UNIVERSITY OF SUSSEX

THE ROYAL SOCIETY

MC generators: Experimental bottlenecks and future needs

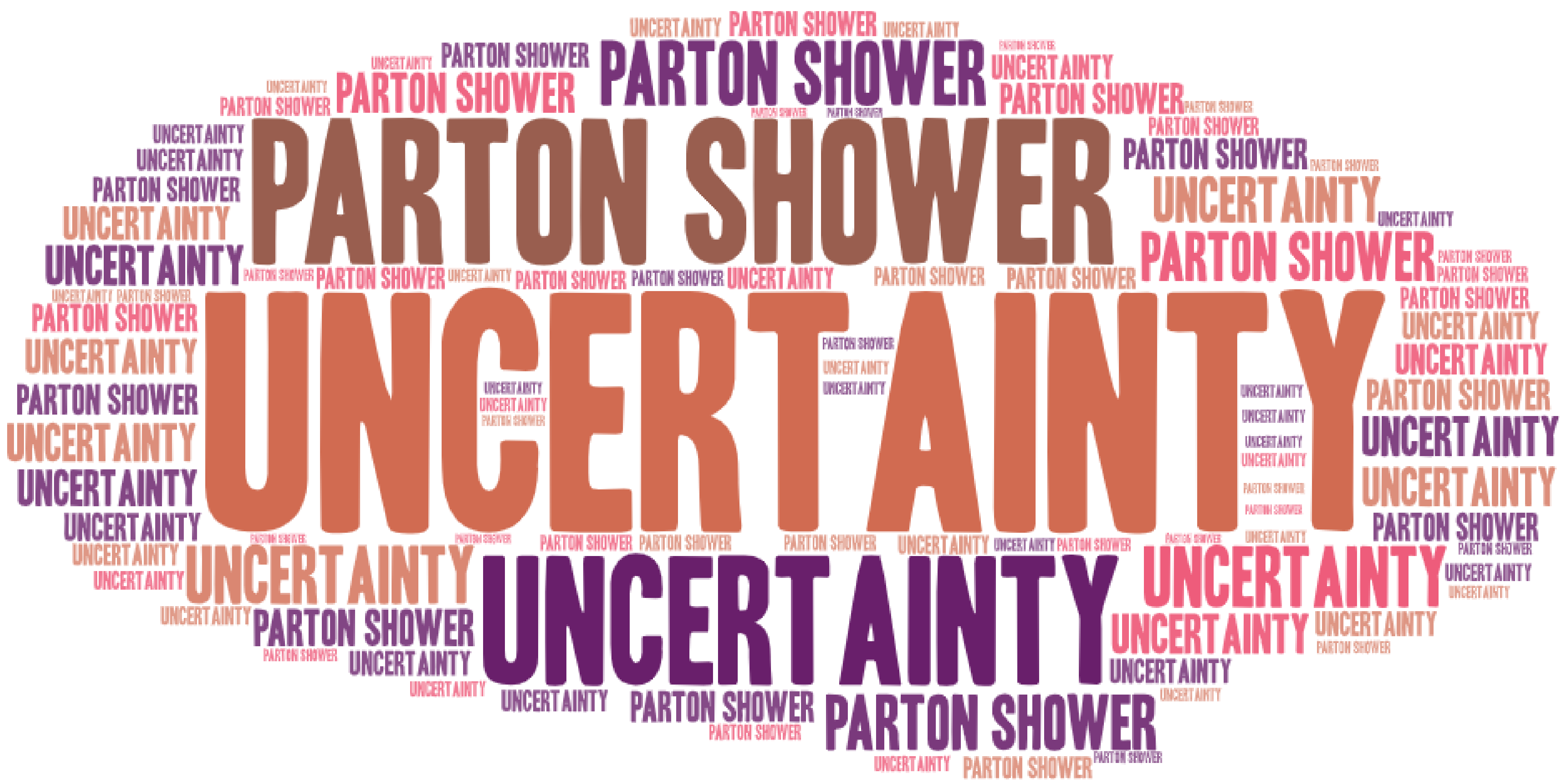
Physics at TeV Colliders 2021

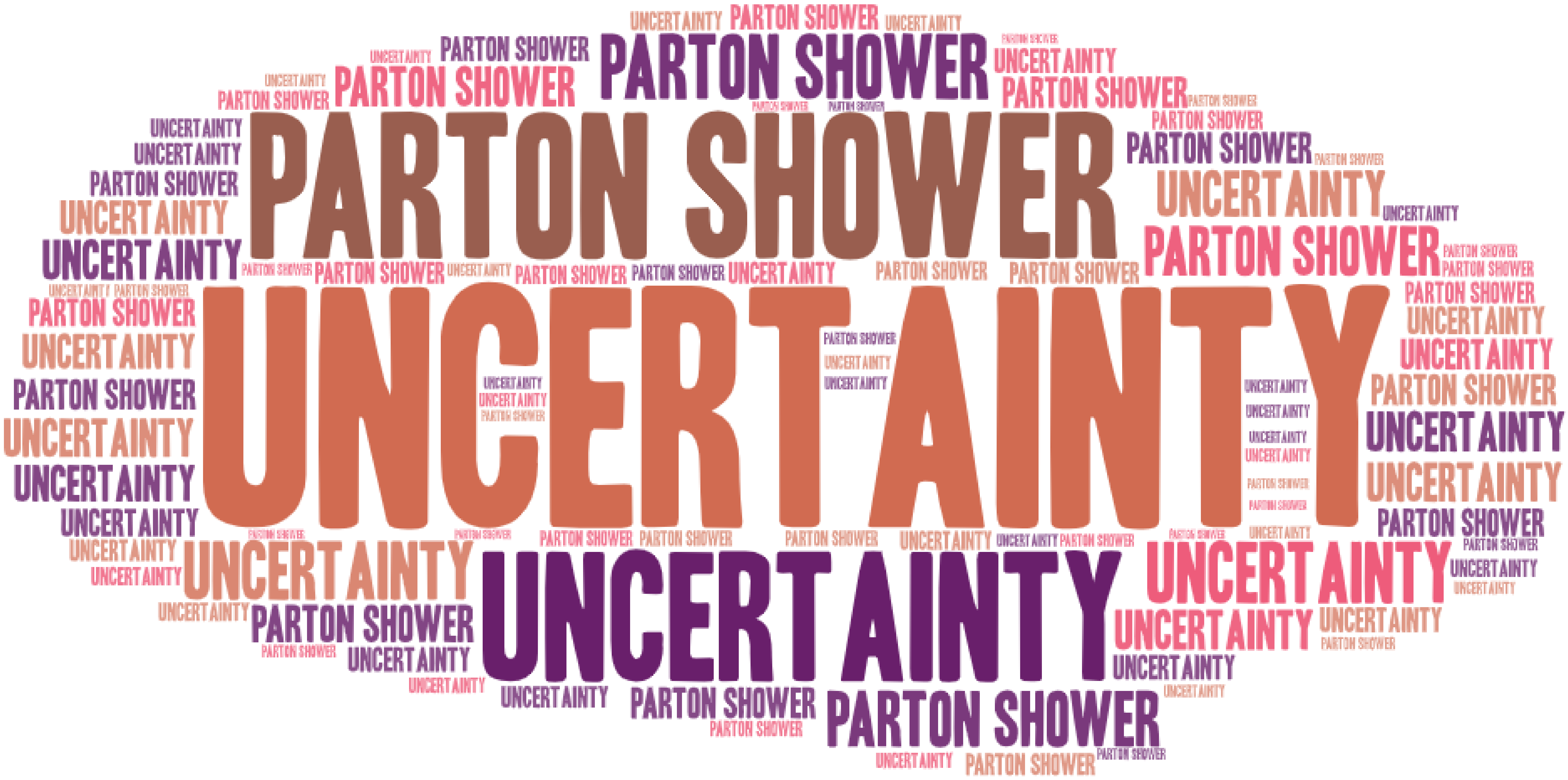
16/06/2021

@JoshMcFayden
cern.ch/mcfayden

Topics of interest

UNIVERSITY OF SUSSEX

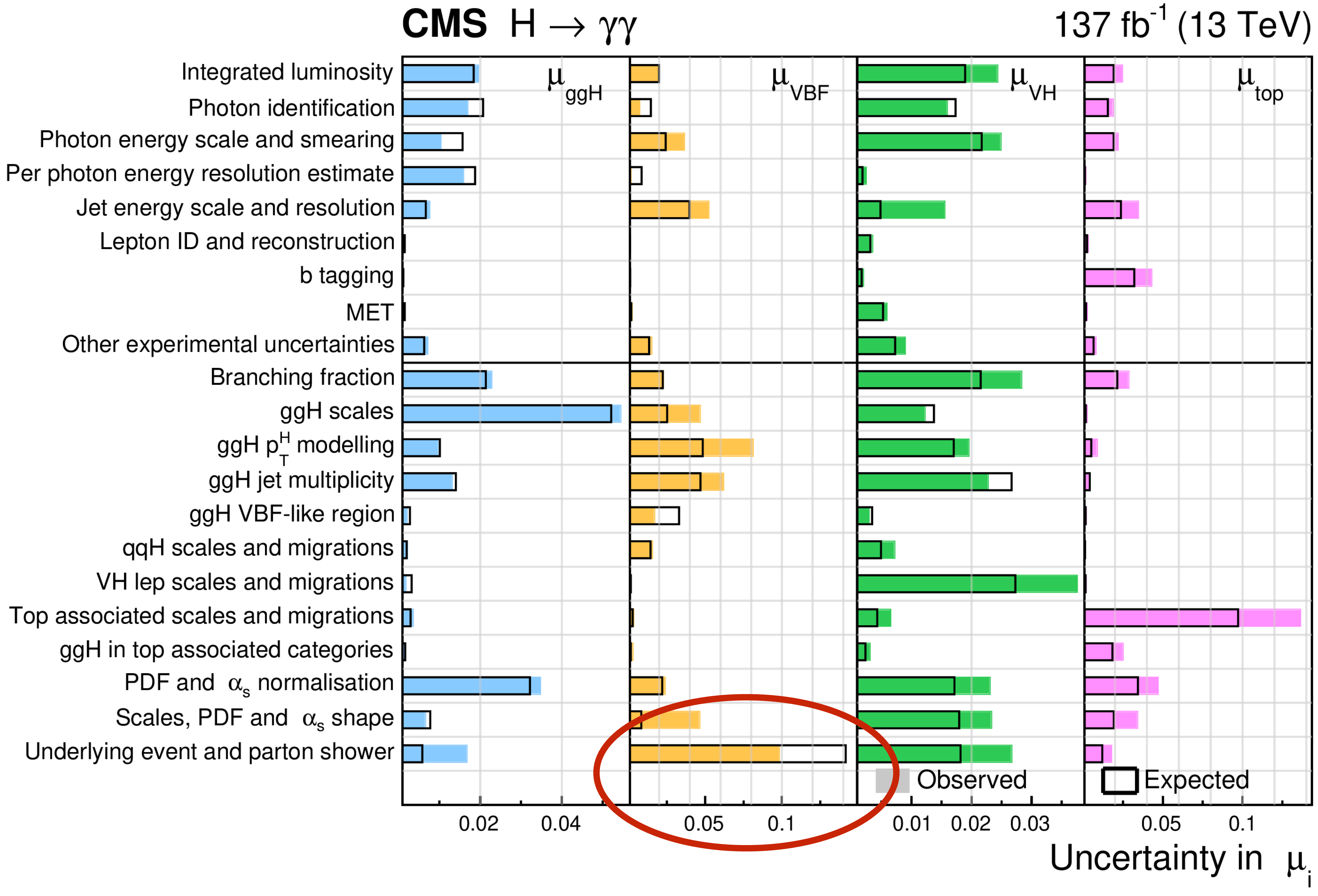



UNCERTAINTY
PARTON SHOWER

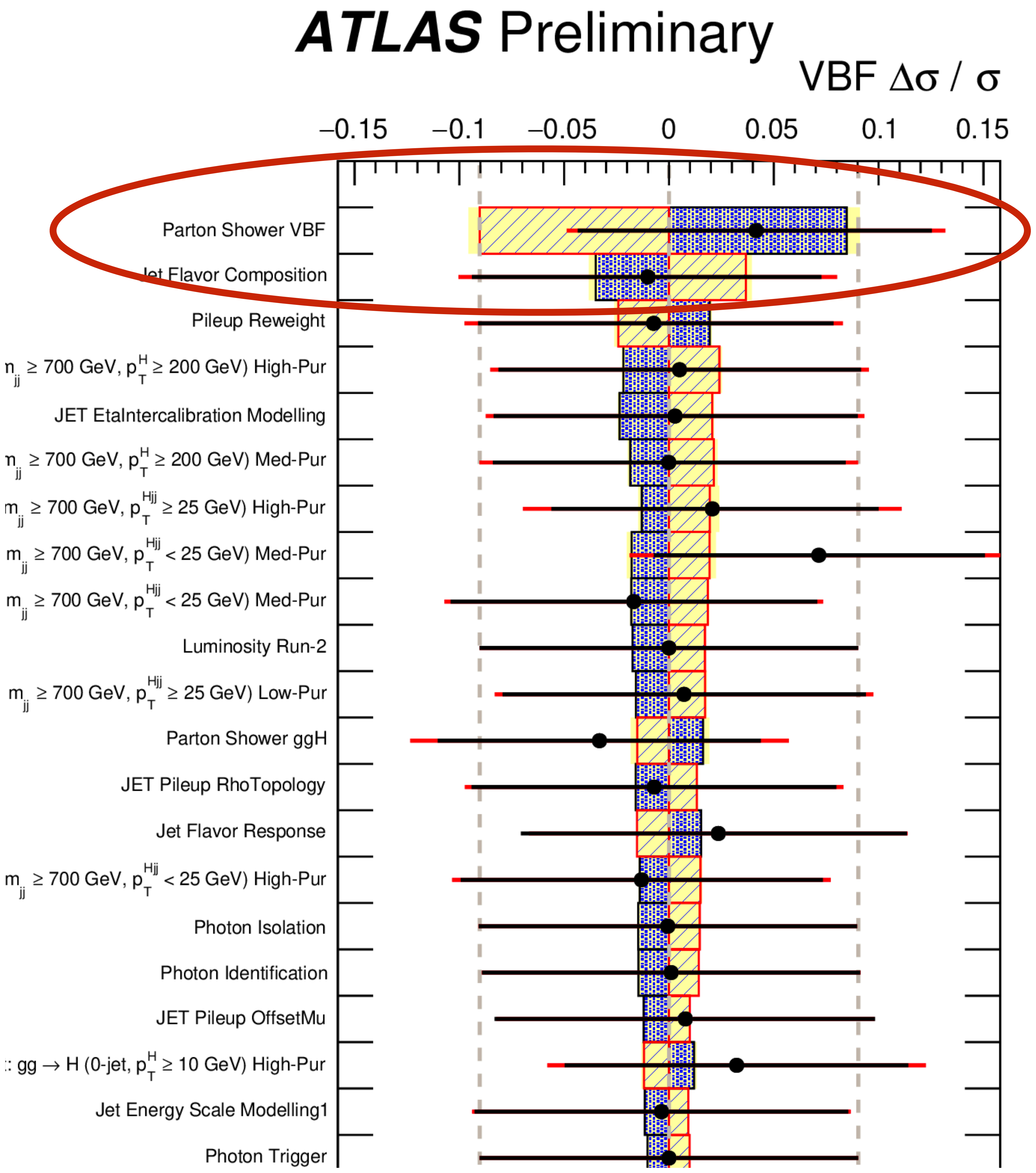
PS/Had/UE | H→γγ VBF

ATLAS-CONF-2020-026

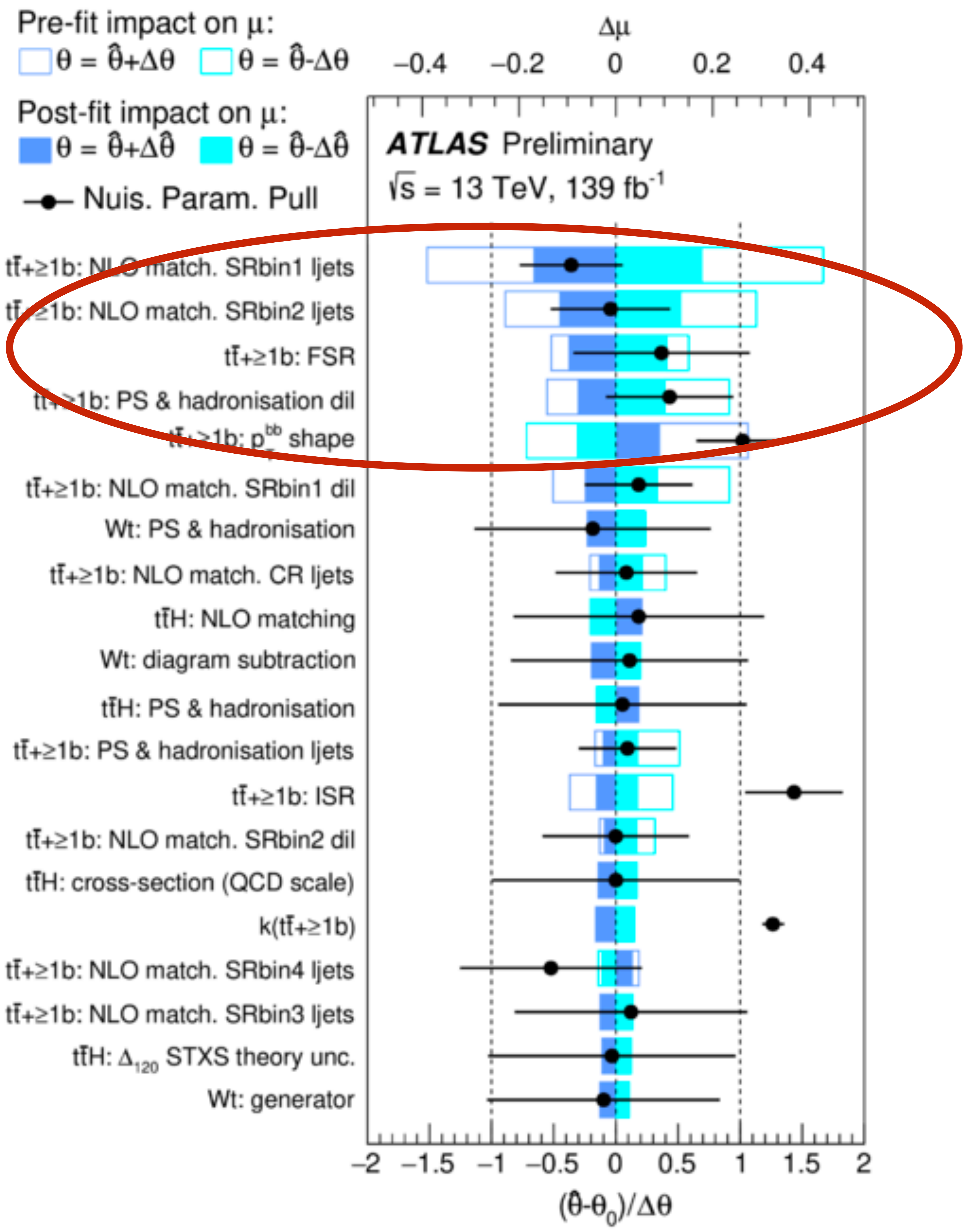
UNIVERSITY OF SUSSEX



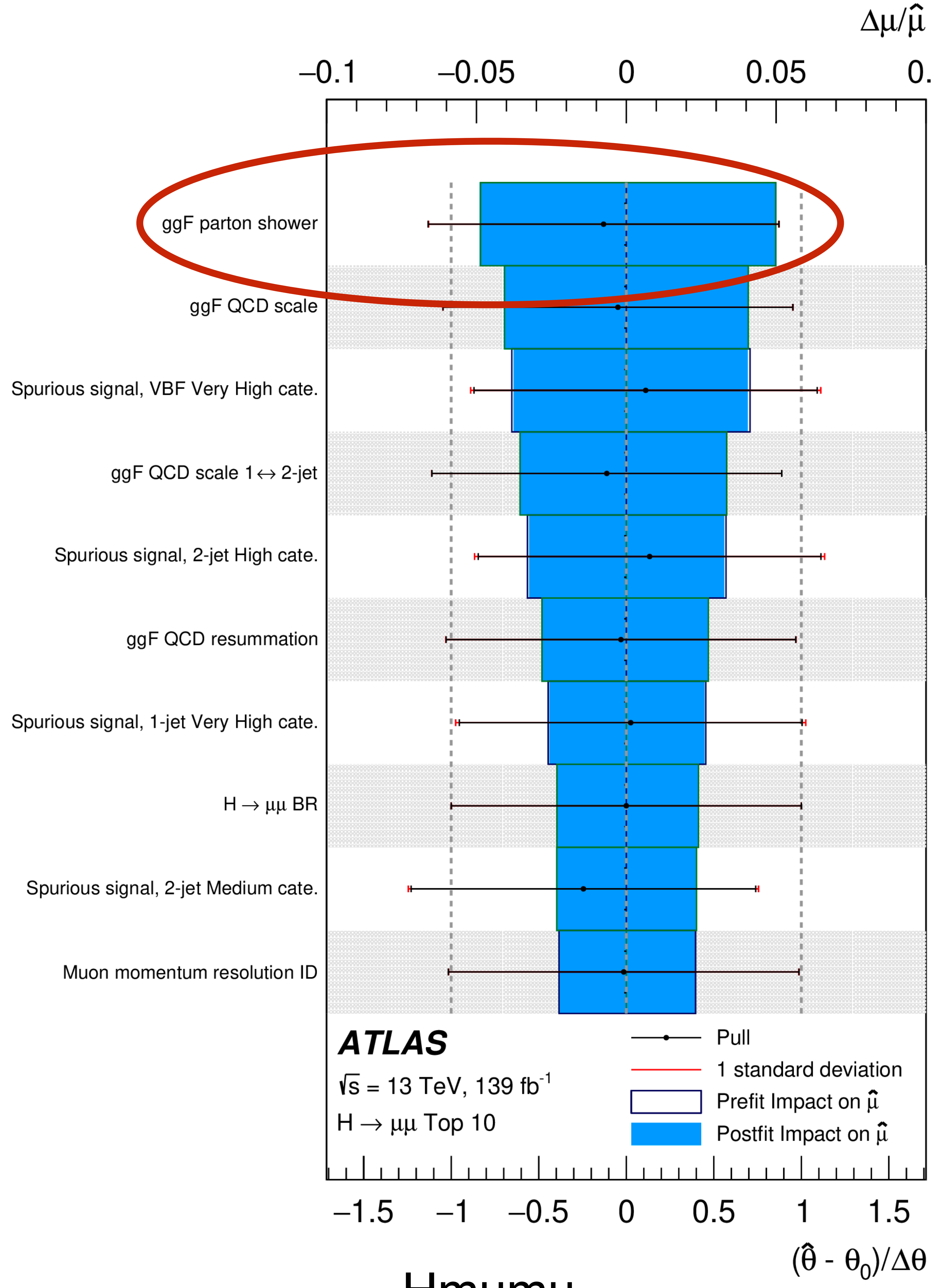
HIG-19-015



PS/Had/UE | ttH & H→μμ



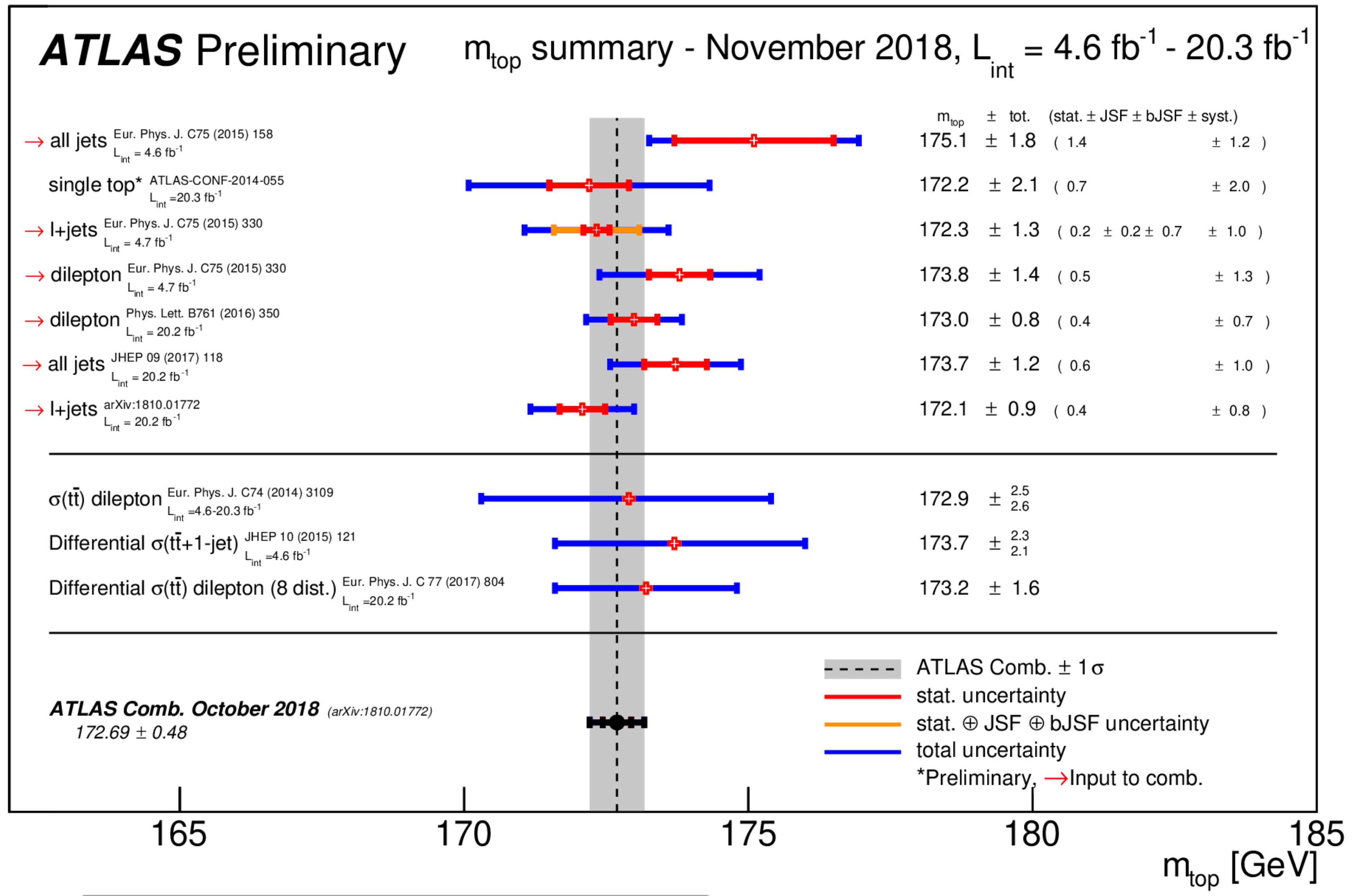
ttH



Hmumu

PS/Had/UE | Top mass

UNIVERSITY OF SUSSEX



Mass scheme	m_t^{pole} [GeV]	$m_t(m_t)$ [GeV]
Value	**171.1**	**162.9**
Statistical uncertainty	[illegible]	**0.5**
Simulation uncertainties		
Shower and hadronisation	0.4	0.3
Colour reconnection	0.4	[illegible]
Underlying event	0.3	0.2
Signal Monte Carlo generator	0.2	0.2
Proton PDF	0.2	0.2
Initial- and final-state radiation	0.2	0.2
Monte Carlo statistics	0.2	0.2
Background	<0.1	<0.1
Detector response uncertainties		
Jet energy scale (including b-jets)	0.4	0.4
Jet energy resolution	0.2	0.2
Missing transverse momentum	0.1	0.1
b-tagging efficiency and mistag	0.1	0.1
Jet reconstruction efficiency	<0.1	<0.1
Lepton	<0.1	<0.1
Method uncertainties		
Unfolding modelling	0.2	0.2
Fit parameterisation	0.2	0.2
Total experimental systematic	**0.9**	**1.0**
Scale variations	(+0.6, −0.2)	(+2.1, −1.2)
Theory PDF⊕α_s	0.2	0.4
Total theory uncertainty	**(+0.7, −0.3)**	**(+2.1, −1.2)**
Total uncertainty	**(+1.2, −1.1)**	**(+2.3, −1.6)**

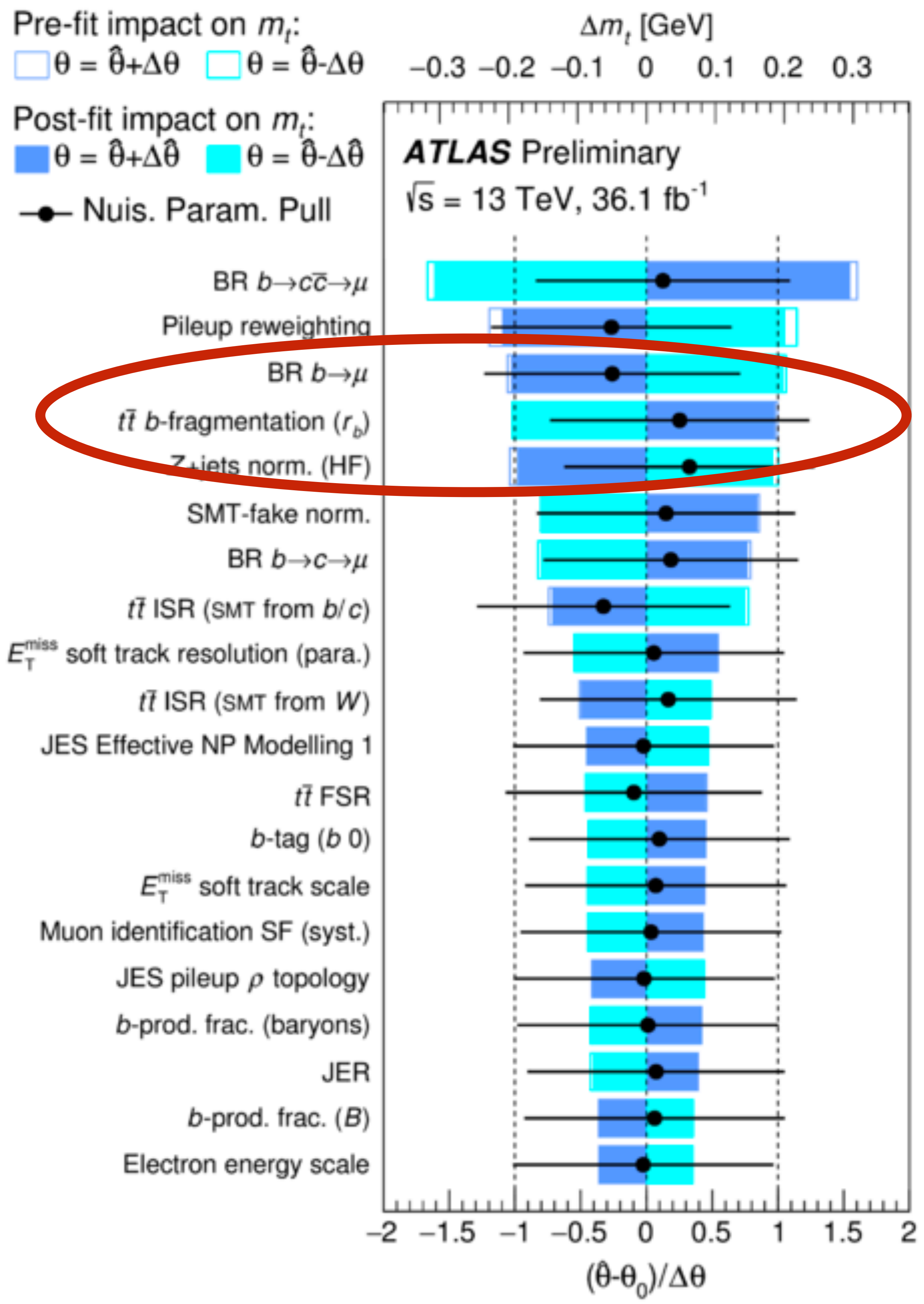
TOPQ-2017-09

TOPQ-2017-03

	$\sqrt{s}$ = 7 TeV	$\sqrt{s}$ = 8 TeV	
Event selection	Standard	Standard	BDT
m_{top} result [GeV]	172.33	171.90	172.08
Statistics	0.75	0.38	0.39
– *Stat. comp.* (m_{top})	0.23	0.12	0.11
– *Stat. comp.* (JSF)	0.25	0.11	0.11
– *Stat. comp.* (bJSF)	0.67	0.34	0.35
Method	0.11 ± 0.10	0.04 ± 0.11	0.13 ± 0.11
Signal Monte Carlo generator	0.22 ± 0.21	0.50 ± 0.17	0.16 ± 0.17
Hadronization	0.18 ± 0.12	0.05 ± 0.10	0.15 ± 0.10
Initial- and final-state QCD radiation	0.32 ± 0.06	0.28 ± 0.11	0.08 ± 0.11
Underlying event	0.15 ± 0.07	0.08 ± 0.15	0.08 ± 0.15
Colour reconnection	0.11 ± 0.07	0.37 ± 0.15	0.19 ± 0.15
Parton distribution function	0.25 ± 0.00	0.08 ± 0.00	0.09 ± 0.00
Background normalization	0.10 ± 0.00	0.04 ± 0.00	0.08 ± 0.00
W+jets shape	0.29 ± 0.00	0.05 ± 0.00	0.11 ± 0.00
Fake leptons shape	0.05 ± 0.00	0	0
Jet energy scale	0.58 ± 0.11	0.63 ± 0.02	0.54 ± 0.02
Relative b-to-light-jet energy scale	0.06 ± 0.03	0.05 ± 0.01	0.03 ± 0.01
