Seminar at DESY
1 March 2011

Progress on Event Generation with PYTHIA 8

Torbjörn Sjöstrand

Department of Astronomy and Theoretical Physics
Lund University

Introduction and Overview

Improved Showers

MultiParton Interactions

BSM Physics

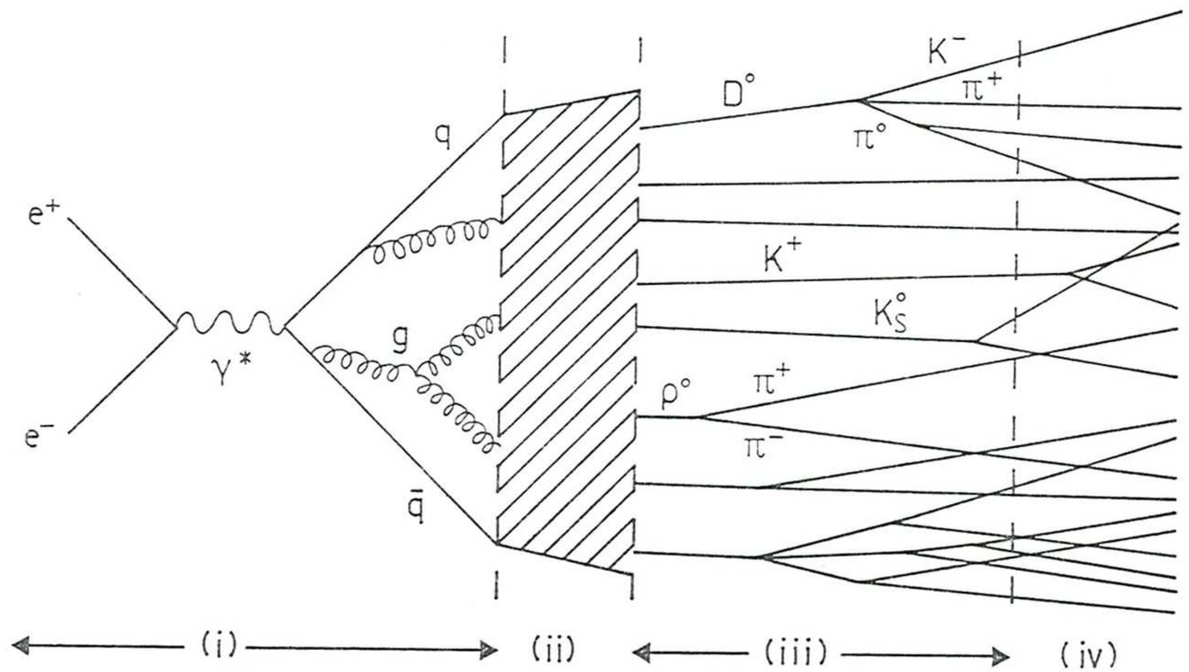
Tunes and Comparisons with LHC Data

Summary and Outlook

Introduction and Overview

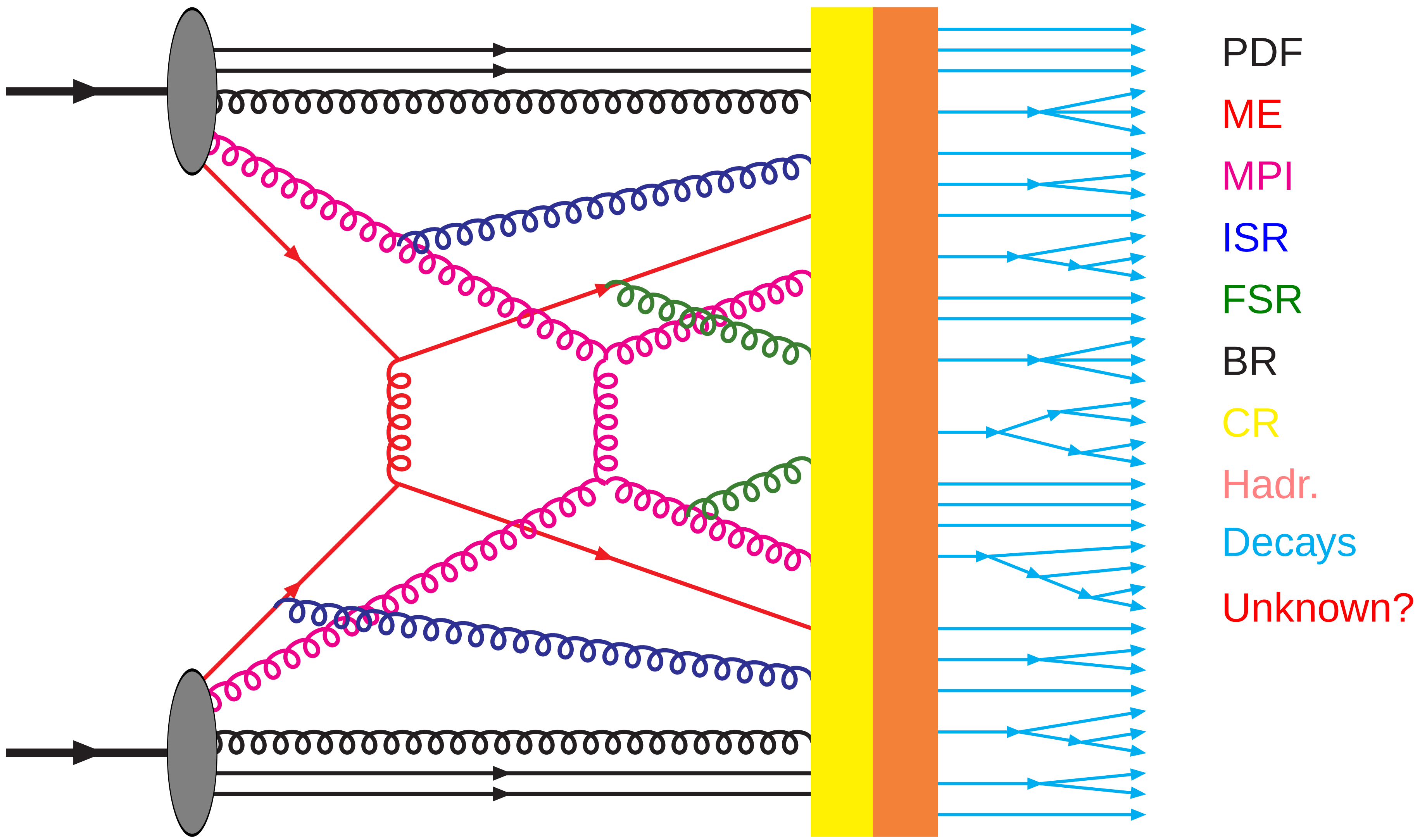
Modern event generators were born at DESY,
for the PETRA e^+e^- collider! $(1978 - 86, 13 - 46 \text{ GeV})$

- Combine perturbative picture of hard processes, involving electroweak and strong interactions, with nonperturbative picture of hadronization.
- Provide "complete" events, with parameters to be tuned to data, and used to study and understand different kinds of physics.



JETSET (PYTHIA predecessor): ~1,000 lines of Fortran code in 1980

Events more messy at the LHC (even when simplified):



General-purpose event generators: PYTHIA, HERWIG, SHERPA

PYTHIA size: $\sim$80,000 lines (Fortran in PYTHIA 6, C++ in PYTHIA 8)

PYTHIA 8 ambition

- Meet experimental request for C++ code.
- Housecleaning $\Rightarrow$ more homogeneous.
- More user-friendly (e.g. settings names).
- Better match to software frameworks (e.g. card files).
- More space for growth.
- Better interfaces to external standards.

Reality

- Work begun autumn 2004.
- 3 years at CERN $\Rightarrow$ good progress.
- First release autumn 2007.
- Since then: slower progress, requests lagging behind.
- Usage slowly taking off.

Team members

Stefan Ask
Richard Corke
Stephen Mrenna
Peter Skands

Contributors

Bertrand Bellenot
Lisa Carloni
Hendrik Hoeth
Philip Ilten
Tomas Kasemets
Mikhail Kirsanov
Ben Lloyd
Marc Montull
Sparsh Navin
MSTW, CTEQ, H1: PDFs
DELPHI, LHCb: D/B BRs
+ several bug reports & fixes

Key differences between PYTHIA 6.4 and 8.1

Old features definitely removed include, among others:

- independent fragmentation
- mass-ordered showers

Features omitted so far include, among others:

- ep, γp and $\gamma\gamma$ beam configurations
- several processes, especially Technicolor, partly SUSY

New features, not found in 6.4:

- $\star$ fully interleaved $p_\perp$-ordered MPI + ISR + FSR evolution
- $\star$ richer mix of underlying-event processes (γ, J/ψ, DY, . . .)
- $\star$ possibility for two selected hard interactions in same event
- $\star$ allow rescattering and x-dependent proton size in MPI framework
- $\star$ full hadron–hadron collision machinery for diffractive systems
- $\star$ several new processes, within and beyond SM
- $\star$ possibility to use one PDF set for hard process and another for rest
- $\star$ τ lepton polarization in production and decay
- $\star$ updated decay data and LO PDF sets

Interleaved evolution

- Transverse-momentum-ordered parton showers for ISR and FSR
- MPI also ordered in $p_\perp$

Allows interleaved evolution for ISR, FSR and MPI

$$\frac{\mathrm{d}\mathcal{P}}{\mathrm{d}p_\perp} = \left(\frac{\mathrm{d}\mathcal{P}_{\mathrm{MPI}}}{\mathrm{d}p_\perp} + \sum \frac{\mathrm{d}\mathcal{P}_{\mathrm{ISR}}}{\mathrm{d}p_\perp} + \sum \frac{\mathrm{d}\mathcal{P}_{\mathrm{FSR}}}{\mathrm{d}p_\perp}\right) \times \exp\left(-\int_{p_\perp}^{p_{\perp\mathrm{max}}} \left(\frac{\mathrm{d}\mathcal{P}_{\mathrm{MPI}}}{\mathrm{d}p'_\perp} + \sum \frac{\mathrm{d}\mathcal{P}_{\mathrm{ISR}}}{\mathrm{d}p'_\perp} + \sum \frac{\mathrm{d}\mathcal{P}_{\mathrm{FSR}}}{\mathrm{d}p'_\perp}\right) \mathrm{d}p'_\perp\right)$$

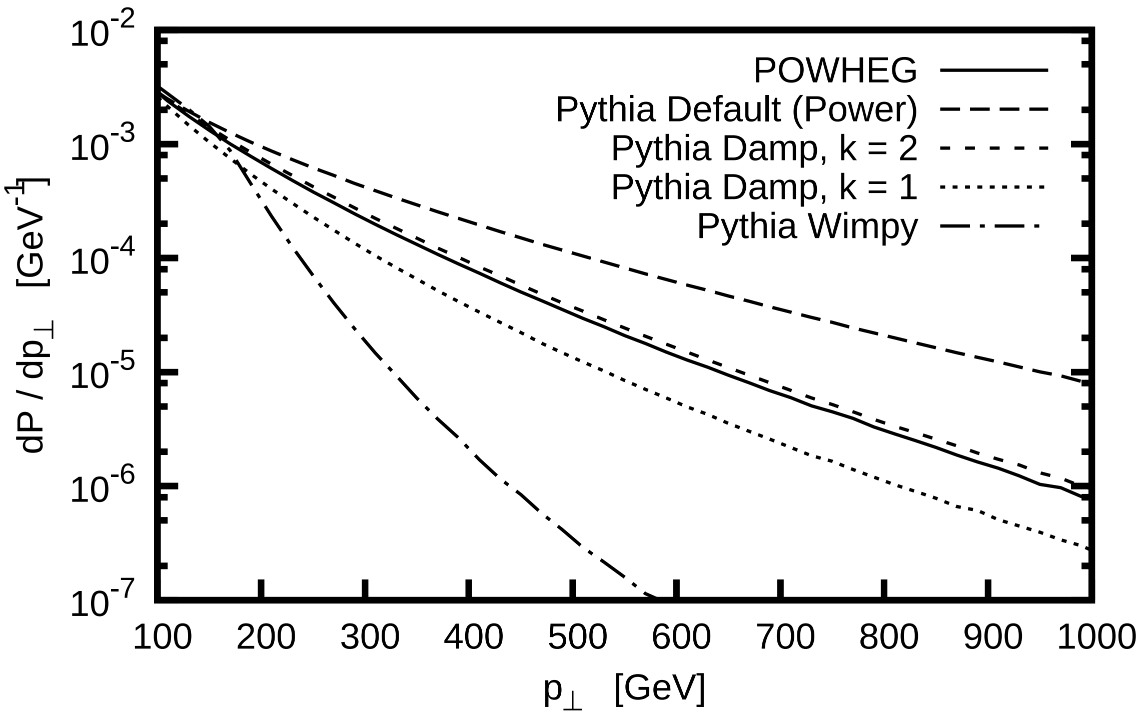
Ordered in decreasing $p_\perp$ using "Sudakov" trick.
Corresponds to increasing "resolution":
smaller $p_\perp$ fill in details of basic picture set at larger $p_\perp$.

Hybrid approach to shower recoils:

- FSR is dipole: nearest colour-connected neighbour
- ISR is traditional: whole hard-scattering system affected
  (since ISR dipole gives wrong answer e.g. for $p_{\perp Z}$)

Shower matching to MEs – I

Aim: provide better default shower behaviour at large $p_\perp$,
to bridge gap between "power" and "wimpy" showers.



$t\bar{t}$ production

$$M^2 = m_{\perp t}^2$$
$$= m_t^2 + p_{\perp t}^2$$

$$\frac{\mathrm{d}P_{\mathrm{ISR}}}{\mathrm{d}p_\perp^2} \propto \frac{1}{p_\perp^2} \frac{k^2 M^2}{k^2 M^2 + p_\perp^2}$$ for coloured final state

No dampening for uncoloured final state (W^+W^-, ..., SUSY).

