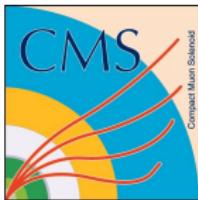


# MC tunning

Timothée Theveneaux-Pelzer

Centre de Physique des Particules de Marseille (CPPM)

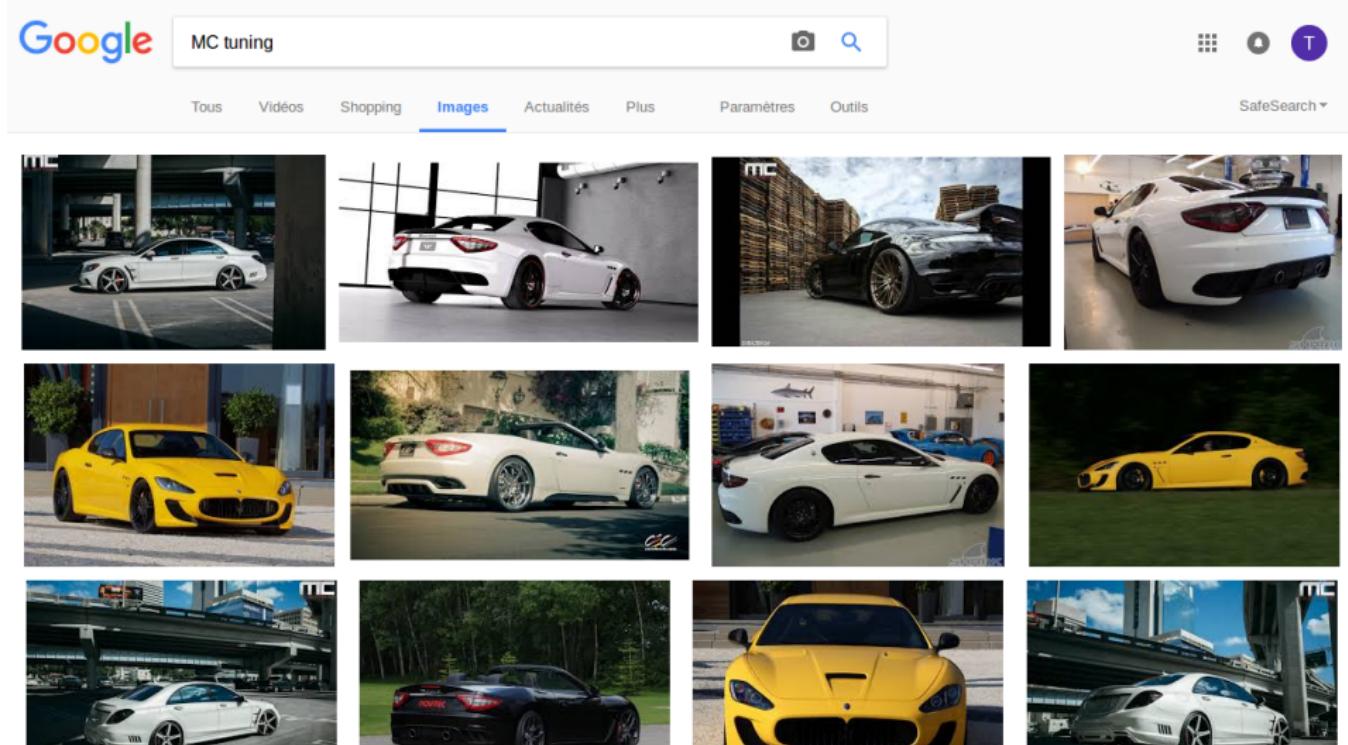
Thursday, May 4th 2017



# Introduction

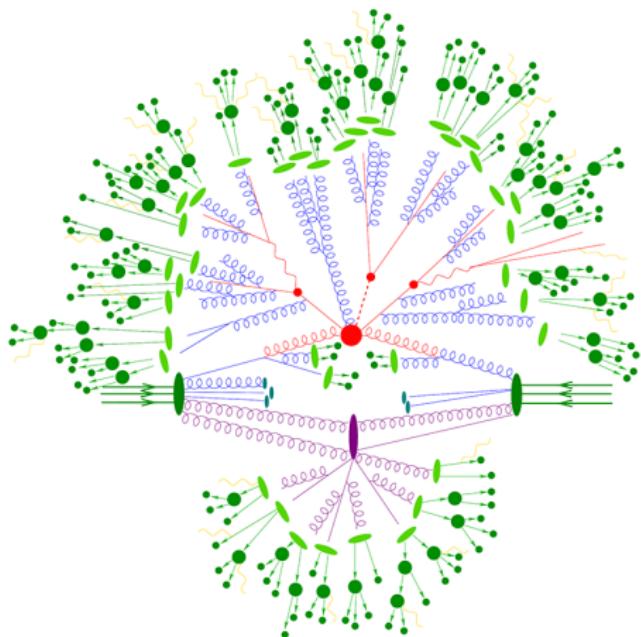
- We need simulated samples to model  $t\bar{t}$ 
  - topology too complicated for fully data-driven techniques !
- MC generators can only be an imperfect modelling
  - several ingredients/approximations are used
- We have to optimise each step of the generation to be closest to reality
  - get a better nominal prediction
  - get better motivated modelling uncertainties - and reduce them
- In this talk : personal view of “MC tuning”
  - speaking mostly of  $t\bar{t}$

# “MC tuning” in google



# Modelling a proton-proton collision

- Hard process
- Parton shower : ISR/FSR
- Hadronisation and fragmentation
- Decay
- Underlying event
- Soft radiations
- Multiple Parton Interactions
- Colour Reconnection



# Several ingredients in our $t\bar{t}$ MC samples

- Parton Density Functions

- extracted from fits to specific data (Hera, Tevatron, LHC run-I etc.)
- mostly using 2014 version of each PDF family
- PDF4LHC15 = MMHT14+CT14+NNPDF3.0 : used for uncertainty propagation

- “Matrix Element” generator

- modelling of the hard collision
- for  $t\bar{t}$  we have NLO predictions
- different methods to match with PS generator, e.g. Powheg and MC@NLO

- “Parton Shower” generator

