August 1986
LU TP 86-16

Status of Fragmentation Models*

Torbjörn Sjöstrand
Department of Theoretical Physics,
University of Lund, Sölvegatan 14A,
S-223 62 Lund, Sweden

Abstract:
The main features of string fragmentation, cluster fragmentation and independent fragmentation are outlined. Some recent developments are summarized, both for the fragmentation models proper, for the connection to perturbative QCD and for comparisons with experimental data.

---

[*] minirapporteur talk in the parallel session on Jets and Fragmentation at the XXIII International Conference on High Energy Physics, July 16-23, 1986, Berkeley, California

Today QCD is generally accepted as the theory of strong interactions, but this does not mean that the properties of QCD are well understood. Indeed, whereas perturbation theory may be used to describe the hard, partonic processes, the subsequent transformation of the outgoing partons into stable hadrons is an unsolved, nonperturbative phenomenon. This vacuum has been filled by the emergence of several different fragmentation models, which are useful tools for experimentalists. Conveniently these models are separated into three main groups, string fragmentation (SF), cluster fragmentation (CF) and independent fragmentation (IF). SF is essentially synonymous with the efforts of the Lund group [1]. There are two current CF models, that of Webber [2] and that of Gottschalk and Morris [3]. IF models are by far the largest group, including Hoyer et al. [4], Ali et al. [5], ISAJET [6], COJETS [7], and so on. Some models, e.g. the fire-string one [8], do not fit into any of the three groups above.

In QCD, a linear confinement is expected at large distances. This provides the starting point for the string model: the simplest Lorentz covariant and causal implementation of a linear potential is provided by the massless relativistic string. The endpoints of the string are associated with quarks and antiquarks. The string pulled out between a q and a $\bar{q}$ may break by the production of additional $q\bar{q}$ pairs. The breaking may be understood as a tunneling phenomenon, automatically providing a suppression of heavy flavour production and a Gaussian transverse momentum spectrum. The q from one breakup will then combine with the $\bar{q}$ from an adjacent one to form the primary hadrons. The different breakup vertices are causally disconnected so that, although slow central particles are the first to be produced, the fragmentation can be described iteratively from the ends of the system inwards. A gluon may be represented as an energy and momentum carrying kink on the string. Since a gluon thus pulls out two string pieces where a quark only pulls out one, a difference of a factor two in asymptotic multiplicity is built in from the start.

The standard philosophy for cluster models has been to use perturbation theory, in the form of parton showers, to describe as much as possible of the expected properties of events. The subsequent fragmentation is assumed to be of a local nature. In the Webber program the shower evolution is therefore followed by a step in which gluons are split, nonperturbatively, into $q\bar{q}$ pairs. This subdivides the original system into a collection of $q\bar{q}$ colour singlet clusters. Each such cluster is assumed to decay isotropically in its rest frame, into a pair of hadrons. The relative probability for each possible

pair, consistent with the original flavours of the cluster, is given by phase space times spin counting factors. These hadrons may be resonances which decay further. A good description of the relative probability to produce different hadrons is obtained, also when compared with the recent HRS results on tensor meson production [9].

There are three known problems with the approach above, which all have been overcome. The first, realized from the onset, is that a cluster may occasionally have a very large mass, in which case allowing it to decay isotropically, and just into two particles, produces funny-looking events. Here the remedy has been to use a string-like scheme to break the cluster into two smaller clusters, each of which may break further if massive enough. Typically less than 10% of all clusters have to be split this way. The second problem was brought to the forefront by the HRS study of the fragmentation function close to $z=1$, Fig. 1 [10]. Since all clusters were supposed to decay into at least two particles, the probability of having a single particle carry almost the full energy of a jet was underestimated (whereas the Lund model, with particles produced one by one, describes the data). This discrepancy disappears if low-mass clusters below the two-particle threshold, constituting maybe 10% of the total number, are allowed to collapse to one single particle. Finally, data from TPC [11] indicates that baryon-antibaryon pairs are more aligned along the thrust axis than if baryons were just pair-produced in isotropic cluster decays, Fig. 2. This is understood in the Lund model, where the pair is pulled apart along the string direction. Agreement can also be obtained in cluster models if not only nonperturbative branchings $g \to q\bar{q}$ are allowed, but also a $\approx 5\%$ admixture of branchings into diquark-antidiquark pairs, so that clusters may have net baryon number.

