LUND UNIVERSITY

CMS Monte Carlo
tuning miniworkshop
CERN, 14 May 2007

PYTHIA – News and Usage

Torbjörn Sjöstrand

CERN/PH and

Department of Theoretical Physics, Lund University

• Timeline •
• Transverse-Momentum-Ordered Showers •
• Multiple Interactions and Beam Remnants •
• Onium Production •
• On The Usage of NLO PDF’s •
• Other News •
• PYTHIA 8 Commercial •
• Future validation/tuning •

Timeline

PYTHIA	6.1	Mar 1997	—	Jul 2001	merge PYTHIA + JETSET
	6.2	Aug 2001	—	Jan 2005	BNV, LNV, junctions
	6.3	Jul 2003	—	Mar 2006	new showers & multiple interactions
	6.4	Mar 2006	—	??? 2008	slow upgrade path (?)
	8.0	Sep 2004	—	fall 2007	C++ experimental
	8.1	fall 2007	—		C++ standard

Currently PYTHIA 6.411 (soon 6.412):

- 75,000 lines of code (including comments/blanks)
- 580 page PYTHIA 6.4 Physics and Manual
  T. Sjöstrand, S. Mrenna and P. Skands,
  JHEP**05** (2006) 026 [hep-ph/0603175]
- New PYTHIA 6 homepage on CEDAR HepForge:
  `http://projects.hepforge.org/pythia6/`
  news: fsplit $\longrightarrow$ tarballs, makefile, SVN,
  bugtracker, wiki, discussion forum, . . . (still trying our way)
- archive of older versions still available on
  `http://www.thep.lu.se/~torbjorn/Pythia.html`

Transverse-momentum-ordered showers

Message from LEP:

ARIADNE dipole ($p_\perp$-ordered) + PYTHIA string fragmentation

$>$ PYTHIA showers (mass-ordered) + string fragmentation

$>$ HERWIG showers (angular-ordered) + cluster fragmentation

Objective:

Incorporate several of the good points of the dipole formalism (like ARIADNE) within the shower approach ($\Rightarrow$ hybrid)

by ordering in

$\boldsymbol{p^2_{\perp\mathrm{evol}} = z(1-z)Q^2 = z(1-z)M^2}$ for FSR

$\boldsymbol{p^2_{\perp\mathrm{evol}} = (1-z)Q^2 = (1-z)(-M^2)}$ for ISR

$\pm$ explore alternative $p_\perp$ definitions

$+$ FSR and ISR cast in similar, but not identical, formalisms

$+$ $p_\perp$ ordering $\Rightarrow$ coherence inherent

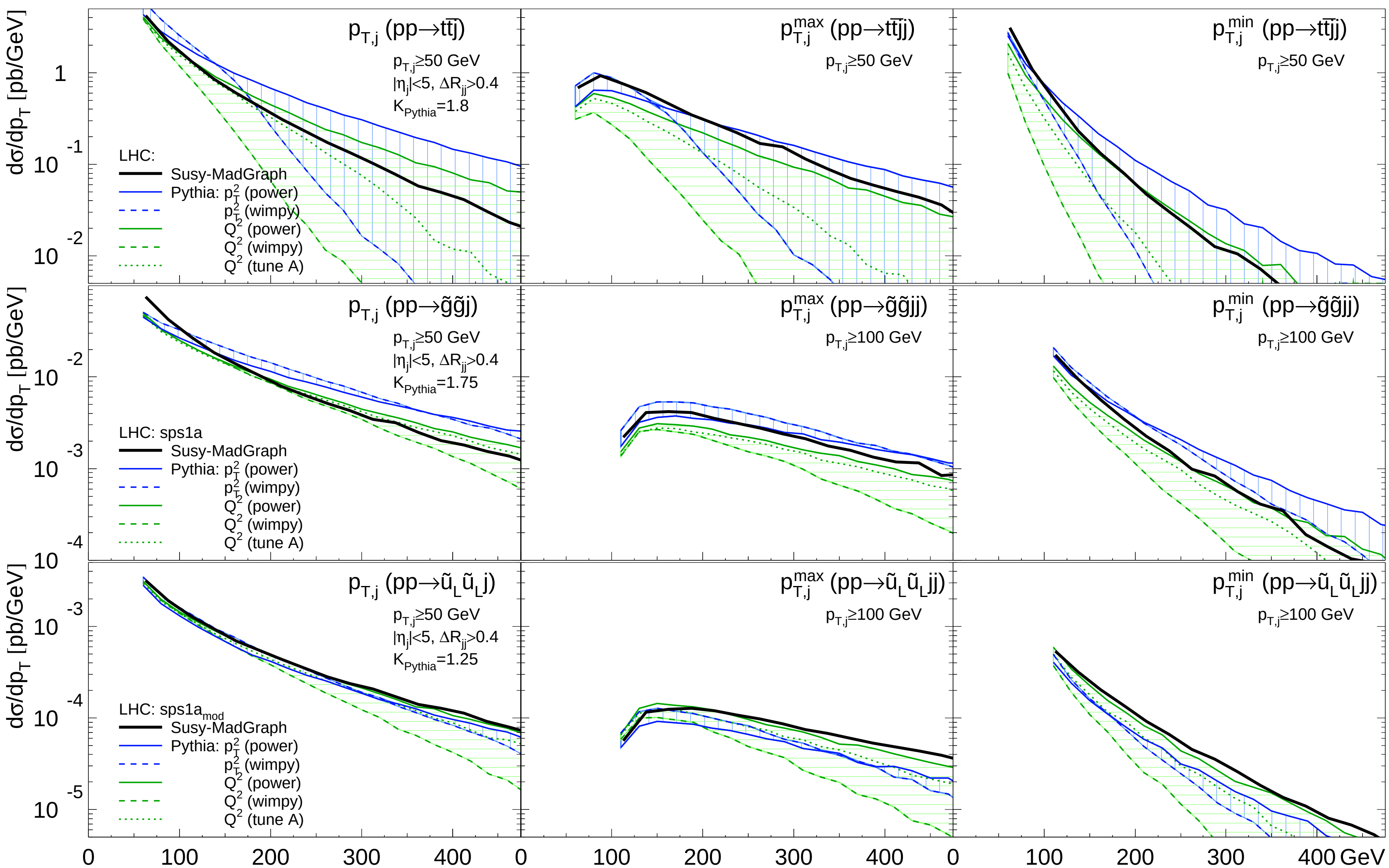
$+$ ME merging works as before (unique $p_\perp^2 \leftrightarrow Q^2$ mapping; same z)