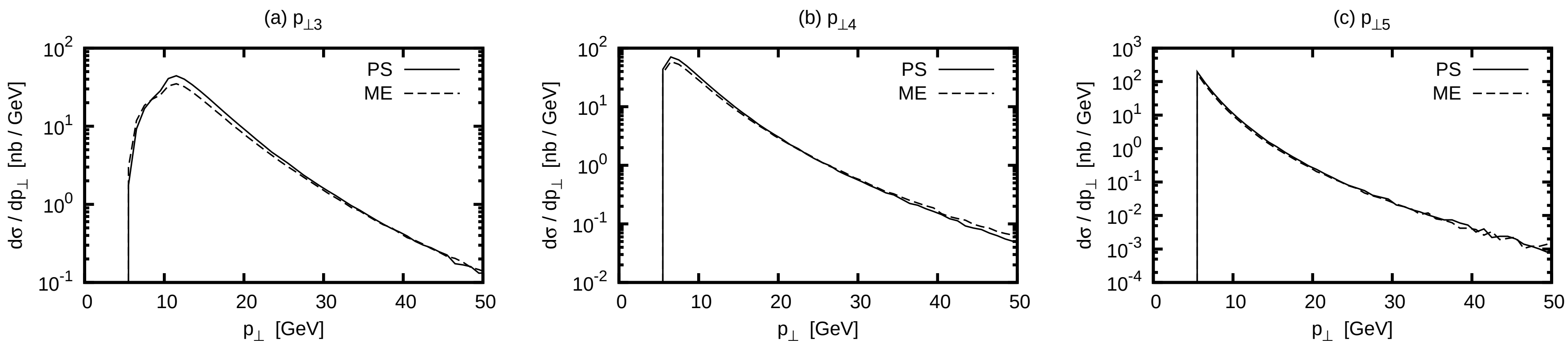
R. Corke & TS, Eur. Phys. J. C69 (2010) 1

(+ improved interfacing to POWHEG, ...)

Shower matching to MEs – II

Must avoid doublecounting for QCD jets:
shower starting scale = $p_\perp$ of hard $2 \to 2$ process.

Study how well the parton shower fills the phase space,
as prelude to full matching to $2 \to 3$ real-emission:

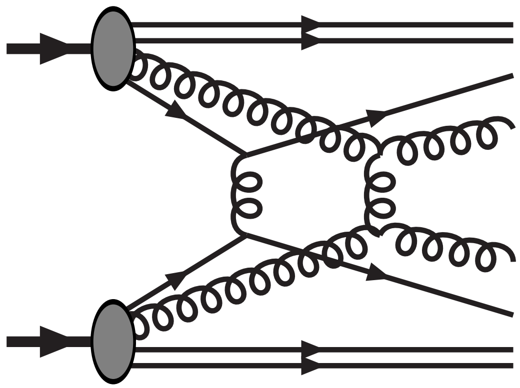


$$p_{\perp 3}^{\text{min}} = 5.0 \text{ GeV}, \; p_{\perp 4}^{\text{min}} = 5.0 \text{ GeV}, \; R_{\text{sep}} = 0.10$$

Obtain good qualitative agreement, best in soft and collinear regions,
but large region of phase space well described, and only corners bad.
No indication for needing a change in starting scale!
R. Corke & TS, arXiv:1011.1759 [hep-ph]

A second hard process

Multiparton interactions key aspect
of PYTHIA since $>$ 20 years.
Central to obtain agreement with data:
Tune A, Professor, Perugia, . . .



Before 8.1 no chance to select character of second interaction.
Now free choice of first process (including LHA/LHEF)
and second process combined from list:

- TwoJets (with TwoBJets as subsample)
- PhotonAndJet, TwoPhotons
- Charmonium, Bottomonium (colour octet framework)
- SingleGmZ, SingleW, GmZAndJet, WAndJet
- TopPair, SingleTop

Can be expanded among existing processes as need arises.

By default same phase space cuts as for “first” hard process
$\Longrightarrow$ second can be harder than first.
However, possible to set $\widehat{m}$ and $\widehat{p_{\perp}}$ range separately.

Multiparton interactions

Regularise cross section with $p_{\perp 0}$ as free parameter

$$\frac{\mathrm{d}\hat{\sigma}}{\mathrm{d}p_{\perp}^2} \propto \frac{\alpha_s^2(p_{\perp}^2)}{p_{\perp}^4} \to \frac{\alpha_s^2(p_{\perp 0}^2 + p_{\perp}^2)}{(p_{\perp 0}^2 + p_{\perp}^2)^2}$$

with energy dependence

$$p_{\perp 0}(E_{\mathrm{CM}}) = p_{\perp 0}^{\mathrm{ref}} \times \left(\frac{E_{\mathrm{CM}}}{E_{\mathrm{CM}}^{\mathrm{ref}}} \right)^{\epsilon}$$

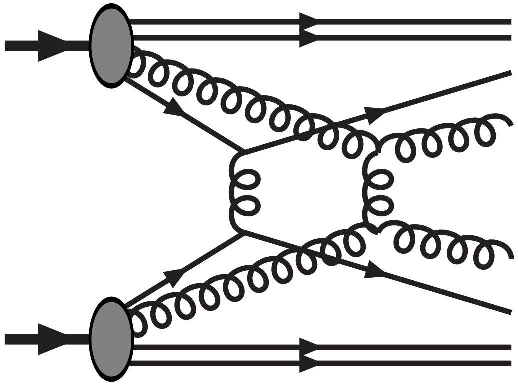
Matter profile in impact-parameter space
gives time-integrated overlap which determines level of activity:
simple Gaussian or more peaked variants

ISR and MPI compete for beam momentum $\longrightarrow$ PDF rescaling
$+$ flavour effects (valence, $\mathrm{q}\overline{\mathrm{q}}$ pair companions, ...)
$+$ correlated primordial $k_{\perp}$ and colour in beam remnant

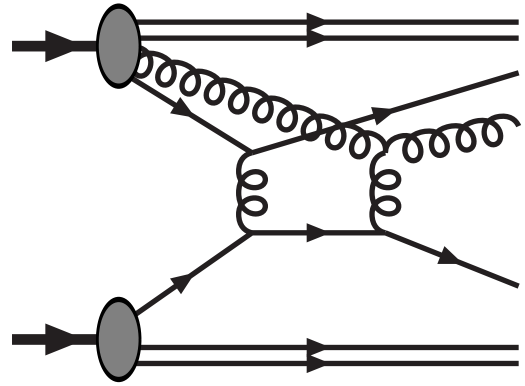
Many partons produced close in space–time $\Rightarrow$ colour rearrangement;
reduction of total string length $\Rightarrow$ steeper $\langle p_{\perp} \rangle (n_{\mathrm{ch}})$

Rescattering

Often assume that MPI =



. . . but should also include



Same order in α_s, $\sim$ same propagators, but

- one PDF weight less $\Rightarrow$ smaller σ
- one jet less $\Rightarrow$ QCD radiation background $2 \to 3$ larger than $2 \to 4$

$\Rightarrow$ will be tough to find direct evidence.

Rescattering grows with number of “previous” scatterings:

	Tevatron		LHC	
	Min Bias	QCD Jets	Min Bias	QCD Jets
Normal scattering	2.81	5.09	5.19	12.19
Single rescatterings	0.41	1.32	1.03	4.10
Double rescatterings	0.01	0.04	0.03	0.15

R. Corke & TS, JHEP 01 (2010) 035 [arXiv:0911.1909]

An x-dependent proton size

Normally assume that PDFs factorize in longitudinal and transverse space:

$$f(x,r) = f(x)\,\rho(r)$$

In contradiction with

- intuitive picture of partons spreading out by cascade to lower x
- formally BFKL, Balitsky-JIMWLK, Colour Glass Condensate, ...
- Mueller's dipole cascade (Lund program: DIPSY; study by Avsar)
- Froissart-Martin $\sigma_{\mathrm{tot}} \propto \ln^2 s$ by Gribov theory related to $r_{\mathrm{p}} \propto \ln(1/x)$
- generalized parton distributions, ...

For now address only inelastic nondiffrative events,
$\sigma_{\mathrm{tot}} = \sigma_{\mathrm{el}} + \sigma_{\mathrm{SD}} + \sigma_{\mathrm{DD}} + \sigma_{\mathrm{ND}}$, with ansatz:

$$\rho(r,x) \propto \frac{1}{a^3(x)} \exp\left(-\frac{r^2}{a^2(x)}\right) \quad \text{with} \quad a(x) = a_0\left(1 + a_1 \ln\frac{1}{x}\right)$$

$a_1 \approx 0.15$ tuned to **rise** of σ_{ND} in Donnachie & Landshoff + Schuler & TS
a_0 tuned to **value** of σ_{ND}, given PDF, $p_{\perp 0}$, ...

R. Corke & TS, arXiv:1101.5953

Convolution of two incoming protons gives impact parameter shape

$$\tilde{\mathcal{O}}(b; x_1, x_2) = \frac{1}{\pi} \frac{1}{a^2(x_1) + a^2(x_2)} \exp\left(-\frac{b^2}{a^2(x_1) + a^2(x_2)} \right)$$

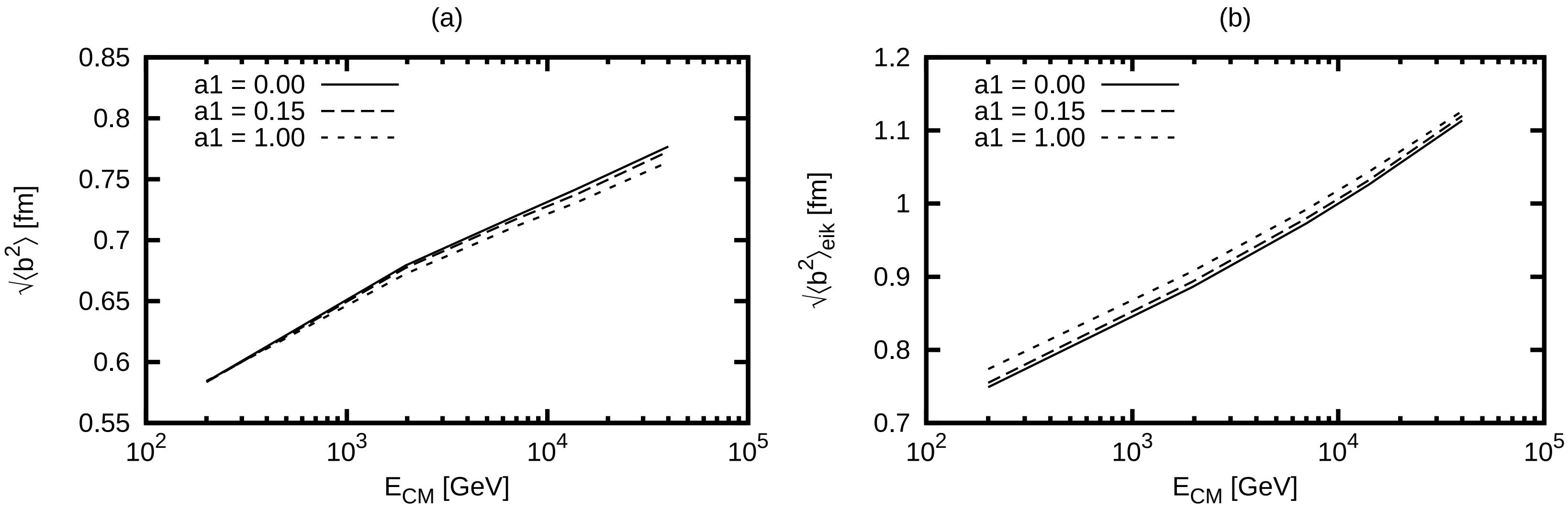
Define $\overline{n}(b)$ as average number of interactions for passage at b

$$\overline{n}(b) = \sum_{i,j} \iiint \mathrm{d}x_1 \, \mathrm{d}x_2 \, \mathrm{d}p_\perp^2 \, f_i(x_1, p_\perp^2) \, f_j(x_2, p_\perp^2) \, \left. \frac{\mathrm{d}\hat{\sigma}_{ij}}{\mathrm{d}p_\perp^2} \right|_{\mathrm{reg}} \tilde{\mathcal{O}}(b; x_1, x_2)$$

