Tuning Herwig 7 with Lund String Model

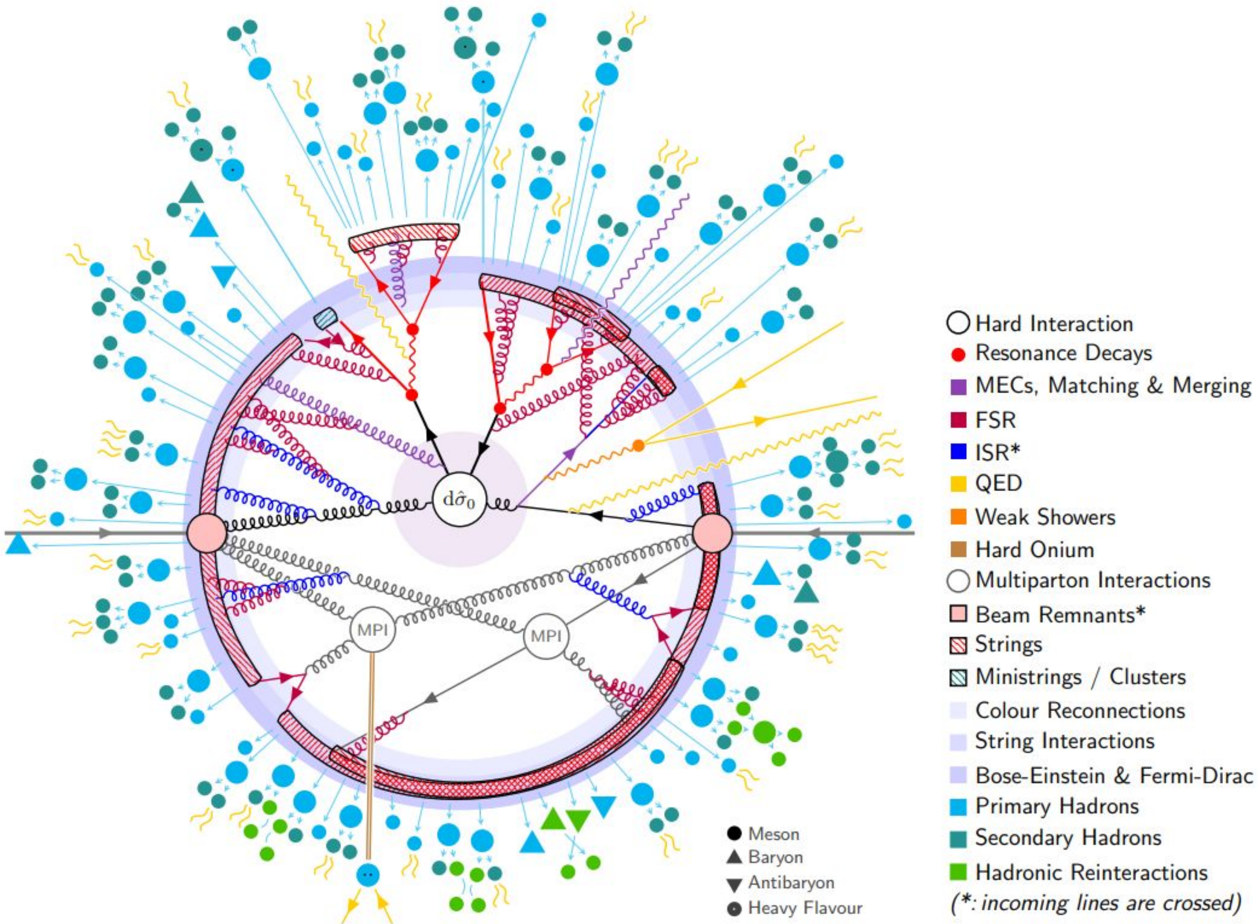
Pratixan Sarmah

NCN: 2019/34/E/ST2/00457

in collaboration with

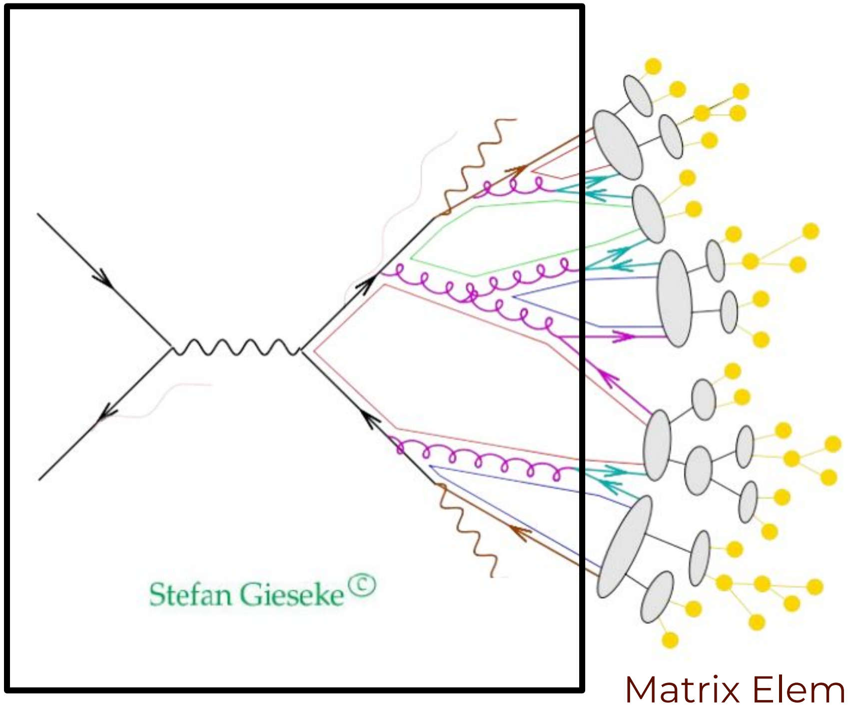
Michaela Divisova, Miroslav Myska and Andrzej Siódmok

- **Monte Carlo generators** are tools that can simulate collisions and bridges the **gap** between theory and experiment. An event generation simulation with the different stages looks like:



High Energy

- **perturbative QCD**
- **in theory we know what to do**
- **in practice very difficult**



Matrix Element + Parton Shower (see J.Whitehead's talk) - well understood

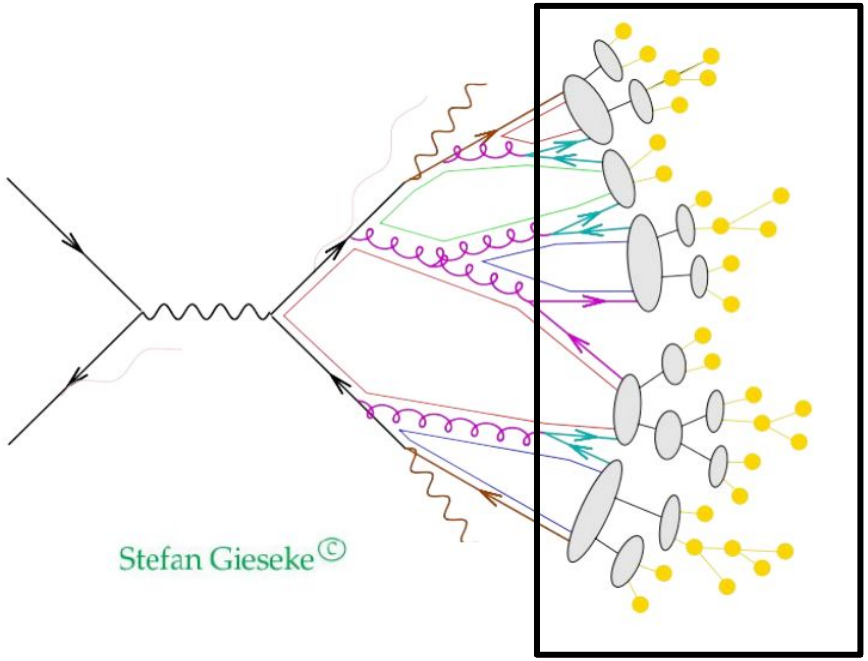
Introduction

High Energy

- **perturbative QCD**
- **in theory we know what to do**
- **in practice very difficult**

Low Energy

- **non-perturbative QCD**
- **we don't know what to do**
- **phenomenological models (with many free parameters)**



Hadronization:
one of the least understood region of MCEG

Motivation

High Energy

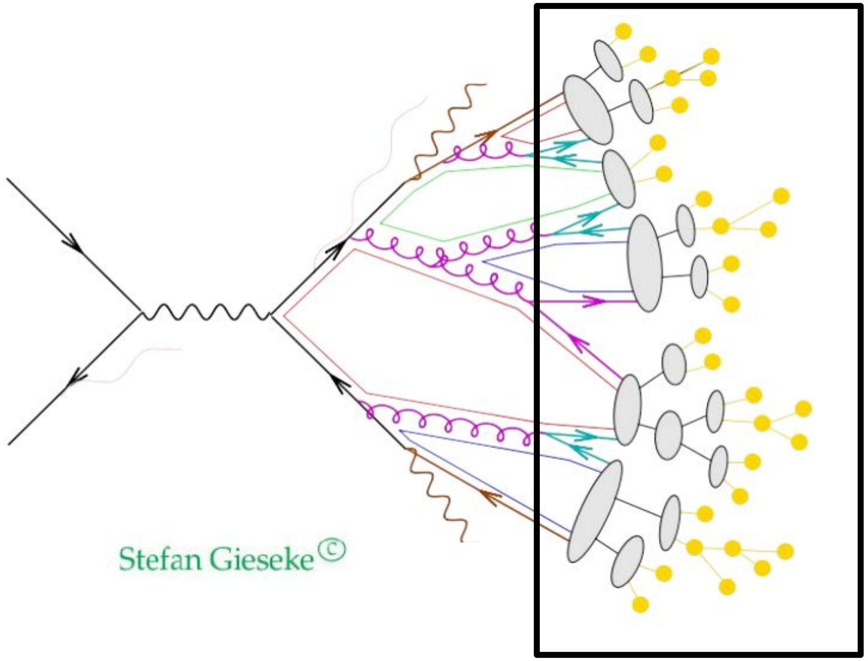
- **Increased control of perturbative corrections**

[see R. Poncelet and J. Whitehead talks]

Low Energy

- **more often LHC measurements limited by non-perturbative components (e.g hadronization)**

- W mass measurement using a new method [Freytsis at al. JHEP 1902 (2019) 003]
- Extraction of the strong coupling [M. Johnson, D. Maître, Phys.Rev. D97 (2018) no.5]
- Top mass [S. Argyropoulos, T. Sjöstrand, JHEP 1411 (2014) 043]
- and more...



