RAL-TR-2000-039

LC-TH-2000-047
August 2000

The implementation of the four-jet matrix elements in HERWIG and elsewhere[1]

S. Moretti
Rutherford Appleton Laboratory, Chilton, Didcot, Oxon OX11 0QX, UK.

Abstract

The new version [1] of the HERWIG event generator [2] allows one to simulate hadronic final states in electron-positron annihilations using the $\mathcal{O}(\alpha_s^2)$ partonic matrix elements (MEs) for $e^+e^- \to q\bar{q}gg$ and $e^+e^- \to q\bar{q}Q\bar{Q}$. The consequent phenomenology is discussed in the case of four-jet signatures of particular relevance to future electron-positron Linear Colliders (LCs), such as in Higgs boson searches. The new approach adopted in HERWIG is compared to the standard $\mathcal{O}(\alpha_s)$ description, as well as to other recently presented event generators also based on $e^+e^- \to 4$ parton MEs.

Four-jet events will play a crucial role at future LCs [3]. Just one example is sufficient to make this point. Such machines are an ideal laboratory where to pursue studies in the Higgs sector [4]. In this respect, it should be recalled that $e^+e^- \to 4$ jet processes from QCD will represent a serious noise in the search for Higgs particles, both in the standard electroweak (EW) theory and in possible extensions. In particular, in the Standard Model (SM), the dominant Higgs production channel proceeds via the 'bremsstrahlung' process $e^+e^- \to HZ$ [5], followed by the hadronic decay $H \to b\bar{b}$ (in the intermediate M_H regime), with the Z boson going into pairs of jets about 70% of the times.

Thus, it is very important that four-jet events from QCD are correctly implemented in the Monte Carlo (MC) programs that are widely used in phenomenological studies of hadron production at e^+e^- colliders. In this connection, it has always been rather worrying that certain aspects of four-jet production were never well described by the standard '$\mathcal{O}(\alpha_s)$ ME + parton shower (PS)' MC programs. As an illustration, one may recall the four-jet studies performed by the ALEPH Collaboration at LEP1 [6], which had revealed a significant disagreement between data and MCs in the typical four-jet angular variables (see Ref. [7] for the definition and the discussion of some of their properties): (i) the Bengtsson-Zerwas angle $\chi_{\rm BZ}$; (ii) the Körner-Schierholz-Willrodt angle $\Phi_{\rm KSW}$; (iii) the (modified) Nachtmann-Reiter angle $\theta^*_{\rm NR}$; (iv) the angle between the two least energetic jets θ_{34}.

[1]Talk given at the 2nd ECFA/DESY Study on Physics and Detectors for a Linear Electron-Positron Collider, Padova, Italy, 4-8 May 2000.

