# Physics of Monte Carlo Event Generators

**Peter Skands** — Monash University

1

## Lecture 1: Perturbative Aspects

From Perturbation Theory to Parton Showers

Fixed-Order Improvements: Matching & Merging

Perturbative Uncertainties

2

## Lecture 2: Nonperturbative Aspects

This Lecture

Hadronization ➤ **Dynamics of Confinement**

Hadronic (pp, pA, AA) Collisions ➤ **"Collective Phenomena"**

New Discoveries ➤ **New Ideas**

# Divide and Conquer Caesar

**Hard LHC collisions** contain **100s of particles**

Need (differential) $\sigma_{pp}$ for that number of "legs"

**Help! Some of them are hadrons!** ← Non-perturbative

And/or have **small opening angles**

And/or are **"soft"** ← Perturbative Infinities

+ Phase Space $\propto \prod_{i=1}^{100} \frac{\mathrm{d}^3 p_i}{2E_i}$ ← Big

How would *you:*

Construct, square, and integrate 100-leg amplitudes (with a lot of IR-divergent + non-pert. structure) over 300-dimensional phase spaces?

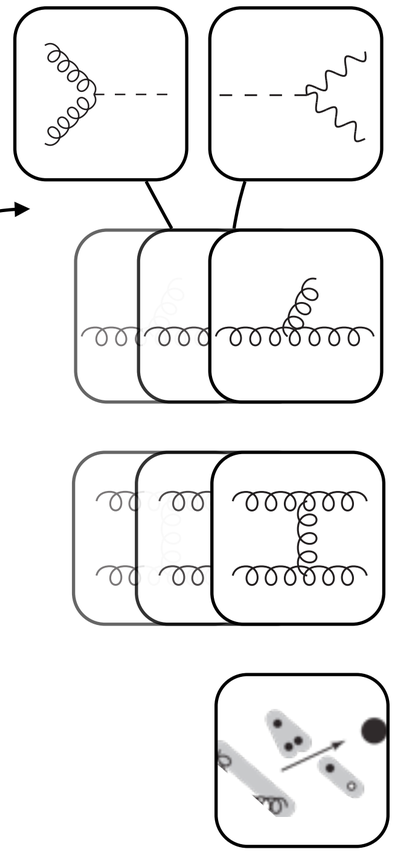
➤ **break it down!**

# Breaking it down

Physics — Separation of time scales ➤ Factorizations — Maths

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

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



Matching & 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

Nonperturbative modeling of partons → hadrons transition

Strings or clusters; followed by hadron and $\tau$ decays

# The Physics of Event Generators

In high-energy processes, need a dynamical process to ensure partons (**quarks and gluons**) become **confined** within hadrons

i.e. a non-perturbative **parton → hadron map**

**Model requirements**

- ➢ Colour neutralisation
- ➢ Dynamical mapping to on-shell hadrons

First Topic:
**Confinement**

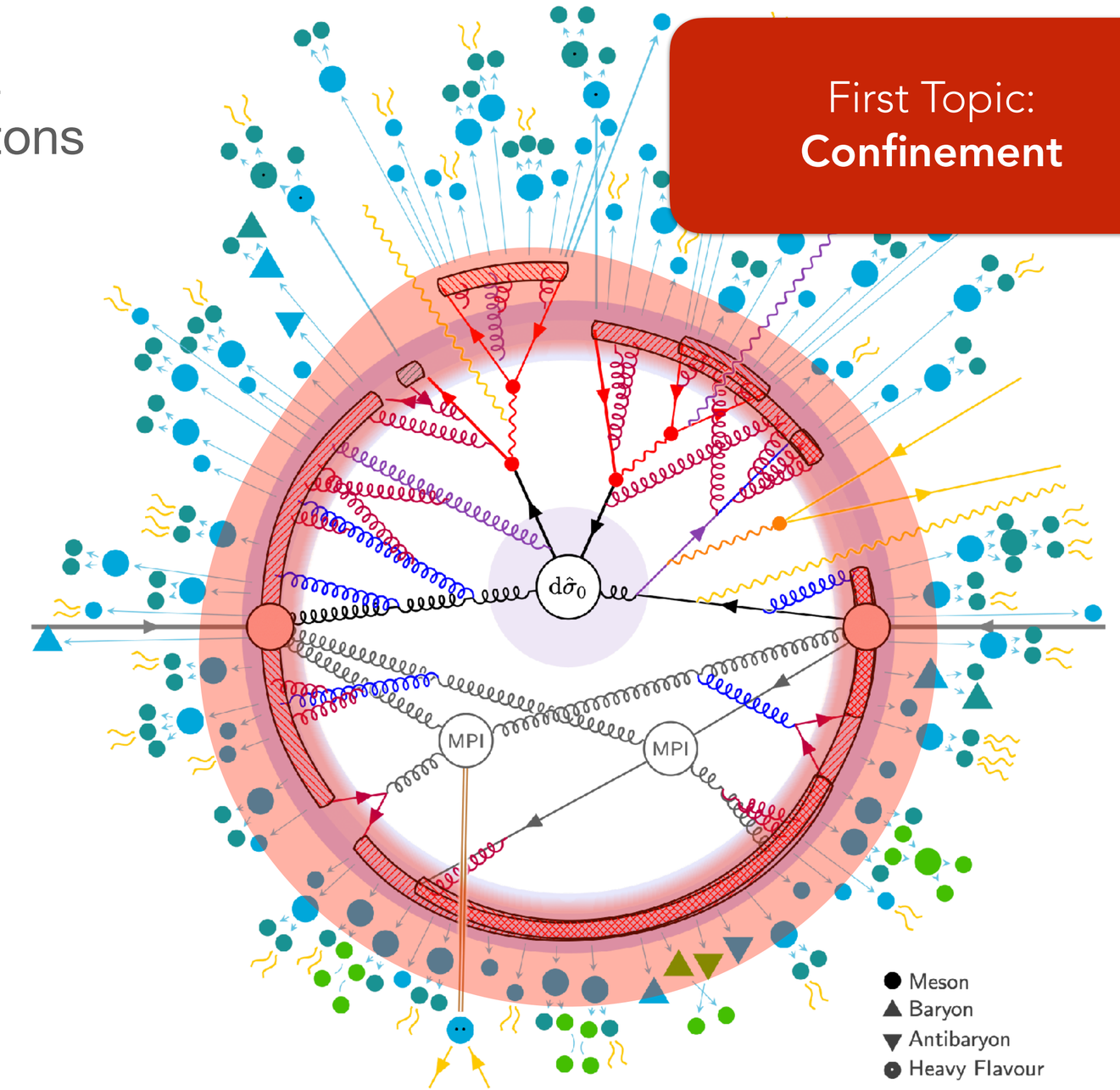



Figure from arXiv:2203.11601

# Requirement #1: **Colour Neutralisation**

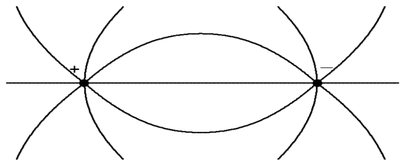
The point of confinement is that partons are **coloured**

A **physical** model needs two or more partons to create **colour-neutral** objects

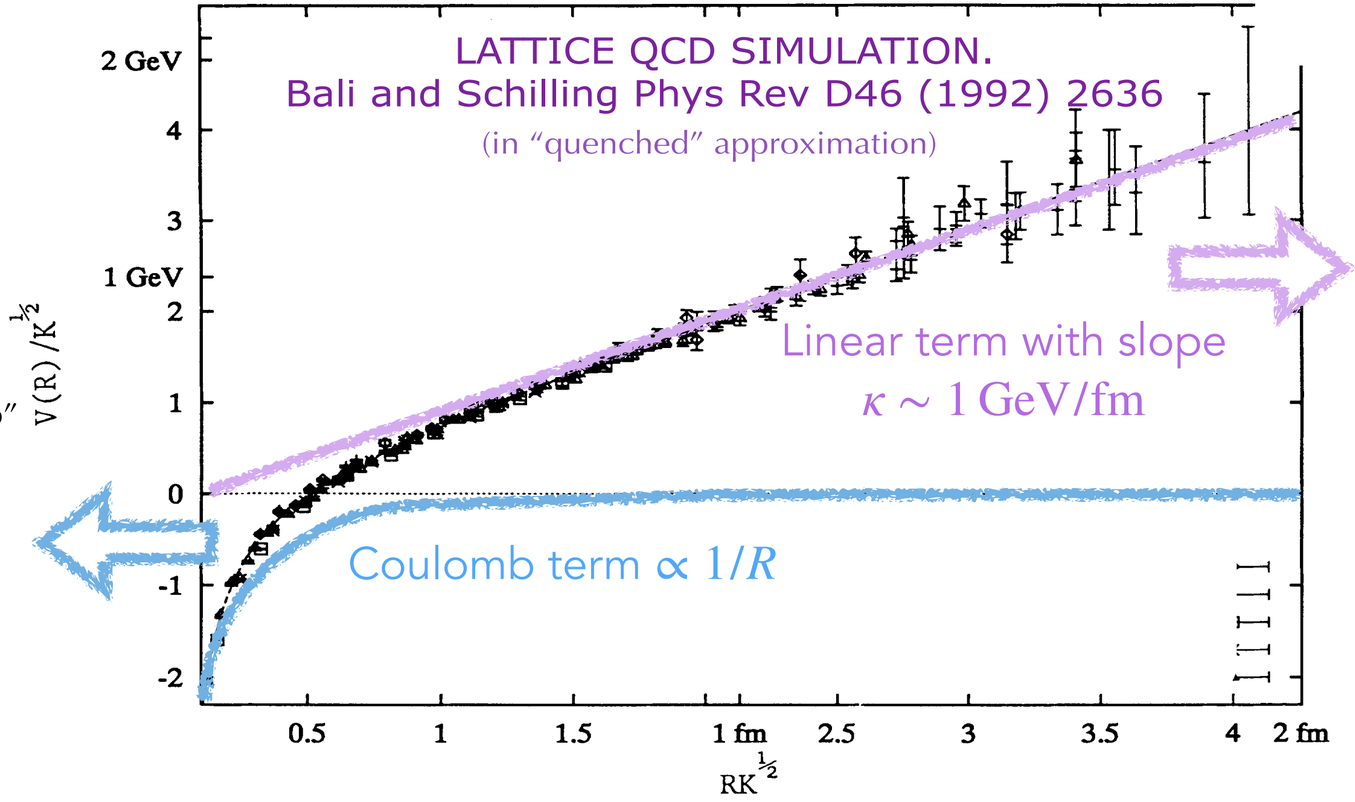
On lattice, compute **potential energy** $V(R)$ of a **colour-singlet** $q\bar{q}$ **state**

as function of the distance, $R$, between the $q$ and $\bar{q}$:

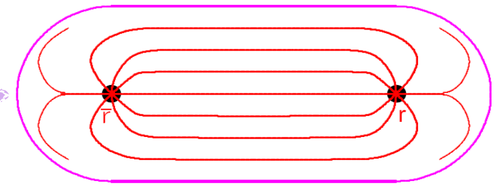
Short Distances ~ "Coulomb"



"Free" Partons

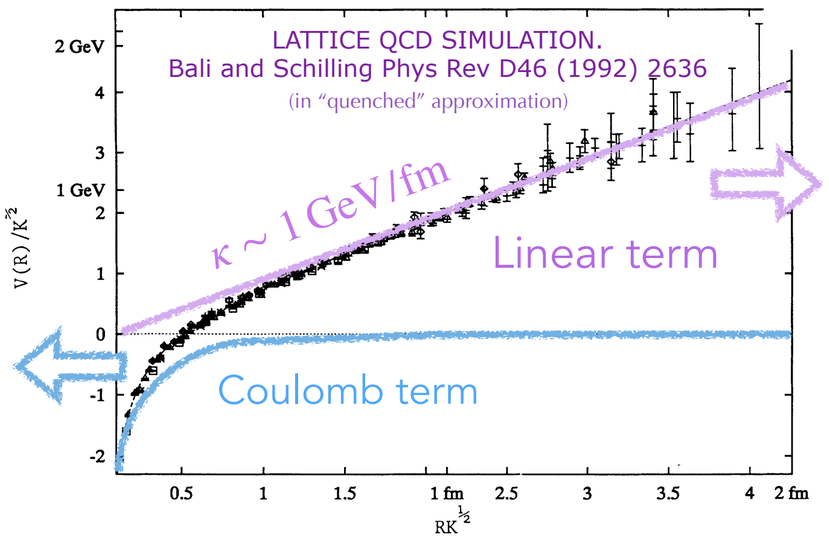



Long Distances ~ Linear Potential



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

# Strings!!




## Lund string model

model the **colour confinement field** as a **string**

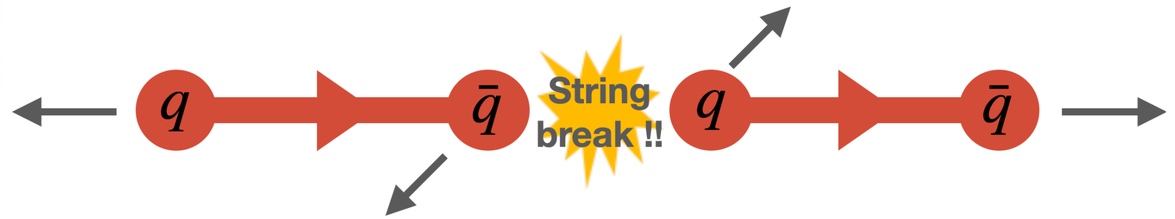
➢ Strings form between partons that form overall **colour-singlet** states




(+ Characteristic Feature of Lund Model: gluons are mapped to transverse kinks)

next slides

**High separation energies $\gtrsim$ 1 GeV**
$\Longrightarrow$ String Breaks (by pair creation):




Modelled by analogy with "Schwinger Mechanism" in QED

$\Longrightarrow$ **Gaussian suppression with "transverse mass":** $\exp\left(\frac{-m_q^2 - p_{\perp q}^2}{\kappa/\pi}\right)$

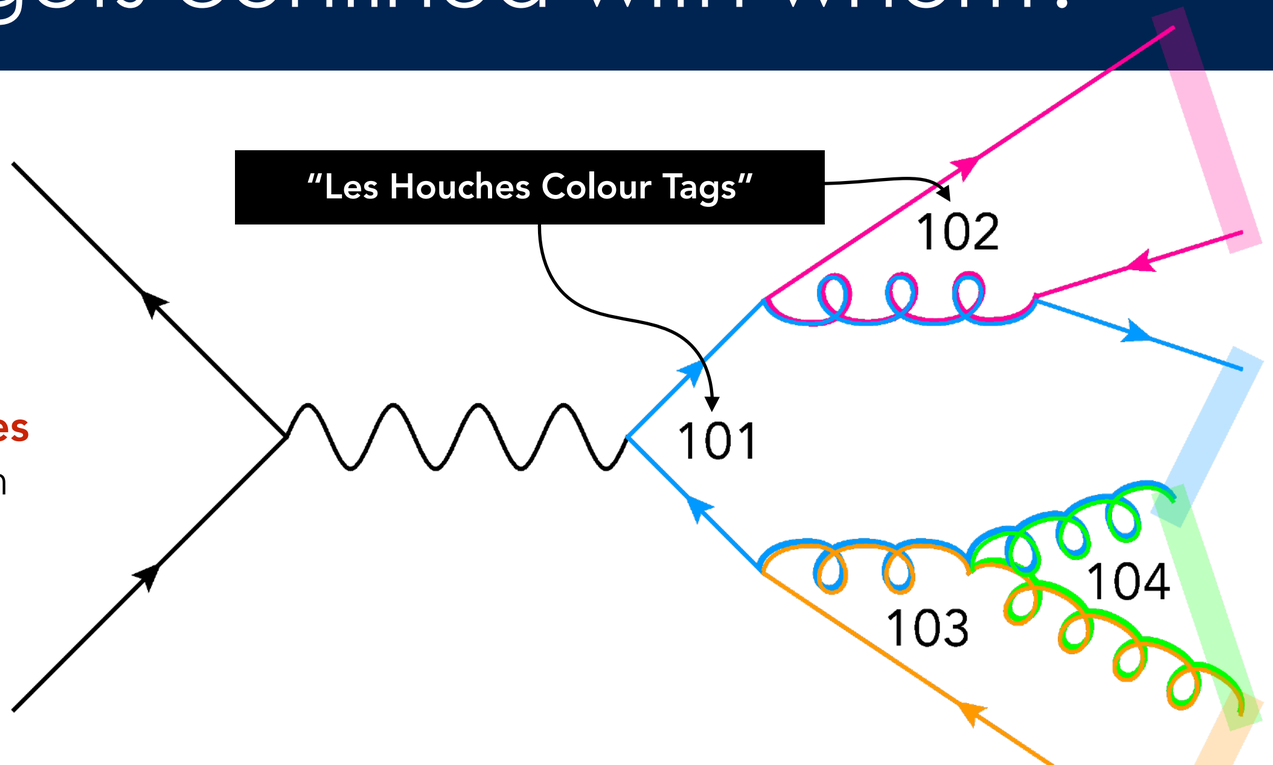
No $b$ or $c$; Suppression of strange

# Who gets confined with whom?

## "Leading Colour"

MCs: $N_C \to \infty$ limit formalised by letting **each "colour line"** be represented by a **unique Les Houches colour tag**† (no interference between different colour lines in this limit)

