The PYTHIA 8 Event Generator

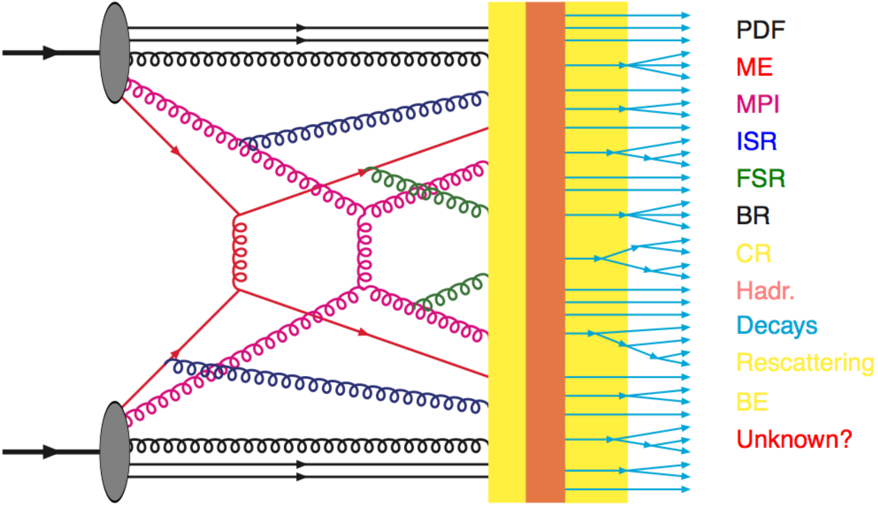
Torbjörn Sjöstrand

Department of Astronomy and Theoretical Physics
Lund University
Sölvegatan 14A, SE-223 62 Lund, Sweden

LHC meets Cosmic Rays school (ISAPP 2018),
CERN, 28 Oct – 2 Nov 2018

The structure of an event

An event consists of many different physics steps to be modelled:



Colliding beams

- e^+e^- (e.g. LEP)
- pp & $p\overline{p}$ (e.g. LHC)
- $e^\pm p$ (e.g. HERA), but either DIS or photoproduction
- $\ell\ell$ & $\ell\overline{\ell}$ ($\ell = e, \mu, \tau, \nu$)
- hh & $h\overline{h}$ ($h = p, n, \pi^{\pm,0}$, limited by PDFs)
- $\ell h, \gamma h, \gamma\gamma$, where γ direct or resolved
- pA, AA (Angantyr; recent)

Limitations:

- Only one beam combination at a time.
- Only one CM energy at a time (or small range).
- $E_{cm} > 10$ GeV.
- No air shower tracking.

Core processes

Available hardcoded internally, almost freely mixable:

- Soft QCD: elastic, single diffractive, double diffractive, central diffractive, nondiffractive (including hard processes)
- Hard QCD: $2 \to 2$ (e.g. $\mathrm{qg} \to \mathrm{qg}$), open heavy flavours, charmonium, bottomonium, top, ($2 \to 3$)
- Electroweak: $\mathrm{f\bar{f}} \to \gamma^*/\mathrm{Z}^0$, $\mathrm{f\bar{f}} \to \mathrm{W}^+\mathrm{W}^-$, $\mathrm{qg} \to \mathrm{q}\gamma$, $\mathrm{f\bar{f}} \to \gamma\gamma$, $\ell\mathrm{q} \to \ell\mathrm{q}$, $\mathrm{q}\gamma \to \mathrm{qg}$, $\gamma\gamma \to \mathrm{f\bar{f}}$, ...
- Higgs in the SM and various extensions
- BSM: SUSY, new gauge bosons, left–right symmetry, leptoquarks, compositeness, hidden valleys, extra dimensions, dark matter

Beyond that: no internal ME generator, so then external input, e.g. MadGraph5_aMC@NLO, PowHeg Box, AlpGen, typically using Les Houches Event Files exchange standard.

Parton Distribution Functions

Trend towards NLO/NNLO parametrizations: need not be positive definite, notably small-x gluon at low Q^2, which is nuisance.
Coming: NNPDF3.1sx+LHCb (N)NLO+NLLx QED.

