High multiplicity processes with BlackHat and Sherpa

Daniel Maître
IPPP, Durham


WHO RULES YOU?
sky
dish
ПРАВДА
BBC
abc
BLAKHAT

Outline

- Ntuples
- Applications
  - Universality
  - FastNLO
  - MiNLO

n-Tuple files [arXiv:1310.7439]

- High multiplicity NLO calculations are computationally intensive
- It would not be possible to rerun a high multiplicity W+jet calculation every time a new interesting observable comes up
- Matrix elements are expensive, while
  - Jet clustering
  - Observables
  - PDF evaluation

  are relatively cheap
- Store each matrix element, PS point and the information necessary to change the factorisation and renormalisation scales in large files we call n-Tuple files
- We use ROOT file as storage

n-Tuple files [arXiv:1310.7439]

- Goodies
  - One can change the analysis cuts, add observables
  - Scale variation
  - PDF errors (otherwise extremely expensive)
  - Easy communication between theorists and experimenters
  - No need for specific know-how of the tool which produced the NLO calculation
  - Easier to “endorse” an event file than a program
- Price to pay
  - Large files
  - Generation cuts need to be loose enough to accommodate many analysis → efficiency cost

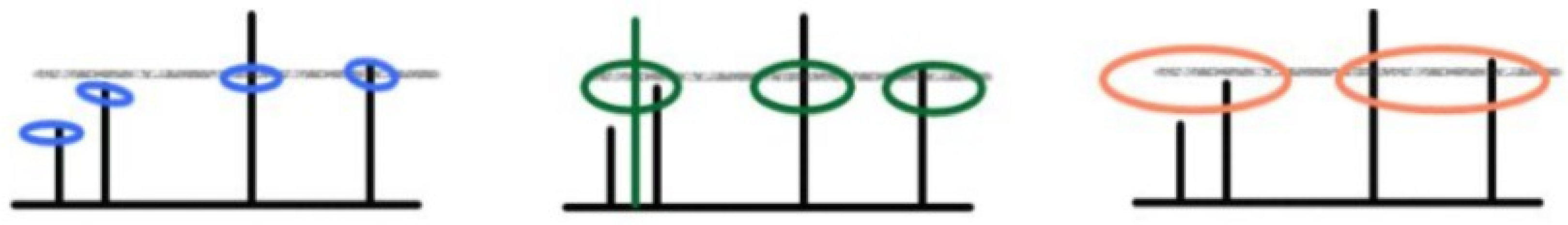
nTupleReader library [arXiv:1310.7439]

- We provide a C++ library to facilitate the use of the n-Tuple files
- Allows:
  - Change of factorisation and renormalisation scales
  - Change of pdf (from LHAPDF set), including error sets
- Has a Python interface
- Template for a customised implementation
- Available on hepforge

  https://blackhat.hepforge.org/trac/wiki/BlackHatSherpaNtuples

NLO event files

- Issues with NLO event files
  - Only fixed number of radii are supported



  - Negative weights
  - Highly anti-correlated weights (by construction!) affect
    - Integration error estimate
    - Rebining and cumulative distributions

NLO events

- Standard formula for integration error:

$$\begin{aligned}\sigma^2 &= \langle(\omega - \langle\omega\rangle)^2\rangle = \langle\omega^2\rangle - (\langle\omega\rangle)^2 \\ &\simeq \frac{\sum \omega_i^2}{N} - \left(\frac{\sum \omega_i}{N}\right)^2 = \frac{1}{N}\left(\sum \omega_i^2 - \frac{(\sum \omega_i)^2}{N}\right)\end{aligned}$$

- If we have only pairs of large anti-correlated weights

$$\omega_+ \simeq L\,,\ \omega_- \simeq -L\,, \qquad \omega_+ + \omega_- \simeq s$$

$$\sigma^2 \simeq \frac{1}{N}\left(NL^2 + \frac{(Ns)^2}{N}\right) = L^2 - s^2 \simeq L^2$$

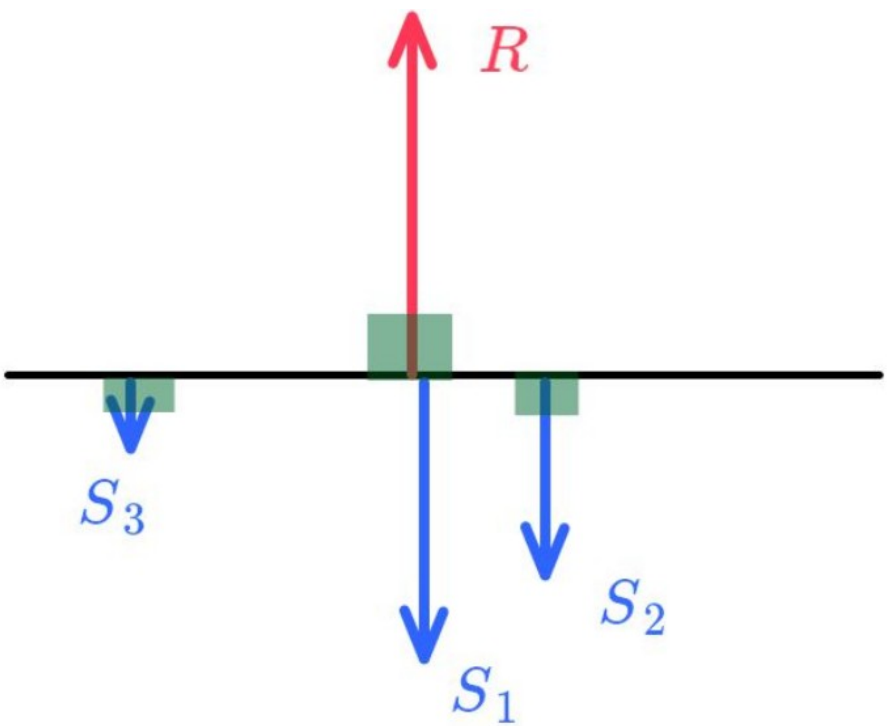
- This is far too large, dependent on technical cutoff, one should use

