Comparisons between Event Generators and Data

Peter Richardson

IPPP, Durham University

Cambridge 19th April

Summary

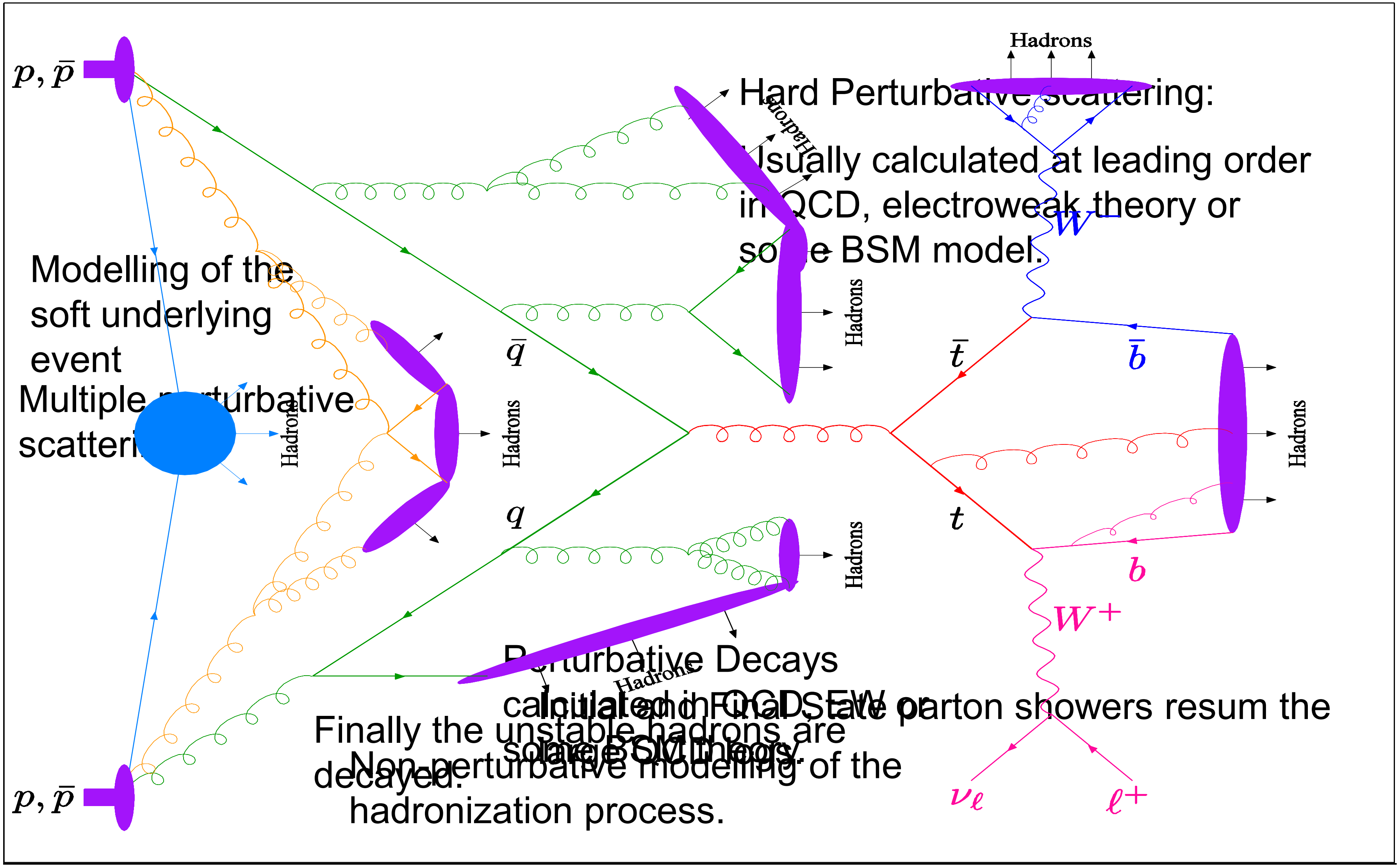
- Introduction
- Basics Of Event Generation
- Multiple Parton-Parton Scattering
- Jets
- Vector Bsons with Jets
- Conclusions

Most plots taken from talks at last weeks SM@LHC and DIS meetings.

Introduction

- Monte Carlo event generators are designed to simulate hadron collisions using a combination of:
  - Fixed order perturbative calculations;
  - Resummation of large QCD logarithms;
  - Phenomenological Models.
- It’s important to understand the different pieces of the simulation.
- Some are on firm theoretical ground and we’d be surprised if they didn’t work, others might break down in the new energy regime of the LHC.

A Monte Carlo Event



Hard Perturbative scattering:

Usually calculated at leading order in QCD, electroweak theory or some BSM model.

Modelling of the soft underlying event

Multiple perturbative scattering

Perturbative Decays calculated in QCD, EW or some BSM theory.

Initial and Final State parton showers resum the large QCD logs.

Finally the unstable hadrons are decayed.

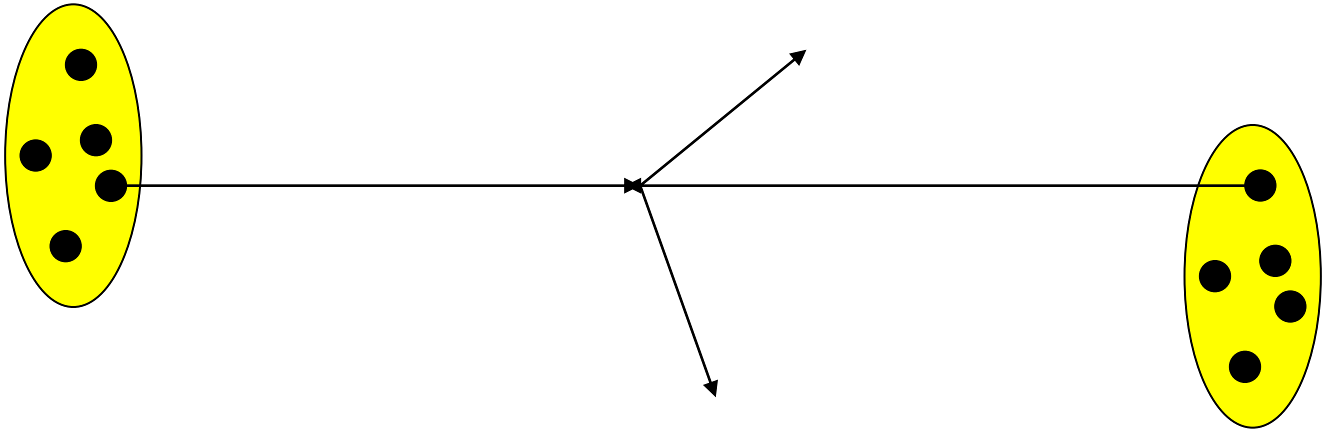
Non-perturbative modelling of the hadronization process.

Introduction

- The different models are generally tuned to different types of data:
  - parameters relating to the final-state parton shower and hadronization are tuned to LEP data;
  - parameters relating to initial-state parton showers and multiple parton-parton interactions are tuned to data from the Tevatron and UA5.
- We expected that the shower and hadronization models would work at LHC energies, less sure about the underlying event.

The Underlying Event

- Protons are extended objects.
- After a parton has been scattered out of each in the hard process what happens to the remnants?



Two Types of Model:

1) Non-Perturbative: Soft parton-parton cross section is so large that the remnants always undergo a soft collision.

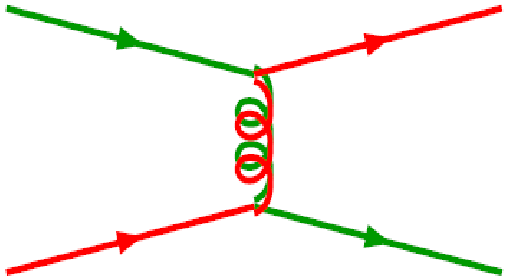
2) Perturbative: 'Hard' parton-parton cross section is huge at low p_T, dominates the inelastic cross section and is calculable.

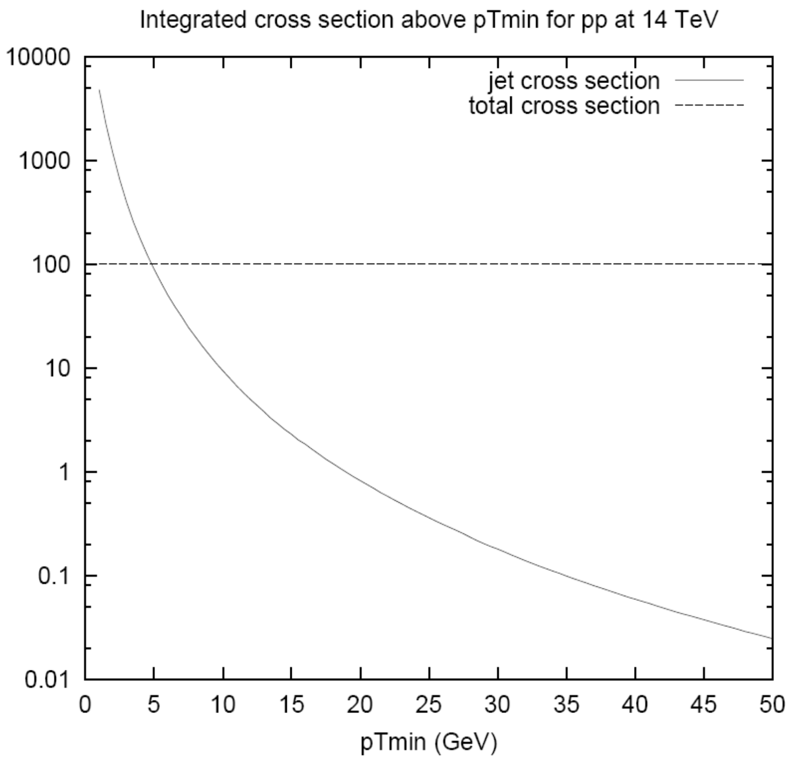
Multiparton Interaction Models

- The cross-section for 2→2 scattering is dominated by t-channel gluon exchange.
- It diverges like

$$\frac{\mathrm{d}\sigma}{\mathrm{d}p_\perp^2} \approx \frac{1}{p_\perp^4} \quad \text{for} \quad p_\perp \to 0$$

- This must be regulated used a cut of p_{Tmin}.
- For small values of p_{Tmin} this is larger than the total hadron-hadron cross section.
- More than one parton-parton scattering per hadron collision



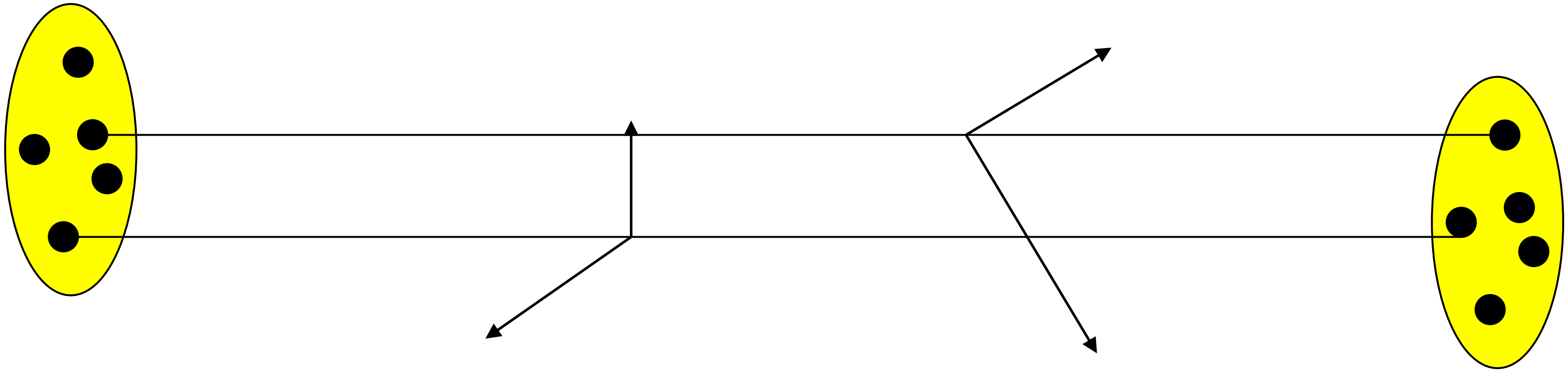


Multiparton Interaction Models

- If the interactions occur independently then follow Poissonian statistics

$$P_n = \frac{\langle n \rangle^n}{n!} e^{-\langle n \rangle}$$

- However energy-momentum conservation tends to suppressed large numbers of parton scatterings.



- Also need a model of the spatial distribution of partons within the proton.

