

Monte Carlo modeling and tuning studies at CMS

Roberto Covarelli^{*†}

University/INFN of Torino, Via Pietro Giuria 1, I-10125 Turin (Italy)

E-mail: roberto.covarelli@cern.ch

The Monte Carlo modeling adopted at the CMS experiment is reviewed. Choices of matrix-element generators are explained and new CMS PYTHIA8 underlying-event tunes are presented. The new tunes are obtained using minimum bias and underlying event observables exploiting Monte Carlo configurations with consistent parton distribution functions and strong coupling constant values in the matrix element and the parton shower, at leading order (LO), next-to-leading order (NLO) and next-to-next-to-leading order (NNLO). Validation and performance studies are presented by comparing the predictions of the new tunes to a wide range of different CMS measurements at 7, 8 and 13 TeV.

*European Physical Society Conference on High Energy Physics - EPS-HEP2019 -
10-17 July, 2019
Ghent, Belgium*

^{*}Speaker.

[†]On behalf of the CMS Collaboration

1. Introduction

Quantum Chromo-Dynamics (QCD) description at the CMS experiment [1] is achieved through the use of well-established Monte-Carlo (MC) tools. In most physics analysis use-cases, i.e. in all cases when QCD precision is required, full-event generators are often replaced by *factorized* MC generation. This is usually comprised of a matrix-element event generation, i.e. up to parton-level, a parton-showering tool (mostly PYTHIA8 [2]) and in some cases specific applications for particle decays, like EvtGen [3] or Tauola [4]. The use of factorized generation requires dealing with physics effects (e.g. “matching and merging”) and more complex interfacing with the CMS software framework.

A very important ingredient of the MC generation is the underlying event (UE), that is a generic name for soft QCD activity occurring in addition to the hard process. It consists of beam-beam remnants and multi-parton interactions (MPI) plus some contribution from the hard process. A good understanding of the UE requires modeling of MPI, hadronisation, initial- and final-state QCD radiation (ISR, FSR) and to tune adjustable parameters in MC tools.

2. Matrix-element generators

Event generation at the next-to-leading order (NLO) in QCD is the most common in CMS MC samples. The tools adopted most frequently are POWHEG [5] and MadGraph5_aMC@NLO [6], while recently SHERPA [7] is becoming more widely used. Specific cases where leading order (LO) is still used, instead, include final states with large particle multiplicities (e.g. vector-boson + 4 jets), new physics models where NLO calculations are not available or implemented in generators, or complicated final-state kinematics, e.g. Higgs-boson couplings with anomalous tensor structure, etc.

Next-to-next-to-leading order (NNLO) calculations and predictions are less common. Several NNLO calculators, like FEWZ [8] or MATRIX [9] are used for comparison with unfolded data. Event generation is at the moment possible only for specific processes, such as the Standard Model Higgs-boson production $gg \rightarrow H$ and are mostly obtained through the NNLOPS method [10] as implemented in POWHEG. A comparison of the $gg \rightarrow H$ transverse momentum spectrum as measured in the CMS $H \rightarrow 4\ell$ analysis with the NNLOPS prediction is shown in Fig. 1.

Matrix-element uncertainties via reweighting of QCD scale and PDF-set variations is now implemented in most commonly-used generators. Significant computing improvements on both MC tools and related CMS implementation allows storing of large number of weights to achieve an accurate estimation of theory uncertainties.

3. Parton showers and past approaches to the UE

PYTHIA8 is used as the main parton-shower tool in CMS. Up to 2016 CMS results, non-uniform recipes were used for parton-shower uncertainties, including difference in relevant variables between PYTHIA8 and HERWIG++ and up- and down-variations of the renormalization scale consistently between matrix-element generator and parton-shower tool. More recently, parton shower weights as determined in [12] have been implemented in MC samples and alternative methods, such as dipole recoil, have been introduced to describe specific processes.