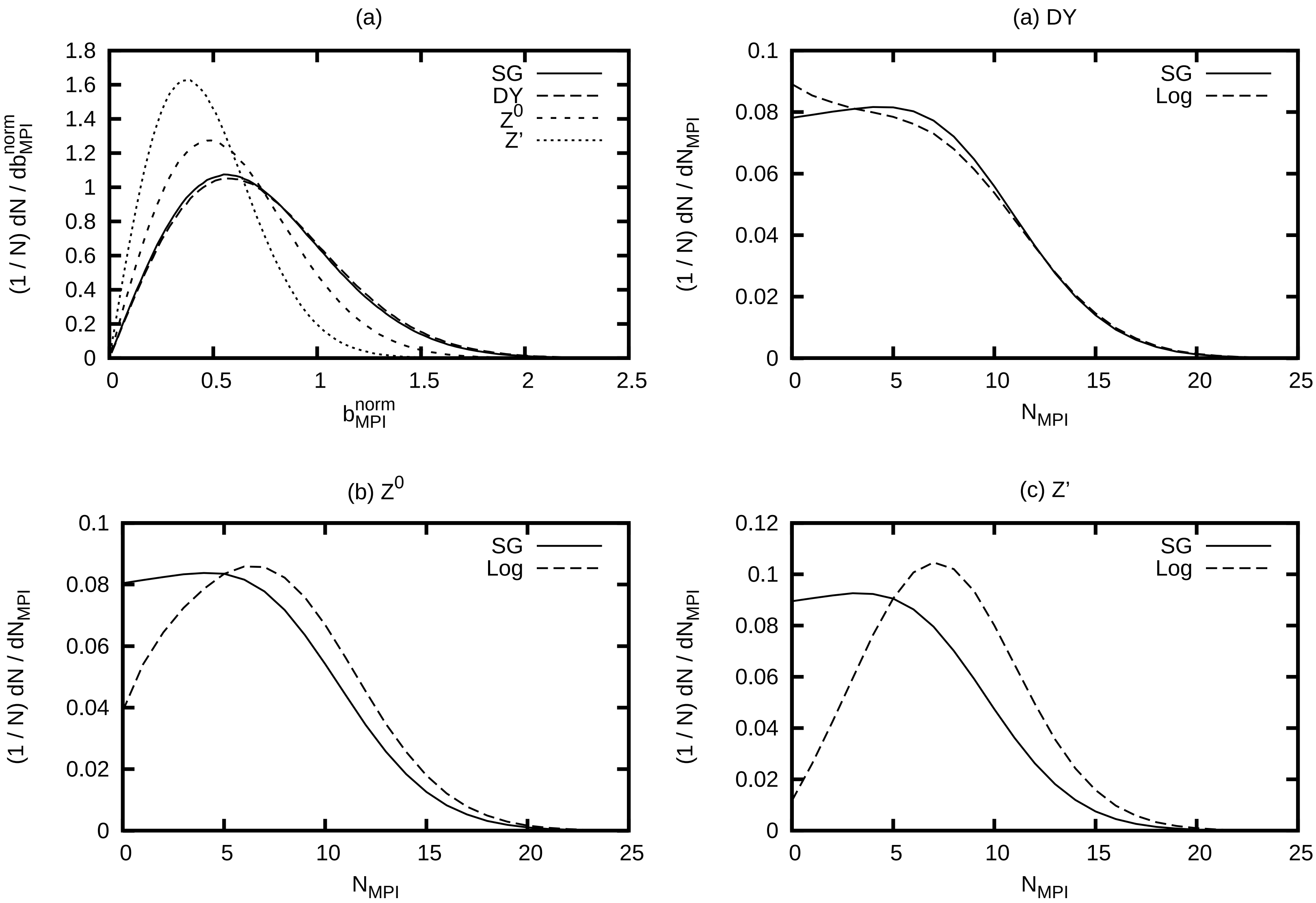
$$\text{such that } \sigma_{\mathrm{hard}} = \int \overline{n}(b) \, \mathrm{d}^2 b$$

$$\sigma_{\mathrm{ND}} = \int P_{\mathrm{int}} \, \mathrm{d}^2 b = \int \left(1 - e^{-\overline{n}(b)} \right) \mathrm{d}^2 b$$

which gives a $\sqrt{\langle b^2 \rangle}$ for σ_{hard} and σ_{ND}:

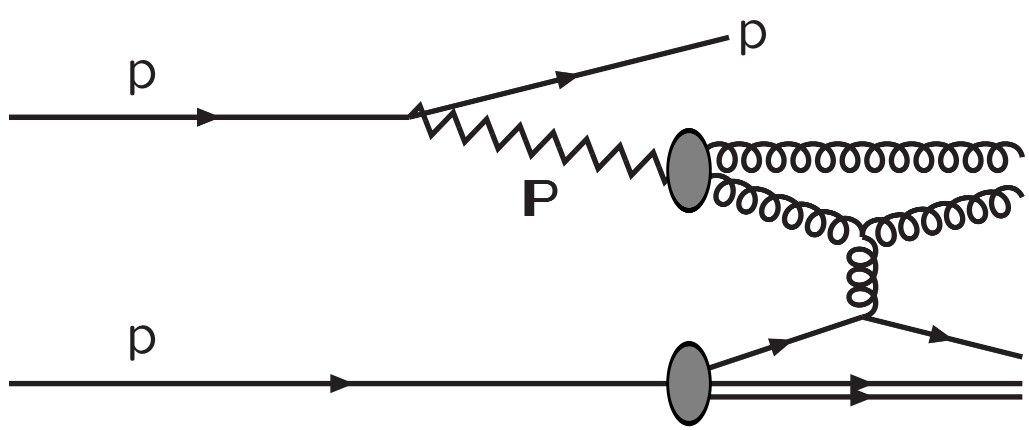


Consequence: collisions at large x will have to happen at small b, and hence further large-to-medium-x MPIs are enhanced, while low-x partons are so spread out that it plays less role.



Diffraction

Ingelman-Schlein: Pomeron as hadron with partonic content
Diffractive event = (Pomeron flux) $\times$ ($\mathbb{P}$p collision)



Used e.g. in
POMPYT
POMWIG
PHOJET

1) σ_{SD} and σ_{DD} taken from existing parametrization or set by user.

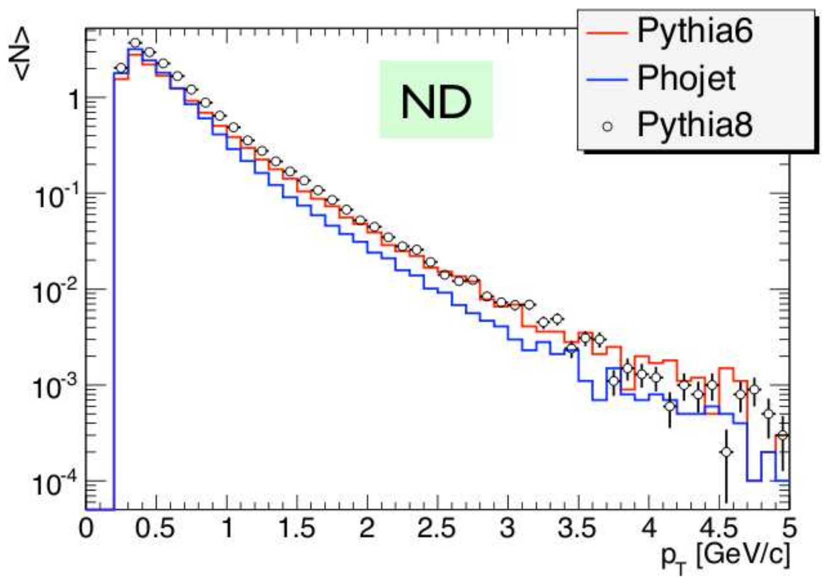
2) Shape of Pomeron distribution inside a proton, $f_{\mathbb{P}/\mathrm{p}}(x_{\mathbb{P}}, t)$ gives diffractive mass spectrum and scattering $p_{\perp}$ of proton.

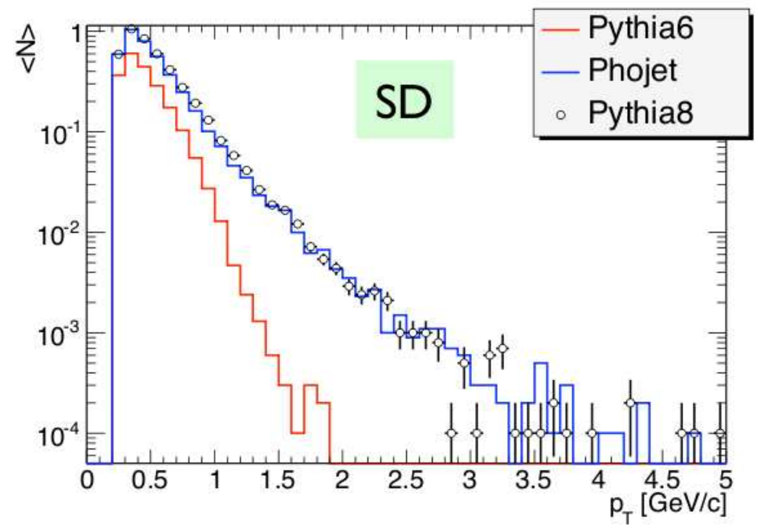
3) At low masses retain old framework, with longitudinal string(s). Above 10 GeV begin smooth transition to $\mathbb{P}$p handled with full pp machinery: multiple interactions, parton showers, beam remnants, ...

4) Choice between 5 Pomeron PDFs. Free parameter $\sigma_{\mathbb{P}\mathrm{p}}$ needed to fix $\langle n_{\mathrm{interactions}} \rangle = \sigma_{\mathrm{jet}}/\sigma_{\mathbb{P}\mathrm{p}}$.

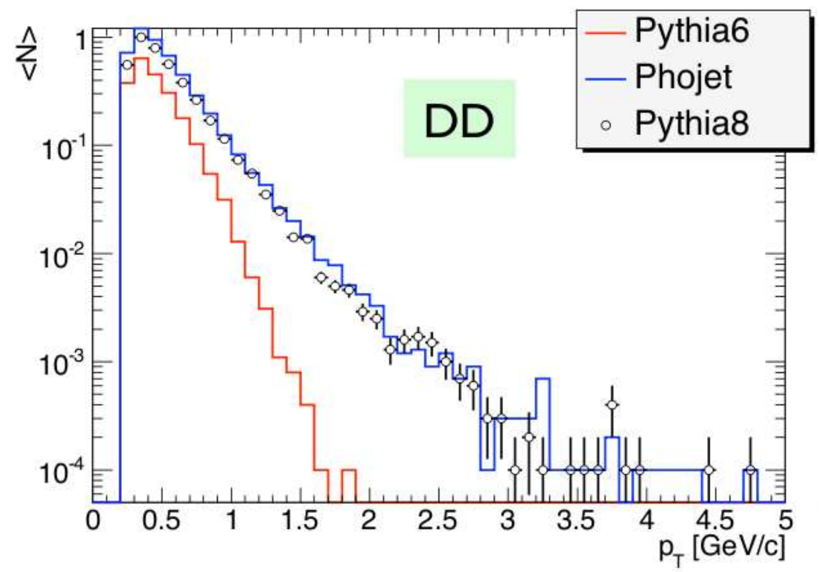
5) Framework needs testing and tuning, e.g. of $\sigma_{\mathbb{P}\mathrm{p}}$.

p_T Distributions ($\sqrt{s}$=0.9 TeV)





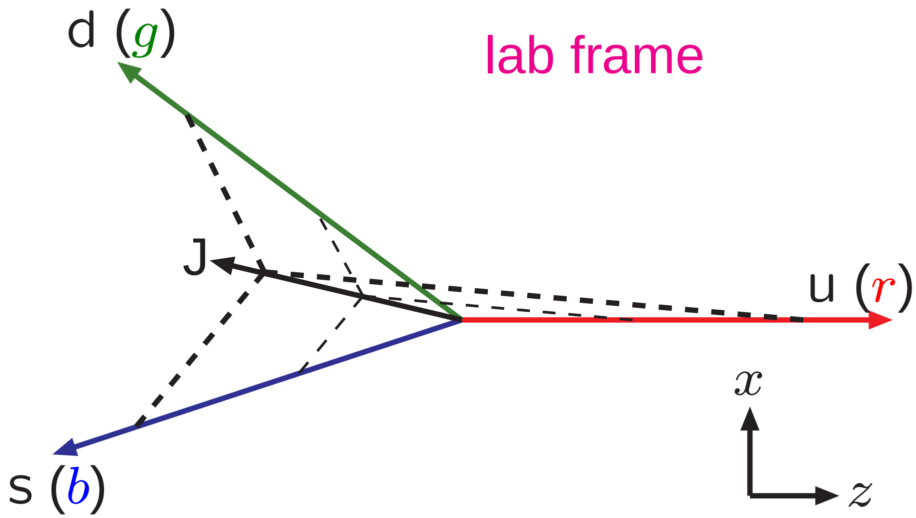
- Softer p_T spectrum in **Pythia6** due to lack of high mass diffraction
- **Pythia8** and **Phojet** agree quite well

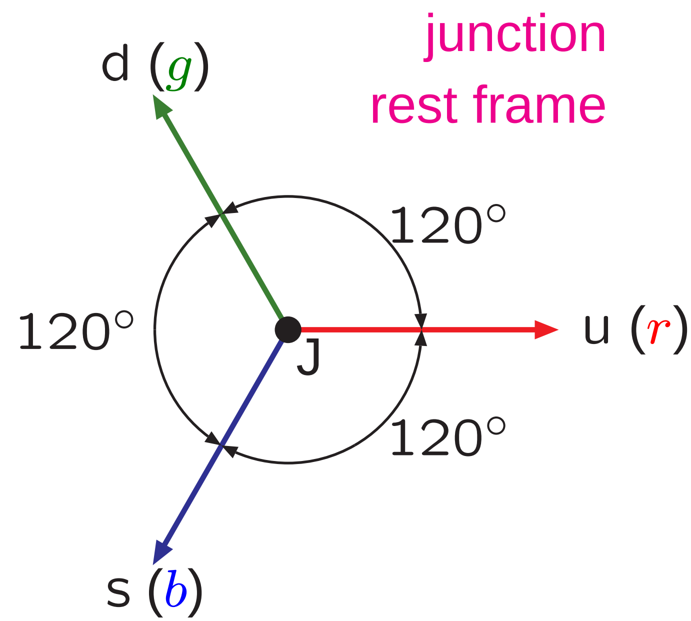


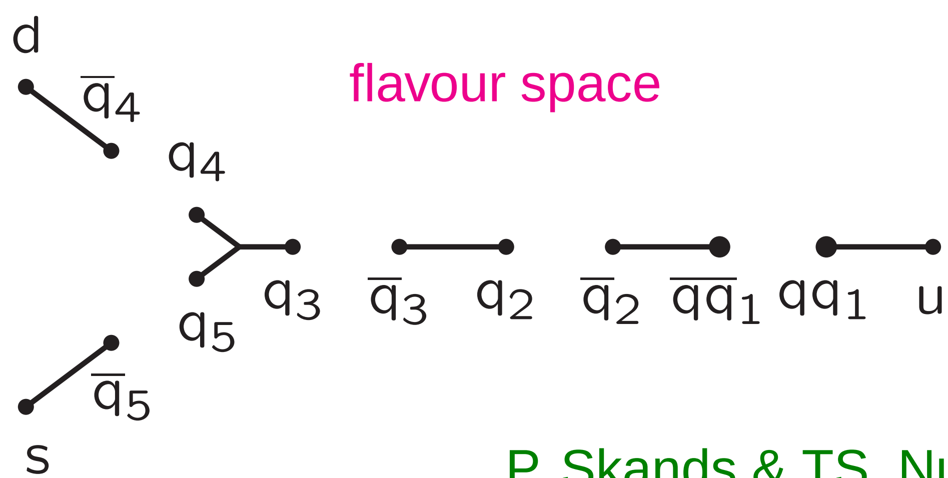
Beate Heinemann, MB/UE Working Group (also Sparsh Navin)

BSM Physics 1: R-parity violation

Encountered in R-parity violating SUSY decays $\tilde{\chi}_1^0 \to$ uds,
or when 2 valence quarks kicked out of proton beam







More complicated
(but ≈solved) with
gluon emission and
massive quarks

P. Skands & TS, Nucl. Phys. B659 (2003) 243

BSM Physics 2: R-hadrons

What if coloured (SUSY) particle like $\tilde{\mathrm{g}}$ or $\tilde{\mathrm{t}}_1$ is long-lived?