- models different non-perturbative things
- shower, hadronisation, fragmentation, UE, MPI, CR
- hadronisation with effective model : string (Herwig) or cluster (Pythia, Sherpa)

- Optional : additional legs in the ME generator

- better modelling of high-pT radiations
- can be LO (Alpgen, MadGraph, Sherpa) or NLO (aMC@NLO, Sherpa) legs
- requires merging procedure, i.e. when the ME stops and when the PS starts

- Optional : afterburners in specific cases

- EvtGen, Tauola, Photos

- NB : some ingredients are multi-purpose

- e.g. we can generate  $t\bar{t}$  NLO+PS with Herwig7 standalone

# Choosing the settings for a MC sample

- All these ingredients have many parameters
  - these parameters can be redundant, or non-physical
- Optimisation driven by different arguments
  - ① tradition - re-using settings from other contexts
  - ② trustable theoretical predictions
  - ③ use of existing measurement of a parameter
  - ④ comparison of predictions to existing measurements
- Comparison of predictions to reliable measurements
  - well-defined observables, unfolded at particle level
  - differential cross-sections as a function of many variables
- For  $t\bar{t}$  MC tuning : mostly ATLAS & CMS data
- Rivet framework to exploit existing measurements
  - reproducing the selection at truth level
  - reference data stored in Yoda histograms
  - newly predicted distributions compared to unfolded data
  - use of Professor framework for tuning



