Experimental review on event generators in ATLAS and CMS

Taming the accuracy of event generators (Part 2)

23rd August 2021

Simone Amoroso, Gurpreet Singh Chahal

Why Monte Carlo Event Generators?

- MC event generators are ubiquitous at the LHC
- Different use cases lead to different requirements
  - **To determine detector efficiencies, optimise analyses, unfold the data for detector effects**:

    need the best possible **description of the data**, irrespective of the formal accuracy of the prediction
  - **To estimate SM backgrounds, extrapolate from control to a signal regions, interpret the data**:

    need the best possible **formal accuracy**, Irrespective to whether the predictions will describe data.

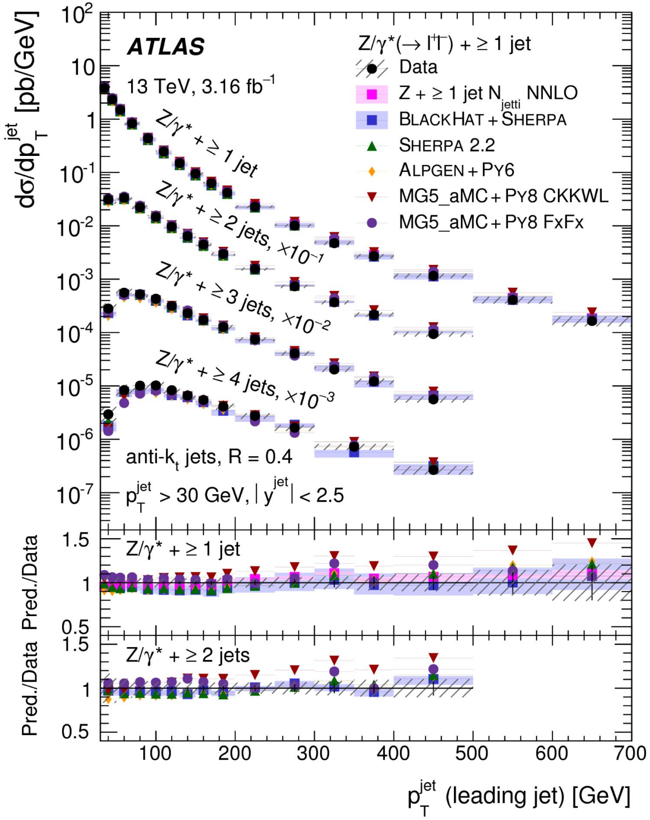
    Essential to have recipes to estimate the **associated theoretical uncertainties**

State-of-the art (in experiments)

- The LHC pushed the automation of NLO+PS and of LO/NLO-merging techniques
  - Needed to describe with decent accuracy high jet multiplicity phase-spaces

- These approaches now constitute the backbone of the MC samples used in experiments
  - **Main background** samples use **NLO-merging** (MEPS@NLO, FxFx, MiNLO)
  - **Inclusive samples** for precision measurements typically at **NLOPS** (Powheg/MC@NLO)
  - **BSM signal** samples rely on **LO-merged** (MLM/CKKW-L) simulations

STDM-2016-01

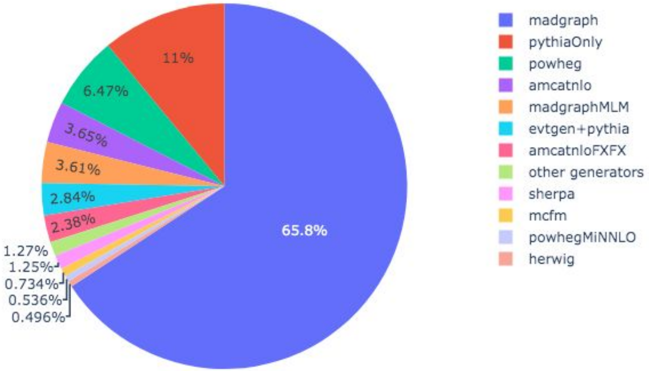


Generator usage in ATLAS/CMS

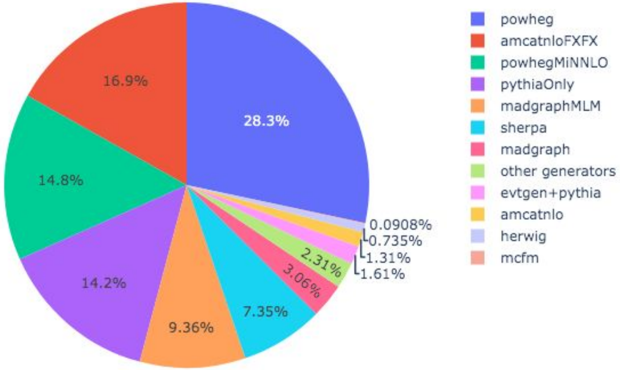
based on Run-2 MC campaign for 2016 data

CMS

number of samples by generator

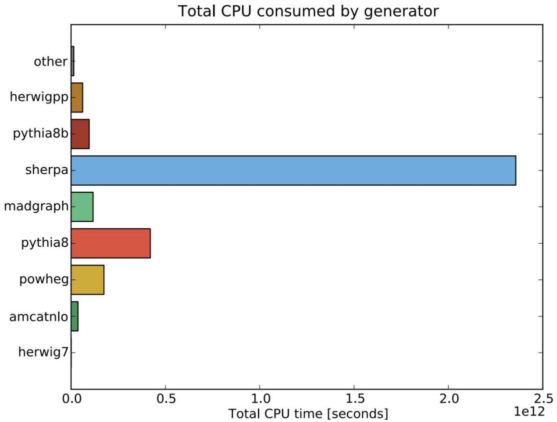


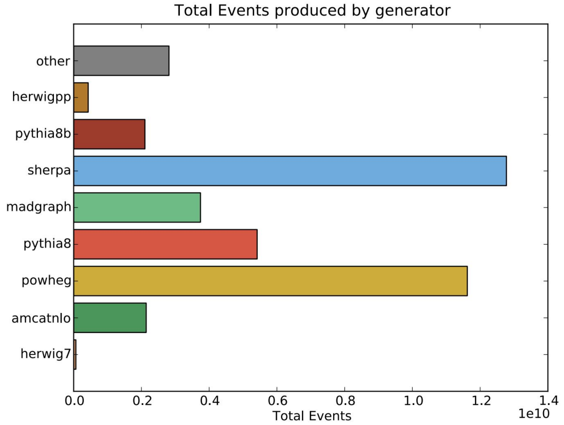
number of events by generator



based on Run-2 MC campaign for 2015 data

ATLAS





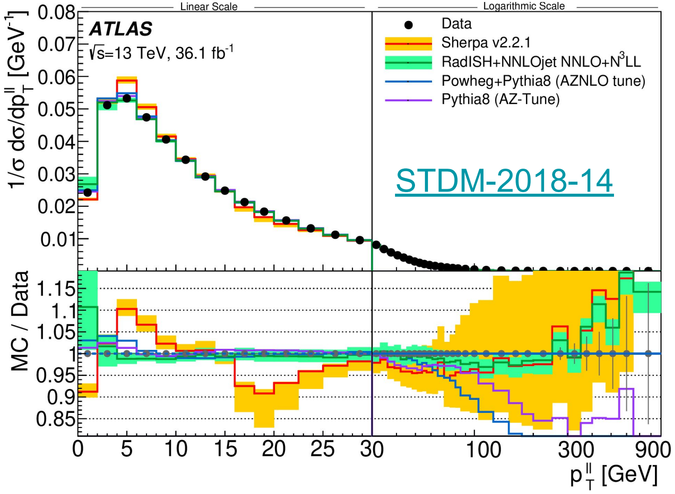
The need for accuracy

- With the increase in the integrated luminosity we are moving towards the analysis of very **exclusive phase-spaces** and **rare processes** while our measurements reach **extreme accuracies**
  - Experimental uncertainties (mostly) scale with the integrated luminosity
  - But theory uncertainties do not!
- Improvements in theory are essential for the successful exploitation of the (HL-) LHC dataset

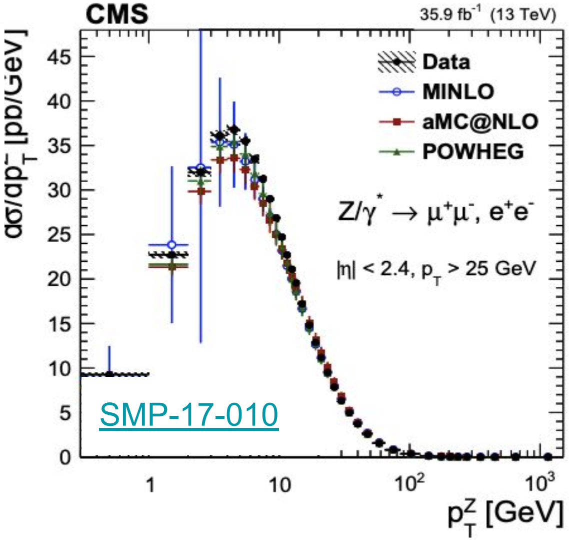
- In the next set of slides we'll go through:
  - a few example cases where the need for higher accuracy is evident already now
  - A wishlist of theoretical (and other more technical) developments

Modelling of color singlet p_T

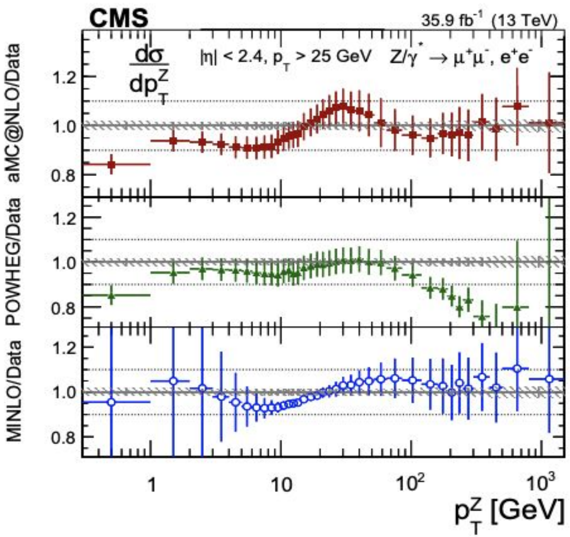
- The p_T of colour singlets is the prime distribution to benchmark our understanding of QCD
    - Can be measured to permill accuracy (in Z events)
    - Relevant for many precise measurement (W-mass, Higgs p_T)
    - Probes transition from non-perturbative physics to resummation and fixed-order



STDM-2018-14

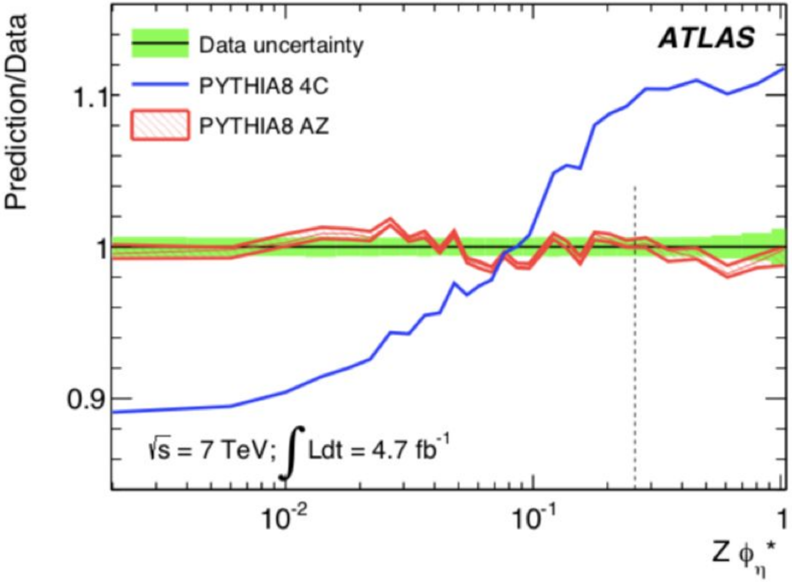


SMP-17-010



Example: W-mass measurement

- W-mass measurement very sensitive to the description of the W boson p_T
  - Relevant region is p_T^w<40 GeV
  - ~2% uncertainties on p_T^w translate into a 10 MeV uncertainty on mW
  - Direct theory uncertainty is significantly larger than this



