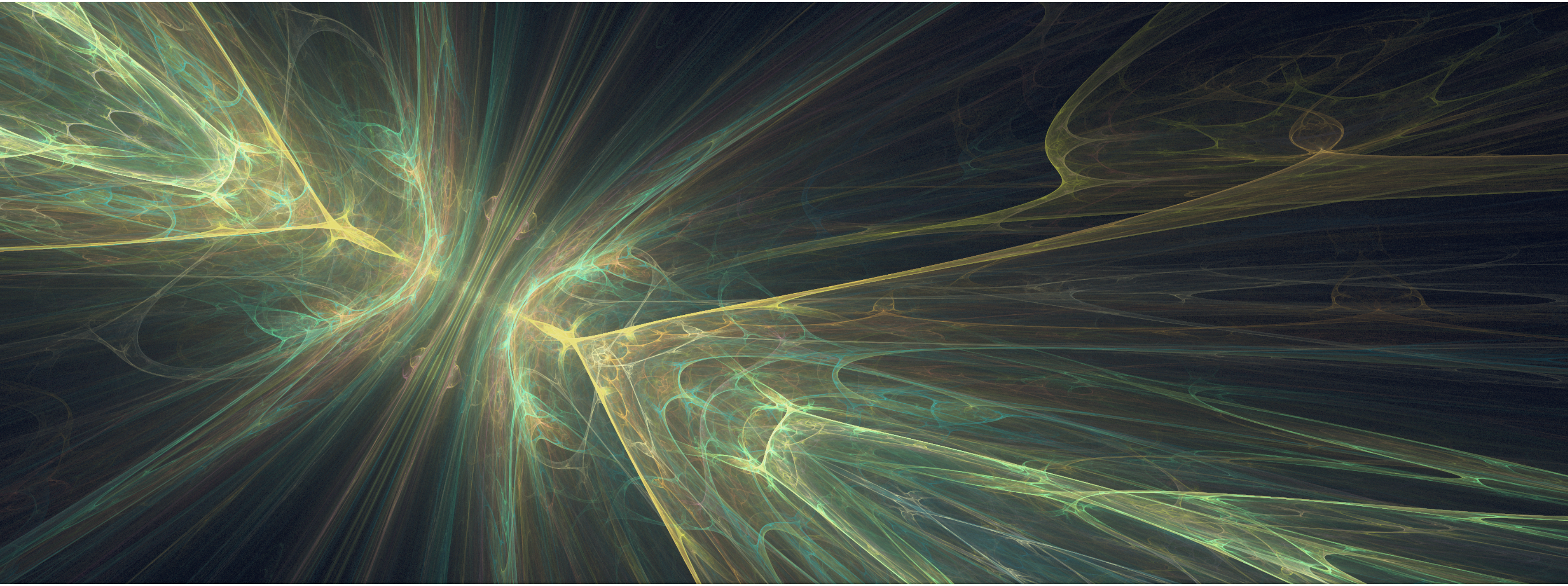
Non-Perturbative Aspects of Event Simulation in pp Collisions

Peter Skands — University of Oxford & Monash University



THE ROYAL SOCIETY

THE UNIVERSITY OF WARWICK

Australian Government
Australian Research Council

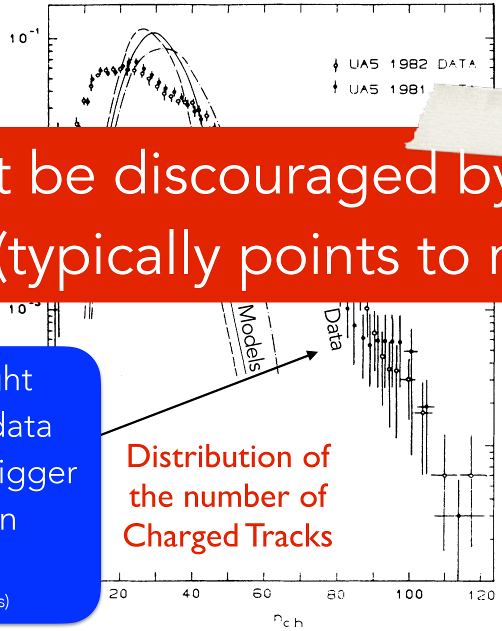
MCnet

MC vs Hadron Collisions

Last Lecture ➜ a model that included hard interactions, parton showers, and string fragmentation. Let's apply it to pp collisions!

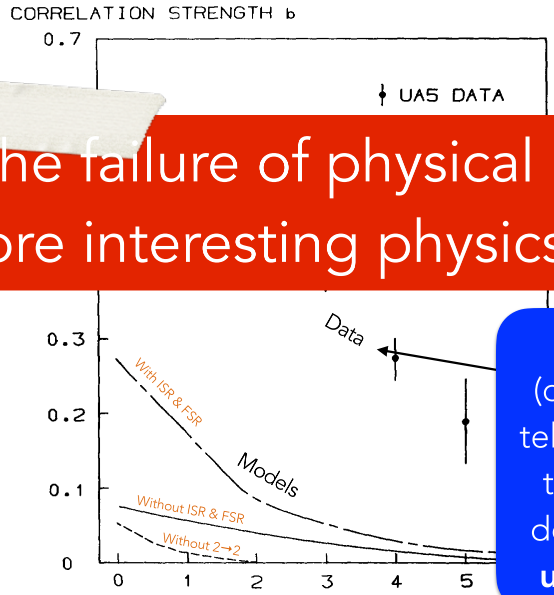
Do not be discouraged by the failure of physical models (typically points to more interesting physics)

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



Distribution of the number of Charged Tracks

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



Correlation Strength

(between # tracks in a bin at $\frac{-|\eta|}{2}$ and one at $\frac{+|\eta|}{2}$)

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

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

Further evidence of additional physics in hadron-hadron collisions

1983: discovery of the "Pedestal Effect"

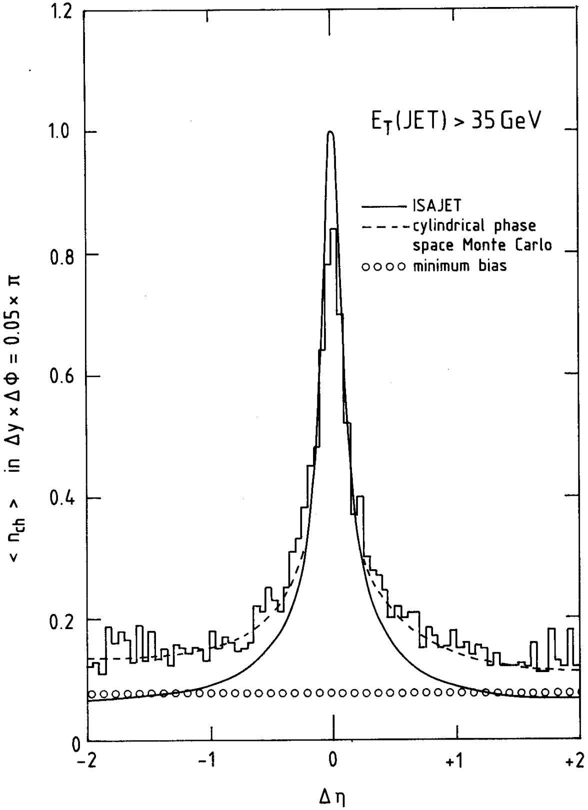
UA1: $p\bar{p}$ at $\sqrt{s} = 540\,\mathrm{GeV}$

Studies of jets with E_T up to 100 GeV

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

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

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



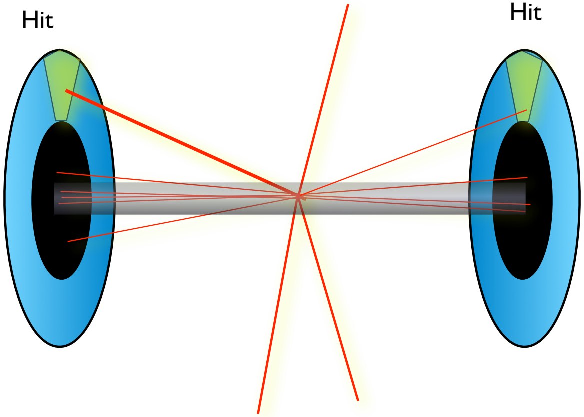
What's "Minimum Bias"?

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

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

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

Example of "Minimum Bias Trigger"



Minimum Bias = Minimal trigger requirement

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

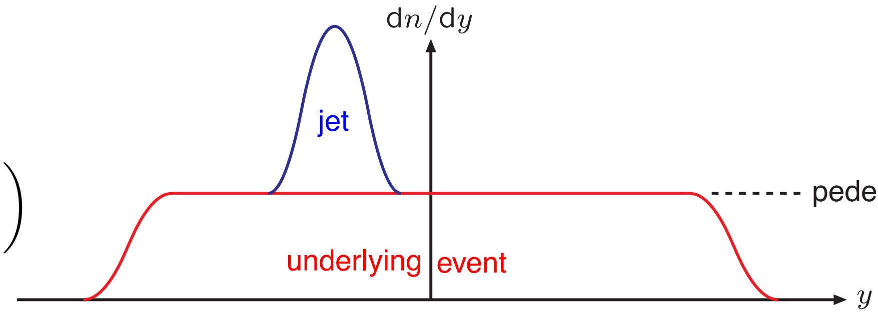
(Double-sided trigger requirement suppresses "single diffraction")

Dissecting the Pedestal

Today, we call the pedestal “the Underlying Event”

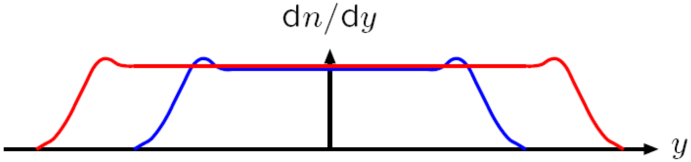
Illustrations by T. Sjöstrand

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



Rapidity (along beam axis)

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



Rapidity (along string axis)

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

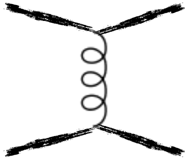
Multiple Strings?

Parton-Parton vs Proton-Proton Cross Sections

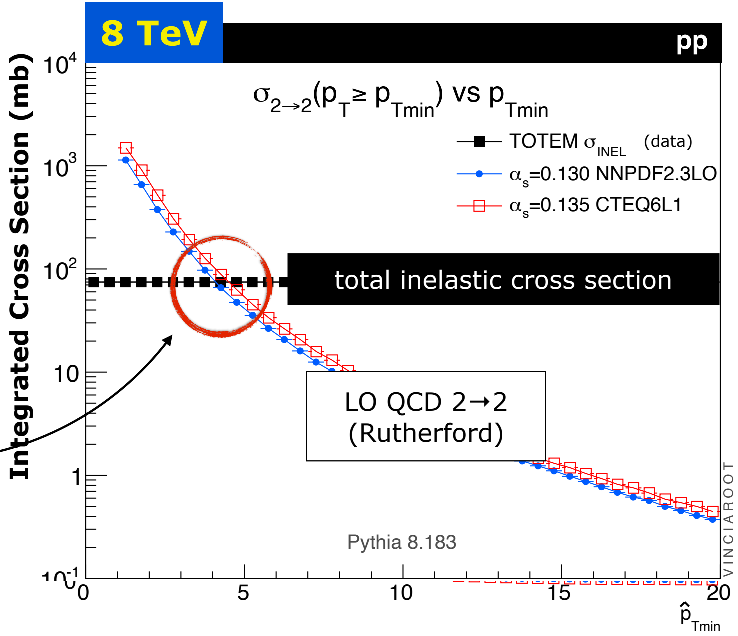
Total inelastic pp cross section @ 8 TeV* ~ 80 mb (measured by TOTEM)

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

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



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



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

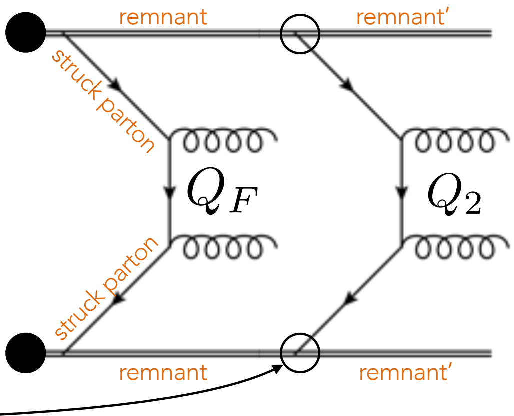
*Note: nothing particularly special about 8 TeV; the crossover point would be lower at lower E_{CM} and higher at higher E_{CM}

Physics of the Pedestal

Recall Factorisation: Subdivide calculation

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

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



More colour exchanges → more strings → more hadrons

+ (mini)-jets from tail with $Q_2 \gg 1$ GeV

Multi-Parton Interactions (MPI)

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

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

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

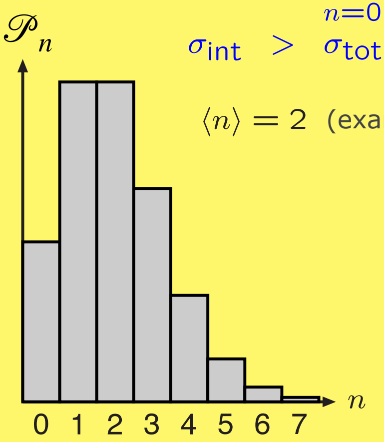
(More issues such as *colour reconnections, saturation*, $3 \rightarrow 4$, *rescattering*, …, not covered here)

How many?