$+$ g $\rightarrow$ q$\overline{\text{q}}$ natural

$+$ well suited for ME/PS matching (still partly to be worked out)

$++$ well suited for *interleaved multiple interactions*

$p_\perp$(1 jet) $\quad$ $p_\perp^{\text{max}}$(2 jets) $\quad$ $p_\perp^{\text{min}}$(2 jets)

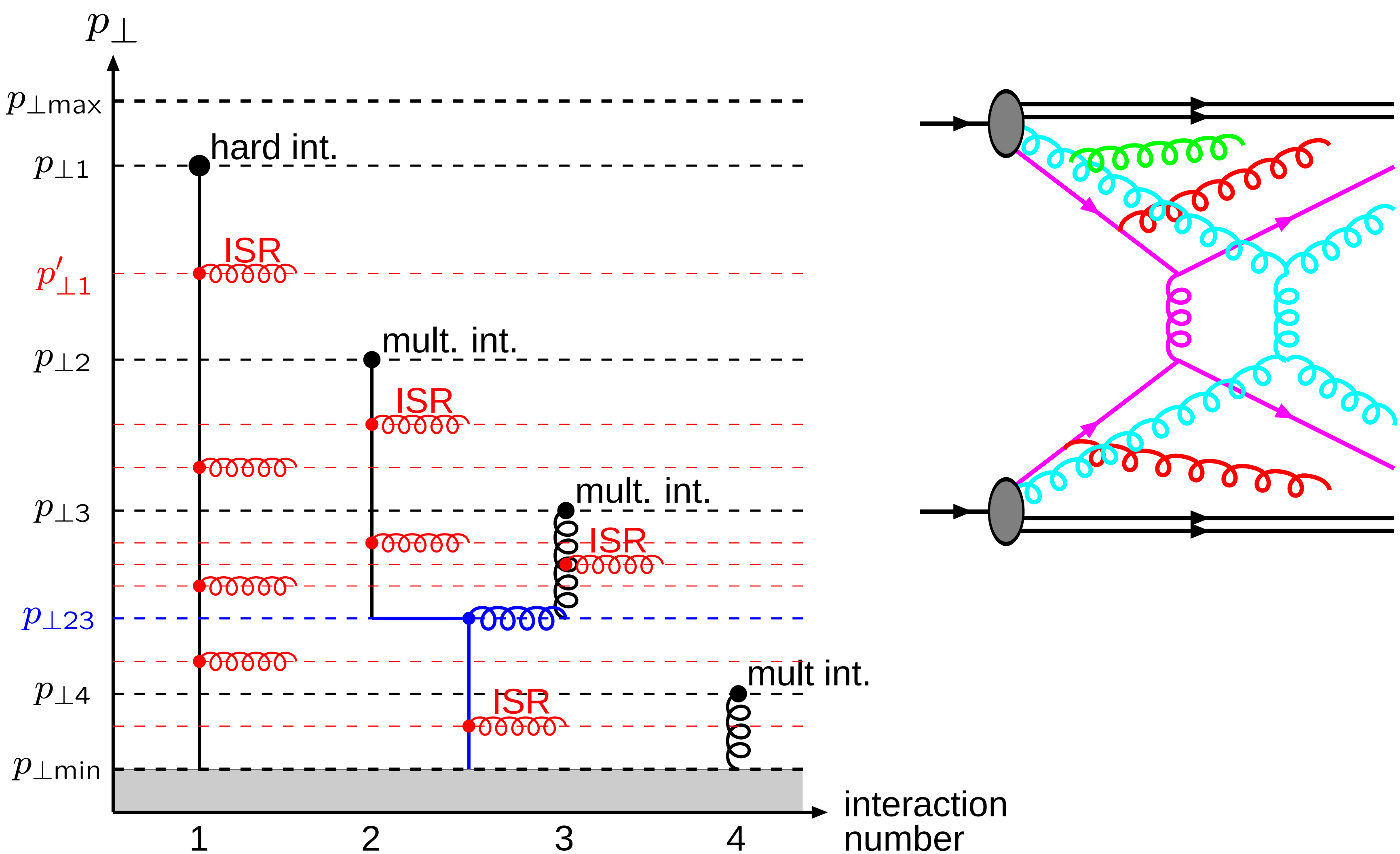


power: $Q^2_{\text{max}} = s$; $\quad$ wimpy: $Q^2_{\text{max}} = m^2_\perp$; $\quad$ tune A: $Q^2_{\text{max}} = 4m^2_\perp$

