

# Multi-Jet Merging in Deep Inelastic Scattering with PYTHIA

Joni Laulainen<sup>1</sup>  
C.T. Preuss<sup>2</sup> I. Helenius<sup>1</sup>

<sup>1</sup>University of Jyväskylä  
Department of Physics

<sup>2</sup>University of Wuppertal  
Department of Physics

DIS2024, April 11th 2024



Multi-Jet Merging  
in DIS with  
PYTHIA

Joni Laulainen

Jet production in  
DIS

Parton Showers

Motivation

Merging

Algorithms

Merging scale

Results

HERA data comparisons

Scale variations

Summary

# Outline

Jet production in DIS

Parton Showers

Motivation

Merging

Algorithms

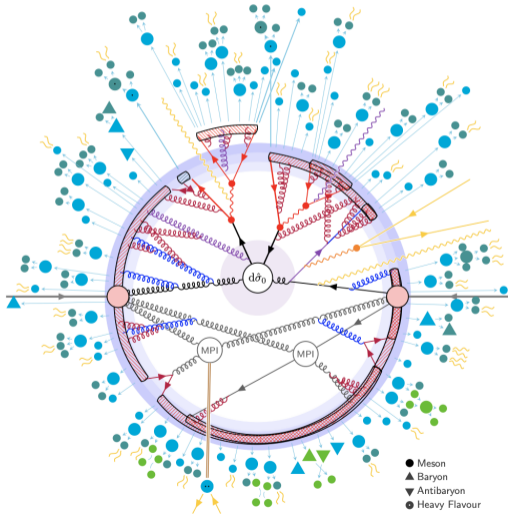
Merging scale

Results

HERA data comparisons

Scale variations

Summary



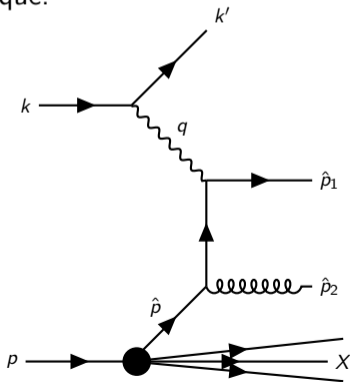
## Jet production in DIS

- ▶ DIS modeling lacking behind LHC-driven improvements in PYTHIA.
- ▶ Good description of DIS jets important with EIC in mind.
- ▶ Multiple hard scales, hardest scale not unique.

$$\mu_F^2 = Q^2$$

$$\mu_F^2 = \left( Q^2 + \left( \frac{1}{N} \sum_{i=1}^N \hat{p}_{T,i} \right)^2 \right) / 2$$

seen success in [J. Currie et al., 1703.05977]  
and [S. Höche et al., 1809.04192]



Jet production in  
DIS

Parton Showers

Motivation

Merging

Algorithms

Merging scale

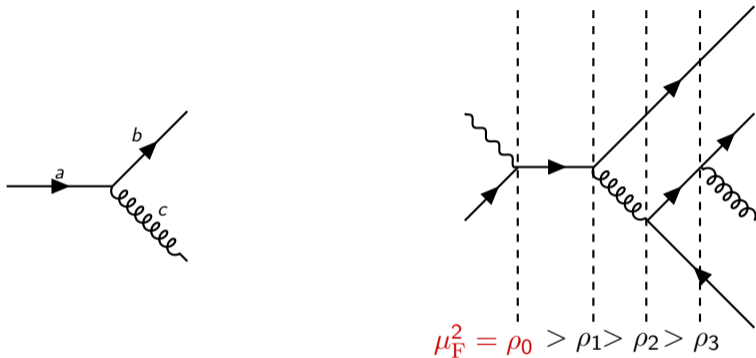
Results

HERA data comparisons

Scale variations

Summary

# Parton Showers



Differential probability for particle  $a$  splitting into  $b$  and  $c$  from DGLAP.  
Backwards evolution from large factorization scale  $\mu_F$  to low hadronization regime  $\Lambda_{\text{QCD}}$ .

$$d\mathcal{P}_a = \frac{d\rho}{\rho} \frac{\alpha_s}{2\pi} \sum_{b,c=\{q,g\}} \hat{P}_{a \rightarrow bc}(z) dz$$

Jet production in  
DIS

Parton Showers

Motivation

Merging

Algorithms

Merging scale

Results

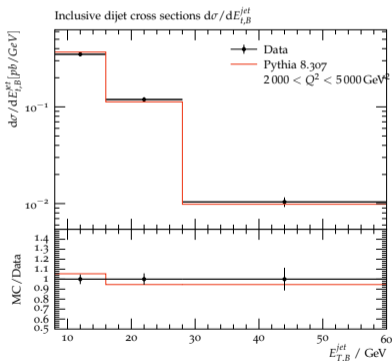
HERA data comparisons

Scale variations

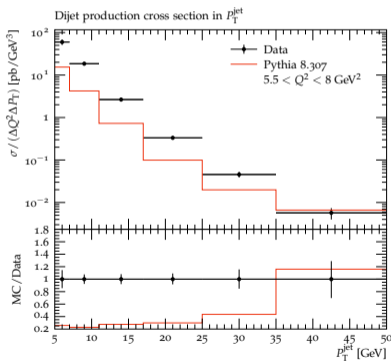
Summary

## Motivation - Parton shower approximation for dijets

- ▶ Good agreement in high- $Q^2$  using Power Shower.
- ▶ Poor description of low- $Q^2$  dijets.



