Chapter 3

Matrix Elements and Shower Matching

Author: Mauro Moretti

Revisors: Paolo Nason

3.1 Introduction

As discussed at length in the previous chapters, final states with many hard jets will play an essential role for LHC physics. These events will hide or strongly modify all possible signals of new physics, which involve the chain decay of heavy coloured particles, such as squarks, gluinos or the heavier partners of the top, which appear in little-Higgs models. Being able to predict their features is therefore essential. To this end it is crucial to describe as accurately as possible both the full matrix elements (ME) for the underlying hard processes, as well as the subsequent development of the hard partons into jets of hadrons.

It is therefore very important to design a strategy to take advantage of the strength (and avoid the drawbacks) of both fixed order calculations and of Parton Shower-like evolution with subsequent hadronization of the partonic event. A given $(n+1)$-jet event can be obtained in two ways: from the collinear/soft-radiation evolution of an appropriate $(n+1)$-parton final state, or from an n-parton configuration where hard, large-angle emission during its evolution leads to the extra jet. A factorization prescription (in this context this is often called a “matching scheme” or “merging scheme”) defines, on an event-by-event basis, which of the two paths should be followed. The primary goal of a merging scheme is therefore to avoid double counting (by preventing some events to appear twice, once for each path), as well as dead regions (by ensuring that each configuration is generated by at least one of the allowed paths). Furthermore, a good merging scheme will optimize the choice of the path, using the one which guarantees the best possible approximation to a given kinematics.

Here we shall briefly review two such merging approaches: the CKKW scheme [1, 2], and the MLM scheme [3]. These two approaches are implemented in currently used matrix element event generators, combined with parton showers tools, like SHERPA [4, 5], ALPGEN [6], MADGRAPH [7] and HELAC [8].

Any merging algorithm is based on one or more *resolution parameters* which split the phase space into two regions one of soft/collinear emissions to be described by Parton Shower (PS) evolution and the other one of hard and large angle emission to be described by fixed order calculations. These resolution parameters play the role of soft/collinear cut-off for fixed order calculations and it is therefore crucial to assess the (in)dependence of the algorithm on these parameters. Notice that if both PS and ME descriptions would provide a *perfect* description of QCD the final result would be independent of the resolution parameters.

For the CKKW scheme, in the context of $e^+e^- \to jets$, it has been shown [1] that the dependence on the resolution parameter is shifted beyond the Next to Leading Log (NLL) accuracy.

Such a proof in the context of *ep* and *pp* collisions is missing and thus for both CKKW (adapted to hadronic collision [2]) and for MLM scheme we don't have any avaliable extimate of the dependence of the final result on the resolution parameters. *Ultimately, at present, such an extimate is possible only empirically*: one has to study the effect of varying the resolution parameters on the widest possible range.

A first series of studies to address both dependence on the resolution parameters and the comparison of the two schemes has been presented in [10].

The internal consistency of CKKW (as implemented in the SHERPA [9] event generator) inspired approach for hadronic collisions has been studied in [11, 12, 13] for Drell-Yan processes at the TEVATRON and at the LHC.

The internal consistency for the MLM approach, as implemented in the ALPGEN [6] event generator, has been addressed in [14] for the process $t\bar{t}$ plus jets.

Monte Carlo event samples for associate productions of jets and W and Z bosons and for jets productions at the TEVATRON colliders have been compared with data [15, 16, 17] finding an overall satisfactory agreement both for the shapes of the distributions and for *relative* jets multiplicities.

Finally an extensive set of comparisons among various codes and matching prescriptions has been presented in [19, 20] where, in addition to a wide range of tests of internal consistency for the variuos codes, a first attempt to assess some of the systematic uncertainties associated to these approaches (α_S and PS scales) is presented. In [19, 20] results are presented also for the event generator ARIADNE [21] and for the Lönnblad matching prescription [22, 23] (a variant to CKKW adapted to the dipole emission approximation which is the root of ARIADNE PS).

3.2 Matching

Let's first try a sort of "pedagogical" introduction to the matching issue. Our goal is to use the capability to compute fairly complex leading order (LO) matrix elements (ME) to describe hard QCD radiation and to complement this description with showering, to include soft and collinear corrections, and hadronization, allowing a realistic description of the event.

The most simple approach is:

- Use the ME to compute the *WEIGHT* of a given event.
- Use the ME computation as a "seed" for the Parton Shower (PS) evolution: the PS needs as inputs the ME weight, the event kinematic, the colour flow associated to the event. (As well as the factorization and renormalization scales chosen for the ME calculation)

This approach, however, leads to double counting: the same final state can arise in many different ways just swapping ME element generated partons and shower generated partons as shown in fig. 3.1.

This effect is formally NLO (indeed any PS emission implies an additional power of α_s) and therefore beyond the accuracy of our computation. However it opens the possibility to particularly harmful events: soft and/or collinear ME partons toghether with hard shower emission to replace the missing hard jets, as shown in figg. 3.1 and 3.2. The ME weight is *divergent* for soft/collinear emissions and those events comes without the Sudakov suppression supplied by the showering algorithms and therefore leads to infrared and collinear sensitivity (it's worth recalling that the *PS algorithm doesn't modify the ME WEIGHT*, it simply dresses the event with soft and collinear radiation). Notice that, as thoroughly discussed in the previous chapters of these proceedings, soft/collinear emissions described by the PS don't exibit the same unphysical behaviour: Sudakov form factors ensure that virtual effects are accounted for (in the NLL approximation) and thus enforce the appropriate dumping of the singularities.

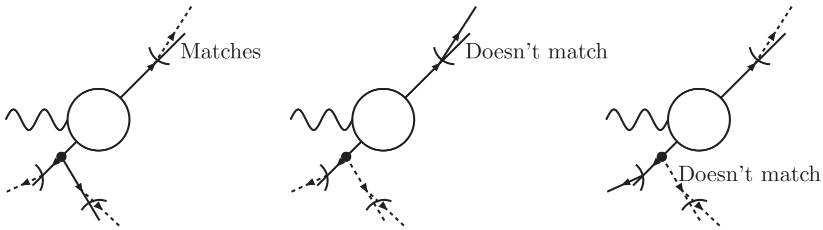


Figure 3.1: Hadron production in e^+e^- collisions via γ^*, Z^* exchange. Example of double counting in ME PS combination. Wiggly line: γ^*, Z^*; solid lines: ME (coloured) partons; dashed lines: PS emissions. The same events obtained in three different ways. Left: hard emissions from ME and soft/collinear ones from PS, Center: one soft emission from ME and one hard emission from PS, Right: one collinear emission from ME and one hard emission from PS.
The second and the third one lack the appropriate Sudakov suppression and lead to a divergent cross section. The first one is the one we would like to retain.
Small arcs denote clusters used in MLM matching prescription.