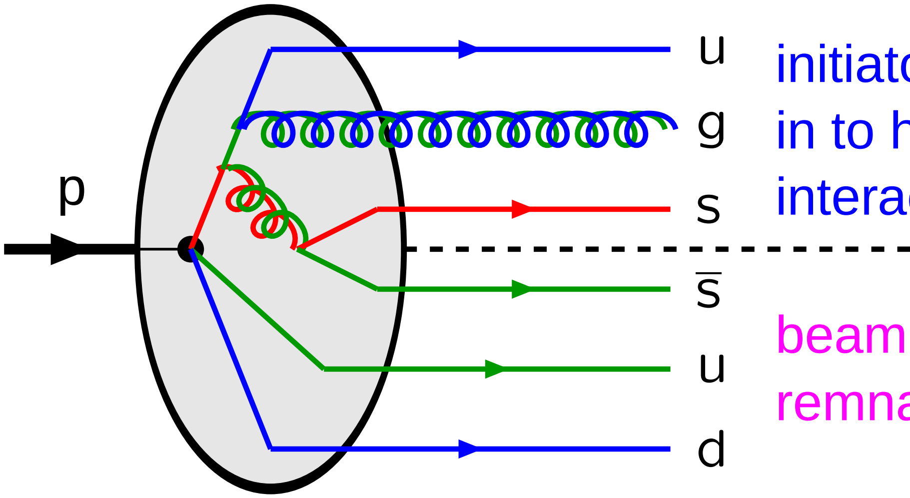
$m_{\text{t}} = 175$ GeV, $\quad m_{\tilde{\text{g}}} = 608$ GeV, $\quad m_{\tilde{\text{u}}_L} = 567$ GeV

(T. Plehn, D. Rainwater, P. Skands)

Interleaved Multiple Interactions



Initiators and Remnants



Need to assign:

- correlated flavours
- correlated $x_i = p_{zi}/p_{z\mathrm{tot}}$
- correlated primordial $k_{\perp i}$
- correlated colours
- correlated showers

- **PDF after preceding MI/ISR activity:**

0) Squeeze range $0 < x < 1$ into $0 < x < 1 - \sum x_i$ (ISR: $i \neq i_{\mathrm{current}}$)

1) Valence quarks: scale down by number already kicked out

2) Introduce companion quark $\mathrm{q}/\overline{\mathrm{q}}$ to each kicked-out sea quark $\overline{\mathrm{q}}/\mathrm{q}$, with x based on assumed $\mathrm{g} \to \mathrm{q}\overline{\mathrm{q}}$ splitting

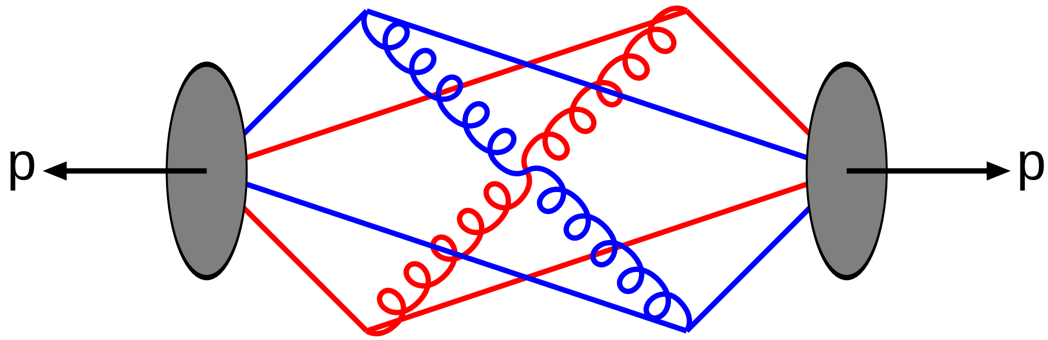
3) Gluon and other sea: rescale for total momentum conservation

PYTUNE

Minimum bias + Underlying event tunes conveniently setup with `CALL PYTUNE(MODEL)`, where possible `MODEL` values are

Old UE, Q^2-ordered showers		
100	A	Rick Field's Tune A
101	AW	Rick Field's Tune AW
102	BW	Rick Field's Tune BW
103	DW	Rick Field's Tune DW
104	DWT	Rick Field's Tune DW with slower UE energy scaling
105	QW	Rick Field's Tune QW (NB: needs CTEQ6.1 PDF's)
106	ATLAS	Arthur Moraes' ATLAS tune
107	ACR	Tune A modified with annealing CR
New UE, Q^2-ordered showers		
200	IM 1	Intermediate 1: annealing CR
New UE, interleaved pT-ordered showers, annealing CR		
300	S0	Sandhoff-Skands Tune 0
301	S1	Sandhoff-Skands Tune 1
302	S2	Sandhoff-Skands Tune 2
303	S0A	S0 with "Tune A" UE energy scaling
304	NOCR	New UE "best try" without colour reconnections
305	Old	New UE, original (primitive) colour reconnections