[ZEUS Collaboration 1010.6167]



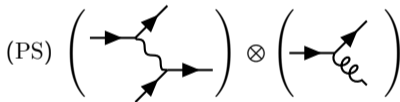
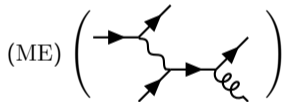
[H1 Collaboration, 1611.03421]

# Multi-Jet Merging

- ▶ Objective: Combine parton showers with matrix elements
- ▶ Separate phase space using merging scale.

Hard jets :  $p_T > t_{MS}$

Soft jets :  $p_T \leq t_{MS}$



SHERPA generated  
parton level events  
(ME) up to +4 jets.  
MADGRAPH5 reliably  
up to +2 jets.

[T. Gleisberg et al., 0811.4622] [J. Alwall  
et al., 1405.0301]

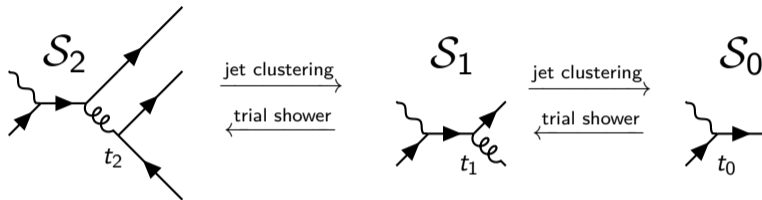


VINCIA as the parton  
shower (PS)

[C.T.Preuss et al., 2003.00702]

```
"PartonShowers:model = 2";
```

## Merging - Algorithm visually



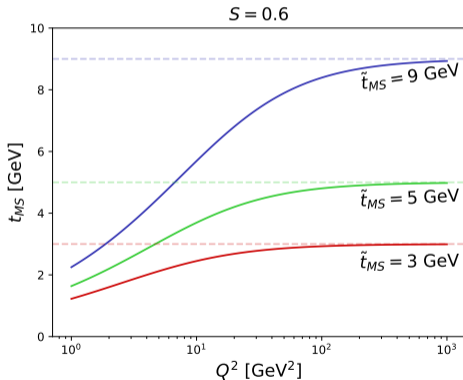
1. Start from ME event which passed the  $t_{MS}$  cut.
2. Cluster from shower state  $\mathcal{S}_n$  to obtain most probable shower history  $\{\mathcal{S}_1, \mathcal{S}_2 \dots, \mathcal{S}_{n-1}\}$  and corresponding scales  $\{t_1, t_2, \dots, t_n\}$ .
3. Starting from  $\mathcal{S}_0$ , generate trial emissions.
4. Reweight with merging weight, consisting of  $\prod_{\mathcal{S}_i}(t_i, t_{i+1}), \frac{f(x_i, t_i)}{f(x_i, t_{i+1})}$  and  $\frac{\alpha_{s,PS}(t_i)}{\alpha_{s,ME}}$ .

## Merging - $t_{MS}$ -dependence

Phase space separation with dynamic merging scale<sup>1</sup>.

- ▶ Attempt to solve problem of low- $Q^2$  jet production.
- ▶ Let MEs contribute more in low- $Q^2$  region.

$$t_{MS} = \tilde{t}_{MS} \frac{1}{\sqrt{1 + \frac{\tilde{t}_{MS}^2}{S^2 Q^2}}}$$



<sup>1</sup>SHERPA DIS jet merging [T. Carli, T. Gehrmann, S. Höche, 0912.3715]

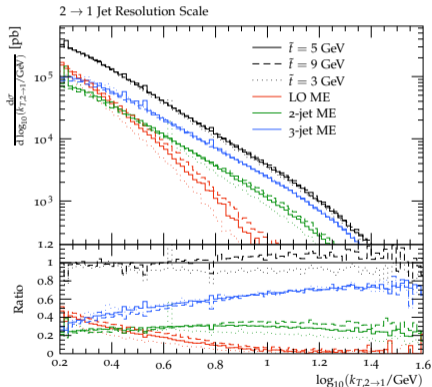


Merging -  $t_{\text{MS}}$ -dependence

Phase space separation with dynamic merging scale<sup>1</sup>.

- ▶ Attempt to solve problem of low- $Q^2$  jet production.
- ▶ Let MEs contribute more in low- $Q^2$  region.

$$t_{\text{MS}} = \tilde{t}_{\text{MS}} \frac{1}{\sqrt{1 + \frac{\tilde{t}_{\text{MS}}^2}{S^2 Q^2}}}$$

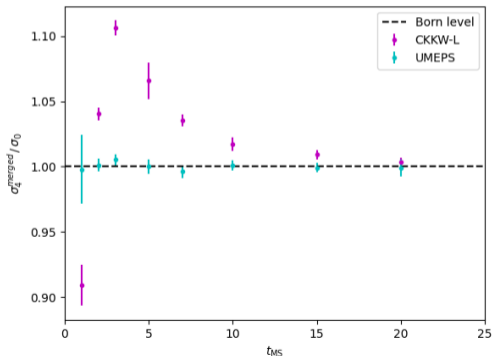


<sup>1</sup>SHERPA DIS jet merging [T. Carli, T. Gehrmann, S. Höche, 0912.3715]

# Merging - Algorithms

Merging algorithm effect on inclusive cross section. DIS +4 jets.