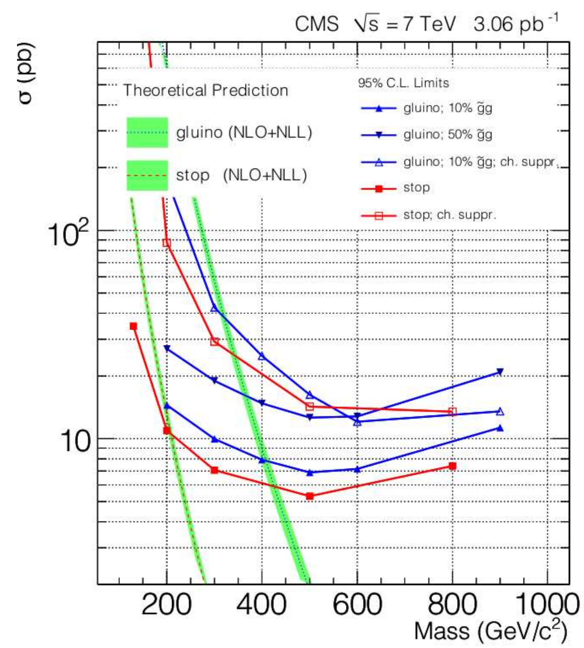
⋆ Formation of R-hadrons

$\tilde{\mathrm{g}}\mathrm{q}\overline{\mathrm{q}}$	$\tilde{\mathrm{t}}_1\overline{\mathrm{q}}$	“mesons”
$\tilde{\mathrm{g}}\mathrm{qqq}$	$\tilde{\mathrm{t}}_1\mathrm{qq}$	“baryons”
$\tilde{\mathrm{g}}\mathrm{g}$		“glueballs”

⋆ Conversion between R-hadrons
by “low-energy” interactions with matter:
$\tilde{\mathrm{g}}\mathrm{u}\overline{\mathrm{d}} + \mathrm{p} \to \tilde{\mathrm{g}}\mathrm{uud} + \pi^+$ irreversible

⋆ Displaced vertices if finite lifetime, or else

⋆ punch-through: $\sigma \approx \sigma_{\mathrm{had}}$ but
$\Delta E \lesssim 1$ GeV $\ll E_{\mathrm{kin},R}$



A.C. Kraan, Eur. Phys. J. C37 (2004) 91;
M. Fairbairn et al., Phys. Rep. 438 (2007) 1

CMS, arXiv:1101.1645

Partly event generation, partly detector simulation.
Public add-on in PYTHIA 6, soon integrated part of PYTHIA 8.

BSM Physics 3: Hidden Valley (Secluded Sector)

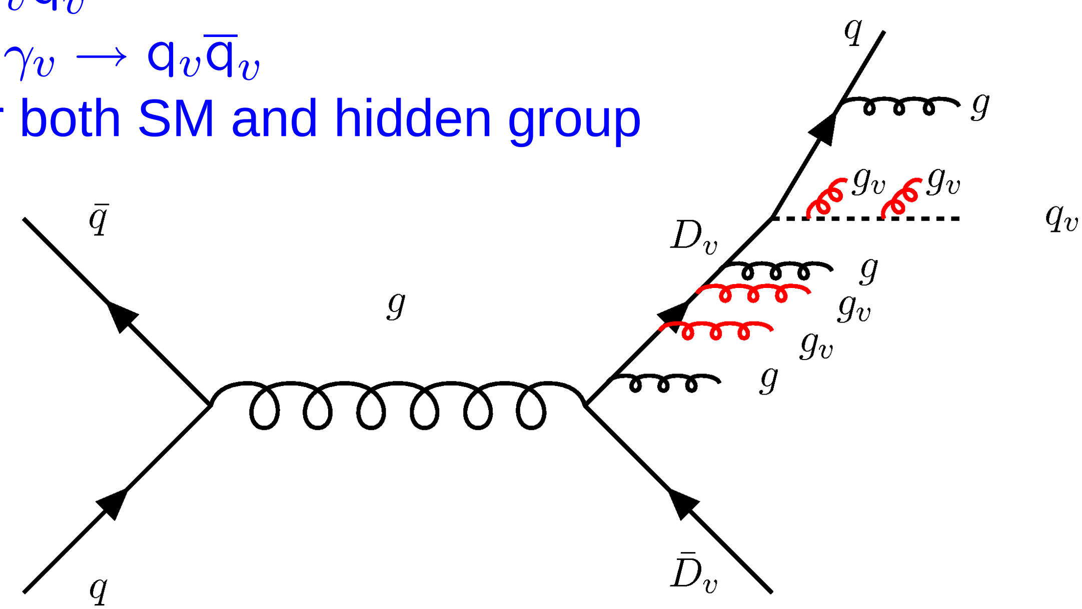
What if new gauge groups at low energy scales, hidden by potential barrier or weak couplings? (M. Strassler & K. Zurek, ...)

Complete framework implemented in PYTHIA:

⋆ New gauge group either Abelian $U(1)$ or non-Abelian $SU(N)$

⋆ 3 alternative production mechanisms
1) massive Z′: $q\overline{q} \to Z' \to q_v\overline{q}_v$
2) kinetic mixing: $q\overline{q} \to \gamma \to \gamma_v \to q_v\overline{q}_v$
3) massive F_v charged under both SM and hidden group

⋆ Interleaved shower in QCD, QED and HV sectors: add $q_v \to q_v\gamma_v$ (and F_v) or $q_v \to q_v g_v$, $g_v \to g_v g_v$, which gives recoil effects also in visible sector



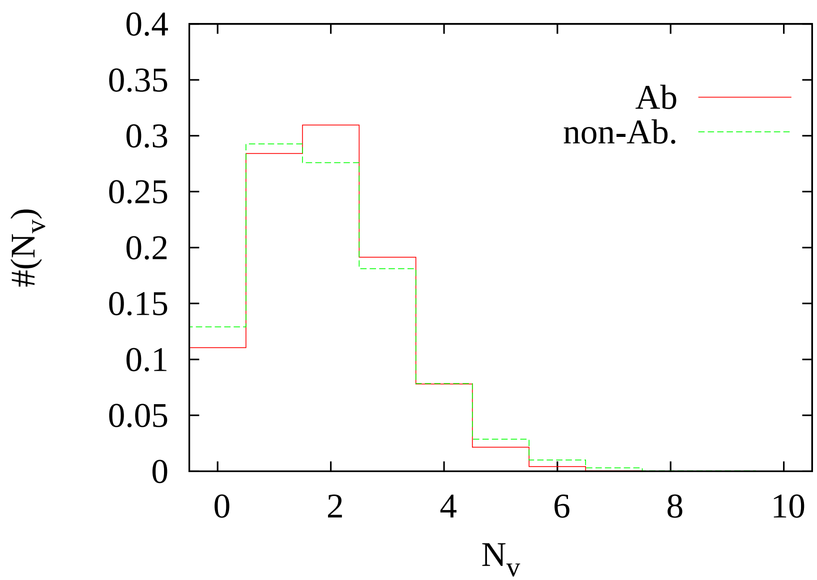
L. Carloni & TS, JHEP 1009 (2010) 105;
L. Carloni, J. Rathsman & TS, arXiv:1102.3795

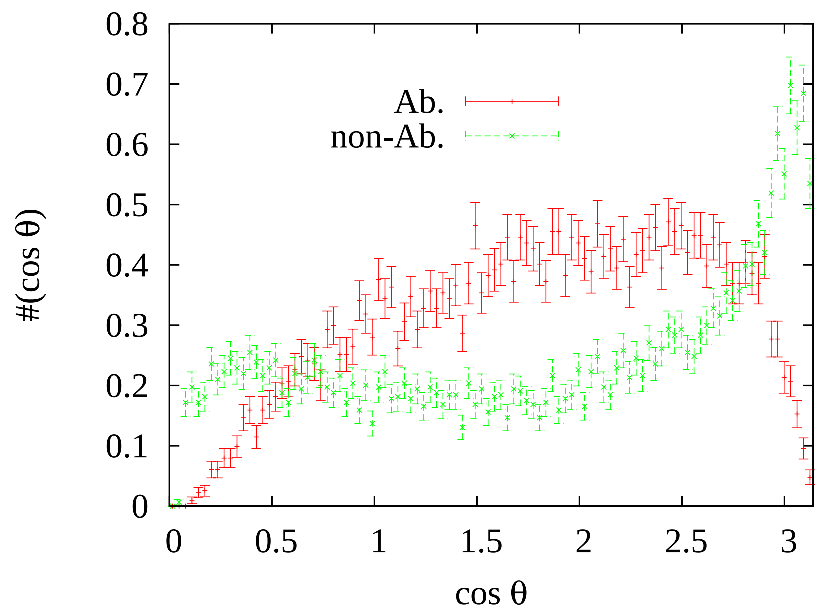
⋆ Hidden Valley particles may remain invisible, or . . .

⋆ Broken $U(1)$: γ_v acquire mass, radiated γ_vs decay back
$\gamma_v \to \gamma \to \mathrm{f}\overline{\mathrm{f}}$ with BRs as photon ($\Rightarrow$ lepton pairs!)

⋆ $SU(N)$: hadronization in hidden sector, with full string fragmentation,
permitting up to 8 different q_v flavours and 64 $\mathrm{q}_v\overline{\mathrm{q}}_v$ mesons,
but for now assumed degenerate in mass, so only distinguish
– off-diagonal, flavour-charged, stable & invisible
– diagonal, can decay back $\mathrm{q}_v\overline{\mathrm{q}}_v \to \mathrm{f}\overline{\mathrm{f}}$

Even when tuned to same average activity, hope to separate $U(1)$ and $SU(N)$:

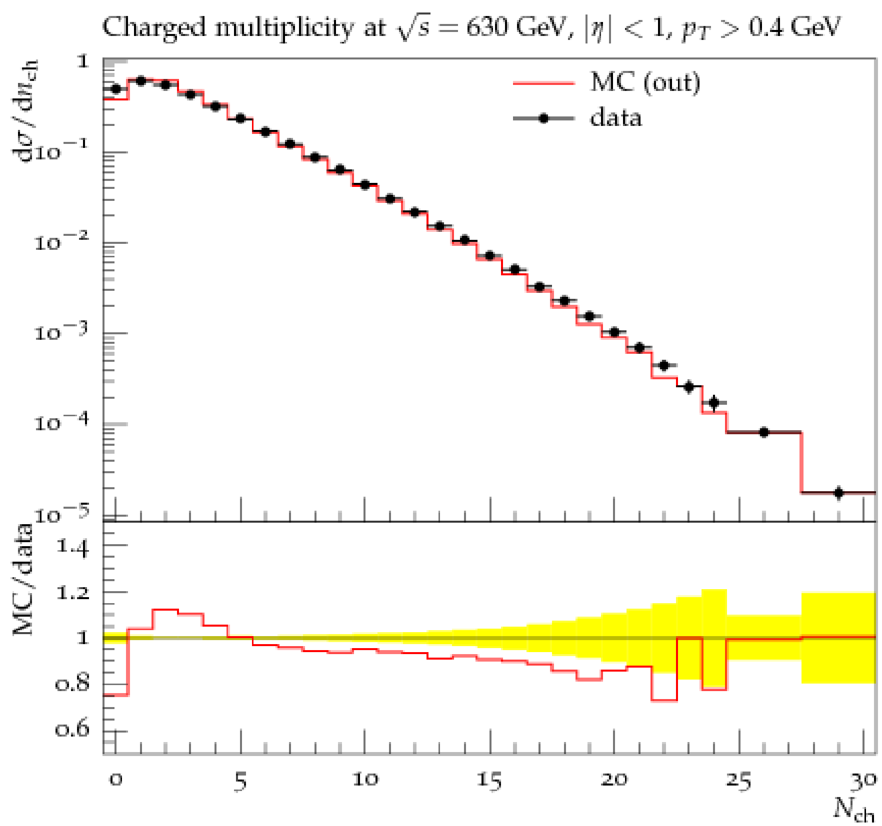


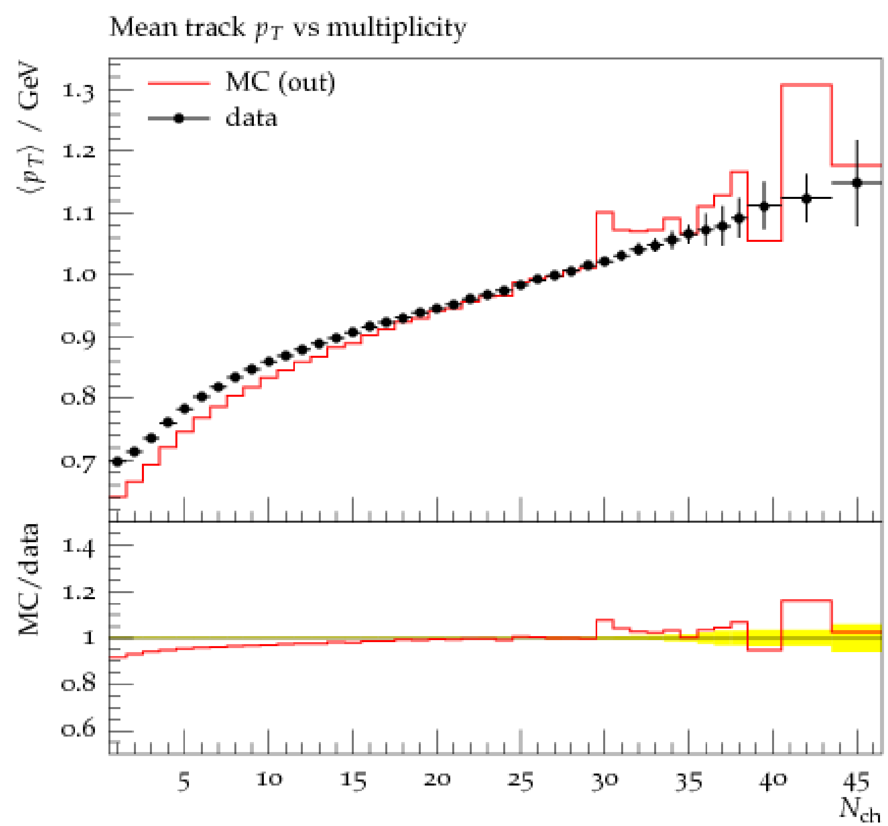


Tuning prospects

Tuning to e^+e^- closely related to $p_\perp$-ordered PYTHIA 6.4;
Rivet+Professor (H. Hoeth) $\Rightarrow$ FSR & hadronization OK (?)

