Durham University

Monte Carlo Event Generators

Lecture 3: Hadronization and Underlying Event Modelling

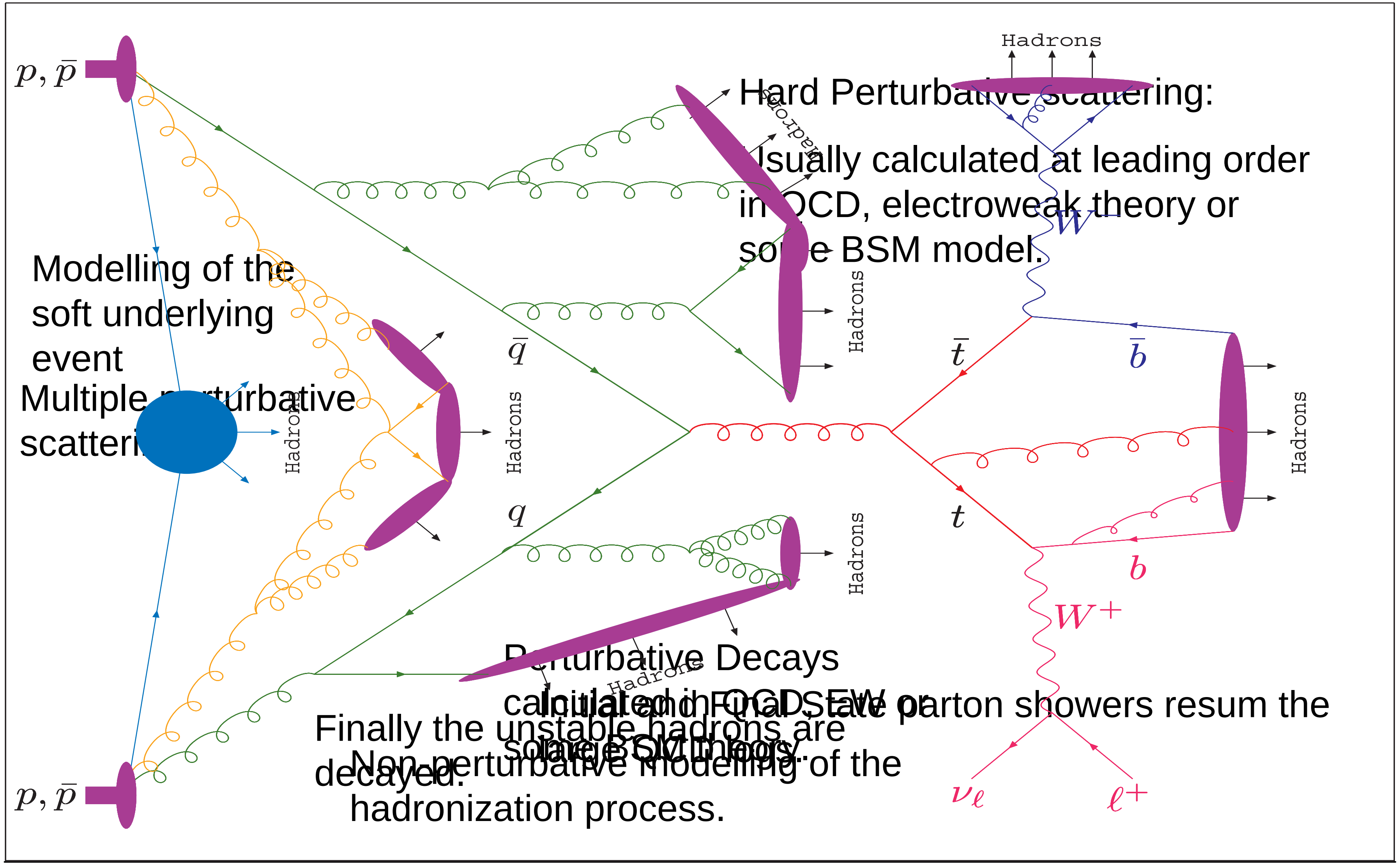
Peter Richardson

IPPP, Durham University

Plan

- Lecture 1: Introduction
  - Basic principles of event generation
  - Monte Carlo integration techniques
  - Matrix Elements
- Lecture 2: Parton Showers
  - Parton Shower Approach
  - Recent advances, CKKW and MC@NLO
- Lecture 3: Hadronization and Underlying Event
  - Hadronization Models
  - Underlying Event Modelling

A Monte Carlo Event



Lecture 3

Today we will cover

- Hadronization Models
  - Independent Fragmentation
  - Lund String Model
  - Cluster Model
- Underlying Event
  - Soft Models
  - Multiple Scattering
- Conclusions

Introduction

- Partons aren't physical particles: they can't propagate freely.
- We therefore need to describe the transition of the quarks and gluons in our perturbative calculations into the hadrons which can propagate freely.
- We need a phenomenological model of this process.
- There are three models which are commonly used.
  - Independent Fragmentation
  - Lund String Model
  - Cluster Model

Independent Fragmentation Model

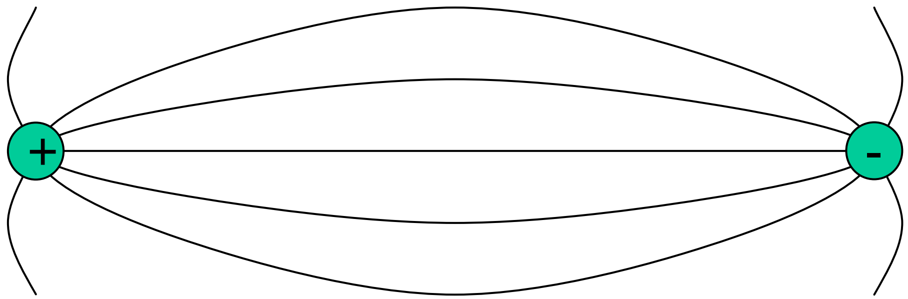
"Feynman-Field"

- The longitudinal momentum distribution is given by an arbitrary fragmentation function which is a parameterization of data.
- The transverse momentum distribution is Gaussian.
- The algorithm recursively splits $q \rightarrow q'$+hadron.
- The remaining soft quark and antiquark are connected at the end.
- The model has a number of flaws
  - Strongly frame dependent
  - No obvious relation with the perturbative physics.
  - Not infrared safe
  - Not a confinement model
  - Wrong energy dependence.

Confinement

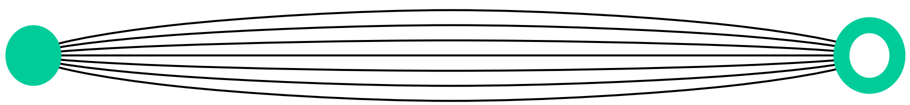
- We know that at small distances we have asymptotic freedom and the force between a quark-antiquark pair is like that between an e^+e^- pair.

QED



- But at long distances the self interactions of the gluons make the field lines attract each other.

QCD



- A linear potential at long distances and confinement.

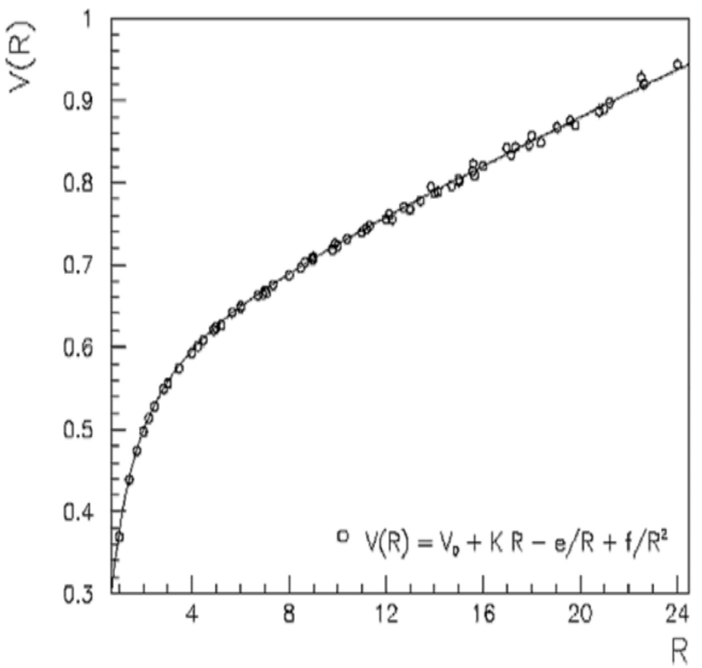
Lund String Model

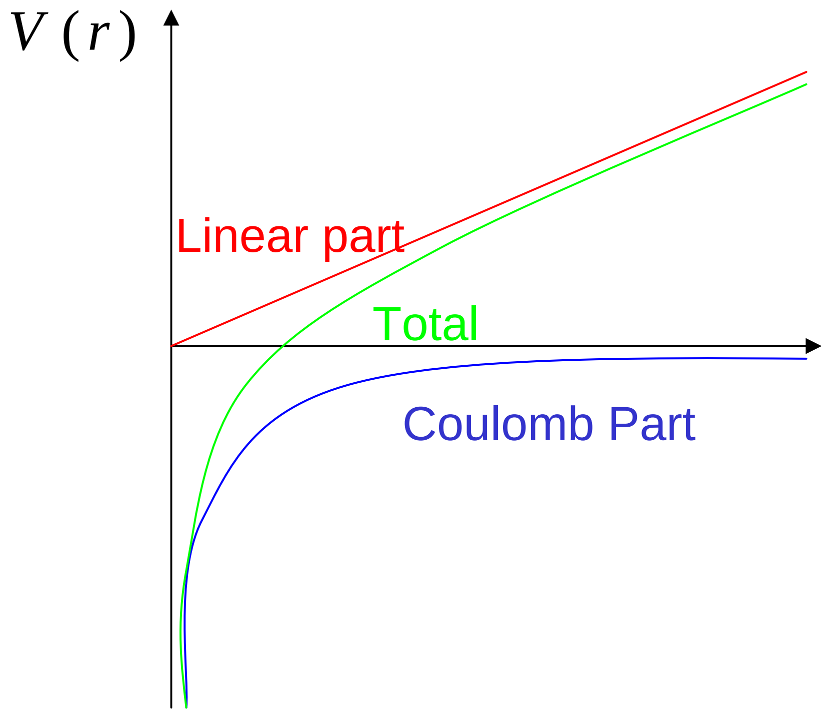
- In QCD the field lines seem to be compress into a tube-like region, looks like a string.
- So we have linear confinement with a string tension.

$$F(r) \approx \text{const} = \kappa \approx 1\,\text{GeV/fm}$$

- Separate the transverse and longitudinal degrees of freedom gives a simple description as a 1+1 dimensional object, the string, with a Lorentz invariant formalism.