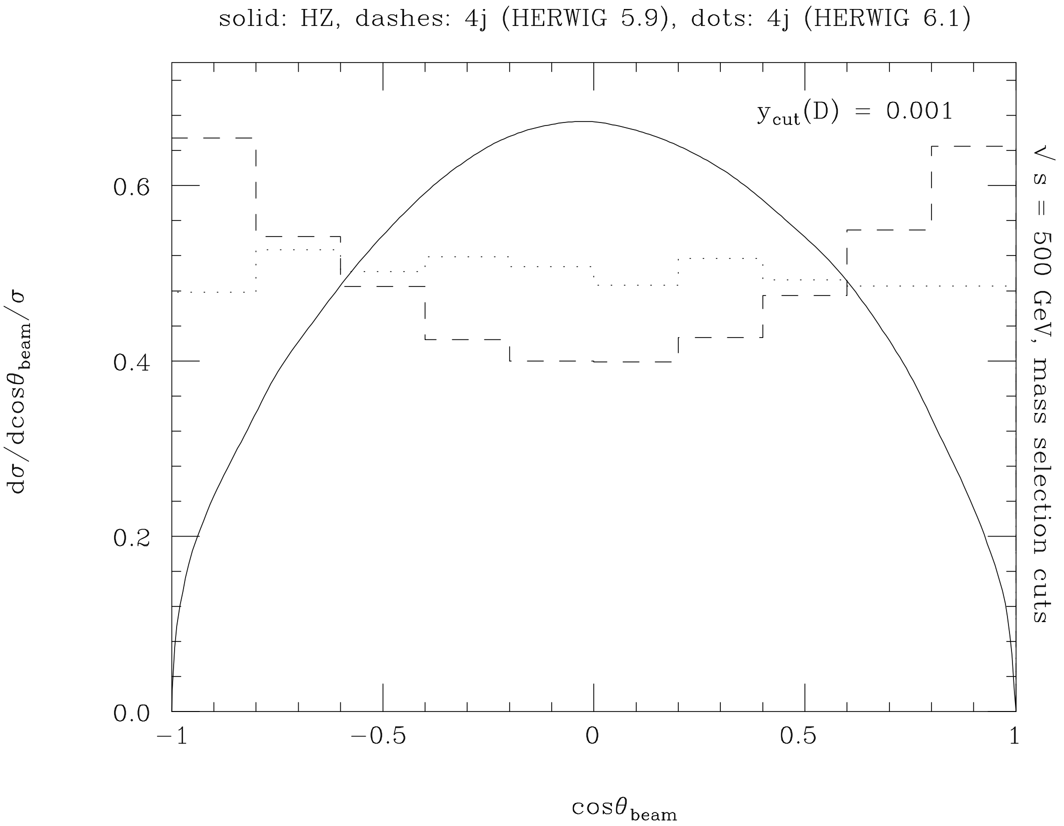


Figure 1: Differential distributions for the polar angle of the reconstructed Z.

In Ref. [8], it was argued that the discrepancy was nothing but the evidence that the standard '$\mathcal{O}(\alpha_s)$ ME + PS' MCs do not provide a correct description of the spin correlations among the various partons, particularly at large jet separations. In fact, the above angular quantities are defined in terms of the three-momenta of the particles in the final state and are precisely the spins of the latter that induce very peculiar orientations in (i)–(iv). Conversely, '$\mathcal{O}(\alpha_s^2)$ ME' programs (for example, the '$\mathcal{O}(\alpha_s^2)$ ME + string fragmentation model' as implemented in JETSET, i.e. with no parton shower) yielded a much better angular description of four-jet final states, see Refs. [6, 9]. In fact, all spin correlations are naturally included in a full $\mathcal{O}(\alpha_s^2)$ ME calculation, but they are not necessarily present in a PS emulation based on the $\mathcal{O}(\alpha_s)$ scattering, although their availability (in the infrared limit, where soft and collinear correlations can be factorised in analytic form) is a feature of some of the above mentioned MC codes (for example, for the HERWIG implementation, see Ref. [10]). However, even an '$\mathcal{O}(\alpha_s^2)$ ME + fragmentation' model, without the PS evolution, is inadequate to describe QCD four-jet production in e^+e^- collisions. The problem is that such ME models contain 'ad-hoc' hadronisation which is adjusted to produce a good agreement with some data [6, 11], but they cannot be extrapolated to other energies. Furthermore, their description of the sub-jet structure

