Experimental Systematics for HL-LHC projections - ATLAS -

Simone Pagan Griso (LBNL)
Sarah Demers (Yale)
Marianna Testa (LNF)
Stephane Jezequel (CNRS)
Marumi Kado (CNRS)

Yellow-Report HL/LHC - WG2

9th May 2018

Introduction

- Current approaches for systematic uncertainties (used in TDRs):
  - assume similar uncertainties as Run-2
  - no systematic (i.e. statistical uncertainty only)

- Clearly we don't want to be over-conservative, nor over-optimistic
  - Some projection clearly limited by systematics with HL-LHC dataset
  - Pile-up effects make projections of systematics harder
  - Systematics may be sub-dominant now but relevant with more data

- How can we project our understanding of systematics?
  - statistics available
  - intrinsic detector limitations
    - new methods may improve our understanding of the detector beyond what we can foresee right now
  - simulation modeling uncertainties rely on theory advances as well

ATLAS projections needs

- Extrapolation from past Run-2 results
  - usually based on existing statistical frameworks
  - capture the full complexity of multi-variables / multi-region analyses
  - use numerous nuisance parameters to capture the deep understanding of systematics with current detector
  - “scale-factors” for event yields can account for expected performance improvements, but no re-optimization possible
  - Best for systematic uncertainties projections is to have <u>projection on individual nuisance parameters</u> (won't go in that detail here, but discuss overall “classes”)
  - Need to pay attention to profiling (over-constraints, correlations,..)

- Truth-based analysis with parametrized detector performance
  - simplified analysis, usually simple cut-and-count but allows re-optimization of selections
  - systematics usually accounted as “flat” numbers (or neglected)
  - Best for systematic uncertainties projections is to have <u>simplified uncertainty estimation for dominant sources</u>

Approach

- Discussion session with performance groups within ATLAS
  - process still ongoing
  - aim to focus on systematics that are most important for the projection studies we need (can't be comprehensive!)
  - derive scaling of main systematic nuisance parameters and overall "uncertainties" as function of X for truth-based projections

- Some common themes/assumptions cross-group
  - statistics-driven sources (data or MC) → 0
  - intrinsic detector limitations stay ~constant
  - simulation modeling uncertainty are halved (?), unless noted
    - specific inputs from theorists very welcome!
  - It was felt that often pile-up challenges will be compensated by algorithmic improvements

- Aim to get a roughly realistic projection
  - i.e. sometimes will be still pessimistic, sometimes may be optimistic

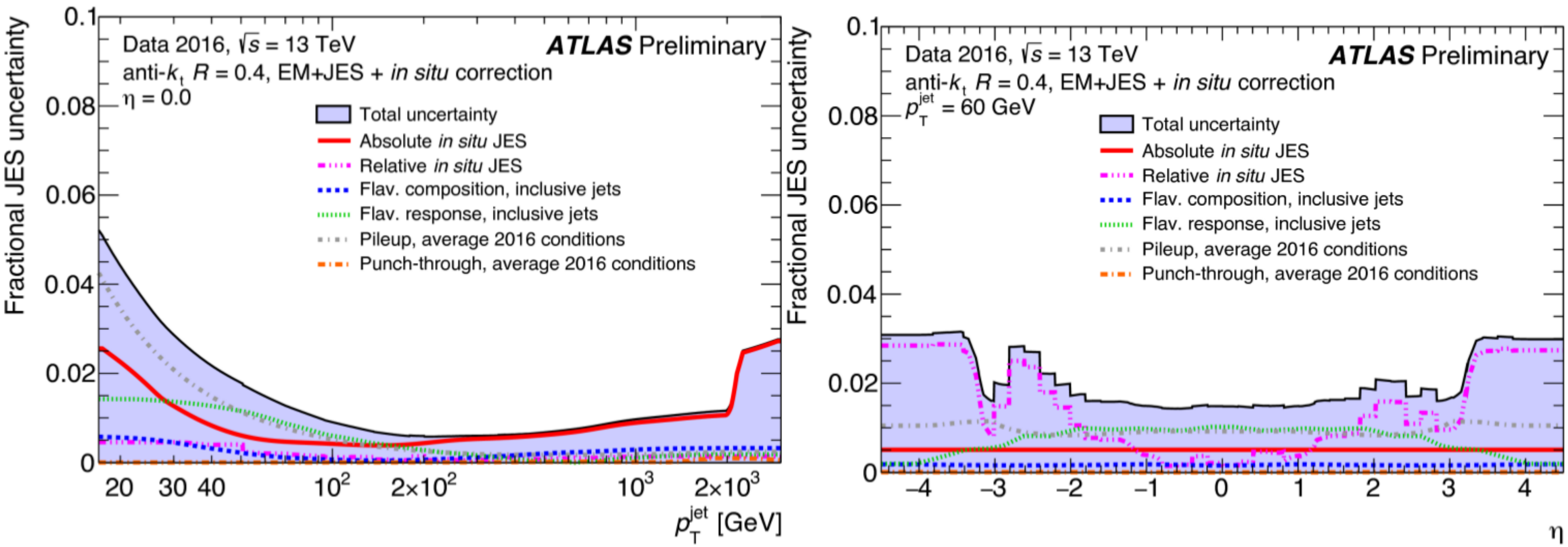
What uncertainties matter?