Internal implementation of several PDFs

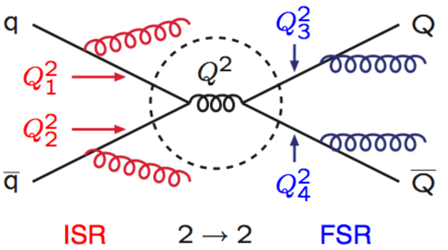
- p: 21 sets, from legacy to new (2017), mainly LO
- n: by isospin (watch out for QED)
- nuclear modification factors (EPS09 LO/NLO, EPPS16 NLO)
- π: GRV 92L (isospin for $\pi^+ \to \pi^0$)
- Pomeron (diffraction): 15 sets
- γ: CJKL
- ℓ: QED (exponentiated)

Can also

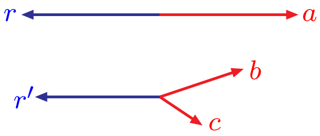
- link to whole LHAPDF library, or
- read single `.dat` file of one PDF set/member.

Parton showers – 1

$2 \to n \;=\; (2 \to 2) \oplus \text{ISR} \oplus \text{FSR}$



Dipole recoil (for FSR):



$p_b + p_c + p'_r = p_a + p_r$

Based on DGLAP evolution equations:

$$\mathrm{d}\mathcal{P}_{a\to bc} = \frac{\alpha_\mathrm{s}}{2\pi}\frac{\mathrm{d}Q^2}{Q^2} P_{a\to bc}(z)\,\mathrm{d}z \cdot (\text{Sudakov})$$

with $p_\perp$ ordering, $Q^2 = p^2_{\perp\mathrm{evol}} \approx p^2_\perp$, and dipole recoils.

ISR by backwards evolution from the hard interaction.

Parton showers – 2

Currently three (main) parton shower options;

- Internal default `SpaceShower + TimeShower`;
- VINCIA antenna shower plugin;
- DIRE dipole shower plugin.

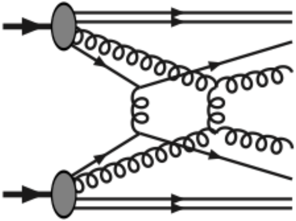
Same basic structure, e.g. MPI + ISR + FSR interleaved evolution:

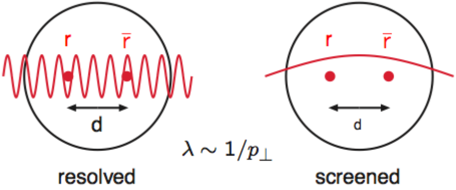
$$\frac{\mathrm{d}\mathcal{P}}{\mathrm{d}p_\perp} = \left(\frac{\mathrm{d}\mathcal{P}_{\mathrm{MPI}}}{\mathrm{d}p_\perp} + \sum \frac{\mathrm{d}\mathcal{P}_{\mathrm{ISR}}}{\mathrm{d}p_\perp} + \sum \frac{\mathrm{d}\mathcal{P}_{\mathrm{FSR}}}{\mathrm{d}p_\perp} \right) \times \exp\left(- \int_{p_\perp}^{p_{\perp\mathrm{max}}} \left(\frac{\mathrm{d}\mathcal{P}_{\mathrm{MPI}}}{\mathrm{d}p'_\perp} + \sum \frac{\mathrm{d}\mathcal{P}_{\mathrm{ISR}}}{\mathrm{d}p'_\perp} + \sum \frac{\mathrm{d}\mathcal{P}_{\mathrm{FSR}}}{\mathrm{d}p'_\perp} \right) \mathrm{d}p'_\perp \right)$$

Support the same facilities, like

- matching and merging with higher-order matrix elements,
- automated uncertainty band from factorization and renormalization scale choices, and finite splitting-kernel terms.

Hadrons are composite ⇒ many partons can interact:





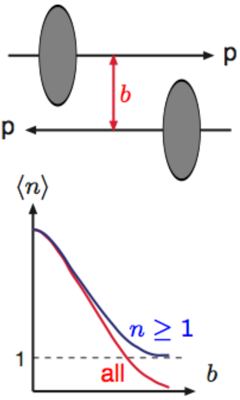
Divergence for $p_\perp \to 0$ in perturbative $2 \to 2$ scatterings;
tamed by unknown colour screening length d in hadron

$$\frac{\mathrm{d}\hat{\sigma}}{\mathrm{d}p_\perp^2} \propto \frac{\alpha_\mathrm{s}^2(p_\perp^2)}{p_\perp^4} \to \frac{\alpha_\mathrm{s}^2(p_{\perp 0}^2 + p_\perp^2)}{(p_{\perp 0}^2 + p_\perp^2)^2}$$

with $p_{\perp 0} \approx$ 2–3 GeV $\simeq 1/d$.

