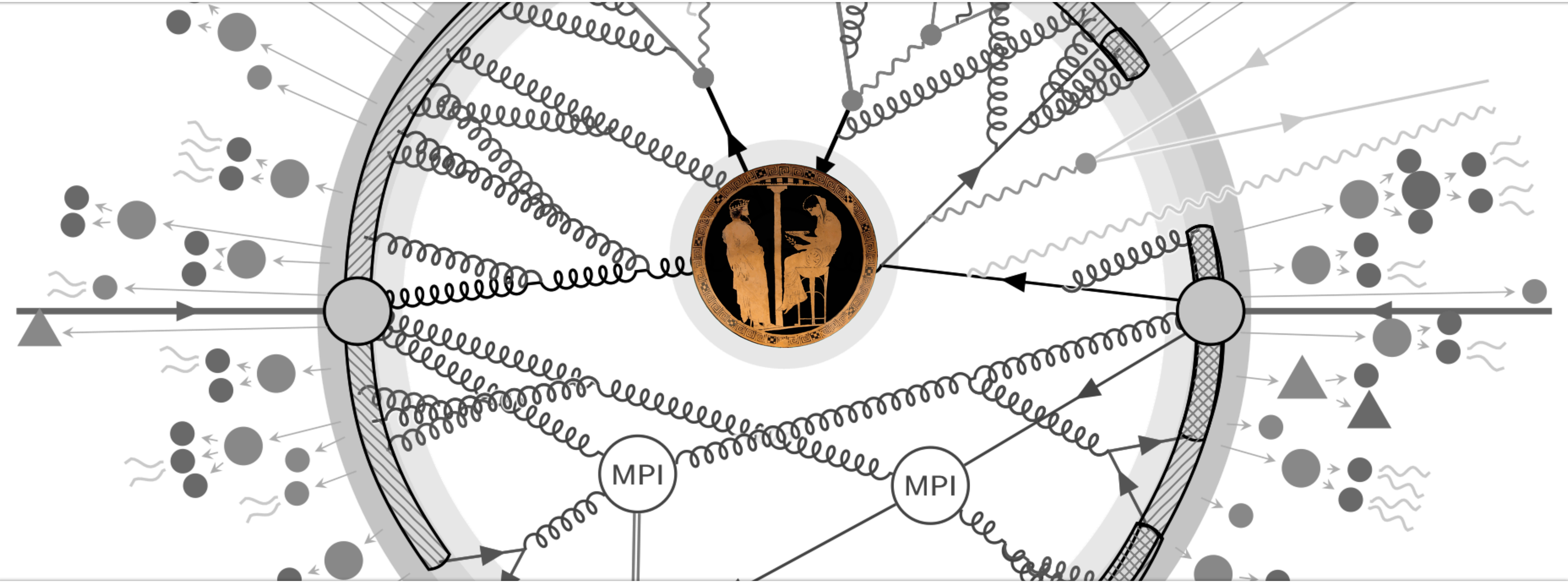
Pythia News & Modelling Uncertainties

Peter Skands — U of Oxford & Monash U.



Overview

Part 1: Perturbative Physics Modelling

1. Perturbative Uncertainties
2. Perturbative Tuning (?)
3. Beating the Factorial: Sectorized CKKW-L Merging in VINCIA

Part 2: Nonperturbative Physics Modelling

4. Automated Hadronization Uncertainties (coming in PYTHIA 8.311)
5. One-Generator Uncertainties: Simple Example (from Dark-Matter Studies)
6. Baryons and Strangeness. Particle Composition ⟷ JES

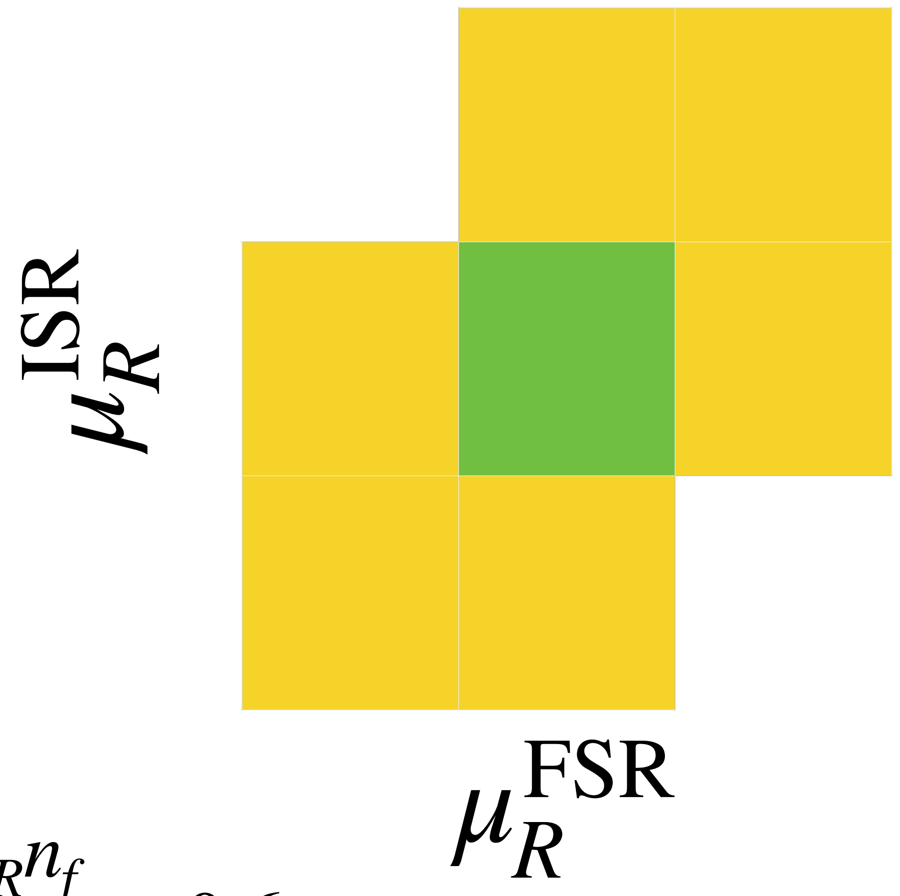
1. Perturbative Uncertainties in PYTHIA

Perturbative Uncertainties

First guess: renormalisation-scale variations,

$\mu_R^2 \to k_\mu \, \mu_R^2$, with constant $k_\mu \in [0.5, 2]$ or $[0.25, 4]$, …

+ e.g., do for ISR and FSR separately → **7-point variations** ⟶



Induces "nuisance" terms beyond calculated orders

Running of $\alpha_s(k\,\mu^2) \;=\; \alpha_s(\mu^2)\dfrac{1}{1 + b_0\alpha_s(\mu^2)\ln(k)}$ with $b_0 = \dfrac{11N_C - 4T_R n_f}{12\pi} \sim 0.6$

$\Longrightarrow$ ME proportional to $\alpha_s^n(\mu^2)\left(\, 1 \;\pm\; \underbrace{b_0\,\alpha_s(\mu^2)\,\ln k^n}_{\text{variation}} \;+\; \ldots \right)$

I think many people suspect this is unsatisfactory and unreliable

Problem: little guidance on what else to do …

What are the issues?

Issue #1: Multiscale Problems (e.g., a couple of bosons + a couple of jets)

Not well captured by **any** variation k_μ around any **single** scale

More of an issue for hard-ME calculations than for showers (which are intrinsically multiscale)

Best single-scale approximation = **geometric mean of all relevant QCD scales**

My recommendation: vary which scales enter geometric mean

Issue #2: Terms that are not proportional to the lower orders

Renormalization-scale variations always proportional to what you already:

$$\mu_R \text{ variations} \Longrightarrow \mathrm{d}\sigma \ \to (1 \pm \Delta\alpha_s)\,\mathrm{d}\sigma$$

No new kinematic dependence

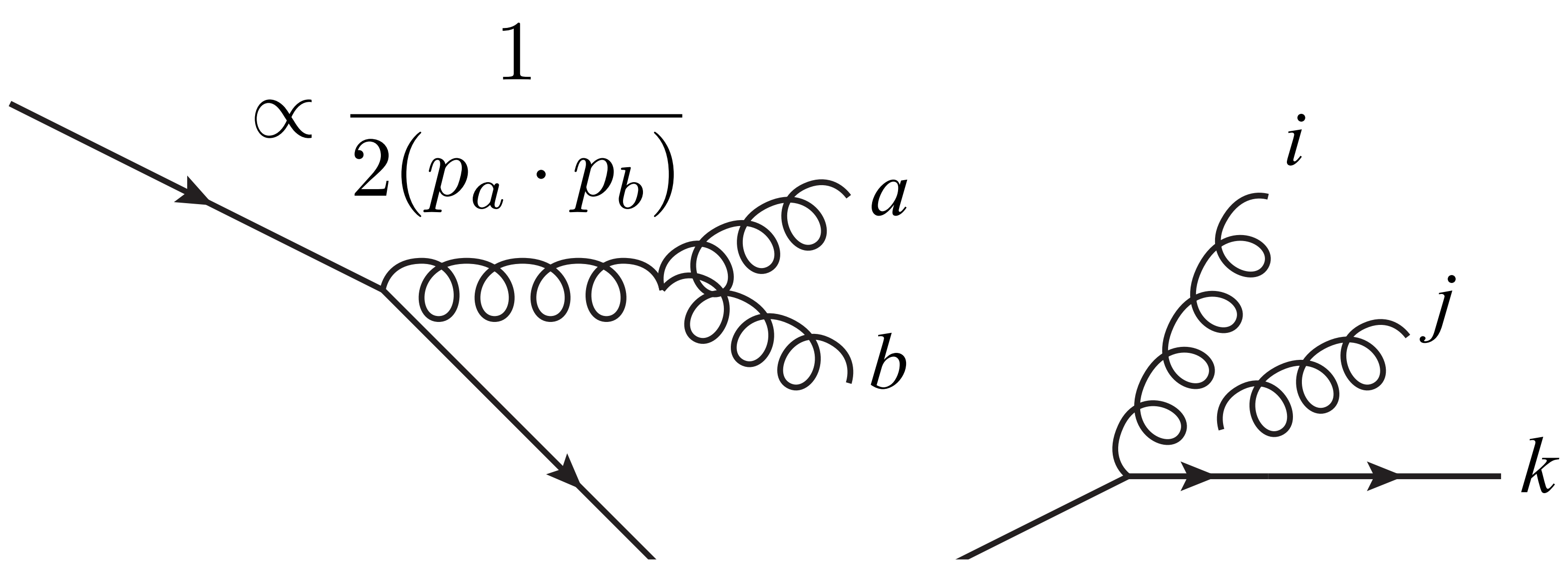
But full higher-order matrix elements will also contain **genuinely new terms** at each order, not proportional to previous orders:

$$\text{More general} \Longrightarrow \mathrm{d}\sigma \to \mathrm{d}\sigma \ \pm \ \Delta\mathrm{d}\sigma$$

Parton Showers: Theory

see e.g PS, *Introduction to QCD*, TASI 2012, arXiv:1207.2389

Most bremsstrahlung is driven by divergent propagators → simple structure



Mathematically, **gauge amplitudes factorize** in **singular limits**

Partons ab → **collinear**:

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

$P(z)$ = **DGLAP splitting kernels**", with $z = E_a/(E_a + E_b)$

Gluon j → **soft**:

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

Coherence → Parton j really emitted by (i,k) "dipole" or **"antenna" (eikonal factors)**

These are the **building blocks of parton showers** (DGLAP, dipole, antenna, …)

(+ running coupling, unitarity, and explicit energy-momentum conservation.)

VINCIA & PYTHIA 8: Non-Singular Variations

OK, so we know the (leading) pole structures of QCD amplitudes; parton-shower approximations are anchored there:

Formally: $|M_{n+1}|^2 \sim \sum_{\text{radiators}} a_{\text{sing}} |M_n|^2$ + iterations/nestings → $|M_{n+m}|^2$

$\uparrow$

$a_{\text{sing}} = 1/Q^2$ poles from singular propagators, with spin-dependent numerators

Renormalization-scale variations only produce terms $a_{\text{sing}} \to (1 + \Delta\alpha_s)\, a_{\text{sing}}$

But genuine matrix elements also have "non-singular terms"

Our solution: Non-singular variations

$$a_{\text{sing}} \to a_{\text{sing}} + \Delta a_{\text{non-sing}}$$

VINCIA (2011): [Giele, Kosower, **PS** PRD84 (2011) 054003]

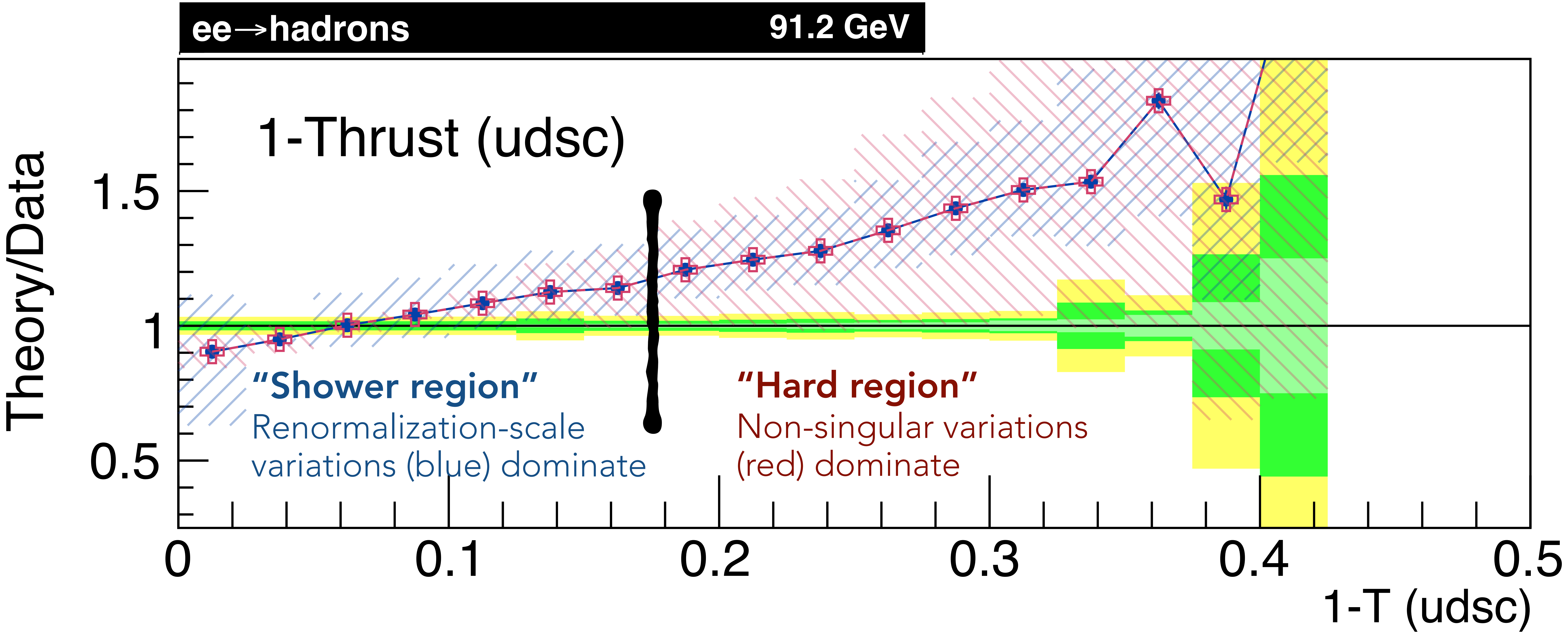
PYTHIA (2016): [Mrenna, **PS** PRD94 (2016) 7]

Can also indicate whether higher matching/merging is needed or not

Non-Singular Variations: Example

*Example from Mrenna & **PS**, "Automated Parton-Shower Variations in Pythia 8", 1605.08352*

Can vary **renormalisation-scale** and **non-singular terms** independently



"Shower region"
Renormalization-scale variations (blue) dominate

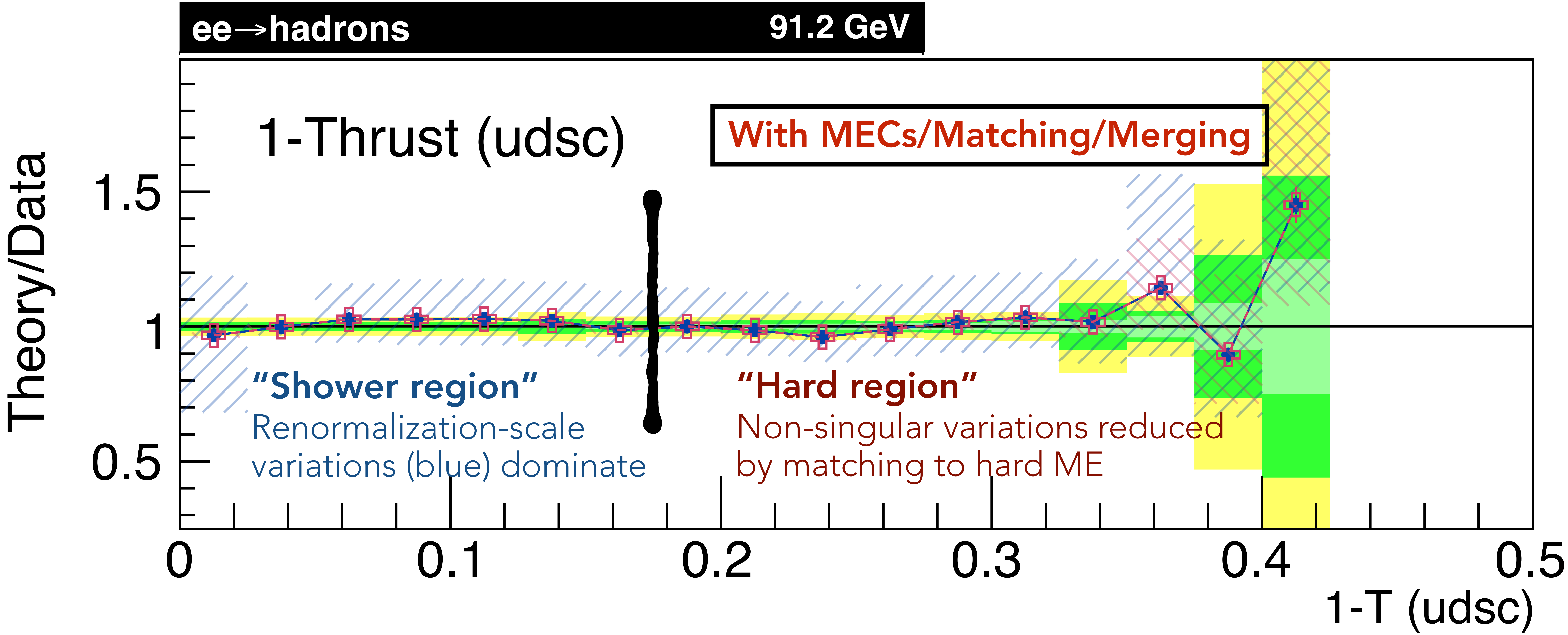
"Hard region"
Non-singular variations (red) dominate

Note: ME corrections were switched off for illustration here. Would reduce red band, but not blue.

Non-Singular Variations: Effect of Matching to Matrix Elements

Example from Mrenna & ***PS****, "Automated Parton-Shower Variations in Pythia 8", 1605.08352*

Can vary **renormalisation-scale** and **non-singular terms** independently



"Shower region"
Renormalization-scale variations (blue) dominate

"Hard region"
Non-singular variations reduced by matching to hard ME

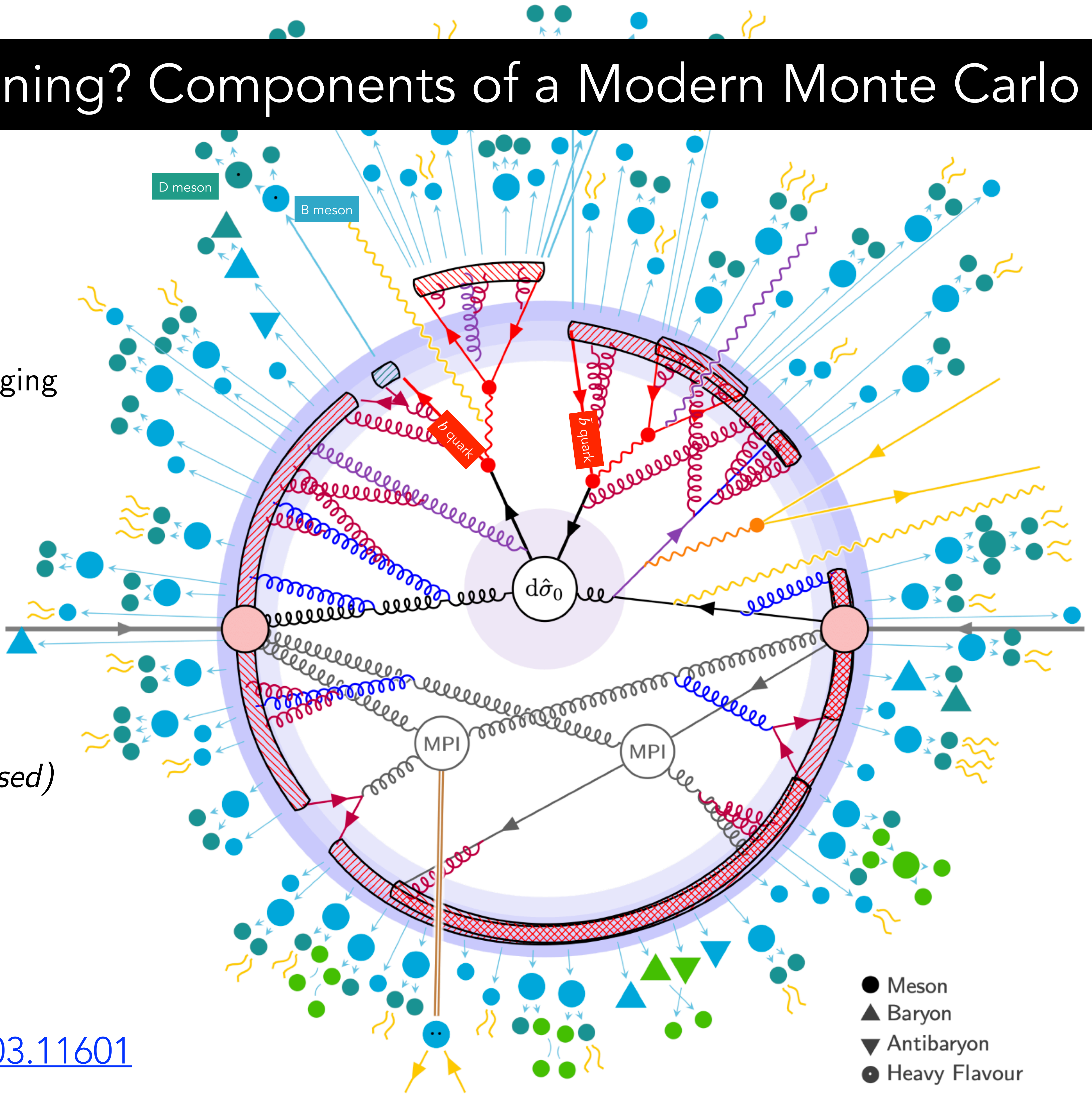
3. Perturbative Tuning (?)

What are we tuning? Components of a Modern Monte Carlo Event Generator:

Parton Level

- ○ Hard Interaction
- ● Resonance Decays
- 🟪 MECs, Matching & Merging
- 🟥 FSR
- 🟦 ISR*
- 🟨 QED
- 🟧 Weak Showers
- 🟫 Hard Onium
- ○ Multiparton Interactions
- ☐ Beam Remnants*

(: incoming lines are crossed)*



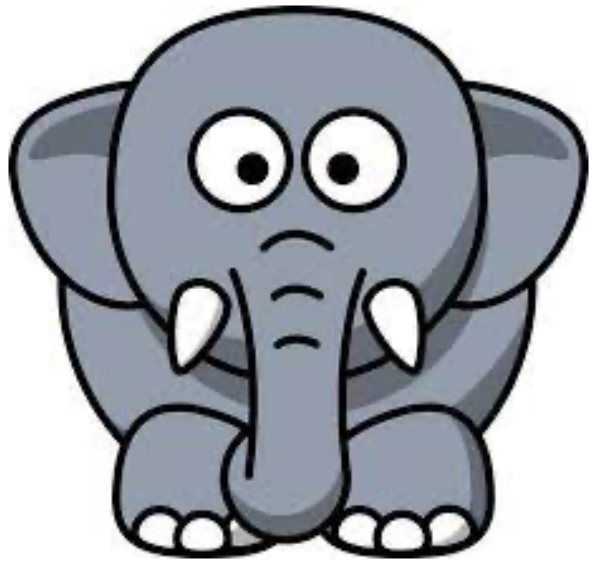
Hadron Level

- ☐ Beam Remnants*
- ▨ Strings
- ▨ Clusters
- ■ Colour Reconnections
- ■ String Interactions
- ■ Bose-Einstein & Fermi-Dirac
- 🟦 Primary Hadrons
- ■ Secondary Hadrons
- 🟩 Hadronic Reinteractions
- ■ QED in Hadron Decays

(: incoming lines are crossed)*

Figure from arXiv:2203.11601

Tuning at Parton Level (?)



The Elephant in the Auditorium:

Purist: you should not "tune" perturbation theory!

Uncalculated orders / coefficients should be set to zero.

Most obvious stance for a theorist to take.

Goal: a theory calculation that delivers a clean simple-to-understand prediction, at a stated accuracy.

It may agree or disagree with data. That's ok, consistent with the stated accuracy.

It may disagree a **lot** with data. Not the theorist's problem.

(ATLAS and CMS may end up with a problem.)

But … Parton Showers *always* generate subleading structures …

Hard to control and generally not possible to set cleanly to zero.

Pythia Philosophy (1)

Vice to Virtue: nothing special about zero as guess for higher orders.

Goal: deliver a description that **faithfully** represents as much data as possible.

Challenge: **avoid doing violence** to the underlying physics model (➜ GIGO).

1) Allow explicit/controlled coefficients to deviate from exact values

Theoretically consistent **if deviation $\lesssim$ uncalculated corrections**.

PYTHIA example: use effective values for $\alpha_s(M_Z)$, consistent with other LO determinations of it.

E.g., : LO PDFs ➜ $\alpha_s(M_Z) \sim 0.14$; LO event shapes at LEP also give $\alpha_s(M_Z) \sim 0.14$.

Slightly extreme: our 1-loop α_s "magic trick" for NLO-level agreement at LEP

Caveat: no guarantee of universality!

