Event Modelling in CMS

Paolo Gunnellini* **on behalf of the CMS Collaboration.**
Deutsches Elektronen-Synchrotron (DESY)
E-mail: paolo.gunnellini@desy.de

Latest tests of double parton scattering, underlying event tunes, minimum bias, and diffraction made by comparing the state-of-the-art theoretical predictions interfaced with up-to-date parton shower codes to the CMS Run I and Run II data are presented. Studies to derive and to test a new CMS PYTHIA 8 event tune obtained through jet kinematics in top quark pair events and global event variables are described.

XXV International Workshop on Deep-Inelastic Scattering and Related Subjects
3-7 April 2017
University of Birmingham, UK

---

*Speaker.

1. Introduction

Hadron production in high-energy proton-proton (pp) collisions originates from multiple scatterings of the partonic constituents of the protons at central rapidities, and from "spectator" partons emitted in the very forward direction. Such hadron-hadron collisions, in which a "hard" 2-to-2 parton scattering has occurred, are modelled by standard QCD Monte Carlo (MC) event generators by means of several components. The "hard scattering" component of the event consists of particles that result from the hadronisation of the two outgoing partons (i.e. the two primary outgoing "jets") and the particles that arise from initial- and final-state radiation (ISR and FSR, respectively). The Underlying Event (UE) consists of the beam-beam remnants (BBR) and particles that arise from multiple parton interactions (MPI). The BBR are what is left over after a parton is knocked out of each of the initial two beam hadrons. MPI are additional "soft" or "semi-hard" 2-to-2 parton-parton scatterings that occur within the same hadron-hadron collision. An accurate understanding of the UE is required for precise measurements of standard model processes at high energies and searches for new physics. The topological structure of pp interactions with a hard scattering can be used to define experimental observables sensitive to the UE. In pp collisions at the LHC, various experiments have carried out UE measurements at $\sqrt{s} = 7$ and 13 TeV using events containing a leading (highest p_T) charged-particle jet [1, 2] or a leading charged particle [3, 4, 5, 6], and in DY [7].

The goal of a MC event generator, such as, e.g., PYTHIA 8 [8], is to simulate everything that happens in a hadron-hadron collision. This includes the production of a hard scattering, the implementation of the parton shower (PS), i.e. ISR and FSR, related to the evolution of the partons and the UE. In general, MC generators differ between each other in the accuracy of the matrix element (ME) calculation used for the hard scattering process, PS, and in the UE description, i.e. MPI simulation. While in leading-order (LO) generators only $2 \rightarrow 2$ processes are calculated and the additional emitted partons are implemented through the PS, there are several cases where higher-order ME are implemented at the Born level. By including additional emissions in the ME, the exact kinematics and all the interference and helicity parton structures are taken into account. On the other hand, to calculate higher-order diagrams and loop corrections is a very time consuming process and a compromise needs to be made. The different components, simulated by MC event generators, are generally governed by many parameters, which can not be a priori defined by first principles. Tuning these parameters in a MC event generator is a very important matter since, in the available models, it is unavoidable that some of the parameters used in the simulation must be determined by looking for the best agreement between the predictions and data from as many processes as possible.

Generally, a factorized approach is followed for the tuning, i.e. one identifies a set of parameters of interest, together with a collection of measurements sensitive to the physics regulated by those parameters, and "freezes" the other parameters, which remain fixed in the fit. For instance, if a tuning effort is setup for extracting the MPI parameters, the parameters related to ISR and FSR [1] are not varied.

The purpose of tuning the MC simulation is, in general, two-fold:

---

[1]The FSR parameters are generally extracted from LEP data on electron-positron collisions, where ISR and MPI parameters do not play any role.

- describe well the measured data, in order to achieve, e.g., a reliable pile-up simulation, a correct evaluation of detector effects treated in the unfolding procedure, and an appropriate estimation of background contributions. For this goal, models are "not allowed" to fail, namely one aims for a best-possible agreement among measurements and predictions;

- provide good physics predictions, based on established theories and/or trustworthy phenomenological models, using meaningful values of free parameters. Following this approach, the models are "allowed" to fail, and their failure actually brings an improvement of the understanding of the physics, implemented in the simulation.