Rivet



Yoda

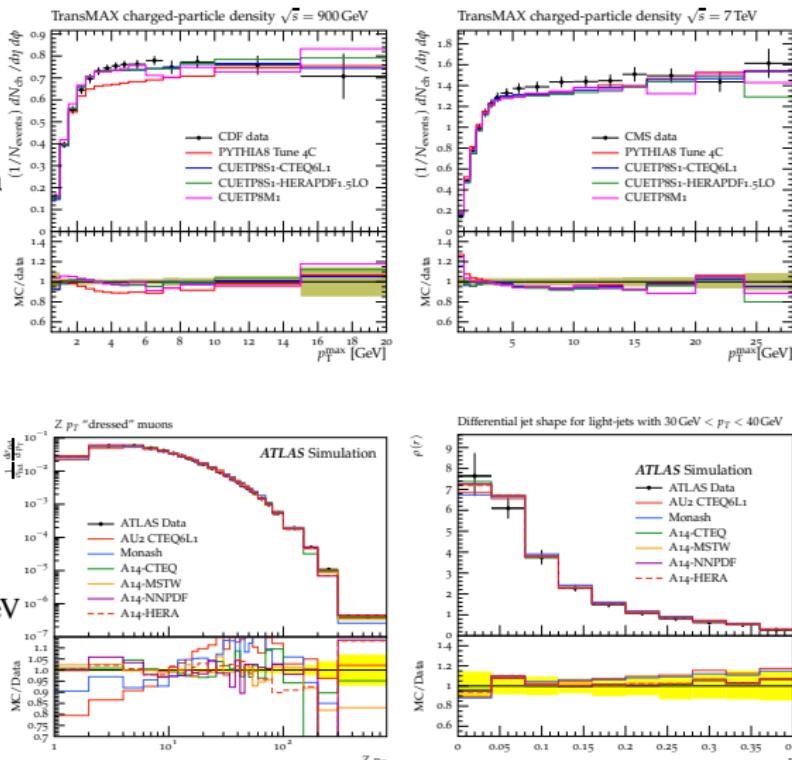


Professor

# Tuning the PS generator

CMS-PAS-GEN-14-001 - ATL-PHYS-PUB-2014-021

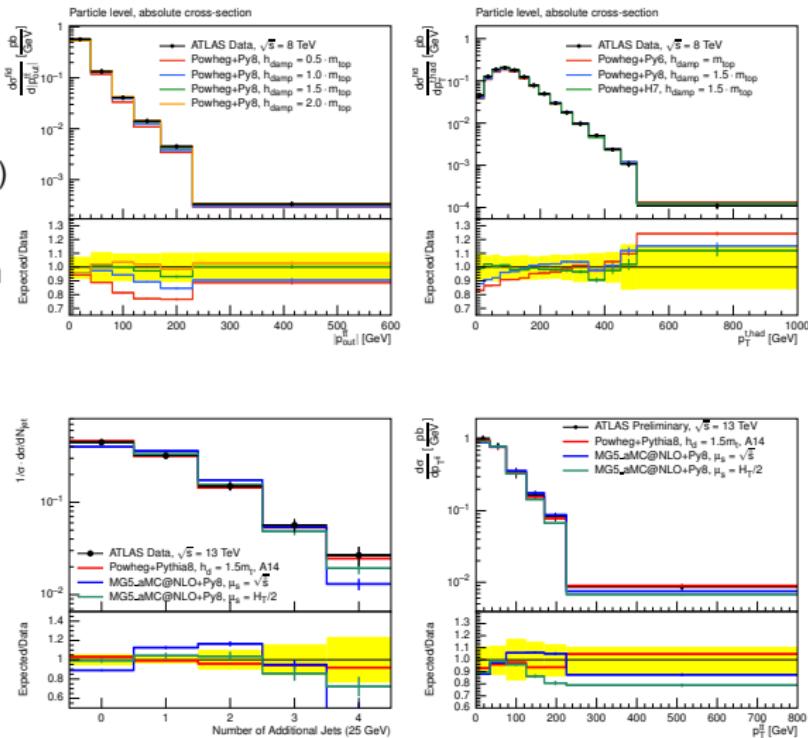
- Example1 : CMS' Pythia8 CUETP8M1
  - parameters : UE/MPI
  - charged particles
  - CDF (0.9, 1.96 TeV), CMS (7 TeV) data
  - 0.3 TeV CDF data not included
- Example2 : ATLAS' Pythia8 A14
  - parameters : UE/MPI, CR, ISR/FSR
  - $Z+jets$ ,  $t\bar{t}$ , multijet, charged particles
  - ATLAS (7 TeV) data
- Many other tunes in the market
  - using LEP data
  - using LHC only data, but non- $t\bar{t}$
- A tune may not work for all  $\sqrt{s}$  values
  - UE very different at 0.3 TeV and 7 TeV
- A tune may not work for all processes
  - CR very different for  $t\bar{t}$  and  $Z+jets$



