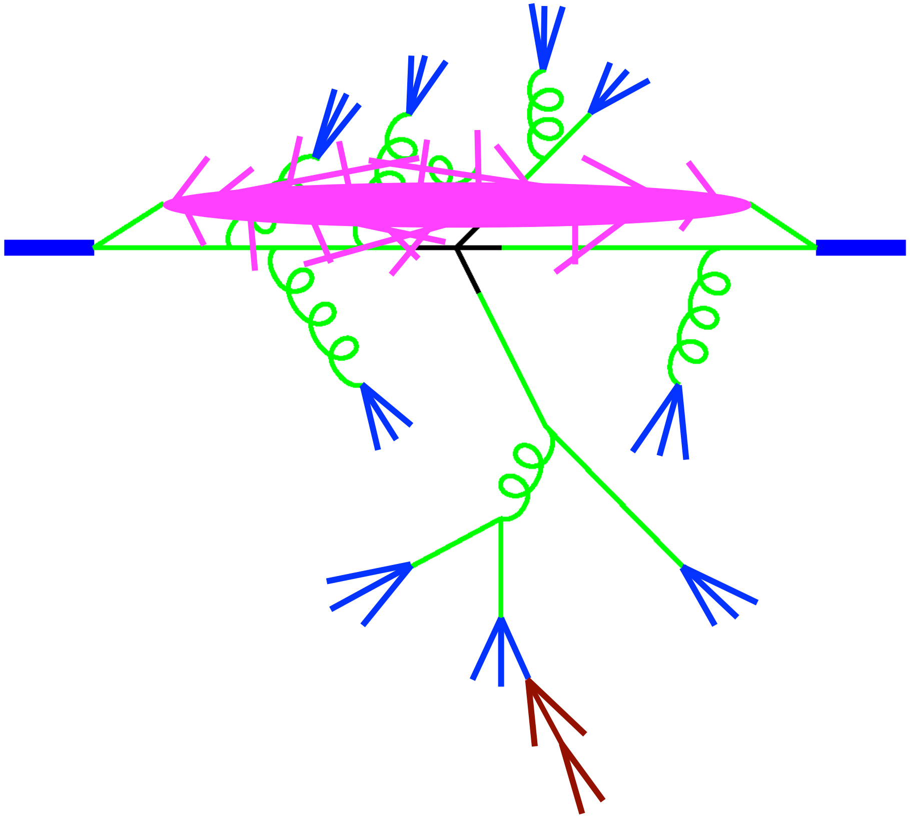
Monte Carlo Event Generators

Mike Seymour
University of Manchester

CTEQ school
University of Pittsburgh,
7–17 July 2015

Structure of LHC Events

1. Hard process
2. Parton shower
3. Hadronization
4. Underlying event
5. Unstable particle decays



Intro to Monte Carlo Event Generators

1. Monte Carlo technique / hard process
2. Parton showers
3. Hadronization
4. Underlying Event / Soft Inclusive Models

Hadronization: Introduction

Partons are not physical particles: they cannot freely propagate.

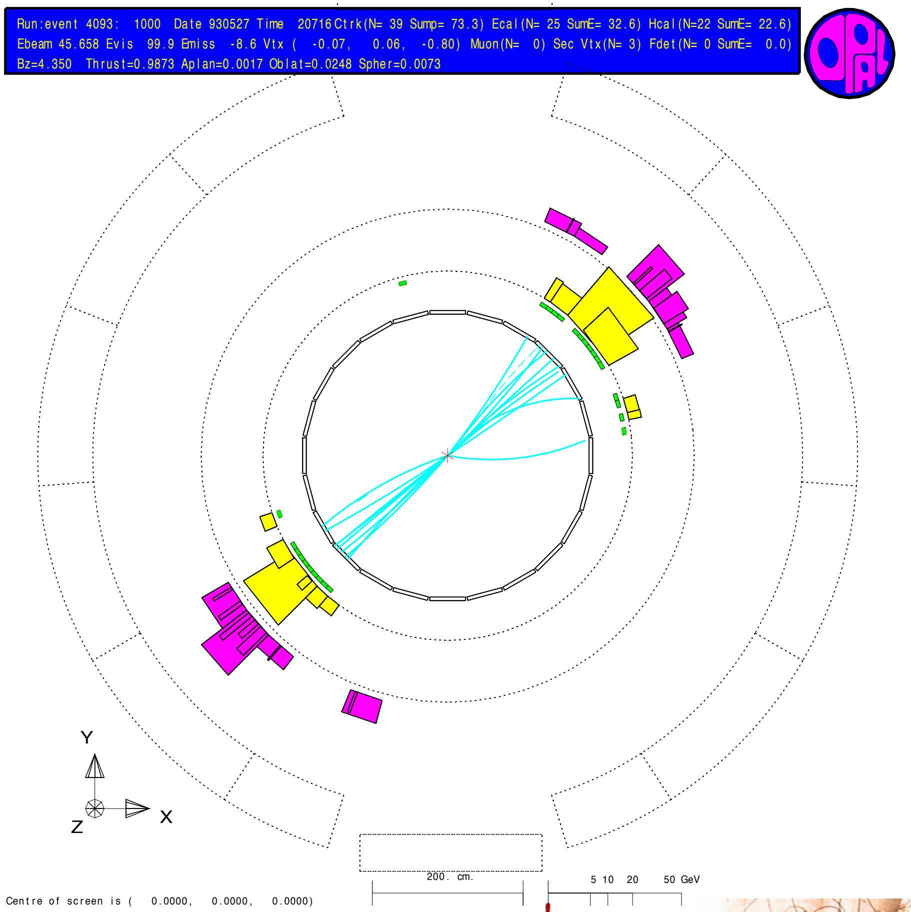
Hadrons are.

Need a model of partons' confinement into hadrons: hadronization.

1. Phenomenological models.
2. Confinement.
3. The string model.
4. Preconfinement.
5. The cluster model.
6. Secondary decays.
7. Underlying event models.

Jet production in $e^+e^- \rightarrow$ hadrons

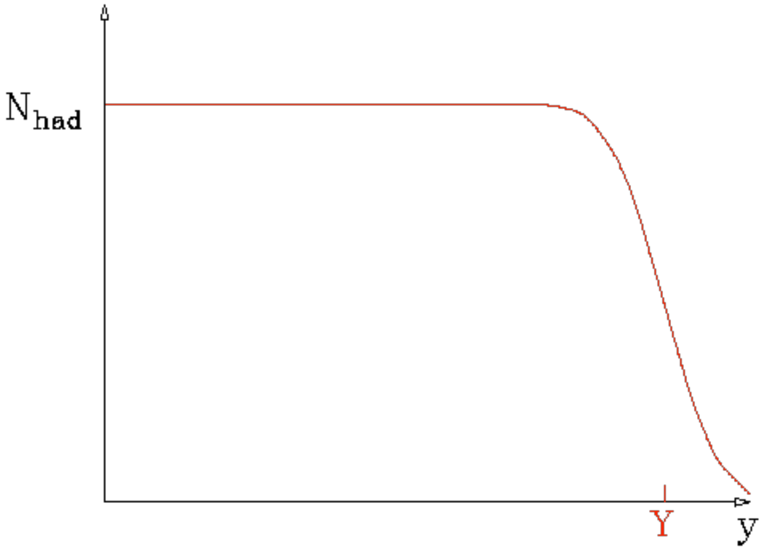
- Most e^+e^- events consist of two back-to-back jets

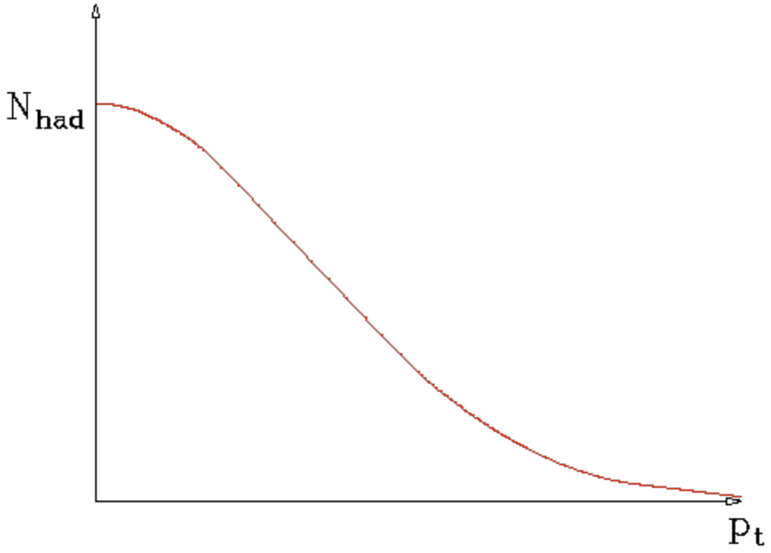


Phenomenological Models

Experimentally, $e^+e^- \rightarrow$ two jets:

Flat rapidity plateau and limited p_t, $\rho(p_t^2) \sim e^{-p_t^2/2p_0^2}$





Estimate of Hadronization Effects

Using this model, can estimate hadronization correction to perturbative quantities.

Jet energy and momentum:

$$E = \int_0^Y dy\, d^2p_t\, \rho(p_t^2)\, p_t\, \cosh y = \lambda \sinh Y$$

$$P = \int_0^Y dy\, d^2p_t\, \rho(p_t^2)\, p_t\, \sinh y = \lambda(\cosh Y - 1) \sim E - \lambda,$$

with $\lambda = \int d^2p_t\, \rho(p_t^2)\, p_t,$ mean transverse momentum.

Estimate from Fermi motion $\lambda \sim 1/R_{had} \sim m_{had}$.

Jet acquires non-perturbative mass: $M^2 = E^2 - P^2 \sim 2\lambda E$

Large: ~ 10 GeV for 100 GeV jets.

Independent Fragmentation Model ("Feynman—Field")

Direct implementation of the above.

Longitudinal momentum distribution = arbitrary fragmentation function: parameterization of data.

Transverse momentum distribution = Gaussian.

Recursively apply $q \to q' + \text{had}$.

Hook up remaining soft q and $\bar{q}$.

Strongly frame dependent.

No obvious relation with perturbative emission.

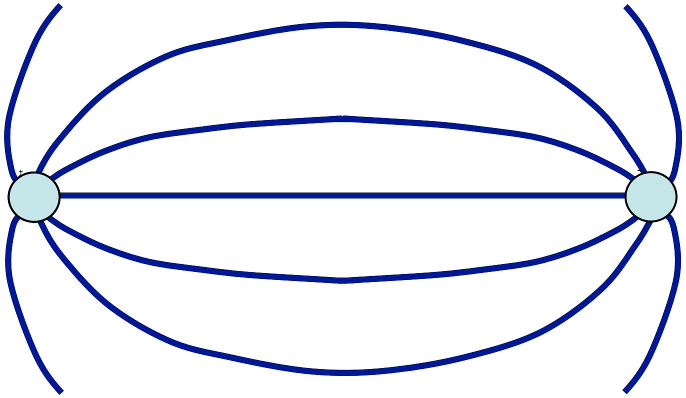
Not infrared safe.

Not a model of confinement.

Confinement

Asymptotic freedom: $Q\bar{Q}$ becomes increasingly QED-like at short distances.

QED:



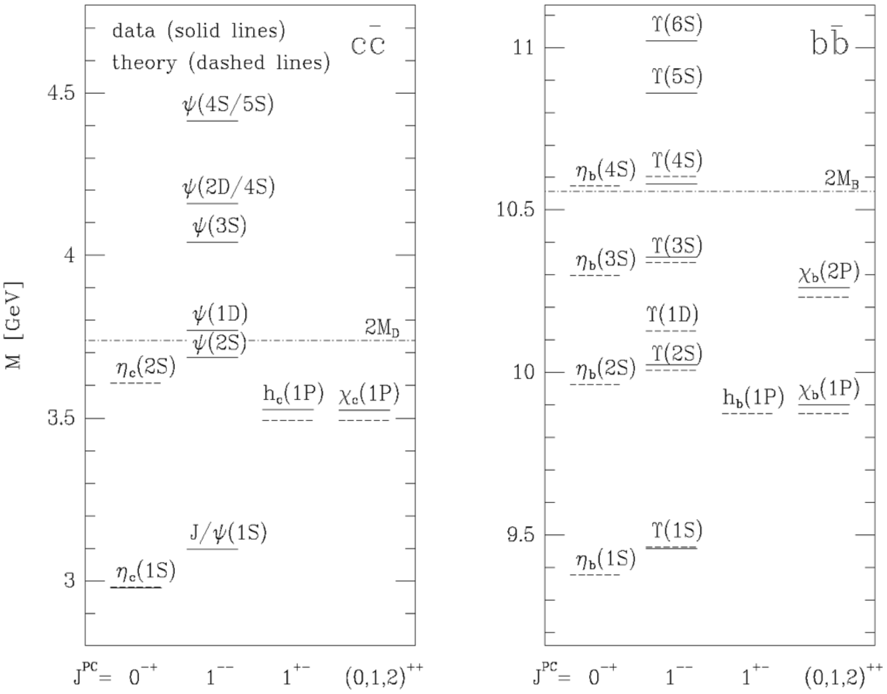
but at long distances, gluon self-interaction makes field lines attract each other:

QCD:

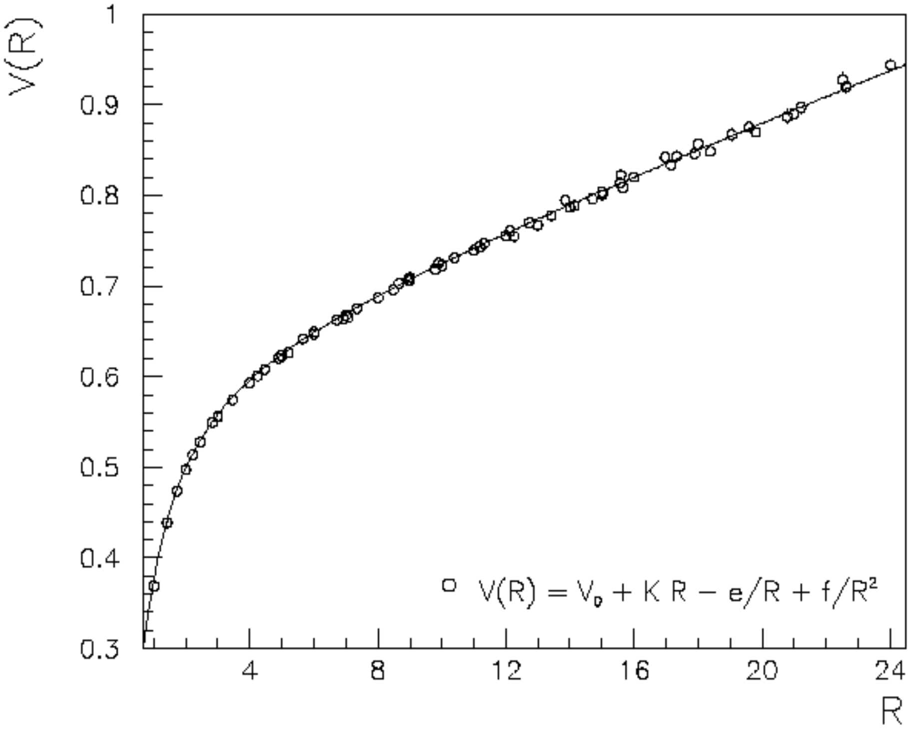
→linear potential → confinement

Interquark potential

Can measure from quarkonia spectra:



or from lattice QCD:



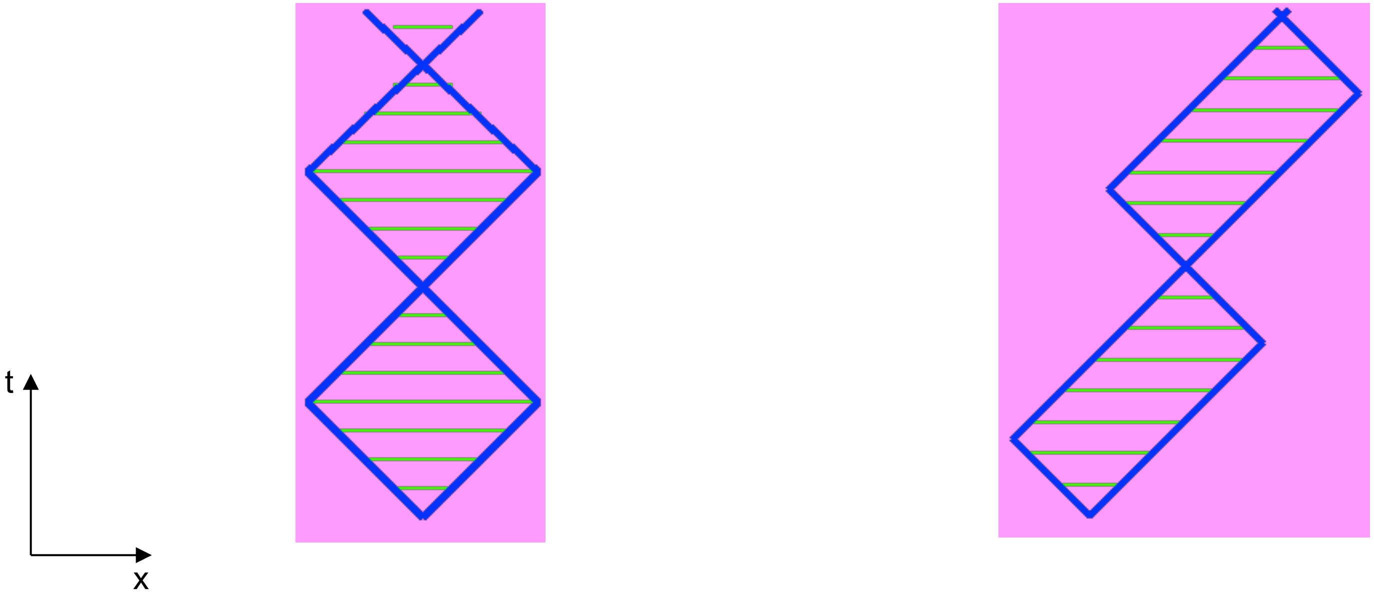
→ String tension

$$\kappa \approx 1 \text{ GeV/fm}.$$

String Model of Mesons

Light quarks connected by string.

L=0 mesons only have ‘yo-yo’ modes:



Obeys area law: $m^2 = 2\kappa^2 \text{ area}$

The Lund String Model

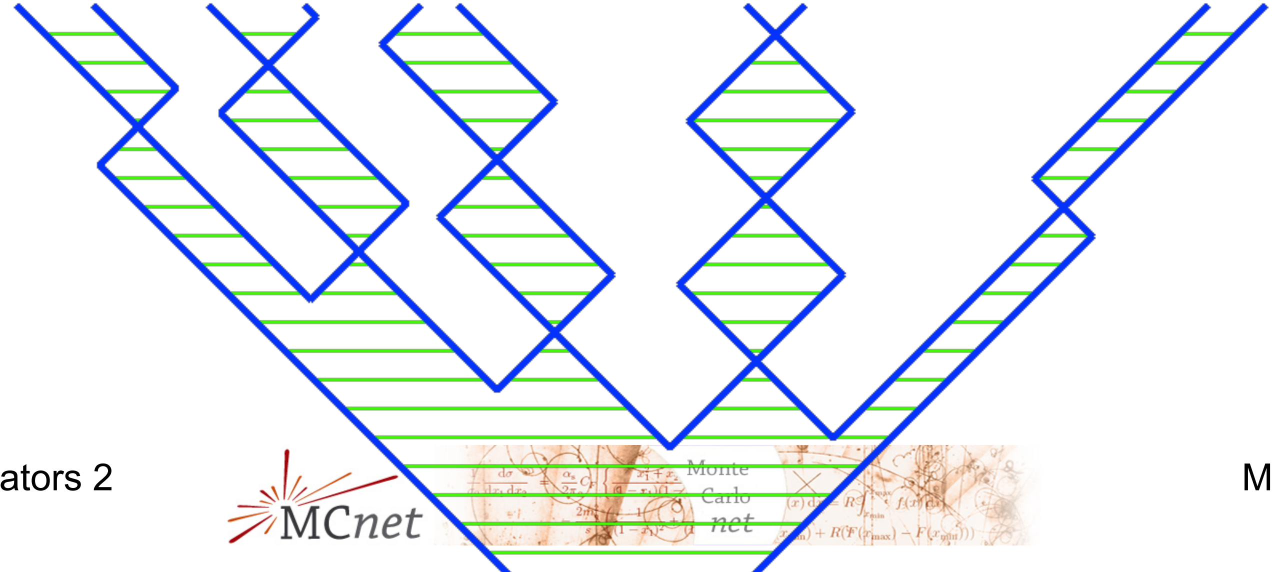
Start by ignoring gluon radiation:

e^+e^- annihilation = pointlike source of $q\bar{q}$ pairs

Intense chromomagnetic field within string → $q\bar{q}$ pairs created by tunnelling. Analogy with QED:

$$\frac{d(\text{Probability})}{dx\ dt} \propto \exp(-\pi m_q^2/\kappa)$$

Expanding string breaks into mesons long before yo-yo point.



Lund Symmetric Fragmentation Function

String picture → constraints on fragmentation function:

- Lorentz invariance
- Acausality
- Left—right symmetry

$$f(z) \propto z^{a_\alpha - a_\beta - 1}(1-z)^{a_\beta}$$

$a_{\alpha,\beta}$ adjustable parameters for quarks α and β.

Fermi motion → Gaussian transverse momentum.

Tunnelling probability becomes

$$\exp\left[-b(m_q^2 + p_t^2)\right]$$

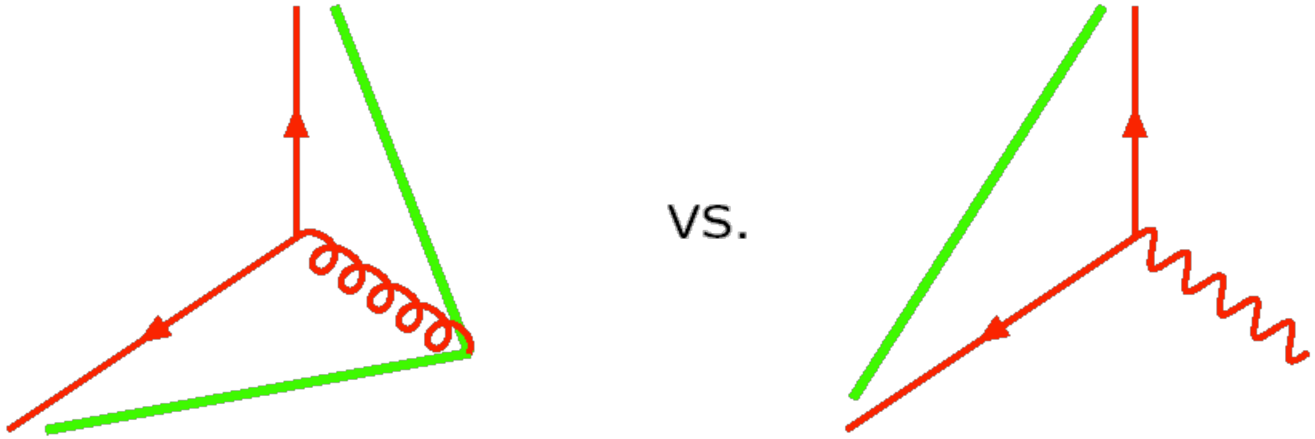
a, b and m_q^2 = main tuneable parameters of model

Three-jet Events

So far: string model = motivated, constrained independent fragmentation!

New feature: universal

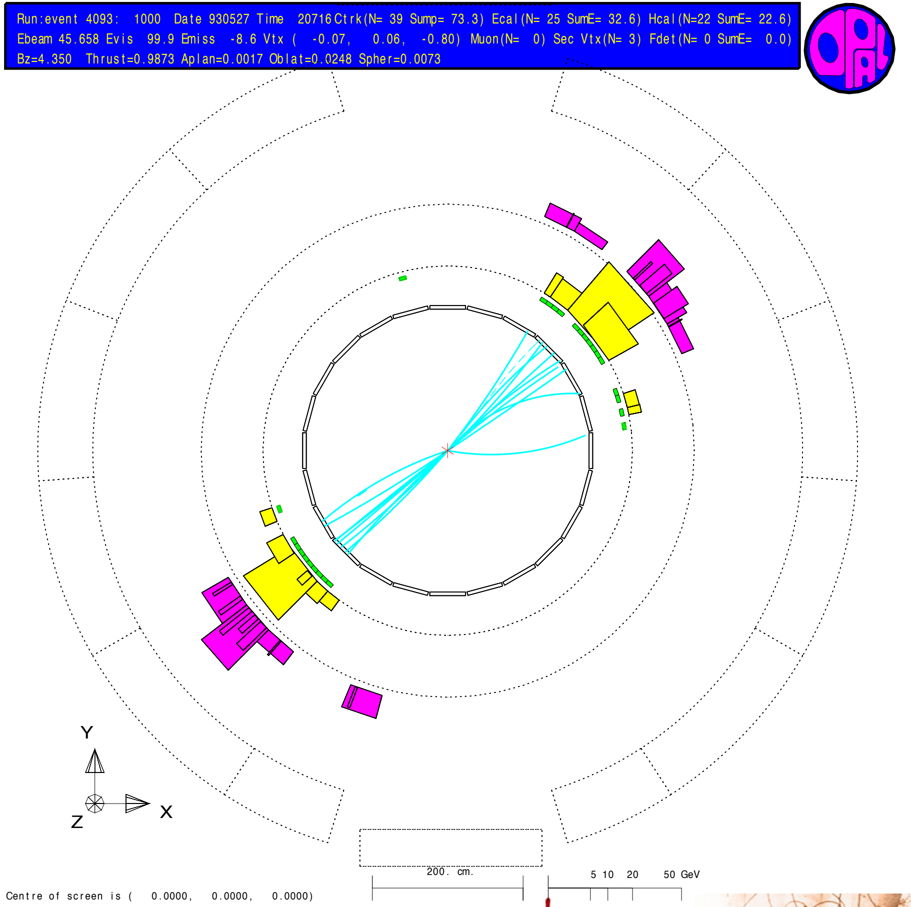
Gluon = kink on string → the string effect



Infrared safe matching with parton shower: gluons with $k_\perp <$ inverse string width irrelevant.