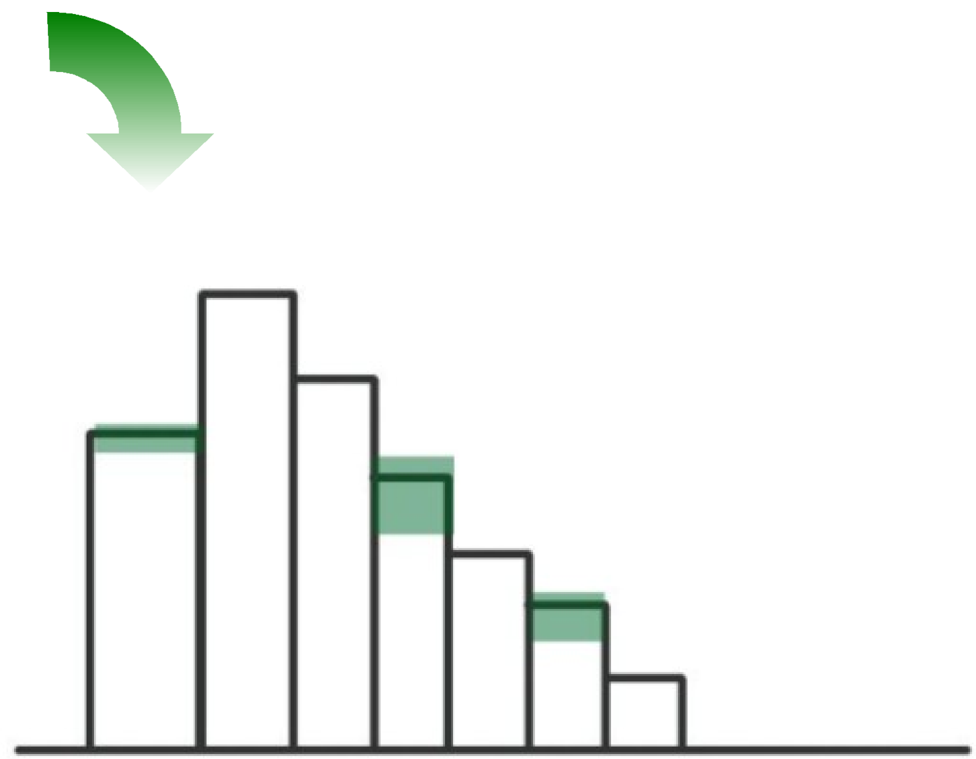
$$\sigma^2 \simeq \frac{1}{N}\left(\sum(\omega_{i,+} + \omega_{i,-})^2 - \frac{(\sum \omega_i)^2}{N}\right)$$

instead

NLO events

- First bin all contribution from the event in auxiliary histogram





- Then transfer to the true histogram

(near) Future

- Rivet is an analysis framework developed to facilitate the validation of MC with experimental data
- Many experimental analyses provide a Rivet analysis
- So far Rivet could not easily be used to analyse NLO predictions because correlated weights were not treated correctly. [Buckley,Grellscheid,Pollard,Schulz,Siegert]
- New version of Rivet will allow for the correct treatment of
  - Correlated events
  - Weight vectors (pdf errors)
- It will be easier to compare NLO predictions with data

Analysis program

- With the new features of Rivet nTuple files can be used as input for the analysis
- New nTupleReader library contain the necessary ingredients to integrate with these new features
- Allows to run

```
> analyse   \
>     RIVET_ANALYSIS \
>     nTupleFile1.root  \
>     nTuplesFile2.root...
```

```
int main(int argc,char *argv[]){

  Rivet::AnalysisHandler rivet;
  Rivet.addAnalysis(argv[1]);

  nTupleReader r;

  for (int ii=2;ii<argc;ii++){
    r.addFile(argv[ii]);
  }

  r.setPDF("CT10nlo.LHgrid");
  r.setCMSEnergy(7000);
  r.setColliderType(pp);

  while(r.nextEntry()){
    HepMC::GenEvent evt;
    r.setHepMC(evt);
    rivet.analyze(evt);
  }

  rivet.setCrossSection(r.getCrossSection());
  rivet.finalize();
  rivet.writeData(std::string(argv[1])+".yoda");
}
```

nTupleFiles

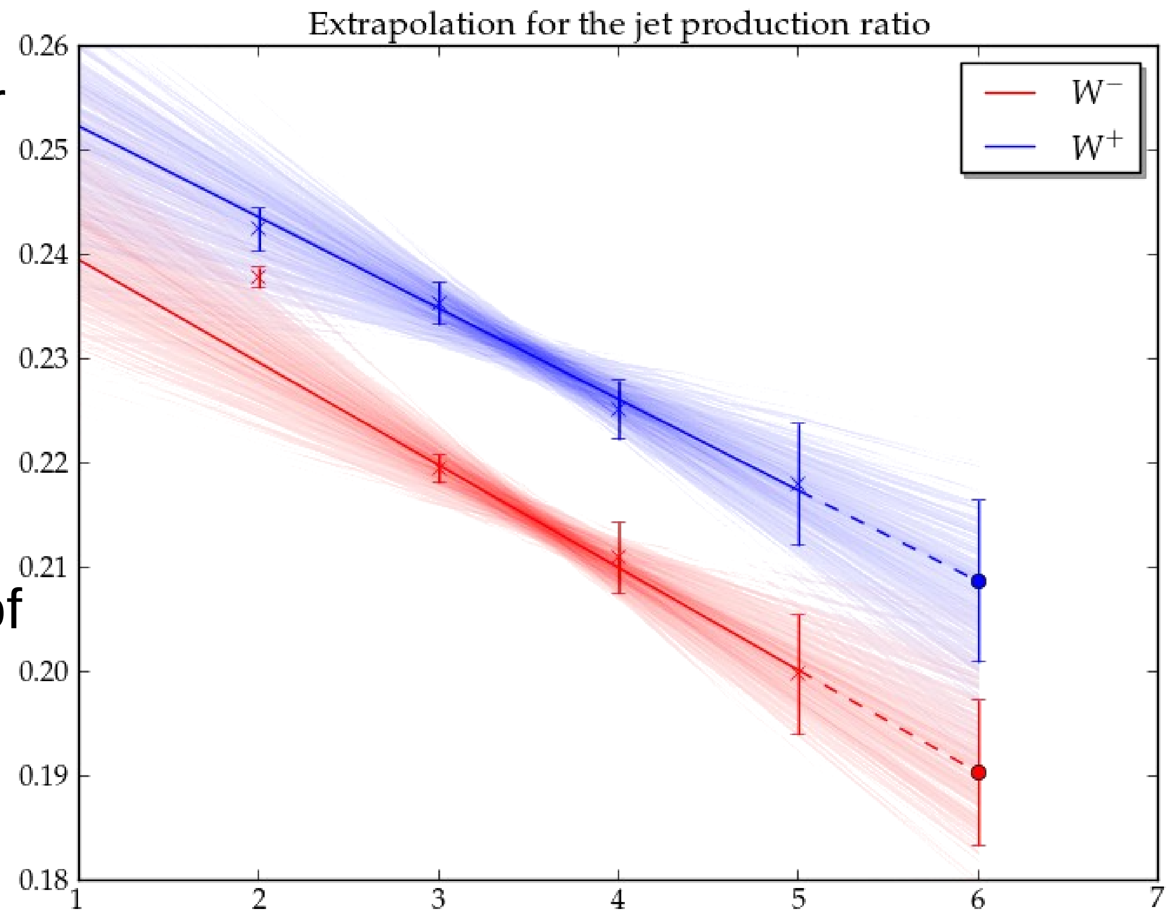
Ntuple-files are your pocket NLO calculation

Towards higher multiplicities?