†: hep-ph/0109068; hep-ph/0609017



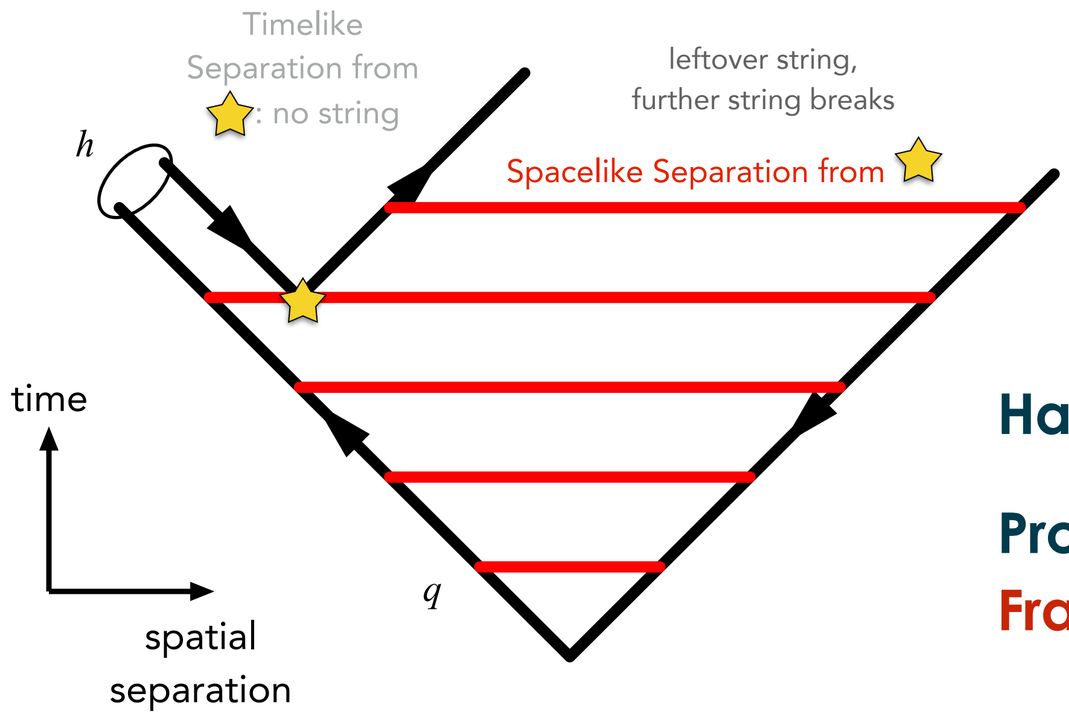

Expect accurate to ~ $1/N_C^2$ ~ 10%

A corresponding event record from PYTHIA, up to the second gluon emission

| # | id | name | status | mothers | | daughters | | colours | | p_x | p_y | p_z | e | m |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 5 | 23 | (Z0) | -22 | 3 | 4 | 6 | 7 | | | 0.000 | 0.000 | 0.000 | 91.188 | 91.188 |
| 6 | 3 | (s) | -23 | 5 | 0 | 10 | 0 | 101 | 0 | -12.368 | 16.523 | 40.655 | 45.594 | 0.000 |
| 7 | -3 | (sbar) | -23 | 5 | 0 | 8 | 9 | 0 | 101 | 12.368 | -16.523 | -40.655 | 45.594 | 0.000 |
| 8 | 21 | (g) | -51 | 7 | 0 | 13 | 0 | 103 | 101 | 9.243 | -9.146 | -29.531 | 32.267 | 0.000 |
| 9 | -3 | sbar | 51 | 7 | 0 | | | 0 | 103 | 3.084 | -7.261 | -10.973 | 13.514 | 0.000 |
| 10 | 3 | (s) | -52 | 6 | 0 | 11 | 12 | 101 | 0 | -12.327 | 16.406 | 40.505 | 45.406 | 0.000 |
| 11 | 21 | g | -51 | 10 | 0 | | | 101 | 102 | -2.834 | -2.408 | 1.078 | 3.872 | 0.000 |
| 12 | 3 | s | 51 | 10 | 0 | | | 102 | 0 | -10.246 | 17.034 | 38.106 | 42.979 | 0.000 |
| 13 | 21 | g | 52 | 8 | 0 | | | 103 | 101 | 9.996 | -7.366 | -28.211 | 30.823 | 0.000 |

# Requirement #2: on-shell hadrons

**Observation:** All string breaks are **causally disconnected**
(➢ independent modulo entanglement from common origin)




Lorentz invariance ⟹ string breaks can be considered in *any order*.

Exploit this to split off "outermost" hadron either from left or right (randomly) — **iteratively!**

**Hadron** $h$ takes a fraction $z$ of the **quark** momentum

**Probability distribution** in $z \in [0,1]$ parametrised by **Fragmentation Function**, $f_{\mathrm{Lund}}(z, Q^2_{\mathrm{HAD}})$

**"Left-right symmetry"** ⟹ **FF** constrained to a form with **two free parameters**, $a$ & $b$
*(constrained by fits to measured hadron spectra)*

[See, e.g., Amoroso et al., JCAP 05 (2019) 007]

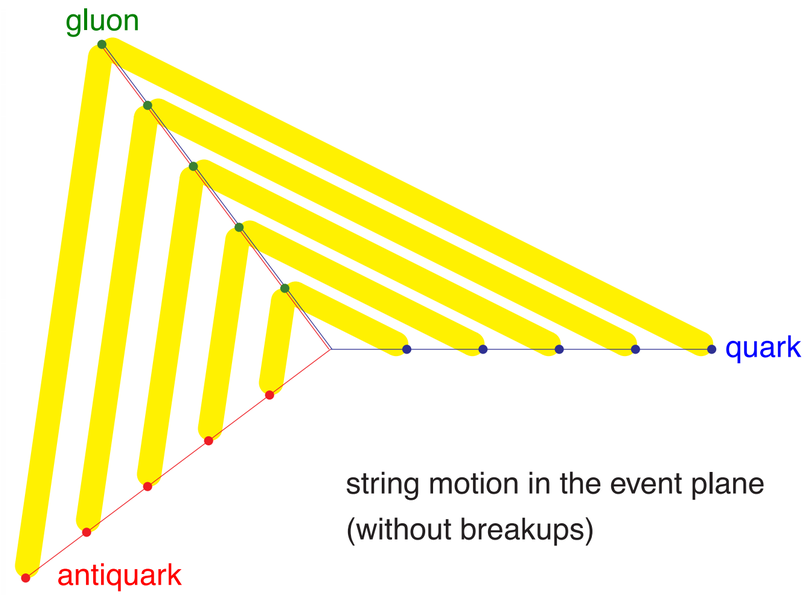
$$f_{\mathrm{Lund}}(z) \propto \frac{1}{z}(1-z)^{a}\exp\left(-\frac{b(m_h^2 + p_{\perp h}^2)}{z}\right)$$

Supresses high $z$ (the $(1-z)^a$ factor)

Supresses low $z$ (the $1/z$ factor)

# **Gluon Kinks:** The Signature Feature of the Lund Model

Gluons are connected to two string pieces

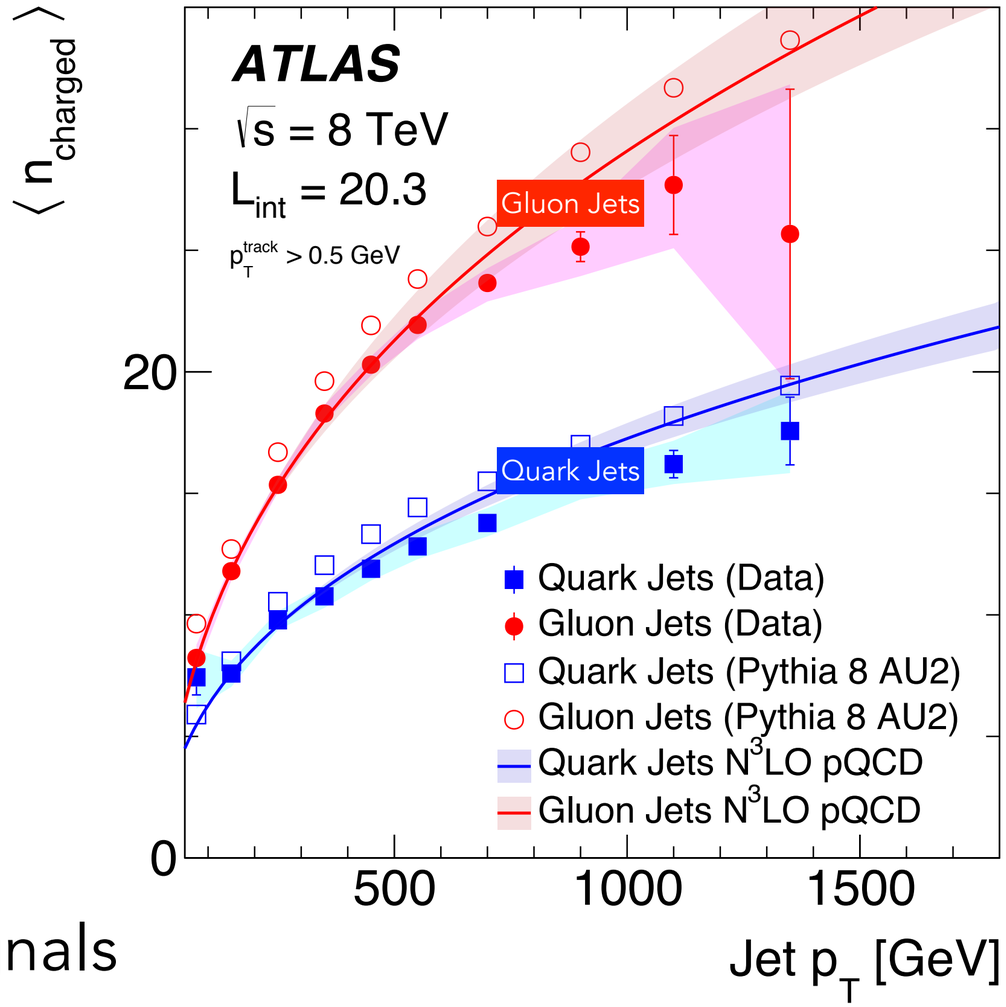



string motion in the event plane
(without breakups)

Each quark connected to **one** string piece

Expect factor ~ $2 \sim C_A/C_F$ more particles in gluon jets

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




Important for discriminating new-physics signals

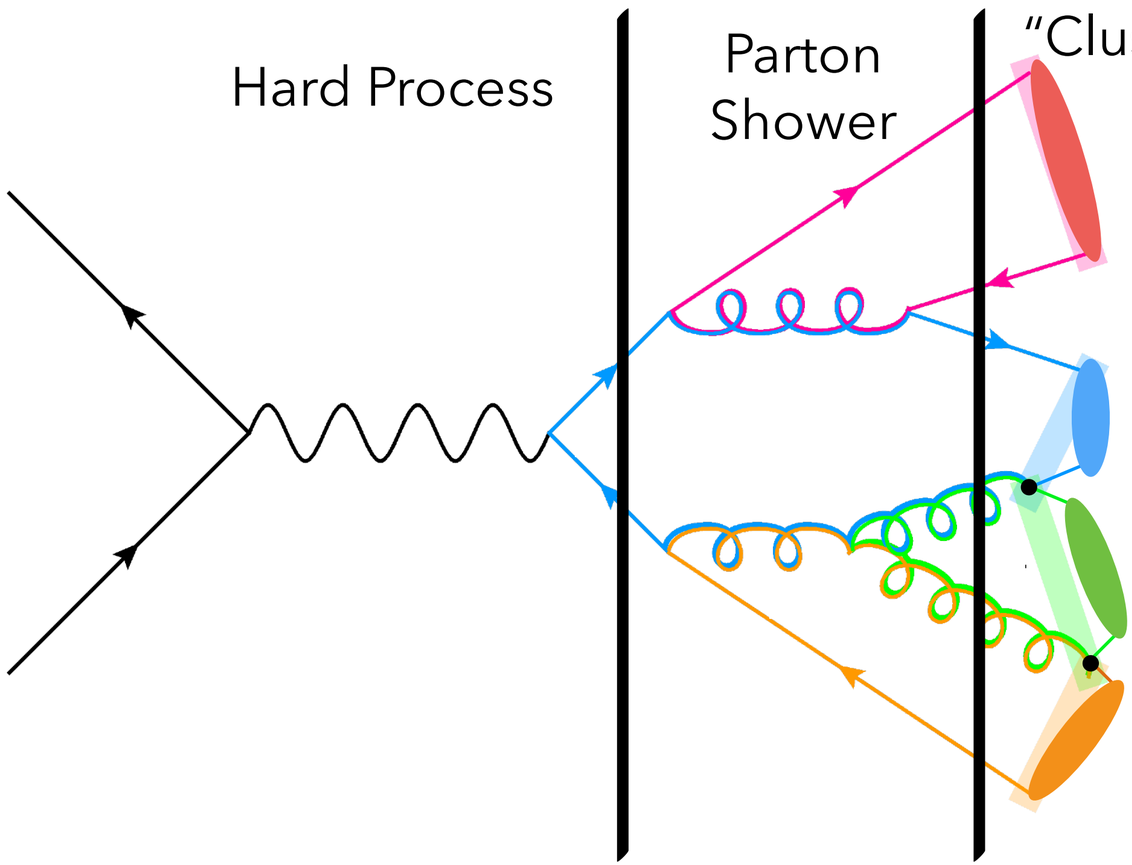
Decays to **quarks** vs decays to **gluons**,

vs composition of **background** and **bremsstrahlung** combinatorics

# Alternative: **The Cluster Model** — Used in HERWIG & SHERPA

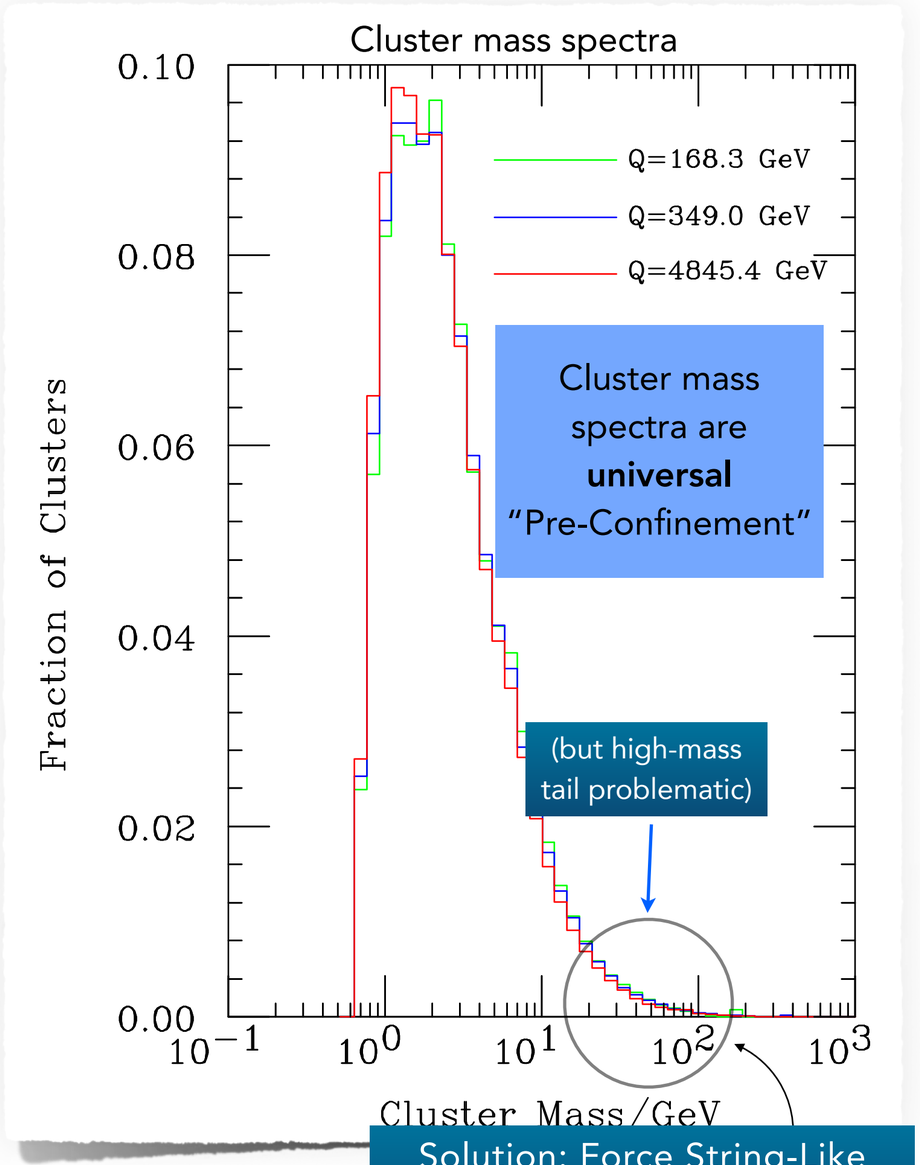
## Alternative to strings:

Force $g \to q\bar{q}$ at end of shower




**Isotropic 2-body decays to hadrons**

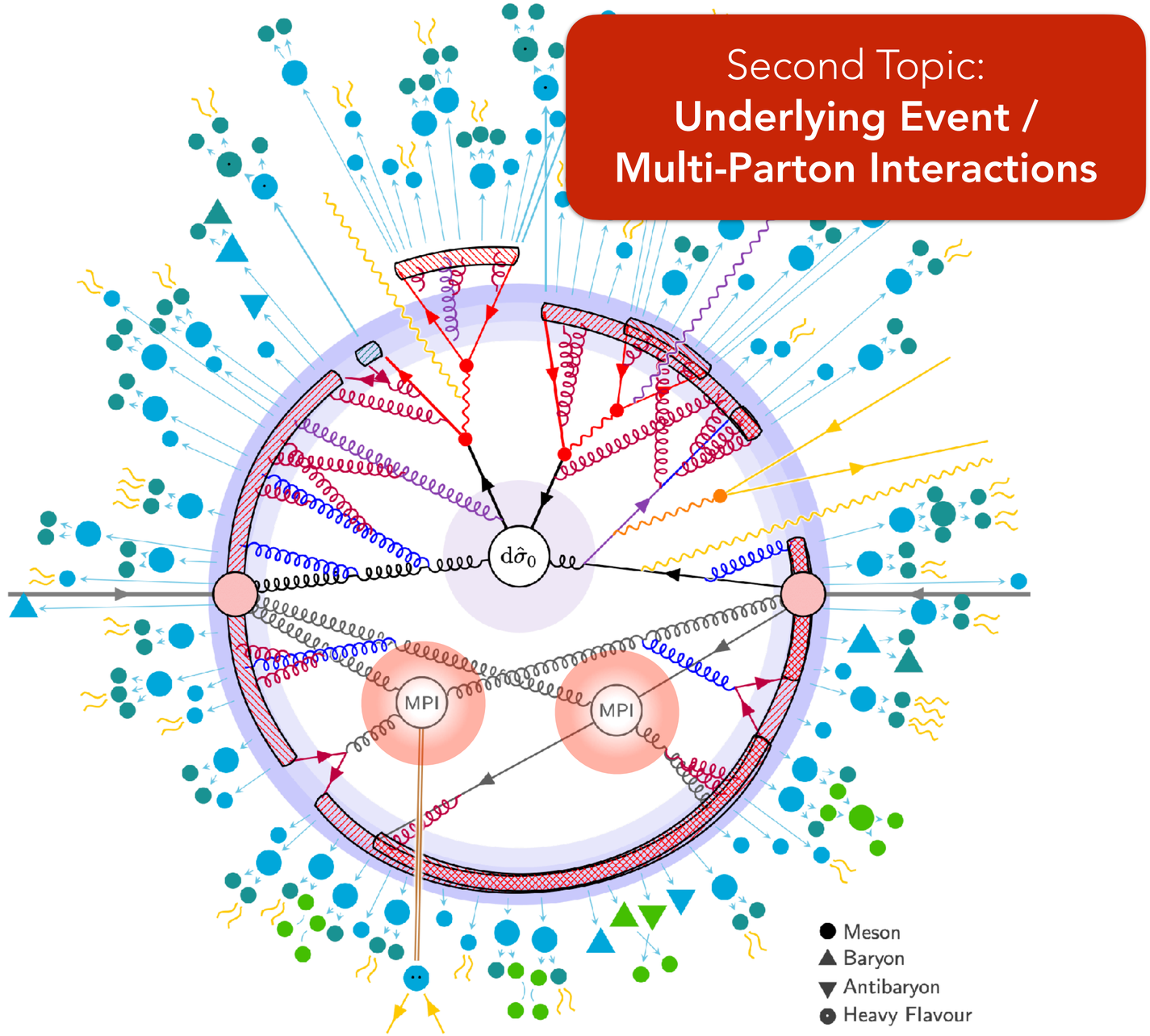
**According to phase space**



# The Physics of Event Generators

- Hard Interaction
- Resonance Decays
- MECs, Matching & Merging
- FSR
- ISR*
- QED
- Weak Showers
- Hard Onium
- Multiparton Interactions
- Beam Remnants*
- Strings
- Ministrings / Clusters
- Colour Reconnections
- String Interactions
- Bose-Einstein & Fermi-Dirac
- Primary Hadrons
- Secondary Hadrons
- Hadronic Reinteractions

*(*: incoming lines are crossed)*




Second Topic:
**Underlying Event / Multi-Parton Interactions**

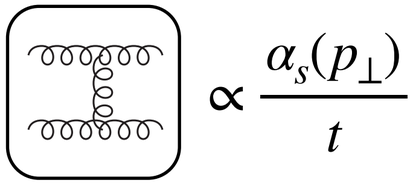
Figure from arXiv:2203.11601

# Hadron Collisions → Multi-Parton Interactions

## Protons are composite

**One** proton = **beam** of partons

+ $\mathrm{d}\sigma_{\mathrm{parton-parton}}$ is dominated by $t$-channel gluon exchange: **diverges for** $\hat{p}_\perp \to 0$ GeV



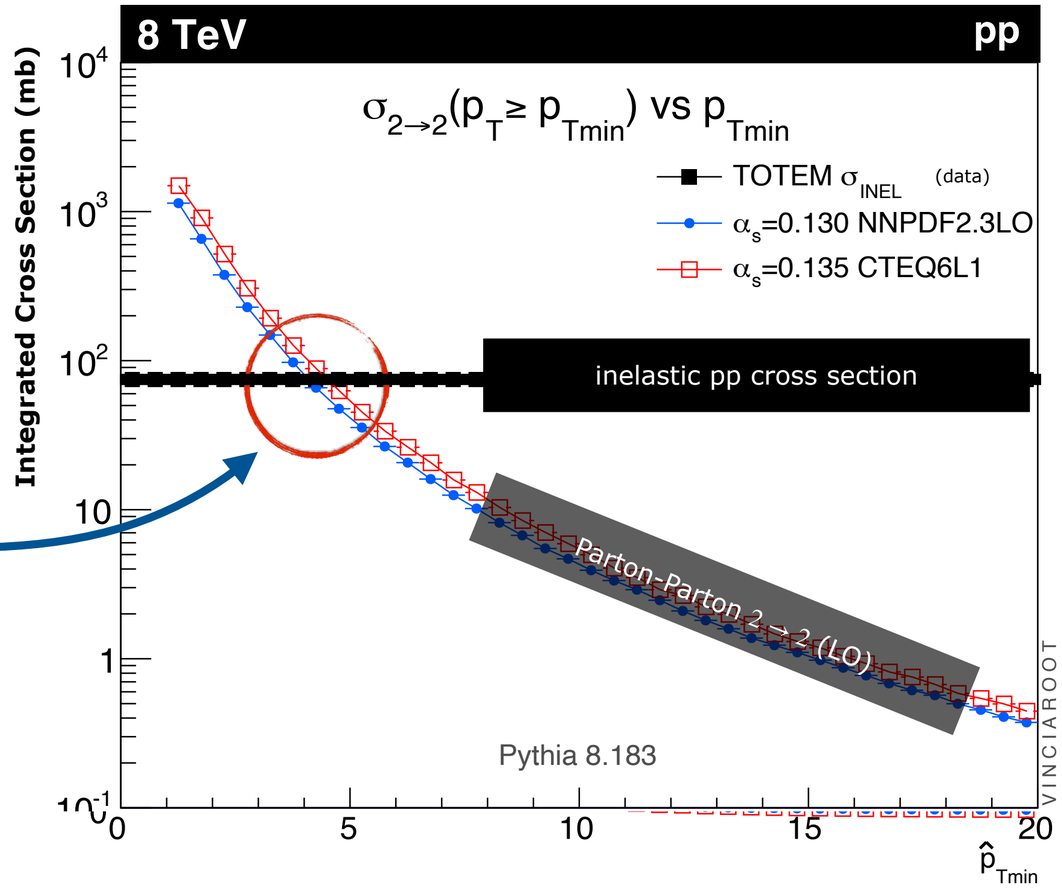
$$\propto \frac{\alpha_s(p_\perp)}{t}$$

$\Longrightarrow$ For sufficiently low $\hat{p}_\perp$(~ 5 GeV at LHC), we will have $\sigma_{\mathrm{parton-parton}}(\hat{p}_\perp) \gtrsim \sigma_{\mathrm{proton-proton}}$

Interpretation: $\dfrac{\sigma_{\textbf{parton-parton}}(\hat{p}_\perp)}{\sigma_{\textbf{hadron-hadron}}} \sim \langle n \rangle_{\mathrm{parton-parton}}(\hat{p}_\perp)$

(Regulated at low $\hat{p}_\perp$ by IR cutoff ~ colour screening)

➜ Multiple Parton-Parton Interactions (MPI)




≈ poor man's saturation

↔ cut pomerons in Regge Theory

# A Brief History of MPI (in PYTHIA)

**1987** [Sjöstrand & van Zijl, Phys.Rev.D 36 (1987) 2019]

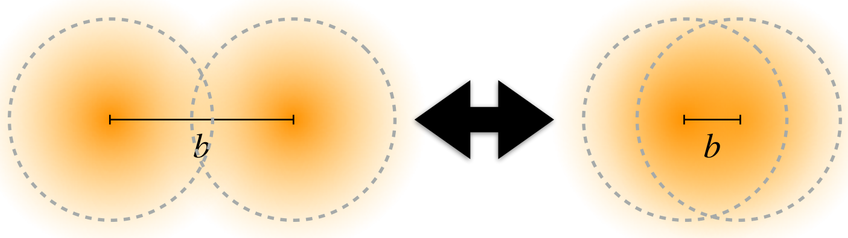
Cast MPI as Sudakov-style evolution:

Analogous to $\sigma_{\mathrm{X+jet}}(p_\perp)/\sigma_{\mathrm{X}}$ for parton showers

$$\frac{dP_{\mathrm{hardest}}}{d^2b\, dx_{T1}} = p(x_{T1},b)\, \exp\left\{-\int_{x_{T1}}^{1} p(x'_T,b)\, dx'_T\right\}$$

with: $\mathrm{p} \propto \sigma_{2\to2}(x_T,b)/\sigma_{pp}$ ; $x_T = 2\hat{p}_\perp/\sqrt{s}$

with Impact-parameter dependence




Crucial to describe **"Underlying Event"**

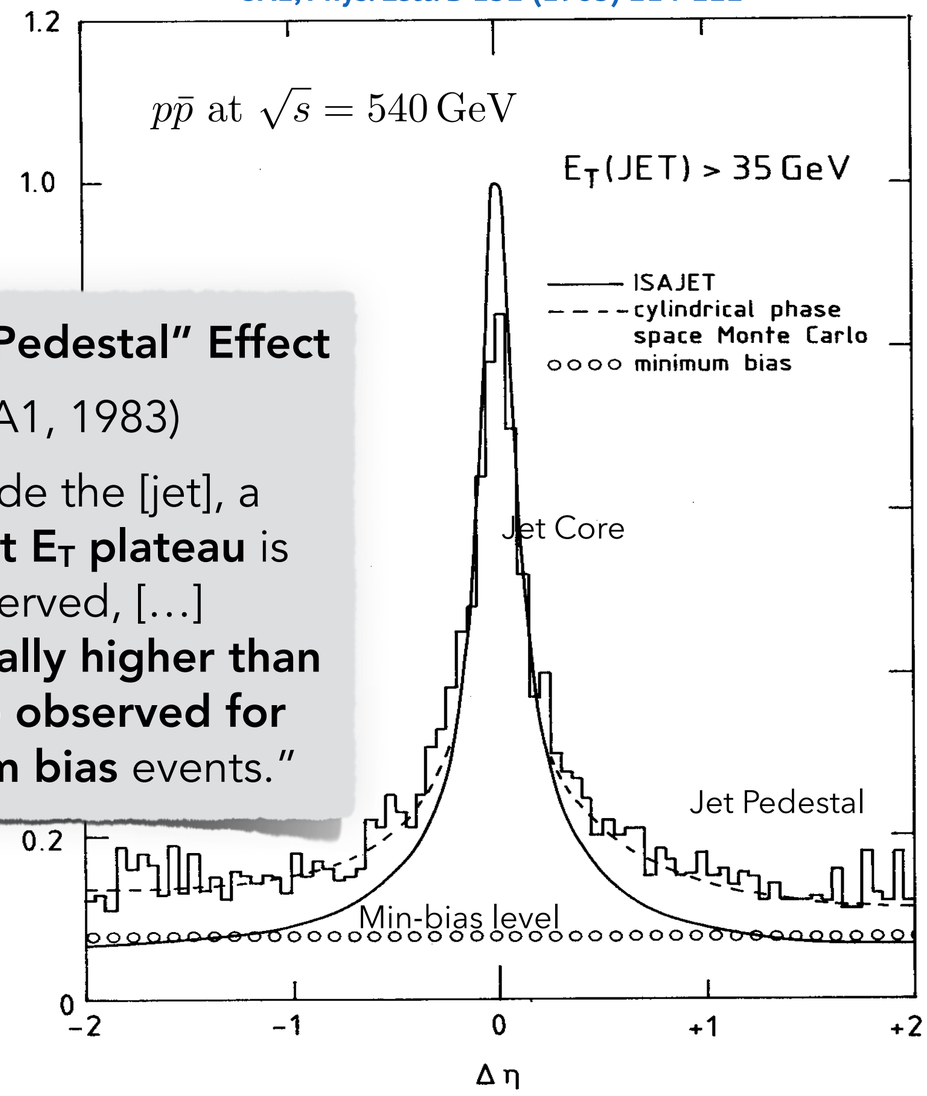
a.k.a. **"Jet Pedestal"**

**The "Jet Pedestal" Effect**

(UA1, 1983)

"Outside the [jet], a **constant $E_T$ plateau** is observed, […] **substantially higher than the one observed for minimum bias** events."

UA1, Phys. Lett. B 132 (1983) 214-222



# Pythia 8 — **Interleaved** Evolution

2005 [Sjöstrand & PS, Eur.Phys.J.C 39 (2005) 129]

**Interleave MPI & ISR** evolutions in one common sequence of pT

➜ ISR & MPI "compete" for the available $x$ in the proton remnant.

2011 [Corke & Sjöstrand, JHEP 03 (2011) 032]

Also include **FSR** in interleaving

**~ Fine-graining of all event structure above hadronization scale in one common sequence of quantum mechanical resolution $\propto p_\perp$**

