# Fragmentation and Hadronization

Bryan Webber

LP99, 13 August 1999

- Jet fragmentation – theory
- Hadronization models
- Single-particle yields and spectra
- Quark and gluon jets
- Event shapes and power corrections
- Current and target fragmentation
- Heavy quark fragmentation
- Bose-Einstein correlations
- WW fragmentation
- Conclusions

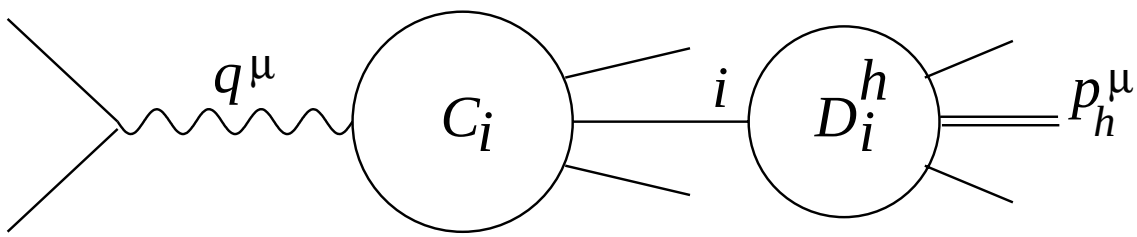
- Apologies for not giving citations in the talk, and for not having time to discuss:

  - Fragmentation function parametrizations
  - Photon fragmentation function
  - Polarization in hadronization
  - Fluctuations and intermittency
  - Two-particle correlations
  - Heavy quark production in jets ($g \to c\bar{c}, b\bar{b}$)
  - Transverse energy flow in deep inelastic scattering
  - Underlying event in DIS and hadron-hadron collisions
  - Identified particle production in DIS
  - Tests of QCD coherence
  - Jet profiles and substructure

- See the fragmentation web pages of the European network "Hadronic Physics with High Energy Electromagnetic Probes" (HaPHEEP) for much useful information:
  `http://www.pv.infn.it/ jakob/fragmentation.html`

# Jet fragmentation – theory

Recall the basic factorization structure of the single-particle inclusive distribution, e.g. in $e^+e^- \to hX$:




$$F^h(x,s) = \sum_i \int_x^1 \frac{dz}{z} C_i(z, \alpha_s(s)) D_i^h(x/z, s)$$

$$s = q^2 \,, \qquad\qquad x = 2p_h \cdot q/q^2 = 2E_h/E_{cm}$$

$C_i$ are the coefficient functions for this particular process (including all selection cuts etc.) and $D_i^h$ is the universal fragmentation function for parton $i \to$ hadron $h$.

The fragmentation functions are not perturbatively calculable but their $s$-dependence (scaling violation) is given by the DGLAP equation:

$$s\frac{\partial}{\partial s} D_i^h(x,s) = \sum_j \int_x^1 \frac{dz}{z} P_{ji}(z, \alpha_s(s)) D_j^h(x/z, s)$$

Thus they can be parametrized at some fixed scale $s_0$ and then predicted at other energies.

In certain kinematic regions, higher-order corrections are enhanced by large logarithms, which need to be resummed.

- At small $x$, leading $\log x$ enhanced terms can be resummed by changing the DGLAP equation to

$$s\frac{\partial}{\partial s}D_i^h(x,s) = \sum_j \int_x^1 \frac{dz}{z} P_{ji}(z, \alpha_{\mathrm{S}}(s)) D_j^h(x/z, z^2 s)$$

  The effect of this is to generate a characteristic hump-backed shape in the variable $\xi = \ln(1/x)$, with a peak at $\xi_p \sim \frac{1}{4}\ln s$.

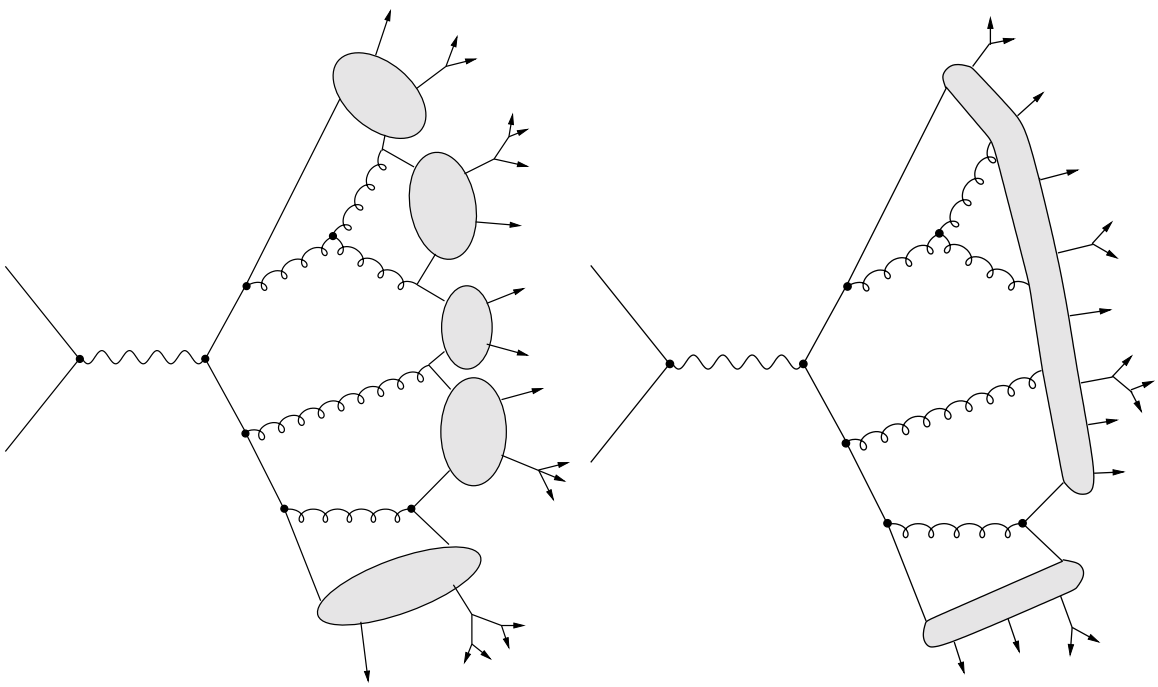
- Large logarithms of ratios of invariants may also appear inside the coefficient functions $C_i$, for example in three-jet events when the angles between jets become small. In some cases these can be absorbed into a change of scale in the fragmentation functions.

Although universal, fragmentation functions are factorization scheme dependent. The splitting functions $P_{ji}$ are also scheme dependent. To specify the scheme requires calculation of the coefficient functions to (at least) next-to-leading order. This has only been done in a few cases. Thus there is need for theoretical work to make full use of the data on fragmentation functions.

# General ideas

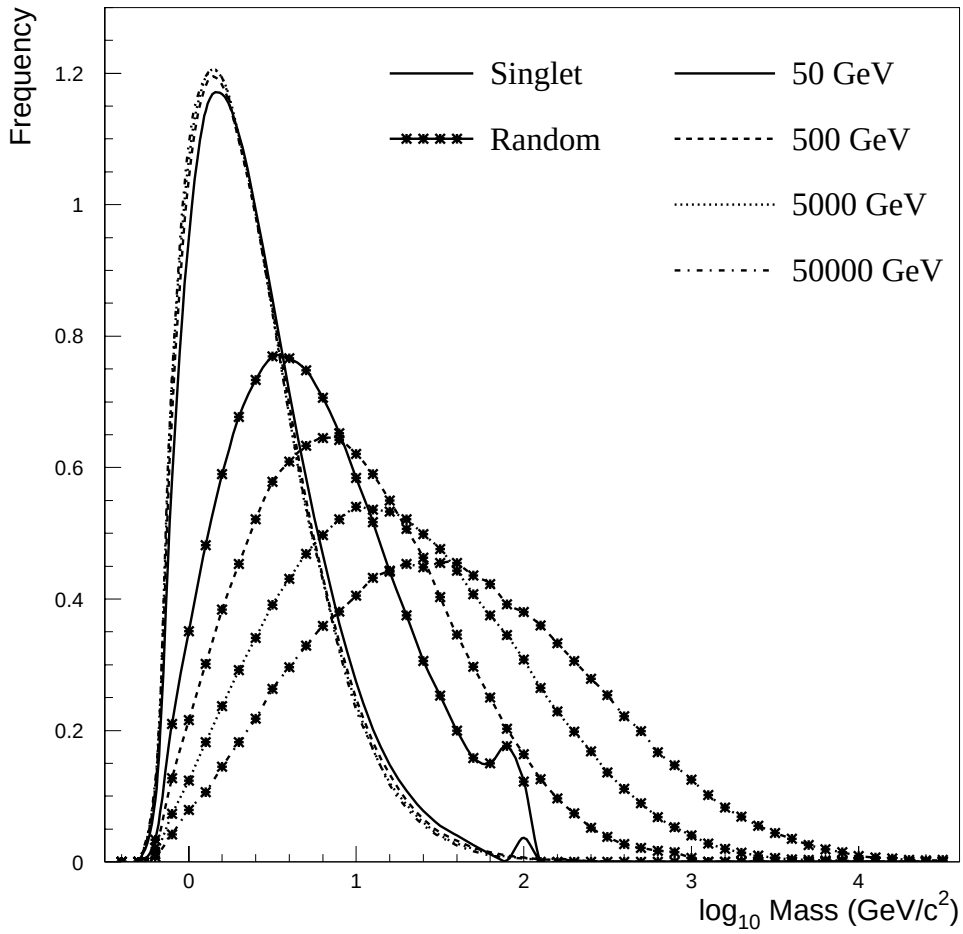
- Local parton-hadron duality
  - Inclusive spectra and multiplicities
- Universal low-scale $\alpha_S$
  - Heavy quark spectra, event shapes

# Specific models



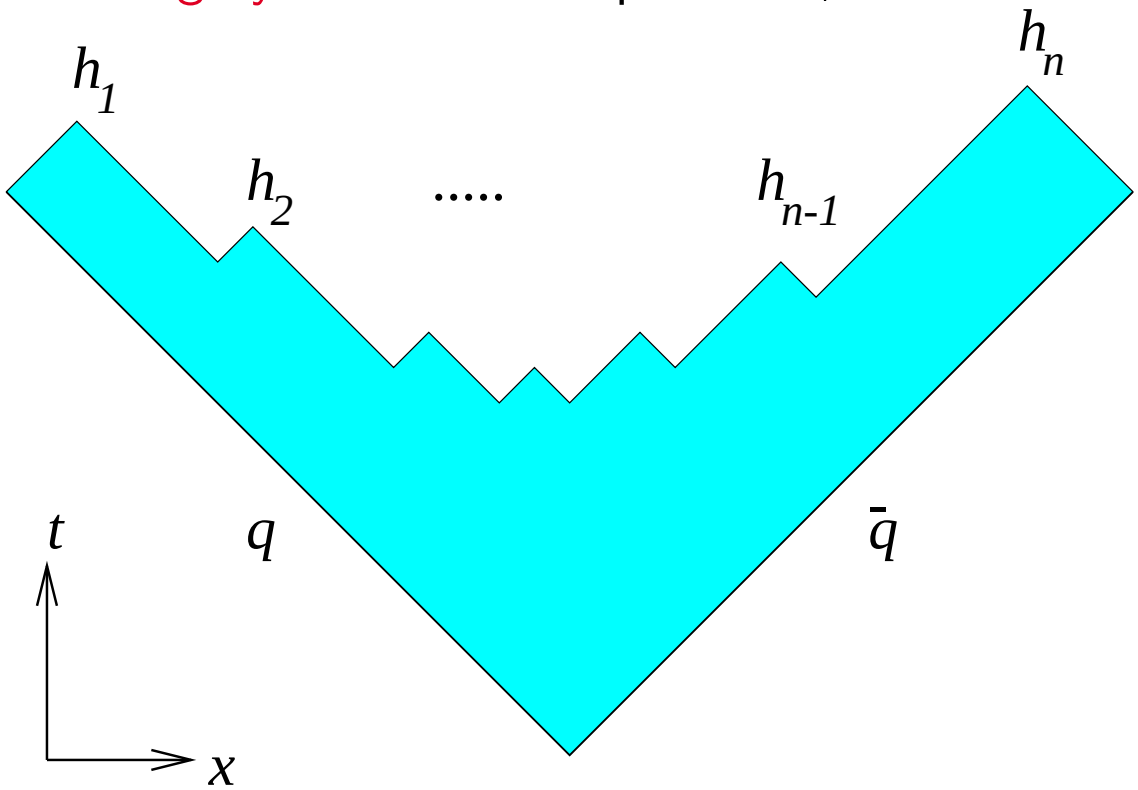
# Cluster (HERWIG)

- Uses preconfinement property of parton shower:




- Few parameters: natural $p_T$ and heavy particle suppression
- Problems with massive clusters, baryons, heavy quarks

- String (JETSET)
  - Uses string dynamics: linear potential, area law




$$|M(q\bar{q} \to h_1 \cdots h_n)|^2 \propto e^{-bA}$$

  - Extra parameters for $p_T$ and heavy particle suppression
  - Some problems with baryons

- String (UCLA)
  - Takes area law more seriously (mass suppression)
  - Extra parameters for $p_T$
  - Some problems with baryons

# Single-particle yields and spectra

## Mesons

| Particle | Multiplicity | HERWIG 5.9* | JETSET 7.4 | UCLA 7.4 | Expts |
|---|---|---|---|---|---|
| Charged | 20.96(18) | 20.95 | 20.95 | 20.88 | ADLMO |
| $\pi^{\pm}$ | 17.06(24) | 17.41 | 16.95 | 17.04 | ADO |
| $\pi^{0}$ | 9.43(38) | 9.97 | 9.59 | 9.61 | ADLO |
| $\eta$ | 0.99(4) | 1.02 | 1.00 | 0.78 | ALO |
| $\rho(770)^{0}$ | 1.24(10) | 1.18 | 1.50 | 1.17 | AD |
| $\omega(782)$ | 1.09(09) | 1.17 | 1.35 | 1.01 | ALO |
| $\eta'(958)$ | 0.159(26) | 0.097 | 0.155 | 0.121 | ALO |
| $f_0(980)$ | 0.155(8) | 0.111 | $\sim 0.1$ | — | ADO |
| $a_0(980)^{\pm}$ | 0.14(6) | 0.162 | — | — | O |
| $\phi(1020)$ | 0.097(7) | 0.104 | 0.194 | 0.132 | ADO |
| $f_2(1270)$ | 0.188(14) | 0.186 | $\sim 0.2$ | — | ADO |
| $f_2'(1525)$ | 0.012(6) | 0.021 | — | — | D |
| $K^{\pm}$ | 2.26(6) | 2.16 | 2.30 | 2.24 | ADO |
| $K^{0}$ | 2.074(14) | 1.98 | 2.07 | 2.06 | ADLO |
| $K^{*}(892)^{\pm}$ | 0.718(44) | 0.670 | 1.10 | 0.779 | ADO |
| $K^{*}(892)^{0}$ | 0.759(32) | 0.676 | 1.10 | 0.760 | ADO |
| $K_2^{*}(1430)^{0}$ | 0.084(40) | 0.111 | — | — | DO |
| $D^{\pm}$ | 0.187(14) | 0.161 | 0.174 | 0.196 | ADO |
| $D^{0}$ | 0.462(26) | 0.506 | 0.490 | 0.497 | ADO |
| $D^{*}(2010)^{\pm}$ | 0.181(10) | 0.151 | 0.242 | 0.227 | ADO |
| $D_s^{\pm}$ | 0.131(20) | 0.115 | 0.129 | 0.130 | O |
| $B^{*}$ | 0.28(3) | 0.201 | 0.260 | 0.254 | D |
| $B_{u,d}^{**}$ | 0.118(24) | 0.013 | — | — | D |
| $J/\psi$ | 0.0054(4) | 0.0018 | 0.0050 | 0.0050 | ADLO |
| $\psi(3685)$ | 0.0023(5) | 0.0009 | 0.0019 | 0.0019 | DO |
| $\chi_{c1}$ | 0.0086(27) | 0.0001 | — | — | DL |

# Baryons

| Particle | Multiplicity | HERWIG 5.9* | JETSET 7.4 | UCLA 7.4 | Expts |
|---|---|---|---|---|---|
| p | 1.04(4) | 0.863 | 1.19 | 1.09 | ADO |
| $\Delta^{++}$ | 0.079(15) | 0.156 | 0.189 | 0.139 | D |
| | 0.22(6) | 0.156 | 0.189 | 0.139 | O |
| $\Lambda$ | 0.399(8) | 0.387 | 0.385 | 0.382 | ADLO |
| $\Lambda(1520)$ | 0.0229(25) | — | — | — | DO |
| $\Sigma^{\pm}$ | 0.174(16) | 0.154 | 0.140 | 0.118 | DO |
| $\Sigma^{0}$ | 0.074(9) | 0.068 | 0.073 | 0.074 | ADO |
| $\Sigma^{\star\pm}$ | 0.0474(44) | 0.111 | 0.074 | 0.074 | ADO |
| $\Xi^{-}$ | 0.0265(9) | 0.0493 | 0.0271 | 0.0220 | ADO |
| $\Xi(1530)^{0}$ | 0.0058(10) | 0.0205 | 0.0053 | 0.0081 | ADO |
| $\Omega^{-}$ | 0.0012(2) | 0.0056 | 0.00072 | 0.0011 | ADO |
| $\Lambda_c^{+}$ | 0.078(17) | 0.0123 | 0.059 | 0.026 | O |

