Event generator overview

MC4EIC

Ilkka Helenius

July 9th, 2025

JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

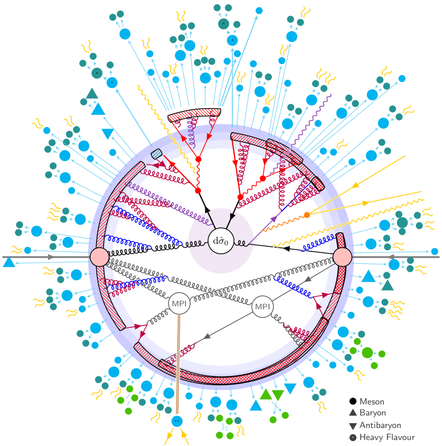
Research Council of Finland

Centre of Excellence in Quark Matter

Outline

Outline

1. Event generators for high energy colliders
2. Electroproduction
3. Photoproduction
4. Diffractive processes
5. Tuning
6. MC4EIC recap



[figure by P. Skands]

Event generators

General purpose event generators

- Aim to provide a full description of a collision event, ie. exclusive hadronic final states, using Monte Carlo methods
- Use perturbative QCD where applicaple, fill in with phenomenologica models
- Main players:
  - Herwig (7.3.0) `https://herwig.hepforge.org` [Eur.Phys.J. C80 (2020) 452]
  - Pythia (8.315) `https://pythia.org` [SciPost Phys. Codebases 8-r8.3 (2022)]
  - Sherpa (3.0.1) `https://sherpa-team.gitlab.io` [JHEP 12 (2024) 156]

Specialized event generators

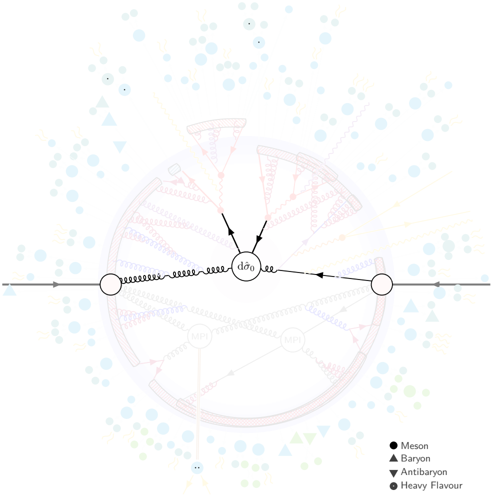
- Matrix-element (Hard-process) generators for higher perturbative accuracy and multiplicities: Madgraph5(_aMC@NLO), POWHEG(-BOX)
- Fixed-order codes: MCFM, NNLOJET, ...

Physics modelled within generators

Classify event generation in terms of "hardness"

1. Hard Process (here $t\bar{t}$)

[figure credit: P. Skands]

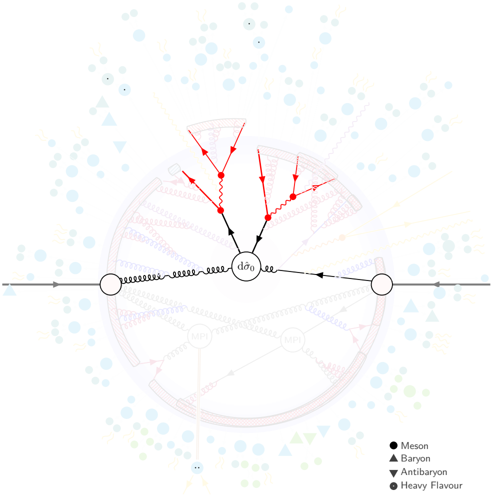


Physics modelled within generators

Classify event generation in terms of "hardness"

[figure credit: P. Skands]

1. Hard Process (here $t\bar{t}$)
2. Resonance decays ($t, Z, \ldots$)

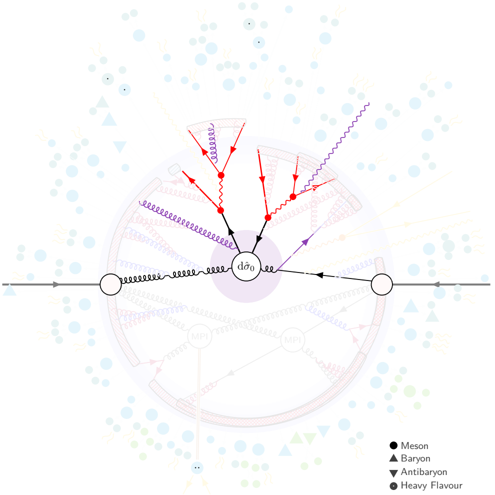


Physics modelled within generators

Classify event generation in terms of “hardness”

[figure credit: P. Skands]