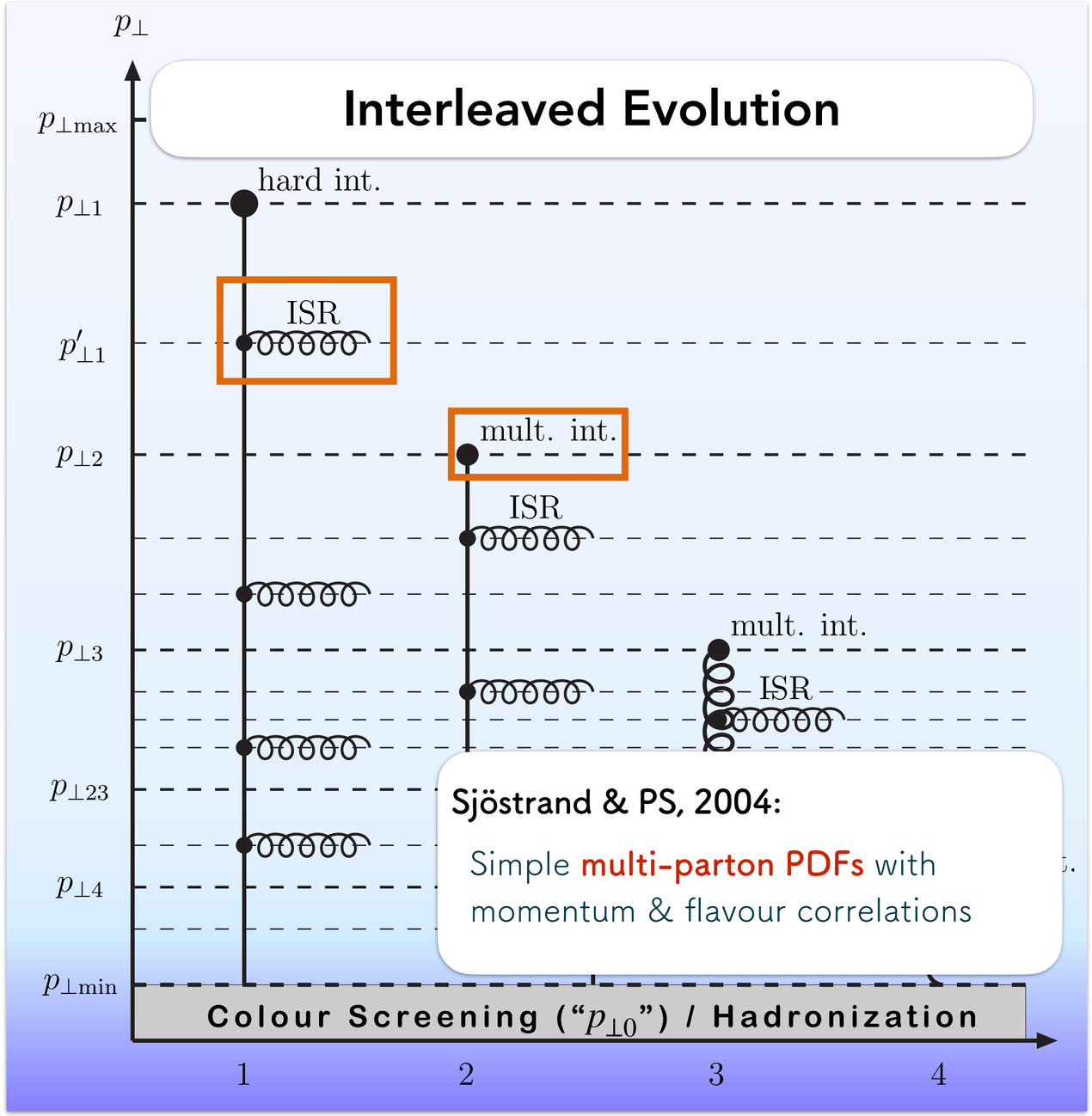



Figure from Sjöstrand & PS, 2005

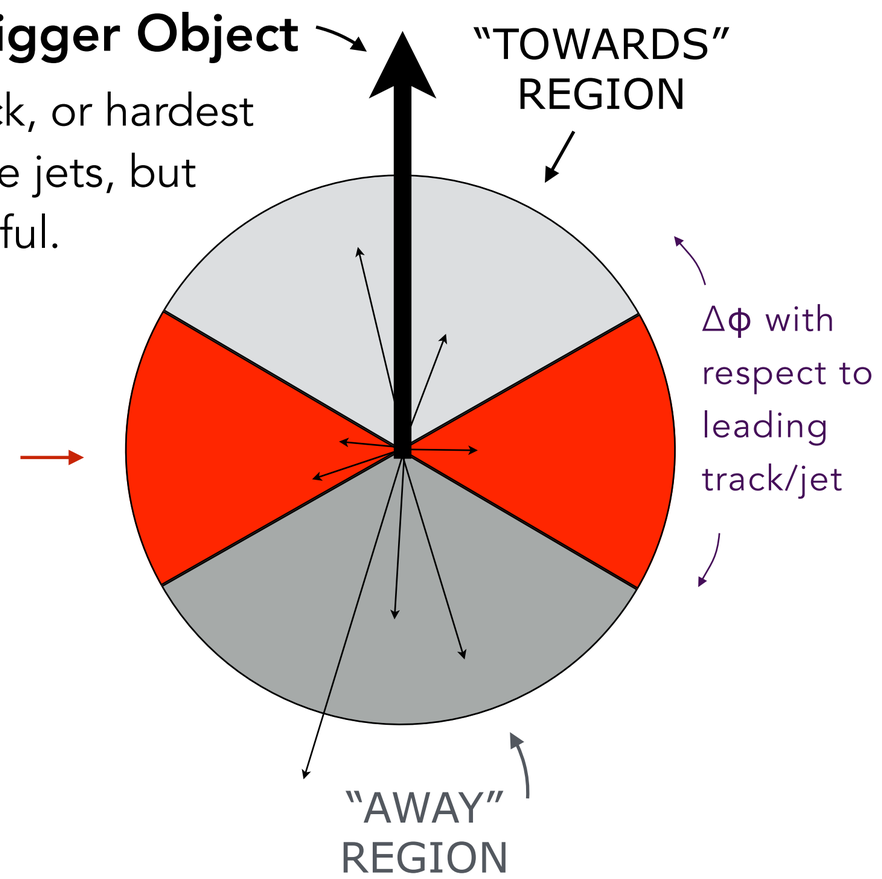
# What do we measure? Characterizing The Underlying Event

(The "Rick Field" UE Plots)

There are many UE variables.
The most important is $\langle \Sigma p_T \rangle$ 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.

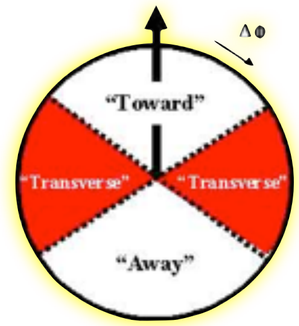



**"Transverse Region" (TRNS)**

Sensitive to activity at right angles to the hardest jets

➜ **Useful definition of Underlying Event**

# Min-Bias VS Underlying Event



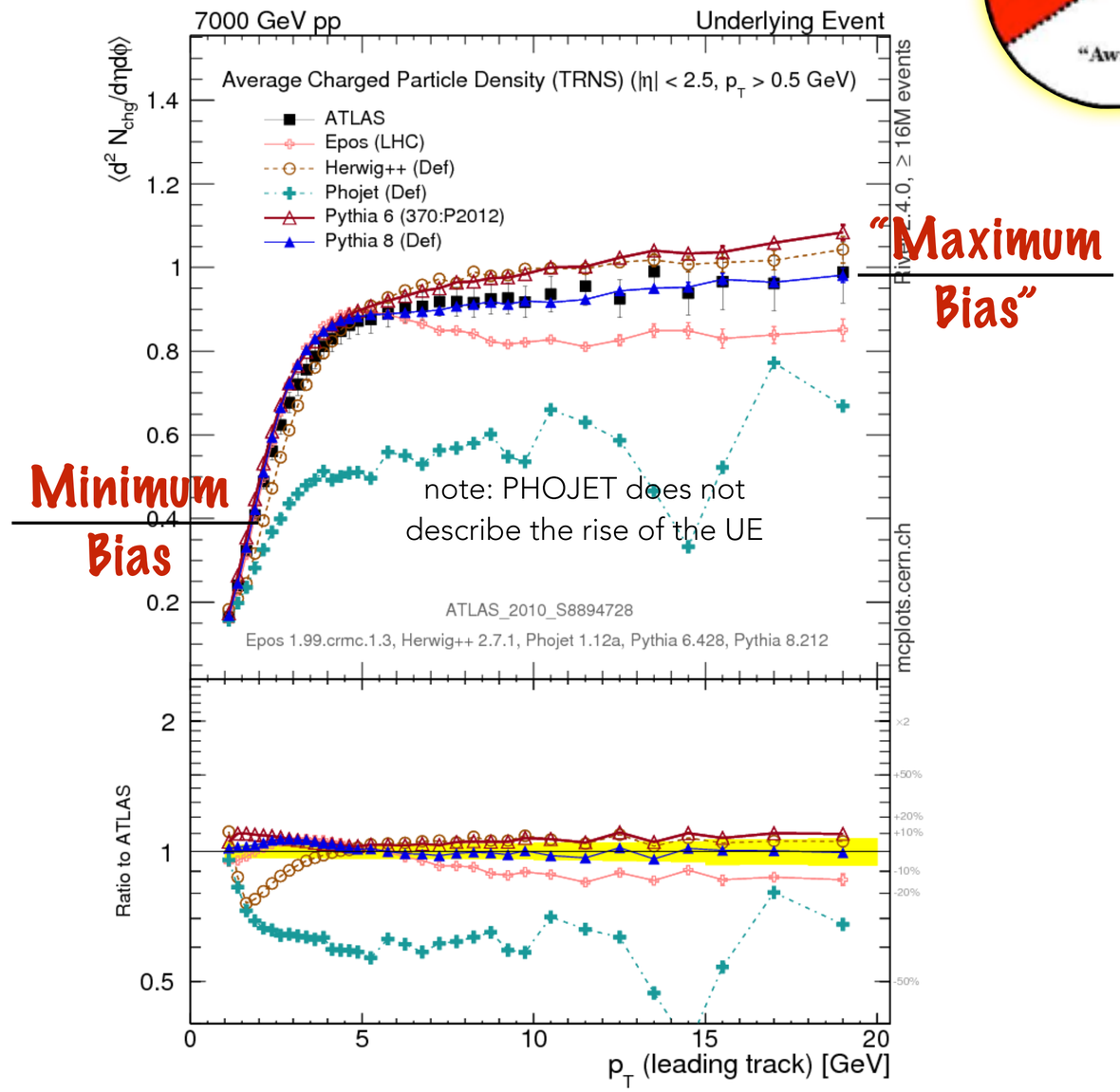

## Tautology:

**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**

# Third Topic: Colour Reconnections

MPI (or cut pomerons) ⇒ lots of coloured partons scattered into final state

Who gets confined with whom?

Each has a **colour ambiguity** $\sim 1/N_C^2 \sim 10\,\%$

E.g.: **random triplet** charge has 1/9 chance to be in **singlet** state with **random antitriplet**:

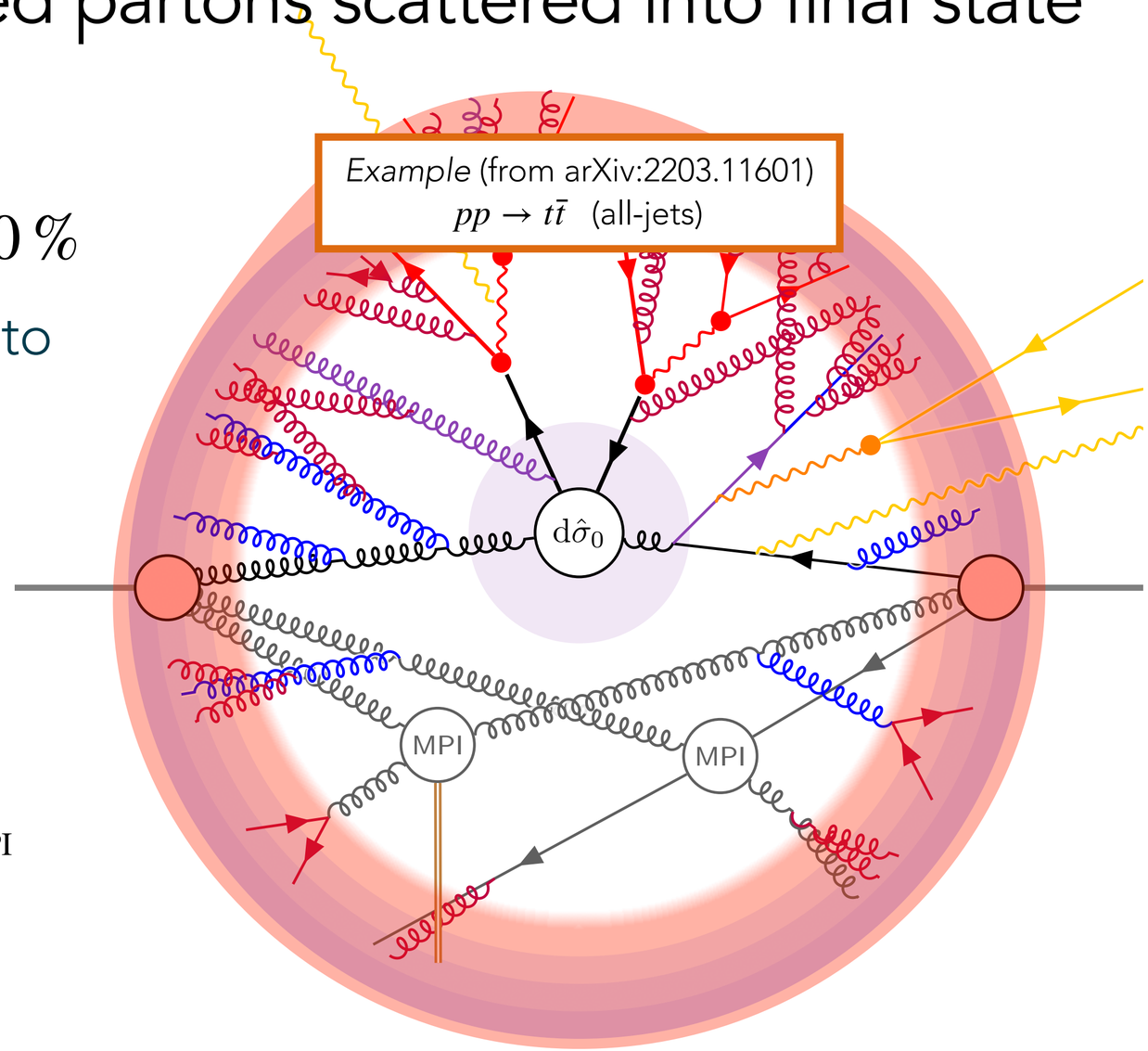
$3 \otimes \bar{3} = 8 \oplus 1,$

$3 \otimes 8 = 15 + 6 + 3,$ etc.

***Many* charges ➜ Colour Reconnections***

**(CR)** more likely than not

$$\text{Expect Prob(no CR)} \propto \left(1 - \frac{1}{N_C^2}\right)^{n_{\text{MPI}}}$$




***):** in this context, QCD CR simply refers to an ambiguity beyond Leading $N_C$, known to exist. The term "CR" can also be used more broadly.*

# Colour Correlations

## Each MPI exchanges colour between the beams

The colour flow determines the hadronizing string topology

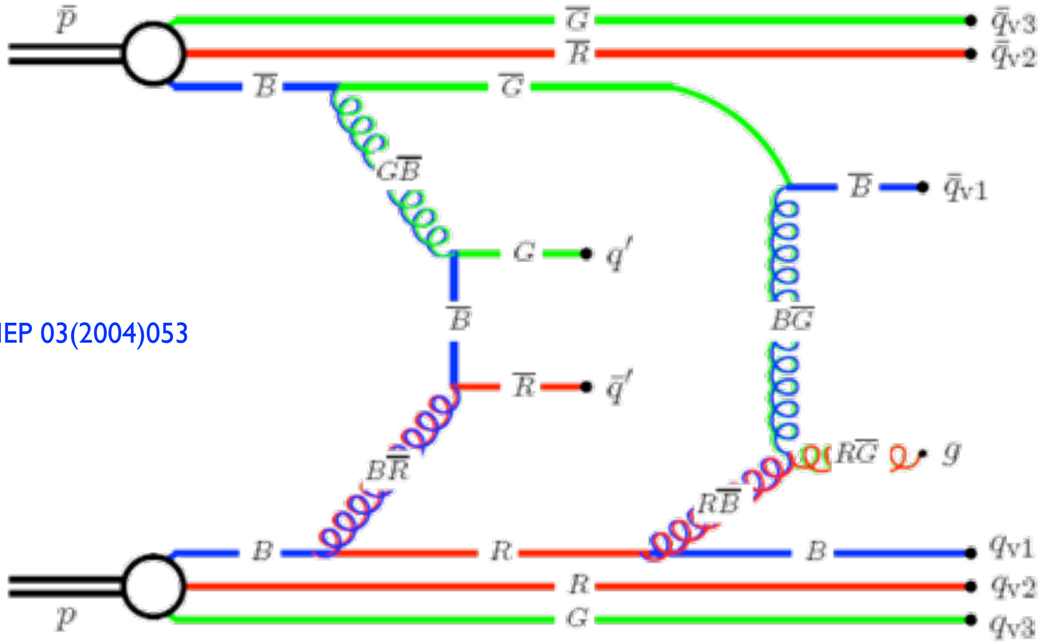
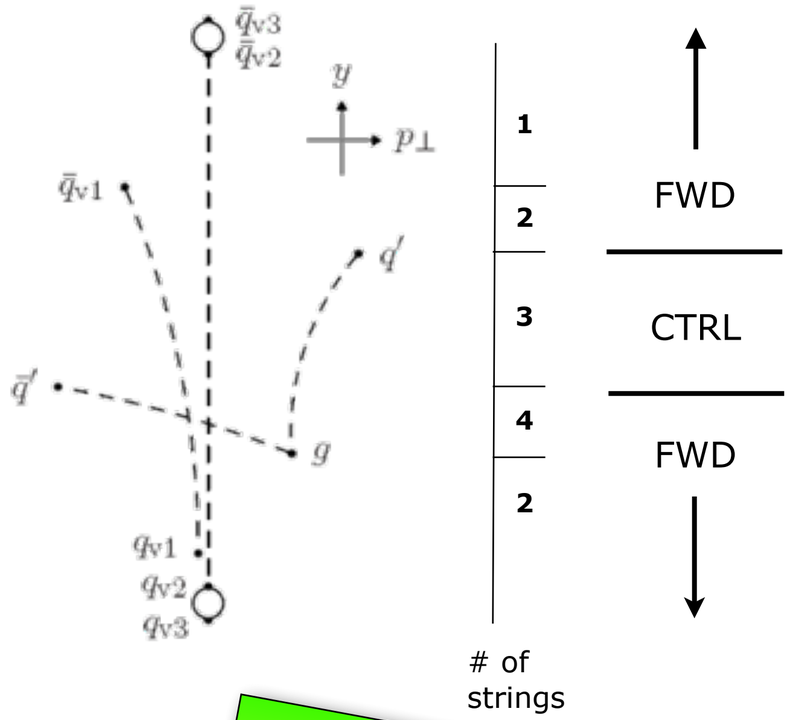