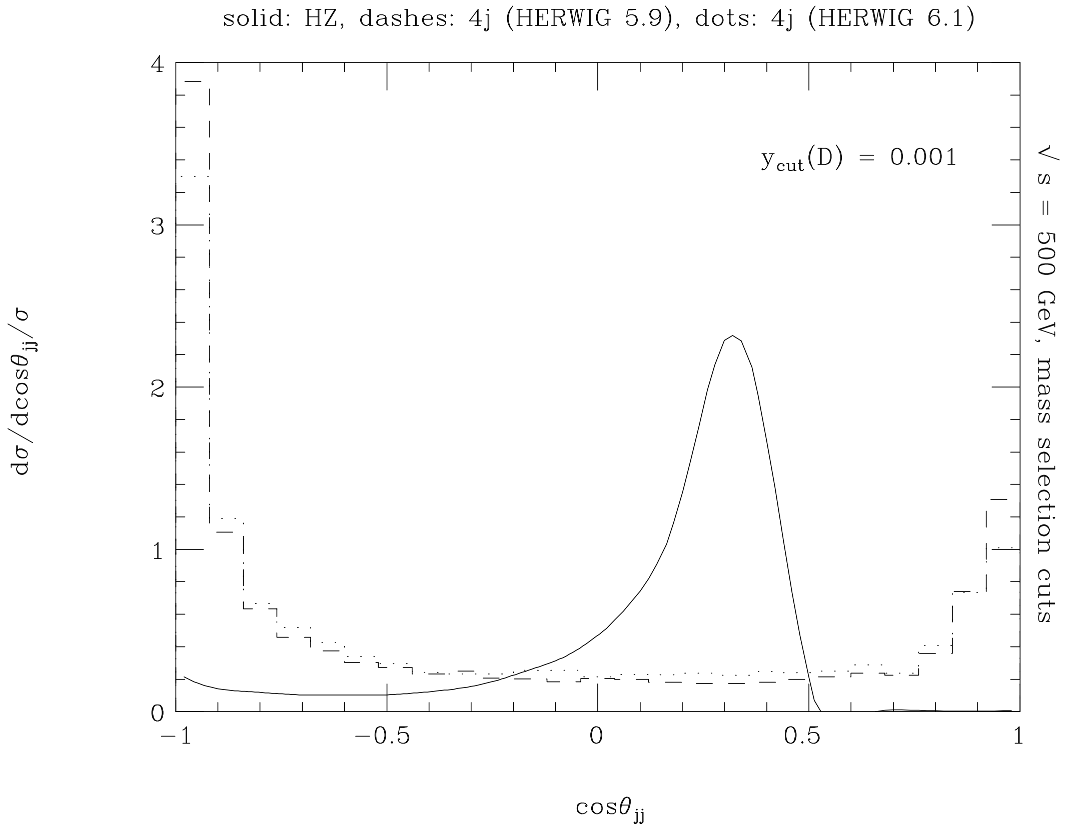


Figure 2: Differential distributions for the opening angle of the jets in the di-jet system reconstructing the Z.

is very poor (see Ref. [12] for an overview).

That such deficiencies in the description of four-jet final states could be cured by an '$\mathcal{O}(\alpha_s^2)$ ME + PS' (plus hadronisation) approach has been clear for some time. Now, several event generators have been made available, that resort to this implementation: HERWIG 6.1 [1], PYTHIA 6.1 [13], 4JPHACT [14] and APACIC++ [15]. It is the intent of this note to describe the impact that the new description of four-jet final states can have in some contexts which are particularly crucial to the physics of future LCs. To do so, we will borrow as an example the case of the HERWIG event generator, though many of its features can equally be appreciated in the case of the other MCs. Eventually, we will also highlight the differences and similarities among these various programs.

Ref. [16] has already made the point, for the case of LEP, that systematic differences exist between the use of the $\mathcal{O}(\alpha_s)$ and $\mathcal{O}(\alpha_s^2)$ MEs in the simulation of QCD four-jet events. This is true not only for the case of QCD studies (e.g., in the determination of the QCD colour factors) but also in the context of $M_{W^\pm}$ measurements. Here, we illustrate similar effects taking place in a higher energy environment, such as a LC operating at $\sqrt{s} = 500$ GeV.

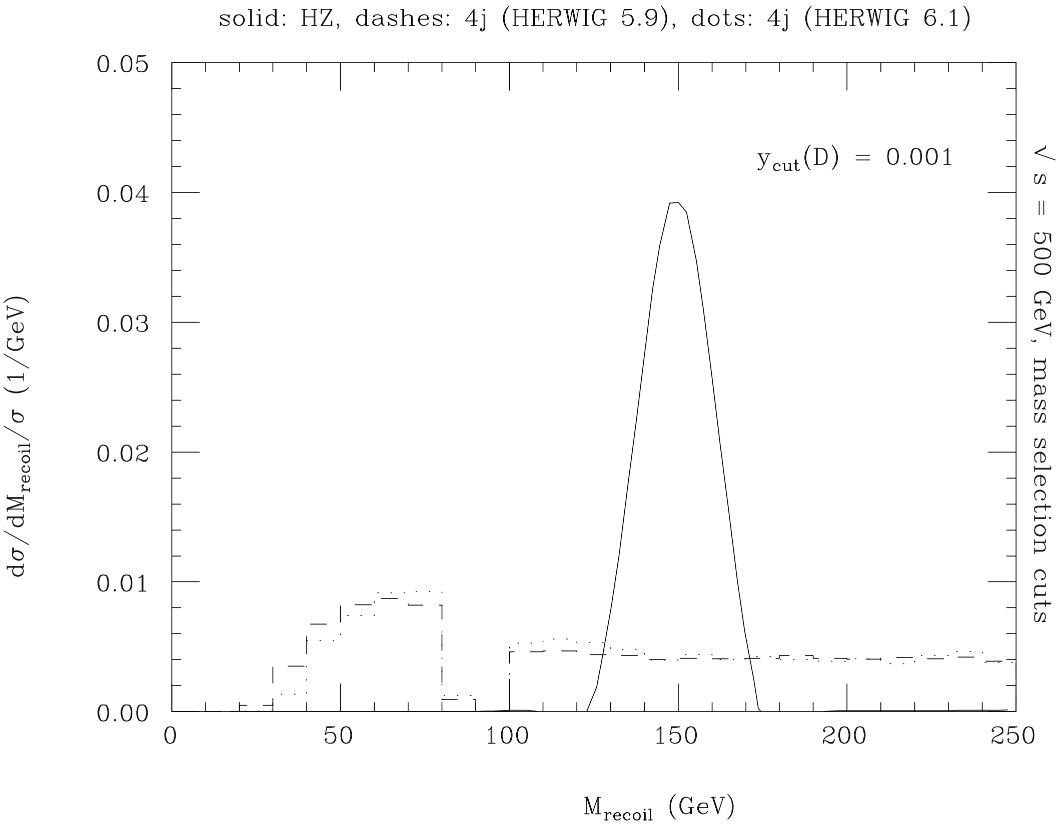
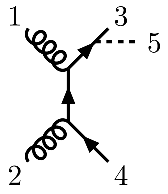
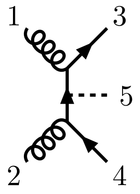