Pythia Philosophy (2)

2) Control for non-universalities

Consider several complementary observables, processes, and contexts

Possibly weighted by how much you care about each

E.g., for the effective FSR α_s value in Pythia

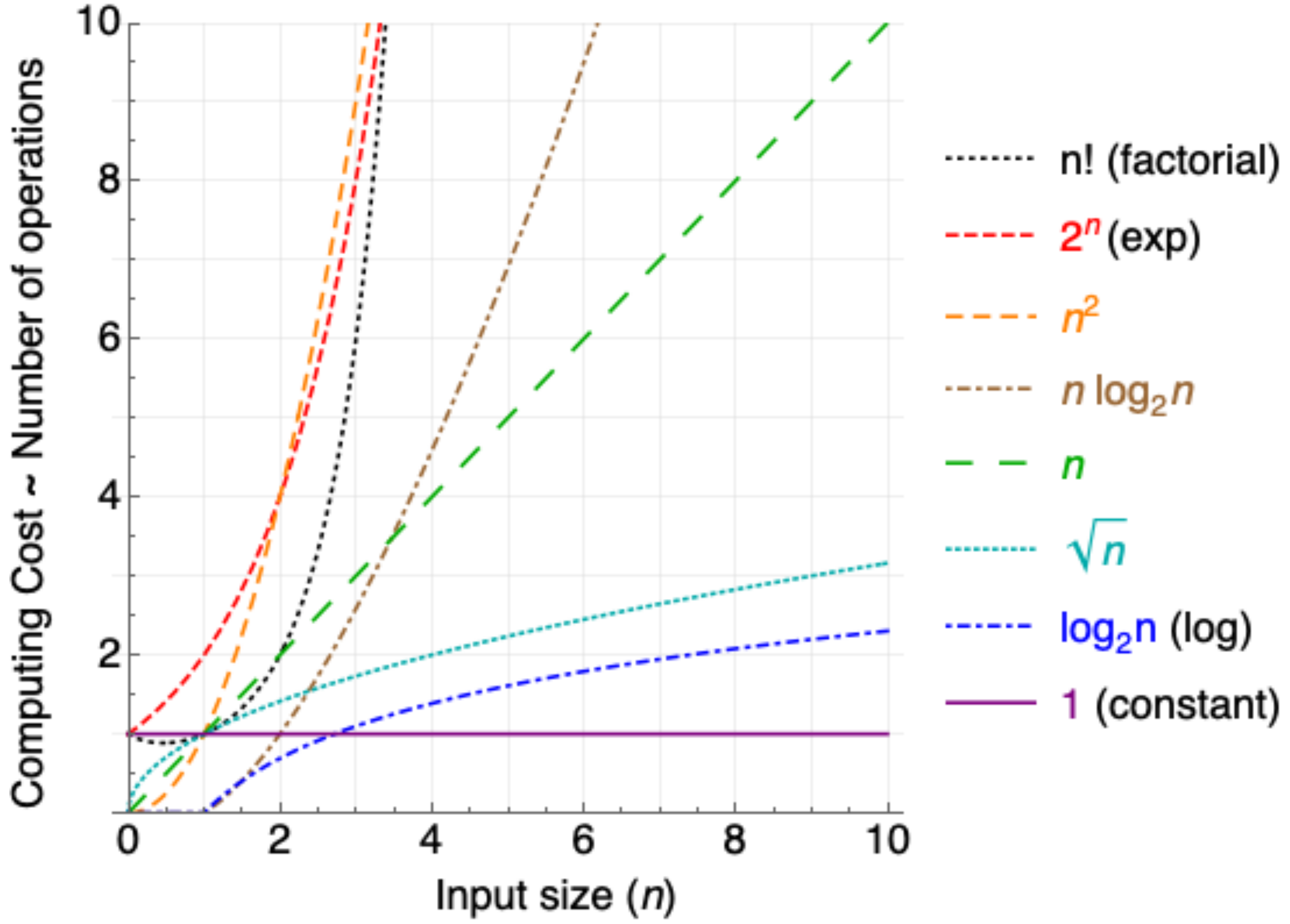
At LEP, we have 3-jet LO MECs and use **3- and 4-jet event shapes** + ditto **jet rates** as main constraints (universality across jet multiplicities)

And then we cross check with **jet shape profiles & jet substructure at the LHC**.

Always a risk that this can fail. E.g., tensions between different processes at LHC (eg top); experiments retune α_s and associated worries.

One thorny example: b-quark fragmentation in the top decay jet.

Hard to be consistent in context of matching and merging $\Longrightarrow$ needs attention & work!



3. Beating the Factorial

Matrix-Element Merging — The Complexity Bottleneck

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

Need to take **all contributing shower histories into account**.

In conventional parton showers (Pythia, Herwig, Sherpa, …)

Each phase-space point receives contributions from many possible branching "histories" (aka "clusterings")

# of histories grows ~ # of Feynman Diagrams, **faster than factorial**

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

Bottleneck for merging at high multiplicities (+ high code complexity)

NEW

Sector Showers (without maths)

PS & Villarejo JHEP 11 (2011) 150
Brooks, Preuss, **PS** JHEP 07 (2020) 032

VINCIA's shower is unique in being a "Sector Shower"

Partition N-gluon Phase Space into N "sectors" (using step functions).

Each sector corresponds to one specific gluon being the "softest" in the event — the one you would cluster if you were running a jet algorithm (ARCLUS)

Inside each sector, **only a single kernel is allowed to contribute** (the most singular one)!

Sector Kernel = the eikonal for the soft gluon and its collinear DGLAP limits for $z > 0.5$.

➜ Unique properties: shower operator becomes **bijective** and is a true **Markov chain**

The crucial aspect:

Only a single history contributes to each phase-space point !

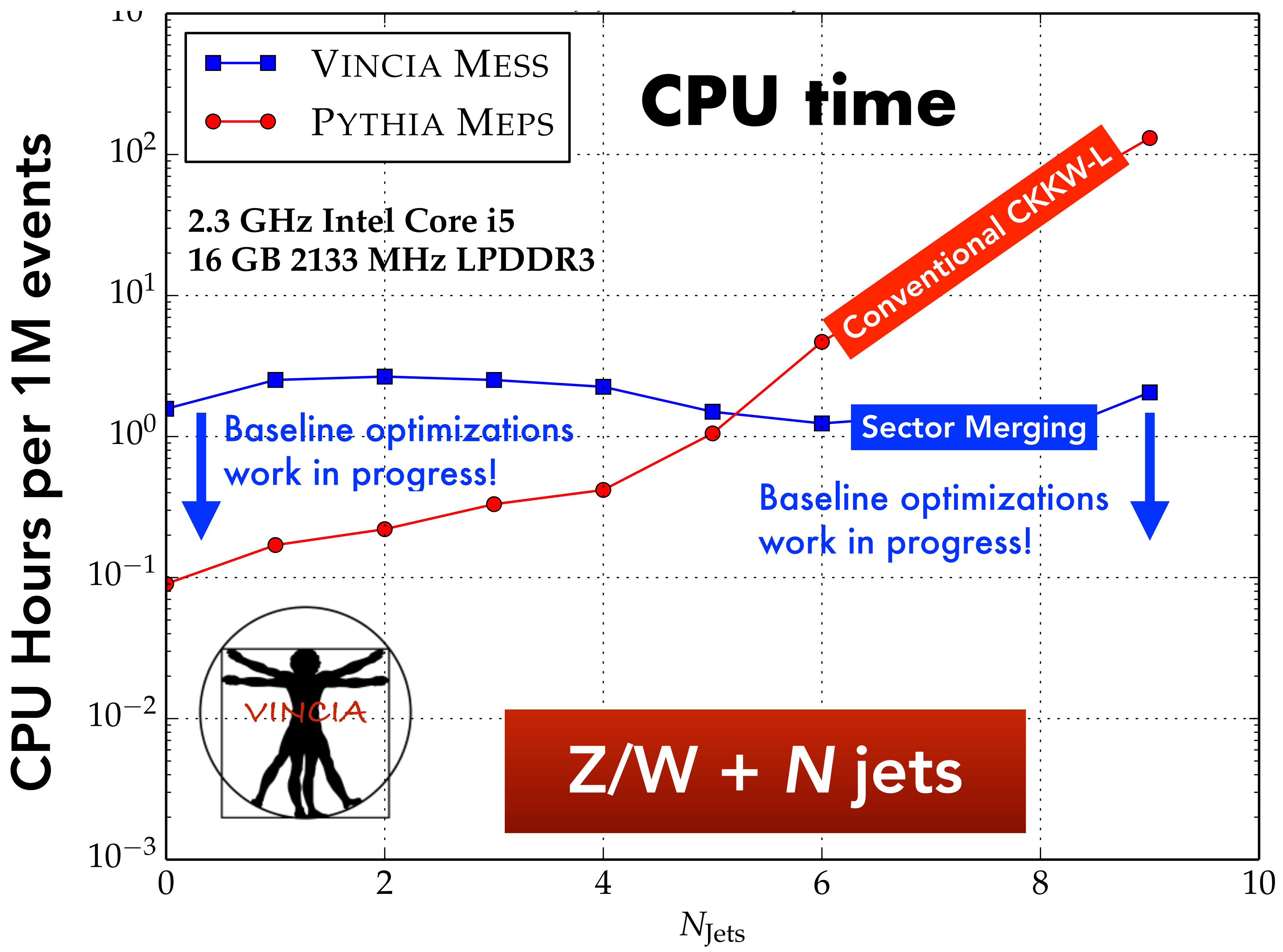
$\Longrightarrow$ **Factorial growth of number of histories reduced to constant!**

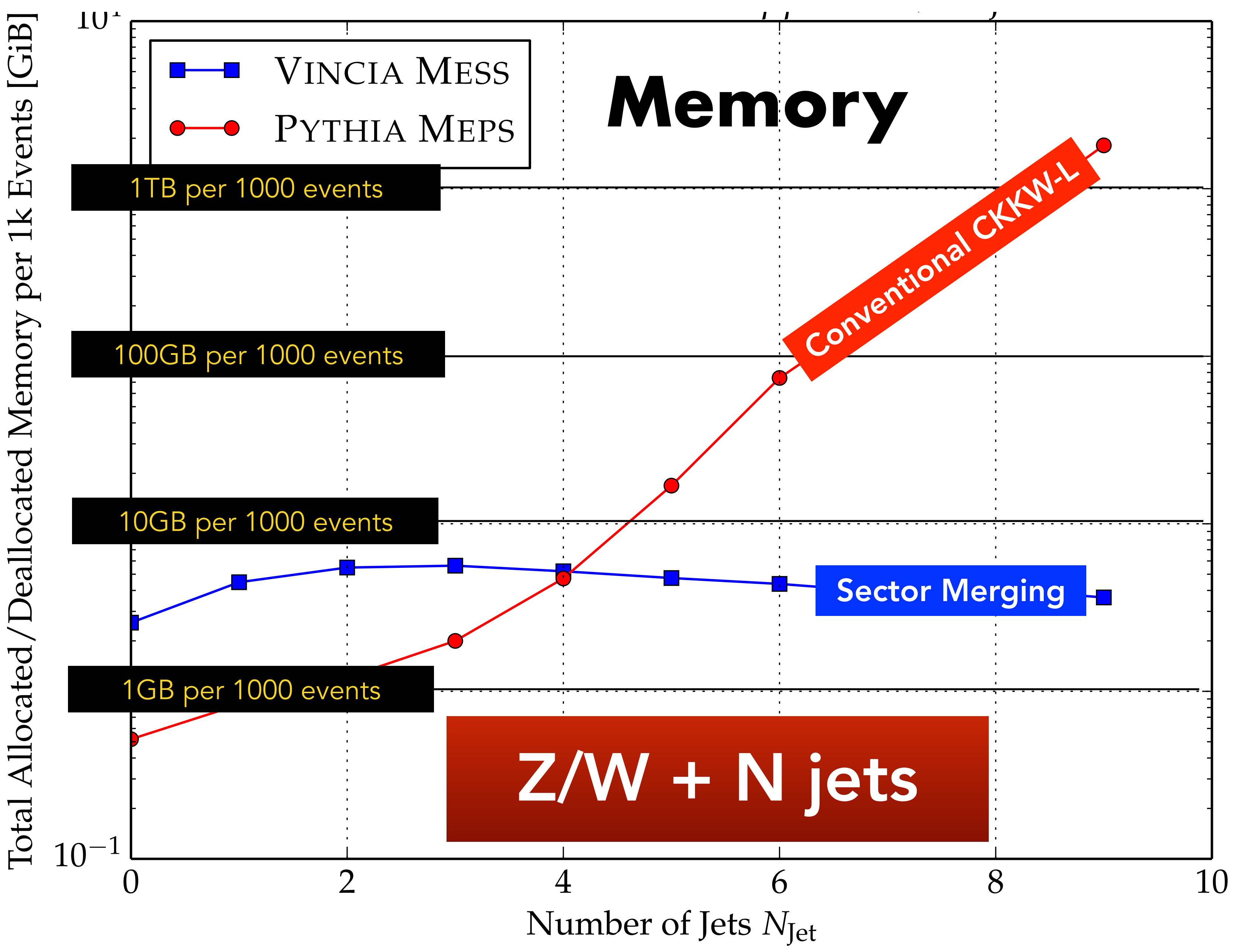
(And the number of sectors only grows linearly with the number of gluons)

($g \to q\bar{q}$ ➜ leftover factorial in number of *same-flavour* quarks; not a big problem)

Sectorized CKKW-L Merging publicly available from Pythia 8.306

Brooks & Preuss, “Efficient multi-jet merging with the VINCIA sector shower”, arXiv:2008.09468





Extensions now pursued:

Sectorized **matching at NNLO** (proof of concepts in arXiv:2108.07133 & arXiv:2310.18671)

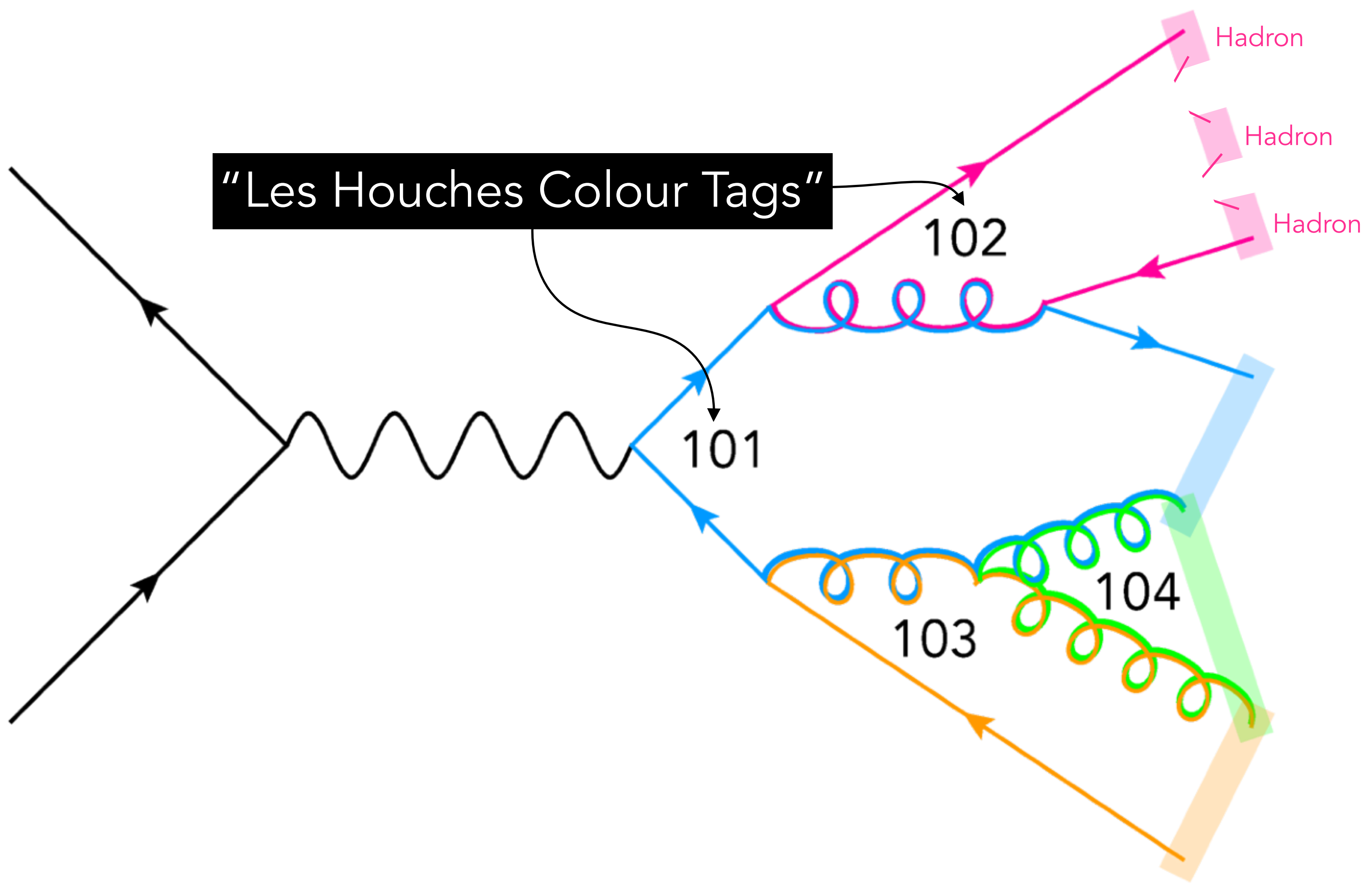
Sectorized **iterated tree-level ME corrections** (demonstrated in PS & Villarejo arXiv:1109.3608)

Sectorized **multi-leg merging at NLO** (active research grants, with **C. Preuss, Wuppertal**)

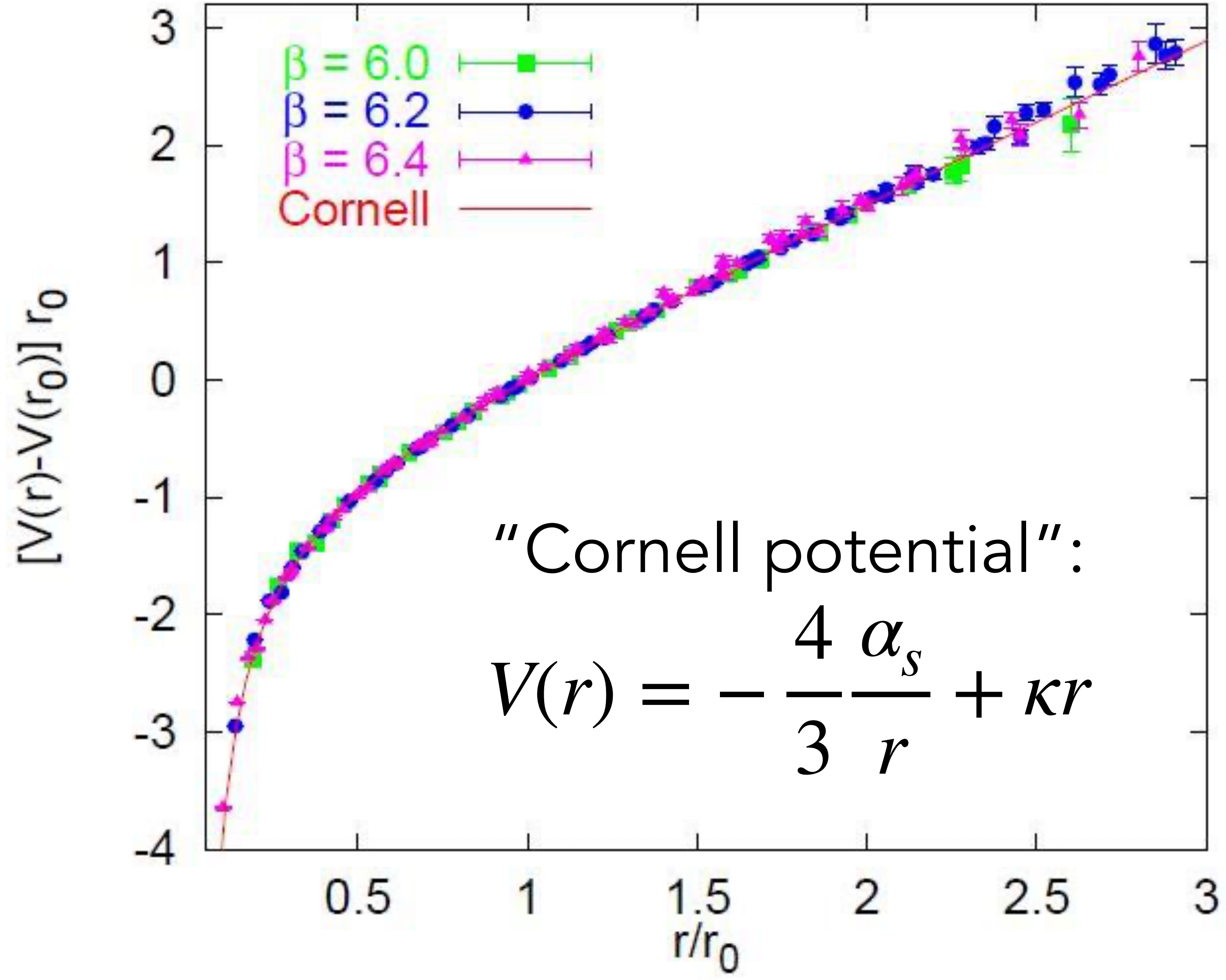
4. Automated Hadronization Uncertainties

Confinement in PYTHIA: *The Lund String Model*

Simplified (leading-N_C) **"colour flow"** ➜ determine between which partons to set up confining potentials



"Linear confinement"
(From Lattice & Hadron Spectroscopy)



"Cornell potential":

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

Map from Partons to Strings:

Quarks ➡ string endpoints; gluons ➡ transverse "kinks"

System then evolves as a string world sheet

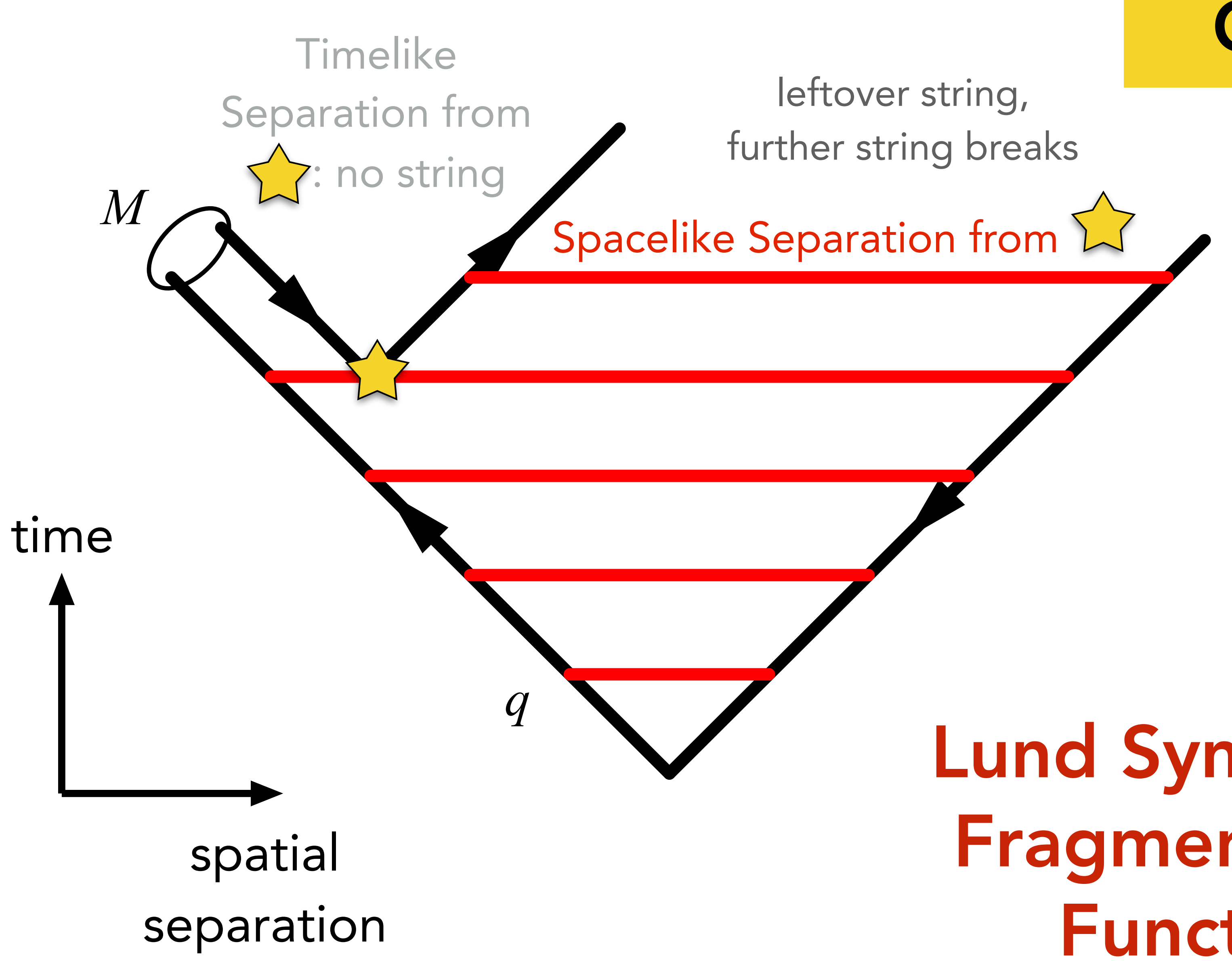
+ **String breaks** via spontaneous $q\bar{q}$ pair creation ("Schwinger mechanism") → **hadrons**

The String Fragmentation Function

Consider a string break ☆, producing a meson M, and a leftover string piece

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

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



Observation: All string breaks are **causally disconnected**

Lorentz invariance $\Longrightarrow$ string breaks can be considered in *any order*. Imposes "left-right symmetry" on the **FF**

$\Longrightarrow$ **FF** constrained to a form with **two free parameters**, a & b: constrained by fits to measured hadron spectra

Lund Symmetric Fragmentation Function

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

Supresses high-z hadrons

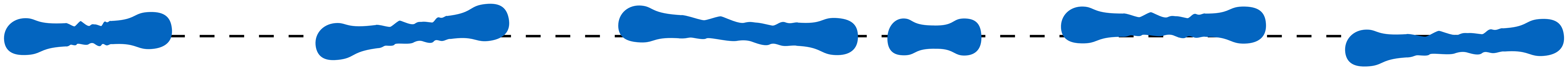
Supresses low-z hadrons

Automated Hadronization Uncertainties

NEW

Problem:

Given a colour-singlet system that (randomly) broke up into a specific set of hadrons:



What is the **relative probability** that same system would have resulted, if the fragmentation parameters had been **different?**

Would this particular final state become **more likely** $(w' > 1)$? Or **less likely** $(w' < 1)$

Crucially: **maintaining unitarity** $\implies$ inclusive cross section remains unchanged!