Hadronization:
one of the least understood region of MCEG

High Energy

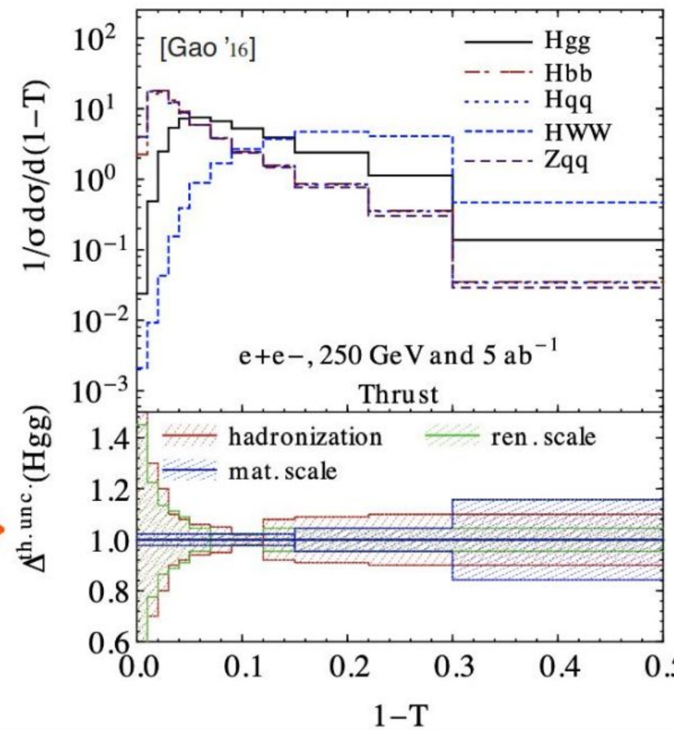
- **Increased control of perturbative corrections**

[see R. Poncelet and J. Whitehead talks]

Low Energy

- **more often LHC measurements limited by non-perturbative components (e.g hadronization)**

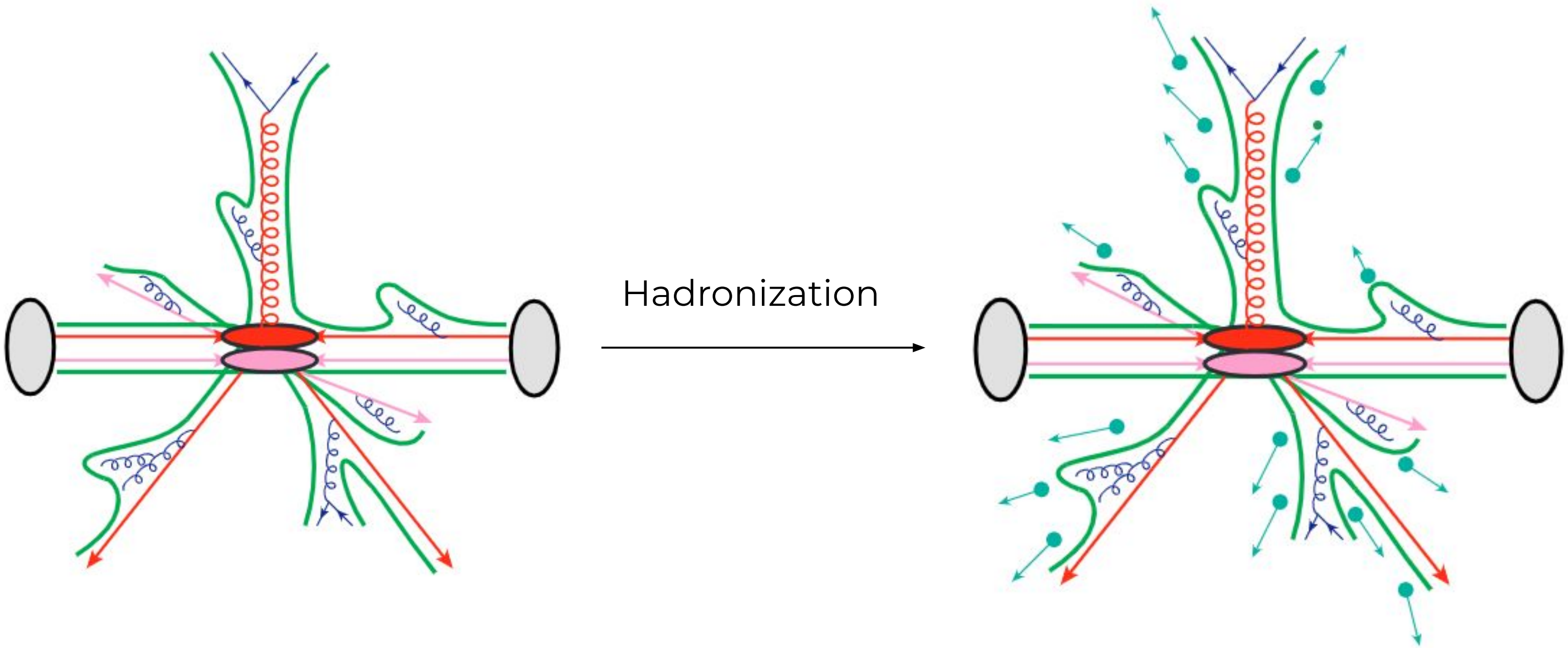
- However, hadronisation remains the main bottleneck
  - e.g. thrust in Higgs decays (MC variation in plot)
- Increase in energy insufficient for suppression ($Q \sim m_H$)
- Runs at lower energies are essential for a robust tuning of NP models in MCs
- Also crucial for training of ML algorithms for jet tagging, instrumental in extraction of Higgs couplings



Pier Moni's talk at FCC Physics Workshop 2023

Motivation

- **Hadronization** - two most commonly used models are **String Model** in Pythia and **Cluster Model** in Herwig and Sherpa



- These models generally contain several free parameters that cannot be inferred from first principles and require **tuning** to experimental data.

Images taken from S.Prestel

Motivation

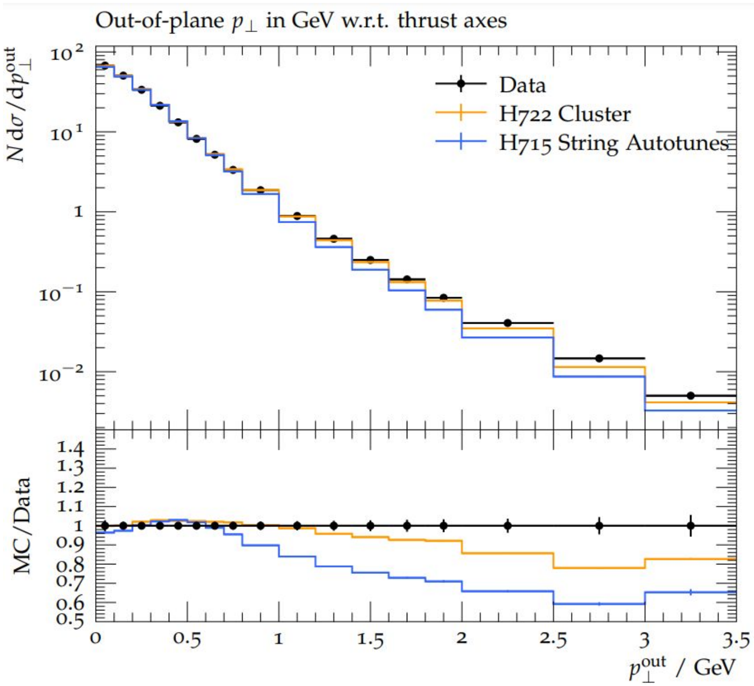
- Several tunes have been performed to LHC data over the years for event generators Pythia and Herwig