Intraquark Potential



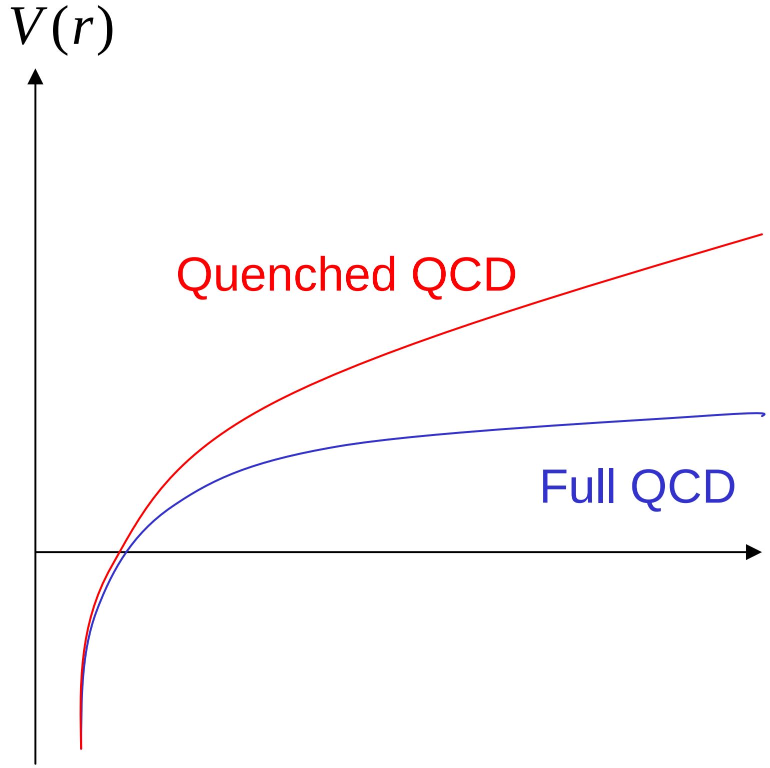


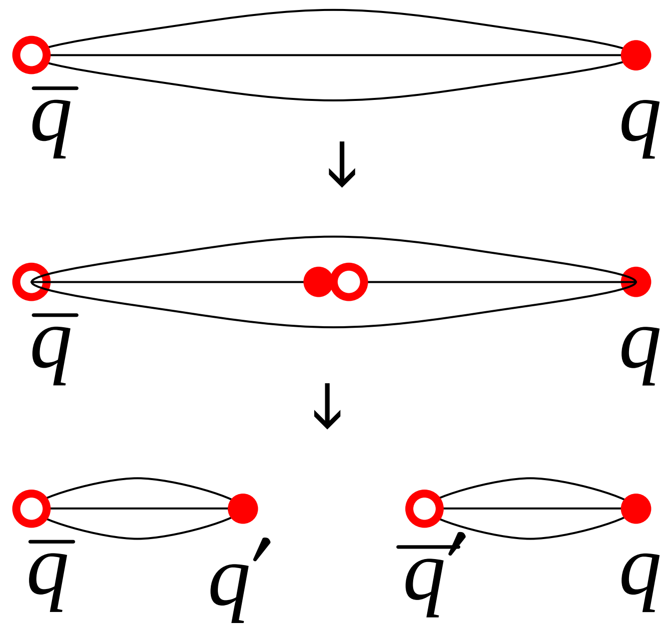
$$V(r) = -\frac{4}{3}\frac{\alpha_s}{r} + \kappa r \approx -\frac{0.13}{r} + r \quad (\text{for } \alpha_s \approx 0.5, \text{r in fm and V in GeV})$$

- The coulomb piece is important for the internal structure of the hadrons but not for particle production.

Intraquark potential

- In the real world the string can break non-perturbatively by producing quark-antiquark pairs in the intense colour field.

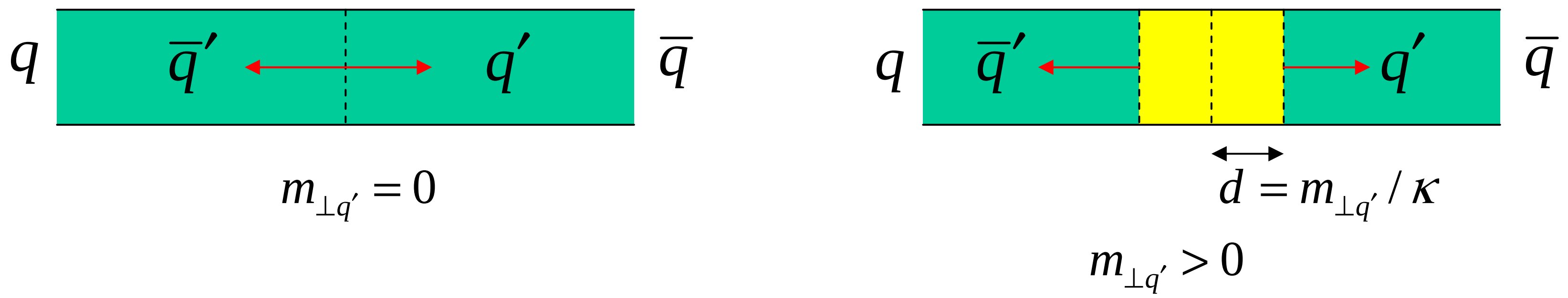




- This is the basic physics idea behind the string model and is very physically appealing.

Lund String Model

- If we start by ignoring gluon radiation and consider e^+e^- annihilation.
- We can consider this to be a point-like source of quark-antiquark pairs.
- In the intense chromomagnetic field of the string $q\bar{q}$ pairs are created by tunnelling.



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

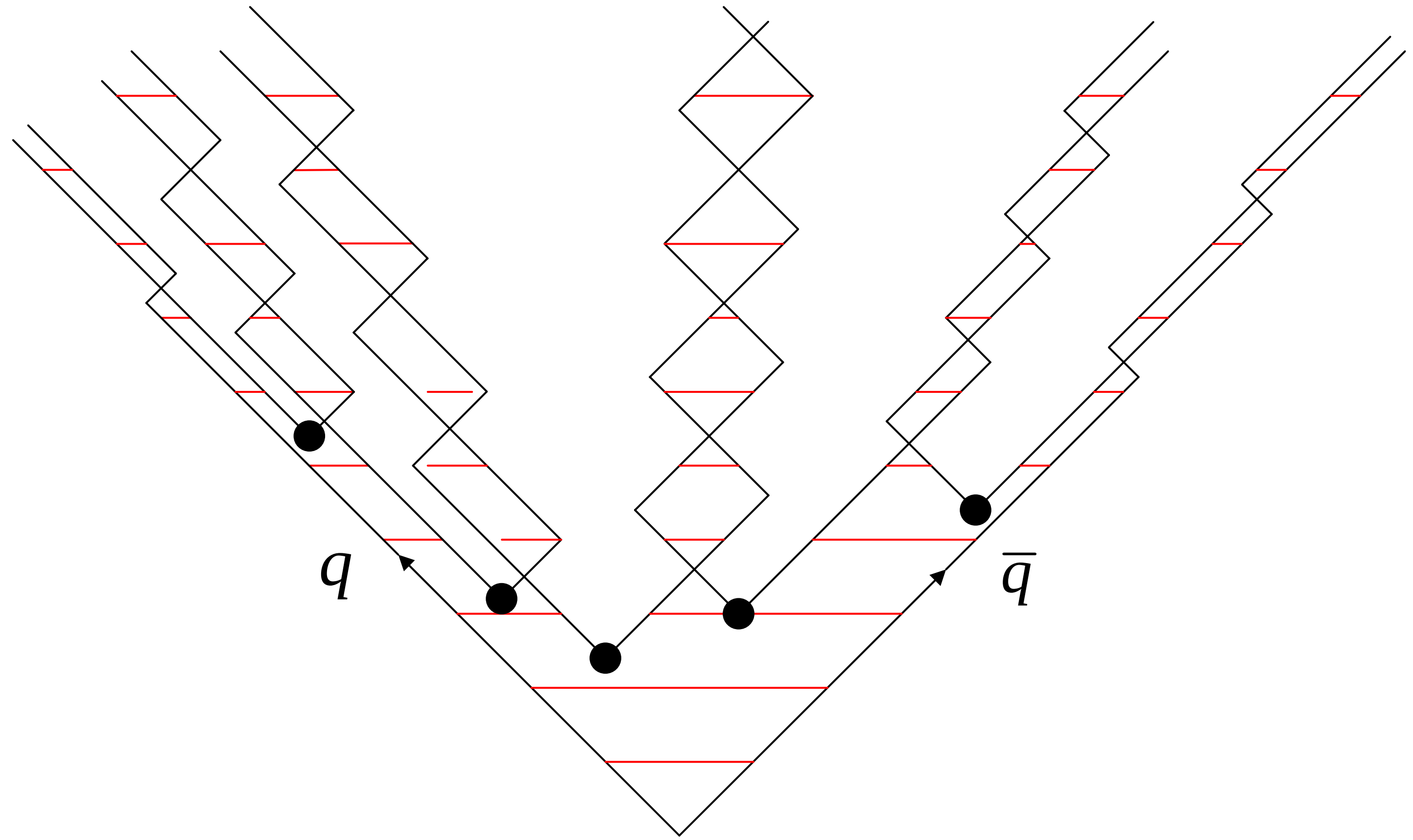
Lund String Model

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

- This gives
  1) Common Gaussian p_T spectrum.
  2) Suppression of heavy quark production $u\bar{u} : d\bar{d} : s\bar{s} : c\bar{c} \approx 1:1:0.3:10^{-11}$
  3) Diquark-antiquark production gives a simple model of baryon production.
- In practice the hadron composition also depends on
  - Spin probabilities
  - Hadronic wave functions
  - Phase space
  - More complicated baryon production models
- Gives many parameters which must be tuned to data.