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

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

$$\begin{aligned}\sigma_\mathrm{tot} &= \sum_{n=0}^{\infty} \sigma_n \\ \sigma_\mathrm{int} &= \sum_{n=0}^{\infty} n\,\sigma_n \\ \sigma_\mathrm{int} &> \sigma_\mathrm{tot} \iff \langle n\rangle > 1\end{aligned}$$



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

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

Real Life

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

Momentum conservation suppresses high-n tail

Impact-parameter dependence

+ physical correlations

→ not simple product

Interleaved Evolution in PYTHIA

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

MPI cast as Sudakov-style evolution in p_T analogous to the one for showers

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

Interleave MPI & ISR evolutions in one common sequence of p_T

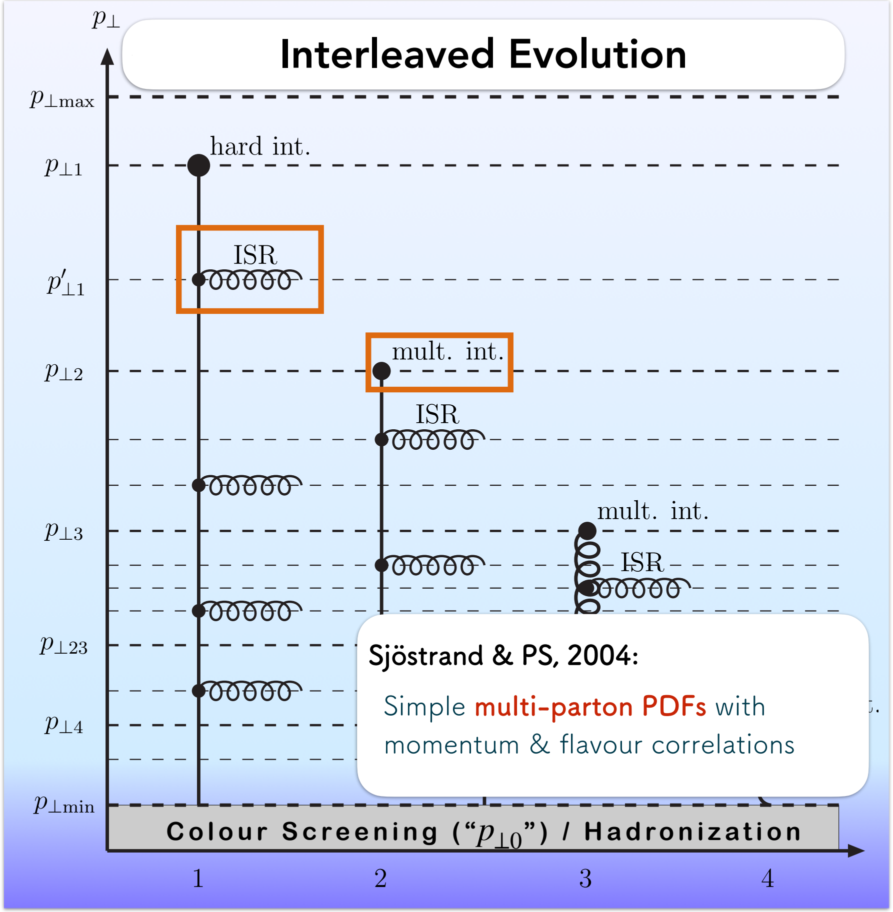
➜ ISR & MPI "compete" for the available x in the proton.

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

Also include **FSR** in interleaving

2021 [Brooks, **PS**, Verheyen, SciPost Phys. 12 (2022) 3]

Also include **Resonance Decays** in interleaving (VINCIA)



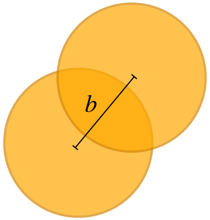
Sjöstrand & PS, 2004:

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

Figure from Sjöstrand & PS, 2005

Impact Parameter Dependence

$\Big($

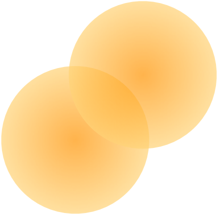


1. **Simple Geometry** (in impact-parameter plane)

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

$\Big)$

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



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

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

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

MC with MPI vs Hadron Collisions

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

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

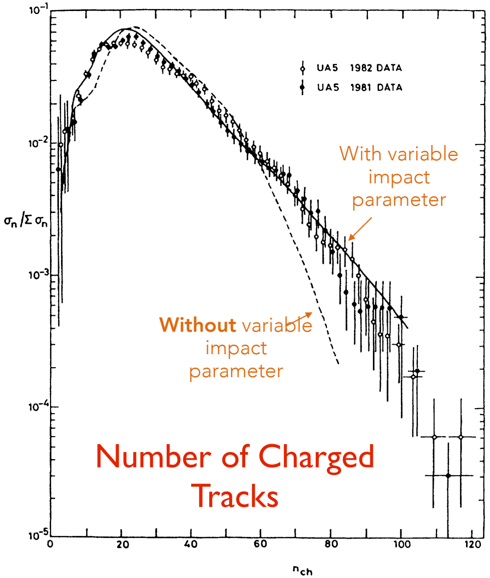
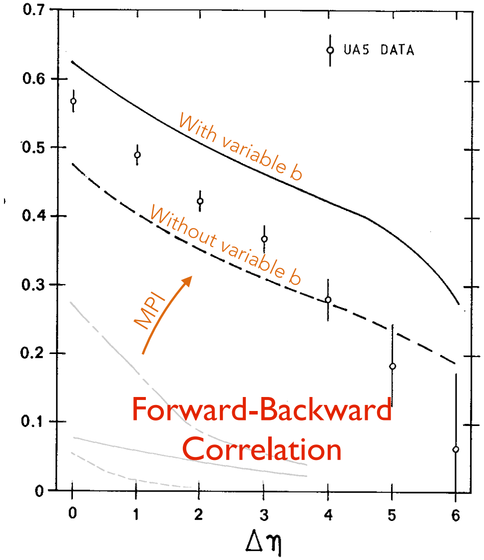
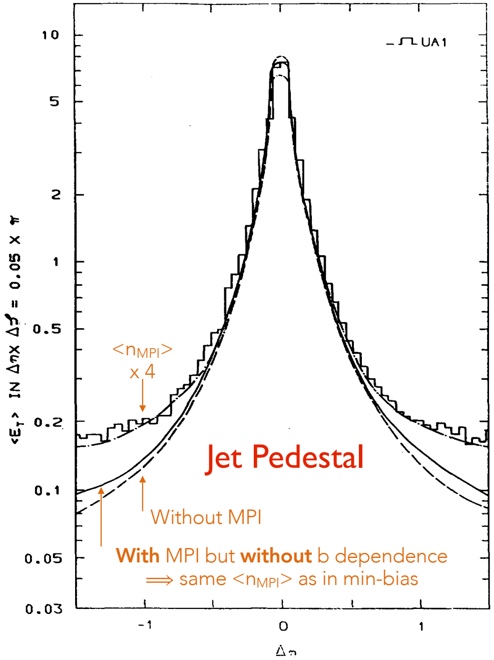


FIG. 12. Charged-multiplicity distribution at 540 GeV, UA5 results (Ref. 32) vs multiple-interaction model with variable impact parameter: solid line, double-Gaussian matter distribution; dashed line, with fix impact parameter [i.e., $\tilde{O}_0(b)$].



MPI → Long-distance correlations in rapidity



Impact-parameter dependence → UE

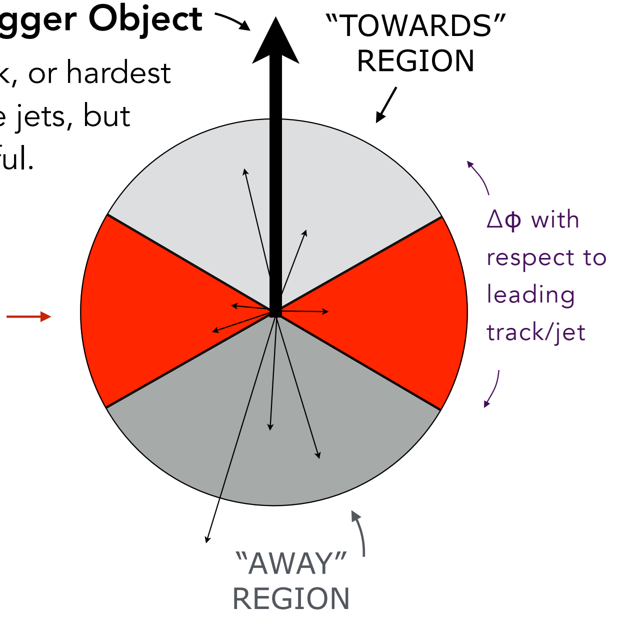
Characterizing The Underlying Event

(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.

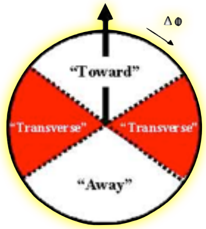


"Transverse Region" (TRNS)

Sensitive to activity at right angles to the hardest jets

➜ Useful definition of Underlying Event

Min-Bias VS Underlying Event



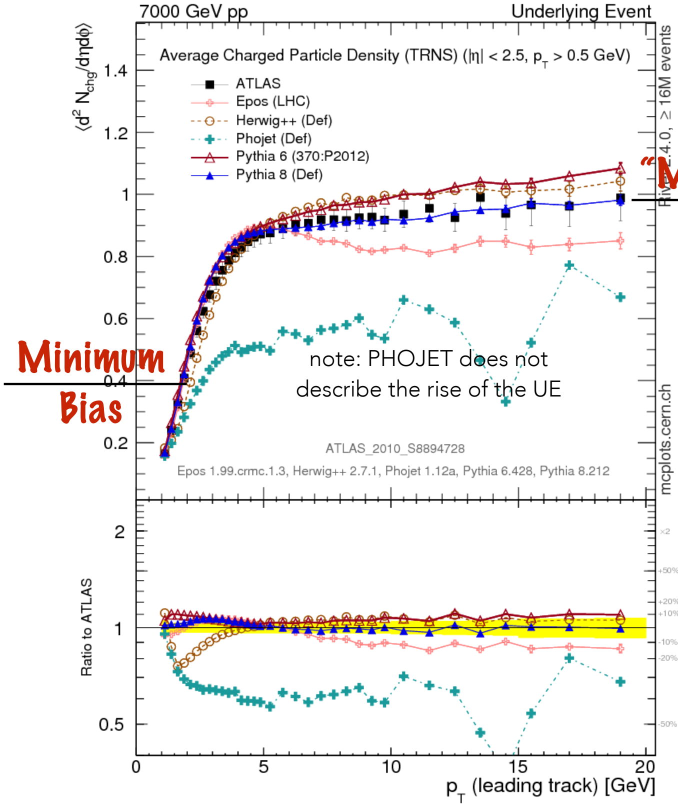
Tautology:

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



note: PHOJET does not describe the rise of the UE

"Maximum Bias"

Plot from mcplots.cern.ch

Colour Space in Hadron Collisions

Colour Correlations

Each MPI exchanges colour between the beams

The colour flow determines the hadronizing string topology

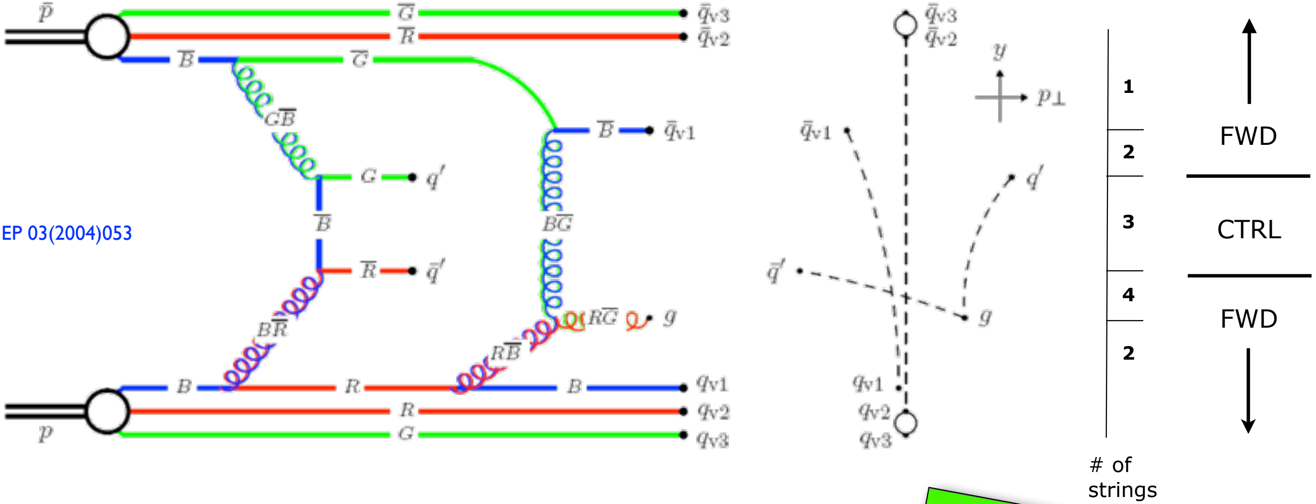


Figure from
Sjöstrand & PS, JHEP 03(2004)053

Different models make different ansätze

Each MPI, even when soft, is a colour spark

Hadron distributions depend on colour space

P. Skands

Colour Correlations

Each MPI exchanges colour between the beams

The colour flow determines the hadronizing string topology

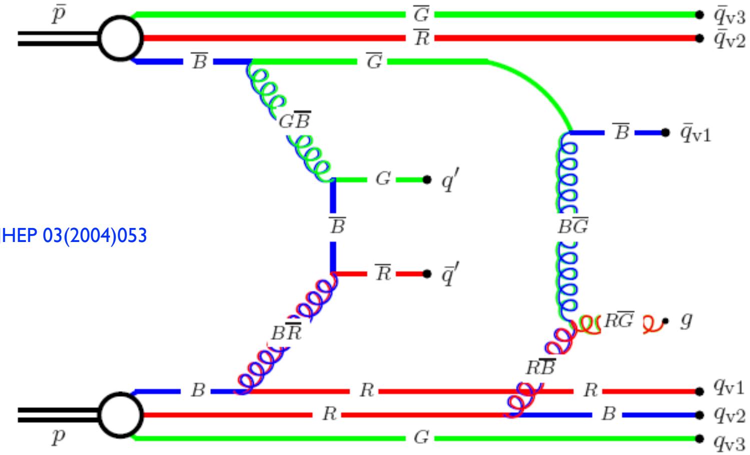
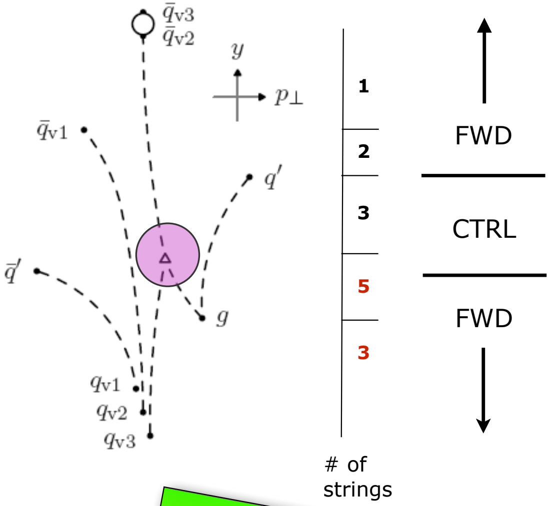


