Soft Physics at the LHC: Insights and Challenges

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

Wien, 18 March 2014
Expect and observe high multiplicities at the LHC.

What are production mechanisms behind this?
The structure of an event

An event consists of many different physics steps, which have to be modelled by event generators:
Hadronization: the Lund string model

Field lines compressed to tubelike regions: \textbf{strings}!
Linear confinement:
\[ V(r) \approx \kappa r, \quad \kappa \approx 1 \text{ GeV/fm} \]
Lorentz invariant.

String breaking by tunneling:
\[ \mathcal{P} \propto \exp\left(-\pi m_{\perp q}^2 / \kappa \right) \]
with adjacent pairs forming mesons (and baryons).

- \( \frac{dn}{dy} \approx 1 \) flat, but with short-range (anti)correlations.
- Common Gaussian \( p_\perp \) spectrum.
- Varied hadron production, with suppression of heavy quarks, and short-range flavour (anti)correlations.

Generally supported by LEP, but some unresolved flavour issues.
A gluon = an energy–momentum-carrying kink on the string. Force ratio $g/q = 2$, cf. QCD $N_C/C_F = 9/4$, $\rightarrow 2$ for $N_C \rightarrow \infty$. No new parameters introduced for gluon jet hadronization!

Predicted for & discovered at PETRA, verified at LEP!
Hadronization: string vs. cluster

<table>
<thead>
<tr>
<th>program model</th>
<th>PYTHIA</th>
<th>HERWIG</th>
</tr>
</thead>
<tbody>
<tr>
<td>energy–momentum picture (space-time picture)</td>
<td>powerful</td>
<td>simple</td>
</tr>
<tr>
<td>parameters</td>
<td>predictive</td>
<td>unpredictive</td>
</tr>
<tr>
<td>flavour composition</td>
<td>few</td>
<td>many</td>
</tr>
<tr>
<td>parameters</td>
<td>messy</td>
<td>simple</td>
</tr>
<tr>
<td></td>
<td>in-between</td>
<td>in-between</td>
</tr>
<tr>
<td></td>
<td>many</td>
<td>few(er)</td>
</tr>
</tbody>
</table>

Overall, string model unsurpassed, > 30 years later (alas!)
Terminology for pp collisions

\[ \sigma_{\text{tot}} = \sigma_{\text{elastic}} + \sigma_{\text{single-diff}} + \sigma_{\text{double-diff}} + \cdots + \sigma_{\text{non-diff}} \]

Minimum Bias (MB) \( \approx \) “all events, with no bias from restricted trigger conditions”.

Experiment: \( \sigma_{\text{min-bias}} \approx \sigma_{\text{non-diff}} + \sigma_{\text{double-diffractive}} \approx \frac{2}{3} \times \sigma_{\text{tot}} \)

Theory (here): \( \sigma_{\text{min-bias}} \approx \sigma_{\text{non-diff}} \)

Underlying Event (UE): particles in an event with a hard process, but not originating from the process as such.

Pedestal effect: the UE contains more activity than a normal MB event does (even discarding diffractive events).

Trigger bias: a jet ”trigger” criterion \( E_{\perp j} > E_{\perp \text{min}} \) is more easily fulfilled in events with upwards-fluctuating UE activity, since the UE \( E_{\perp} \) in the jet cone counts towards the \( E_{\perp j} \).

Not enough!

Pileup: several pp collisions in a bunch crossing; separate issue.
The divergence of the QCD cross section

Cross section for $2 \rightarrow 2$ interactions is dominated by $t$-channel gluon exchange, so diverges like $\frac{d\sigma}{dp_{\perp}^2} \approx \frac{1}{p_{\perp}^4}$ for $p_{\perp} \rightarrow 0$.