Semiperturbative $2 \to 2$ generates whole nondiffractive σ!?

Hadrons are extended, so dependence on impact parameter b.



Overlap of protons during encounter is

$$\mathcal{O}(b) = \int \mathrm{d}^3\mathbf{x}\,\mathrm{d}t\;\rho_1(\mathbf{x},t)\,\rho_2(\mathbf{x},t)$$

where ρ is (boosted) matter distribution in p, e.g. Gaussian or more narrow peak.

Average activity at b proportional to $\mathcal{O}(b)$:

⋆ central collisions more active
 ⇒ $\mathcal{P}_n$ broader than Poissonian;

⋆ peripheral passages normally give
 no collisions ⇒ finite σ_{tot}.

At LHC $\langle n_{\mathrm{MPI}} \rangle \approx 3$ for all events, but $\gtrsim 10$ for central collisions.

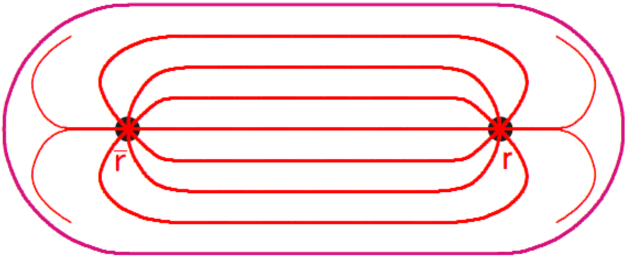
Preselected hard process ⇒ central ⇒ "pedestal effect".

MPIs in PYTHIA

- MPIs are gererated in a falling sequence of $p_\perp$ values; recall Sudakov factor approach to parton showers.
- Core process QCD $2 \to 2$, but also onia, γ's, $\mathrm{Z}^0, \mathrm{W}^\pm$.
- Energy, momentum and flavour conserved step by step: subtracted from proton by all "previous" collisions.
- Protons modelled as extended objects, allowing both central and peripheral collisions, with more or less activity.
- Colour screening increases with energy, i.e. $p_{\perp 0} = p_{\perp 0}(E_{\mathrm{cm}})$, as more and more partons can interact.
- Colour connections: each interaction hooks up with colours from beam remnants, but also correlations inside remnants.
- Colour *reconnections*: many interaction "on top of" each other $\Rightarrow$ tightly packed partons $\Rightarrow$ colour memory loss?

The QCD potential

In QCD, for large charge separation, field lines are believed to be compressed to tubelike region(s) $\Rightarrow$ **string(s)**



Gives force/potential between a q and a $\overline{\mathrm{q}}$:

$F(r) \approx \mathrm{const} = \kappa \quad \Longleftrightarrow \quad V(r) \approx \kappa r$

$\kappa \approx 1\ \mathrm{GeV/fm}$ $\approx$ potential energy gain lifting a 16 ton truck.

Flux tube parametrized by center location as a function of time $\Rightarrow$ simple description as a 1+1-dimensional object – a string.

String motion

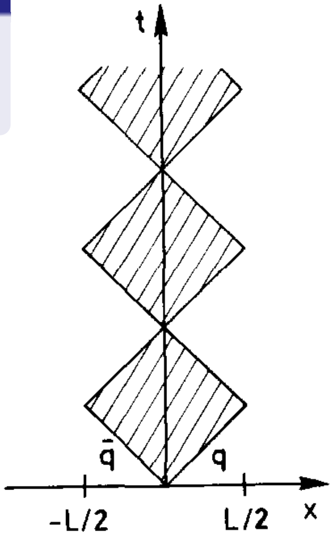
The Lund Model: starting point

Use only linear potential $V(r) \approx \kappa r$ to trace string motion, and let string fragment by repeated $\mathrm{q\overline{q}}$ breaks.

