Monte Carlo Modelling and Tuning in CMS

Vieri Candelise
University of Trieste and INFN

11th International Workshop on Multiple Partonic Interactions
Prague, Czech Republic

19/11/2019
Outline

- Monte Carlo Modelling and Tuning in CMS: overview
- QCD parameters and Tuning Procedure
- Extraction of 5 new Pythia8 Tunes
- Validation and Comparisons on Several Physics Processes
Modelling and Tuning

Modelling Monte Carlo Generators at the LHC

useless to say: we (often) rely on Monte Carlo doing LHC physics

- Heavily dominated QCD-radiated / jets environment
- irreducible backgrounds - MC for searches
- achieve precision (NNLO) for SM measurements
- need higher accuracy precision on phenom. parameters

In CMS we use LO and NLO multi-leg generators matched with parton shower

- MadGraph5_aMC@NLO+Pythia8 MLM-merging LO up to 4 partons W+jets, Z+jets...
- MadGraph5_aMC@NLO+Pythia8 FxFx-merging NLO up to 2 partons
- Powheg MiNLO +Pythia8 NNLO up to 2 jets top, V+2J...
- Sherpa, Herwig7... NLO
Modelling and Tuning

Modelling Monte Carlo Generators at the LHC

strong coupling and parton density functions appear in many physics stages of pp collisions and can be evaluated at different orders in QCD

- PDFs and strong coupling orders ME, PS and MPI
- this poses problems of $\alpha_S$ orders ME/PS/MPI
- Order of strong coupling at FSR/ISR
- rapidity ordering effects

It is important to have tunes at different $\alpha_S$ and PDF orders (they can be very different!) and compare to data to understand their effect

why tuning?
The good old underlying event sketch

LHC’s  

\[ pp \rightarrow \]

Underlying Event at the LHC consists of a lot of processes at different scales

- hadronization modelling
- initial/final state radiation (PS)
- multiple parton interactions (UE)
- beam remnants (primordial \(k_T\))

- Double Parton Scattering
- Diffraction
- Semi-hard MPI processes

A rich physics relying on phenomenological parameters. We need data to understand QCD.
Several sensitive QCD parameters

LHC’s

\[ pp \rightarrow \]

MPI

\[ p_T^0 = p_T^{ref} \cdot (E/E_{ref})^\epsilon \]

smooth cut to regularize \( pt \rightarrow 0 \) divergences

initial/final state radiation (PS)
strong coupling, regularization upper scale…

hadronization modelling
length of fragmentation string, strange baryon suppression…

beam remnants (primordial \( kT \))
width of the gaussian of the primordial \( kT \) in the proton…

we use Professor as main tool to choose parameters ranges, make interpolation and minimization to extract the tunes
Tunes Extraction Procedure: tools

PROFESSOR
for tuning MC generators

RIVET
for comparisons

https://professor.hepforge.org/
[v.2.5.2]

https://rivet.hepforge.org/
[v.1.4.0]

common strategy, experiment
independent, flexible and reliable

“social” drawback: we need RIVET routines from analyses to make comparisons!
Tunes Extraction Procedure

We use CDF data at $\sqrt{s}=1.96$ TeV and CMS data with $\sqrt{s}=7,13$ TeV:

- charged particle and $p_T^{\text{sum}}$ densities in $\text{transMIN}$, $\text{transMax}$ vs $p_T^{\text{max}}$
- charged-particle multiplicity vs $\eta$

Generate predictions points in a grid parameter space

Tune data by minimizing

$$\chi^2(p) = \sum w_o \sum \frac{f^b(p) - R(b)^2}{\Delta^2_b}$$

**MPI modelling:** The overlap between two protons modelled by a double-gaussian
The 7 TeV Tunes Era

tested on 13 TeV data

7 TeV CMS tunes were based on Monash tune

CMS official MC production used CUETP8M1 as default tune

none of the tunes describes both low $p_T$ and high $p_T$ particle multiplicity from UE!

we need to explore deeper the QCD and evaluate the tune dependence on PDF/ $\alpha_S$ at higher orders…

TransMAX charged-particle density $\sqrt{s} = 13$ TeV

MC/Data

$$(1/N_{\text{events}}) dN_{\text{ch}} / d\eta d\phi$$

$P_T^{\text{max}}$ [GeV]
The CMS “Tune Paper”

Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements

The CMS Collaboration*

Abstract

New sets of CMS underlying-event parameters ("tunes") are presented for the PYTHIA8 event generator. These tunes use the NNPDF3.1 parton distribution functions (PDFs) at leading (LO), next-to-leading (NLO), or next-to-next-to-leading (NNLO) orders in perturbative quantum chromodynamics, and the strong coupling evolution at LO or NLO. Measurements of charged-particle multiplicity and transverse momentum densities at various hadron collision energies are fit simultaneously to determine the parameters of the tunes. Comparisons of the predictions of the new tunes are provided for observables sensitive to the global underlying event, soft multiparton interactions, and double-parton scattering contributions. In addition, comparisons are made for observables measured in various specific processes, such as multijet, Drell-Yan, and top quark-antiquark pair production. The simulation of the underlying event provided by the new tunes is interfaced to a higher-order matrix-element calculation. For the first time, predictions from PYTHIA8 obtained with tunes based on NLO or NNLO PDFs are shown to reliably describe minimum-bias and underlying-event data with a similar level of agreement to predictions from tunes using LO PDF sets.