Jet energy resolution	0.22 ± 0.11	0.23 ± 0.03	0.20 ± 0.04
Jet reconstruction efficiency	0.12 ± 0.00	0.04 ± 0.01	0.02 ± 0.01
Jet vertex fraction	0.01 ± 0.00	0.13 ± 0.01	0.09 ± 0.01
b-tagging	0.50 ± 0.00	0.37 ± 0.00	0.38 ± 0.00
Leptons	0.04 ± 0.00	0.16 ± 0.01	0.16 ± 0.01
Missing transverse momentum	0.15 ± 0.04	0.08 ± 0.01	0.05 ± 0.01
Pile-up	0.02 ± 0.01	0.14 ± 0.01	0.15 ± 0.01
Total systematic uncertainty	1.04 ± 0.08	1.07 ± 0.10	0.82 ± 0.06
Total	1.28 ± 0.08	1.13 ± 0.10	0.91 ± 0.06

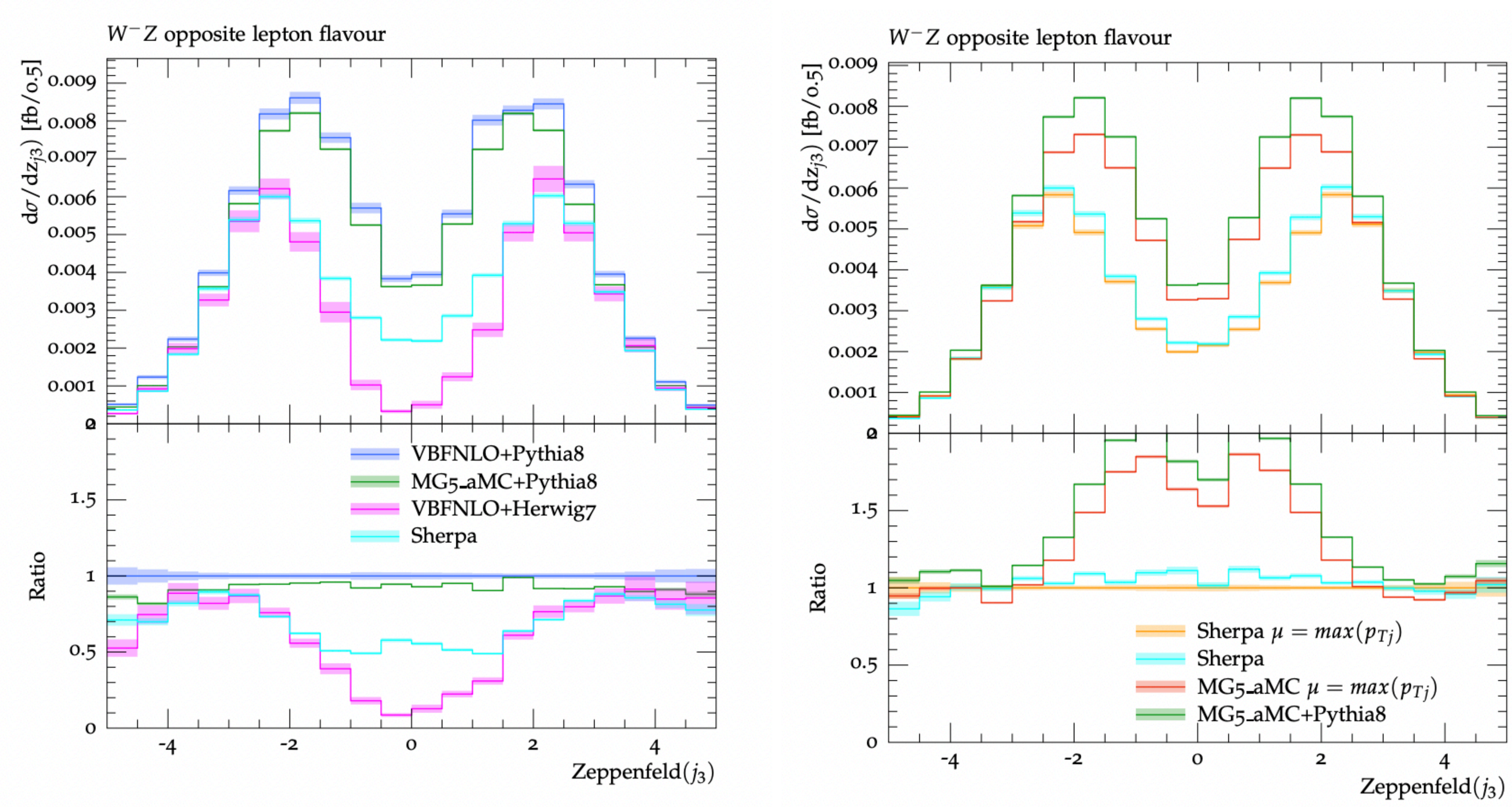
	2D		1D	hybrid	
	δm_t^{2D} [GeV]	δJSF^{2D} [%]	δm_t^{1D} [GeV]	δm_t^{hyb} [GeV]	δJSF^{hyb} [%]
Experimental uncertainties					
Method calibration	0.03	0.0	0.03	0.03	0.0
JEC (quad. sum)	0.12	0.2	0.82	0.17	0.3
– Intercalibration	−0.01	0.0	+0.16	+0.04	+0.1
– MPFInSitu	−0.01	0.0	+0.23	+0.07	+0.1
– Uncorrelated	−0.12	−0.2	+0.77	+0.15	+0.3
Jet energy resolution	−0.18	+0.3	+0.09	−0.10	+0.2
b tagging	0.03	0.0	0.01	0.02	0.0
Pileup	−0.07	+0.1	+0.02	−0.05	+0.1
All-jets background	0.01	0.0	0.00	0.01	0.0
All-jets trigger	+0.01	0.0	0.00	+0.01	0.0
ℓ+jets Background	−0.02	0.0	+0.01	−0.01	0.0
ℓ+jets Trigger	0.00	0.0	0.00	0.00	0.0
Lepton isolation	0.00	0.0	0.00	0.00	0.0
Lepton identification	0.00	0.0	0.00	0.00	0.0
Modeling uncertainties					
JEC flavor (linear sum)	−0.39	+0.1	−0.31	−0.37	+0.1
– light quarks (uds)	+0.11	−0.1	−0.01	+0.07	−0.1
– charm	+0.03	0.0	−0.01	+0.02	0.0
– bottom	−0.31	0.0	−0.31	−0.31	0.0
– gluon	−0.22	+0.3	+0.02	−0.15	+0.2
b jet modeling (quad. sum)	0.08	0.1	0.04	0.06	0.1
– b frag. Bowler–Lund	−0.06	+0.1	−0.01	−0.05	0.0
– b frag. Peterson	−0.03	0.0	0.00	−0.02	0.0
– semileptonic b hadron decays	−0.04	0.0	−0.04	−0.04	0.0
PDF	0.01	0.0	0.01	0.01	0.0
Ren. and fact. scales	0.01	0.0	0.02	0.01	0.0
ME/PS matching	−0.10 ± 0.08	+0.1	+0.02 ± 0.05	+0.07 ± 0.07	+0.1
ME generator	+0.16 ± 0.21	+0.2	+0.32 ± 0.13	+0.21 ± 0.18	+0.1
ISR PS scale	+0.07 ± 0.08	+0.1	+0.10 ± 0.05	+0.07 ± 0.07	0.1
FSR PS scale	[illegible]	[illegible]	[illegible]	+0.12 ± 0.06	−0.3
Top quark p_T	+0.01	−0.1	−0.06	−0.01	[illegible]
Underlying event	−0.06 ± 0.07	+0.1	+0.00 ± 0.05	−0.04 ± 0.06	+0.1
Early resonance decays	−0.20 ± 0.08	+0.7	+0.42 ± 0.05	−0.01 ± 0.07	+0.5
CR modeling (max. shift)	+0.37 ± 0.09	−0.2	+0.22 ± 0.06	+0.33 ± 0.07	−0.1
– "gluon move" (ERD on)	+0.37 ± 0.09	−0.2	+0.22 ± 0.06	+0.33 ± 0.07	−0.1
– "QCD inspired" (ERD on)	−0.11 ± 0.09	−0.1	−0.21 ± 0.06	[illegible]	[illegible]
Total systematic	[illegible]	[illegible]	[illegible]	0.61	0.7
Statistical (expected)	0.08	0.1	0.05	0.07	0.1
Total (expected)	0.72	1.0	1.08	0.61	0.7

TOP-17-008



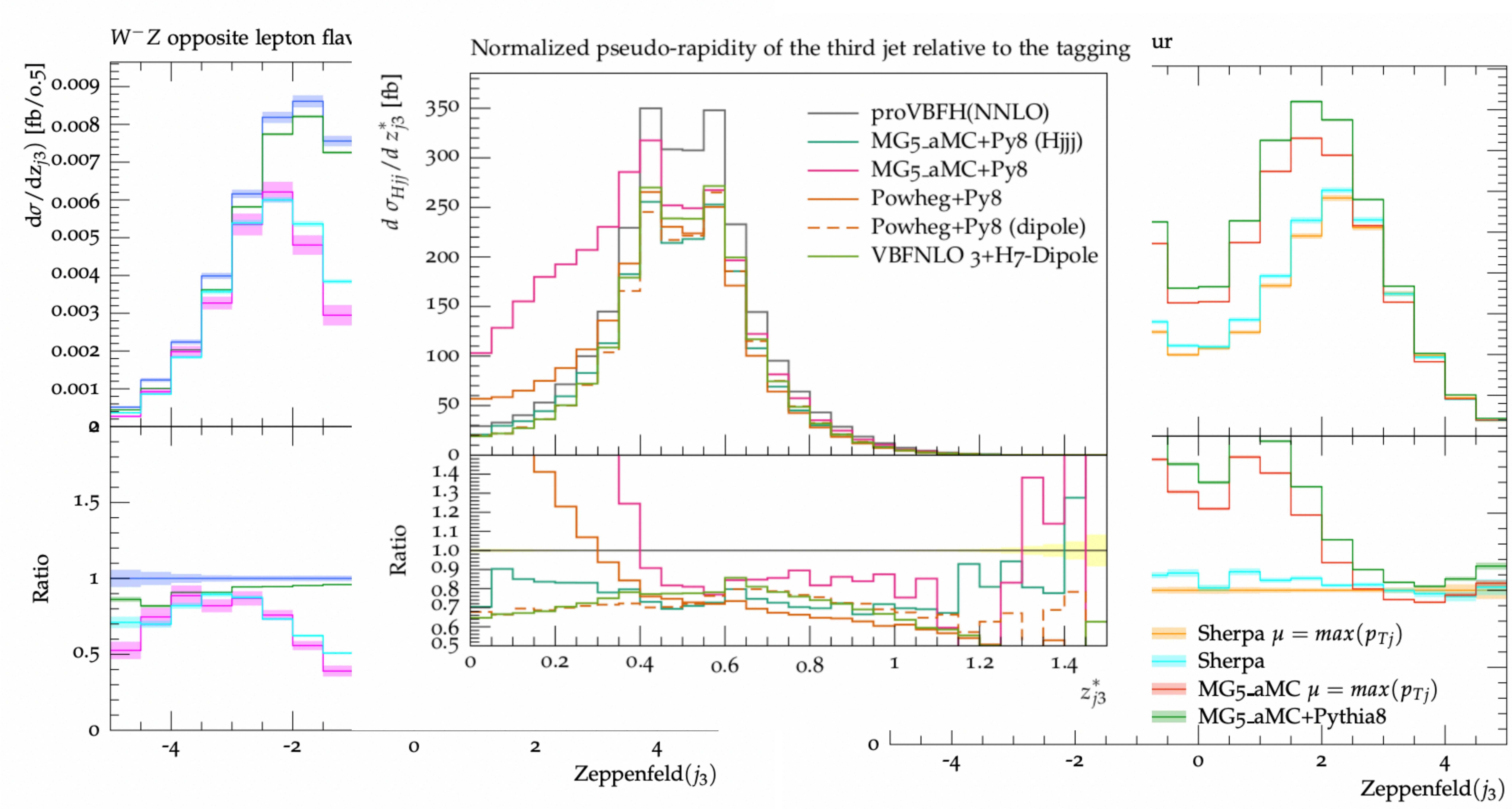
top mass (soft lepton)