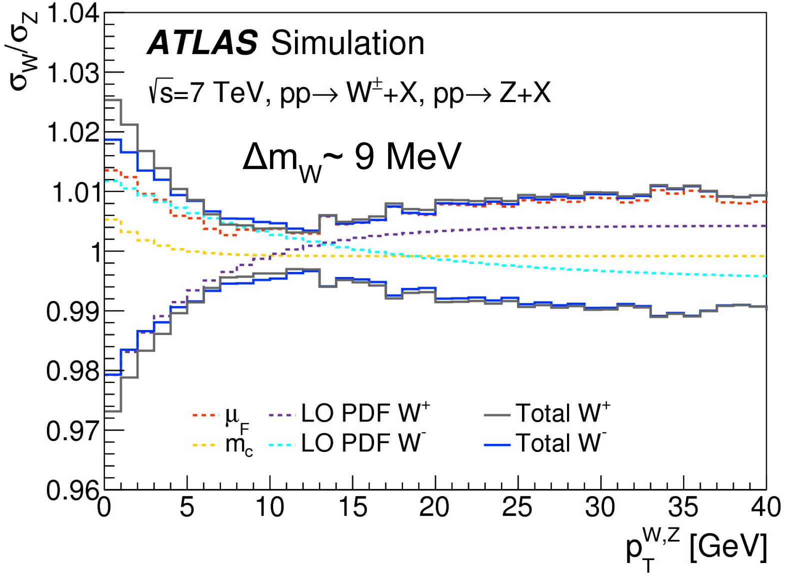
- Exploit precise measurements of Z bosons p_T to get best possible description of p_T^w

$$p_T^W = R_{W/Z} \cdot p_T^Z$$

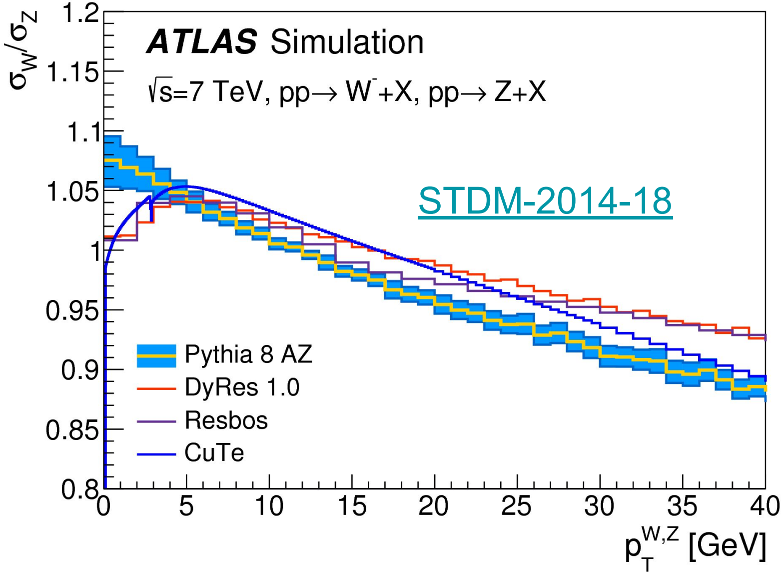
- Model of Z p_T obtained tuning a flexible MC prediction (Pythia8) to data
- But crucial to get an accurate estimate of effects which decorrelate between W and Z

W/Z ratio and uncertainties

- Attempt to describe the ratio through higher accuracy analytic calculations or MC generators
    - Only shower and NLOPS predictions able to describe the data



Δm_W~ 9 MeV

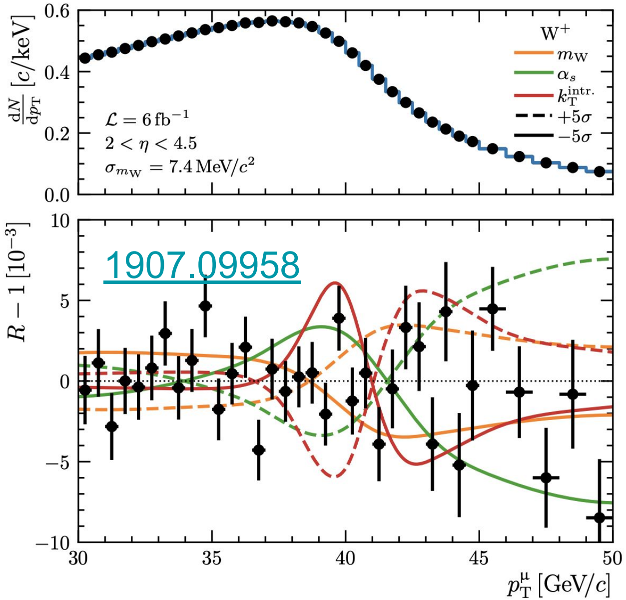


STDM-2014-18

- Pythia8 used also to obtain an uncertainty on the W/Z ratio
    - Decorrelate μ_F between light and HF contributions as a proxy of HF matching scale variations

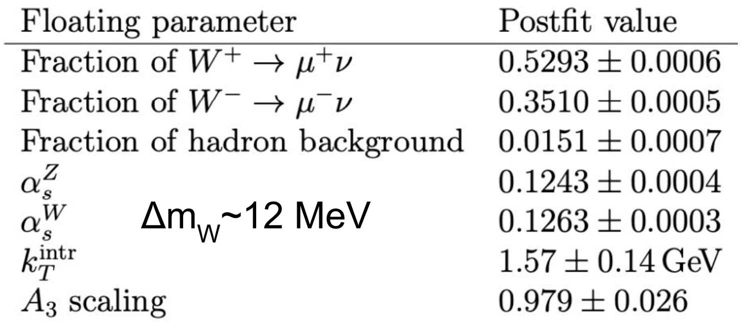
A QCD model for the W pT

- LHCb presented last month their first W-mass measurements using 13 TeV data
    - Obtained performing template fits to the p_T^μ distribution
- The issues with the W/Z extrapolation are avoided by simultaneously fitting to the p_T^μ distribution the W boson mass and the parameters of a QCD model of the W p_T



1907.09958

- Method stress-tested by successfully fitting pseudo-data from different predictions
- But to which accuracy can we believe the correlations in p_T given by these two Pythia8 parameters?

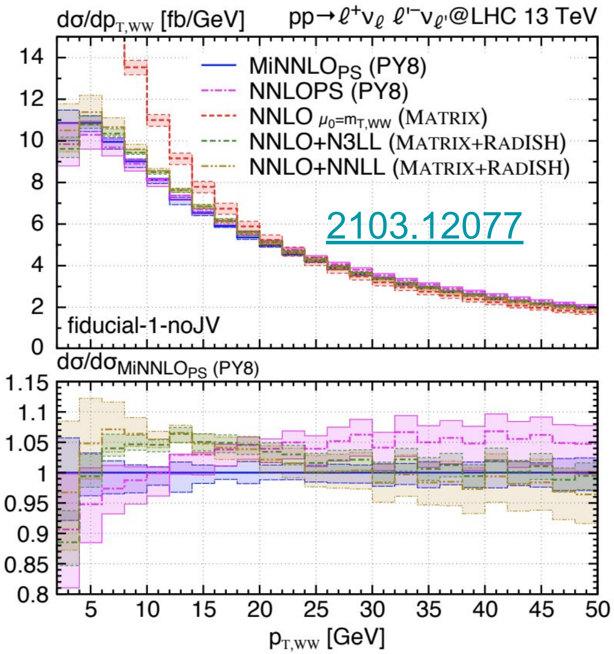


Floating parameter	Postfit value
Fraction of $W^+ \to \mu^+\nu$	0.5293 ± 0.0006
Fraction of $W^- \to \mu^-\nu$	0.3510 ± 0.0005
Fraction of hadron background	0.0151 ± 0.0007
α_s^Z	0.1243 ± 0.0004
α_s^W	0.1263 ± 0.0003
$k_T^{\rm intr}$	1.57 ± 0.14 GeV
A_3 scaling	0.979 ± 0.026

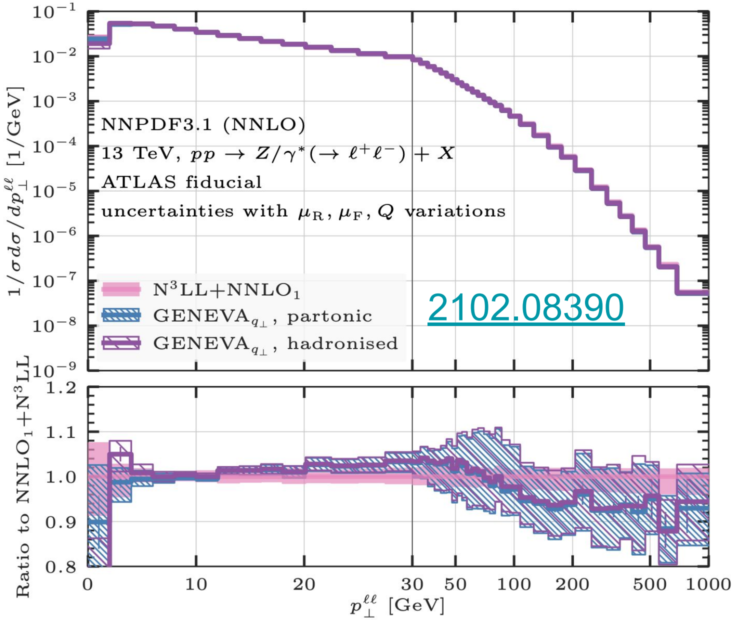
Δm_W~12 MeV

NNLO+PS, the new standard?

- NNLOPS formally developed since a while (UN2LOPS), in the past couple of years an explosion of new results
- Mostly Geneva (only beta release public) and MiNNLOPS (now going beyond color singlets)
  - Good agreement with N3LL analytic resummation, better than LL shower accuracy



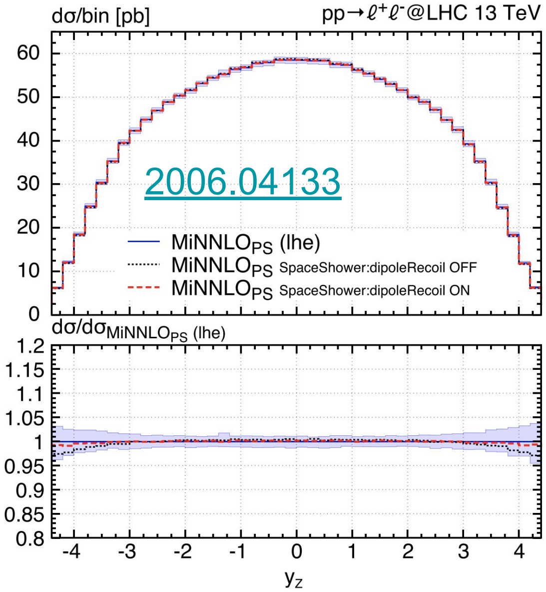
2103.12077



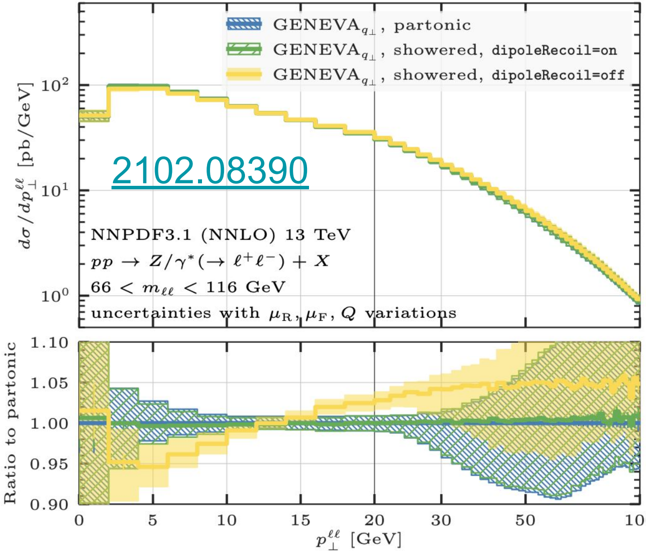
2102.08390

NNLO+PS and shower recoils

- So far MiNNLO/Geneva have only been interfaced to the Pythia8 shower
  - Surprisingly large impact of ISR shower recoil: real uncertainty or is global recoil just wrong?
  - Can we match NNLO to other showers at LL accuracy? What about NLL?