August 2023: Bierlich, Ilten, Menzo, Mrenna, Szewc, Wilkinson, Youssef, Zupan

[Reweighting MC Predictions & Automated Fragmentation Variations in Pythia 8, 2308.13459]

Method is general; demonstrated on variations of the 7 main parameters governing longitudinal and transverse fragmentation functions in PYTHIA 8

https://gitlab.com/uchep/mlhad-weights-validation

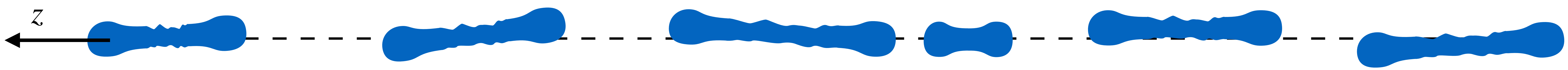
Pythia 8.311

NEW

Demonstration

[Reweighting MC Predictions & Automated Fragmentation Variations in Pythia 8, 2308.13459]

Example: Longitudinal Fragmentation Function (Lund Symmetric FF)



$f(z)$ ~ ***scaled*** **light-cone hadron momentum fraction**

$$\propto \frac{1}{z^{1+r_Q b m_Q^2}} (1-z)^{a} \exp\left(-\frac{b m_\perp^2}{z}\right)$$

variations

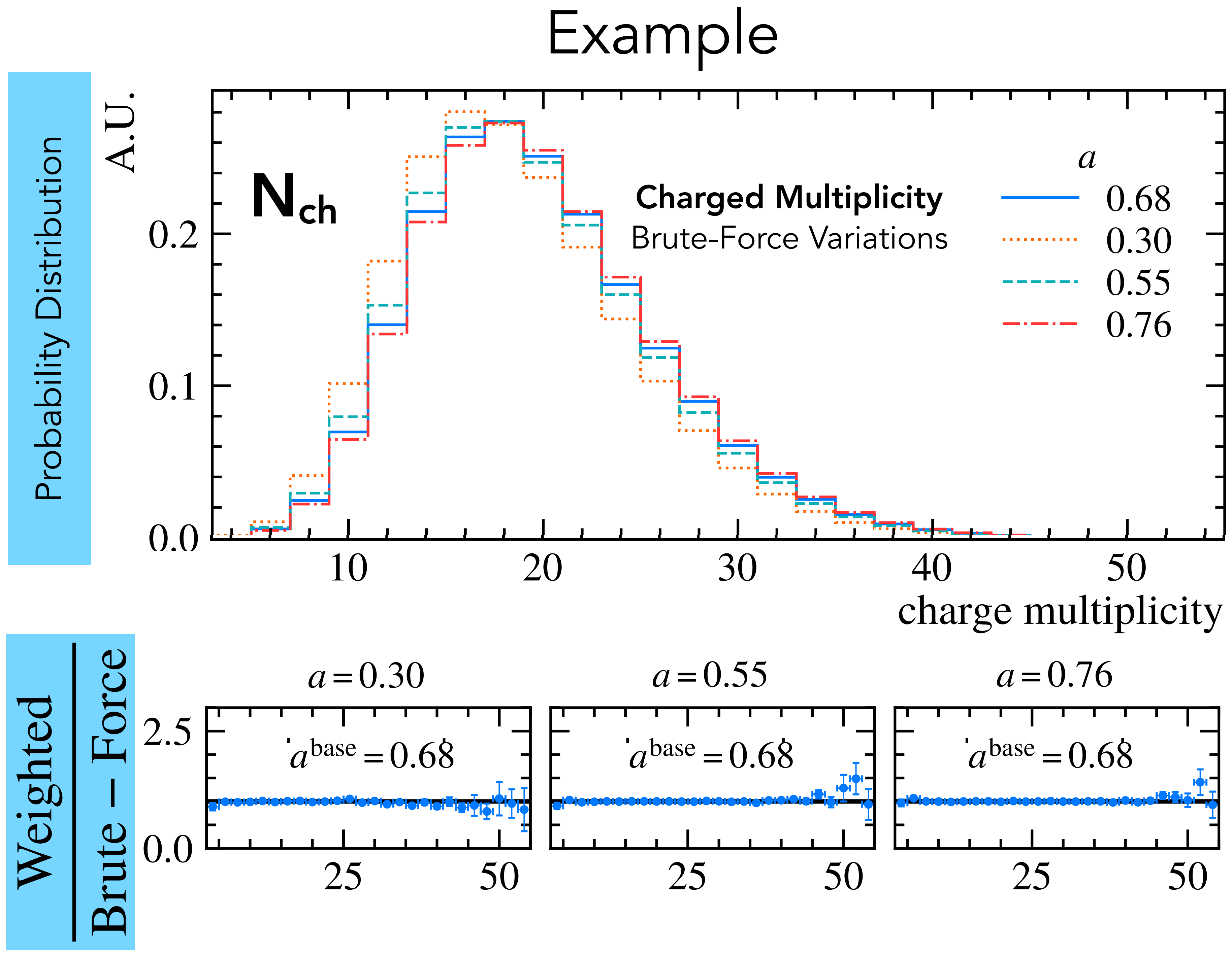
Reweighting Methodology:

Accept-Reject Algorithm (analogous to shower variations):

$$w' = w \prod_{i \in \text{accepted}} R'_{i,\text{accept}}(z) \prod_{j \in \text{rejected}} R'_{j,\text{reject}}(z),$$

with

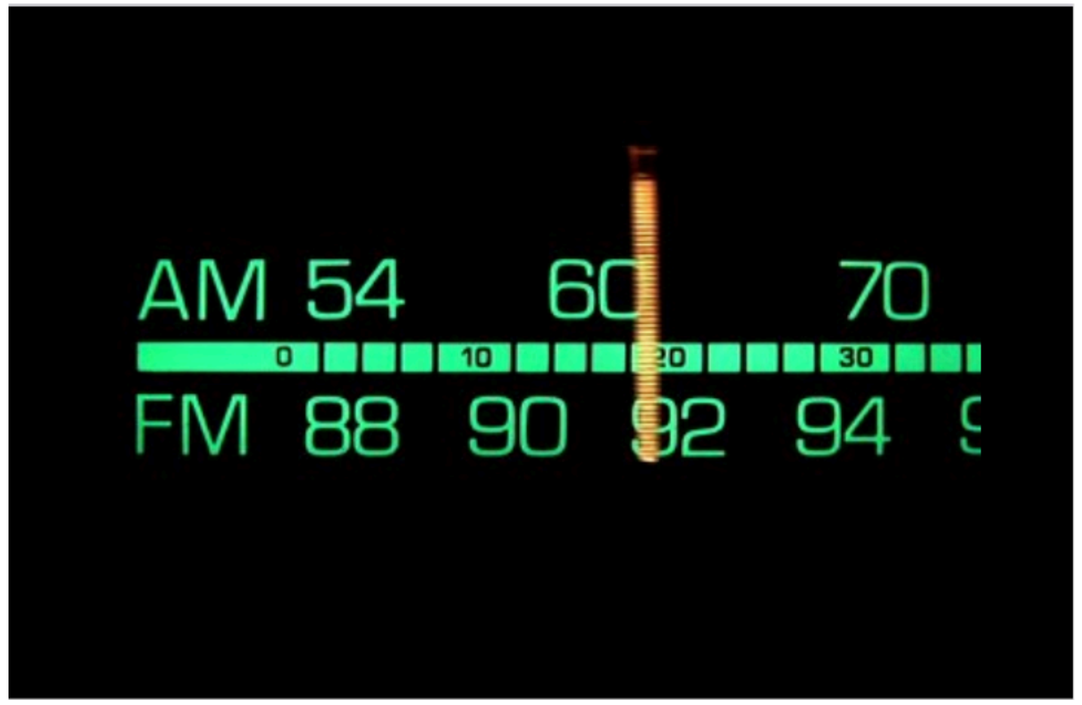
$$R'_{\text{accept}}(z) = \frac{P'_{\text{accept}}(z)}{P_{\text{accept}}(z)} \qquad R'_{\text{reject}}(z) = \frac{P'_{\text{reject}}(z)}{P_{\text{reject}}(z)} = \frac{1 - P'_{\text{accept}}(z)}{1 - P_{\text{accept}}(z)}$$

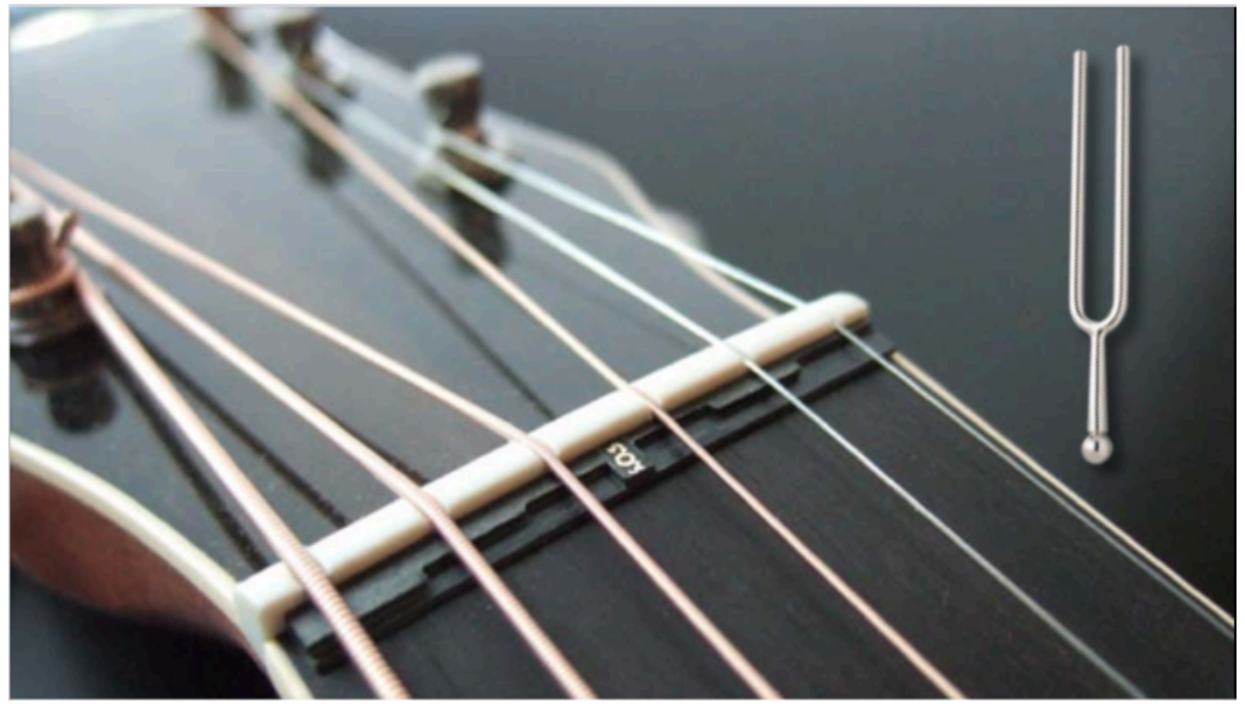

Example

5. One-Generator Hadronization Uncertainties

(Simple Example from Dark-Matter Studies)

Tuning





Tuning: PROFESSOR — a powerful tool for (semi)automated tuning

Inspired by idea pioneered by DELPHI (Hamacher et al., 1995):

Bin-wise interpolation of MC generator response and χ^2 minimization

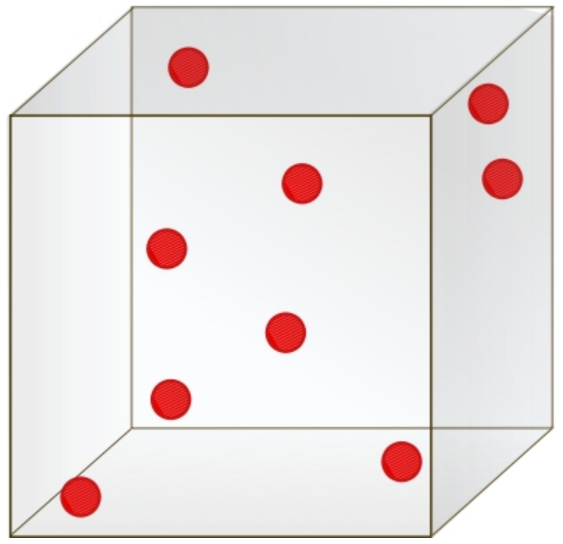
2^{nd}-order polynomials account for parameter correlations.

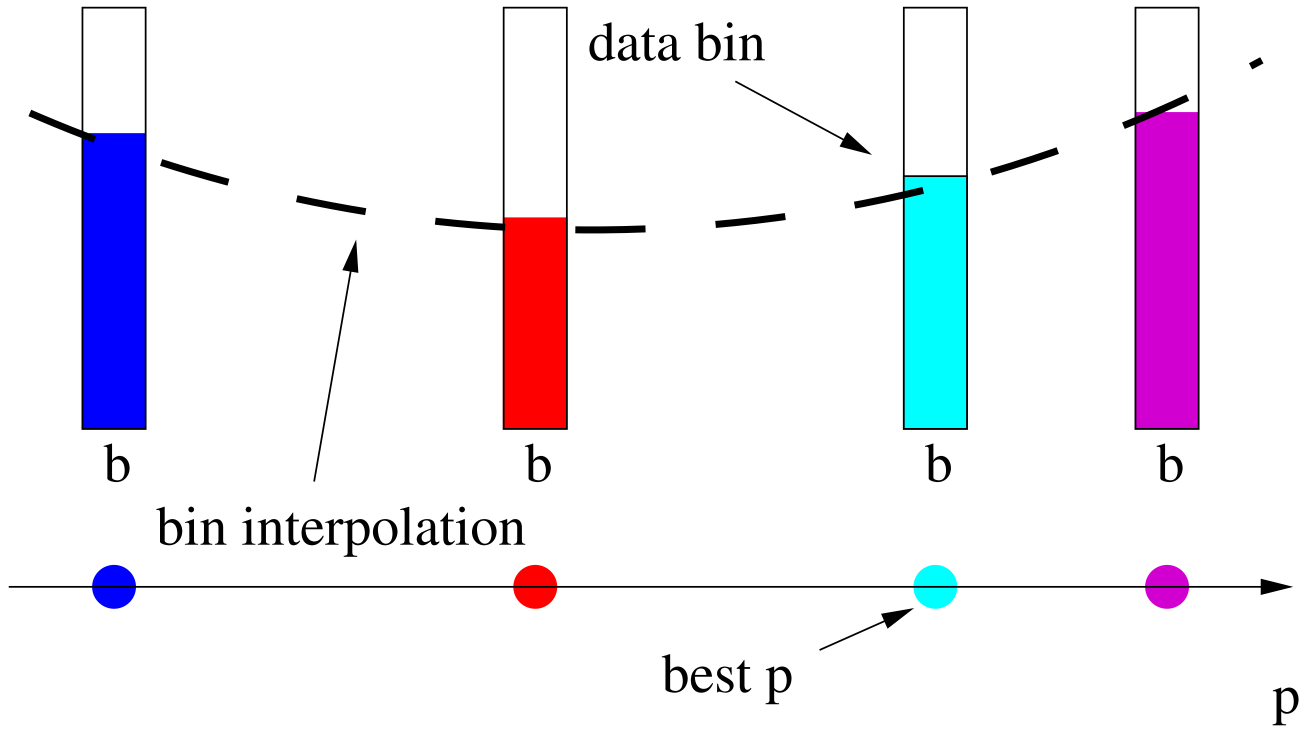
→

Modern Python Package
with much more functionality,
tutorials, etc.
https://professor.hepforge.org/

Professor
Tuning procedure

1. Random sampling: N parameter points in n-dimensional space
2. Run generator and fill histograms
3. For each bin: use N points to fit interpolation (2^{nd} or 3^{rd} order polynomial)
4. Construct overall (now trivial) $\chi^2 \approx \sum_{bins} \frac{(interpolation - data)^2}{error^2}$
5. and Numerically *minimize* `pyMinuit, SciPy`





PROFESSOR — Some Caveats

A. Buckley et al., *EPJC*65 (2010) 331

https://professor.hepforge.org/

Fitting an imperfect theory model — with unknown uncertainties

Overfitting: very precisely measured data points can generate large χ^2 values

Even if MC gets within what one would naively consider "reasonable" agreement

Fit reacts by **sacrificing agreement elsewhere** (typically in tails) to improve χ^2 in peaks.

PROFESSOR now has facility to include a "sanity limit" (e.g., 5%) "theory uncertainty"

➤ Fit not rewarded (much) for improving agreement beyond that point. More freedom in tails.

Also tends to produce $\chi^2_{5\%}$ values ~ unity → better uncertainty bands?

Incompatibilities: MC unable to agree with (some part of) **a given measurement**

Fit reacts by trying to reduce huge differences in bins it shouldn't have been asked to fit in the first place, at cost of everything else.

Choose measurements carefully ~ within domain of applicability of physics model

(+ PROFESSOR now has facility to not penalise χ^2 beyond some max deviation)

Practical Example: Uncertainties on Dark-Matter Annihilation Spectra

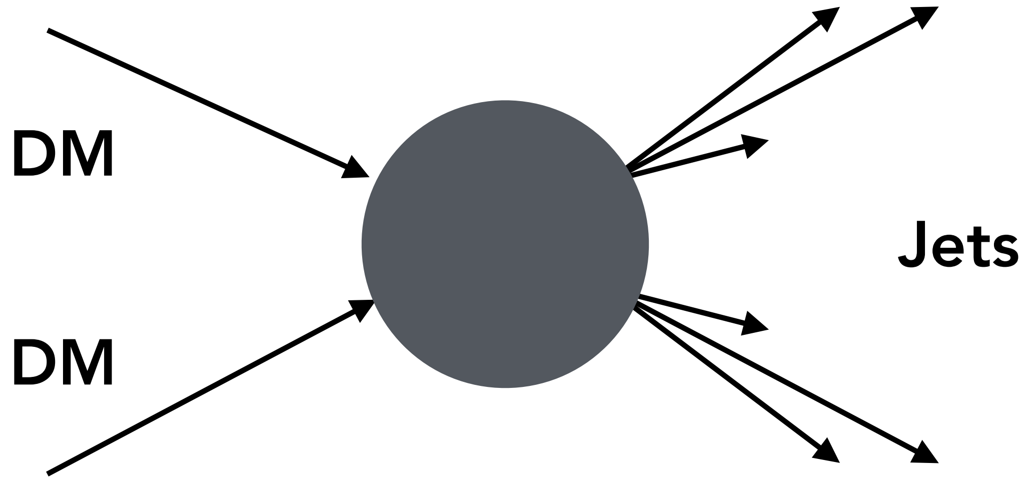
Based on A. Jueid et al., 1812.07424 (gamma rays, *eg for GCE*) and 2202.11546 (antiprotons, *eg for AMS*) + **2303.11363 (all)**

Compare different generators?

E.g., HERWIG – PYTHIA

Problem: tuned to ~ same data

Difference not guaranteed to span genuine uncertainties



Instead, did parametric refittings of LEP data within PYTHIA's modelling

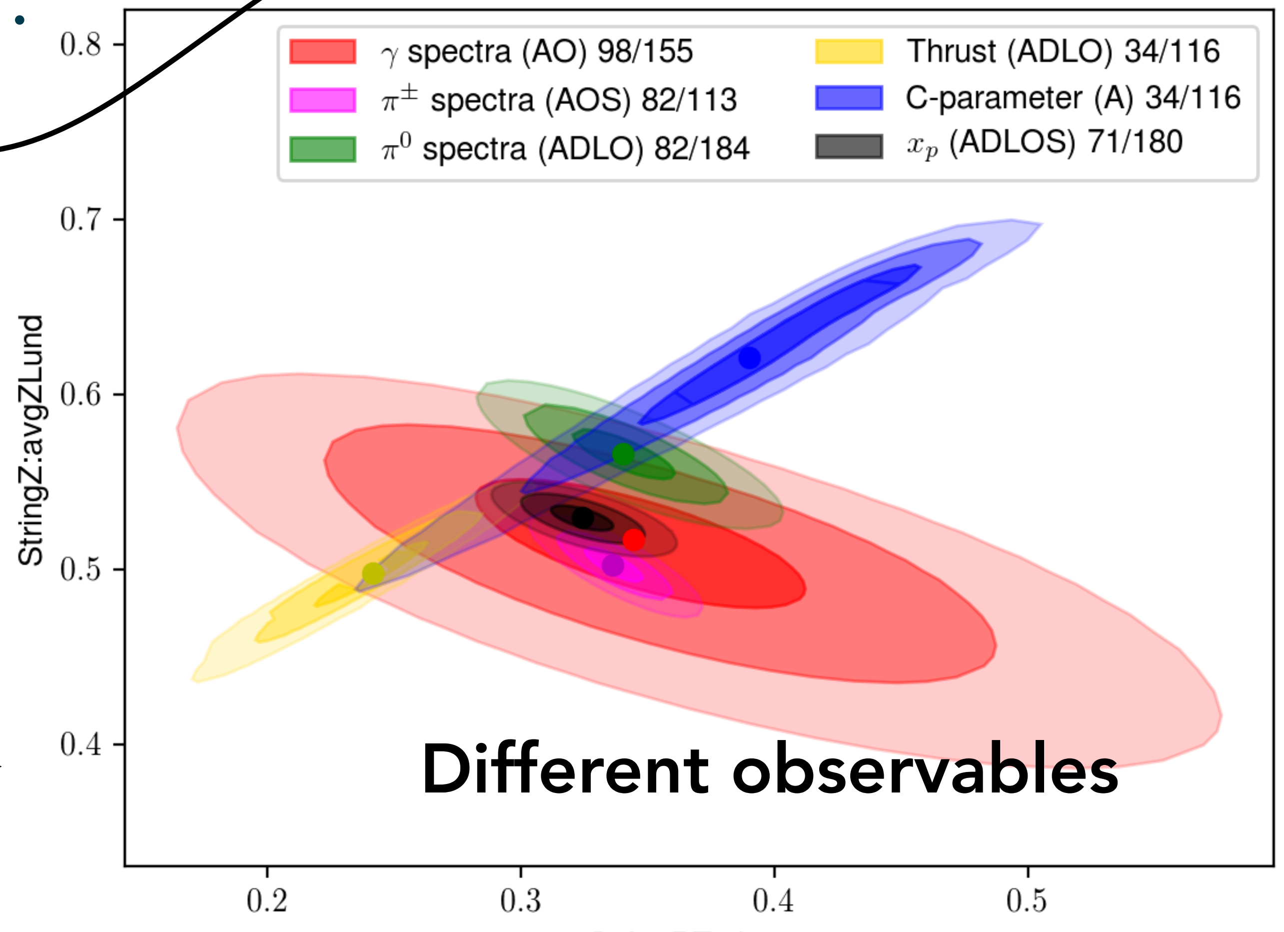
Simple sanity limit / overfit protection / tension resolution:

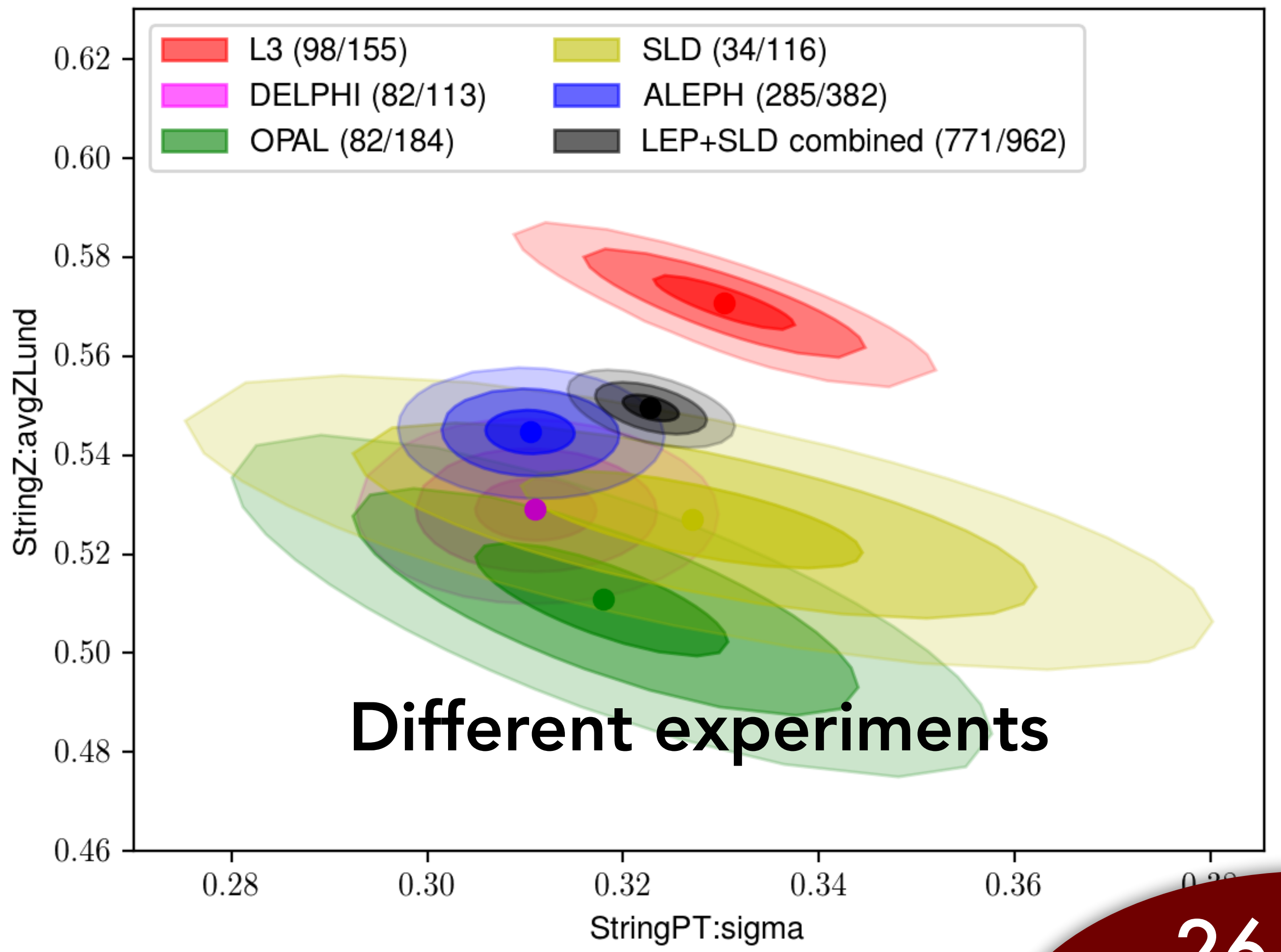
Added blanket 5% baseline uncertainty

(+ excluded superseded measurements)