- ▶ CKKW-L<sup>1</sup>, dynamic merging
  - ▶ phase space separation  $t_{MS}$  and reweighting using PS histories to account for double-counting.
- ▶ UMEPS<sup>2</sup>, fixed merging
  - ▶ similar, but preserves unitarity.

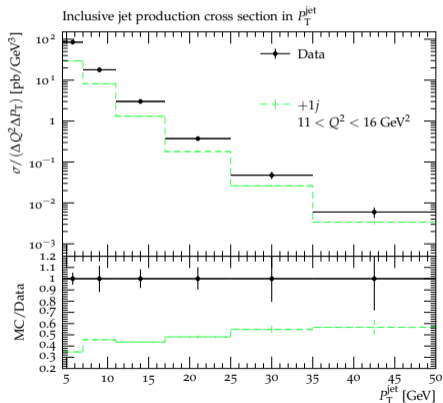
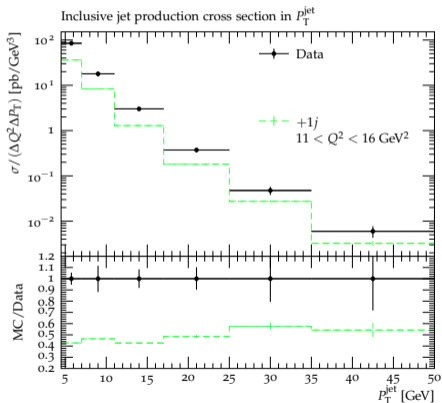


<sup>1</sup>[L. Lönnblad, 0112284]

<sup>2</sup>[L. Lönnblad, S. Prestel, 1211.4827]

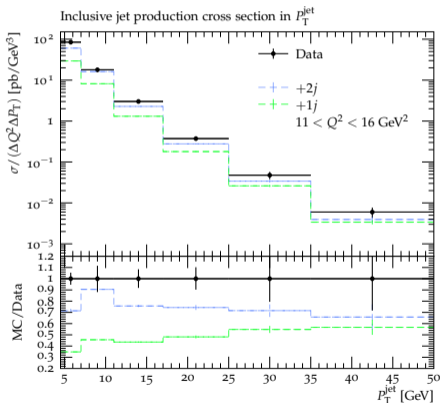
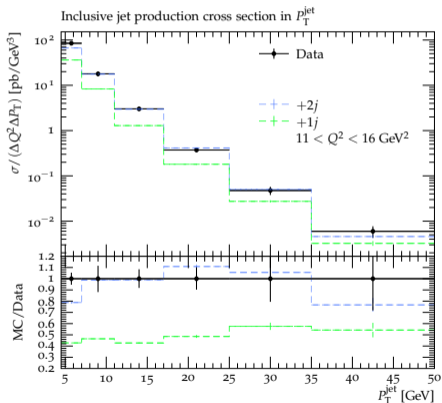
# Results

We compare the results to HERA data. [\[H1 Collaboration, 1611.03421\]](#)



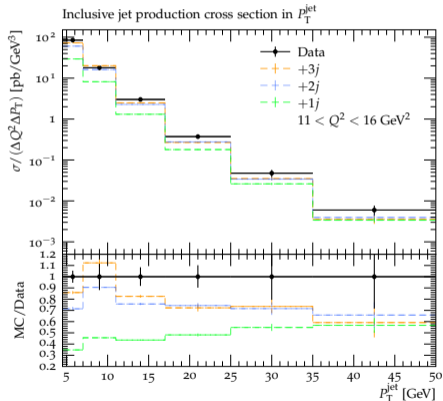
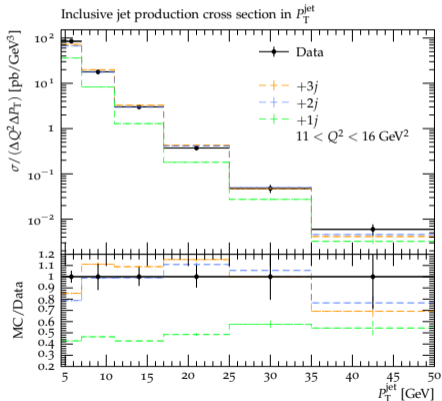
# Results

We compare the results to HERA data. [H1 Collaboration, 1611.03421]



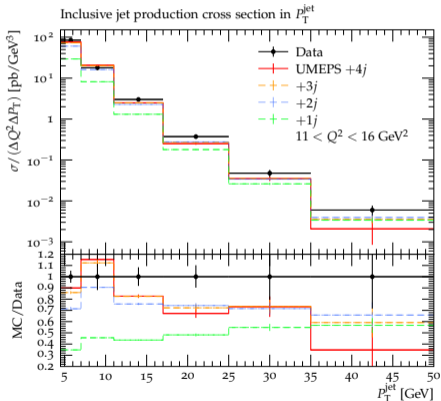
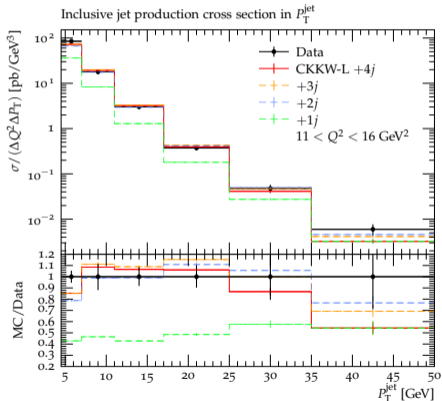
# Results

