Showering at the LHC

Deepak Kar
University of Glasgow/Witwatersrand

UCL HEP Seminar
17th March, 2015

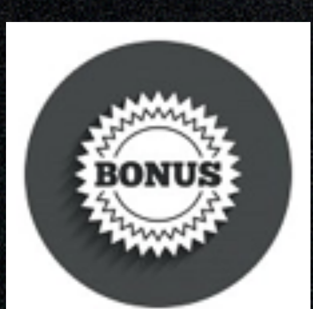
Outline

Tuning

Phenomenological Models

Compare with model predictions

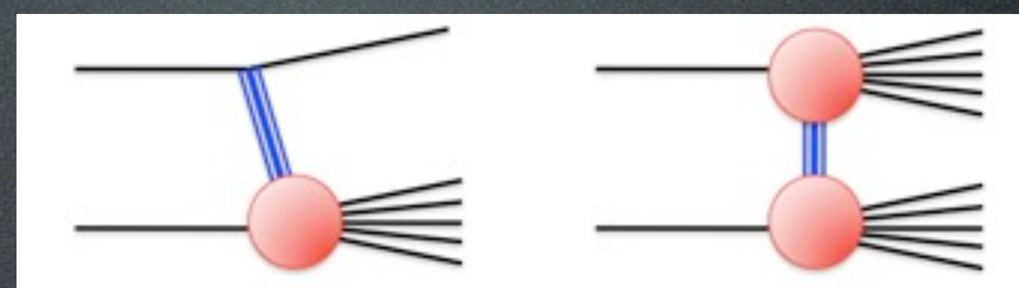
Measure distributions sensitive to softQCD

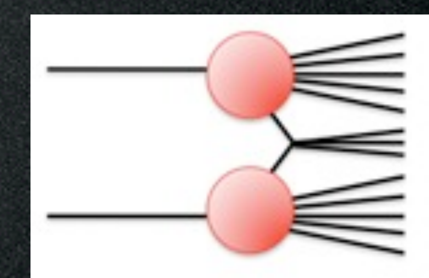


Application in jet substructure/top tagging

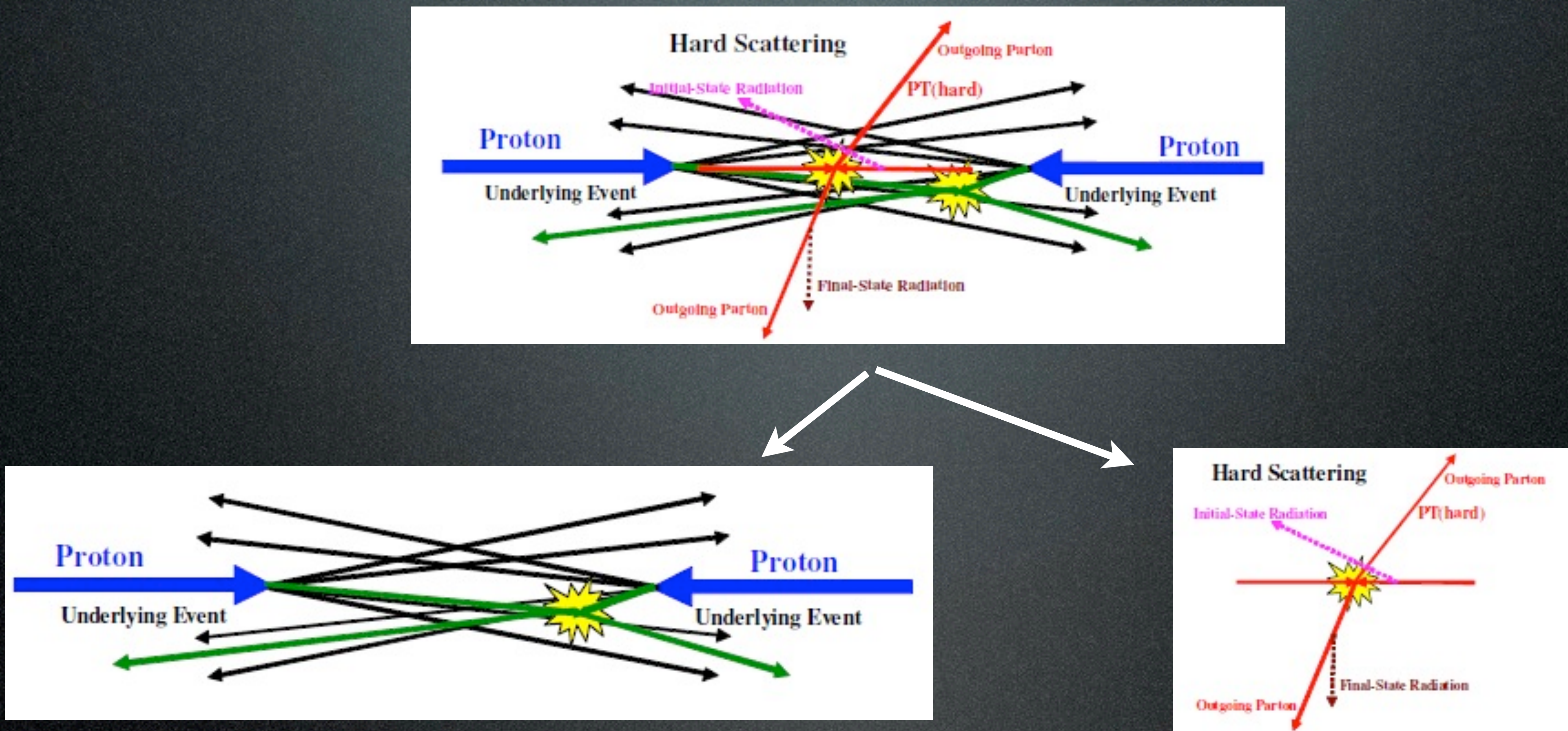
Soft-QCD

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





Soft-QCD



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

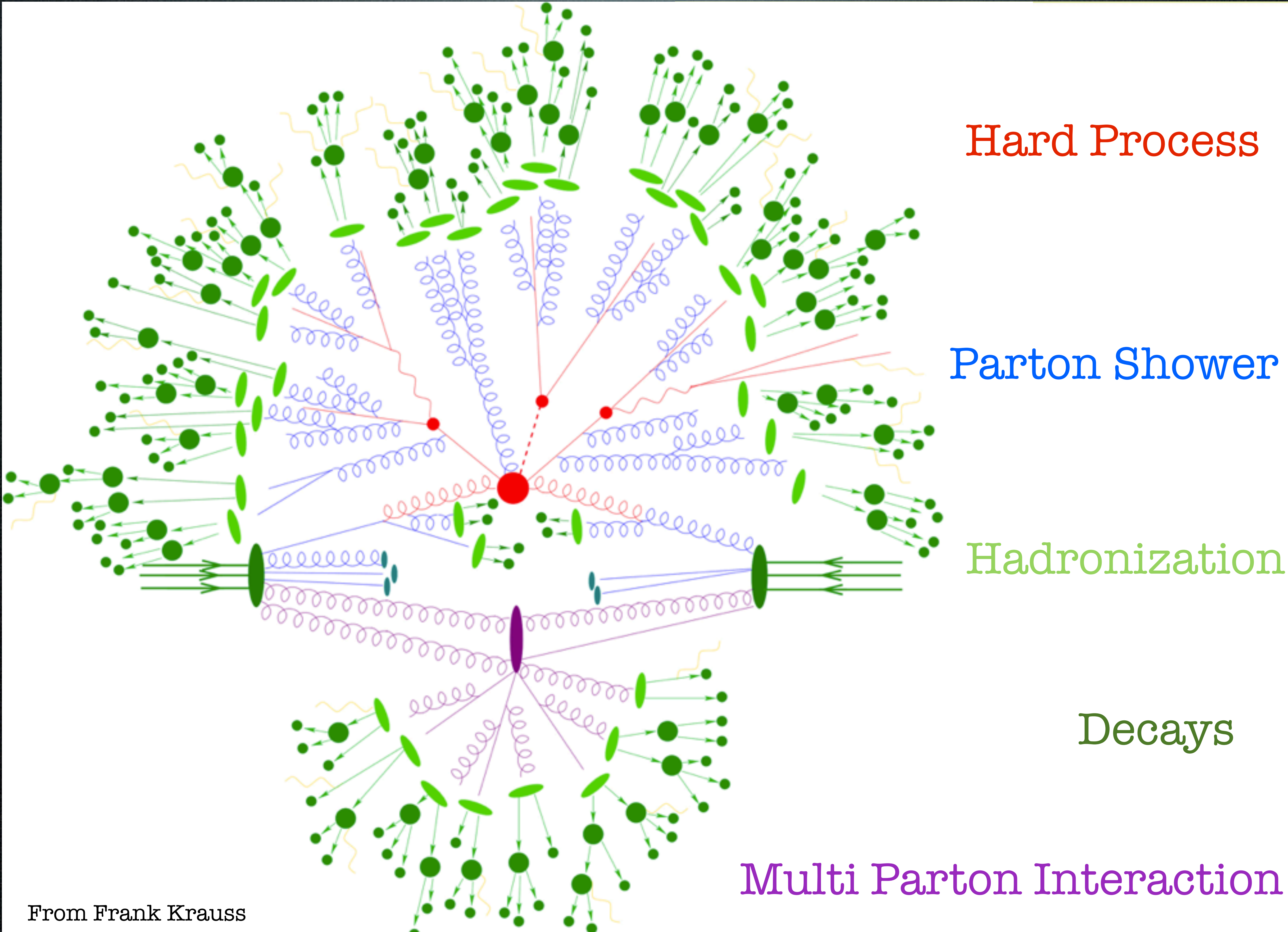
BBR: Beam-beam remnants
MPI: Multiple Parton interactions
ISR/FSR: Initial/Final state radiation

Glossary

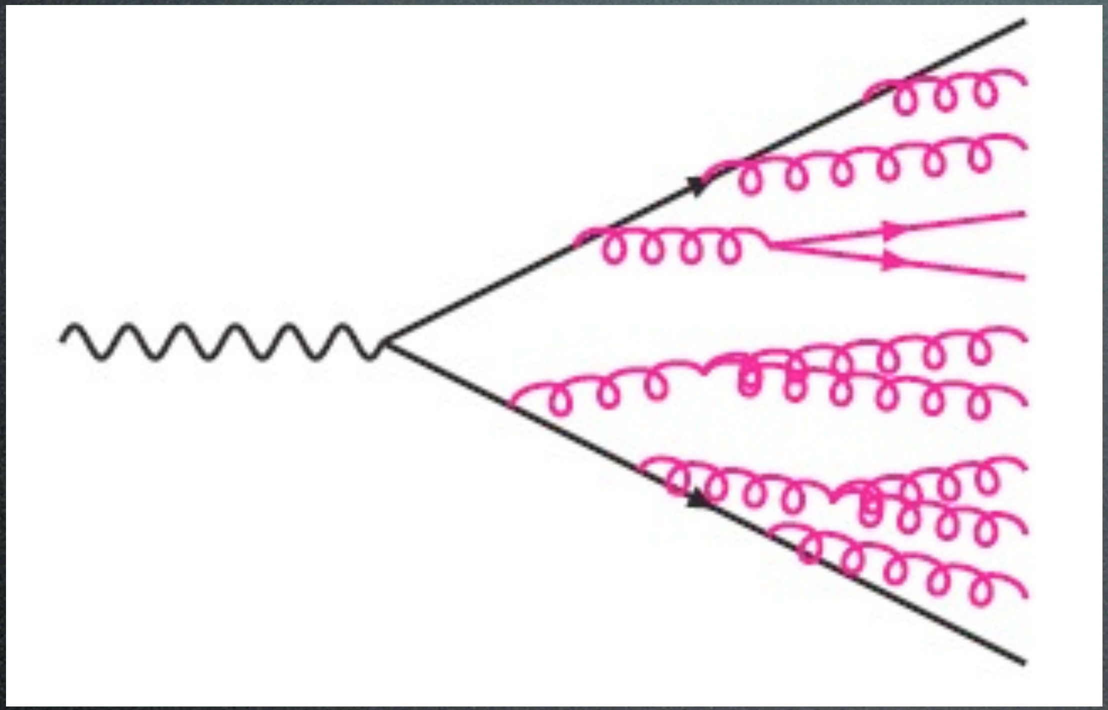
- Minimum-bias (MB): Pretty much everything, exact definition trigger dependent.
- Underlying event (UE): background to events with an identified hard scatter (more like the actual interesting events we want to look at)
- Pileup (PU): (uncorrelated) separate collisions within the same/different bunch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

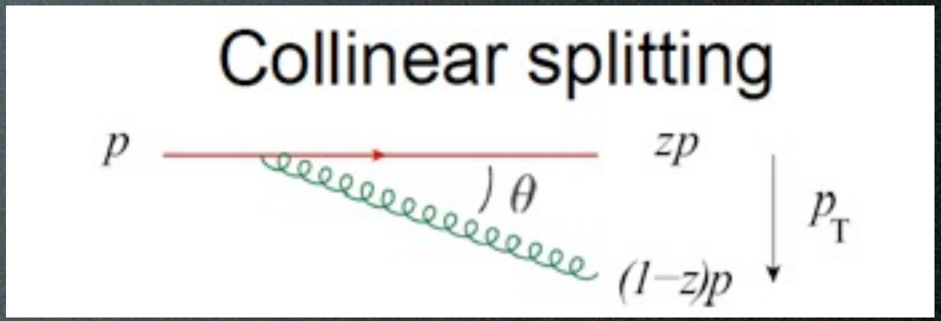
Monte Carlo Models

- 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, aMC@NLO, Madgraph, Powheg ...
- Generators used mostly for a specific process: Phojet (diffraction), HIJING (heavy ion), AcerMC (top), JHU (spin and polarization information)...


Hard Process
Parton Shower
Hadronization
Decays
Multi Parton Interaction
From Frank Krauss

Parton Shower





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

- Each branching governed by DGLAP equation.

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

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 . . . "

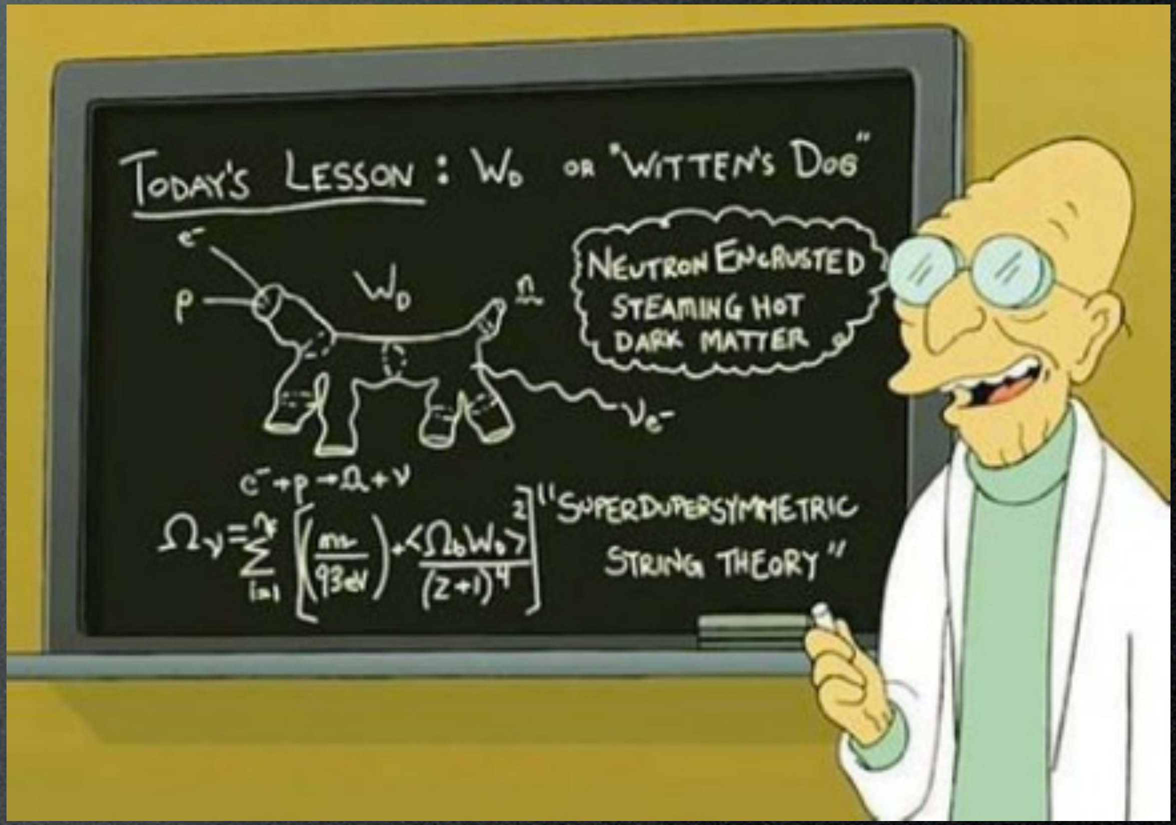
– Richard Feynman and Rick Field, 1978

Why do we care?

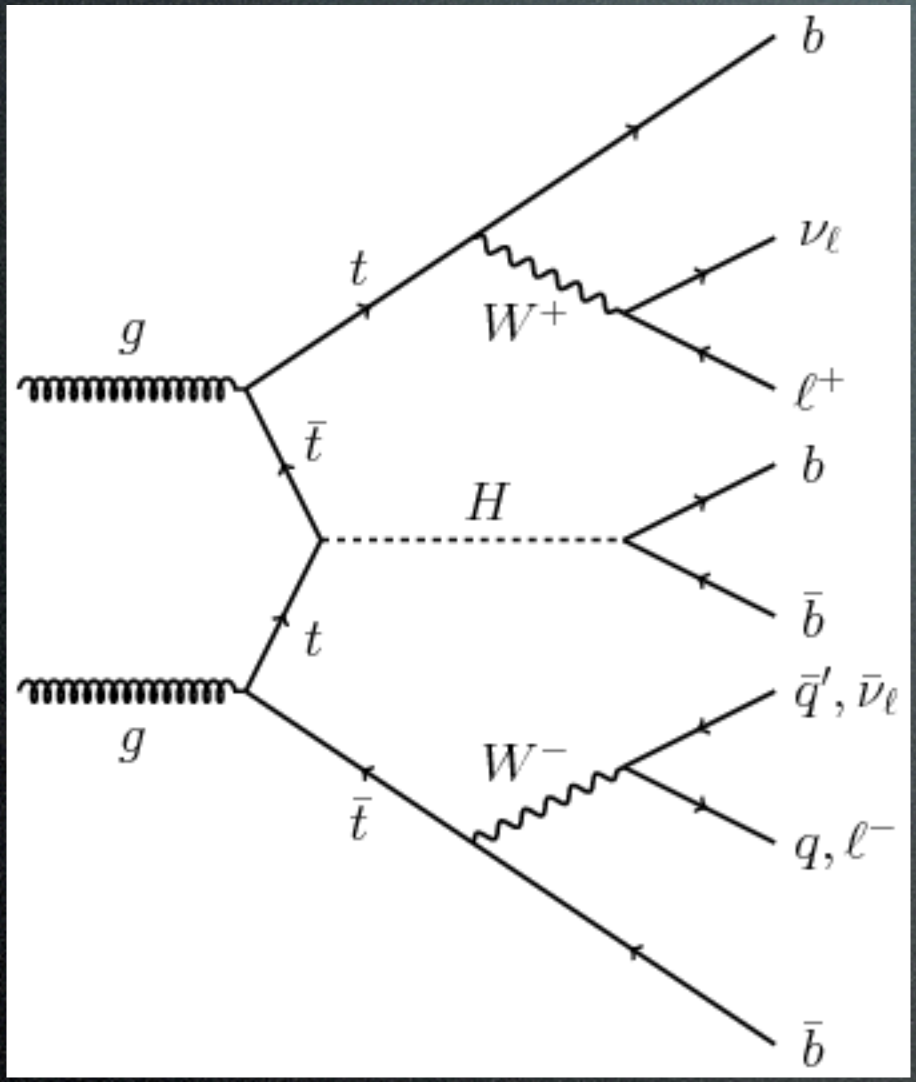
- The “background” to the interesting physics signals are the SM processes.

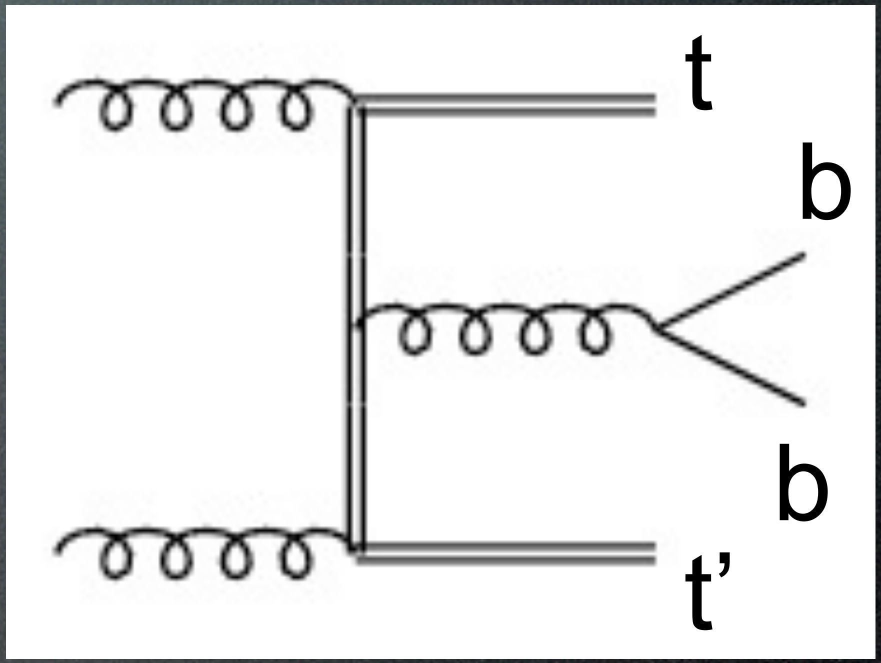
- The hard scattering part can be calculated theoretically (in some order) by QCD matrix elements.

- The soft part is not calculable, so we use phenomenological models implemented in Monte Carlo event generators.



An Example



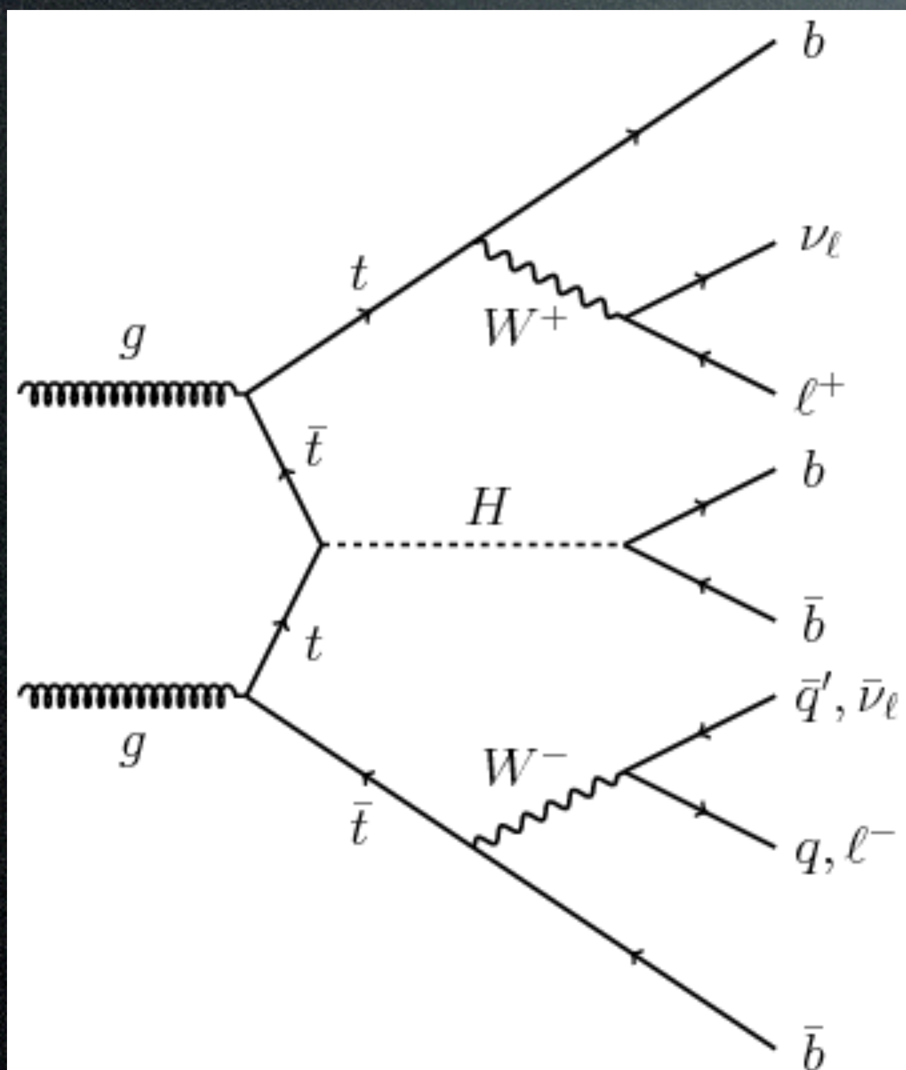


Signal: ttH(bb)

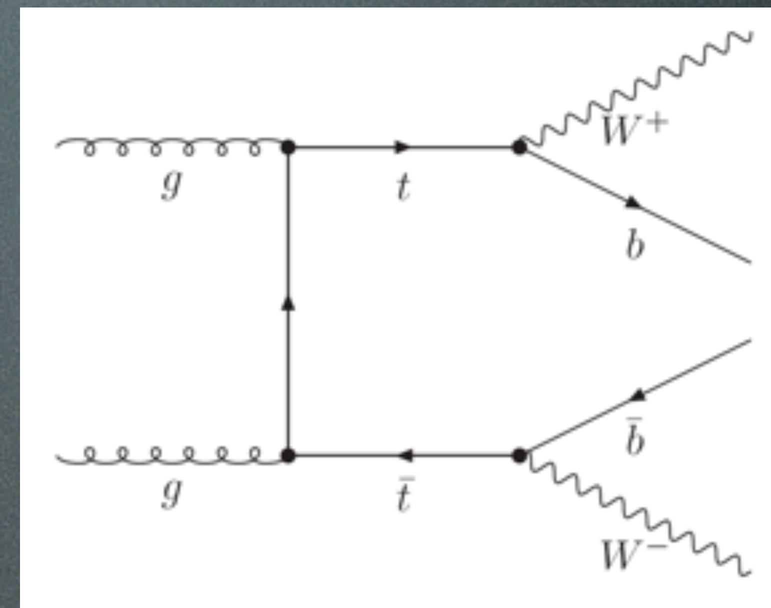
BG: ttbb

important for measuring Yukawa couplings

One of the hardest measurements now...

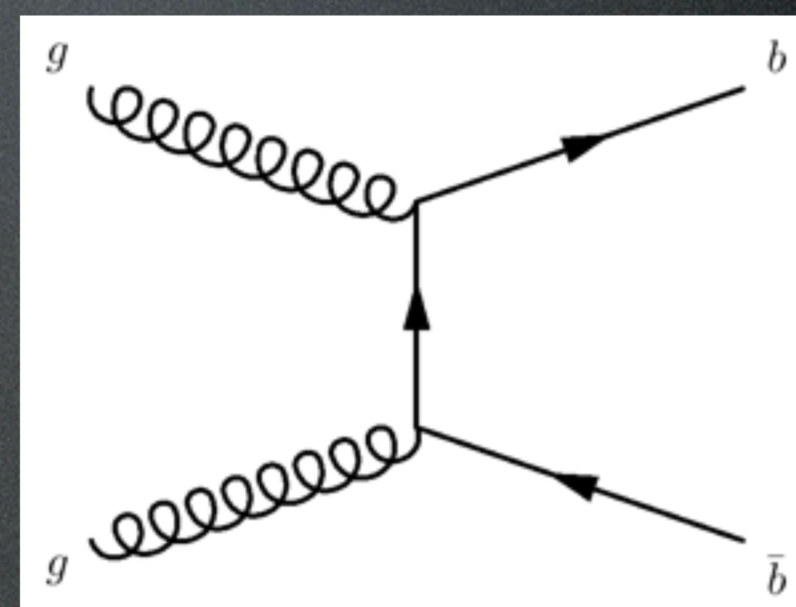


Signal: ttH(bb)

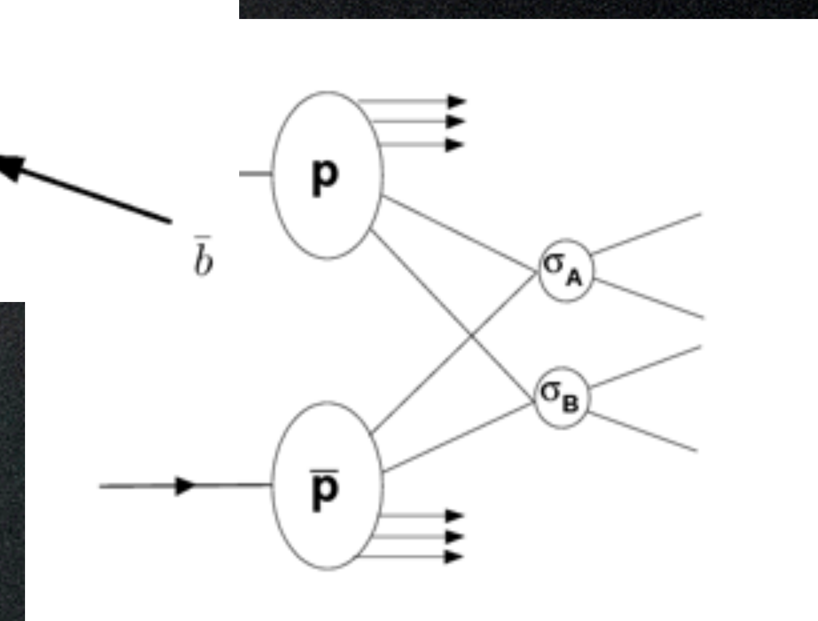


tt

+



(DPI) bb



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.

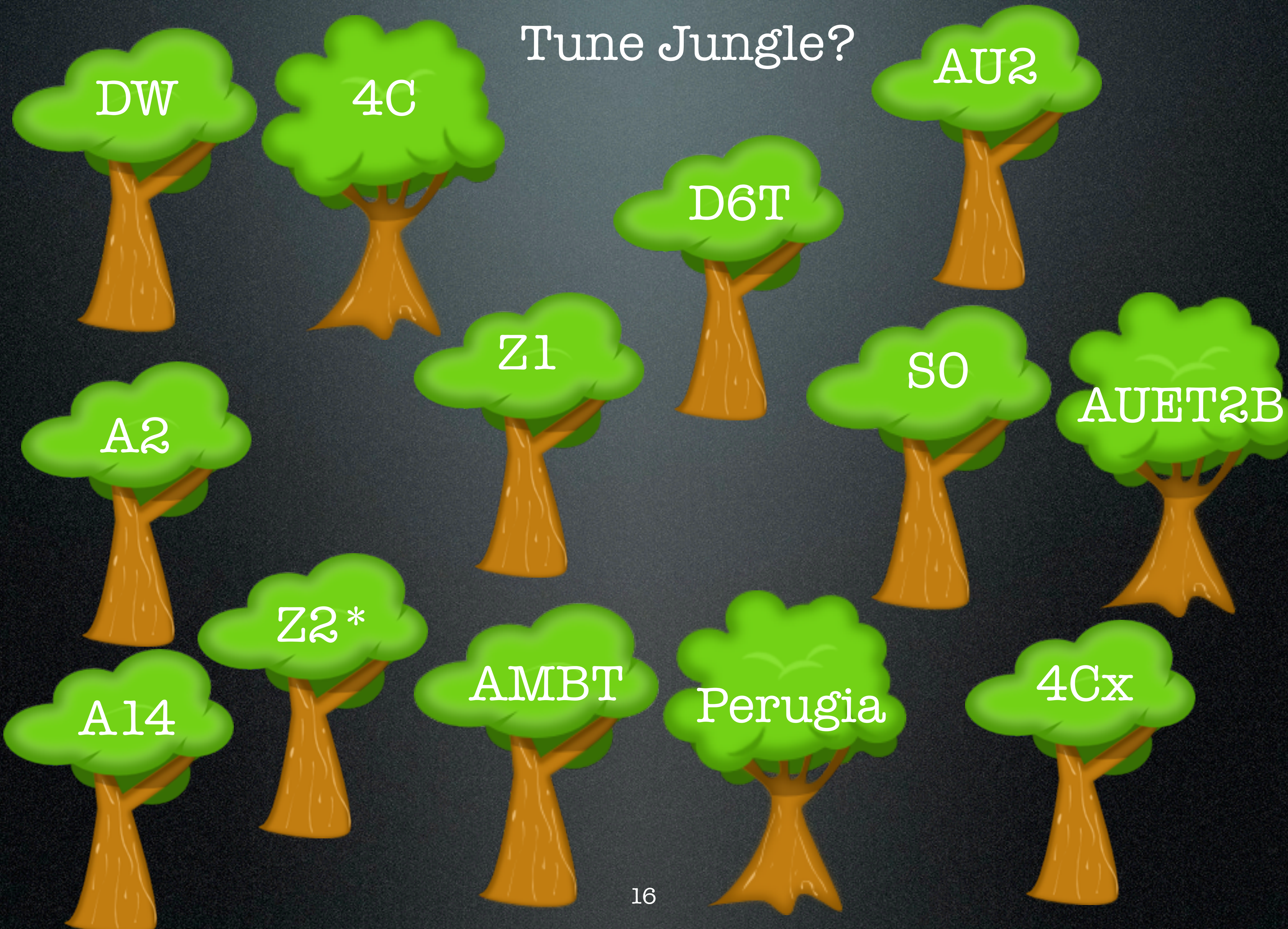
Apollo's priestess, Pythia, performing the duty of the oracle

- 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 intution/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?



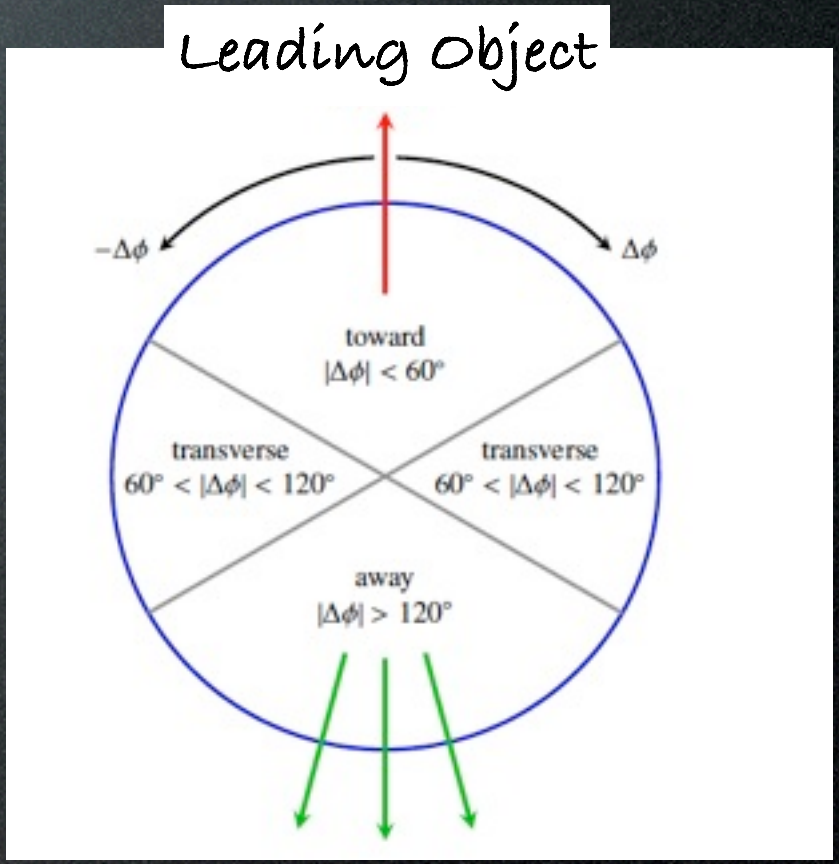
Tevatron Era Tunes

- 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).

