An Analytic Initial-State Parton Shower

Wolfgang Kilian,[a] **Jürgen Reuter,**[b] **Sebastian Schmidt**[b] **and Daniel Wiesler**[b]

[a] *Department Physik, Universität Siegen, D–57068 Siegen, Germany*

[b] *DESY Theory Group, Notkestrasse 85, D–22607 Hamburg, Germany*

E-mail: kilian@physik.uni-siegen.de, juergen.reuter@desy.de, sebastian.t.schmidt@desy.de, daniel.wiesler@desy.de

ABSTRACT: We present a new algorithm for an analytic parton shower. While the algorithm for the final-state shower has been known in the literature, the construction of an initial-state shower along these lines is new. The aim is to have a parton shower algorithm for which the full analytic form of the probability distribution for all branchings is known. For these parton shower algorithms it is therefore possible to calculate the probability for a given event to be generated, providing the potential to reweight the event after the simulation. We develop the algorithm for this shower including scale choices and angular ordering. Merging to matrix elements is used to describe high-energy tails of distributions correctly. Finally, we compare our results with those of other parton showers and with experimental data from LEP, Tevatron and LHC.

Contents

1 Introduction

Parton Showers are an indispensable part in the simulation of hadronic collisions in today's high-energy colliders, like the now-running LHC at CERN. A precise simulation of these collisions demands a coordinated interplay of such diverse elements as the calculation of hard matrix elements, the simulation of parton showers, a matching procedure to combine matrix elements and parton shower, the modeling of the underlying event and the hadronization and the simulation of the detector response. The main task of parton showers in this framework is to describe collinear and soft emissions off incoming or outgoing partons from the hard matrix element, thereby affecting the jet substructure and possibly increasing the number of resolved jets. The parton shower implementation most commonly used is the one in PYTHIA [1, 2]. Alternatively, programs such as Herwig++ [3] and SHERPA [4] implement parton shower algorithms as components for their event generation frameworks. Moreover programs that are tailored exclusively to parton showers and work as plug-ins to event generators are available, the most prominent being Vincia [5].

Recently it was shown that the formulation of parton showers can be reproduced by the soft-collinear effective field theory [6]. This led to the formulation of so-called *analytic parton showers* [7]. In this work, we intend to use the analytic parton shower for event generation. We extend the analytic final-state parton shower presented by Bauer et al. [8] and develop a new analytic parton shower for initial-state radiation. During this work, we implemented both these parton showers in the event generator WHIZARD [9]. We succeeded in reproducing event shape distributions simulated using PYTHIA's parton shower, as well as distributions measured at LEP, Tevatron and LHC very well, given that very little tuning was done.

Section 2 describes the theory of the parton shower derived from the analytic approach, as well as the extensions and improvements implemented in the final-state parton shower and the newly implemented parton shower for the initial state. Section 3 describes the steps to be taken for a comparison of the implementation of our algorithms to data and other showers. These include the implementation of a MLM-type matching, described section 3.1, an interface to an external hadronization routine, in our case PYTHIA, in section 3.2 and the handling of beam-remnants, described in section 3.3. In section 4, we show distributions obtained using our shower or PYTHIA's shower for various event and jet shapes. Furthermore we compare those to the respective measurements at LEP, Tevatron and LHC. Finally, in section 5 we summarize our findings, conclude and give an outlook on future developments.

2 The Analytic Parton Shower - Introduction and Algorithm

2.1 Concept of Parton Showers

Parton showers are commonly formulated using branchings of one particle into two, which can either be considered as one parton splitting into two partons or one parton emitting a new parton[1]. The central entity of the parton shower is the Sudakov form factor Δ – originally described in [11] – its simplest form is given by

$$\Delta(t_1, t_2) = \exp\left[\int\limits_{t_1}^{t_2} dt \int\limits_{z_-}^{z_+} dz \frac{\alpha_s}{2\pi t} P(z)\right] \tag{2.1}$$

giving the probability for a parton to evolve from scale t_2 to t_1 without emitting a further parton. Therein, the variable z describes the relation of the two partons after the branching, the most common choice is to take z to be the ratio of one parton's energy after the branching and the parton's energy prior to the branching. The functions $P(z)$ are called splitting functions. They describe the probabilities for the respective branchings and can be inferred from approximate calculations of matrix elements. The variable t is called the scale. There is a certain ambiguity how to define the scale in an implementation. Here we use the virtuality that is defined as the square of the four-momentum $t = p^2$. The simulation of the parton shower is an evolution in the scale. For final-state radiation, the branchings that occur after the hard interaction, the evolution is from a scale corresponding to the hard interaction, $t \sim \hat{s}$, down to a cut-off scale $t = t_{cut}$, that symbolizes the transition to the non-perturbative physics encapsulated in the hadronization. For initial-state radiation[2], branchings that appear before the hard interaction, the evolution is from a cut-off $t = -t_{cut}$, representing the factorization scale and thus the parton density functions, down to a scale corresponding to the negative of the center of mass energy, $t \sim -\hat{s}$[3]. In the implementation, this evolution, that corresponds to an evolution in physical time, is replaced by an evolution starting at the hard interaction and ending at the cut-off scale. This is the so-called backwards evolution, originally described first in [12], its most prominent consequence being the appearance of parton density functions in the Sudakov factor.

2.2 The Analytic Parton Shower

The parton shower is a well-defined approximation to the full matrix element. Therefore it should be preferable to be able to reconstruct the matrix element from the parton shower.

---

[1] The exception to this are the showers based on the dipole/antenna picture and their formulation using two to three splittings, originally developed in [10].

[2] For now, only initial-state radiation for partons stemming from hadrons are considered.

[3] This treatment is only correct for s-channel resonance decays, like the processes considered in this paper. For general processes a more sophisticated ansatz will be implemented in the future.

In common parton shower algorithms, this ability is lost due to the formulation of the parton shower as a Markov chain in such a way, that branchings that fail to respect correct kinematics can be produced and are subsequently rejected or manually modified to respect momentum conservation. It is these branchings that prevent the probability for a branching to be calculated analytically after the branching is generated. Therefore in developing an analytic parton shower, care was taken to avoid branchings that need to be rejected or manually modified, thereby preserving the ability to reconstruct the matrix element. The two main changes are the simultaneous simulation of the branchings of sisters and replacing the splitting variable z, that normally is the ratio of the first daughter's energy or light-cone momentum to the mothers' energy or light-cone momentum.

The first modification is to replace the simulation of individual branchings by the simulation of so-called double branchings. A double branching consists of the simultaneous branching (or no-branching) of the two daughter-partons of one parton (cf. figure 1). So instead of taking one parton a and letting it branch into two partons, $a \rightarrow bc$, an existing branching $a \rightarrow bc$ is replaced by the double branching $a \rightarrow bc \rightarrow defg$ with the new partons d, e, f, g in case both partons b and c branch. The corresponding situations where one or both of the daughters do not branch are also taken into account. The advantage is that the energy-conservation[4]

$$\sqrt{t_a} \geq \sqrt{t_b} + \sqrt{t_c} \tag{2.2}$$

can be included in the generation of the branchings, avoiding the production of complicated interconnections between different single branchings. The sequence of steps is then

- Pick a branch with unprocessed daughters b and c.

- Generate $\{t_b, v_b\}$ and $\{t_c, v_c\}$ for both daughters independently with the probability given by the single branching probability. (v_i stands for the values needed to describe the branching apart from the virtuality t_i, like the opening angle $\cos \vartheta$, the azimuthal angle ϕ and the type of the daughter parton.)

- Keep the branch of the daughter with the higher scale $t_{max} = \max (t_b, t_c)$. Discard the branching of the other daughter.

- Determine new values for the other daughter with the maximum scale set to $t_* = \min \left[t_{max}, \left(\sqrt{t_a} - \sqrt{t_{max}} \right)^2 \right]$.

For the different cases, the double branch probabilities can be constructed from the single branching probabilities [7] $\mathcal{P}_i^{br}(t_i, v_i)$ and Sudakov factors $\Delta_i(t_a, t)$ for a branching at the

---

[4]The following equation can best be understood in the rest frame of the mother. Then $\sqrt{t_a}$ is the mother's mass and energy and trivially the sum of the daughters' masses $\sqrt{t_b} + \sqrt{t_c}$ has to be less.

scale t_i and the remaining values v_i and the probability $\mathcal{P}_i^{nb}$ for no branching above the cut-off. The double branch probabilities for the case in which both daughters branch is

$$\mathcal{P}^{br,br}(t_b, v_b, t_c, v_c) = \theta(t_b - t_c)\mathcal{P}_b^{br}(t_b, v_b)\Delta_c(t_a, t_b) \tag{2.3}$$
$$\mathcal{P}_c^{br}(t_c, v_c; t_a = t_*)$$
$$+ \theta(t_c - t_b)\mathcal{P}_c^{br}(t_c, v_c)\Delta_b(t_a, t_c)$$
$$\mathcal{P}_b^{br}(t_b, v_b; t_a = t_*),$$

while in the case that only one daughter branches

$$\mathcal{P}^{br,nb}(t_b, v_b) = \mathcal{P}_b^{br}(t_b, v_b)\Delta_c(t_a, t_b)\Delta_c(t_*, t_{cut}) \tag{2.4}$$
$$\mathcal{P}^{nb,br}(t_c, v_c) = \mathcal{P}_c^{br}(t_c, v_c)\Delta_b(t_a, t_c)\Delta_b(t_*, t_{cut}), \tag{2.5}$$

and in the case no parton branches

$$\mathcal{P}^{nb,nb} = \Delta_b(t_a, t_{cut})\Delta_c(t_a, t_{cut}). \tag{2.6}$$

Taking all different combinations into account, the double branch probability can be composed in the following way

$$\mathcal{P}(t_b, v_b, t_c, v_c) = \mathcal{P}^{br,br}(t_b, v_b, t_c, v_c)$$
$$+ \mathcal{P}^{br,nb}(t_b, v_b) \quad \delta(t_c)$$
$$+ \mathcal{P}^{nb,br}(t_c, v_c) \quad \delta(t_b)$$
$$+ \mathcal{P}^{nb,nb} \quad \delta(t_b)\delta(t_c). \tag{2.7}$$

The second step is replacing the kinematic ratio z with the angle $\cos\theta$ in the mother's rest frame between the momentum of the first daughter and the boost axis. This leads to simple phase space limits

$$-1 \leq \cos\theta \leq 1.$$

There is a direct correspondence between the $\cos\theta$ angle and the energy splitting z [7] as a function of the masses of the daughters t_b and t_c:

$$z = \frac{1}{2}\left[1 + \frac{t_b}{t_a} - \frac{t_c}{t_a} + \beta_a \cos\theta_a \lambda(t_a, t_b, t_c)\right] \tag{2.8}$$

with the boost β_a and the phase space factor λ:

$$\beta_a = \sqrt{1 - \frac{t_a}{E_a^2}} \quad \text{and} \quad \lambda(t_a, t_b, t_c) = \frac{1}{t_a}\sqrt{(t_a - t_b - t_c)^2 - 4t_b t_c}$$

The important distinction between common and analytic parton showers is that in the analytic parton shower every branching is generated with a calculable probability. Every source for vetoing branchings where the probability for the veto cannot be calculated has therefore been avoided.
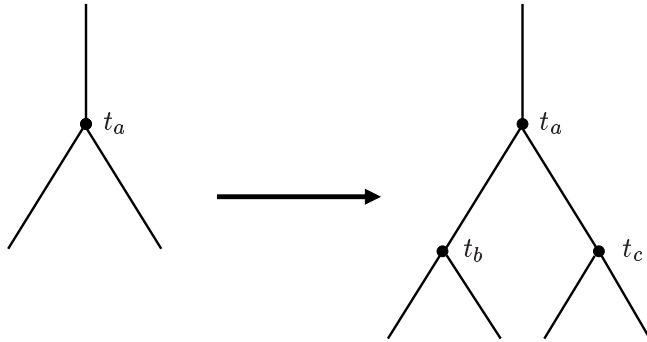
Figure 1: Schematic view of a double branching. Before the double branching (left): A parton has branched at a scale t_a into two on-shell daughter partons. After the simulation of the double branching the branching scales t_b, t_c for the daughter partons are known. In case the daughters branch themselves the needed values are generated as well. The case in which both daughters branch, $t_b > t_{cut}, t_c > t_{cut}$, is shown on the right. For the next step the double branchings of the two daughter partons will be simulated, the branchings at t_b and t_c, respectively, replace the parton branching at t_a on the left.

2.3 Improved Analytic Final State Parton Shower

The `GenEvA` framework [8, 13] is an event generation framework designed to combine matrix elements and parton showers during event generation. It uses its parton shower to distribute events over phase space, in order to reweight them to a corrected distribution later. Therefore only a simplified implementation of parton showers was included in the framework, as the reweighting would later reintroduce the correct distributions. We, on the other hand, will use the analytic parton shower to generate physical events and cannot defer anything to a reweighting procedure, we are therefore forced to implement the full theory of parton showers. The two main simplifications made in the `GenEvA` framework are the omission of the running of the coupling constant and the omittance of color coherence. Our extensions to the parton shower are as follows.

The running of the coupling constant $\alpha_S(Q^2)$ was implemented, the inclusion is straightforward. The coupling was chosen to be

$$\alpha_S = \alpha_S \left(z(1-z)Q^2 \right) = \alpha_S(z(\theta), Q^2) \tag{2.9}$$

in agreement with most parton shower generators.

As color coherence is approximated by demanding that the angles of subsequent emissions decrease – this is known as angular ordering – the resulting phase space cuts have to be implemented in the parton shower. The opening angle[5] $\cos\vartheta$ is given by

$$\cos\vartheta = 1 - \frac{t}{2z(1-z)E^2} \tag{2.10}$$

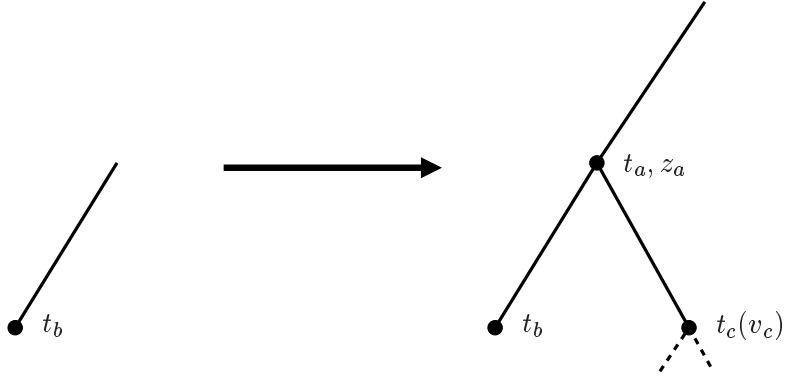[5] not to be confused with the angle $\cos\theta$ used in the description of branchings

Figure 2: Schematic view of a double branching in ISR: Before the double branching (left): At the scale t_b a parton b exists. For this parton, the scale t_a of the branching that produced this parton, the corresponding energy ratio z_a and the scale t_c where the emitted parton c branches, and if necessary the remaining quantities (v_c) are simulated (on the right).

in the approximation for massless children. Using $z(1-z) \leq \frac{1}{4}$ this can be used to give a cut on the scale of a next branching

$$t \leq E^2 \frac{1 - \cos \vartheta_{cut}}{2} \tag{2.11}$$

for the branching to have an opening angle less than $\cos \vartheta_{cut}$. An additional cut on z [7] is necessary

$$\left| z - \frac{1}{2} \right| \leq \frac{\beta}{2} \sqrt{1 - \frac{t}{\beta^2 E^2} \frac{1 + \cos \vartheta_{cut}}{1 - \cos \vartheta_{cut}}}. \tag{2.12}$$

With these phase space cuts angular ordering is enforced in the approximation of massless daughter partons. However the inclusion of this constraint demands keeping track of the used energy E and the used angle $\cos \vartheta_{cut}$ either by explicitly storing their values for every branching or by using a distinct rule to calculate them for every branching.

