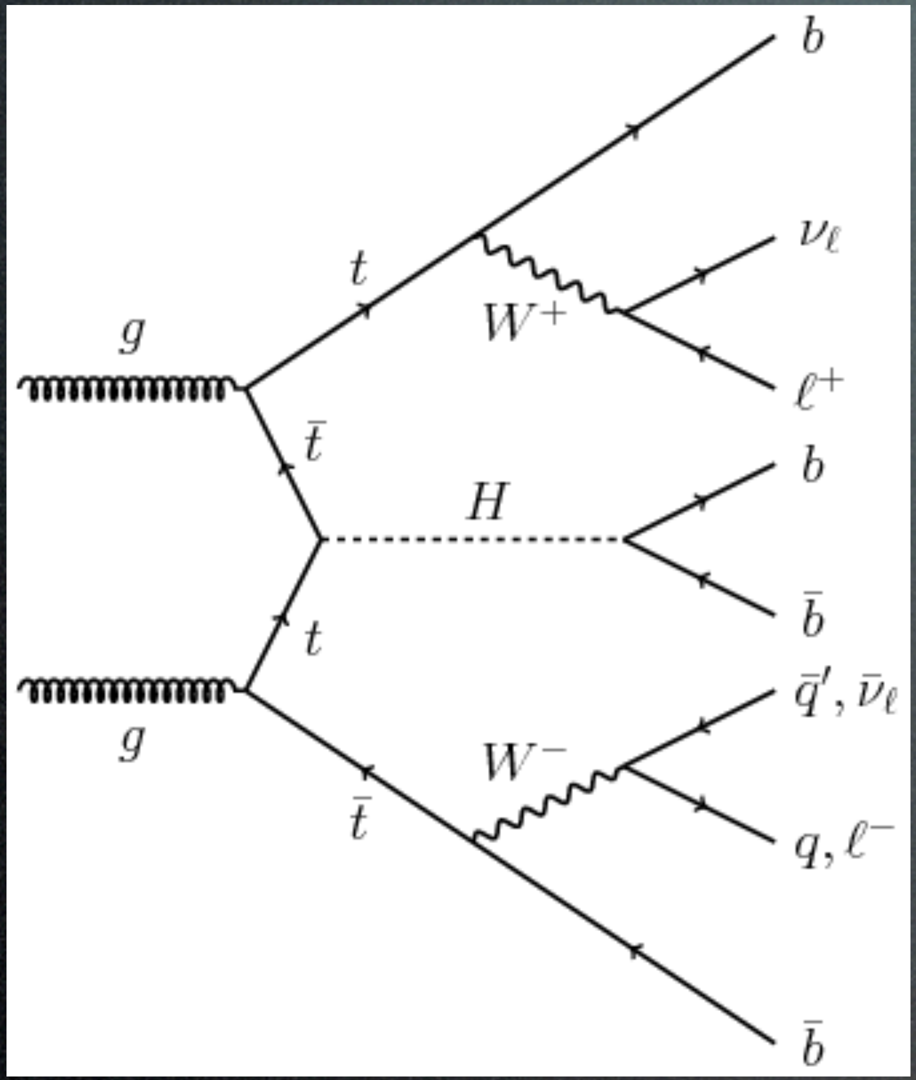
Underlying Event and Monte Carlo Generators

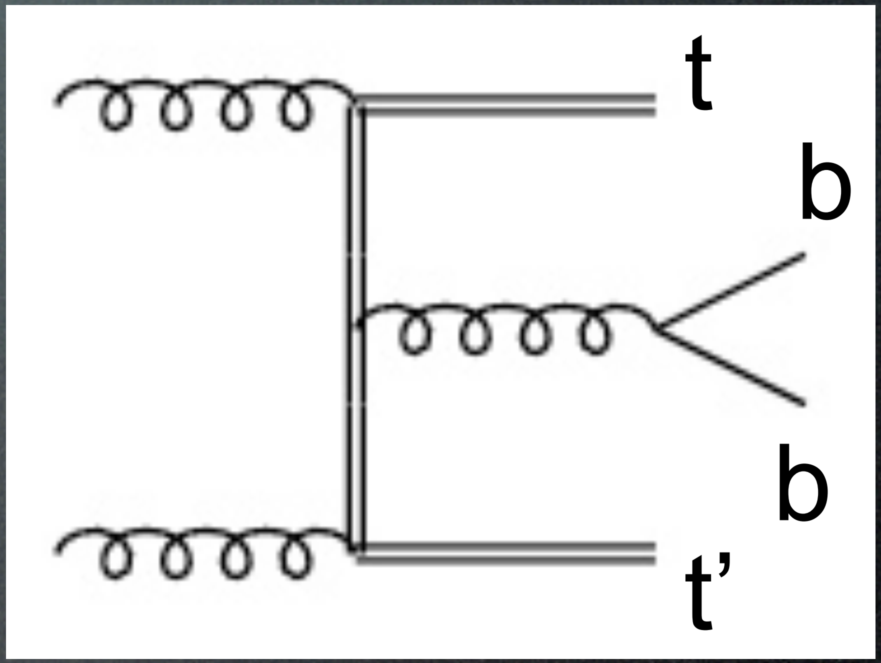
Deepak Kar
University of Glasgow

Conference on What Next at LHC
6-8 January, 2014
TIFR, India

Prologue

One of the hardest measurements now:



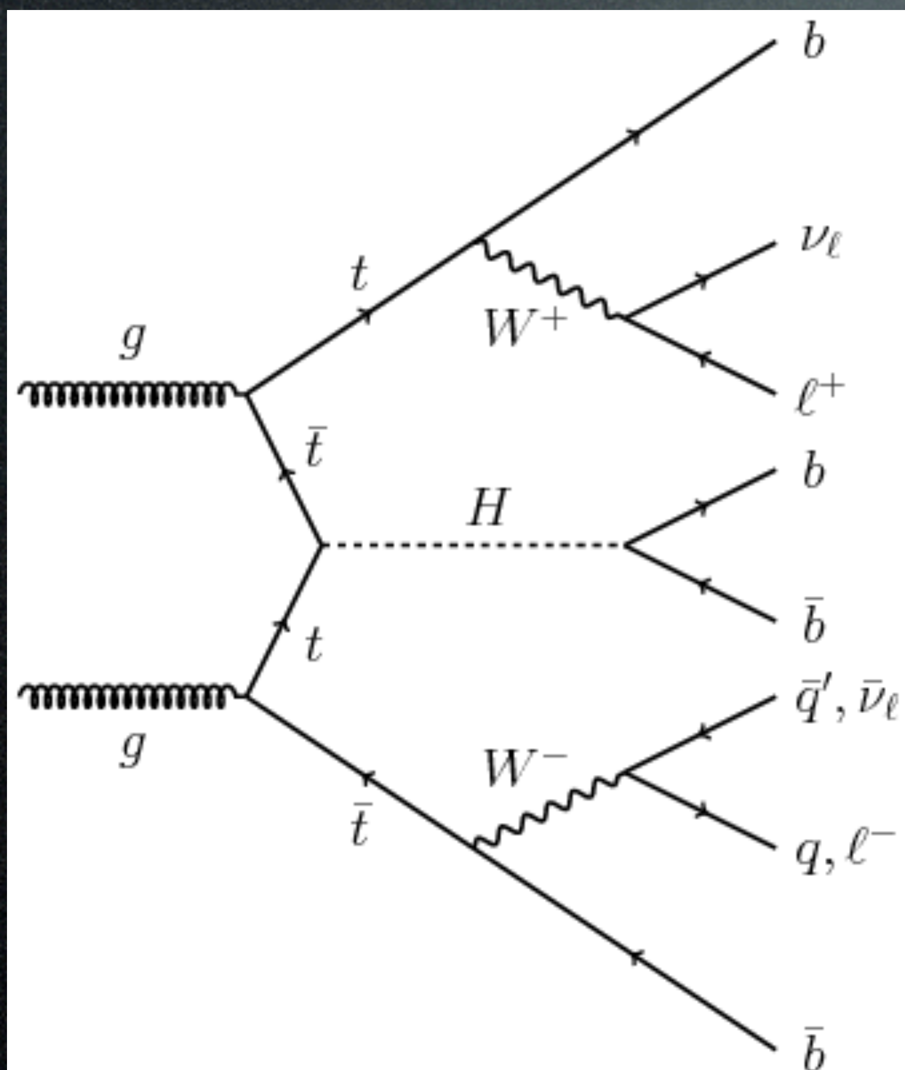


Signal: ttH(bb)

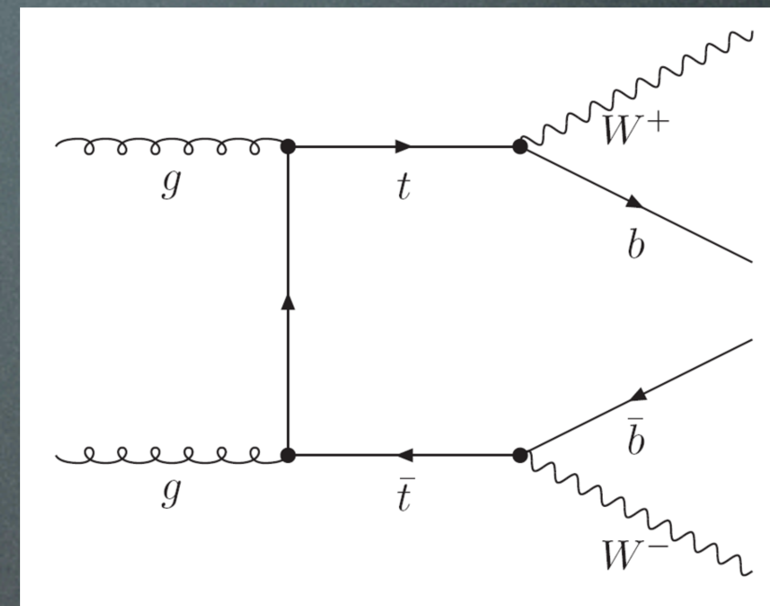
BG: ttbb

important for measuring Yukawa couplings

One of the hardest measurements now:

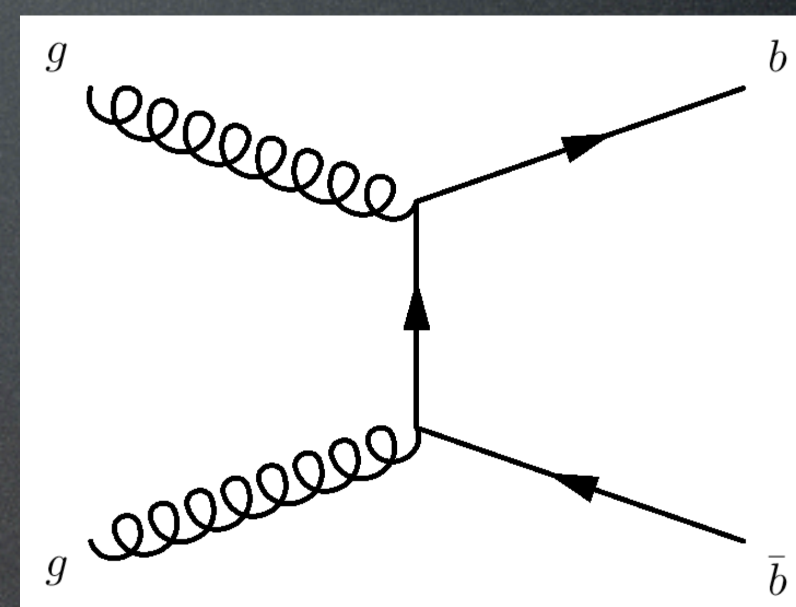


Signal: ttH(bb)

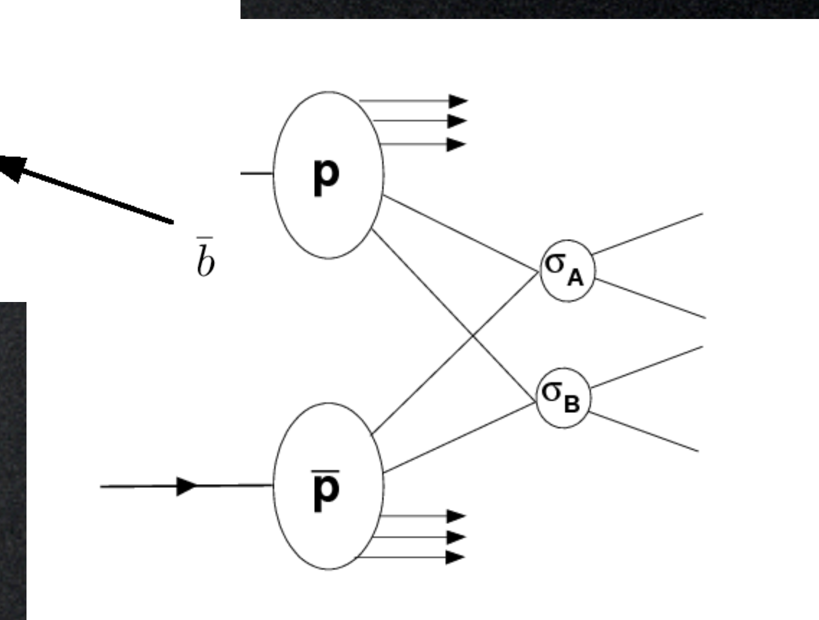


tt

+



(DPI) bb



Glossary

- Minimum-bias (MB): Pretty much everything, exact definition trigger dependent.
- Underlying event (UE): background to events with an identified hard scatter (more like the actual interesting events we want to look at)
- Pileup (PU): (uncorrelated) separate collisions within the same/different bunch crossing we can't differentiate because of our finite detector resolution (more like "isotropic" min-bias events).

Underlying Event

Why do we care?

- The process of interest at hadron colliders are mostly the hard scattering events.
- These hard scattering events are contaminated by the underlying event.
- The underlying event is an unavoidable background to most collider observables.
- The underlying event is not well predicted since non-perturbative physics is involved.
- And from an experimental point of view, on an event by event basis, it is impossible to separate the UE component.

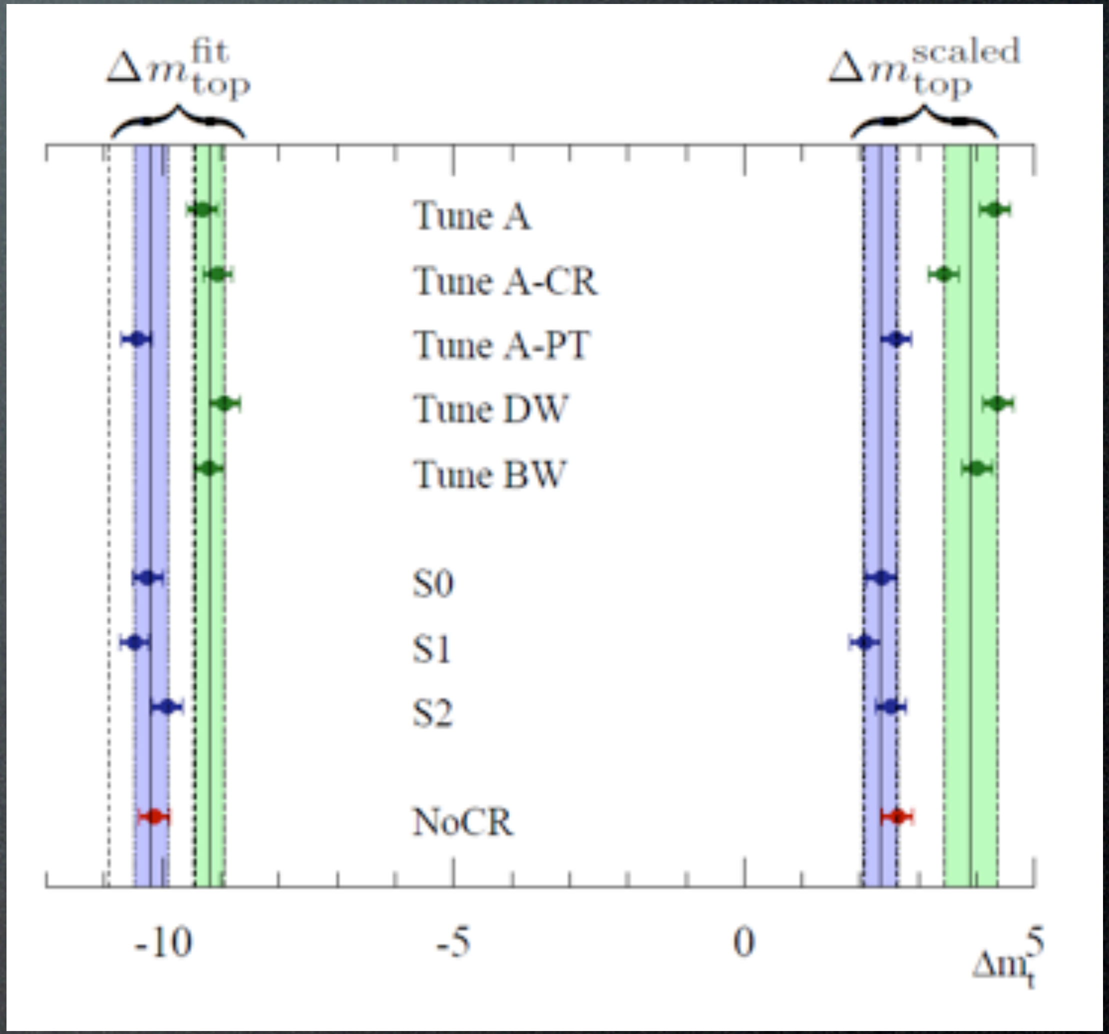
NEW PREDICTIONS (10 years)

1. QCD tests & applications will greatly improve, incorporating NLO, NNLO,...and a theory of fragmentation and hadronization.
2. Atlas and CMS will discover a candidate Higgs particle.
3. There will be convincing evidence for Susy particles.
4. Plans will be underway to build a LC (at Cern) to explore the superworld and the US will join CERN.
5. There will be direct detection of the Dark Matter wind.
6. Alice will see a crossover to the perturbative quark-gluon plasma.
7. Some new Z mesons will be discovered.
8. Gravitational waves and B modes will be observed.
9. String theory will start to be a **theory** with predictions.
10. We will have a plausible explanation of why Λ is so small.

David Gross at EPS 2012

An Example

- Top mass is a very important "free" parameter.
- Measured experimentally, uncertainty dependent on the type of UE model used (more on the models later).



Eur.Phys.J.C52:133-140, 2007

Our Strategy

- We have to use the underlying event and other softQCD distributions to test the phenomenological models and "tune" the Monte-Carlo event generators to give the best description of the data.
- We gain deeper insight if data does not match up with Monte-Carlo predictions, which reflect our current understanding of these processes.

Monte Carlo Models

- Leading order/Parton shower models: Trying to build up a complex 2->N final state by showers.
- Pieces of a Parton-Shower MC Generator: (2->2 hard scattering), ISR, FSR, MPI, Fragmentation, Hadronization.
- Examples: Pythia, Herwig family.
- Higher order/Multileg generators: Sherpa, Alpgen, MC@NLO, Madgraph, Powheg ...
- Generators used mostly for a specific process: Phojet (diffraction), HIJING (heavy ion), AcerMC (top), JHU (spin and polarization information)...

A Note on the Models

"The predictions of the model are reasonable enough physically that we expect it may be close enough to reality to be useful in designing future experiments and to serve as a reasonable approximation to compare to data. We do not think of the model as a sound physical theory . . . "

– Richard Feynman and Rick Field, 1978

Tuning

- Ultimate goal: models need to describe real data.
- "Free" parameters control all these aspects of the models, which cannot be derived analytically.
- A bunch of correlated (or anti-correlated) parameters describe one aspect, so have to change them simultaneously.

Tune: A particular optimized parameter setting in a particular MC generator to match the simulation with available data. Differ according to which datasets are included.

A Brief History of Tuning

- Historically most effort has been devoted to tuning (Fortan) Pythia6, even at LEP/CDF.
- ATLAS did tune (Fortran) Herwig+Jimmy(which adds MPI), and now (C++) Pythia8.
- (C++) Herwig++, Sherpa has so far been tuned by authors.

Apollo's priestess, Pythia, performing the duty of the oracle

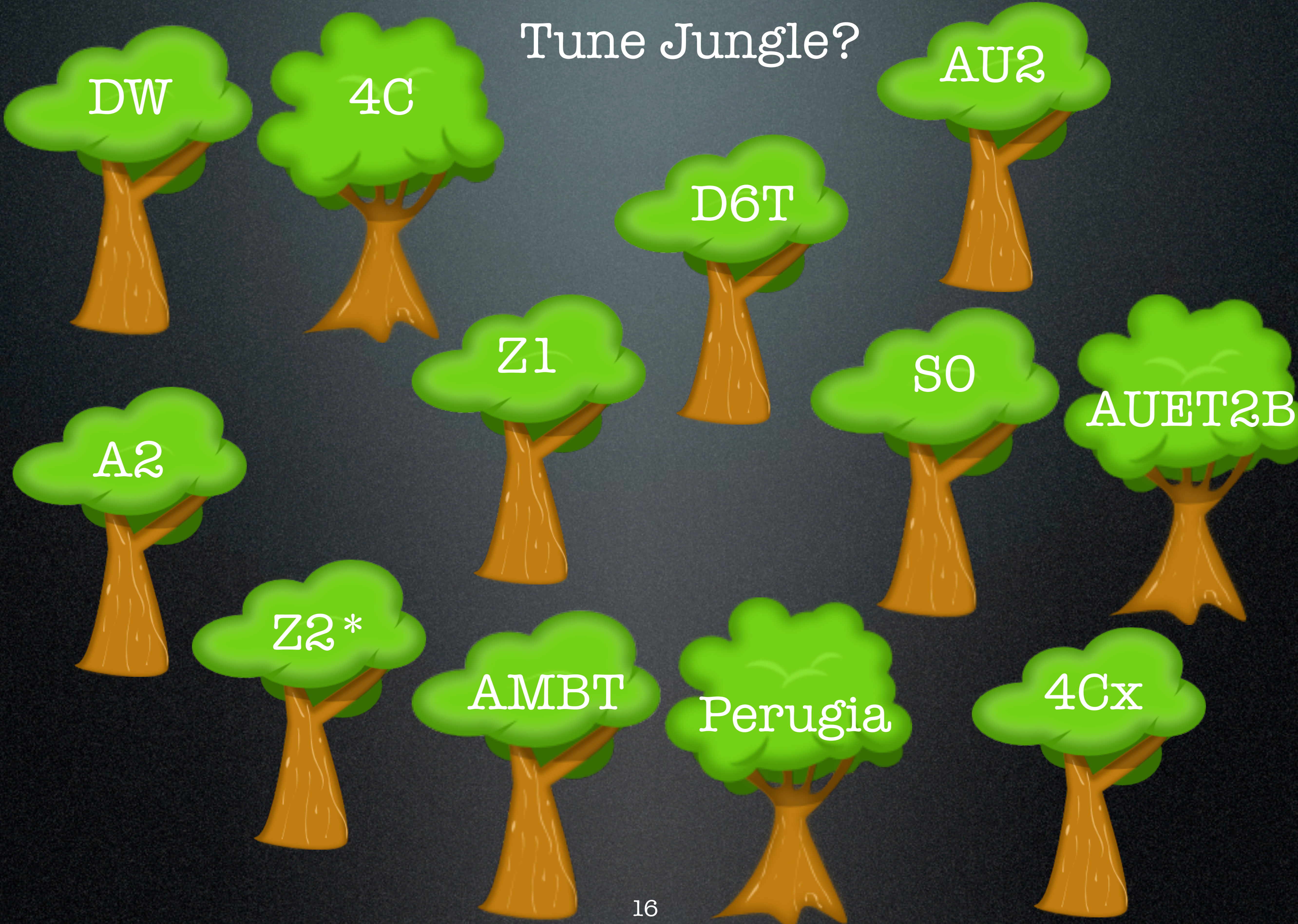
- Hadronization and FSR: LEP
- ISR and MPI: Hadron colliders

Tuning Procedure

- Tuning-by-eye: the classical approach. Stare at a few distributions, think hard, change some parameters, hope those are better, nothing else is broken. Very intuition/experience dependent.

- Automated tuning tool/Professor: pioneered by ATLAS. Essentially generate lot of samples covering the parameter space. Interpolate the generator response, get the best fit by minimization. (and burn a lot of CPU)

Tune Jungle?

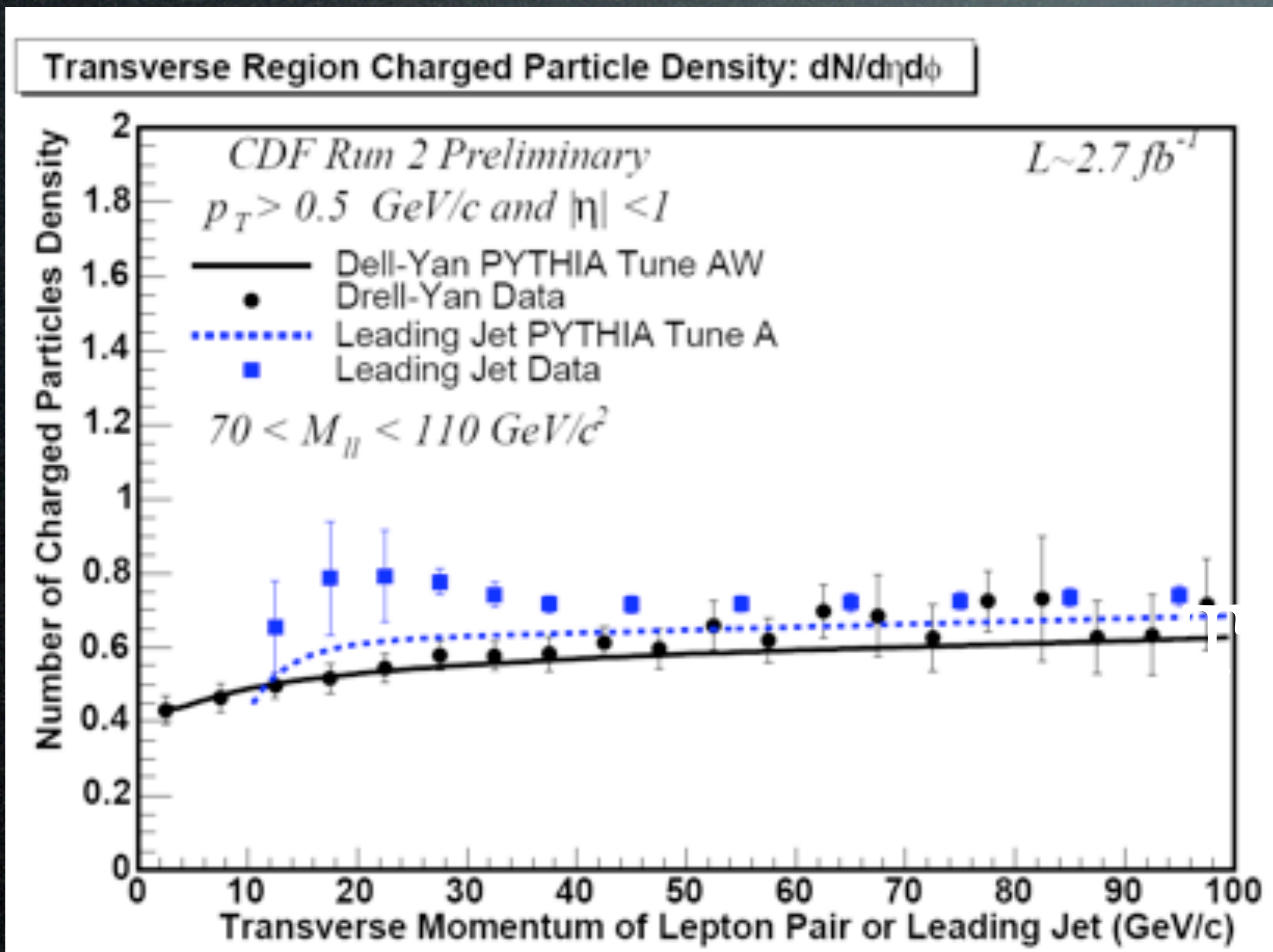


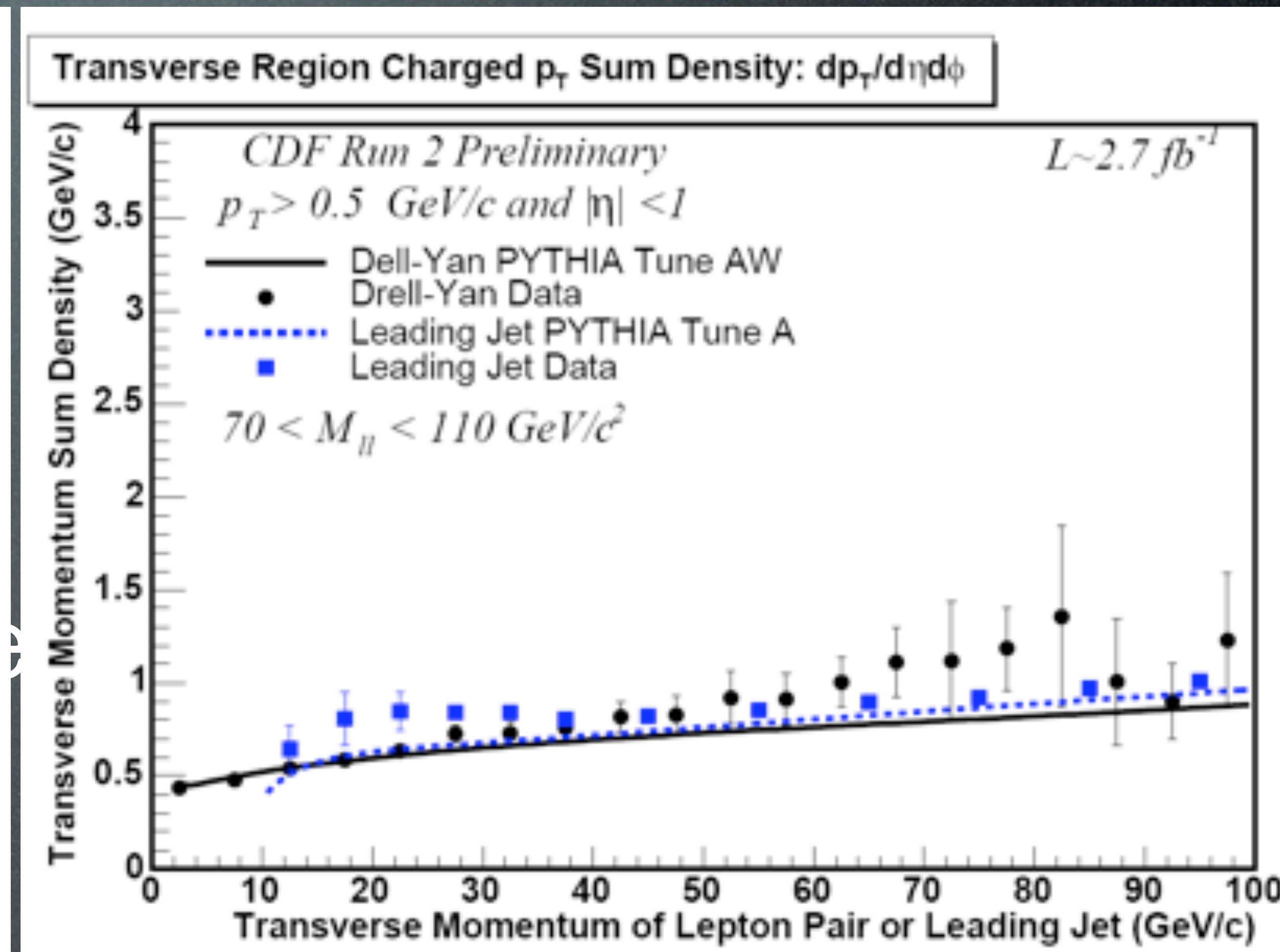
Tevatron Era Tunes

- CDF/Rick Field tunes: Pythia6 tune A, AW, DW, DWT, D6, D6T.
- ATLAS: DC2, CSC/MC08, MC09, Mc09c.
- Perugia/Peter Skands tunes: S0, Perugia0, Perugia10 (soft, hard, no colour reconnection variants).