Figure from
Sjöstrand & PS, JHEP 03(2004)053




## Each MPI, even when soft, is a colour spark

Hadron distributions depend on colour space

Different models make different ansätze

# Colour Correlations

## Each MPI exchanges colour between the beams

The colour flow determines the hadronizing string topology

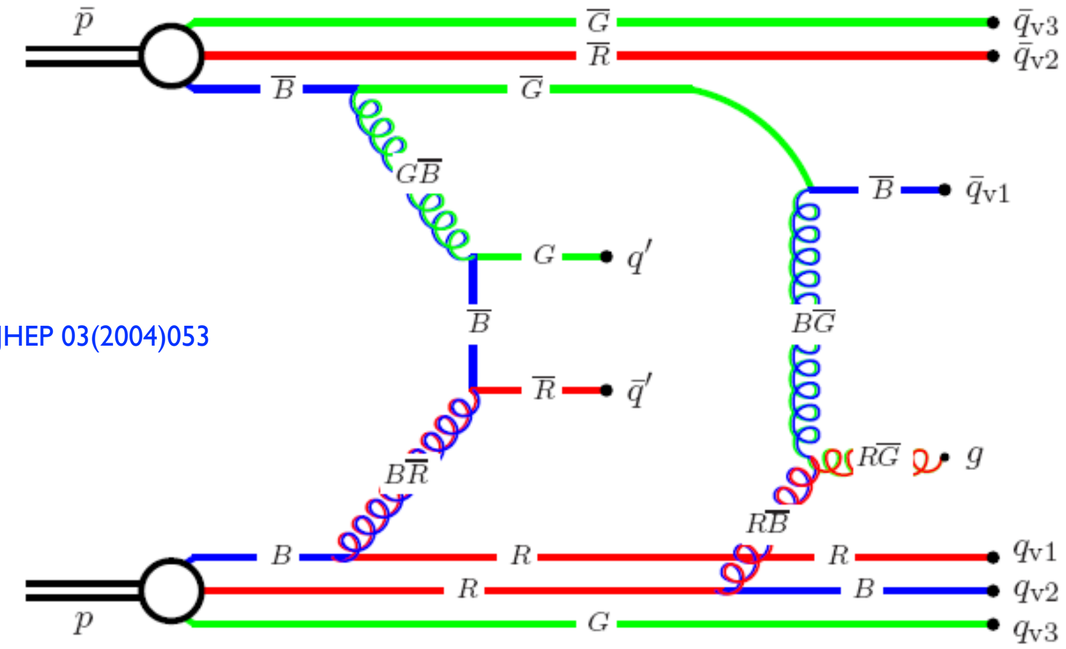
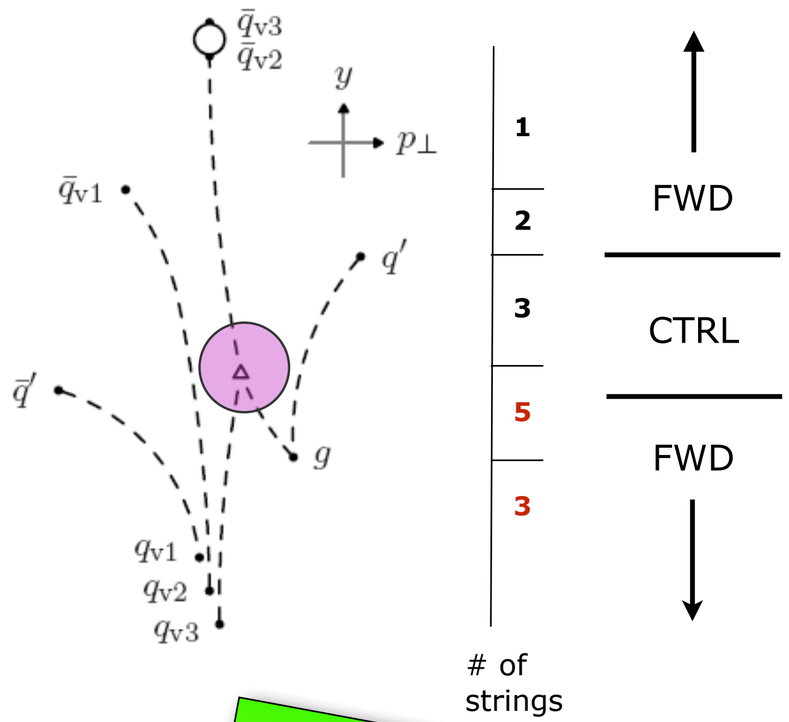



Figure from
Sjöstrand & PS, JHEP 03(2004)053



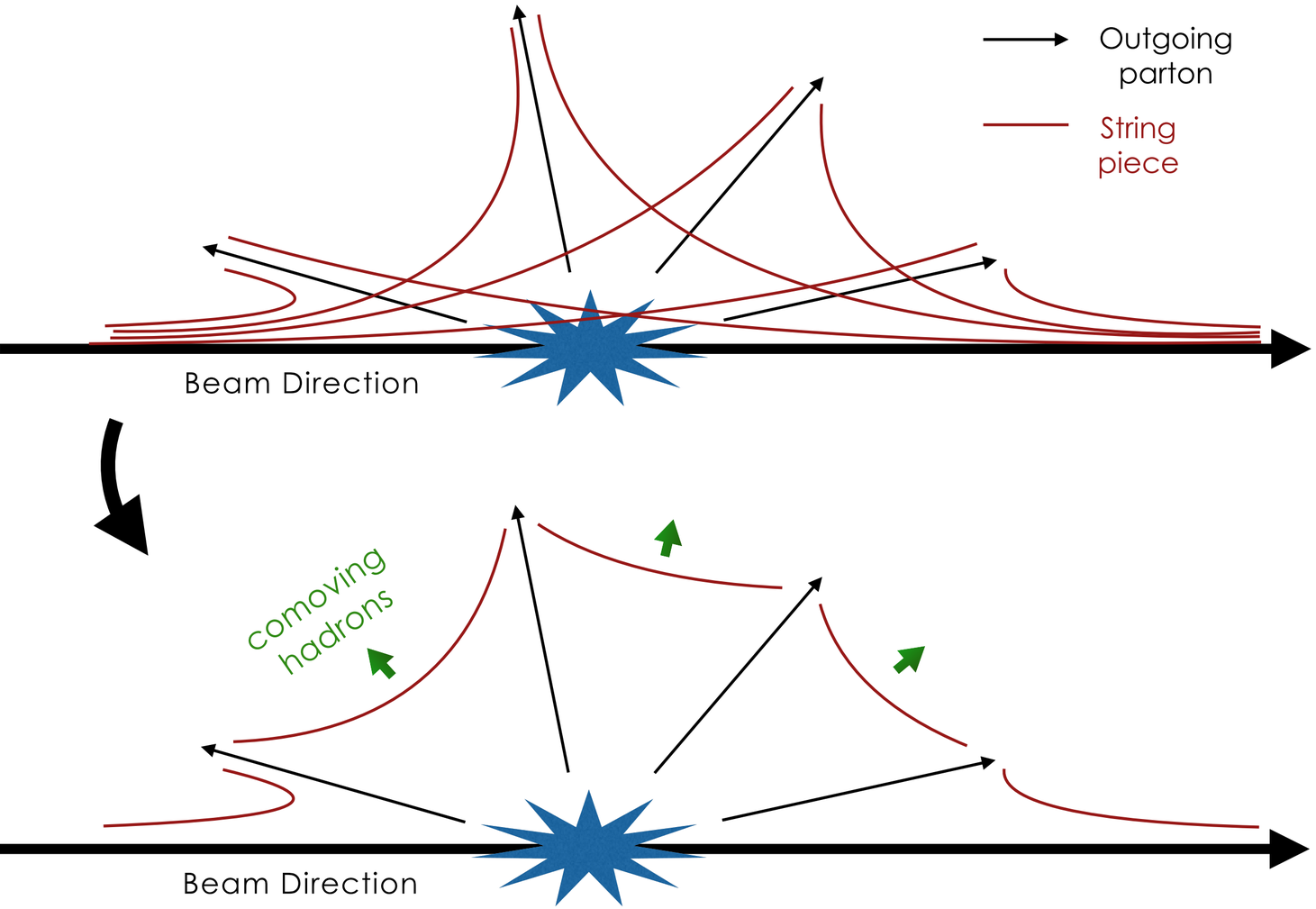

## Each MPI, even when soft, is a colour spark

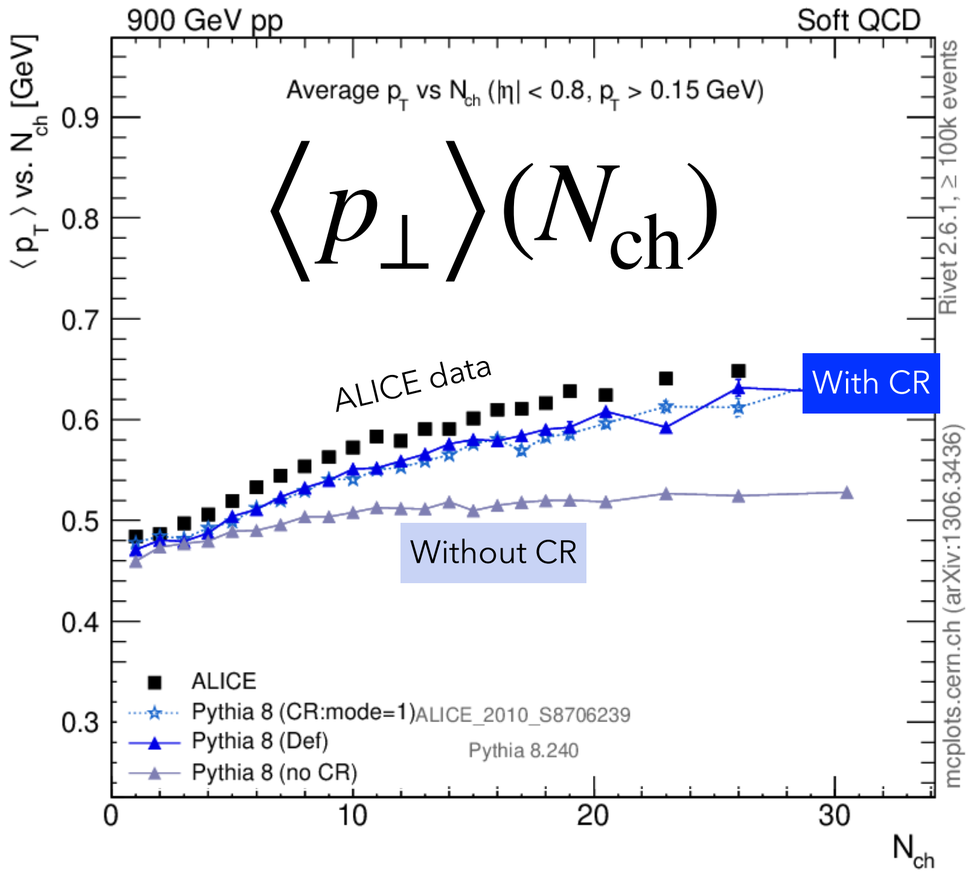
Hadron distributions depend on colour space

Different models make different ansätze

# String-length minimisation and <pT>($N_{ch}$)

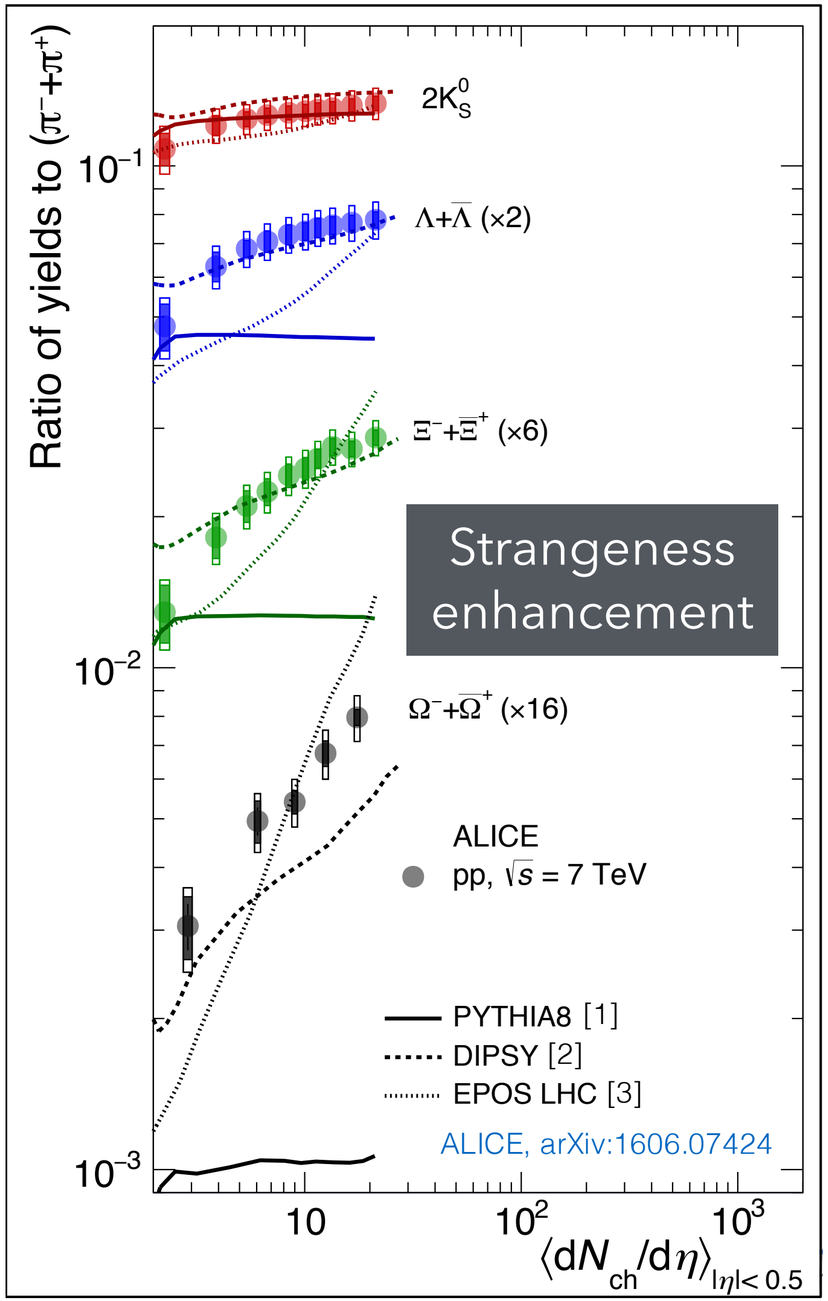
When many string configurations are possible, assume nature picks the one with smallest potential energy ~ "string length"

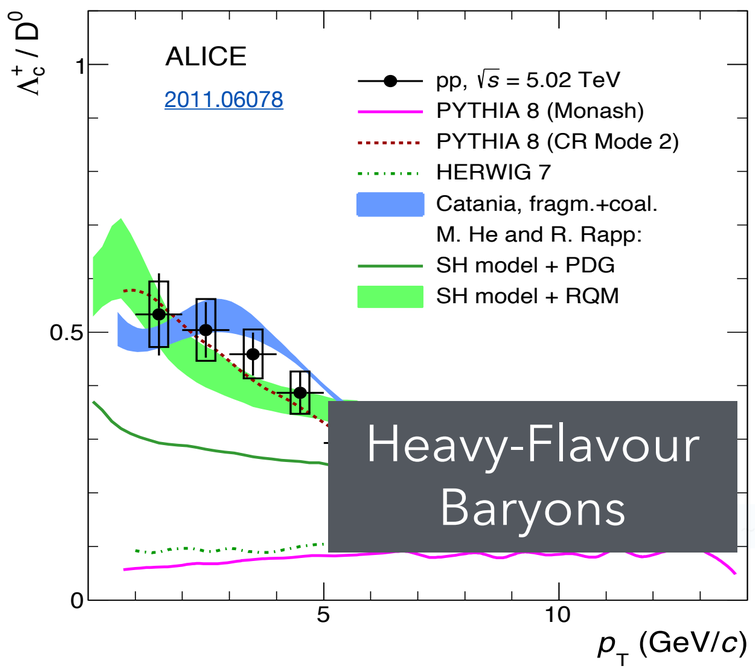


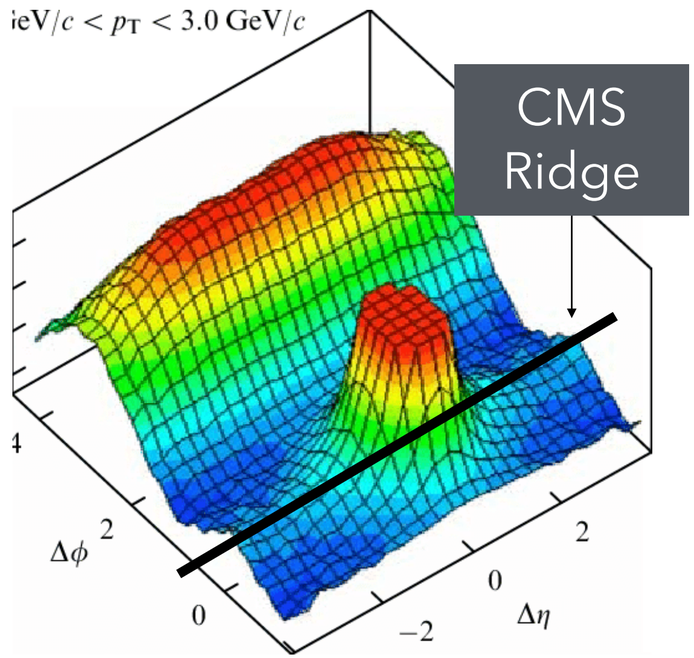



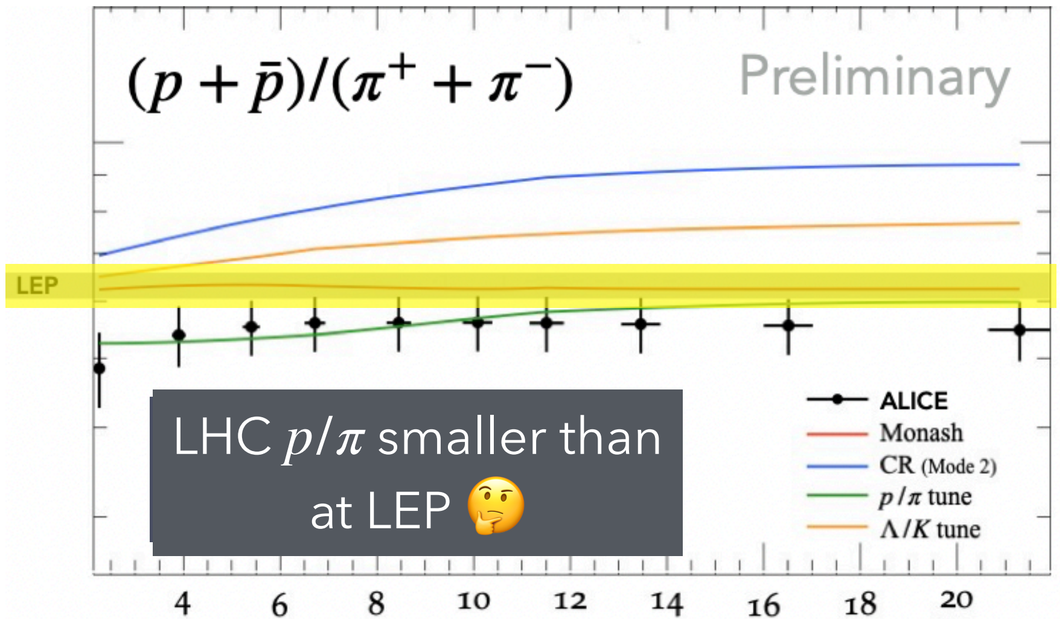

[See also Ortiz et al., Phys.Rev.Lett. 111 (2013) 4, 042001]

