



Institut für Theoretische Physik

**Simulation of final states
in
Vector Boson Fusion**

**Simulation der Endzustände
von Vektorboson-Fusions-Reaktionen**

Diplomarbeit

Christoph Hackstein

November 2007

Referent: Prof. Dr. Zeppenfeld
Korreferent: Prof. Dr. Klinkhamer

Ich versichere, dass ich diese Arbeit selbständig verfasst und keine anderen als die angegebenen Hilfsmittel verwendet habe

Christoph Hackstein
Karlsruhe, den 7. November 2007

Als Diplomarbeit anerkannt.

Prof. Dr. Dieter Zeppenfeld
Karlsruhe, den 7. November 2007

Contents

1	Introduction	1
2	Theory	3
2.1	The Higgs mechanism	3
2.1.1	Higgs production at the Large Hadron Collider	4
2.1.2	Vector boson fusion as a tool to measure the properties of the Higgs boson	5
2.2	Quantum chromodynamics	7
2.2.1	QCD – a non abelian gauge theory	7
2.2.2	Running coupling and confinement	8
2.2.3	Jets and observables	9
2.2.4	QCD description of particle interactions	9
2.3	Parton shower	10
2.3.1	Parton branching	11
2.4	Hadronization	18
2.4.1	Feynman-Field model	19
2.4.2	The Lund string model	20
2.4.3	Cluster model	22
2.5	Underlying event	24
2.6	Decays	26
2.7	Jet definitions	26
2.7.1	Cone-type algorithms	27
2.7.2	Clustering algorithms	29
3	Analysis	33
3.1	Programs used	33
3.1.1	The program <code>vbfno</code>	33
3.1.2	The Les Houches Accord	34
3.1.3	The program <code>Herwig++</code>	36
3.2	Analysis	38
4	Vector boson fusion observables	41
4.1	Jet observables	41
4.1.1	p_T and E_T distributions of the tagging jets	42
4.2	Tagging lepton distributions	48

4.3	Cut efficiency and additional hadronic activity	49
5	Third jets	53
5.1	Rapidity of third jet relative to the tagging jets	54
5.1.1	$y^* = y_3 - \frac{1}{2}(y_1 + y_2)$	54
5.1.2	$ y^* = y^*/ y_1 - y_2 $	56
5.2	R separation of third jets to nearest tagging jet	58
5.3	Transverse momentum	59
5.4	Rapidity and transverse momentum	60
6	Central jets	63
6.1	Rapidity distributions of central jet	63
6.1.1	Rapidity of central jet relative to tagging jets	65
6.2	p_T distribution of hardest central jet	68
6.2.1	p_T distribution with respect to $ y^* $	69
6.3	Veto efficiency	71
7	p_T shifts	73
7.1	p_T shift of hardest tagging jet	73
7.2	p_T shift of 2 nd hardest tagging jet	80
7.3	p_T distributions for correlated events	88
8	Conclusion	93
A	Sample Les Houches File	97
	Bibliography	99
	Zusammenfassung	105
	Danksagung	109

Chapter 1

Introduction

The Standard Model of particle physics is a theory which describes three of the four known fundamental interactions between the elementary particles that make up all matter. It is a quantum field theory based on a $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge symmetry yielding the gauge boson fields A_μ^a , W_μ^i and B_μ , corresponding to the gluon, the electroweak gauge bosons W^\pm , Z and the photon. These gauge fields mediate the forces between the elementary particles. The matter content of the model consists of three generations of quarks and leptons. To date, almost all experimental tests of the three forces described by the Standard Model have agreed to a very high degree with its predictions.

However, the Standard Model falls short of being a complete theory of fundamental interactions, primarily because of its lack of inclusion of gravity, the fourth known fundamental interaction, but also because of the large number of numerical parameters such as masses and coupling constants that must be put into the theory "by hand", rather than being derived from first principles.

A problem arising from the description of the Standard Model as a gauge theory is the fact that the field quanta of the weak interaction, the W and Z bosons, have been found to be massive particles. Thus in order to preserve the gauge invariance of the theory, the Higgs mechanism has to be introduced. This part of the Standard Model has not been verified experimentally, especially the Higgs boson, the field quantum of the postulated Higgs field, which is a neutral scalar field has not been discovered yet.

The electroweak precision data obtained by the LEP collider between 1989 and 2000 in combination with the negative results from the direct search seem to favor a light Higgs, $114 \text{ GeV} \lesssim m_H \lesssim 166 \text{ GeV}$. A major task for the upcoming Large Hadron Collider (LHC) currently being built at CERN in Geneva, Switzerland, is the search for the Higgs boson and the measurement of its properties.

The methods for detecting the Higgs boson depend crucially on the various decay channels which are available for a given mass. For the favored mass region, the most dominant production process is gluon fusion. This is followed by vector boson fusion which has a very clear signal structure and also provides the possibility to measure the couplings of the Higgs boson to gauge bosons and fermions independently.

Predictions based on the Standard Model for hadronic particles are obtained from

matrix element calculations at parton level. These partons, the quarks, antiquarks and gluons, are not color neutral and therefore not measurable as free physical states, a phenomenon known as *confinement*. The final state of a hadron collider experiment consists of a large number of mostly hadronic particles that derive from the original partons involved in the process.

In this work the influence of a full event simulation on the observables for vector boson fusion is examined. Therefore, the first chapter gives an overview of the theoretical framework in which this work takes place.

In order to compare the predictions obtained from matrix element calculations for hadronic processes, higher order contributions, hadronization and underlying event have to be included. For this, a class of programs, called *event generators*, is used.

After a short discussion of the Higgs sector in the Standard Model the basic principles and models needed to make complete event simulations are described. Since the used models are based on Quantum Chromodynamics (QCD), the theory of the strong interaction, the basic principles of this theory are described.

The matrix elements for a process with outgoing partons form the hard process of an event simulation. Since they have to be calculated at a fixed order but higher order contributions are quite large, the most dominant subleading terms can be resummed to improve the precision of the result. This offers a possibility to include the most dominant effects of the higher order QCD corrections in an event simulation by a Markov process, the parton shower. The final state of an event thus obtains a large number of additional partons, mostly soft or collinear with the original final state particles of the matrix element. The basic principles and techniques used to implement this in an event generator are explained.

These additional particles however are also partons and therefore not color neutral either. Thus the quarks and gluons have to be turned into hadrons. This process, the hadronization, has to be modeled. Since the scale of the particles entering the hadronization does not allow perturbation theory any more, phenomenological models have to be used here. An overview of the most important models is given.

Most of the particles obtained from hadronization are unstable, so the simulation of an event needs to include their decays. In addition, for hadron colliders, the remnant of the original hadron is a colored object and must be taken care of. Thus another model must be used to treat this remnant, called the underlying event, here defined as all the activity not stemming from the hard process in the simulation.

To compare the final state of the matrix element and the final state of the full event simulation, observables have to be defined. This analysis is based on the use of jets, so an overview of the different jet definitions used in particle physics is given.

After that the actual analysis done in this work is reviewed, the used programs are described and the proceedings explained.

Then follows the discussion of the results, in chapter 4 the influence of the event simulations on the observables is investigated. In chapter 5 the additional jets stemming from the models in the event simulation are examined. After this, in chapter 6, the impact of additional jets on the vector boson fusion events is investigated. Finally, in chapter 7 the changes in the tagging jets are explored.

Chapter 2

Theory

2.1 The Higgs mechanism

The standard model of elementary particle physics is a quantum field theory¹ in which interactions between particles are described by the exchange of field quanta. These interactions are subject to the principle of gauge invariance. Electromagnetic interactions in the standard model are described by the theory of quantum electrodynamics, QED, where electrons are represented by spinors Ψ and the interaction between the charged particles by the exchange of vector bosons, the quanta of the photon field A_μ . The Lagrangian \mathcal{L}_{QED} is invariant under a local transformation $\lambda(x)$:

$$\Psi \rightarrow e^{ie\lambda(x)}\Psi, \quad A_\mu \rightarrow A_\mu + \partial_\mu\lambda(x) .$$

The photon has to be massless, since a mass-term $m^2 A^\mu A_\mu$ would break gauge invariance.

The problem is that the gauge fields for the weak interaction, the Z, W^\pm bosons, are massive and thus the symmetry of the Lagrangian has to be broken dynamically.

A possible solution to this is the Higgs mechanism, where a scalar field is postulated and the mass terms arise from the coupling of the particles to the Higgs field.

The introduced field is a scalar $SU(2)$ doublet

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \quad \text{with the potential} \quad V(\Phi) = \mu^2|\Phi^\dagger\Phi| + \lambda|\Phi^\dagger\Phi|^2 .$$

In the case $\mu^2 < 0$ the symmetry is broken spontaneously and the minimum of the potential is at

$$|\langle\Phi_0\rangle| = \sqrt{\frac{-\mu^2}{2\lambda}} \equiv \frac{v}{\sqrt{2}} .$$

In the unitary gauge Φ can be expressed by

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix} ,$$

¹For a detailed introduction see for example [1] and [2].

where H is an elementary neutral scalar field, the Higgs boson. This leads to the Lagrangian density

$$\mathcal{L}_{Higgs} = (D_\mu \Phi)^\dagger (D^\mu \Phi) - g_d \bar{Q}_L \Phi d_R - g_u \bar{Q}_L \Phi_c u_R - V(\Phi) ,$$

where

$$iD_\mu = i\partial_\mu - g_2 \vec{I}_W \vec{W}_\mu - g_1 Y B_\mu$$

is the covariant derivative with respect to the gauge fields \vec{W}, B for the weak isospin \vec{I} and the hypercharge Y and $\Phi_c = i\sigma_2 \Phi^\dagger$.

The covariant derivative describes gauge invariant couplings of the Higgs boson to the gauge fields producing mass terms for gauge bosons and fermions. The acquired mass is proportional to the coupling of the gauge boson.

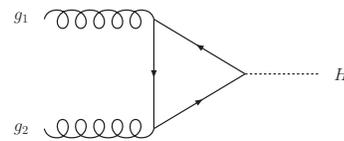
To establish the Higgs mechanism, the Higgs has to be discovered experimentally and its couplings have to be measured in order to test whether the particle really has the properties assigned to the Standard Model Higgs particle.

2.1.1 Higgs production at the Large Hadron Collider

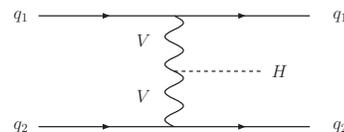
The overall picture for Higgs production at the Large Hadron Collider (LHC) at CERN depends on the mass of the Higgs boson. In case of a light Higgs, the production rate is high but the process has a lot of backgrounds whereas heavy Higgs bosons have smaller production cross sections with more manageable backgrounds.

There is only a limited number of production processes relevant for standard model Higgs production at the LHC. These are **gluon fusion** ($gg \rightarrow H$), **vector boson fusion** ($qq \rightarrow Hqq$ via W or Z exchange) and **associated production** with vector bosons or with top quarks ($q\bar{q} \rightarrow WH, ZH$ or $gg, q\bar{q} \rightarrow t\bar{t}H$).

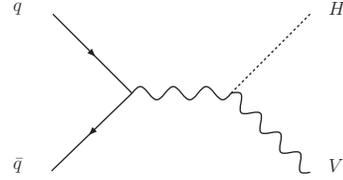
Gluon fusion, $gg \rightarrow H$ via an intermediate top-quark loop is, with the possible exception of very heavy Higgs bosons, the dominant production mechanism. This holds especially for center-of-mass energies \sqrt{s} far above the threshold due to the increase of the gluon distribution for small x . The leading order result is enhanced by the next-to-leading order contributions by a factor of ~ 2 , depending on m_H , [3].



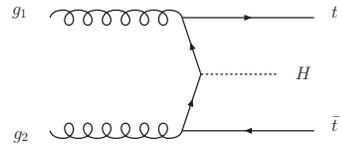
Vector boson fusion. Although gluon fusion is the dominant process for Higgs production at the LHC, the contribution from vector boson fusion (VBF) becomes comparable for very large Higgs masses. It can be visualized as elastic scattering of two quarks via the exchange of W and Z bosons. The next-to-leading order QCD corrections to these processes give only small enhancements of 5% – 10%, [4].



Associated production with vector bosons only has small cross sections for Higgs + W, Z production but may be a useful channel to find light Higgs bosons decaying to $b\bar{b}$ or $\gamma\gamma$ since the weak boson can be used to tag the Higgs. The cross section for this process is larger for $p\bar{p}$ colliders than for pp colliders. The next-to-leading order QCD corrections increase the leading order cross sections by about 20%, [5].



Associated production with top quarks As in associated production with vector bosons, a Higgs can be radiated off a top quark. The cross section depends sensitively on the top quark mass and even though the top-Higgs coupling is large for large m_t , the corresponding $Ht\bar{t}$ final state is heavy and therefore suppressed. The next-to-leading order calculations yield K -factors around 1.2 – 1.4, [6].



2.1.2 Vector boson fusion as a tool to measure the properties of the Higgs boson

The vector boson fusion process in hadron-hadron collisions is a promising Higgs discovery process at the LHC [7]. In addition, once the Higgs boson has been found and its mass has been determined, the vector boson fusion processes will be of great importance in the measurement of its couplings to gauge bosons and fermions, since it allows for independent observation of different decay channels like $H \rightarrow \tau^+\tau^-$, $H \rightarrow WW$, $H \rightarrow \gamma\gamma \dots$

The VBF processes can be measured quite exactly at the LHC since the estimated relative error in the cross section times decay branching ratio, $\sigma \cdot B$ is in the range of 5% – 10%. On the other hand the theoretical calculations of the next-to-leading order QCD corrections for these processes are quite precise with K factors² around 1.05 – 1.1. The other important process for Higgs production, gluon fusion, is more dominant with respect to the branching ratio but even the next-to-next-to-leading order predictions for the production cross section remain with uncertainties of about 10% – 20%.

A defining feature of VBF events in hadron-hadron collisions is the presence of two forward tagging jets, corresponding, at leading order, to the two outgoing quarks [8]. Since the VBF process can be considered as elastic scattering of two quarks via vector boson exchange, the transverse momentum of these jets is expected to be about half the mass of the exchanged bosons. Characteristically, one of the jets is in the forward and one in the backward region of the detector.

In order to distinguish vector boson fusion events and especially the VBF Higgs boson signal from the background, hard cuts have to be applied on the Higgs boson

²The K factor is defined as the ratio between the NLO and LO cross sections, σ_{NLO}/σ_{LO}

decay products and the two tagging jets. These tagging jets are taken to have each a transverse momentum $p_{T,j} \geq 20$ GeV and a rapidity $|y_j| \leq 4.5$.

The Higgs boson decay products are required to fall between the tagging jets in rapidity, $y_{\text{jet},\min} < \eta_l < y_{\text{jet},\max}$ and be well observable, to take this into account they are also required to pass the following cuts:

$$p_{T,l} \geq 10 \text{ GeV}, \quad |\eta_l| \leq 2.5, \quad \Delta R_{jl} \geq 0.6 .$$

For the leading order calculation of a hard process, there are only two outgoing quarks, thus these are identified as the tagging jets, but at next-to-leading order there may be more partons and of course experimental results in hadron colliders have many particles in their final states but still should show the characteristic jet behaviour. In order to compare this, a jetfinder is run on the final state particles and recombines them into jets³. By this, bundles of particles stemming from one parton in the matrix element are recombined to yield jets with the properties of the final state partons in the matrix element. In case of more than two final state jets in the event, there are several possibilities to define the tagging jets. One would be to take the two jets with the highest transverse momentum, ensuring that the tagging jets are part of the hard scattering, another would be to take the two jets with the highest energy, which favors the very energetic forward jets typical for VBF processes.

To further suppress the backgrounds, especially the ones from QCD processes for Higgs production, further cuts are applied to the two tagging jets.

Due to the fact that the exchanged particle in vector boson fusion is a color singlet state, there is no color connection between the final state jets. Thus there is no radiation of partons off the exchanged particle and therefore no jet activity in the central region stemming from the hard process. This leads to a rapidity gap between the two tagging jets. Results from Tevatron show that a possible color singlet exchange in the one-loop level gluon fusion is suppressed by a factor of at least 10^{-2} . So a further cut requiring a rapidity gap $|y_1 - y_2| > 4$ between the tagging jets causes a big improvement in reducing the QCD background. Requiring the two tagging jets additionally to be in opposite detector hemispheres ($y_1 \cdot y_2 < 0$) improves the signal to background ratio as well [9].

The above mentioned requirement for the Higgs decay products to fall between the tagging jets also removes some QCD background.

The tagging jets tend to follow the kinematics of the initial state partons, therefore the pair of tagging jets obtains a very high invariant mass $m_{j_1 j_2}$. This is in contrast to the background QCD processes, where soft and collinear gluon radiation predominates. Hence a cut on the minimum value of the invariant mass of the two tagging jets is a sensible requirement.

³More on jet definitions in chapter 2.7

2.2 Quantum chromodynamics

Experiments have shown that hadrons are not pointlike particles but consist of particles called quarks, antiquarks and gluons, the partons. The properties of these quarks gave rise to tight constraints on the nature of their interactions. The description of these interactions, called the strong interaction, gave rise to many different ideas but could only be satisfactorily described as a non-abelian gauge theory. The theory for the strong interaction is known as quantum chromodynamics, QCD. It has been tested to an overwhelming precision in experiments.

A complete discussion of QCD can for example be found in [10], [11], so here only a brief overview of the most important features is given.

2.2.1 QCD – a non abelian gauge theory

QCD is a non-abelian gauge theory with gauge group $SU(3)$. $SU(3)$ is a Lie group and so the generators of the group τ^a satisfy the corresponding Lie-algebra $su(3)$ ⁴

$$[\tau^a, \tau^b] = if^{abc}\tau^c$$

where the numbers f^{abc} are called the *structure constants*. They obey the *Jacobi identity*

$$f^{ade}f^{bcd} + f^{bde}f^{cad} + f^{cde}f^{abd} = 0 ,$$

which follows from the generators satisfying the identity

$$[\tau^a, [\tau^b, \tau^c]] + [\tau^b, [\tau^c, \tau^a]] + [\tau^c, [\tau^a, \tau^b]] = 0 .$$

The generators t^a in the fundamental representation of $su(N)$ obey the relations

$$t_N^a t_N^a = C_N \cdot, \quad C_N = \frac{N^2 - 1}{2N}, \quad \text{Tr}[t_N^a t_N^b] = \frac{1}{2}\delta^{ab} .$$

In the adjoint representation the generators T^a are given by the totally antisymmetric structure constants as

$$(T^c)_{ab} = -if^{abc}, \quad \text{Tr}[T^c T^d] = f^{abc}f^{abd} = C_A \delta^{cd}, \quad C_A = N$$

Constructing a Lagrangian density that is invariant under local $SU(3)$ transformations yields

$$\begin{aligned} \mathcal{L}_{\text{QCD}} &= \bar{q}(i\cancel{\partial} - m_q)q - g_s \bar{q}\gamma^\mu T^a q A_\mu^a - \frac{1}{4}F^{\mu\nu,a}F_{\mu\nu}^a \\ &\text{with } F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a - g_s f^{abc}A_\mu^b A_\nu^c . \end{aligned}$$

⁴here and in the following no distinction between upper and lower indices is made, since in $su(N)$ the Cartan-metric can always be chosen as δ_{ab}

The quantization of this theory is usually done by means of path integrals which needs the insertion of a gauge-fixing term which in covariant gauges can be chosen to be

$$\mathcal{L}_{\text{gauge fixing}} = -\frac{1}{2\xi}(\partial_\mu A^\mu)^2$$

as well as the insertion of Grassman-valued ghost-fields c_a to absorb unphysical polarizations of the gluon field:

$$\mathcal{L}_{\text{ghost}} = \partial_\mu \bar{c}^a \tilde{D}_{ab}^\mu c^b$$

with \tilde{D}_{ab}^μ being the covariant derivative in the adjoint representation

$$\tilde{D}_{ab}^\mu = \partial^\mu \delta_{ab} - g f_{abc} A^{c,\mu} .$$

The difference to abelian theories, like quantum electrodynamics, QED, lies in the structure of the field strength tensor $F_{\mu\nu}^a$. In abelian theories the commutator of the fields vanishes by definition whereas in non-abelian theories the commutator yields additional terms leading to self-interactions of the gauge field so the gauge bosons themselves carry the charge associated with the interaction.

2.2.2 Running coupling and confinement

The fact that gauge bosons in QCD, the gluons, also carry color charge and therefore have self-interactions leads to significant differences between QCD and QED.

In QED the strength of the interaction is given by $\alpha = \frac{g^2}{4\pi}$ and for long distances the interaction between an electron and a positron can be described by the Coulomb limit

$$V(R) = -\frac{\alpha}{R} ,$$

in the uncertainty principle one may associate

$$R \simeq \frac{1}{\sqrt{-Q^2}} ,$$

where Q^2 is the typical momentum transfer in the process considered. This approximation is valid for distances $R > 1/m_e \simeq 10^{-13}m$. For smaller $R < 1/m_e$ quantum effects become important and the potential is changed due to vacuum polarization

$$\begin{aligned} V(R) &= -\frac{\alpha}{R} \left(1 + \frac{2\alpha}{3\pi} \ln \frac{1}{m_e R} + \mathcal{O}(\alpha^2) \right) \\ &= -\frac{\bar{\alpha}}{R} \end{aligned}$$

where $\bar{\alpha}_R$ is called the *effective* or *running coupling*. An explanation for this behavior is that for large R (or small Q^2) the charge is screened by electron positron pairs created out of the vacuum.

In QCD this screening effect is also present but due to the self-coupling there is also an anti-screening effect which is stronger, so for small distances R and large Q^2 the coupling gets small. This phenomenon is known as *asymptotic freedom*. For large distances however, the potential gets dominated by a term $\propto R$ for large distances R . This may be interpreted as the reason for *confinement*. Confinement explains why quark and gluon degrees of freedom are not observed as states which propagate over macroscopic distances. A rigorous proof to this has to rely on lattice QCD.

2.2.3 Jets and observables

Quarks and gluons determine the dynamics of QCD reactions at high energies. However, these partons do not exist as free particles but only in bound states, in hadrons. Therefore a method to compare the final state of a QCD calculation, which consists of partons, and the final state of a corresponding experiment, consisting of hadrons has to be found. For this reason a *jet algorithm* is introduced, a procedure to classify any given final state according to the number of jets it contains. This measure must give cross sections which are free of soft and collinear divergences, just like total cross sections calculated in perturbation theory.

More generally, every physical observable \mathcal{O}_n must be collinear and infrared safe, so the following two properties must hold:

$$\begin{aligned} \mathcal{O}_n(p_1, p_2, \dots, p_n) &\xrightarrow[p_1 \parallel p_2]{} \mathcal{O}_{n-1}(p_1 + p_2, \dots, p_n) && \text{collinear} \\ \mathcal{O}_n(p_1, p_2, \dots, p_n) &\xrightarrow[E_1 \rightarrow 0]{} \mathcal{O}_{n-1}(p_2, \dots, p_n) && \text{soft} \end{aligned}$$

From these criteria jet algorithms as well as event shapes can be constructed. While a jet algorithm will always find jets, a shape variable only measures some particular aspect of the shape of a hadronic final state. These shapes then characterize an event for example as pencil-like, planar, spherical etc. They also can be compared with theoretical predictions.

2.2.4 QCD description of particle interactions

In general, one uses QCD to describe a reaction by ways of perturbation theory. This is a purely practical restriction since with the exception of lattice gauge theory there are no non-perturbative methods to calculate properties of QCD. Perturbation theory of course is only applicable if the strong coupling α_s is small. As mentioned above an important property of QCD is that the size of α_s varies with the size of the characteristic momentum transfer Q in a process. In leading order one has

$$\alpha_s(Q^2) \equiv \frac{g_s^2(Q^2)}{4\pi} = \frac{1}{\beta_0 \ln(Q^2/\Lambda_{\text{QCD}})}$$

where g_s is the coupling introduced in the QCD Lagrangian, Λ_{QCD} the energy scale at which non-perturbative effects become important and $\beta_0 > 0$ is given by the perturbative expansion. Λ_{QCD} has been found experimentally to be $\mathcal{O}(200 \text{ MeV})$. More

qualitatively, Λ_{QCD} indicates the order of magnitude of the scale at which $\alpha_s(Q^2)$ becomes strong, thus the found value indicates a breakdown of perturbation theory for scales comparable with the masses of light hadrons.

The concept of infrared safe observables is based on the reliability of perturbative predictions. Given that the coupling is small enough to justify perturbative calculations, finite predictions are possible if for every parton emission a corresponding virtual correction is taken into account as well. In that case the infrared singularities stemming from soft and collinear gluon emission cancel.

In conclusion, the description of a scattering process in QCD is divided in an infrared safe perturbative part and an infrared sensitive part. The scale separating the two parts is called the factorization scale μ_F . The quantities describing the non-perturbative parts of the process are on the one hand the parton-density functions (PDF) describing the parton distributions in the incoming hadron and on the other hand the fragmentation function describing the distribution of hadrons emerging from a parton.

In the case that, like in QCD, the soft and collinear divergences can be factorized from an infrared safe observable one can argue as in the renormalization that by evaluating the PDFs or fragmentation functions at a given scale μ_F the divergences can be subtracted. This leads to an evolution of the distributions between two scales which is calculable perturbatively, the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi equations (DGLAP equations, [12]) for parton distributions (and analogous equations for the fragmentation functions)

$$t \frac{\partial}{\partial t} f_\alpha(x, t) = \frac{\alpha_s(t)}{2\pi} \sum_{\alpha \rightarrow \beta\gamma} \int_x^1 \frac{dz}{z} P_{\alpha \rightarrow \beta\gamma} \left(\frac{x}{z} \right) f_\beta(z, t) ,$$

where α, β, γ describe the partons, namely the gluons and the n_f quarks, and the $P_{\alpha \rightarrow \beta\gamma}$ are the regularized splitting functions, which can be obtained from perturbative QCD.

2.3 Parton shower

Since the work for higher order QCD calculations increases roughly factorially with the order and the higher dimensional phase space gets arbitrarily complex, only a limited number of high order terms have been calculated. But there are cases where the higher-order terms are enhanced in some phase space regions, for example in deep inelastic scattering when considering collinear parton emission from the struck parton. This leads to logarithmic scaling violation in the structure functions ([10], Chapter 4.3).

Instead of making precise predictions at a fixed order in perturbation theory one can try to get an approximate result in which enhanced terms are taken into account to all orders. The concept can be illustrated by a comparison of QCD and QED. In electrodynamics accelerated charges radiate photons whereas in QCD accelerated colors radiate and create gluons, which carry color as well and therefore also radiate,

leading to a cascade of partons, the *parton shower* [13], [14]. The parton shower represents an approximative perturbative treatment of QCD dynamics above a given momentum transfer squared t , typically $\mathcal{O}(1 \text{ GeV}^2)$. In principle the parton shower converts highly virtual, primary partons into low virtuality, final state partons, either of positive virtuality just prior to hadronization or negative virtuality partons emerging from a beam hadron entering the hard scattering process. The generation of the incoming, space-like, and outgoing, time-like, parton showers is an iterative Markov branching process based on the use of the the DGLAP equations of the fragmentation functions. These equations sum the leading effects of repeated parton branchings to all orders and thus improve the convergence of the perturbative calculation. The enhanced higher order terms appear for processes like soft gluon emission and the splitting of a gluon or a light quark into two almost collinear partons. These terms are associated with large logarithms of the scale, thus the inclusion of these enhanced configurations is called the *Leading Log Approximation*.

2.3.1 Parton branching

Collinear enhancements are associated with *parton branching* on an incoming or outgoing line of a QCD Feynman diagram. The branchings are given by the splitting functions which can be calculated from the vertices in the QCD Lagrangian. They form the evolution kernels for the DGLAP equations.

For an outgoing parton the branching is timelike because then the momentum transfer t defining the scale of the process is greater than zero, $t > 0$. The opening angle for this branching can be calculated in terms of t , E and z , where z denotes the energy fraction of the outgoing partons b, c with respect to the energy of the incoming parton a , $z = E_b/E_a = 1 - E_c/E_a$.

For the case that the three partons are gluons, $g \rightarrow gg$, the splitting is given by the triple gluon vertex V_{ggg} from the Lagrangian, a factor $1/t$ from the propagator of the incoming gluon, and the contribution of the polarization. Since the vertex factor is proportional to the opening angle θ and $t \propto 1/\theta^2$, the amplitude has a singularity proportional to $1/\theta$. In the small angle region the amplitude can be factorized

$$|\mathcal{M}_{n+1}|^2 \sim \frac{4g^2}{t} P_{gg} |\mathcal{M}_n|^2$$

where the DGLAP kernel P_{gg} , called gluon splitting function, is given by the color factor $C_A = 3$, the energy fraction z of the outgoing parton and the polarization of the gluons:

$$P_{gg}(z) = C_A \left[\frac{1-z}{z} + \frac{z}{1-z} + z(1-z) \right].$$

The enhancements for this splitting come from the cases where $z \rightarrow 0$ or $z \rightarrow 1$, corresponding to soft emission of a gluon.

In the case of a gluon branching into a quark-antiquark pair $g \rightarrow q\bar{q}$ the spinor structure of the vertex has to be taken into account as well. The amplitude for small

angles can again be factorized and the splitting function is given by

$$P_{qg}(z) = T_R [z^2 + (1 - z)^2] ,$$

where the color factor $T_R = 1/2$. Here no soft singularities appear ($z \rightarrow 0$ or 1).