We compare the results to HERA data. [H1 Collaboration, 1611.03421]



# Results

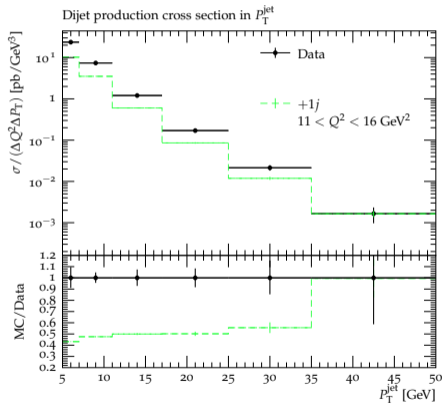
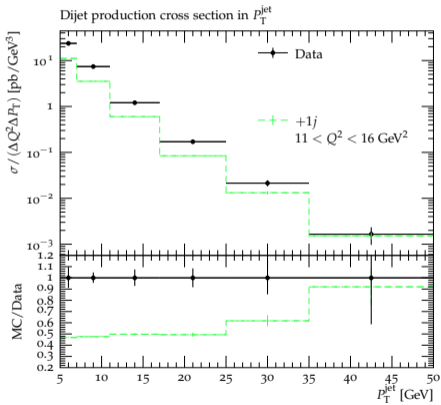
We compare the results to HERA data. [H1 Collaboration, 1611.03421]



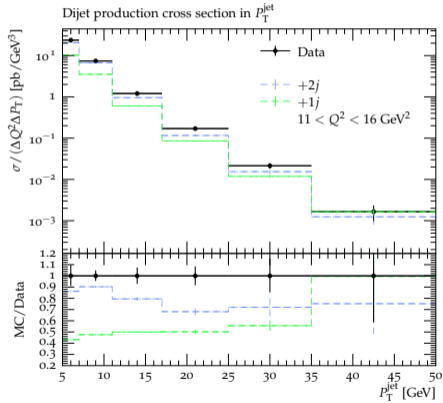
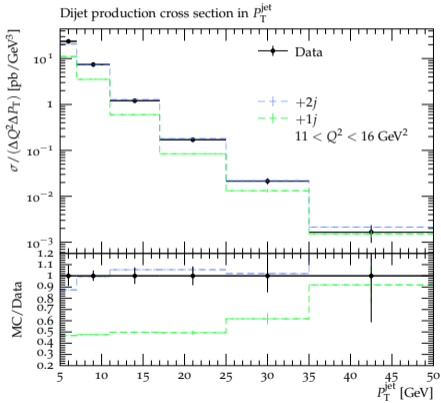
Satisfactory result with just 2 additional jets.

Result converges with addition of higher multiplicity samples.

# Results

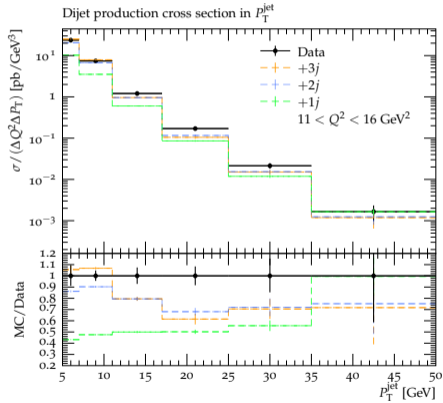
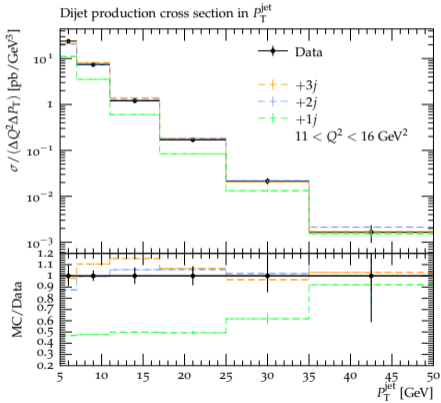


# Results

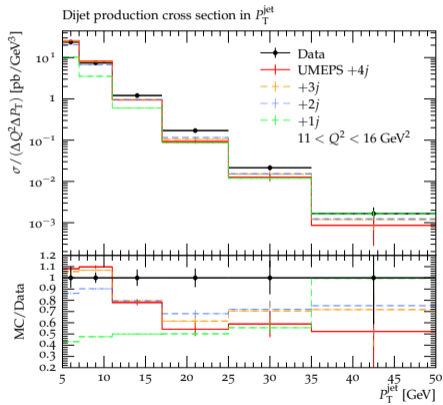
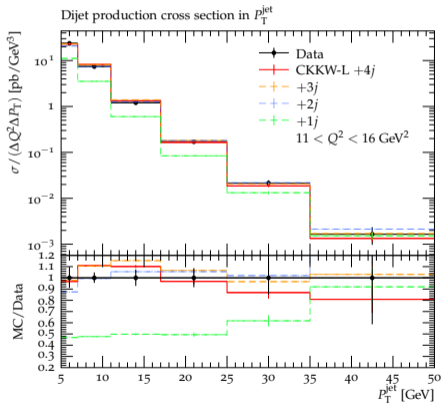