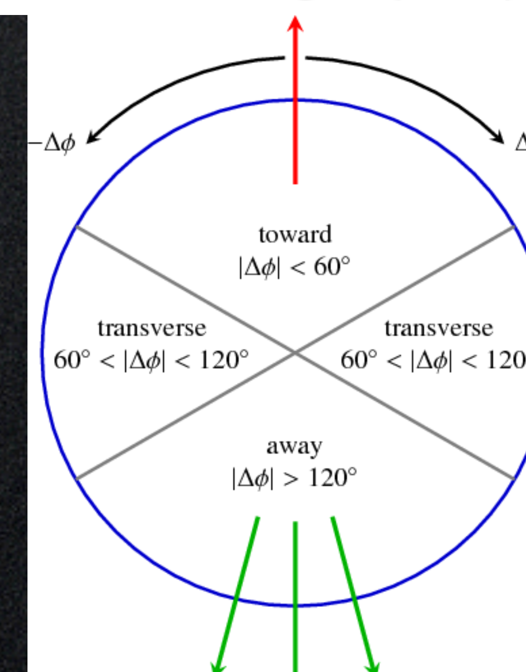
Leading Jet and Z UE Results

Phys.Rev. D82 (2010) 034001



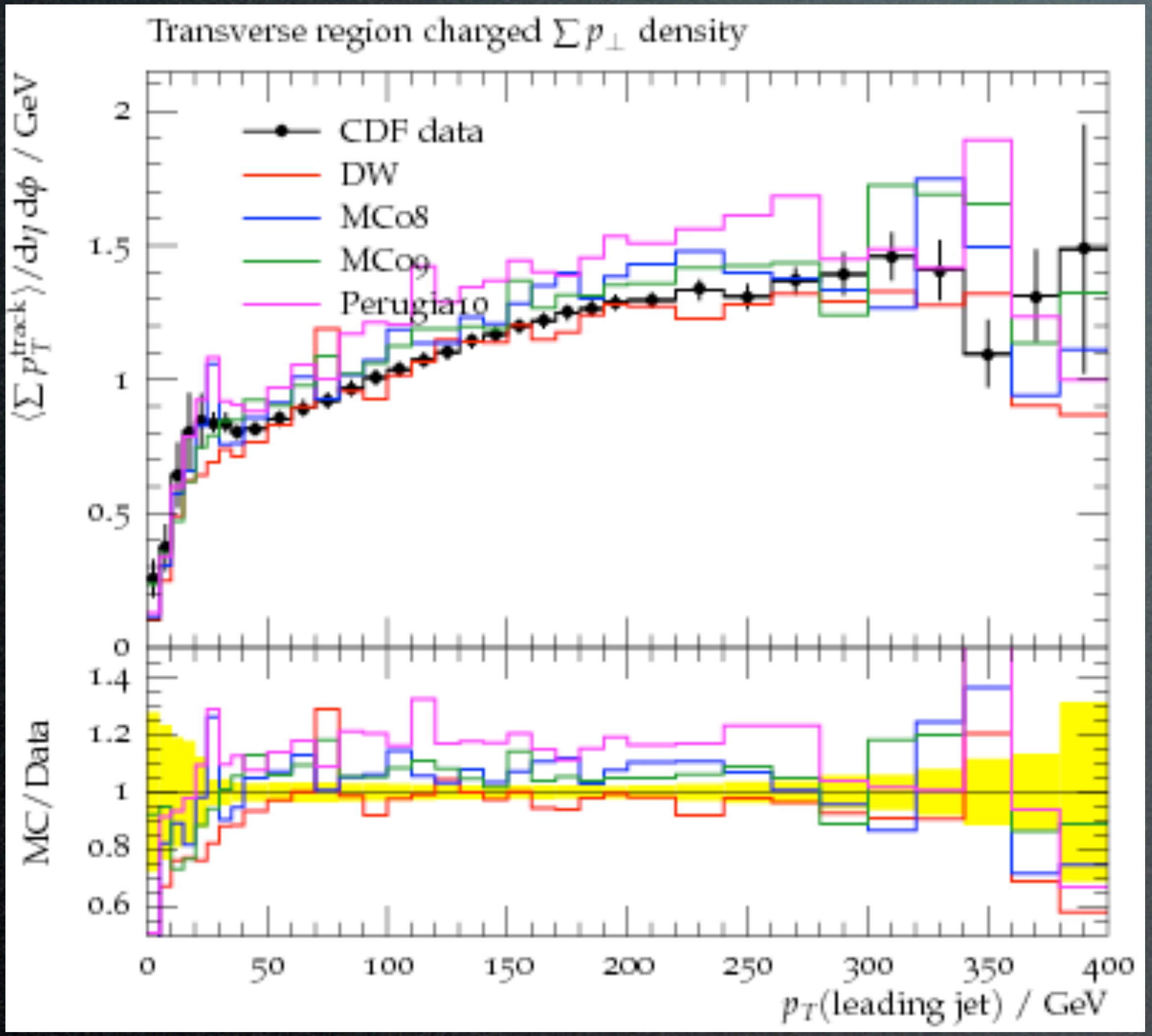


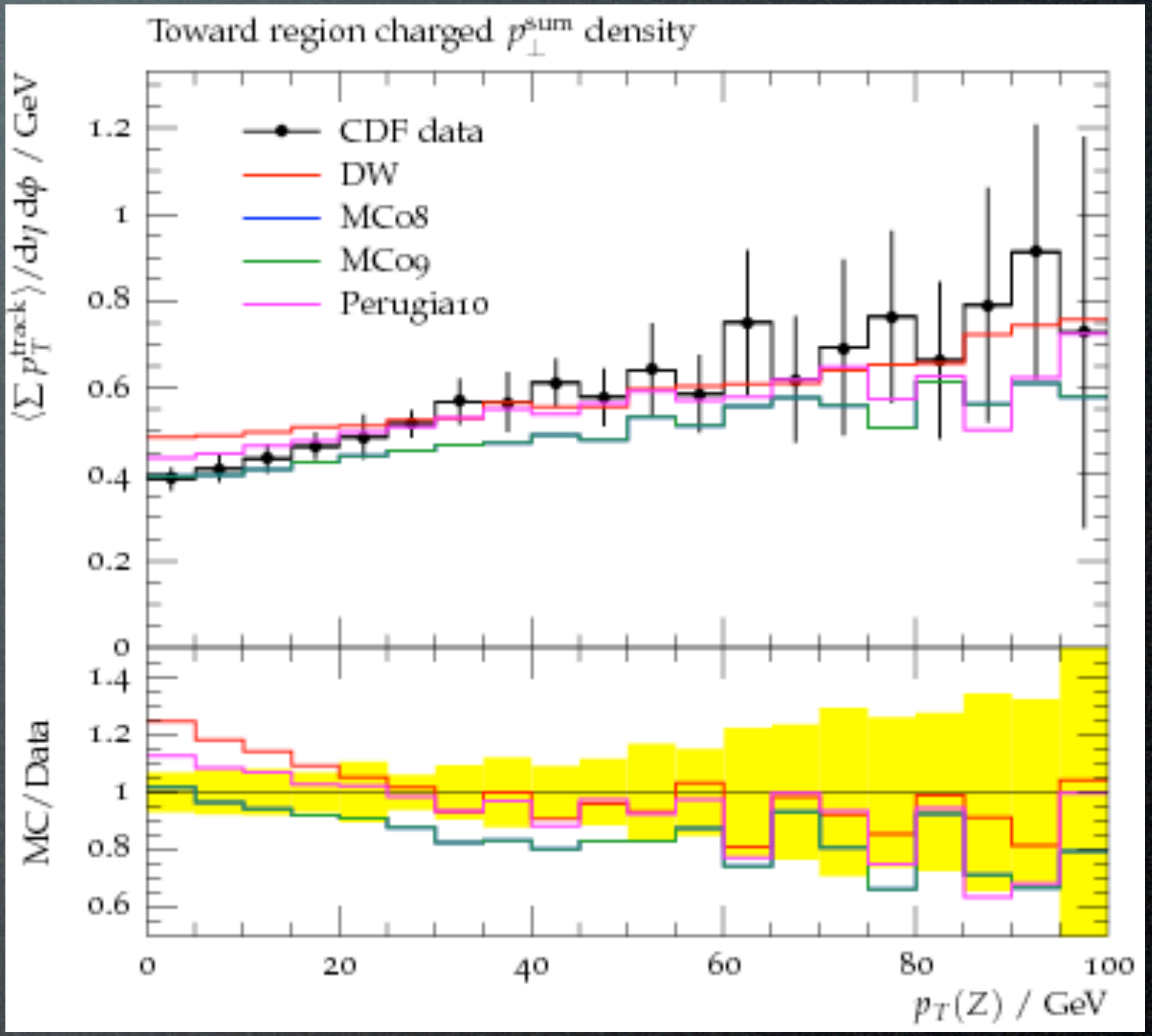
UE activity in Z-boson and jet events fairly similar in Tevatron.



Is it still the case at the LHC?

Pre-LHC tunes





Phys.Rev. D82 (2010) 034001

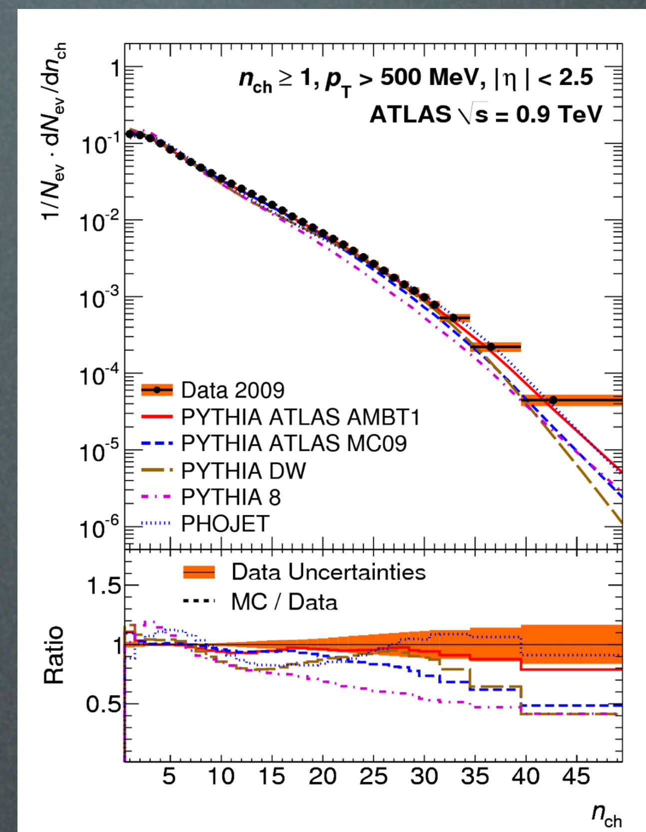
The tunes do quite well ...

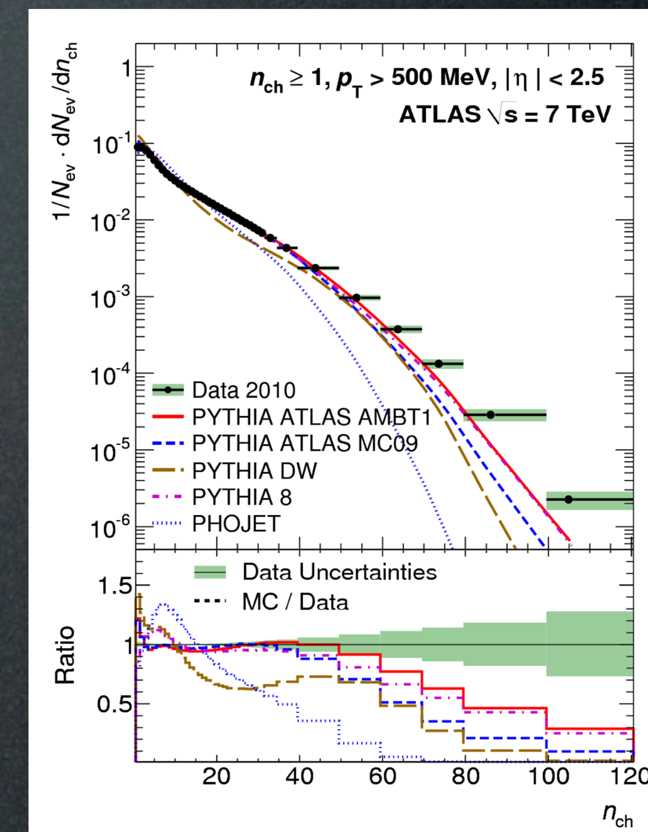
Did they work at the LHC?

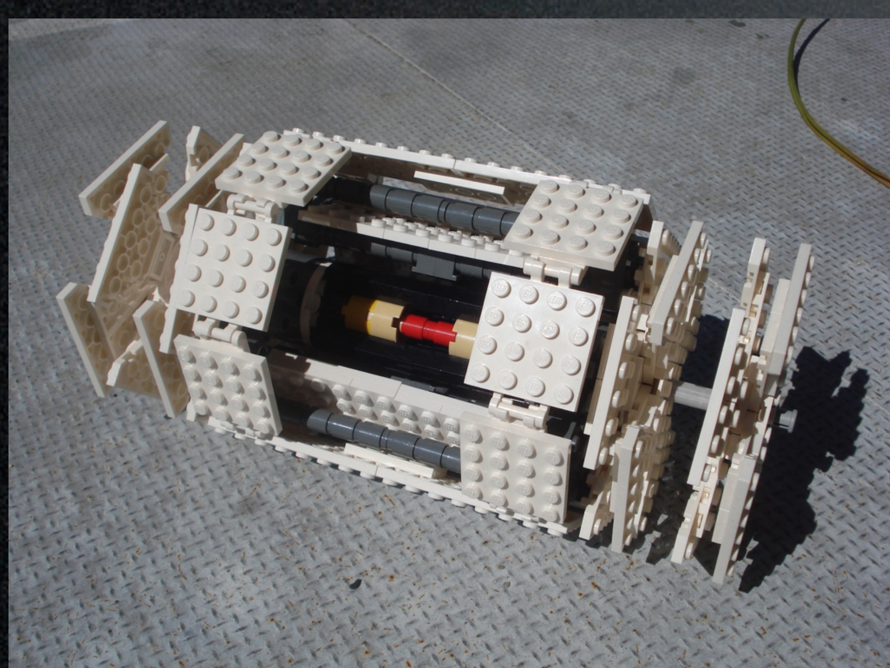
Then Came the LHC

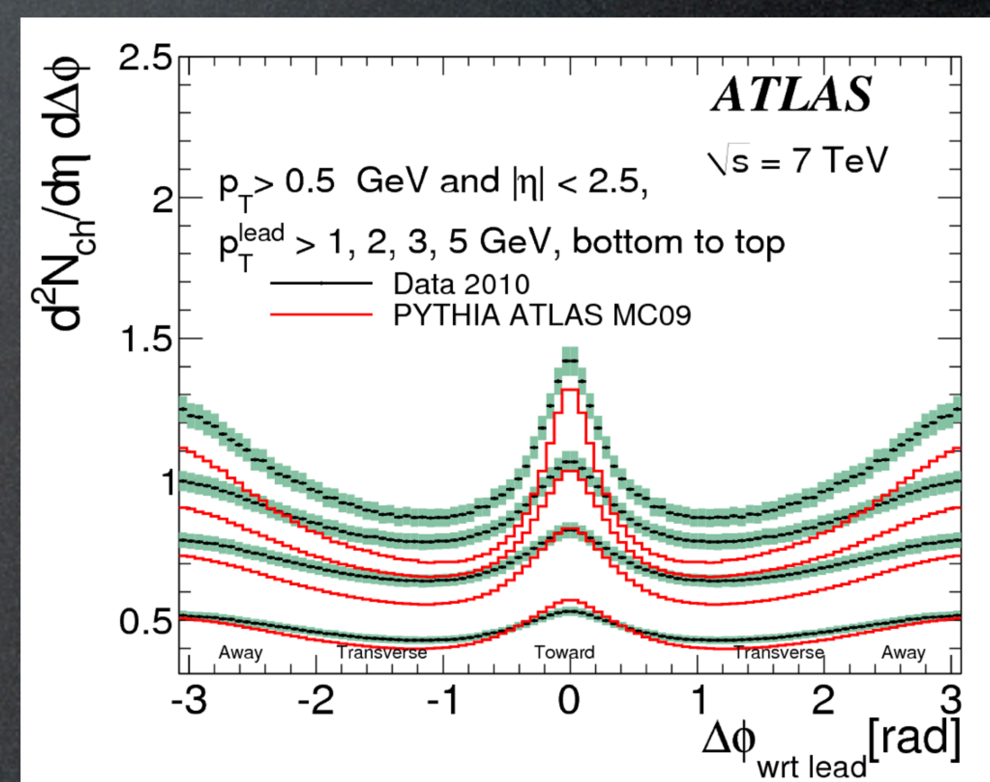
- Tevatron tunes did not agree with the early minbias and underlying event data.

- Not just at 7 TeV, but also at 900 GeV!

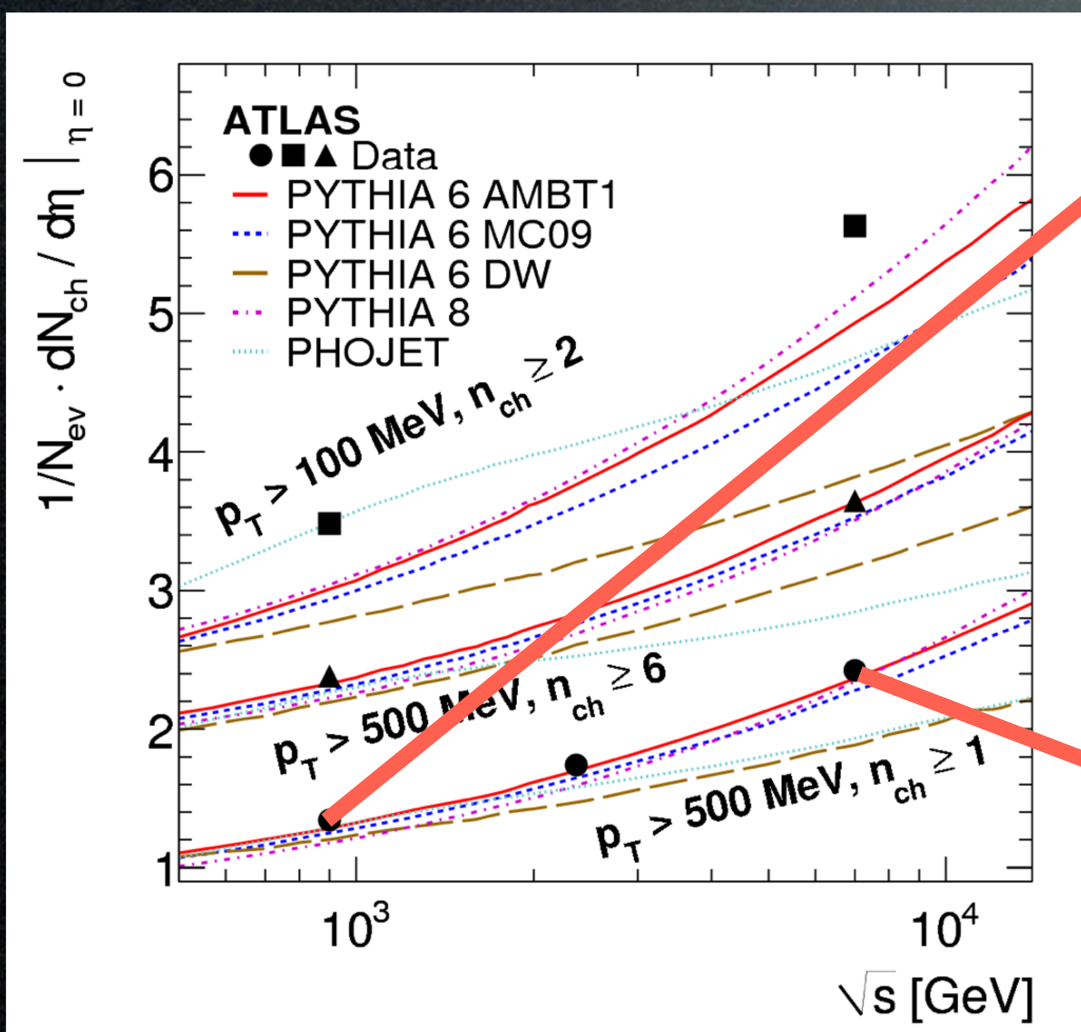








A slight detour: comparsion between UE and MB



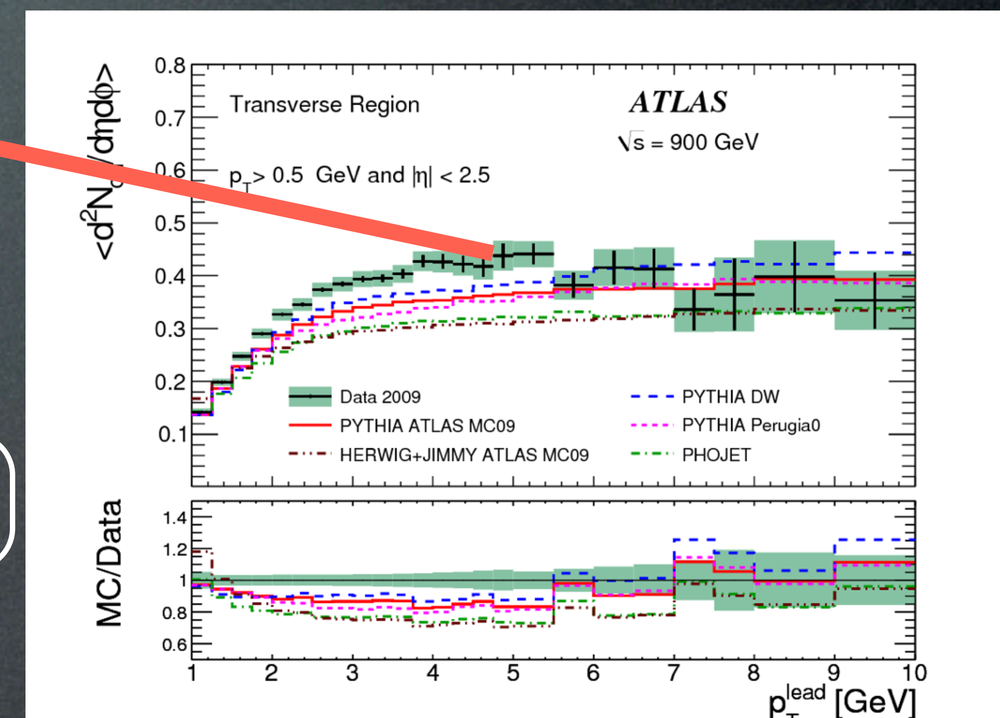
0.4 ->2.5

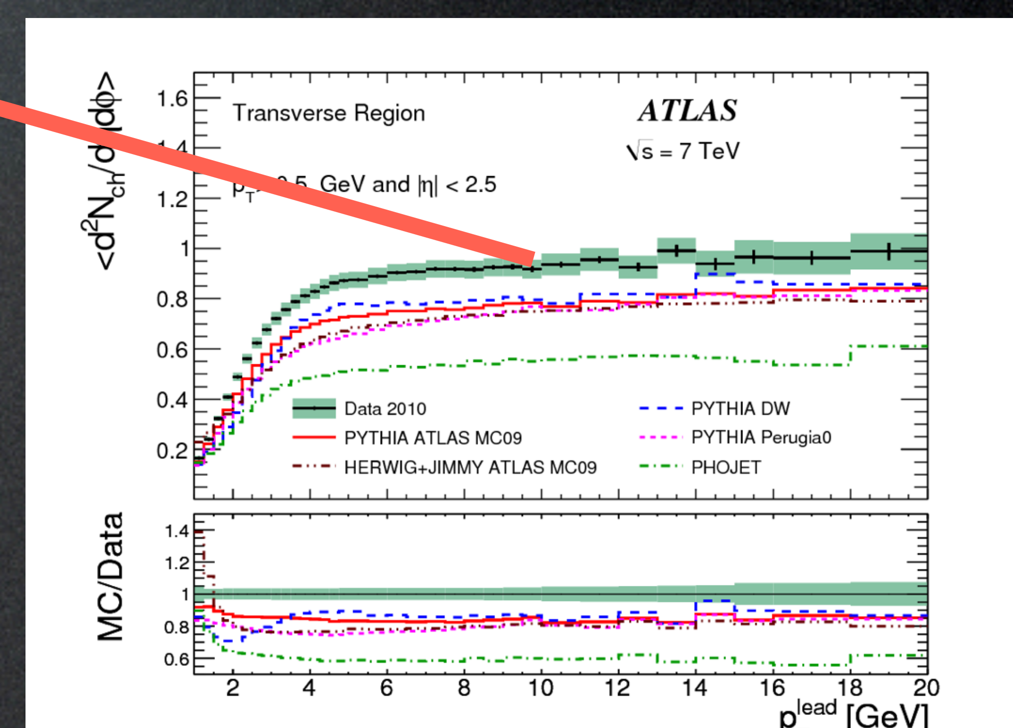
~1.3

factor of 2!