For the branching where a quark emits a gluon, $q \rightarrow qg$, the small angle matrix element is again factorizable and the splitting function P_{qq} is given by

$$P_{qq}(z) = C_F \frac{1 + z^2}{1 - z} ,$$

with the color factor $C_F = 4/3$. Here again the splitting shows a soft singularity as $z \rightarrow 1$ which is strongly dependent on the polarization, as only the amplitude for gluon polarization in the plane of branching develops this singularity.

These angular correlations can be measured in e^+e^- collisions, where the Bengtson-Zerwas-angle in fourjet-events, defined as the angle between the planes of the two hardest and two softest jets [16], is sensitive to these contributions⁵.

To get the complete cross sections for the different processes, the phase space for the n -parton final state $d\Phi_n$ and the flux factor \mathcal{F} have to be taken into account as well. The result is

$$d\sigma_{n+1} = d\sigma_n \frac{dt}{t} dz \frac{\alpha_s}{2\pi} P_{ji}(z) ,$$

where P_{ji} is the appropriate splitting function.

As mentioned above, for multiple branching the enhancement of higher order contributions is associated with multiple small angle parton emission and is summed up by the DGLAP equations.

It is now possible to introduce regularized splitting functions and solve the DGLAP equations analytically via Mellin transformations or in order to study more detailed features of the branching process and structure of the final state a numerical Monte Carlo approach can be used.

In an axial gauge the DGLAP equations correspond to a sum of ladder diagrams [17] whereas interference or crossed-rung diagrams give sub-leading contributions. This makes it possible to describe the process in the parton model language. Basically, each parton is assigned a set of probabilities for its possible branchings, including the possibility for no branching as well, and is randomly evolved. This probabilistic development can be described by a classical Markov process.

The problem is that the splitting kernels P_{ji} are singular and have to be regularized before the DGLAP equations can be treated numerically. This can be achieved by introducing suitable cut-offs on the soft gluon momenta providing correctly normalized branching probabilities. In an analytic calculation the regularization would be achieved due to the fact that the virtual and real emission divergences at a fixed order cancel, yielding infrared finite observables. In the same manner the subtraction terms here follow from unitarity and the infrared finiteness of inclusive observables.

⁵This is also a test for the non-Abelian nature of QCD

Timelike branching

A time-like Monte Carlo algorithm can be constructed by introducing the Sudakov form factor

$$\Delta_i(t, t_0^i) \equiv \exp \left[- \sum_j \int_{t_0^i}^t \frac{dt'}{t'} \int_{\epsilon(t')}^{1-\epsilon(t')} dz \frac{\alpha_s(t, z)}{2\pi} P_{ji}(z) \right]$$

where all possible branching processes $i \rightarrow j$ are summed up, and t_0^i is a cut-off associated with the parton i . The cut-off $z < 1 - \epsilon(t')$ is introduced to take the infra-red singularities into account.

Using this form factor an integral equation for the parton distribution $f(x, t)$ can be obtained

$$f_i(x, t) = \Delta_i(t, t_0^i) f_i(x, t_0^i) + \int_{t_0^i}^t \frac{dt'}{t'} \Delta_i(t, t') \int_x^{1-\epsilon(t')} \frac{dz}{z} \frac{\alpha_s(t', z)}{2\pi} P_{ji}(z) f(x/z, t')$$

where the relations $\Delta_i(t_0^i, t_0^i) = 1$ and $\Delta_i(t, t_0^i) / \Delta_i(t', t_0^i) = \Delta_i(t, t')$ were used.

This set of inhomogeneous integral equations can be solved by repeated back substitutions, yielding a von-Neumann series solution which forms the basis of the Monte Carlo implementation of final state time-like parton cascades.

$$\begin{aligned} f_i(x, t) &= \Delta_i(t, t_0^i) f_i(x, t_0^i) \\ &+ \int_{t_0^i}^t \frac{dt_1}{t_1} \Delta_i(t, t_1) \int_x^{1-\epsilon(t_1)} \frac{dz_1}{z_1} \frac{\alpha_s(t_1, z_1)}{2\pi} P_{ji}(z_1) \Delta_j(t_1, t_0^j) f\left(\frac{x}{z_1}, t_0^j\right) \\ &+ \int_{t_0^i}^t \frac{dt_1}{t_1} \Delta_i(t, t_1) \int_x^{1-\epsilon(t_1)} \frac{dz_1}{z_1} \frac{\alpha_s(t_1, z_1)}{2\pi} P_{ji}(z_1) \int_{t_0^j}^{t_1} \frac{dt_2}{t_2} \Delta_j(t_1, t_2) \\ &\quad \times \int_{x/z_1}^{1-\epsilon(t_2)} \frac{dz_2}{z_2} \frac{\alpha_s(t_2, z_2)}{2\pi} P_{jk}(z_2) \Delta_k(t_2, t_0^k) f\left(\frac{x}{z_1 z_2}, t_0^k\right) \\ &+ \dots \end{aligned}$$

This formula for the parton distribution has a direct probabilistic interpretation. The first term on the right-hand side is the contribution from configurations where no branching between the scales t and t_0 occurs, so the Sudakov form factor $\Delta_i(t, t_0^i)$ gives the probability for a parton i to evolve between these two scales without resolvable radiation.

The second term represents the case that the initial parton i evolves to an intermediate scale t_1 at which a resolvable branching $i \rightarrow jj'$ occurs. Parton j receives the fraction $z_1 > x$ of the momentum of parton i and evolves without further branching to its cut-off scale t_0^j .

The same holds for the further terms, where the contributions for additional resolvable partons are taken into account. The sum of all these configurations gives a solution to the DGLAP equations.

The implementation for this process in a Monte Carlo simulation is done by sampling the Markov process defined by the DGLAP equations. The process starts with an off-shell parton i of scale $t > t_0^i$ and the Sudakov form factors $\Delta_{i \rightarrow jj'}(t, t_0^i)$ are used to select an intermediate scale at which a specific branching $i \rightarrow jj'$ can occur. In case this scale is below t_0^i , the parton is set on mass-shell, $p_i = t_0^i$ and the evolution stops, if a branching occurs, the momentum fraction of the daughter particles is selected by $\alpha_s(t, z)P_{ji}(z)$ and their scales t_j, t'_j are derived from t, z . This procedure is repeated until all particles are below the threshold and on mass-shell.

Spacelike branching

Up until now only forward evolution was considered where an outgoing parton with timelike momentum emits further partons and moves to lower virtual mass-squared. This is a good method for timelike evolution like the final state of e^+e^- collisions.

For spacelike cascades it is more convenient to start by specifying the momentum fraction x_n of the partons entering the hard process that is used for the matrix element calculation. Evolving the parton backwards from the hard-scattering scale guided by the PDFs already used to choose the hard process to the low-scale incoming hadron gives in every case the desired final state of the cascade, unlike the case of forward evolution, which would only in a few cases yield the right configuration and thus make a simulation rather inefficient.

If the same description as for the forward evolution is used for the backward evolution, the cases where the backward evolution from a given x_2 to x_1 yields a fraction of the time unphysical configurations where $x_1 > 1$. These have to be rejected and thus lead to lower efficiency.

The correct description for backward evolution uses a modified form factor, taking the local parton density $f(x, t)$ into account when choosing the next value of the evolution variable t

$$\Pi_i(t, t_s^i; x) = \frac{f_i(x, t_s^i)}{f_i(x, t)} \Delta_i(t, t_s^i)$$

This can be interpreted as the probability that the parton i in a hadron evolves from the scale t backwards to the scale t_s^i with the same momentum fraction x and without

resolvable parton emission. This leads to the equation

$$\begin{aligned}
1 &= \Pi_i(t, t_s^i; x) \\
&+ \int_{t_s^i}^t \frac{dt_1}{t_1} \int_x^{1-\epsilon_{j'}^s(t_1)} \frac{dz_1}{z_1} \Pi_i(t, t_1; x) \frac{\alpha_s(t_1, z_1)}{2\pi} P_{ij}(z_1) \frac{f_{j/h}(\frac{x}{z_1}, t_1)}{f_{i/h}(x, t_1)} \Pi_j(t_1, t_s^j; x) \\
&+ \int_{t_s^i}^t \frac{dt_1}{t_1} \int_x^{1-\epsilon_{j'}^s(t_1)} \frac{dz_1}{z_1} \Pi_i(t, t_1; x) \frac{\alpha_s(t_1, z_1)}{2\pi} P_{ij}(z_1) \frac{f_{j/h}(\frac{x}{z_1}, t_1)}{f_{i/h}(x, t_1)} \\
&\quad \times \int_{t_s^j}^{t_1} \frac{dt_2}{t_2} \int_{x/z_1}^{1-\epsilon_{k'}^s(t_2)} \frac{dz_2}{z_2} \Pi_j\left(t_1, t_2; \frac{x}{z_1}\right) \frac{\alpha_s(t_2, z_2)}{2\pi} P_{jk}(z_2) \\
&\quad \times \frac{f_{k/h}(\frac{x}{z_1 z_2}, t_2)}{f_{j/h}(\frac{x}{z_1}, t_2)} \Pi_k\left(t_s, t_s^k; \frac{x}{z_1 z_2}\right) \\
&+ \dots
\end{aligned}$$

The PDF fractions accompanying the splitting functions guide the evolution towards the correct parton content. This equation can now be interpreted as a normalized sum of the probabilities for all chains of branchings that take a given parton i at scale t back to an initial parton at scale t_s^i . The first term on the right hand side gives the probability that the parton evolved from t_s^i without resolvable emission, the second term gives the probability that the parton i evolved to t from the scale t_1 without resolvable radiation, where it had been produced in the branching of a parton b of momentum fraction x/z_1 which had come from the scale t_s^b without resolvable emission, and so on.

The numerical implementation of this is similar to that in the timelike case. Given a parton i at scale t , the modified Sudakov form factor $\Pi_i(t, t_s^i)$ is used to select a branching scale. If this scale is below the cut-off, no resolvable branching is assumed to have occurred, the particle is set on mass-shell, $p_i = t_s^i$ and the branching stops. Else in case of branching the type $j \rightarrow ii'$ and the momentum fraction are chosen according to $\alpha_s(t, z) P_{ij}(z) f_{j/h}(x/z, t)$. The above procedure is repeated for j whereas i' undergoes timelike branching.

In the case of gluons or strange quarks entering the hard process, a minimum number of branchings has to occur in order to give the correct flavors⁶. In case the scale is already below the cut-off scale before this amount of branchings has taken place, a non-perturbative model for this must be used.

Coherent (soft) branching

So far only logarithmic enhancements stemming from soft-collinear and collinear enhancements have been considered. But there are also enhancements associated with soft gluon emission. This can be seen from the singularities in the small-angle parton splitting functions for soft gluon emission.

⁶Consider for example p: $(uu)d \rightarrow dg, g \rightarrow s\bar{s}$

The emission of a gluon of momentum q from an external line with momentum p of a QCD Feynman graph gives a propagator factor

$$\frac{1}{(p \pm q)^2 - m^2} = \frac{\pm 1}{2p \cdot q} = \frac{\pm 1}{2\omega E(1 - v \cos \theta)}$$

where ω is the energy of the gluon, E and v the energy and velocity of the emitting parton and θ the angle of emission. In addition to the collinear enhancement for $\theta \rightarrow 0$ there is also one for $\omega \rightarrow 0$ for any velocity and angle.

This soft enhancement corresponds to a color factor times a universal, spin-independent factor of $F_{soft} = p \cdot \epsilon / p \cdot q$ in the amplitude⁷, where ϵ is the polarization vector of the emitted gluon.

This enhancement factor in the amplitude for each external line implies that the cross section has a factor which is the sum over all pairs of external lines i, j

$$d\sigma_{n+1} = d\sigma_n \frac{d\omega}{\omega} \frac{d\Omega}{2\pi} \frac{\alpha_s}{2\pi} \sum_{i,j} C_{ij} W_{ij}$$

where $d\Omega$ is the solid angle of the gluon emission, C_{ij} a color factor and W_{ij} is the radiation function given in the case of massless partons by

$$W_{ij} = \frac{\omega^2 p_i \cdot p_j}{p_i \cdot q p_j \cdot q} = \frac{1 - \cos \theta_{ij}}{(1 - \cos \theta_{iq})(1 - \cos \theta_{jq})}.$$

The color weighted sum in the cross section is called the *antenna pattern* of the process.

The radiation function can be separated into two parts containing the leading collinear singularities for emissions from particles i and j , respectively,

$$W_{ij} = W_{ij}^{[i]} + W_{ij}^{[j]},$$

with
$$W_{ij}^{[i]} = \frac{1}{2} \left(W_{ij} + \frac{1}{1 - \cos \theta_{iq}} - \frac{1}{1 - \cos \theta_{jq}} \right).$$

This function has the property of *angular ordering*. Carrying out the azimuthal part of the angular integration, taking the parton i as reference direction yields

$$\int_0^{2\pi} \frac{d\phi_{iq}}{2\pi} W_{ij}^{[i]} = \begin{cases} \frac{1}{1 - \cos \theta_{iq}} & \text{if } \theta_{iq} < \theta_{ij} \\ 0 & \text{else} \end{cases}$$

This means that soft radiation in $W_{ij}^{[i]}$ is only emitted inside a cone of opening angle θ_{ij} . The same applies for the contribution $W_{ij}^{[j]}$ with i and j exchanged.

The angular ordering property is a coherence effect common to all gauge theories, in electrodynamics it accounts for the suppression of soft bremsstrahlung from

⁷This enhancement occurs only for on-shell partons since for internal propagators the denominator factor $(p + q)^2 - m^2 \rightarrow p^2 - m^2 \neq 0$

electron-positron pairs, the *Chudakov effect*, which can also be explained heuristically by time-ordered perturbation theory.

In a suitable representation of the color algebra color charges of the partons can be represented by vectors \mathbf{Q}_i such that $\mathbf{Q}_i^2 = C_F$ for a quark, $\mathbf{Q}_i^2 = C_A$ for a gluon and $\mathbf{Q}_i^2 = 0$ for a singlet. With this, the color factor in the antenna pattern becomes $C_{ij} = -\mathbf{Q}_i \cdot \mathbf{Q}_j$.

For the case $e^+e^- \rightarrow q\bar{q}g$ with $\mathbf{Q}_i + \mathbf{Q}_j + \mathbf{Q}_k = 0$ this leads to

$$W = -\mathbf{Q}_i\mathbf{Q}_jW_{ij} - \mathbf{Q}_j\mathbf{Q}_kW_{jk} - \mathbf{Q}_i\mathbf{Q}_kW_{ik}$$

In case the partons i, j are close in angle, they form a system l with resulting net color charge $\mathbf{Q}_l = \mathbf{Q}_i + \mathbf{Q}_j = -\mathbf{Q}_k$. Using the above decomposition into leading collinear singularities and introducing the terms $\tilde{W}_{jk}^{[i]} = \frac{1}{2}(W_{ik}^{[i]} - W_{ij}^{[i]})$, W can be approximated as

$$W \simeq \mathbf{Q}_i^2 W_{ij}^{[i]} + \mathbf{Q}_j^2 W_{ij}^{[j]} + \mathbf{Q}_k^2 W_{lk}^{[k]} + \mathbf{Q}_l^2 \tilde{W}_{lk}^{[ij]}$$

The interpretation for this is straightforward: each parton i, j and k radiates proportional to its color charge squared. When two partons i and j are close in angle, after azimuthal averaging their incoherent contributions are limited to cones of half-angle θ_{ij} . At larger angles, out of the direction of k the coherent contribution is proportional to the combined color charge squared \mathbf{Q}_l^2 which is the same contribution as from an internal line with momentum $p_l = p_i + p_j$. So k is only able to resolve the net color charge \mathbf{Q}_l .

This can be extended to higher orders, yielding the *coherent parton branching* formalism. With this it is possible to calculate soft gluon enhancement to all orders. To do this, the parton shower needs to be modified a little in order to impose angular ordering in the shower. This means that partonic emissions are only allowed inside the cone specified by the previous emission. Instead of using the virtuality t as evolution variable, now

$$\zeta = \frac{p_b \cdot p_c}{E_b E_c} \simeq 1 - \cos \theta$$

is used for the branching $a \rightarrow bc$ with $dt/t = d\zeta/\zeta$. Imposing angular ordering on the shower algorithm then simply translates to $\zeta' < \zeta$ for successive branching. An angular cut-off ζ_0 specifying the end of branching and removing infrared divergences has to be introduced as well, a good choice for this is $\zeta_0 = t_0/E^2$, keeping t_0 as minimum mass-squared.

Using also the appropriate splitting function P_{ba} instead of the soft approximation $\mathbf{Q}_a^2 d\omega/\omega$ the formalism treats both soft and collinear enhancement correctly. With this the formula for coherent branching becomes

$$d\sigma_{n+1} = d\sigma_n \frac{d\zeta}{\zeta} dz \frac{\alpha_s}{2\pi} \hat{P}_{ba}(z) .$$

The angular cut-off implies a more convenient evolution variable $\tilde{t} = E^2\zeta \geq t_0$. The angular ordering condition $\zeta_b, \zeta_c < \zeta_a$ for $a \rightarrow bc$ then reads

$$\begin{aligned} \tilde{t}_b < z^2\tilde{t}_a, \quad \tilde{t}_c < (1-z)^2\tilde{t}_a \\ \sqrt{\frac{t_0}{\tilde{t}}} < z < 1 - \sqrt{\frac{t_0}{\tilde{t}}} \end{aligned}$$

Putting all this together, the coherent, angular-ordered Sudakov form factor is

$$\tilde{\Delta}_q(\tilde{t}) = \exp \left[- \int_{4t_0}^{\tilde{t}} \frac{dt'}{t'} \int_{\sqrt{t_0/t'}}^{1-\sqrt{t_0/t'}} dz \frac{\alpha_s(z^2(1-z)^2t')}{2\pi} P_{qq}(z) \right]$$

For large \tilde{t} it falls more slowly than the original Sudakov form factor, which implies less branching due to the suppression of soft gluon emission.

2.4 Hadronization

After the partons that were produced in the hard process far off mass shell have evolved in the parton shower, one is left with the final state of a cascade of partons, all near mass-shell at the cut-off scale t_0 . The next step in the simulation of an event is however not so clear, since now, due to the running of the strong coupling α_s , perturbation theory does not work any more and further calculations involve the long-distance, non-perturbative properties of the theory. These somehow lead to the confinement of partons into colorless bound states, the hadrons, which are the final state particles measured in an experiment.

Since non-perturbative techniques are far from providing enough understanding of confinement to allow the calculation of final state hadron distributions in jets from first principles, one has to resort to phenomenological models for this. A number of approaches for this have been tried with some success, the **Feynman-Field model** [18], the **Lund string model** [19] and the **Cluster model** [22]. In addition, the unstable hadrons produced by these models are decayed subsequently into stable hadrons, leptons and photons, according to pretabulated decay tables, measured in experiments. So the fragmentation function for finding a hadron h in a jet initiated by an outgoing primary parton a at a scale Q carrying the fraction x of the initial parton's energy is given by

$$\begin{aligned} D_a^h(x, Q^2) = & (QCD \text{ evolution} : Q^2 \rightarrow t_0) \otimes (\text{hadronization model} : b \rightarrow H)|_{t_0} \\ & \otimes (\text{decay tables} : H \rightarrow h, h', \dots) \end{aligned}$$

In principle, the shower cut-off t_0 is an arbitrary parameter and not connected with hadronization, for which the intrinsic energy scale is presumably of $\mathcal{O}(\Lambda_{\text{QCD}})$. Of course the outcome of the hadronization, i.e. the configuration of final state particles should not depend on the exact value of t_0 . The problem is now that for an increased

value of t_0 the shower terminates earlier and thus produces less partons with higher virtualities to hadronize. So ideally the hadronization model should as well have a parameter t_0 whose effect cancels when both the shower and the hadronization are combined.

The present understanding of hadronization and its effects is only based on models and not QCD as such. Since hadronization is expected to be a local effect without involving large momentum transfers. So it seems reasonable that the hard process and the parton shower calculations are the dominant parts in determining the overall features of the process, like energy dependences, event shapes and so on. Nevertheless the effects of hadronization are not negligible, as can be seen from the results in e^+e^- -collisions for event shape variables that are sensitive to out-of-plane activity.

2.4.1 Feynman-Field model

This is the simplest model simulating hadron production and is based on independent fragmentation. The assumption is that each parton fragments independently into hadrons, based on the observations of quark jets produced in e^+e^- colliders at moderate energies.

The model is constructed explicitly to reproduce the limited transverse momenta and scaling of energy fraction distributions in these experiments.

Observations from $e^+e^- \rightarrow 2$ jets experiments show that the number of produced hadrons forms a flat rapidity plateau and has limited transverse momentum, $\rho(p_t^2) \propto \exp(-p_t^2/2p_0^2)$ so using this simple model one can estimate the hadronization correction to perturbative quantities, arriving at a rather large non-perturbative mass M^2 for the jets, about 10 GeV for a 100 GeV jet.

The model combines each fragmenting quark a with an anti-quark from a $b\bar{b}$ pair created out of the vacuum, due to the color field present. This results in a 'first generation' meson $a\bar{b}$ carrying the energy fraction z of the

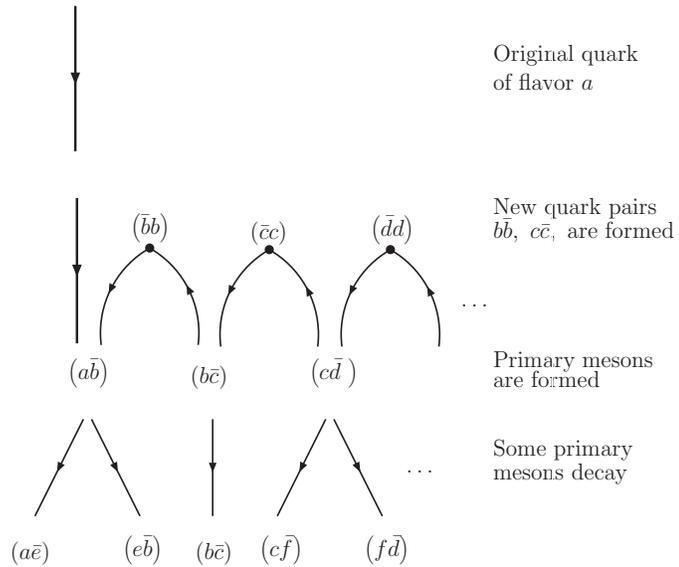


Figure 2.1: Schematic representation of the feynman-field model.

original quark which then may either be observed directly as a pseudoscalar meson or it may be a vector or higher spin unstable resonance which subsequently decays into the observed mesons. The remaining quark b carrying the energy fraction $(1 - z)$ is fragmented in the same way. The underlying chain-decay ansatz of the model now as-

sumes that the hadrons resulting in the remaining cascade that starts with the quark b are distributed in exactly the way as the hadrons stemming from a jet originated by a quark of type b with initial momentum given by the fraction $(1 - z)$ of the original momentum. Furthermore for very high momenta, all distributions are taken to scale such that they depend only on ratios of the hadron momenta to the quark momenta.

This procedure is done with all further quarks created in that cascade until the energy fraction of the leftover quark is below some cut-off.

In the case of gluons the fragmentation starts with splitting the gluon into a $q\bar{q}$ pair, either assigning all the gluons energy to one of the two quarks or using the Altarelli-Parisi splitting function for $g \rightarrow q\bar{q}$.

The resulting model then determines the structure of quark jets only using three parameters used to describe flavor, primary meson spin and transverse momentum and a function $f(\eta)$ defined by the property that $f(\eta)d\eta$ is the probability that the 'first generation' meson leaves the momentum fraction η to the remaining cascade, with $\int_0^1 f(\eta)d\eta = 1$. Taking $f(\eta)$ to be a gaussian distribution, the model gives good agreement with data from e^+e^- experiments at moderate energies.

A problem in this hadronization model is that the fragmentation of the partons depends on their energy rather than their virtuality, with the particles assumed to be on mass shell the fragmentation leads to violations of momentum-conservation that have to be corrected by rescaling the final state momenta, making this model also strongly frame dependent. Also each jet ends with a leftover quark, whose color and flavor has to be neutralized, so quantum number conservation has to be built in additionally.

Furthermore, the model has no obvious relation with perturbative emission, is no model of confinement and is not infrared or collinear safe as jets with a small opening angle do not merge smoothly to one jet if the included angle goes to zero but remain distinguishable as two jets.

2.4.2 The Lund string model

The string model is motivated by the properties of the color field between two colored particles. For example a quark and an antiquark, produced in a e^+e^- annihilation move out in opposite directions and loose energy to the colorfield between them. This field is supposed to collapse due to self-interactions into a stringlike configuration with a uniform energy per unit length, or string tension κ with $\kappa \approx 1 \text{ GeV}/\text{fm}$, as is described by the linear quark potential described by confinement in QCD. Since the transverse size of the string $\langle r_{\perp}^2 \rangle = \pi/(2\kappa)$ is negligible, it is plausible to describe its dynamics by a massless, one-dimensional, relativistic string without transverse excitations. The equations of motion for the string then lead to the so called 'yo-yo modes', where a massless quark-antiquark pair forming the endpoints of the string, seen from the string's center of mass frame, oscillate repeatedly outwards and inwards at the speed of light, passing through each other and transferring energy to and from the string [20].

The string then breaks due to spontaneous $q\bar{q}$ pair production in the intense color

field, subsequently forming hadron-sized pieces. The choice which hadrons are formed that way is based on an attempt to model, via flavor and spin selection rules, the supposed dynamics of the process.

For quark-antiquark pairs, the model works similar to the independent fragmentation scheme, the string between the quarks may be broken at either the quark or the antiquark end, or at both ends simultaneously, proceeding iteratively by the creation of $q\bar{q}$ pairs like in the Feynman-Field model [21]. Since the breaking points have spacelike separation, the progression depends on the frame in which it is done. The improvement compared to the Feynman-Field model is that the string model, together with constraints on the fragmentation function, gives a more consistent and covariant picture of hadronization, ensuring independence of the starting point for the fragmentation as well as the transverse momentum distribution which is related to the quantum mechanic tunneling mechanism of $q\bar{q}$ pair production in the string. The model provides several tuneable parameters, namely in the functions defining the left- right symmetry and tunneling probability.

Meson production now is rather straightforward from a quark-antiquark pair connected by a string without any further breaking whereas baryon production is somewhat more difficult and not so well understood. In principle it can be pictured as a configuration with three quarks attached by strings to a common center but there are several different ways to do this, most commonly either the 'diquark' and the popcorn model are used for that. In the diquark model the string breaking is not imposed by a $q\bar{q}$ pair but by a $qq'\bar{q}\bar{q}'$ or diquark pair with the same color configuration as the endpoint quarks. Alternatively, if a quark-antiquark pair created in the middle of

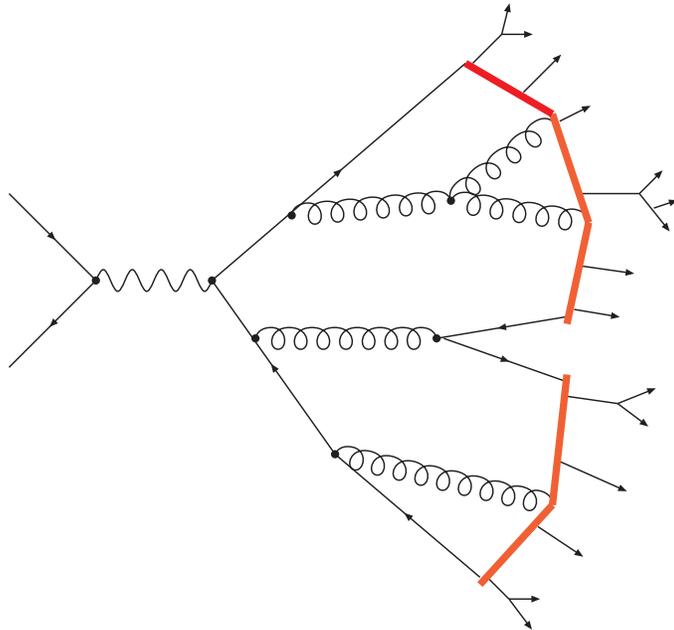


Figure 2.2: Schematic representation of the string model. Quarks form endpoints of strings whereas gluons produce kinks in them.

the string has a different color configuration, then an anti-aligned, color triplet gluon field remains between them, inducing the possibility for further $q\bar{q}$ pairs to form which allows the string to break before the first virtual $q\bar{q}$ pair recombines. This is the popcorn model. In both cases the remaining string fragment has a three-quark configuration and therefore nonzero baryon number. The two models illustrate the ambiguity whether a baryon is regarded as a quark-diquark bound state or a three

quark configuration.

With the introduction of a gluon, the string model becomes more than a motivated and constrained independent fragmentation model or improved Feynman-Field model. Now gluons are not split into $q\bar{q}$ pairs at the beginning but are represented as kinks on the string, each carrying localized energy and momentum, given by that of the initial gluon. These kinks in the string lead to changes in the angular distribution of the produced hadrons compared to the Feynman-Field model and give for the case of e^+e^- three jet final states a better agreement with experiment. A second aspect of this treatment of gluons is that it makes infrared matching with the parton shower possible, gluons with a transverse momentum smaller than the inverse string width are irrelevant for the hadronization process because these low energetic gluons produce only small kinks on the string that are negligible.