Topic	Channel	Method	Dominant systematic
Diboson	hyy	Parametrized comb.	PES/PER, JES/JER
Diboson	hWW	Parametrized comb.	
Diboson	hZZ	Parametrized comb.	ggF:leptons, others:JES/JER
Differential	Hbb and STXS	-	
Differential	Hyy and STXS	Run2 extrapolation (TDR)	PES/PER
Differential	H4l and STXS	Parametrized old	
Fermion	VHbb	Partial par.	Jet/MET, BTag
Fermion	Htautau	Partial par.	Jet/MET, Tau
Non-resonant HH	bbyy		Small (E. Petit study)
Non-resonant HH	ttHH (bbbb)	Parametrized	
Non-resonant HH	bbbb	Run2 extrapolation	Multi-jet shape (TH)
Non-resonant HH	bbtautau	Run2 extrapolation	Tau fake
Rare decay	HZy	Parametrized	
Rare decay	Hmumu	Parametrized	
Top yukawa	ttH (all channels)	-	JES/JER, BTag
Top yukawa	ttH (bb)	-	BTag, JES/JER

To be revised

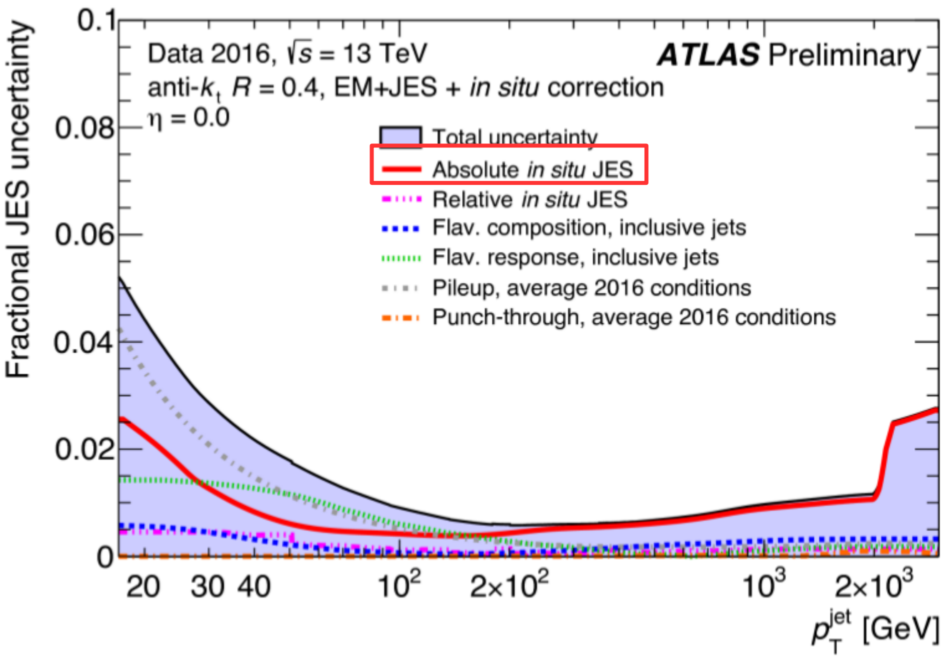
- Example above for a subset of Higgs projections planned
- Most “wanted”: Jet/γ Energy Scale/Resolution, MET, B-tagging, Tau
- Less critical: leptons (e,μ), (hadronic) tracking