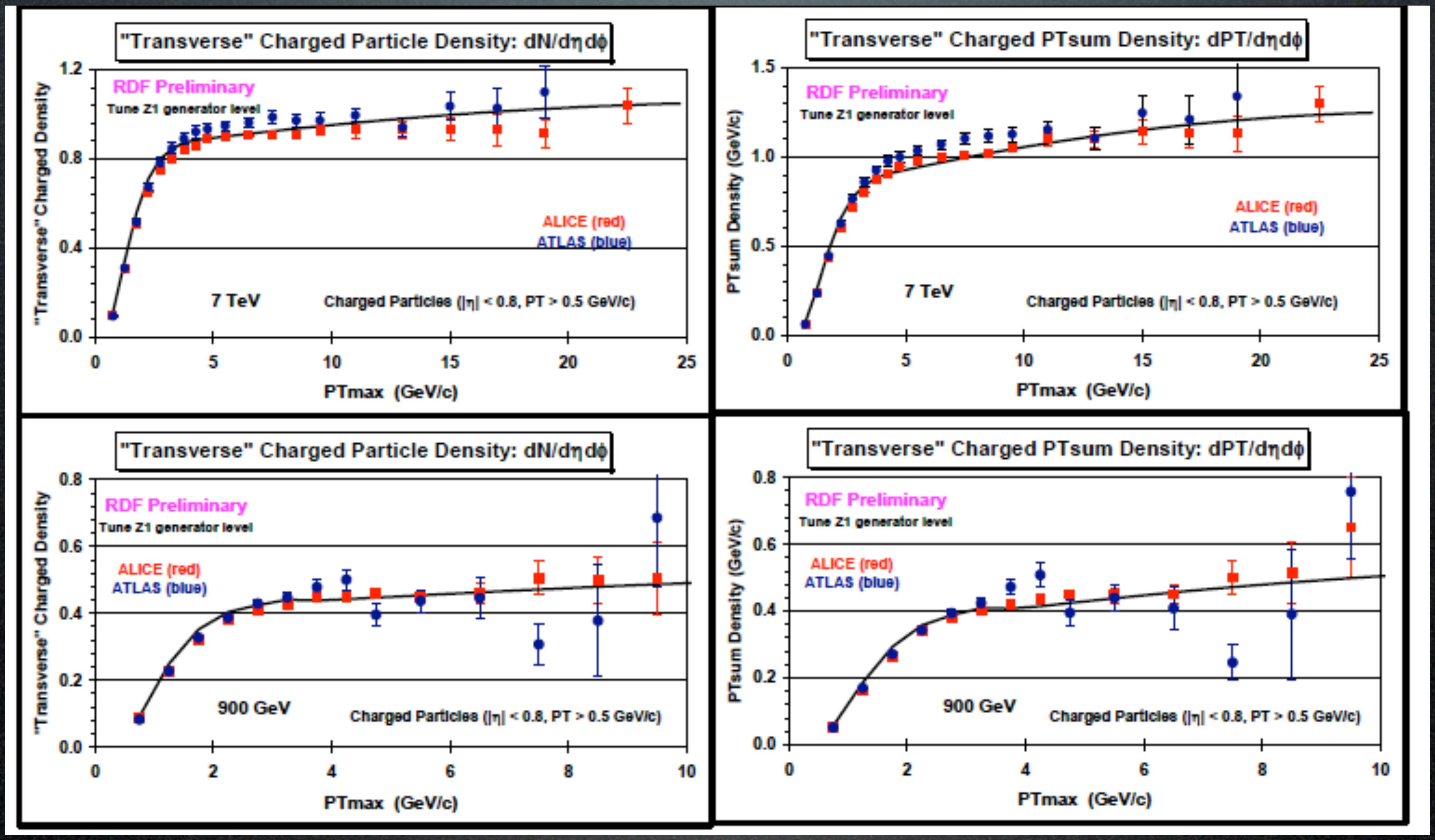
0.8 -> 5

~2.4





LPCC UE&MB WG

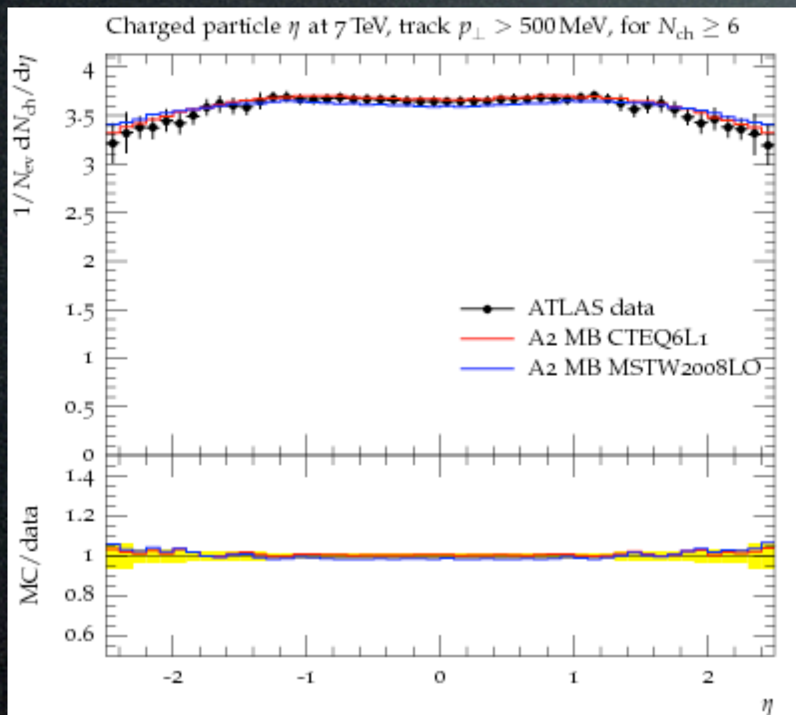


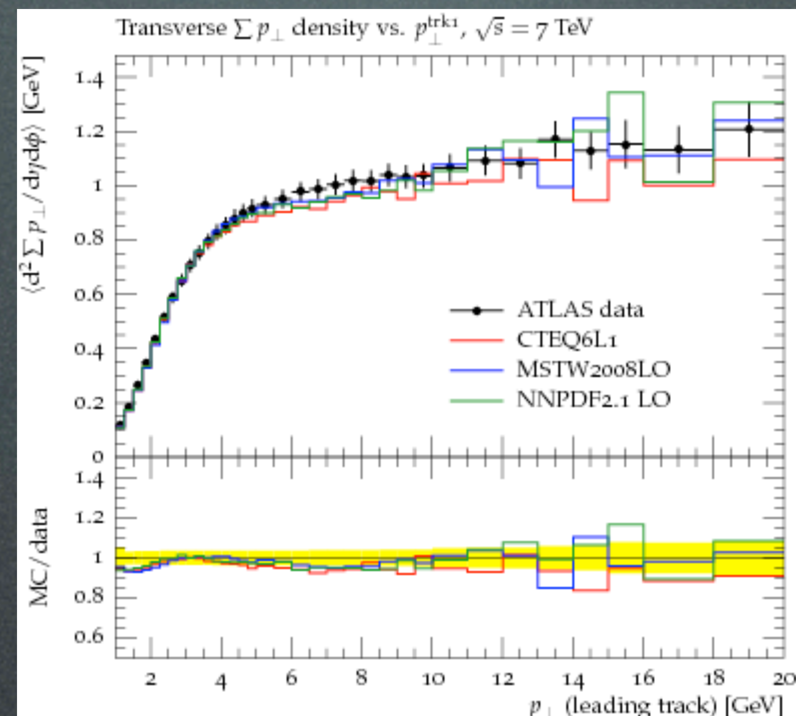
Rick Field: WG meeting, 17th June 2011

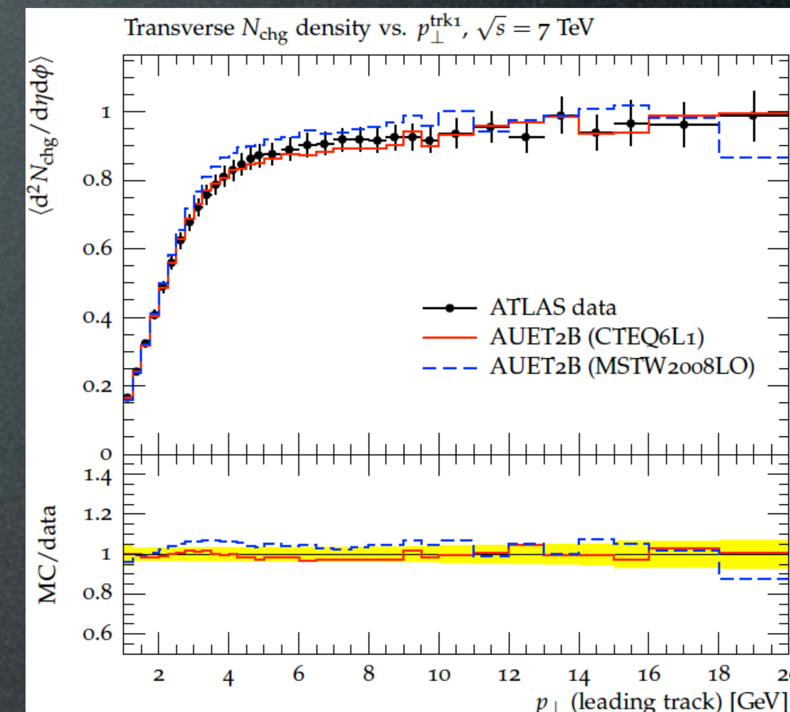
Post-LHC Tunes

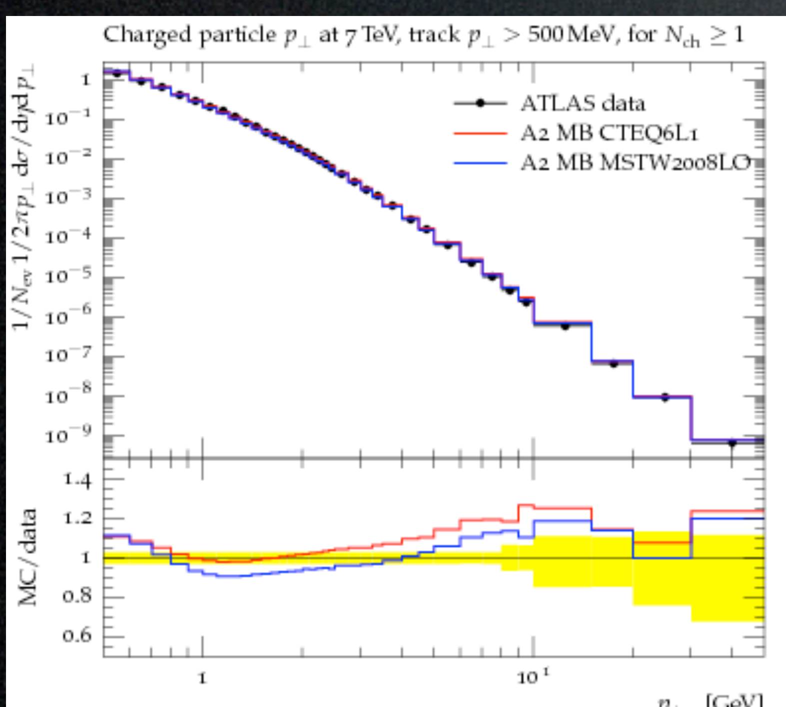
- (Pythia 6)ATLAS Tunes: AMBT1, AMBT2, AMBT2B, AUET2, AUET2B. [First separate MB/UE Tunes. also for many PDFs.]
- (Pythia 6) CMS Tunes: Z1, Z2, Z2*.
- (Pythia 6) Perugia 2011 tunes.
- (Pythia 8) author tunes: 4C, 4Cx.
- (Pythia 8) ATLAS tunes: A2, AU2.

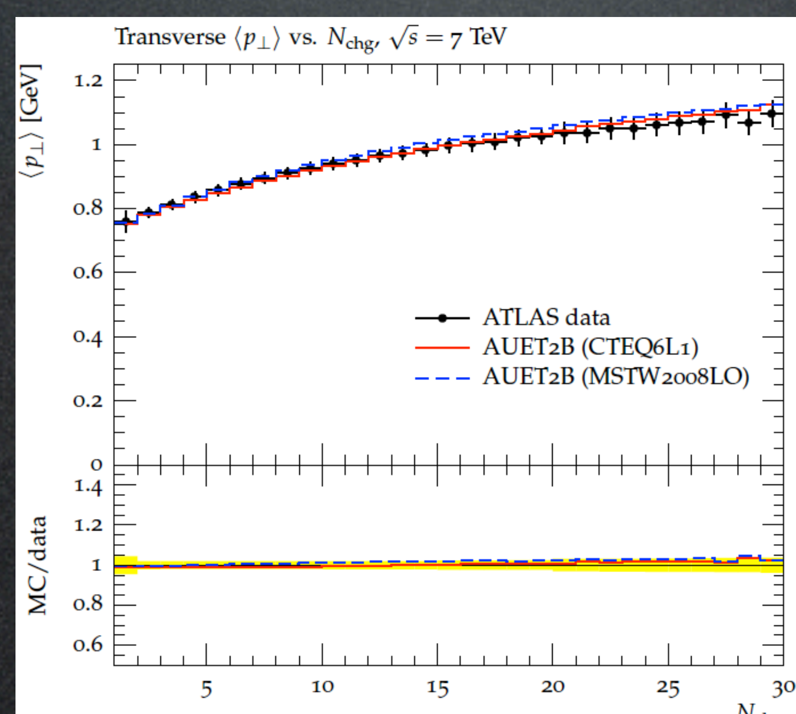
How do they do?





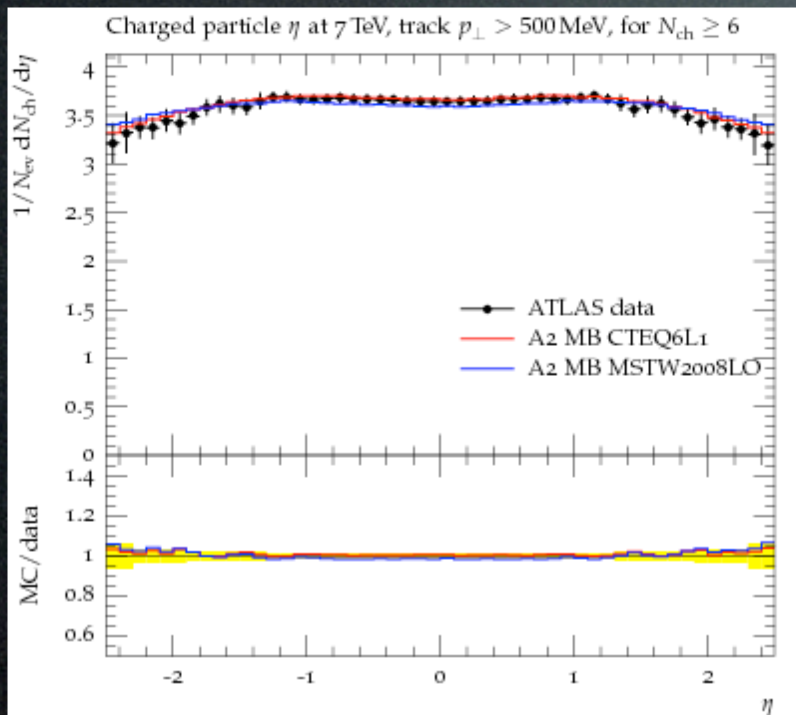


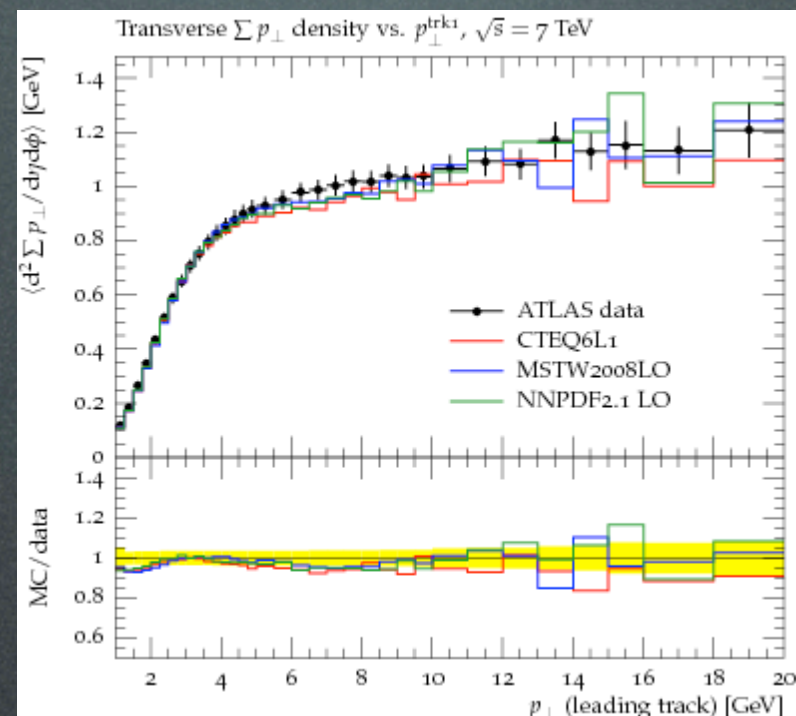


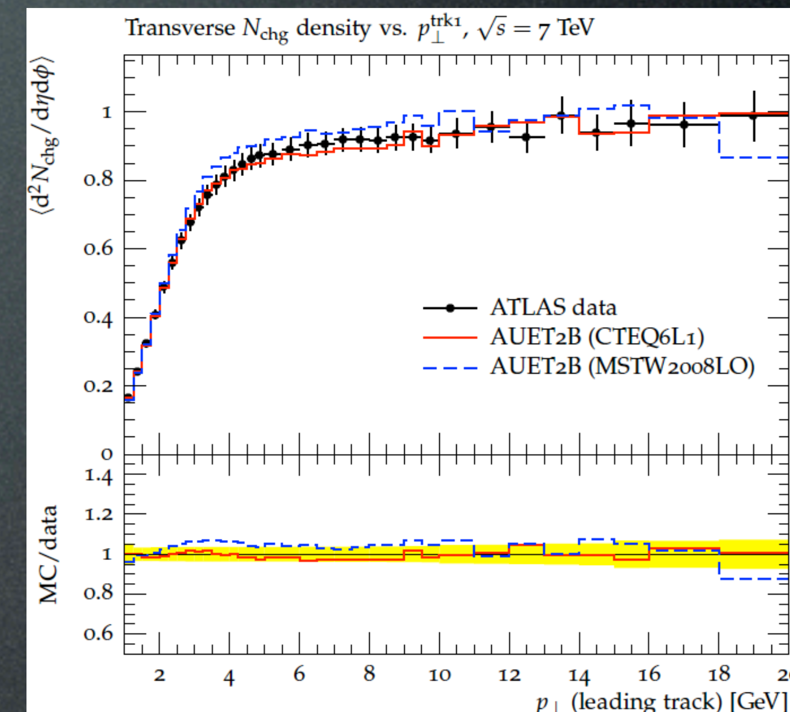


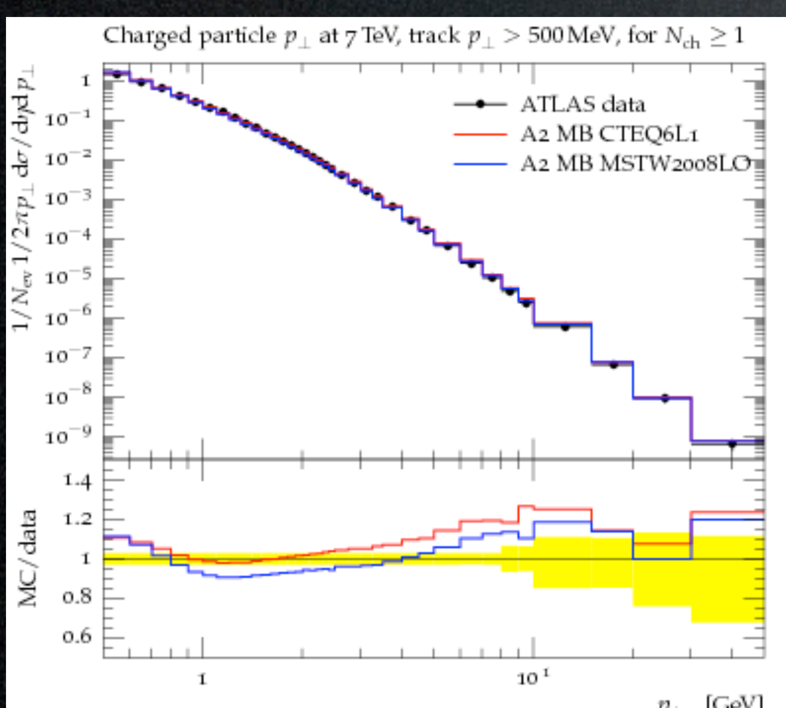
How good is good?

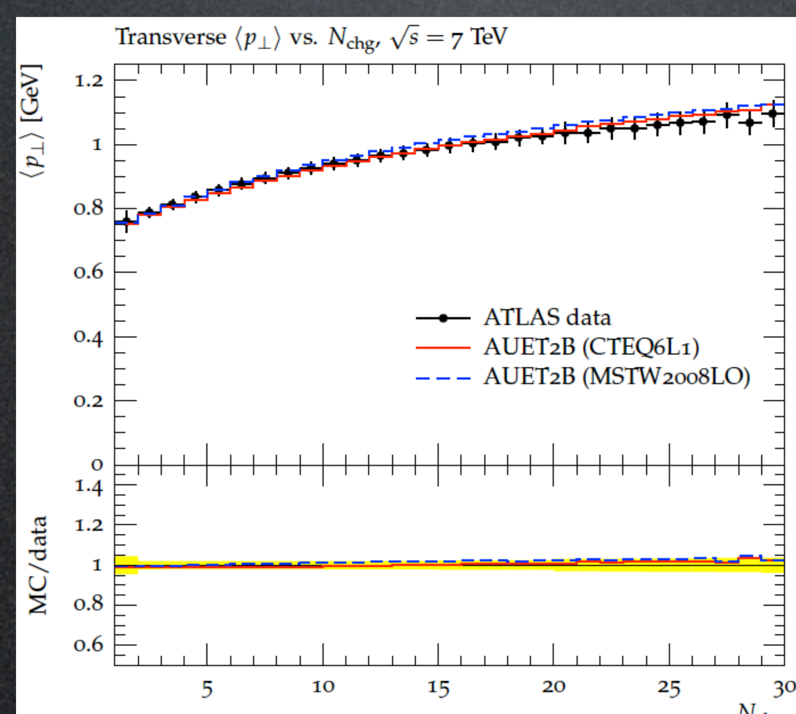
How do they do?





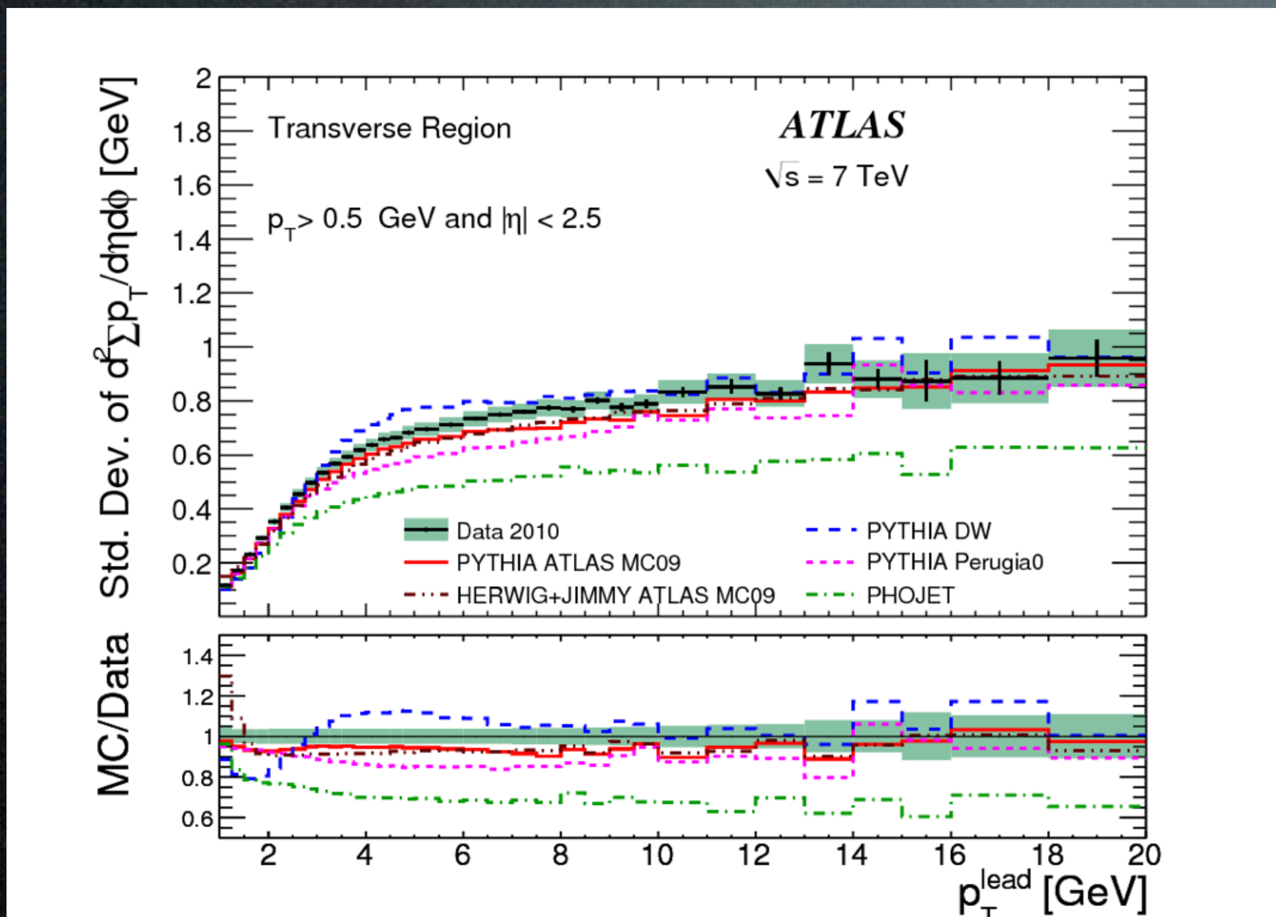




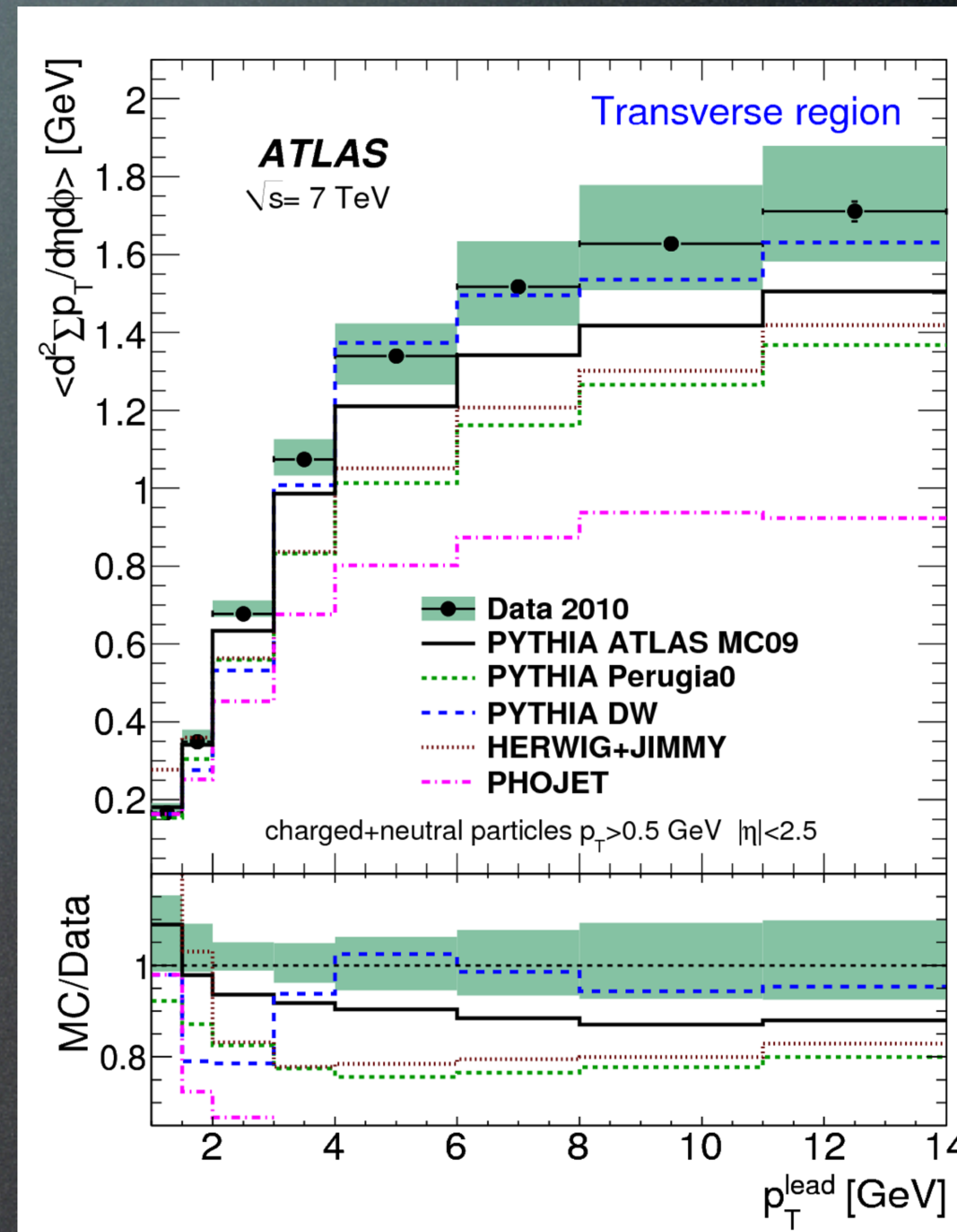


How good is good?
Is not it amazing that the models are doing so well?