In a previous study [9], several PYTHIA 8 UE tunes, including the CUETP8M1 tune, were constructed, before 13 TeV data were available. In particular, the PYTHIA 8 CUETP8M1 tune, performed within the CMS Collaboration [10], started with the Monash tune [11] was obtained by fitting CDF UE data at 900 GeV and 1.96 TeV together with CMS UE data at 7 TeV. A tension was observed when trying to simultaneously reproduce data sensitive to the UE and to the hard spectrum of the MPI, e.g. to double parton scattering (DPS). The retuning effort, described in [12], was motivated by two main reasons. On the one hand, the first data published by CMS and ATLAS collaborations on UE observables revealed a suboptimal description of the available tunes [2], and showed the need for a more accurate tune at 13 TeV. On the other hand, also the measurements of the additional jet multiplicity in $t\bar{t}$ events using 8 and 13 TeV data showed poor agreement when compared to predictions from POWHEG [13]+PYTHIA 8 CUETP8M1 tune [14, 15, 16]. These two main reasons drove the extraction of a new tune and motivated the tuning strategy, which is described in Sect. 2.

2. Tuning strategy

The new event tune is extracted using the PROFESSOR software [17] and RIVET [18], by generating sets of MC predictions in an automatized way. The set of tuned parameters that best fits the input measurements defines the new event tune. The tuning procedure consists of two main steps. In a first step [12], the strong coupling for the ISR, α_s^{ISR}, is tuned by using $t\bar{t}$ jet multiplicity and leading additional jet p_T distributions. The jet activity mainly constrains those parameters that control the probability of parton emission and the interplay between hard and soft parton emission. The jet activity, however, does not strongly constrain the global production of hadrons, i.e. the UE contribution. For the first step, a next-to-leading (NLO) $t\bar{t}$ matrix element produced by the POWHEG event generation, matched to PYTHIA 8 for PS and UE simulation, is considered. In detail, the following parameters are considered for tuning:

- h_{damp}: the model parameter that controls matrix-element/parton-shower matching and effectively regulates the high-p_T radiation by damping real emissions generated by POWHEG. The value used in the default simulation is equal to the top-quark mass $m_t = 172.5$ GeV;

---

[2]Since CMS was particularly interested in a correct energy extrapolation when performing the pre-13-TeV tunes and the data showed suboptimal performance at 13 TeV, one might question the accuracy of the energy extrapolation function implemented in e.g. PYTHIA 8, but this goes a bit beyond the tuning activities.

- α_s^{ISR}: the value of the strong coupling at m_Z used for the initial-state shower. Data from the event shapes, measured at LEP [19], yielded $\alpha_s^{FSR} = 0.1365$ for the final-state shower. The values used in the default simulation for both α_s^{FSR} and α_s^{ISR} are set to 0.1365.

Once the α_s^{ISR} is extracted from the aforementioned measurements, the second step of the tuning procedure extracts the full UE tune accordingly, using independent data sets and fixing the α_s value to the optimized one. In this step, leading-order predictions from standalone PYTHIA 8 are used. The parameters considered for the event tune are related to the amount of multi-parton interactions (MPI), of proton overlap in the hadronic collisions and of colour reconnection. A change of the event parameters is expected, if a different value of ISR α_s is chosen; this is due to the complex interplay between the MPI and PS contributions. The MPI-related parameters, that are let free in in the new tune called CUETP8M2T4 are listed below:

- PT0REF: the parameter that regularises the partonic QCD cross section and sets the lower scale of the MPI contribution;

- EXPPOW: the parameter in the overlap function that describes the convolution of the matter distributions of the two incoming hadrons;

- RANGE, it quantifies the amount of colour reconnection which is performed in the simulation between the hard scattering and the partonic MPI systems.

Other parameters of the simulation that affect, for example, the PS, the fragmentation, and the intrinsic-parton p_T are fixed to the values of the CUETP8M1 tune [9]. The full UE tune is determined by fitting event observables measured at 13 TeV [5], which are the charged-particle multiplicity and p_T sum (Σp_T) in the TRANSMIN and TRANSMAX regions, as defined in [9, 5], as a function of the leading track p_T. The charged-particle multiplicity as a function of η at 13 TeV [20] is also included in the fit. The adopted procedure is similar to the one described in [9]. Since only measurements performed at 13 TeV are considered, the parameter relative to the energy dependence of the partonic cross section cutoff (ECMPOW) is fixed to the value of CUETP8M1.