Measuring the Underlying Event

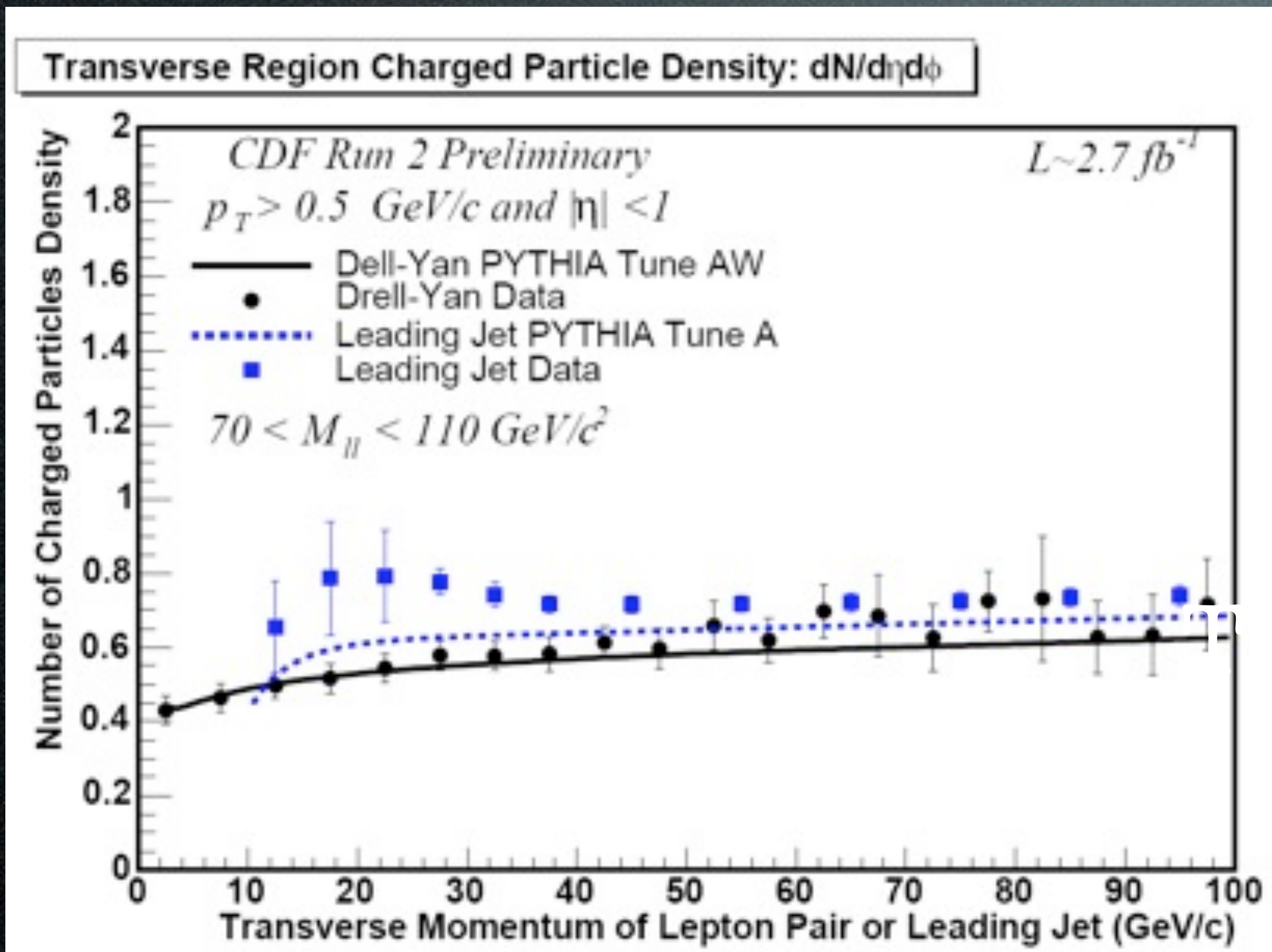
UE Measurements

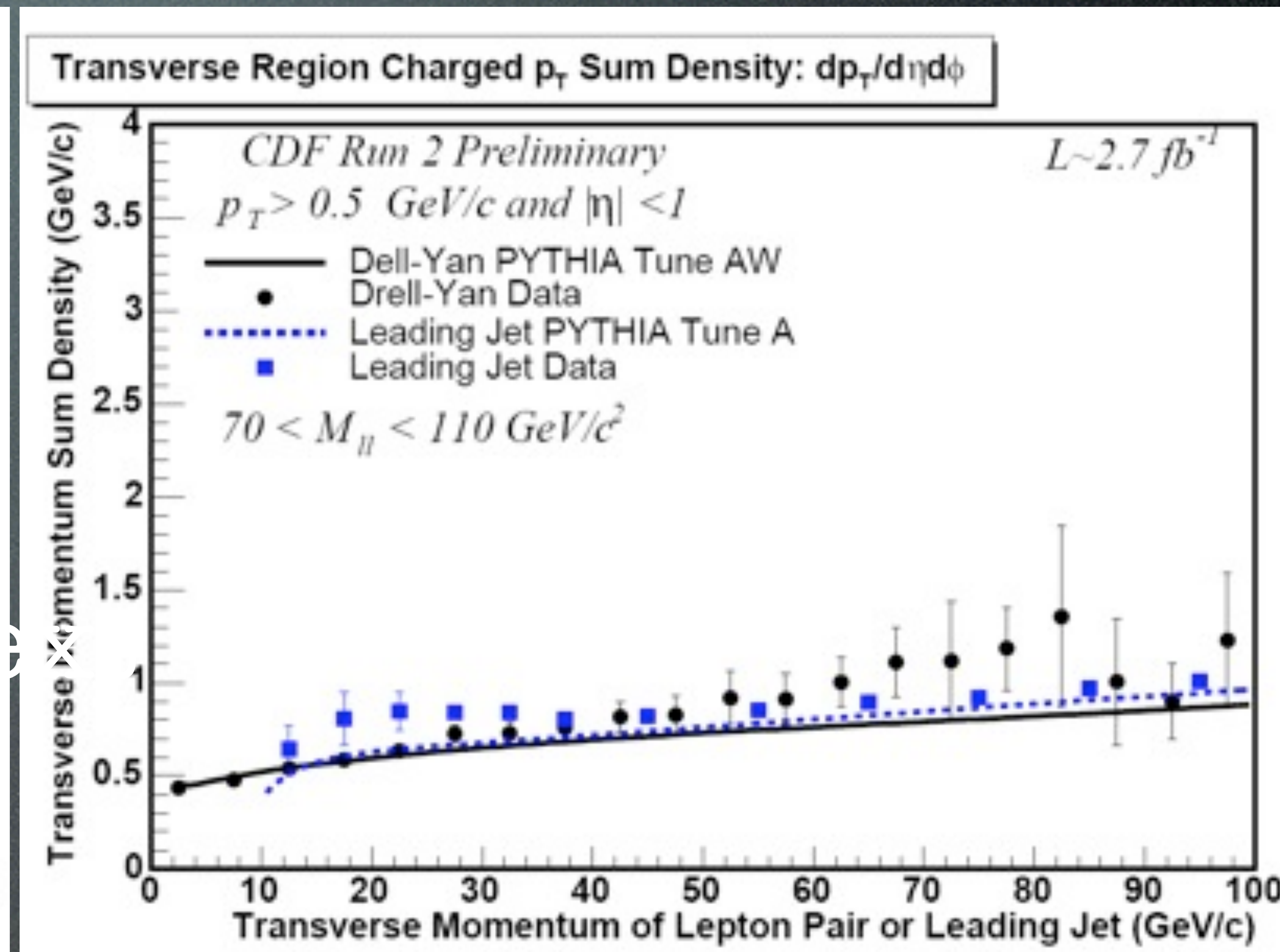
- Many results from ATLAS and CMS.
- In busy LHC environments, how much of "UE" is UE?
- Sensitive to DPI contribution.
- Not just comparing with PS models, but with ME+PS setups too.



Leading Jet and Z UE Results

Phys.Rev. D82 (2010) 034001

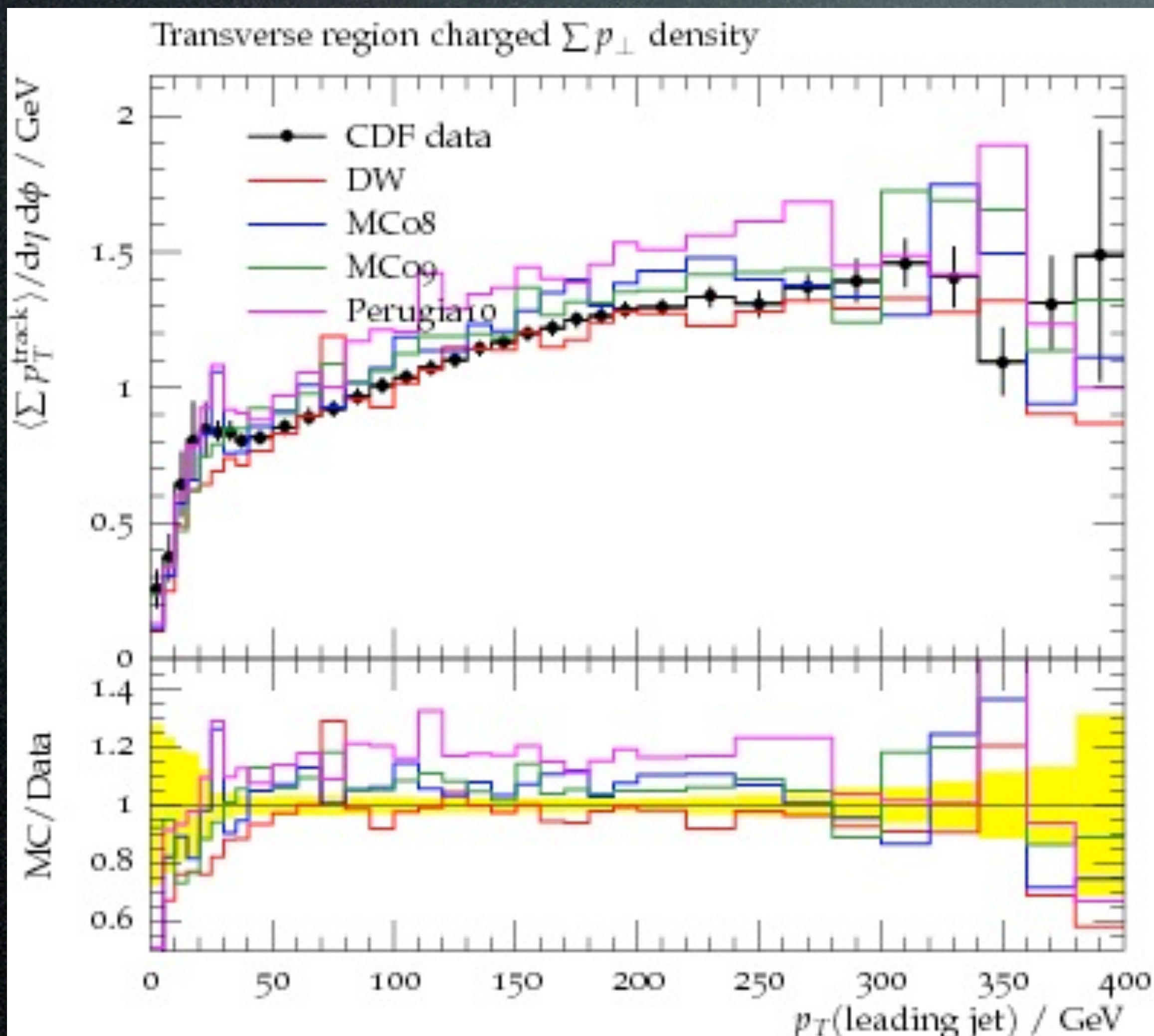


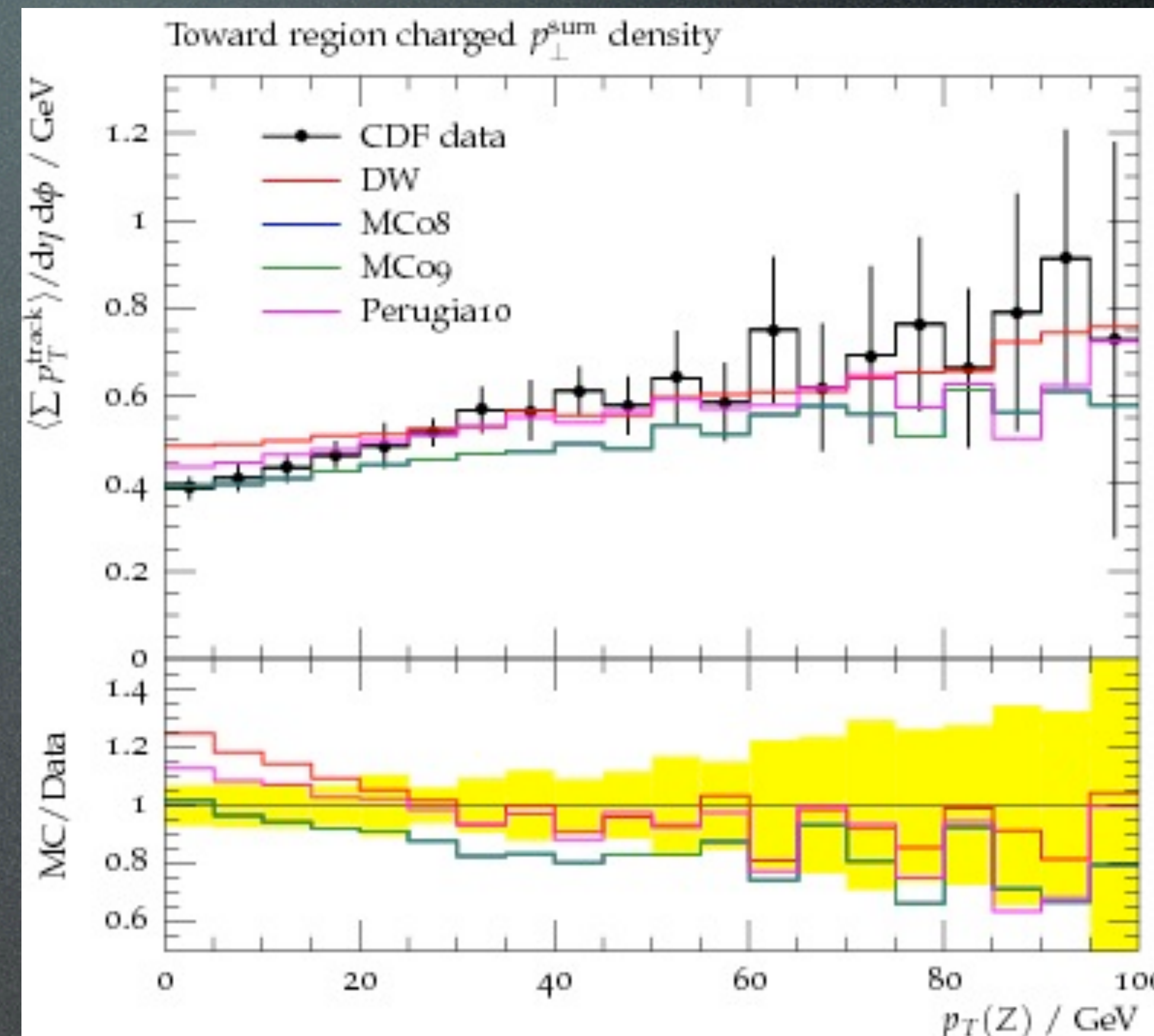


UE activity in Z-boson and jet events fairly similar in Tevatron.

Is it still the case at the LHC?

Pre-LHC tunes





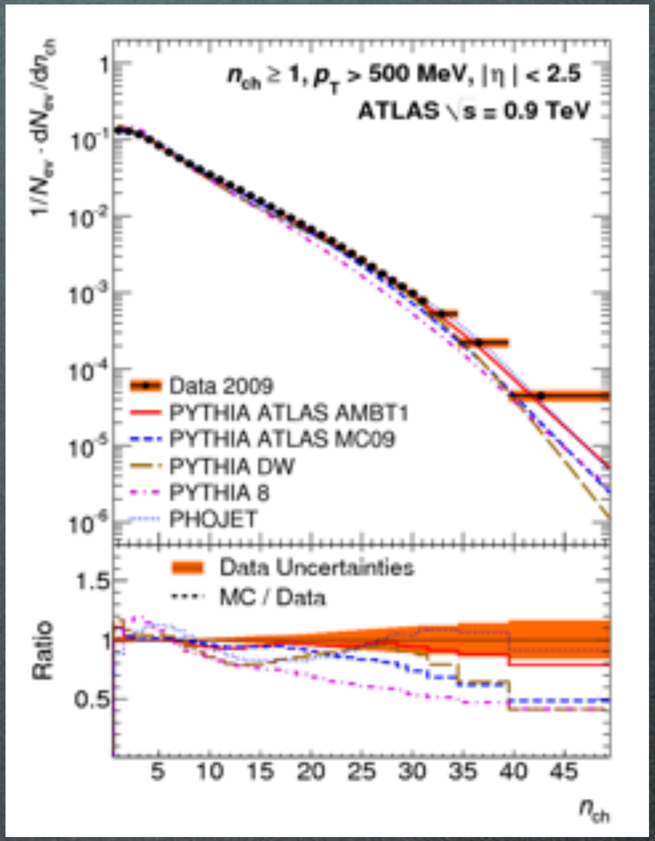
Phys.Rev. D82 (2010) 034001

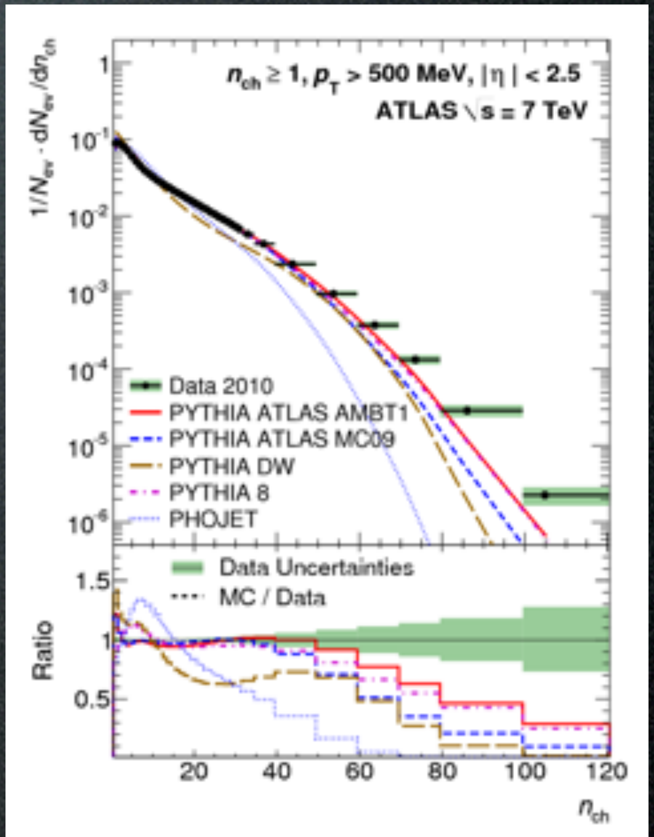
The tunes do quite well ...

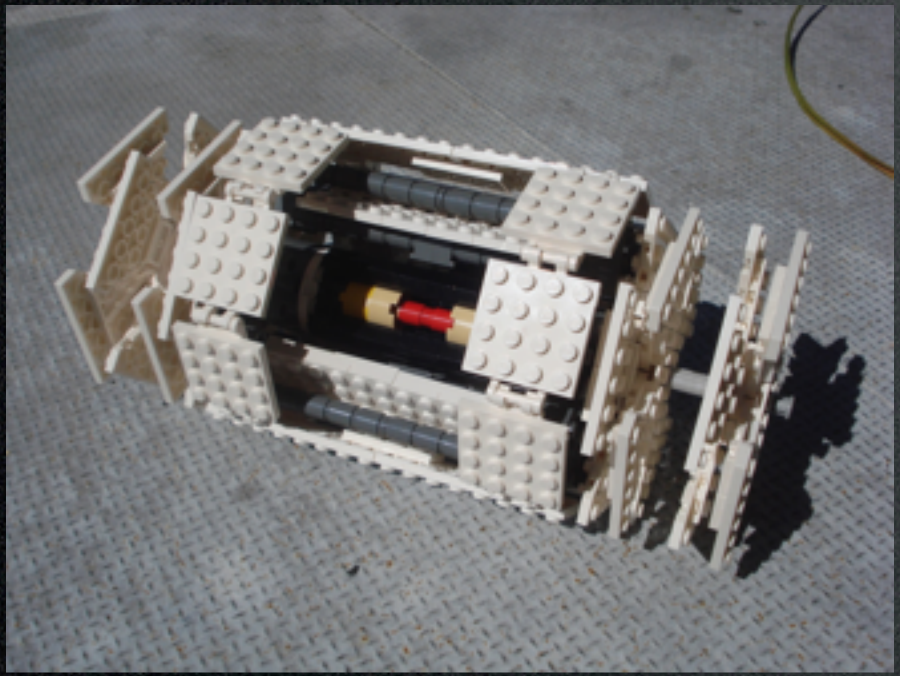
Did they work at the LHC?

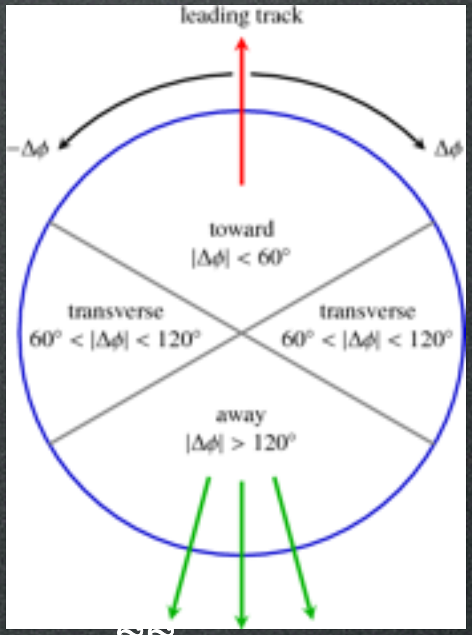
Then Came the LHC

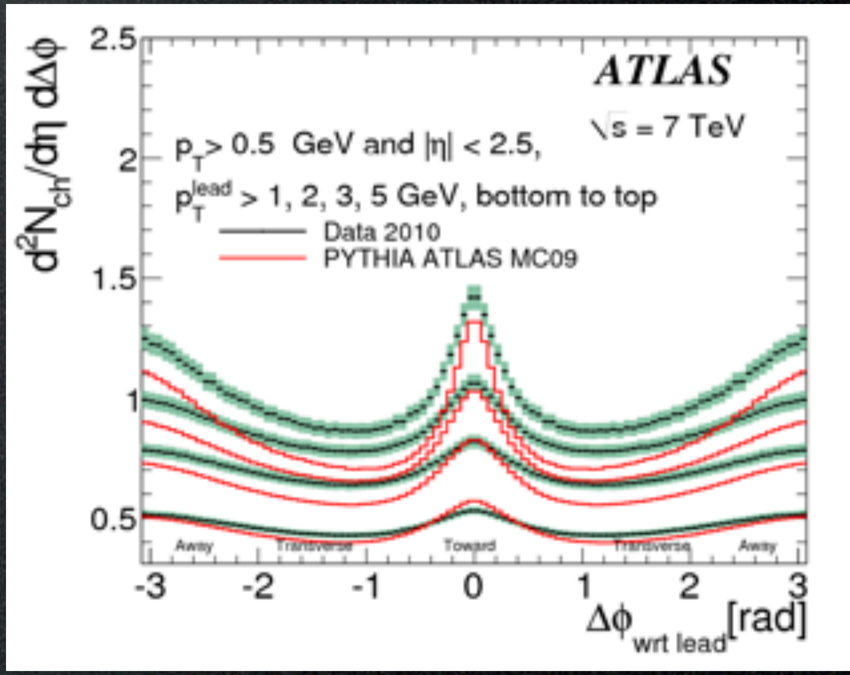
- Tevatron tunes did not agree with the early minbias and underlying event data.
- Not just at 7 TeV, but also at 900 GeV!





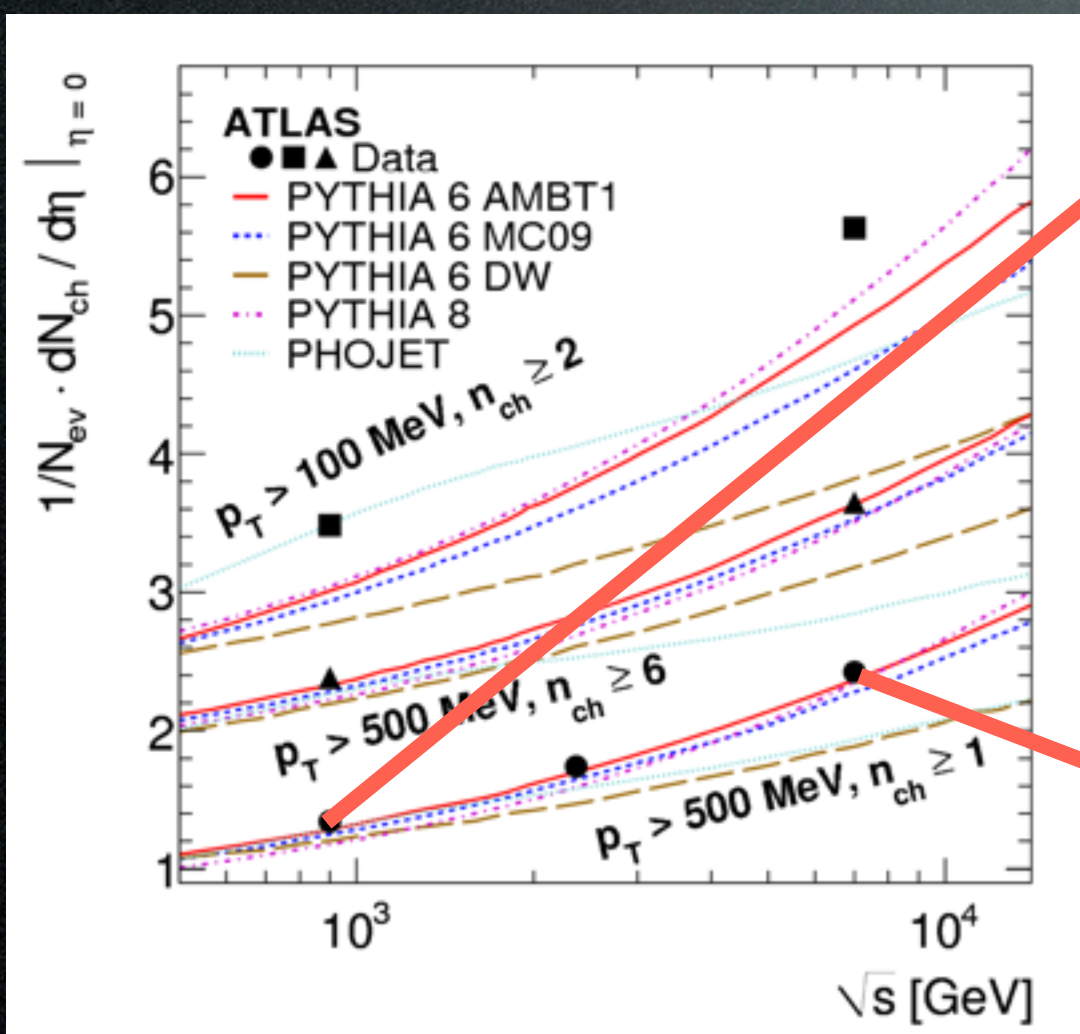






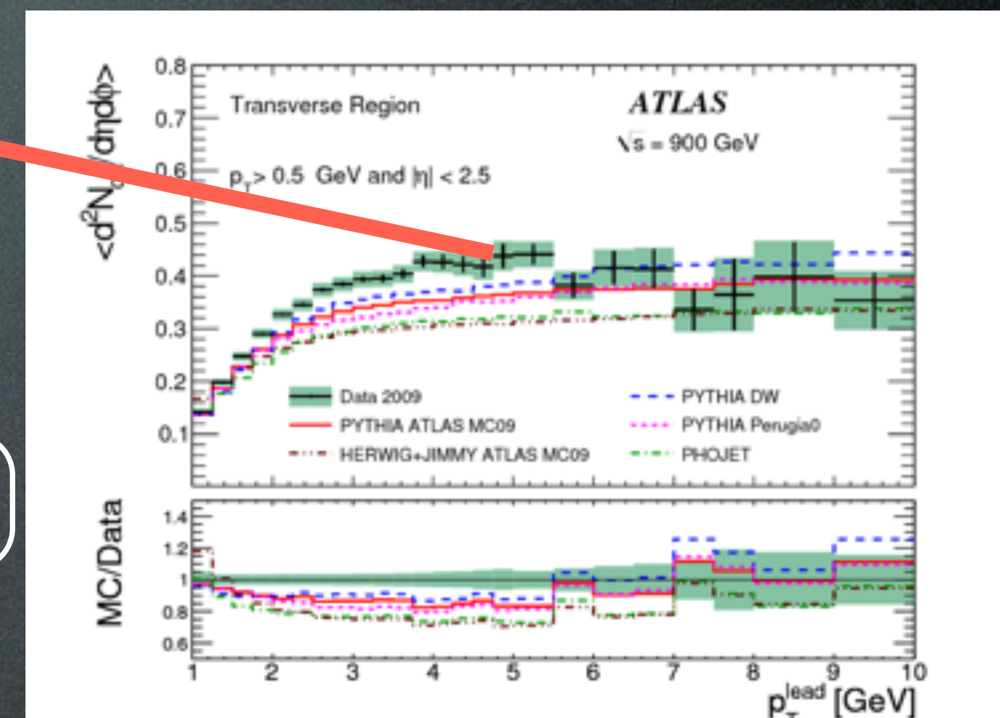
A slight detour: comparsion between UE and MB

0.4 ->2.5



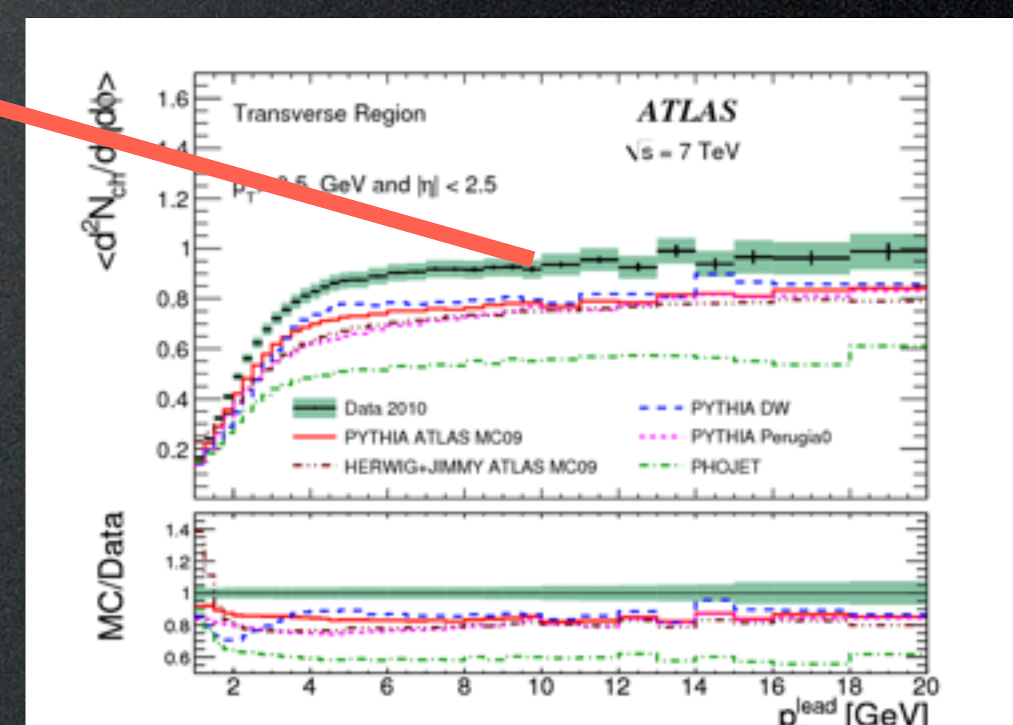
~1.3

factor of 2!



0.8 -> 5

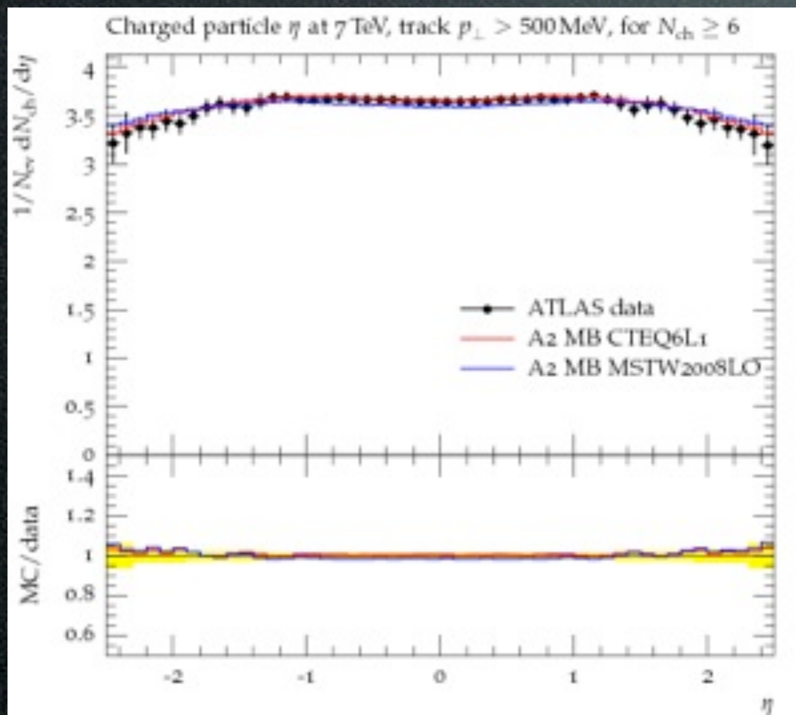
~2.4

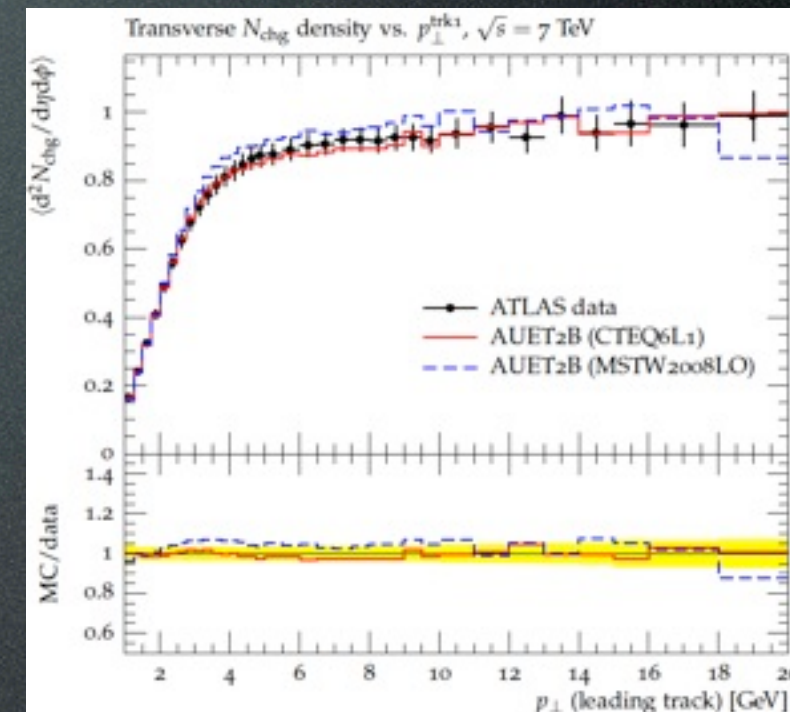


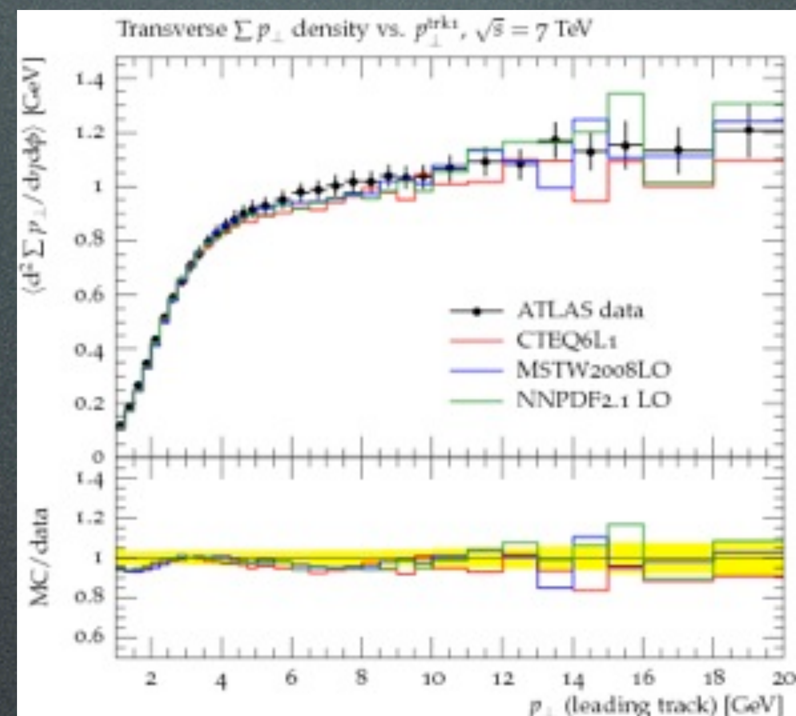
Post-LHC Tunes

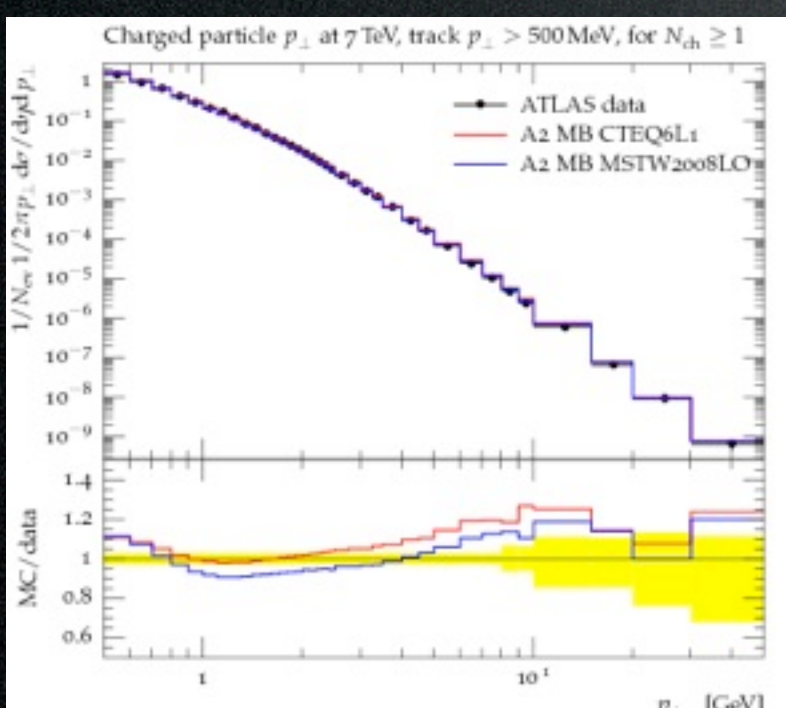
- (Pythia 6)ATLAS Tunes: AMBT1, AMBT2, AMBT2B, AUET2, AUET2B. [First separate MB/UE Tunes. also for many PDFs.]
- (Pythia 6) CMS Tunes: Z1, Z2, Z2*.
- (Pythia 6) Perugia 2011 tunes.
- (Pythia 8) author tunes: 4C, 4Cx.
- (Pythia 8) ATLAS tunes: A2, AU2, A14.