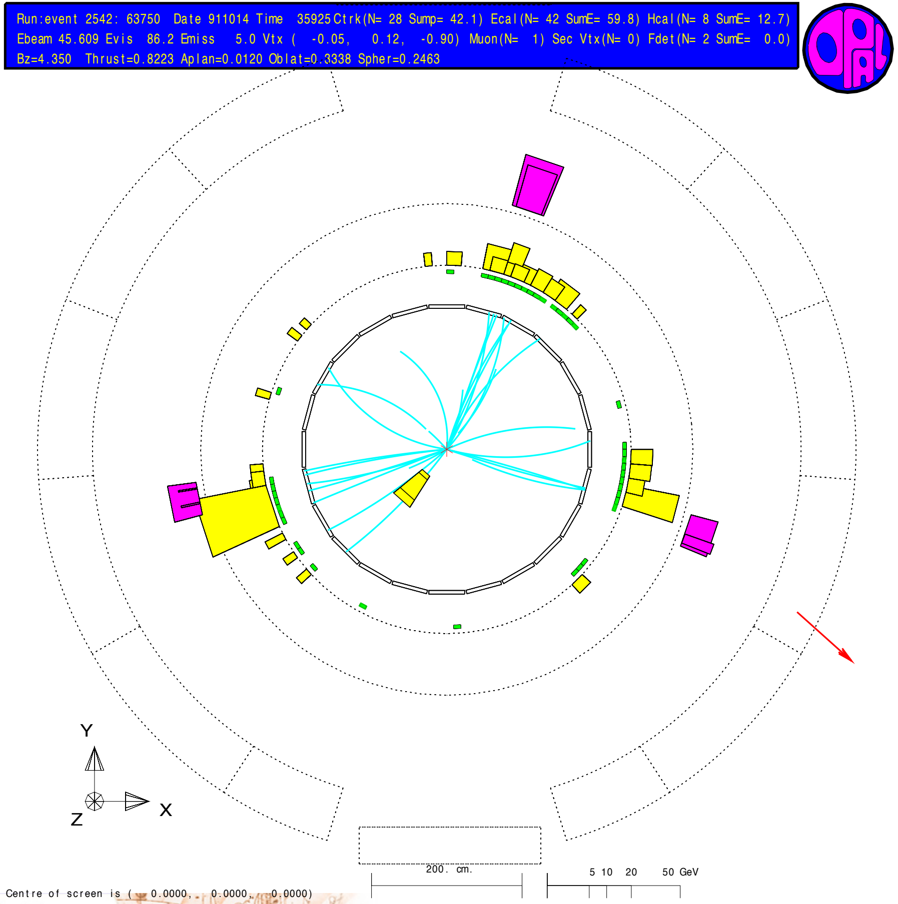
Jet production in $e^+e^- \rightarrow$ hadrons

- Most e^+e^- events consist of two back-to-back jets
- But some consist of three (or more) jets → gluons



```
Run:event 4093:  1000  Date 930527 Time  20716 Ctrk(N= 39 Sump= 73.3) Ecal(N= 25 SumE= 32.6) Hcal(N=22 SumE= 22.6)
Ebeam 45.658 Evis  99.9 Emiss  -8.6 Vtx (  -0.07,   0.06,  -0.80) Muon(N=  0) Sec Vtx(N= 3) Fdet(N= 0 SumE=  0.0)
Bz=4.350  Thrust=0.9873 Aplan=0.0017 Oblat=0.0248 Spher=0.0073
```

```
Centre of screen is (   0.0000,   0.0000,   0.0000)
```



```
Run:event 2542: 63750  Date 911014 Time  35925 Ctrk(N= 28 Sump= 42.1) Ecal(N= 42 SumE= 59.8) Hcal(N= 8 SumE= 12.7)
Ebeam 45.609 Evis  86.2 Emiss   5.0 Vtx (  -0.05,   0.12,  -0.90) Muon(N=  1) Sec Vtx(N= 0) Fdet(N= 2 SumE=  0.0)
Bz=4.350  Thrust=0.8223 Aplan=0.0120 Oblat=0.3338 Spher=0.2463
```

```
Centre of screen is (   0.0000,   0.0000,   0.0000)
```

String Summary

- String model strongly physically motivated.
- Very successful fit to data.
- Universal: fitted to e^+e^- little freedom elsewhere.

- How does motivation translate to prediction?

  ~ one free parameter per hadron/effect!

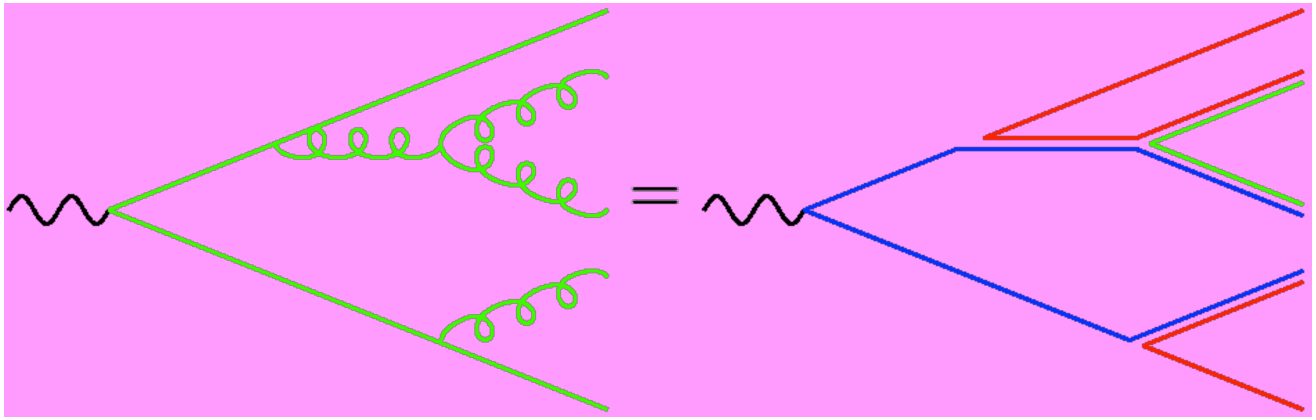
- Blankets too much perturbative information?

- Can we get by with a simpler model?

Preconfinement

Planar approximation: gluon = colour—anticolour pair.

Follow colour structure of parton shower: colour-singlet pairs end up close in phase space

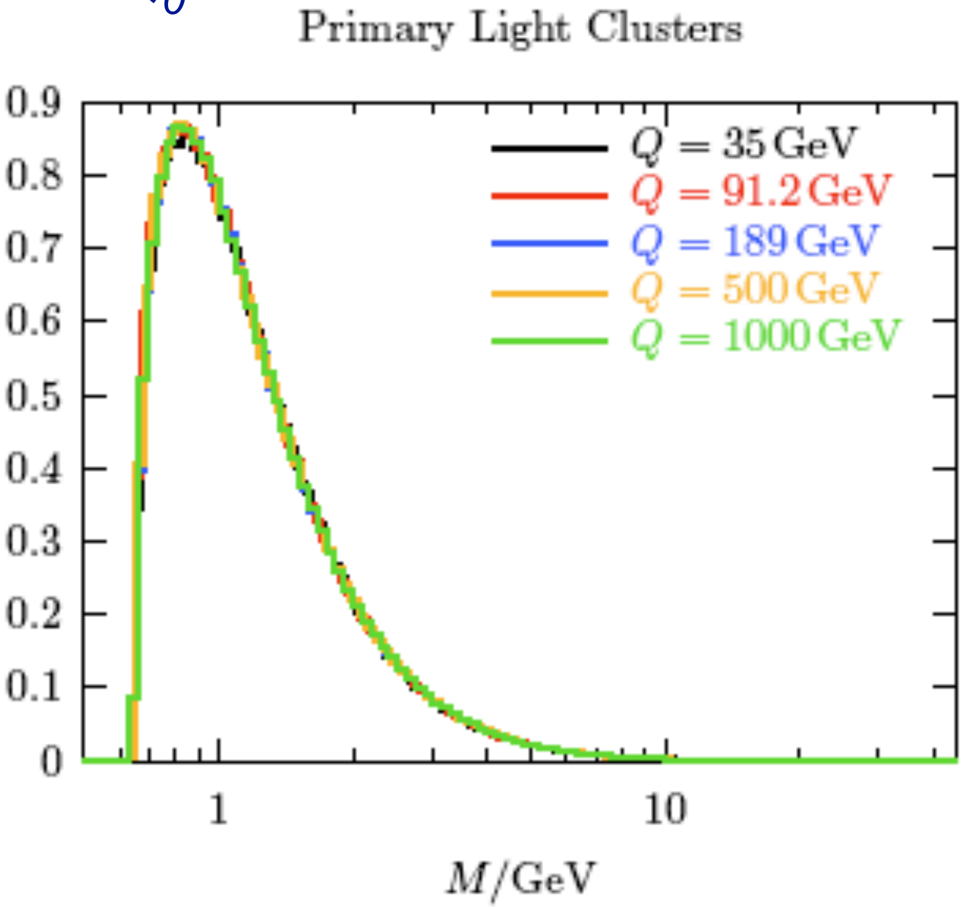


Mass spectrum of colour-singlet pairs asymptotically independent of energy, production mechanism, …

Peaked at low mass $\sim Q_0$.

Cluster mass distribution

- Independent of shower scale *Q*
  - depends on Q_0 and Λ



The Naïve Cluster Model

Project colour singlets onto continuum of high-mass mesonic resonances (=clusters). Decay to lighter well-known resonances and stable hadrons.

Assume spin information washed out:

decay = pure phase space.

→ heavier hadrons suppressed

→ baryon & strangeness suppression 'for free' (i.e. untuneable).

Hadron-level properties fully determined by cluster mass spectrum, i.e. by perturbative parameters.

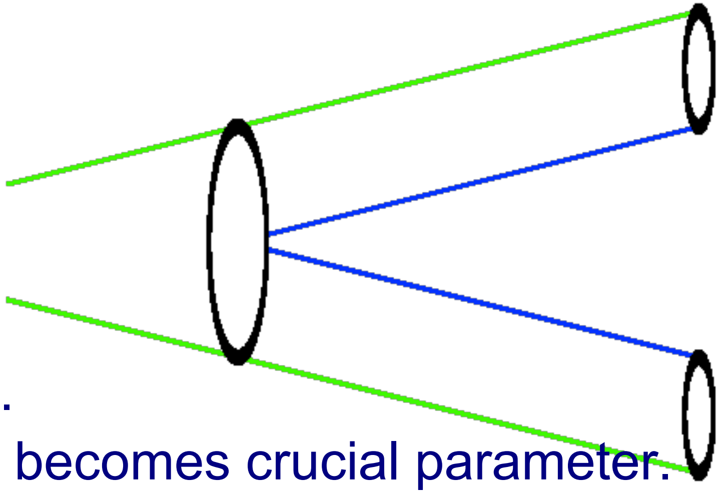
Q_0 crucial parameter of model.

The Cluster Model

Although cluster mass spectrum peaked at small m, broad tail at high m.

“Small fraction of clusters too heavy for isotropic two-body decay to be a good approximation”.

Longitudinal cluster fission:



Rather string-like.

Fission threshold becomes crucial parameter.

~15% of primary clusters get split but ~50% of hadrons come from them.

The Cluster Model

“Leading hadrons are too soft”

→ ‘perturbative’ quarks remember their direction somewhat

$$P(\theta^2) \sim \exp(-\theta^2/2\theta_0^2)$$

Rather string-like.

Extra adjustable parameter.

Strings

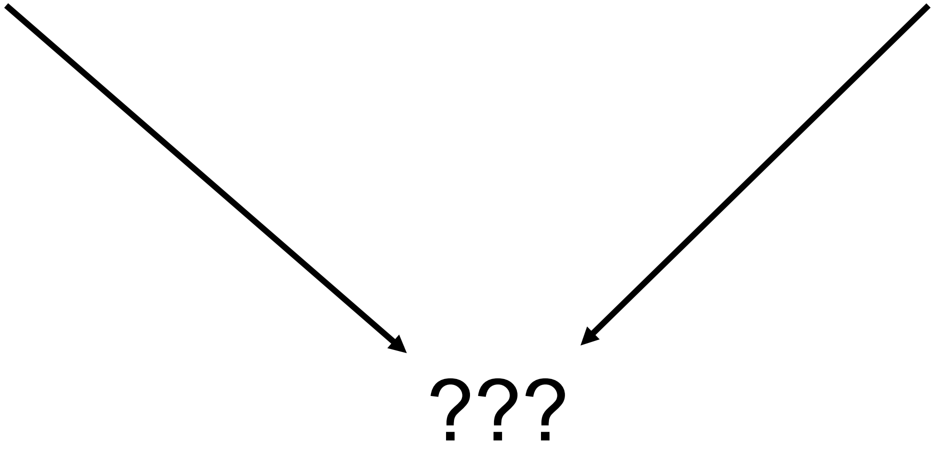
“Hadrons are produced by hadronization: you must get the non-perturbative dynamics right”

Improving data has meant successively refining perturbative phase of evolution…

Clusters

“Get the perturbative phase right and any old hadronization model will be good enough”

