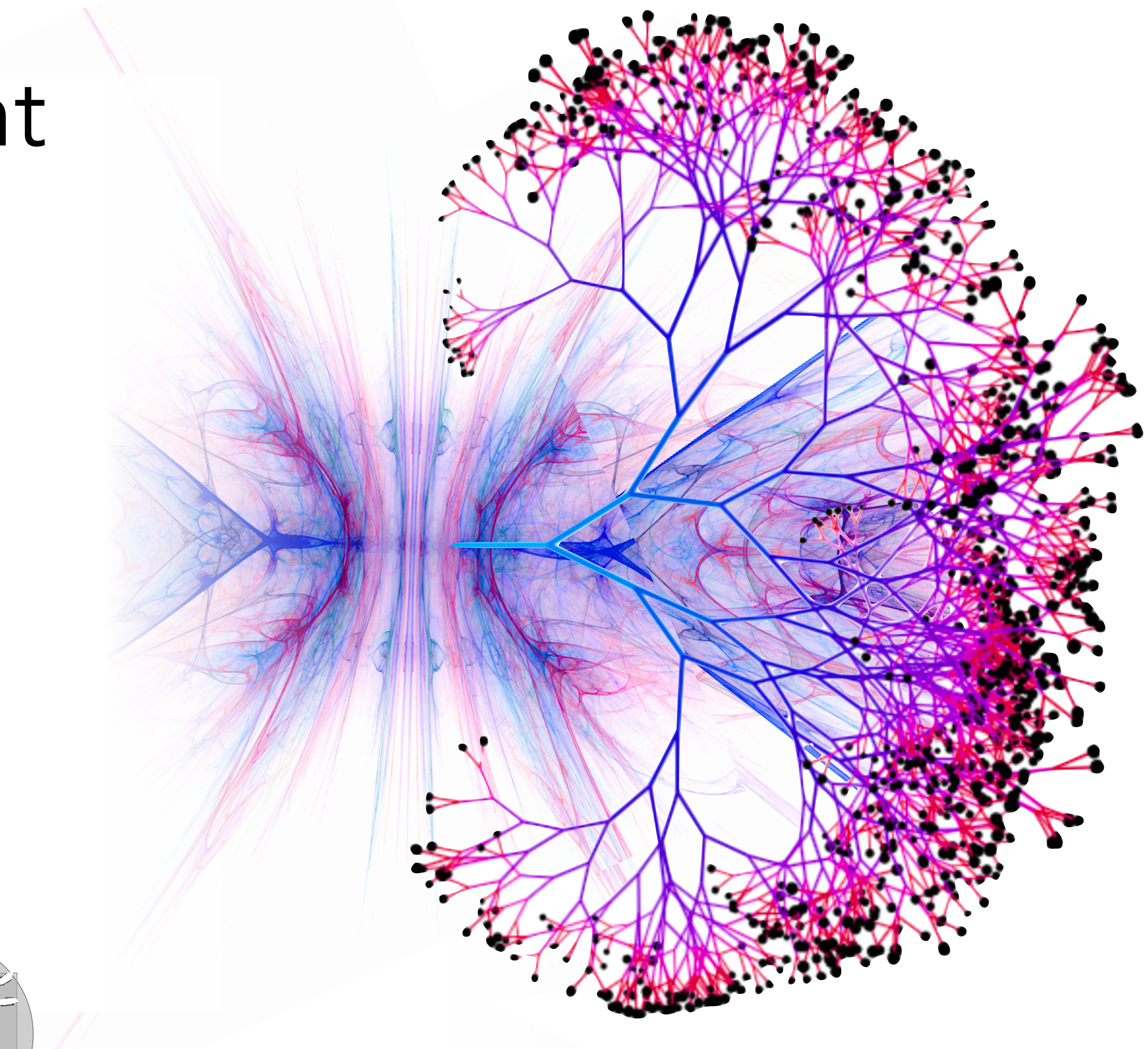
Hadronization
& Underlying Event

QCD and Event Generators
Lecture 3 of 3

Peter Skands
Monash University
(Melbourne, Australia)

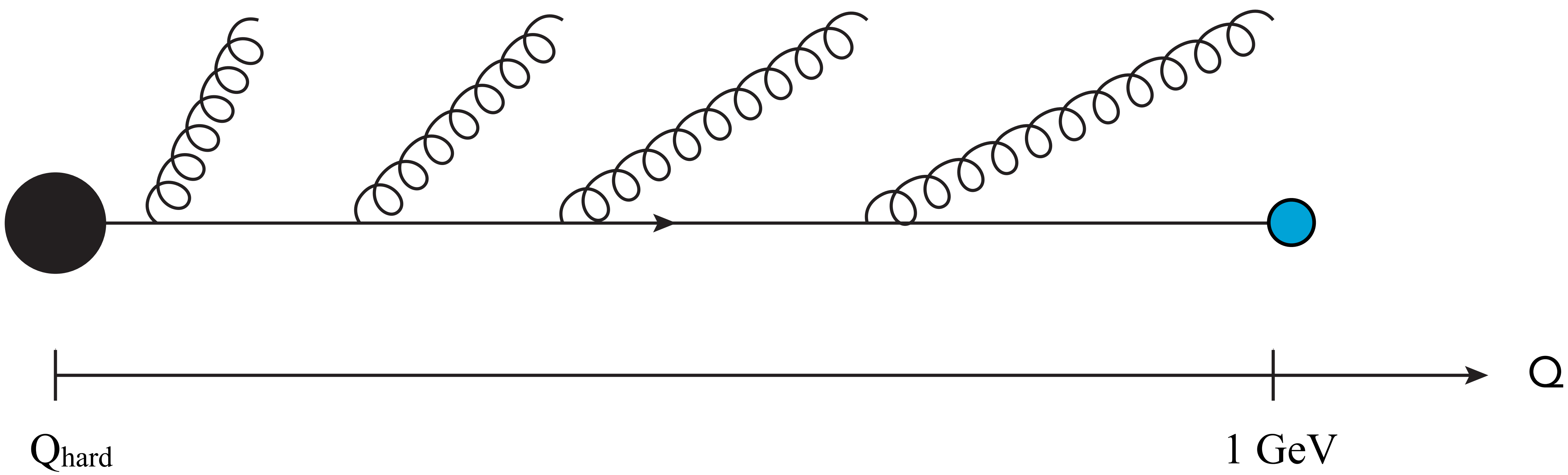


Consider a parton emerging from a hard scattering (or decay) process

It starts at a high factorization scale $Q = Q_F = Q_{hard}$

It showers (bremsstrahlung)

It ends up at a low effective factorization scale $Q \sim m_\rho \sim 1$ GeV



How about I just call it a hadron?

→ "Local Parton-Hadron Duality"

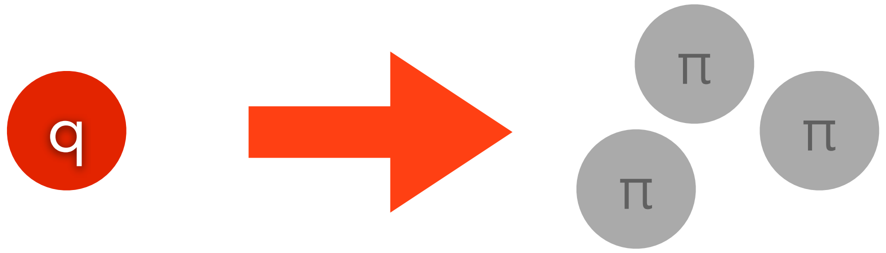
Parton → Hadrons?

Early models: "Independent Fragmentation"

Local Parton Hadron Duality (**LPHD**) can give useful results for inclusive quantities in collinear fragmentation

Motivates a simple model:

"Independent Fragmentation"



But …

The point of confinement is that partons are **coloured**

Hadronisation = the process of **colour neutralisation**

→ Unphysical to think about independent fragmentation of a single parton into hadrons

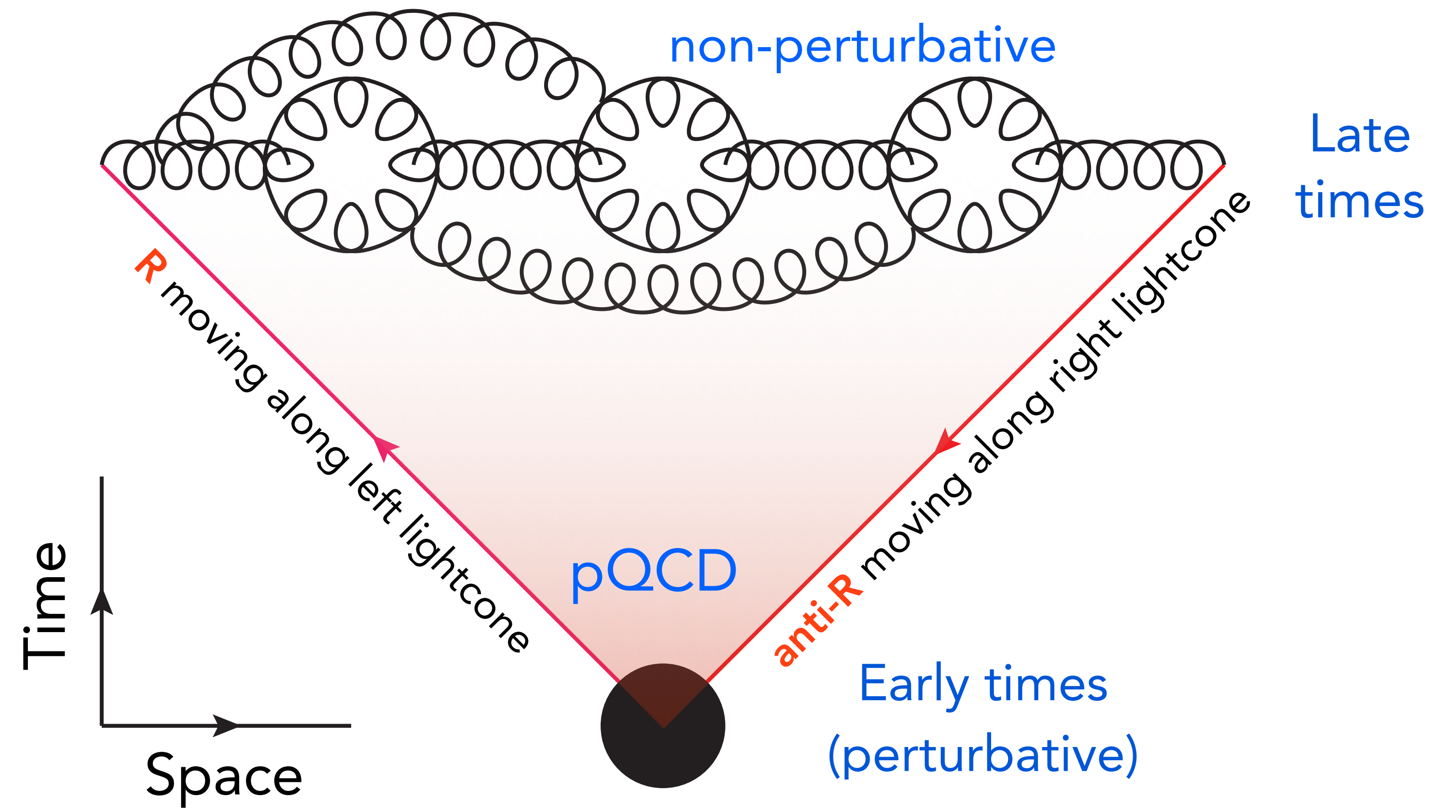
→ Too naive to see **LPHD** (inclusive) as a justification for Independent Fragmentation (exclusive)

→ More **physics** needed

A **physical** hadronization model

Should involve at least **two** partons, with opposite color charges*

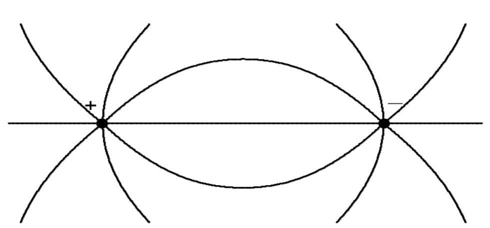
A strong **confining field** emerges between the two when their separation $\gtrsim$ 1fm



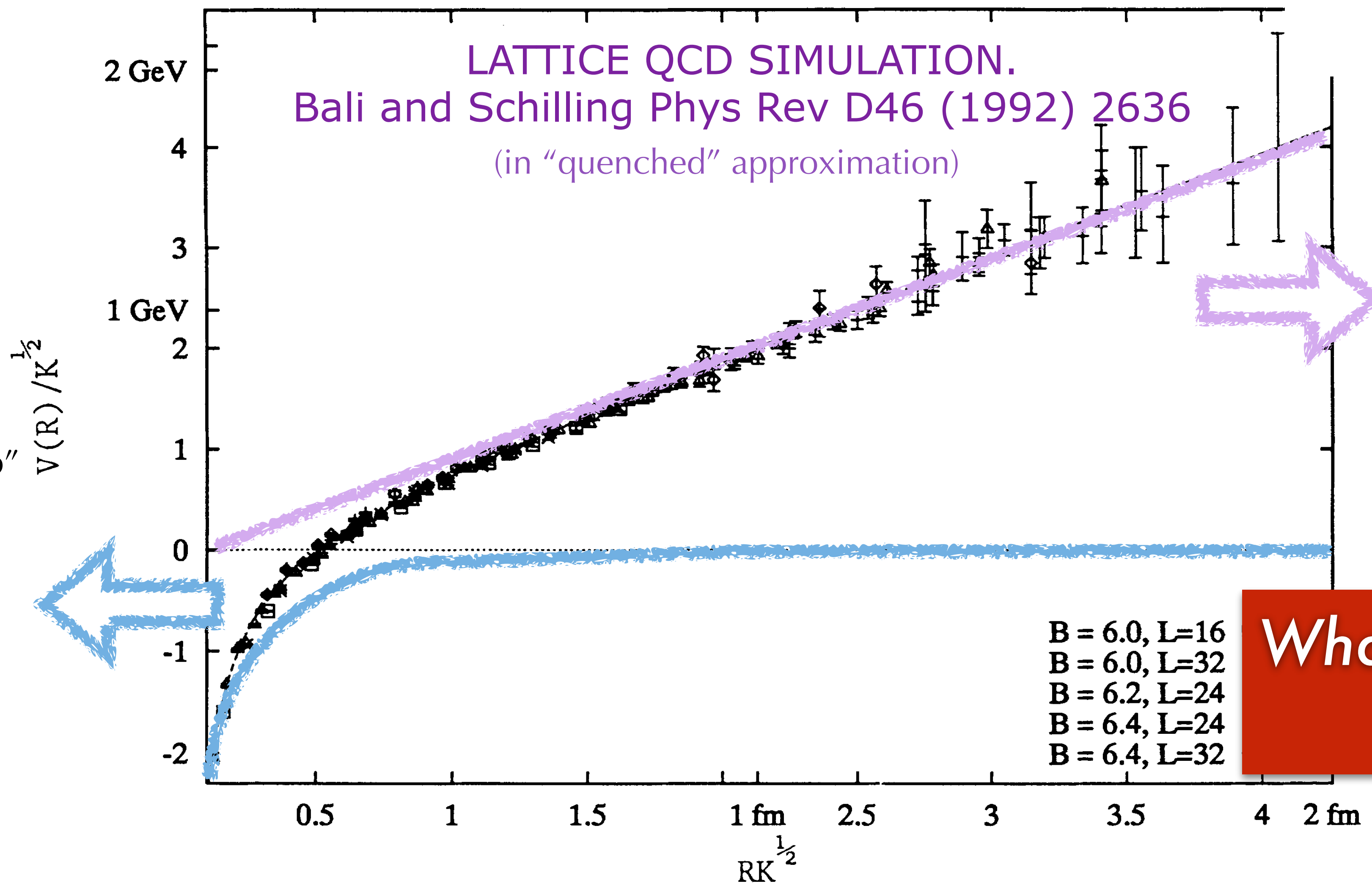
*) Really, a colour singlet state $\frac{1}{\sqrt{3}}\left(|R\bar{R}\rangle + |G\bar{G}\rangle + |B\bar{B}\rangle\right)$; the LC **colour flow rules** discussed in lecture 1 allow us to tell which partons to pair up (at least to LC; see arXiv:1505.01681)

Using explicit computer simulations of QCD on a 4D "lattice" (lattice QCD), one can compute the potential energy of a colour-singlet $q\bar{q}$ state, as a function of the distance, r, between the q and $\bar{q}$

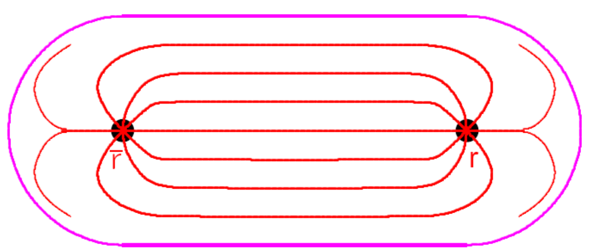
Short Distances ~ "Coulomb"



"Free" Partons



Long Distances ~ Linear Potential



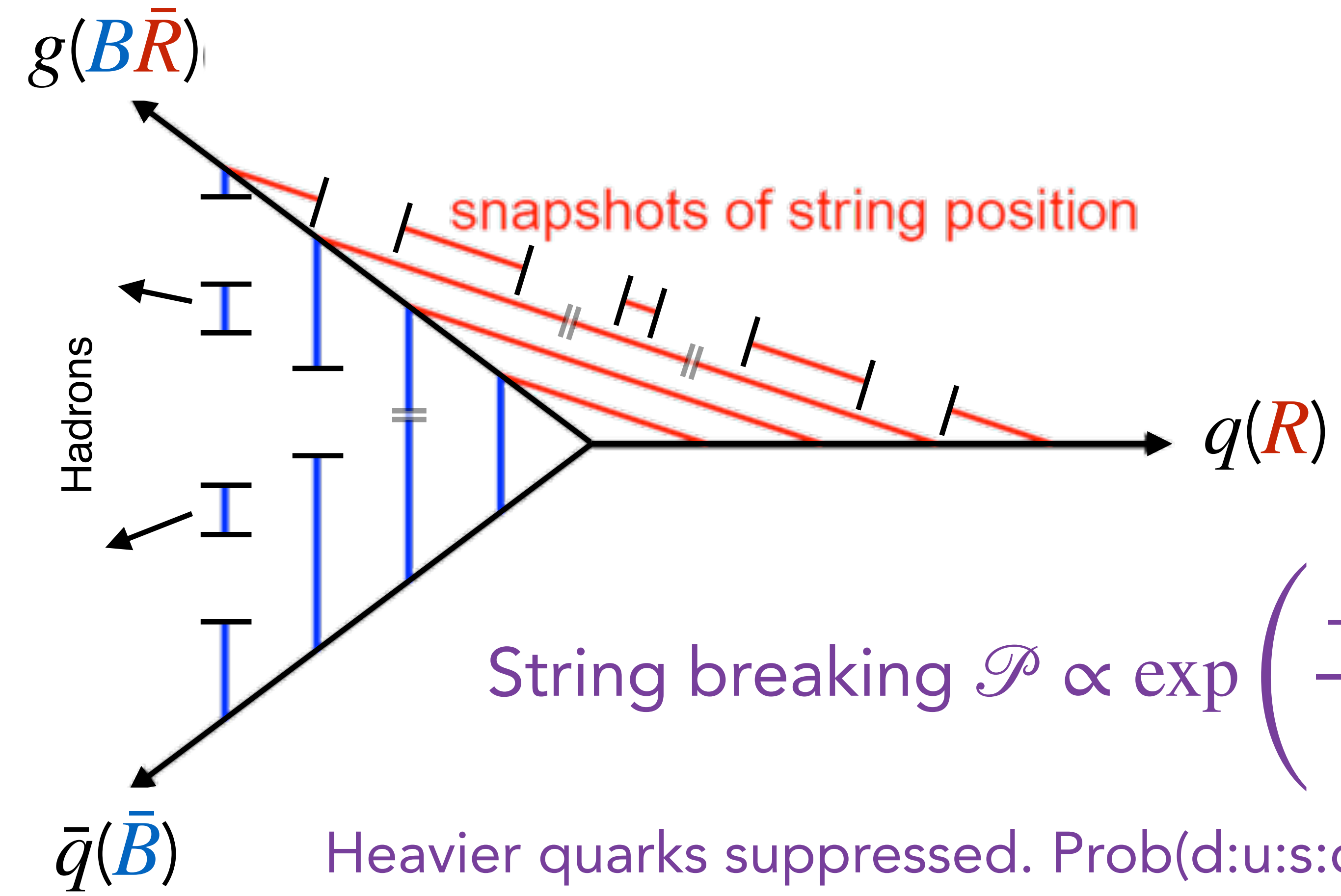
"Confined" Partons
(a.k.a. Hadrons)

What physical system has a* linear *potential?