3. Results

The first step of the tuning procedure performed on the $t\bar{t}$ observables yields a result of:

$$h_{\text{damp}} = 1.581^{+0.658}_{-0.585} \times m_t, \qquad\qquad \alpha_s^{\text{ISR}} = 0.1108^{+0.0145}_{-0.0142} \qquad\qquad (3.1)$$

implemented in an intermediate tune, called CUETP8M1T4. A retuning of α_s^{ISR} alone to the same $t\bar{t}$ data yields $\alpha_s^{\text{ISR}} = 0.115^{+0.021}_{-0.019}$ and the corresponding tune is called CUETP8M1T1 [21].

Predictions obtained with these optimized settings are shown in Fig. 1. Predictions based on the CUETP8M1T4 tune, with a significantly lower value of α_s^{ISR}, cure the overshoot observed for the predictions from CUETP8M1 at high jet multiplicities.

The resulting parameters, obtained in the second step of the tuning procedure for the tune CUETP8M2T4, are listed in Table 1, compared to the old CUETP8M1 tune [9]. Note that the main changes between the parameters of the two tunes are α_s, PDF set, MPI infrared regularization scale
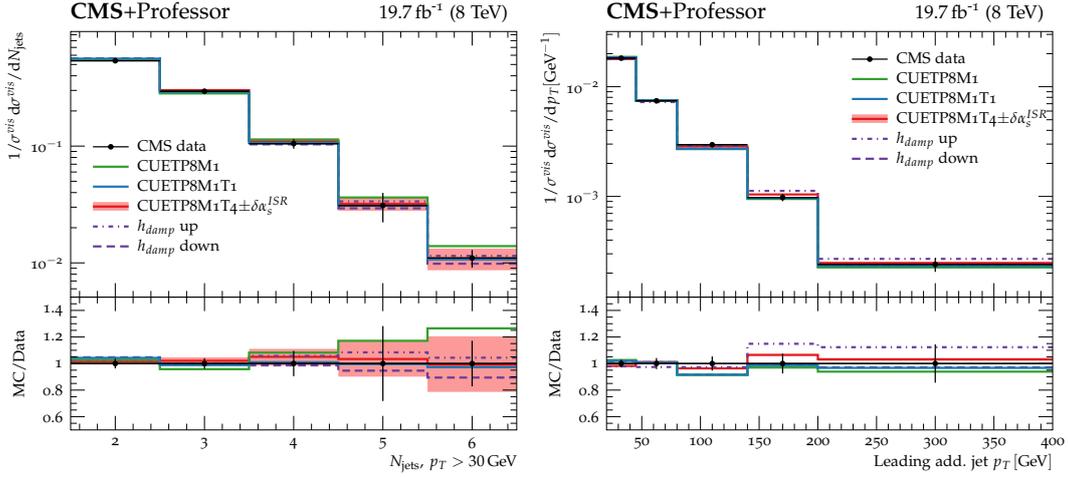
Figure 1: (Left) CMS data [21] at $\sqrt{(s)} = 8$ TeV on (left) additional jet multiplicity (with jets in $|\eta| < 2.5$ and $p_T > 30$ GeV) and (right) transverse momentum of the leading additional jet in $t\bar{t}$ fully-leptonic events. Predictions of POWHEG+PYTHIA 8 with different UE tunes are compared to the data. The ratios of MC predictions to data are given below each panel.

(p_{T0}^{ref}) and the amount of color reconnection. The large change in the amount of color reconnection is a result of the different PDF set and the α_s value used in the different tunes.

Table 1: The parameters of the old tune (CUETP8M1) and the new tune (CUETP8M2T4). The parameter (in italic) relative to the energy dependence of the partonic cross section cutoff (ECMPOW) is fixed to the value of CUETP8M1 in the fit of the new tune. In the new tune, the ISR α_s is fixed to the value extracted from the $t\bar{t}$ events.