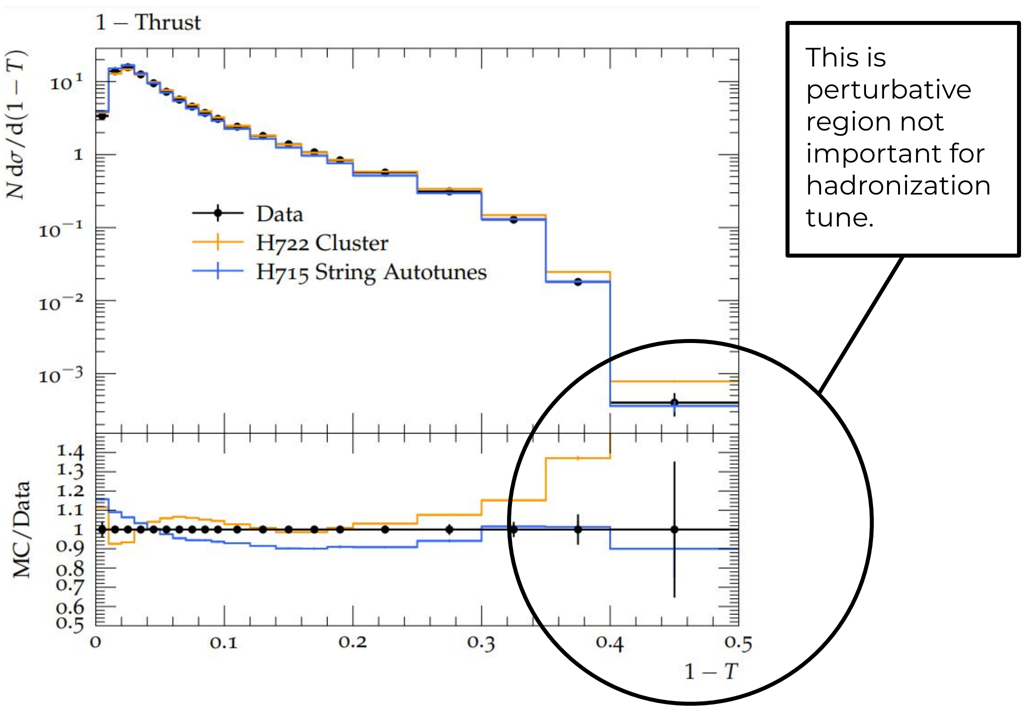
- It has been seen, the **Lund String** model in Pythia does better describing data sensitive to hadronization process and the **Angular Ordered (AOPS)** shower in Herwig7 does better in regions sensitive to perturbative calculations

- A first attempt was made to tune Herwig7 with Lund String in a new framework named AutoTunes [J.Bellm, L.Gellersen, EPJC 2020]

- **TheP8I**, written by L. Lonnblad allowed the internal use of Pythia 8 strings with Herwig 7 events (with some additional modifications)

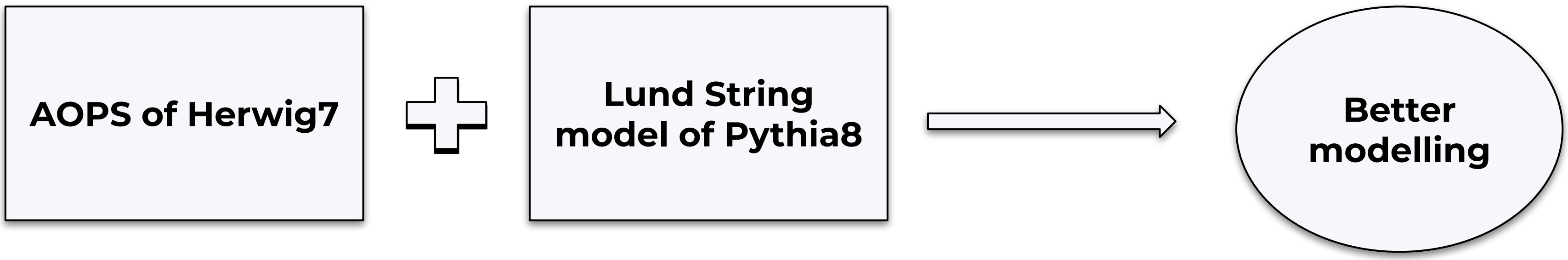
- However, the AutoTunes paper focused more on a new automatized tuning procedure than getting the best tune.
- Thus, we reproduced the results based on their tuned parameters
  [there were no plots comparing the tune to the data in the paper]
- The results didn't show much improvement and it also performed **worse** in many regions when compared to the default Herwig7 + cluster tune. This could be due to AutoTunes not taking correlations into account.





Motivation

- We would naively expect:



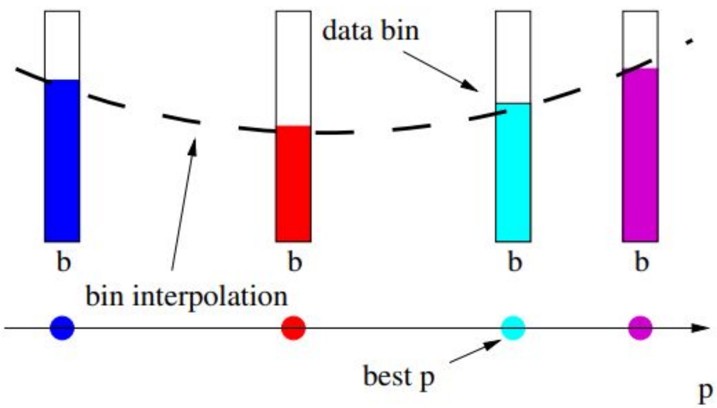
- Therefore we want to understand if the gap between our expectation and previous results are due to the tuning setup or physics understanding of the NP* region.

- A successfully combined model would also provide a setup for a dedicated study of the systematics of NP effects.

*NP - Non-Perturbative

Tuning strategy

- We follow the **PROFESSOR** approach (as done in Pythia 6 tune) [A.Buckley et al, EPJC 2010] to tune the AOPS of Herwig7 + Lund String model.
  Two stages of tuning:
  - Fragmentation parameters and
  - Flavour parameters



- The **PROFESSOR** framework parameterizes the generator response to a n-order polynomial and minimizes a Goodness-of-Fit function (χ^2) to get the set of best parameter values.

$$\chi^2(\boldsymbol{p}) = \sum_{\mathcal{O}} w_{\mathcal{O}} \sum_{b \in \mathcal{O}} \frac{(f^{(b)}(\boldsymbol{p}) - \mathcal{R}_b)^2}{\Delta_b^2},$$