As a minor extension we allow for parton masses, although these are only taken into account when distributing momenta, the splitting functions are still taken for massless daughter partons.

2.4 Introducing the Analytic Initial State Parton Shower

For physics at the LHC, a parton shower has to be able to describe both, initial and final-state radiation. We therefore implement an initial-state parton shower satisfying the requirement of analyticity analogously to the parton shower for the final state.

The changes applied to the final-state shower cannot be transferred to the initial-state shower[6]. A different set of changes is needed to reformulate the initial-state parton shower

---

[6] Due to the negative virtualities all momenta and energies would be imaginary in the mothers restframe.

in order to fulfill the demand of analyticity. The known Sudakov factor for initial-state radiation is

$$\Delta_b^{ISR}(t_a, t_b) = \exp\left[-\int\limits_{|t_a|}^{|t_b|} \mathrm{d}t' \int\limits_0^1 \frac{\mathrm{d}z}{z} \frac{\alpha_S}{2\pi t'} \sum_{a,c} \left(P_{a\to bc}(z) + P_{a\to cb}(z) \right) \frac{f_a\left(\frac{x_b}{z}, t'\right)}{f_b(x_b, t')} \right] \quad (2.13)$$

with the splitting function $P_{a\to bc}(z)$ for a parton of type a branching into two partons of type b and c and the parton density functions $f_a(x, t)$. The conservation of momentum can be enforced by explicitly vetoing momentum conservation-violating branchings directly in the Sudakov factor. To do this, the branching of the mother parton and the branching of the emitted parton have to be simulated simultaneously, cf. figure 2. The simulated branching therefore effectively becomes a $1 \to 2$ (if the emitted parton does not branch) or a $1 \to 3$ (if the emitted parton branches) branching. The Sudakov factor that takes the emitted parton's branching into account can be written in the form

$$\Delta_b^{ISR}(t_a, t_b) = \exp\left[-\int\limits_{|t_a|}^{|t_b|} \mathrm{d}t' \int\limits_0^1 \frac{\mathrm{d}z}{z} \frac{\alpha_S}{2\pi t'} \sum_{a,c} \int\limits_0^{t'} \mathrm{d}t_c\, \mathcal{P}_c(t_c| - t', z) \right.$$

$$\left. \bar{\Theta}\left(-t', t_b, t_c, z_a, E_a \right) \left(P_{a\to bc}(z) + P_{a\to cb}(z) \right) \frac{f_a\left(\frac{x_b}{z}, t'\right)}{f_b(x_b, t')} \right] \quad (2.14)$$

with the veto function

$$\bar{\Theta}\left(t_a, t_b, t_c, z_a, E_a\right) = \Theta\left(|\vec{p}_b| + |\vec{p}_c| - |\vec{p}_a|\right) \cdot \Theta\left(|\vec{p}_a| - ||\vec{p}_b| - |\vec{p}_c||\right), \quad (2.15)$$

and the one parton branching distribution function for the emitted parton c

$$\mathcal{P}_c(t_c| - t', z), \quad (2.16)$$

giving the probability distribution for the branching of the emitted timelike parton as a function of the branching this parton was produced, described by $-t'$ and z. The veto function ensures that the three partons a, b, c can be combined in a branching that conserves momentum by enforcing the triangle inequality. By adding more terms it can also be used to impose cuts for angular ordering or a minimum energy for the emitted timelike parton. If the emitted parton branches, its final-state parton shower can now be simulated further by the use of the known double branching probabilities from the analytic final-state radiation.

However there is a slight difference in the interpretation of the known one-branching Sudakov factor as used for example in PYTHIA and the supplemented one in equation (2.14). In the former, the probability for a branching is independent of the available allowed branchings of the emitted parton, while in the latter the probability for a branching is reduced when the emitted parton has a restricted phase space for branchings. Therefore the supplemented Sudakov factor rather resembles a conditional probability.

Using these prescriptions, double branch probability distributions can be formulated, analogously to the ones formulated for final-state radiation. The probability for no earlier branching, the parton being directly emitted by the hadron and therefore being on-shell, $t_a \to m_a^2$, consists of the Sudakov factor $\Delta_b^{ISR}(-t_{cut}, t_b)$ and a δ-distribution forcing the parton to be on-shell and thus can be formulated in the form

$$\mathcal{P}_b^{nb}(t_a; t_b, t_{cut}) = \Delta_b^{ISR}(-t_{cut}, t_b)\delta\left(t_a - m_a^2\right). \qquad (2.17)$$

In case an earlier branching is found, the common single branch probability would be

$$\begin{aligned}
\mathcal{P}_{a \to bc}(t_a, z_a; t_b, t_{cut}) = {} & \frac{\alpha_S}{2\pi t_a}\frac{1}{z_a}P_{a \to bc}(z_a)\frac{f_a\left(x_a, t_a\right)}{f_b\left(x_b, t_a\right)} \\
& \cdot \Delta_b^{ISR}(t_a, t_b)\Theta\left(t_a - t_b\right)\Theta\left(-t_a - t_{cut}\right)
\end{aligned} \qquad (2.18)$$

with the Sudakov factor $\Delta_b^{ISR}(t_a, t_b)$, a relative weight, given by the ratio of parton density functions, $\frac{f_a(x_a, t_a)}{f_b(x_b, t_a)}$, the probability for the branching itself, $\frac{\alpha_S}{2\pi t_a}\frac{1}{z_a}P_{a \to bc}(z_a)$ and two step functions that force the parton to be in the correct range of virtuality.

For the transition to analytic showers, a dependence on the scale of the emitted parton t_c is introduced. Thus two different cases have to be considered. In the case of the emitted parton not branching further, the corresponding probability distribution is supplemented by the no-branching-probability $\mathcal{P}_c^{nb}$ for the emitted parton c and the veto function $\bar{\Theta}$. It can be written in the way

$$\begin{aligned}
& \mathcal{P}_{a \to bc}^{br,nb}(t_a, t_c, z_a; t_b, t_{cut}) \\
& = \frac{\alpha_S}{2\pi t_a}\bar{\Theta}\left(t_a, t_b, t_c, z_a, E_a\right)\frac{1}{z_a}P_{a \to bc}(z_a)\frac{f_a\left(x_a, t_a\right)}{f_b\left(x_b, t_a\right)} \\
& \cdot \Delta_b^{ISR}(t_a, t_b)\Theta\left(t_a - t_b\right)\Theta\left(-t_a - t_{cut}\right)\mathcal{P}_c^{nb}\left(t_c; |t_a|, t_{cut}\right).
\end{aligned} \qquad (2.19)$$

In the case, the emitted parton undergoes another branching, the distribution is supplemented by the single branch probability for the emitted parton $\mathcal{P}_{c \to de}^{br}$ and has the form

$$\begin{aligned}
& \mathcal{P}_{a \to bc \to bde}^{br,br}(t_a, t_c, z_a, v_c; t_b, t_{cut}) \\
& = \frac{\alpha_S}{2\pi t_a}\bar{\Theta}\left(t_a, t_b, t_c, z_a, E_a\right)\frac{1}{z_a}P_{a \to bc}(z_a)\frac{f_a\left(x_a, t_a\right)}{f_b\left(x_b, t_a\right)} \\
& \cdot \Delta_b^{ISR}(t_a, t_b)\Theta\left(t_a - t_b\right)\Theta\left(-t_a - t_{cut}\right)\mathcal{P}_{c \to de}^{br}\left(t_c, v_c; |t_a|, t_{cut}\right).
\end{aligned} \qquad (2.20)$$

Using these expressions the probability distribution for the scale t_a can therefore be

written analogously to equation (2.7) in the form

$$
\begin{aligned}
\mathcal{P}_b(t_a; t_b, t_{cut}) &= \mathcal{P}_b^{nb}(t_a; t_b, t_{cut}) \\
&+ \sum_{a,c} \int dz_a \int dt_c\, \mathcal{P}_{a \to bc}^{br,nb}(t_a, t_c, z_a; t_b, t_{cut}) \\
&+ \sum_{a,c} \int dz_a \int dt_c\, \mathcal{P}_{a \to cb}^{br,nb}(t_a, t_c, z_a; t_b, t_{cut}) \\
&+ \sum_{a,c,d,e} \int dz_a \int dt_c \int dv_c\, \mathcal{P}_{a \to bc \to bde}^{br,br}(t_a, t_c, z_a, v_c; t_b, t_{cut}) \\
&+ \sum_{a,c,d,e} \int dz_a \int dt_c \int dv_c\, \mathcal{P}_{a \to cb \to deb}^{br,br}(t_a, t_c, z_a, v_c; t_b, t_{cut}).
\end{aligned}
\tag{2.21}
$$

The probability distributions for the parton species and the energy fractions z follow directly from this equation.

One aspect of initial-state parton showers that is a rather critical technical point, is the assignment of momenta for the first branchings in the initial state. By the first branchings we mean the respective branchings closest to the hard interaction for the two incoming partons in the matrix element. As these partons are on the mass-shell and often assumed to be massless, any branching would be kinematically forbidden. Therefore the partons have to be set off-shell in order to allow for kinematically allowed branchings. This is done by simultaneously scaling the partons' momenta, until the four-momentum squared reaches the scale of the first branching,

$$
t = p^2 = E^2 - \vec{p}^2 = t_{first} \ll 0.
$$

The distribution of t_{first} is obtained by solving a Sudakov factor similar to the one given in equation (2.14), but with the terms corresponding to the emitted parton removed,

$$
\Delta_b^{ISR}(t_{first}, t_b) = \exp \left[- \int_{|t_a|}^{|t_b|} dt' \int_0^1 \frac{dz}{z} \frac{\alpha_S}{2\pi t'} \sum_{a,c} \right.
$$
$$
\left. \left(P_{a \to bc}(z) + P_{a \to cb}(z) \right) \frac{f_a\left(\frac{x_b}{z}, t'\right)}{f_b(x_b, t')} \right].
\tag{2.22}
$$

Thus the initial-state parton shower is not started from the two partons in the initial state of the matrix element, but from copies of them that have their momenta assigned in the following way:

$$
t_1 = t_{first\,1} \qquad t_2 = t_{first\,2}
\tag{2.23}
$$
$$
E_1 = \frac{\hat{s} + t_1 - t_2}{4\sqrt{\hat{s}}} \qquad E_2 = \frac{\sqrt{\hat{s}}}{2} - E_1
\tag{2.24}
$$
$$
|\vec{p}_1| = |\vec{p}_2| = \sqrt{E_1^2 - t_1} \qquad \vec{p}_1 = -\vec{p}_2
\tag{2.25}
$$

By doing so, both partons are set off-shell so that branchings are kinematically allowed, while conserving the total energy and momentum. Another possibility would be to enlarge the three-momenta so that the scales are equal to the negative partonic center-of-mass energy, $t \to -\hat{s}^2$, and then start the shower from there, but this starting configuration has the disadvantage that the three-momenta of the initial partons tend to be very large, so that it becomes very hard to find kinematically allowed branchings.

3 Prerequisites for a realistic description

In this section we discuss the technical prerequisites for a realistic implementation of our analytic parton shower algorithm and the preparations needed to compare results with results from other parton showers and with experimental data. To do so, we chose to implement the shower algorithm within the event generator WHIZARD [9] which contains highly optimized (tree-level) matrix elements by the matrix-element generator O'Mega [14], a very efficient phase-space parametrization and a multi-channel adaptive Monte-Carlo integration [15]. WHIZARD has been developed and has found a wide range of application to lepton collider physics (cf. e.g. [16–18]). The program has been completely recast for hadron collider physics and been successfully applied to BSM and jet physics (cf. e.g. [19–23]). Several steps to include NLO corrections in a semi-automatic way have been undertaken [24–27].

To make contact with experimental distributions, one has to cover the whole of phase space to access the high-energy tails of distributions. One possibility is using the so-called *power-shower* concept where one artificially opens up more phase space than physically available to generate hard and/or non-collinear jets from the parton shower. We decided not to use this concept, but to implement a matching procedure of matrix elements with explicit additional jets with the showered Born process. This is done in the first part of this section. In the second part we discuss our treatment of hadronization to compare a realistic event simulation with experimental data, while in the last section we explain how beam remnants are dealt with in our framework.

The extensions presented in this section will become publicly available in all future releases from WHIZARD 2.1 on.

3.1 Matching

A matching procedure is a procedure to combine the description of up to a certain number of multiple, widely separated jets by the matrix element and the description of possible additional jets and the substructure of the jets by a parton shower. The main approaches for matching to leading order calculations are the CKKW [28], CKKW-L [29] and MLM [30, 31] schemes, for a general overview see [32–34]. In the process of implementing the analytic

parton shower, we also implemented a matching procedure according to the principles of the MLM approach in `WHIZARD` with the use of the `KTCLUS` clustering package [35].

The steps as implemented in `WHIZARD` are:

1. The cross-sections for the main process and processes with up to N additional partons in the hard matrix-element are calculated. The phase space has to satisfy the additional cuts enforced by the matching procedure

$$p_T > p_{T\,min}, \qquad |\eta| < |\eta_{max}|, \qquad \Delta R_{jj} > R_{min} \tag{3.1}$$

   with the transverse momentum p_T, the pseudo-rapidity η and the $\eta - \phi$-distance between two jets ΔR_{jj}. The values $p_{T\,min}$, η_{max} and R_{min} can be set in the `WHIZARD` input file.

2. According to the relative probability $P(i)$ given by the relative size of the corresponding cross-sections,

$$P(i) = \frac{\sigma_i}{\sum_j \sigma_j} \tag{3.2}$$

   a matrix-element event with i additional partons is generated.

3. These events are then showered with the analytic shower[7].

4. After the shower evolution, a k_T-clustering jet algorithm [35] is applied to the showered, but not yet hadronized event, taking only colored partons with a pseudo-rapidity $|\eta| < \eta_{max\,clus}$ into account. Jets are defined by a minimum jet-jet separation y_{cut}[8]

$$\eta_{max\,clus} = \eta_{clus\,factor} * \eta_{max} \tag{3.3}$$
$$y_{cut} = \Big[p_{T\,min} + \max\left(E_{Tclus\,min\,E}, \right.$$
$$\left. E_{T\,clus\,factor} * p_{T\,min} \right) \Big]^2 \tag{3.4}$$

   The factors and hence the clustering variables can be varied as part of the systematics assessment, the defaults for these factors are chosen to be 1.

5. If the jet algorithm in the matching procedure undershoots the number of matrix element jets the event is discarded. When the event after the jet merging overshoots the number of matrix element jets, the event is rejected as well, unless the number of matrix element jets is equal to the maximum number of matrix element jets. In that case the scale y_{cut} is adapted such that the number of reconstructed jets is reduced to the number of matrix element jets, i.e. the jet resolution is lowered accordingly.