"Cornell Potential" fit: $V(r) = -\frac{a}{r} + \kappa r$ **with** $\kappa \sim 1\ \mathrm{GeV/fm}$ (→ could lift a 16-ton truck)

A high-energy quark-gluon-antiquark system is created and starts to fly apart

- **Quarks** → String Endpoints
- **Gluons** → Transverse Excitations (kinks)
- **Physics** then in terms of 1+1-dim **string "worldsheet"** evolving in spacetime
- Probability of **string break** (by quantum tunneling) constant per unit space-time area



String breaking $\mathscr{P} \propto \exp\left(\frac{-m^2 - p_\perp^2}{(\kappa/\pi)}\right)$

Heavier quarks suppressed. Prob(d:u:s:c) ≈ 1 : 1 : 0.2 : 10^{-11}
and Gaussian pT spectrum (transverse to local string axis)

Computer algorithms to model this process began to be developed in late 70'ies and early 80'ies

➜ **Monte Carlo Event Generators**

Modern MC **hadronization models**: PYTHIA (string), HERWIG (cluster), SHERPA (cluster)

The (Lund) String Hadronization Model ↔ PYTHIA (org JETSET)

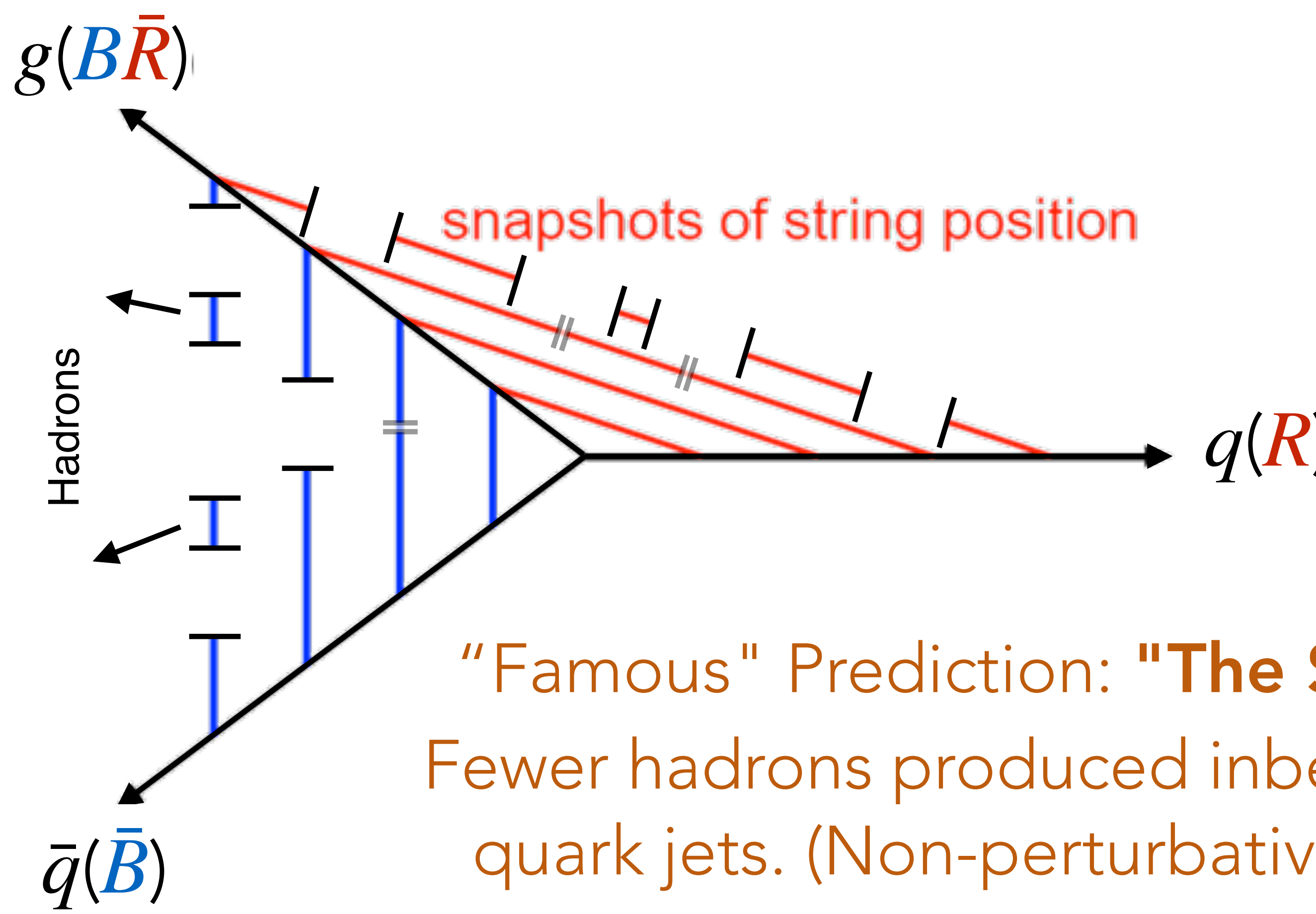
Simple space-time picture

Highly predictive, few free parameters

Causality and Lorentz invariance $\Longrightarrow$ "Lund Symmetric Fragmentation Function" with two free parameters *a* and *b*:

$$f(z) \propto \frac{(1-z)^a}{z} \exp(-b m_\perp^2 / z)$$

with $z \sim E_{\mathrm{hadron}}/E_{\mathrm{quark}}$



"Famous" Prediction: **"The String Effect"**
Fewer hadrons produced inbetween the two quark jets. (Non-perturbative coherence.)
Confirmed by JADE in 1980.

Details of string breaks more complicated

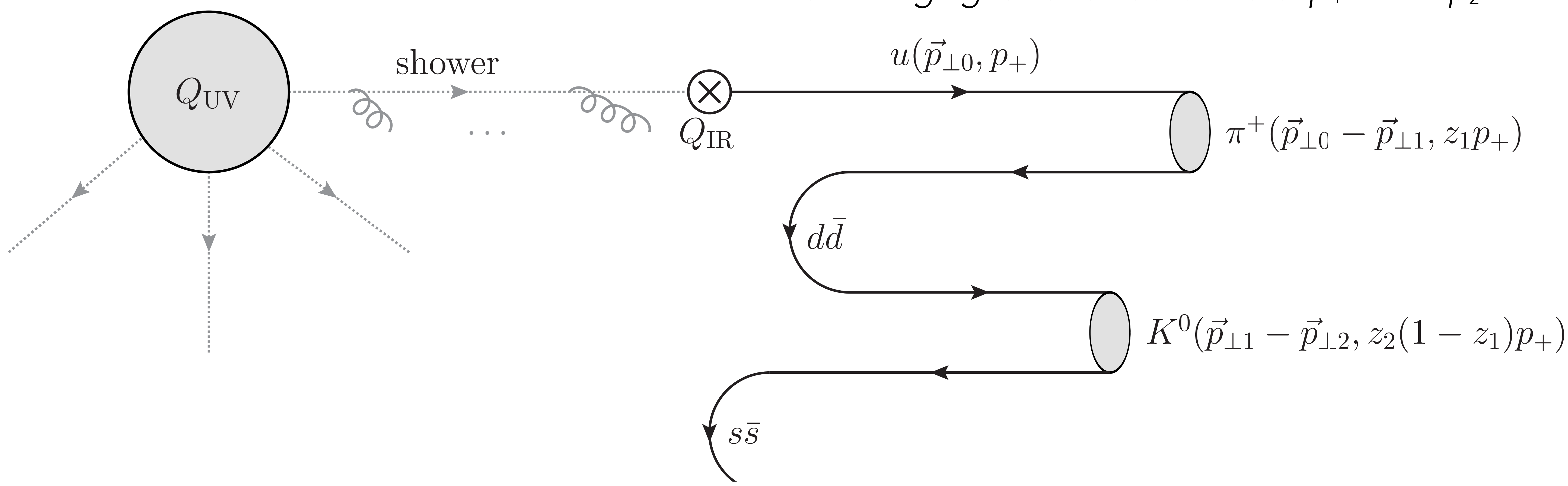
Many free parameters for **flavour & spin** of produced hadrons ➜ **fit to** $e^+e^- \to \mathrm{hadrons}$

Iterative String Breaks

String breaks are separated by **spacelike** intervals ➜ **causally disconnected**

➜ We do not have to consider the string breaks in any specific time order ➜ choose the most convenient order for us: **starting from the endpoints** ("outside-in")

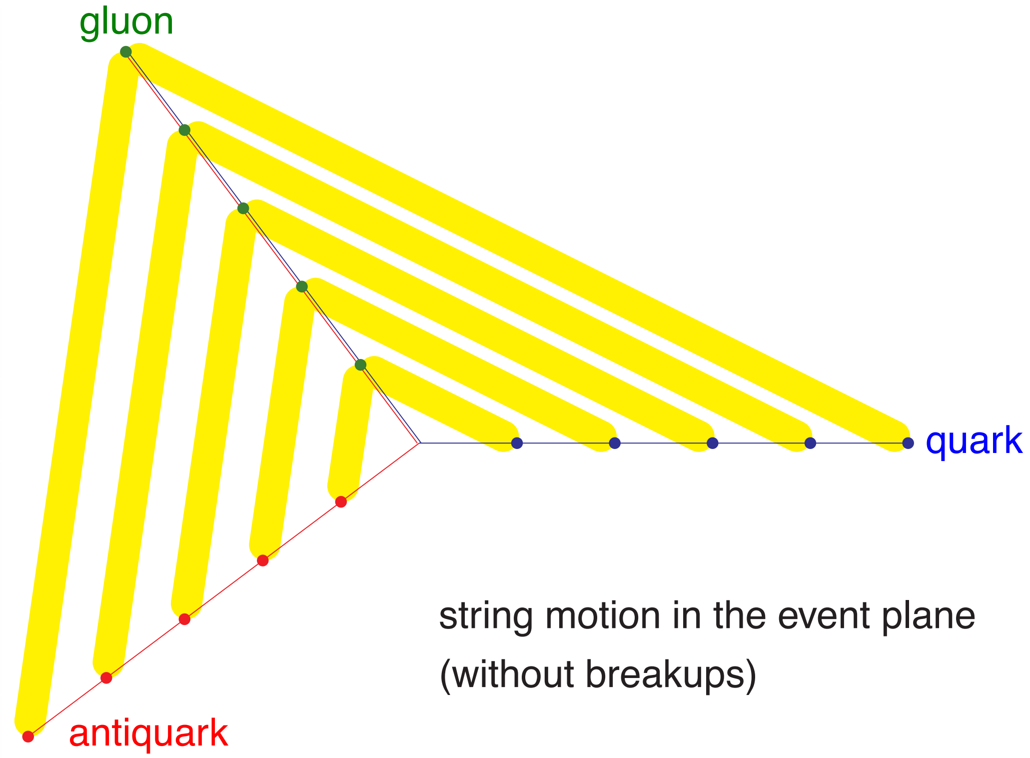
Note: using light-cone coordinates: $p_+ = E + p_z$



On average, expect energy of n$^{\mathrm{th}}$ "rank" hadron ~ E_n ~ $<z>^n E_0$

Hallmark feature of Lund string model:

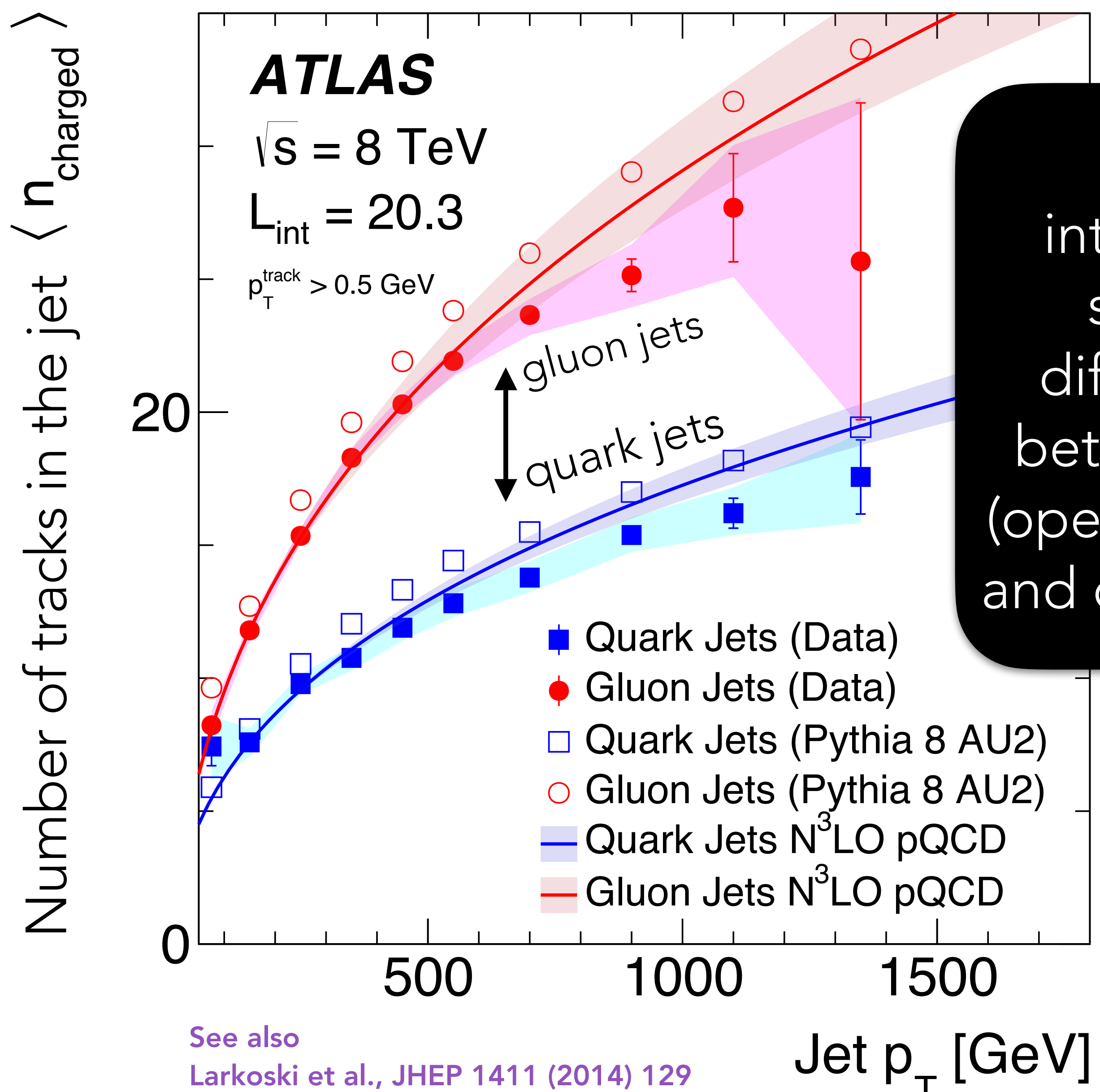
Gluon connected to **two** string pieces



Each quark connected to **one** string piece

→ **expect factor 2 ~ C_A/C_F larger particle multiplicity in gluon jets vs quark jets**

ATLAS, Eur.Phys.J. C76 (2016) no.6, 322



Note: interesting smaller differences between MC (open symbols) and data (filled)

See also
Larkoski et al., JHEP 1411 (2014) 129
Thaler et al., Les Houches, arXiv:1605.04692

Can be important for discriminating new-physics signals (decays to quarks vs decays to gluons, vs composition of background and bremsstrahlung combinatorics)

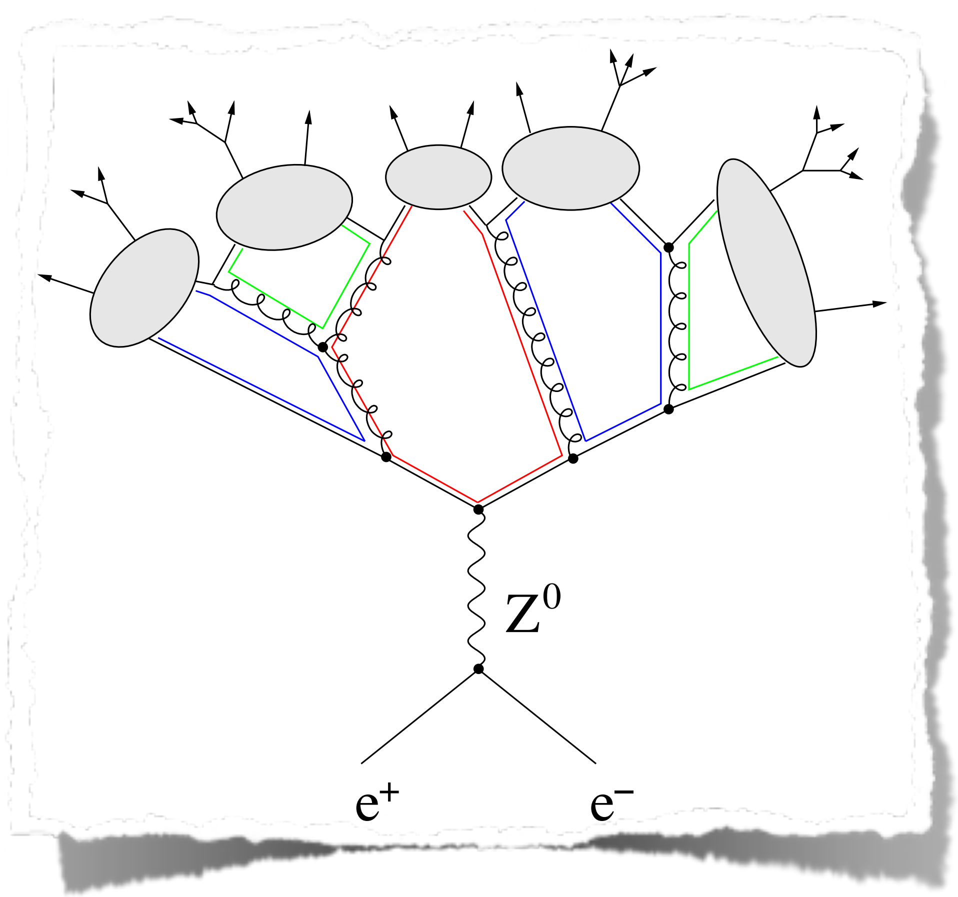
Starting observation: "Preconfinement"