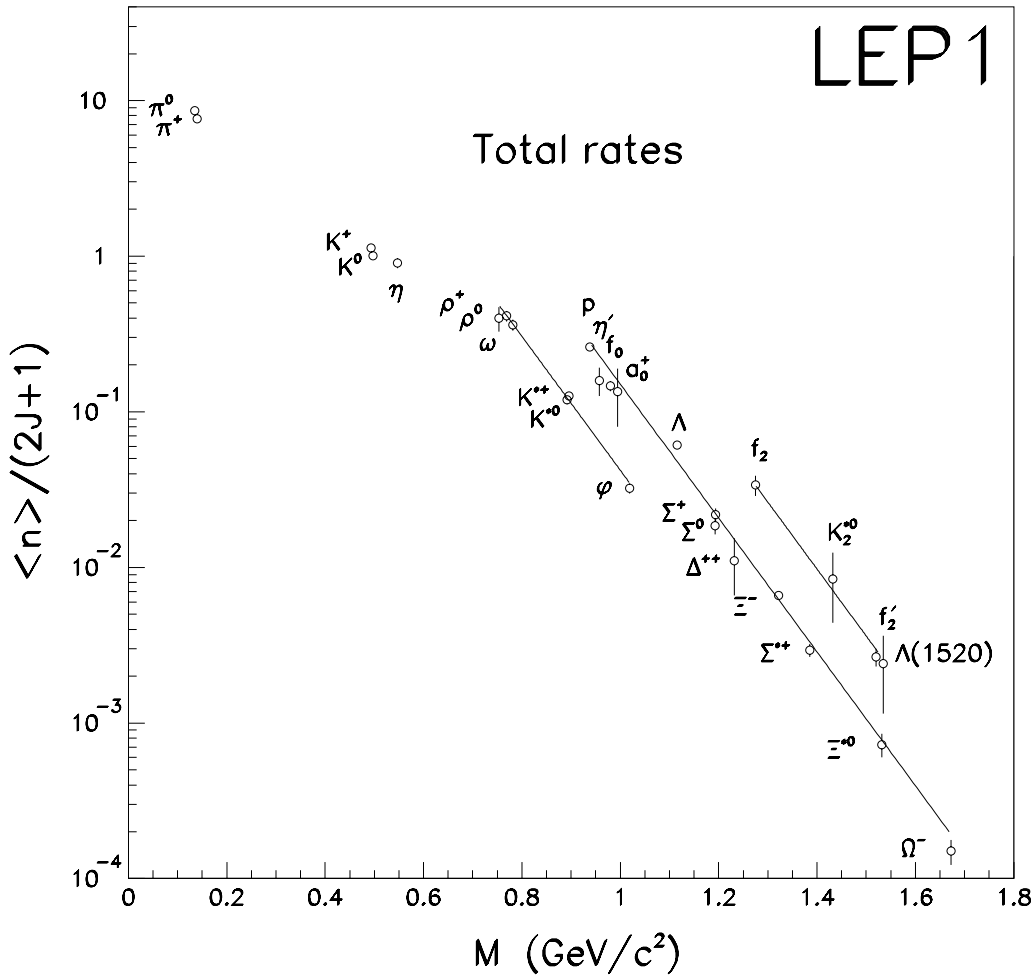
* Recent ALEPH tuning with strangeness suppression 0.8 (G. Rudolph)

- The orbitally-excited $J = \frac{3}{2}$ baryon $\Lambda(1520)$ is produced almost as much as the unexcited $J = \frac{3}{2}$ baryon $\Sigma(1385)$.

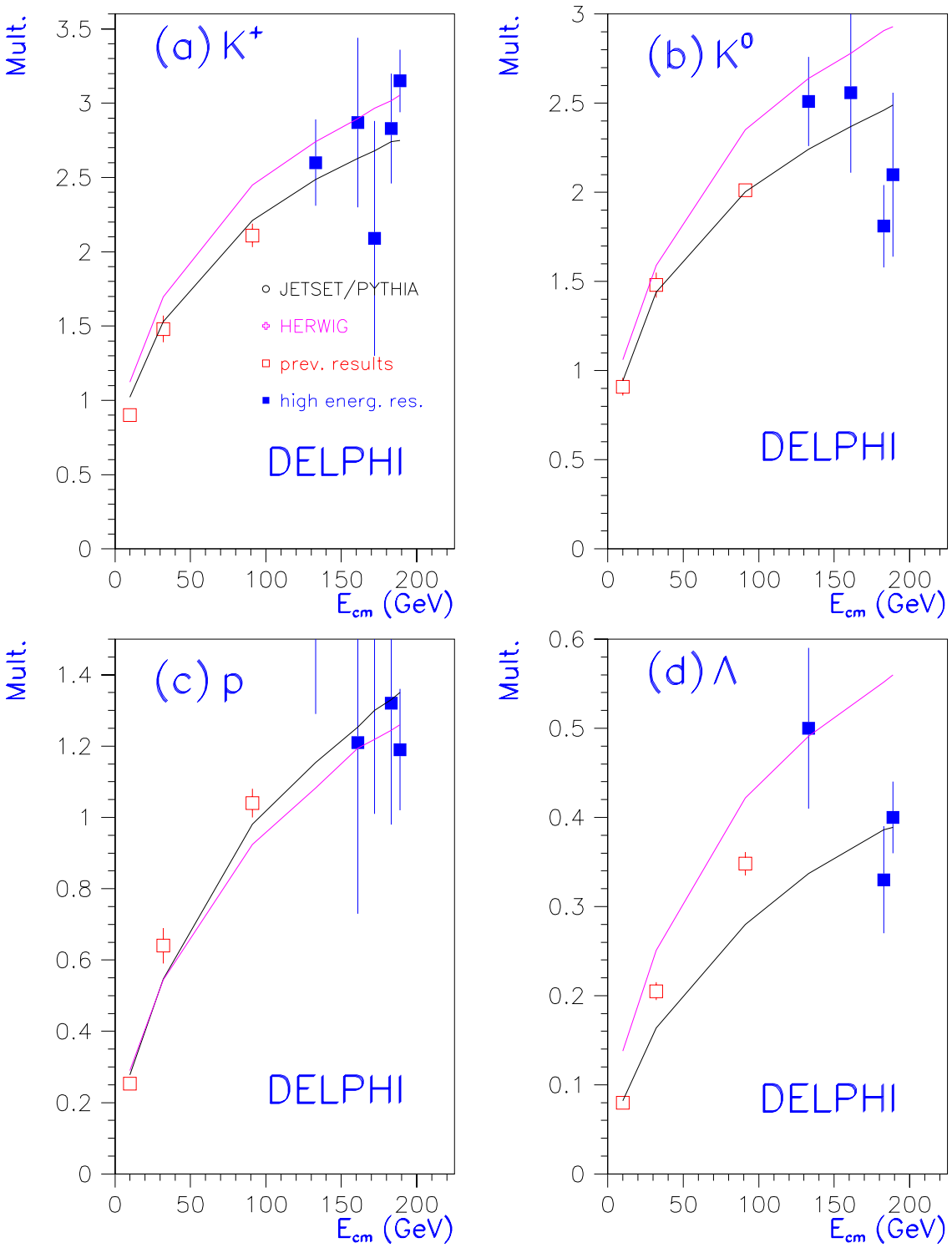
- It is remarkable that most measured yields (except for the $0^-$ mesons) lie on the family of curves

$$\frac{\langle n \rangle}{2J+1} = a e^{-M/T}$$

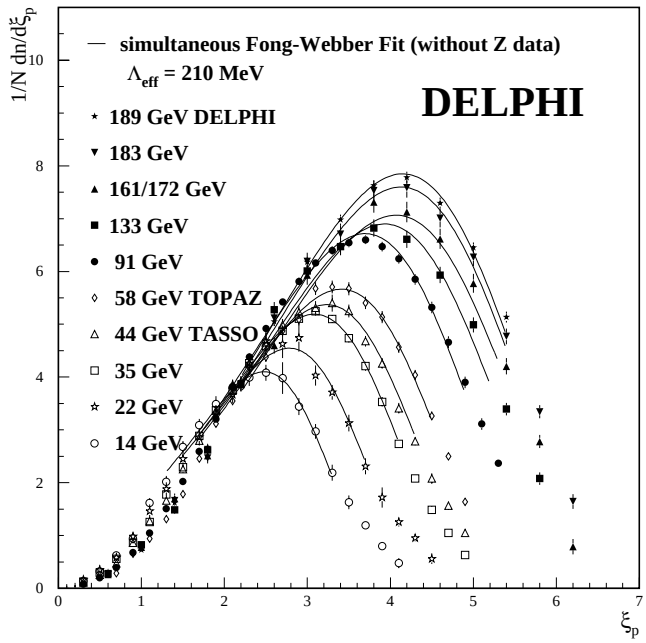
with $T \simeq 100$ MeV (Chliapnikov).

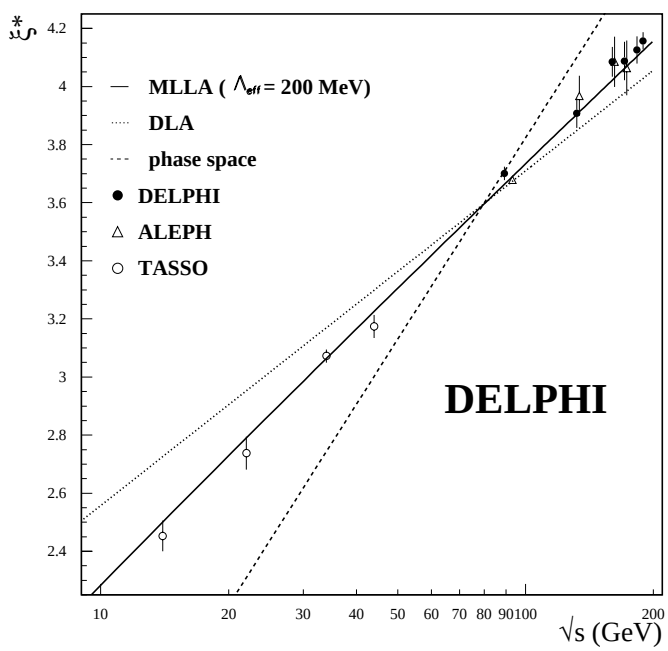


- Higher-energy yields are in broad agreement with models

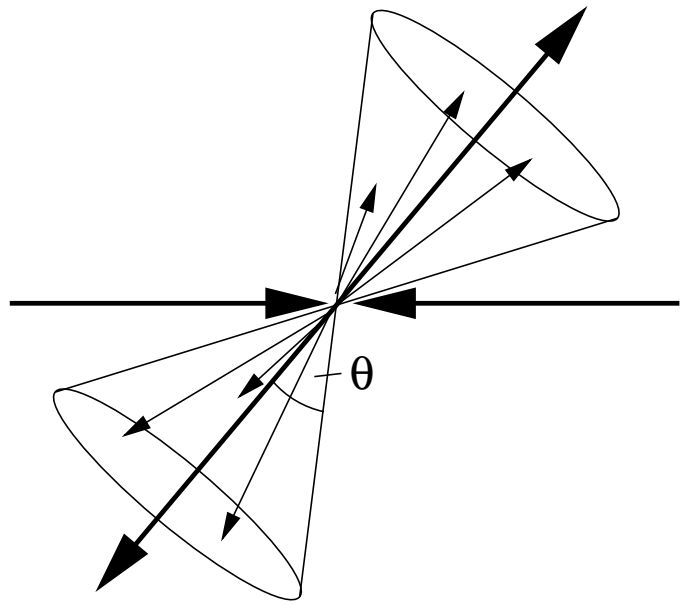


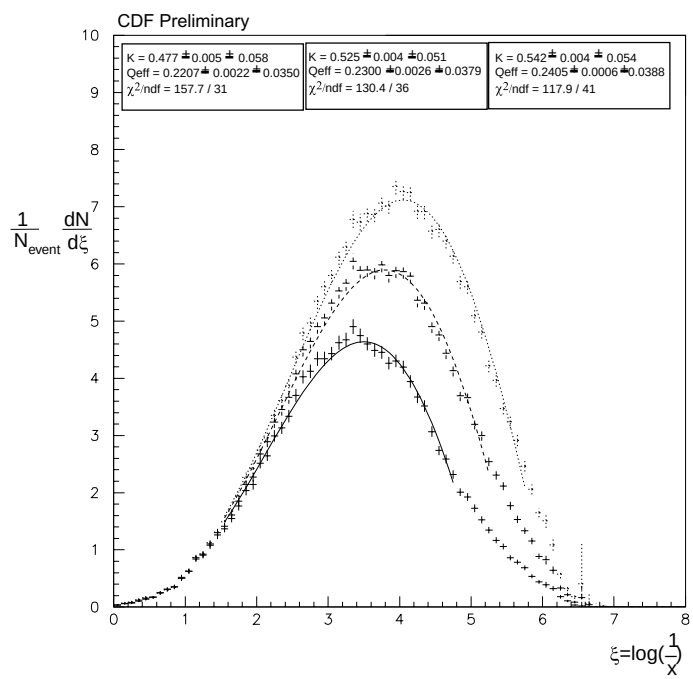
- In $e^+e^-$ annihilation, low-$x$ fragmentation functions agree remarkably well with the resummed (NLLA or MLLA) predictions over a wide energy range

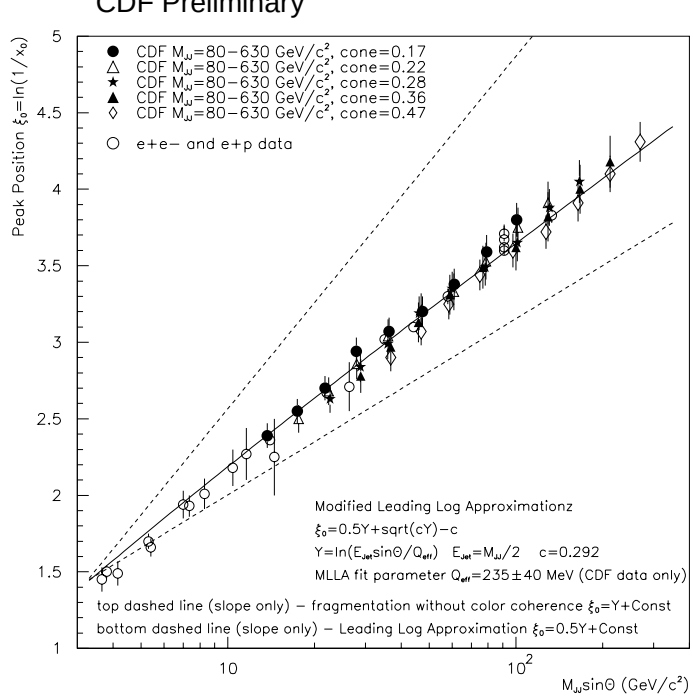


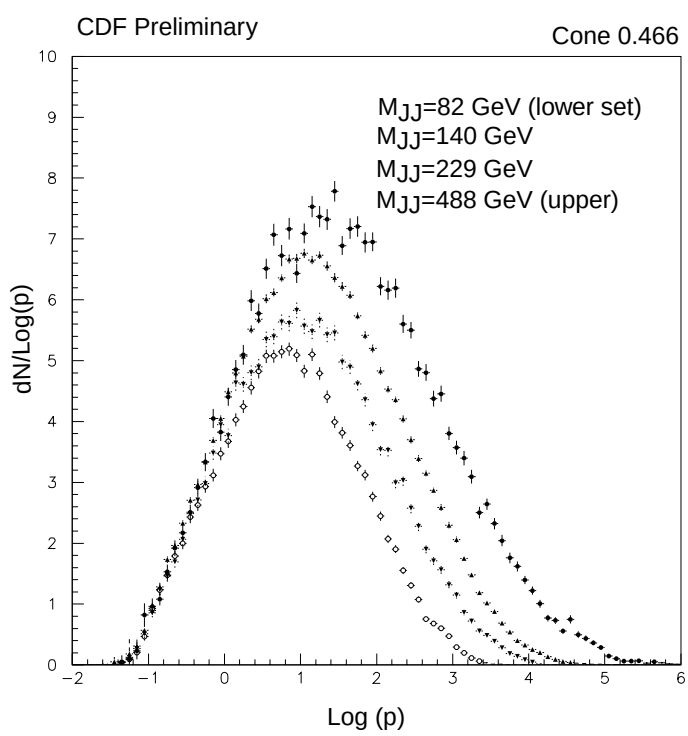
- In $pp \to$ dijets the relevant scale is taken to be $M_{JJ} \sin\theta$ where $M_{JJ}$ is the dijet mass and $\theta$ is the jet cone angle