1. Hard Process (here $t\bar{t}$)
2. Resonance decays $(t, Z, \ldots)$
3. Matching, Merging and matrix-element corrections

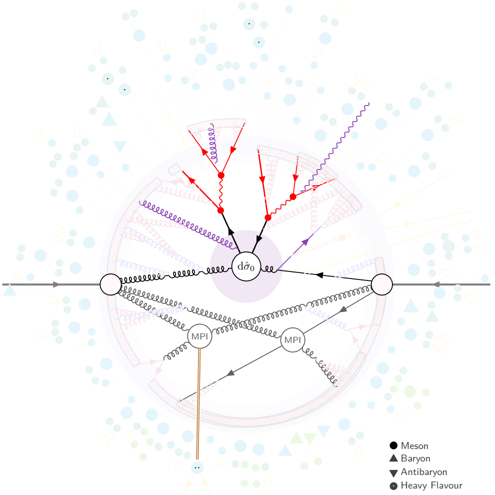


Physics modelled within generators

Classify event generation in terms of "hardness"

1. Hard Process (here $t\bar{t}$)
2. Resonance decays ($t, Z, \ldots$)
3. Matching, Merging and matrix-element corrections
4. Multiparton interactions

[figure credit: P. Skands]

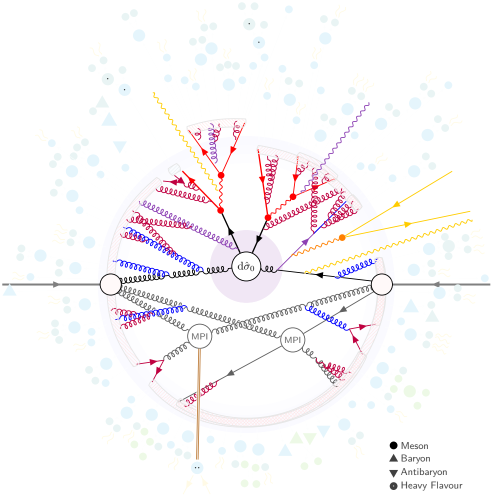


Physics modelled within generators

Classify event generation in terms of "hardness"

1. Hard Process (here $t\bar{t}$)
2. Resonance decays ($t, Z, \ldots$)
3. Matching, Merging and matrix-element corrections
4. Multiparton interactions
5. Parton showers: ISR, FSR, QED, Weak

[figure credit: P. Skands]

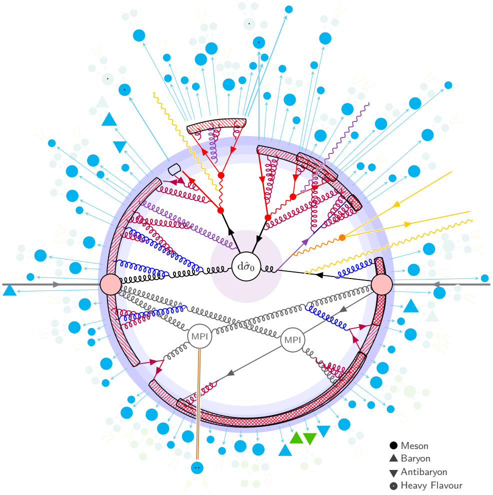


Physics modelled within generators

Classify event generation in terms of "hardness"

[figure credit: P. Skands]

1. Hard Process (here $t\bar{t}$)
2. Resonance decays ($t, Z, \ldots$)
3. Matching, Merging and matrix-element corrections
4. Multiparton interactions
5. Parton showers: ISR, FSR, QED, Weak
6. Hadronization, Beam remnants

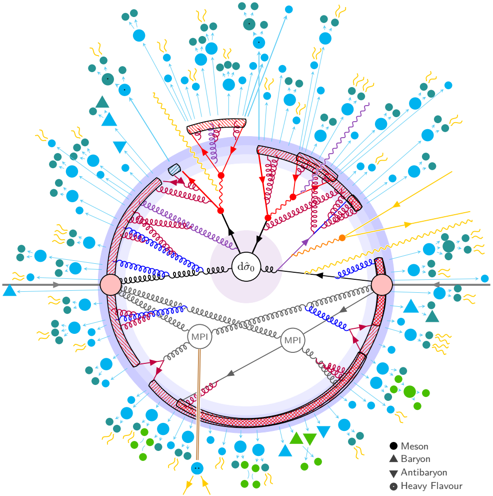


Physics modelled within generators

Classify event generation in terms of "hardness"

1. Hard Process (here $t\bar{t}$)
2. Resonance decays ($t, Z, \ldots$)
3. Matching, Merging and matrix-element corrections
4. Multiparton interactions
5. Parton showers: ISR, FSR, QED, Weak
6. Hadronization, Beam remnants
7. Decays, Rescattering

[figure credit: P. Skands]



