients fit to a Monte Carlo simulation valid in the large N_C limit, obtained by means of a dipole evolution code [24]. It was found that in inclusive calculations[6] the effects of both

[5] The colour structure of soft gluon resummation in a multi jet environment has been studied in [40,41] and resummed calculations for the case of five hard partons in the context of jet production with a central jet veto can be found in Refs. [34,35,42,43]

[6] We refer to inclusive calculations if no requirements were made on the number of additional jets in the selection of the event.

the soft wide-angle radiation and the NGLs, both of which affect the $g_2^{(\delta)}$ series in Eq. (1), play a relevant role even at relatively small values of jet radius such as $R = 0.6$ and hence in general cannot be neglected.

A restriction on the number of additional jets could be implemented, for instance, by vetoing additional jets with $p_T > p_T^{\text{cut}}$. The presence of a jet veto modifies the calculation in several ways. First of all, it affects the argument of the non-global logarithms: $\ln^n(m_J^2/p_T^2) \to \ln^n(m_J^2/(p_T\, p_T^{cut}))$. Thus p_T^{cut} could, in principle, be used to tame the effect of NGLs. However, if the veto scale is chosen such that $p_T^{\text{cut}} \ll p_T$, logarithms of this ratio must be also resummed. Depending on the specific details of the definition of the observable, this further resummation can be affected by a new class of NGLs [44,45].

An obstacle to inclusive predictions in the number of jets is that the constant term $g_0^{(\delta)}$ in Eq. 1 receives contributions from topologies with higher jet multiplicities that are not related to any Born configurations. For instance, the jet mass in the Z+jet process would receive contributions from Z+2jet configurations, which are clearly absent in the exclusive case. The full determination of the constant term to $\mathcal{O}(\alpha_s)$ and the matching to NLO is ongoing.

2.5 Resummation in SCET

There have been several recent papers in SCET directly related to substructure in hadron collisions[7]. Reference [46] discusses the resummation of jet mass by expanding around the threshold limit, where (nearly) all of the energy goes into the final state jets. Expanding around the threshold limit has proven effective for other observables, see Ref. [47] and references in Ref. [46]. The large logarithms for jet mass are mainly due to collinear emission within the jet and soft emission from the recoiling jet and the beam. These same logarithms are present near threshold and the threshold limit automatically prevents additional jets from being relevant, simplifying the calculation. The study in Ref. [46] performs resummation at the NNLL level, but does not include NGLs. Instead, their effect is estimated and found to be subdominant in the peak region, where other effects, such as nonperturvative corrections, are comparable. Thus NGLs could be safely ignored where the calculation was most accurate.

An alternative approach using SCET is found in Ref. [48]. Beam functions are used to contain the collinear radiation from the beam remnants. The jet mass distribution in Higgs+1jet events is studied via the factorization formula for the 1-jettiness event shape [49], that is calculated to NNLL accuracy. Using 1-jettiness, the jet boundaries are defined by the distance measure used in 1-jettiness itself, instead of a more commonly employed jet algorithm, although generalizations to arbitrary jet algorithms are possible.

For a single jet in hadron collisions, 1-jettiness can be used as a means to separate the in-jet and out-of-jet radiation (see for a review the BOOST2011 report [10]). The observable studied in Ref. [48] is separately differential in the jet mass and the beam thrust. The in-jet component is related to the jet mass, and can be converted directly, up to power corrections that become negligible for higher p_T (up to about 3% for $p_T = 300$ GeV in the peak of the distribution of the in-jet contribution to 1-jettiness which is smaller than NNLL uncertainties). The beam thrust[8] is a measure of the out-of-jet contributions, equivalent to a rapidity-weighted veto scale p_{cut} on extra jets. The calculation can be made exclusive in the number of jets by making the out-of-jet contributions small. Where Ref. [46] ensures a fixed number of jets by expanding around the threshold limit, Ref. [48] includes an explicit jet veto scale.

Exclusive calculations in the number of jets avoid some of the issues mentioned in Sect. 2.4. An important property of 1-jettiness is that, when considering the sum of the in- and out-of-jet contributions, no NGLs are present, and when considering these contributions separately, only the ratio p_{cut}/m_J of these two scales is non-global. A smart choice of the veto scale may then allow to minimize the NGL and make the resummation unnecessary. This corresponds to the NGLs discussed in Sect. 2.4 that are induced in going from the inclusive to the exclusive case. These are the only NGLs present; the additional NGLs of the measured jet p_T to their mass discussed for the observable of Sect. 2.4 are absent in this case. By using an exclusive observable, with an explicit veto scale, NGLs are controlled. For comparison with inclusive jet mass measurements, such as those discussed in Sects. 2.1 and 2.2, the uncertainty associated with the veto scale can be estimated in a similar fashion as the NGL estimate in Ref. [46].

It was argued in Ref. [48] that the NGLs induced by imposing a veto on both the p_T and jet mass are smaller than the resummable logarithms of the measured jets over a range of veto scales. In contrast, in the inclusive case the corresponding p_T value that appears in the NGLs is of the order of the measured jet p_T (since all values less than this are allowed), making it a large scale and the NGLs as large as other logarithms. For a fixed veto cut, it was argued that the effect of these NGLs (at least of those that enter at the first non-trivial order, $\mathcal{O}(\alpha_s^2)$), can be considered small enough to justify avoiding resummation for a calculation up to NNLL accuracy for $1/\sqrt{8} < m_J^{\text{cut}}/p_{\text{cut}} < \sqrt{8}$ (cf. Ref. [53]) in the peak region where a majority of events lie. It is also worth noting that the effect of normalizing the distribution by the total

[7] We consider here only research made publicly available at the time of BOOST 2012 or soon after.

[8] The resummation of beam thrust is analogous to that of thrust in e^+e^- collisions [50–52].

rate up to a maximum m_J^{cut} and p_{cut} has several advantages and in particular has a smaller perturbative uncertainty than the unnormalized distribution, in addition to having smaller experimental uncertainties.

We also note that while jet mass is now theoretically the best understood substructure observable, experiments often apply much more complex techniques in their analyses of the data. There has also been progress in understanding more complicated measurements using SCET, and in particular a calculation of the signal distribution in $H \to b\bar{b}$ was performed in Ref. [54]. While it is probably fair to say that our theoretical understanding (or at least the numerical accuracy) of such measurements are currently not at the same level as that of the jet mass, this is a nice demonstration that reasonably accurate calculations of realistic substructure measurements can be performed with the current technologies and that it is not unreasonable to expect related studies in the near future.

2.6 Discussion and recommendations for further substructure measurements

We have presented a status report for the two main approaches to the resummation of jet substructure observables, with a focus on their potential to predict the jet invariant mass at hadron colliders. In both approaches recent work has shown important progress

We hope that providing predictions beyond the accuracy of parton showers may help both discovery and measurement. Beyond the scope of improving our understanding of QCD, gaining intuition for which treatments work best is an important step towards adopting such predictions as an alternative to parton showers. Non-perturbative corrections like hadronization are more complicated at the LHC due to the increased colour correlations. Entirely new perturbative and semi-perturbative effects such as multiple-particle interactions appear. Monte Carlo simulations suggest that these have a significant impact.

The treatments of non-perturbative corrections and NGLs are often different in pQCD and SCET[9] and this leads to slight differences in which measurements are best suited for comparison to predictions. The first target for the next year should be a phenomenological study of the jet mass distribution in Z+jet, for which we encourage ATLAS and CMS measurements. Ideally, since the QCD and SCET literature have emphasized a difference in preference for inclusive or exclusive measurements (in the number of jets), both should be measured to help our understanding of the two techniques.

[9] We have focused on differences in our discussion, but typically both communities have the option to adopt the treatments commonly employed in the other community. That is, the treatments typically utilized are not features inherent to the approach.

The importance of boosted-object taggers in searches for new physics will increase strongly in the near future in view of the higher-energy and higher-luminosity LHC runs. However, the theoretical understanding of these tools is in its infancy. Analytic calculations must be performed in order to understand the properties of the different taggers and establish which theoretical approaches (MC, resummation or even fixed order) are needed to accurately compute these kind of observables[10].

3 Monte Carlo generators for jet substructure observables

Section prepared by the Working Group: 'Monte Carlo predictions for jet substructure', *A. Arce*, *D. Bjergaard, A. Buckley, M. Campanelli*, *D. Kar*, *K. Nordstrom.*

In order to use boosted objects and substructure techniques for measurements and searches, it is important that Monte Carlo generators describe the jet substructure with reasonable precision, and that variations due to the choice of parton shower models and their parameters are characterized and understood. We study jet mass, before and after several jet grooming procedures, a number of popular jet substructure observables, colour flow and jet charge. For each of these we compare the predictions of several parton shower and hadronisation codes, not only in *signal-like* topologies, but also in background or calibration samples.

3.1 Monte Carlo samples and tools

Three processes in pp collisions are considered at $\sqrt{s} = 7$ TeV: semileptonic $t\bar{t}$ decays, boosted semileptonic $t\bar{t}$ decays, and $(W^{\pm} \to \mu\nu)$+jets. These processes provide massive jets coming from hadronic decays of a colour-neutral boson as well as jets from heavy and light quarks.

Like Z+jets, the $(W \to \mu\nu)$+jets process provides a well-understood source of quarks and gluons, and additionally allows an experimentally accessible identification ("away-side-tag") of the charge of the leading jet. Assuming that the charge of this jet is opposite to the muon's charge leads to the same charge assignment as a conventional parton matching scheme in approximately 70% of simulated events in leading order Monte Carlo simulation; in the remaining 30% of cases, the recoiling jet matches a (charge-neutral) gluon.

[10] Before the completion of this manuscript, two papers appeared [55,56] which perform analytic resummed calculations for boosted-object methods, such as trimming, pruning and mass drop, and energy correlations were computed and used for quark and gluon discrimination in Ref. [57].

The selection of t, $W^{\pm}$, and quark jet candidates for the distributions compared below include event topologies that can be realistically collected in the LHC experiments, with typical background rejection cuts, so that these studies, based on simulation, could be reproduced using LHC data.

The most commonly used leading order Monte Carlo simulation codes are the PYTHIA and HERWIG families. Here, predictions from the Perugia 2011 [58,59] tune with CTEQ5L [60] parton density function (PDF) and corresponding NOCR tunes of PYTHIA6 [61,17] (version 6.426), tune 4C [62] with CTEQ6L1 PDF [63] of the newer C++ PYTHIA8 generator [18] (version 8.170), and the LHC-UE-EE-4 [64] tune of HERWIG++ [20,65] (version 2.6.1) with CTEQ6L1 PDF are compared. The default parameter tune of the next-to-leading order (NLO) parton shower model implemented in SHERPA [23] (version 1.4.2) with CT10 PDF [66] is also included in comparisons. The PYTHIA6 generator with the Perugia2011 tune is taken as a reference in all comparisons. For each generator, tune and process one million proton-proton events at $\sqrt{s} = 7$ TeV are produced.

The analysis relies on the FastJet 3.0.3 package [67,68] and Rivet analysis framework [69]. All analysis routines are available on the conference web page [70]. In the boosted semileptonic $t\bar{t}$ analysis, large-radius jets were formed using the anti-k_t algorithm [71] with a radius parameter of 1.2 using all stable particles within pseudorapidity $|\eta| < 4$. The jets are selected if they passed the following cuts: $p_T^{\text{jet}} > 350$ GeV, 140 GeV $< m^{\text{jet}} < 250$ GeV. Only the leading and subleading jets were selected if more than two jets passed the cuts. The subjets were formed using the Cambridge-Aachen algorithm [72,73] with radius 0.3.

3.2 Jet mass

The jet mass distribution for the leading jet in the boosted semi-leptonic $t\bar{t}$ sample is shown in Fig. 1. The parton shower models in PYTHIA6, PYTHIA8, HERWIG++ and SHERPA yield significantly different predictions. Important differences are observed in the location and shape of the top quark mass peak. The largest deviations of the normalized cross section in a given jet mass bin amount to approximately 20%. Much better agreement is obtained for predictions with different tunes of a single generator.

The effect of different grooming techniques on jet mass is also shown in Fig. 1. For filtering, three hardest subjets with $R^{sub} = 0.3$ are used. The trimming uses all subjets over 3% of p_T^{jet} and $R^{sub} = 0.3$. For pruning, $z = 0.1$ and $D = m^{jet}/p_T^{jet}$ is used. As expected, a much narrower top quark mass peak is obtained, with a particularly strong reduction of the high-mass tail. The grooming procedure improves the agreement among the different Monte Carlo tools, as expected from previous Monte Carlo studies with a more limited set of generators [9] and comparison with data [13].

3.3 Jet substructure observables

We investigate the spread among generators for a number of other substructure observables on the market:

– The Angular Correlation Function [74] measures the ΔR distance scale of the radiation in the jet:

$$\mathcal{G}(R) = \frac{1}{\sum p_{T,i}\, p_{T,j} \Delta R_{i,j}^2} \sum p_{T,i}\, p_{T,j} \Delta R_{i,j}^2 \Theta \times (R - \Delta R_{i,j})$$

where the sum runs over all pairs of particles in the jet, and $\Theta(x)$ is the Heaviside step function. The Angular Structure Function is defined as the derivative:

$$\Delta\mathcal{G}(R) = \frac{d \log \mathcal{G}(R)}{d \log R}$$

A peak in $\Delta\mathcal{G}(R)$ at a given ΔR indicates that radiation in the jet separated by ΔR contributes significantly to the jet mass. Only prominent peaks, with prominence $h > 4$ are retained[11]. The variable r_{1*} studied here corresponds to the location of the first peak in the angular structure function. Jets with a total number of prominent peaks n_p greater than 1 are discarded.