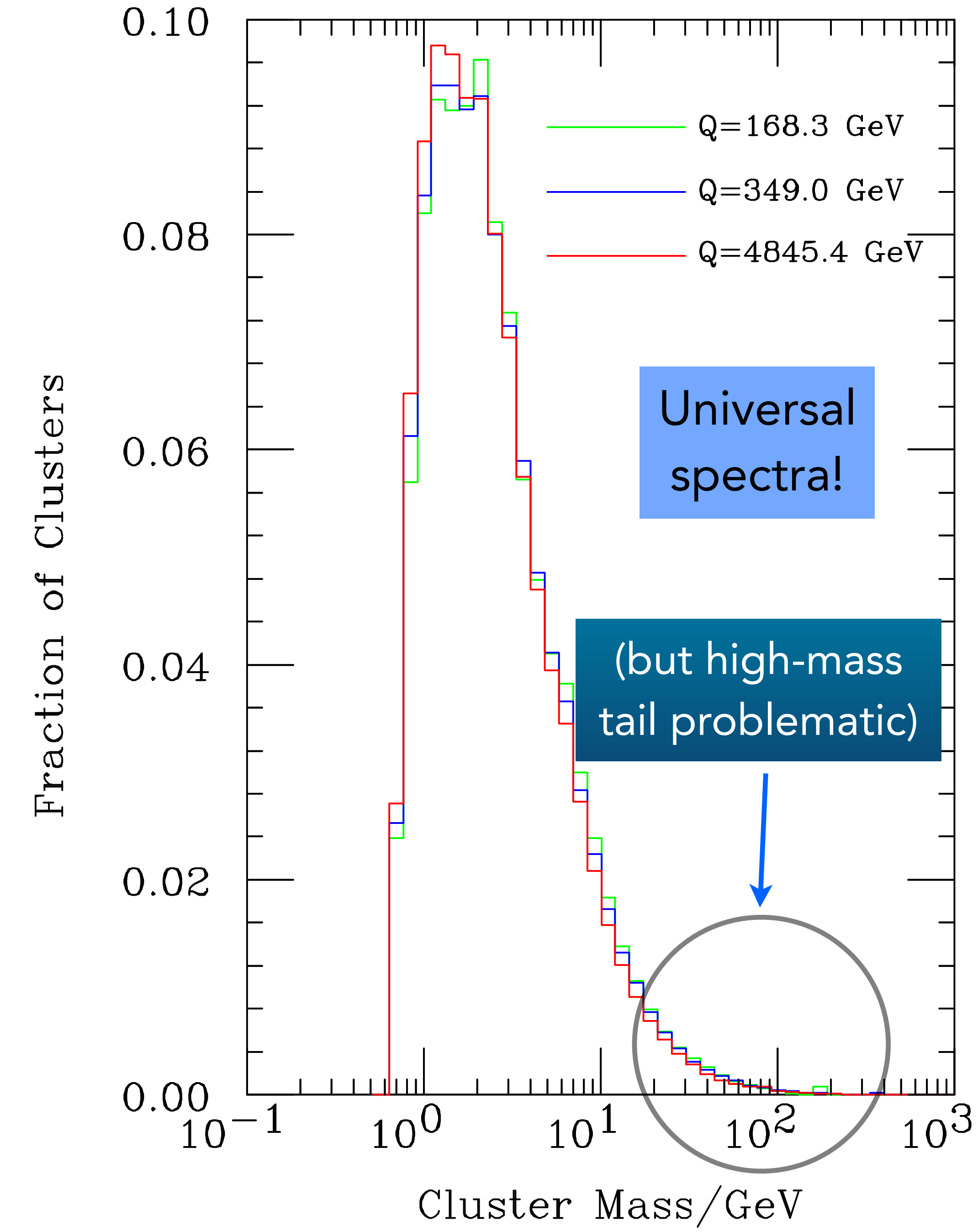
+ Force g→qq splittings at Q_0

→ high-mass q-qbar "clusters"

Isotropic 2-body decays to hadrons

according to PS ≈ $(2s_1+1)(2s_2+1)(p^*/m)$





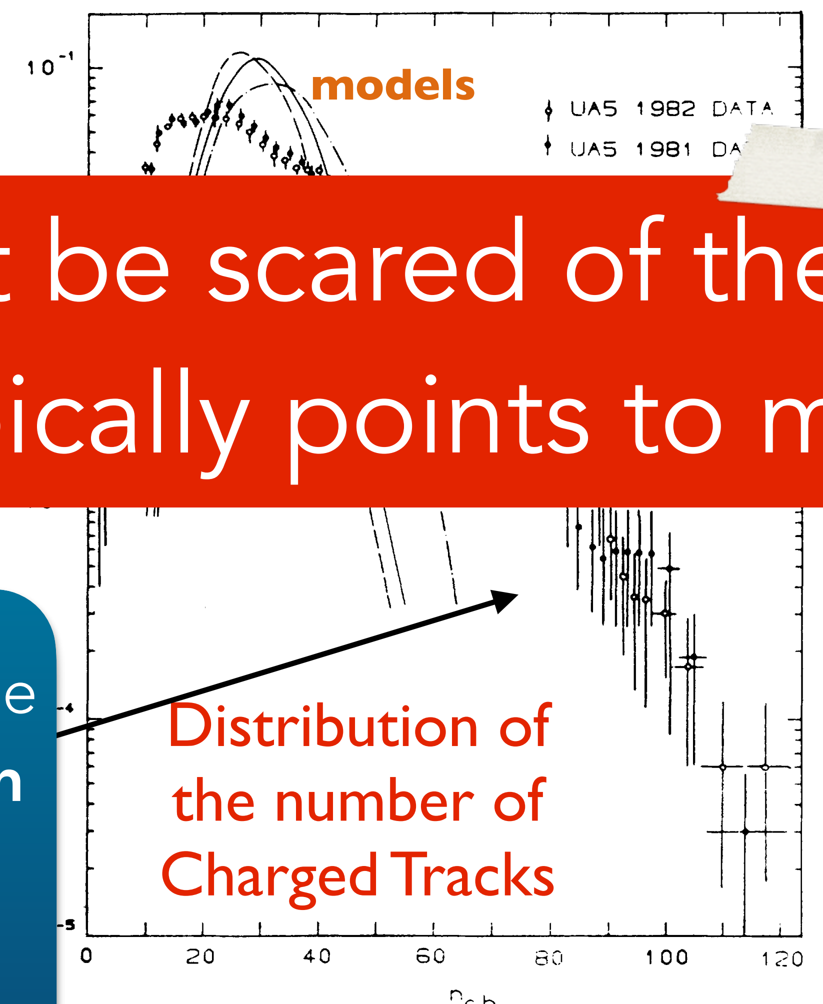
Large clusters → string-like. (In PYTHIA, small strings → cluster-like).

MC vs Hadron Collisions

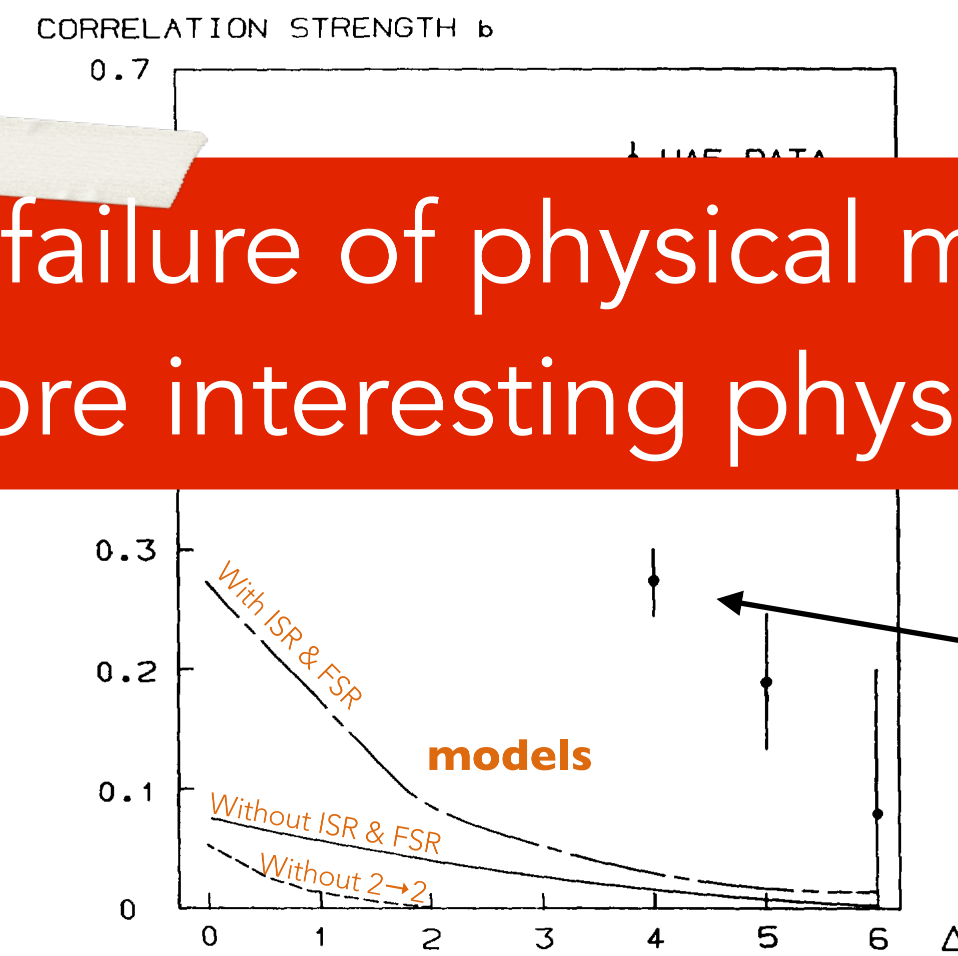
Now that we have a model that includes hard interactions, showers, and string fragmentation, **let's apply it to pp collisions**!

Do not be scared of the failure of physical models
(typically points to more interesting physics)

Can get ~ right average but data exhibits **much** bigger **fluctations** in multiplicity
(here: of charged tracks)



3. Charged-multiplicity distribution at 540 GeV, UA5 (Ref. 32) vs simple models: dashed low p_T only, full including hard scatterings, dash-dotted also including initial- and final-state radiation.



Correlation Strength (forward-backward)

some global (quantum) number tells the entire event to fluctuate up or down across many units of rapidity?

Sjöstrand & v. Zijl, Phys.Rev.D36 (1987)2019

Further evidence of additional physics in hadron-hadron

1983: discovery of the **"Pedestal Effect"**

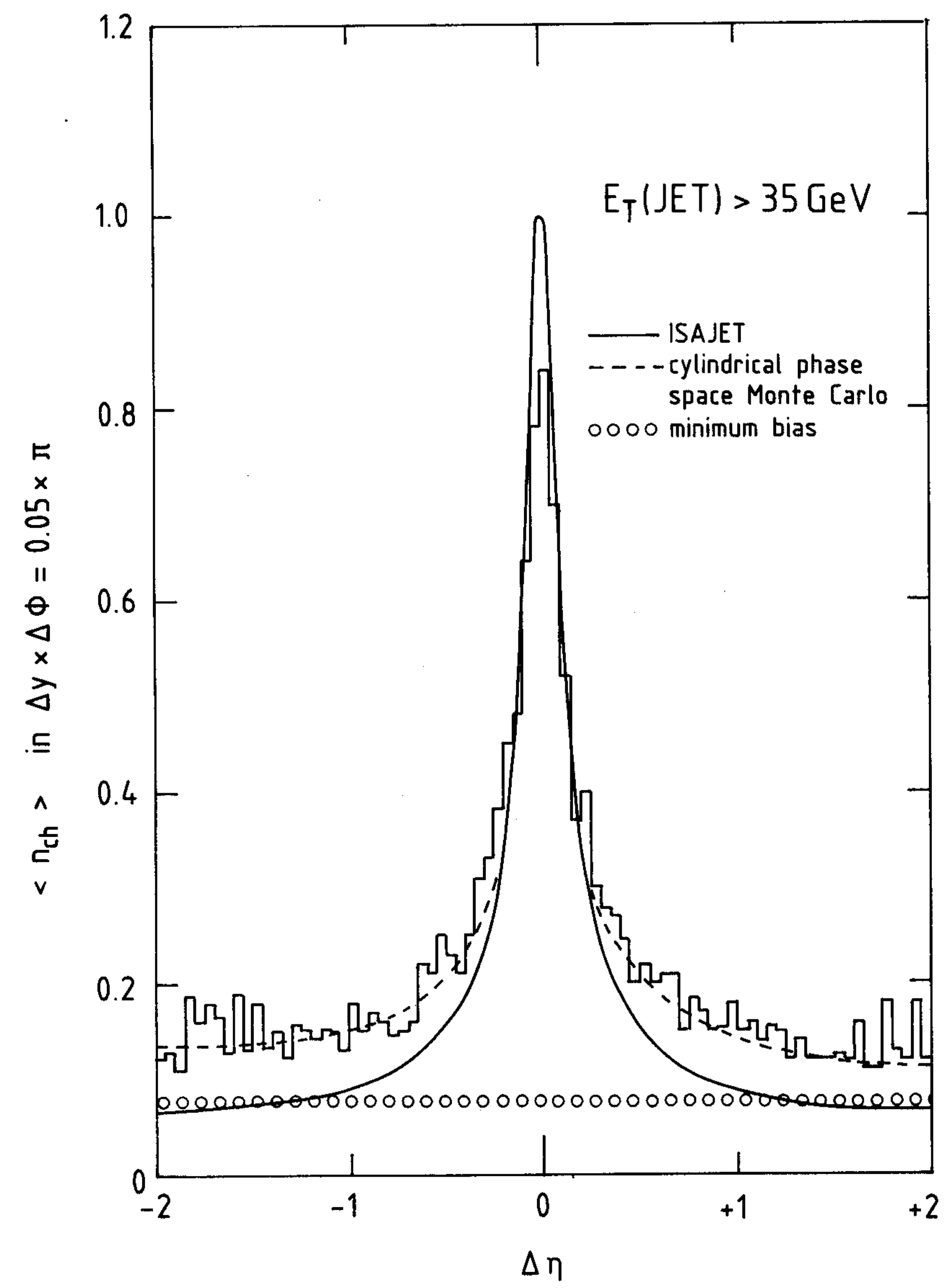
UA1: $p\bar{p}$ at $\sqrt{s} = 540\,\mathrm{GeV}$

Studies of jets with E_T up to 100 GeV

Phys. Lett. B 132 (1983) 214-222

"Outside the [jet], a constant E_T plateau is observed, whose height is independent of the jet E_T. Its value is substantially higher than the one observed for minimum bias events."

In hadron-hadron collisions, **hard jets** sit on **"pedestals"** of increased particle production **extending** far from the jet cores.



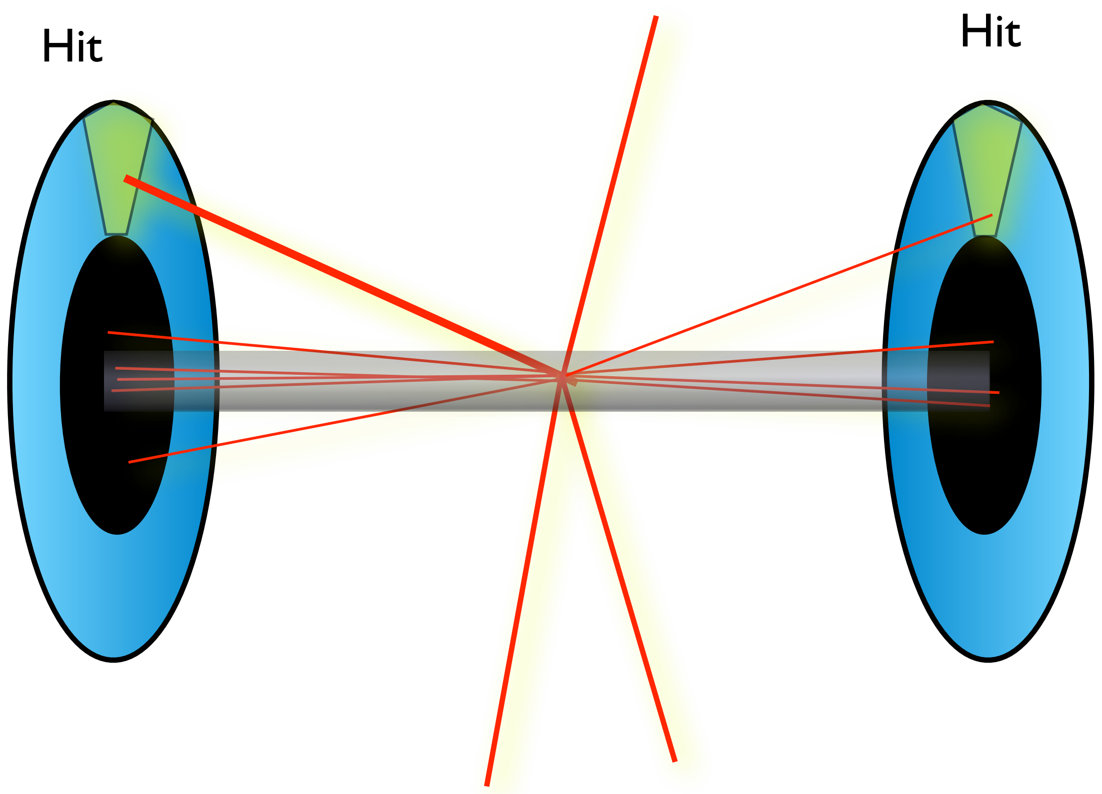
What's "Minimum-Bias"?

Simple question: what does the ***average*** LHC collision look like?

First question: how many are there? What is $\sigma_{tot}(pp)$ at LHC ?

