- I was asked to discuss b/c/τ simulation issues for the LC
- What I know something about is b/c/τ issues for hadron colliders
- I will focus on this, since I believe it is still relevant

b

MSTJ(42) :

(D=2) branching mode, especially coherence level, for time-like showers in PYSHOW.

= 1 :

conventional branching, i.e. without angular ordering.

= 2 :

coherent branching, i.e. with angular ordering.

= 3 :

in a branching $a \to bg$, where m_b is nonvanishing, the decay angle is reduced by a factor $(1 + (m_b^2/m_a^2)(1-z)/z)^{-1}$, thereby taking into account mass effects in the decay [Nor01]. Therefore more branchings are acceptable from an angular ordering point of view. In the definition of the angle in a $g \to \mathrm{q\overline{q}}$ branchings, the naive massless expression is reduced by a factor $\sqrt{1-4m_q^2/m_g^2}$, which can be motivated by a corresponding actual reduction in the $p_\perp$ by mass effects. The requirement of angular ordering then kills fewer potential $\mathrm{g} \to \mathrm{q\overline{q}}$ branchings, i.e. the rate of such comes up. The $\mathrm{g} \to \mathrm{gg}$ branchings are not changed from =2. This option is fully within the range of uncertainty that exists.

= 4 :

as =3 for $a \to bg$ and $\mathrm{g} \to \mathrm{gg}$ branchings, but no angular ordering requirement conditions at all are imposed on $\mathrm{g} \to \mathrm{q\overline{q}}$ branchings. This is an unrealistic extreme, and results obtained with it should not be overstressed. However, for some studies it is of interest. For instance, it not only gives a much higher rate of charm and bottom production in showers, but also affects the kinematical distributions of such pairs.

= 5 :

new `intermediate' coherence level [Nor01], where the consecutive gluon emissions off the original pair of branching partons is not constrained by angular ordering at all. The subsequent showering of such a gluon is angular ordered, however, starting from its production angle. At LEP energies, this gives almost no change in the total parton multiplicity, but this multiplicity now increases somewhat faster with energy than before, in better agreement with analytical formulae. (The PYSHOW algorithm overconstrains the shower by ordering emissions in mass and then vetoing increasing angles. This is a first simple attempt to redress the issue.) Other branchings as in =2.

= 6 :

`intermediate' coherence level as =5 for primary partons, unchanged for $\mathrm{g} \to \mathrm{gg}$ and reduced angle for $\mathrm{g} \to \mathrm{q\overline{q}}$ and secondary $\mathrm{q} \to \mathrm{qg}$ as in =3.

= 7 :

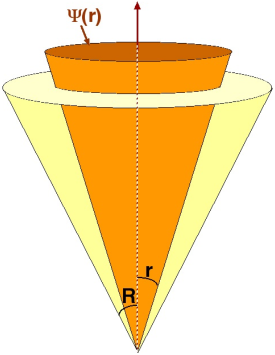
`intermediate' coherence level as =5 for primary partons, unchanged for $\mathrm{g} \to \mathrm{gg}$, reduced angle for secondary $\mathrm{q} \to \mathrm{qg}$ as in =3 and no angular ordering for $\mathrm{g} \to \mathrm{q\overline{q}}$ as in =4.

b-jet Shapes

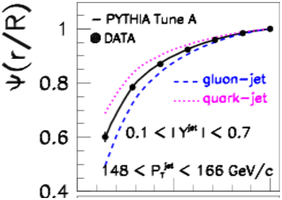
Document(s)	Web Page Public Note
Contact(s)	A. Lister
Abstract	We present preliminary results on the integrated jet shapes of b-jets in inclusive b-jet production in p-pbar collisions at sqrt{s} = 1.96 TeV. The data used for this analysis were collected between February 2002 and September 2004 and represent an integrated luminosity of about 300 pb-1. The measurements are carried out for jets with rapidity | yjet | < 0.7 and transverse momentum between 52 and 300 GeV/c. The measured b-jet shapes are corrected to the particle level and compared to PYTHIA-Tune A and HERWIG predictions.
Comments	Last Update: October 2006 Dataset: 300 pb-1