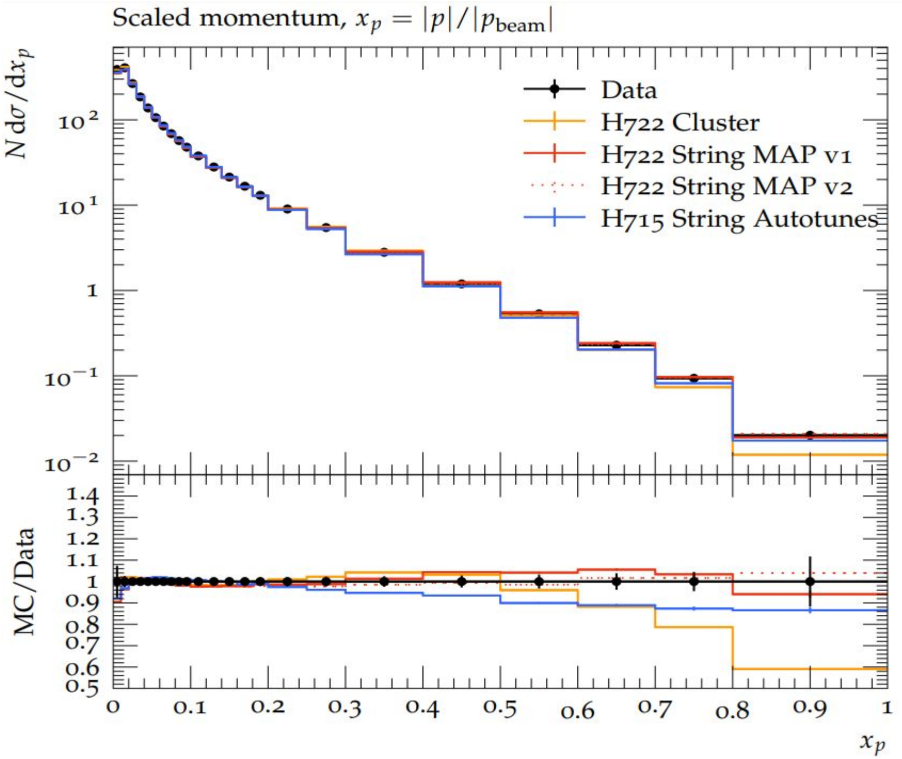
Image taken from H.Schulz

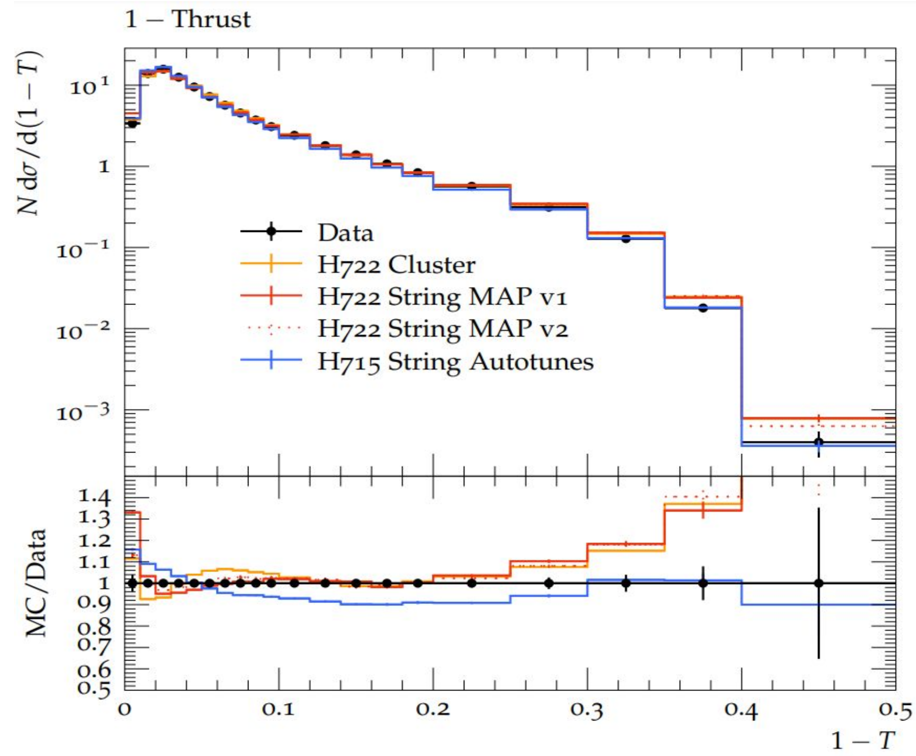
Tuning to LEP data

Herwig 7 + Lund String Model

Event Shape Observables

- Measurements from DELPHI

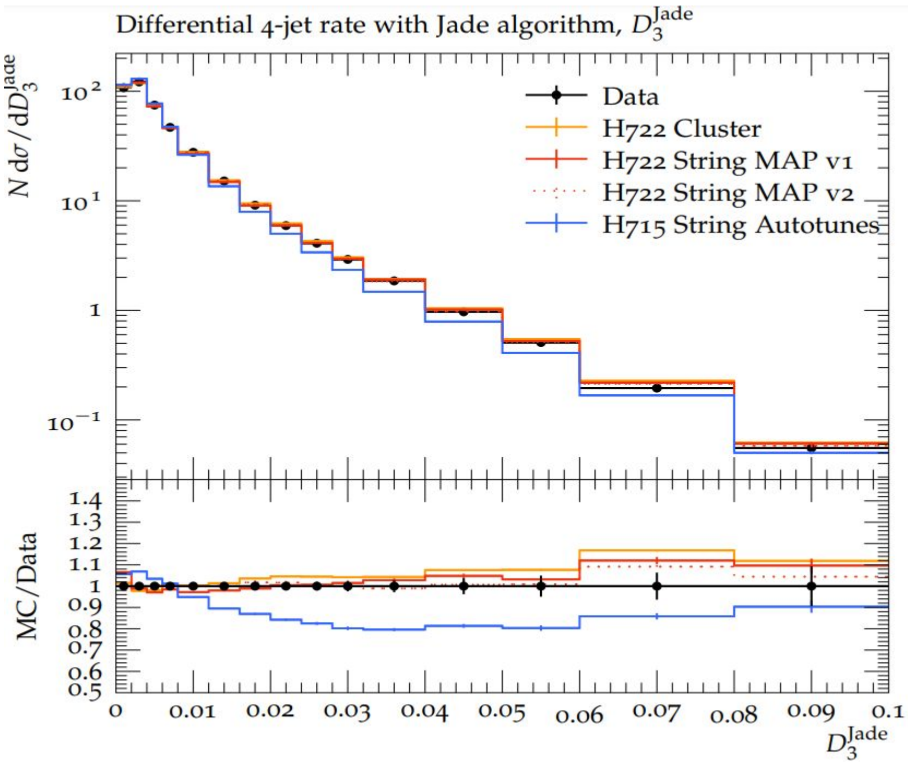


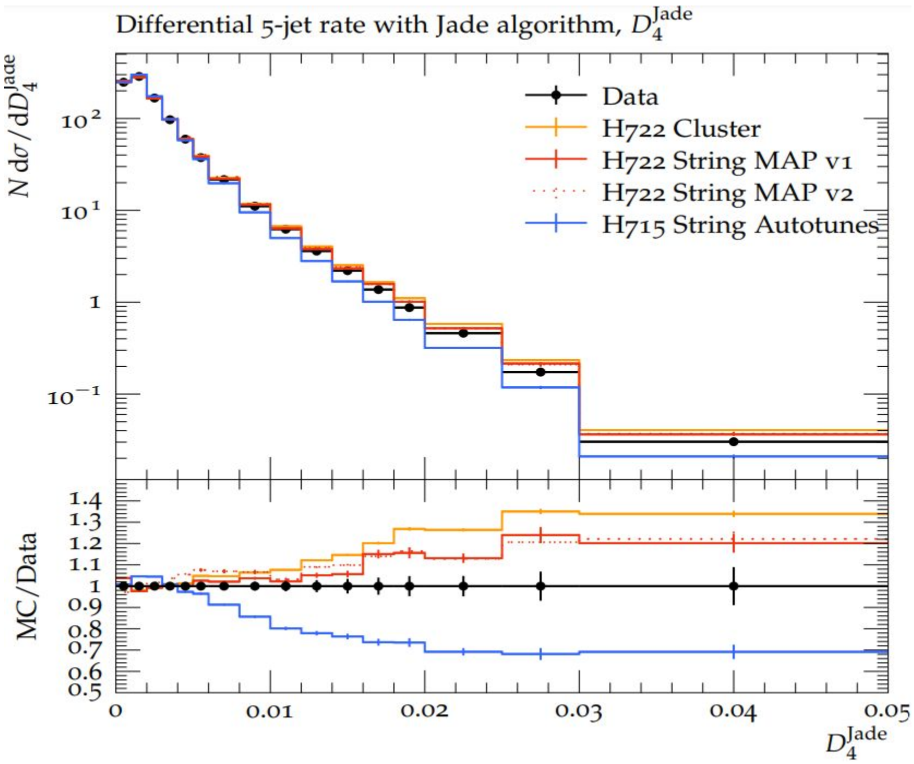


- Both these distributions were used in tune.

Jet Observables

- Measurements from DELPHI

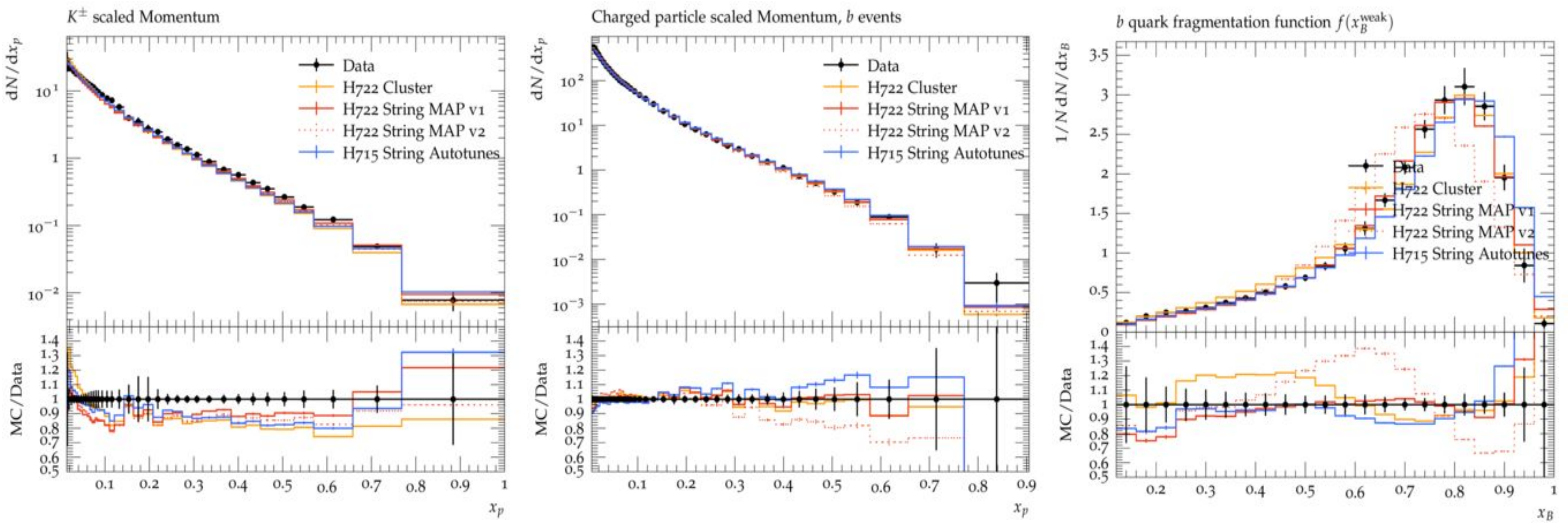




- Both these distributions were **not** used in the tune.

Flavour sensitive Observables

- Measurements from SLD.



- MAP tune v1 and v2 shows much differences with v1 performing better over a wider range

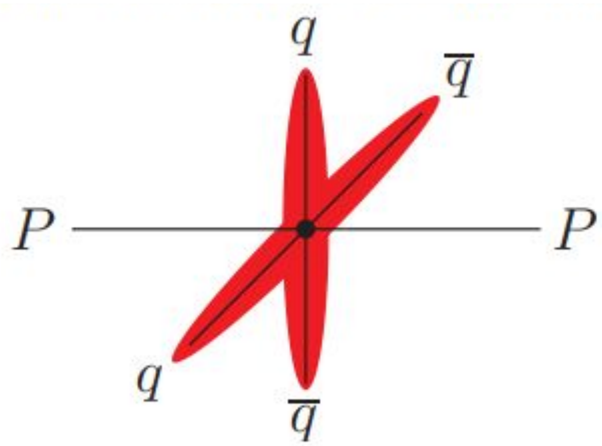
Tuning to LHC data

Herwig7 + Lund String Model
with Colour Reconnection

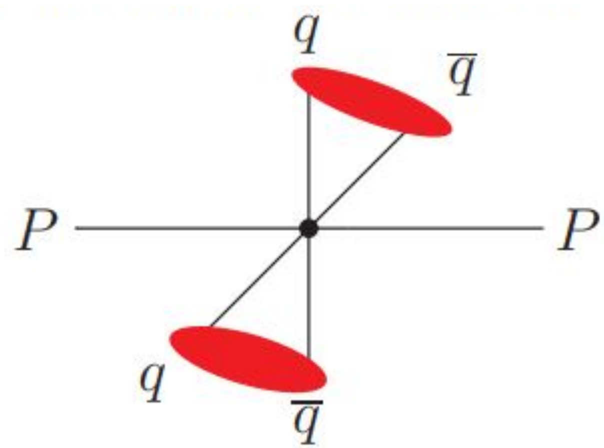
Colour Reconnection

- **Colour Reconnection (CR)** reshuffles the colours just prior to hadronization such that strings form between favourable partons
- This is needed to explain the rise of $\langle p_T \rangle$ with multiplicity as observed experimentally
- We implemented **CR** in the **P8I** interface and switched it on when tuning to LHC data which is particularly sensitive to it due to large partonic activity in the final state

