Soft QCD models and general-purpose Monte Carlo simulation

Andrzej Siódmok

CERN, Theory Division & IFJ, Cracow

General-purpose Monte Carlo

- Monte Carlo simulations are used by all experimental collaborations both to compare their data and theoretical predictions, and in data analysis.
- Unfortunately they are often treated as black boxes ...
  J. D. Bjorken
  "But it often happens that the physics simulations provided by the the MC generators carry the authority of data itself. They look like data and feel like data, and if one is not careful they are accepted as if they were data."
- It's important to understand the assumptions and approximations involved in these simulations.
- It is important to understand what is inside the programs to be able to answer the following type of questions.
  - Is the effect I'm seeing due to different models, or approximations, or is it a bug?
  - Am I measuring a fundamental quantity or merely a parameter in the simulation code?

"General-purpose event generators for LHC physics", MC authors [arXiv:1101.2599]

What do general-purpose Monte Carlo generators do?

- An “event“ is a list of particles (pions, protons, ...) with their momenta.
- The MCs generate events.
- The probability to generate an event is proportional to the (approximate!) cross section for such an event.
- Calculate Everything $\sim$ solve QCD $\rightarrow$ requires compromise!
- Improve lowest-order perturbation theory, by including the ”most significant“ corrections $\rightarrow$ complete events (can evaluate any observable you want)

The Workhorses: What are the Differences?

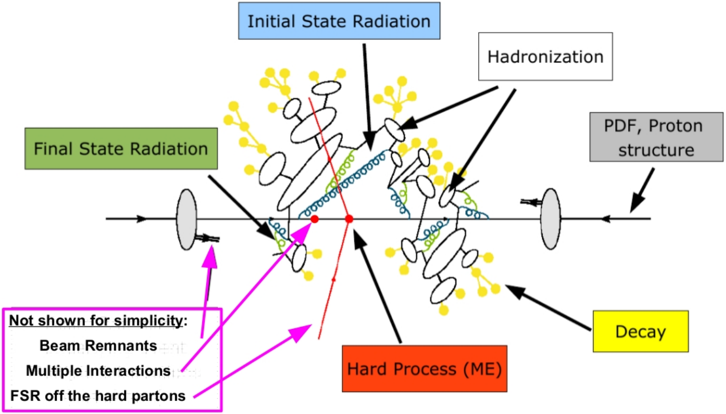
All offer convenient frameworks for LHC physics studies, but with slightly different emphasis:

PYTHIA: Successor to JETSET (begun in 1978). Originated in hadronization studies: Lund String.

HERWIG: Successor to EARWIG (begun in 1984). Originated in coherence studies: angular ordering parton shower. Cluster model.

SHERPA: Begun in 2000. Originated in ”matching” of matrix elements to showers: CKKW.

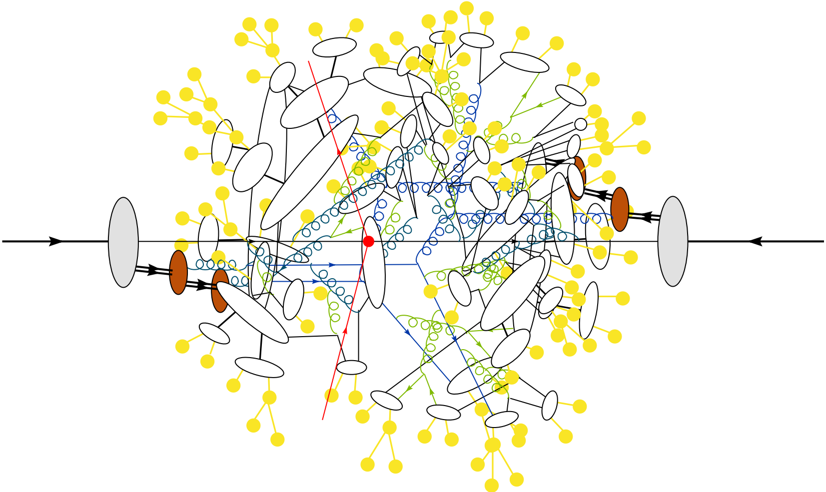
Basics of Monte Carlo Generators



taken from Stefan Gieseke©

The general approach is the same in different programs but the models and approximations used are different.

