Introduction to Event Generators 4

Torbjörn Sjöstrand

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

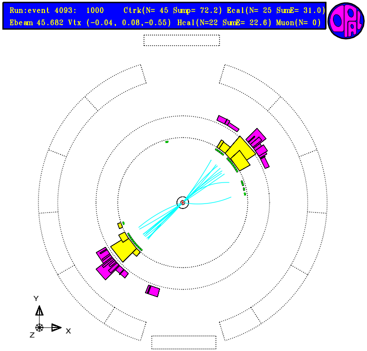
CTEQ/MCnet School, DESY, 12 July 2016

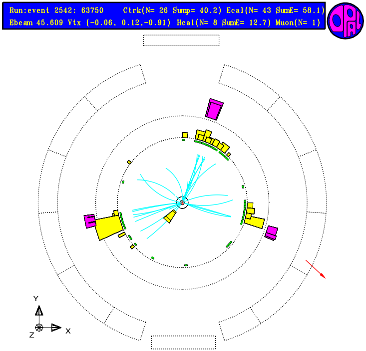
Hadronization

Hadronization/confinement is nonperturbative $\Rightarrow$ only models.

Main contenders: **string** and **cluster** fragmentation.

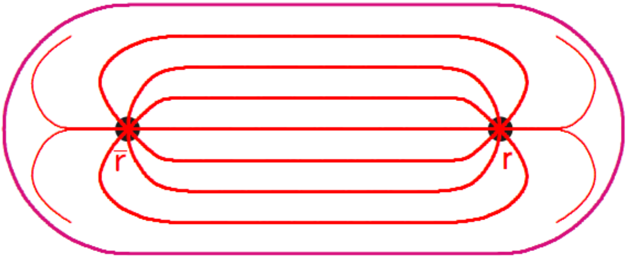
Begin with $\mathrm{e}^+\mathrm{e}^- \to \gamma^*/\mathrm{Z}^0 \to \mathrm{q}\overline{\mathrm{q}}$ and $\mathrm{e}^+\mathrm{e}^- \to \gamma^*/\mathrm{Z}^0 \to \mathrm{q}\overline{\mathrm{q}}\mathrm{g}$:





The QCD potential – 1

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 $\overline{\mathrm{q}}$:

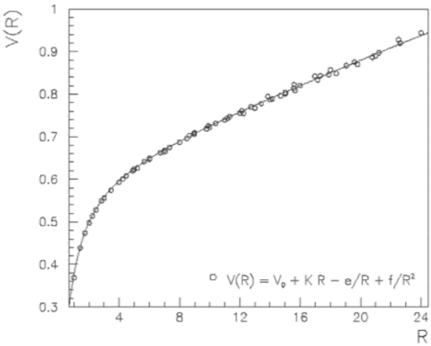
$F(r) \approx \mathrm{const} = \kappa \quad \Longleftrightarrow \quad V(r) \approx \kappa r$

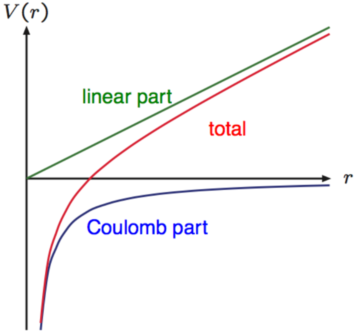
$\kappa \approx 1\ \mathrm{GeV/fm}$ $\approx$ potential energy gain lifting a 16 ton truck.

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

The QCD potential – 2

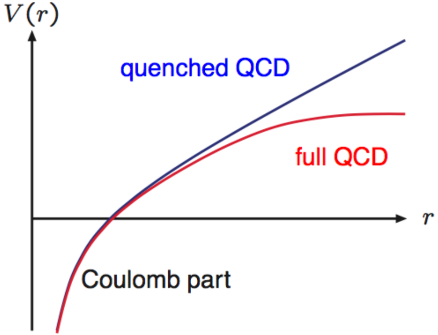
Linear confinement confirmed e.g. by lattice QCD calculation of gluon field between a static colour and anticolour charge pair:

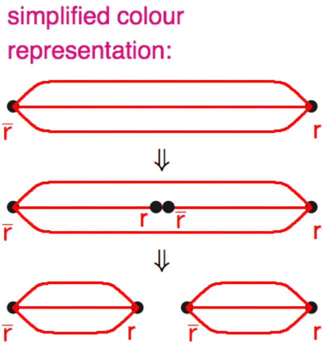




At short distances also Coulomb potential, important for internal structure of hadrons, but not for particle production (?).

Full QCD = gluonic field between charges ("quenched QCD")
plus virtual fluctuations $\mathrm{g} \to \mathrm{q\overline{q}}\,(\to \mathrm{g})$
$\Longrightarrow$ nonperturbative string breakings $\mathrm{gg}\ldots \to \mathrm{q\overline{q}}$