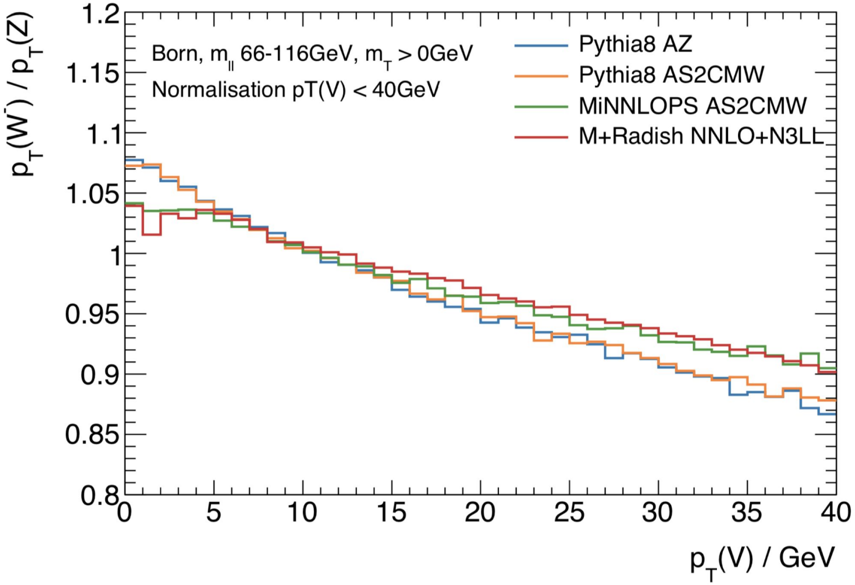
2006.04133

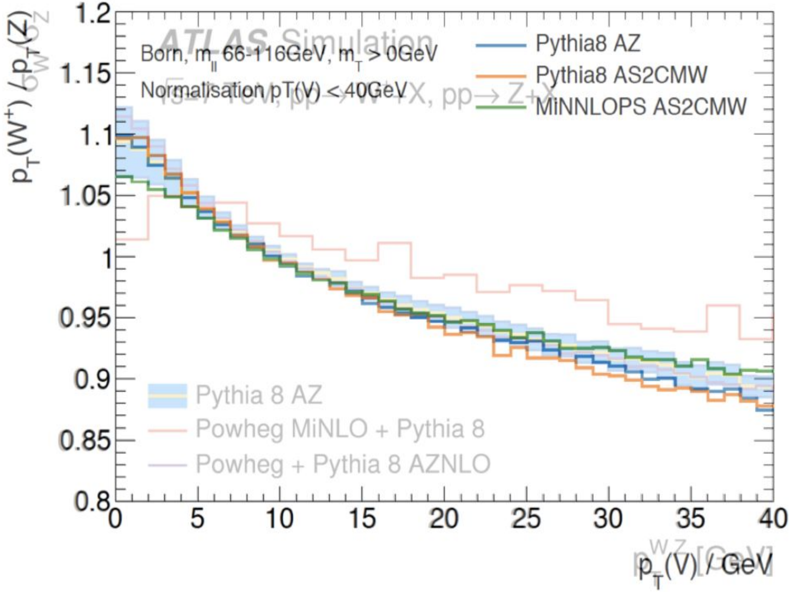


2102.08390

MiNNLOPS and the W/Z ratio

- New MiNNLOPS within 1-sigma of Pythia8 AZ (and so of data) in the W/Z ratio
    - Fixes problems with MiNLO and analytic resummation codes
    - Perfect agreement with NNLO+N3LL from Matrix+Radish
- Can we construct a W pT model with a more sophisticated/QCD driven model than with Pythia8





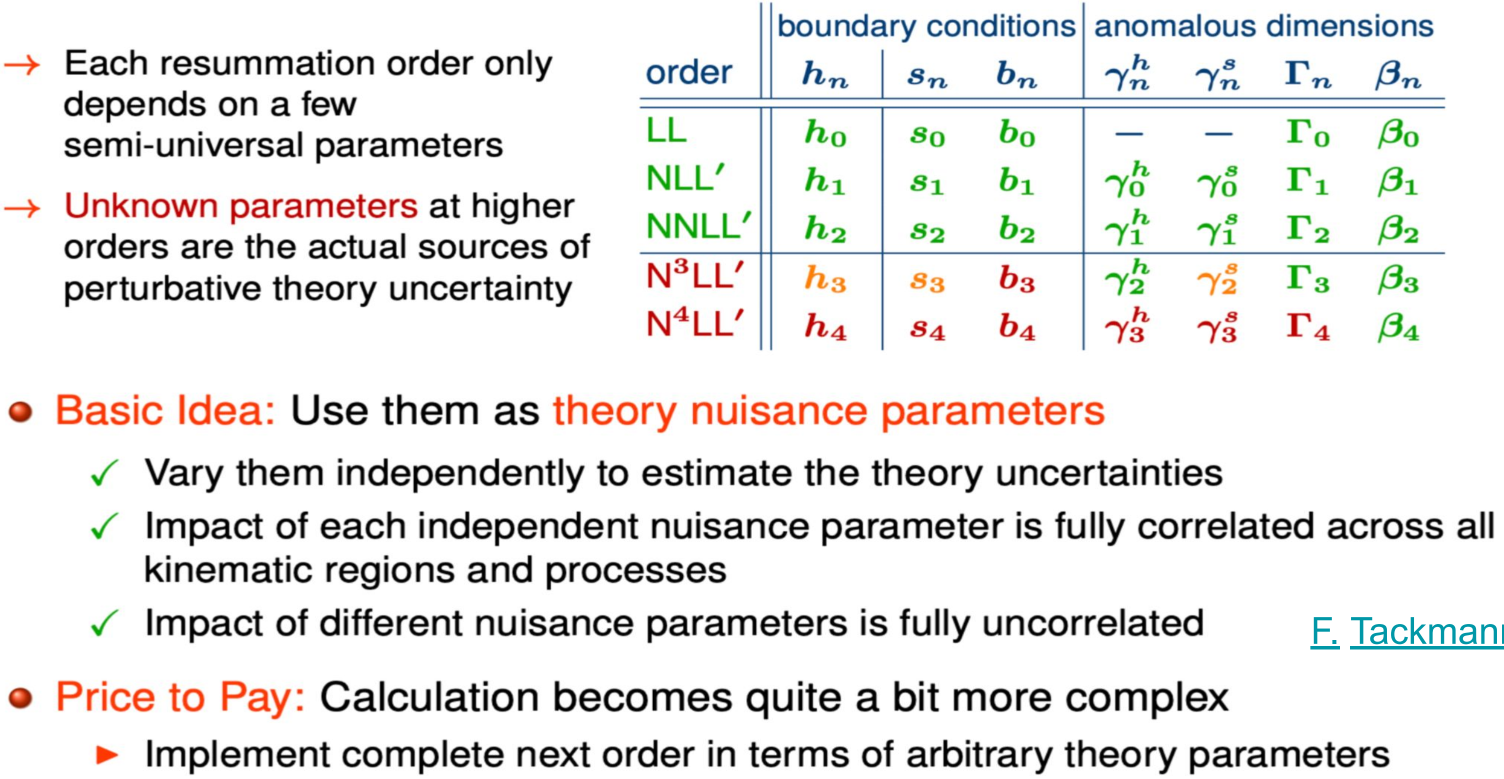
Uncertainties for ME predictions

- A prediction is only as good as its associated uncertainty
- More and more LHC analyses are moving towards **complicated fits**, in which **theory uncertainties are incorporated in the likelihood** and determined in situ together with the parameter of interest
    - Was the case only for Higgs and BSM, but SM measurements are catching up
    - General belief that data can constrain theory beyond its validity range
- We need a proper and **reliable model of theory uncertainties,** including their correlations across different phase-spaces, observables and processes!

- This is notably a very complicated problem for missing higher order uncertainties, which we estimate with scale variations
    - Several attempts for fixed-order uncertainties in the past years (2006.16293, 2106.04585)
- Any progress in this area would be highly welcome

Uncertainties for ME predictions

- Using the resummation formalism can exploit the known structure of missing higher-orders to parametrise them in terms of nuisance parameters



→ Each resummation order only depends on a few semi-universal parameters

→ Unknown parameters at higher orders are the actual sources of perturbative theory uncertainty

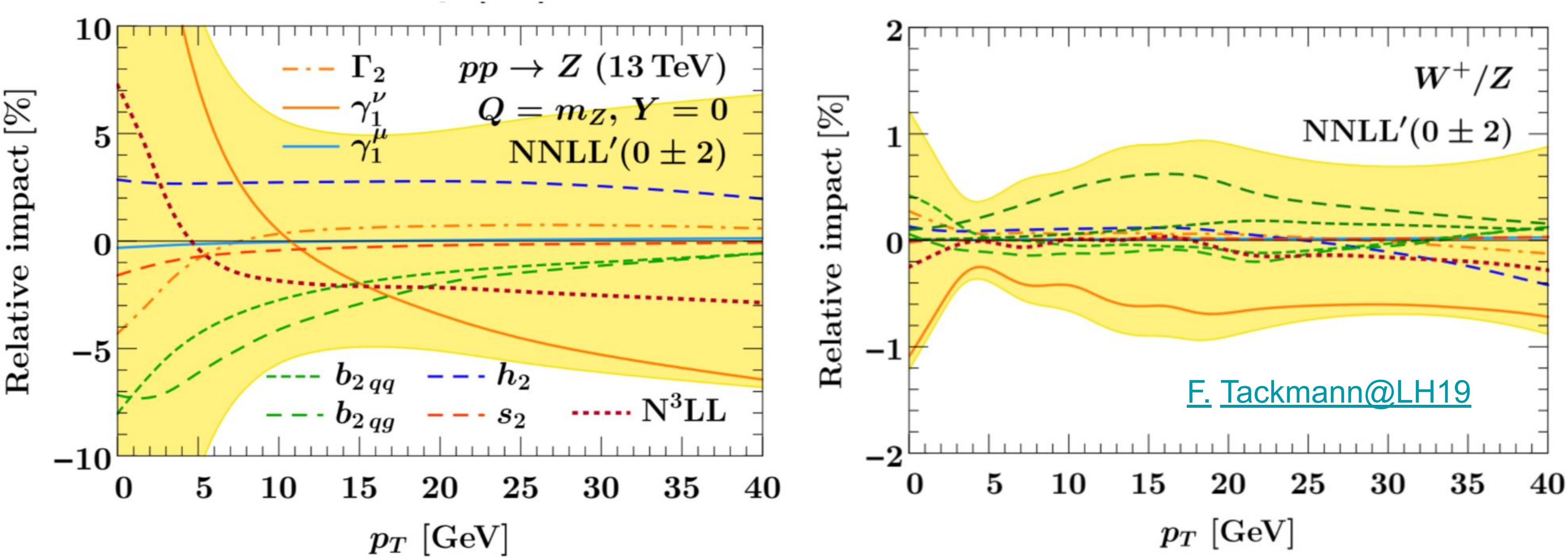
	boundary conditions			anomalous dimensions			
order	h_n	s_n	b_n	γ_n^h	γ_n^s	Γ_n	β_n
LL	h_0	s_0	b_0	—	—	Γ_0	β_0
NLL′	h_1	s_1	b_1	γ_0^h	γ_0^s	Γ_1	β_1
NNLL′	h_2	s_2	b_2	γ_1^h	γ_1^s	Γ_2	β_2
N^{3}LL′	h_3	s_3	b_3	γ_2^h	γ_2^s	Γ_3	β_3
N^{4}LL′	h_4	s_4	b_4	γ_3^h	γ_3^s	Γ_4	β_4

- **Basic Idea:** Use them as theory nuisance parameters
  - ✓ Vary them independently to estimate the theory uncertainties
  - ✓ Impact of each independent nuisance parameter is fully correlated across all kinematic regions and processes
  - ✓ Impact of different nuisance parameters is fully uncorrelated
- **Price to Pay:** Calculation becomes quite a bit more complex
  - ► Implement complete next order in terms of arbitrary theory parameters

F. Tackmann@LH19

Uncertainties for ME predictions

- Obvious use-case in the W->Z extrapolation at small p_T for the W-mass

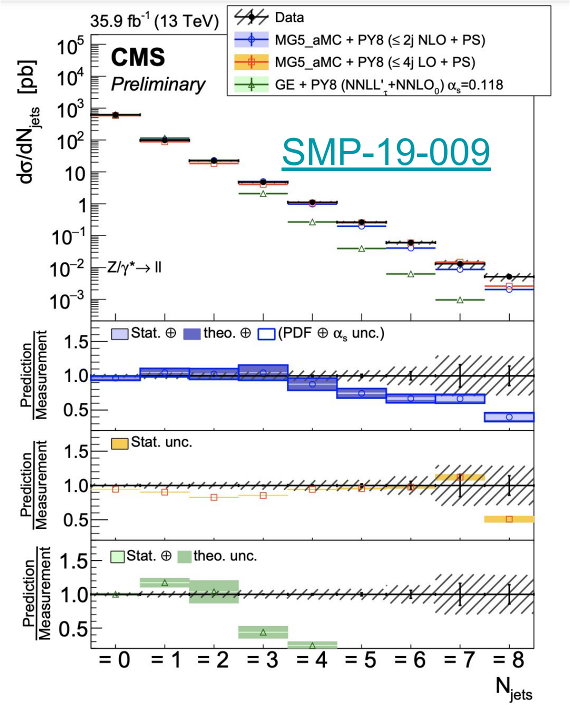


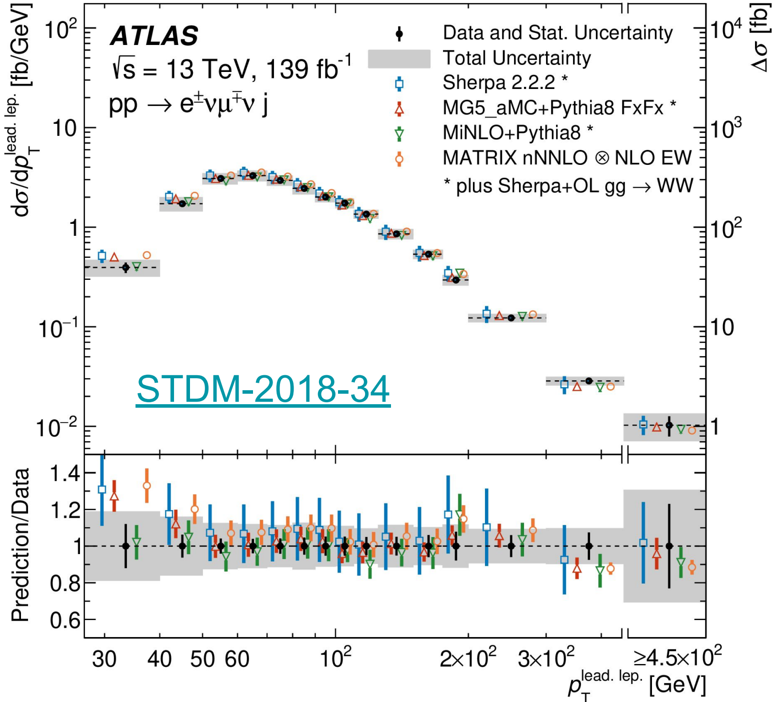
F. Tackmann@LH19

- Can something of this sort be incorporated in MC generators?
  - i.e. MiNNLO/Geneva exploit similar ingredients from resummation