★ a huge and complex collective effort!

★ under EPJC review

★ first attempt to deliver a P8 tune “tested” over several physics processes

★ Output of the work is a new set of 5 tunes

one tune will be chosen for the official CMS MC production based on performances
The CMS Pythia8 Tunes Family

<table>
<thead>
<tr>
<th>PYTHIA parameter</th>
<th>CP1</th>
<th>CP2</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDF Set</td>
<td>NNPDF3.1 LO</td>
<td>NNPDF3.1 LO</td>
</tr>
<tr>
<td>$\alpha_s(M_Z)$</td>
<td>0.130</td>
<td>0.130</td>
</tr>
<tr>
<td>SPACESHOWERRAPIDITYORDER</td>
<td>off</td>
<td>off</td>
</tr>
<tr>
<td>MULTIPARTONINTERACTIONS:EcmRef [GeV]</td>
<td>7000</td>
<td>7000</td>
</tr>
<tr>
<td>$\alpha_s^{ISR}$ value/order</td>
<td>0.1365/LO</td>
<td>0.130/LO</td>
</tr>
<tr>
<td>$\alpha_s^{ISR}$ value/order</td>
<td>0.1365/LO</td>
<td>0.130/LO</td>
</tr>
<tr>
<td>$\alpha_s^{MPI}$ value/order</td>
<td>0.130/LO</td>
<td>0.130/LO</td>
</tr>
<tr>
<td>$\alpha_s^{ME}$ value/order</td>
<td>0.130/LO</td>
<td>0.130/LO</td>
</tr>
<tr>
<td>MULTIPARTONINTERACTIONS:pT0Ref [GeV]</td>
<td>2.4</td>
<td>2.3</td>
</tr>
<tr>
<td>MULTIPARTONINTERACTIONS:eCMpow</td>
<td>0.1543</td>
<td>0.1391</td>
</tr>
<tr>
<td>MULTIPARTONINTERACTIONS:coreRADIUS</td>
<td>0.5436</td>
<td>0.3755</td>
</tr>
<tr>
<td>MULTIPARTONINTERACTIONS:coreFraction</td>
<td>0.6836</td>
<td>0.3269</td>
</tr>
<tr>
<td>COLORRECONNECTION:range</td>
<td>2.633</td>
<td>2.323</td>
</tr>
<tr>
<td>$\chi^2$/dof</td>
<td>0.89</td>
<td>0.538</td>
</tr>
</tbody>
</table>

The Leading Order Tunes CP1 and CP2

CP Tunes = CMS Pythia Tunes
### The CMS Pythia8 Tunes Family

The next-to-leading order tunes CP3, CP4, CP5.

<table>
<thead>
<tr>
<th>PYTHIA parameter</th>
<th>CP3</th>
<th>CP4</th>
<th>CP5</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDF Set</td>
<td>NNPDF3.1 NLO</td>
<td>NNPDF3.1 NNLO</td>
<td>NNPDF3.1 NNLO</td>
</tr>
<tr>
<td>(\alpha_s(M_Z))</td>
<td>0.118</td>
<td>0.118</td>
<td>0.118</td>
</tr>
<tr>
<td>SPACESHOWER:RAPIDITYORDER</td>
<td>off</td>
<td>off</td>
<td>on</td>
</tr>
<tr>
<td>MULTIPARTONINTERACTIONS:ECMRef [GeV]</td>
<td>7000</td>
<td>7000</td>
<td>7000</td>
</tr>
<tr>
<td>(\alpha_S^{ISR}) value/order</td>
<td>0.118/NLO</td>
<td>0.118/NLO</td>
<td>0.118/NLO</td>
</tr>
<tr>
<td>(\alpha_S^{ISR}) value/order</td>
<td>0.118/NLO</td>
<td>0.118/NLO</td>
<td>0.118/NLO</td>
</tr>
<tr>
<td>(\alpha_S^{MPI}) value/order</td>
<td>0.118/NLO</td>
<td>0.118/NLO</td>
<td>0.118/NLO</td>
</tr>
<tr>
<td>(\alpha_S^{ME}) value/order</td>
<td>0.118/NLO</td>
<td>0.118/NLO</td>
<td>0.118/NLO</td>
</tr>
<tr>
<td>MULTIPARTONINTERACTIONS:PT0Ref [GeV]</td>
<td>1.516</td>
<td>1.483</td>
<td>1.41</td>
</tr>
<tr>
<td>MULTIPARTONINTERACTIONS:ecmPow</td>
<td>0.02266</td>
<td>0.02012</td>
<td>0.03344</td>
</tr>
<tr>
<td>MULTIPARTONINTERACTIONS:coreRadius</td>
<td>0.5396</td>
<td>0.5971</td>
<td>0.7634</td>
</tr>
<tr>
<td>MULTIPARTONINTERACTIONS:coreFraction</td>
<td>0.3869</td>
<td>0.3053</td>
<td>0.63</td>
</tr>
<tr>
<td>COLORRECONNECTION:range</td>
<td>4.727</td>
<td>5.613</td>
<td>5.176</td>
</tr>
<tr>
<td>(\chi^2/\text{dof})</td>
<td>0.759</td>
<td>0.803</td>
<td>1.04</td>
</tr>
</tbody>
</table>