String motion

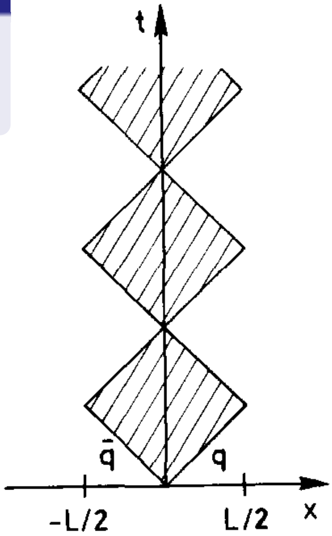
The Lund Model: starting point

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

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

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

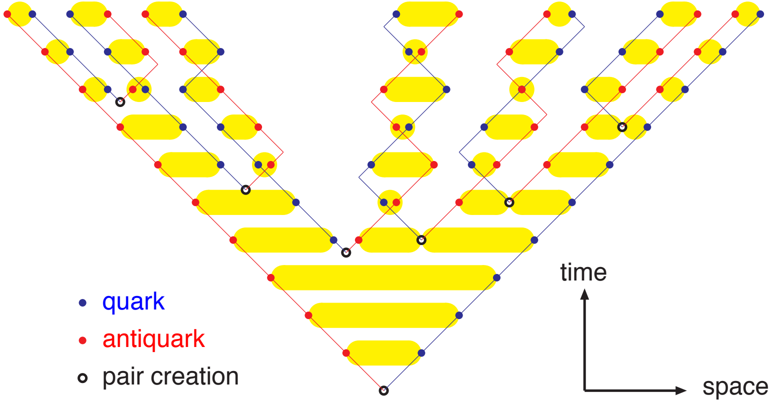
($c = 1$) for a $\mathrm{q}\overline{\mathrm{q}}$ pair flying apart along the $\pm z$ axis.
But signs relevant: the q moving in the $+z$ direction has $\mathrm{d}z/\mathrm{d}t = +1$ but $\mathrm{d}p_z/\mathrm{d}t = -\kappa$.



The Lund Model

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

Motion of quarks and antiquarks with intermediate string pieces:

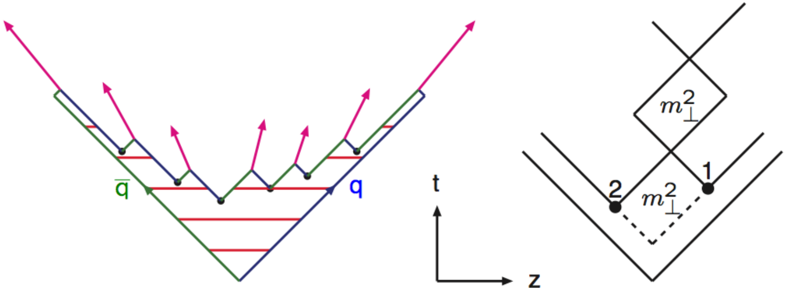


A q from one string break combines with a $\overline{q}$ from an adjacent one.

Gives simple but powerful picture of hadron production.

Where does the string break?

Fragmentation starts in the middle and spreads outwards:



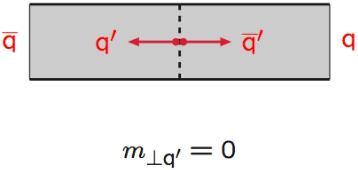
Corresponds to roughly same invariant time of all breaks,
$\tau^2 = t^2 - z^2 \sim$ constant,
with breaks separated by hadronic area $m_\perp^2 = m^2 + p_\perp^2$.

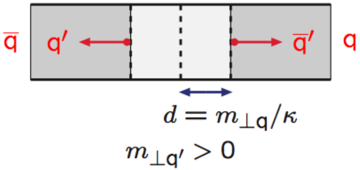
Hadrons at outskirts are more boosted.

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?



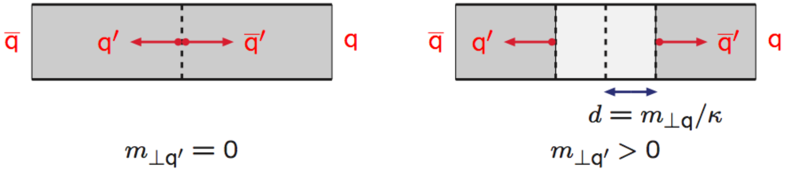


String breaking modelled by tunneling:

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

- 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.

How does the string break?



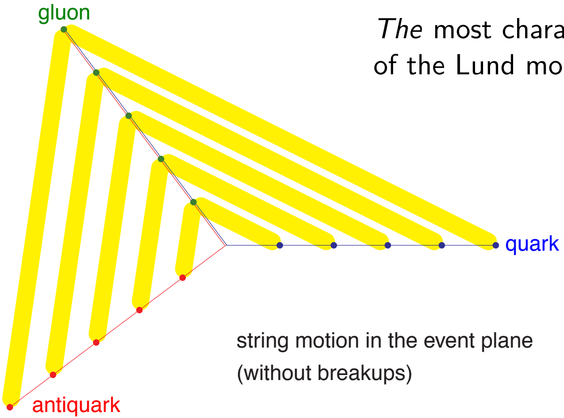
String breaking modelled by tunneling:

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

- 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 – 1



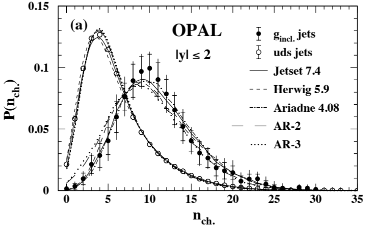
The most characteristic feature of the Lund model:

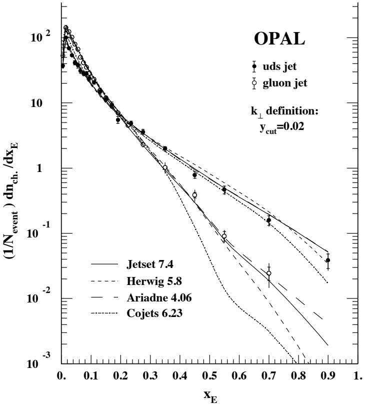
Gluon = kink on string

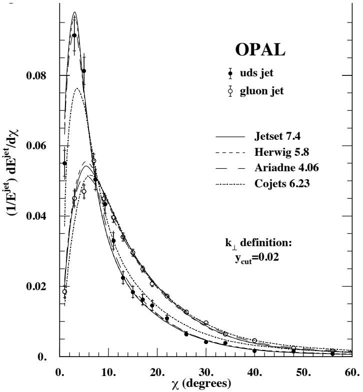
Force ratio gluon/ quark = 2,
cf. QCD $N_C/C_F = 9/4$, $\rightarrow 2$ for $N_C \rightarrow \infty$
No new parameters introduced for gluon jets!

The Lund gluon picture – 2

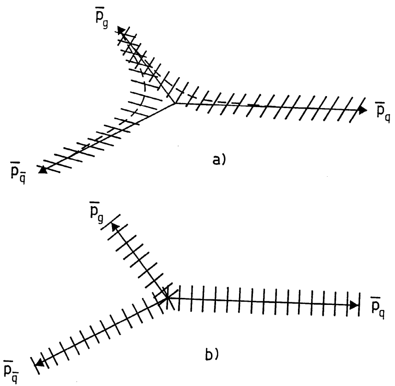
Energy sharing between two strings makes hadrons in gluon jets softer, more and broader in angle:

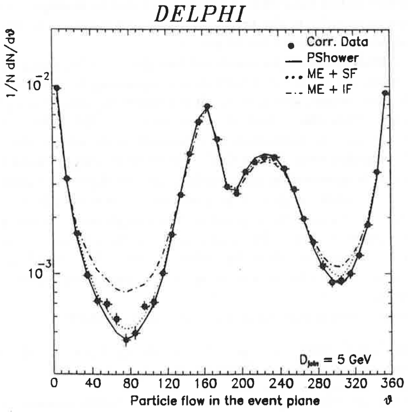






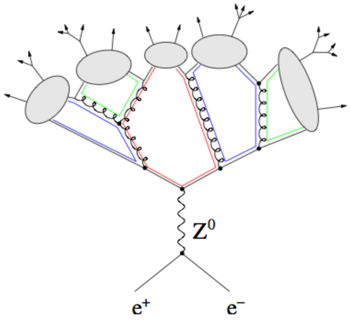
Particle flow in the $q\bar{q}g$ event plane depleted in q–$\bar{q}$ region owing to boost of string pieces in q–g and g–$\bar{q}$ regions:

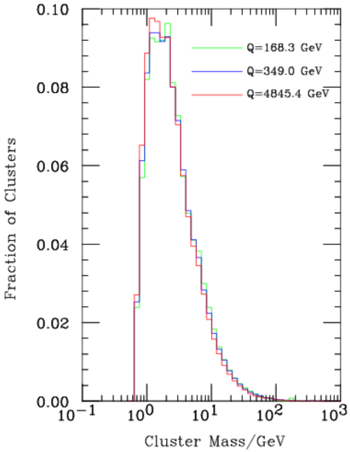




String fragmentation (SF) vs. independent fragmentation (IF), latter (nowadays) straw model of symmetric jet profile.

The Herwig Cluster Model

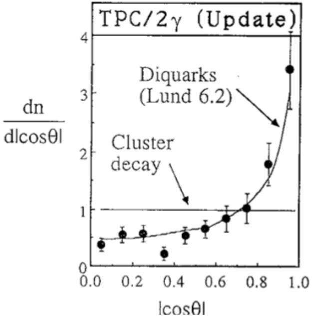




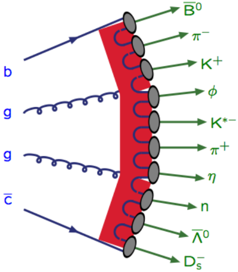
1. Introduce forced $g \to q\bar{q}$ branchings
2. Form colour singlet clusters
3. Clusters decay isotropically to 2 hadrons according to phase space weight $\sim (2s_1+1)(2s_2+1)(2p^*/m)$

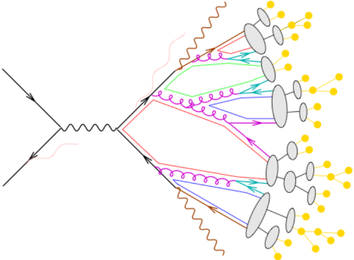
Cluster Model issues

1. Tail to very large-mass clusters (e.g. if no emission in shower); if large-mass cluster $\to$ 2 hadrons then incorrect hadron momentum spectrum, crazy four-jet events
   $\Longrightarrow$ split big cluster into 2 smaller along "string" direction; daughter-mass spectrum $\Rightarrow$ iterate if required;
   $\sim$ 15% of primary clusters are split,
   but give $\sim$ 50% of final hadrons
2. Isotropic baryon decay inside cluster
   $\Longrightarrow$ splittings $\mathrm{g} \to \mathrm{qq} + \overline{\mathrm{qq}}$
3. Too soft charm/bottom spectra
   $\Longrightarrow$ anisotropic leading-cluster decay
4. Charge correlations still problematic
   $\Longrightarrow$ all clusters anisotropic (?)
5. Sensitivity to particle content
   $\Longrightarrow$ only include complete multiplets