Colour correlations



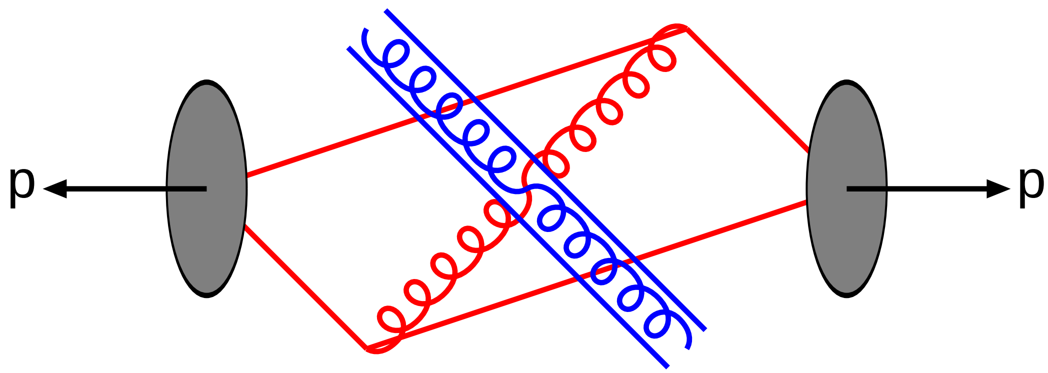
long strings to remnants

$\Rightarrow$ much n_{ch}/interaction

$\Rightarrow$ few interactions

$\Rightarrow$ little $p_{\perp\text{pert}}$

$\Rightarrow$ $\langle p_\perp \rangle(n_{\text{ch}}) \sim$ flat



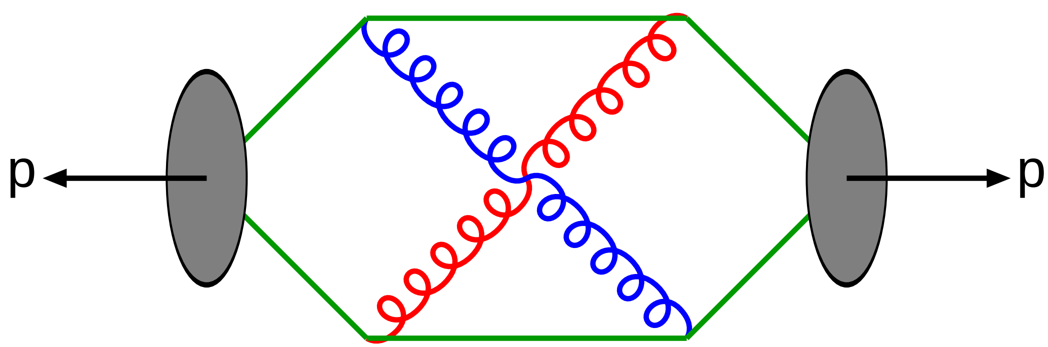
PYTHIA4 simplification:

after hardest interaction

stretch further strings

between scattered gluons

$\Rightarrow$ $\langle p_\perp \rangle(n_{\text{ch}})$ flattens out



short strings (more central)