We are therefore forced to find a way to avoid double counting or at least to push its impact below the accuracy of our prediction. The final goal is to split the phase space in two regions: one, of soft and/or collinear emissions, to be covered from the PS algorithm and the other one, of hard and large angle emissions, to be described by the matrix element. The separation among these two regions is achieved introducing one or more "resolution parameters" which discriminate among "resolved" jets (to be described by the ME) and "non resolved" jets to be described by the PS. Notice that the solution has to fulfill three main requirements

- It should avoid (minimize) double counting and ensure full phase space coverage
- It should ensure a smooth (as much as possible) transition among the PS and ME description
- It should ensure that the ME weight is reweighted with the appropri-

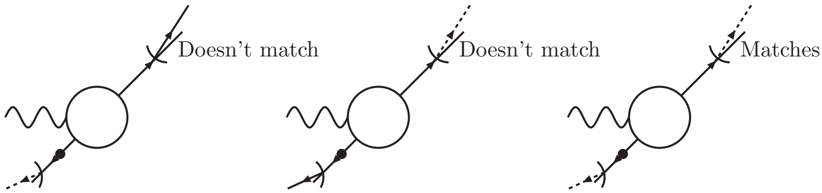


Figure 3.2: Same symbols (and process) as in fig. 3.1. The same events obtained in three different ways. Left: one collinear emission from ME and soft/collinear emissions from PS, number of jets smaller than number of ME partons; Center: one soft emission from ME and soft/collinear emissions from PS, number of jets smaller than number of ME partons; Right: hard emissions from ME and soft/collinear emissions from PS. The first and the second one lack the appropriate Sudakov suppression and lead to a divergent cross section. The third one is the one we would like to retain.

ate *Sudakov form factor*, where by appropriate we mean that it should reabsorb the divergencies of the ME weight. [1]

[1] If we denote with $\mathcal{R}^{(n)}_{ME}$ the real radiative correction to the $n-$jets squared matrix element X^n_{ME} with l and L the soft and collinear logarithms respectively and with ξ an infrared/collinear finite quantity we shall have

$$\mathcal{R}^{(n)}_{ME} = X_{ME}\alpha_S(c_1Ll + c_2L + c_3l) + \alpha_S\xi$$

and the corresponding "Sudakov form factor" (to be used to reweight X_{ME}) Δ has to be

$$\Delta = \exp[-\alpha_S(c_1Ll + c_2L + c_3l)]$$

Notice that with a wrong choice of Δ (different c_j) one still obtains infrared/collinear finitness (*for the reweighted* X_{ME}), the result however will exhibit a strong dependence on the chosen soft/collinear cut-off.

3.3 Matching ME and PS: a practical perspective

Let's now have a look at the practical implications of the double countig problem. One has in mind an event generator which combines the benefit of (fixed-order) ME calculation and showering (+ hadronization).

Let's first attempt the more naive approach:

- Use the ME to compute the *WEIGHT* of a given event.
- Use the ME computation as input for the PS

One immediately faces the problem to determine the appropriate parton level cuts required to build up the event sample. Notice that this is *mandatory* if one has final state coloured partons (emitted by coulored partons): in the absence of cuts the ME *diverges*.

A first attempt is to use, as parton level cuts, the same cuts used to define a jet in the analysis.

Let's have a look at the consequences. We analyze the answer of our event generator (after ME computation and showering) looking at *jets observables*.

To reconstruct *jets* out of final state partons (namely those found after the showering stage) we shall use a simplified cone algorithm as provided by the GETJET package [24], which represents a simplified jet cone algorithm a la UA1. Jets are defined requiring that jet p_T has to be at least 20 Gev, the cone size is $R = 0.4$ and the calorimeter coverage is $|\eta| < 2.5$.

Ultimately we shall study the signal $p\bar{p} \to e^+e^- + 2$ *jets with at least two jets with* $p_T > 40$ *GeV and with* $\Delta R > 0.7$ at the LHC COLLIDER.

We start by generating $p\bar{p} \to e^+ + e^- + 2$ *partons* ($parton \equiv g, u, d, c, s$) with $p_T > 40$ Gev, $|\eta| < 2.5$ and $\Delta R_{p_j p_k} < 0.7$. After ME computation the event is showered with PITHYA PS and the jets are reconstructed according to the chosen jet algorithm.

In fig. 3.3 we display the p_T of the second leading jet (jets ordered according to p_T) for the events that, *at parton level (ME)*, have the second highest p_T parton with a p_T between 40 and 50 GeV and with a p_T between 50 and 60 GeV. The effect of the shower is to smear the parton p_T: some of the partons have their energy degraded by radiating energy, other partons actually originate a harder jet collecting soft energy (mostly originated by initial state radiation). We are now facing a problem: by imposing generation cuts equal to the jet resolution parameters we are loosing the contribution of ME partons with a p_T just below threshold which after showering would anyhow make up a jet with a p_T larger than that chosen in the analysis. A similar "edge effect" occur for "close" ($R \simeq 0.7$), see fig. 3.3 or large rapidity partons ($\eta \simeq 2.5$), see fig. 3.3.

An obvious solution to this problem is to soften the generation cuts. In this way we loose efficency since many of the soft/collinear partons don't originate resolved jets, however we recover the event which we were missing in the previous analysis. We however face another problem: our prediction is not stable against generation level cuts. To see the effect we study the subsample of events that, *after showering* have at least two jets with $p_T > 40GeV$ and $\Delta R > 0.7$. As it is seen in fig. 3.4 the cross section increases as parton level generation cuts are softened and also distributions are affected. *Notice that resolution parameters for jets, as well as the event selection criteria, are unchanged and therefore the results, after showering should remain unchanged.*

The reason of this behaviour can be traced back to the problem of double counting associated with soft/collinear ME emission. In the soft/collinear limit the ME weight diverge, the PS can supply a hard and large angle emission:[2] this is suppressed by a factor of α_S but enhanced by soft/collinear logarithms which (as opposite to soft/collinear PS emission) *are not dumped* by Sudakov suppression.

3.4 Catani, Krauss, Kuhn and Webber algorithm

A solution has been proposed in [1] in the context of e^+e^- collisions. The dependence on the resolution parameter is shifted beyond NLL. In [2] an extension of the procedure to *ep* and *pp* environments has been proposed, without however a proof that the dependence on the resolution parameter is below NLL. This algorithm is implemented [5] in the `SHERPA` MC [4] and has been studied in [10] for `HERWIG` and `PITHYA` showers.