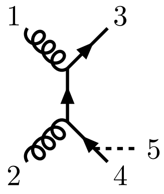
Figure 3: Differential distributions for the invariant mass recoiling against the reconstructed Z.

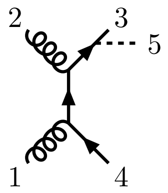
Before proceeding to the numerical part of our note, we would like to spend a few words about some technical aspects of the HERWIG implementation. The new process describing electron-positron annihilation into four jets can be simulated by setting `IPROC=600+IQ`, where a non-zero value for `IQ` guarantees production of quark flavour `IQ` whereas `IQ=0` corresponds to the natural flavour mix. (`IPROC=650+IQ` is as above but without those terms in the MEs which orient the event with respect to (w.r.t) the lepton beam direction.) The MEs are based on those of Ref. [17] with orientation terms from [18]. Infrared (soft and collinear) divergences are avoided by imposing a minimum separation among the four partons (hereafter, denoted by $y_{\rm inf}$ – `Y4JT` in the code, by default equal to 0.01). At present, the jet separation is calculated using either the Durham or Jade metrics (see, e.g. Ref. [19], for definitions and a review on jet clustering algorithms). This choice is governed by the logical variable `DURHAM` (whose default is `.TRUE.`). Note that the phase space is for massless partons, as are the MEs, though a mass threshold cut is applied. The scale `EMSCA` for the parton showers is set equal to $\sqrt{sy_{\rm min}}$, where $y_{\rm min}$ is the least distance, according to the selected metric, between any two partons. Finally, the jet resolution $y_{\rm cut}$ – used to select a multi-jet final state – should be chosen such that it is always larger than $y_{\rm inf}$.

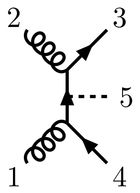
An ambiguity arises in the treatment of interference terms between diagrams with different colour structures. In fact, apart from the trivial case $e^+e^- \to q\bar{q}Q\bar{Q}$, with $q \neq Q$, the MEs for $e^+e^- \to q\bar{q}gg$ and $e^+e^- \to q\bar{q}q\bar{q}$ (i.e. $q = Q$) receive contributions from two distinct colour flows each. In general, the interference terms between the latter are not positive definite and a choice has to be made, on how to assign them to one or the other of the two leading colour contributions (although suppressed by $1/N_C^2$ w.r.t. the latter, they cannot be neglected without inducing sizable effects in both total and differential rates). For illustration, we describe here the treatment adopted in HERWIG for $e^+e^- \to q\bar{q}gg$ and refer the reader to the forthcoming new HERWIG manual [20] (see also [21]) for details on the $e^+e^- \to q\bar{q}q\bar{q}$ subprocess. Things go as follows. Two-quark-two-gluon final states are produced via the following eight Feynman graphs (here, 5 labels the $e^+e^- \to \gamma, Z$ current):

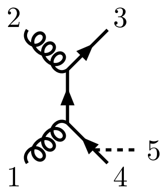
$T_1 =$ 

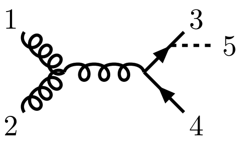
$T_2 =$ 

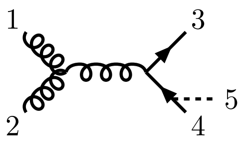
$T_3 =$ 

$T_4 =$ 

$T_5 =$ 

$T_6 =$ 

$T_7 =$ 

$T_8 =$ 

(1)