It should be emphasized that the issue of cluster versus string is not a simple either-or question. Rather, it spans a multidimensional space, with "pure string" and "pure cluster" at opposite corners. Allowing large clusters to decay into smaller ones along the string direction, or having small clusters collapse into a single particle, makes cluster models more "stringy". An intermediate example is provided by the new CALTECH-II cluster model of Gottschalk and Morris [3], in which the nonperturbative $g \to q\bar{q}$ branchings are rejected, and a string is spanned continuously from a q to a $\bar{q}$ via a number of gluons, just like in the Lund model. This string then breaks into variable-mass clusters (somewhat like in the Artru-Mennessier model [12]), however, where the Lund string breaks into fixed-mass particles. Further, the flavour composition in cluster decays is in CALTECH-II given by parametrizations of low-energy data, rather than by phase space like in the Webber model. A remaining dividing line is that both cluster models use the phase space of the cluster decays to generate transverse momenta, where the string model invokes the tunneling mechanism.

While not a part of fragmentation models proper, perturbative QCD is needed to provide the parton configuration that is to be fragmented. Two main possibilities exist for obtaining this information, matrix elements and parton showers. In principle, matrix elements is the way to go, since we believe that one will come arbitrarily close to the "exact" answer if matrix elements are only calculated to sufficiently high orders in perturbation theory. In contrast, parton showers are based on leading log approximations to the full answer, and are therefore less accurate in describing the production of well separated jets.

There are a number of factors which today seem to make parton showers better than matrix elements for Monte Carlo applications, however. Calculations to higher orders in α_s are increasingly difficult. In e^+e^- annihilation one has come to full second order [13], i.e. including the Born term for $q\bar{q}gg$ and $q\bar{q}q'\bar{q}'$ production and the one-loop corrections to $q\bar{q}g$ production. Considering the complexity of that calculation, availability of a full $O(\alpha_s^3)$ answer does not seem imminent. On the other hand, data at PETRA/PEP energies already indicate that the $O(\alpha_s^2)$ matrix elements are less than adequate. A first indication of this was obtained in the JADE analysis [14] of the energy-energy correlation asymmetry [15]. If one desires to describe the asymmetry at small angles, matrix element cutoffs have to be chosen so small that essentially no two-parton events remain [16], at which point higher order corrections can no longer be neglected. At this conference, JADE showed that the number of four-jet events in data is roughly a factor of two larger than predicted by using second order matrix elements [17]. (Such large higher order corrections may well be expected based on the choice of Q^2 scale for the perturbative expansion [16]). Finally, the Mark II comparison of quark and gluon jet fragmentation [18] shows that the q/g difference is much larger than obtained with matrix elements + Lund string, whereas using parton shower + cluster or string gives good agreement.

Turning to parton showers, these have become more sophisticated over the last few years. In particular, coherence effects [19] have been included, thus ensuring a more reliable description of the internal structure of jets, and providing more confidence for extrapolations to higher energies. The

pioneering model in this field is the one by Marchesini and Webber [20]; today similar (but different in details) programs are also available from Gottschalk [3] and, within the Lund Monte Carlo, from Bengtsson and Sjöstrand [21]. Recently, emphasis has been put on also getting the wide-angle emission of gluons in $q\bar{q}g$ events "right", thus (hopefully) providing a good overall description of event structures. One interesting oservation is that the use of a low cutoff mass in the parton shower evolution, typically $Q_0 \approx 1 - 1.5$ GeV, seems to be mandatory for a good description of data [21]. Such a low value has to be used anyhow for the Webber cluster fragmentation scheme to work well, but is in no sense obvious when shower evolution is combined with the Lund string. Like in the matrix element case, the punch line seems to be that, at PETRA/PEP energies, very few events remain that are "pure two-jets", i.e. that do not contain any extra gluons at all.