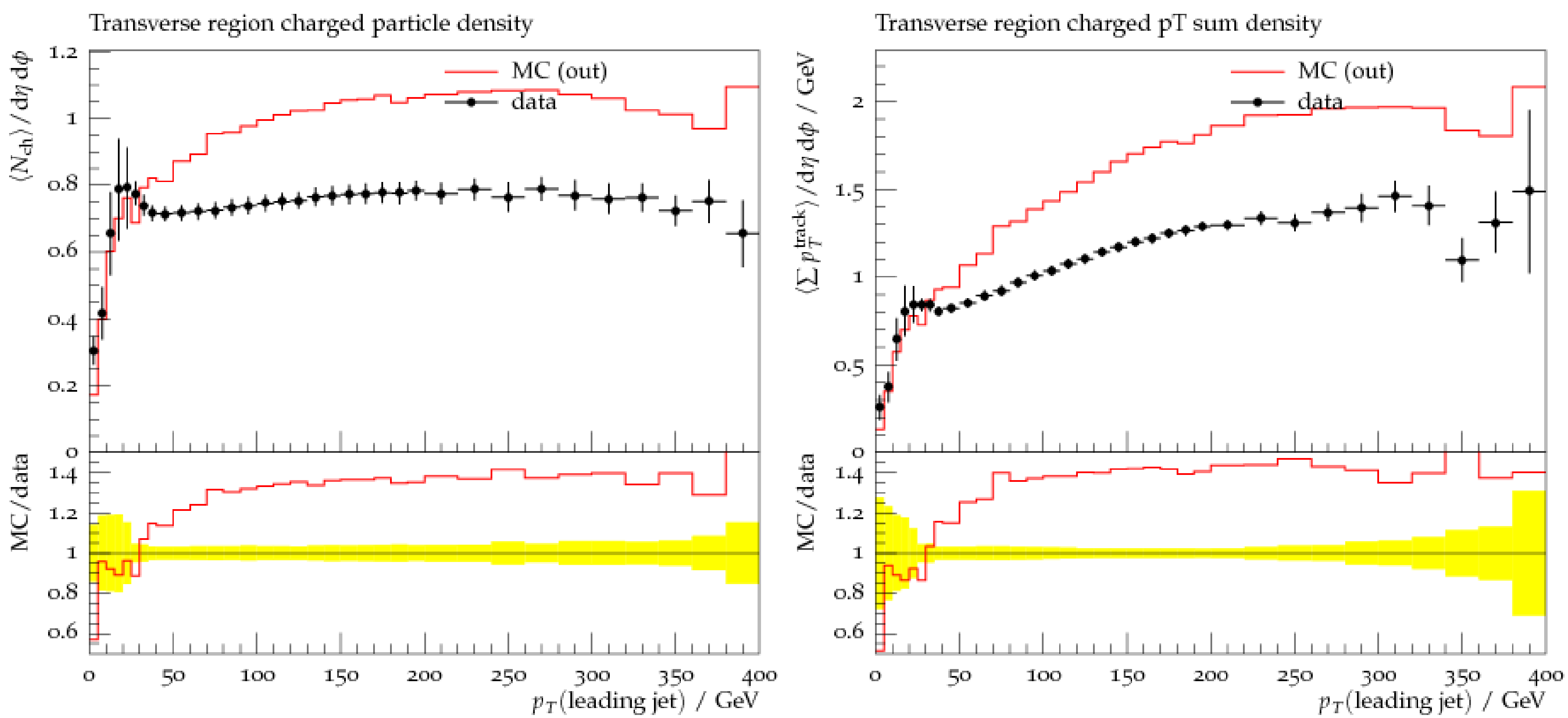
First tuning to MB data by P. Skands:





$\Rightarrow$ MPI & colour reconnection OK (?)

But Rivet+Professor (H. Hoeth) shows it fails miserably for UE
(Rick Field's transverse flow as function of jet $p_\perp$):



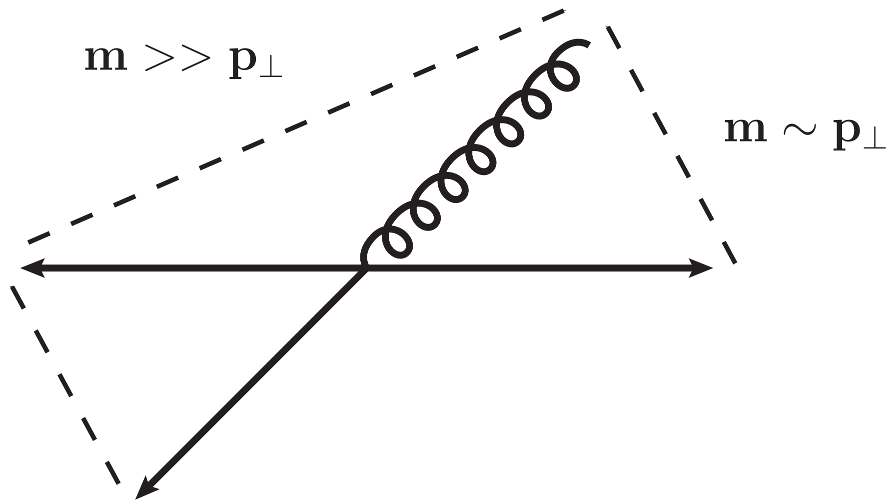
No universal tune MB + UE!

Where did we go wrong?

Final-state parton may have colour partner in the initial state.

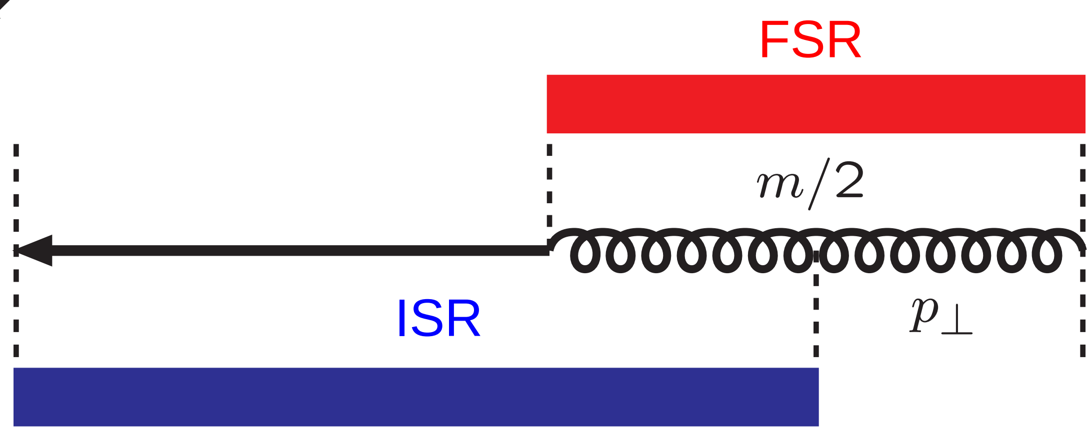
How to subdivide FSR and ISR in this kind of dipole?

Large mass $\longrightarrow$ large rapidity range for emission:



In dipole rest frame

(think rapidity space)

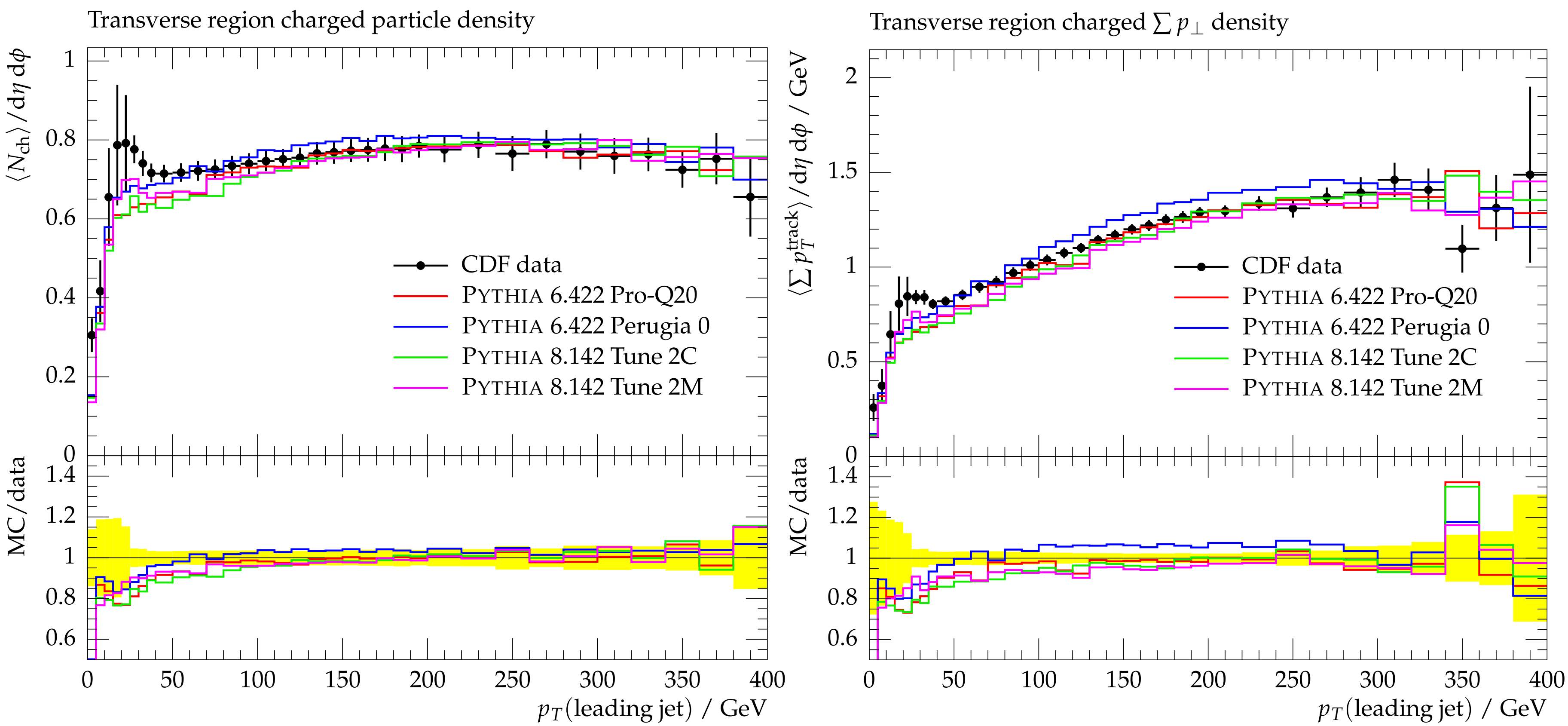


Solution: suppress final-state radiation in double-counted region

Is a simultaneous MB/UE Tevatron tune now possible?

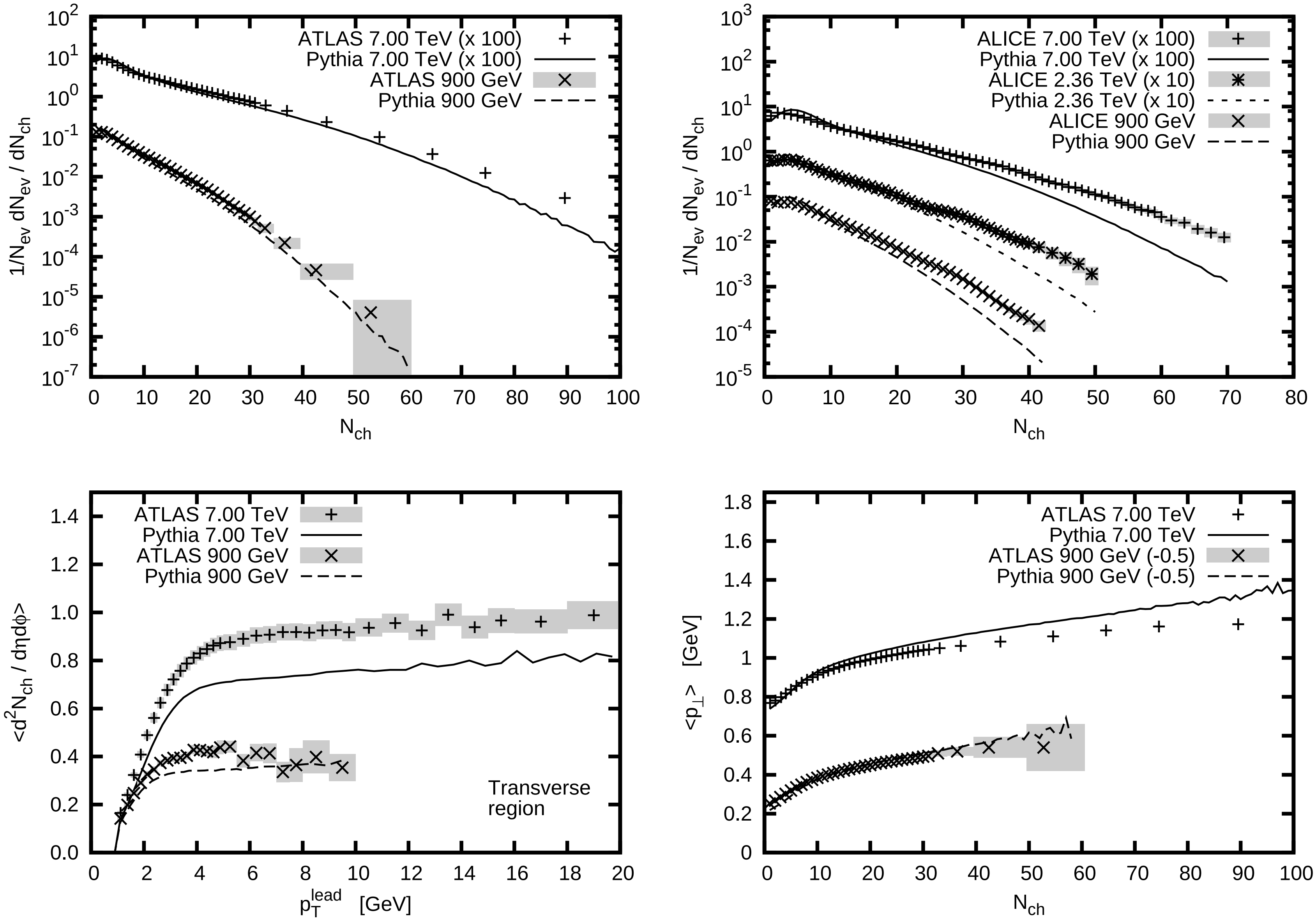
Tunes 2C and 2M done “by hand” (= using Rivet, but not Professor), using the CTEQ6L1 and MRST LO** PDF sets, respectively, to MB data (n_{ch}, $\langle p_\perp \rangle(n_{\mathrm{ch}})$, ...