Back to (early)UE Results

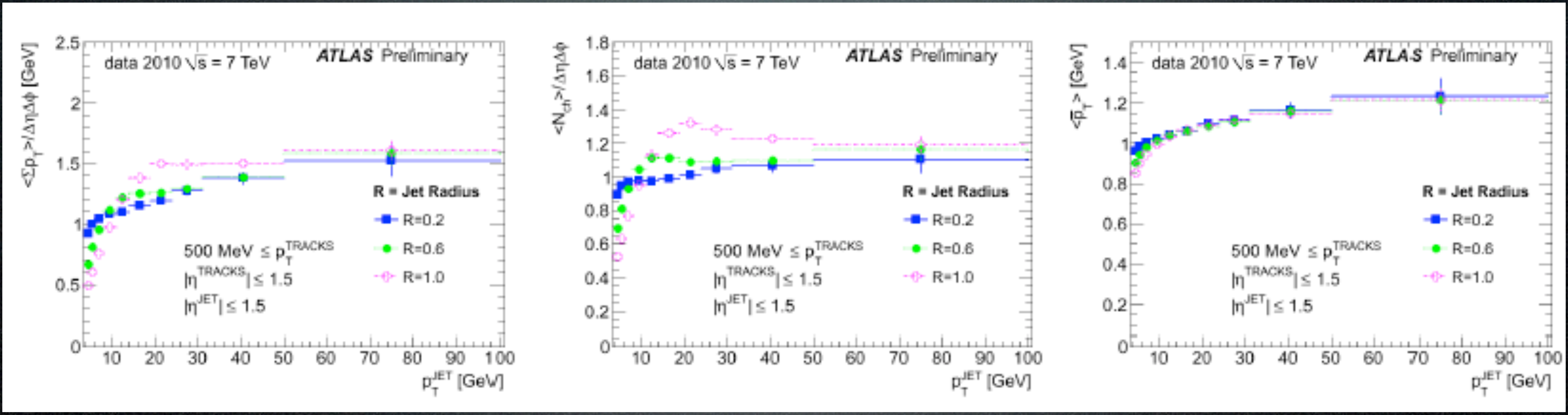


shows UE activity can not be subtracted as an average "pedestal" from each event.



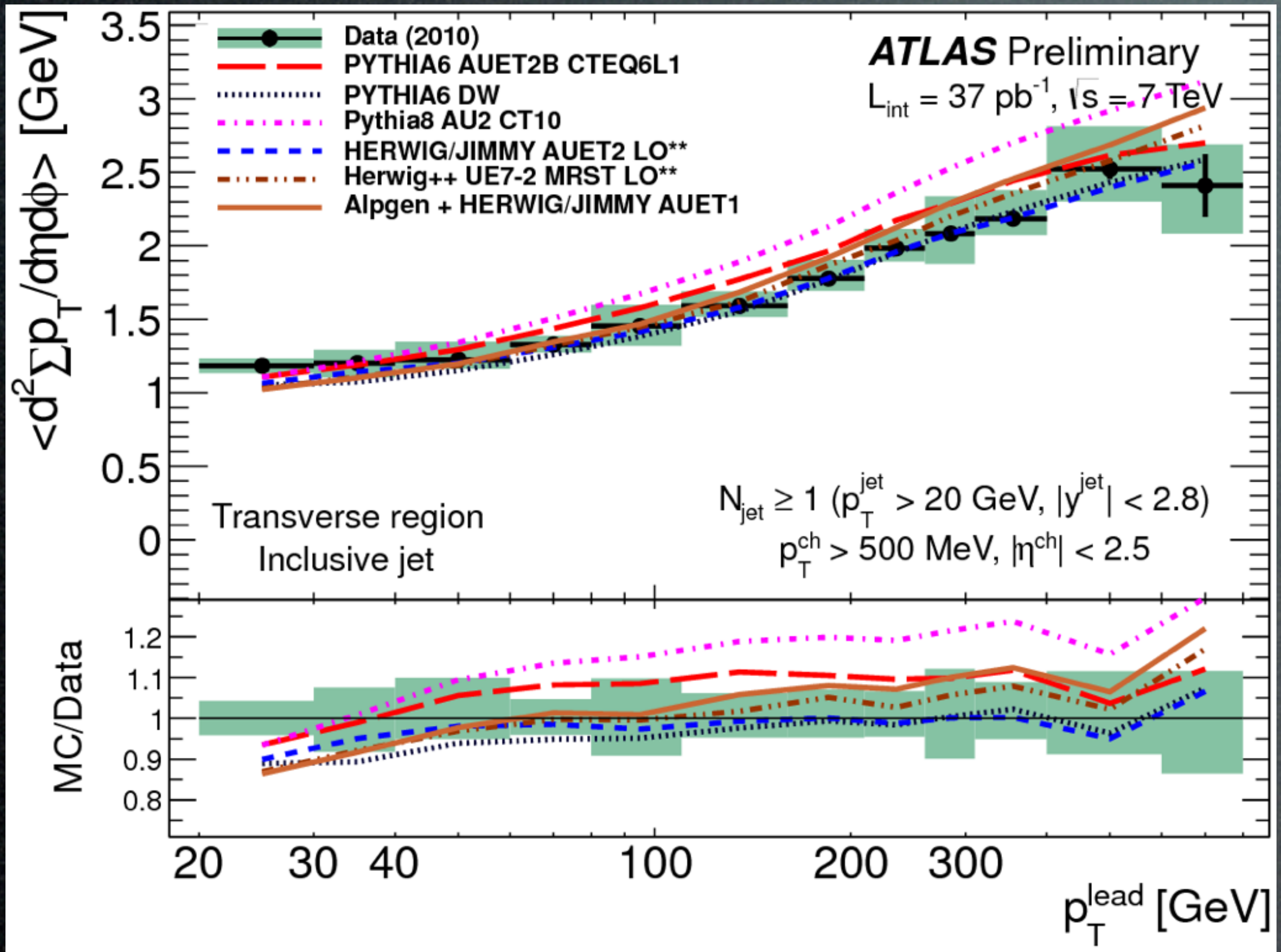
Sensitive to both charged and neutral component of UE.

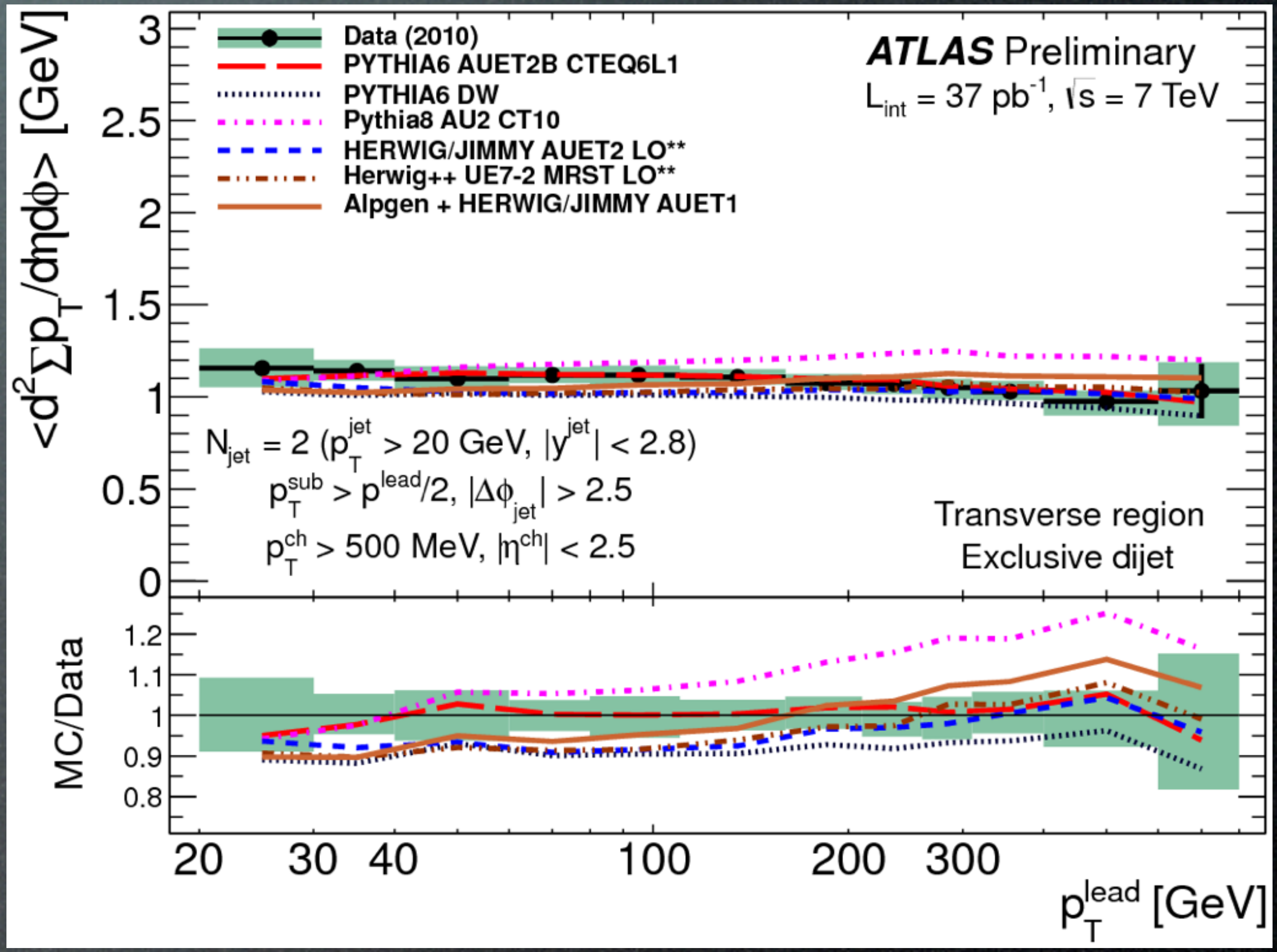
Newer ATLAS UE Results



More UE activity for higher jet radius. Why?