# Plugging a PS generator on a ME generator

ATL-PHYS-PUB-2016-020 - ATL-PHYS-PUB-2017-007

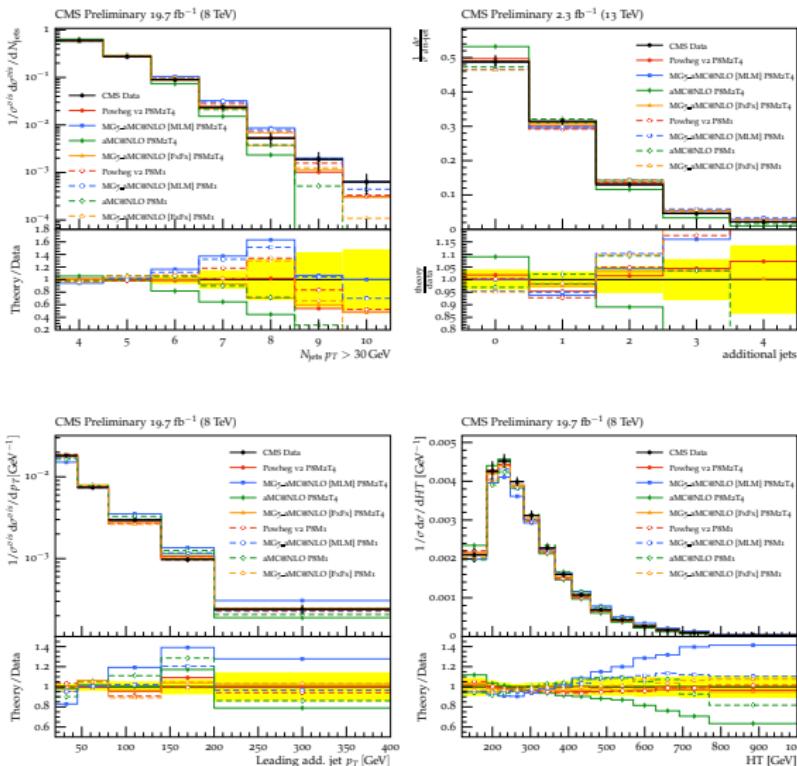
- NLO predictions matched to PS
  - Powheg (Powheg-Box)
  - MC@NLO (aMC@NLO, Sherpa)
- hdamp parameter in Powheg
  - damping of the hardest radiation
  - not used years ago
  - better modelling around  $1.5 \cdot m_{top}$
- Matching scale in aMC@NLO
  - issue in Njets,  $p_T^{\bar{t}\bar{t}}$  with default
  - trying better functional forms
  - problem not fixed with multi-leg
- PS generator changes kinematics
  - (much) better  $p_T^{top}$  with Herwig7