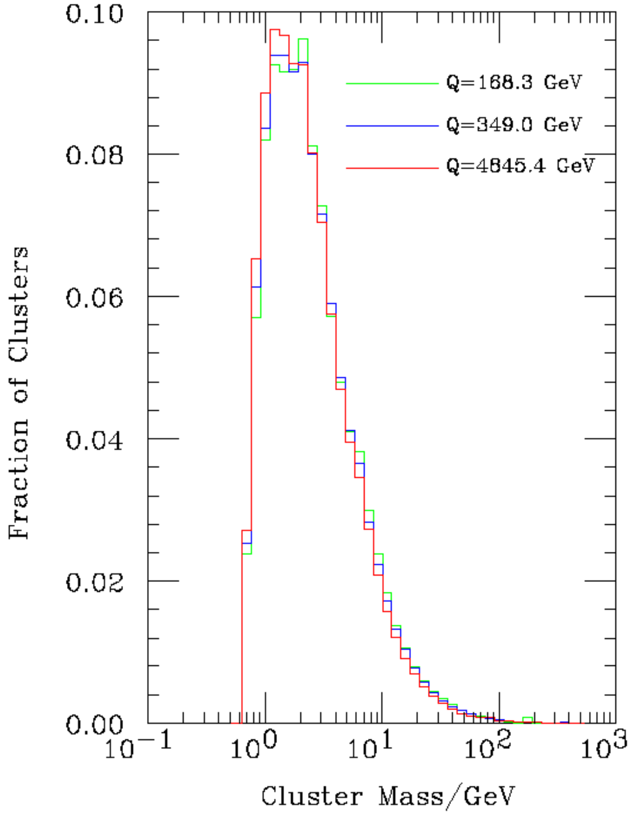
Improving data has meant successively making non-perturbative phase more string-like…

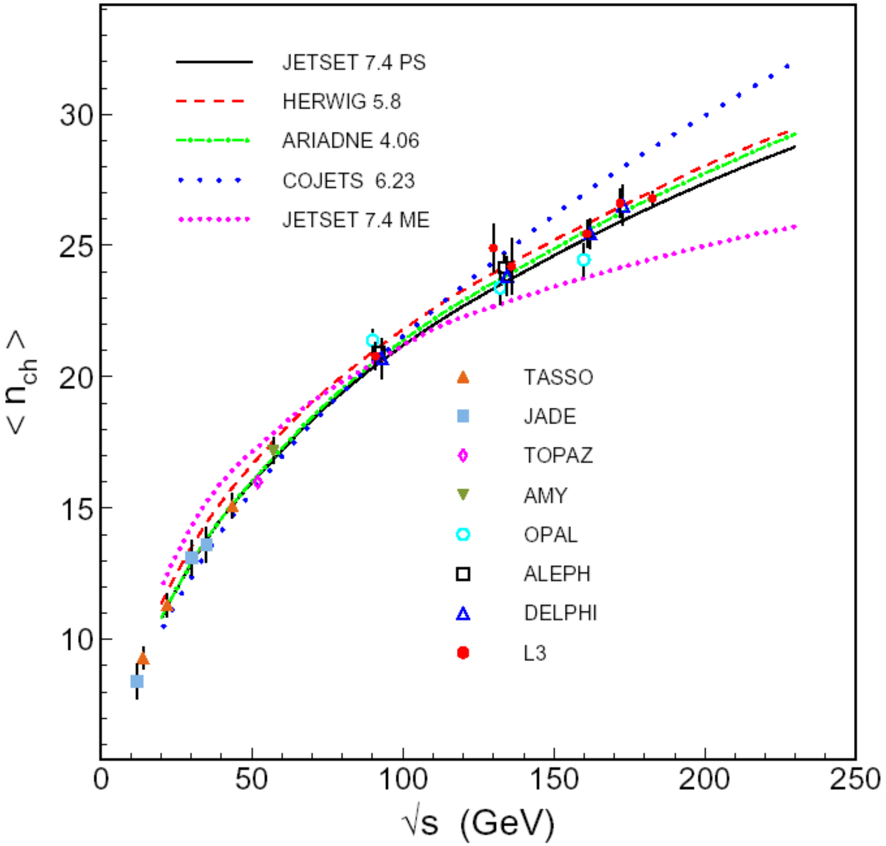


???

Universality of Hadronization Parameters

- Is guaranteed by preconfinement: do not need to retune at each energy

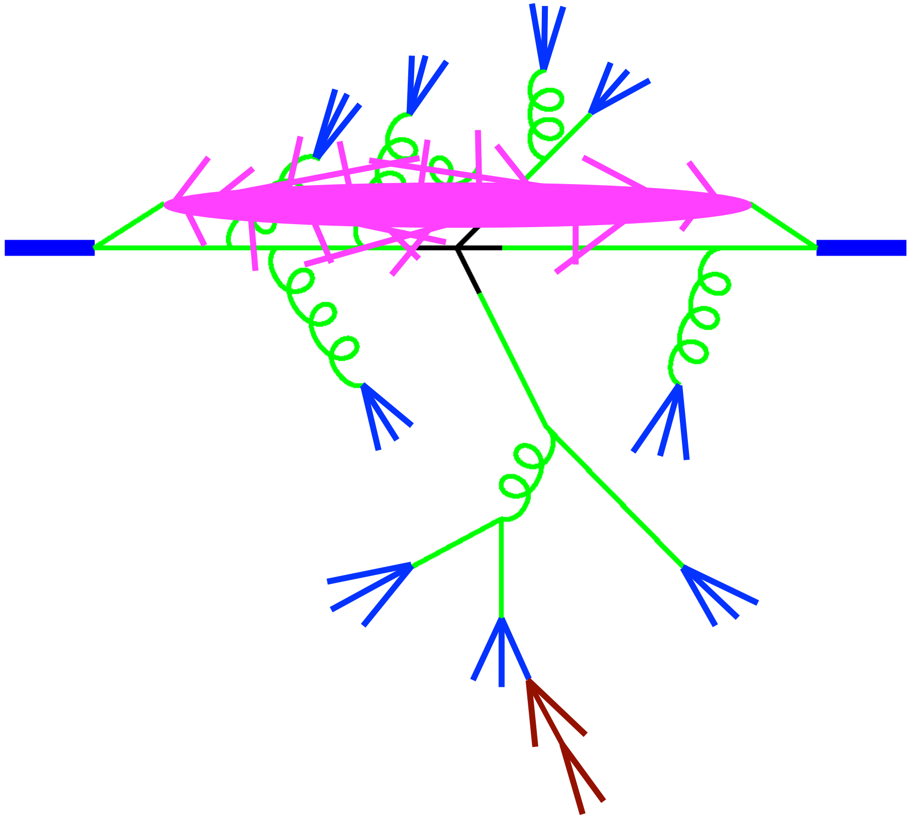




→ Only tune what's new in hadron—hadron collisions

Structure of LHC Events

1. Hard process
2. Parton shower
3. Hadronization
4. Underlying event
5. Unstable particle decays



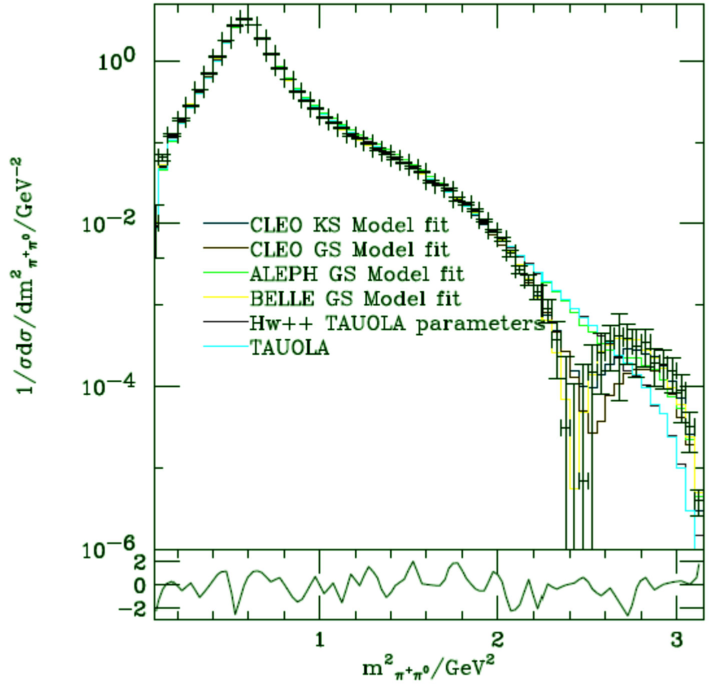
Secondary Decays and Decay Tables

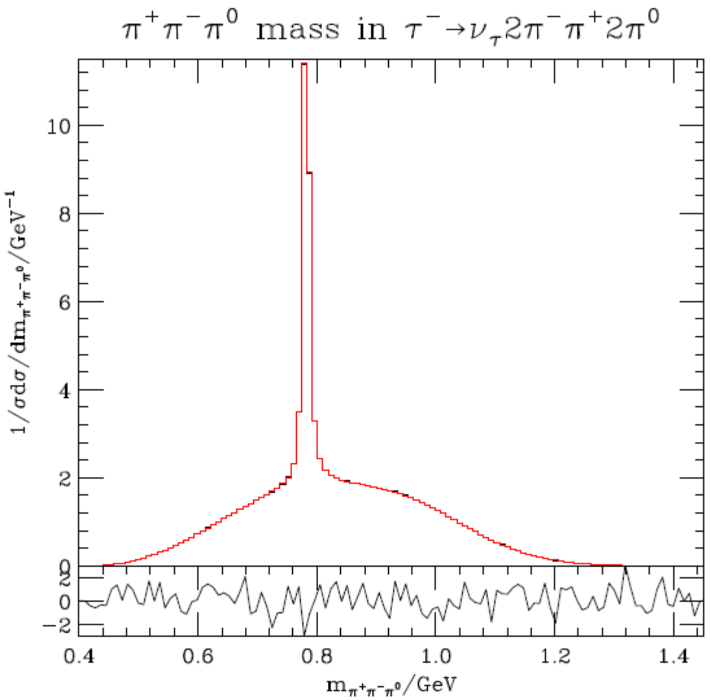
- Often forgotten ingredient of event generators:
  - String and cluster decay to some stable hadrons but mainly unstable resonances
  - These decay further “according to PDG data tables”
    - Matrix elements for n-body decays
  - But…
    - Not all resonances in a given multiplet have been measured
    - Measured branching fractions rarely add up to 100% exactly
    - Measured branching fractions rarely respect isospin exactly
  - So need to make a lot of choices
  - Has a significant effect on hadron yields, transverse momentum release, hadronization corrections to event shapes, …
  - Should consider the decay table choice part of the tuned set

Secondary particle decays

- Previous generations typically used external packages, e.g. TAUOLA, PHOTOS, EVTGEN
- Sherpa & Herwig++ contain at least as complete a description in all areas…
- without interfacing issues (c.f. τ spin)

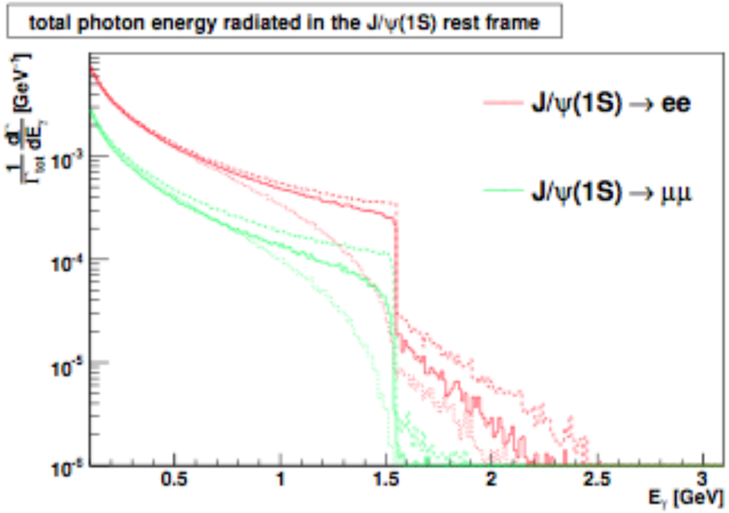
Tau Decays



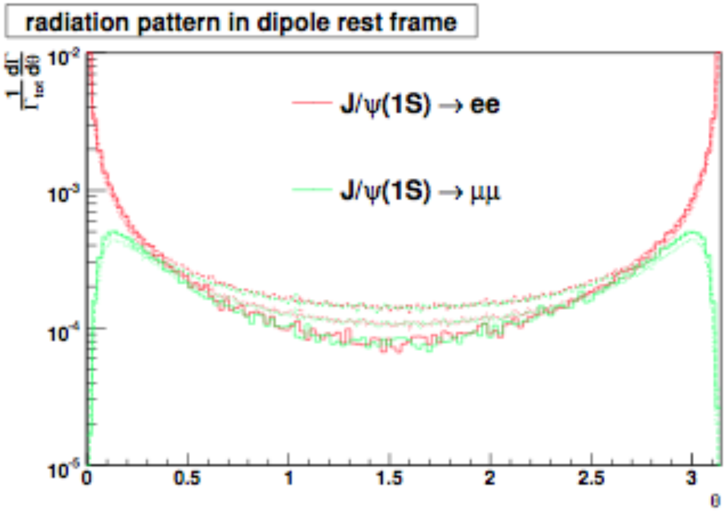


Mass spectrum of $\pi\pi$ in $\tau\to\pi\pi\nu_\tau$ for various models and example of mass distribution in $\tau\to5\pi\nu_\tau$ comparing Herwig++ and TAUOLA.