Merging at high multiplicities

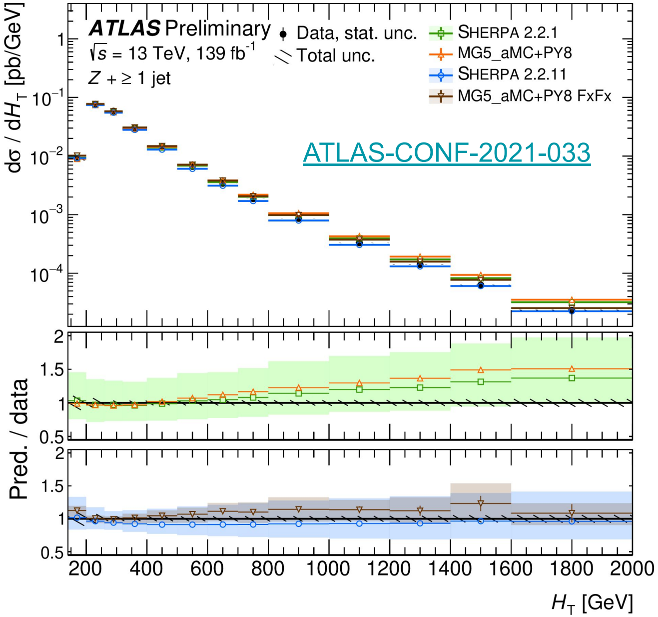
- NLO-merging is the workhorse of ATLAS/CMS MC samples
    - Provides an excellent description of the data in exclusive phase-spaces


SMP-19-009


STDM-2018-34

Merging at high multiplicities

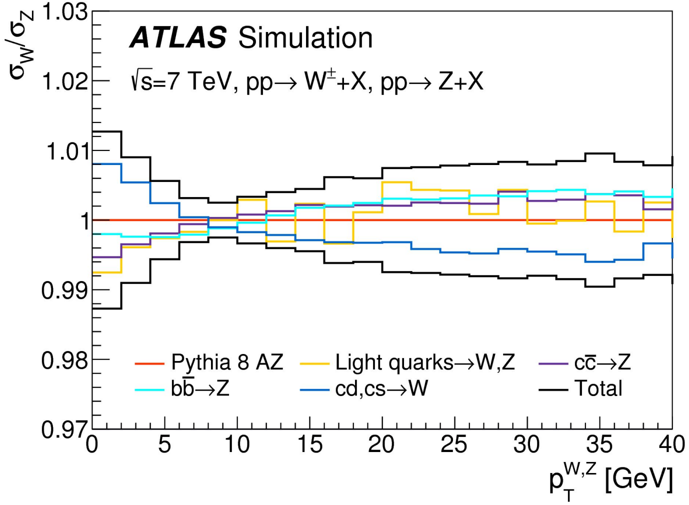
- Current applications only limited by computing considerations. But work still ongoing towards:
  - Optimisation of subleading choices, to get better description of data
  - Optimisation of computational resources (negative weight fraction, phase-space biasing)
- Example ATLAS optimisation of Sherpa 2.2.11 samples:
  - V+0,1,2jet@NLO+3,4,5@LO
  - Inclusion of ME+PS scales, PDFs, NLO EW corrections and 4FS/5FS fusing as weights
  - Reduced (from ~20% to ~10%) neg. weight Fraction thanks to an approximated treatment of subleading color (no change in phenomenology)
  - Faster generation due to different scale choice
  - Optimised phase-space biasing to only generate Events where they are going to be used

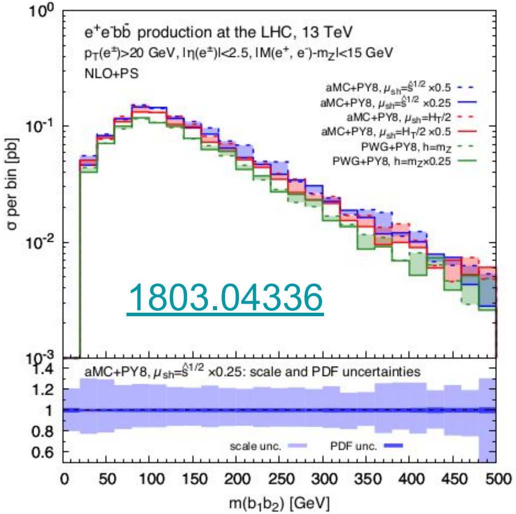


ATLAS-CONF-2021-033

Heavy Flavour production processes

- Heavy flavor production constitute a special class of processes
  - HF-initiated contributions to Z, W, Higgs p_T distributions
  - Z/W+HF as background to VH->bb, tt+bb background to ttH,4-top
- Described with 5FS samples, as shower contribution cannot be neglected
  - Complex reweightings often needed to obtain decent description of the data



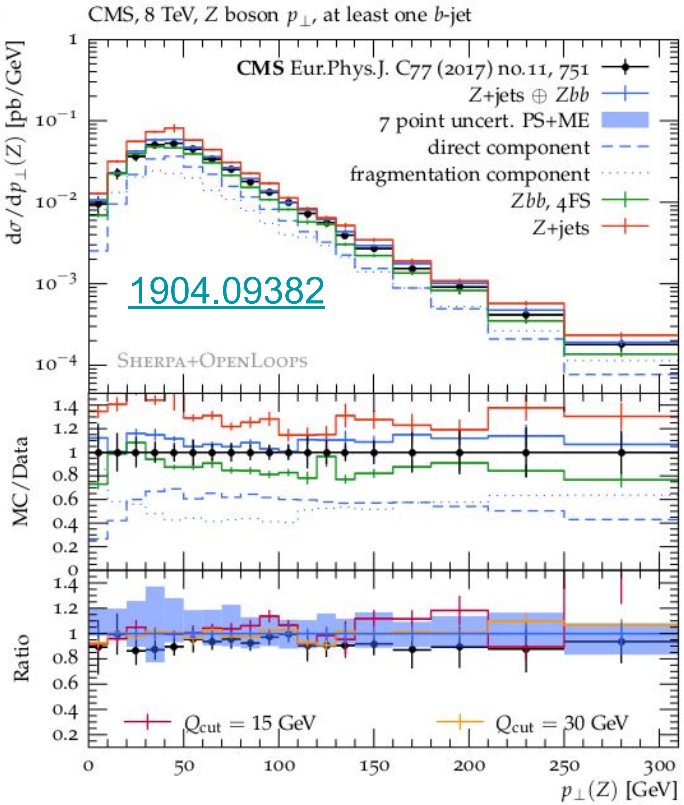


1803.04336

Heavy Flavour production processes

- Variable flavour-number scheme (FONLL-like) recently implemented by Sherpa in the context of MEPS@NLO merging
    - Merging of a 5FS massless calculation with a 4FS massive one (applied to Z+jets/Z+bb)
    - Double counting of events can conveniently be removed through an event-weight
    - Already being benchmarked in ATLAS

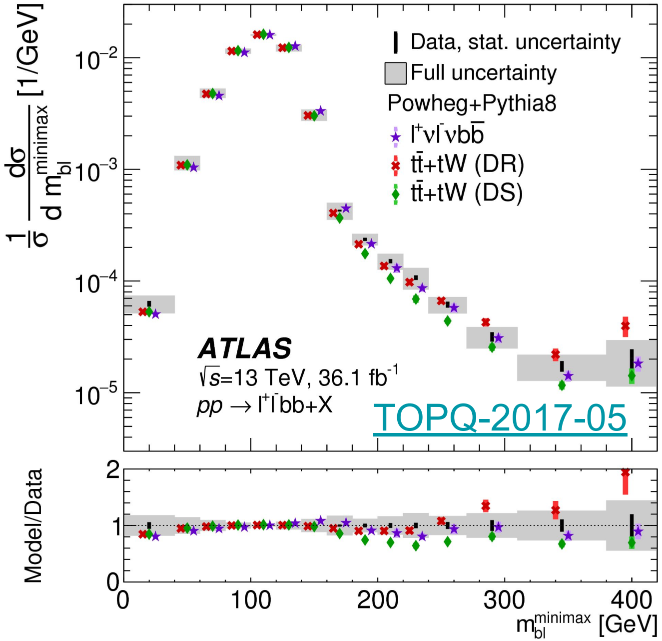
- Opens up new interesting possibilities:
    - Extension to other processes (tt/tt+bb, jets?)
    - Extension to include charm AND bottom thresholds?
    - How hard is it to extend it to other NLO-merging schemes: FxFx, MiNLO (MiNNLO)?



1904.09382

Resonance-aware matching

- ttbar/Wt diagram removal/subtraction prescription is the dominant uncertainty in many analyses
- Resonance aware matching developed within Powheg and applied to bb4l production (1607.04538)
    - Long awaiting update to same-flavour leptons and semileptonic decays (all-had possible?)
    - Matching with showers cumbersome, again need extension of LHE standard?
    - Do we need uncertainties on the width regularisation? (alternatives to the complex-mass-scheme)



TOPQ-2017-05

- Possible also with MC@NLO matching (1305.7088), but negative weight fraction >40%. Perspectives?
- NLO for a few other processes exist (ttj, ttW). Could we get them matched to the shower?

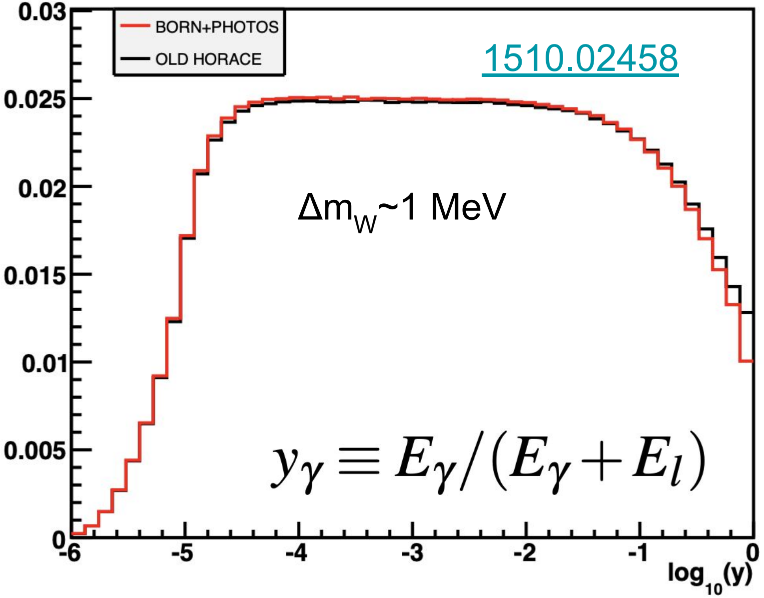
Not only QCD (higher EW orders)

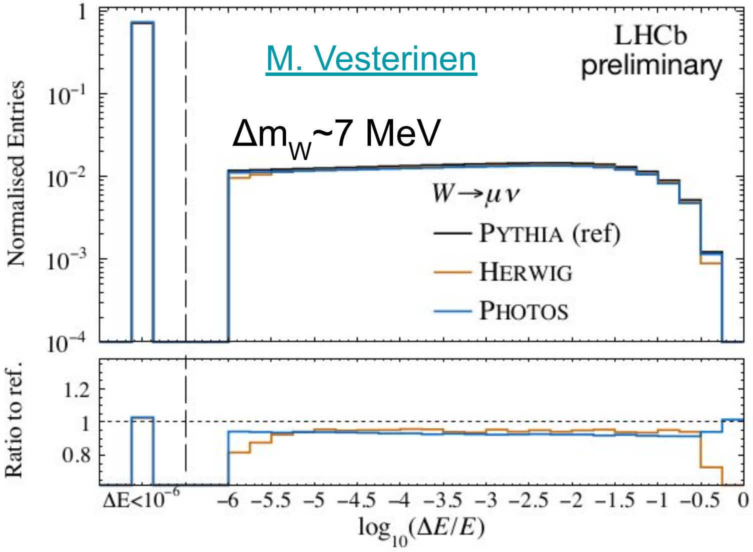
- EW effects beyond LO can become important with the increasing accuracy of our measurement, or when they get enhanced in specific phases-spaces

- Different level of approximations available in MC generator codes
    - **QED FSR** typically included through shower approximation (i.e. Photos, YSF)
    - **EW Sudakovs** logarithms at high energy
    - Full **NLOPS at QCD+EW** for selected processes and analyses
- All needed and being explored by experiments

QED FSR

- EW corrections can be enhanced by collinear final-state QED radiation, leading to large shape effects
  - Typically resummed within the parton shower approximation (i.e. Photos, YSF) + ME corrections
  - Several comparisons in the past, but codes/models evolve. Do we need a thorough benchmarking?