---

[7] As an alternative, one could also use the `PYTHIA` shower.
[8] see section A.2 for a short introduction to jet clustering

6. Then it is tested if the reconstructed jets match the matrix element partons. This is done in an iterative way: The clustering is reapplied to a set consisting of the reconstructed jets and one matrix element parton. If this additional parton leads to an additional jet above the scale y_{cut}, the parton is assumed not to be matched to any of the reconstructed jets and the event is discarded. Otherwise the matched reconstructed jet is removed from the set and this step is repeated for the next matrix element parton. If and only ifall matrix element partons can be matched in this way, the event is accepted.

7. The remaining steps of event generation, like multiple interactions, hadronization, and pile-up, can then be applied.

3.2 Hadronization

Hadronization of the generated events was delegated to PYTHIA. An interface between WHIZARD and PYTHIA was written for this purpose. A detailed description of the interface will be given in the WHIZARD manual once the analytic parton shower is released as an official part of the WHIZARD package.

3.3 Handling of beam remnants

We implemented a very rudimentary treatment of beam remnants, with the main purpose of being able to provide a color-neutral input to the hadronization. In dependence of the emitted particle the beam remnant is assumed to consist of one or two partons, the procedure for determining these partons' flavours and momenta is given below.

The given procedure obviously only applies in the case of only one emitted parton per proton, that is in the case of only one hard interaction. As an implementation of an interleaved multiple interactions/initial-state radiation evolution along the lines of the *Interleaved Evolution* approach [36] is in preparation, this simple treatment of beam-remnants will become inapplicable. Thus a more sophisticated treatment will be implemented in the future.

3.3.1 Flavours

The flavours of the beam remnant are chosen according to a simplified version of PYTHIA's procedure [1, section 11.1.1]. Depending on the flavour of the emitted parton the flavours of the beam remnant are chosen (These rules apply for protons as the initial hadrons, with obvious substitutions for antiprotons.):

- A valence quark of the hadron is assumed to leave behind a diquark beam remnant. A ud-diquark[9] is assumed to be a ud_1 in 25% and a ud_0 in 75% of the cases, while a uu-diquark is always a uu_1.

- A gluon is assumed to leave behind a colour octet state, that is divided into a colour triplet quark and an anti-colour triplet antiquark. The division into $u + ud_1$ for 1/6 of the cases, into $u + ud_0$ for 1/2 and into $d + uu_1$ for 1/3 of the cases.

- A sea quark, for example a s, leaves behind an $uud\bar{s}$ state, that is subdivided into a meson and a diquark. The relative probabilities are 1/6 into $u\bar{s} + ud_1$, 1/2 into $u\bar{s} + ud_0$ and 1/3 into $d\bar{s} + uu_1$.

- An antiquark $\bar{q}$ leaves behind a $uudq$ state, that is divided into a baryon and a quark. Since mostly the $q\bar{q}$ pair comes from an emission of a gluon, the subdivision $uud + q$ is not allowed as it would correspond to a color singlet gluon. The subdivision is therefore in 2/3 of the cases into $udq + u$ and in 1/3 of the cases into $uuq + d$. The three quark state uuq or udq is then replaced by the corresponding baryon of lowest spin.

3.3.2 Momenta

The total momentum of the beam remnant is given by the remaining momentum of the hadron after the emitted particle has been removed. In case the beam remnant consists of only one parton, this parton is assigned the complete momentum, if the beam remnant consists of a diquark and a quark the momentum is distributed in equal parts to both the constituents. If the beam remnant consists of two constituents with one of them being a meson or baryon, the energy is distributed in equal parts but the three-momentum is distributed so that the hadron is on-shell and the quark is assigned the remaining momentum. This procedure generates on-shell colorless particles and off-shell coloured particles so that the coloured particles off-shellness is absorbed in the hadronization.

4 Results and Validation

We compared the predictions for the process $e^+e^- \rightarrow hadrons$ at LEP with an energy of $\sqrt{s} = 133\,\text{GeV}$ and for Z production, $p\bar{p}/pp \rightarrow Z + X$, at the Tevatron and the LHC at energies of $\sqrt{s} = 1.96\,\text{TeV}$ and $\sqrt{s} = 7\,\text{TeV}$, respectively. All event sets were generated using WHIZARD, which means the hard interaction was simulated by WHIZARD/O'Mega, the parton shower was either simulated using PYTHIA's virtuality-ordered shower or WHIZARD's own

---

[9] Diquarks are given in the notation qq_S, where the q are the building quarks and S is the total spin.

analytic shower, denoted in the plots by either PYTHIA PS or WHIZARD PS. For WHIZARD's parton shower we used a first-order running α_S in the $\overline{\text{MS}}$-scheme, given by

$$\alpha_S\left(Q^2\right) = \frac{4\pi}{\left(11 - \frac{2}{3}n_f\right)\log\left(Q^2/\Lambda^2\right)}, \tag{4.1}$$

taking n_f and Λ as constants, neglecting the influence of flavor thresholds for the moment. For PYTHIA's parton shower we used the same Λ and n_f values for a first-oder running α_S, but with threshold effects enabled. If not stated otherwise, the values for Λ were chosen to be $\Lambda = 0.19\,\text{GeV}$ for WHIZARD's parton shower and $\Lambda = 0.29\,\text{GeV}$ for PYTHIA's parton shower. The Λ value for WHIZARD was chosen by hand to improve agreement with PYTHIA's distributions. The hadronization, if activated, was simulated using PYTHIA with the hadronization tune from [37, table 10, Dec. 93]. This tune was of course made using PYTHIA's parton shower, but will be used in here together with WHIZARD's parton shower as well. As the hadronization tune depends on the parton shower, using a tuning obtained with a different parton shower can lead to unsubstantial deviations in the results. As there is no tune with WHIZARD's parton shower available, we cannot give an estimate for the deviations. The possible tuning of our shower is beyond the scope of the present paper, presenting merely the algorithm, and will be left for future work. Given the fact that no tuning has been done, the shower already describes data in a QCD environment reasonably well.

The definitions of all observables are given in section A in the appendix.

4.1 Final State Radiation at parton level

Figures 3 to 6 show a comparison of distributions of event shapes at parton level. For both parton shower programs the events were showered with a cut-off virtuality $Q^2_{min} = 1\,\text{GeV}$, hadronization was disabled. The plots for thrust, thrust major and thrust minor show that WHIZARD's parton shower generates more spherical events compared to PYTHIA's parton shower. Nonetheless, they show a satisfactory agreement as WHIZARD's parton shower was not tuned at all for these plots. However, it is unclear if the discrepancies can be tuned away. Moreover, as distributions at parton level are not observable in an experiment, it is doubtful if they need to be.

4.2 Final State Radiation at hadron level

4.2.1 Event shapes

For hadronized events, we can compare the generated distributions with experimental data. We compared the distributions for several event shapes with data from the DELPHI collaboration [38]. The measurement was performed using e^+e^- collisions at center-of-mass energies of $\sqrt{s} = 130\,\text{GeV}$ and $136\,\text{GeV}$. The simulated hard interaction was chosen to be $e^+e^- \rightarrow u\bar{u}$ at a center-of-mass energy of $\sqrt{s} = 133\,\text{GeV}$.
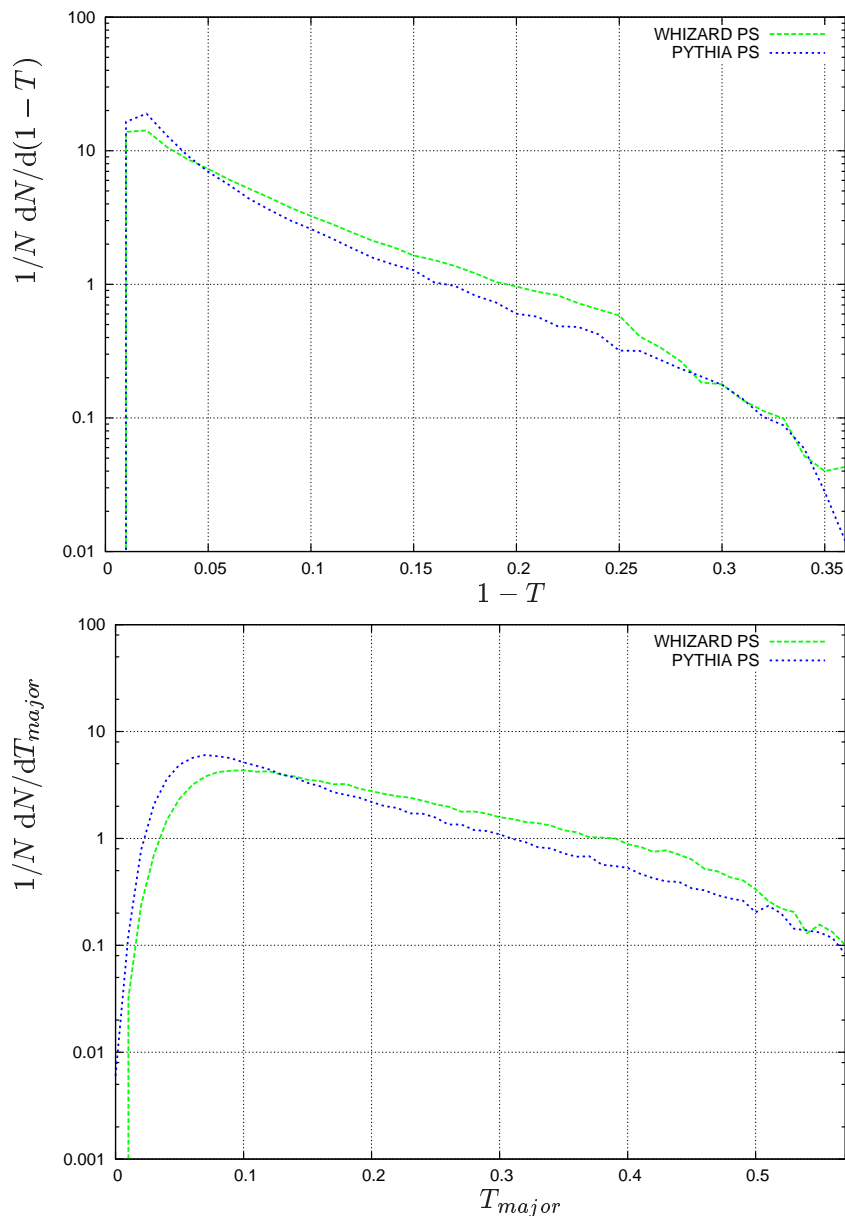
Figure 3: Plots for thrust T and thrust major T_{major} (without hadronization). The dashed/green/bright line is WHIZARD, the dotted/blue/dark line is PYTHIA.

The results are shown in figures 7 to 10. Both parton showers show good agreement, especially if one takes into account that the events showered with WHIZARD's parton shower where hadronized with the PYTHIA hadronization tuned to data using events showered with PYTHIA. As for the unhadronized samples, events showered with WHIZARD tend to populate the regions corresponding to more spherical configurations compared to events generated using the PYTHIA shower. The plot for thrust major T_{maj} shows a slight undershooting of the WHIZARD curve with respect to the data in the two bins from 0.04 to 0.08. However,
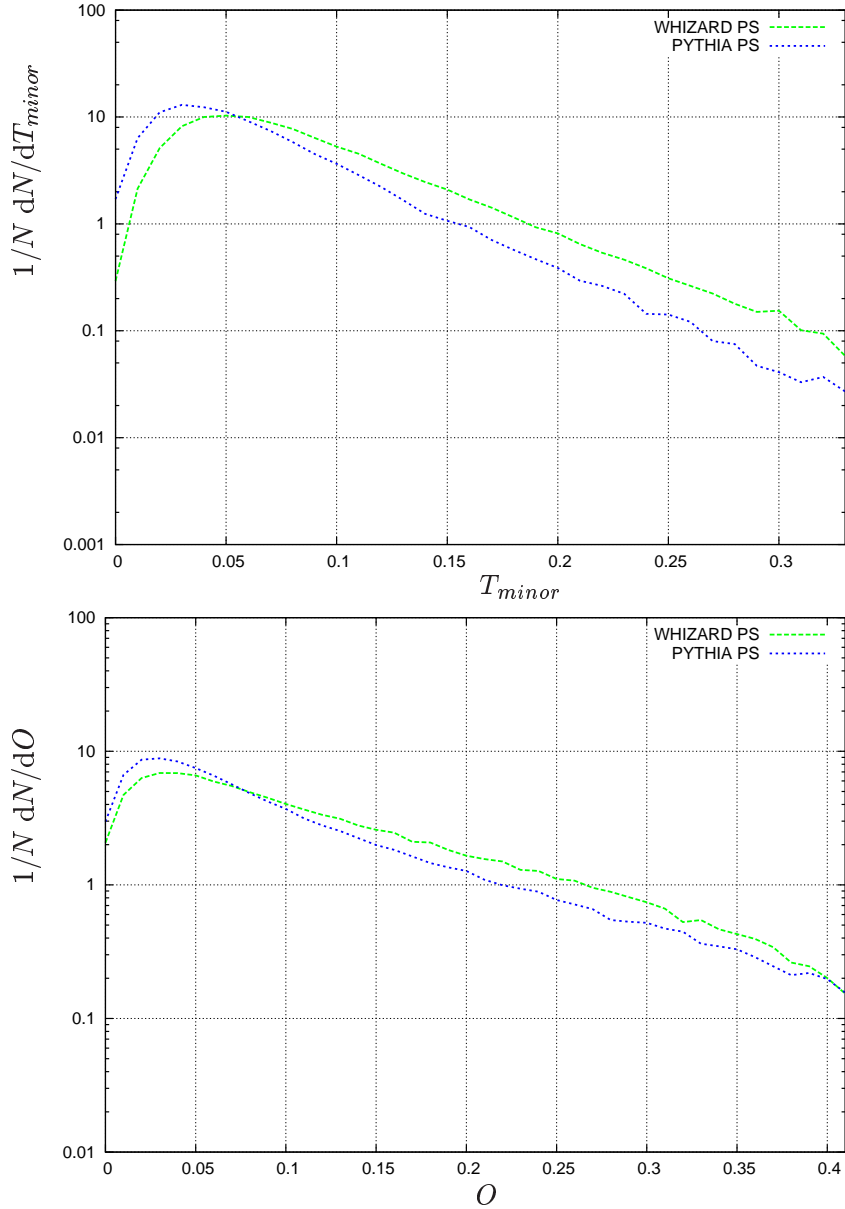
Figure 4: Plots for thrust minor T_{minor} and Oblateness O (without hadronization). The dashed/green/bright line is WHIZARD, the dotted/blue/dark line is PYTHIA.

both distributions are mostly consistent with the data.