Lund String Model

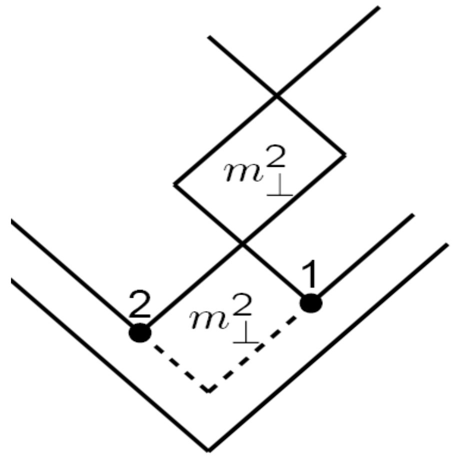
- Motion of quarks and antiquarks in a $q\overline{q}$ system.



- Gives a simple but powerful picture of hadron production

Lund String Model

- The string picture constrains the fragmentation function
  - Lorentz Invariance
  - Acausality
  - Left-Right Symmetry $P(1,2) = P(1)\times P(1\rightarrow 2) = P(2)\times P(2\rightarrow 1)$



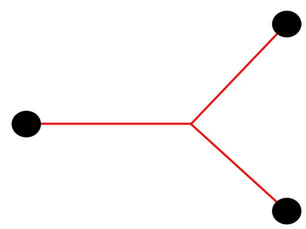
- Give the Lund symmetric fragmentation function.

$$f(z) \propto \frac{(1-z)^a}{z}\exp\left(\frac{-bm_\perp^2}{z}\right)$$

- a, b and the quark masses are the main tuneable parameters of the model.

Baryon Production

- In all hadronization models baryon production is a problem.
- Earliest approaches were to produce diquark-antidiquark pairs in the same way as for quarks.
- Later in a string model baryons pictured as three quarks attached to a common centre.

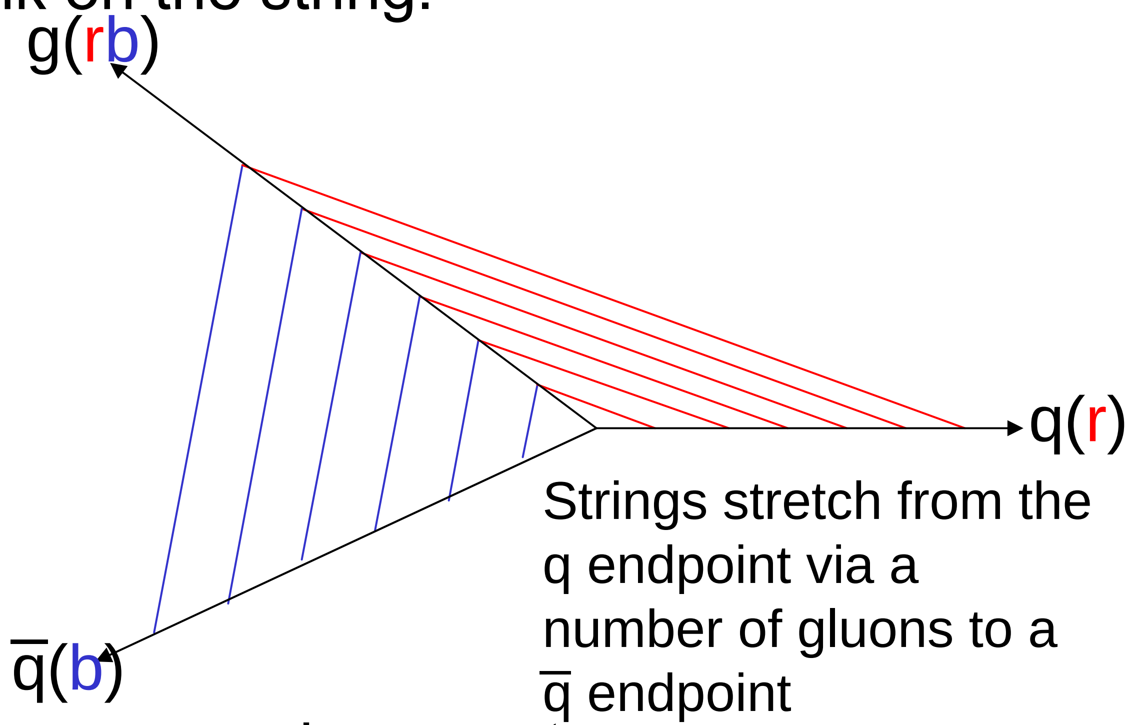


- At large separation this can be considered as two quarks tightly bound into a diquark.

- Two quarks can tunnel nearby in phase space: baryon-antibaryon pair.
- Extra adjustable parameter for each diquark.

Three Jet Events

- Gluons give a kink on the string.



- The kink carries energy and momentum.
- There are no new parameters for gluon jets.
- Few parameters to describe the energy-momentum structure.
- Many parameters for the flavour composition.

Summary of the String Model

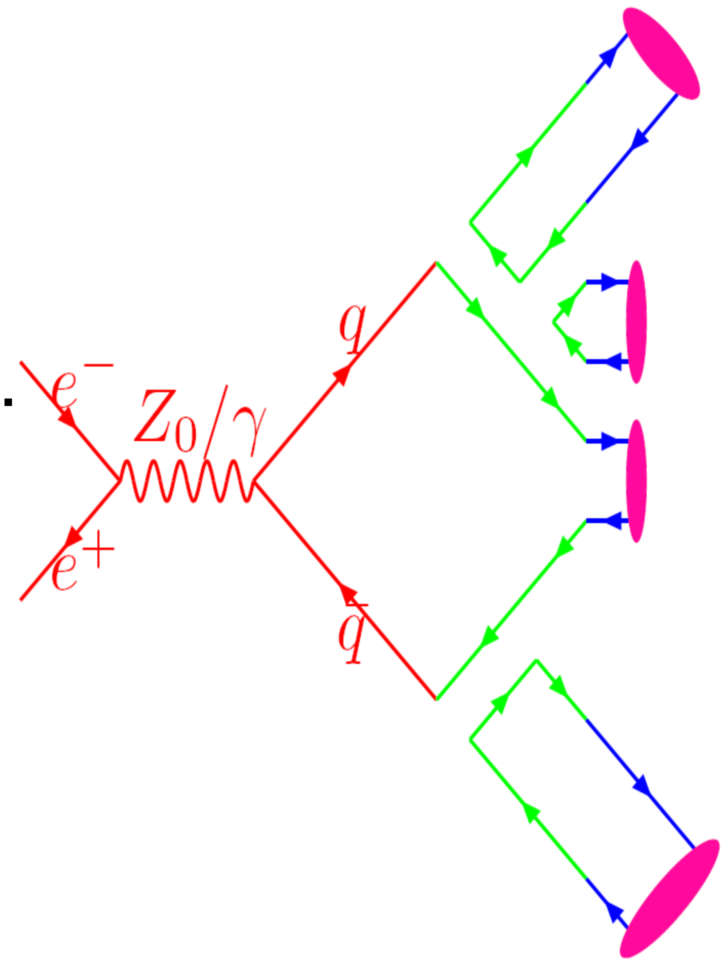
- The string model is strongly physically motivated and intuitively compelling.
- Very successful fit to data.
- Universal, after fitting to e^+e^- data little freedom elsewhere.
- But
  - Has many free parameters, particularly for the flavour sector.
  - Washes out too much perturbative information.
- Is it possible to get by with a simpler model?

Preconfinement

- In the planar approximation, large number of colours limit

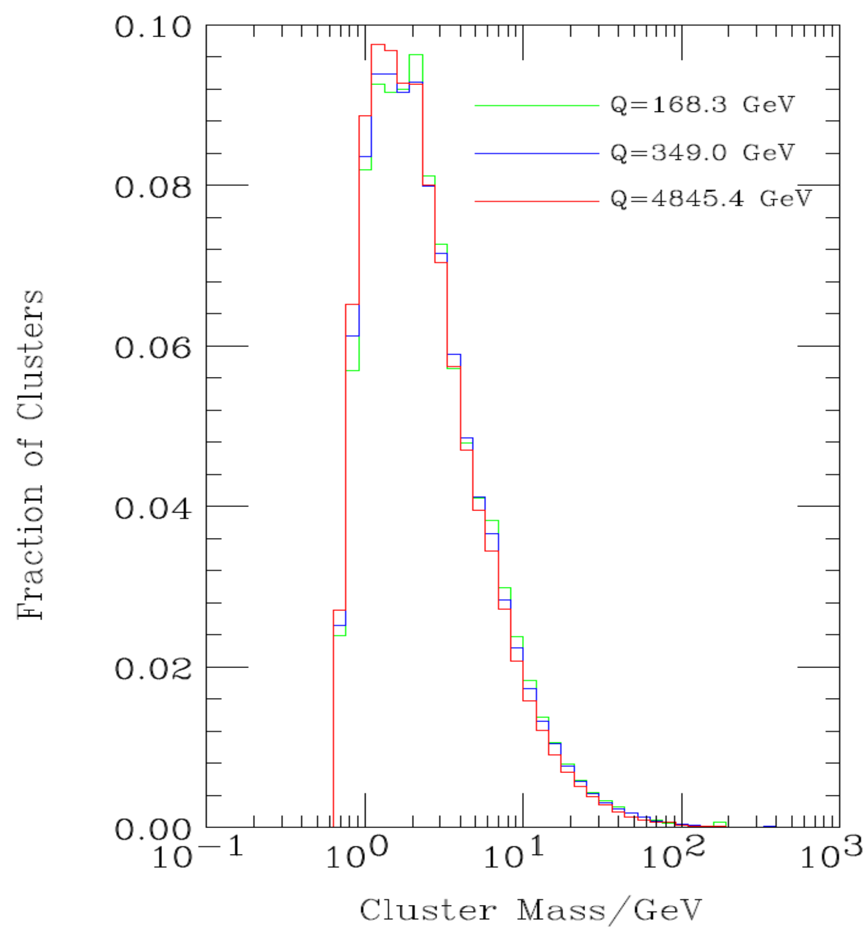
  Gluon= colour-anticolour pair

- We can follow the colour structure of the parton shower.
- At the end colour singlet pairs end up close in phase space.
- Non-perturbatively split the gluons into quark-antiquark pairs.