Leptonic hadron decays: $J/\psi \to \ell\bar{\ell}$



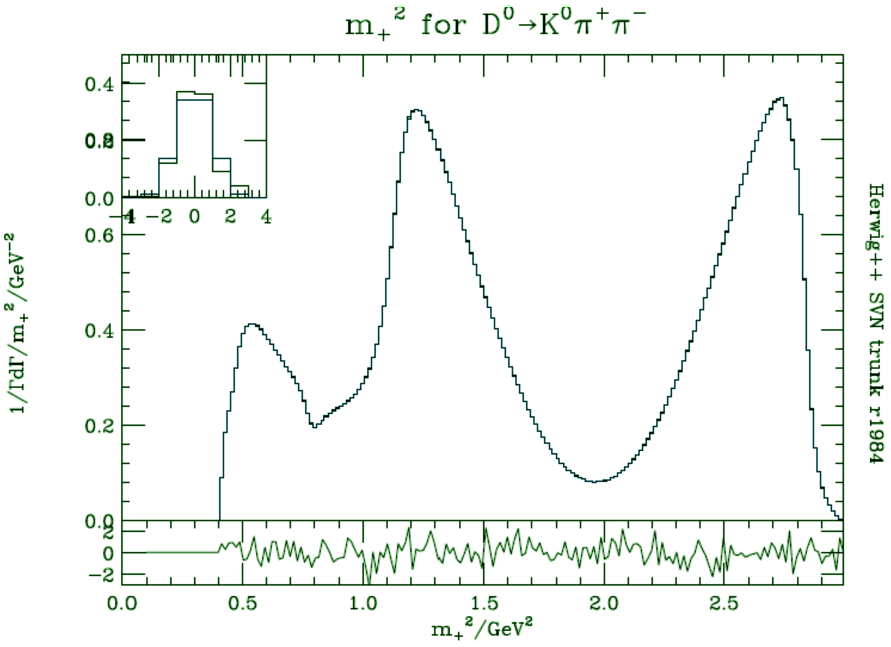
total radiated energy in the J/ψ rest frame

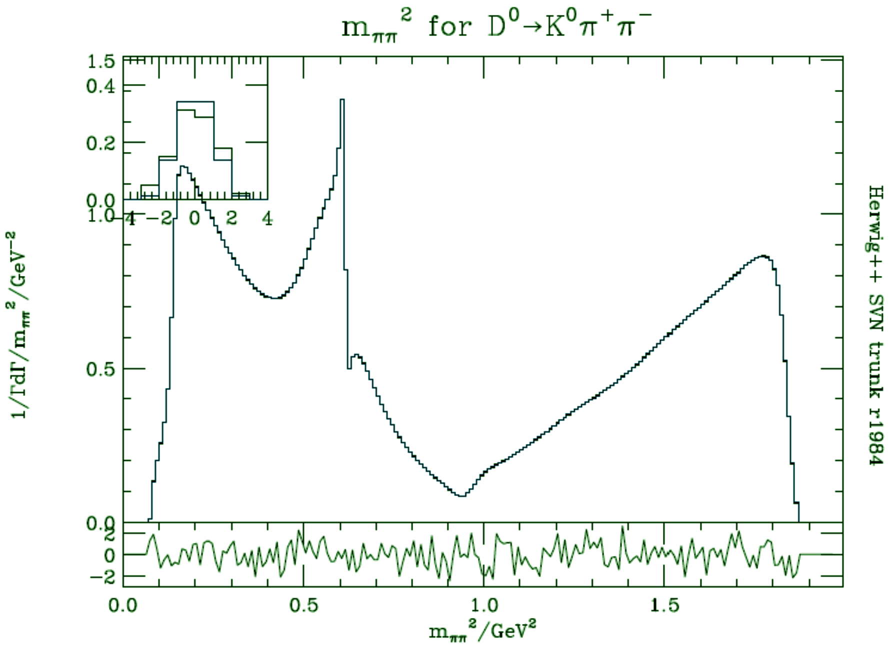


angular spectrum in the rest frame of the dipole

- soft only (dotted)
- collinear approximated ME (dashed)
- exact ME (solid)

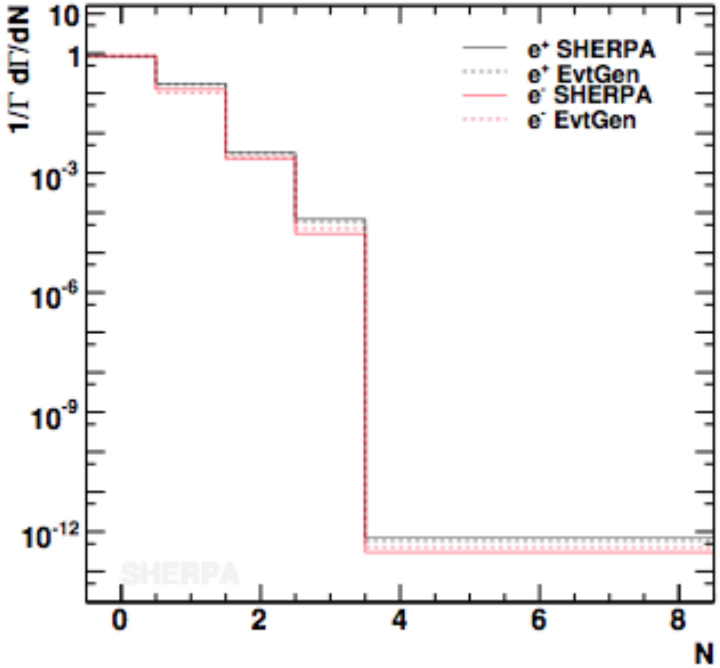
D→Kππ



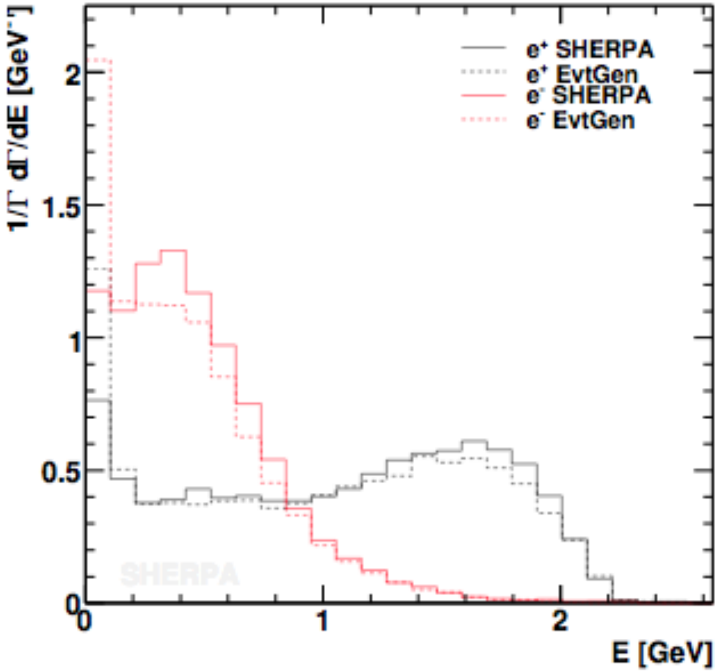


Comparison of Herwig++ and EvtGen implementations of the fit of Phys. Rev. D63 (2001) 092001 (CLEO).

Inclusive observables for B^+ decay



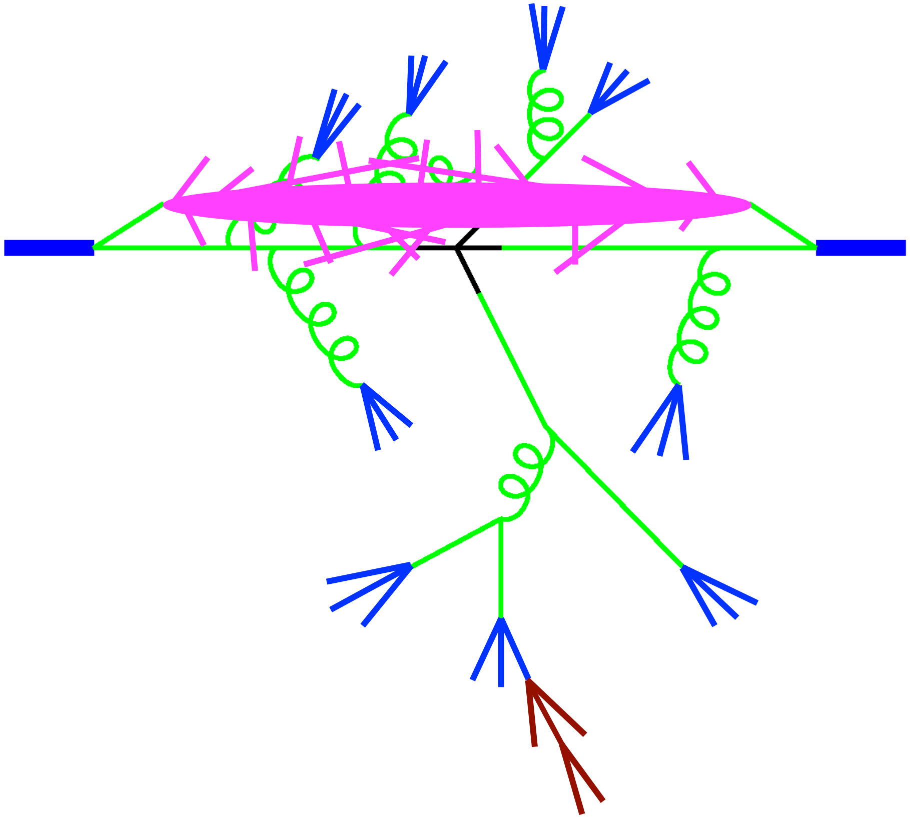
Electron multiplicity



Electron energy spectrum

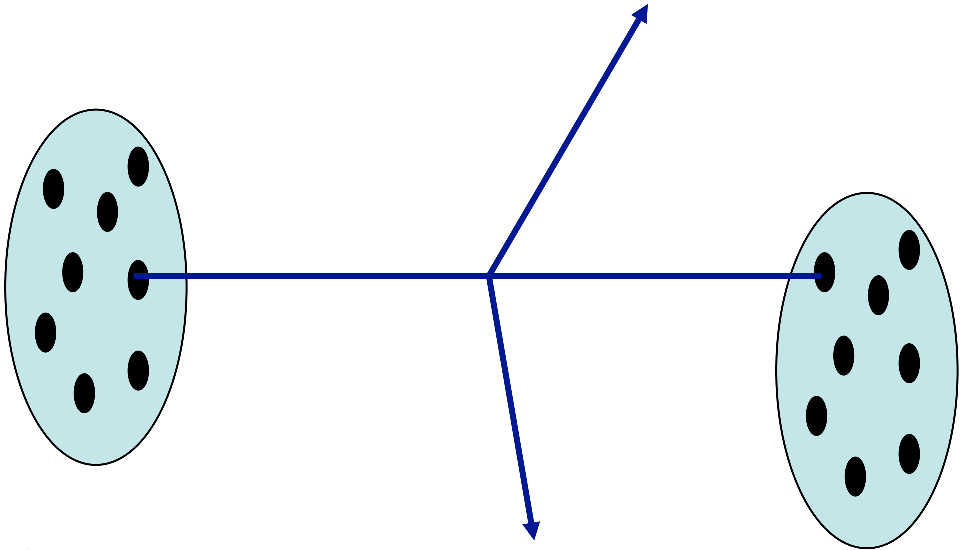
Structure of LHC Events

1. Hard process
2. Parton shower
3. Hadronization
4. Underlying event
5. Unstable particle decays



The Underlying Event

- Protons are extended objects
- After a parton has been scattered out of each, what happens to the remnants?

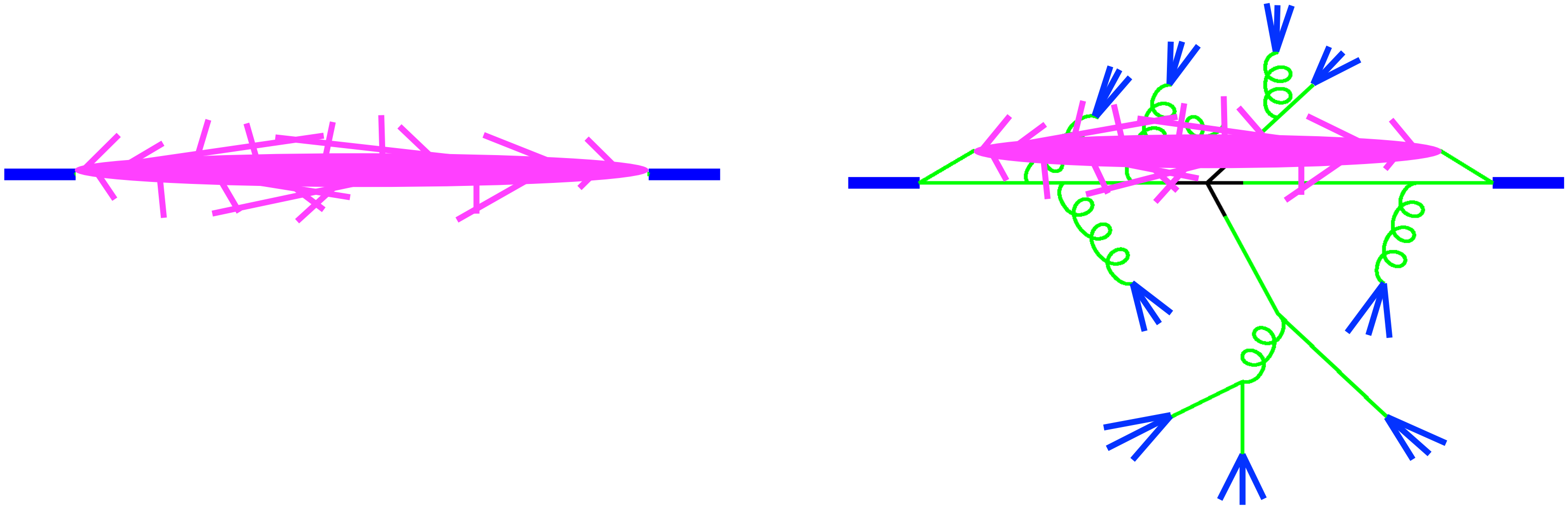


Two models:

- Non-perturbative: Soft parton—parton cross section is so large that the remnants always undergo a soft collision.
- Perturbative: ‘Hard’ parton—parton cross section huge at low p_t, high energy, dominates inelastic cross section and is calculable.