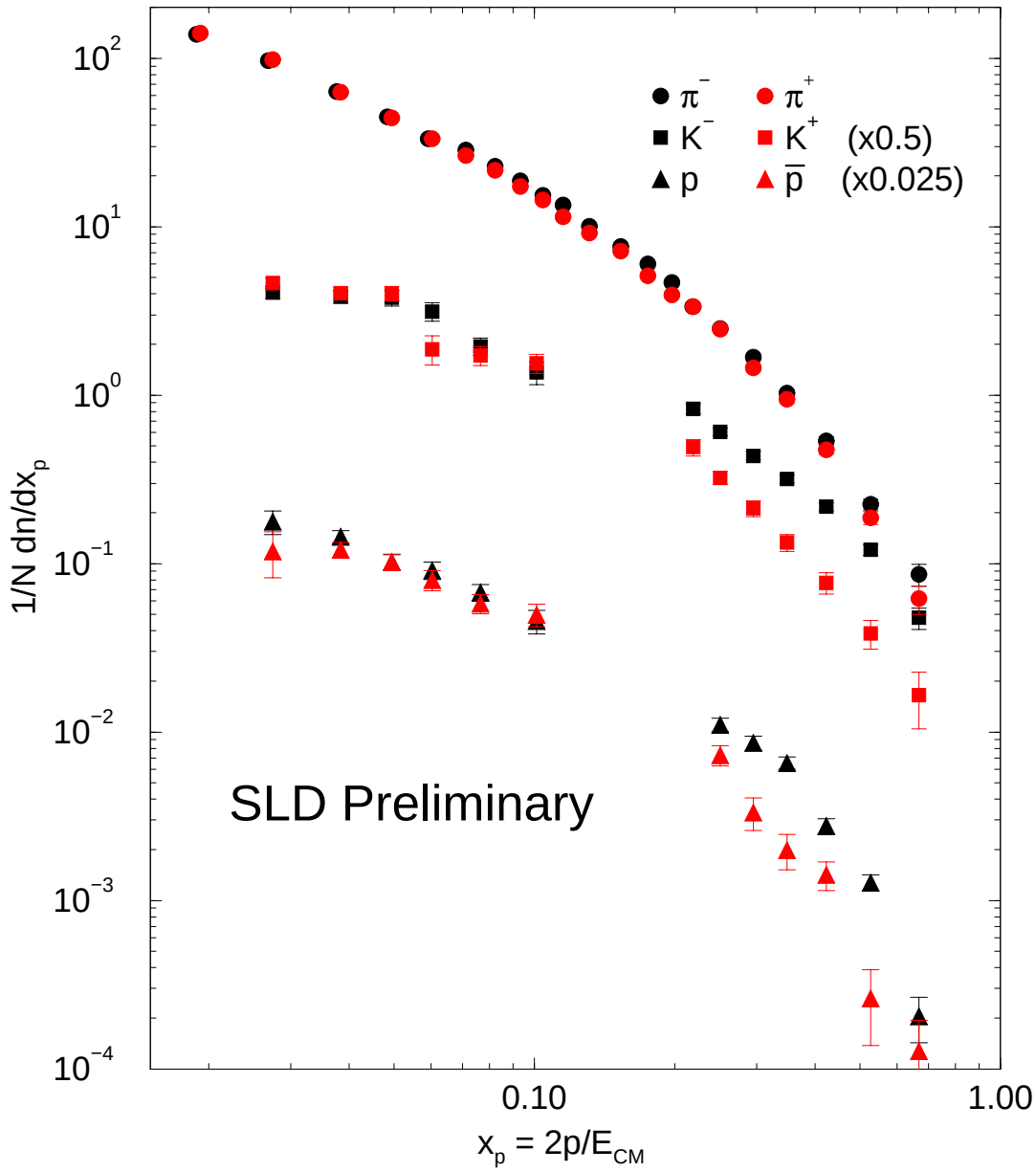





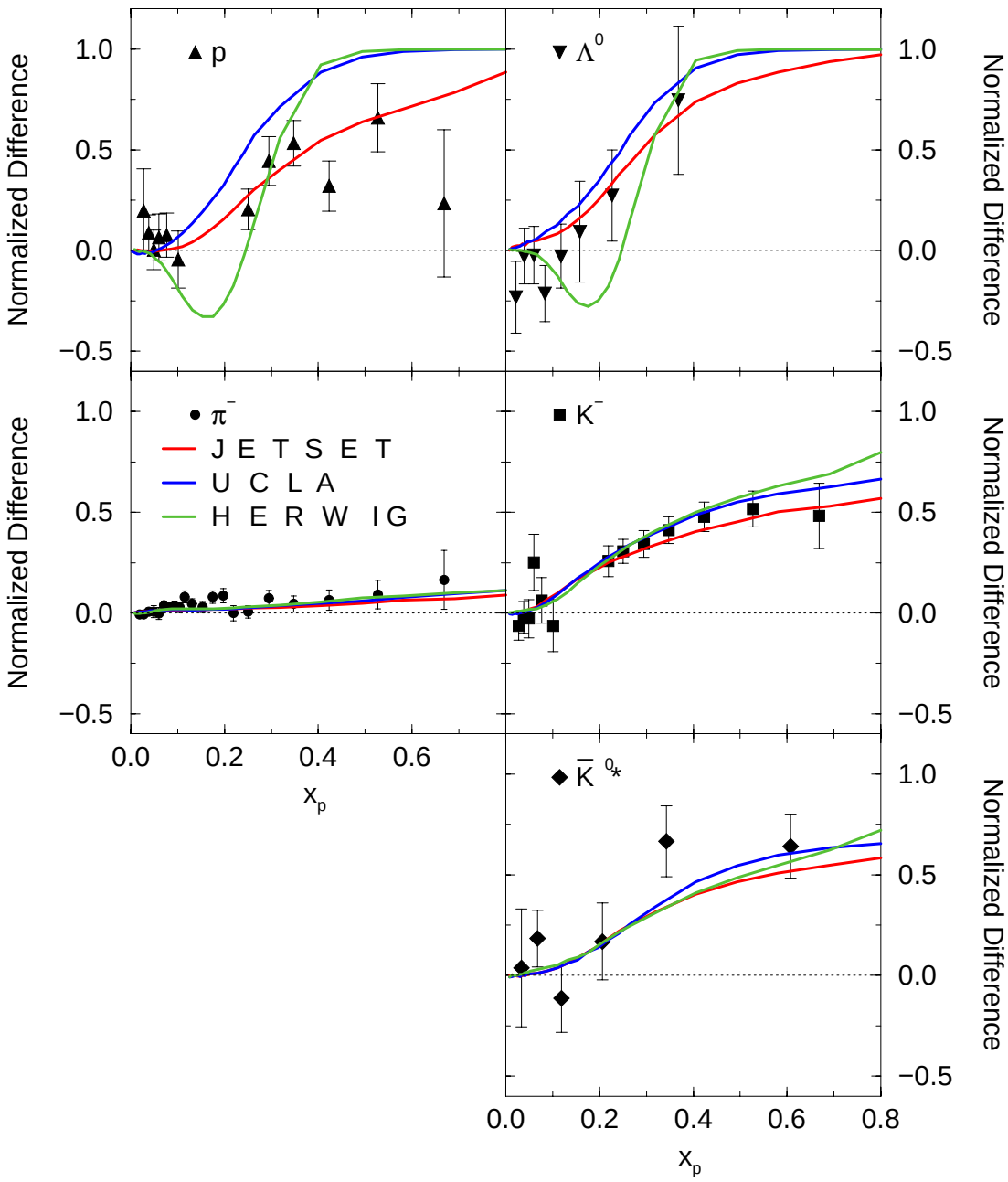

- When plotted against $\ln p$ instead of $\ln(1/x)$, the CDF data show clearly the scale-independence of the soft particle distribution (a test of colour coherence).



- New SLD data include hadron spectra in light quark (rather than antiquark) fragmentation, selected by hemisphere using SLC beam polarization:

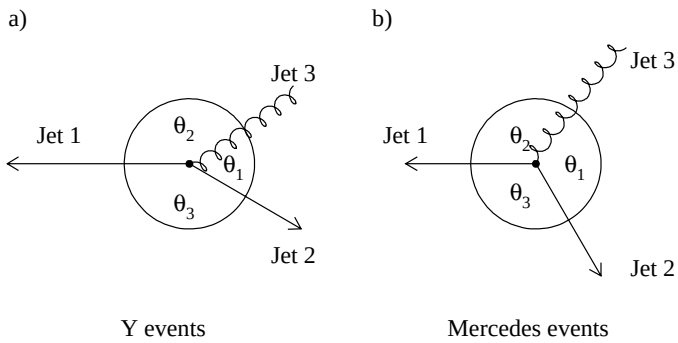


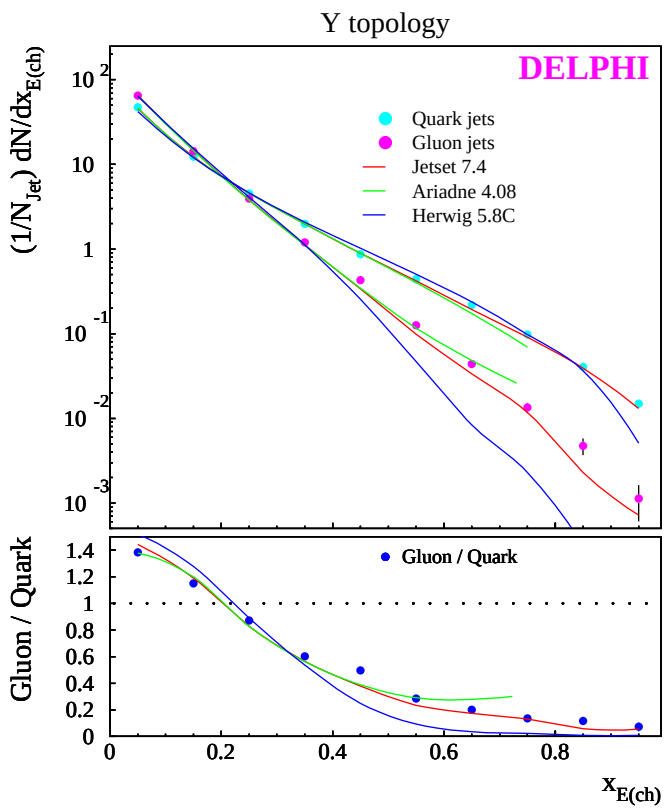
- Here one sees strong particle/antiparticle differences in the expected directions:

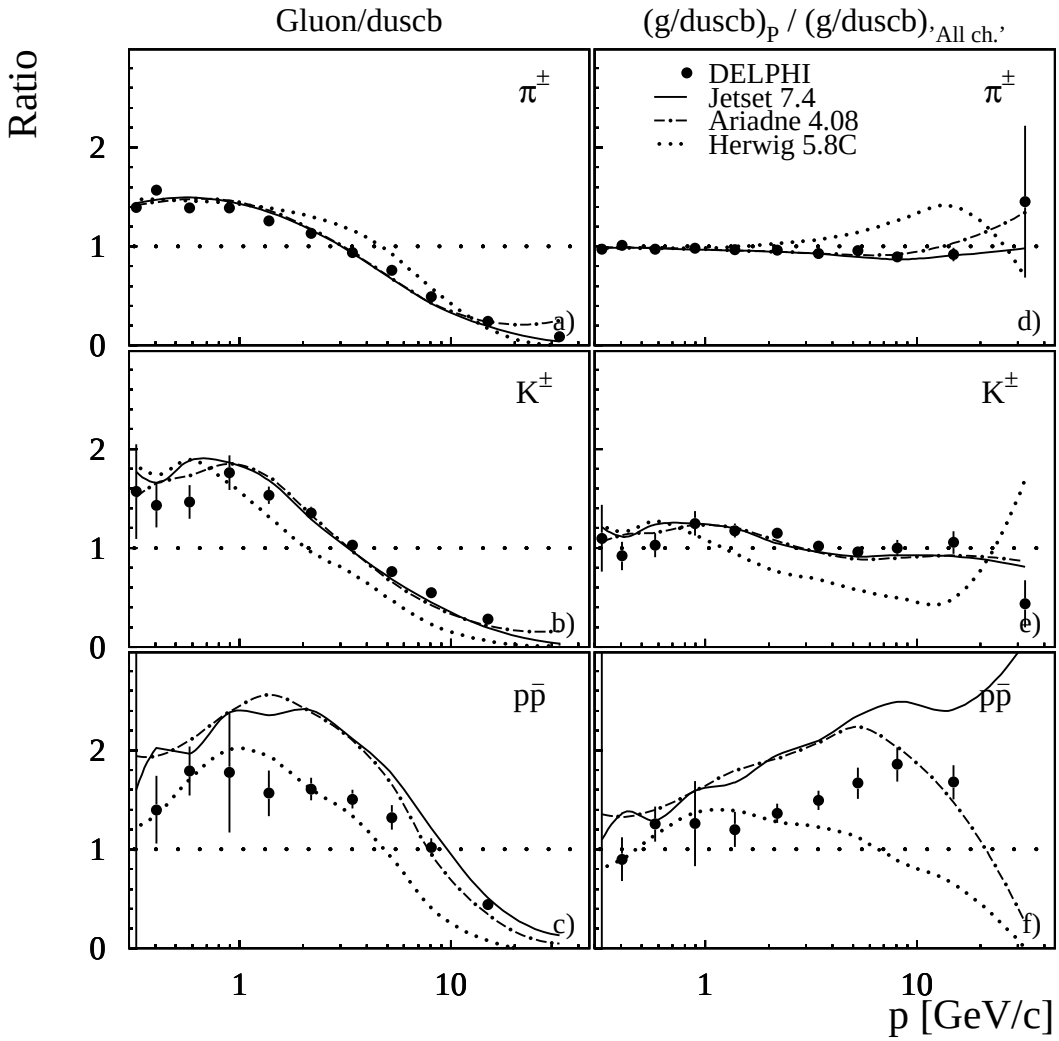


# Quark and gluon jets

- DELPHI select gluon jets by anti-tagging heavy quark jets in Y and Mercedes three-jet events



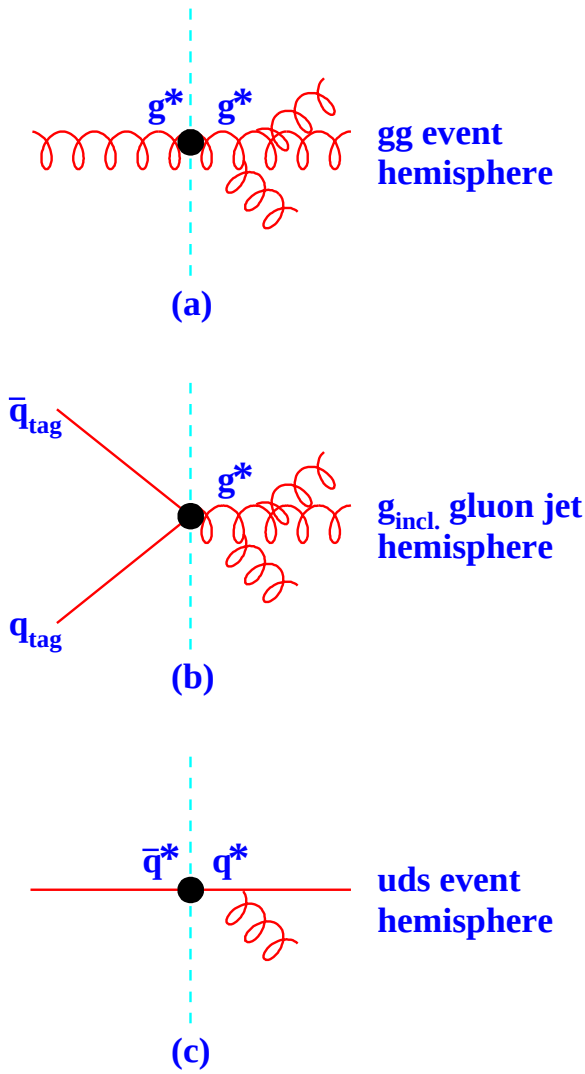


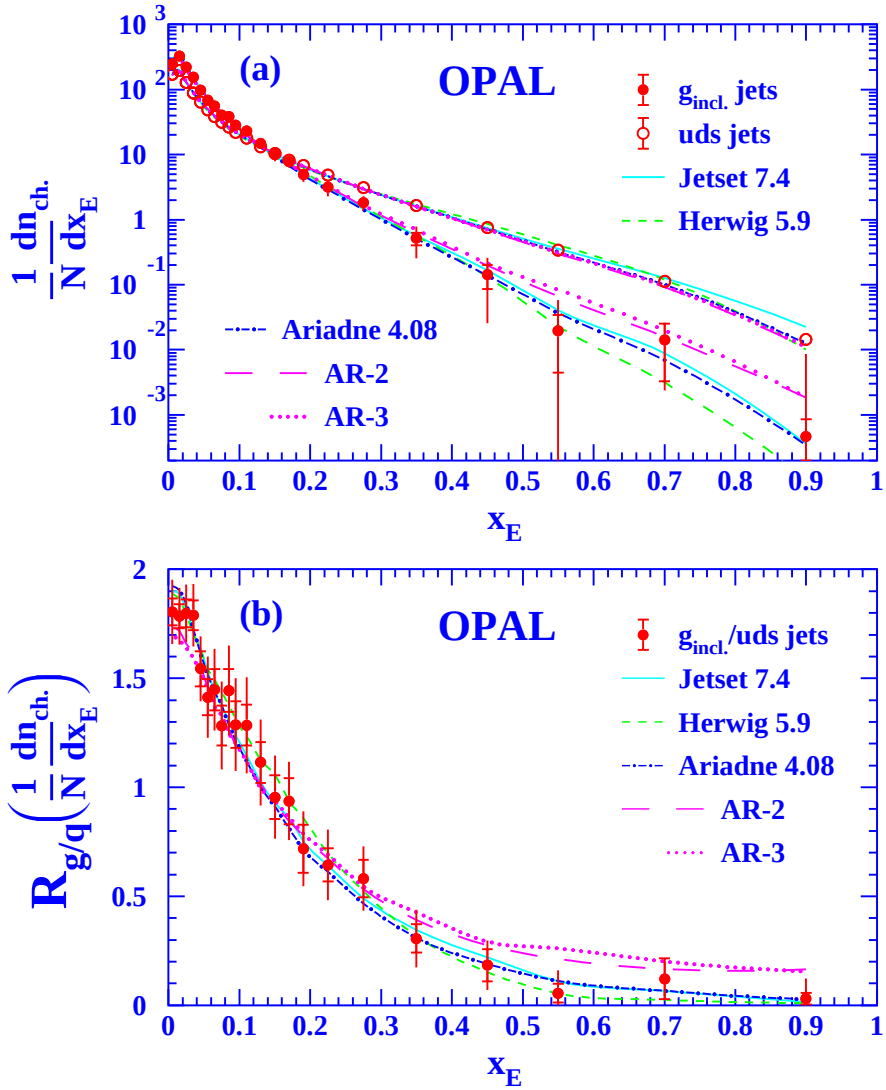

- In general the relative multiplicities of identified particles are consistent with those of all charged, with no clear excess of certain species in gluon jets (ALEPH, DELPHI, SLD)