Preconfinement

- The mass spectrum of colour-singlet pairs is asymptotically independent of energy and the production mechanism.
- It peaks at low mass, of order the cut-off Q_0.



The Cluster Model

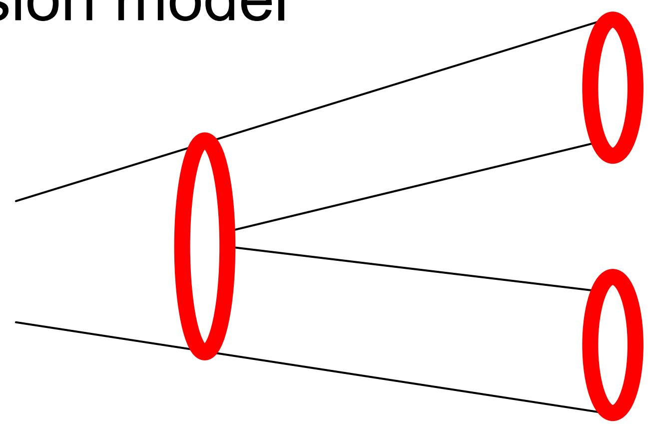
- Project the colour singlet clusters onto the continuum of high-mass mesonic resonances (=clusters).
- Decay to lighter well-known resonances and stable hadrons using
  - Pure 2 body phase-space decay and phase space weight

$$(2s_1+1)(2s_2+1)\frac{2p^*}{m}$$

- The hadron-level properties are fully determined by the cluster mass spectrum, i.e. by the properties of the parton shower.
- The cut-off Q_0 is the crucial parameter of the model.

The Cluster Model

- Although the cluster spectrum peaks at small masses there is a large tail at high mass.
- For this small fraction of high mass clusters isotropic two-body is not a good approximation.
- Need to split these clusters into lighter clusters using a longitudinal cluster fission model



- This model
  – Is quite string-like
  – Fission threshold is a crucial parameter
  – ~15% of clusters get split but ~50% of hadrons come from them

The Cluster Model: Problems

1) Leading Hadrons are too soft
   - Perturbative quarks remember their direction

$$P(\theta^2) \sim \exp\left(\frac{-\theta^2}{2\theta_0^2}\right)$$

   - Rather string like
   - Extra adjustable parameter

2) Charm and Bottom spectra too soft
   - Allow cluster decays into one meson for heavy quark clusters.
   - Make cluster splitting parameters flavour dependent.
   - More parameters

The Cluster Model: Problems

3) Problems with baryon production
4) Some problems with charge correlations.
5) Sensitive to the particle content.
    - Only include complete multiplets.

The 'Beliefs'

- There are two main schools of thought in the event generator community.

PYTHIA

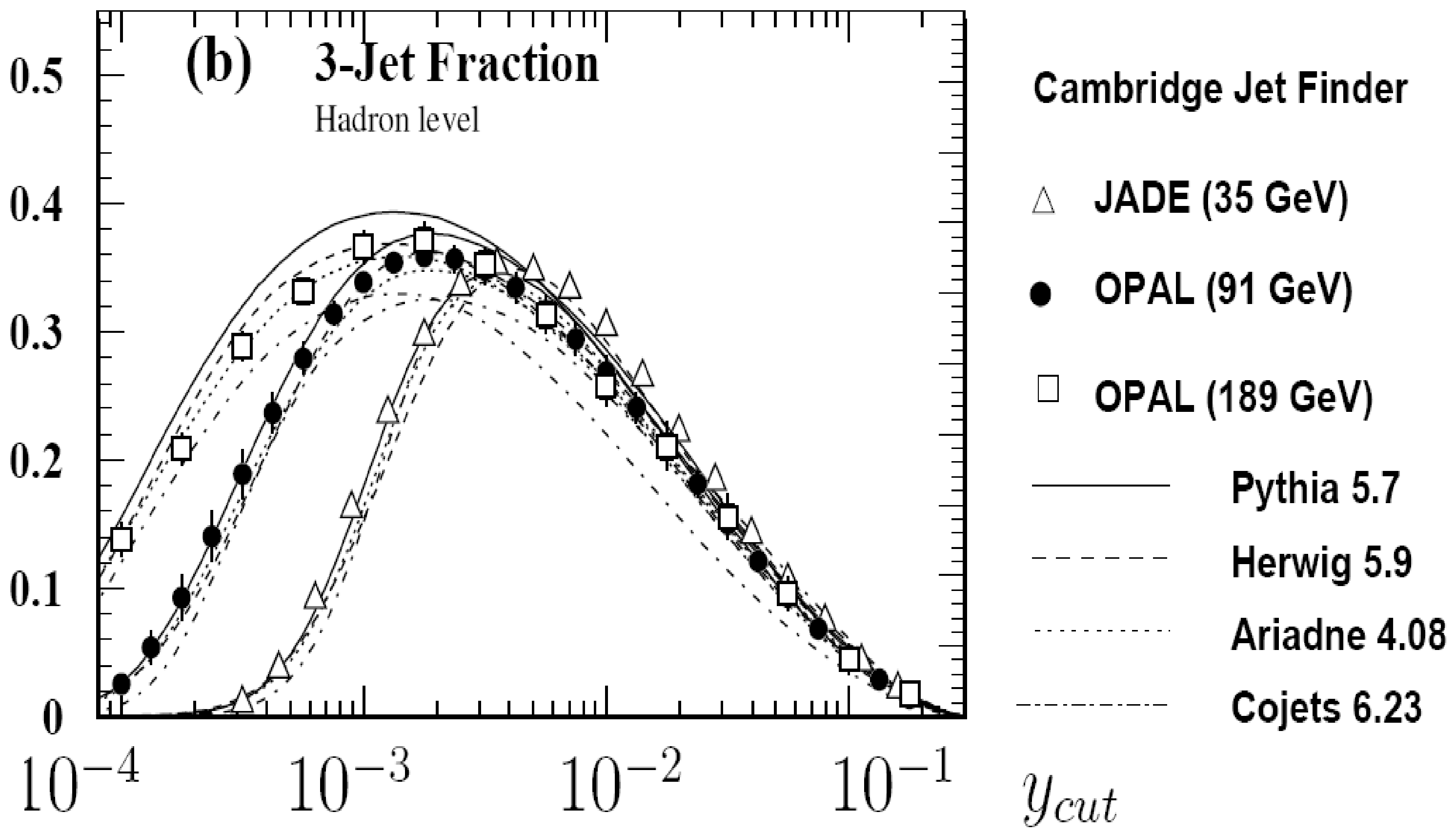
- "Hadrons are produced by hadronization. You must get the non-perturbative dynamics right."
- Better data has required improvements to the perturbative simulation.

HERWIG

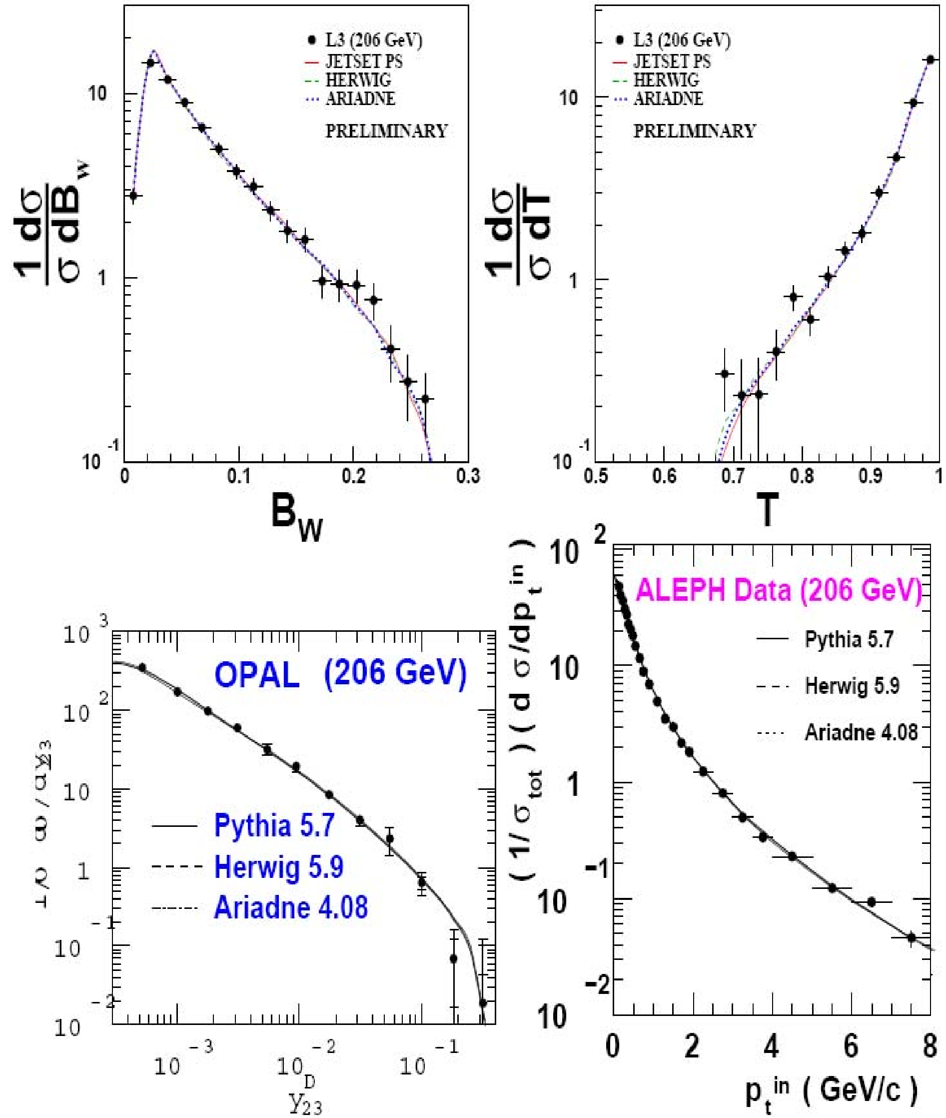
- "Get the perturbative physics right and any hadronization model will be good enough"
- Better data has required changes to the cluster model to make it more string-like

- There ain't no such thing as a good parameter-free description.