The Basics: event classes

‘Minimum bias’ collision and underlying event



Minimum bias = experimental statement

Models = zero bias? i.e. inclusive sample of all inelastic (non-diffractive?) events

CERN-PH-TH-2010-298 KA-TP-40-2010
Cavendish-HEP-10/21 DCPT/10/202
MAN/HEP/2010/23 IPPP/10/101
SLAC-PUB-14333 LU TP 10-28
HD-THEP-10-24 MCnet-11-01

General-purpose event generators for LHC physics

Andy Buckley[a], Jonathan Butterworth[b], Stefan Gieseke[c],
David Grellscheid[d], Stefan Höche[e], Hendrik Hoeth[d], Frank Krauss[d],
Leif Lönnblad[f,g], Emily Nurse[b], Peter Richardson[d], Steffen Schumann[h],
Michael H. Seymour[i], Torbjörn Sjöstrand[f], Peter Skands[g], Bryan Webber[j]

[a] *PPE Group, School of Physics & Astronomy, University of Edinburgh, EH25 9PN, UK*
[b] *Department of Physics & Astronomy, University College London, WC1E 6BT, UK*
[c] *Institute for Theoretical Physics, Karlsruhe Institute of Technology, D-76128 Karlsruhe*
[d] *Institute for Particle Physics Phenomenology, Durham University, DH1 3LE, UK*
[e] *SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA*
[f] *Department of Astronomy and Theoretical Physics, Lund University, Sweden*
[g] *PH Department, TH Unit, CERN, CH-1211 Geneva 23, Switzerland*
[h] *Institute for Theoretical Physics, University of Heidelberg, 69120 Heidelberg, Germany*
[i] *School of Physics and Astronomy, University of Manchester, M13 9PL, UK*
[j] *Cavendish Laboratory, J.J. Thomson Avenue, Cambridge CB3 0HE, UK*

Abstract

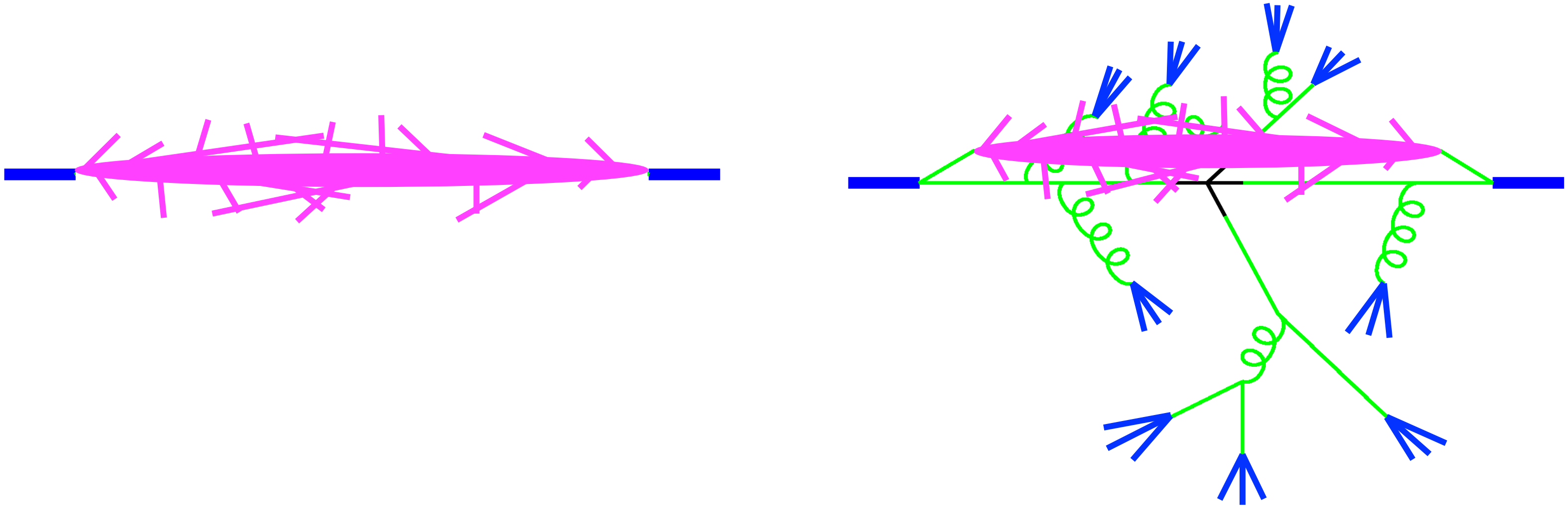
We review the physics basis, main features and use of general-purpose Monte Carlo event generators for the simulation of proton-proton collisions at the Large Hadron Collider. Topics included are: the generation of hard-scattering matrix elements for processes of interest, at both leading and next-to-leading QCD perturbative order; their matching to approximate treatments of higher orders based on the showering approximation; the parton and dipole shower formulations; parton distribution functions for event generators; non-perturbative aspects such as soft QCD collisions, the underlying event and diffractive processes; the string and cluster models for hadron formation; the treatment of hadron and tau decays; the inclusion of QED radiation and beyond-Standard-Model processes. We describe the principal features of the ARIADNE, Herwig++, PYTHIA 8 and SHERPA generators, together with the Rivet and Professor validation and tuning tools, and discuss the physics philosophy behind the proper use of these generators and tools. This review is aimed at phenomenologists wishing to understand better how parton-level predictions are translated into hadron-level events as well as experimentalists wanting a deeper insight into the tools available for signal and background simulation at the LHC.

Preprint submitted to Physics Reports *January 13, 2011*

Event Generators 2

The Basics: event classes

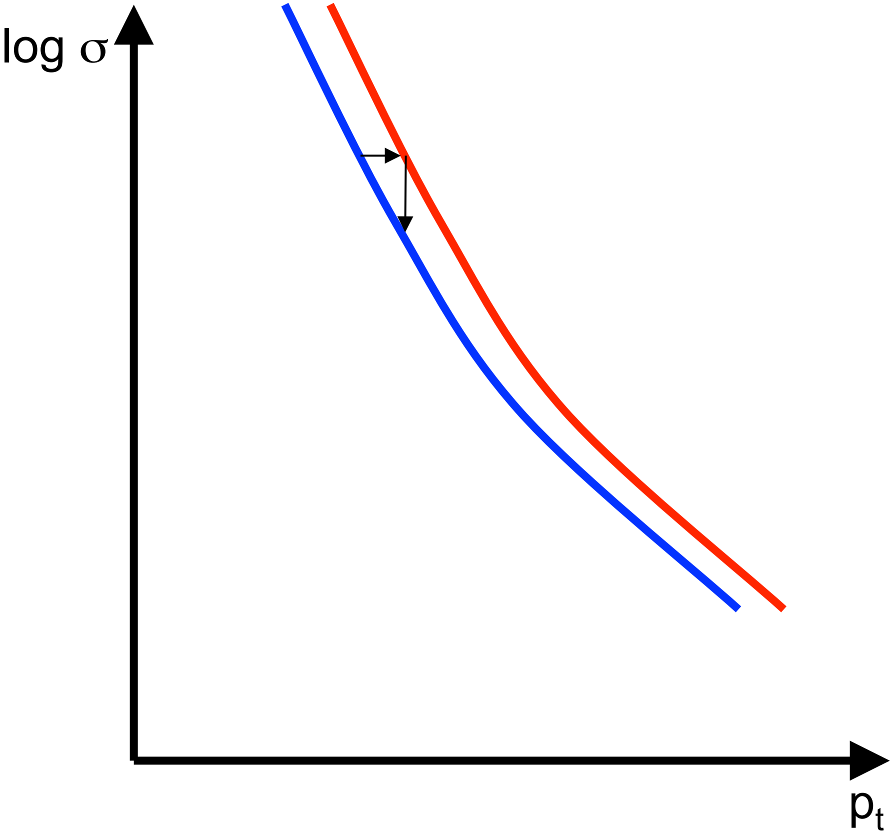
‘Soft inclusive’ events and the underlying event



How similar are they?

Fluctuations and correlations play crucial role

Fluctuations and correlations



Steep distribution ⇒ small sideways shift = large vertical

Rare fluctuations can have a huge influence

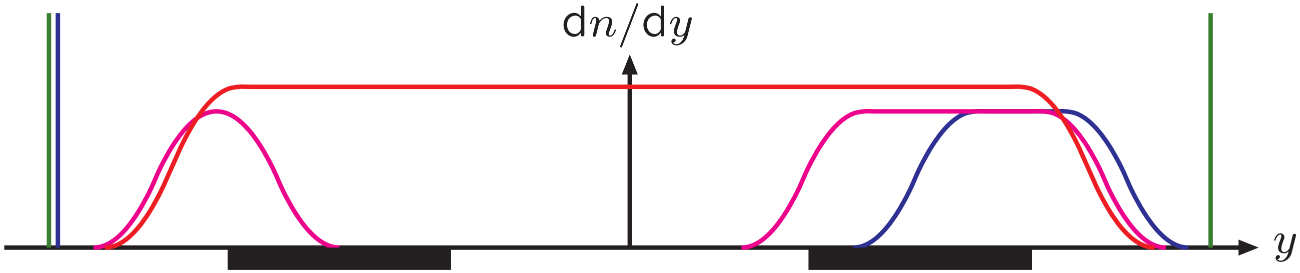
$1/p_t^n \to n$th moment

⇒ corrections depend on physics process

The Basics – what's what

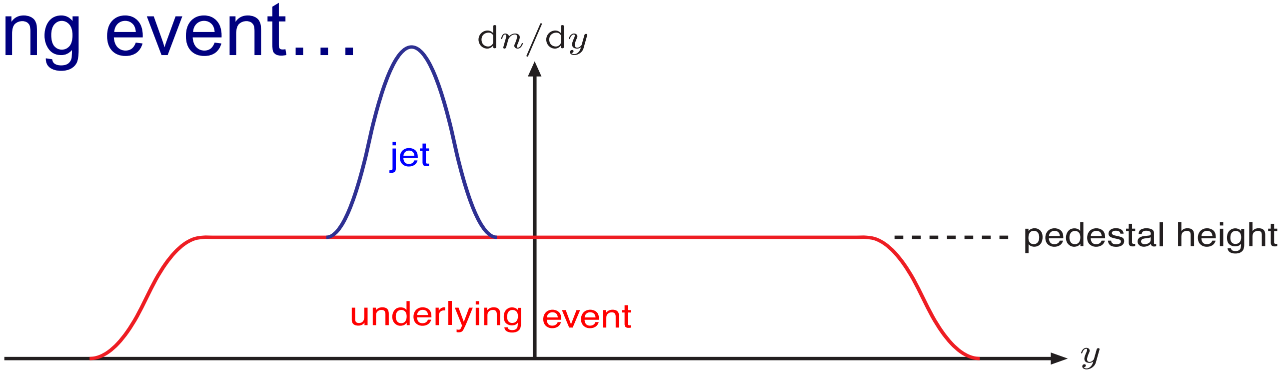
- Soft inclusive collisions…

$$\sigma_{\text{tot}} = \sigma_{\text{elastic}} + \sigma_{\text{single-diffractive}} + \sigma_{\text{double-diffractive}} + \ldots + \sigma_{\text{non-diffractive}}$$



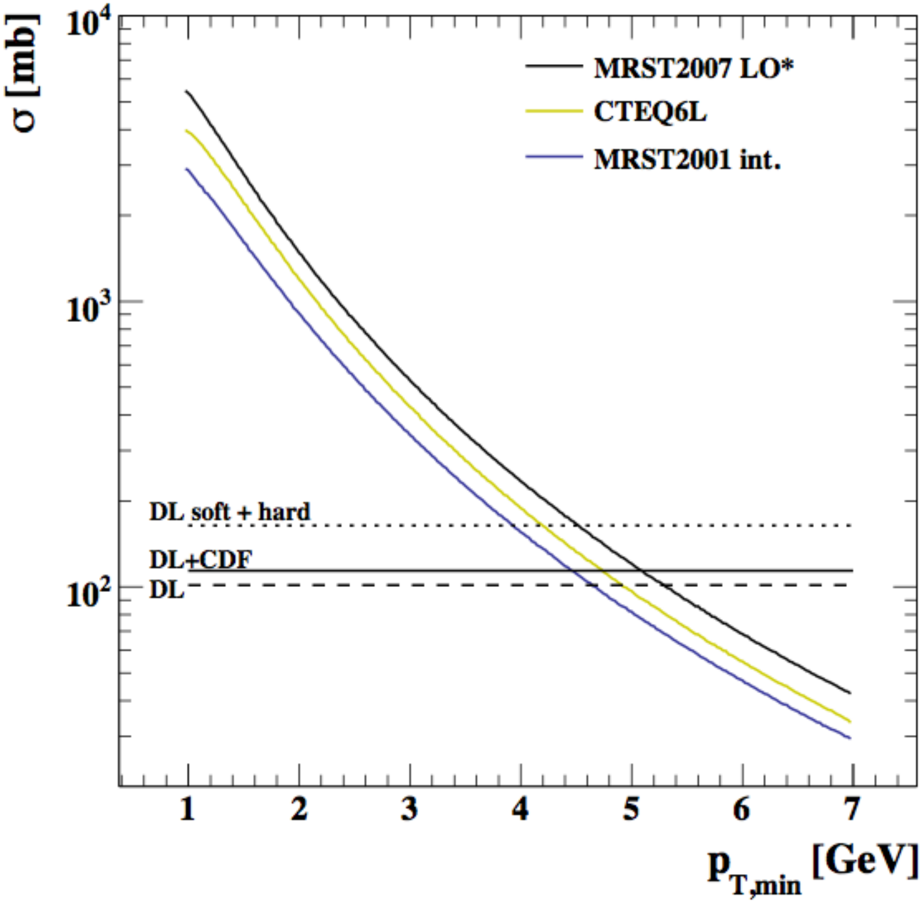
reality: $\sigma_{\text{min-bias}} \approx \sigma_{\text{non-diffractive}} + \sigma_{\text{double-diffractive}} \approx 2/3 \times \sigma_{\text{tot}}$