  In particular there is no enhanced $\eta$ or $\phi(1020)$ production:

  OPAL: $N_g(\eta)/N_q(\eta) = 1.29 \pm 0.11$
  DELPHI: $N_g(\phi)/N_q(\phi) = 0.7 \pm 0.3$

- OPAL select gluon jets recoiling against tagged b-jets in the same hemisphere. Monte Carlo studies indicate that such jets should be similar to those emitted by a point source of gluon pairs (e.g. a $^1S_0$ $Q\bar{Q}$ state).

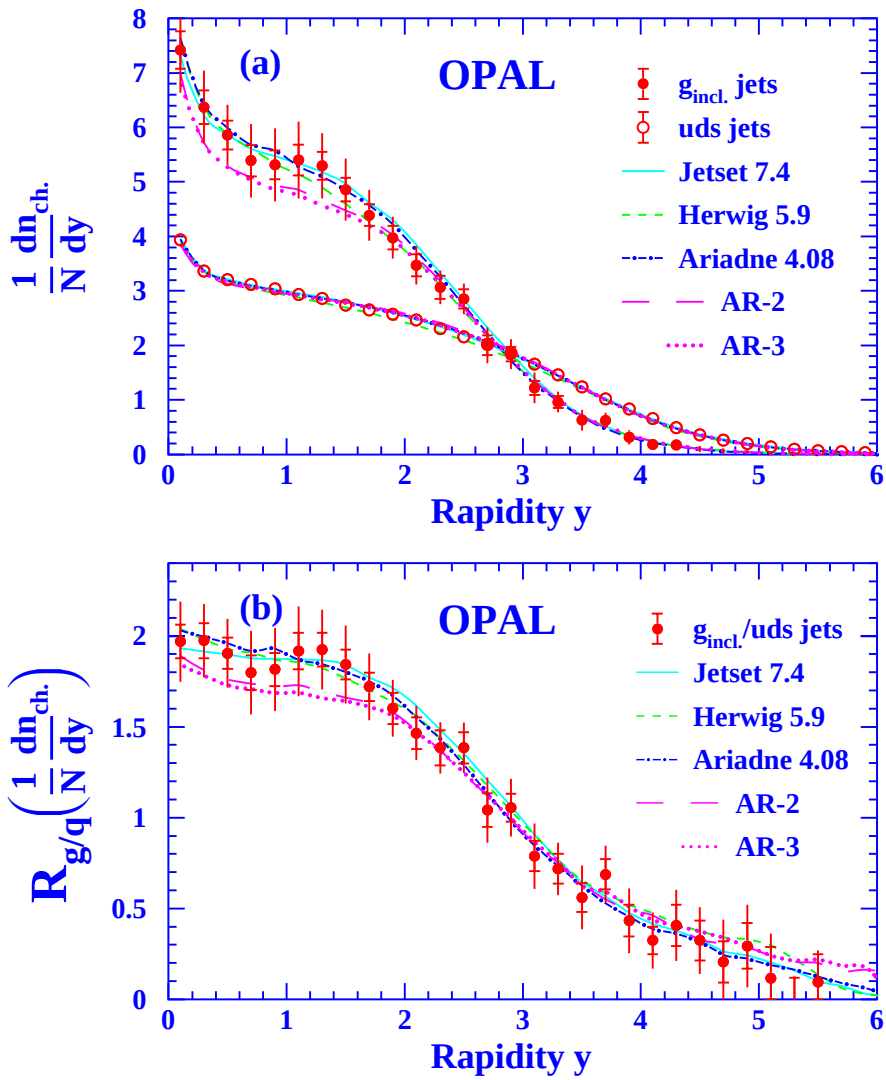





- Next-to-leading order calculations of the relevant coefficient functions are needed to check universality. However the qualitative message is clear:

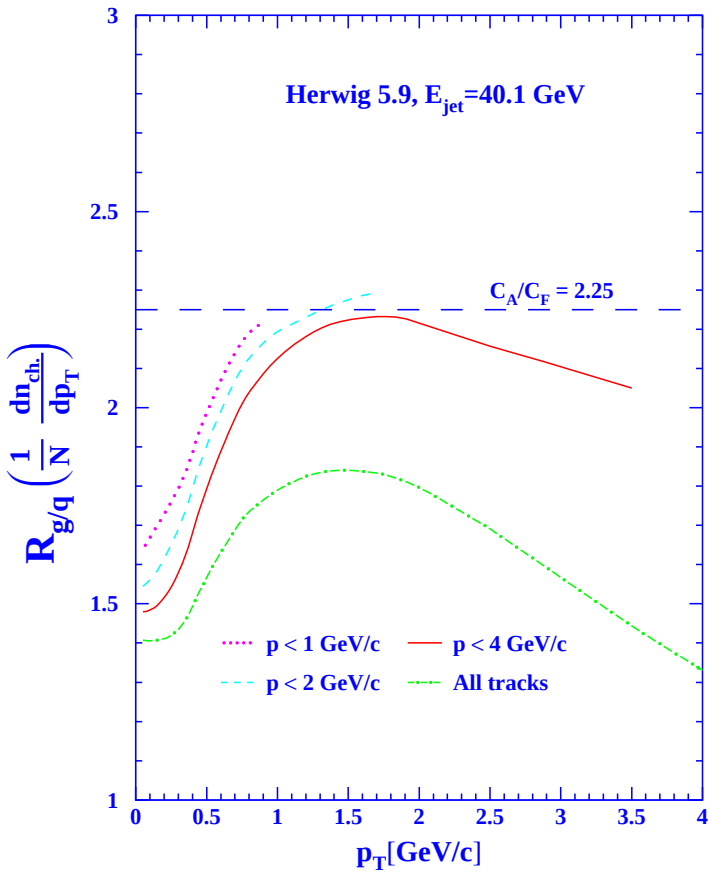
  Gluon jets have softer fragmentation functions than light quark jets, and higher charged multiplicity.

- The ratio of charged multiplicities at low rapidity (large angle) is close to the ratio of colour factors $r \equiv C_A/C_F = 2.25$, in agreement with Local Parton-Hadron Duality (LPHD).




OPAL: $r_{ch}(|y| < 1) = 1.919 \pm 0.047 \pm 0.095$

- Monte Carlo studies suggest that a better measure of $C_A/C_F$ is obtained by selecting low-momentum hadrons with relatively large $p_T$ (i.e. low rapidity)

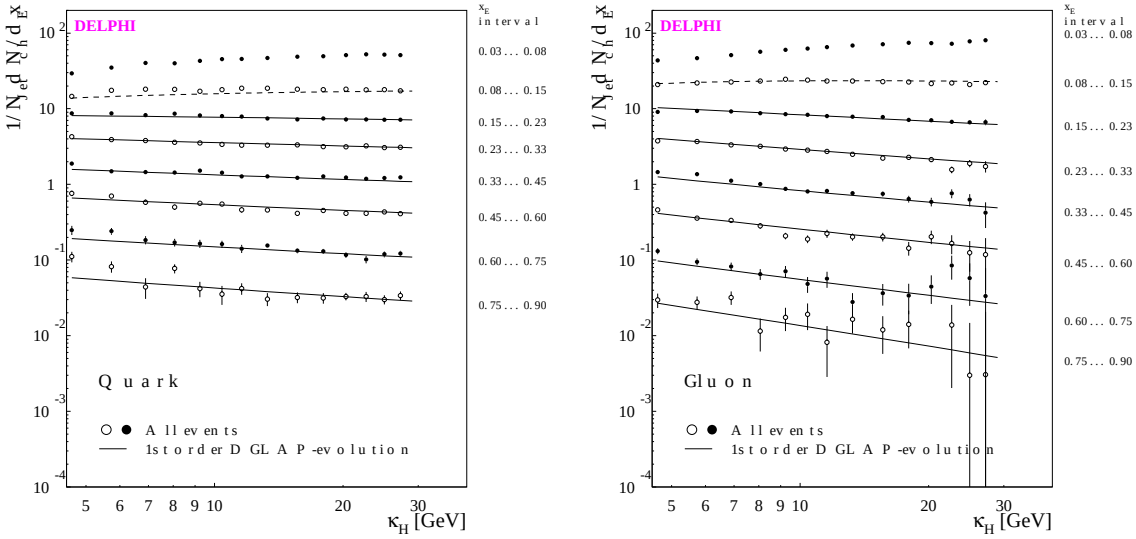


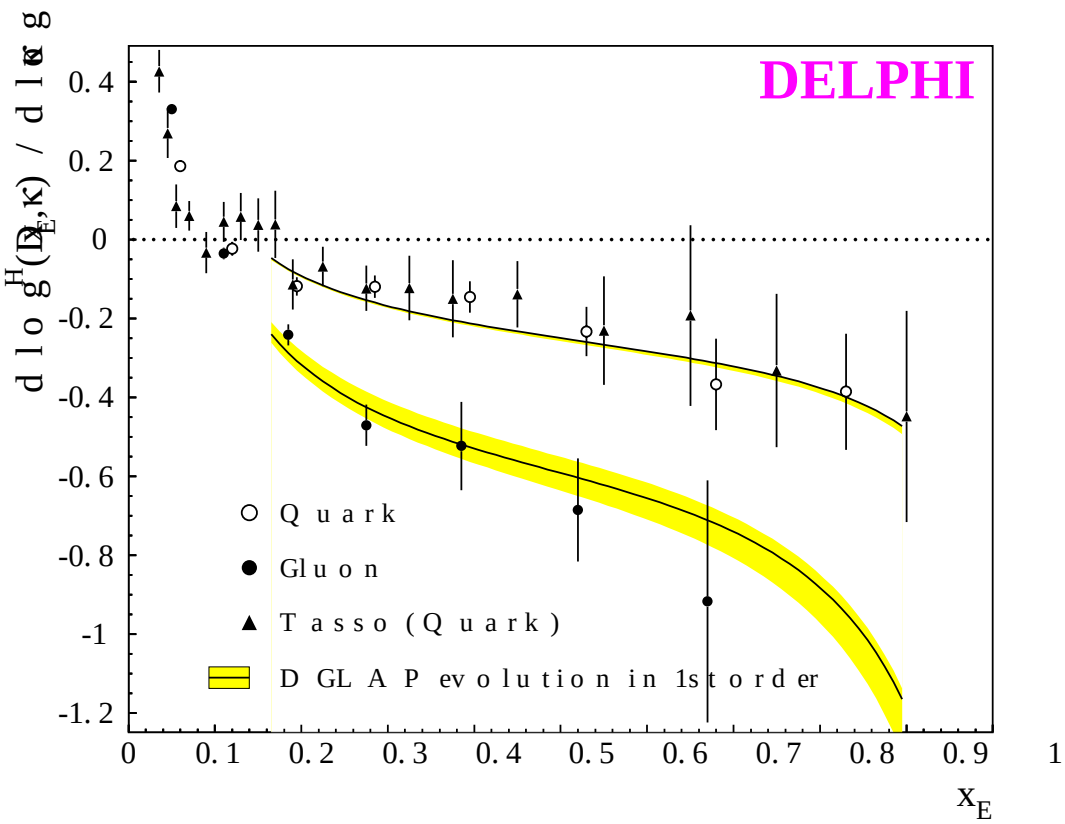

OPAL: $r_{ch}(p < 4,\ 0.8 < p_T < 3\text{GeV}) = 2.29 \pm 0.09 \pm 0.015$

- DELPHI have also observed scaling violation in gluon jet fragmentation by studying the dependence on

$$\kappa_H = E_{jet}\ \sin(\theta/2)\ \ (\simeq \tfrac{1}{2}\sqrt{sy_3})$$

where $\theta$ is the angle to the closest jet. This is expected to be the relevant scale, at least when $\theta$ gets small.



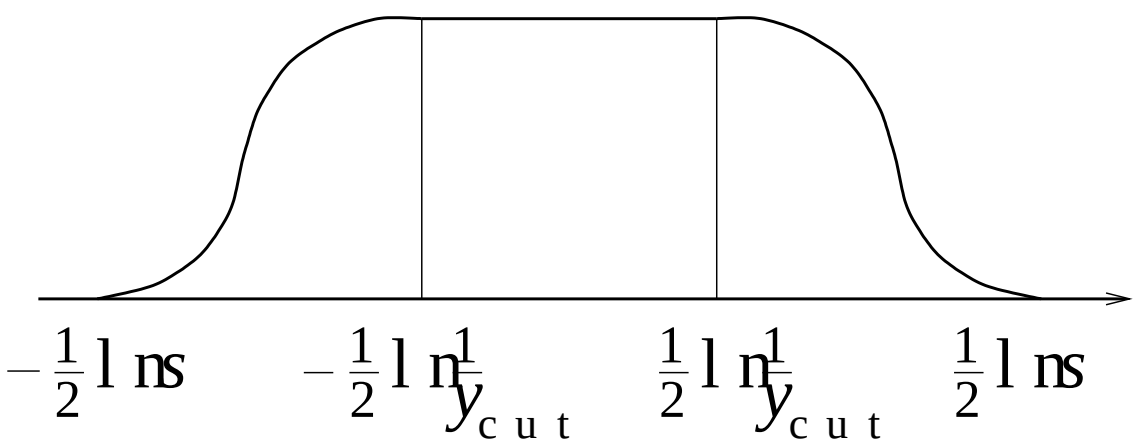




The ratio of scaling violations in quark and gluon jets provides another measure of $C_A/C_F$:

DELPHI: $r_{\text{sc.viol.}} = 2.23 \pm 0.09 \pm 0.06$

- A crucial point in the above study is that 3-jet events were not selected using a fixed jet resolution $y_{\text{cut}}$, but rather each event was clustered to precisely 3 jets. This avoids 'biasing' the gluon jet sample (by preventing further jet emission above $y_{\text{cut}}$).