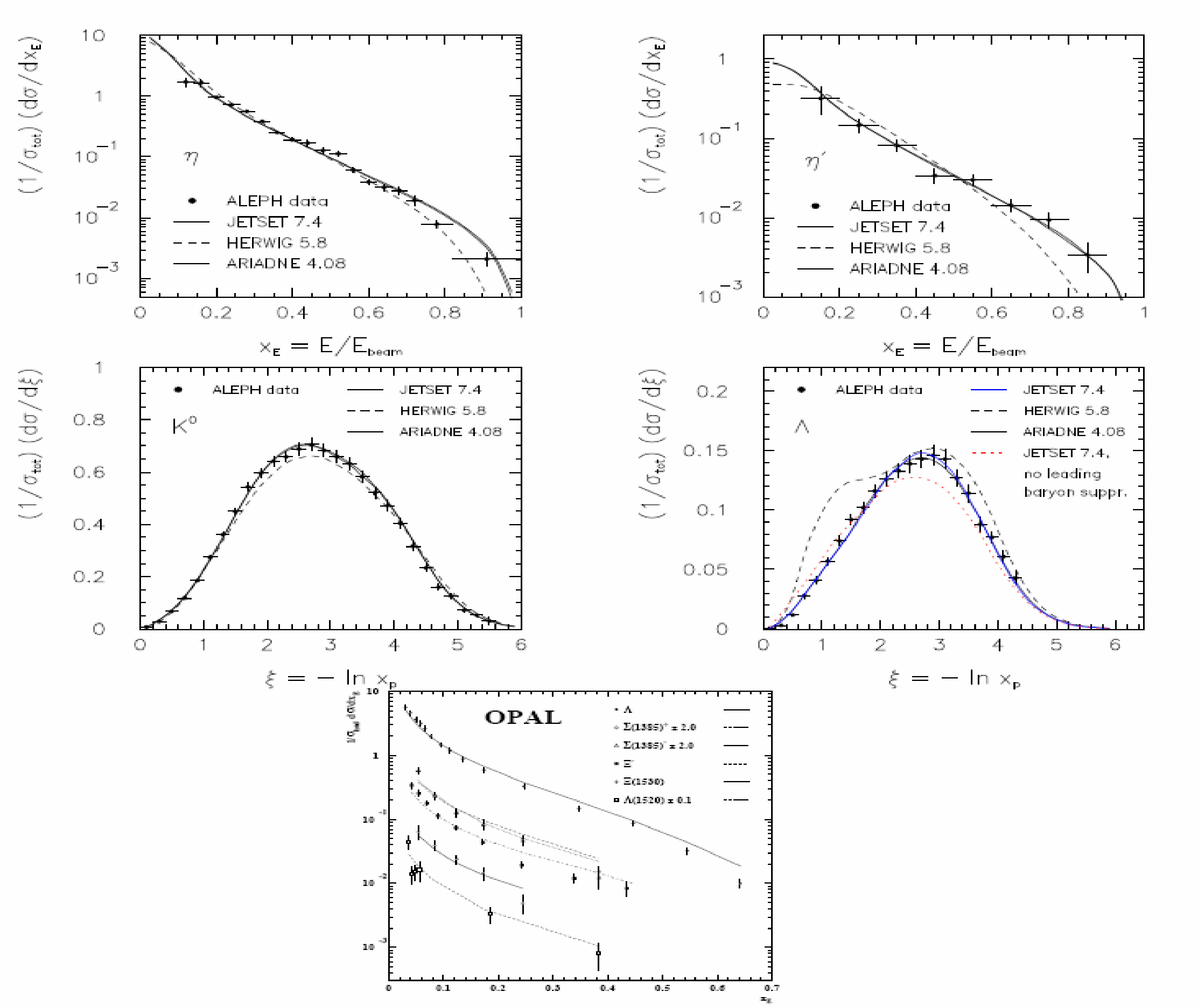
Energy Dependence



Event Shapes



Identified Particle Spectra

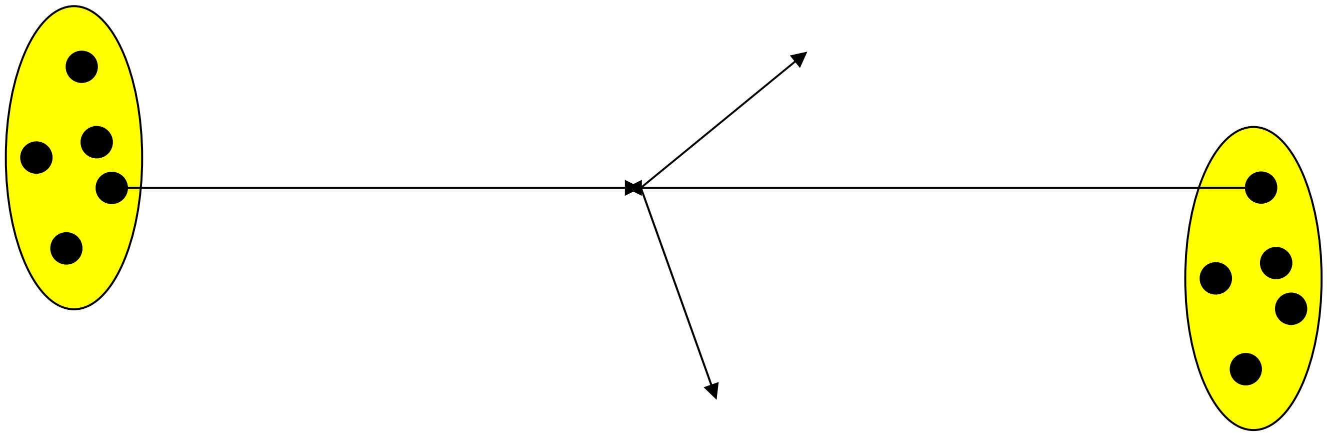


The facts?

- Independent fragmentation doesn’t describe the data, in particular the energy dependence.
- All the generators give good agreement for event shapes
- HERWIG has less parameters to tune the flavour composition and tends to be worse for identified particle spectra.

The Underlying Event

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



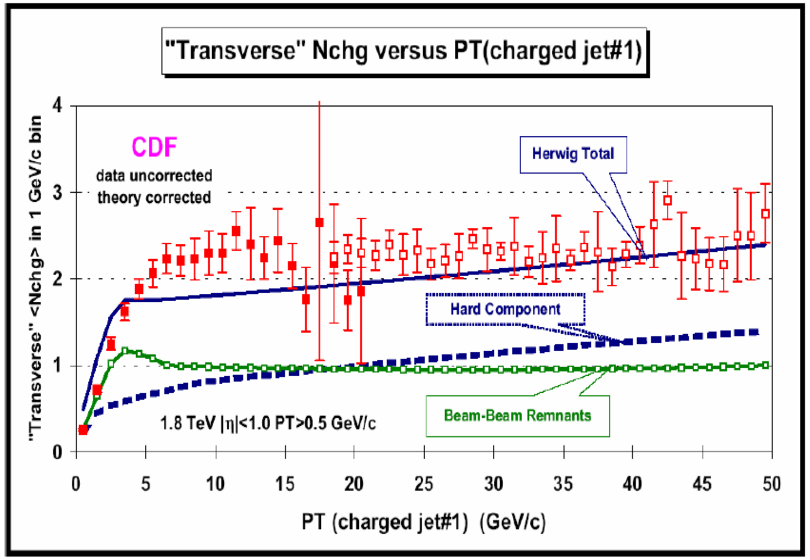
Two Types of Model:

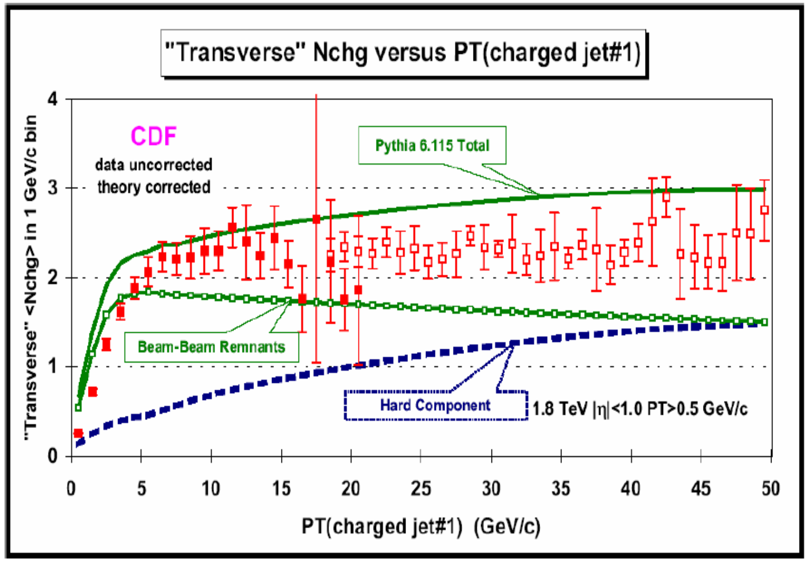
1) Non-Perturbative: Soft parton-parton cross section is so large that the remnants always undergo a soft collision.

2) Perturbative: 'Hard' parton-parton cross section is huge at low p_T, dominates the inelastic cross section and is calculable.

Minimum Bias and Underlying Event

- Not everyone means the same thing by “underlying event”
- The separation of the physics into the components of a model is of course dependent on the model.
- Minimum bias tends to mean all the events in hadron collisions apart from diffractive processes.
- Underlying event tends to mean everything in the event apart from the collision we are interested in.