Assume negligibly small quark masses. Then linearity between space–time and energy–momentum gives

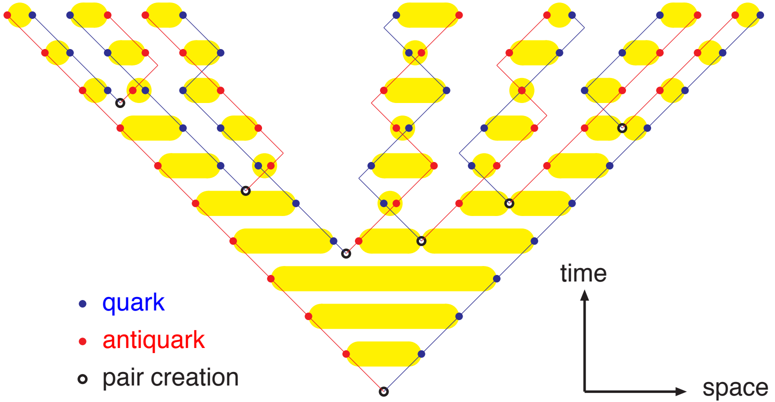
$$\left|\frac{\mathrm{d}E}{\mathrm{d}z}\right| = \left|\frac{\mathrm{d}p_z}{\mathrm{d}z}\right| = \left|\frac{\mathrm{d}E}{\mathrm{d}t}\right| = \left|\frac{\mathrm{d}p_z}{\mathrm{d}t}\right| = \kappa$$

$(c = 1)$ for a $\mathrm{q\overline{q}}$ pair flying apart along the $\pm z$ axis.
But signs relevant: the q moving in the $+z$ direction has $\mathrm{d}z/\mathrm{d}t = +1$ but $\mathrm{d}p_z/\mathrm{d}t = -\kappa$.



Combine yo-yo-style string motion with string breakings!

Motion of quarks and antiquarks with intermediate string pieces:

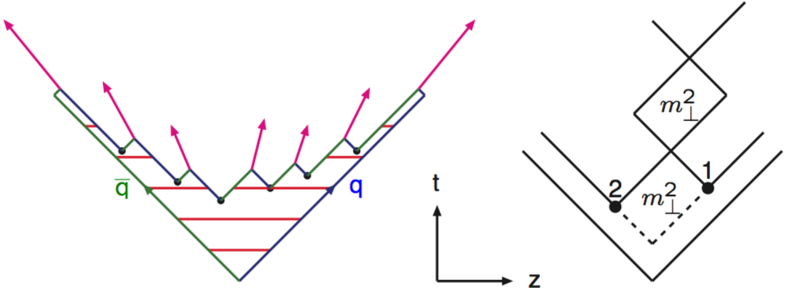


A q from one string break combines with a $\overline{\mathrm{q}}$ from an adjacent one.

Gives simple but powerful picture of hadron production.

Where does the string break?

Fragmentation starts in the middle and spreads outwards:



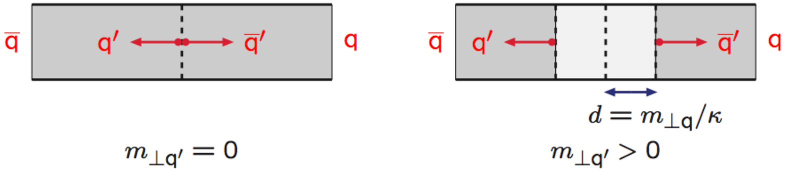
Corresponds to roughly same invariant time of all breaks,
$\tau^2 = t^2 - z^2 \sim$ constant,
with breaks separated by hadronic area $m_\perp^2 = m^2 + p_\perp^2$.

Hadrons at outskirts are more boosted.

Approximately flat rapidity distribution, $\mathrm{d}n/\mathrm{d}y \approx$ constant

$\Rightarrow$ total hadron multiplicity in a jet grows like $\ln E_{\mathrm{jet}}$.

How does the string break?