Figure from
Sjöstrand & PS, JHEP 03(2004)053



Each MPI, even when soft, is a colour spark

Hadron distributions depend on colour space

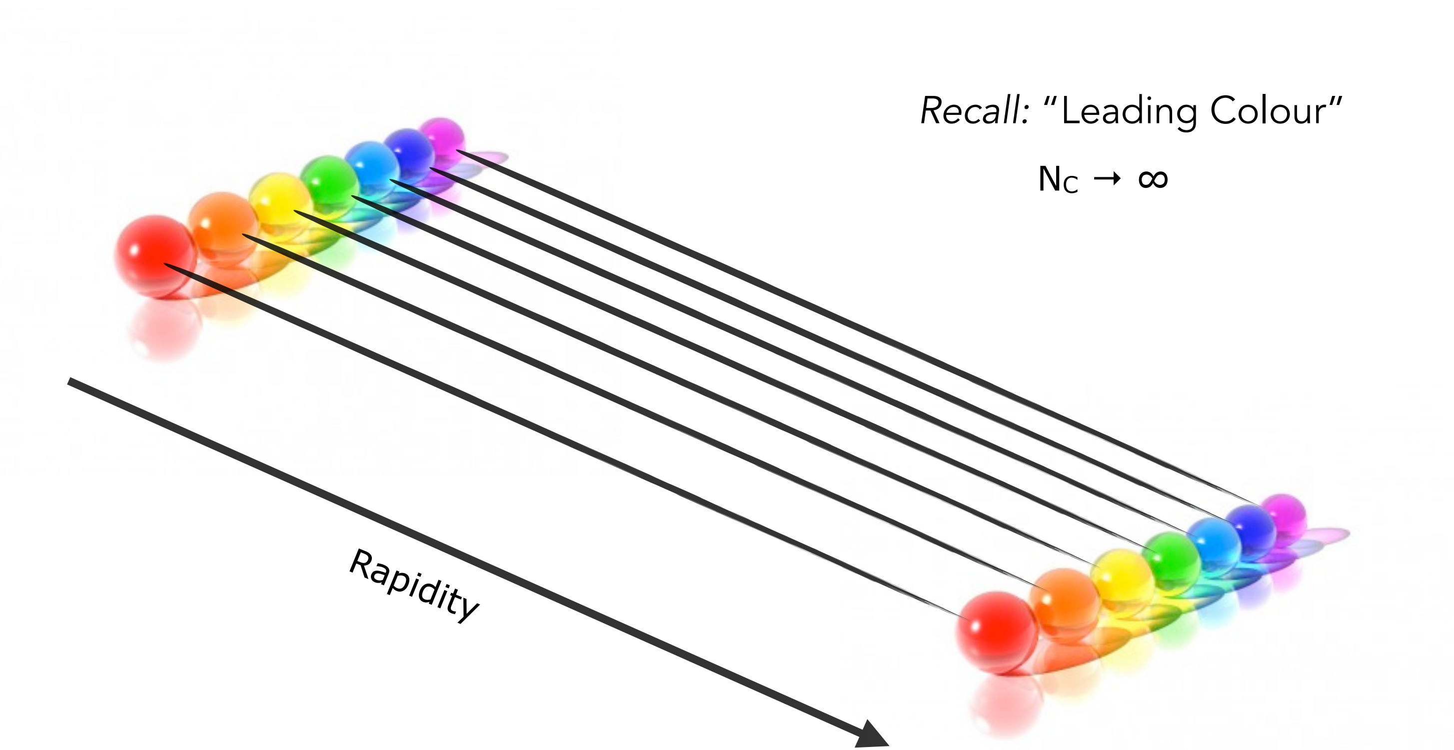
Different models make different ansätze

Colour Connections

Multi-Parton Interactions ➜ Expect Multiplicity ∝ N_{MPI} ?

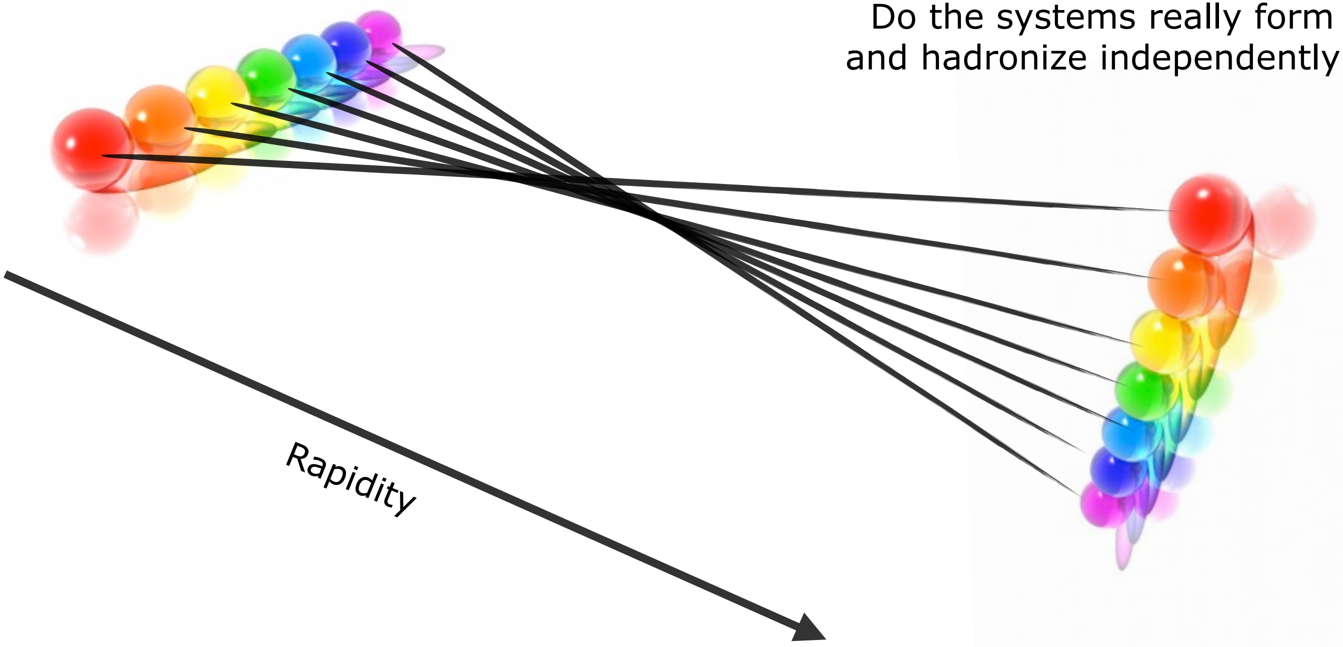
Recall: "Leading Colour"

$N_C \to \infty$



Colour Reconnections?

Multi-Parton Interactions + Colour coherence ➜ Multiplicity $\overset{<}{\not\propto} N_{\text{MPI}}$?

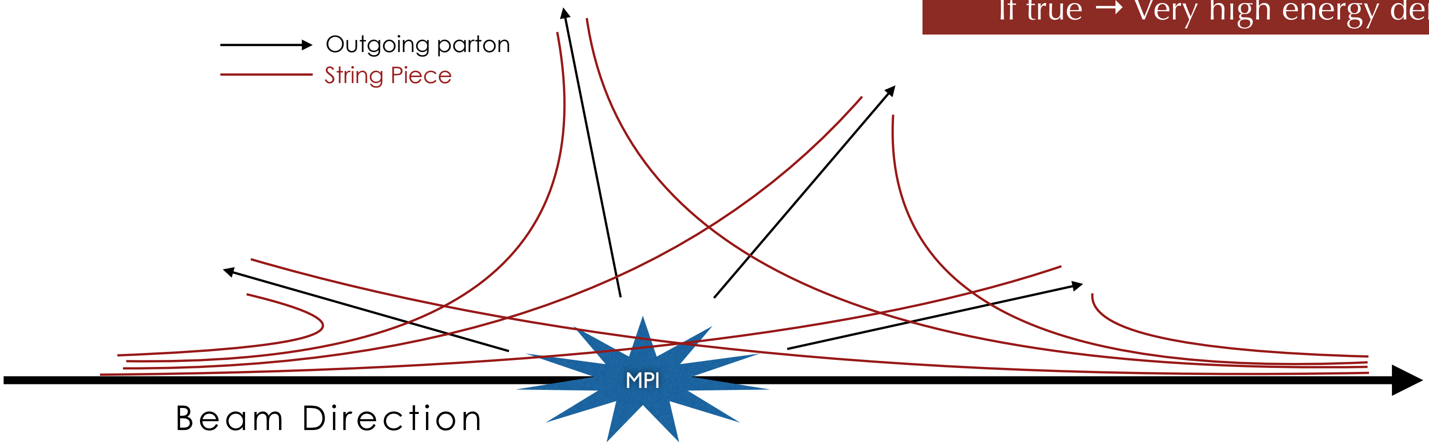


Do the systems really form
and hadronize independently?

Hadronization — with MPI

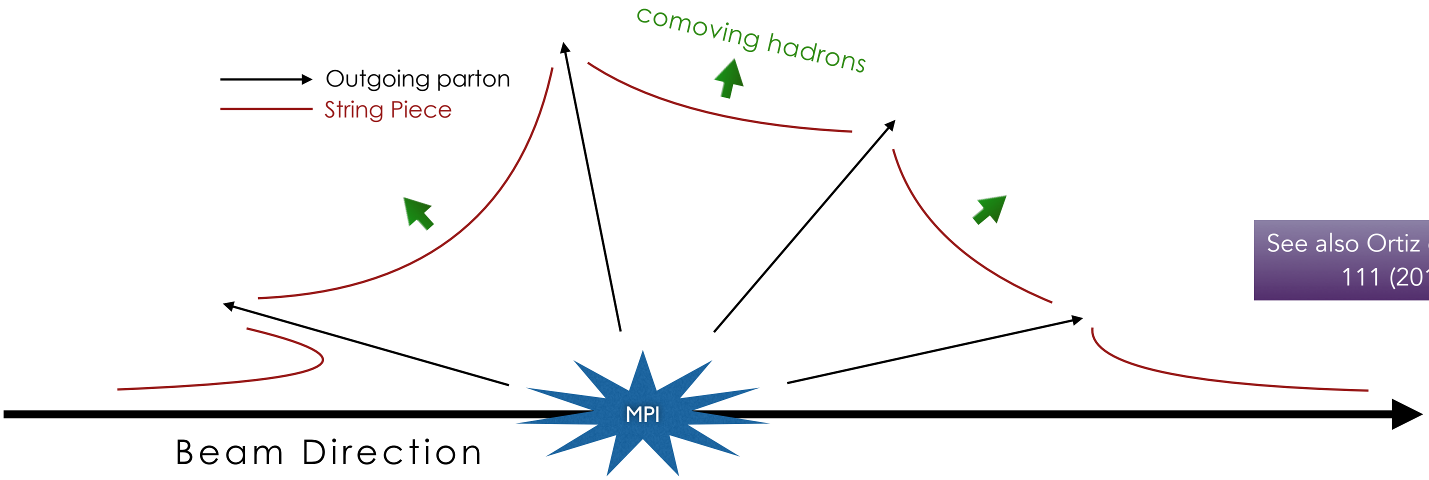
“Leading Colour”
➜ Each MPI hadronizes **independently** of all others
~ the equivalent of “independent fragmentation” for MPI

So many strings in so little space
If true → Very high energy densities



Colour Reconnections (CR)

With Colour Reconnections
MPI hadronize **collectively**
But how do we know which partons should be confined with which?



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

Confinement in LHC Collisions

High-energy pp collisions — with ISR, MPI, and Beam Remnants

Final states with **very many** coloured partons

Who gets confined with whom?