Around 100mb (of which about half is "inelastic, non-diffractive")

Example of "Minimum Bias Trigger"



Minimum Bias = Minimal trigger requirement

At least one hit in some simple and efficient hit counters (typically at large η)

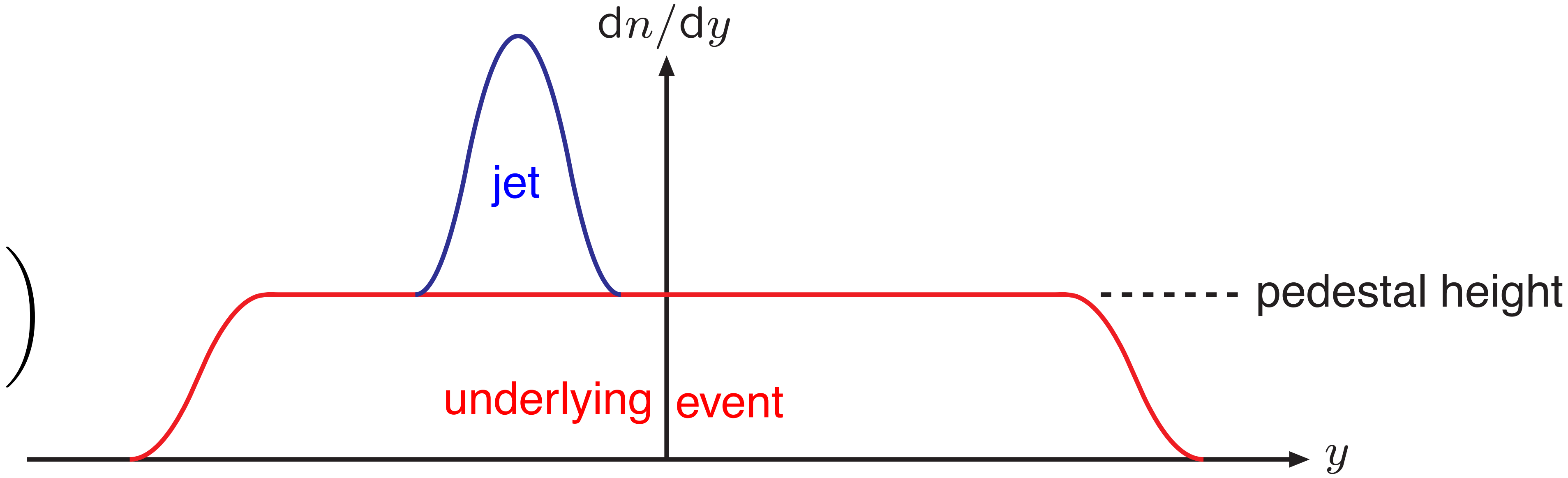
(Double-sided trigger requirement suppresses "single diffraction")

Dissecting the Pedestal

Illustrations by
T. Sjöstrand

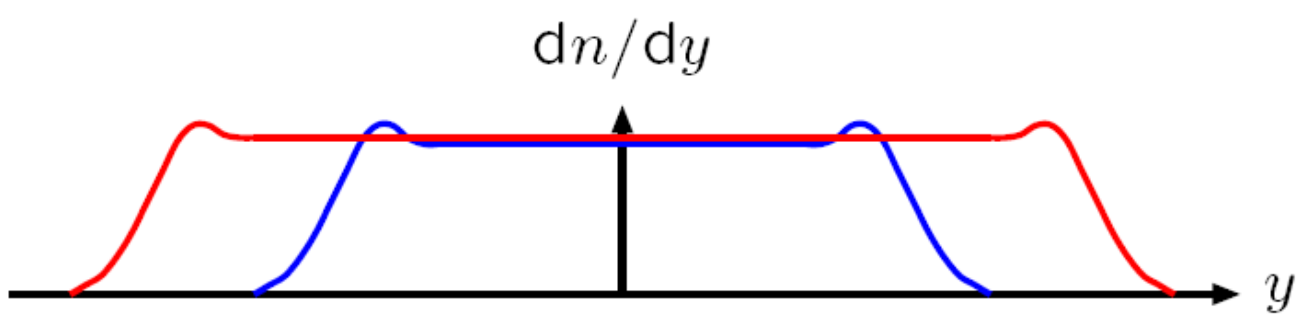
Today, we call the pedestal **"the Underlying Event"**

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right)$$



Rapidity (along beam axis)

A uniform (constant) particle density per rapidity unit is just what a string produces …



Rapidity (along string axis)

but the **height** of the pedestal was much larger than that of **one** string…

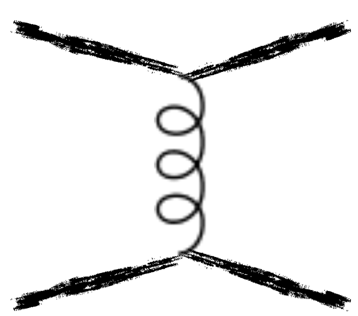
Multiple Interactions?

Parton-Parton vs Proton-Proton Cross Sections

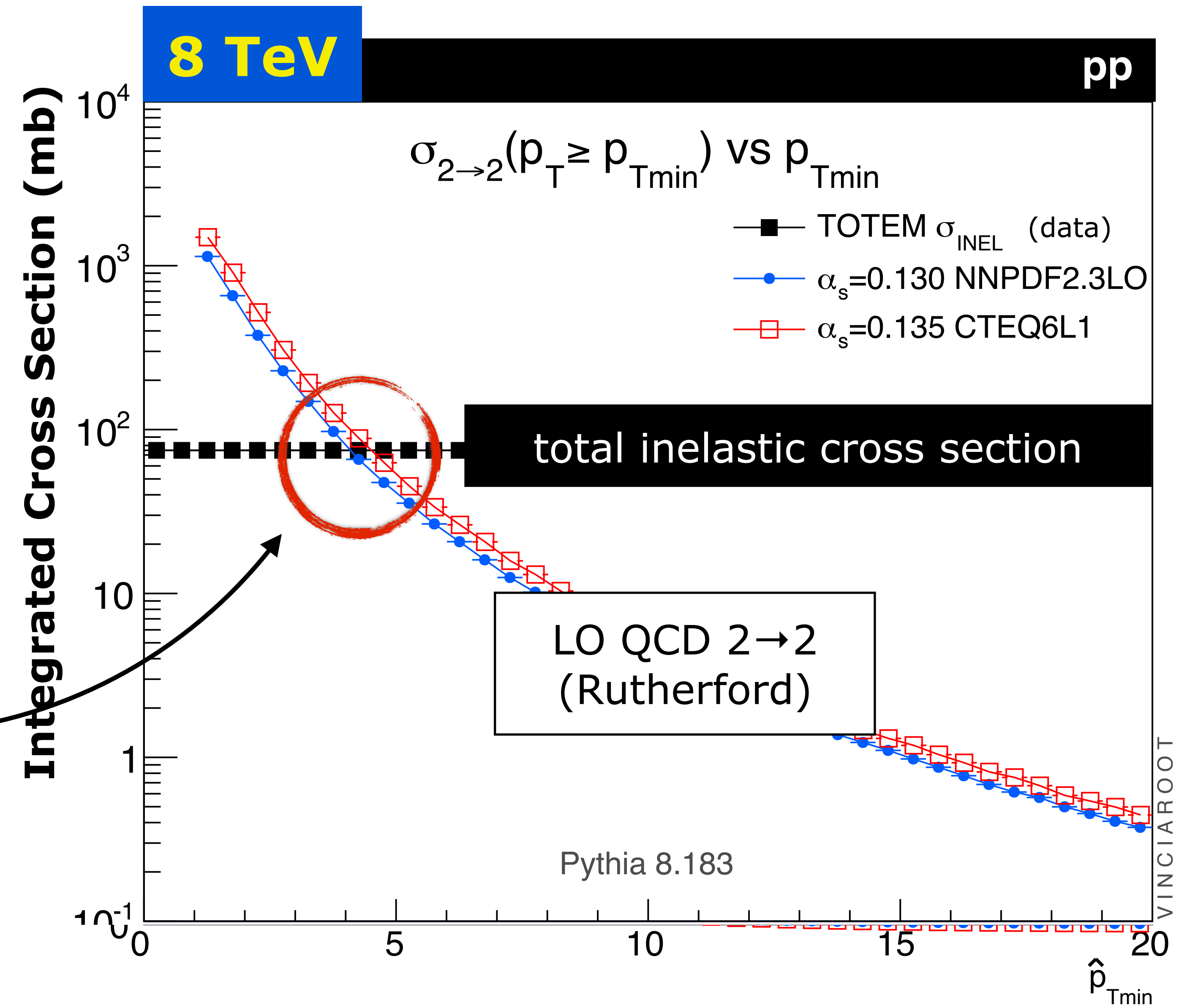
Total **inelastic** pp cross section @ 8 TeV* ~ 80 mb (measured by TOTEM)

Compare this to perturbative calculation of QCD $2 \to 2$ scattering cross section (mainly t-channel gluon exchange; divergent for $p_T \to 0$)

QCD $2 \to 2$ cross section dominated by t-channel gluon exchange



Larger than total pp cross section for $\hat{p}_\perp \leq 4\,\mathrm{GeV}$



Interpret to mean that **every** pp collision has **more than one $2 \to 2$** QCD scattering with $\hat{p}_\perp \leq 4\,\mathrm{GeV}$

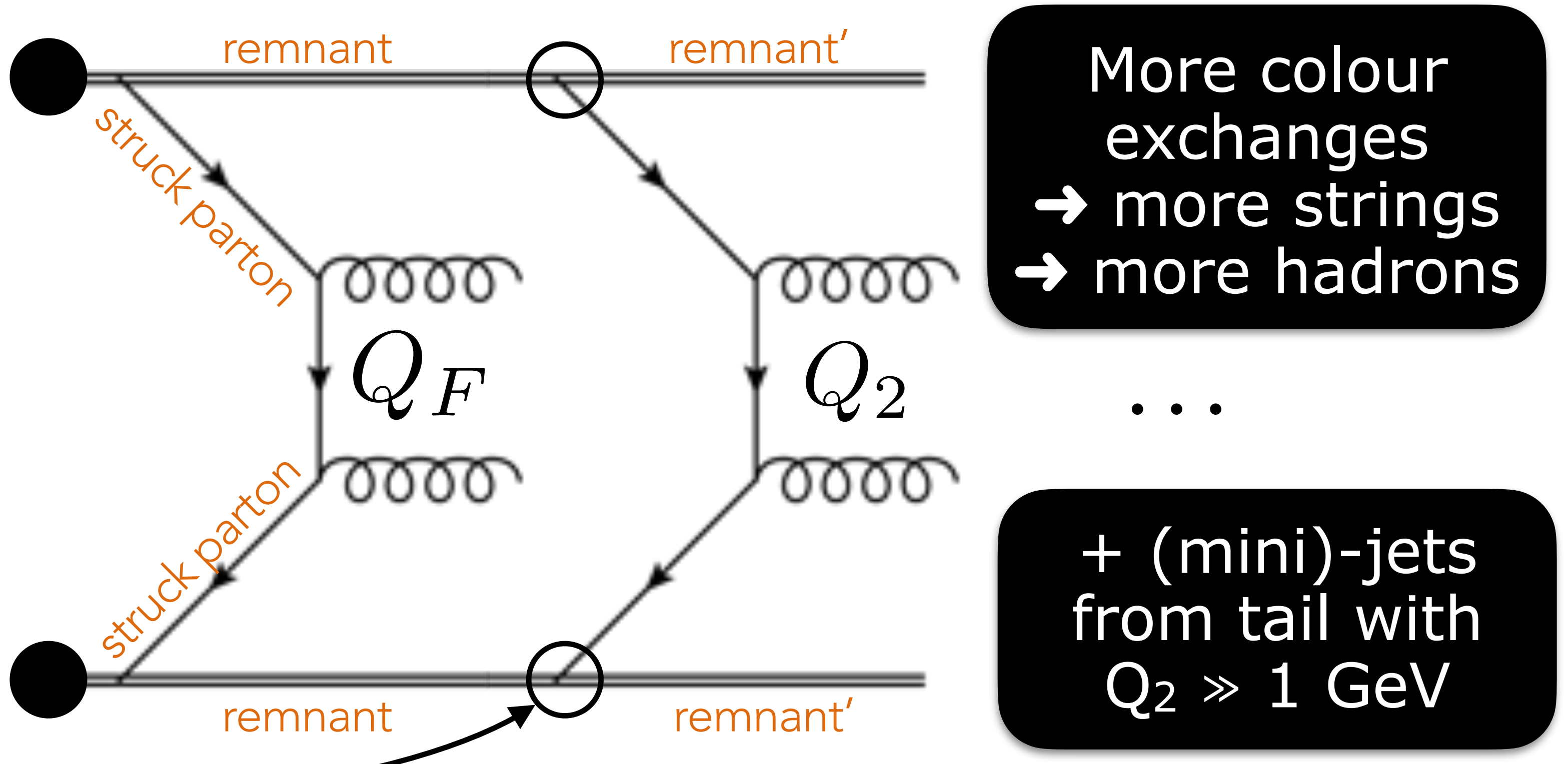
*Note: nothing particularly special about 8 TeV; the crossover point would be lower at lower E_{CM} and higher at higher E_{CM}

Physics of the Pedestal

Recall **Factorisation:** Subdivide calculation

Hard scattering: parton-parton cross section $d\hat{\sigma}$ independent of non-pert. dynamics

x **PDF factors** $f(x, Q_F^2)$
representing: partitioning of proton into struck **parton** + unresolved **remnant**, at *factorisation* scale Q_F^2



Multi-Parton Interactions (MPI)

Several QCD 2→2 in **one** pp collision

$\Longrightarrow$ need **Multi-parton PDFs** (PYTHIA, e.g., Sjöstrand & PS *JHEP* 03 (2004) 053 • hep-ph/0402078)

Constructed using **momentum** and **flavour conservation**; goes beyond existing factorisation theorems (though some work on special case Double Parton Scattering)

(More issues such as *colour reconnections, saturation, rescattering, higher twist,* not covered here)

How many?

Naively $$\langle n_{2\to2}(p_{\perp\min})\rangle = \frac{\sigma_{2\to2}(p_{\perp\min})}{\sigma_{\mathrm{tot}}}$$

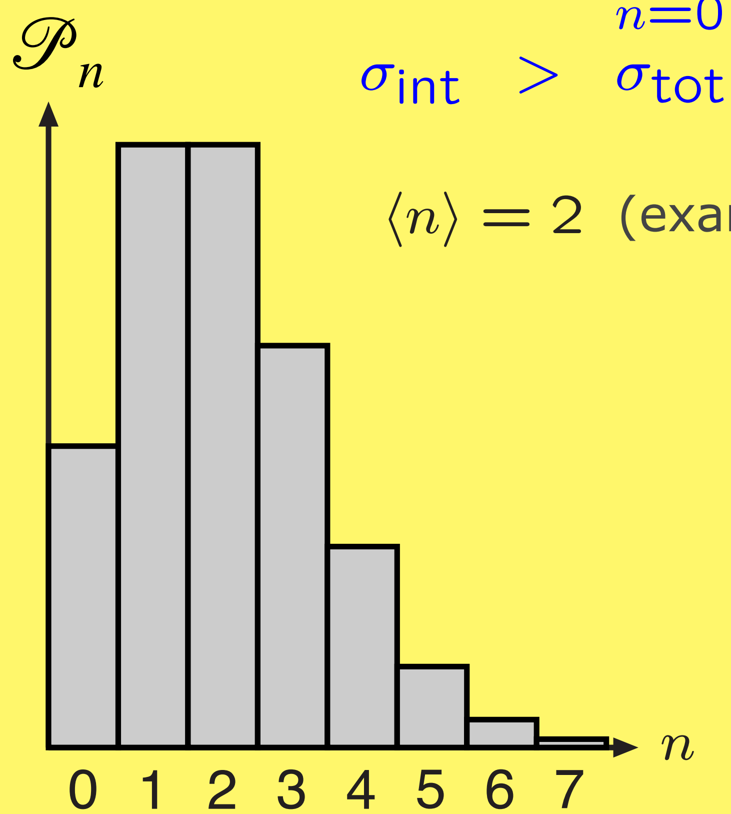
If the interactions are assumed ~ independent (naive factorisation) → Poisson

$$\sigma_{\mathrm{tot}} = \sum_{n=0}^{\infty} \sigma_n$$

$$\sigma_{\mathrm{int}} = \sum_{n=0}^{\infty} n\,\sigma_n$$

$$\sigma_{\mathrm{int}} > \sigma_{\mathrm{tot}} \iff \langle n\rangle > 1$$

$\langle n\rangle = 2$ (example)



$$\mathcal{P}_n = \frac{\langle n\rangle^n}{n!}\, e^{-\langle n\rangle}$$

Real Life

Color screening: $\sigma_{2\to2}\to0$ for $p_\perp\to0$

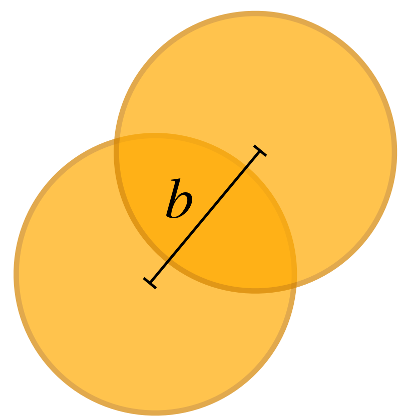
Momentum conservation suppresses high-n tail

Impact-parameter dependence

+ physical correlations

→ not simple product

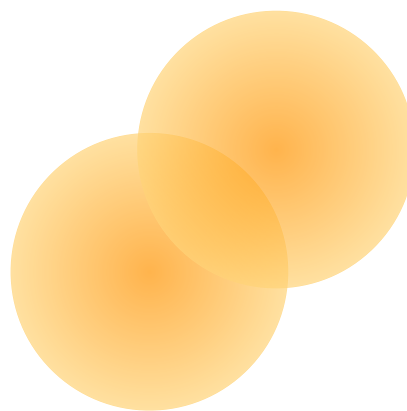
Impact Parameter Dependence



1. **Simple Geometry** (in impact-parameter plane)

Simplest idea: smear PDFs across a **uniform disk** of size πr_p^2
→ simple geometric overlap factor ≤ 1 in dijet cross section
Some collisions have the full overlap, others only partial
$\Longrightarrow$ Poisson distribution with different mean $\langle n_{\mathrm{MPI}} \rangle$ at each b

2. More realistic **Proton b-shape** (used by all modern MPI models)



Smear PDFs across a non-uniform disk
E.g., Gaussian(s), or **more**/less peaked (e.g., EM form factor)
Overlap factor = convolution of two such distributions

→ Poisson distribution with different mean <n> at each b
"Lumpy Peaks" → large matter overlap enhancements, higher <n>

Note: this is an *effective* description. Not the actual proton mass density.
E.g., peak in overlap function (»1) can represent unlikely configurations with huge overlap enhancement. Typically use total σ_{inel} as normalization.