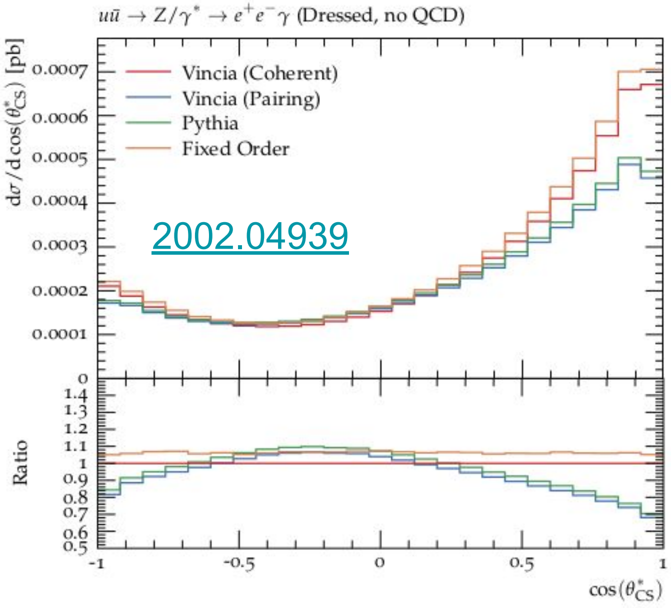




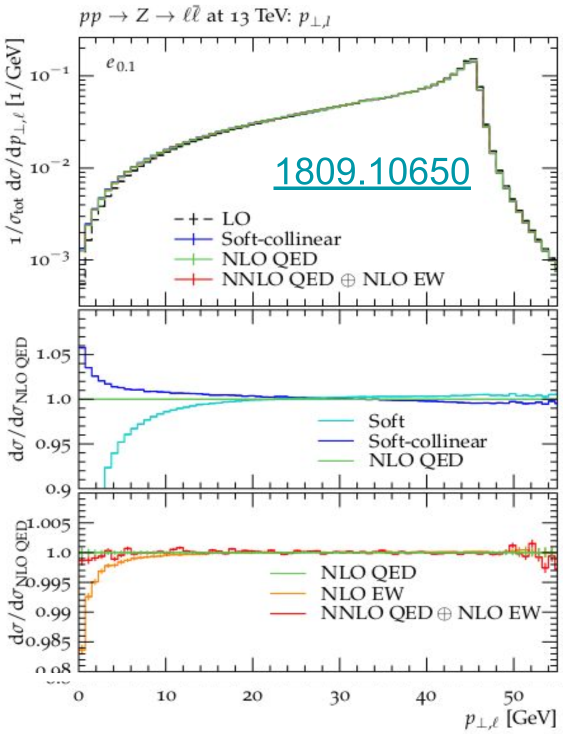
The reweighting variable where ΔE is the difference in energy between the final-state lepton pair before and after QED FSR.

QED FSR

- Several improvements already exist. How much would they affect precise EW measurements (m_W, θ_W)?



2002.04939



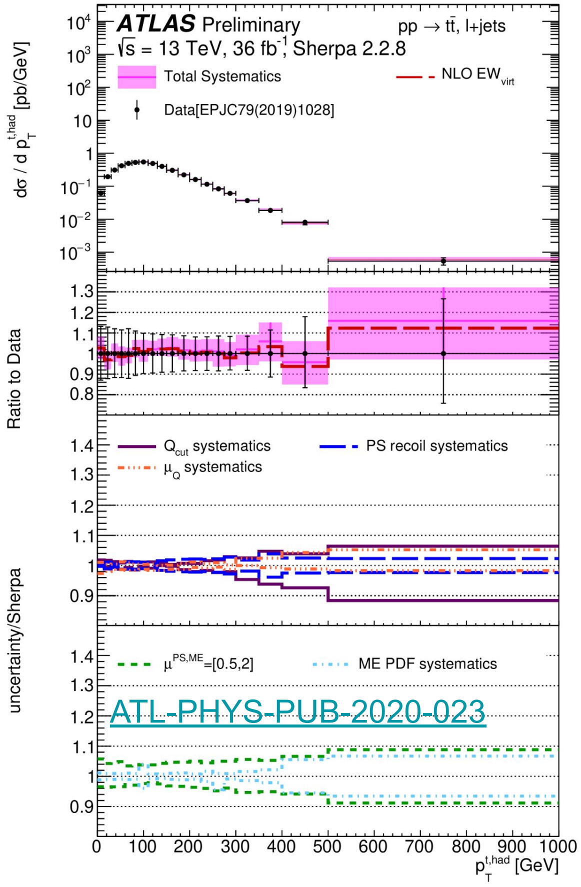
1809.10650

- Interleaved QCD and QED evolution in Vincia
- Large effect on ϑ_{CS}, impacts Weinberg angle?

○ NNLO QED FSR in Sherpa

EW Sudakov logs

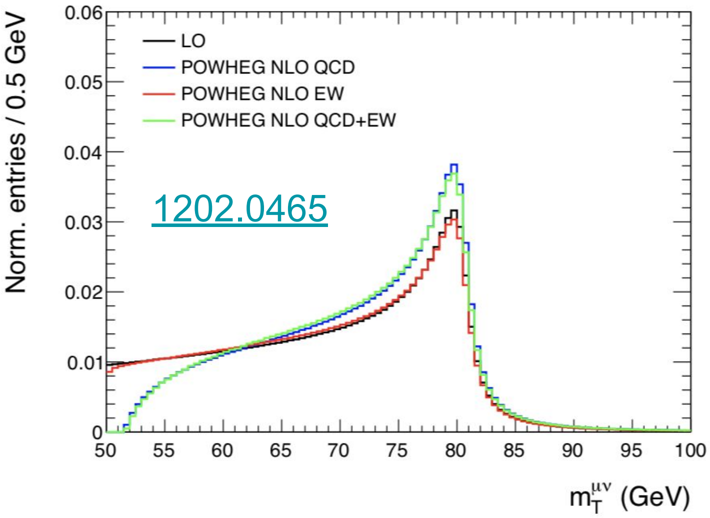
- Virtual weak corrections can reach negative tens of percent in the Sudakov region
  - Essential to consider them in the searches phase-space
- Sherpa now allows for them to be included in an approximated approach within its MEPS@NLO QCD merging
  - EW corrections can be combined with an additive, multiplicative or exponentiated prescription to evaluate uncertainties
  - Conveniently available as weights on top of the QCD-merged prediction (no need for separate samples)
- Will be part of upcoming ATLAS/CMS MC samples



ATL-PHYS-PUB-2020-023

NLOPS QCD + EW

- Few selected Powheg processes allow for NLOPS QCD+EW interfaced to a QED+QCD shower
    - So far available for W, Z, ZH/WH, diboson production
    - Not widely used, but W,Z being commissioned for precision EW measurements



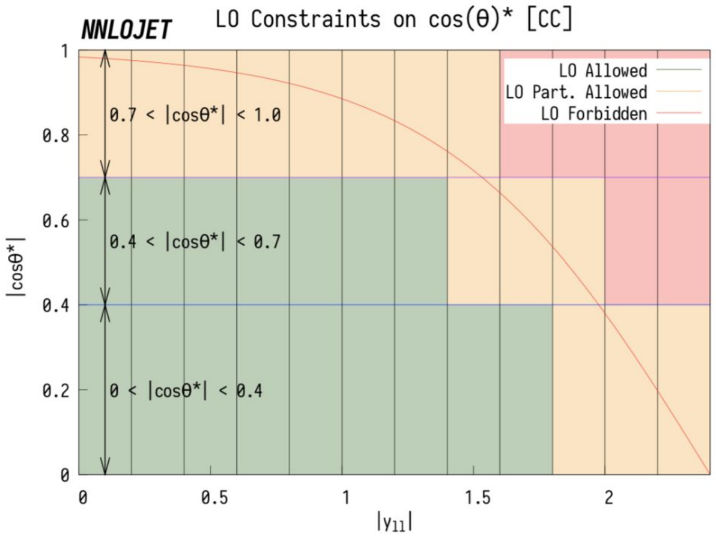
1202.0465

- Can only be interfaced with Pythia8, PHOTOS QED showers. What are the perspectives to extend this to other QED showers?
- Shall we foresee an extension of the LHE standard to facilitate more complex shower veto algorithms?

NLOPS QCD + EW is not enough

- The fiducial selections on the decay kinematic can restrict some measurement bins to be zero at LO
- In Drell-Yan, lepton rapidity cuts induce a LO constraint

$$\cos\theta^* \leq \frac{\sinh(2(y^l_{\max} - |y_{ll}|))}{1+\cosh(2(y^l_{\max} - |y_{ll}|))}$$

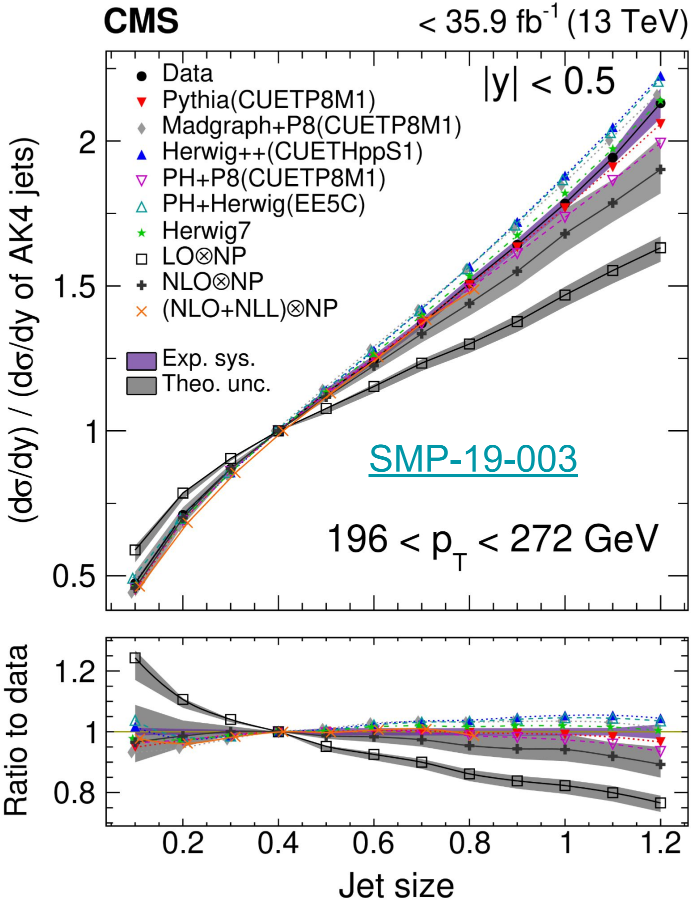


D.Walker Ph.D. thesis

- An NLOPS QCD+EW calculation will add the NLO EW to the Born and spread it across the Sudakov region.
- But to add EW corrections to these LO forbidden bins one would need Z+jet NLO QCD +EW predictions?
- Would it be possible to extend the MiN(N)LO/Geneva prescriptions to a QCD+EW merging?

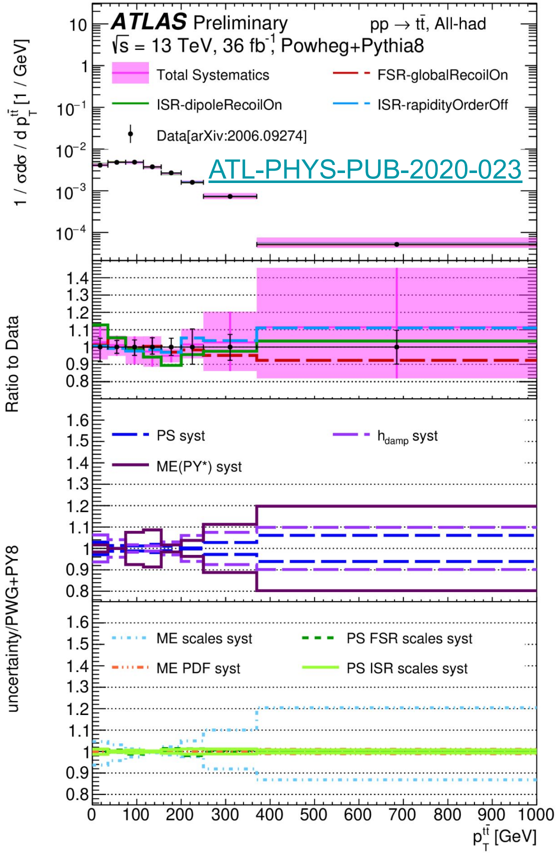
Parton showers accuracy

- The accuracy and uncertainties in parton showers are typically subleading for analyses which use the NLO-merged samples
- The notable exceptions are precision analyses in top, Higgs, DY and jets, which need very large MC samples providing an accurate description of observables inclusive in radiation and sensitive to resummation
    - NLO-merging often introduces "artifacts" and cannot be used (see Z pT case)
    - NNLOPS will certainly help, provided no large "matching" ambiguities are introduced
- The default Pythia8, Herwig7 and Sherpa shower make up 99% of our MC productions.
- High interest in new/better showers (Dire/Vincia/H7-dipole), but little explored and often not supported for matching/merging



Parton shower uncertainties

- Shower uncertainties evaluated with (x0.5,x2) **variations of scales** at which the emissions are performed
    - Available as weight, included in most samples
- Other ambiguities (ordering variable, recoils, …) are ignored, or included through **2-point shower comparisons** (Pytha/Herwig)
    - Need expensive dedicated runs
    - ATLAS study looked at variations of these choices in NLOPS ttbar. Impact not large, but comparable to scale variations
    - Also dedicated studies on recoils impact on log-accuracy from Herwig7 (1904.11866, 2107.04051)
- Another recurrent question is whether hard scattering and parton shower scales should be correlated or not
- Could we get some general recommendation on how to construct a parton shower uncertainty band?