- We have a lot of prediction for high multiplicity processes at NLO
- We can try to find 'universal' properties/features
- Usually need to discard 0-jet and 1-jet because new partonic channels open
- Usually these features are more easily seen in ratios between multiplicities
- [Bern,Dixon,Febres Cordero,Höche,Ita,Kosower, Maître; arXiv:1412.4775]

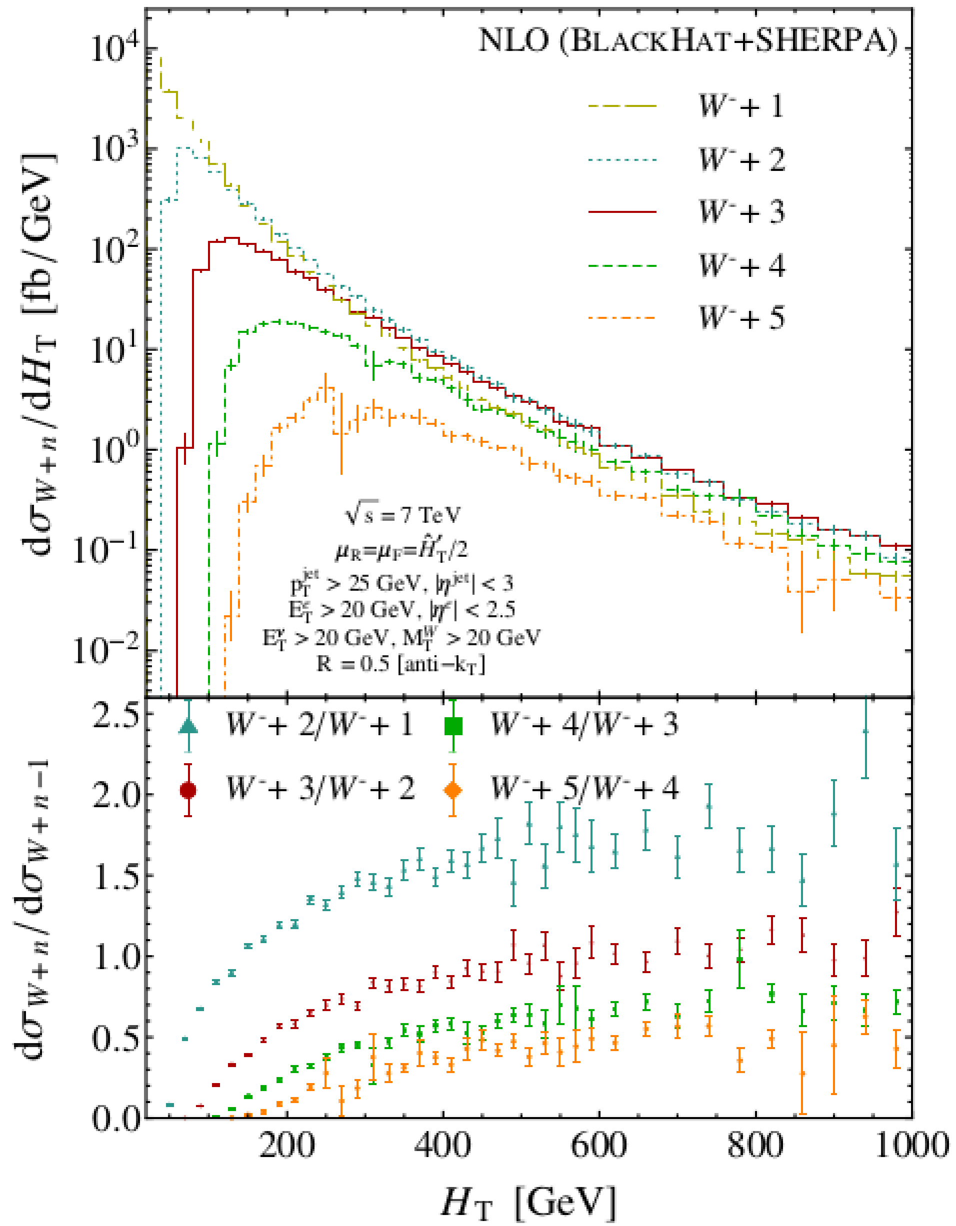
Extrapolation for ratios

- Ratio V+n jets/(V+n-1 jets)
- Consistent with straight line for n>2
- Use extrapolation for 6 jets:
- W^- : 0.15 ± 0.01 pb
- W^+ : 0.30 ± 0.03 pb
- Consistent with extrapolation of charge asymmetry
- Error estimates through Monte Carlo method



Distributions

- What about distributions?
- Look at sum of transverse energies of the jets (HT)
- Cannot extrapolate the value of each bin separately
  - Statistical errors are too large
  - Different thresholds
  - Different peak positions



arXiv:1412.4775

Distributions

- Instead find a parametrisation and extrapolate the parameters of the parametrisation
- Ansatz for the HT distribution:

$$\frac{d\sigma_{V+n}}{d\,H_T} = \left(\frac{N_C\alpha_s}{2\pi}\right)^n f(H_T)\mathcal{N}_n \ln^{\tau_n}\rho_{H,n}\left(1 - H_T/H_T^{\max}\right)^{\gamma_n}$$

$$\rho_{H,n} = H_T/(np_T^{min})$$

HT Distribution

- Instead find a parametrisation and extrapolate the parameters of the parametrisation
- Ansatz for the HT distribution:

$$\frac{d\sigma_{V+n}}{d\,H_T} = \left(\frac{N_C\alpha_s}{2\pi}\right)^n f(H_T)\mathcal{N}_n \ln^{\tau_n}\rho_{H,n}\left(1 - H_T/H_T^{\max}\right)^{\gamma_n}$$