The two aforementioned colour flows are identified by the terms proportional to $(t^A t^B)_{i_3 i_4}$ and $(t^B t^A)_{i_3 i_4}$, where A and B are the colours of the gluons labelled 1 and 2, respectively. In fact, it should be recalled that the colour pieces associated with the triple-gluon vertices are nothing but the structure constants f^{ABX} of the $SU(N_C)$ gauge group, for which

$$[t^A, t^B]_{i_3 i_4} \equiv (t^A t^B)_{i_3 i_4} - (t^B t^A)_{i_3 i_4} = \mathrm{i} f^{ABX} t^X_{i_3 i_4}, \tag{2}$$

X being the label of the intermediate gluon and i_3, i_4 the outgoing quark colour indices. Therefore, one can conveniently group the original eight Feynman amplitudes as

$$M_+ = \sum_{i=1}^{3} T_i + \sum_{i=7}^{8} T_i, \qquad\qquad M_- = \sum_{i=4}^{6} T_i - \sum_{i=7}^{8} T_i. \tag{3}$$

The two positive definite contributions N_+ and N_- corresponding to the two fundamental colour connections needed for the interface to the `HERWIG` parton shower (see Ref. [22]) can, for example, be obtained as:

$$N_+ = D\,(|M_+|^2 - \frac{1}{N_C^2}|M_+ + M_-|^2 \frac{|M_+|^2}{|M_+|^2 + |M_-|^2}),$$

$$N_- = D\,(|M_-|^2 - \frac{1}{N_C^2}|M_+ + M_-|^2 \frac{|M_-|^2}{|M_+|^2 + |M_-|^2}), \tag{4}$$

where the colour pre-factor is ($N_C = 3$)

$$D = \frac{N_C}{4}(N_C^2 - 1), \tag{5}$$

so that the total squared amplitude can be recovered from

$$M^2(q\bar{q}gg) = \sum_{i=\pm} N_i. \tag{6}$$

In the right-hand sides of eq. (4), the first terms are the 'planar' whereas the second are the 'non-planar' ones [22]. Here, N_+ would correspond to the t-channel colour flow ($4 \to 2, 2 \to 1, 1 \to 3$) and N_- to the u-channel one ($4 \to 1, 1 \to 2, 2 \to 3$), that is, `2413` and `4123` in the notation of Ref. [22]. This is the default procedure adopted in HERWIG 6.1 that we have maintained in our analysis (see Ref. [1] for other possible choices).

Several kinematic quantities can be used to disentangle Higgs events from the backgrounds at a LC. One example is the polar angle of the reconstructed Z boson. In fact, in the case of the signal, one expects the gauge vector to be very central, owning to the s-channel and spin dynamics of the Higgs process [4]. In contrast, no such a behaviour would appear in the QCD noise. Fig. 1 shows this variable as obtained by using $\mathcal{O}(\alpha_s)$ (from HERWIG 5.9, `IPROC=105`) and $\mathcal{O}(\alpha_s^2)$ (from HERWIG 6.1, `IPROC=605`) MEs, supplemented by the subsequent PS and the hadronisation stage. The four-jet final state has been selected by using the Durham (D) jet-clustering algorithm with resolution $y_{\rm cut} = 0.001$[2]. Furthermore, we have imposed that one pair of jets reproduce the Z mass within 10 GeV, i.e. $|M_{\rm jj} - M_Z| < 10$ GeV, as illustration of a typical selection procedure of $e^+e^- \to HZ \to 4$ jet events. These are the two jets that we will consider to be produced in the $Z \to 2$ jet decay. For reference, we also have superimposed the same angular spectrum as obtained from the Higgs process, again generated by using HERWIG (`IPROC=305`). As self-evident from the choice of the `IPROC` numbers, we always require two b-jets in the final state, for both QCD and Higgs processes. The HERWIG default choice $M_H = 150$ GeV is used for the SM Higgs mass. Also notice that we normalise the

[2]Notice that exactly four jets are required to be reconstructed at hadron level. To allow for $n \geq 4$ jets in the final state (eventually clustered into four) would not alter our conclusions. Note that we use the so-called E recombination scheme: two tracks are clustered by summing their four-momenta.