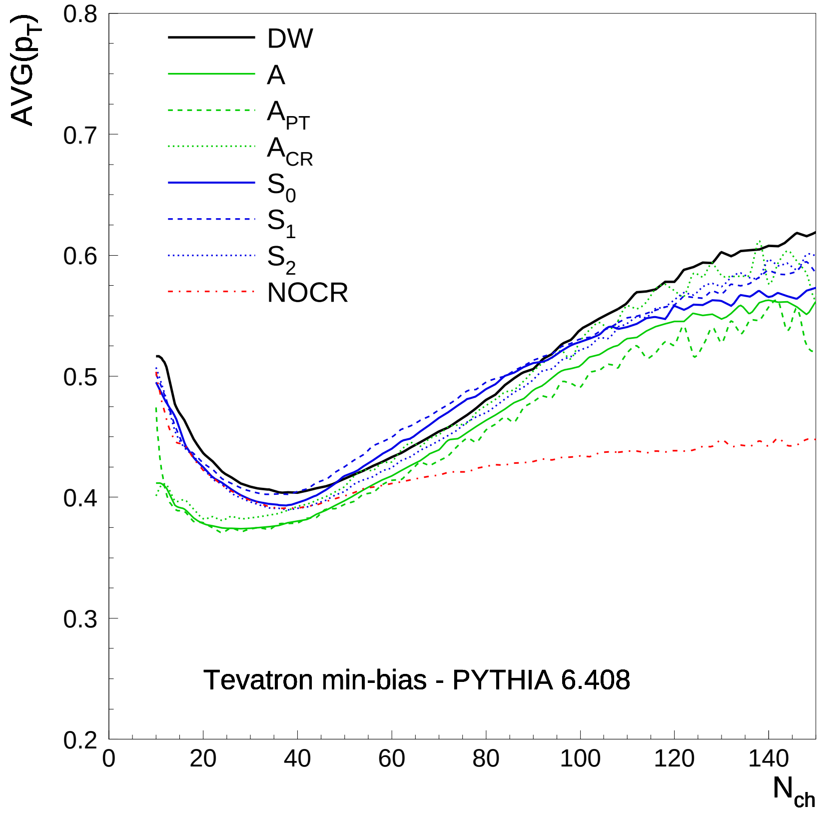
$\Rightarrow$ less n_{ch}/interaction

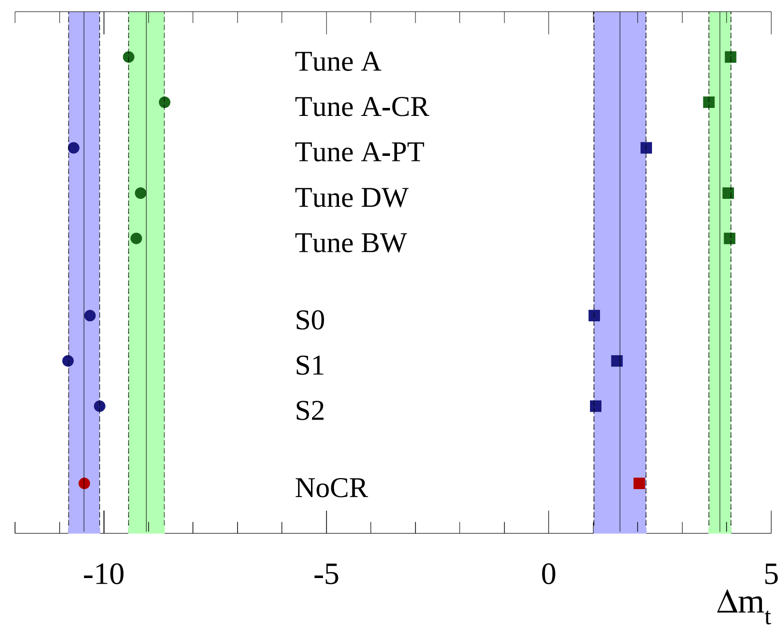
$\Rightarrow$ more interactions

$\Rightarrow$ more $p_{\perp\text{pert}}$

$\Rightarrow$ $\langle p_\perp \rangle(n_{\text{ch}})$ rising

Well-known issue over the years — unsolved *and* interesting!
2007 (Skands & Wicke): also top mass determinations affected

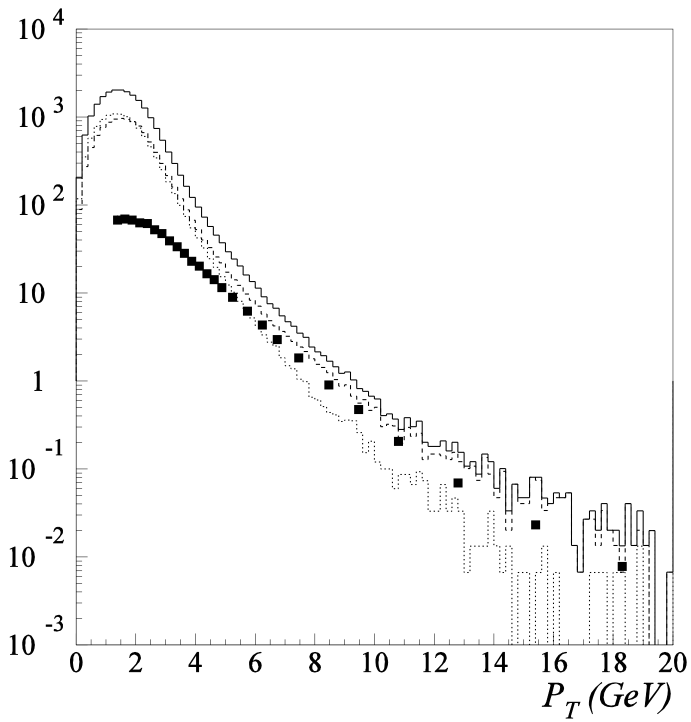


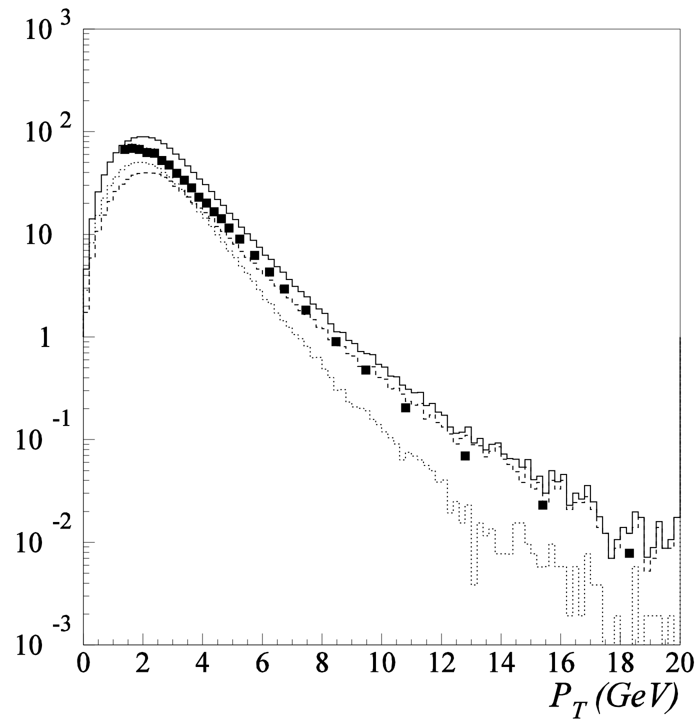


Also several other interesting issues, like energy dependence
$p_{\perp 0}(s) = p_{\perp 0}(s_0)(s/s_0)^{\epsilon}$,
$\epsilon \approx 0.08$ (DL) or bigger?

Onium Production

- Colour-singlet production of $\mathrm{J}/\psi/\Upsilon$ since long
- Stefan Wolf (Valencia postdoc) wrote colour-octet implementation;
  at the time (2002) no interest in having it in standard PYTHIA
- 2005 LHC-B interest (V. Vagnoni), integrated (421–439, 461–479)
- 2006 validated by M. Bargiotti; compared with CDF data:





sharp cut $p_{\perp \mathrm{min}} = 1$ GeV

dampening $p_{\perp 0} = 2.85$ GeV

+ CTEQ 6L PDF + maximal showers + . . .

Why dampening?