– N-subjettiness [75,76] measures how much of a jet's radiation is aligned along N subjet axes in the plane formed by the rapidity y and azimuthal angle ϕ. It is defined as:

$$\tau_N = \frac{1}{\sum_k p_{T,k} R_{\text{jet}}^{\beta}} \sum_k p_{T,k} \min(\Delta R_{1,k}^{\beta}, \Delta R_{2_k}^{\beta}, \ldots)$$

where $\Delta R_{n,k}$ is the distance from k to the nth subjet axis in the $y - \phi$ plane, R_{jet} is the radius used for clustering the original jet, and β is an angular weighting exponent[12].

[11] The prominence of the highest peak is defined as its height and the prominence of any lower peak as the minimum vertical descent that is required in descending from that peak before ascending a higher, neighboring peak. In this analysis the derivatives are smoothed using a Gaussian in the numerator and an error function in the denominator, both with $\sigma = 0.06$.

[12] To improve the performance of N-subjettiness it is possible to use a k-means clustering algorithm to find (locally) optimal locations for the subjet axes. In this analysis $\beta = 1$ is used to find the subjet axes by reclustering with the k_t algorithm. The k-means clustering algorithm is run once, as with this angular weighting exponent it finds a local minimum immediately. No attempt is made to find the global minimum.

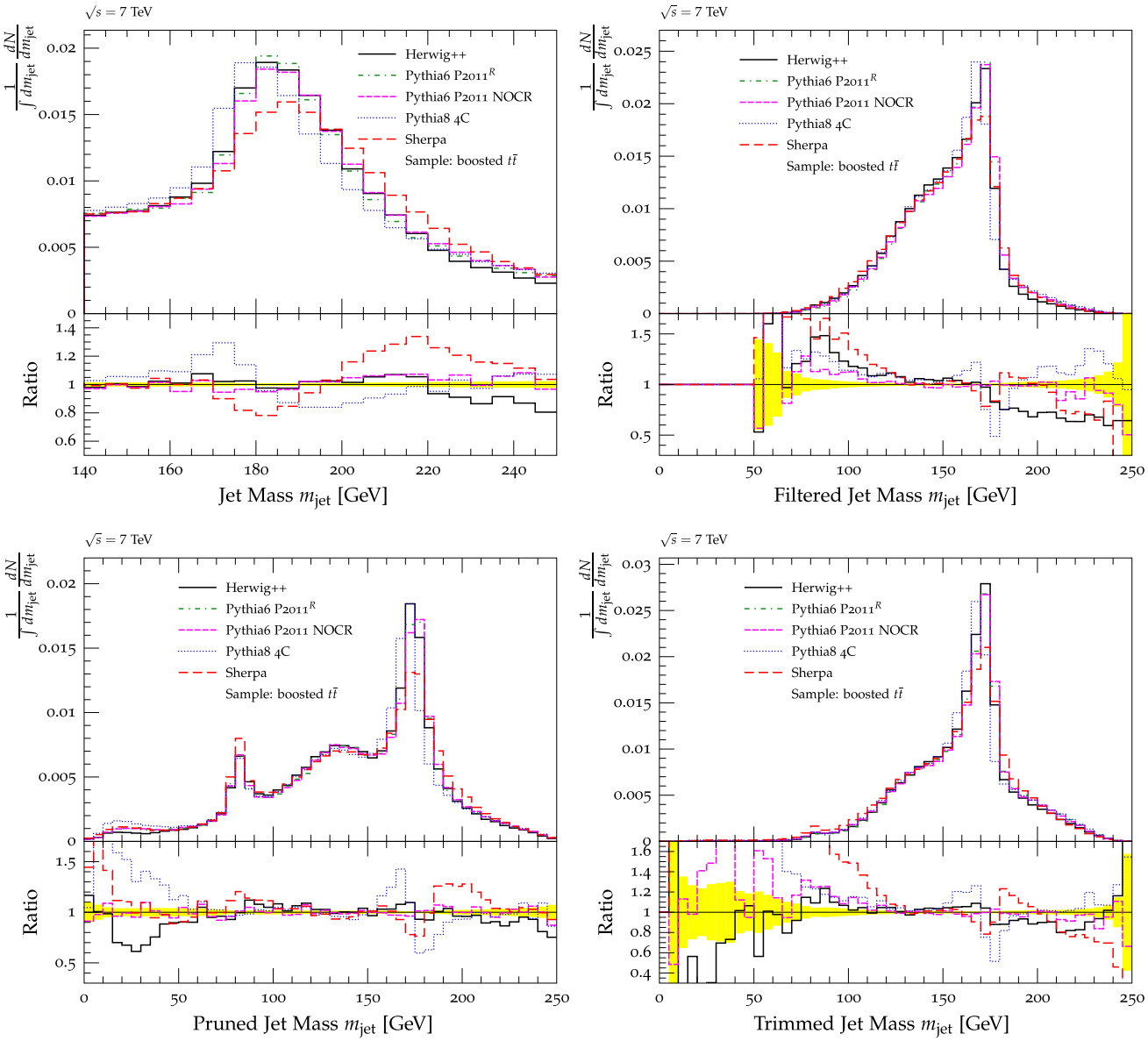

Fig. 1 Simulated jet invariant mass distribution for the leading jet in the boosted semileptonic $t\bar{t}$ event sample, before and after jet grooming. The ratio of the predictions of the different generators to that of the PYTHIA6 code with the Perugia 2011 tune is shown in the *insets*. The *yellow band* indicates the statistical error

– Angularity [77] introduces an adjustable parameter a that interpolates between the well-known event shapes thrust and jet broadening. Jet angularity is an IRC safe variable (for $a < 2$) that can be used to separate multijet background from jets containing boosted objects [78]. It is defined as:

$$\tau_a = \frac{1}{m_{jet}} \sum_{i \in jets} \omega_i \sin^a \theta_i (1 - \cos\theta_i)^{1-a}$$

where ω_i is the energy of a constituent of the jet.

– Eccentricity [79] of jets is defined by $1 - v_{\max}/v_{\min}$, where v_{max} and v_{min} are the maximum and minimum values of the variances of jet constituents along the principle and minor axes[13].

The comparison of the several Monte Carlo generators in Fig. 2 show that most models predict very similar behavior

[13] Eccentricity is strongly correlated with the planar flow, and it is a measure of jet elongation ranging from 0 for perfectly circularly symmetric jet shapes to 1 for infinitely elongated jet shapes. This is primarily useful for identifying high p_T merged jets.

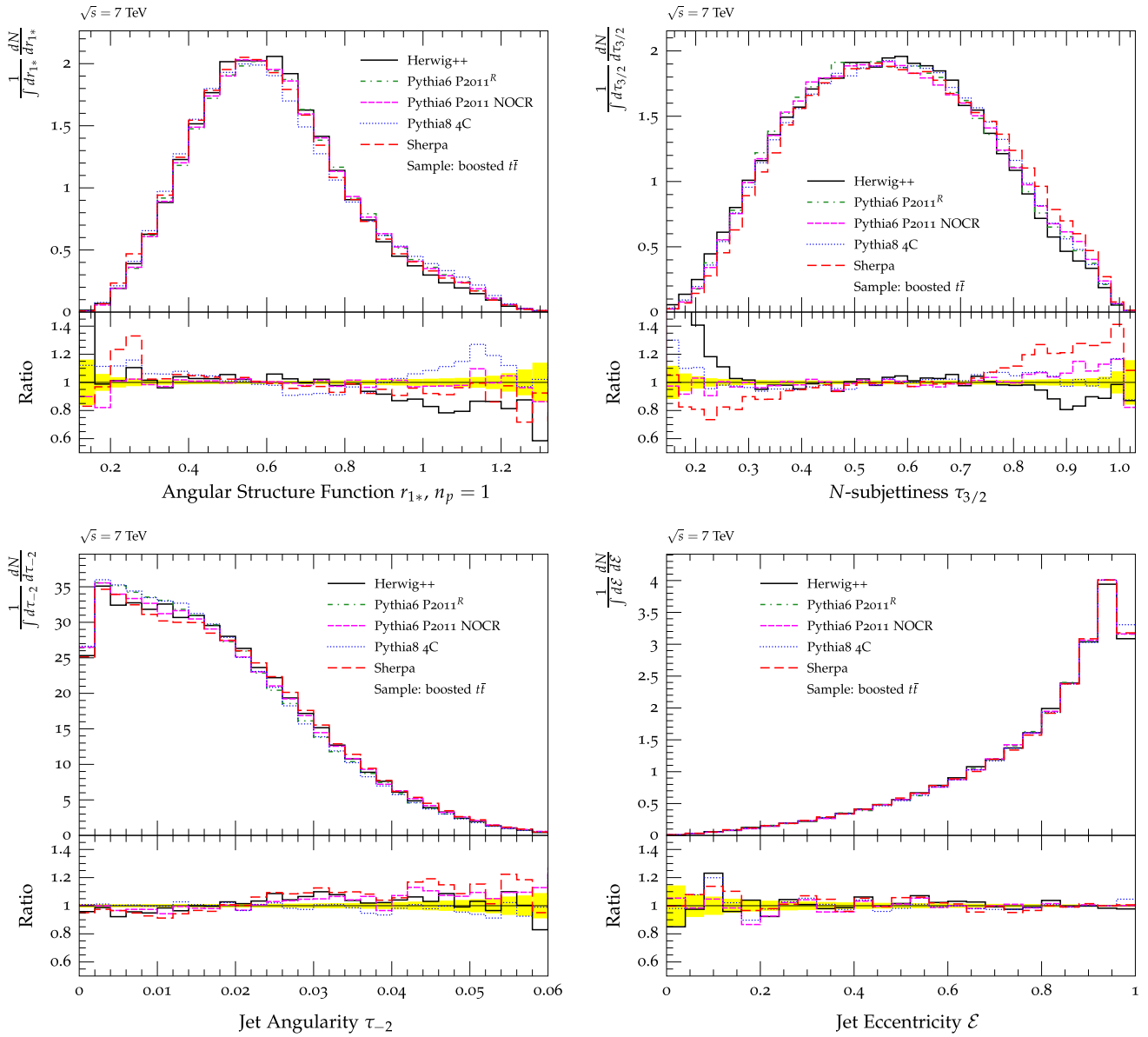

Fig. 2 Simulated distribution of four different measures of jet substructure. *Upper row* the angular structure function r_{1*} (*left*) and the ratio of 3-jettiness and 2-jettiness $\tau_{3/2}$ (*right*). *Lower row* the jet angularity τ_{-2} and the jet eccentry ϵ for leading jet of a boosted semileptonic top sample. The C-A algorithm is used in reclustering, as mentioned in the text. The ratio of the predictions of the different generators to that of the PYTHIA6 code with the Perugia 2011 tune is shown in the *insets*. The *yellow band* indicates the statistical error

for angularity, eccentricity and the ΔR scale of the peak in the $n_p = 1$ bin for the angular structure function. Deviations are typically below 10% for these observables. The harder jet mass distribution in SHERPA and the softer spectrum in PYTHIA8 are reflected in the edges of the $\tau_{3/2}$ distribution.

3.4 Colour flow

Colour flow observables offer a complementary way to probe boosted event topologies. Pull [80] is a p_T-weighted vector in $\eta-\phi$ space that is constructed so as to point from a given jet to its colour-connected partner(s). The pull is measured with respect to the other $W^{\pm}$ daughter jet. The W-boson is selected kinematically in 4-jet events with 2 b-quarks, and flavors are labelled using the highest p_T cone. In Fig. 3, the top left plot shows this variable for a background-like distribution. The comparisons demonstrate that Herwig produces a different colour flow structure.

Dipolarity [81] can distinguish whether a pair of subjets arises from a colour singlet source. In the top right plot of Fig. 3, the dipolarity predictions are seen to be similar for all models considered.