For configurations with several final state partons, there is an ambiguity about the different ways the strings should be connected between the various possible endpoints given by the quarks and antiquarks and the kinks, given by the gluons. In leading order in N^2 however, where $N = 3$ is the number of colors, it is always possible to arrange the partons in a planar configuration where every parton has an equal and opposite color than its neighbour (or neighbours in the gluon case). It now seems most reasonable to stretch the string between the color connected neighbours. The planar approximation simplifies the calculations by discarding *nonplanar* graphs, i.e. graphs that cannot be drawn without any particle lines overlapping. These graphs are suppressed relative to the *planar* ones by powers of $1/N$, this corresponds to the above mentioned assumption that there are N^2 gluons instead of $N^2 - 1$.

The String model is a strongly physically motivated model which is very successful in fitting data. It is rather universal, once it is fitted to data from e^+e^- -collisions, there is not much freedom elsewhere to fit. The motivation for the model translates rather good to prediction, for every hadron of effect described by the model, one free parameter is used. A possible drawback is that it hides too much perturbative information, because the string formation in leading color distorts the configuration.

2.4.3 Cluster model

The cluster model makes use of a property of the parton branching process called preconfinement. The central idea is the formation of color singlet clusters of partons stemming from the hard process and the subsequent parton shower. These clusters then decay into hadrons. Preconfinement implies that pairs of colorconnected neighbouring particles have the tendency to be arranged with limited extension in both coordinate and momentum space. These color-singlet 'clusters' have an asymptotic mass distribution that falls rapidly at high masses, is asymptotically independent of the overall energy Q^2 and universal. In the planar approximation a gluon is represented by a color-anticolor pair, so the simplest way for color-singlet clusters to form is by means of a non-perturbative low q^2 enhancement of the splitting of gluons into quark-antiquark pairs, which is relatively uncommon in the perturbative branching process. Neighbouring quarks can then form color singlets.

With the right approximation [23] for the $g \rightarrow q\bar{q}$ form-factor, it is possible to include such an enhancement automatically, with a proper set of parameters all gluons will eventually decay into quark-antiquark pairs.

The mass spectrum of the formed clusters is again universal, peaking at a low mass and falling off rapidly at high masses. The precise form is shaped by the value assigned to the QCD scale

Λ_{QCD} , the cut-off scale t_0 and also, though only to a lesser extent, on the gluon-splitting mechanism. It is independent, however, of the center of mass energy of the collision. The spectrum is likely to represent a smeared version of "primordial resonances", formed in the early stage of confinement of real jets. Hence it seems reasonable to treat the cluster fragmentation as an averaged resonance decay with very simplified dynamics, which leads to quasi-two-body decay, usually predominating for known resonances and branching ratios determined by density of states, (phase space times spin degeneracy) and no spin correlations, leading to isotropic decay. The reduced phase space for cluster decay into heavy mesons and baryons is then sufficient to model the multiplicities of the various hadrons produced in e^+e^- collisions. So the probability for a cluster to decay into a given pair of hadrons h_1 and h_2 is given by

$$\mathcal{P}(Cl \rightarrow h_1 + h_2) = (2J_{h_1} + 1)(2J_{h_2} + 1) p(m_{Cl}, m_{h_1}, m_{h_2})\theta(m_{Cl} - m_{h_1} - m_{h_2})$$

where the J_i are the spin of the hadrons, m_i the masses, p the c.o.m. frame three-momentum in the two-body decay and the Heavyside function guarantees that the decay is allowed physically.

Even without the introduction of an adjustable fragmentation function the energy and transverse momentum distributions of the hadrons agree quite well with experimental data and if soft gluon interference is taken into account the angular distribution in e^+e^- three-jet events is described effectually, as in the string model.

For very heavy clusters, typically $m_{Cl} > 4 \text{ GeV}$, isotropic decay is an unreasonable assumption, so for these clusters an anisotropic fission mode is used. Since only a fraction of about 15% of the clusters lie beyond the fission threshold, the explicit form of the fission algorithm is not crucial and a simple one can be chosen, like the "symmetrical string breaking". This corresponds to a string with a given energy density between quark and diquark, which breaks in the middle by quark antiquark pair production, where the flavour of the produced pair is taken to be u, d , or s . This is repeated, until all decay products are below the fission threshold, where the above mentioned cluster decay takes place. The threshold for cluster fission becomes

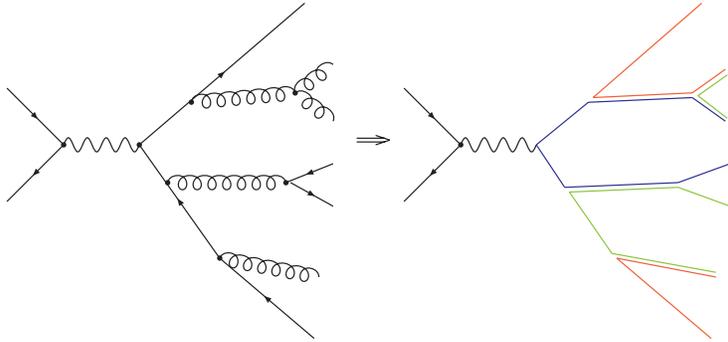


Figure 2.3: Preconfinement. Quarks can be represented by one, gluons by two color lines.

a rather crucial parameter in this model, since even though only 15% of the clusters get split about 50% of the hadrons in the final state stem from these clusters.

For very light clusters which are too light to undergo two-body decay one assumes one-body decay and redistributes the excess momentum amongst neighboring clusters.

In case the clusters contain heavy quarks, namely c or b quarks, the decay model must be modified since the data of heavy flavor resonance multiplets is incomplete. The heavy quark is assumed to undergo free-particle β -decay, giving two hadronic clusters or a cluster and leptons. In the case of $b \rightarrow c$ the procedure is repeated on the charmed cluster. This model is rather crude but gives roughly correct multiplicities.

Cluster model in Monte Carlo simulations

The procedure implementing the cluster model for Monte Carlo simulations works as follows. For clusters formed of a quark- antiquark pair with flavours $q_1\bar{q}_2$ a third flavour q_3 or d_3 is chosen randomly, where $q_3 = u, d,$ or s and d_3 is one of the six diquarks formed from these quarks. The decay products then are taken to be of flavour $q_1\bar{q}_3, q_3\bar{q}_2$ in the case of a two meson decay and $q_1d_3, \bar{q}_2\bar{d}_3$ in the case of baryon-antibaryon decay. For the chosen flavor combinations, the possible decay products are taken from a list of resonances, weighted with the spin degeneracy $(2S + 1)$ of these resonances. For the chosen decay product, the available phase space is tested against a random number, if the test fails, the procedure starts again by choosing a new flavour q_3 or d_3 and selecting the possible decay products. In this procedure, the produced quark-antiquark or diquark-antidiquark pair is only used as a flavour label and does not have any dynamical role.

2.5 Underlying event

For hadron colliders an event is determined by the hard process taking place at parton level. In order to do a full simulation the initial state partons for the hard process or initial state shower have to be extracted from the incoming hadrons. The remaining part of the incoming particle, the remnant, has to be taken into account. Since its behavior is dominated mostly by small momentum transfers, the relevant processes occur at non-perturbative scales, thus instead of analytical methods models have to be used.

A very simple approach to an underlying event model which is based on the **string hadronization** can be made if the color information of the remnant is kept and the remnant particles thus can be connected to the rest of the event via strings which then can hadronize as usual.

Analogous to this in the **cluster hadronization** the remnant can be treated as a cluster with a diquark flavor that then is treated like all the other clusters during hadronization.

A further possibility is a **Pomeron physics** based model. The basic unit of simulation here is a cut Pomeron giving rise to a chain of hadrons with a uniform rapidity and a Poissonian multiplicity distribution. The number and transverse momentum of

these Pomerons can be adjusted for hard and soft scatterings according to fits to experimental data. These Pomerons are then fragmented in their respective center of mass frame with an independent fragmentation function taking the necessary energy dependence into account.

Widely used is the approach based on the **parameterization of experimental data**. The UA5 collaboration developed a Monte Carlo program and fitted the treatment of the soft underlying event to the data taken in the experiment [24]. The remnants can be turned into clusters that decay into a varying number of subsequent clusters with an average of total sixteen clusters. These clusters are assumed to have a flat, central rapidity plateau with Gaussian tails and limited transverse momenta. These clusters decay isotropically, with different functions determining charged multiplicity, particle composition, quantum number and energy-momentum conservation. The parameters for these distributions are taken from fits to experimental data.

Another way of modeling the underlying event is via **multiparton interactions**, MPI, [25]. When the centre-of-mass energy \sqrt{s} is large enough, additional interactions in the same proton-proton collision are likely to happen. The mean number of interactions depending on the impact parameter b and s is taken to be

$$\begin{aligned}\langle n(b, s) \rangle &= \mathcal{L}_{\text{partons}} \otimes \hat{\sigma}_H \\ &= \frac{A(b)}{P_{\text{res}}} \sigma_H^{\text{inc}}(s)\end{aligned}$$

where $\mathcal{L}_{\text{partons}}$ is the parton luminosity, $\hat{\sigma}$ the cross section for a pair of partons to produce a pair of jets, $A(b)$ a function specifying the parton distributions depending on the impact parameter, $P_{\text{res}} \approx 1/300$ a parameter describing the ρ -meson dominance and σ_H^{inc} is the *inclusive* cross section.

From this the probability for m scatters in an event can be derived assuming that the separate scatters are uncorrelated

$$P_m = \frac{(\langle n(b, s) \rangle)^m}{m!} \exp(-\langle n(b, s) \rangle)$$

Based on this, the total cross section is

$$\begin{aligned}\sigma_H(s) &= \pi P_{\text{res}} \int db^2 \sum_{m=1}^{\infty} P_m \\ &= \pi P_{\text{res}} \int db^2 [1 - \exp(-\langle n(b, s) \rangle)]\end{aligned}$$

The implementation of this model in a Monte Carlo event generator has to deal with the assumption of uncorrelated scatters, especially taking care of energy-momentum conservation. Since the outgoing particles in the MPI model are partons, they as well as the hard process outgoing particles have to be showered by the parton shower. This means that the implementation has to be done in the same step as the shower in order to treat the outgoing partons properly.

2.6 Decays

Since most of the hadrons produced in hadronization have a short lifetime, the real final state of an experiment consists of the decay products of these particles. An event generator therefore must include a part in which these decays are taken into account. Since the decays as well can not be simulated analytically, measured values of decay widths, branching fractions and lifetimes are stored in tables and the event generator chooses a decay channel with the weight according to the corresponding branching fraction.

The phase space for these decay processes is rather complicated, especially if the final state consists of a very large number of particles with non-negligible masses. To deal with this special algorithms have been developed, like the MAMBO algorithm [26] [27], where the N body final state momentum configuration first is generated by using an exponentially damped phase space integration for the process and then boosting and rescaling it to reach overall energy-momentum conservation.

2.7 Jet definitions

Jet algorithms start from a list of "particles" that can be calorimeter cells, particle tracks, etc. in an experimental study, the output of hadrons in a full event simulation or the partons in a perturbative QCD calculation. One wants a jet definition which is able to specify a jet configuration unambiguously, be it in a theoretical calculation or in experimental data analysis.

The role of the algorithm is to associate clusters of these final state particles into jets such that the kinematic properties of the jets can be related to the corresponding properties of the partons in the hard process. Thus the jet algorithm gives a direct insight into the process at QCD level, so for a good jet algorithm there should not be large differences in the properties of jets at parton level to the jets at hadron or calorimeter level.

Many qualitative features of hadron production can be described by the intuitive jet definition of "a large amount of hadronic energy/activity in a small angular region" but for a detailed qualitative analysis a more precise definition is needed. A proper algorithm should satisfy the following set of conditions [28]:

fully specified All the steps should be fully specified so that the jet selection, the kinematic variables and all possible corrections are clearly and completely defined. Additional algorithms like preclustering, splitting or merging must as well be specified completely.

Theoretically well behaved In QCD calculations, all observables must be infrared and collinear safe in order to cancel infrared divergences. Also important for jet algorithms in hadron colliders is longitudinal boost invariance because in that case the center of mass of the individual parton-parton collision is normally boosted with respect to the hadron-hadron center of mass.

Detector independence The outcome of the algorithm should not depend strongly on the detector specifications like segmentation or resolution and not amplify the effects of resolution smearing and angle biases.

Straightforward implementation The algorithm should be easy to implement for perturbative calculations and experimental data, as well as provide an efficient use of computer resources.

In principle a jet algorithm works in two steps. In the first step, the actual **jet algorithm**, it selects a set of particles which are typically close to each other in the $y-\phi$ -plane and then in a second step combines their momenta to form the momentum of a jet according to a specific **recombination scheme**.

There are essentially two classes of jet algorithms, **cone-type** [29] and **clustering** [30] algorithms.

2.7.1 Cone-type algorithms

In cone-type algorithms a jet of radius R consists of all the particles whose trajectories lie in a given area $A = \pi R^2$ of the $y-\phi$ space and the axis of the cone coincides with the jet direction defined by the the E_T weighted centroid of the particles in the cone, E_T being the transverse energy, $E_T = E \sin \theta$. Jets are then defined by maximizing the amount of energy which can be covered by such cones.

In order to save computation time, the iteration used to search for the stable cones (i.e. the cones where the cone central axis coincides with the jet direction) in experimental data starts with the cones centered about the most energetic particles, the **seeds**. For all particles in a seed cone, the E_T weighted centroids are calculated and used as centers for new cones. This is iterated until the cone is stable.

A problem for the cone algorithm arises from the fact that nothing in the algorithm prevents the cones from overlapping, so that particles may belong to several jets at once. While it is no problem to have particles which do not belong to any jet, the case where particles are not assigned to jets unambiguously has to be avoided, else the particles' energy would be counted twice. This can be resolved by adding a procedure in the algorithm specifying how to split or merge overlapping cones.

Another problem encountered when using a cone algorithm with seeds arises when the algorithm is applied in theoretical calculations. The algorithm is neither stable under the emission of soft or collinear gluons [28]. The presence of soft gluons can change the outcome of the jet finding because it can change the cone axis and prevent merging when in a configuration without soft gluons merging would have taken place or vice versa.

SIS cone - a seedless infrared safe cone algorithm

For final states with a low number N of particles a soft and collinear safe algorithm can be constructed as follows.

- All subsets \mathcal{S} of the total set of particles are identified and for each subset the the rapidity $y_{\mathcal{S}}$ and azimuth $\phi_{\mathcal{S}}$ of the combined momentum of the subset $p_{\mathcal{S}} = \sum_{i \in \mathcal{S}} p_i$ are calculated.
- All subsets \mathcal{S} are checked whether a cone centered on $y_{\mathcal{S}}, \phi_{\mathcal{S}}$ contains all particles in \mathcal{S} and no more. If this is the case, \mathcal{S} corresponds to a stable cone.

This method finds all possible jets, but is very limited in its use, since there are about 2^N subsets and establishing whether a given subset corresponds to a stable cone takes $\mathcal{O}(N)$ calculations, leading to a total of $\mathcal{O}(N2^N)$ calculations per event, which is way too much for a hadron-hadron collision final state with hundreds of particles.

A useful seedless algorithm [34] can be constructed by identifying all distinct circular enclosures, also called distinct cones, where distinct means having a different particle content. This can be done as follows.

- All final state particles are put into the list of particles.
- For the current set of particles, all stable cones are found:
 - Merge any group of collinear particles into a single one.
 - For every particle i
 - * find all particles j within a distance $2R$ of i , if there are none, i gives a stable cone.
 - * For all j find the two circles of radius R for which i and j lie on the circumference, compute the angle of its centre C relative to i , $\theta = \arctan \frac{\Delta\phi_{iC}}{\Delta y_{iC}}$ and sort these circles by increasing angle θ .
 - * For the first circle in this order, calculate the total momentum and get the particle content⁸ for the cones that it defines. All four permutations of edge points being included or excluded are considered and called the current cones.
 - * For each of these four current cones the cones that have not been found yet are added to the list of distinct cones, the cones that have not yet been labeled as unstable are checked whether the in/out status of the edge particles with respect to the cone momentum axis is the same as when the cone was defined. If this is not the case, the cone is labeled as unstable⁹.
 - * Move to the next circle in order, differing from the previous by one particle entering or leaving the circle. Calculate the momentum and particle content of the new circle and current cones¹⁰.

⁸This is labeled efficiently by assigning a random integer tag to each particle from which a tag for combinations of particles can be defined by taking the logical exclusive-OR of all the tags of the individual particles. Thus the particle content of two cones can be compared by comparing their tags rather than their full content list.

⁹This is a test to reduce the number of cones to be tested for stability.

¹⁰This can be done by adding or removing the momentum and label of the particle entering or leaving.

- Check all cones not labeled as unstable for stability and add the stable to the list of protojets.
- All particles that are in stable cones are removed from the list of particles.
- The stable cones for the remaining set of particles are found.
- If there are no more stable cones or the algorithm has gone around the loop for a given number of times the finder stops and a split-merge procedure is applied on the full list of protojets.

A full estimate on the amount of calculations to be undertaken by this algorithm can be found in [34], the overall amount of calculations required is $\mathcal{O}(N^2 \ln N)$ which is much lower than the brute force implementation and in the range of the midpoint implementation of a cone algorithm with seeds.

2.7.2 Clustering algorithms

In clustering algorithms the particles are assigned to jets in an iterative way, starting on a list of "protojets" consisting of the final state particles of interest. The algorithm then merges protojets with nearly parallel momenta into a new protojet, removing the original two protojets from the list. A merging of two protojets to a new one takes place if given energy-angle resolution variable y_{ij} is smaller than a fixed resolution parameter y_{cut} . The algorithm then starts again with the updated list of protojets. It also determines when, for a particular protojet, the joining should cease, in which case that protojet is labeled a "jet" and is removed from the list of protojets and not manipulated further.

k_T algorithm

1. For every protojet h_k compute the resolution variable d_{kB} . It has the property that in the small angle limit it reduces to the squared relative transverse momentum of the protojet with respect to the beam direction. The actual definition for d_{kB} may differ depending on the process considered. For hadron-hadron collisions the most common choice is the $\Delta\mathcal{R}$ scheme, where $d_{kB} = p_{T,k}^2 \cdot R$, where R is a dimensionless parameter, playing a radius-like role defining the extent of the jets.
2. For every pair h_k and h_l of protojets calculate the resolution variable d_{kl} . It has the property that in the small angle limit it reduces to the squared relative transverse momentum of the two protojets. Again there are several possible choices for this variable, in the $\Delta\mathcal{R}$ scheme it would be $d_{kl} = \min(p_{T,k}^2, p_{T,l}^2) \cdot R_{kl}^2$, where $R_{kl}^2 = (y_k - y_l)^2 + (\phi_k - \phi_l)^2$. This definition corresponds to the one used in cone algorithms.
3. Find the smallest value d_{\min} among the d_{kB} and d_{kl} .

- if a d_{kl} is the smallest value, remove h_k and h_l from the list of protojets and combine them into a new protojet with momentum $p_{(kl)}$ according to a recombination scheme. Again there are several possible schemes, for example the E -scheme corresponding to vector addition of the fourmomenta, this is a good choice because it is an exact way of merging and also favorable with regard to simplicity.
 - If a d_{kB} is smallest, objekt h_k is defined to be a jet and removed from the list of protojets.
4. Repeat until all objects have been included in jets.

In step 2., it is possible to include an additional parameter d_{cut} [31] and check if $d_{\text{min}} > d_{\text{cut}}$. In that case, all remaining objects are classified as jets and the algorithm is complete. This so called "exclusive mode" separates the hard final state explicitly from the soft beam remnants, the d_{min} defines the hard scale of the process. This can be done in two ways, either a fixed d_{cut} (and therefore scale) is used to find the jets with $p_{T,i}^2 > d_{\text{cut}}$, or the d_{cut} is set in each event, defined by the d_{min} obtained after clustering the list of protojets into a given number of final state jets.

Without this stopping parameter, the resulting "inclusive mode" bears a similarity with the cone algorithm and is a good choice for hadron-hadron colliders.

FastJet Algorithm - a faster k_T algorithm

The above definition for the k_T jet finder for an event with N final state objects requires $\mathcal{O}(N^3)$ operations. Explicitly, the number of calculations needed is:

1. Constructing a table of all the d_{kl} , d_{kB} for the initial set of particles - $\mathcal{O}(N^2)$ operations, done once.
2. Find the minimal value d_{min} of the d_{kl} , d_{kB} in the table - $\mathcal{O}(N^2)$ operations, done N times.
3. Merge or remove the particles corresponding to the d_{min} as appropriate - $\mathcal{O}(1)$ operations, done N times.
4. Update the table of the d_{kl} , d_{kB} to take the merging or removal into account, proceed at step two if particles remain - $\mathcal{O}(N)$ operations, done N times.

The dominant step is the second one, requiring $\mathcal{O}(N^2 \times N = N^3)$ operations. This can be improved, if one considers the geometrical aspects of the problem. It can be proven [33] that if a pair h_k , h_l forms the smallest d_{kl} and $k_{T,k} < k_{T,l}$ then $R_{kl} < R_{kj}$ for all $j \neq l$, so l is the geometrical nearest neighbour of particle k .

This means if for each particle the geometrical nearest neighbor can be found, one does not have to calculate a $\mathcal{O}(N^2)$ table of d_{kl} but only an array of size N of the $d_{i\mathcal{G}_i}$ where \mathcal{G}_i is the nearest neighbor of particle i . The resulting algorithm then is:

1. Find the geometrical nearest neighbour \mathcal{G}_i for each particle i - scan through all $\mathcal{O}(N)$ other particles ($\mathcal{O}(N^2)$ operations).
2. Calculate the table of all the $d_{i\mathcal{G}_i}, d_{iB}$ - $\mathcal{O}(N)$ operations.
3. Find the minimal value d_{\min} of the $d_{i\mathcal{G}_i}, d_{iB}$ - $\mathcal{O}(N)$ operations, repeated N times.
4. Merge or remove the particles corresponding to the d_{\min} as appropriate - $\mathcal{O}(1)$ operations, done N times.
5. Find which particles' nearest neighbour has changed and update the table of the $d_{i\mathcal{G}_i}, d_{iB}$, proceed at step three if particles remain - $\mathcal{O}(N)$ operations, repeated N times.¹¹.

By this, the complexity of the problem is reduced to N^2 , but it can be done even faster by using solutions which have been devised in computational algebraic geometry for similar problems. The calculation of the geometrical problems – finding the initial state nearest neighbours, finding d_{\min} at each iteration and updating the nearest neighbour information - can be improved by using a Voronoi diagram. In a Voronoi diagram the plane with N vertices is divided up into N cells, one per vertex, in such way that every point in the cell surrounding vertex i has i as its nearest vertex. This structure makes the nearest neighbour location easy, because the nearest vertex \mathcal{G}_i for a given vertex i is one of the $\mathcal{O}(1)$ cells sharing an edge with the cell of vertex i . Such a Voronoi diagram for N vertices can be constructed with $\mathcal{O}(N \ln N)$ operations, finding the nearest neighbours can be done with $\mathcal{O}(N)$ operations.

Inserting or removing a point in the Voronoi diagram and updating all the nearest neighbour information can be done in $\mathcal{O}(\ln N)$ operations, which is done N times. Since the array of the $d_{i\mathcal{G}_i}$ changes only by entries per iteration, it can be represented with a binary tree structure, taking $\mathcal{O}(N \ln N)$ operations to create, in which finding the minimal value, insertion and removal of an element are guaranteed to take not more than $\mathcal{O}(\ln N)$ operations. The tree is constructed at the beginning, following $\mathcal{O}(N)$ updates and searches, resulting with a total of $\mathcal{O}(N \ln N)$ operations.

In conclusion both the geometrical and the minimum- finding aspects of the k_t jet-finder can be solved by relating them to known problems whose solutions require $\mathcal{O}(N \ln N)$ operations.

Based on these geometrical considerations a pure geometrical variation of the k_T algorithm has been proposed, where the recombination parameter is defined as $R_{ij}^2 = \Delta y_{ij}^2 + \Delta \varphi_{ij}^2$, so the jet definition is based on *angular ordering*. This Cambridge-Aachen algorithm is described in more detail in [32].

¹¹After removal of particle i the nearest neighbour information for particles which had i as nearest neighbour has to be updated, the number of particles that have i as nearest neighbour can be shown to be of $\mathcal{O}(1)$. Also one has to check if any particles acquire the newly created particle l as new nearest neighbour, $\mathcal{O}(N)$: comparing each particle's current nearest neighbour distance with its distance from l .

Chapter 3

Analysis

3.1 Programs used

3.1.1 The program vbfno¹

The program `vbfno` is a parton level Monte Carlo (MC) program for various **V**ector **B**oson **F**usion processes at **NLO** QCD. It calculates amplitudes by means of helicity amplitudes using `HELAS` [35] and `MADGRAPH` [36]. The MC integration and stratified sampling is done with a modified version of `VEGAS` [37] using a optimized phasespace for up to seven final state particles.

The output of the simulation can be either leading or next-to-leading order QCD cross sections and arbitrary distributions allowing to calculate differential K-factors or event files in the format specified by the Les Houches Accord (LHA files).

It is possible to implement various different cuts as well as other parameters such as the scales and PDF sets used. In addition anomalous couplings for Higgs and vector bosons can be used and set.

The PDF set to be used in the calculation can be specified, either the built-in CTEQ6 tables [38] can be used or the LHAPDF library [39] can be linked into the program.

Additionally it is possible to make parallelized runs using Condor clusters.

A wide variety of processes is implemented, on the one hand Higgs production and various decays such as

$$\begin{aligned} pp &\rightarrow Hjj \\ &\rightarrow Hjj \rightarrow \gamma\gamma jj \\ &\rightarrow Hjj \rightarrow b\bar{b}jj \\ &\rightarrow Hjj \rightarrow ll + jj \\ &\rightarrow Hjj \rightarrow W^+W^-jj \rightarrow l^+l^-\nu_l\bar{\nu}_l + jj \\ &\rightarrow Hjj \rightarrow ZZjj \rightarrow 4l + jj , \end{aligned}$$

¹<http://www-itp.physik.uni-karlsruhe.de/~vbfno/loweb>

at next-to-leading order, the same processes are implemented with an additional jet in the final state at leading order.

Additionally, background processes to Higgs production in vector boson fusion are implemented

$$\begin{aligned}
 pp &\rightarrow Zjj \rightarrow ll + jj \\
 &\rightarrow W^\pm jj \rightarrow l\nu_l + jj \\
 &\rightarrow W^+W^-jj \rightarrow l^+l^-\nu_l\bar{\nu}_l \\
 &\rightarrow ZZjj \rightarrow 4l + jj ,
 \end{aligned}$$

at NLO and with an additional final state jet at leading order.

3.1.2 The Les Houches Accord

The events created in `vbfnlo` do not have physically measurable final states, since the output contains quarks and gluons which are not color singlet states. So in order to get a full event simulation, a second program has to be used, a showering and hadronization generator (SHG). To transfer the results from `vbfnlo`, a specific file format is used that is defined in the Les Houches Accord [40], [41].

The communication between a matrix element generator (MEG) like `vbfnlo` and a SHG like `Herwig++` occurs at several stages of a full event simulation.

First, at initialization basic parameters about the experiment and the calculation have to be transferred as well as information about the considered processes. Additionally, "meta information" needs to be conveyed, how to deal with the weights and information about the used MC and so on.

Then the data for every created event has to be transferred as well in order to provide the SHG with the information needed to perform the simulation of the shower and hadronization. This includes the parameters for the event like scales, the weight and the couplings as well as information about the particles in the event.

Normally, different parts of the SHG are used to read these informations, so there are two `common` blocks used for this. Their form is similar to the `HEPEVT common` blocks used to transfer information about SHG final states to detector simulation or analysis programs.

User process run information: HEPRUP

The general information about the generated events is stored in this `common` block.

```

integer          MAXPUP
parameter ( MAXPUP=100 )
integer          IDBMUP, PDFGUP, PDFSUP, IDWTUP, NPRUP, LPRUP
double precision EBMUP, XSECUP, XERRUP, XMAXUP
common /HEPRUP/  IDBMUP(2), EBMUP(2), PDFGUP(2), PDFSUP(2),
+               IDWTUP, NPRUP, XSECUP(MAXPUP), XERRUP(MAXPUP),
+               XMAXUP(MAXPUP), LPRUP(MAXPUP)

```

First, there are informations about the beam particles, what particles there are (IDBMUP), their energy (EBMUP) and in the case of hadrons the PDF sets used in the calculation (PDFGUP, PDFSUP).

Then there is a switch (IDWTUP) telling the SHG if the events are weighted and if the SHG is supposed to produce unweighted events or not.

Finally, there is a list of the NPRUP different processes stored in the file and for each process the cross section (XSECUP), its error (XERRUP), the maximal weight (XMAXUP) and a number LPRUP identifying the process is stored.

Only the most general informations are included here in order to obtain a general structure.

User process event information: HEPEUP

For each event in the final state of the calculation there is the HEPEUP common block in which the necessary information concerning this event is stored.

```
integer      MAXNUP
parameter (  MAXNUP=500  )
integer      NUP, IDPRUP, IDUP, ISTUP, MOTHUP, ICOLUP
double precision  XWGTUP, SCALUP, AQEDUP, AQCDUP,
                  PUP, VTIMUP, SPINUP

common /HEPEUP/  NUP, IDPRUP, XWGTUP, SCALUP, AQEDUP, AQCDUP,
+                IDUP(MAXNUP), ISTUP(MAXNUP), MOTHUP(2,MAXNUP),
+                ICOLUP(2,MAXNUP), PUP(5,MAXNUP), VTIMUP(MAXNUP),
+                SPINUP(MAXNUP)
```

First, there is the size of the common block given by the number of particles NUP. This is followed by information about the event, the process IDPRUP, the weight XWGTUP, the scale SCALUP at which the running couplings and PDFs have been evaluated and the values of the coupling constants used (AQEDUP, AQCDUP).