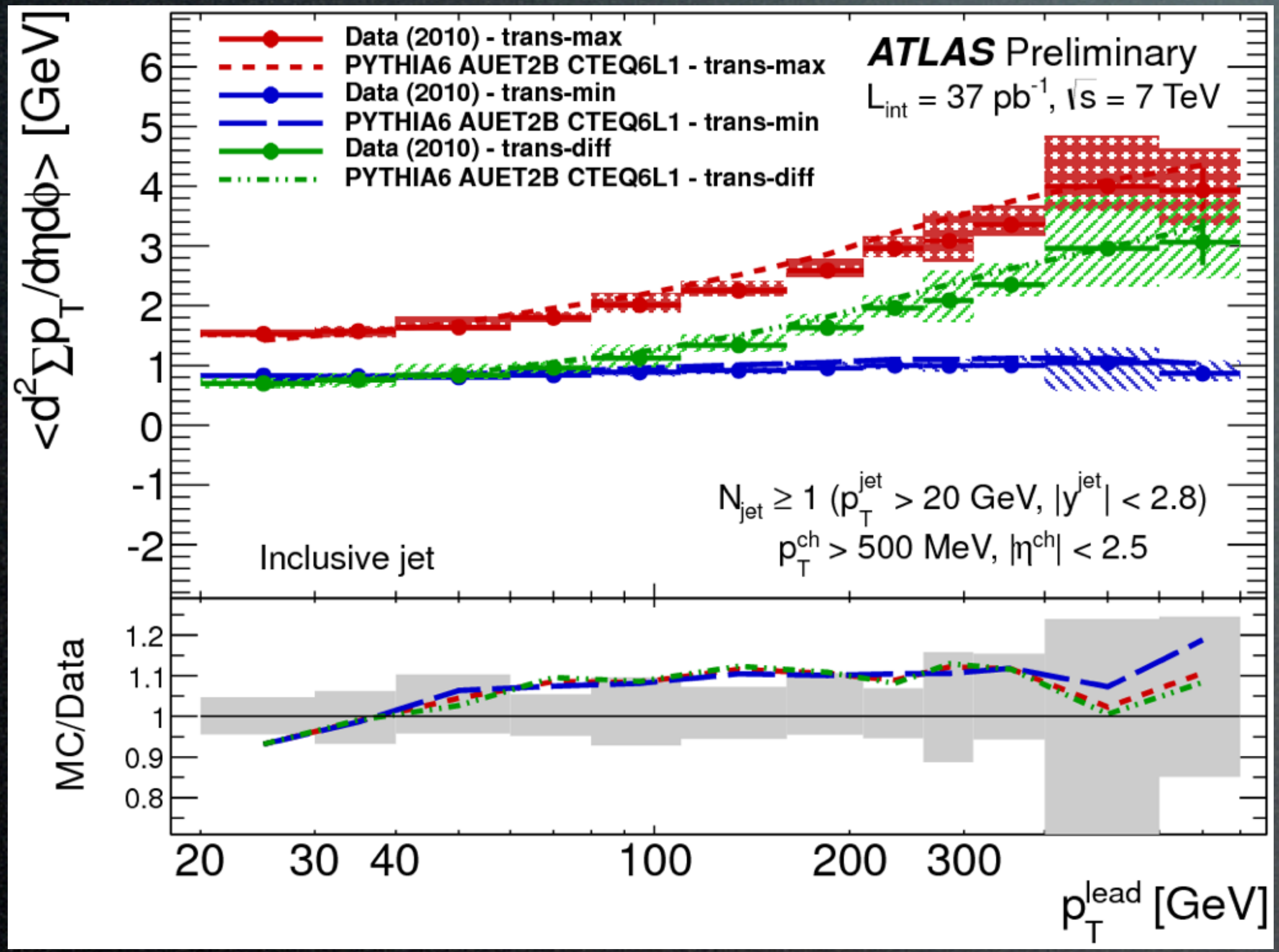
Newest ATLAS UE Results

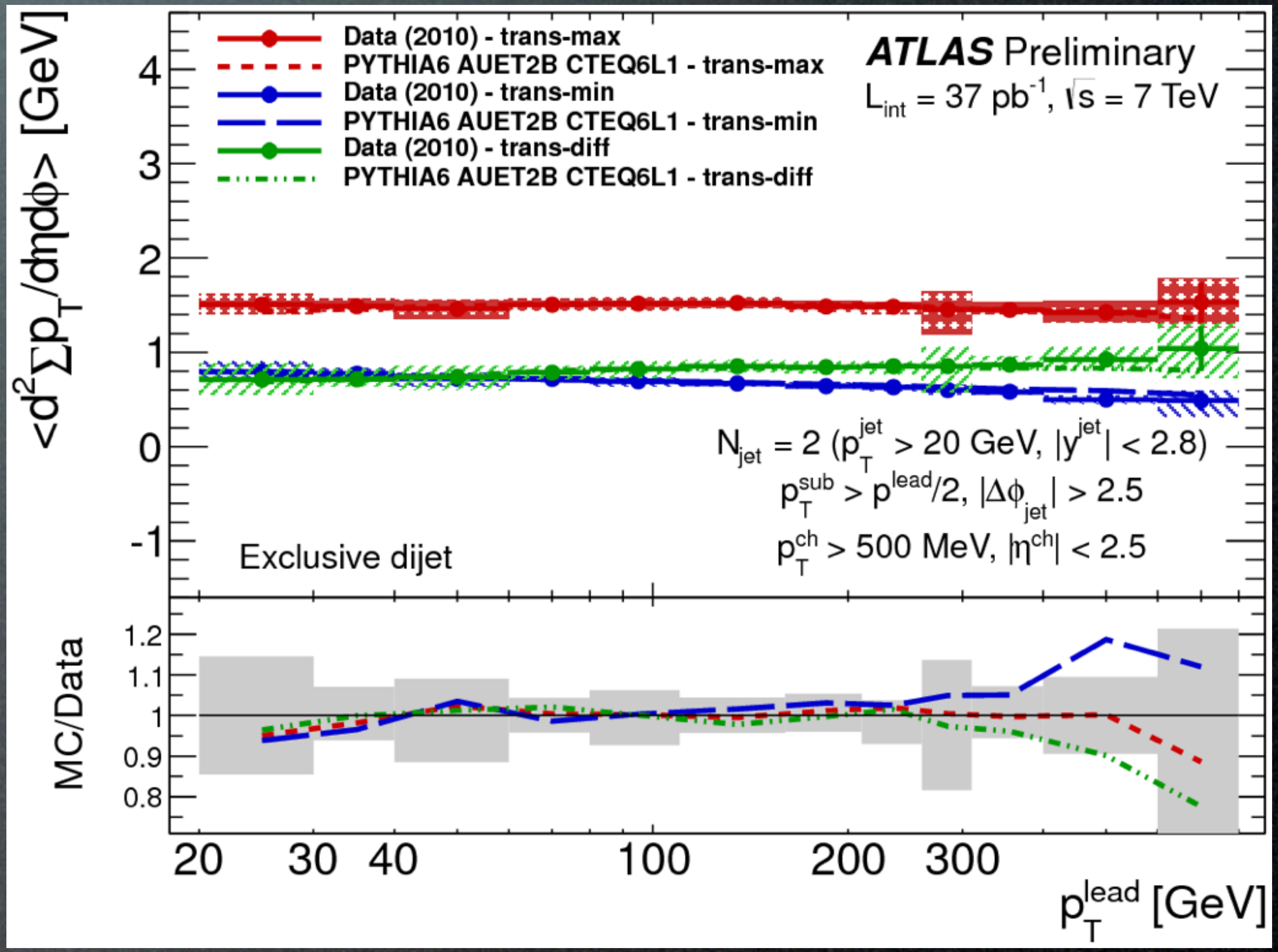




Rise in inclusive, almost flat in exclusive

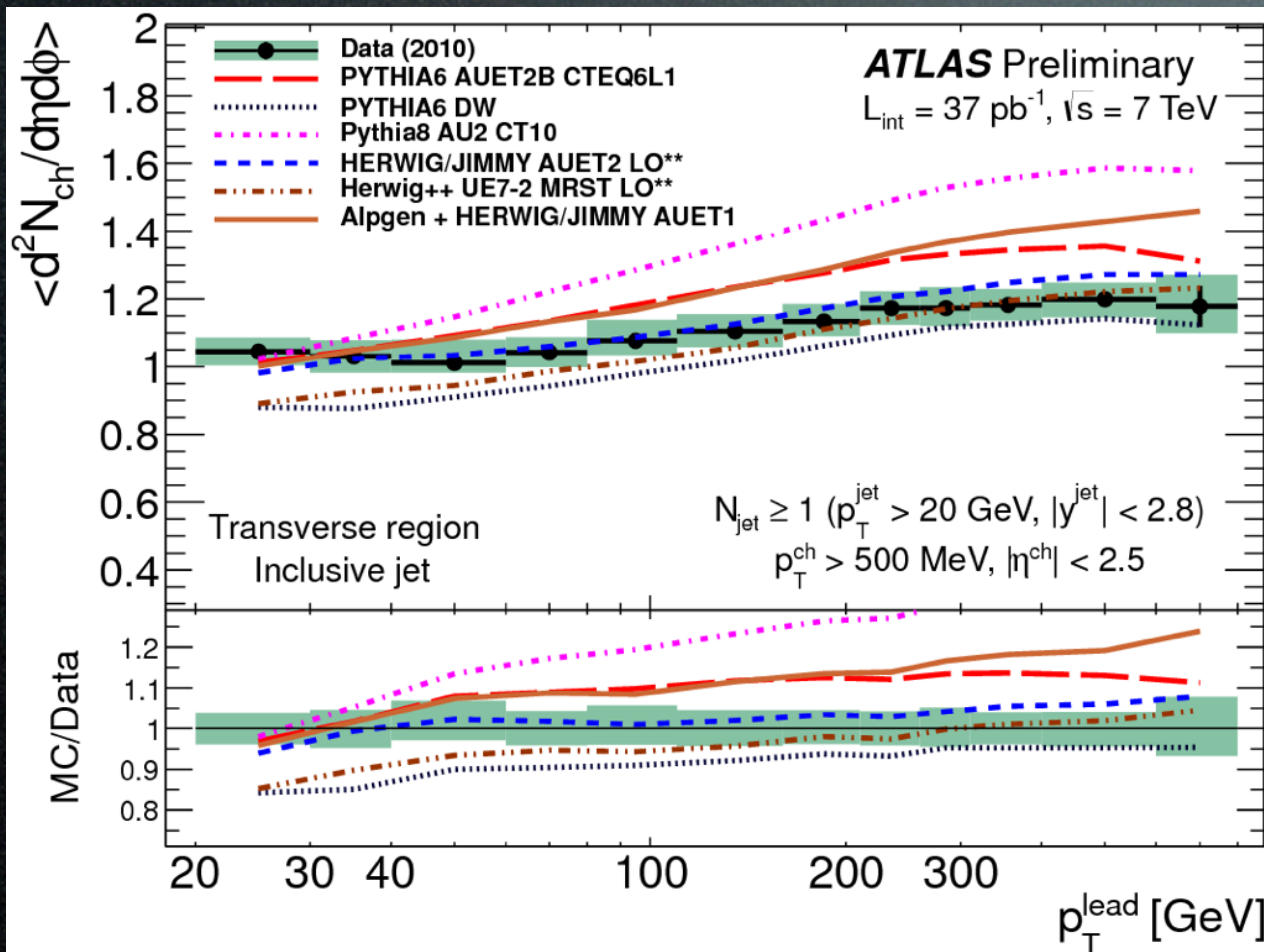
Transmax/min

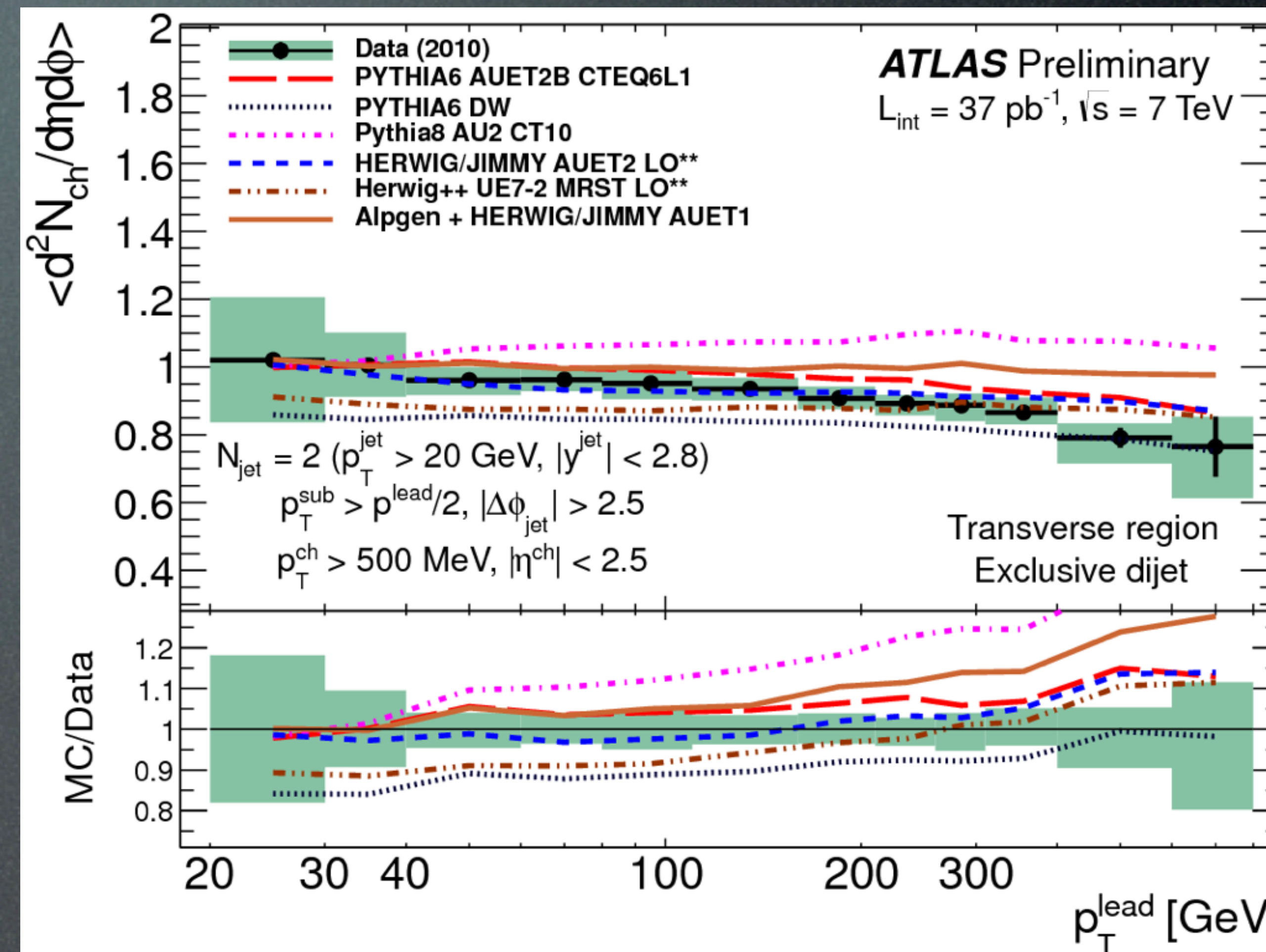




Max/min gets closer in exclusive

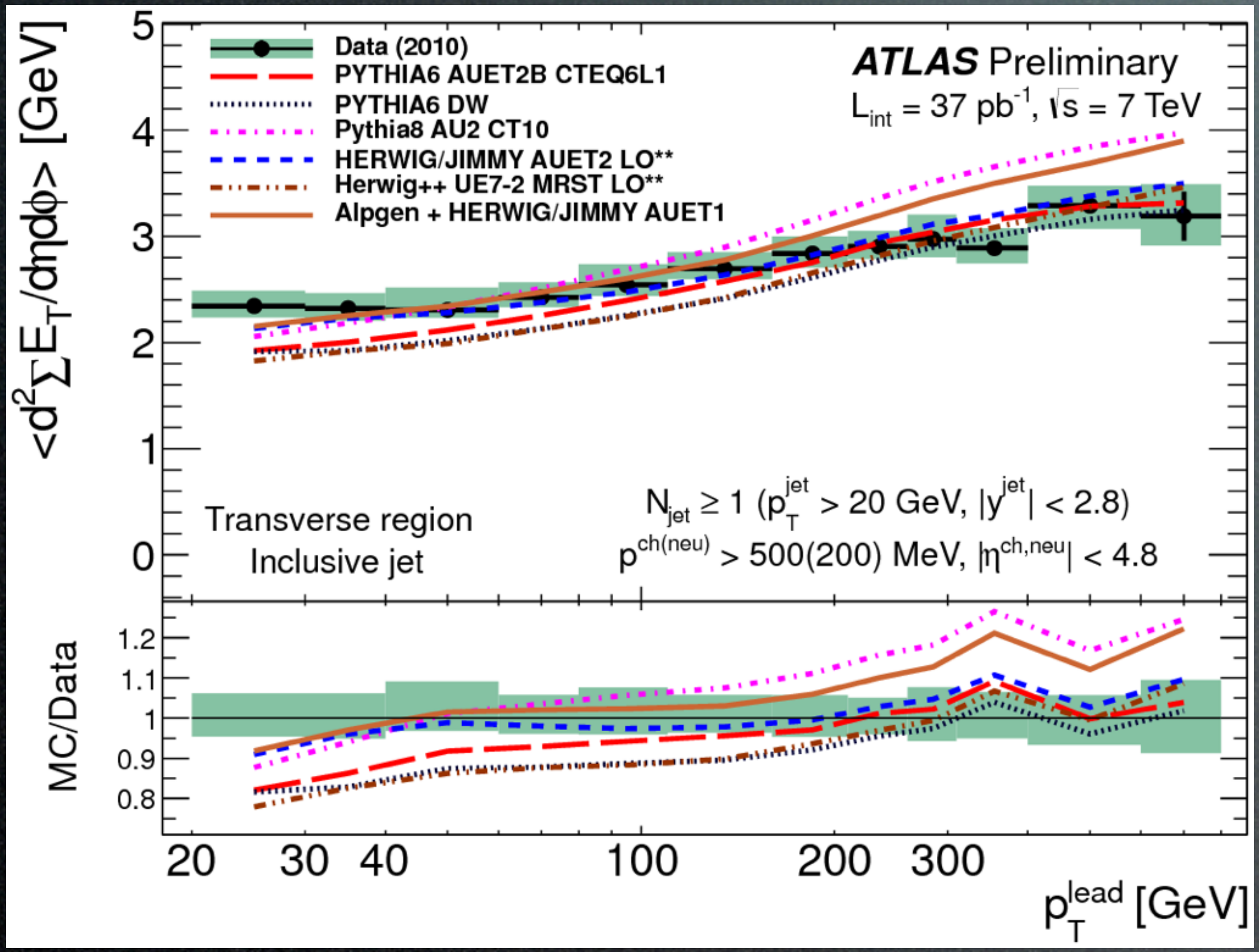
...continued

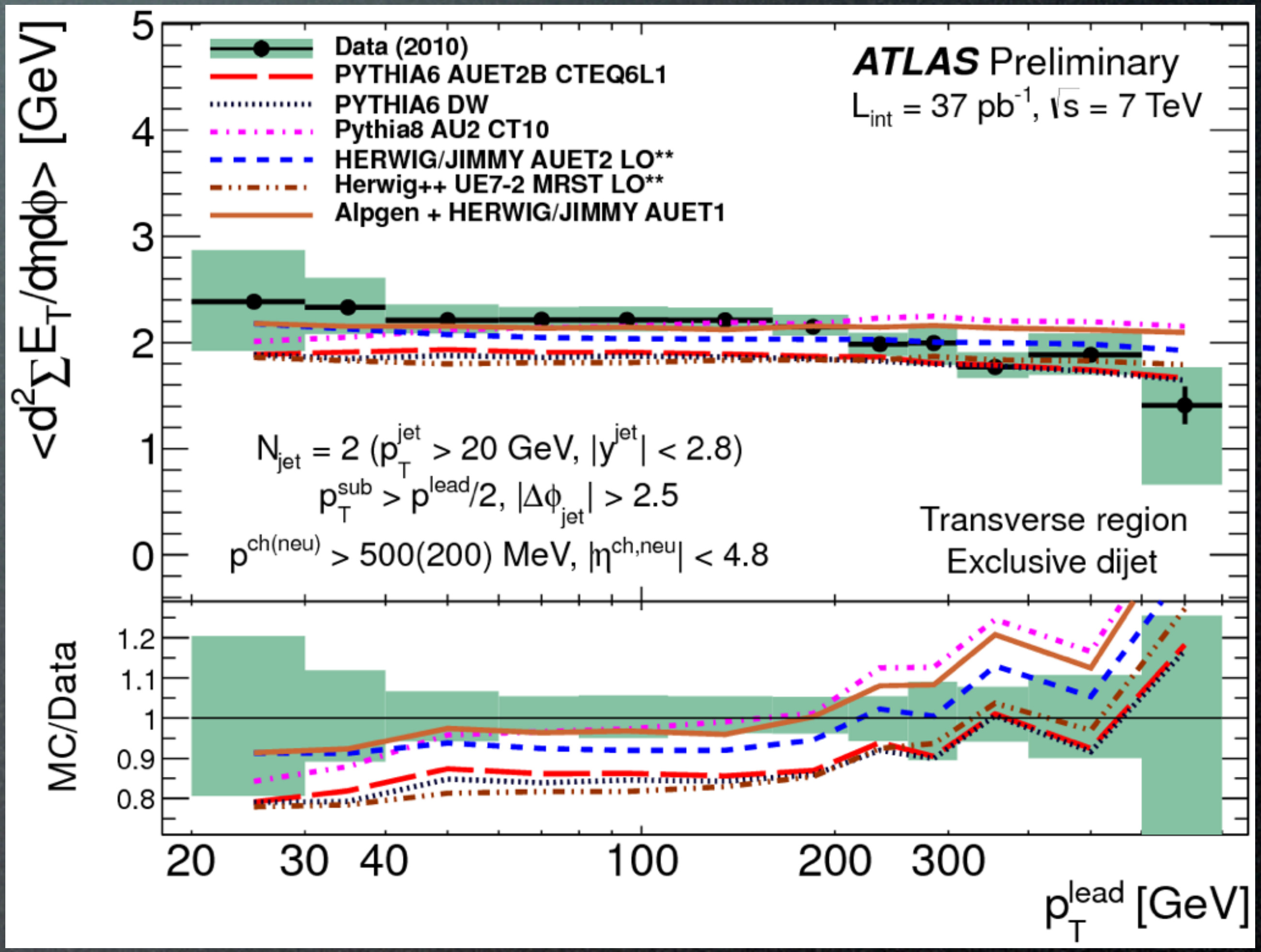




Same trend as in sum p_T

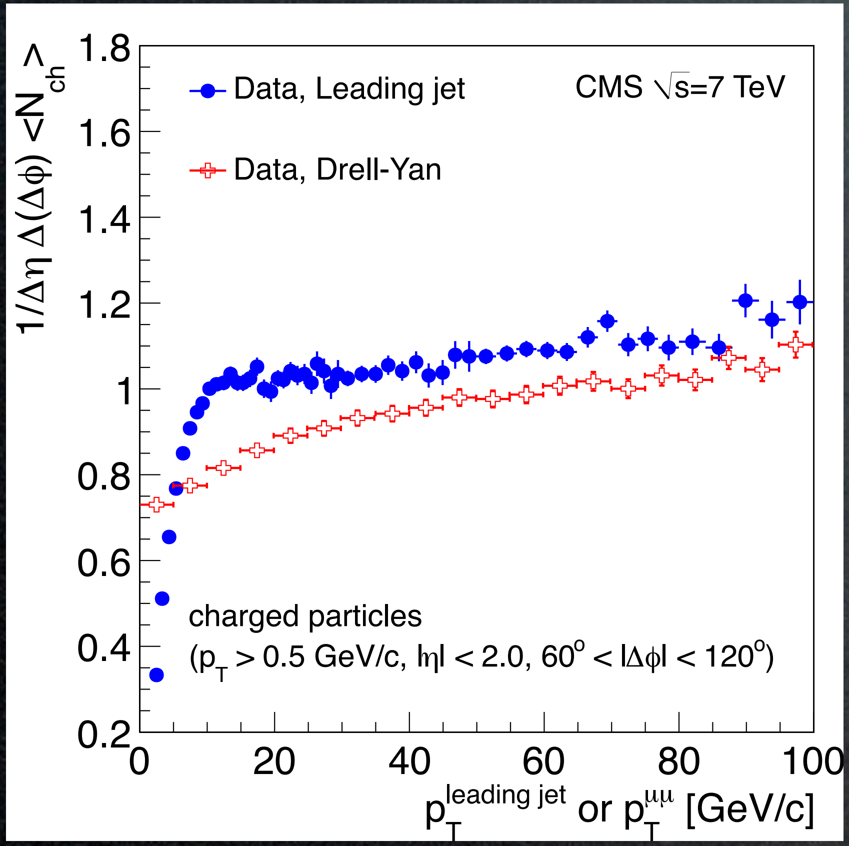
Including Neutrals

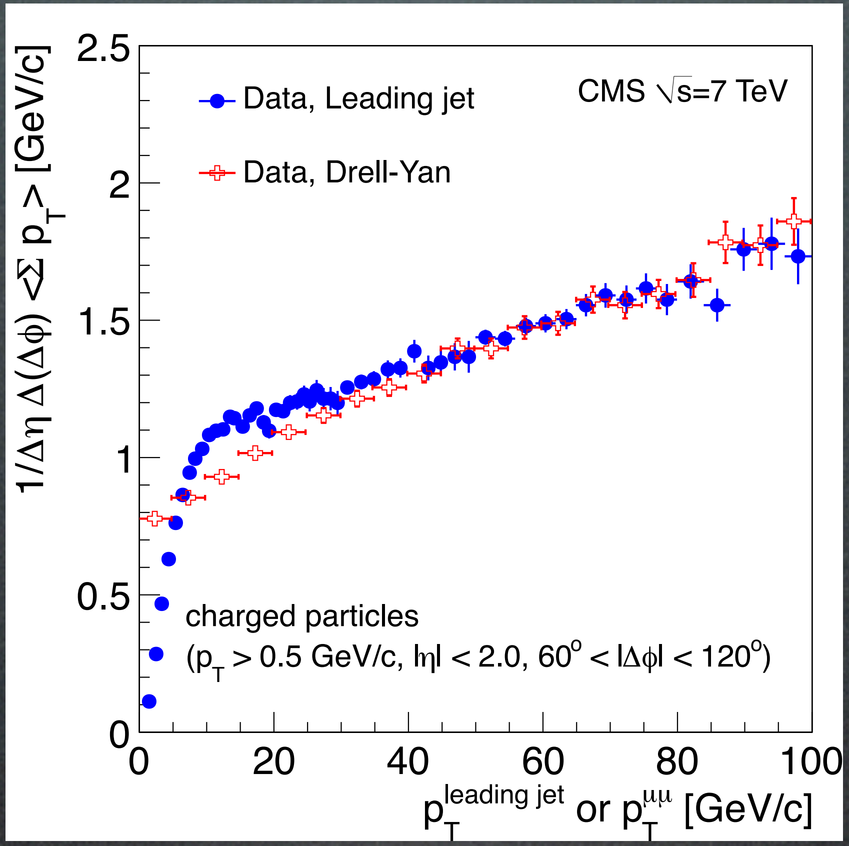


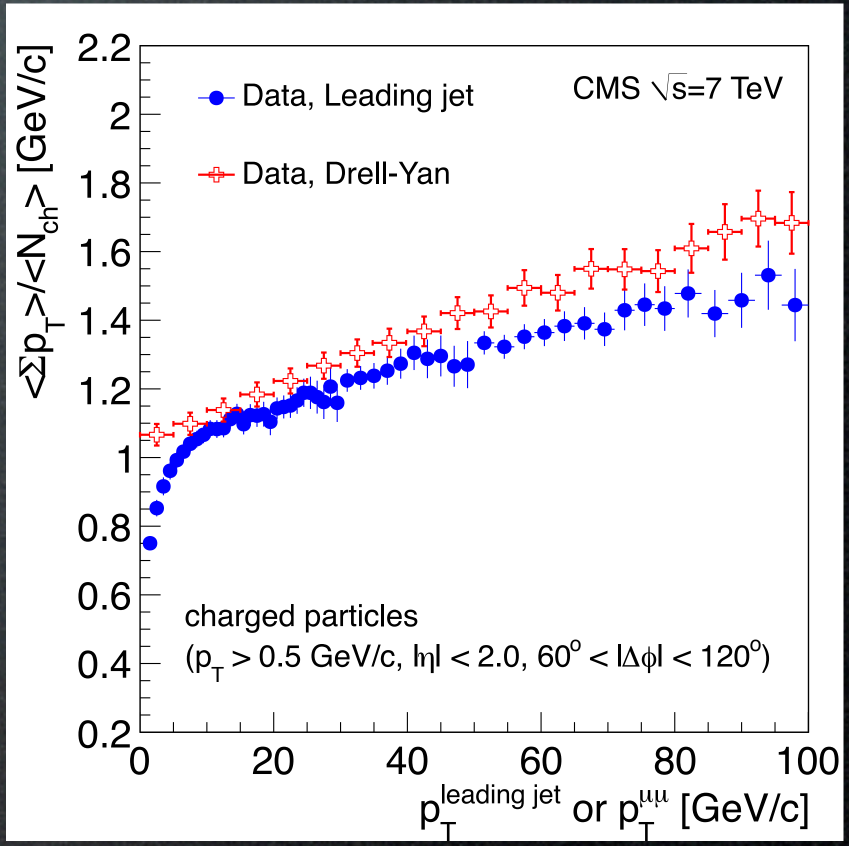


We can do charged to neutral ratio

Coming back to Z-jet UE difference



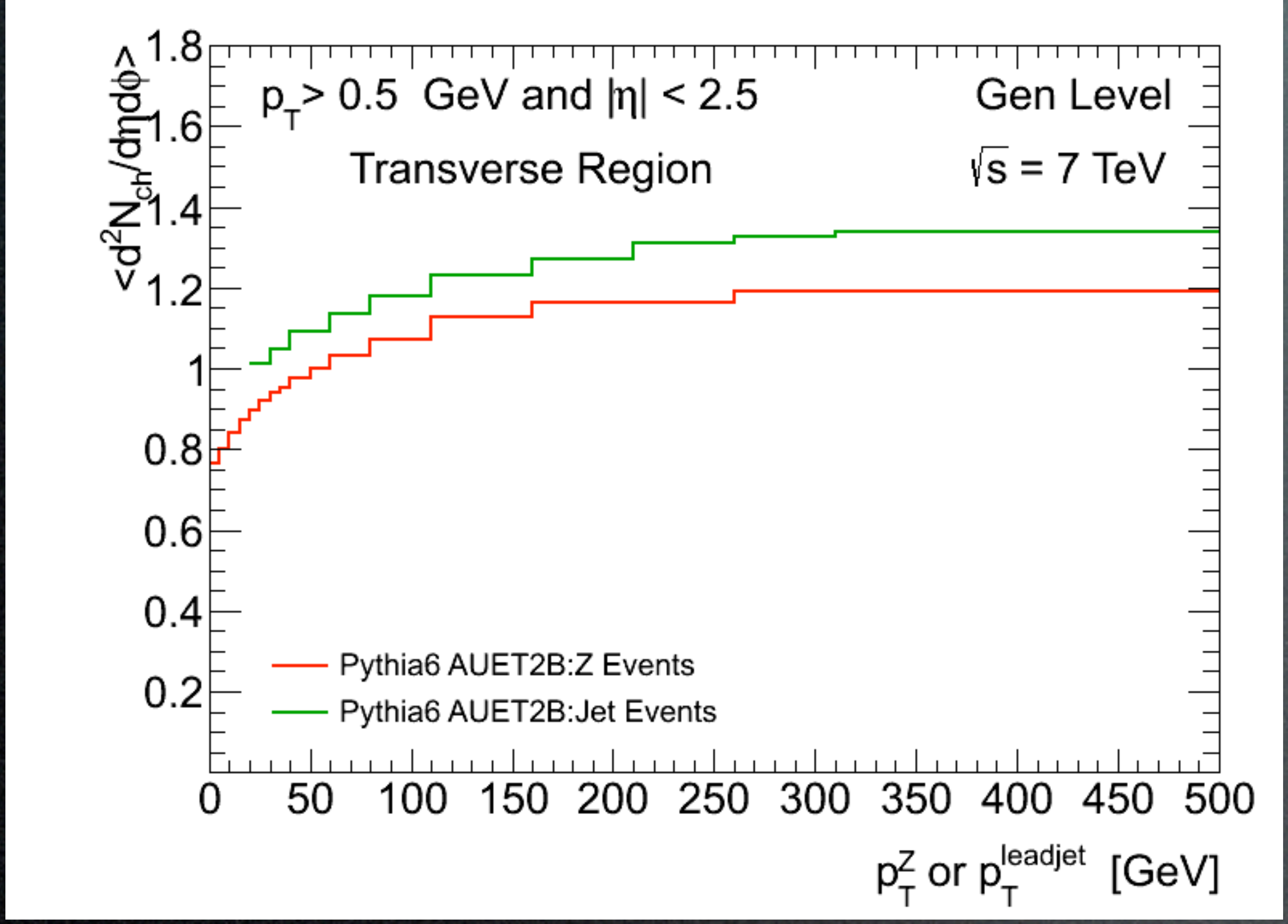


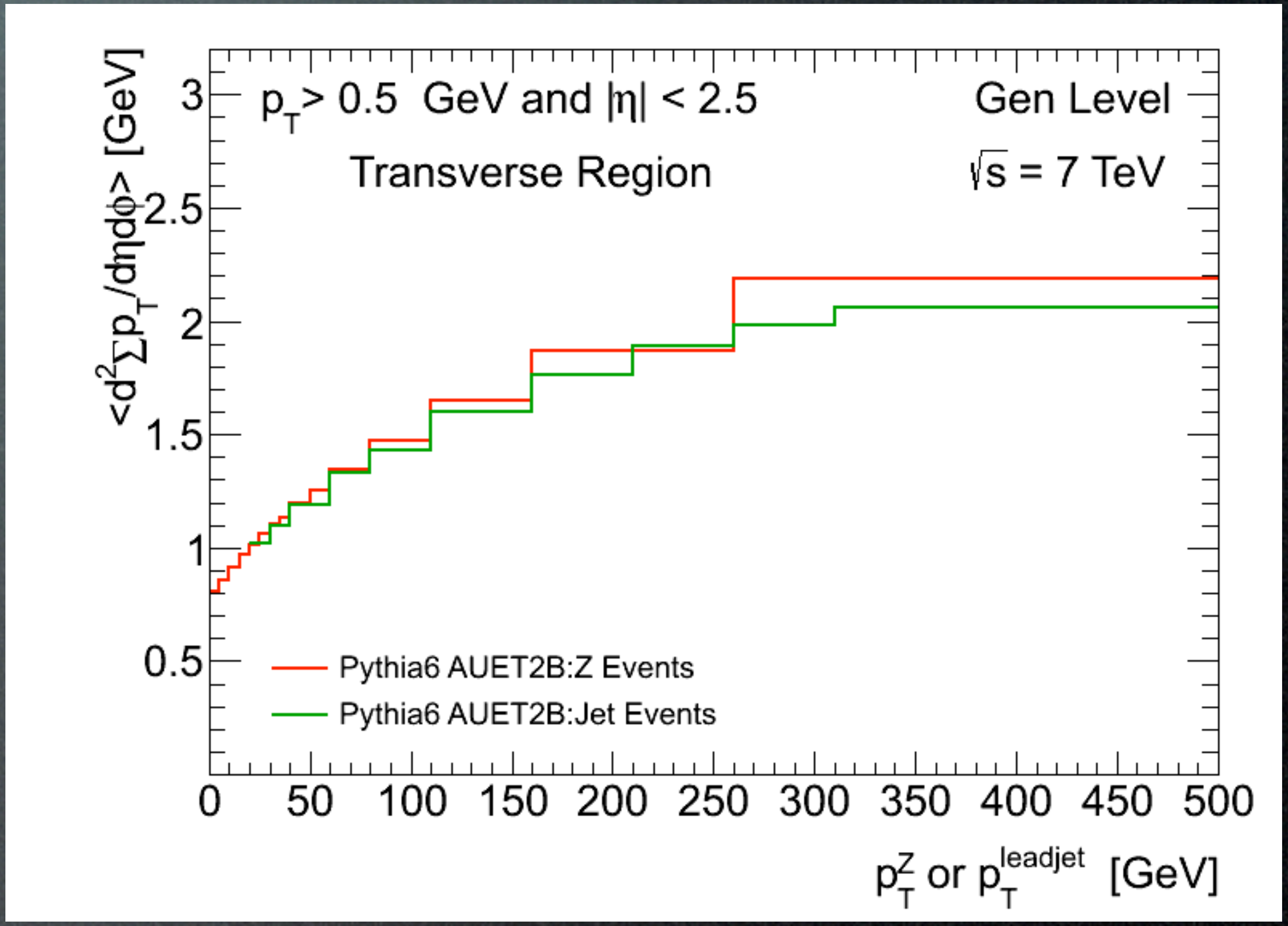


Eur. Phys. J. C 72 (2012) 2080

CMS results show they are still similar

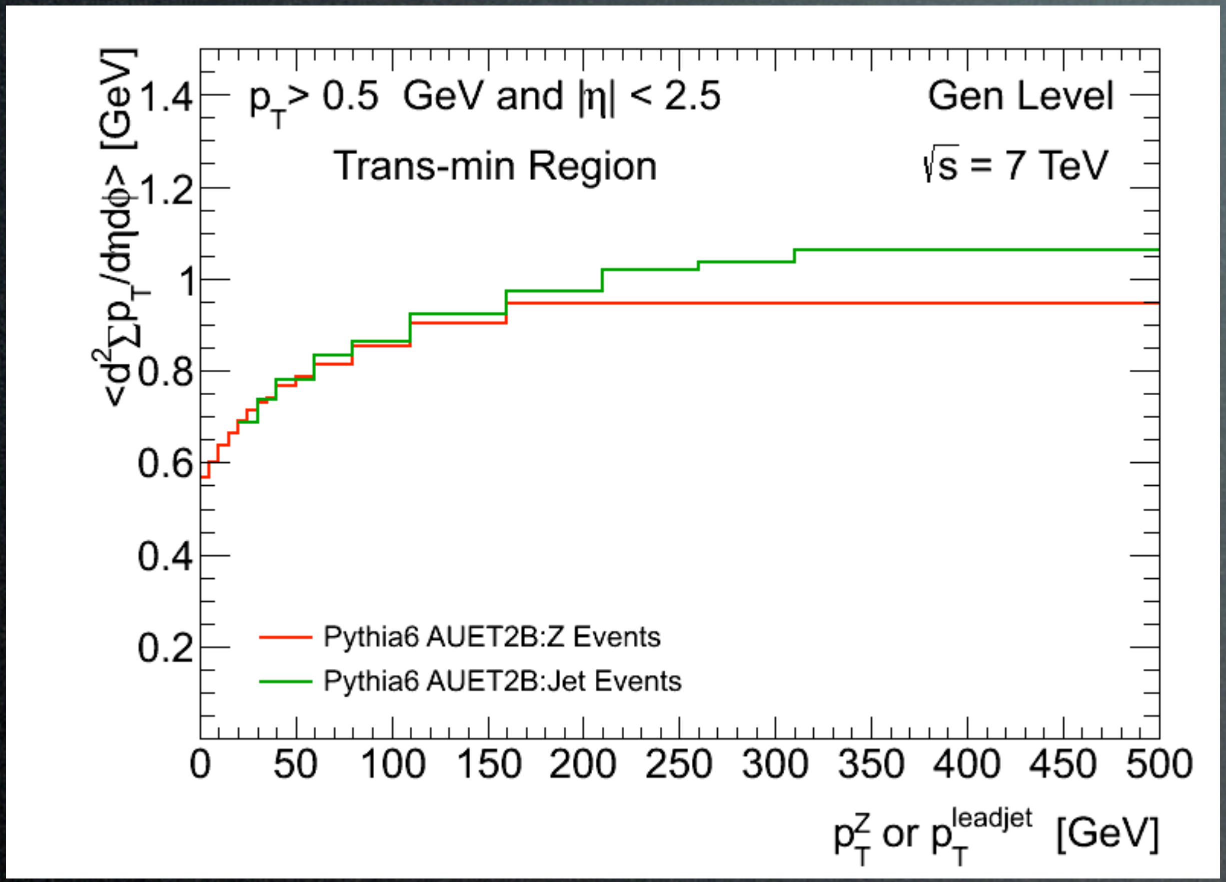
Extend to Higher Energy Scale

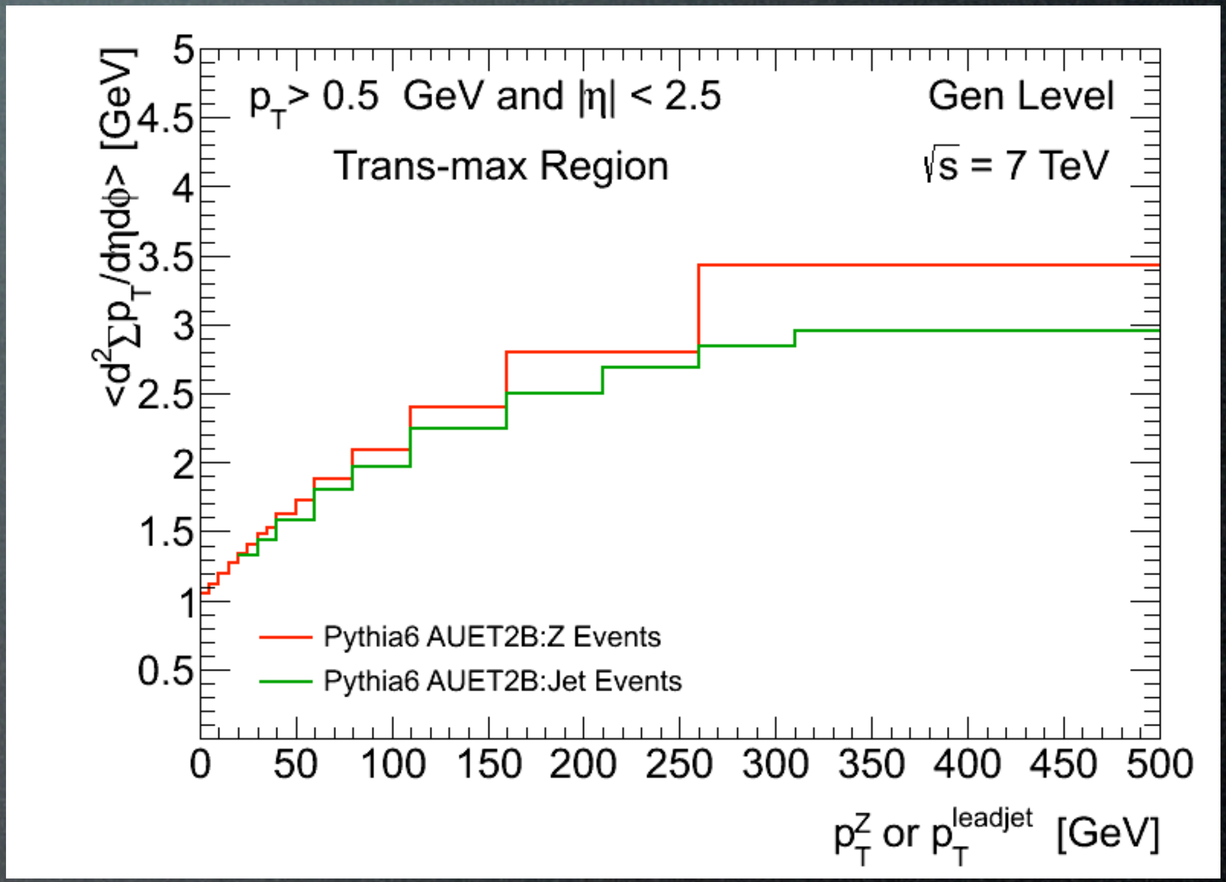




Generator level (Pythia6): Not much difference

Trans-max/min Regions

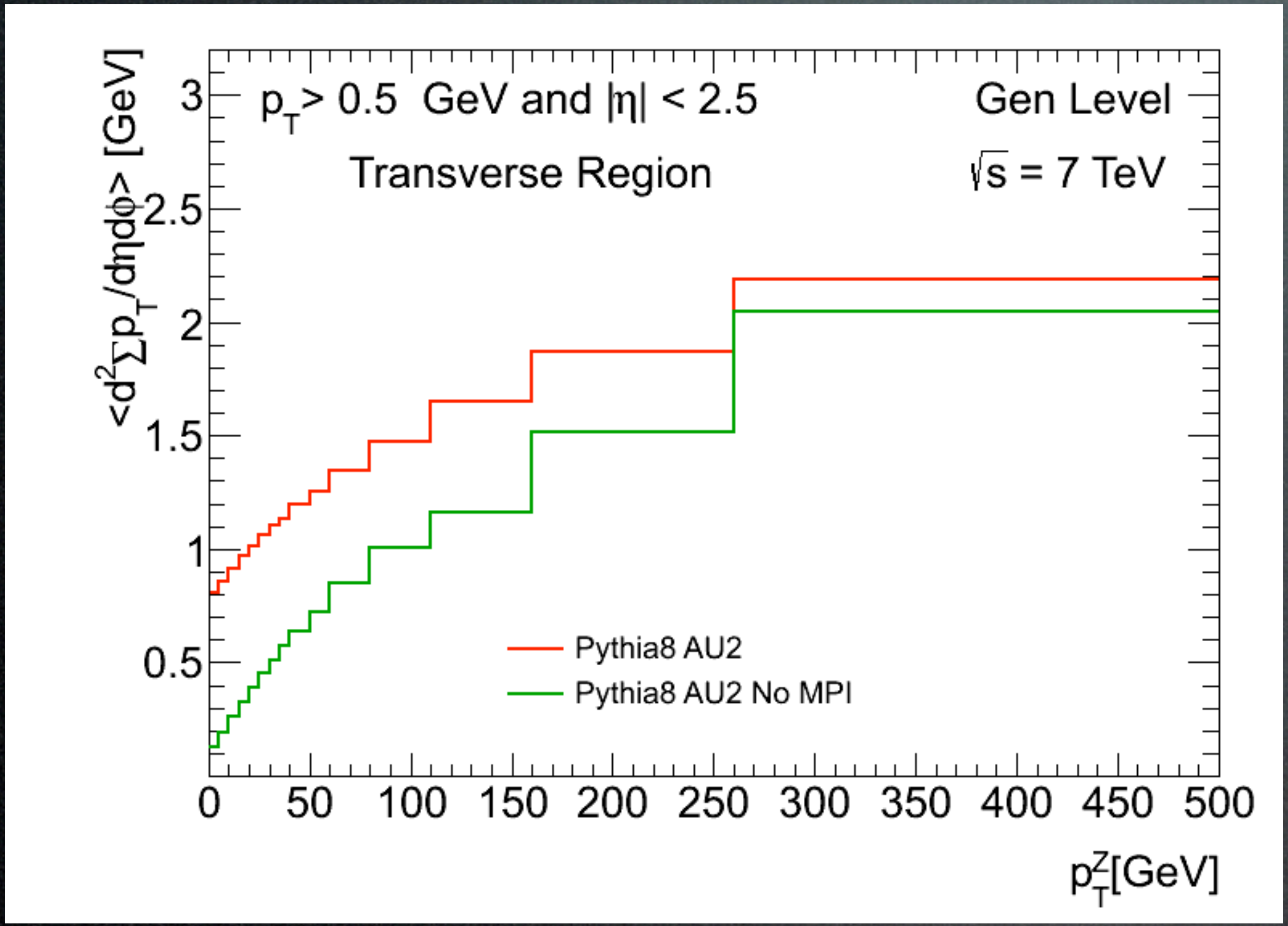




There is a pronounced difference here!

The activities are still similar, with a caveat.

How much of the UE is UE?



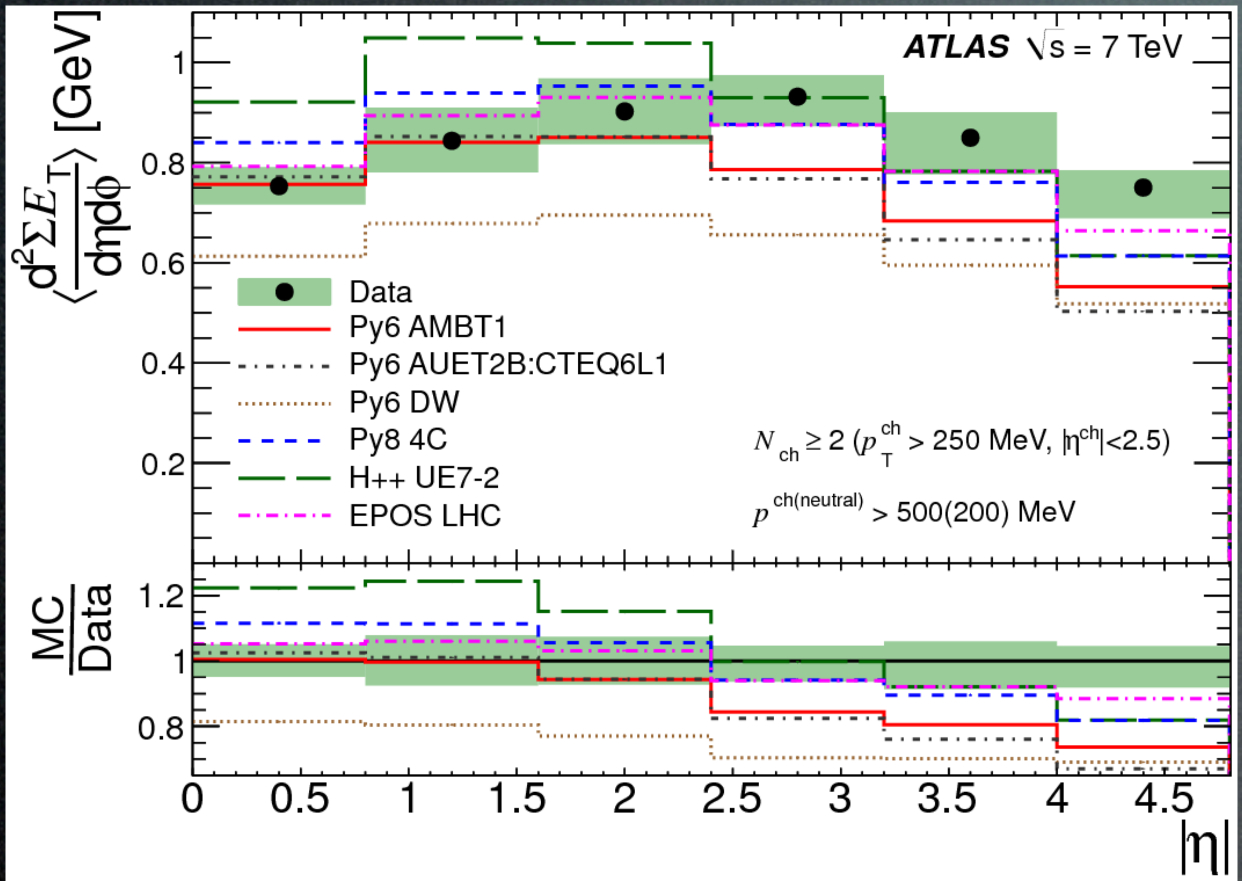
Even without MPI, the "UE" activity is catching up.