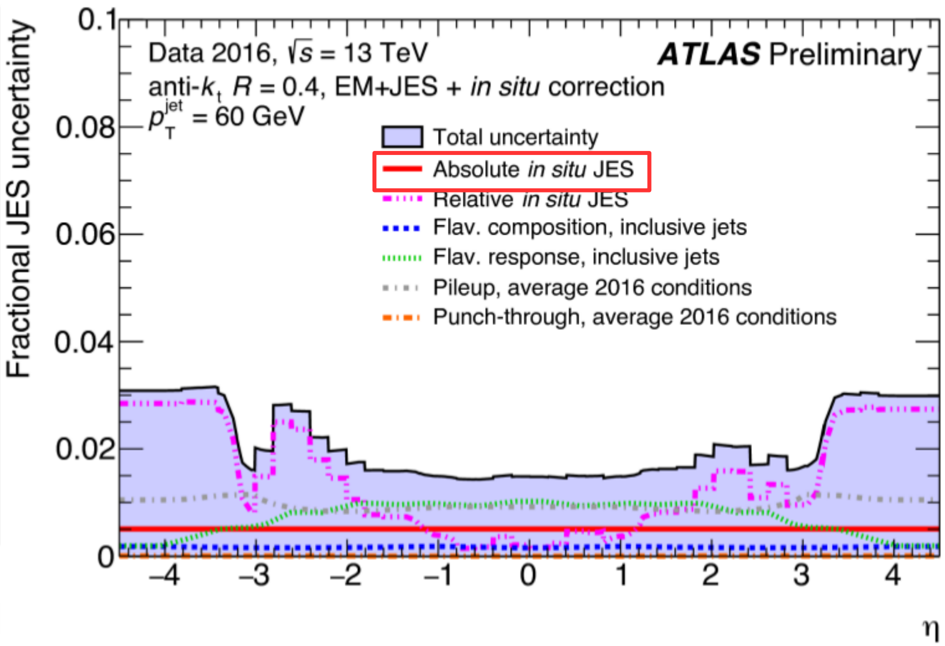
Jet energy scale



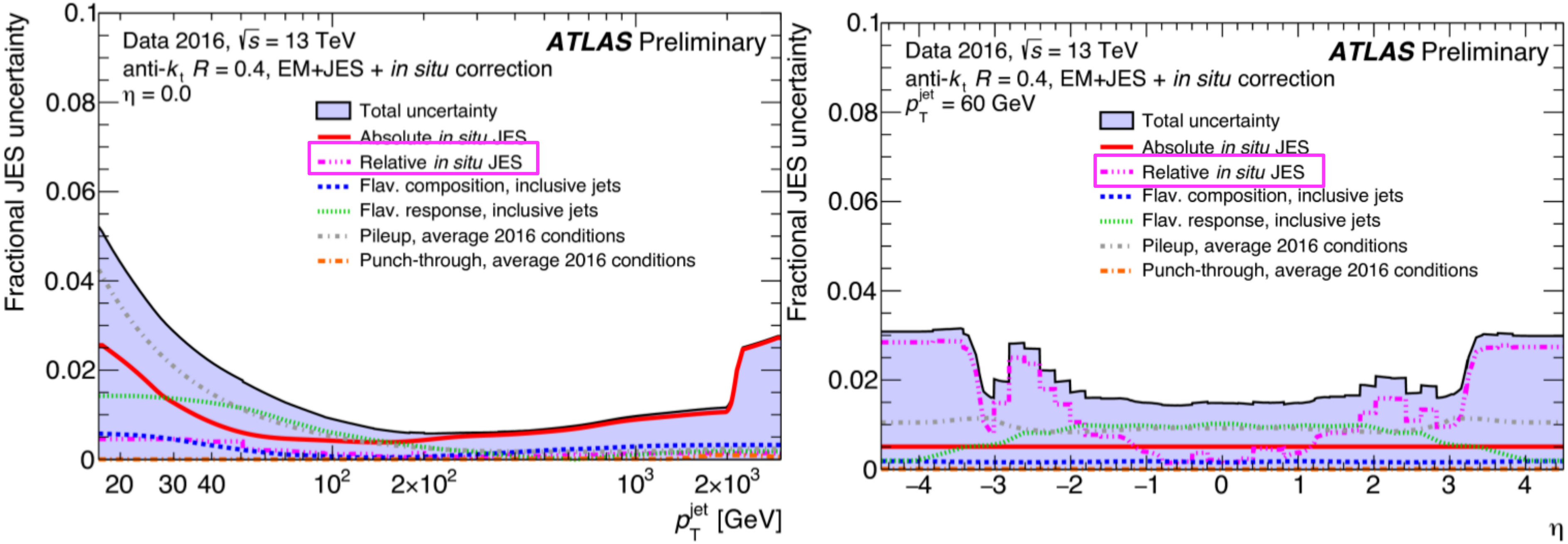
- Starting point: latest run-2 public results
- Will go in a bit more detail for this important systematic