Parton Showers provide leading-log resummation

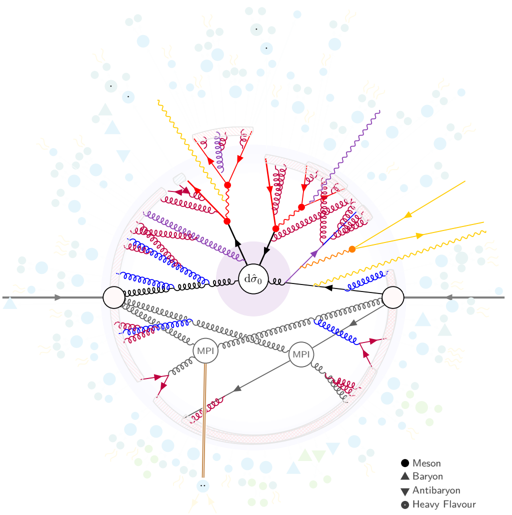
Dress the partons by generating explicit branchings iteratively

- Start from highly-virtual partons, evolve down to low scales with DGLAP
- Splitting probabilities from

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

where $P_{a\to bc}(z)$ splitting kernels

- Different choices in ordering variable and phase-space mapping lead to some differences between different implementations



Improve precision: Matching and merging

Combine multi-jet (fixed-order) calculations with each other and with PS

Matrix element corrections (MECs):

- Correct first PS splitting ($2 \to 2 + 1$) with the full matrix element ($2 \to 3$)

Matching:

- Combine $\{n, n + 1\}$-parton states from NLO ME generator with parton shower
- Exclude overlap by subtraction or by correction factors
- NLO precision for n-parton observables

Merging:

- Combine $\{n, n + 1, \ldots, n + m\}$ events from ME generators with each other and parton shower
- Overlap removed by applying cuts and vetoes

NLO merging:

- As above but with NLO MEs, overlap removed by subtraction
- NLO precision for inclusive $(n + i)$-parton observables

Multiparton interactions (MPIs)

- MPIs from $2 \to 2$ QCD cross sections

$$\frac{d\mathcal{P}_{\mathrm{MPI}}}{dp_T^2} = \frac{1}{\sigma_{\mathrm{nd}}(\sqrt{s})} \frac{d\sigma^{2\to 2}}{dp_T^2}$$

$\sigma_{\mathrm{nd}}(\sqrt{s})$ is the non-diffractive cross section

- Partonic cross section diverges at $p_T \to 0$

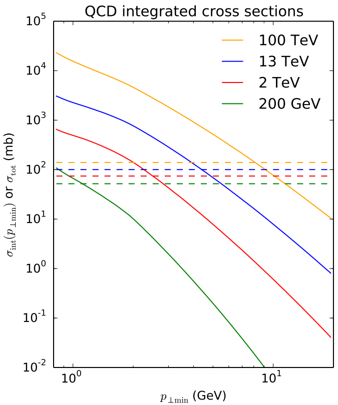
$\Rightarrow$ Introduce a screening parameter p_{T0}

$$\frac{d\sigma^{2\to 2}}{dp_T^2} \propto \frac{\alpha_s(p_T^2)}{p_T^4} \to \frac{\alpha_s(p_{T0}^2 + p_T^2)}{(p_{T0}^2 + p_T^2)^2}$$

- Energy-dependent parametrization:

$p_{T0}(\sqrt{s}) = p_{T0}^{\mathrm{ref}}(\sqrt{s}/\sqrt{s_{\mathrm{ref}}})^{\alpha}$

- Number of interactions: $\langle n \rangle = \sigma_{\mathrm{int}}(p_{T0})/\sigma_{\mathrm{nd}}$



- $\sigma_{\mathrm{int}}(p_{T,\mathrm{min}})$ exceeds σ_{tot}

$\Rightarrow$ Several interactions

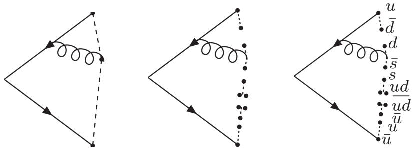
Hadronization models

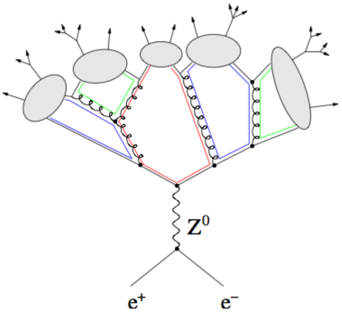
String hadronization

- Implemented in Pythia, can be interfaced from Sherpa
- Colour string between colour charges, hadrons formed from string breaking

Cluster model

- Implemented in Herwig and Sherpa
- Gluons are forced to make $q\bar{q}$ pairs
- Form colour-singlet clusters, these decay isotropically into hadrons