3.4.1 PS and ME phase space boundaries

The first ingredient of the algorithm is the *measure of parton-parton separation.* To this pourpose the $k_\perp$ jet algorithm [25, 26, 27] is used: *the distance among two final state partons* is defined as

$$y_{ij} = \frac{2\min\{E_i^2, E_j^2\}(1-\cos\theta_{i,j})}{s} \tag{3.1}$$

s being the center of mass squared energy, $E_{i,j}$ the parton energies and $\theta_{i,j}$ their relative angles. *The "distance" between a parton and the incoming partons (the*

[2]Notice that, if one or more ME parton are "soft", there must be a corresponding number of "hard" PS emissions in order to preserve the number of "observed" hard jets

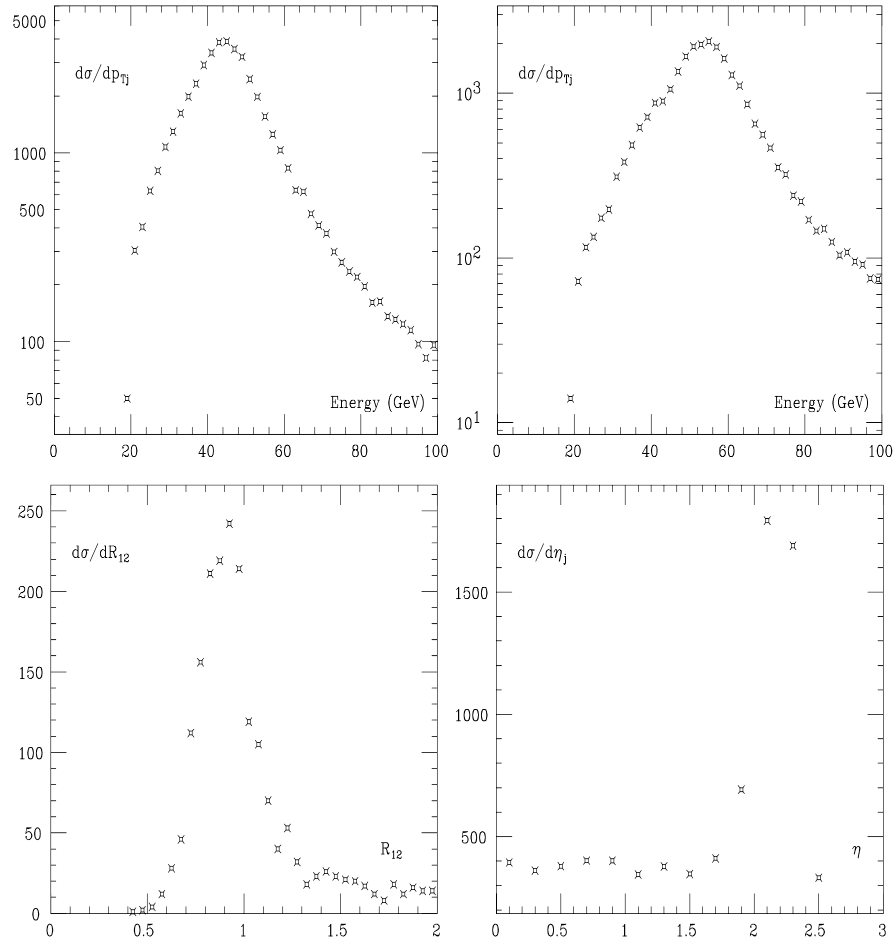


Figure 3.3: **Upper-left panel:** $p_T^{j_2}$ of the next to leading jet (jets ordered in p_T) for *showered* events initiated by partonic event $Z^*/\gamma^* + 2$ partons, subject to the constraint $40 < p_{T2} < 50$ GeV, p_{T2} being the p_T of the next to leading parton (partons ordered in p_T). **Upper-right panel:** Same as Upper-left panel but $50 < p_{T2} < 60$ GeV **Lower-left panel:** distance ΔR_{12} among the two leading jets for *showered* events initiated by partonic events $Z^*/\gamma^* + 2$ partons, subject to the constraint $0.7 < \Delta R_{12}^{partonic} < 1$. **Lower-right panel:** rapidity $|\eta^{j_1}|$ of the leading jet for *showered* events initiated by partonic event $Z^*/\gamma^* + 2$ partons, subject to the constraint $2.5 > |\eta^1| > 2.0$, η^1 being the rapidity of the leading parton. *All plots are for the LHC, and the normalizations are arbitrary*

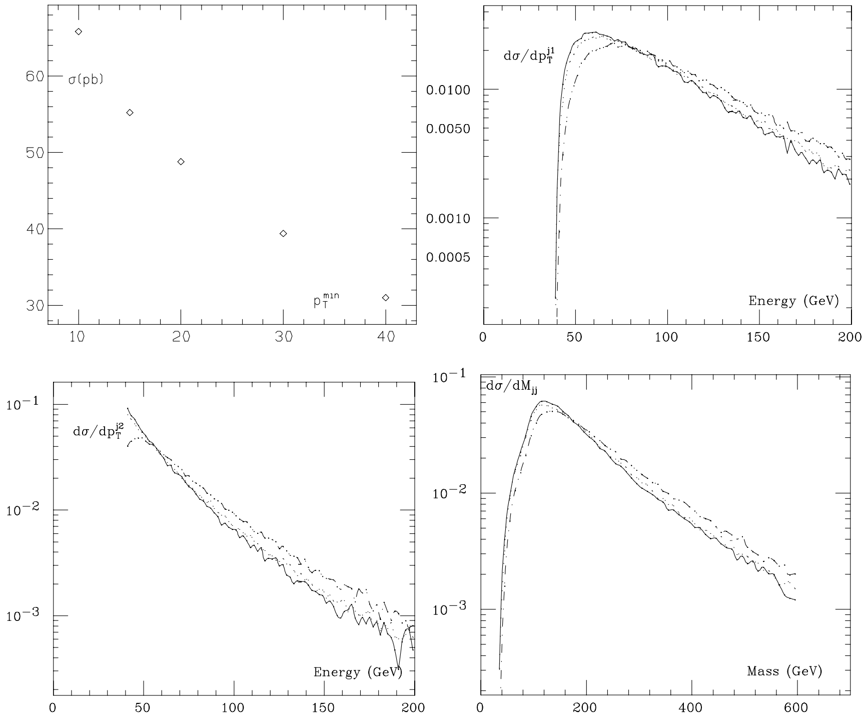


Figure 3.4: **Upper-left panel:** Cross section in pb for $pp \to e^+e^- + 2$ jets at the LHC as a function of the *partonic* p_T *at the generation level.* Both jets, *after showering* are required to have $p_T > 40$ GeV, $|\eta_j| < 2.5$ and $\Delta R_{j_1 j_2} > 0.7$. **Upper-right panel:** transverse momentum p_T^{j1} of the leading jet as a function of parton level cuts. Continuos: $p_T^{part} > 10$ GeV; dots: $p_T^{part} > 20$ GeV; dot-dash: $p_T^{part} > 40$ GeV; **Lower-left panel:** transverse momentum p_T^{j2} of the next to leading jet as a function of parton level cuts. **Lower-right panel:** invariant mass $m_{j_1 j_2}$ of the two leading jets as a function of parton level cuts.

beam) is defined as

$$y_i = \frac{p_{\perp i}^2}{s} \tag{3.2}$$

The separation among ME partons and PS partons is achieved introducing a resolution parameter Y_{sep} and

- requiring that ME partons are *resolved*:

$$y_{i,j}\,,\ y_i > Y_{sep}$$

- vetoing PS emissions at a scale harder than Y_{sep}

This ensures that a given phase space configuration is covered only once

Notice that in the region described by the PS dead zones are still present and thus one has to choose Y_{sep} in such a way to minimize these effects in the regions relevant for the analysis of interest.