Jet energy scale



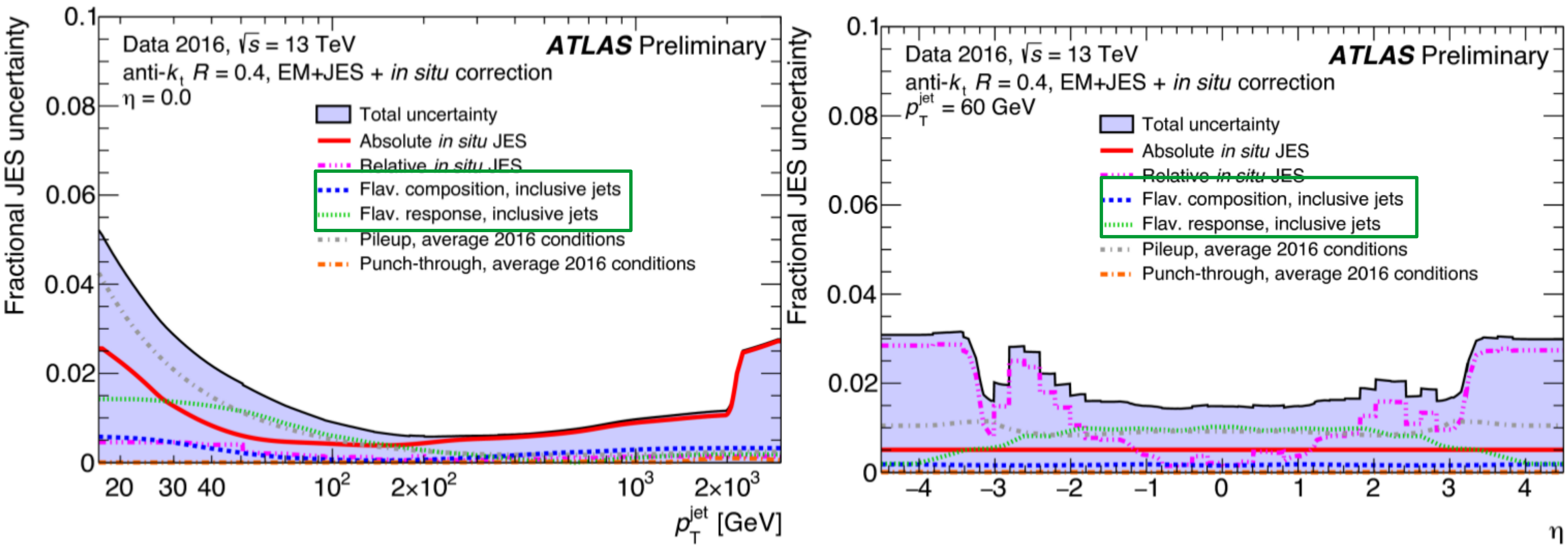


- Absolute "in-situ" JES
  - low-medium p_T from Z+jets balance study
    - dominated by generator differences, pile-up rejection, radiation
    - overall expect improvements to balance challenges → **keep same**
  - high-p_T dominated by photon energy scale in γ+jets balance
    - Expect better accuracy with large statistics → **halved**
  - Other components will be neglected, based on current experience



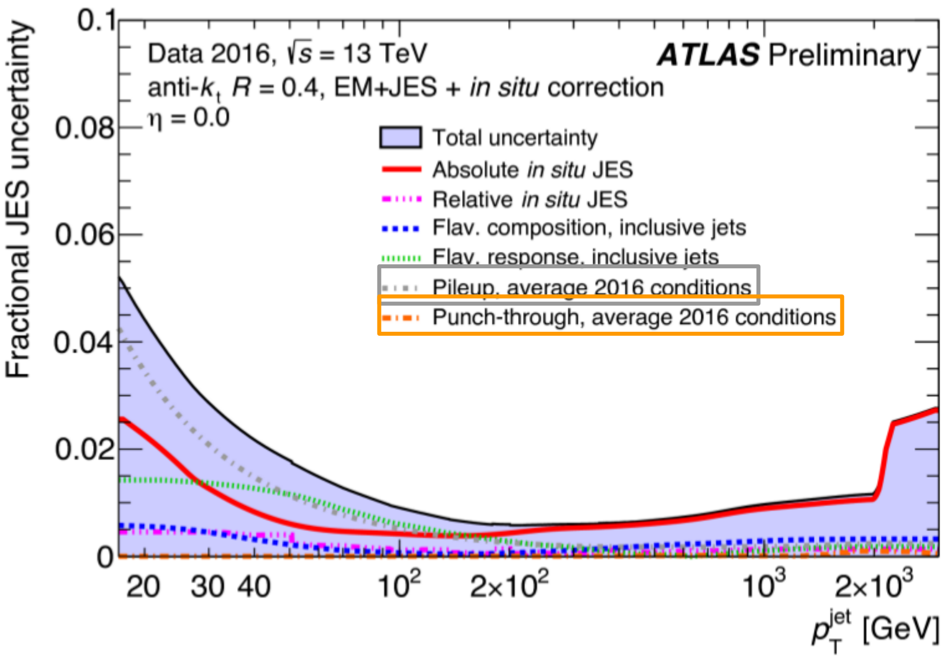
- Relative "in-situ" JES
  - dominated by statistics and simulation modeling
  - in this case it was felt advances in modeling can be substantial
  - Expect it will become negligible → **0**