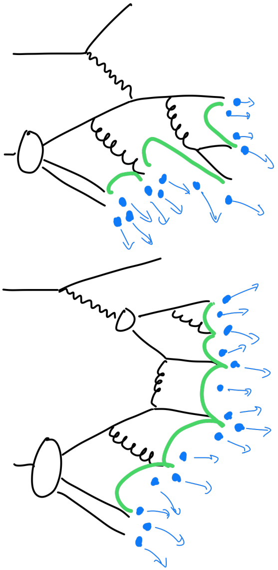
Electron-hadron collisions

Electroproduction (deep inelastic scattering, DIS)

- Lepton scatters off a parton by exchanging a highly virtual photon
- High virtuality, $Q^2 > $ a few GeV^2
- Hard process + Parton showers

Photoproduction (PhP)

- Low virtuality, $Q^2 \rightarrow 0\ \text{GeV}^2$
- Photon may fluctuate into a hadronic state, resolved in the interaction $\Rightarrow$ MPIs
- Factorize photon flux, evolve γp system
- Also soft QCD processes, diffraction



Electroproduction

Event generation in DIS

Hard scattering

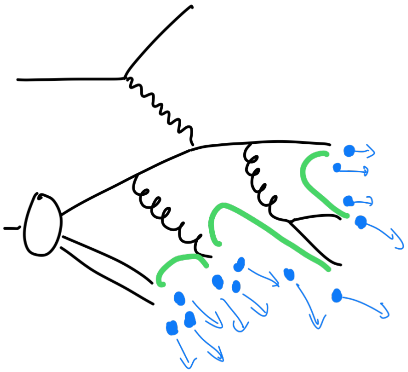
- Convolution between PDFs and matrix element (ME) for partonic scattering

Parton shower

- Final state radiation (FSR)
- Initial state radiation (ISR) for hadron
- QED emissions from leptons

Hadronization

- String/cluster hadronization with colour reconnections
- Decays to stable hadrons



H1 data for 1-jettiness

Pythia

- Default shower with the dipole-recoil option
- Vincia sectorized parton shower
- Dire parton shower

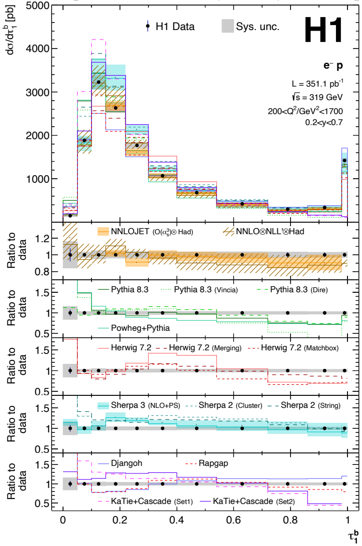
Herwig

- Default angular-ordered shower
- Matching (Matchbox) and merging

Sherpa

- With cluster and string hadronization
- NLO matching

[H1: EPJC 84 (2024) 8, 785]



Photoproduction

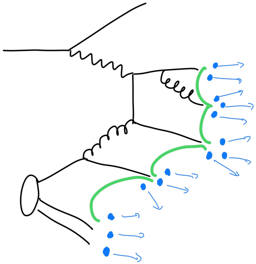
Photoproduction in electron-proton collisions

Direct processes

- Convolute photon flux f_γ with proton PDFs f_i^{p} and $\mathrm{d}\hat{\sigma}$

$$\mathrm{d}\sigma^{\mathrm{ep}\to kl+X} = f_\gamma^{\mathrm{e}}(x, Q^2) \otimes f_j^{\mathrm{p}}(x_{\mathrm{p}}, \mu^2) \otimes \mathrm{d}\hat{\sigma}^{\gamma j\to kl}$$

- Generate FSR and ISR for proton side

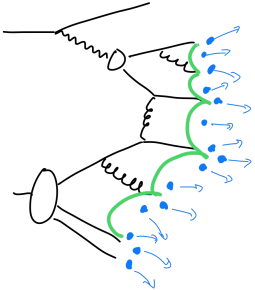


Resolved processes

- Convolute also with photon PDFs

$$\mathrm{d}\sigma^{\mathrm{ep}\to kl+X} = f_\gamma^{\mathrm{e}}(x, Q^2) \otimes f_i^{\gamma}(x_\gamma, \mu^2) \otimes f_j^{\mathrm{p}}(x_{\mathrm{p}}, \mu^2) \otimes \mathrm{d}\sigma^{ij\to kl}$$

- Sample x and Q^2, setup $\gamma\mathrm{p}$ sub-system with $W_{\gamma\mathrm{p}}$
- Evolve $\gamma\mathrm{p}$ as any hadronic collision (including MPIs)