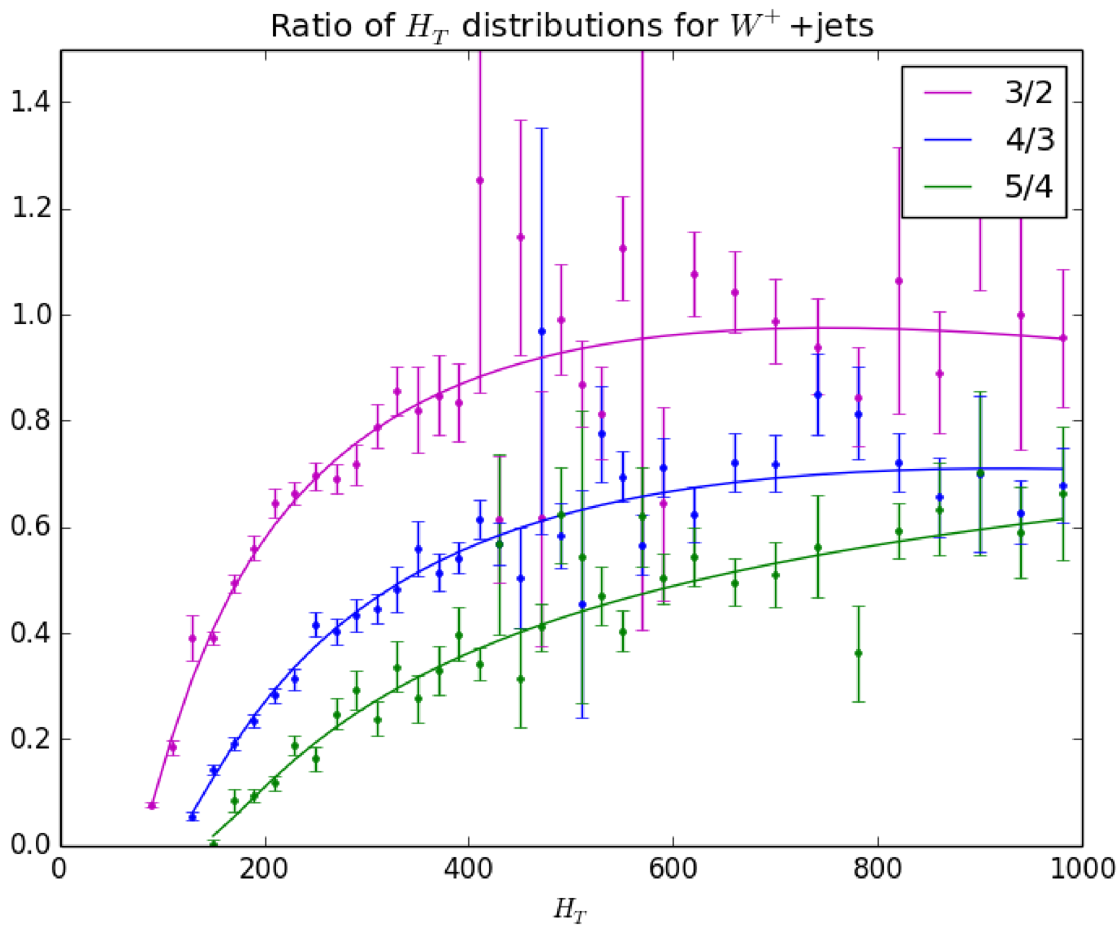
Independent of *n*

parameters

$$\rho_{H,n} = H_T/(np_T^{min})$$

HT distribution

- Fit ratios to get the parameters



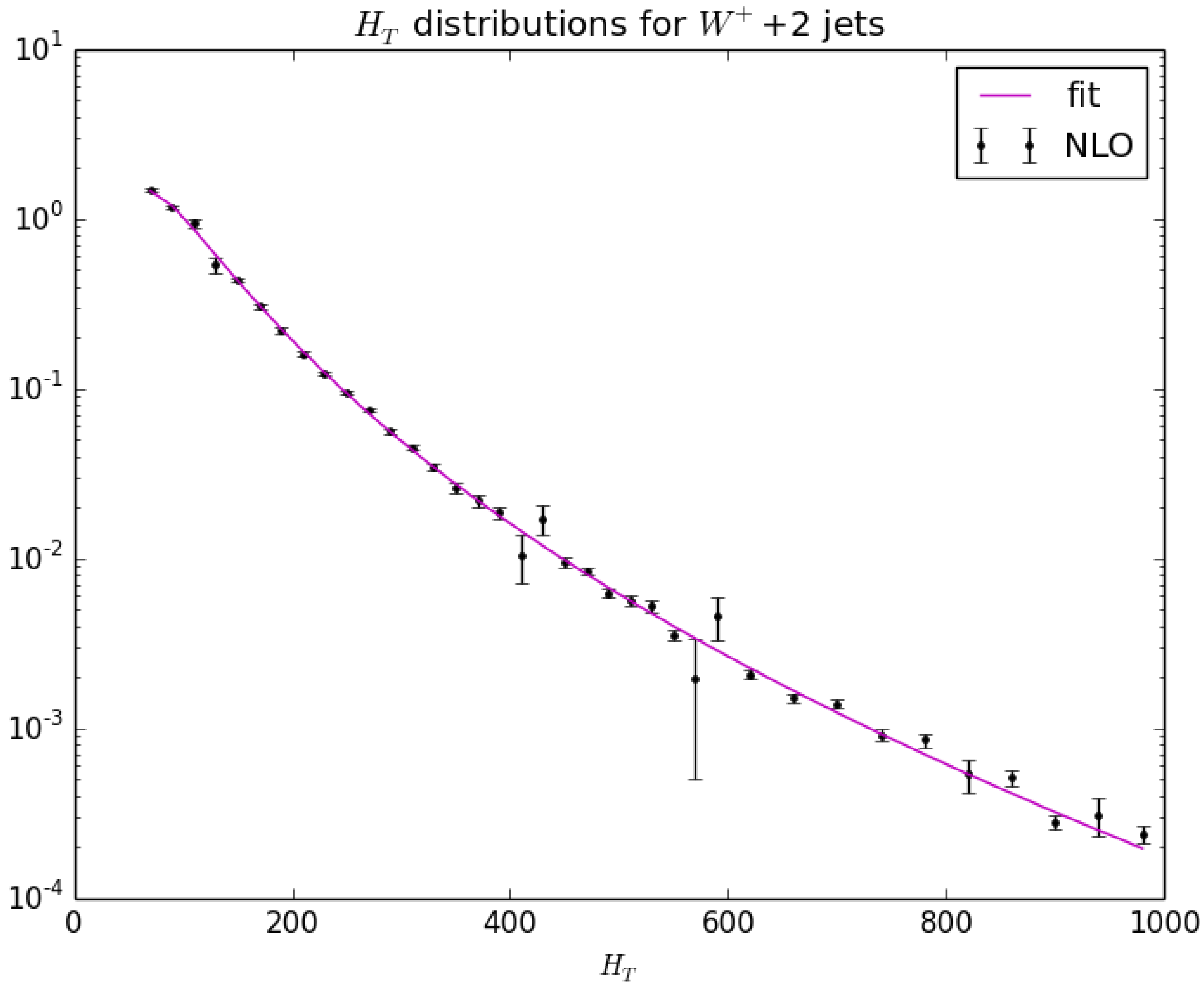
HT distribution

- With the parameters one can extract *f(H)* from a distribution
- But it is more convenient to have an analytical form for it
- We can use the following form

