

## Boosted top production in ATLAS and CMS

MARINO ROMANO

*On behalf of the ATLAS and CMS Collaborations,  
INFN, Sezione di Bologna*

### ABSTRACT

An overview of the boosted top production analyses using data collected by the ATLAS and CMS experiments at  $\sqrt{s} = 8$  TeV and 13 TeV of proton-proton collisions at the LHC is presented. These analyses use techniques for the reconstruction of boosted objects to measure the production of top quarks at high transverse momenta. The measurements are optimized for the different final states and for different ranges of the transverse momenta of the particles involved, improving on measurements with traditional objects reconstruction based on the combination of resolved objects.

### PRESENTED AT

The Fifth Annual Conference  
on Large Hadron Collider Physics  
Shanghai Jiao Tong University, Shanghai, China  
May 15-20, 2017



# 1 Introduction

The top quark is the heaviest quark, significantly heavier than its partner, the bottom quark. Once the bottom quark was experimentally discovered in 1977, the existence of a charge  $2/3$  quark in the third quark generation was expected to preserve the Standard Model (SM) renormalizability. The top quark was discovered by both the CDF [1] and DØ [2] collaborations in 1995 at the Tevatron collider.

The top quark is special not only due to its large mass, but also due to its short lifetime which prevents it from hadronizing before decaying, i.e. there are no bound state hadrons made of top quarks. This allows to experimentally test the properties of the “bare” top quark itself through its decay products without diluting information in the hadronization process.

As the properties of the top quark are precisely predicted by the SM, top quark physics provides sensitive probes to test the validity of the SM and a tool to investigate the Higgs boson properties and to potentially discover physics beyond the SM.

Top quark pairs,  $t\bar{t}$ , are produced in proton-proton collisions via the strong interaction of the colliding partons. In a leading order picture, at  $\sqrt{s} = 8$  TeV about 85% of  $t\bar{t}$  pairs are produced by gluon-gluon interactions. The remaining 15% is produced in quark and anti-quark annihilation processes.

A top quark decays almost exclusively into a  $W$  boson and  $b$  quark. Therefore,  $t\bar{t}$  events can be classified through the decay products of the two  $W$  bosons.  $W$  bosons can decay into a quark pair or a lepton-neutrino pair, so it is possible to define three final state topologies: full hadronic, when both  $W$  bosons decay in a quark pair; dileptonic, when both  $W$  bosons decay in a lepton-neutrino pair; lepton+jets, when a  $W$  boson decays in a quark pair and the other in a lepton-neutrino pair.

When the top quark is produced with a very high momentum, its decay products are collimated, affecting the reconstruction efficiency of each individual decay product. This specific phase space is referred to as “boosted regime”, and can be particularly interesting when probing the Standard Model at extreme energies or when searching for heavy particles decaying in pair of top quarks. In addition, a boosted topology study can complement a resolved analysis by recovering events otherwise mis-reconstructed, or can provide additional advantages, such as a reduction of the combinatorial background due to less final state objects.

## 2 Boosted top reconstruction in ATLAS and CMS

The defining property of boosted object decays is the fact that their decay products appear collimated in the momentum direction of the boosted mother particle in the rest frame of the detector. Therefore, the decay products merge into a single large radius (large- $R$ ) jet with a characteristic substructure that can be exploited to distinguish these jets from those initiated from gluons or light-quarks.

At high-luminosity hadron colliders, a major obstacle for analyses relying on large- $R$  jets is the presence of pile-up and the underlying event, both of which lead to soft, wide-angle contaminations that dilute the jet substructure. Their impact is mitigated using the so-called “grooming” techniques. A large number of combinations of large- $R$  jet reconstruction algorithms, substructure variables and grooming algorithms to identify boosted top-quarks have been studied in the ATLAS [3] and CMS [4] experiments. Table 1 summarizes the jet reconstruction and top tagging tools generally used by each experiment.

The rapid growth of the field makes it impossible to cover every single technique within the scope of this document. The full extent of those studies is documented in [5, 6] for ATLAS and [7, 8] for CMS. The top-tagging efficiencies as a function of the transverse momentum of the reconstructed jets for several working points used by ATLAS and CMS are shown in Figure 1.