# Combined tune of the ME+PS generator

CMS-PAS-TOP-16-021

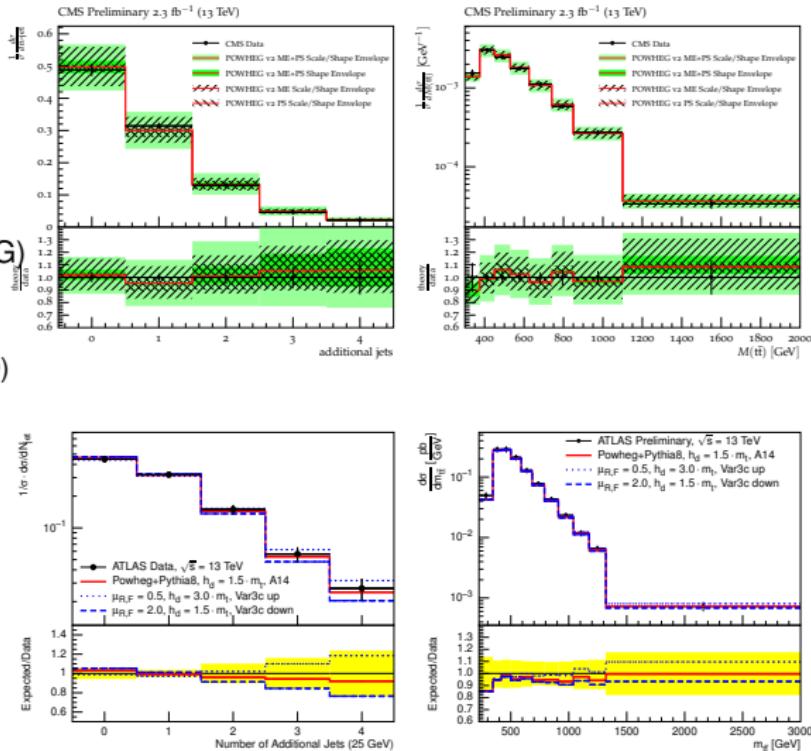
- CMS CUETP8M2T4 tune
  - combined Powheg+Pythia8 tune
- Tuning hdamp and  $\alpha_s$ 
  - using 8 TeV  $t\bar{t}$  data
  - Njets and lead. add. jet  $p_T$
- Then re-tune of UEMPI
  - fix hdamp and  $\alpha_s$
  - same as CUETP8M1
- Pythia8 plugged on several MEs
- Similar strategy in ATLAS
  - ATTBAR : ATL-PHYS-PUB-2015-007
  - 7 TeV  $t\bar{t}$  data
- Issues of this approach
  - hdamp only for Powheg- $t\bar{t}$
  - UE/CR different in other processes



# Systematic uncertainties

CMS-PAS-TOP-16-021 - ATL-PHYS-PUB-2017-007

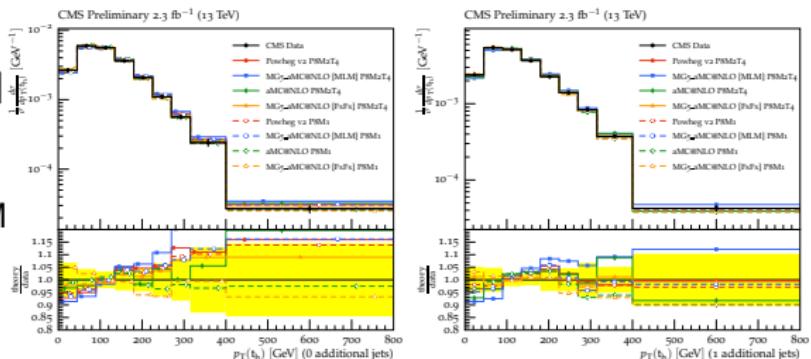
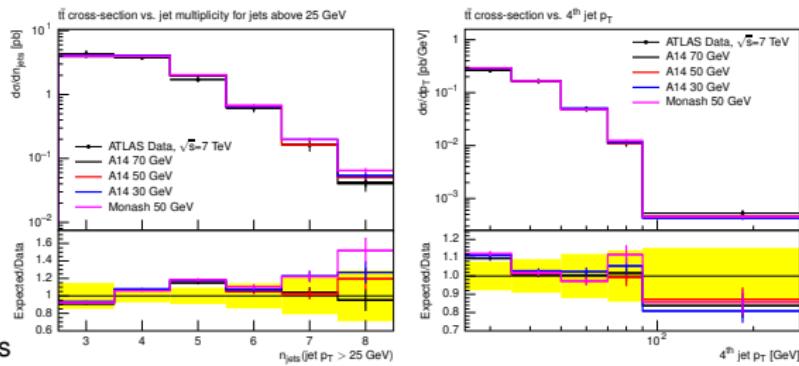
- Different ingredients to model  $t\bar{t}$ 
  - different uncertainties to reflect our understanding or each step
  - comparing different realistic setups
- Current split (ATLAS/CMS/LHCTopWG)
  - PDFs
  - hadronisation (Pythia vs. Herwig)
  - NLO match. (Powheg vs. MC@NLO)
  - radiation (both in ME and PS)
  - ad hoc ( $m_t$ , CR, UE/MPI)
- CMS radiation systematic
  - vary  $\mu_{R,F}$  (ME and PS)
- ATLAS radiation systematic
  - vary  $\mu_{R,F}$  (ME), A14 tune, hdamp
  - brackets Njets experimental uncert.