"QCD Colour Reconnection" Model: [Christiansen & **PS** JHEP 08 (2015) 003]

Stochastically sample ~ all possibilities

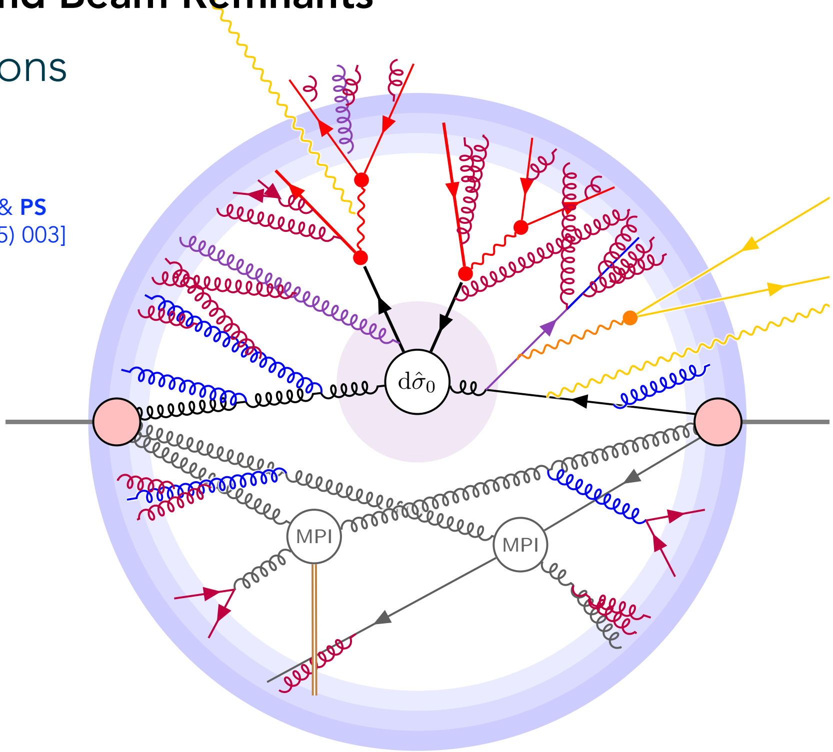
E.g.: **random triplet** charge has 1/9 chance to be in **singlet** state with **random antitriplet**:

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

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

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

Choose between allowed string configurations: smallest world-sheet area (a.k.a. "string-length" minimization)



**): 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.*

How to confront with measurements?

Can't measure n_{MPI} directly

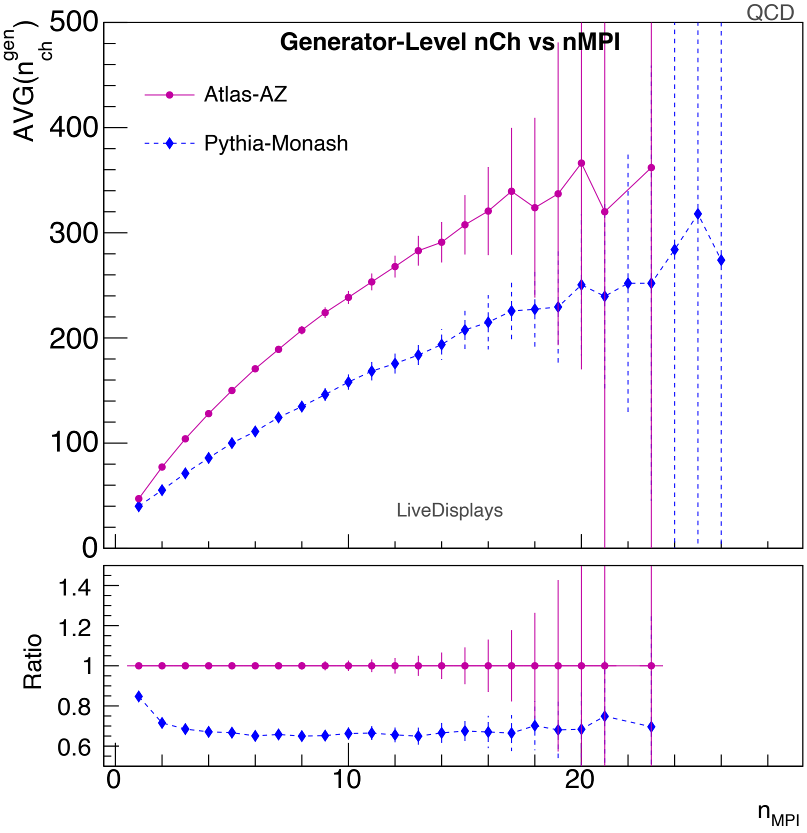
Use number of particles produced ~ rough indicator of how much colour gets kicked around

$\Longrightarrow$ study event properties as a function of "N_{ch}" = N_{tracks}

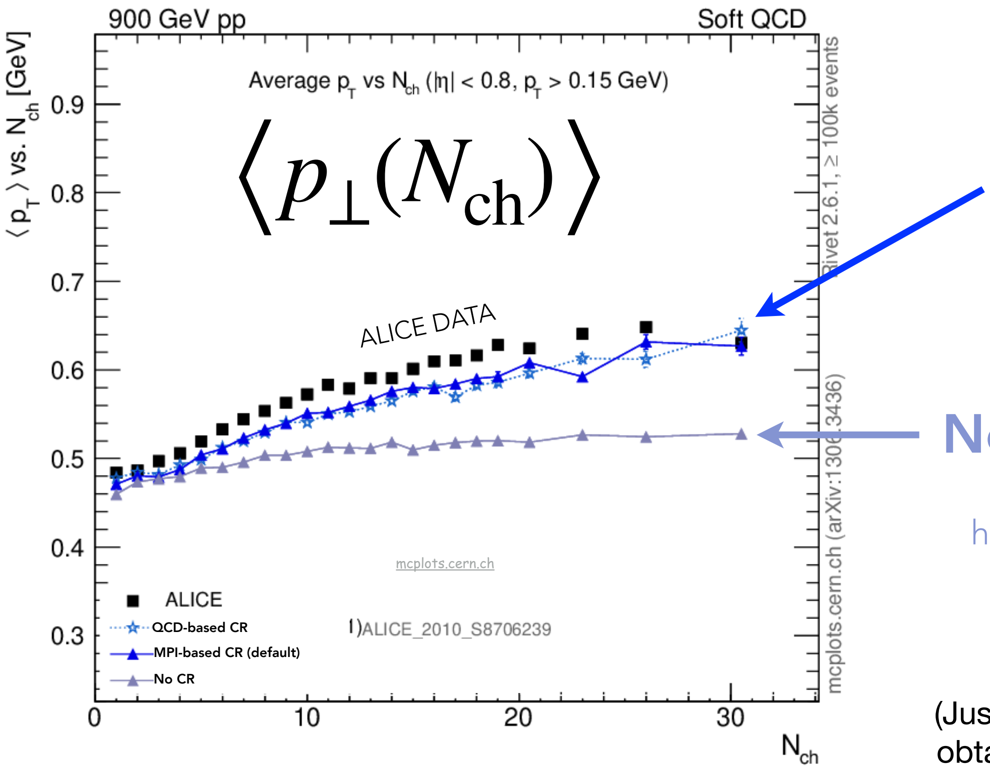
Different models/tunes

Predict different number of charged particles "per" MPI

But all predict a strong correlation ➜ useful indicator



Consequences of CR: <p_T>(nch)



Different CR Models

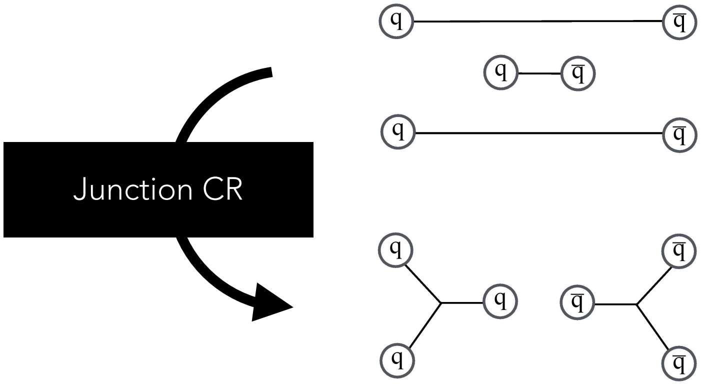
Both **MPI-based** (default) and **QCD-based** CR can reproduce the rising trend of <pT>(N_{ch})

No CR ⟹ <p_T> ~ same for all N_{ch}

(Many MPI just produce more hadrons, but with ~ same spectra)

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

+ New junction-type CR ⟹ Increased Baryon-to-Meson ratios



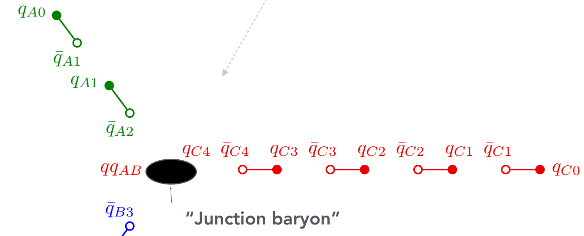
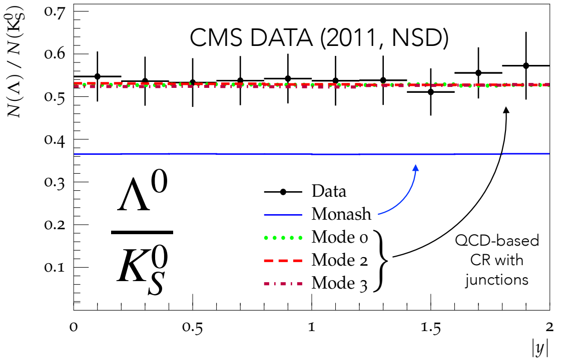


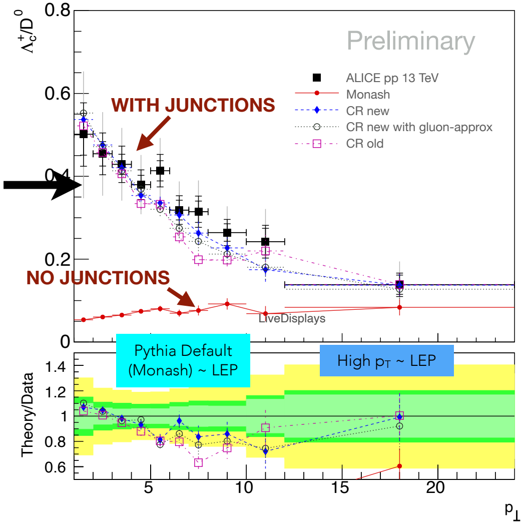
Illustration from Pythia 8.3 manual

Λ/K^0_S versus rapidity at $\sqrt{s} = 7$ TeV [Christiansen & PS, arXiv:1505.01681]



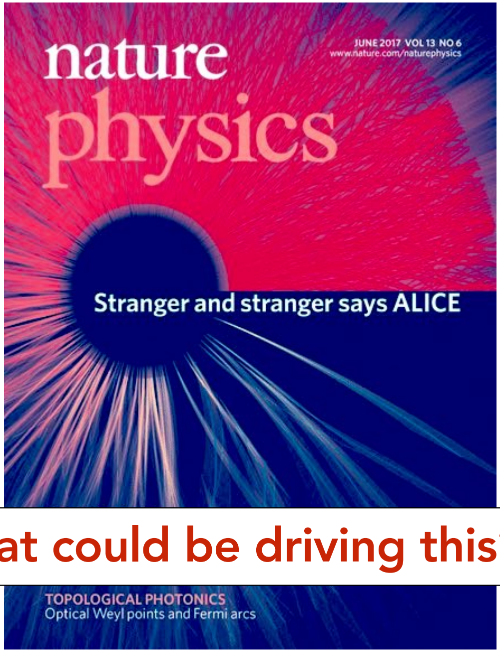
ALICE 2021: also in charm

QCD CR model(s):
Junctions drive
order-of-magnitude
increase in Λ_c/D^0 at
low $p_\perp$



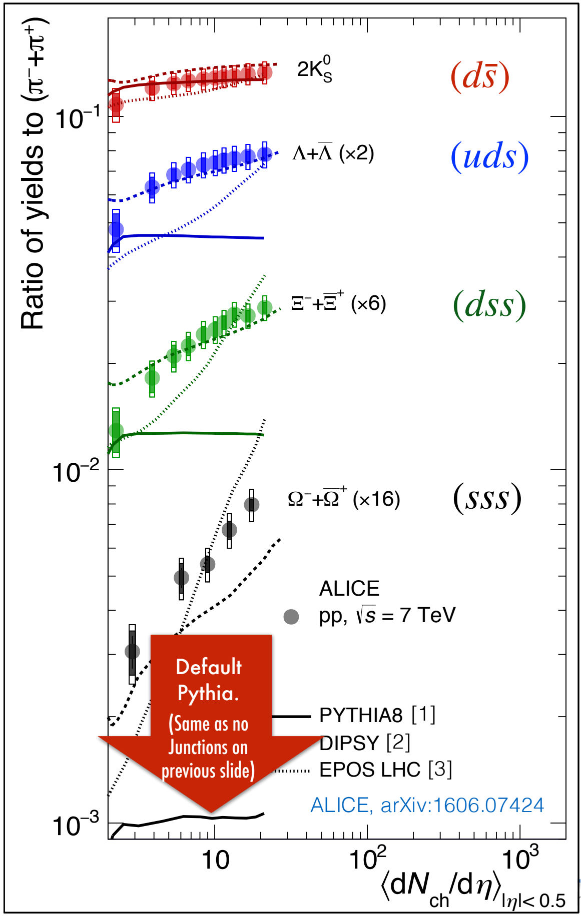
What a **strange** world we live in, said Alice

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



June 2017

What could be driving this?



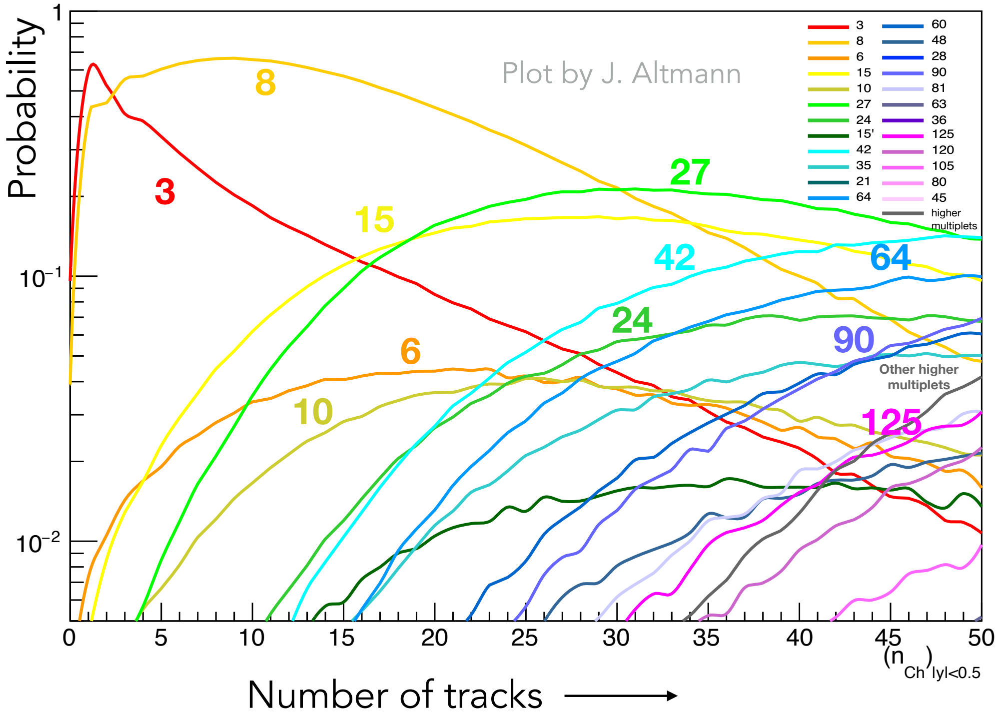
ALICE, arXiv:1606.07424

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

→ Non-Linear String Dynamics?