Soft QCD models - Multiple Partonic Interactions (MPI)



Motivation - how do we know MPI exists?

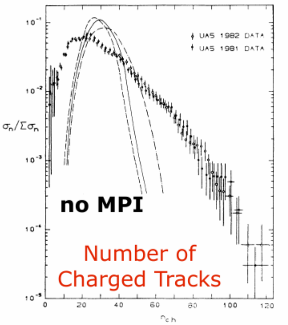


FIG. 3. Charged-multiplicity distribution at 540 GeV, UA5 results (Ref. 32) vs simple models: dashed low p_T only, full including hard scatterings, dash-dotted also including initial- and final-state radiation.

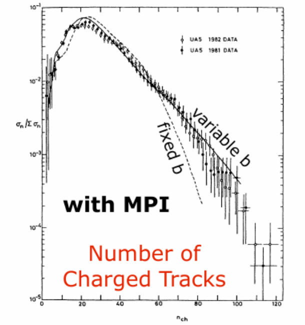


FIG. 12. Charged-multiplicity distribution at 540 GeV, UA5 results (Ref. 32) vs multiple-interaction model with variable impact parameter: solid line, double-Gaussian matter distribution; dashed line, with fix impact parameter [i.e., $\bar{O}_0(b)$].

Sjöstrand & v. Zijl, Phys.Rev.D36(1987)2019

Direct evidence: measurement of momentum imbalance in multijet events at CERN ISR and $\gamma + 3$ jet at TVT

[AFS (1987), UA2 (1991), CDF (1993, 1997), D0 (2009)]

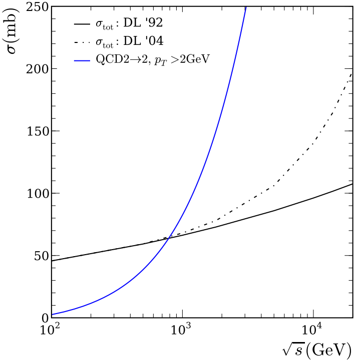
Motivation - is it really important?

Motivation:

- The minimum bias/underlying event is an unavoidable background to most collider observables and having good understand of it leads to more precise collider measurements!
- First LHC results are Minimum Bias and Underlying Event! Alice: [0911.5430], CMS [1002.0621], ATLAS [1003.3124] so it must be important ;)
- These will be particularly relevant for the LHC as, when it is operated at design luminosity, rare signal events will be embedded in a background of more than 20 near-simultaneous minimum-bias collisions.
- Any realistic experiment simulation event generator needs to be able to model these effects.
- “Don’t worry, we will measure and subtract it” But... fluctuations and correlations on an event-by-event basis are crucial.

Inclusive hard jet cross section in pQCD:

$$\sigma^{\text{inc}}(s, p_t^{\min}) = \sum_{i,j} \int_{p_t^{\min 2}} \mathrm{d}p_t^2 \int \mathrm{d}x_1 \mathrm{d}x_2 \; f_i(x_1, Q^2)\, f_j(x_2, Q^2) \; \frac{\mathrm{d}\hat{\sigma}_{ij}}{\mathrm{d}p_t^2}$$



$\sigma^{\text{inc}} > \sigma_{\text{tot}}$ eventually

Interpretation:

- σ^{inc} counts all partonic scatters in a single *pp* collision
- more than a single interaction

$$\sigma^{\text{inc}} = \langle n_{\text{dijets}} \rangle \sigma_{\text{inel}}$$

MPI model basics (Herwig++)

Assumptions:

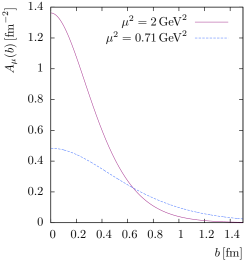
- the distribution of partons in hadrons factorizes with respect to the b and x dependence ⇒ average number of parton collisions:

$$
\begin{aligned}
\bar{n}(\vec{b}, s) &= L_{\text{partons}}(x_1, x_2, \vec{b}) \otimes \sum_{ij} \int \mathrm{d}p_t^2 \frac{\mathrm{d}\hat{\sigma}_{ij}}{\mathrm{d}p_t^2} \\
&= \sum_{ij} \frac{1}{1+\delta_{ij}} \int \mathrm{d}x_1 \mathrm{d}x_2 \int \mathrm{d}^2\vec{b}' \int \mathrm{d}p_t^2 \frac{\mathrm{d}\hat{\sigma}_{ij}}{\mathrm{d}p_t^2} \\
&\quad \times D_{i/A}(x_1, p_t^2, |\vec{b}'|) D_{j/B}(x_2, p_t^2, |\vec{b} - \vec{b}'|) \\
&= \sum_{ij} \frac{1}{1+\delta_{ij}} \int \mathrm{d}x_1 \mathrm{d}x_2 \int \mathrm{d}^2\vec{b}' \int \mathrm{d}p_t^2 \frac{\mathrm{d}\hat{\sigma}_{ij}}{\mathrm{d}p_t^2} \\
&\quad \times f_{i/A}(x_1, p_t^2) G_A(|\vec{b}'|) f_{j/B}(x_2, p_t^2) G_B(|\vec{b} - \vec{b}'|) \\
&= A(\vec{b}) \sigma^{\text{inc}}(s; p_t^{\min}) \,.
\end{aligned}
$$

- at fixed impact parameter b, individual scatterings are independent (leads to the Poisson distribution)

Underlying event in Herwig++ - key components

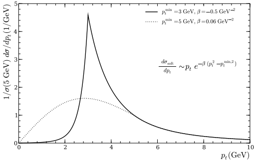
Matter distribution (μ^2)



Based on electromagnetic form factor

(radius of the proton free parameter)

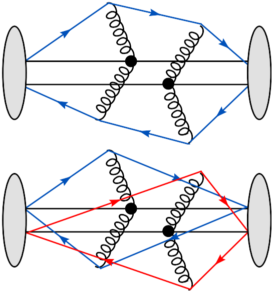
Extension to soft MPI ($p_t < p_t^{\min}$)



Gaussian extension below $p_t^{\min}$

Energy dependent $p_t^{\min}$

Colour structure (p_{reco}, p_{CD})



Possibility of change of color structure (color reconnection)

[Gieseke, Röhr, AS, EPJC 72 (2012)]

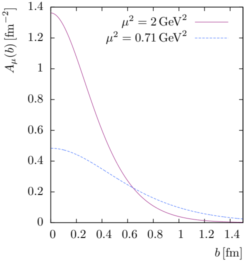
The least understood part of modeling

Main parameters:

- μ^2 - inverse hadron radius squared (parametrization of overlap function)
- $p_t^{\min}$ - transition scale between soft and hard components $\Rightarrow p_t^{\min} = p_{t,0}^{\min}\left(\frac{\sqrt{s}}{E_0}\right)^b$
- p_{reco} - colour reconnection
- p_{CD} - colour structure of the Soft UE

Underlying event in Herwig++ - key components

Matter distribution (μ^2)



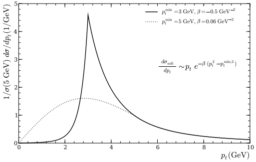
Based on electromagnetic form factor (radius of the proton free parameter)

Pythia:

- Many options including double Gaussian (similar shape to EE)

- x-depended overlap [Corke, Sjostrand, JHEP 1105:009]

Extension to soft MPI ($p_t < p_t^{\mathrm{min}}$)



Gaussian extension below p_t^{min}

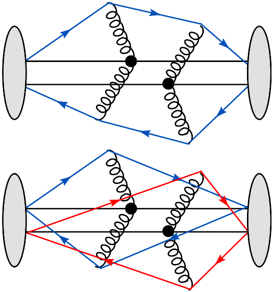
Energy dependent p_t^{min}

Pythia:

Regularise cross section with p_t^{min} as free parameter:

$$\frac{d\sigma}{dp_T^2} \propto \frac{\alpha^2(p_T^2)}{p_T^4} \rightarrow \frac{\alpha^2(p_T^2 + p_t^{\mathrm{min}\,2})}{(p_T^2 + p_t^{\mathrm{min}\,2})^2}$$

Colour structure (p_{reco}, p_{CD})



Possibility of change of color structure (color reconnection)
[Gieseke, Röhr, AS, EPJC 72 (2012)]
The least understood part of modeling (very active area research)

Pythia:

The most recent development: String Formation Beyond Leading Colour J. Christiansen, P. Skands [arXiv:1505.01681]

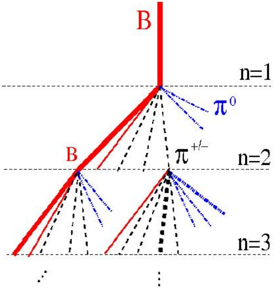
Min Bias/Underlying Event

Herwig++ MPI model with independent hard and soft processes, showered and with colour reconnection. Just few parameters. Min bias without integrated diffraction (work in progress).

