Available on CMS information server **CMS CR -2017/204**

The Compact Muon Solenoid Experiment

Conference Report

Mailing address: CMS CERN, CH-1211 GENEVA 23, Switzerland

03 August 2017 (v2)

Top quark event modelling and generators in CMS

Efe Yazgan for the CMS Collaboration

Abstract

Recent top quark event modelling studies done using CMS proton-proton data collected at a centre of mass energies of 8 and 13 TeV and state-of-the-art theoretical predictions accurate to next-to-leading order QCD interfaced with PYTHIA and HERWIG event generators are summarised. The particle-level top quark (pseudo-top), underlying event measurement in events and parton shower tuning using events are discussed.

Presented at *EPS-HEP2017 EPS Conference on High Energy Physics*

PROCEEDINGS OF SCIENCE

Top quark modelling and generators in CMS

Efe Yazgan on behalf of the CMS Collaboration[*]
Chinese Academy of Sciences, Institute of High Energy Physics, Beijing, China
E-mail: `efe.yazgan@cern.ch`

Recent top quark event modelling studies done using CMS proton-proton data collected at a centre of mass energies of 8 and 13 TeV and state-of-the-art theoretical predictions accurate to next-to-leading order QCD interfaced with PYTHIA8 and HERWIG++ event generators are summarised. The particle-level top quark (pseudo-top), underlying event measurement in $t\bar{t}$ events and parton shower tuning using $t\bar{t}$ events are discussed.

EPS-HEP 2017, European Physical Society conference on High Energy Physics
5-12 July 2017
Venice, Italy

---

[*] Speaker.

1. Introduction

Top quark measurements provide important tests of QCD. Better understanding of perturbative and non-perturbative effects is required to obtain the highest possible precision in top quark mass, its interpretation and in other properties. Top quark measurements are also important to improve the accuracy of predictions in different phase space regions in searches for beyond the Standard Model (SM) effects. Precise top quark measurements are also be used in direct searches for new physics as well, e.g. with effective field theories expanding the SM Lagrangian. For these purposes, the uncertainties in the measurements and the predictions need to be at a level where deviations of the predictions of the Monte Carlo (MC) codes or deviations due to new physics effects be visible. State-of-the-art next-to-leading order (NLO) matrix element (ME) event generators interfaced to new parton shower (PS) codes used in LHC Run II may provide better modelling and eventually reduce the major theoretical uncertainties. The theory uncertainties can partially be tested and improved with datasets that allow differential measurements with well-defined top-quark objects. In this note, a selection of recent top quark event modelling studies from CMS[1] are discussed.

2. Particle Level Top Quark

Simulations at NLO take the finite width of the top quark into account which is important to accurately model the off-shell production of top quarks and their interference with the backgrounds. In these calculations the concept of top quark as a particle is not well-defined and MC dependent. One can only use the kinematics of the final-state particles unambiguously. A particle-level top quark (also called pseudo-top quark) can be constructed from the final-state objects after hadronisation. Using particle-level top quarks would yield smaller uncertainties from non-perturbative effects and from acceptance corrections because of the similar phase definitions at the particle and detector levels minimising MC dependence. The details of particle-level top quark definitions and their adoption in the RIVET [2] framework in the official CMS reconstruction code are discussed in [3] as a fundamental aspect for current and future measurements of differential production cross sections in both $t\bar{t}$ and single-top quark production. The results reported in [3] indicate that the particle-level top quark definition needs to be optimised depending on the production mode, the final state or the variable and the phase space being investigated.