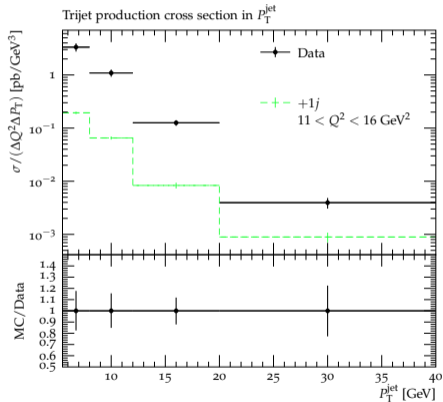
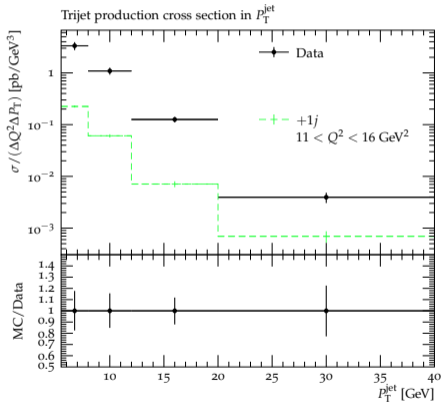
# Results



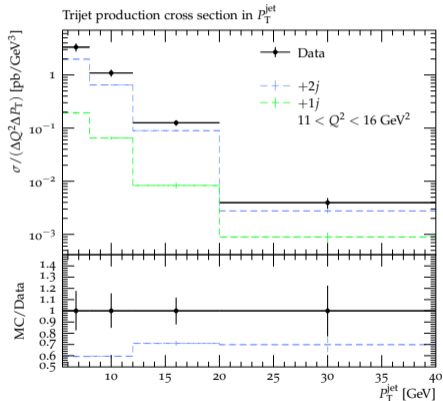
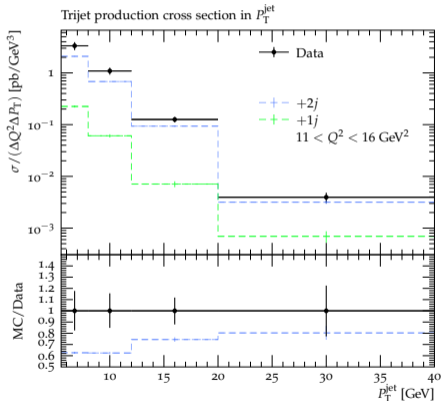
# Results



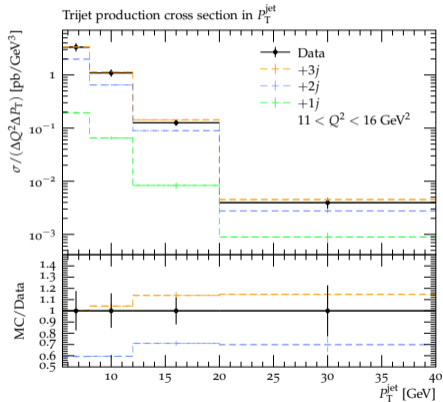
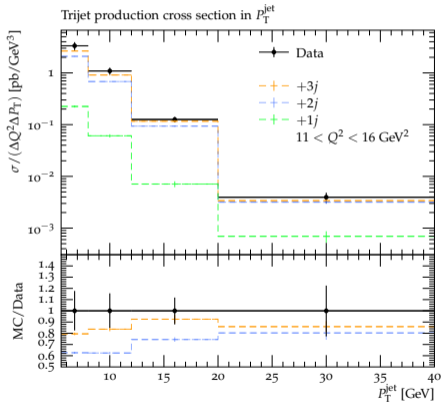
# Results



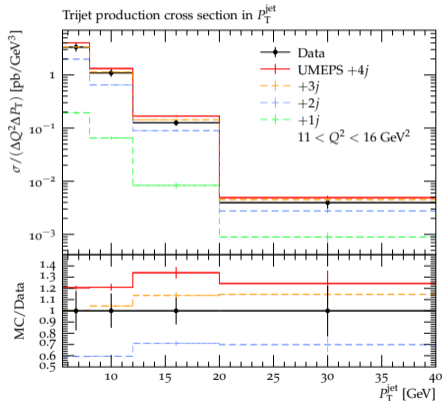
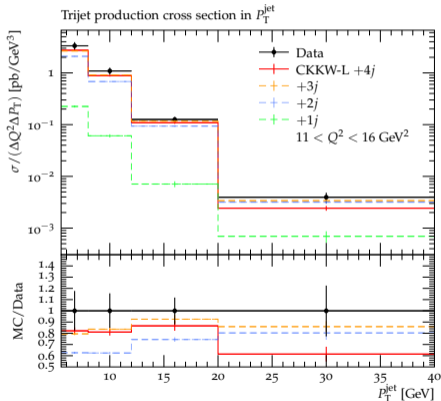
# Results