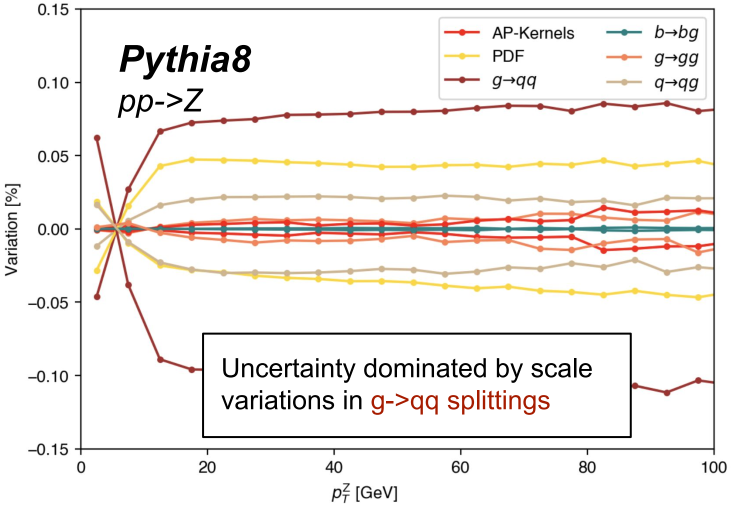
ATL-PHYS-PUB-2020-023

Parton shower nuisance parameters

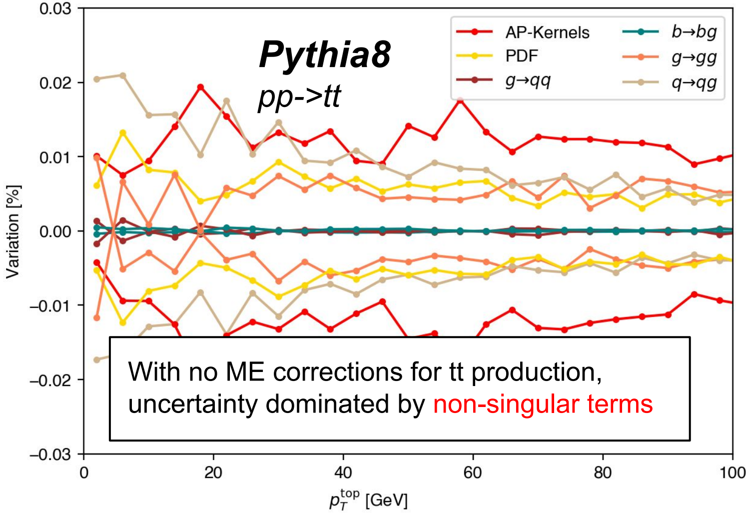
- Pythia8 allows for **decorrelated variations of scales** in the LO DGLAP splitting for the ISR/FSR shower, and for separate variations of non-singular terms (1605.08352)

- Decomposition lends itself to a nuisance-parameters interpretation.
  - Allow the **universal singular terms** to be constrained by data (or higher log-accuracy predictions)
  - **Process and phase-space specific non-singular terms** provide a limit to the possible improvements

- Is this something we could adopt as a **general recommendation?**



Uncertainty dominated by scale variations in g->qq splittings

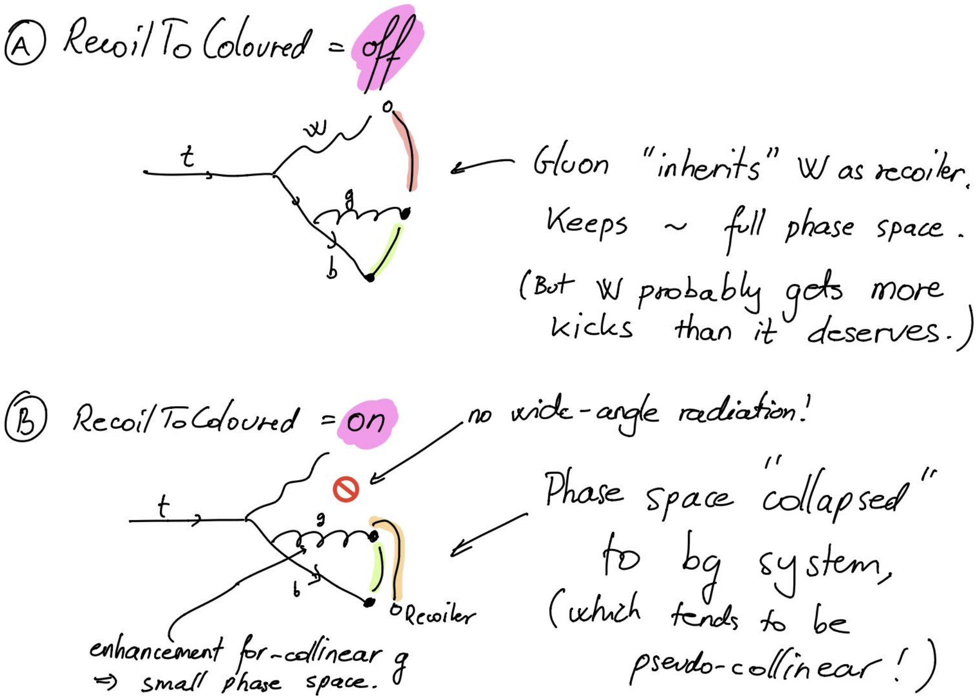


With no ME corrections for tt production, uncertainty dominated by non-singular terms

Shower recoils and top decays

- Recently observed how the Pythia8 shower recoil for the second emission in a top-quark decay can have a huge impact on the reconstructed top observables and on the top mass
  - Pythia8 defaults to assign the Shower **recoil to the b-quark**
  - Can choose as alternative to **recoil against the W-boson**
  - Neither of the two available recoils (b-quark or W-boson) is “correct”, but W-recoil likely better

- Following discussions with the authors, new, more appropriate option implemented, with the **recoil given to the top-quark** through an eikonal reweighting factor



(A) RecoilToColoured = off

Gluon "inherits" W as recoiler.
Keeps ~ full phase space.
(But W probably gets more kicks than it deserves.)

(B) RecoilToColoured = on

no wide-angle radiation!

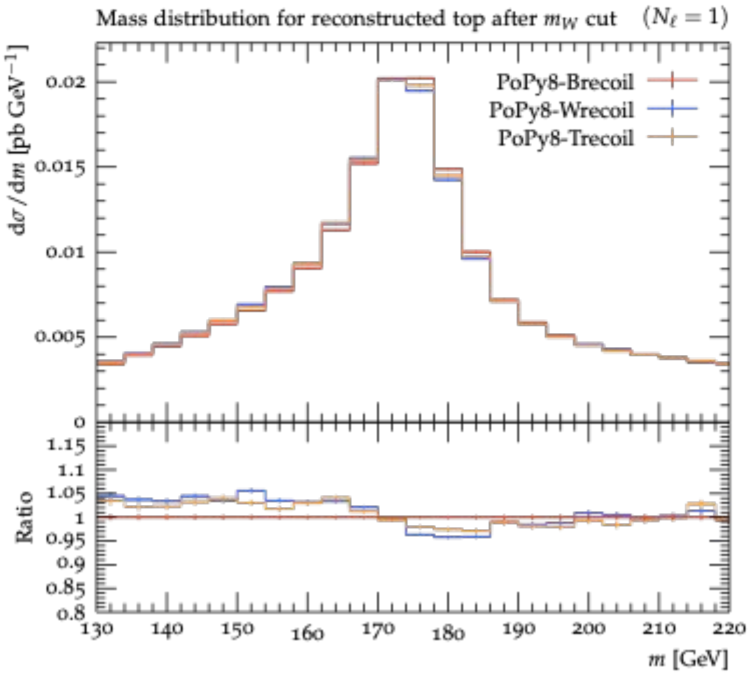
Phase space "collapsed" to bg system, (which tends to be pseudo-collinear!)

enhancement for collinear g
⇒ small phase space.

[from P. Skands]

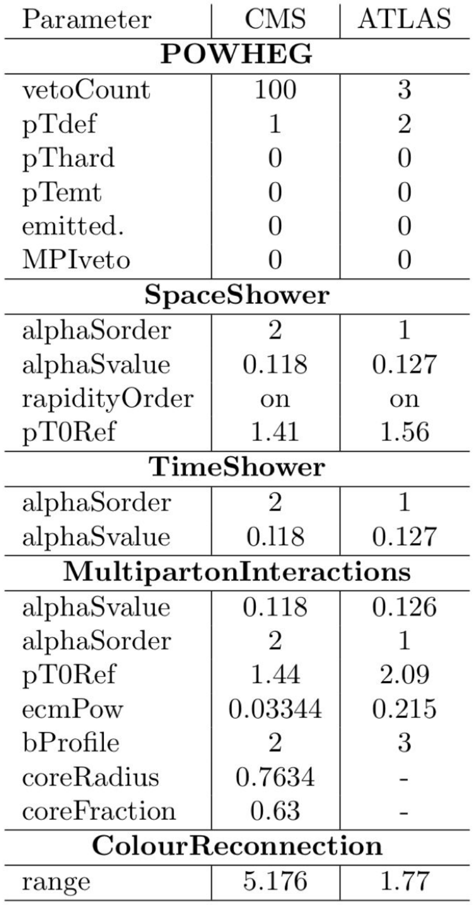
Shower recoils in top decays

- Large impact on top mass, at the level of ~300 MeV. Comparable to FSR uncertainties
  - Top- and W-recoils numerically very similar
- Vincia has an improved treatment of coherence in resonance decays which avoids this issue
  - What about other parton showers?
- A few codes allow for top decays at NLOPS (Powheg, H7-Matchbox) but this issue arise with the second emission
- What are the perspectives for a top decay at NNLOPS?
- How many other cases of "unknown uncertainties" are currently being ignored in LHC analyses?
  - i.e. color flow and shower recoil prescriptions in VBF topologies



NP parameters tuning

- ATLAS/CMS have a tradition to derive their own tunes of Pythia8 (CMS now also H7) parameters
    - Standard Monash includes little LHC data
      And does not provide uncertainties
    - Baseline Monash parameters for fragmentation
      and retuning shower + MPI parameters
    - Different choice of (own) input measurements and processes
- In general, the resulting parameters are mostly consistent
    - α_S 0.118->0.126 from CMW rescaling
    - CR strength very different across experiments
- Yet the small parameter differences make it difficult to compare Analogue Pythia samples across experiments
    - Should we move to joint tunes with the author's help/feedback?

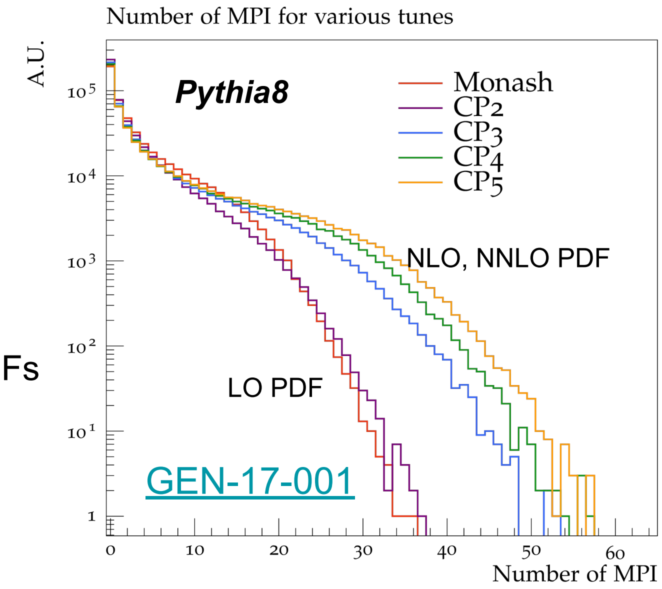


Parameter	CMS	ATLAS
POWHEG		
vetoCount	100	3
pTdef	1	2
pThard	0	0
pTemt	0	0
emitted.	0	0
MPIveto	0	0
SpaceShower		
alphaSorder	2	1
alphaSvalue	0.118	0.127
rapidityOrder	on	on
pT0Ref	1.41	1.56
TimeShower		
alphaSorder	2	1
alphaSvalue	0.118	0.127
MultipartonInteractions		
alphaSvalue	0.118	0.126
alphaSorder	2	1
pT0Ref	1.44	2.09
ecmPow	0.03344	0.215
bProfile	2	3
coreRadius	0.7634	-
coreFraction	0.63	-
ColourReconnection		
range	5.176	1.77

PDFs and parton showers

- Which PDFs should be used in the parton shower evolution?
  - Pythia (and Herwig) advocate for LO PDFs, to get a positive gluon at low-x probed by MPI models
  - Sherpa uses NNLO PDFs, to be consistent with the PDFs in the hard Matrix-Elements
  - Is using LO PDFs only for MPI a better option?