PS/Had/UE | VBF modelling



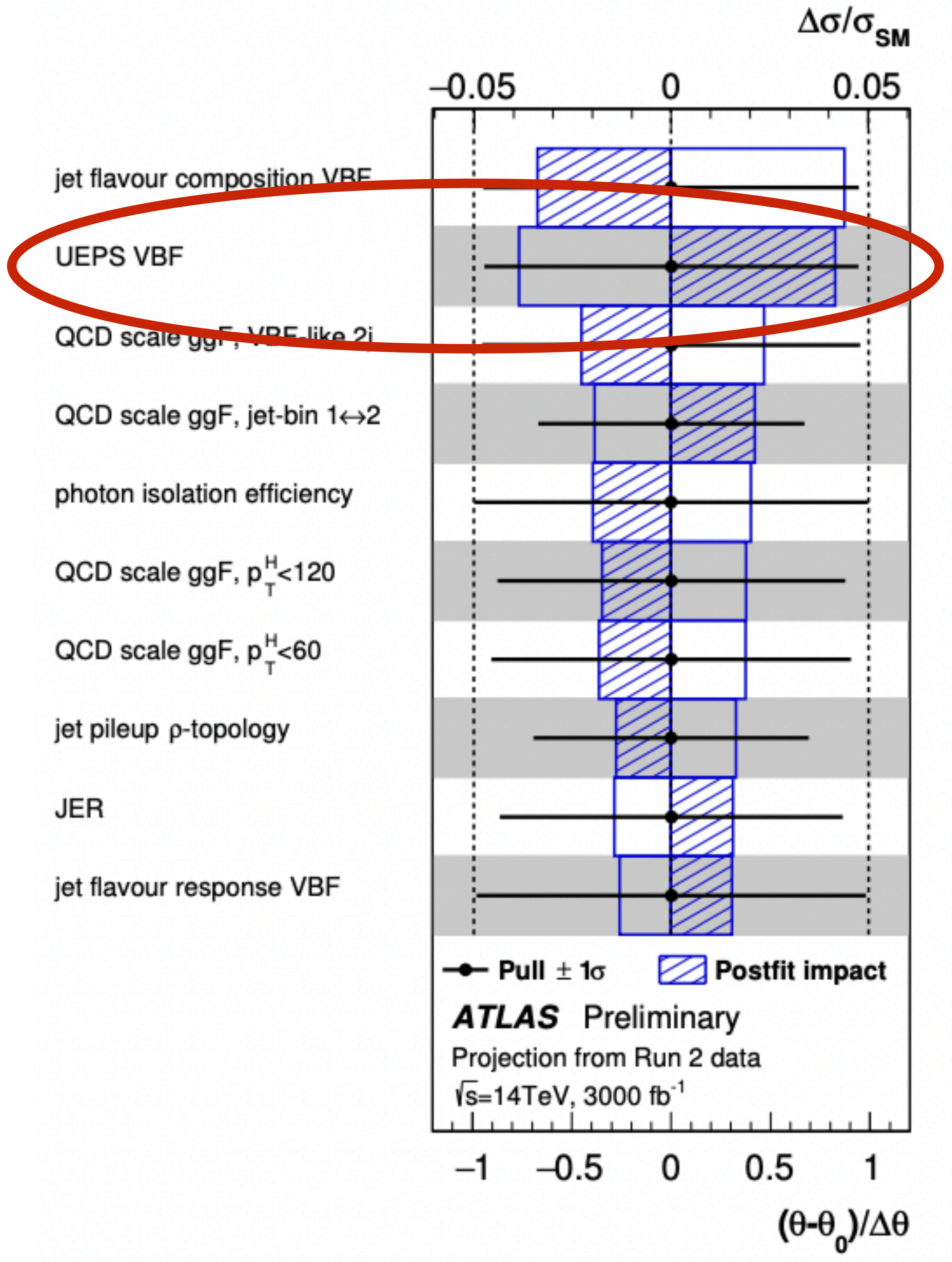
LH2017 Proceedings

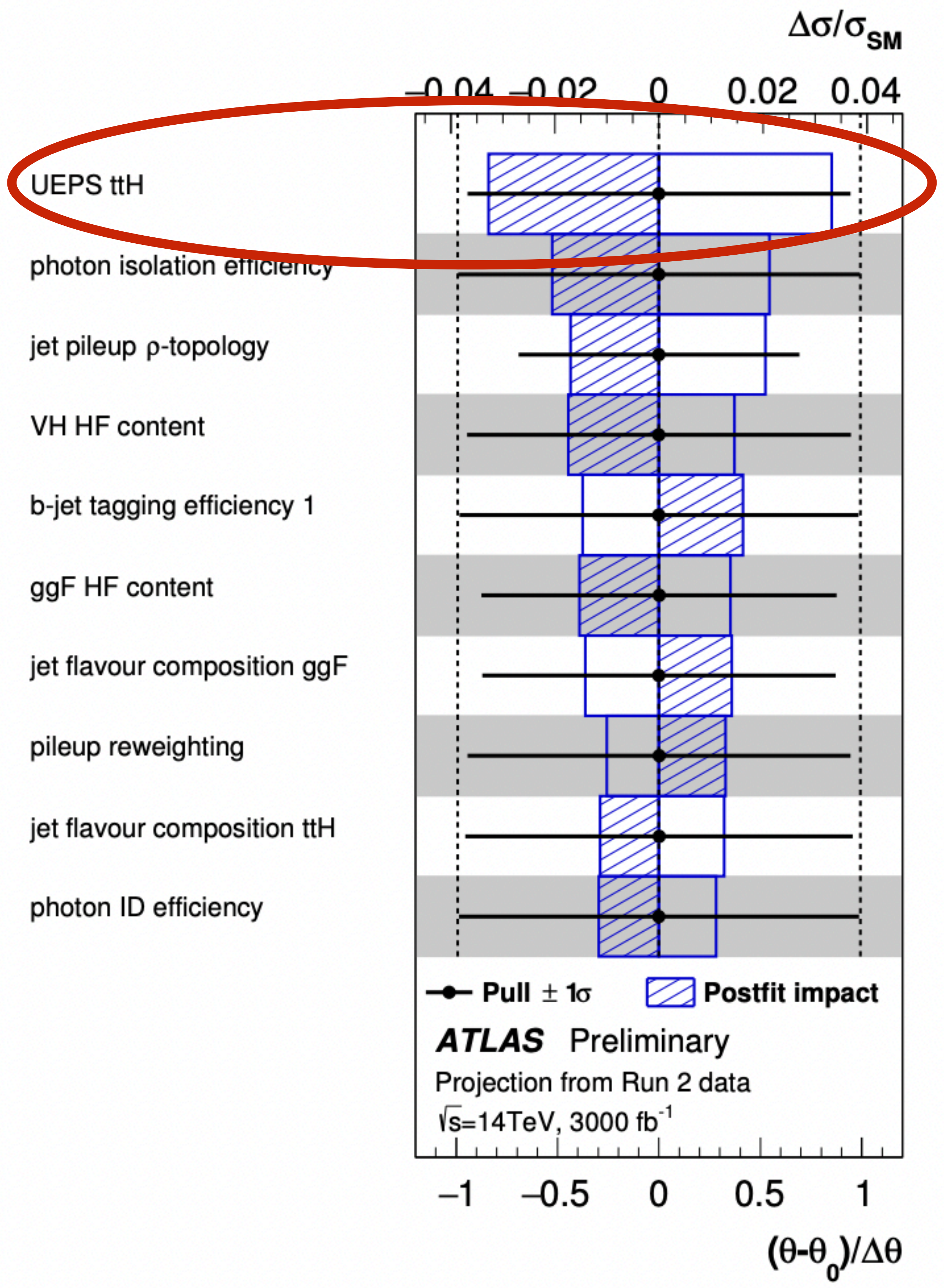
PS/Had/UE | VBF modelling

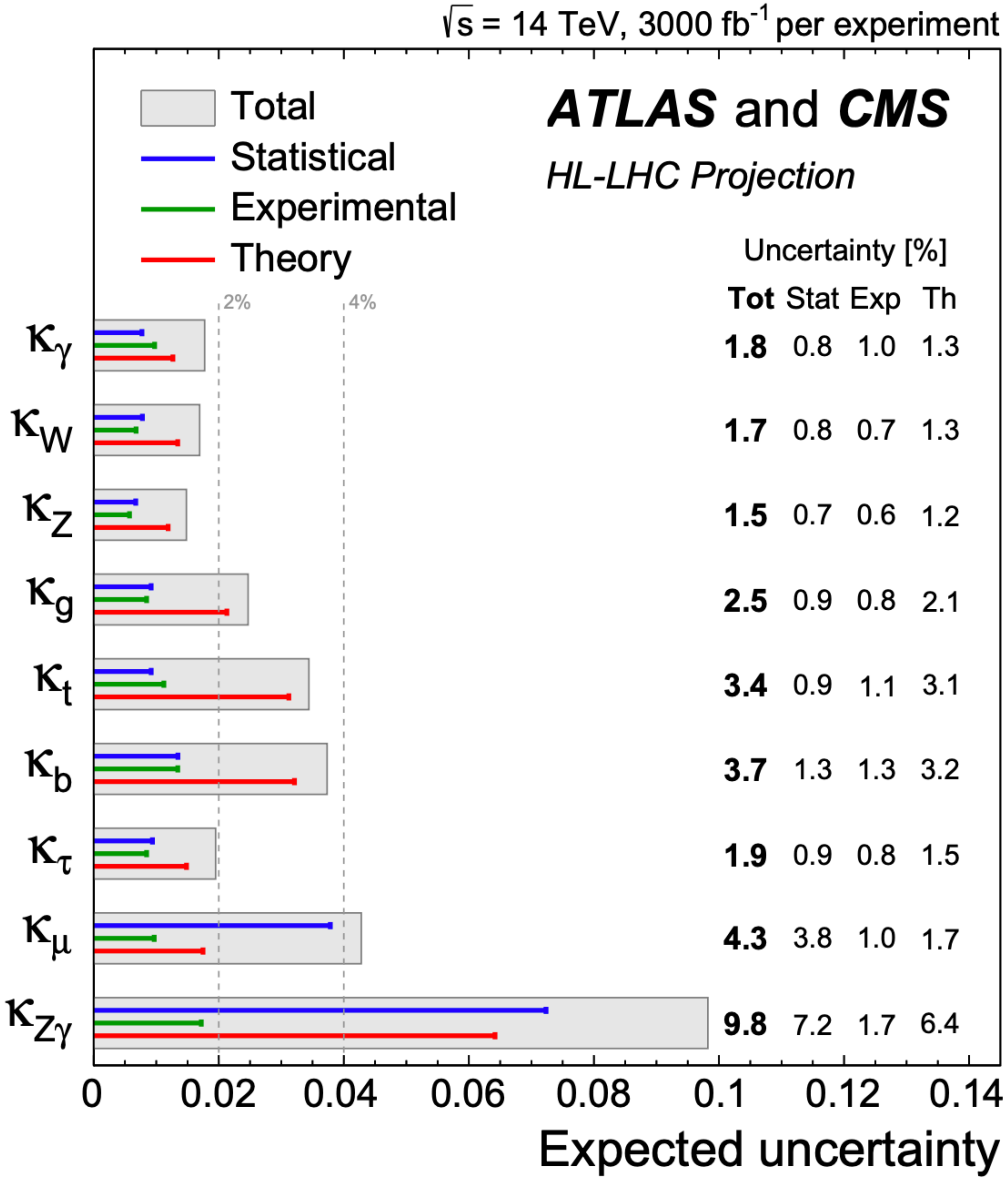


LH2017 Proceedings

PS/Had/UE | HL-LHC







What to do?

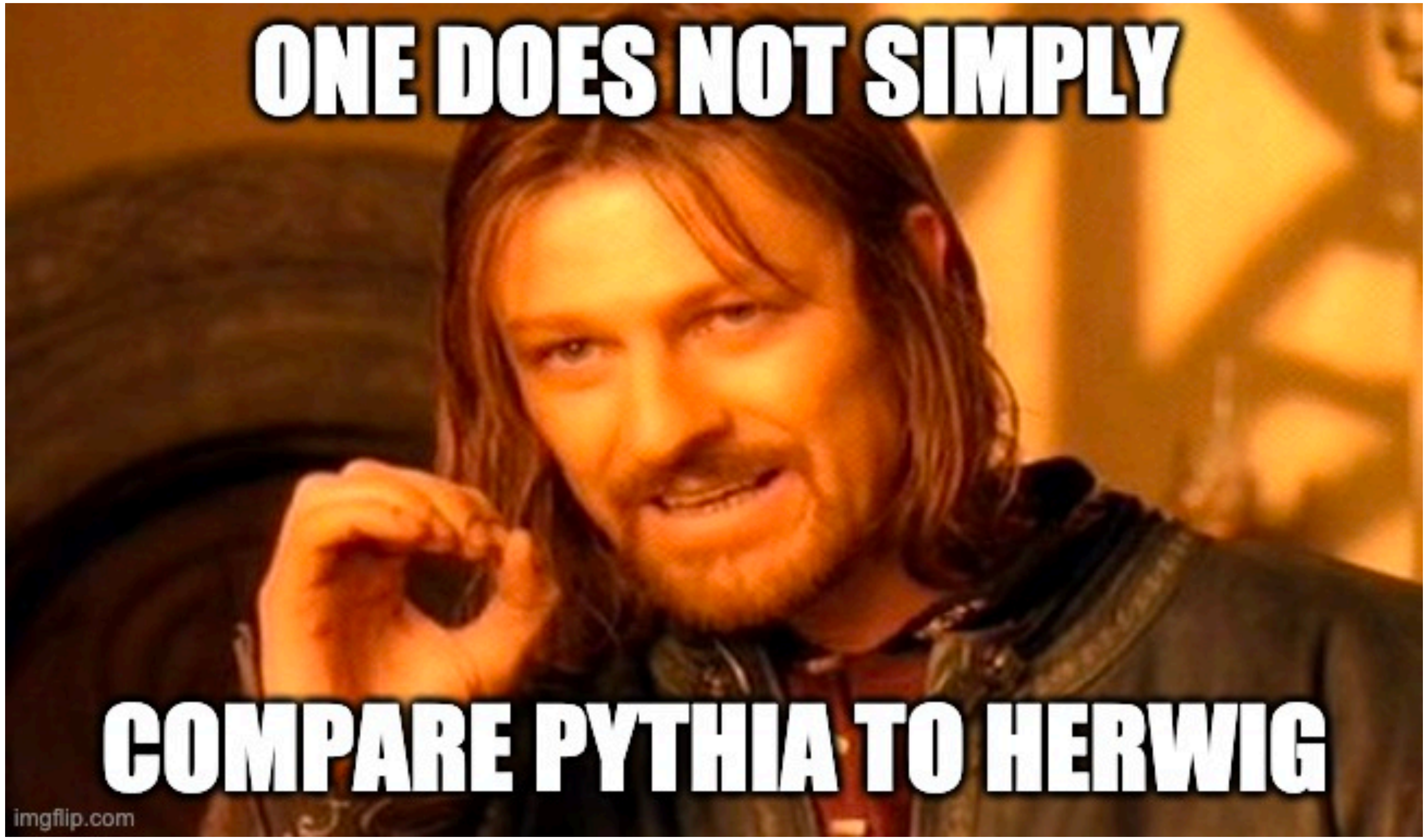
- *"Why are these uncertainties dominant in so many LHC analyses; can this situation be improved?"*

  \- Joey Huston

- Well let's start at the beginning
  - Do we understand where the large uncertainties are coming from?

What to do?

- Comparing Pythia to Herwig
  - We know it's bad from a physics point of view
  - We know it's bad from a statistical point of view
  - We have strong indications that it's over-estimating our uncertainties



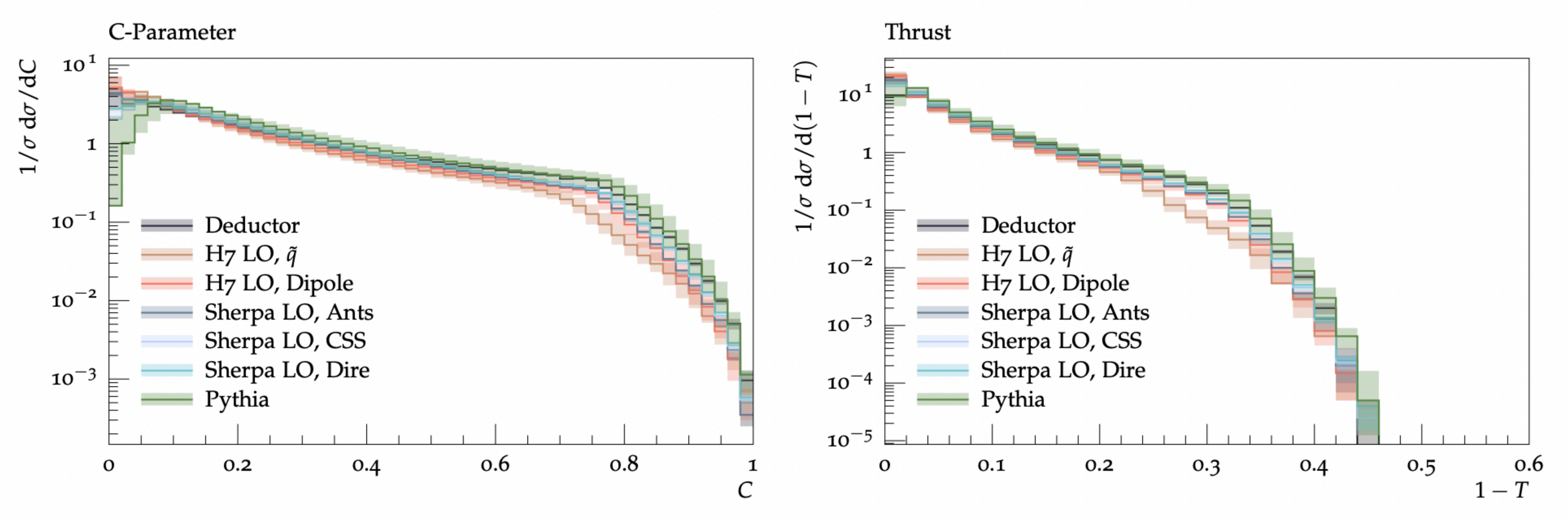
- So why do we continue?!
  - Because it's easy and everything else is hard..?
  - Because the tools are complicated?
  - Because we're attached to our tunes?

Factorisation

- Of course what we really want to do is factorise the different physics/ sources of uncertainty
  - Shower evolution
  - Hadronisation
  - Underlying event
  - Color reconnection
  - Details of matching setups
- Do we have the tools to do this?