Photon flux from EPA

$$f_\gamma^{\mathrm{e}}(x, Q^2) = \frac{\alpha_{\mathrm{em}}}{2\pi}\frac{1}{Q^2}\frac{(1+(1-x)^2)}{x}$$

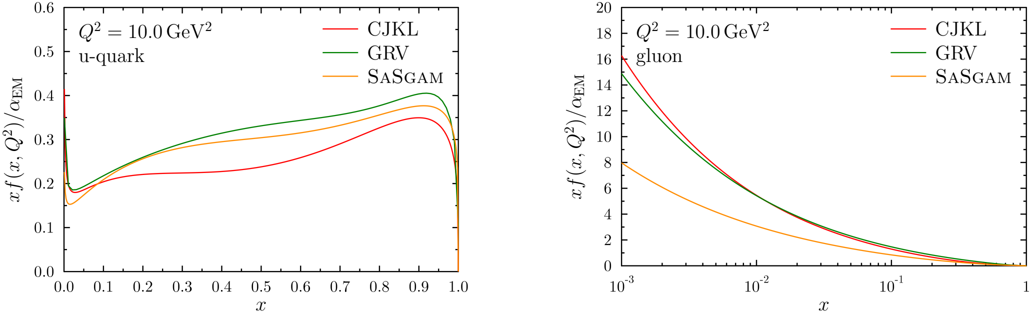
PDFs for resolved photons

DGLAP equation for photons

- Additional term due to $\gamma \to q\bar{q}$ splittings

$$\frac{\partial f_i^{\gamma}(x,Q^2)}{\partial \log(Q^2)} = \frac{\alpha_{\mathrm{em}}}{2\pi} e_i^2 P_{i\gamma}(x) + \frac{\alpha_s(Q^2)}{2\pi} \sum_j \int_x^1 \frac{dz}{z} P_{ij}(z)\, f_j(x/z, Q^2)$$

where $P_{i\gamma}(x) = 3\,(x^2 + (1-x)^2)$ for quarks, 0 for gluons (LO)



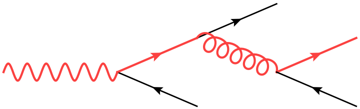
Evolution equation and ISR for resolved photons

ISR probability based on DGLAP evolution

- Add a term corresponding to $\gamma \to q\bar{q}$ to (conditional) ISR probability

$$\mathrm{d}\mathcal{P}_{a\leftarrow b} = \frac{\mathrm{d}Q^2}{Q^2}\frac{\alpha_\mathrm{s}}{2\pi}\frac{x' f_a^{\gamma}(x', Q^2)}{x f_b^{\gamma}(x, Q^2)} P_{a\to bc}(z)\,\mathrm{d}z + \frac{\mathrm{d}Q^2}{Q^2}\frac{\alpha_\mathrm{em}}{2\pi}\frac{e_b^2\, P_{\gamma\to bc}(x)}{f_b^{\gamma}(x, Q^2)}$$

- Corresponds to ending up to the beam photon during evolution
  $\Rightarrow$ Parton originated from the point-like (anomalous) part of the PDFs
  - No further ISR or MPIs below the scale of the splitting
  - Implemented for the default Simple Shower in Pythia 8

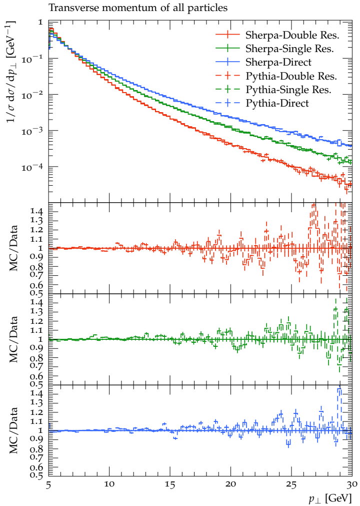


Comparisons between Pythia, Sherpa and Herwig

[I. Helenius, P. Meinzinger, S. Plätzer, P. Richardson: arXiv:2406.08026 [hep-ph]]

Compare different generators for photoproduction

- Good agreement at ME-level
- Differences build up from inputs and modelling
- Scale variations large at LO



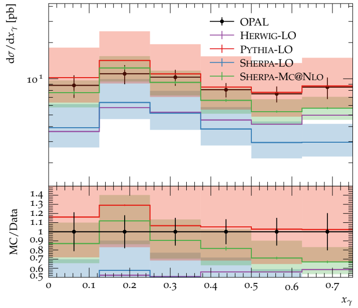
Comparisons between Pythia, Sherpa and Herwig

[I. Helenius, P. Meinzinger, S. Plätzer, P. Richardson: arXiv:2406.08026 [hep-ph]]

Compare different generators for photoproduction

- Good agreement at ME-level
- Differences build up from inputs and modelling
- Scale variations large at LO

Dijets in $\gamma\gamma$ (LEP)