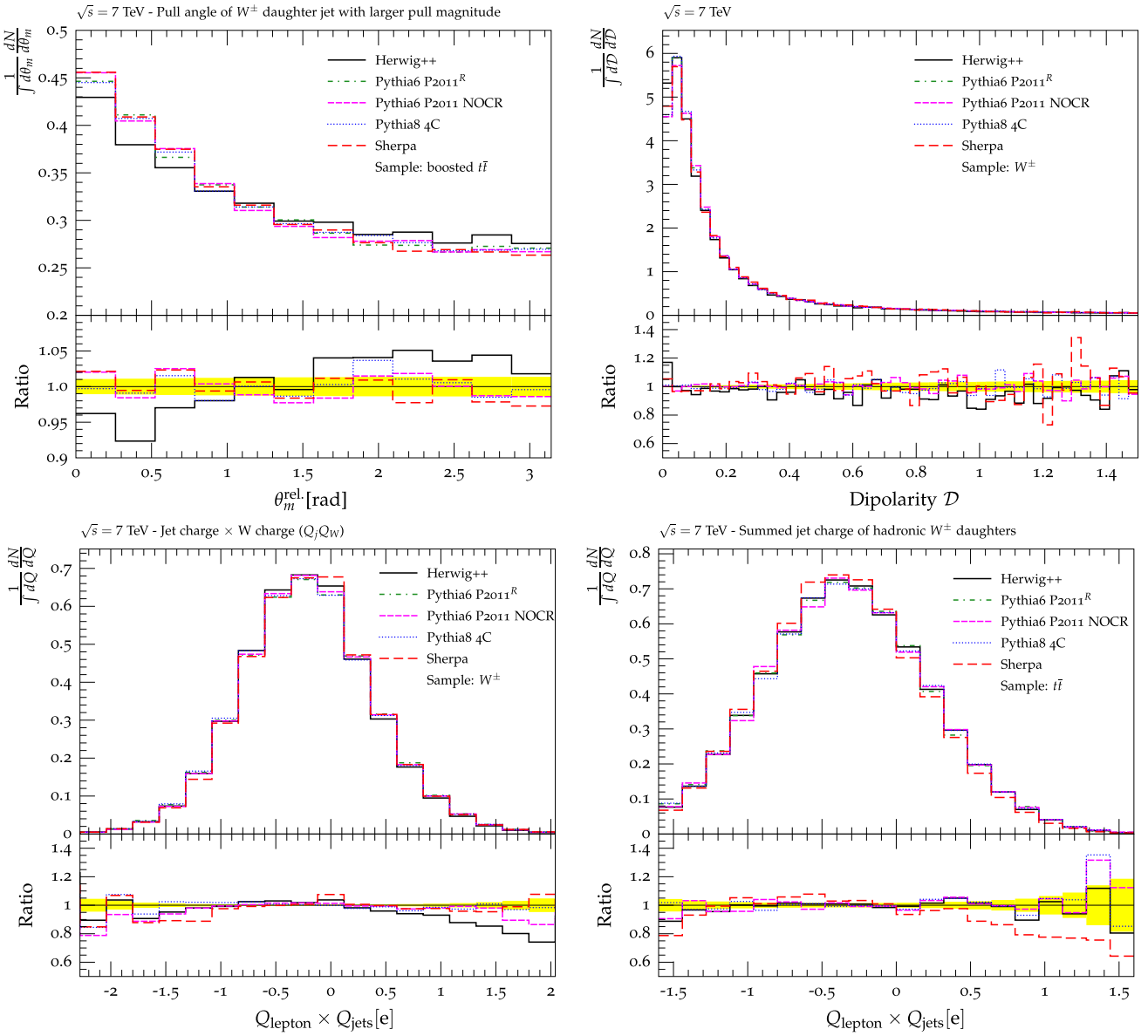

Fig. 3 *Upper row* Comparison of simulations for colour-flow observables: the pull angle of the leading jet attributed to the hadronic W decay in $t\bar{t}$ events, and the dipolarity of the leading jet produced in association with a leptonically decaying W. *Lower row* Comparison of simulated jet charge observables ($\kappa = 0.3$): the jet charge for the leading jet produced in association with a leptonically decaying W (*left panel*), and the sum of jet charge observables for the two jets attributed to the hadronic W decay in $t\bar{t}$ events (*right panel*). The ratio of the predictions of the different generators to that of the PYTHIA6 code with the Perugia 2011 tune is shown in the *insets*. The *yellow band* indicates the statistical error

3.5 Jet charge

Jet charge [82–84] is constructed in an attempt to associate a jet-based observable to the charge of the originating hard parton. The p_T-weighted jet charge

$$Q_j = \frac{1}{{p_T}_j^{\kappa}} \sum_{i \in T} q_i \times (p_T^i)^{\kappa}$$

is shown with $\kappa = 0.3$ in Fig. 3, using anti-kt 0.6 jets. The comparison displays the most relevant distributions for typical quark tagging and boson tagging analyses. Different MC models are seen to have very similar predictions for this observable too.

3.6 Summary

We have prepared the Rivet routines to evaluate the predictions of Monte Carlo generators for the internal structure of large area jets. The normalized predictions from several mainstream Monte Carlo models are compared. Several

aspects of jet substructure are evaluated, from basic jet invariant mass to colour flow observables and jet charge.

We find that for jet mass large variations are observed between the various MC models. However, for groomed jets the deviations between different model predictions are smaller. The differences between several recent tunes of the PYTHIA generator are much smaller. The MC model predictions are similar for N-subjettiness, angularity and eccentricity. The colour flow model recently implemented in HERWIG++ yields different predictions for colour flow observables than the models in other generators.

4 The impact of multiple proton-proton collisions on jet reconstruction

Section prepared by the Working Group: 'Jet substructure performance at high luminosity', P. Loch, D. Miller, K. Mishra, P. Nef, A. Schwartzman, G. Soyez.

The first LHC analyses exploring the experimental response to jet substructure demonstrated that the highly granular ATLAS and CMS detectors can yield excellent performance. They also confirmed the susceptibility of the invariant mass of large-radius jets, with a very large catchment area, to the energy flow from the additional proton-proton interactions that occur each bunch crossing. And, finally, they provided a first hint that jet grooming could be a powerful tool to mitigate the impact of pile-up. Since then, the LHC collaborations have gained extensive experience in techniques to correct for the impact of pile-up on jets. In this Section these tools are deployed in an extreme pile-up environment. We simulate pile-up levels as high as $\langle\mu\rangle = 200$, such as may be expected in a future high-luminosity phase of the LHC. We evaluate the impact on jet reconstruction, with a focus on the (substructure) performance.

4.1 Pile-up

Each LHC bunch crossing gives rise to a number of proton-proton collisions and typically the hard scattering (signal) interaction is accompanied by several additional pile-up proton-proton collisions. The total proton-proton cross-section is about $\sigma_{\text{tot}} = 98$ mb (inelastic $\sigma_{\text{inel}} = 72.9$ mb) at $\sqrt{s} = 7$TeV [85], and even slightly higher at $\sqrt{s} = 8$TeV in 2012. With a peak instantaneous luminosity of about 7.7×10^{33} cm^{-2} s^{-1} in 2012, the resulting average number of pile-up collisions reached $\langle\mu\rangle = 20$ at the highest intensities. The 2012 data set has a rather flat μ distribution extending from $\mu = 5$ to $\mu = 35$. In future LHC running even higher $\langle\mu\rangle$ are expected.

Pile-up manifests itself mostly in additional hadronic transverse momentum flow, which is generated by overlaid and statistically independent, predominantly soft proton-proton collisions that we refer to as "minimum bias" (MB). This diffuse transverse energy emission is superimposed onto the signal of hard scattering final state objects like particles and particle jets, and typically requires corrections, in particular for particle jets. In addition, pile-up can generate particle jets (*pile-up jets*). We distinguish two types: *QCD jets*, where the particles in the jet stem from a single MB collision, and *stochastic jets* that combine particles from different vertices in the high density particle flow.

4.2 Monte Carlo event generation

We model the pile-up with MB collisions at $\sqrt{s} = 8$TeV and a bunch spacing of 50 ns, generated with the PYTHIA Monte Carlo (MC) generator [86,87], with its 4C tune [62]. All inelastic, single diffractive, and double diffractive processes are included, with the default fractions as provided by PYTHIA(tune 4C).

Overall 100×10^6 MB events are available for pile-up simulation. The corresponding data are generated in samples of 25000 MB collisions, with the largest possibly statistical independence between samples, including new random seeds for each sample. To model pile-up for each signal interaction, the stable particles[14] generated in a number μ of MB collisions, with μ being sampled from a Poisson distribution around the chosen $\langle\mu\rangle$, are added to the final state stable particles from the signal. This is done dynamically by an event builder in the analysis software, and is thus not part of the signal or MB event production. All analysis is then performed on the merged list of stable particles to model one full collision event at the LHC.

The example signal chosen for the Monte Carlo simulation based studies presented in this Section is the decay of a possible heavy Z' boson with a chosen $M_{Z'} = 1.5$ TeV to a (boosted) top quark pair, at $\sqrt{s} = 8$TeV. The top- and anti-top-quarks then decay fully hadronically ($t \to Wb \to jj\,b$-jet) or semi-leptonically ($t \to Wb \to \ell\nu\,b$-jet). The PYTHIA generator [86,87] is used to generate the signal samples. The soft physics modeling parameters in both cases are from the pre-LHC-data tune 4C [62]. The pile-up is simulated by overlaying generated minimum bias proton-proton interactions at $\sqrt{s} = 8$TeV using Poisson distributions with averages $\langle\mu\rangle = \{30, 60, 100, 200\}$, respectively, thus focusing on the exploration of future high intensity scenarios at LHC.

All analysis utilizes the tools available in the FASTJET [67] package for jet finding and jet sub-structure analysis. The larger jets used to analyze the final state are reconstructed with the anti$-k_{\text{T}}$ algorithm [71] with $R = 1.0$, to assure that most of the final state top-quark decays can be collected into one jet. This corresponds to top-quarks generated with

[14] A particle is considered stable if its lifetime τ in the laboratory frame of reference passes $c\tau > 10$ mm.

$p_{\mathrm{T}} \gtrsim 400$ GeV. The configurations for jet grooming are discussed in Sect. 4.6.

4.3 Investigating jets from pile-up

Stable particles emerging from the simulated proton-proton collisions are clustered into anti$-k_{\mathrm{T}}$ jets [71] with a radius parameter $R = 0.4$, using the FastJet [67] implementation:

Truth jets are obtained by clustering all stable particles from a single MB interaction. For an event containing μ pile-up interactions, jet finding is therefore executed μ times. The resulting truth jets are required to have $p_{\mathrm{T}} \geq 5$ GeV.

Pile-up jets are obtained by clustering the stable particles from all MB interactions forming the pile-up event. They are subjected to the kinematic cuts described below.

Jets with rapidity $|y| < 2$ are accepted.

The contribution of pile-up to jets can be corrected using the jet area based method in Ref. [88]. It employs the median transverse momentum density ρ, which here is determined using k_{T} jets with $R = 0.4$ within $|y| < 2$. To evaluate the effect of this correction, the transverse momentum ratio R_{p_T} is introduced as

$$R_{p_T} = \frac{p_T^{match}}{p_{\mathrm{T}} - \rho A} = \frac{p_T^{match}}{p_T^{corr}}. \quad (3)$$

Here A is the catchment area [68] of the pile-up jet, and p_T^{match} is the matching truth jet p_{T}. The matching criterion is similar to the one suggested in Ref. [89], where the truth jet matching uses the constituents shared between the truth jet and the pile-up jet. The jets are considered matched if the fraction of constituents of the truth jet that are also contained in the pile-up jet contribute to at least 50% of the truth jet p_{T}. In the following, pile-up jets are only considered if their corrected transverse momentum is $p_T^{corr} \geq 20$ GeV, and they are matched to at least one truth jet.

The contribution of particles from any vertex to a given pile-up jet can be measured using the jet vertex fraction ($\mathcal{F}_{\mathrm{jvf}}$). It is defined as

$$\mathcal{F}_{\mathrm{jvf}}(V_i) = \frac{\sum_{k=1}^{N_{\mathrm{part}}(V_i)} p_{\mathrm{T},k}}{\sum_{i=1}^{N_{\mathrm{coll}}} \sum_{k=0}^{N_{\mathrm{part}}(V_i)} p_{\mathrm{T},k}} = \frac{1}{p_{\mathrm{T}}} \sum_{k=1}^{N_{\mathrm{part}}(V_i)} p_{\mathrm{T},k}, \quad (4)$$

where $N_{\mathrm{part}}(V_i)$ is the number of particles from a given vertex V_i, and N_{coll} is the number of collision vertices contributing particles to the jet. $\mathcal{F}_{\mathrm{jvf}}$ is calculated for each of these vertices. Note that p_{T} corresponds to the uncorrected jet transverse momentum and consequently, the value of each component of $\mathcal{F}_{\mathrm{jvf}}(V_i)$ depends on μ.

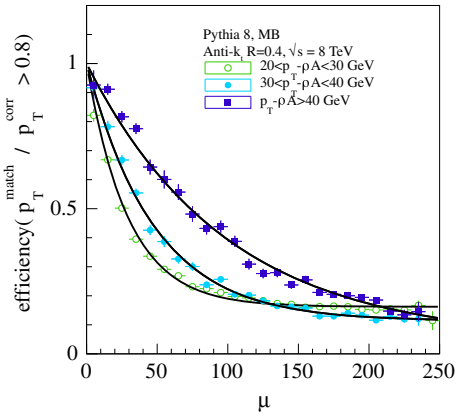

Fig. 4 The simulated fraction of pile-up jets with $R_{p_T} > 0.8$ (QCD-like) in a 8 TeV operation of the LHC, as a function of the number of minimum-bias interactions and for different values of p_T^{corr}. A fit of the exponential form $f = c_0 + c_1 \exp(c_2 \cdot N_{\mathrm{PV}})$ is superposed where one degree of freedom is fixed via the constraint $f(0) = 1$, i.e. $c_1 = (1 - c_0)$

4.4 Evaluation of the pile-up jet nature

It follows from the definition of R_{p_T} that pile-up jets with values of R_{p_T} close to unity are matched to a truth jet with $p_{\mathrm{T}} \approx p_T^{corr}$ of the pile-up jet itself. Consequently, there is a single MB interaction which predominantly contributes to the jet. On the other hand, jets with a small value of R_{p_T} are mostly stochastic, as no single minimum-bias collision contributes in a dominant way to the pile-up jet. We characterize jets as stochastic if R_{p_T} is smaller than 0.8. This threshold value is arbitrary and the fraction of QCD-like and stochastic jets depends on the exact choice. The conclusion of our study holds for a broad range of cut values.