- Underlying event…



The Basics: Multiparton Interaction Model

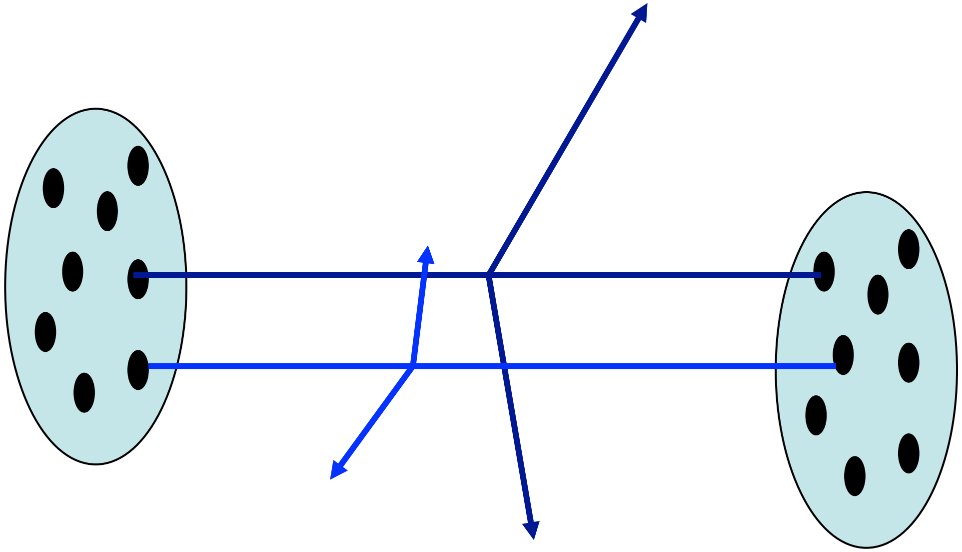
For small $p_{t\ min}$ and high energy inclusive parton—parton cross section is larger than total proton—proton cross section.



The Basics: Multiparton Interaction Model

For small $p_{t\,\mathrm{min}}$ and high energy inclusive parton—parton cross section is larger than total proton—proton cross section.

→ More than one parton—parton scatter per proton—proton



Sjöstrand, van Zijl,
Phys. Rev. D36
(1987) 2019

Need a model of spatial distribution within proton

→ Perturbation theory gives you n-scatter distributions

Matter Distributions

- Usually assume *x* and *b* factorize

$$n_i(x,b;\mu^2,s) = f_i(x;\mu^2)\,G(b,s)$$

- and *n*-parton distributions are independent

$$n_{i,j}(x_i,x_j,b_i,b_j) = n_i(x_i,b_i)\,n_j(x_j,b_j)$$

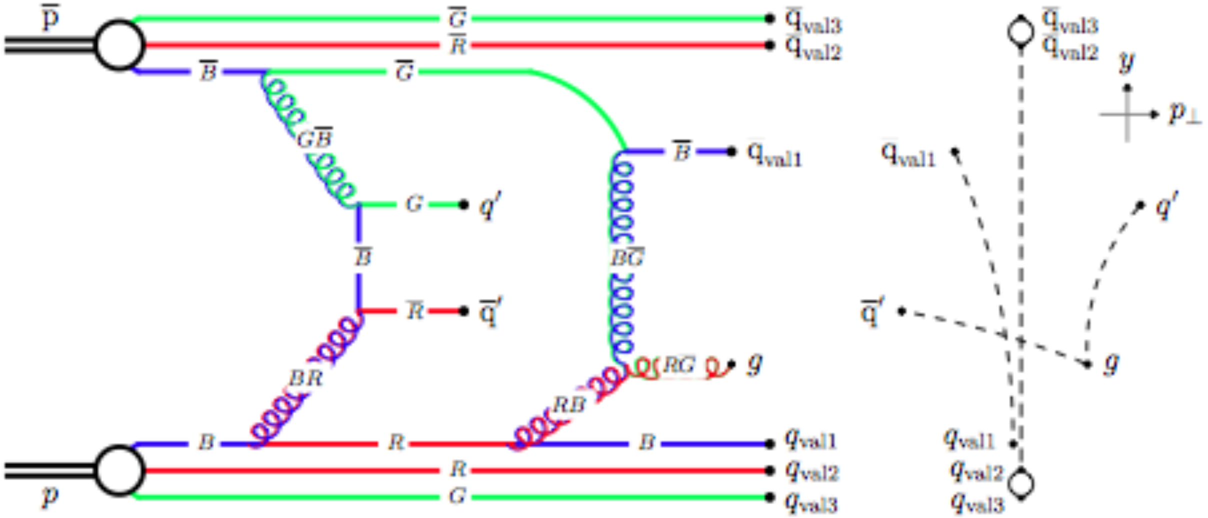
⇒ scatters Poissonian at fixed impact parameter

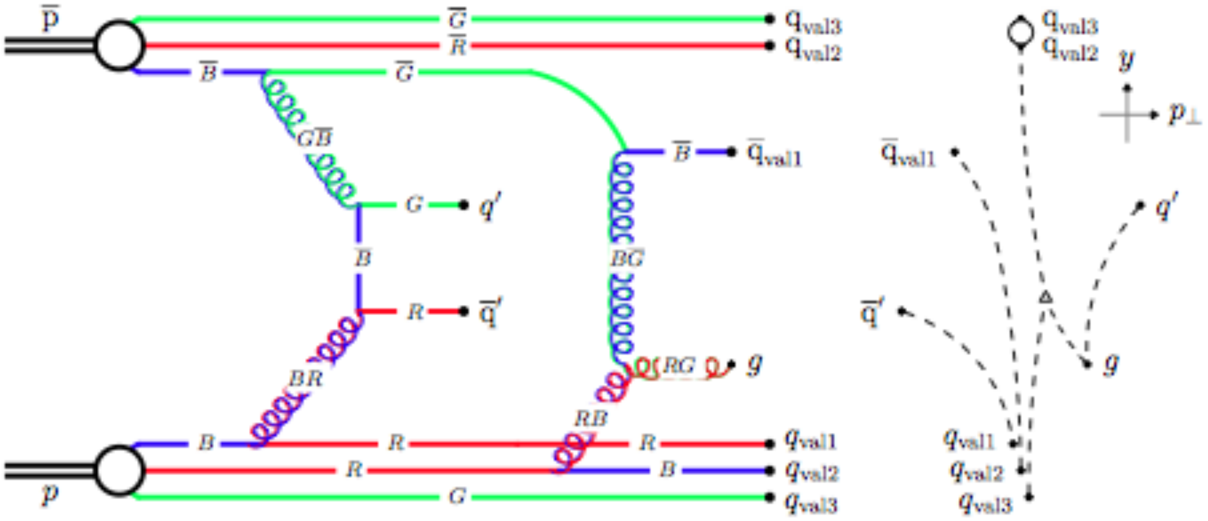
$$\sigma_n = \int d^2b\,\frac{(A(b)\sigma^{inc})^n}{n!}\,\mathrm{exp}(-A(b)\sigma^{inc})$$

$$A(b) = \int d^2b_1G(b_1)\,d^2b_2G(b_2)\,\delta(b-b_1+b_2)$$

Colour correlations

Can have a big influence on final states





→ see later

The Herwig++ Model (formerly known as Jimmy+Ivan)

- Take eikonal+partonic scattering seriously

$$\sigma_{tot} \;=\; 2\int d^2b\left(1-e^{-\frac{1}{2}A(b)\sigma_{inc}}\right)$$

$$B=\left[\frac{d}{dt}\left(\ln\frac{d\sigma_{el}}{dt}\right)\right]_{t=0} = \frac{1}{\sigma_{tot}}\int d^2b\,b^2\left(1-e^{-\frac{1}{2}A(b)\sigma_{inc}}\right)$$

- given form of matter distribution ⇒ size and σ_{inc}

Bähr, Butterworth & MHS, JHEP 0901:067, 2009

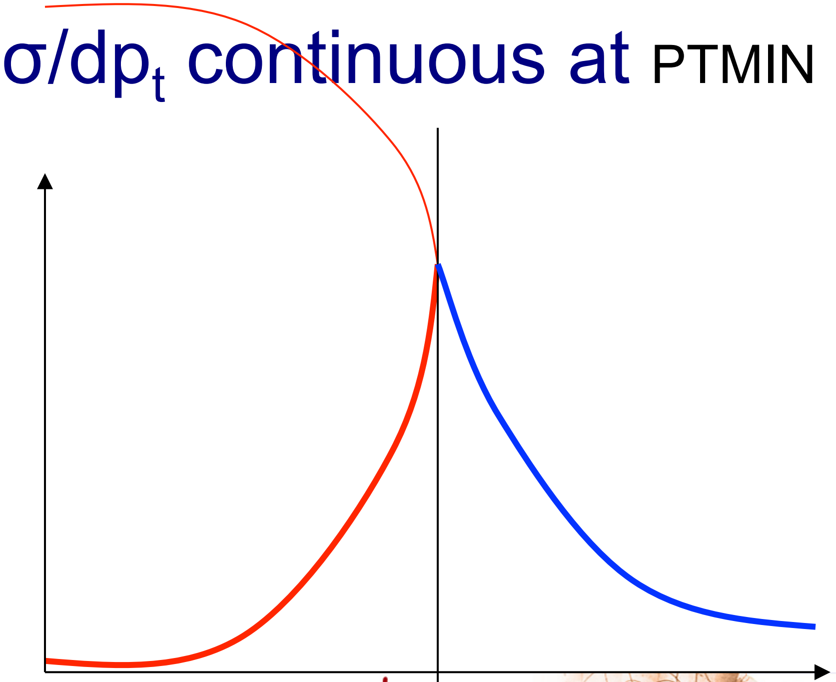
- too restrictive ⇒

$$\sigma_{tot} \;=\; 2\int d^2b\left(1-e^{-\frac{1}{2}(A_{\mathrm{soft}}(b)\sigma_{\mathrm{soft,inc}}+A_{\mathrm{hard}}(b)\sigma_{\mathrm{hard,inc}})}\right)$$

- ⇒ two free parameters

Final state implementation

- Pure independent perturbative scatters above PTMIN
- Gluonic scattering below PTMIN with total $\sigma_{soft,inc}$ and Gaussian distribution in p_t
- $d\sigma/dp_t$ continuous at PTMIN



→ possibility that entire process could be described perturbatively?

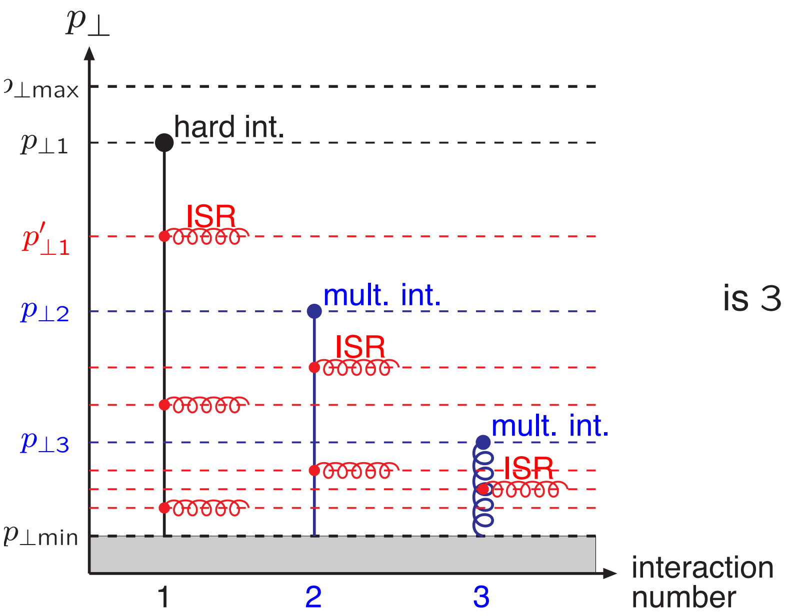
Pythia implementation

(4) Evolution interleaved with ISR (2004)

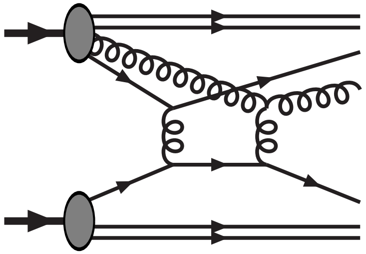
- Transverse-momentum-ordered showers

$$\frac{\mathrm{d}\mathcal{P}}{\mathrm{d}p_\perp} = \left(\frac{\mathrm{d}\mathcal{P}_{\mathrm{MI}}}{\mathrm{d}p_\perp} + \sum \frac{\mathrm{d}\mathcal{P}_{\mathrm{ISR}}}{\mathrm{d}p_\perp}\right) \exp\left(-\int_{p_\perp}^{p_{\perp i-1}} \left(\frac{\mathrm{d}\mathcal{P}_{\mathrm{MI}}}{\mathrm{d}p'_\perp} + \sum \frac{\mathrm{d}\mathcal{P}_{\mathrm{ISR}}}{\mathrm{d}p'_\perp}\right) \mathrm{d}p'_\perp\right)$$