Factorisation | Shower algorithm

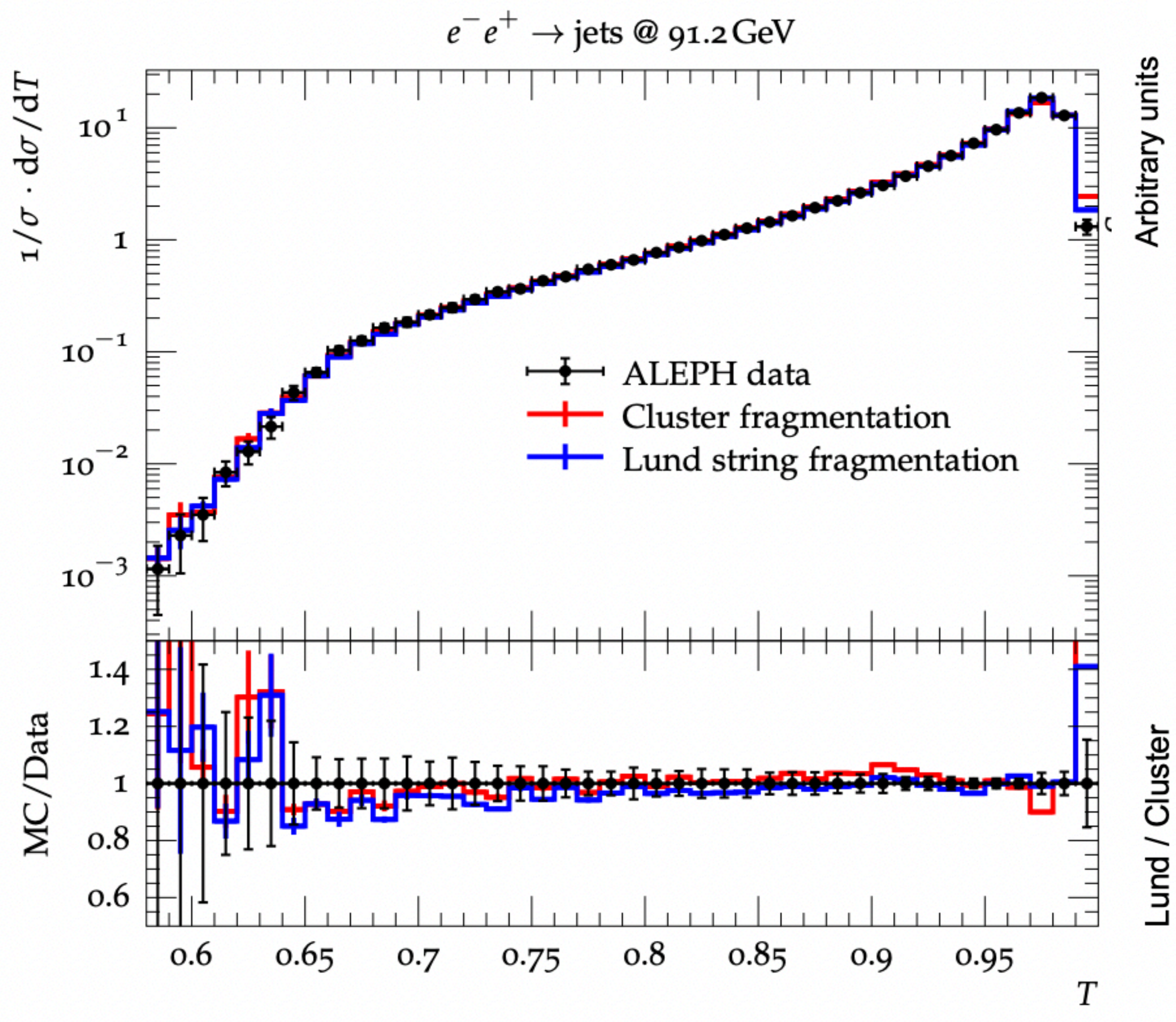
- Even in Les Houches 2015 this was being studied:

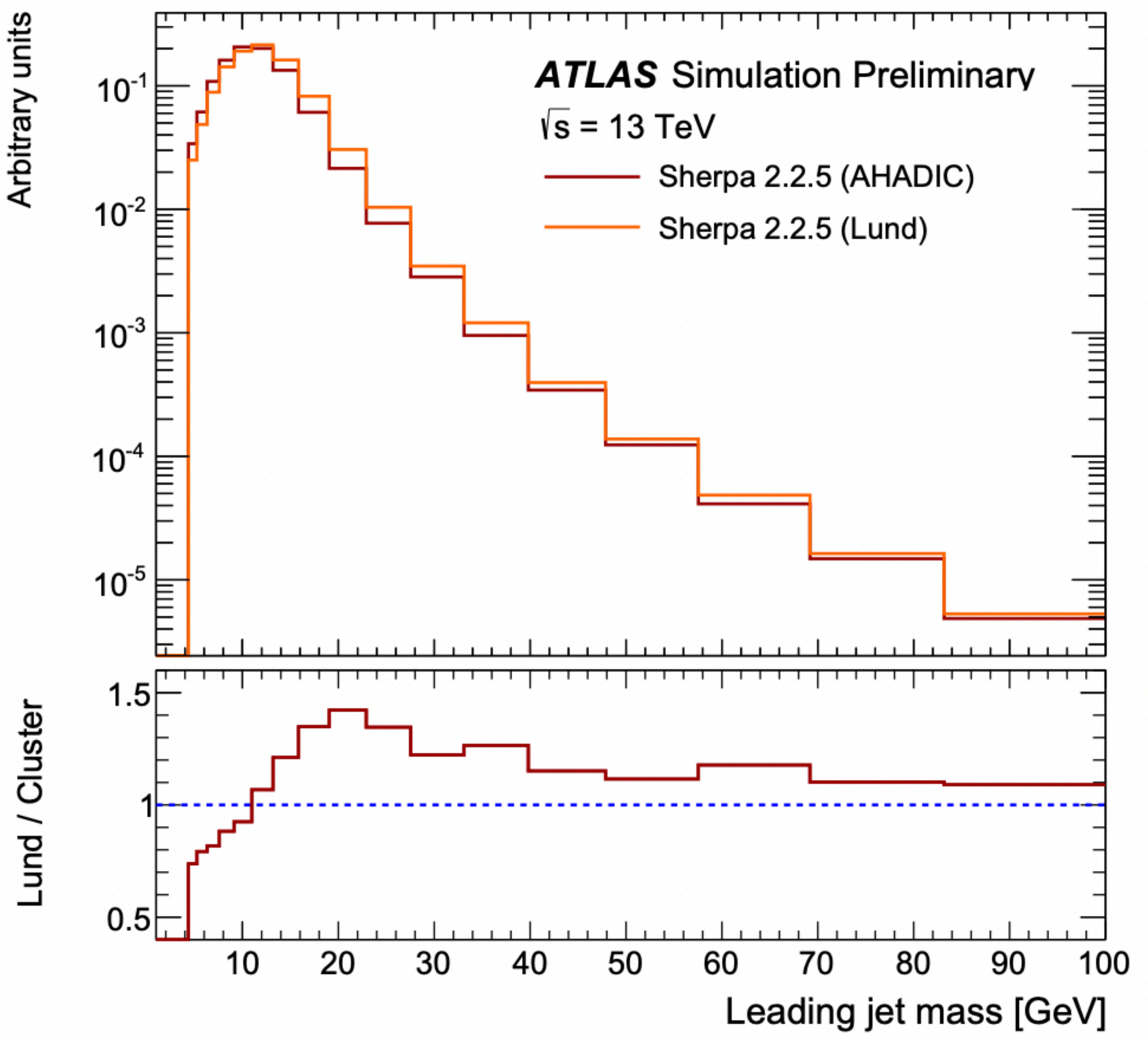


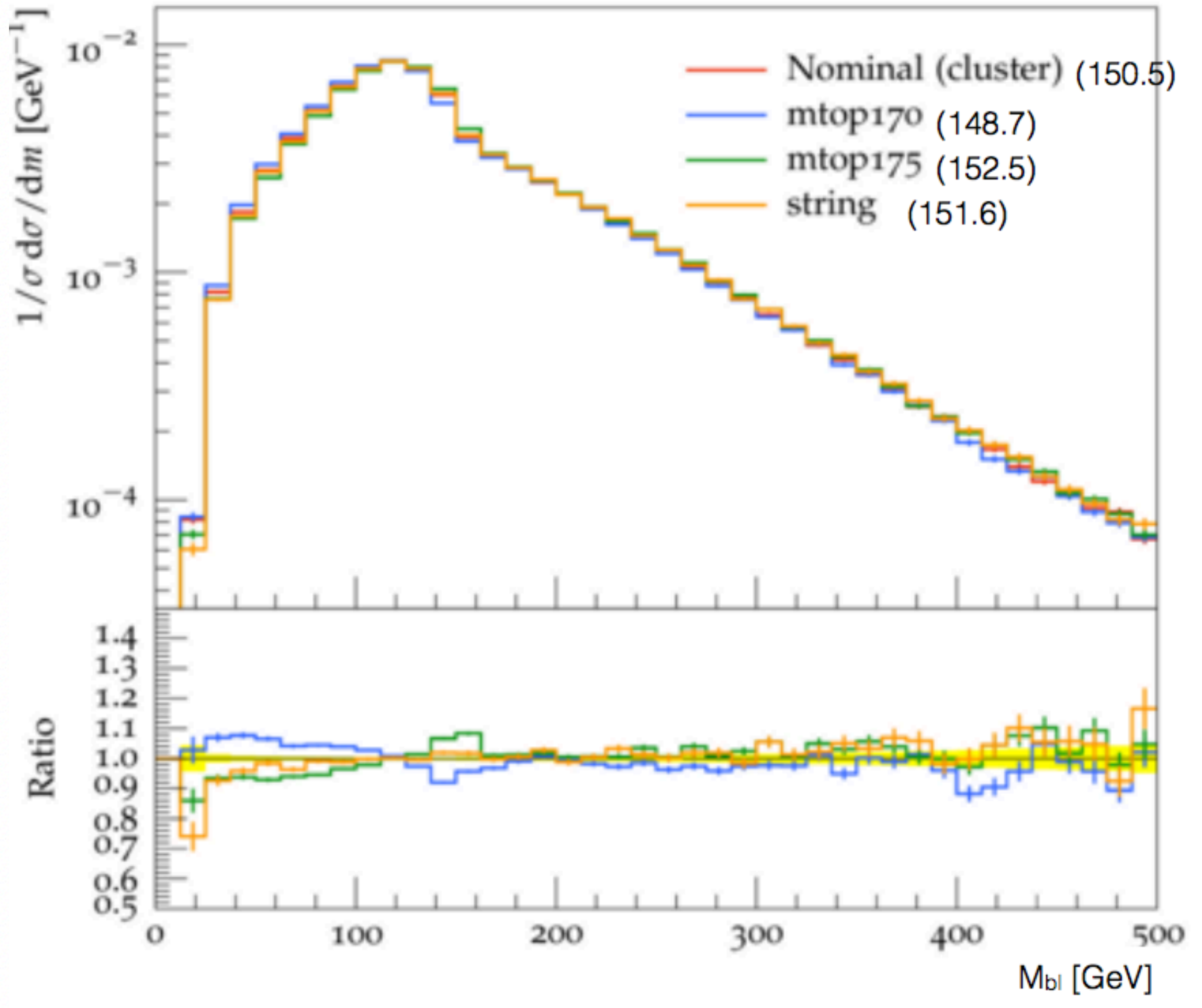
LH2015 Proceedings

Factorisation | Hadronisation

- We do at least in Sherpa have a switchable hadronisation model:

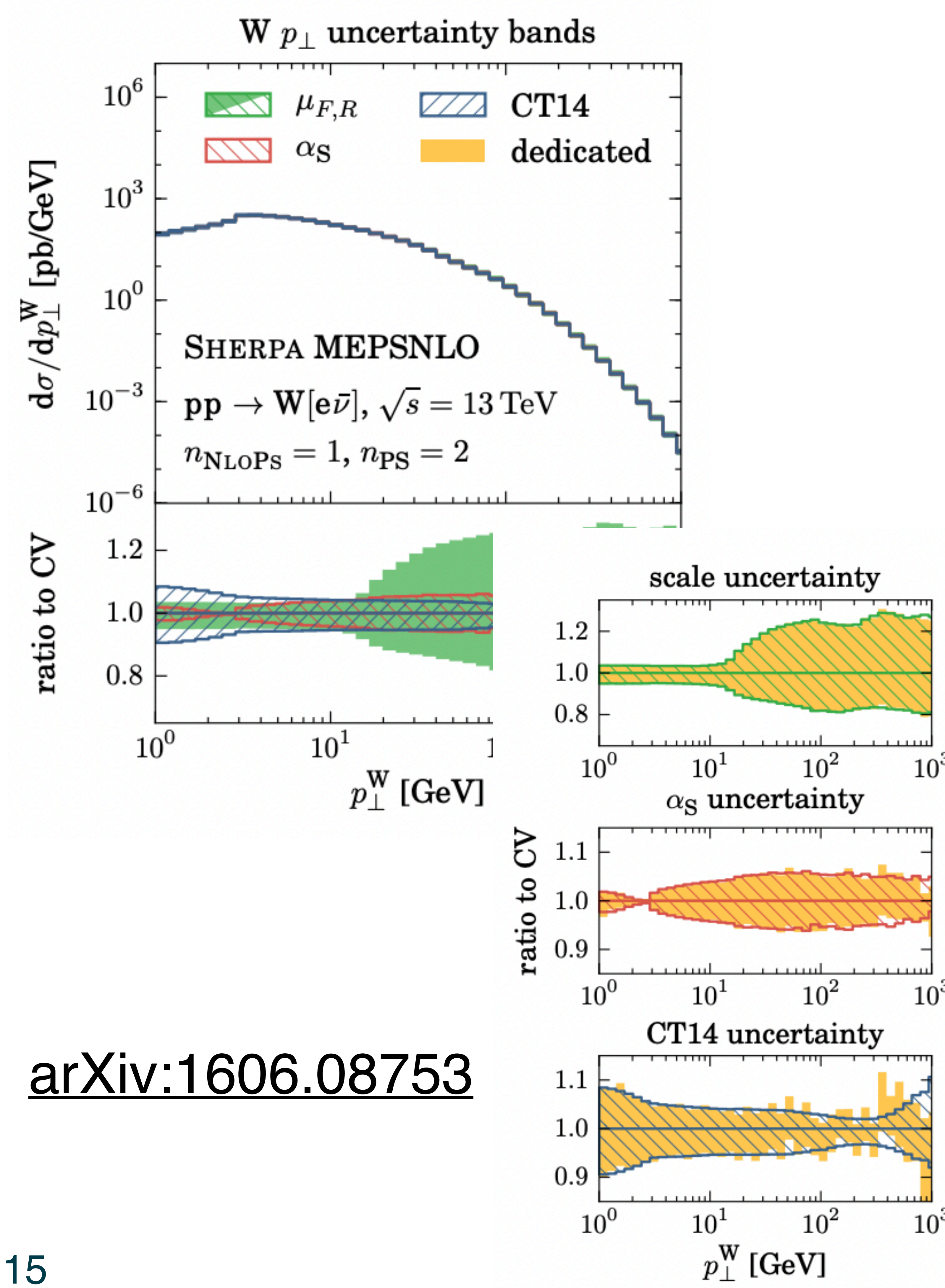




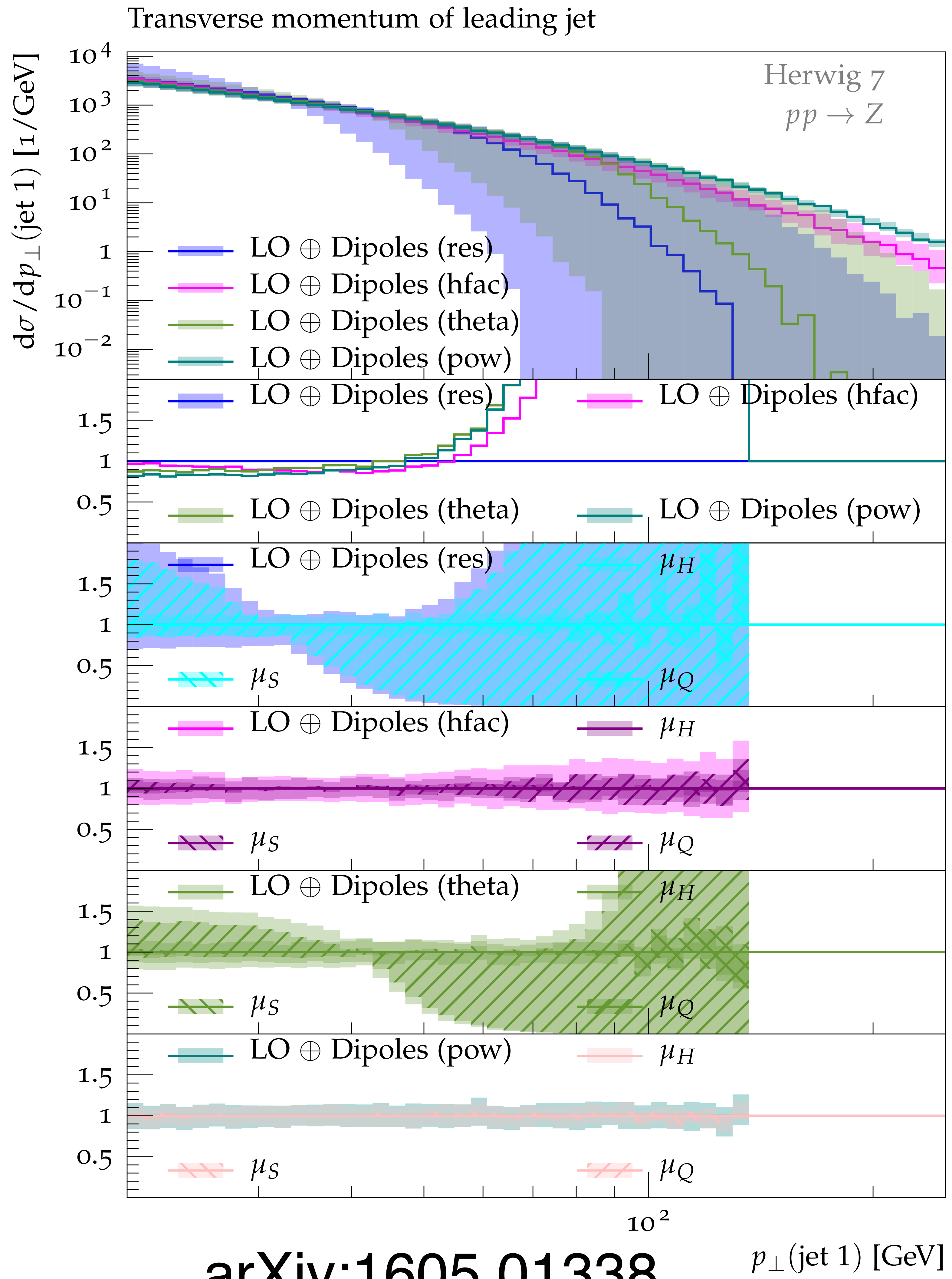


Factorisation | Shower scale

- We have shower uncertainty weights:

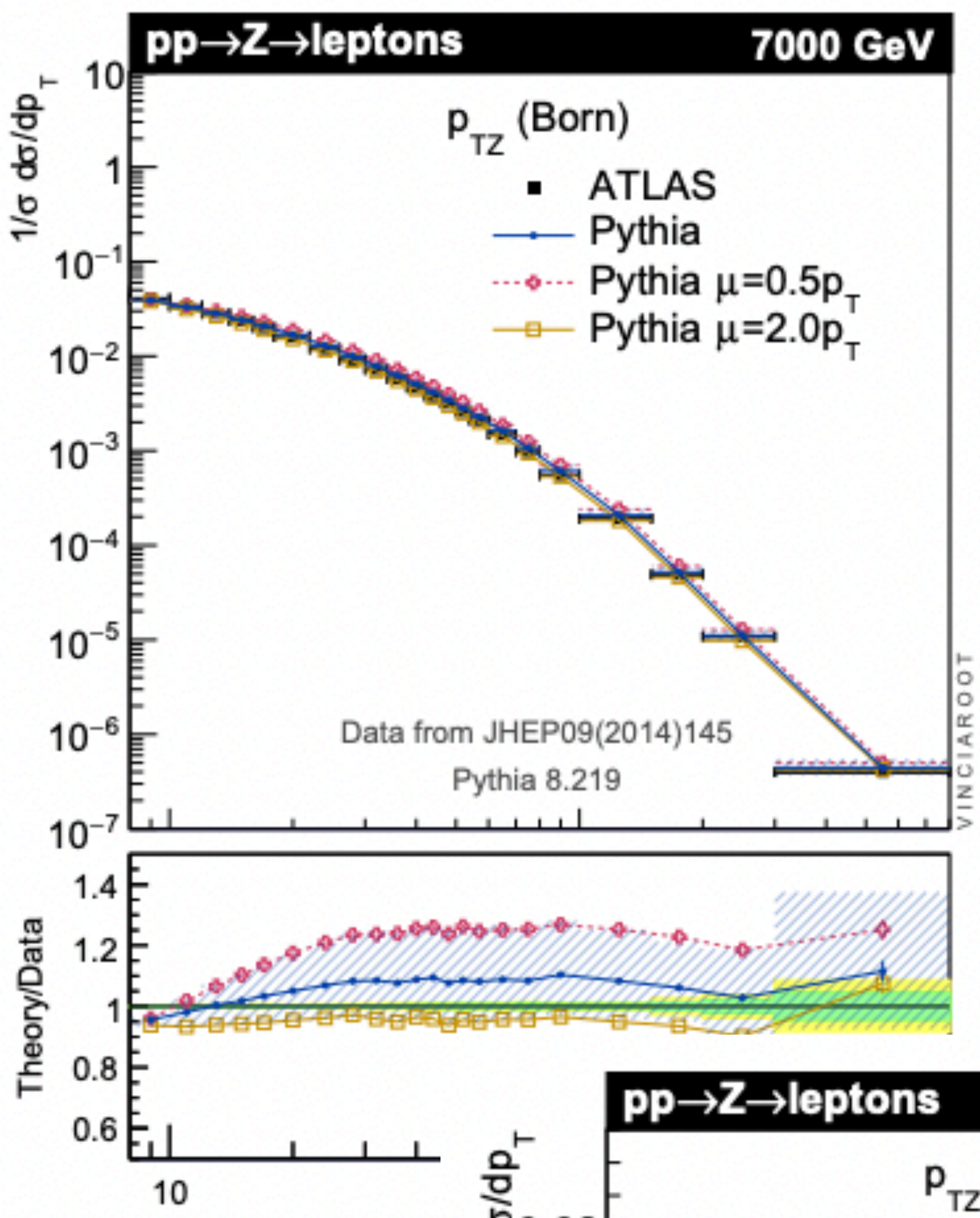


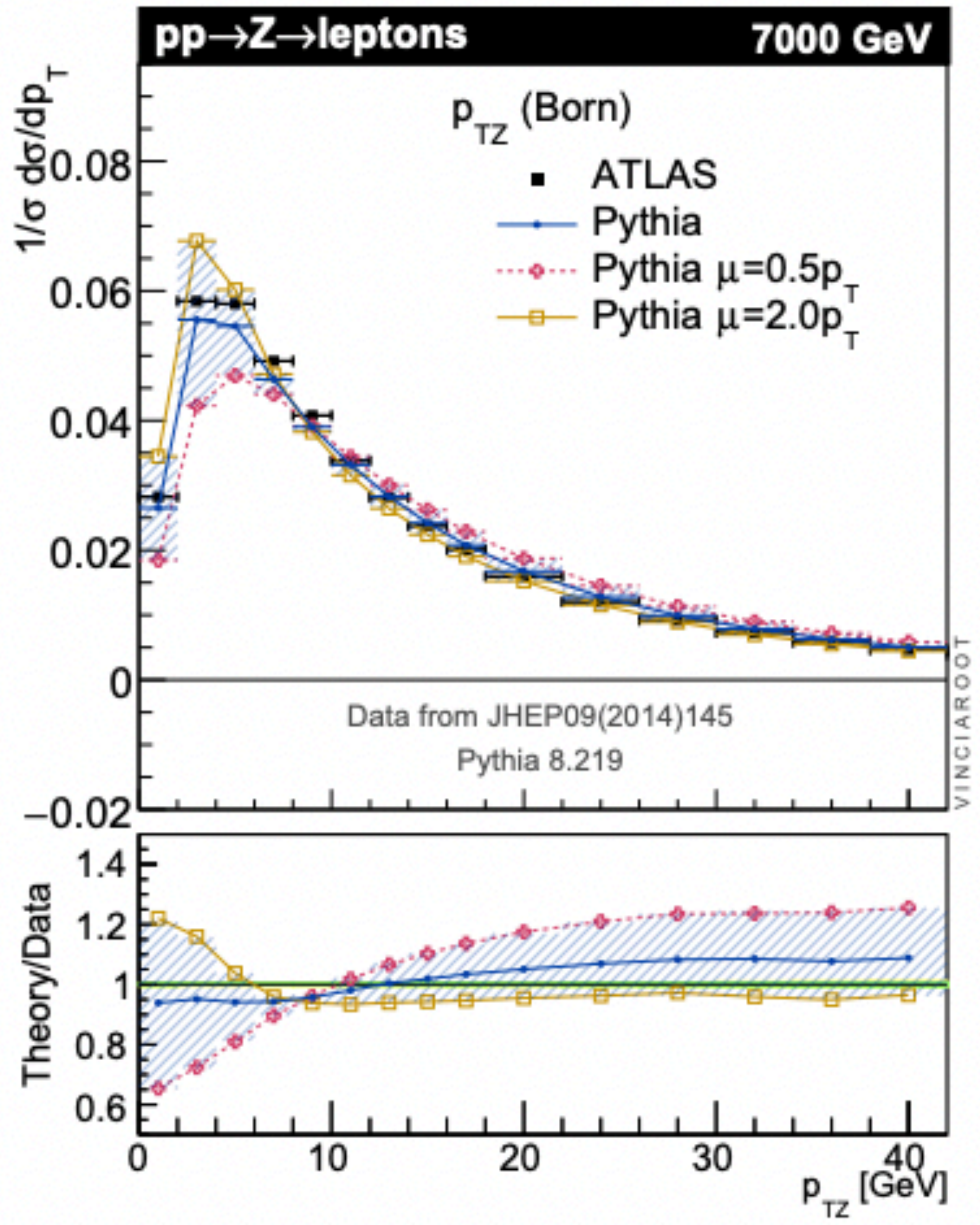
arXiv:1606.08753



arXiv:1605.01338

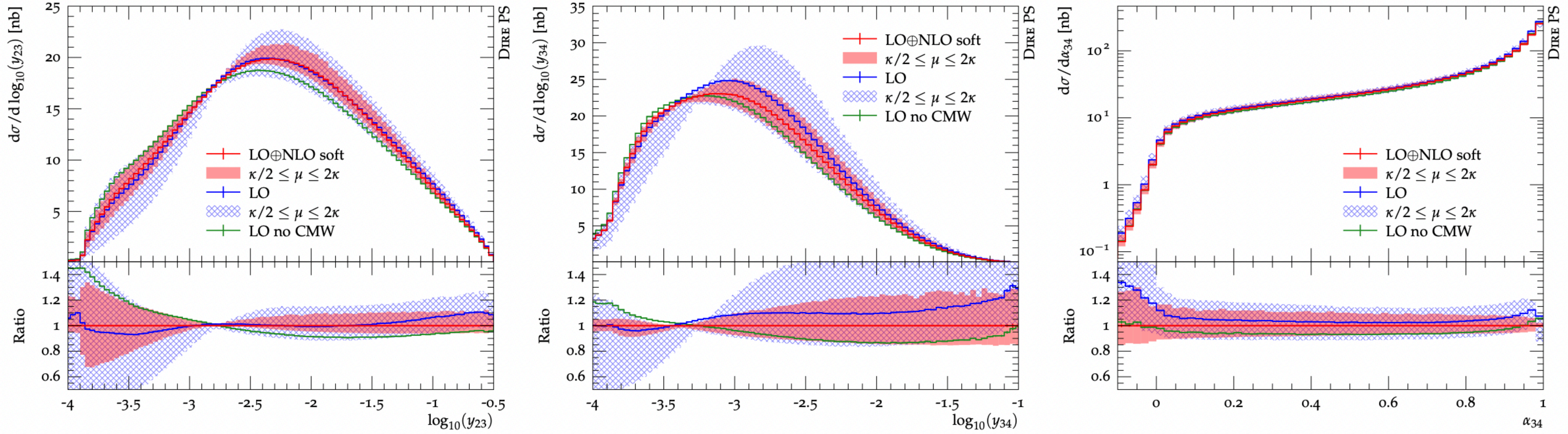
arXiv:1605.08352





Factorisation | Shower scales @ NLO

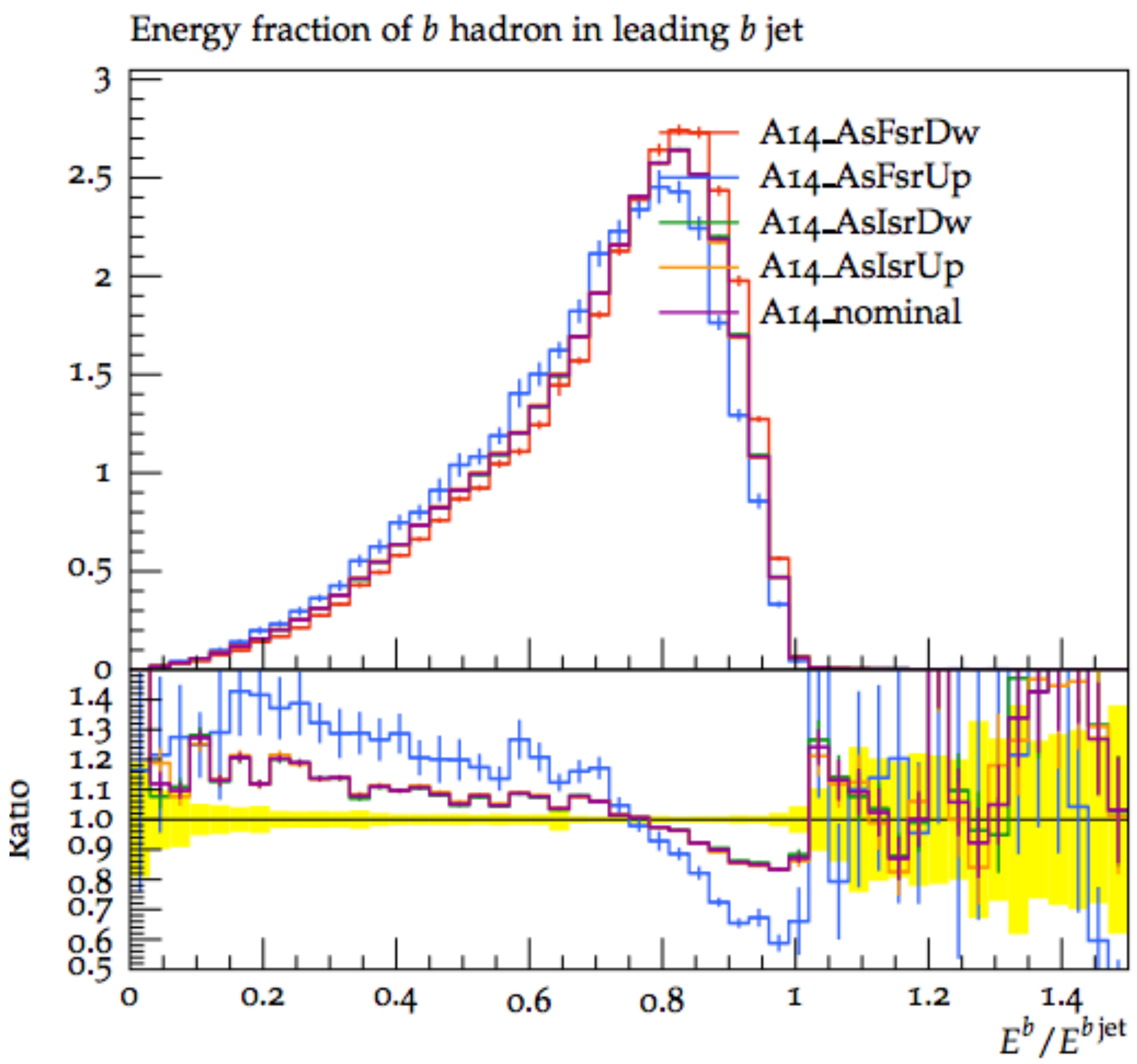
- We are even getting these variations at NLO:

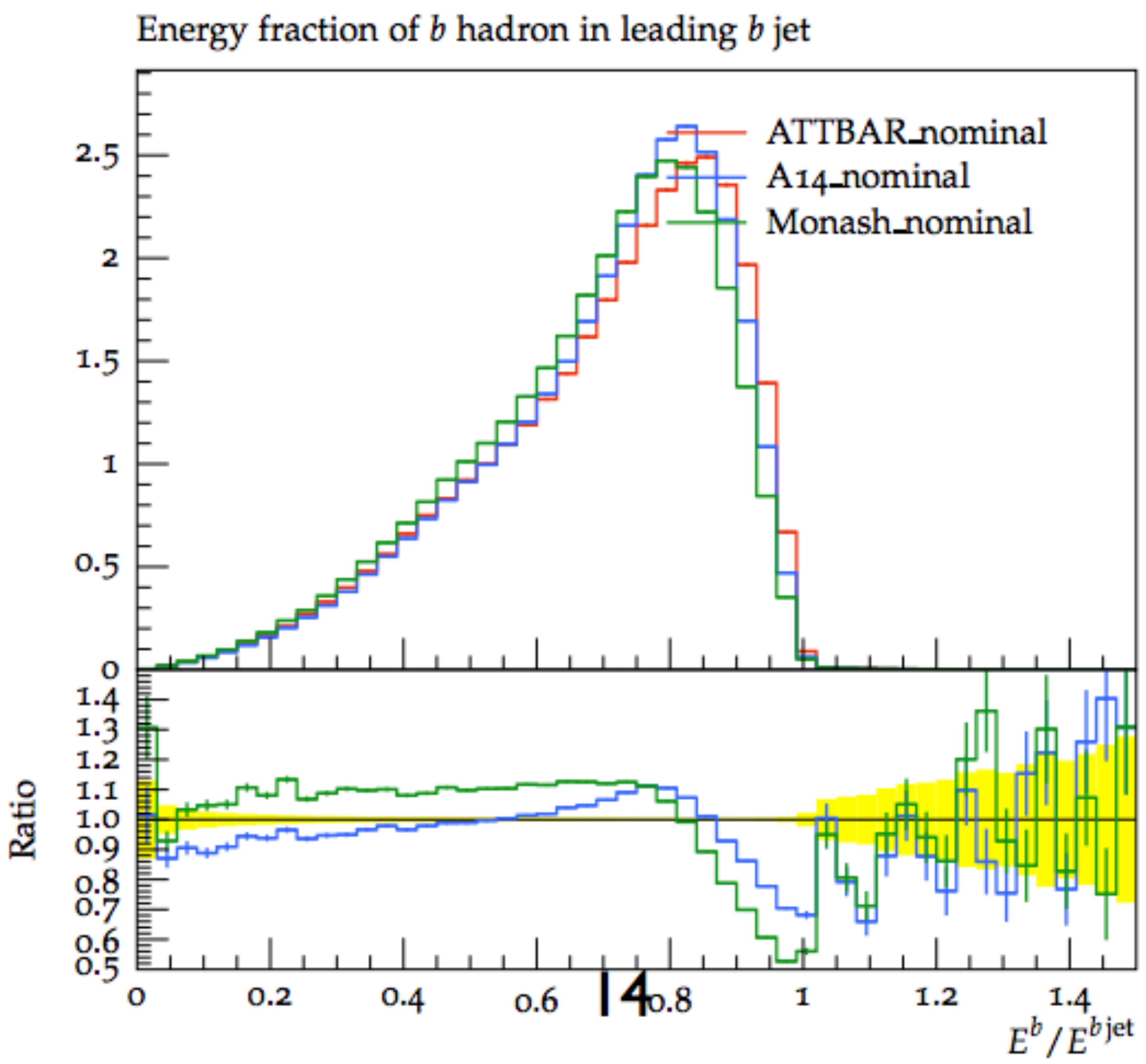


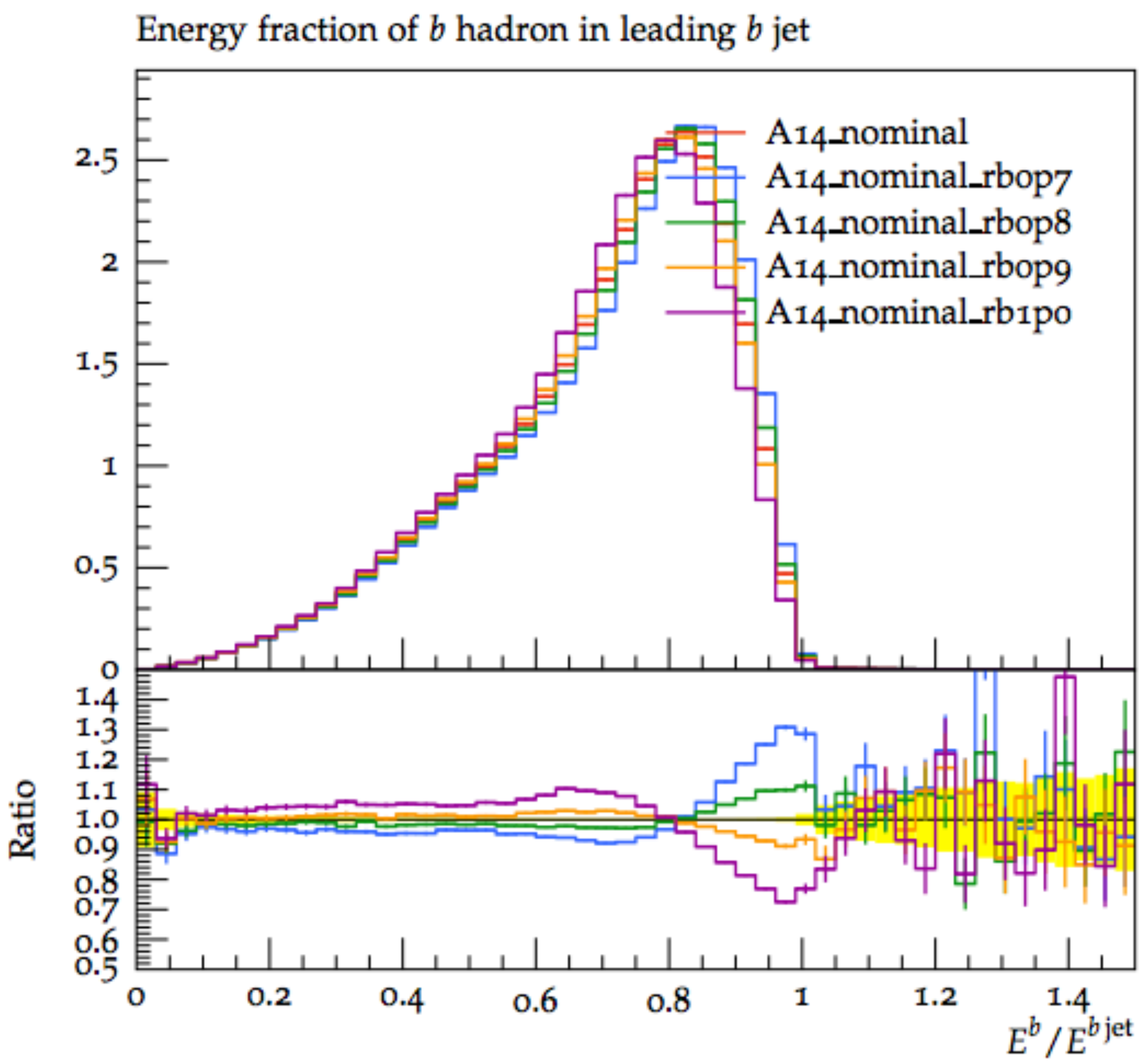
arXiv:1805.03757

Factorisation | Fragmentation

- We have ways to assess the fragmentation uncertainties:

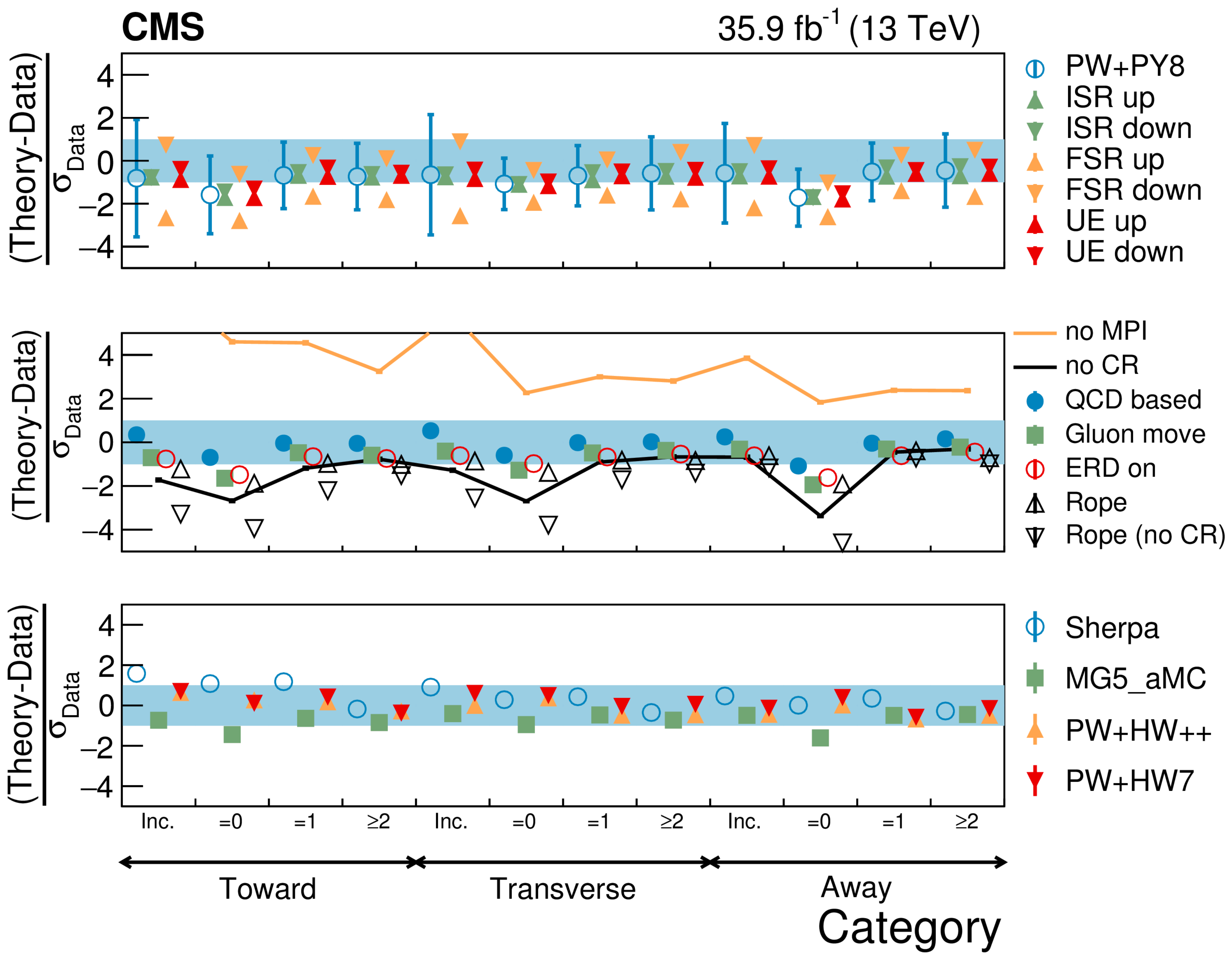


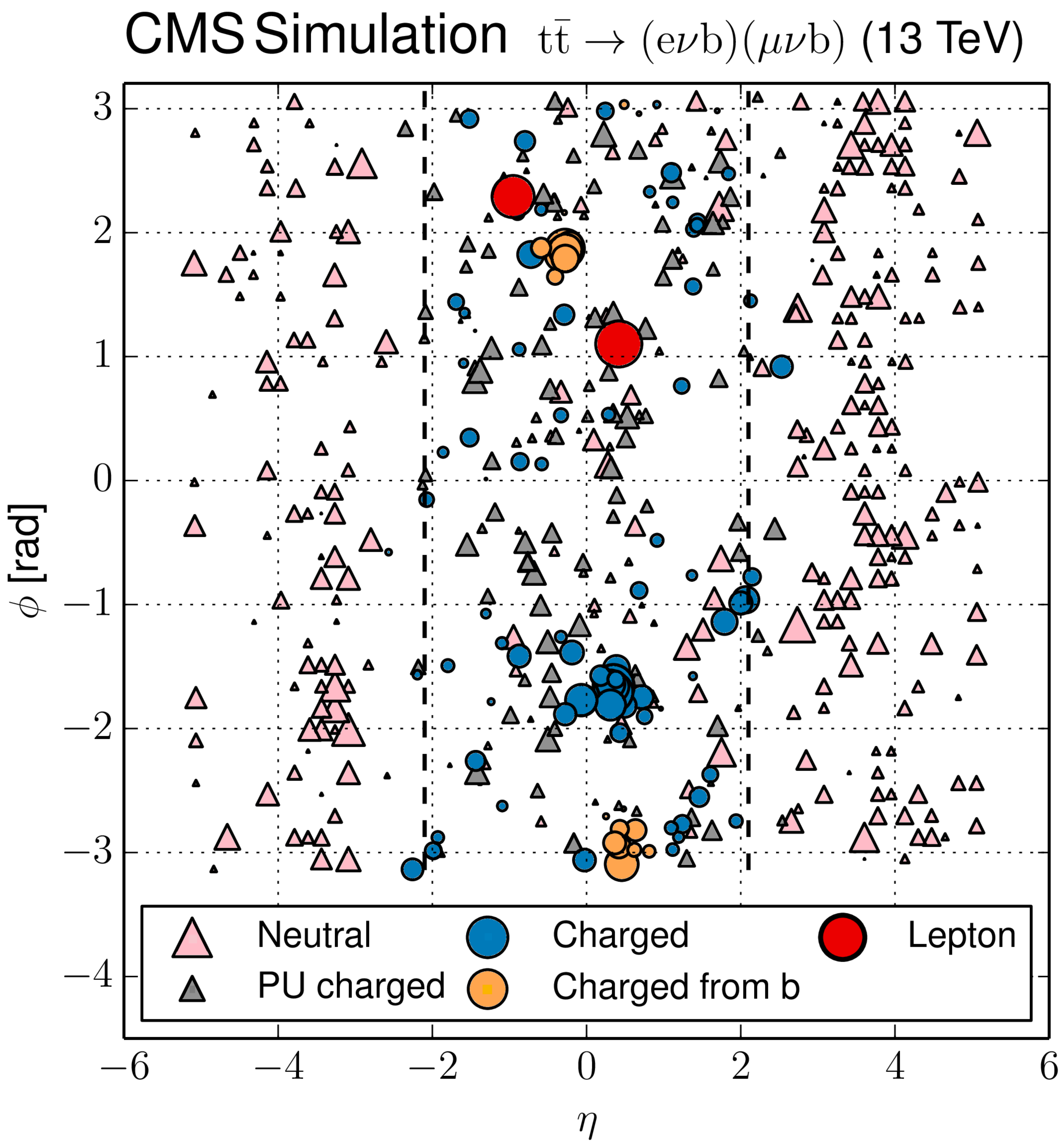




Factorisation | UE / Colour reconnection

- We have tools to vary MPI and colour reconnection:





What to do?