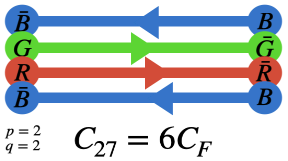
MPI $\Longrightarrow$ lots of coloured partons scattered into the final states

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

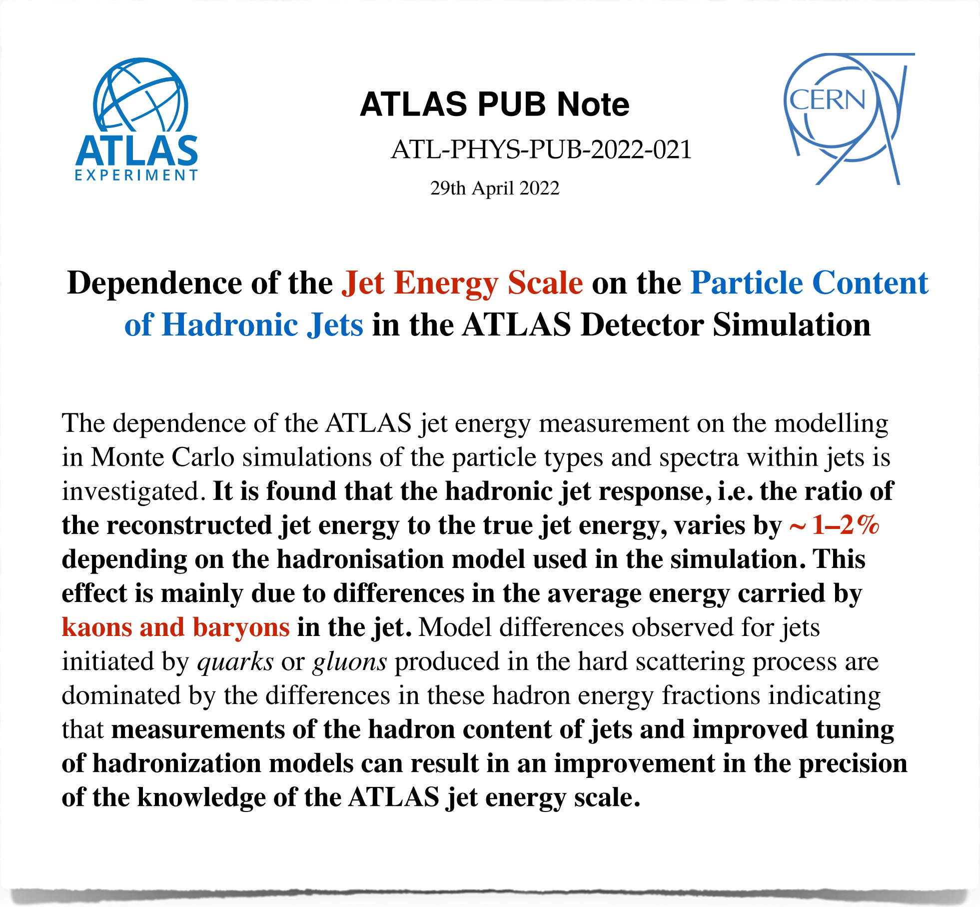
$C_{27} = 6C_F$

Two approaches in PYTHIA:

1) Colour Ropes (Lund)

2) Close-Packing (Monash)

Particle Composition: Impact on Jet Energy Scale



ATLAS 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

Thank you for your attention!

Hard Process

- ○ Hard Interaction
- ● Resonance Decays
- 🟪 MECs, Matching & Merging

Parton Showers

- 🟥 QCD Final-State Radiation
- 🟦 QCD Initial-State Radiation*
- 🟨 Electroweak Radiation

Underlying Event

- ○ Multiparton Interactions
- ☐ Beam Remnants*

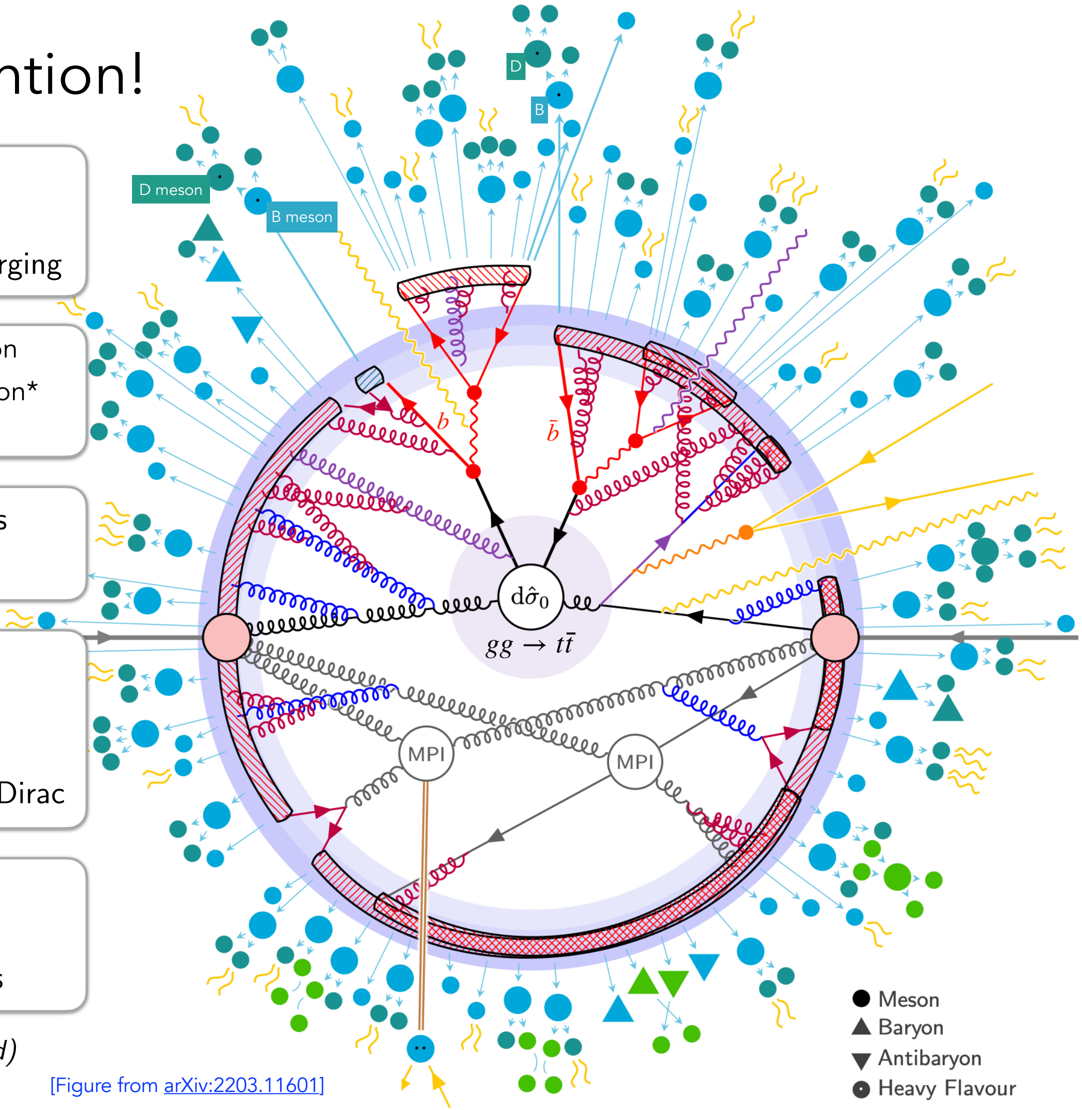
Hadronization

- ▨ Strings
- ☐ Colour Reconnections
- ☐ String Interactions
- ☐ Bose-Einstein & Fermi-Dirac

Hadron (& τ) Decays

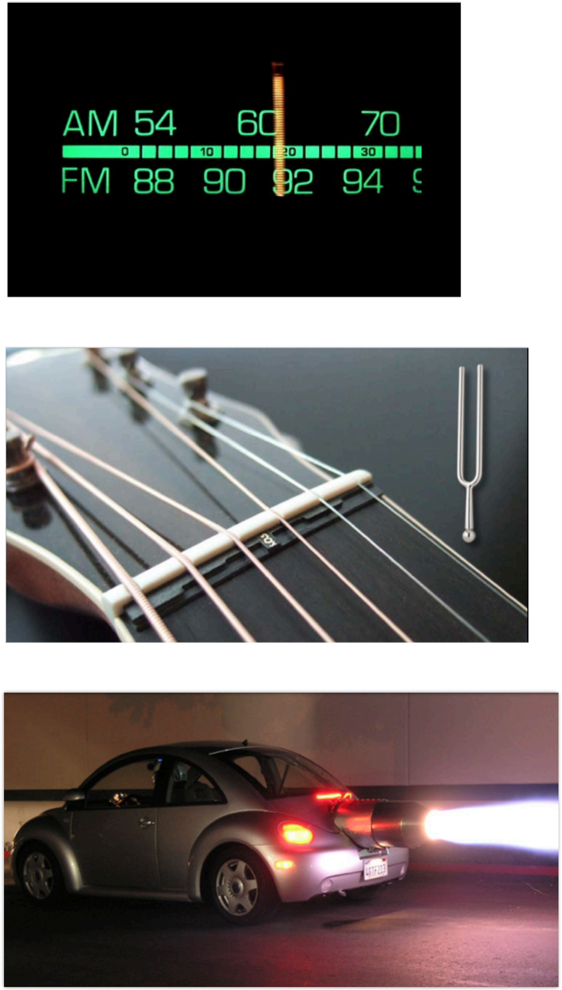
- 🟦 Primary Hadrons
- 🟩 Secondary Hadrons
- 🟩 Hadronic Reinteractions

(: incoming lines are crossed)*



[Figure from arXiv:2203.11601]

Extra Slides



PART 2

Tuning & Validation

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

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

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

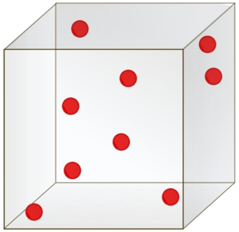
2^{nd}-order polynomials account for parameter correlations.

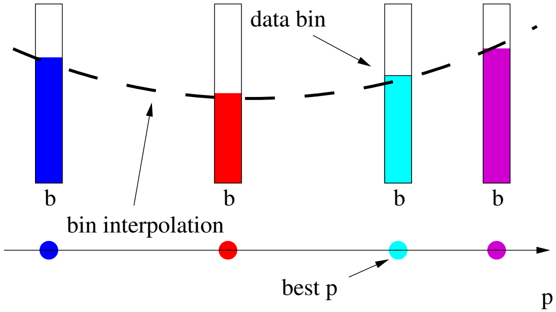
Modern Python Package
with much more functionality,
tutorials, etc.
https://professor.hepforge.org/

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`





Caveat 1: Tensions vs Incompatibilities ?

Physics Model may not be able to simultaneously agree with all measurements

Not immediately a concern. Consider overall physics/consistency ↔ your priorities.

Physics Model may be *unable* to agree with (some part of) a given measurement

Fit reacts by desperately trying to reduce order-of-magnitude differences in bins it shouldn't have been asked to fit in the first place

At cost of everything else ➤ total garbage.

Choose measurements carefully

Within context of physics model ⟷ domain of applicability

- This can also apply to *bins* of a histogram, e.g., if part of a measurement goes outside domain of validity of theory model
- E.g., professor allows to put zero (or very small) weights for some bins

Consider whether you should effectively "give up" on some measurements

Caveat 2: Sensitivities and Observable Hierarchies

For each observable and/or MC parameter you want to consider:

What is/are the most salient MC parameter(s) which **that** observable is sensitive to

PROFESSOR can help with this ➜ sensitivity and correlation analyses

What is a **full set of observables** that **span** constraints on those parameters?

Are some of those observables/parameters more important than others?

Do some parameters control **larger aspects of the modelling** ➜ Cross checks.

Are some observables more important **to you** than others? ➜ Weightings.

Example: a measurement reveals the kaon yield is too low in the MC

You can increase the production of *all* particles, including kaons

Or you could increase just the strangeness *fraction* keeping the total constant

If you don't know and don't think about this, you risk tuning to agree with kaons while mistuning agreement on the overall level of particle production. Is that what you want?

➜ include an observable sensitive to the total number of particles (or kaon fraction)

Caveat 3: 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.

Still bad overall fit, typically not spanning uncertainties (only on one side)

My recommendation:

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?

Some Helper Tools

Wouldn't it be nice if there was a tool:

That could automatically detect correlations between parameters and observables.

And tell you which "groups" they fall into naturally : which parameter sets you should ideally tune together, and which are more nicely factorised.