# Fourth Topic: New Discoveries at LHC




**+ Many more …**

Baryon correlations

$D_s$ asymmetries

Exotica

…

# **New Directions** in String Fragmentation

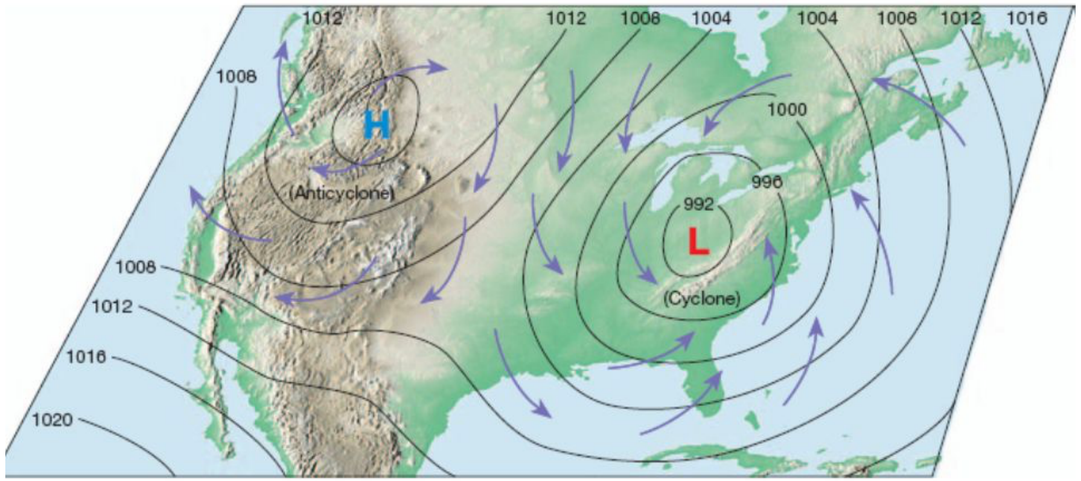
Regard tension $\kappa$ as an emergent quantity (not fundamental strings)

May depend on (invariant) time $\tau$?

*E.g.*, hot strings which cool down

[Hunt-Smith & PZS EPJ C 80 (2020) 11]

Cyclonic and Anticyclonic Winds




May depend on spatial coordinate $\sigma$?

Work in progress with J. Altmann (Monash), and E. Carragher & J. March-Russell (Oxford).

May depend on environment? (e.g., other strings nearby)

Two approaches (so far) within Lund string-model context:

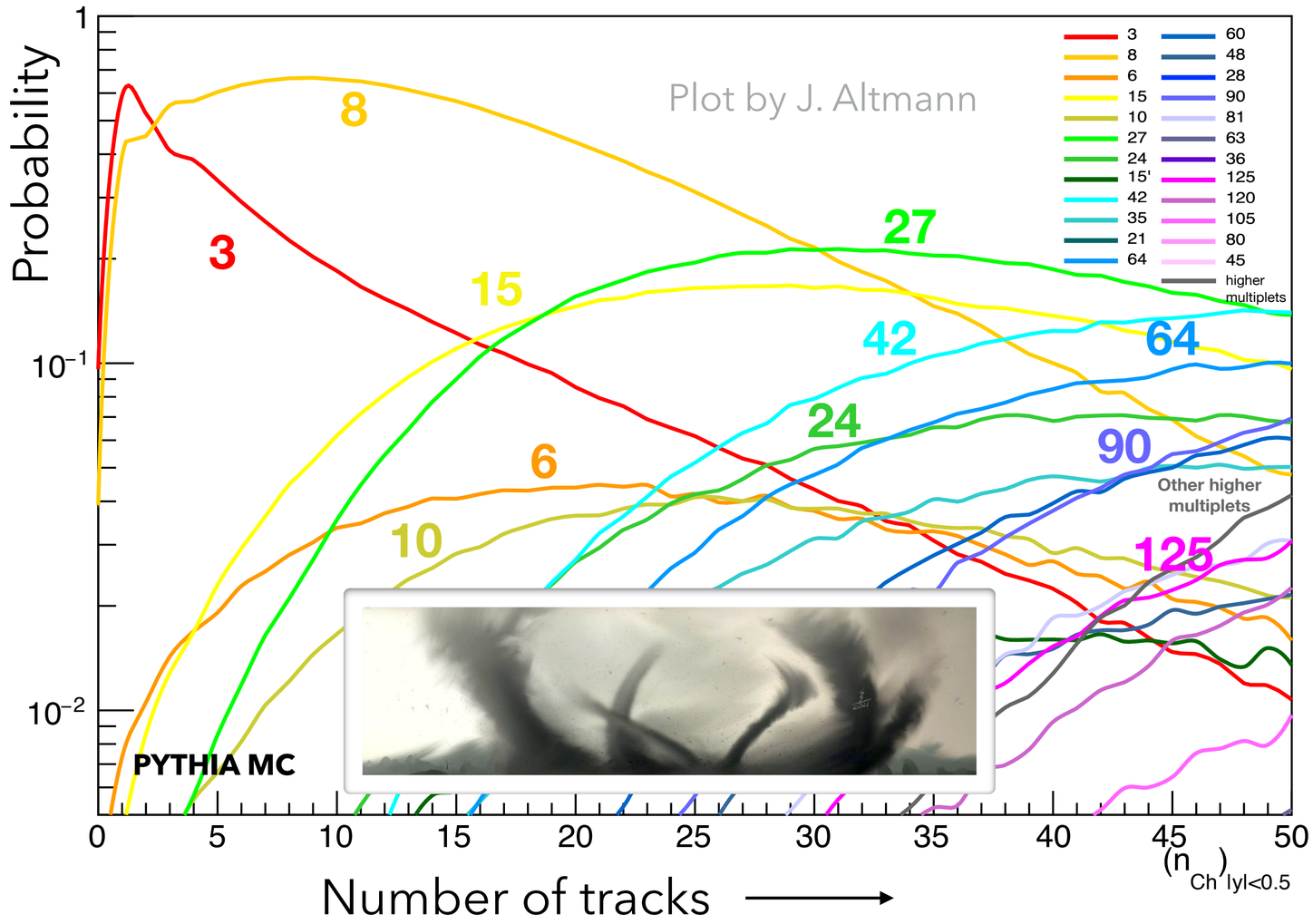
**Colour Ropes** [Bierlich et al. 2015] + several more recent

**Close-Packing** [Fischer & Sjöstrand 2017] + Work in progress with L. Bernardinis & V. Zaccolo (Trieste)

# Non-Linear String Dynamics?

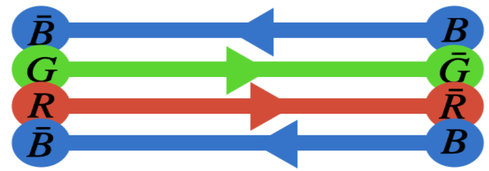
Count # of (oriented) flux lines crossing $y = 0$ in pp collisions at LHC

(according to PYTHIA) — And classify by SU(3) multiplet:




**Confining fields** may be reaching **higher effective representations** than simple $q\bar{q}$ (3) ones.

*E.g.:* $27 \otimes 27$




$p = 2$, $q = 2$ $\quad C_{27} = 6C_F$

**➜ Is "emergent tension" driving strangeness enhancement in pp?**

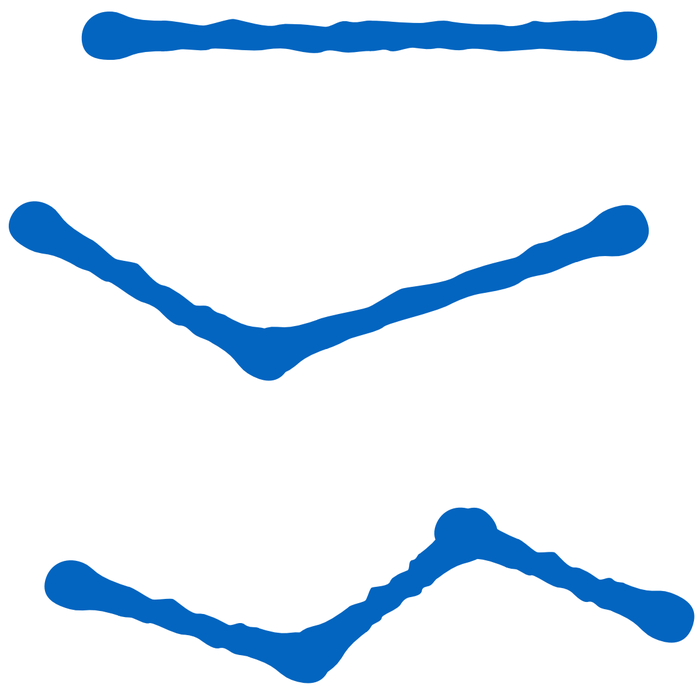
Colour Ropes (Bierlich et al.),

+ Close-Packing: Altmann, Bernardinis, Jueid, Kreps, **PS**, Zaccolo (in progress)

# What about Baryon Number?

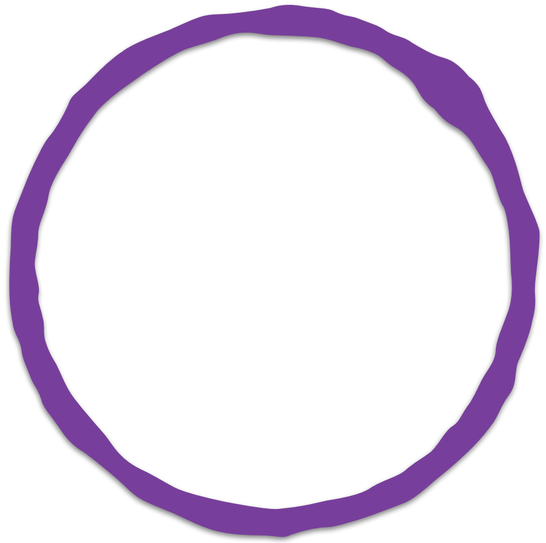
Types of string topologies:

**Open Strings**



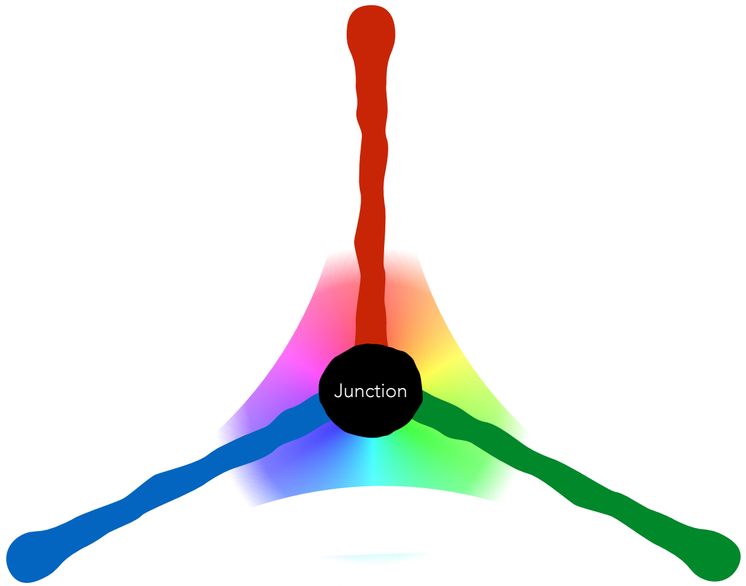
$$(3 \otimes \bar{3})_{\text{singlet}} = \frac{1}{9}$$

**Closed Strings**



$$(8 \otimes \bar{8})_{\text{singlet}} = \frac{1}{64}$$

**SU(3) String Junction**

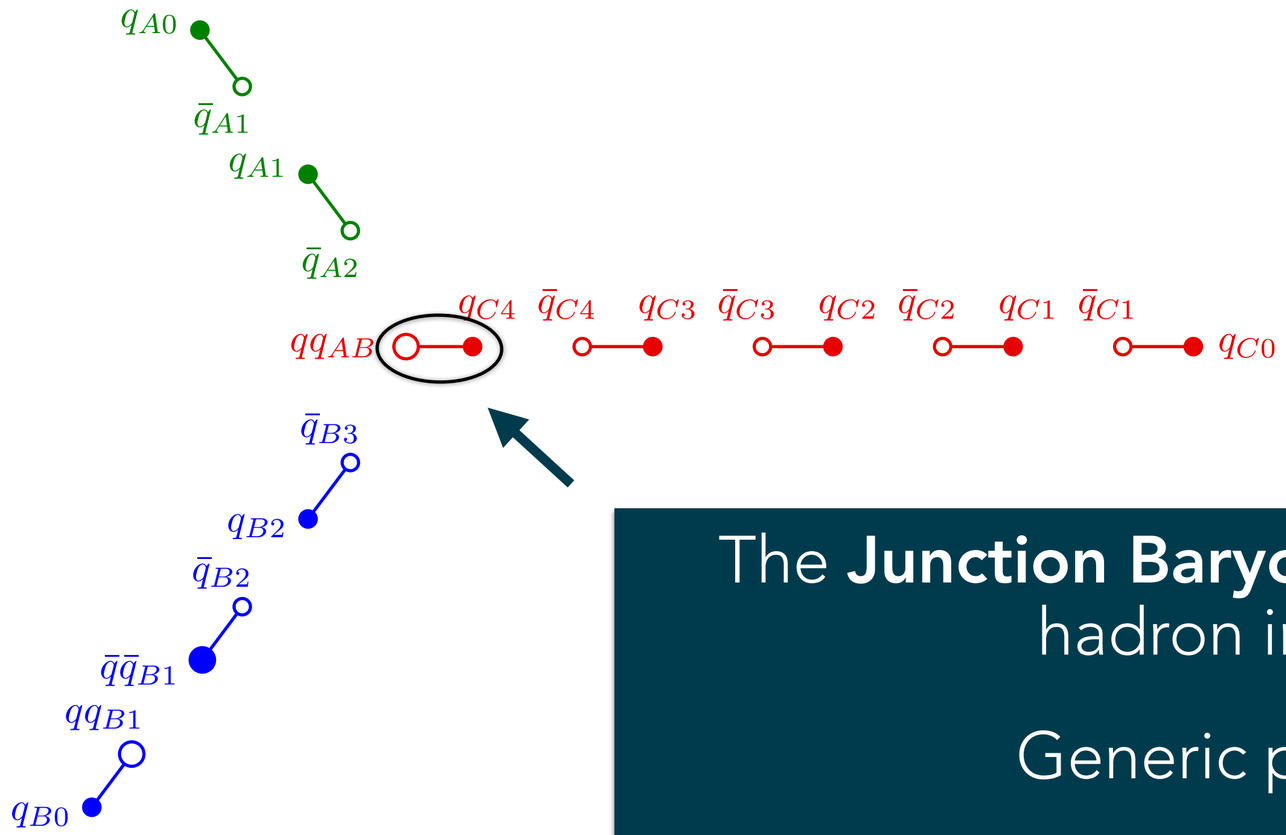



$$(3 \otimes 3 \otimes 3)_{\text{singlet}} = \frac{1}{27}$$

Could we get these at LHC?

# Fragmentation of String Junctions

Assume Junction Strings have same properties as ordinary ones (u:d:s, Schwinger $p_T$, etc) ➤ No new string-fragmentation parameters




[Sjöstrand & PS, NPB 659 (2003) 243]

[+ Altmann & PS, JHEP 07 (2024) 238]

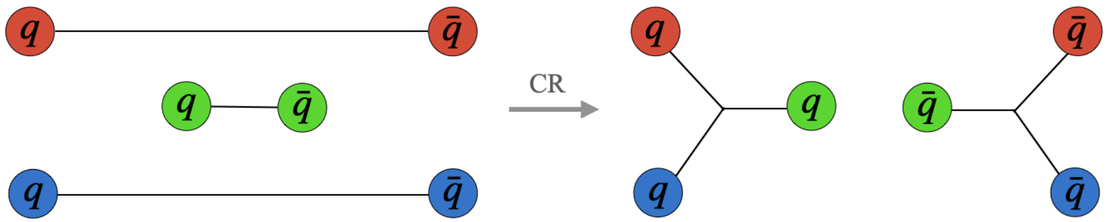
The **Junction Baryon** is the most "subleading" hadron in all three "jets".