- The same point is well illustrated by the analysis of average multiplicities in 2- and 3-jet events. If $N_{q\bar{q}}(s)$ is the 'unbiased' $q\bar{q}$ multiplicity, then in events with precisely 2 jets at resolution $y_{\text{cut}}$ we get a 'rapidity plateau':

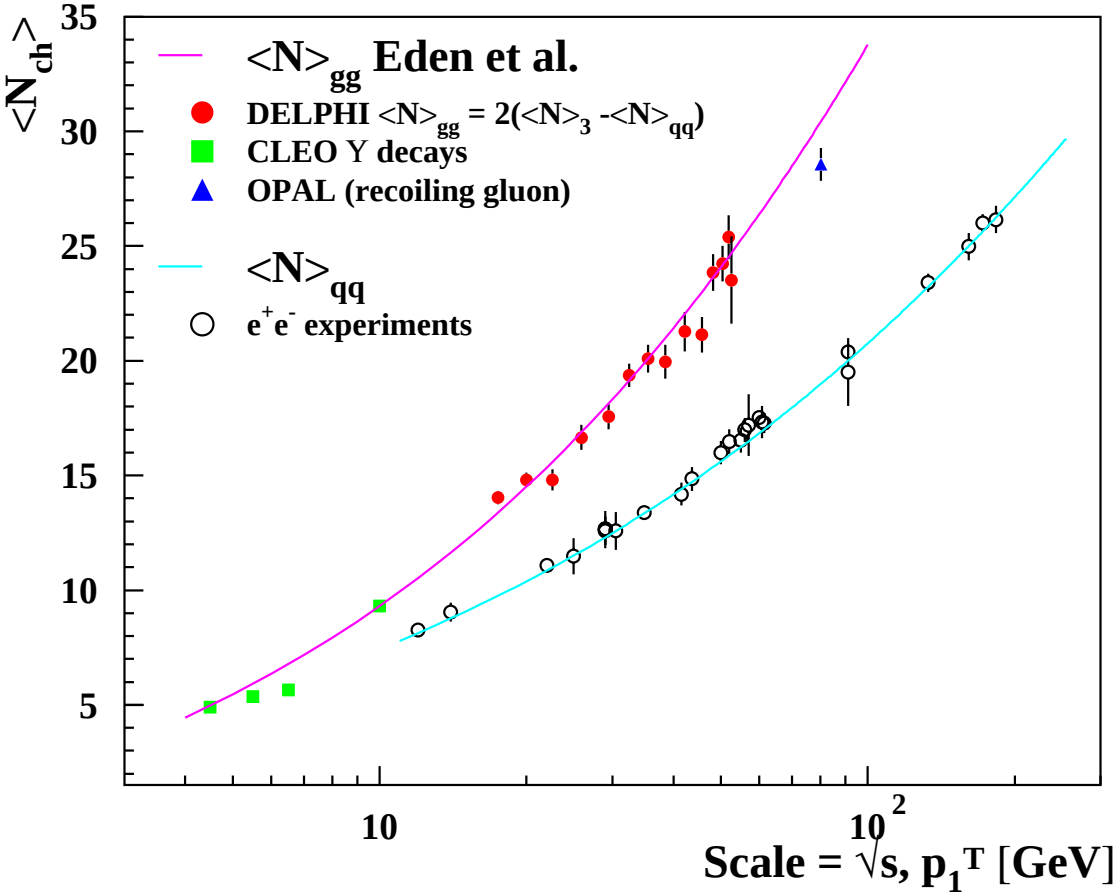



$$N_2(s, y_{\text{cut}}) \simeq N_{q\bar{q}}(s y_{\text{cut}}) + \ln(1/y_{\text{cut}}) N'_{q\bar{q}}(s y_{\text{cut}})$$

where $N'(s) \equiv s dN/ds$. Clustering each event to 3 jets we get this with $y_3$ in place of $y_{\text{cut}}$, plus an unbiased gluon jet:

$$N_3(s) \simeq N_2(s, y_3) + \tfrac{1}{2} N_{gg}(s y_3)$$

Thus one can obtain the unbiased $gg$ multiplicity (plotted here vs. $p_1^T \sim \sqrt{sy_3}$).

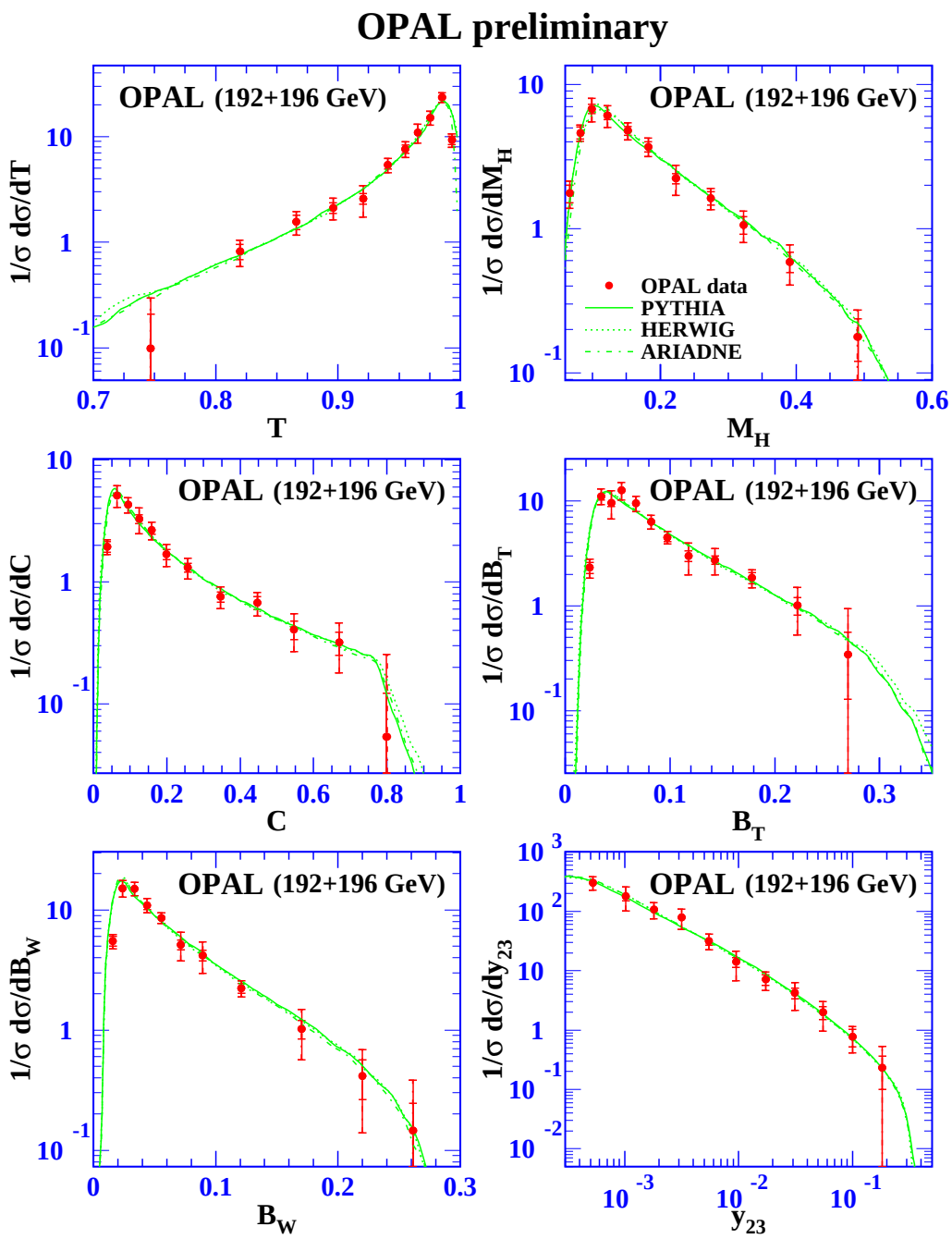



The ratio of slopes gives yet another measure of $C_A/C_F$:

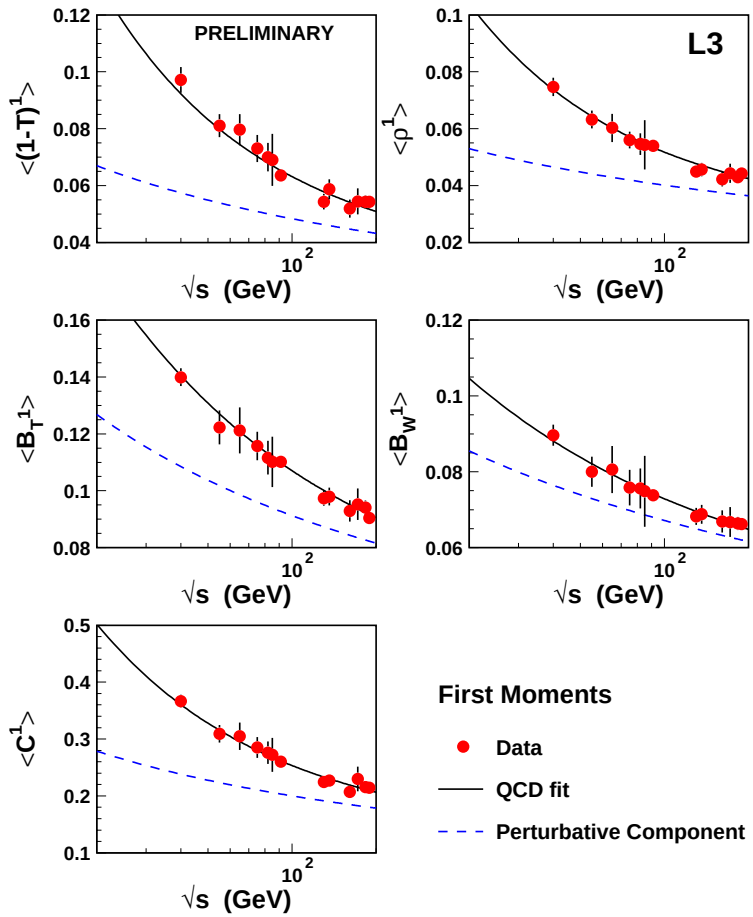
$$r_{\text{mult}} = 2.246 \pm 0.062(\text{stat.}) \pm 0.080(\text{sys.}) \pm 0.095(\text{theo.})$$

# Event shapes and power corrections

- New $e^+e^-$ event shape data at $\sqrt{s} \geq 189$ GeV confirm good agreement with models.



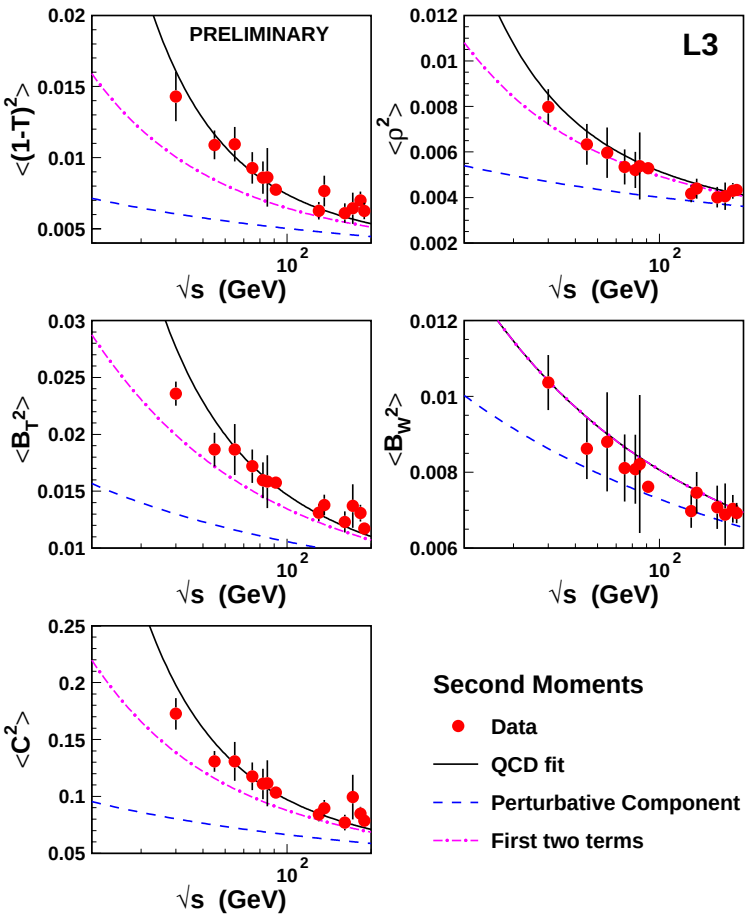
- Mean values of many event shapes show large power-behaved $(1/Q)$ discrepancies with NLO perturbation theory




Monte Carlo studies and models based on a universal low-scale form for $\alpha_S$ suggest that for some shape variables $(f = 1 - T, C, \rho)$, the dominant non-perturbative effect is a shift in the distribution by an amount $c_f \mathcal{P}$ where $\mathcal{P} \sim 1/Q$. This implies that

$$\begin{aligned} \langle f \rangle &= \langle f \rangle_{\text{pert}} + c_f \mathcal{P} + \mathcal{O}(1/Q^2) \\ \langle f^2 \rangle &= \langle f^2 \rangle_{\text{pert}} + c_f \langle f \rangle_{\text{pert}} \mathcal{P} + \mathcal{O}(1/Q^2) \end{aligned}$$

- The data on the second moment suggest that significant $1/Q^2$ terms are needed.

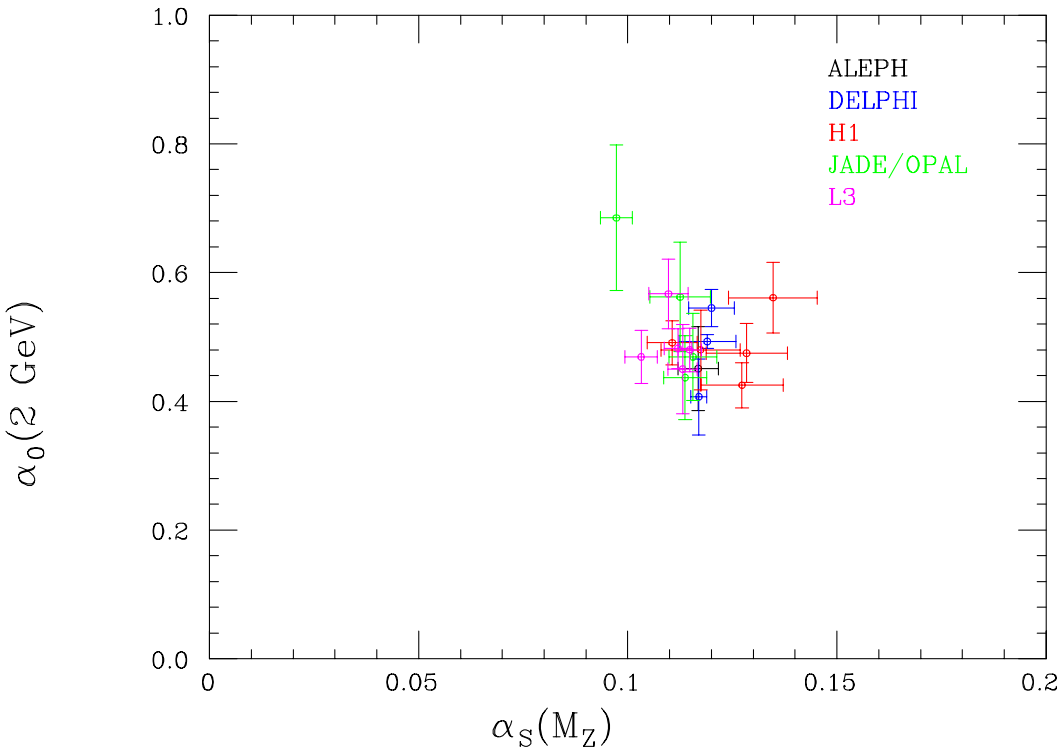



- According to the universal low-scale $\alpha_S$ model, the value of $\mathcal{P}$ can be used to measure the mean value of $\alpha_S$ below some infrared matching scale $\mu_I$,

$$\alpha_0(\mu_I) \equiv \frac{1}{\mu_I}\int_0^{\mu_I} d\mu\, \alpha_S(\mu^2)$$

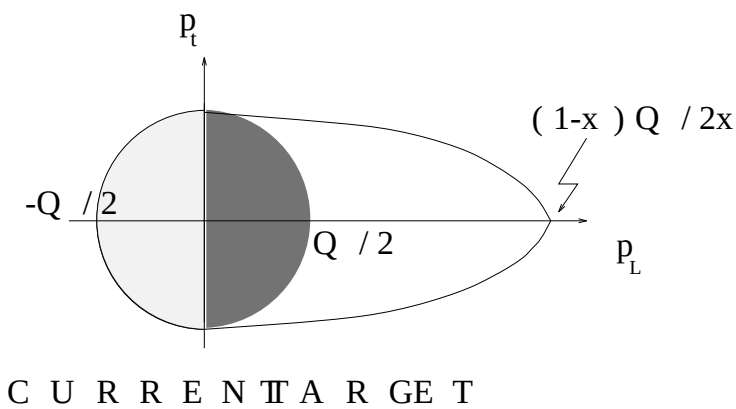
- Fits to HERA, LEP and lower-energy data suggest that