Pythia MPI interleaved with showering. MPI ordered in p_T. Many options and parameters (Pythia has strong emphasis on NP physics) $\Rightarrow$ many tune families.

Sherpa New model - SHRiMPS with integrated diffraction based on KMR (Khoze-Martin-Ryskin model). Model in development - currently not suitable for UE studies (work in progress - this winter?). Currently for UE there is "cheap version of Pythia's UE model" (F. Krauss)

Air Shower Simulation

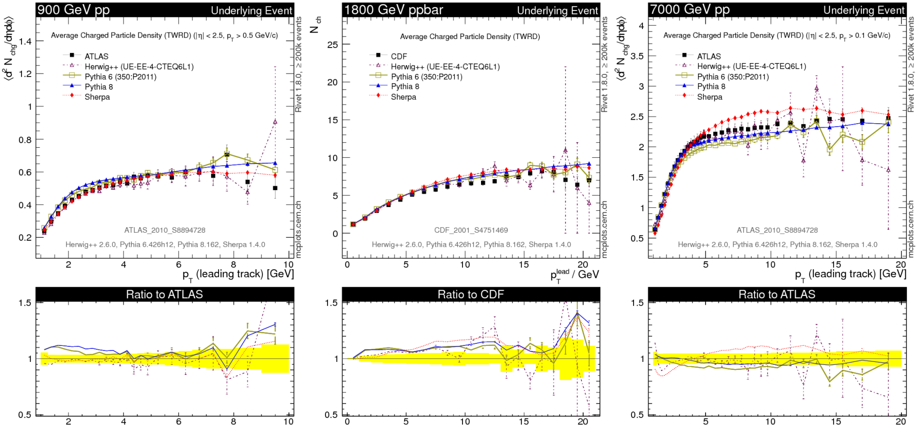


Thickness = amount of energy

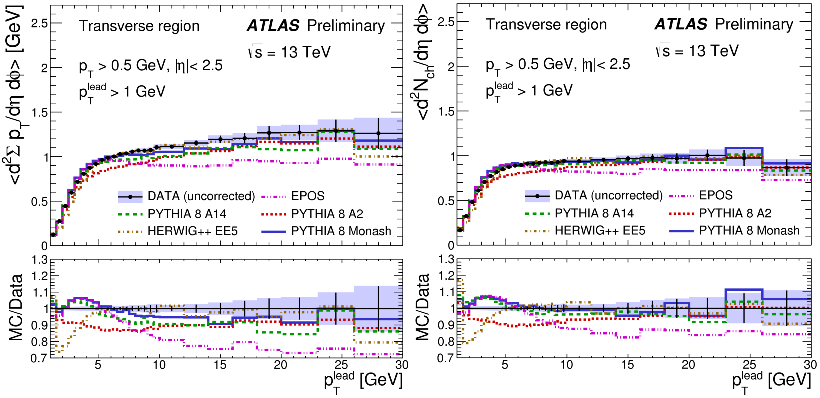
- **Hadronic models for simulations :**
  - mainly soft physics + diffraction (forward region)
  - should handle p-, π-Air, K-Air and A-Air interactions
  - should be able to run at 10^6 GeV center-of-mass energy
  - models used for EAS analysis :
    - QGSJET01/II
    - SIBYLL 2.1
    - EPOS
    - ...

Quite different model to Pythia/Herwig, for example no color reconnection but collective hadronization instead.