Integrate QCD $2 \rightarrow 2$

$qq' \rightarrow qq'$
$q\bar{q} \rightarrow q'\bar{q}'$
$q\bar{q} \rightarrow gg$
$qg \rightarrow qg$
$gg \rightarrow gg$
$gg \rightarrow q\bar{q}$
(with CTEQ 5L PDF's)
Multiple Partonic Interactions (MPI)

Note that $\sigma_{\text{int}}(p_{\perp \text{min}})$, the number of $(2 \rightarrow 2 \text{ QCD})$ interactions above $p_{\perp \text{min}}$, involves integral over PDFs,

$$\sigma_{\text{int}}(p_{\perp \text{min}}) = \int\int\int_{p_{\perp \text{min}}} d{x}_1 \, d{x}_2 \, d{p}_{\perp}^2 \, f_1(x_1, p_{\perp}^2) \, f_2(x_2, p_{\perp}^2) \, \frac{d\hat{\sigma}}{d{p}_{\perp}^2}$$

with $\int d{x} \, f(x, p_{\perp}^2) = \infty$, i.e. infinitely many partons.

So half a solution to $\sigma_{\text{int}}(p_{\perp \text{min}}) > \sigma_{\text{tot}}$ is

**many interactions per event: MPI** (historically MI or MPPI)

\[
\sigma_{\text{tot}} = \sum_{n=0}^{\infty} \sigma_n
\]
\[
\sigma_{\text{int}} = \sum_{n=0}^{\infty} n \sigma_n
\]
\[
\sigma_{\text{int}} > \sigma_{\text{tot}} \iff \langle n \rangle > 1
\]
Other half of solution is that perturbative QCD is not valid at small $p_\perp$ since $q, g$ are not asymptotic states (confinement!).

Naively breakdown at

$$p_{\perp \text{min}} \sim \frac{\hbar}{r_p} \approx \frac{0.2 \text{ GeV} \cdot \text{fm}}{0.7 \text{ fm}} \approx 0.3 \text{ GeV} \approx \Lambda_{\text{QCD}}$$

... but better replace $r_p$ by (unknown) colour screening length $d$ in hadron:

\[ \lambda \sim \frac{1}{p_\perp} \]
Regularization of low-$p_{\perp}$ divergence

so need **nonperturbative regularization for** $p_{\perp} \to 0$, e.g.

$$\frac{d\hat{\sigma}}{dp_{\perp}^2} \propto \frac{\alpha_s^2(p_{\perp}^2)}{p_{\perp}^4} \rightarrow \frac{\alpha_s^2(p_{\perp}^2)}{p_{\perp}^4} \frac{\theta(p_{\perp} - p_{\perp\min})}{p_{\perp}} \quad \text{(simpler)}$$

or $$\rightarrow \frac{\alpha_s^2(p_{\perp0}^2 + p_{\perp}^2)}{(p_{\perp0}^2 + p_{\perp}^2)^2} \quad \text{(more physical)}$$

where $p_{\perp\min}$ or $p_{\perp0}$ are free parameters, empirically of order **2 GeV**.

Typically 2 – 3 interactions/event at the Tevatron, 4 – 5 at the LHC, but may be more in “interesting” high-$p_{\perp}$ ones.
Energy dependence of $p_{\perp \text{min}}$ and $p_{\perp 0}$

HERA PDF data, with steeper small-$x$ rise, introduced need for energy-dependent $p_{\perp \text{min}}$ and $p_{\perp 0}$:

- Larger collision energy $\Rightarrow$ probe parton ($\approx$ gluon) density at smaller $x$
- $\Rightarrow$ smaller colour screening length $d$
- $\Rightarrow$ larger $p_{\perp \text{min}}/p_{\perp 0}$
- $\Rightarrow$ dampened multiplicity rise

$$p_{\perp 0} \propto s^\epsilon, \epsilon \approx 0.08$$
in DL tune to $\sigma_{\text{tot}}$
Indirect evidence for MPI – 1

without MPI:

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.

(TS & M. van Zijl, 1987)

FIG. 4. Forward-backward multiplicity correlation at 540 GeV, UA5 results (Ref. 33) vs simple models; the latter models with notation as in Fig. 3.
Indirect evidence for MPI – 2

with MPI included:

(TS & M. van Zijl, 1987)
Indirect evidence for MPI – 3

At the LHC:
Need fine-tuning of MPI, but without it way off
All modern general-purpose generators are built on MPI concepts. but details differ, both physics and technology, e.g.