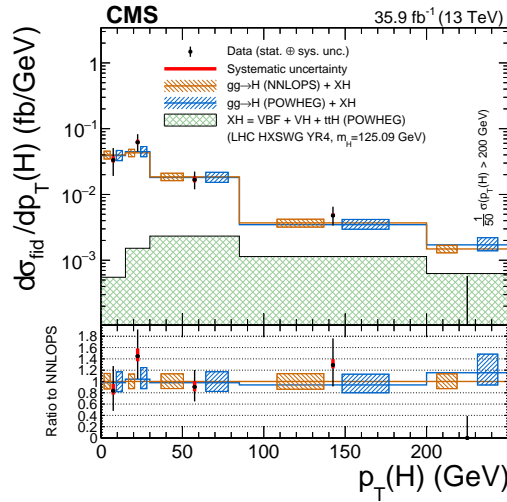


Figure 1: Comparison of the $gg \rightarrow H$ transverse momentum spectrum as measured in the CMS $H \rightarrow 4\ell$ analysis, with different generators at the NLO and NNLO accuracy in Higgs boson production [11].

UE tuning is an important ingredient in the PYTHIA8 generation parameters. Up to 2016, the CMS Collaboration used the CUETP8M1 tune [13], which is derived from CMS minimum-bias data up to 7 TeV and is based on LO parton distribution functions (PDFs). However, with the analysis of 13 TeV data, a not so good agreement was found between minimum-bias data and basic UE variables, as well as in jet multiplicity of $t\bar{t}$ events [14].

4. New approaches to the UE: the CP tunes

CMS PYTHIA8 tunes were traditionally based on LO PDFs, matching the parton-shower accuracy. As such, a value of $\alpha_s(M_Z) = 0.130$ was used for ISR, FSR, and MPI consistently. Following the evidence that UE measurements in $t\bar{t}$ events [14] prefer lower values of $\alpha_s(M_Z)$, a new approach has been devised.

In this new approach, PDF and $\alpha_s(M_Z)$ orders are used consistently in matrix-elements generators, parton-shower, and MPI. The CMS Collaboration uses now NNLO PDFs as default for matrix-elements generators. Following the recommendation of using the same PDF set and $\alpha_s(M_Z)$ value in both components of matched configurations, therefore, this PDF order could be used to determine tunes, with $\alpha_s(M_Z)$ at (N)NLO equaling 0.118.

In order to improve data/modeling agreement at 13 TeV, and compare to past results, a set of new 13 TeV PYTHIA8.226 UE tunes is obtained with different choices of $\alpha_s(M_Z)$ values used in the modeling of the ISR, FSR, hard scattering, and MPI, as well as the order of its evolution as a function of the squared momentum transfer q^2 . We distinguish the new tunes according to the order of the NNPDF3.1 PDF set [16] used: LO, NLO, or NNLO. The tunes are labeled as CP_i , where $i = 1, 2, \dots, 5$: CP1 and CP2 are based on LO, CP3 on NLO, and CP4 and CP5 on the NNLO NNPDF3.1 set.

Input variables for the PYTHIA8 parameter fits include charged-particle and Σp_T densities, as measured in the “transMin” and “transMax” regions, which are available either as a function of

the leading-track or leading-jet p_T [17], as well as the charged-particle multiplicity as a function of pseudorapidity η , measured by CMS with no magnetic field at $\sqrt{s} = 13$ TeV [18]. The charged-particle and Σp_T densities as a function of the leading-track p_T are also available from CMS at $\sqrt{s} = 7$ TeV [19] and from CDF at $\sqrt{s} = 1.96$ TeV [20] (in $p\bar{p}$ collisions) and are included in the fit.

The results of the comparison of the fit results for the CP3, CP4, and CP5 tunes with the 13 TeV CMS minimum-bias data is shown in Fig. 2. NNLO-based tunes have similar or better performances than LO ones, with some residual disagreement at low p_T in the transMin region, and the CP5 tune is the most stable as a function of \sqrt{s} .