UE measurements - Energy Overview

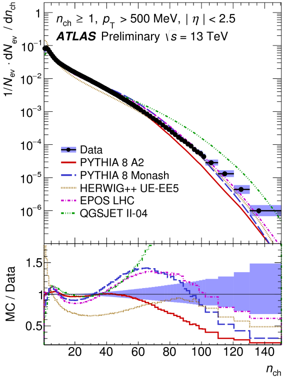


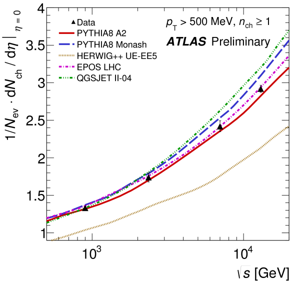
Run II results - UE [ATL-PHYS-PUB-2015-019]



- Many LHC UE observables (not tuned since not available) and well described by most of the models!
- EPOS has no hard component - not surprise that it does not describe UE (authors are working to improve it)

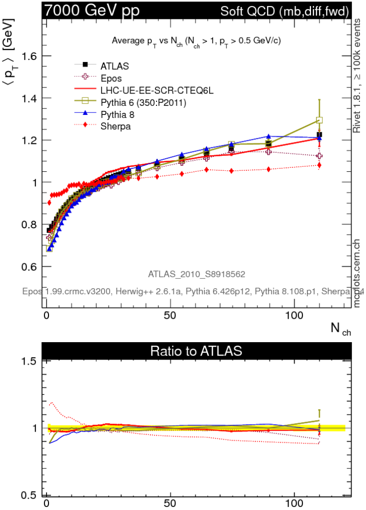
Run II results - MB [ATLAS-CONF-2015-028]





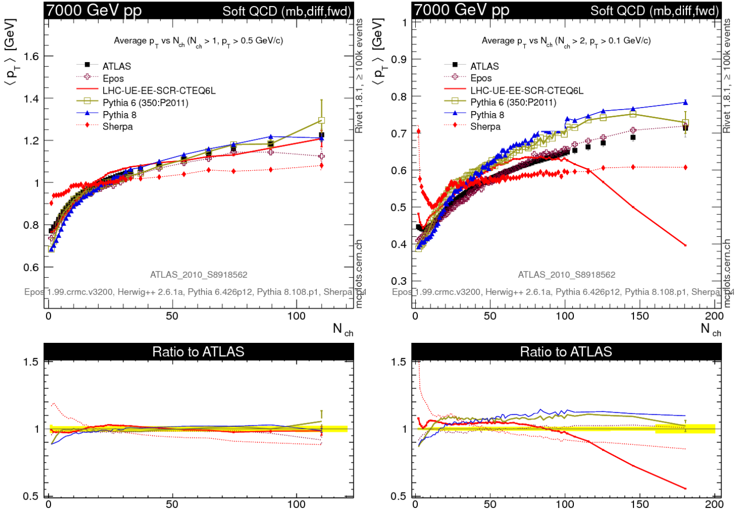
- Hard to get very good description of both MB and UE data!
- However as you will see on Friday (talks by S. Martin-Haugh and E. Kuznetsova) Epos and Pythia do reasonable job.
- Not surprising that Hw++ does not describe $dN_{ch}/d\eta$ vs $\sqrt{(s)}$ - no diffraction model (as expected there is a space for it!) also not tuned to MB data (tuned only to UE data)

Problems - very soft MinBias ATLAS



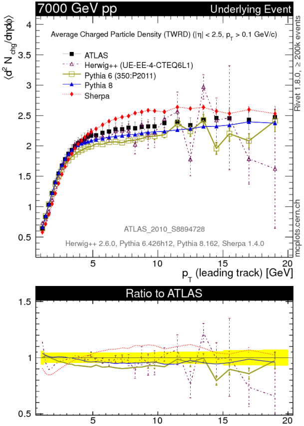
Need of the colour reconnection.

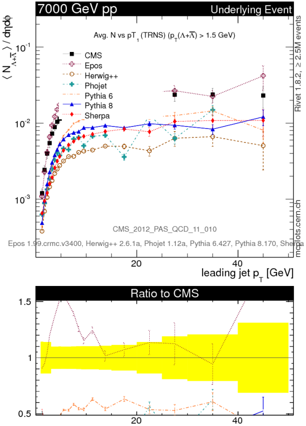
Problems - very soft MinBias ATLAS



Need of the colour reconnection.
MB 7000 TeV, problem at low p_T, high Nch
Epos seems to describe MB data but fails to describe UE data.