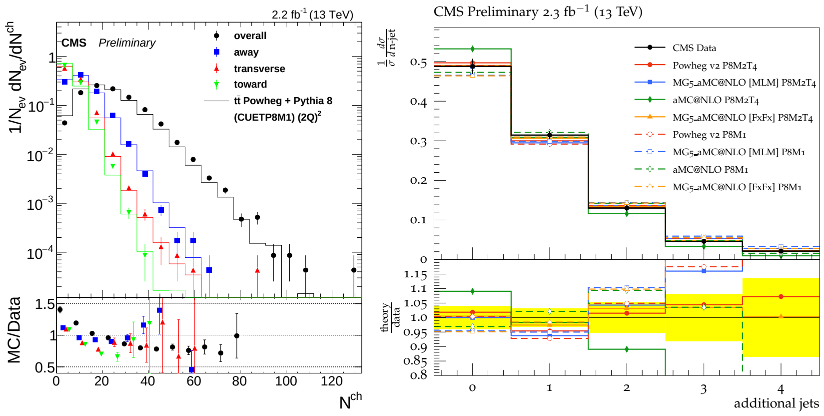
3. Underlying event and Parton shower tuning in $t\bar{t}$ events

The CMS Run I combination of direct top quark mass measurements at 7 and 8 TeV, in lepton+jets, dilepton, and all-hadronic channels yields a precision of 0.3% [4]. In this result, the dominant uncertainties are related to the event modelling. Therefore, to improve top quark mass measurements, dedicated measurements and theory studies are required. The the b quark from the top quark decay carries the colour flow. To become colourless, the b quark "connects" with the beam remnants or other coloured final particles produced in the event. A b jet in the final state can be constructed however the uncertainty in the origin of all the final states in the jet results in "odd clusters" (e.g. see [5]). Therefore, it is important to have an accurate description of b quark fragmentation and hadronisation, as well as UE is needed.

UE measurements use the highest p_T charged-particle jet, the highest E_T calorimeter jet, ad the Z-boson direction as the leading-object to define regions of $\eta - \phi$ space, in the toward, away, and transverse regions. The transverse region is particularly sensitive to the modelling of the UE. The PYTHIA8 tune CUETP8M1 and the HERWIG++ tune EE5C [6] are reconstructed by fitting the UE data at several centre-of-mass energies, where the leading object is the highest p_T charged particle or the charged-particle jet in the event. These tunes describe well the UE as measured in Z-boson production. However, very little is known about UE in heavy quark production. Ref. [7] compared detector-level top-quark production data at 13 TeV with the PYTHIA8 CUETP8M1 tune and the HERWIG++ EE5C tune after detector simulation in $t\bar{t}$ enriched events in the lepton+jets channel. Both of the parton shower models are interfaced with POWHEG V2. Fair agreement is observed between POWHEG V2+PYTHIA8 CUETP8M1 tune predictions. It is also observed that UE is sensitive to QCD scales. Fig. 1a displays the charged-particle multiplicity when the PS scale is increased from its default value. A complete measurement of UE in $t\bar{t}$ events at particle-level may lead to more precise top quark mass with better understood systematic uncertainties.