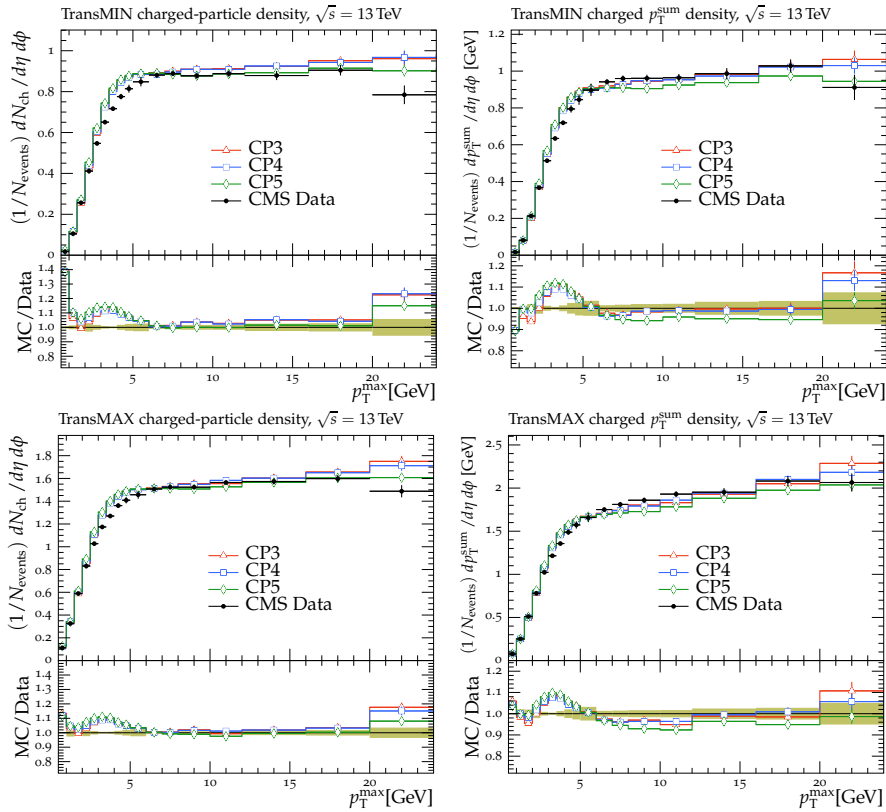


Figure 2: The transMin (upper left) charged-particle and (upper right) Σp_T densities and the transMax (lower left) charged-particle and (lower right) Σp_T densities, as a function of the transverse momentum of the leading track from [17]. The data are compared with the CMS PYTHIA 8 (N)NLO-PDF tunes CP3. The PYTHIA8/data ratios are also shown, where the shaded band indicates the total experimental uncertainty in the data [15].

5. CP tune validation

CP tunes have been validated on several benchmark cross-section measurements at CMS.

$t\bar{t}$ distributions from the CMS analysis at 13 TeV [21] are compared to NLO generators: POWHEG inclusive and MadGraph5_aMC@NLO with FxFx merging and ≤ 2 partons. PDFs used are NNPDF3.1 NNLO, with $\alpha_s(M_Z) = 0.118$ and same choice of dynamical QCD scales in

both cases. A lepton+jets selection (with jet b-tagging) is used and unfolded data are compared to tune predictions. In the POWHEG description NNLO-PDF-based tunes show a large improvement in the number of additional jets (Fig. 3 left).

Z-boson distributions from the CMS analysis at 13 TeV [22] are compared to LO and NLO MadGraph5_aMC@NLO predictions: at the LO with MLM merging, ≤ 4 partons, and with FxFx merging, ≤ 2 partons. PDFs used are NNPDF3.1 NNLO, with $\alpha_s(M_Z)=0.118$ and same choice of dynamical QCD scales in both cases. Z-boson selection is based on identification of isolated pairs of electrons and muons with an invariant mass within 20 GeV of the Z mass. The NLO generator with the new PYTHIA8 tunes is found to describe well the N_{jets} and $p_T(Z)$ above 5 GeV (Fig. 3 right).

CMS 4-jet and 2b-2jet events, which are signature of double-parton scattering are used to test CP tunes through the correlation variable ΔS , the angle between the vectors of the hard di-jet (2-b) and soft di-jet (2-jet) systems. For the NNLO-PDF tunes a good agreement with the fitted DPS σ_{eff} is found but the tunes, especially those featuring rapidity ordering, do not describe the data well.