Then for each of the NUP particles the relevant informations are stored. These consist of the information what kind of particle there is (IDUP²), the status of the particle (ISTUP), specifying whether the particle is in- or outgoing or an internal resonance, the mother particles (MOTHUP) and the colorflow (ICOLUP). Then there is the five-momentum³ of the particle (PUP), the invariant lifetime VTIMUP and the spin information SPINUP.

XML⁴ structures in Les Houches event files

The data in the HEPRUP and HEPEUP common blocks has to be stored in one single file. In addition to the data stored in these blocks a user might want to add additional program- specific information on how the events were generated or other totally different data. For this an XML-like framework is used to store this additional data and

²Numbered by the convention of the particle data group

³The first four components are the usual components of a four-momentum vector, p_x, p_y, p_z, E , the fifth is the invariant mass squared

⁴XML = extensible markup language, for an introduction see for example [42]

to distinguish it from the compulsory one in the `common` blocks. The Les Houches File created that way is no "real" XML document by the standard of this file format, since the `common` blocks contain data which is not structured by XML standard, so no validating is possible.

The resulting structure for the files is:

```
<LesHouchesEvents version="1.0">
<!--
  # optional information in completely free format,
  # except for the reserved endtag (see next line)
-->
<header>
  <!-- individually designed XML tags, in fancy XML style -->
</header>
<init>
  compulsory initialization information - HEPRUP data
  # optional initialization information
</init>
<event>
  compulsory event information - HEPEUP data
  # optional event information
</event>
(further <event> blocks, one for each event)
</LesHouchesEvents>
```

The tags must all be alone in their respective line, especially the `<init>`, `<event>` tags, since their position defines the points after which the compulsory data in the `common` blocks is stored. The `LesHouchesEvents` block defines the root element of the XML document, it shows the standard and version of the file. Additional information can be written in the `<!-- ... -->` and `<header> ... </header>` blocks. The data in the `<!-- ... -->` block is XML "comment" style, so it can be arbitrarily formatted, since it is ignored by XML parsers, whereas data in the `<header> ... </header>` block has to be based on XML syntax.

This structure is open to future as well as individual extensions, it is for example possible to add additional information in the attribute part of the elements containing the `common` blocks, since `<init` , `<event` followed by a blank and further text on the same line should be treated equivalent to `<init>`, `<event>` tags, any comment or information can follow after the compulsory initialization or event information before the respective end tag.

3.1.3 The program Herwig++

Herwig++ [43] is a Monte Carlo event generator based on the highly successful HERWIG event generator used during LEP/HERA and Tevatron [44]. HERWIG, which was written in FORTRAN, has been improved with the progress of understanding the physics involved in event simulation, but it has reached the limit of being manageable.

New theoretical ideas for the LHC make major changes necessary but since the original event generator evolved into a large-scale programming process, this is rather difficult to accomplish. Therefore a new approach was made to redevelop the original program in the object-orientated programming language C++, including new developments wherever possible.

The outline was to recode the cluster hadronization model making minor improvements to fix known problems and to write a new angular-ordered parton shower with improved treatment of mass effects and Lorentz invariance properties. To accomplish this the angular ordered shower based on the DGLAP equations and the evolution variable

$$\tilde{q}^2 = \frac{p_T^2}{z^2(1-z)^2} = \frac{q^2}{z(1-z)}$$

was changed to a shower algorithm based on quasi-collinear splitting introduced in [15]. To accomplish this, a different definition of z and p_T was used [45] to obtain invariance under boosts along the jet direction and the evolution variable was changed to

$$\tilde{q}^2 = \frac{q^2 - m^2}{z(1-z)} = \frac{p_T^2}{z^2(1-z)^2} + \frac{m_i^2}{z^2(1-z)} + \frac{m_j^2}{z(1-z)^2} + \frac{m^2}{z(1-z)} .$$

The definitions for the masses can be seen from figure 3.1.

The main improvements aimed for are evolution down to zero p_T radiation from massive particles, thus avoiding the 'dead-cone' and the overlap in the soft region of the 'final state + gluon' phase space that appeared in the shower used in the FORTRAN program. To avoid a 'dead' region of phase space not covered by the parton shower due to the choice of evolution variables, hard matrix element corrections are applied, thus taking the most important subleading higher-order corrections into account. Additionally the azimuthal angle φ is not chosen randomly but as a result of planned azimuthal spin correlations [46].

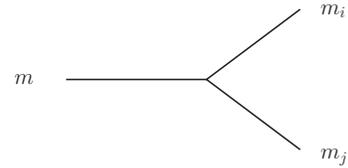


Figure 3.1: Kinematics of the shower splitting.

The hadronization in **Herwig++** is modeled by the cluster model that was described in section 2.4.3, as it was done in the FORTRAN version, but with an improved algorithm for the cluster decay.

There are several possible choices for the underlying event, see 2.5, in **Herwig++** it is possible to choose between an implementation of the UA5 parametrization and a multiparton interaction (MPI) model as well as no model at all, where the beam remnants are treated as normal clusters.

Herwig++ is based on **ThePEG** [47], a toolkit for implementing physics models of event generation. **ThePEG** provides a general structure of event generation in terms of abstract base classes for the different tasks appearing in the simulation of an event. The implementation of different models is then done by creating classes that inherit from the respective base classes and implementing sets of pre-defined virtual functions.

3.2 Analysis

The program `vbfnlo` was used to generate Les Houches event files for Higgs production and decay in vector boson fusion. The Higgs mass was set to be $m_{\text{Higgs}} = 120 \text{ GeV}$ and some VBF cuts were applied to the events, especially on the rapidity gap between the hardest two partons. The factorization scale of the process was chosen to be set dynamically by the momentum transfer of the exchanged W^\pm and Z bosons⁵. The calculation depends on several electroweak observables which are related to each other. These are the masses of the W^\pm and Z bosons M_W, M_Z , the Weinberg angle θ_W , the electromagnetic coupling constant g , the fine structure constant α and the Fermi constant, G_F . In order to gain numerical accuracy, instead of assigning values to all of them, some were calculated from others. The used relations are:

$$\begin{aligned}\cos^2 \theta_W &= \frac{M_W^2}{M_Z^2} \\ e &= g \sin \theta_W \quad \text{where } e \text{ is the charge of the electron} \\ \alpha &= \frac{e^2}{4\pi} \\ G_F &= \frac{g^2}{4\sqrt{2}M_W^2}\end{aligned}$$

The scheme chosen was the G_μ scheme⁶, so the values for the boson masses and G_F were set to be the actual experimentally found values and θ_W, α were calculated:

$$\begin{aligned}\cos^2 \theta_W &= \frac{M_W^2}{M_Z^2} \\ \alpha &= \sqrt{G_F 4\sqrt{2}M_W}\end{aligned}$$

To test the influence of a full event simulation on the VBF observables, the different steps performed by an event generator have to be examined. For `Herwig++` this is rather straightforward, since the different parts of the simulation are done by special objects, the handlers, which are set in the repository. This is an object database where all the settings for the event generator are stored at initialization to be called during the run.

An event simulation then goes step by step where the different steps are done by the respective handler, using the output particles of the previous step as input. All these internal steps can be analyzed independently by extracting their respective final state particles.

The analysis was done with a comparison of several different underlying event models, and since the handlers for these models work in different ways, the steps undertaken by `Herwig++` were dependent on the underlying event model used.

⁵by setting `id_muf=12` in `vbfnlo.dat`.

⁶`EWScheme=3` in `vbfnlo.dat`.

Three different possibilities for the underlying event were considered, the UA5 model, the MPI underlying event model and no extra underlying event model, here the proton remnants are just considered as clusters and enter the usual hadronization.

The different steps undertaken by `Herwig++` are

- **The hard process.** This was done by using Les Houches files created with `vbfnlo`. So in this simulation this part solely consisted of combining two incoming protons with the incoming partons specified in the Les Houches file, thus creating a "remnant" for every proton to be treated by an appropriate handler.
- **The shower.** The initial and final state partons in the hard process undergo a forward or respectively backward evolution resulting in a cascade of partons. In case of the Jimmy underlying event model, the underlying event was created here, since the additional interactions are processes involving partons, so their final states have to enter the shower as well.
- **Hadronization** To prepare the showered particles for hadronization, the gluons are split into quark-antiquark pairs, as described in section 2.4.3, these quarks are combined into colorless clusters, heavy clusters undergo cluster fission and quark-antiquark pairs are created to assign flavor to the clusters. The clusters then decay into hadrons, according to their quantum numbers and their flavour.
- **Soft underlying event** In case the UA5 underlying event model is used, the beam remnants decay according to the parametrization of this model in an extra step.
- **Decay of unstable particles** The unstable hadrons created in the cluster decay and the underlying event model decay into stable particles.

The final state particles of each step were extracted and examined individually. First, the detectable leptons were searched for, that is neutrinos were discarded. The remaining charged leptons had to pass several weak cuts taking into account detector effects: since very forward and very backward particles cannot be resolved, a pseudo-rapidity of absolute value of five or less was required $|\eta| \leq 5$, as well as a minimum energy of $E_{l,min} = 1$ GeV, to reject particles which would not reach a calorimeter in a detector due to magnetic deflection in the magnetic field present in a detector.

Then, the hadronic particles were collected. These particles of course depend on the simulation done in the present step, for the first two steps, the hard process and the shower, the final state consists of quarks and gluons, for the other steps the final state is made up of hadrons and, in the last step, when the unstable hadrons have been decayed, also photons. These particles are also checked for detectability in the same way as the leptons ($|\eta| \leq 5$, $E_{charged, min} = 1$ GeV) and then turned into jets by a jetfinder. In this analysis `fastjet 2.1.0` [33] was used, which allows to use different recombination methods. The jet finding algorithm used was the k_T algorithm, in the fast implementation mentioned in section 2.7.2 and `SISCone`, a fast infrared-safe cone algorithm, which is available as a plug-in for `fastjet`, described in section 2.7.1. The

R parameter for `fastjet` was set to be $R = 0.7$, a custom value in experimental analysis, in the case of `SISCone` [34], a cone radius of $R = 0.5$ was chosen. The resulting jets were then sorted in p_T or E_T and stored in `ROOT` files.

For the analysis, cuts were imposed on the generated events. The first requirement was that the two hardest jets are of

$$p_T \geq 20 \text{ GeV} \quad \text{respectively} \quad E_T \geq 20 \text{ GeV} ,$$

depending on the sorting of the jets and have a rapidity

$$|y| \leq 4.5 .$$

This is chosen in order to avoid distortions at the threshold of the detector's rapidity sensitivity.

To get a definite vector boson fusion signal, the two hardest jets were required to exhibit a rapidity gap and lie in opposite detector hemispheres

$$|y_1 - y_2| \geq 4 \quad \text{and} \quad y_1 \cdot y_2 < 0 .$$

In order to suppress background from QCD processes, a cut on the invariant mass of the two tagging jets was imposed

$$M_{j_1 j_2} \geq 600 \text{ GeV} .$$

In addition cuts were imposed on the leptonic decay products of the Higgs boson in case the two hardest jets passed the above cuts.

The leptons were required to be in the central region of the detector

$$|\eta_l| \leq 2.5$$

and hard enough to be measured

$$\begin{aligned} p_{T,l} &\geq 10 \text{ GeV} && \text{in case of two leptons,} \\ p_{T,l} &\geq 20 \text{ GeV} && \text{in case of one lepton.} \end{aligned}$$

In addition leptons and jets should be well resolvable, so the leptons were required to fall in the rapidity gap of the two tagging jets

$$y_1 \geq \eta_l \geq y_2^7$$

and a minimum separation of the leptons and jets in the rapidity- Φ (lego plot) plane

$$\Delta R_{jl} \geq 0.4 \quad \Delta R_{ll} \geq 0.1$$

The events passing these cuts were then accepted as measurable vector boson fusion events.

⁷Here y_1 denotes just the rapidity of the jet with $y_1 > y_2$ without implying that this jet is the hardest jet in the event.

Chapter 4

Vector boson fusion observables

The vector boson fusion observables did not turn out to be modified strongly by the full event simulation. A comparison of these shapes at matrix element level and after the event simulation does not show a large deviation. The final states of all the different intermediate steps were compared but in most cases the VBF observables were very similar. In these cases, only the matrix element level distribution and the final state distribution are shown.

To see the influence of the underlying event model on the VBF observables a comparison between the final state of the samples created with the UA5 parametrization and MPI model was made. Here again the distributions did not show a large deviation.

4.1 Jet observables

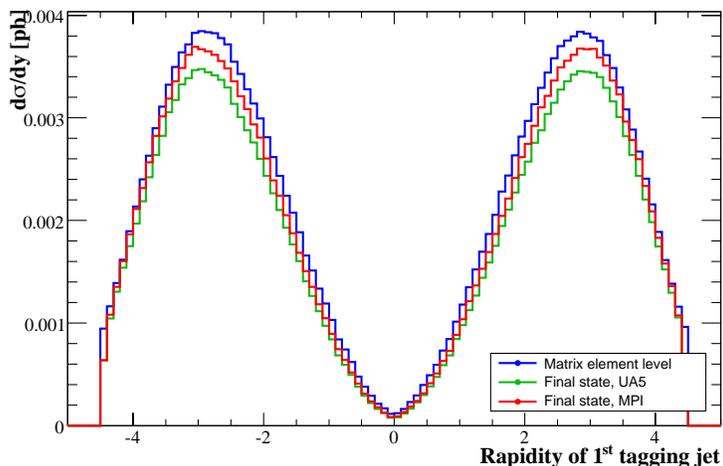


Figure 4.1: Rapidity distribution of the hardest tagging jet, Matrix element level and hadronic final state, UA5 parametrization and MPI model.

The rapidity distributions, Fig. 4.1 and 4.2 of the tagging jets turned out to be very stable, the distributions for the shower and final state are just a fraction lower

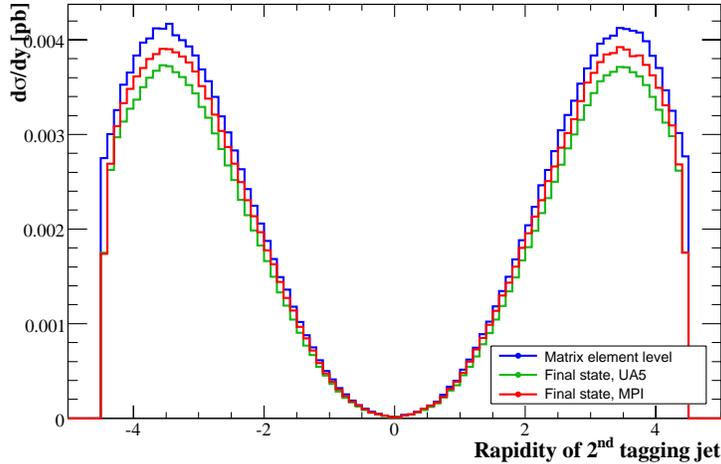


Figure 4.2: Rapidity distribution of the 2nd hardest tagging jet, Matrix element level and hadronic final state, UA5 parametrization and MPI model.

than that at matrix element level, since after the full simulation less events pass the VBF cuts and are rejected.

The characteristic features of the distributions remain even after the full simulation, the tagging jets exhibit a rapidity gap and the peaks of the distribution for the second tagging jet, Fig. 4.2, are a little further away from the origin than that of the first tagging jet, Fig. 4.1.

For the sample created with the UA5 parametrization the rapidity distributions peak at the same positions as the distribution at matrix element level, whereas the distributions from the MPI model have a slight shift in the peak position, the hardest tagging jet tends to lie more in the outbound region of the detector, the second hardest tagging jet on the other hand is a little bit more central.

4.1.1 p_T and E_T distributions of the tagging jets

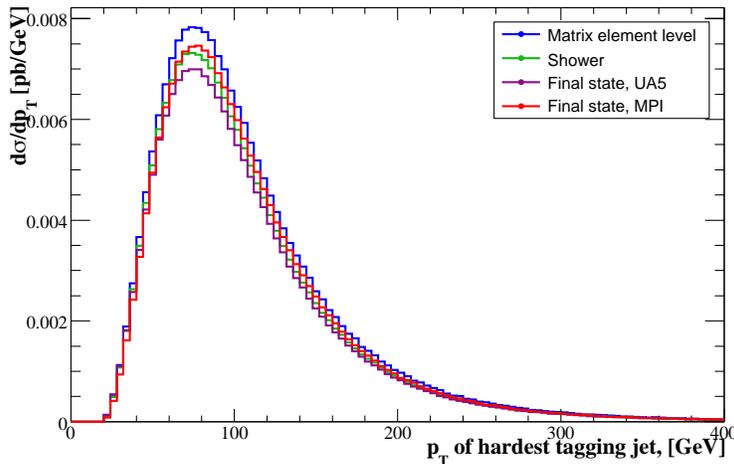


Figure 4.3: p_T distribution of hardest tagging jet, Matrix element level, shower and hadronic final state, UA5 parametrization and MPI model.

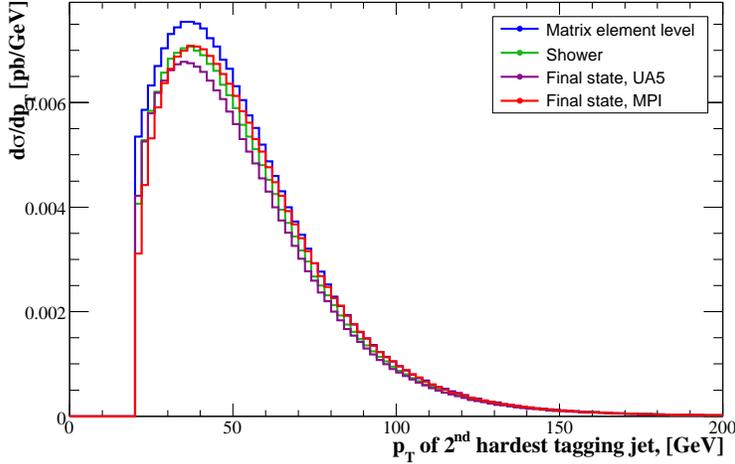


Figure 4.4: p_T distribution of 2nd hardest tagging jet, Matrix element level, shower and hadronic final state, UA5 parametrization and MPI model.

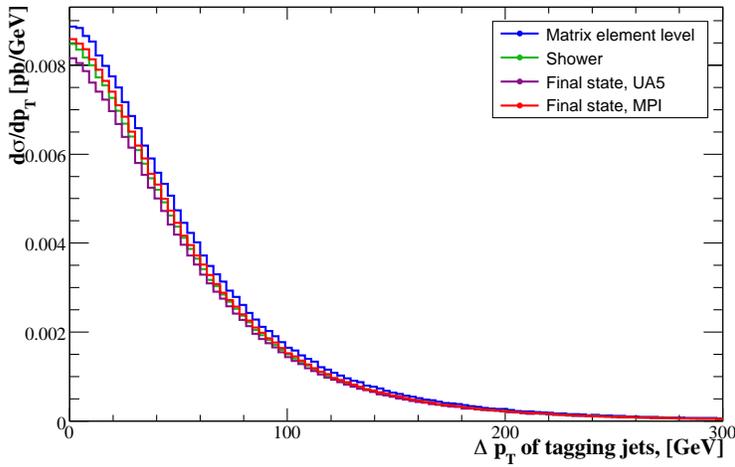


Figure 4.5: Δp_T distribution of tagging jets. Matrix element level, shower and hadronic final state, UA5 parametrization and MPI model. It can be seen that the tagging jets tend to have the same p_T .

The overall picture of the p_T distributions of the tagging jets is similar to the situation for the rapidity distributions. The distributions are very similar for all the steps, the only difference is that at matrix element level some more events pass the VBF cuts than at later steps. Different to the rapidity distributions a slight difference between the shower level and the final state is recognizable, the peak for the p_T distributions, Fig. 4.3, 4.4 of the two tagging jets are shifted by the underlying event, very little by the UA5 parametrization but quite a lot by the MPI model, where the peak position tends to be around 10 GeV higher in the hadronic final state than for the shower. This is very likely due to the fact that the MPI model produces a lot of additional low- p_T activity and some of the particles stemming from the additional scatterers are likely to be recombined into the original jet. Thus the jets gain energy.

The p_T difference between the two tagging jets, Fig. 4.5, however has the same shape for all the considered samples, so all effects seem to influence both tagging jets together.

Comparison of p_T and E_T ordering

Instead of defining the tagging jets as the two jets with the highest p_T value, it is possible to define them as the two jets with highest E_T . For the transverse momentum and energy the following definitions were used:

$$p_T^2 = p_x^2 + p_y^2$$

$$E_T^2 = E^2 \cdot \frac{p_T^2}{p_T^2 + p_z^2}$$

For parton level configurations, where the jets are in principle massless, this yields the same results as the previous definition, since for $m \rightarrow 0$ also $E_T \rightarrow p_T$.

The E_T distributions for the tagging jets defined this way, Fig. 4.6, 4.7, 4.8, show no perceivable deviation to the p_T distributions of the p_T ordered tagging jets, for both underlying event models used.

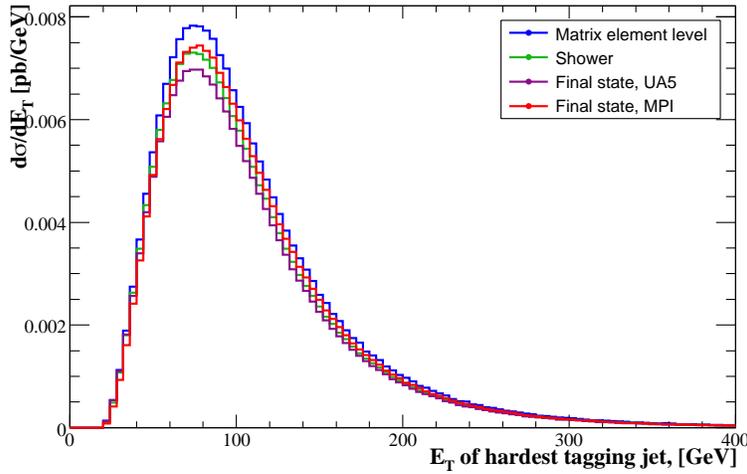


Figure 4.6: E_T distribution of hardest tagging jet, Matrix element level, shower and hadronic final state, UA5 parametrization and MPI model.

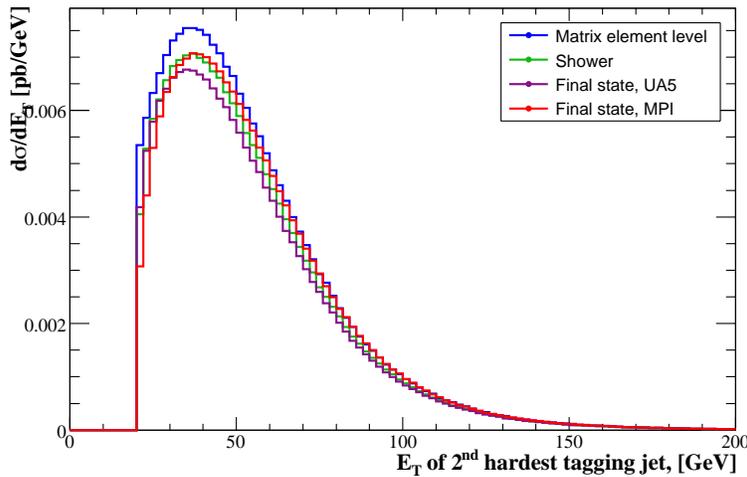


Figure 4.7: E_T distribution of 2nd hardest tagging jet, Matrix element level, shower and hadronic final state, UA5 parametrization and MPI model.

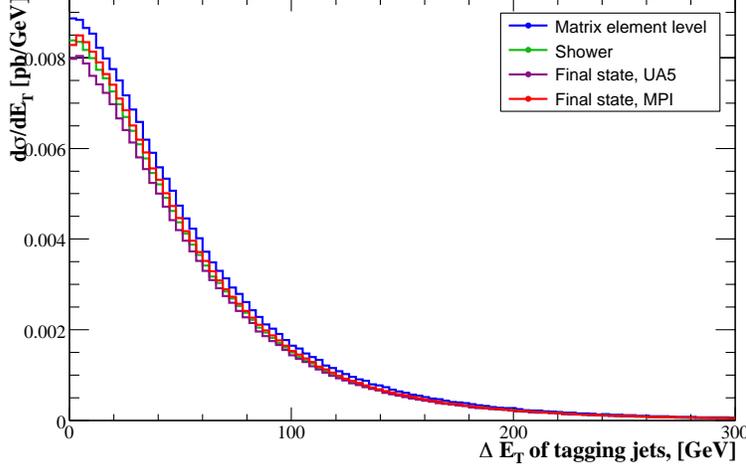


Figure 4.8: ΔE_T distribution of tagging jets. Matrix element level, shower and hadronic final state, UA5 parametrization and MPI model.

Since for the massive hadronic final state p_T and E_T are not exactly the same any more, the ratios of the p_T and E_T distributions are considered to see how big this difference is, Fig. 4.9. The mass effects turned out to be very small, in the low p_T range between 20 and 50 GeV the p_T distribution is about 5% above the E_T distribution. So sorting the jets by transverse momentum or transverse energy does not have a huge impact on the observables.

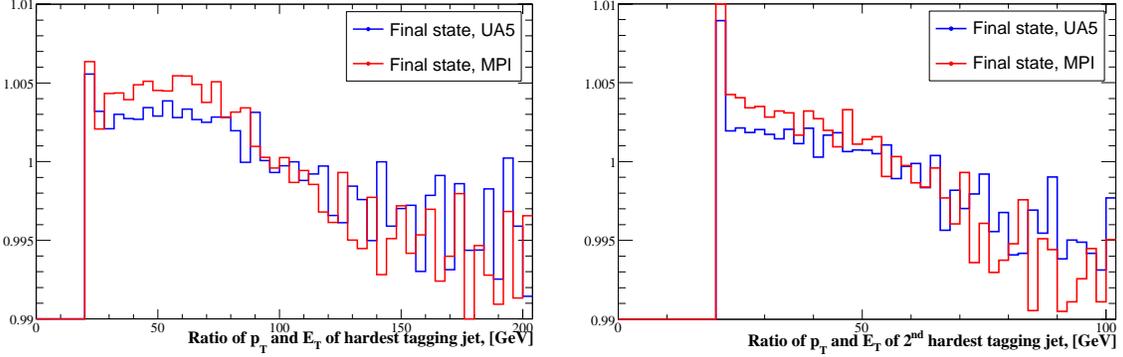


Figure 4.9: Ratio of p_T and E_T distributions of the two tagging jets, UA5 parametrization and MPI model.

This can be understood by regarding the mass distributions of the final state jets. The tagging jets tend to have a very low invariant mass, see Fig. 4.10, 4.11, for the hardest tagging jet the mass distribution for the UA5 parametrization peaks at 13 GeV, for the MPI UE model at 18 GeV and second hardest jet this distribution peaks at 9 GeV and 11 GeV respectively. With the above mentioned definition for E_T^2 ,

$$E_T^2 = E^2 \cdot \frac{p_T^2}{p_T^2 + p_z^2} = \frac{(m^2 + \vec{p}^2)p_T^2}{\vec{p}^2}$$

it is clear that these values for m_{jet} can be neglected, since the masses of the tagging jets are rather small compared to their transverse momentum.

The ratios in Fig. 4.9 show the mass difference between the UA5 and the MPI sample, the effect of the mass is slightly stronger for the MPI sample, taking the slightly higher mass of the tagging jets into account.

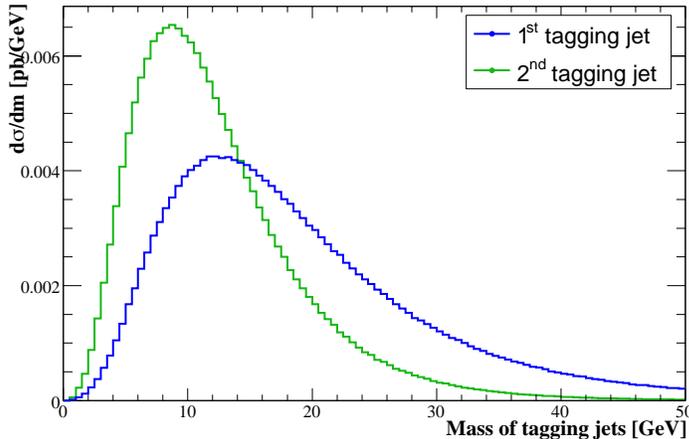


Figure 4.10: Invariant mass distribution of the two tagging jets, hadronic final state, UA5 parametrization. For the matrix element level, the jets are massless, since the quarks were taken to be massless in the calculation.