Multiparton Interaction Models

- In general there are two options for regulating the cross section.

$$\frac{\mathrm{d}\hat{\sigma}}{\mathrm{d}p_\perp^2} \propto \frac{\alpha_s^2\left(p_\perp^2\right)}{p_\perp^4} \rightarrow \frac{\alpha_s^2\left(p_\perp^2\right)}{p_\perp^4}\theta\left(p_\perp - p_{\perp\mathrm{min}}\right) \quad \text{simpler}$$

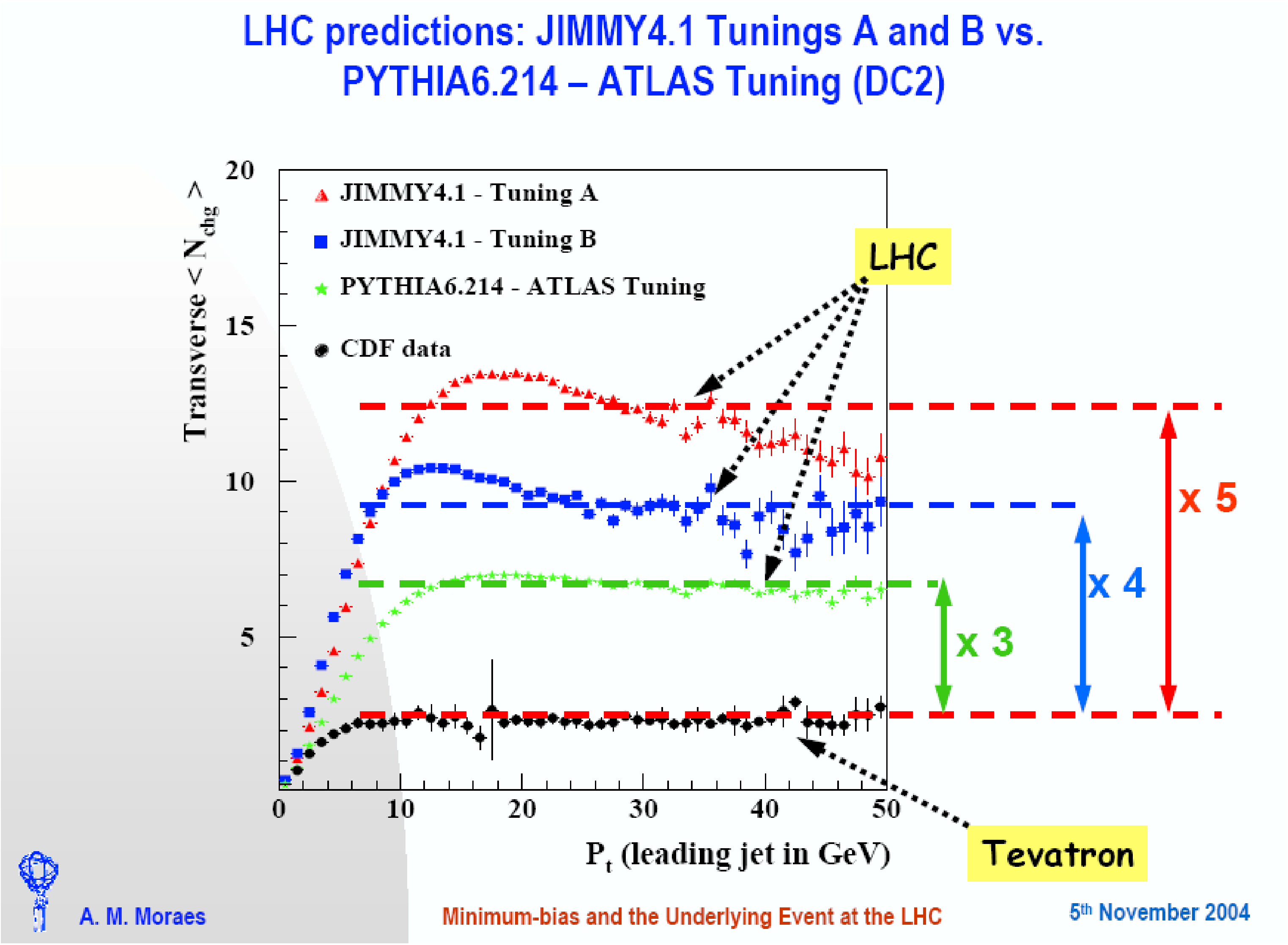
$$\text{or} \quad \rightarrow \frac{\alpha_s^2\left(p_\perp^2 + p_{\perp 0}^2\right)}{\left(p_\perp^2 + p_{\perp 0}^2\right)^2} \qquad \text{more complicate d}$$

  where $p_{\perp\mathrm{min}}$ or $p_{\perp 0}$ are free parameters of order 2 GeV.
- Typically 2-3 interactions per event at the Tevatron and 4-5 at the LHC.
- However tends to be more in the events with interesting high p_T ones.

Prior to LHC

- Before the LHC data from:
  - UA5 experiment;
  - CDF at 630, 1800 and 1960 GeV.

  were used to constrain the parameters of the underlying event model.
- The data at the higher Tevatron energies is the best for tuning the parameters at a specific energy.
- Need the other points to extrapolate the parameters to LHC energies.



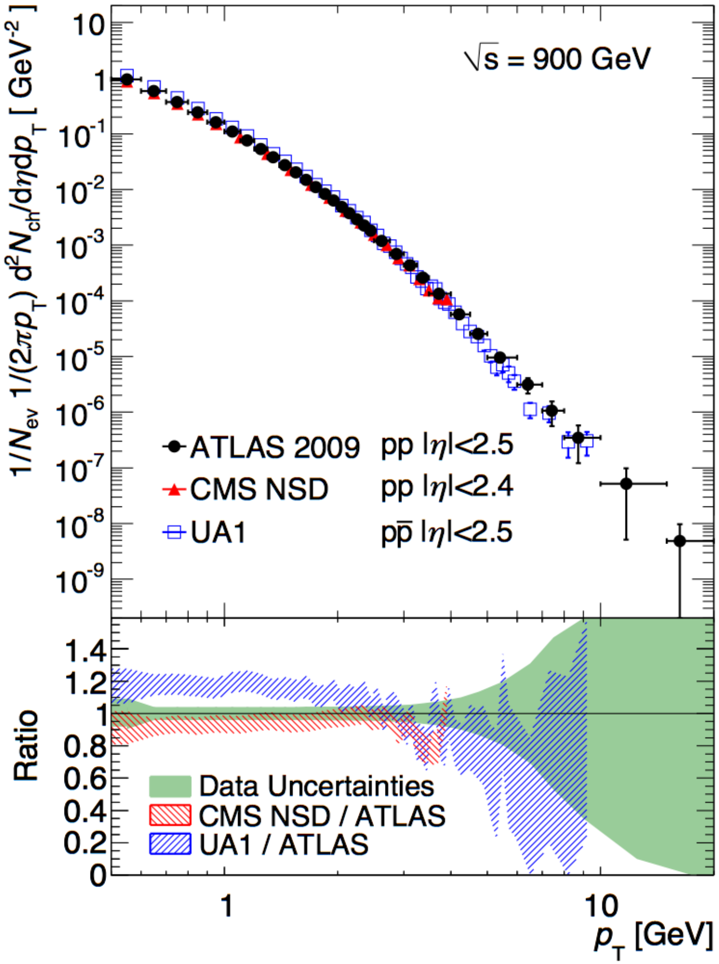
LHC predictions: JIMMY4.1 Tunings A and B vs. PYTHIA6.214 – ATLAS Tuning (DC2)

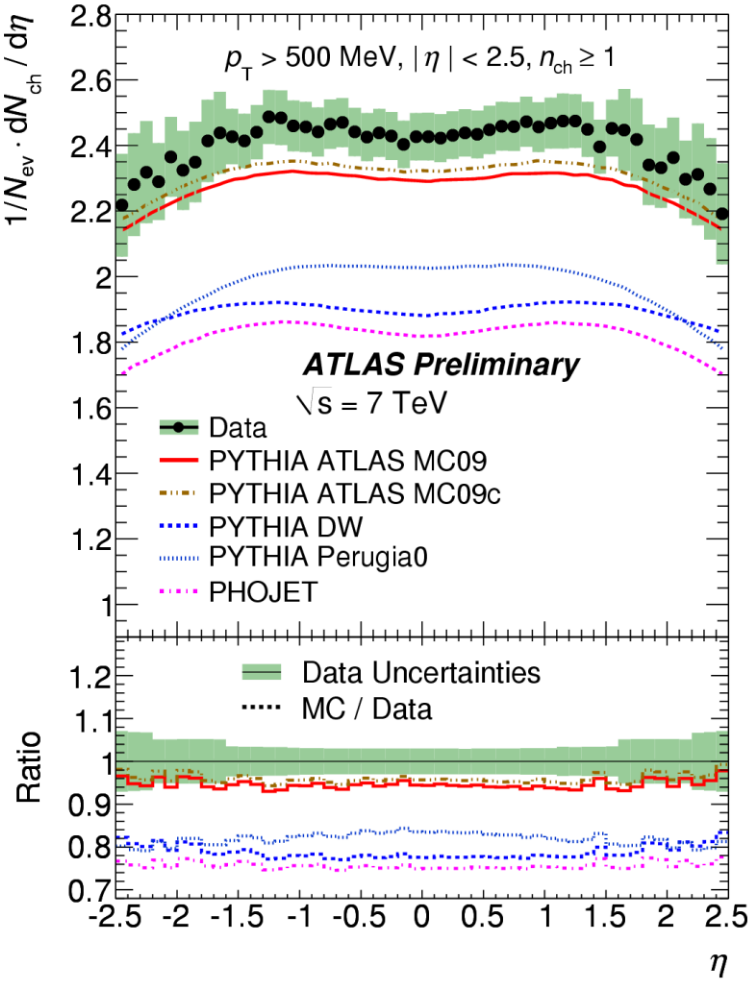
A. M. Moraes

Minimum-bias and the Underlying Event at the LHC

5th November 2004

Charged Particle Multiplicities at √s=0.9, 7 TeV



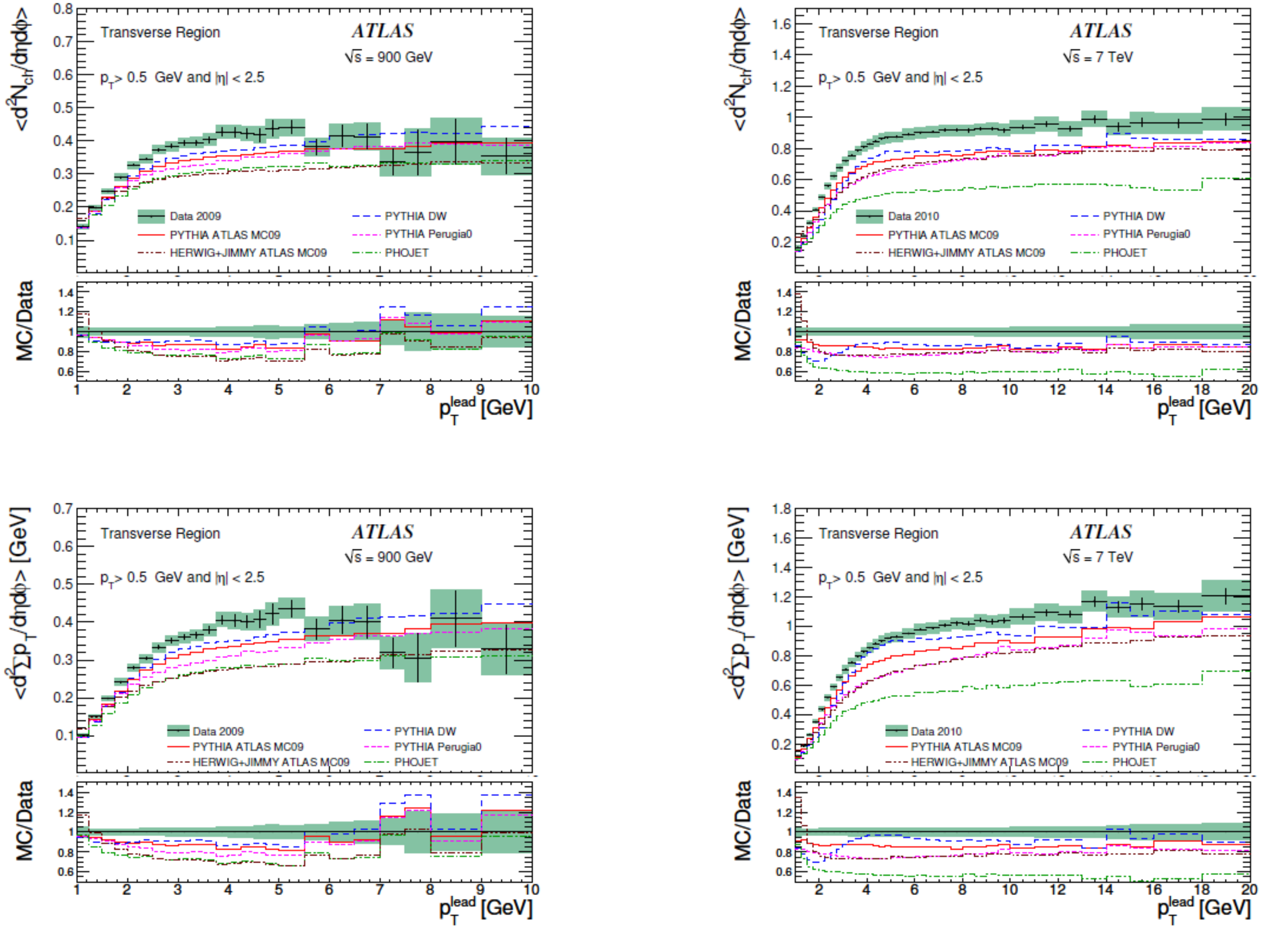


Monte Carlo underestimates the track multiplicity seen in ATLAS

ATLAS underlying event results

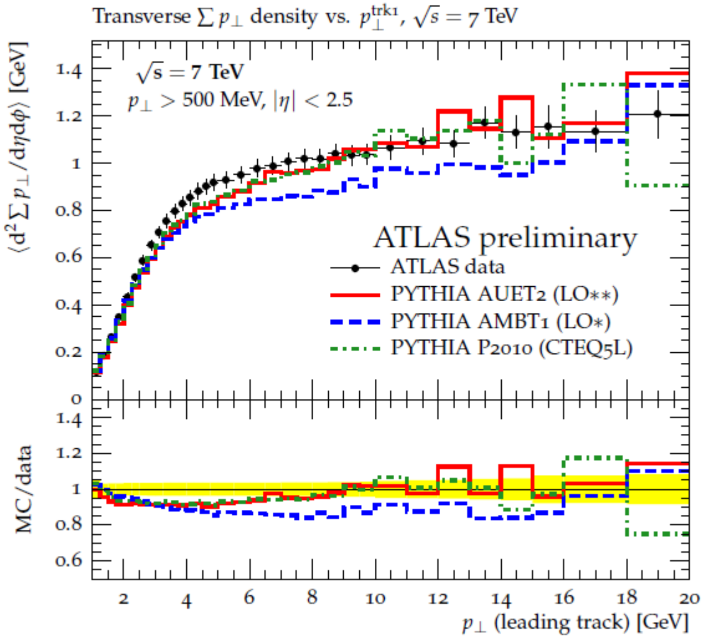
Leading charged track N_{ch} and $\sum p_T$, $p_T > 500$ MeV (arXiv:1012.0791)