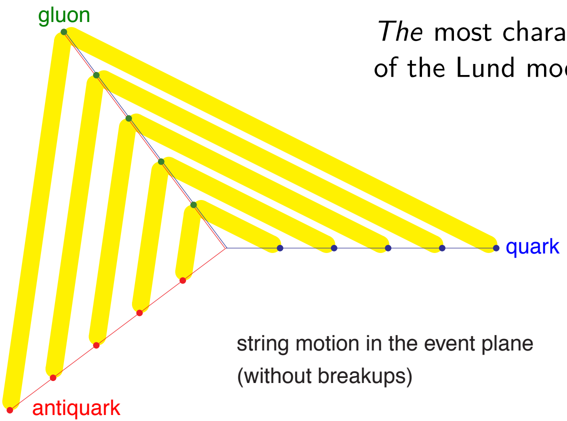
String breaking modelled by tunneling:

$$\mathcal{P} \propto \exp\left(-\frac{\pi m_{\perp q}^2}{\kappa}\right) = \exp\left(-\frac{\pi p_{\perp q}^2}{\kappa}\right) \exp\left(-\frac{\pi m_q^2}{\kappa}\right)$$

- Common Gaussian $p_\perp$ spectrum, $\langle p_\perp \rangle \approx 0.4$ GeV.
- Suppression of heavy quarks,
  $\mathrm{u\overline{u}} : \mathrm{d\overline{d}} : \mathrm{s\overline{s}} : \mathrm{c\overline{c}} \approx 1 : 1 : 0.3 : 10^{-11}$.
- Diquark $\sim$ antiquark $\Rightarrow$ simple model for baryon production.

String model unpredictive in understanding of hadron mass effects
$\Rightarrow$ many parameters, 10–20 depending on how you count.

The Lund gluon picture



string motion in the event plane
(without breakups)

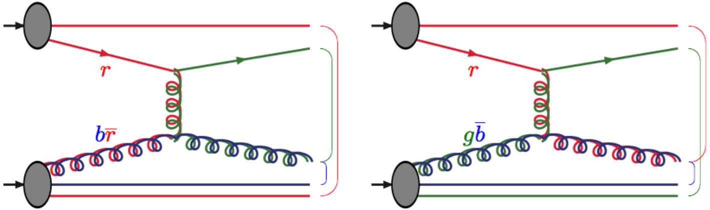
The most characteristic feature of the Lund model:

Gluon = kink on string

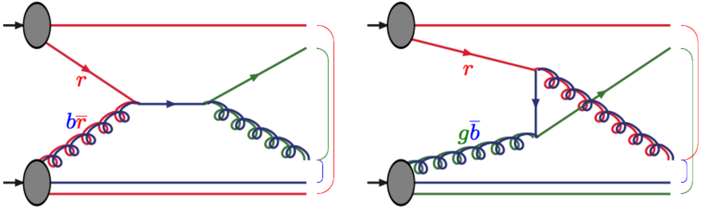
Force ratio gluon/ quark = 2,
cf. QCD $N_C/C_F = 9/4$, $\rightarrow 2$ for $N_C \rightarrow \infty$
No new parameters introduced for gluon jets!

Colour flow in hard processes

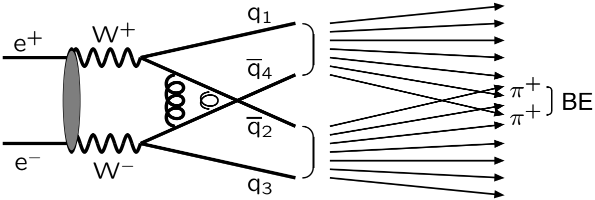
One Feynman graph can correspond to several possible colour flows, e.g. for $\mathrm{qg} \to \mathrm{qg}$:



while other $\mathrm{qg} \to \mathrm{qg}$ graphs only admit one colour flow:



Interference terms with indeterminate colour flow $\propto 1/N_C^2$.



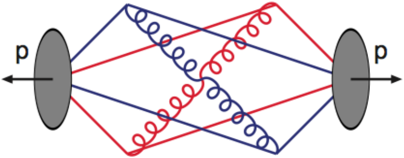
At LEP 2 search for effects in $\mathrm{e^+e^-} \to \mathrm{W^+W^-} \to \mathrm{q_1\overline{q}_2\, q_3\overline{q}_4}$:

- perturbative $\langle \delta M_{\mathrm{W}} \rangle \lesssim$ 5 MeV : negligible!
- nonperturbative $\langle \delta M_{\mathrm{W}} \rangle \sim$ 40 MeV :

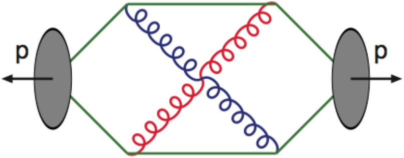
  favoured; no-effect option ruled out at 99.5% CL.