[OPAL: PLB 651 (2007) 92-101]

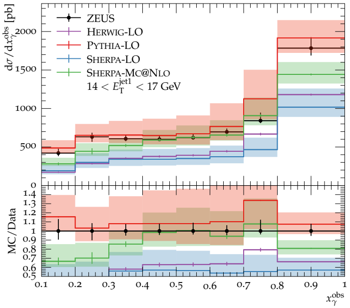
Comparisons between Pythia, Sherpa and Herwig

[I. Helenius, P. Meinzinger, S. Plätzer, P. Richardson: arXiv:2406.08026 [hep-ph]]

Compare different generators for photoproduction

- Good agreement at ME-level
- Differences build up from inputs and modelling
- Scale variations large at LO

Dijets in γp (HERA)



[ZEUS: EPJC 23 (2002) 615-631]

Comparisons between Pythia, Sherpa and Herwig

[I. Helenius, P. Meinzinger, S. Plätzer, P. Richardson: arXiv:2406.08026 [hep-ph]]

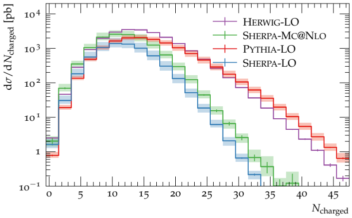
Compare different generators for photoproduction

- Good agreement at ME-level
- Differences build up from inputs and modelling
- Scale variations large at LO

Solid predictions for EIC require

- Validated inputs: (γ)PDFs, accurate flux
- Improved modelling for PS and remnant handling
- Tuning of models to HERA and LEP data

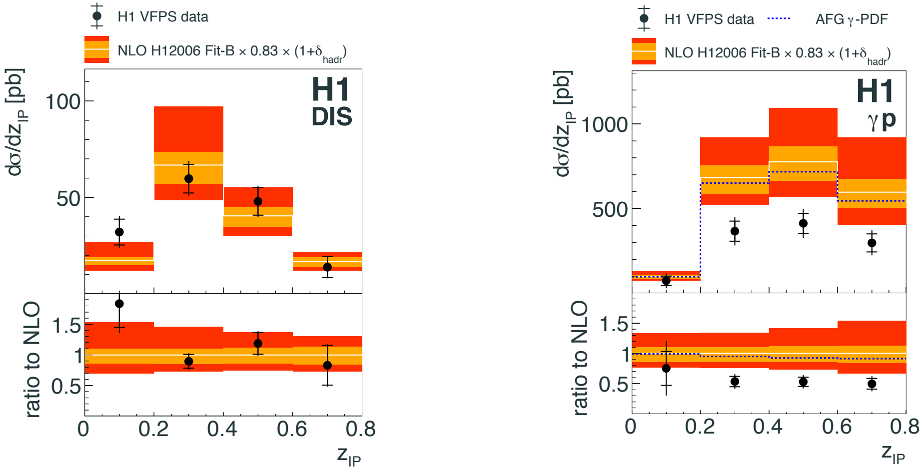
Predictions for multiplicity distributions in EIC



Diffractive processes

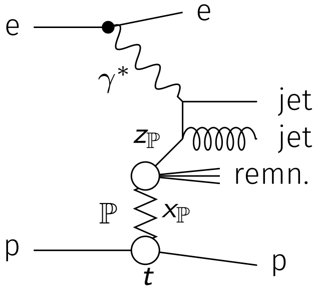
Diffractive dijets at HERA

[H1: JHEP 1505 (2015) 056]



- Good agreement between H1 data and NLO calculation in DIS regime (high-Q^2)
- NLO overshoot the data in photoproduction (low-Q^2)

Hard diffraction in DIS



Diffractive dijets

- Virtual photon interacts with Pomeron from proton producing jets
- Signature: scattered proton or a rapidity gap between proton and Pomeron remnant

Factorized cross section for diffractive dijets

- DIS: $d\sigma^{2\text{jets}+X} = f_i^{IP}(z_{IP}, \mu^2) \otimes f_{IP}^{p}(x_{IP}, t) \otimes d\sigma^{ie\rightarrow 2\text{jets}}$

  where f_{IP}^{p} is Pomeron flux and f_i^{IP} diffractive PDF (dPDF)
- Factorization verifed by H1 and ZEUS at HERA

Hard diffraction in photoproduction

Factorization-based approach

- Direct:

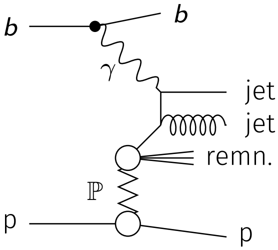
$$d\sigma^{2\text{jets}} = f_\gamma^b(x) \otimes d\sigma^{\gamma j\to 2\text{jets}} \otimes f_j^{\mathbb{P}}(z_{\mathbb{P}}, \mu^2) \otimes f_{\mathbb{P}}^{p}(x_{\mathbb{P}}, t)$$