One could also use a different measure of the *parton-parton* distance, it is however necessary that it preserve the properties of $k_\perp$ algorithm if one wishes to retain NLL accuracy.

3.4.2 Matching ME and PS weight

The second key ingredient is *ME reweighting*. The ME weight is infrared and collinear divergent and thus will diverge as Y_{sep} becomes small. On the other hand the PS is well behaved in this limit due to soft and collinear emission resummation. The ME is thus reweighted in order to ensure a smooth transition among ME and PS description:

- for a given ME phase space point a *branching tree* is reconstructed by clustering toghether the two *closest partons* (according to y measure given in eqns. (3.1,3.2)) and iterating the procedure until when the "leading order" process is reached: for $pp \to W + \text{n} - jets$ we proceed until $qq' \to W$ is reached, for $pp \to t\bar{t} + \text{n} - jets$ until $pp \to t\bar{t}$ is reached[3]
- for each branching reweight the squared ME by $\alpha_S(k_\perp)/\alpha_S(Q_{ME})$

[3]some qualification is actually required: if the scale of some QCD emission is larger than the typical scale for the LO process the clustering is done in a different way. We refer to [2] for a more thorough discussion.

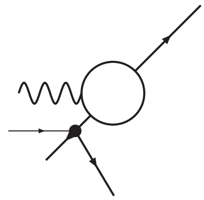


Figure 3.5: Same symbols (and process) as in fig. 3.1. ME final state partons (originating from small dots) are clustered toghether to reduce the process to the leading order $2 \to 2$ process. Small dots represents would be branchings in a PS-like picture of the event. The $k_\perp$ separation among the clustered partons is the appropriate PS scale for α_S evaluation at the given branching.

- to each internal and external line of the branching tree associate the proper combination of *Sudakov*[4] *form factor*: defining

$$Y_{sep} = \frac{Q_1^2}{Q^2}$$
$$y_j = \frac{q_j^2}{Q^2}$$

  where Q^2 is the hard process scale, to each *internal line*, connecting a branching at a scale q_j and a branching at a scale q_k, associate a reweighting factor

$$\frac{\Delta(Q_1, q_k)}{\Delta(Q_1, q_j)}$$

  where $\Delta(Q_1, q_j)$ is the appropriate Sudakov Form Factor. To each external line, originating from a branching at a scale q_j associate a reweighting factor

$$\Delta(Q_1, q_j)$$

- also the PS needs to be modified:

[4]for a thorough discussion of Sudakov form factors meaning and definitions refer to the previous chapters of this proceedings.

1. *the scale* for the PS evolution is given, for each parton, by the scale at which the parton was produced (the hard process scale for initial state partons)
2. *resolved* PS emissions ($y > Y_{sep}$) are inhibited. Notice that this is done simply *rejecting* those emissions without affecting the event weight

3.4.3 Building the event sample

- Finally one has to build event samples with up to ∞ ME partons (each normalized to the same luminosity, at least in principle) and sum them up toghether.
- One obviously has to stop to some *finite* number of ME partons. The highest multiplicity sample needs to be treated separately: for a given ME the smallest $k_\perp$ separation is computed and the PS is allowed to produce branching up to this scale. In this way the higher parton multiplicities are supplied by the shower emissions.

In [1] it is shown that, with the above prescriptions, the NLL resummed exclusive $e^+e^- \to n$ jets is reproduced.

A few remarks are in order

- the proof of NLL accuracy holds only for e^+e^- collisions;
- even in the e^+e^- framework, to achieve NLL accuracy, it is crucial that the employed PS *correctly describes the soft sructure of the ME, including interferences*: this is the case for PS incorporating coherent branching like HERWIG or based on dipole emission like ARIADNE but not for virtuality ordered PS like PITHYA. Notice that APACIC (SHERPA) provides both options: virtuality ordered and angular ordered PS[5] and thus it provides the opportunity to study the numerical impact of the two approaches.
- ultimately the smoothness of the interpolation must be judged inspecting the stability of the relevant (for the analysis) ditribution over at least a sizable range for the resolution parameter.
- the sample with the highest multiplicity of ME emissions is also the one with the larger systematics. One should care to minimize its weight on the inclusive sample and anyway to check “indipendence” of the predictions from the maximum number of ME partons used to build up the sample.

[5] actually the first emission is not described by coherent branching

Let's finally add a few words of caution

- *NLL accuracy is already ensured by the PS standalone (if coherent effects are included).*
- *the ultimate goal of ME-PS merging is to correctly describe* hard and large angle emissions *toghether with soft/collinear resummation. This class of events is suppressed by at least* two powers of log *and thus the proof of ([1]) doesn't ensure that they are dealt with correctly.*
- *in particular if the Sudakov reweighting of* hard and large angle emissions *is not correct the hard tail of the distributions will suffer of* LL dependence on the resolution parameter *and thus of artificial enhancement/dumping.*

3.4.4 Implementation and comparison with TEVATRON data

The CKKW algorithm for *pp* collisions, according to the proposal in ([2]), is implemented in SHERPA [9] and has been studied in [11, 12, 13]. The overall consistency looks good:

- the overall rate is stable against sizable changes of the resolution parameters.
- the distributions doesn't show large discontinuities around the resolution parameters.
- stability is achieved with a moderately small number of ME partons.
- there is a nice agreement with `MC@NLO` ([28])

There is ongoing experimental activity in testing SHERPA predictions expecially for jet related quantities. D0 collaboration has studied $Z + jets$ production. A thorough account can be found in [15], the overall agreement looks pretty good. In fig. 3.6 we show the comparison of SHERPA prediction and data for the p_T of the Z boson and of the two leading jets and for the jet multiplicity.

D0 collaboration has also studied [16] dijet azimutal correlation in pure jet sample and compared DATA to SHERPA predictions, again finding good agreement as shown in fig. 3.7.

