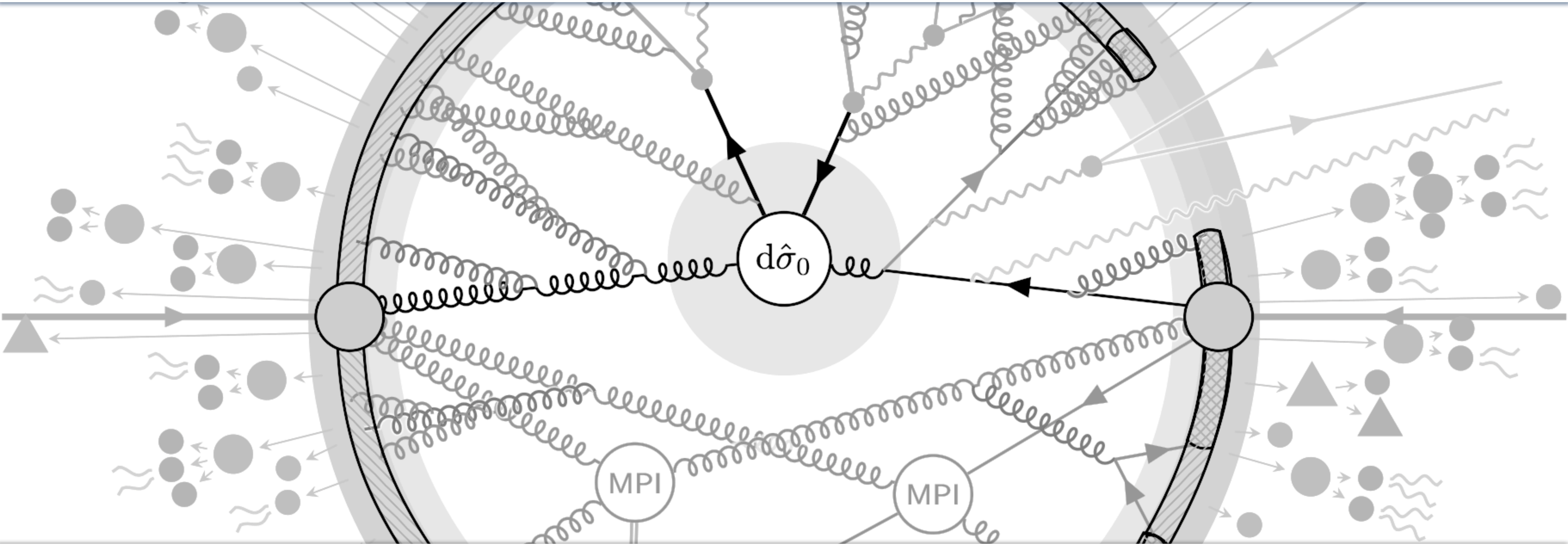
Anatomy of an LHC Collision — and Challenges for the Future



Peter Skands (Monash University)

June 2022

LHC Collisions – Theory vs Real Life

Theory Goal: Use LHC measurements to test hypotheses about Nature.

The ATLAS Experiment at the LHC

ATLAS collision event at 7 TeV from March 2010

http://atlas.ch

But we have no exact solutions to (B)SM Quantum Field Theories.

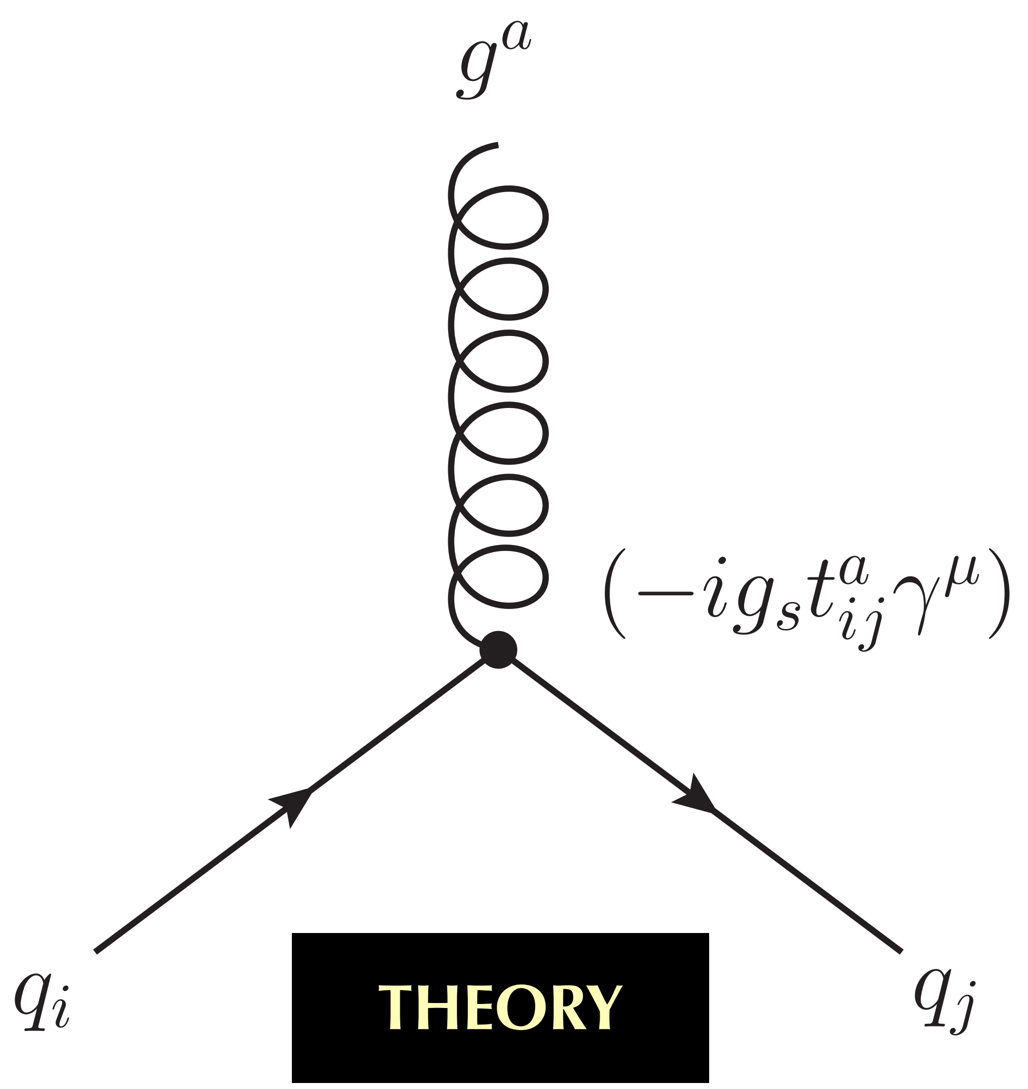
How to make predictions to form (reliable) conclusions?

Confounded by Confinement

We are colliding — and observing — hadrons

Strongly bound states of quarks and gluons (non-perturbative QCD)

How do we connect this…



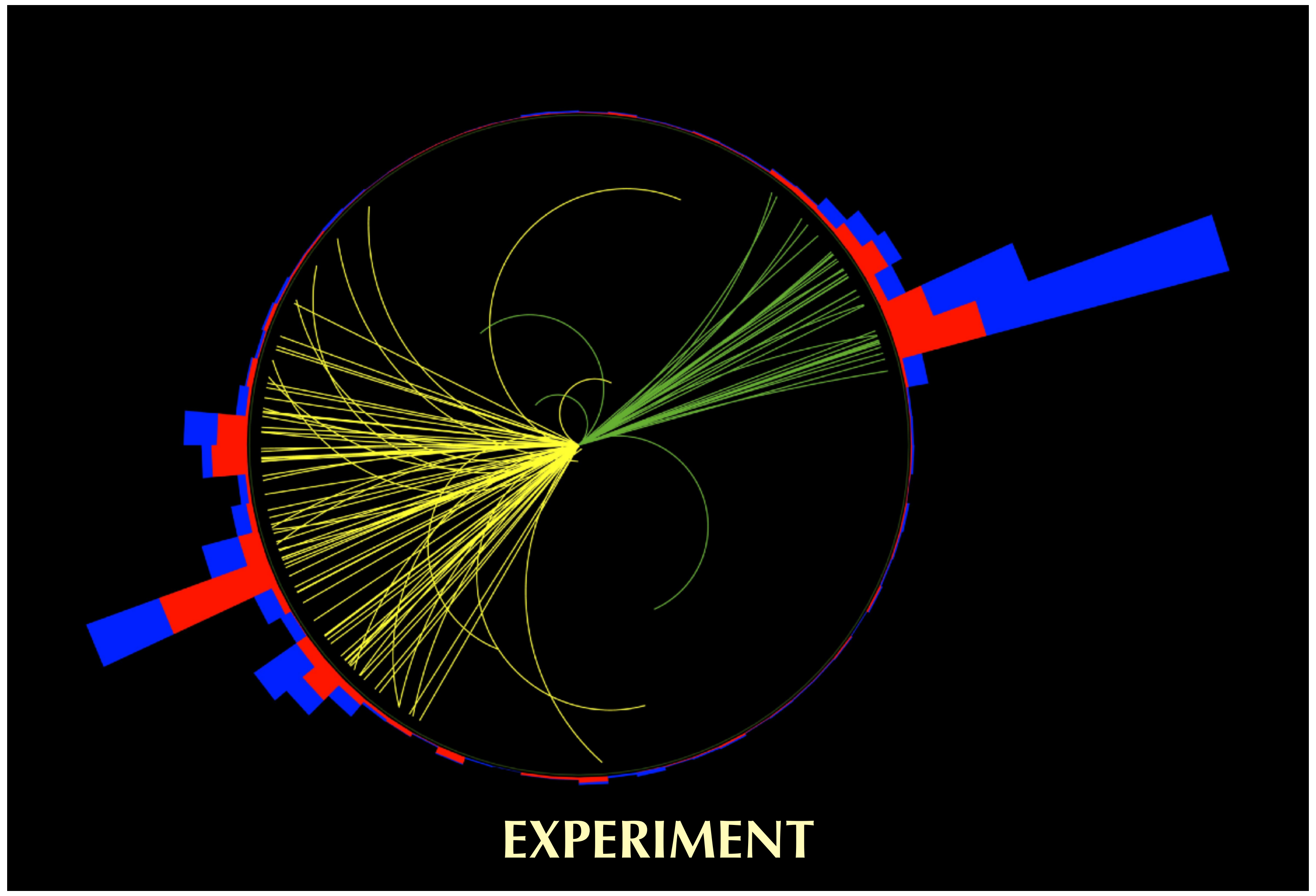
Elementary Fields & Symmetries

"Fundamental" parameters.

Asymptotic freedom, perturbative QFT

CONFINEMENT

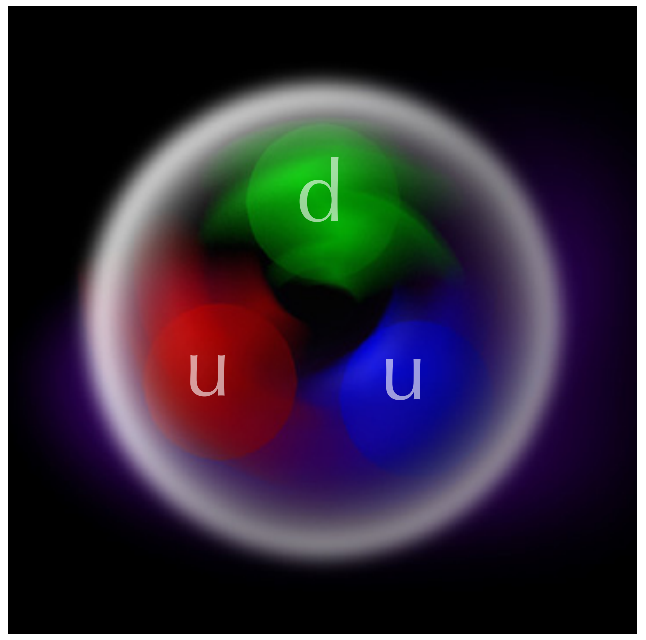
… with this?



"Emergent" degrees of freedom

Jets of hadrons

Consider a hadron; why is it complicated?



Textbook "quark-model" proton:

"Three quarks for muster Mark" (Gell-Mann/Joyce)

Quark-model flavour ⊗ spin wave functions

Real-life hadrons

Are composite & strongly bound, with time-dependent structure

For wavelengths ~ confinement scale:

quark & gluon plane waves are not going to be good approximations

⟹ **forget about the interaction picture and perturbation theory**

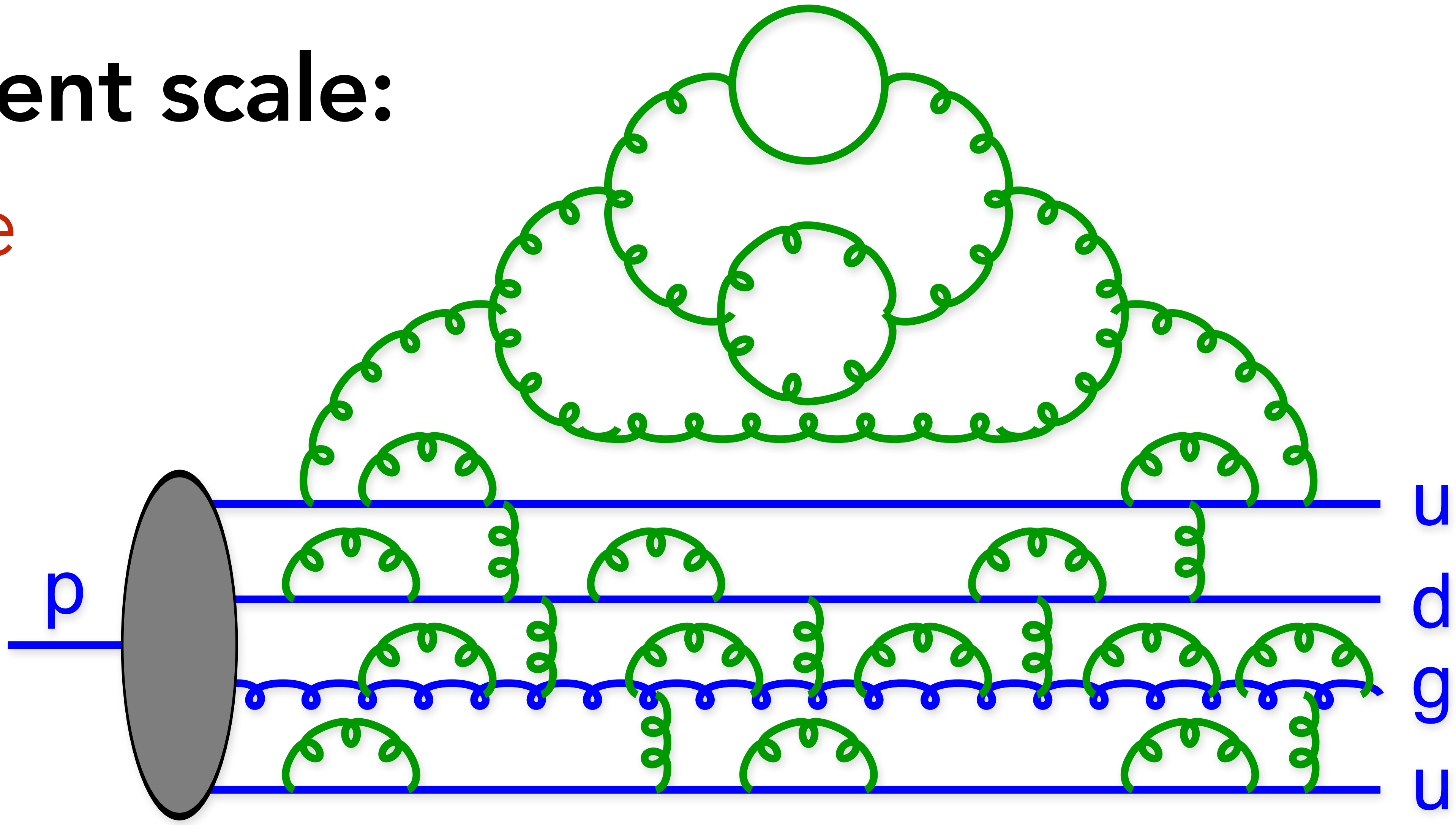


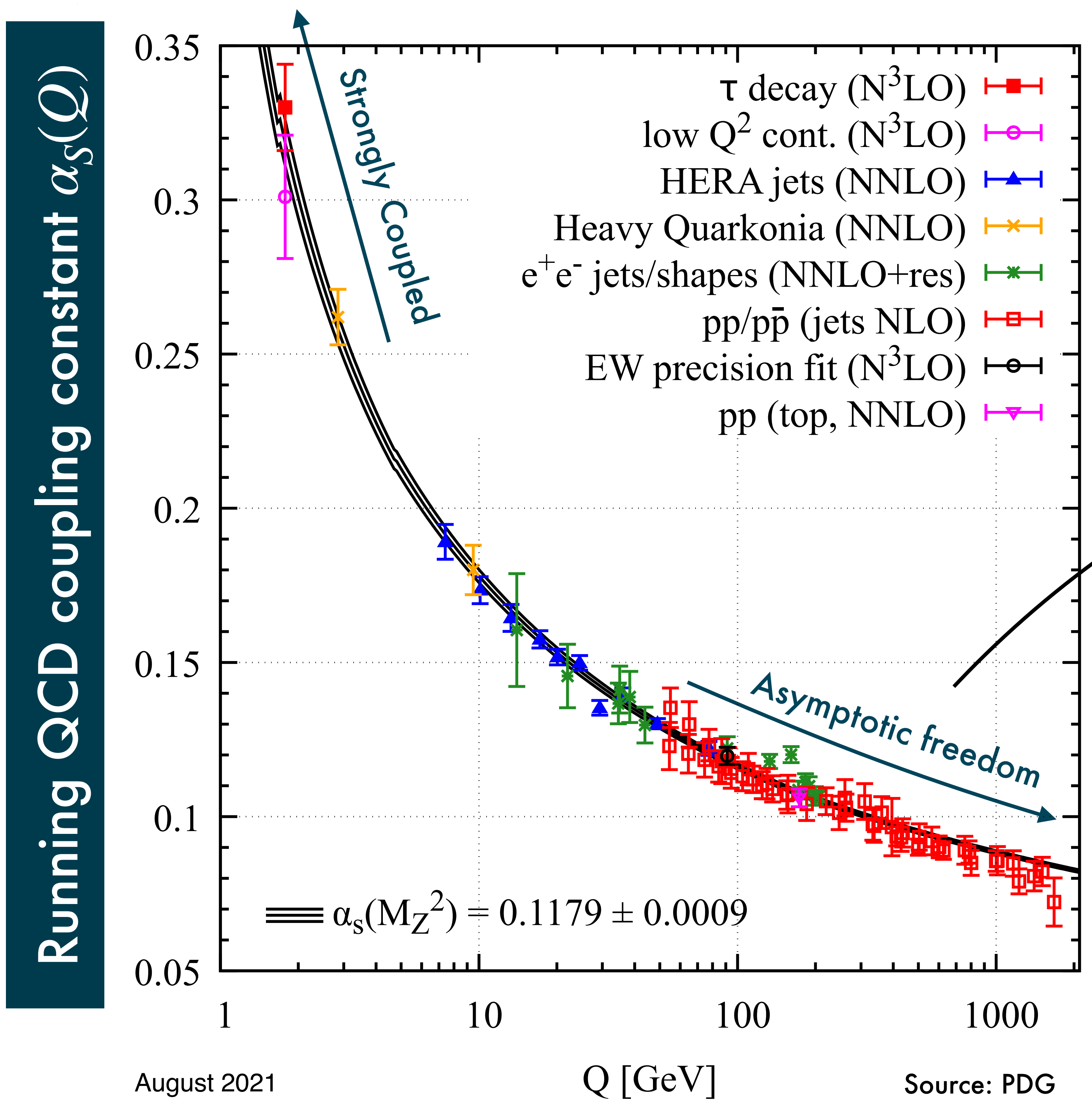
Figure by T. Sjöstrand

What about shorter wavelengths?

Nobel Prize 2004: Asymptotic Freedom in QCD (Gross, Politzer, Wilczek)

Over **short** distances, quarks and gluons **do** behave like *almost* free particles

Then it's OK to start from free-field solutions (plane waves) and treat interactions as perturbations $\Longrightarrow$ The interaction picture and perturbation theory are saved!