Problems - Identified particles



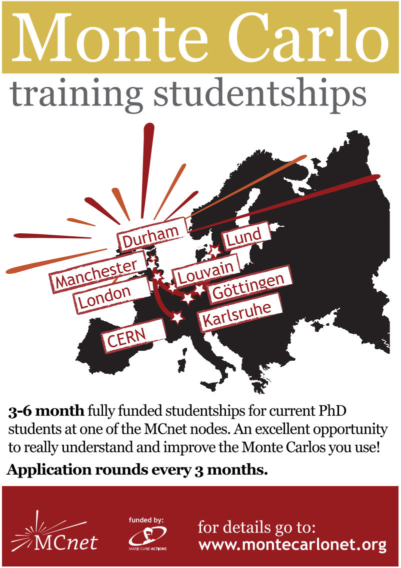


More plots: mcplots.cern.ch

Summary:

- Non perturbative regime -> need for models with several parameters, no unique way -> few models on the market
- Minimum bias/underlying event/diffraction under constant improvement (new MPI model Shrimps in Sherpa, improvements in Pythia and Herwig, Epos for LHC)!
- Good first round of LHC data well described...
- ... but still a lot space for improvements.
- Not-too-soft not-too-high-multiplicity physics under good control (if you use modern models with modern tunes).
- As LHC needs to study more rare phenomena and more subtle effects, generators must keep up by increased precision.

MCnet Short-term studentships



MC*net* projects

Pythia
Herwig
Sherpa
MadGraph
Ariadne
CEDAR

Thank you for the attention!

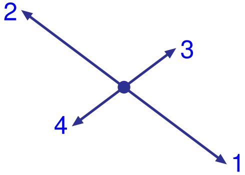
Motivation - how do we know MPI exists?

Direct observation of multiple interactions

Five studies: AFS (1987), UA2 (1991), CDF (1993, 1997), D0 (2009)

Order 4 jets $\mathbf{p}_{\perp 1} > \mathbf{p}_{\perp 2} > \mathbf{p}_{\perp 3} > \mathbf{p}_{\perp 4}$ and define φ
as angle between $\mathbf{p}_{\perp 1} \mp \mathbf{p}_{\perp 2}$ and $\mathbf{p}_{\perp 3} \mp \mathbf{p}_{\perp 4}$ for AFS/CDF

Double Parton Scattering

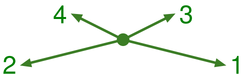


$|\mathbf{p}_{\perp 1} + \mathbf{p}_{\perp 2}| \approx 0$

$|\mathbf{p}_{\perp 3} + \mathbf{p}_{\perp 4}| \approx 0$

$\mathrm{d}\sigma/\mathrm{d}\varphi$ flat

Double BremsStrahlung



$|\mathbf{p}_{\perp 1} + \mathbf{p}_{\perp 2}| \gg 0$

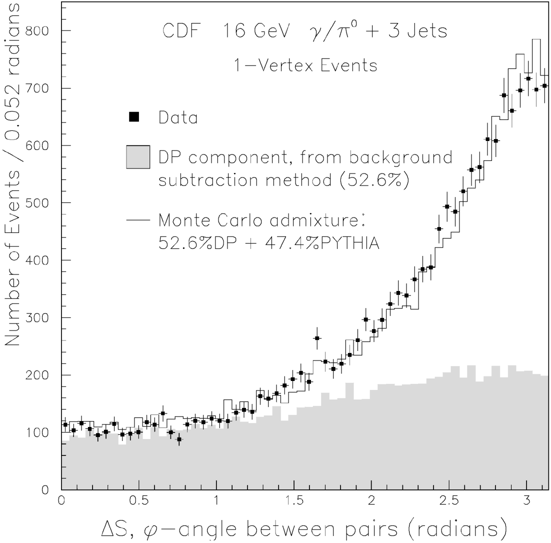
$|\mathbf{p}_{\perp 3} + \mathbf{p}_{\perp 4}| \gg 0$

$\mathrm{d}\sigma/\mathrm{d}\varphi$ peaked at $\varphi \approx 0/\pi$ for AFS/CDF

Motivation - how do we know MPI exists?