**CP Tunes = CMS Pythia Tunes**
Test of the CP tunes

Underlying Events

all results taken from arXiv:1903.12179
leading charged particle jet in the **TransMIN** region

**TransMIN charged-particle density $\sqrt{s} = 13$ TeV**

- CMS preliminary

- CP1
- CP2
- Data

**Soft region:** higher order PDF tunes (CP3, CP4, CP5) better agreement

**Soft region:** up to 15% disagreement with LO tunes agreement ~ok after 10 GeV
leading charged particle jet in the TransMAX region

TransMAX charged $p_T^{\text{sum}}$ density $\sqrt{s} = 13$ TeV

- soft region: higher order PDF tunes (CP3, CP4, CP5) better agreement
- soft region: up to 15% disagreement with LO tunes agreement ~ok after 10 GeV
Test of the CP tunes

*Double Parton Scattering*

all results taken from arXiv:1903.12179
\[
\Delta S = \arccos \frac{\vec{p}_{T,1} \cdot \vec{p}_{T,2}}{|\vec{p}_{T,1}| \cdot |\vec{p}_{T,2}|}
\]

- DPS observable tested over inclusive 4j and 2b+2j final states
- LO tunes better: different pT0ref value to simulate MPI
- DPS is very sensitive to rapidity ordering: need a specific tune
Test of the CP tunes

Top quark

all results taken from arXiv:1903.12179
Normalized $t\bar{t}$ cross section in bins of top quark mass and jet multiplicity

matrix element generation made by madgraph_aMC@NLO[FxFx] and Powheg

CP5 does the best job describing top data (effect of rapidityOrder ON)
CP4 overestimates higher multiplicities (effect of rapidityOrder OFF)
(all tunes describe higher multiplicities using MG5)
Test of the CP tunes

Jets & EW bosons

all results taken from arXiv:1903.12179
CP tunes in high $p_T$ jets & EW bosons final states

**UE in Z+jets**  
matrix element generation made by madgraph_aMC@NLO [FxFx]

Toward charged-particle density $\sqrt{s} = 13$ TeV

<table>
<thead>
<tr>
<th>$p_T(\mu\mu)$ [GeV]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data</td>
</tr>
<tr>
<td>MG5_aMC+PYTHIA8 [FxFx] CUETM1</td>
</tr>
<tr>
<td>MG5_aMC+PYTHIA8 [FxFx] CP2</td>
</tr>
<tr>
<td>MG5_aMC+PYTHIA8 [FxFx] CP4</td>
</tr>
<tr>
<td>MG5_aMC+PYTHIA8 [FxFx] CP5</td>
</tr>
</tbody>
</table>

Toward charged $p_T^{\text{sum}}$ density $\sqrt{s} = 13$ TeV

<table>
<thead>
<tr>
<th>$p_T(\mu\mu)$ [GeV]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data</td>
</tr>
<tr>
<td>MG5_aMC+PYTHIA8 [FxFx] CUETM1</td>
</tr>
<tr>
<td>MG5_aMC+PYTHIA8 [FxFx] CP2</td>
</tr>
<tr>
<td>MG5_aMC+PYTHIA8 [FxFx] CP4</td>
</tr>
<tr>
<td>MG5_aMC+PYTHIA8 [FxFx] CP5</td>
</tr>
</tbody>
</table>

$p_T(\mu)>20/10$ GeV  
$|M_{\mu\mu}-91| < 20$ GeV  
eta < 2

old-Monash based tune CUETPM1 not able to describe the Z
CP5 gives the average best description in all regions  
all tunes describe data
CP tunes in high $p_T$ jets & EW bosons final states

hard QCD: $Z$+jets events with jet $p_T$>30 GeV
matrix element generation made by madgraph_aMC@NLO[FxFx and MLM]