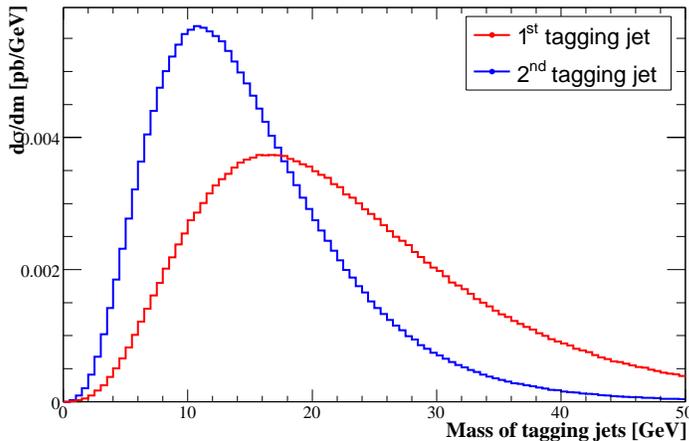


Figure 4.11: Invariant mass distribution of the two tagging jets, hadronic final state, MPI UE model. For the matrix element level, the jets are massless, since the quarks were taken to be massless in the calculation.

Further jet observables

Since the two tagging jets are very forward and very backward in rapidity, the invariant mass of the pair of tagging jets gets much larger than in typical QCD processes. Therefore it is possible to reduce that background by applying the above mentioned m_{jj} cut.

Here the simulation has a bigger impact on the distribution, Fig. 4.12. At matrix element level the cut on the invariant mass of the jet pair appeared to be on the peak of the distribution but the position of the peak moves to higher masses, from 600 GeV to 700 GeV for the UA5 sample and to 800 GeV for the MPI sample. So

this observable is influenced by the hadronization and to a large extent as well by the underlying event model used.

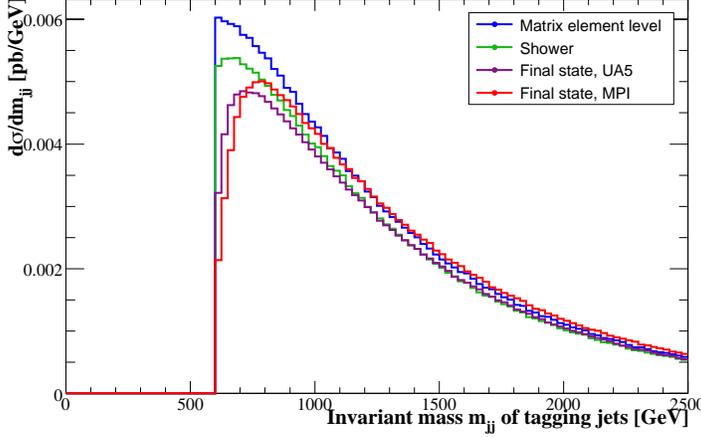


Figure 4.12: Invariant mass m_{jj} of the tagging jet pair. Matrix element, shower and final state, UA5 parametrization and MPI model.

A further observable of interest is the rapidity ordered azimuthal angle between the tagging jets. It is defined as

$$\Phi_{jj} = \begin{cases} \Phi_{j_1} - \Phi_{j_2} & \text{if } y_{j_1} \geq y_{j_2} \\ \Phi_{j_2} - \Phi_{j_1} & \text{if } y_{j_1} < y_{j_2} \end{cases}$$

The resulting angle is mapped to the interval $(-\pi, \pi]$. This distribution is sensitive to the nature of the HVV vertex. The most general tensor structure of this vertex which can contribute to VBF in the massless quark limit can be written as

$$T^{\mu\nu}(q_1, q_2) = a_1(q_1, q_2)g^{\mu\nu} + a_2(q_1, q_2)[q_1 \cdot q_2 g^{\mu\nu} - q_2^\mu q_1^\nu] + a_3(q_1, q_2)\epsilon^{\mu\nu\rho\sigma} q_{1,\rho} q_{2,\sigma}$$

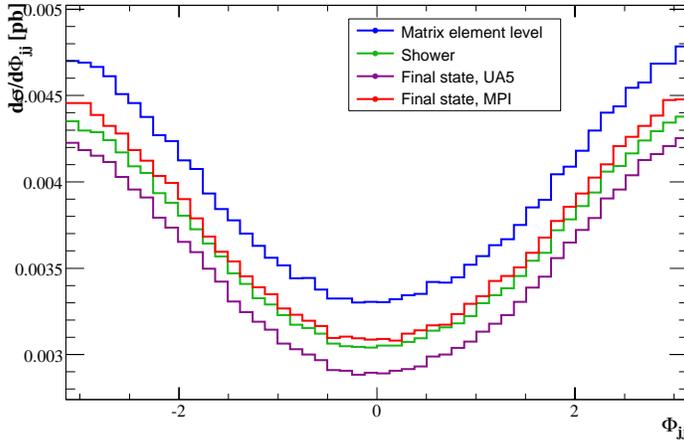


Figure 4.13: Φ_{jj} distribution of the two tagging jets. The observable is defined in the text. Matrix element and shower level and final state, UA5 parametrization and MPI model

For the standard model, $a_1 = 1$ and $a_2 = a_3 = 0$. The Φ_{jj} variable now is sensitive to the couplings, a CP-even and a CP-odd coupling yield a different shape. This has been studied in detail in [48]. Here it can be seen that the shape of the distribution, in this simulation a standard model coupling, is not affected by the simulation, Fig.4.13.

4.2 Tagging lepton distributions

The tagging leptons in this analysis were muons. Due to their comparatively long lifetime, muons normally do not decay while being in the range of a calorimeter. So they were set to be stable in the simulation, thus passing all the steps undergone by the event unchanged except for Bremsstrahlung effects.

Therefore all the distributions conceivable with tagging leptons are expected to be very stable. This turned out to be the case, of course the total amount of events was scaled down, since fewer events passed the VBF cuts at the final state of the simulation than at matrix element level.

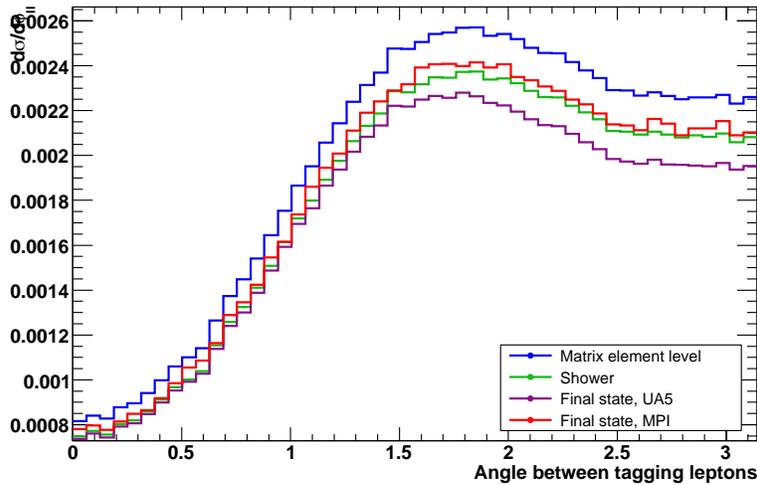


Figure 4.14: ϕ_U distribution of the two tagging leptons. Here ϕ denotes just the azimuthal angle between the leptons, without rapidity ordering as in the case of the jets.

Exemplary two distributions are shown, the azimuthal angle between the two tagging leptons, Fig.4.14, and the p_T distribution of the hardest tagging lepton, Fig.4.15.

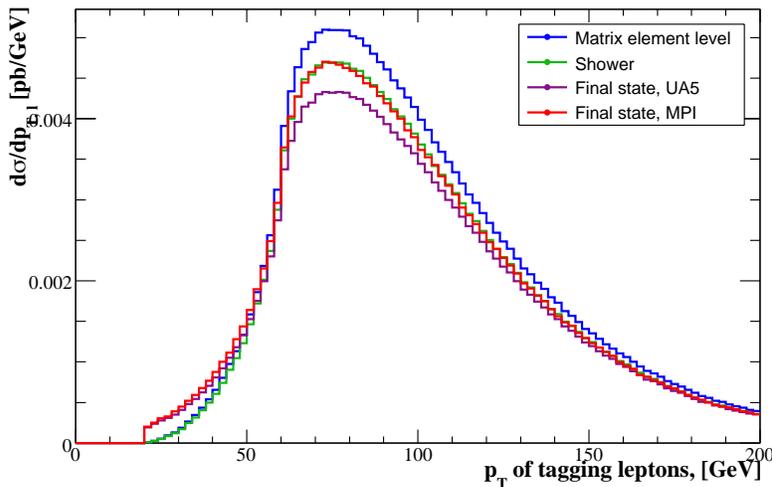


Figure 4.15: p_T distribution of the hardest tagging lepton.

4.3 Cut efficiency and additional hadronic activity

The cut efficiency of the used sample was rather high, which stems from the fact that the sample was created with almost all the VBF cuts applied as well, just a lower $p_{T,\min}$ cut was set in `vbfnlo`.

To test the influence of the used underlying event model, a comparison between the cut efficiency for these two models was made. A problem with a direct comparison emerges from the different ways the underlying event is simulated in the event, since for the multi parton interaction the underlying event produces additional outgoing partons, this step is included in the shower handler, whereas the underlying event given by the UA5 parametrization is almost independent of the rest of the simulation, its outgoing particles already being hadrons and therefore is done in an extra step after hadronization, just before the decay step.

So a direct comparison for the cut efficiencies can only be done at matrix element level (where of course the two samples are identical, using the same event files) and at the final state, Tab. 4.1.

Events in run:	UA5 parametrization		MPI model	
	$4 \cdot 10^6$		$4 \cdot 10^6$	
p_T ordered, hard process	$3.864 \cdot 10^6$	96.61 %	$3.864 \cdot 10^6$	96.59 %
p_T ordered, after shower	$3.559 \cdot 10^6$	88.97 %	$3.576 \cdot 10^6$	89.39 %
p_T ordered, after hadronization	$3.416 \cdot 10^6$	85.41 %	$3.649 \cdot 10^6$	91.24 %
p_T ordered, after ue and decays	$3.416 \cdot 10^6$	85.40 %	$3.625 \cdot 10^6$	90.63 %
E_T ordered, hard process	$3.864 \cdot 10^6$	96.61 %	$3.864 \cdot 10^6$	96.59 %
E_T ordered, after shower	$3.559 \cdot 10^6$	88.97 %	$3.574 \cdot 10^6$	89.36 %
E_T ordered, after hadronization	$3.416 \cdot 10^6$	85.40 %	$3.643 \cdot 10^6$	91.07 %
E_T ordered, after ue and decays	$3.416 \cdot 10^6$	87.39 %	$3.624 \cdot 10^6$	90.59 %

Table 4.1: Cut efficiency for the different simulation steps, UA5 and MPI UE.

The efficiency after the full event simulation in the UA5 parametrization is reduced by about eleven percentage points, where the biggest loss of about eight percentage points occurs during the shower step, followed by a loss of three percentage points during hadronization. The underlying event and the hadronic decays do not have a big impact on the acceptance of events.

In case of the multi parton interaction model, the cut efficiency is reduced by six percentage points from matrix element level to final state, the major loss of events occurring during hadronization.

The fact that the UA5 model produces more events that do not pass the VBF cuts can be explained by the different ways these models work. For the UA5 parametrization the events get smeared out by the different steps, the shower, the hadronization and especially the underlying event tend to distribute their outgoing particles around the positions of their incoming particles. That way, the event signature gets smeared out and the event does not pass the cuts any more.

Here again it can be seen that the way of defining the jets either via p_T or E_T does not have a huge impact on the structure of the events, the events pass the cuts regardless of the way the hardest jets are defined. This holds for both models that were examined.

Hadronic background

At matrix element level, the two tagging jets are the only outgoing hadronic particles. After the full event simulation, the number of hadronic particles has increased a lot, as a result of the shower and particularly the underlying event. Most of these particles get recombined into a relatively small number of jets by the jet finder. In addition, in the central region a lot of single, low energetic particles are not recombined with other particles and form a set of very low energetic jets.

To gain an insight into the hadronic activity beyond the two tagging jets, the transverse energy distribution of the background was examined. This background is defined here as all jets except the two tagging jets. In order to analyse this, all jets except the tagging jets are combined by fourvector addition to get a 'background fourvector'. For this no minimum p_T cut was set on the jets, so that no background information was lost. Of this background vector the transverse energy E_T was calculated. This again was done for both underlying event models.

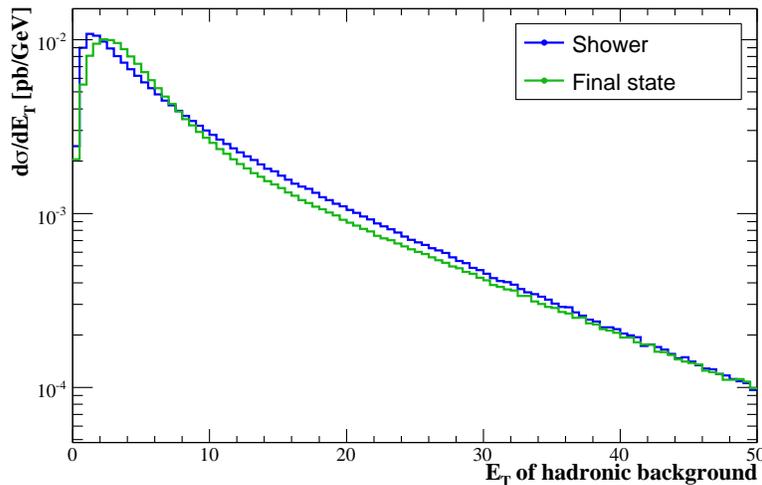


Figure 4.16: E_T distribution of the background. UA5 parametrization.

The two distributions, Fig. 4.16, 4.17, are rather similar. In both cases, the distributions peak for a small value of E_T and fall exponentially for high E_T . The peak for the UA5 parametrization, Fig. 4.16 is more narrow than that for the MPI model, Fig. 4.17. The UA5 sample as well shows a more significant difference between the shower and the final state distributions, the peak for the shower level is around $E_{T,\text{background}} = 1$ GeV whereas for the final state it is around $E_{T,\text{background}} = 3$ GeV. The peak for the final state gets wider as well.

For the MPI model, the peak is wider and peaks around $E_{T,\text{background}} = 4$ GeV, this holds for both the shower level and the final state.

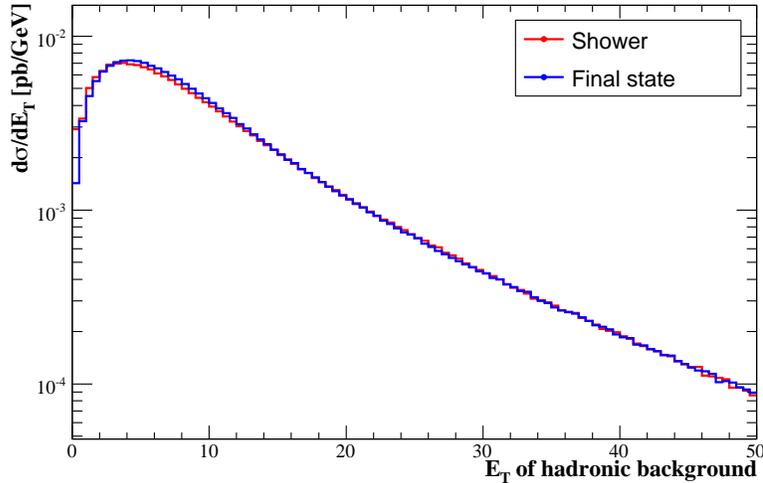


Figure 4.17: E_T distribution of the background. MPI underlying event model.

For $E_{T,\text{background}}$ between 10 GeV and 40 GeV, the UA5 sample shows less activity for the final state than after the shower. For the MPI model however, both distributions show exactly the same behavior.

The overall hadronic background in the final state seems to be rather small and the influence of the underlying event in this observable is not too important. This is studied more in the subsequent chapters.

To compare the two distributions the mean value of the distributions was investigated as well, Tab. 4.2. This confirmed the trends already seen by comparing the distributions, the MPI model has a slightly higher value than the UA5 parametrization and the final state of the UA5 parametrization has a slightly softer background than the shower.

	Shower level	Final state
UA5 parametrization	$E_T = 9.42$ GeV	$E_T = 9.14$ GeV
MPI model	$E_T = 10.36$ GeV	$E_T = 10.39$ GeV

Table 4.2: $E_{T,\text{mean}}$ of hadronic background.

Chapter 5

Third jets

Since in QCD processes like the parton shower additional jets can be generated by radiation of additional partons, the different steps are likely to contain more than the two tagging jets. In each step additional jets can be generated and thus increase the number of jets in the final state. Since collinear and soft emission is more likely than other, the tagging jets at matrix element level are expected to stay the leading jets after all the steps of the simulation and further jets are taken to be softer than the tagging jets and located near them in the lego-plot plane.

In order to investigate the properties of the additional jets, the third jet was examined as representative since it is the hardest and therefore most dominant jet of the background. The major sources for additional jet production were expected to be the shower and the underlying event, therefore in the analysis the distributions for the 3rd hardest jet at shower level and for the final states of the two underlying event models were compared. In addition, a comparison with a real next-to-leading order calculation was made to see the difference between the third jet calculated by the shower and the full matrix element calculation.

Not all VBF events were considered threejet events. Besides the VBF cut it was also requested that a third jet with $|y_3| < 4.5$ and $p_T > 20$ GeV was present.

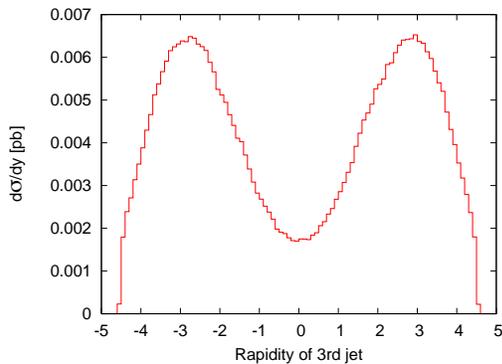


Figure 5.1: Rapidity of third jet in NLO calculation.

The rapidity distribution of the 3rd hardest jet after the shower resembles the rapidity distributions of the two tagging jets and is similar to the distribution obtained by the matrix element calculation for Higgs production in VBF at next to leading order, Fig. 5.2. Especially the central region, where a VBF event is expected to exhibit a rapidity gap stays free. A difference between the third jet produced by the shower and the one from the NLO calculation can be seen, Fig. 5.1, the peaks for the NLO jet are closer to the origin and here the central region shows more activity as well.

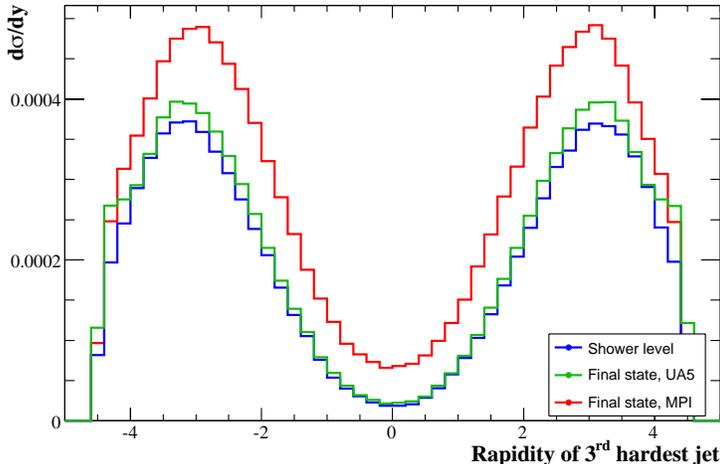


Figure 5.2: Rapidity of the 3rd hardest jet, shower level and final states for UA5 parametrization and MPI model.

In the UA5 parametrization the rapidity distribution for third jet in the final state of the simulation stays very close to that at shower level. It is slightly above it but leaves the shape itself unchanged, preserving the free central region as well. This is in contrast to the distribution given by the MPI underlying event model. While the overall shape here again is similar to that of the distribution after the shower, there is much more activity in the central region for the MPI model. This is to be expected, since the MPI model introduces additional QCD processes that lead to jets in the central region of the detector.

5.1 Rapidity of third jet relative to the tagging jets

To see the position of the 3rd jet with respect to the tagging jets and to investigate the impact of the background generated in the event simulation on the rapidity gap, several observables are introduced.

5.1.1 $y^* = y_3 - \frac{1}{2}(y_1 + y_2)$

y^* centers the position of the third jet in rapidity on the mean rapidity of the two tagging jets, Fig. 5.3. For events where the third jet is in the central region, y^* is about zero and for events where the third jet is close to the tagging jets y^* is around ± 3 , where the rapidity distributions for the tagging jets peak.

For the shower particles the distribution indeed shows a free central region, as is expected, since the shower enhances soft and collinear radiation and thus additional jets are close to the jets at matrix element level. The distribution shows two peaks around ± 3 , as expected as well. It has to be noted though that on the edges towards the central region smaller bumps can be found. This can be explained by the fact that the distribution for y^* can be viewed as a superposition of two distributions, one for the jets lying outside the rapidity gap and one for the third jet lying between the

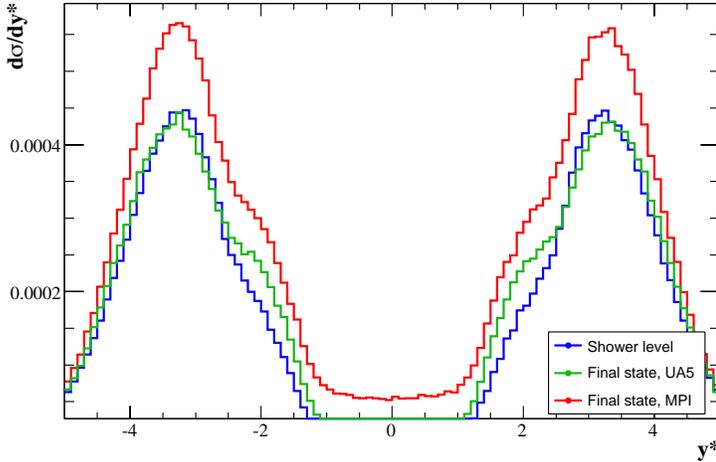


Figure 5.3: Distribution of $y^* = y_3 - \frac{1}{2}(y_1 + y_2)$, centering the rapidity of the 3rd jet on the mean value of the tagging jet rapidities.

two tagging jets. Third jets with the same rapidity as the tagging jets are suppressed, since in a circle around the tagging jets with radius given approximately by the R parameter of the jet finder hardly any additional jets are.

This is due to the fact that these jets have a high probability to be recombined into the normally much harder tagging jet. Therefore a separation of the set of third jets can be made, into one subset where the jets are between the tagging jets and one where they are outside, and a certain region in rapidity around the tagging jets where only little third jet activity appears.

The positions of the bumps on the inward edges of the shower distributions mark the peaks of the distribution for the jets between the tagging jets and the overall peaks are the peaks of the distribution for the jets outside the rapidity gap. This is illustrated in Fig. 5.4, where the distributions for the two possibilities for the position of the third jet and the distribution for all events are plotted and the latter can be seen as the combination of the former two.

The comparison with the third jet in the NLO calculation shows the same trend that could already be seen in the rapidity distribution for the third jet, Fig. 5.5. In the NLO case the peaks of the y^* distribution are more central and the central region as well shows a certain amount of activity, whereas the third jet created by the shower exhibited a region without any activity at all. For the NLO sample the inward slopes of the distributions show slight bumps as well, stemming from the jets between the tagging jets as well as explained above.

The distribution for the final state of the simulation with the UA5 parametrization

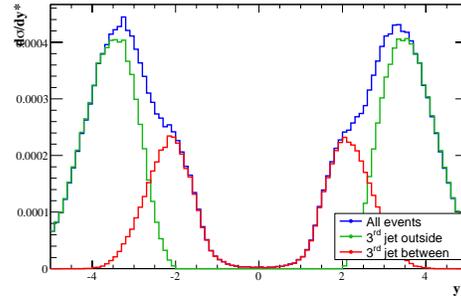


Figure 5.4: Distributions for the events where the 3rd jet lies between and outside the tagging jets and the complete sample.

again shows a very similar behavior to the third jet after the shower. Especially on the outside edges, the curves for shower and UA5 final state are very similar. A difference can be seen in the central region, around the bumps given by the peaks of the third jets lying between the tagging jets more activity can be found. This stems from the fact that the UA5 parametrization favors additional activity in the central region.

For the sample generated with the MPI underlying event model, the y^* distribution is rather similar to the other two distributions. It is clearly visible that more events are considered threejet events, thus the overall shape of the distribution is higher. The peaks are at about the same position as they are for the UA5 parametrization, this holds for the high peak for the jets outside as well as for the bumps indicating the peaks for the jets between the tagging jets. For this sample however the central region shows activity, evenly distributed between the tagging jets. This stems from the uncorrelated additional interactions in the MPI model that is spaced independently.

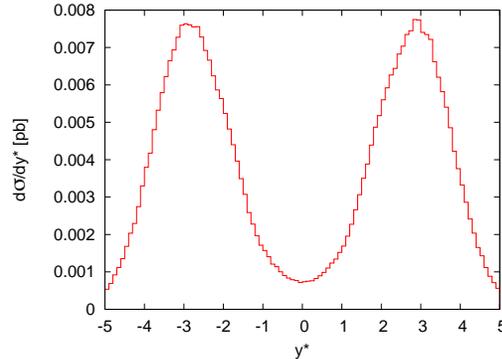


Figure 5.5: y^* in NLO calculation.

5.1.2 $|y^*| = y^*/|y_1 - y_2|$

To see the relative position of the third jet with respect to the tagging jets more clearly, the observable

$$|y^*| = y^*/|y_1 - y_2|$$

is introduced. This variable in addition normalizes the above observable y^* on the rapidity gap present in the event, in case the 3rd jet has the same rapidity as one of the tagging jets, $|y^*|$ will be $\pm\frac{1}{2}$, if the third jet is between the two tagging jets, the absolute value of $|y^*|$ will be smaller than $\frac{1}{2}$ and bigger if the 3rd jet lies outside.

The distributions for $|y^*|$ show sharper peaks than the y^* distributions but have a similar behavior in the overall shape, Fig. 5.7. For the shower, two high peaks can be seen around ± 0.62 and two lower peaks around ± 0.38 . The position of these peaks can be explained by the properties of the jet finder. As mentioned above, the jet finder causes the region around the hard tagging jets to be free of other jets, so in the rapidity region of the tagging jets only a reduced amount of activity is found. The rapidity

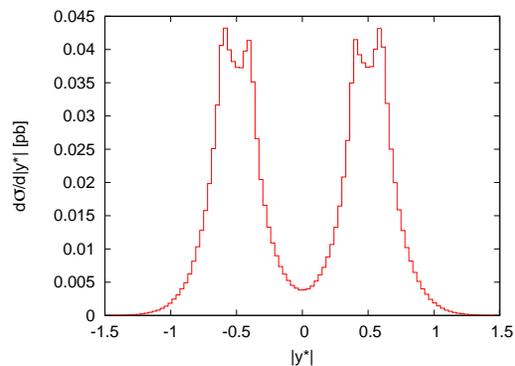


Figure 5.6: $|y^*|$ for the NLO calculation.

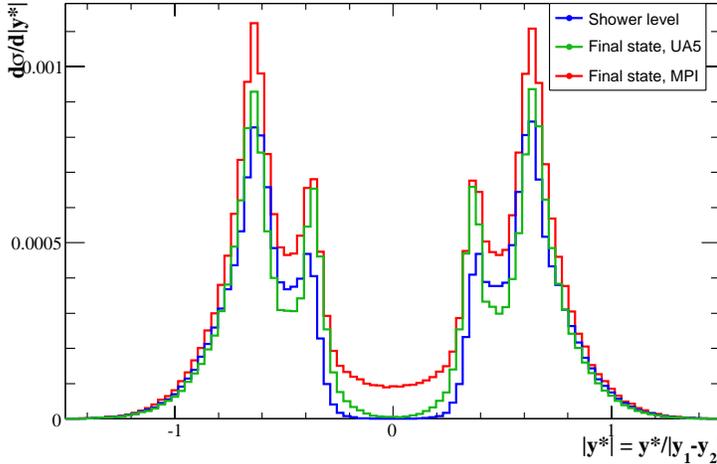


Figure 5.7: Distribution of $|y^*| = y^*/|y_1 - y_2|$, normalizing y^* to the length of the rapidity gap.

difference between the tagging jets by which y^* was divided to obtain $|y^*|$ is about four. The jet measure was set to be $R = 0.7$. The peaks in the $|y^*|$ distribution now can be found within the distance d to the tagging jet positions, which in this distribution were found to be $\pm\frac{1}{2}$ with

$$d \approx \frac{R}{\langle |y_1 - y_2| \rangle} \approx \frac{0.7}{4} = 0.175 ,$$

which fits quite well with the found peak positions.

Since this observable makes a distinction between inside and outside jets, it can be seen directly that the third jets after the shower tend to be close to the tagging jets in rapidity, as long as they are able to be resolved as individual jets. In addition, it can be seen directly that after the shower the third jets tend to be outside the rapidity gap between the tagging jets. This is again in contrast to the result for the NLO calculation, where the peaks for the events with the 3rd jet between the tagging jets are almost as high as the peaks for the outward jets, Fig. 5.6.

So in general it can be said for the comparison between the third jet stemming from the shower in comparison with the third jet in the NLO calculation that the shower underestimates the central region to some extent.

As in the case for y^* there is no big difference between the distributions at the shower level and the UA5 final state. The peaks in the distributions are almost identical and for the jets outside the rapidity gap almost no additional activity can be perceived. For the region $-0.5 < |y^*| < 0.5$ the UA5 parametrization gives more activity than the shower which again can be explained by the preference of the UA5 parametrization to generate particles in the central region.