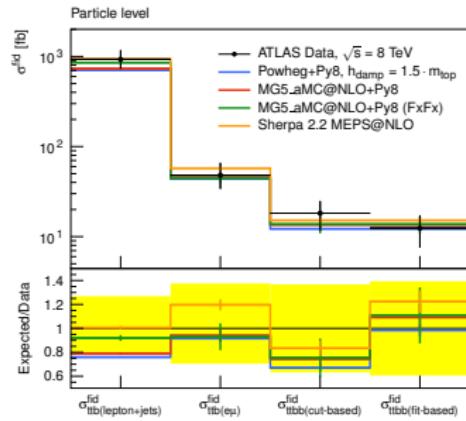
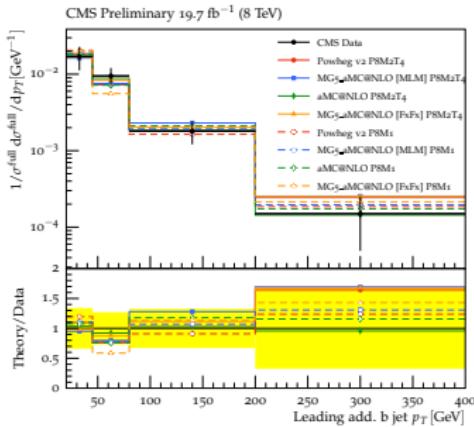
# $t\bar{t}$ +jets

ATL-PHYS-PUB-2016-016 - CMS-PAS-TOP-16-021

- Bigger challenge :  
good  $t\bar{t}$  modelling vs. N jets
- Use of multi-leg ME generators
  - better model of hard radiations
  - more parameters to tune
- Many options with aMC@NLO
  - MLM matching : up to 4 LO legs
  - FxFx merging : up to 2 NLO legs
  - UNLOPS merging : up to 2 NLO legs
- Sherpa/OpenLoops (own shower)
  - 1 leg @NLO, 4 legs @LO
- ATLAS : testing aMC@NLO[FxFx]
  - optimisation of merging scale
  - no clear improvement
- CMS : testing both FxFx and MLM
  - FxFx favoured over MLM
- Tuning to be eventually made  
in exclusive jet bins



- Even more challenging : modelling of  $t\bar{t}$  in association with HF jets
  - QCD at two very different scales ( $m_t, m_b$ )
- Some unfolded data already available to tune against
  - CMS differential vs. additional  $b$  and  $b\bar{b}$  kinematics
  - ATLAS inclusive in 3 fiducial volumes
- Not sensitive enough - need more measurements
- Critical topic for  $t\bar{t}H(\rightarrow b\bar{b})$  analysis (and other searches)



# Conclusions

- Wide topic !
- Very different ingredients, many parameters to tune
- Key is to chose which parameters to tune, and with which measurements
- Good MC modelling of  $t\bar{t}$  relies on good measurements
- Full run-2 dataset should lead to many more improvements
- Multi-leg NLO generators already explored, and should/will be better optimised