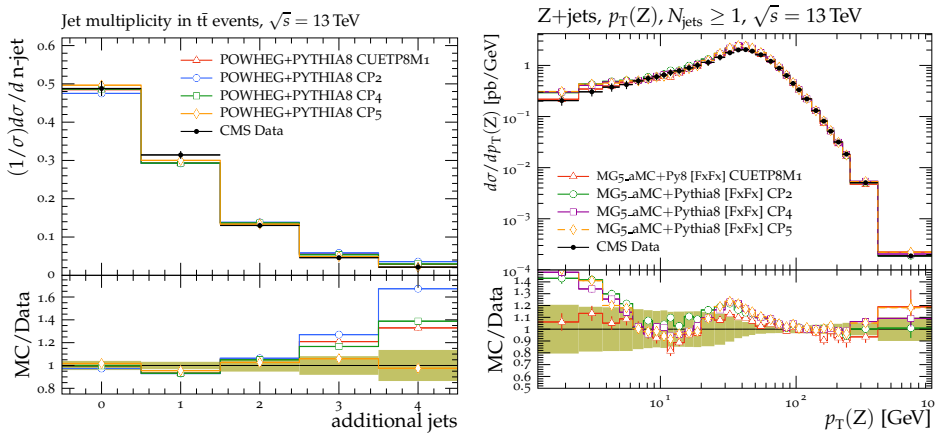


Figure 3: Left: The normalized $t\bar{t}$ cross section in the lepton+jets channel, in bins of number of additional jets. The data [21] are compared with the predictions of POWHEG, the parton-shower simulation is done with the PYTHIA 8 tunes CUETP8M1, CP2, CP4, or CP5. Right: Comparison of the measurement of the $p_T(Z)$ distribution [22] with these predicted by MadGraph5_aMC@NLO + PYTHIA8 with FxFx merging for the different tunes [15].

6. Conclusions

Choices of matrix-element generators in the CMS experiment have been reviewed and new PYTHIA8 UE tunes are presented. The new tunes are obtained using minimum bias observables using consistent parton distribution functions and strong coupling constant values in the matrix element and the parton shower, at leading order (LO), next-to-leading order (NLO) and next-to-next-to-leading order (NNLO). For the first time, predictions based on higher-order PDF sets are shown to give a reliable description of UE measurements, with a similar level of agreement as predictions from tunes using LO PDF sets. Predictions from PYTHIA 8 obtained with tunes based on NLO or NNLO PDFs are shown to reliably describe data in a wide range of different measurements.

References

- [1] The CMS Collaboration, *JINST* 3 (2008) S08004.
- [2] T. Sjöstrand *et al.*, *Comput. Phys. Commun.* 191 (2015), 159.
- [3] D. J. Lange, *Nucl. Instrum. Meth.* A462 (2001), 152.
- [4] N. Davidson *et al.*, *Comput. Phys. Commun.* 183 (2012) 821.
- [5] P. Nason, *JHEP* 0411 (2004) 040; S. Frixione, P. Nason and C. Oleari, *JHEP* 0711 (2007) 070; S. Alioli, P. Nason, C. Oleari and E. Re, *JHEP* 1006 (2010) 043.
- [6] J. Alwall *et al.*, *JHEP* 1407 (2014) 079.
- [7] E. Bothmann *et al.*, *SciPost Phys.* 7 (2019), 034.
- [8] R. Gavin *et al.*, *Comput. Phys. Commun.* 182 (2011), 2388.
- [9] M. Grazzini, S. Kallweit and M. Wiesemann, *Eur. Phys. J. C* 78 (2018) 537.
- [10] K. Hamilton *et al.*, *JHEP* 1310 (2013) 222.
- [11] The CMS Collaboration, *JHEP* 1711 (2017) 047.
- [12] S. Mrenna, *Phys. Rev. D* 94 (2016) 074005.
- [13] The CMS Collaboration, *Eur. Phys. J. C* 76 (2016) 52.
- [14] The CMS Collaboration, *Eur. Phys. J. C* 79 (2019) 123.
- [15] The CMS Collaboration, [arXiv:1903.12179](https://arxiv.org/abs/1903.12179), submitted to *Eur. Phys. J. C*.
- [16] R. D. Ball *et al.*, *Eur. Phys. J. C* 77 (2017) 603.
- [17] The CMS Collaboration, CMS Physics Analysis Summary CMS-PAS-FSQ-15-007, <http://cds.cern.ch/record/2104473>.
- [18] The CMS Collaboration, *Phys. Lett. B* 751 (2015) 143.
- [19] The CMS Collaboration, CMS Physics Analysis Summary CMS-PAS-FSQ-12-020, <http://cds.cern.ch/record/1478982>.
- [20] The CDF Collaboration, *Phys. Rev. D* 92 (2015) 092009.
- [21] The CMS Collaboration, *Phys. Rev. D* 95 (2017) 092001.
- [22] The CMS Collaboration, *Eur. Phys. J. C* 78 (2018) 965.