For the MPI underlying event model the distribution is quite similar again to the UA5 parametrization. The peaks are almost at the same height, so here the same amount of activity can be found. A difference can be found for the regions where $0.4 < |y^*| < 0.6$ and $-0.6 < |y^*| < -0.4$ and the central region. Here the MPI model lies quite a lot above the other two distributions. From this can be concluded that the

additional scatterers also produce an excess of activity in the rapidity region of the tagging jets that is not recombined into the tagging jets. The activity in the central region is to be expected from the distribution of y^* , Fig. 5.3.

5.2 R separation of third jets to nearest tagging jet

The R separation ΔR of the third jet to its nearest tagging jet shows the distance in the lego-plot plane of the third jet to its nearest tagging jet. It is calculated from the rapidity and azimuthal differences

$$\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta y)^2}$$

All considered distributions show almost no activity for values of $\Delta R < 0.6$ and then a very steep rise between 0.6 and 0.9, where they peak, Fig. 5.8. Again this can be explained by the property of the jet finder, jets too close to the tagging jets get recombined into them. Since the dominant source of additional jets is radiation from the shower, most jets on the other hand are close to the tagging jet. Therefore most of the activity lies in an annulus around the tagging jets with inner radius ≈ 0.6 given by the R parameter in the jet finder.

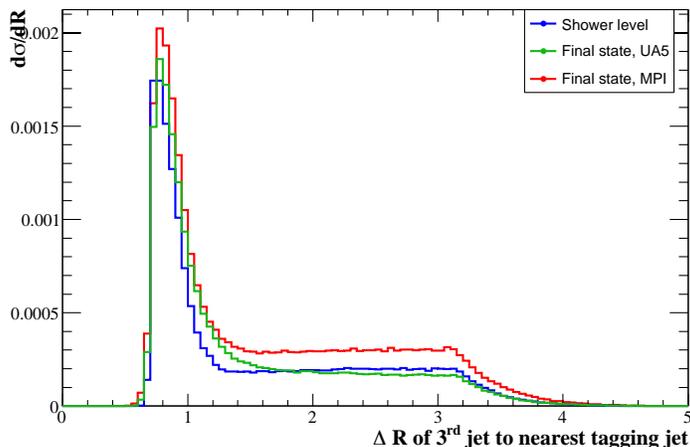


Figure 5.8: R separation of 3rd jet to the nearest tagging jet.

The distribution for the shower then falls again very steeply, until it reaches a rather low value around $\Delta R \approx 1.2$. From here follows a slight linear ascent until $\Delta R \approx \pi$, from where the distribution falls exponentially to zero around $\Delta R = 4$. The very low peak around $\Delta R \approx \pi$ can be explained by momentum conservation, in some cases a tagging jet with an angular direction φ emits a jet in the same rapidity region and the third jet tends to balance that in the opposite direction, $\varphi - \pi$.

The third jet in the NLO calculation has again a similar behavior to the third jet in the shower, but instead of falling down very steeply, an exponential decline between the peak at $\Delta R \approx 0.6$ and $\Delta R = \pi$ can be seen, where the distribution reaches an

edge and then again falls exponentially down to about $\Delta R \approx 4$ like in the shower case, Fig. 5.9.

The distribution for the MPI model peaks at a slightly higher value than the shower, but then shows a similar behavior. It reaches, after a steep decline a low value around $\Delta R \approx 1.4$ significantly higher than for the shower and then as well increases slightly linearly until $\Delta R \approx \pi$ from where it falls exponentially to zero.

The UA5 parametrization shows a little deviation to the other two distributions, here after the peak it falls of steeply as well, but instead of rising again it decreases slightly. From $\Delta R \approx \pi$ the behavior is the same as for the other two samples, an exponential fall to zero. This difference is probably due to the fact that the for the UA5 parametrization the jets after the shower are just smeared and thus the distribution does not show the more subtle structures as in the shower.

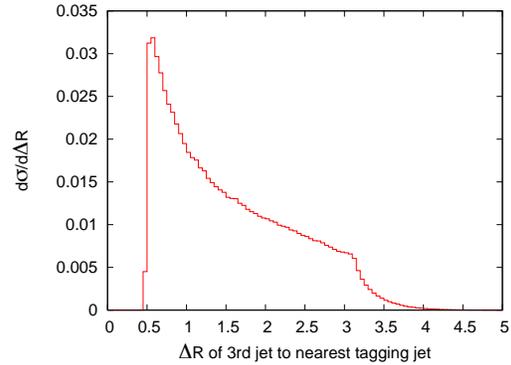


Figure 5.9: ΔR in the NLO calculation.

5.3 Transverse momentum

The transverse momentum spectrum of the third hardest jet turned out to be quite simple and stable in the simulation, Fig. 5.10.

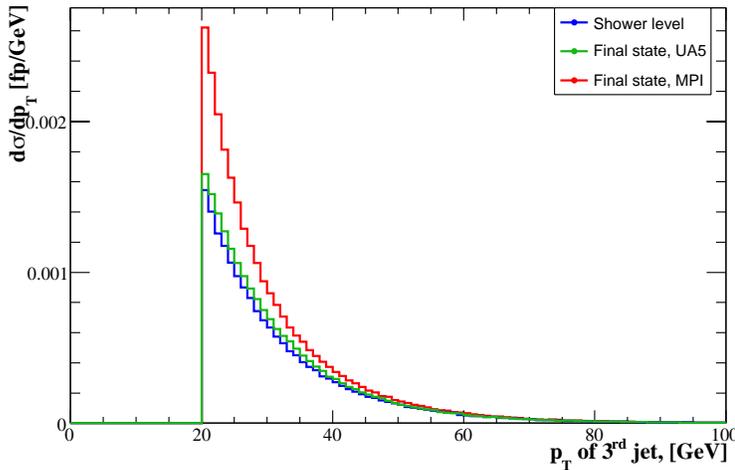


Figure 5.10: Transverse momentum of 3rd jet. Shower level and final state, UA5 parametrization and MPI UE model.

In all the distributions an exponential decrease can be found, starting from the $p_{T,min}$ cut on the third jet at 20 GeV and falling towards zero around 80 GeV in the case for the third jets in the simulation and around 100 GeV for the third jet in the next-to-leading order calculation.

Here again it can be seen that in the MPI sample more events have a third jet hard enough to be resolved as a jet than in the UA5 parametrization, which in turn shows an almost identical behavior to the shower distribution.

In the matrix element calculation for vector boson fusion with an additional jet, the cross section diverges if no minimum p_T cut is imposed on the third jet, for a full NLO calculation this divergence however cancels exactly with the divergence stemming from the virtual corrections, Fig. 5.11.

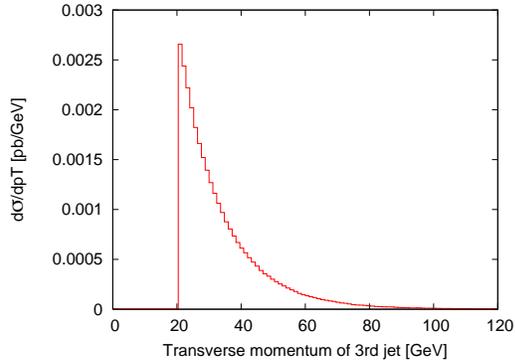


Figure 5.11: p_T of 3rd jet at NLO.

5.4 Rapidity and transverse momentum

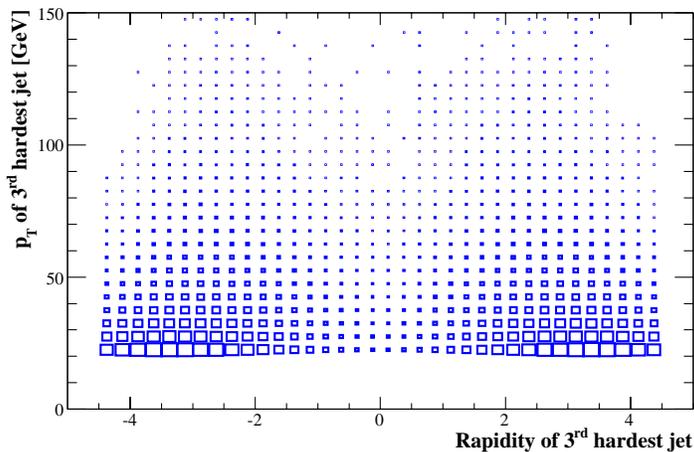


Figure 5.12: Rapidity and transverse momentum of 3rd jet. **Shower.**

the transverse momentum the less events. This holds for the complete distribution, as could be seen from the previous chapter 5.3, as well as for each bin in rapidity individually.

The rapidity showed for the lower p_T bins two peaks, symmetrical to the origin of the rapidity axis. For the lowest p_T region, $p_T < 30$ GeV these peaks are at ± 3.5 . This is expected since the rapidity distribution for the third jet peaks around this value.

To gain more insight into the third jets the distribution in rapidity and transverse momentum was examined. To see if there are correlations between these quantities, two-dimensional histograms were used. For the shower the majority of the events was found in the low p_T region with rapidity around ± 3.5 , Fig. 5.12. Regarding the distribution in view of the transverse momentum it can be found that the higher

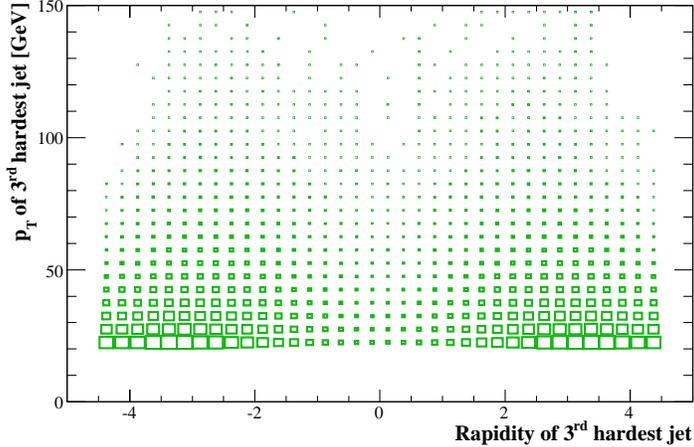


Figure 5.13: Rapidity and transverse momentum of 3^{rd} jet. Final state, **UA5 parametrization**.

particles entering the shower are two hard jets and during the parton shower a lot of emission takes place that leads to resolvable additional jets. Since in the shower the emission of soft and collinear partons is preferred, the majority of events has low p_T jets in the same rapidity region than the tagging jets.

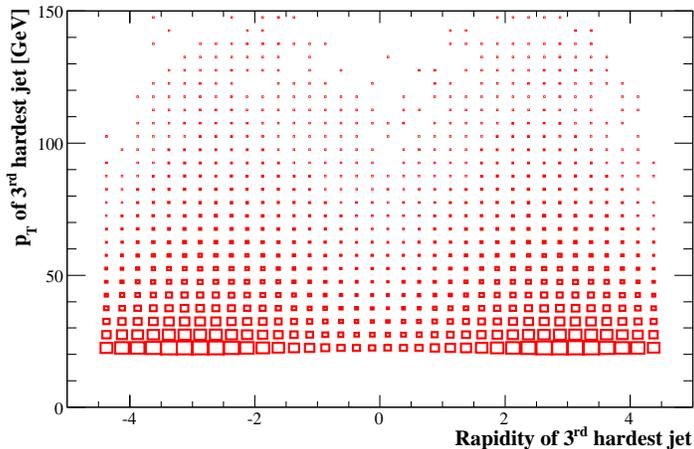


Figure 5.14: Rapidity and transverse momentum of 3^{rd} jet. Final state, **MPI underlying event model**.

region of rapidity for low transverse momentum and as well around the tagging jet rapidities for transverse momentum $p_T > 30$ GeV, Fig. 5.14.

From the comparison of these three distributions it can be seen that for high p_T

For higher values of the transverse momentum the position of the peaks shifts nearer to the origin, for the region where $40 \text{ GeV} < p_T < 100 \text{ GeV}$ the peaks are around ± 3 . For $p_T > 100 \text{ GeV}$ the peaks are not visible any more in the printout, since the number of events in this region is very small but can be found at $\approx \pm 2.5$.

The behavior found in this distribution is to be expected for VBF events. The final state

For the final state of the simulation with the UA5 parametrization the distribution is very similar to the shower distribution, Fig. 5.13. Only in the regions where there is already a lot of activity from the shower, the low p_T region around the tagging jet rapidities, a slight increase of activity can be found.

The sample with the MPI underlying event model is as well very similar to the shower distribution, but has more activity in the central region

the behavior of the third hardest jet is dominated by the shower, jets with a transverse momentum of $p_T > 50$ GeV do not seem to be influenced much by the hadronization, underlying event and decays of instable particles. The largest differences can be seen for small transverse momentum in the rapidity region around the tagging jets.

Chapter 6

Central jets

The VBF signal is characterized by a rapidity gap, so the central region between the tagging jets should be free of further jets. Since during the shower, hadronization and especially from the underlying event further jets can emerge, some of these jets will fall between the two tagging jets. In order to get a clear signal, events with hard jets in the central region should be rejected. Therefore events containing a central jet with a p_T above some cut $p_{T,\min}$ should be vetoed. This is also denoted "central jet veto".

To investigate the properties of the jets in the central region all the jets between the tagging jets for each event passing the VBF cuts were collected and sorted by p_T . Since the activity in the central region is expected to be dominated by the underlying event, a comparison of the observables created with different underlying event models and without any extra model was made.

6.1 Rapidity distributions of central jet

To see the orientation of the central jet in rapidity, a comparison of the rapidity distributions of the tagging jets and the central jet was made.

Here no $p_{T,\min}$ cut was imposed on the central jet to gain a more complete insight on the activity in the central region.

The rapidity distributions of the tagging jets are very similar, independent of the underlying event model used, if any, as could be seen in chapter 4. The distributions of the hardest central jet are very similar as well if a more sophisticated model for the underlying event is used, in the case of the MPI underlying event model, Fig. 6.2, there is a bit more activity than in the case of the UA5 parametrization, Fig. 6.1, especially the peaks are more pronounced for the MPI model.

Without an underlying event model, the central jet distribution is very similar to the rapidity distribution of the tagging jets, the central jets seem to stem mostly from the shower step, showing the rapidity distribution of a third jet at matrix element level, Fig. 6.3.

A direct comparison of the hardest central jets in the different simulations shows the similarities of the distributions from the underlying event models with a slight

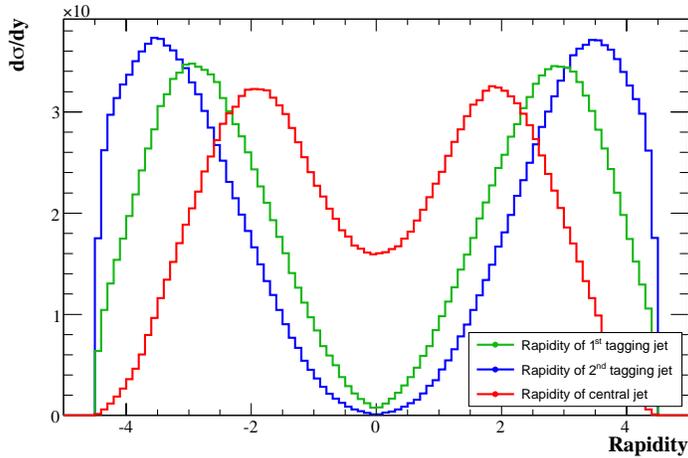


Figure 6.1: Rapidity of the two tagging jets and the hardest central jet. UA5 underlying event model. No $p_{T,\min}$ cut on 3rd jet.

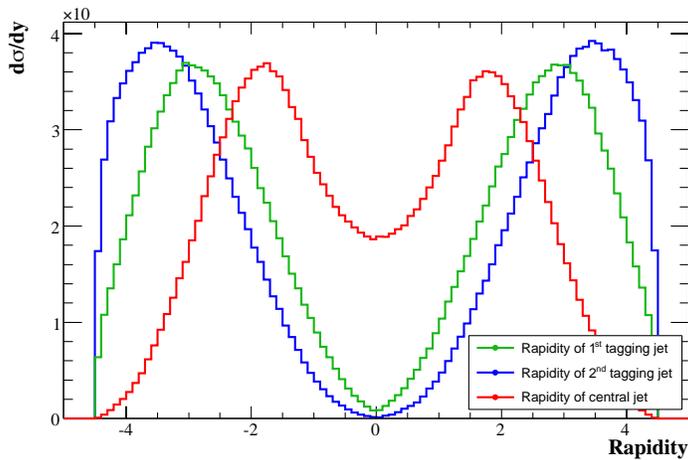


Figure 6.2: Rapidity of the two tagging jets and the hardest central jet. MPI underlying event model. No $p_{T,\min}$ cut on 3rd jet.

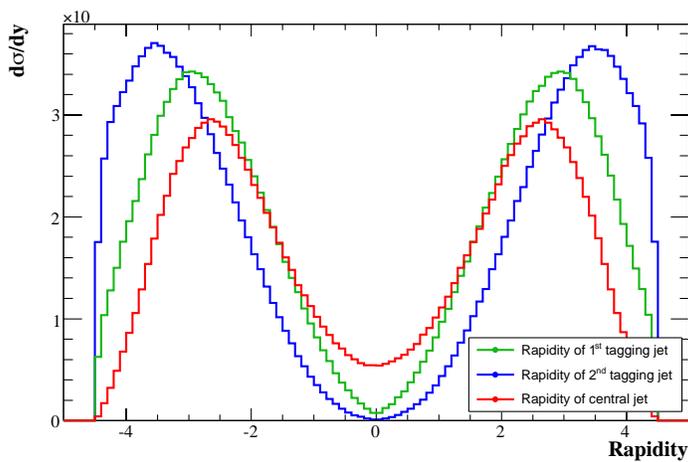


Figure 6.3: Rapidity of the two tagging jets and the hardest central jet. No underlying event model. No $p_{T,\min}$ cut on 3rd jet.

excess of activity in the central region for the MPI model and the deviation of the distribution without an elaborate model, Fig. 6.4.

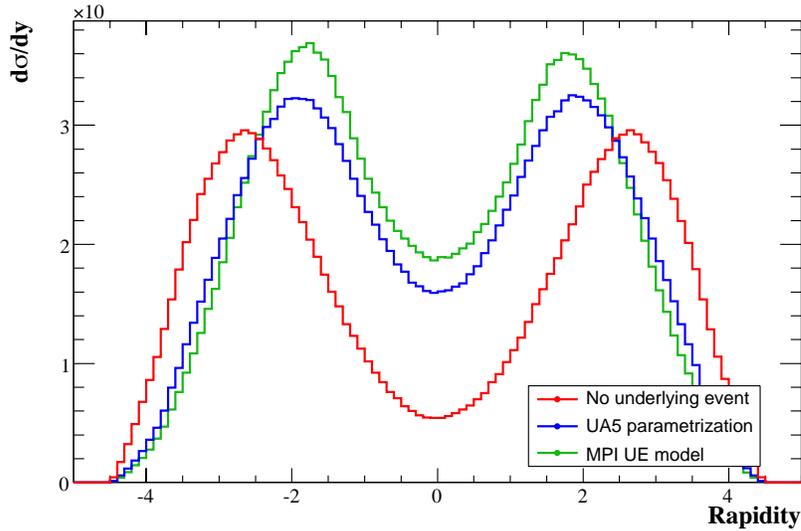


Figure 6.4: Rapidity of the hardest central jet. UA5, MPI and no underlying event model. No $p_{T,\min}$ cut on 3^{rd} jet.

6.1.1 Rapidity of central jet relative to tagging jets

To see more precisely the position of the hardest central jet in rapidity with respect to the tagging jets, the observable y^* was investigated as well. $y^* = y_{\text{central}} - \frac{1}{2}(y_1 + y_2)$ gives the absolute deviation of the rapidity of the central jet y_{central} from the mean value of the tagging jet rapidities y_1, y_2 .

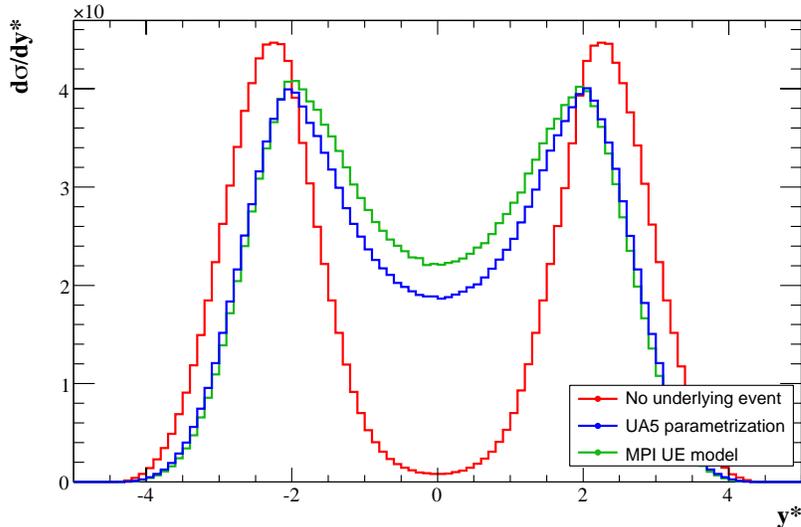


Figure 6.5: Distribution of y^* showing the rapidity of the hardest central jet with respect to the tagging jets. UA5, MPI and no underlying event model. No $p_{T,\min}$ cut on 3^{rd} jet.

Here again the similarities between the two underlying event models can be seen and the difference to the distribution obtained without an underlying event model, Fig. 6.5. The underlying event models produce quite a lot of activity in the central

region, the MPI model giving somewhat more activity here than the UA5 parametrization, whereas the sample without an extra underlying event model shows a rather free central region. All distributions peak at rapidities around ± 2 , the peaks in the distributions given by the MPI and UA5 models lying slightly more central than the peaks in the distribution given by the sample without underlying event model. Since the distributions of the tagging jets were rather similar for all three considered samples, this seems to indicate that in the case where no underlying event model was used the hardest central jet is closer to the tagging jets than in the cases where such a model is used.

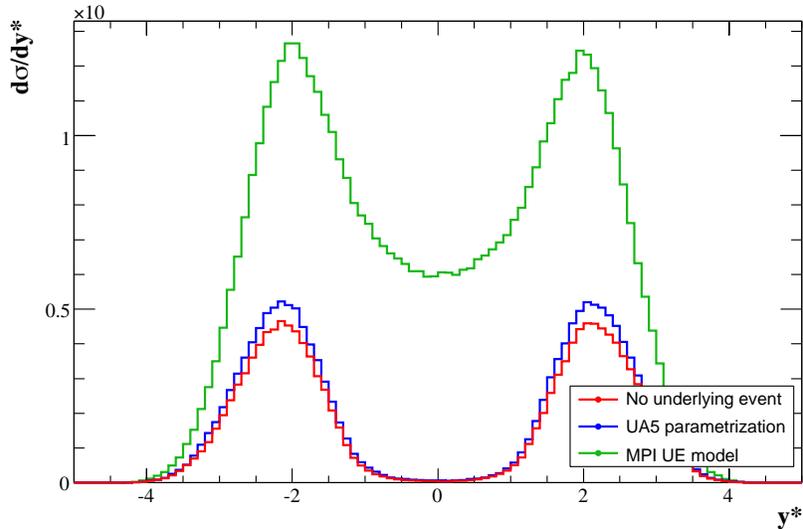


Figure 6.6: Distribution of y^* showing the rapidity of the hardest central jet with respect to the tagging jets, $p_{T,\min, \text{central}} = 10$ GeV. UA5, MPI and no underlying event model.

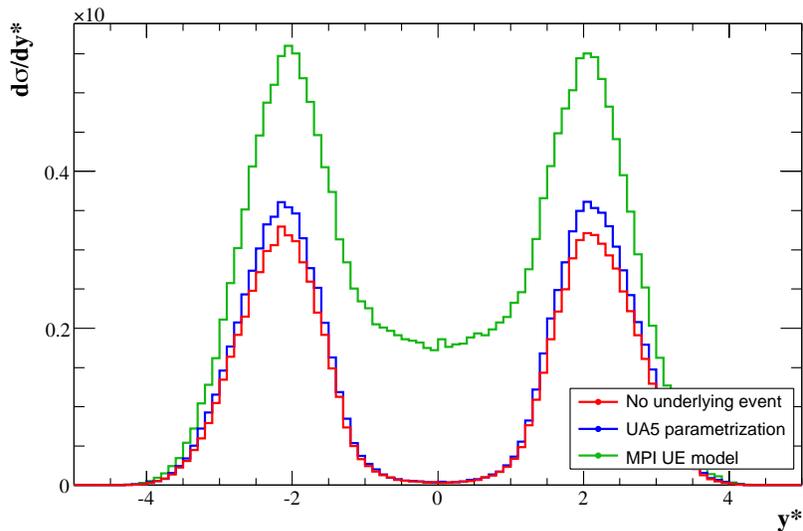


Figure 6.7: Distribution of y^* showing the rapidity of the hardest central jet with respect to the tagging jets, $p_{T,\min, \text{central}} = 15$ GeV. UA5, MPI and no underlying event model.

The figures Fig. 6.6, 6.7 and 6.8 show the differences in this distribution for several possible values for $p_{T,\min}$, the cut-off on the third hardest jet.

The higher the value for the cut-off, the less activity can be found in the central region, especially the UA5 parametrization shows a free central region in the y^* distri-

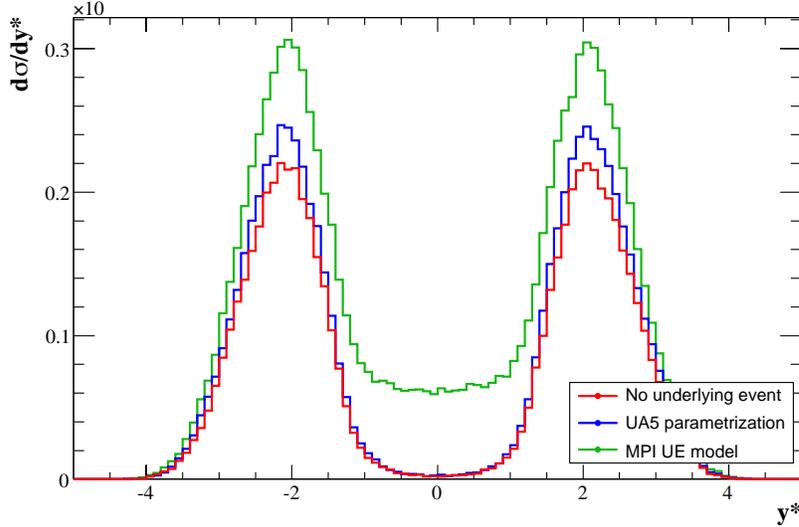


Figure 6.8: Distribution of y^* showing the rapidity of the hardest central jet with respect to the tagging jets, $p_{T,\min, \text{central}} = 20$ GeV. UA5, MPI and no underlying event model.

bution even for $p_{T,\min} = 10$ GeV, Fig. 6.6, whereas for the MPI sample the activity in the central region consists of harder jets, most of them with a transverse momentum above 10 GeV.

The peak positions for the underlying event samples remain and are not influenced by the value of $p_{T,\min}$, only for the sample without an underlying event model the peaks shift more inwards for higher $p_{T,\min}$ and have the same position as the peaks in the underlying event samples.

The higher the $p_{T,\min}$ cut, the more the distributions look alike, but even for $p_{T,\min} = 20$ GeV the MPI sample shows more activity than the other two.

To see the position of the central jets more precisely, the distribution of $|y^*| = y^*/|y_1 - y_2|$ is investigated. $|y^*|$, being normalized to the size of the rapidity gap between the tagging jets, shows the relative proximity of the hardest central jet to the tagging jets. Since the total rapidity difference of the tagging jets is not relevant any more, the proximity of the central jet to the tagging jets can be seen from this distribution. As described in section 5.1, for jets with rapidity equal to the rapidity of the first tagging jet, $y_3 = y_1$, $|y^*| = \frac{1}{2}$, for the case where $y_3 = y_2$, $|y^*| = -\frac{1}{2}$ and if the additional jet is between the tagging jets, $-\frac{1}{2} < |y^*| < \frac{1}{2}$.

Here again the difference in the central region can be seen, the two underlying event models producing quite a lot of activity, the MPI sample more so than the UA5 sample, whereas without an underlying event model there is only little activity, Fig. 6.9.

In addition, $|y^*|$ shows the region close to the tagging jets in more detail. It can be seen clearly that without an underlying event model most of the activity is along the tagging jets, the central jet thus stemming mostly from the shower. The smaller peaks around ± 0.38 can be explained as a relic of the jet finder, since particles that are close to a jet are more likely to be recombined into it, the area in the lego plot around the tagging jets is expected to be quite free of other jets, approximately a circle of radius R , the jet recombination parameter.