	ATLAS	CMS
Jet reconstruction	anti- $k_T$ $R = 1.0$	C/A $R = 1.5$ and anti- $k_T$ $R = 0.8$
Grooming	Trimming, $R_{\text{subject}} = 0.2$ , $f_{\text{cut}} < 5\%$	Soft-drop, $z_{\text{cut}} = 0.1$ and $\beta = 0$
Top tagging	Jet mass, $d_{12}$ (8 TeV only) and $\tau_{3/2}$	Jet mass, $\tau_{3/1}$ and $\tau_{3/2}$

Table 1: Summary of the jet reconstruction, grooming and top tagging criteria employed by ATLAS and CMS.

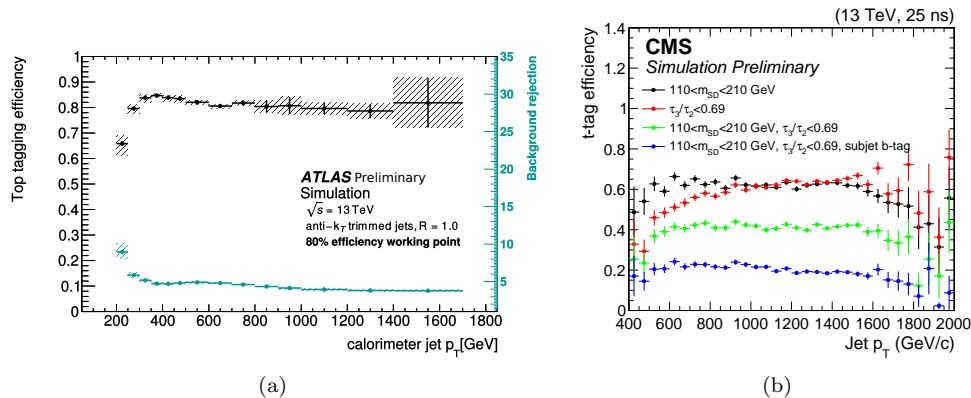


Figure 1: (a) top-tagging efficiencies [6] as a function of the  $p_T$  of the reconstructed jet for the “80%” working points used in the differential cross section measurements performed by ATLAS in the lepton+jets channel [9]. (b) top-tagging efficiencies as a function of the  $p_T$  of the reconstructed jet for different top-tagging criteria employed by the CMS experiment [8].

### 3 Top pair differential cross section measurements

Since the predictions for the inclusive  $t\bar{t}$  production cross section are in very good agreement with the measurements and considering the rapid increase of the integrated luminosity, the focus is starting to switch to the “differential” measurements of the top properties, such as the cross section as a function of the  $t\bar{t}$  system kinematic variables. Such experimental measurements, performed in different channels, allow precision tests of the predictions of perturbative QCD. New physics may also give rise to additional  $t\bar{t}$  production mechanisms or modifications of the top quark decay channels, that can be discovered looking at the differential spectra.

Top pair differential cross section in the boosted regime has been measured in the lepton+jets and full hadronic channels by ATLAS [9, 10, 11], and CMS [12, 13, 14] at a center of mass energy of 8 and 13 TeV.

#### 3.1 Top pair differential cross sections in the lepton+jets channel at 13 TeV

The measurement of  $t\bar{t}$  differential cross sections in the lepton+jets decay channel has been performed by ATLAS, in both the resolved and boosted regimes, at a center-of-mass energy of  $\sqrt{s} = 13$  TeV [9]. The integrated luminosity corresponds to  $L = 3.2 \text{ fb}^{-1}$ , recorded in 2015 in proton-proton collisions. In the boosted regime, the absolute ( $\frac{d\sigma}{dX}$ ) and relative ( $\frac{1}{\sigma} \frac{d\sigma}{dX}$ ) differential cross sections have been measured as a function of the transverse momentum and absolute rapidity of the hadronically decaying top ( $X = p_T^{t_{\text{had}}}, |y^{t_{\text{had}}}|$ ).