Before Colour Reconnection



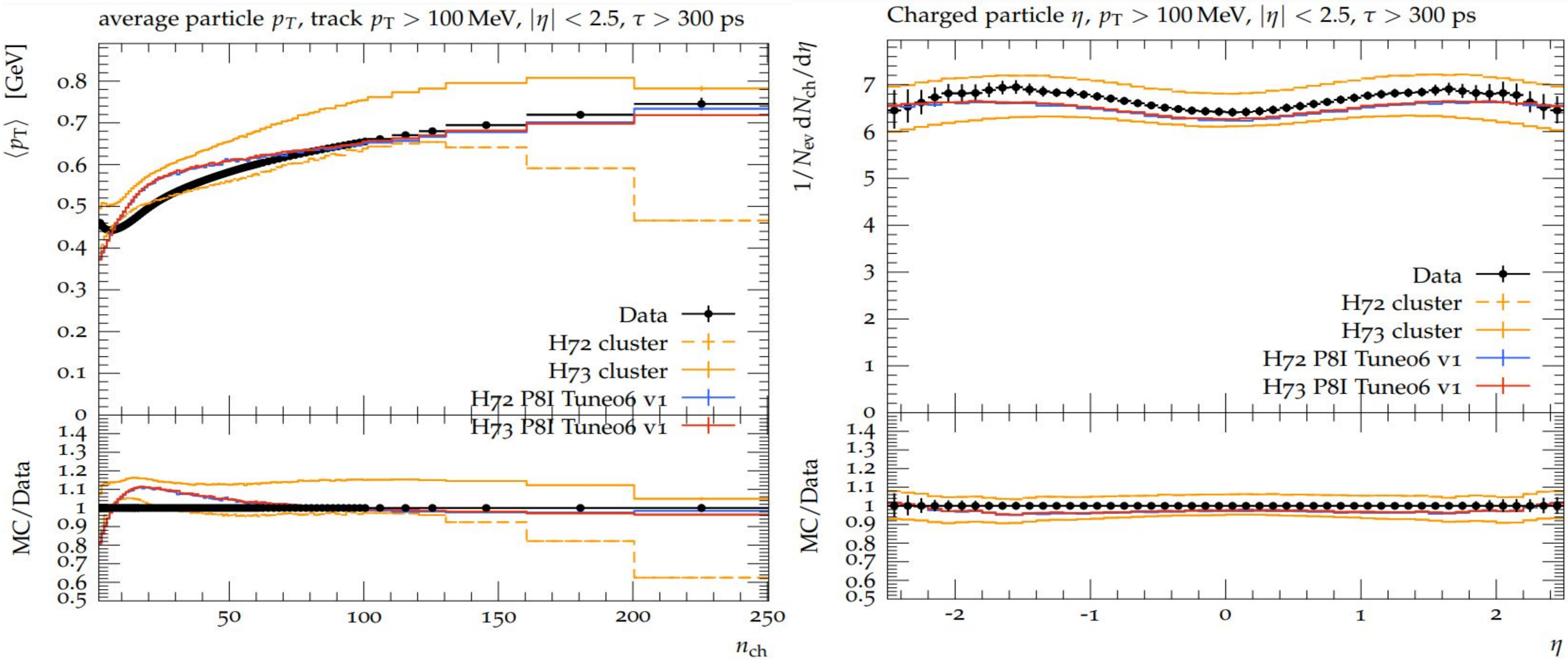
After Colour Reconnection



Images taken from J.R. Christiansen

Tuning to LHC Data (at 13 TeV)

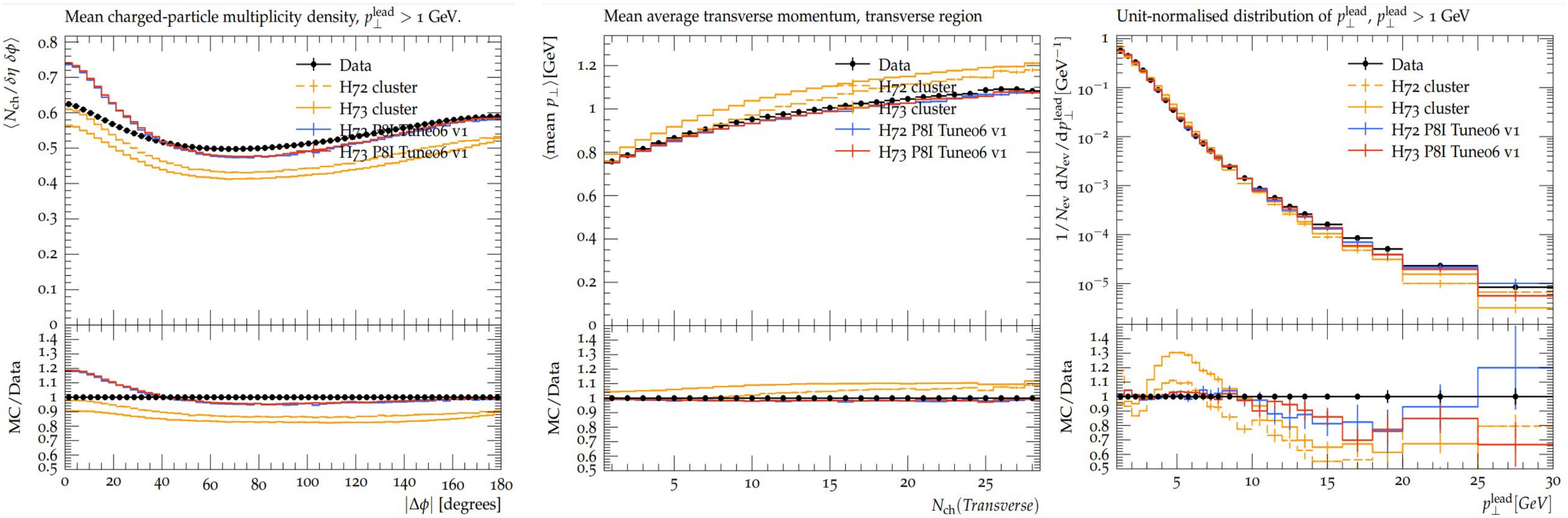
- Measurements from ATLAS



- Average particle pT distribution (left) well modelled in the high multiplicity region by our tune compared to default cluster tune
- Better modelling of charged particle pseudorapidity as well (right)

Tuning to LHC Data (at 13 TeV)

- Measurements from ATLAS (Track based underlying events)



- More examples of observables well modelled by our tune over large regions in multiplicity and mean transverse momenta distributions

Summary

- We obtain a robust setup for using both strings and cluster within Herwig for systematic studies of Non-perturbative effects.

- Our tune does significantly better than AutoTunes and does reasonably well compared to the default cluster tunes for both LEP and LHC data.

- Tunes were done with Herwig 7.2.2 initially but have been extended and validated with the latest unreleased version of Herwig 7.3 (will be out soon) as well.

- We do see some tensions in some of the flavour sensitive observables which may need model development rather than tuning.

Thank you!

Backup Slides

Parameter Values (MAP Tune v1)

Pythia8 parameters:

Fragmentation related Parameters

StringZ_aLund	0.750043
StringZ_bLund	0.898493
StringZ_sigma	0.309940
StringZ_aExtraSQuark	0.175699
StringZ_bExtraDiquark	0.053576
StringZ_rFactC	0.680046
StringZ_rFactB	1.265325

Flavour related Parameters

StringFlav_probStoUD	0.190037
StringFlav_probQQtoQ	0.079770
StringFlav_probSQtoQQ	0.998577
StringFlav_probQQ1toQQ0	0.022243
StringFlav_etaSup	0.512356
StringFlav_etaPrimeSup	0.184657
StringFlav_popcornRate	0.734028
StringFlav_mesonUDvector	0.329842
StringFlav_mesonSvector	0.676458
StringFlav_mesonCvector	1.065635
StringFlav_mesonBvector	1.849178

Herwig7 parameters:

/Herwig/Shower/AlphaQCD:AlphaIn	0.125772
/Herwig/Shower/NLOAlphaS:input_alpha_s	0.125772
/Herwig/Shower/PTCutOff:pTmin	1.027660

Parameter Values (MAP Tune v2)

Pythia8 parameters:

Fragmentation related Parameters

StringZ_aLund	0.394348
StringZ_bLund	0.688822
StringZ_sigma	0.308203
StringZ_aExtraSQuark	0.863354
StringZ_bExtraDiquark	1.901871
StringZ_rFactC	0.522541
StringZ_rFactB	1.577757