August 2021

Source: PDG

Parametrise "mess" in terms of (measurable) probability densities for each type of plane wave:
Parton Distribution Functions (PDFs)

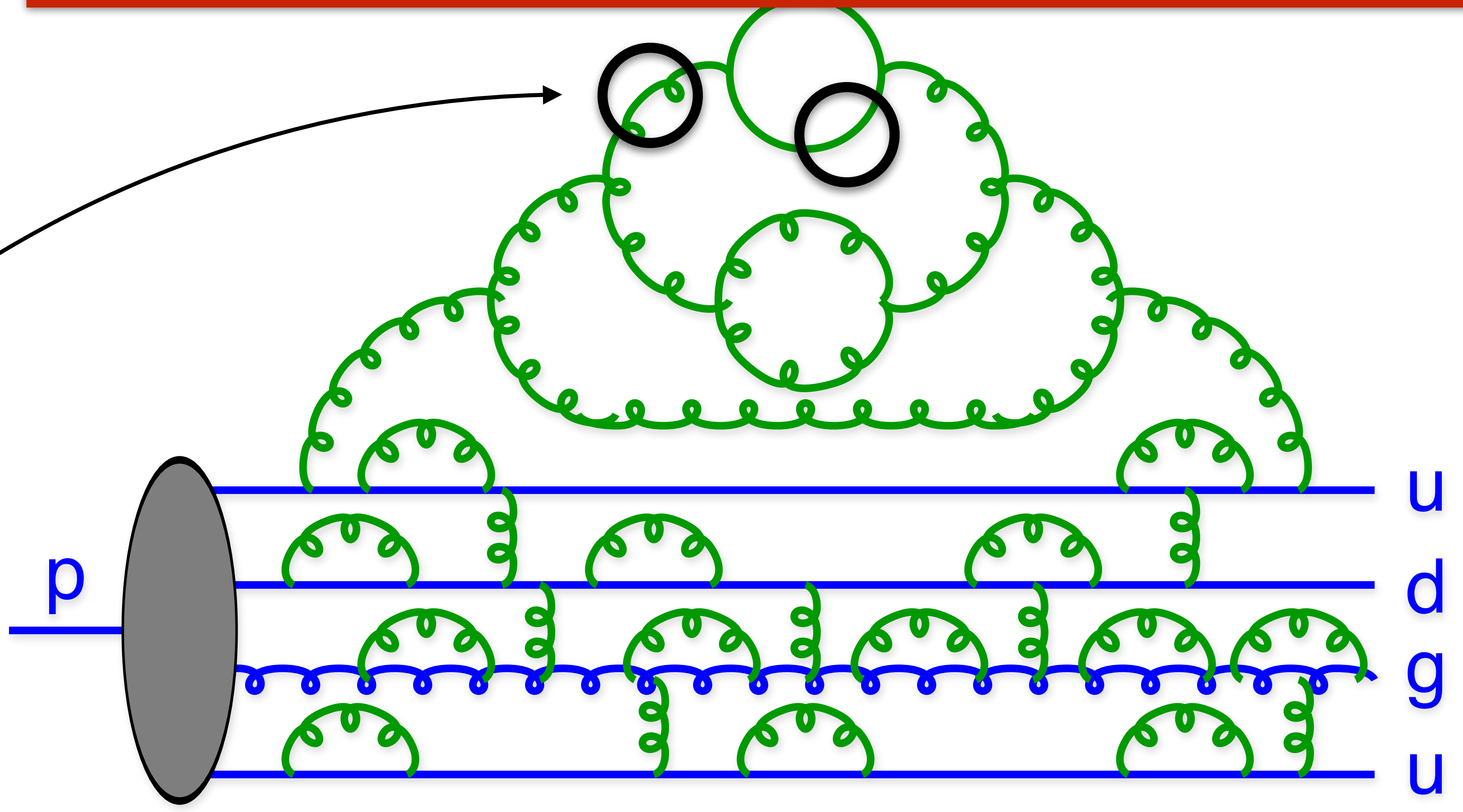


Figure by T. Sjöstrand

Mathematically, the cross section **factorises**

(Collins, Soper, '87)

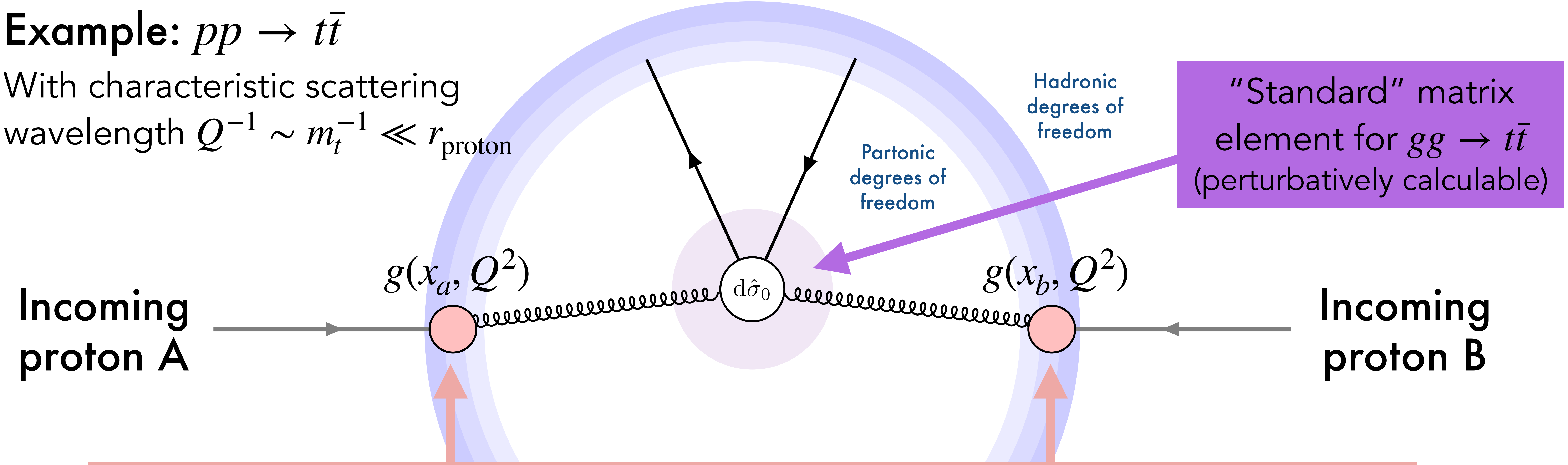
Hadron-level cross sections can be computed as (sums over):

Perturbative Parton-level cross sections $\otimes$ **Parton Distribution Functions**

Thus, we can compute, e.g., the total top-quark-pair cross section we expect at LHC:

Example: $pp \to t\bar{t}$

With characteristic scattering wavelength $Q^{-1} \sim m_t^{-1} \ll r_{\text{proton}}$



"Standard" matrix element for $gg \to t\bar{t}$ (perturbatively calculable)

Probability densities for finding gluons inside protons A and B

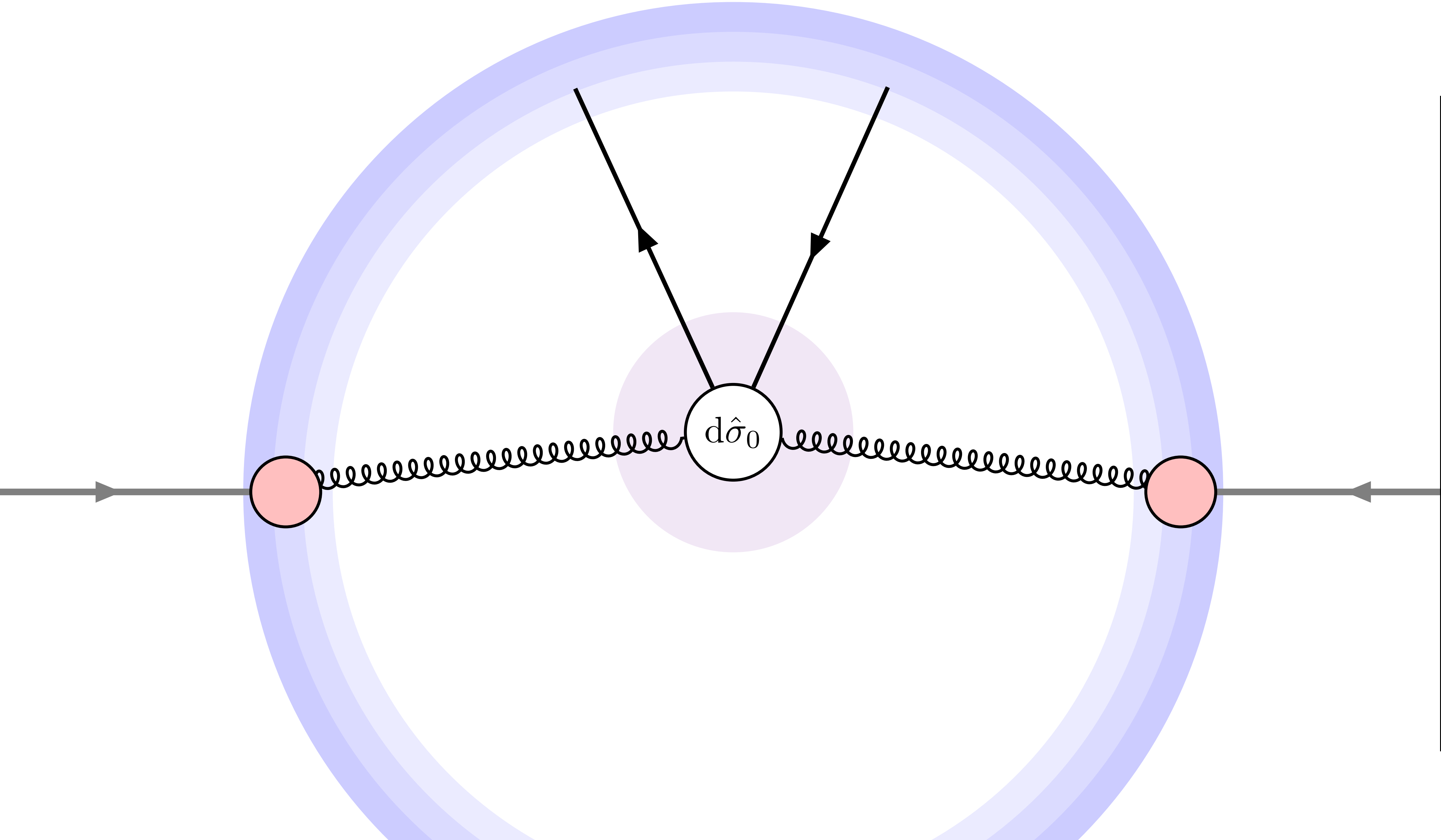
(carrying fractions x_a and x_b of the respective proton energies)

These (+ their quark equivalents) have been extensively measured at previous colliders (esp. HERA); increasingly now also at LHC itself.

Compare with measurements

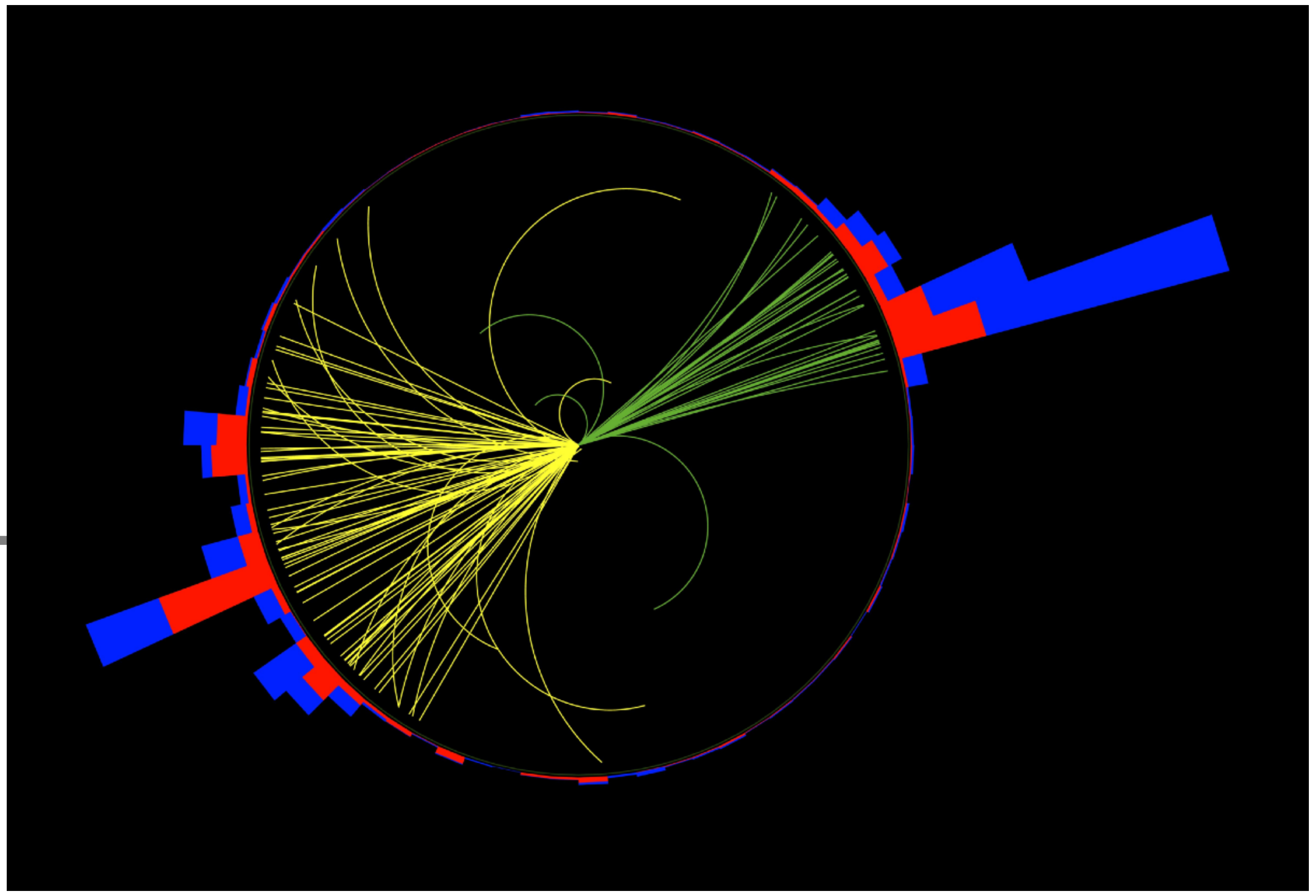
Theorist:

This is a $t\bar{t}$ event



Experimentalist:

Is this a $t\bar{t}$ event?



With factorisation, we recover the use of perturbation theory (for high-Q processes*)

But we also lose a lot of detail (and still cannot address low Q)

---

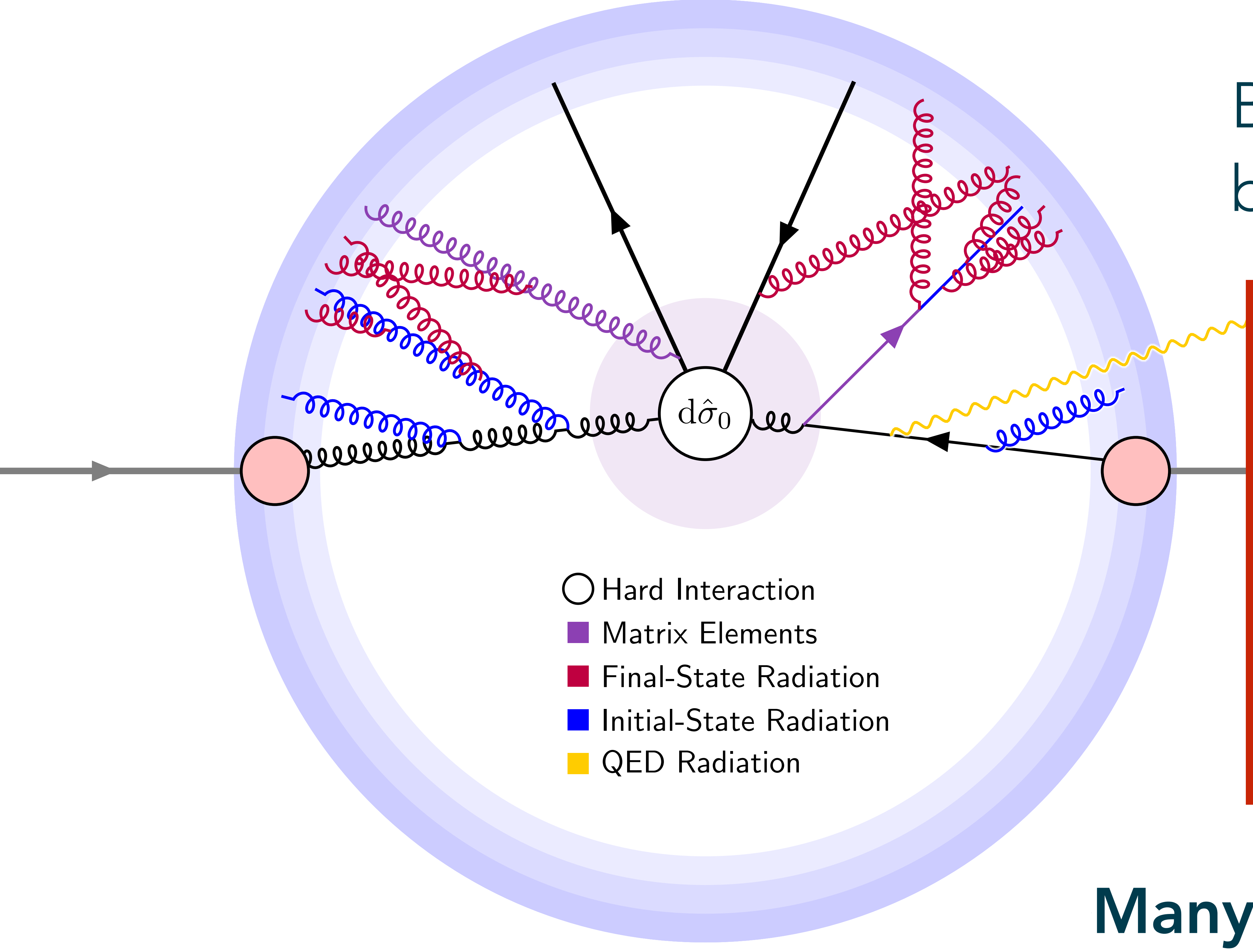
*for so-called Infrared and Collinear Safe Observables

Accuracy & Detail 1: Radiative Corrections

The scattered partons carry QCD and/or electric charges

Will give off bremsstrahlung radiation, at wavelengths > 1/Q.

Probabilities can be computed order by order in perturbation theory



But the leading (~classical) effects can also be (re)summed to ∞ perturbative order.

Can be achieved numerically by Markov-Chain Monte Carlo algorithms which iterate **factorised** emission probabilities:

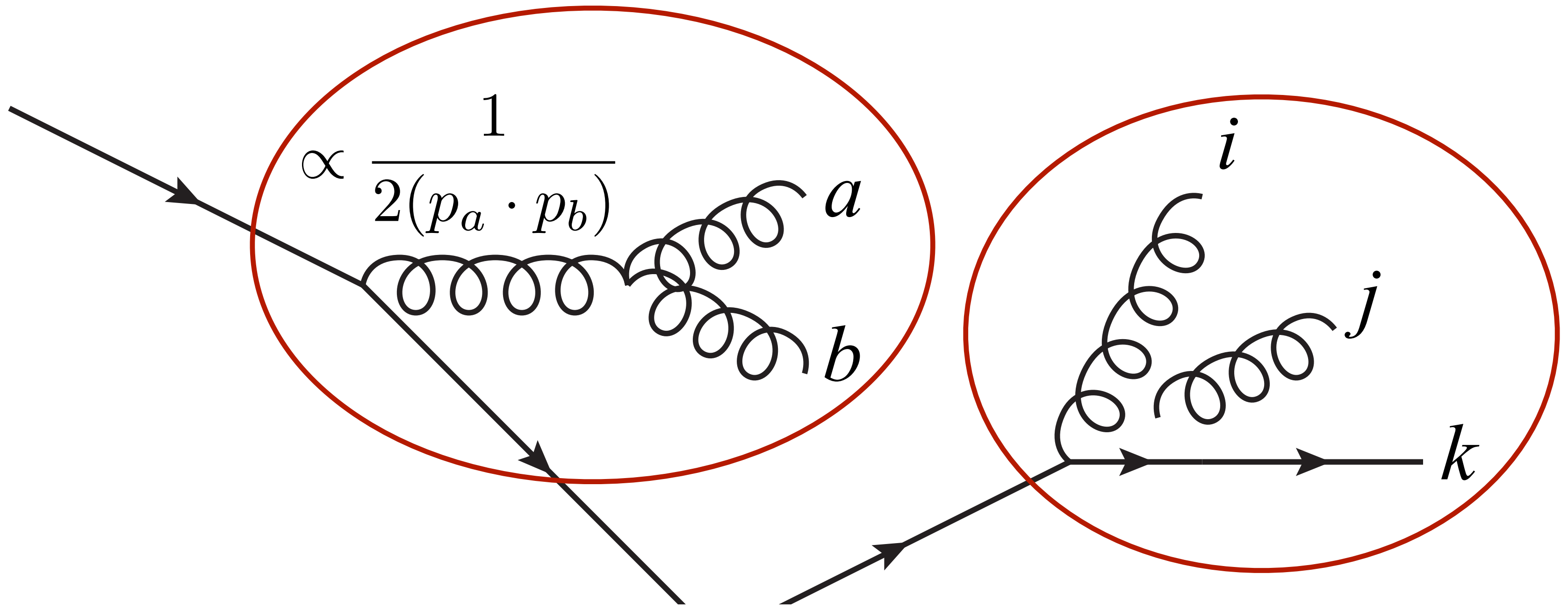
➤ **Parton Showers**

E.g.: Sjöstrand ('85, '86, '87), Marchesini & Webber ('84, '87, '88), Gustafson ('88) + many more recent

Many new efforts over the past decade!

Parton Showers = Iterated Sums over "Radiation Kernels"

Most bremsstrahlung is driven by **divergent propagators** → simple universal structure, independent of process details



Amplitudes *factorise* in singular limits

In **collinear** limits, we get so-called DGLAP splitting kernels:

$$|\mathcal{M}_{F+1}(\ldots,a,b,\ldots)|^2 \overset{a||b}{\to} g_s^2\mathcal{C}\,\frac{P(z)}{2(p_a\cdot p_b)}|\mathcal{M}_F(\ldots,a+b,\ldots)|^2$$

In **soft** limits ($E_g/Q \to 0$), we get dipole factors (same as classical):

$$|\mathcal{M}_{F+1}(\ldots,i,j,k\ldots)|^2 \overset{j_g\to 0}{\to} g_s^2\mathcal{C}\,\frac{(p_i\cdot p_k)}{(p_i\cdot p_j)(p_j\cdot p_k)}|\mathcal{M}_F(\ldots,i,k,\ldots)|^2$$

These limits are not independent; they overlap in phase space.

How to treat the two consistently has given rise to **many** individual approaches:

Angular ordering, angular vetos, dipoles, global antennae, sector antennae, …

After 40 years of development, how far have we got?

In fixed-order perturbative QCD (pQCD):

LO → NLO → NNLO → N^{3}LO ⟵ State of the art for simple processes

State of the art for complex processes

Translates to accuracies of order a few per cent or better

For all-orders showers, it makes no sense to count "orders"

Instead, people count "logarithms" (arising from $1/Q^2$ propagators on previous slide integrated over phase spaces $\propto \mathrm{d}Q^2$)

Counting logs is not the only way to judge (ignores other important aspects), but:

Angular ordering (80s): (N)LL

Modern **dipole/antenna** showers: (N)LL

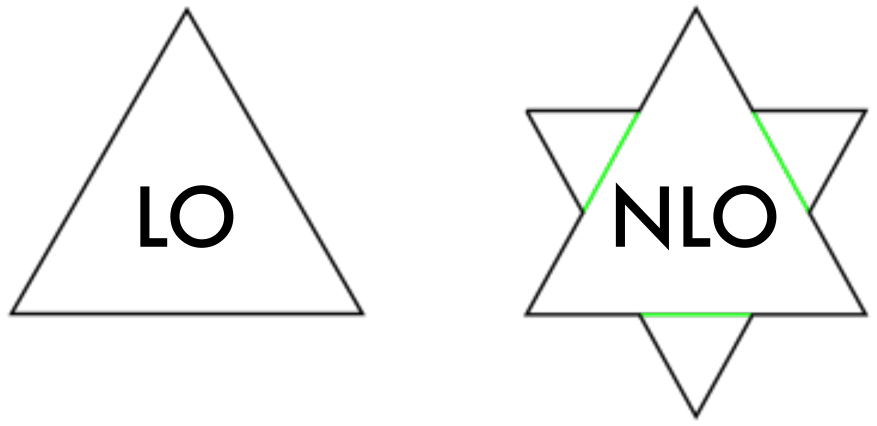
Colour flow also still "leading colour"

(with small refinements)

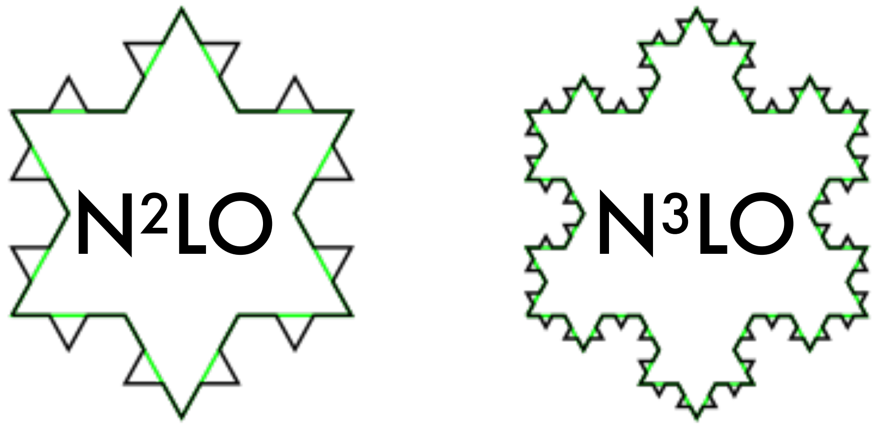
Last remaining "leading" frontiers in pQCD

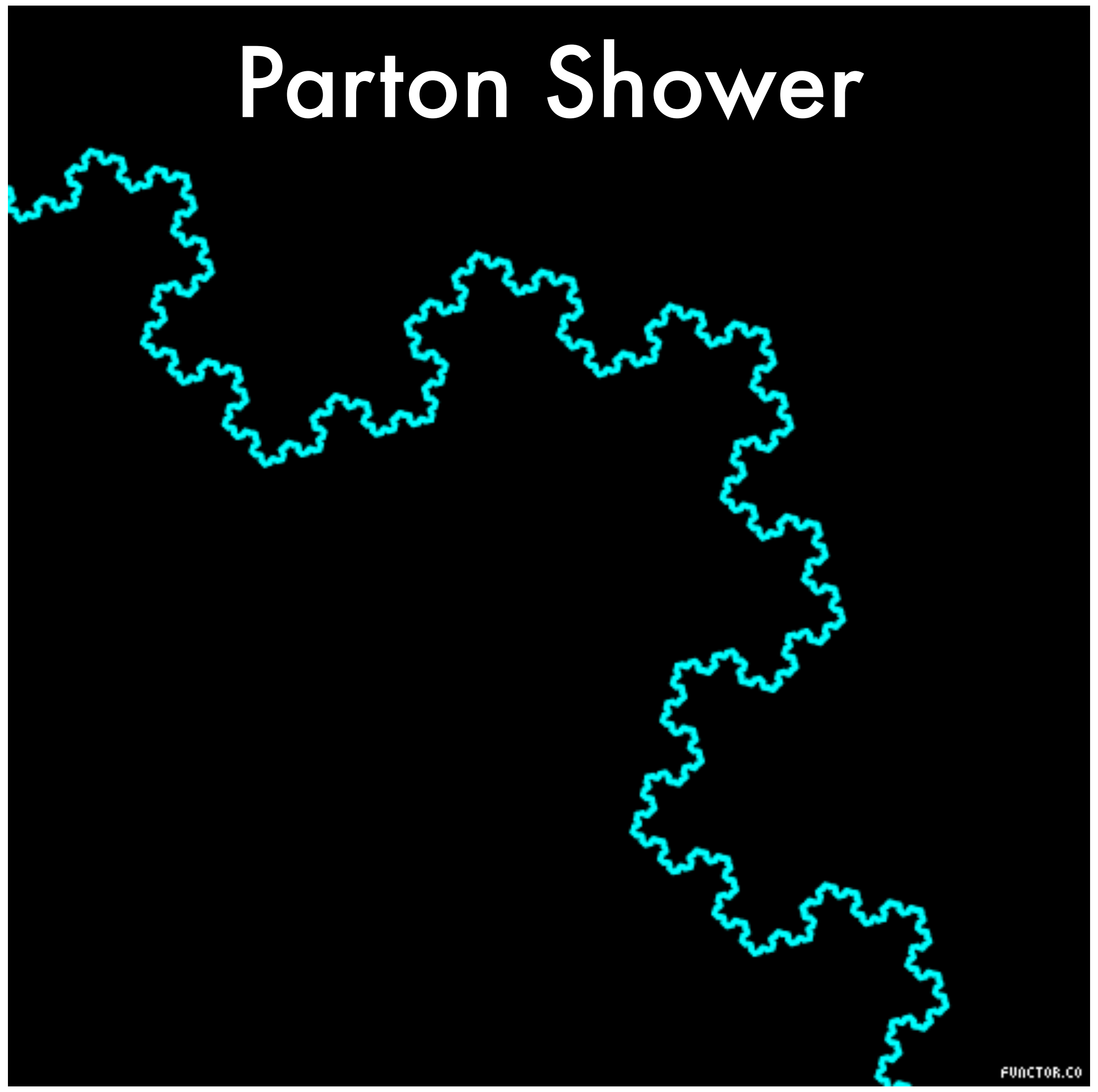
Why is that hard?