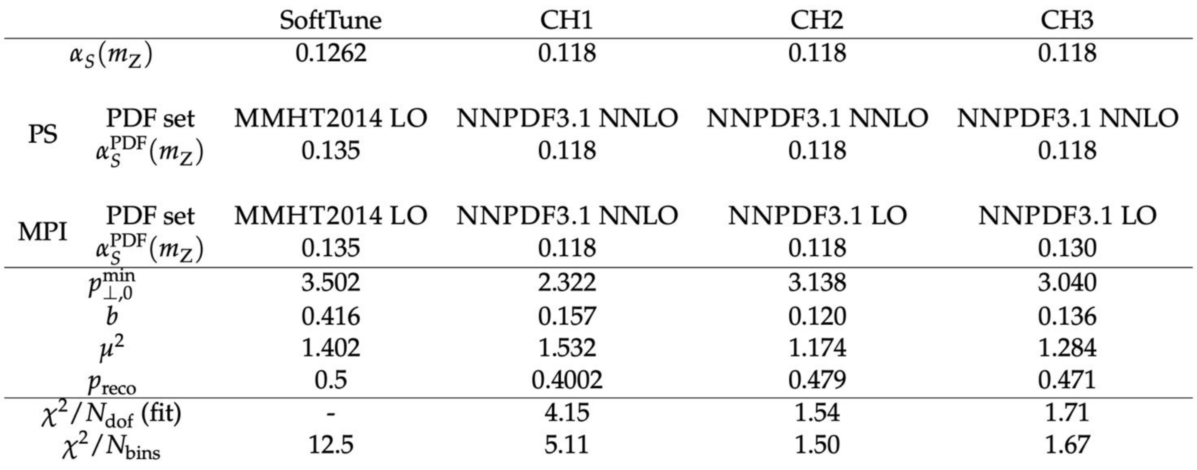
- ATLAS currently using LO PDFs, CMS tuned for different orders and using NNLO PDFs for the nominal tune
- Recent studies (see 2002.04125 and 2003.01700) showed that shower backward evolution does not preserve DGLAP evolution
  - Violation larger with NLO PDFs, but present even for LO PDFs
- Which implications for current shower tunes?
  And for ongoing NLO/NLL shower developments?



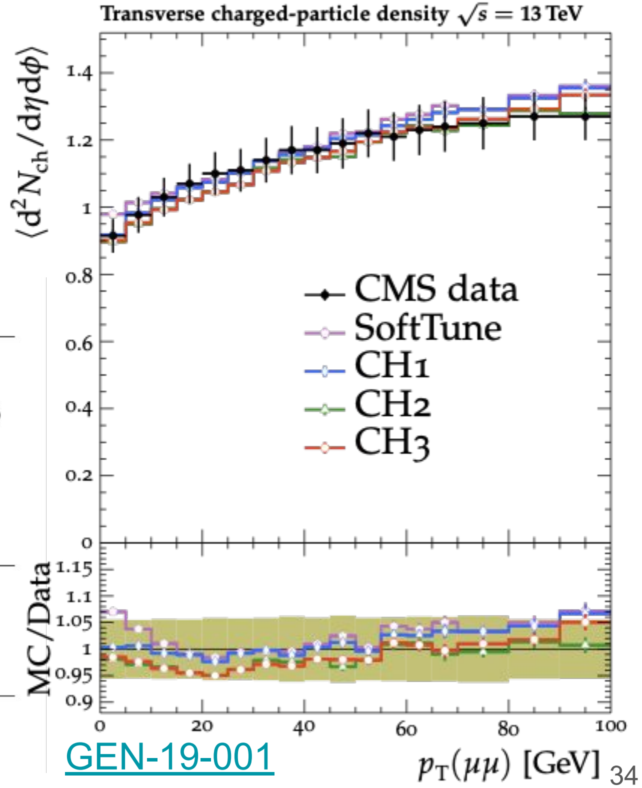
GEN-17-001

CMS tunes and PDFs

- New CMS tunes of Herwig7 explored the description of Minimum Bias data with PDFs of different orders
    - "SoftTune" fitting α_S with LO PDF
    - Other tunes keep α_S=0.118 for consistency with hard ME, and vary α_S and the order of the PDFs in MPI
    - Similar description of the data, but better chi2 with LO PDFs in MPI



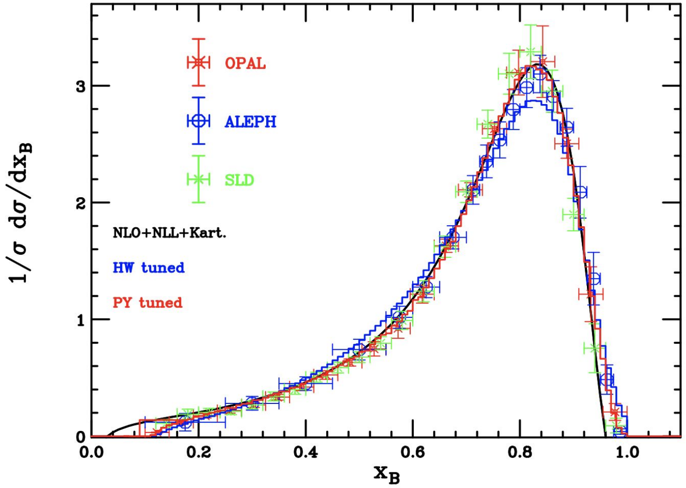
		SoftTune	CH1	CH2	CH3
	$\alpha_S(m_Z)$	0.1262	0.118	0.118	0.118
PS	PDF set	MMHT2014 LO	NNPDF3.1 NNLO	NNPDF3.1 NNLO	NNPDF3.1 NNLO
	$\alpha_S^{\mathrm{PDF}}(m_Z)$	0.135	0.118	0.118	0.118
MPI	PDF set	MMHT2014 LO	NNPDF3.1 NNLO	NNPDF3.1 LO	NNPDF3.1 LO
	$\alpha_S^{\mathrm{PDF}}(m_Z)$	0.135	0.118	0.118	0.130
	$p_{\perp,0}^{\mathrm{min}}$	3.502	2.322	3.138	3.040
	b	0.416	0.157	0.120	0.136
	μ^2	1.402	1.532	1.174	1.284
	p_{reco}	0.5	0.4002	0.479	0.471
	χ^2/N_{dof} (fit)	-	4.15	1.54	1.71
	χ^2/N_{bins}	12.5	5.11	1.50	1.67

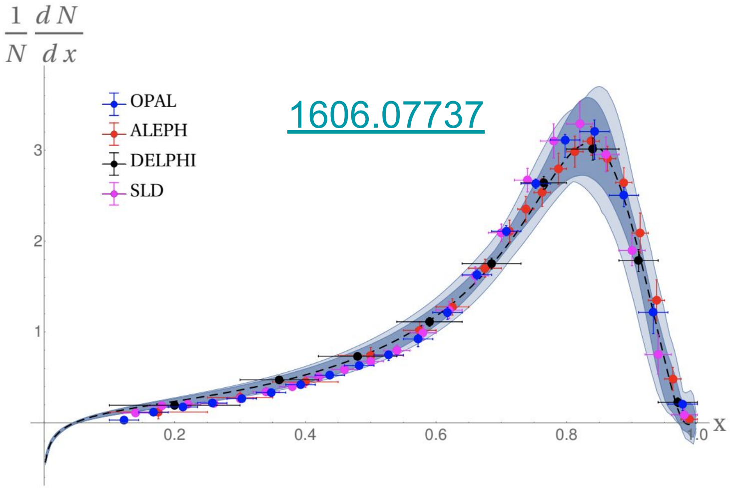


GEN-19-001

Heavy Flavour fragmentation

- The description of heavy-flavour fragmentation is one of the limiting uncertainty in top mass analyses, and important in studies of W, Z+HF and for H->bb/cc decays
- Perturbative fragmentation described by the parton shower, at which accuracy?
    - Old studies comparing NLL resummation with Pythia6/herwig++, to be repeated?
    - Can we learn something from recent computations at NNLO (2102.08267)?
- Non-perturbative component through phenomenological model fitted to LEP/SLD legacy data

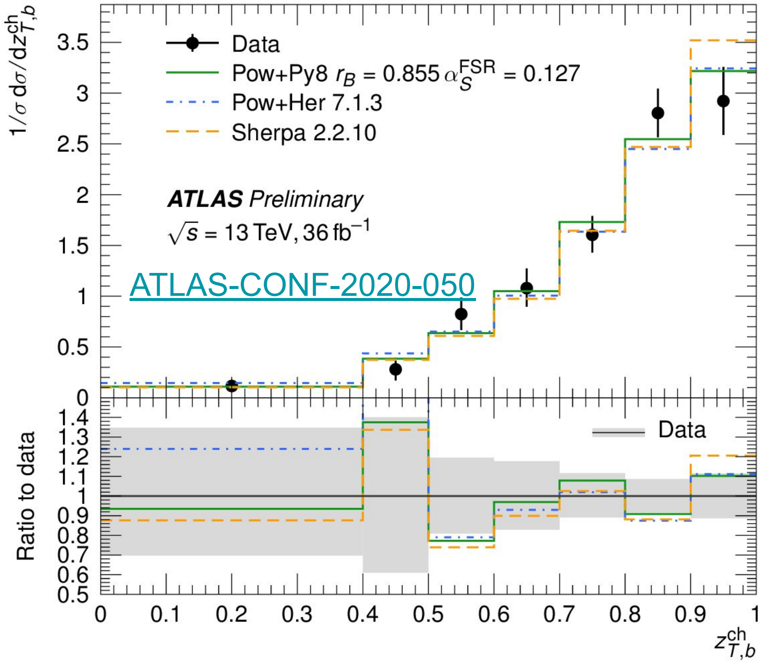




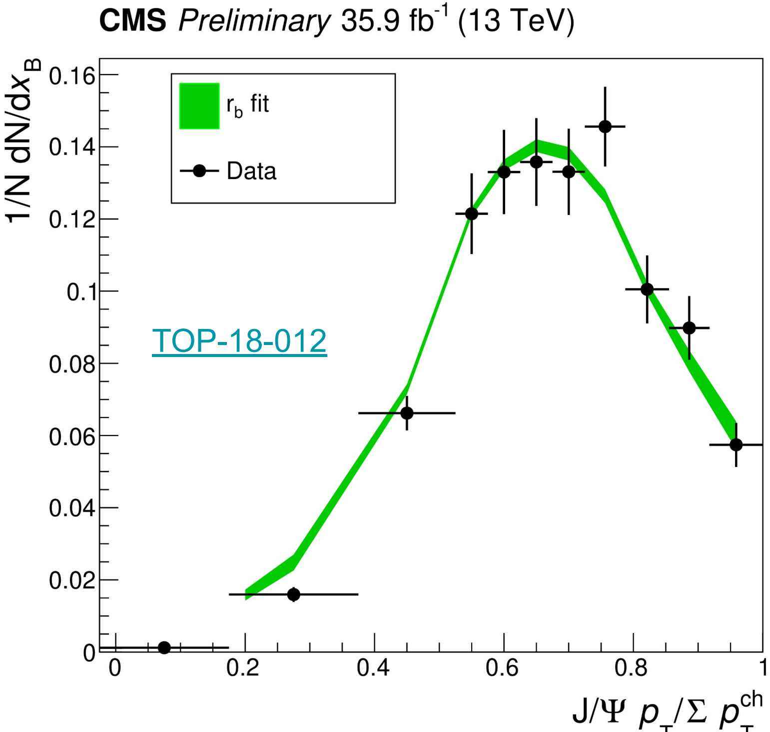
1606.07737

Heavy Flavour fragmentation

- Recent measurements sensitive to b-quark fragmentation in ttbar production from ATLAS/CMS
    - Sensitivity not huge, but can we use them to test LEP/LHC fragmentation universality?