  Best description for reconnection in $\approx$ 50% of the events.
- Bose-Einstein $\langle \delta M_{\mathrm{W}} \rangle \lesssim$ 100 MeV : full effect ruled out (while models with $\sim$ 20 MeV barely acceptable).

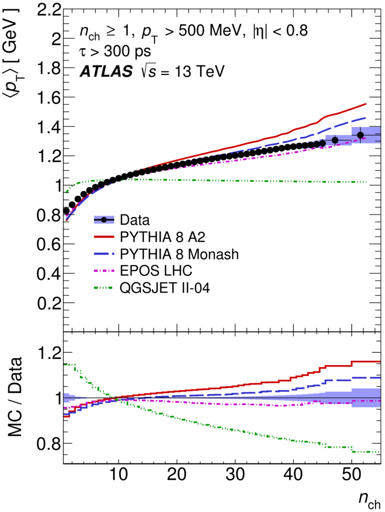
$\langle p_\perp \rangle(n_{\mathrm{ch}})$ is very sensitive to colour flow



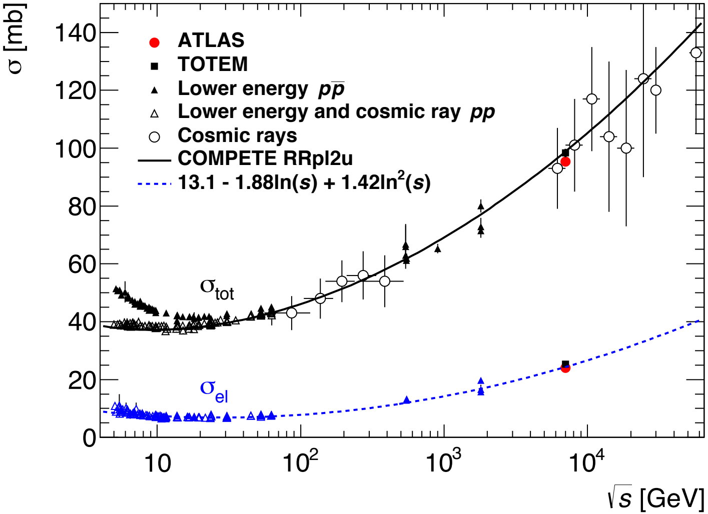
long strings to remnants $\Rightarrow$ much n_{ch}/interaction $\Rightarrow \langle p_\perp \rangle(n_{\mathrm{ch}}) \sim$ flat



short strings (more central) $\Rightarrow$ less n_{ch}/interaction $\Rightarrow \langle p_\perp \rangle(n_{\mathrm{ch}})$ rising



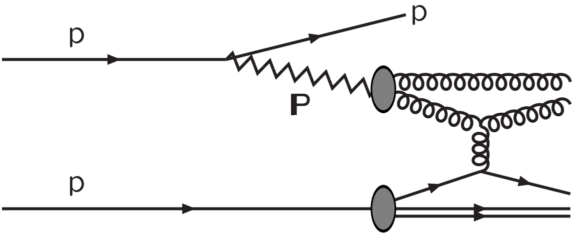
Total cross section



Several options for total and partial pp & $\mathrm{p\bar{p}}$ cross sections: DL/SaS, MBR, ABMST, RPP2016.

Diffraction

Ingelman-Schlein: Pomeron as hadron with partonic content
Diffractive event = (Pomeron flux) × ($\mathbb{P}$p collision)



Used e.g. in
POMPYT
POMWIG
PHOJET

1) σ_{SD}, σ_{DD} and σ_{CD} set by Reggeon theory.

2) $f_{\mathbb{P}/\mathrm{p}}(x_{\mathbb{P}}, t) \Rightarrow$ diffractive mass spectrum, $p_{\perp}$ of proton out.

3) Smooth transition from simple model at low masses to $\mathbb{P}$p with full pp machinery: multiparton interactions, parton showers, etc.

4) Choice between different Pomeron PDFs.

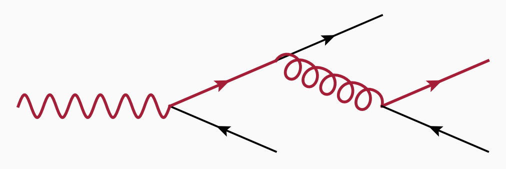
5) Free parameter $\sigma_{\mathbb{P}\mathrm{p}}$ needed to fix $\langle n_{\mathrm{interactions}} \rangle = \sigma_{\mathrm{jet}}/\sigma_{\mathbb{P}\mathrm{p}}$.