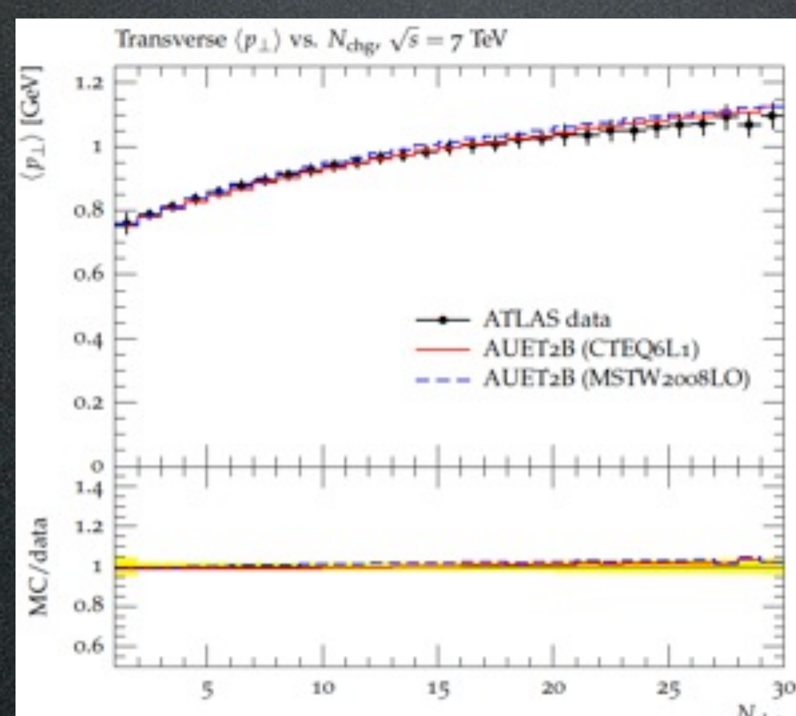
How do they do?





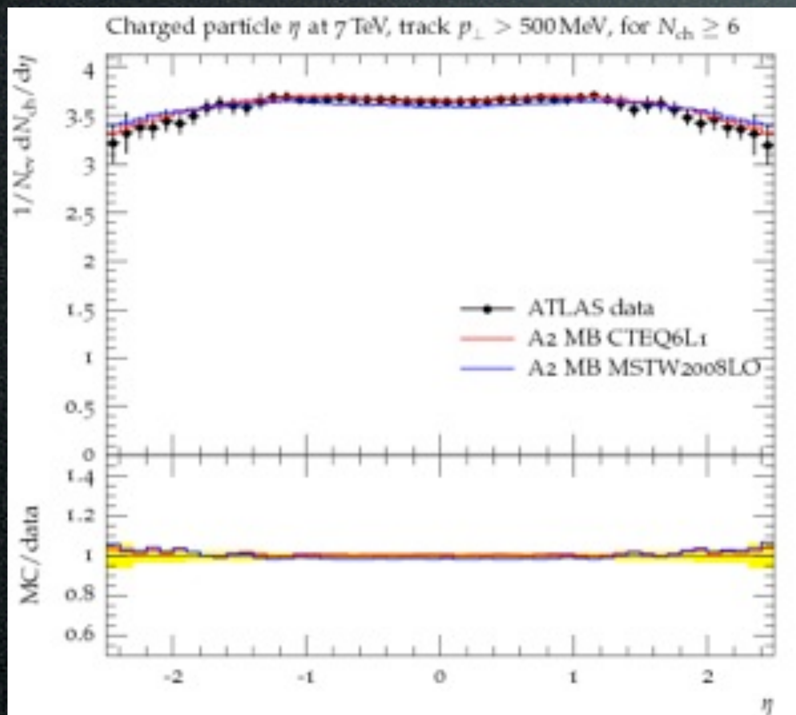


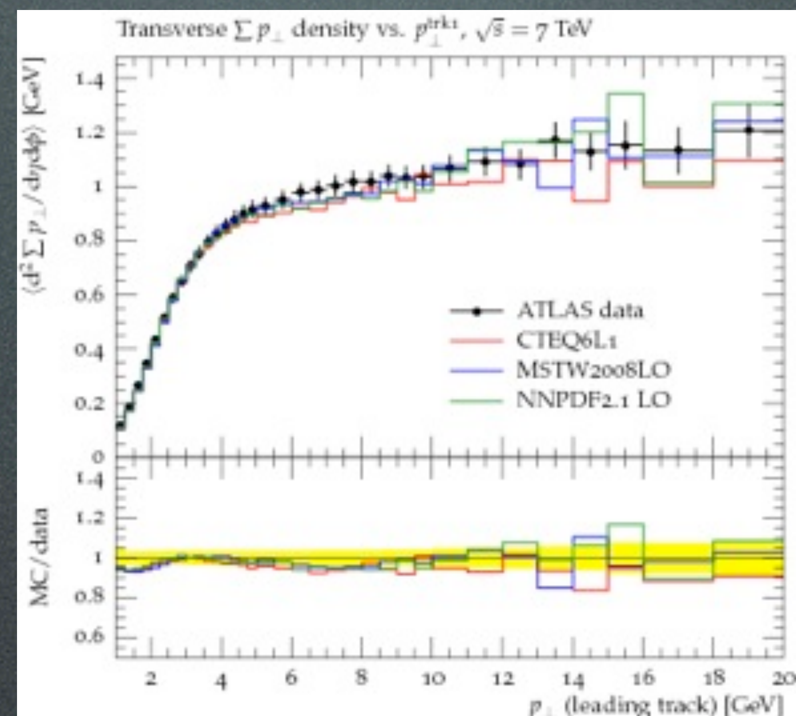


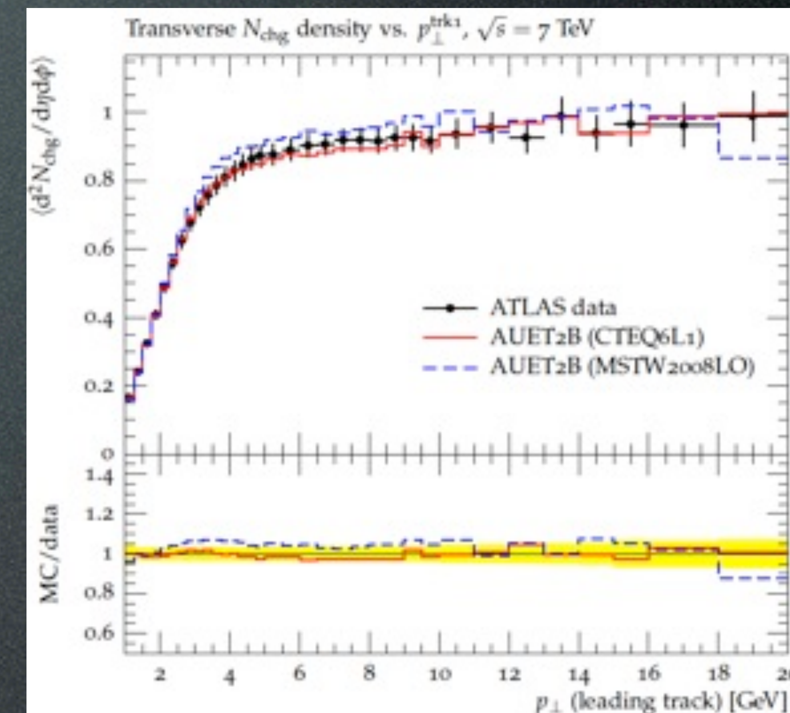


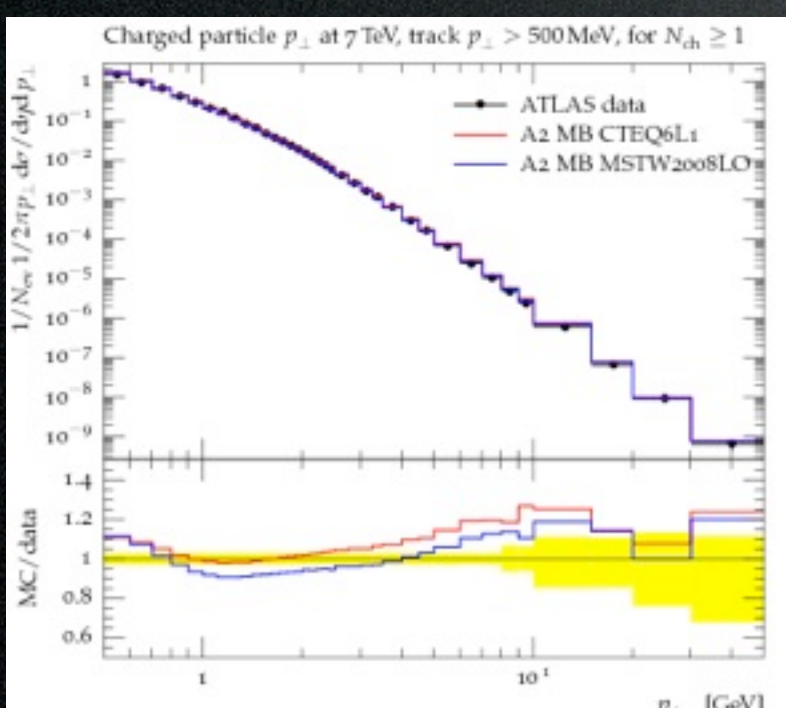
How good is good?

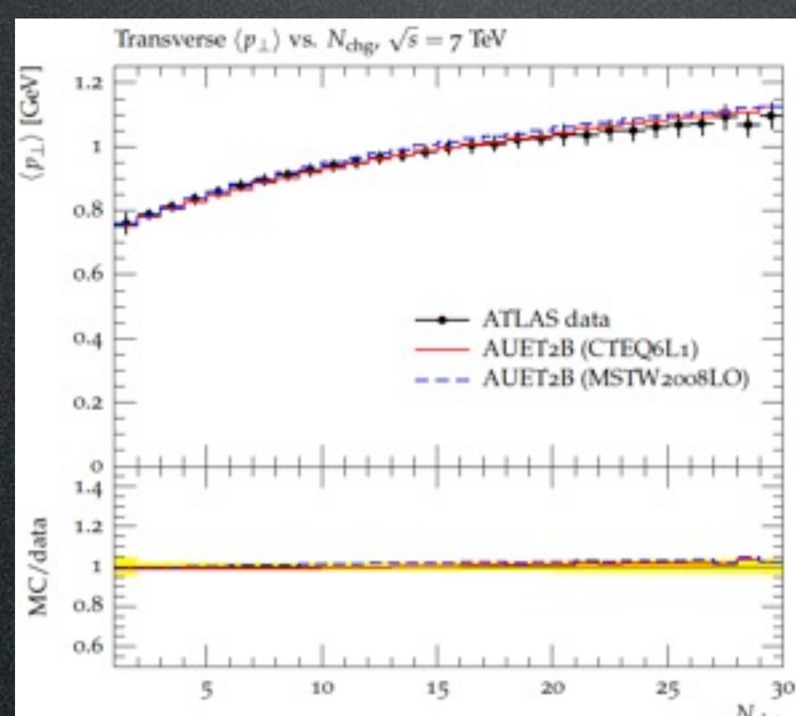
How do they do?





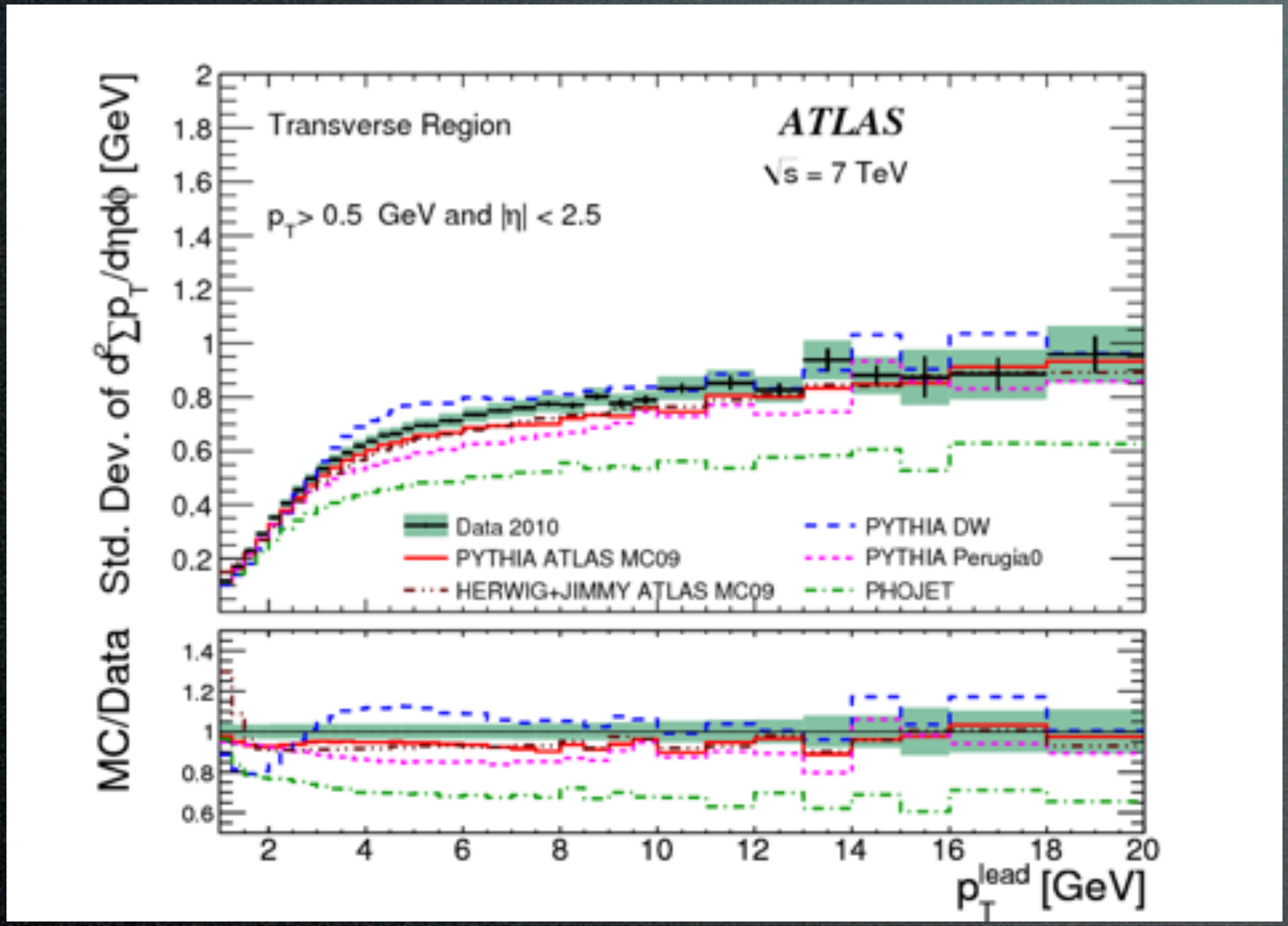




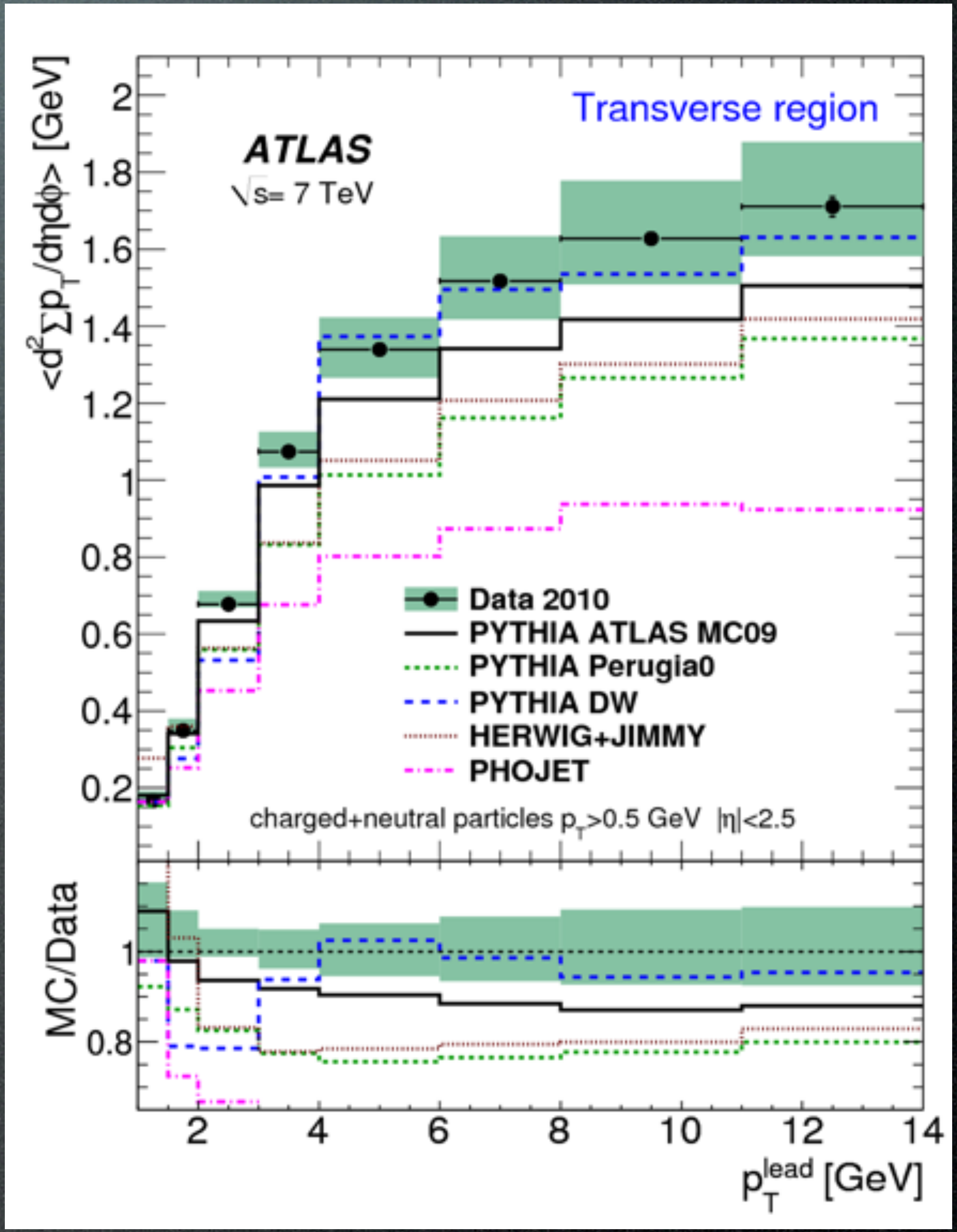


How good is good?
Is not it amazing that
the models are doing
so well?

Back to (early)UE Results

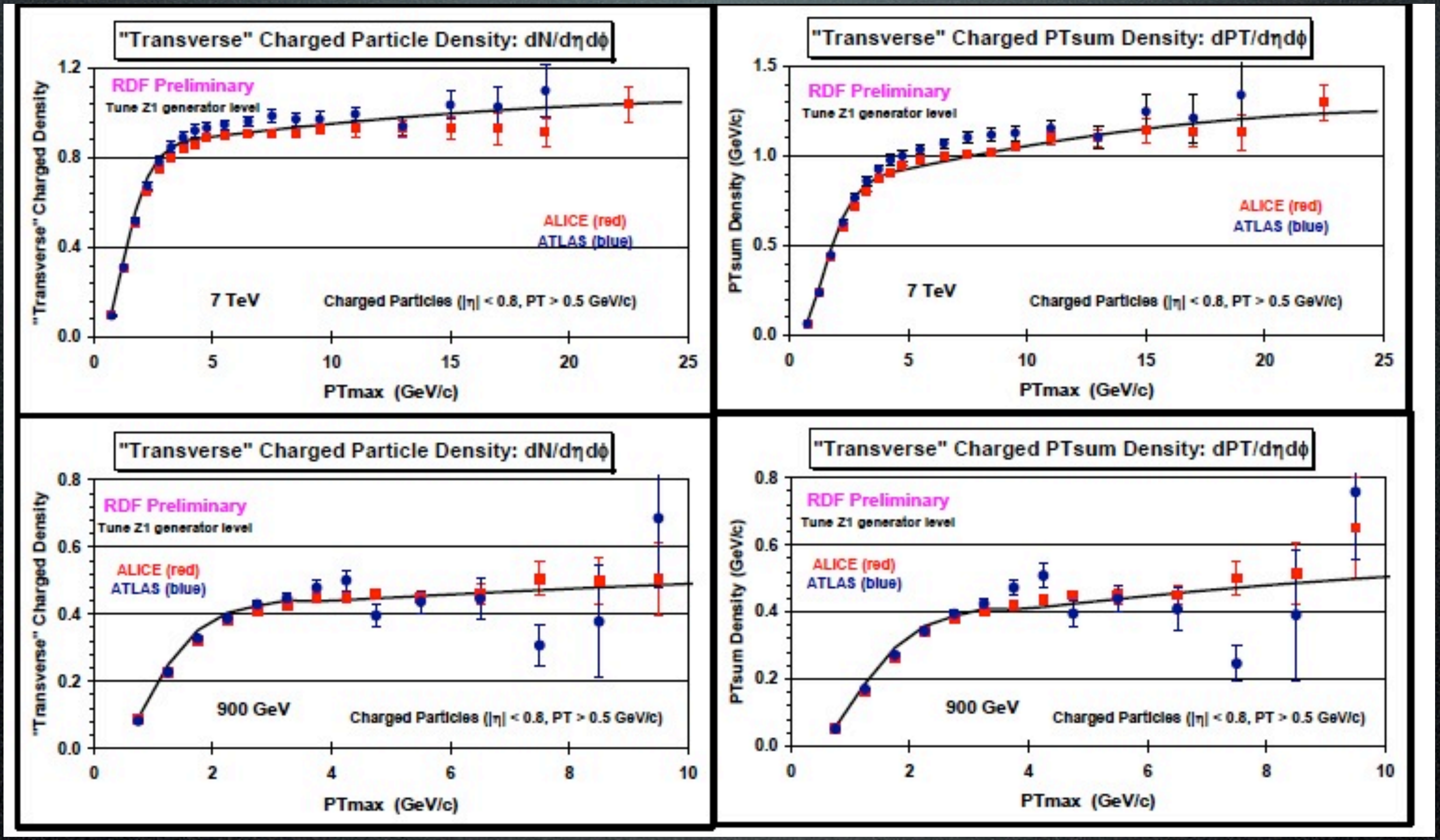


shows UE activity can not be subtracted as an average “pedestal” from each event.



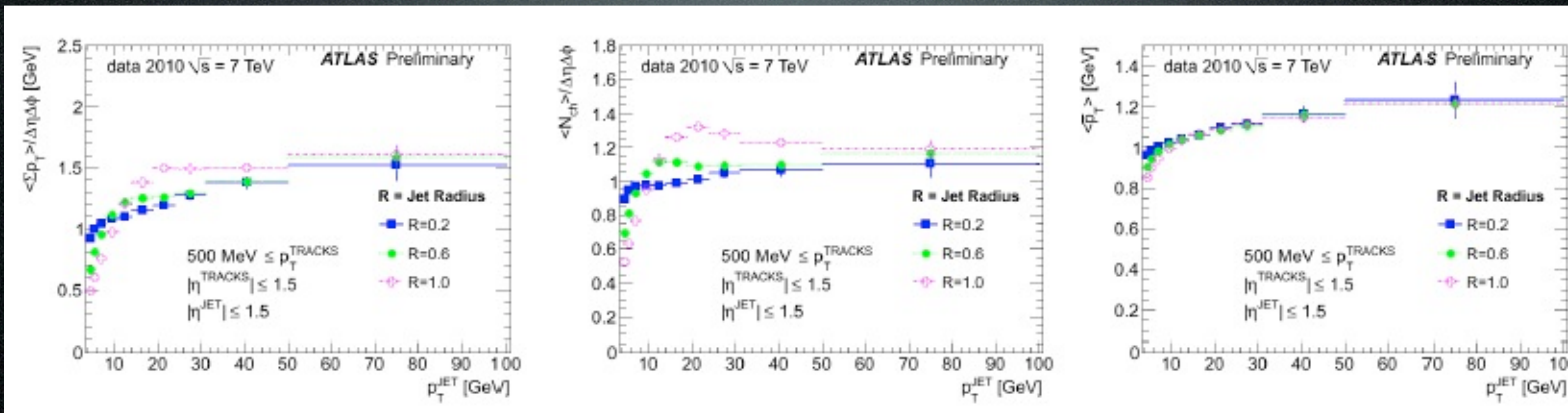
Sensitive to both charged and neutral component of UE.

LPCC UE&MB WG



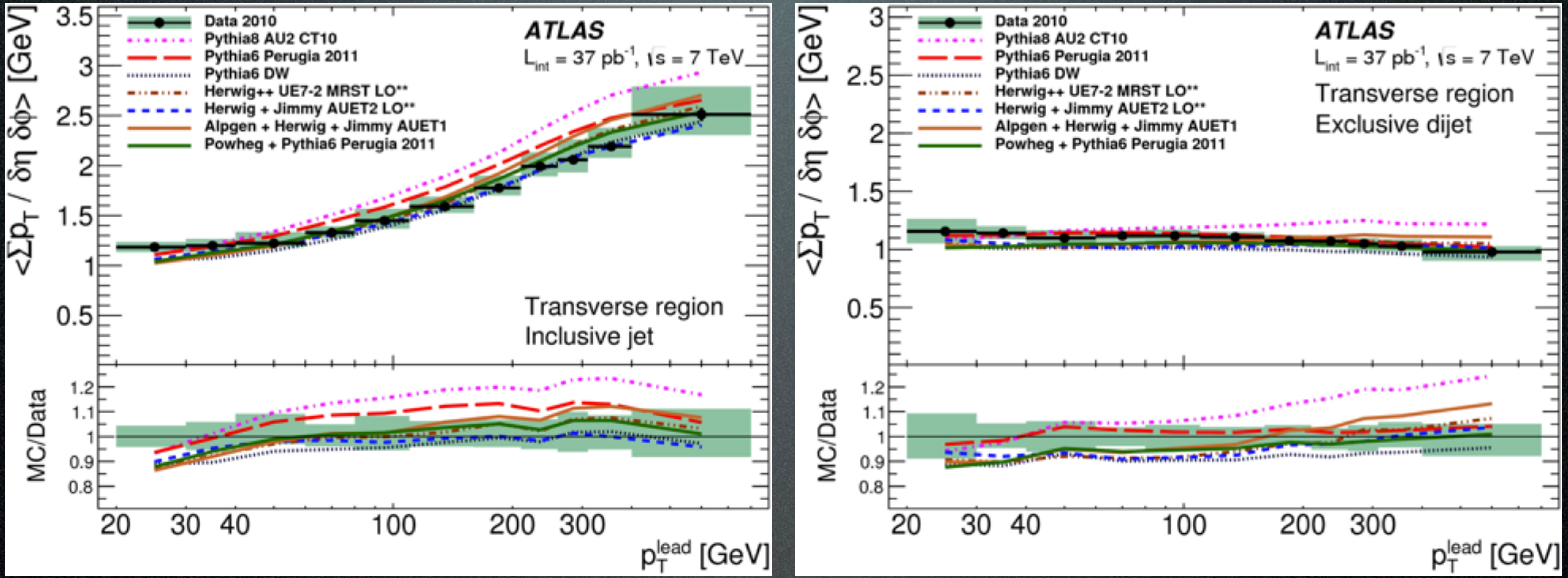
Rick Field: WG meeting, 17th June 2011

Jet Radius Dependence



More UE activity for higher jet radius.
Why?

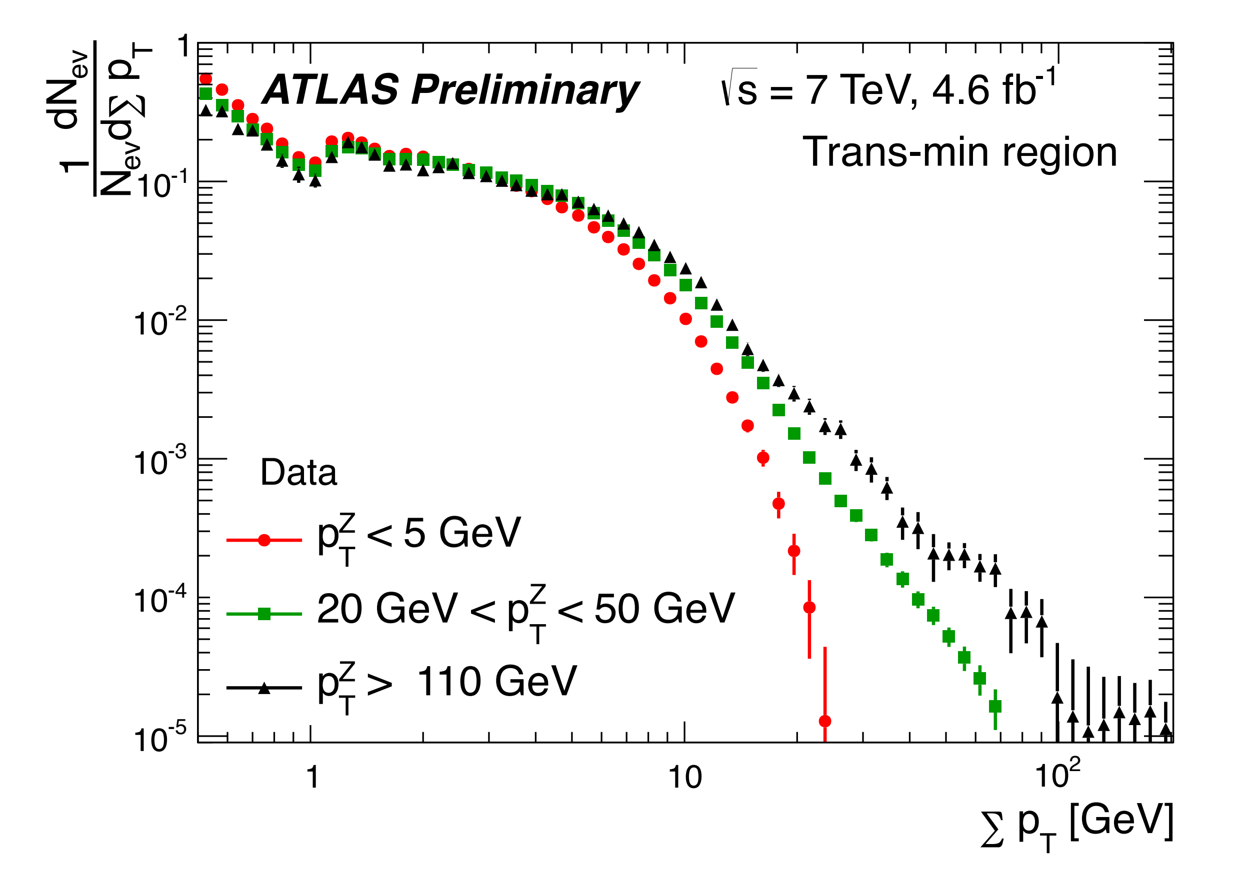
ATLAS Jet UE

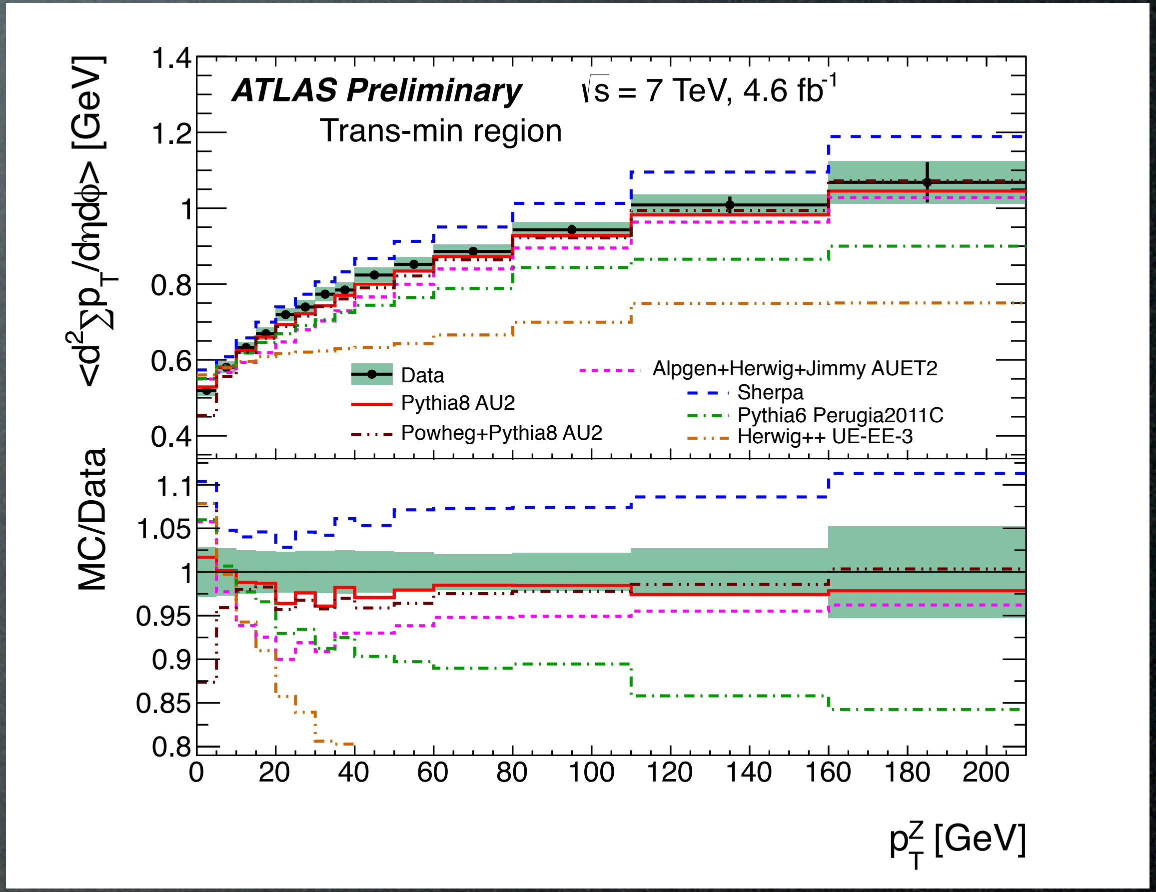


Rise in inclusive, almost flat in when requiring exactly 2 jets .
Models better describe exclusive profile.