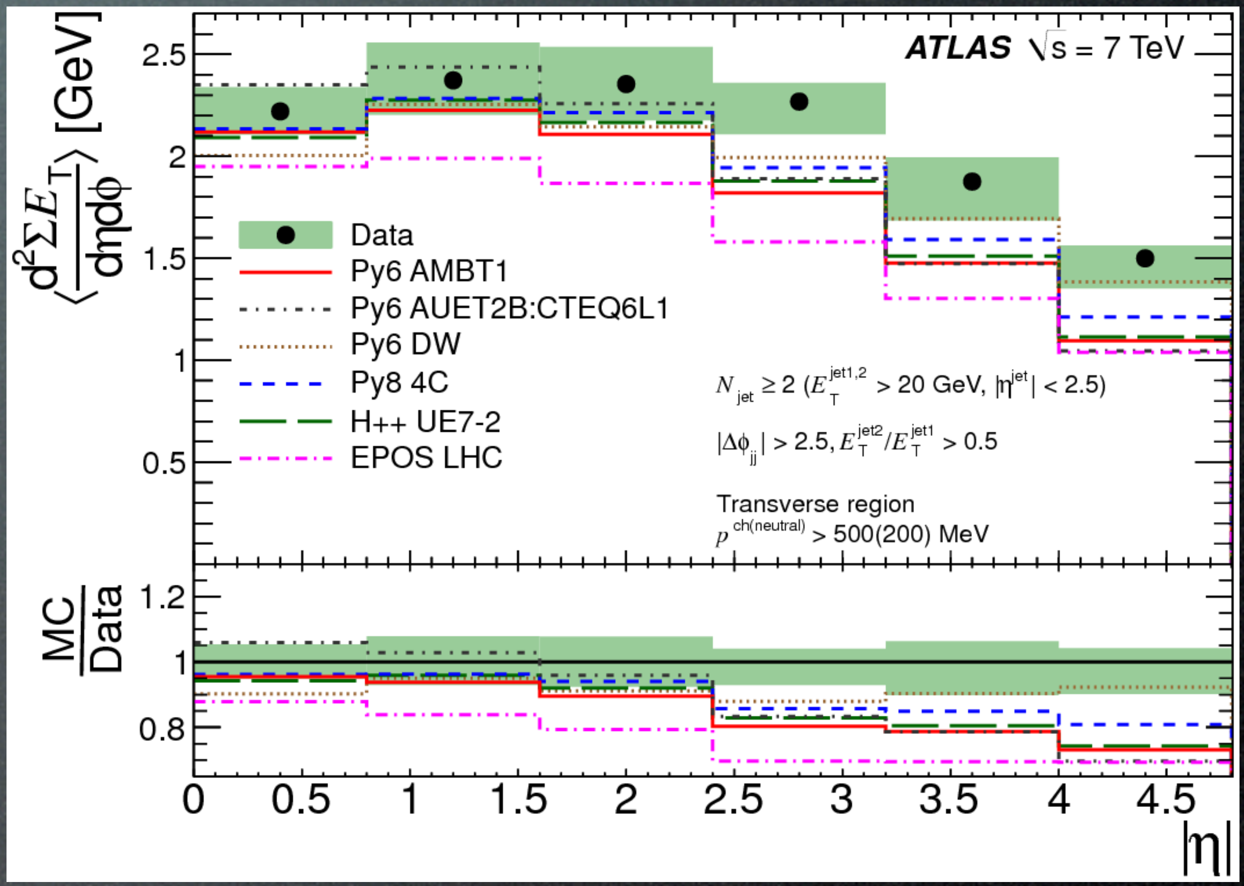
Indicative of additional hard (non-MPI) hard jets

Isolating the UE

- Full transverse (or trans-max) regions are described better by NLO or multileg generators than pure LO ones.
- Trans-min (and towards region for Z-boson events) were thought to be populated by "pure" UE.
- But at LHC, even those are not flat.

UE-sensitive Observables

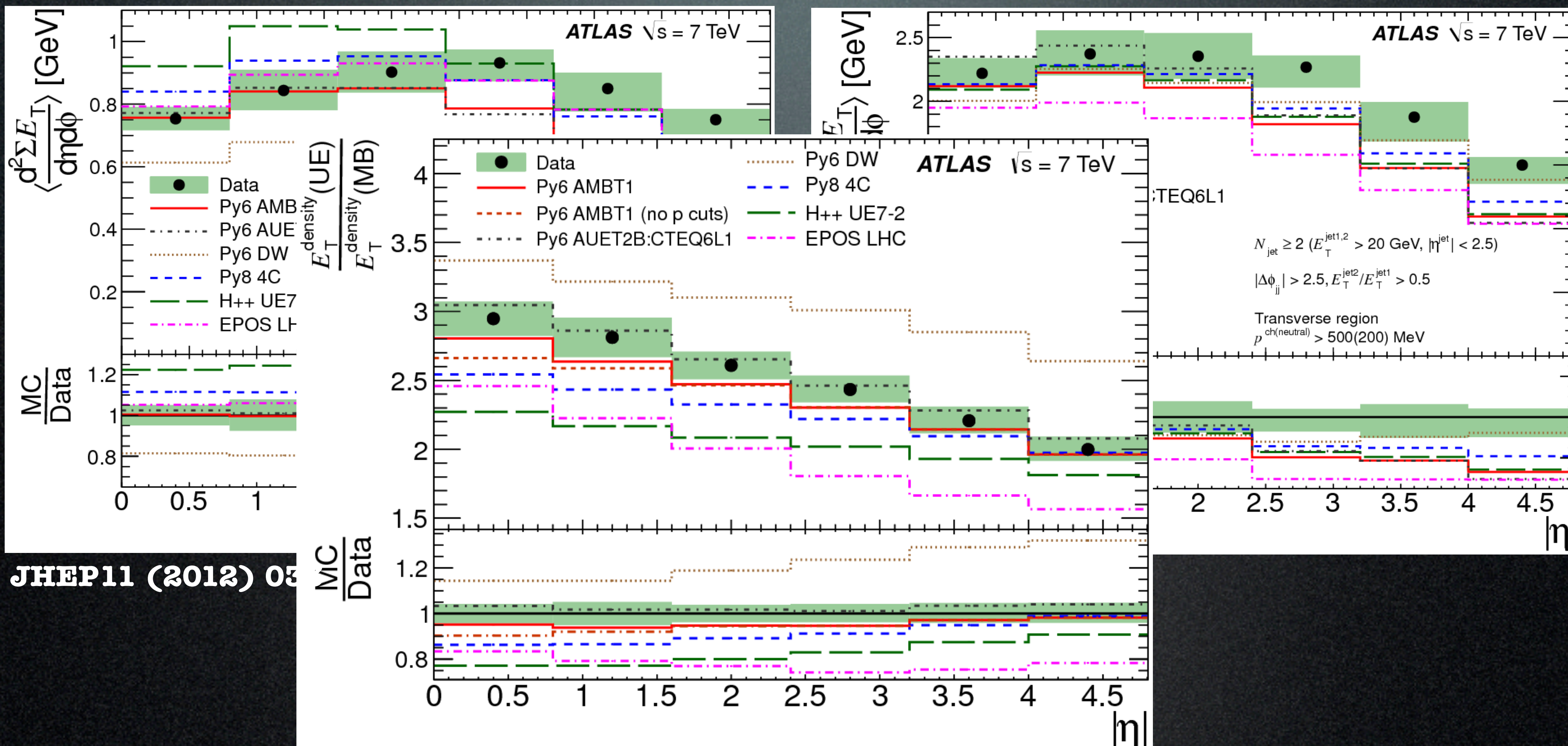




JHEP11 (2012) 033

Transverse energy flow: all models bad in forward region

UE-sensitive Observables

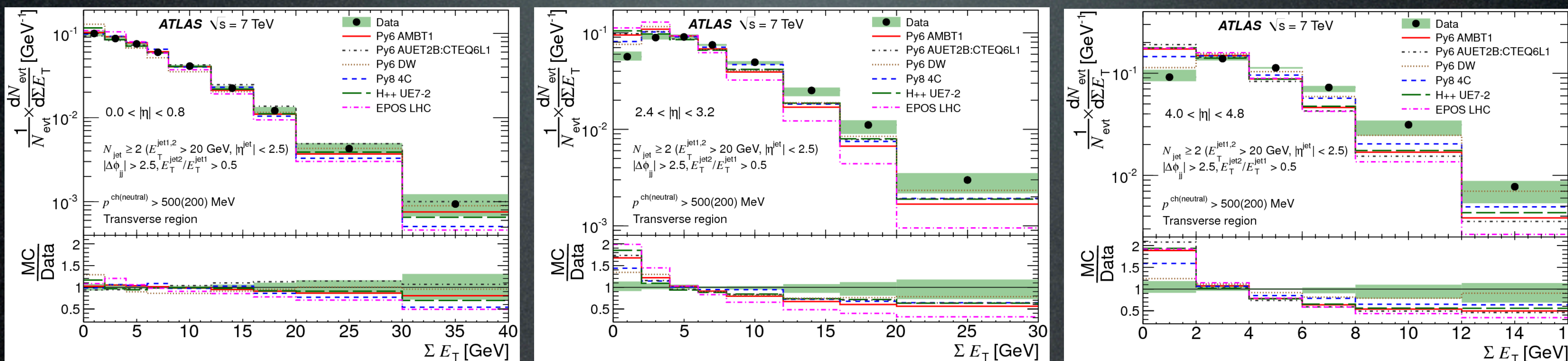


JHEP11 (2012) 03

More energy in dijet events!

From Central to Forward

low η → **high η**

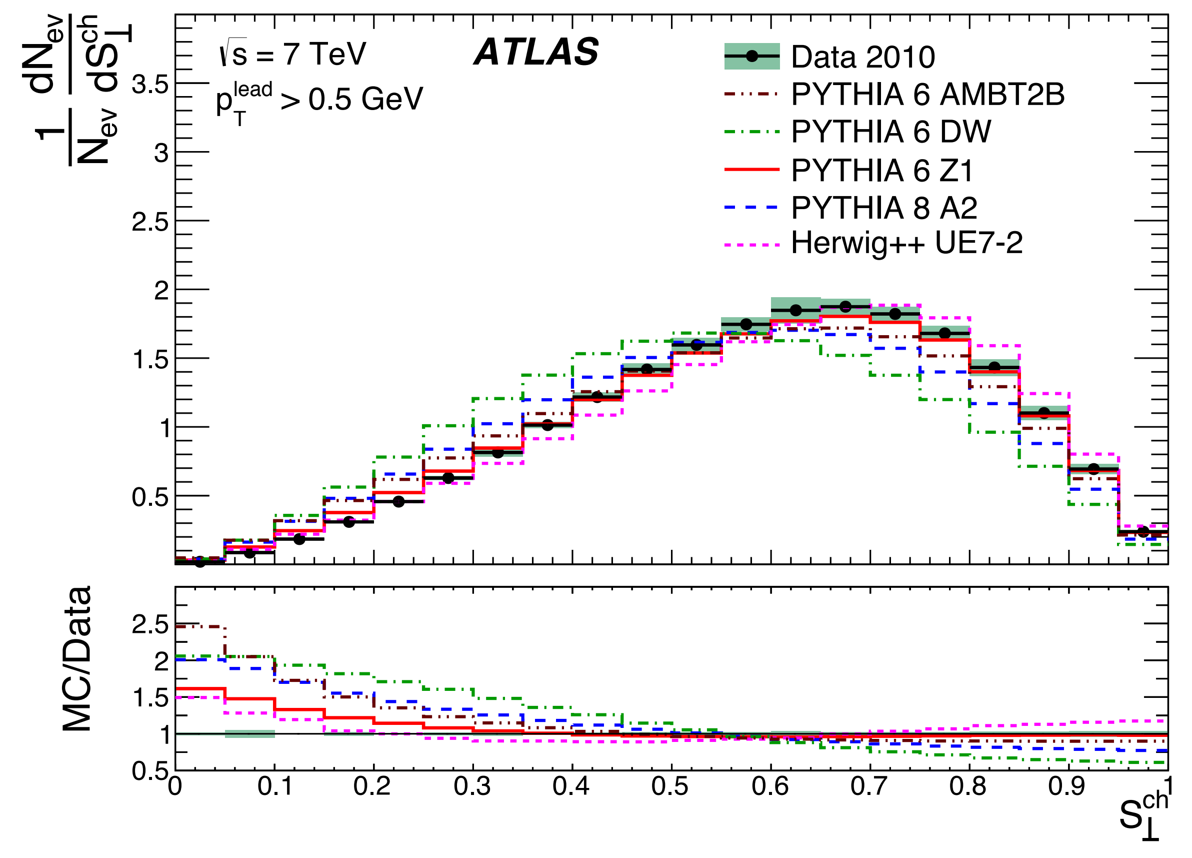


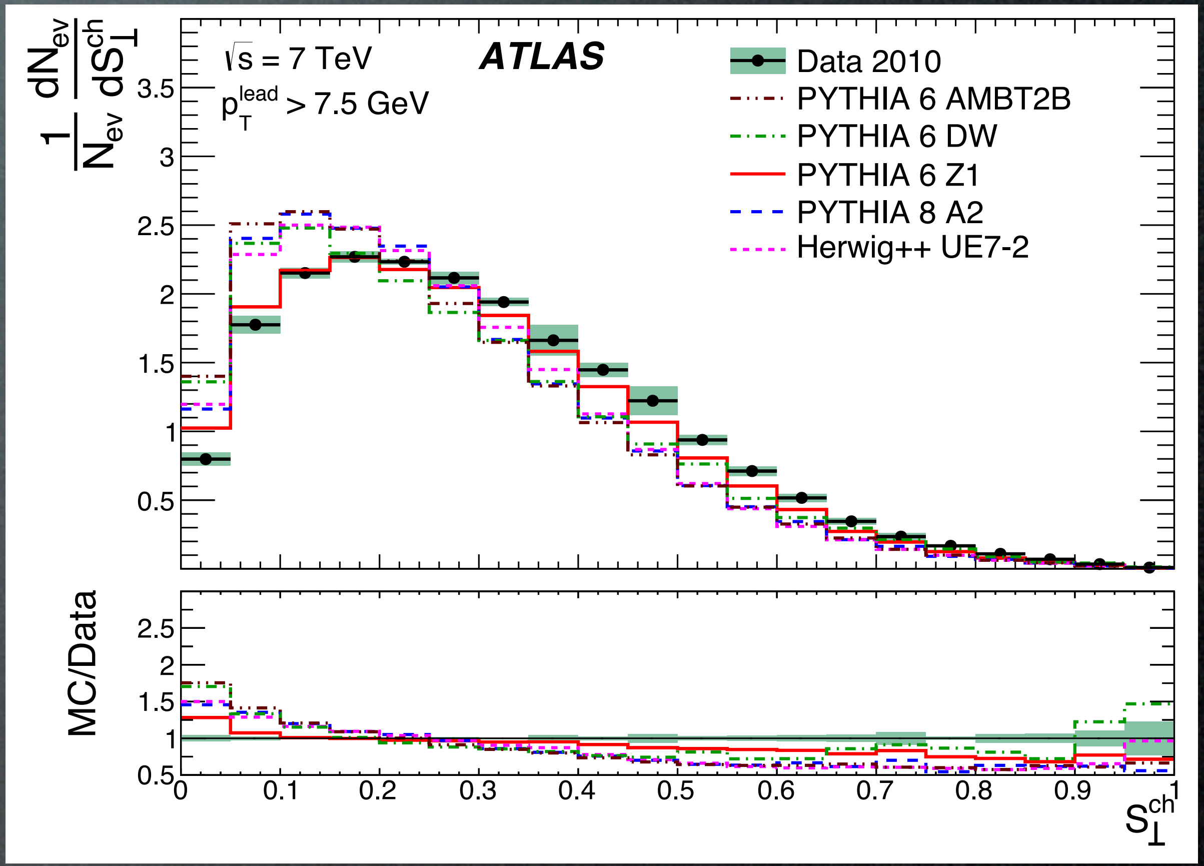
JHEP11 (2012) 033

UE tunes do better overall

Event Shapes

Low lead p_T → High lead p_T

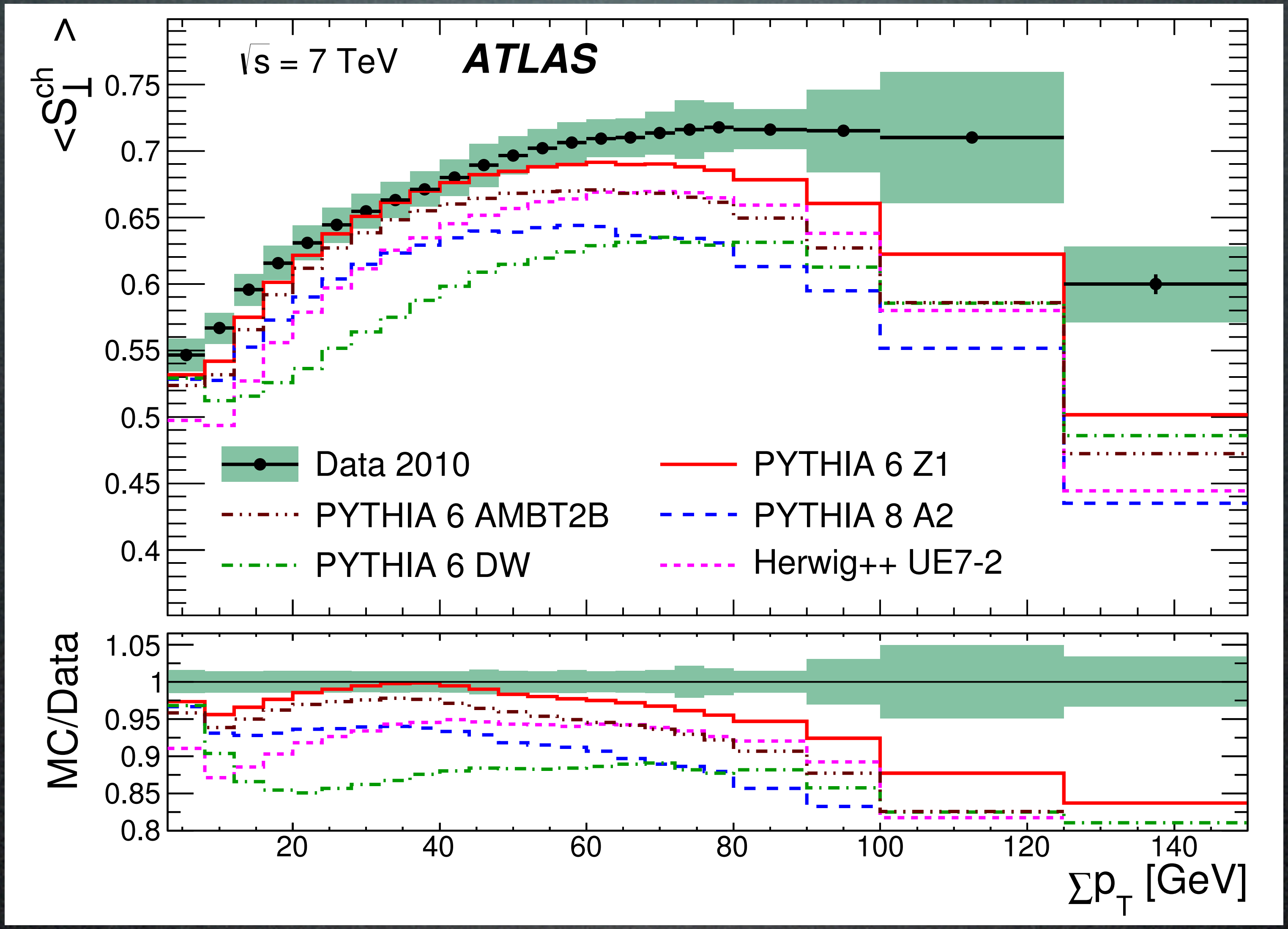




Phys. Rev. D 88, 032004 (2013)

UE starts taking over....

Event Shape Profile



Phys. Rev. D 88, 032004 (2013)

Emergence of jets?

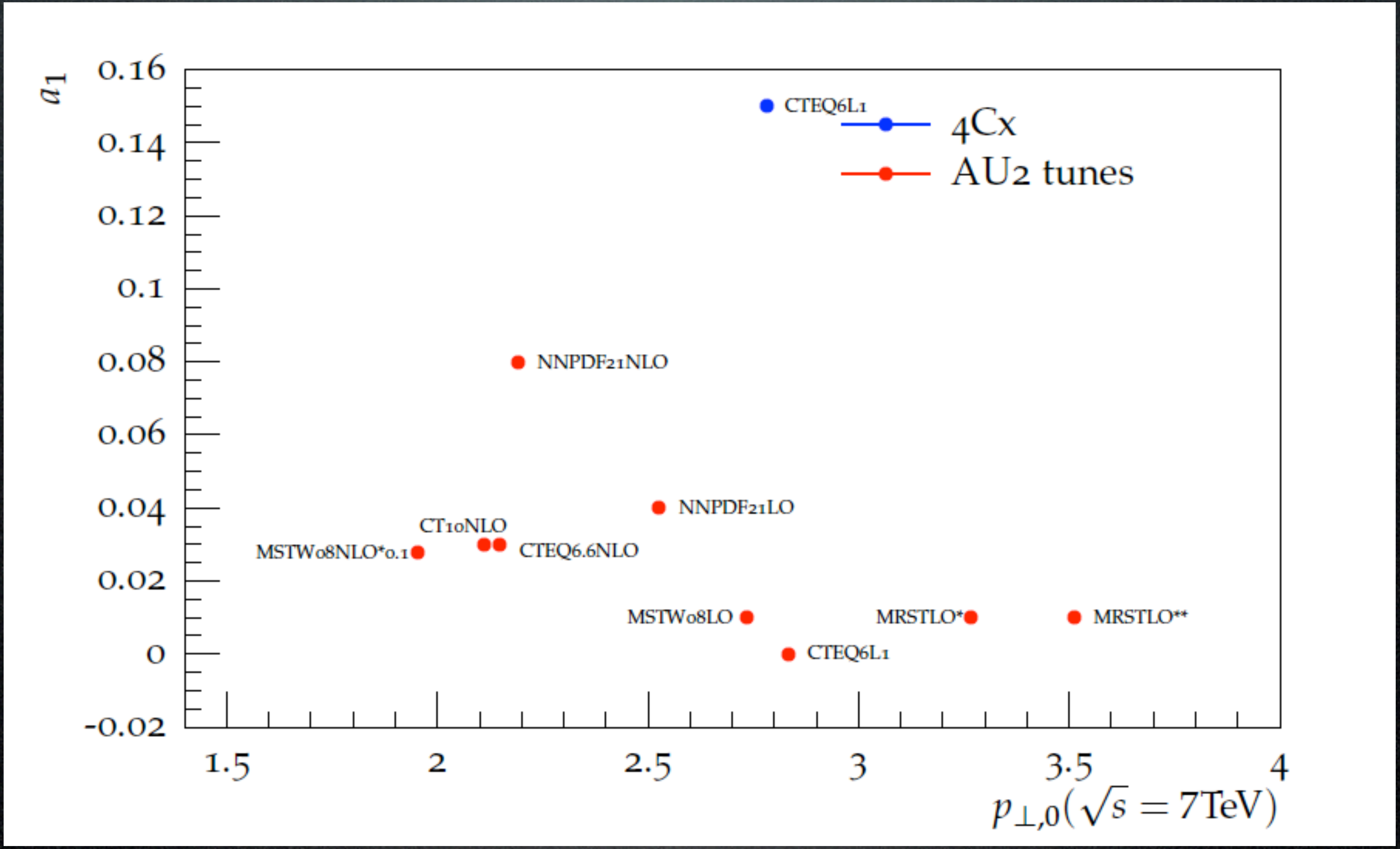
Back to Tuning

- All new data are being used in tuning
- Effects of PDF and matching important
- Problem with AUET2B tunes: tuned away the lack of dipole radiation in FSR.

PDF Dependence of Tunes

- Changing PDFs change gluon density, so re-tuning is needed.
- ATLAS systematically explored the effect of NLO and modified LO PDFs on the tunes.
- Many matrix element generators use NLO/mLO PDFs, so it is important to understand the effect on matched parton-shower generators.
- LO PDFs generally give the best description, with mLO ones the worst.
- NLO PDFs require less MPI cross-section screening and stronger color reconnection.

PDF Dependence



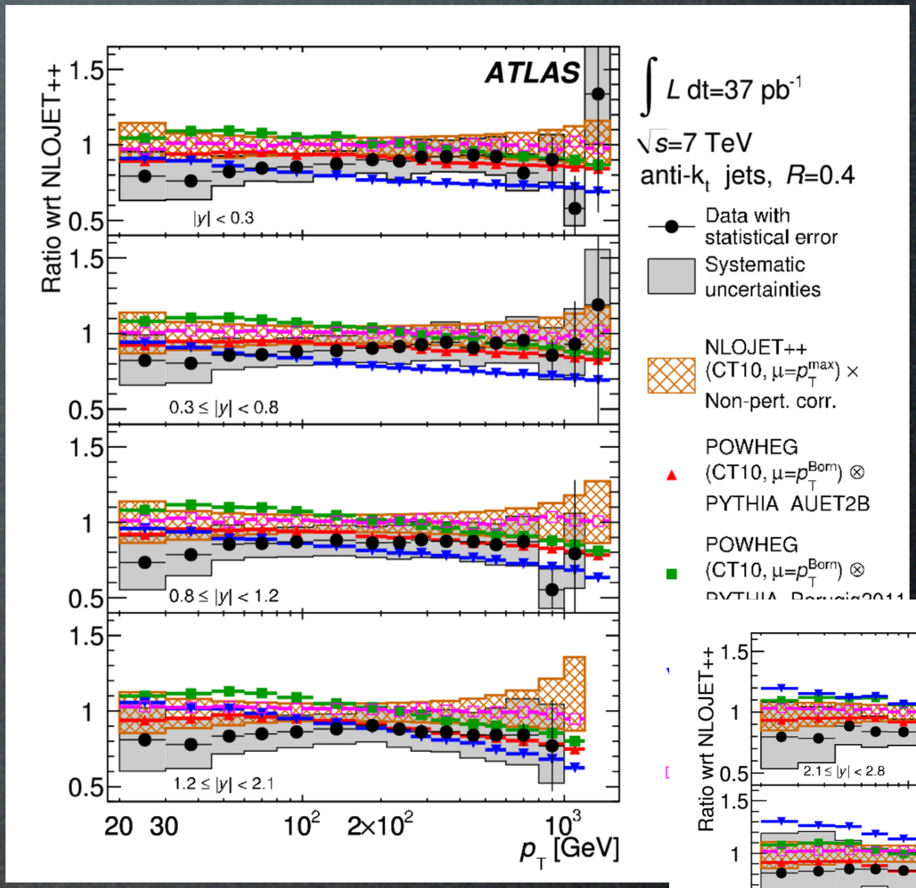
Matching

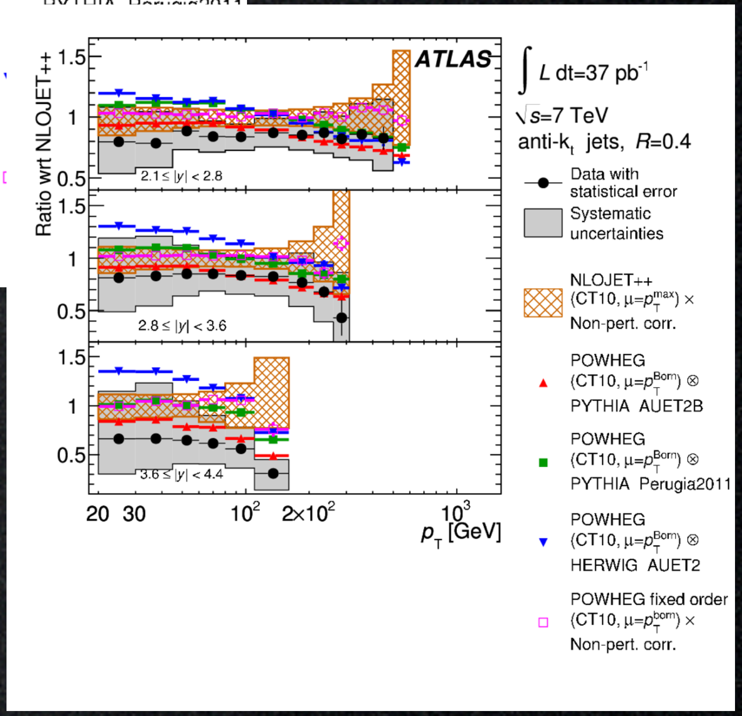
- The LO parton shower generators cannot predict the radiation of one or more hard jets (among other things), but do well in soft collinear regime.
- Use NLO matrix elements to improve description of the hardest jet.
- LO matrix elements with higher legs to improve description of many hard jets.
- Combine all these?