Recently, A. Petersen made a comparison between Mark II data and the very latest versions of the Webber, CALTECH-II and Lund Monte Carlos, all now using parton showers plus CF (first two) or SF (last one) [22]. Some of the resulting distributions are shown in Figs. 3 - 5. Generally speaking, the agreement between data and models is good, at least if viewed by the eye rather than the more unforgiving χ^2 measure. Most problems are noted with the CALTECH-II program, which not only is the youngest and therefore least well tuned, but also has a higher parton shower cutoff than the two others. The larger fraction of pure two-jets is reflected e.g. in the thrust distribution close to T=1. The largest difference between the models is found in the minor (the third member of the thrust family) distribution, where the Lund model agrees with the data, but both cluster models deviate significantly in the low-minor region. Unfortunately this discrepancy is not easily interpreted, so it is unknown whether this points to any inherent problem with cluster models or just to the need of some fine-tuning.

So far we have not mentioned the third major class of fragmentation models, that of independent fragmentation. In this approach it is assumed that the fragmentation of each parton can be described independently of all other partons in the system. The actual fragmentation can, like in the Lund string model, be described iteratively from the ends of the system, normally with particles being produced directly without any intermediate clusters. This framework is extremely easy to work with, and usually provides a good description of events, but it does suffer from a number of very serious deficiencies, like lack of Lorentz covariance, and nonconservation of energy, momentum and flavour. The difference between SF and IF is most easily observed in $q\bar{q}g$ events, where the region between the q and $\bar{q}$ is depleted from particles in the string case [23], since the string is stretched from the q via the g to the $\bar{q}$. This "string effect" was first observed by JADE in 1980 [24], and has since been confirmed and expanded on in a number of publications [25]. Also the Webber and CALTECH-II programs with coherent parton showers and cluster fragmentation can explain the effect, whereas the previous non-coherent cluster program of Gottschalk could not. In fact, the studies of the Leningrad group [26] suggest that an approximate duality may exist between a perturbative approach with multiple gluon emission, in which coherence effects are properly taken into account, and a nonperturbative description in terms of strings. New data on this kind of issues have here been presented by TPC and Mark II [27], which compare the energy flow in $q\bar{q}\gamma$ and $q\bar{q}g$ events, with results that can be well described either by strings or by perturbation theory.

Finally, a note on comparisons between quark and gluon jets. Naive asymptotic expectations are for a factor 9/4 higher multiplicity in gluon than in quark jets, just from the higher colour charge of gluons. There are a number of well understood reasons why real differences are much smaller than this, including higher order perturbative QCD corrections [28], contamination by charm and bottom jets in the quark jet sample, that gluon jets are broader in angle and therefore more particles are lost out of a jet cone around a gluon than around a quark [29], and that nonvanishing hadron masses at present energies provide a real enough cutoff to probing the low-z end of the fragmentation spectrum. Such problems notwithstanding, sensible differences have by now been observed by JADE [29], Mark II [18], HRS [30] and, contrary to original claims, UA1 [31], see e.g. Fig. 6.

To summarize, no major breakthroughs have taken place in the study or understanding of fragmentation, but small steps are made all the time. Certainly we today have string and cluster models that provide quite decent descriptions of the data, and that can be used with some confidence to predict TRISTAN/SLC/LEP results. Over the last few years, we have seen some convergence between the CF and SF classes of models, driven by experimental data. Cluster models have come to include special cases in which string fragmentation ideas are used. On the other hand, the Lund string model, traditionally associated with matrix elements and thus fairly high cutoff scales, is now more used together with a parton shower approach, with a low cutoff scale for the perturbative evolution. This does not mean that all differences have disappeared, but only that the task of distinguishing the

present generation of models will require even more sophisticated experimental analysis.