4.2.2 Jet rates

A comparison of the Monte Carlo results for the process $e^+e^- \Rightarrow q\bar{q}$ at $\sqrt{s} = 91\,\text{GeV}$ with measurements from the JADE and OPAL collaborations given in [39] is shown in figures 11 and 12. Shown are differential jet rates as a function of the resolution parameter in the k_T-clustering algorithm $y_{i\,i+1}$, where the event turns from being a $i + 1$-jet event into
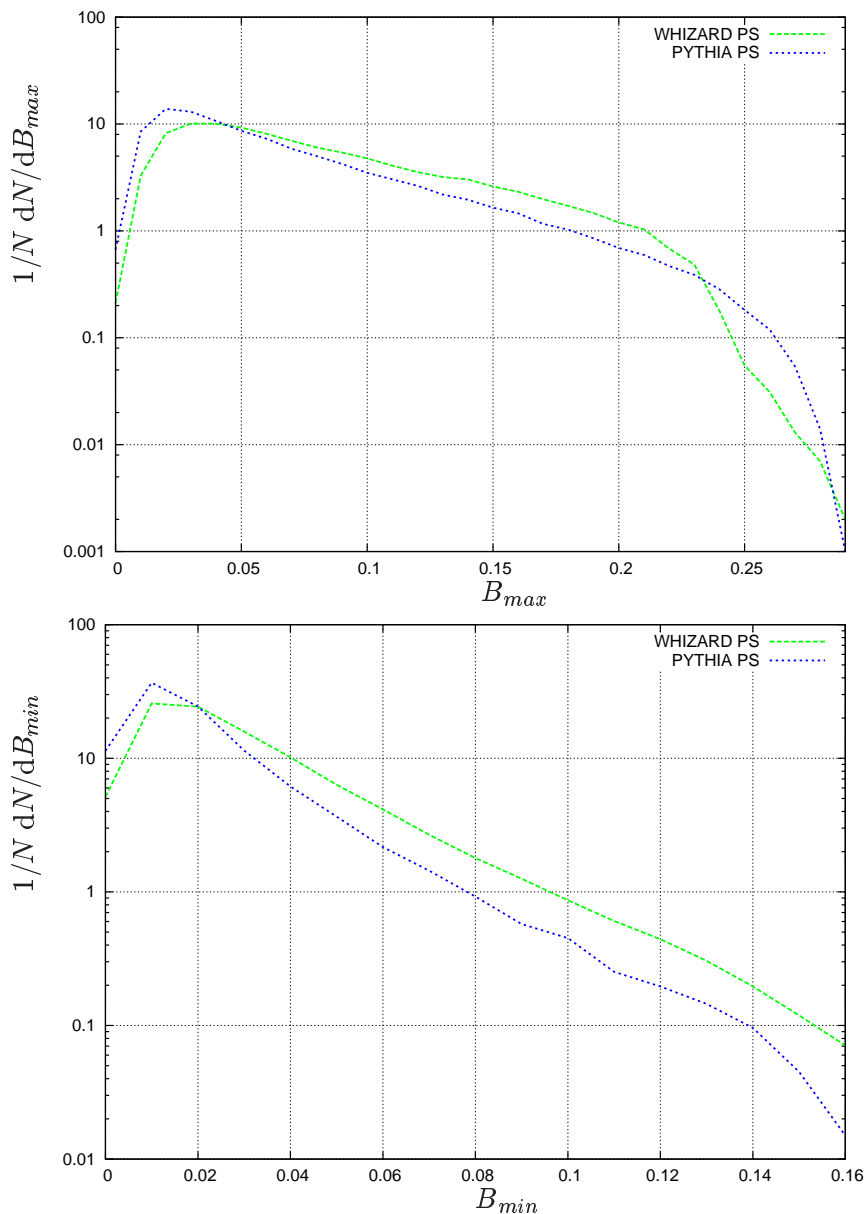
Figure 5: Plots for jet broadenings B_{max} and B_{min} (without hadronization). The dashed/green/bright line is WHIZARD, the dotted/blue/dark line is PYTHIA.

a i-jet event. The definition of the clustering variable is given in equation (A.1) in the appendix. The comparison is equivalent to the one in [40, 41], where a tuning of some parton shower and hadronization parameters was performed. The only tuning applied to the parton shower in the comparison was a by-hand adjustment of α_S, setting Λ_{QCD} to a value of $0.15\,\mathrm{GeV}$. In general, the plots confirm the tendency of WHIZARD's parton shower to generate more spherical events compared to PYTHIA as the bins with higher values of $y_{i\,i+1}$ are populated more. Note that small values of $y_{i\,i+1}$ correspond to small invariant
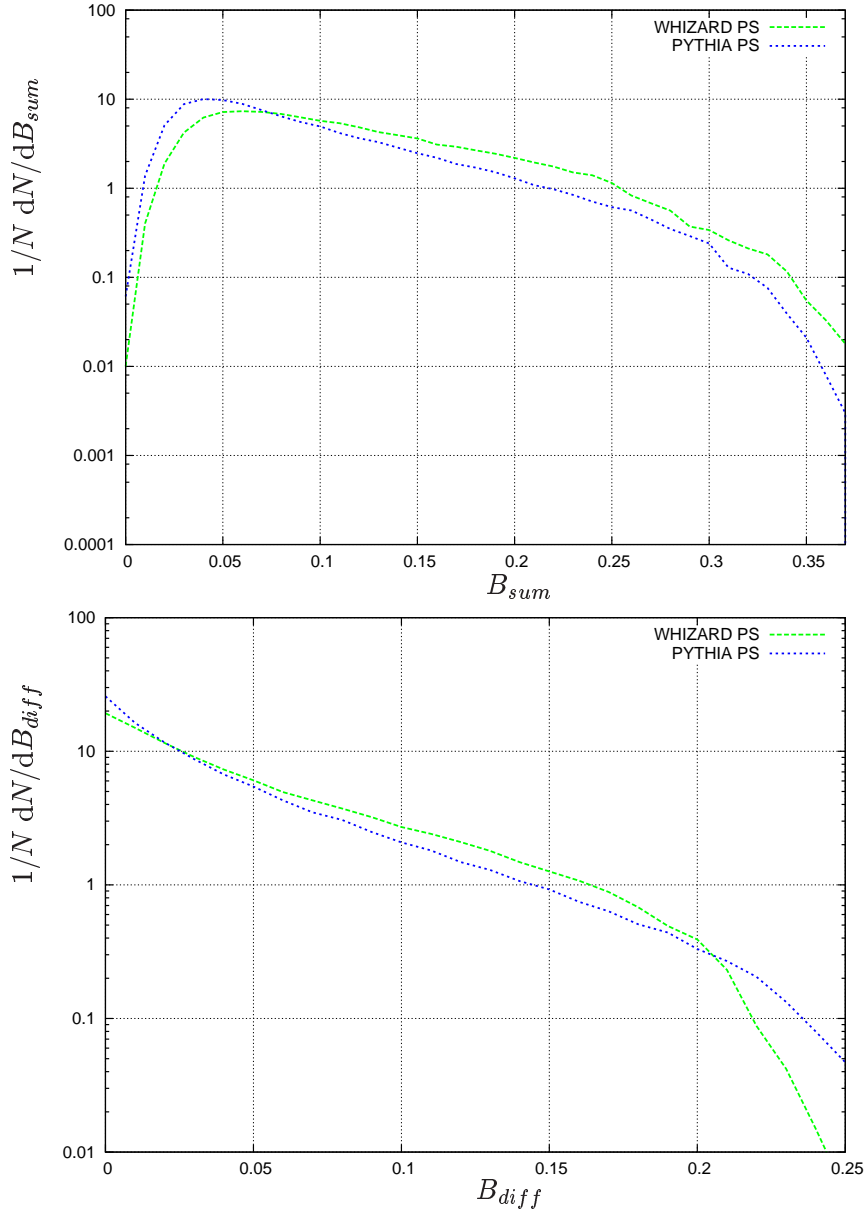
Figure 6: Plots for jet broadenings B_{sum} and B_{diff} (without hadronization). The dashed/green/bright line is WHIZARD, the dotted/blue/dark line is PYTHIA.

masses and that these regions are described by the hadronization model and not the parton shower. So the differences in the left parts of the plots can stem from two sources. They might be caused by normalization effects due to over-estimation in the right parts. Any remaining difference would show that the hadronization tune obtained with PYTHIA is not suitable to describe these regions when used with WHIZARD's shower.
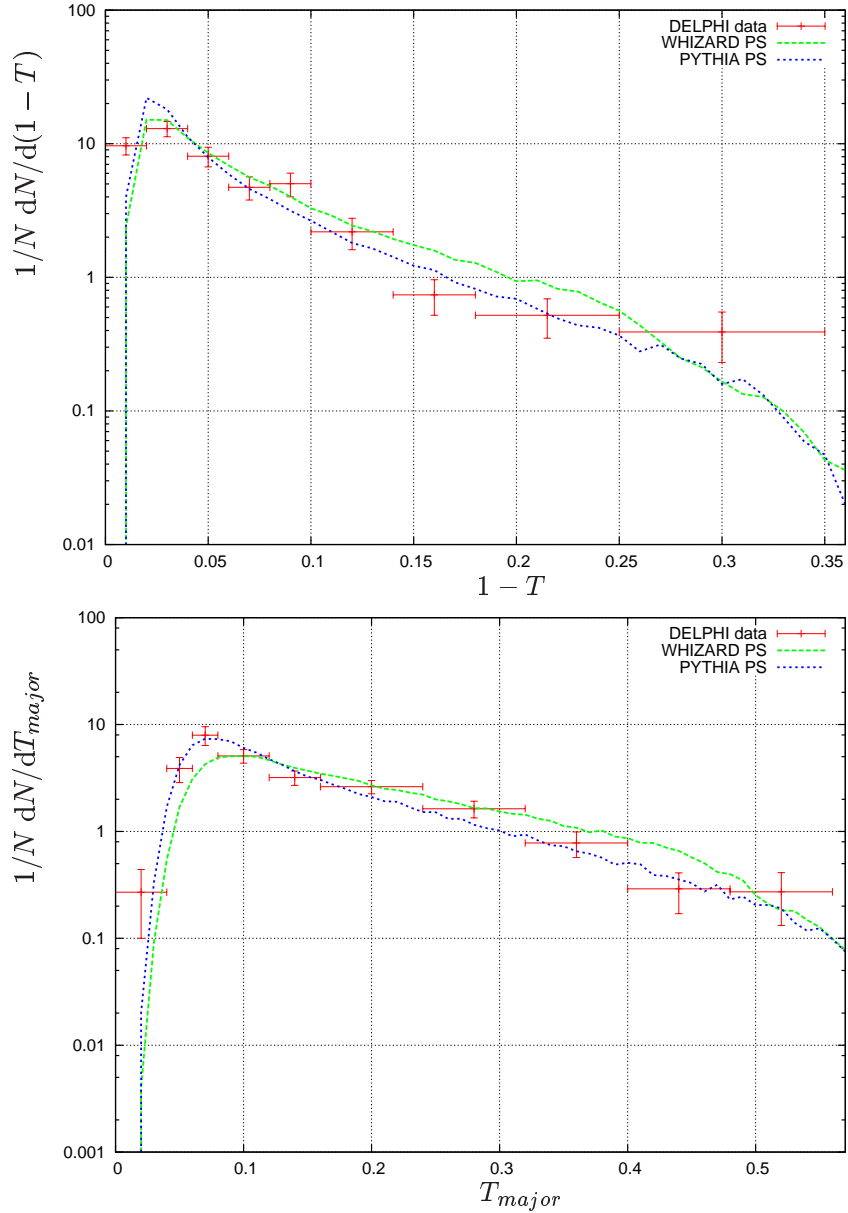
Figure 7: Plots for thrust T and thrust major T_{major} (with hadronization, data from [38]). The dashed/green/bright line is WHIZARD, the dotted/blue/dark line is PYTHIA.

4.3 Initial State Radiation

A plot for the transverse momentum of a Z-Boson produced in $p\bar{p}$-collisions at $\sqrt{s} = 1.96$ TeV is given in figure 13. The simulation with PYTHIA was done using Rick Field's CDF Tune D6 with CTEQ6L1 parton distribution functions. The simulation using WHIZARD's parton shower was done using the same PDFs, multiple interactions were disregarded in both simulations. The data obtained from WHIZARD's initial-state parton shower shows two distinct features: first of all, the curve in the low-p_T region shows a slight deviation with
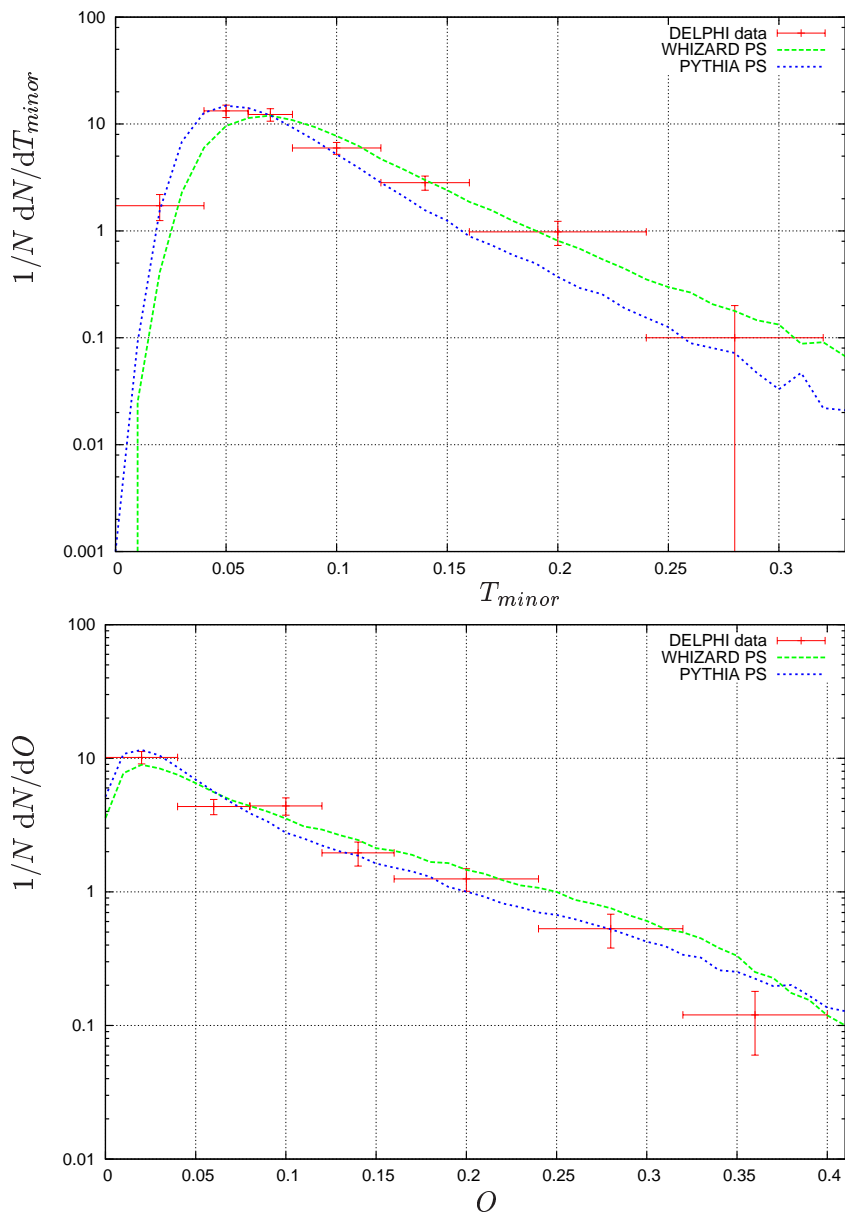
Figure 8: Plots for thrust minor T_{minor} and Oblateness O (with hadronization, data from [38]). The dashed/green/bright line is WHIZARD, the dotted/blue/dark line is PYTHIA.

respect to the corresponding PYTHIA curve. However, as we will see later, this is still in agreement with data. Second of all, it shows the known phase space cut at $p_T \lesssim m_Z$ [42]. For comparison, the plot is supplemented by a p_T-histogram for the unshowered process $u\bar{u} \to Zg$. PYTHIA's description uses the power-shower and matching and closely resembles the result for the partonic process.