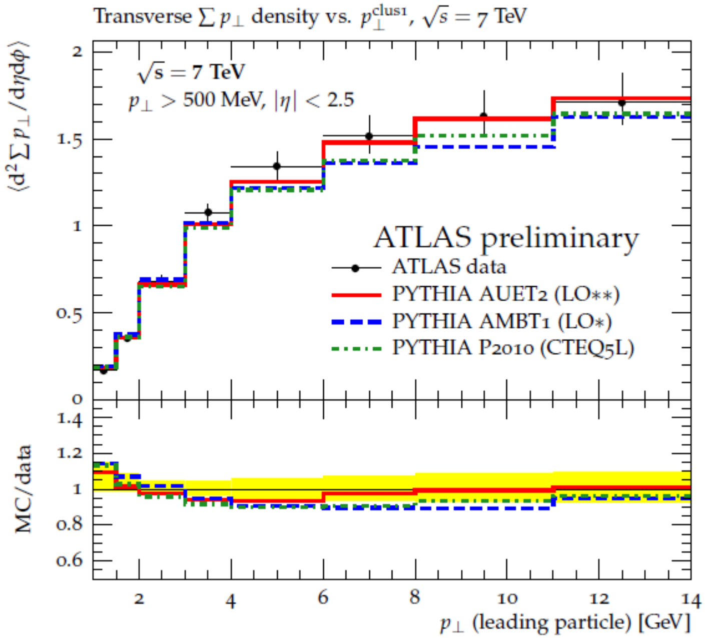
900 GeV 7 TeV



A. Buckley SM@LHC, Durham

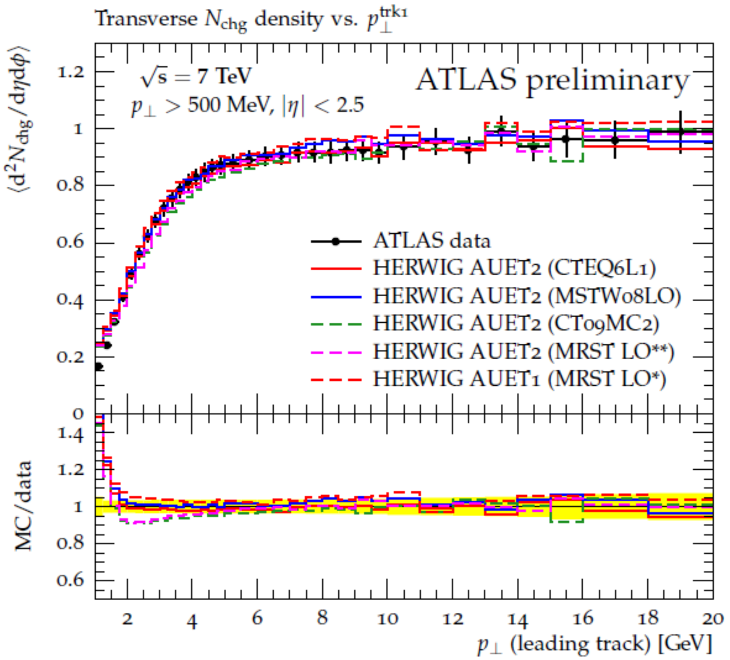
New Pythia 6 tune to ATLAS data: AUET2

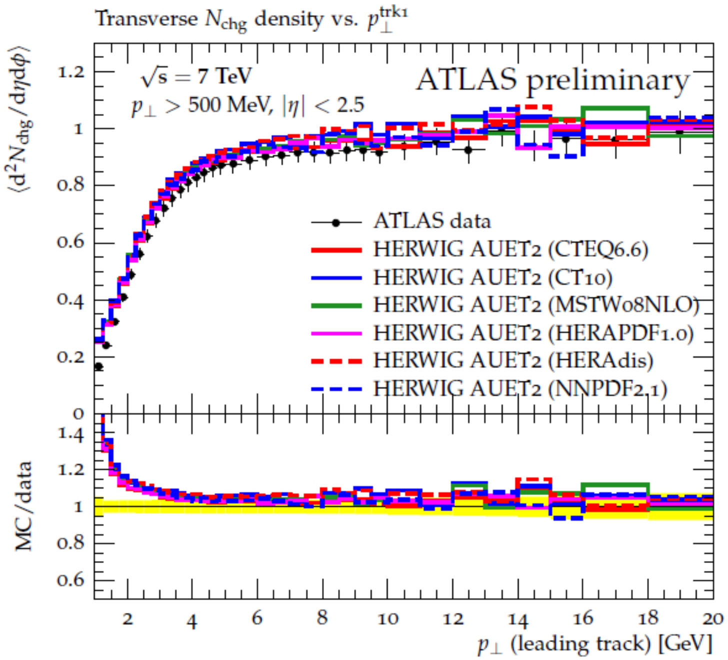




- Improvement w.r.t. AMBT1, “turn-over region” undershot
- Similar agreement for track- and calorimeter based UE measurements

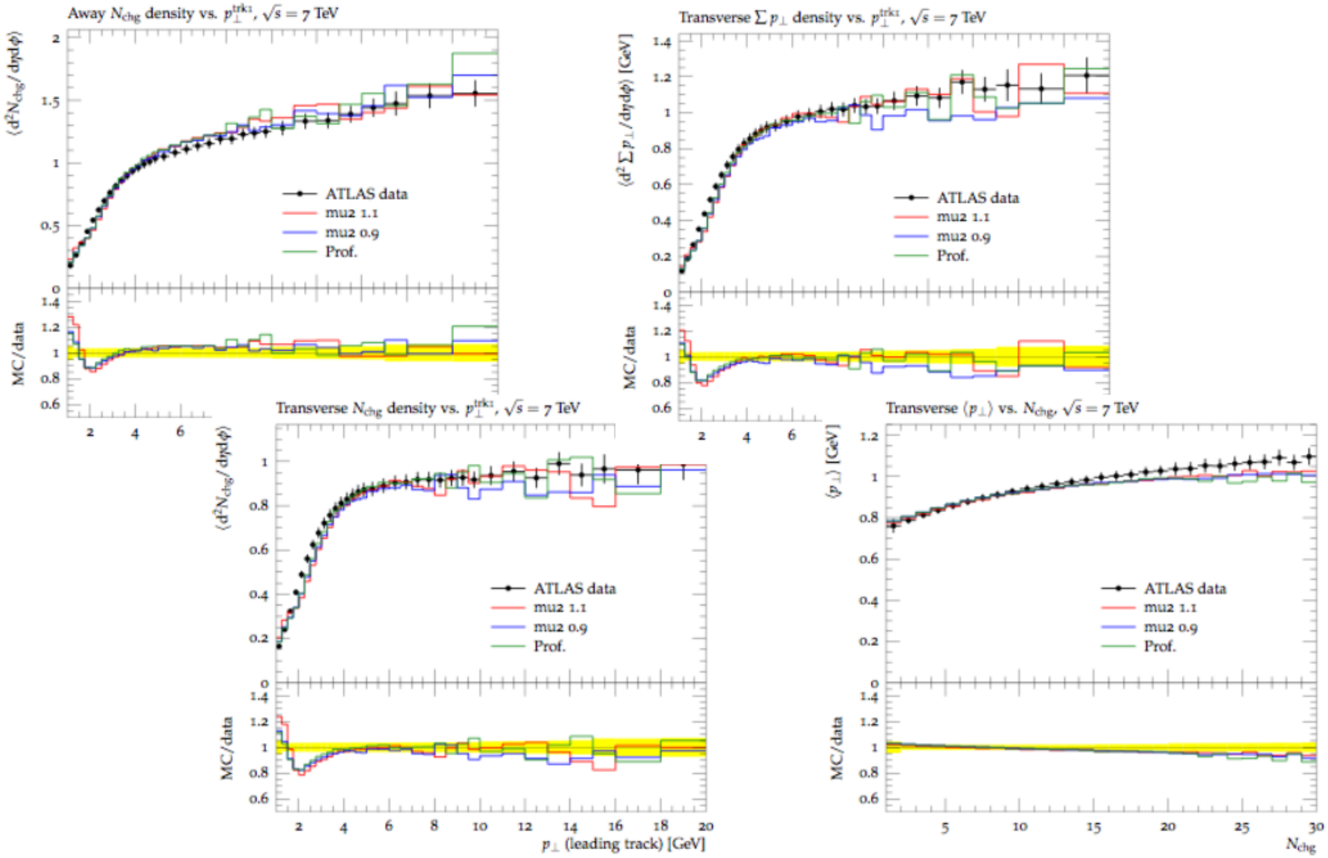
7 TeV ATLAS UE, $p_\perp^{\text{TRACK}} > 500$ MeV TRACKS





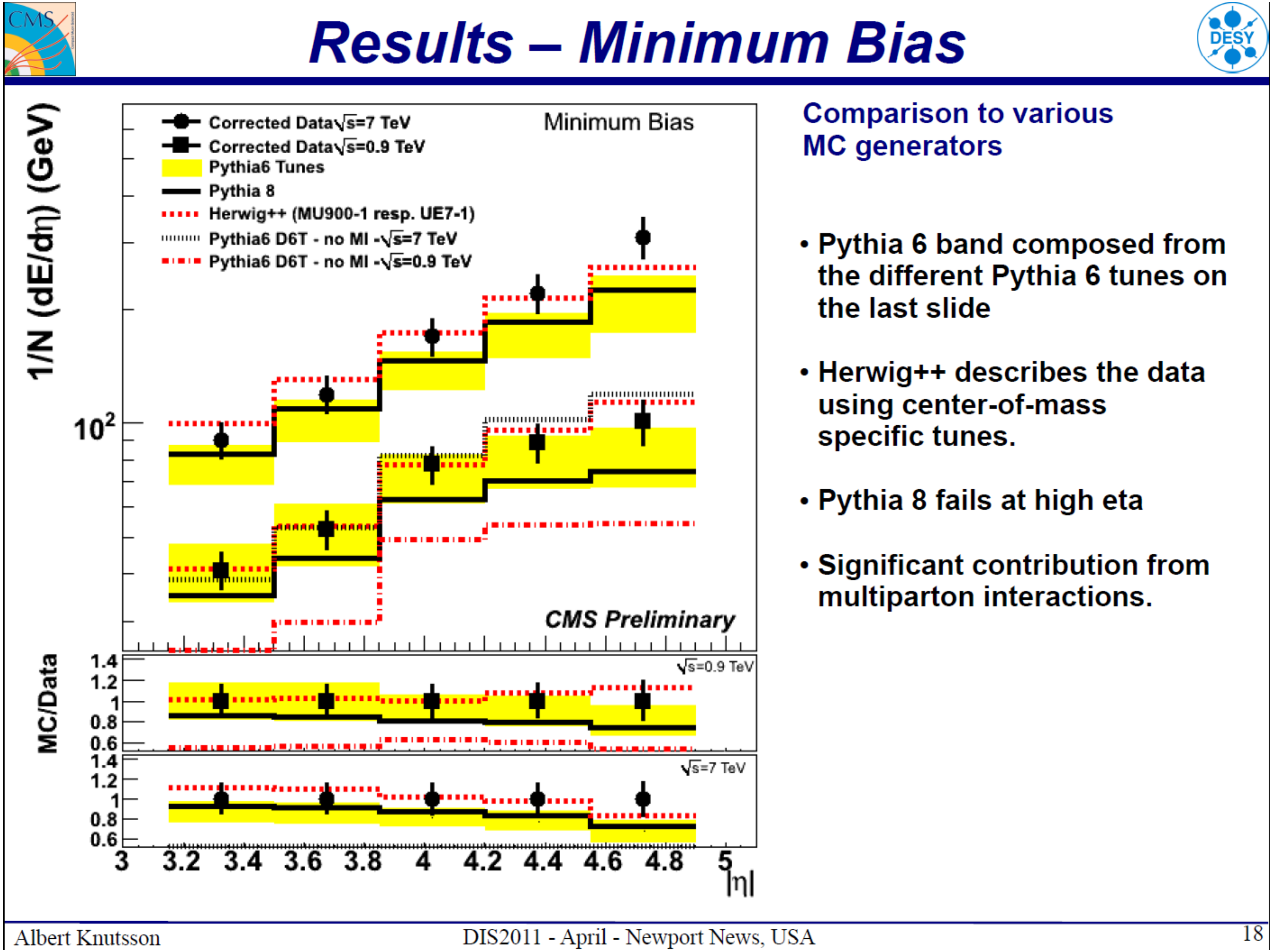
- l.h.s: LO and mLO PDFs r.h.s: NLO PDFs
- PDF can almost be “tuned away” in UE
- Tuning results cluster in MPI cut-off `PTJIM` as function of PDF-type