The fractions of QCD-like and stochastic pile-up jets change as a function of pile-up jet p_T and μ. This can be seen in Fig. 4, where QCD jet-like samples are defined by $R_{p_T} > 0.8$ for each $pile-up$ level. The fraction of these jets at a given p_T^{corr} decreases exponentially with μ. The exponential decrease is slower for larger p_T^{corr}. At a $pile-up$ activity of $\mu = 100$, the fraction of pile-up jets that are QCD-like is about 40% (20%) for $p_T^{corr} > 40$ GeV($20 < p_T^{corr} < 30$ GeV). At $\mu = 150$, these numbers decrease to about 25 and 15%, respectively.

4.5 Pile-up jet multiplicity

The mean number of pile-up jets per event, as a function of jet p_T^{corr} and N_{PV}, is indicative of the efficiency of the jet area based method to suppress jets generated by pile-up. It is shown in Fig. 5 for the inclusive pile-up jets and separately for the subsample of QCD-like pile-up jets satisfying $R_{p_T} > 0.8$. It is observed that the average inclusive number $\langle N \rangle$ of low ($p_T^{corr} \simeq 20$ GeV) pile-up jets per event increases rather linearly with μ, i.e. $\partial \langle N \rangle / \partial \mu \approx$ const. For higher pile-up

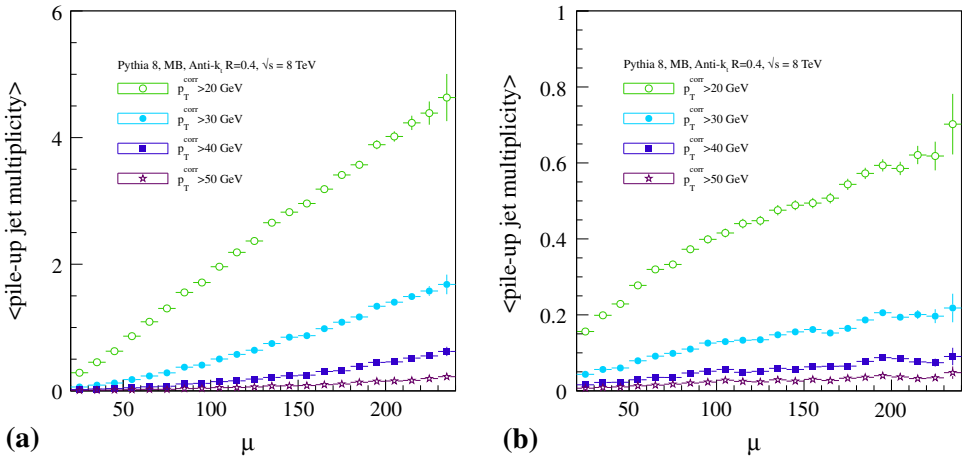

Fig. 5 Simulation results for mean number of pile-up jets per event in the 8 TeV LHC, inclusively (**a**) and for QCD-like pile-up jets with $R_{p_T} > 0.8$ (**b**), as a function of μ and p_T^{corr}

jet p_{T}, $\partial\langle N\rangle/\partial\mu$ is significantly smaller, and displays an increase with increasing μ.

The sub-sample of QCD-like jets in the inclusive pile-up jet sample shows a different behavior, as indicated in the righmost panel of Fig. 5. In this case $\partial\langle N\rangle/\partial\mu$ decreases with increasing μ in all considered bins of p_T^{corr}. This contradicts the immediate expectation of an increase following the inclusive sample, but can be understood from the fact that with increasing μ the likelihood of QCD-like jets to overlap with (stochastic) jets increases as well. The resulting (merged) pile-up jets no longer display features consistent with QCD jets (e.g., loss of single energy core), and thus fail the $R_{p_T} > 0.8$ selection.

The pile-up jet multiplicity shown in Fig. 5 is evaluated as a function of the pile-up corrected transverse momentum of the jet (p_T^{corr}). This means that after the correction approximately two pile-up jets with $p_T^{corr} > p_{\mathrm{T}}^{\mathrm{min}} = 20$ GeV can be expected for $\langle\mu\rangle \simeq 100$. This number decreases rapidly with increasing $p_{\mathrm{T}}^{\mathrm{min}}$. The mean number of QCD jets is small, at about 0.4 at $\langle\mu\rangle = 100$, for $p_{\mathrm{T}}^{\mathrm{min}} = 20$ GeV.

4.6 Jet grooming configurations

Three jet grooming techniques are used by the LHC experiments:

Jet trimming Trimming is described in detail in Ref. [2]. In this approach the constituents of the large anti$-k_{\mathrm{T}}$ jet formed with $R = 1.0$ are re-clustered into smaller jets with $R_{\mathrm{trim}} = 0.2$, using the anti$-k_{\mathrm{T}}$ algorithm again. The resulting sub-jets are only accepted if their transverse momentum is larger than a fraction f (here $f = 0.03$) of a hard scale, which was chosen to be the p_{T} of the large jet. The surviving sub-jets are recombined into a groomed jet.

Jet filtering Filtering was introduced in the context of a study to enhance the signal from the Higgs boson decaying into two bottom-quarks, see Ref. [1]. In its simplified configuration without mass-drop criterion [90] applied in this study it works similar to trimming, except that in this case the sub-jets are found with the Cambridge-Aachen algorithm [73,91] with $R_{\mathrm{filt}} = 0.3$, and only the three hardest sub-jets are retained. The groomed jet is then constructed from these three sub-jets.

Jet pruning Pruning was introduced in Ref. [92]. Contrary to filtering and trimming, it is applied during the formation of the jet, rather than based on the recombination of sub-jets. It dynamically suppresses small and larger distance contributions to jet using two parameters, Z_{cut} for the momentum based suppression, and $D_{cut} = D_{\mathrm{cut,fact}} \times 2m/p_{\mathrm{T}}$ (here m and p_{T} are the transverse momentum and mass of the original jet) for the distance based. Pruning vetoes recombinations between two objects i and j for which the geometrical distance between i and j is more than D_{cut} and the p_T of one of the objects is less than $Z_{\mathrm{cut}} \times p_T^{i+j}$, where p_T^{i+j} is the combined transverse momentum of i and j. In this case, only the hardest of the two objects is kept. Typical values for the parameters are: $Z_{\mathrm{cut}} = 0.1$ and $D_{\mathrm{cut,fact}} = 0.5$.

In this study, trimming and filtering are applied to the original anti$-k_{\mathrm{T}}$ jets with size $R = 1.0$. We study the interplay between jet grooming and area-based pile-up correction. The subtraction is applied directly on the 4-momentum of the jet using:

$$p^{\mu}_{\mathrm{jet,sub}} = p^{\mu}_{\mathrm{jet}} - [\rho A^{x}_{\mathrm{jet}},\ \rho A^{y}_{\mathrm{jet}},\ (\rho+\rho_m)A^{z}_{\mathrm{jet}},\ (\rho+\rho_m)A^{E}_{\mathrm{jet}}], \tag{5}$$

with

$$\rho = \underset{\text{patches}}{\mathrm{median}} \left\{ \frac{p_{t,\mathrm{patch}}}{A_{\mathrm{patch}}} \right\}, \quad \rho_m = \underset{\text{patches}}{\mathrm{median}} \left\{ \frac{m_{\delta,\mathrm{patch}}}{A_{\mathrm{patch}}} \right\}, \tag{6}$$

$m_{\delta,\text{patch}} = \sum_{i\in\text{patch}} \left(\sqrt{m_i^2 + p_{t,i}^2} - p_{ti}\right)$, and A^μ is the active area of the jet as defined in Ref. [68] and computed by FastJet. The ρ term, mentioned above is the standard correction typically correcting the transverse momentum of the jet. The ρ_m term corrects for contamination to the total jet mass due to the pile-up particle. When applying this subtraction procedure, we discard jets with negative transverse momentum or (squared) mass of the jet.

The estimation of ρ and ρ_m is performed with FastJet using[15] k_t jets with $R = 0.4$. Corrections for the rapidity dependence of the pile-up density ρ are applied using a rapidity rescaling.

When we apply this background subtraction together with trimming or filtering, the subtraction is performed directly on the subjets, before deciding which subjets should be kept, so as to limit the potential effects of pile-up on which subjets are to be kept.

4.7 Jet substructure reconstruction

The various methods and configurations discussed in the previous section are applied to the jets reconstructed with the anti$-k_\text{T}$ algorithm with $R = 1.0$ in the $Z' \to t\bar{t}$ final state in the presence of pile-up. For the studies presented in this report we require jet p_T before grooming and pile-up subtraction to be greater than 100 GeV and consider the two hardest p_T-jets in the event. We further require that the rapidity difference between the two jets $|y_1 - y_2|$ is less than one. The immediate expectation for the reconstructed jet mass m is the top mass, i.e. $m \approx 175$ GeV, and no residual dependence on the pile-up activity given by $\langle\mu\rangle$, after the pile-up subtraction. The two plots in the upper row of Fig. 6 show the distributions of the reconstructed jet masses without any grooming and with the pile-up subtraction discussed in Sect. 4.6 applied. The effect of pile-up on the mass scale and resolution is clearly visible. Applying only the pile-up subtraction, without changing the composition of the jets, already improves the mass reconstruction significantly. All $\langle\mu\rangle$ dependence is removed from the jet mass spectrum, as shown in Fig. 6. In particular, the position of the mass peak is recovered. With increasing pile-up, the mass peak gets more and more smeared, an effect due to the fact that the pile-up is not perfectly uniform. These point-to-point fluctuations in an event lead to a smearing $\pm\sigma\sqrt{A}$ in (5). For very large pile-up, this smearing extends all the way to $m = 0$ as seen in Fig. 6.

The effect of the other grooming techniques on the reconstructed jet mass distributions is summarized in Fig. 6, with and without the pile-up subtraction applied first. The spectra show that both trimming and filtering can improve the mass reconstruction. The application of the pile-up subtraction in addition to trimming or filtering further improves the mass reconstruction performance.

The findings from the spectra in Fig. 6 are quantitatively summarized in Fig. 7 for the mass scale and resolution. Here the resolution is measured in terms of the mass range in which 67% of all jet masses can be found ($Q_{67\%}(m_\text{jet})$ quantile). Maintaining the jet mass scale around the expectation value of 175 GeV works well for trimming and filtering with and without pile-up subtraction, see Fig. 7. The same figure indicates that for very high pile-up ($\langle\mu\rangle = 100 - 200$), the jet mass after trimming and filtering without pile-up subtraction shows increasing sensitivity to the pile-up. The additional pile-up subtraction tends to restore the mass scale with better quality.

Both trimming and filtering improve the mass resolution to different degrees, but in any case better than pile-up subtraction alone, as expected. Applying the additional pile-up subtraction to trimming yields the least sensitivity to the pile-up activity in terms of mass resolution and scale.

These effects can be explained as follows. As discussed earlier, pile-up has mainly two effects on the jet: a constant shift proportional to ρA and a smearing effect proportional to $\sigma\sqrt{A}$, with σ a measure of the fluctuations of the pile-up within an event. In that language, the subtraction corrects for the shift leaving the smearing term untouched. Grooming, to the contrary, since it selects only part of the subjets, acts as if it was reducing the area of the jet[16]. This reduces both the shift and the dispersion. Combining grooming with subtraction thus allows to correct for the shift leftover by grooming and reduce the smearing effects at the same time. All these effects are observed in Fig. 7.

4.8 Concluding remarks

The source of jets produced in minimum bias collisions in the presence of pile-up is analyzed using a technique relating the single collision contribution in the jet to its transverse momentum after pile-up correction in particle level Monte Carlo. The rate of pile-up jets surviving after application of the jet area based pile-up subtraction is about two with $p_T^{corr} > 20$ GeV and within $|y| < 2$, at a pile-up activity of $\langle\mu\rangle = 100$. It rises about linearly with increasing pile-up for this particular selection. Higher p_T jets occur at a much reduced rate, but with a steeper than linear rise with increasing μ.

The rate of QCD-like jets is significantly smaller, and shows a less-than-linear increase with increasing μ even

[15] Ghosts are placed up to $y_\text{max} = 3$ and explicit ghosts are enabled.

[16] Note that grooming techniques do more than reducing the catchment area of a jet. Noticeably, the selection of the hardest subjets introduces a bias towards including upwards fluctuations of the background. This positive bias is balanced by a negative one related to the perturbative radiation discarded by the grooming. These effects go beyond the generic features explained here.