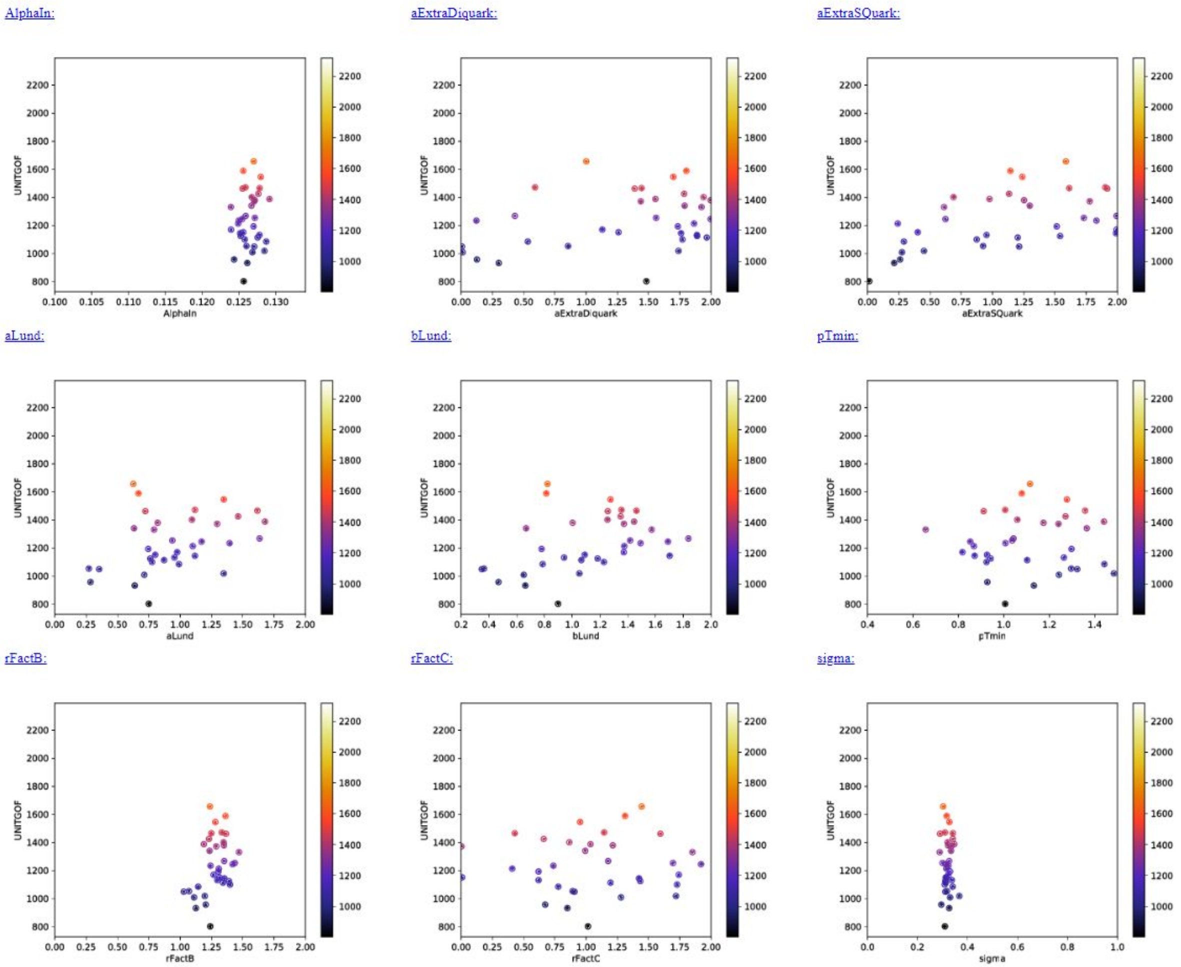
Flavour related Parameters

StringFlav_probStoUD	0.185213
StringFlav_probQQtoQ	0.076194
StringFlav_probSQtoQQ	0.998577
StringFlav_probQQ1toQQ0	0.063338
StringFlav_etaSup	0.537677
StringFlav_etaPrimeSup	0.135651
StringFlav_popcornRate	0.002145
StringFlav_mesonUDvector	0.330021
StringFlav_mesonSvector	0.688189
StringFlav_mesonCvector	1.151696
StringFlav_mesonBvector	2.257771

Herwig7 parameters:

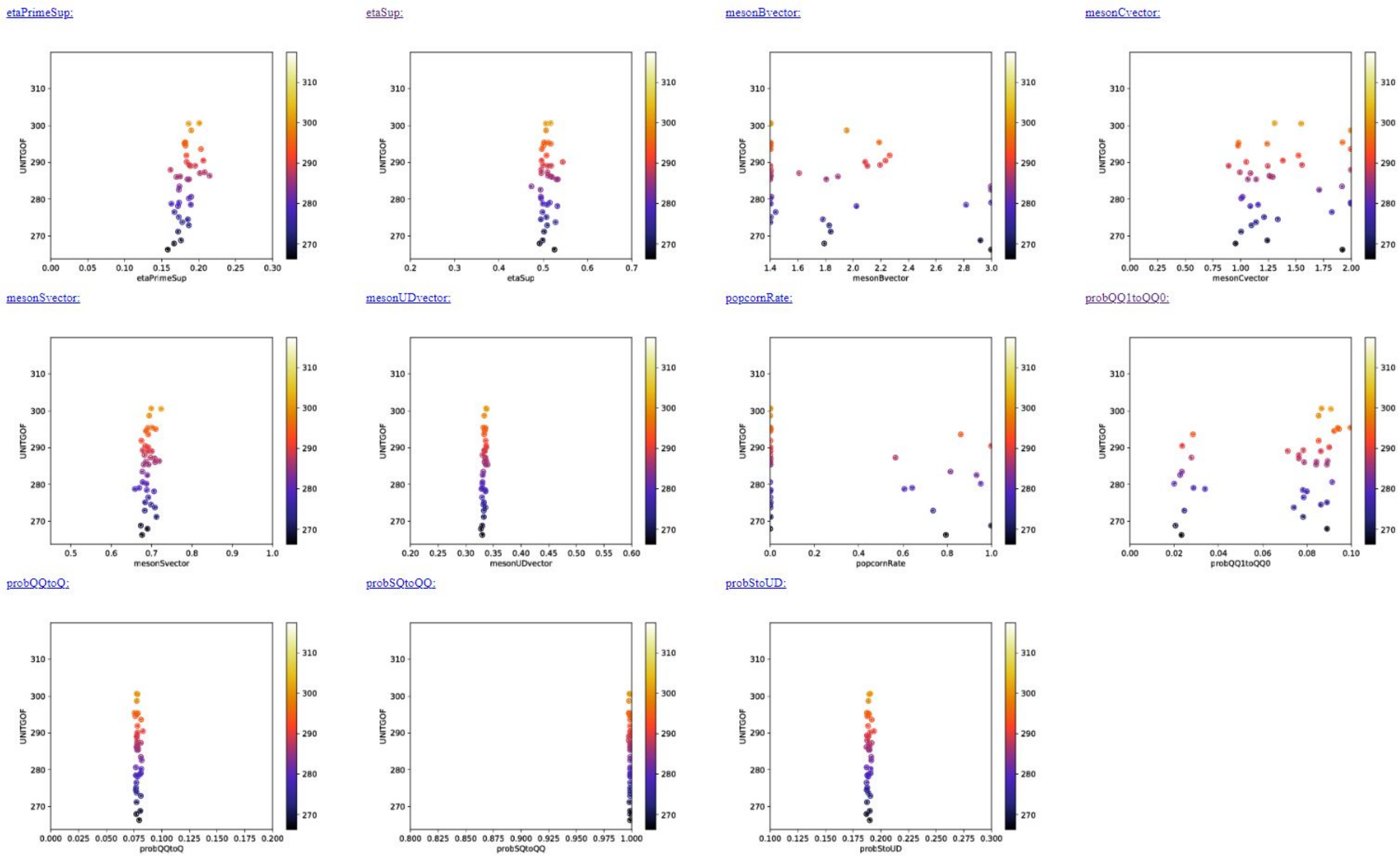
/Herwig/Shower/AlphaQCD:AlphaIn	0.124523
/Herwig/Shower/NLOAlphaS:input_alpha_s	0.124523
/Herwig/Shower/PTCutOff:pTmin	0.894570

Scatter plots (Fragmentation parameters)



- Parameters which are constrained in a narrow range are fixed and the tune is run again until maximum number of parameters are fixed.
- This reduces the number of parameters to be tuned after each step.

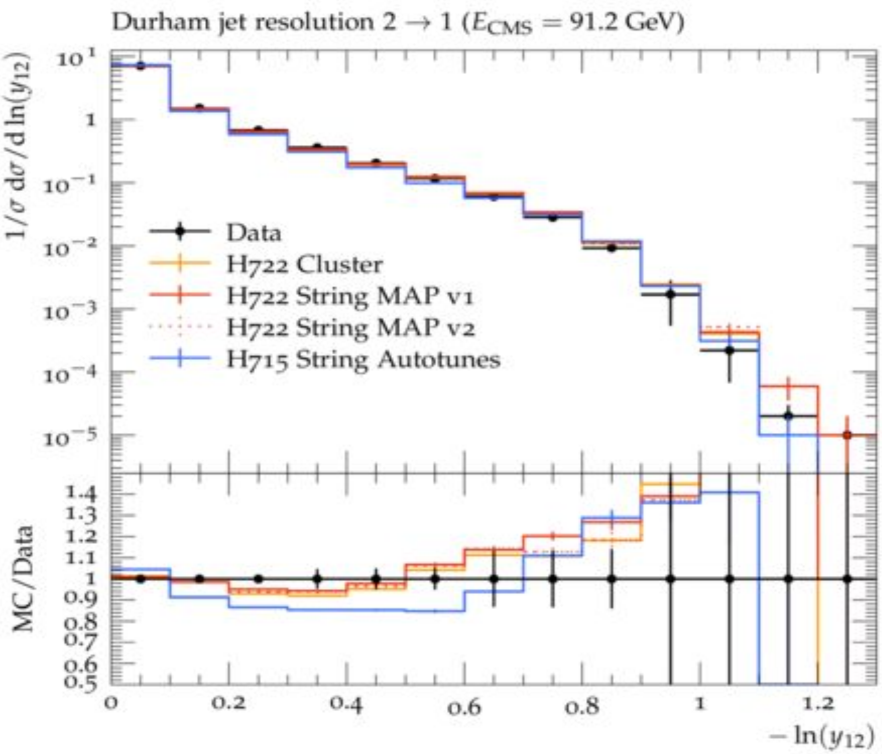
Scatter plots (Flavour Parameters)