Underlying event at 7000 GeV



SM@LHC

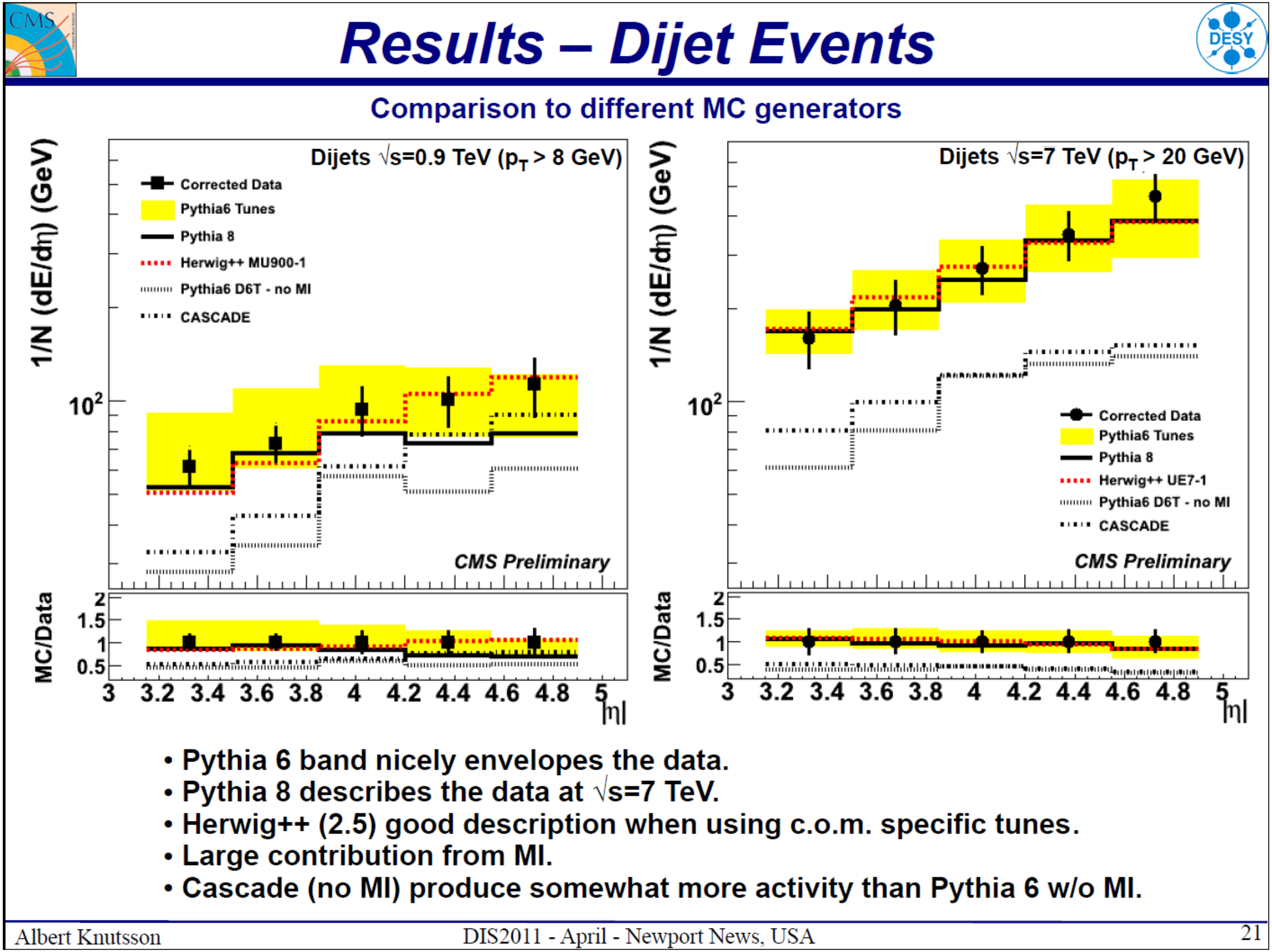
Mike Seymour



Results – Minimum Bias

Comparison to various MC generators

- Pythia 6 band composed from the different Pythia 6 tunes on the last slide
- Herwig++ describes the data using center-of-mass specific tunes.
- Pythia 8 fails at high eta
- Significant contribution from multiparton interactions.



Results – Dijet Events

Comparison to different MC generators

- Pythia 6 band nicely envelopes the data.
- Pythia 8 describes the data at √s=7 TeV.
- Herwig++ (2.5) good description when using c.o.m. specific tunes.
- Large contribution from MI.
- Cascade (no MI) produce somewhat more activity than Pythia 6 w/o MI.

Albert Knutsson DIS2011 - April - Newport News, USA 21

Multiple Parton Scattering

- Results are encouraging.
- The results of the tunes made before data taking don’t exactly agree with the data but aren’t orders of magnitude off.
- Including the new results in the fitting gives good agreement.
- The models therefore seem reasonable, although some theoretical tweaking, e.g. colour reconnection in Herwig++ required, but not a major rethink of the whole approach.

Improving the Simulations

- Prior to the LHC there was a lot of theoretical work designed to improve parton showers by merging the results:
  - with NLO calculations giving the correct NLO cross section and description of the hardest emission (MC@NLO Frixione and Webber, POWHEG Nason);
  - with LO matrix elements to give the correct description of many hard emissions (MLM and CKKW);
  - Combining both approaches MENLOPS Hamilton and Nason.

NLO Simulations

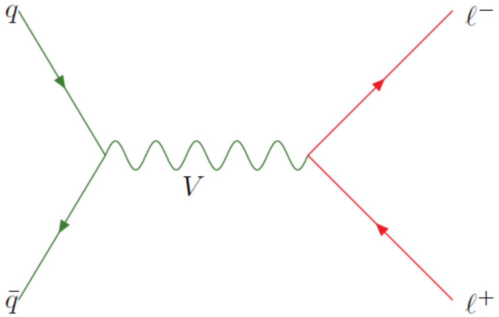
- NLO simulations rearrange the NLO cross section formula.

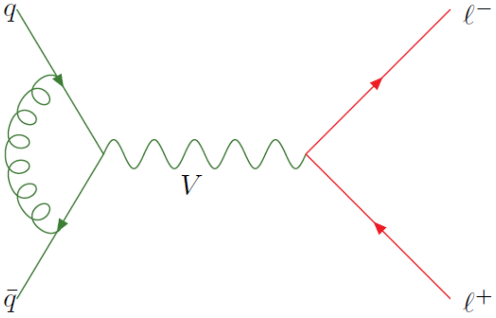
$$d\sigma = B(v)d\Phi_v + (V(v) + C(v,r)d\Phi_r)d\Phi_v$$
$$+ (R(v,r) - C(v,r))d\Phi_v d\Phi_r$$

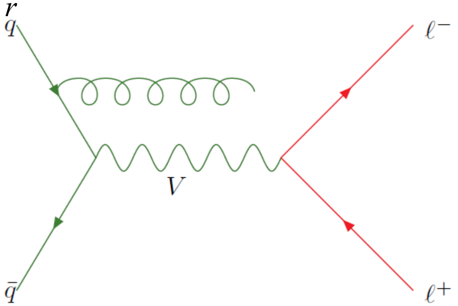
- Either choose C to be the shower approximation

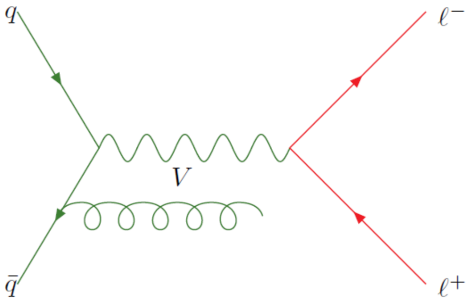
$$d\sigma = B(v)d\Phi_v + (V(v) + C_{\text{shower}}(v,r)d\Phi_r)d\Phi_v$$
$$+ (R(v,r) - C_{\text{shower}}(v,r))d\Phi_v d\Phi_r$$

MC@NLO (Frixione, Webber)









NLO Simulations

- Or a more complex arrangement POWHEG(Nason)

$$d\sigma = \bar{B}(v)d\Phi_v\left[\Delta_R^{NLO}(0)+\Delta_R^{NLO}(p_T)\frac{R(v,r)}{B(v)}d\Phi_r\right]$$

where

$$\bar{B}(v) = B(v)+V(v)+C(v,r)\ \ d\Phi_r+\int R(v,r)-C(v,r)\ \ d\Phi_r$$

$$\Delta_R^{NLO}(p_T) = \exp\left[-\int d\Phi_r\frac{R(v,r)}{B(v)}\theta\left(k_T(v,r)-p_T\right)\right]$$

Pros and Cons

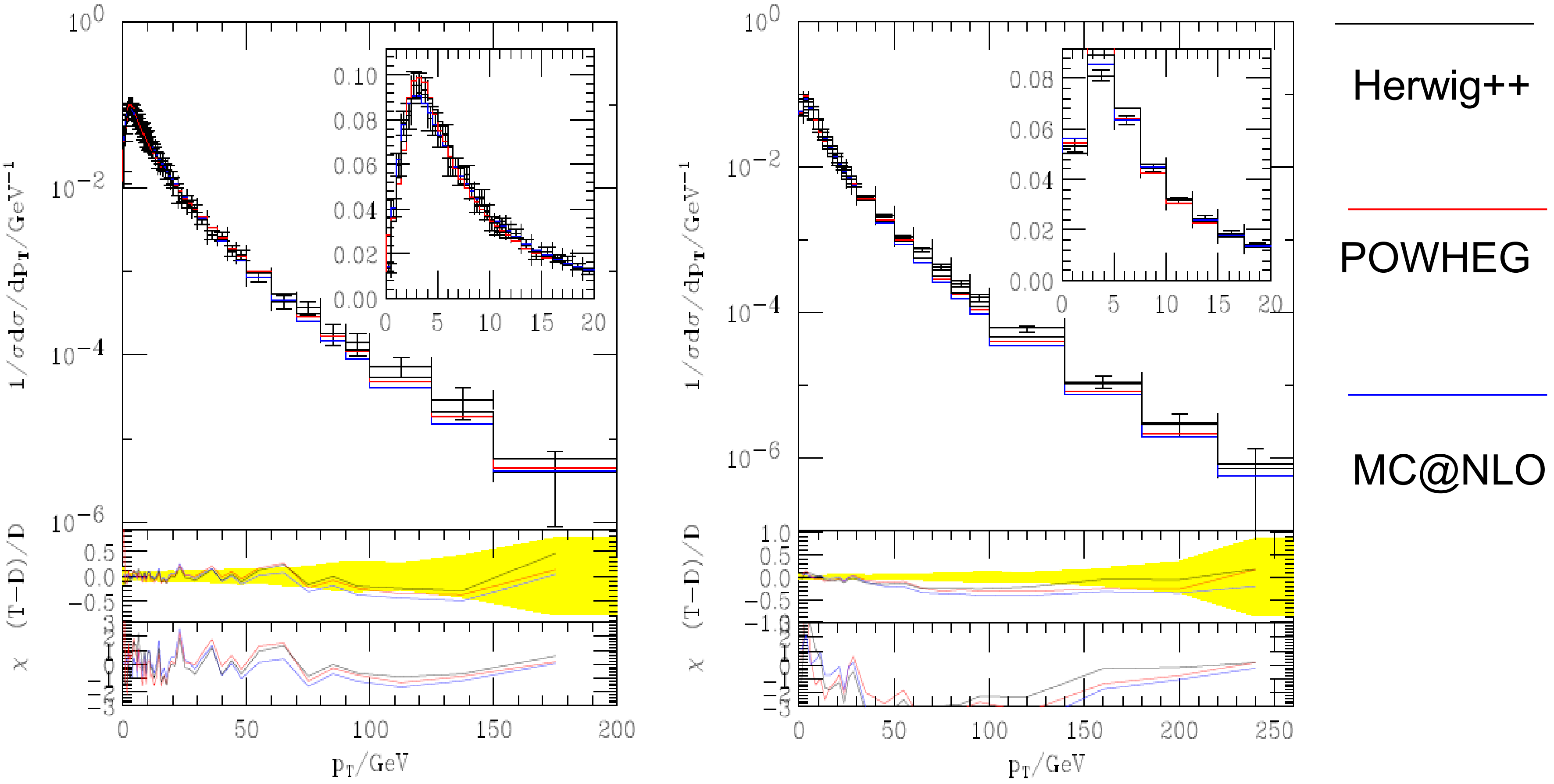
POWHEG

- Positive weights.
- Implementation doesn’t depend on the shower algorithm.
- Needs changes to shower algorithm for non-p_T ordered showers.
- Differs from shower and NLO results, but changes can be made to give NLO result at large p_T.