String vs. Cluster

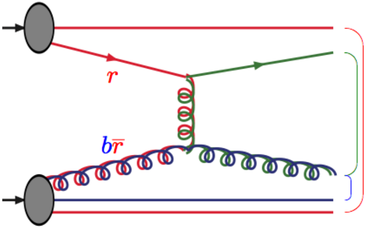




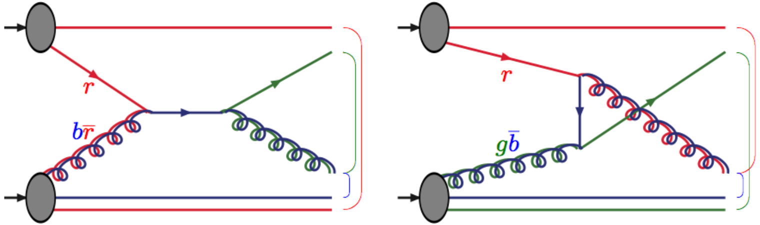
program	PYTHIA	Herwig
model	string	cluster
energy–momentum picture	powerful	simple
	predictive	unpredictive
parameters	few	many
flavour composition	messy	simple
	unpredictive	in-between
parameters	many	few

Colour flow in hard processes – 1

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:



Colour flow in hard processes – 2

so nontrivial mix of kinematics variables $(\hat{s}, \hat{t})$
and colour flow topologies I, II:

$$\begin{aligned} |\mathcal{A}(\hat{s},\hat{t})|^2 &= |\mathcal{A}_{\mathrm{I}}(\hat{s},\hat{t}) + \mathcal{A}_{\mathrm{II}}(\hat{s},\hat{t})|^2 \\ &= |\mathcal{A}_{\mathrm{I}}(\hat{s},\hat{t})|^2 + |\mathcal{A}_{\mathrm{II}}(\hat{s},\hat{t})|^2 + 2\,\mathcal{R}e\left(\mathcal{A}_{\mathrm{I}}(\hat{s},\hat{t})\mathcal{A}^*_{\mathrm{II}}(\hat{s},\hat{t})\right) \end{aligned}$$

with $\mathcal{R}e\left(\mathcal{A}_{\mathrm{I}}(\hat{s},\hat{t})\mathcal{A}^*_{\mathrm{II}}(\hat{s},\hat{t})\right) \neq 0$
$\Rightarrow$ indeterminate colour flow, while
• showers *should* know it (coherence),
• hadronization *must* know it (hadrons singlets).
Normal solution:

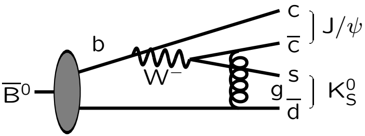
$$\frac{\mathrm{interference}}{\mathrm{total}} \propto \frac{1}{N_{\mathrm{C}}^2 - 1}$$

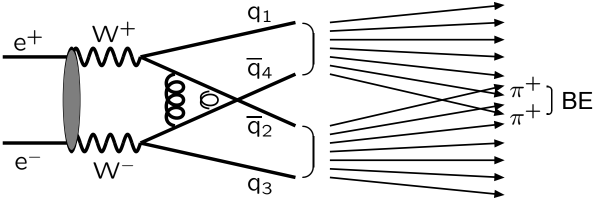
so split $\mathrm{I} : \mathrm{II}$ according to proportions in the $N_{\mathrm{C}} \to \infty$ limit, i.e.

$$\begin{aligned} |\mathcal{A}(\hat{s},\hat{t})|^2 &= |\mathcal{A}_{\mathrm{I}}(\hat{s},\hat{t})|^2_{\mathrm{mod}} + |\mathcal{A}_{\mathrm{II}}(\hat{s},\hat{t})|^2_{\mathrm{mod}} \\ |\mathcal{A}_{\mathrm{I(II)}}(\hat{s},\hat{t})|^2_{\mathrm{mod}} &= |\mathcal{A}_{\mathrm{I}}(\hat{s},\hat{t}) + \mathcal{A}_{\mathrm{II}}(\hat{s},\hat{t})|^2 \left(\frac{|\mathcal{A}_{\mathrm{I(II)}}(\hat{s},\hat{t})|^2}{|\mathcal{A}_{\mathrm{I}}(\hat{s},\hat{t})|^2 + |\mathcal{A}_{\mathrm{II}}(\hat{s},\hat{t})|^2} \right)_{N_{\mathrm{C}} \to \infty} \end{aligned}$$

Colour Reconnection Revisited

Colour rearrangement well established e.g. in B decay.





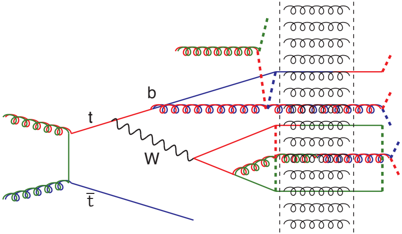
At LEP 2 search for effects in $\mathrm{e^+e^- \to W^+W^- \to q_1\overline{q}_2\, q_3\overline{q}_4}$:

- 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 2.8σ.
- Bose-Einstein $\langle \delta M_{\mathrm{W}} \rangle \lesssim 100$ MeV : full effect ruled out (while models with ~ 20 MeV barely acceptable).