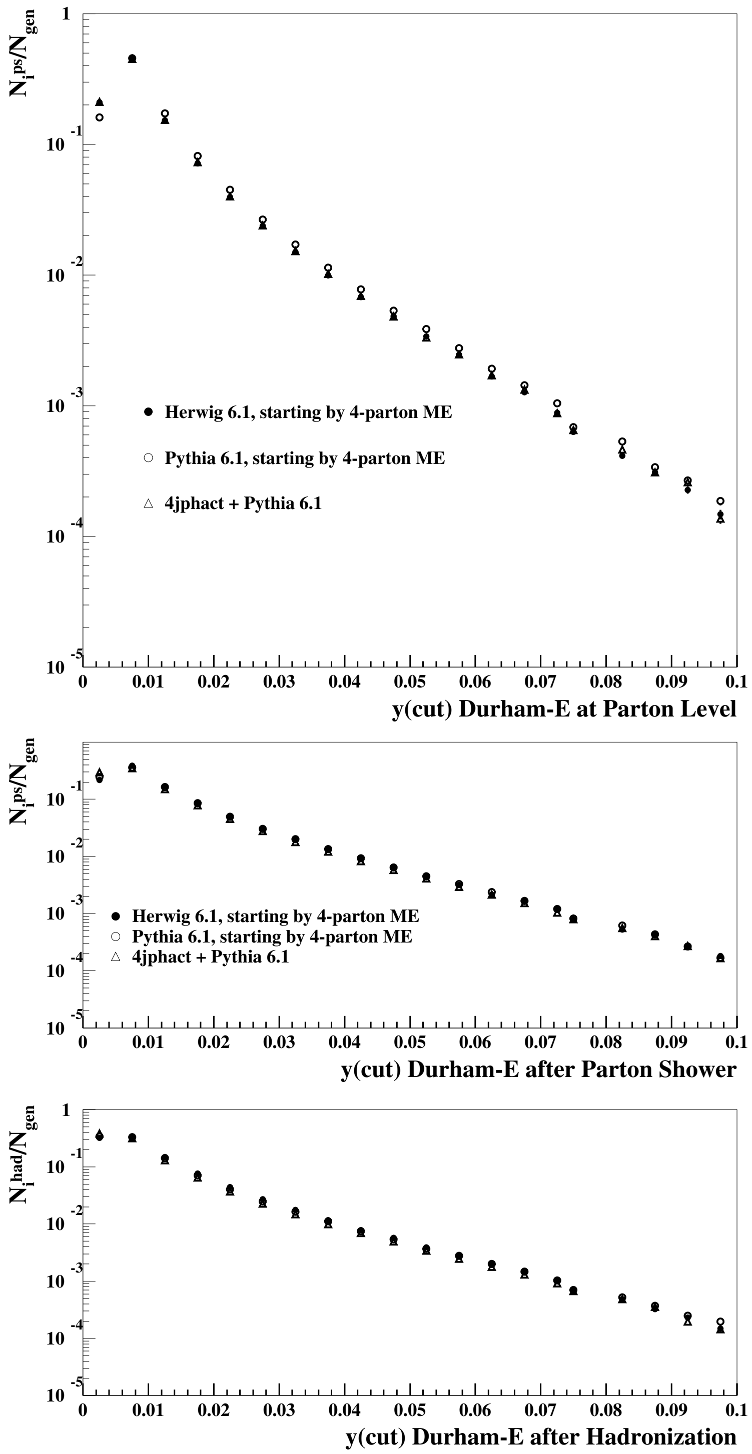


Figure 4: The y_{cut} dependence of four-jet events at parton level before (top) and after (middle) showering effects as well as at hadron level (bottom).

cross sections of all processes to unity. In other terms, we implicitly assume that the two QCD descriptions produce the same event rate, for a given choice of α_s, jet-clustering scheme and $y_{\rm cut}$. In reality, there exist differences in the total production rates between the two QCD implementations, that could be source of further systematic uncertainties. However, we leave these aside for the time being, as we only concentrate on the kinematic effects. It is clear from Fig. 1 that not only the two QCD descriptions show significant differences between each other in the polar angle distribution of the di-jet system emulating a Z decay, but these also persist where the Higgs process accumulates. Similar effects can be appreciated in the distribution of the relative angle between the two jets reconstructing the gauge vector, see Fig. 2. This quantity too is used in the selection of Higgs events at a LC [4]. Thus, although the behaviour of the two QCD implementations is rather similar in the vicinity of the Higgs peak, see Fig. 3, the effect of cuts enforced on $\cos\theta_{\rm beam}$ and/or $\cos\theta_{\rm jj}$ can be dramatically different for the actual number of background events falling around $M_{\rm recoil} \approx M_H$. This can be appreciated in Tab. 1.