# Results



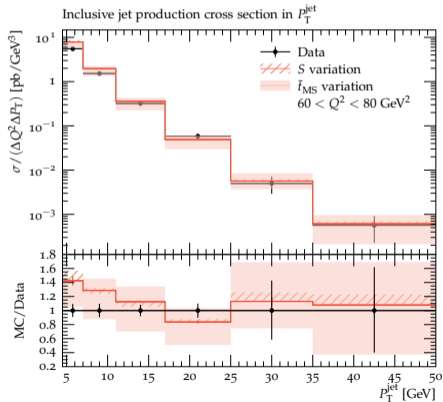
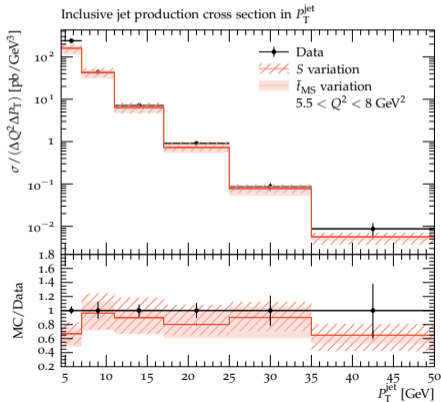
# Results



Competent trijet modeling requires at least +3-jet merging.

# Results - CKKW-L parameter variations

Variation of  $S = 0.8 \pm 0.1$  and  $t_{\text{MS}}$  between 3 and 9 GeV.



$S$ -parameter mostly influences low- $Q^2$  region.

Jet production in  
DIS

Parton Showers

Motivation

Merging

Algorithms

Merging scale

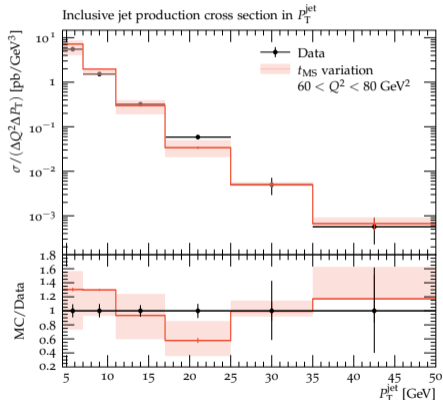
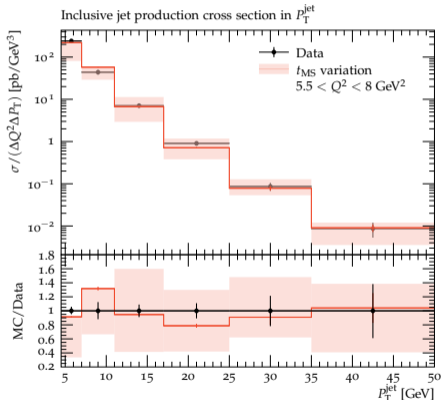
Results

HERA data comparisons

Scale variations

Summary

## Results - UMEPS parameter variations



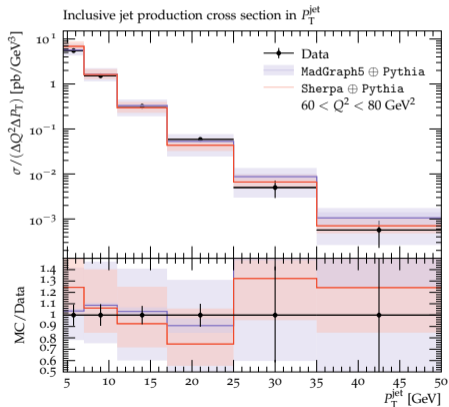
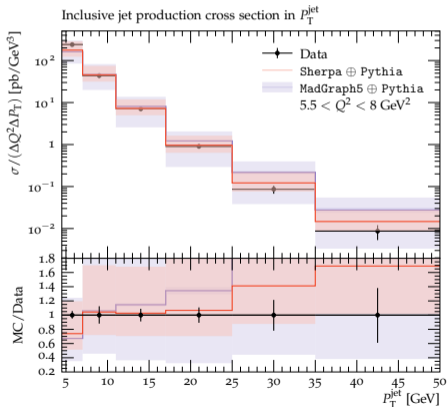
Moderate sensitivity to fixed merging scale variations.



## Results - Scale variations

Also possible to use MG5 input. CKKW-L +2 jets

Uncertainty bands by varying the scales  $\mu_R$  and  $\mu_F$  by a factor of 2.



Jet production in  
DIS

Parton Showers

Motivation

Merging

Algorithms

Merging scale

Results

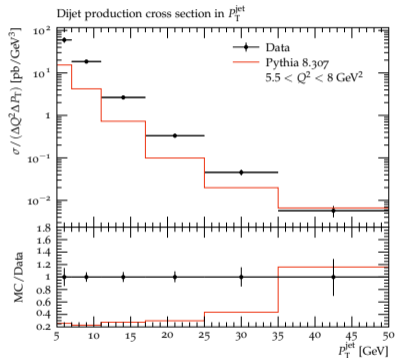
HERA data comparisons

Scale variations

Summary

## Summary

- ▶ Starting point: default PYTHIA shower not enough to describe jets in DIS  
→ merging implementation.
- ▶ Problem: Hardest scale not unique  
→ use dynamic merging scale or different factorization scale choice.
- ▶ Multi-jet merging provides good description of HERA data also in low- $Q^2$  region.