A top mass puzzle

$$\left.\begin{array}{l}\Gamma_t \approx 1.5 \text{ GeV} \\ \Gamma_W \approx 2 \text{ GeV} \\ \Gamma_Z \approx 2.5 \text{ GeV}\end{array}\right\} \Rightarrow c\tau \approx 0.1 \text{ fm} :$$

p "pancakes" have passed,
MPI/ISR/FSR for $p_\perp \geq 2$ GeV,
inside hadronization colour fields.

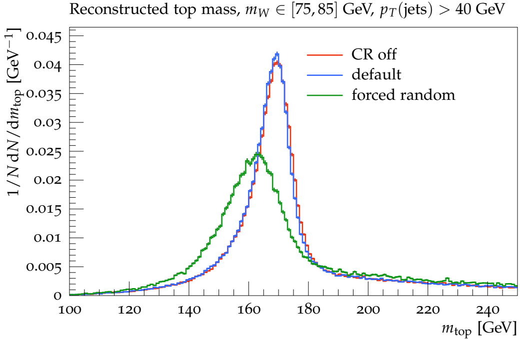


Experiment	m_{top} [GeV]	Error due to CR	Reference
World comb.	173.34±0.76	310 MeV (**40%**)	arXiv:1403.4427
CMS	172.22±0.73	150 MeV (**20%**)	CMS-PAS-TOP-14-001
D0	174.98±0.76	100 MeV (**13%**)	arXiv:1405.1756

(S. Argyropoulos)

1. Great job in reducing the errors.
2. CR is one of the dominant systematics.
3. Why is the CR uncertainty going down when there are
   - no advances in theoretical understanding, and
   - no measurements to constrain it?

Effects on top mass before tuning



Δm_{top} relative to no CR:

model	Δm_{top} [GeV]	Δm_{top} rescaled
default (late)	-0.415	$+0.209$
default early	$+0.381$	$+0.285$
forced random	-6.970	-6.508

Asymmetric spread:
$\Delta m_{top} < 0$ easy,
$\Delta m_{top} > 0$ difficult.

Parton showers already prefer minimal λ.

Main effect from jet broadening, some from jet–jet angles.

Effects on top mass after tuning

No publicly available measurements of UE in top events.

- Afterburner models tuned to ATLAS jet shapes in $\mathrm{t\bar{t}}$ events
  $\Rightarrow$ high CR strengths disfavoured.
- Early-decay models tuned to ATLAS minimum bias data
  $\Rightarrow$ maximal CR strengths required to (almost) match $\langle p_\perp \rangle (n_\mathrm{ch})$.

model	Δm_top rescaled
default (late)	$+0.239$
forced random	-0.524
swap	$+0.273$

Δm_top relative to no CR

Excluding most extreme (unrealistic) models

$\mathbf{m_{top}^{max} - m_{top}^{min} \approx 0.50}$ **GeV**

(in line with Sandhoff, Skands & Wicke)

New: $\Delta m_\mathrm{top} \approx 0$ in QCD-based model

Studies of top events could help constrain models:

- jet profiles and jet pull (skewness)
- underlying event

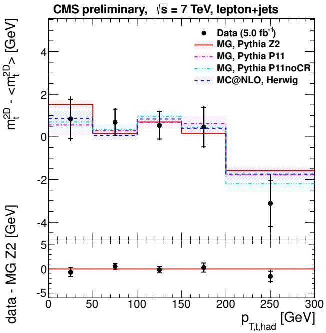
Probing reconnection through the top mass

Dependence of Top Mass on Event Kinematics

CMS-PAS-TOP-12-029

NEW

	Fig.	Observable
color recon.	1	$\Delta R_{q\bar{q}}$
	2	$\Delta\phi_{q\bar{q}}$
	3	$p_{T,t,had}$
	4	$\lvert\eta_{t,had}\rvert$
ISR/FSR	5	H_T
	6	$m_{t\bar{t}}$
	7	$p_{T,t\bar{t}}$
	8	Jet multiplicity
b-quark kin.	9	$p_{T,b,had}$
	10	$\lvert\eta_{b,had}\rvert$
	11	$\Delta R_{b\bar{b}}$
	12	$\Delta\phi_{b\bar{b}}$



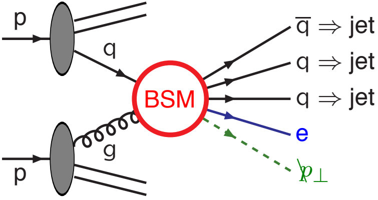
- *First top mass measurement binned in kinematic observables.*
- Additional validation for the top mass measurements.
- With the current precision, no mis-modelling effect due to
  - color reconnection, ISR/FSR, b-quark kinematics, difference between pole or MS~ masses.