QCD processes by t-channel gluon exchange diverge roughly like

$$\frac{\mathrm{d}\hat{\sigma}}{\mathrm{d}p_\perp^2} \propto \frac{\alpha_\mathrm{s}^2(p_\perp^2)}{p_\perp^4}$$

Assumed regularized by colour screening distance inside hadron, $p_{\perp 0} \approx \hbar / r_{\mathrm{screen}}$, implemented by multiplicative factor

$$F(p_{\perp 0}) = \frac{p_\perp^4}{(p_{\perp 0}^2 + p_\perp^2)^2} \frac{\alpha_\mathrm{s}^2(p_{\perp 0}^2 + p_\perp^2)}{\alpha_\mathrm{s}^2(p_\perp^2)}$$

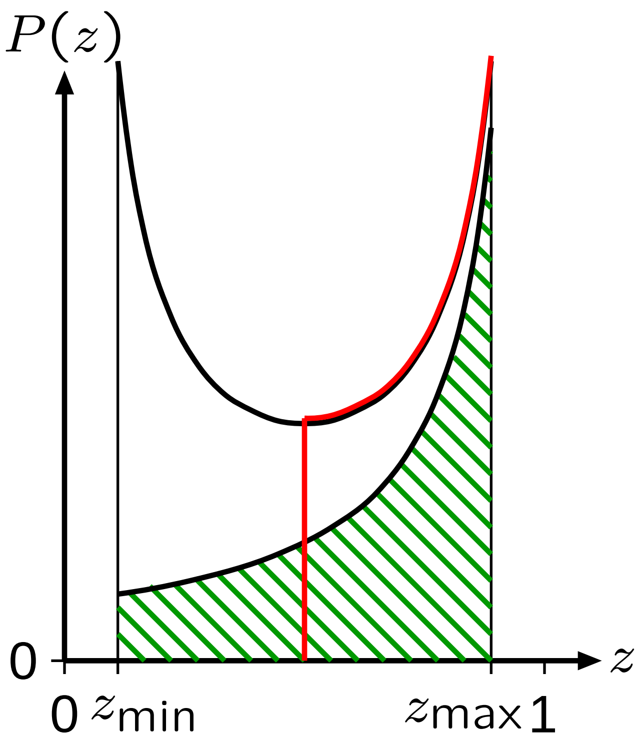
(TS & M. van Zijl, Phys. Rev. D36 (1987) 2019)

Empirically $p_{\perp 0} \approx 2$ GeV at Tevatron energies.

Onium production also involves t-channel gluon exchange,
so *assume universality* and apply $F(p_{\perp 0})$ to onia cross section.
Success if same $p_{\perp 0}$ works everywhere.
Half success that *almost* same $p_{\perp 0}$ works everywhere, even for Υ.
Hope: could converge in future.

What showering?

Recall $P_{g\to gg}(z) \propto \frac{1}{z(1-z)} = \frac{1}{z} + \frac{1}{1-z} \sim \frac{2}{1-z} \propto 2P_{q\to qg}(z)$



In code: decided from onset to produce onium, for efficiency reasons, but . . .

. . . physical picture (my understanding):

a) gluon produced in hard interaction

b) showering $g \to gg$ symmetric around $z = 1/2$
 but $z > 1/2$ gluon kinematically favoured

c) branching $g \to c\bar{c}$ (or $g \to b\bar{b}$)

d) further cascading $c \to cg$ and $\bar{c} \to \bar{c}g$

e) formation of semibound $c\bar{c}$ octet state

f) soft gluon radiation to singlet state

Newly found mismatch description $\Leftrightarrow$ code ("**bug**"):

3) `MSTP(148) = 1, MSTP(149) = 1` : radiation like $g \to gg$ $\Leftrightarrow$ OK

2) `MSTP(148) = 1, MSTP(149) = 0` : radiation like $g \to gg$, but $z > 1/2$
 $\Leftrightarrow$ like $P_{g\to gg}(z) \propto 2/(1-z)$ (≈ the description in PYTHIA 8)

1) `MSTP(148) = 0` : no radiation $\Leftrightarrow$ like $P_{q\to qg}(z) \propto 1/(1-z)$

On The Usage of NLO PDF's

General-purpose event generators provide
$\hat{\sigma}(\mathrm{LO}) \otimes \mathrm{PDF}(\mathrm{LO}) \otimes$ showers.
Each component separately is positive definite.

Showers are often matched to $\hat{\sigma}(\mathrm{NLO})$ in hard-emission region,
but handles soft/collinear regions by Sudakovs $\Longrightarrow$ positivity preserved.
Exception: external NLO input, e.g. MC@NLO or other Les Houches.