$\gamma\gamma$ and γp physics

Dual nature of photon: direct (pointlike) and resolved (hadronlike).
DGLAP evolution has additional term from $\gamma \to q\bar{q}$:

$$\frac{\mathrm{d}f_i^{\gamma}(x, Q^2)}{\mathrm{d}\ln Q^2} = \frac{\alpha_{\mathrm{em}}(Q^2)}{2\pi}\, e_i^2\, P_{i/\gamma}(x) + \frac{\alpha_{\mathrm{s}}(Q^2)}{2\pi} \sum_j \int_x^1 \frac{\mathrm{d}z}{z}\, f_j\left(\frac{x}{z}\right)\, P_{i/j}(z)$$

so backwards evolution can find photon beam.



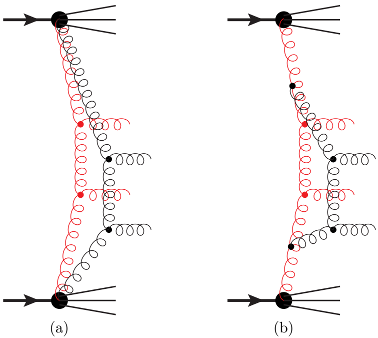
Have implemented combined direct + resolved for $\gamma\mathrm{p}$ and $\gamma\gamma$, for hard and soft processes; elastic and diffractive to come.

Also ep and $\mathrm{e}^+\mathrm{e}^-$ in quasi-real Equivalent Photon Approximation.

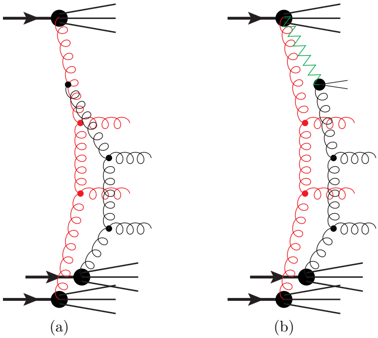
To come: Photon flux from hadrons (p and A), nuclear PDFs.

Heavy-ion collisions – 1

Angantyr (from 19th century Norse-style poem, like Fritiof.)

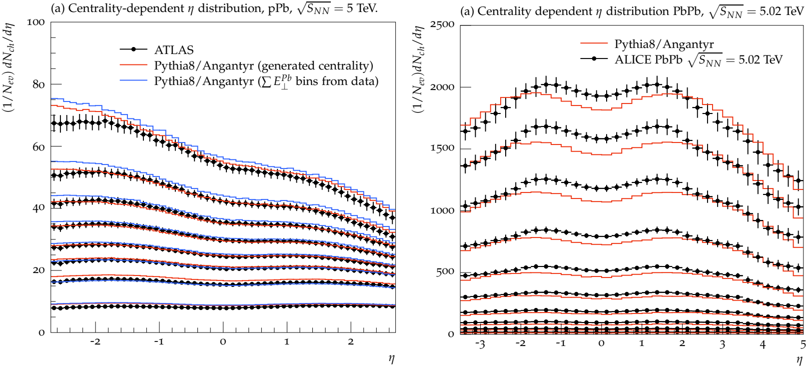


pp: MPIs naively attach multiple colour chains to remnants, but CR used/needed to reduce activity at large $|y|$.



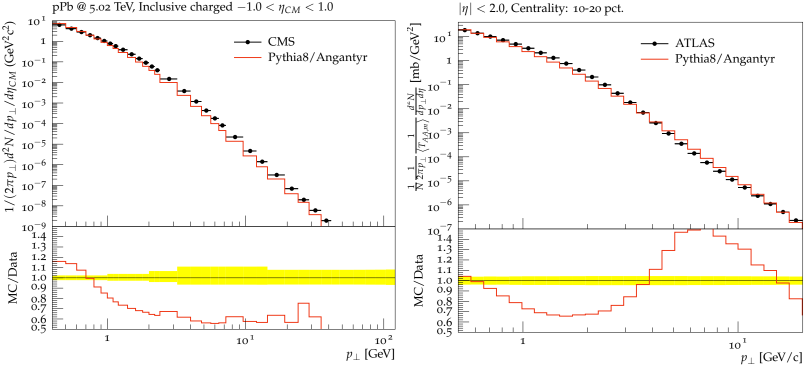
pd: similarly CR will reduce activity in p hemisphere; $\approx$ as one normal and one diffractive scattering.