E. Yazgan

(Moriond 2013)

1(

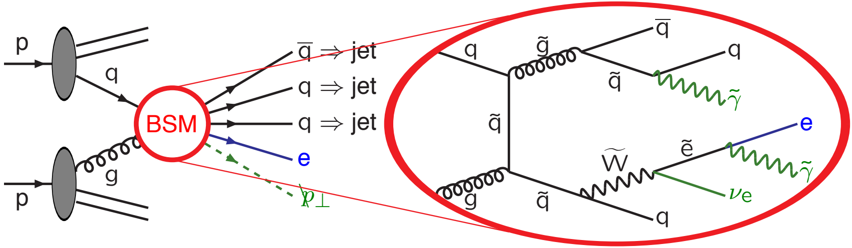
BSM particles usually short-lived, or weakly interacting (like DM). Then visible final state consists of hadrons, leptons and photons, just like ordinary processes.



As easy to model as SM processes.

BSM at the LHC

BSM particles usually short-lived, or weakly interacting (like DM). Then visible final state consists of hadrons, leptons and photons, just like ordinary processes.



As easy to model as SM processes.

Original structure hidden, but traces of it may be left in terms of invariant masses and angular distributions.

Discovery requires detailed understanding of rare signals and huge backgrounds.

BSM and QCD at the LHC

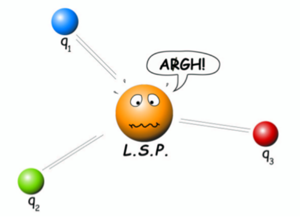
LHC is a QCD machine:

- hard processes initiated by quarks and gluons,
- final state almost always dominated by hadrons,
- underlying event by QCD mechanisms (showers, MPIs, ...),
- even in scenarios for physics Beyond the Standard Model (BSM) production of new coloured states often favoured (squarks, KK gluons, excited quarks, leptoquarks, ...).

In addition, BSM physics can raise "new", specific QCD aspects:

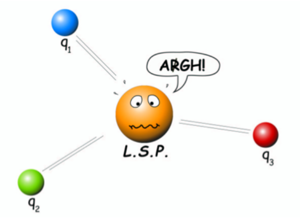
- new production mechanisms
- new parton-shower aspects
- new decay channels
- new hadronization phenomena
- new correlations with rest of the event

Examples of nontrivial BSM physics

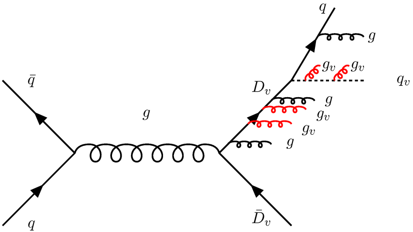


BNV ⇒ junction topology
⇒ special handling of
showers and hadronization

Examples of nontrivial BSM physics

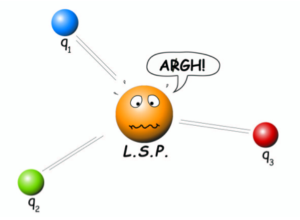


BNV $\Rightarrow$ junction topology
$\Rightarrow$ special handling of
showers and hadronization

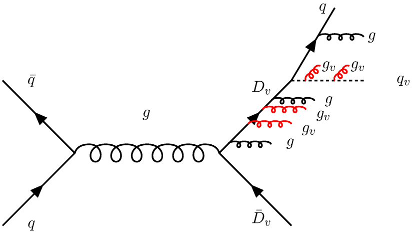


Hidden valleys:
showers potentially interleaved
with normal ones;
hadronization in hidden sector;
decays back to normal sector

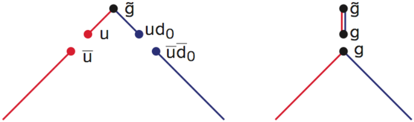
Examples of nontrivial BSM physics



BNV ⇒ junction topology
⇒ special handling of
showers and hadronization



Hidden valleys:
showers potentially interleaved
with normal ones;
hadronization in hidden sector;
decays back to normal sector

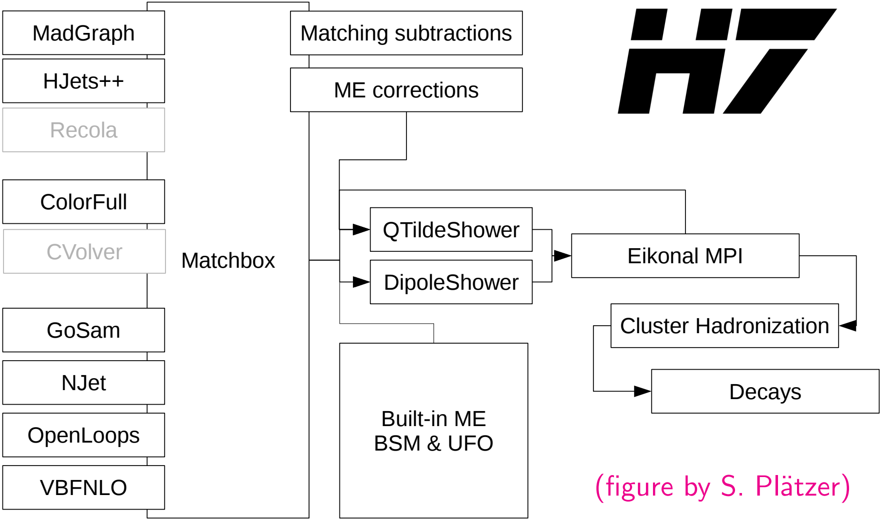


R-hadrons: long-lived $\tilde{g}$ or $\tilde{q}$;
new: hadronization of massive object "inside" the string

Herwig 7.0 news

- Herwig++ 3.0 ⇒ **Herwig 7.0** (December 2015).
  Concludes 16 years effort to replace Fortran Herwig 6.
- **NLO** matched to parton showers **default** for hard process.
  - Fully **automated**: no external codes to run, no intermediate event files.
  - Choice of **subtractive** (MC@NLO type) or **multiplicative** (PowHeg type) matching.