HZ	$\mathcal{O}(\alpha_s)$	$\mathcal{O}(\alpha_s^2)$
72. %	9.8 %	12. %

Table 1: Percentage of four-jet events in a window $|M_{\rm recoil} - M_H| < 30$ GeV, after the cuts $|\cos\theta_{\rm beam}| < 0.6$ and $\cos\theta_{\rm jj} < 0.5$.

As mentioned earlier on, at least three other MCs exist that have the option of generating multi-jet final states starting from the $\mathcal{O}(\alpha_s^2)$ MEs. Among these, we consider here PYTHIA 6.1 and 4JPHACT. The former makes use of the same (massless) MEs implemented in HERWIG (with a similar approximate procedure to account for mass effects [12]), the main differences being in the treatment of the PS and hadronisation. The latter, while using showering and fragmentation dynamics borrowed from PYTHIA, employs the fully massive MEs of Ref. [23] and a massive four-body phase space as well. A comparison among these MCs has been performed by Silvia Bravo for ALEPH and we present here just a selection of the results she obtained[3]. This is done in Fig. 4, where the $y_{\rm cut}$ dependence of the four-jet rates (all flavour combinations), as obtained in the three MC environments, is presented at three different stages. First, at the four-parton level before showering effects, then after the latter have taken place, finally after hadronisation. The comparison is performed at $\sqrt{s} = M_Z$ and exactly four jets are required to be selected by the Durham jet finder (using E-scheme recombination), for a given $y_{\rm cut}$ value. Normalisation is to the number of generated events: 1,000,000 in each case. The overall agreement is quite good, particularly at the experimentally more relevant hadron level. In the end, systematic errors induced by the different treatment of MEs, phase space, PS and hadronisation in the three MCs are typically much smaller than those between any $\mathcal{O}(\alpha_s^2)$ description and the standard $\mathcal{O}(\alpha_s)$ ones.

[3] Others can be found at `http://www.pd.infn.it/ecfa/`.

In summary, we have shown that there exist sizable differences in the MC simulation of key LC phenomenology in four-jet final states, depending upon whether $e^+e^- \to q\bar{q}g$ or $e^+e^- \to q\bar{q}gg$ plus $e^+e^- \to q\bar{q}Q\bar{Q}$ partonic scatterings are used to initiate the PS of the event generator. However, several reliable MC programs exist by now, each exploiting a different implementation of either approach. Thus, new, more sophisticated analyses of four-jet events will soon be available, these allowing one to pin down to even better accuracy the physic potential of a future e^+e^- collider operating at the near-TeV scale.

Before concluding, we would like to briefly remark on other aspects that enter the description of hadronic final states with arbitrary jet multiplicity, that were beyond the scope of this note. Here, we have been concerned with what we should like to call the procedure of 'interfacing' MEs to the subsequent PS, that is, of generating hadronic events through higher-order MEs, but limiting the event generation to a phase space region where the former are reliable, without any attempt to resort to the latter otherwise. In fact, another important issue is that of 'matching' MEs and PS, by which we mean the procedure of exploiting a combined approach that uses MEs for large phase space separations, PS in the infrared limit and modified MEs (by Sudakov form factors) in the overlapping region [24]. This is what is currently done within the APACIC++ event generator [15], which uses exact leading-order (LO) MEs for $e^+e^- \to n$ partons, with $n \leq 5$[4]. Note that the procedure of Ref. [24] is correct to next-to-leading logarithmic (NLL) accuracy and could possibly be extended to next-to-leading order (NLO) as well [26]. The interface/matching to the PS of higher-order MEs including loops [17, 27] (hence, generating negative weights) remains instead a challenge for the future [28].

Acknowledgements: I would like to thank Silvia Bravo for the kind permission to reproduce here some of her unpublished material. Useful conversation with A. Ballestrero, M.H. Seymour, T. Sjöstrand and B.R. Webber are also acknowledged.

References

[1] G. Corcella, I.G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M.H. Seymour and B.R. Webber, preprint Cavendish-HEP-99/17, December 1999, `hep-ph/9912396`.

[2] G. Marchesini, B.R. Webber, G. Abbiendi, I.G. Knowles, M.H. Seymour and L. Stanco, *Comp. Phys. Commun.* **67** (1992) 465.