$Z$-J balance: hadronic activity not clustered in jets generate imbalance
(!) Only below 20 GeV we see differences among tunes

$$p_T^{bal} = |p_T(Z) + \sum_{jets} p_T^j|$$
Inclusive Multiplicity in W+jets

\[ \sigma(W \rightarrow \mu\nu + \geq N \text{ jets}) \text{ [pb]} \]

Inclusive Multiplicity in Z+jets

\[ \sigma(Z/\gamma^* \rightarrow ll + \geq N \text{ jets}) \text{ [pb]} \]

Inclusive Jet Multiplicity, \( \sqrt{s} = 13 \text{ TeV} \)

- MC/Data
- \( N_jets \)
- fairly good agreement on all the tunes
- \( N_j > 3 \) only CP2 describes data
- describes data
- CMS preliminary
Conclusions & Perspectives

- The extraction of the 5 CMS Pythia8 Tunes has been described.

- The performances of the tunes have been tested over several physics processes: top, V+Jets, UE, DPS.

- Strong coupling, PDF ordering in QCD play a role as well as the rapidity ordering in many cases.

- The wide physics under test showed interesting inputs on the dependence of the tunes on QCD showering.

- CMS is very active in this field! we are working on specific tunes: Herwig7, DPS, ColorReconnection and more…
STAY TUNED!
backups
CP tunes in high $p_T$ jets & EW bosons final states

Benchmark matrix elements used in V+jets in CMS
MadGraph5_aMC@NLO + Pythia8

$\text{LO-Mode}$ kT–MLM merging scheme
0,1,2,3 partons

$\text{NLO-Mode}$ FxFx merging scheme
up to 2 partons NLO

PDF: NNPDF3.1 @NNLO
$\alpha_s = 0.118$

UE in DY

$|\Delta\phi| \leq 60^\circ$

$60^\circ < |\Delta\phi| \leq 120^\circ$

charged particle density
charged $p_T$ sum

as a function of

Z boson properties
A brief CMS Tunes Chronology

In CMS we developed a set of tunes based on UE in Pythia8.

The (six) tunes differ by PDF order, $\alpha_s$ evolution and are extracted by fitting simultaneously charged-particle multiplicity and transverse momentum densities at various hadron collision energies.

The idea was to obtain a tune for the official MC production that was tested over a variety of different final states, from soft to hard physics.

One of the six tunes was selected as the main "official" one for CMS Run2 MC production.

A huge paper containing the validation of the tune, methodology and cross checks was prepared (arxiv:903.12179 under EPJC review).
CP tunes in high $p_T$ jets & EW bosons final states

Towards region

old-Monash-LO tune

band: total experimental uncertainty of the data
CP tunes in high $p_T$ jets & EW bosons final states

**Transverse region:** CPs are improving the description wrt CUETP8M (but still ~10% off in the central region)
**Z-events**: leptons $p_T > 20$ GeV and $|y| < 2.4$ and the dilepton mass lies in a ±20 GeV window around 91 GeV.

**W-events**: leptons $p_T > 20$ GeV and $|y| < 2.4$ and $M_T > 50$ GeV

**Jets** anti-$kT$04 clustered with $p_T > 30$ GeV and $|y| < 2.4$

**observables** data VS MG5aMC@NLO+P8CPn both MLM&FxFx

- **Jet Multiplicity** (inclusive, $\geq N_j$ & exclusive $= N_j$)
  
  $kT$-MLM predictions of the jet multiplicity have little sensitivity to the UE and PS tunes

- **$p_T$ balance Z-jet**
  
  sensitive to the UE and PS tunes

- **Z boson momentum** (for $Z+\geq 1j$)
  
  interesting to see the effect of the tune especially at low momentum
Jet-Z balance in Z+jets

Fairly good agreement on all the tunes

Nj>3 only CP2 describes data

Jet-Z balance in W+jets

Fairly good agreement on all the tunes

Nj>3 no tunes describe data
CP tunes in high $p_T$ jets & EW bosons final states

Jet-Z balance in Z+jets

(!) Only below 20 GeV we see differences among tunes
CP tunes in high $p_T$ jets & EW bosons final states

hard QCD: $Z$ momentum in $Z$+jets events with jet $p_T$>30 GeV
matrix element generation made by madgraph_aMC@NLO[FxFx and MLM]

\[ p_T(Z), N_{\text{jet}} \geq 1 \sqrt{s} = 13 \text{ TeV} \]

\[ \frac{d\sigma}{dp_T(Z)} \text{ [pb/GeV]} \]

- MLM: overall not able to describe the $Z$ pt
- FxFx: all tunes give reasonable agreement around 5 GeV
- $p_T(Z)$$<$$10$ GeV: description by Monash-based CUETP8M1 is better