- a single regularized hard component
  or separate hard + soft components
- multiparton PDFs
- MPIs generated ordered in $p_{\perp}$ or not
- MPIs before ISR+FSR, after ISR+FSR, or interleaved
- energy/momentum/flavour conservation
- impact-parameter profile
- energy dependence, e.g. of $p_{\perp 0}$
- beam-remnant structure
- colour connection & reconnection strategies
- other collective effects: ropes, QGP, flow, BE

Will concentrate on Pythia 8 in the following.
Basic generation of MPI

- Differential probability for interaction at $p_\perp$ is
  \[
  \frac{dP}{dp_\perp} = \frac{1}{\sigma_{nd}} \frac{d\sigma}{dp_\perp}
  \]
  considering nondiffractive events only.

- Apply to sequence of decreasing $p_\perp^1 > p_\perp^2 > p_\perp^3 > \ldots$
  \[
  \mathcal{P}(p_\perp = p_\perp^i) = \frac{1}{\sigma_{nd}} \frac{d\sigma}{dp_\perp} \exp \left[ - \int_{p_\perp}^{p_\perp^{(i-1)}} \frac{1}{\sigma_{nd}} \frac{d\sigma}{dp_\perp'} dp_\perp' \right]
  \]
  starting from $p_{\perp max} = E_{cm}/2$, and including “Sudakov” factor (cf. radioactive decay).

- Use rescaled PDF’s taking into account already used momentum and flavours.

- Require $\geq 1$ interaction in an event or else pass through without anything happening.
Interleaved evolution

- Transverse-momentum-ordered parton showers for ISR and FSR.
- MPI also ordered in $p_\perp$.

$\Rightarrow$ Allows interleaved evolution for ISR, FSR and MPI:

$$\frac{dP}{dp_\perp} = \left( \frac{dP_{\text{MPI}}}{dp_\perp} + \sum \frac{dP_{\text{ISR}}}{dp_\perp} + \sum \frac{dP_{\text{FSR}}}{dp_\perp} \right)$$

$$\times \exp \left( -\int_{p_\perp}^{p_{\perp \text{max}}} \left( \frac{dP_{\text{MPI}}}{dp'_\perp} + \sum \frac{dP_{\text{ISR}}}{dp'_\perp} + \sum \frac{dP_{\text{FSR}}}{dp'_\perp} \right) \, dp'_\perp \right)$$

Ordered in decreasing $p_\perp$ using “Sudakov” trick.

Corresponds to increasing “resolution”:
smaller $p_\perp$ fill in details of basic picture set at larger $p_\perp$.

- Start from fixed hard interaction $\Rightarrow$ underlying event
- No separate hard interaction $\Rightarrow$ minbias events
- Possible to choose two hard interactions, e.g. $W^- W^-$
Often assume that MPI = ... but should also include the same order in $\alpha_s$, similar propagators, but

- one PDF weight less $\Rightarrow$ smaller $\sigma$
- one jet less $\Rightarrow$ QCD radiation background $2 \rightarrow 3$ larger than $2 \rightarrow 4$

$\Rightarrow$ will be tough to find direct evidence.

Rescattering grows with number of “previous” scatterings:

<table>
<thead>
<tr>
<th></th>
<th>Tevatron</th>
<th></th>
<th>LHC</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Min Bias</td>
<td>QCD Jets</td>
<td>Min Bias</td>
<td>QCD Jets</td>
</tr>
<tr>
<td>Normal scattering</td>
<td>2.81</td>
<td>5.09</td>
<td>5.19</td>
<td>12.19</td>
</tr>
<tr>
<td>Single rescatterings</td>
<td>0.41</td>
<td>1.32</td>
<td>1.03</td>
<td>4.10</td>
</tr>
<tr>
<td>Double rescatterings</td>
<td>0.01</td>
<td>0.04</td>
<td>0.03</td>
<td>0.15</td>
</tr>
</tbody>
</table>
Events are distributed in impact parameter $b$

$$\mathcal{O}(b) = \int d^3x \, dt \, \rho^{\text{boosted}}_{1, \text{matter}}(x, t) \rho^{\text{boosted}}_{2, \text{matter}}(x, t)$$