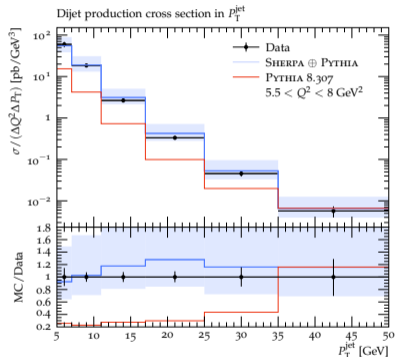


Outlook: Working DIS jet merging setup for an upcoming PYTHIA release, possibly QED-clusterings to VINCIA merging.

Upcoming projects: Matrix element corrections.

## Summary

- ▶ Starting point: default PYTHIA shower not enough to describe jets in DIS → merging implementation.
- ▶ Problem: Hardest scale not unique → use dynamic merging scale or different factorization scale choice.
- ▶ Multi-jet merging provides good description of HERA data also in low- $Q^2$  region.



Outlook: Working DIS jet merging setup for an upcoming PYTHIA release, possibly QED-clusterings to VINCIA merging.  
 Upcoming projects: Matrix element corrections.

## Fixed order calculations / Matrix elements

Why not just use exact fixed order calculations all the way to a given final state multiplicity? Why stop at +4 jets?

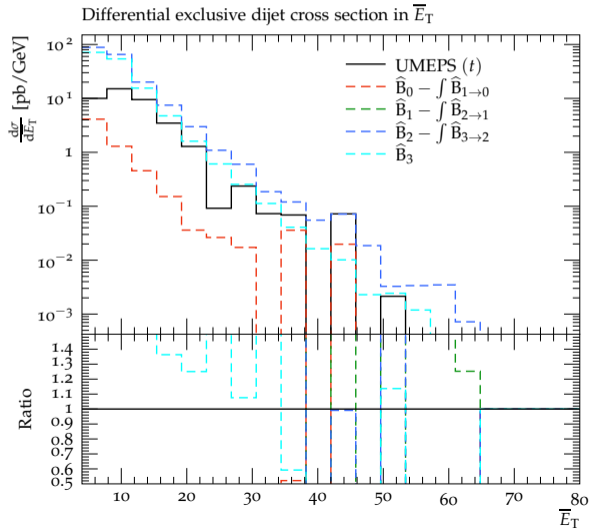
Number of diagrams grows almost **factorially** as a function of  $N$ .

Final states are observed to contain **hundreds** of particles.

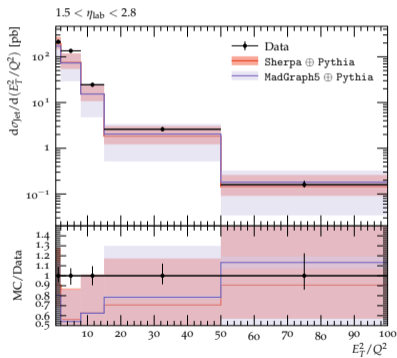
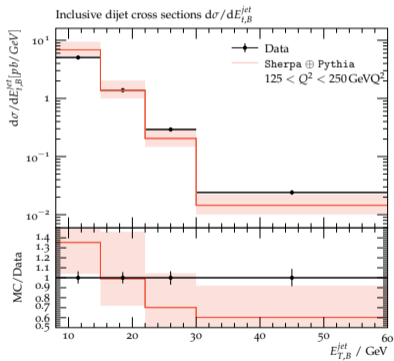
Process	# of diagrams
$e^- p \rightarrow e^- j$	20
$e^- p \rightarrow e^- 2j$	60
$e^- p \rightarrow e^- 3j$	720
$e^- p \rightarrow e^- 4j$	5 340
$e^- p \rightarrow e^- 5j$	71 080

Access description of up to 5 hard jets with matrix elements.

## UMEPS dynamical merging scale



## Other analyses: ZEUS 2010 and HERA 2002



Hard events are generated at leading order in QCD using a HPC-enabled version of the event generator<sup>2</sup> and stored using the efficient and scalable HDF5 event-file format. We use the  $G$ -scheme with electroweak input parameters

$$m_W = 80.419 e, \quad m_Z = 91.188 e, \quad G_F = 1.16639 \times 10^{-5} e^{-2}. \quad (1)$$

Quark masses are taken in the four-flavour scheme and the  $b$ -quark mass is set to  $m_b = 4.7 e$ . The PDF choice is NNPDF40\_lo\_pch\_as\_01180.

---

<sup>2</sup>Available at <https://gitlab.com/hpcgen/me>.

I. Produce *Les Houches* event files (LHEF) [25] with a matrix element generator for  $n = 0, 1, \dots, N$  extra jets with a regularisation cut-off,  $\rho_{\text{MS}}$ , typically using a fixed factorisation scale,  $\mu_F$ , and a fixed  $\alpha_s(\mu_R)$ .

II. Pick a jet multiplicity,  $n$ , and a state  $S_{+n}$  according to the cross sections given by the matrix element generator.

1. Find all shower histories for the state  $S_{+n}$ , pick a sequence according to the product of splitting probabilities. Only pick un-ordered sequences if no ordered sequence was found. Only pick incomplete paths if no complete path was constructed.