This measurement shows that, despite relatively large systematic uncertainties, the measured b-quark jet shapes are significantly different from those expected from Pythia Tune A and Herwig Monte Carlo simulations. This difference seems to be in part explained by the fact that the fraction of b-quark jets that originate from flavour creation (where a single b-quark is expected inside the same jet cone) over those that originate from gluon splitting (where two b-quarks are expected to be inside the same jet cone) is slightly different in Monte Carlo predictions than in data. This measurement can help in the tuning of the fraction of gluon splitting to flavour creation b-quark jets in the Monte Carlo simulation. This tuning is particularly important for the extrapolation up to LHC energies where many searches will involve b-quark jets.

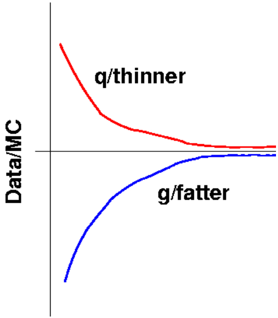
Jet Shapes



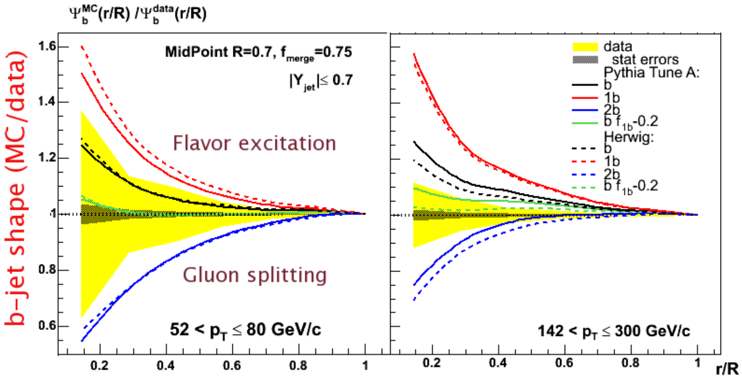
Ψ(r)
r
R



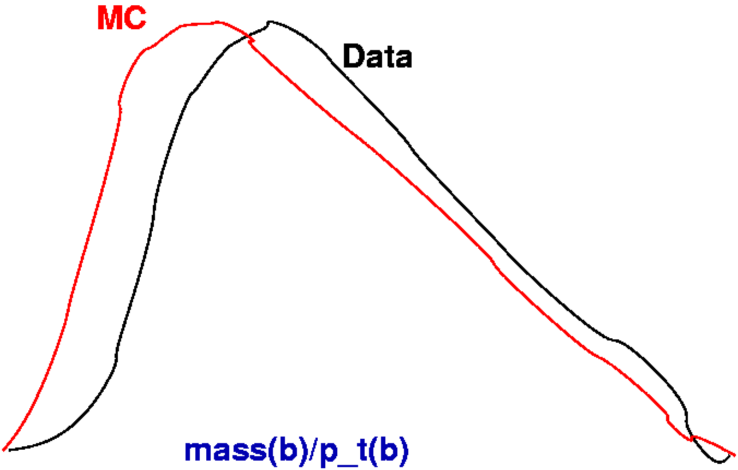
Ψ(r/R)
1
0.8
0.6
0.4
– PYTHIA Tune A
DATA
gluon–jet
quark–jet
0.1 < | Y^jet | < 0.7
148 < P_T^jet < 166 GeV/c



Data/MC
q/thinner
g/fatter



- Even after correcting $g \rightarrow b\bar{b}$, jet shapes differ
- MC/data > 1 means jets are thinner



Vista cartoon

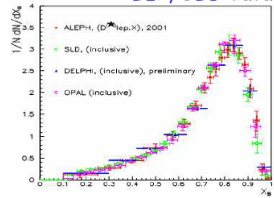
Tevatron Top Mass

- JES corrections are derived for q jets
- What is the systematic in applying this to b jets?