In the case of the two underlying event models, the central jet also tends to lie

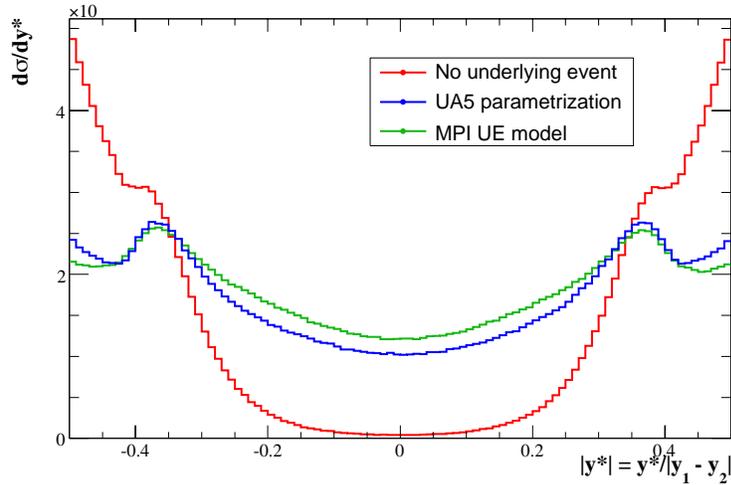


Figure 6.9: Distribution of $|y^*|$ showing the relative position of the rapidity of the hardest central jet relative to the tagging jets. UA5, MPI and no underlying event model.

close to the tagging jets, but significantly less so than in the case without a model. The distribution for the UA5 model is higher than that for MPI in the regions close to the tagging jets, thus reversing the behavior of the central region. Both distributions develop small peaks in the same region as the one without an underlying event model, slightly more towards the center.

6.2 p_T distribution of hardest central jet

The above distributions made clear that in a full event simulation of a VBF process there is quite a lot of activity in the central region, despite the fact that at matrix element level there is none. In order to investigate how much influence this activity has on the observables, the transverse momentum distribution of the hardest central jet is studied.

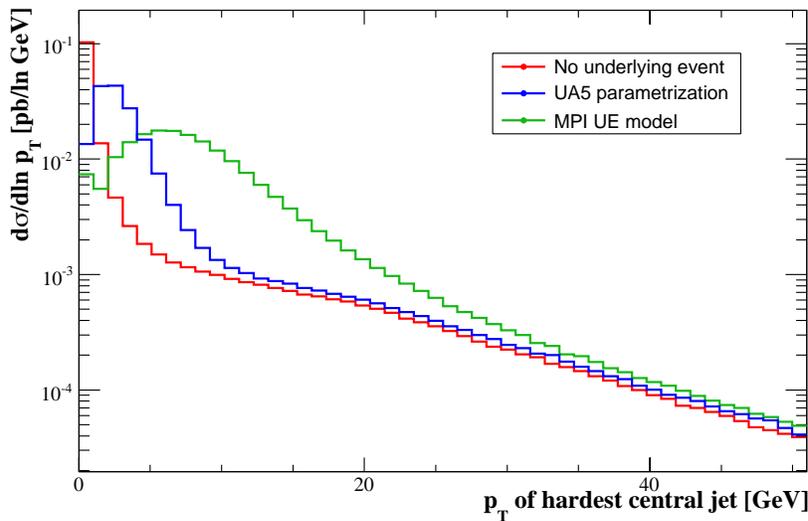


Figure 6.10: p_T distributions of the hardest central jet. UA5, MPI and no underlying event model.

The distribution without an underlying event model peaks at zero and then falls

of rapidly, for $p_T \geq 2$ GeV lying below the distributions for the two underlying event models, Fig. 6.10.

The distribution for the UA5 model peaks around 3 GeV, then falls off steeply, lying above the distribution for the sample without an underlying event model, from about 10 GeV the drop goes less steeply and the distribution is almost identical to the one from the sample without a model. From 20 GeV on the two samples cannot be distinguished any more.

The distribution for the MPI underlying event model exhibits a dip around 2 GeV. This kink originates from the cut on the cross section for the additional hard processes introduced by the model. This $p_{T,\min}$ cut is the main tuning parameter for the MPI model. The activity below that cut stems from the shower, washing out the steep cut.

After the dip the distribution rises and peaks around 6 GeV. For higher p_T it falls off approximately exponentially between the peak value and 15 GeV. From 20 GeV on it behaves identically to the other two distributions, falling slowly, but slightly above them.

Here it can be seen clearly that the underlying event models mostly add activity in the region of low transverse momentum and especially the difference between the two models compared, the MPI model working in a higher region of transverse momentum than the UA5 parametrization.

6.2.1 p_T distribution with respect to $|y^*|$

To gain more insight into the p_T distribution of the central jets a further distribution showing the p_T of the central jet depending on the tagging jets is investigated. For this a two dimensional histogram was used, where the p_T of the hardest central jet and $|y^*|$ of the event were booked. The boxes show the activity in the respective bin, the bigger a box is, the more events fall into it.

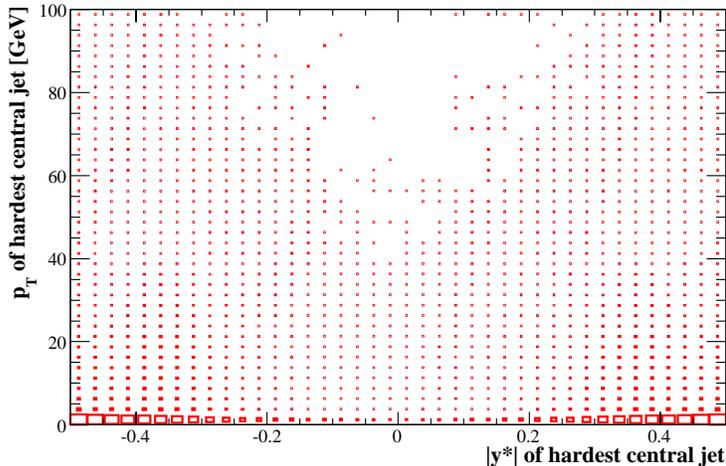


Figure 6.11: $|y^*|$ vs. p_T . No underlying event model.

It can be seen that without an underlying event model, Fig. 6.11, a lot of activity is in the low p_T region, very high activity can be found near the tagging jets and a

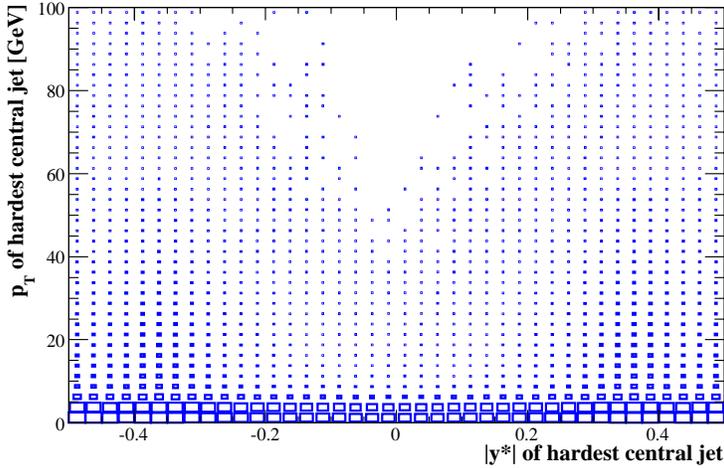


Figure 6.12: $|y^*|$ vs. p_T . UA5 underlying event model.

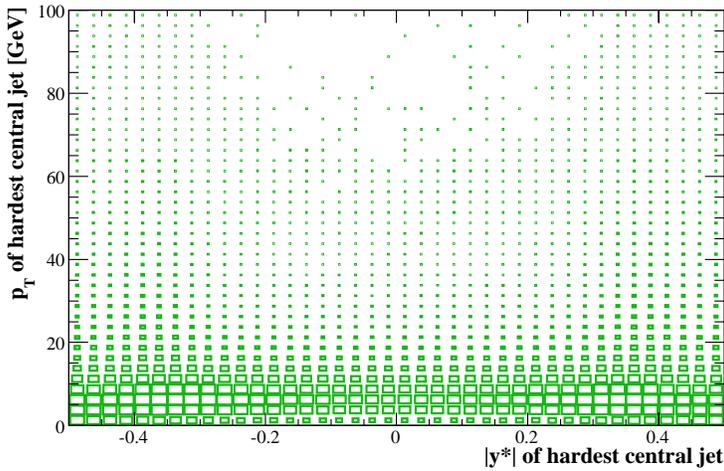


Figure 6.13: $|y^*|$ vs. p_T . MPI underlying event model.

diminishing amount towards the center. The same applies for central jets with higher p_T , so that the central region is more or less free of jets with p_T above 50 GeV. The histogram clearly shows that the major contribution to the central jets is very soft.

The underlying event models produce more activity in the whole central region, the MPI model, Fig. 6.13, more so than the UA5 parametrization, Fig. 6.12. Both models exhibit a similar behavior that is quite different to the distribution without a model. It can be seen clearly that the structure for the activity in the central region is significantly changed compared to the sample without an underlying event model. Especially the activity for higher transverse momentum can be perceived.

On the other hand for all samples the region above 50 GeV is more or less equal, almost no activity in the central region and some activity towards the tagging jets. It can be seen that the harder the jets get, the closer they are to the tagging jets. This is a typical behavior for jets emitted from hard jets during the parton shower.

In conclusion the underlying event models seem to produce a large amount of

low- p_T activity over the whole rapidity range, as seen before the MPI model is more dominant than the UA5 model, and the high- p_T region is quite unaffected by them.

6.3 Veto efficiency

Events with very hard central jets do disturb the VBF signature, so a sensible strategy is to veto these events. A possible way of doing this can be by rejecting events with a central jet with transverse momentum above some cut-off $p_{T,\min}$.

To see the percentage of events passing such a veto, the p_T distributions of the hardest central jets can be consulted.

Since for a given cut-off $p_{T,\min}$ all events with a central jet with $p_{T,jet} \geq p_{T,\min}$ will be vetoed, the number of events being vetoed will be the sum of all these events. By summing the bins in the p_T distribution starting from high p_T , a distribution showing the number of events being vetoed by a veto with $p_{T,\min}$ is created.

Dividing the bins of this distribution by the number of events passing the VBF cuts, the histogram shows the fraction of the events passing the veto depending on the $p_{T,\min}$, to get the number of events passing the veto, a histogram with $1-(bin\ content)$ is plotted, Fig. 6.14.

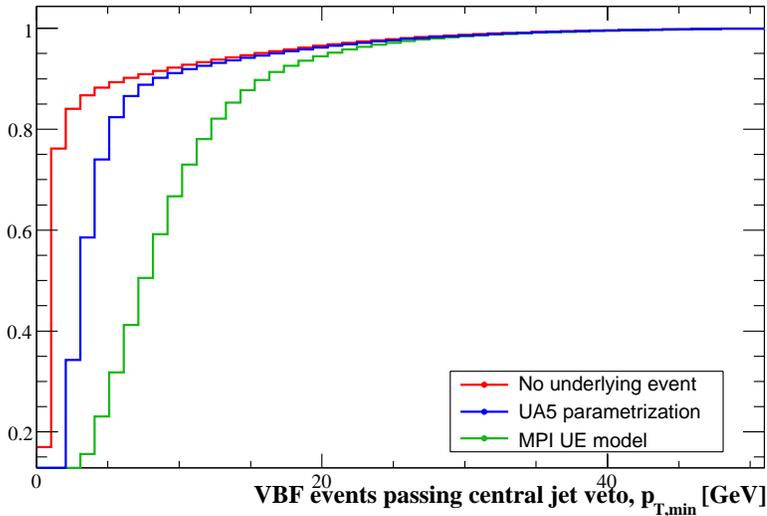


Figure 6.14: Veto efficiency. The fraction of events passing a central jet veto for a given $p_{T,\min}$ is plotted. MPI, UA5 and no underlying event model.

From this distribution it is clear that without an underlying event model only a very small fraction of events will be vetoed, with an underlying event model the efficiency for $p_{T,\min} < 25$ GeV is less than without a model, for $p_{T,\min} > 25$ GeV only a very small fraction of events will be vetoed.

The distributions for the sample without an underlying event model and the UA5 parametrization meet for a veto at $p_{T,\min} = 15$ GeV, the efficiency being at about 94%, below that the UA5 model shows a considerably lower veto efficiency than the sample without a model.

The MPI sample yields a smaller efficiency for low $p_{T,\min}$, for $p_{T,\min} = 10$ GeV the veto efficiency is about 73%, but the efficiency approaches that of the other two

samples around 25 GeV, where 96% of the events pass the veto.

As a result it can be stated that a reasonable veto can be imposed with $p_{T,\min}$ between 10 GeV and 20 GeV without losing a large fraction of the events of interest.

For the sample created with the MPI model the results have to be treated with caution, since the tuning for this model so far only has been done with data from the Tevatron collider. The main tuning parameter for this model is the $p_{T,\min}$ of the additional scatterers, this defines the total cross section obtained. The influence of this cut-off on the considered observables was not investigated. Tevatron data imposes hard restrictions on the allowed values for this tuning parameter, and since no data for higher center-of-mass energies is available, this tune was used.

Chapter 7

p_T shifts

To gain more insight into the structure of the events at the different steps in the simulation the tagging jets in the steps were examined in comparison. Since the observables were found to be stable under the whole event simulation, the jets themselves are supposed to be stable as well. This means that the direction of a jet should not change much as well as its transverse momentum.

To get an insight into that, the two tagging jets at matrix-element level were compared with the tagging jets after the shower and after the complete simulation. This was done by calculating the p_T difference between the jets at matrix element and at shower level respectively in the final state.

$$p_{T,\text{shift}} = p_{T,\text{shower}} - p_{T,\text{ME}}$$

This was compared to the same calculation with the additional condition that the jets at shower level respectively in the final state lie in the same direction as the jet at matrix element level. To ensure this, the jets were required to have a maximum R -separation of $\Delta R = 0.6$. These events are called *correlated* henceforth.

To see the influence of the underlying event model, a comparison for the two used models, the UA5 parametrization and the MPI model was made.

7.1 p_T shift of hardest tagging jet

For the hardest tagging jet, the p_T shift from matrix element level to shower level in case of the UA5 parametrization turned out to be rather moderate, for the majority of events the p_T of the hardest jet was unchanged or shifted only a little bit, Fig. 7.1. In case the p_T was changed, the jets tended to loose transverse momentum rather than gain it, which is to be expected in the shower, where especially hard jets can emit partons outside the soft and collinear region that then are resolved as additional jets rather than be recombined into the original jet.

The sample created with the MPI underlying event model the distribution looked very similar, the p_T for the majority of events was unchanged as well, and more events loose transverse momentum rather than gain it. Only a slight difference to the UA5 sample is perceivable, in comparison more events gain transverse momentum in the

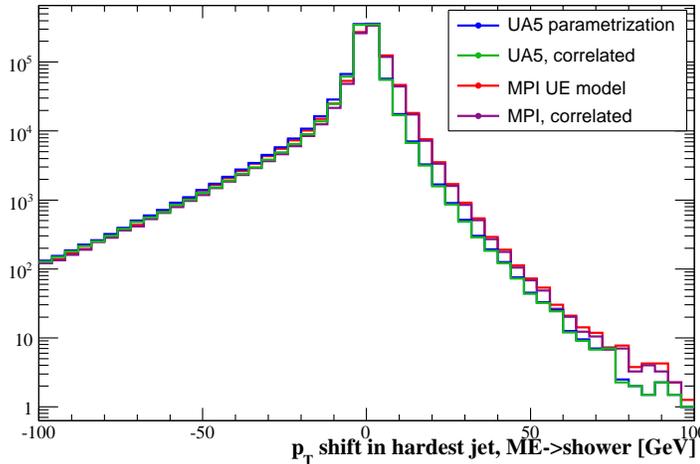


Figure 7.1: p_T shift of hardest tagging jet, matrix element level to shower level. Distribution for all jets and correlated jets, UA5 parametrization and MPI UE model.

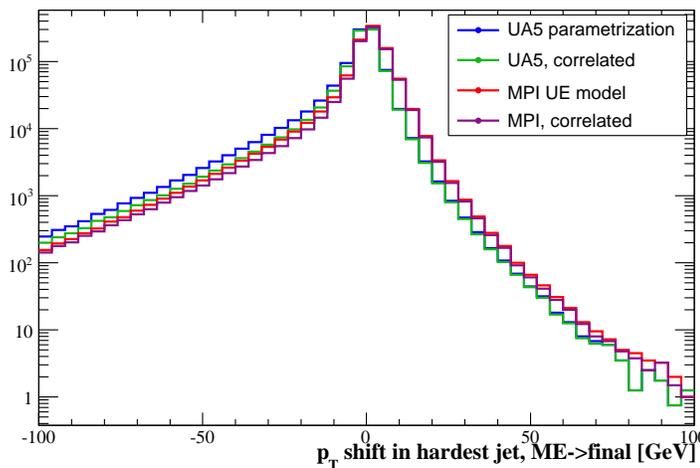


Figure 7.2: p_T shift of hardest tagging jet, matrix element level to final state. Distribution for all jets and correlated jets, UA5 parametrization and MPI UE model.

MPI sample, due to particles from the additional scatterings that are recombined into the jet and thus increase its energy and transverse momentum.

For both underlying event models, the distribution for the correlated jets are more or less exactly the same as in the case for all jets, only for a minority of events the hardest jet at matrix element level seems not to be correlated with the hardest jet after the shower. This can be seen from the absolute values as well, Tab. 7.1

The p_T shift distribution from matrix element level to the final state for the UA5 parametrization is almost the same as the shift from matrix element level to shower level, Fig. 7.2. Only a few events more produce a shift, and the distribution for the correlated jets is again rather close to the distribution for all the jets.

The same distribution for the MPI sample as well is quite similar to that for the shift from matrix element level to shower, but in comparison with the UA5 sample less events lose transverse momentum, which is due to the fact that the events rather gain energy by the additional scatterers.

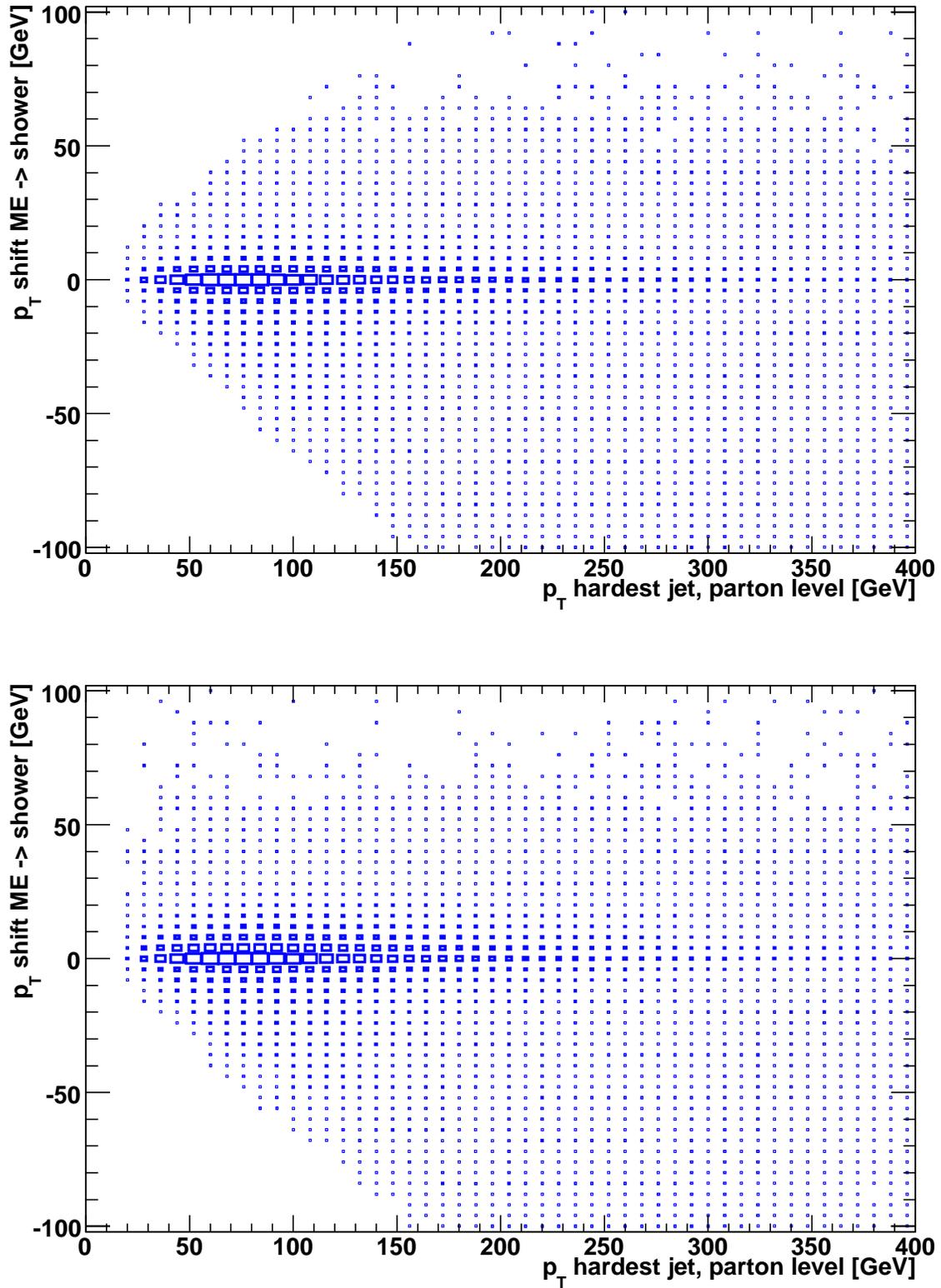


Figure 7.3: p_T shifts of hardest jet, matrix element to shower, relative to p_T at matrix element level, all jets, UA5 parametrization (above) and MPI model (below).

To see more in detail in which regions the p_T shifts occur, the distribution were plotted as well depending on the transverse momentum of the jet at matrix element level. Since there are no big differences between the correlated events and all events, only the distributions for all events are shown.

	UA5 parametrization		MPI model	
	All events	Correlated	All events	Correlated
Complete:				
Total	$3.842 \cdot 10^6$	$3.682 \cdot 10^6$	$3.841 \cdot 10^6$	$3.656 \cdot 10^6$
No shift	0.567	0.561	0.452	0.462
$ p_T \text{ shift} < 10 \text{ GeV}$	0.901	0.910	0.875	0.88274
$p_{T,\text{hard}} < 40 \text{ GeV}$:				
Total	$0.035 \cdot 10^6$	$0.029 \cdot 10^6$	$0.035 \cdot 10^6$	$0.028 \cdot 10^6$
No shift	0.755	0.789	0.605	0.650
$ p_T \text{ shift} < 10 \text{ GeV}$	0.996	0.998	0.974	0.979
$40 \text{ GeV} < p_{T,\text{hard}} < 120 \text{ GeV}$:				
Total	$2.409 \cdot 10^6$	$2.274 \cdot 10^6$	$2.274 \cdot 10^6$	$2.406 \cdot 10^6$
No shift	0.584	0.601	0.483	0.498
$ p_T \text{ shift} < 10 \text{ GeV}$	0.925	0.936	0.901	0.911
$120 \text{ GeV} < p_{T,\text{hard}} < 200 \text{ GeV}$:				
Total	$1.082 \cdot 10^6$	$1.065 \cdot 10^6$	$1.083 \cdot 10^6$	$1.063 \cdot 10^6$
No shift	0.498	0.504	0.406	0.411
$ p_T \text{ shift} < 10 \text{ GeV}$	0.868	0.875	0.839	0.846
$p_{T,\text{hard}} > 200 \text{ GeV}$:				
Total	$0.316 \cdot 10^6$	$0.315 \cdot 10^6$	$0.315 \cdot 10^6$	$0.315 \cdot 10^6$
No shift	0.443	0.445	0.359	0.360
$ p_T \text{ shift} < 10 \text{ GeV}$	0.824	0.826	0.839	0.846

Table 7.1: p_T shift of the hardest tagging jet from matrix element level to shower. All events and $p_{T,\text{ME}}$ regions. Number of all events in considered region, fraction of events with shift in 'zero bin' ($|\Delta p_T| \leq 2\text{GeV}$) and fraction of events with shift ($|\Delta p_T| \leq 10\text{GeV}$).

First it can again be seen from the values in Tab. 7.1, that for most of the events the hardest jet at matrix element level is correlated with the hardest jet after the shower, since only a few per cent of the events get vetoed by this restriction. This holds for the total amount of events as well as for the different regions of transverse momentum at matrix element level examined in Tab. 7.1.

The p_T shifts for the hardest tagging jet from matrix element level to shower are very small. About half of the events have a shift of absolute value less then 2 GeV, corresponding to the zero bin in the histogram, 56% for the UA5 parametrization and 45% for the MPI model. In the overall distribution, 90% respectively 88% of the events have a shift with absolute value less than 10 GeV. For the different p_T intervals the picture is rather similar, for small transverse momenta at matrix element level the shifts tend to be smaller and for large transverse momenta the shifts are larger. This difference is not surprising, since a jet with small p_T only can radiate a small amount

of energy.

This is very similar for the correlated events, from Tab. 7.1 it can be seen that the shifts tend to be a little smaller for correlated events since the percentage of low shift events is slightly higher. This can be expected, since the radiation of high p_T partons tends to change the direction of the original jet due to energy-momentum conservation.

The distribution for the change in transverse momentum of the hardest jet from matrix element level to the final state of the simulation with the UA5 parametrization modeling the underlying event is very similar to the distribution for the change from matrix element level to shower level, Fig. 7.4. This can be confirmed by the values in Tab. 7.2. The changes in transverse momentum tend to get higher, since the percentage of events with $|\Delta p_T| < 2$ GeV is sixteen and for $|\Delta p_T| < 10$ GeV six percentage points lower than for the change between hard process and shower, but this effect is rather small. A similar picture holds for the shifts in the different p_T intervals examined separately. The shift for events with low p_T at matrix element level tends to be smaller than for events with high p_T .

The changes for correlated events are again smaller than for all events, for this distribution more significantly than for the change to shower level, but again only a rather small fraction of events was rejected. The distribution showing the change in transverse momentum relative to the transverse momentum at matrix element level for correlated events hardly shows a change to the distribution of all events.

In the case of the MPI model the overall picture for the changes in transverse momentum of the hardest jet is similar to the result of the UA5 parametrization, Fig. 7.5. But some differences can be perceived. First, quite a lot of events show a slight gain in transverse momentum. Additionally there are more events in the MPI model that get a large change to higher p_T than in the UA5 parametrization, while for the UA5 sample events with changes up to 70 – 80 GeV can be found only for higher p_T values at Matrix element level, shifts like that already occur for low values of $p_{T,ME}$ in the MPI model. For the shifts to smaller transverse momentum the distributions are almost identical.

This can be seen as well from the values in Tab. 7.2, where the amount of events with no or low shift is significantly lower in the MPI sample.

In these samples the difference between the correlated and uncorrelated events is again not very large but more significant than at shower level. This can be seen directly from the amount of events rejected by the correlation condition.

Comparing the percentage of correlated events for both underlying event models used it can be seen that the numbers are very alike, especially for the 'zero bin', the bin where the events with $|\Delta p_T| < 2$ GeV are booked.

In conclusion it can be said that the hardest jet in the VBF events are not affected much by the full event simulation, the majority keeps their direction and transverse momentum, the biggest impact on this stems from the underlying event.

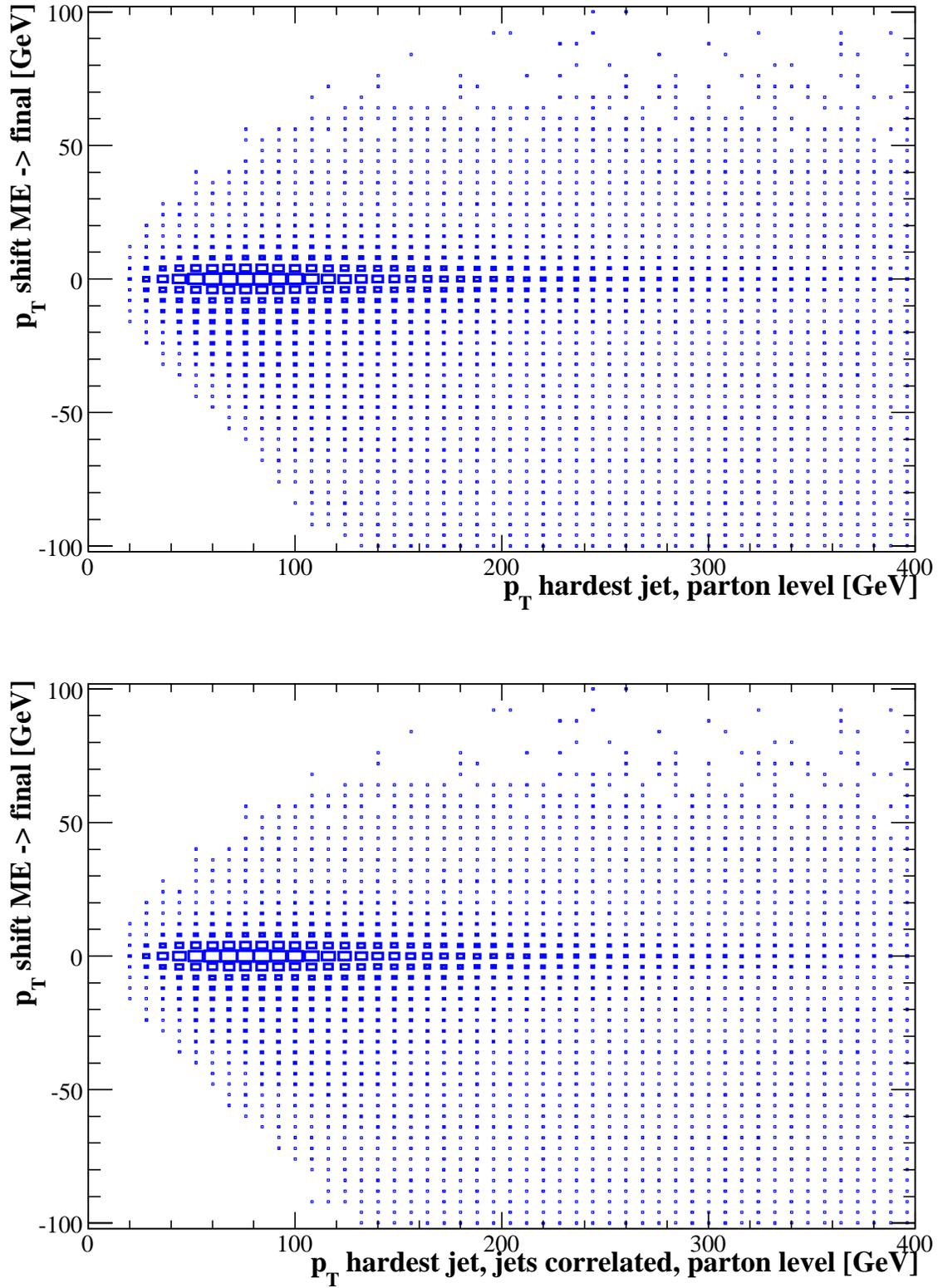


Figure 7.4: p_T shifts in hardest jet, ME to final state, relative to p_T at ME level, UA5 parametrization, all jets (above) and correlated (below).