arXiv:1406.0392 [hep-ex]

ATLAS Z UE

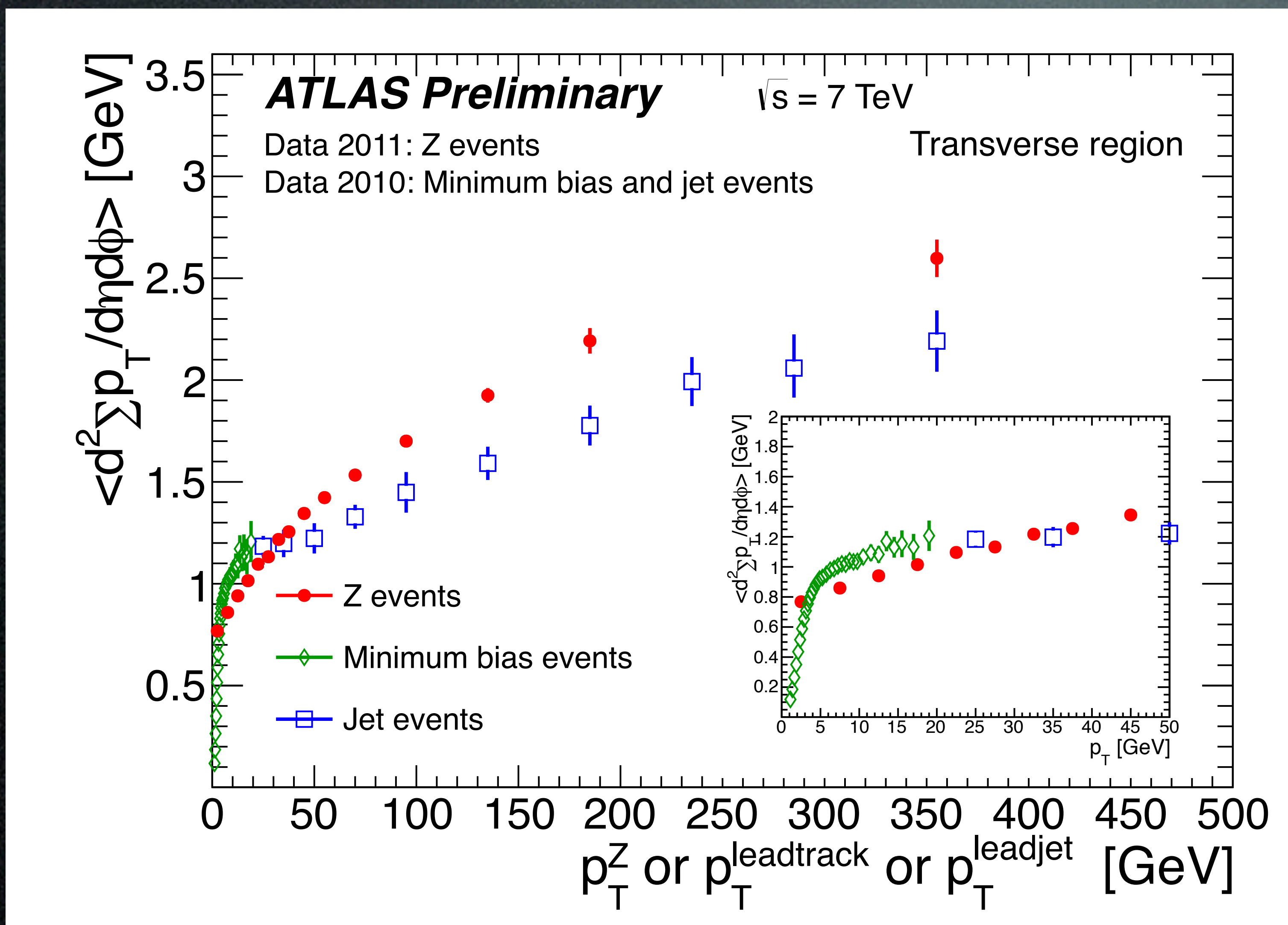


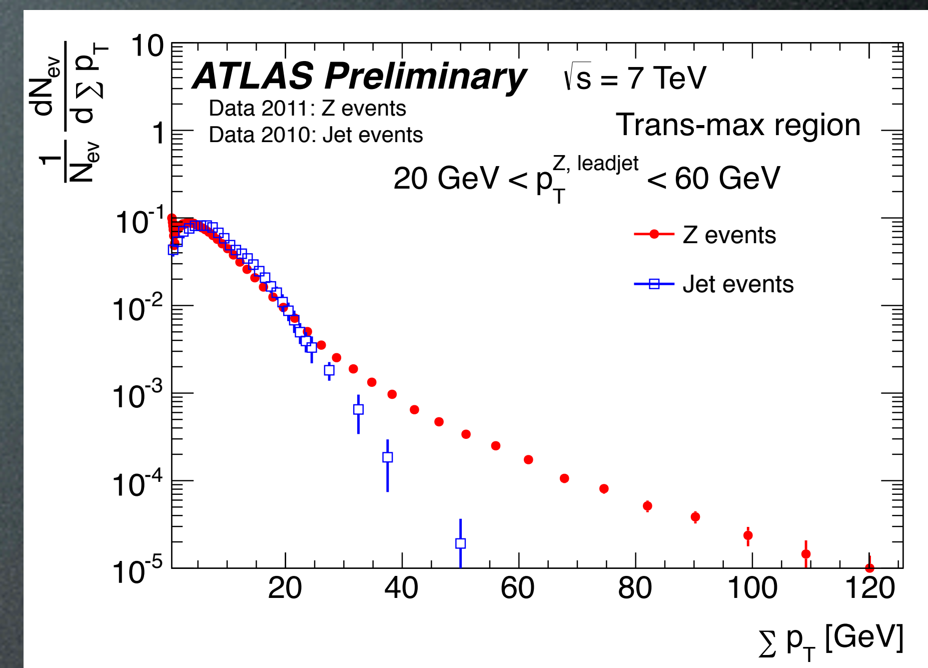


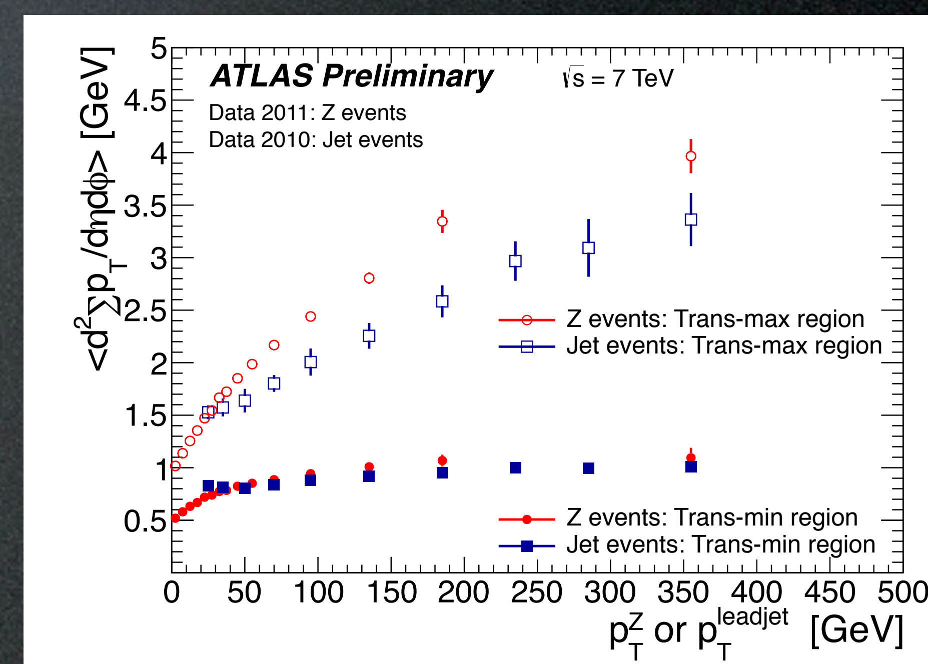
Transmin independent of Z p_T till about 10 GeV, profile best described by Pythia8 and Powheg+Pythia8

However full transverse (or trans-max) regions are described better by NLO or multileg generators than pure LO ones.

Z/Jet UE Comparison







Discrepancy due to selection bias,
trans-min identical.

Isolating the UE

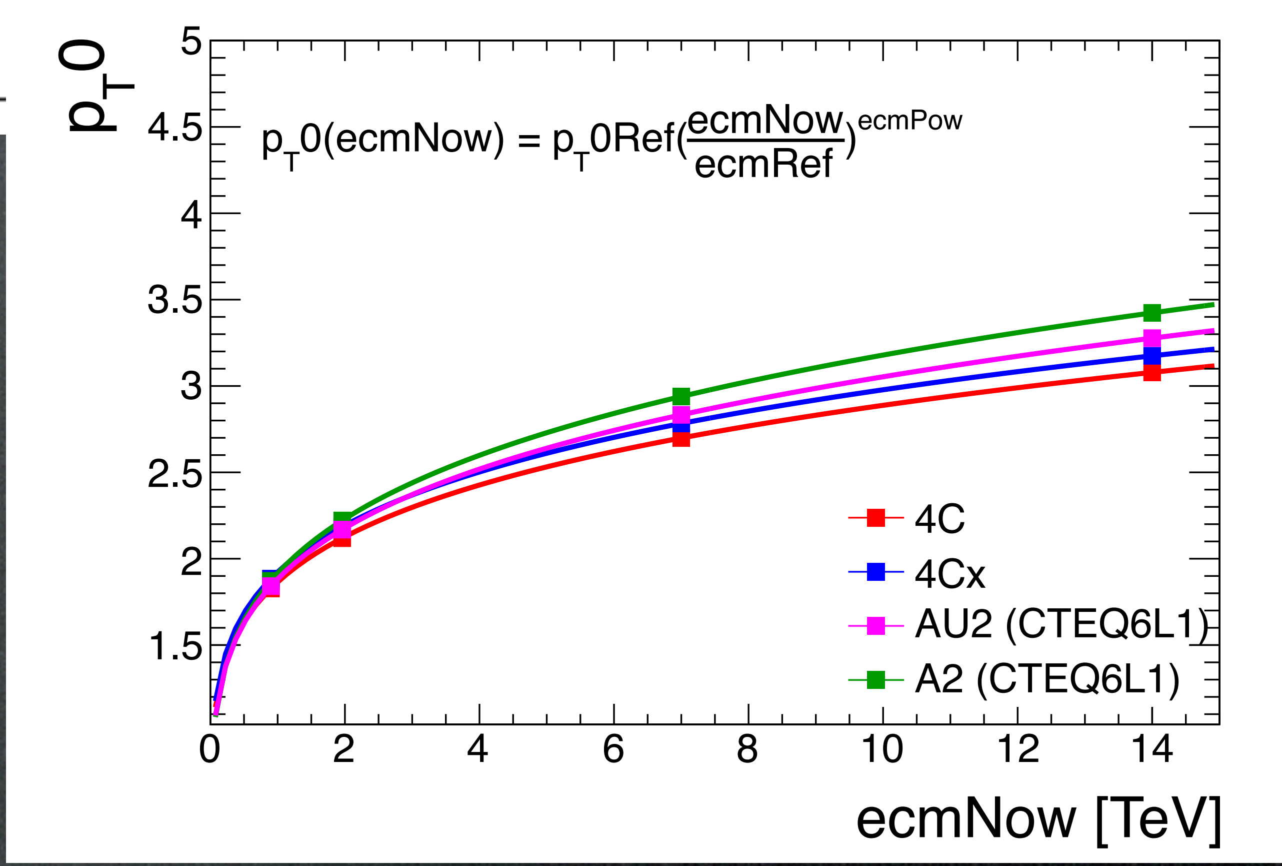
- Full transverse (or trans-max) regions are described better by NLO or multileg generators than pure LO ones.
- Trans-min (and towards region for Z-boson events) were thought to be populated by “pure” UE.
- But at LHC, even those are not flat.

Tuning Shower & MPI Together

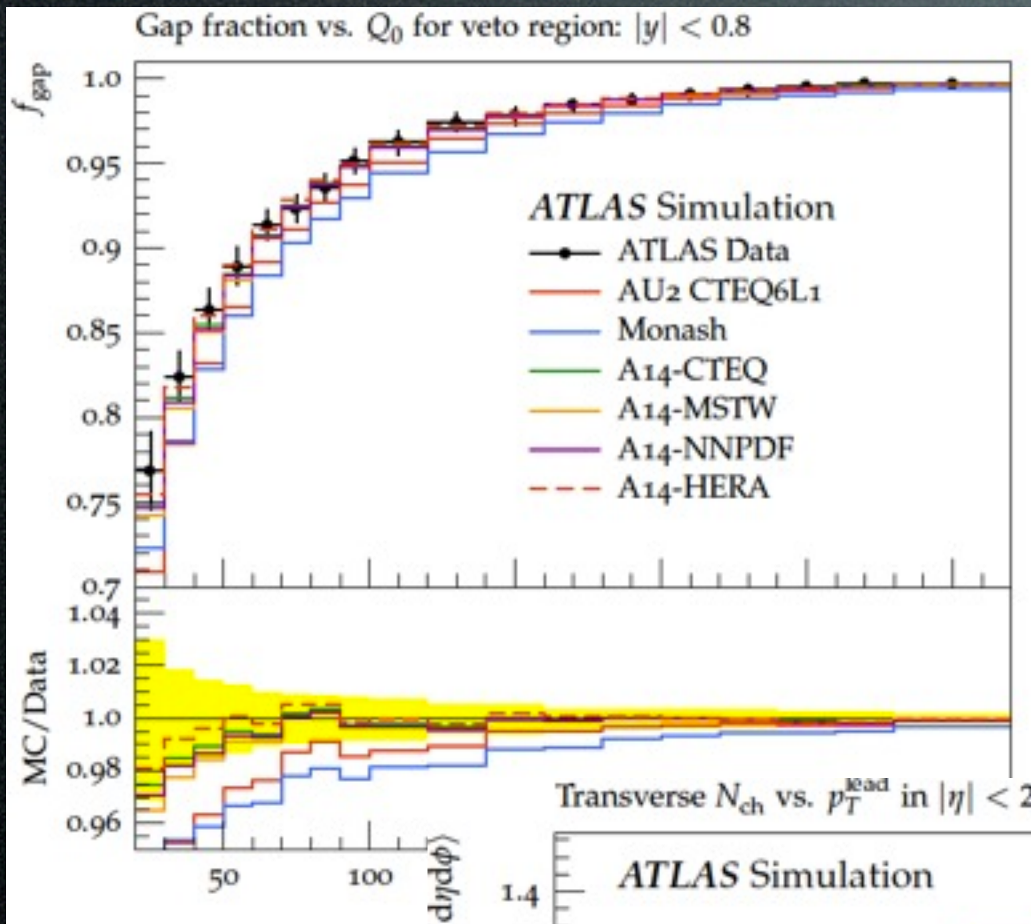
Parameter	Definition
`SigmaProcess:alphaSvalue`	The α_S value at scale $Q^2 = M_Z^2$
`SpaceShower:pT0Ref`	ISR p_T cutoff
`SpaceShower:pTmaxFudge`	Mult. factor on max ISR evolution scale
`SpaceShower:pTdampFudge`	Factorisation/renorm scale damping
`SpaceShower:alphaSvalue`	ISR α_S
`TimeShower:alphaSvalue`	FSR α_S
`BeamRemnants:primordialKThard`	Hard interaction primordial $k_\perp$
`MultipartonInteractions:pT0Ref`	MPI p_T cutoff
`MultipartonInteractions:alphaSvalue`	MPI α_S
`BeamRemnants:reconnectRange`	CR strength

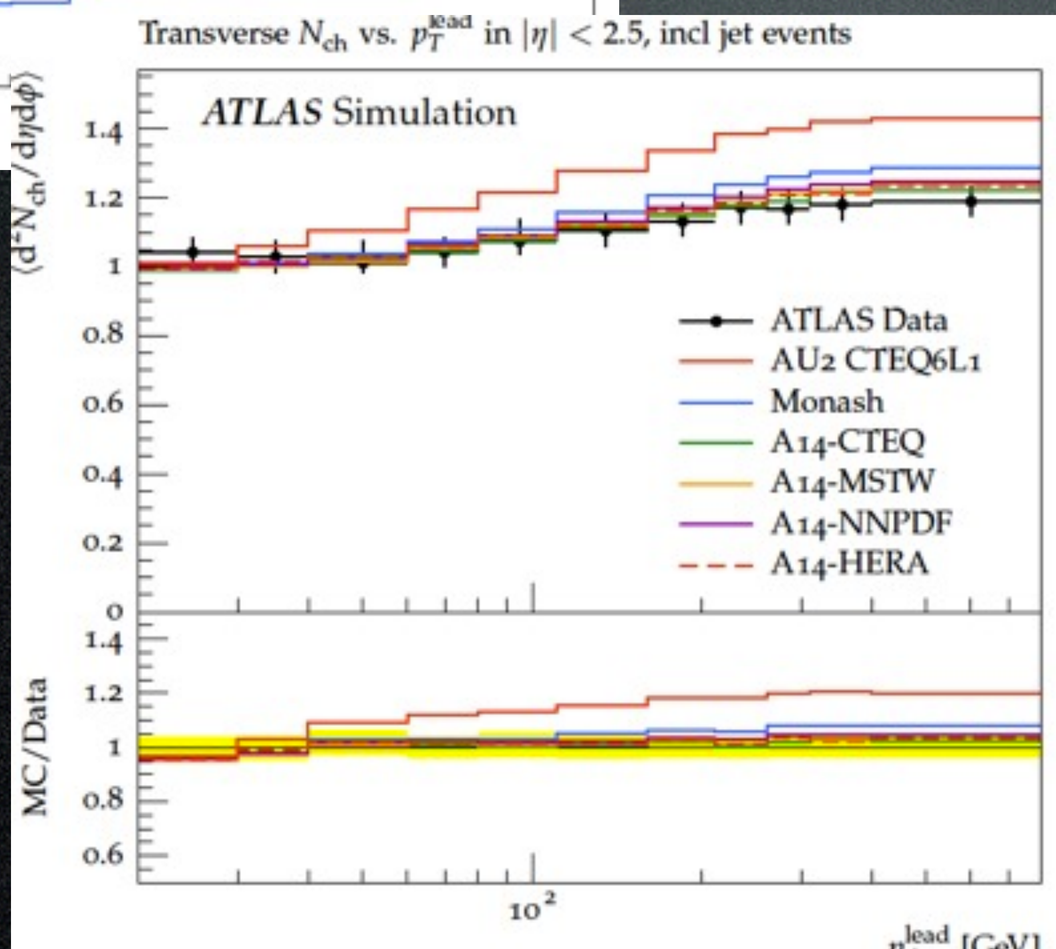
Used exponential matter overlap

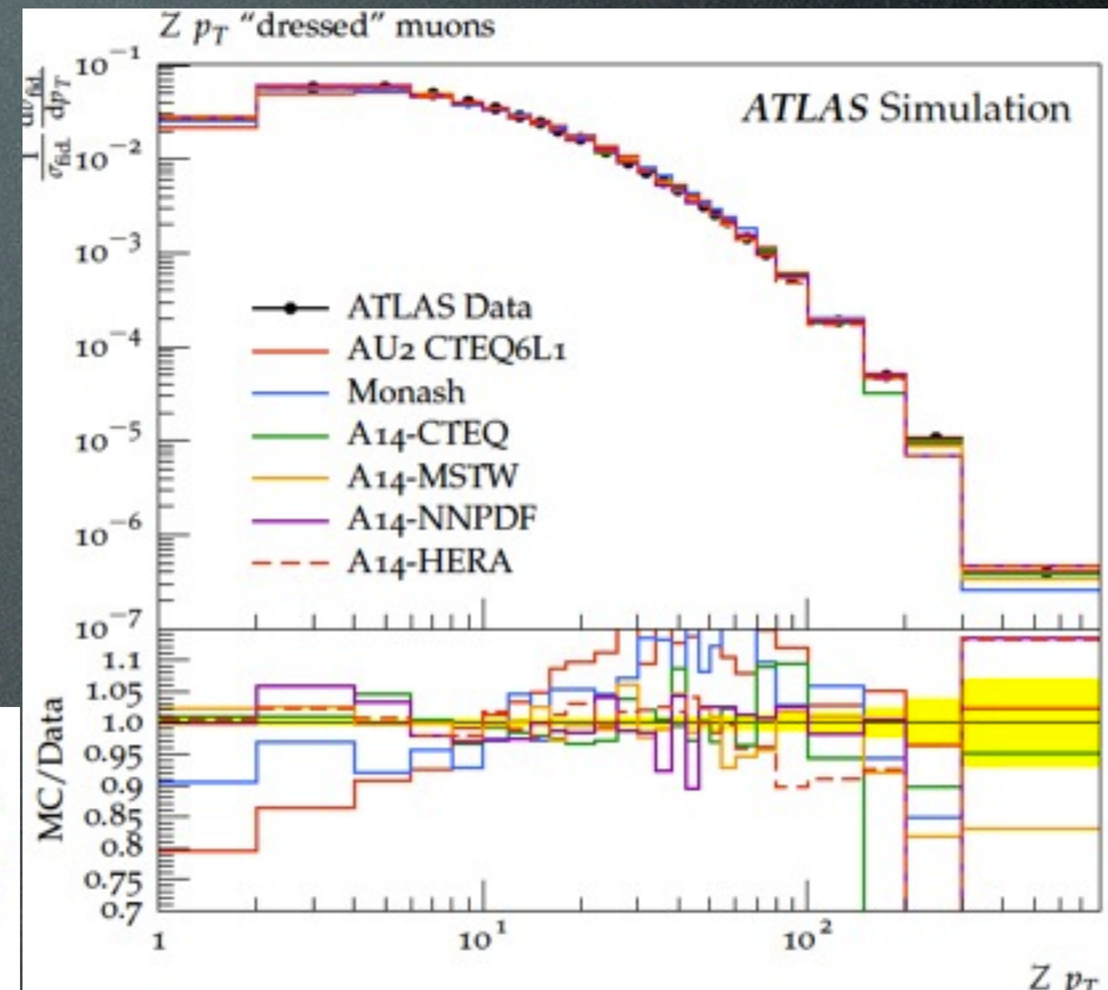
Also:
MulipartonInteractions:ecmPow



Tuning Shower & MPI Together



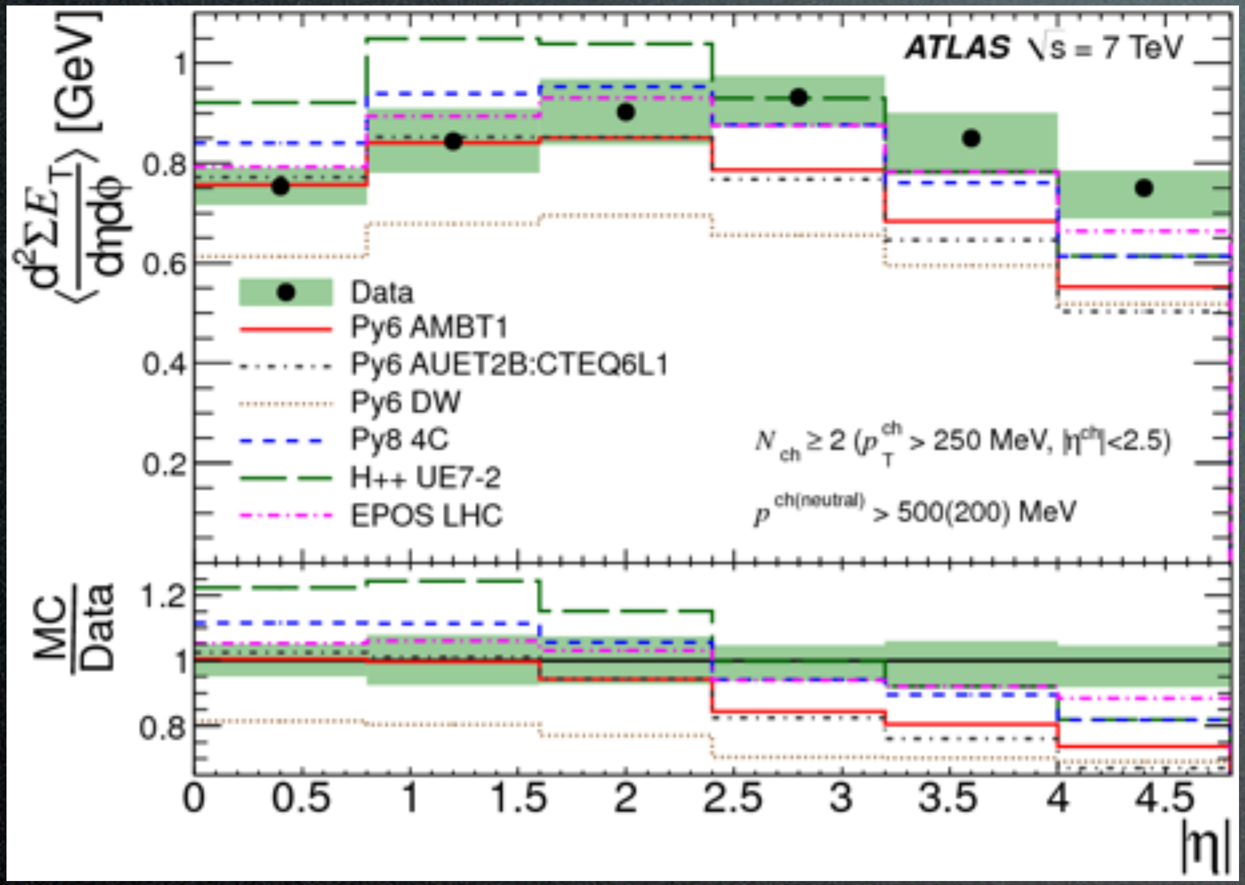


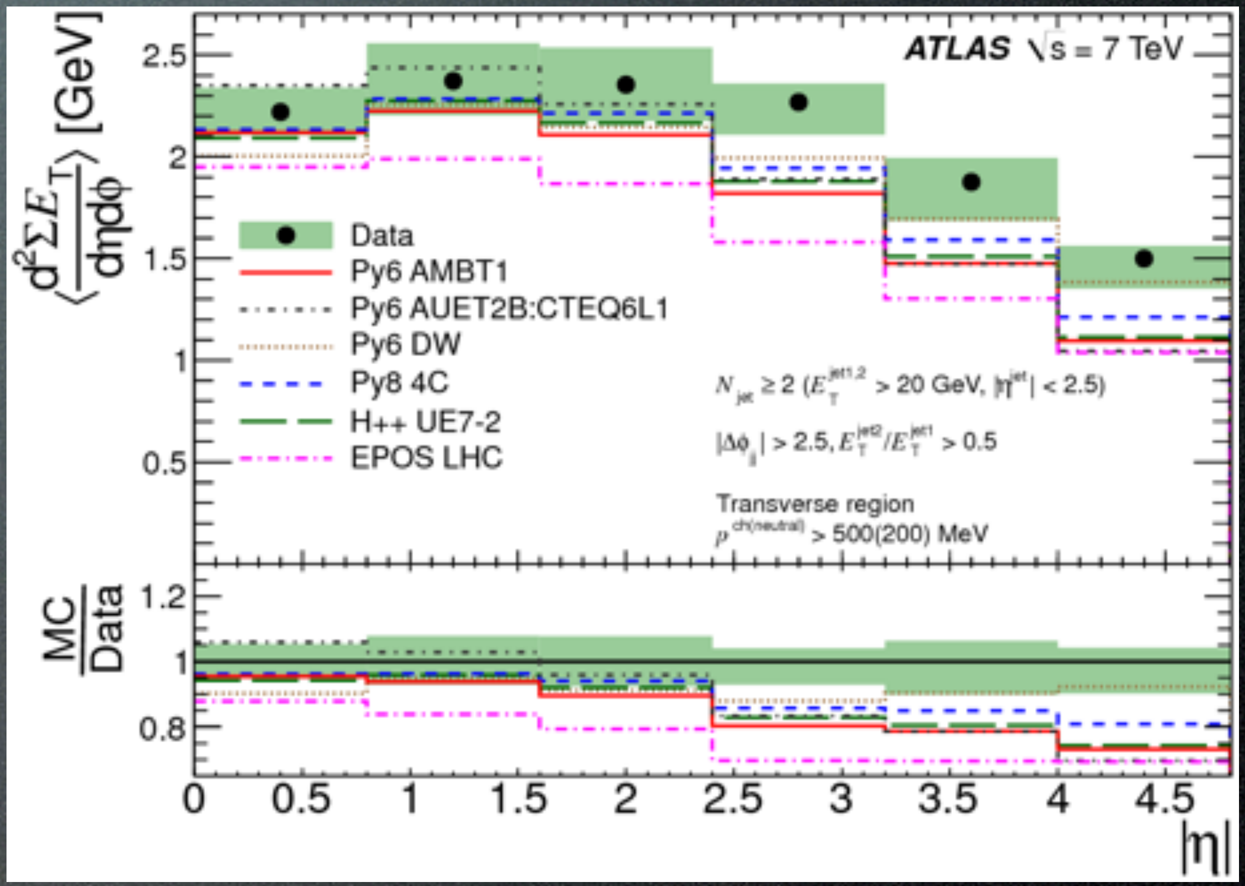


New A14 tune!
(with systematic variation tunes)

Extend to matched setups...

UE-sensitive Observables

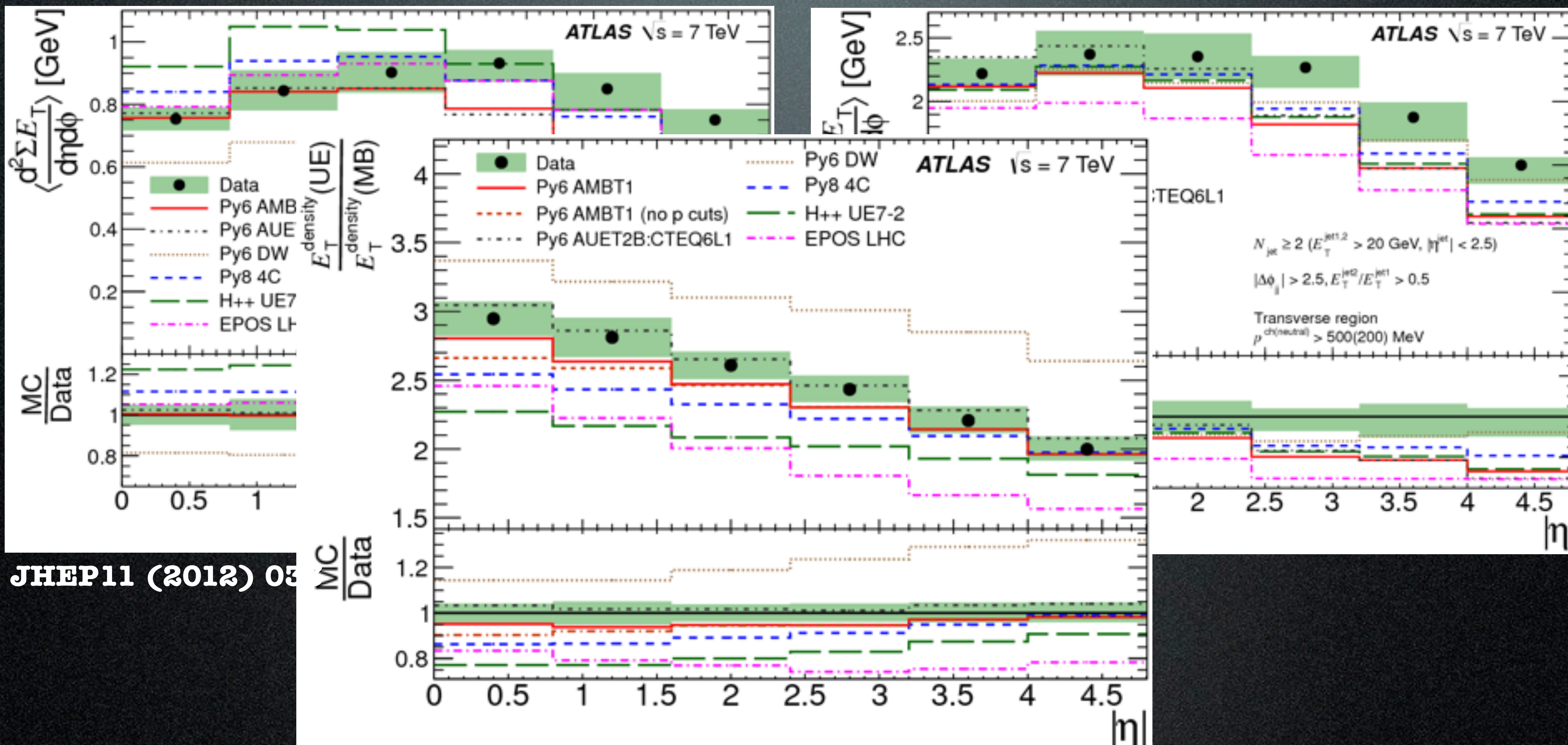




JHEP11 (2012) 033

Transverse energy flow: all models bad in forward region

UE-sensitive Observables

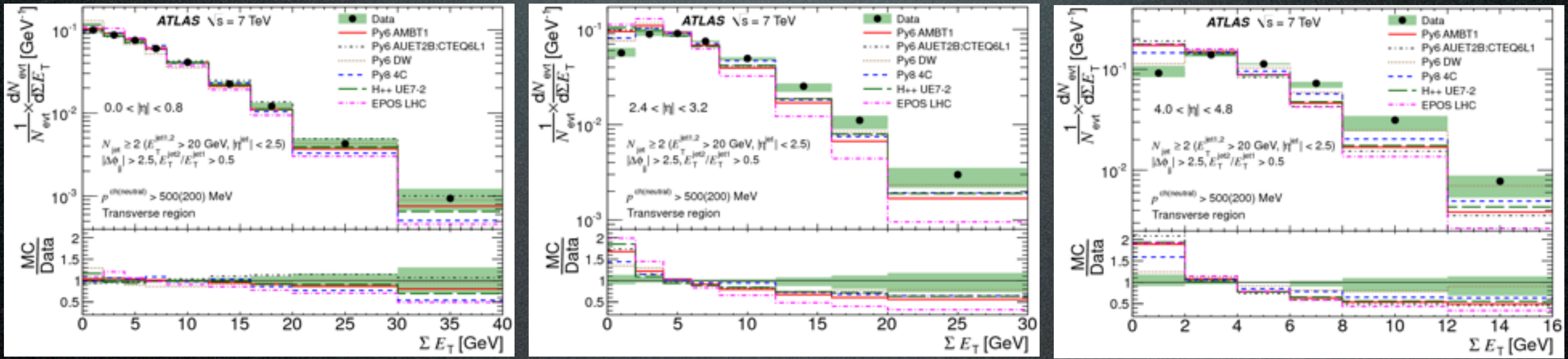


JHEP11 (2012) 03

More energy in dijet events!