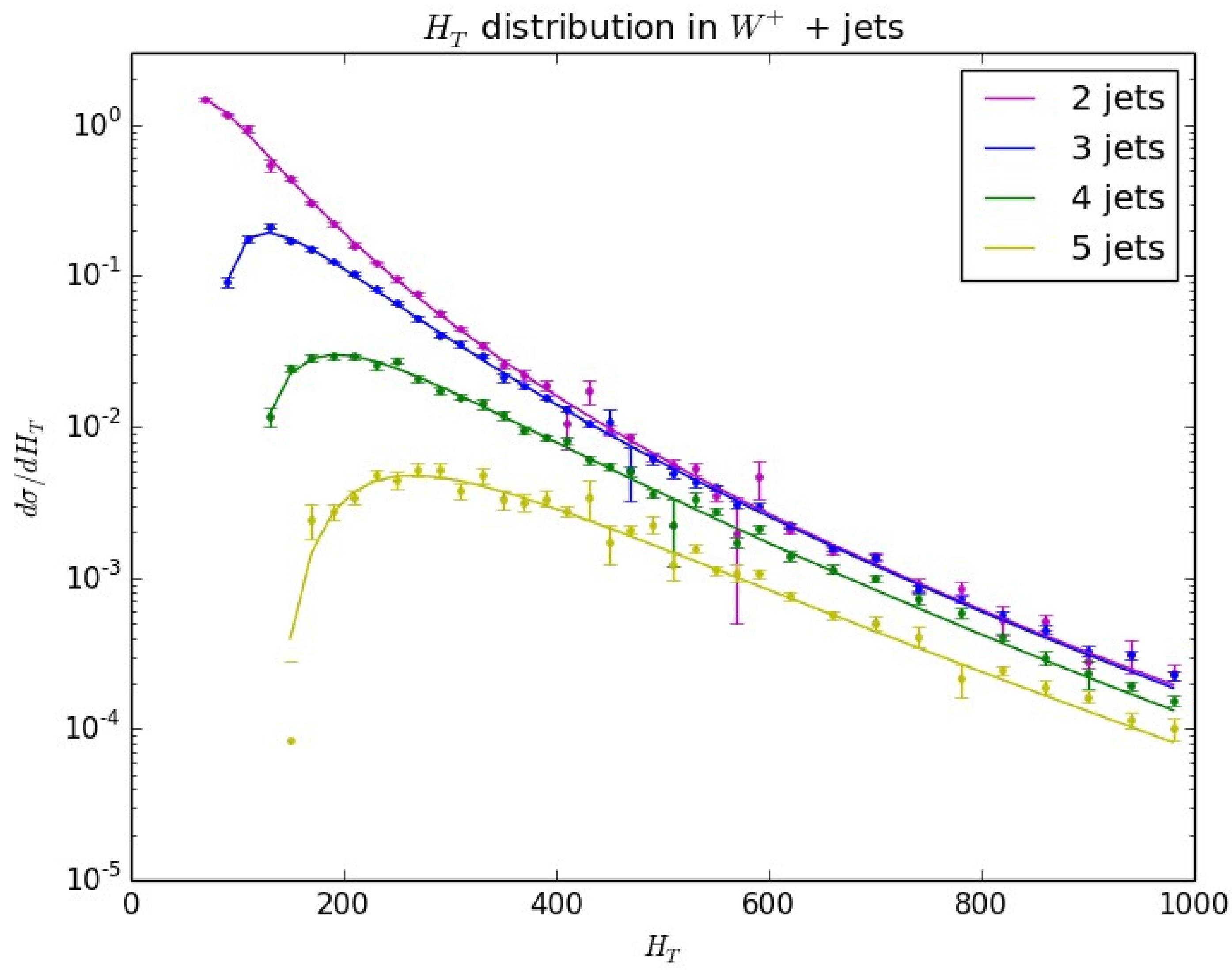
$$f(H) = c \ln^r(H/10) \left(\frac{H}{2p_T^{\min}} \right)^{\omega_2} e^{-h_* H} \,,$$