$$\alpha_0(\mu_I) \simeq 0.5 \qquad \text{for } \mu_I = 2 \text{ GeV}$$

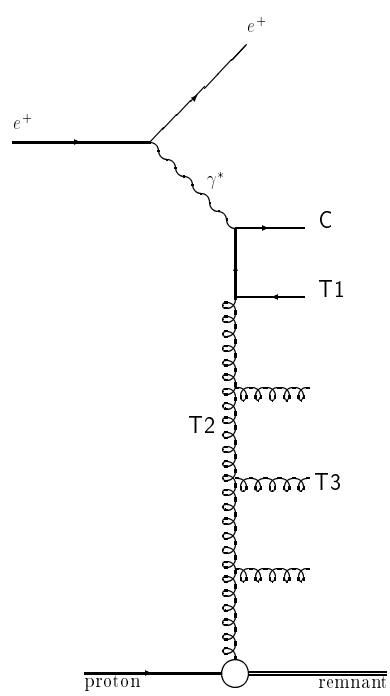


# Current and target fragmentation

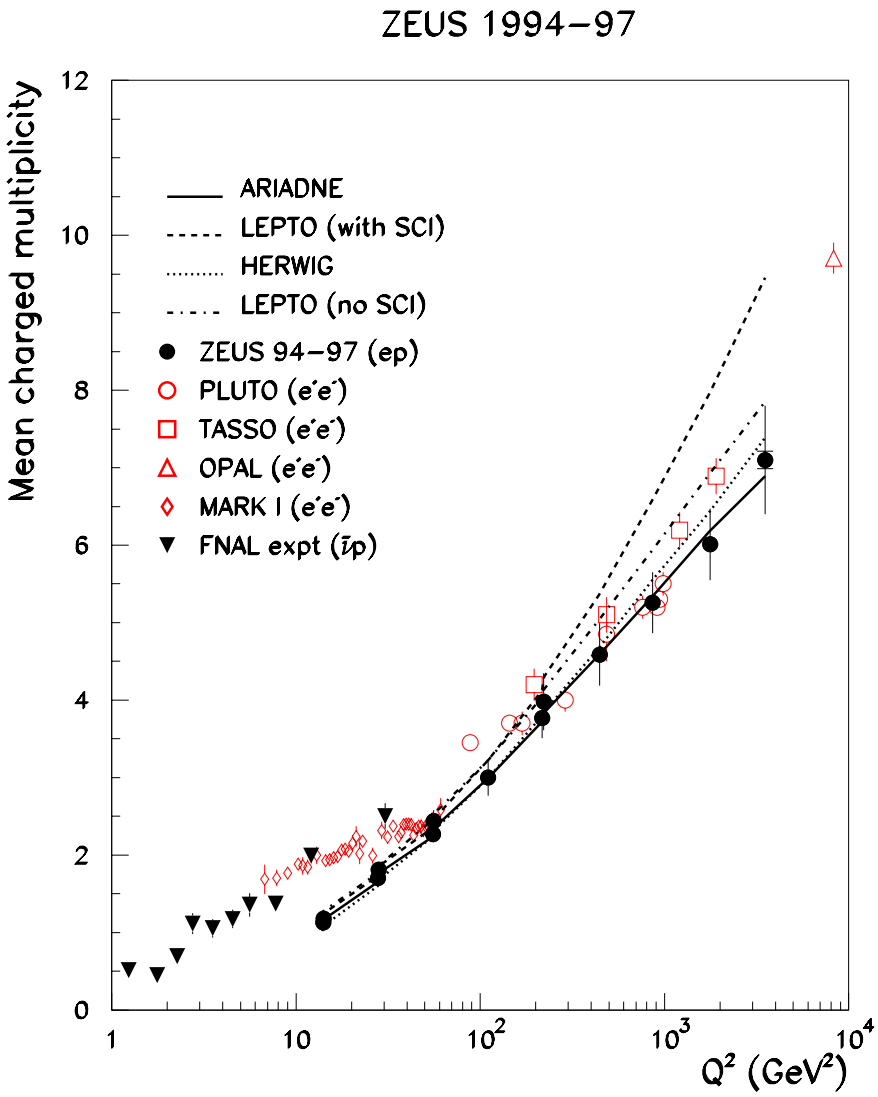
H1 and ZEUS study distributions of $x_p = 2|\boldsymbol{p}|/Q$ in the current and target hemispheres in the Breit frame.



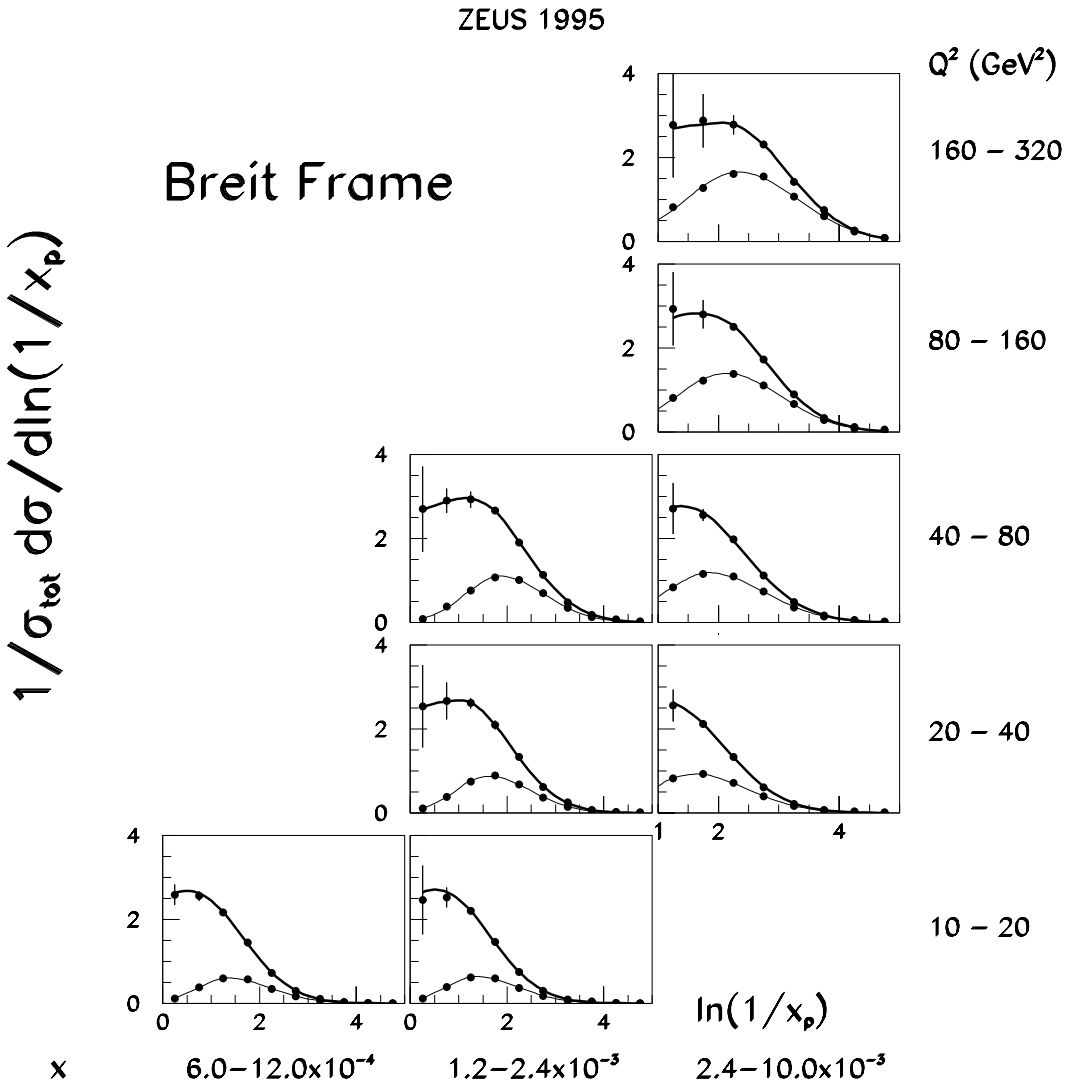

- In the current hemisphere one expects fragmentation of the current jet (C), similar to $e^+e^-$. In the target hemisphere, contribution T1 is similar to C, T2 gives extra particles with $x_p < 1$, while T3 gives $x_p \gtrsim 1$, generally outside acceptance.



- The charged multiplicity in the current hemisphere is indeed similar to $e^+e^-$. Differences at low $Q^2$ are consistent with the boson-gluon fusion contribution at low $Q^2$.

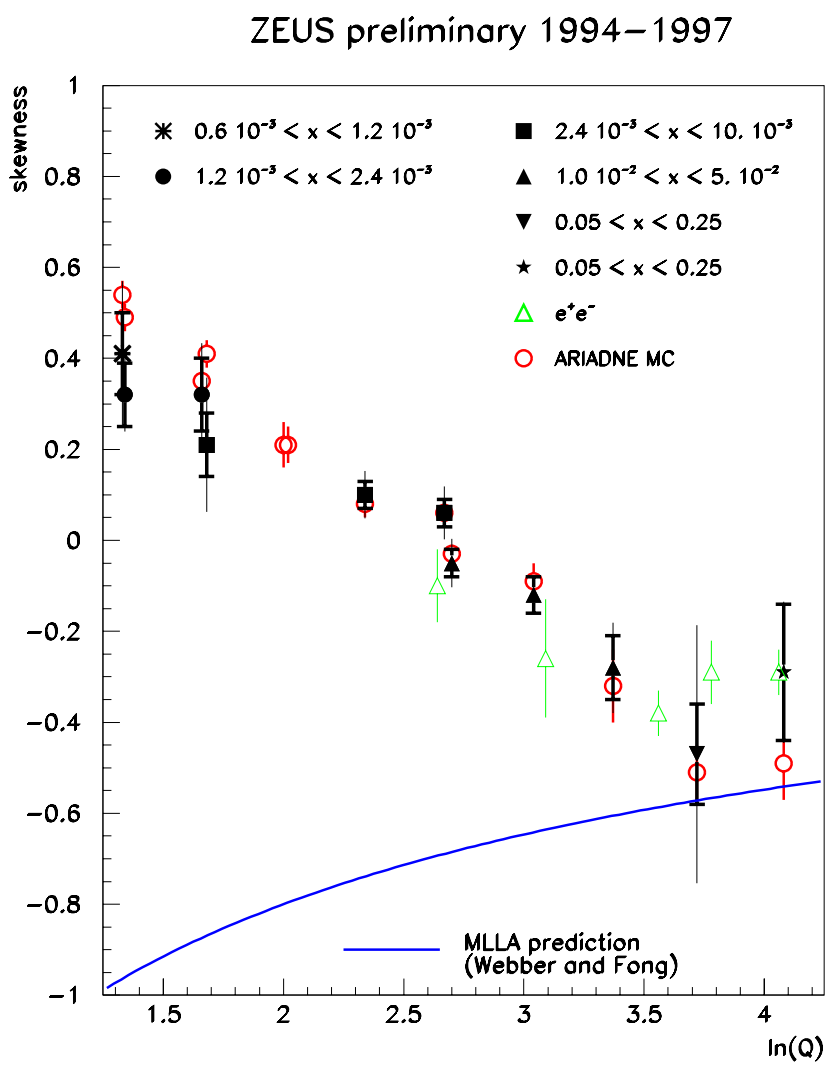


❖ The current-hemisphere distribution of $\xi = \ln(1/x_p)$ is also similar to $e^+e^-$, i.e. close to gaussian with little Bjorken $x$ dependence.

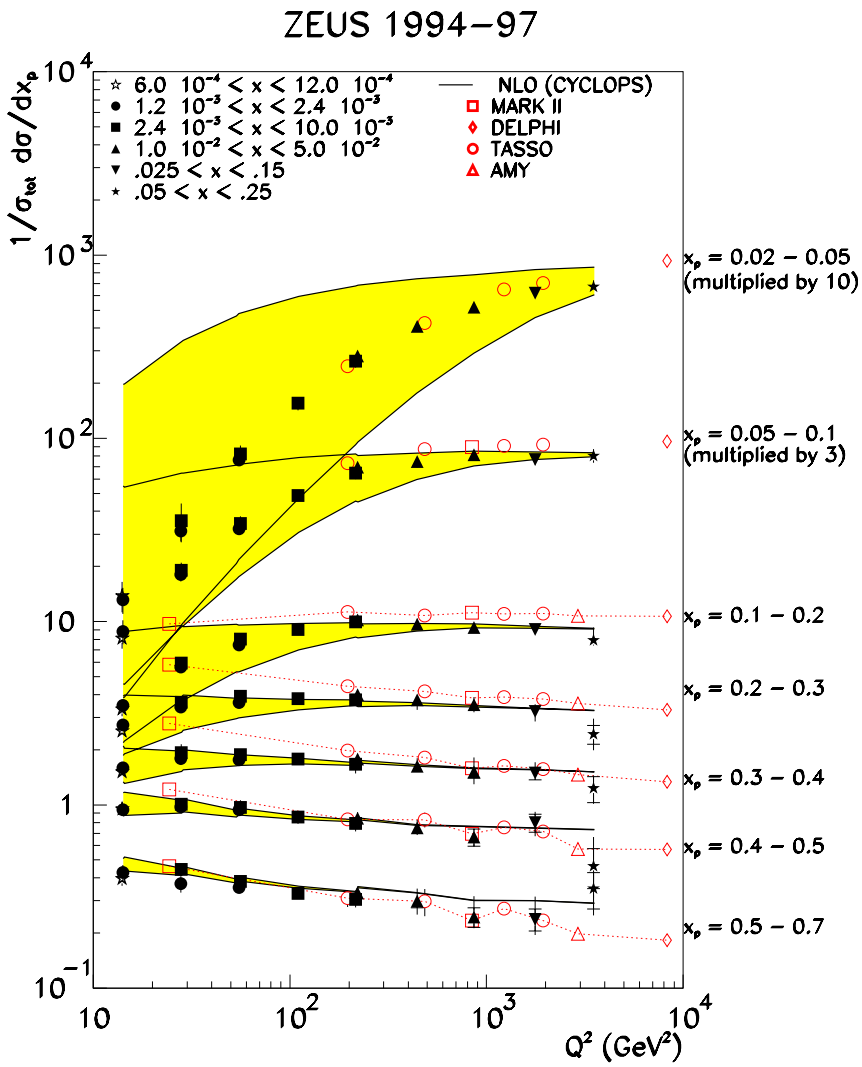


❖ At low $Q^2$ there is evidence of strong subleading corrections. The distribution is skewed towards higher values of $\xi$ (smaller $x_p$), contrary to NLLA predictions.

$$\text{Skewness} \equiv \left\langle (\xi - \bar{\xi})^3 \right\rangle / \left\langle (\xi - \bar{\xi})^2 \right\rangle^{\frac{3}{2}}$$



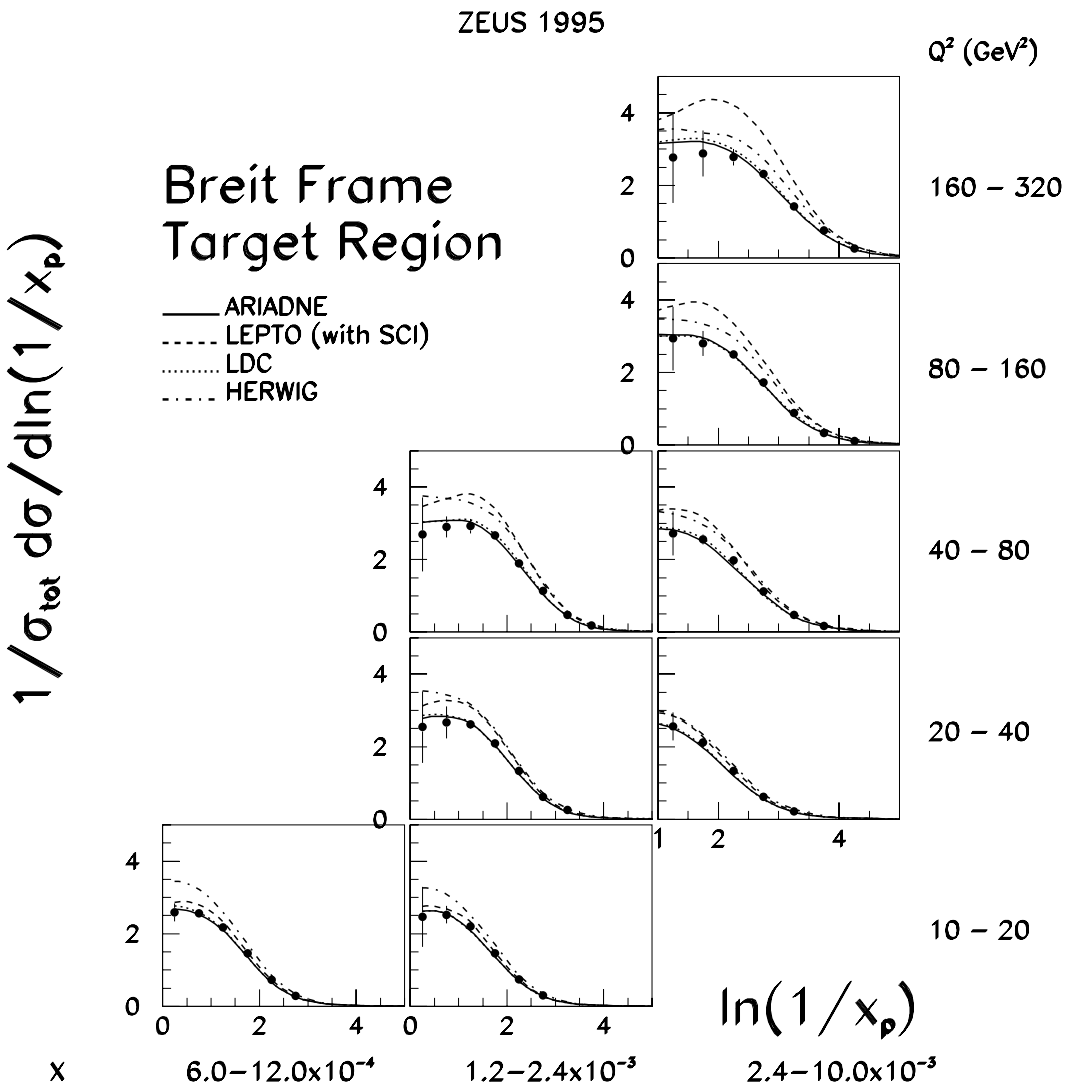
- On the other hand, the data lie well below the fixed-order perturbative prediction (CYCLOPS) at low $x_p$ and $Q^2$.