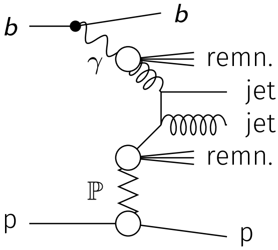
- Resolved:

$$d\sigma^{2\text{jets}} = f_\gamma^b(x) \otimes f_i^{\gamma}(x_\gamma, \mu^2) \otimes d\sigma^{ij\to 2\text{jets}} \otimes f_j^{\mathbb{P}}(z_{\mathbb{P}}, \mu^2) \otimes f_{\mathbb{P}}^{p}(x_{\mathbb{P}}, t)$$

Factorization breaking

- Suppression wrt. factorized approach around 10%–50% at HERA
- Even larger effects seen in pp (and p$\bar{\text{p}}$)
- Potential explanation additional interactions between photon remnants and the proton covering the rapidity gap





Hard diffraction in photoproduction

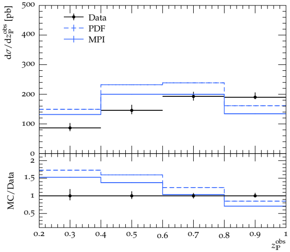
Pythia [I.H., C. O. Rasmussen, EPJC (2019) 79:413]

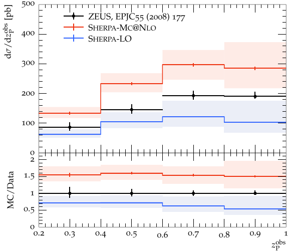
- Based on diffractive PDFs at LO
- Factorization breaking effects with dynamical rapidity gap survival with MPI rejection
  $\Rightarrow$ Suppression in line with the HERA data

Sherpa [F. Krauss, P. Meinzinger, EPJC 84 (2024) 9, 894]

- Both LO and NLO available
- NLO tend to overshoot the data
- Factorization breaking effects studies by scaling resolved and direct components

Three Rivet routines available (2 for H1, 1 ZEUS)





Tuning

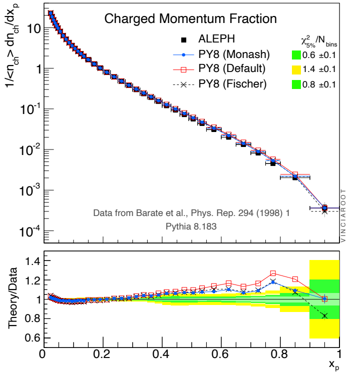
MC tuning

Need for tuning

- Modelling complete collision events require phenomenological models
- These involve parameters that have to fixed using experimental data
- Should be “global” to retain predictability of a given model (eg. energy dependence)

Tools for automated tuning

- Rivet provides easy comparison between data and simulations
- Professor 2 provides Rivet-based framework to optimize parameters by minimizing χ^2



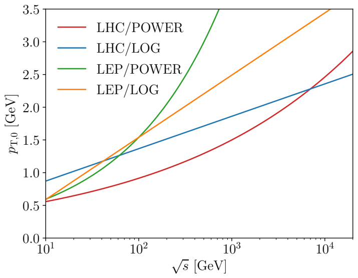
[P. Skands, S. Carrazza, J. Rojo, EPJC 74(8), 3024 (2014)]

An example: Tuning MPIs in photoproduction

[J.M. Butterworth, I. H., J.J. Juan Castella, B. Pattengale, S. Sanjrani, M. Wing: SciPost Phys. 17 (2024) 6, 158]

Systematic comparisons of existing MPI tunes

- Vary $p_{T,0}$ parametrization
- pp at LHC and Tevatron and for $\gamma\gamma$ from LEP
- Data for jet and charged-particle production for pp, γp and $\gamma\gamma$ (10 data sets in total)



An example: Tuning MPIs in photoproduction

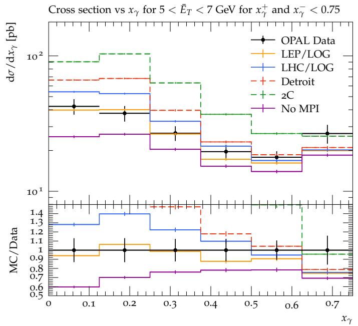
[J.M. Butterworth, I. H., J.J. Juan Castella, B. Pattengale, S. Sanjrani, M. Wing: SciPost Phys. 17 (2024) 6, 158]

Systematic comparisons of existing MPI tunes

- Vary $p_{\mathrm{T},0}$ parametrization
- pp at LHC and Tevatron and for $\gamma\gamma$ from LEP
- Data for jet and charged-particle production for pp, γp and $\gamma\gamma$ (10 data sets in total)