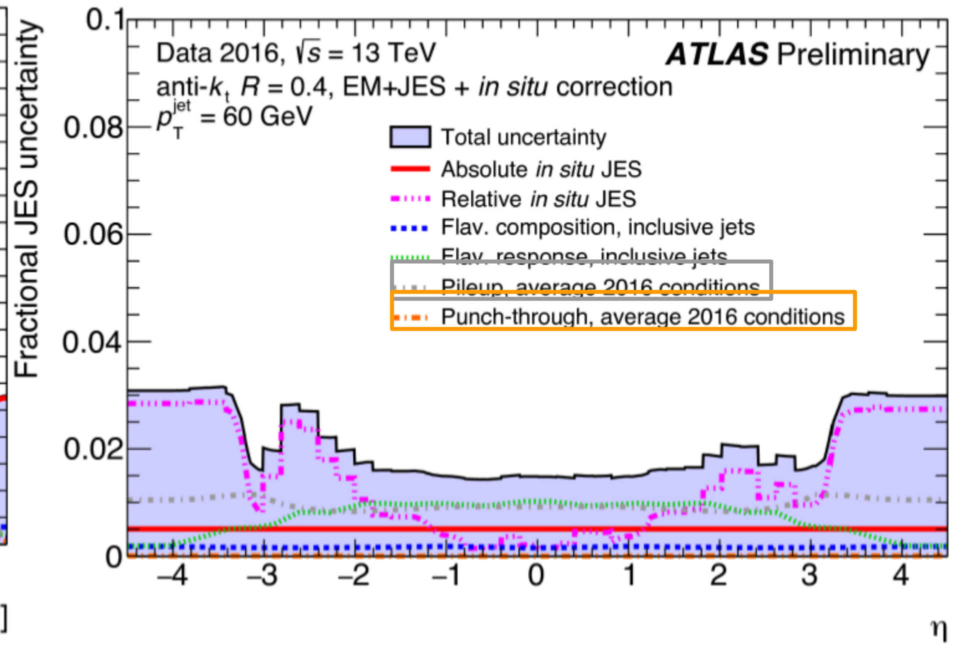
Jet energy scale



- Flavor composition and response
  - mainly comes from how generators model gluon jet radiation
  - rely on fragmentation measurements and re-tuning of parton shower generators
  - Propose to have two scenarios:
    - Optimistic → **halved**
    - Baseline → **keep same**