Compare against Pro-Q20 and Perugia 0 (PYTHIA6)



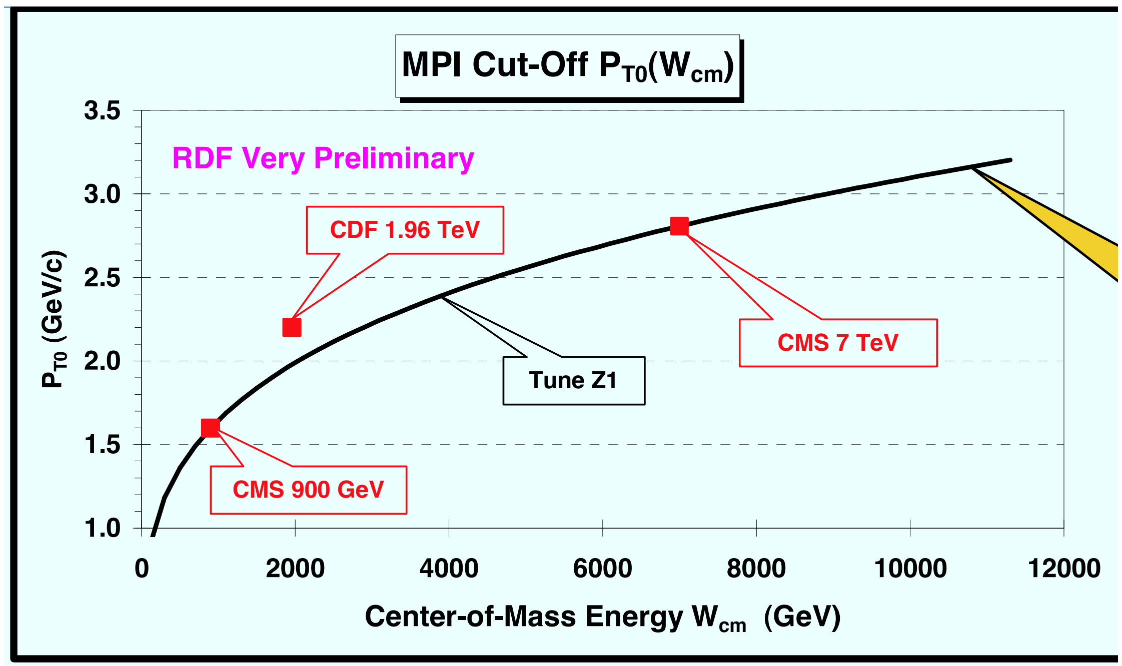
Now generally comparably good description to old tunes at the Tevatron.

Tune 2C applied to LHC data does not do so good:



Tension between Tevatron and LHC data?

Rick Field: if $p_{\perp 0}(E_{\mathrm{CM}}) \propto E_{\mathrm{CM}}^{\epsilon}$ tuned to LHC data,
then gives too much UE activity at Tevatron
($\Rightarrow$ need higher $p_{\perp 0}$ to compensate)



Pick some key LHC data sets, use Tune 2C as starting point:

- slightly dampen diffractive cross section (ATLAS)
- only vary MPI and colour reconnection parameters

New tunes

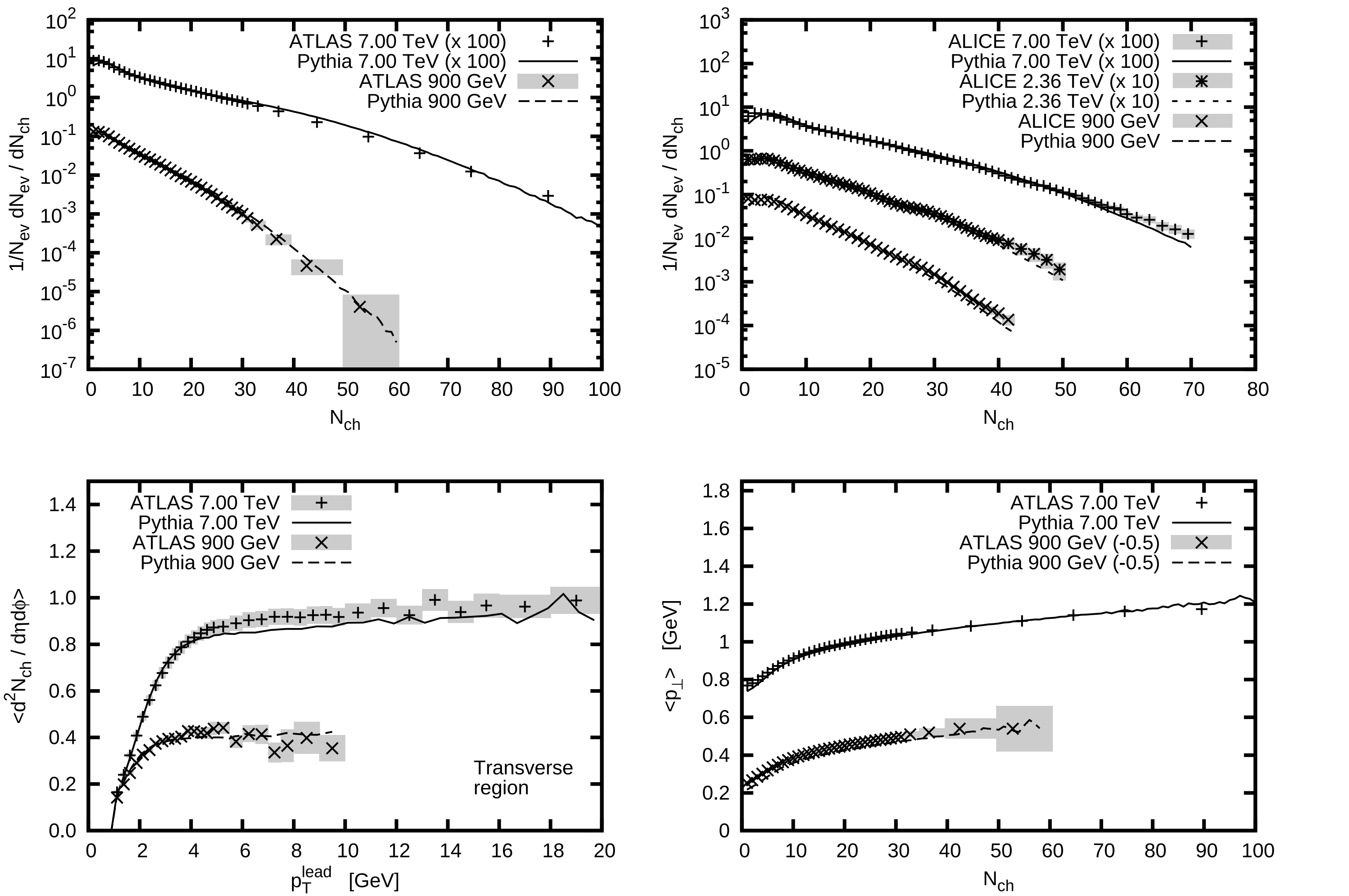
. . . while waiting for LHC data conveniently implemented in Rivet.

Parameter	**Tune 2C**	**Tune 2M**	**Tune 4C**
`SigmaProcess:alphaSvalue`	0.135	0.1265	0.135
`SpaceShower:rapidityOrder`	on	on	on
`SpaceShower:alphaSvalue`	0.137	0.130	0.137
`SpaceShower:pT0Ref`	2.0	2.0	2.0
`MultipleInteractions:alphaSvalue`	0.135	0.127	0.135
`MultipleInteractions:pT0Ref`	2.320	2.455	2.085
`MultipleInteractions:ecmPow`	0.21	0.26	0.19
`MultipleInteractions:bProfile`	3	3	3
`MultipleInteractions:expPow`	1.60	1.15	2.00
`BeamRemnants:reconnectRange`	3.0	3.0	1.5
`SigmaDiffractive:dampen`	off	off	on
`SigmaDiffractive:maxXB`	N/A	N/A	65
`SigmaDiffractive:maxAX`	N/A	N/A	65
`SigmaDiffractive:maxXX`	N/A	N/A	65

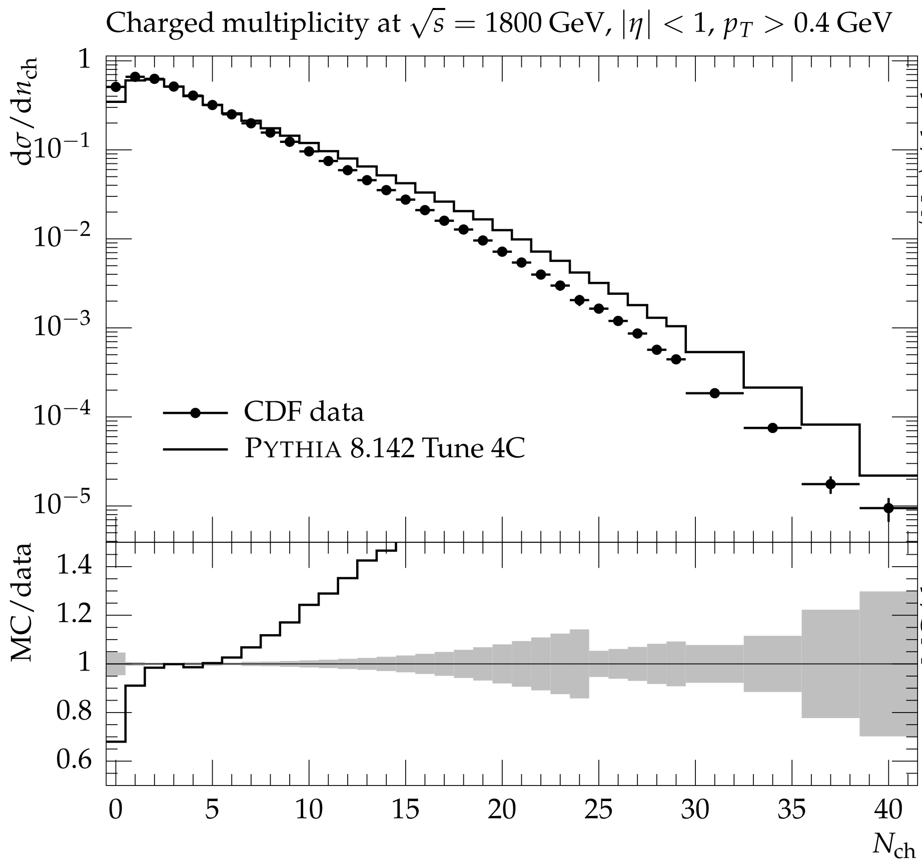
R. Corke & TS, arXiv:1011.1759 [hep-ph]

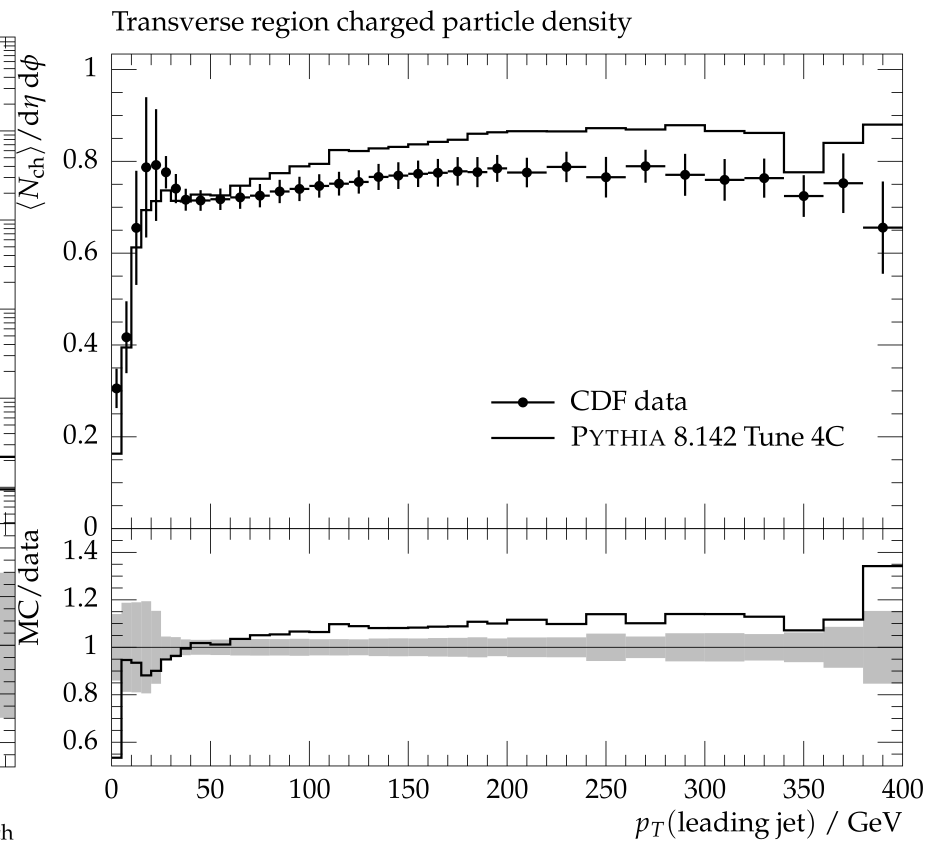
(recently also variant of 4C with x-dependent proton size)