Parameter	without 5%	with 5%
`StringPT:Sigma`	$0.3151^{+0.0010}_{-0.00010}$	$0.3227^{+0.0028}_{-0.0028}$
`StringZ:aLund`	$1.028^{+0.031}_{-0.031}$	$0.976^{+0.054}_{-0.052}$
`StringZ:avgZLund`	$0.5534^{+0.0010}_{-0.0010}$	$0.5496^{+0.0026}_{-0.0026}$
χ^2/ndf	5169/963	778/963

+ Universality Tests:





6. Such Stuff as Jets are Made Of

Particle Composition in PYTHIA — Baryons & Strangeness

Confinement — in PP Collisions

High-energy pp collisions — with ISR and Multi-Parton Interactions

Final states with **very many** coloured partons

With significant overlaps in phase space

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**:

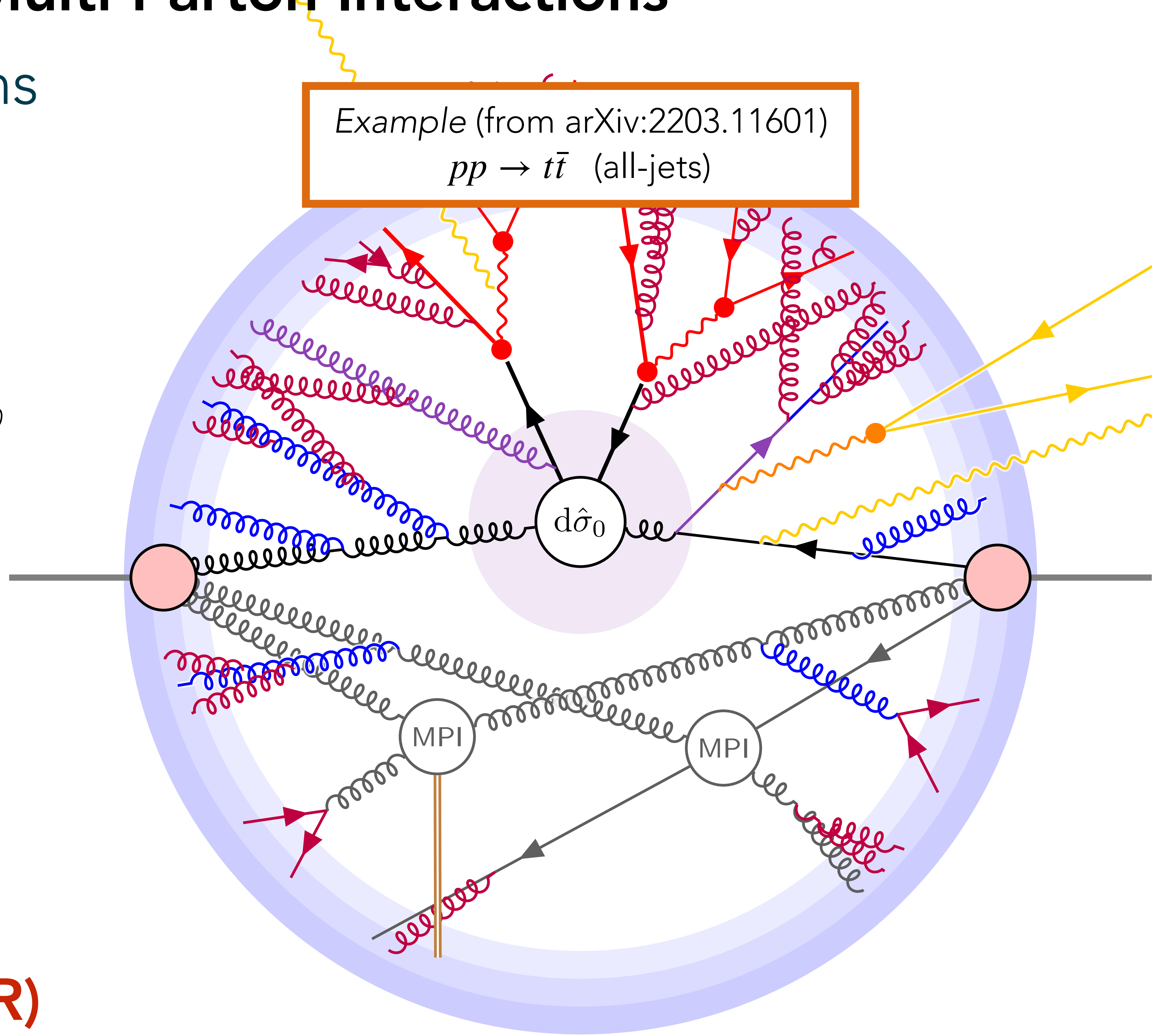
$$3 \otimes \bar{3} = 8 \oplus 1$$

$$3 \otimes 3 = 6 \oplus \bar{3} \quad ; \quad 3 \otimes 8 = 15 \oplus 6 \oplus 3$$

$$8 \otimes 8 = 27 \oplus 10 \oplus \overline{10} \oplus 8_S \oplus 8_A \oplus 1$$

Many charges ➜ Colour Reconnections* (CR)

more likely than not — "Colour Promiscuity!" [J. Huston]



Example (from arXiv:2203.11601)
$pp \to t\bar{t}$ (all-jets)

"Parton Level"
(Event structure before confinement)

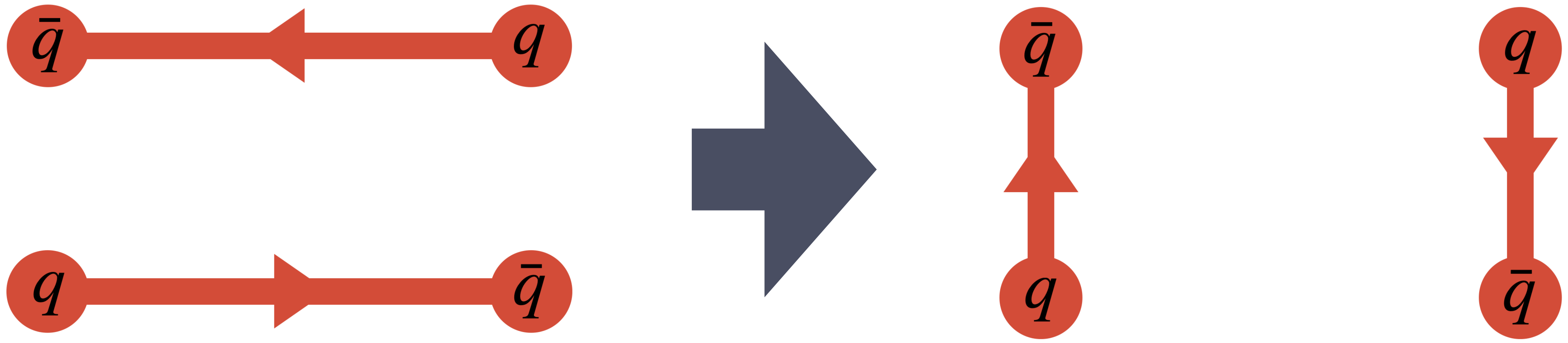
**): in this context, QCD CR simply refers to an ambiguity beyond Leading N_C, known to exist.*
Note the term "CR" can also be used more broadly to incorporate further physics concepts.

QCD Colour Reconnections ⟷ String Junctions

Stochastically restores colour-space ambiguities according to **SU(3) algebra**

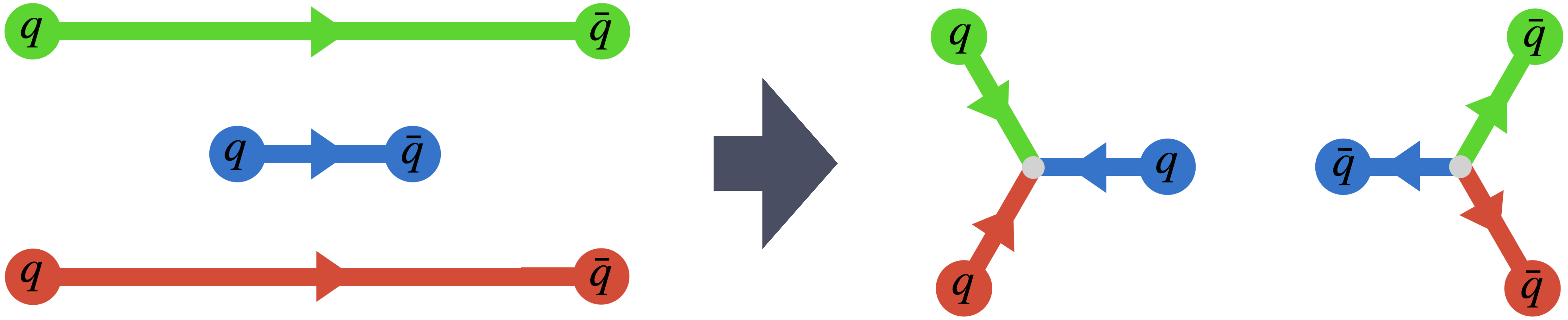
[Christiansen & **PS**
JHEP 08 (2015) 003]

- ➢ Allows for reconnections to minimise string lengths



Dipole-type reconnection

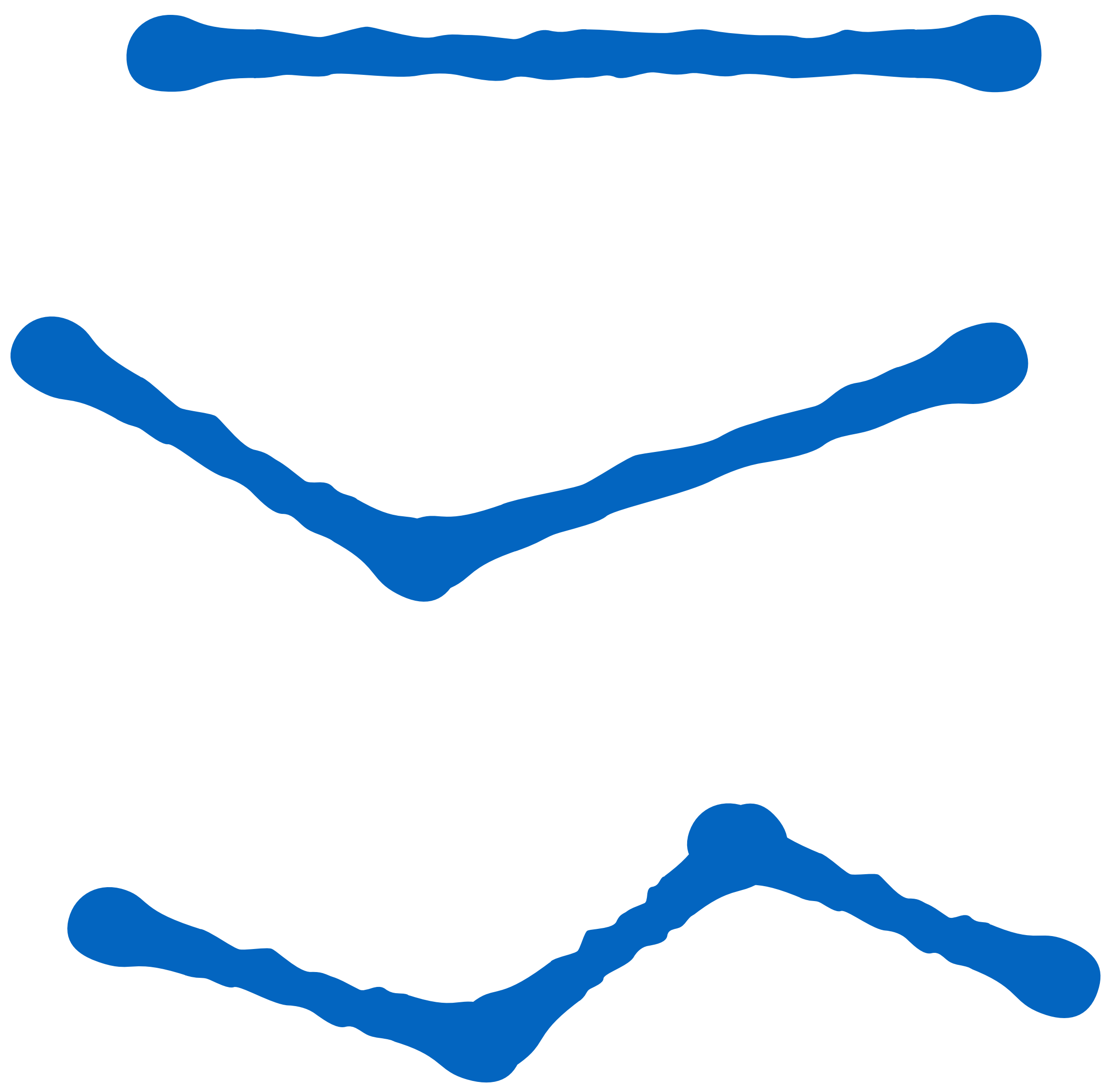
What about the **red**-**green**-**blue** colour singlet state?



Junctions!

(Types of String Topologies)

Open Strings

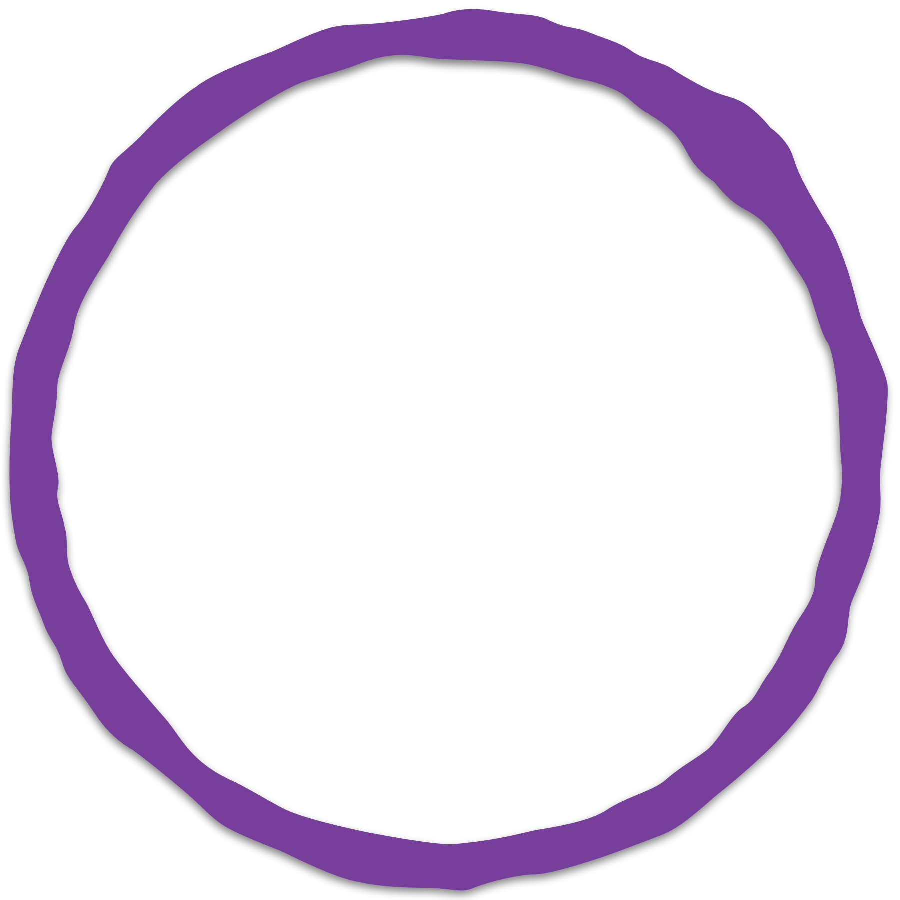


$q\bar{q}$ strings (with gluon kinks)

E.g., $Z \to q\bar{q}$ + shower

$H \to b\bar{b}$ + shower

Closed Strings



Gluon rings

E.g., $H \to gg$ + shower

$\Upsilon \to ggg$ + shower

SU(3) String Junction

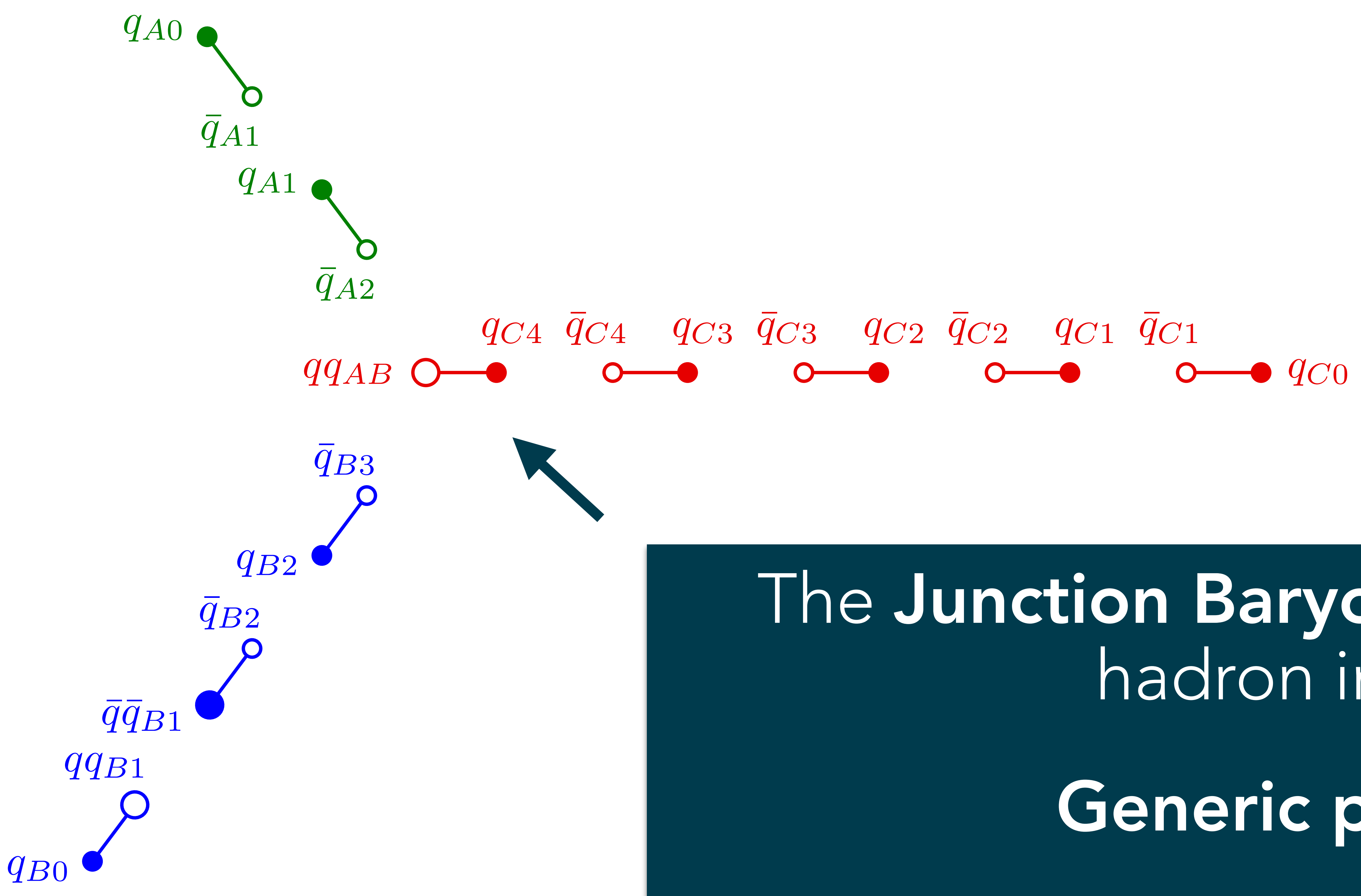
Open strings with $N_C = 3$ endpoints

E.g., Baryon-Number violating neutralino decay $\tilde{\chi}^0 \to qqq$ + shower

What do String Junctions do?

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]
[+ J. Altmann & **PS**, in progress]

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

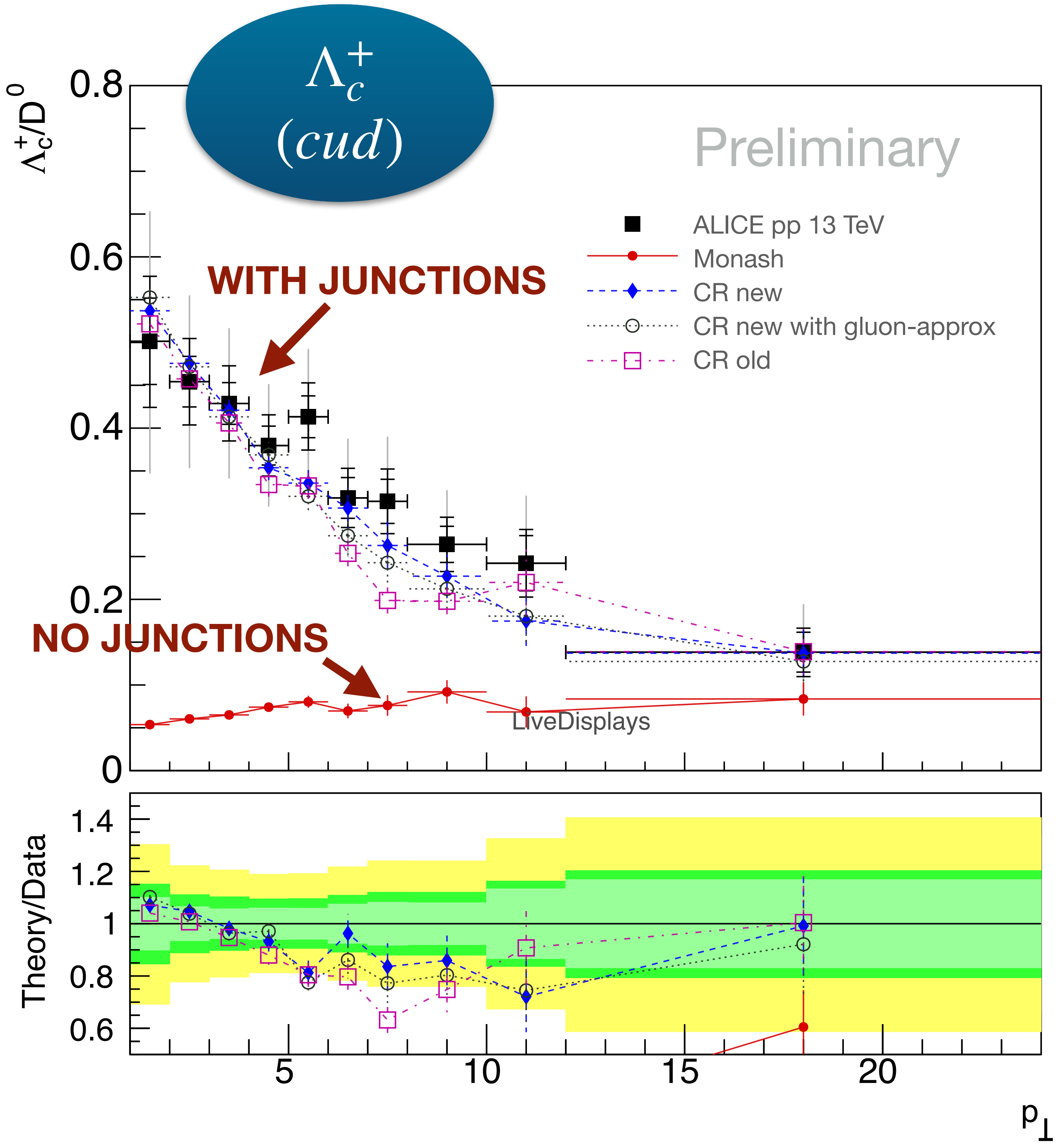
Generic prediction: low p$_T$

A Smoking Gun for String Junctions: Baryon enhancements at low p$_T$

NEW

Confront with Measurements

LHC experiments report very large (factor-10) enhancements in heavy-flavour baryon-to-meson ratios at low p_T!



WITH JUNCTIONS

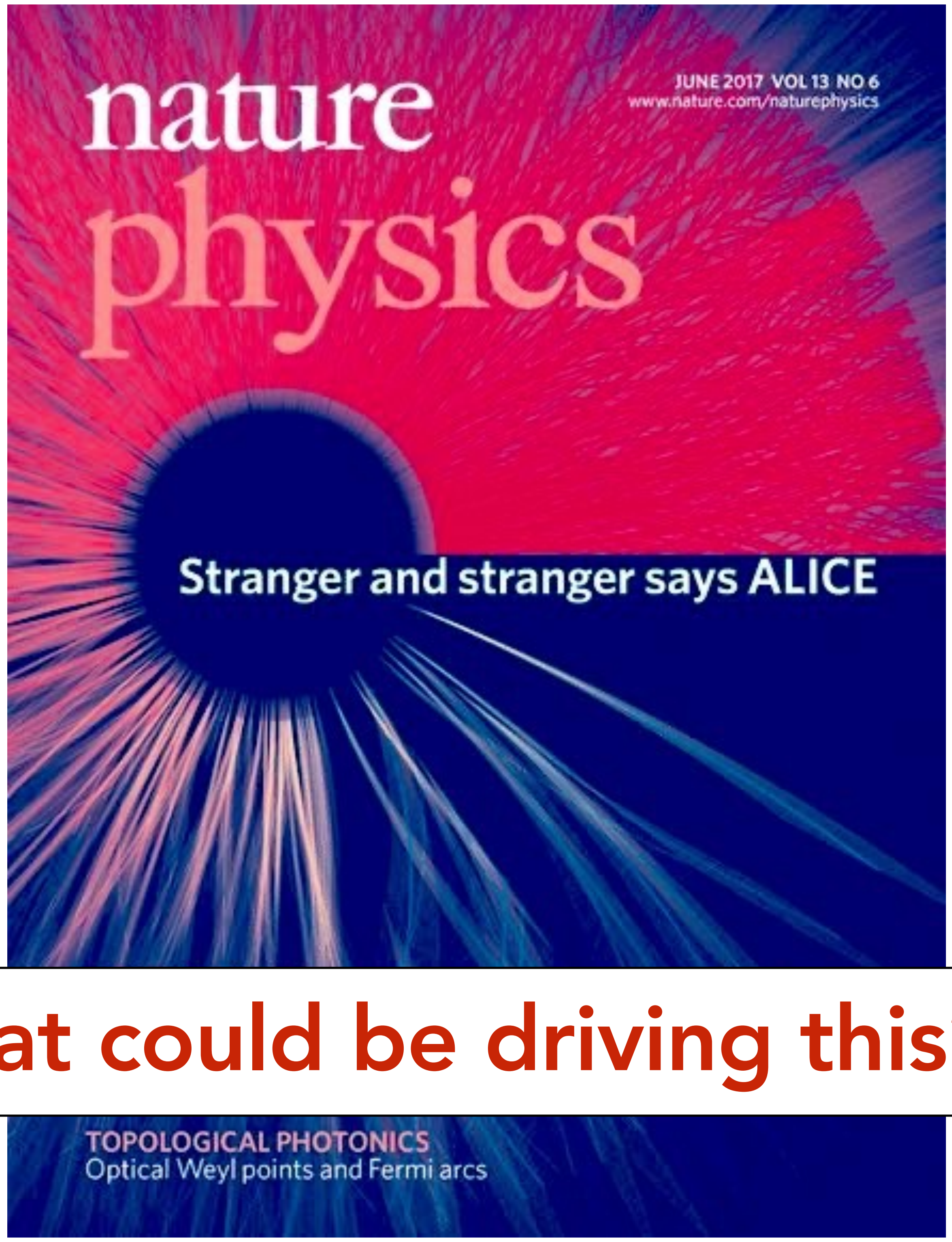
NO JUNCTIONS

[J. Altmann & **PS**, in progress]

Very exciting!