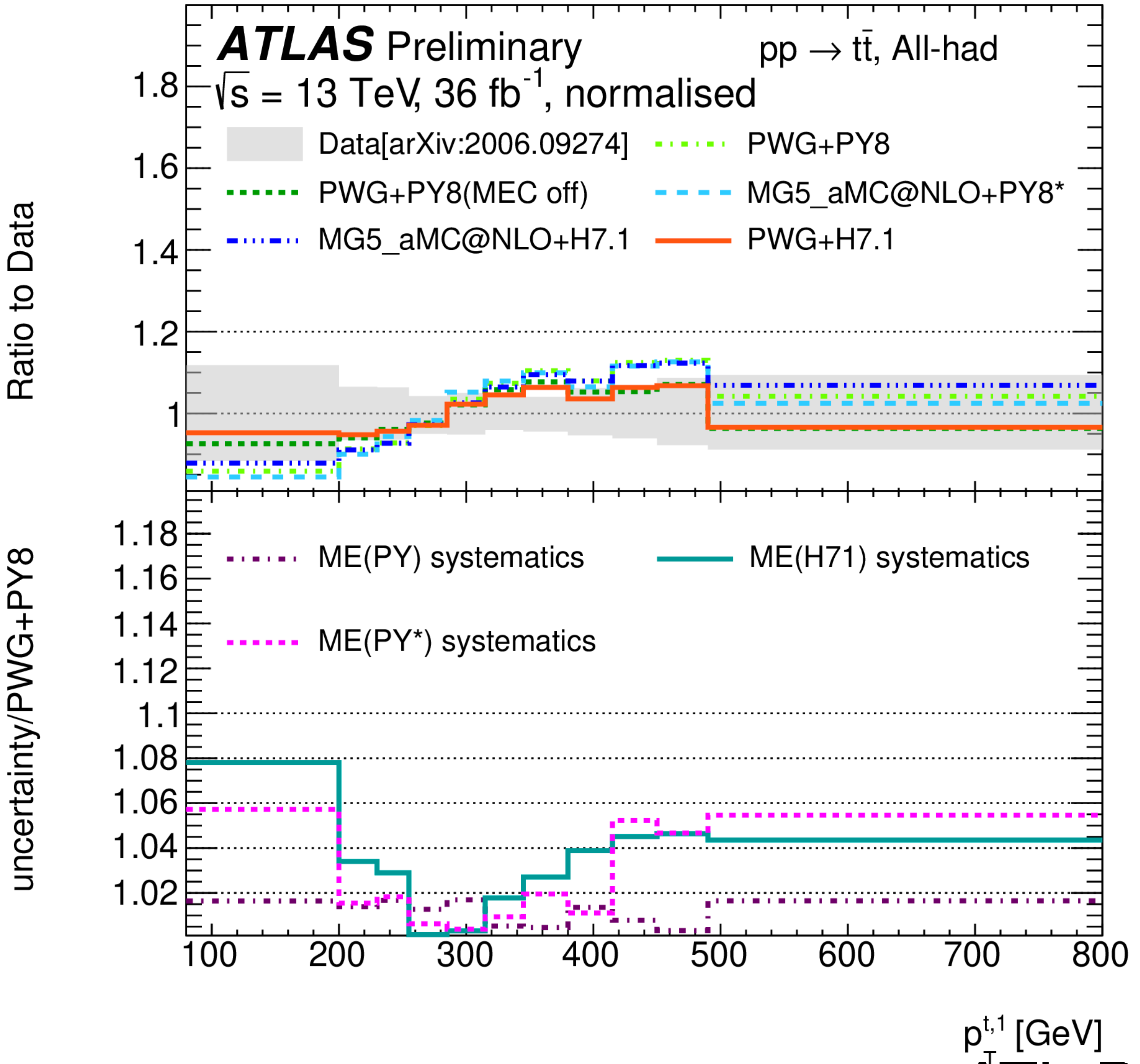
- So what's the problem?

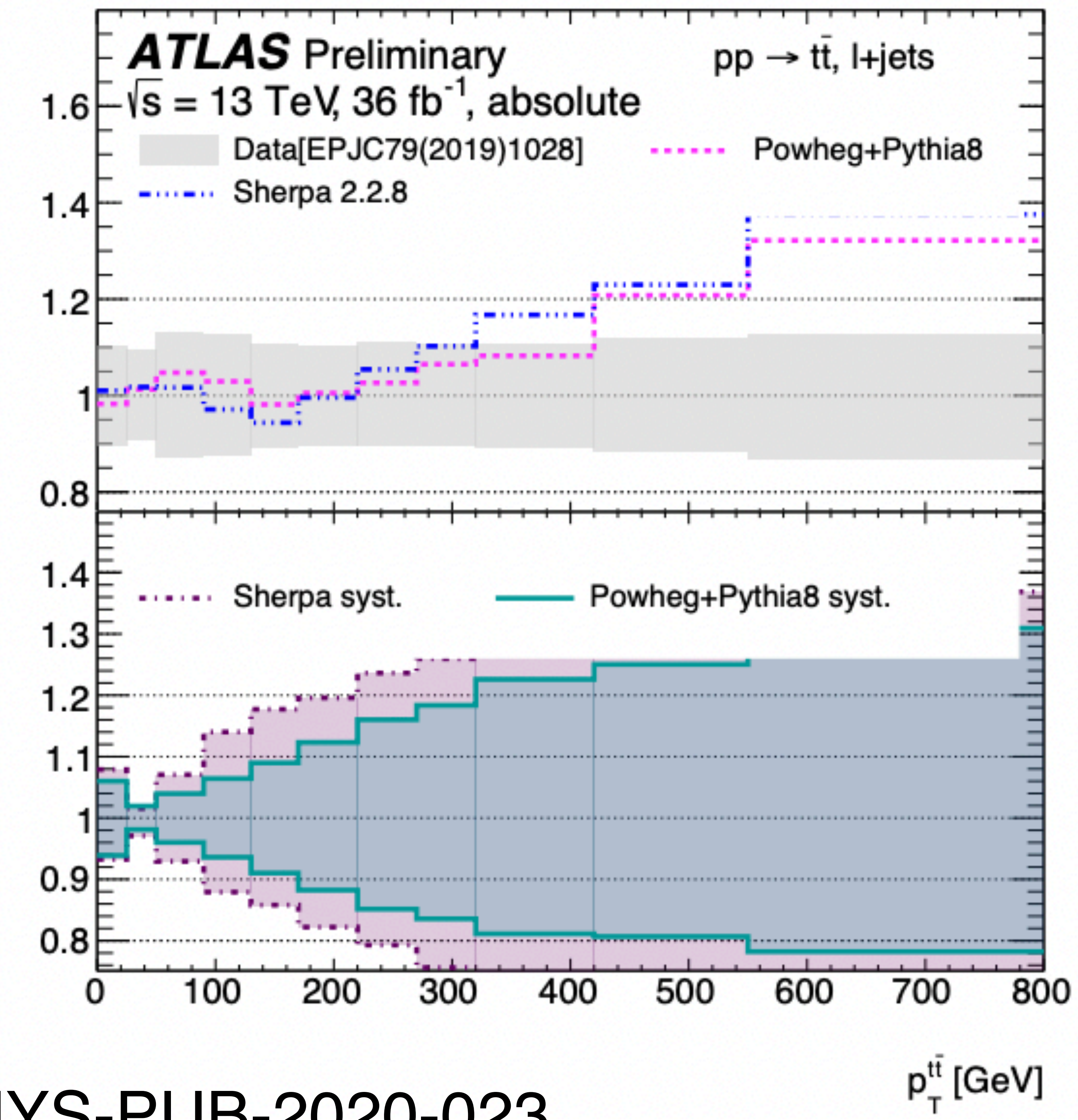
- In many cases new generator configurations take a long time to setup and validate
- It can also be very costly in terms of grid resources to produce completely separate samples

- This implies the need for:
  - More guidance on what variations are most important.
  - Simpler interfaces?
  - Maximal re-use of events?

Existing Issues

- Issues that have been discovered in the past related to this:
  - Incoherent parameter settings for apples-to-apples comparisons
  - Incoherent assessment of uncertainties leading to illogical results.





ATL-PHYS-PUB-2020-023

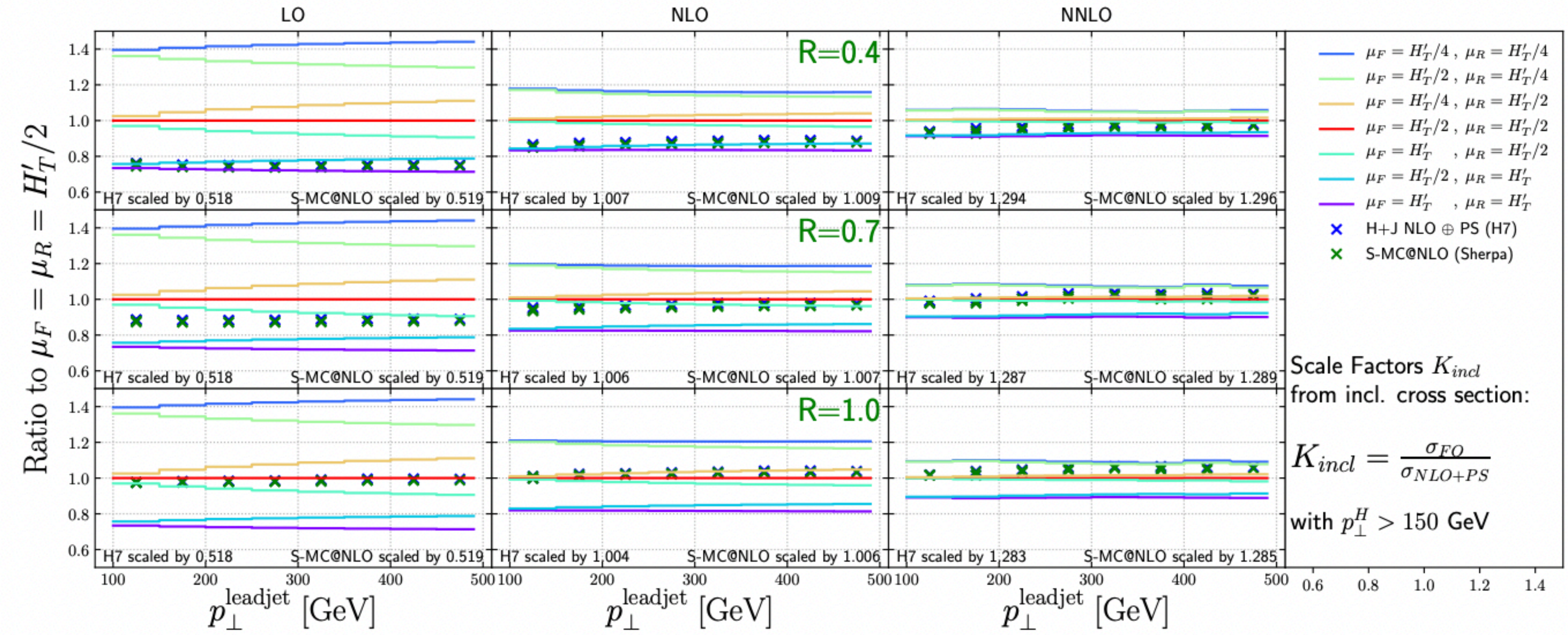
Existing Studies

- So what work was there already in this area?

- It turns out quite a lot…

Existing Studies | NLO+PS vs FO

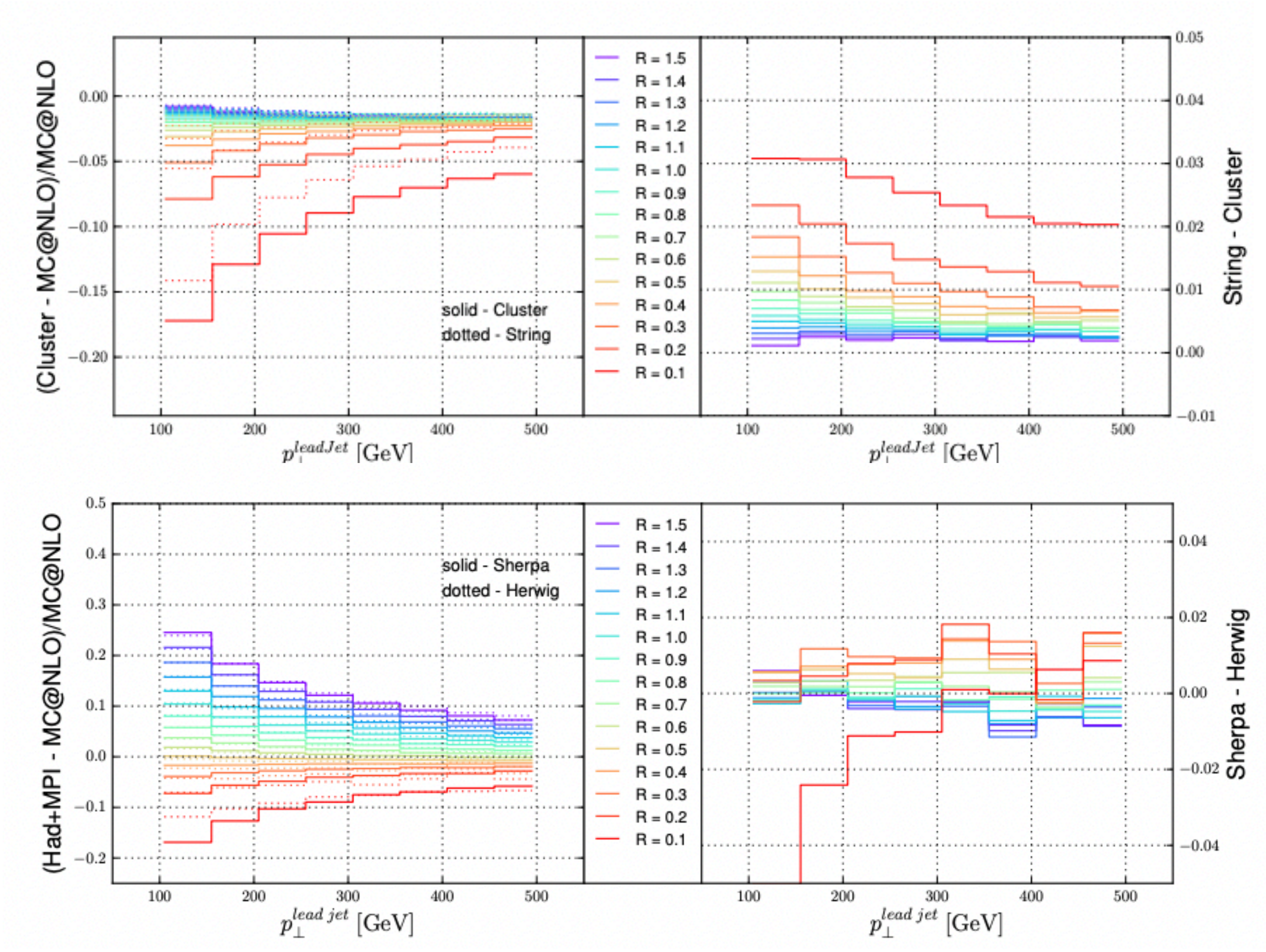
- NLO+PS predictions agree very well with each other and with the FO predictions (to varying degrees depending on the radius parameter).



arXiv:1903.12563

Existing Studies | NP corrections

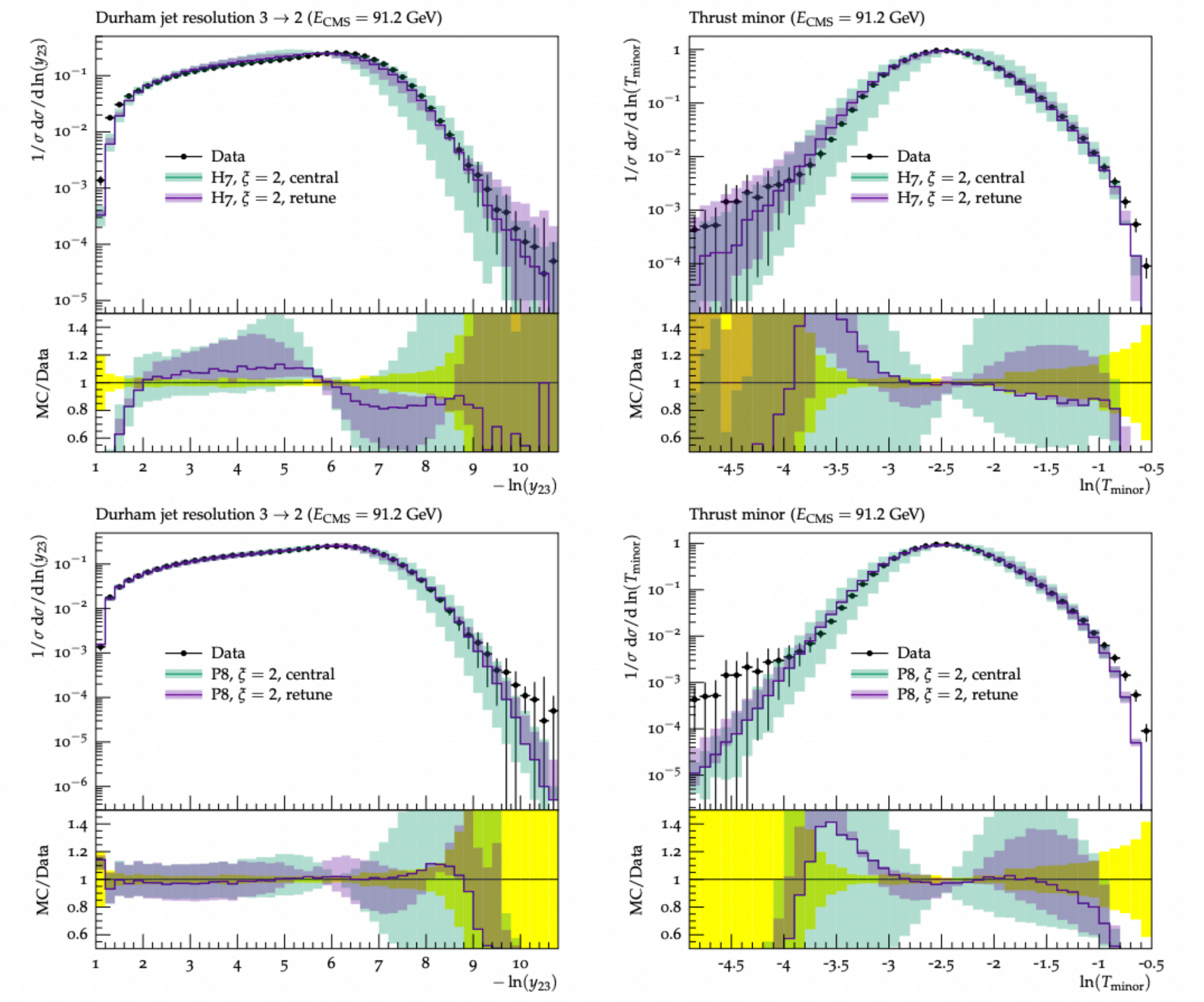
- Non-perturbative predictions compared as a function of jet size R and jet p_T
- Very good agreement between string and cluster fragmentation
- Also between the full non-perturbative corrections, with fragmentation and MPI, between Sherpa and Herwig.



arXiv:1903.12563

Existing Studies | Tunes and scale vars

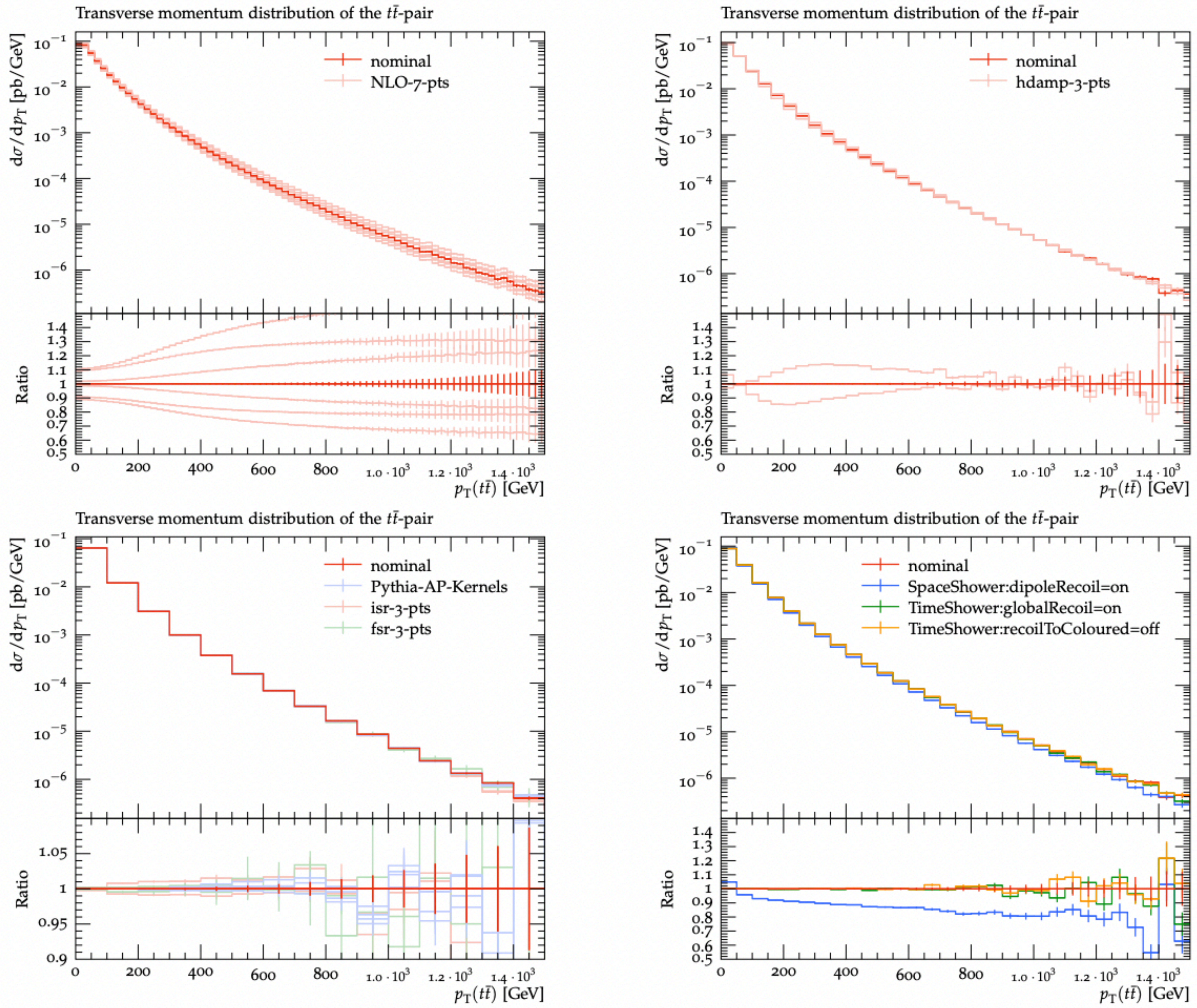
- The naive ”central“ variation bands of HERWIG and PYTHIA are different
- The size of the retuned uncertainties is comparable between the two simulations.
  - This convergence of two very different models suggests that comparing retuned variations might provide a better assessment of MCEG uncertainties in the future.