	CUETP8M1	CUETP8M2T4
MultipartonInteractions:ecmPow	0.2521	0.2521
SpaceShower:alphaSvalue	0.1365	0.1108
PDF pSet LHAPDF6	NNPDF23_lo_qed_as_0130	NNPDF30_lo_as_0130
MultipartonInteractions:PT0REF	2.40	2.20
MultipartonInteractions:EXPPOW	1.6	1.6
ColourReconnection:RANGE	1.8	6.6

Figure 2 shows the charged-particle multiplicity as a function of the p_T of the leading charged particle, measured in the TRANSMIN region, and the charged-particle multiplicity as a function of pseudorapidity η. The measurements are compared to predictions of the old (CUETP8M1) and the new (CUETP8M2T4) tunes. Predictions of the new tune reproduce very well the data points and perform much better than the predictions of the old one, which tend to underestimate the considered observables by about 10-15%. The discrepancy in the rising part of the spectrum of the charged-particle multiplicity as a function of the p_T of the leading charged particle, in the region up to 3 GeV, observed for both predictions, might be due to diffractive contributions, which were not considered in the tuning.
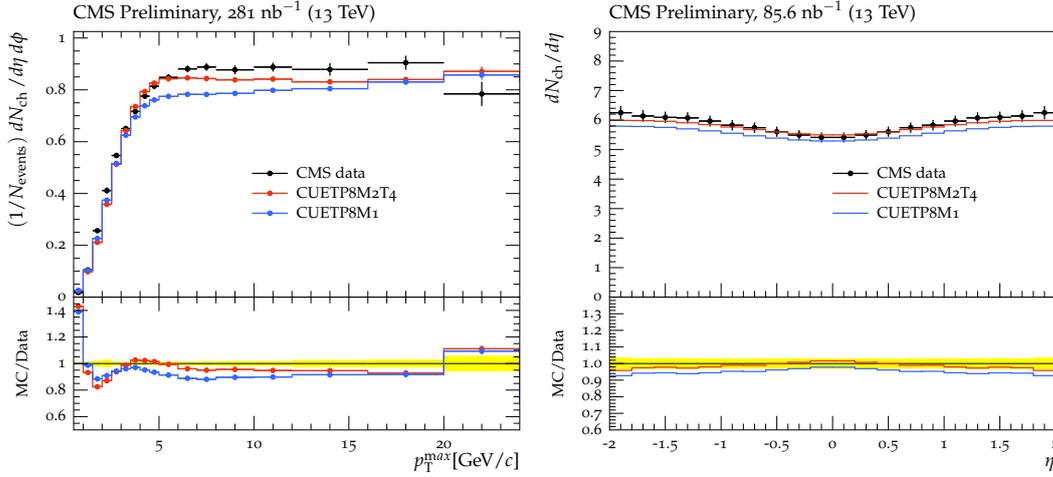
Figure 2: (Left) CMS data [5] at $\sqrt{(s)}$ = 13 TeV on particle densities for charged particles with $p_T >$ 0.5 GeV and $|\eta| < 0.8$ in the TransMIN region as defined by the leading charged particle, as a function of the transverse momentum of the leading charged-particle $p_T^{\max}$. The data are compared to predictions of PYTHIA 8 Tune CUETP8M1 and CUETP8M2T4. The ratios of MC events to data are given below each panel. (Right) CMS data [20] at $\sqrt{(s)}$ = 13 TeV for the charged-particle pseudorapidity distribution, $dN_{ch}/d\eta$, in inelastic proton-proton collisions. The data are compared to predictions of PYTHIA 8 Tune CUETP8M1 and CUETP8M2T4. The ratios of MC events to data are given below each panel.

4. Summary and conclusions

This study has produced an improved PYTHIA 8 tune at 13 TeV, CUETP8M2T4, with a lower parton shower α_s, with respect to the one used in the CUETP8M1 [9, 11] tune along with a new h_{damp} value in the POWHEG matrix element simulation, that improves the description of $t\bar{t}$ kinematics and describes the UE sensitive observables in a hard scattering process, as well at the global event observables in MB collisions. The CUETP8M2T4 tune uses the NNPDF3.0LO [22] PDF set, which improves the constraint on the low-x gluon distribution, resulting in a smaller low-x gluon contribution than the NNPDF2.3LO PDF set.