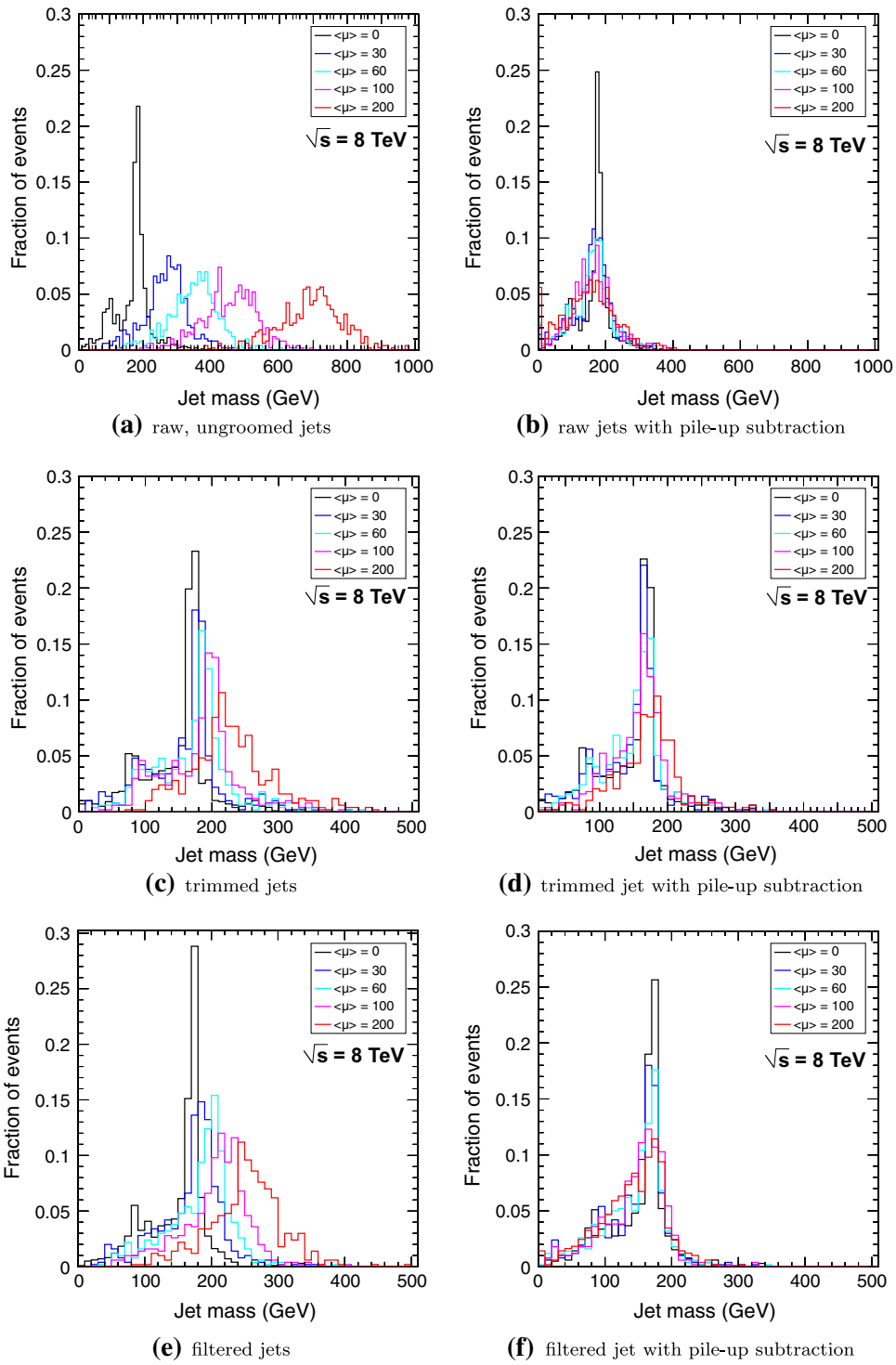

Fig. 6 Simulation results for the impact of pile-up on the jet mass distribution. *Top row* the raw jet mass distribution for jets reconstructed with the anti$-k_{\mathrm{T}}$ algorithm and $R = 1.0$ in $Z' \to t\bar{t}$ final states with $m_{Z'} = 1.5$ TeV, in the presence of pile-up with $\langle\mu\rangle = 30$, 60, 100, and 200, before and after pile-up subtraction. The second and third row show the same result after trimming (*middle row*) and filtering (*lower row*)

for $p_{\mathrm{T}}^{\mathrm{min}} = 20$ GeV. This can be understood as a sign of increased merging between QCD-like jets and stochastic jets. The merged jets are less likely to display features characteristic for QCD-like jets, and therefore fail the selection.

The fraction of QCD-like jets with a core of energy arising from a single proton-proton interaction of at least $0.8\,p_T^{corr}$ is found to decrease rapidly with increasing μ. At $\mu = 50$ about 60% of the pile-up jets with $p_T^{corr} > 50$ GeV are found to be QCD-like, whereas at $\mu = 200$ this number is decreased to about 20%.

A brief Monte Carlo study of the effect of jet grooming techniques on the jet mass reconstruction in $Z' \to t\bar{t}$ final states has been conducted. Jet trimming and filtering are used by themselves, or in combination with the pile-up subtraction

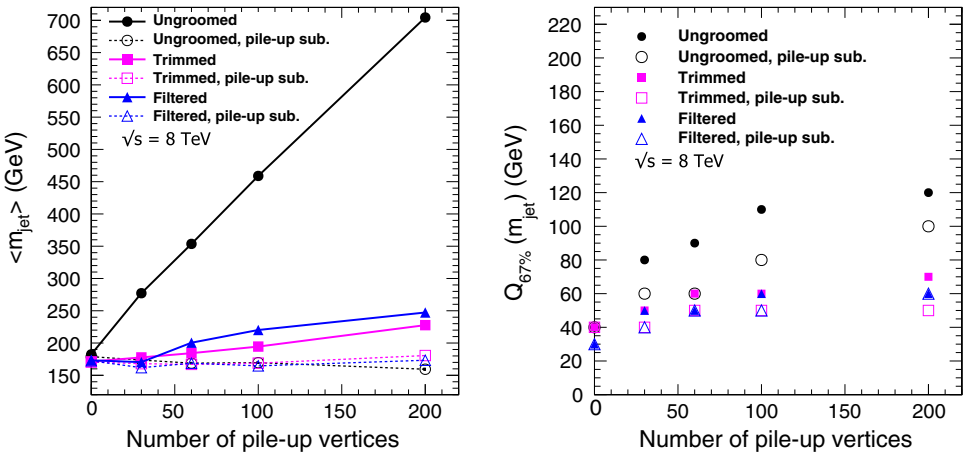

Fig. 7 Simulated average (leftmost figure) and RMS (rightmost figure) of the reconstructed jet mass distribution in $Z' \to t\bar{t}$ final states, as a function of the pile-up activity $\langle\mu\rangle$, for various jet grooming techniques

using the four-vector area, to reconstruct the single jet mass and evaluate the stability of the mass scale and resolution at pile-up levels of 30, 60, 100, and 200 extra proton-proton collisions, in addition to the signal event. It is found that for this particular final state trimming and filtering work well for maintaining the mass scale and resolution, provided they are applied together with pile-up subtraction so as to benefit both from the average shift correction from subtraction and noise reduction from the grooming.

The studies presented here are performed with Monte Carlo simulated signal and pile-up (minimum bias) interactions. No considerations have been given to detector sensitivities and other effects deteriorating the stable particle level kinematics and flows exploited here. With this respect the conclusions of this study are limited and can be considered optimistic until shown otherwise.

Note also that comparing the performance of filtering and trimming would require varying their parameters and that this goes beyond the scope of this study.

5 The potential of boosted top quarks

Section prepared by the Working Group: 'Prospects for boosted top quarks', A. Altheimer, J. Ferrando, J. Pilot, S. Rappoccio, M. Villaplana, M. Vos.

Many applications of the strategies for boosted objects have been proposed (the bibliography of this paper and those included in Refs. [9,10] are a good starting point to navigate the extensive literature). Among these, the study of highly energetic top quarks forms the case that has been studied in greatest detail by the experiments. Several studies of the production of boosted top quarks have set limits on new physics scenarios. The first sample of boosted top quarks has also been used to understand the modelling of the parton shower and the detector response. In this section we present a summary of achievements so far, discuss how existing analyses could benefit from an improved understanding of jet substructure, and explore possible directions for future work.

5.1 Boosted top quark production

The top quark decay topology observed in the detector depends strongly on the kinematic regime. The decay products of top quarks produced nearly at rest ($p_T < 200$ GeV/c) are well-separated, leading to experimental signatures such as isolated leptons and a relatively large number of clearly resolved jets. With increasing transverse momentum, the decay products of the top quark will become collimated and possibly reconstructed in the same final state object. For intermediate boosts ($200 < p_T < 400$ GeV), the daughters of the W boson from a fully-hadronic top decay will be close enough to be clustered into the same jet. At this point, the use of jet substructure techniques becomes important to efficiently identify these decay signatures. At even larger p_T top quarks become truly boosted objects: all decay products of the top will be strongly collinear, with the $\Delta R \sim 2m_{\text{top}}/p_T$. Hadronic top quarks can be reconstructed in a single jet, and top quarks with leptonic decays generally contain non-isolated leptons due to the overlap with the b-quark jet.

Table 1 presents the expected numbers of boosted top quark pairs according to the Standard Model at past, present and future colliders. The numbers show clearly how the study of boosted top quarks becomes viable only with the start of the LHC. The first phase of operation yields a sample of several tens of thousands of boosted top quark pairs. The next-to-last column indicates the size of the sample expected in a 13 or 14 TeV run of the LHC, that is to start by the end of 2014. The increase in the centre-of-mass energy and the larger integrated luminosity each bring an increase of an order of magnitude in the production of boosted top quarks.

We expect, therefore, that boosted topologies will gain considerable importance as the LHC program develops. To exploit the LHC data to their full potential it is critical that

Table 1 The top pair production rate at past, present and future colliders, calculated with the MCFM code [93]. The inclusive production rate is given in the first row. The expected number of events with boosted top quarks ($M_{t\bar{t}} > 1$ TeV) and highly boosted top quarks ($M_{t\bar{t}} > 2$ TeV) is given in the second and third row, respectively

Collider & phase process & energy, integrated luminosity	Tevatron run II $p\bar{p}$ at $\sqrt{s} = 1.96$ TeV $\mathcal{L} = 10$ fb^{-1}	LHC 2012 pp at $\sqrt{s}$ =8 TeV $\mathcal{L} = 20$ fb^{-1}	LHC phase II pp at $\sqrt{s}$ =13 TeV $\mathcal{L} = 300$ fb^{-1}	HE-LHC pp at $\sqrt{s}$ =33 TeV $\mathcal{L} = 300$ fb^{-1}
Inclusive $t\bar{t}$ production	6×10^4	4×10^6	2×10^8	1.4×10^9
Boosted production	23	6×10^4	5.2×10^6	7.1×10^7
Highly boosted	0	500	1.1×10^5	3.9×10^6

existing experimental strategies are adapted to this challenging kinematical regime. Before we turn to the results of analyses of boosted object production, we discuss a number of new tools that were developed to identify and reconstruct boosted top quarks efficiently.

5.2 Top tagging

Excellent reviews of top tagging algorithms exist [94]. Previous BOOST reports have compared their performance for simulated events (at the particle level). In this Section we present a very brief review for completeness.

The Johns Hopkins (JHU) tagger [95] identifies substructure by reversing the last steps of the jet clustering. Hard subjets are selected using several criteria—the ratio of their individual p_T to the original jet p_T must be above a given threshold, and the subjets must be spatially separated from each other to give a valid decomposition. In this way, a jet can be deconstructed into up to four subjets, and jets with three or more subjets are analyzed further, requiring the invariant mass of the identified subjets to be in the range [145, 205] GeV, and two of the subjets to be consistent with m_W, in the range [64,94] GeV. There is an additional cut on the W boson helicity angle, $\cos\theta_h < 0.7$.

The variant of the JHU tagger used by CMS [96] uses a similar jet decomposition, with slight differences in the selections of top quark and W boson masses from the subjets. Additionally, the CMS top tagger does not apply the W boson helicity angle requirement, but instead selects jets with the minimum pairwise mass of the subjets larger than 50 Gev. The JHU and CMS top tagging algorithms have been developed with jet distance parameters up to $R = 0.8$, and therefore are only efficient for top quarks with p_T above approximately 400 GeV/c.

The HEP top tagger [97], is designed to use jets with distance parameter $R = 1.5$, thereby extending the reach of the tagging algorithm to lower jet p_T values. The algorithm uses a mass drop criterion to identify substructure within the jet, but also uses a filtering algorithm to remove soft and large-angle constituents from the individual subjets. The three subjets with a combined mass closest to m_t are then chosen for further consideration. Cuts are then applied to masses of subjet combinations to ensure consistency with m_W and m_t. Specifically, for the three subjets sorted in order of subjet p_T, having masses m_1, m_2, m_3, the quantities m_{23}/m_{123} and $\arctan m_{13}/m_{12}$ are computed. Geometrical cuts can be applied in the phase space defined by these two quantities to select top jets and reject quark or gluon jets.