B fragmentation

- The problem:
  - Fraction of parton energy carried by B hadron (x_B) is described by a special function.
  - Bowler (cf. Peterson) parametrization describes LEP data well, but parameters are uncertain.
  - Results in an "extra" uncertainty on the energy scale specific to b jets.
- So far @ CDF:
  - Vary Peterson parameter in LEP-allowed range, run PEs in TMT-1D, take 0.5*largest shift as systematic for all analyses.

LEP, SLD data



- So far @ D0:
  - Reweight events using alternate Bowler & Peterson params, run PEs, add shifts in quadrature.

B Fragmentation

- Recommendation:
  - Compare ADO, SLD param values for syst.
  - Complication: pythia default is nowhere in the ballpark! So compare pythia default vs ADO/SLD? Or try to use ADO as default?
  - Reweight events by their relative probability: maximize stat power.

Bowler parameter values

	r	a	b
Pythia default	1.0	0.3	0.58
ADO	0.897 ± 0.013	1.03 ± 0.08	1.31 ± 0.08
SLD	0.98 ± 0.01	1.30 ± 0.09	1.58 ± 0.09

- To do:
  - Correlations in uncertainties above?
  - Study pythia default vs ADO/SLD. How important is it?
  - Exact prescription for reweighting.

Tuning of b-quark Fragmentation

Typically use $x_{\rm B} = \dfrac{2E_{\rm B}}{E_{\rm CM}}$ to tune:

1. theorist's model (nothing to do with event generators)
2. Petersen, etc. in string (other) fragmentation
3. Bowler or Lund fragmenation form (some universal features)

$$f(z) \propto \frac{1}{z^{1+r_Q b m_Q^2}} z^{a_\alpha} \left(\frac{1-z}{z}\right)^{a_\beta} \exp\left(-\frac{b m_\perp^2}{z}\right) .$$

DØ Strategy

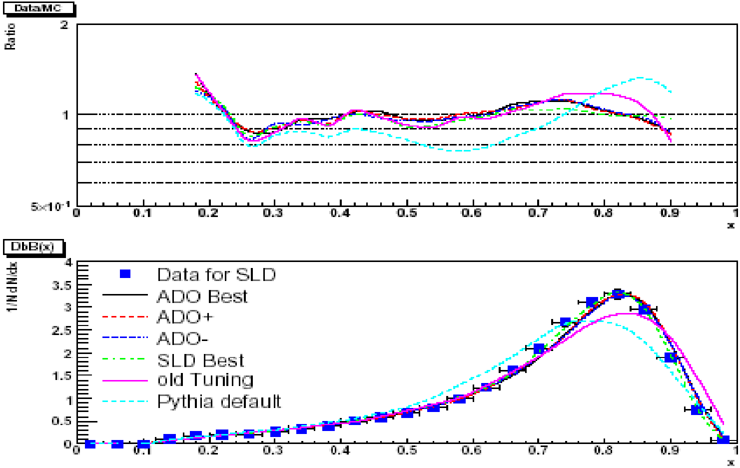
Claim:

1. Default `Pythia` tune is not adequate
2. q-fragmentation not sensitive to $a - b$

Strategy:

1. Fix $a - b$
2. extract data from LEP
3. use SLD as a cross check

Preliminary and for demonstration purposes only!



Improvements:

- consistently fit q-frag as well
- “proper” errors (can we implement inside the generators?)
- public tunes

This would all be facilitated if the LEP/SLD data were public

B semileptonic decays

- The problem:
  - B hadron branching ratios are not known perfectly.
  - Different decay f.s. affect energy scale through f_{EM}, $<p_T>$, etc.
  - Most importantly (?), semileptonic decays contain ν and μ, with large effect on cal response.
- So far @ D0:
  - Reweight events using 1σ exp. uncertainties on semi-lep BRs.
  - Run PEs for variation of
    - B->lep BRs
    - B->D->lep BRs
    - D->lep BRs
- So far @ CDF:
  - Determine average b-jet energy scale in s.l. vs non-s.l. b jets.
  - Derive 1-sigma shift in <b-JES>, convert to shift in M_{top}.

B semileptonic decays

- Recommendation:
  - Shift semi-leptonic decay BRs by up-to-date 1σ, reweight events by relative probability.