LH2017 Proceedings

Existing Studies | Perturbative uncs in ttbar

- Aim to establish if perturbative variations from different PS algorithms obtained with different generators are compatible
  - First step in systematically comparing full PS models
- A systematic look at all possible/reasonable variations for a limited setup
  - Powheg+Pythia8 w/o hadronisation or MPI



LH2019 Proceedings

What to do?

- Try to lay out minimal/maximal set of variations needed for a proper assessment of the uncertainties
  - Two levels:
    - Wishlist - write down what you want ideally want for each generator
    - Pheno studies - comparing setups
  - Essentially an extension of existing LH studies.

- Link to specific pheno study?
  - VBS/VBF?
  - Heavy flavour?

- Tools: A common hadronisation interface for Herwig/Pythia/Sherpa?
  - Base class in HepMC or another common package?

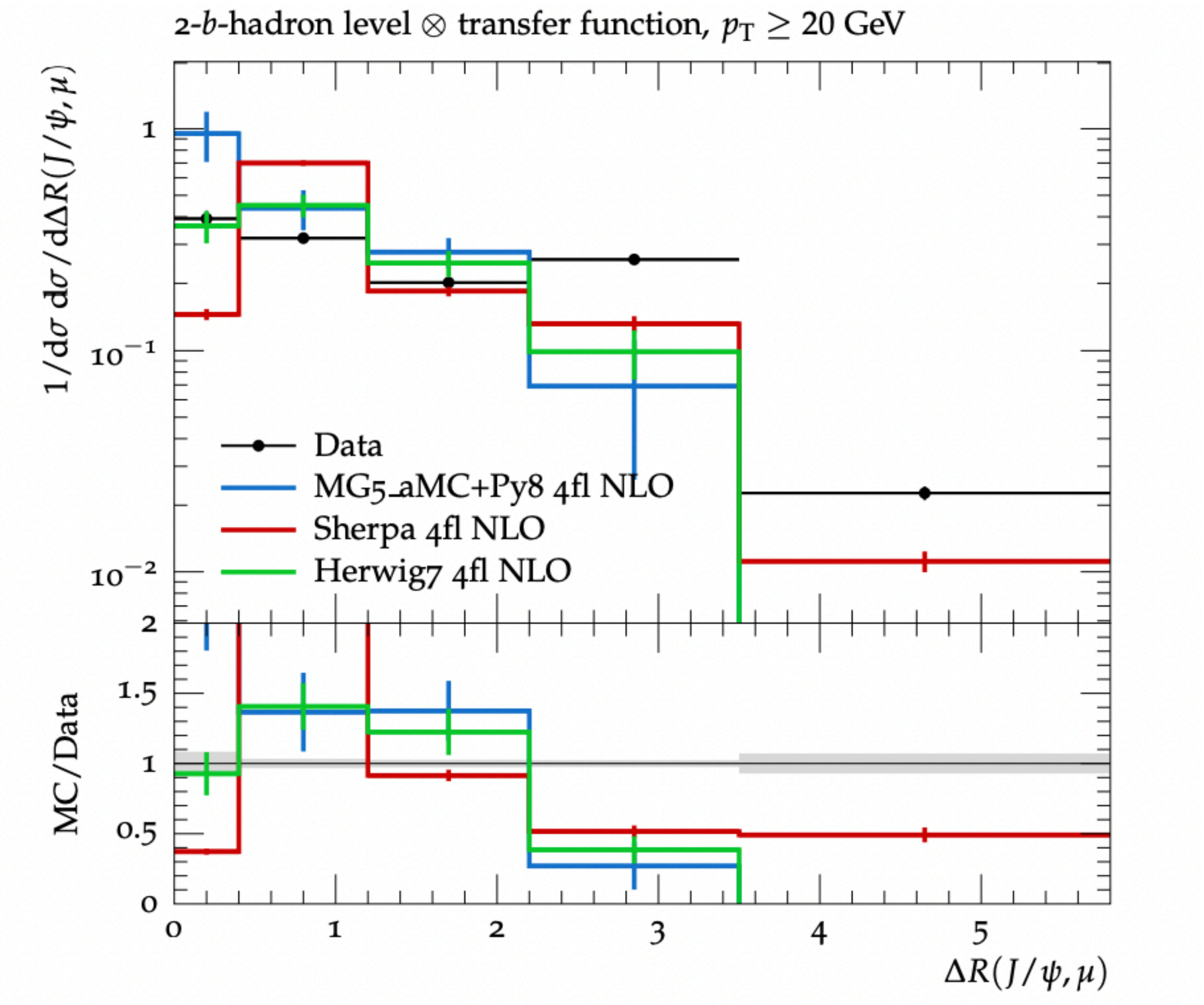
Heavy Flavour

Heavy Flavour

- There are several issues surrounding heavy quark/flavour modelling that are of high priority for experiments
  - ttbb modelling for ttH
  - Vbb for VHbb
  - VVbb for Top+X
  - W+D for strange PDF and eventually Hcc
  - Treatment of off-shell production and top decays for top quark measurements

Heavy Flavour | Global analysis?

- At LH2017 we started trying to initiate a "global analysis" using all available heavy hadron measurements
  - Is this something we want to revisit?!
- Two potential goals:
  - Determine systematic uncertainty recipe
  - Tune to the data

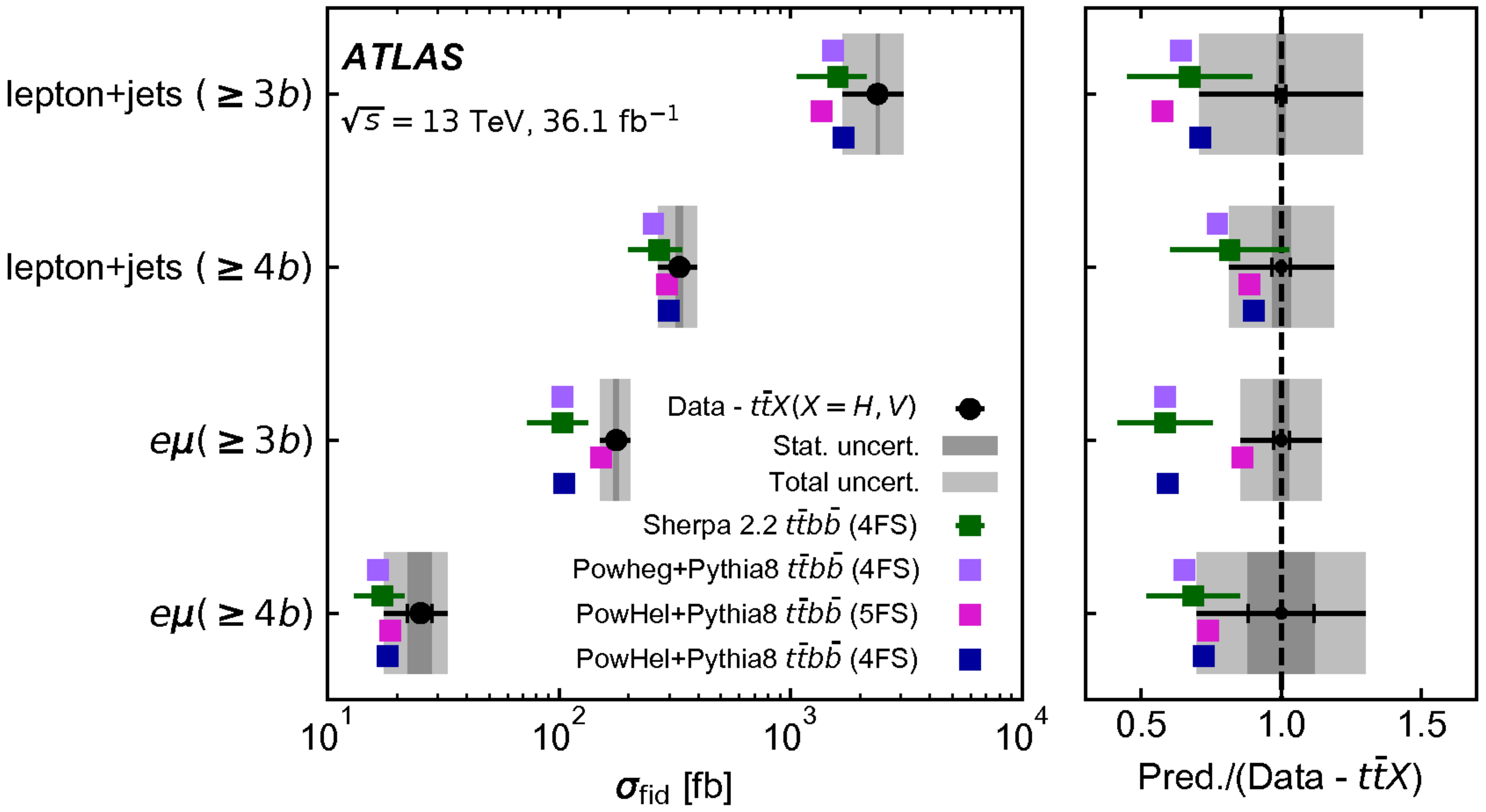


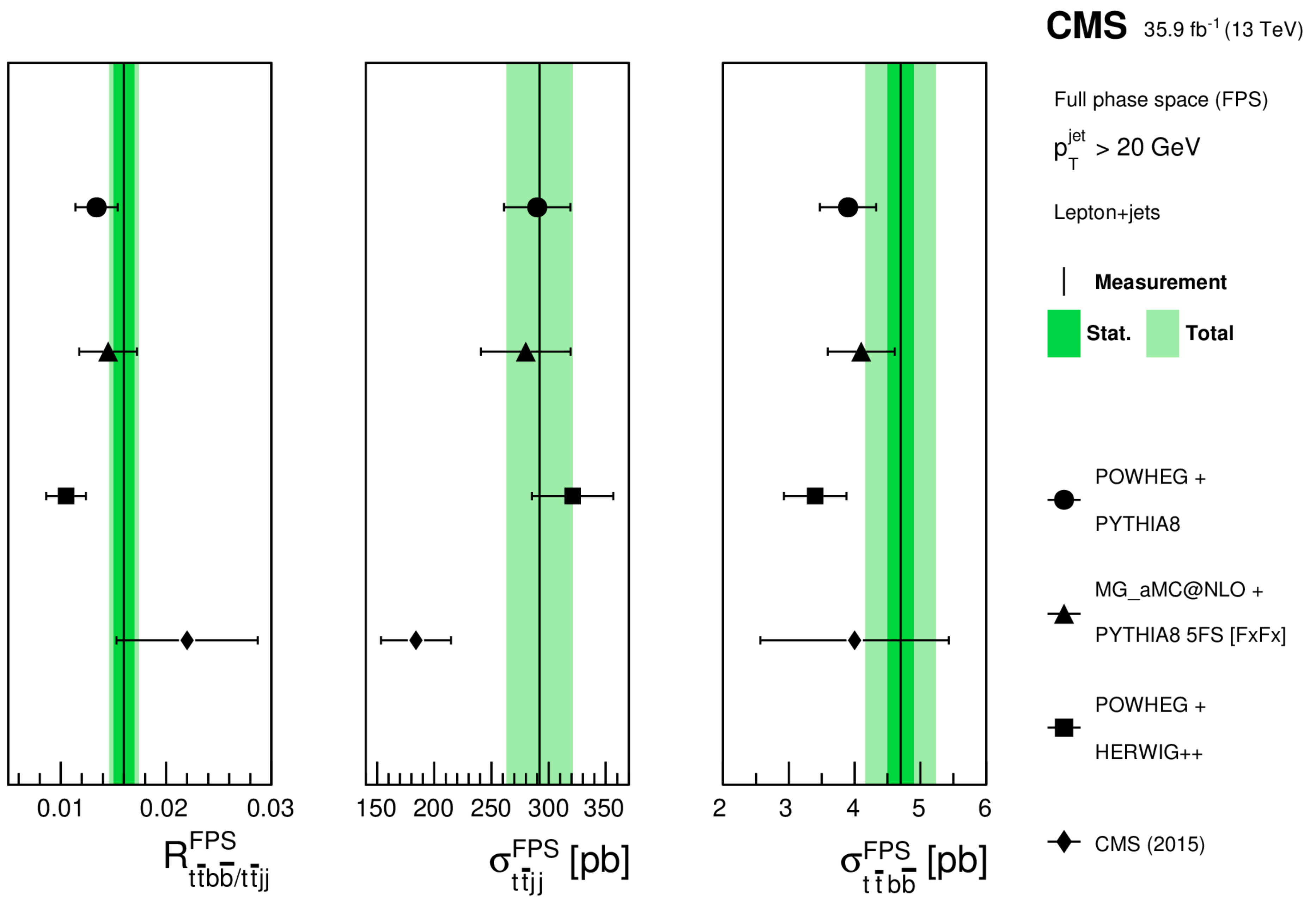
LH2017 Proceedings

Heavy Flavour | Global analysis?

- At LH2017 we started trying to initiate a "global analysis" using all available heavy hadron measurements
  - Is this something we want to revisit?!
- Two potential goals:
  - Determine systematic uncertainty recipe
  - Tune to the data
- Worth coming back to?
  - New measurements

TOPQ-2017-12



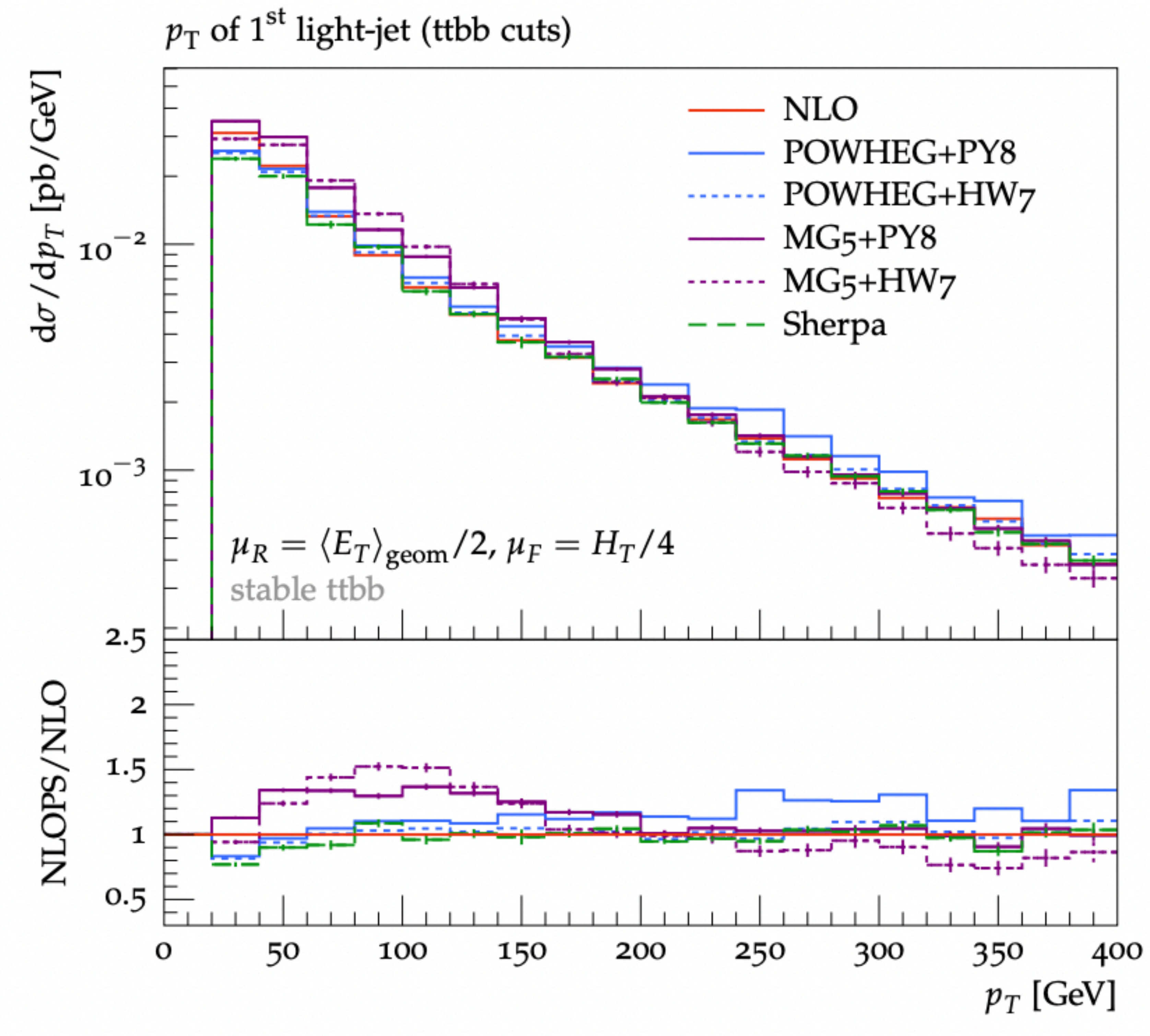


TOP-18-002

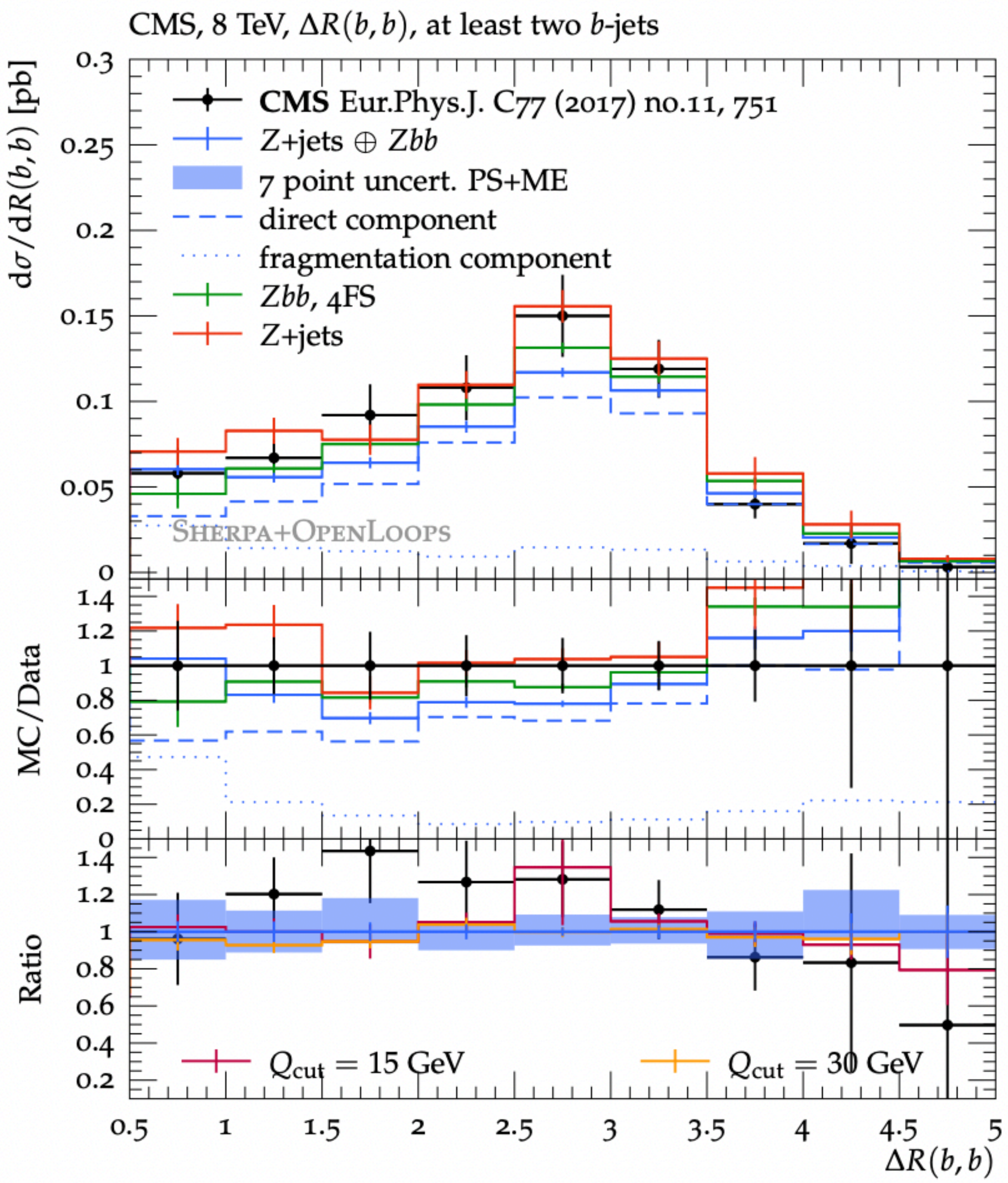
Heavy Flavour | Global analysis?