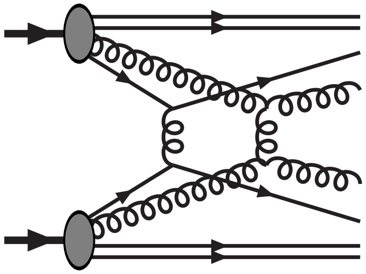
with ISR sum over all previous MI



(5) Rescattering

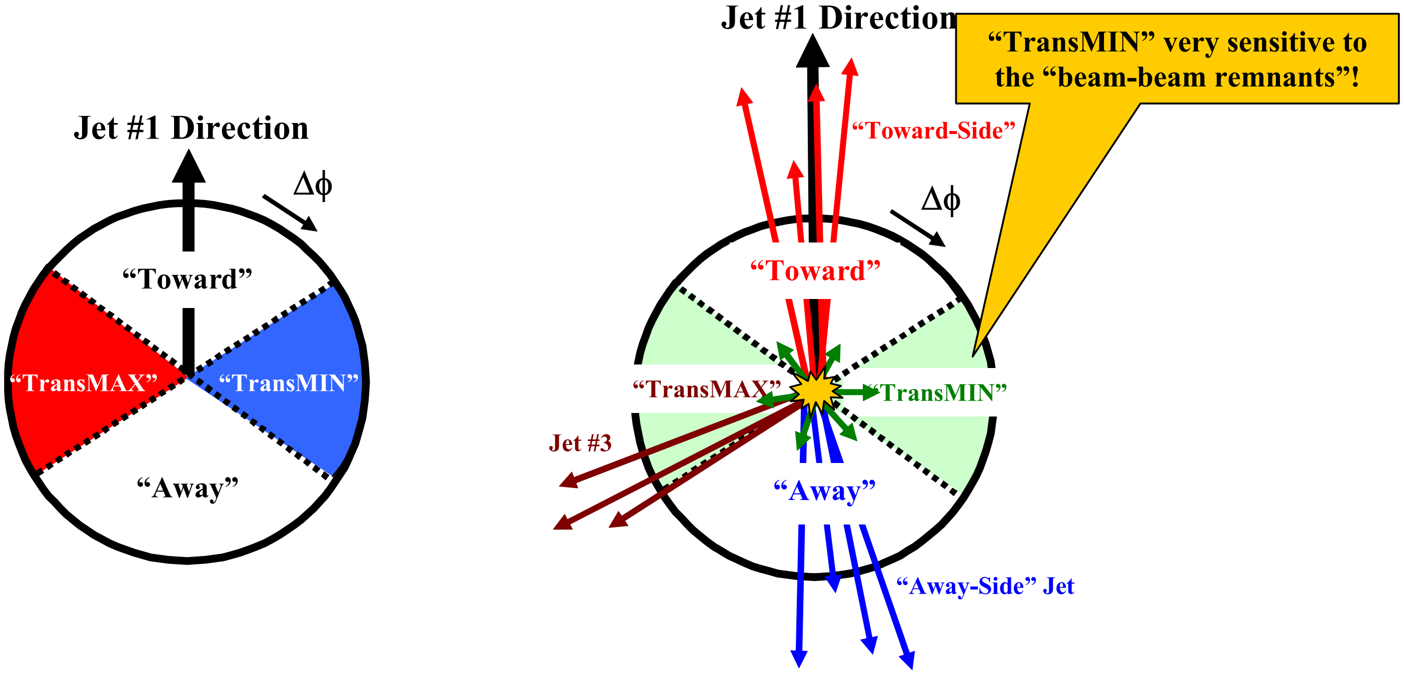


is 3 → 3 instead of 4 → 4:



Underlying event measurements

“MAX/MIN Transverse” Densities

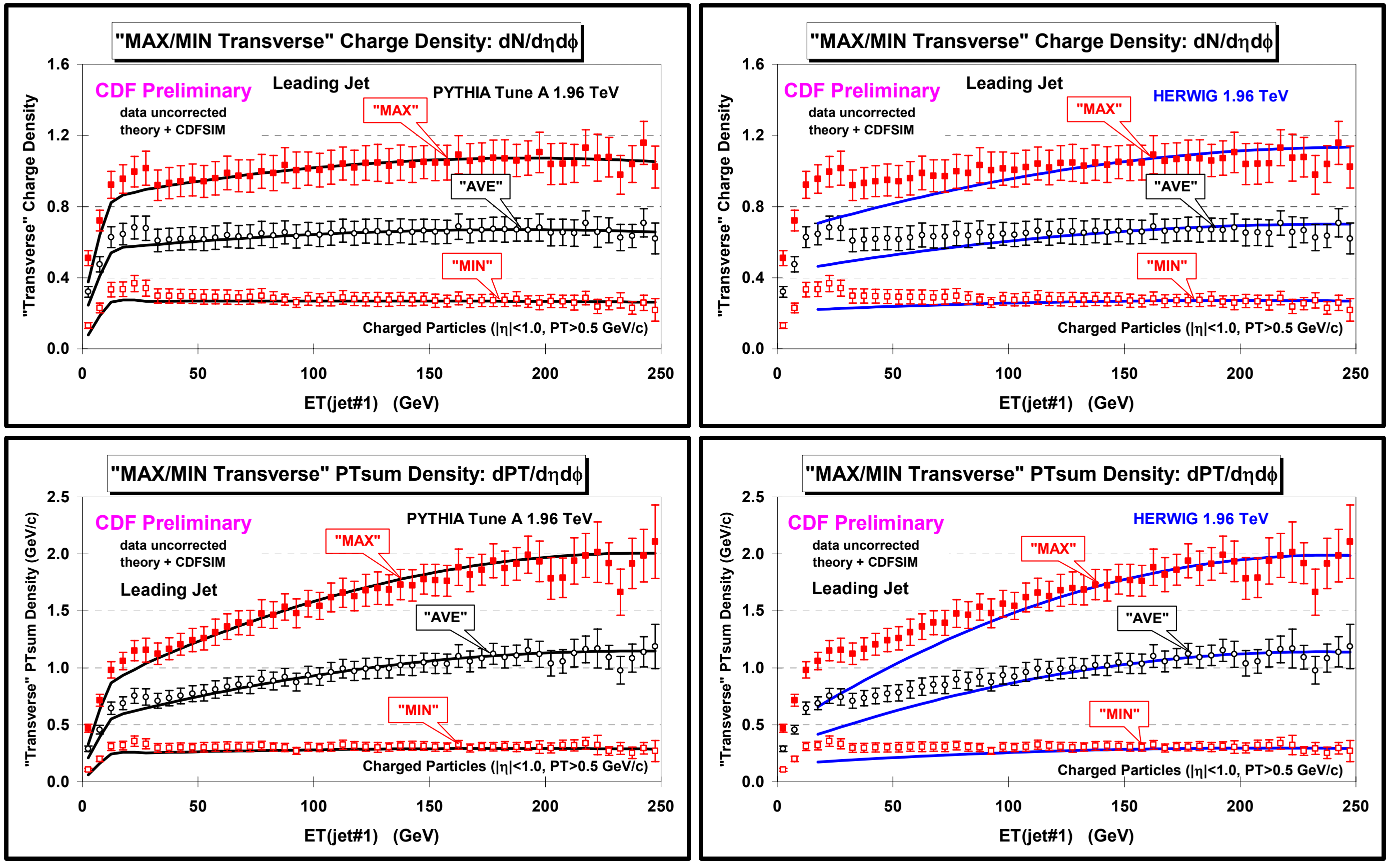


“TransMIN” very sensitive to the “beam-beam remnants”!

- **Define the MAX and MIN “transverse” regions on an event-by-event basis with MAX (MIN) having the largest (smallest) density.**

Underlying event measurements

PYTHIA Tune A **HERWIG**



Charged particle density and PTsum density for “leading jet” events versus E_T(jet#1) for PYTHIA Tune A and HERWIG.

Conclusions on UE/MB

- Despite ~25 year history, multi-parton interaction models are still in their infancy
- LHC experiments'
  - step up in energy
  - high efficiency, purity and phase space coverage
  - emphasis on physical definition of observables

  have given us a huge amount of *useful* data
- existing models describe data well with tuning
- need more understanding of correlations/corners of phase space/relations between different model components

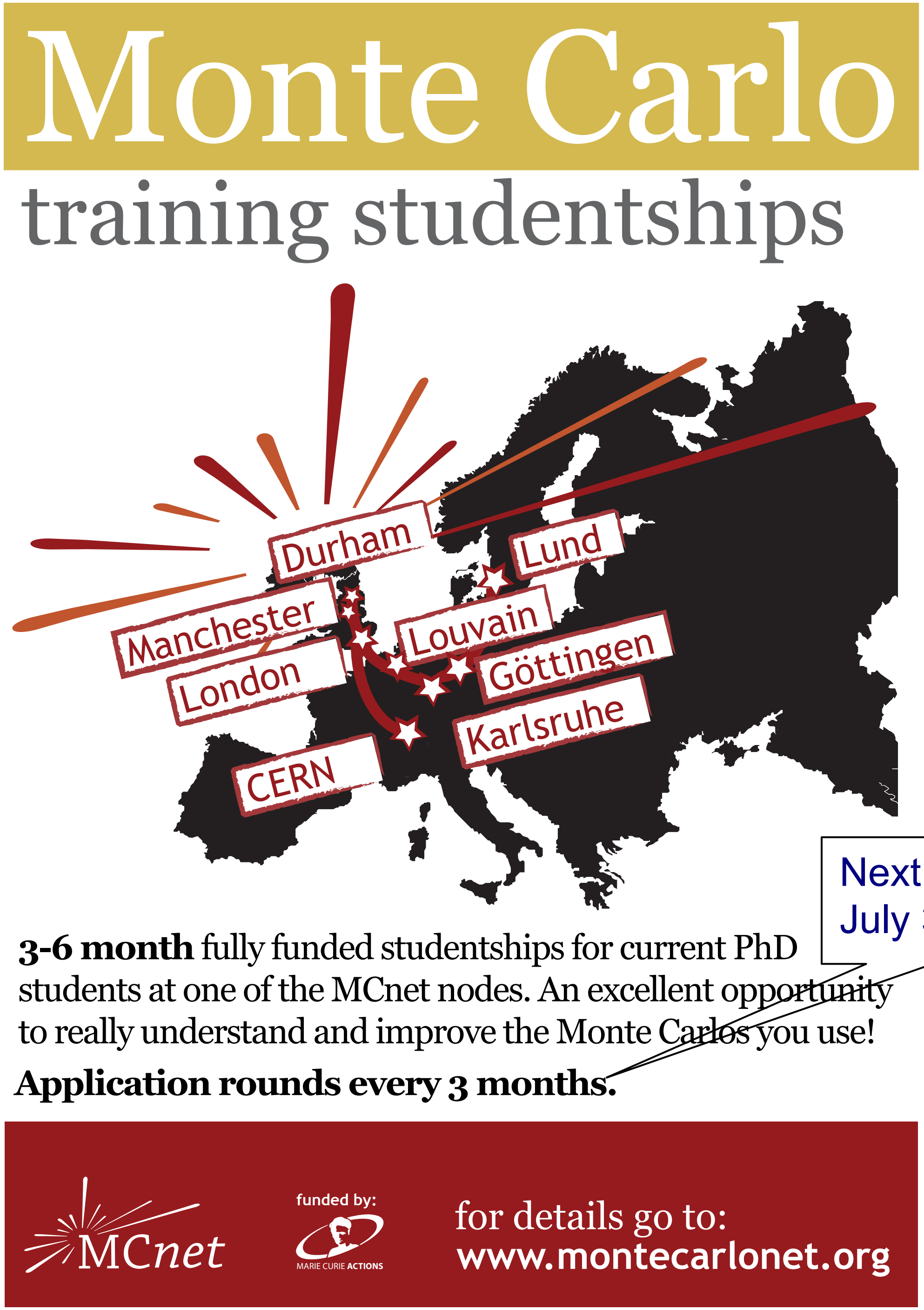
Conclusions on UE/MB

- don’t forget that jet corrections depend on correlations and high moments of distributions and are physics-process dependent

Summary

- Hard Process is very well understood: firm perturbative basis
- Parton Shower is fairly well understood: perturbative basis, with various approximations
- Hadronization is less well understood: modelled, but well constrained by data. Extrapolation to LHC ~ reliable.
- Underlying event least understood: modelled and only weakly constrained by existing data. Extrapolation?

- Always ask "What physics is dominating my effect?"



Monte Carlo

training studentships

3-6 month fully funded studentships for current PhD students at one of the MCnet nodes. An excellent opportunity to really understand and improve the Monte Carlos you use!

Application rounds every 3 months.

MCnet

funded by: MARIE CURIE ACTIONS

for details go to:
www.montecarlonet.org

Next closing:
July 31st

MC*net* projects
- Pythia
- Herwig
- Sherpa
- MadGraph
- Ariadne
- CEDAR

Industry:
- blue yonder
- d-fine
- IBA

Outreach:
- LHC@home