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

We also know ratios of strange hadrons to pions strongly increase with event activity



nature physics

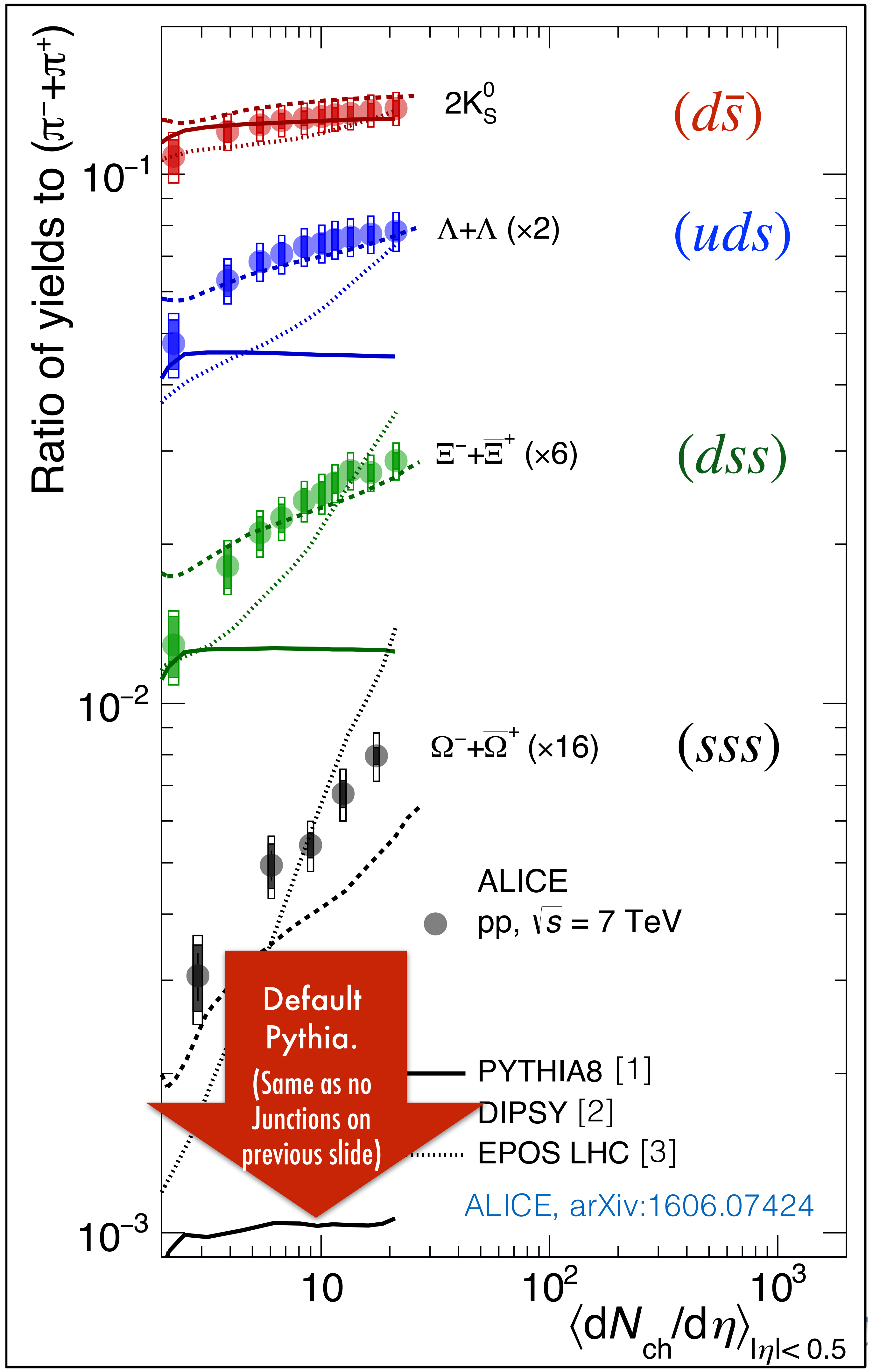
JUNE 2017 VOL 13 NO 6
www.nature.com/naturephysics

Stranger and stranger says ALICE

TOPOLOGICAL PHOTONICS
Optical Weyl points and Fermi arcs

June
2017

What could be driving this?



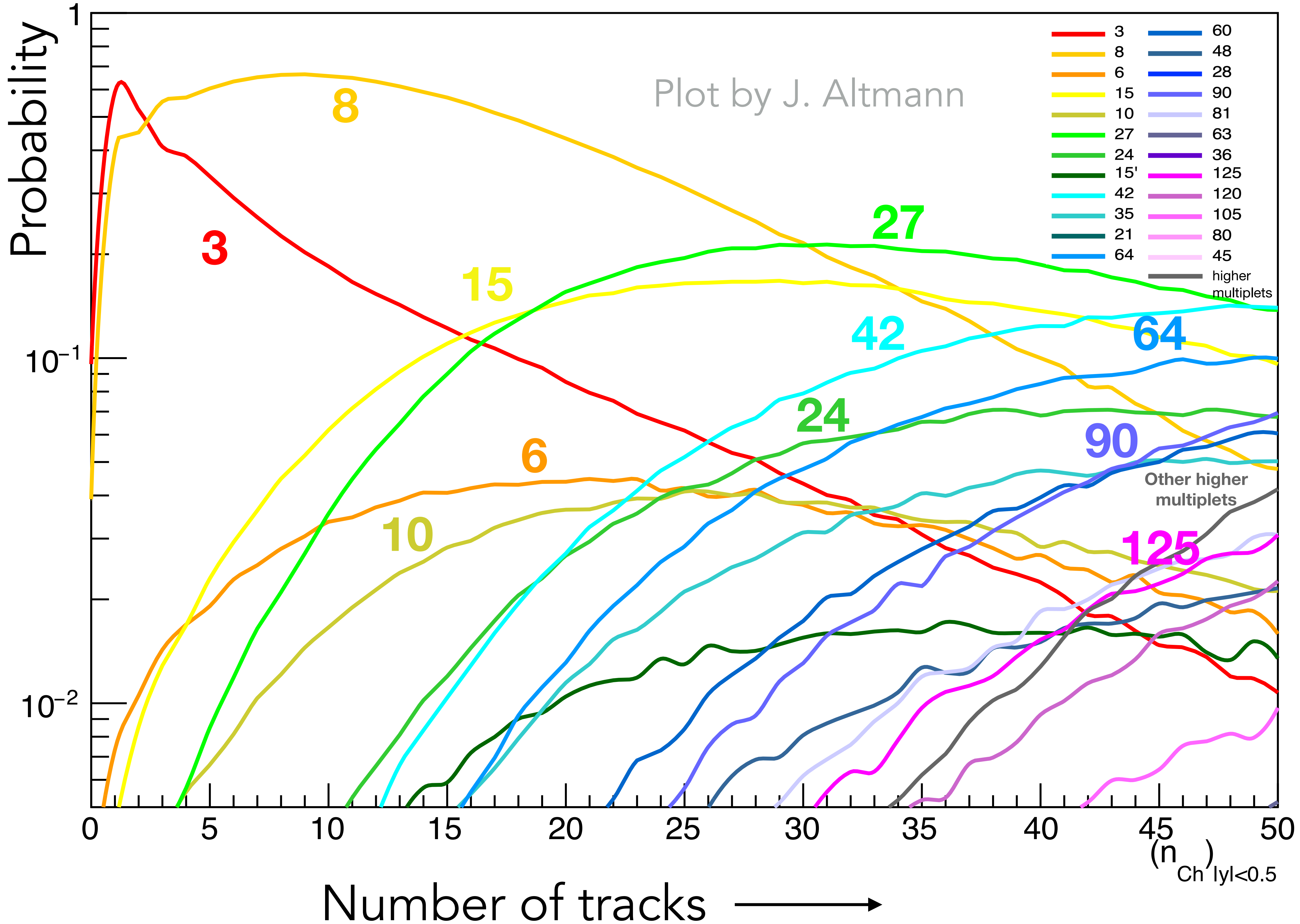
ALICE, arXiv:1606.07424

Default Pythia.
(Same as no Junctions on previous slide)

→ Non-Linear String Dynamics?

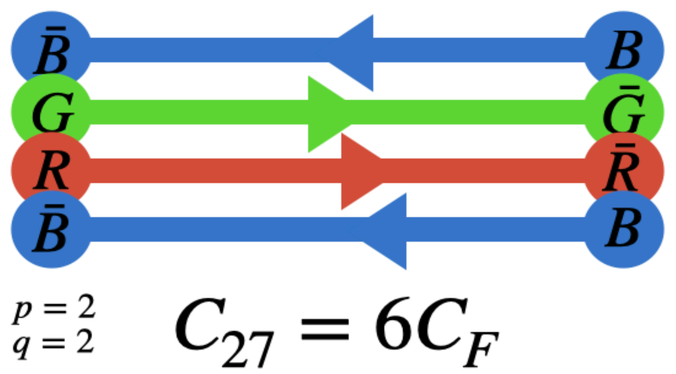
MPI ⟹ lots of coloured partons scattered into the final states

Count **# of flux lines** crossing $y = 0$ in pp collisions (according to PYTHIA):



Confining fields may be reaching much **higher effective representations** than simple quark-antiquark (3) ones.

E.g.: 27



27

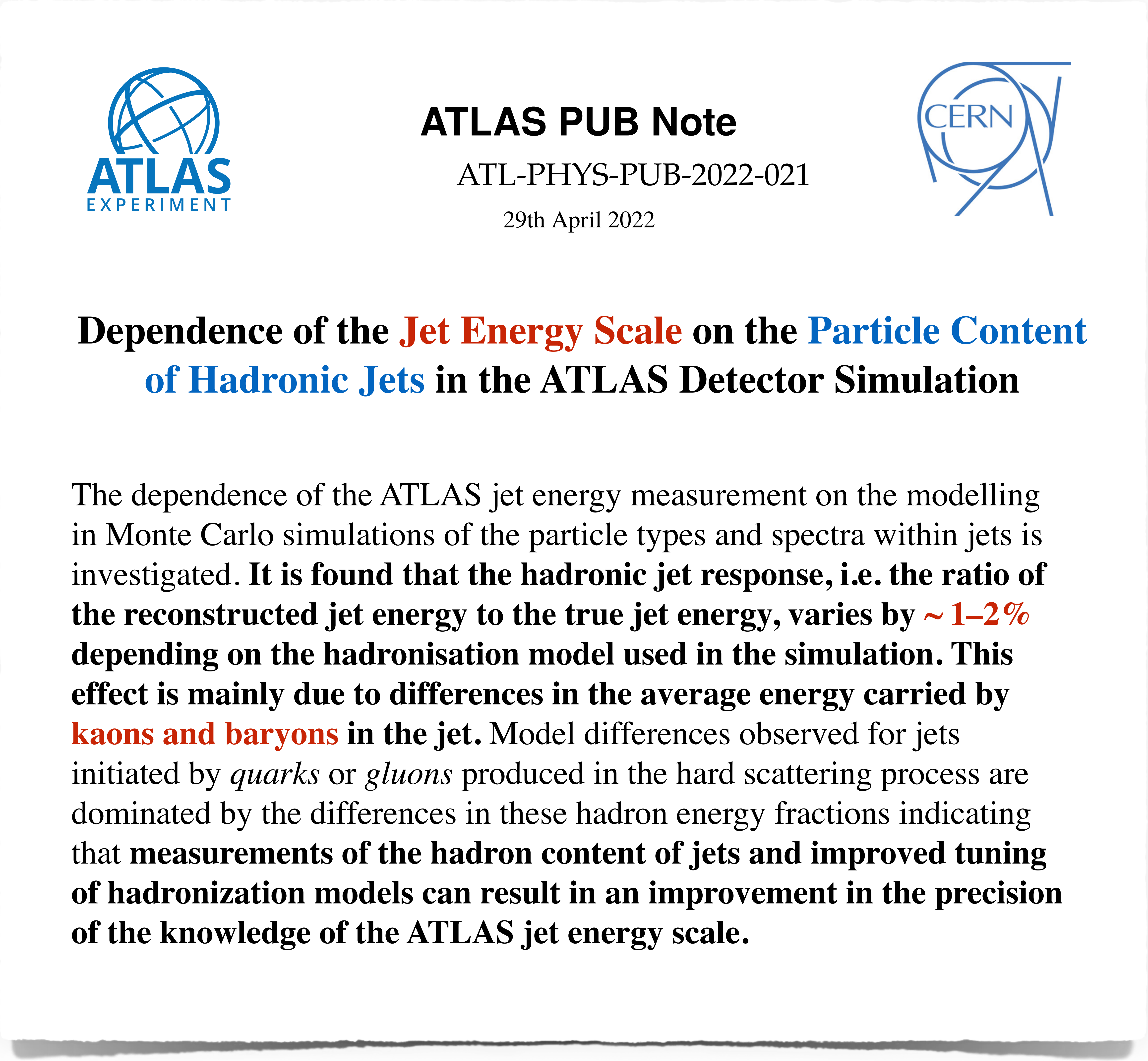
$C_{27} = 6C_F$

Two approaches in PYTHIA:

1) Colour Ropes (Lund)

2) Close-Packing (Monash)

Particle Composition: Impact on Jet Energy Scale



ATLAS EXPERIMENT

ATLAS PUB Note

ATL-PHYS-PUB-2022-021

29th April 2022

CERN

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

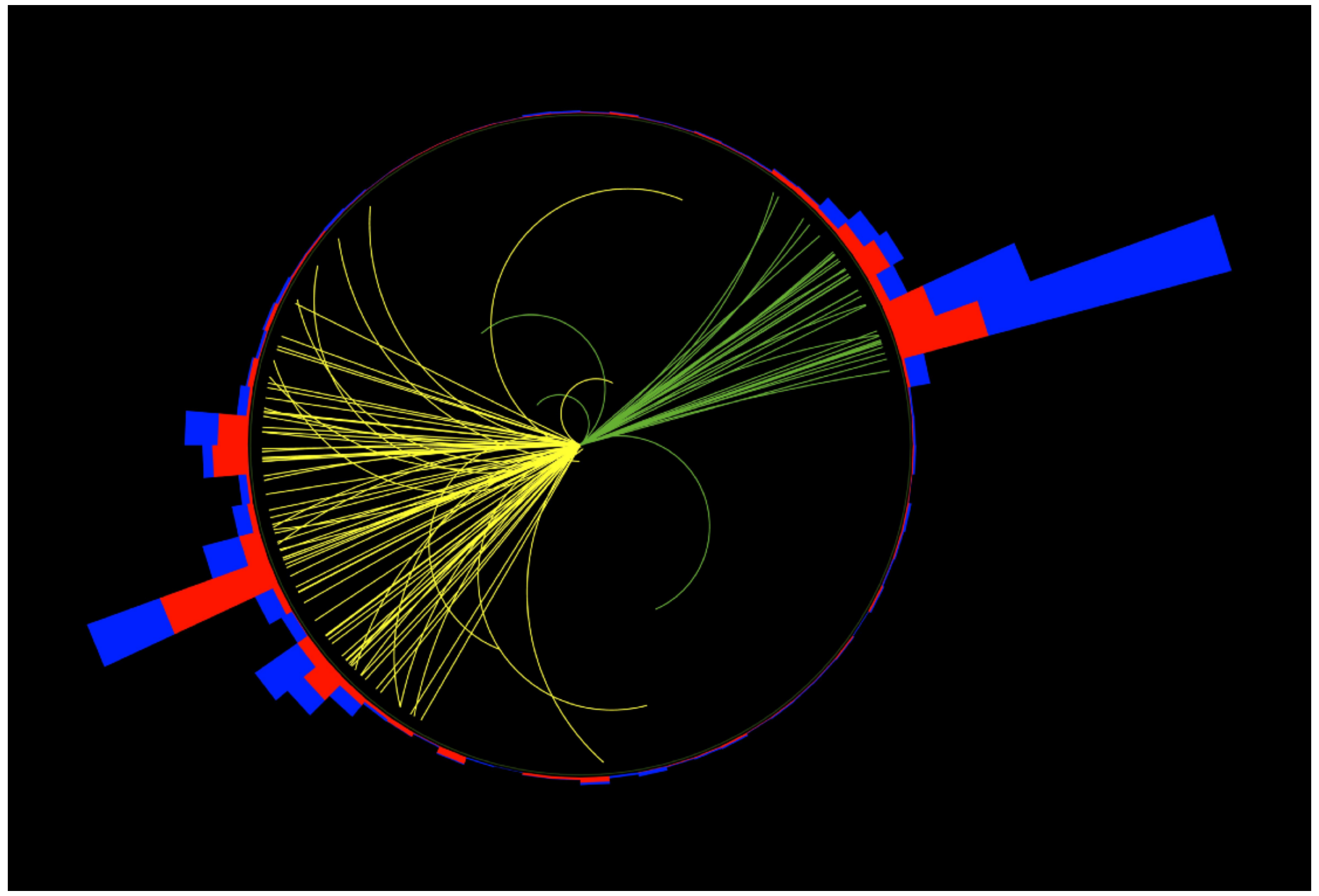
- Reweighting to force similar baryon and kaon fractions
- Max variation ➜ [1.2%, 0.8%, 0.5%]
- Significant potential for improved Jet Energy Scale uncertainties!

Motivates Careful Models & Careful Constraints

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

Summary

MC generators connect theory with experiment

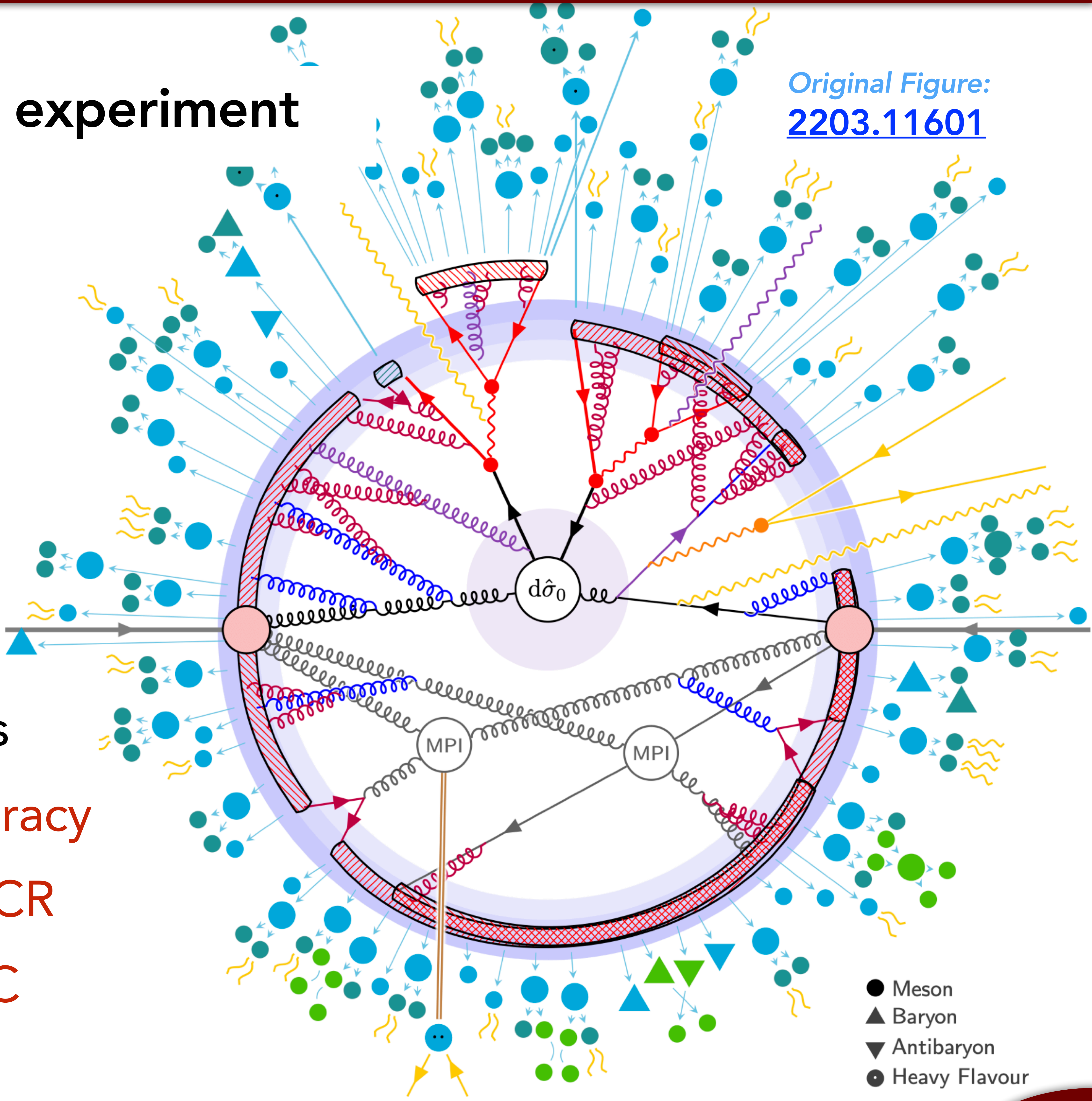


Plan for NNLO+NNLL accurate MCs

➜ era of percent-level perturbative accuracy

+ much new work on hadronization & CR

Driven by **new measurements** at LHC



Original Figure: 2203.11601

Extra Slides

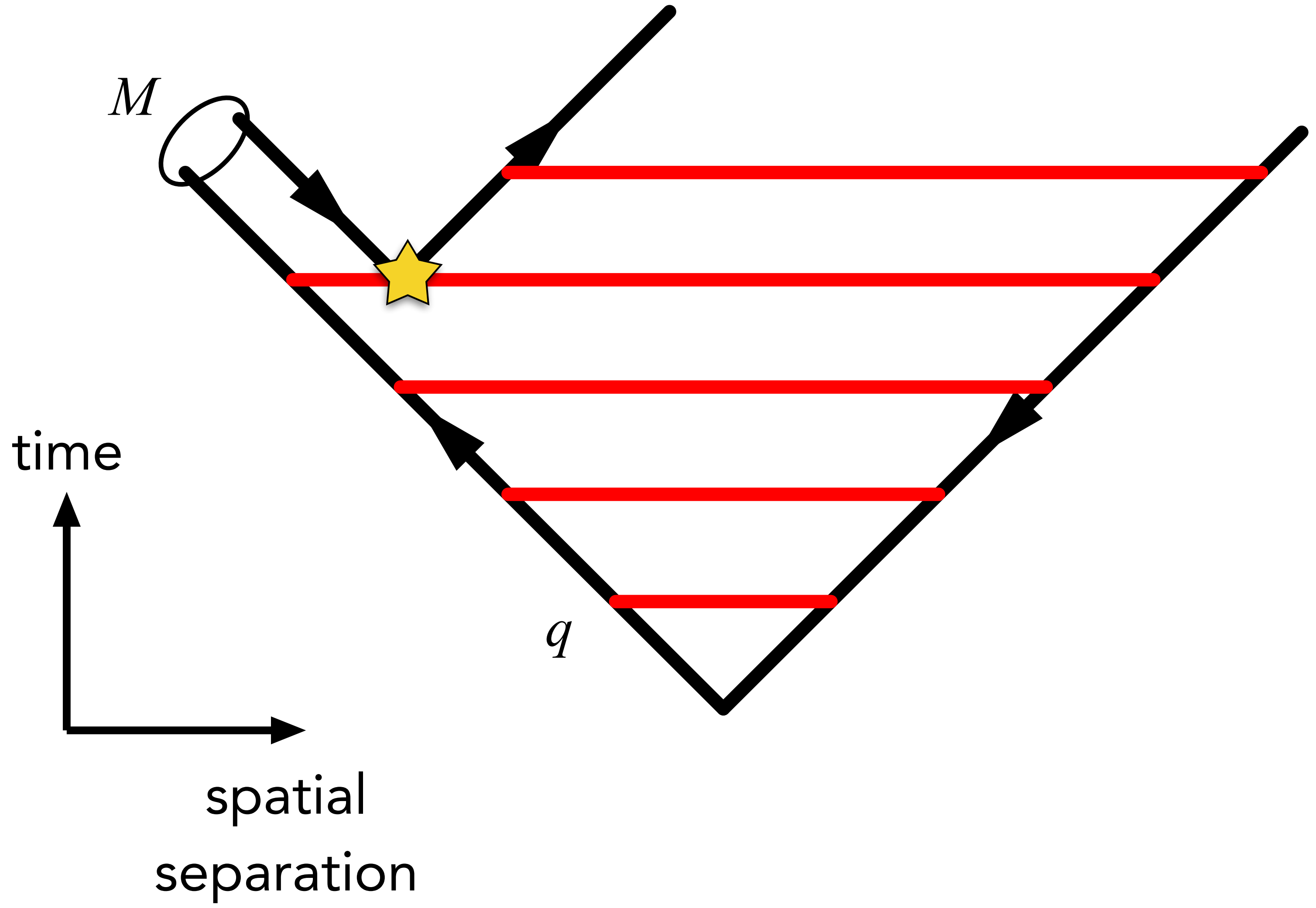
String Breaking

In “unquenched” QCD

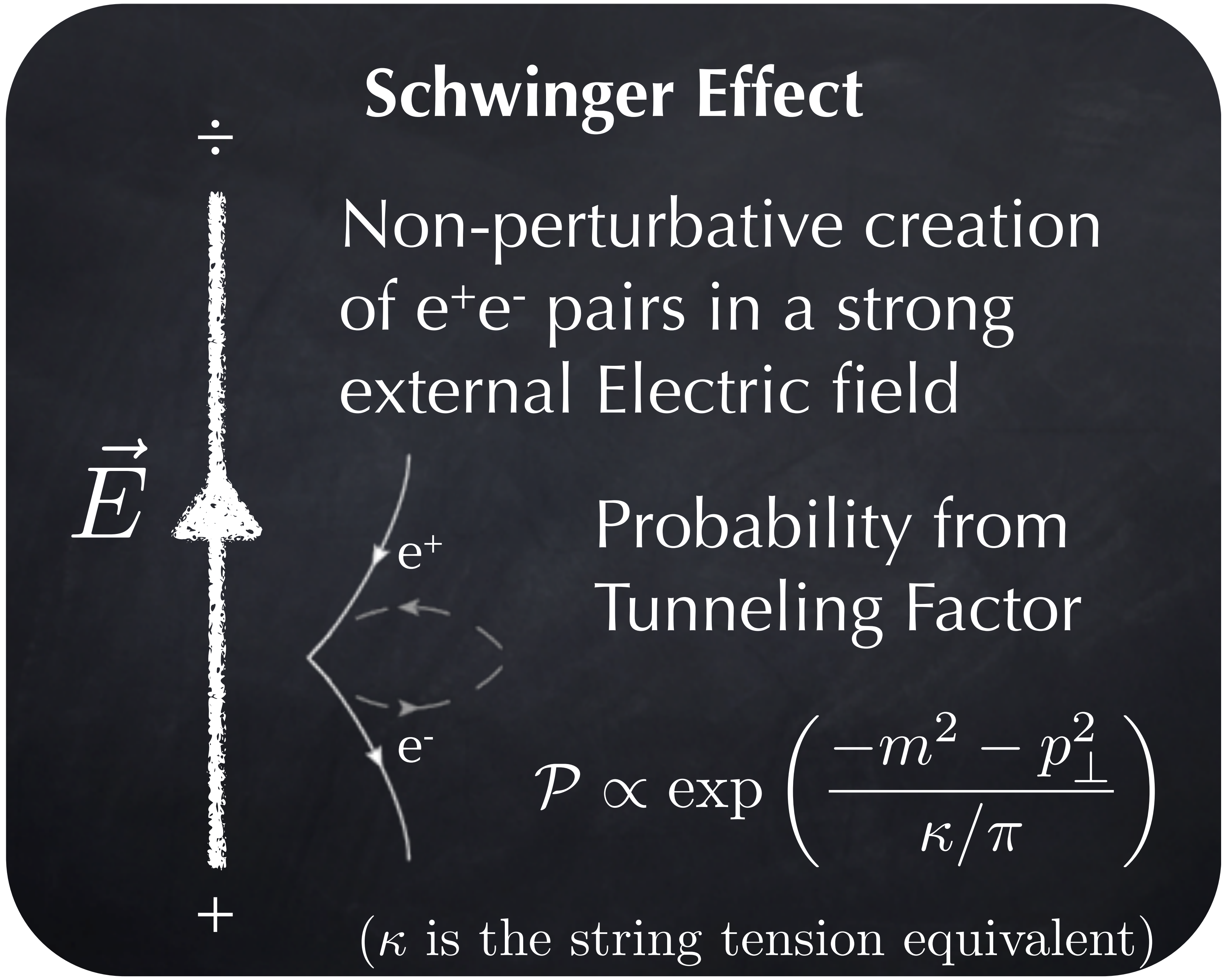
$g \to q\bar{q} \implies$ The strings will “break”