Boosted top quark events in the lepton+jets channel are selected by requiring exactly one charged lepton (electron or muon), at least one small- $R$  jet and exactly one large- $R$  jet. The large- $R$  jet is tagged as a top

using the working point corresponding to an efficiency of 80% of the tagger described in [6]. The measurement is performed at particle level: the reconstructed distributions are corrected to the particle-level by means of a bayesian-inspired unfolding technique [15].

In Figure 2(a) and Figure 2(b) are presented the measured absolute differential cross sections as a function of the transverse momentum and absolute rapidity of the hadronically-decaying top quark. The results are in fair agreement with the predictions over a wide kinematic range for the different observables. However, most generators predict an higher cross section at high values of  $p_T^{\text{had}}$  compared to the observations in data, as it has already been shown by the different Run 1 measurements by both ATLAS and CMS at 8 TeV [10, 12].

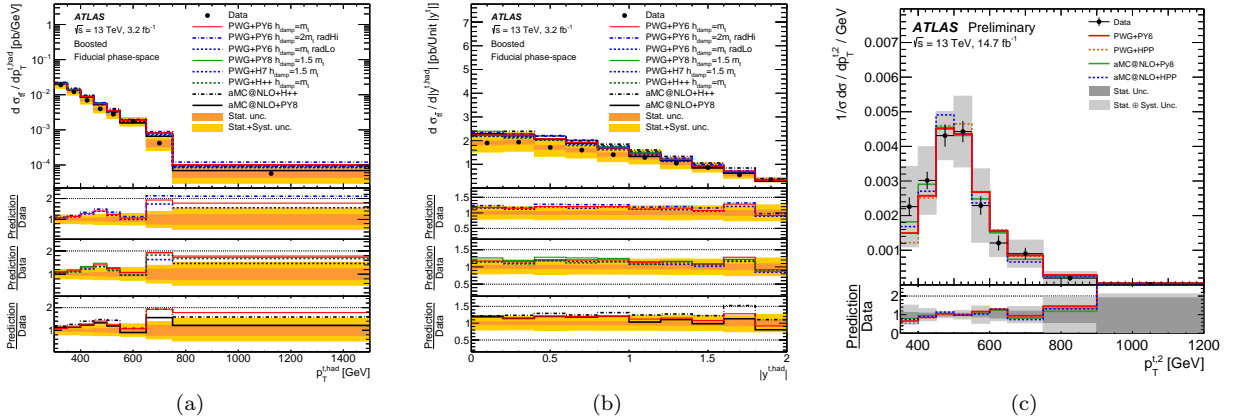


Figure 2: Fiducial phase-space differential cross sections as a function of (a) the transverse momentum ( $p_T^{\text{had}}$ ) and (b) the absolute value of the rapidity ( $|y^{\text{had}}|$ ) of the hadronic top quark measured by the ATLAS experiment at  $\sqrt{s} = 13$  TeV in the lepton+jets channel [9]. (c) normalized fiducial phase-space differential cross sections as a function of the transverse momentum of the second-leading top-quark jet measured by the ATLAS experiment at  $\sqrt{s} = 13$  TeV in the full hadronic channel [11].

### 3.2 Top pair differential cross sections in the full hadronic channel at 13 TeV

The measurement of  $t\bar{t}$  differential cross sections in the full hadronic channel has been performed by both ATLAS [11] and CMS [13] at a centre-of-mass energy  $\sqrt{s} = 13$  TeV. The integrated luminosities for these analyses correspond to  $L = 14.7 \text{ fb}^{-1}$  (recorded in 2015 and 2016 in proton-proton collisions) and  $L = 2.53 \text{ fb}^{-1}$  (recorded in 2015), respectively.

Boosted top quark events in full hadronic channels are selected by requiring two high- $p_T$  top-tagged jets, according the procedures described in [6, 8].

For these analyses, the main challenge is represented by the estimation of the multijet background. Both measurements employ data-driven methods to estimate this background:

- the ATLAS measurement makes use of an extension of the ABCD method to extract the yield in the signal region based on the yields in control regions using the formula

$$S = \frac{1}{2} \left( \frac{G}{A} + \frac{H}{B} \right) C, \quad (1)$$

where the control regions  $A, B, C, G, H$  are defined in terms of the number of  $b$ - and  $t$ -tagged jets;

- the CMS measurement makes use of a template fit of the distribution of the soft-drop groomed jet mass to extract both the signal and multijet event yields. The signal and QCD templates are taken from the simulation and the control sample in data, respectively, corrected with a MC transition factor.