Our approach to solve the shortcomings of WHIZARD's parton shower was not to include the power shower ansatz, but instead accept this as a deficiency of the parton shower and
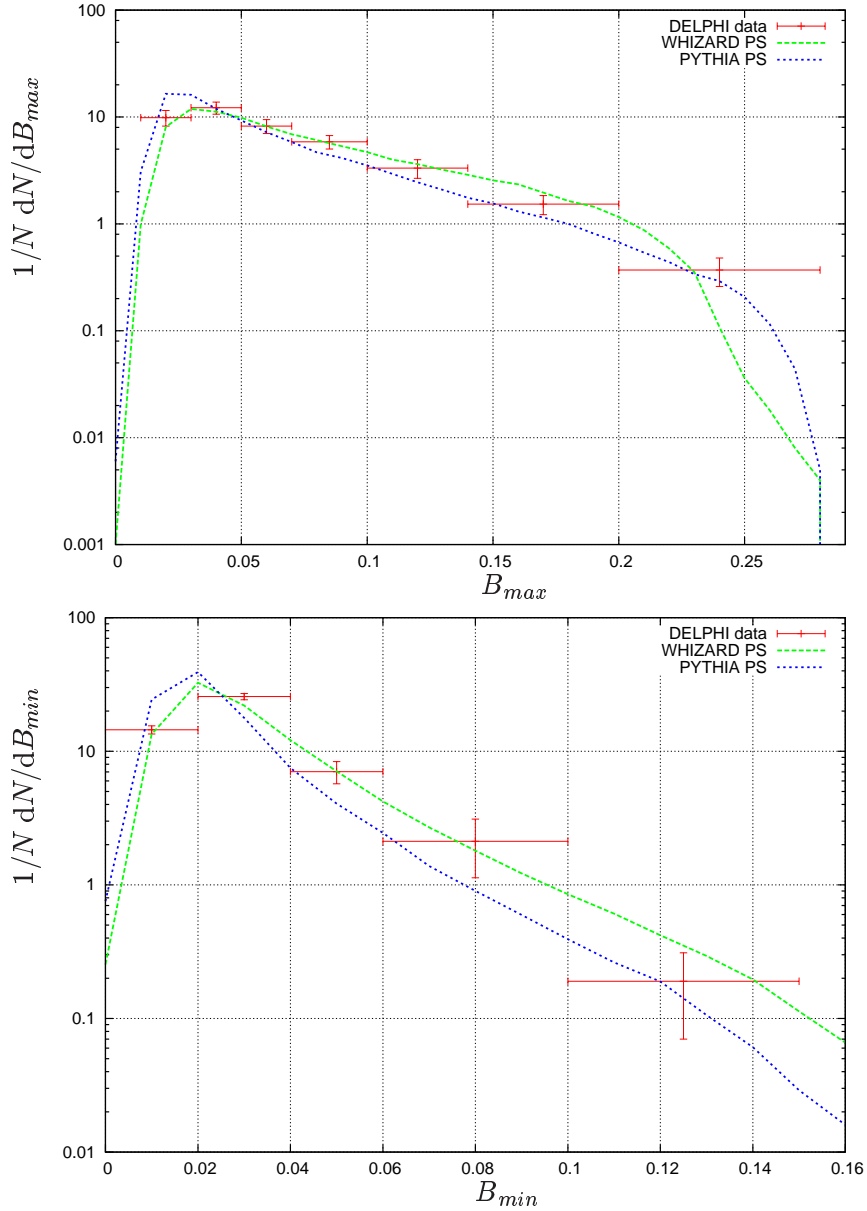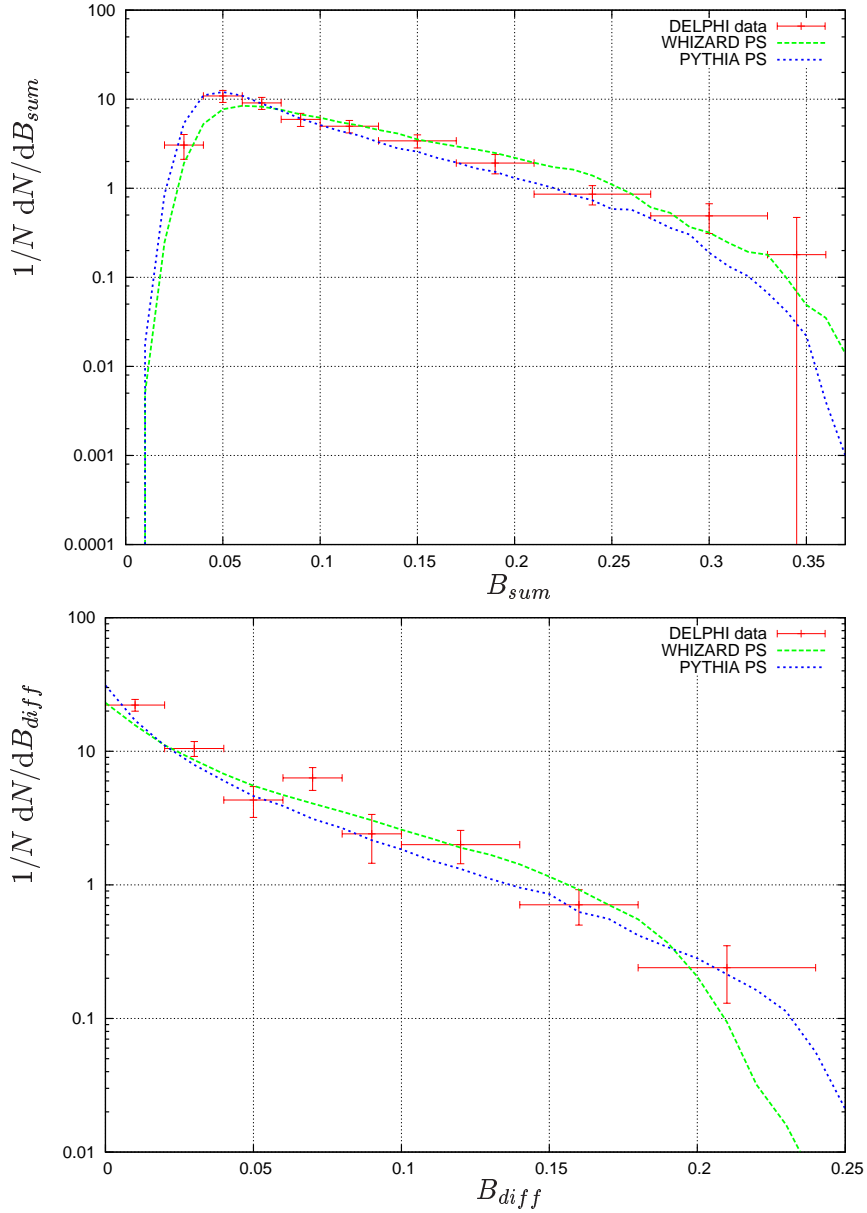
Figure 9: Plots for jet broadenings B_{max} and B_{min} (with hadronization, data from [38]). The dashed/green/bright line is WHIZARD, the dotted/blue/dark line is PYTHIA.

delegate the task of describing the high-p_T region to a matching algorithm.

4.4   Matched Final State Radiation

Plots for results obtained with the MLM matching for the final-state parton shower are shown in figures 14 and 15 for events showered with WHIZARD's parton shower and figures 16 and 17 for PYTHIA. The process under consideration is $e^+e^- \rightarrow u\bar{u}$ at a center of mass energy of 91 GeV, hadronization was switched off. The process was simulated in

Figure 10: Plots for jet broadenings B_{sum} and B_{diff} (with hadronization, data from [38]). The dashed/green/bright line is `WHIZARD`, the dotted/blue/dark line is `PYTHIA`.

five different ways, first without any matching at all and then with a variable number of additional jets from zero to three, where each additional jet could be a gluon or a u,d,s or c quark. For the unmatched case and for each jet multiplicity an event set consisting of 150000 events was simulated. The plots show normalized distributions for event shapes obtained from these samples.

The plots show some common features. The line for the (moot) case of no additional jets closely resembles the line for the unmatched event sample, except in the region of low
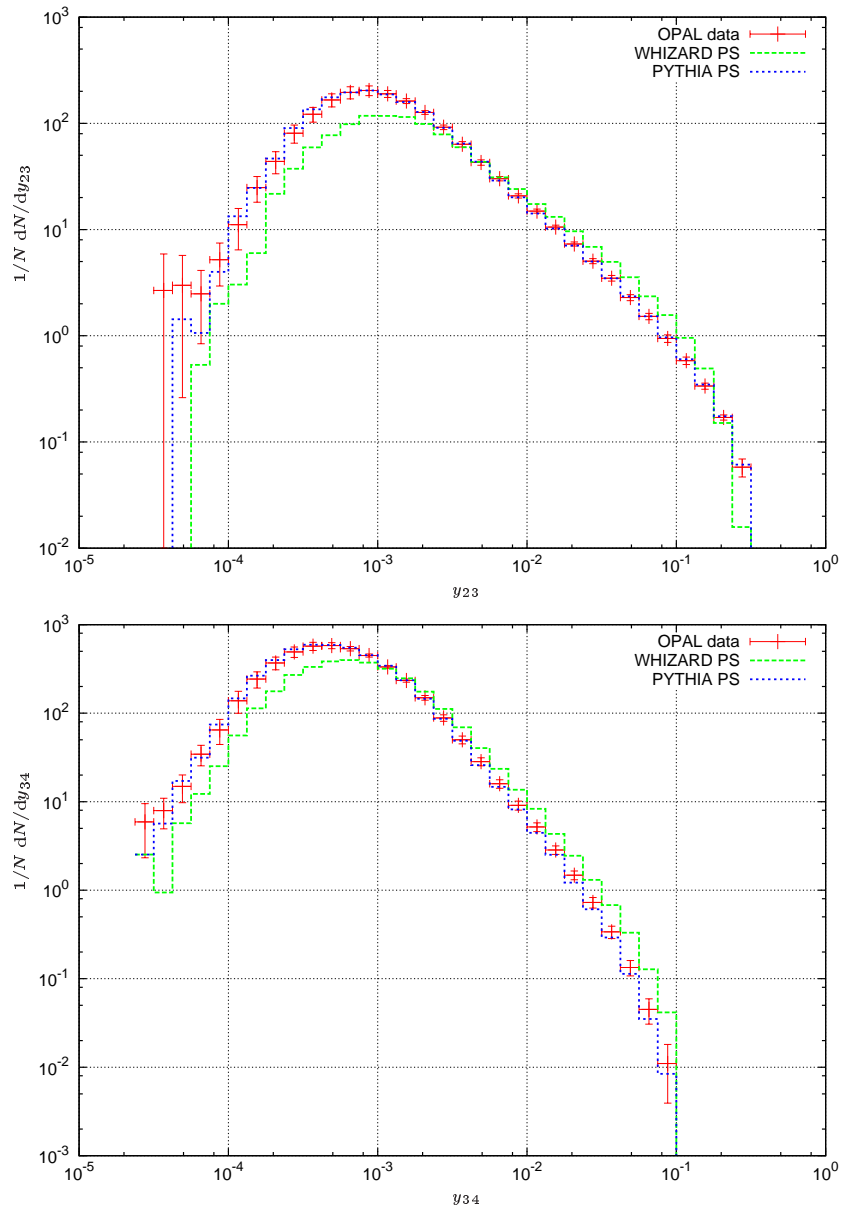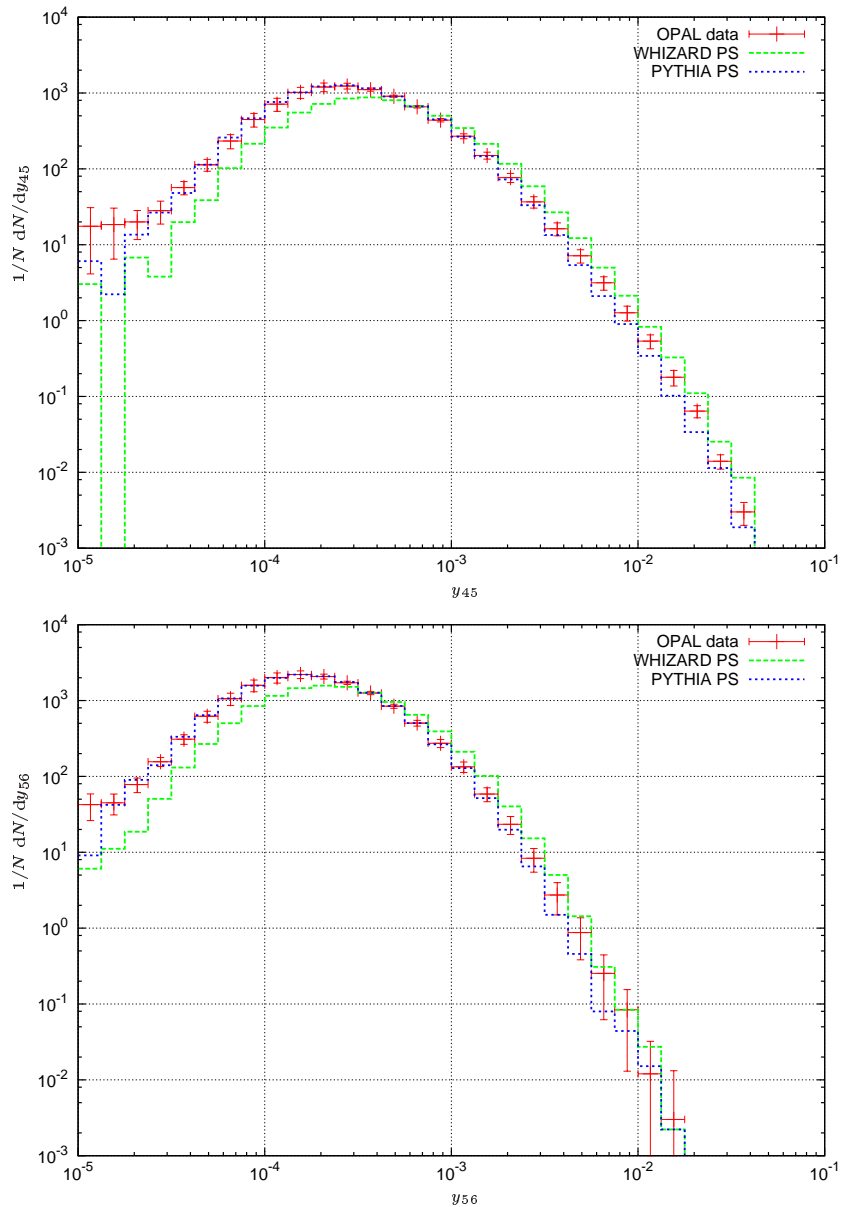
Figure 11: Plots for differential jet rates y_{23} and y_{34}. The dashed/green/bright line is WHIZARD, the dotted/blue/dark line is PYTHIA.

thrust (right part of the upper image in figures 14 and 15). The missing events are events where the parton shower splits a hard jet into two separated jets, so that the matching procedure can not cluster any of the two jets to the original parton and therefore rejects the event.