and extract the parameters from the W+2 jets distribution

HT distribution

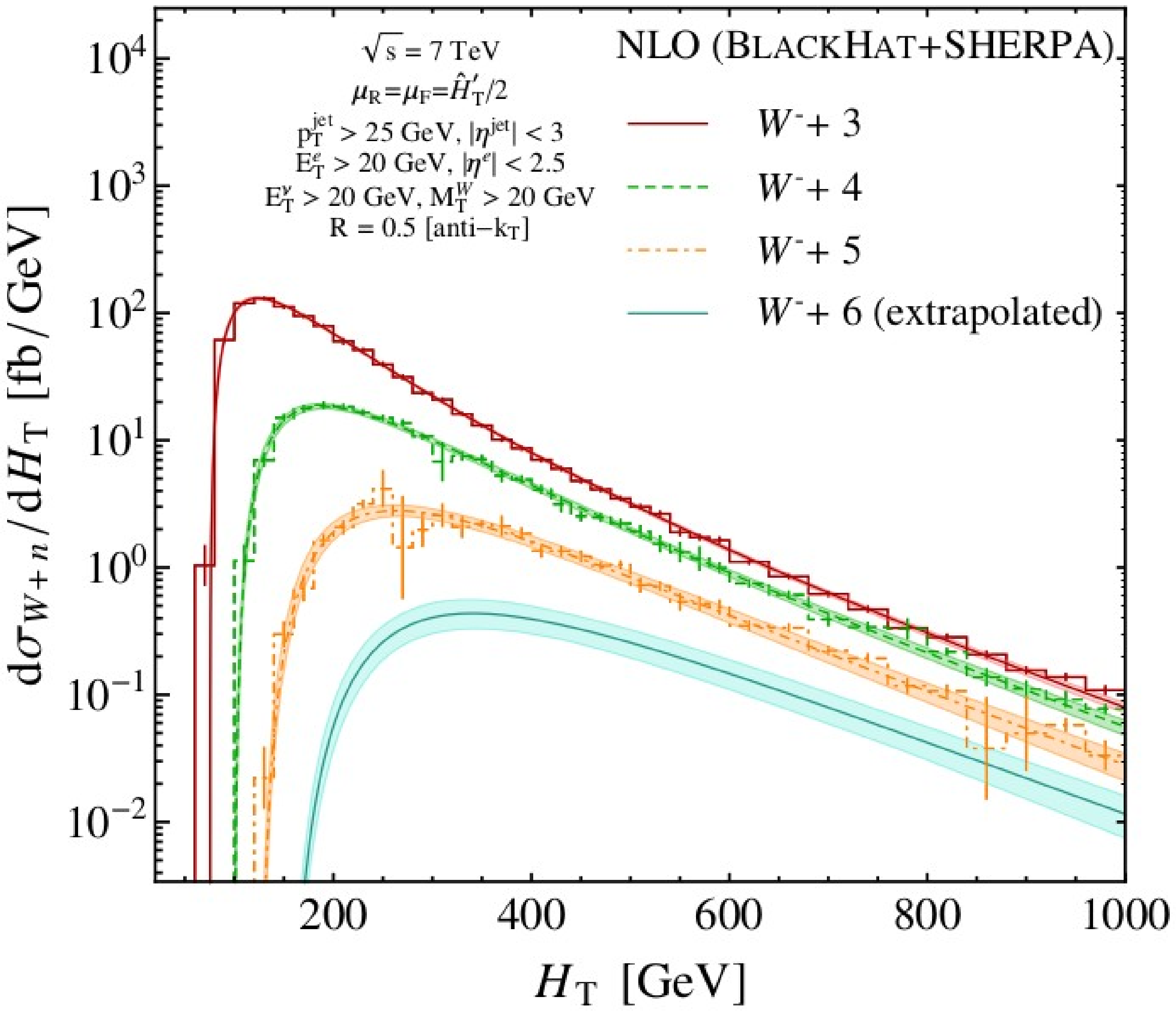


HT distribution



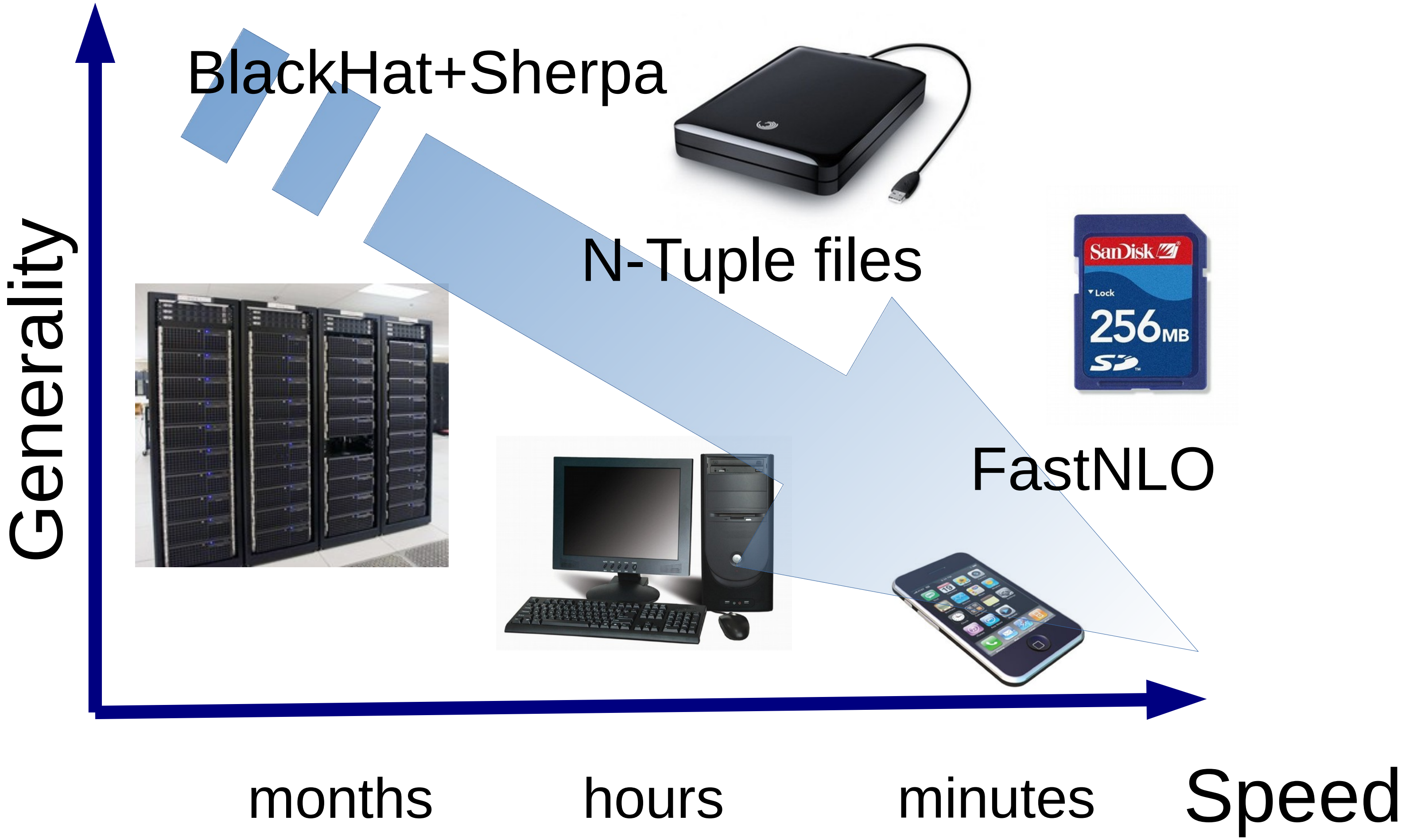
Distributions

- Extrapolated HT distribution
- Uncertainty bands are estimated using a MC method