Conclusions

- Can find good agreement for $\gamma\gamma$



[OPAL: EPJC 31, 307 (2003)]

An example: Tuning MPIs in photoproduction

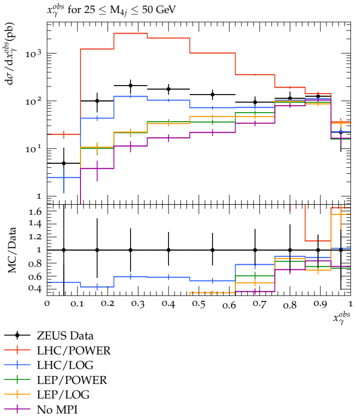
[J.M. Butterworth, I. H., J.J. Juan Castella, B. Pattengale, S. Sanjrani, M. Wing: SciPost Phys. 17 (2024) 6, 158]

Systematic comparisons of existing MPI tunes

- Vary $p_{\mathrm{T},0}$ parametrization
- pp at LHC and Tevatron and for $\gamma\gamma$ from LEP
- Data for jet and charged-particle production for pp, γp and $\gamma\gamma$ (10 data sets in total)

Conclusions

- Can find good agreement for $\gamma\gamma$ and γp



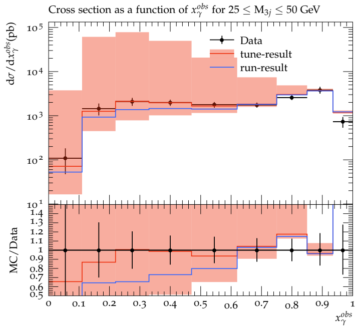
[ZEUS: NPB 792 1 (2008)]

An example: Tuning MPIs in photoproduction

[J.M. Butterworth, I. H., J.J. Juan Castella, B. Pattengale, S. Sanjrani, M. Wing: SciPost Phys. 17 (2024) 6, 158]

Systematic comparisons of existing MPI tunes

- Vary $p_{T,0}$ parametrization
- pp at LHC and Tevatron and for $\gamma\gamma$ from LEP
- Data for jet and charged-particle production for pp, γp and $\gamma\gamma$ (10 data sets in total)



[In progress]

Conclusions

- Can find good agreement for $\gamma\gamma$ and γp
- Published new Rivet analyses enabling dedicated tunes for each beam configuration
- Automatized parameter optimization with Professor 2

MC4EIC recap

MC4EIC

- Monte Carlo event generators required for detector planning and analysis
- Follow up MC development relevant to EIC with MC working group in the EICUG

Previous workshops, following MCEGs in 2018 and 2019

- MC4EIC 2021, Remote, hosted by CFNS
  - Kick of to review experimental needs for theory and event generators
- MC4EIC 2022, Remote, hosted by BNL
  - Reports from MC developers and experimentalists, live notes
- MC4EIC 2024, In-person meeting in Durham
  - Reports from general purpose and specialized event generators
  - Reviews on the existing data relevant to validation
- MC4EIC 2025, Hybrid in JLab (connected to EICUG meeting the following week)
  - Overview talks and generator updates, focus on Rivet and validation
  - Draft a report from the validation efforts

MC4EIC

- Monte Carlo event generators required for detector planning and analysis
- Follow up MC development relevant to EIC with MC working group in the EICUG

Previous workshops, following MCEGs in 2018 and 2019

- MC4EIC 2021, Remote, hosted by CFNS
  - Kick of to review experimental needs for theory and event generators
- MC4EIC 2022, Remote, hosted by BNL
  - Reports from MC developers and experimentalists, live notes
- MC4EIC 2024, In-person meeting in Durham
  - Reports from general purpose and specialized event generators
  - Reviews on the existing data relevant to validation
- MC4EIC 2025, Hybrid in JLab (connected to EICUG meeting the following week)
  - Overview talks and generator updates, focus on Rivet and validation
  - Draft a report from the validation efforts

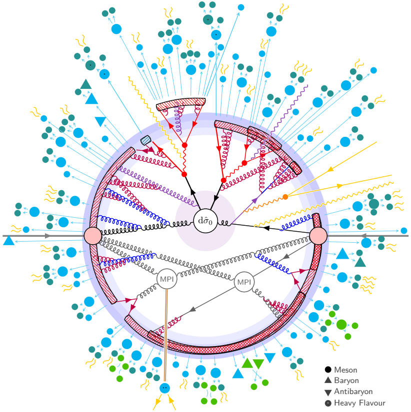
Summary & Outlook

Many recent developments in different areas relevant for EIC

- Extend capabilities for different processes in event generators
- Precision improvements with matching and merging
- First validation and tuning efforts completed/ongoing

Things to work on

- Radiative effects, nuclear targets, diffraction
- What else? How to communicate experimental needs? Rivet analyses?



[figure by P. Skands]