This is (at least partly) **what the tool AutoTunes does** Bellm, Gellersen, Eur.Phys.J.C 80 (2020)

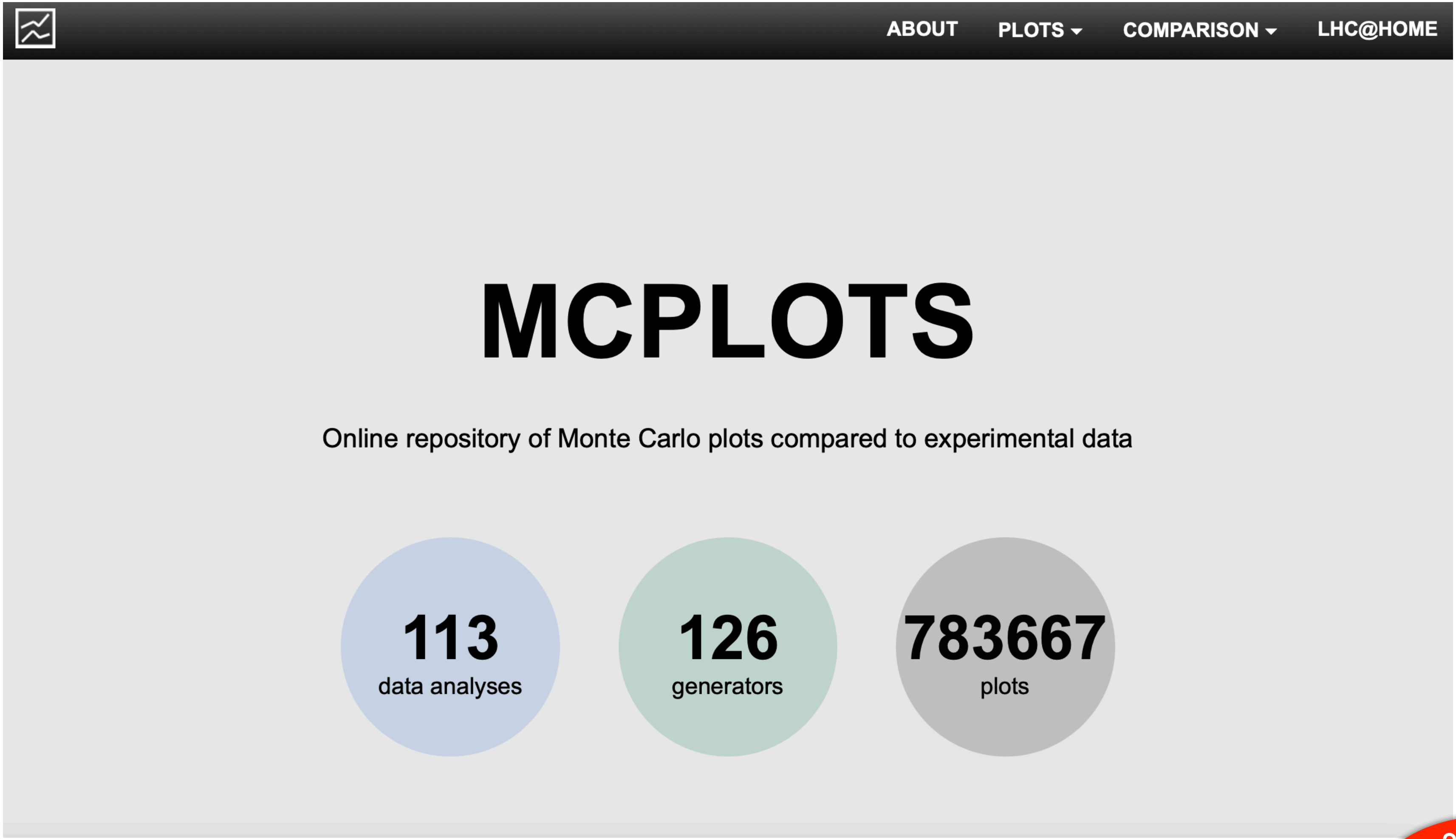
I won't have time to discuss that today, but I think it looks promising

I encourage you to study it and use it

You may also be interested in Apprentice Krishnamoorthy et al., EPJ Web Conf. 251 (2021) 03060

Variance reduction to semi-automate how to weight observables & bins

mcplots.cern.ch



MCPLOTS

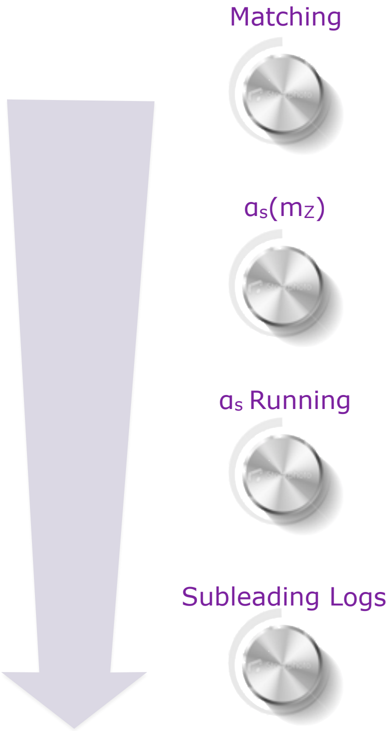
Online repository of Monte Carlo plots compared to experimental data

113 data analyses

126 generators

783667 plots

Parameters (in PYTHIA): FSR pQCD Parameters



Additional Matrix Elements included?

At tree level / one-loop level? Using what matching scheme?

The value of the strong coupling

In PYTHIA, you set an effective value for $\alpha_s(m_Z^2) \Leftrightarrow$ choice of k in $\alpha_s(kp_\perp^2)$

Renormalization Scheme and Scale for α_s

1- vs 2-loop running, MSbar / CMW scheme, choice of k in $\alpha_s(kp_\perp^2)$, cf

Ordering variable, coherence treatment, effective 1→3 (or 2→4), recoil strategy, …

Branching Kinematics (z definitions, local vs global momentum conservation), hard parton starting scales / phase-space cutoffs, masses, non-singular terms, …

Parameters (in PYTHIA): String Tuning

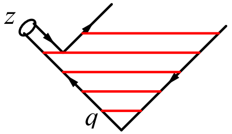
Hadron energy fractions

Fragmentation Function

The “Lund a and b parameters”

Or use a and $\langle z \rangle$ instead (less correlated) A. Jueid et al., *JCAP* 05 (2019) 007

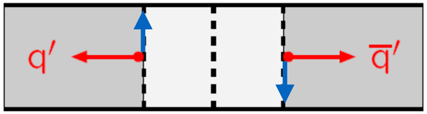
+ $\Delta a_{\mathrm{diquark}}$ for baryons



p_T in string breaks

Scale of string-breaking process

Shower cutoff and $\langle p_\perp \rangle$ in string breaks



Meson Multiplets

Mesons

Strangeness suppression, **Vector/Pseudoscalar**, η, η′, …

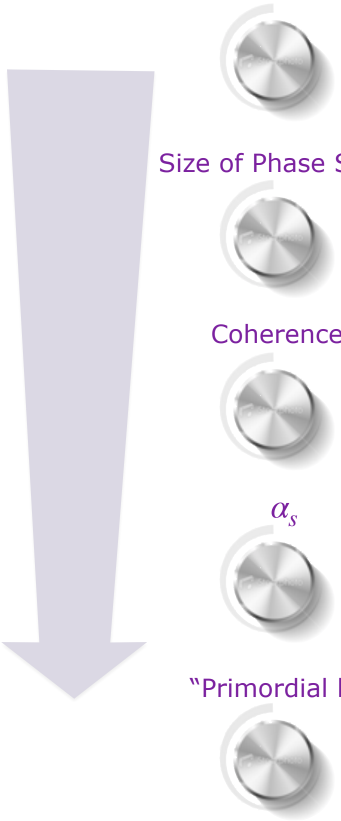
Baryon Multiplets

Baryons

Baryon-to-meson ratios, **Spin-3/2 vs Spin-1/2**,
“popcorn”, colour reconnections (junctions), … ?

Parameters (in PYTHIA): Initial-State Radiaton

+ PDF Choice



Matching & Merging

Additional Matrix Elements included?

At tree level / one-loop level? What matching scheme?

Size of Phase Space

Starting scale

Relation between Q_{PS} and Q_F (Vetoed showers? Suppressed? cf matching)

Coherence

Initial-Final interference

I-F colour-flow interference effects (eg VBF & Tevatron $t\bar{t}$ asym) & interleaving

α_s

Value and running of the strong coupling

Governs overall amount of radiation (cf FSR)

"Primordial kT"

A small additional amount of "unresolved" kT

Fermi motion + unresolved ISR emissions + low-x effects?

Minimum-Bias & Underlying Event

Number of MPI

Infrared Regularization scale $p_{\perp 0}$ for the QCD 2→2 (Rutherford) scatterings used for multiple parton interactions

→ average number of MPI, sets size of overall UE activity

Note: **strongly correlated with choice of PDF set!** (low-x gluon)

Pedestal Rise

Proton transverse mass distribution → difference between central (more active) **vs peripheral** (less active) **collisions**

Strings per Interaction

Color correlations between multiple-parton-interaction systems (aka colour *reconnections* — relative to LC)

→ shorter or longer strings → less or more hadrons per MPI

Affect $<p_T>$ vs N_{ch} balance: High CR ➜ fewer particles, each carrying more p_T

$\sqrt{s}$ scaling

Evolution of UE, $\langle \mathrm{d}N/\mathrm{d}\eta \rangle$, … with collider CM energy

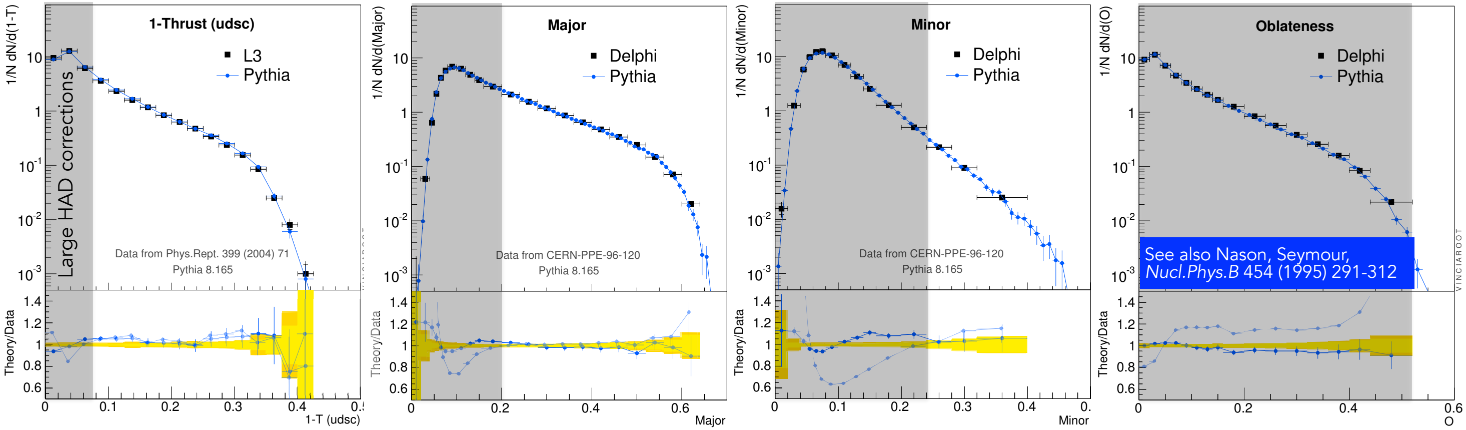
Cast as energy evolution of p_{T0} parameter.

IR Safe Observables: Sensitivity to Hadronization Parameters

PYTHIA 8 (hadronization on) **Vs** (hadronization off)

Important point: These observables are **IR safe ➜ minimal hadronisation corrections**

Big differences in **how** sensitive each of these are to hadronisation & over what **range**



Large HAD corrections

See also Nason, Seymour, *Nucl.Phys.B* 454 (1995) 291-312

The shaded bins provide constraints for the non-perturbative tuning stage.
You want your hadronization power corrections to do the "right thing" eg at low Thrust.

Hadronization Corrections: Fragmentation Tuning

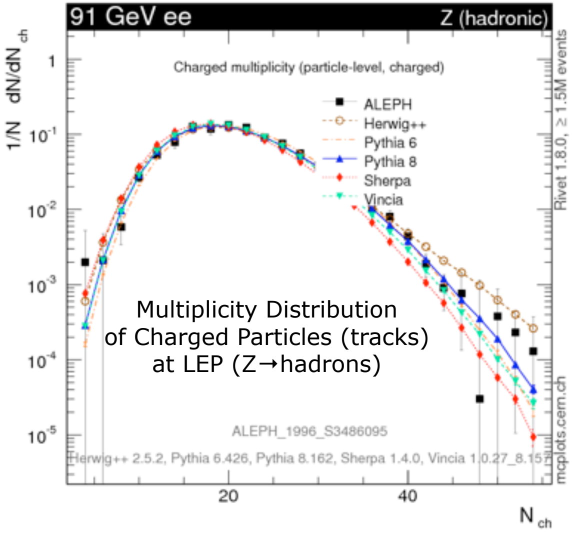
Now use infrared **sensitive** observables - sensitive to hadronization +
first few bins of previous (IR safe) ones

Longitudinal FF parameters *a* and *b*.

Transverse p_T broadening in string breaks (curtails high-N tail, and significantly affects event shapes)

Further parameter a_{diquark} requires looking at a baryon spectrum

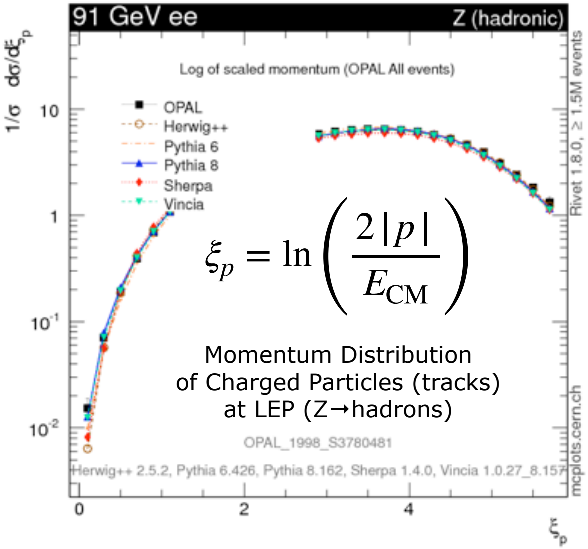
How many hadrons do you get?