Soft Underlying Event Models

- There are essentially two types of model
- Pomeron Based
  - Based on the traditional of soft physics of "cut Pomerons" for the $p_T \to 0$ limit of multiple interactions.
  - Used in ISAJET, Phojet/DTUJet
- UA5 Parameterization
  - A parameterization of the UA5 experimental data on minimum bias collisions.
  - Used in HERWIG

UA5 Model

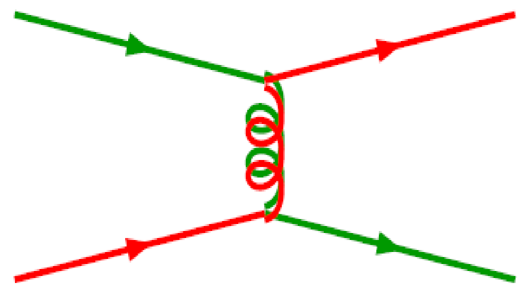
- UA5 was a CERN experiment to measure minimum bias events.
- The UA5 model is then a simple phase-space model intended to fit the data.
  - Distribute a number of clusters independently in rapidity and transverse momentum according to a negative binomial.
  - Conserve overall energy and momentum and flavour.
- Main problem is that there is no high p_T component and the only correlations are due to cluster decays.

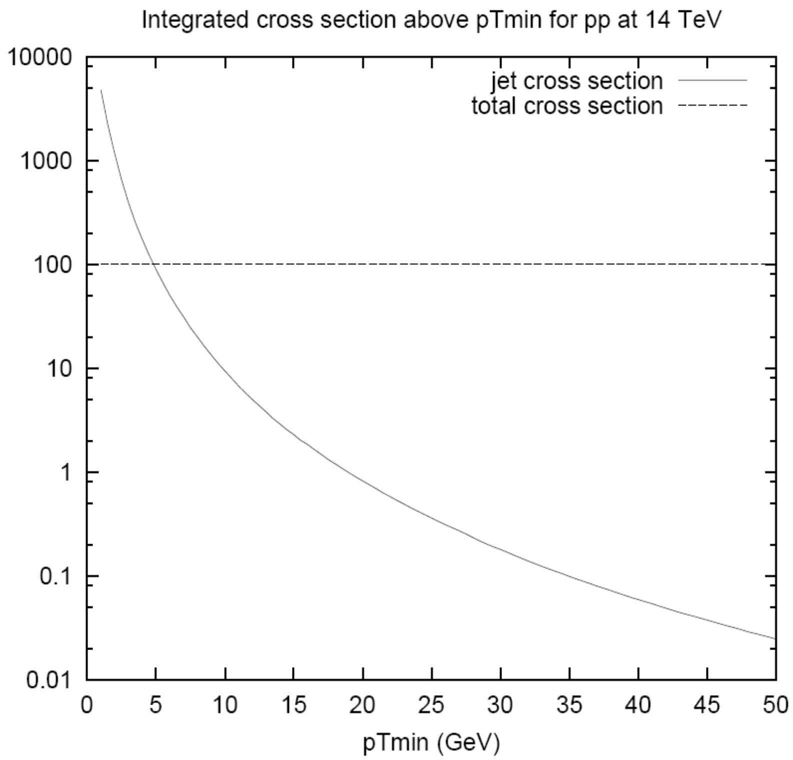
Multiparton Interaction Models

- The cross-section for 2→2 scattering is dominated by t-channel gluon exchange.
- It diverges like

$$\frac{\mathrm{d}\sigma}{\mathrm{d}p_\perp^2} \approx \frac{1}{p_\perp^4} \quad \text{for} \quad p_\perp \to 0$$

- This must be regulated used a cut of p_{Tmin}.
- For small values of p_{Tmin} this is larger than the total hadron-hadron cross section.
- More than one parton-parton scattering per hadron collision



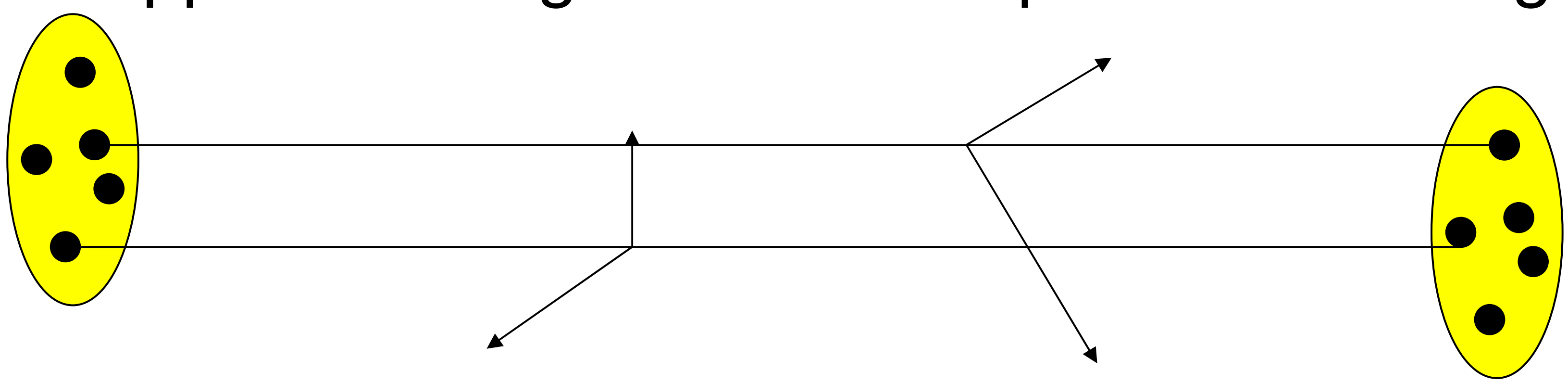


Multiparton Interaction Models

- If the interactions occur independently then follow Poissonian statistics

$$P_n = \frac{\langle n \rangle^n}{n!} e^{-\langle n \rangle}$$

- However energy-momentum conservation tends to suppressed large numbers of parton scatterings.



- Also need a model of the spatial distribution of partons within the proton.

Multiparton Interaction Models

- In general there are two options for regulating the cross section.

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

$$\text{or} \to \frac{\alpha_s^2(p_\perp^2 + p_{\perp 0}^2)}{(p_\perp^2 + p_{\perp 0}^2)^2} \qquad \text{more complicated}$$

where $p_{\perp\min}$ or $p_{\perp 0}$ are free parameters of order 2 GeV.

- Typically 2-3 interactions per event at the Tevatron and 4-5 at the LHC.
- However tends to be more in the events with interesting high p_T ones.

Simple Model

- T. Sjostrand, M. van Zijl, PRD36 (1987) 2019.
- Sharp cut-off at p_{Tmin} is the main free parameter.
- Doesn't include diffractive events.
- Average number of interactions is

$$\langle n \rangle = \sigma_{\mathrm{int}}(p_{\perp\min}) / \sigma_{\mathrm{non-diffractive}}$$

- Interactions occur almost independently, i.e. Poisson $P_n = \langle n \rangle^n e^{-\langle n \rangle} / n!$
- Interactions generated in ordered p_T sequence
- Momentum conservation in PDF's reduces the number of collisions.

More Sophisticated

- Use a smooth turn off at p_{T0}.
- Require at least 1 interaction per event
- Hadrons are extended objects, e.g. double Gaussian ("hot spots"):

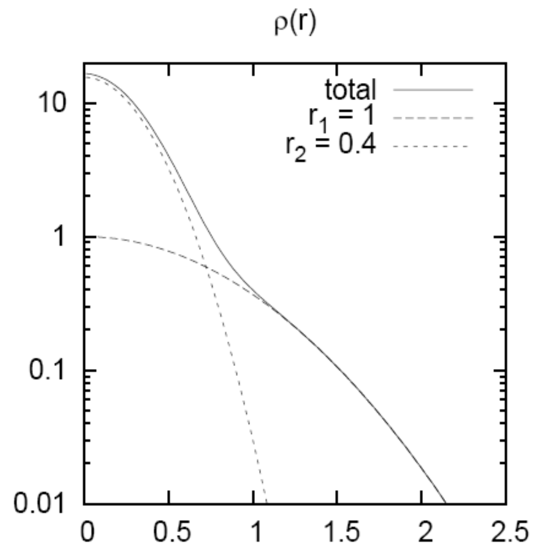
$$\rho_{\text{matter}} = N_1 \exp\left(-\frac{r^2}{r_1^2}\right) + N_2 \exp\left(-\frac{r^2}{r_2^2}\right)$$

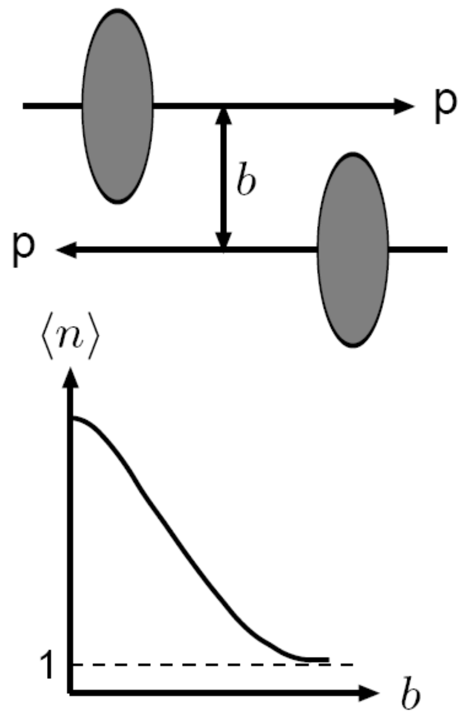
where $r_2 \neq r_1$ represents "hot spots"

- Events are distributed in impact parameter *b*.
- The hadrons overlap during the collision

$$O(b) = \int \mathrm{d}^3x \mathrm{d}t \rho_{1,\text{matter}}^{\text{boosted}}(x,t) \rho_{2,\text{matter}}^{\text{boosted}}(x,t)$$

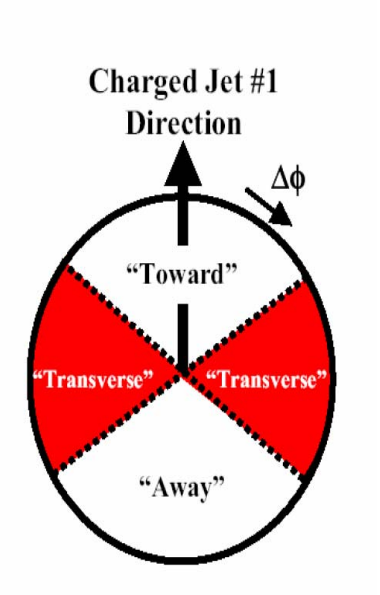
- Average activity at b proportional to O(b).
  - Central collisions normally more active
  - more multiple scattering.