Powheg+Pythia8 Matching

(ongoing Les Houches Study)

- PoWHEG provides a scale (SCALUP) that is an indication of where the shower should take over from the perturbative calculation.
- What should be this scale?
- Imperfection in transition region



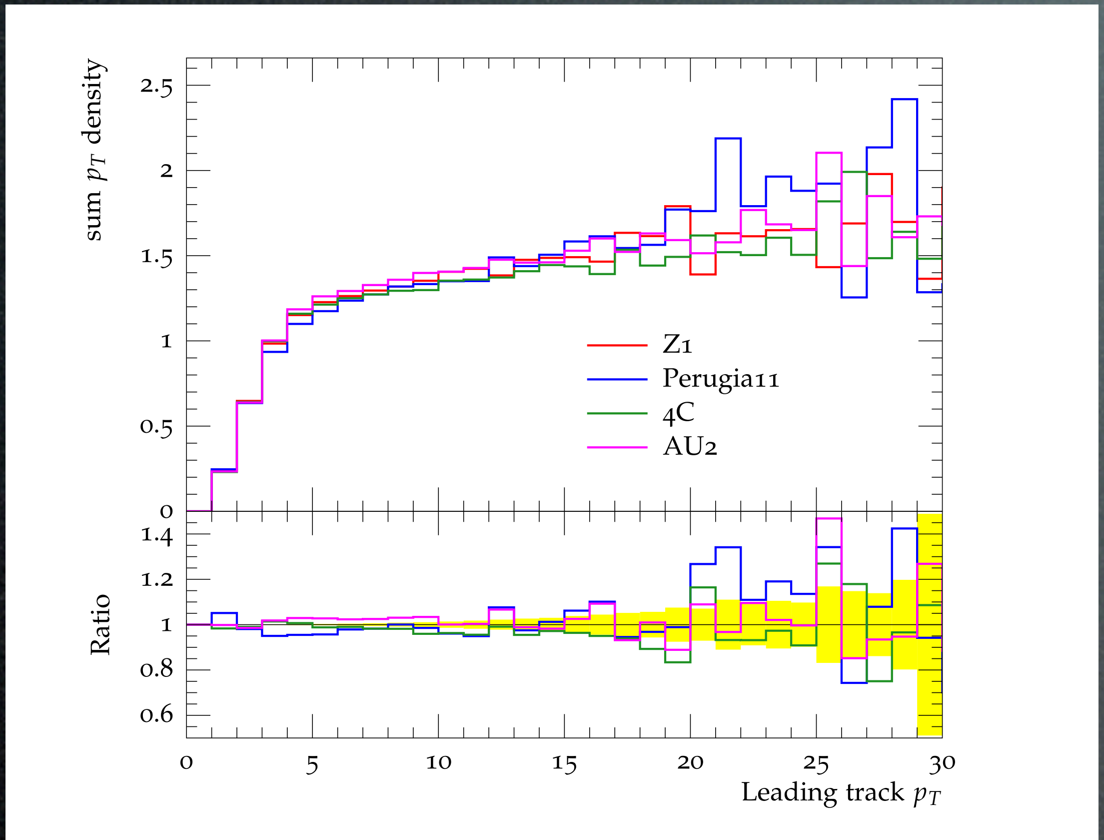


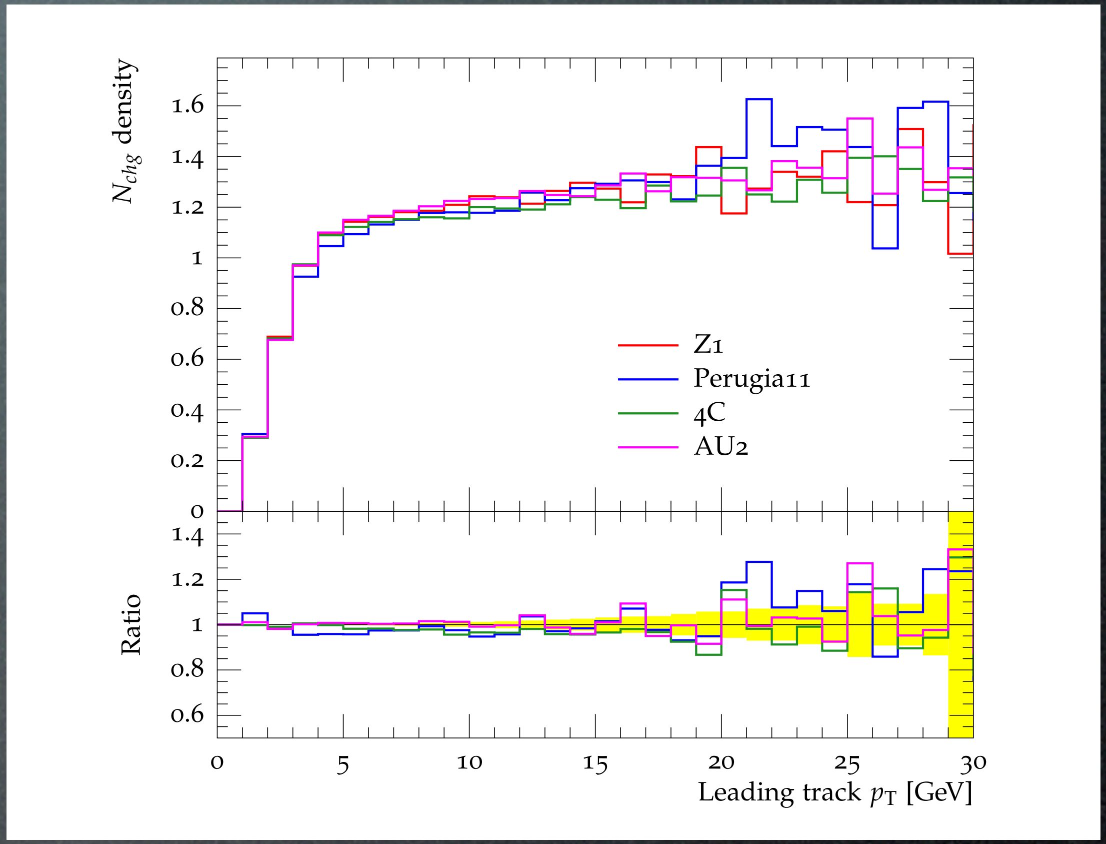
Summary

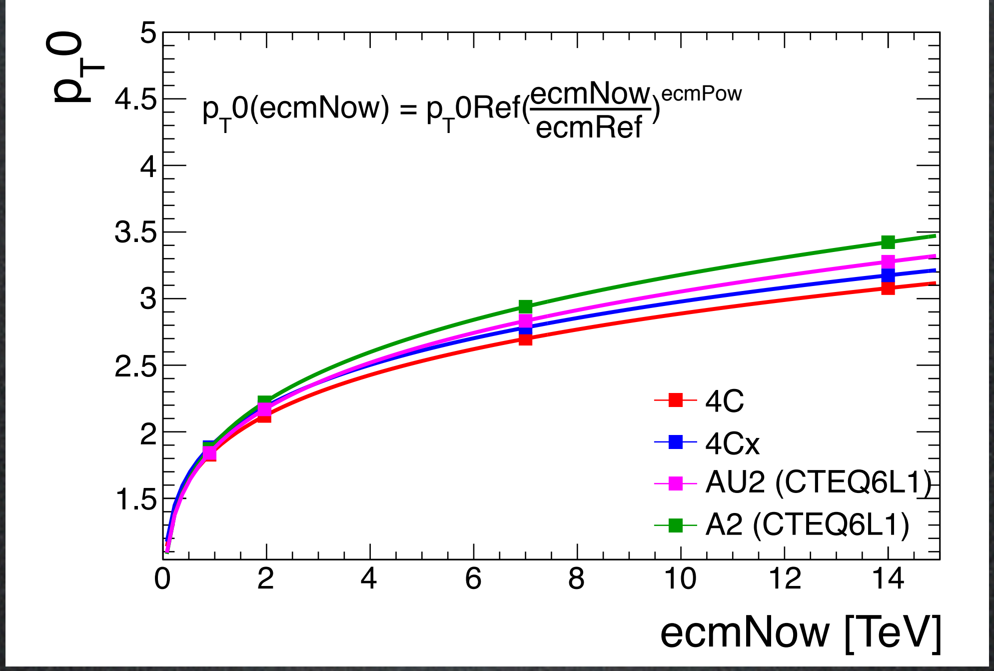
- Soft QCD is fun (and useful).
- Tuning is fun too, but hard to get everything right.
- Generators contain a lot under their hood, and it is good to have some understanding of it.
- The improved modelling of low p_T processes is feeded back to full event generation, where it affects high p_T part of the event, especially for precision measurements.
- Many analyses/data are available in Rivet/Hepdata, but experiments should try to have MC-independent final results, and make sure they are included in Rivet/Hepdata.

Epilogue

14 TeV UE Predictions







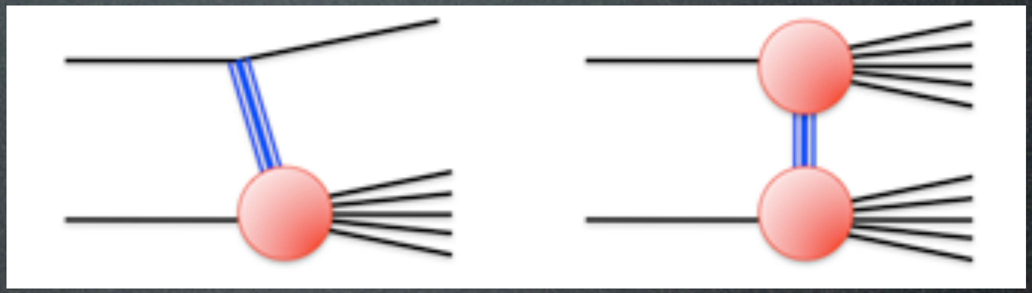
Supporting Material

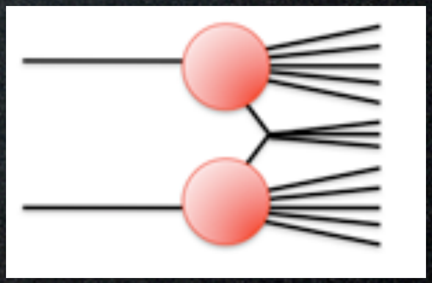
Soft-QCD

$$\sigma_{total} = \sigma_{el} + \sigma_{inel}$$

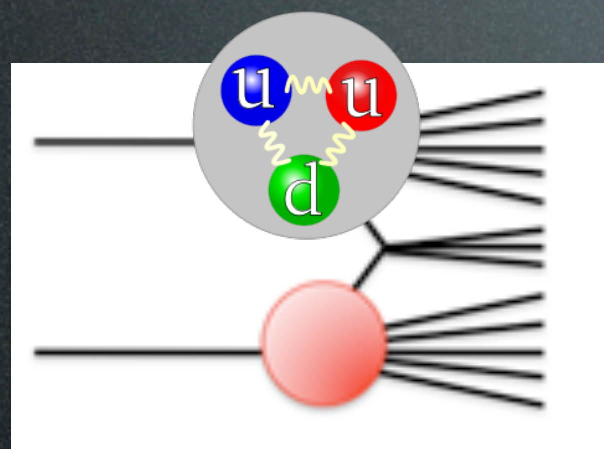
Soft-QCD

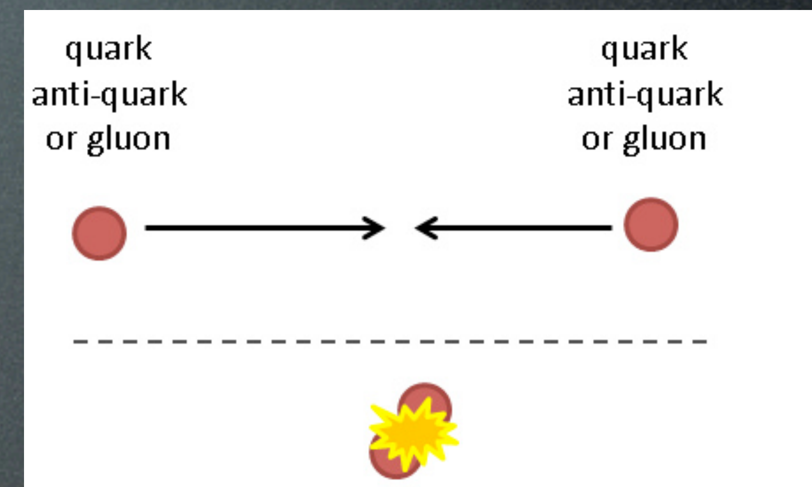
$$\sigma_{total} = \sigma_{el}+\sigma_{sd}+\sigma_{dd}+\sigma_{nd}$$





Soft-QCD

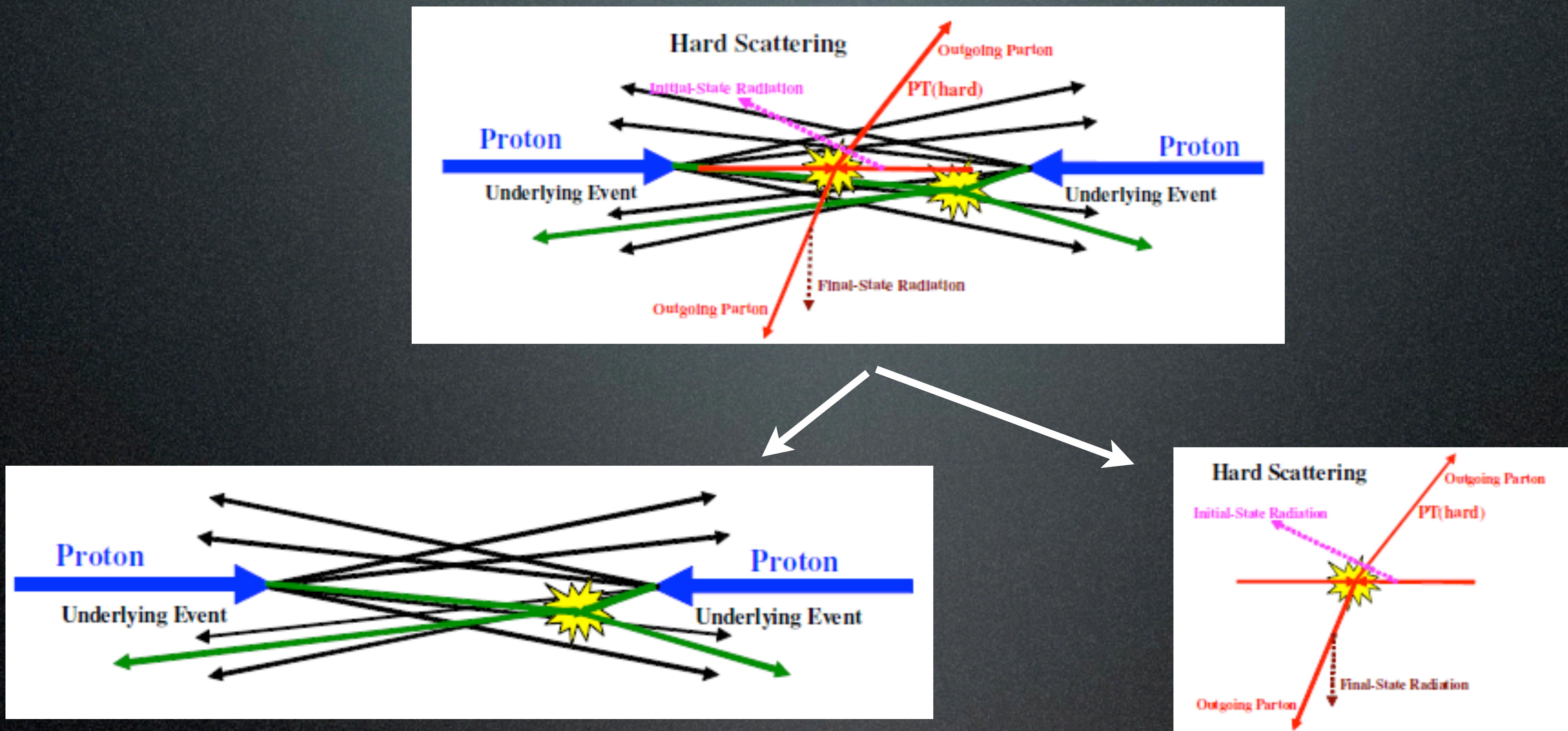




Interesting part!

No hard scatter

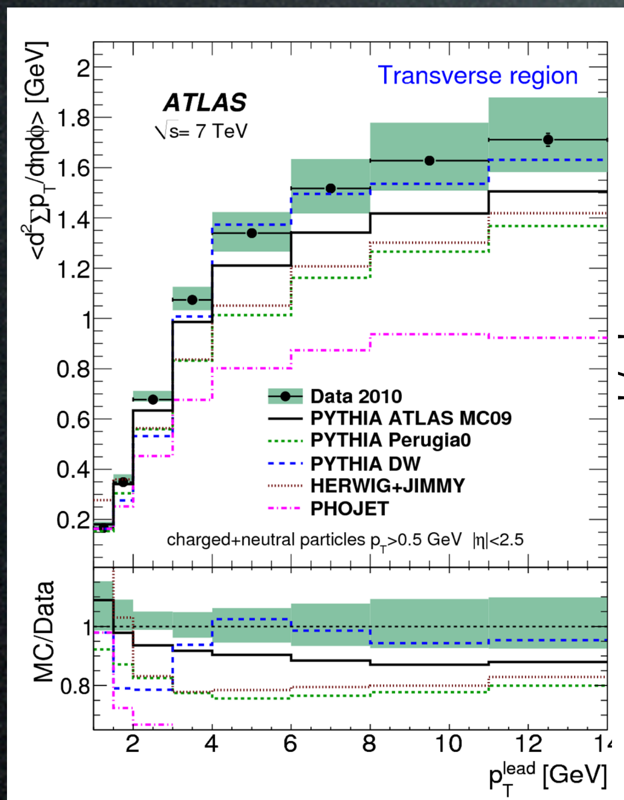
Soft-QCD

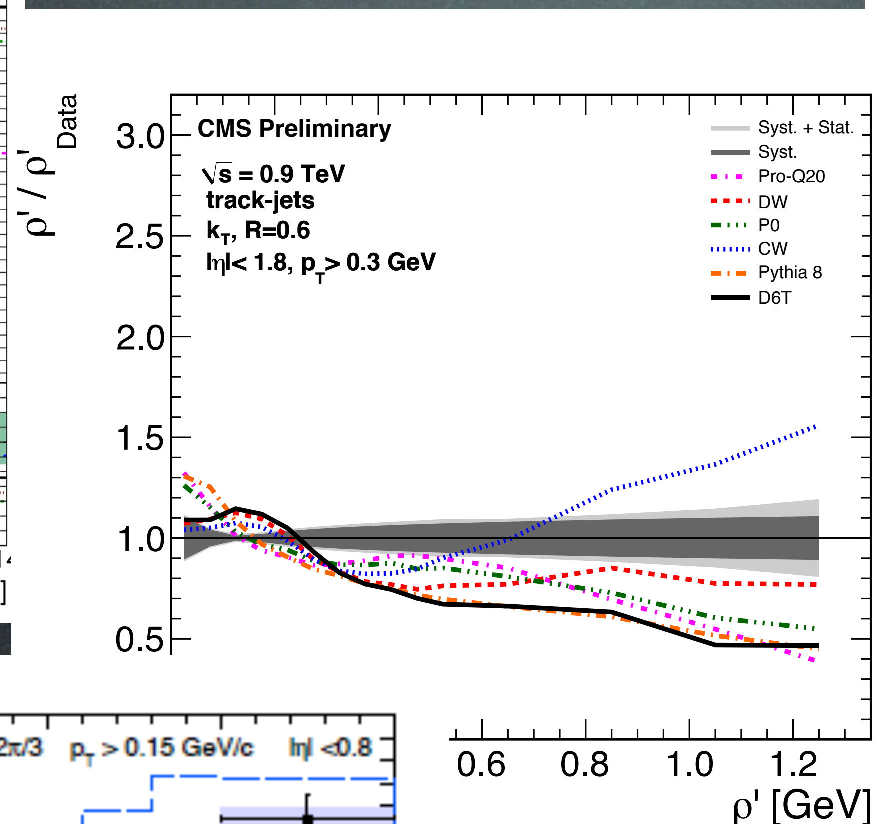


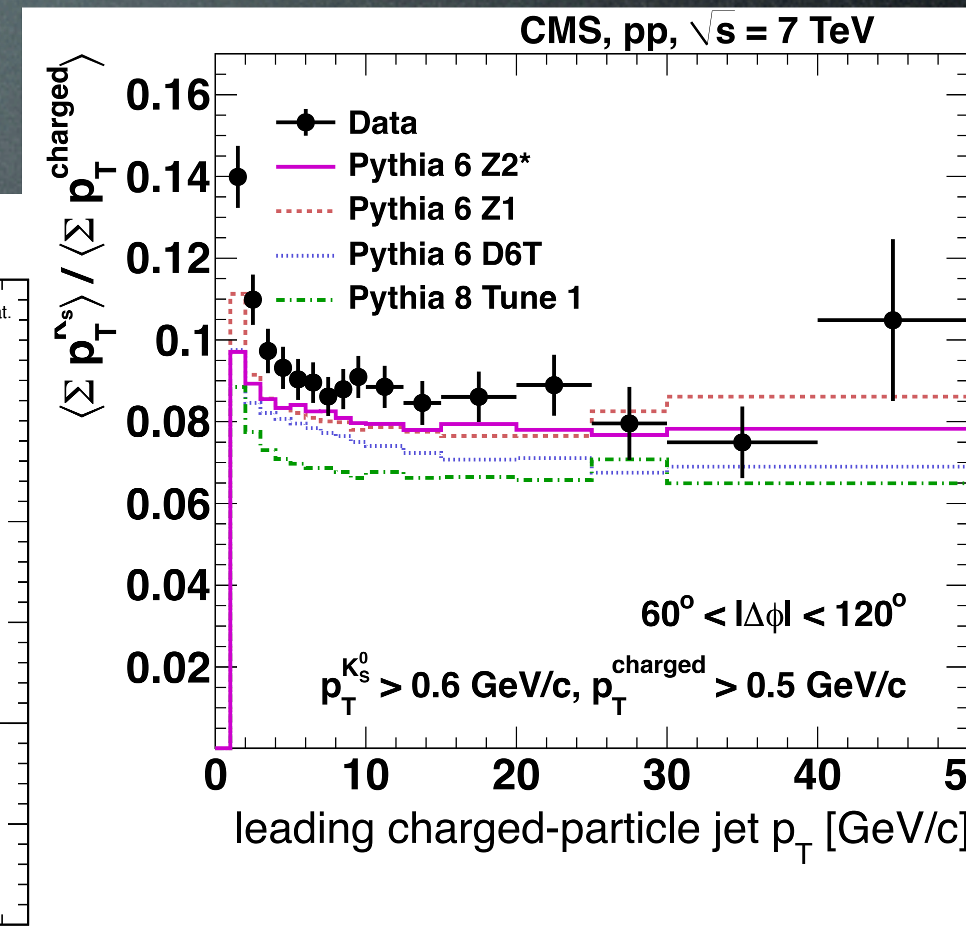
Underlying event = BBR+ MPI+ (ISR+FSR)

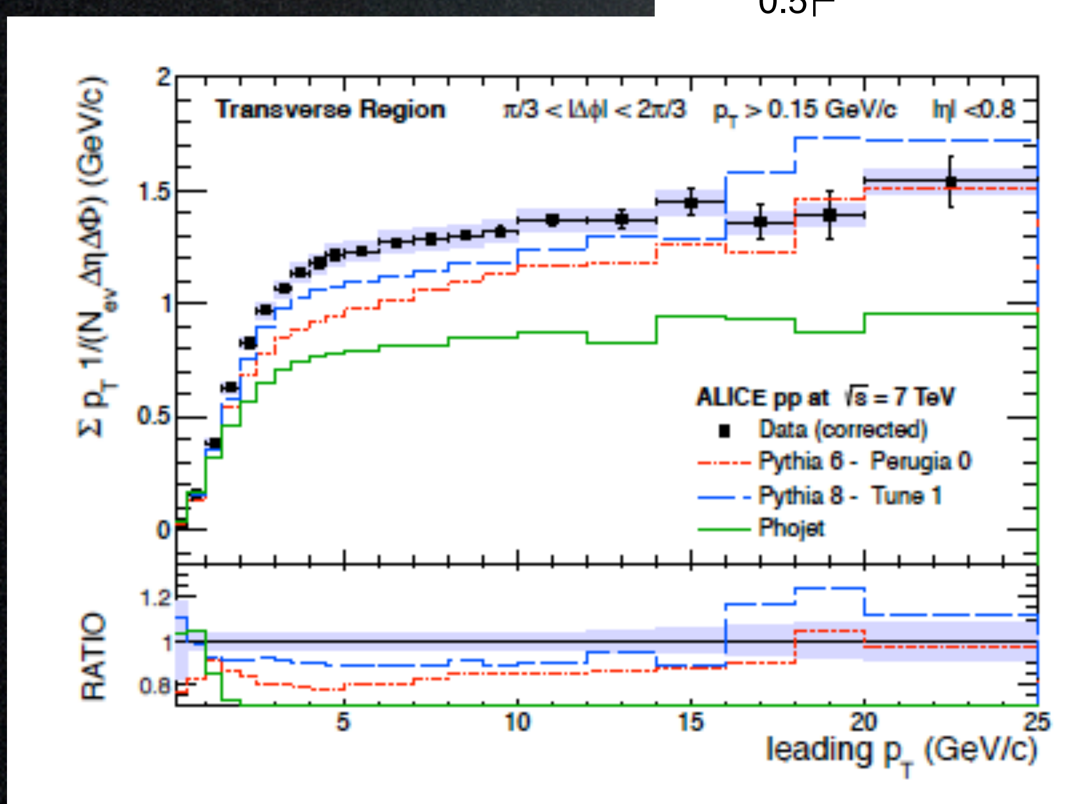
BBR: Beam-beam remnants
MPI: Multiple Parton interactions
ISR/FSR: Initial/Final state radiation

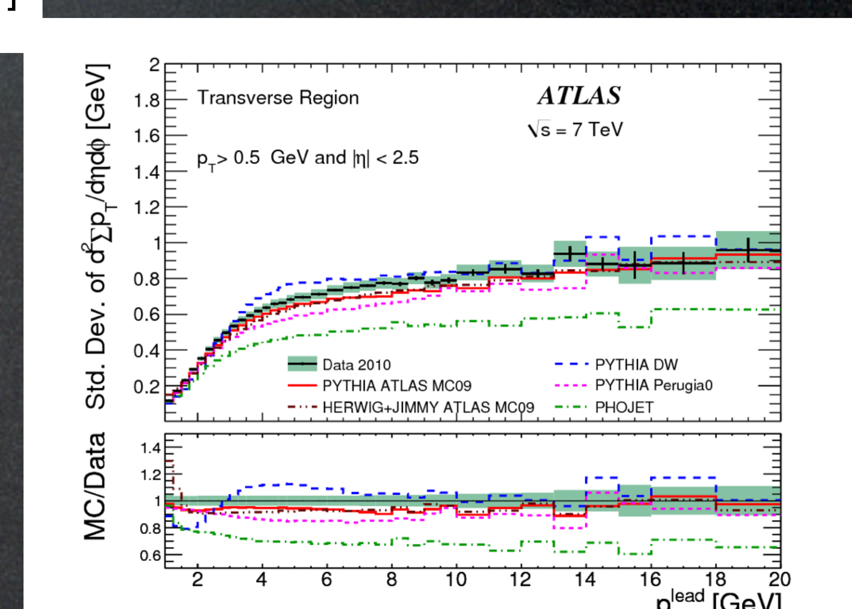
Many LHC UE Analyses











Distribution of observable: O

In production of X + anything

Fixed Order (all orders)

$$\frac{d\sigma}{d\mathcal{O}}\Big|_{ME} = \sum_{k=0} \int d\Phi_{X+k} \left| \sum_{\ell=0} M^{(\ell)}_{X+k} \right|^2 \delta(\mathcal{O} - \mathcal{O}(\{p\}_{X+k}))$$