From Central to Forward

low η → **high η**

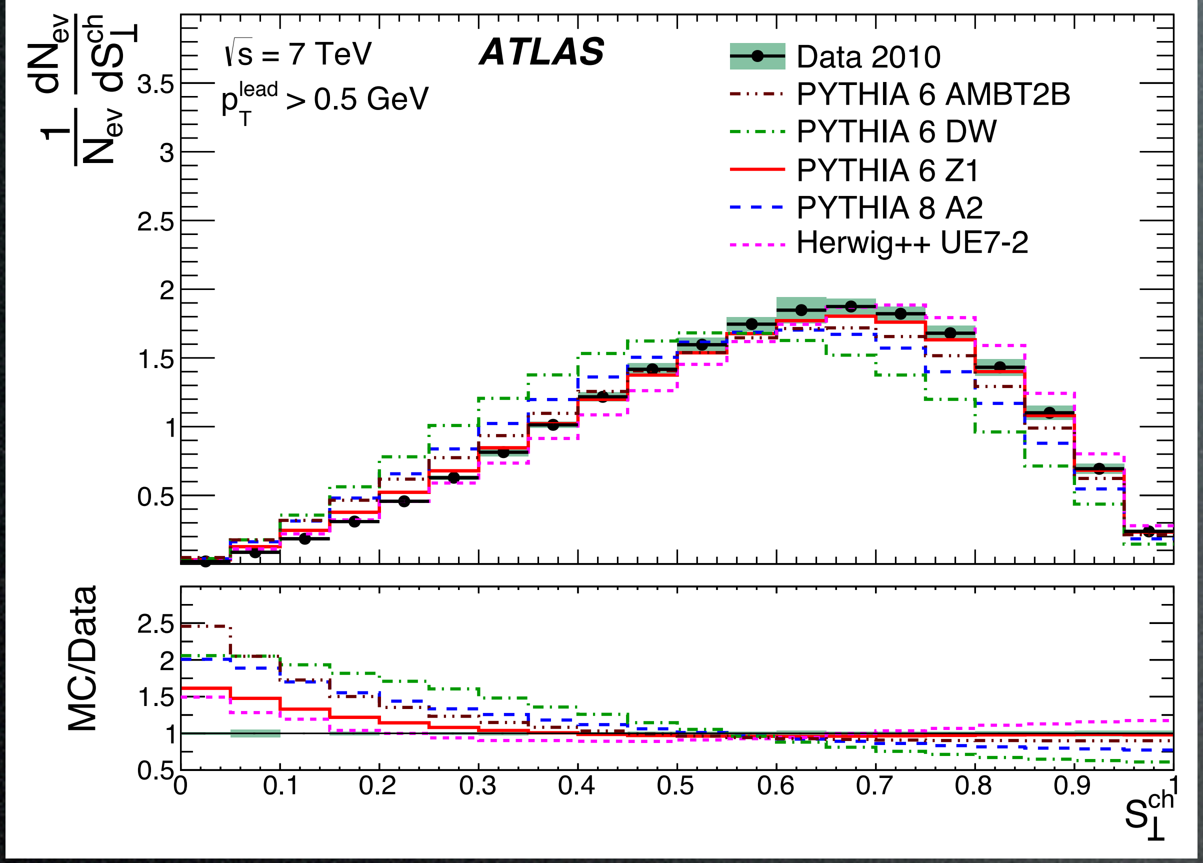


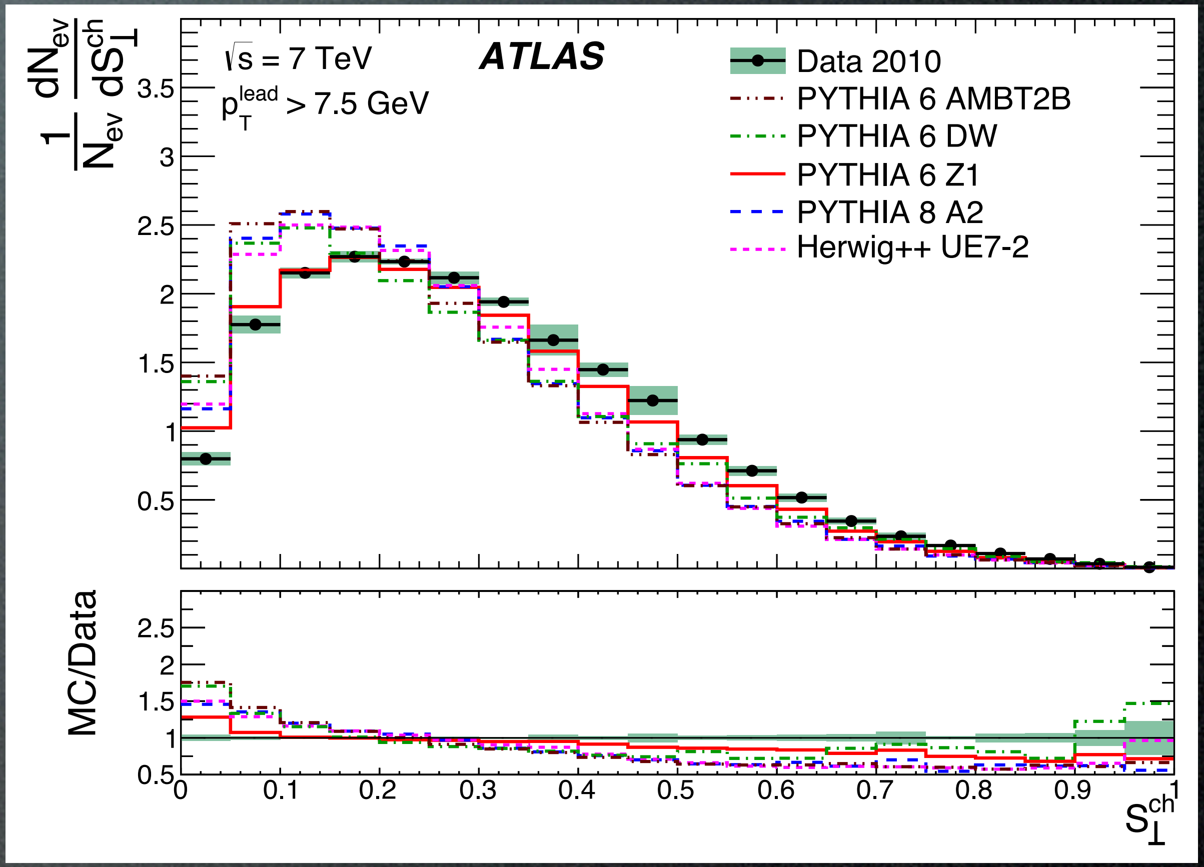
JHEP11 (2012) 033

UE tunes do better overall

Event Shapes

Low lead p_T → High lead p_T

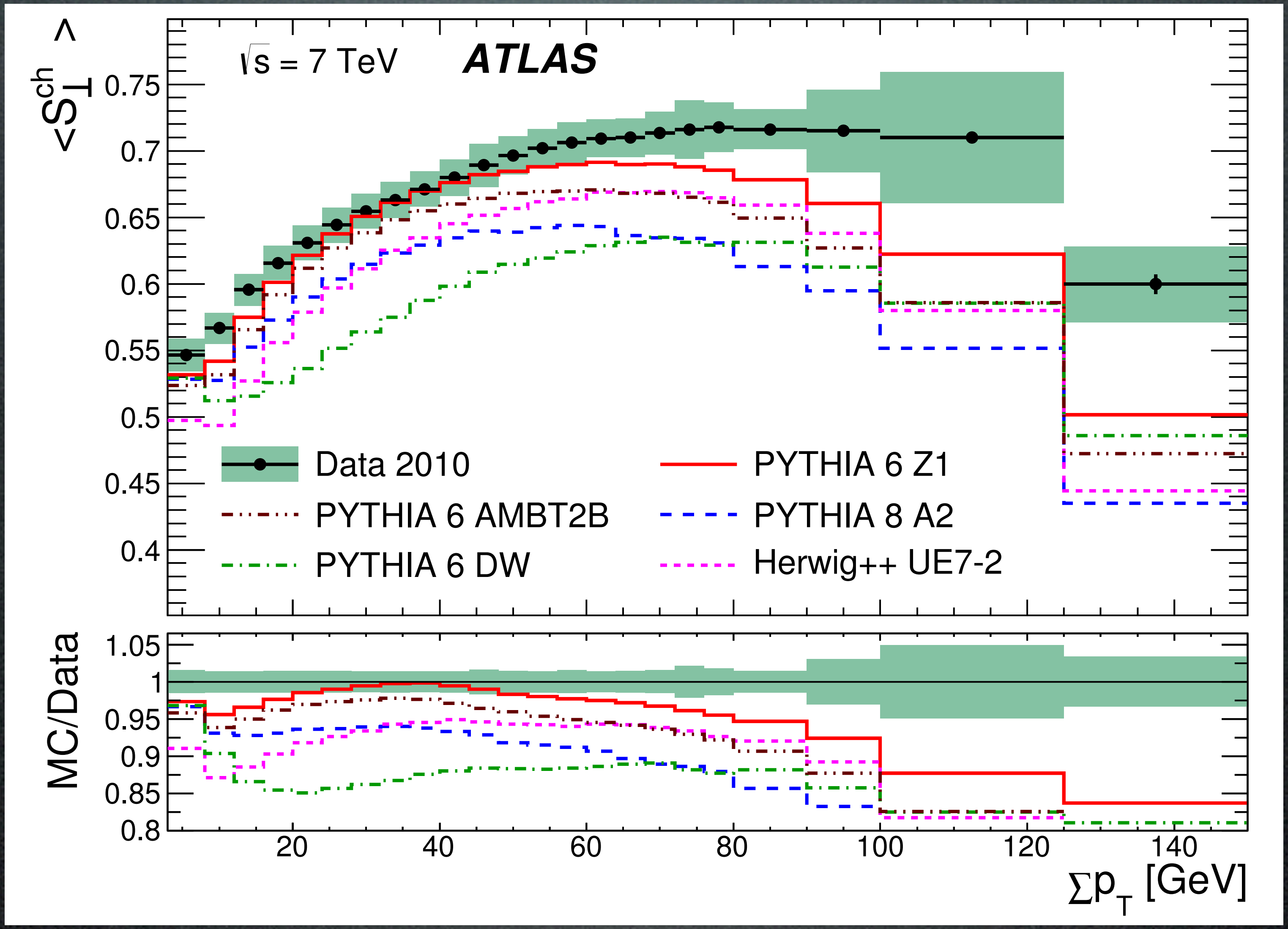




Phys. Rev. D 88, 032004 (2013)

UE starts taking over....

Event Shape Profile

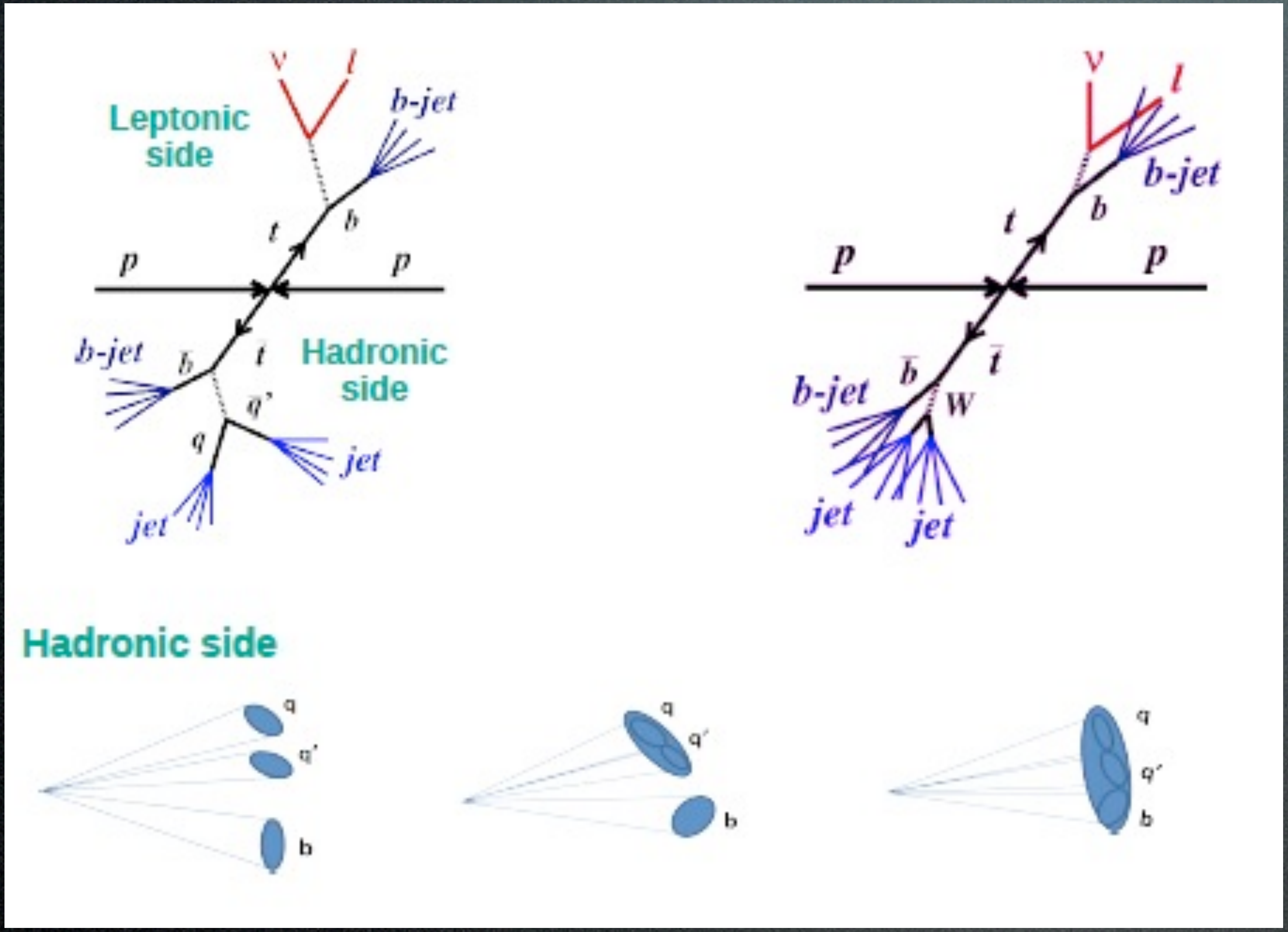


Phys. Rev. D 88, 032004 (2013)

Emergence of jets?

Jet Substructure in Brief

Large radius jets



The angular resolution of the decay products: $\Delta R \approx 2m/p_T$

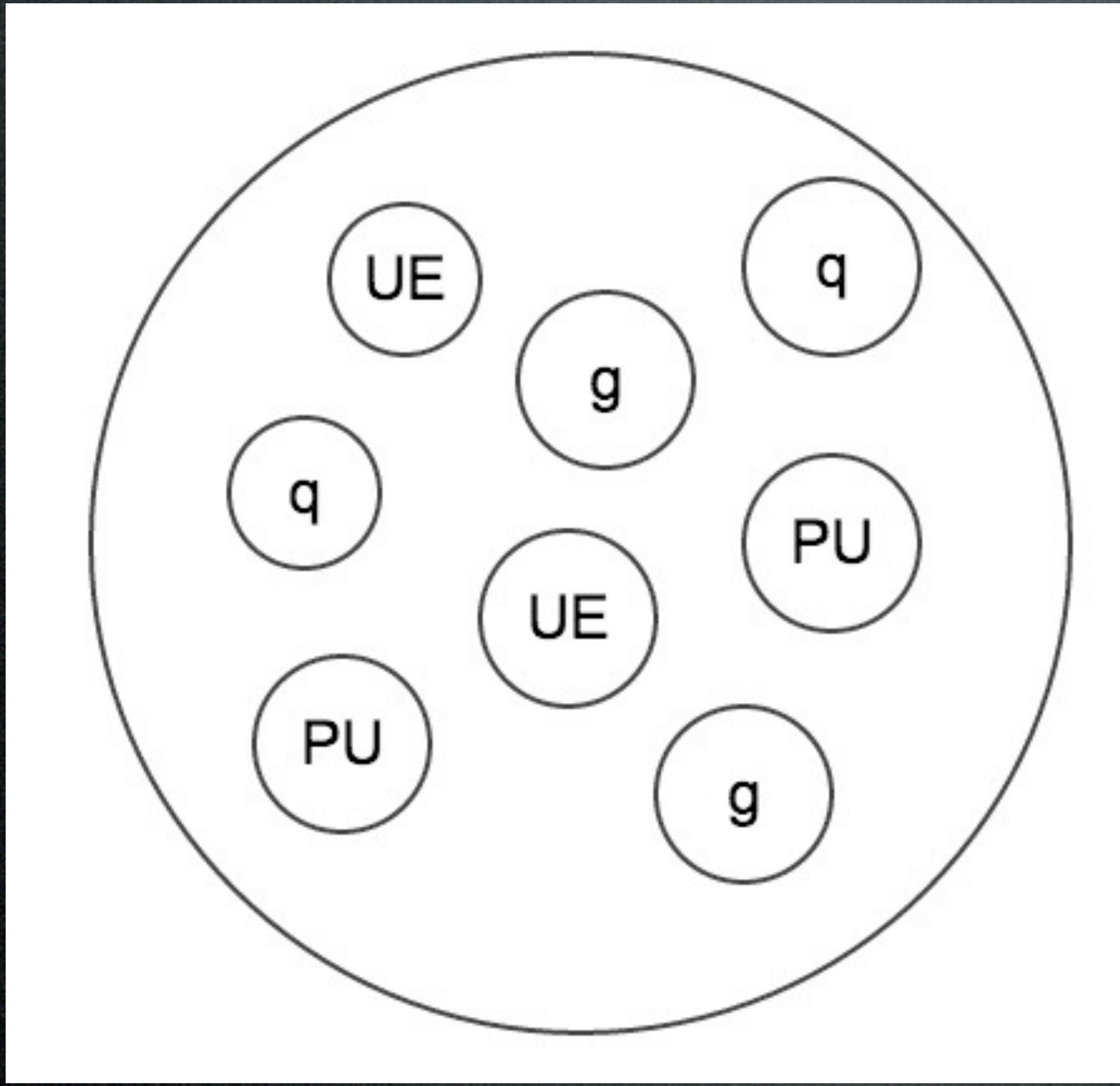
So for a top quark (of mass 173 GeV) with $p_T > 350$ GeV, we have $\Delta R \sim 1$.

With increasing c.m energy: collimated decay products from boosted heavy particles result in a single massive jet.

So when you take apart a jet, what does it look like?

So when you take apart a jet, what does it look like?

So when you take apart a jet, what does it look like?



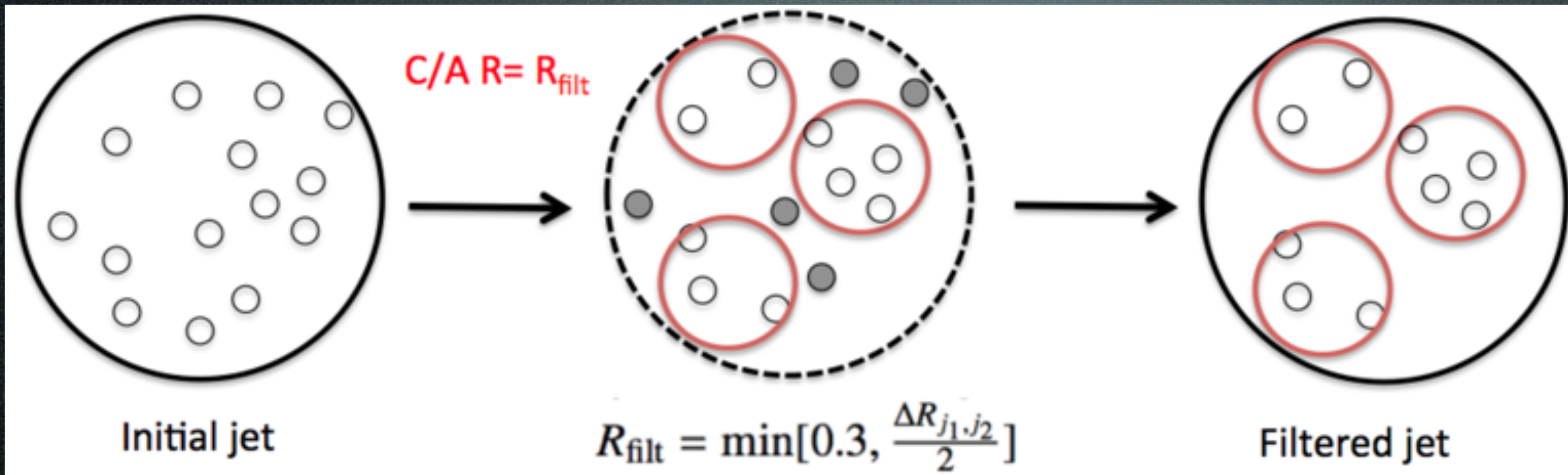
We want to exploit the "substructure" of the large-radius jet to identify original particles

Substructure Techniques

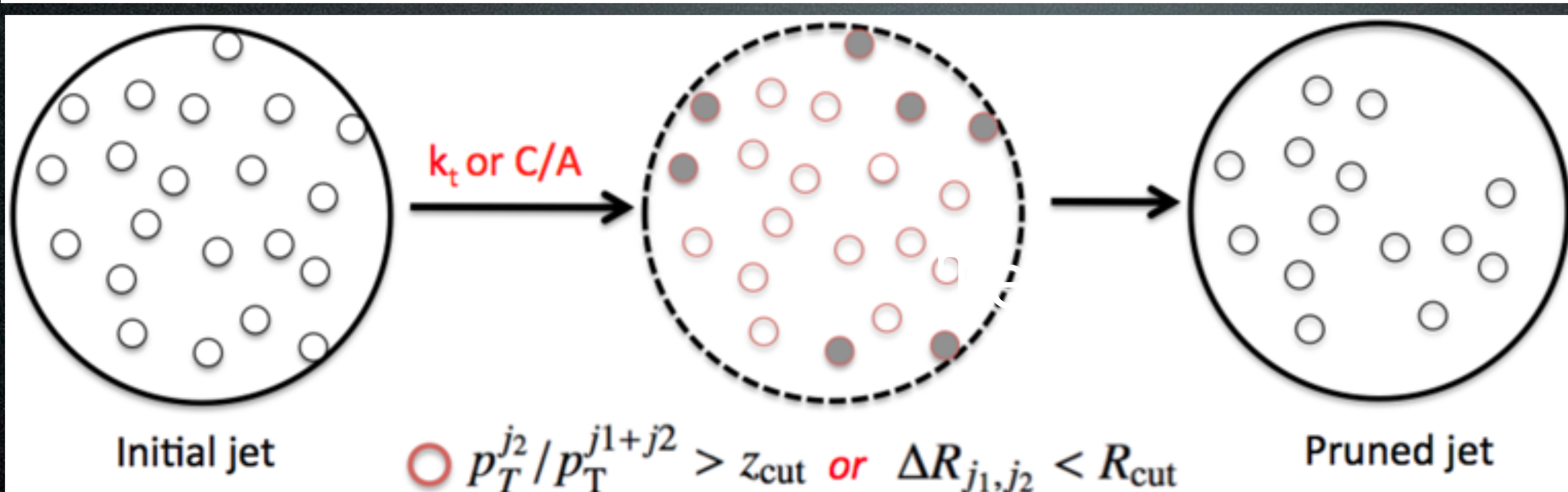
- Jets need to be “groomed”.

The large-radius jets not only include particles coming from the interesting decays, but also from pileup, underlying event ….

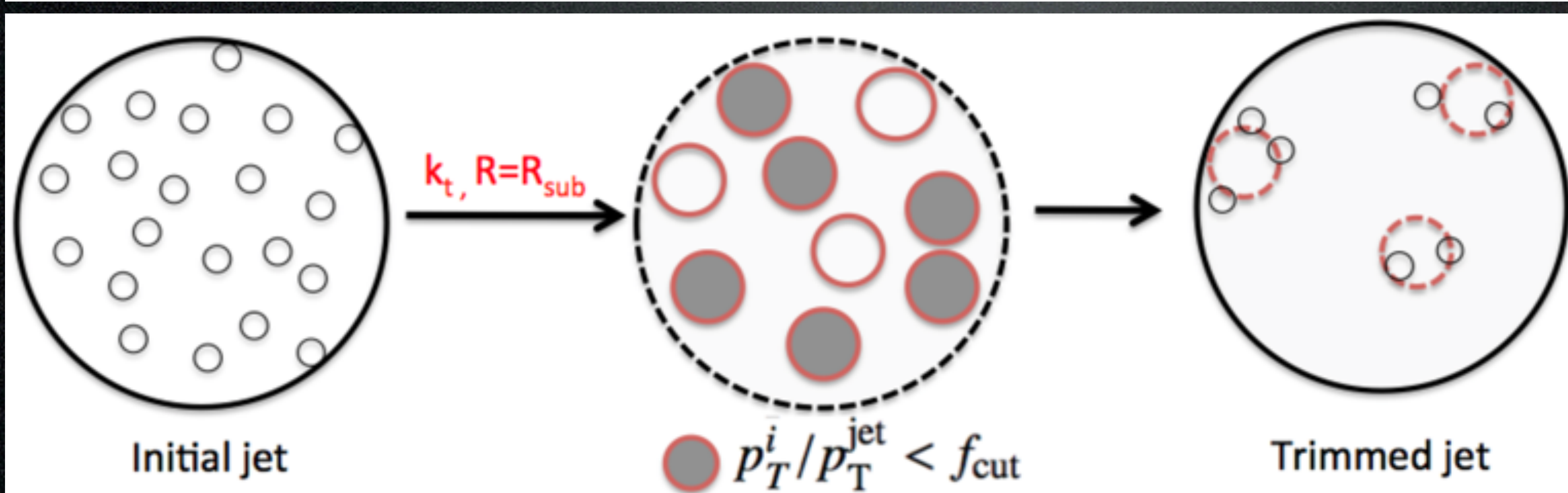
- Need observables which would be sensitive to signal-like or background-like nature of these jets.



Filtering

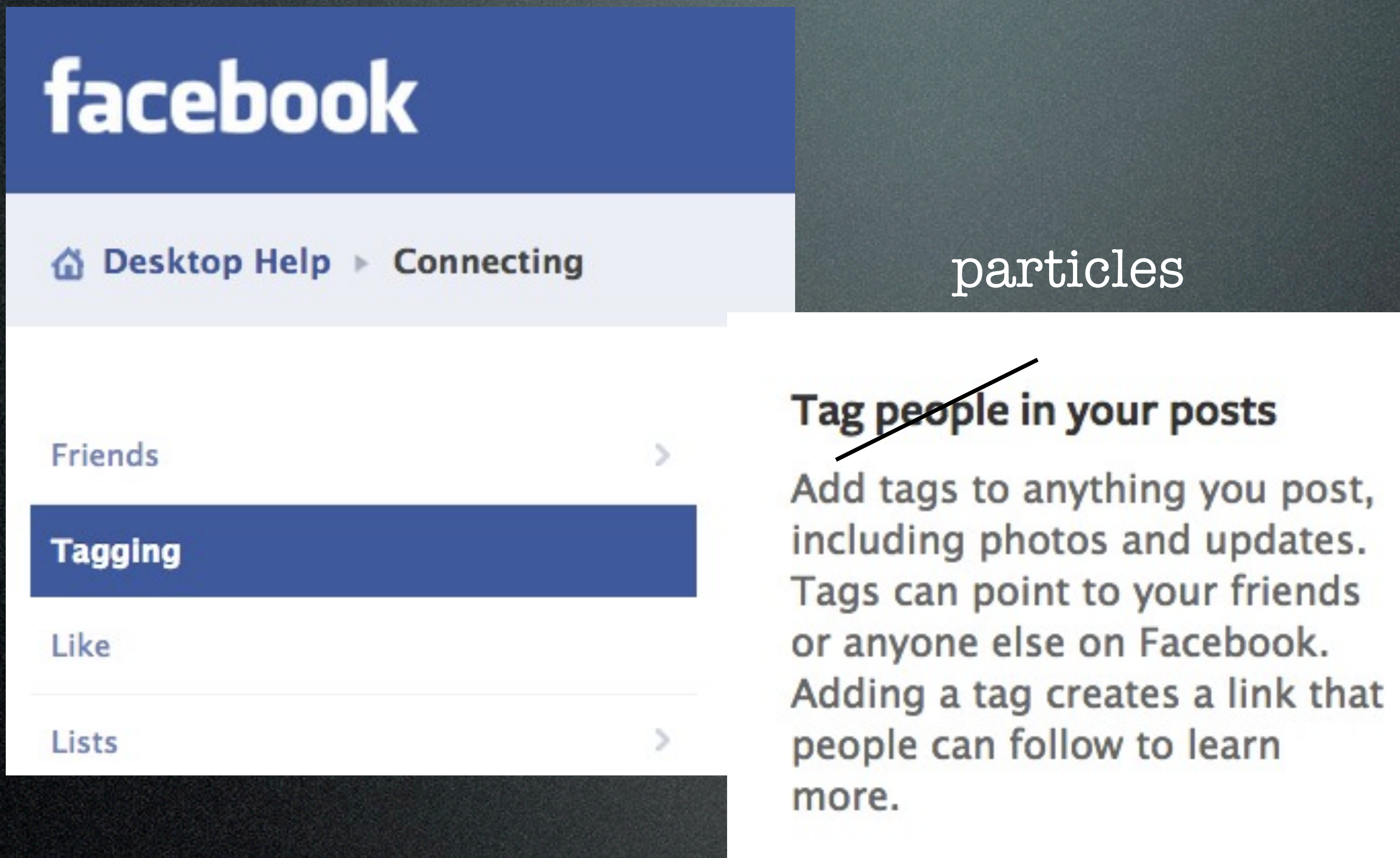


Prunning



Trimming

Tagging Top or Higgs

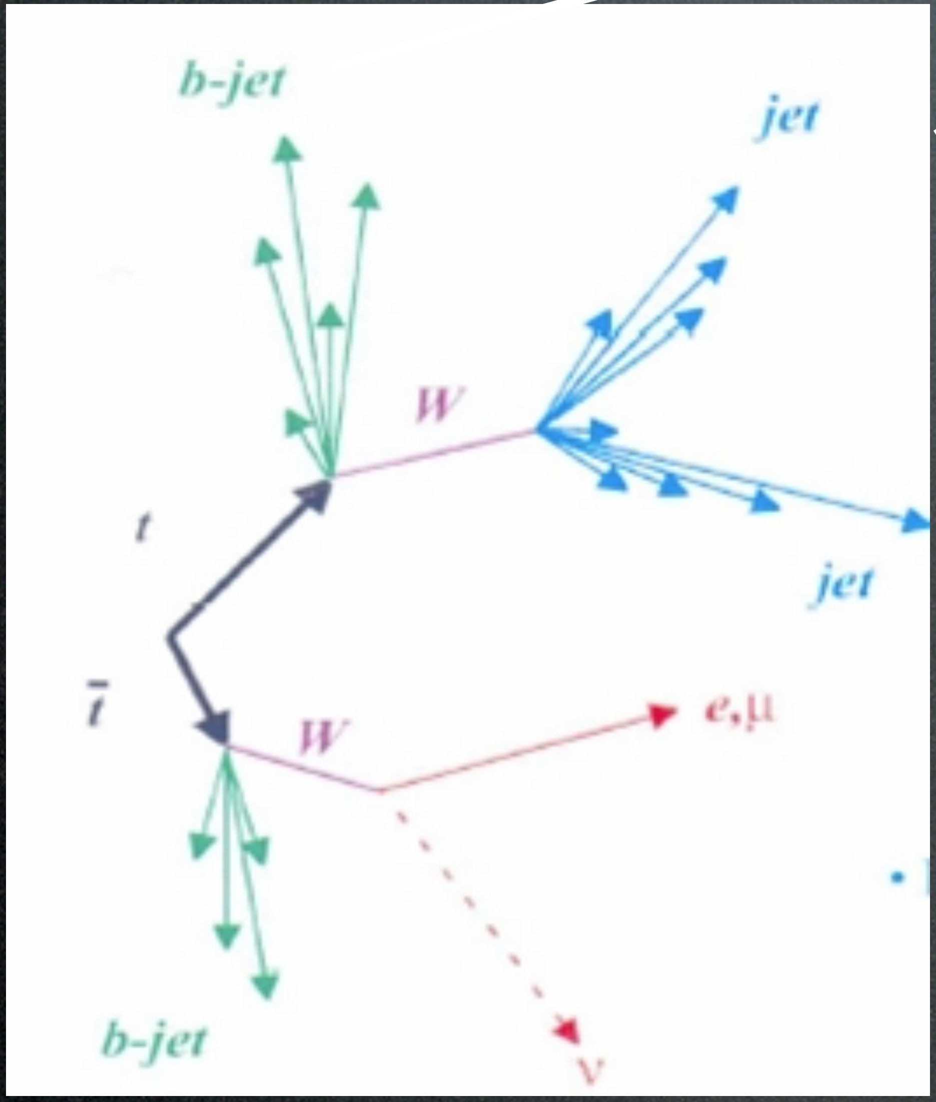


- Target is to identify jets resulting from the decay of top quark or Higgs against jets coming from light quark/gluons.

Playing with the Shower

Recall

Top quark decay:

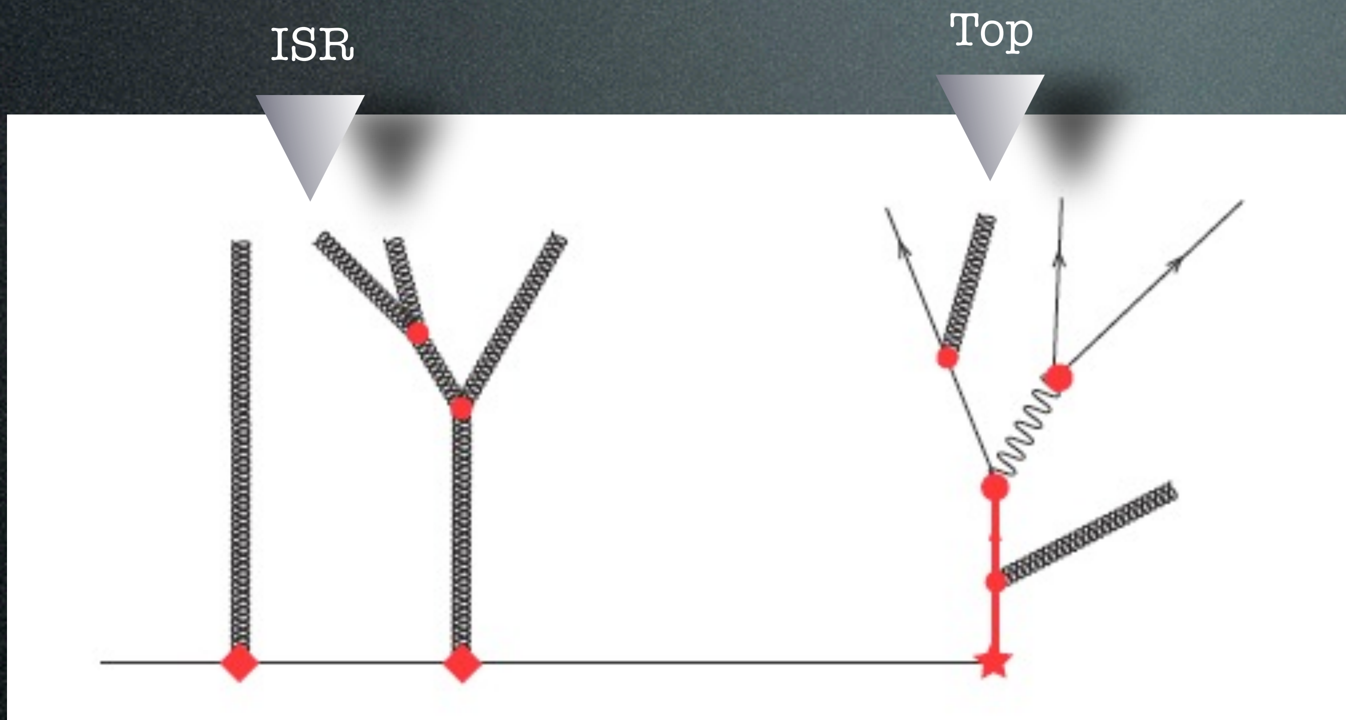


Parton Shower!

Reversing the Shower

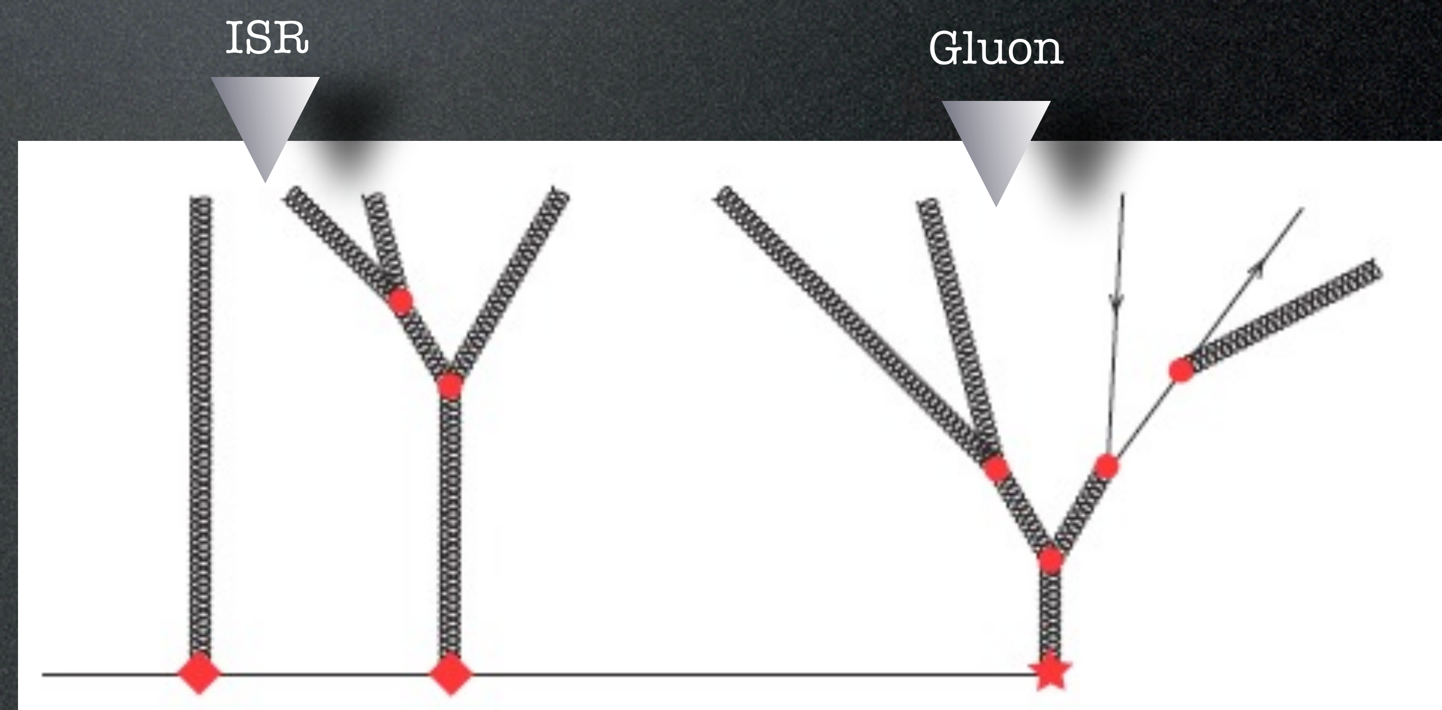
Davison E. Soper, Michael Spannowsky; arXiv:1102.3480, arXiv:1211.3140

Shower Deconstruction



Top quark jet shower history

vs.