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

- central collisions more active $\Rightarrow P_n$ broader than Poissonian
- peripheral passages normally give no collisions $\Rightarrow$ finite $\sigma_{\text{tot}}$
An $x$-dependent proton size

Reasonable to assume that low-$x$ partons are more spread out:

$$\rho(r, x) \propto \frac{1}{a^3(x)} \exp \left( - \frac{r^2}{a^2(x)} \right) \quad \text{with} \quad a(x) = a_0 \left( 1 + \alpha_1 \ln \frac{1}{x} \right)$$

$a_1 \approx 0.15$ tuned to rise of $\sigma_{ND}$
$a_0$ tuned to value of $\sigma_{ND}$, given PDF, $p_{\perp 0}$, ...

Consequence: collisions at large $x$ will have to happen at small $b$, and hence further large-to-medium-$x$ MPIs are enhanced.

$a_1 > 0$ not favoured by tunes so far!
Direct observation of MPI – 1

Basic idea for direct observation goes back to AFS (1987): Order 4 jets $p_{\perp 1} > p_{\perp 2} > p_{\perp 3} > p_{\perp 4}$ and define $\varphi$ as angle between $p_{\perp 1} \pm p_{\perp 2}$ and $p_{\perp 3} \pm p_{\perp 4}$ for AFS/CDF

Later extended to $\gamma + 3$ jet, $W + 2$ jet.
Direct observation of MPI – 2

\[ \sigma_{a+b} = \frac{1}{1 + \delta_{ab}} \frac{\sigma_a \sigma_b}{\sigma_{\text{eff}}} \]

with Kronecker \( \delta \) for identical processes.

\[ f = \frac{\sigma_{nd}}{\sigma_{\text{eff}}} > 1 \]

from impact parameter. Beware of experimental issues!
Interpretation not as trivial as it seems!
(\( > 2 \) MPI, \( x \) dependence, \ldots)
Events with hard scale (jet, $W/Z$) have more underlying activity! Events with $n$ interactions have $n$ chances that one of them is hard, so “trigger bias”: hard scale $\Rightarrow$ central collision $\Rightarrow$ more interactions $\Rightarrow$ larger underlying activity.

Studied in particular by Rick Field, with CDF/CMS data:

- Define the MAX and MIN “transverse” regions on an event-by-event basis with MAX (MIN) having the largest (smallest) density.
"Transverse" Charged Particle Density: $dN/d\eta d\phi$

- **RDF Preliminary**
- **Corrected Data**

**Charged Particles (PT>0.5 GeV/c)**
- 1.96 TeV
- 300 GeV
- 900 GeV
- 7 TeV
- PYTHIA Tune Z1

**PTmax (GeV/c)**
- 0.0
- 0.4
- 0.8
- 1.2

**Torbjörn Sjöstrand**
**Colour Reconnection**
slide 25/44
One Feynman graph can correspond to several allowed colour flows, corresponding to different string drawings, e.g. for $qg \rightarrow qg$:

while other $qg \rightarrow qg$ graphs only admit one colour flow:

Interference term suppressed $\propto \frac{1}{(N_C^2 - 1)}$, so split cross section in proportions given by $N_C \rightarrow \infty$ limit.
Also parton showers conveniently traced in $N_C \to \infty$ limit.

But hooking it all up at beam remnants relies on $N_C = 3$:

\[ q \to qg \quad g \to gq \]

\[ g \to gg \]

Sjöstrand & PS, JHEP 03(2004)053
The Density of particle production

multiplicities in nondiffractive events (8 TeV LHC)

strings crossing $y = 0$
primary hadrons in $|y| < 0.5$
charged particles in $|y| < 0.5$

String width $\sim$ hadronic width $\Rightarrow$ Overlap factor $\sim 10!$
Reconnection in B decays

Colour operators in B decay $\Rightarrow$ some $\eta_c$:

$B \rightarrow J/\psi \rightarrow \mu^+ \mu^-$ good way to find B mesons:
... soon confirmed by experiment

$g^* \rightarrow c\bar{c} \rightarrow J/\psi$ production mechanism in pp ("colour octet")
more complicated to test (at the time, later "confirmed")
Interconnection at LEP 2

\[ e^+ e^- \rightarrow W^+ W^- \rightarrow q_1 \bar{q}_2 q_3 \bar{q}_4 \] reconnection limits \( m_W \) precision!