One should also note that this minireview, for reasons of space, has been dealing essentially exclusively with the field of e^+e^- annihilation, which is the place where fragmentation issues are easiest to study. Based on a general belief in jet universality, we expect that experiences from e^+e^- can be applied to deep inelastic leptoproduction and high-p_T jets in hadron physics. This leaves open the issue of providing a complete description for hadronic events, also including low-p_T beam jets. Here much work has been carried out in recent years under the aegis of SSC studies; some useful papers with further references may be found in [32,33]. Still, it is abundantly clear that this field is in its infancy, and that much work remains to be done before it reaches the relative maturity of e^+e^- physics.

Acknowledgements

The author wishes to thank M. Bengtsson, S. Bethke, T. Gottschalk, A. Petersen and B. Webber for useful discussions on the subjects covered in this talk.

References

1. B. Andersson, G. Gustafson, G. Ingelman, T. Sjöstrand, Phys. Rep. 97 (1983) 33

   T. Sjöstrand, Computer Phys. Comm. 39 (1986) 347
2. B. R. Webber, Nucl. Phys. B238 (1984) 492
3. T. D. Gottschalk, in [33], p. 122

   T. D. Gottschalk, D. Morris, in preparation
4. P. Hoyer, P. Osland, H. G. Sander, T. F. Walsh, P. M. Zerwas, Nucl. Phys. B161 (1979) 349
5. A. Ali, E. Pietarinen, G. Kramer, J. Willrodt, Phys. Lett. 93B (1980) 155
6. F. E. Paige, S. D. Protopopescu, in [32], p. 41
7. R. Odorico, Nucl. Phys. B228 (1983) 381
8. L. Angelini, L. Nitti, M. Pellicoro, G. Preparata, G. Valenti, Riv. Nuovo Cimento 6 (1983) 1
9. HRS Collaboration, S. Abachi et al., contributed paper to this conference
10. HRS Collaboration, M. Derrick et al., Phys. Lett. 164B (1985) 199
11. TPC Collaboration, H. Aihara et al., Phys. Rev. Lett. 55 (1985) 1047
12. X. Artru, G. Mennessier, Nucl. Phys. B70 (1974) 93
13. R. K. Ellis, D. A. Ross, A. E. Terrano, Nucl. Phys. B178 (1981) 421
14. JADE Collaboration, W. Bartel et al., Z. Physik C25 (1984) 231
15. C. Basham, L. Brown, S. Ellis, S. Love, Phys. Rev. Lett. 41 (1978) 1585

16. T. Sjöstrand, Z. Physik C26 (1984) 93
17. JADE Collaboration, W. Bartel et al., DESY 86-086
18. Mark II Collaboration, A. Petersen et al., Phys. Rev. Lett. 55 (1985) 1954
19. A. H. Mueller, Phys. Lett. 104B (1981) 161
B. I. Ermolaev, V. S. Fadin, JETP Lett. 33 (1981) 269
A. Bassetto, M. Ciafaloni, G. Marchesini, Phys. Rep. 100 (1983) 201
20. G. Marchesini, B. R. Webber, Nucl. Phys. B238 (1984) 1
21. M. Bengtsson, T. Sjöstrand, in preparation
22. A. Petersen (Mark II Collaboration), private communication
23. B. Andersson, G. Gustafson, T. Sjöstrand, Phys. Lett. 94B (1980) 211
24. A. Petersen, in Elementary Constituents and Hadronic Structure, ed. J. Tran Thanh Van (Editions Frontieres, Dreux, 1980), p. 505
25. JADE Collaboration, W. Bartel et al., Phys. Lett. 101B (1981) 129, ibid. 134B (1984) 275, ibid. 157B (1985) 340
TPC Collaboration, H. Aihara et al., Z. Physik C28 (1985) 31
TASSO Collaboration, M. Althoff et al., Z. Physik C29 (1985) 29
26. Ya. I. Azimov, Yu. L. Dokshitzer, V. A. Khoze, S. I. Troyan, Phys. Lett. B165 (1985) 147
27. TPC Collaboration, H. Aihara et al., contributed paper to this conference
Mark II Collaboration, P. D. Sheldon et al., SLAC-Pub-4030 (1986)
28. J. B. Gaffney, A. H. Mueller, Nucl. Phys. B250 (1985) 109
29. JADE Collaboration, W. Bartel et al., Z. Physik C21 (1983) 37
30. HRS Collaboration, M. Derrick et al., contributed paper to this conference
31. UA1 Collaboration, G. Arnison et al., CERN-EP/86-55
32. Supercollider Physics, ed. D. E. Soper (World Scientific, Singapore, 1986)
33. Observable Standard Model Physics: Monte Carlo Simulation and Detector Capabilities, eds. H.-U. Bengtsson, C. Buchanan, T. Gottschalk, A. Soni (World Scientific, Singapore, 1986)