arXiv:1412.4775

Speed vs Generality

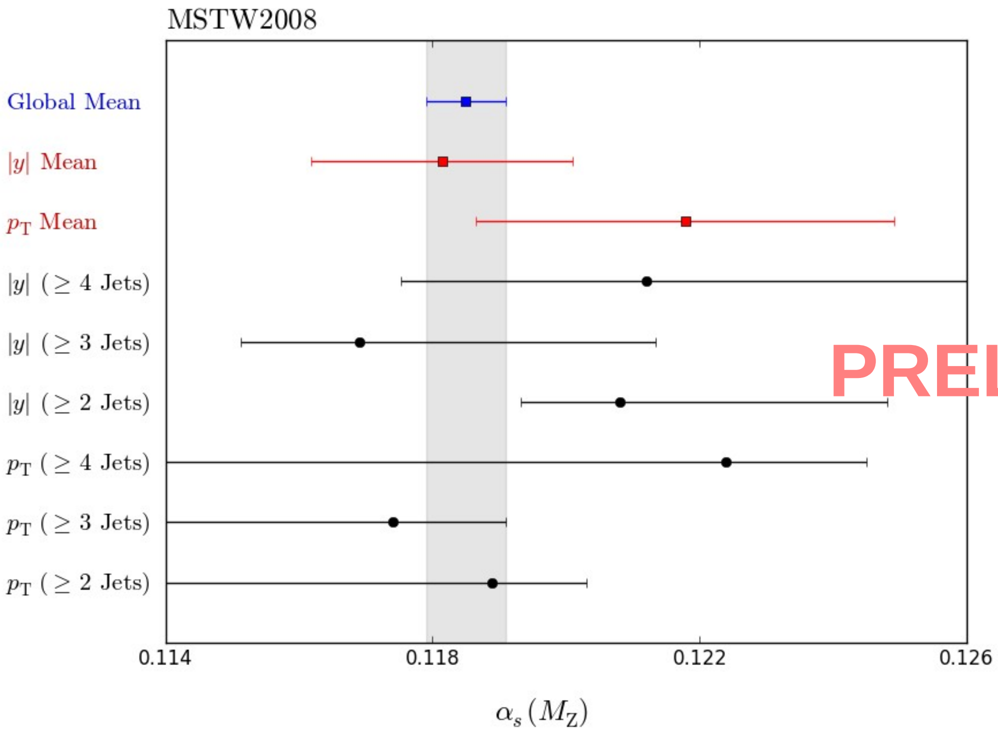


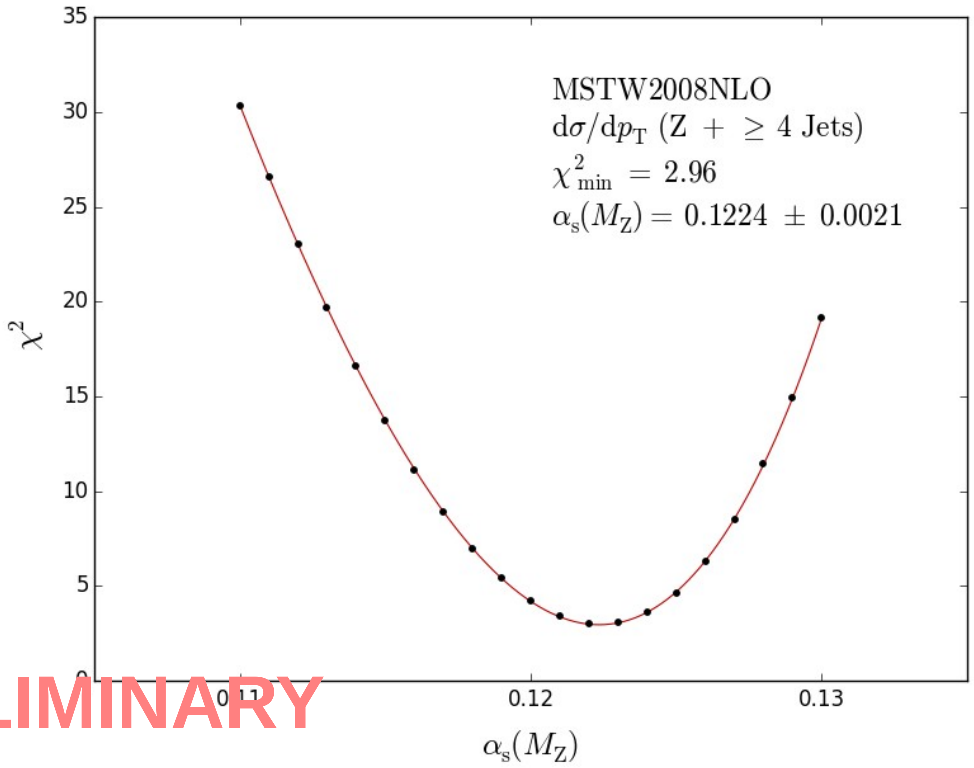
fastNLO

- A direct analysis using n-Tuple files can still be too slow for some applications
  - PDF4LHC-type error require several PDF sets with errors
  - alphas determination
  - PDF fitting
- We can use the n-Tuple files to create fastNLO tables for very quick evaluation
  - Very fast
  - Need one table for each measured distribution
- Similar to the APPLgrid and other projects

Strong coupling determination

- Use Atlas Z+2,3,4 jets data to extract the value of the strong coupling constant
- Use PDF fit for various values of $\alpha(M_Z)$



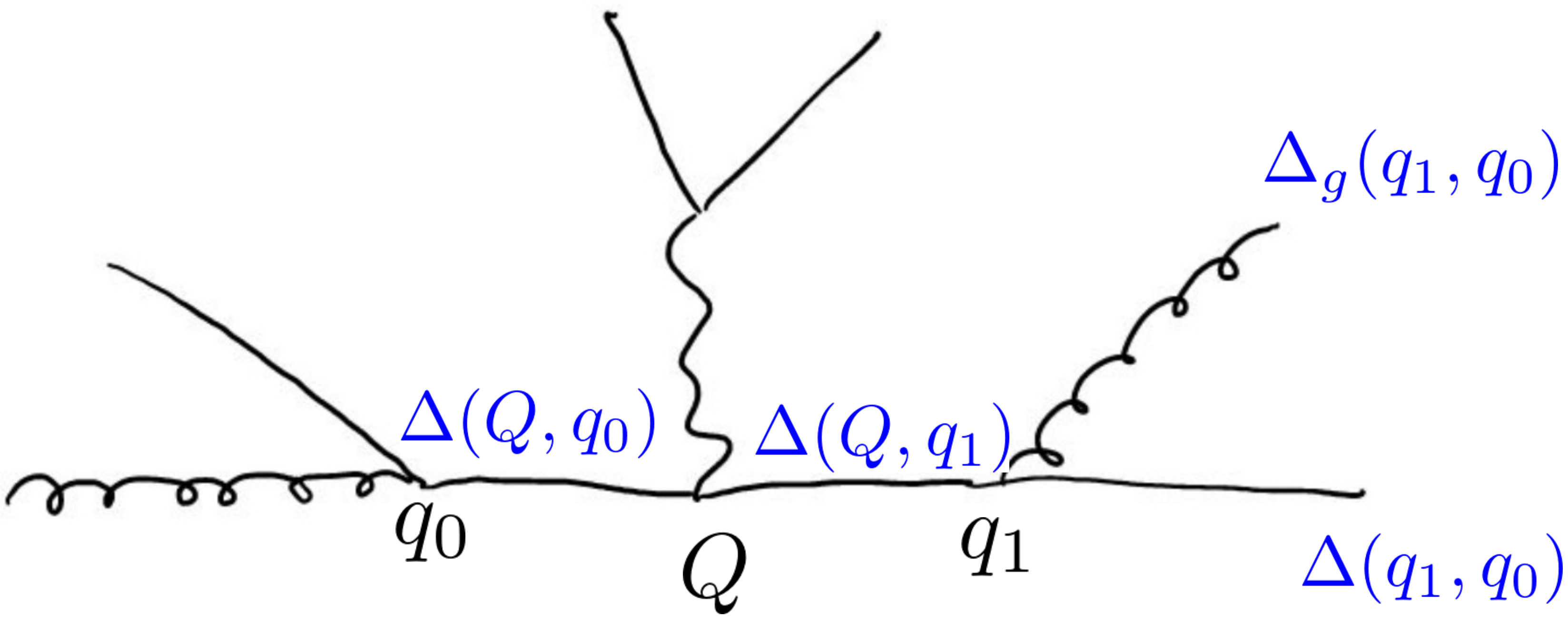


[Plots from Mark Johnson]

MiNLO scale setting

- Having nTuple files allows to experiment with scale settings
- MiNLO method allows to evaluate NLO cross sections where such a cross section would diverge.
- The basic idea is to recluster the final state and use the nodal scales for the coupling constant and to add Sudakov form factor to account for the non-emission between these scales. [Hamilton,Nason,Zanderighi,arXiv:1206.3572]
- We can reweight our nTuple files to use such a scheme for setting the scales.
- Procedure has to be infrared safe: in the collinear and soft limits the reweighting of a real matrix element and its associated counter-events have to be the same.