Generic prediction: **low $p_T$**

**A Smoking Gun for String Junctions: Baryon enhancements at low $p_T$**

# Colour Reconnections ➢ String Junctions

[Christiansen & PS 2015, Altmann & PS 2024]

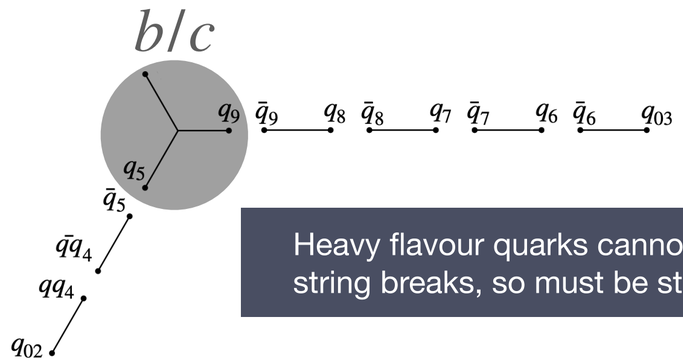



Mechanism for **baryon production**

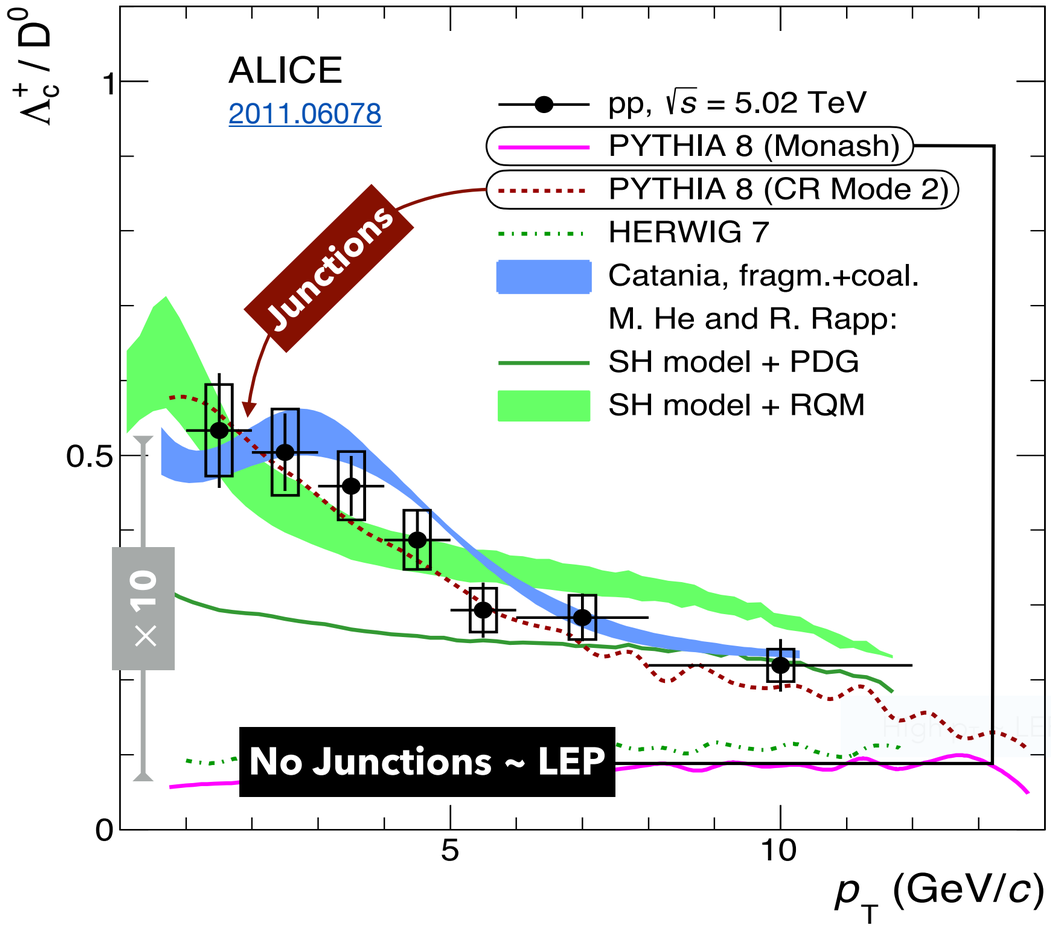
➢ ~40% of baryons are from junctions in PYTHIA

**Heavy flavour baryons**

➢ **~70% of heavy baryons** are from junctions in PYTHIA




Heavy flavour quarks cannot be made from string breaks, so must be string endpoints




**Next Steps:** put it all together *(+ "Altmann mechanism" for diquark disruption in octet fields)*

See how close we can get to describing light, strange, and heavy-flavour mesons + baryons in *pp*

+ Lund group developing extensions/applications to heavy-ion collisions!

# Heavy-Ion Physics

**Disclaimer:** I am not an expert

Also for HI, there are of course **event generators**

E.g., ANGANTYR, EPOS, HIJING, JEWEL, QGSJET, SIBYLL, …

Lots of recent activity **!**
Also in PYTHIA
Led by **Jyväskylä** & **Lund**

Another big class of models: **statistical hadronization**

Differ in **how much detail** you aim for, how multi-differential and/or event-by-event you want to be able to go …

You may **want** to focus on macroscopic properties, not the microphysics

Or you may **want** to pursue a microscopic description, without all macroscopic aspects

Most of us specialise, but I don't think the point is to pick a winner

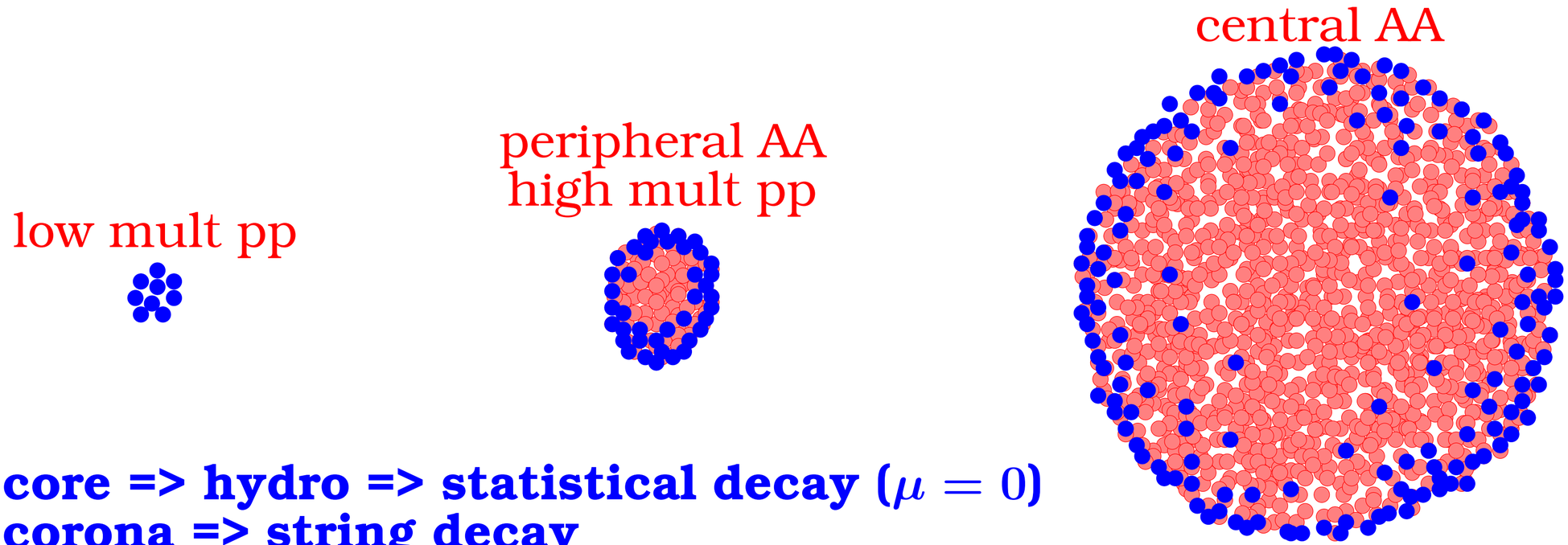
As a physicist, I'd like to understand ***both:*** what are the **macroscopic** properties? what is the **microphysics**? **How** do the former **emerge** from the latter?

Which paradigms are compatible / incompatible? How to form **clear** conclusions from **data?**

# Beyond Strings — QGP?

Currently most realistic complete approach for pp ↔ pA ↔ AA?

The core-corona solution [Werner 2007]: **mix** discrete **strings** with continuous **QGP**




**core => hydro => statistical decay ($\mu = 0$)**
**corona => string decay**

Allows smooth transition between string and hydro descriptions. Implemented in **EPOS MC**

Qualitatively agrees with ALICE strangeness data (but too steep rise with multiplicity?)

# **Conversely:** Collective flow from strings? (without QGP)

**Strings should push each other transversely**

Colour-electric fields ➜ Classical force

**Model string radial shape & shoving physics**

$\Longrightarrow$ force $$f(d_\perp) = \frac{g\kappa d_\perp}{R^2}\exp\left(-\frac{d_\perp^2}{4R^2}\right)$$

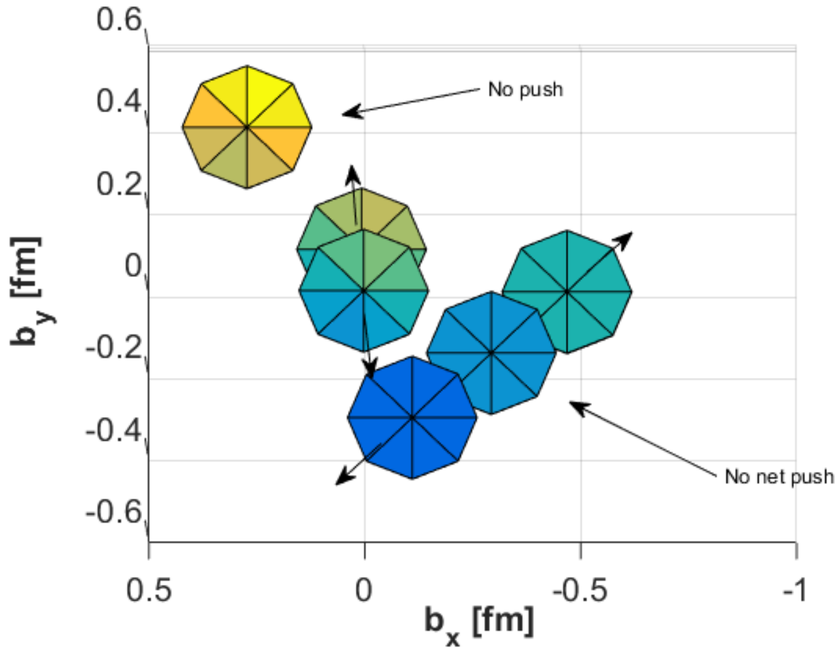
$g$: fraction of energy in chromo-electric field (as opposed to in condensate or magnetic flux)

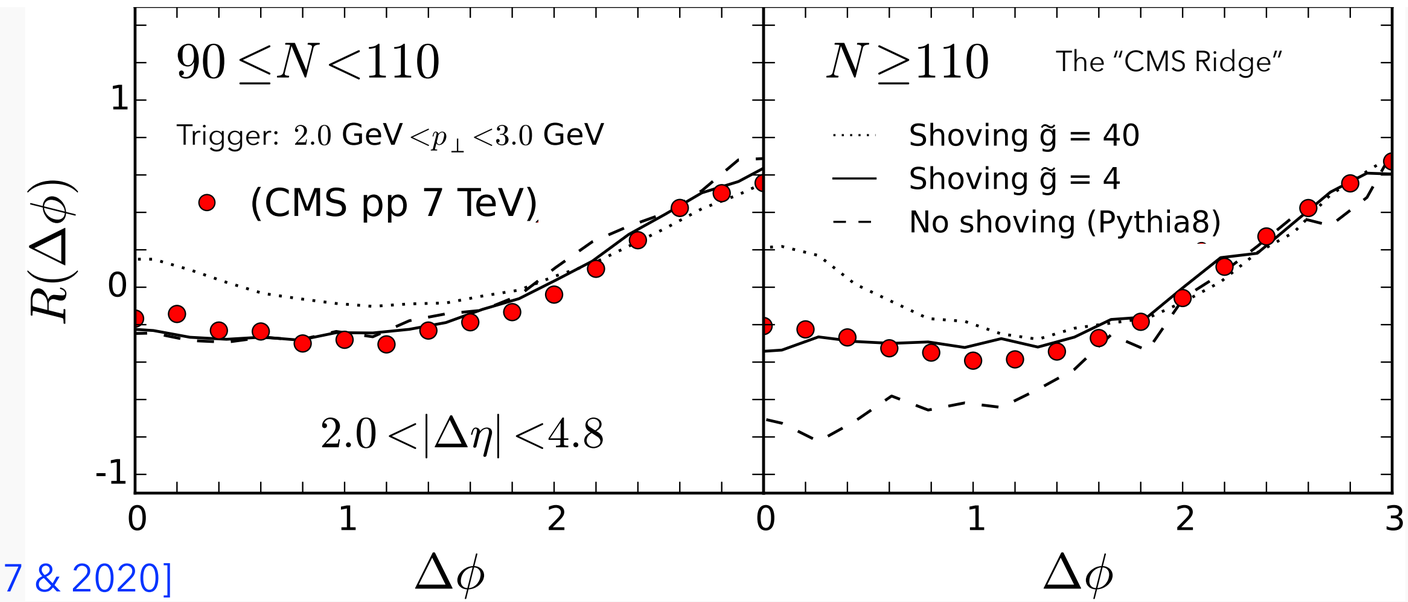
$d_\perp$: transverse distance (in string-string "shoving frame")

$R$: string radius

$\kappa$: string tension ~ $1\,\mathrm{GeV/fm}$

[Bierlich, Chakraborty, Gustafson, Lönnblad, 2017 & 2020]




CMS 1009.4122. Also: ATLAS 1906.08290, ALICE 2101.03110

# Towards a unified picture?

NB: not a community consensus, just my own current best guess

$e^+e^- \to Z/\gamma^* \to$ hadrons

Single String (with kinks)

$e^+e^- \to W^+W^- \to$ hadrons

Two strings

Average multiplicity *pp* min-bias

A handful of strings?

(Abrikosov State)

High-mult. *pp*

UE

AA

pA

Quench?

Central Heavy Ion Collisions

QGP Core

(+ String Corona?)

**PYTHIA**

Lund Model

COLOUR RECONNECTIONS

ROPES + SHOVING

CLOSEPACKING

**ANGANTYR**

(NB: no QGP component)

**EPOS** 100% Corona

Interpolation

**EPOS** 100% Core

# Pythia for Cosmic Rays ↔ Corsika 8



Based on 2 articles by **Marius Utheim** & TS:
"A Framework for Hadronic Rescattering in pp Collisions",
Eur. Phys. J. C80 (2020) 907, arXiv:2005.05658
"Hadron Interactions for Arbitrary Energies and Species, with Applications to Cosmic Rays",
Eur. Phys. J. C82 (2022) 21, arXiv:2108.03481

- Models arbitrary hadron–hadron collisions at low energies.
- Models arbitrary hadron-p/n collisions at any energy.
- Initialization slow, $\sim$ 15 minutes,
  - $\star$ but thereafter works for any hadron–p/n at any energy, and
  - $\star$ initialization data can be saved, so only need to do once.
- The ANGANTYR nuclear geometry part used to extend to hadron-nucleus at any energy.
- Native C++ simplifies interfacing PYTHIA 8 $\leftrightarrow$ CORSIKA 8.
- So far limited comparisons with data.

**+ Extension with ANGANTYR** (→incoming nuclei) ➢ **PYTHIA 8.313**

# Extra Slides

# (Note on the Length of Strings)

In Spacetime:

String tension ≈ 1 GeV/fm → a 50-GeV quark can travel 50 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)

The MC implementation is formulated in momentum space

Lightcone momenta $p_\pm = E \pm p_z$ along string axis

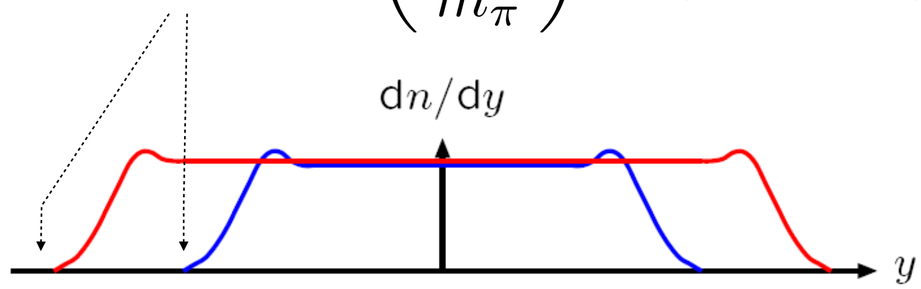
➜ Rapidity (along string axis) and $p_\perp$ transverse to it

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

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




## Particle Production:

Scaling in $z \Longrightarrow$ flat in rapidity (long. boost invariance)

**"Lightcone scaling"**

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

# Particle Composition: Impact on Jet Energy Scale

**ATLAS PUB Note**

ATL-PHYS-PUB-2022-021

29th April 2022

**Dependence of the Jet Energy Scale on the Particle Content of Hadronic Jets in the ATLAS Detector Simulation**

The dependence of the ATLAS jet energy measurement on the modelling in Monte Carlo simulations of the particle types and spectra within jets is investigated. **It is found that the hadronic jet response, i.e. the ratio of the reconstructed jet energy to the true jet energy, varies by ~ 1–2% depending on the hadronisation model used in the simulation. This effect is mainly due to differences in the average energy carried by kaons and baryons in the jet.** Model differences observed for jets initiated by *quarks* or *gluons* produced in the hard scattering process are dominated by the differences in these hadron energy fractions indicating that **measurements of the hadron content of jets and improved tuning of hadronization models can result in an improvement in the precision of the knowledge of the ATLAS jet energy scale.**

Variation largest for gluon jets

- For $E_T$ = **[30, 100, 200] GeV**
- Max JES variation = **[3%, 2%, 1.2%]**

Fraction of jet $E_T$ carried by baryons (and kaons) varies significantly

- Reweighting to force similar baryon and kaon fractions
- Max variation ➜ **[1.2%, 0.8%, 0.5%]**
- Significant potential for improved Jet Energy Scale uncertainties!