Discrepancies could be due to power-suppressed $(1/Q^2)$ corrections, of dynamical and/or kinematic origin. The bands correspond to a correction factor

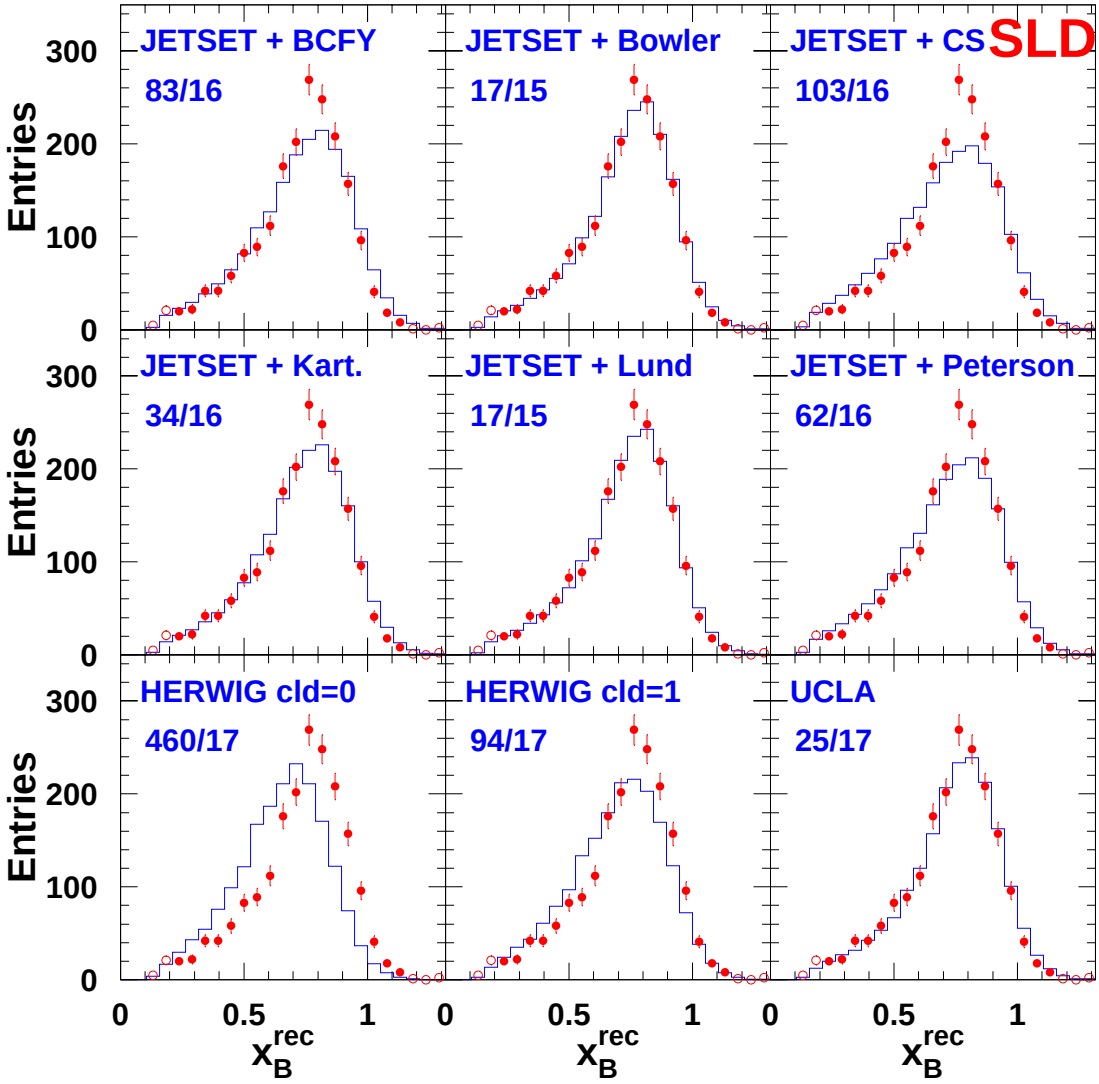
$$\left[1+\left(\frac{m_{\text{eff}}}{Qx_p}\right)^2\right]^{-1} \qquad (0.1 < m_{\text{eff}} < 1 \text{ GeV})$$

- In the target hemisphere there is also disagreement with MLLA, possibly due to the T3 contribution "leaking" into the region $x_p < 1$. If anything, models predict too much leakage. Little $Q^2$ dependence is observed.



# Heavy quark fragmentation

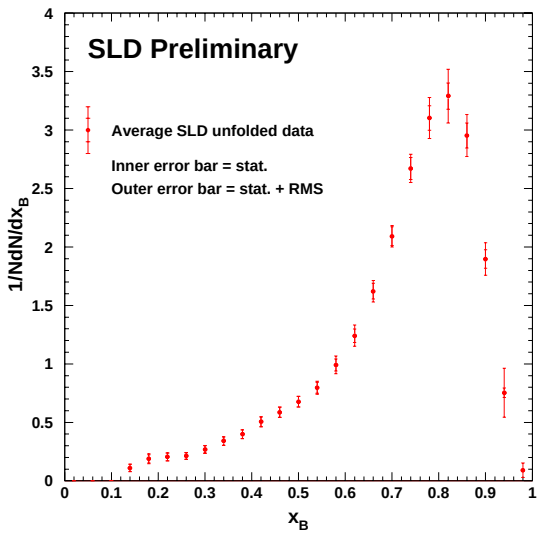
- New data on $b \to$ B fragmentation from SLD, using high-precision vertexing, discriminate between parton-shower plus hadronization models. [N.B. Uncorrected data.]

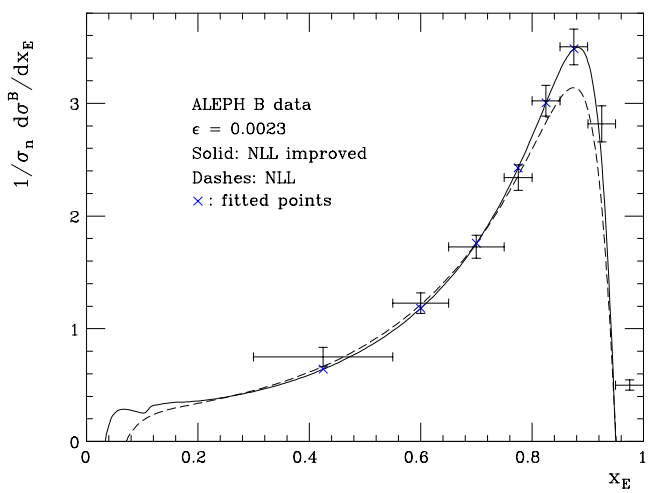


- Including more perturbative QCD leads to a reduction in the amount of non-perturbative smearing required.

Peterson function: $$\frac{1}{z}\left(1-\frac{1}{z}-\frac{\epsilon_b}{1-z}\right)^{-2}$$

- Pure Peterson: $\epsilon_b = 0.036$
  JETSET ($\simeq$ LLA QCD) + Peterson : $\epsilon_b = 0.006$
  NLLA QCD + Peterson (Nason & Oleari): $\epsilon_b = 0.002$

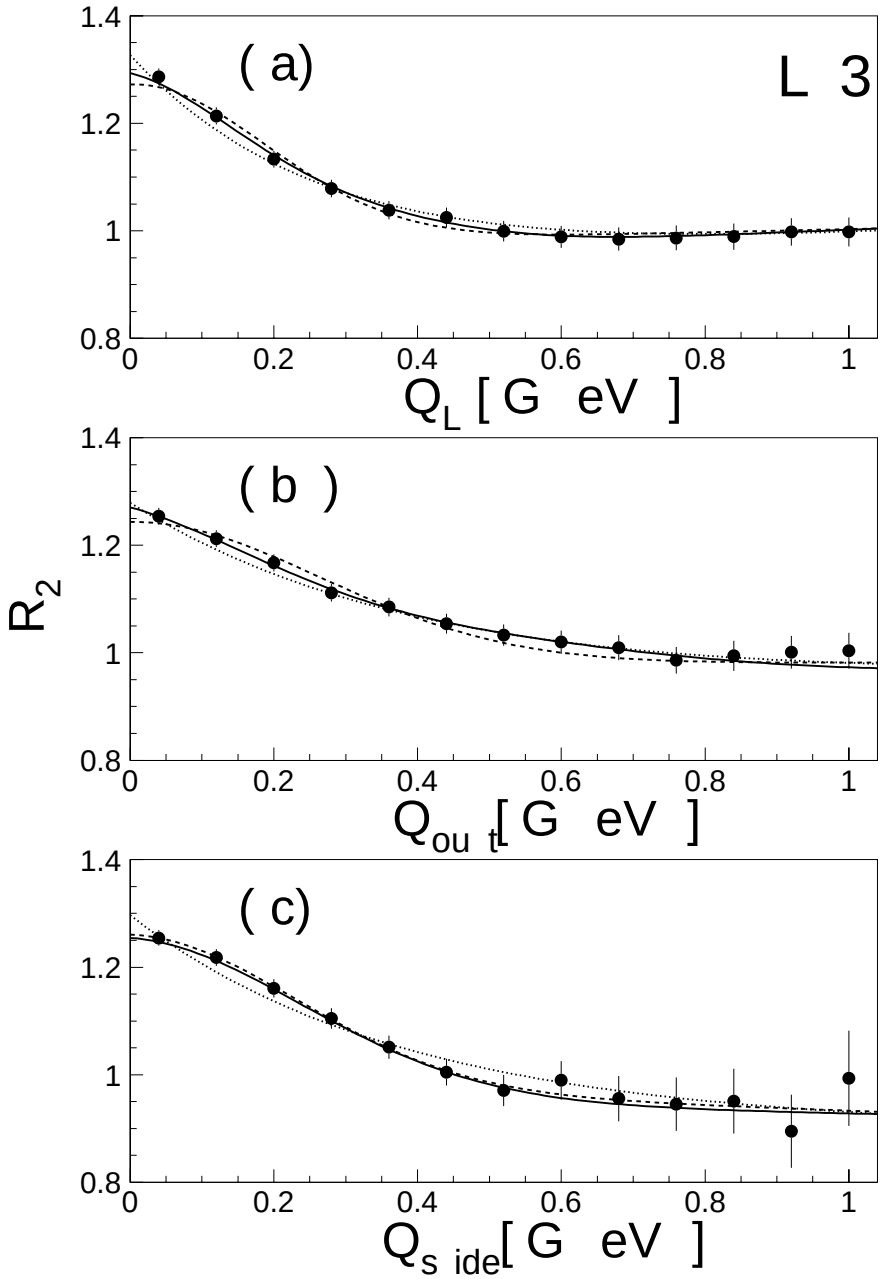



- In the universal low-scale $\alpha_S$ model, the perturbative prediction is extrapolated smoothly to the non-perturbative region, with no Peterson function (Dokshitzer, Khoze & Troyan)

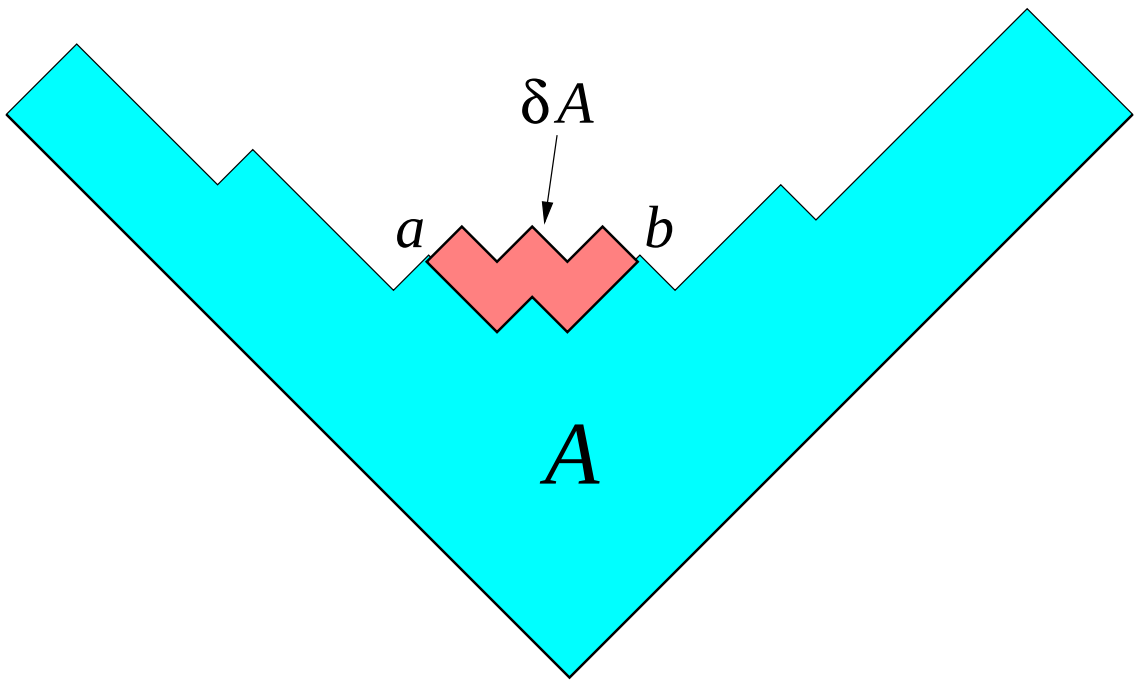
# Bose-Einstein correlations

- Definite evidence for elongation of the $\pi^{\pm}\pi^{\pm}$ source region along the thrust axis (DELPHI, L3, OPAL)

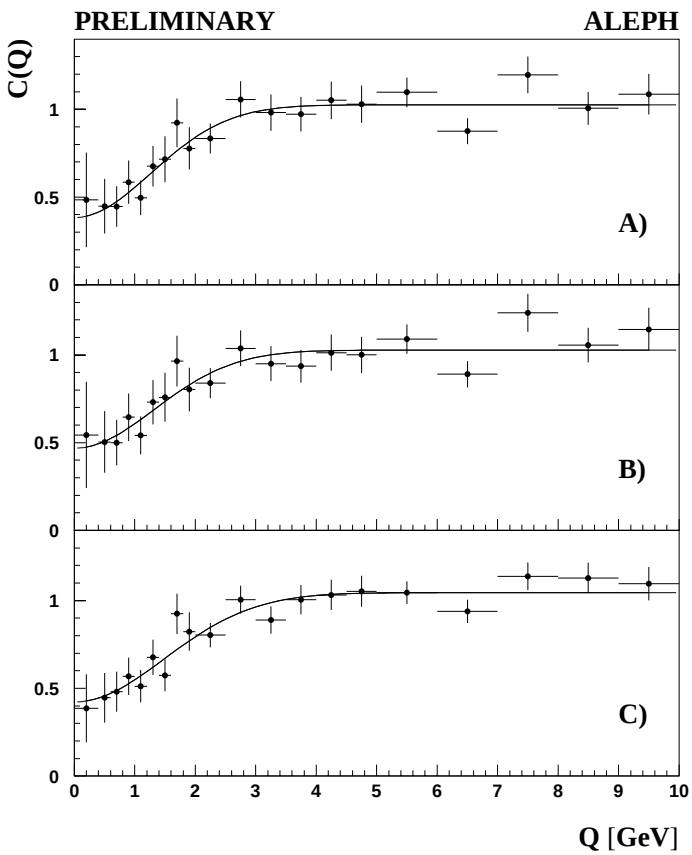


| Expt | $R_L$ (fm) | $R_T$ (fm) | $R_L/R_T$ |
|---|---|---|---|
| DELPHI | $0.85 \pm 0.04$ | $0.53 \pm 0.04$ | $1.61 \pm 0.10$ |
| L3 | $0.74 \pm 0.04$ | $0.56^{+0.03}_{-0.06}$ | $1.23 \pm 0.03^{+0.40}_{-0.13}$ |
| OPAL | $0.935 \pm 0.029$ | $0.720 \pm 0.045$ | $1.30 \pm 0.12$ |

This has a good explanation in the Lund string model (Andersson & Ringnér)



- ALEPH now has good evidence of Fermi-Dirac correlation in $\Lambda\Lambda$ (S=1). Plots A–C correspond to different comparison (no-correlation) samples.