Simplified analogy:



Using a "Koch snowflake" as a stand-in for perturbation theory





Some Complications:

Showers are **quantum stochastic processes**, not deterministic rules

Several branching types, on **multiparton phase spaces** (beware overlaps/double-counting/dead zones)

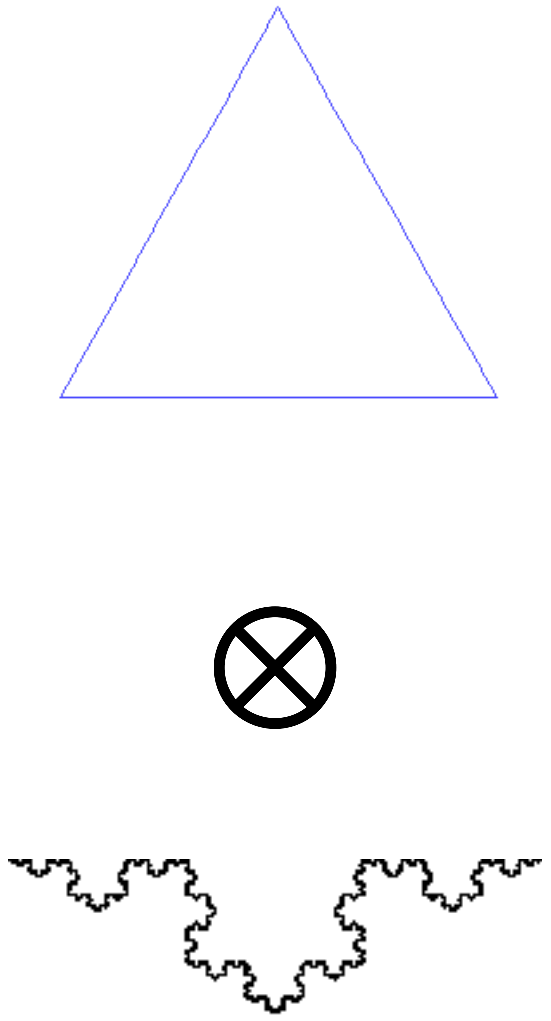
With SU(3) **colour** structure, **spin/polarisation** structure, and quantum **interference**

Universality: start from *any* hard process (~ starting "shape"); + **scaling violation**.

Conservation Laws: must be momentum conserving, and Lorentz & gauge invariant.

Unitarity: must have *perfect* cancellations between (singular) real and virtual corrections.

Well Established for **First Few Orders**



Matching, Merging, and Matrix-Element Corrections

Essentially: use exact rule for first few orders; then let shower approximation take over

- LO matrix-element corrections (➤ Sjöstrand et al., 80s)
- LO merged calculations (➤ CKKW, Lönnblad, '00s + more recent)
- NLO matched calculations (➤ MC@NLO, POWHEG '00s)

State of the art (for LHC phenomenology right now):

Merging several NLO + PS matched calculations (➤ UNLOPS, FxFx, …)

Intense activity; here just using "my" projects as representative examples:

NNLO + PS matching (Proof of concept ➤ Campbell, Hoeche, Li, Preuss, PS, '21)

Iterated LO matrix-element corrections (➤ soon…)

Iterated NLO matrix-element corrections (➤ in a while 🐊)

Limiting factors are **complexity growth** & **shower accuracy**

Complexity Growth: a bottleneck for matching and merging

In conventional ("global") **showers, each phase-space point receives contributions from many possible branching "histories"** (="clusterings")

~ sum over (singular) diagrams $\Longrightarrow$ full singularity structure ✅

Number of Histories for n Branchings (Starting from a single $q\bar{q}$ pair)

	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$	$n=7$
CS Dipole	2	8	48	384	3840	46080	645120
Global Antenna	1	2	6	24	120	720	5040

Fewer partial-fractionings, but still factorial growth

For CKKW-L style merging: (incl UMEPS, NL3, UNLOPS, …)

Need to take all contributing shower histories into account.

Bottleneck at high multiplicities (+ high code complexity)

Sector Showers

New in Pythia 8.304: Sectorized Antenna Showers in Vincia

PartonShowers:Model = 2 Brooks, Preuss & PS 2003.00702 (+ Lopez-Villarejo & PS 1109.3608)

Sector antennae: no partial-fractioning of any singularities.

Divide the *n*-gluon phase space up into *n* non-overlapping sectors, inside each of which **only the most singular** (~"classical") kernel is allowed to contribute.

Kosower, hep-ph/9710213
hep-ph/0311272 (+ Larkoski & Peskin 0908.2450, 1106.2182)

Lorentz-invariant def of most singular gluon based on ARIADNE p_T:

Gustafson & Pettersson, NPB 306 (1988) 746

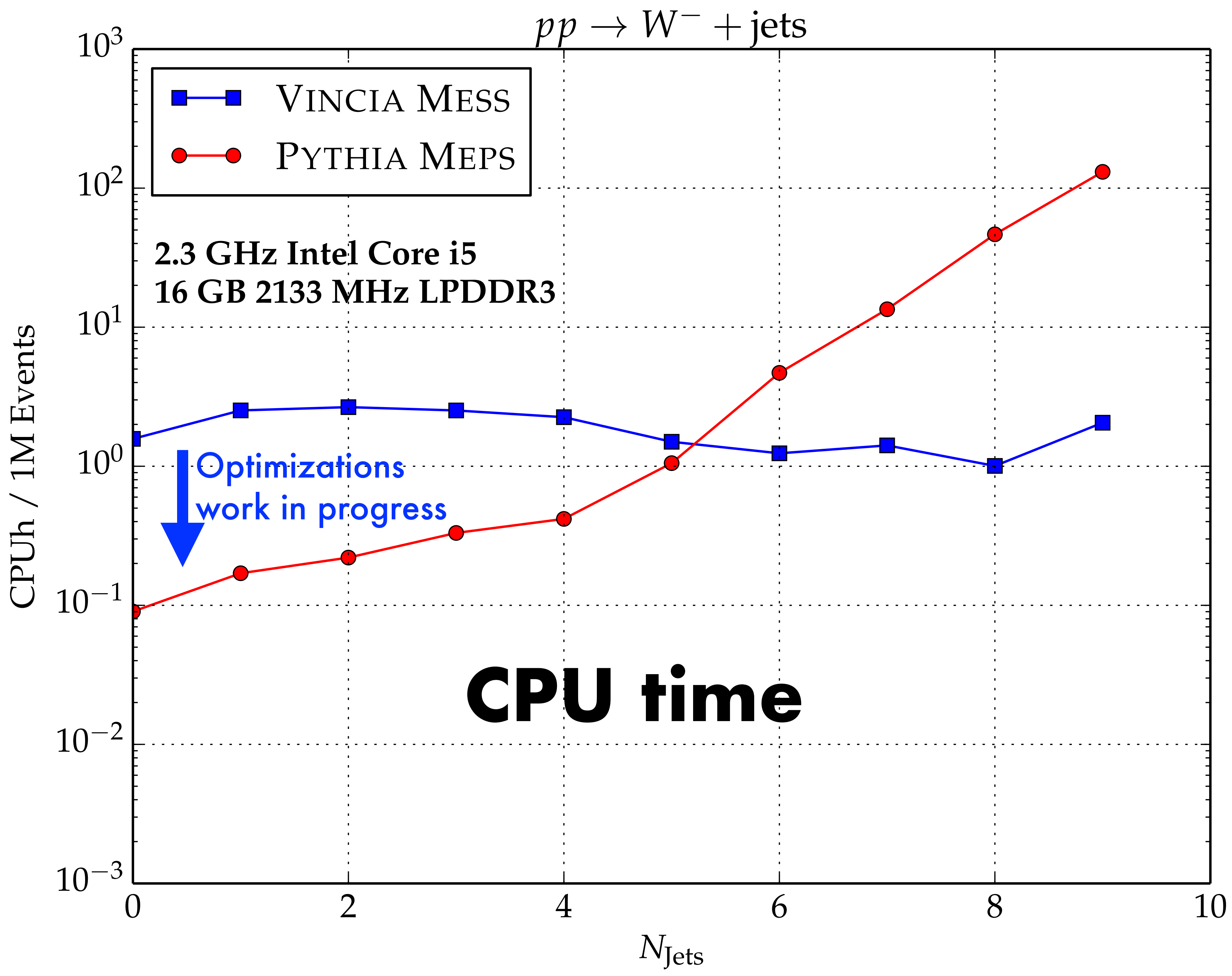
$$p_{\perp j}^2 = \frac{s_{ij}s_{jk}}{s_{ijk}}$$ with $s_{ij} \equiv 2(p_i \cdot p_j)$ (+ generalisations for heavy-quark emitters)

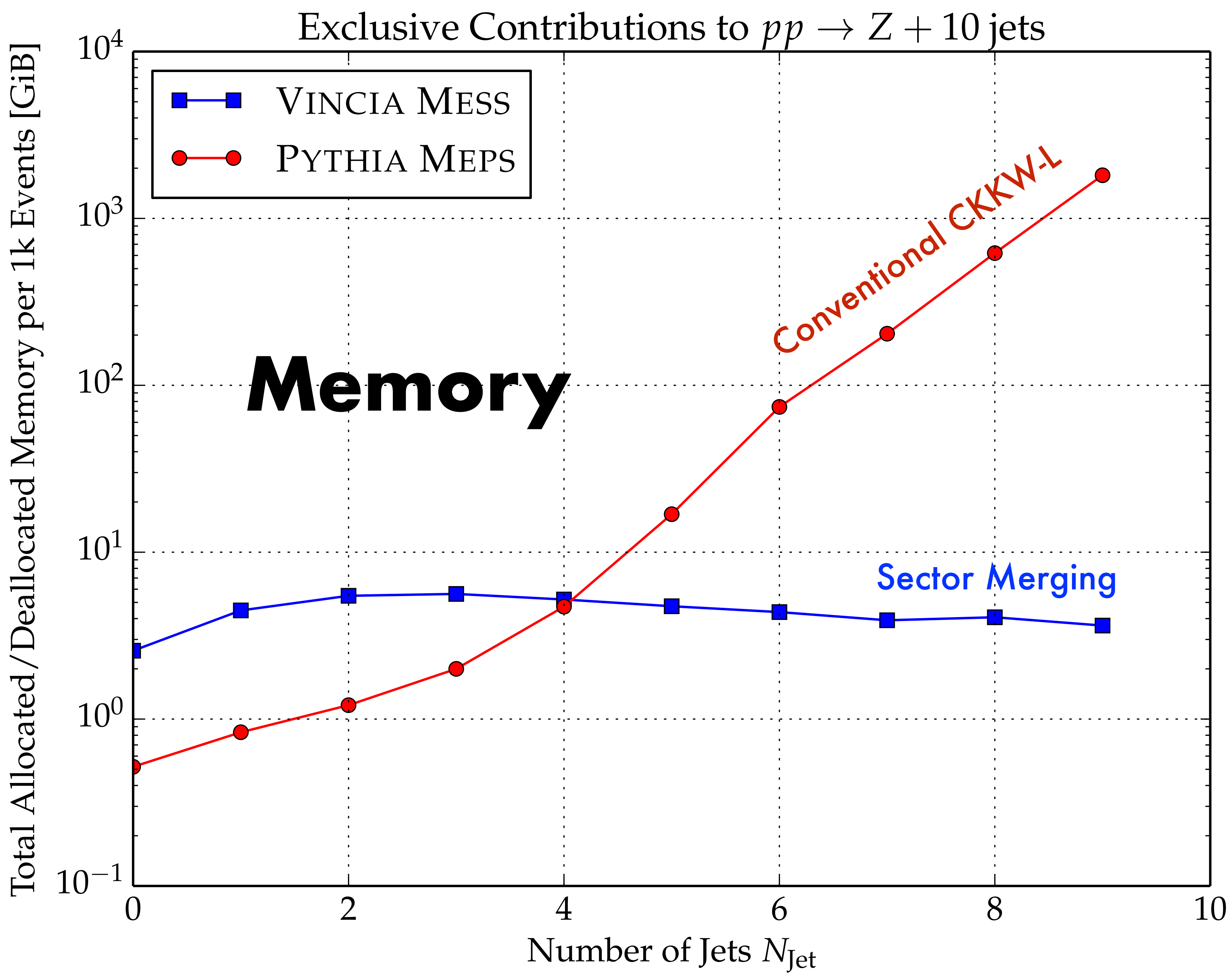
Achieves (N)LL with a single history.

Factorial → **constant scaling** in number of gluons.

Generalisation to $g \to q\bar{q} \Longrightarrow$ factorial in number of same-flavour quark pairs.

New: Sectorized CKKW-L Merging in Pythia 8.306





Brooks & Preuss, "Efficient multi-jet merging with the VINCIA sector shower", 2008.09468

Ready for serious applications (Note: Vincia also has dedicated POWHEG hooks)

Work ongoing to optimise baseline algorithm.

Work at Fermilab: **NNLO** matching, 2 → 4 sector antennae, **MCFM** interface, …

Vincia tutorial: http://skands.physics.monash.edu/slides/files/Pythia83-VinciaTute.pdf

The Final Frontier: **Shower Accuracy**

Second-order radiative corrections

Iterating **only** single-emission probabilities will ultimately fail to describe multi-emission correlations & interferences

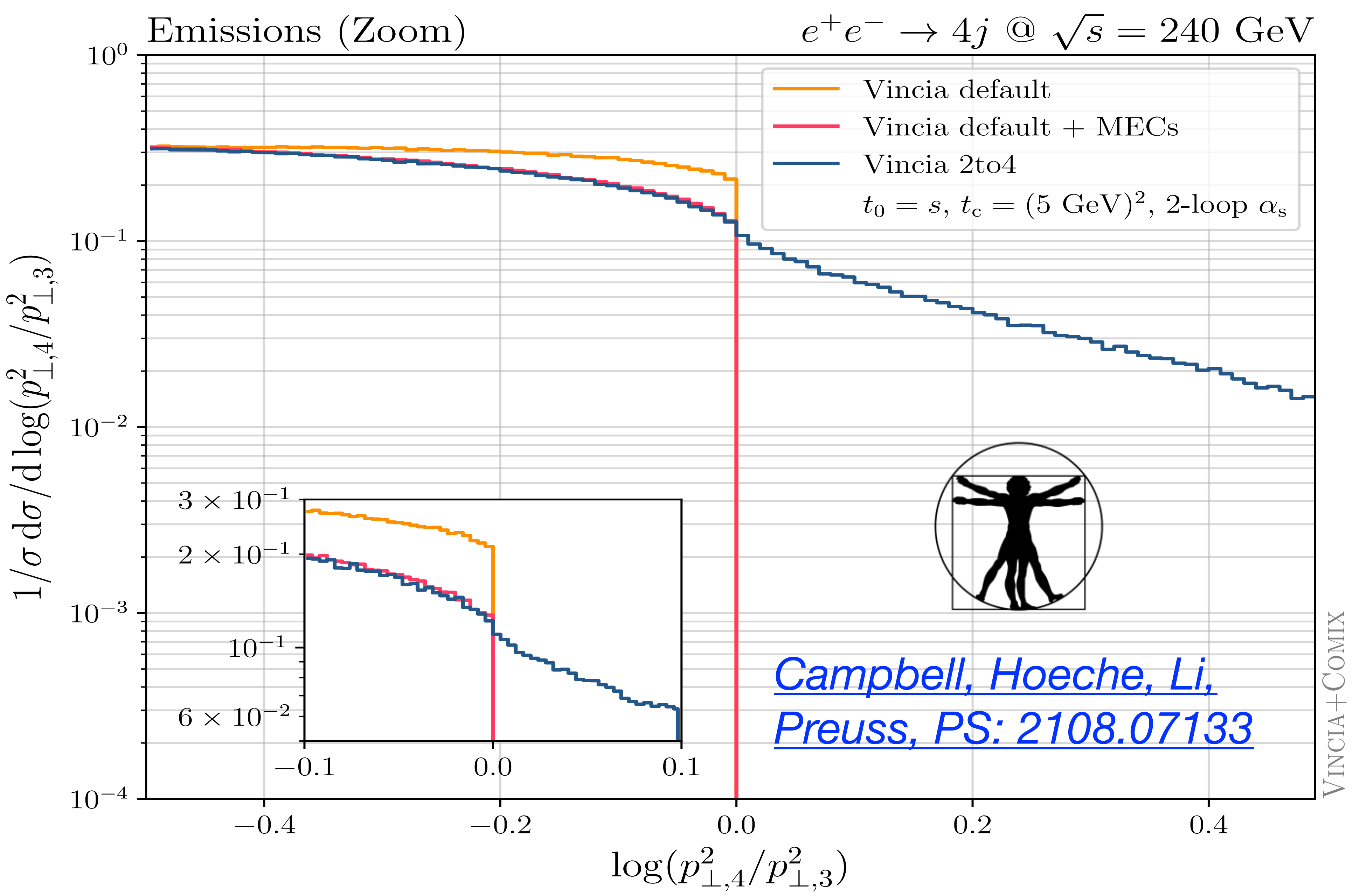
Hard to iterate single **and** double emissions without any overlaps.

VINCIA sector approach

Clean separation of phase space into "ordered" and "unordered" sectors (using single-valued resolution Q)

Pieces look ready. Proof of concept for iterated single emissions (augmented by virtual & double-emission MEC factors) + "direct" double-emissions

Goal: iterate full structure; not there yet



Campbell, Hoeche, Li, Preuss, PS: 2108.07133

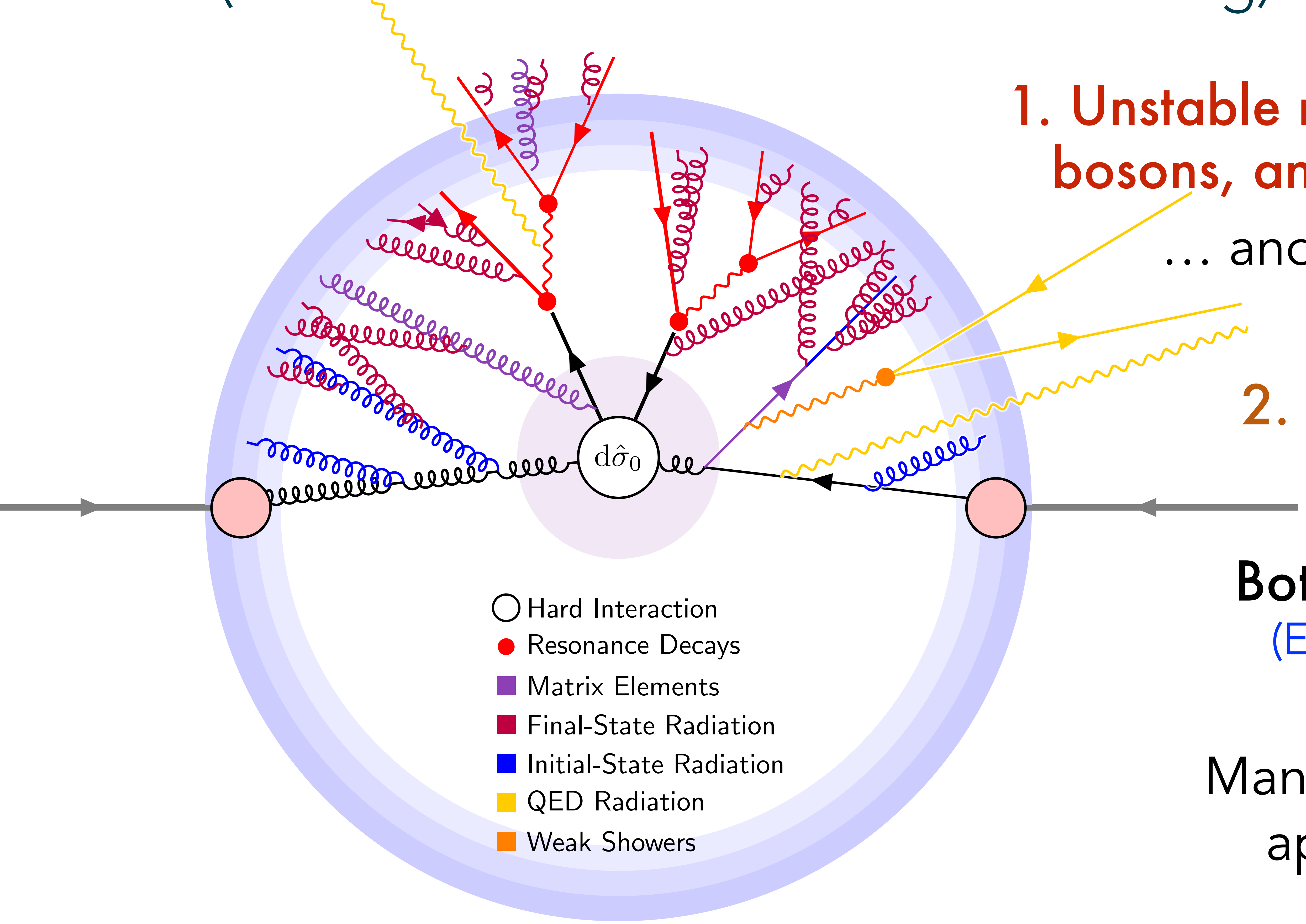
Active research field; alternative approaches also hotly pursued