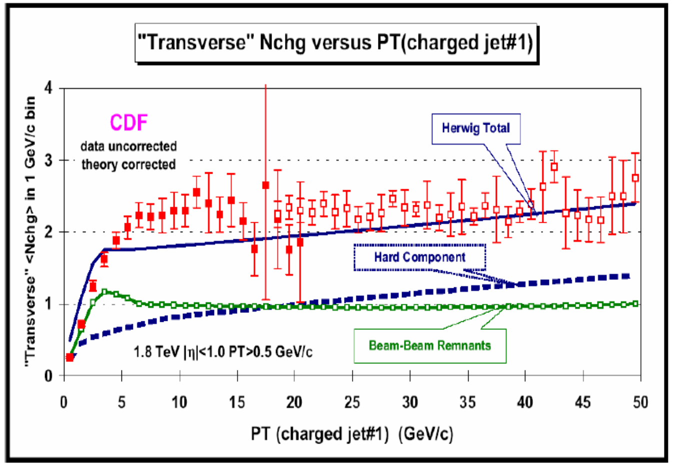


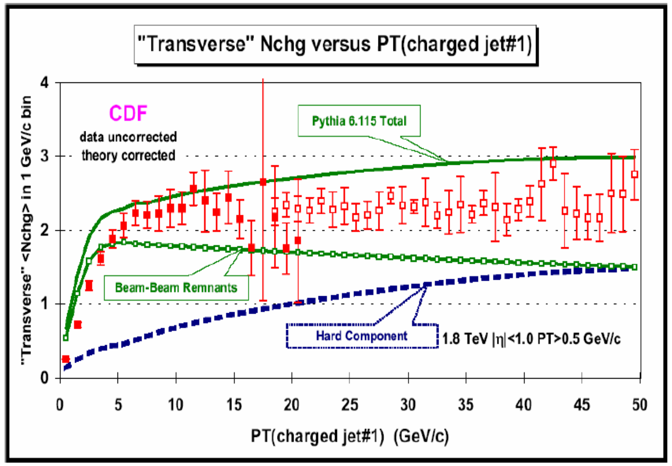


Data

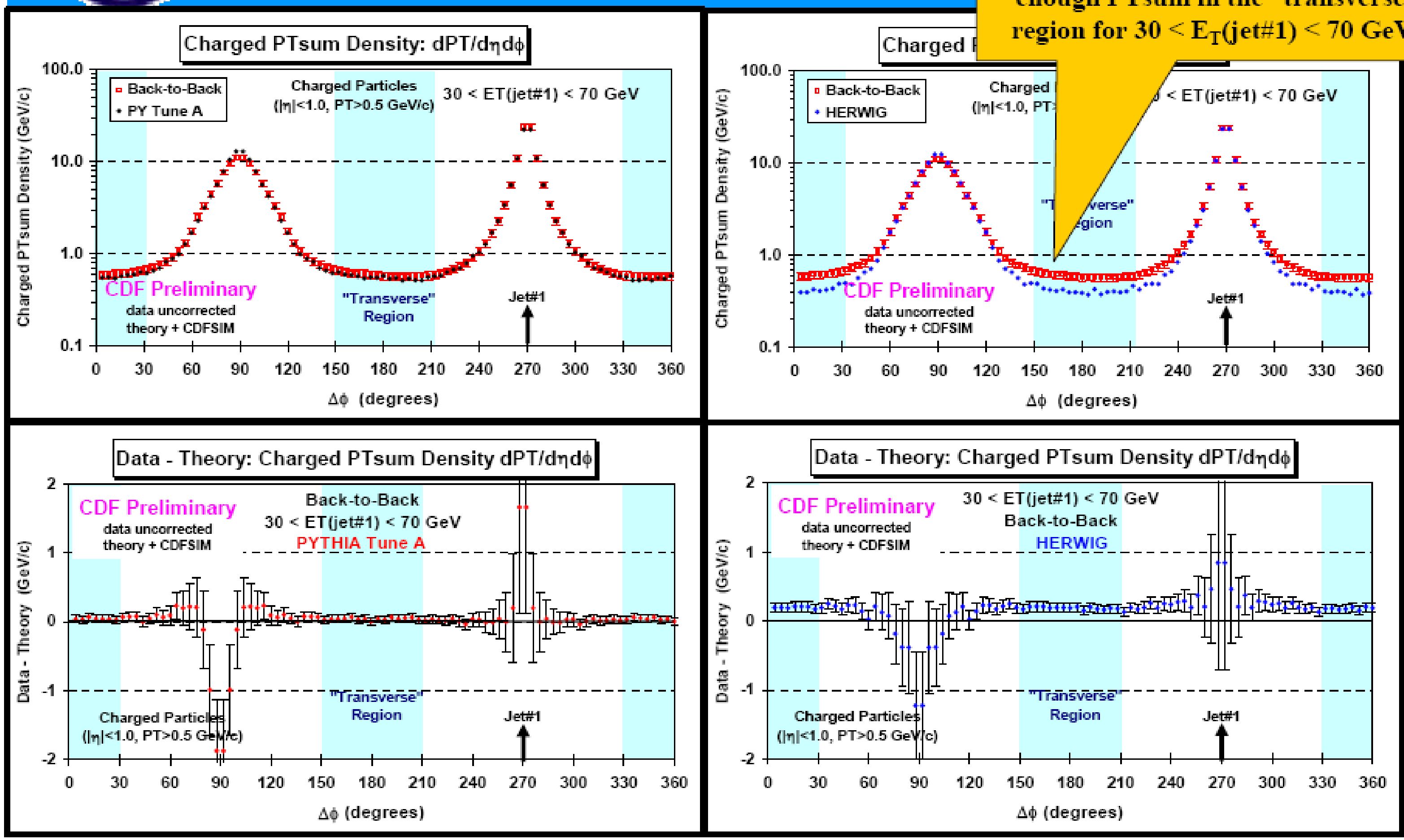
- There has been a lot of work in recent years comparing the models with CDF data by Rick Field.







Charged PTsum Density PYTHIA Tune A vs HERWIG

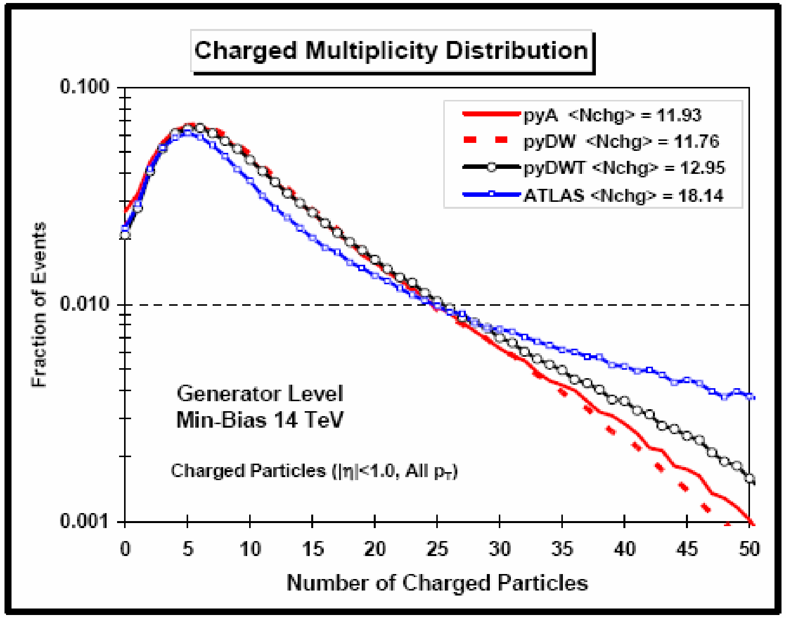


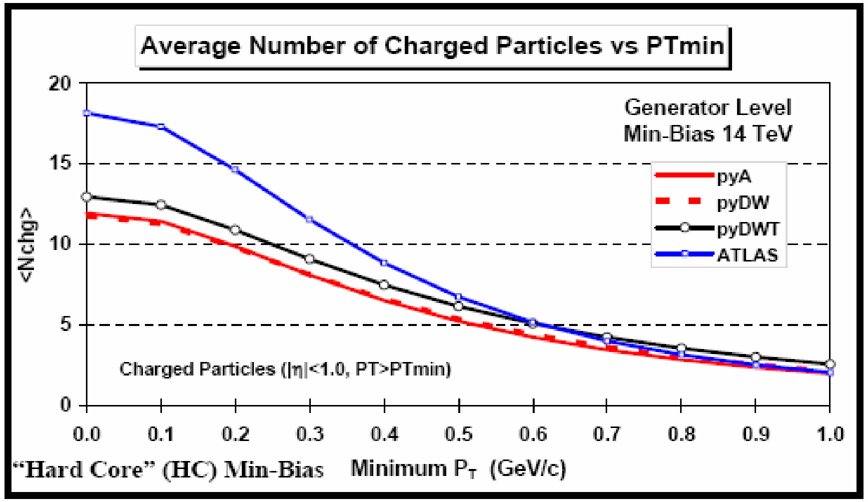
KITP Collider Workshop
February 17, 2004
Rick Field - Florida/CDF
Page 28

Underlying Event

- There is strong evidence for multiple interactions.
- In general the PYTHIA model (Tune A) gives the best agreement with data although there has been less work tuning the HERWIG multiple scattering model JIMMY(although there seem to be problems getting agreement with data).
- However taking tunes which agree with the Tevatron data and extrapolating to the LHC gives a wide range of predictions.