- To do:
  - Prescription for shift in BRs, calculation of relative weights.
  - Could other BRs also be important??

τ

- At the Tevatron, τ's are skinny jets
- Skinny jets come from fragmentation to a leading particle

Knuteson-Mrenna Lemma

- All fake τ's come from q fragmentation

Model:

$$\frac{P(g \to \tau)}{P(q \to \tau)} \sim \frac{\exp\left(-2C_F \int_{(\mathrm{GeV})^2}^{p_T^2} \mathcal{P}\right)}{\exp\left(-C_F \int_{(\mathrm{GeV})^2}^{p_T^2} \mathcal{P}\right)} << 1$$

Almost all single particle fakes come from the same source:

$$
\begin{aligned}
p(j \to \gamma) \sim & \quad p(q \to \pi^0)\epsilon(\pi^0 \to \gamma(\gamma)) \\
p(j \to \tau) \sim & \quad p(q \to \pi^{\pm})\epsilon(\pi^{p}m \to \tau) \\
p(j \to e) \sim & \quad p(q \to \pi^0)\epsilon(\gamma \to e(e)) \\
p(j \to \mu) \sim & \quad p(q \to \pi^{\pm})\epsilon(\pi^{p}m \to \mu)
\end{aligned}
$$

ϵ's are calculable

Understanding Fakes

Knuteson, Culbertson, *et al.*

	e^+	e^-	μ^+	μ^-	τ^+	τ^-	γ	j
e^+	62154	33	0	0	1161	1	3749	25913
e^-	24	62300	0	0	0	1156	3730	25817
μ^+	0	0	50330	0	15	0	0	596
μ^-	0	1	0	50294	0	11	0	573
γ	1381	1326	0	0	8	14	67732	21372
π^0	1196	1208	0	0	25	34	59727	31651
π^+	266	0	115	0	72113	42	117	23908
π^-	1	352	0	88	80	71491	169	24499
K^+	150	1	272	1	73333	36	49	21670
K^-	1	249	0	163	112	71701	151	23654

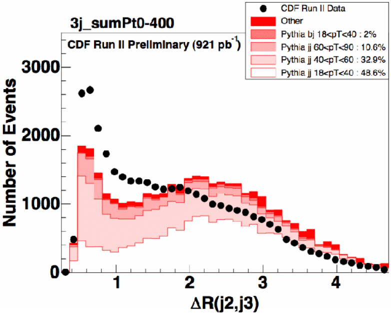
TABLE XIV: Central single particle misidentification matrix. Using a single particle gun, 10^5 particles of each type shown at the left of the table were shot with $p_T = 25$ GeV into the central CDF detector, uniformly distributed in θ and in ϕ.

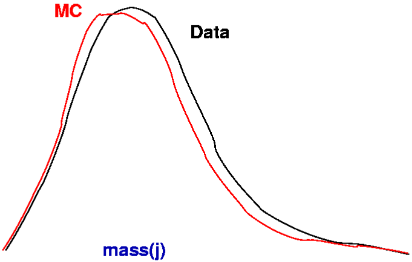
A consistent understanding of fakes in such a model has been obtained in the Vista analysis at CDF

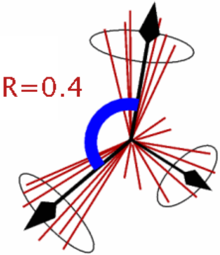
Skinny Jets

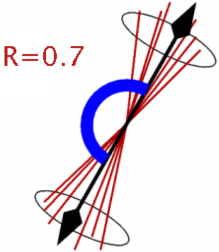
- In general, jet *masses* have never been modelled well
- M_j is mostly set by $\Lambda_{\rm FSR}$
- One is constrained by LEP shape measurements
- However, jet masses not directly available using cone (any?) algorithm
- Would like to appeal to the (non-public) LEP data to resolve the issue

Case in Point: dR(j2,j3) and minMass(j)