E.g.: PanScales, Dire…

(Resonance Decays and Weak Showers)

I will add a few further details without much comment

(Otherwise this talk would be too long)



1. Unstable resonances (top quarks, Z/W bosons, and Higgs bosons) will decay

… and their decay products will shower

2. Weak SU(2) bremsstrahlung

Both are topics of active research

(E.g.: R. Verheyen & PS, 2108.10786)

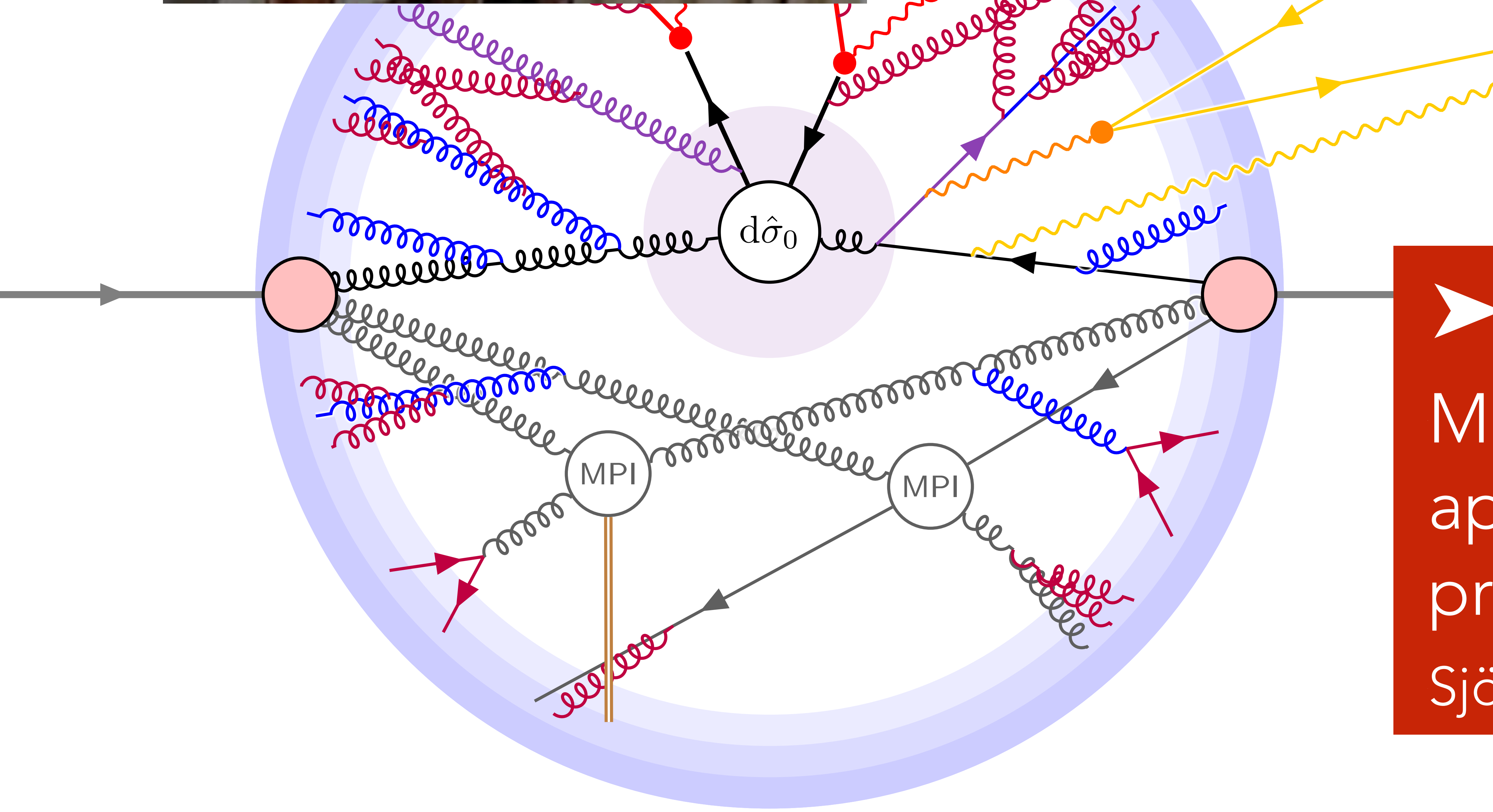
Many interesting questions and applications (but no major revolutions expected).

Such Stuff as Beams are Made Of

Before we talk about confinement

Recall that the protons were composite

Who said only a *single* pair of partons collided?



As they pass through each other, the two protons present a **beam** of partons to each other

➤ Multi-Parton Interactions (MPI)

MCMC algorithms with iterated application of factorised scattering probabilities. Around since 80s.

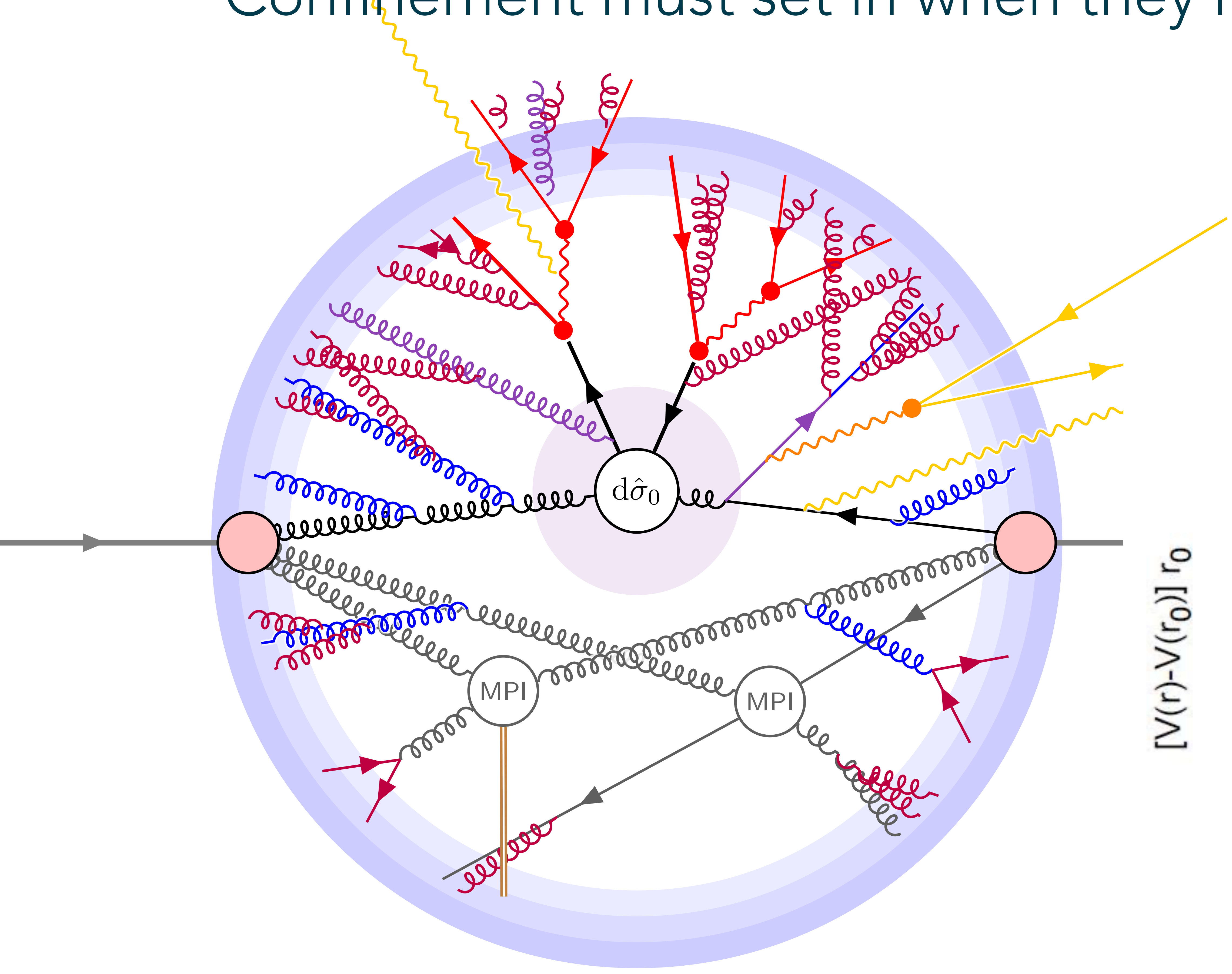
Sjöstrand ('85) + a few more recent

Crucial to describe event structure at hadron colliders

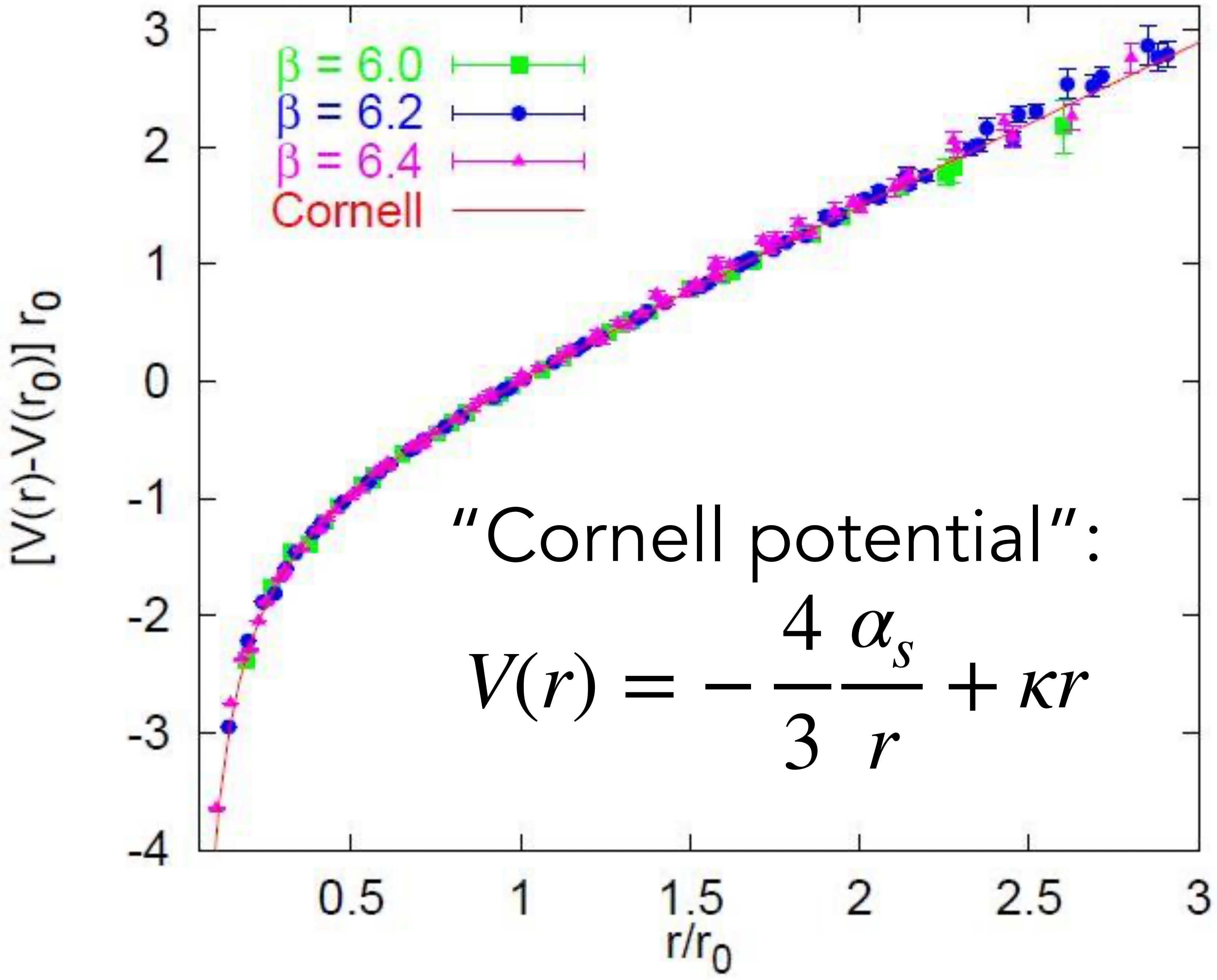
Confinement

Event structure still in terms of (colour-charged) quarks & gluons

Confinement must set in when they reach O(1fm) relative distances.



Between a single quark-antiquark pair, we know the long-distance behaviour is a linear potential



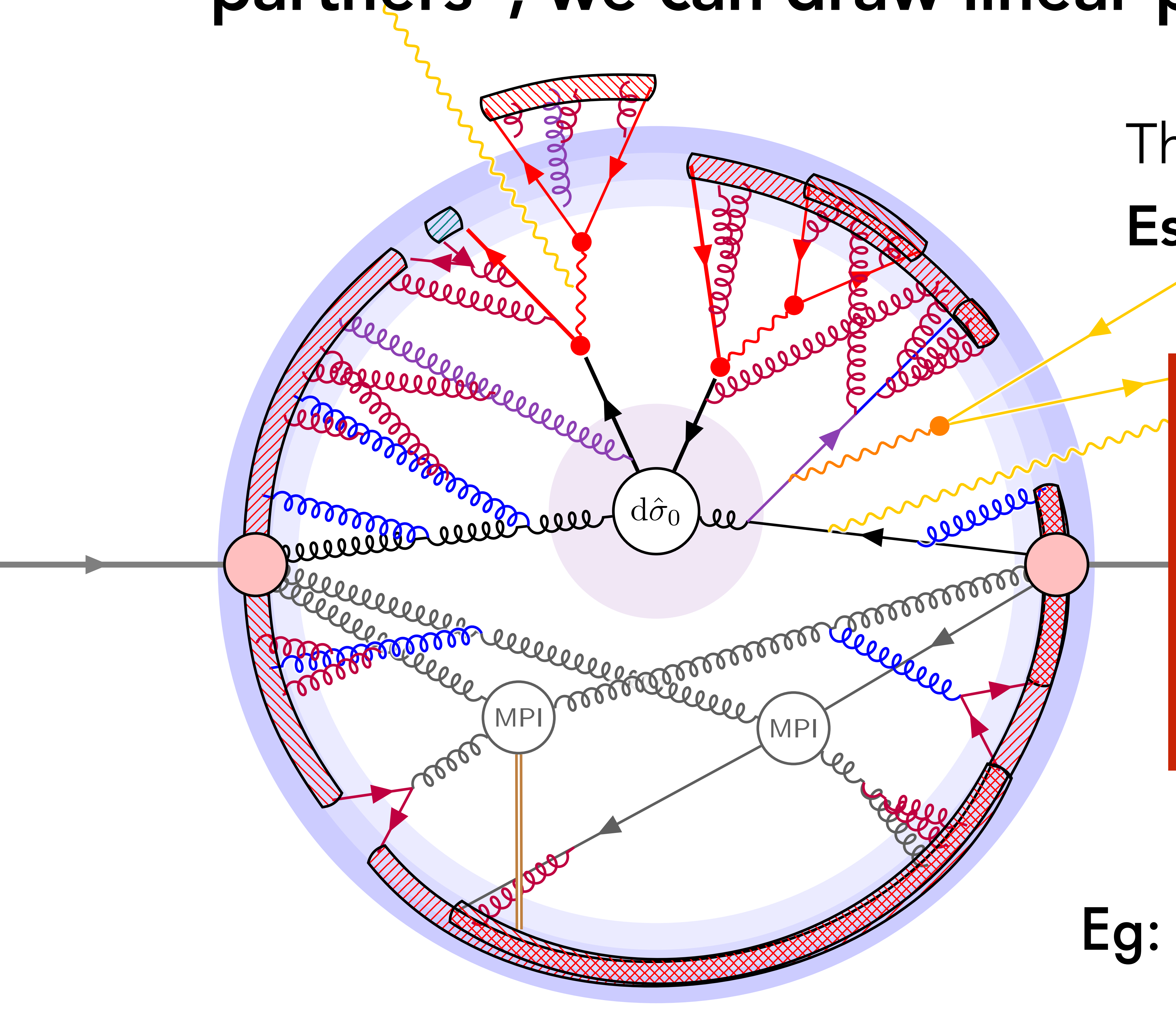
"Cornell potential":

$$V(r) = -\frac{4}{3}\frac{\alpha_s}{r} + \kappa r$$

$r_0 = 0.5\text{fm}$ $\kappa \sim 0.9\,\text{GeV/fm}$

It's all about connections

So *if* we know which partons are each others' "colour partners", we can draw linear potentials between them:



There are, however, **ambiguities**

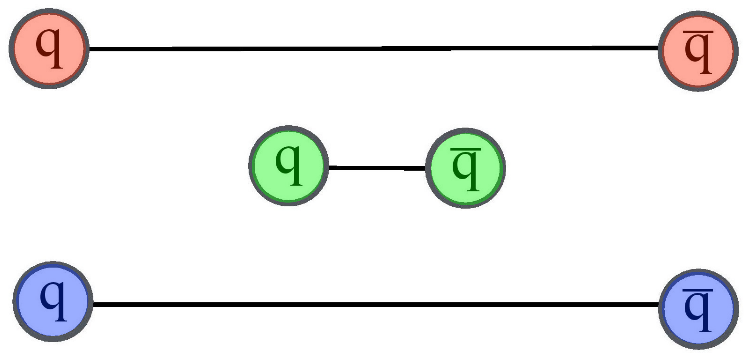
Especially in *complex* events with *many* MPI

➤ Colour Reconnections (CR)

Represented by inner blue shaded band. Generally thought to act to **minimise** the **total** linear potential.

Sjöstrand & v. Zijl ('85), Christiansen & PS ('15) + …

Eg:



Vs

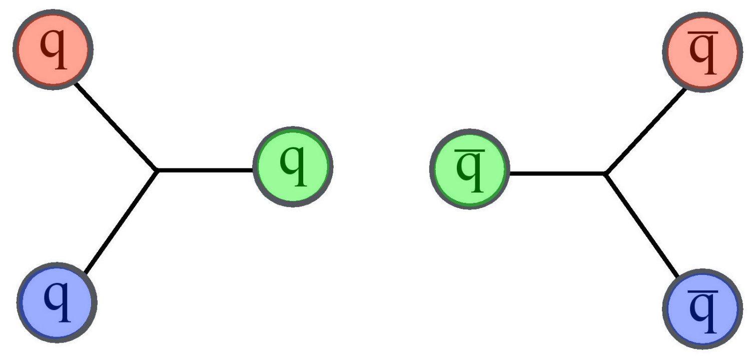
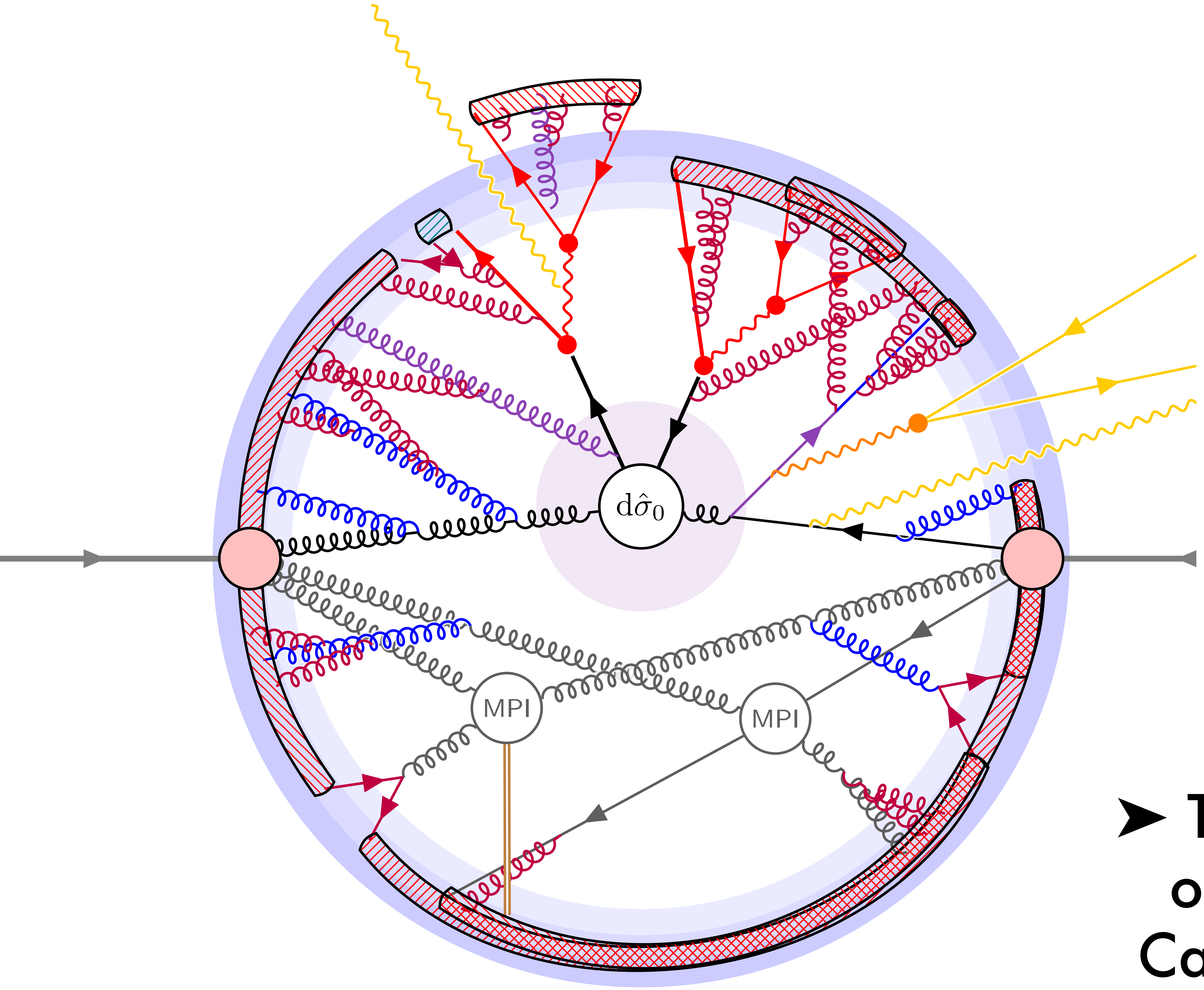


Illustration by J. Altmann

Christiansen & PS ('15)

Time to call a string a string

What physical system has a linear potential? A string.