Non-perturbative so can’t use $P_{g\to q\bar{q}}(z)$

Model: Schwinger mechanism



J. Schwinger, Phys. Rev. **82** (1951) 664



Schwinger Effect

Non-perturbative creation of e$^+$e$^-$ pairs in a strong external Electric field

Probability from Tunneling Factor

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

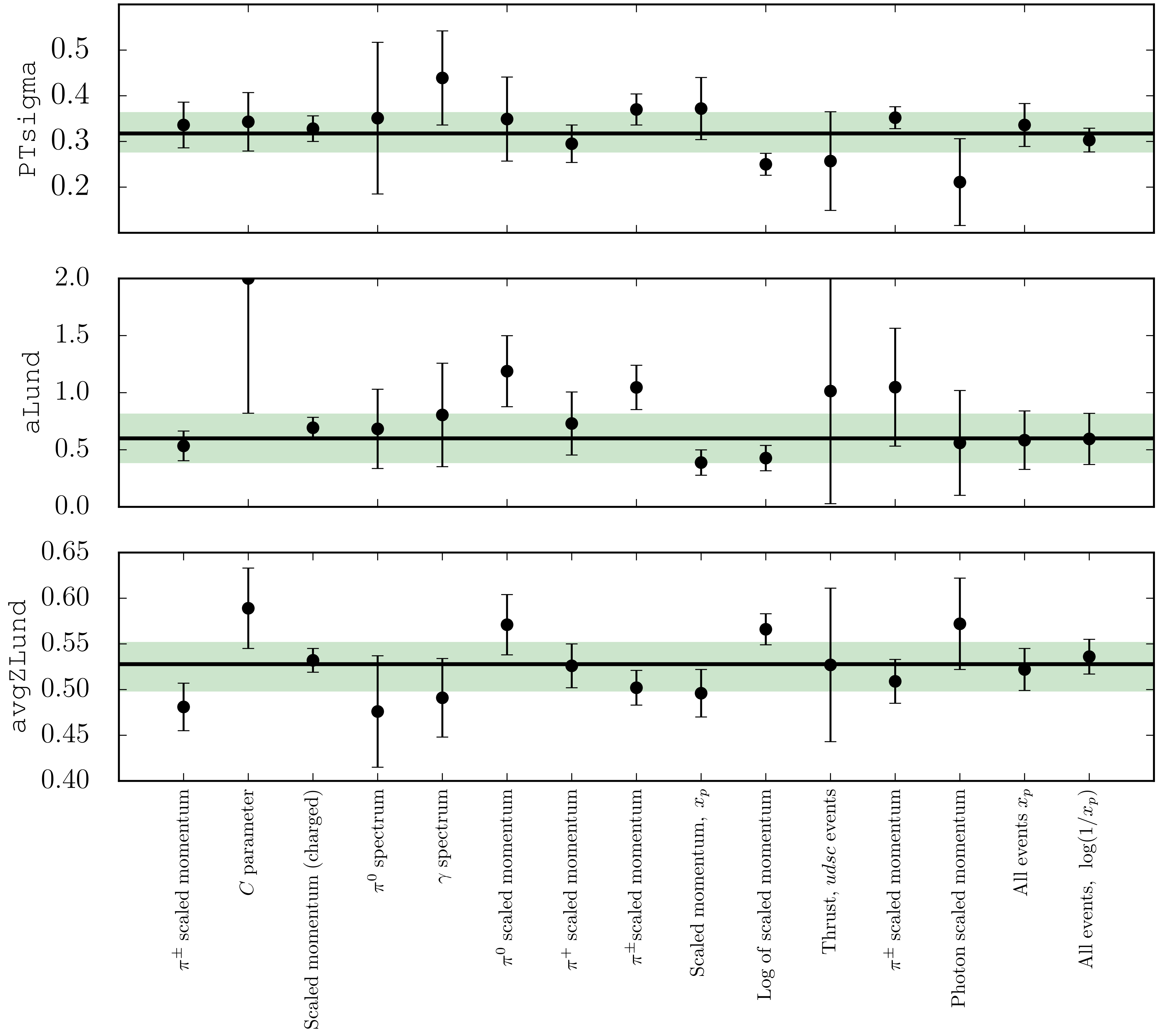
(κ is the string tension equivalent)

$\implies$ Gaussian suppression of high $m_\perp = \sqrt{m_q^2 + p_\perp^2}$

Assume probability of string break constant per unit world-sheet area

Practical Example: Uncertainties on Dark-Matter Annihilation Spectra

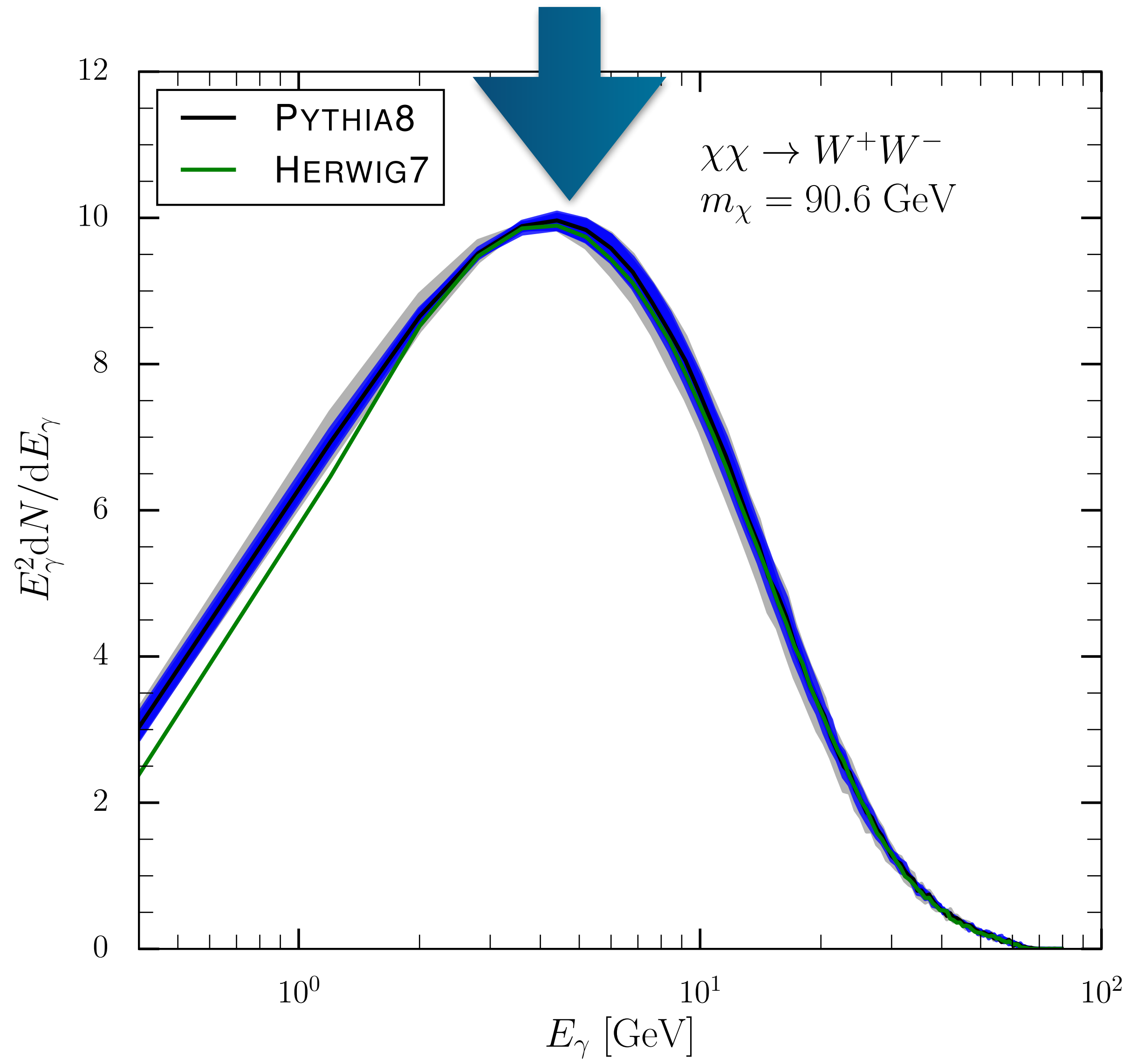
Based on A. Jueid et al., 1812.07424 (gamma rays, *eg for GCE*) and 2202.11546 (antiprotons, *eg for AMS*) + **2303.11363 (all)**



Weighted Average: good consistency across observables

Expensive?

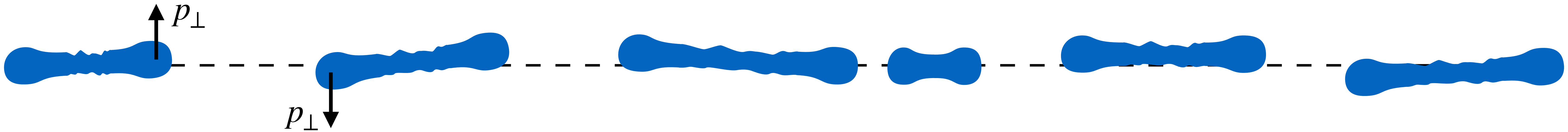
10-point variations ➤ **Fairly convincing** uncertainty bands?



Examples with Pythia 8

[Reweighting MC Predictions & Automated Fragmentation Variations in Pythia 8, 2308.13459]

Transverse Fragmentation Function (Gaussian)



$$\frac{1}{2\pi\sigma_{p_T}^2}\exp\left(-\frac{(\Delta p_x)^2+(\Delta p_y)^2}{2\sigma_{p_T}^2}\right)$$

`StringPT:sigma`

variations

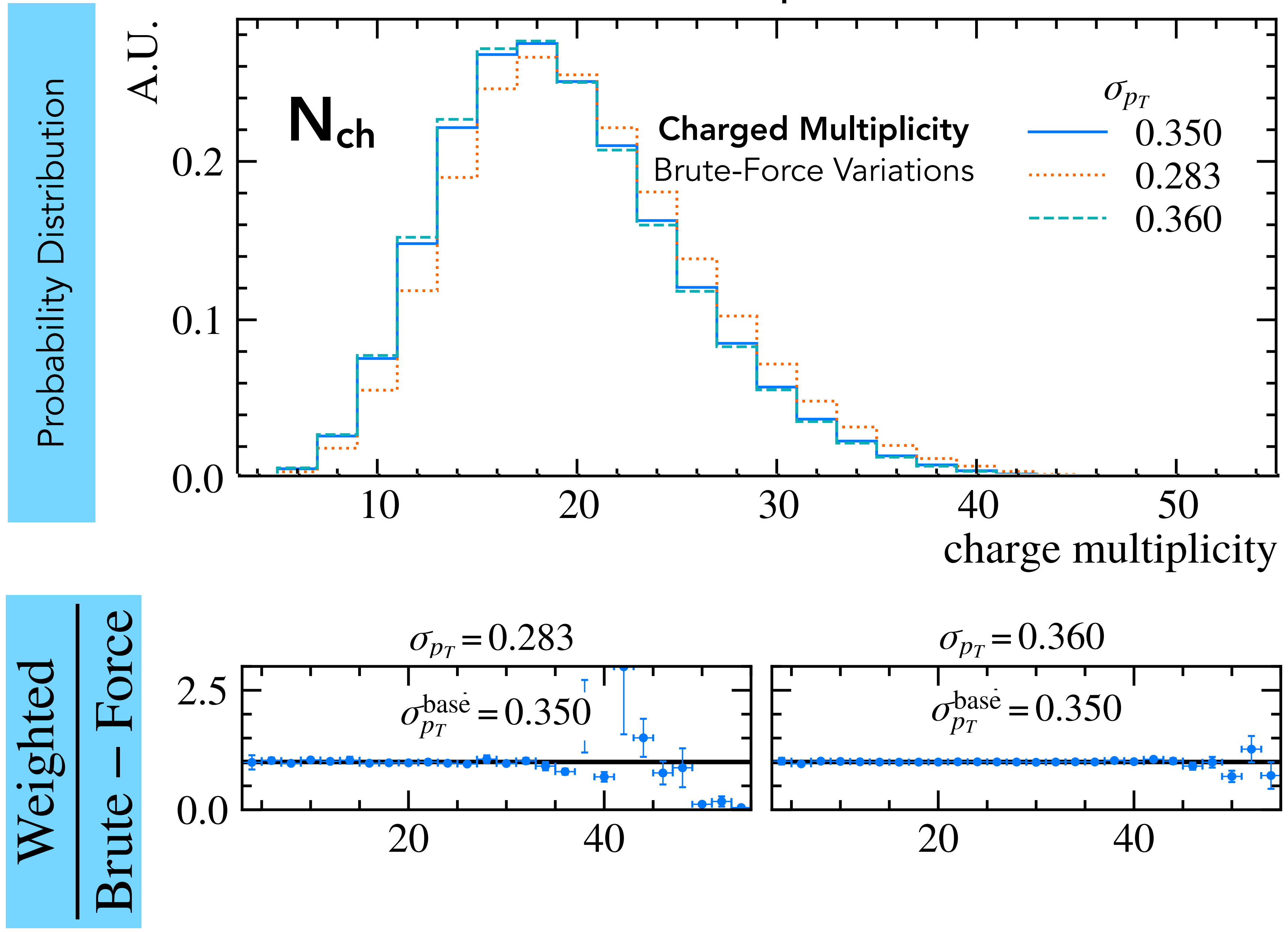
Reweighting Methodology:

For each p_T (Box-Muller transform):

$$w' = \frac{\sigma^2}{\sigma'^2}\exp\left(-\kappa\left(\frac{\sigma^2}{\sigma'^2}-1\right)\right)$$

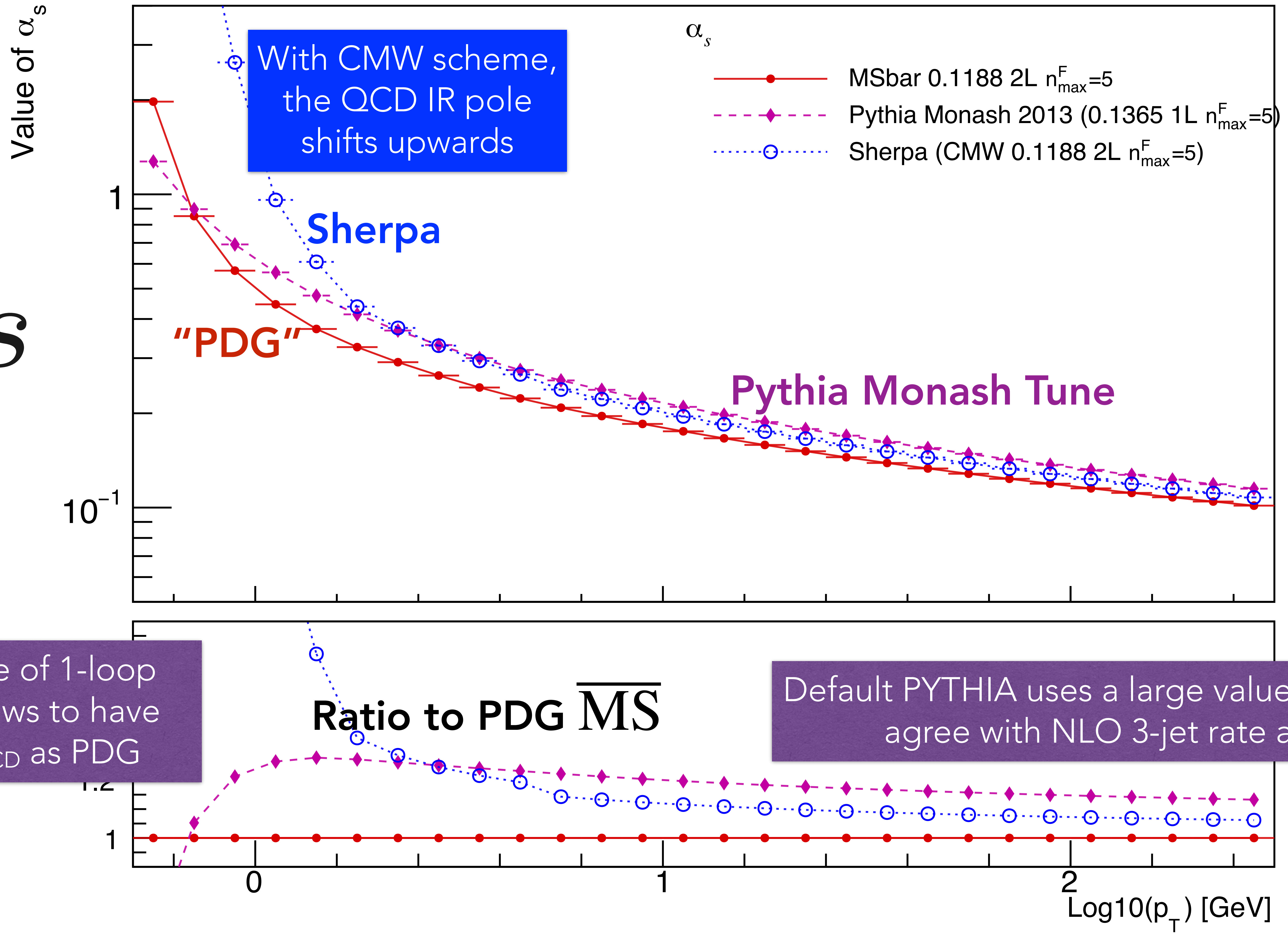
$\kappa = (n_1^2+n_2^2)/2$ and n_i are normally distributed random variates

Example



Note on Different alpha(S) Choices

α_s



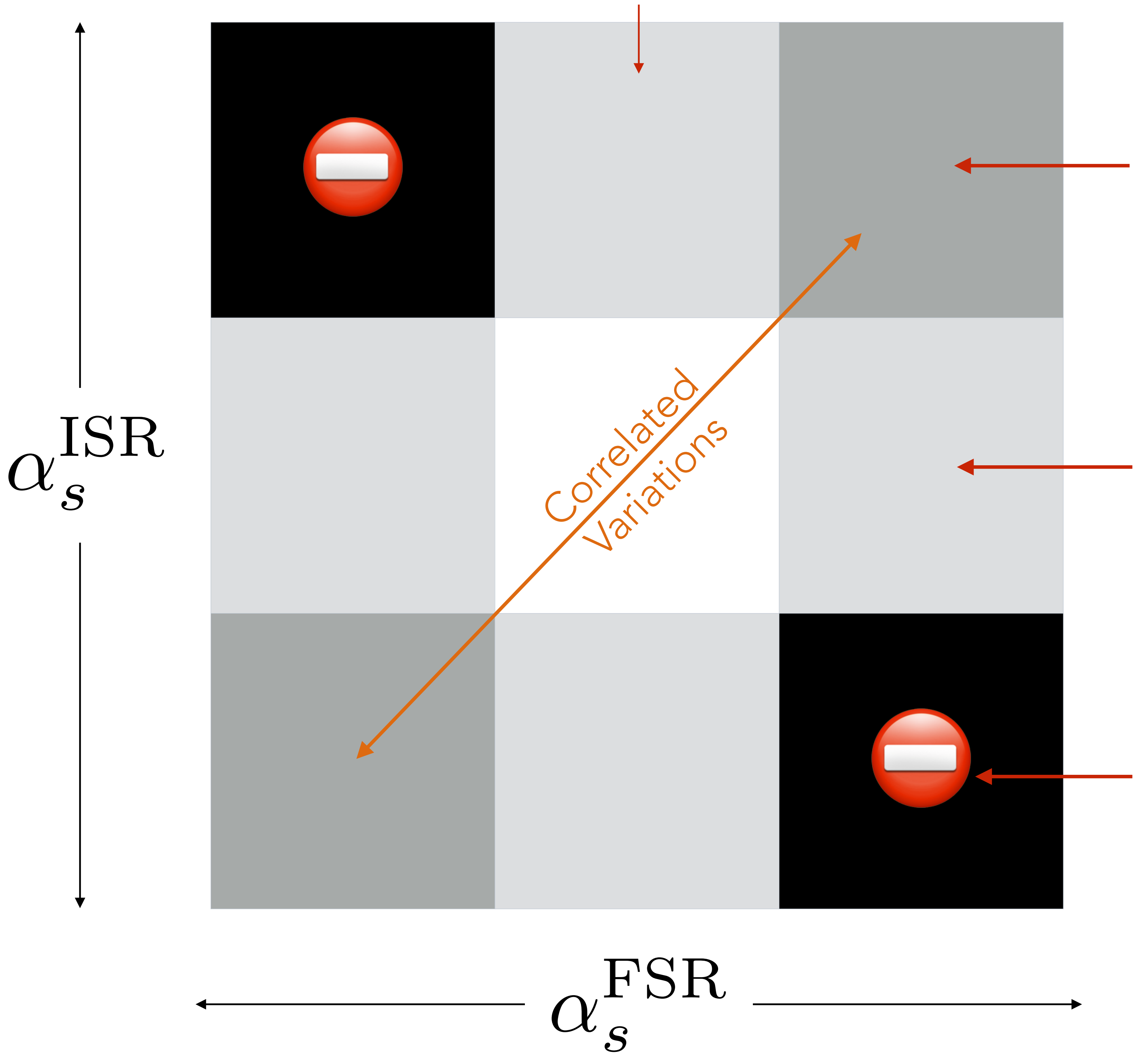
With CMW scheme, the QCD IR pole shifts upwards

Slower pace of 1-loop running allows to have similar Λ_{QCD} as PDG

Default PYTHIA uses a large value of $\alpha_s(M_Z)$ to agree with NLO 3-jet rate at LEP

Correlated or Uncorrelated?

What I would do: **7-point variation** (resources permitting → use the automated bands?)



Increasing only ISR

- More H_T and N_{jets}; similar core jet shapes

Increasing both ISR and FSR

- More H_T in the events.
- More OOC loss (from FSR) but also more H_T and more hard ISR jet seeds → partial cancellation in N_{jets}?

Increasing only FSR

- More OOC loss (FSR jet broadening), acting on similar number of seed partons (no increase in ISR).
- Similar H_T

Increasing FSR, Decreasing ISR -> Exclude?

- Double counting? Fewer ISR partons, **and** more smearing of those that remain. *(Easy to rule out?)*
- Also from theoretical/mathematical point of view, the artificially induced discrepancy is now proportional to ln(16) = 2.8 instead of ln(4) = 1.4.

Scale Variations: How Big?

Scale variations induce 'artificial' terms beyond truncated order in QFT ~ Allow the calculation to float by (1+O(α_s)).

$$\frac{\alpha_s(k_1^2\mu^2)}{\alpha_s(k_2^2\mu^2)} \sim 1 - \underset{\uparrow}{b_0}\,\underline{\ln(k_1^2/k_2^2)}\,\alpha_s(\mu^2) \longleftarrow$$

Flavour-dependent slope of order 1
$b_0 \sim 0.65 \pm 0.07$

Proportionality to $\alpha_s(\mu)$ ⟹ can get a (misleadingly?) small band if you choose central μ scale very large.

E.g., some calculations use $\mu \sim H_T$ ~ largest scale in event ?!