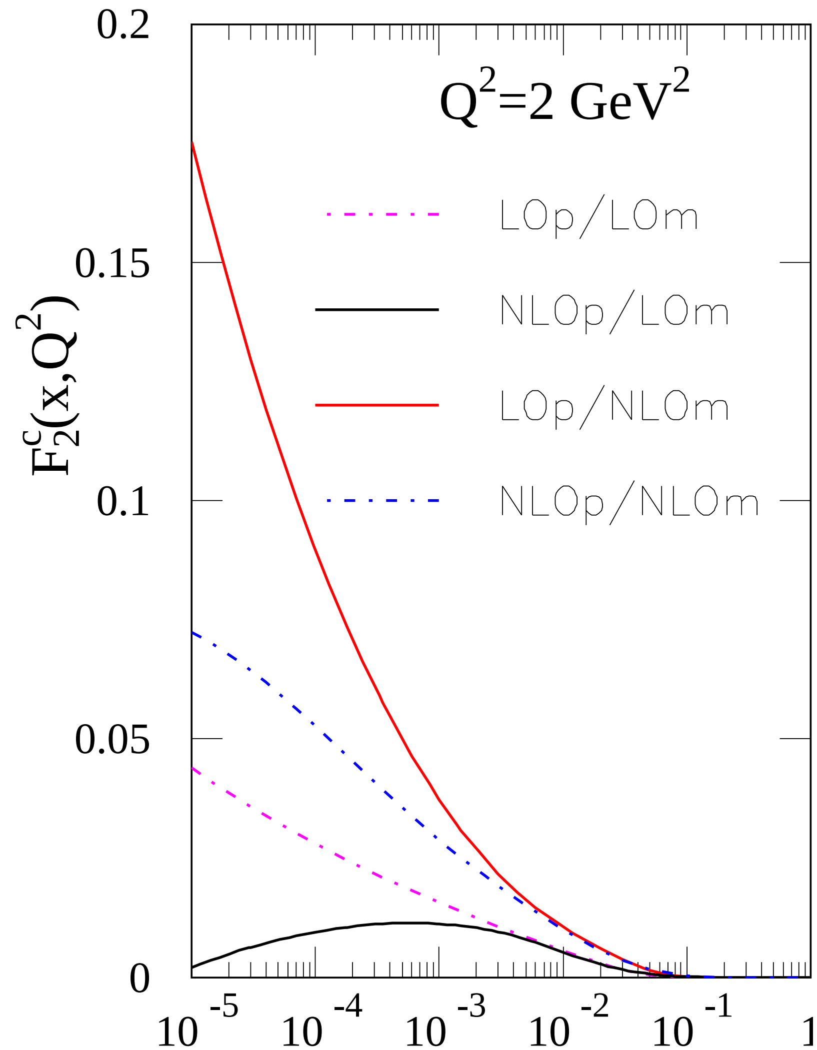
In NLO only $\hat{\sigma}(\mathrm{NLO}) \otimes \mathrm{PDF}(\mathrm{NLO})$ required to be positive,
PDF(NLO) on its own not physical (e.g. $\overline{\mathrm{MS}}$ vs. DIS).
Therefore $\hat{\sigma}(\mathrm{LO}) \otimes \mathrm{PDF}(\mathrm{NLO})$ *may* give crap.

Especially problematical at small x and Q^2.
E.g. $p_\perp = 1.4$ GeV scattering at LHC:
$Q^2 \approx p_\perp^2 \approx 2$ GeV^2, $x \geq 4p_\perp^2/s = 4 \cdot 10^{-8}$.
Not uncommon values for multiple-interactions description!

This is the region we need to test before we let the genie out of the bottle.
($\hat{\sigma}(\mathrm{LO}) \otimes \mathrm{PDF}(\mathrm{LO})$ at least is "sensible" here, even if not "correct".)

From Robert Thorne's presentation at ATLAS meeting, 14 December 2006:

- Large logarithms in $\ln(1-x)$ and $\ln(1/x)$ tend to lead to enhanded NLO matrix elements in these regions.
- Due to this, partons extracted in these regions at NLO are smaller.
- By momentum sum rule NLO PDF's can be larger at intermediate $x \sim 0.1 - 0.01$, i.e. often of most interest at hadron colliders.
- Conclusion: really is best to think a bit each time rather than to have fixed prescription of either LO or NLO partons with Monte Carlos. Each choice will be wrong in some cases.



Bad χ^2 for LO fits from sum rules forcing undershoot at intermediate x?

$\Rightarrow$ Homework: produce LO PDF set without sum rule constraint.

The proof of the pudding lies in the eating

Assume the best description of physics is obtained with
(a) $\hat{\sigma}(\mathrm{NLO}) \otimes \mathrm{PDF}(\mathrm{NLO})$.

Interesting comparisons would then be with the scenarios:
(b) $\hat{\sigma}(\mathrm{LO}) \otimes \mathrm{PDF}(\mathrm{LO})$.
(c) $\hat{\sigma}(\mathrm{LO}) \otimes \mathrm{PDF}(\mathrm{LO}) \otimes$ showers.
(d) $\hat{\sigma}(\mathrm{LO}) \otimes \mathrm{PDF}(\mathrm{NLO})$.
(e) $\hat{\sigma}(\mathrm{LO}) \otimes \mathrm{PDF}(\mathrm{NLO}) \otimes$ showers.
Only if (e) is a better approximation to (a) than is (c) would the use of NLO PDF's be motivated in a general-purpose generator.

Comparisons based on jet clustering have problem that
NLO = 2 + 3 jets, LO = 2 jets, LO + shower $\geq$ 2 jets,
so complement with inclusive clustering-independent quantities such as

$$E_{\perp} = \sum_i |p_{\perp i}| \qquad \text{or} \qquad Y = \frac{1}{E_{\perp}} \sum_i |p_{\perp i}| \, y_i$$

where $p_{\perp}$-weighting ensures infrared and collinear safety.
Especially interesting in low-$E_{\perp}$ region!

"Homework" for Joey Huston for Les Houches?

Other News

Bug from PYTHIA 6.321 - PYTHIA 6.406:
procedure for setting maximum scale of shower rewritten from scratch,
but did not consider gauge-boson-fusion or prompt-photon processes;
so there scale became s, which gave too much extra jet activity

- Les Houches Event Files provides standard input fomat,
  e.g. from MadGraph
- PDF reweighting at hard interaction simplified:
  $xf_i(x, Q^2)$ now stored in `PARI(29), PARI(30)`.
  Thus optionally allow LHEF extra output line
  `#pdf id1 id2 x1 x2 scalePDF xpdf1 xpdf2`
  which here translates into
  `#pdf MSTI(15) MSTI(16) PARI(33) PARI(34) PARI(23) PARI(29) PARI(30)`
- `UPVETO` routine allows rejection of parton-level events
  for MLM matching to Alpgen (MadGraph on the way)
- Many improvements in BSM physics, especially SLHA
- Many further small additions and bug fixes

Switch on/off decay channels with PYGIVE

Problem: addressing by decay-channel-number unstable and error-prone.