It is observed that the predictions of the NLO MC ME generators + PYTHIA8 CUETP8M1 tune [8] overshoot the $\sqrt{s} = 8$ [9, 10] and 13 TeV [10, 11] data for large jet multiplicities when out of the box parameters are used, while all other distributions are modelled well (except top quark p_T). The tune CUETP8M1 is based on the Monash tune [12]. Accurate predictions of this observable is particularly important in measurements of the Higgs boson and many new physics search analysis. To improve the description of high jet multiplicities in $t\bar{t}$ events, a number of parameters have been studied and the most sensitive ones to jet kinematics in $t\bar{t}$ events are selected and optimised. The strong coupling parameter at m_Z for initial-state radiation in the PS, α_s^{ISR} , and the h_{damp} parameter that controls the jet matching in the POWHEG V2+PYTHIA8 [13, 14, 15] setup are tuned using Run 1 data on jet activity in $t\bar{t}$ events. The Monash tune for α_s^{ISR} adopts the α_s^{FSR} value tuned to LEP event shapes. This is found to be the main cause of overproduction of jets in the MC. We tuned the values of α_s^{ISR} and h_{damp} using the jet multiplicity and leading additional jet p_T distributions in the dilepton final state measured at $\sqrt{s} = 8$ TeV using the PROFESSOR tool [16]. In this procedure, all other PYTHIA8 parameters are kept fixed to the ones in CUETP8M1 tune. It is observed that α_s^{ISR} impacts mostly $N_{jets} > 3$, while h_{damp} affects the ratio of 2-to-3-jet events and the leading additional jet p_T. This is in agreement with the fact that the leading additional jet, in the POWHEG V2+PYTHIA8 configuration, stems from the real radiation calculated by the POWHEG V2 generator. The tuning procedure yields h_{damp} $=1.581^{+0.658}_{-0.585} \times m_t$ and α_s^{ISR} $=0.1108^{+0.0145}_{-0.0142}$. The tuned α_s^{ISR} value agrees with the PDG value of $\alpha_s(M_Z) = 0.1181 \pm 0.0011$ [17] well within uncertainties. Fixing h_{damp} to its default value of m_t, a re-tuning of α_s^{ISR} alone to the same data yields α_s^{ISR} $=0.115^{+0.021}_{-0.019}$ [18] again in agreement with the PDG value. POWHEG V2+PYTHIA8 with these optimised parameters cures the overshoot of the CUETP8M1 at high jet multiplicities.

The jet activity mainly constrains those parameters that control the probability for parton emission and the interplay between hard and soft parton emission. The jet activity, however, does not strongly constrain the global production of hadrons known as the underlying event (UE). Therefore, α_s^{ISR} as determined from $t\bar{t}$ jet kinematics can be used as a fixed parameter to tune the UE. See ref. [10] for the details of the CUETP8M2T4 tune derived fixing α_s^{ISR} to 0.1108. The performance of the CUETP8M2T4 PYTHIA8 tune is evaluated in different configurations. It is found that both POWHEG V2+PYTHIA8 and MG5_aMC@NLO + PYTHIA8 with FxFx merging [19] describe the top quark data well (except for top quark p_T independent of the tune), while MG5_aMC@NLO +PYTHIA8 with MLM matching [20] and the inclusive aMC@NLO +PYTHIA8 does not describe the data in general (e.g. see Fig. 1b). It is also observed that the global event variables such as H_T or S_T do not get modified significantly with the change of α_s^{ISR} (except for MG5_aMC@NLO + PYTHIA8 [MLM] and aMC@NLO +PYTHIA8 independent of the tune for some variables such as the jet multiplicity) [10]. The comparisons of predictions of POWHEG V2+PYTHIA8 with the CUETP8M2T4 tune for six different differential cross-sections to the ones measured with 35.9 fb^{-1} data at $\sqrt{s} = 13$ TeV and yield an overall p-value from < 0.01 when theory uncertainties are ignored to 0.91 when theory uncertainties are included [21].



(a) (b)

Figure 1: The charged particle multiplicity distributions for the away, transverse and toward regions as well as for the overall sample (a). Distributions are obtained with the increased scale, i.e. $(2Q)^2$ that matches the data better. The points correspond to the data at the detector level and the lines represent the POWHEG V2+ PYTHIA8 predictions with the CUETP8M1 tune. Each distribution is normalised to one. CMS data at 13 TeV on the normalised cross section in the lepton+jets channel, as a function of the number of additional jets (b). The data are compared with the predictions of POWHEG V2, MG5_aMC@NLO , either with MLM matching or FxFx merging, and with aMC@NLO interfaced with PYTHIA8 with CUETP8M1 and CUETP8M2T4 tunes. Also shown is the ratio of the theory and the data (theory/data), where the yellow band indicates the total experimental uncertainty of the data.

References

[1] CMS Collaboration, *The CMS Experiment at the CERN LHC*, *JINST* **3** (2008) S08004.

[2] A. Buckley et al., *RIVET hepforge homepage*, `https://rivet.hepforge.org/`.