MC@NLO

- Negative weights
- Implementation depends on the specific shower algorithm used.
- No changes to parton shower.
- Reduces to the exact shower result at low p_T and NLO result at high p_T

Drell Yan



CDF Run I Z p_T

D0 Run II Z p_T

JHEP 0810:015,2008 Hamilton, PR, Tully

Different Approaches

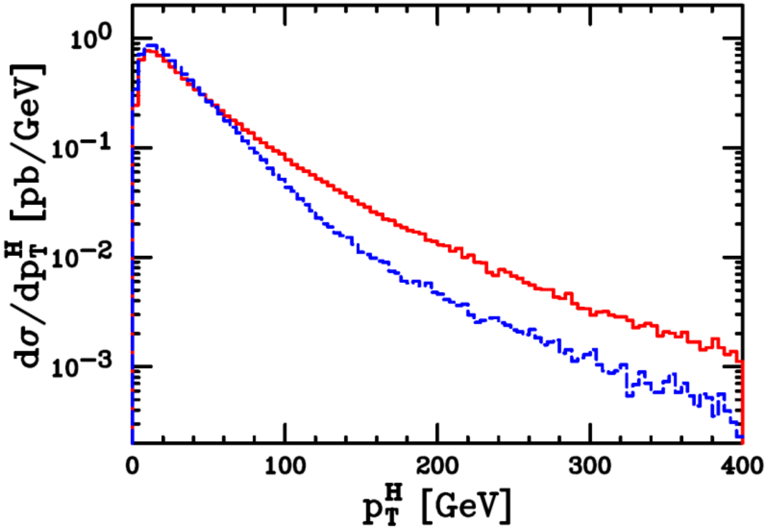
- The two approaches are the same to NLO.
- Differ in the subleading terms.
- In particular at large p_T

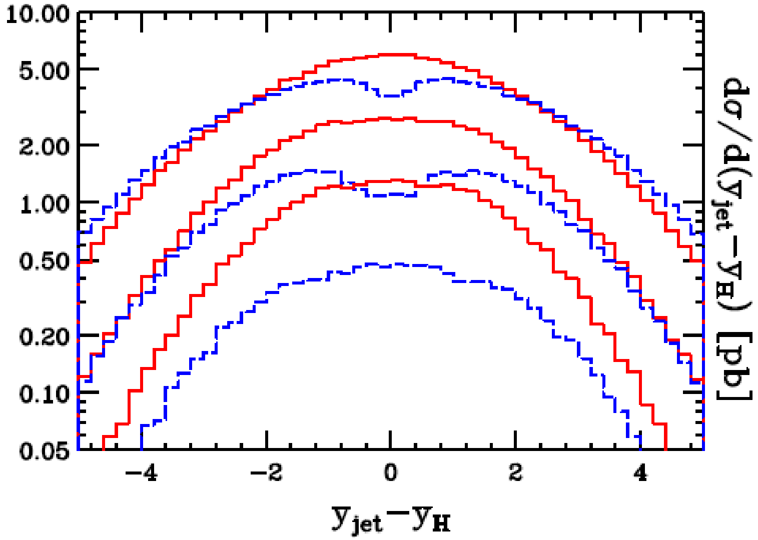
$$d\sigma \approx R(v,r)\,d\Phi_v\,d\Phi_r$$

MC@NLO

$$d\sigma \approx \frac{\bar{B}(v)}{B(v)}R(v,r)\,d\Phi_v\,d\Phi_r$$

POWHEG





JHEP 0904:002,2009 Alioli et. al.

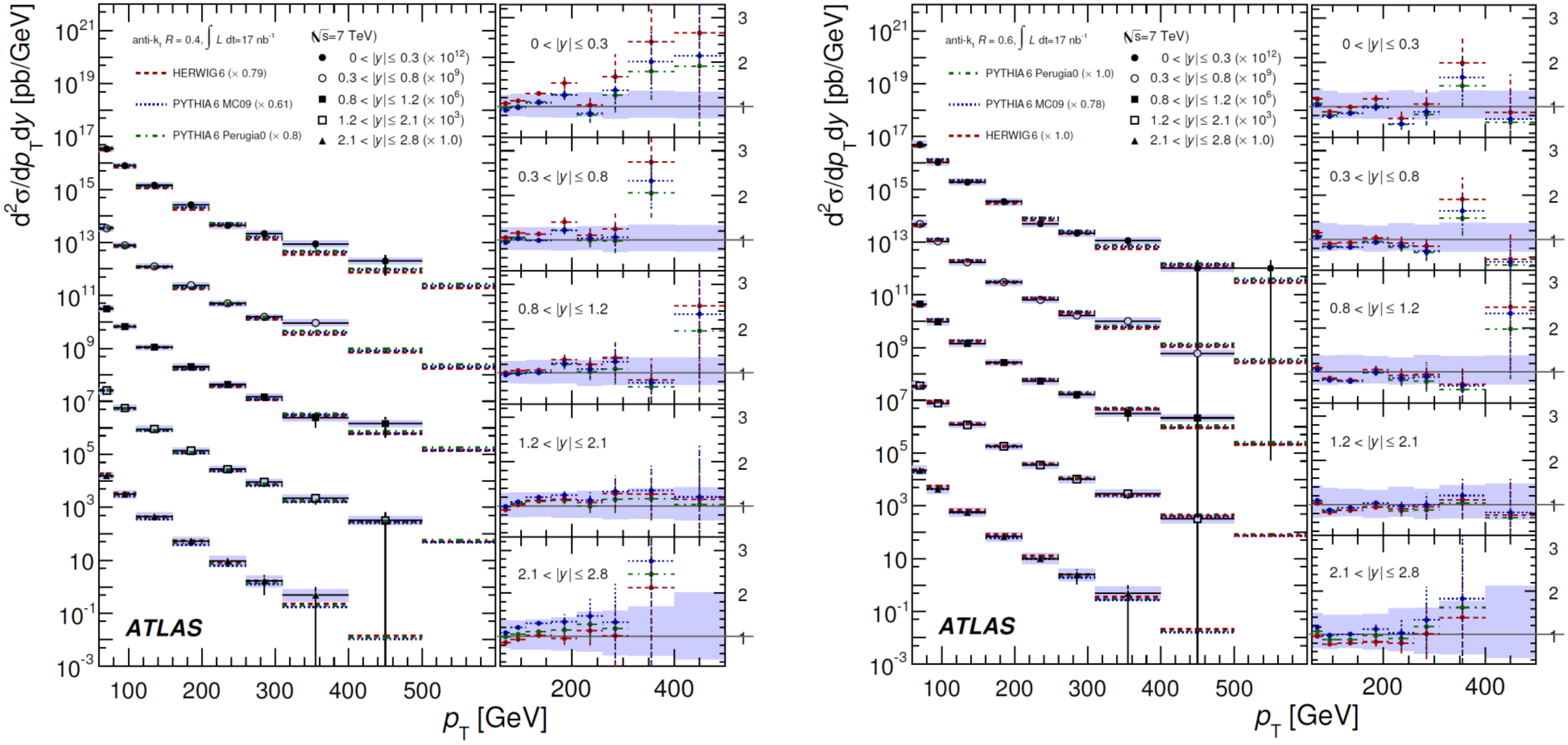
Multi-Jet Leading Order

- While the NLO approach is good for one hard additional jet and the overall normalization it cannot be used to give many jets.
- Therefore to simulate these processes use matching at leading order to get many hard emissions correct.
- The most sophisticated approaches are variants of the CKKW method (Catani, Krauss, Kuhn and Webber JHEP 0111:063,2001)
- Recent new approaches in SHERPA(Hoeche, Krauss, Schumann, Siegert, JHEP 0905:053,2009) and Herwig++(JHEP 0911:038,2009 Hamilton, PR, Tully)

Jets

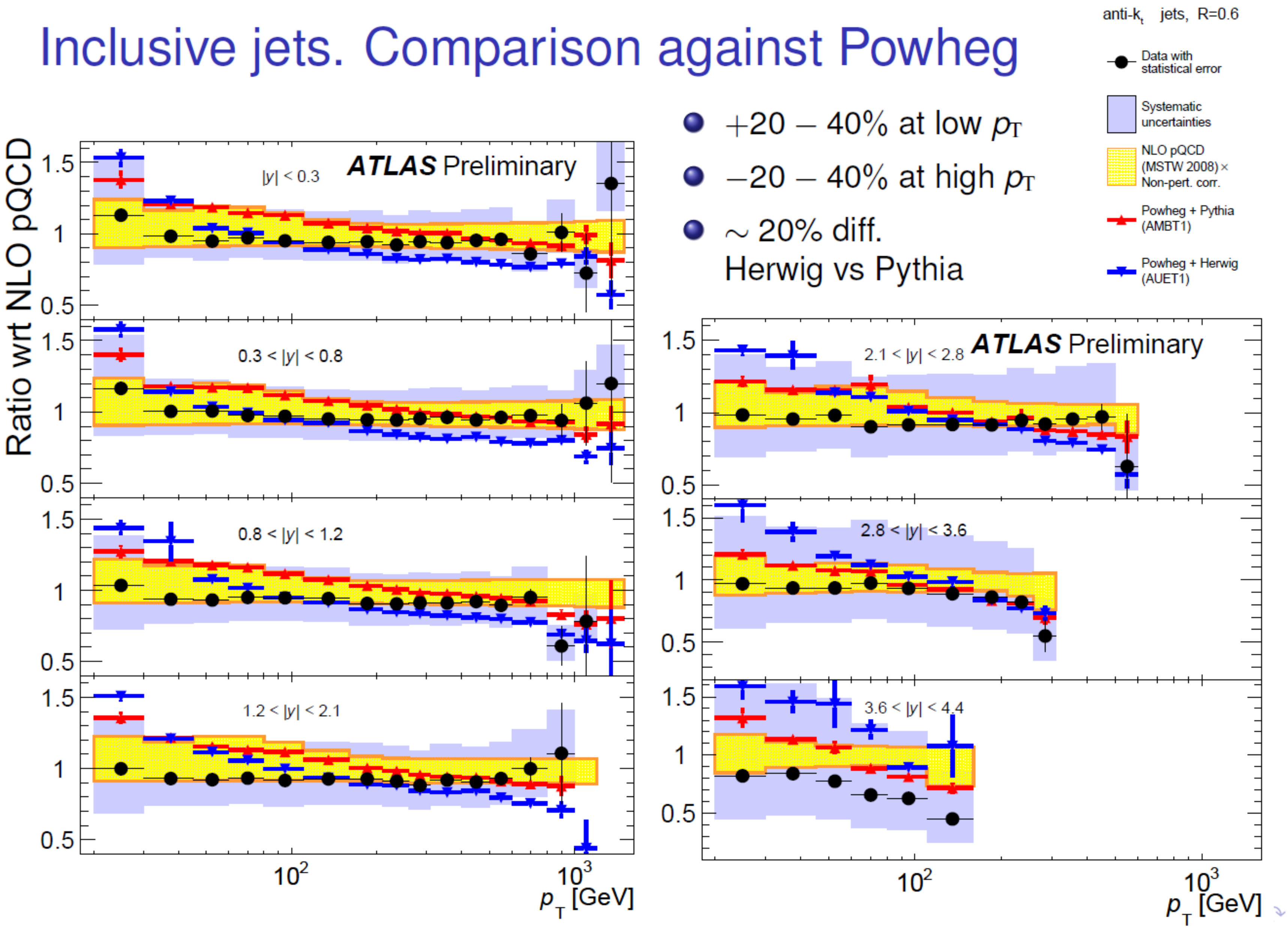
- We would expect the parton shower simulations to describe most properties of up to dijet systems, apart from the total cross section.
- For higher jet multiplicities need either CKKW/MLM or the recent POWHEG simulation of jet production.
- For the NLO rate the only option is the POWHEG simulation.

Inclusive Jet Production



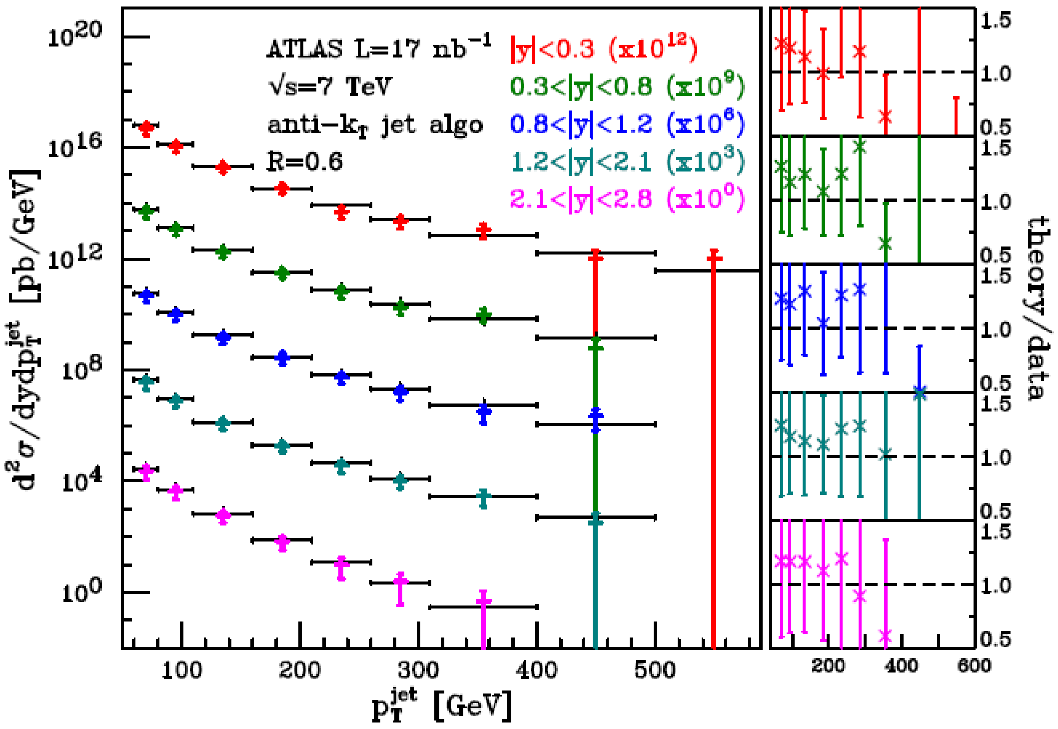
Taken from 1009.5908 ATLAS

Inclusive jets. Comparison against Powheg



- $+20 - 40\%$ at low p_T
- $-20 - 40\%$ at high p_T
- $\sim 20\%$ diff.
  Herwig vs Pythia