Heavy-ion collisions – 2



- Glauber formalism for geometry and number of NN collisions.
- Good–Walker formalism for diffractive cross sections.
- Full MPI machinery for NN collisions, also diffractive.
- $\mathrm{d}x_{\mathrm{Pom}}/x_{\mathrm{Pom}}$ spectrum for energy taken from beam remnant.
- Energy–momentum–flavour conservation.

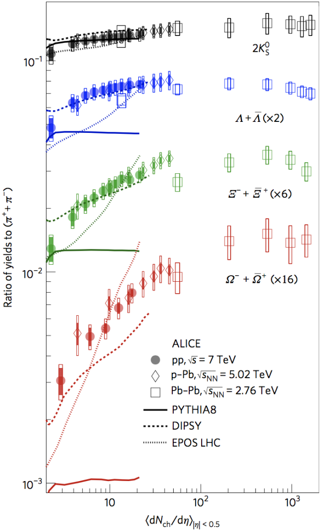
Heavy-ion collisions – 3

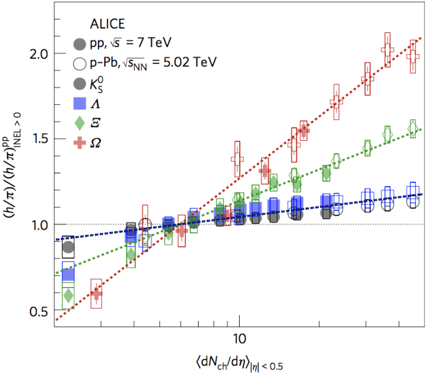


- Possibility to preselect one "trigger" event, e.g. Z^0 production.
- No quark-gluon plasma, for better or worse: which data can be explained without QGP?
- No explicit collective effects, for now.
- Under active evolution to improve agreement with data.

The ALICE revelation: goodbye jet universality!

Several unexpected collective effects at LHC pp, like ridge, and

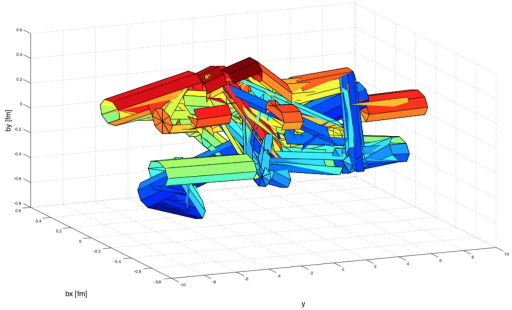




Signs of QGP in high-multiplicity pp collisions? If not, what else? A whole new game!

Ropes and shove

DIPSY: initial-state dipole evolution in transverse coordinates and longitudinal momenta.
Strong string overlap!



Ropes: combination of several overlapping strings into higher colour multiplets $\Rightarrow$ higher string tension favour strangeness, notably multistrange baryons.

Shove: overlap pushes strings apart $\Rightarrow$ ridge effects etc.

Currently not in PYTHIA, but ropes and shove will come.
Also other collective-event alternatives available or coming.

The PYTHIA collaboration

Current members: Christian Bierlich, Nishita Desai, Ilkka Helenius, Philip Ilten, Leif Lönnblad, Stephen Mrenna, Stefan Prestel, Christine Rasmussen, Torbjörn Sjöstrand, Peter Skands

. . . but many have other projects as their main research interest.

Significant code pieces contributed by ~ 30 more persons.

Comments and bug reports from > 100 persons.

Summary and outlook

- Core PYTHIA program is small and self-contained:
  $\sim$ 160k lines code, $\sim$ 20 MB gzipped tarball.
- Quick & easy to install, well documented and many examples,
  download from http://home.thep.lu.se/Pythia/
- Feasible to do simple standalone analyses, e.g. with jet finders.
- Possible to link to various external libraries.
- Used by **many** other programs, notably string fragmentation.
- Steady progress, e.g. heavy ions, γ beams, NLL showers.
- **Not** well structured for complete cosmic ray shower.