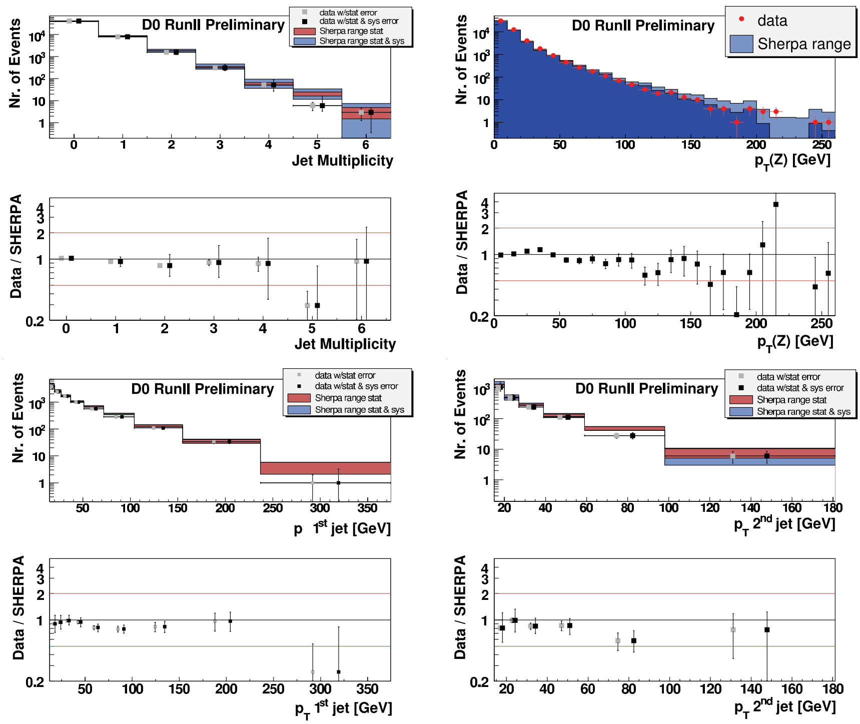


Figure 3.6: **Upper-left panel:** Jet multiplicity in inclusive Z production **Upper-right panel:** $p_T^{j_1}$ of the leading jet (jets ordered in p_T) in $Z + jets$ production **Lower-left panel:** $p_T^{j_1}$ of the leading jet (jets ordered in p_T) in $Z + jets$ production **Lower-right panel:**$p_T^{j_2}$ of the next to leading jet (jets ordered in p_T) in $Z + jets$ production *All plots are for the Tevatron and the normalization of SHERPA prediction is fitted to the data. Both absolute values and SHERPA to DATA ratio are shown. Figures from* [15])

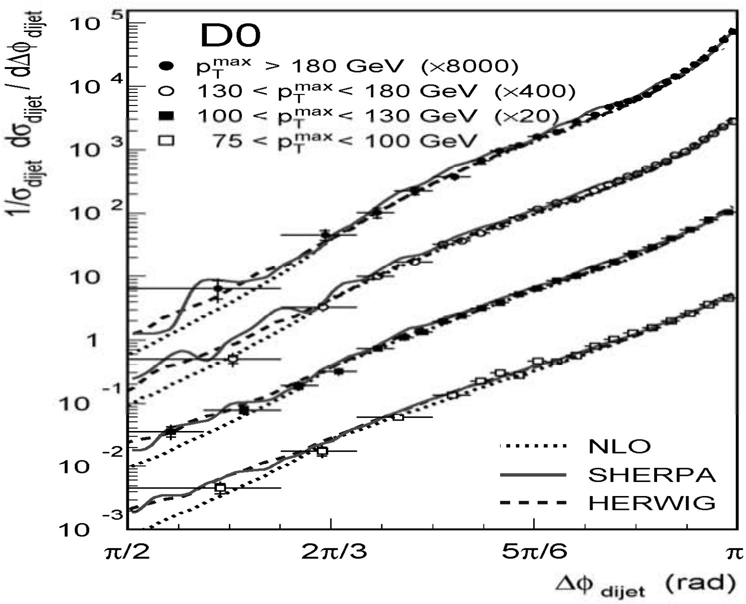


Figure 3.7: Angular separation $\Delta\Phi$ in the transverse plane among the two leading jets (jets ordered in p_T) in inclusive two jets sample. D0 data versus SHERPA prediction are shown. Figure from [16].

3.5 Michelangelo Mangano matching prescritpion

An alternative prescription has been proposed by M. Mangano in [3].

The pourpose is to build up an inclusive event sample summing up "exclusive" event samples with different *m-clusters* multiplicities. Clusters are just partons clustered toghether according to some arbitrary jet finding algorithm and doesn't need to be identified with experimental jets, once the event sample is built the user can apply any kind of analysys to the resulting events.

To produce an event sample with *m-clusters*

- produce a sample of unweighted *partonic* events with

  $$P_T > P_{Tgen} \qquad \Delta R_{j_1 j_2} > R_{gen} \qquad |\eta| < \eta_{gen}$$

  Notice that, in principle, p_{Tgen}, R_{gen} and η_{gen} *are not parameters of the matching prescription.* One should generate completely inclusive (no cuts at the generation level) and, once the matching step is performed, unwanted kinematics configuration will be rejected. However this is not possible since it will lead to null unweighting efficency and therefore one has to find a satisfactory balance: generation cuts should be soft enough to avoid edge effects and hard enough to obtain a good unweigthing efficency.

- for the given kinematic configuration a PS-like branching tree is obtained clustering (see fig. 3.5) the final partons according to $k_\perp$ [25, 26, 27] algorithm until when the LO process is obtained. Namely for $pp \to t\bar{t}+m$-jets cluster until when $pp \to t\bar{t}$ is reached, for $pp \to W + m$-jets cluster until when $pp \to W$ is reached, ... Then at each "branching" assign the proper $\alpha_S(k_\perp)$ factor. In this way the ME is reweighted to mimic more closely the PS weight. This step is the same as in CKKW algorithm.

- perform the shower and merge toghether the obtained partons (ME + PS) to reconstruct cluster of partons according to a jet finding algorithm, see figg. 3.1 and 3.2 where small arcs denote clusters. In `ALPGEN` Paige's `GETJET` algorithm is used. The minimum jet transverse momentum p_{Tmin}, and separation R_{min} toghether with maximum rapidity η_{max} are *the genuine matching parameters.* Notice that, to avoid edge effects due to the smearing of jet momenta induced by the shower, matching parameters should be *harder* than generation cuts

  $$p_{Tmin} > p_{Tgen} \qquad R_{min} > R_{gen} \qquad \eta_{max} < \eta_{gen}$$

  The larger the difference the smaller the edge effects and the unweighting efficency. Actually for η_{max} there is an additional subtlety to be discussed later.

- now reject the event if the number of clusters is not equal to the number of ME generated partons. These events will be generated in other event samples with different parton multiplicities and this prescription avoids double counting. Notice that by performing the PS till the very end and applying the rejection criteria to the final PS generated partons we achieve, at least in the limit of no cuts at the generation level, a net separation among PS and ME generated events, indeed there is no chance that the same event can be generated by ME with different multiplicities.