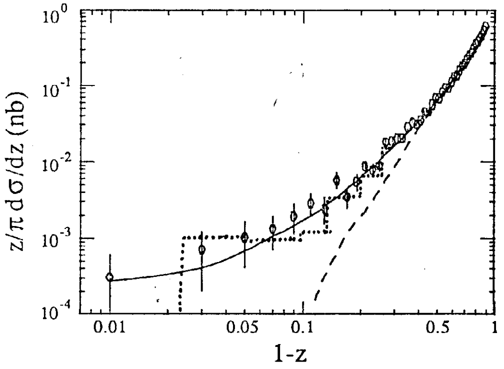

Fig. 1: Charged particle spectrum, HRS data points [10] compared with Lund model (full line), old Webber model (dashed line) and Webber model with small clusters collapsing to one particle (dotted histogram).

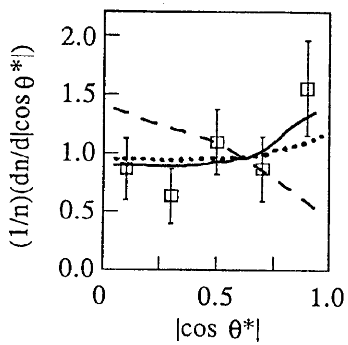

Fig. 2: Relative angular distribution of $p\bar{p}$ pairs, TPC data points [11] compared with Lund model (full), old Webber model (dashed) and Webber model with 5% $g \to qq + \bar{q}\bar{q}$ (dotted).

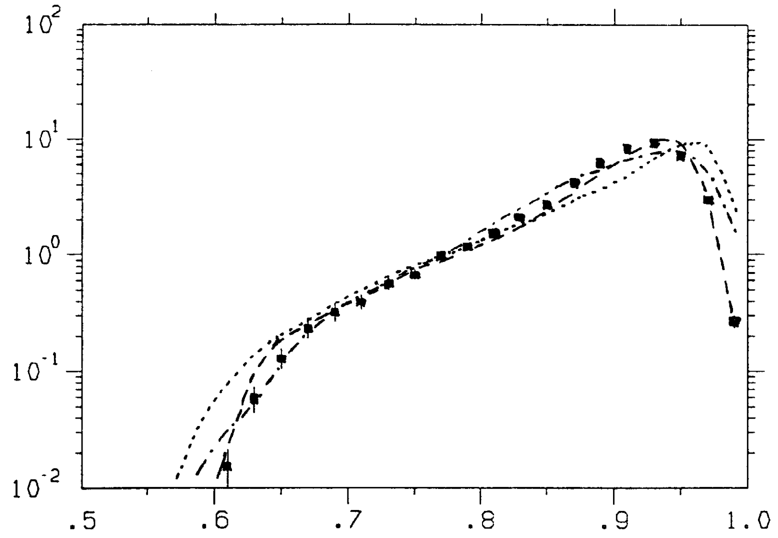

Fig. 3: Thrust distribution, Mark II data points [22] compared with Lund model (dashed), Webber model (dash-dotted) and CALTECH-II (dotted).