[3] See for example: Proceedings of the Workshop 'e^+e^- Collisions at 500 GeV. The Physics Potential', Munich, Annecy, Hamburg, 3–4 February 1991, ed. P.M. Zerwas, DESY 92–123A/B, August 1992, DESY 93–123C, December 1993.

[4] See for example: 'Higgs Particles' sections, in Ref. [3].

[5] J.D. Bjorken, Proceedings of the ' Summer Institute on Particle Physics', *SLAC Report* 198 (1976); B.W. Lee, C. Quigg and H.B. Thacker, *Phys. Rev* **D16** (1977) 1519; B.L. Ioffe and

[4] A similar implementation, for $n \leq 4$, is also provided in the HERWIG context [25].

V.A. Khoze, *Sov. J. Part. Nucl.* **9** (1978) 50; J. Ellis, M.K. Gaillard and D.V. Nanopoulos, *Nucl. Phys.* **B106** (1976) 292.

[6] ALEPH Collaboration, *Z. Phys.* **C76** (1997) 1.

[7] S. Bethke, A. Ricker and P.M. Zerwas, *Z. Phys.* **C49** (1991) 59; S. Moretti and J.B. Tausk, *Z. Phys.* **C69** (1996) 635.

[8] A. Ballestrero *et al.*, *J. Phys.* **G24** (1998) 365.

[9] G. Cowan, *J. Phys.* **G24** (1998) 307.

[10] I.G. Knowles, *Nucl. Phys.* **B310** (1988) 571, *J. Phys.* **G17** (1991) 1562.

[11] OPAL Collaboration, *Z. Phys.* **C65** (1995) 367.

[12] J. André and T. Sjöstrand, *Phys. Rev.* **D57** (1998) 5767; J. André, *preprint* LU-TP-97-12, June 1997, `hep-ph/9706325`.

[13] T. Sjöstrand, *Comp. Phys. Commun.* **39** (1984) 347; M. Bengtsson and T. Sjöstrand, *Comp. Phys. Commun.* **43** (1987) 367; for the latest PYTHIA version, see: `http://www.thep.lu.se/staff/torbjorn/Pythia.html`.

[14] A. Ballestrero, in preparation.

[15] F. Krauss, R. Kuhn and G. Soff, *J. Phys.* **G26** (2000) L11, *Acta Phys. Polon.* **B30** (1999) 3875.

[16] S. Moretti and W.J. Stirling, *Eur. Phys. J.* **C9** (1999) 81.

[17] R.K. Ellis, D.A. Ross and A.E. Terrano, *Nucl. Phys.* **B178** (1981) 421.

[18] S. Catani and M.H. Seymour, *Phys. Lett.* **B378** (1996) 287; see also: `http://hepwww.rl.ac.uk/theory/seymour/nlo/`.

[19] S. Moretti, L. Lönnblad, T. Sjöstrand, *JHEP* **08** (1998) 001.

[20] G. Corcella, I.G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M.H. Seymour and B.R. Webber, in preparation.

[21] K. Odagiri, *JHEP* **10** (1998) 006.

[22] G. Marchesini and B.R. Webber, *Nucl. Phys.* **B310** (1988) 461.

[23] A. Ballestrero, E. Maina and S. Moretti, *Phys. Lett.* **B294** (1992) 425, *Nucl. Phys.* **B415** (1994) 265, Proceedings of the 'XXIXth Rencontres de Moriond', Méribel, France, March 1994, ed. J. Trân Thanh Vân (ed. Frontières, Gif-sur-Yvette, 1994), page 367.

[24] B.R. Webber, Talk given at 'XXXVth Rencontres de Moriond', Les Arcs, France, March 2000, preprint Cavendish-HEP-00/05, May 2000, `hep-ph/0005035`.

[25] See: `http://webber.home.cern.ch/webber/`.

[26] J. Collins, *JHEP* **0005** (2000) 004.

[27] Z. Bern, L. Dixon, D.A. Kosower and S. Weinzierl, *Nucl. Phys.* **B489** (1997) 3; Z. Bern, L. Dixon, D.A. Kosower, *Nucl. Phys.* **B513** (1998) 3; L. Dixon and A. Signer, *Phys. Rev. Lett.* **78** (1997) 811; E.W.N. Glover and D.J. Miller, *Phys. Lett.* **B396** (1997) 257; J.M. Campbell, E.W.N. Glover and D.J. Miller, *Phys. Lett.* **B409** (1997) 503.

[28] B. Pötter, preprint MPI/PhT/2000-24, July 2000, `hep-ph/0007172`.