The HEP top tagger obtains tagging efficiencies of up to 37% for lower p_T top quarks ($p_T > 200$ GeV/c), with an acceptable mistag rate. It has been used by the ATLAS $t\bar{t}$ resonance search in the fully hadronic channel [98], where no *resolved* analysis has been performed. At high jet p_T, the efficiencies for the HEP Top Tagger and JHU Top Tagger selections are comparable.

Boosted top quarks were also studied using both $R = 1.0$ anti-k_t jets and jets identified by the HEPTopTagger [97] algorithm as candidate "top-jets." Kinematic and substructure distributions were compared between data and MC simulation and were found to be in agreement. Furthermore, the efficiency with which top quarks were identified as such was found to be significantly increased in both cases, and the HEPTopTagger was shown to reduce the backgrounds to such searches dramatically, even with a relatively relaxed transverse momentum selection.

Overall, the results from ATLAS suggests that, among the jet grooming configurations tested, the trimming algorithm exhibited an improved mass resolution and smaller dependence of jet kinematics and substructure observables on pile-up (such as N-subjettiness [75,76] and the k_t splitting scales [99]) compared to the pruning configurations examined. For boosted top quark studies, the anti-k_t algorithm with a radius parameter of $R = 1.0$ and trimming parameters $f_{\text{cut}} = 0.05$ and $R_{\text{sub}} = 0.3$ was found to be optimal, where a minimum p_T requirement of 350 GeV is typical. It is important to note that only the k_t-pruning for $R = 1.0$ jets was tested and that since the performance does depend somewhat on this parameter, further studies are necessary to optimize for other jet size. Lastly, Cambridge-Aachen jets with $R = 1.2$ using the mass-drop filtering parameter $\mu_{\text{frac}} = 0.67$ were found to perform well for boosted two-pronged analyses such as $H \to b\bar{b}$ or searches involving boosted $W \to q\bar{q}$ decays.

A final algorithm that is currently being investigated is the N-subjettiness algorithm [75] presented in Sect. 3.

Several new techniques and ideas are emerging, that aim to improve boosted top identification and reconstruction.

One such technique is that of shower deconstruction [100]. This method aims to identify boosted hadronic top quarks by computing the probability for a top quark decay to produce the observed jet, including its distribution of constituents. The probability for the same jet to have originated from a background process is also computed. These probabilities are computed by summing over all possible shower formations resulting in the observed final state, accounting for different gluon splittings and radiations, among other processes. This is done both for the signal shower processes and background shower processes. A likelihood ratio is formed from the signal and background probabilities and used to discriminate boosted top quarks from generic QCD jets. The process of evaluating all shower histories can be computationally intensive, so certain requirements are made on the number of constituents used in the method to make the problem tractable. The results presented in Ref. [101] show an improvement on the top taggers described previously. Specifically, the shower deconstruction method reduces the top mistag rate by a factor of 3.6 compared to the JHU top tagger, while maintaining the same signal acceptance. This method is also applicable to the lower p_T regime, and there improves upon the top mistag rate from the HEP top tagger by a factor of 2.6, again keeping identical signal efficiency.

Another algorithm under development is the template overlap method [116]. The template overlap method is designed for use in boosted top identification as well as boosted Higgs identification. The method is similar to that of shower deconstruction, in that it attempts to quantify how well a given jet matches a certain expectation such as a boosted top quark or boosted Higgs decay. However, this method uses only final state configurations, whereas the shower deconstruction method takes into account the showering histories. A catalog of templates is formed by analyzing signal events. Once this is in place, individual jets can be analyzed by evaluating an overlap function which evaluates how well the current jet matches the templates from the signal process of interest. For example, a template for hadronic boosted top quark decays would consist of three energy deposits within the jet. The background of high-p_T QCD jets is reduced by two orders of magnitude. One additional feature of this template overlap method is the automatic inclusion of additional parton radiation into the template catalog, such as for Higgs decays to bottom quark pairs, where there is commonly an additional gluon radiated, resulting in 3 energy deposits instead of the 2 from the b quarks.

Finally, the Q-jets [117] scheme could be used for top-tagging. This is a method to remove dependence of analysis results on the choice of clustering algorithm used to reconstruct jets. For example, one could use either the Cambridge-Aachen algorithm or the k_t algorithm to cluster jets, and may obtain significantly different results in the jet masses. The Q-jets algorithm attempts to use all possible "trees" to cluster constituents, rather than using the single tree provided by the specific clustering algorithm used. In this way, each jet now has a distribution of possible masses instead of a single jet. This provides additional information which can enhance signal discrimination. For example, the variance of the jet mass between individual clustering trees can be examined, rather than relying on just a single value. The statistical stability is also enhanced when using the Q-jets algorithm.

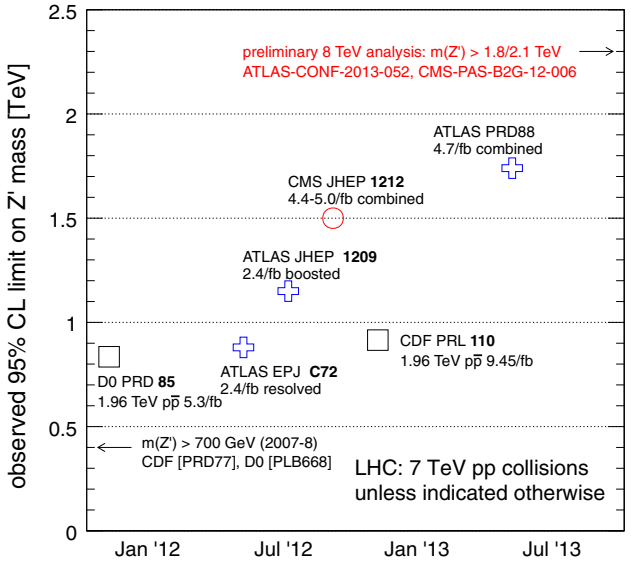

Fig. 8 Overview of evolution of the sensitivity of $t\bar{t}$ resonance searches in the first years of LHC operation. The sensitivity is presented in terms of the lower limit on the mass of a narrow Z' boson. The production rate for this new state is given by a benchmark model that is common to all experiments (a leptophobic topcolor Z' boson)

5.3 Searches with boosted top quarks

The first area where new tools developed specifically for the selection and reconstruction of boosted top quarks have shown their value is in searches for massive new states decaying to top quark pairs. The first application of techniques specifically aimed at boosted top decays was the CMS $t\bar{t}$ resonance search in the all-hadronic channel [111]. The evolution of the mass reach[17] of $t\bar{t}$ resonance searches in the more sensitive "lepton+jets" channel is shown in Fig. 8. By the start of the LHC program the Tevatron experiments had excluded a Z' boson mass lower than 700 GeV [102,103]. In the course of 2011 and 2012 the limit was extended to 800

[17] The sensitivity to massive particles is expressed in terms of the observed 95% CL lower limit on the mass of a leptophobic topcolor Z' boson. The motivation of this particular model may not have survived recent advances in particle physics, but to monitor the sensitivity of searches it is still the best benchmark on the market.

Table 2 Exclusion limits at 95% confidence level for a narrow Z' boson, as obtained in $t\bar{t}$ resonance searches at the Tevatron and the first years of operation of the LHC

CDF and D0 References	[102]	[103]	[104]	[105]	[106]
Final state &	l+jets	l+jets	fully had.	l+jets	l+jets
Reconstruction	resolved	resolved	resolved	resolved	resolved
$\sqrt{s}$ [TeV]	1.96	1.96	1.96	1.96	1.96
$\int L$ [fb^{-1}]	1 fb^{-1}	1 fb^{-1}	4 fb^{-1}	4 fb^{-1}	10 fb^{-1}
Z' mass [TeV]	<0.7	<0.720	<0.805	<0.835	<0.915
ATLAS Reference	[107]	[108]	[98]	[109]	[110]
Final state &	l+jets	l+jets	fully had.	l+jets	l+jets
Reconstruction	resolved	boosted	boosted	combined	combined
$\sqrt{s}$ [TeV]	7	7	7	7	8
$\int L$ [fb^{-1}]	2.04 fb^{-1}	2.04 fb^{-1}	4.07 fb^{-1}	4.07 fb^{-1}	14 fb^{-1}
Z' mass [TeV]	0.5 − 0.88	0.6 − 1.15	0.7−1, 1.28−1.32	<1.74	<1.8
g_{KK} mass [TeV]	0.5 − 1.13	0.6 − 1.5	0.7 − 1.62	<2.07	<2.0
CMS Reference	[111]	[112]	[113]	[114]	[115]
Final state &	fully hadronic	l+jets	di-lepton	fully hadronic	l+jets
Reconstruction	boosted	combined		boosted	combined
$\sqrt{s}$ [TeV]	7	7	7	8	8
$\int L$ [fb^{-1}]	5.0 fb^{-1}	4.4−5.0 fb^{-1}	5.0 fb^{-1}	19.6 fb^{-1}	19.6 fb^{-1}
Z' mass [TeV]	1.3 − 1.5	<1.49	<1.3	<1.7	<2.1
g_{KK} mass [TeV]	1.4 − 1.5	<1.82	<1.8	<1.8	<2.5

GeV by a D0 search on nearly 5 fb^{-1} [105] and to approximately 900 GeV by a CDF analysis of the complete Tevatron data set [106]. An ATLAS search on 2.4 fb^{-1} of 7 TeV LHC data [107] collected in 2011 reached a similar precision. All these analyses followed the conventional, *resolved* approach that is based on the assumption that the six fermions from the decay of the top quark pair ($t \to W^+b \to l^+\nu_l b$ and the charge conjugate process) can be resolved individually.

In some cases ATLAS and CMS analyses specifically designed for boosted top quarks [108,112] scrutinized the same data set that had been used by the *resolved* approach. A direct comparison of these results demonstrates that the novel approach has considerably better sensitivity for massive states [108]. The final analyses on 2011 data [109,112] combine *resolved* and *boosted* methods to attain good sensitivity over the complete mass spectrum. The excluded mass range is pushed up to 1.74 TeV.

Searches in the "lepton+jets" channel are complemented by analyses of the fully hadronic ($t\bar{t} \to 6$ jets) and di-lepton ($t\bar{t} \to b\bar{b}l^+\nu_l l'^-\bar{\nu}'_l$) decay chains. Only one fully hadronic $t\bar{t}$ resonance search was performed at the Tevatron [104]. At the LHC, with a daunting multi-jet background, these searches are even more challenging. The advent of new algorithms has, however, greatly boosted their potential. The mass reach of the CMS [111] and ATLAS search [98] are compared to that of the "lepton+jets" searches in Table 2.

The prospects for progress are good. Preliminary results on the 2012 data set [110,114,115] have significantly extended previous limits.

5.4 Jet substructure performance and searches

The results in the previous Section form the proof-of-principle: the addition of jet substructure analysis techniques to the experimentalists' tool-box boosts the sensitivity of searches for new physics at the LHC. It is clear, however, that these tools are still in their infancy. In all searches discussed in the previous Section large systematic uncertainties are assigned to the large-R jets. It is natural to suspect that further progress could be made with better (and, especially, better understood) tools.

To quantify the impact of the jet-related systematics on the sensitivity we have evaluated expected limits on the narrow Z' boson with all sources of systematic uncertainty, except one (so-called $N-1$ limits) in several iterations of the ATLAS searches in the lepton+jets final state. The uncertainties associated with the large-R jet that captures the hadronic top decay are always the dominant source of uncertainty. Their

impact is considerably larger than that of systematics associated with the narrow jets, even at relatively low resonance mass. The limits over a large mass range (1–2 TeV) would improve by approximately 5–10% if only the uncertainty on the scale and resolution of mass and energy of anti-k_t jets with $R = 1$ were removed.

If we apply an ad hoc scale factor of two to this uncertainty (representing a failure to bring these uncertainties under control) we find that the sensitivity is further degraded. A significant reduction of large-R jet uncertainties, on the other hand, brings the $N - 1$ limits with no jet-related systematics and the limits with reduced large-R jet systematics to within 2%.

CMS has not published the $N - 1$ results for their searches, but qualitatively the same picture emerges. In the fully hadronic searches the jet-related uncertainties have the largest impact on the limits.

We conclude that further progress understanding jet substructure still has substantial potential to increase our sensitivity to massive new states decaying to top quarks.

5.5 Further applications

The selection for boosted top quarks, in the lepton+jets and fully hadronic channels, have proven their value in $t\bar{t}$ resonance searches, but are more generally applicable.

The obvious direction to extend the range of applications is to other searches with boosted top quarks. The $W' \to tb$ that are currently performed in the channel where the top quark decays to a charged lepton, neutrino and b-jet. We expect, however, that, ultimately the highest mass reach should be obtained in the hadronic decay (with a factor two large branching ratio if τ-leptons are not considered).