[3] CMS Collaboration, *Object definitions for top quark analyses at the particle level*, CERN-CMS-NOTE-2017-004 2017.

[4] CMS Collaboration, *Measurement of the top quark mass using proton-proton data at $\sqrt{s}$ = 7 and 8 TeV*, *PRD* **93** (2016) 072004 [`arXiv:1509.04044`]

[5] S. Argyropoulos, and S. Torbjorn, *Effects of color reconnection on $t\bar{t}$ final states at the LHC*, *JHEP* **11** (2014) 043 [`arXiv:1407.6653`]

[6] M. H. Seymour, and A. Siodmok, *Constraining MPI models using σ_{eff} and recent Tevatron and LHC Underlying Event data*, *JHEP* **10** (2013) 113 [`arXiv:1307.5015`]

[7] CMS Collaboration, *Underlying event measurement with $t\bar{t}$ +X events with p-p collision data at $\sqrt{s}$=13 TeV*, *CMS-PAS-TOP-15-017* [`https://cds.cern.ch/record/2114809`] 2015.

[8] CMS Collaboration, *Event generator tunes obtained from underlying event and multiparton scattering measurements*, *EPJ-C* **76** (2016) 155 [`arXiv:1512.00815`]

[9] CMS Collaboration, *Measurement of $t\bar{t}$ production with additional jet activity, including* b *quark jets, in the dilepton decay channel using pp collisions at $\sqrt{s}$ = 8 TeV*, *EPJ-C* **76** (2016) 379 [`arXiv:11510.03072`]

[10] CMS Collaboration, *Investigations of the impact of the parton shower tuning in* PYTHIA8 *in the modelling of $t\bar{t}$ at $\sqrt{s} = 8$ and 13 TeV*, *CMS-PAS-TOP-16-021* [`https://cds.cern.ch/record/2235192`] 2016.

[11] CMS Collaboration, *Measurement of differential cross sections for top quark pair production using the lepton+jets final state in proton-proton collisions at 13 TeV*, *Phys. Rev. D* **95** (2017) 092001 [`arXiv:1610.04191`]

[12] P. Skands, S. Carrazza, and J. Rojo, *Tuning* PYTHIA8.1*: the Monash 2013 Tune*, *EPJ-C* **74** (2014) 3024 [`arXiv:1404.5630`]

[13] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040 [`arXiv:ep-ph/0409146`]

[14] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the* POWHEG V2 *method*, *JHEP* **11** (2007) 070 [`arXiv:hep-ph/0709.2092`]

[15] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the* POWHEG V2 *BOX*, *JHEP* **06** (2010) 043 [`arXiv:1002.2581`]

[16] A. Buckley, et al., *Systematic event generator tuning for the LHC*, *EPJ-C* **65** (2010) 331 [`arXiv:0907.2973`]

[17] C. Patrignani et al. (Particle Data Group), *Review of Particle Physics*, *Chin. Phys. C* **40** (2016) 100001 [`http://pdg.lbl.gov/`]

[18] CMS Collaboration, *Measurement of $t\bar{t}$ production with additional jet activity, including* b *quark jets, in the dilepton decay channel using pp collisions at $\sqrt{s}$ = 8 TeV*, *EPJ-C* **76** (2016) 379 [`supplementary material`]

[19] R. Frederix, and S. Frixione, *Merging meets matching in MC@NLO*, *JHEP* **12** (2012) 061 [`arXiv:1209.6215`]

[20] J. Alwall, et al., *Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions*, *EPJ-C* **53** (2008) 473 [`arXiv:0706.2569 [hep-ph]`]

[21] CMS Collaboration, *Measurement of the differential cross sections of top quark pair production as a function of kinematic event variables in pp collisions at $\sqrt{s} = 13$ TeV*, *CMS-PAS-TOP-16-14* [`https://cds.cern.ch/record/2273459`] 2017.