This is the basis for the **Lund String Fragmentation Model**

Andersson, Gustafson, Pettersson, Sjöstrand, … ('78 - '83)

A *comparatively* simple 1+1 dimensional model of massless relativistic strings, with tension $\kappa \sim 1$ GeV/fm

➤ The signature feature of the Pythia Monte Carlo event generator

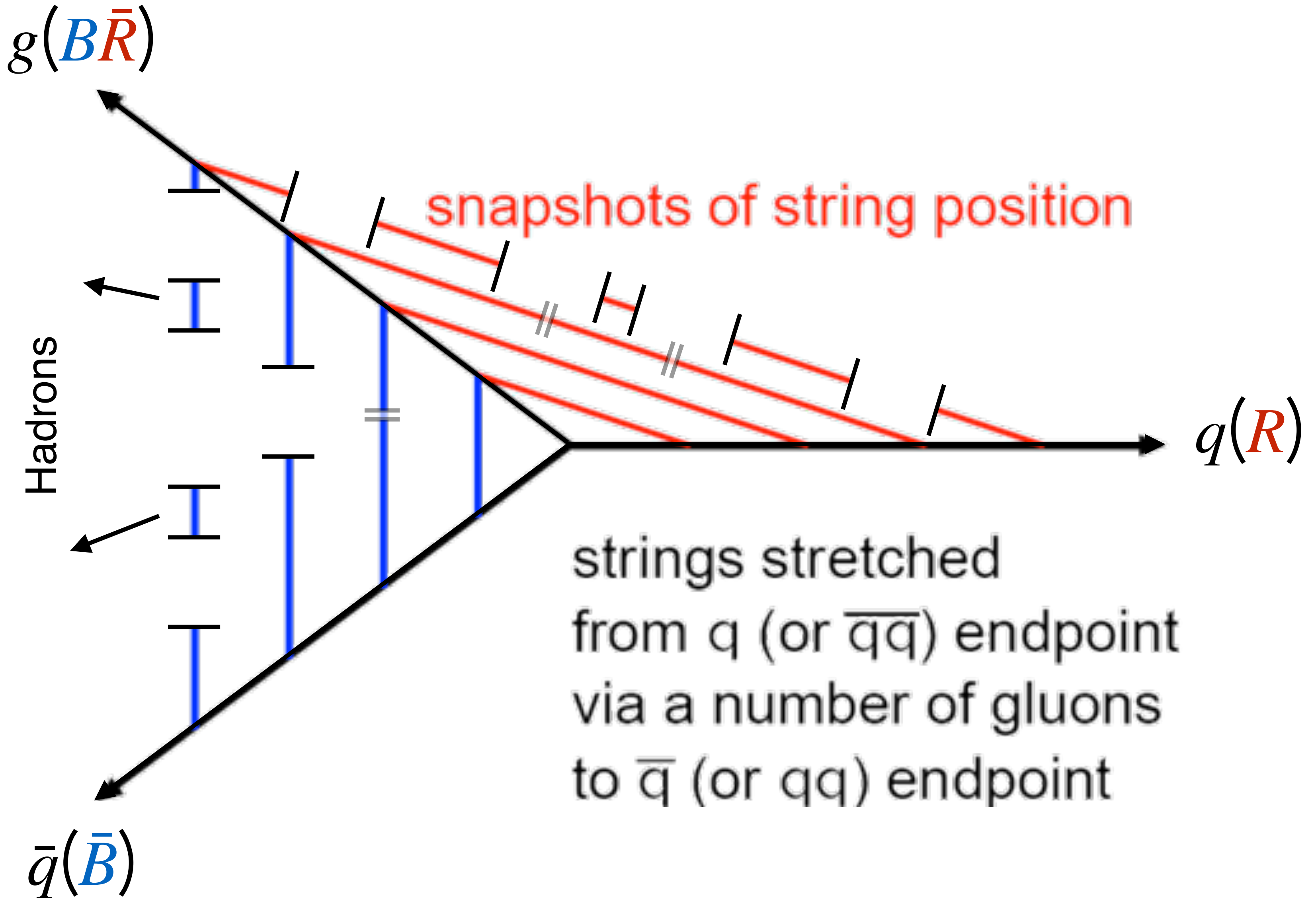
A New Set of Degrees of Freedom

The string model provides a mapping:

Quarks ➤ String endpoints

Gluons ➤ Kinks on strings

Further evolution then governed by string world sheet (area law)



strings stretched
from q (or $\overline{qq}$) endpoint
via a number of gluons
to $\overline{q}$ (or qq) endpoint

+ string breaks by tunnelling

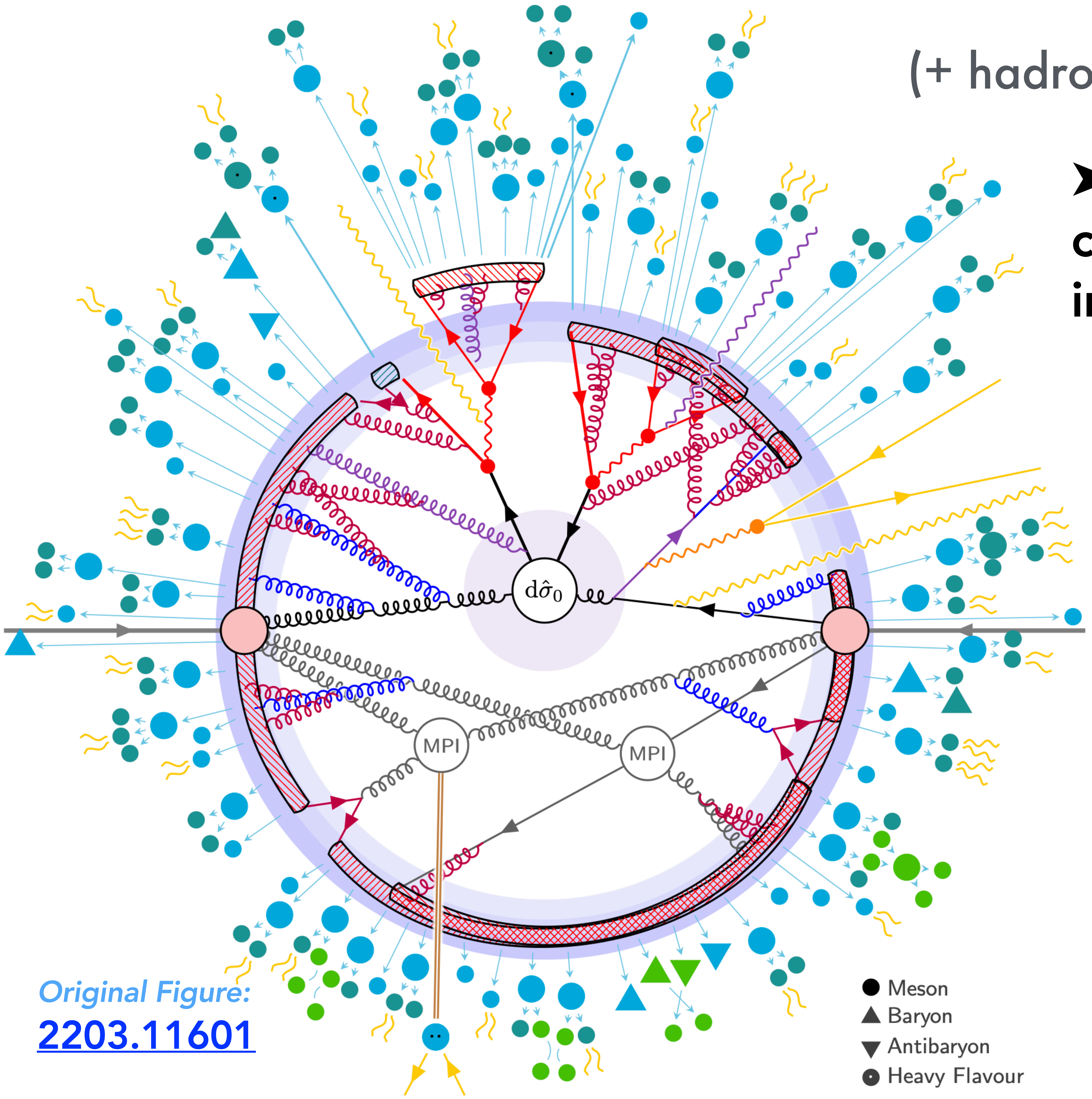
By analogy with "Schwinger mechanism" in QED (electron-positron pair production in strong electric field)

Predictive for phase-space distribution of hadrons (but not for their spin/flavour composition ➤ **Bierlich, Chakraborty, Gustafson, Lönnblad '22**)

Hyperfine splitting effects in string hadronization

➤ **Jets of Hadrons!**

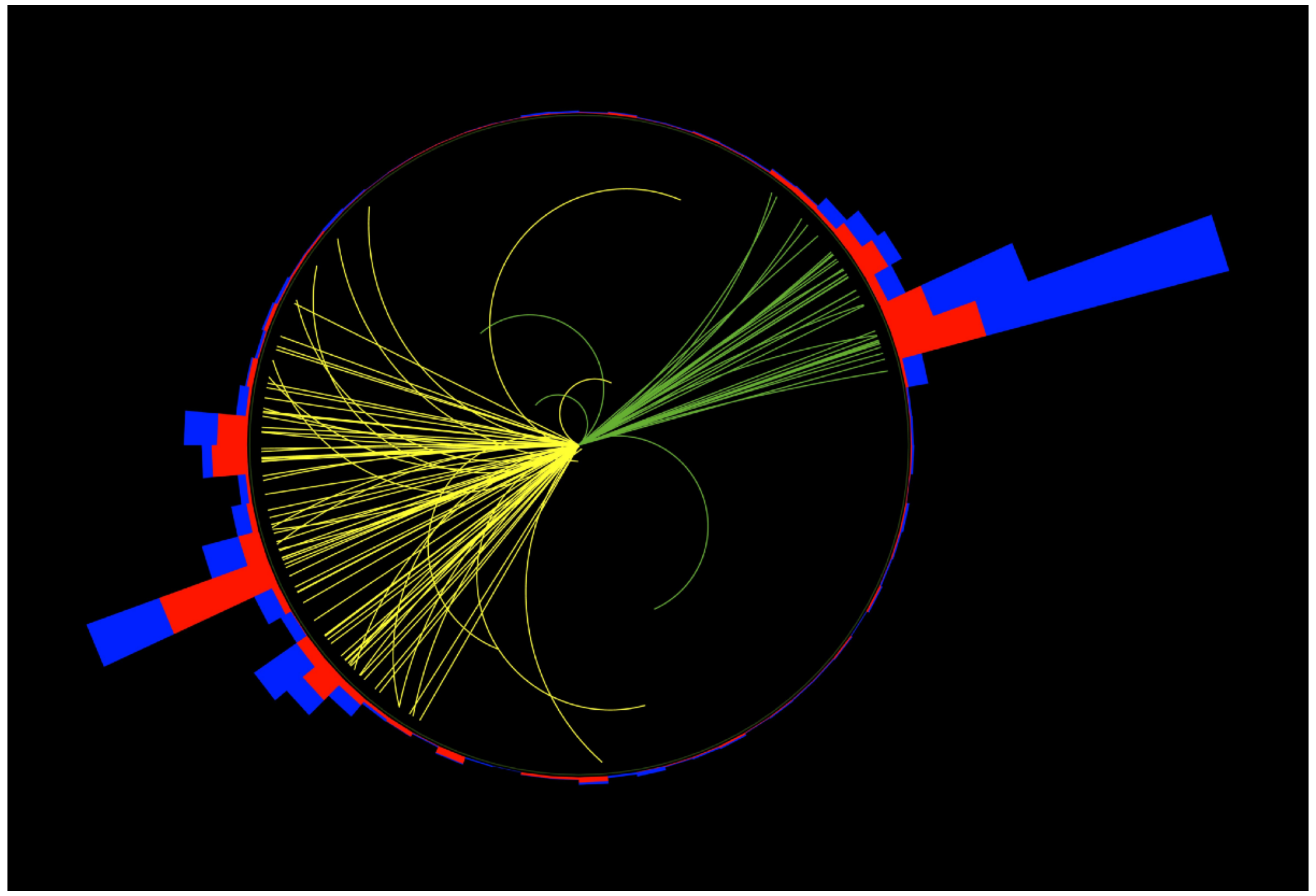
Hadronisation



Original Figure: 2203.11601

(+ hadron decays; added without comment)

- **We finally have a model that can be compared to experiments in full detail …**



I can only show you a few hand-picked measurements I find particularly interesting

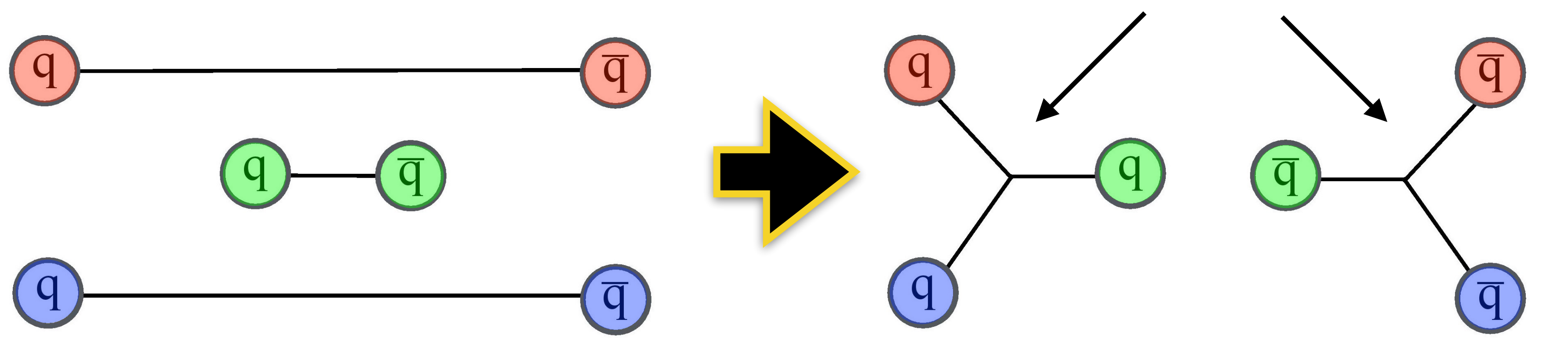
Unique feature of SU(3): Y-Shaped 3-String "Junctions" ➤ Baryons

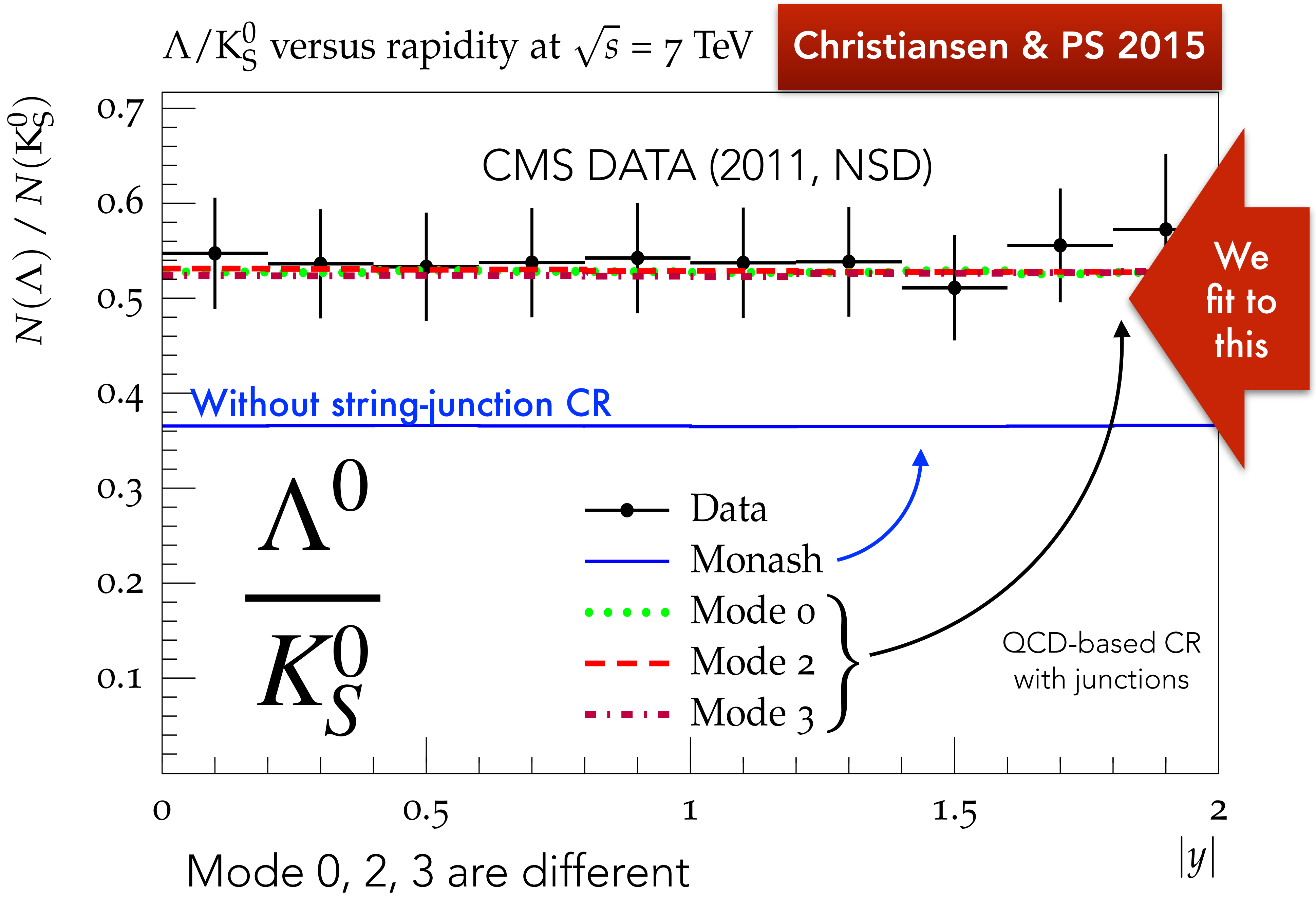
Baryon Number Violation & String Topologies: Sjöstrand & PS hep-ph/0212264

String Formation Beyond Leading Colour: Christiansen & PS 1505.01681

"Colour reconnection" modelling based on stochastic sampling of SU(3) group probabilities: allows for random (re)connections

For example:

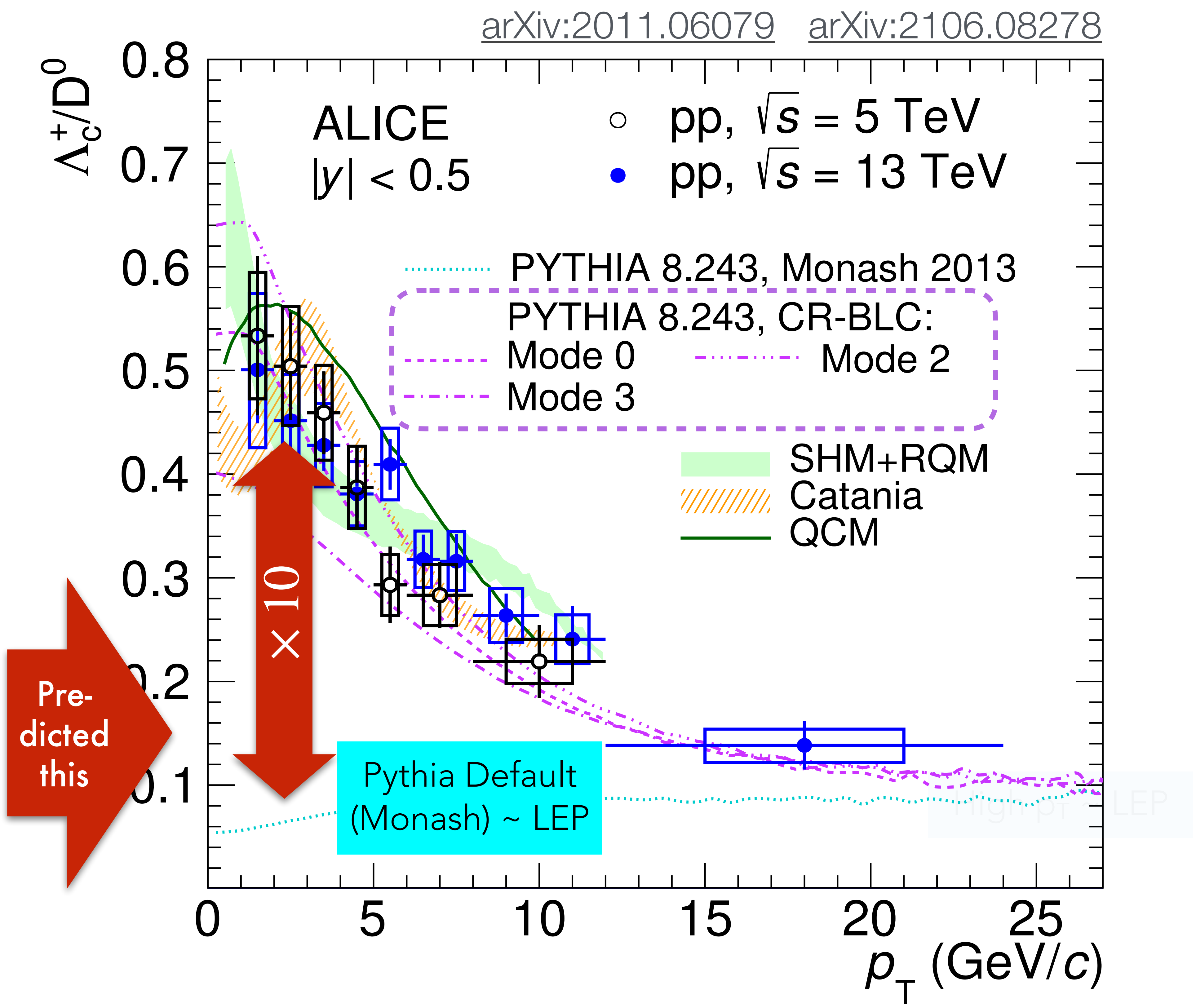

Extra baryon-antibaryon production


Christiansen & PS 2015

We fit to this

Mode 0, 2, 3 are different causality restrictions (0 = none)