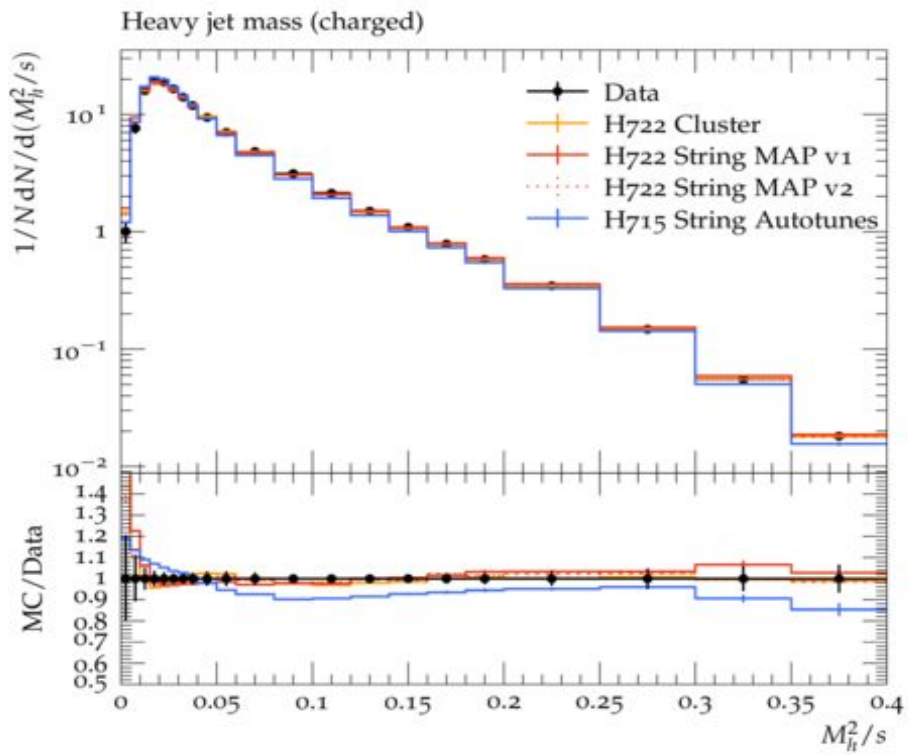


- Many of the flavour parameters are strongly constrained which are fixed and tuned.
- The probQQ1toQQ0 parameter shows interesting constraints near two values, but converges to the lower value for MAP tune v1 and the higher value for MAP tune v2 when other parameters are fixed.

Jet Observables (Tune for Herwig 7.2.2)

- Measurements from ALEPH

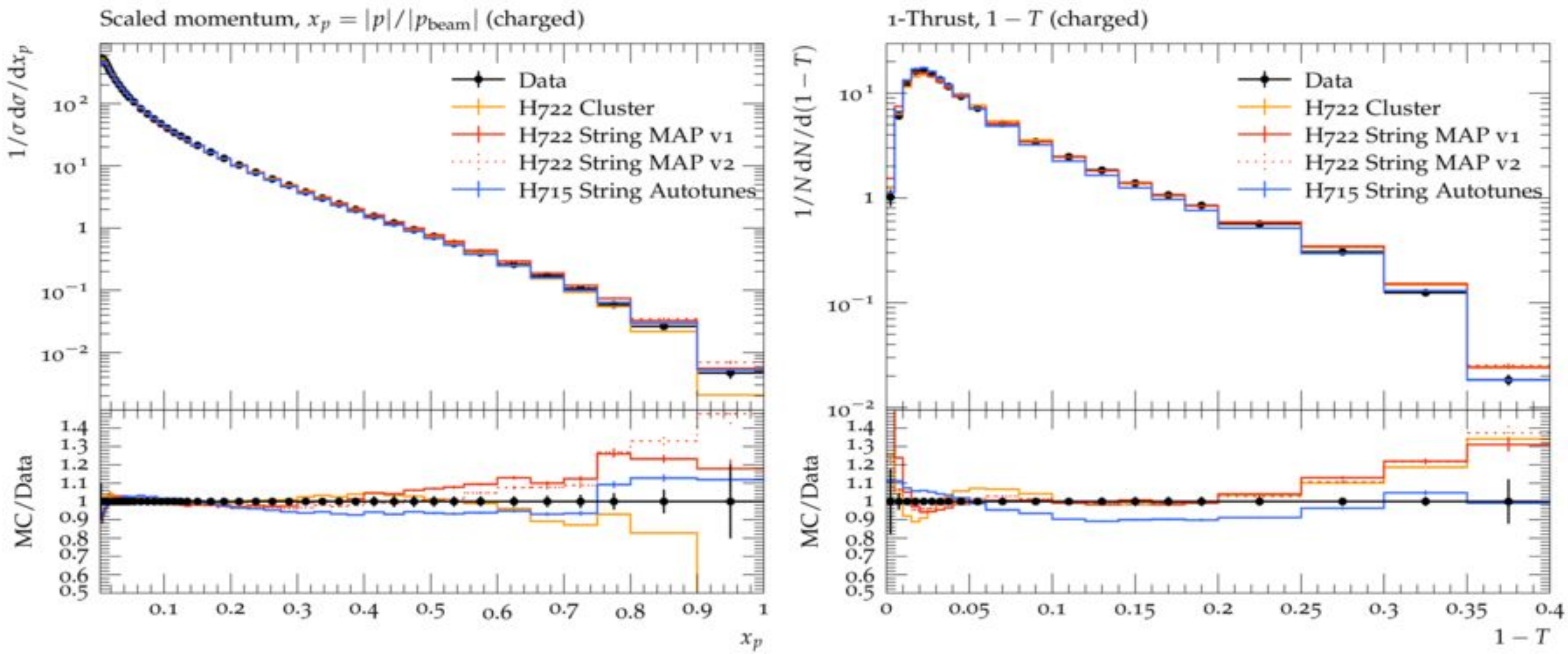




- Both these distributions were **not** used in the tune.
- MAP tune still does better than AutoTunes tune in these distributions and does similar to the default tune - shows robustness of our tune.

Event Shapes (not used in tuning but doing well for Herwig 7.2.2 tune)

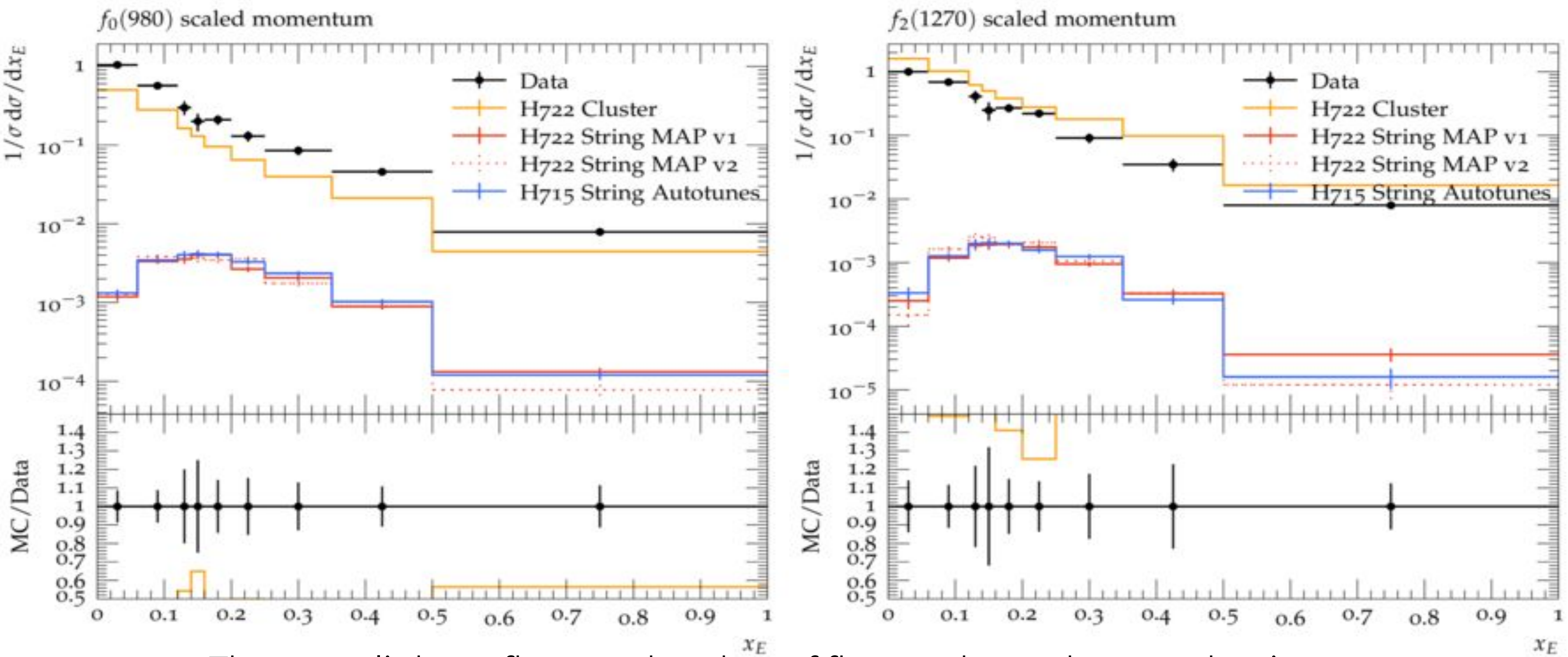
- Measurements from ALEPH



- Both these distributions were **not** used in tune.
- MAP tune does relatively well even though these distributions are no weighted in the tune
- For x_p distribution, MAP tune does better than the default but the AutoTunes is better near the tail.