Motivates Careful Models & Careful Constraints

- Interplay with advanced UE models
- In-situ constraints from LHC data
- Revisit comparisons to LEP data

# Work in Progress: **Strangeness Enhancement from Close-Packing**

Idea: each string exists in an effective background produced by the others

## Close-packing




$C_8 = 2.25C_F$



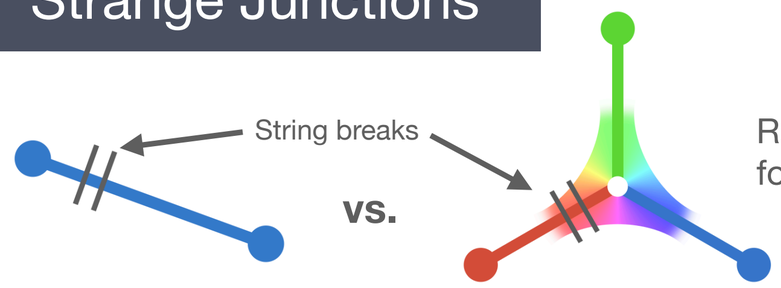

$C_6 = 2.5C_F$

Dense string environments

→ Casimir scaling of **effective string tension**

→ Higher probability of strange quarks

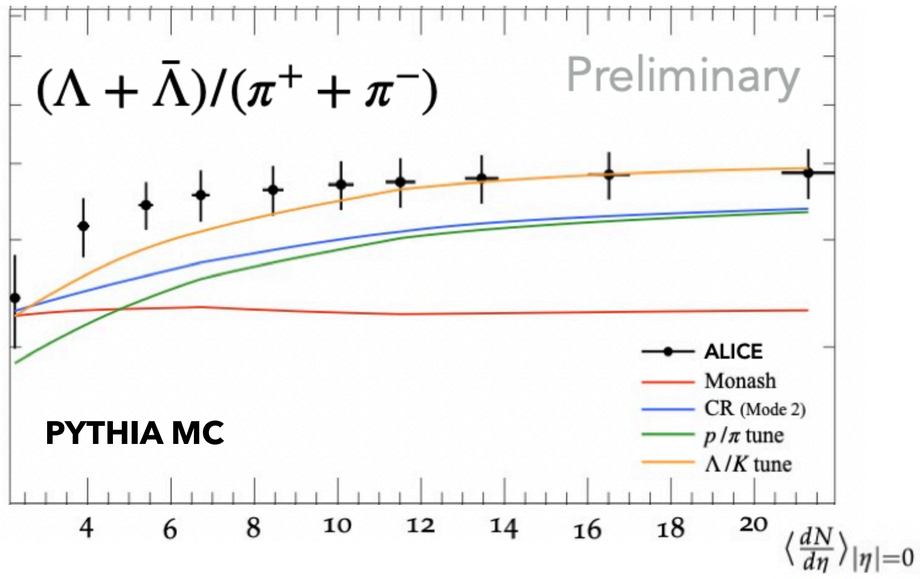
## Strange Junctions

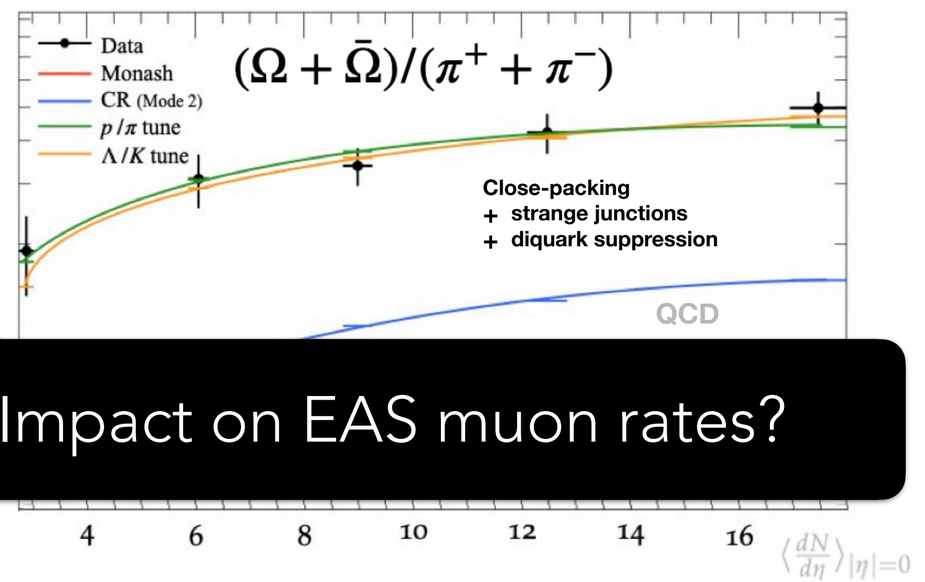



Results in strangeness enhancement focused in baryon sector

String tension could be different from the vacuum case compared to near a junction

Altmann, Bernardinis, Jueid, PS, Zaccolo (in progress)




↔ Impact on EAS muon rates?

*Slide adapted from J. Altmann*

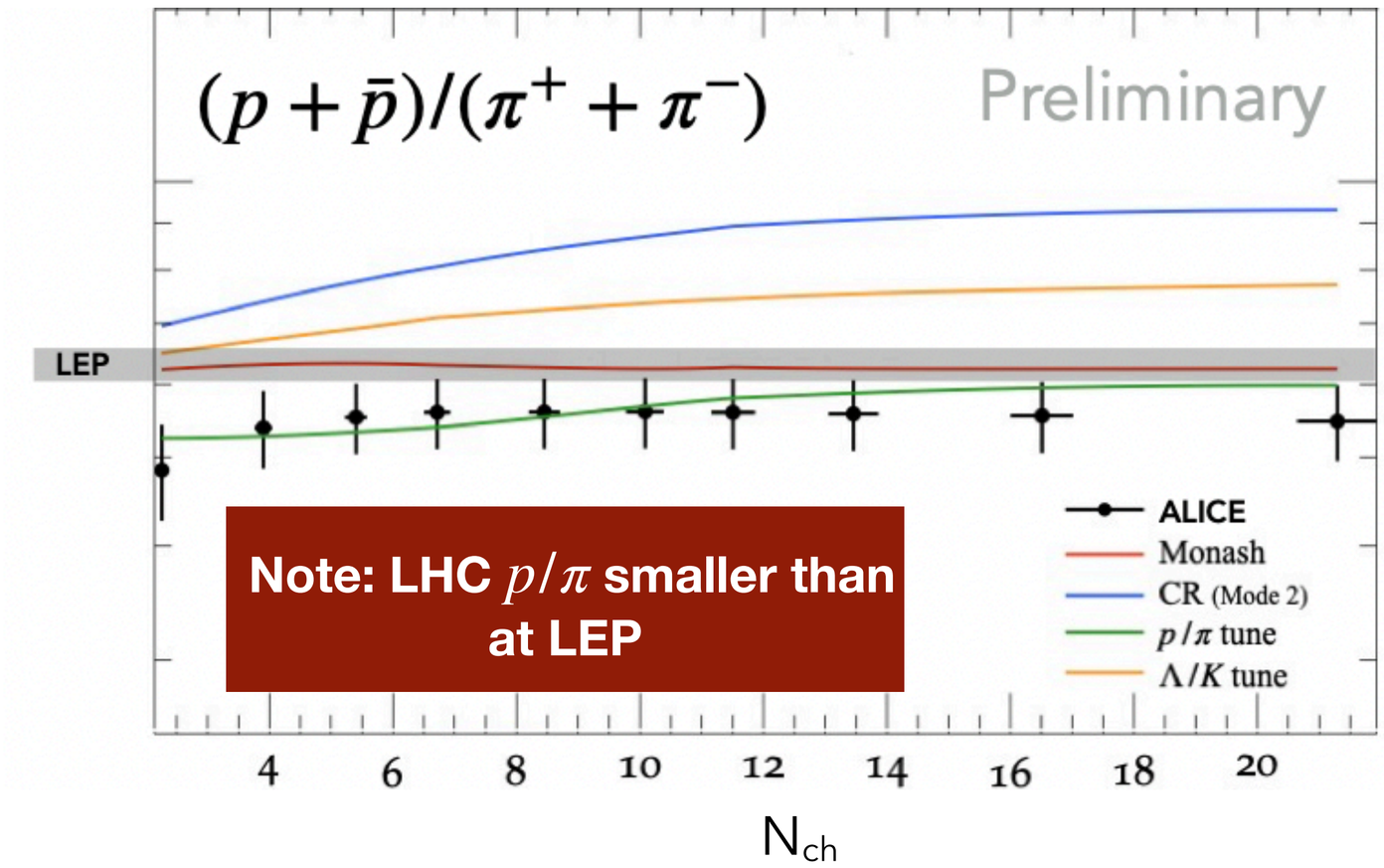
# Thorny Issue ⚠️ The **Proton-to-Pion** Ratio

**Note:**

Observed $p/\pi$ in pp collisions at LHC is ***lower*** than in $e^+e^-$ ones (LEP).

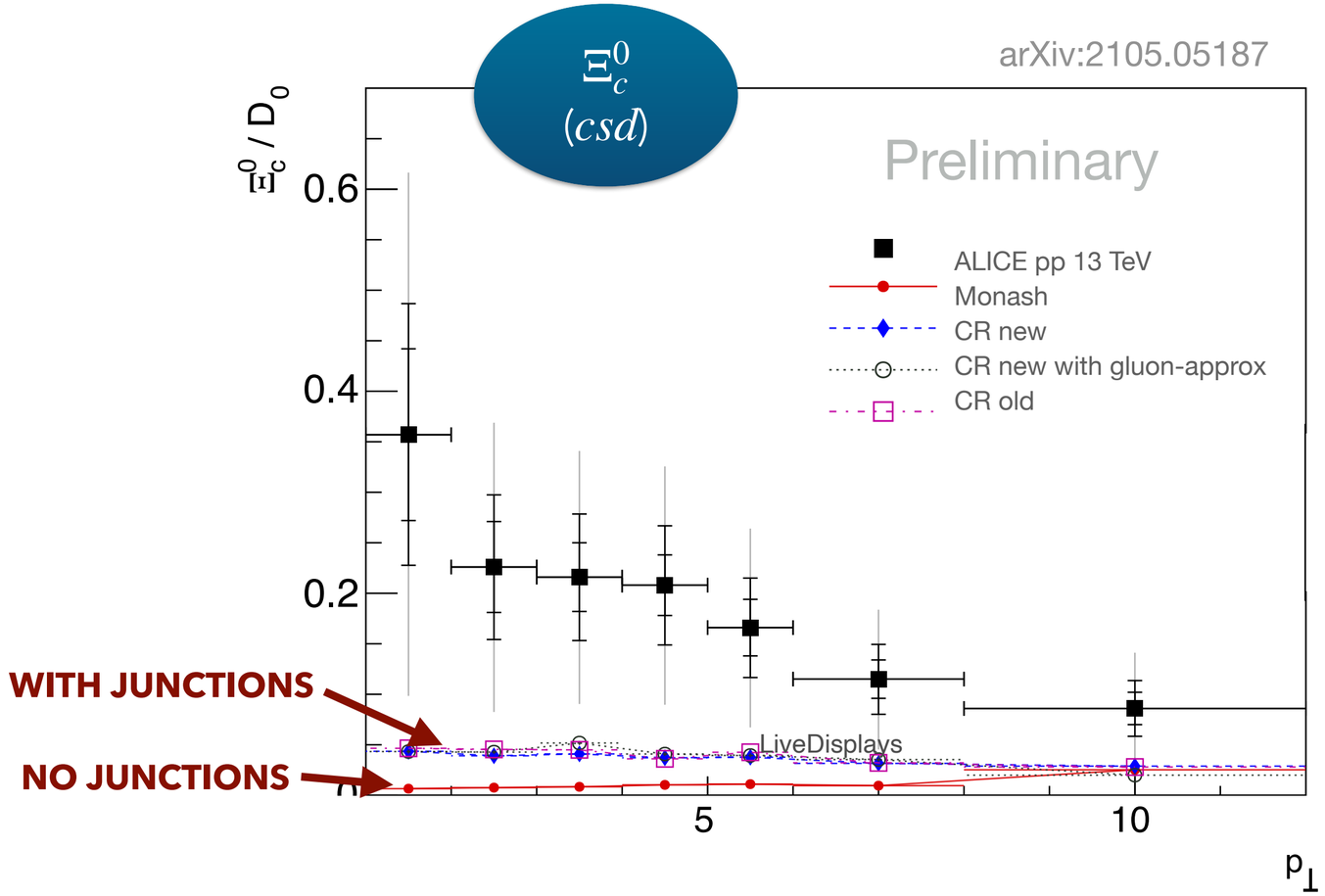
**I think** this is now the **main challenge** for strangeness-enhancement models

*Interactions? Upscattering/Annihilation? Octet vs Triplet fragmentation? …?*

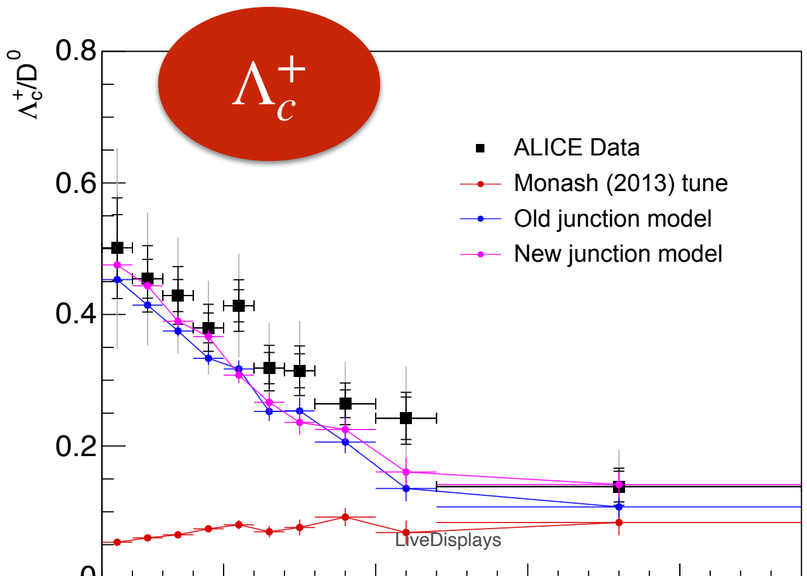


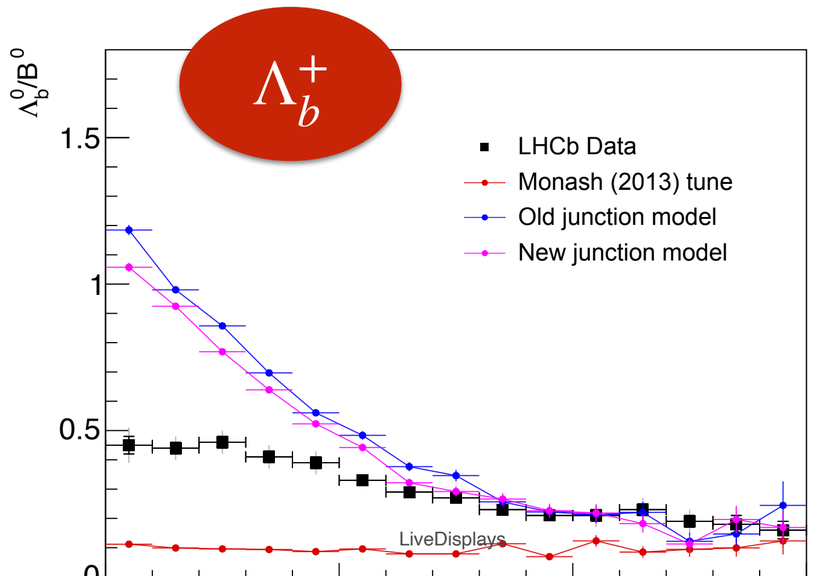
# Confront with Measurements: **Strangeness**

What about Strange heavy-flavour baryons ?



# LambdaB




***String Formation Beyond Leading Colour,* Christiansen & PZS, 1505.01681**

**New: *String Junctions Revisited,* Altmann & PZS, 2404.12040**

Also: baryon asymmetry diluted by extra baryon pairs