Worth keeping in mind when considering (uncertainty on) central μ choice

Expansion around μ only sensible if this stays ≲ 1

Mainstream view:

Regard scale dependence as unphysical / leftover artefact of our mathematical procedure to perform the calculations.

Dependence on it has to vanish in the 'ultimate solution' to QFT

→ Terms beyond calculated orders must sum up to at least kill μ dependence

Such variations are thus regarded as a useful indication of the size of uncalculated terms.
(Strictly speaking, only a lower bound!)

Typical choice **(in fixed-order calculations):** k ~ [0.5,1,2]

Note: In PYTHIA you specify k^2

TimeShower:renormMultFac

SpaceShower:renormMultFac

Scale Variations: How big?

What do parton showers do?

In principle, LO shower kernels proportional to α_s

Naively: do the analogous factor-2 variations of μ_{PS}.

There are at least 3 reasons this could be **too** conservative

1. **For soft gluon emissions,** we know what the NLO term is

   → even if you do not use explicit NLO kernels, you are effectively NLO (in the soft gluon limit) **if** you are coherent and use $\mu_{PS} = (k_{CMW}\, p_T)$, with 2-loop running and $k_{CMW} \sim 0.65$ (somewhat n_f-dependent). *[Though there are many ways to skin that cat; see next slides.]*

   Ignoring this, a **brute-force** scale variation **destroys** the NLO-level agreement.
2. Although hard to quantify, showers typically achieve better-than-LL accuracy by accounting for **further physical effects** like (E,p) conservation
3. We see empirically that (well-tuned) showers tend to stay inside the envelope spanned by factor-2 variations in **comparison to data**

Scale variations: How Big?

Poor man's recipe: Use $\sqrt{2}$ instead?

Sure … but still somewhat arbitrary

Instead: add compensation term to preserve soft-gluon limit at $O(\alpha_s^2)$

Still allowing full factor-2 outside that limit.

Pythia includes such a compensation term, at least in context of automated uncertainty bands

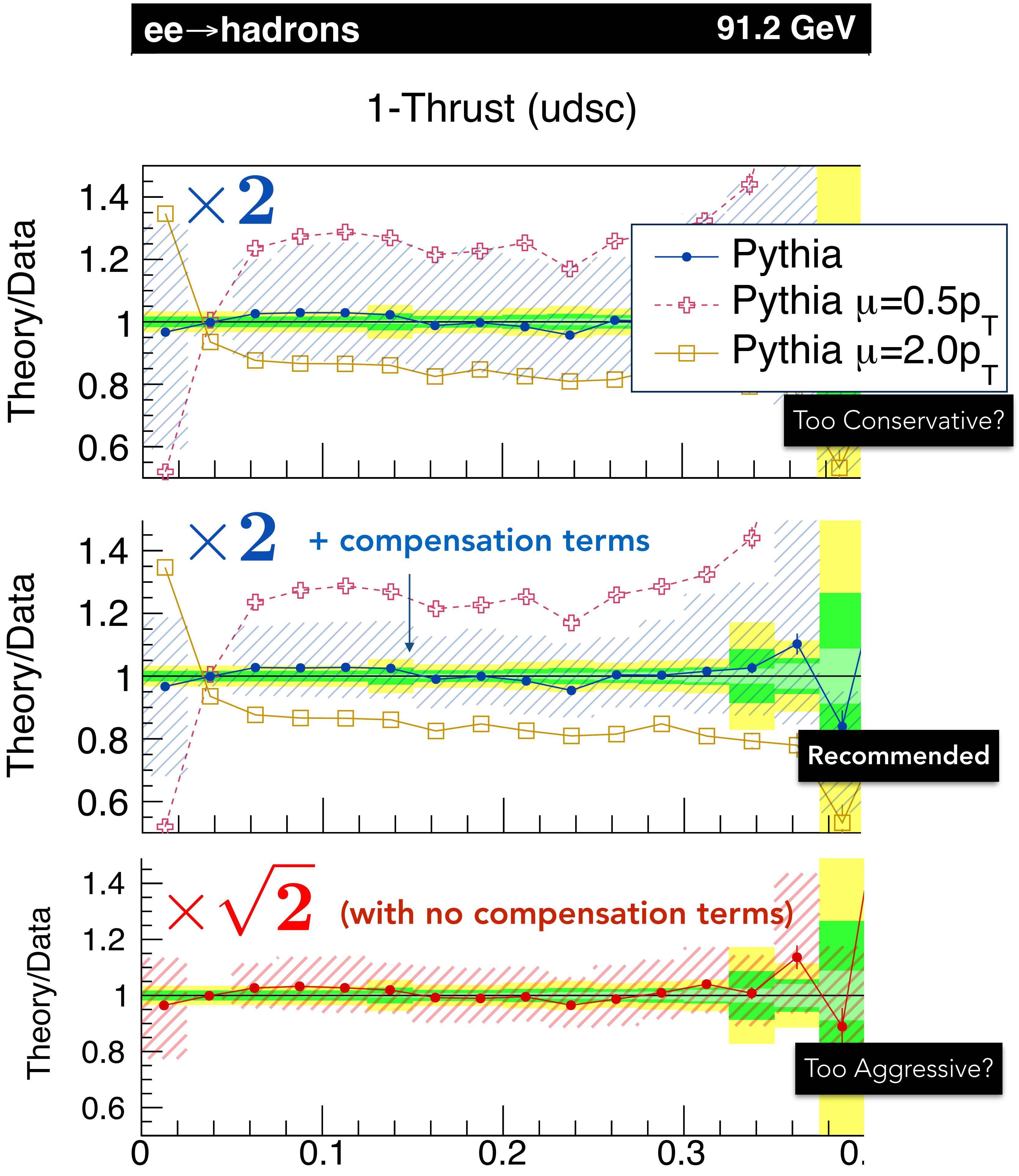
Since aggressive definitions can lead to overcompensation / **extremely** optimistic predictions → *very* small uncertainty bands, we chose a rather conservative definition for PYTHIA → larger bands.

$$P'(t,z) = \frac{\alpha_s(kp_\perp)}{2\pi}\left(1+(1-\zeta)\frac{\alpha_s(\mu_{\mathrm{max}})}{2\pi}\beta_0\ln k\right)\frac{P(z)}{t}$$

Kills the compensation outside the soft limit

Small absolute size of compensation

$$\zeta = \begin{cases} z & \text{for splittings with a } 1/z \text{ singularity} \\ 1-z & \text{for splittings with a } 1/(1-z) \text{ singularity} \\ \min(z,1-z) & \text{for splittings with a } 1/(z(1-z)) \text{ singularity} \end{cases}$$

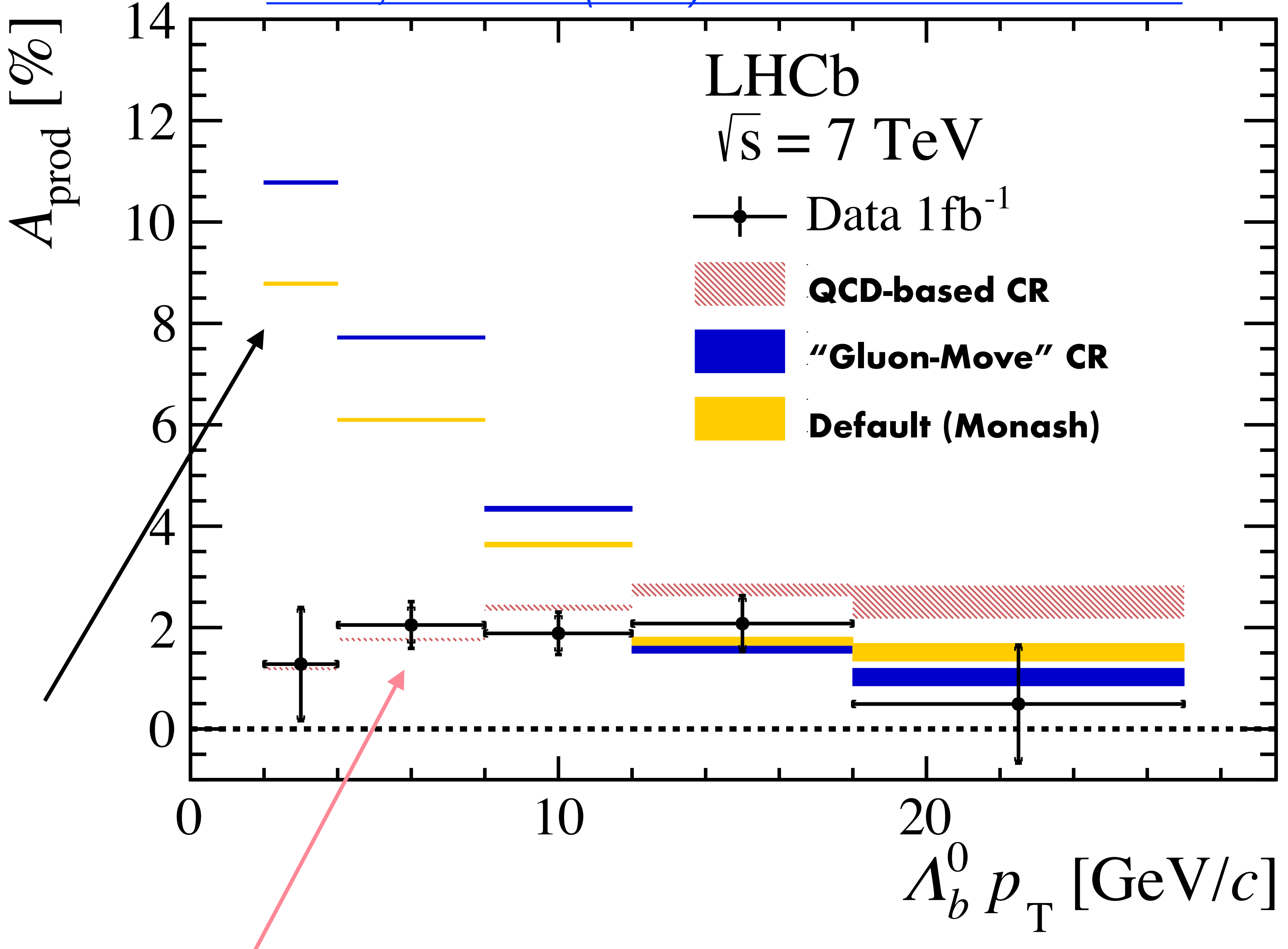


S. Mrenna & PS: PRD94(2016)074005; arXiv:1605.08352

LHCb: also in Bottom

Λ_b asymmetry

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



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

Without junction CR, an important source of low-p$_\mathrm{T}$ 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$_\mathrm{T}$ junction Λ_b and $\bar{\Lambda}_b$, in equal amounts. Dilutes asymmetry!

(Illustration of the “Magic Trick”)

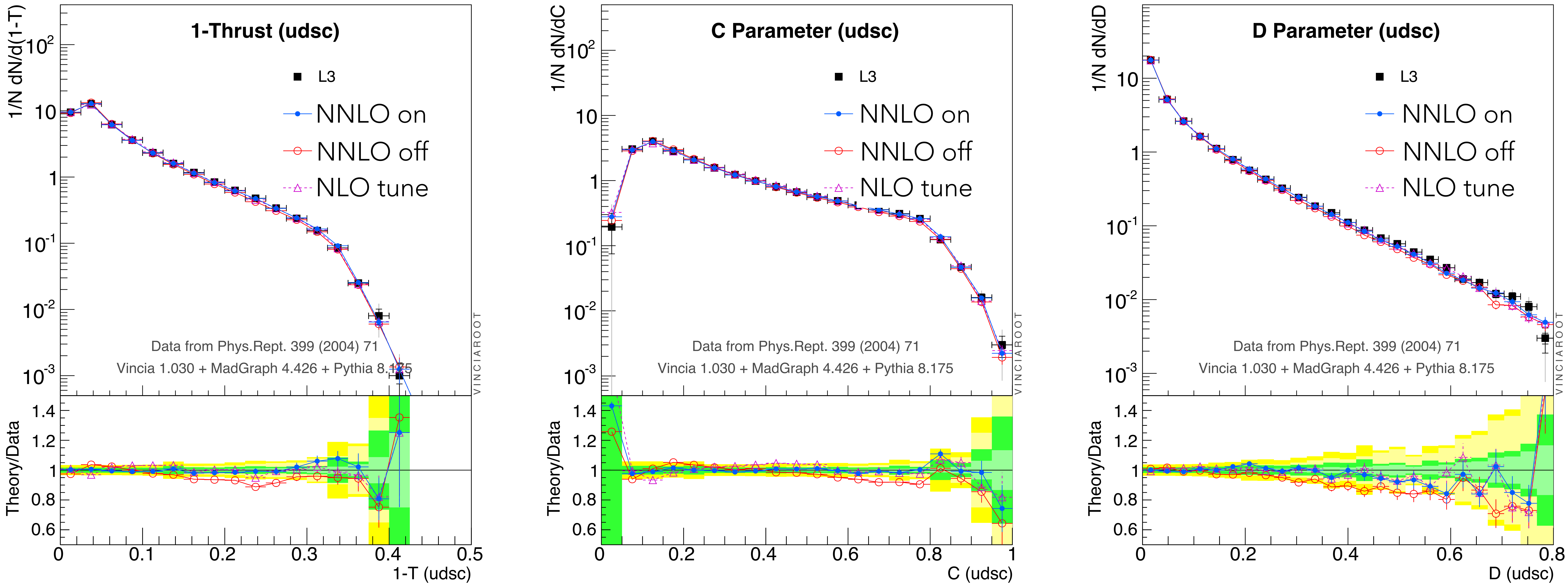
Hartgring, Laenen, **PS**, arXiv:1303.4974

Proof-of-Concept NNLO LEP tune (NNLO Z Decay, ie with NLO 3-jet corrections — using VINCIA)

NNLO tune (3-jet NLO) with $\alpha_s(M_Z) = 0.122$ (2-loop running, CMW)

NLO tune ~ Monash (3-jet LO) with $\alpha_s(M_Z) = 0.139$ (1-loop running, MSbar)

} Comparable values for Λ_{QCD}



1-Thrust (udsc)

C Parameter (udsc)

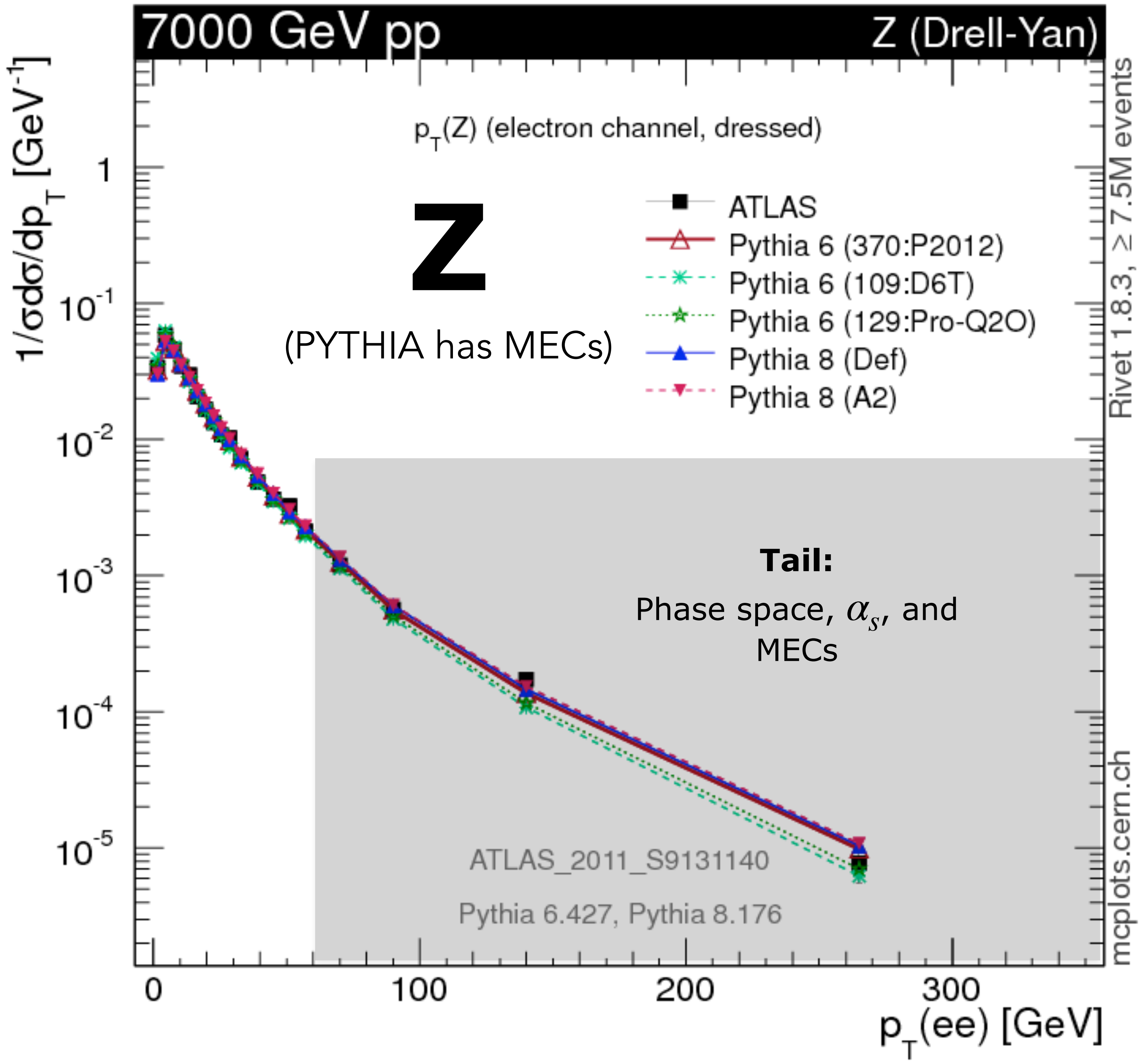
D Parameter (udsc)

Data from Phys.Rept. 399 (2004) 71

Vincia 1.030 + MadGraph 4.426 + Pythia 8.175

Controlling for Process Dependence!

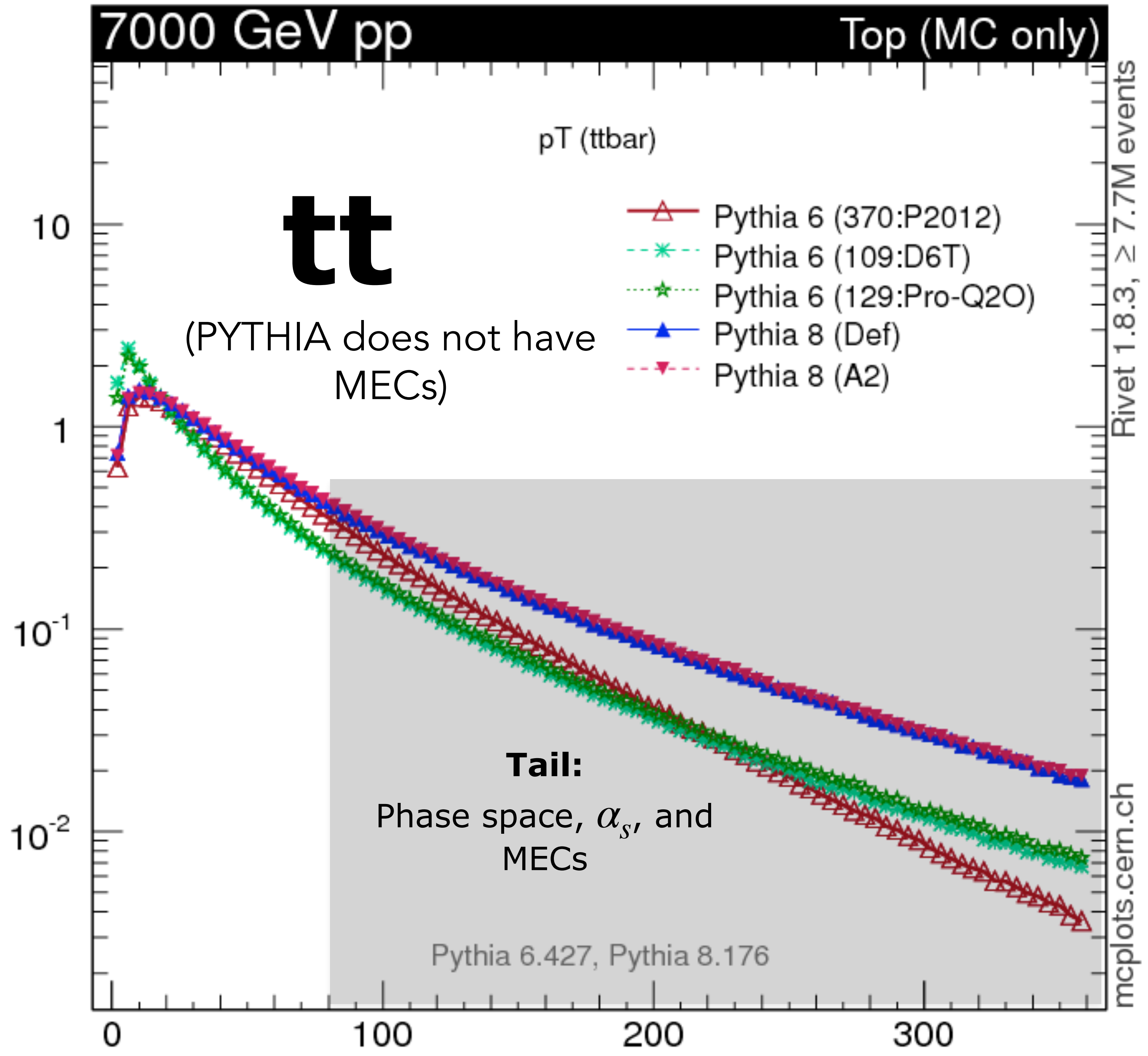
Note: these distributions rely on Pythia's "Power Showers"



Z

(PYTHIA has MECs)

Tail:
Phase space, α_s, and MECs



tt

(PYTHIA does not have MECs)

Tail:
Phase space, α_s, and MECs

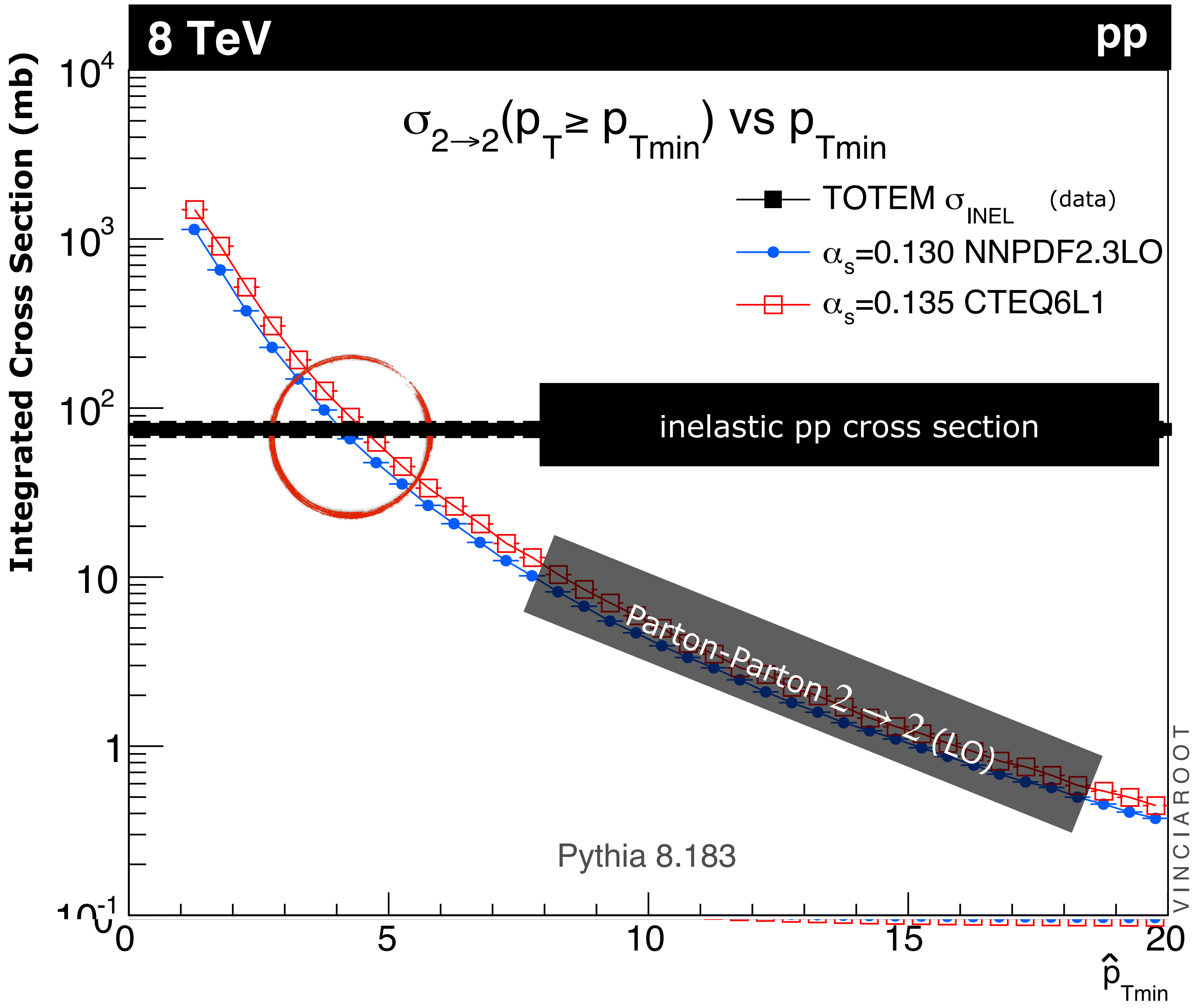
These points are quite sensitive to MECs / Matching / Merging.

➜ we should ensure we do MECs / matching / merging if we want to use them (or something equivalent to that.)