- if the number of cluster is equal to the number of ME generated partons define the matching of a *parton* and a *cluster* as follows. A parton matches a cluster is the relative separation is smaller than R_{min}, namely if the parton is inside the jet cone. If more than one parton matches the same cluster (collinear ME partons) or if a parton doesn't match to any cluster (soft ME partons), reject the event. With this prescription we avoid double counting and we *reweight* the ME with the appropriate *Sudakov form factor* [6]. Indeed (with this prescription toghether with the requirement imposed at the previous step) a ME "event" will be accepted according to the probability that the PS doesn't emit any "hard" (above the chosen resolution) radiation [7]. An important point has to be noticed here regarding η_{gen} and η_{max}. We have already noticed that to avoid edge effects we should have $\eta_{gen} > \eta_{max}$. There is an additional subtlety here. If one is not inclusive in η_{max} we don't obtain the proper Sudakov form factor. This is due to the fact that, not being inclusive in η_{max} we reweight the ME with the probability that the PS doesn't produce any hard emission *inside the given rapidity range.* This probability, with shrinking rapidity range, obviously approaches one rather than the Sudakov form factor which we wish. Therefore strictly speaking both η_{gen} and η_{max} should go to ∞. Taking smaller values increases the unweighting efficency and again the actual choice is a matter of balance among

[6] Actually a residual infrared sensitivity is left: a soft partons might accidentally fall inside the cone of a cluster originated from a hard PS emission. This is suppressed by the small avaliable phase space and in the studies performed insofar we havn't found any appreciable effect even pushing the generation cuts close the soft/colliner PS cut-off.

[7] Actually the prescription overestimate the Sudakov form factors: two "soft" partons can be clustered even if they can't be traced back to a single splitting. If the resulting cluster is hard enough the event is vetoed. The lower multiplicity sample will not return this PS history it will simply return the contribution of the production and subsequent splitting of the hard parton. This is again a phase space suppressed Log term.

the increasing efficency and the increasing systematic effects. Notice that whereas for p_{Tgen} and R_{gen} we are indeed *forced* to choose non zero values to avoid null unweighting efficency, for η_{gen} there is actually a natural maximum allowed value once p_{Tgen} is chosen and therefore, at least in principle, it's possible to avoid completely this problem.

- the cross section of the event sample is simply the input, parton level, cross section times the ratio between the number of accepted events and the total number of processed events.
- we repeat the above steps for ME with 0 up to ∞ light quarks and jets and we sum up the various event samples
- actually, since it is impossibile to compute ME with an arbitrary number of legs, we shall stop at a definite number n_{max} of light quarks or gluons ($n_{max} = n_{light\ quarks} + n_{gluons}$). For the corresponding matrix element the matching procedure has to be modified, to define an *inclusive event sample* (see fig. 3.8), as follows

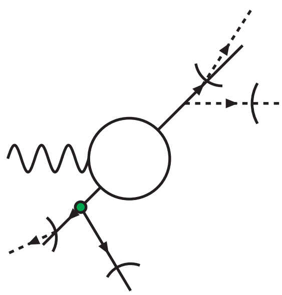


Figure 3.8: Same symbols (and process) as in fig. 3.1. Hard emissions from ME and one hard emission from PS, the number of reconstructed clusters will be greater than the number of ME partons. This event will be retained only in inclusive samples, namely events initiated by ME with the highest particle multiplicity.

1. events with a number of reconstructed clusters *equal or larger* than n_{max} are accepted.
2. events are accepted only if n_{max}-ME partons match the hardest m-clusters (ordinated by p_T).

The major advantage of the above prescription is to be independent from the PS algorithm which is employed and to require minimal interaction with the PS code itself: it is enough to have access to the final partonic configuration after the shower.

From a theoretical poin of view it has the disadvantage that a clean classification of the Logarithmic structure accounted for or missed is very hard. It's hard to work out a closed analytical form for the "Sudakov" reweighting imposed by the algorithm and ultimately it rests on the empirical evidence provided by the smooth behaviour of the distributions and their (in)dependence from the matching parameters. On the other hand it has the advantage that the "Sudakov" reweighting is borrowed from the PS: assuming that indeed this is done exactly (a strong and undemonstrated assumption), this would be the best possible recipe. In fact if, for the given kinematical configuration, the PS reproduces correctly the divergent structure of the ME the two descriptions will merge correctly, otherwise it will be anyway impossible to achieve simultaneously a correct infrared/collinear damping and a smooth interpolation among PS and ME description.

3.5.1 Implementation and comparisons with Tevatron data

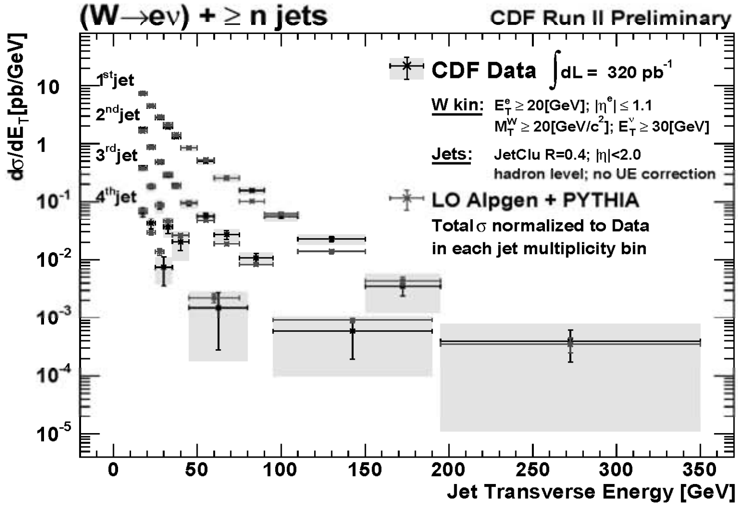
The algorithm described in the previous section is implemented into the ME event generators `ALPGEN` [6], `HELAC` [8] and `MADGRAPH` [7].

A fairly extensive exploration of the matching prescription, for the case of $t\bar{t}$+jets production is reported in [14]. The overall consistency looks good, the prediction is stable against sizable variations of the matching parameters and also the comparison with `MC@NLO` description is good, once the appropriate K-factor rescaling is imposed.

The prescrition has also been tested against Tevatron data mostly looking at jets productions.

CDF has looked at jets production [17] in Drell-Yan processes finding a satisfactory agreement between data and `ALPGEN` +`PITHYA` predictions, once MC predictions are normalized to the data. Preliminary results are shown in fig. 3.9 (left panel, from [17]). Once the overall normalization is fitted to data also jet multiplicities are well reproduced as shown in fig. 3.9 (right panel from [18]).

D0 collaboration has studied [16] dijet azimutal correlation in pure jet sample and compared DATA to ALPGEN+PITHYA predictions predictions, again finding good agreement as shown in fig. 3.10.