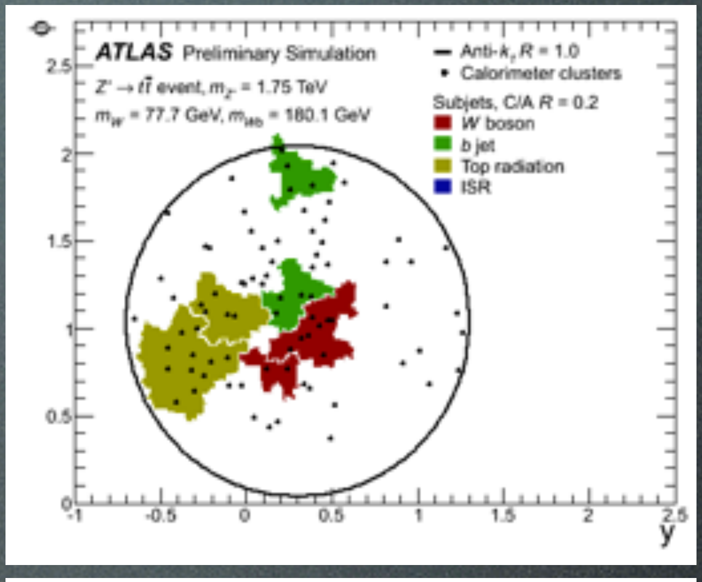


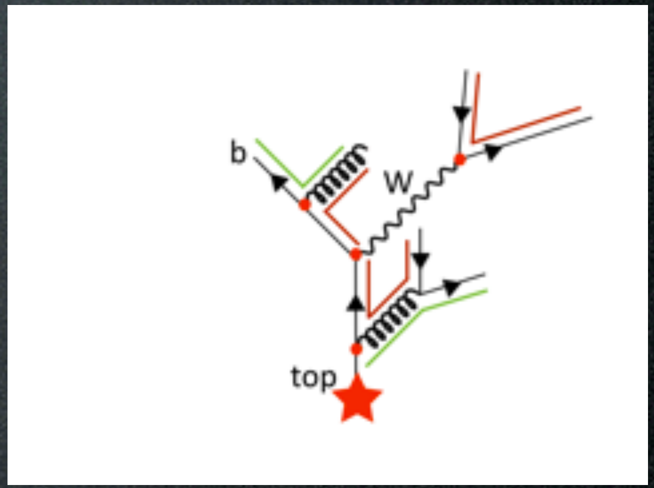
Light quark jet shower history

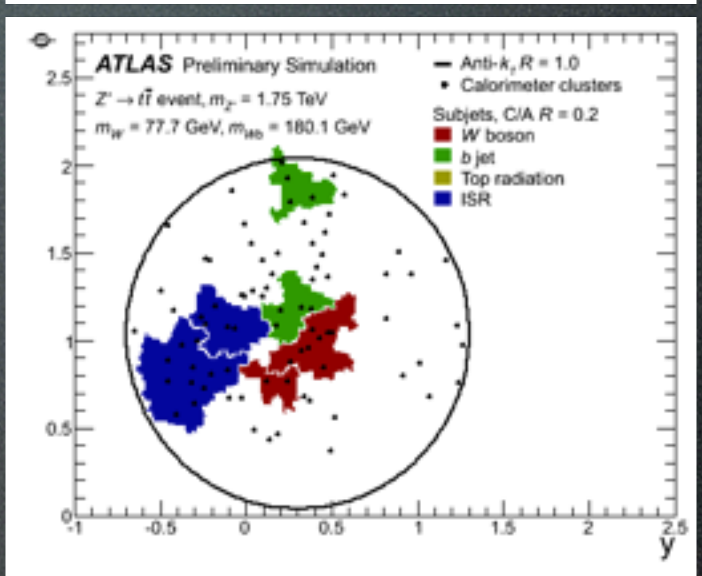
Shower Deconstruction

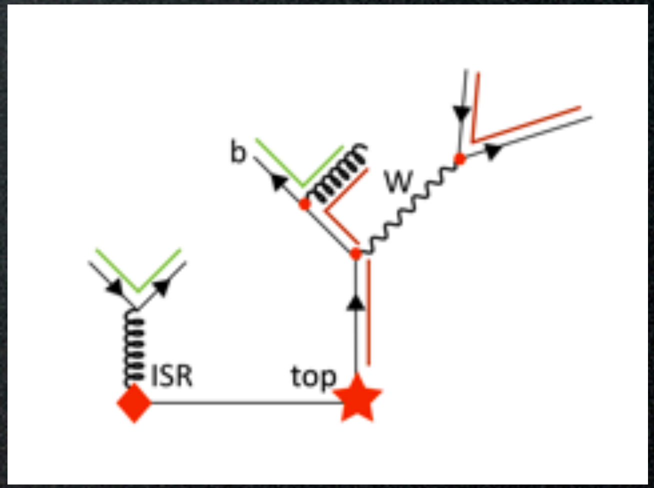
- Decompose the large-radius jet into small radius **sub/microjets.**
- Build all possible shower histories with the microjets.
- Assign probability whether signal-like or background-like.
- A single analytic function:

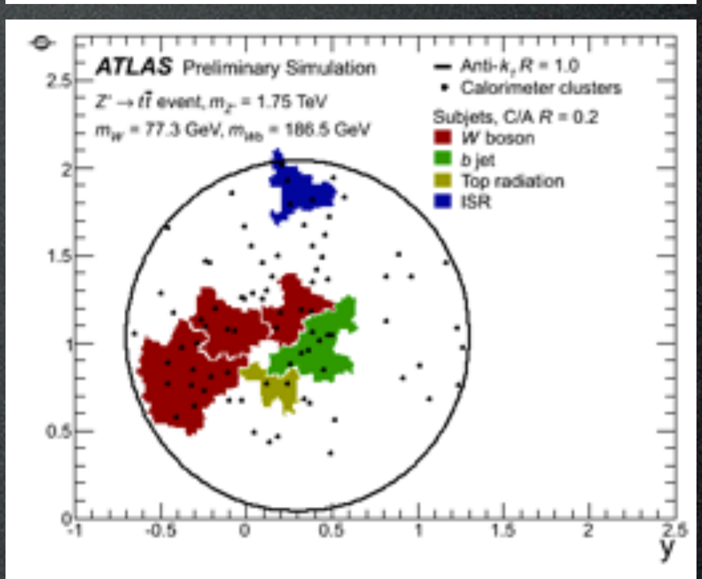
$$\chi(\{p\}_N) = \frac{P(\{p\}_N|\mathrm{S})}{P(\{p\}_N|\mathrm{B})}$$

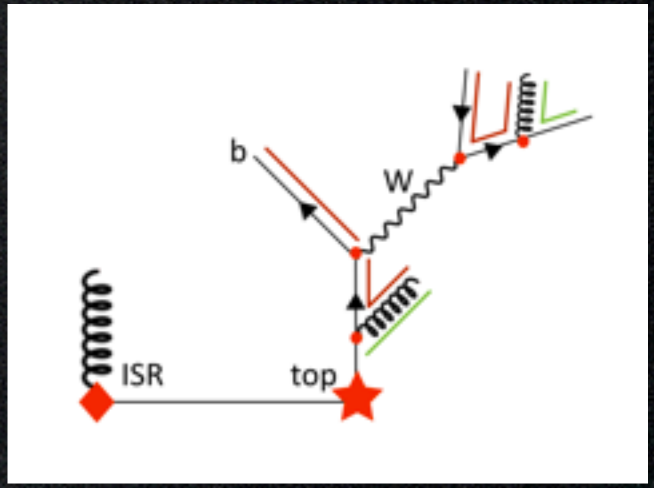






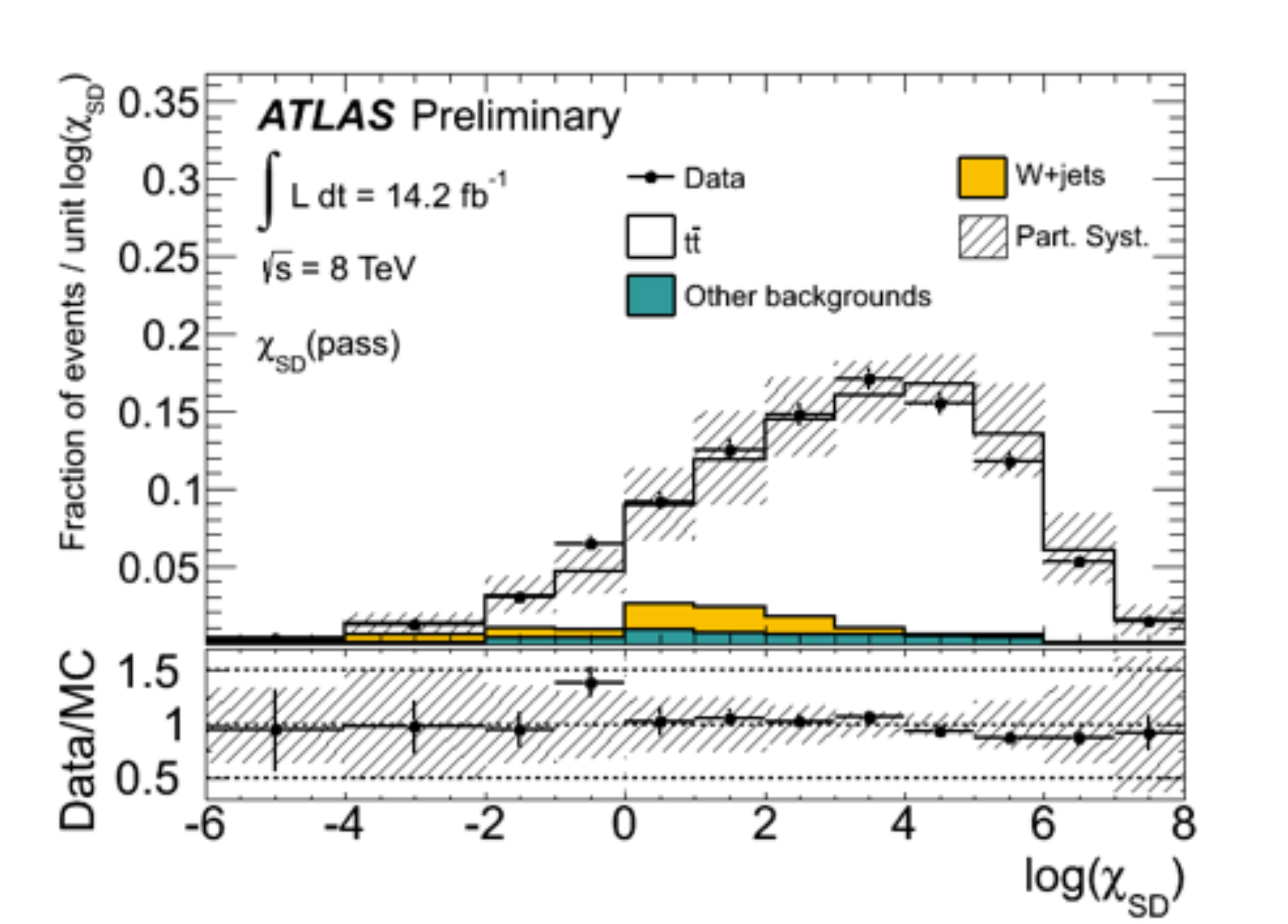




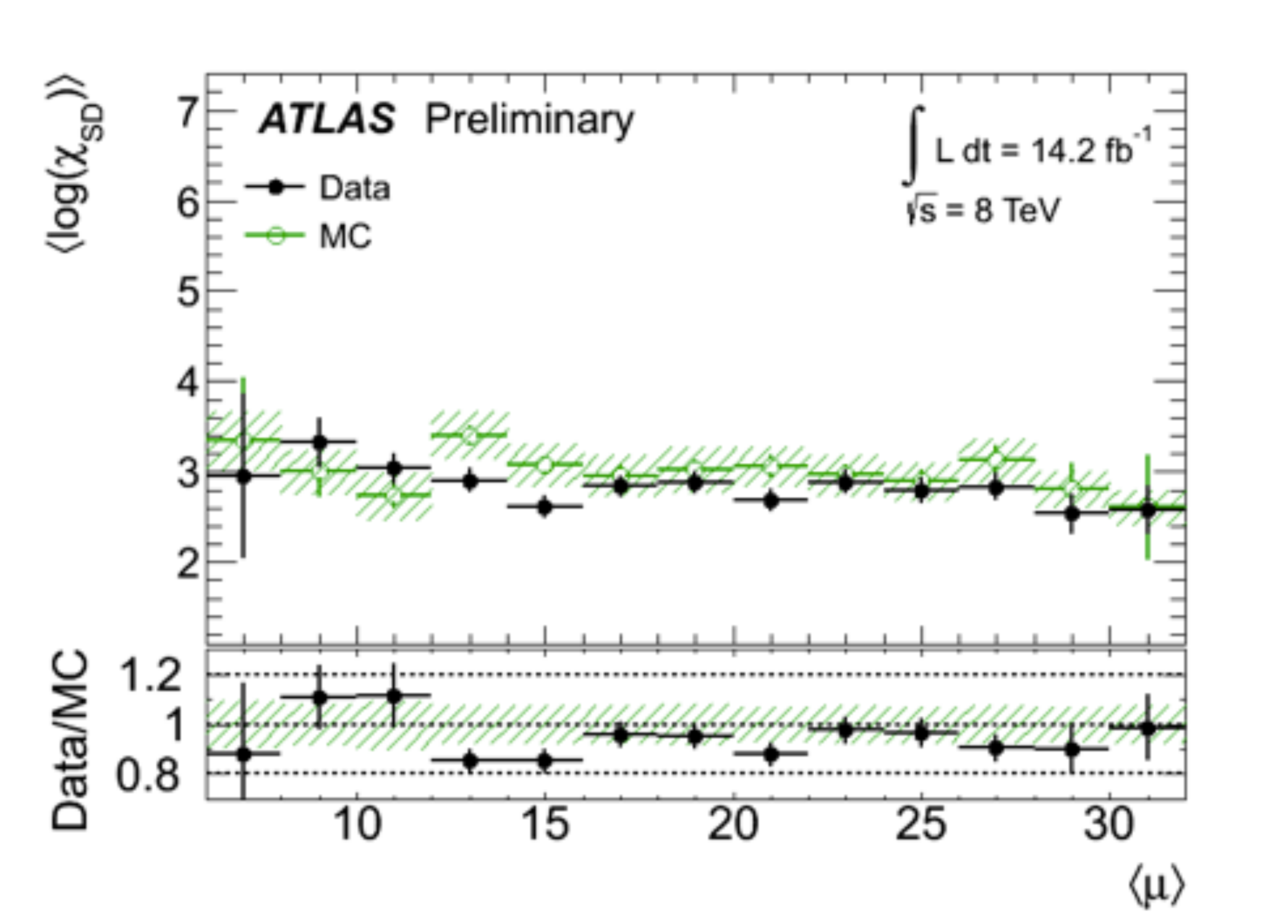


ATLAS-CONF-2014-003

Looking at our Data



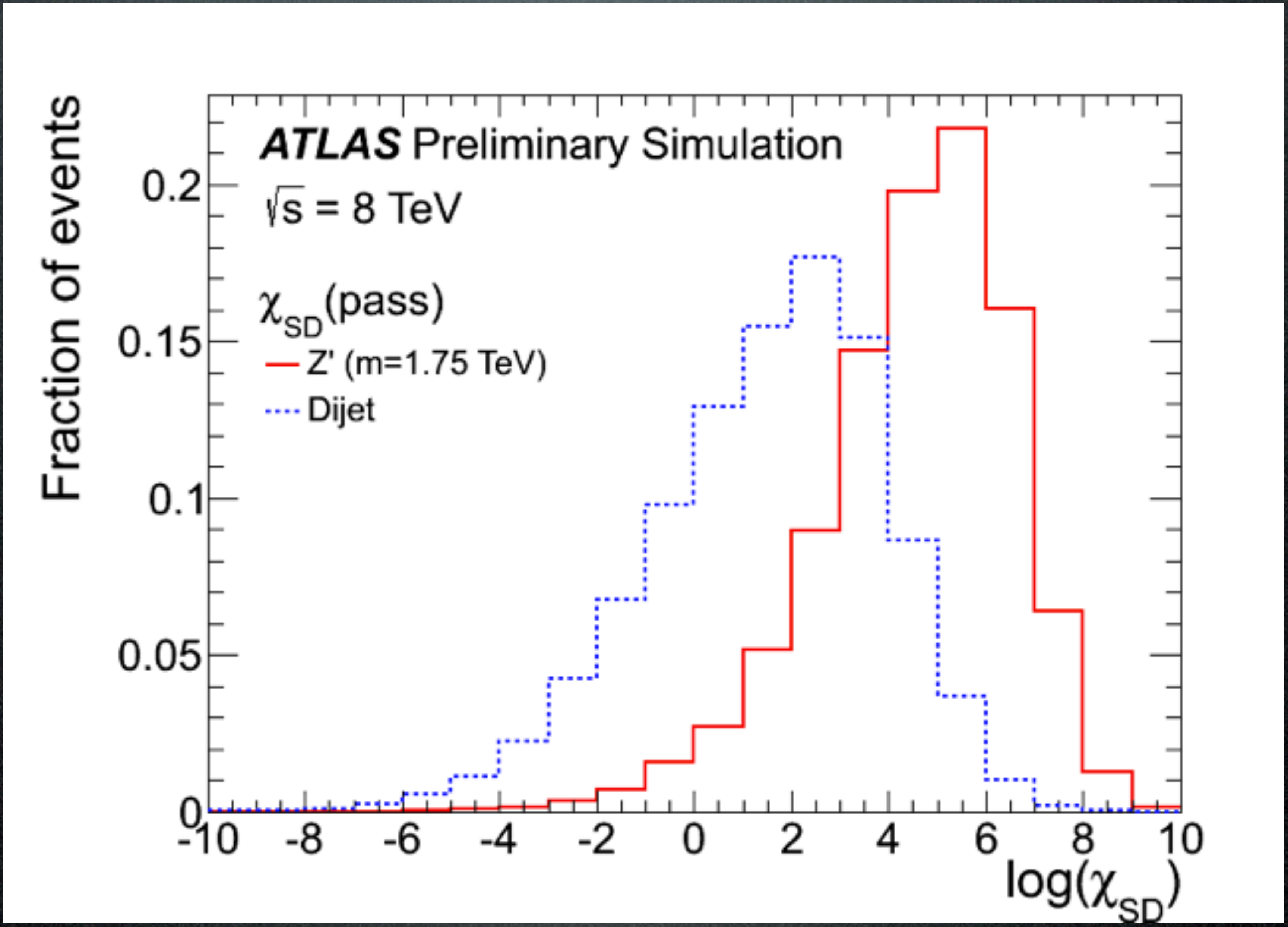
LogChi modelled well by MC



LogChi robust against pileup

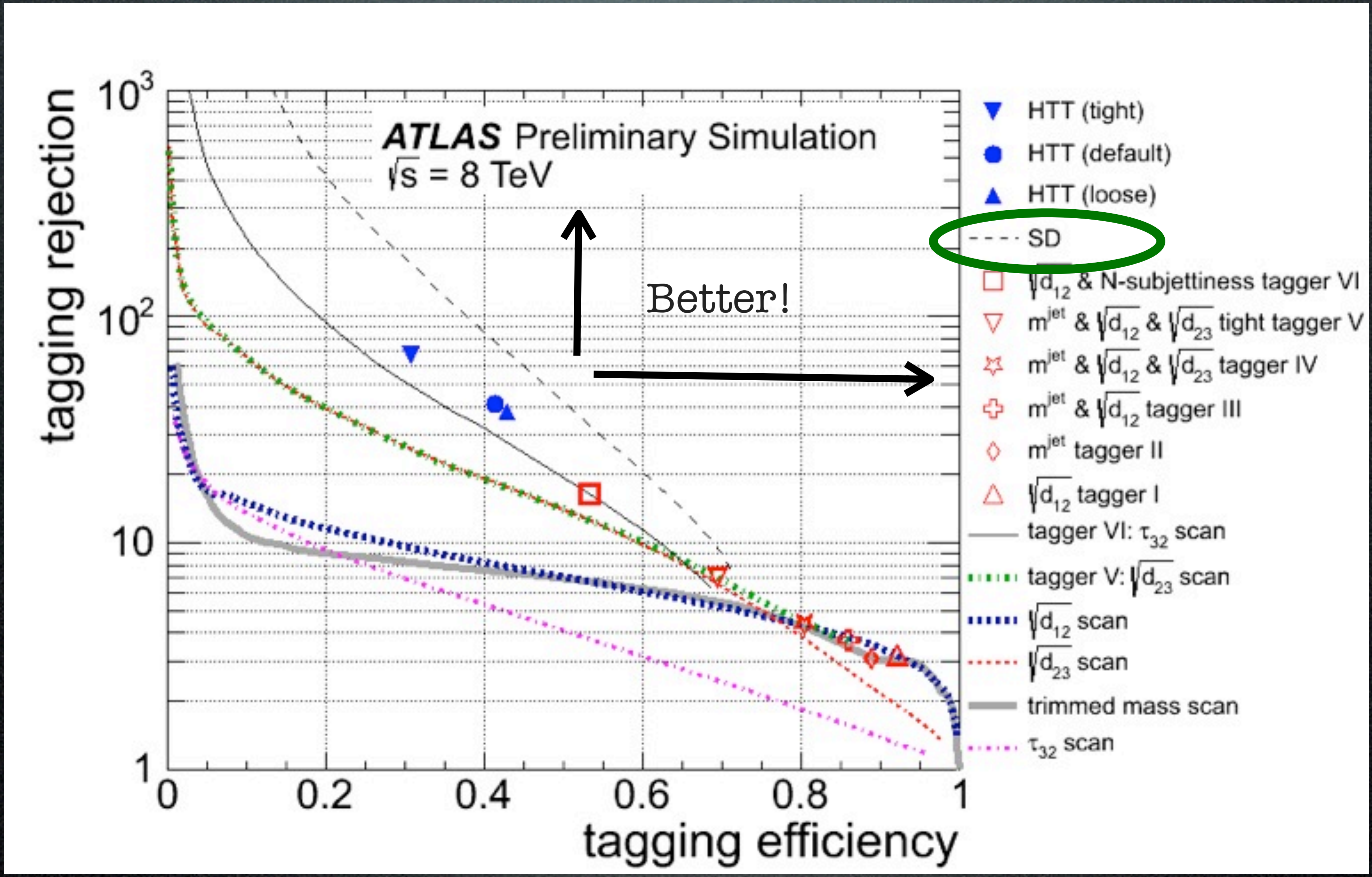
ATLAS-CONF-2014-003

Signal and BG Discrimination



ATLAS-CONF-2014-003

Top-Tagging Comparison



Better top quark finding efficiency at the same rejection of multijets when compared to the HEPTopTagger.

ATLAS-CONF-2014-003

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.
- Jet substructure techniques utilise the knowledge and modelling of shower, so direct application in searches as well.

Supporting Material

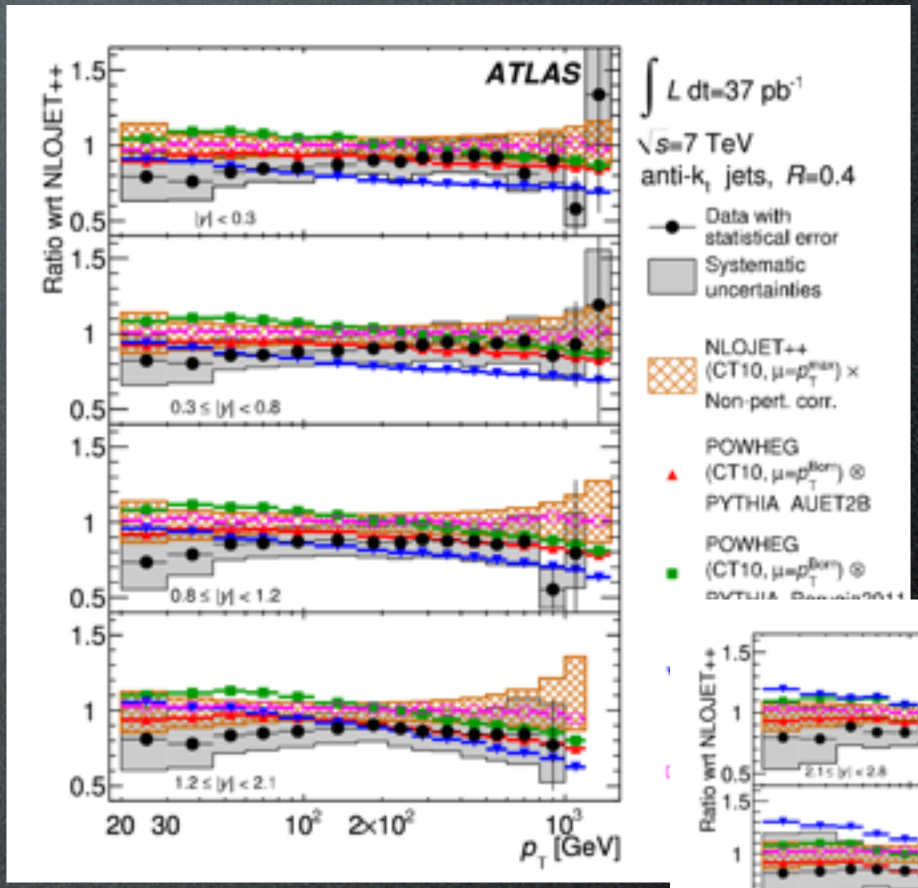
Powheg+Pythia8 Matching

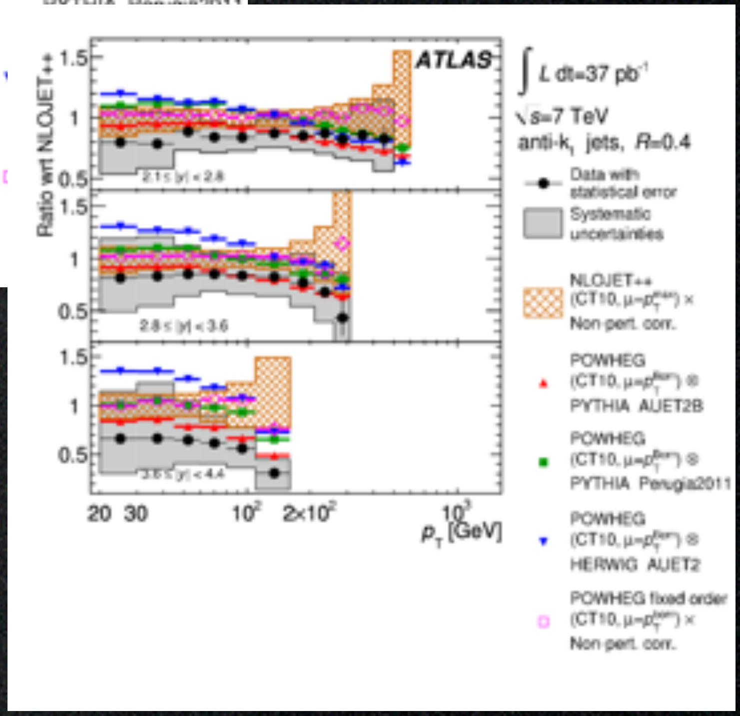
(ongoing Les Houches Study)

- PoWHEG provides a scale (SCALUP) that is an indication of where the shower should take over from the perturbative calculation.

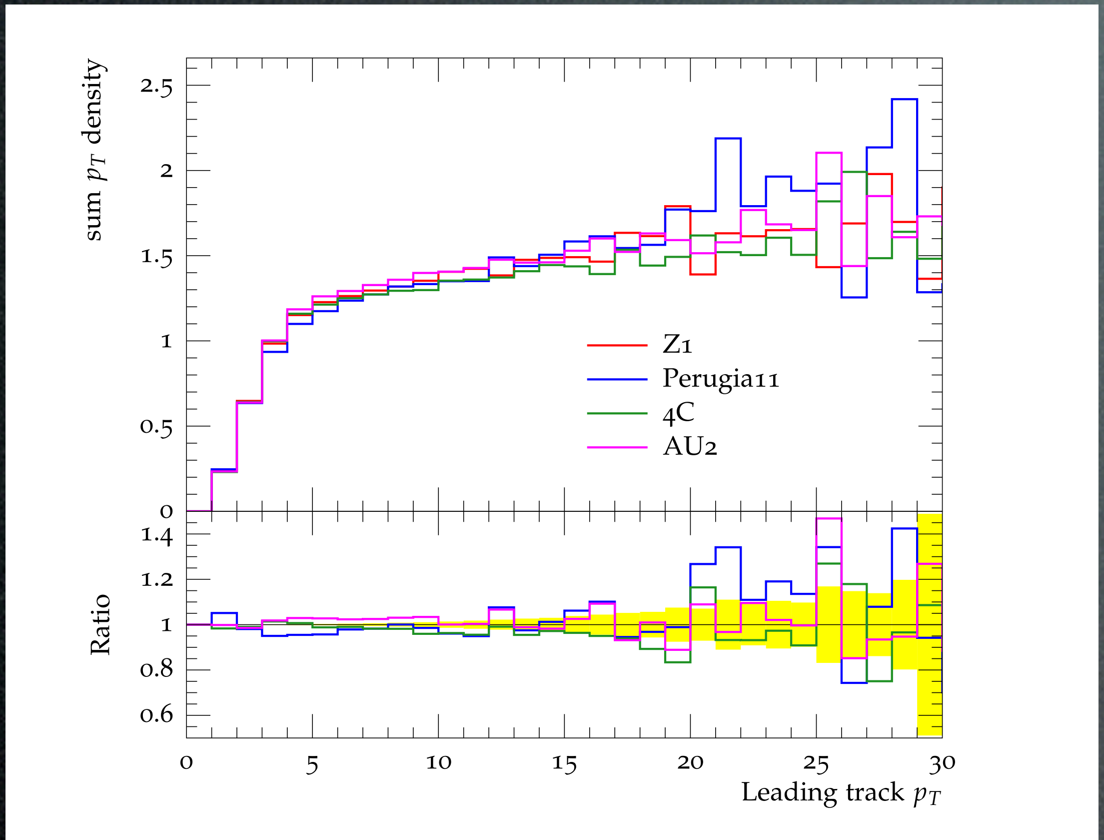
- What should be this scale?