Solution (from PYTHIA 8): allow decisions based on match to final state

```
CALL PYGIVE("KF:PROPERTY = LIST")
```

where

`KF` is the PDG code of the decaying particle

`PROPERTY` tells which action to take

- `= ALLON, ALLOFF` all decay channels are switched on or off; expects no `= LIST`.
- `= ONIFANY, OFFIFANY` a decay channel is switched on or off if one of its decay products matches one of the particles in the list.
- `= ONIFALL, OFFIFALL` a decay channel is switched on or off if each of the particles in the list is matched to one particle in the channel. The channel may contain extra unmatched particles.
- `= ONIFMATCH, OFFIFMATCH` works as `ONIFALL, OFFIFALL`, but there must be a perfect match, i.e. no excess particles.

`LIST` is a list of PDG particle codes (at most 10, blank- or comma-separated). Only absolute value relevant for match.

Examples:

```
CALL PYGIVE("23:alloff")
CALL PYGIVE("23:onifany = 1 2 3 4 5")
```

first switches off all Z^0 decay and then switches back on those to quarks.

```
CALL PYGIVE("15:offifany = 311 321 310 130 313 323")
CALL PYGIVE("15:onifmatch = 16 323")
CALL PYGIVE("15:offifall = 211 211")
```

first switches off all decay modes of the $\tau^{\pm}$ that contains a K or K^* meson,
then switches back on the specific $\nu_\tau K^{*-}$ channel,
and finally switches off all channels with at least two charged pions.

Some notes:

1) Only works for simple on/off, not e.g. to set B^0 and $\overline{B}^0$ differently.
2) Input case-insensitive. Maxlength 100 characters.
3) Provides echo-back and warns if failure to parse.
4) Commands carried out in sequence $\Rightarrow$ order relevant.
Check result with `CALL PYLIST(12)` or `CALL PYSTAT(2)`.

News after 6.409

In 4.610:

- minor bug correction for colour tags (especially for ACT = `PYTUNE(107)`)
- much SUSY, especially for SLHA

In 6.411:

- further much SUSY
- `PYGIVE` additions as above

For 6.412 (?):

- `STOP` $\longrightarrow$ `CALL PYSTOP(ICODE)`
- `PYVETO` for new showers
- ME corrections for $\mathrm{b\overline{b}} \to \mathrm{H}$ to $\mathrm{bg} \to \mathrm{bH}$
- New option for `PYSPLI`: with `MSTP(98) = 1` a proton is assumed to radiate a photon without breaking up, $\mathrm{p} \to \mathrm{p}\gamma$ (CMS diffractive group)
- Option for FSR in old MI framework (??) (S. Mrenna: essential for improved tune to old MI!) (P. Skands, TS: what we strive to do well with new MI!!)

PYTHIA 8 Commercial

PYTHIA 8 not yet ready for production run, but

PYTHIA 8.080 first realistic development release, with

- limited set of internal processes, useful LHA/LHEF interfaces
- initial- and final-state $p_{\perp}$-ordered showers (with $\gamma \to \mathrm{f\overline{f}}$)
- fully interleaved with multiple interactions
  (also with prompt photons and $\mathrm{J}/\psi/\Upsilon$ in underlying event)
- string fragmentation and particle decays (with updated tables)
- HEPMC (M. Kirsanov) and LHAPDF interfaces
- brief introduction, PHP-based interactive manual (B. Lloyd)
- sample main programs, student worksheet
- download `pythia8080.tgz` from
  `http://www.thep.lu.se/~torbjorn/Pythia.html`, link "`Future`"
- very simple to install, at least standalone (M. Kirsanov)
- in GENSER and interfaced to CMS framework (M. Kirsanov)
- ROOT interface under way (B. Bellenot)

Future validation/tuning

Need more experimentalists who understand, analyze and communicate:
fixed-target $\rightarrow$ ISR $\rightarrow$ **RHIC(pp)** $\rightarrow$ Sp$\overline{\text{p}}$S $\rightarrow$ **Tevatron** $\rightarrow$ LHC
with special responsibility for running experiments

Need reproducible data, not necessarily corrected but with fully specified comparison procedures (HZtool $\rightarrow$ Rivet)

Examples of useful distributions:

- charged multiplicity distributions
- flavour composition
- single-particle and jet $p_{\perp}$ spectra
- $\langle p_{\perp} \rangle (n_{\text{charged}})$
- $p_{\perp}$ spectra of Drell-Yan pairs as a function of mass and energy
- jet profiles (mass, angular shape, substructure, . . .)
- underlying-event activity
- number of minijets as a function of cone size and jet energy
- forward–backward and other correlations

Future work in collaboration between:

- Monte Carlo authors (few, at least for PYTHIA)
- other theorists (e.g. PDF people: tunes for small x & Q^2)
- GENSER people (GENSER meetings clearinghouse?)
- MCnet postdocs:
  Lars Sonnenschein (Jan 2007, from D0, at CERN)
  Hendrik Hoeth (Oct 2007 (?), from DELPHI & D0, in Lund)
- experimental collaborations: bulk of manpower!

Need to start over from LEP and work upwards.

Recommendation 1: need procedure inside CMS/ATLAS/. . . to show (non-Nobel-Prize) data to MC people in a timely fashion. Do not repeat LEP secrecy of trivial data and rumour-mongering of non-trivial one.

Recommendation 2: do not spend too much effort on tunes with NLO PDF's, especially not before relevant homework has been done.

Recommendation 3: start to move resources away from further old-MI/ISR/FSR tunes/tests to new-framework ones, either in 6.4 or, even better for the longer run, in 8.0.