- perturbative \( \langle \delta M_W \rangle \lesssim 5 \text{ MeV} \) : negligible!
  (killed by dampening from off-shell W propagators)
- nonperturbative \( \langle \delta M_W \rangle \sim 40 \text{ MeV} \) : inconclusive.
  (but more extreme models from other authors ruled out)
- Bose-Einstein \( \langle \delta M_W \rangle \lesssim 100 \text{ MeV} \) : full effect ruled out.
  (but models with \( \sim 40 \text{ MeV} \) barely acceptable)

Colour rearrangement models for LEP 2

Colour rearrangement studied in several models, e.g.

**Scenario II: vortex lines.**
Analogy: type II superconductor.
Strings can reconnect only if central cores cross.

**Scenario I: elongated bags.**
Analogy: type I superconductor.
Reconnection proportional to space–time overlap.

In both cases favour reconnections that reduce total string length.

(schematic only; nothing to scale)
Reconnection at $Spp\bar{S}$


$\langle p_\perp \rangle(n_{\text{ch}})$ sensitive to colour flow.

long strings to remnants $\Rightarrow$ comparable $n_{\text{ch}}$/interaction

$\Rightarrow$ $\langle p_\perp \rangle(n_{\text{ch}}) \sim$ flat.

shorter extra strings for each consecutive interaction $\Rightarrow$ $\langle p_\perp \rangle(n_{\text{ch}})$ rising.

FIG. 27. Average transverse momentum of charged particles in $|\eta| < 2.5$ as a function of the multiplicity. UA1 data points (Ref. 49) at 900 GeV compared with the model for different assumptions about the nature of the subsequent (nonhardest) interactions. Dashed line, assuming $q\bar{q}$ scatterings only; dotted line, $gg$ scatterings with “maximal” string length; solid line $gg$ scatterings with “minimal” string length.
Reconnection at the LHC

\[ \langle p_{\perp} \rangle (n_{\text{ch}}) \text{ effect alive and kicking:} \]

Reconnection important also for other generators, e.g. Herwig++
Space–time models too complicated
⇒ simplified (in Pythia)

Common aspect: reduce string length
\[ \lambda = \sum \ln \left( \frac{m_{ij}^2}{m_0^2} \right) \sim \text{multiplicity} \]

Ingelman, Rathsman: reduce \( \sum m_{ij}^2 \);

Generalized Area Law

In total 12 scenarios in Pythia 6, mainly annealing:
- \( P_{\text{reconnect}} = 1 - (1 - \chi)^{n_{\text{MPI}}} \) with \( \chi \) strength parameter.
- Random assignment by \( P_{\text{reconnect}} \) for each string piece.
- Choose new combinations that reduce \( \lambda \) (with restrictions).

Pythia 8 still only primitive:
let each MPI either form a separate system, or attach its partons to a higher-\( p_\perp \) MPI where it gives minimal \( \lambda \) increase.

Much room for improvement; work ongoing.
MC: close to pole mass, in the sense of Breit–Wigner mass peak.
\[ c \tau \approx 0.1 \text{ fm} < r_p. \]

At the Tevatron: \[ m_t = 173.20 \pm 0.51 \pm 0.71 \text{ GeV} = \text{PMAS}(6,1) \]
At the LHC (CMS): \[ m_t = 173.54 \pm 0.33 \pm 0.96 \text{ GeV} \]

Now severely limited by colour reconnection uncertainty
Studies for the Tevatron.

Green bands: old virtuality-ordered showers.

Blue bands: new $p_\perp$-ordered showers.

In total $\pm 1.0$ GeV, whereof $\pm 0.7$ GeV perturbative, and $\pm 0.5$ GeV nonperturbative.

(M. Sandhoff and P. Z. Skands, FERMILAB-CONF-05-518-T;)
First top mass measurement binned in kinematic observables.

- Additional validation for the top mass measurements.
- With the current precision, no mis-modelling effect due to
  - color reconnection, ISR/FSR, b-quark kinematics, difference between pole or MS~ masses.