MC with MPI vs Hadron Collisions

Plots from: Sjöstrand & v. Zijl, Phys.Rev.D36 **(1987)** 2019

Fluctuations in n_{mpi} → Bigger (global) fluctuations

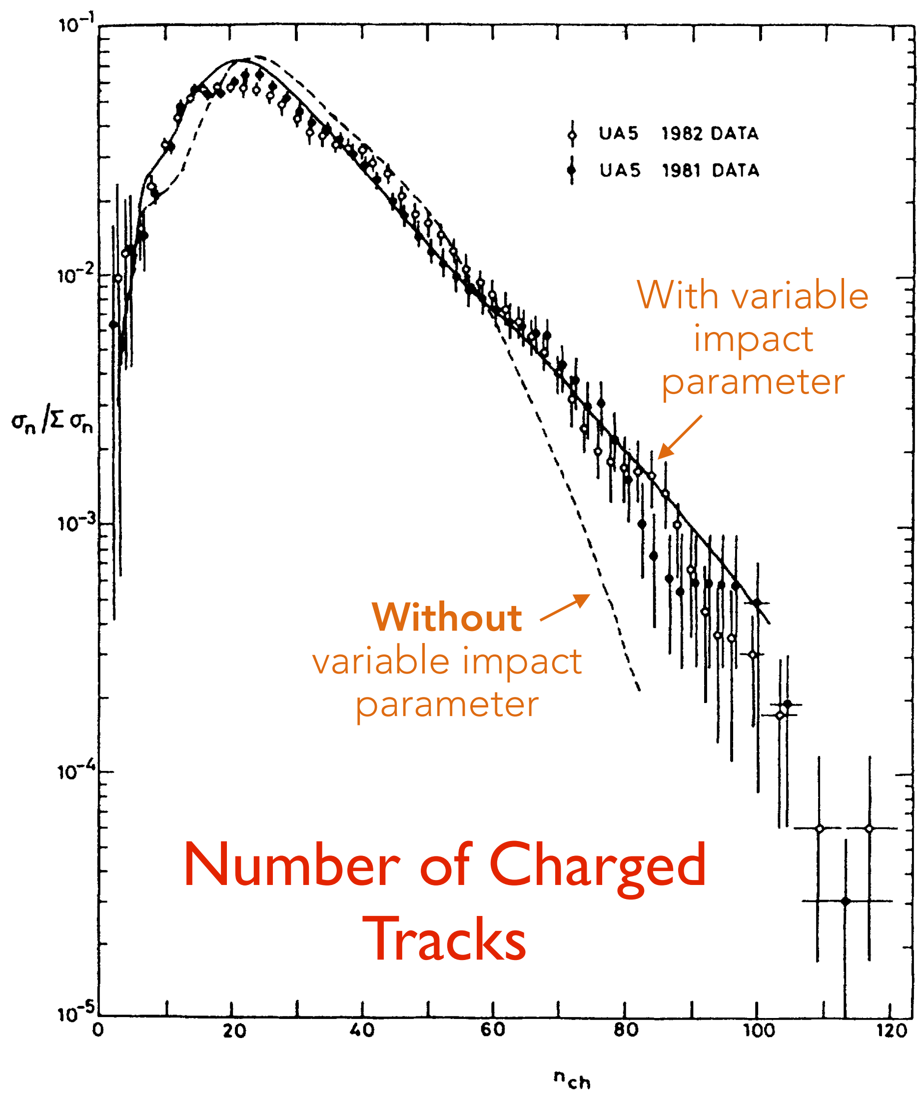
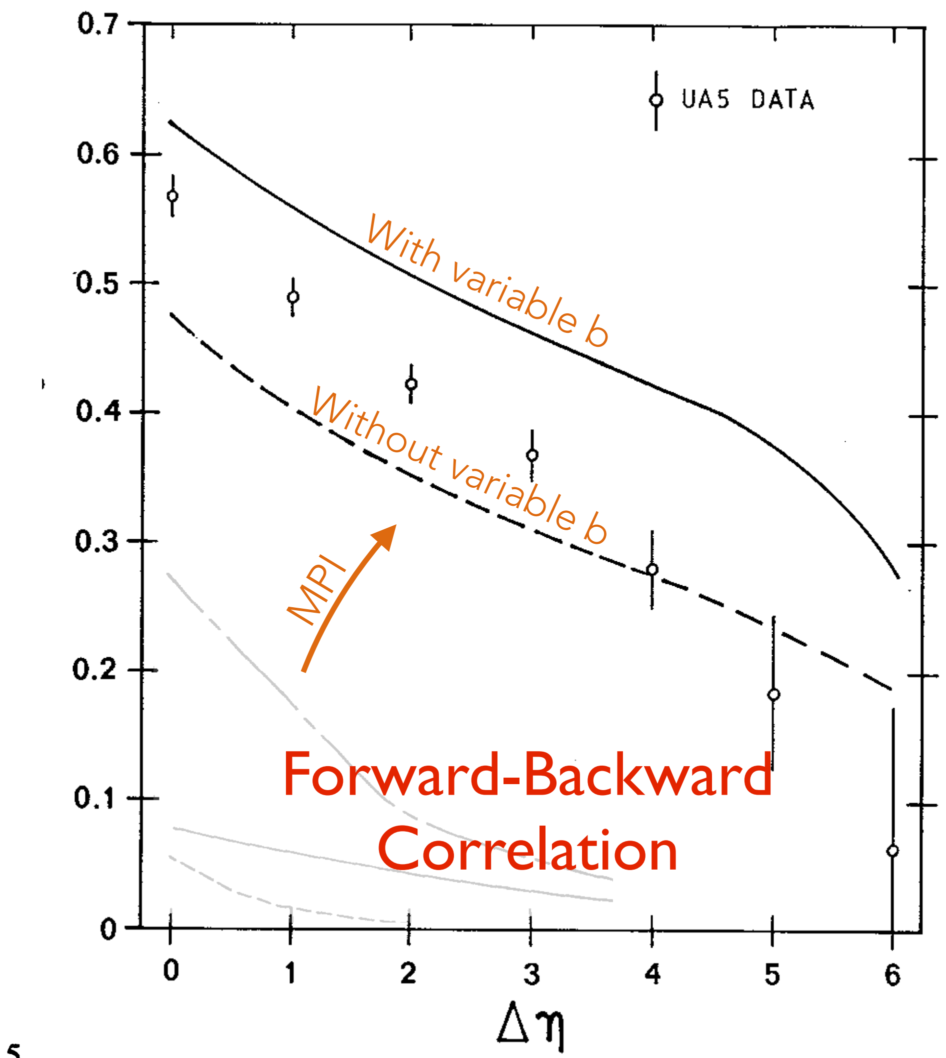
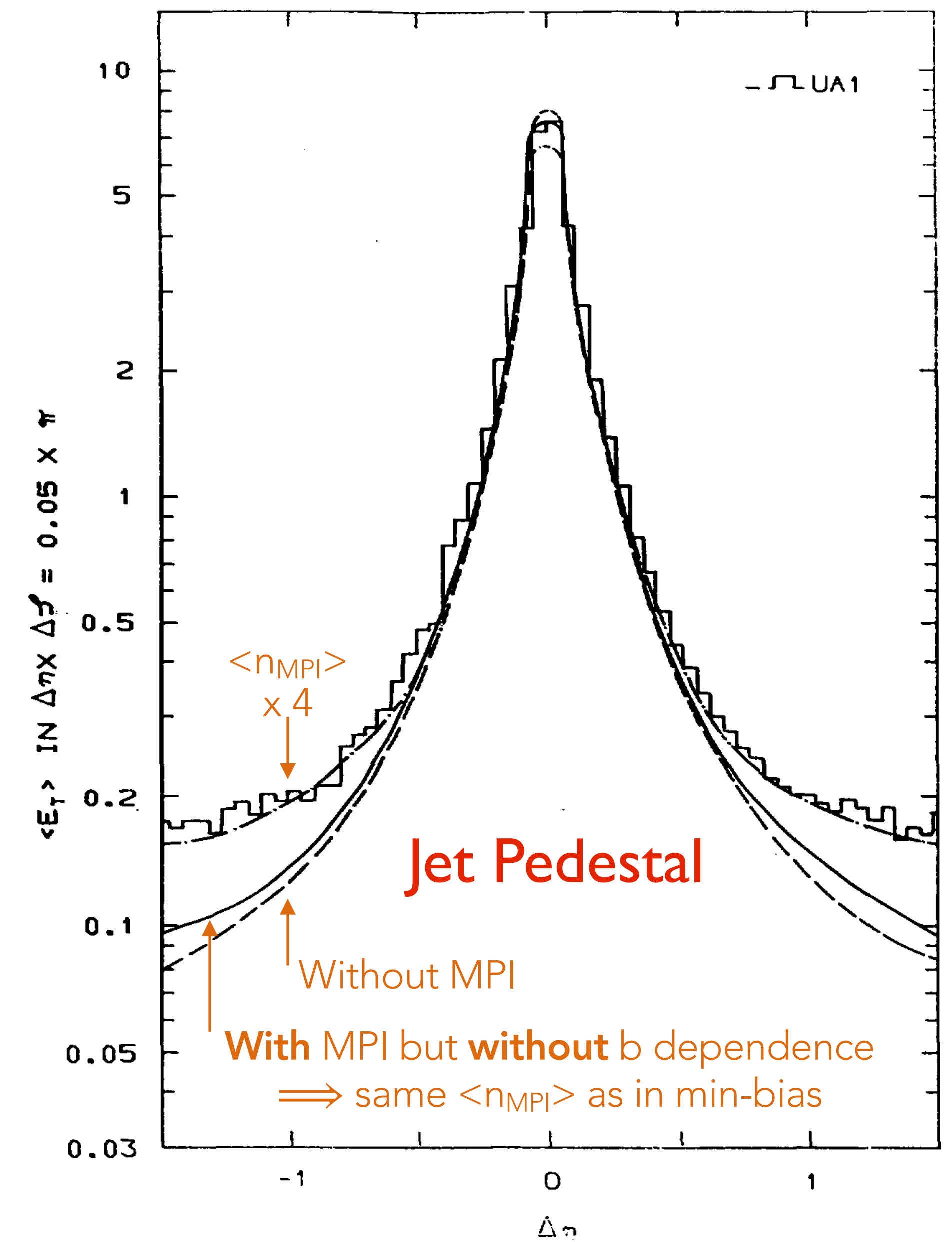


FIG. 12. Charged-multiplicity distribution at 540 GeV, UA5 results (Ref. 32) vs multiple-interaction model with variable impact parameter: solid line, double-Gaussian matter distribution; dashed line, with fix impact parameter [i.e., $\tilde{O}_0(b)$].



MPI → Long-distance correlations in rapidity



Impact-parameter dependence → UE

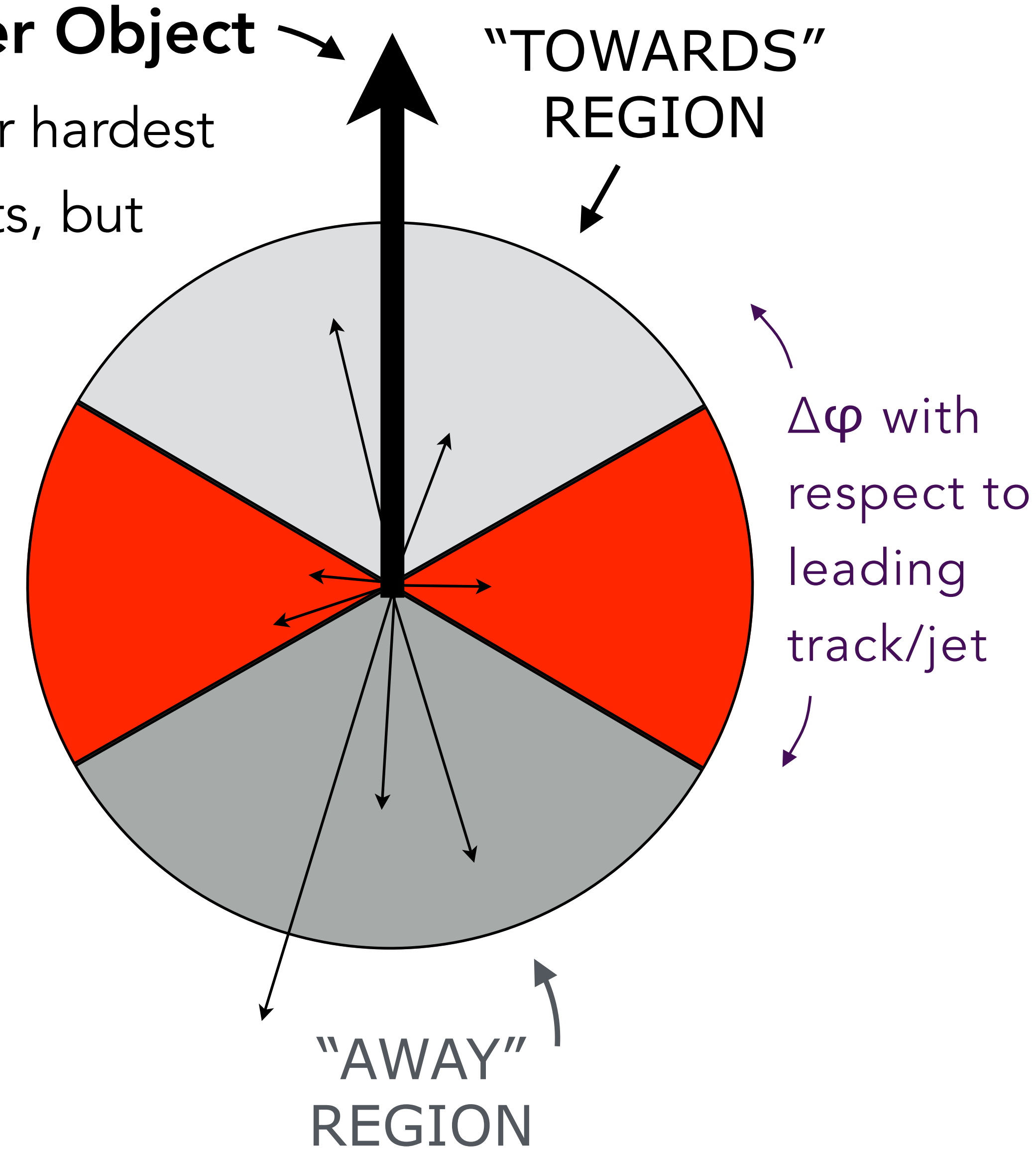
Characterising The Underlying Event

(The "Rick Field" UE Plots)

There are many UE variables.
The most important is <Σp_T> in the "Transverse Region"

Leading Trigger Object

E.g., **hardest jet**, hardest track, or hardest track-jet; more inclusive to use jets, but track-based analyses also useful.



"Transverse Region" (TRNS)

Sensitive to activity at right angles to the hardest jets

→ Useful definition of Underlying Event

Min-Bias VS Underlying Event

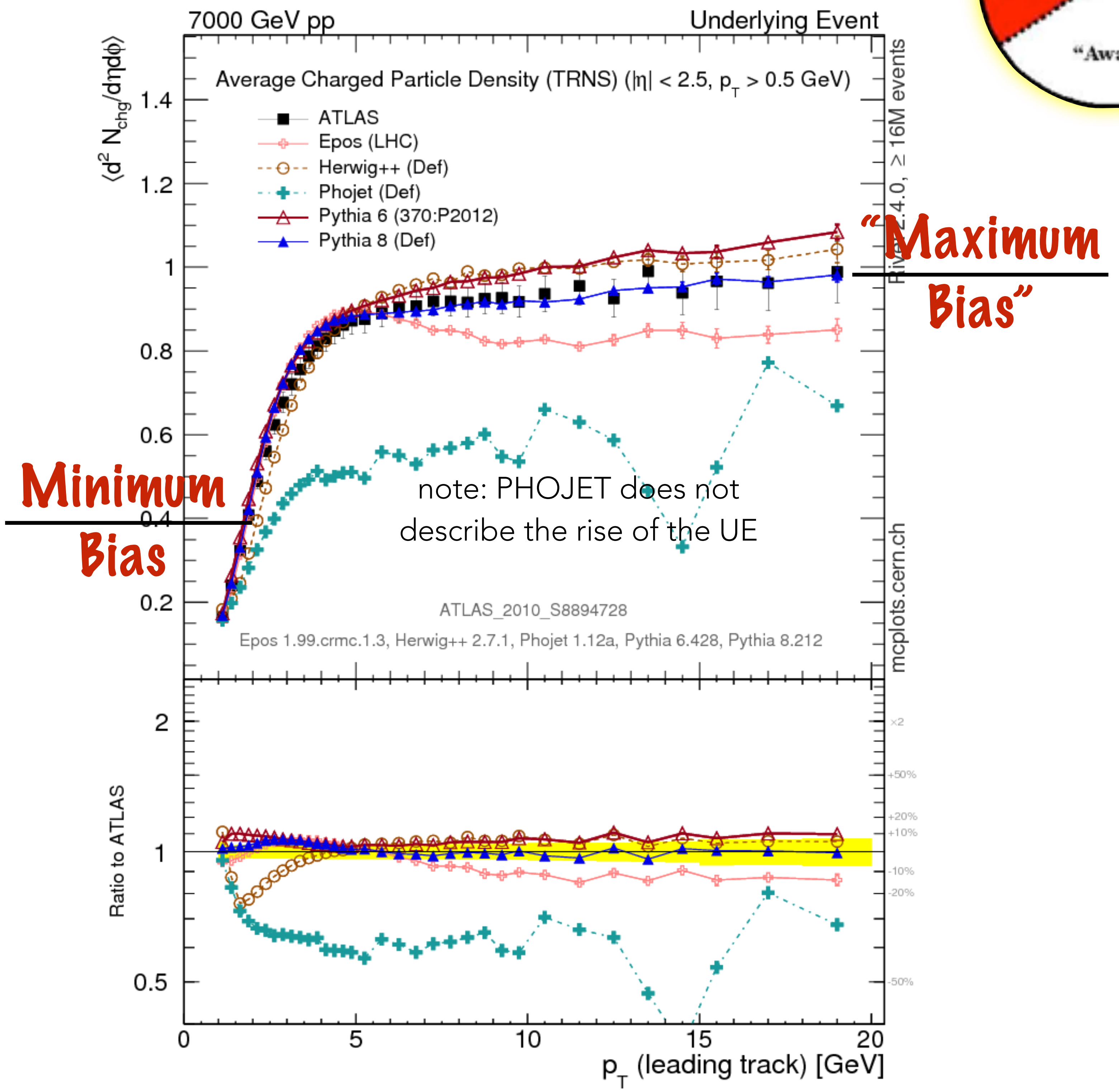
Tautology:

A jet trigger provides a bias
(→subsample of minimum-bias)

Pedestal effect:

Events with a hard jet trigger are accompanied by a higher plateau of ambient activity

MPI: interpreted as a biasing effect. Small pp impact parameters → larger matter overlaps → more MPI → higher chances for a hard interaction