Multiplicity Distribution
of Charged Particles (tracks)
at LEP (Z→hadrons)

$\langle N_{ch}(M_Z)\rangle \sim 21$

And how much momentum do they carry?



$$\xi_p = \ln\left(\frac{2|p|}{E_{\mathrm{CM}}}\right)$$

Momentum Distribution
of Charged Particles (tracks)
at LEP (Z→hadrons)

Practical Example: Uncertainties on Dark-Matter Annihilation Spectra

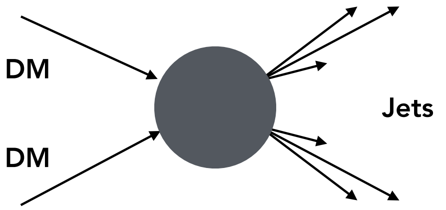
Based on A. Jueid et al., 1812.07424 (gamma rays, *eg for GCE*) and 2202.11546 (antiprotons, *eg for AMS*) + **2303.11363 (all)**

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

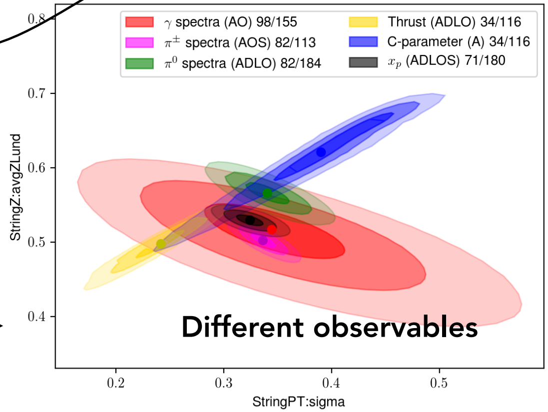
Simple sanity limit / overfit protection / tension resolution:

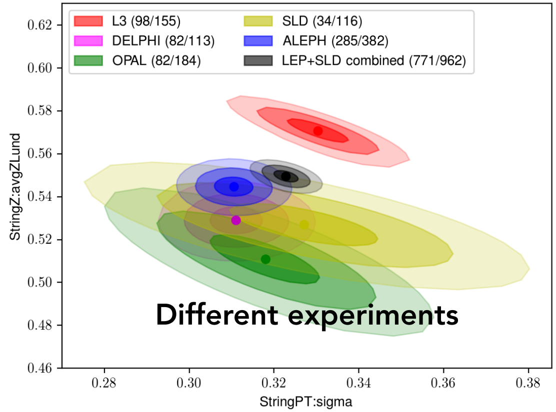
Add blanket 5% baseline uncertainty

(+ exclude 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

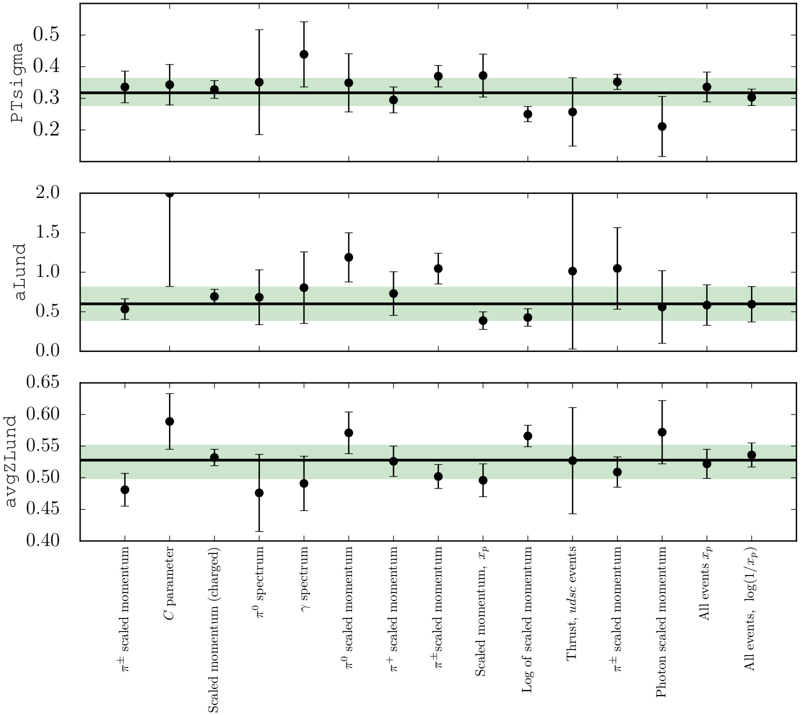
+ Universality Tests:





Practical Example: Uncertainties on Dark-Matter Annihilation Spectra

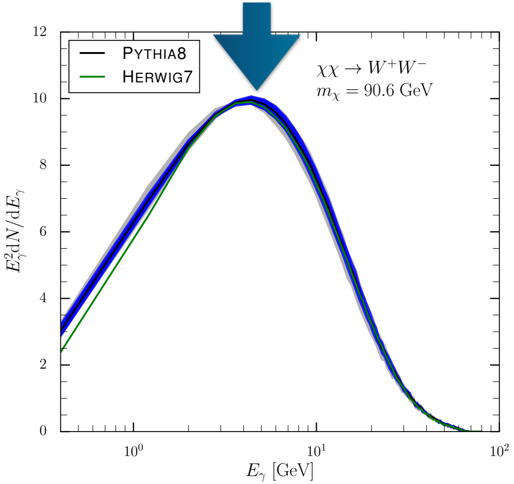
Based on A. Jueid et al., 1812.07424 (gamma rays, *eg for GCE*) and 2202.11546 (antiprotons, *eg for AMS*) + **2303.11363 (all)**



Weighted Average: good consistency across observables

Expensive?

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



Main Contact: adil.jueid@gmail.com

Same done for antiprotons, positrons, antineutrinos

Tables with uncertainties available on request. Also the spanning tune parameters of course.

Fragmentation Tuning — Know what Physics Goes In

Somewhat sensitive to particle composition:
heavier hadrons are harder!

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

+ particle decays

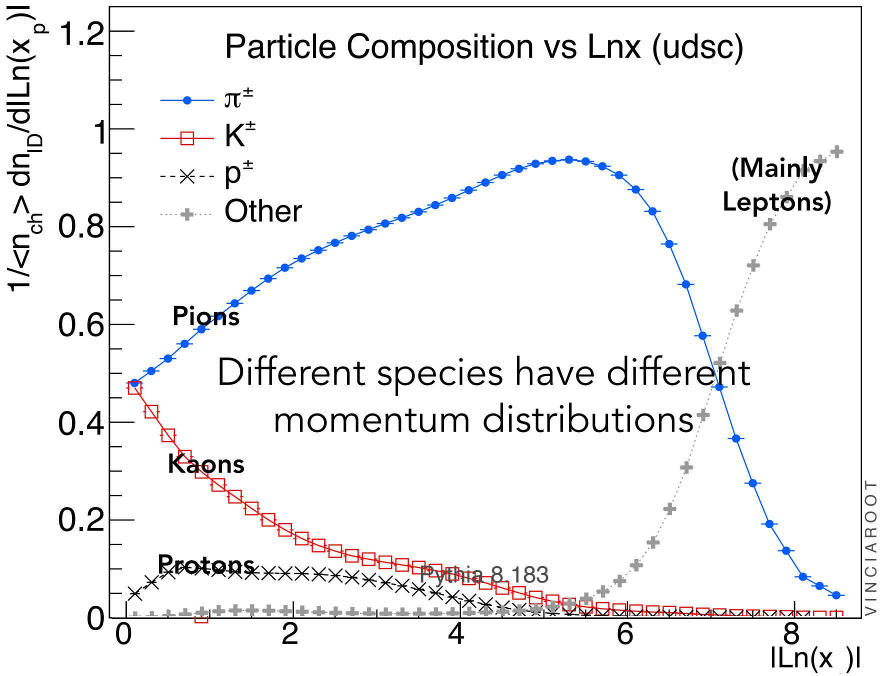
→ effects of feed-down!

$$\rho \to \pi\pi$$

$$K^* \to K\pi$$

$$\eta \to \pi\pi\pi$$

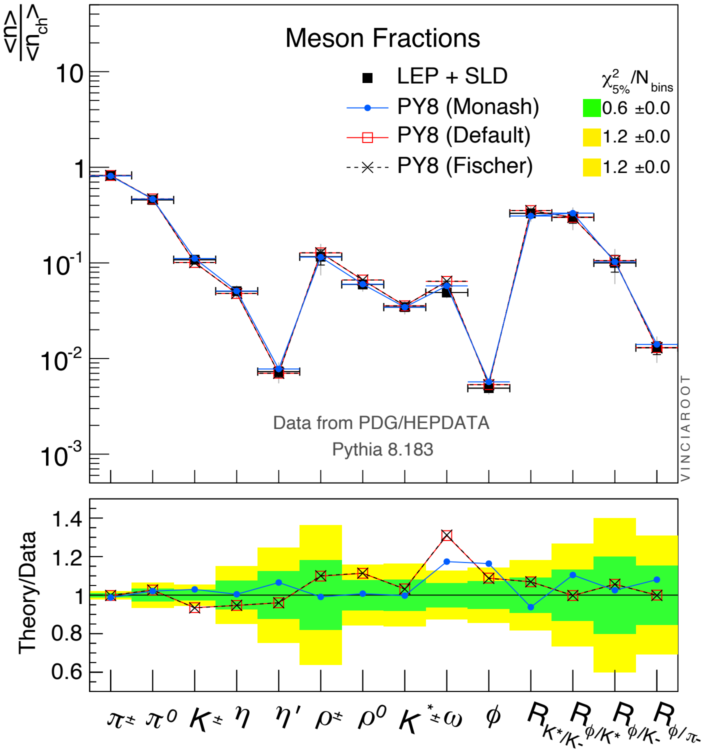
…

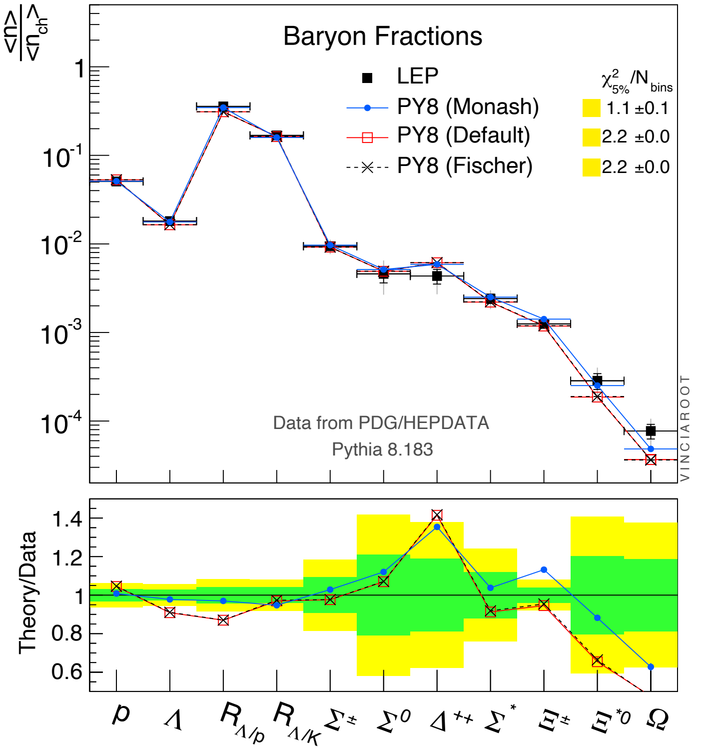


Different species have different momentum distributions

Meson and Baryon Rates and Ratios

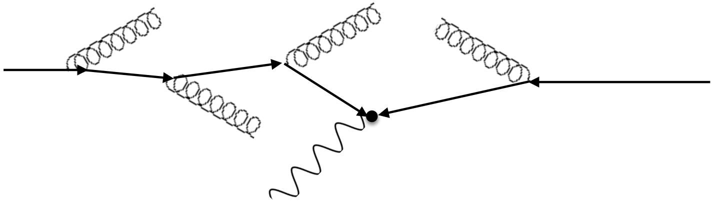
From PS et al., "Tuning PYTHIA 8.1: the Monash 2013 Tune", Eur.Phys.J.C 74 (2014) 8