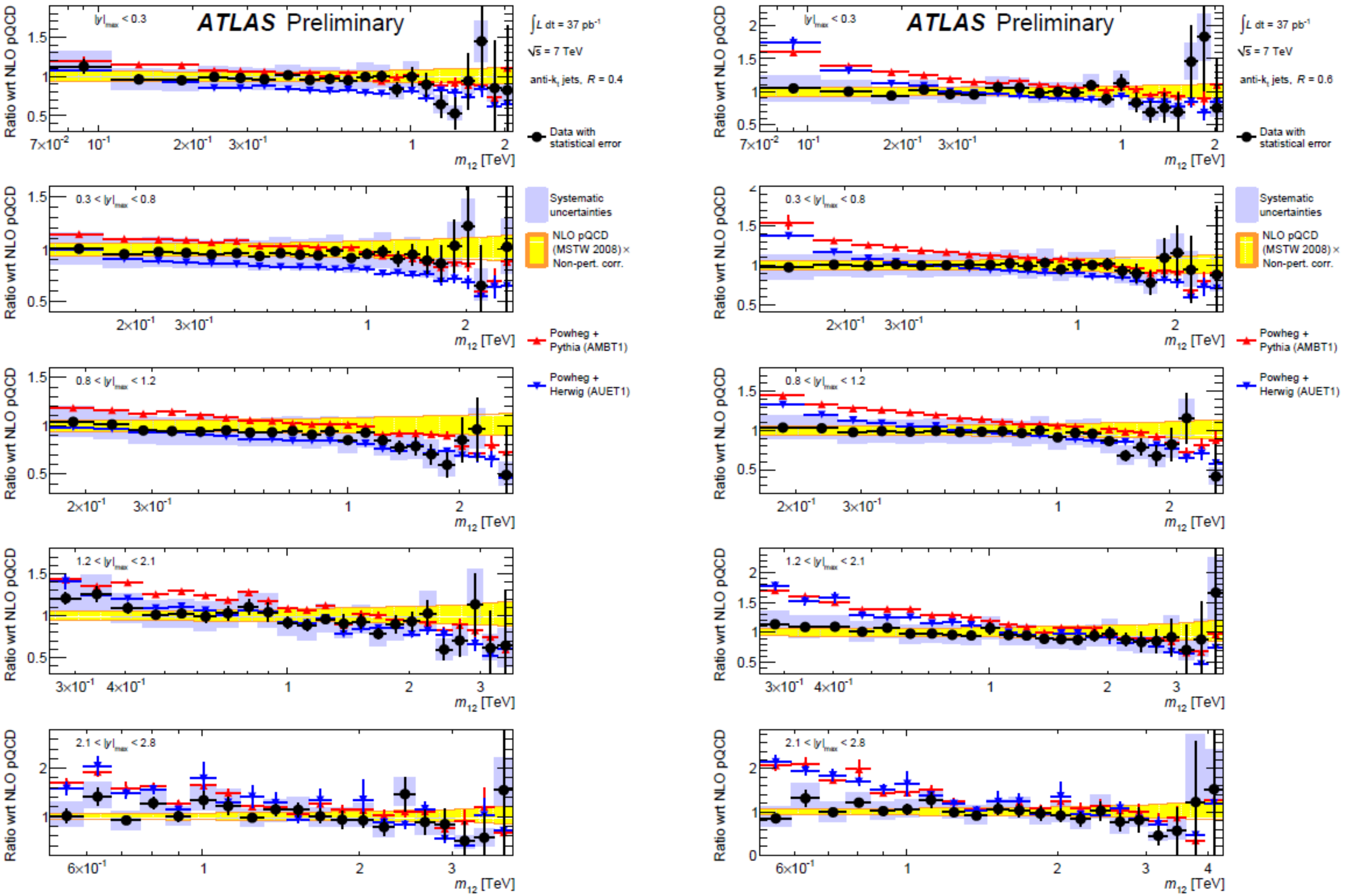
NLO Jet Production

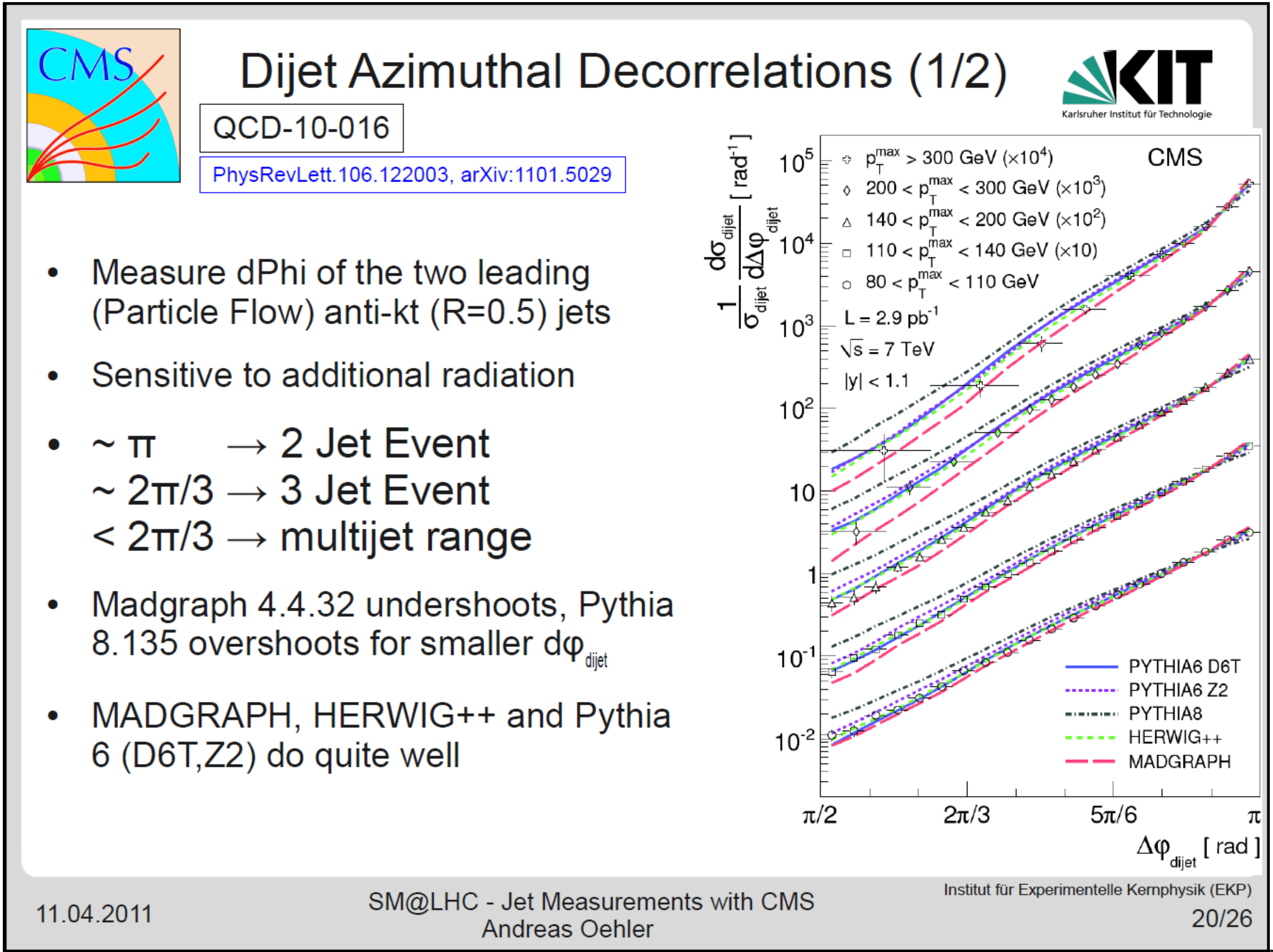


POWHEG compared to ATLAS data arXiv:1012.3380 Alioli et. al.

Dijet Cross Section: Comparison with NLO MC

Anti-k_T $R = 0.4$ Anti-k_T $R = 0.6$



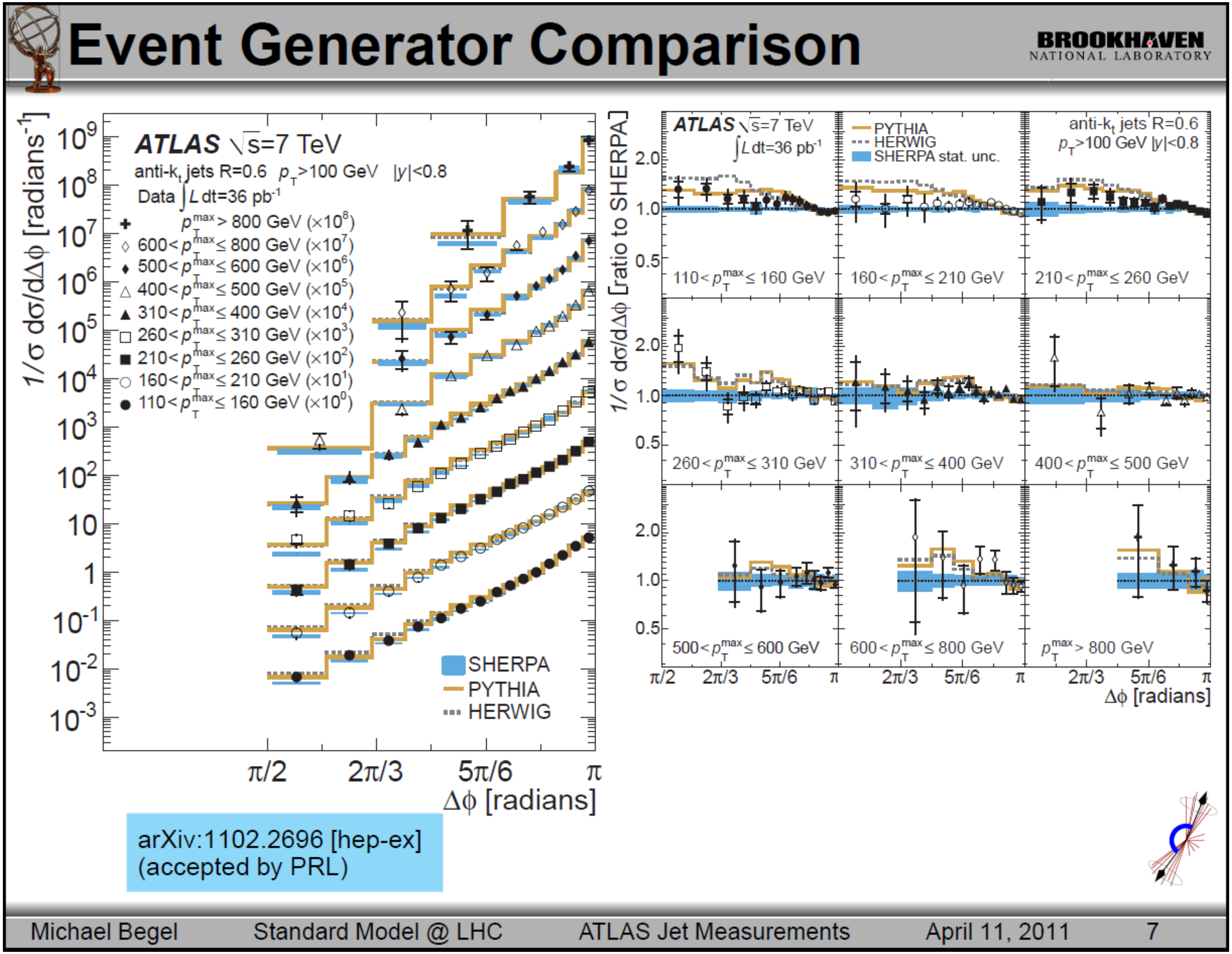


Dijet Azimuthal Decorrelations (1/2)

QCD-10-016

PhysRevLett.106.122003, arXiv:1101.5029

- Measure dPhi of the two leading (Particle Flow) anti-kt (R=0.5) jets
- Sensitive to additional radiation
- ~ π → 2 Jet Event
  ~ 2π/3 → 3 Jet Event
  < 2π/3 → multijet range
- Madgraph 4.4.32 undershoots, Pythia 8.135 overshoots for smaller $d\varphi_{dijet}$
- MADGRAPH, HERWIG++ and Pythia 6 (D6T,Z2) do quite well



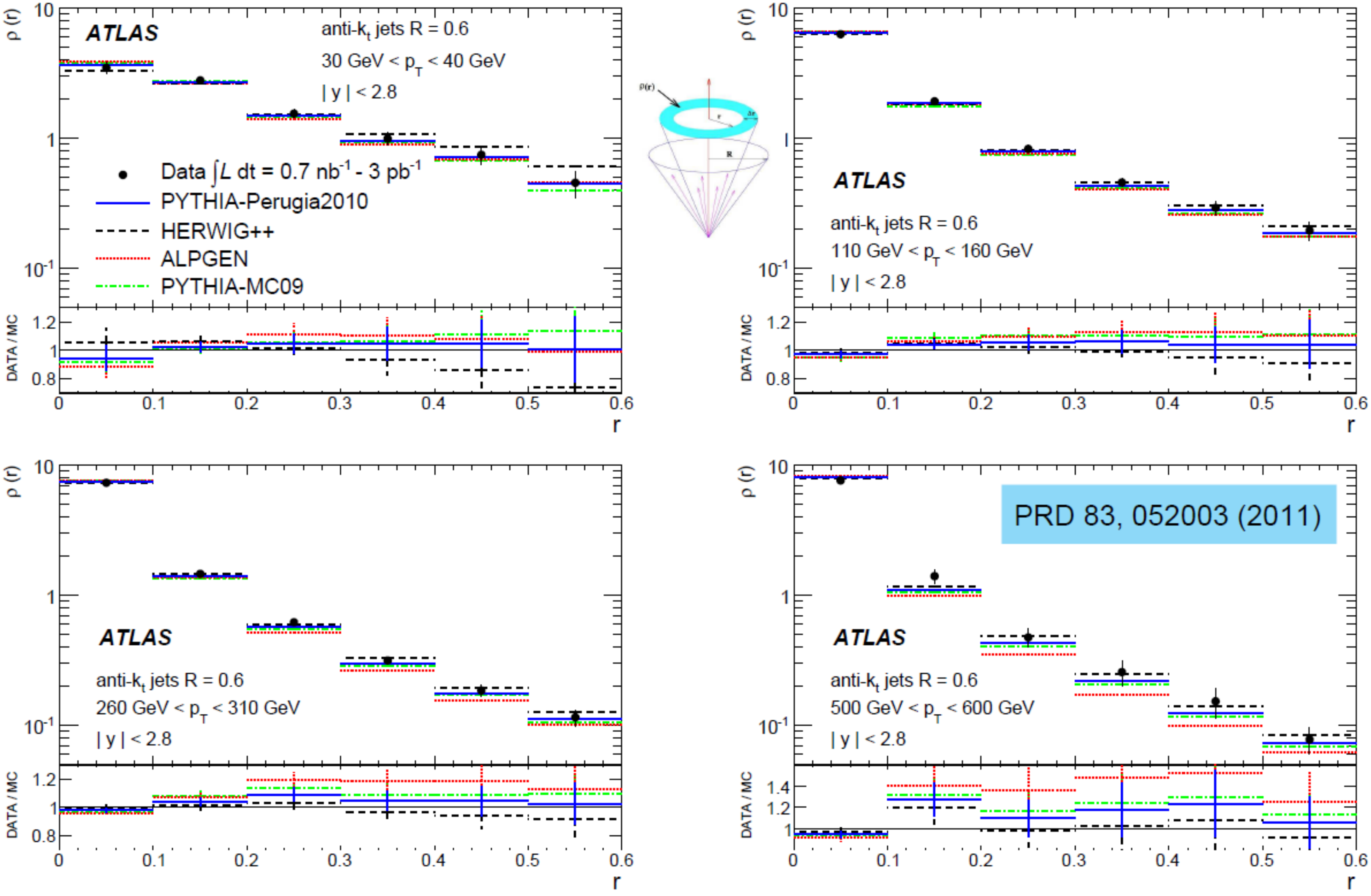
Event Generator Comparison

BROOKHAVEN
NATIONAL LABORATORY

arXiv:1102.2696 [hep-ex]
(accepted by PRL)