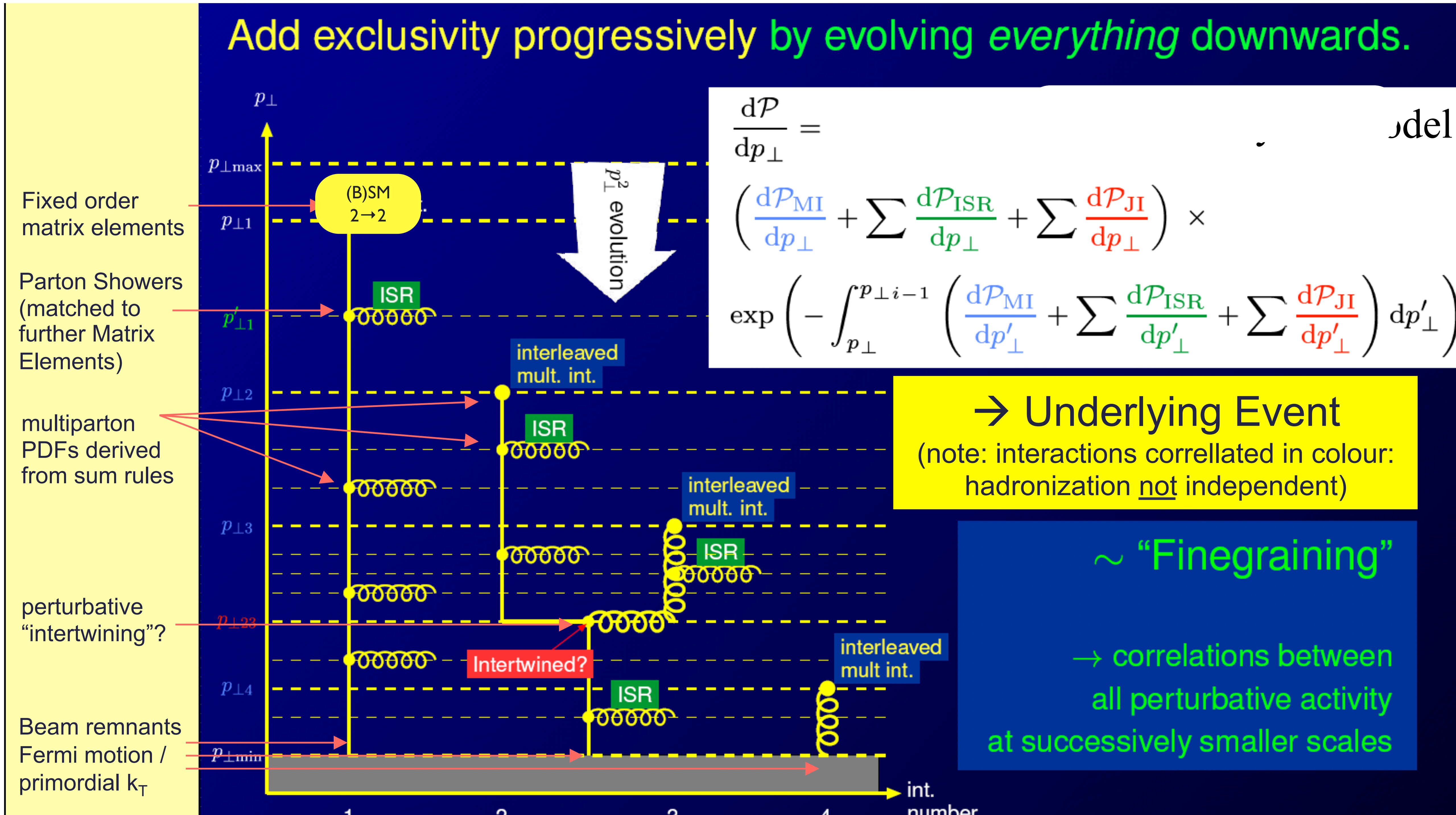


Plot from mcplots.cern.ch

Interleaved Evolution

The model in Pythia 8

Sjöstrand, P.S., JHEP 0403 (2004) 053; EPJ C39 (2005) 129

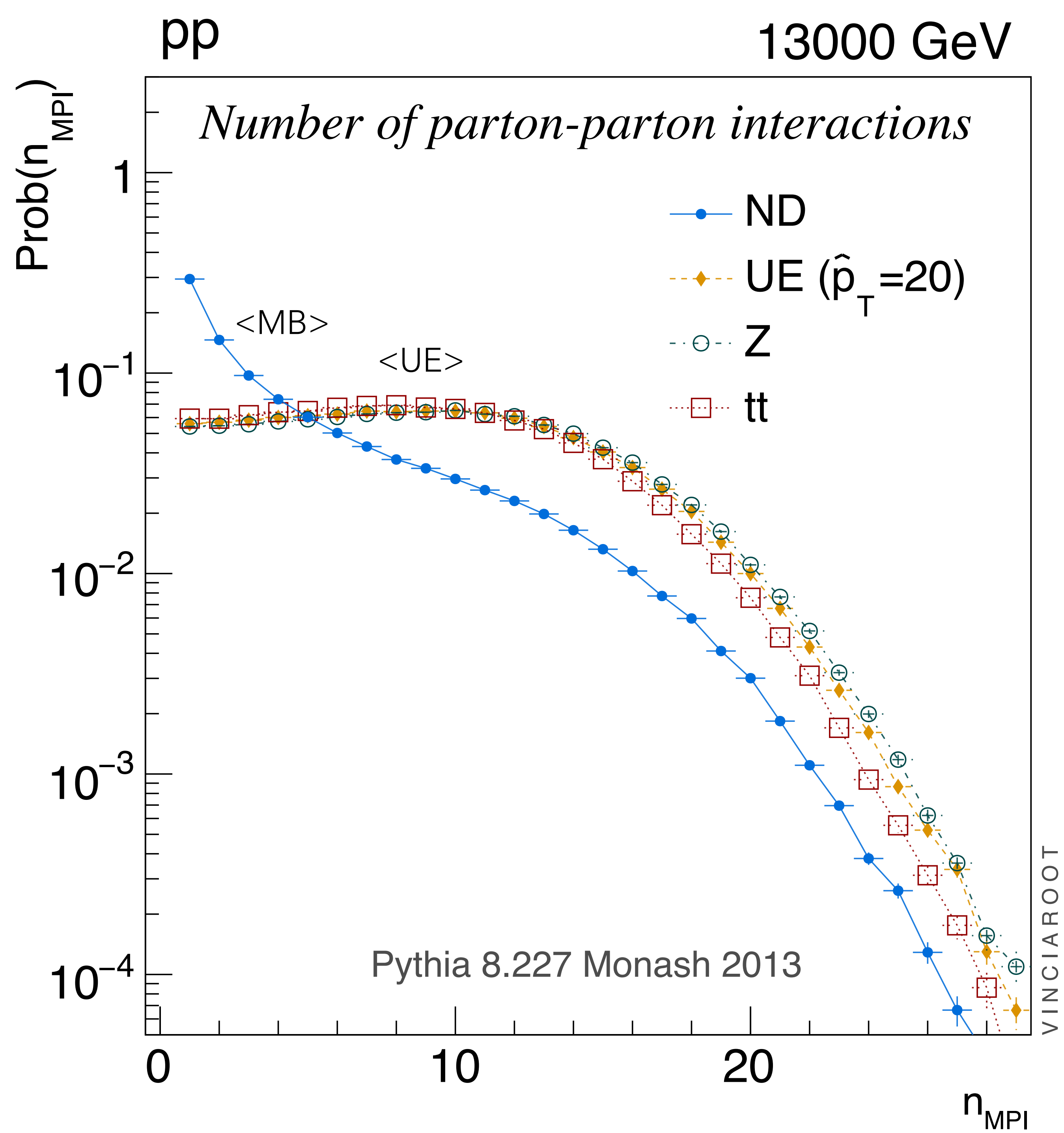


How many MPI are there? *

Example for pp collisions at 13 TeV — PYTHIA's default MPI model

Averaged over all pp impact parameters

(Really: averaged over all pp overlap enhancement factors)



*note: can be arbitrarily soft

Summary — Divide and Conquer

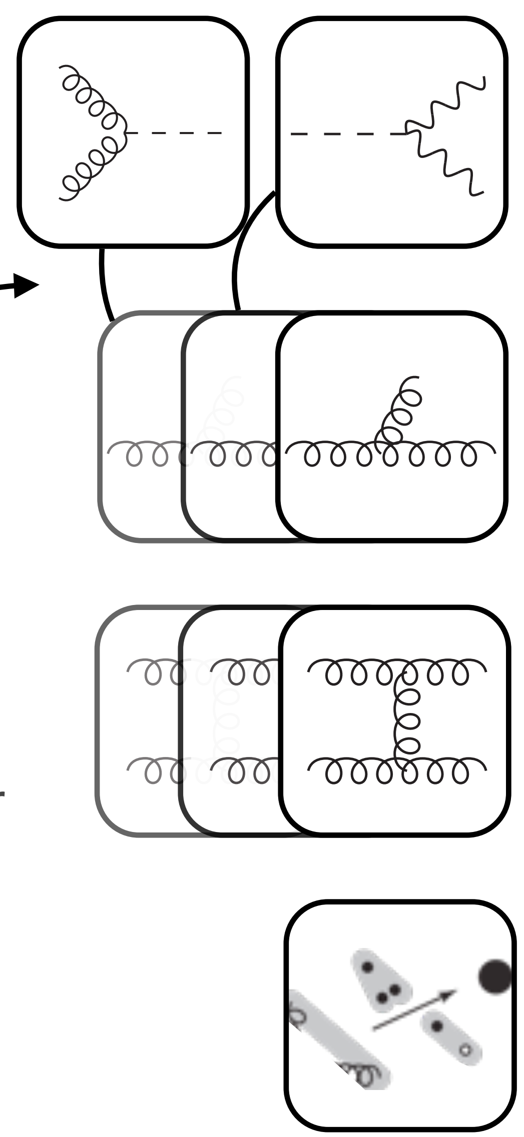
Physics Separation of time scales ➤ Factorisations Maths

➜ Can split **big** problem into many (nested) pieces + make random choices (MC)^2 ~ like in nature

$$\mathcal{P}_{\text{event}} = \mathcal{P}_{\text{hard}} \otimes \mathcal{P}_{\text{dec}} \otimes \mathcal{P}_{\text{ISR}} \otimes \mathcal{P}_{\text{FSR}} \otimes \mathcal{P}_{\text{MPI}} \otimes \mathcal{P}_{\text{Had}} \otimes \ldots$$

Merging

Eliminate double-counting between fixed-order and shower corrections



Hard Process & Decays:

Use process-specific (N)LO matrix elements (e.g., $gg \to H^0 \to \gamma\gamma$)

→ Sets "hard" resolution scale for process: Q_{HARD}

ISR & FSR (Initial- & Final-State Radiation):

Driven by differential (e.g., DGLAP) evolution equations, dP/dQ^2, as function of resolution scale; from Q_{HARD} to Q_{HAD} ~ 1 GeV

MPI (Multi-Parton Interactions)

Protons contain lots of partons → can have additional (soft) parton-parton interactions → Additional (soft) "Underlying-Event" activity

Hadronisation

Non-perturbative modeling of partons → hadrons transition

Final Remarks — Why study QCD?

We tend to focus on how useful it is, essential even, to collider phenomenology (MC tools, N^{n}LO, etc)**.**

But let's not forget how pretty it is

And how little we still know about it

… especially beyond fixed order

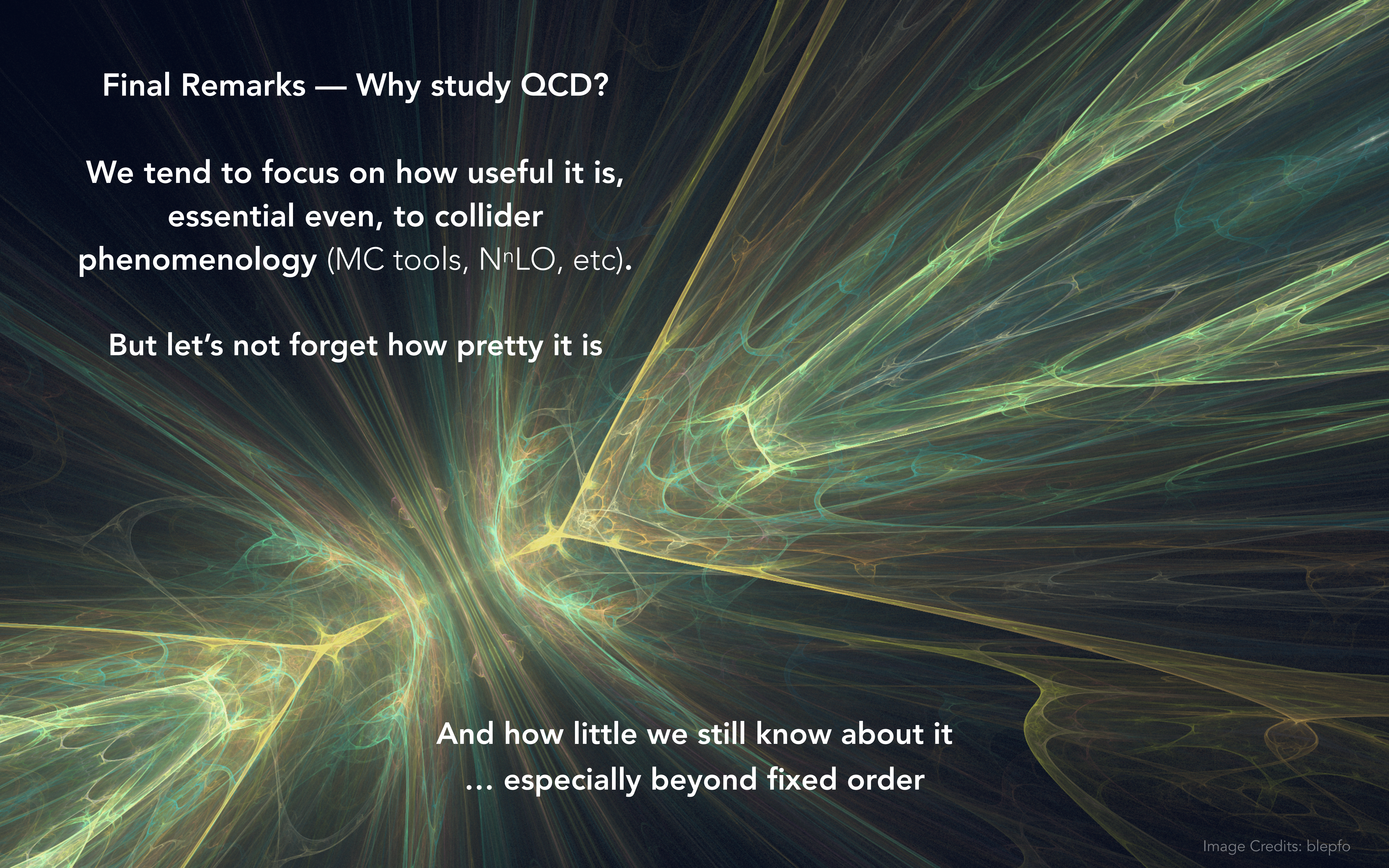


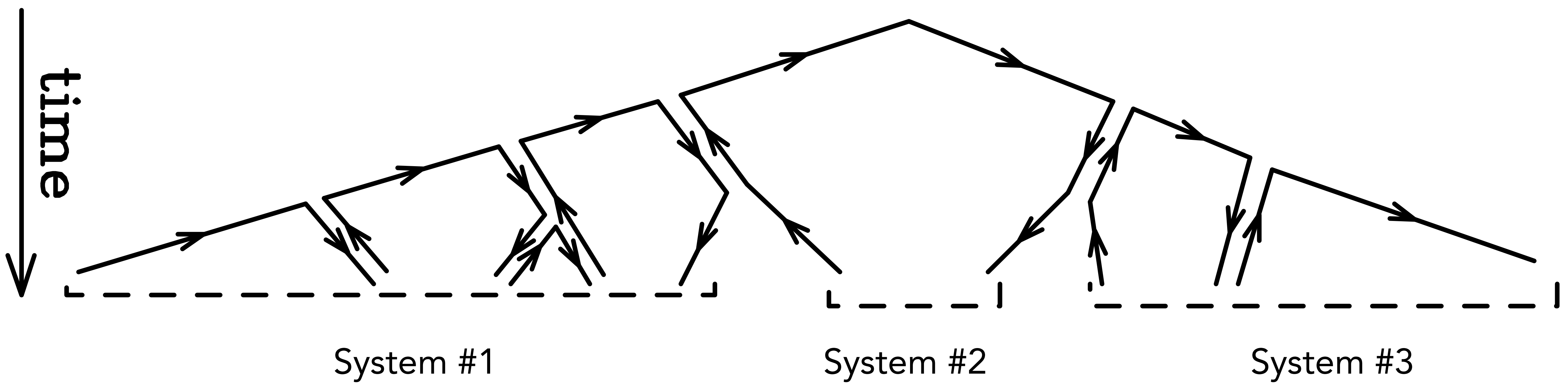
Image Credits: blepfo

Extra Slides

RECAP: Colour Flow

Colour flow in parton showers (in leading-colour approximation)

Example: $Z^0 \rightarrow q\bar{q}$



Coherence of pQCD cascades → not much "overlap" between systems
→ Leading-colour approximation pretty good

(LEP measurements in $e^+e^- \rightarrow W^+W^- \rightarrow$ hadrons confirm this (at least to order 10% ~ $1/N_c^2$))

Note: (much) more color getting kicked around in hadron collisions. More tomorrow.

In **Spacetime**:

String tension ≈ 1 GeV/fm → a 5-GeV quark can travel 5 fm before all its kinetic energy is transformed to potential energy in the string.