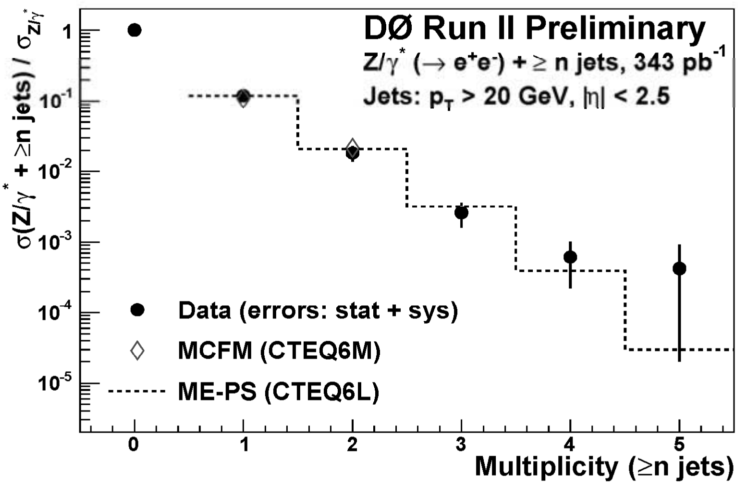


Figure 3.9: **Left:** differential cross section $d\sigma(W \to e\nu+ \geq n-\text{jets})/dE_T^{jet}$ (Right) for the first, second, third and fourth inclusive jet sample. Data are compared to `Alpgen`+`PYTHIA` predictions normalized to the measured cross section in each jet multiplicity sample. **Right:** Measured cross section for Z+jets production as a function of inclusive jet multiplicity compared to MADGRAPH + PITHYA. Absolute cross section normalized to data.

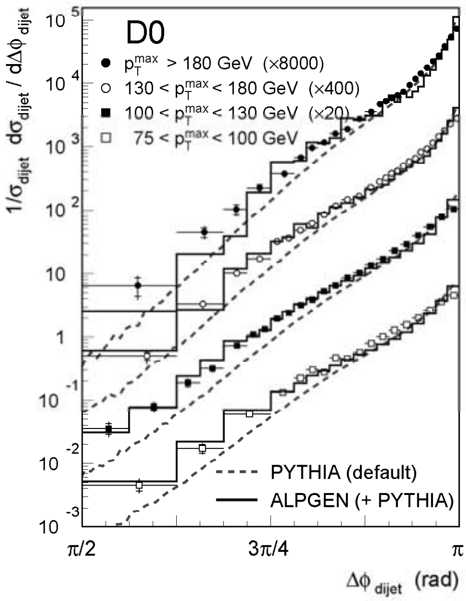


Figure 3.10: Angular separation $\Delta\Phi$ in the transverse plane among the two leading jets (jets ordered in p_T) in inclusive two jets sample. D0 data versus ALPGEN + PITHYA prediction are shown. Figure from [16].

3.6 Comparison among matching prescriptions

We refer to [10, 20] for a more complete account of a detailed series of comparisons.

These comparisons have been performed for the Drell-Yan process both at the Tevatron and at the LHC.

The overall agreement is relatively good as shown in fig. 3.11 (from [20]) and the differencies are compatible with the effect of the factorization scale variation for a LO calculation. ARIADNE exhibits larger variation mostly due to the different approach to the shower evolution.

In [20] it is also provided an extimate of (at least some of) the systematic uncertainties associated to the approach varying α_S scale and, for SHERPA and ARIADNE, also the PS scale. In fig. 3.12, from [20], we show an example of this exploration for MADEVENT.

We address to [10, 20] for a more thorough discussion. Here we want just make a few remarks

- As step zero, to gain confidence on an event sample, one should first investigate the dependence on the resolution parameters looking at the impact of moving away from the various codes default setting. We emphasize once again that this is the *only way to extimate* this dependence since we lack an analitical extimate.
- Some of the differencies among the various recipes can be minimized adjusting the resolution parameters and/or α_S scale. This doesn't make much sence in the absence of data. However once data are avaliable all these parameters provide an handle to improve the description of data.
- Having performed step zero one should also move to step one: investigating the impact of scale variation on the prediction (expecially to assess the impact on the *shapes* of the various observables).
- As a final remark let's outline that if one is interested in a fairly exclusive region of phase space one should repeat the above steps for the region of interest: *an overall stable and satisfactory picture for l^+l^- production doesn't gurantee that the same holds in the hard mass tail, say $m_{l^+l^-} > 1$ TeV.*

References

1. S. Catani, F. Krauss, R. Kuhn, and B. R. Webber, JHEP**0111**, 063 (2001).
2. F. Krauss, JHEP**0208**, 015 (2002).

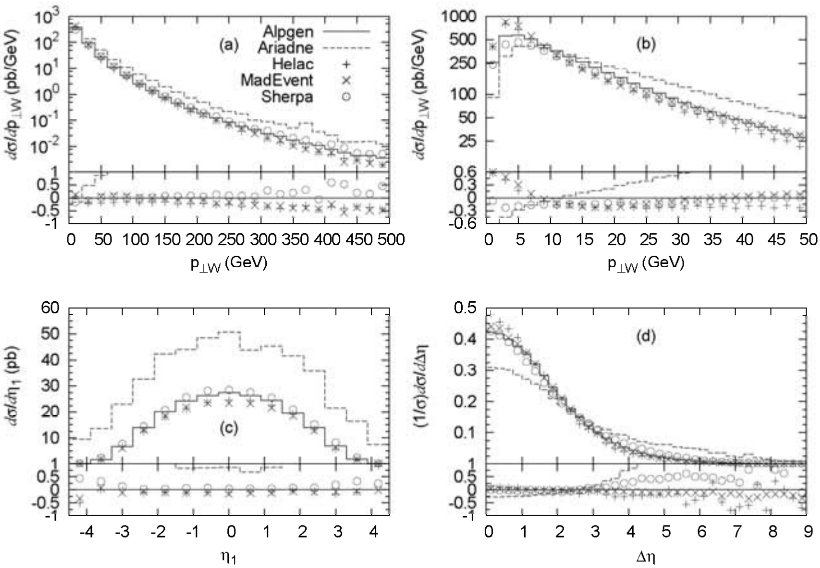


Figure 3.11: (a) and (b) $p_\perp$ spectrum of W^+ bosons at the LHC (pb/GeV). (c) η spectrum of the leading jet, for $p_\perp^{jet1} > 100$ GeV; absolute normalization (pb). (d) Pseudo-rapidity separation between the W^+ and the leading jet, $\Delta\eta = |\eta_{W^+} - \eta_{\rm jet1}|$, for $p_\perp^{jet1} > 40$ GeV, normalized to unit area. In all cases the full line gives the ALPGEN results, the dashed line gives the ARIADNE results and the "+", "x" and "o" points give the HELAC, MADEVENT and SHERPA results respectively. In the lower frame relative deviation with respect to ALPGEN predictions are shown.