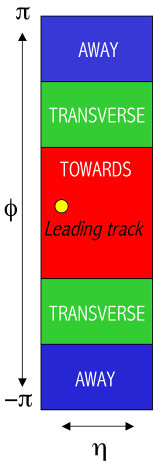
CDF: Double parton scattering in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$

[Phys. Rev. D 56, 3811-3832 (1997)]



Motivation - how do we know MPI exists?

CDF Run II

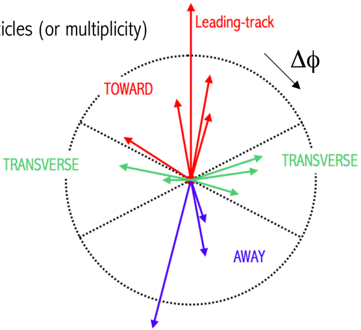


On event-by-event basis:

1) Identify the leading object in the event

2) Build **TRANSVERSE REGIONS** w.r.t. it

3) Compute Σp_T of charged particles (or multiplicity) in the different regions



SETTINGS:

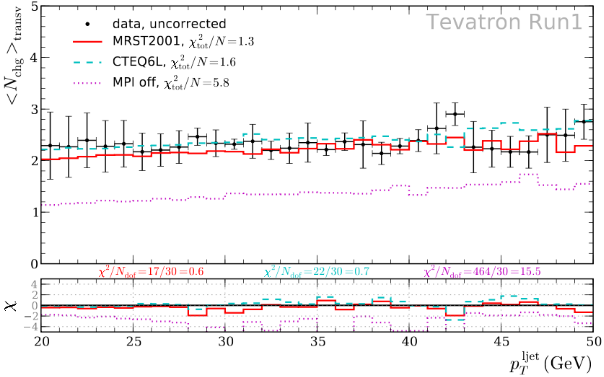
- pT > 0.5 GeV/c (tracks and leading-track)

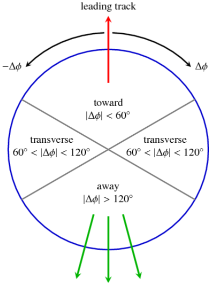
- leading-track not included in distributions

Motivation - how do we know MPI exists?

CDF Run II

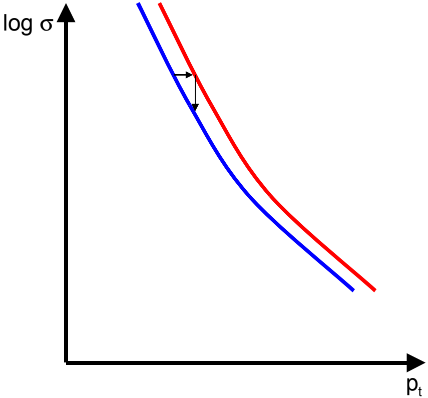
Good description of Run I Underlying event data ($\chi^2 = 1.3$).





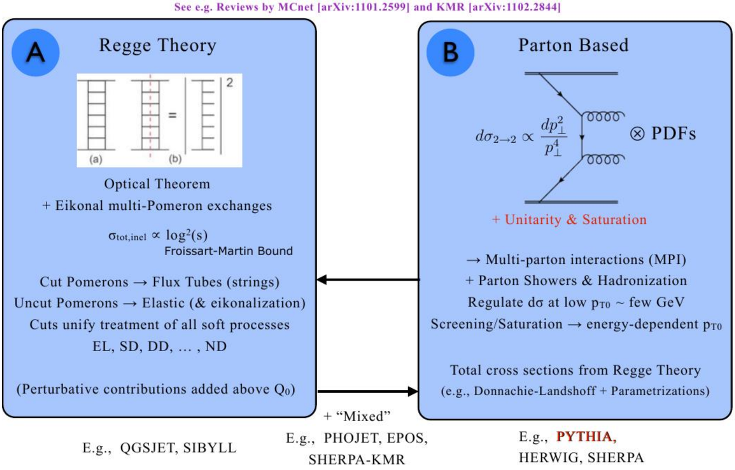
Only $p_T^{ljet} > 20\,\text{GeV}$.

Motivation - is it really important?



- Steep distribution ⇒ small sideways shift = large vertical
- Rare fluctuations can have a huge influence

MPI models overview and comparison with data



EPOS, Herwig++, Pythia and Sherpa mostly used at the LHC.