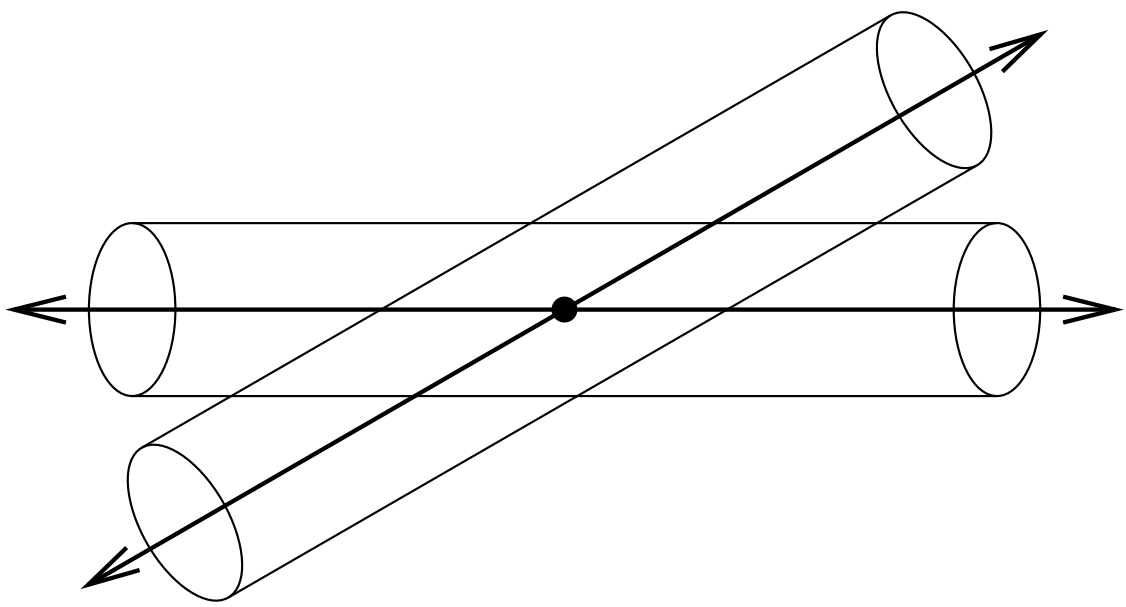



The source size appears to decrease with increasing particle mass:

| Particles | $R_{\text{source}}$ (fm) |
|---|---|
| $\pi\pi$ | $0.65 \pm 0.04 \pm 0.16$ |
| KK | $0.48 \pm 0.04 \pm 0.07$ |
| $\Lambda\Lambda$ | $0.11 \pm 0.02 \pm 0.01$ |

# WW fragmentation

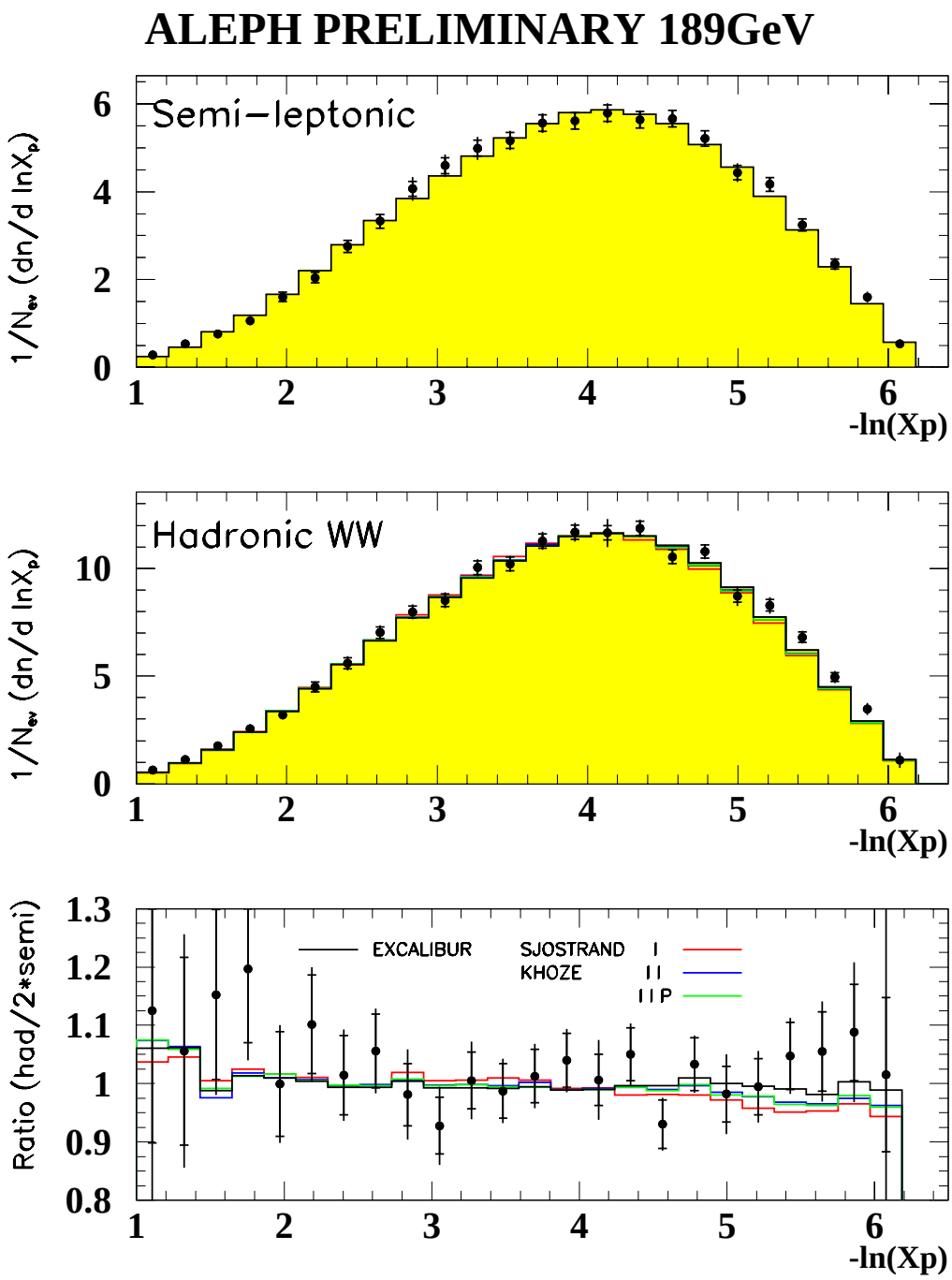
- In general we would expect correlations between W hadronic decays due to overlap of hadronization volumes. This occurs mainly in the central region, and is orientation-dependent.



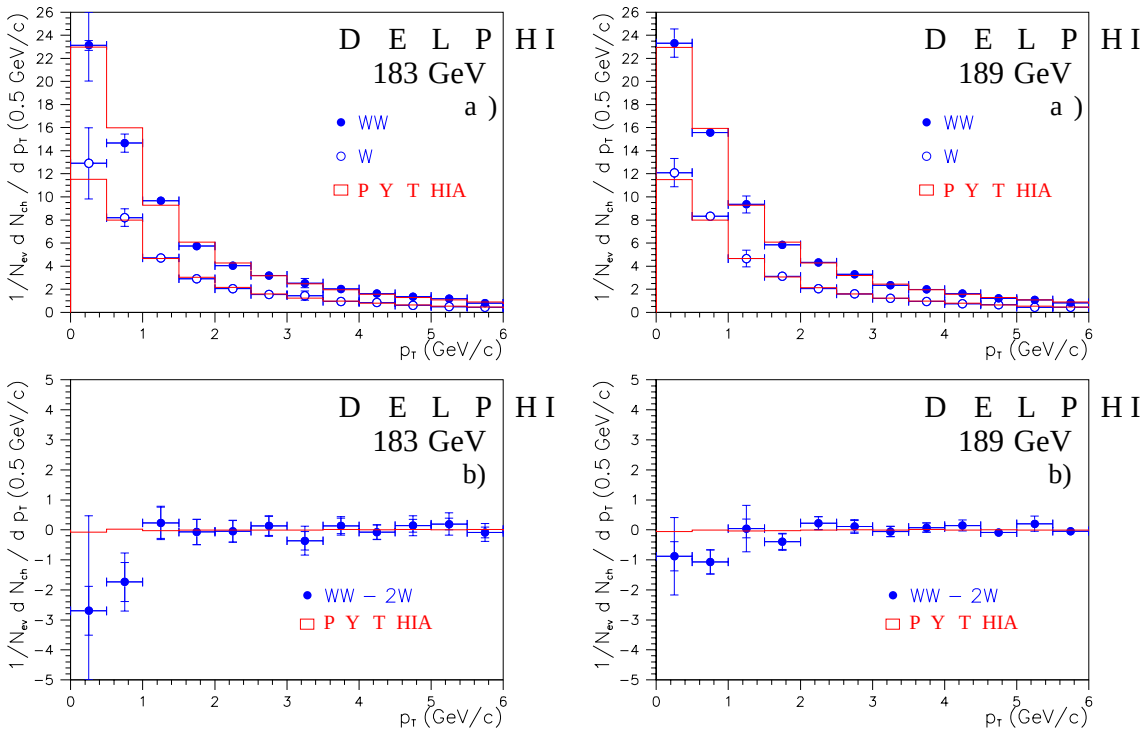
- Colour reconnection effects have been searched for in single-particle distributions.

- Bose-Einstein correlations between hadrons from different W's are also being looked for.

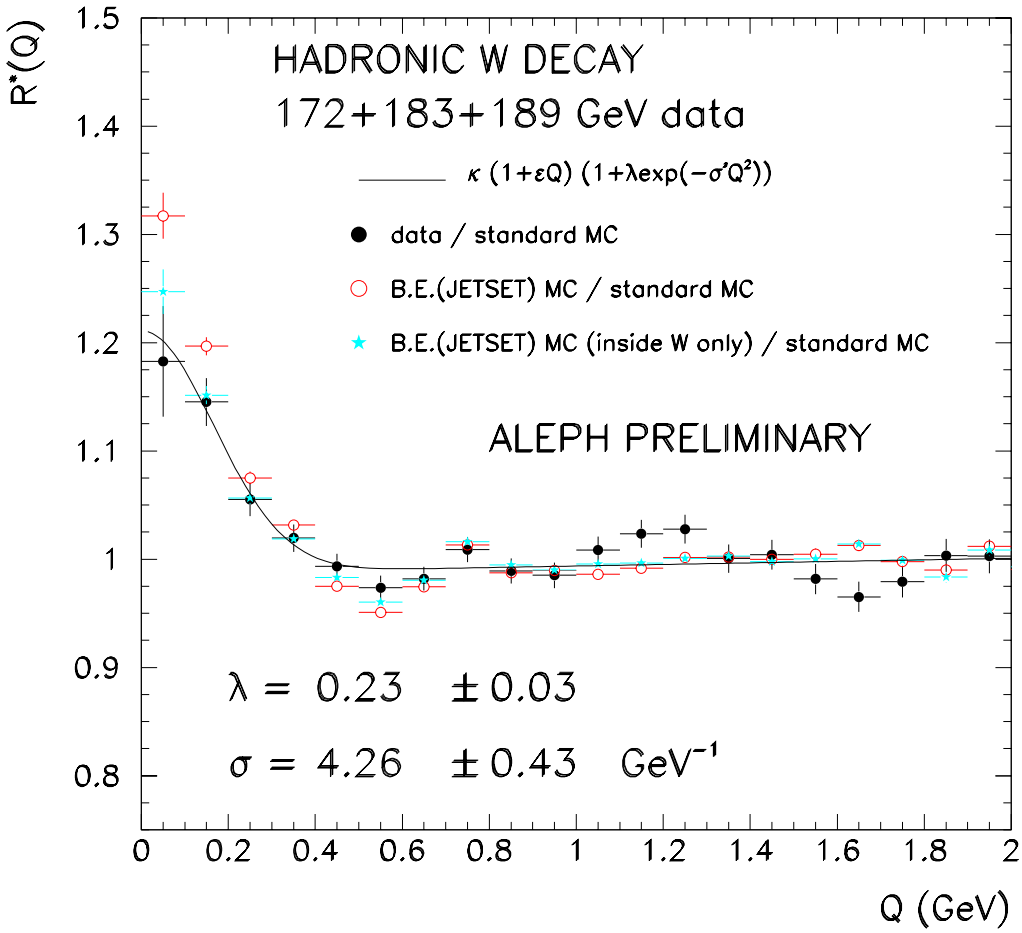
- No firm evidence yet for colour reconnection effects in charged particle spectra:



- DELPHI report a possible small ($\sim 2\sigma$?) effect in the distribution of $p_T$ (relative to the thrust axis):



- Bose-Einstein correlations between pions from different W's also seem to be suppressed:

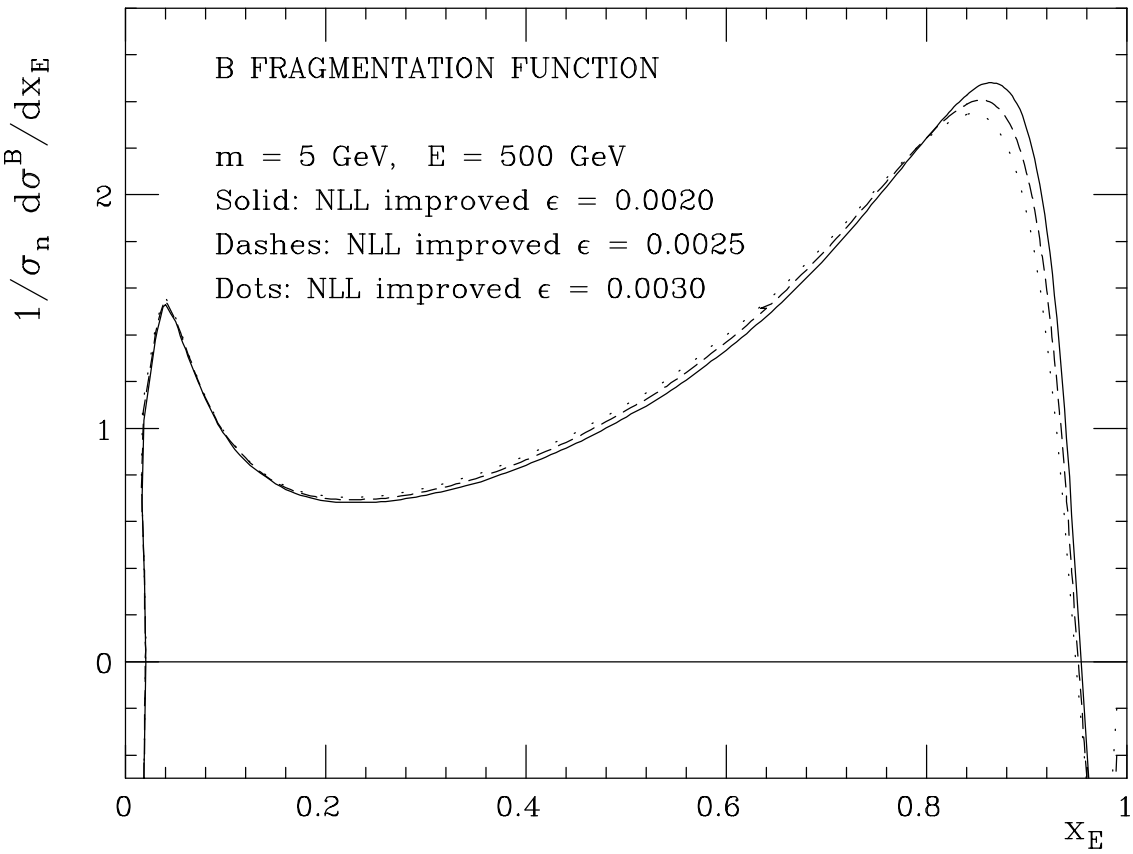


# Conclusions

- Detailed fragmentation studies need more theoretical input (coefficient functions).
- Hadronization studies suggest that mass is the dominant factor in heavy particle suppression, but baryon production is not well modelled.
- Gluon jets have the expected properties and can be used to measure ratio of colour factors $C_A/C_F$. However, no strong evidence of different particle content.
- $1/Q$ power corrections appear universal at the $\sim 20\%$ level. The corresponding low-scale $\alpha_s$ is $\sim 0.5$.
- Fragmentation in DIS shows disagreements with perturbative predictions: higher-order and/or non-perturbative?
- New $b$ quark fragmentation data test models and suggest perturbative effects dominate.
- Bose-Einstein (Fermi-Dirac) correlations show elongation of source, and source 'shrinkage' with increasing mass.
- WW fragmentation still shows no firm evidence for correlation between the decay products of the two W's.

# (Evolution of b fragmentation to NLC)

Fragmentation function for $b \to B$ evolved from $\sqrt{s} = M_Z$ to 500 GeV (Nason & Oleari)



# (Evolution of proton spectrum to UHE)

Decay of superheavy ($M \sim 10^{21}$ eV) metastable relic particles in the galactic halo has been suggested as a possible source of ultra-high energy cosmic rays.

Assuming no new physics, evolution of the proton spectrum from $\sqrt{s} = M_Z$ could give a sufficiently hard spectrum (N. Rubin)