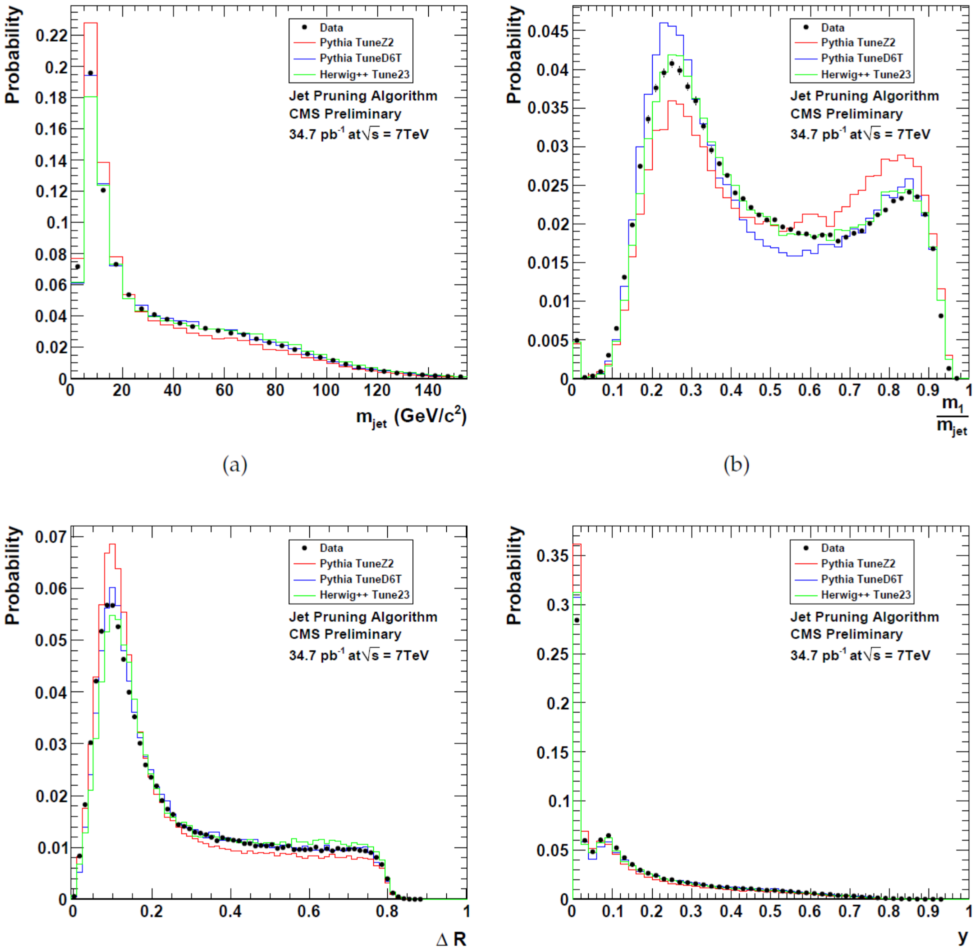
Michael Begel Standard Model @ LHC ATLAS Jet Measurements April 11, 2011 7

Differential Jet Shapes



PRD 83, 052003 (2011)

Jet Substructure



(a) (b)

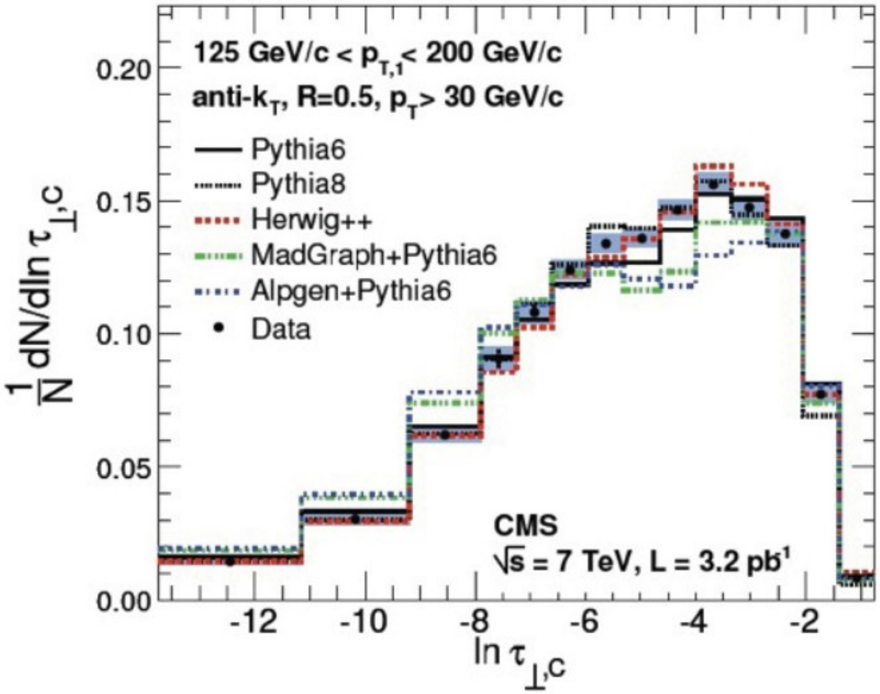
(c) (d)