ATLAS measured the normalized differential cross section in the fiducial phase space as a function of the  $p_T$  and absolute rapidity of the leading and subleading top, the  $p_T$ , mass and absolute rapidity of the top-antitop system and additional variables sensitive to effects of initial- and final-state radiation, to the different parton distribution functions (PDF), and to non-resonant processes including particles beyond the SM. The differential cross section as a function of the  $p_T$  of the subleading top is shown in Figure 2(c).

CMS measured the inclusive cross section, detector and parton level differential cross sections as a function of the  $p_T$  of the leading top quark, shown in Figure 3(a) and Figure 3(b). The measured inclusive cross section is

$$\sigma_{t\bar{t}} = 727 \pm 46(\text{stat.})_{-112}^{+115}(\text{syst.}) \pm 20(\text{lumi.}) \text{ pb};$$

in agreement, within the uncertainties, with the theoretical cross section  $\sigma_{t\bar{t}}^{th} = 832_{-29}^{+20}(\text{scale}) \pm 35(\text{PDF} + \alpha_s) \text{ pb}$ , as calculated with the TOP++ program [16] at next-to-next-to-leading order (NNLO) in perturbative QCD, including soft-gluon resummation at next-to-next-to-leading-log order.

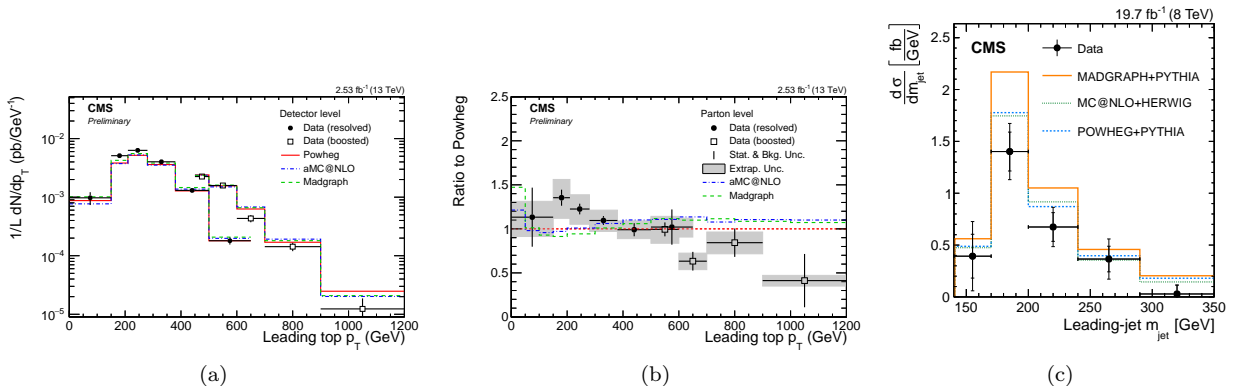


Figure 3: (a) detector-level differential cross section as a function of the transverse momentum of the leading top measured by the CMS experiment at  $\sqrt{s} = 13$  TeV in the full hadronic channel using resolved (black dots) and boosted (empty boxes) techniques [13]. (b) ratio between the parton-level differential cross section as a function of the transverse momentum of the leading top measured by the CMS experiment at  $\sqrt{s} = 13$  TeV in the full hadronic channel using resolved (black dots) and boosted (empty boxes) techniques and the prediction obtained with the POWHEG+PYTHIA event generator [13]. (c) particle-level absolute differential cross section as a function of the leading jet mass measured by the CMS experiment at  $\sqrt{s} = 8$  TeV in the lepton+jets channel [14].

### 3.3 Measurement of the jet mass distribution in boosted top pair production at 8 TeV

A detailed understanding of jet substructure observables, most importantly the jet mass  $m_{jet}$ , is crucial for LHC analyses in boosted topologies. CMS has performed, for the first time, a measurement of the differential  $t\bar{t}$  production cross section as a function of the leading jet mass in fully-merged top quark decays [14]. Data collected with the CMS detector in  $pp$  collisions at  $\sqrt{s} = 8$  TeV are used, corresponding to an integrated luminosity of  $19.7 \text{ fb}^{-1}$ . The measurement is carried out in the lepton+jets channel.