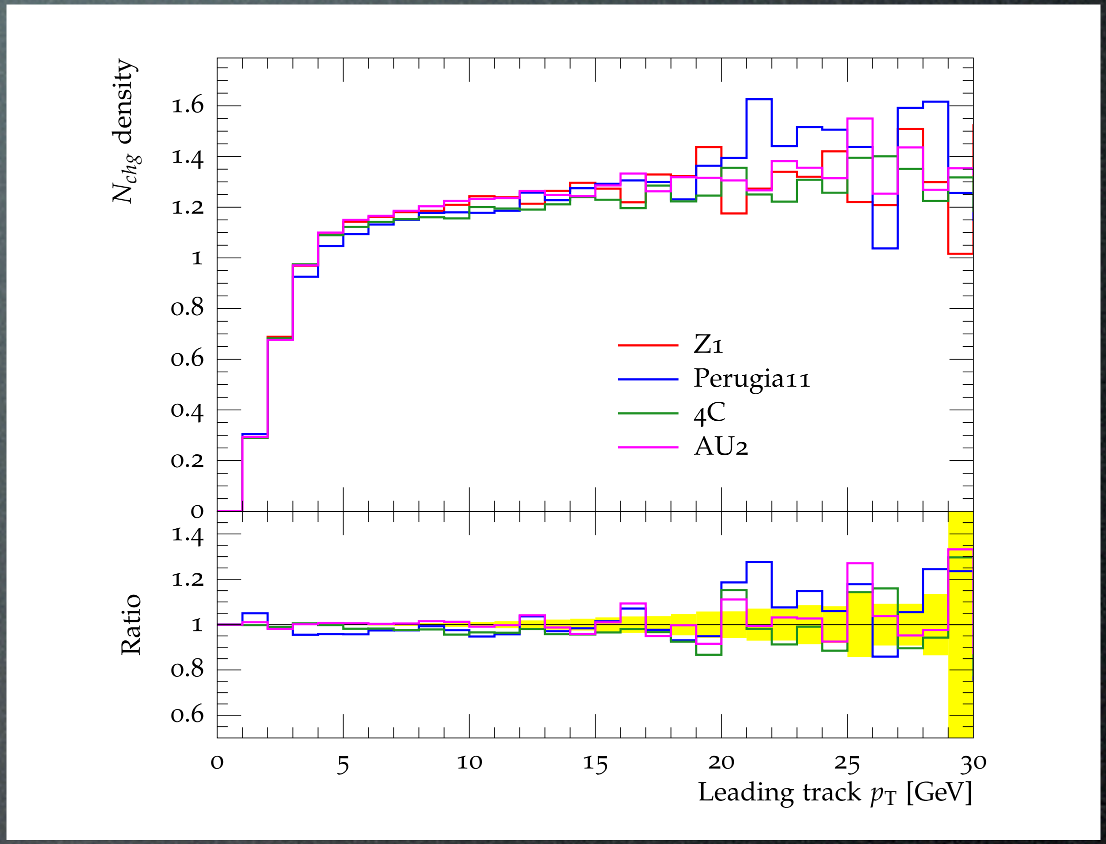
- Imperfection in transition region

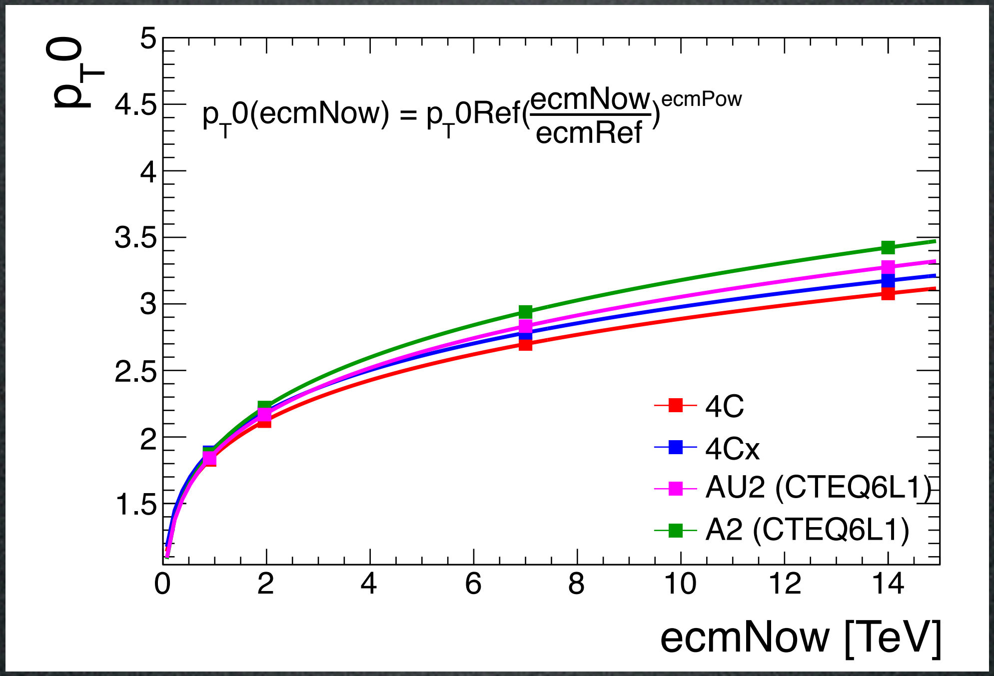




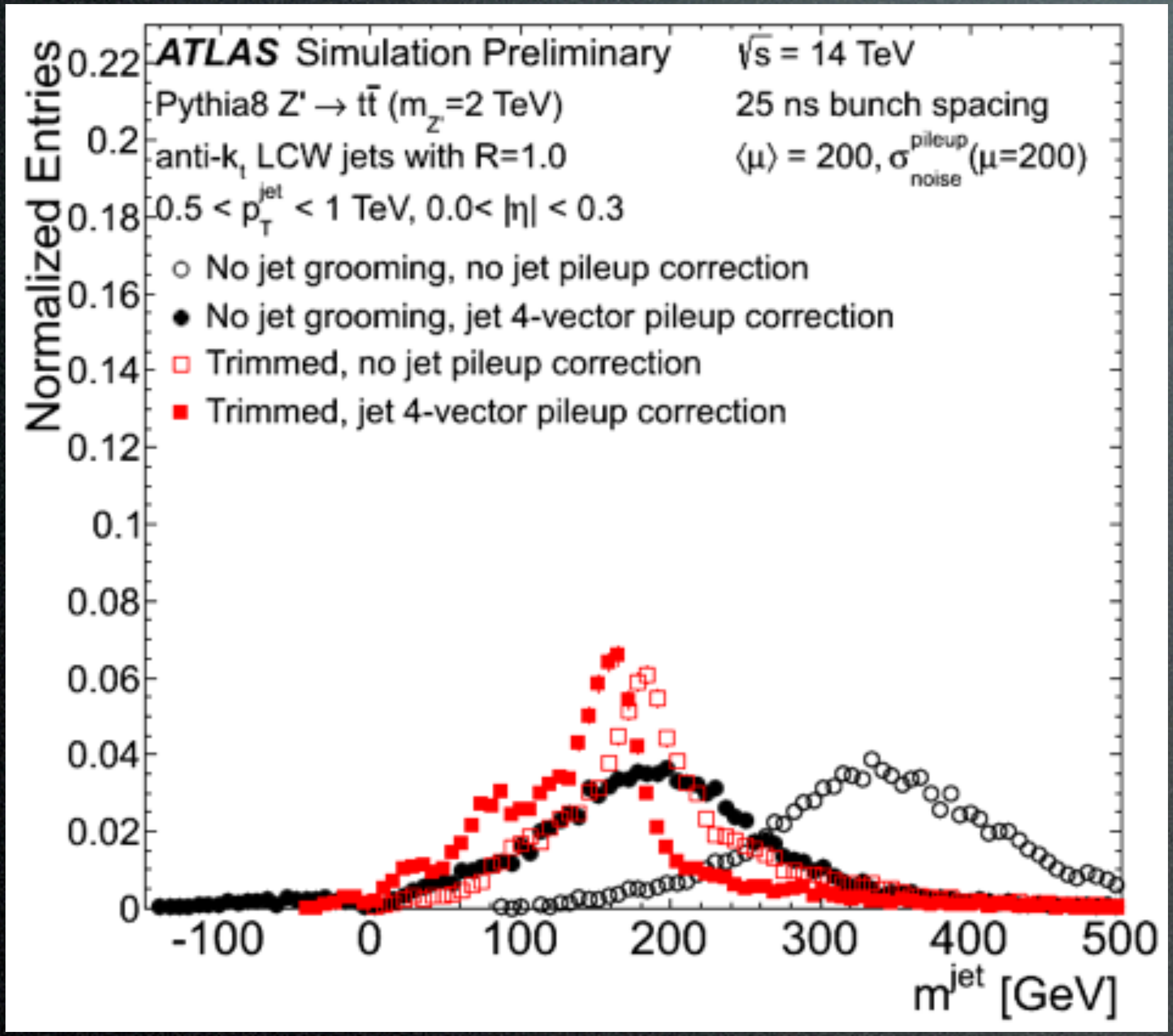
14 TeV UE Predictions

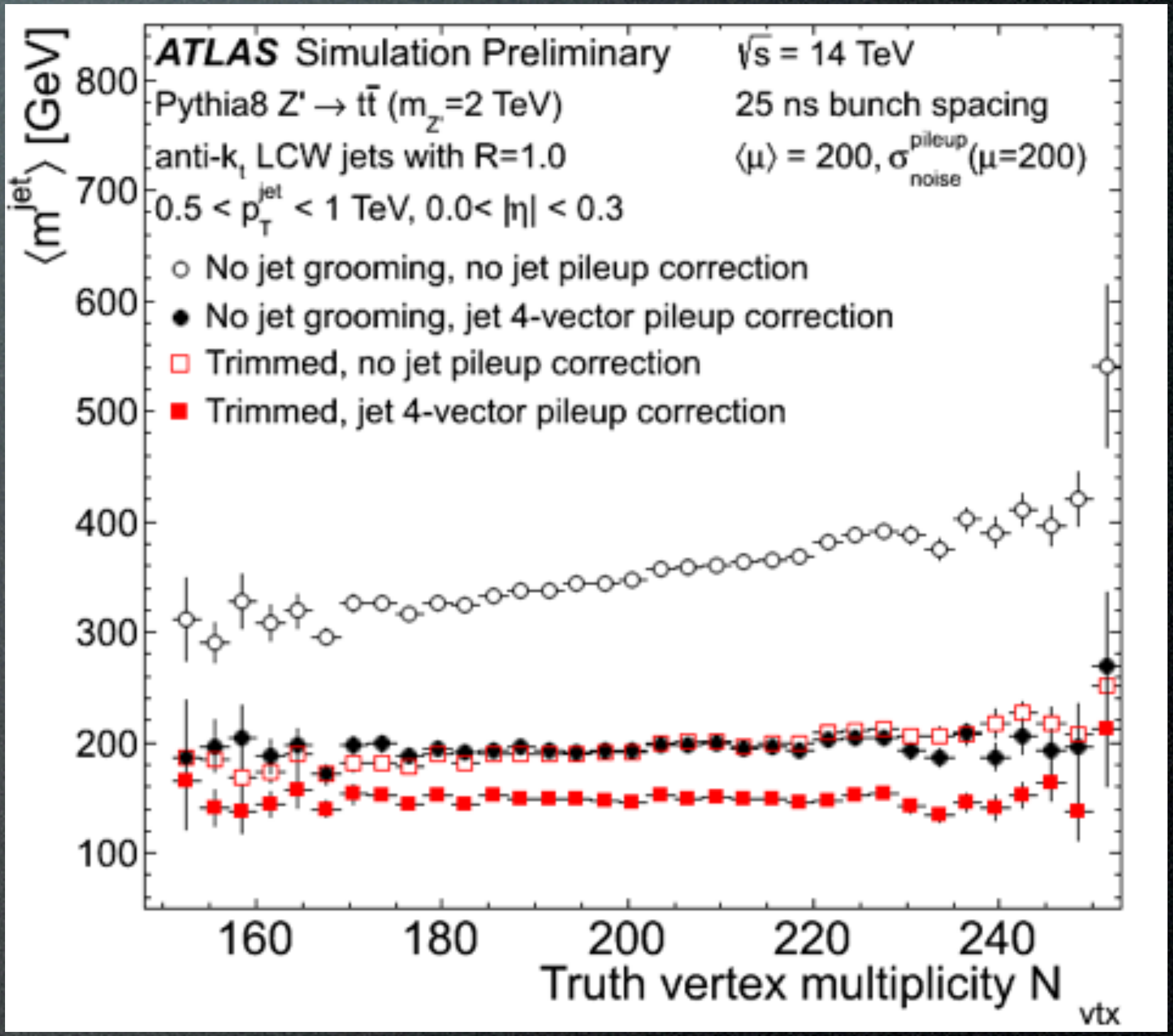






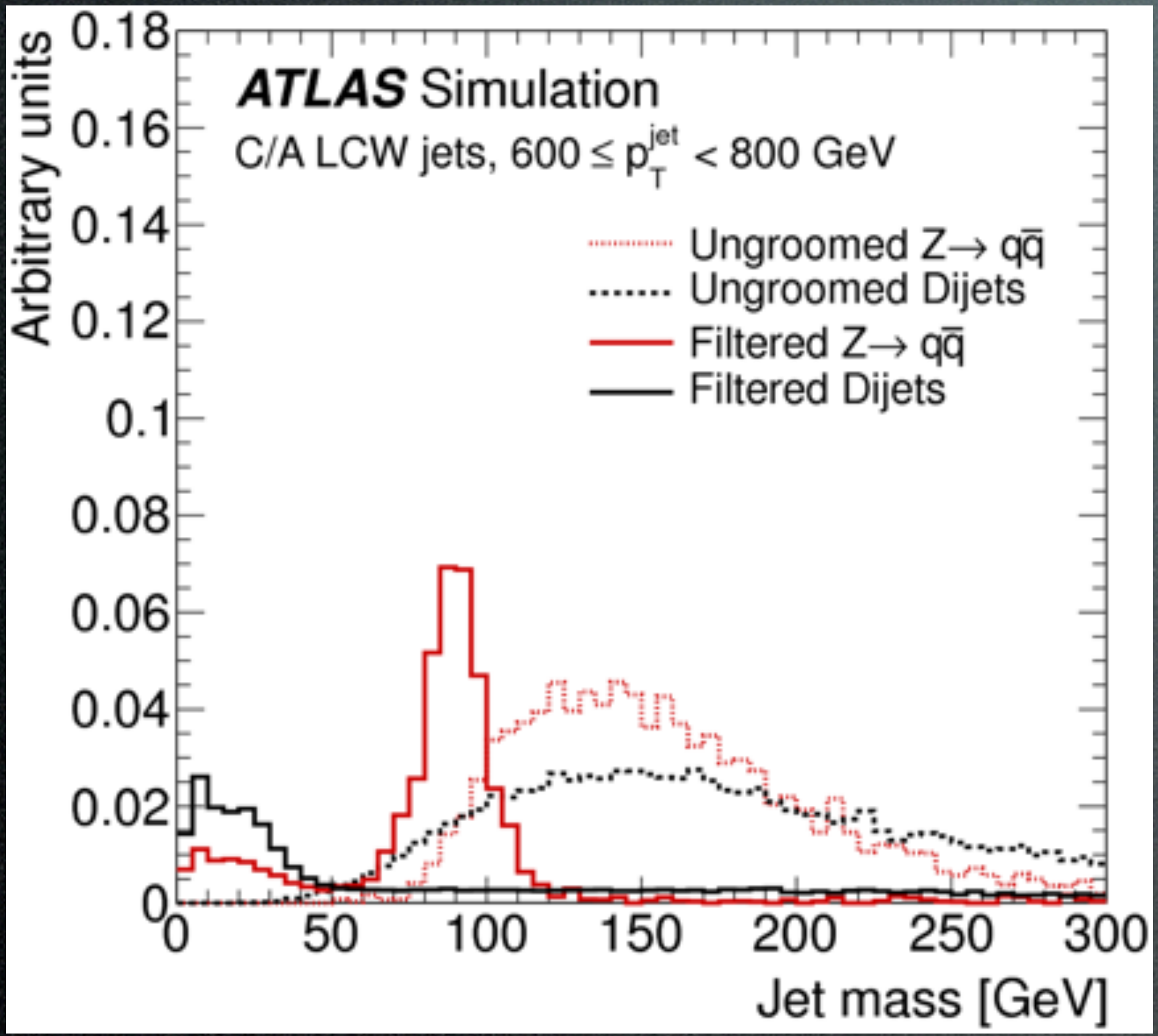
Effect of Grooming on Pileup



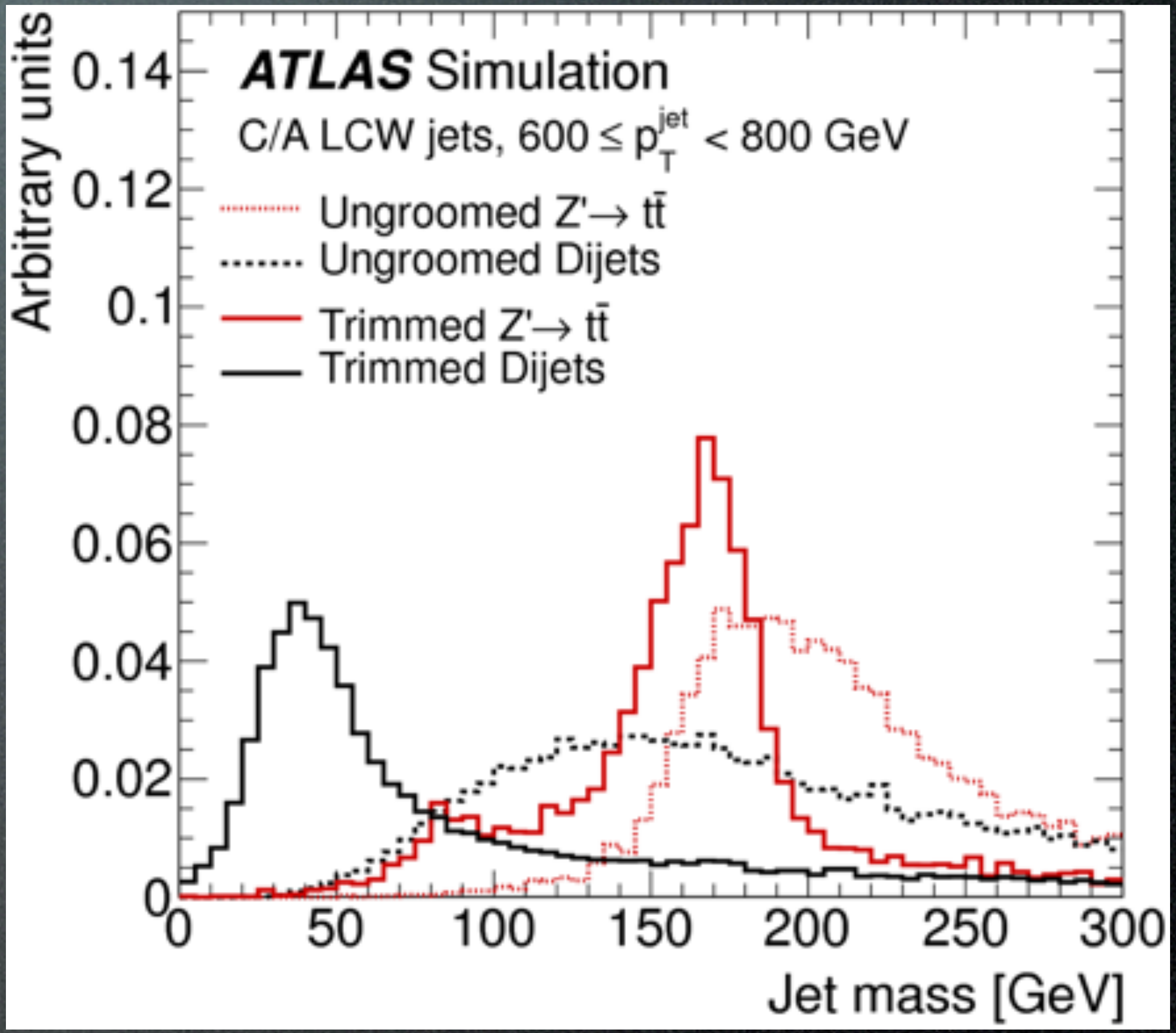


https://twiki.cern.ch/twiki/bin/view/AtlasPublic/JetEtmissApproved2013HighMuSubstructure

Jet Mass



arXiv:1306.4945v1



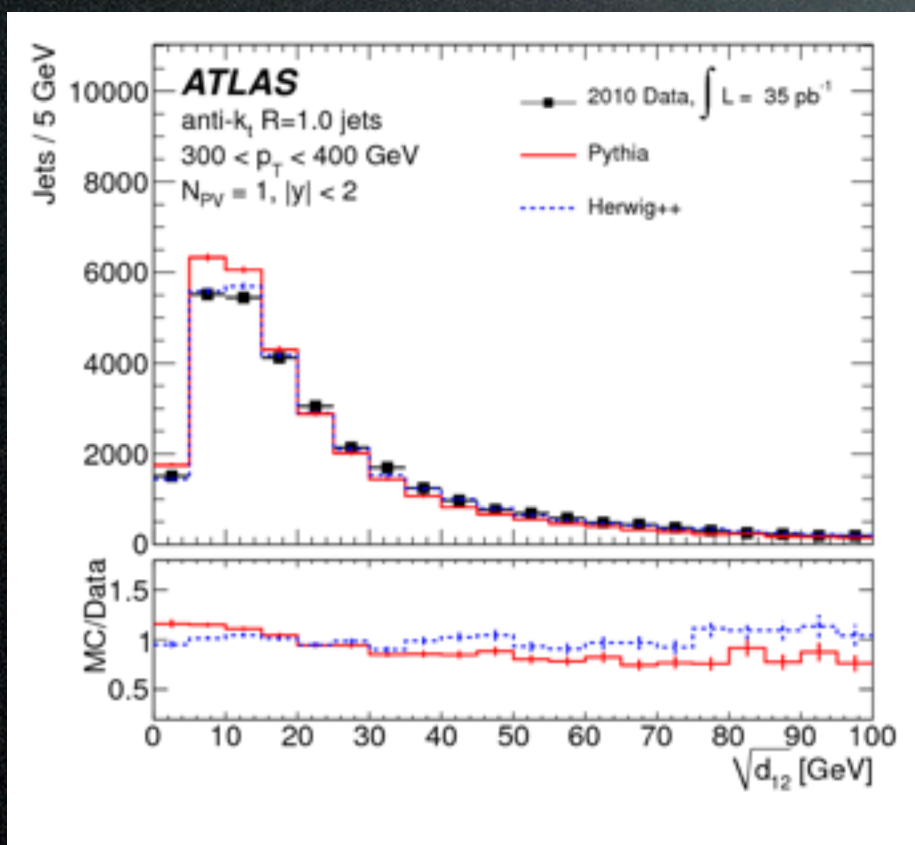
arXiv:1306.4945v1

Clear peak visible after grooming

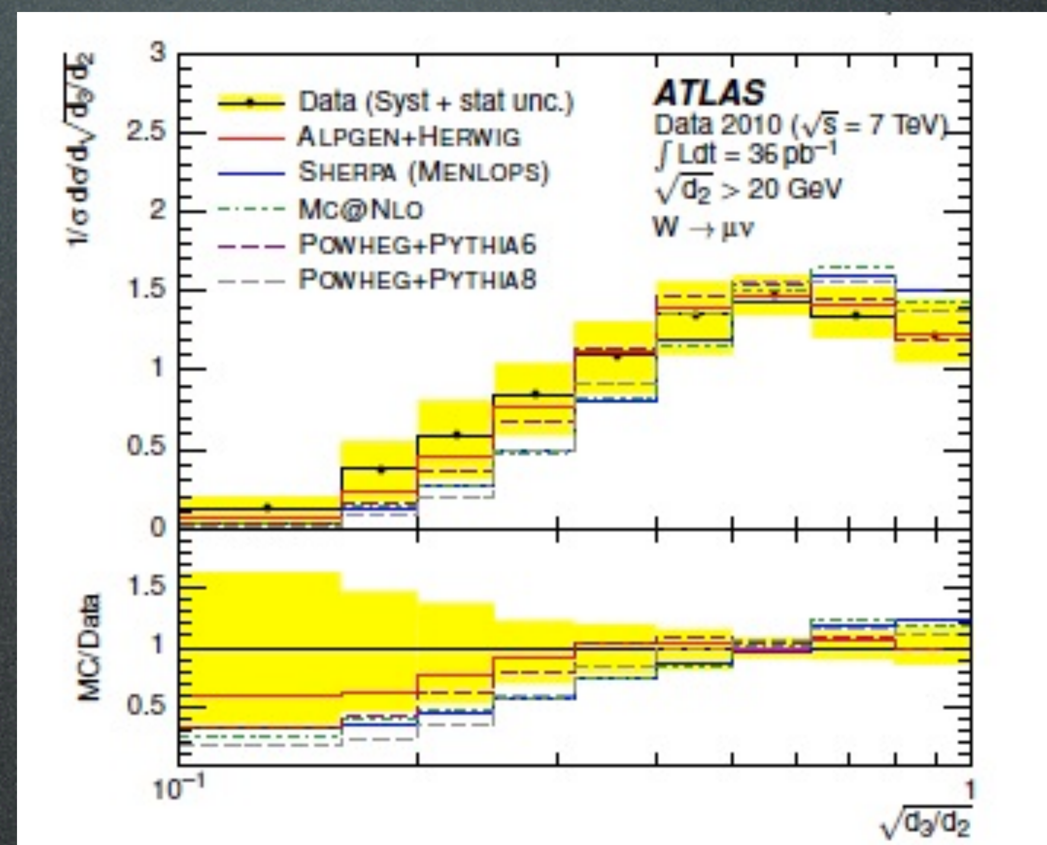
k_t Splitting Scale

$$\sqrt{d_{ij}} = \min(p_{\mathrm{T}i}, p_{\mathrm{T}j}) \times \Delta R_{ij}$$

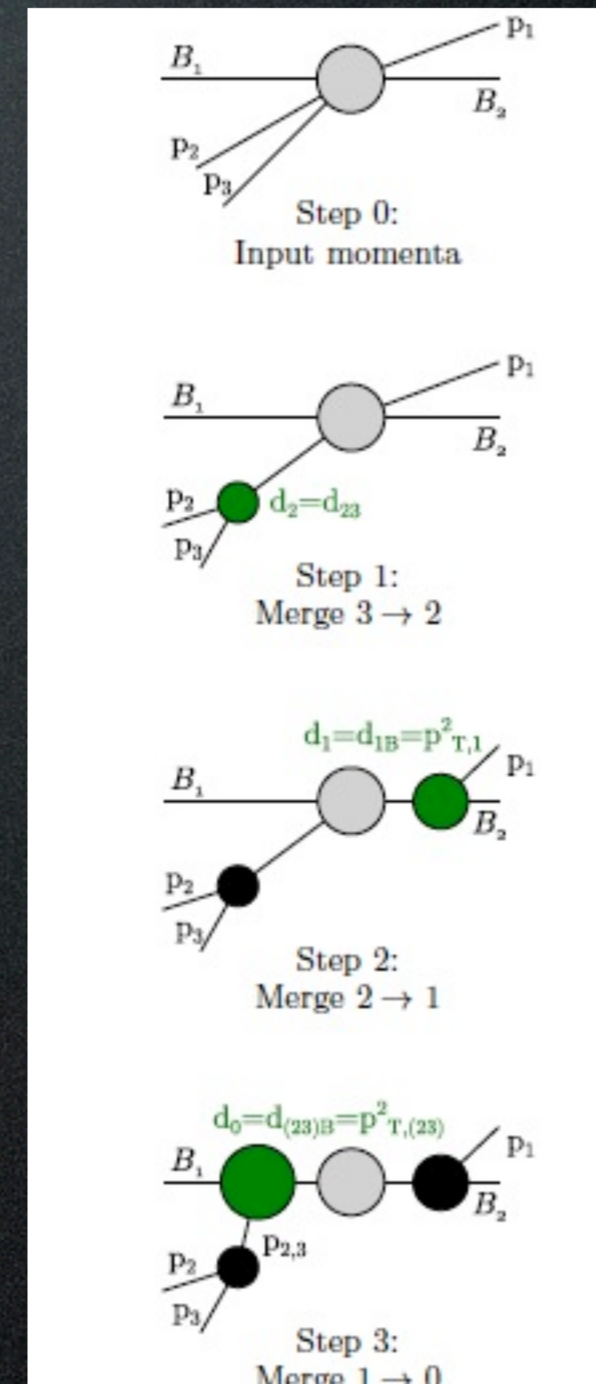
When combining two subjets with k_t algorithm:



arXiv:1203.4606



arXiv:1302.1415



arXiv:1302.1415

Symmetric for heavy particle two body decay

Jesse Thaler and Ken Van Tilburg; arXiv:1011.2268

N-Subjettiness

Quantify the degree to which jet radiation is aligned along specific subjet axes.

$$\tau_N \equiv \frac{1}{d_0} \sum_{k=1}^{M} \left(p_{\mathrm{T},k} \times \underbrace{\Delta R_{\min,k}}_{\text{distance to nearest subjet}} \right)$$

$d_0 = R \times$ sum of p_{T} of all constituents

Smaller values: N or less energy deposits

Larger values: more than N energy deposits

No of Subjets: ≤N >N

$\tau_N = 0$ $\tau_N = 1$

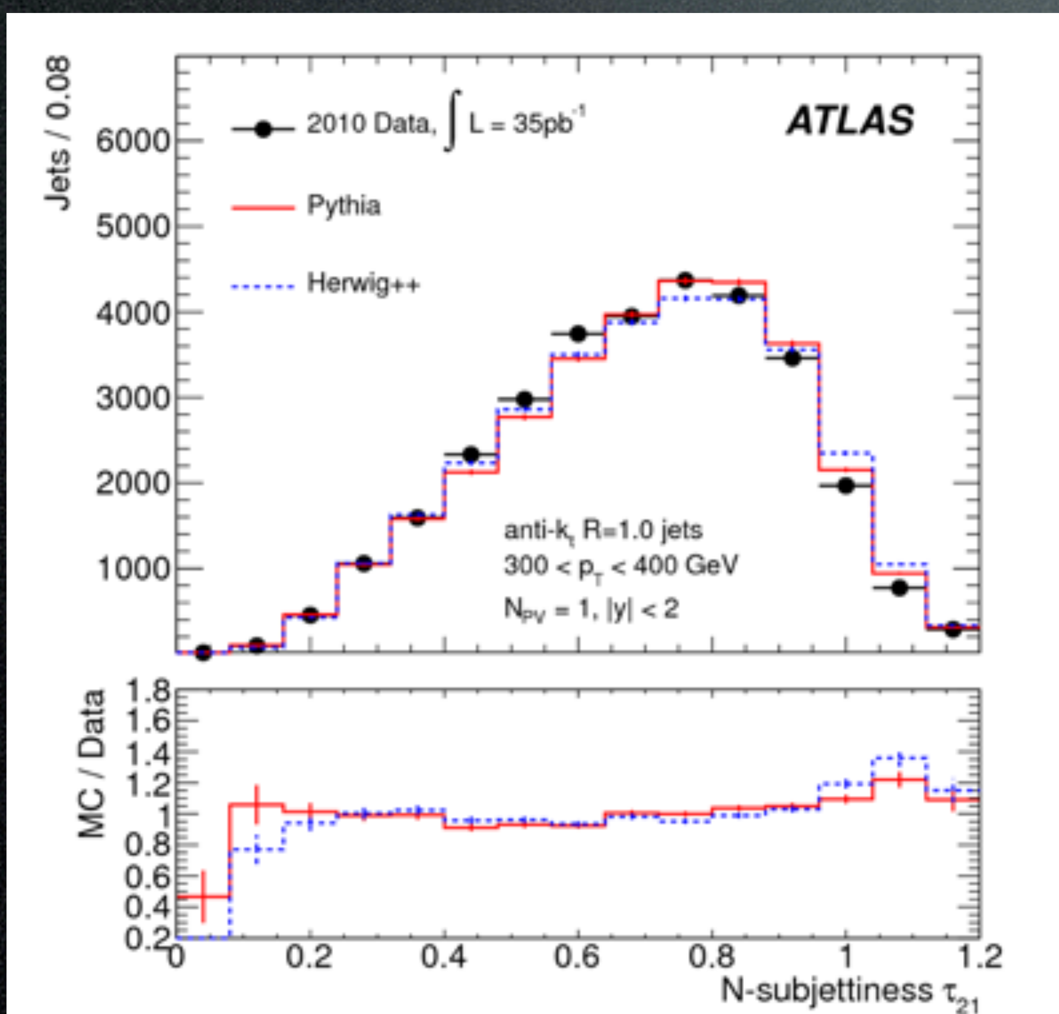
Calculated by k_t clustering the constituents, and requiring exactly N subjets

N-Subjettiness

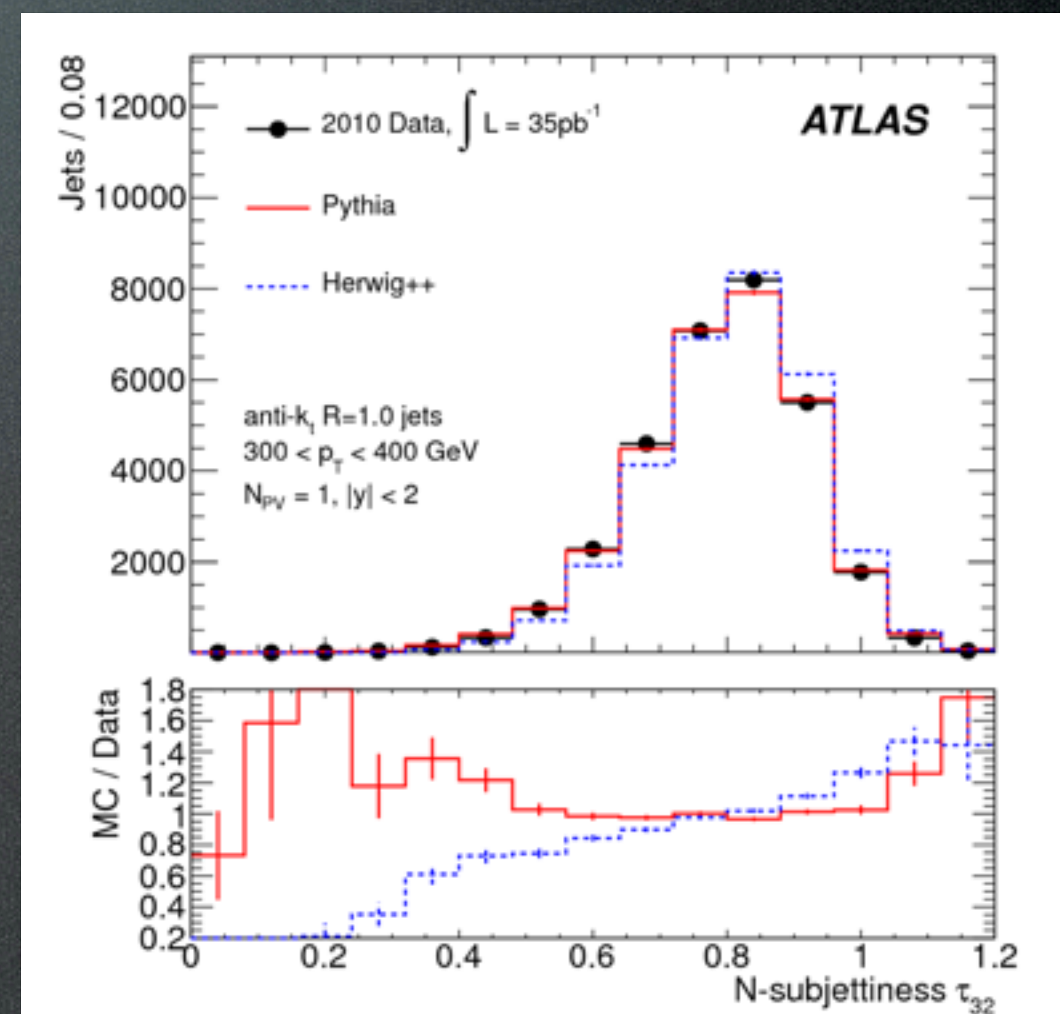
The ratio τ_N/τ_{N-1} is used as discriminant

More like 2 subjets than 1

More like 3 subjets than 2



arXiv:1203.4606



arXiv:1203.4606

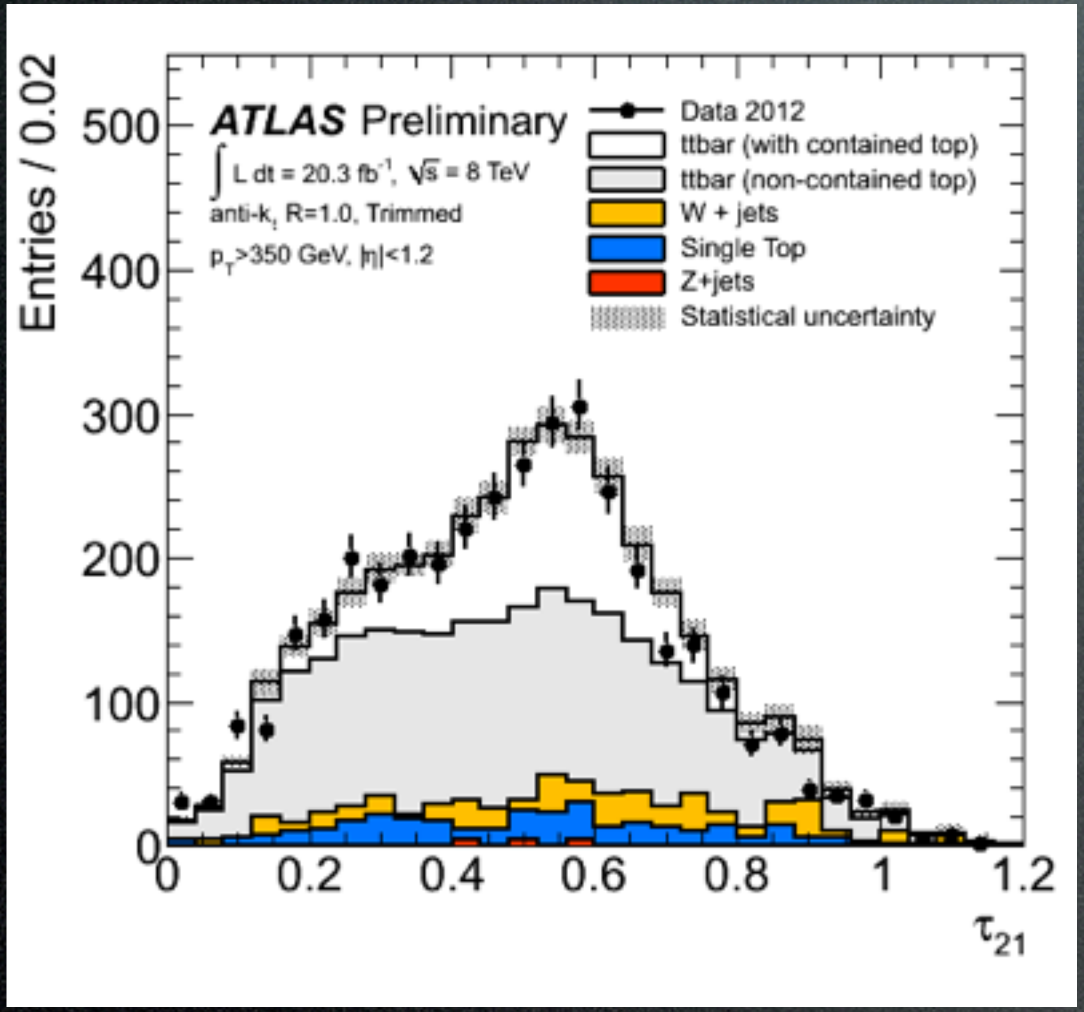
W-like ⟷ QCD-like

N-Subjettiness

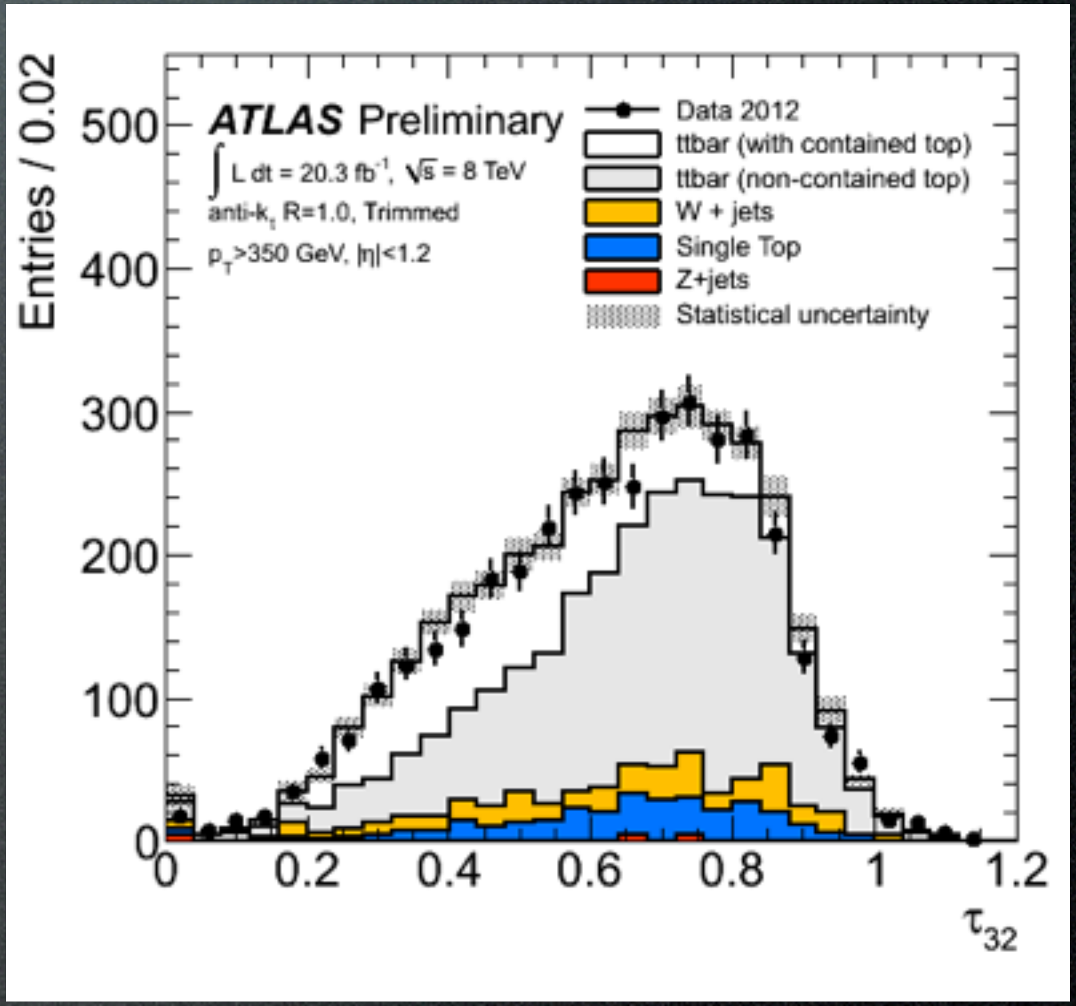
The ratio of τ_N/τ_{N-1} is used as discriminant