Then it must start moving the other way (→ "yo-yo" model of mesons. Note: string breaks → several mesons)

In **Rapidity** :

(convenient variable in momentum space)

$$y = \frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right) = \frac{1}{2}\ln\left(\frac{(E+p_z)^2}{E^2-p_z^2}\right)$$

$$\left(\text{for } m \to 0 \ : \ \frac{1}{2}\ln\left(\frac{1+\cos\theta}{1-\cos\theta}\right) \ = \ -\ln\tan(\theta/2) \ = \ \eta \right)$$

"Pseudorapidity"

Rapidity is useful because it is additive under Lorentz boosts (along the rapidity axis)

$$y' = y + \ln\sqrt{\frac{1-\beta}{1+\beta}}$$

⇛ Δy **difference** is invariant

Particle Production:

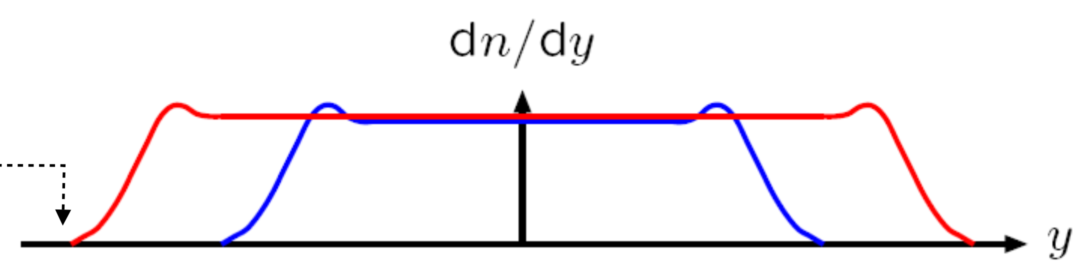
If the quark gives all its energy to a single pion traveling along the z axis

$$y_{\max} \sim \ln\left(\frac{2E_q}{m_\pi}\right)$$

Increasing E_q → logarithmic growth in rapidity range

Scaling in lightcone $p_\pm = E \pm p_z$

⇛ flat central rapidity plateau (+ some endpoint effects)



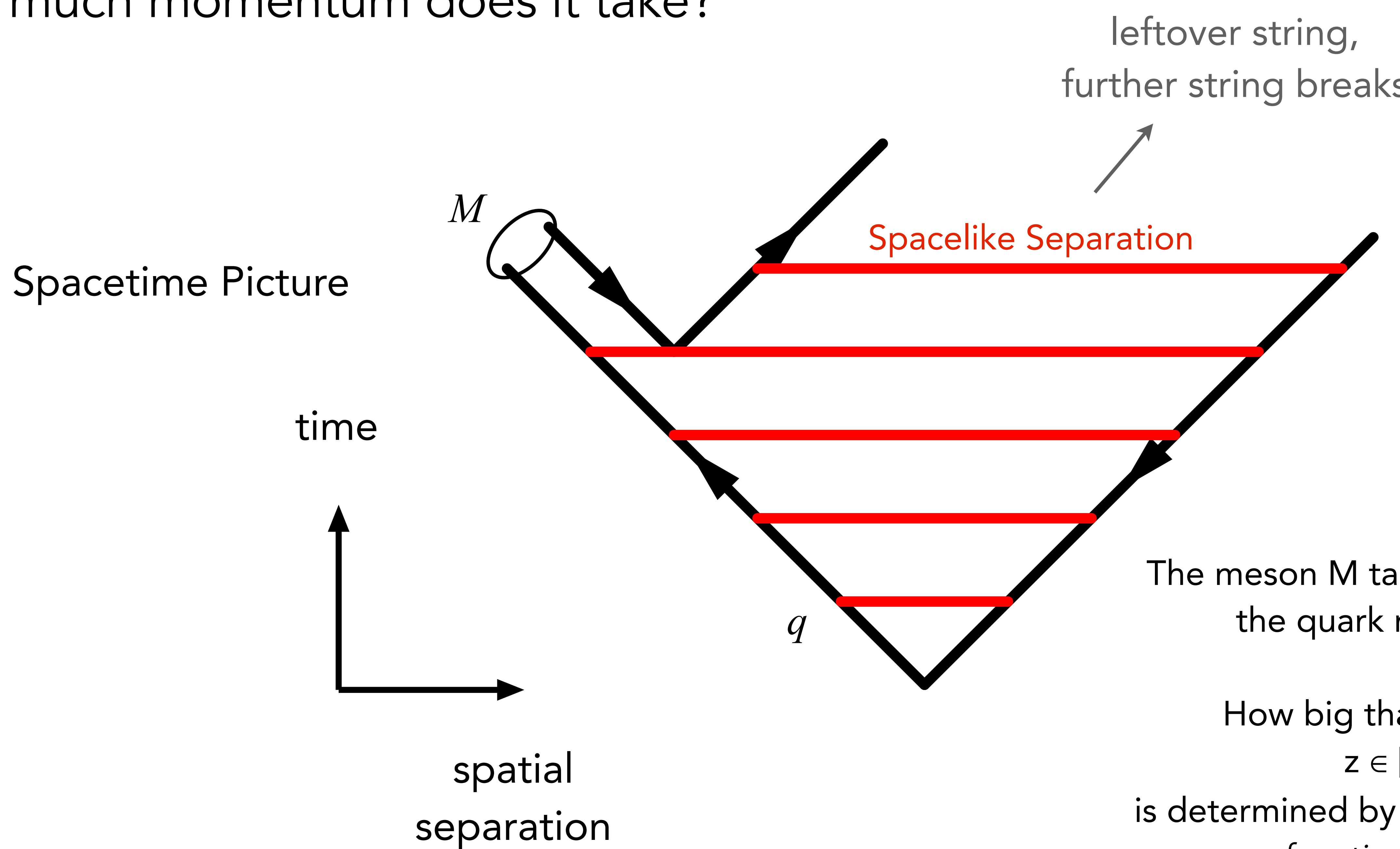
$\langle n_{\mathrm{ch}} \rangle \approx c_0 + c_1 \ln E_{\mathrm{cm}}$, ~ Poissonian multiplicity distribution

Fragmentation Function

Having selected a hadron flavor

(see lecture notes for how selection is made between different spin/excitation states)

How much momentum does it take?

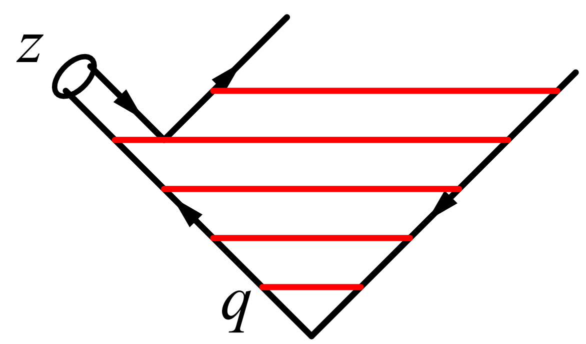


The meson M takes a fraction z of the quark momentum,

How big that fraction is,
$z \in [0,1]$,
is determined by the *fragmentation function*, $f(z, Q_0^2)$

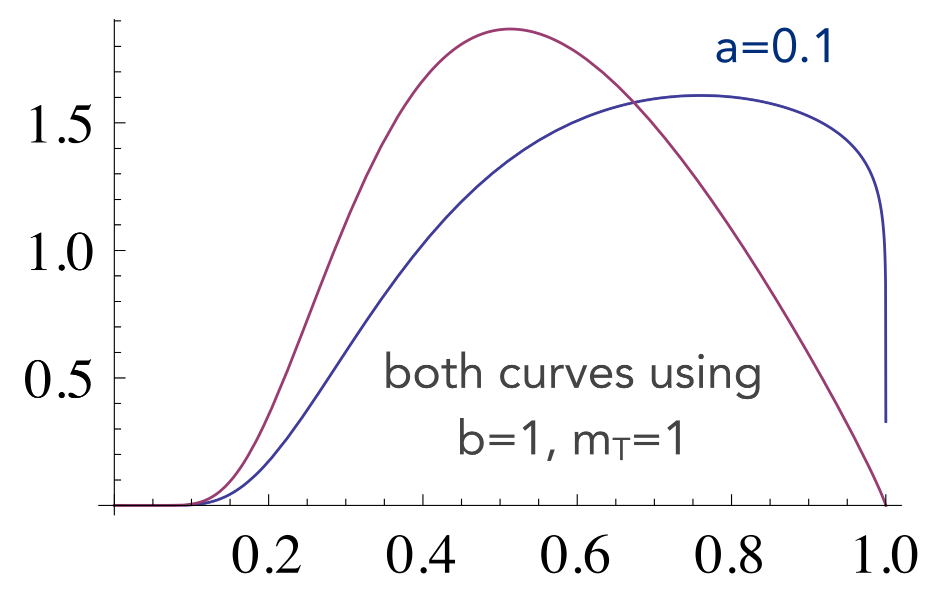
Causality → Left-Right Symmetry

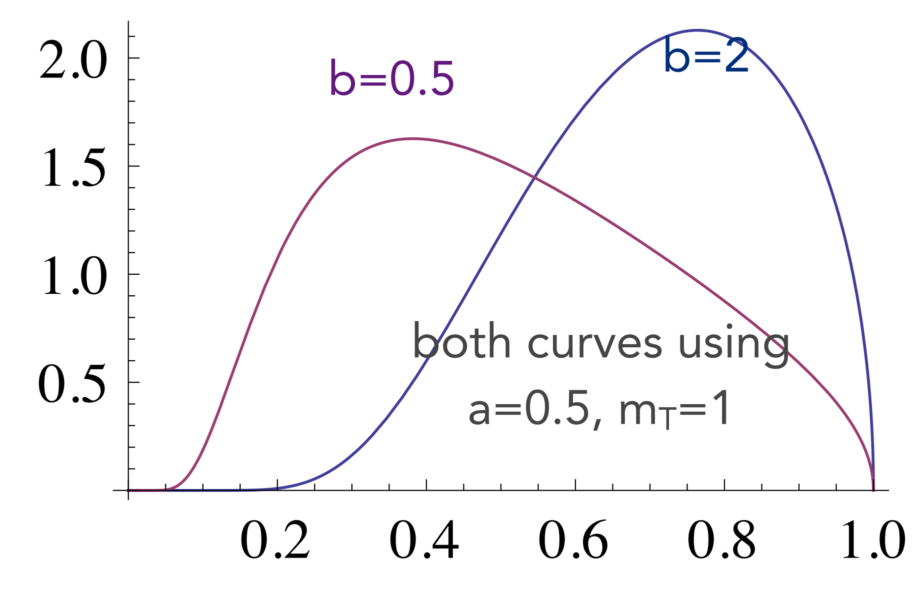
→ Constrains form of fragmentation function!



→ Lund Symmetric Fragmentation Function

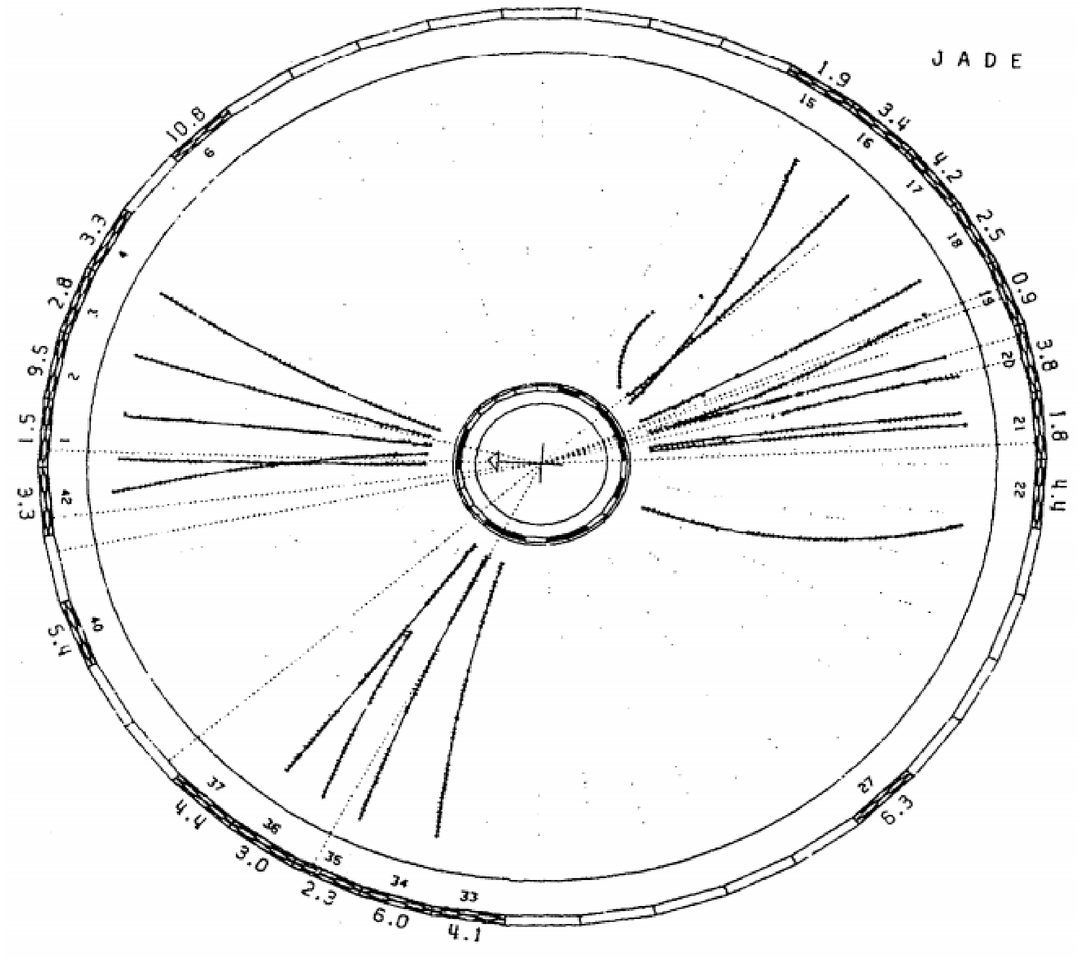
$$f(z) \propto \frac{1}{z}(1-z)^a \exp\left(-\frac{b\,(m_h^2+p_{\perp h}^2)}{z}\right)$$

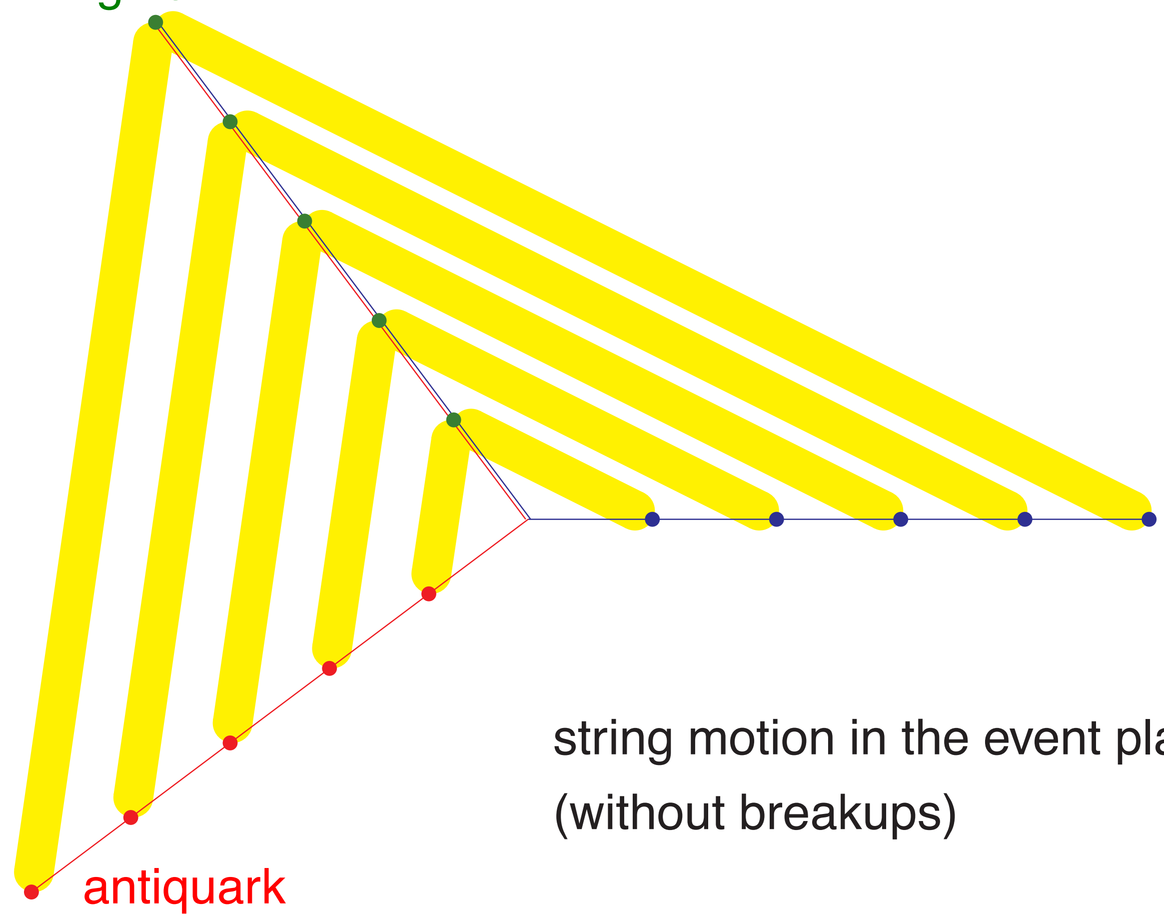




Note: In principle, a can be flavour-dependent. In practice, we only distinguish between baryons and mesons