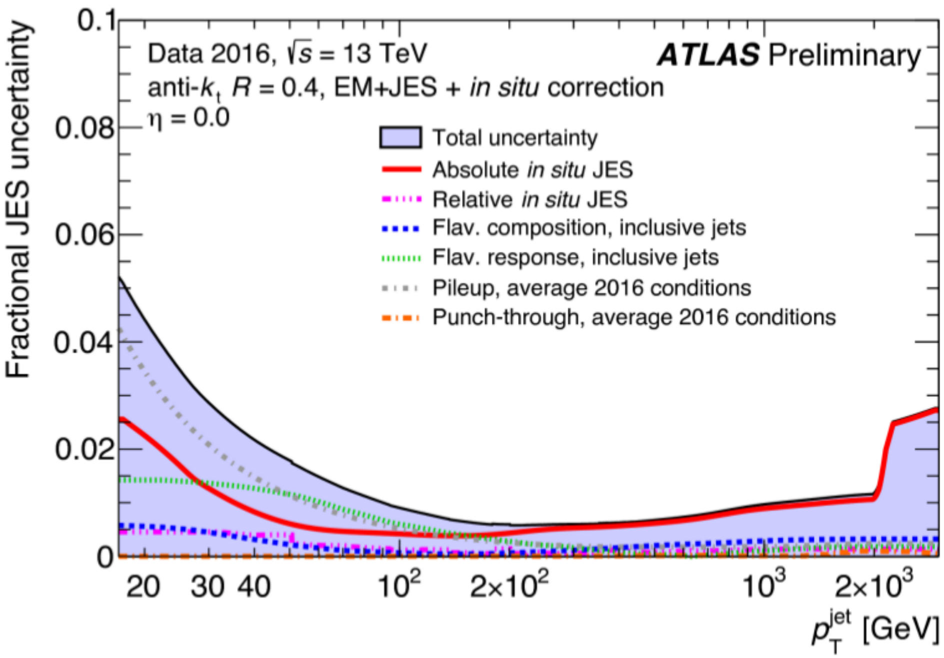
Jet energy scale

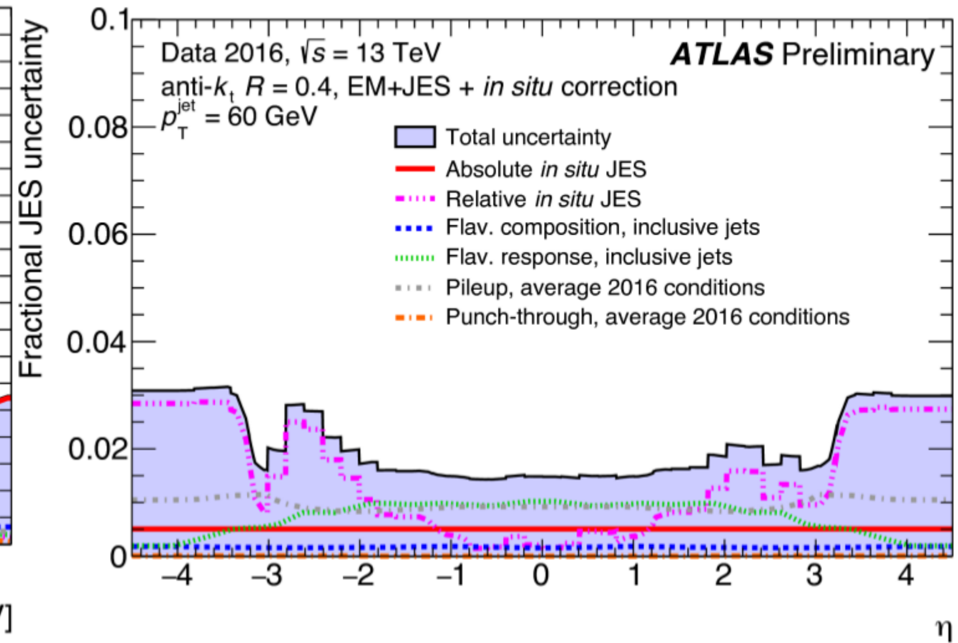




- Pile-up
  - Current method bring an increase uncertainty with pile-up
  - Expect new methods will be developed to at least compensate
  - Two scenarios:
    - Baseline → **keep same**
    - Optimistic → **halved**
- Punch-through, high-pT
  - single particle response but kicks in when we run out of statistics in the multijet balance
  - expect large statistics will allow us to make this negligible → **0**

Jet energy scale

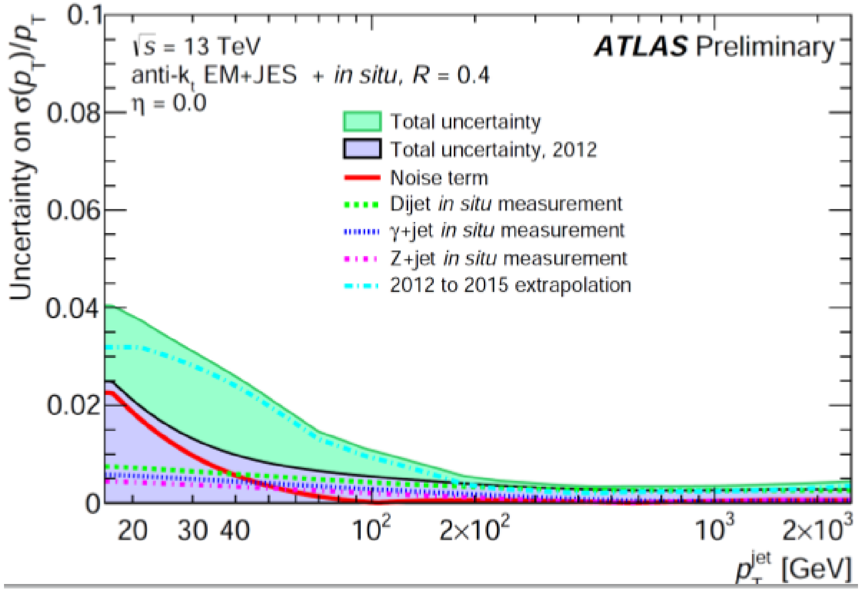




- Overall we have now a clear recipe to scale all individual nuisance parameters for run-2 extrapolations
- We will then produce “summary” plots (like the ones above) for truth-based projections for which this is an important uncertainty
  - Applied as additional shift to the true (smeared) value

Jet Energy Resolution / MET

- JER currently larger than Run-1
- Expect to reduce it to similar levels
  → **halved** (reaching run-1 values)



- MET systematics driven by object scale/resolution uncertainties
- Soft-term uncertainties are rarely dominant and hard to extrapolate
  → **keep same**
  - discuss exceptions on a case-by-case