The products of the hadronic decay are reconstructed with a single Cambridge/Aachen jet with distance parameter  $R = 1.2$ , and transverse momentum  $p_T > 400$  GeV. No top tagging algorithm is applied to the large- $R$  jet, in order to avoid any bias to the  $m_{jet}$  distribution. A bias would for example be introduced by selecting the leading jet based on the number of subjets or requiring a certain maximum value of the N-subjettiness as applied in top-tagging algorithms.

The measured absolute differential cross section is shown in Figure 3(c). The measurement gives a total cross section of  $103.5 \pm 11.2(\text{stat.}) \pm 10.9(\text{syst.}) \pm 9.4(\text{model}) \text{ fb}$ . The predicted cross sections from the

MADGRAPH+PYTHIA and POWHEG+PYTHIA  $t\bar{t}$  simulations, normalized to a total  $t\bar{t}$  cross section of 252.9 pb, are 160 fb and 134 fb respectively.

The peak position of the  $m_{\text{jet}}$  distribution is also sensitive to the top quark mass  $m_t$ . The value of  $m_t$  is determined using the normalized measurement since only the shape of the  $m_{\text{jet}}$  distribution can be reliably calculated. The measured value of  $m_t$  is extracted from data using a fit of the  $\chi^2$  evaluated comparing the measured  $m_{\text{jet}}$  distribution with templates obtained from MADGRAPH+PYTHIA predictions with different values of  $m_t$ . The result is

$$m_t = 171.8 \pm 5.4 \text{ (stat.)} \pm 3.0 \text{ (syst.)} \pm 5.5 \text{ (model)} \pm 4.6 \text{ (theory)} \text{ GeV} = 171.8 \pm 9.5 \text{ GeV},$$

consistent with the current LHC+Tevatron average,  $173.34 \pm 0.27 \text{ (stat.)} \pm 0.71 \text{ (syst.)} \text{ GeV}$  [17].

## 4 Charge asymmetry measurement in highly boosted top pair production at 8 TeV

In the  $pp \rightarrow t\bar{t}$  process the angular distributions of top and anti-top quarks are expected to present a subtle difference, which could be enhanced by processes not included in the Standard Model. In the SM, a forward-backward asymmetry ( $A_{\text{FB}}$ ), of order  $\alpha_s$ , is expected at a proton-antiproton ( $p\bar{p}$ ) collider such as the Tevatron.

At the LHC, the forward-backward asymmetry is not present due to the symmetric initial state, but a related charge asymmetry ( $A_{\text{C}}$ ) is expected in the distribution of the difference of absolute rapidities of the top and anti-top quarks:

$$A_{\text{C}} = \frac{N(\Delta|y| > 0) - N(\Delta|y| < 0)}{N(\Delta|y| > 0) + N(\Delta|y| < 0)}, \quad (2)$$

where  $\Delta|y| = |y_t| - |y_{\bar{t}}|$  and  $y$  denotes the rapidity of the top and anti-top quarks.

The ATLAS experiment has performed a measurement of the rapidity-dependent charge asymmetry for boosted top-quark pair production in the lepton+jets channel [18] using the boosted top reconstruction and tagging techniques described in [5].

The analysis is performed using  $20.3 \text{ fb}^{-1}$  of  $pp$  collision data at  $\sqrt{s} = 8 \text{ TeV}$  collected by the ATLAS experiment at the LHC. The charge asymmetry in a fiducial region with large invariant mass of the top-quark pair ( $m_{t\bar{t}} > 0.75 \text{ TeV}$ ) and an absolute rapidity difference of the top and antitop quark candidates within  $-2 < |y_t| - |y_{\bar{t}}| < 2$  is measured to be  $A_{\text{C}} = (4.2 \pm 3.2) \%$ , in agreement with the Standard Model prediction at next-to-leading order  $A_{\text{C}}^{\text{th}} = (1.60 \pm 0.04) \%$  [19].