The lines for one, two and three additional jets lie on top of each other so that it can be concluded that for these observables, the inclusion of one additional jet is sufficient.

Figure 12: Plots for differential jet rates y_{45} and y_{56}. The dashed/green/bright line is WHIZARD, the dotted/blue/dark line is PYTHIA.

The deviations between the unmatched and the matched event samples exhibit different behaviour: for PYTHIA the number of spherical events is larger for the matched sample, stemming from the better description of large angle emissions. For WHIZARD the deviations are opposite, the number of more pencil-like events are enhanced, while especially the number of events with medium values of $1-T$ and T_{maj} is decreased. This can be regarded as correcting the tendency to favour more spherical events mentioned in section 4.2. The
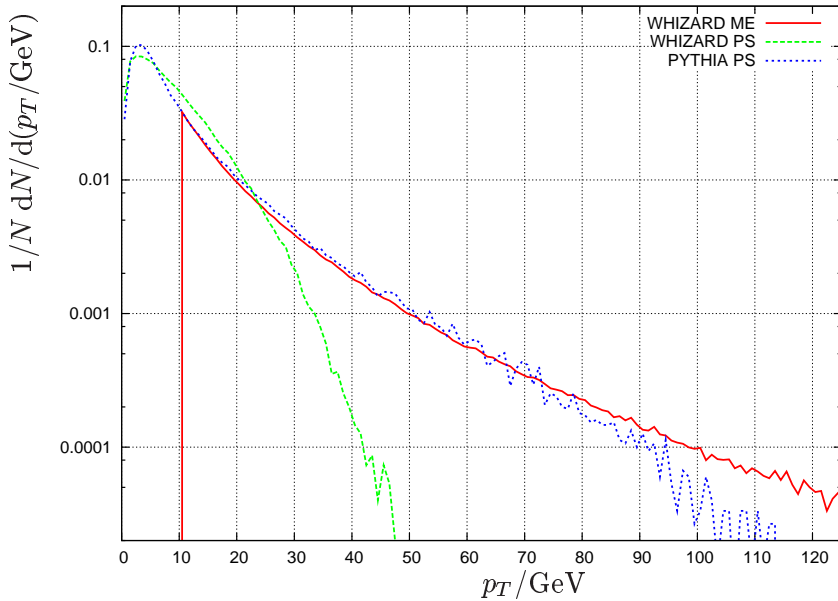
Figure 13: Transverse momentum of a Z-Boson in various schemes. The normalization for events from `WHIZARD`'s matrix element was chosen manually to fit `PYTHIA`'s PS result in the range $10\,\mathrm{GeV} < p_T < 20\,\mathrm{GeV}$.

differences between the distributions for matched and unmatched event samples have to be taken into account when tuning the combination of shower and matching to data. Therefore this can be seen as an example for an observable which is sensitive to regions enriched by hard jet emission, and not so much dominated by universal logarithmic terms. For such an observable, a tuning obtained without matching cannot be reliably used to generate matched samples.

We also did a comparison to data corresponding to the comparison for the unmatched showers in figure 7. We used the curve for the $e^+e^- \to 5jets$ as the sample for the matched shower. The plot is shown in figure 18. The curve for thrust T is slightly altered, most prominent differences to figure 7 is a less pronounced peak with both showers and an increase for the `PYTHIA` curve for values $1 - T > 0.1$. The curves for Thrust major T_{maj} show similar behaviour to the unmatched curves. Both reproduce the data, except for `WHIZARD`'s parton shower's tendency to more spherical configurations and the small number of events in the lower T_{maj}-bins. Both these deficiencies have already been visible in the unmatched event samples.

4.5 Matched Initial State Radiation

To test the matching procedure for the initial state we simulated the process $q\bar{q} \to Z$ and the additional corrections $j\,j \to Z\,j\,(j)$ for one (two) additional jets, $j = u, \bar{u}, d, \bar{d}, s, \bar{d}, c, \bar{c}, g$. The resulting distributions for the Z boson transverse momentum are given in figure 19 for
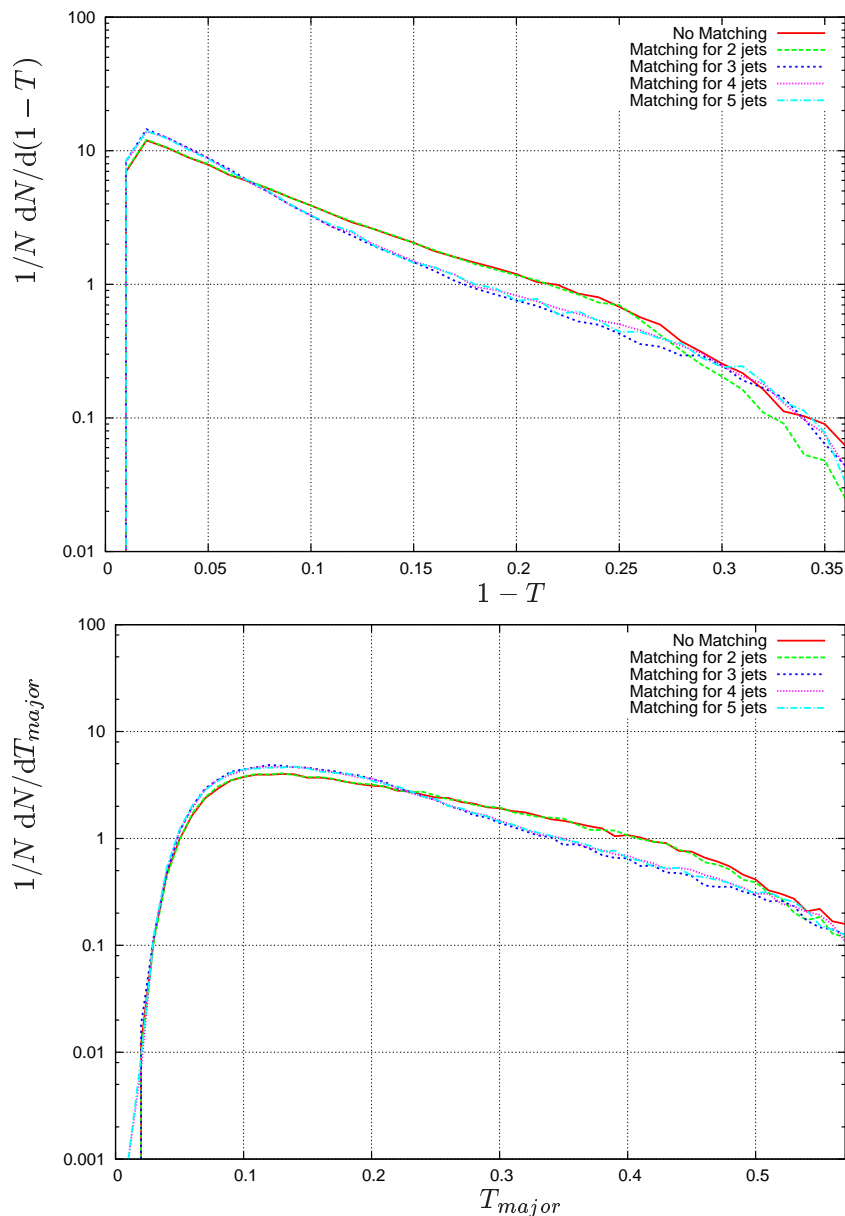
Figure 14: Plots for thrust T and thrust major T_{major} (`WHIZARD` ME + `WHIZARD` PS with matching).

`WHIZARD` and figure 20 for `PYTHIA`. For comparison the measured distribution from D0 [43] was included. Note that all simulated distributions were obtained with disabled primordial k_T.

As expected, the results for `PYTHIA` do not depend much on the application of matching as its power shower approach already generates a p_T-distribution close to the correct distribution [42][10]. The plot for `WHIZARD` shows the expected addition of high-p_T events,

---

[10]`PYTHIA`'s own matching was disabled during this simulation.
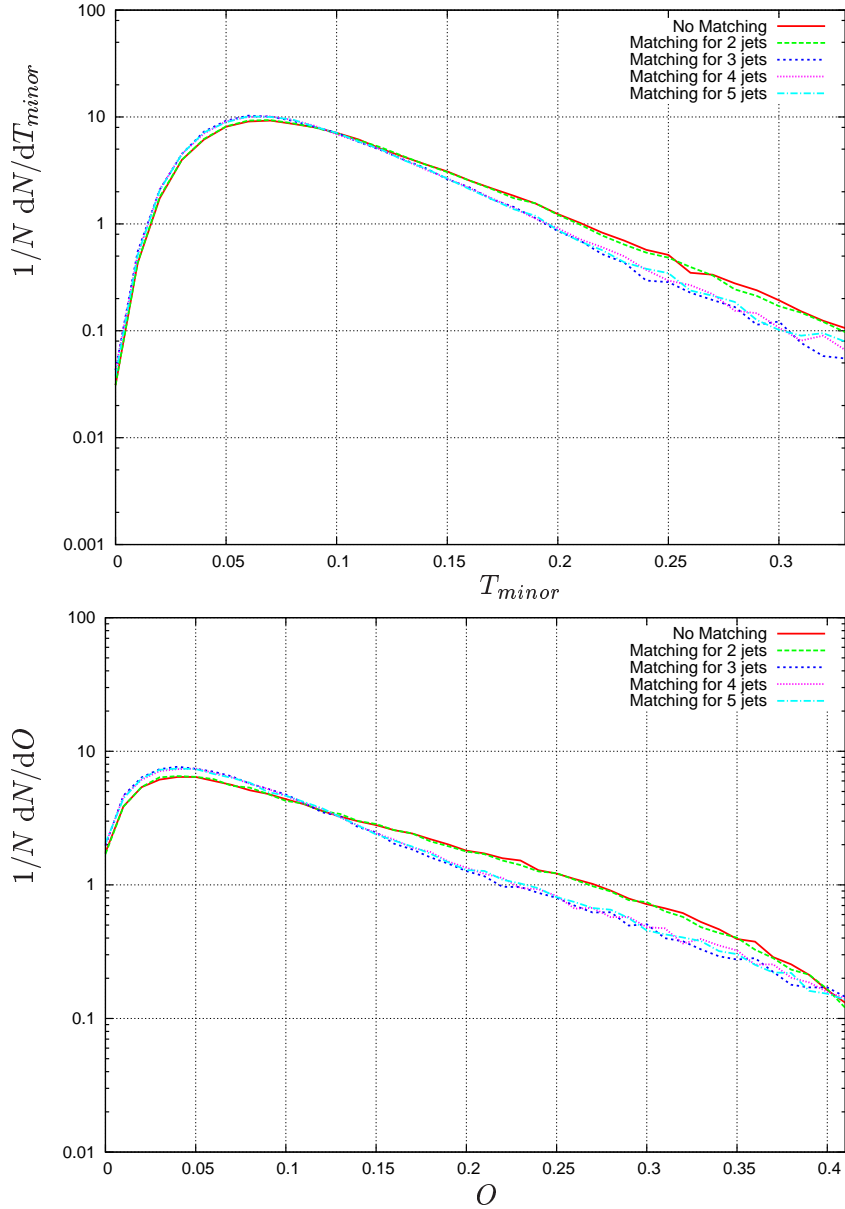
Figure 15: Plots for thrust minor T_{minor} and Oblateness O (`WHIZARD` ME + `WHIZARD` PS with matching).

the concavity is weakend. Adding a second jet described by the matrix element does change the distribution only marginally for both showers.

Figure 21 shows the dependence of the p_T-spectrum on the MLM-matching parameter $p_{T\,min}$. The distribution should be independent of $p_{T\,min}$, however a small difference is visible in the range $10\,\mathrm{GeV} \lesssim p_T \lesssim 80\,\mathrm{GeV}$. The high-$p_T$-tail remains stable when changing $p_{T\,min}$, the shape at the peak does not change as well. The differences are within the expected dependence on the matching parameters.
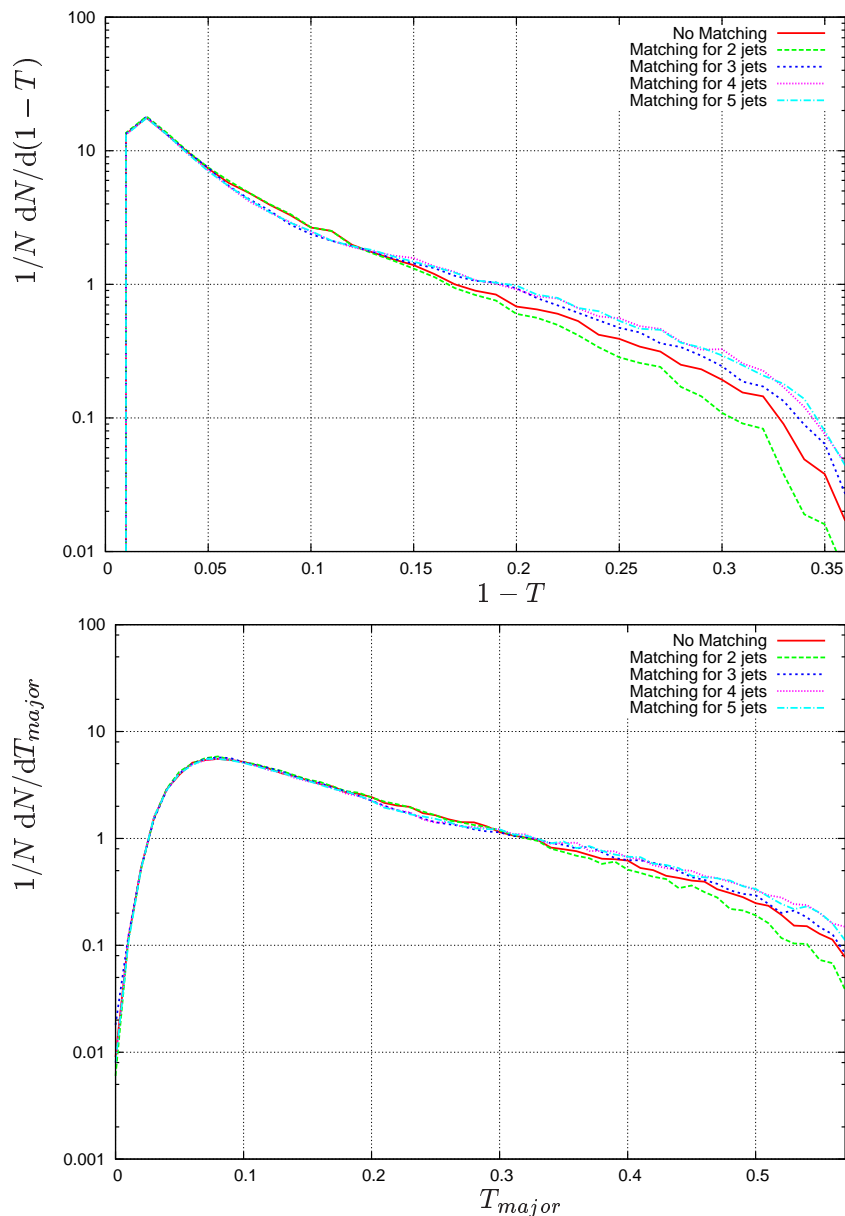
Figure 16: Plots for thrust T and thrust major T_{major} (`WHIZARD` ME + `PYTHIA` PS with matching).