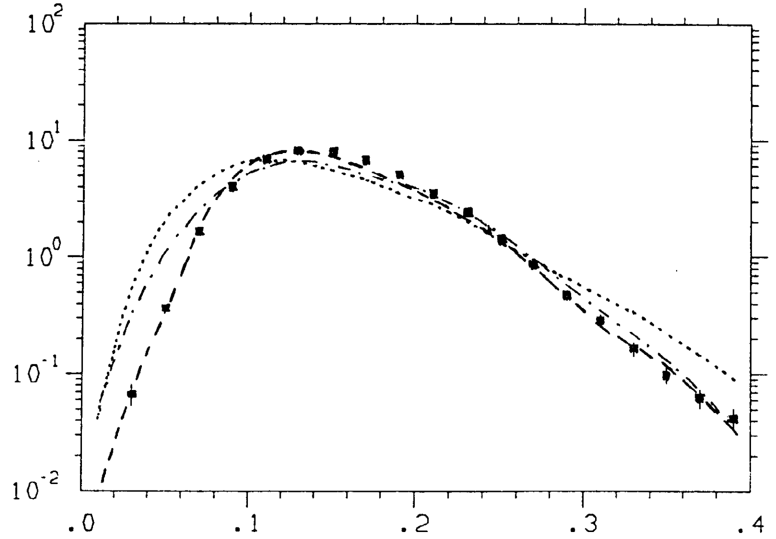

Fig. 4: Minor distribution, notation as Fig. 3.

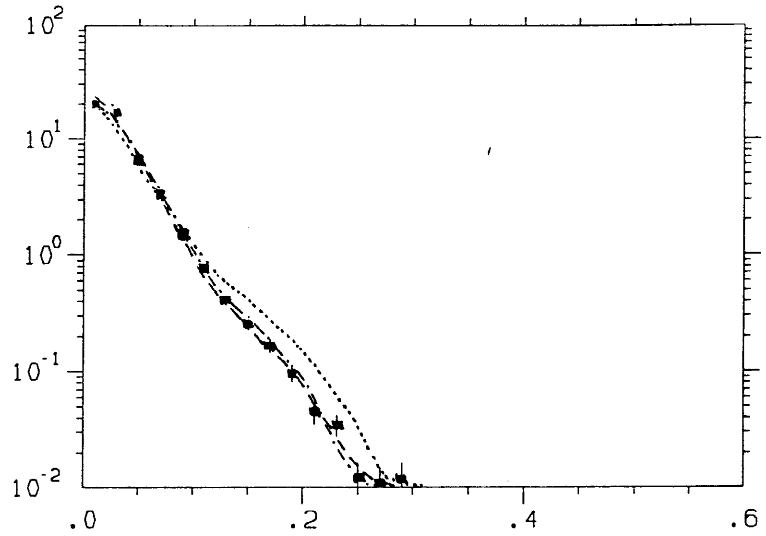

Fig. 5: Aplanarity distribution, notation as Fig. 3.

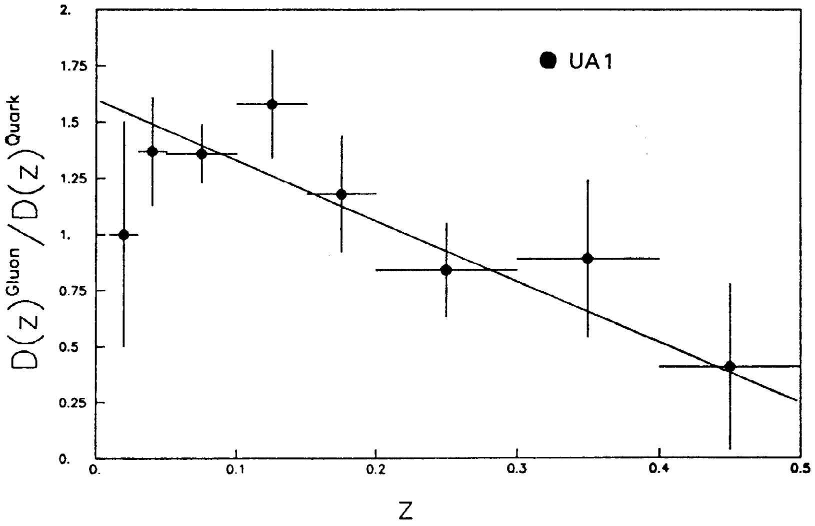

Fig. 6: Ratio of gluon to quark fragmentation functions, UA1 data points [31] compared with shower evolution + Lund strings (full line).