A differential measurement has been also performed in three  $t\bar{t}$  mass bins ( $0.75 \text{ TeV} < m_{t\bar{t}} < 0.9 \text{ TeV}$ ,  $0.9 \text{ TeV} < m_{t\bar{t}} < 1.3 \text{ TeV}$  and  $m_{t\bar{t}} > 1.3 \text{ TeV}$ ). The most significant deviation from the SM prediction (around  $1.6\sigma$ ) is observed in the mass bin that ranges from 0.9 TeV to 1.3 TeV. The other two mass bins yield values compatible with the SM prediction within  $1\sigma$ .

## 5 Summary and outlook

LHC offers a unique opportunity to explore extreme topologies through boosted tops. Several boosted top reconstruction algorithms have been (and are being) developed by ATLAS and CMS: jet grooming procedures allow for stability in high pileup conditions and the use of substructure variables improves the background discrimination.

Boosted tops have been used in SM measurements and BSM searches (not presented in this paper): ATLAS and CMS have measured the  $t\bar{t}$  differential cross sections using boosted tops in both the lepton+jets and full hadronic channel and the charge asymmetry in the lepton+jets channel. We are entering an era where data statistics is not the limiting factor: the measurements are already limited by uncertainties related to the large- $R$  jets and signal modelling.

The measurements presented in this paper can be used to improve future analyses:

- by adding better constraints to the generator parameters;

- by improving the understanding of the jet mass distribution.

In the light of the ongoing data taking at the LHC at  $\sqrt{s} = 13$  TeV, boosted techniques will gain even more importance for measurements of top quarks with high transverse momenta but also for searches for new phenomena at the TeV scale. The experience gained with boosted techniques in Run-1 and early Run-2 analyses is a promising basis for the upcoming challenges.

## References

- [1] CDF Collaboration, Phys. Rev. Lett., **74** (1995) 2632 [arXiv:hep-ex/9503002].
- [2] DØ Collaboration, Phys. Rev. Lett., **74** (1995) 2626 [arXiv:hep-ex/9503003].
- [3] ATLAS Collaboration, JINST **3**, S08003 (2008).
- [4] CMS Collaboration, JINST **3**, S08004 (2008).
- [5] ATLAS Collaboration, JHEP **1606** (2016) 093 [arXiv:1603.03127 [hep-ex]].
- [6] ATLAS Collaboration, ATL-PHYS-PUB-2015-053, <https://cds.cern.ch/record/2116351>.
- [7] CMS Collaboration, CMS-PAS-JME-13-007, <https://cds.cern.ch/record/1647419>.
- [8] CMS Collaboration, CMS-PAS-JME-16-003, <https://cds.cern.ch/record/2256875>.
- [9] ATLAS Collaboration, arXiv:1708.00727 [hep-ex].
- [10] ATLAS Collaboration, Phys. Rev. D **93** (2016) no.3, 032009 [arXiv:1510.03818 [hep-ex]].
- [11] ATLAS Collaboration, ATLAS-CONF-2016-100, <https://cds.cern.ch/record/2217231>.
- [12] CMS Collaboration, Phys. Rev. D **94** (2016) no.7, 072002 [arXiv:1605.00116 [hep-ex]].
- [13] CMS Collaboration, CMS-PAS-TOP-16-013, <https://cds.cern.ch/record/2161138>.
- [14] CMS Collaboration, Eur. Phys. J. C **77**, no. 7, 467 (2017) [arXiv:1703.06330 [hep-ex]].
- [15] G. D’Agostini, Nucl. Instrum. Meth. A **362** (1995) 487.
- [16] M. Czakon, P. Fiedler and A. Mitov, Phys. Rev. Lett. **110**, 252004 (2013) [arXiv:1303.6254 [hep-ph]].
- [17] ATLAS, CDF, CMS, DØ Collaborations, arXiv:1403.4427.
- [18] ATLAS Collaboration, Phys. Lett. B **756**, 52 (2016) [arXiv:1512.06092 [hep-ex]].
- [19] J. H. Kuhn and G. Rodrigo, JHEP **01** (2012) 063, arXiv:1109.6830 [hep-ph].