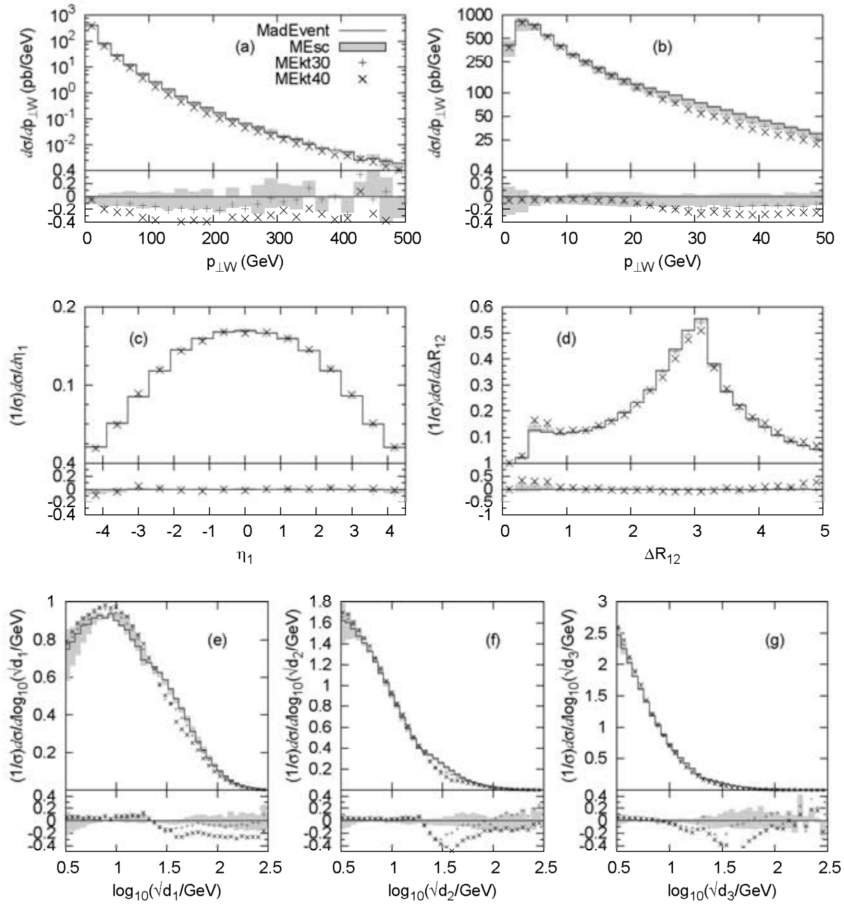


Figure 3.12: MADEVENT systematics at the LHC. (a) and (b) show the $p_\perp$ spectrum of the W, (c) shows the pseudorapidity distribution of the leading jet, (d) shows the ΔR separation between the two leading jets, and (e)–(g) show the d_i $(i = 1, 2, 3)$ spectra, where d_i is the scale in a parton-level event where i jets are clustered into $i - 1$ jets using the $k_\perp$-algorithm. The full line is the default settings of MADEVENT, the shaded area is the range between MEscL and MEscH, while the points represent MEkt30 and MEkt40.

3. M.L. Mangano, presentation at the FNAL Matrix Element/Monte Carlo Tuning Working Group, 15 Nov 2002, http://www-cpd.fnal.gov/personal/mrenna/tuning/nov2002/mlm.pdf .

4. T. Gleisberg, S. Höche, F. Krauss, A. Schälicke, S. Schumann, and J. Winter, JHEP**0402**, 056 (2004).

5. A. Schälicke and F. Krauss, JHEP**07**, 018 (2005).

6. M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau and A. D. Polosa *JHEP* **07** (2003) 001

7. T. Stelzer and W. F. Long *Comput. Phys. Commun.* **81** (1994) 357 F. Maltoni and T. Stelzer *JHEP* **02** (2003) 027

8. A. Kanaki and C. G. Papadopoulos *Comput. Phys. Commun.* **132** (2000); A. Cafarella, C. G. Papadopoulos and M. Worek, e-Print: arXiv:0710.2427 [hep-ph]

9. F. Krauss, A. Schalicke and G. Soff, Comput. Phys. Commun. **174** (2006) 876

10. S. Mrenna and P. Richardson, JHEP**05**, 040 (2004).

11. F. Krauss, A. Schalicke, S. Schumann and G. Soff, Phys. Rev. D **72** (2005) 054017

12. F. Krauss, A. Schalicke, S. Schumann and G. Soff, Phys. Rev. D **70** (2004) 114009

13. T. Gleisberg, F. Krauss, A. Schalicke, S. Schumann and J. C. Winter, Phys. Rev. D **72** (2005) 034028

14. M. L. Mangano, M. Moretti, F. Piccinini, M. Treccani; JHEP 0701:013,2007.

15. *see for example,* D0 collaboration (Jeroen Hegeman for the collaboration, *Jet Production in the D0 experiment*, in QCD06, Montpellier July 2006.

16. *see for example,* By D0 Collaboration (Marek Zielinski for the collaboration). Published in AIP Conf.Proc.842:265-267,2006. Also in *Santa Fe 2005, Particles and nuclei* 265-267
e-Print: hep-ex/0602019

17. *see for example,* CDF collaboration, (Andrea Messina for the collaboration), Braz.J.Phys.37:840-842,2007.
e-Print: arxiv:0708.1380 [hep-ex]

18. *see for example*; D0 collaboration (Gavin Hesketh for the collaboration), Proceedings of the XLI Rencontres de Moriond, QCD and Hadronic Interactions.

19. S. Höche *et. al.* [http://arXiv.org/abs/hep-ph/0602031hep-ph/0602031].

20. J. Alwall et al. Eur.Phys.J.C53 (2008) 473.

21. L. Lönnblad *Comput. Phys. Commun.* **71** (1992) 15–31.

22. L. Lönnblad *JHEP* **05** (2002) 046

23. N. Lavesson and L. Lönnblad *JHEP* **07** (2005) 054

24. F.E. Paige and S.D. Protopopescu, in *Physics of the SSC*, Snowmass, 1986, Colorado, edited by R. Donaldson and J. Marx.

25. S. Catani, Y. L. Dokshitser, M. Olsson, G. Turnock, and B. R. Webber, Phys. Lett. **B269**, 432 (1991)

26. S. Catani, Y. L. Dokshitser, and B. R. Webber, Phys. Lett. **B285**, 291 (1992)

27. S. Catani, Y. L. Dokshitser, and B. R. Webber, Nucl. Phys. **B406**, 187 (1993)

28. S. Frixione and B. R. Webber, JHEP **0206** (2002) 029