Tune 4C now describes LHC data reasonably well:



. . . but at the expense of Tevatron agreement:

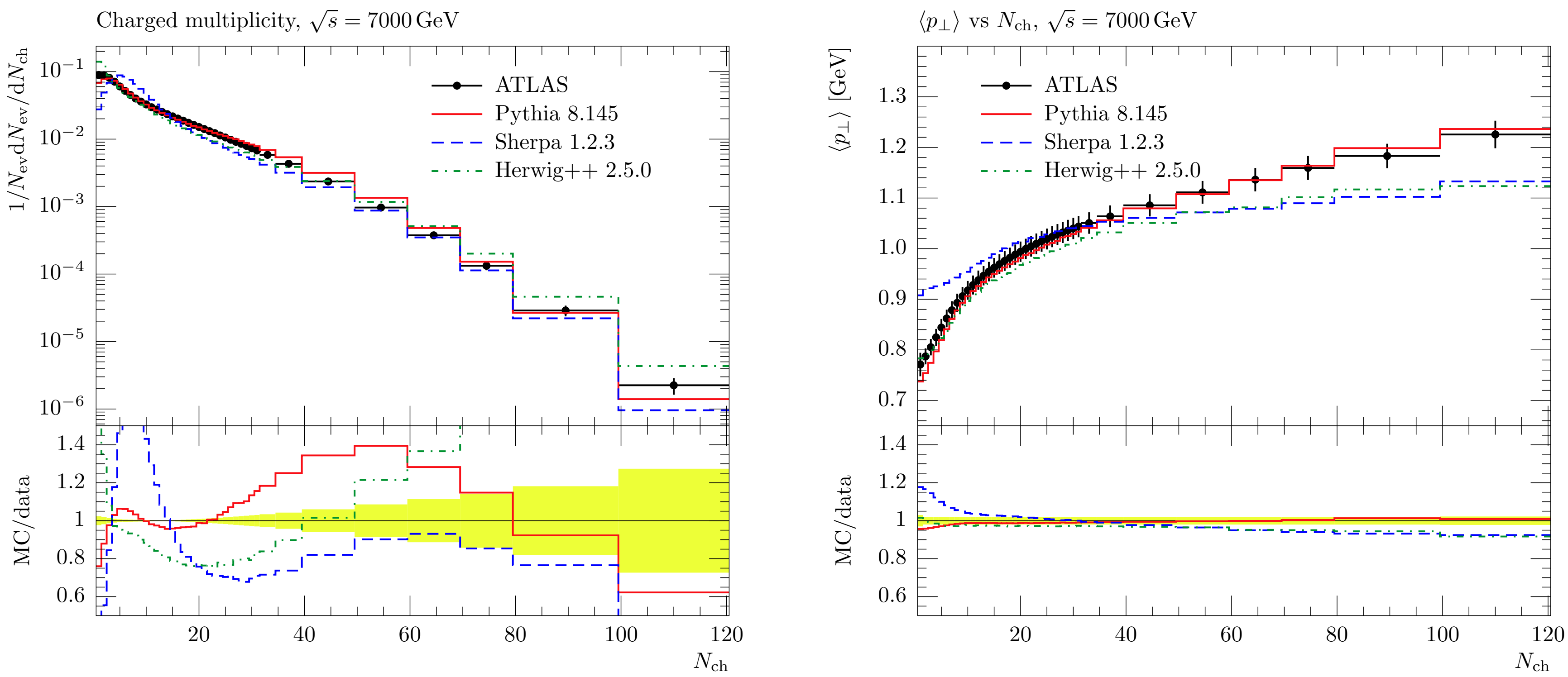




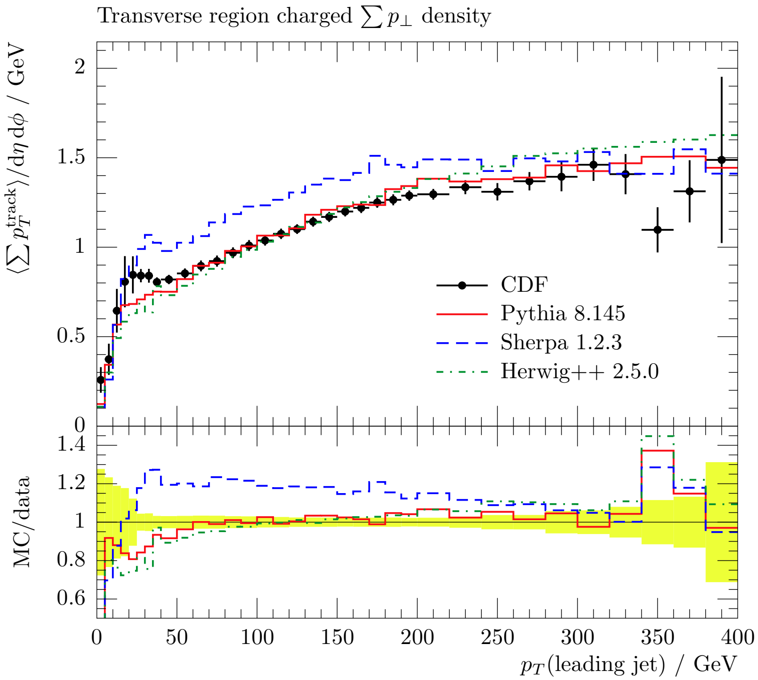
Future:

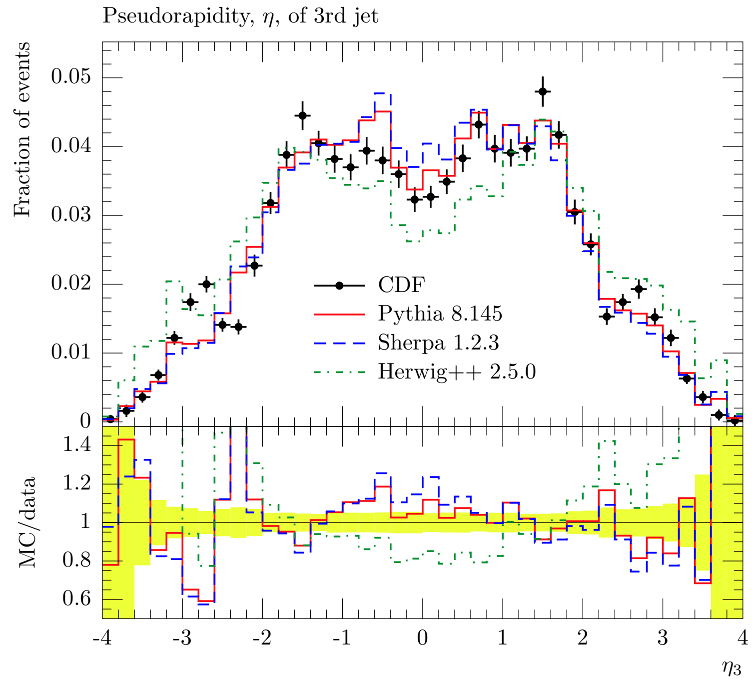
- better understanding of data?
- official/validated inclusion in Rivet?
- combined tune Tevatron + LHC?

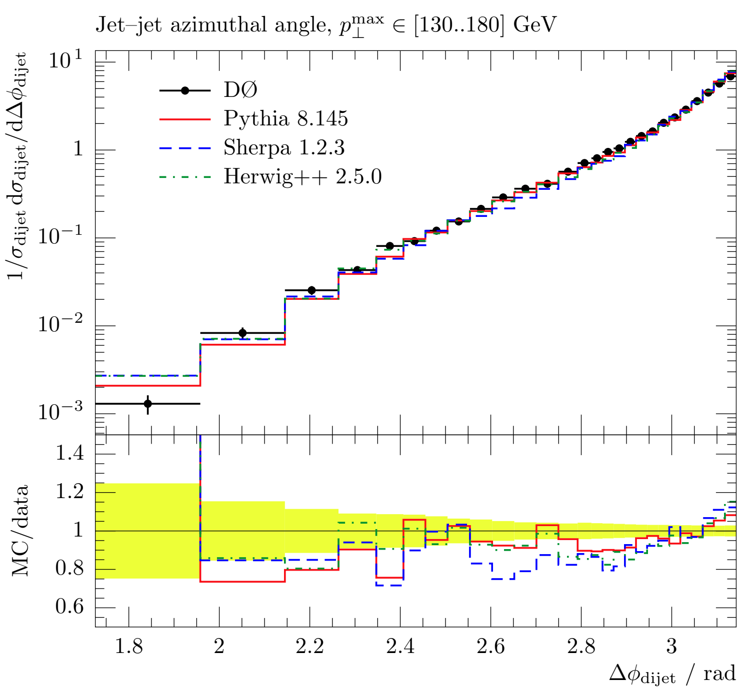
Generally fares well when compared with HERWIG++ and SHERPA:

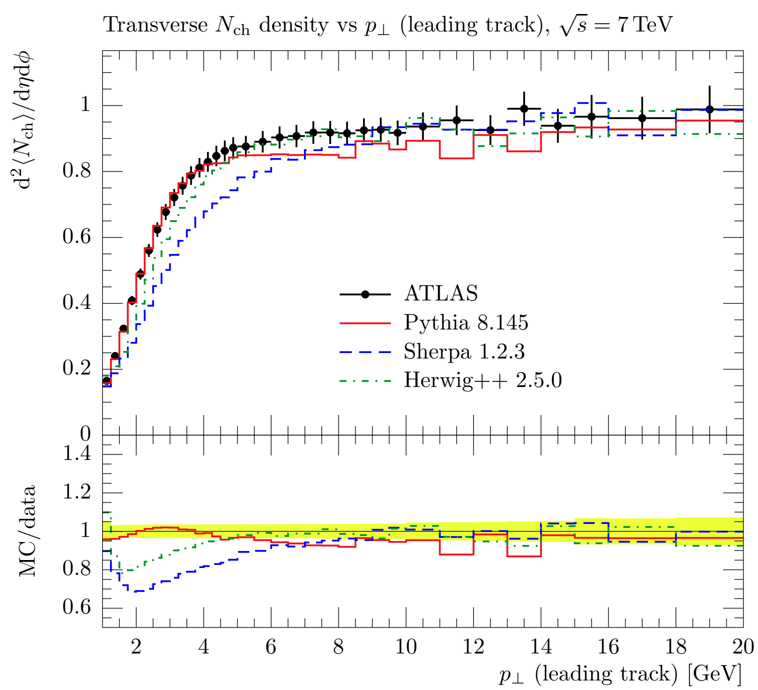


A. Buckley et al., MCnet/11/01, arXiv:1101.2599


Transverse region charged $\sum p_\perp$ density
$\langle \sum p_T^{\mathrm{track}} \rangle / \mathrm{d}\eta\, \mathrm{d}\phi$ / GeV
CDF
Pythia 8.145
Sherpa 1.2.3
Herwig++ 2.5.0
MC/data
p_T(leading jet) / GeV

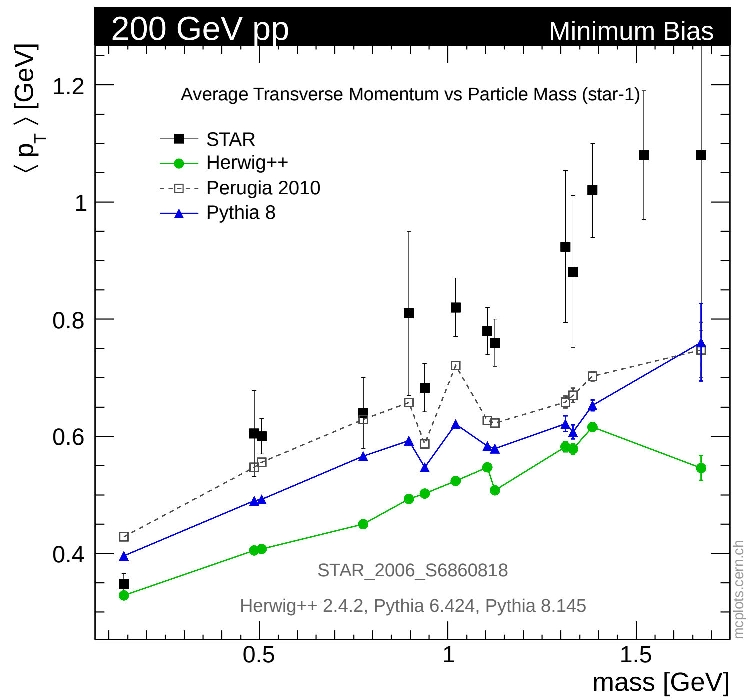

Pseudorapidity, η, of 3rd jet
Fraction of events
CDF
Pythia 8.145
Sherpa 1.2.3
Herwig++ 2.5.0
MC/data
η_3