References

[1] S. Chatrchyan *et al.* [CMS Collaboration], JHEP **09** (2011) 109 doi:10.1007/JHEP09(2011)109 [arXiv:1107.0330 [hep-ex]].

[2] G. Aad *et al.* [ATLAS Collaboration], Eur. Phys. J. C **74** (2014) no.8, 2965 doi:10.1140/epjc/s10052-014-2965-5 [arXiv:1406.0392 [hep-ex]].

[3] V. Khachatryan *et al.* [CMS Collaboration], Eur. Phys. J. C **70** (2010) 555 doi:10.1140/epjc/s10052-010-1453-9 [arXiv:1006.2083 [hep-ex]].

[4] G. Aad *et al.* [ATLAS Collaboration], Phys. Rev. D **83** (2011) 112001 doi:10.1103/PhysRevD.83.112001 [arXiv:1012.0791 [hep-ex]].

[5] CMS Collaboration, CMS-PAS-FSQ-15-007.

[6] CMS Collaboration, CMS-PAS-FSQ-12-020.

[7] S. Chatrchyan *et al.* [CMS Collaboration], Eur. Phys. J. C **72** (2012) 2080 doi:10.1140/epjc/s10052-012-2080-4 [arXiv:1204.1411 [hep-ex]].

[8] T. SjÃűstrand *et al.*, Comput. Phys. Commun. **191** (2015) 159 doi:10.1016/j.cpc.2015.01.024 [arXiv:1410.3012 [hep-ph]].

[9] V. Khachatryan *et al.* [CMS Collaboration], Eur. Phys. J. C **76** (2016) no.3, 155 doi:10.1140/epjc/s10052-016-3988-x [arXiv:1512.00815 [hep-ex]].

[10] S. Chatrchyan *et al.* [CMS Collaboration], JINST **3** (2008) S08004. doi:10.1088/1748-0221/3/08/S08004

[11] P. Skands, S. Carrazza and J. Rojo, Eur. Phys. J. C **74** (2014) no.8, 3024 doi:10.1140/epjc/s10052-014-3024-y [arXiv:1404.5630 [hep-ph]].

[12] CMS Collaboration, CMS-PAS-TOP-16-021.

[13] S. Frixione, P. Nason and G. Ridolfi, JHEP **09** (2007) 126 doi:10.1088/1126-6708/2007/09/126 [arXiv:0707.3088 [hep-ph]].

[14] CMS Collaboration, CMS-PAS-TOP-16-011.

[15] CMS Collaboration, CMS-PAS-TOP-16-007.

[16] CMS Collaboration, CMS-PAS-TOP-16-008.

[17] A. Buckley, H. Hoeth, H. Lacker, H. Schulz and J. E. von Seggern, Eur. Phys. J. C **65** (2010) 331 doi:10.1140/epjc/s10052-009-1196-7 [arXiv:0907.2973 [hep-ph]].

[18] A. Buckley, J. Butterworth, L. Lonnblad, D. Grellscheid, H. Hoeth, J. Monk, H. Schulz and F. Siegert, Comput. Phys. Commun. **184** (2013) 2803 doi:10.1016/j.cpc.2013.05.021 [arXiv:1003.0694 [hep-ph]].

[19] P. Achard *et al.* [L3 Collaboration], Phys. Rept. **399** (2004) 71 doi:10.1016/j.physrep.2004.07.002 [hep-ex/0406049].

[20] V. Khachatryan *et al.* [CMS Collaboration], Phys. Lett. B **751** (2015) 143 doi:10.1016/j.physletb.2015.10.004 [arXiv:1507.05915 [hep-ex]].

[21] V. Khachatryan *et al.* [CMS Collaboration], Eur. Phys. J. C **76** (2016) no.7, 379 doi:10.1140/epjc/s10052-016-4105-x [arXiv:1510.03072 [hep-ex]].

[22] R. D. Ball *et al.* [NNPDF Collaboration], JHEP **04** (2015) 040 doi:10.1007/JHEP04(2015)040 [arXiv:1410.8849 [hep-ph]].