We expect differential cross-section measurements for $t\bar{t}$ to benefit from these techniques at large transverse momentum and invariant mass of the $t\bar{t}$ pair. Apart from the better selection efficiency in algorithms designed for this kinematic regime, the better truth-to-reconstructed mapping of p_T and $m_{t\bar{t}}$ is expected to be an important advantage. We are looking forward to such measurements from the ATLAS and CMS experiments. Also analyses that rely strongly on the reconstruction of the top quark direction, such as the charge asymmetry measurement, should benefit.

Finally, several authors [118] have commented on the potential of events with mildly boosted top quarks for the observation of $t\bar{t}H$ and a measurement of the production rate.

5.6 Summary

Over the last five years, many ideas have been proposed to cope with the challenge of boosted top quark reconstruction. Since then, these ideas have been implemented by the experiments and put to the test, primarily in searches for massive new states decaying to $t\bar{t}$ pairs. The overview we presented in Table 2; Fig. 8 is a testimony to the increase of sensitivity for such states fuelled by the performance of the LHC. Such progress would not have been possible if novel techniques for the study of boosted top quarks had not been developed. We expect the selection developed for the lepton+jets and fully hadronic to find further applications in searches and measurements.

6 Summary and conclusions

This report of the BOOST2012 workshop addresses a number of important questions concerning the use of jet substructure for the study of boosted object production at the LHC.

We evaluated the current limitations in the description of jet substructure, both at the analytical level and in Monte Carlo generators. Impressive progress is being made for the former and we expect a meaningful comparison to LHC data to be a reality soon. Two approaches—perturbative QCD and Soft Collinear Effective Theory—to a first-principle resummation of the jet invariant mass are producing mature results. Measurements of the jet mass in Z+jet events are proposed, both inclusively and exclusively in the number of jets. We hope that in the not-too-distant future these calculations can enhance our understanding of the internal structure in jets.

Monte Carlo predictions remain crucial to searches and measurements employing jet substructure. We have compared the predictions of several mainstream generators for a number of substructure observables a and for several signal and background topologies. While jet mass is still poorly described by several generators, several ways of introducing the inherent uncertainties become evident. Jet grooming reduces the spread among Monte Carlo models, as do several alternative jet substructure observables.

We also studied potential experimental limitations that could check further progress, in particular the impact of the large number of simultaneous proton-proton interactions. We find that, even if the substructure of large-radius jets is quite sensitive to pile-up, a combination of a state-of-the-art correction technique and jet grooming can effectively restore the jet mass scale and strongly mitigate the impact on the jet mass resolution.

Finally, we reviewed top-tagging techniques deployed in the LHC experiments and assessed their impact on the sensitivity to new physics. A series of $t\bar{t}$ resonance searches performed by ATLAS and CMS provide clear proof of the power of techniques specifically designed for boosted top quarks. Through an evaluation of the impact of all sources of systematic uncertainties, we show that further progress can still be made with an enhanced understanding of jet substructure. We expect to see these techniques applied in further searches involving boosted top quarks and in measurements of the boosted top production rate.

Acknowledgments We thank the Spanish Center for Particle Physics, Astroparticle and Nuclear Physics (CPAN), the regional government (Generalitat Valenciana), Heidelberg University and IFIC (U. Valencia/CSIC) for their generous support of the BOOST2012 conference. We furthermore thank the Fundación Cultural Bancaja for putting at our disposal the "Centro Cultural" that hosted the workshop, the IT teams at IFIC, UW and LPTHE for the facilities offered to host the BOOST samples, TotNou for the organization of the workshop, Isidoro García of CPAN for organizing the outreach event and coordinating the contacts with the media and Pilar Ordaz for the design of the poster and logotype.

References

1. J.M. Butterworth, A.R. Davison, M. Rubin, G.P. Salam, Jet substructure as a new Higgs search channel at the LHC. Phys. Rev. Lett. **100**, 242001 (2008). arXiv:0802.2470
2. D. Krohn, J. Thaler, L.-T. Wang, Jet trimming. JHEP **1002**, 084 (2010). arXiv:0912.1342
3. S.D. Ellis, C.K. Vermilion, J.R. Walsh, Recombination algorithms and jet substructure: pruning as a tool for heavy particle searches. Phys. Rev. **D81**, 094023 (2010). arXiv:0912.0033
4. Boost 2012, IFIC Valencia, 23–27 July 2012. http://ific.uv.es/boost2012
5. Boost 2009, giving physics a new boost, SLAC National Accelerator Laboratory, 9–10 July, 2009. http://www-conf.slac.stanford.edu/Boost2009
6. Boost 2010, University of Oxford, 22–25 June 2010. http://www.physics.ox.ac.uk/boost2010
7. Boost 2011, Princeton University, 22–26 May 2011. http://boost2011.org
8. Boost 2013, University of Arizona, 12–17 August 2013. http://w3atlas.physics.arizona.edu/boost2013
9. A. Abdesselam, E.B. Kuutmann, U. Bitenc, G. Brooijmans, J. Butterworth et al., Boosted objects: a Probe of beyond the Standard Model physics. Eur. Phys. J. **C71**, 1661 (2011). arXiv:1012.5412
10. A. Altheimer, S. Arora, L. Asquith, G. Brooijmans, J. Butterworth et al., Jet substructure at the tevatron and LHC: new results, new tools, new benchmarks. J. Phys. **G39**, 063001 (2012). arXiv:1201.0008
11. Atlas Collaboration, Study of jet shapes in inclusive jet production in pp collisions at $s = 7$ TeV using the ATLAS detector. Phys. Rev. **D83**, 052003 (2011). arXiv:1101.0070
12. C.M.S. Collaboration, Shape, transverse size, and charged hadron multiplicity of jets in pp collisions at 7 TeV. JHEP **1206**, 160 (2012). arXiv:1204.3170
13. ATLAS Collaboration, Jet mass and substructure of inclusive jets in $s = 7$ TeV pp collisions with the ATLAS experiment. JHEP **1205**, 128 (2012). arXiv:1203.4606
14. ATLAS Collaboration, ATLAS measurements of the properties of jets for boosted particle searches. Phys. Rev. **D86**, 072006 (2012). arXiv:1206.5369
15. ATLAS Collaboration, Performance of jet substructure techniques for large-R jets in proton-proton collisions at $s = 7$ TeV using the ATLAS detector. Submitted to JHEP (2013). arXiv:1306.4945
16. C.M.S. Collaboration, Studies of jet mass in dijet and W/Z + jet events. JHEP **1305**, 090 (2013). arXiv:1303.4811
17. T. Sjostrand, S. Mrenna, P.Z. Skands, PYTHIA 6.4 physics and manual. JHEP **0605**, 026 (2006). hep-ph/0603175
18. T. Sjostrand, S. Mrenna, P.Z. Skands, A brief introduction to PYTHIA 8.1. Comput. Phys. Commun. **178**, 852–867 (2008). arXiv:0710.3820
19. G. Corcella, I. Knowles, G. Marchesini, S. Moretti, K. Odagiri et al., HERWIG 6: an event generator for hadron emission reactions with interfering gluons (including supersymmetric processes). JHEP **0101**, 010 (2001). hep-ph/0011363
20. M. Bahr et al., Herwig++ physics and manual. Eur. Phys. J. **C58**, 639–707 (2008). arXiv:0803.0883
21. J.M. Campbell, R.K. Ellis, Next-to-leading order corrections to W^+ 2 jet and Z^+ 2 jet production at hadron colliders. Phys. Rev. **D65**, 113007 (2002). hep-ph/0202176
22. Z. Nagy, Next-to-leading order calculation of three jet observables in hadron hadron collision. Phys. Rev. **D68**, 094002 (2003). hep-ph/0307268
23. T. Gleisberg et al., Event generation with SHERPA 1.1. JHEP **02**, 007 (2009), arXiv:0811.4622
24. M. Dasgupta, G. Salam, Resummation of nonglobal QCD observables. Phys. Lett. **B512**, 323–330 (2001). hep-ph/0104277
25. Y. Hatta, T. Ueda, Resummation of non-global logarithms at finite N_c. arXiv:1304.6930
26. M. Dasgupta, G.P. Salam, Accounting for coherence in interjet E(t) flow: a case study. JHEP **0203**, 017 (2002). hep-ph/0203009
27. A. Banfi, G. Marchesini, G. Smye, Away from jet energy flow. JHEP **0208**, 006 (2002). hep-ph/0206076
28. A. Banfi, M. Dasgupta, Problems in resumming interjet energy flows with k_t clustering. Phys. Lett. **B628**, 49–56 (2005). hep-ph/0508159
29. Y. Delenda, R. Appleby, M. Dasgupta, A. Banfi, On QCD resummation with k(t) clustering. JHEP **0612**, 044 (2006). hep-ph/0610242
30. R. Kelley, J.R. Walsh, S. Zuberi, Disentangling clustering effects in jet algorithms. arXiv:1203.2923
31. Y. Delenda, K. Khelifa-Kerfa, On the resummation of clustering logarithms for non-global observables. JHEP **1209**, 109 (2012). arXiv:1207.4528
32. S. Catani, D. de Florian, G. Rodrigo, Space-like (versus time-like) collinear limits in QCD: is factorization violated ? JHEP **1207**, 026 (2012). arXiv:1112.4405
33. J.R. Forshaw, M.H. Seymour, A. Siodmok, On the breaking of collinear factorization in QCD. JHEP **1211**, 066 (2012). arXiv:1206.6363
34. A. Kyrieleis, M. Seymour, The colour evolution of the process $qq \to qqg$. JHEP **0601**, 085 (2006). hep-ph/0510089
35. J. Forshaw, A. Kyrieleis, M. Seymour, Super-leading logarithms in non-global observables in QCD: colour basis independent calculation. JHEP **0809**, 128 (2008). arXiv:0808.1269
36. A. Banfi, G.P. Salam, G. Zanderighi, Phenomenology of event shapes at hadron colliders. JHEP **1006**, 038 (2010). arXiv:1001.4082
37. V. Mateu, I.W. Stewart, J. Thaler, Power corrections to event shapes with mass-dependent operators. Phys. Rev. **D87**, 014025 (2013). arXiv:1209.3781
38. H.-n. Li, Z. Li, C.-P. Yuan, QCD resummation for jet substructures. Phys. Rev. Lett. **107**, 152001 (2011). arXiv:1107.4535
39. M. Dasgupta, K. Khelifa-Kerfa, S. Marzani, M. Spannowsky, On jet mass distributions in Z+jet and dijet processes at the LHC. JHEP **1210**, 126 (2012). arXiv:1207.1640
40. M. Sjodahl, Color evolution of $2 \to 3$ processes. JHEP **0812**, 083 (2008). arXiv:0807.0555
41. M. Sjodahl, Color structure for soft gluon resummation: a General recipe. JHEP **0909**, 087 (2009). arXiv:0906.1121