Matchbox in Herwig 7

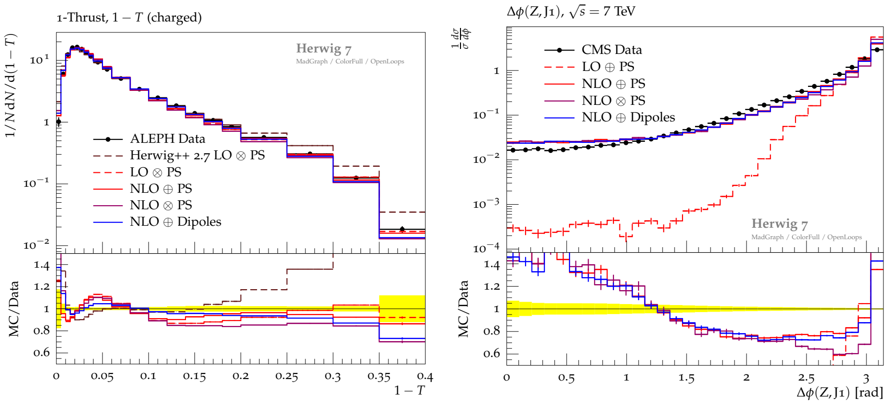


(figure by S. Plätzer)

script downloads & sets up external libraries (above + more)

Herwig 7.0 news

- Herwig++ 3.0 ⇒ **Herwig 7.0** (December 2015).
  Concludes 16 years effort to replace Fortran Herwig 6.
- **NLO** matched to parton showers **default** for hard process.
  - Fully **automated**: no external codes to run,
    no intermediate event files.
  - Choice of **subtractive** (MC@NLO type)
    or **multiplicative** (PowHeg type) matching.
- Two showers: angular ordered or **dipole**.
  Spin correlations and QED radiation in the former.
- Facilities for parton-shower uncertainties.
- New tunes, including MB/UE.
- Vastly improved documentation, usage and installation.
- Several parallelization options.



LO → NLO ⇒ major improvements in e^+e^- and pp alike.

Subtractive or multiplicative matching less important.

Ditto angular-ordered or dipole shower.

Future of Herwig 7

Herwig 7.1 later this year:

- **NLO multijet merging** (unitarized merging ideas).
- Loop-induced processes.
- Extended UFO-model support.
- Extended reweighting: weight vectors in HepMC files.
- Improved top decay in dipole shower.
- Interface to HEJ.
- Soft interactions and diffraction.

In the longer run:

- Code now 500k lines $\Rightarrow$ need for significant restructuring.
- Amplitude-based parton showers.

Sherpa 2.2 news and activities

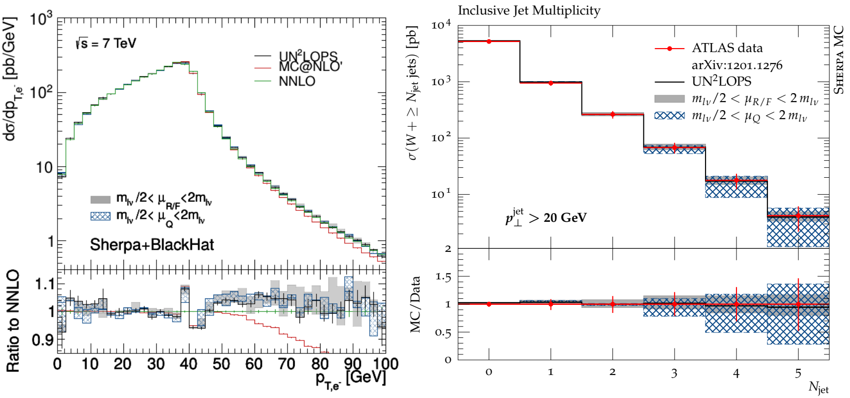
Recent news:

- DIRE shower (see lecture 2).
- UNNLOPS - first results on NNLO merging.

Sherpa NNLO QCD with parton showers

W production @ NNLO+PS with SHERPA +BLACKHAT

[Höche et al. arXiv:1507.05325]



↪ fully differential hadron-level NNLO+PS simulation

- inclusive (born-like) distribution NNLO accurate
- 0-jet bin NNLO, 1-jet bin NLO, 2-jet bin LO, ≥ 3-jets shower accuracy

↪ small corrections away from Born kinematics

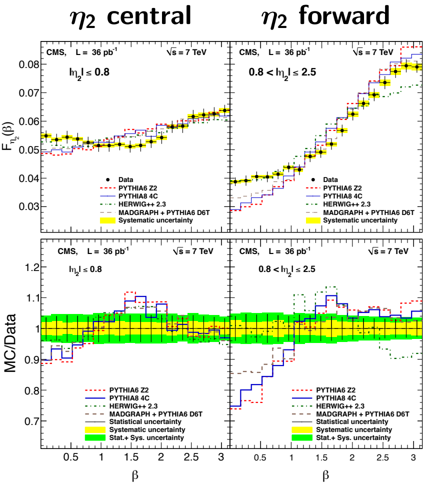
Sherpa 2.2 news and activities

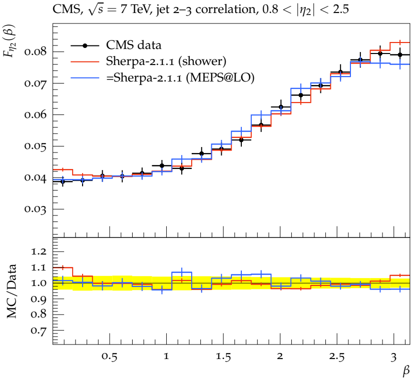
Recent news:

- DIRE shower (see lecture 2).
- UNNLOPS - first results on NNLO merging.
- On-the-fly scale variations of NLO ME + PS.
  ME observables through interpolating grids
  (ApplGrid, FastNLO, MCgrid, . . .).
- Electroweak NLO corrections, together with OpenLoops.
- Merging for loop-induced processes.

Sherpa QCD coherence test

Study events with two hard and one further softer third jets.
Angular distribution of third around second probes colour coherence:





PYTHIA/Herwig does not quite describe data, whereas Sherpa fares much better.

Sherpa 2.2 news and activities

Recent news:

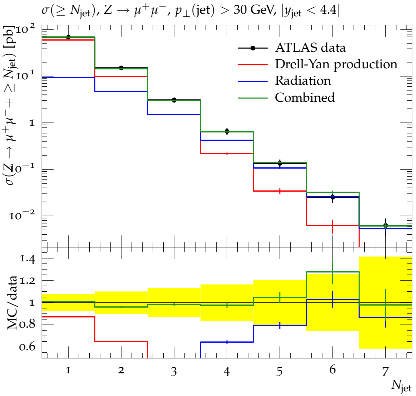
- DIRE shower (see lecture 2).
- UNNLOPS - first results on NNLO merging.
- On-the-fly scale variations of NLO ME + PS.
  ME observables through interpolating grids
  (ApplGrid, FastNLO, MCgrid, . . .).
- Electroweak NLO corrections, together with OpenLoops.
- Merging for loop-induced processes.

Ongoing work and plans:

- Full NNLO QCD + NLO EW (for $2 \rightarrow 1$, $2 \rightarrow 2$).
- Higher-order shower
  (one-loop splitting functions, sub-leading colour).
- Automated N-jettiness slicing.

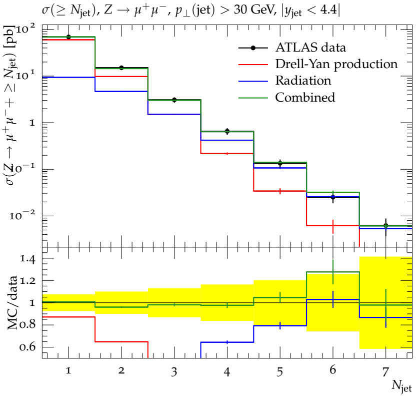
PYTHIA 8.2 news

- New match&merge schemes (now 8) and options.
- Weak showers: $q \to qZ^0$, $q \to q'W^{\pm}$ (also merged).



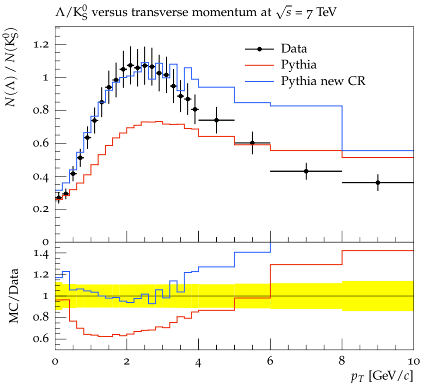
PYTHIA 8.2 news

- New match&merge schemes (now 8) and options.
- Weak showers: $q \to qZ^0$, $q \to q'W^{\pm}$ (also merged).
- Allow reweighting of rare shower branchings.
- Automated parton-shower uncertainty bands.
- Extended interface for external shower plugins, like VINCIA and DIRE.
- Complete LHEF v3 support.
- Can run Madgraph5_aMC@NLO and POWHEG BOX from within PYTHIA.
- Complete Python interface.



PYTHIA 8.2 news

- Many new colour reconnection models.
- Double onium production.
- New model for hard diffraction.
- Several new tunes; Monash new default.



Ongoing work and plans:

- $\gamma\gamma$, γp and ep.
- Total, elastic and diffractive cross sections.
- Improved showers (including VINCIA and DIRE).
- New approaches to hadronization, in response to pp/pA/AA similarities.

Summary and Outlook

- Increased ME calculational capability: legs and loops.
- Match and merge approaches still steadily developing.
- Continued/increased interest in parton shower development, with each generator offering several options.
- Many challenges remaining in soft physics, pA, AA: diffraction, colour reconnection, collective effects, ...
- **Generators have gone from fringe activity for a few to a mainstream part of phenomenology research.**

Summary and Outlook

- Increased ME calculational capability: legs and loops.
- Match and merge approaches still steadily developing.
- Continued/increased interest in parton shower development, with each generator offering several options.
- Many challenges remaining in soft physics, pA, AA: diffraction, colour reconnection, collective effects, ...
- **Generators have gone from fringe activity for a few to a mainstream part of phenomenology research.**