ALICE 2021: also in charm


arXiv:2011.06079 arXiv:2106.08278

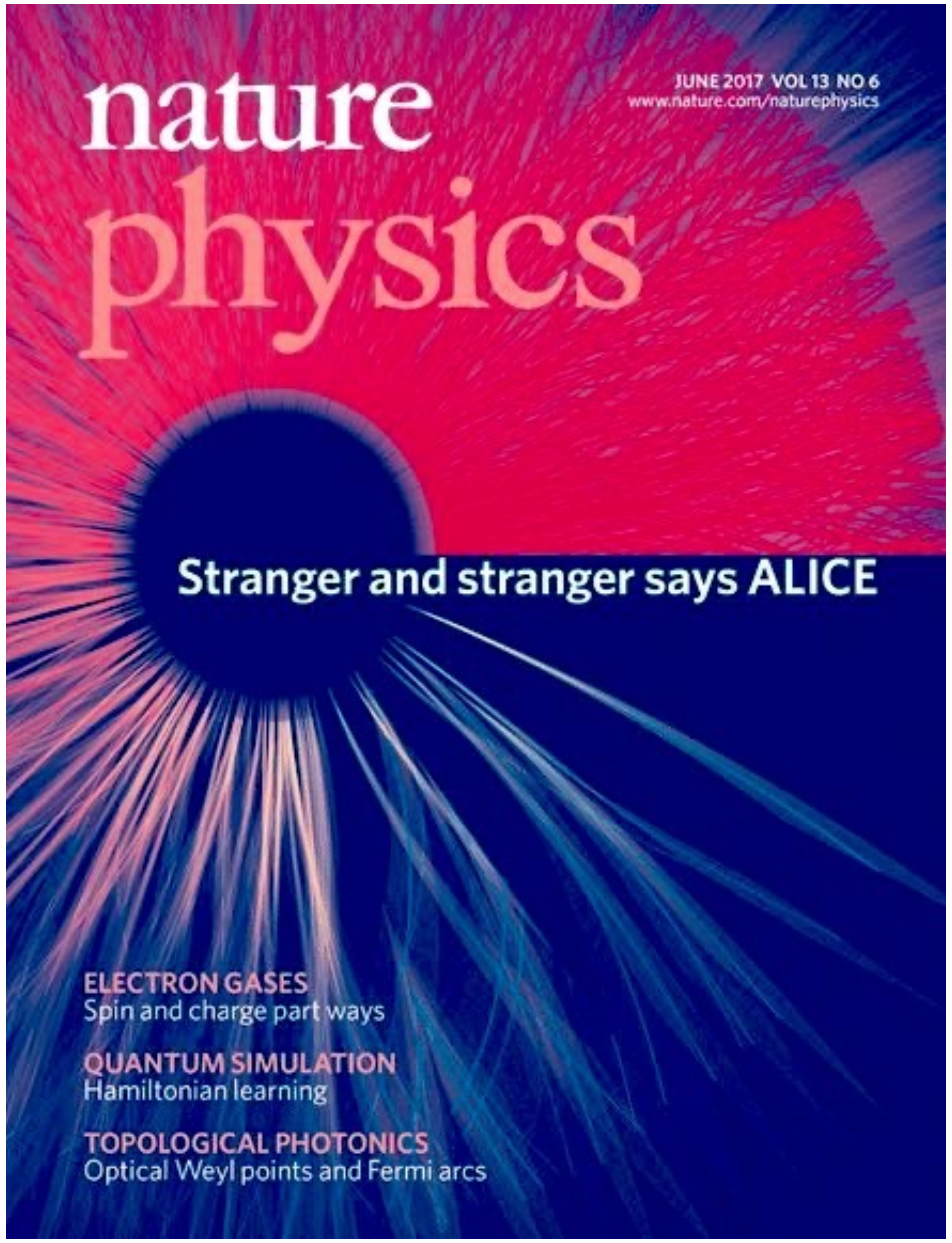
Predicted this

Pythia Default (Monash) ~ LEP

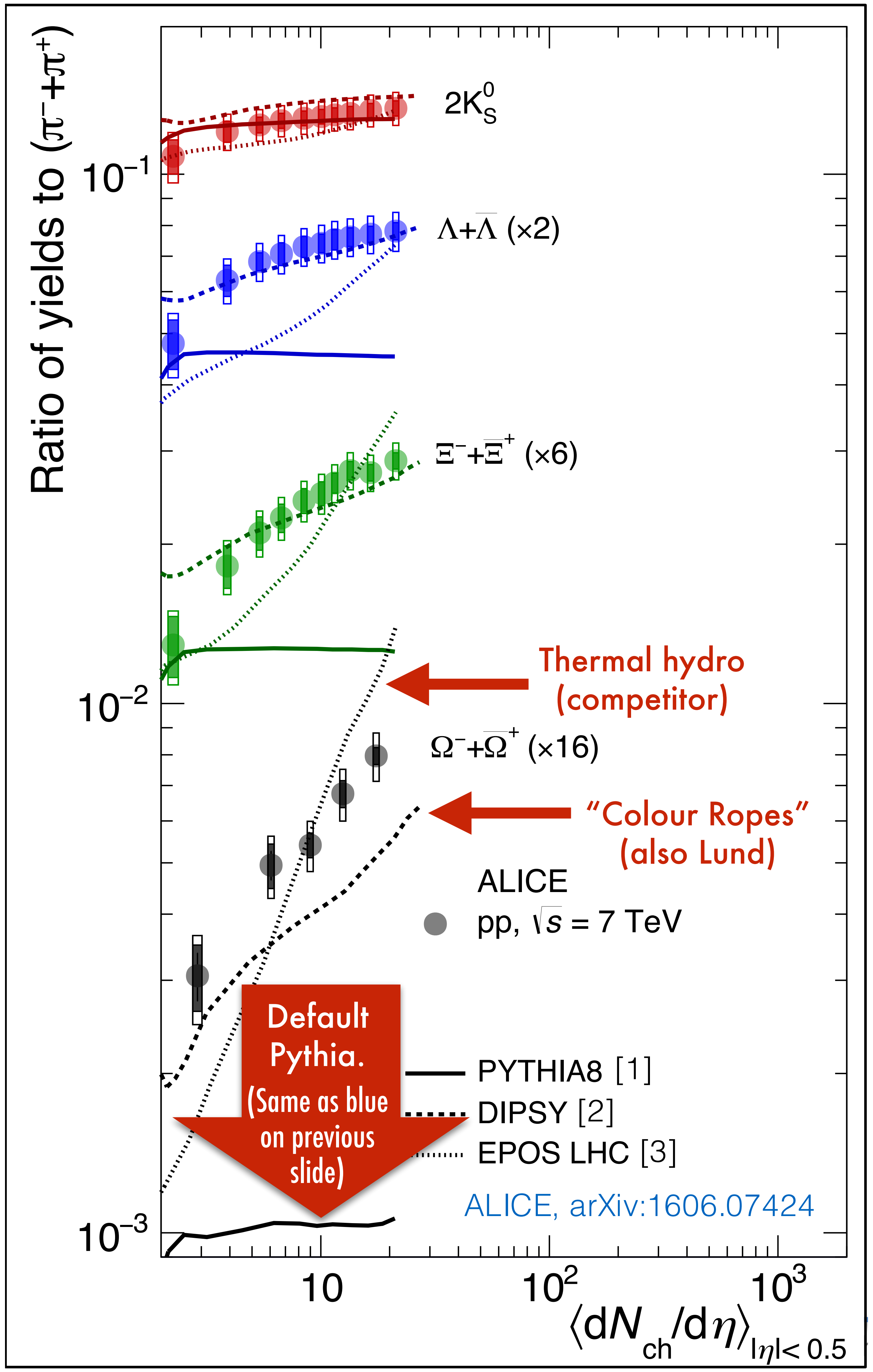
What a **strange** *world we live in, said Alice*

Landmark measurement by ALICE ('17)

Ratios of strange hadrons to pions



June 2017



Other signs of "collectivity"

"CMS ridge" (CMS '10):

Long-distance correlations between particles at same azimuthal angle, in "busy" events — not predicted!

Interpreted as sign of a **"collective flow"** along common (transverse) axis

By now many follow-up measurements confirming same features

Taken together: string junctions, strangeness enhancement, flow

I think indicates that we are seeing **QCD string interactions**

Strings have physical properties of vortex lines. Strings with same flux orientation repel each other, like two co-rotating tornadoes.

Lund group has implemented a model of "string shoving".

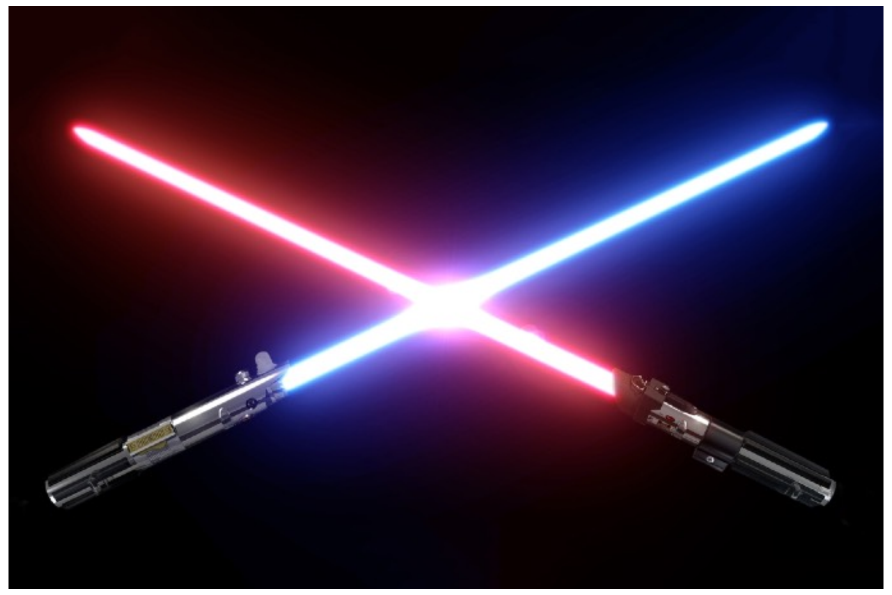
The interaction energy also increases the string tension ➤ more **strangeness**

These new measurements, and our growing understanding of them, are ushering in a new era of exploration of emergent non-perturbative phenomena

Apologies: Many things not mentioned …

Photon-induced processes (photoproduction) ⟶

Photons can appear pointlike, or with partonic substructure ~ hadrons



Flavour Physics, Neutrino Physics, Cosmic Rays, …

Hadrons, Heavy ions, ropes, shoving, diffraction, …

Heavy Ions, ropes, shoving ➤ Much work in Lund & Jyväskyla (+ Monash)

Hadronic Rescattering ➤ Sjöstrand, Utheim 2005.05658

Bose-Einstein & Fermi-Dirac Correlations (➤ N-particle correlations, Femtoscopy)

New Physics …

Dark Matter and Dark Sectors / Hidden Valleys ➤ Desai, Sjöstrand

➤ **Brand new Comprehensive Guide: 2203.11601**

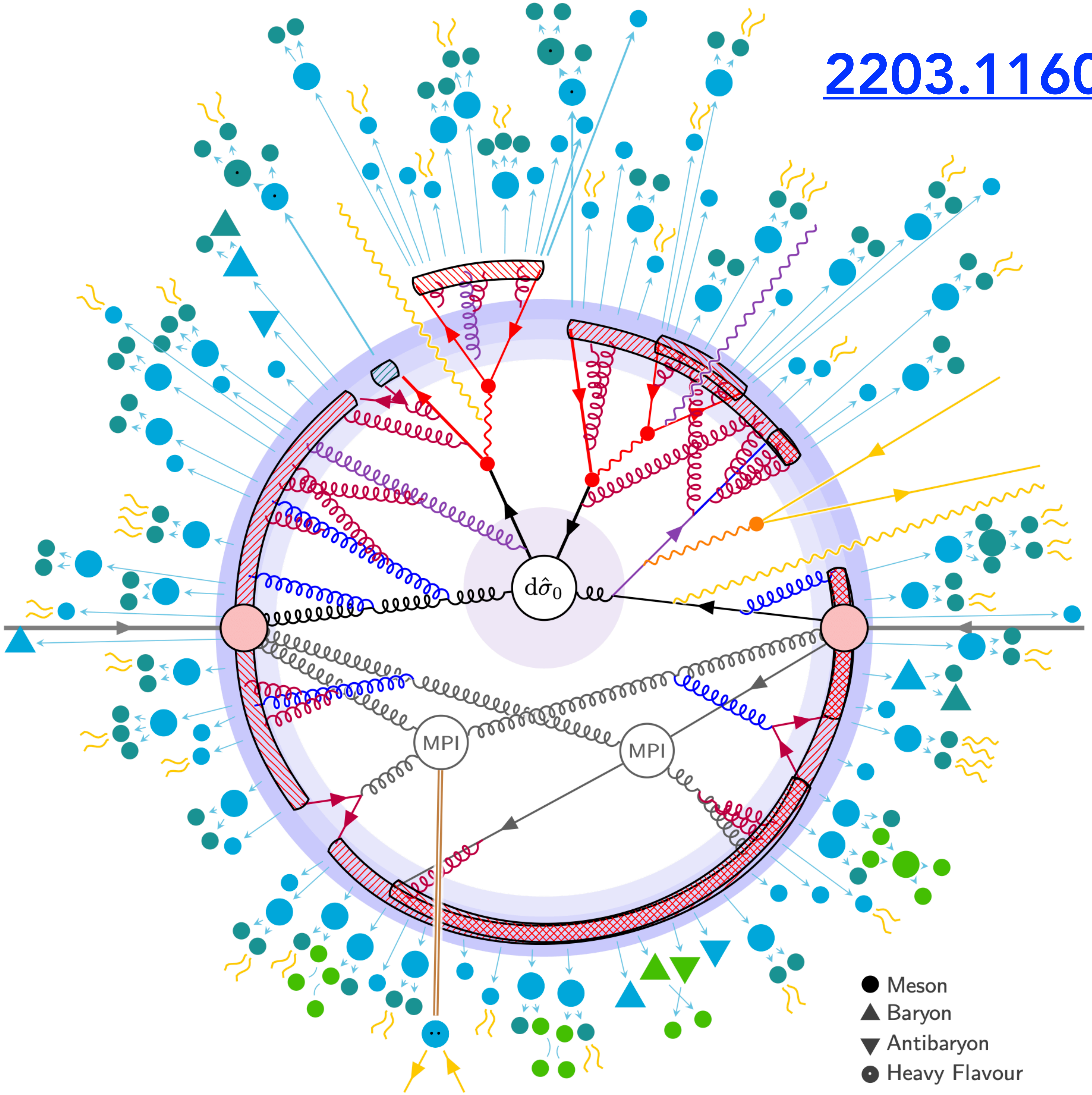
315 pages: "A comprehensive guide to the physics and usage of Pythia 8.3"

Thank you!

Anatomy of an LHC Collision

- ○ 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)*



2203.11601

Re-examations of String Basics? Time dependence?

Cornell potential

Potential V(r) between **static** (lattice) and/or **steady-state** (hadron spectroscopy) colour-anticolour charges:

$$V(r) = -\frac{a}{r} + \kappa r$$

Coulomb part

String part
Dominates for $r \gtrsim 0.2\,\mathrm{fm}$

Lund string model built on the asymptotic large-*r* linear behaviour

But intrinsically only a statement about the late-time / long-distance / steady-state situation. Deviations at early times?

Coulomb effects in the grey area between shower and hadronization?

Low-*r* slope > κ favours "early" production of quark-antiquark pairs?

+ Pre-steady-state thermal effects from a (rapidly) **expanding string?**

Berges, Floerchinger, and Venugopalan JHEP 04(2018)145)

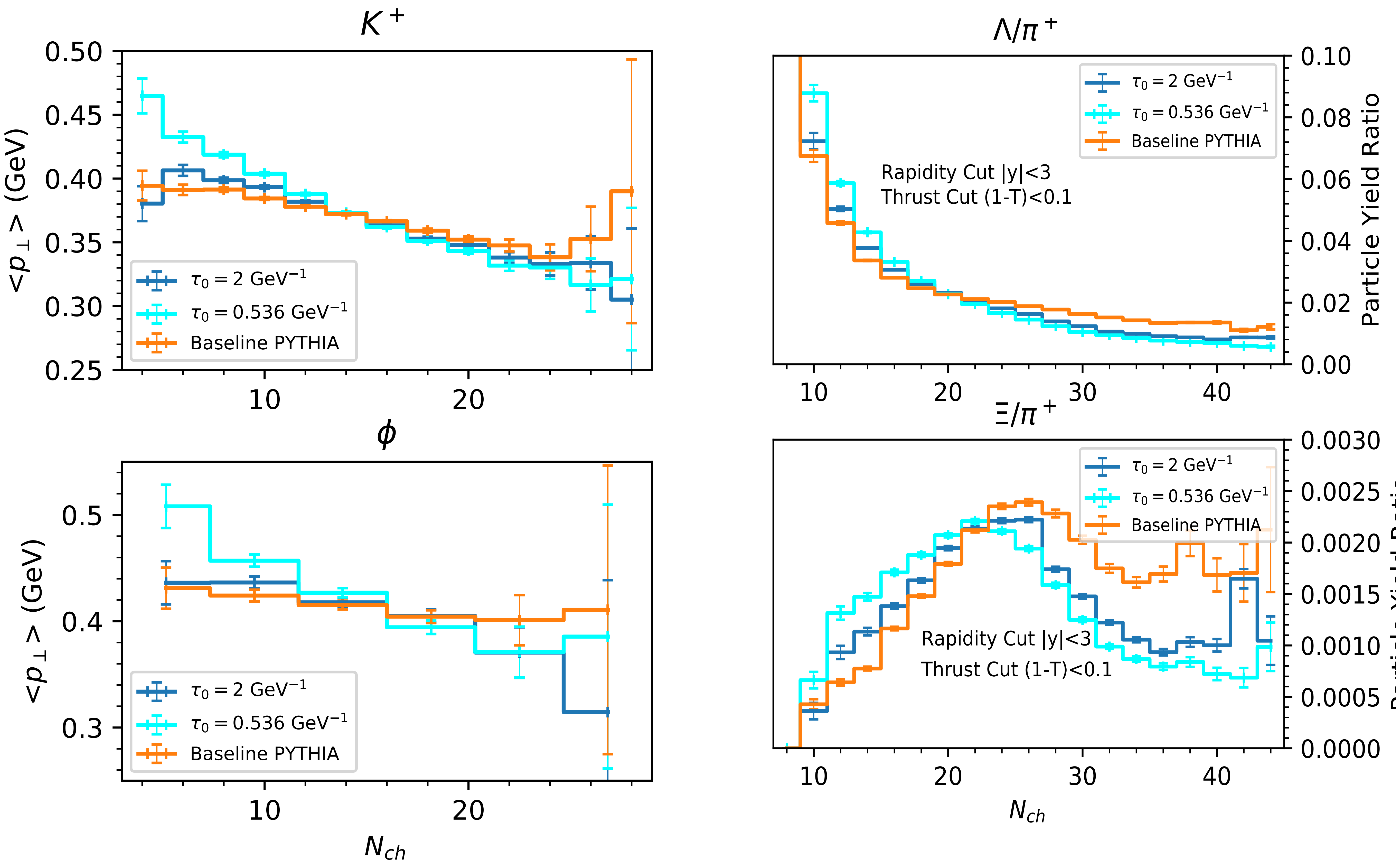
Toy Model with Time-Dependent String Tension

N. Hunt-Smith & PS arxiv:2005.06219

Model constrained to have same average tension as Pythia's default "Monash Tune"

➤ same average N_{ch} etc ➤ main LEP constraints basically unchanged.

But expect different fluctuations / correlations, e.g. with multiplicity N_{ch}.



- ➤ Want to study (suppressed) tails with very low and very high N_{ch}.
- ➤ These plots are for LEP-like statistics.
- ➤ Would be crystal clear at CEPC/FCC-ee

Colour Connections: Between which partons do confining potentials form?

High-energy collisions with QCD bremsstrahlung + multi-parton interactions
➤ final states with very many coloured partons

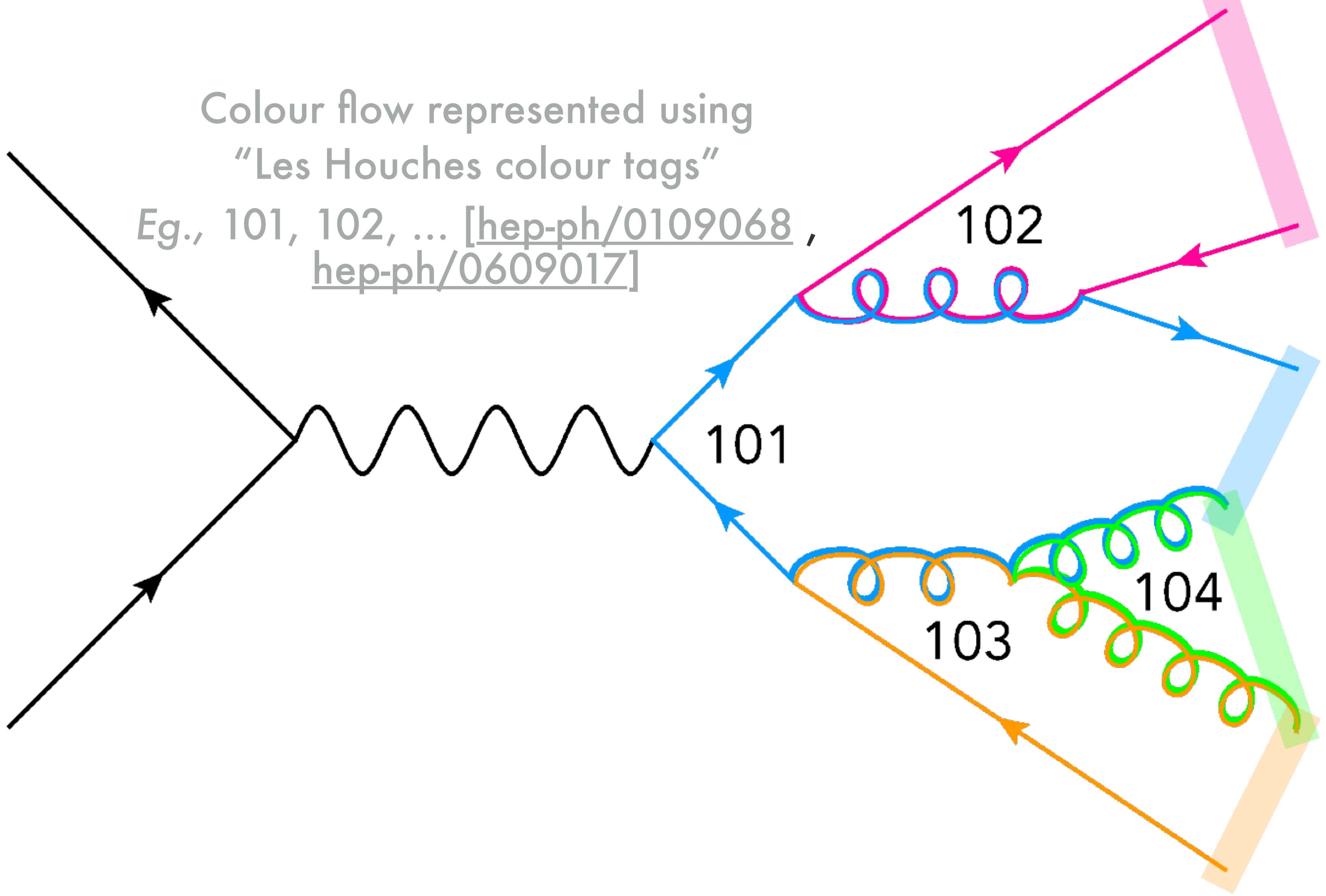
Who gets confined with whom?

Starting point for MC generators = Leading Colour limit $N_C \to \infty$

$\implies$ Probability for any given colour charge to accidentally be same as any other $\to 0$.

$\implies$ Each colour appears only once & is matched by a unique anticolour.