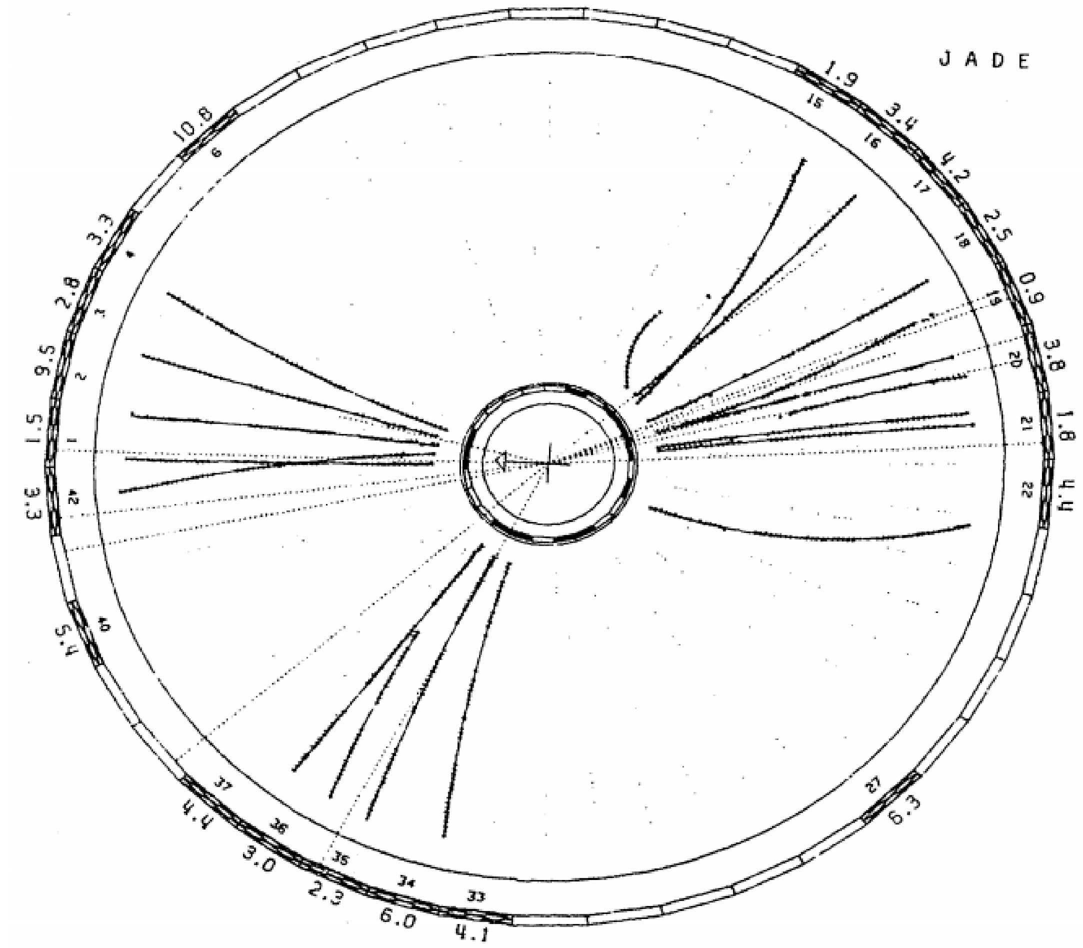
1980: string (colour coherence) effect

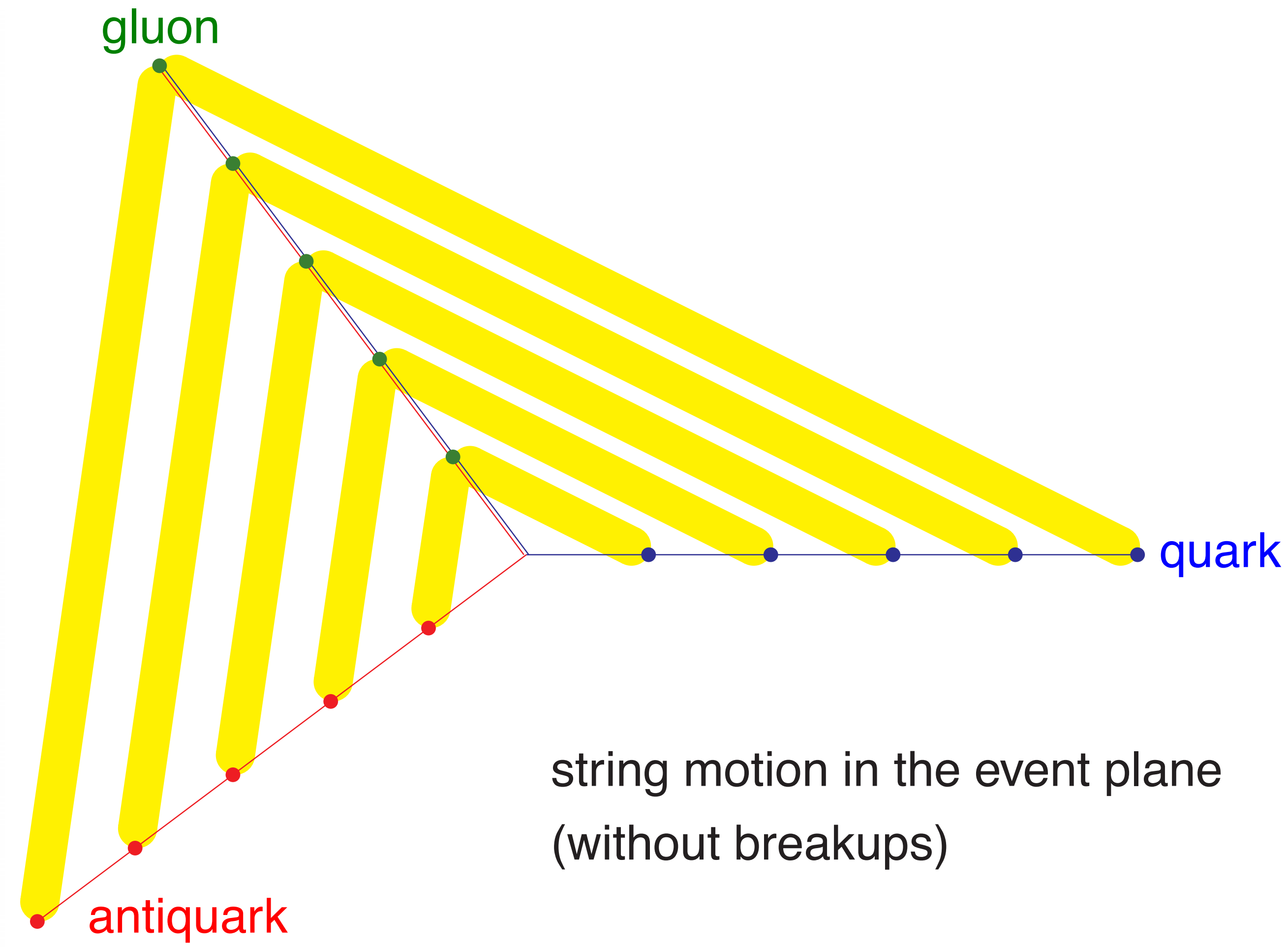


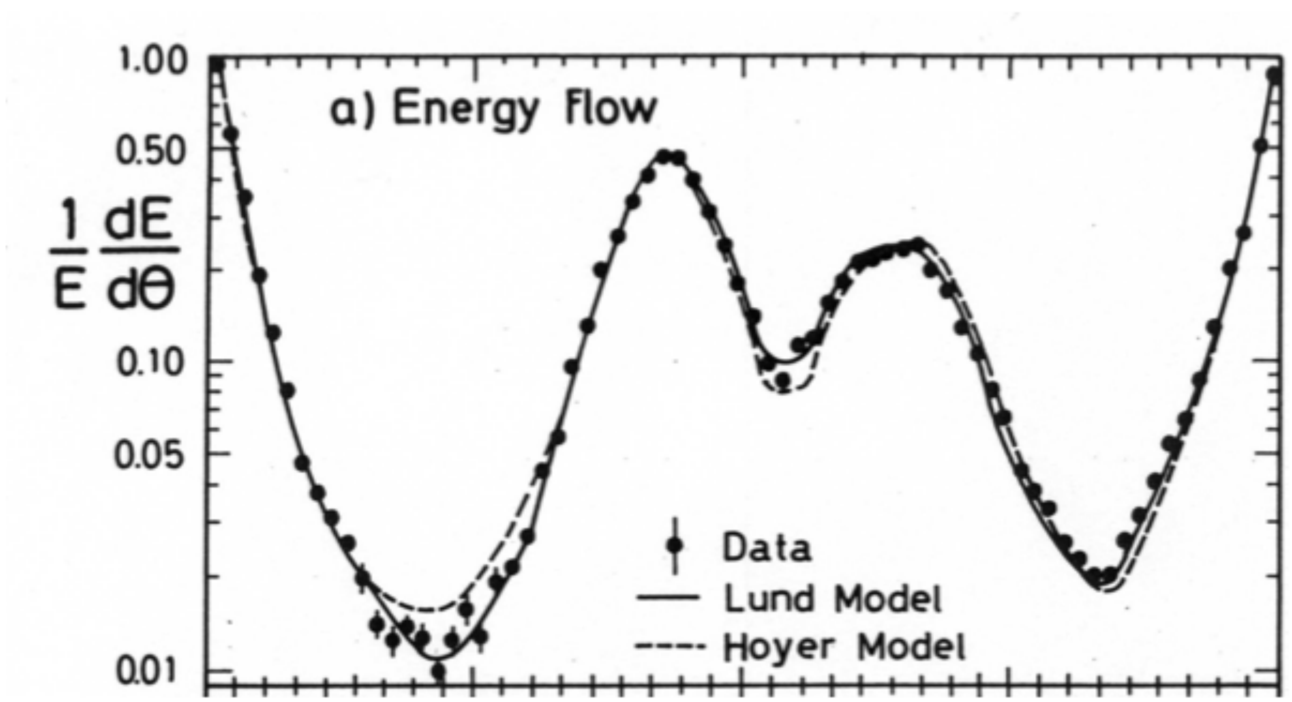


string motion in the event plane
(without breakups)

1980: string (colour coherence) effect







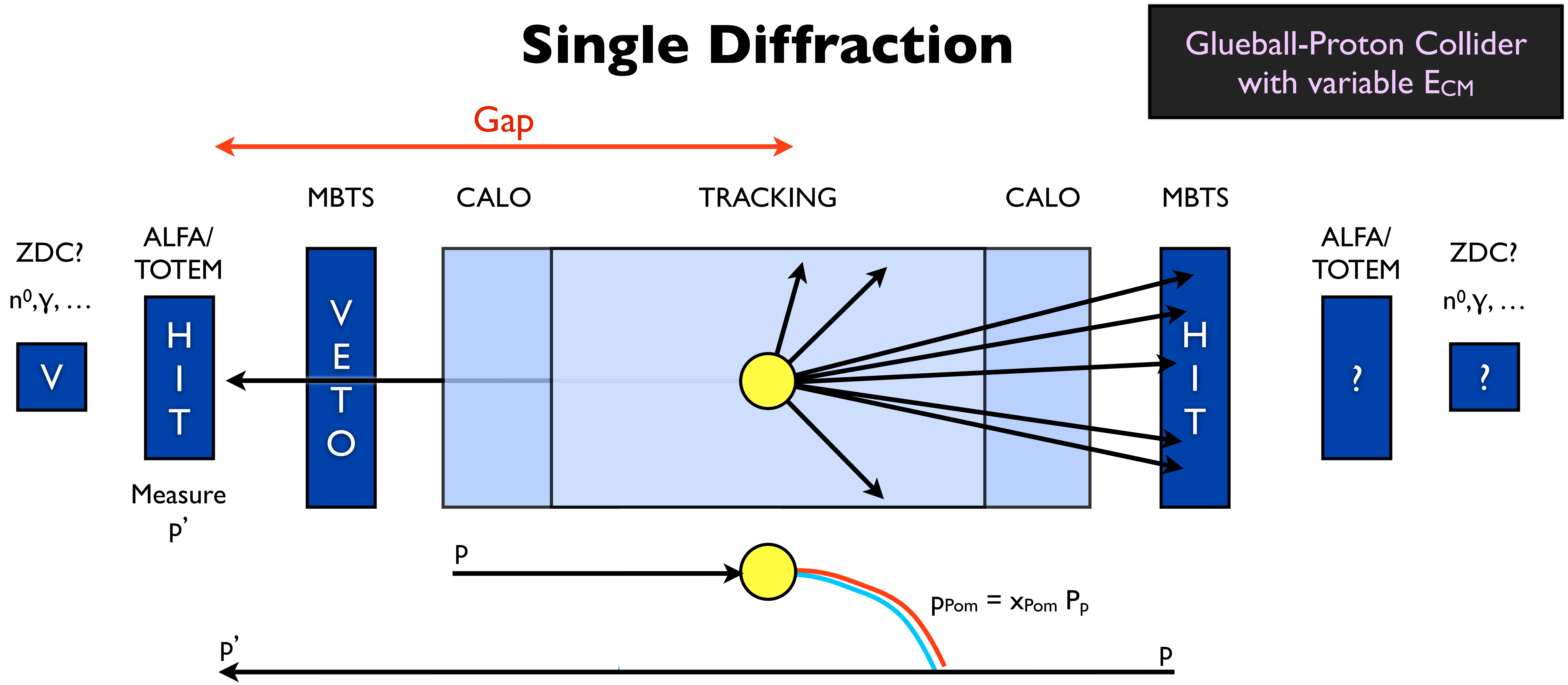
Predicted unique event structure; inside & between jets.

Confirmed first by JADE 1980.

Generator crucial to sell physics!

(today: PS, M&M, MPI, . . .)

(Aside: What is diffraction?)



Also:
"Double Diffraction": both protons explode; defined by gap inbetween
"Central Diffraction": two protons + a central (exclusive) system

1: A Simple Model

A minimal model incorporating single-parton factorization, perturbative unitarity, and energy-and-momentum conservation

Take literally

$$\sigma_{2\to2}(p_{\perp\min}) = \langle n\rangle(p_{\perp\min})\ \sigma_{\mathrm{tot}}$$

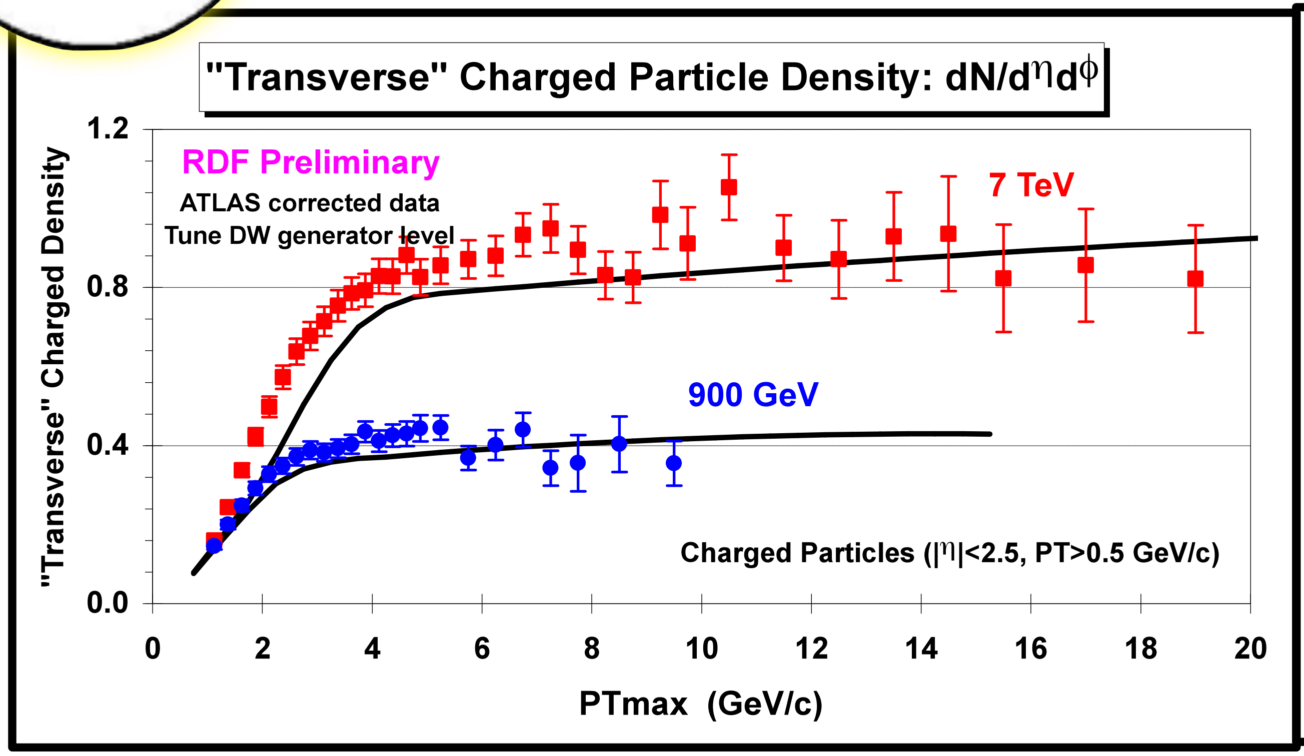
Parton-Parton Cross Section　　　　Hadron-Hadron Cross Section

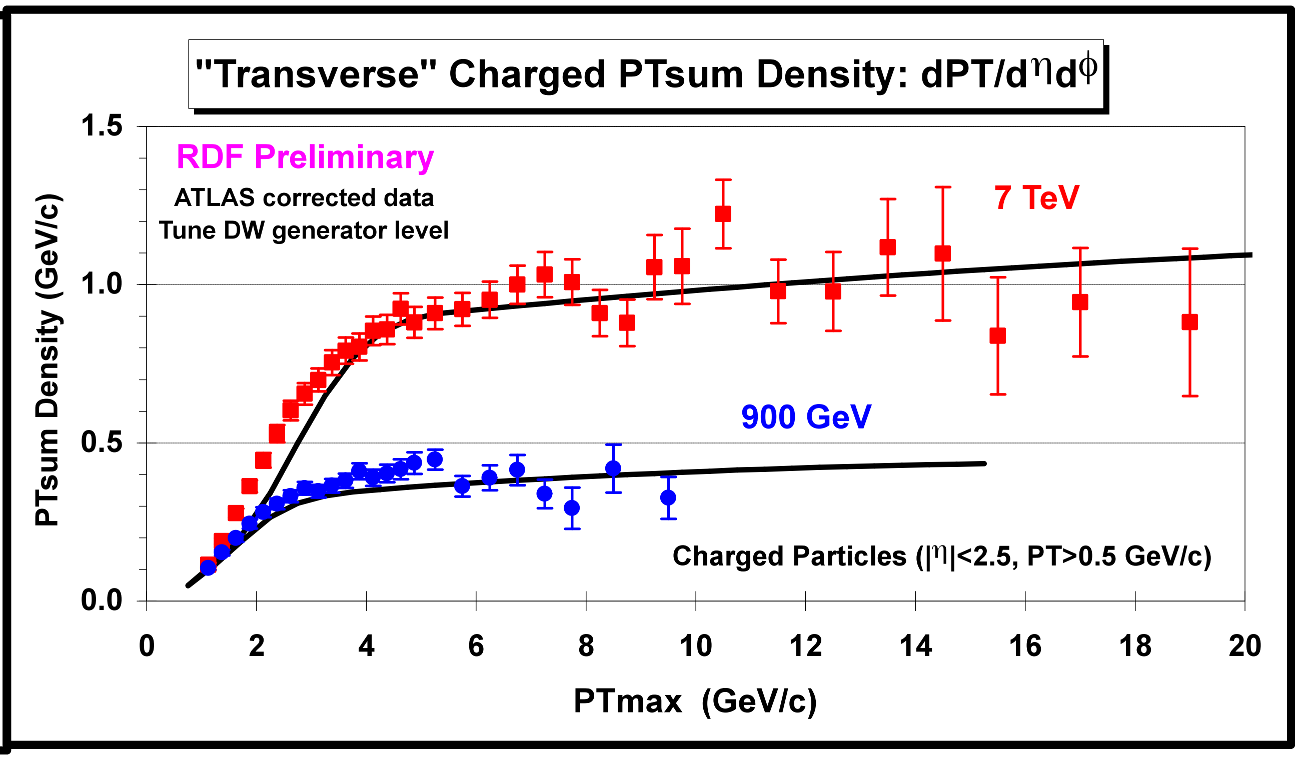
1. Choose $p_{T\min}$ cutoff
   = main tuning parameter
2. Interpret $<n>(p_{T\min})$ as mean of Poisson distribution
   Equivalent to assuming all parton-parton interactions equivalent and independent ~ each take an instantaneous "snapshot" of the proton
3. Generate n parton-parton interactions (pQCD 2→2)
   Veto if total beam momentum exceeded → overall (E,p) cons
4. Add impact-parameter dependence → $<n> = <n>(b)$
   Assume factorization of transverse and longitudinal d.o.f., → PDFs : $f(x,b) = f(x)g(b)$ ← Ordinary CTEQ, MSTW, NNPDF, …
   b distribution ∝ EM form factor → JIMMY model (F77 Herwig) ⟵ Butterworth, Forshaw, Seymour Z.Phys. C72 (1996) 637
   Constant of proportionality = second main tuning parameter
5. Add separate class of "soft" (zero-p_T) interactions representing interactions with $p_T < p_{T\min}$ and require $\sigma_{\mathrm{soft}} + \sigma_{\mathrm{hard}} = \sigma_{\mathrm{tot}}$
   → **Herwig 7 model**　Bähr et al, arXiv:0905.4671

The Pedestal

(now called the Underlying Event)

LHC from 900 to 7000 GeV - ATLAS





Track Density (TRANS)

(Not Infrared Safe)

Large Non-factorizable Corrections

Prediction off by ≈ 10%

Sum(pT) Density (TRANS)

(more) Infrared Safe

Large Non-factorizable Corrections

Prediction off by < 10%

Truth is in the eye of the beholder:

R. Field: "See, I told you!"

Y. Gehrstein: "they have to fudge it again"