As a further test, we compared the Z-Boson p_T at the LHC. We used the recently published measurement by CMS [44]. Except for the change from proton-antiproton beams to proton-proton beams and the increased center of mass energy $\sqrt{s} = 7\,\text{TeV}$, all other settings were the same as for the Tevatron simulation. This holds particularly for the chosen `PYTHIA` tune, that was obtained from measurements at Tevatron and usage at the LHC cannot be regarded as trustworthy. Nevertheless, the data can be reproduced very well, except for an overshoot in the lowest bins. As for `WHIZARD` there are no available tunes yet, so the dependency on a particular tune is not an issue. Note that the simulation was
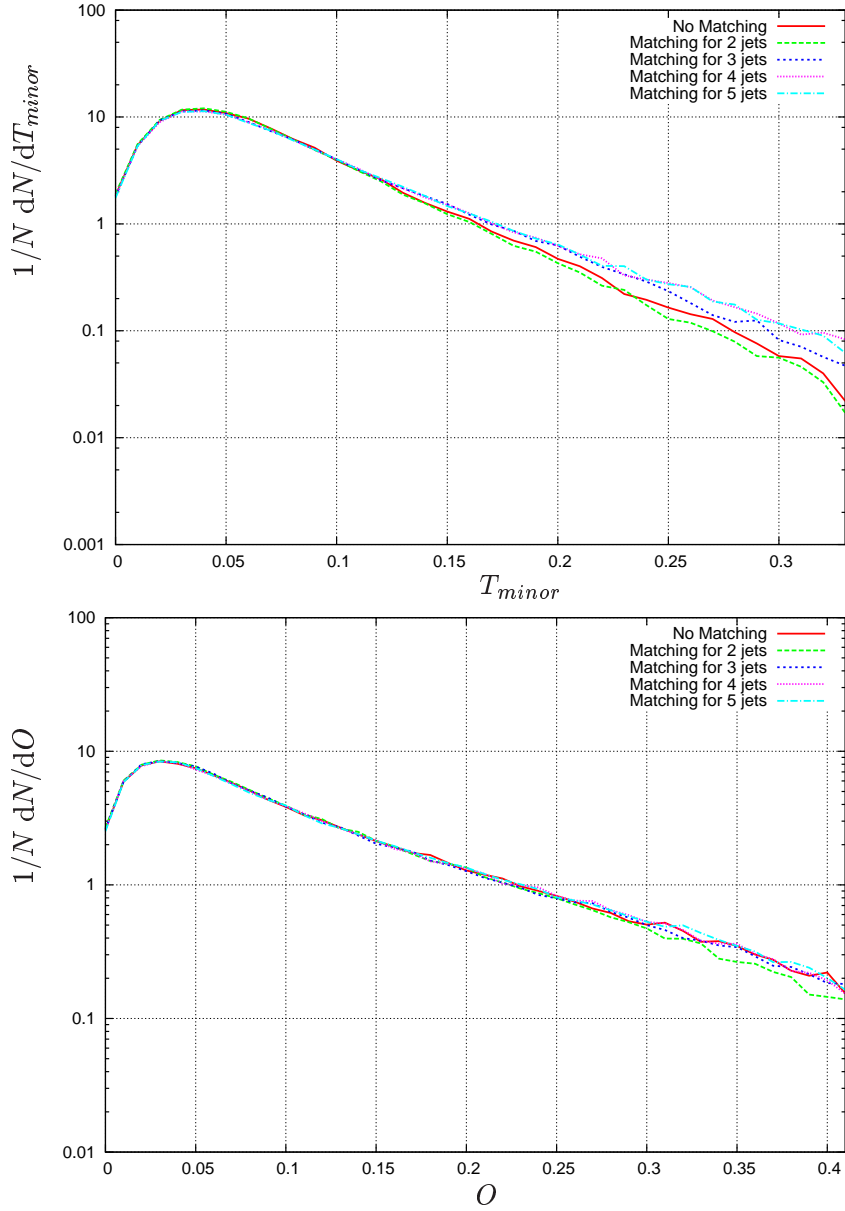
Figure 17: Plots for thrust minor T_{minor} and Oblateness O (`WHIZARD` ME + `PYTHIA` PS with matching).

done with primordial k_T disabled, so that the lowest bins are expected to be overpopulated. Apart from this difference, the simulation using `WHIZARD`'s hard interaction, parton shower and matching procedure reproduces the data as good as the simulation performed using `WHIZARD`'s hard interaction and matching, but `PYTHIA`'s parton shower.
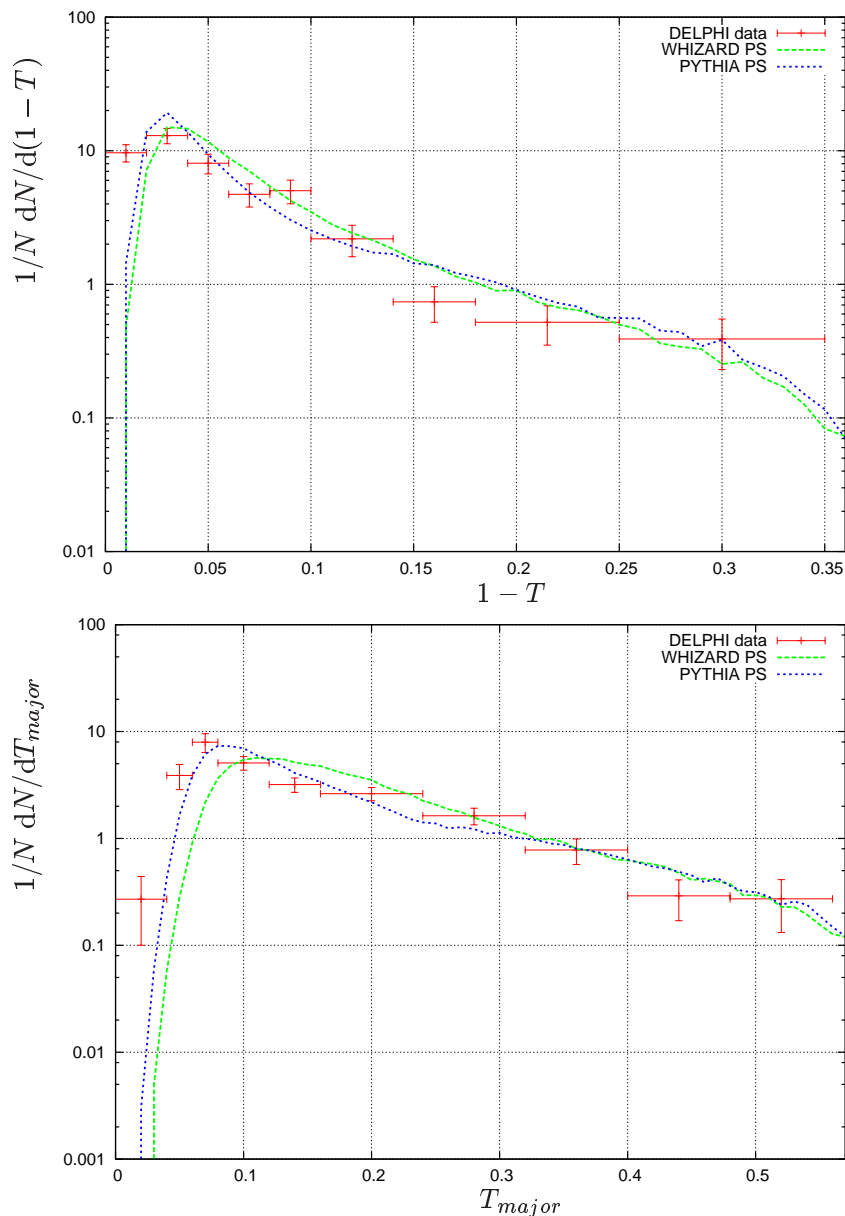
Figure 18: Comparison of the predictions for parton showers matched to the process $e^+e^- \rightarrow u\bar{u} + 3jets$.

5   Conclusion and Outlook

In this paper, we presented an algorithm for an analytic parton shower for both initial and final-state radiation. While this parton shower algorithm for the final state has been known for quite some time, the extension to the initial state had not been done up to now.

Analytic parton showers are especially interesting for conceptual development in a theoretical description of QCD in a hadron (but also hadronic final state lepton) collider environment, as they allow to determine the corresponding shower weights from the com-
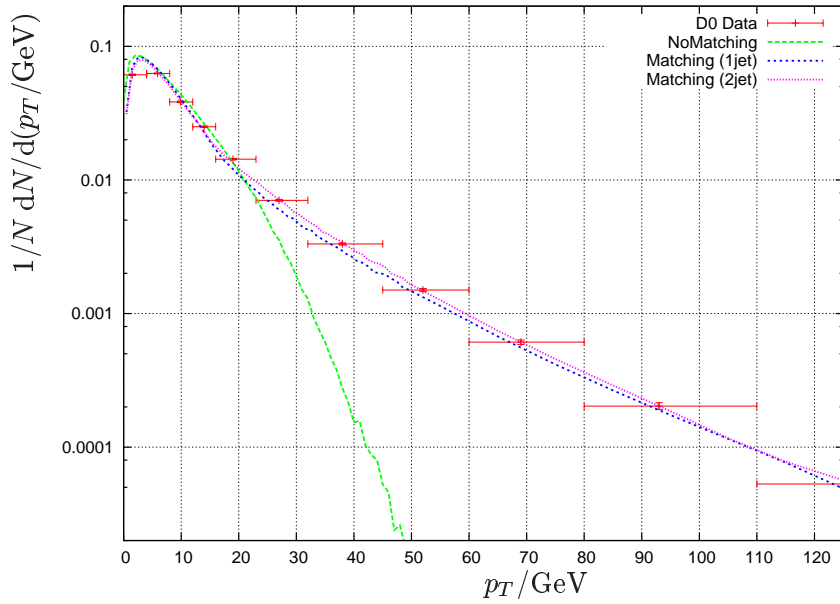
Figure 19: *Z*-Boson transverse momentum, simulated with `WHIZARD` ME and PS without and with matching.
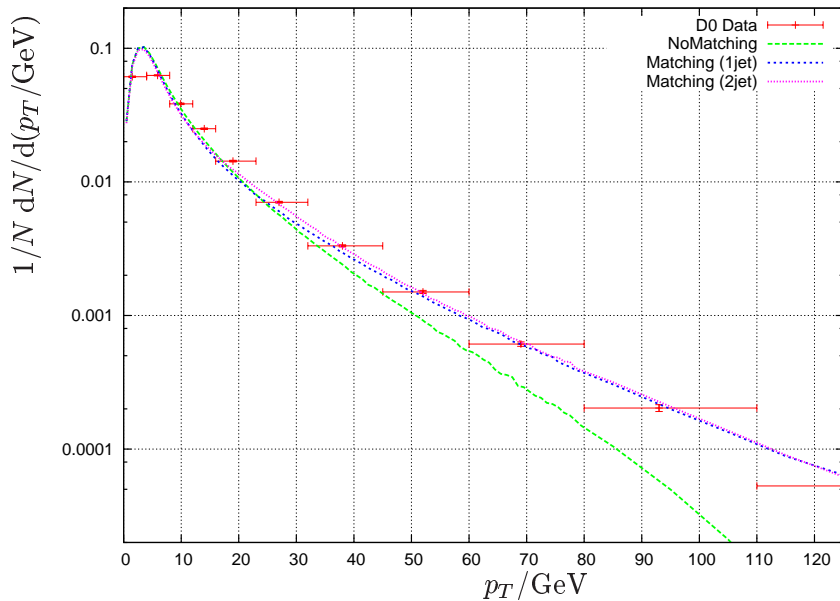

Figure 20: *Z*-Boson transverse momentum, simulated with `WHIZARD` ME and `PYTHIA` PS without and with matching for one and two additional jets.

plete shower histories. As there are also veto mechanism for probabilistic showers, such a task is not viable there. The knowledge of complete shower histories and weights enables one to e.g. change the hard scattering matrix element or the PDFs and reweight the showered events to the new hard scattering process. Furthermore, analytic parton show-
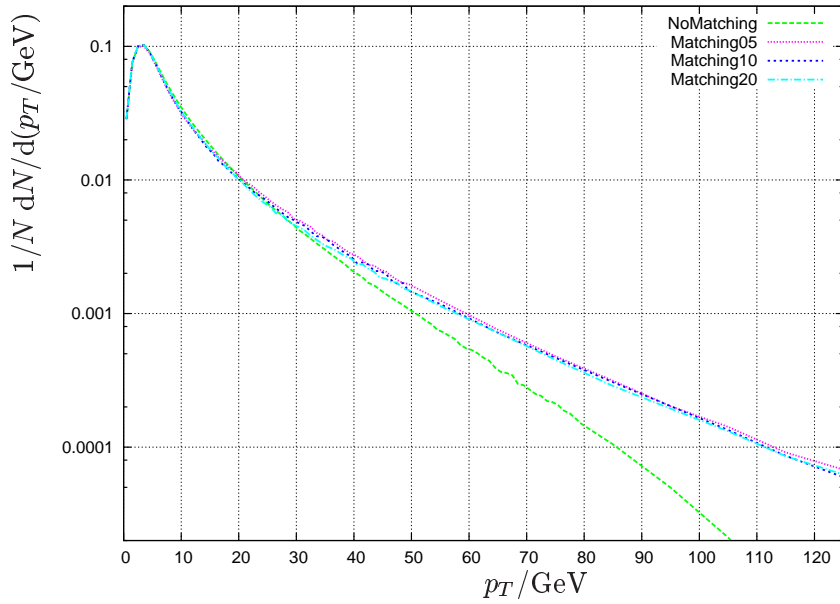
Figure 21: The "NoMatching" curve and the "Matching (1jet)" (cf. figure 20) curve for three different values of $p_{T\,min}$, $5\,\mathrm{GeV}$, $10\,\mathrm{GeV}$ and $20\,\mathrm{GeV}$.
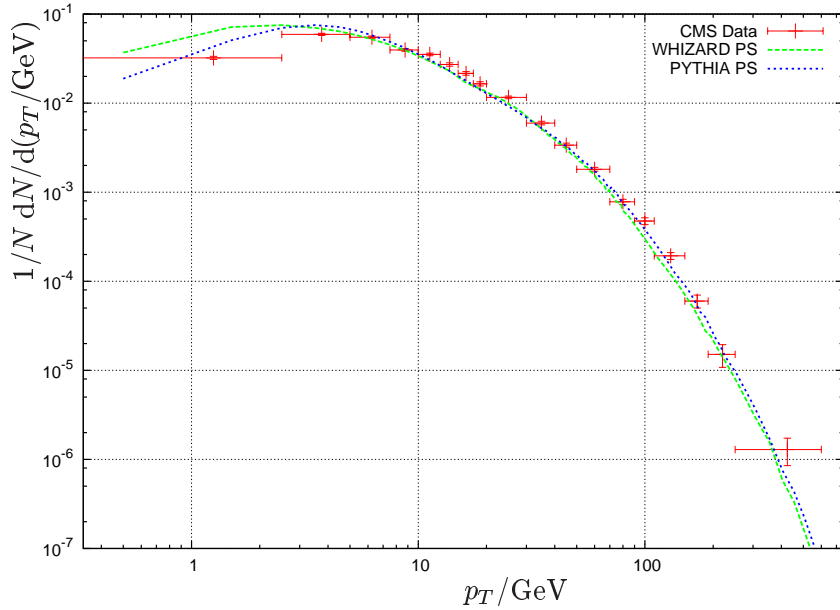

Figure 22: Z-Boson transverse momentum, simulated with WHIZARD ME and PS and WHIZARD ME + PYTHIA PS. A similar figure with the same data appeared in [44]. The dashed/green/bright line is WHIZARD, the dotted/blue/dark line is PYTHIA.

ers might offer the possibility to determine systematic uncertainties from a parton shower approximation in a reliable and theoretically well-defined way[11]

---

[11] A complementary approach has recently been presented in [45]. There the probability for an event is

Also, it might be achievable – using analytic parton showers – to systematically construct higher-order corrections to the parton shower approximation.