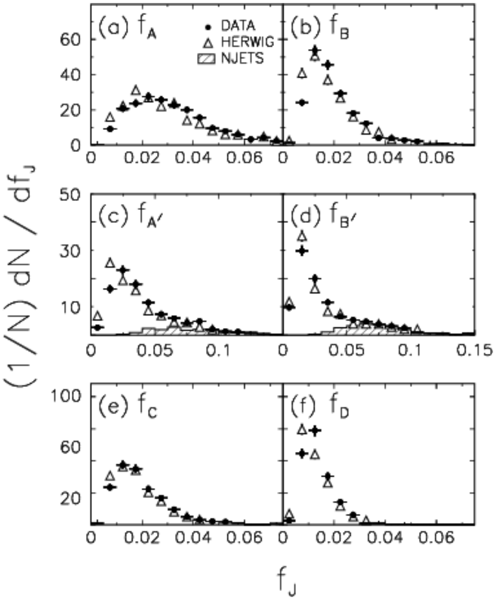




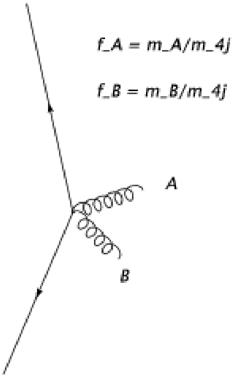
How is this possible?

Steve Geer led an effort to study multijets in Run I

- Nice analysis of 3-, 4-, and 5-jet production
- Comparison of Herwig and simple models to the data
- Some notable discrepancies (f_i = mass fractions)
- These are not hidden in the text
- Main discrepancies dropped at the end when quoting overall goodness-of-fit



hep-ex/9605004, CDF-RunI

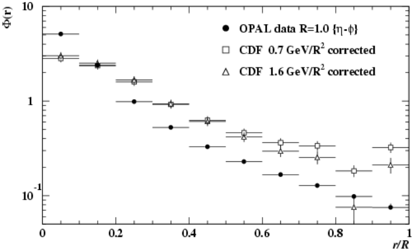


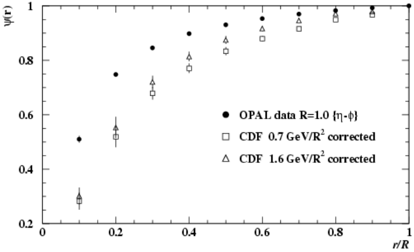
Kinematic set-up

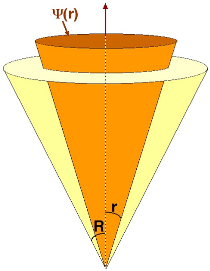
is χ^2/NDF = 1.21 (63 degrees of freedom). The observed distributions are described less well by the HERWIG parton shower Monte Carlo predictions, for which the X_4, $\cos\theta_{3'}$, $\psi_{3'}$, and $\cos\theta_{3''}$ distributions have χ^2s significantly poorer than those for the corresponding NJETS predictions. Restricting the comparison to those distributions predicted by both the NJETS and HERWIG calculations (i.e. all distributions except the single-body mass fraction distributions) we find the overall χ^2 per degree of freedom for the HERWIG comparison of the combined three-jet distributions is χ^2/NDF = 1.58 (45 degrees of freedom), for the combined four-jet distributions χ^2/NDF = 1.63 (63 degrees of freedom), and for the combined five-jet distributions χ^2/NDF = 1.52 (63 degrees of freedom).

f_i removed from the overall fit

no NJETS prediction for small f_i







- LEP Jets .NE. TeV Jets
- Attributed to either UE or gluon jets
- Implies TeV Jets fatter!
- Would be useful to have access to the Z pole data

Notes

- Problem is rather "universal"
- High statistics
- Doesn't seem to depend on jet definition
- Doesn't seem to depend on detector
- Doesn't seem to depend on generator
- Reproducible in orthogonal analyses

Conclusions I Draw

Work needs to be done on:

- global properties of b fragmentation in event generators including correlation with q frag
- measurement and theoretical understanding of q/g fragmentation near the endpoints where they are fakes

A great step forward for LC Sim efforts would be:

- Collecting LEP information on event shapes and fragmentation *along with the expertise to use it*
- Work in conjunction with Run2 analyses to reanalyze the LEP data