Sum over identical amplitudes, then square

Momentum configuration

Phase Space

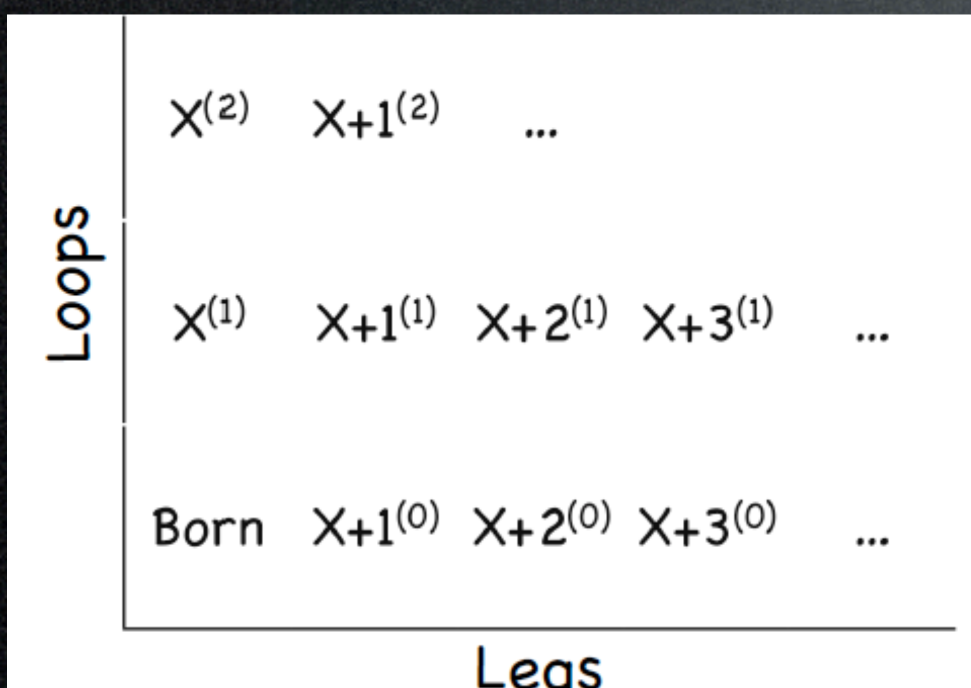
Cross Section differentially in O

Sum over "anything" ≈ legs

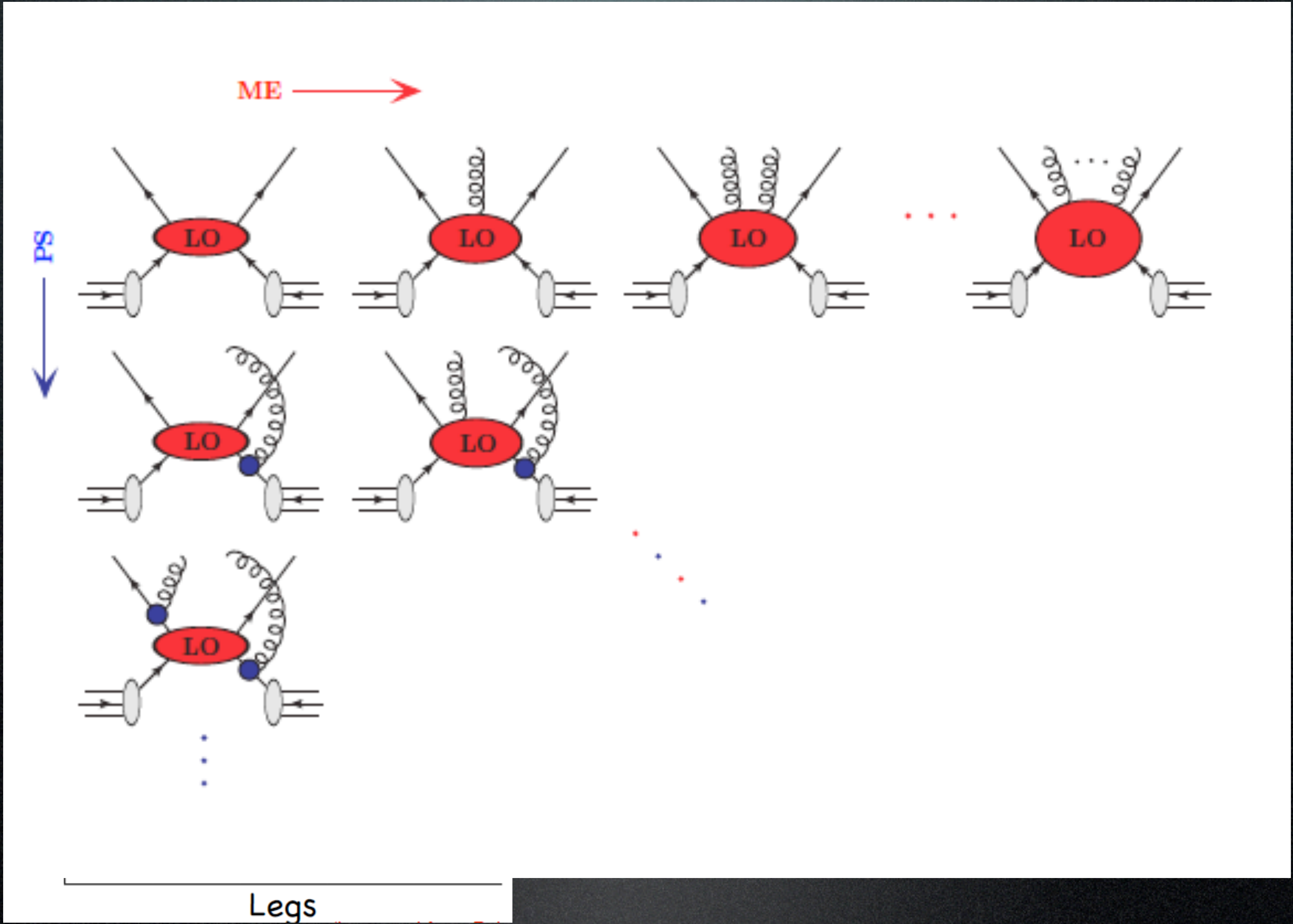
Matrix Elements for X+k at (l) loops

Evaluate observable → differential in O

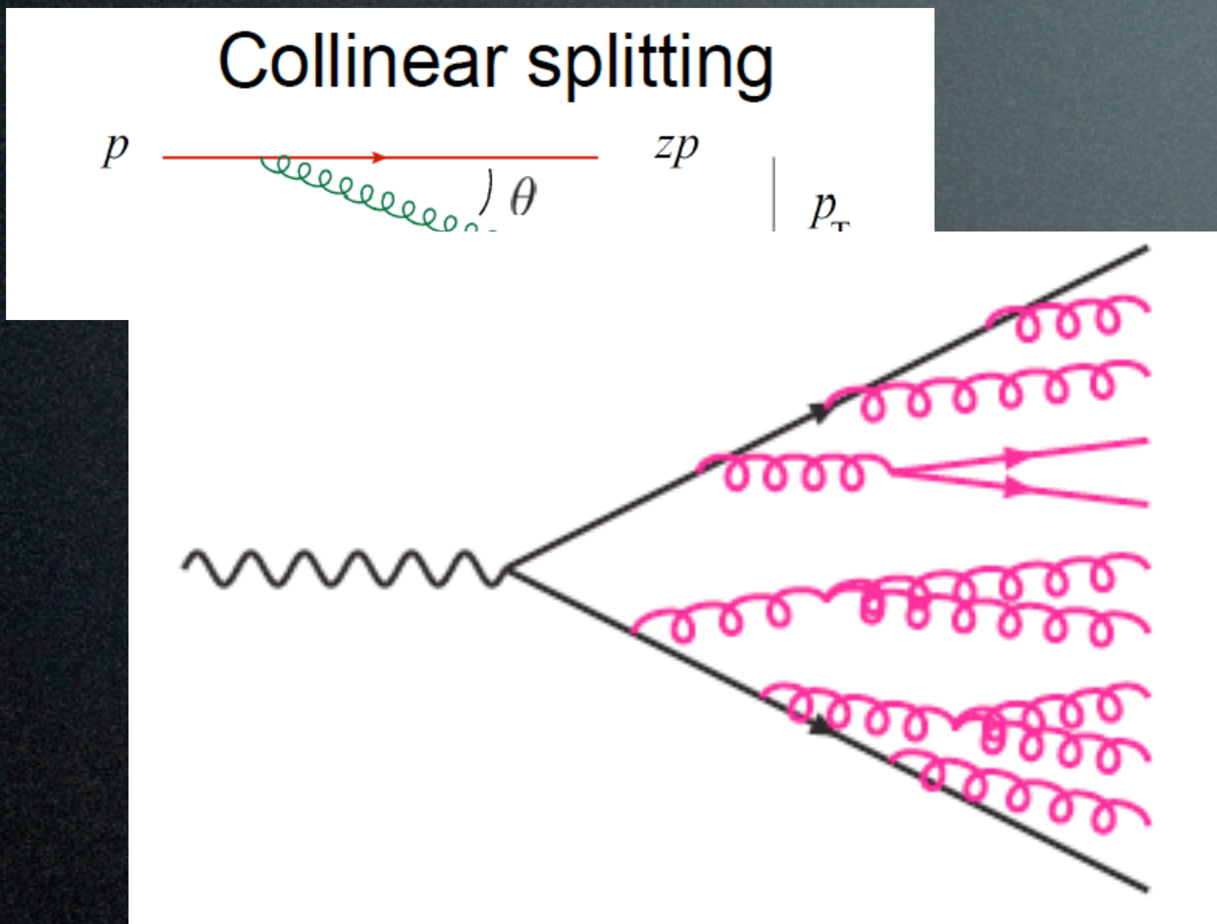
Leading Order: truncate at l=0.



Leading order for X+n: truncate at l=0, k=n. (Lowest order at which X+n happens)


ME
PS
LO
LO
LO
LO
LO
LO
LO
Legs

Parton Shower



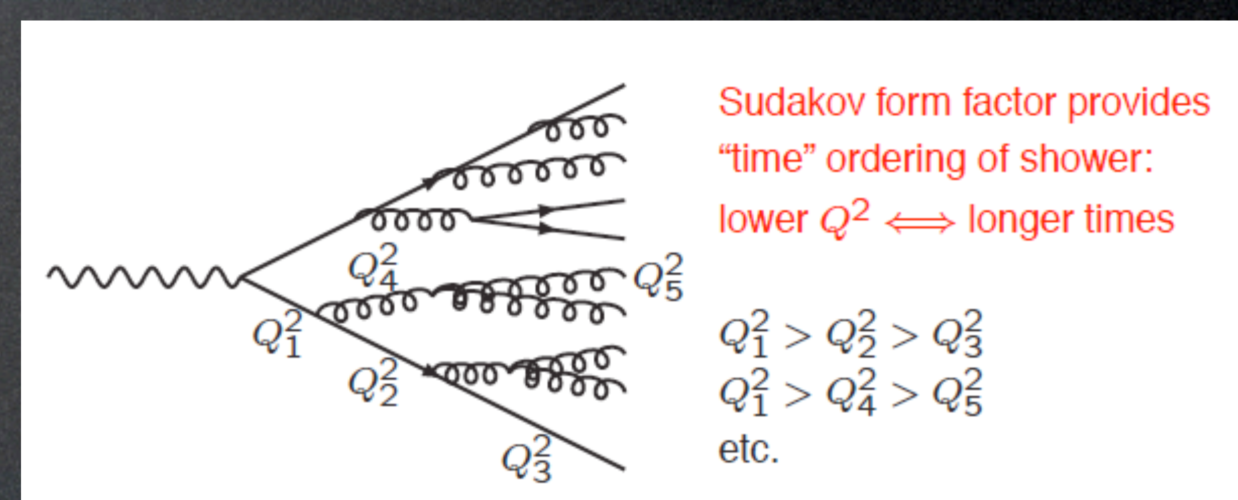
Q=energy scale

Each branching governed by DGLAP equation

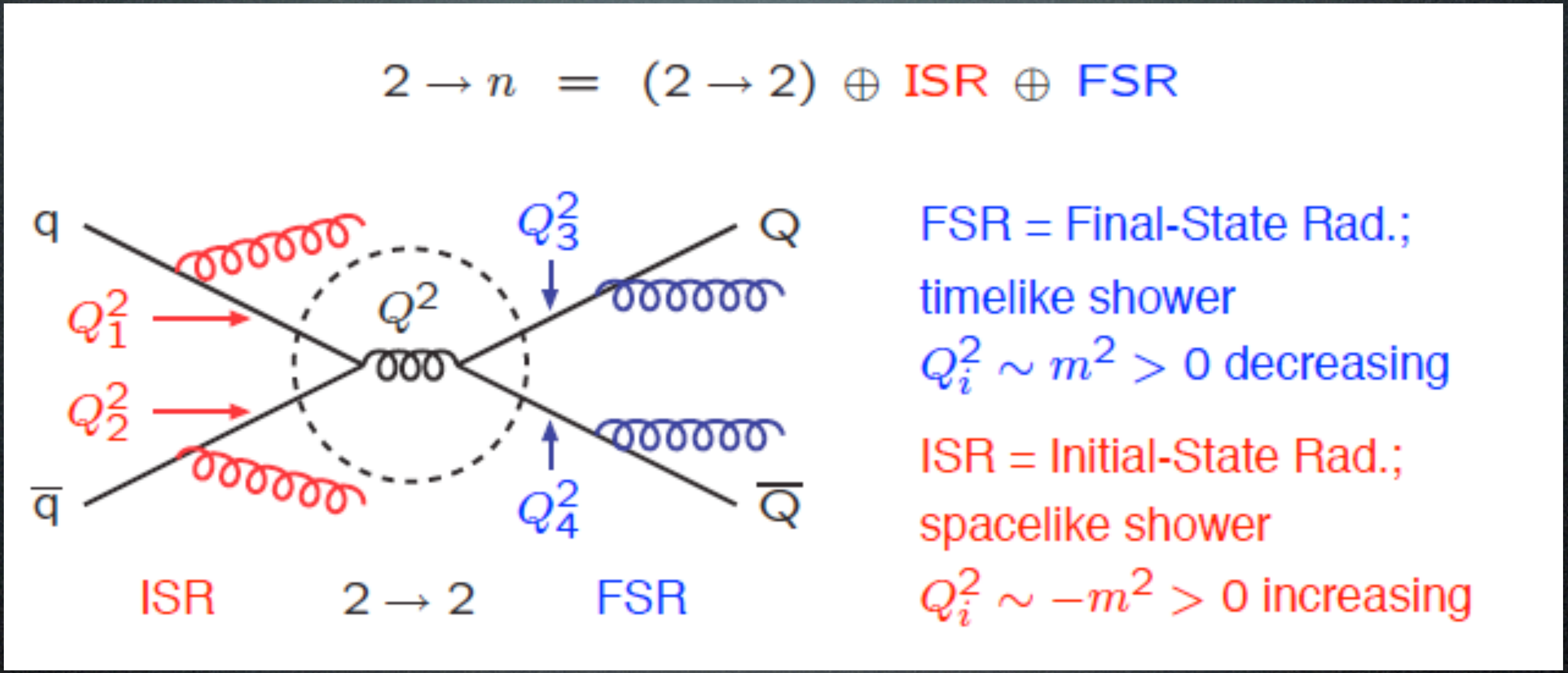
$$d\mathcal{P}_{a\to bc} = \frac{\alpha_s}{2\pi}\frac{dQ^2}{Q^2} P_{a\to bc}(z)\, dz$$

Braching continues until each parton finally undergoes transitions to hadrons that can be observed.

Probability that a branching happens at a given time is given by Sudakov Form Factor



Time/Space-like Showers

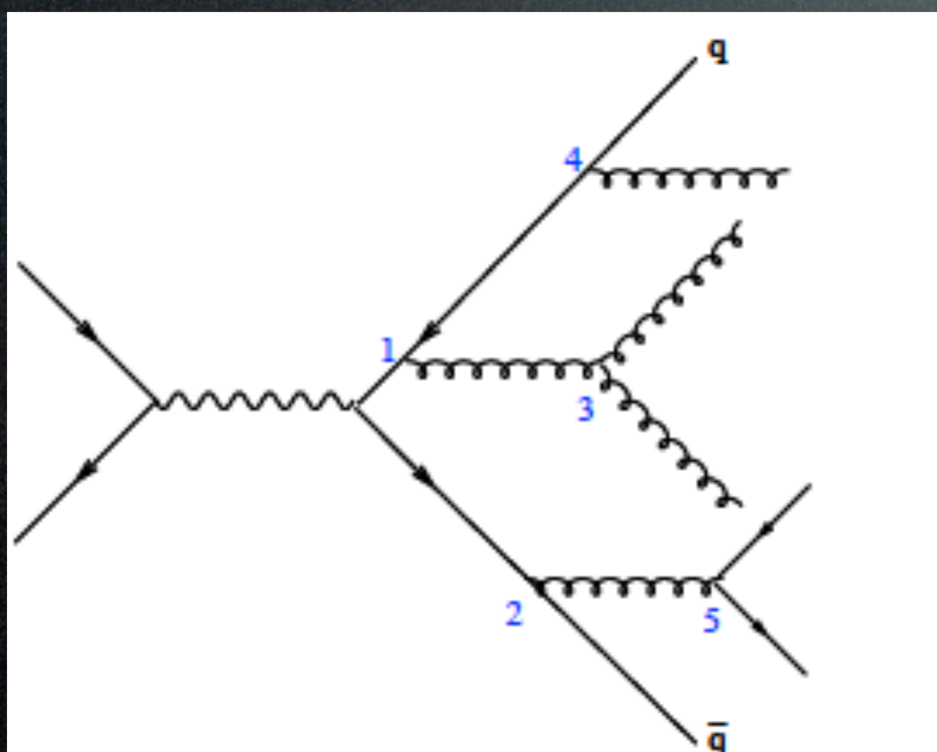


ISR: incoming partons to the hard process initiate a parton emission cascade (or shower) where in each branching one parton becomes increasingly off-shell with a space-like virtuality.

FSR: outgoing partons, including the non-colliding partons emitted from the initial state, have timelike virtualities that decrease in the cascade down to on-shell partons.

Shower Ordering

Tree-level matrix element for an n-parton state is approximated by a product of splitting functions corresponding to a sequence of one-parton emissions from the zeroth order state.



Order the emissions according to some resolution scale:

HERWIG: $Q^2 \approx E^2(1-\cos\theta) \approx E^2\theta^2/2$
PYTHIA: $Q^2 = m^2$ (timelike) or $= -m^2$ (spacelike)

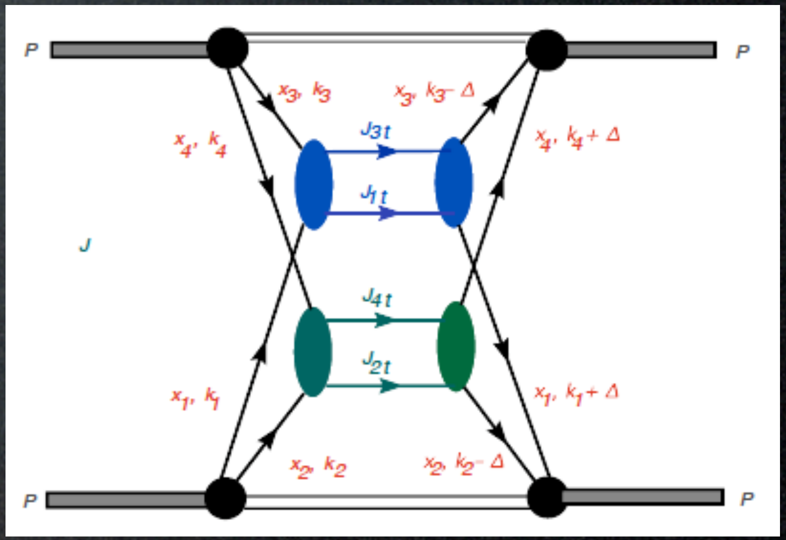
Pythia: Large mass first/Large p_T first.
Herwig: Large angle first.

Multiple Parton Interactions

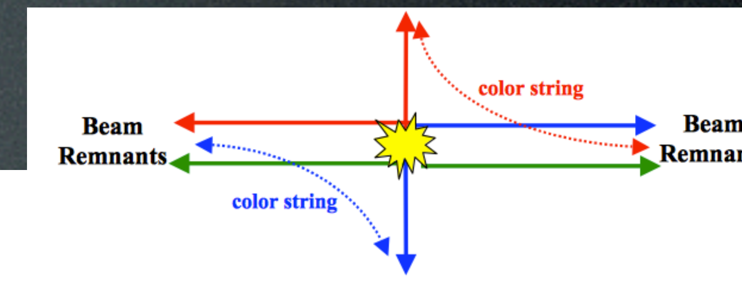
- Cross-section for 2->2 interactions is dominated by t channel gluon exchange. Diverges like $1/(p_T)^4$ as $(p_T)\text{->}0$. Regularize by a smooth cutoff, $(p_T)\text{->}(p_T{}^2+p_{T0}{}^2)^2$.

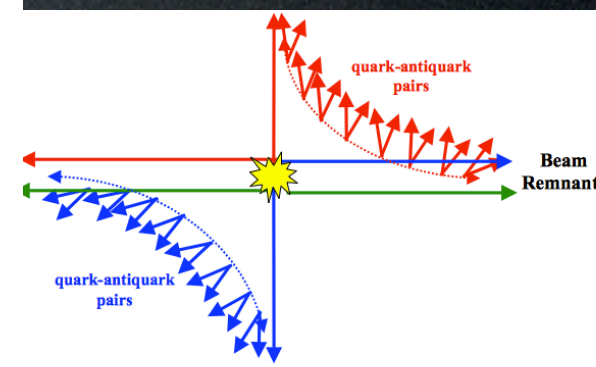
- If average number of interactions $<n>(P_T=P_{T0})=\sigma^{integrated}(p_{T0})/\sigma^{total}$, then reconciled.

- Protons are composite objects, physically corresponds to several parton pairs interacting.

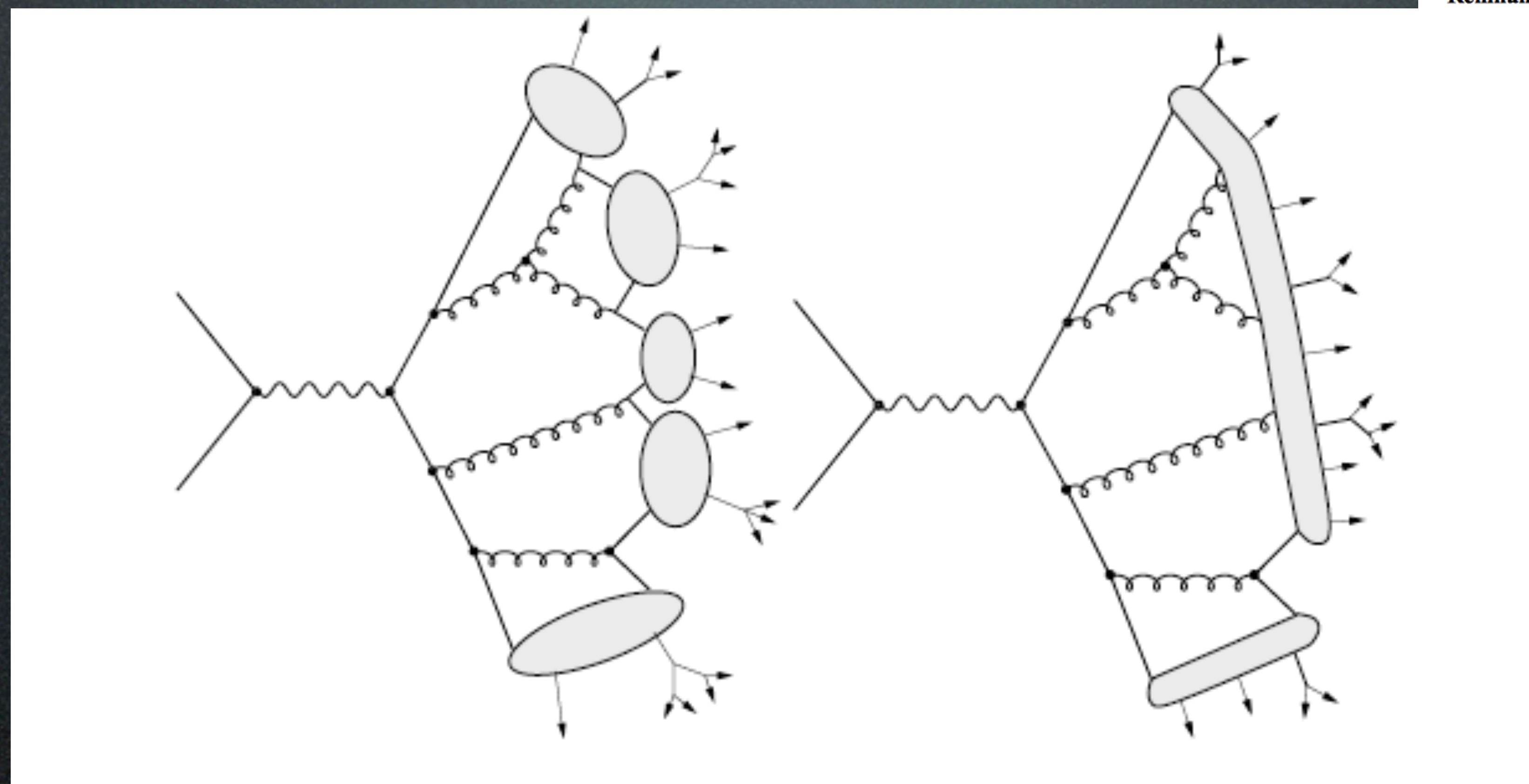


Hadronization





Herwig family



Pythia family

Cluster: neighbouring pairs form color neutral clusters which (usually) decay into two hadrons.

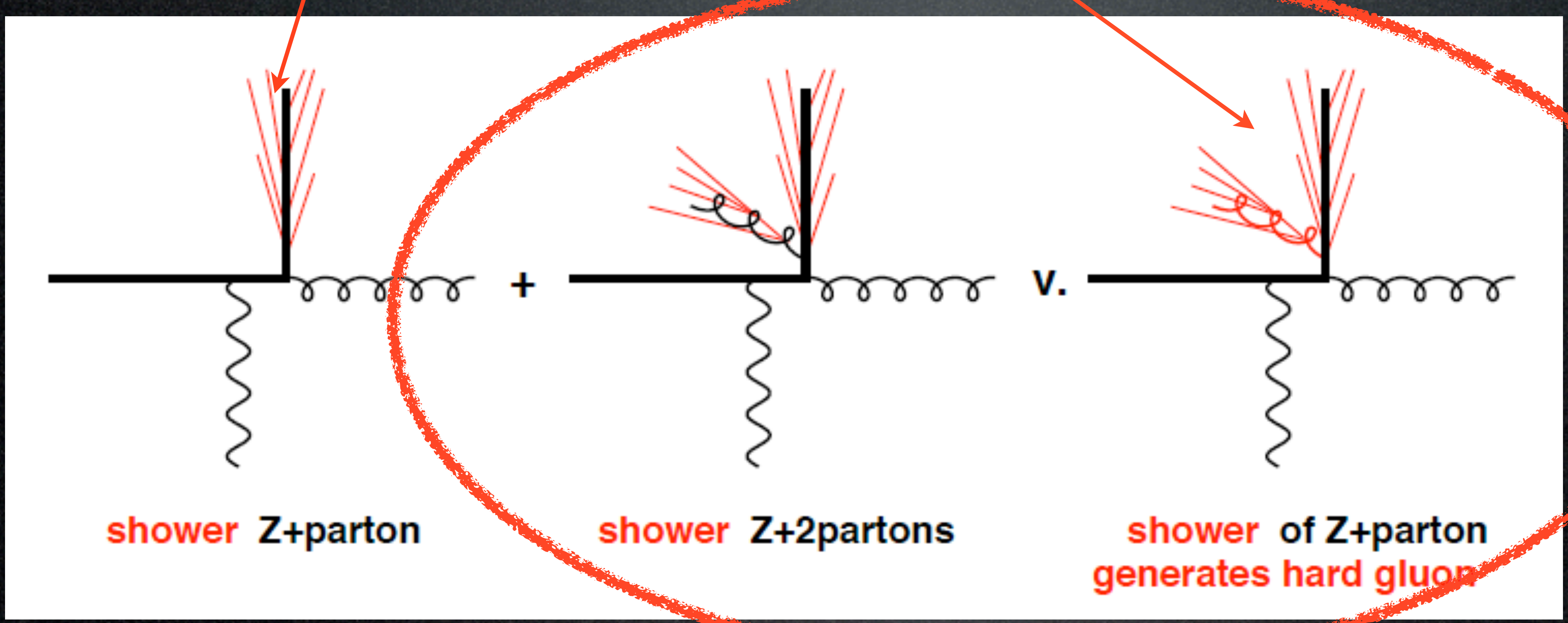
String: partons connected by massless relativistic string, breaks up as they move apart.

Matching Issues

For a Z+jets event:

PS: Z+1 jet + shower

Multileg: Z+ 2 jets; then shower from each leg



double counting!

My very biased/incomplete guide on what (pythia) tunes to use*:

- Usual high p_T LHC processes: AUET2B with LO PDF. Compare to DW (old shower ordering, old UE model), Perugia 2011 (identified particle production), Z1 (older PDF).
- To avoid: DWT/D6T (wrong energy extrapolation), Z2 (no energy scaling with PDF change), AUET2 (incorrect alpha_S propagation), any MB tune, MC08, CSC, MC09...
- Or use AU2 (Pythia8), 4C.
- For soft processes, use AMBT2B or (Pythia8) A2(M).

* subject to change anytime. No guarantee assumed if your analysis fails ;-)