- At LH2017 we started trying to initiate a “global analysis” using all available heavy hadron measurements
  - Is this something we want to revisit?!
- Two potential goals:
  - Determine systematic uncertainty recipe
  - Tune to the data
- Worth coming back to?
  - New measurements
  - New tools

arXiv:1904.09382



LHC Higgs WG Studies

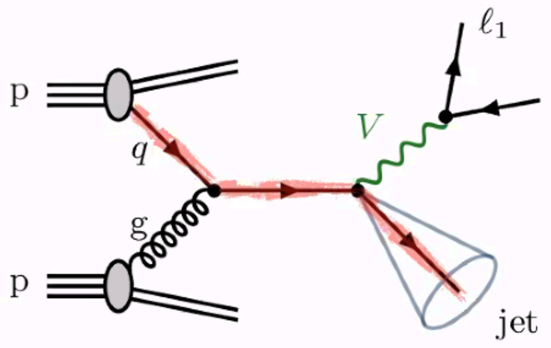


Heavy Flavour | W+D

- Have new NNLO calculations on the market for e.g. W+c
- But measurements at Hadron level are of high priority
  - Need to connect precise analytical calculations to fragmentation functions (and possibly decay packages from MC generators?)
- Otherwise can't be used in PDF fits
  - See **Miha's talk** for more information.

V + jet Production @ NNLO QCD with Flavour

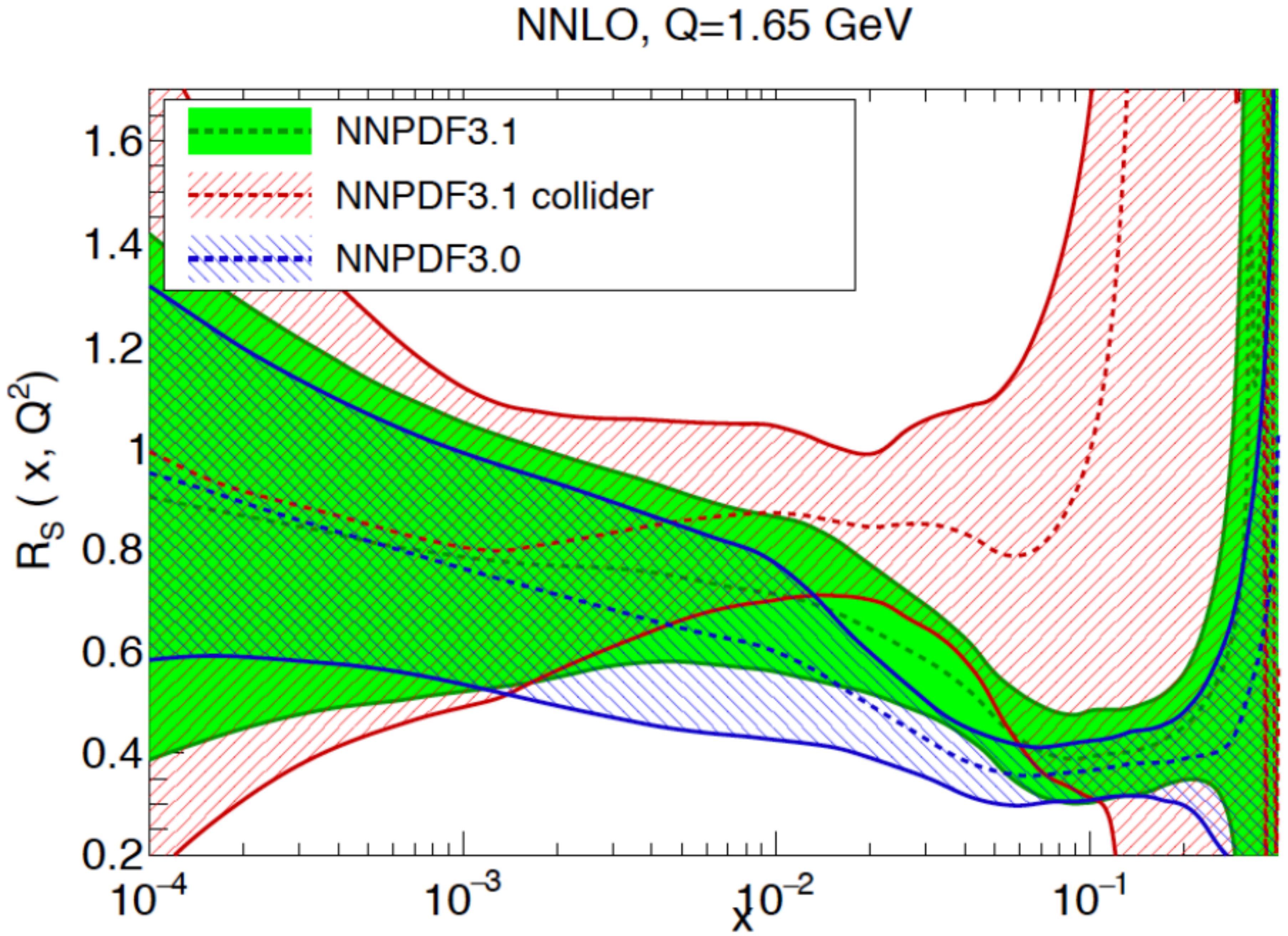
NNLO QCD now well-established with 2 independent calculations:



- Z+jet:
  - Antenna: [Gehrmann-De Ridder, Gehrmann, Glover, AH, Morgan '15]
  - N-jettiness: [Boughezal, Campbell, Ellis, Focke, Giele, Liu, Petriello '15]
- W+jet:
  - N-jettiness: [Boughezal, Liu, Petriello '15]
  - Antenna: [Gehrmann-De Ridder, Gehrmann, Glover, AH, Walker '17]
- γ+jet:
  - N-jettiness: [Campbell, Ellis, Williams '16]
  - Antenna: [Chen, Gehrmann, Glover, Höfer, AH '19]

… now comes in different flavours:

- Z+b-jet [Gauld, Gehrmann-De Ridder, Glover, AH
- W+c-jet [Czakon, Mitov, Pellen, Poncelet '20]
- identify flavour of a jet ("tag")



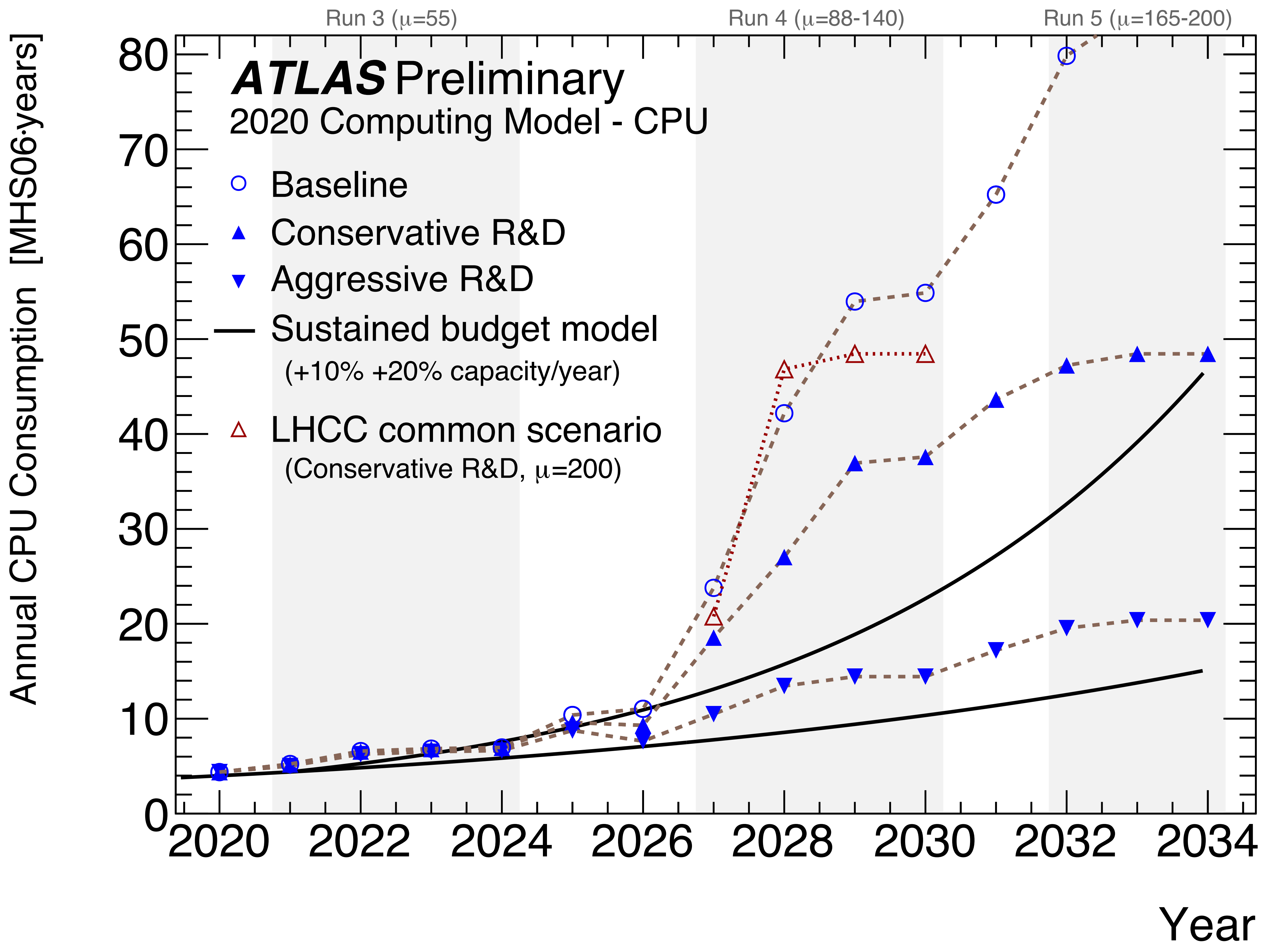

MCStatistics

Computing considerations

- We know that resource projections for HL-LHC look very challenging

- HEP Software Foundation Event Generators WG recently published: *"Challenges in MC event generator software for HL-LHC"* [2004.13687]
  - Details quite a lot of the areas for further focus
    - More efficient software
    - Porting to GPUs
    - Use of ML tools

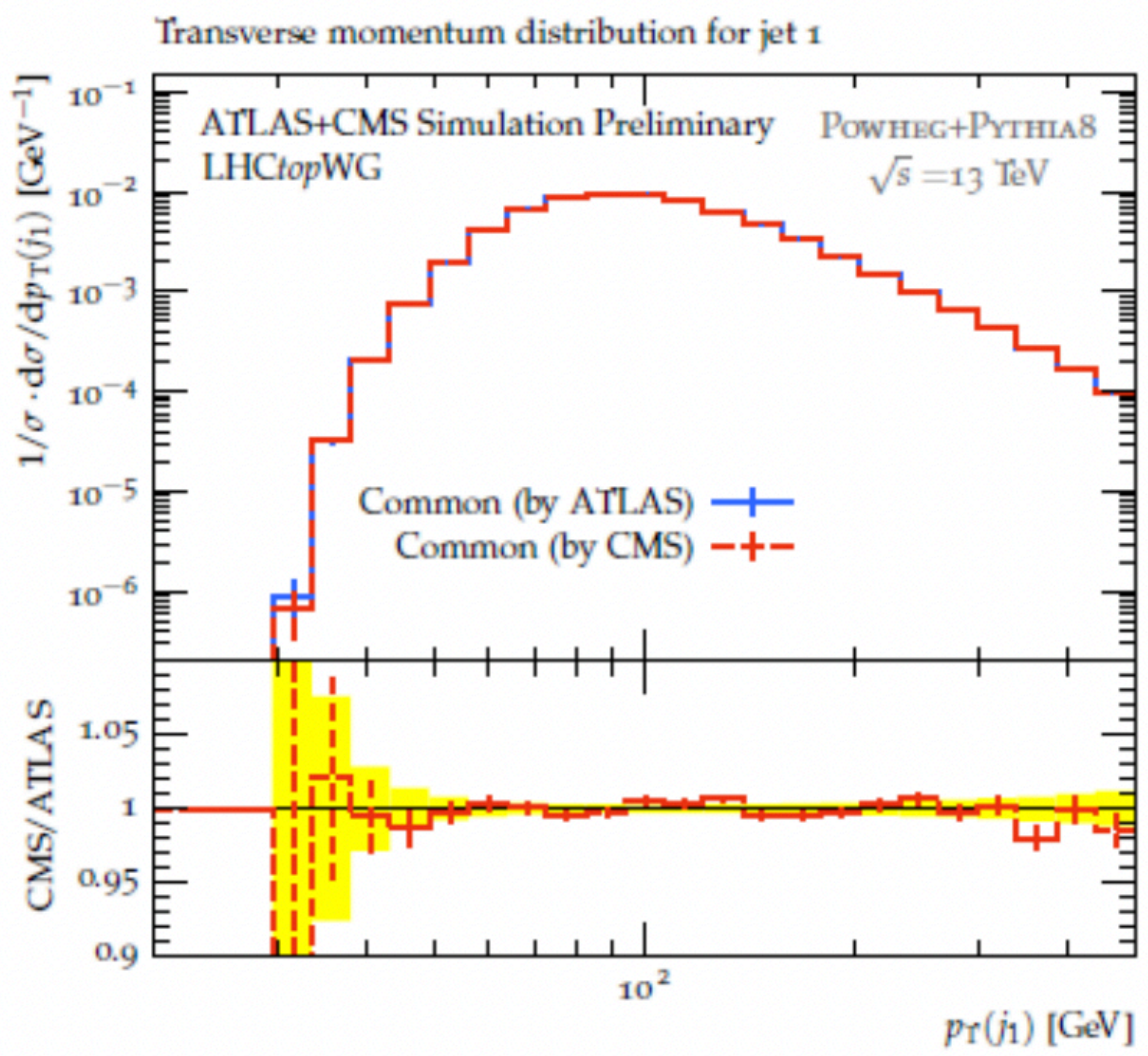
- LH-specific possibilities
  - Benchmarking and link to ML tools
  - Common samples



Computing considerations

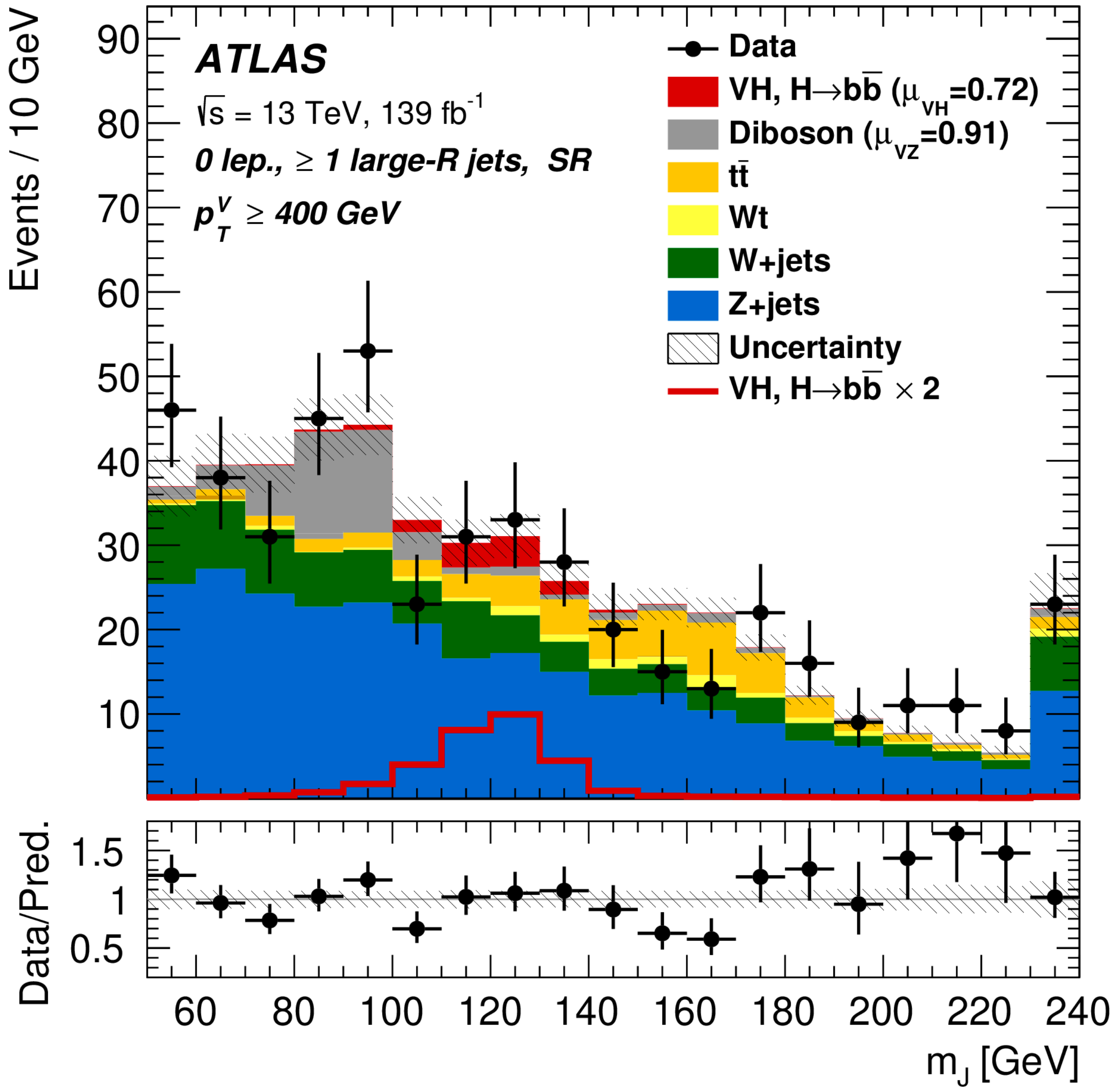
- Benchmarking
  - Create code kernels for different generation steps for benchmarking/porting?
    - E.g. simple case and realistic case for each step
  - Use for classification of ML tools/approaches
    - Critical assessment/comparison against existing MC techniques (e.g. Vegas, FOAM, EXSAMPLE, …)
  - Isolated fragments could also be used for porting to different architectures.

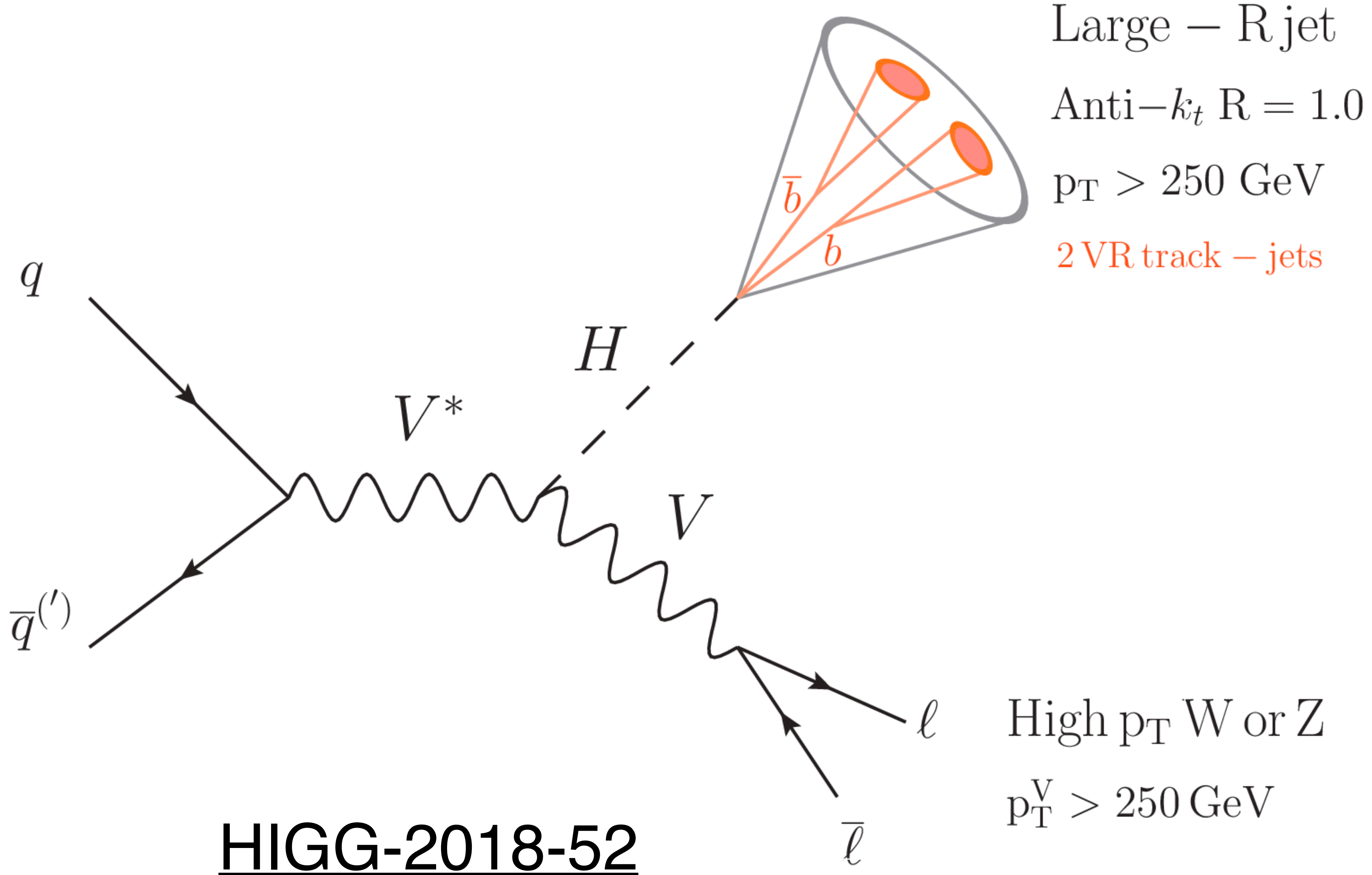
- Common samples
  - Sharing events between ATLAS and CMS
    - Progress in ttbar: Top LHC WG talk
  - Sharing between hadronisation algorithms?



Back-ups

Limited MC stats





HIGG-2018-52

Source of uncertainty		Avg. impact
Total		0.372
Statistical		0.283
Systematic		0.240
Experimental uncertainties		
Small-R jets		0.038
Large-R jets		0.133
E_T^{miss}		0.007
Leptons		0.010
b-tagging	b-jets	0.016
	c-jets	0.011
	light-flavour jets	0.008
	extrapolation	0.004
Pile-up		0.001
Luminosity		0.013
Theoretical and modelling uncertainties		
Signal		0.038
Backgrounds		0.100
↪ Z + jets		0.048
↪ W + jets		0.058
↪ $t\bar{t}$		0.035
↪ Single top quark		0.027
↪ Diboson		0.032
↪ Multijet		0.009
MC statistical		0.092