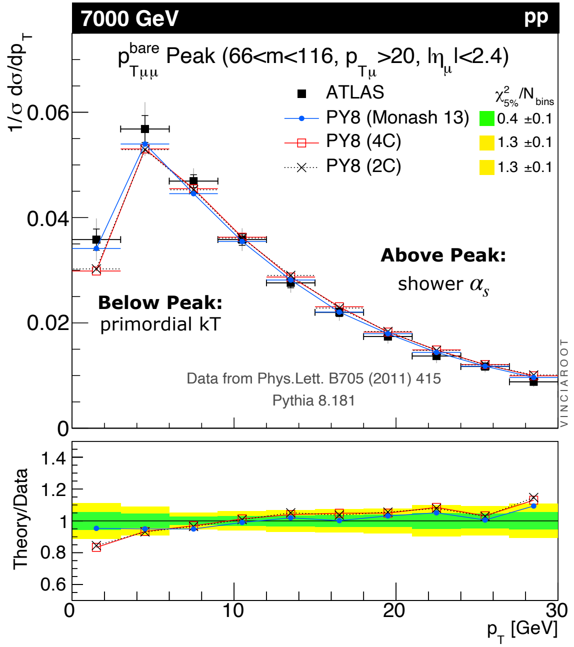




ISR + Primordial kT

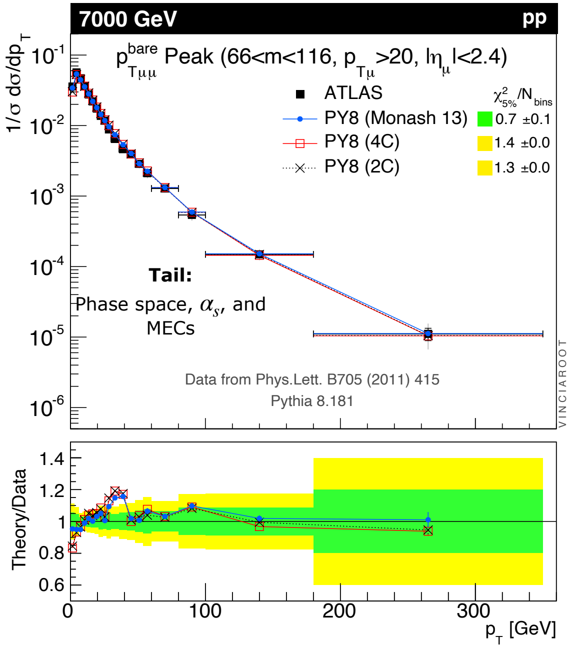
Drell-Yan pT distribution





Below Peak: primordial kT

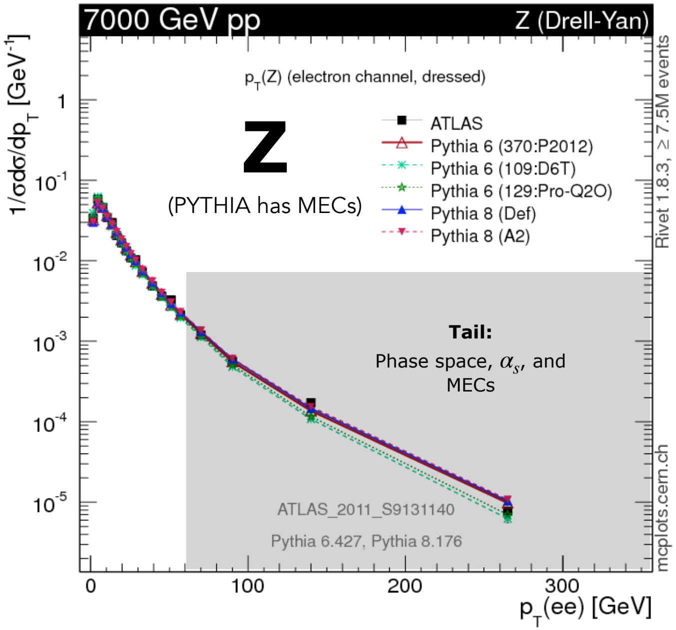
Above Peak: shower α_s

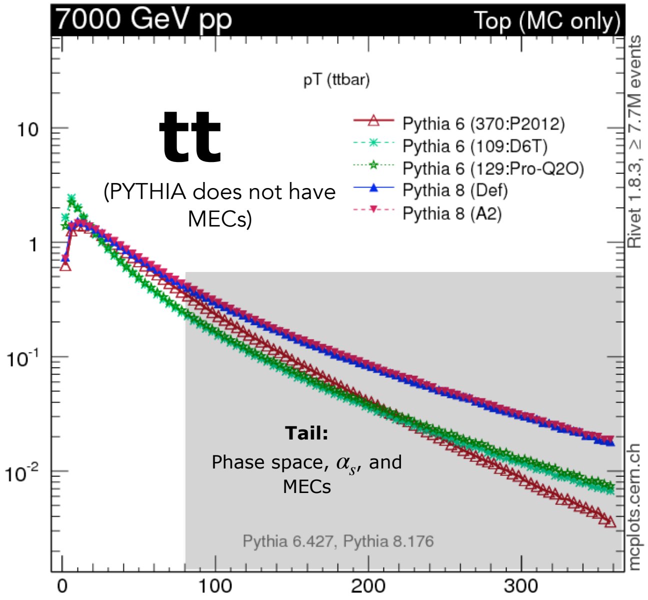


Tail:
Phase space, α_s, and MECs

Controlling for Process Dependence!

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





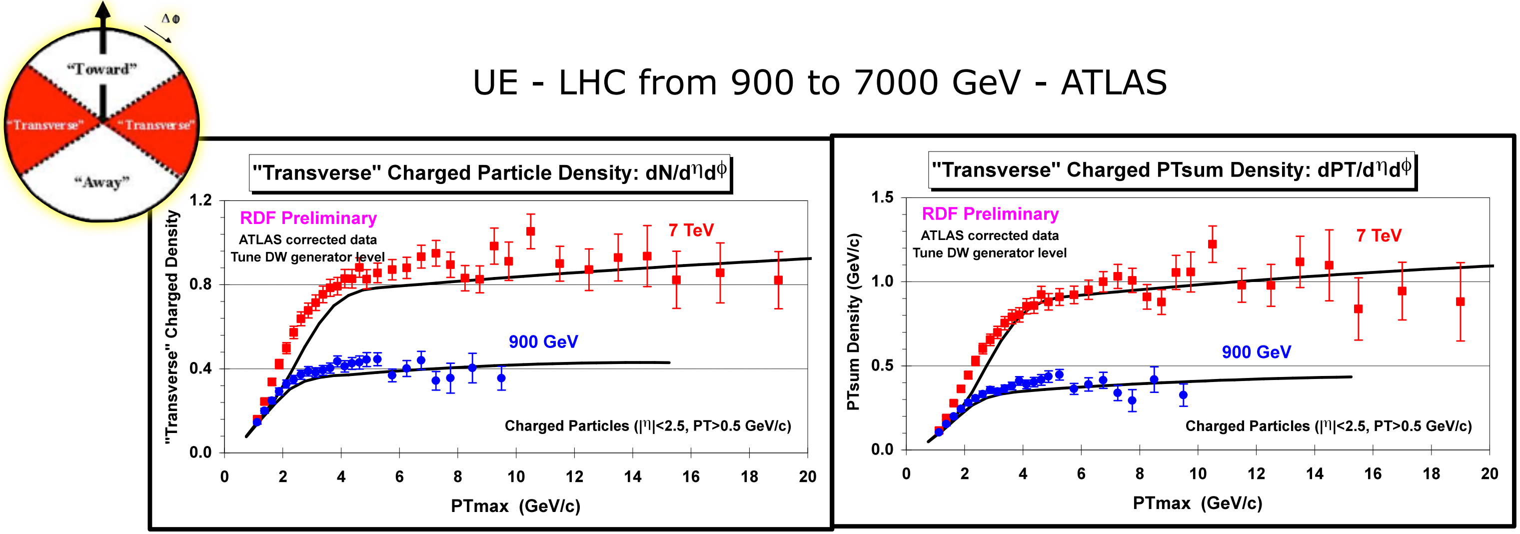
These points are quite sensitive to MECs / Matching / Merging.

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

Underlying Event

Same thing as before: how many particles do you get? And how much p_T do they carry?

UE - LHC from 900 to 7000 GeV - ATLAS



As you trigger on progressively higher p_T, the entire event increases …

… until you reach a plateau ("max-bias") also called the "jet pedestal" effect

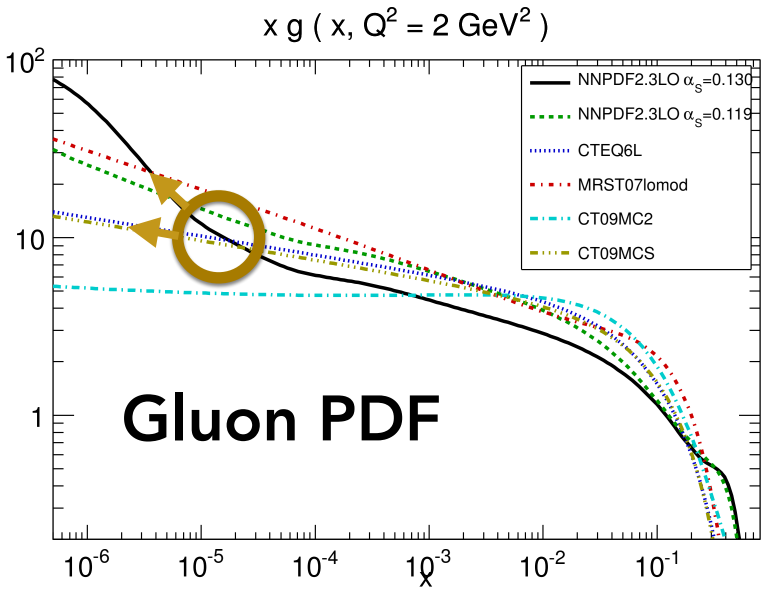
Interpreted as impact-parameter effect

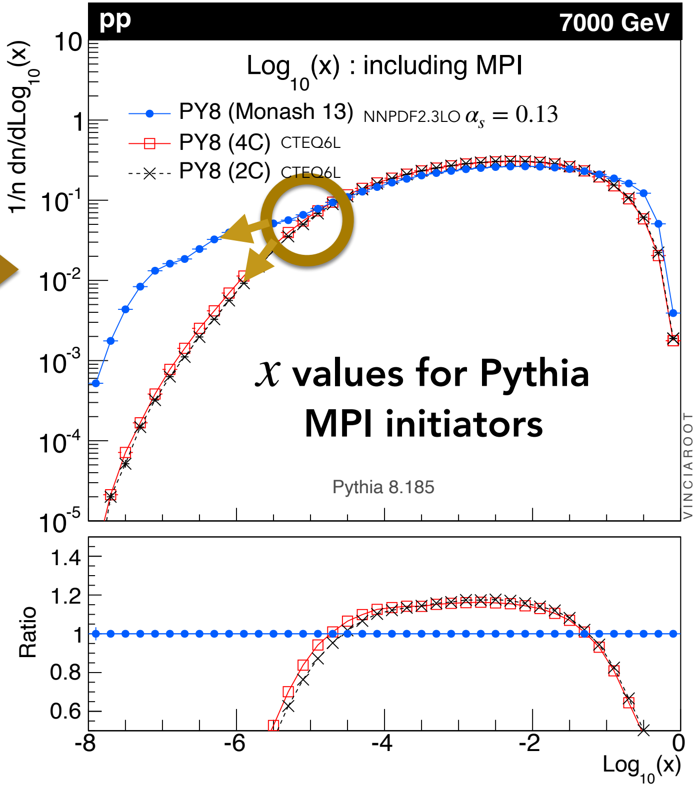
Qualitatively reproduced by MPI models

Relative size of this plateau / min-bias depends on pT0, PDF, *and* b-profile

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 Tyranny of Carlo” [J. D. Bjorken, ca. 1990]

> “Another change that I find disturbing is the rising tyranny of Carlo. No, I don’t mean that fellow who runs CERN [Rubbia], but the other one, with first name Monte.
>
> The simultaneous increase in detector complexity and in computation power has made simulation techniques an essential feature of contemporary experimentation. The MC simulation has become the major means of visualization of not only detector performance but also of physics phenomena. **So far so good.**
>
> But it often happens that the physics simulations provided by the MC generators carry the authority of data itself. They look like data and feel like data, and if one is not careful they are accepted as if they were data. **All Monte Carlo codes come with a GIGO* warning label**. But that warning label is just as easy for a physicist to ignore as that little message on a packet of cigarettes is for a chain smoker to ignore. I see nowadays experimental papers that claim agreement with QCD (translation: someone’s simulation labeled QCD) and/or disagreement with an alternative piece of physics (translation: an unrealistic simulation), without much evidence of the **inputs into those simulations**.”

Treat Tuning & Validation Studies with **same scientific rigour**
as any other scientific endeavour

---

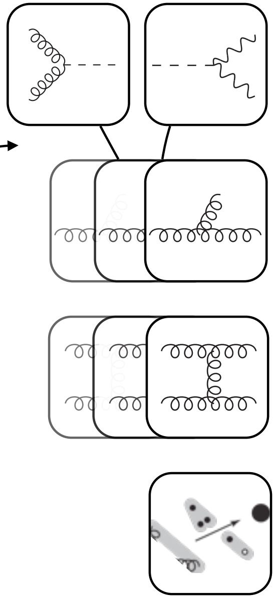
*GIGO: Garbage In, Garbage Out

Event Simulation — Summary

Physics Separation of time scales ➤ Factorizations Maths

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

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



Merging

Eliminate double-counting between fixed-order and shower corrections

Hard Process & Decays:

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

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

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

Driven by differential (e.g., DGLAP) evolution equations, dP/dQ², as function of resolution scale; from $Q_{\rm HARD}$ to $Q_{\rm HAD}$ ~ 1 GeV

MPI (Multi-Parton Interactions)

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

Hadronisation

Non-perturbative modeling of partons → hadrons transition

Strings or clusters; followed by hadron and τ decays

Final Words

MCs *can* be treated as black boxes, without knowing what's in them.

Best Case: Limited Sophistication

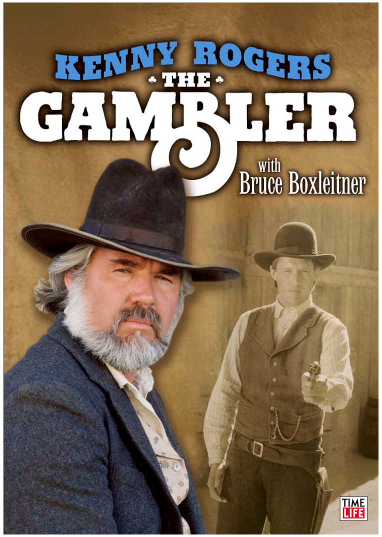
Worst Case: Not your lucky day

The key to successful Monte Carlo:

In the words of Kenny Rogers

Knowing what to throw away

Knowing what to keep



Kenny Rogers "The Gambler", first recorded in 1978

Same year as the first version of PYTHIA (JETGEN)