Example (from upcoming big Pythia 8.3 manual):
$e^+e^- \to Z^0 \to q\bar{q}$ + parton shower



Colour flow represented using
"Les Houches colour tags"
Eg., 101, 102, ... [hep-ph/0109068 , hep-ph/0609017]

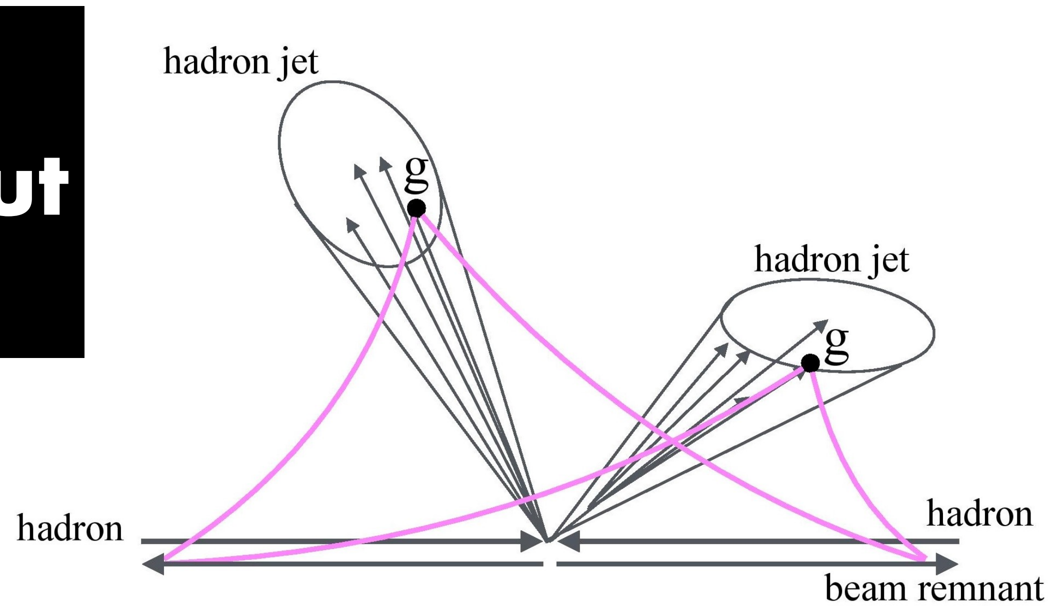
Naively, corrections suppressed by $1/N_C^2 \sim 10\,\%$

But in pp collisions, multi-parton interactions $\implies$ many such systems

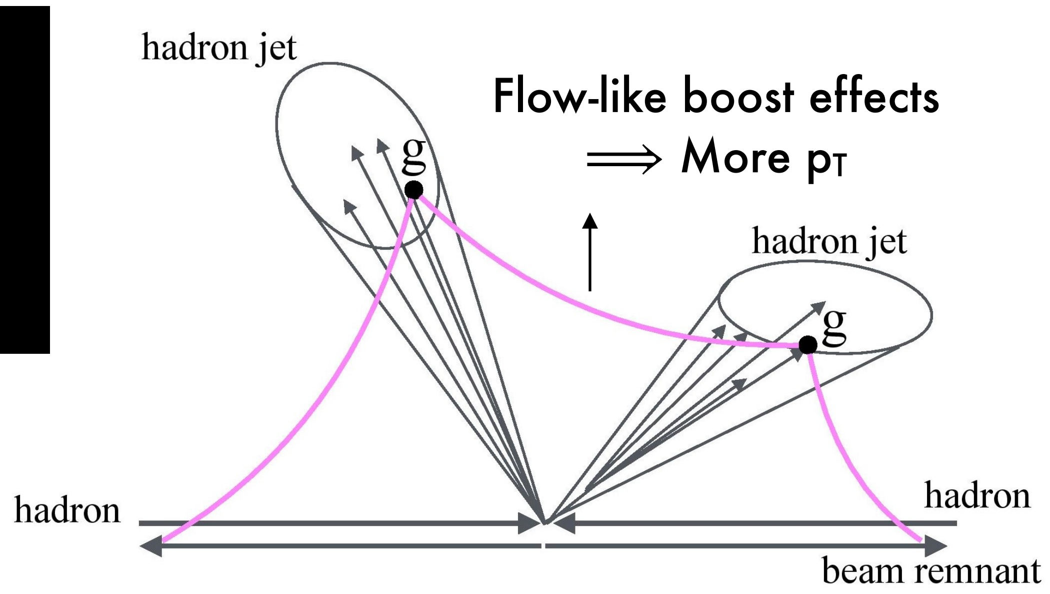
Each has probability ~ 10% + significant overlaps in phase space $\implies$ CR more likely than not

Colour **Reconnections** Original Goal: describe observables like $\langle p_T \rangle (n_{ch})$

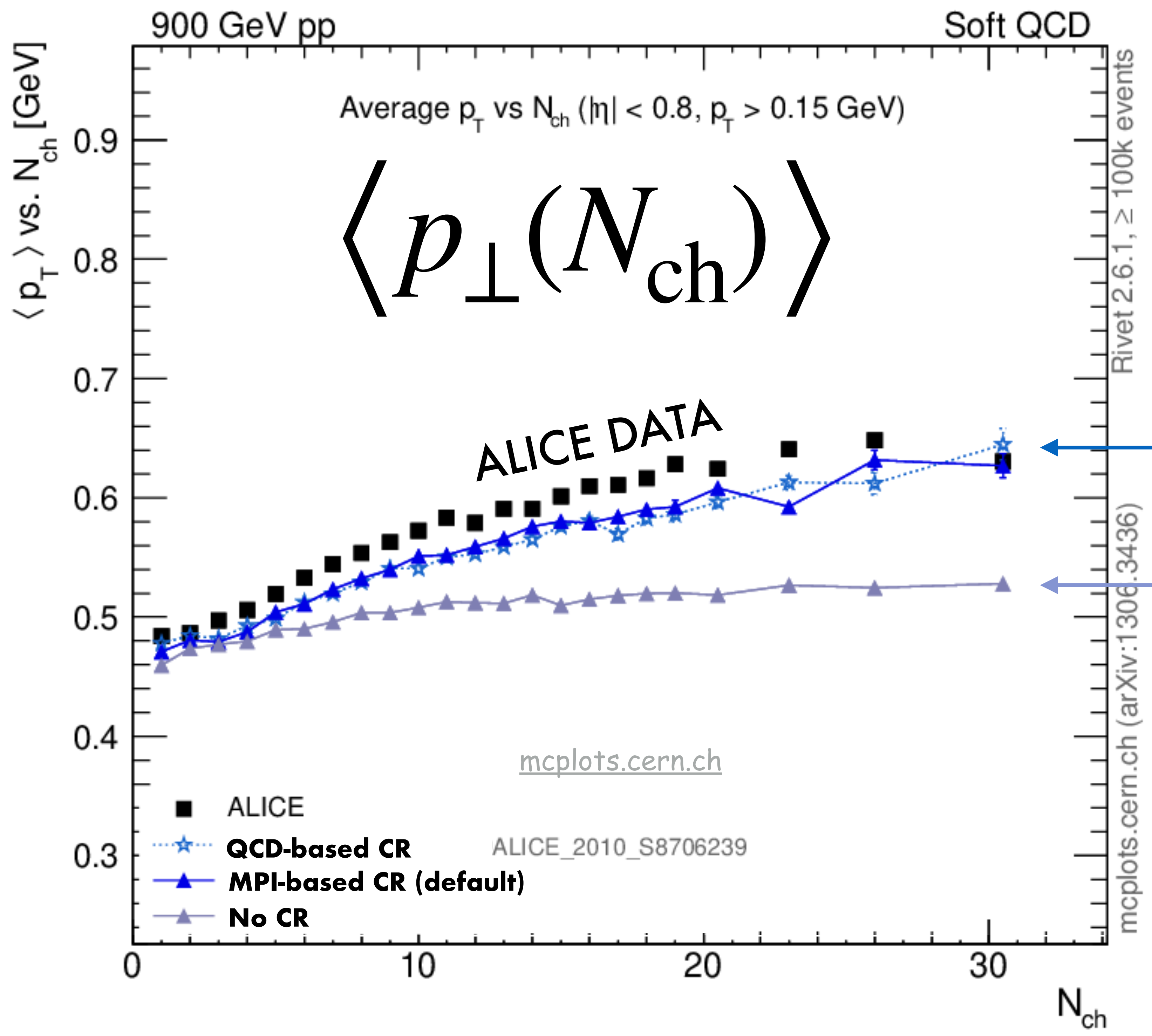
MPI
without
CR:



MPI
with
CR:



Note: for more on flow-like effects from CR, see also, e.g., Ortiz Velasquez et al. arXiv:1303.6326



Both **MPI-based** (default) and **QCD-based** CR [1505.01681] reproduce the rising trend of $\langle pT \rangle (N_{ch})$

No CR $\Longrightarrow$ $\langle p_T \rangle$ approximately the same for all N_{ch} (Many MPI just produce more hadrons, but with ~ same

(Just one example here, that I could easily obtain from mcplots.cern.ch; with minor differences all other CM energies and fiducial cuts show same trend)

QCD-based CR Model: Rules of the Game

Christiansen & PS [1505.01681]

MPI + showers $\Longrightarrow$ partons with LC connections

Idea: stochastically allow ($1/N_C^2$) colour correlations, using SU(3) rules:

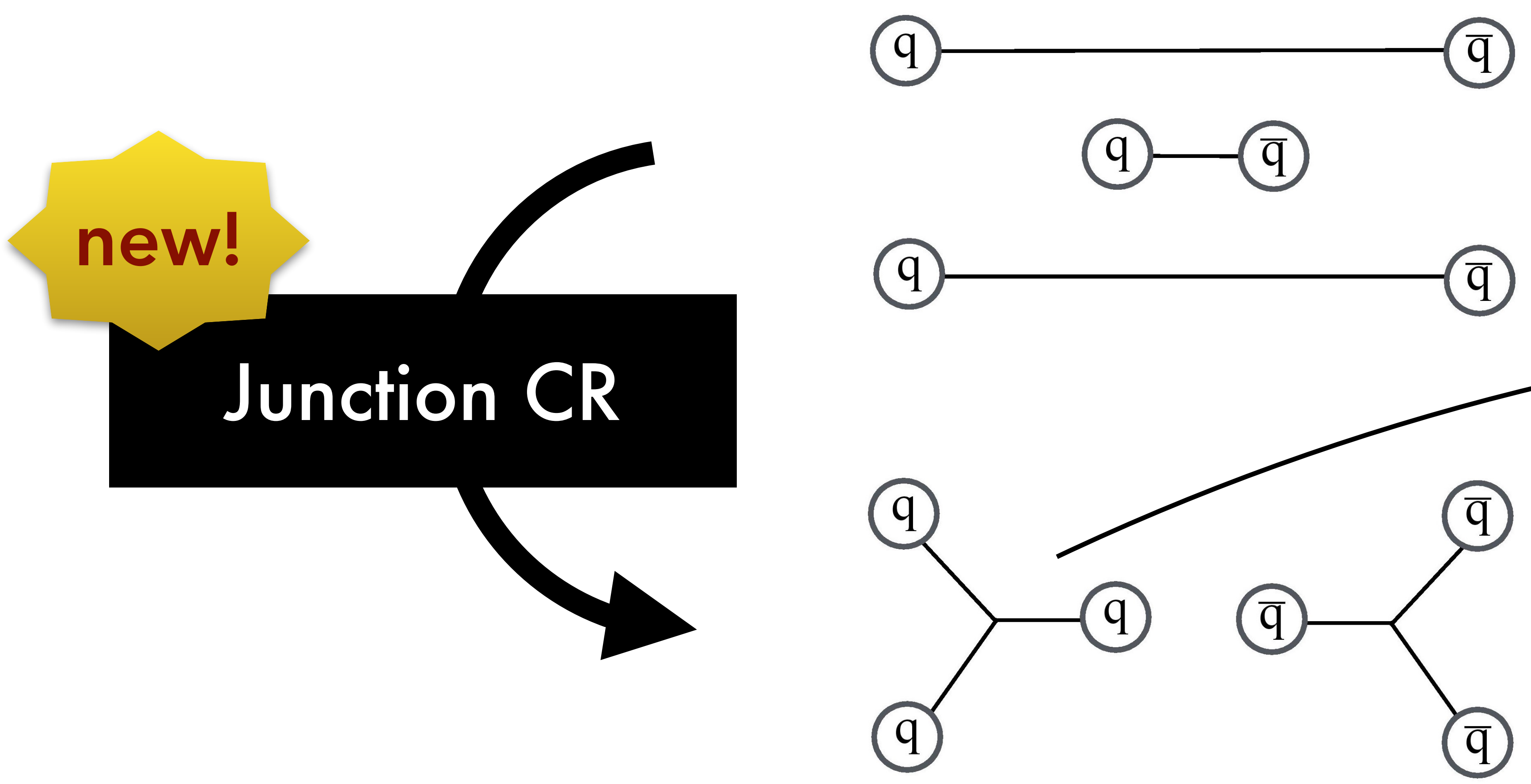
(1) $3 \otimes \bar{3} = 8 \oplus 1$ **for uncorrelated colour-anticolour pairs (allows "dipole CR")**

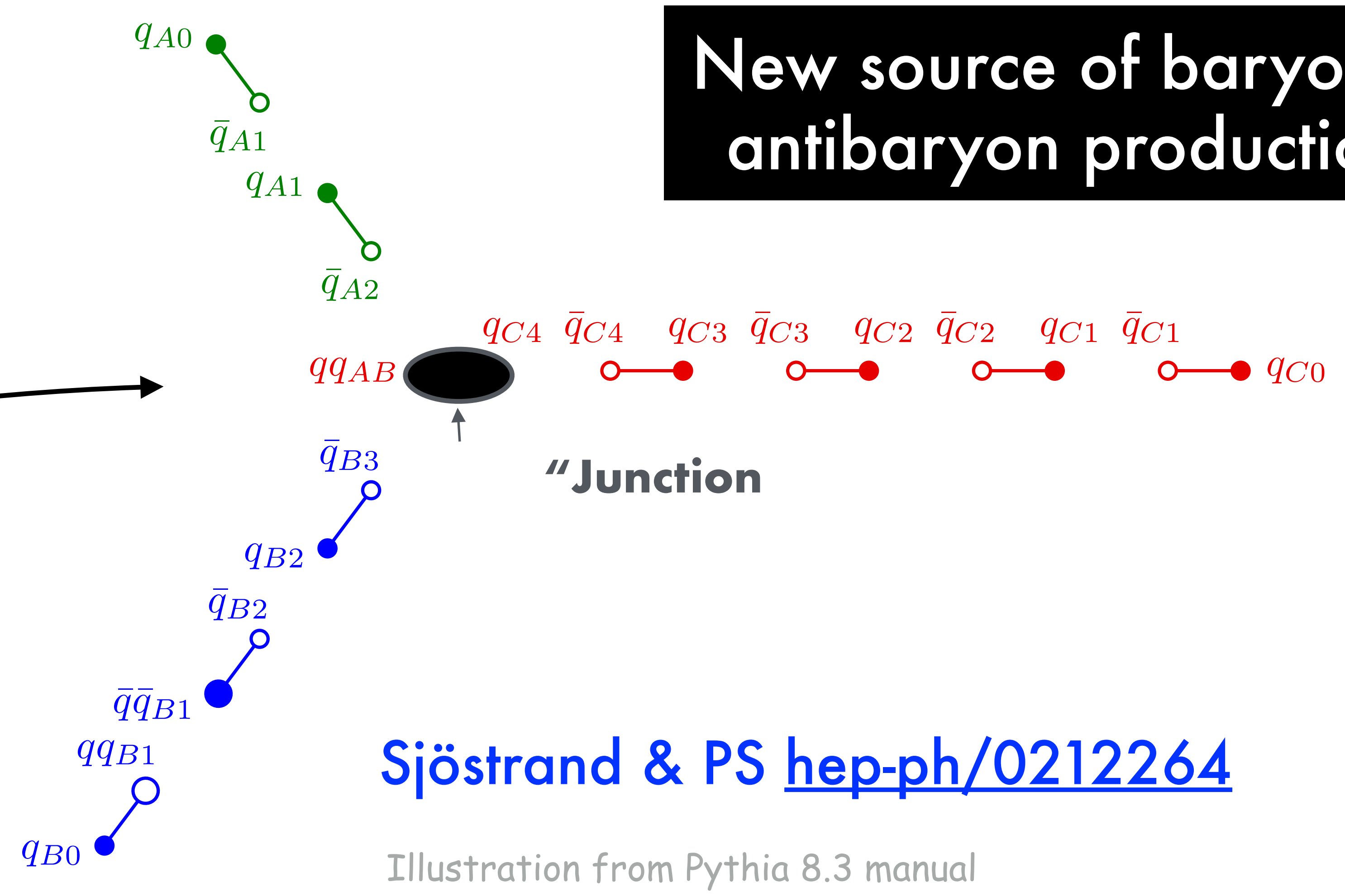
(2) $3 \otimes 3 = 6 \oplus \bar{3}$ **for uncorrelated colour-colour pairs (allows "junction CR")**

Then choose between which ones to realise confining potentials

Smallest measure of "invariant string length" $\propto$ number of hadrons

Illustrations by J. Altmann





New source of baryon + antibaryon production

Sjöstrand & PS [hep-ph/0212264]

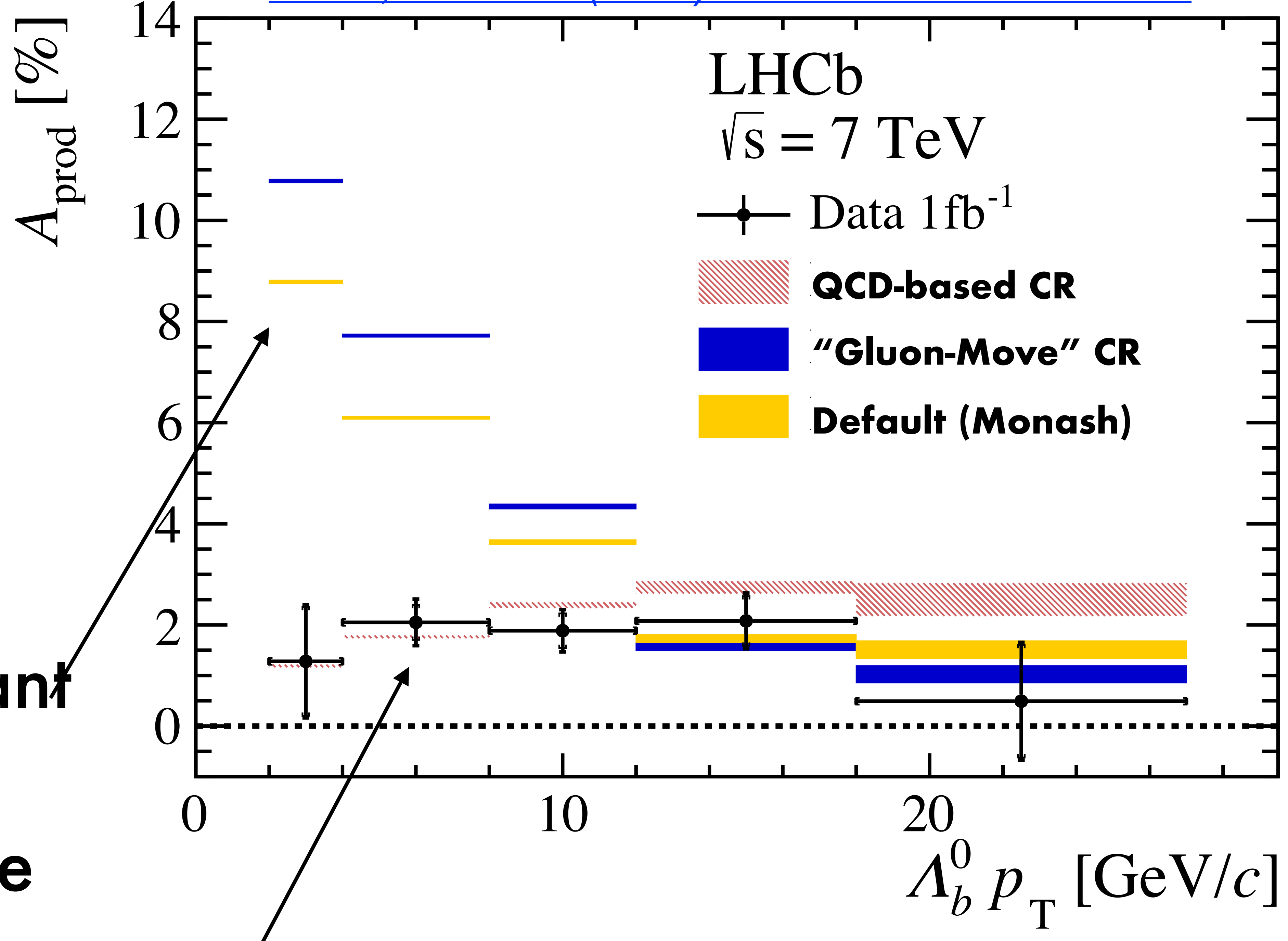
Illustration from Pythia 8.3 manual

LHCb: also in Bottom

Λ_b **asymmetry**

$$A = \frac{\sigma(\Lambda^0_\mathrm{b}) - \sigma(\overline{\Lambda}^0_\mathrm{b})}{\sigma(\Lambda^0_\mathrm{b}) + \sigma(\overline{\Lambda}^0_\mathrm{b})}$$

LHCb, JHEP 10 (2021) 060 • arXiv: 2107.09593



Without junction CR, an important source of low-p_T Λ_b production is when a b quark combines with the proton beam remnant.

Not possible for $\bar{\Lambda}_b$ (no $\bar{p}$ remnant at LHC)

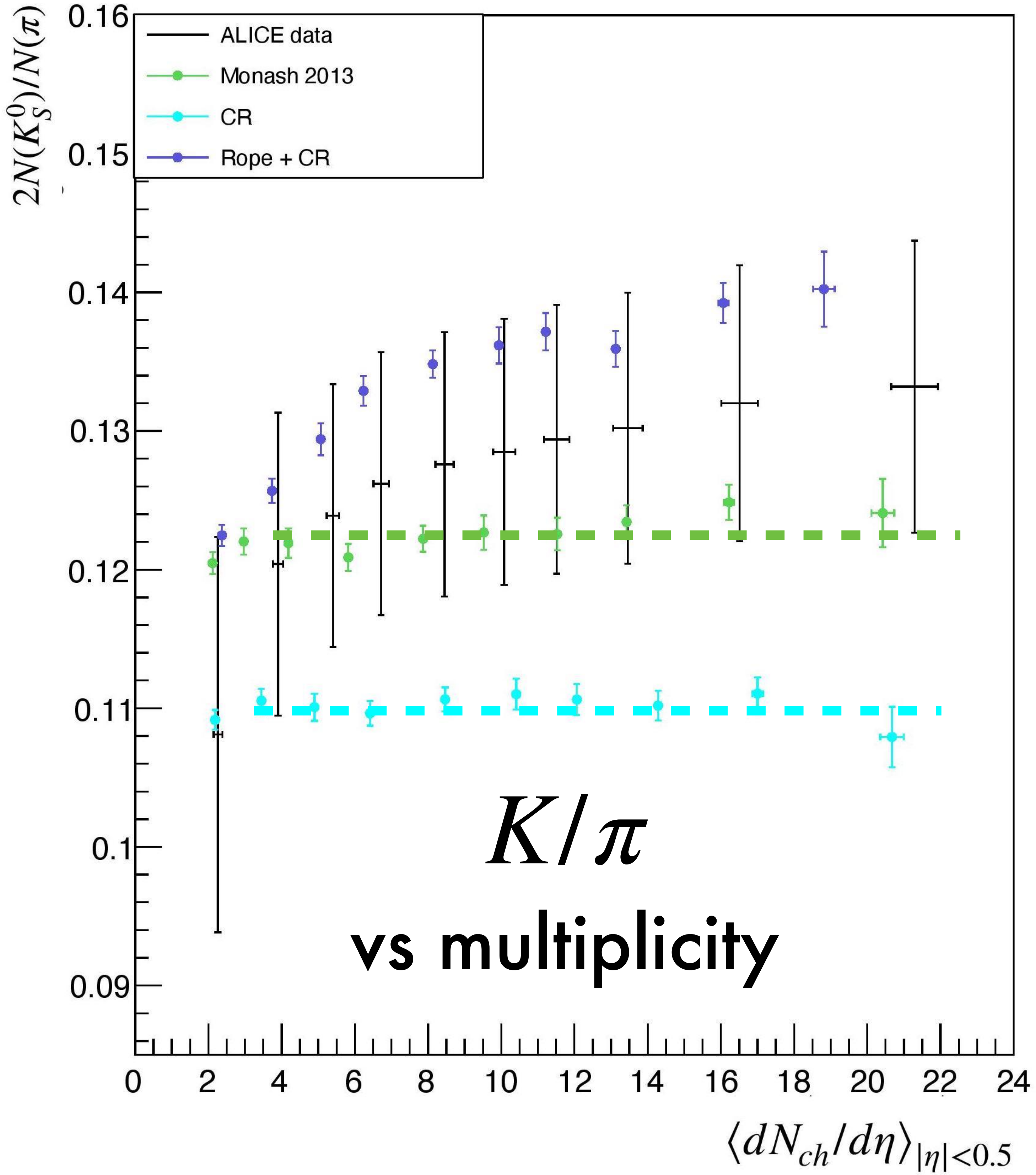
QCD CR adds large amount of low-p_T junction Λ_b and $\bar{\Lambda}_b$, in equal amounts. Dilutes asymmetry!