More like 2 subjets than 1

More like 3 subjets than 2



arXiv:1203.4606



arXiv:1203.4606

W-like QCD-like

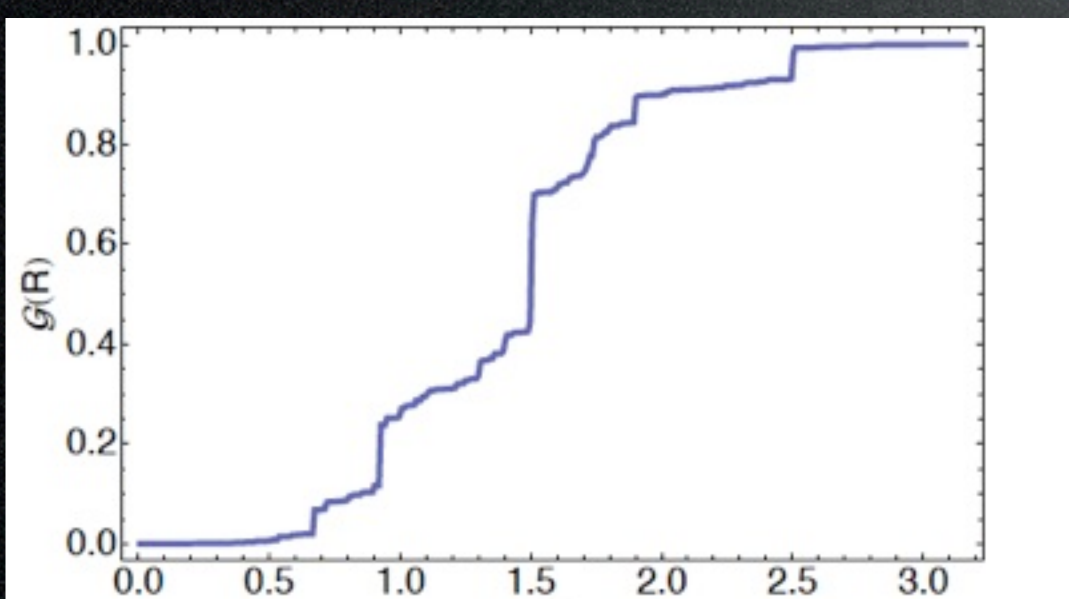
Martin Jankowiak, Andrew J. Larkoski; arXiv:1104.1646

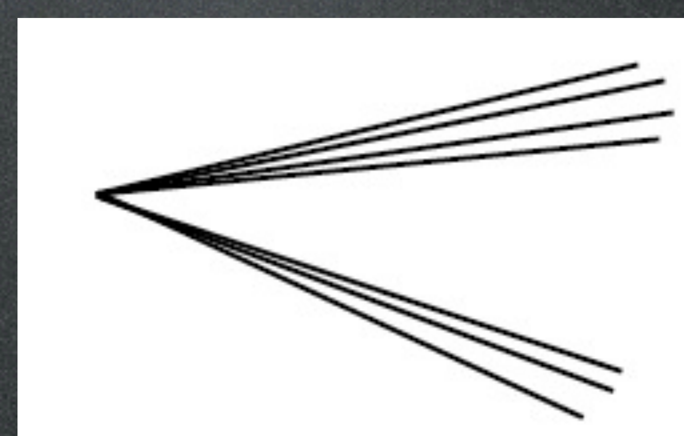
Angular Correlation Function

(or jet substructure without trees)

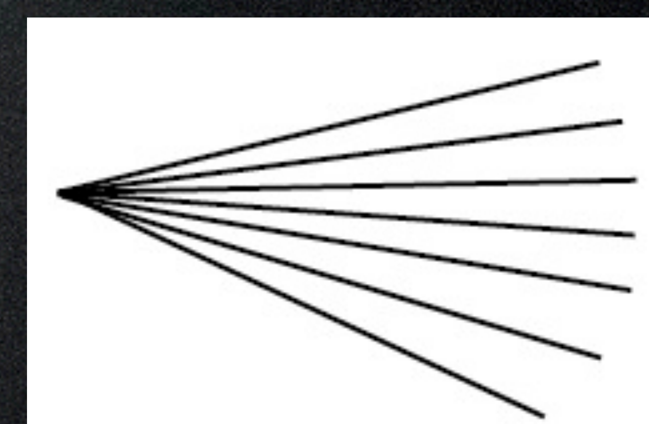
$$\mathcal{G}(R) \equiv \sum_{i \neq j} p_{\perp i} p_{\perp j} \Delta R_{ij}^2 \Theta[R - \Delta R_{ij}]$$

$$\Delta R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$$



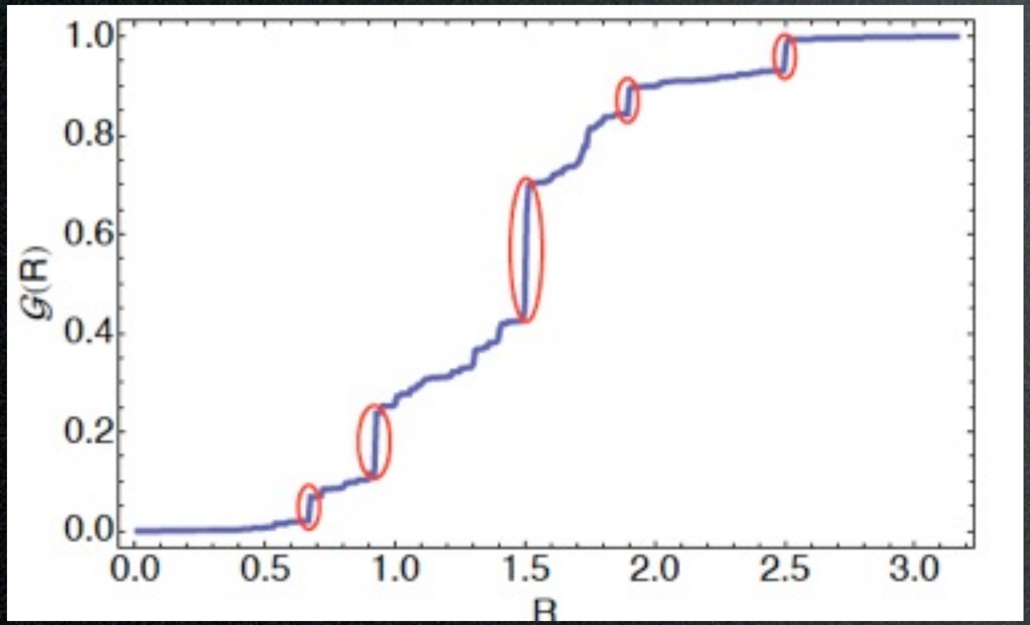


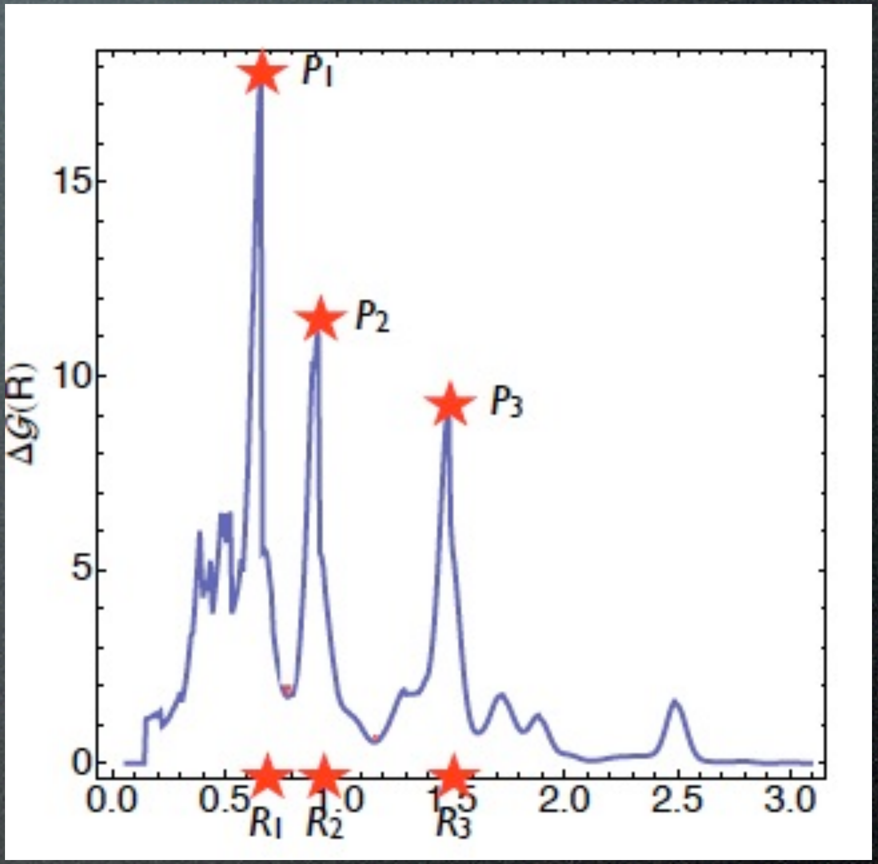
vs.



Angular Structure Function

$$\Delta\mathcal{G}(R) \equiv \frac{d\log\mathcal{G}(R)}{d\log R}$$



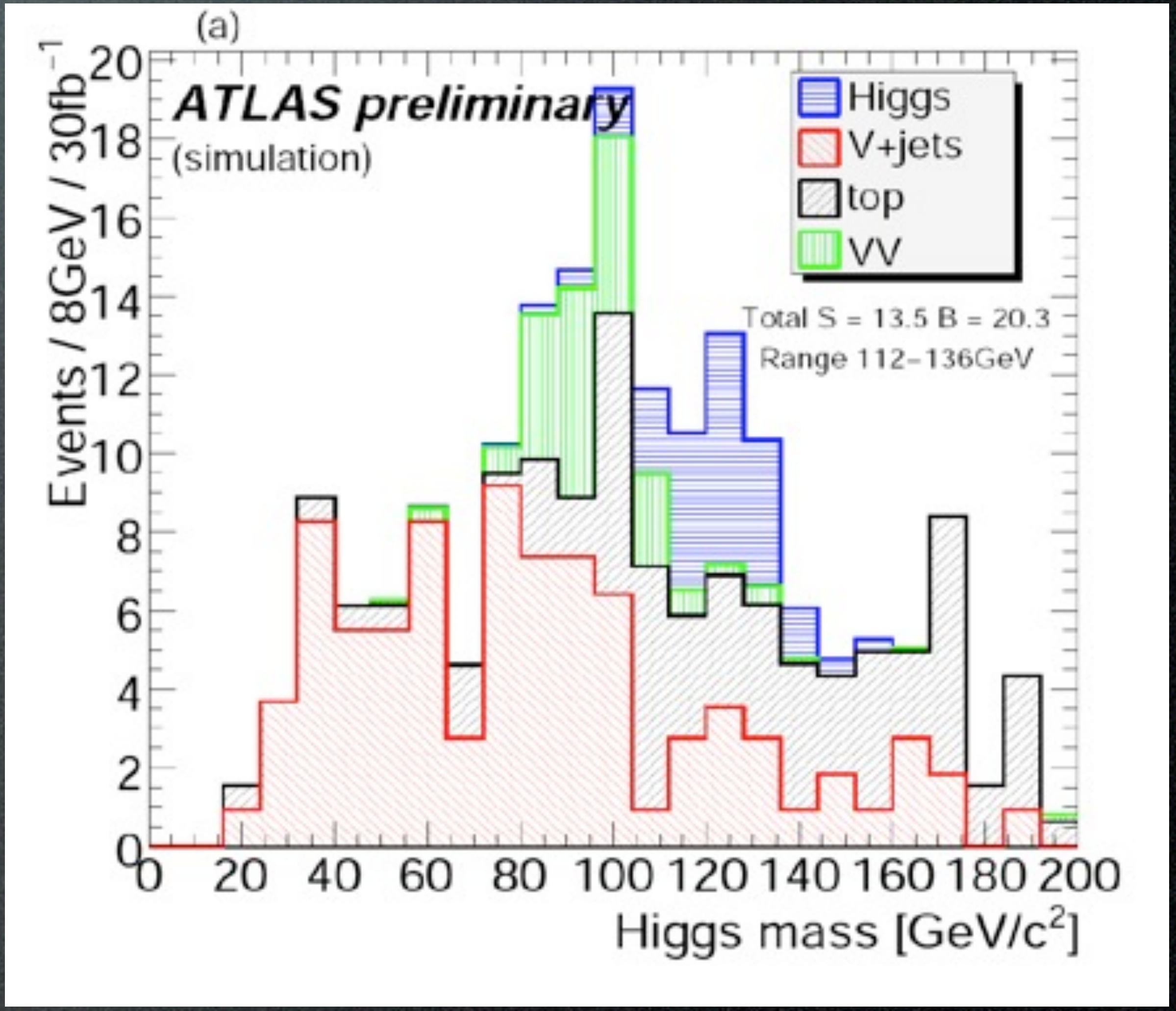


- Location of the peaks
- Height of the peaks
- Number of peaks

J. Butterworth, A. Davidson, M. Rubin, G. Salam; http://arxiv.org/abs/0802.2470

Where it all started: **B**utterworth-**D**avison-**R**ubin-**S**alam Higgs to bb tagger (2008)

- Start with fat (C-A 1.2) boosted ($p_T > 200$) b-tagged jet.
- De-cluster the jet. At each stage, mass drop and symmetric splitting requirement.
- Continue till an interesting splitting has been found.
- Higgs candidate from two hardest b-tagged subjets among the three hardest.



Tilman Plehn, Michael Spannowsky, Michihisa Takeuchi, Dirk Zerwas; arXiv:1006.2833

HEPTopTagger

Browsing through all the branches of jet recombination history

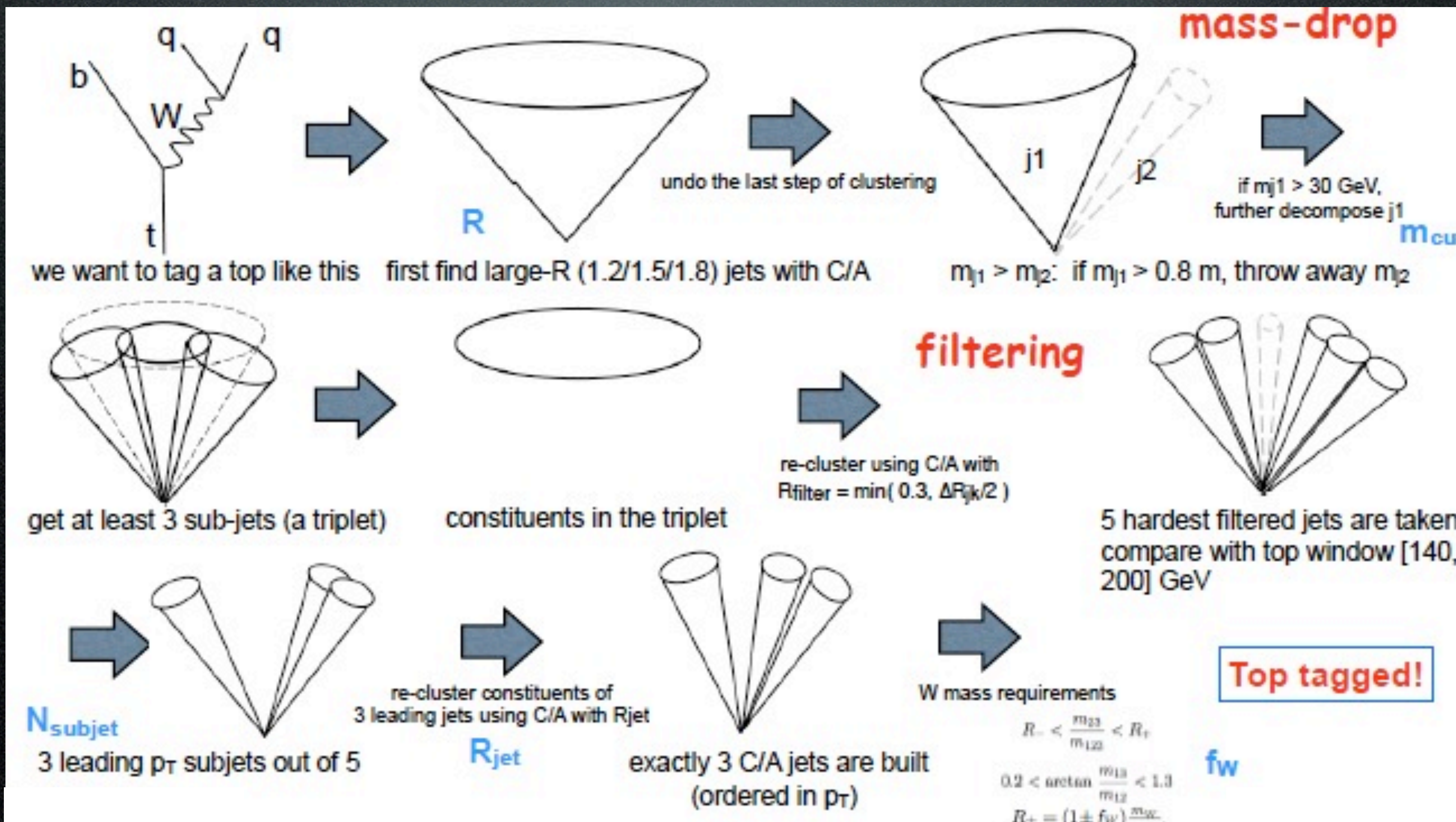
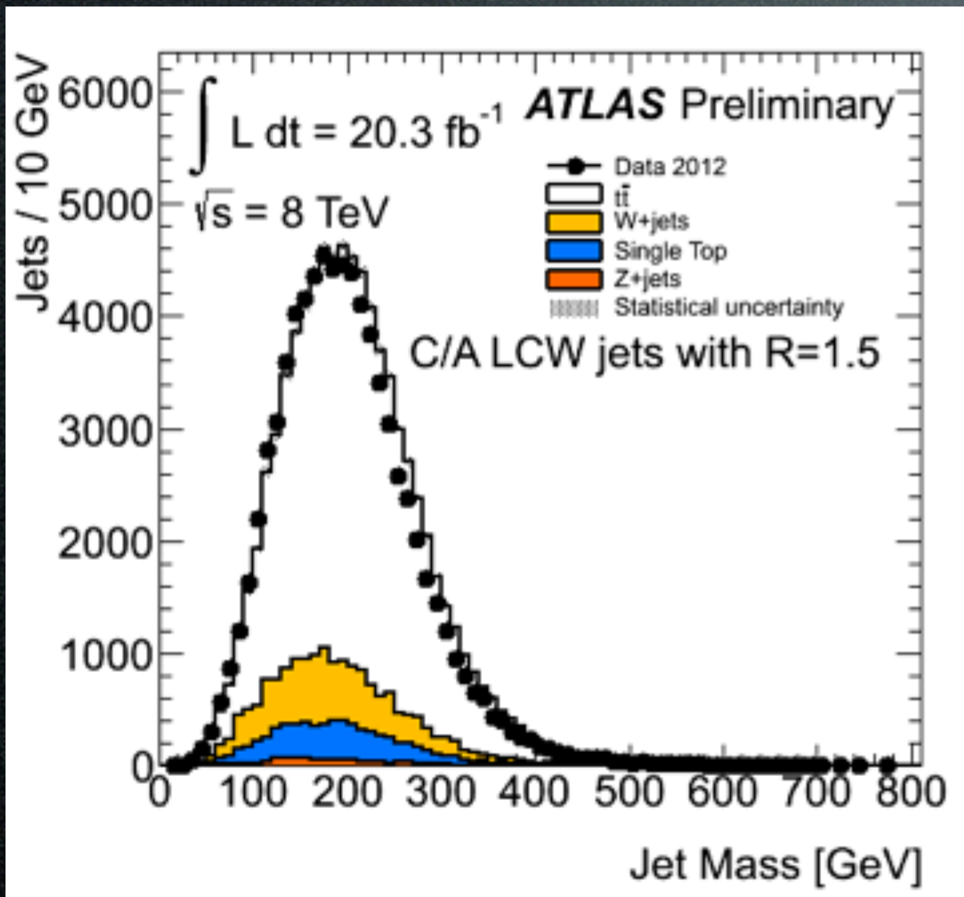
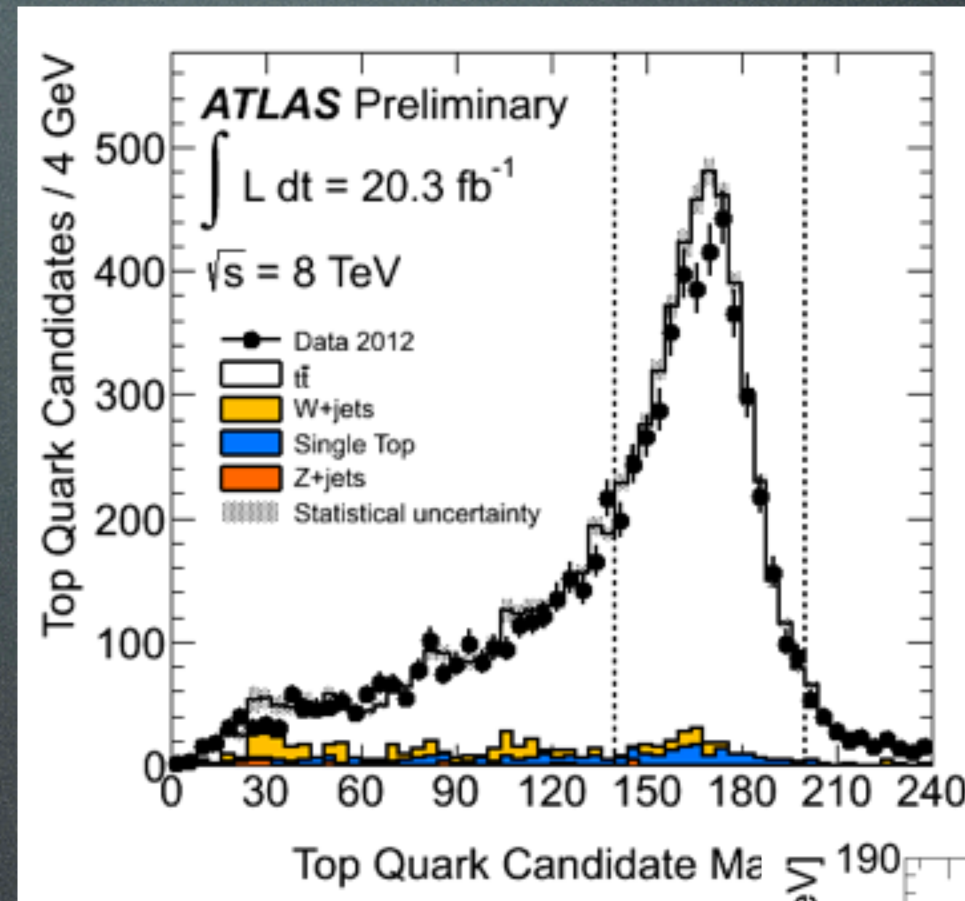


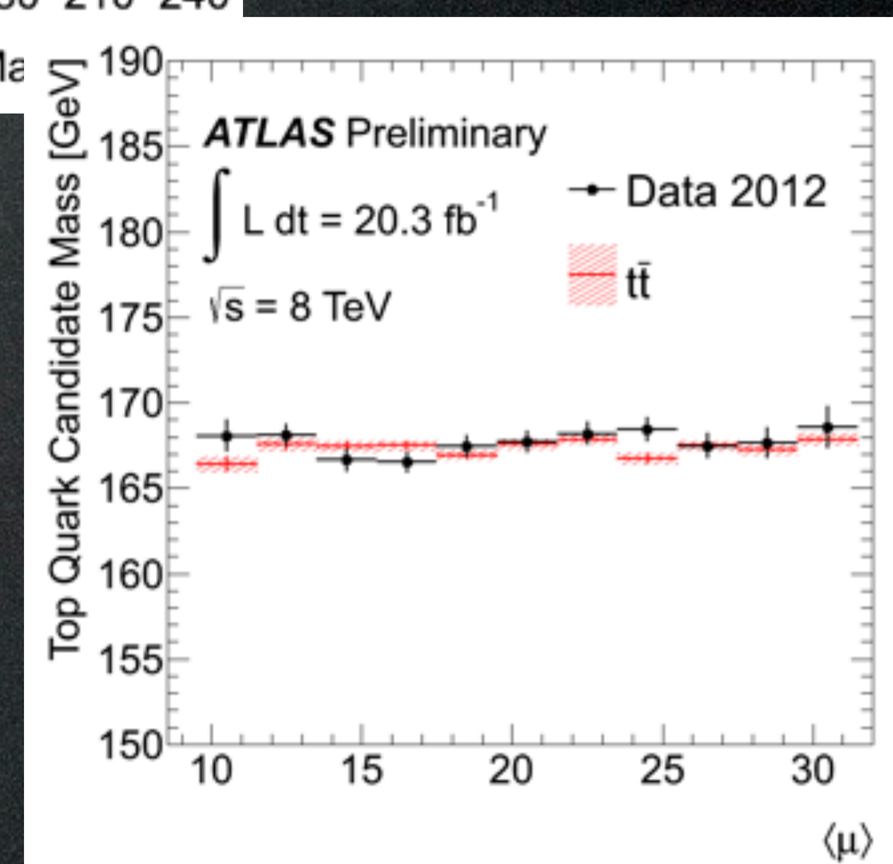
Figure by Xiaoxiao Wang

HEPTopTagger Performance





Before and after tagging by HepTopTagger



Pileup resilience

ATLAS-CONF-2013-084

Stephen D. Ellis, Davision E. Soper; arXiv:hep-ph/9305266

Matteo Cacciari, Gavin P. Salam, Gregory Soyez; arXiv:0802.1189

Detour: Jet Clustering

Distance between two input objects

Distance between each input object and beam

$$d_{ij} = min(k_{ti}^{2p}, k_{tj}^{2p})\frac{\Delta y^2 + \Delta\phi^2}{R^2}; \quad d_{iB} = k_{ti}^{2p}; \quad p = \begin{cases} 1 & k_t \\ 0 & \text{Cambridge/Aachen} \\ -1 & \text{anti-}k_t \end{cases}$$

Intrinsic transverse momentum

Fixed “radius” parameter

- Find the smallest of all $\{d_{ij}, d_{iB}\}$
- If this is one of the d_{ij} values, inputs i and j are merged.
- If it is one of the d_{iB} values, i^{th} input is considered a jet.
- Continue till all inputs are merged into jets.

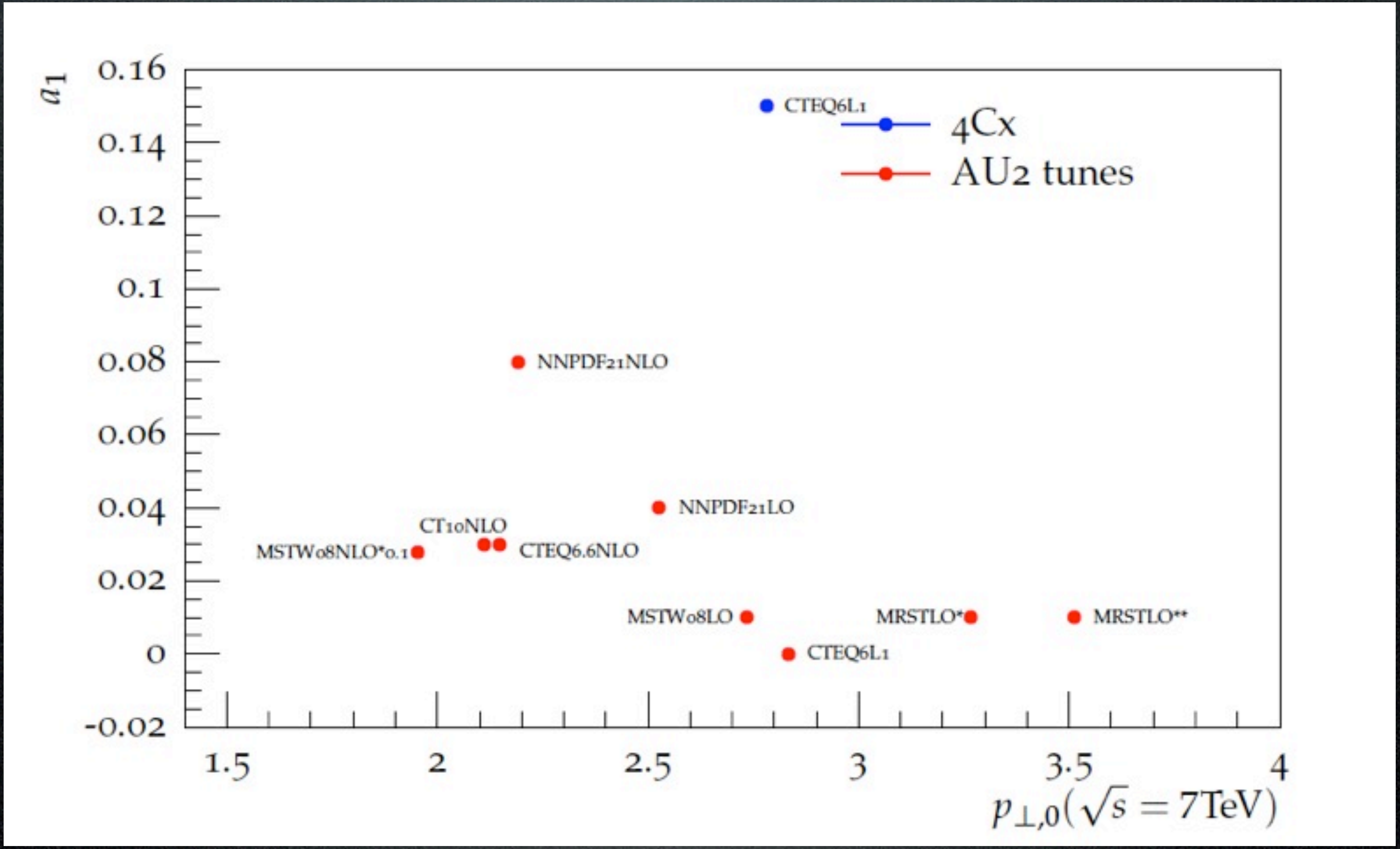
Milestones and Prospects

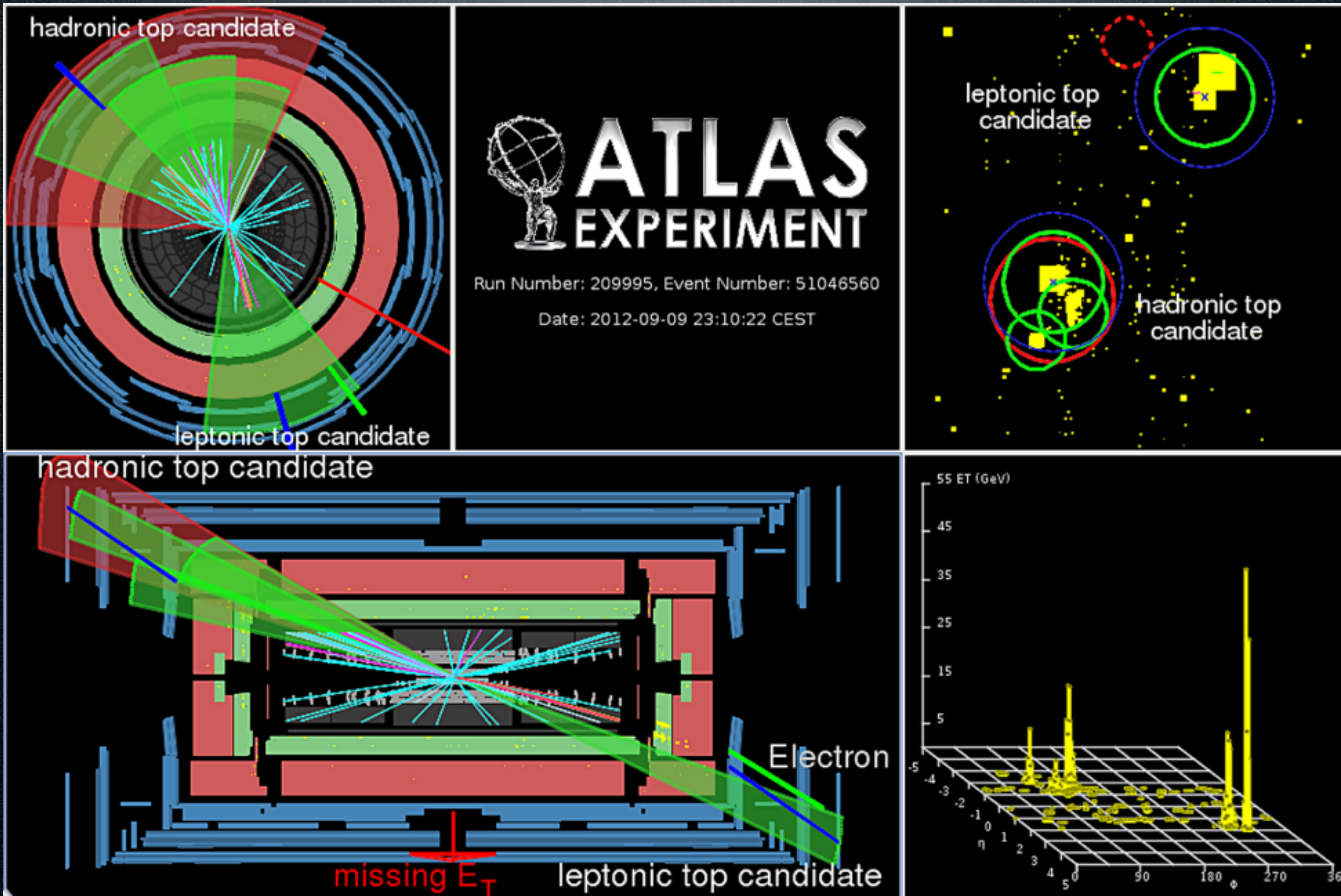
Run 1	Commissioning the tools
Run 2: 100 fb^{-1}	and... Improve precision of top/W/Higgs mass measurements. Exclude/severely constraint many of the new physics models with the higher energy reach
Run 3: 300 fb^{-1}	and... Directly test the coupling of the Higgs boson to fermions
HL-LHC: 3000 fb^{-1}	and... Measure Higgs self coupling Measure vector boson scattering Observe rare Higgs decays

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




hadronic top candidate
leptonic top candidate
ATLAS EXPERIMENT
Run Number: 209995, Event Number: 51046560
Date: 2012-09-09 23:10:22 CEST
leptonic top candidate
hadronic top candidate
hadronic top candidate
Electron
missing E_T
leptonic top candidate
55 ET (GeV)
45
35
25
15
5
-5
-4
-3
-2
-1
0
1
2
3
4
5
η
0
90
180
270
360
φ