Flavor tagging

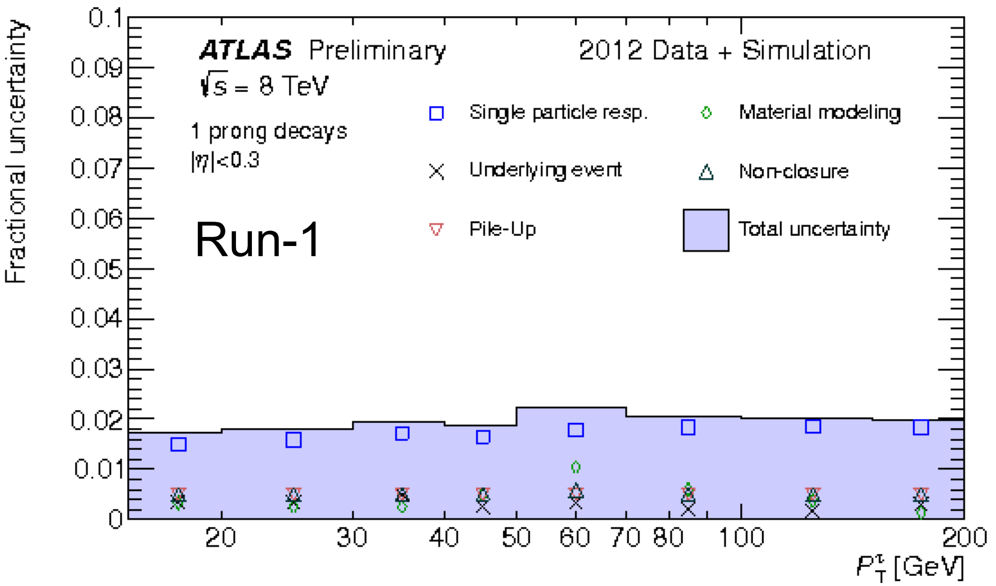
- Data/MC scale-factor measured using in-situ techniques

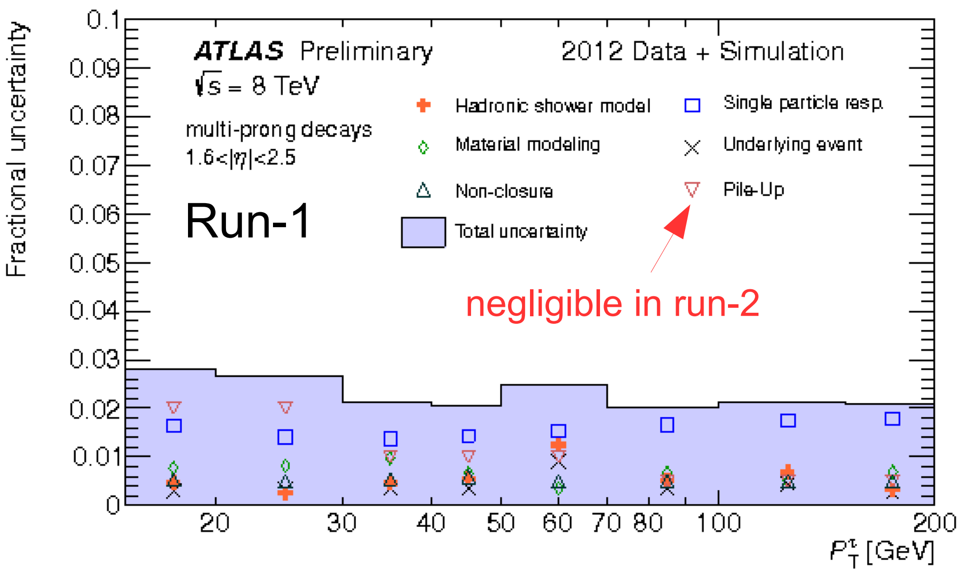
Calibration method	Dominating uncertainties
b-tagging ttbar PDF	Modeling: Pythia vs PowHeg vs aMC@NLO..
b-tagging ttbar tag-n-probe	Modeling: parton showering, matrix element, ISR/FSR
c-tagging W+c	Jet/MET reconstruction, modeling (c-flavor decays), MC statistics
c-tagging ttbar→c	Modeling, MC+data statistics, light tag SF
mistag negative tagging	track uncertainties (IP, fakes), HF fraction, MC+data statistics (low jet pT)
mistag adjusted MC	track uncertainties (IP, fakes)

- Statistics-dominated uncertainties → 0
- MC modeling uncertainties → **halved** (optimistic) [,**same** (baseline)]
  - currently mainly from generator comparisons
- Expect new methods to perform better at high-pT/η
  - current uncertainties too large → **under discussed** (CMS?)