42. J.R. Forshaw, A. Kyrieleis, M. Seymour, Super-leading logarithms in non-global observables in QCD. JHEP **0608**, 059 (2006). hep-ph/0604094
43. R.M. Duran Delgado, J.R. Forshaw, S. Marzani, M.H. Seymour, The dijet cross section with a jet veto. JHEP **1108**, 157 (2011). arXiv:1107.2084
44. A. Banfi, M. Dasgupta, K. Khelifa-Kerfa, S. Marzani, Non-global logarithms and jet algorithms in high-pT jet shapes. JHEP **1008**, 064 (2010). arXiv:1004.3483
45. K. Khelifa-Kerfa, Non-global logs and clustering impact on jet mass with a jet veto distribution. JHEP **1202**, 072 (2012). arXiv:1111.2016
46. Y.-T. Chien, R. Kelley, M.D. Schwartz, H.X. Zhu, Resummation of jet mass at Hadron Colliders. Phys. Rev. **D87**, 014010 (2013). arXiv:1208.0010
47. E. Laenen, G. Oderda, G.F. Sterman, Resummation of threshold corrections for single particle inclusive cross-sections. Phys. Lett. B **438**, 173 (1998). hep-ph/9806467
48. T.T. Jouttenus, I.W. Stewart, F.J. Tackmann, W.J. Waalewijn, Jet mass spectra in Higgs + one jet at NNLL. Phys. Rev. **D88**, 054031 (2013). doi:10.1103/PhysRevD.88.054031. arXiv:1302.0846
49. I.W. Stewart, F.J. Tackmann, W.J. Waalewijn, N-jettiness: an inclusive event shape to veto jets. Phys. Rev. Lett. **105**, 092002 (2010). arXiv:1004.2489
50. C.F. Berger, C. Marcantonini, I.W. Stewart, F.J. Tackmann, W.J. Waalewijn, Higgs production with a central jet veto at NNLL+NNLO. JHEP **1104**, 092 (2011). arXiv:1012.4480
51. I.W. Stewart, F.J. Tackmann, W.J. Waalewijn, The beam thrust cross section for Drell-Yan at NNLL order. Phys. Rev. Lett. **106**, 032001 (2011). arXiv:1005.4060
52. I.W. Stewart, F.J. Tackmann, W.J. Waalewijn, Factorization at the LHC: from PDFs to initial state jets. Phys. Rev. D **81**, 094035 (2010). arXiv:0910.0467
53. A. Hornig, C. Lee, I.W. Stewart, J.R. Walsh, S. Zuberi, Non-global structure of the O(α_s^2) Dijet soft function. JHEP **1108**, 054 (2011). arXiv:1105.4628
54. I. Feige, M.D. Schwartz, I.W. Stewart, J. Thaler, Precision jet substructure from boosted event shapes. Phys. Rev. Lett. **109**, 092001 (2012). arXiv:1204.3898
55. M. Dasgupta, A. Fregoso, S. Marzani, G.P. Salam, Towards an understanding of jet substructure. JHEP **1309**, 029 (2013). doi:10.1007/JHEP09(2013)029. arXiv:1307.0007
56. M. Dasgupta, A. Fregoso, S. Marzani, A. Powling, Jet substructure with analytical methods. Eur. Phys. J. C **73**, 2623 (2013). doi:10.1140/epj/s10052-013-2623-3. arXiv:1307.0013
57. A.J. Larkoski, G.P. Salam, J. Thaler, Energy correlation functions for jet substructure. JHEP **1306**, 108 (2013). arXiv:1305.0007
58. P. Skands, The Perugia tunes, FERMILAB-CONF-09-113-T. arXiv:0905.3418
59. P.Z. Skands, Tuning Monte Carlo generators: the Perugia Tunes. Phys. Rev. **D82**, 074018 (2010). arXiv:1005.3457
60. CTEQ Collaboration Collaboration, H. Lai et al., Global QCD analysis of parton structure of the nucleon: CTEQ5 parton distributions. Eur. Phys. J. **C12**, 375–392 (2000). hep-ph/9903282
61. T. Sjostrand, P. Eden, C. Friberg, L. Lonnblad, G. Miu et al., High-energy physics event generation with PYTHIA 6.1. Comput. Phys. Commun. **135**, 238–259 (2001). hep-ph/0010017
62. R. Corke, T. Sjostrand, Interleaved parton showers and tuning prospects. JHEP **03**, 032 (2011). arXiv:1011.1759
63. J. Pumplin et al. New generation of parton distributions with uncertainties from global QCD analysis. JHEP **07**, 012 (2002). hep-ph/0201195
64. S. Gieseke, C. Rohr, A. Siodmok, Colour reconnections in Herwig++. Eur. Phys. J. **C72**, 2225 (2012). arXiv:1206.0041
65. S. Gieseke et al. Herwig++ 2.5 release note. arXiv:1102.1672
66. H.-L. Lai et al., New parton distributions for collider physics. Phys. Rev. **D82**, 074024 (2010). arXiv:1007.2241
67. M. Cacciari, G.P. Salam, G. Soyez, FastJet user manual. Eur. Phys. J. **C72**, 1896 (2012). arXiv:1111.6097
68. M. Cacciari, G.P. Salam, G. Soyez, The catchment area of jets. JHEP **0804**, 005 (2008). arXiv:0802.1188
69. A. Buckley, J. Butterworth, L. Lonnblad, H. Hoeth, J. Monk, et al. Rivet user manual. Comput. Phys. Commun. **184**, 2803–2819 (2013). doi:10.1016/j.cpc.2013.05.021. arXiv:1003.0694
70. Boost 2012 MC Working Group web page. http://ific.uv.es/boost2012/workingroups.html
71. M. Cacciari, G.P. Salam, G. Soyez, The Anti-k(t) jet clustering algorithm. JHEP **0804**, 063 (2008). arXiv:0802.1189
72. Y.L. Dokshitzer, G.D. Leder, S. Moretti, B.R. Webber, Better jet clustering algorithms. JHEP **08**, 001 (1997). hep-ph/9707323
73. M. Wobisch, T. Wengler, Hadronization corrections to jet cross sections in deep- inelastic scattering. hep-ph/9907280
74. M. Jankowiak, A.J. Larkoski, Jet substructure without trees. JHEP **1106**, 057 (2011). arXiv:1104.1646
75. J. Thaler, K. Van Tilburg, Identifying boosted objects with N-subjettiness. JHEP **1103**, 015 (2011). arXiv:1011.2268
76. J. Thaler, K. Van Tilburg, Maximizing boosted top identification by minimizing N-subjettiness. JHEP **1202**, 093 (2012). arXiv:1108.2701
77. C.F. Berger, T. Kucs, G.F. Sterman, Event shape/energy flow correlations. Phys. Rev. **D68**, 014012 (2003). hep-ph/0303051
78. L.G. Almeida, S.J. Lee, G. Perez, G.F. Sterman, I. Sung et al., Substructure of high-p_T jets at the LHC. Phys. Rev. **D79**, 074017 (2009). arXiv:0807.0234
79. S. Chekanov, J. Proudfoot, Searches for TeV-scale particles at the LHC using jet shapes. Phys. Rev. **D81**, 114038 (2010). arXiv:1002.3982
80. J. Gallicchio, M.D. Schwartz, Seeing in color: jet superstructure. Phys. Rev. Lett. **105**, 022001 (2010). arXiv:1001.5027
81. A. Hook, M. Jankowiak, J.G. Wacker, Jet dipolarity: top tagging with color flow. JHEP **1204**, 007 (2012). doi:10.1007/JHEP04(2012)007. arXiv:1102.1012
82. R. Feynman, R. Field, A parameterization of the properties of Quark Jets. Nucl. Phys. B **136** 1 (1978), in *Selected papers of Richard Feynman*, ed. by L.M.Brown, pp. 763–838. doi:10.1016/0550-3213(78)90015-9
83. W.J. Waalewijn, Calculating the charge of a jet. Phys. Rev. D **86**, 094030 (2012)
84. D. Krohn, T. Lin, M.D. Schwartz, W.J. Waalewijn, Jet charge at the LHC. Phys. Rev. Lett. **110**, 212001 (2013). doi:10.1103/PhysRevLett.110.212001. arXiv:1209.2421
85. TOTEM Collaboration, G. Antchev et al., Luminosity-independent measurements of total, elastic and inelastic cross-sections at $\sqrt{s} = 7$ TeV. Europhys. Lett. **101**, 21004 (2013)
86. T. Sjostrand, S. Mrenna, P.Z. Skands, PYTHIA 6.4 physics and manual. JHEP **05**, 026 (2006). hep-ph/0603175. For older versions, see http://home.thep.lu.se/torbjorn/Pythia.html
87. T. Sjostrand, S. Mrenna, P.Z. Skands, A brief introduction to PYTHIA 8.1. Comput. Phys. Commun. **178**, 852–867 (2008). arXiv:0710.3820
88. M. Cacciari, G.P. Salam, Pileup subtraction using jet areas. Phys. Lett. **B659**, 119–126 (2008). arXiv:0707.1378
89. M. Cacciari, J. Rojo, G.P. Salam, G. Soyez, Jet reconstruction in heavy ion collisions. Eur. Phys. J. C **71**, 1539 (2011). arXiv:1010.1759
90. M. Cacciari, J. Rojo, G.P. Salam, G. Soyez, Quantifying the performance of jet definitions for kinematic reconstruction at the LHC. JHEP **0812**, 032 (2008). arXiv:0810.1304
91. M. Wobisch, Measurement and QCD analysis of jet cross-sections in deep inelastic positron proton collisions at $\sqrt{s} = 300$ GeV.

PhD thesis, University of Hamburg, 2000. DESY-THESIS-2000-049
92. S.D. Ellis, C.K. Vermilion, J.R. Walsh, Techniques for improved heavy particle searches with jet substructure. Phys. Rev. **D80**, 051501 (2009). arXiv:0903.5081
93. J.M. Campbell, R. Ellis, MCFM for the Tevatron and the LHC. Nucl. Phys. Proc. Suppl. **205–206**, 10–15 (2010). arXiv:1007.3492
94. T. Plehn, M. Spannowsky, Top tagging. J. Phys. **G39**, 083001 (2012). arXiv:1112.4441
95. D.E. Kaplan, K. Rehermann, M.D. Schwartz, B. Tweedie, Top tagging: a method for identifying boosted hadronically decaying top quarks. Phys. Rev. Lett. **101**, 142001 (2008). arXiv:0806.0848
96. CMS Collaboration, A cambridge-aachen (c-a) based jet algorithm for boosted top-jet tagging. Tech. Rep. CMS PAS JME-09-001, CERN, Geneva, 2009
97. T. Plehn, M. Spannowsky, M. Takeuchi, D. Zerwas, Stop reconstruction with tagged tops. JHEP **1010**, 078 (2010). doi:10.1007/JHEP10(2010)078. arXiv:1006.2833
98. ATLAS Collaboration, Search for resonances decaying into top-quark pairs using fully hadronic decays in pp collisions with ATLAS at $\sqrt{s} = 7$ TeV. JHEP **1301**, 116 (2013). arXiv:1211.2202
99. J.M. Butterworth, B.E. Cox, J.R. Forshaw, WW scattering at the CERN LHC. Phys. Rev. **D65**, 096014 (2002). hep-ph/0201098
100. D.E. Soper, M. Spannowsky, combining subjet algorithms to enhance ZH detection at the LHC. arXiv:1005.0417
101. D.E. Soper, M. Spannowsky, Finding top quarks with shower deconstruction. Phys. Rev. **D87**, 054012 (2013). arXiv:1211.3140
102. C.D.F. Collaboration, T. Aaltonen et al., Limits on the production of narrow $t\bar{t}$ resonances in $p\overline{p}$ collisions at $\sqrt{s} = 1.96 - TeV$. Phys. Rev. **D77**, 051102 (2008). arXiv:0710.5335
103. D0 Collaboration, V. Abazov et al., Search for$t\bar{t}$ resonances in the lepton plus jets final state in $p\overline{p}$ collisions at $\sqrt{s} = 1.96 - TeV$. Phys. Lett. **B668**, 98–104 (2008). arXiv:0804.3664
104. C.D.F. Collaboration, T. Aaltonen et al., Search for resonant production of tt decaying to jets in $t\bar{t}$ collisions at $\sqrt{s} = 1.96 - TeV$. Phys. Rev. **D84**, 072003 (2011). arXiv:1108.4755
105. D0 Collaboration, V.M. Abazov et al., Search for a narrow $t\bar{t}$ resonance in $p\overline{p}$ collisions at $\sqrt{s} = 1.96 - TeV$. Phys. Rev. **D85**, 051101 (2012). arXiv:1111.1271
106. CDF Collaboration Collaboration, T. Aaltonen et al., Search for resonant $t\bar{t}$ production in the semi-leptonic decay mode using the full CDF data set. Phys. Rev. Lett. **110**, 121802 (2013). arXiv:1211.5363
107. ATLAS Collaboration, A search for $t\bar{t}$ resonances with the ATLAS detector in 2.05 $fb^{-}1$ of proton-proton collisions at $\sqrt{s} = 7TeV$. Eur. Phys. J. **C72**, 2083 (2012). arXiv:1205.5371
108. ATLAS Collaboration, A search for $t\bar{t}$ resonances in lepton+jets events with highly boosted top quarks collected in $p\overline{p}$ collisions at$\sqrt{s} = 7TeV$ with the ATLAS detector. JHEP **1209**, 041 (2012). arXiv:1207.2409
109. ATLAS Collaboration, A search for ttbar resonances in the lepton plus jets final state with ATLAS using 4.7 $fb^{-}1$ of pp collisions at s = 7 TeV. Phys. Rev. **D88**, 012004 (2013). doi:10.1103/PhysRevD.88.012004. arXiv:1305.2756
110. ATLAS Collaboration, A search for t t-bar resonances in the lepton plus jets final state with atlas using 14 $fb^{-}1$ of pp collisions at s = 8 TeV. Tech. Rep. ATLAS-CONF-2013-052, May, 2013
111. CMS Collaboration, Search for anomalous t t-bar production in the highly-boosted all-hadronic final state. JHEP **1209**, 029 (2012). arXiv:1204.2488
112. CMS Collaboration, Search for resonant tt production in lepton+jets events in pp collisions at s=7 TeV. JHEP **1212**, 015 (2012). arXiv:1209.4397
113. CMS Collaboration, Search for Z' resonances decaying to $t\bar{t}$ in dilepton+jets final states in $p\overline{p}$ collisions at $\sqrt{s} = 7TeV$. Phys. Rev. **D87**, 072002 (2013). arXiv:1211.3338
114. CMS collaboration, Search for anomalous top quark pair production in the boosted all-hadronic final state using pp collisions at $\sqrt{s} = 8$ TeV. Tech. Rep. CMS-PAS-B2G-12-005, CERN, Geneva, May, 2013
115. CMS collaboration, Search for t t-bar resonances in semileptonic final states in pp collisions at $\sqrt{s} = 8$ TeV. Tech. Rep. CMS-PAS-B2G-12-006, CERN, Geneva, April, 2013
116. M. Backovic, J. Juknevich, G. Perez, Boosting the standard model Higgs signal with the template overlap method. JHEP **1307**, 114 (2013). arXiv:1212.2977
117. S.D. Ellis, A. Hornig, T.S. Roy, D. Krohn, M.D. Schwartz, Qjets: a non-deterministic approach to tree-based jet substructure. Phys. Rev. Lett. **108**, 182003 (2012). arXiv:1201.1914
118. T. Plehn, G.P. Salam, M. Spannowsky, Fat jets for a light Higgs. Phys. Rev. Lett. **104**, 111801 (2010). arXiv:0910.5472