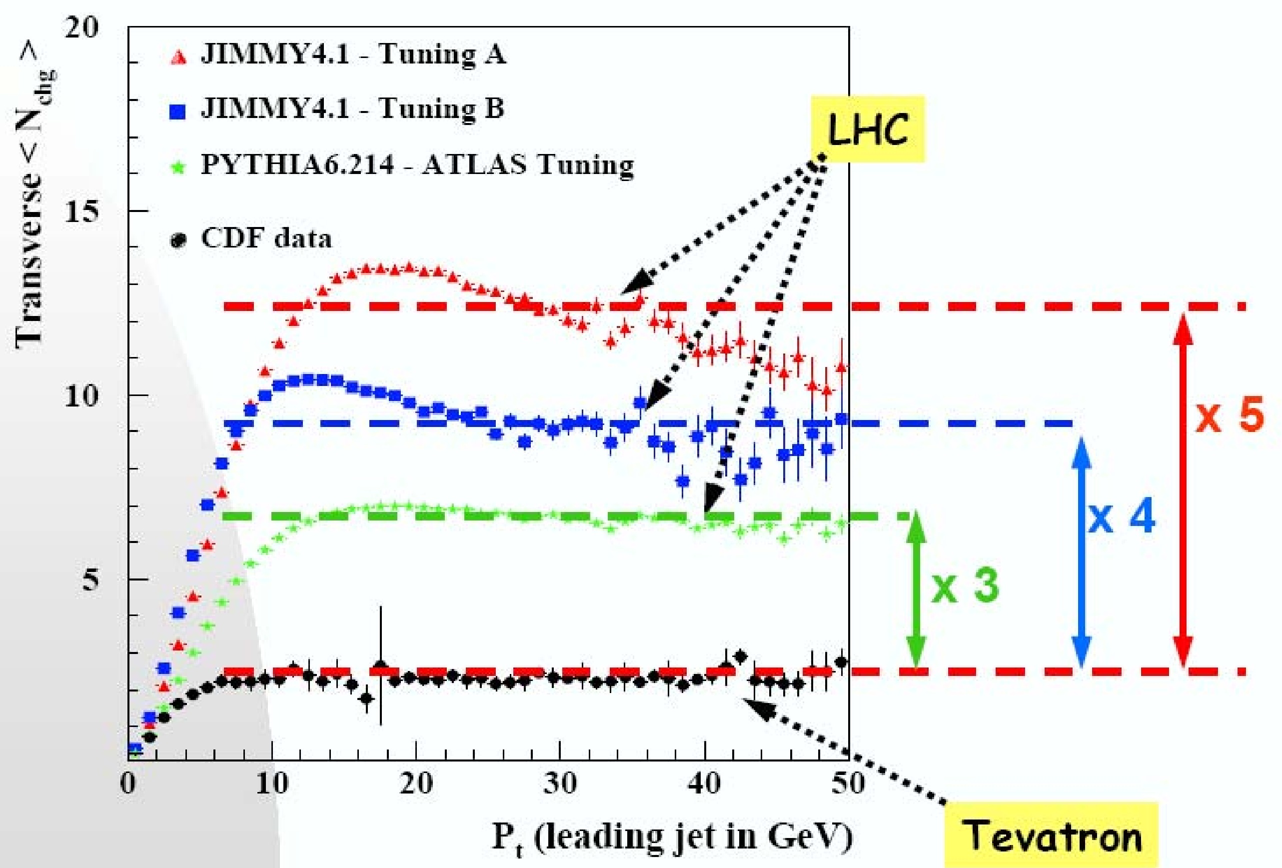
PYTHIA 6.2 Tunes
LHC HC Min-Bias Predictions





- Shows the predictions of **PYTHIA Tune A**, Tune DW, Tune DWT, and the ATLAS tune for the charged particle multiplicity distribution for "hard core" (HC) Min-Bias at 14 TeV ($|\eta| < 1$) and the average number of charged particles with $p_T > p_T^{min}$ ($|\eta| < 1$).
- The ATLAS tune has many more "soft" particles than does any of the CDF Tunes. The ATLAS tune has $<N_{chg}> = 18.14$ ($|\eta| < 1$) while Tune A has $<N_{chg}> = 11.93$ ($|\eta| < 1$).

LHC predictions: JIMMY4.1 Tunings A and B vs. PYTHIA6.214 – ATLAS Tuning (DC2)



A. M. Moraes

Minimum-bias and the Underlying Event at the LHC

5th November 2004

Improvements to PYTHIA

- One of the recent changes to PYTHIA

  T. Sjostrand and P.Z. Skands Eur.Phys.J.C39:129-154,2005

- Interleave the multiply scattering and initial-state shower.
- Order everything in terms of the p_T giving competition between the different processes.
- Also changes to the colour structure of the remnant.

Improvements to JIMMY

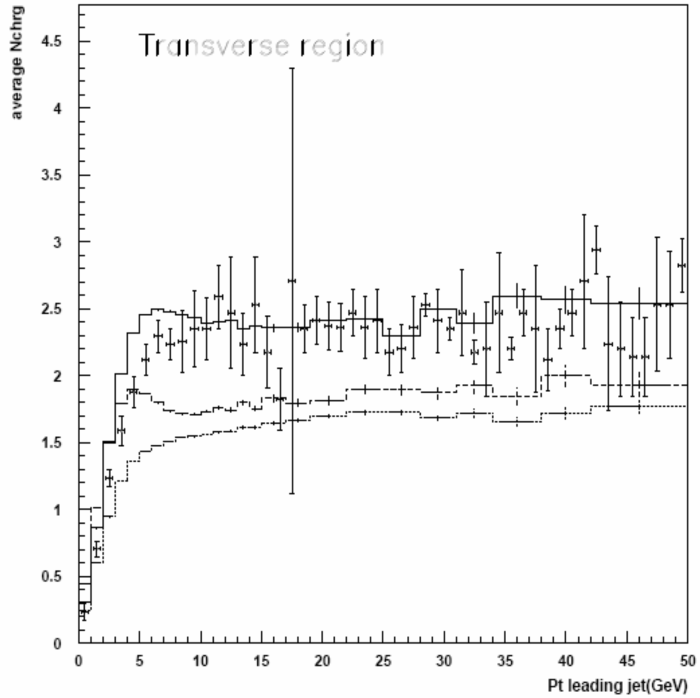


Figure 9: The average number of charged particles as a function of P_t (leading charged jet) in the transverse region. HERWIG + Eikonal model (solid line), HERWIG + Underlying Event model (solid dashed), HERWIG + Multiparton Hard model (dotted), experimental data solid circles.

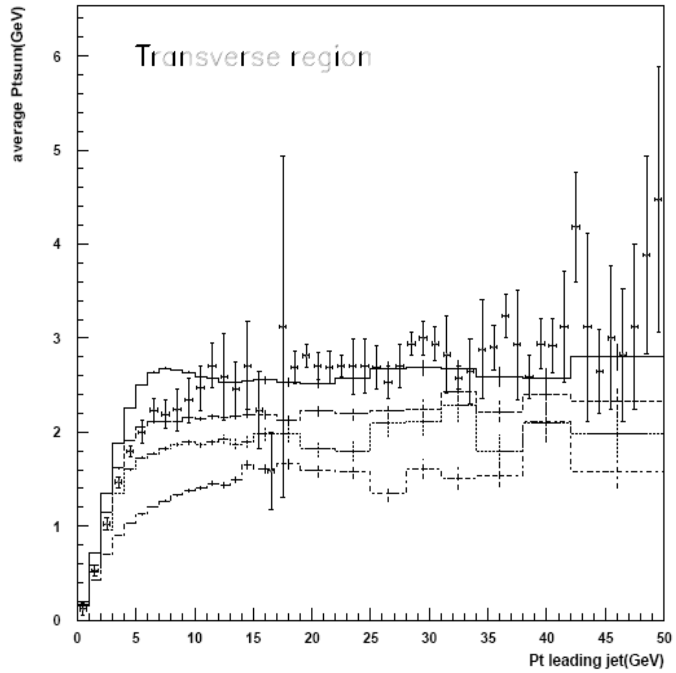


Figure 10: The average P_t sum of charged particles as a function of P_t (leading charged jet) in the transverse region. HERWIG + Eikonal Model for the two sets of p_{tmin}=3.0 GeV (solid line) and 2.0 GeV (dashed), HERWIG + Multiparton Hard Model p_{tmin}=2.0 GeV (dotted), p_{tmin}=3.0 GeV (dott dashed), experimental data points.

•One problem with JIMMY is the hard cut on the p_T. One idea is to include a soft component below the cut-off

I. Borozan, M.H. Seymour JHEP 0209:015,2002.

Hadron Decays

- The final step of the event generation is to decay the unstable hadrons.
- This is unspectacular/ungrateful but necessary, after all this is where most of the final-state particles are produced.
- There’s a lot more to it than simply typing in the PDG.
- Normally use dedicated programs with special attention to polarization effects:
- EVTGEN: B Decays
- TAUOLA: τ decays
- PHOTOS: QED radiation in decays.

The Future

- Most of the work in the generator community is currently devoted to developing the next generation of C++ generators.
- We needed to do this for a number of reasons
  - Code structures needed rewriting.
  - Experimentalists don’t understand FORTRAN any more.
  - Couldn’t include some of the new ideas in the existing programs.

The Future

- A number of programs
  - ThePEG
  - Herwig++
  - SHERPA
  - PYTHIA8
- While it now looks likely that C++ versions of HERWIG and PYTHIA won’t be used for early LHC data this is where all the improvements the experimentalists want/need will be made and will have to be used in the long term.

Outlook

- The event generators are in a constant state of change, in the last 5 years
  - Better matrix element calculations.
  - Improved shower algorithms.
  - Better matching of matrix elements and parton showers.
  - First NLO processes.
  - Improvements to hadronization and decays.
  - Improved modelling or the underlying event.
  - The move to C++.
- Things will continue to improve for the LHC.

Summary

- Hopefully these lectures will help you understand the physics inside Monte Carlo event generators.
- If nothing else I hope you knew enough to start think about what you are doing when running the programs and questions like
  - Should the simulation describe what I'm looking for?
  - What is the best simulation for my study?
  - What physics in the simulation affects my study?
  - Is what I'm seeing physics or a bug?