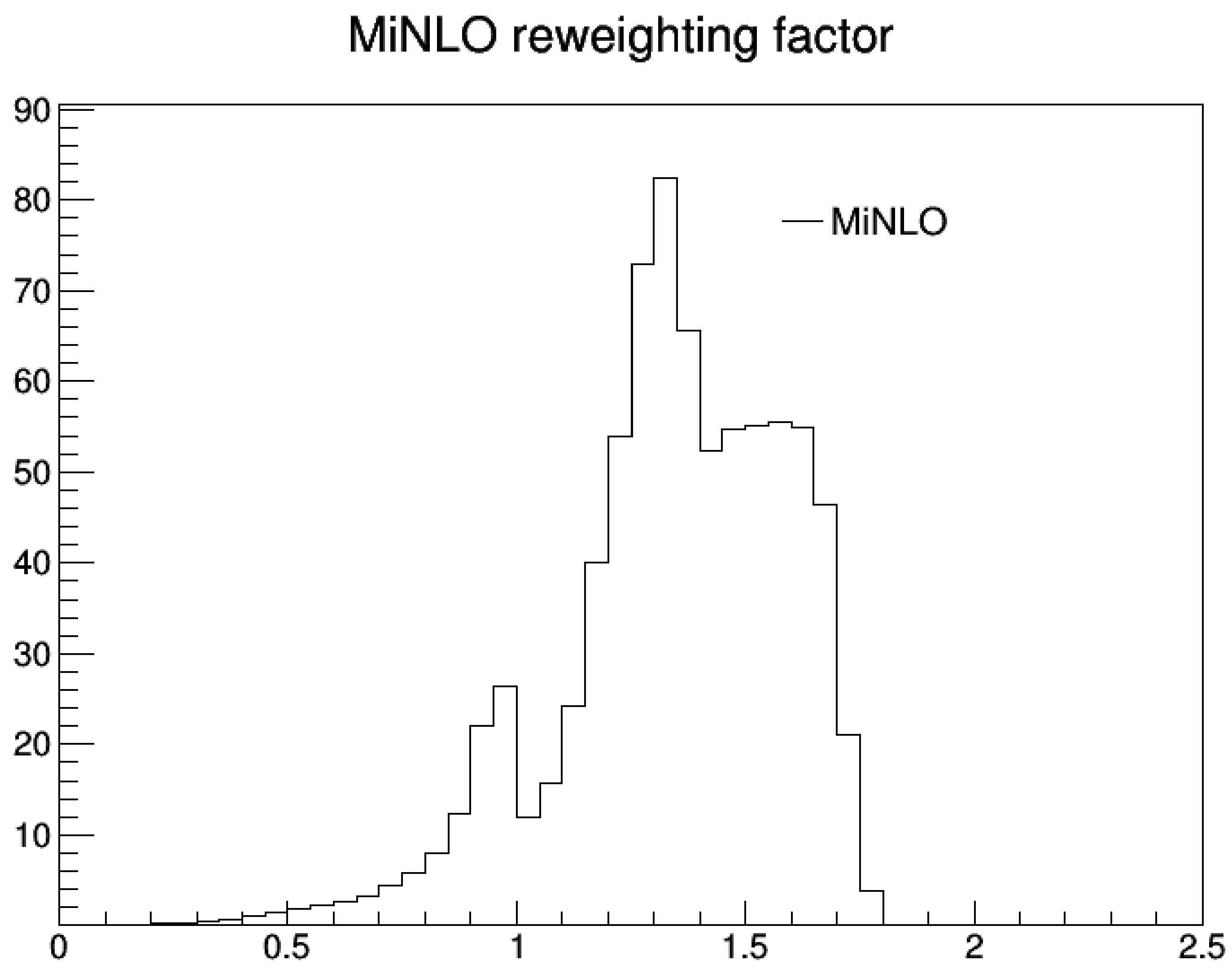
MiNLO



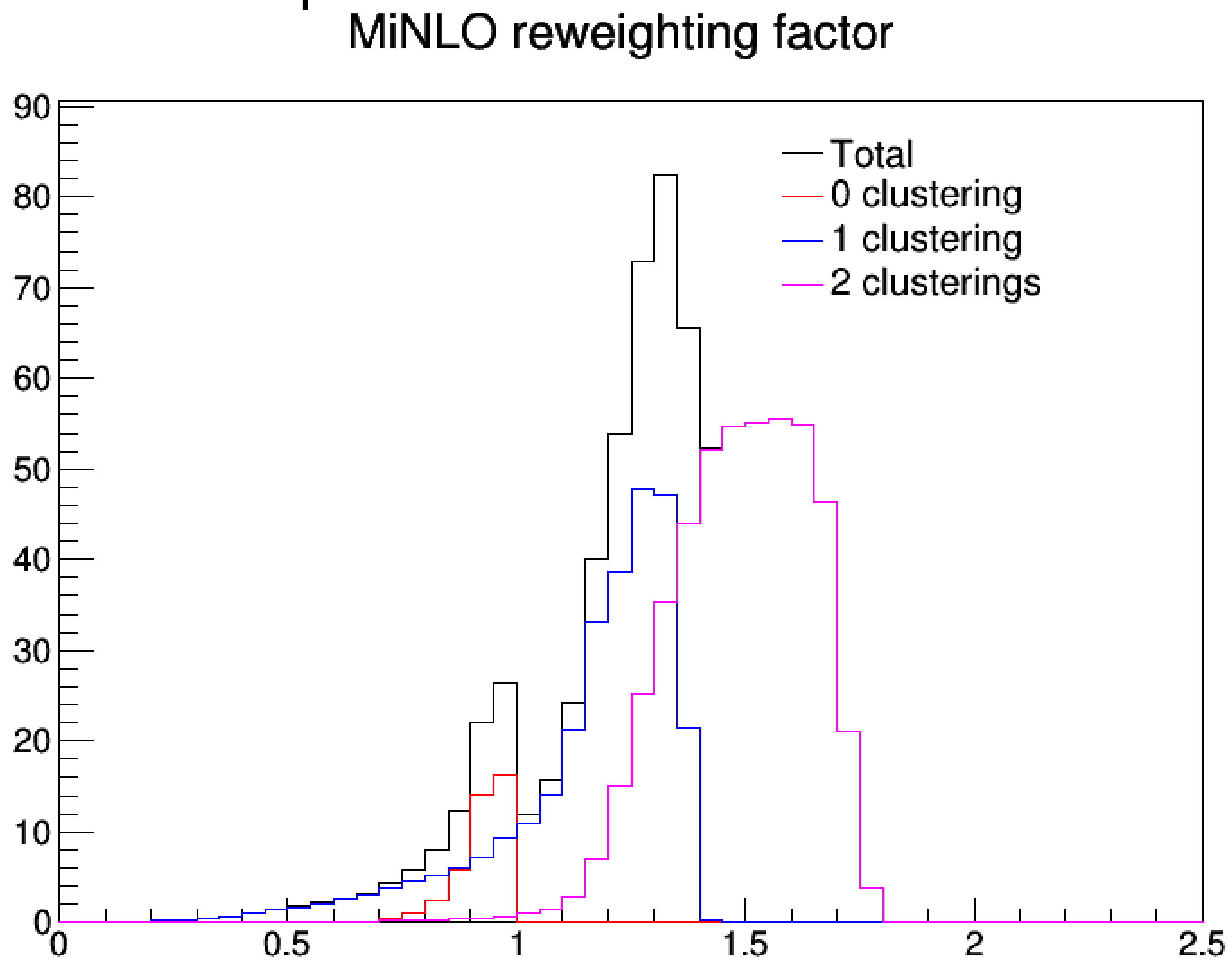
- Recluster event
- First clustering scale is the factorisation scale
- Each clustering scale is used as the argument of one power of α_s
- Apply Sudakov for each internal line
- Apply Sudakov for each external line
- Stop clustering when the next clustering is not credible (violates flavour, clustering on the “wrong side” beam)

Minlo

- Look at the effect reweighting factor we get for W+2j at LO



- We can understand the structure if we separate the contribution depending on how many clusterings the MiNLO method performed.



Conclusions

- The nTuple format is a very useful analysis tool for high multiplicity processes at NLO
- It allows to experiment with new scale setting methods
- A large set of jet multiplicities for a process allows to investigate universal properties as a function of the multiplicity