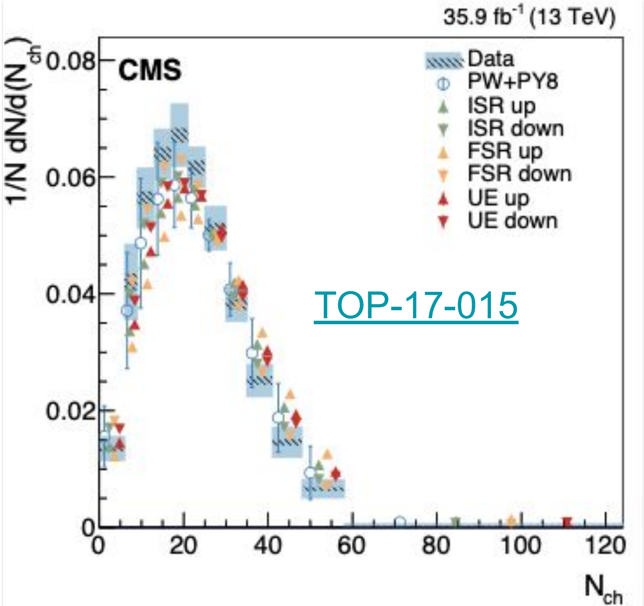
ATLAS-CONF-2020-050



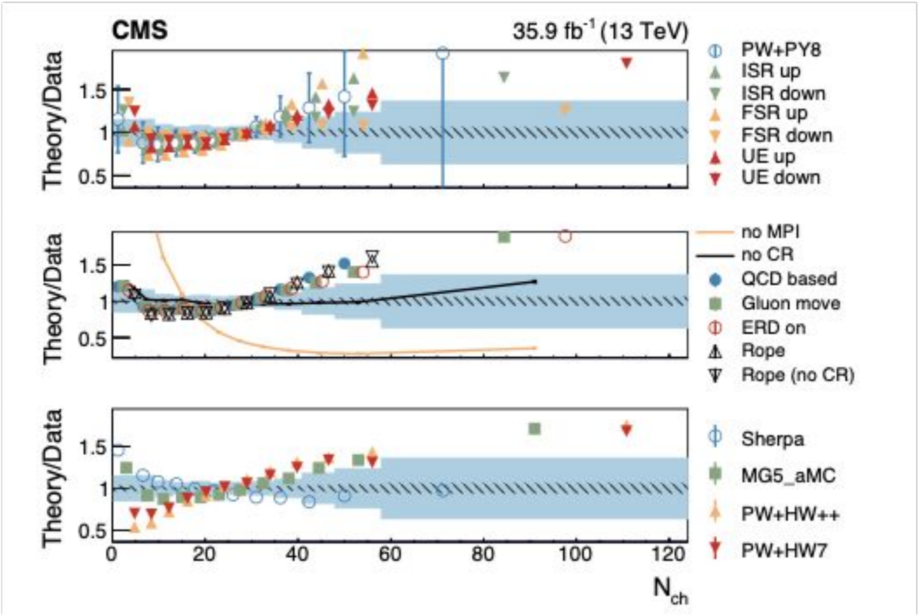
TOP-18-012

Color-reconnection effect

- The model of color-reconnection effects is a large source of uncertainty in top mass measurements
  - New models have now been implemented in Pythia8 and are being considered for uncertainties
  - Run2 direct mass measurement from CMS has a ~0.4 GeV uncertainty from CR
- Measurements of charged particles in ttbar exists, but sensitivity to CR parameters not very large



TOP-17-015



Collective effects in pp

- One of the surprising findings of the LHC is the presence of “collective effects” even in pp collisions
- These are not included in standard MC generators
- What is the possible impact for LHC pp physics
    - Description of pile-up
    - Particle composition of jets could affect detector response

Summary

- Event generators remain the essential (and unavoidable) tool for LHC analyses
- Successful analysis of Run1 and early Run2 thanks to the “NLO revolution”

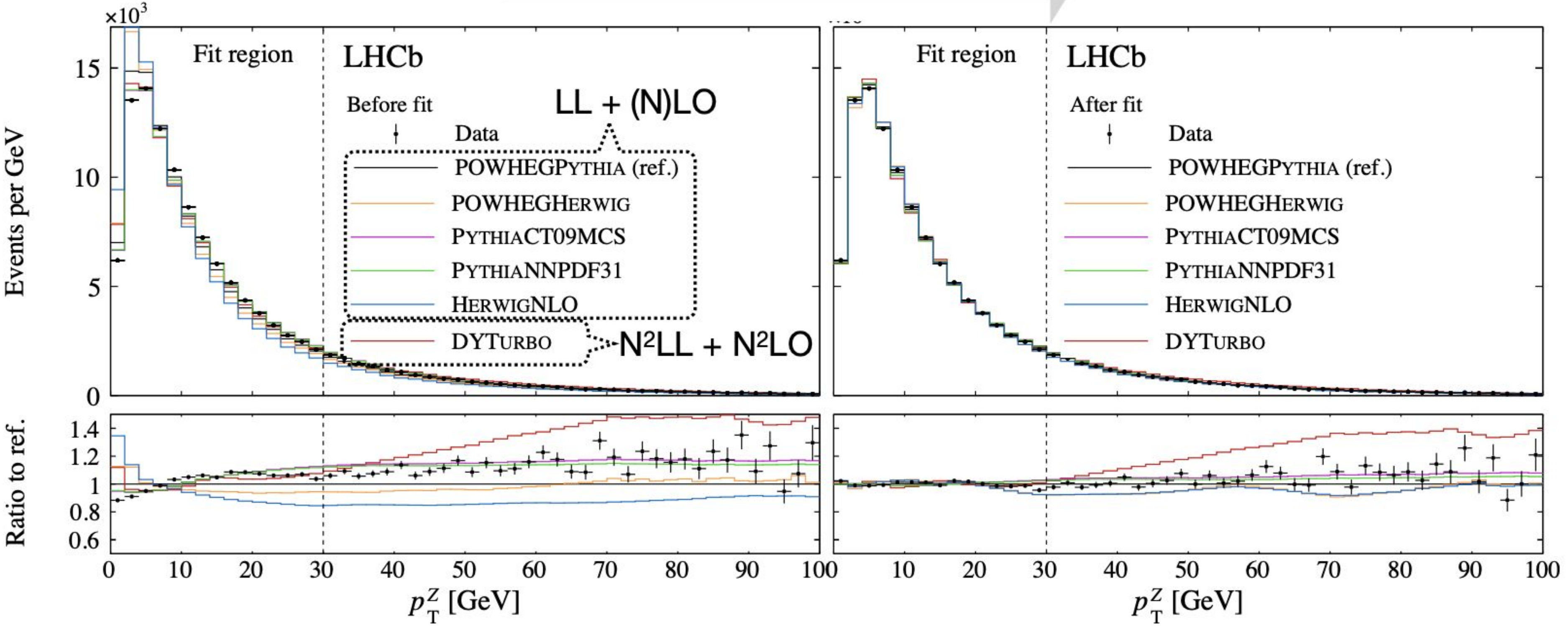
- New developments are now needed to achieve the Run3 and HL-LHC precision targets
    - Matching to **higher QCD orders** (NNLOPS, NNLO merging?)
    - Coherent inclusion of **higher EW orders** in QCD matching/merging and in the shower
    - **Higher accuracy parton showers** which can be matched to NNLO
    - A **better understanding of non-perturbative aspects** (including PDFs)
    - **Tuning with realistic uncertainties** of soft/non-perturbative parameters
- And all of the above without impacting too much the experiments computing budget

Thanks!
(and good luck!)

Back up

LHCb Z pT validation

Tuning of α_s and intrinsic k_T



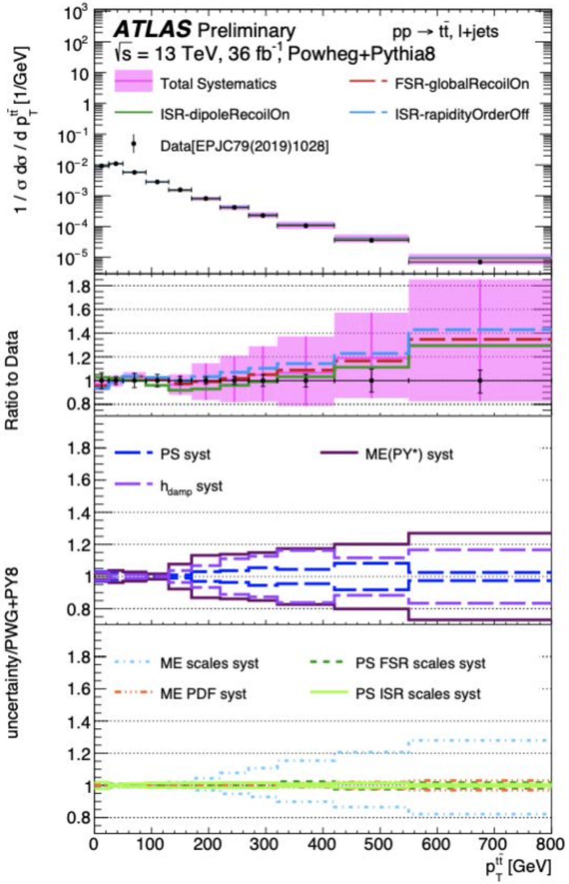
LL + (N)LO

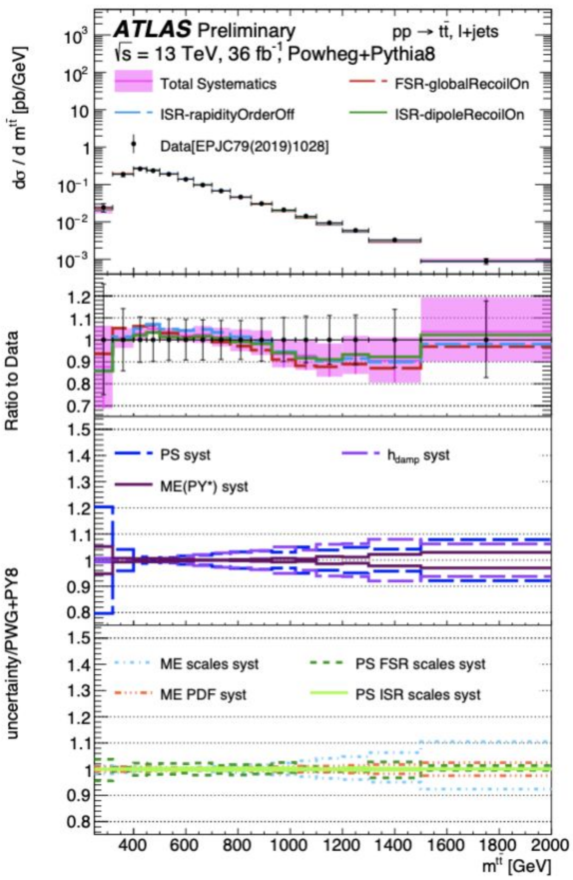
N²LL + N²LO

ATLAS/CMS ttbar uncertainties

Systematic	**ATLAS**	**CMS**
Nominal	PowhegPythia8	
PDFs	PDF4LHC recommendations	
NLO matching	Powheg vs MC@NLO	MC@NLO as cross-check but reweights top p_T to NNLO
Initial State Radiation	7-point variations of $\mu_R^{\text{ME}}, \mu_F^{\text{ME}}$ + independent variations of $h_{\text{damp}}, \mu_R^{\text{PS,ISR}}$	
Final State Radiation	Variations of $\mu_R^{\text{PS,FSR}}$	
Underlying Event	Tune variations (***A14***/***CP5***) + different CR models	
B-fragmentation	Variations of r_B parameter in Pythia8 (CMS also compares to Peterson fragmentation)	
Fragmentation/ Hadronisation	Pythia8 vs Herwig7	Pythia6 vs Herwig++ (only impact on jet response)
ttbar/Wt interference	DR vs DS in Powheg	

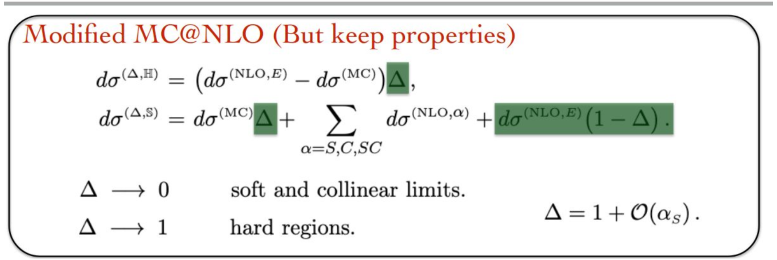
Pythia recoils in ttbar





Negative Weights

MC@NLO-Δ



Modified MC@NLO (But keep properties)

$$d\sigma^{(\Delta,\mathbb{H})} = \left(d\sigma^{(\mathrm{NLO},E)} - d\sigma^{(\mathrm{MC})}\right)\Delta,$$

$$d\sigma^{(\Delta,\mathbb{S})} = d\sigma^{(\mathrm{MC})}\Delta + \sum_{\alpha=S,C,SC} d\sigma^{(\mathrm{NLO},\alpha)} + d\sigma^{(\mathrm{NLO},E)}(1-\Delta)\,.$$

$\Delta \longrightarrow 0$ soft and collinear limits.

$\Delta \longrightarrow 1$ hard regions.

$$\Delta = 1 + \mathcal{O}(\alpha_S)\,.$$

- Negative weights strongly reduce statistical power
- For weighted events w_i, effective events N_{eff} for fraction of negative weights f: $N_{eff} = \frac{(\sum_i w_i)^2}{\sum w_i^2} = N(1-2f)^2$
- for 35% negative weights (common at for high jet-mulitplicity/ high pt) $\Longrightarrow$ 9% effective events compared to $w_i = 1$

- Analysis of a parton shower in the vicinity of the soft and collinear regions allows to formulate a modified MC@NLO-matching prescription that reduces the number of negative-weight events
- CMS is testing this modified MC@NLO scheme
  - involves adjusting Pythia interface in CMSSW

Sherpa 3.0 will have internal features to reduce negative weights as explained here

A Positive Resampler for Monte Carlo Events with Negative Weights

- Turns negative event weights into positive ones
- Preserves distributions
- Applied successfully to complicated process: W production at NLO + PS with multijet merging
  - Already discussed and plan to explore in CMS