E. Yazgan
(Moriond 2013)
Top mass shift in Pythia 8

- Semileptonic top decay.
- Find jets with anti-$k_{\perp}$, $R = 0.5$, $p_{\perp\text{min}} = 20 \text{ GeV}$.
- Request $n_{\text{jet}} \geq 4$.
- Find two jets closest to $m_W$.
- Kill if $|m_{12} - m_W| > 5 \text{ GeV}$.
- Find third jet closest to $m_t$.
- Kill if $|m_{123} - m_t| > 20 \text{ GeV}$.

$t/W$ decay after $\rightarrow$ before CR:
$\langle \delta n_{\text{ch}} \rangle = -0.26 \pm 0.09$
$\langle \delta m_t \rangle = +0.060 \pm 0.020 \text{ GeV}$

t/$W$ decay after $\rightarrow$ no CR:
$\langle \delta n_{\text{ch}} \rangle = +36.44 \pm 0.09$
$\langle \delta m_t \rangle = +0.149 \pm 0.020 \text{ GeV}$
Reconnection and collective flow

Transverse boosts ⇒
≈ collective particle velocity.
More common with reconnection.

A. Ortiz Velasquez et al.,
Ongoing work

P. Skands, J. Christiansen, (TS): colour wavefunction of incoming partons for $N_C = 3$ gives new connection model, as start for reconnection

S. Argyropoulos, TS: worst-case scenarios for top mass

C. Bierlich, G. Gustafson, L. Lönnblad: DIPSY contains space–time dipole picture, colour ropes in high-density regions, higher string tension gives more strangeness and baryons:

$$K \frac{d n}{d \eta} \quad \Lambda \frac{d n}{d \eta} \quad \Xi^− \frac{d n}{d \eta}$$

How can we understand the strong force 36

Gösta Gustafson
Lund University
Ingelman-Schlein: Pomeron as hadron with partonic content
Diffractive event = (Pomeron flux) × (IPp collision)

1) $\sigma_{SD}$ and $\sigma_{DD}$ taken from existing parametrization or set by user.
2) $f_{IP/\bar{P}}(x_{IP}, t) \Rightarrow$ diffractive mass spectrum, $p_\perp$ of proton out.
3) Smooth transition from simple model at low masses to IPp with full pp machinery: multiple interactions, parton showers, etc.
4) Choice between 5 Pomeron PDFs, incl. Min Bias Rockefeller.
5) Free parameter $\sigma_{IPp}$ needed to fix $\langle n_{\text{interactions}} \rangle = \sigma_{\text{jet}}/\sigma_{IPp}$. 

Used e.g. in
POMPYT
POMWIG
PHOJET
Diffraction data

\[ \Delta \eta_F \]

as a function of \( \Delta \eta_F \)

- non-diffractive events dominate at small gaps
- diffractive plateau observed for large gaps

CMS Preliminary, \( \sqrt{s} = 7 \text{ TeV}, L = 20.3 \, \mu b^{-1} \)

- increasing particle threshold requirement results in more ND events with large gaps; confirms that inclusive events are dominated by low \( p_T \) production

ATLAS Coll., EPJ C72 (2012) 1926

PYTHIA8 models provide reasonable description

(C. Gwenlan, EPSHEP 2013)
Summary and Outlook

- Lund string still best hadronization description?
- MPI mechanism well established, with $b$ dependence.
- Low-$p_\perp$ regularization by colour screening?
  Energy dependence related to low-$x$ PDFs and $\sigma_{\text{tot}}$ rise?
- Rescattering and $x$-dependent $b$ profile marginal?
- Reconnection well established, from $B \rightarrow J/\psi$ to $\langle p_\perp \rangle (n_{\text{ch}})$.
- No understanding of detailed mechanism!
  One step: better space–time picture & colour algebra?
- Missed chance for clean tests at LEP 2. TLEP ideal?
- Multitude of algorithms for Pythia 6 $\Rightarrow$ uncertainty band.
  More algorithms on the way!
- Predict (possibility of) significant effects on $m_t$.
- High string density may preclude any definitive answers?
  Or breakthrough from new precision differential data?
Let’s aim for more than a flying circus . . .