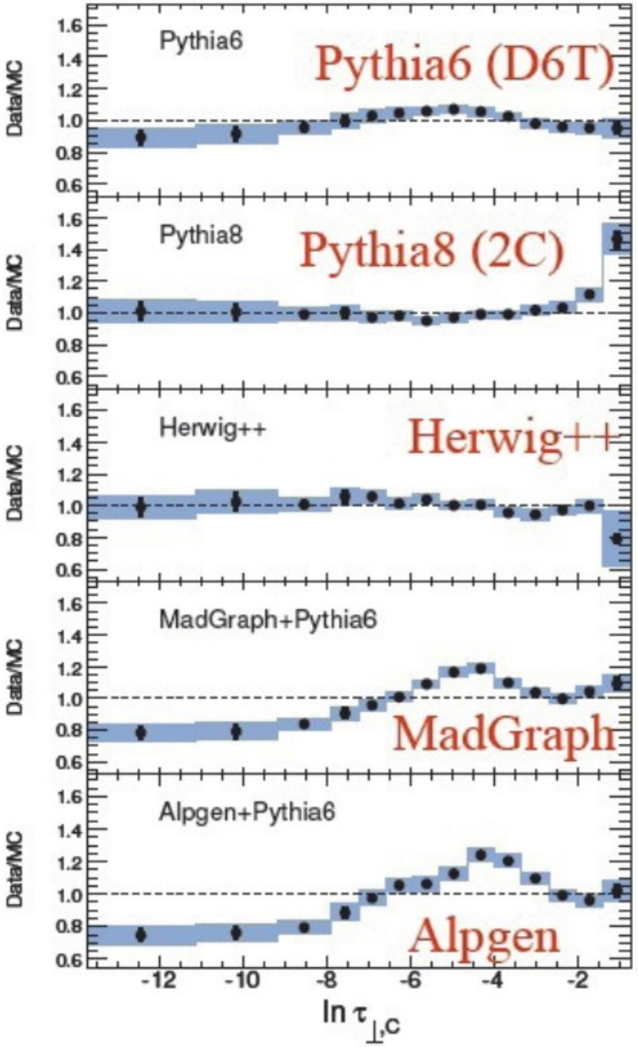
Taken from CMS PAS JME-10-013

Central Transverse Thrust

ETH Institute for Particle Physics

Blue error band = systematic + statistical errors

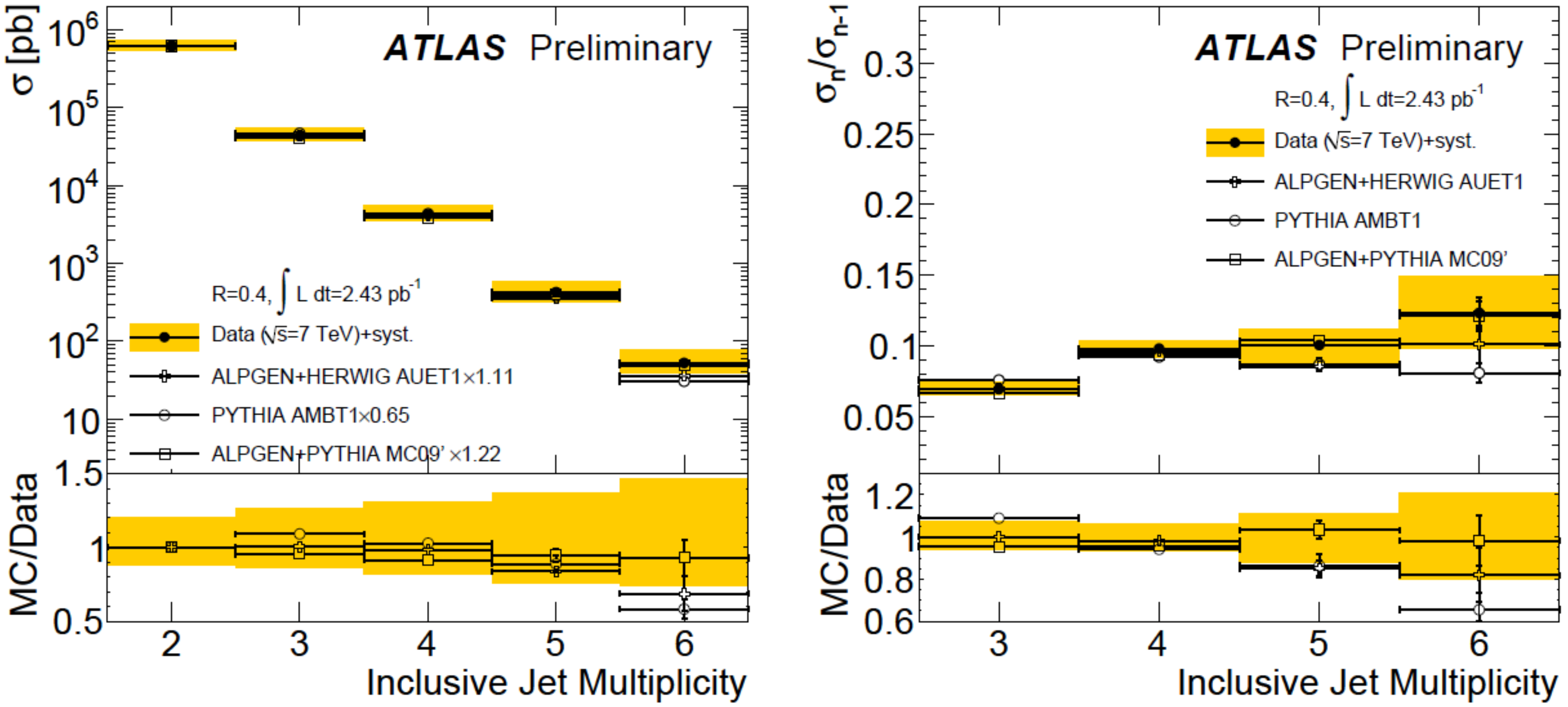




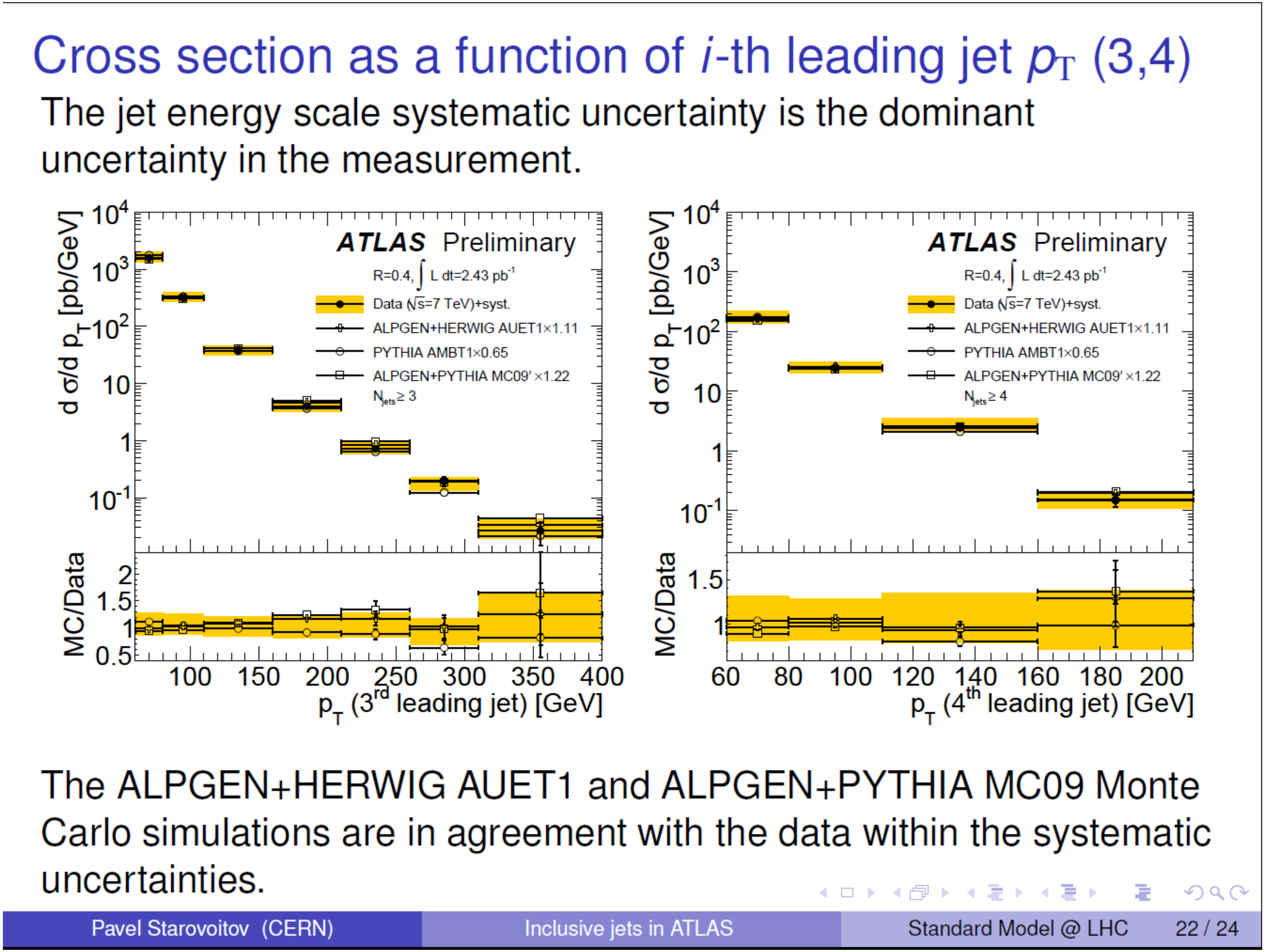
Pythia6, Pythia8 and Herwig++ close to the data, **Alpgen**, **MadGraph** show discrepancies (with CMS parameter choice)

J.Weng

Total inclusive jet cross section as a function of multiplicity



systematic uncertainty in $\sigma(n_{jets})$ is dominated by JES.
JES and unfolding systematics are comparable in $\sigma(n_{jets})/\sigma(n_{jets}-1)$



Cross section as a function of *i*-th leading jet p_T (3,4)

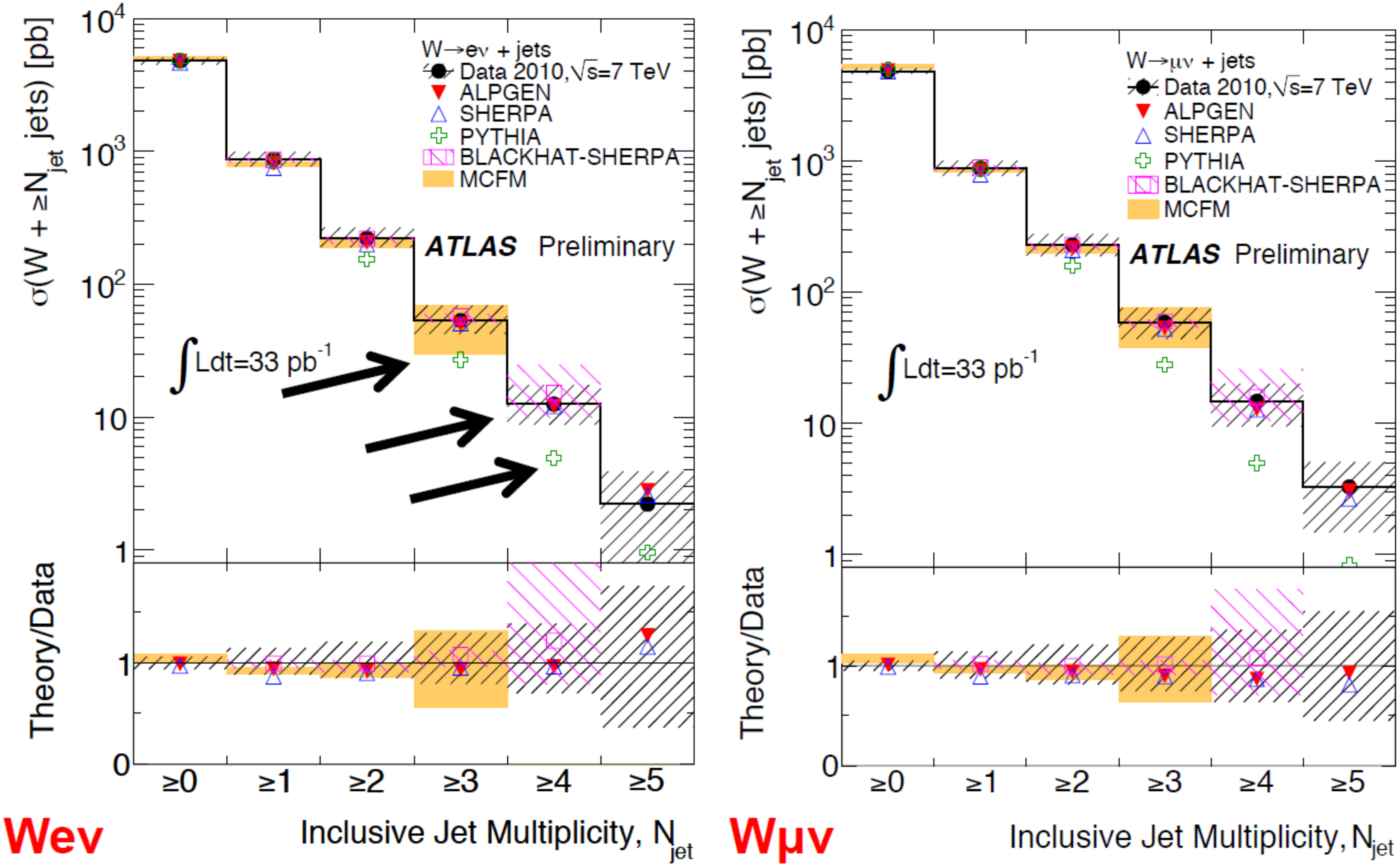
The jet energy scale systematic uncertainty is the dominant uncertainty in the measurement.

The ALPGEN+HERWIG AUET1 and ALPGEN+PYTHIA MC09 Monte Carlo simulations are in agreement with the data within the systematic uncertainties.

V+jet production

- Traditionally the production of W/Z bosons in association with many jets has been an important test of improvements to the parton shower, e.g. CKKW and MLM.
- Easier to calculate than pure jet production and has the advantage of a large scale from the mass of the boson.

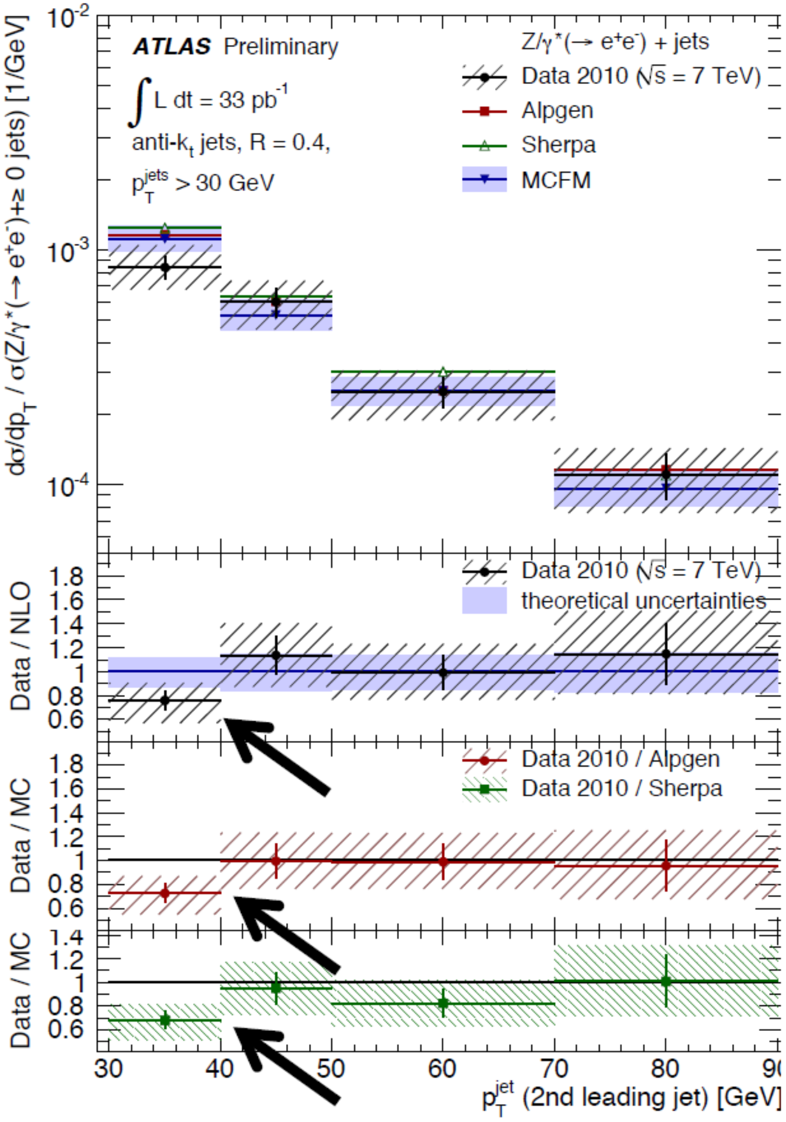
dσ/dN$_{jet}$



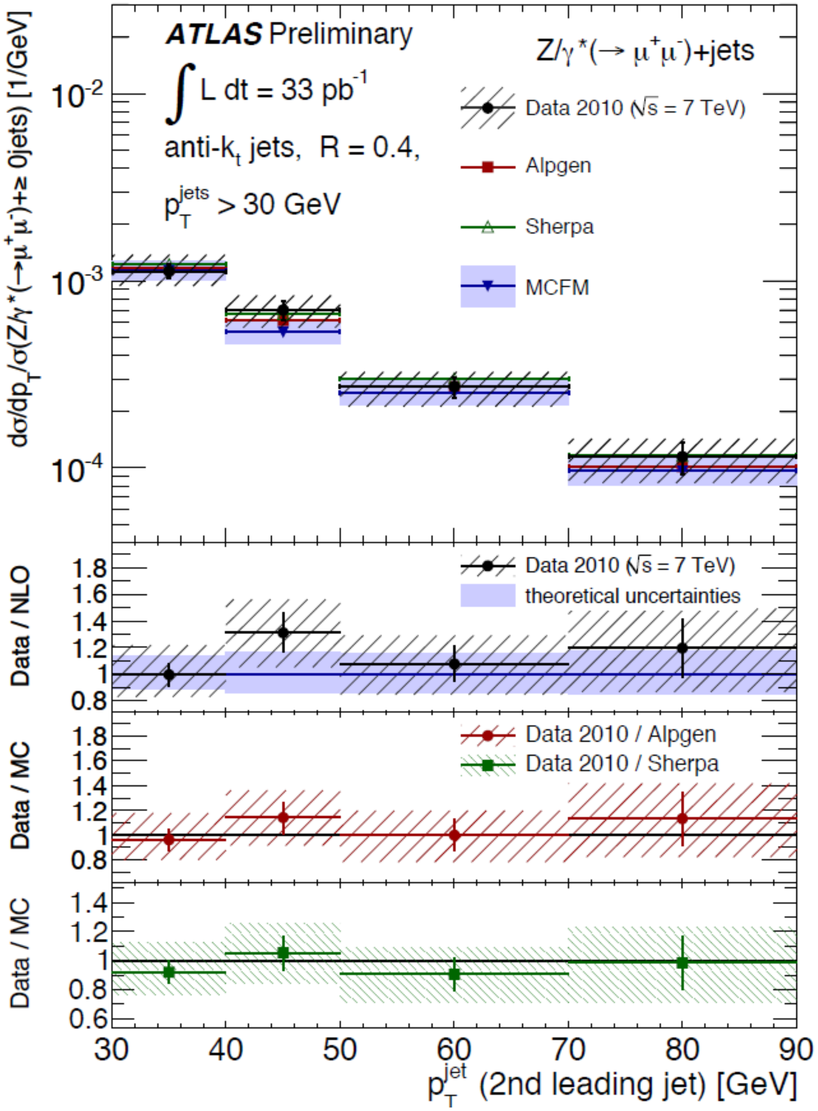
Wev

Wμν

Zee



Zμμ



E. Dobson SM@LHC

Summary

- We’ve spent a long time developing a new generation of simulations for the LHC.
- So far things look O.K. but that may well change as statistics improve and systematic errors reduce.
- A tune of PYTHIA can describe pretty much anything, not clear that there’s a tune of PYTHIA that can describe everything.
- Limited use of the new generation of tools, hopefully this will improve as higher statistics requires more accurate predictions.