2. Perform reweighting: For each  $0 \leq i < n$ ,

i. Start the shower off the state  $S_{+i}$  at  $\rho_i$ , generate a trial state  $R$  with scale  $\rho_R$ . If  $\rho_R > \rho_{i+1}$ , veto the event and start again from II.

ii. Calculate the weight factor

$$w_i = \frac{\alpha_s(\rho_{i+1}) x_i^+ f_i^+(x_i^+, \rho_i) x_i^- f_i^-(x_i^-, \rho_i)}{\alpha_s(\mu_R) x_i^+ f_i^+(x_i^+, \rho_{i+1}) x_i^- f_i^-(x_i^-, \rho_{i+1})} \quad (4.16)$$

3. Start the shower from  $S_{+n}$ .

i. If  $n < N$ , start the shower at  $\rho_n$ , veto any shower emission producing an additional resolved jet.

ii. If  $n = N$ , start the shower at  $\rho_n$ .

III. If the event was not rejected, multiply the event weight by

$$w'_n = \frac{x_n^+ f_n^+(x_n^+, \rho_n)}{x_n^+ f_n^+(x_n^+, \mu_F)} \times \frac{x_n^- f_n^-(x_n^-, \rho_n)}{x_n^- f_n^-(x_n^-, \mu_F)} \times \prod_{i=0}^{n-1} w_i \quad (4.17)$$

V. Start again from II.

The second part, i.e. producing  $\int_s \widehat{\text{B}}_{n \rightarrow m}$ -events to effect lower-multiplicity PS resummation, requires only two changes:

II.3 Replace the matrix-element state by  $S_{+n-1}$ , or the first state  $S_{+l}$  with all  $l \leq n-1$  partons above the merging scale. If no integrated state can be constructed, i.e. if only incomplete paths were found, reject the event. For valid events, start the shower at  $\rho_n$ , veto any shower emission producing an additional resolved jet.

1. Calculate cross sections and generate events according to step 1 and 2 in section 2.1.  $Q_{\text{MS}}$  denotes the merging scale which is equal to the matrix element cutoff and may be defined using any choice of scale. The events are generated using a fixed strong coupling,  $\alpha_{\text{sME}}$ , and a maximum parton multiplicity,  $N$ .

2. Construct a full cascade history by considering all possible ordered histories and selecting one randomly with a probability proportional to the product of the branching probabilities. If no ordered histories can be constructed, unordered ones are considered. This results in a set of intermediate states ( $S_2, S_3, \dots, S_n$ ) and scales ( $\rho_2 = \rho_{\text{max}}, \rho_3, \dots, \rho_n$ ).  $S_2$  denotes here denotes the constructed  $2 \rightarrow 2$  process and  $S_n$  is the state given by the matrix element.  $n$  is the parton multiplicity in the event,  $\rho_{\text{max}}$  is the maximum scale of the process and  $\rho_i$  is the constructed scale where the state  $S_{i-1}$  emits a parton to produce the state  $S_i$ .

3. Reweight the events with  $\prod_{i=3}^n \alpha_s(\rho_i) / \alpha_{\text{sME}}^{n-2}$ .

4. For each state  $S_i$  (except  $S_n$ ), generate an emission with  $\rho_i$  as starting scale and if this emission occurred at a scale larger than  $\rho_{i+1}$  reject the event. This is equivalent to reweighting with a factor  $\prod_{i=2}^{n-1} \Delta_{S_i}(\rho_i, \rho_{i+1})$ .

5. For the last step there are two cases.

- If the event does not have the highest multiplicity  $n < N$ , generate an emission from the state  $S_n$  with  $\rho_n$  as starting scale. If the emission is above the merging scale  $Q_{\text{MS}}$ , reject the event. Otherwise accept the event and continue the cascade.
- If the event has the highest possible multiplicity  $n = N$ , accept the event and start the cascade from the state  $S_n$  with the scale  $\rho_n$ .



For CKKW-L 1-jet merging to preserve unitarity, the first PS emission has to exactly match the 1-jet ME. This can be done (matching), but then there would be no need for 1-jet CKKW-L. For 2-jet merging, "the splitting kernels need to exactly reproduce the matrix element, phase space must be fully covered by the parton shower, and the no-emission probabilities need to be produced identically in both cases. Particularly the requirement that the phase space is completely covered is problematic, since parton showers commonly fill only phase space regions in which consecutive emissions are ordered in a decreasing evolution variable."<sup>3</sup>

---

<sup>3</sup>[L. Lönnblad, S. Prestel, 1211.4827]

UMEPS preserves the Born level cross section by integrating over  $n + 1$ -jet phase space to induce resummation in  $n$ -jet cross section. The idea is to subtract in  $n$  what you add in  $n - 1$ .

$$\sigma^{\text{inc}} = \int d\phi_0 \left( \frac{d\sigma_0^{\text{exc}}}{d\phi_0} + \frac{d\sigma_1^{\text{inc}}}{d\phi_0} \right) \quad (2)$$

$$= \int d\phi_0 \left( \frac{d\sigma_0^{\text{inc}}}{d\phi_0} - \frac{d\sigma_{1 \rightarrow 0}^{\text{inc}}}{d\phi_0} + \frac{d\sigma_1^{\text{inc}}}{d\phi_0} \right) \quad (3)$$