The important point for a successful algorithm for an analytic initial-state parton shower is the scale choice, specifically the starting condition of the backward shower evolution, together with the prescriptions for energy and momentum projections in the splittings.

We also improved on the original algorithm for the final-state shower, where e.g. running couplings constants within the shower evolution had not been taken into account. For the description of complete kinematical distributions at hadron colliders, including the high-energy tails, we refrained from the power-shower concept, where also hard and/or non-collinear jets are being produced by means of the shower. Instead we use an MLM-type matching of the analytic parton shower with matrix elements containing one or more additional hard jets explicitly.

Together with the development of the algorithm, we made a thorough comparison, where for now we restricted ourselves to the parton shower from `PYTHIA`. We also made an extended validation of our parton shower algorithm with jet and event shape data from the LEP experiments, from the Tevatron Run II measurements as well as first results from the 2010/11 LHC run. For this task we integrated an implementation of our parton shower algorithm into the event generator `WHIZARD`, while hadronization needed for the comparison, is performed by means of external packages. Our setup allows for a direct comparison of the `PYTHIA` and our own parton shower using the same hard matrix elements within the framework of the `WHIZARD` generator.

Without performing an overly sophisticated tuning of the shower, we reproduced the gross features of a big number of jet and event shape variables at lepton and hadron colliders and found in all cases good agreement.

This paper serves as a proof of concept that an analytic parton shower for the initial state is viable to describe QCD in a realistic collider environment. Future lines of developments will contain a more extensive tuning and validation of the shower as well as the matching and merging prescription. We will also be investigating a possible exchange of the evolution variable for the transverse momentum, p_T, which would guarantee angular ordering and color coherence right from the beginning, which might simplify or even improve on the parton shower description given in our algorithm. A development of an interleaved multiple interaction algorithm connected with a properly color-connected analytic initial-state parton shower together with its implementation is in preparation and will be part of a future publication.

---

not calculated analytically. Instead a veto algorithm is used that keeps track of all accepted and vetoed branchings. For an event generated with unit weight in one setup, its weight in a different setup is given by the ratio of the joint probabilities to accept all accepted branchings while vetoing all vetoed branchings.

Acknowledgements

References

[1] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, *JHEP* **05** (2006) 026, [hep-ph/0603175].

[2] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, *Comput.Phys.Commun.* **178** (2008) 852–867, [arXiv:0710.3820].

[3] M. Bahr *et. al.*, *Herwig++ Physics and Manual*, *Eur. Phys. J.* **C58** (2008) 639–707, [arXiv:0803.0883].

[4] T. Gleisberg *et. al.*, *Event generation with SHERPA 1.1*, *JHEP* **02** (2009) 007, [arXiv:0811.4622].

[5] W. T. Giele, D. A. Kosower, and P. Z. Skands, *A Simple shower and matching algorithm*, *Phys.Rev.* **D78** (2008) 014026, [arXiv:0707.3652].

[6] C. W. Bauer and M. D. Schwartz, *Event generation from effective field theory*, *PHYS.REV.D* **76** (2007) 074004.

[7] C. W. Bauer and F. J. Tackmann, *Gaining analytic control of parton showers*, *Phys. Rev.* **D76** (2007) 114017, [arXiv:0705.1719].

[8] C. W. Bauer, F. J. Tackmann, and J. Thaler, *GenEvA (II): A phase space generator from a reweighted parton shower*, *JHEP* **12** (2008) 011, [arXiv:0801.4028].

[9] W. Kilian, T. Ohl, and J. Reuter, *WHIZARD: Simulating Multi-Particle Processes at LHC and ILC*, *Eur.Phys.J.* **C71** (2011) 1742, [arXiv:0708.4233].

[10] G. Gustafson and U. Pettersson, *Dipole Formulation of QCD Cascades*, *Nucl. Phys.* **B306** (1988) 746.

[11] V. V. Sudakov, *Vertex parts for very high energies in quantum electrodynamics*, *Ž. Eksper. Teoret. Fiz.* **30** (1956) 87–95.

[12] T. Sjostrand, *A Model for Initial State Parton Showers*, *Phys.Lett.* **B157** (1985) 321.

[13] C. W. Bauer, F. J. Tackmann, and J. Thaler, *GenEvA (I): A New framework for event generation*, *JHEP* **12** (2008) 010, [arXiv:0801.4026].

[14] M. Moretti, T. Ohl, and J. Reuter, *O'Mega: An Optimizing matrix element generator*, hep-ph/0102195.

[15] T. Ohl, *Vegas revisited: Adaptive Monte Carlo integration beyond factorization*, *Comput.Phys.Commun.* **120** (1999) 13–19, [hep-ph/9806432].

[16] T. Ohl and J. Reuter, *Testing the noncommutative standard model at a future photon collider*, *Phys.Rev.* **D70** (2004) 076007, [hep-ph/0406098].

[17] M. Beyer, W. Kilian, P. Krstonosic, K. Monig, J. Reuter, *et. al.*, *Determination of New Electroweak Parameters at the ILC - Sensitivity to New Physics*, *Eur.Phys.J.* **C48** (2006) 353–388, [hep-ph/0604048].

[18] J. Kalinowski, W. Kilian, J. Reuter, T. Robens, and K. Rolbiecki, *Pinning down the Invisible Sneutrino*, *JHEP* **0810** (2008) 090, [arXiv:0809.3997].

[19] K. Hagiwara, W. Kilian, F. Krauss, T. Ohl, T. Plehn, *et. al.*, *Supersymmetry simulations with off-shell effects for CERN LHC and ILC*, *Phys.Rev.* **D73** (2006) 055005, [hep-ph/0512260].

[20] W. Kilian, D. Rainwater, and J. Reuter, *Distinguishing little-Higgs product and simple group models at the LHC and ILC*, *Phys.Rev.* **D74** (2006) 095003, [hep-ph/0609119].

[21] A. Alboteanu, W. Kilian, and J. Reuter, *Resonances and Unitarity in Weak Boson Scattering at the LHC*, *JHEP* **0811** (2008) 010, [arXiv:0806.4145].

[22] N. D. Christensen, C. Duhr, B. Fuks, J. Reuter, and C. Speckner, *Exploring the golden channel for HEIDI models using an interface between WHIZARD and FeynRules*, arXiv:1010.3251.

[23] J. Reuter and D. Wiesler, *Distorted mass edges at LHC from supersymmetric leptoquarks*, *Phys.Rev.* **D84** (2011) 015012, [arXiv:1010.4215].

[24] W. Kilian, J. Reuter, and T. Robens, *NLO Event Generation for Chargino Production at the ILC*, *Eur.Phys.J.* **C48** (2006) 389–400, [hep-ph/0607127].

[25] T. Robens, J. Kalinowski, K. Rolbiecki, W. Kilian, and J. Reuter, *(N)LO Simulation of Chargino Production and Decay*, *Acta Phys.Polon.* **B39** (2008) 1705–1714, [arXiv:0803.4161].

[26] T. Binoth, N. Greiner, A. Guffanti, J. Reuter, J.-P. Guillet, *et. al.*, *Next-to-leading order QCD corrections to pp → b anti-b b anti-b + X at the LHC: the quark induced case*, *Phys.Lett.* **B685** (2010) 293–296, [arXiv:0910.4379].

[27] N. Greiner, A. Guffanti, T. Reiter, and J. Reuter, *NLO QCD corrections to the production of two bottom-antibottom pairs at the LHC*, *Phys.Rev.Lett.* **107** (2011) 102002, [arXiv:1105.3624].

[28] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber, *QCD Matrix Elements + Parton Showers*, *JHEP* **11** (2001) 063, [hep-ph/0109231].

[29] L. Lonnblad, *Correcting the colour-dipole cascade model with fixed order matrix elements*, *JHEP* **05** (2002) 046, [hep-ph/0112284].

[30] M. Mangano, *Merging multijet ME's with shower MC's: some studies of systematics*, talk given at *ME/MC Tuning WG Meeting*, Fermilab, Batavia, U.S.A., 15 November 2002.

[31] M. L. Mangano, M. Moretti, and R. Pittau, *Multijet matrix elements and shower evolution in hadronic collisions: W b anti-b + n jets as a case study*, Nucl. Phys. **B632** (2002) 343–362, [hep-ph/0108069].

[32] J. Alwall *et. al.*, *Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions*, Eur. Phys. J. **C53** (2008) 473–500, [arXiv:0706.2569].

[33] S. Hoeche, F. Krauss, N. Lavesson, L. Lonnblad, M. Mangano, *et. al.*, *Matching parton showers and matrix elements*, hep-ph/0602031.

[34] N. Lavesson and L. Lonnblad, *Merging parton showers and matrix elements – back to basics*, JHEP **04** (2008) 085, [arXiv:0712.2966].

[35] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, *Longitudinally-invariant k[perpendicular]-clustering algorithms for hadron-hadron collisions*, Nuclear Physics B **406** (1993), no. 1-2 187 – 224.

[36] T. Sjostrand and P. Z. Skands, *Transverse-momentum-ordered showers and interleaved multiple interactions*, Eur. Phys. J. **C39** (2005) 129–154, [hep-ph/0408302].

[37] **DELPHI Collaboration** Collaboration, P. Abreu *et. al.*, *Tuning and test of fragmentation models based on identified particles and precision event shape data*, Z.Phys. **C73** (1996) 11–60.

[38] **DELPHI Collaboration** Collaboration, P. Abreu *et. al.*, *Measurement of event shape and inclusive distributions at S**(1/2) = 130-GeV and 136-GeV*, Z.Phys. **C73** (1997) 229–242.

[39] **JADE collaboration, OPAL Collaboration** Collaboration, P. Pfeifenschneider *et. al.*, *QCD analyses and determinations of alpha(s) in e+ e- annihilation at energies between 35-GeV and 189-GeV*, Eur.Phys.J. **C17** (2000) 19–51, [hep-ex/0001055].

[40] S. Plätzer and S. Gieseke, *Dipole Showers and Automated NLO Matching in Herwig++*, arXiv:1109.6256.

[41] S. Plätzer, *Parton Showers and Radiative Corrections in QCD*. PhD thesis, Karlsruhe Institute of Technology, 2010.

[42] G. Miu and T. Sjostrand, *W production in an improved parton shower approach*, Phys. Lett. **B449** (1999) 313–320, [hep-ph/9812455].

[43] **D0 Collaboration** Collaboration, V. M. Abazov *et. al.*, *Measurement of the normalized $Z/\gamma^* \to \mu^+\mu^-$ transverse momentum distribution in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV*, Phys.Lett. **B693** (2010) 522–530, [arXiv:1006.0618].

[44] **CMS** Collaboration, S. Chatrchyan *et. al.*, *Measurement of the Rapidity and Transverse Momentum Distributions of Z Bosons in pp Collisions at sqrt(s)=7 TeV*, arXiv:1110.4973.

[45] W. T. Giele, D. A. Kosower, and P. Z. Skands, *Higher-Order Corrections to Timelike Jets*, Phys. Rev. **D84** (2011) 054003, [arXiv:1102.2126].

[46] G. P. Salam, *A Practical Seedless Infrared Safe Cone Algorithm*, arXiv:0705.2696.

A Definitions of Observables

A.1 Event shapes

The summations are always over all final state partons.

- Thrust T:
$$T = \max_{\vec{n}} \frac{\sum_i |\vec{p}_i \cdot \vec{n}|}{\sum_i |\vec{p}_i|},$$

- Thrust major T_{major}:
$$T_{major} = \max_{\vec{n}, \vec{n} \cdot \vec{n}_T = 0} \frac{\sum_i |\vec{p}_i \cdot \vec{n}|}{\sum_i |\vec{p}_i|},$$
  with the thrust axis $\vec{n}_T$.

- Thrust minor T_{minor}:
$$T_{minor} = \frac{\sum_i |\vec{p}_i \cdot \vec{n}|}{\sum_i |\vec{p}_i|},$$
  with $\vec{n}$ perpendicular to the thrust axis $\vec{n}_T$ and the thrust major axis $\vec{n}_{T_{major}}$.

- Oblateness O:
$$O = T_{major} - T_{minor}$$

- Hemisphere broadenings B:
$$B_{\pm} = \frac{\sum\limits_{\pm \vec{p}_i \cdot \vec{n}_{Thrust} > 0} |\vec{p}_i \times \vec{n}_{Thrust}|}{2 \sum\limits_i |\vec{p}_i|}$$
$$B_{max} = \max(B_+, B_-) \qquad B_{min} = \min(B_+, B_-)$$
$$B_{sum} = B_+ + B_- \qquad\qquad B_{diff} = |B_+ - B_-|$$

A.2 Jet rates

Jet algorithms are tools to organize the plethora of particles produced in a collision. This is done by grouping "similar" particles into one pseudo-particle called *jet*. The criteria can be the closeness in the geometry of the detector, leading to *cone-jet algorithms*, where all particles within a cone of "radius" R are assumed to be one jet. The measure R is given by $R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ with the pseudo-rapidity η and the azimuthal angle ϕ. For a further discussion of the problems arising from this approach see e.g. SISCone [46]. A different approach is to sequentially remove one particle after another. The procedure is to find the minimum value of the jet separations y_{ij}, y_{ib} where i and j denote the particles and then, if the smallest value is a y_{ij} both particles are removed and replaced by a combination of the

two particles. If the smallest value is y_{ib}, the particle i is removed and implicitly clustered to the beam axis. The values y_{ij}, y_{ib} are given by

$$y_{ij} = 2\min\left(E_i, E_j\right)^2 \left(1 - \cos\theta_{ij}\right).$$
$$y_{ib} = 2E_i^2 \left(1 - \cos\theta_{i\,beam}\right). \tag{A.1}$$

For hadronic collisions another popular definition is

$$y_{ij} = \left(\Delta R_{ij}\right)^2 \min\left(p_{\perp\,i}^2, p_{\perp\,j}^2\right)$$
$$y_{ib} = p_{\perp\,i}^2. \tag{A.2}$$

By consecutively applying this prescription, every event can be gradually clustered to a $2 \rightarrow 2$ process. For each step of the clustering, the y value of the last clustering gives the jet separation for the corresponding number of jets. These prescriptions for the distance-measures compose the so-called k_T-*algorithm* [35]. The algorithm can be varied by replacing the 2 in the exponent of (ΔR_{ij}) in equation (A.1). Other values that have been studied are 0 and -2, changing to 0 leads to the *Cambridge-Aachen* algorithm, while changing to -2 produces the *anti-k_T-algorithm*.