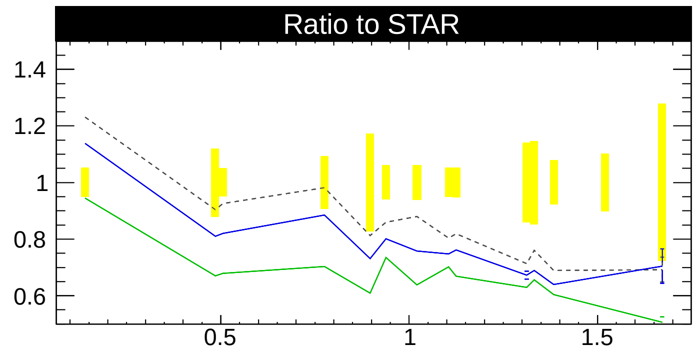

Jet–jet azimuthal angle, $p_\perp^{\max} \in [130..180]$ GeV
$1/\sigma_{\mathrm{dijet}}\, \mathrm{d}\sigma_{\mathrm{dijet}} / \mathrm{d}\Delta\phi_{\mathrm{dijet}}$
DØ
Pythia 8.145
Sherpa 1.2.3
Herwig++ 2.5.0
MC/data
$\Delta\phi_{\mathrm{dijet}}$ / rad

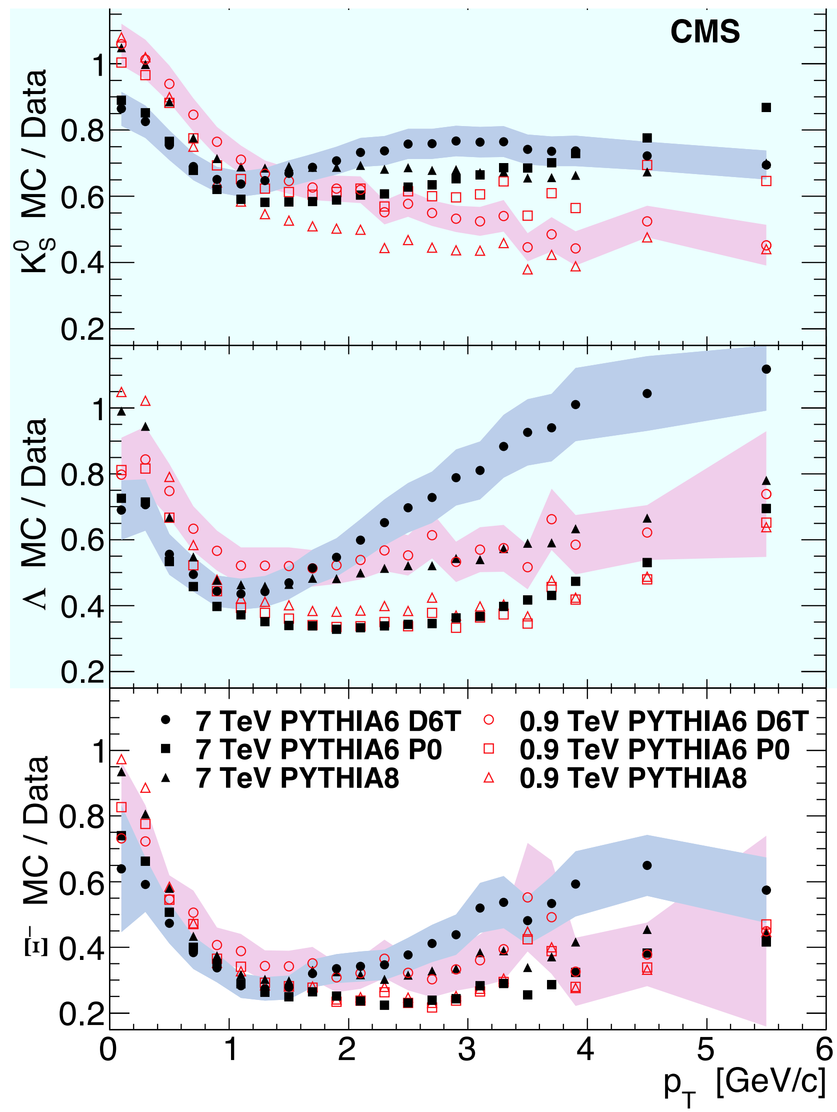

Transverse N_{ch} density vs $p_\perp$ (leading track), $\sqrt{s} = 7$ TeV
$\mathrm{d}^2 \langle N_{\mathrm{ch}} \rangle / \mathrm{d}\eta \mathrm{d}\phi$
ATLAS
Pythia 8.145
Sherpa 1.2.3
Herwig++ 2.5.0
MC/data
$p_\perp$ (leading track) [GeV]

What next?

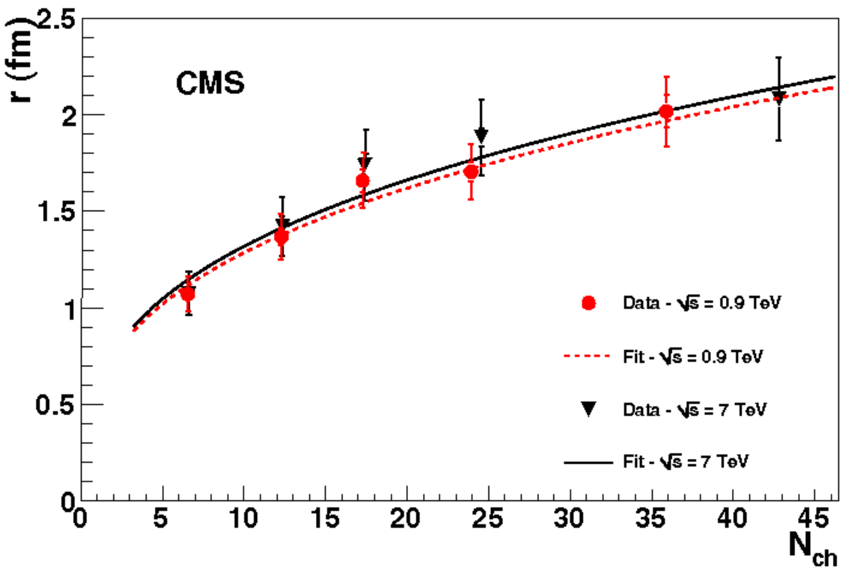
Need more $p_\perp$ for K, p, Λ, . . . , relative to $\pi^\pm$:

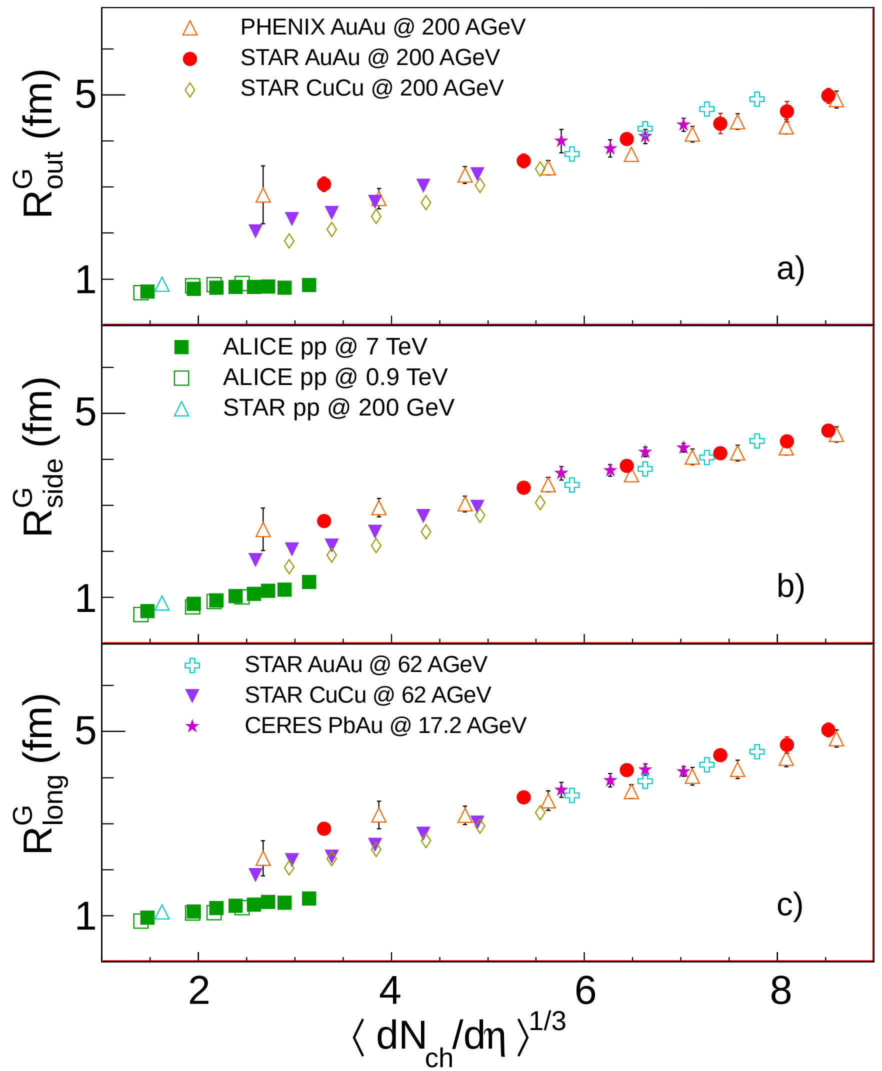






Bose-Einstein $r(N_{\mathrm{ch}}) \propto N_{\mathrm{ch}}^{1/3}$
cannot be accommodated
in PYTHIA effective description
that worked at LEP



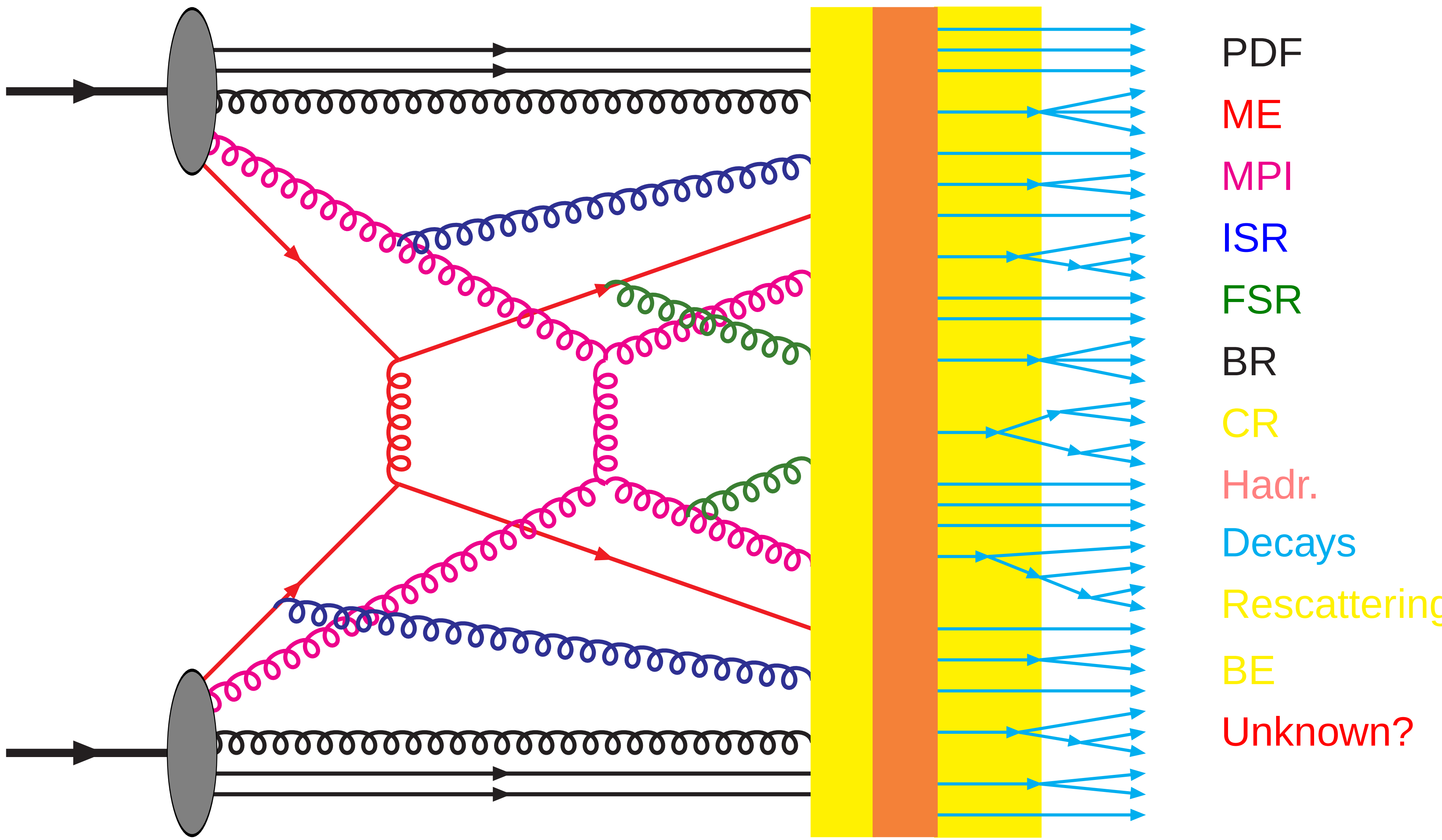


Multiple overlapping fragmenting strings $\Rightarrow$ dense hadron gas!

Target: model for rescattering as afterburner.

Do not need to change particle composition (significantly),
only (partial) collective flow?

So the event picture at the LHC will have to become even more messy:



The road leads to more complexity, not less.
Another example: matching to NLO, NNLO (VINCIA; Lönnblad & Prestel).
The need for experiment ↔ generators ↔ theory remains.

Summary and Outlook

- PYTHIA6 is winding down:
  - ⋆ is supported but not developed;
  - ⋆ still main option for current run (sigh),
  - ⋆ *but not after long shutdown 2013*!
- PYTHIA8 is the natural successor,
  - ⋆ is (sadly!) not yet quite up to speed in *all* respects,
  - ⋆ but for much already better than PYTHIA6,
  - ⋆ is starting to have competitive tunes,
  - ⋆ and will continue to move ahead.
  - ⋆ Version 8.150 coming soon with the features mentioned here.
- Advise to experimentalists:
  - ⋆ step up PYTHIA8 usage to gain experience;
  - ⋆ if you want new features then be prepared to use PYTHIA8;
  - ⋆ provide feedback, both what works and what does not;
  - ⋆ do your own tunes to data and tell outcome.

We are living in an exciting time, and more is to come!