Flavour sensitive Observables

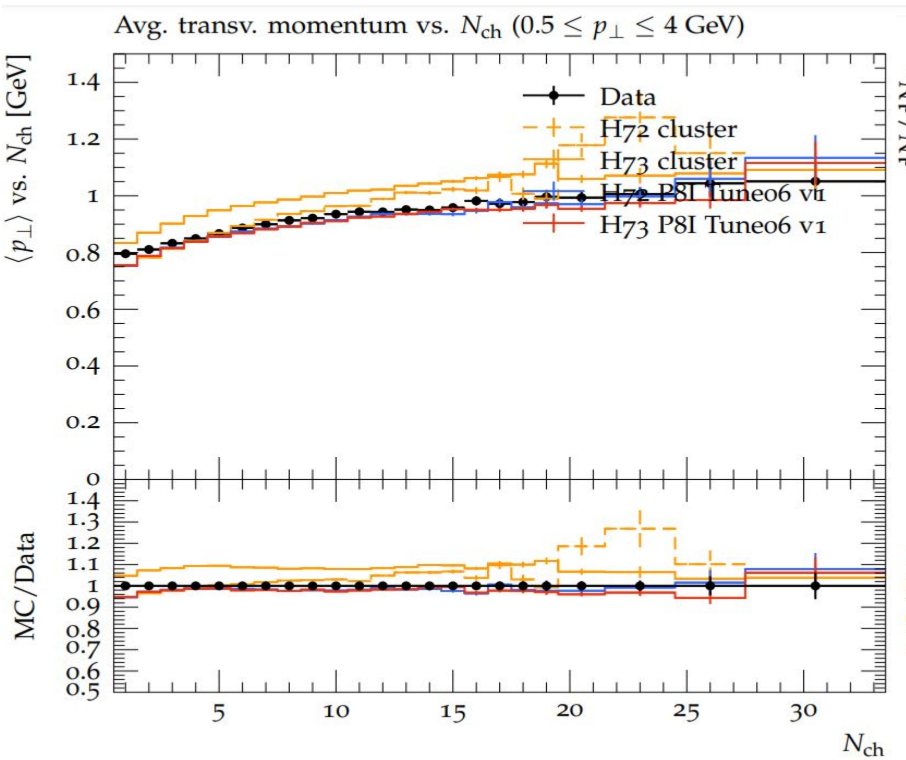
- Measurements from OPAL

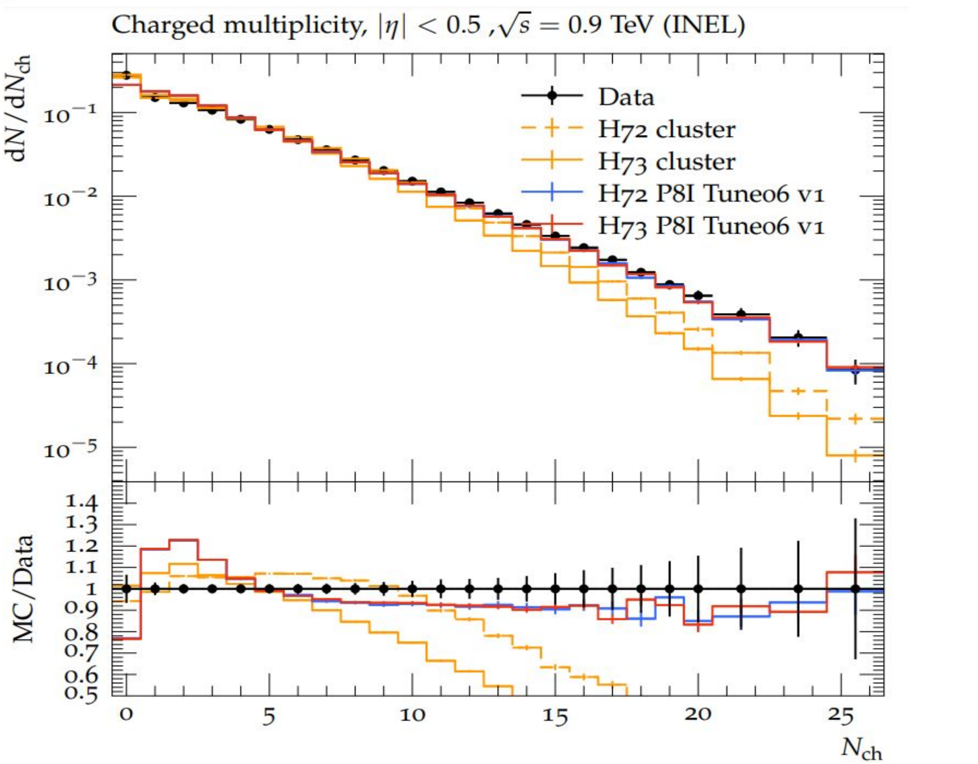


- These are light unflavoured probes of flavour dependent production.
- These distributions were **not** used in the tune
- It would also be interesting to see how Pythia does for the observables.
- The cluster model could be insensitive to the flavour as it does reasonably well in these distributions.

LHC data at 900 GeV

- Measurements from ALICE

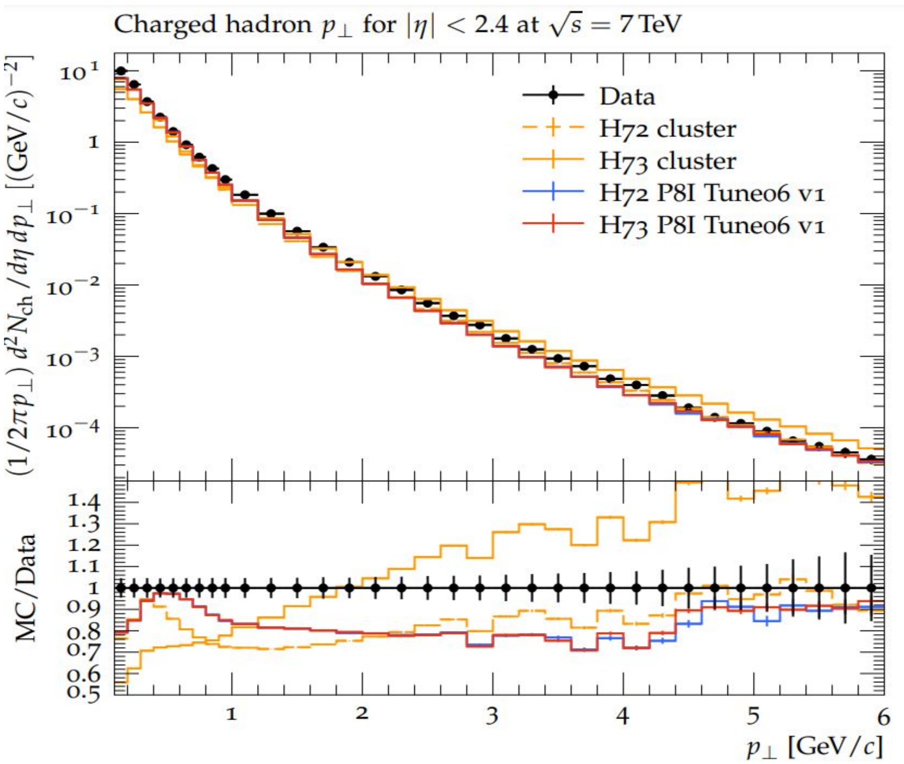


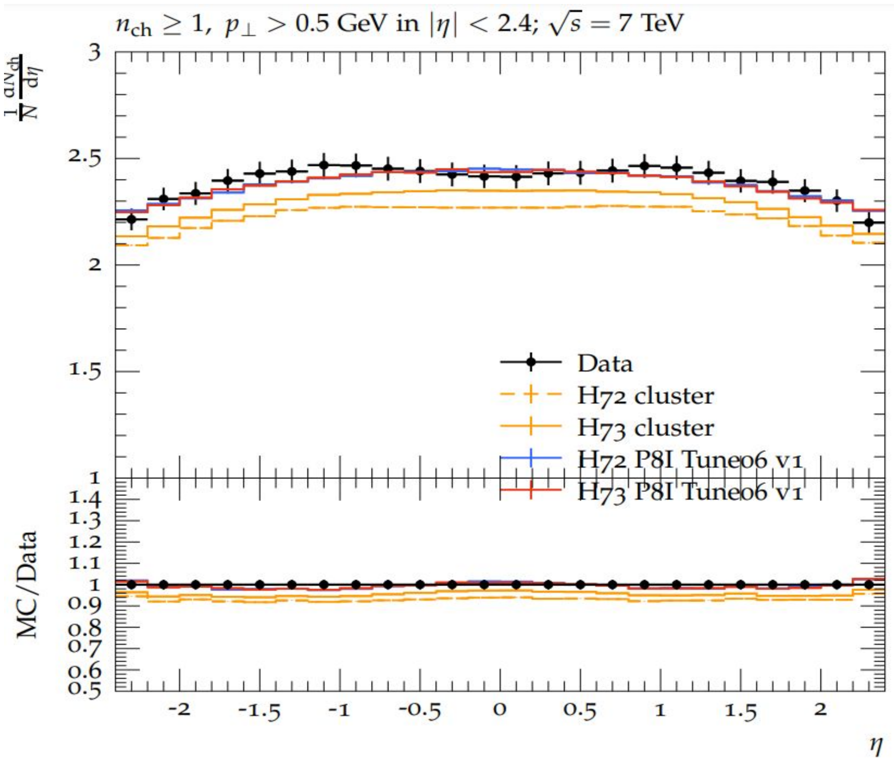


- Our tune doing well at different energies

LHC data at 7 TeV

- Measurements from CMS





- Overall improvements have been made to the default cluster tunes with our tunes for a large set of observables