Tau

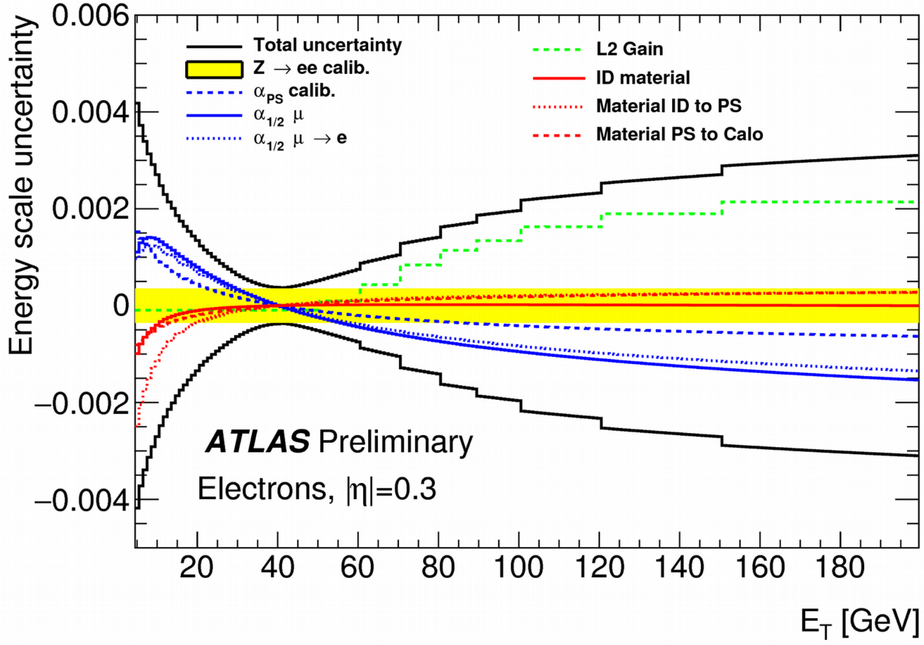
- Most important components:
  - ID efficiency
  - Tau Energy Scale
  - others less important
    → **neglected**
- Tau ID
  - Mostly limited by systematics
    - Simulation τ modeling
    - Fakes background
  - Expect **"floor" of ~ 5%**
  - Under discussion p_T > 250 GeV
- Tau Energy Scale
  - Theory modeling, detector, in-situ
  - Expect **"floor" of ~ 2-3%**
  - Under discussion for high pT





Photon/Electrons

- ID efficiency for e/γ already at 0.5% → **keep same**
  - however low-p_T regime has larger uncertainties
- Energy scale calibration 0.1%(0.2%) to 0.3%(0.5%) for e (γ) → **keep same**
  - larger dataset will help in monitoring detector stability
  - expect to be able to mitigate larger pile-up effects
  - critical understanding of detector seems difficult to go much furth



- Energy resolution
  - About 10% @ 60 GeV, ultimately dominated by basic detector knowledge
  - Pile-up modeling will play a role, but expect to mitigate it → **keep same**

Muons

- Reconstruction + ID efficiency well known (~0.1%)
- Scale and resolution also well measured

- Most of these measurements are systematically limited but robust against pile-up

- Baseline plan is to keep current uncertainties
  - if a particular analysis is limited by this, can be revised
  - note: measurements as m(W) will rely on dedicated low-μ datasets

- Very high momentum muons (~TeV) have larger uncertainties
  - under investigation impact on analyses (e.g. Z')
  - if relevant, will probably try to be more optimistic than now (e.g. assuming dedicated toroid-off runs for alignment, ...)

Background modeling uncertainties

- MC modeling enters directly into uncertainties on backgrounds too
  - likely too wide of a problem, it needs discussion on a case-by-case
  - or could attempt to define a common assumption?
    → **as-is vs ½ comparison**

- Data-driven background can be limited by
  - statistics in control region → will get better with **~sqrt(L)**
  - closure of method → harder to improve, **keep same**

- MC statistical uncertainty is assumed → **0**

- Reality is less black&white than above and requires some judgments to be done on a case by case, but guidelines above could still be useful

Conclusions

- Wide range of **experimental systematics**, but no need to cover all

- Extrapolation of existing analyses:
  - assumption on scaling of systematics as nuisance parameters
  - care to be taken to ensure no over-constraint, if it happens can think of applying scaling “post-fit” (manually)
- Truth-based analyses:
  - simplified parametrization vs pT, η, … for dominant uncertainties

- Overall approach philosophy (with exceptions) for discussion:
  - statistical (data/MC) uncertainty usually → 0
  - physics simulation modeling → ½ (comparing w/ no scaling)
  - often new methods are expected to compensate for increased pile-up effects