Strangeness

QCD-CR is not a mechanism for strangeness enhancement

When we look at "steps in strangeness", we see disagreements

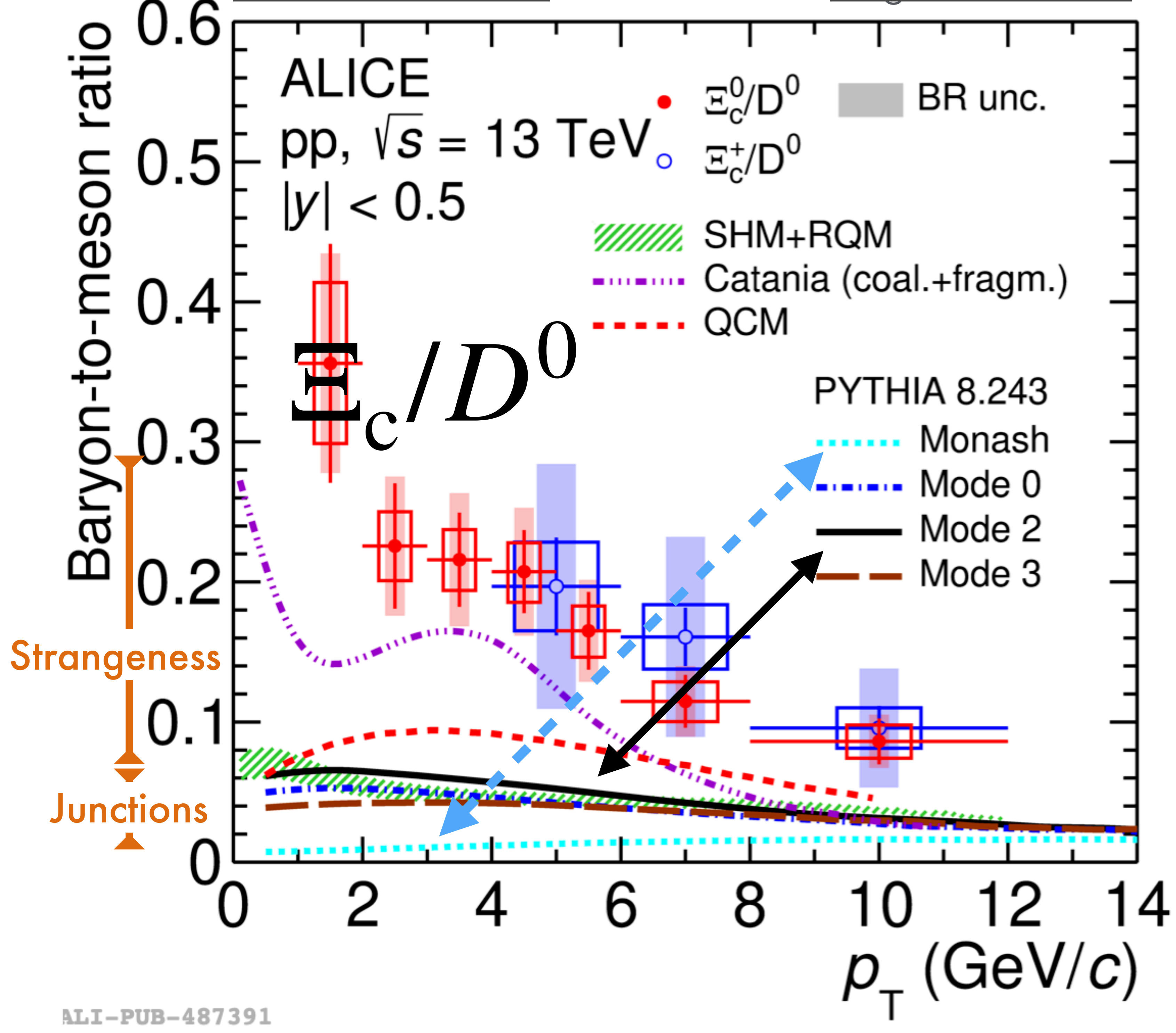
Ratio of yields of K_S to π vs charged multiplicity



Similarly, $\Xi/\Lambda, \ldots$

ALICE 2021: also in charm

arXiv:2105.05187 talk Luigi Dello Stritto



Enter: Close-Packing

"Close Packing" of strings

Fischer & Sjöstrand, 1610.09818

Even with CR, high-multiplicity events still expected to involve multiple overlapping strings.

Interaction energy ⟹ higher effective string tension **(similar to "Colour Ropes")**

⟹ **strangeness** (& baryons & <p_T>)

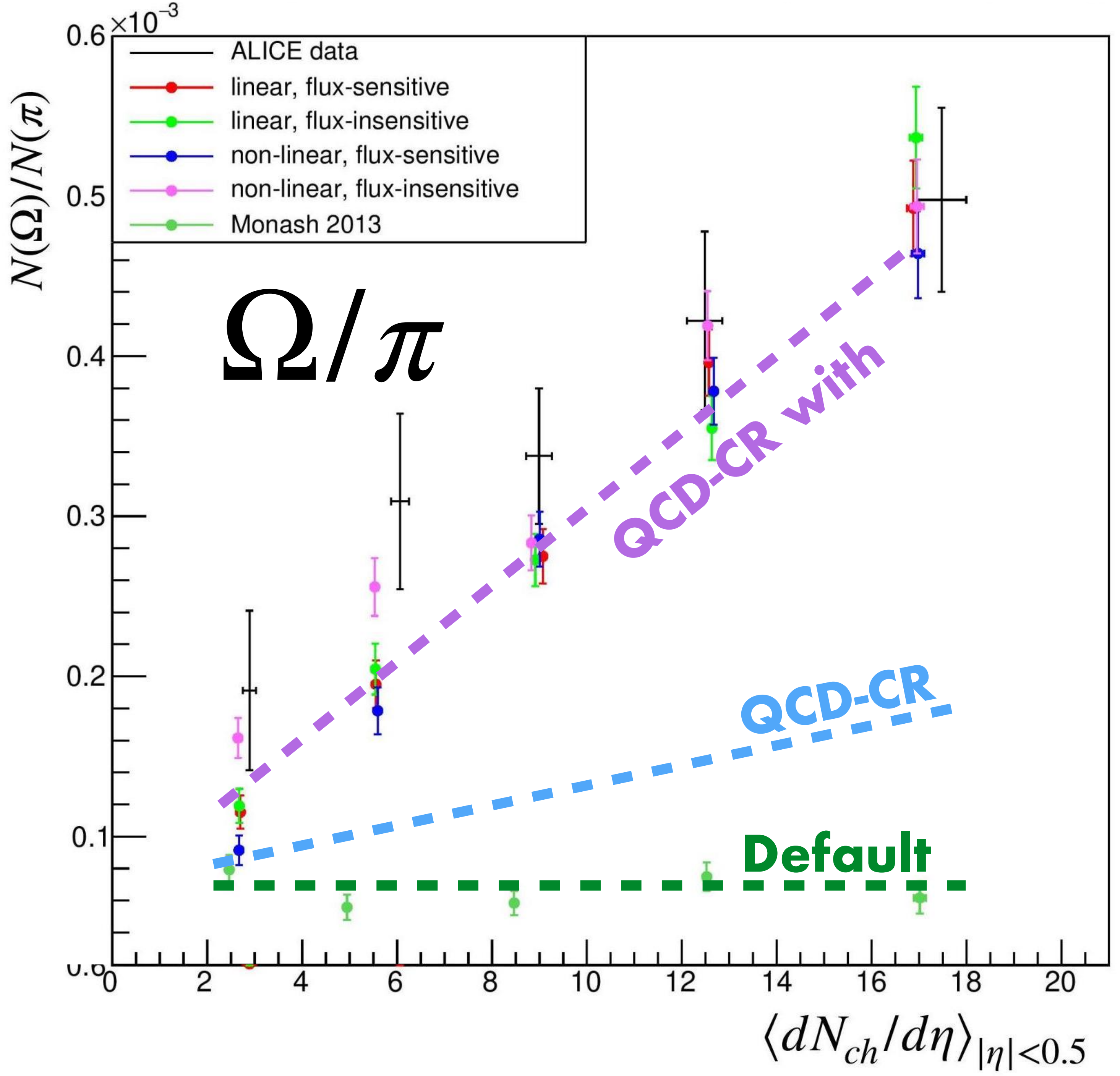
Current close-packing model in Pythia only for "thermal" string-breaking model

Interesting in its own right!

2021: Monash student J. Altmann extended it to conventional string-breaking model and began the (complicated) work to extend to junction topologies. **Work in progress!**

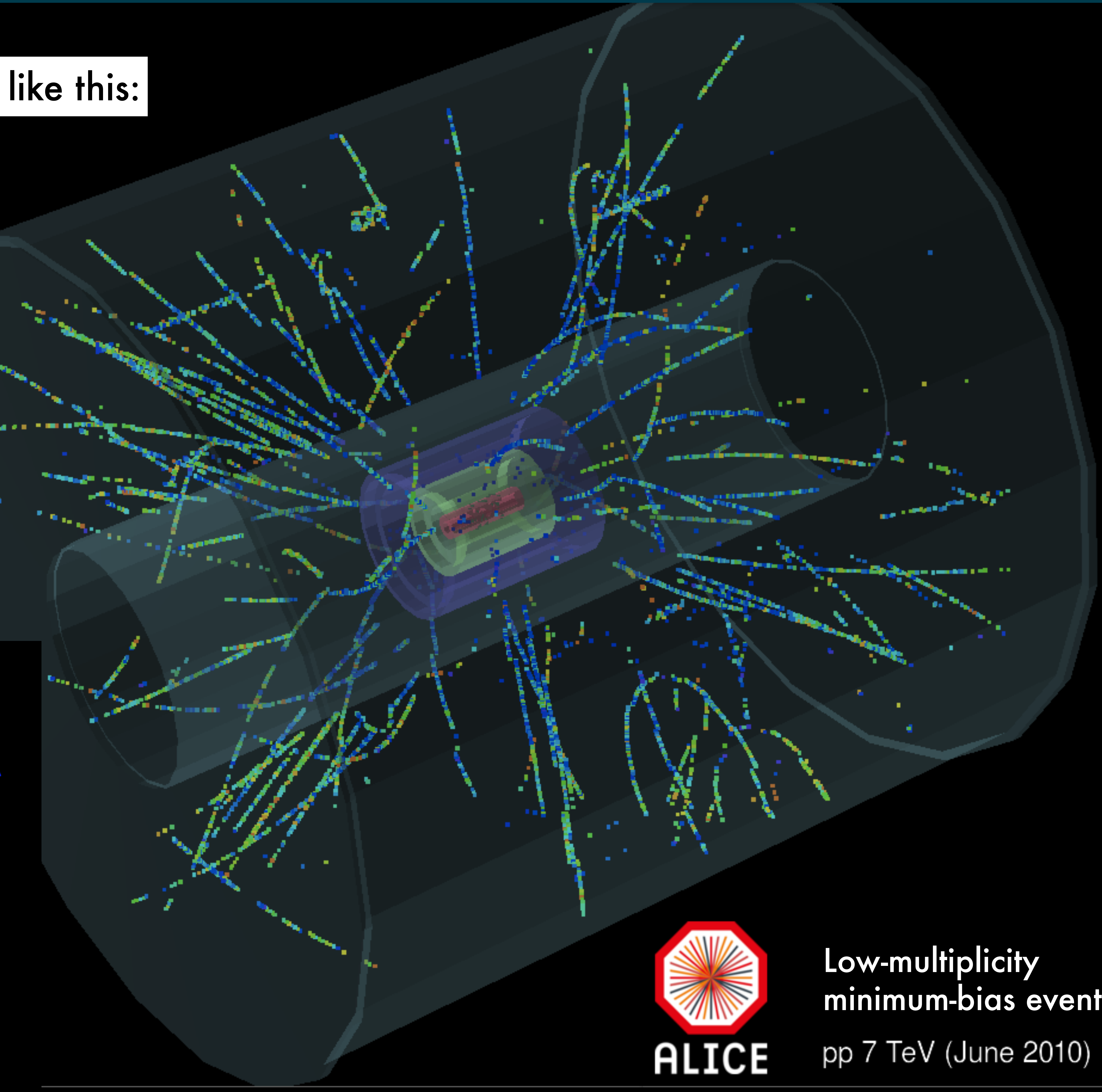
Intended as a simple alternative to rope model.

Preliminary results (J. Altmann)



What do LHC collisions look like?

Most of them look like this:

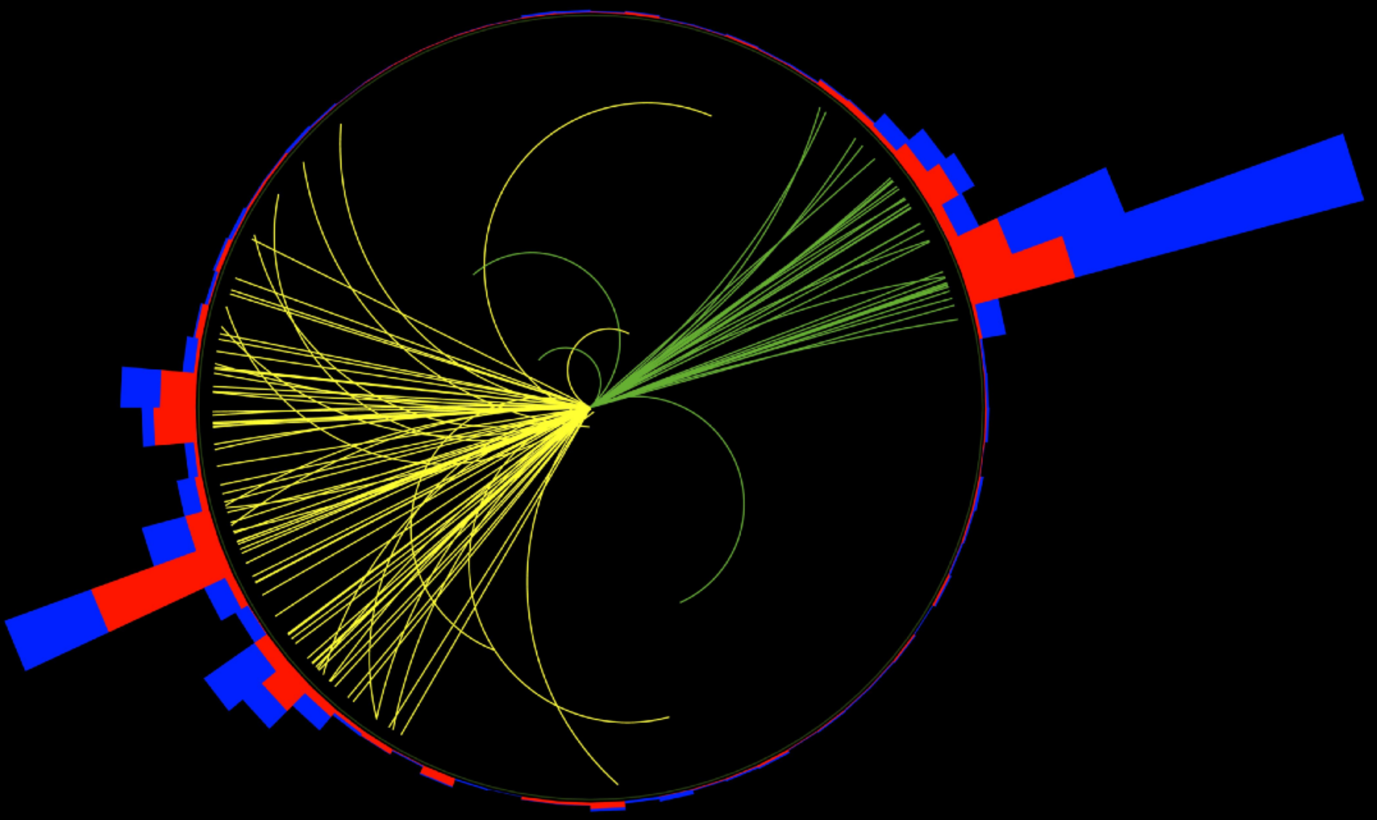


ALICE

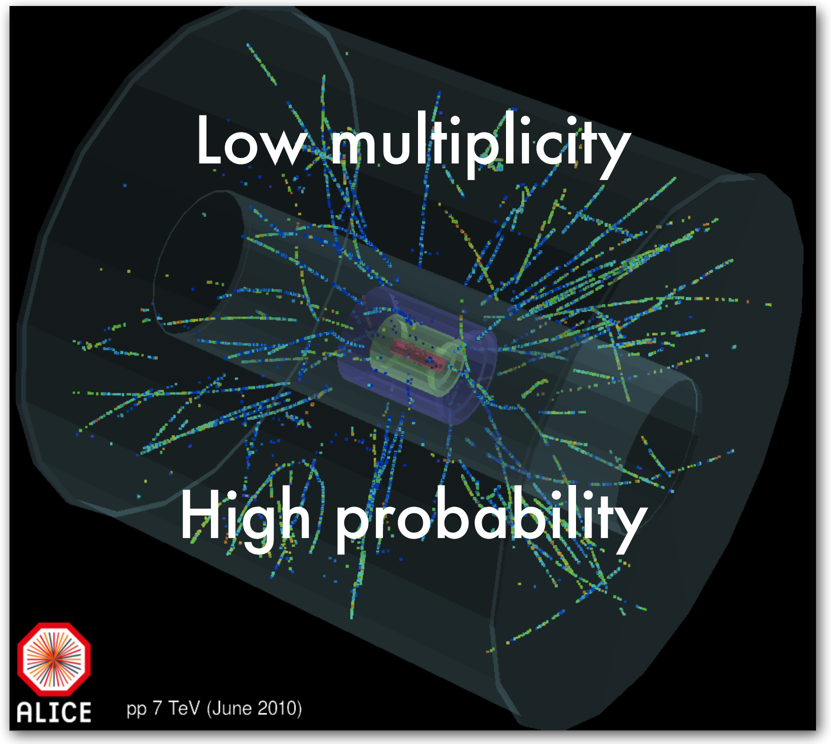
Low-multiplicity minimum-bias event

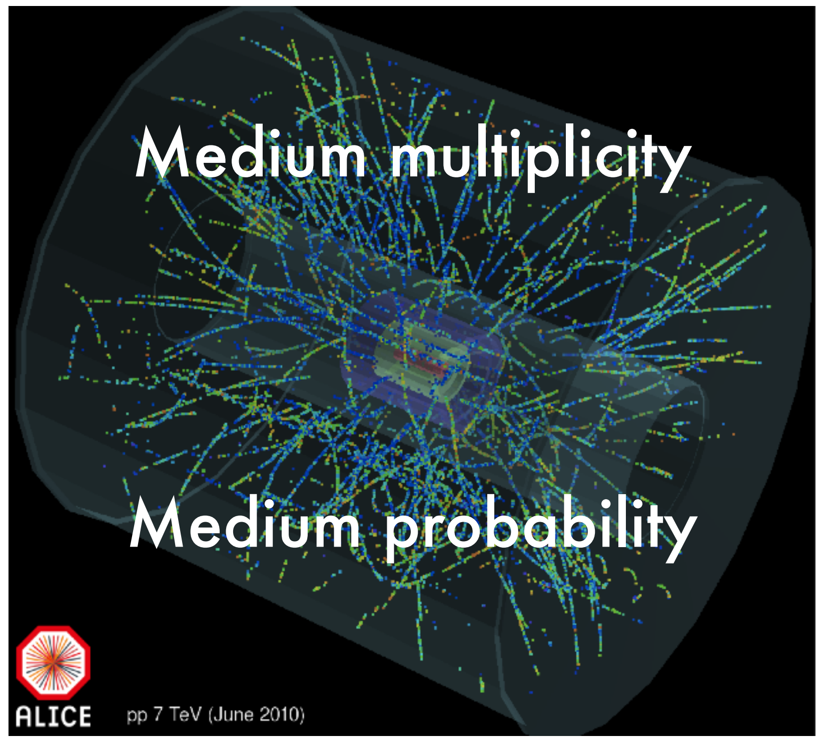
pp 7 TeV (June 2010)

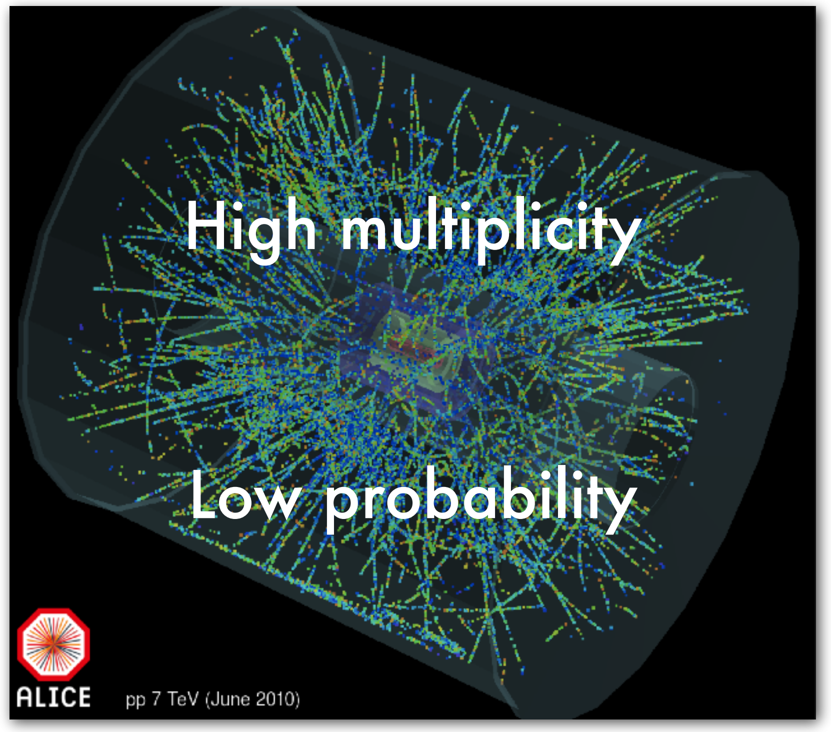
Some look like this:

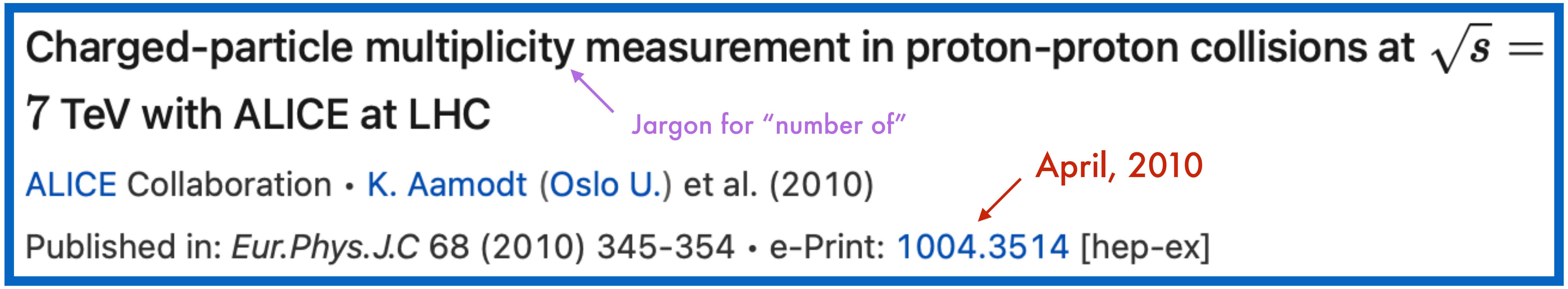


First Physics at Colliders = Counting Tracks









Charged-particle multiplicity measurement in proton-proton collisions at $\sqrt{s} =$ 7 TeV with ALICE at LHC

Jargon for "number of"

ALICE Collaboration • K. Aamodt (Oslo U.) et al. (2010)

April, 2010

Published in: *Eur.Phys.J.C* 68 (2010) 345-354 • e-Print: 1004.3514 [hep-ex]

First 7-TeV LHC measurement

Probability distribution for the **number of charged particles** (illustrated to the left with real collisions)

Experimentally: simple to measure.

Count number of **"tracks"** left by ionising charged particles

& correct for imperfect reconstruction of those tracks.

Theoretically: impossible to predict (in perturbative QFT)…

Why? Can we predict **anything at all?**

We were still able to make predictions within **~10%**; How?