A Brief History of MPI in PYTHIA

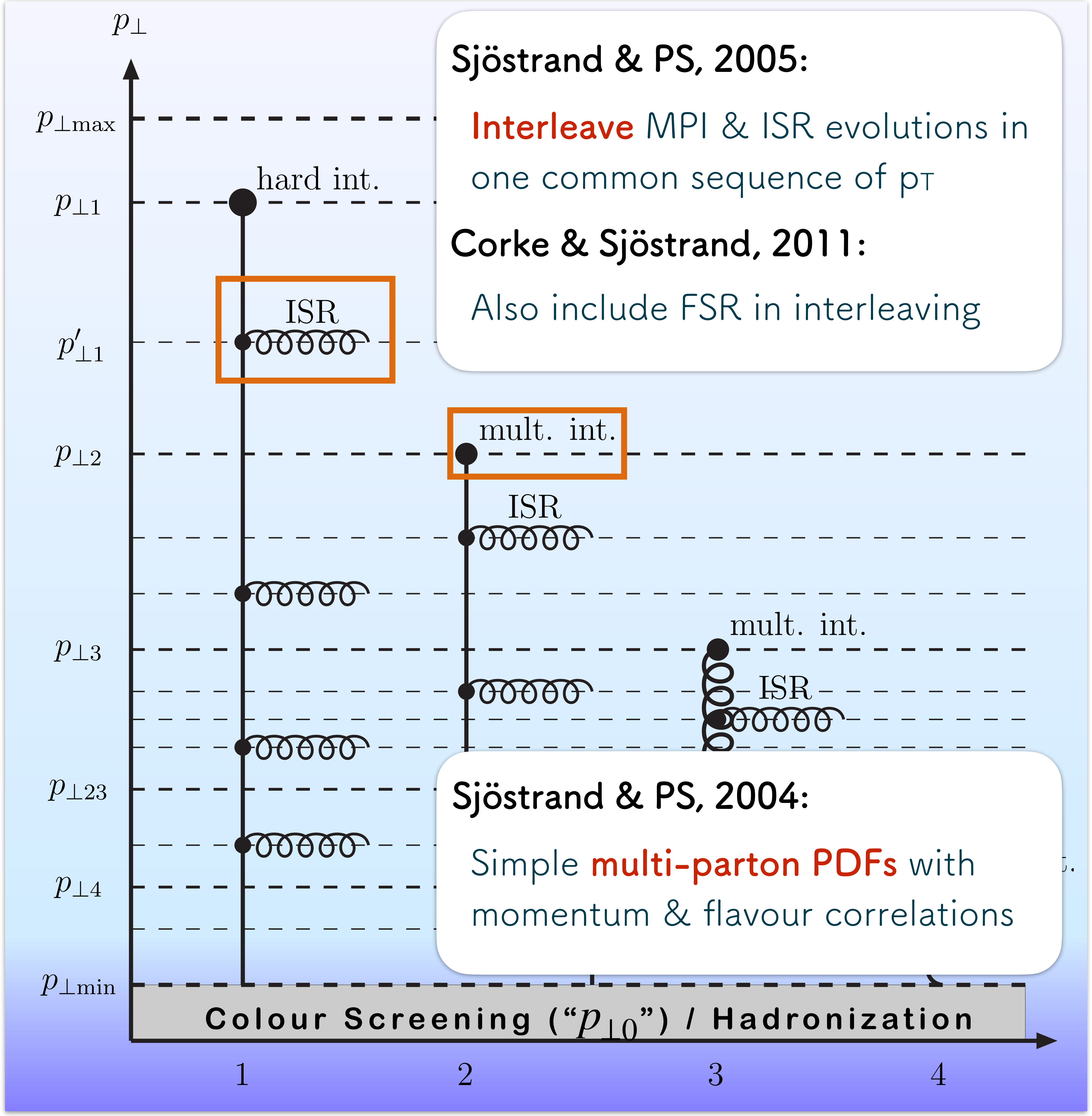
$$\frac{\sigma_{\text{parton-parton}}(\hat{p}_\perp)}{\sigma_{\text{hadron-hadron}}} > 1$$

$\Longrightarrow$ several parton-parton interactions *per* hadron-hadron interaction: **MPI**



Sjöstrand & van Zijl, 1985:

Cast as **Sudakov-style evolution equation**, analogous to the $\sigma_{X+\text{jet}}(p_\perp)/\sigma_X$ one of showers



Sjöstrand & PS, 2005:

Interleave MPI & ISR evolutions in one common sequence of pT

Corke & Sjöstrand, 2011:

Also include FSR in interleaving

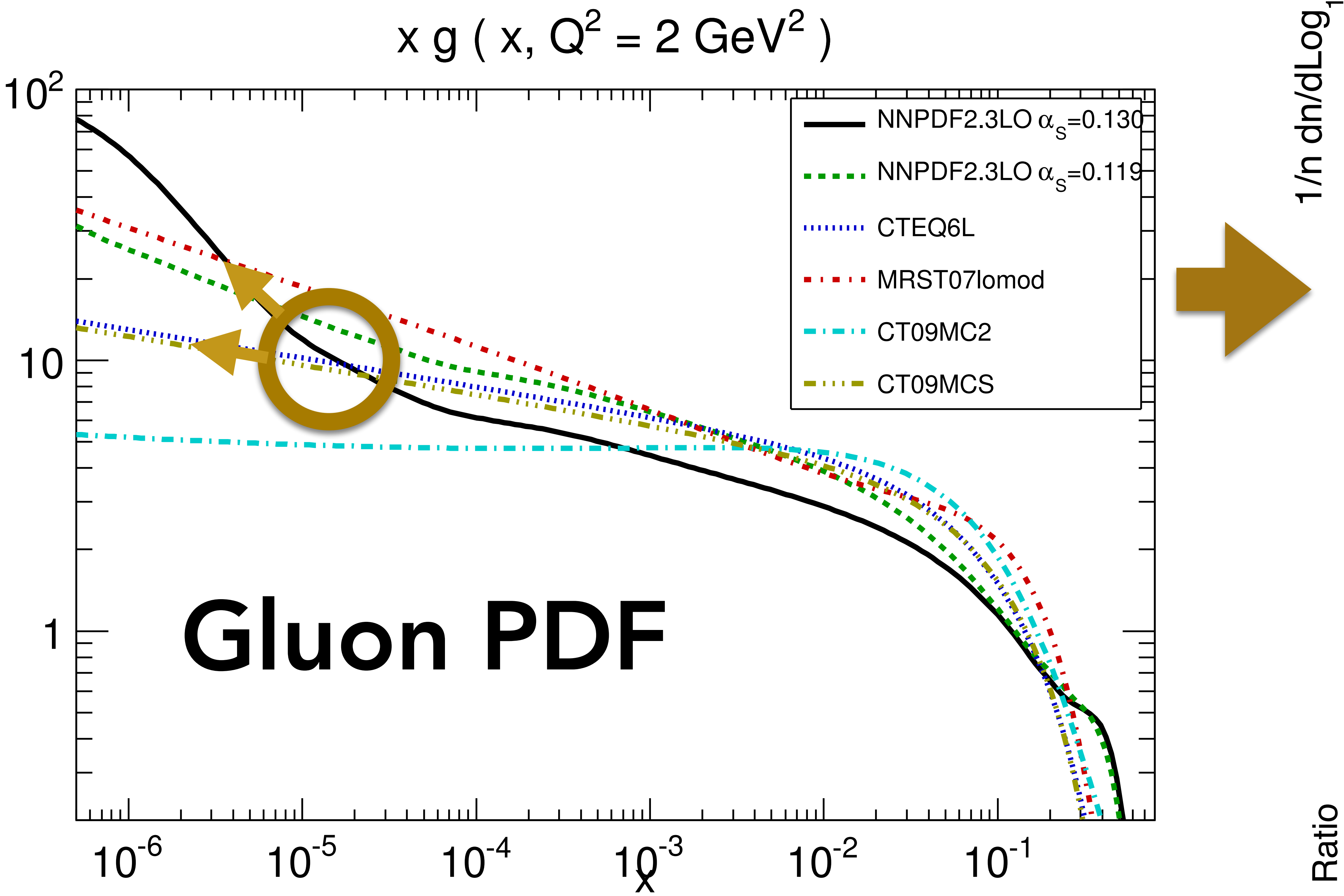
Sjöstrand & PS, 2004:

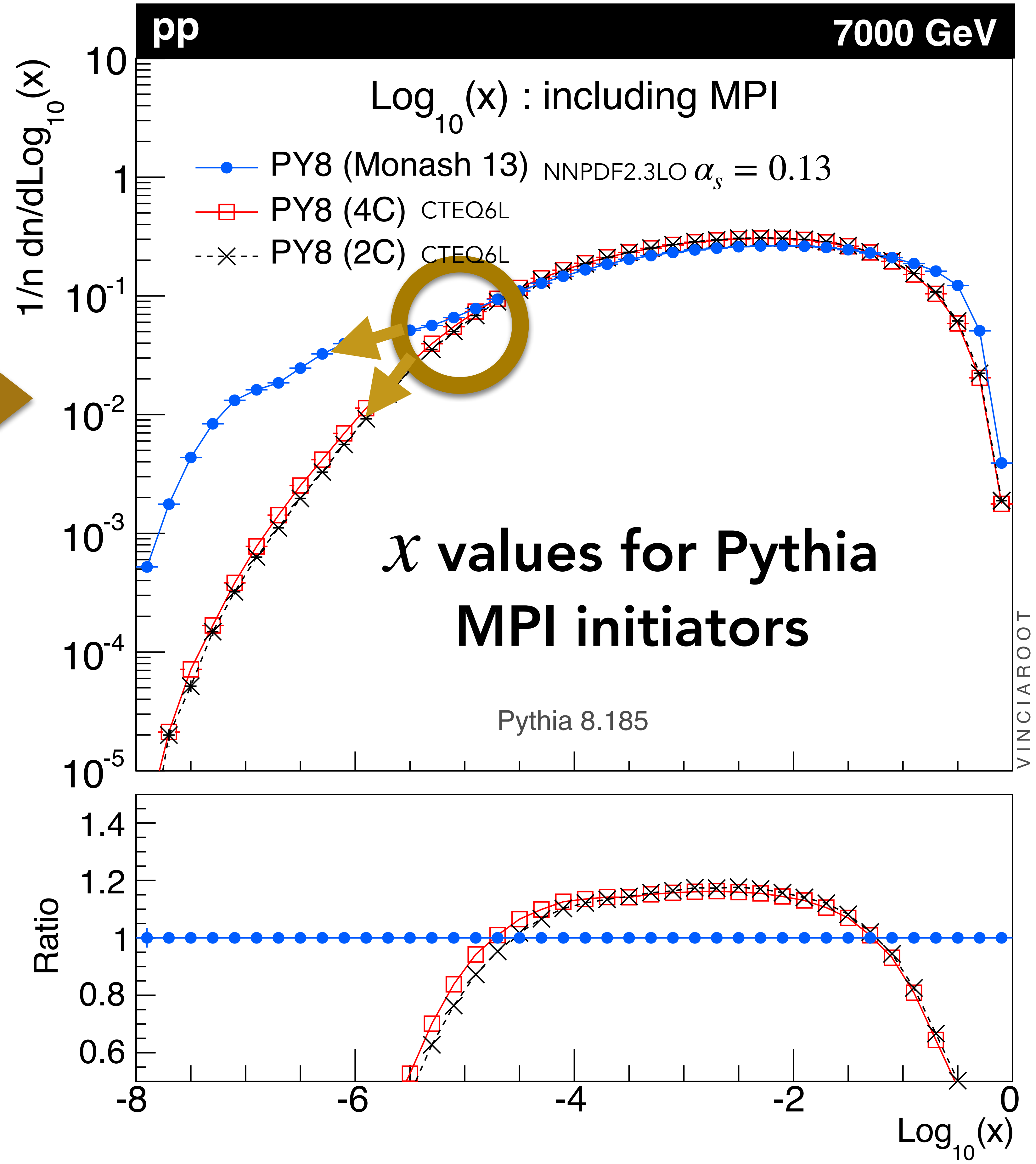
Simple **multi-parton PDFs** with momentum & flavour correlations

Figure from Sjöstrand & PS, 2005

Interplay between MPI and PDF set

Some PDFs that were available at the time of the Monash tune





Need sensible behaviour down to very low x,

and very low Q ~ ISR/MPI cutoff ~ 1 GeV

Negative PDFs not an option. Shower and MPI kernels are LO.

The issue with NLO gluons at low x

(Summary of note originally written by T. Sjöstrand, from discussions with R. Thorne though any oversimplifications or misrepresentations are our own)

Low-x gluon

Key constraint: DIS F_2

Low x: $\mathrm{d}F_2/\mathrm{d}\ln(Q^2)$ driven by $g \to q\bar{q}$

LO $P_{q/g}(z)$ ~ flat $\Longrightarrow$ x of measured quark closely correlated with x of mother gluon.

NLO Integral over $P_{q/g}(z) \propto 1/z$ for small z $\Longrightarrow$ approximate $\ln(1/x)$ factor.

➤ Effectively, the NLO gluon is probed more "non-locally" in x.

$\mathrm{d}\ln F_2/\mathrm{d}Q^2$ at small x becomes too big unless positive contribution from medium-to-high-x gluons (derived from $\mathrm{d}\ln F_2/\mathrm{d}Q^2$ in that region, and from other measurements) is combined with a negative contribution from low-x gluons.

Mathematically (toy NLO Calculation with just one x):

$$\frac{\mathrm{ME_{NLO}}}{\mathrm{ME_{LO}}} = 1 + \alpha_\mathrm{s}(A_1 \ln(1/x) + A_0)$$

$\ln(1/x)$ largely compensated in def of NLO PDF:

$$\frac{\mathrm{PDF_{NLO}}}{\mathrm{PDF_{LO}}} = 1 + \alpha_\mathrm{s}(B_1 \ln(1/x) + B_0)$$

➤ Product well-behaved at NLO if we choose $B_1 \approx A_1$

Cross term at $\mathcal{O}(\alpha_s^2)$ is beyond NLO accuracy …

For large x and small $\alpha_s(Q^2)$, e.g. $\alpha_s A_1 \ln(1/x) \sim 0.2$:

$$\frac{\mathrm{ME_{NLO}\,PDF_{NLO}}}{\mathrm{ME_{LO}\,PDF_{LO}}} = (1+0.2)(1-0.2) = 0.96$$

👍 log terms cancel

But if x and Q^2 are small, say $\alpha_s A_1 \ln(1/x) \sim 2$:

$$\frac{\mathrm{ME_{NLO}\,PDF_{NLO}}}{\mathrm{ME_{LO}\,PDF_{LO}}} = (1+2)(1-2) = -3$$

👎 Cross term dominates;
The PDF becomes negative

Not so important for high-p_T processes because 1) DGLAP evolution fills up low-x region, 2) kinematics restricted to higher x, 3) smaller α_s

Some Desirable Properties for PDFs for Event Generators

General-Purpose MC Generators are used to address very diverse physics phenomena and connect (very) high and (very) low scales ➤ **Big dynamical range!**

1. Stable (& positive) evolution to **rather low Q^2 scales**, e.g. $Q_0 \lesssim 1\,\mathrm{GeV}$

   ISR shower evolution and MPI go all the way down to the MC IR cutoffs ~ 1 GeV

2. **Extrapolates sensibly to very low $x \sim 10^{-8}$** (at LHC), especially at low $Q \sim Q_0$.

   "Sensible" ~ positive and smooth, without (spurious) structure

   Constraint for perturbative MPI: $\hat{s} \geq (1\,\mathrm{GeV})^2 \implies x_{\mathrm{LHC}} \gtrsim 10^{-8}$ $(x_{\mathrm{FCC}} \geq 10^{-10})$

   Main point: MPI can probe a **large range of x**, beyond the usual $\sim 10^{-4}$

   (Extreme limits are mainly relevant for ultra-forward / beam-remnant fragmentation)

3. **Photons** included as partons

   Bread and butter for part of the user community

4. **LO** or equivalent in some form (possibly with α_s^{eff}, relaxed momentum sum rule, …)

   Since MPI Matrix Elements are LO; ISR shower kernels also LO (so far)

5. Happy to have **N^{n}LO** ones in a similar family.

   E.g., for use with higher-order MEs for the hard process.

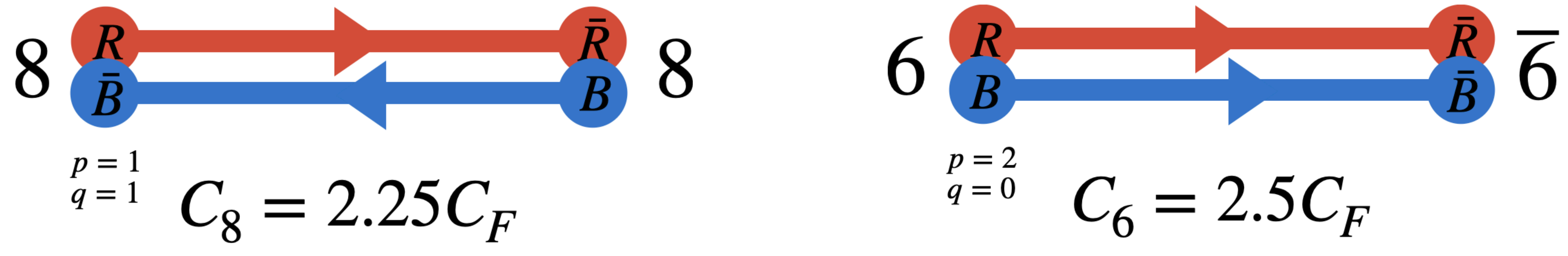
   Useful (but possible?) for these to satisfy the other properties too?

NEW

In Progress: Strangeness Enhancement from Close-Packing

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

Close-packing

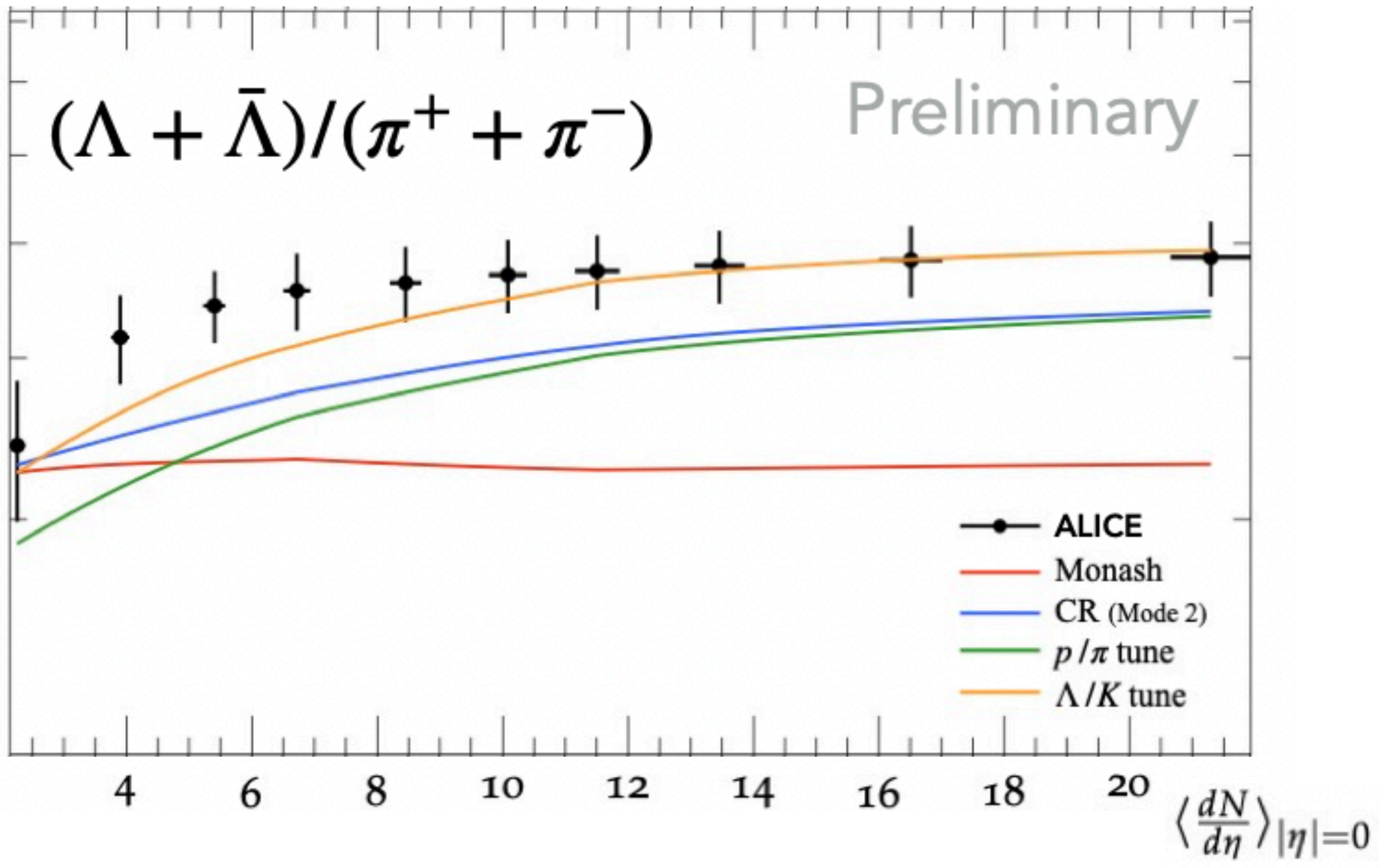


$C_8 = 2.25 C_F$ $C_6 = 2.5 C_F$

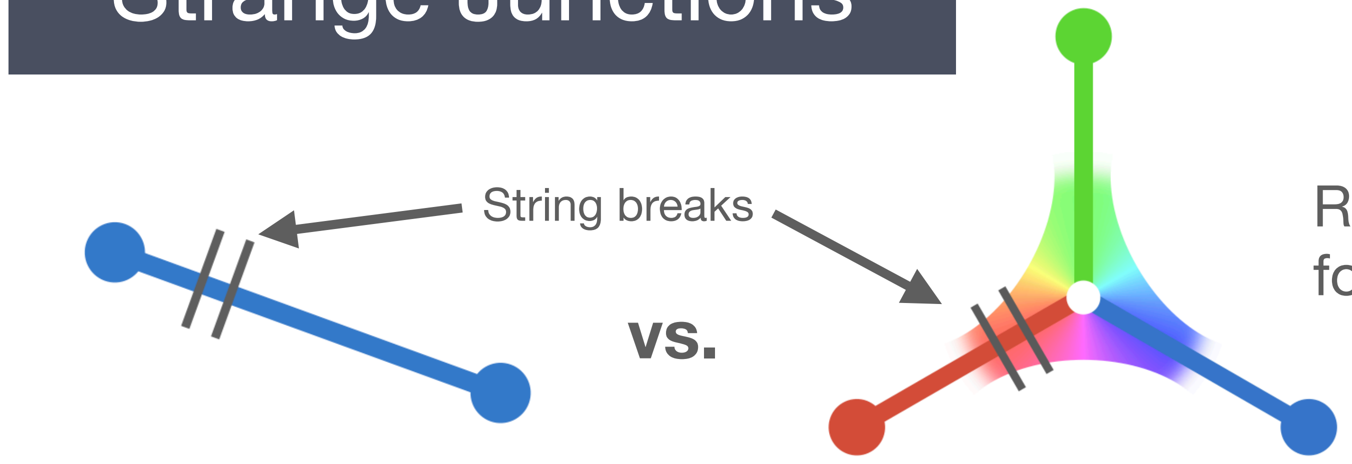
Dense string environments

→ Casimir scaling of **effective string tension**

→ Higher probability of strange quarks

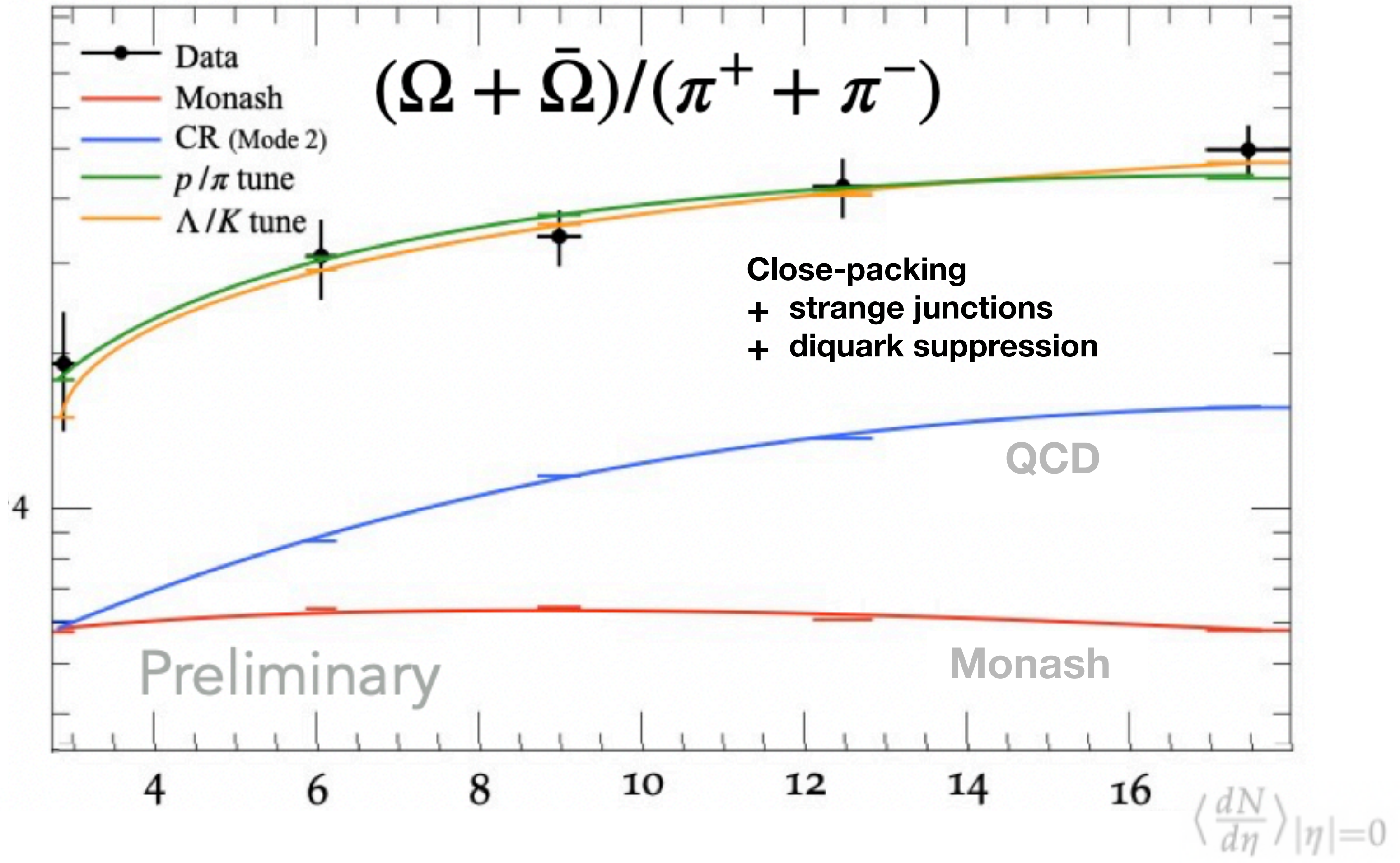


Strange Junctions



Results in strangeness enhancement focused in baryon sector

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



Slide adapted from J. Altmann