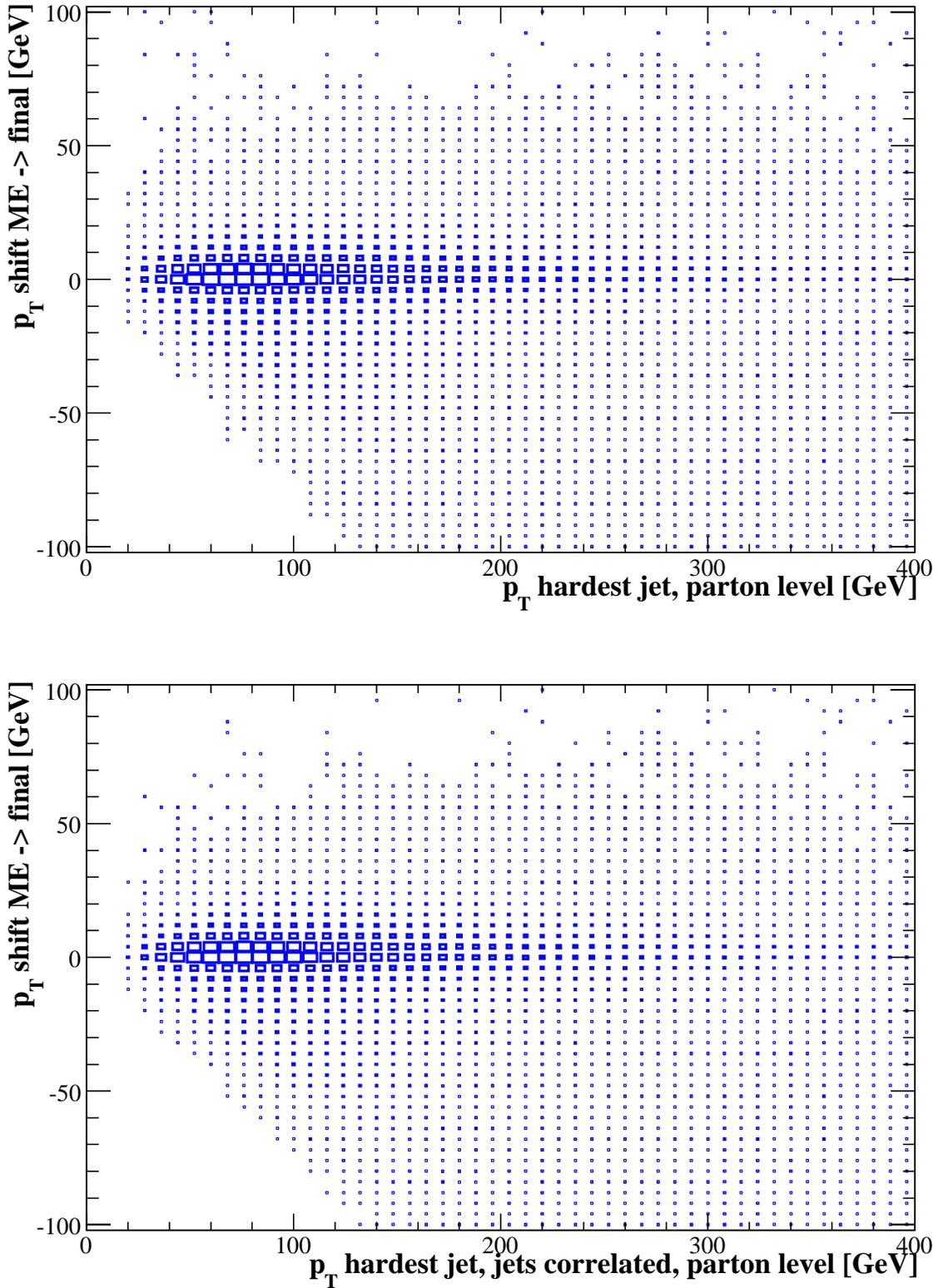


Figure 7.5: p_T shifts in hardest jet, ME to final state, relative to p_T at ME level, MPI model, all jets (above) correlated (below).

	UA5 parametrization		MPI model	
	All events	Correlated	All events	Correlated
complete				
Total	$3.842 \cdot 10^6$	$3.565 \cdot 10^6$	$3.812 \cdot 10^6$	$3.616 \cdot 10^6$
No shift	0.406	0.423	0.345	0.352
$ p_T \text{ shift} < 10 \text{ GeV}$	0.847	0.870	0.855	0.866
$p_{T,\text{hard}} < 40 \text{ GeV}$				
Total	$0.035 \cdot 10^6$	$0.025 \cdot 10^6$	$0.035 \cdot 10^6$	$0.026 \cdot 10^6$
No shift	0.548	0.610	0.472	0.497
$ p_T \text{ shift} < 10 \text{ GeV}$	0.966	0.980	0.971	0.976
$40 \text{ GeV} < p_{T,\text{hard}} < 120 \text{ GeV}$				
Total	$2.409 \cdot 10^6$	$2.175 \cdot 10^6$	$2.406 \cdot 10^6$	$2.216 \cdot 10^6$
No shift	0.427	0.453	0.366	0.378
$ p_T \text{ shift} < 10 \text{ GeV}$	0.867	0.900	0.882	0.897
$120 \text{ GeV} < p_{T,\text{hard}} < 200 \text{ GeV}$				
Total	$1.082 \cdot 10^6$	$1.051 \cdot 10^6$	$1.083 \cdot 10^6$	$1.059 \cdot 10^6$
No shift	0.374	0.382	0.312	0.317
$ p_T \text{ shift} < 10 \text{ GeV}$	0.817	0.833	0.817	0.826
$p_{T,\text{hard}} > 200 \text{ GeV}$				
Total	$0.316 \cdot 10^6$	$0.313 \cdot 10^6$	$0.318 \cdot 10^6$	$0.316 \cdot 10^6$
No shift	0.337	0.339	0.280	0.281
$ p_T \text{ shift} < 10 \text{ GeV}$	0.777	0.782	0.769	0.772

Table 7.2: p_T shift of the hardest tagging jet from matrix element level to final state. UA5 parametrization and MPI UE model. Number of all events in considered region, fraction of events with shift in 'zero bin' ($|\Delta p_T| \leq 2\text{GeV}$) and fraction of events with shift ($|\Delta p_T| \leq 10\text{GeV}$).

7.2 p_T shift of 2nd hardest tagging jet

The second hardest jet is in the majority of events not affected much by the shower as well. In most cases there is no or only little change in transverse momentum for the 2nd hardest jet from matrix element level to shower level as well as to the final state.

From the distributions of the overall shifts, Fig. 7.6 and 7.7 it can be seen that here the difference between all events and only correlated events is much larger than for the hardest jet. While the curves for changes to smaller transverse momentum looks similar to the case of the hardest jet, a large difference for the changes to larger p_T can be found, here a lot of events are not correlated with the events at matrix element level.

The difference between the distributions for all events and the correlated events is very likely an effect imposed by the shower. For events where the second hardest jet is of small transverse momentum, it is possible that during the shower another jet gets generated that is harder than the 2nd tagging jet. This new second hardest jet

then is normally not in the same direction than the original second tagging jet. So events where this happens do not pass the correlation requirement.

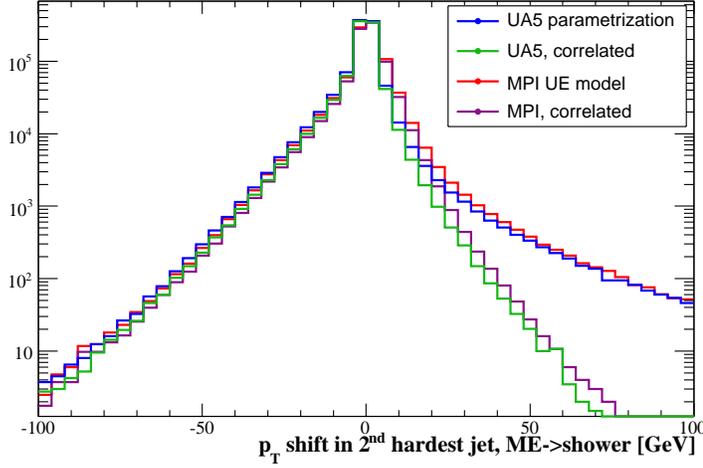


Figure 7.6: p_T shift of 2nd hardest tagging jet, matrix element level to shower level. Distribution for all jets and correlated jets, UA5 parametrization and MPI UE model.

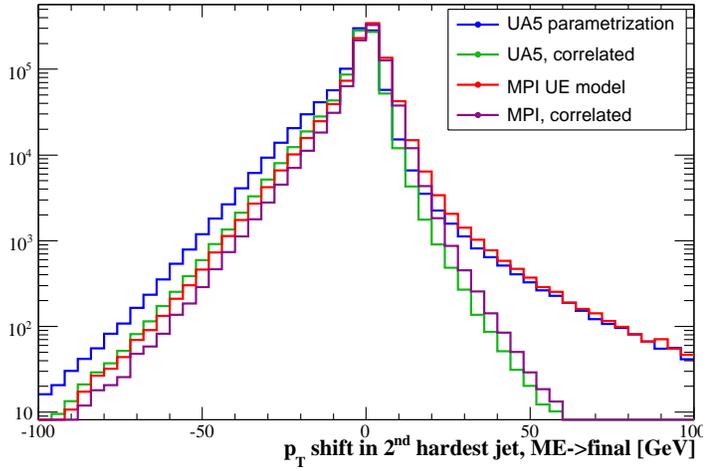


Figure 7.7: p_T shift of 2nd hardest tagging jet, matrix element level to final state. Distribution for all jets and correlated jets, UA5 parametrization and MPI UE model.

This can be seen more in detail in the distributions where the change in transverse momentum is plotted in combination with the distribution for the transverse momentum of the second hardest jet at matrix element level, Fig. 7.8 and Fig. 7.9. It is clearly visible that without the correlation in quite a lot of events with $p_T < 100$ GeV the transverse momentum increases very much, by as much as 100 GeV even for events with $p_T = 20$ GeV at matrix element level.

Comparing this with the distribution for the correlated events it can be seen that for all these events with a high change in transverse momentum there is no correlation between the 2nd hardest jet at matrix element level and shower.

Tab. 7.3 confirms this, especially for the lower p_T region at matrix element level the correlated events the absolute numbers of events decrease about 6.3% resp. 7.9%

in total and about the same fraction for the lower p_T regions, whereas the fraction with no or small changes increases by about three percentage points.

	UA5 parametrization			
	All events	Correlated	All events	Correlated
complete:				
Total	$3.857 \cdot 10^6$	$3.599 \cdot 10^6$	$3.857 \cdot 10^6$	$3.552 \cdot 10^6$
No shift	0.583	0.608	0.486	0.510
$ p_T \text{ shift} < 10 \text{ GeV}$	0.905	0.925	0.889	0.910
$p_{T,\text{hard}} < 40 \text{ GeV}$				
Total	$1.052 \cdot 10^6$	$0.980 \cdot 10^6$	$1.053 \cdot 10^6$	$0.958 \cdot 10^6$
No shift	0.661	0.695	0.549	0.584
$ p_T \text{ shift} < 10 \text{ GeV}$	0.944	0.969	0.929	0.958
$40 \text{ GeV} < p_{T,\text{hard}} < 120 \text{ GeV}$				
Total	$2.694 \cdot 10^6$	$2.517 \cdot 10^6$	$2.693 \cdot 10^6$	$2.492 \cdot 10^6$
No shift	0.558	0.580	0.466	0.486
$ p_T \text{ shift} < 10 \text{ GeV}$	0.893	0.911	0.877	0.895
$120 \text{ GeV} < p_{T,\text{hard}} < 200 \text{ GeV}$				
Total	$0.111 \cdot 10^6$	$0.102 \cdot 10^6$	$0.111 \cdot 10^6$	$0/101 \cdot 10^6$
No shift	0.444	0.466	0.370	0.389
$ p_T \text{ shift} < 10 \text{ GeV}$	0.818	0.843	0.798	0.822

Table 7.3: p_T shift of the 2^{nd} hardest tagging jet from matrix element level to shower. Number of all events in considered region, fraction of events with shift in 'zero bin' ($|\Delta p_T| \leq 2 \text{ GeV}$) and fraction of events with shift ($|\Delta p_T| \leq 10 \text{ GeV}$).

As in the case for the hardest tagging jet, the change from matrix element level to the final state of the event simulation is very similar to the change between matrix element level and the shower, Fig. 7.7.

This holds especially for the distribution for the UA5 parametrization, for small p_T at matrix element level quite large increases can be seen for all events, whereas for events with correlated jets these events were rejected, Fig. 7.10.

In the overall picture, it can be seen that the transverse momentum tends to get smaller, in the $p_{T, \text{ME}}$ region between 20 GeV and 60 GeV, where the majority of events lies, a shift to lower transverse momentum seems more favored than an increase. From the distribution for the correlated events it can be seen that the maximum increase in transverse momentum depending on the p_T at matrix element level increases first with p_T up to $p_{T, \text{ME level}} \approx 100 \text{ GeV}$ where $(\Delta p_T)_{\text{max}} \approx 70 \text{ GeV}$ and then, with some exceptions decreases again to $(\Delta p_T)_{\text{max}} \approx 30 \text{ GeV}$. For the hardest jet, the maximum increase stayed at around 70 GeV for all higher values of $p_{T, \text{ME level}}$. A possible explanation might be that the second hardest jet is not as stable as the hardest jet and is more influenced by the shower, hadronization and underlying event.

The simulation done with the MPI underlying event model shows a similar behavior for the second hardest jet, Fig. 7.11, for the uncorrelated events some events seem to gain a large amount of transverse momentum but these events are rejected by the

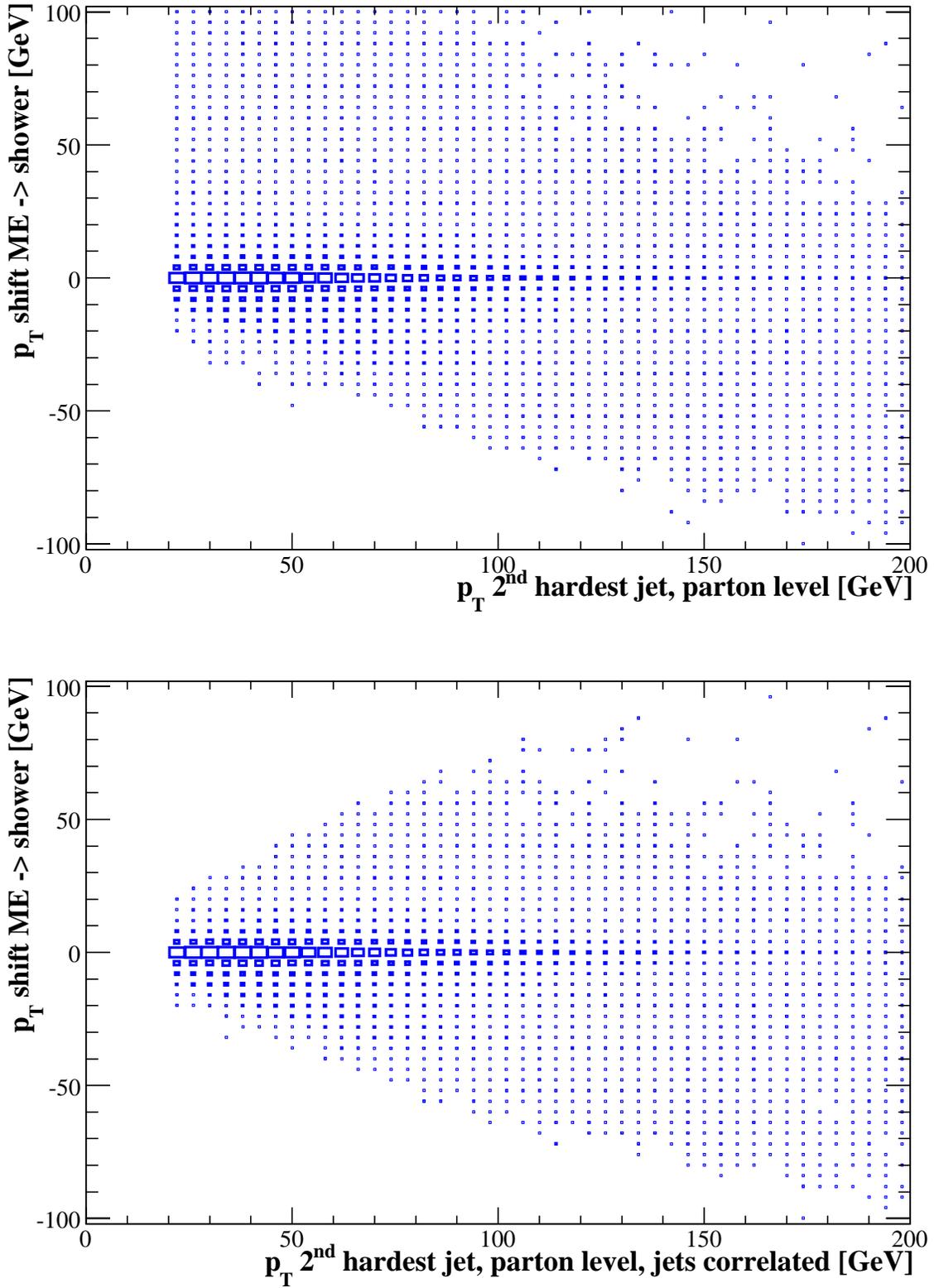


Figure 7.8: p_T shifts of 2nd hardest jet, ME to shower, relative to p_T at ME level, all (above) and and correlated jets (below). UA5 parametrization.

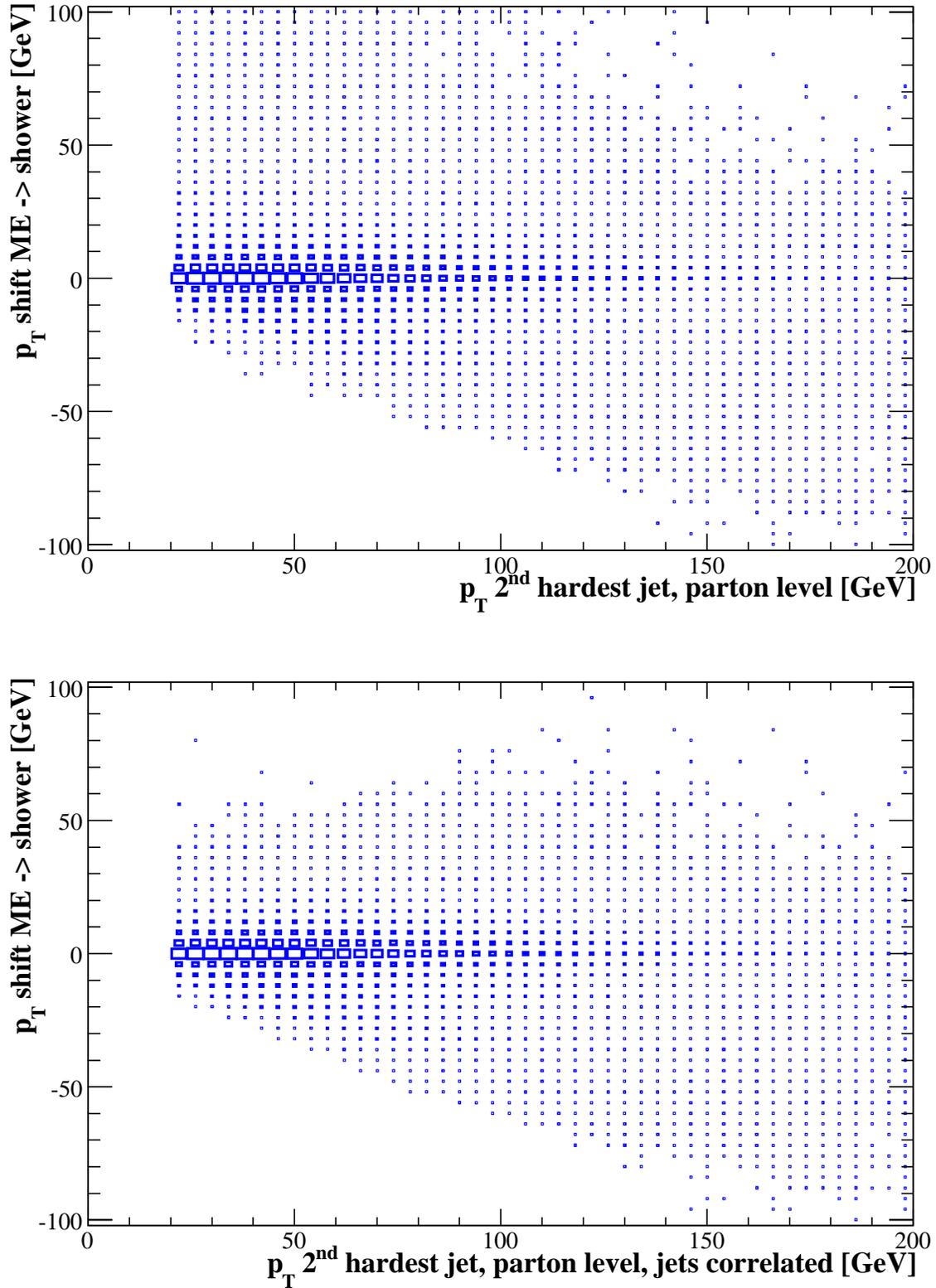


Figure 7.9: p_T shifts of 2nd hardest jet, ME to shower, relative to p_T at ME level, all (above) and and correlated jets (below). MPI model.

correlation condition. The 2nd hardest jets in this sample do not gain much transverse momentum, the maximum increase found actually decreases with $p_{T,ME}$ level, from $(\Delta p_T)_{\max} \approx 40$ GeV it first increases to $(\Delta p_T)_{\max} \approx 70$ GeV for $p_{T,ME}$ level ≈ 100 GeV to $(\Delta p_T)_{\max} \approx 30$ GeV at $p_{T,ME}$ level = 200 GeV.

The comparison of the number of correlated events and all events in Tab. 7.4 shows that for a large fraction of events the direction of the 2nd hardest jet is not stable, the overall number of events decreases by about 13% in the UA5 sample, for the MPI sample by 10%. At the same time the fraction of events with no or only small changes in transverse momentum increases with the correlation, about 4 and 2 percentage points for the 'zero shift bin' resp. the shifts with $|\Delta p_T| < 10$ GeV for the UA5 sample and 5.5 resp. 3 percentage points for the MPI sample.

A similar picture holds for the intervals considered separately, the proportion between all events and the correlated events is more or less the same for all intervals.

	UA5 parametrization		MPI model	
	All events	Correlated	All events	Correlated
Complete:				
Total	$3.857 \cdot 10^6$	$3.350 \cdot 10^6$	$3.857 \cdot 10^6$	$3.479 \cdot 10^6$
No shift	0.397	0.439	0.365	0.386
$ p_T \text{ shift} < 10$ GeV	0.810	0.865	0.861	0.891
$p_{T,\text{hard}} < 40$ GeV				
Total	$1.052 \cdot 10^6$	$0.884 \cdot 10^6$	$1.053 \cdot 10^6$	$0.933 \cdot 10^6$
No shift	0.431	0.492	0.406	0.438
$ p_T \text{ shift} < 10$ GeV	0.836	0.905	0.901	0.940
$40 \text{ GeV} < p_{T,\text{hard}} < 120$ GeV				
Total	$2.694 \cdot 10^6$	$2.367 \cdot 10^6$	$2.693 \cdot 10^6$	$2.446 \cdot 10^6$
No shift	0.387	0.422	0.352	0.370
$ p_T \text{ shift} < 10$ GeV	0.802	0.853	0.849	0.875
$120 \text{ GeV} < p_{T,\text{hard}} < 200$ GeV				
Total	$0.111 \cdot 10^6$	$0.099 \cdot 10^6$	$0.111 \cdot 10^6$	$0.100 \cdot 10^6$
No shift	0.332	0.354	0.288	0.302
$ p_T \text{ shift} < 10$ GeV	0.757	0.795	0.775	0.802

Table 7.4: p_T shift of the 2nd hardest tagging jet from matrix element level to final state. UA5 parametrization and MPI UE model. Number of all events in considered region, fraction of events with shift in 'zero bin' ($|\Delta p_T| \leq 2\text{GeV}$) and fraction of events with shift ($|\Delta p_T| \leq 10\text{GeV}$).

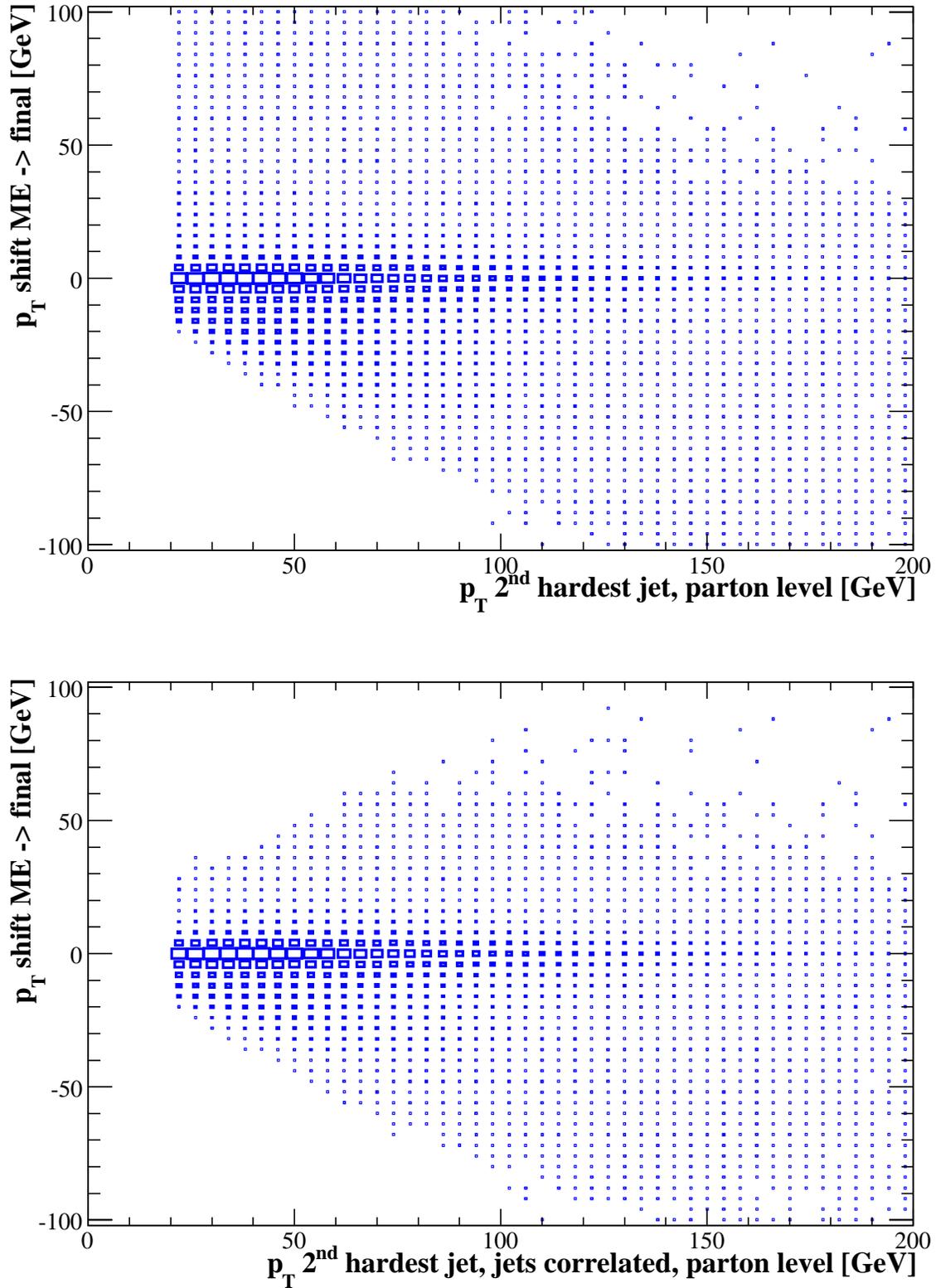


Figure 7.10: p_T shifts ME to final state, relative to p_T at ME level, UA5 parametrization, all (above) and correlated jets (below).

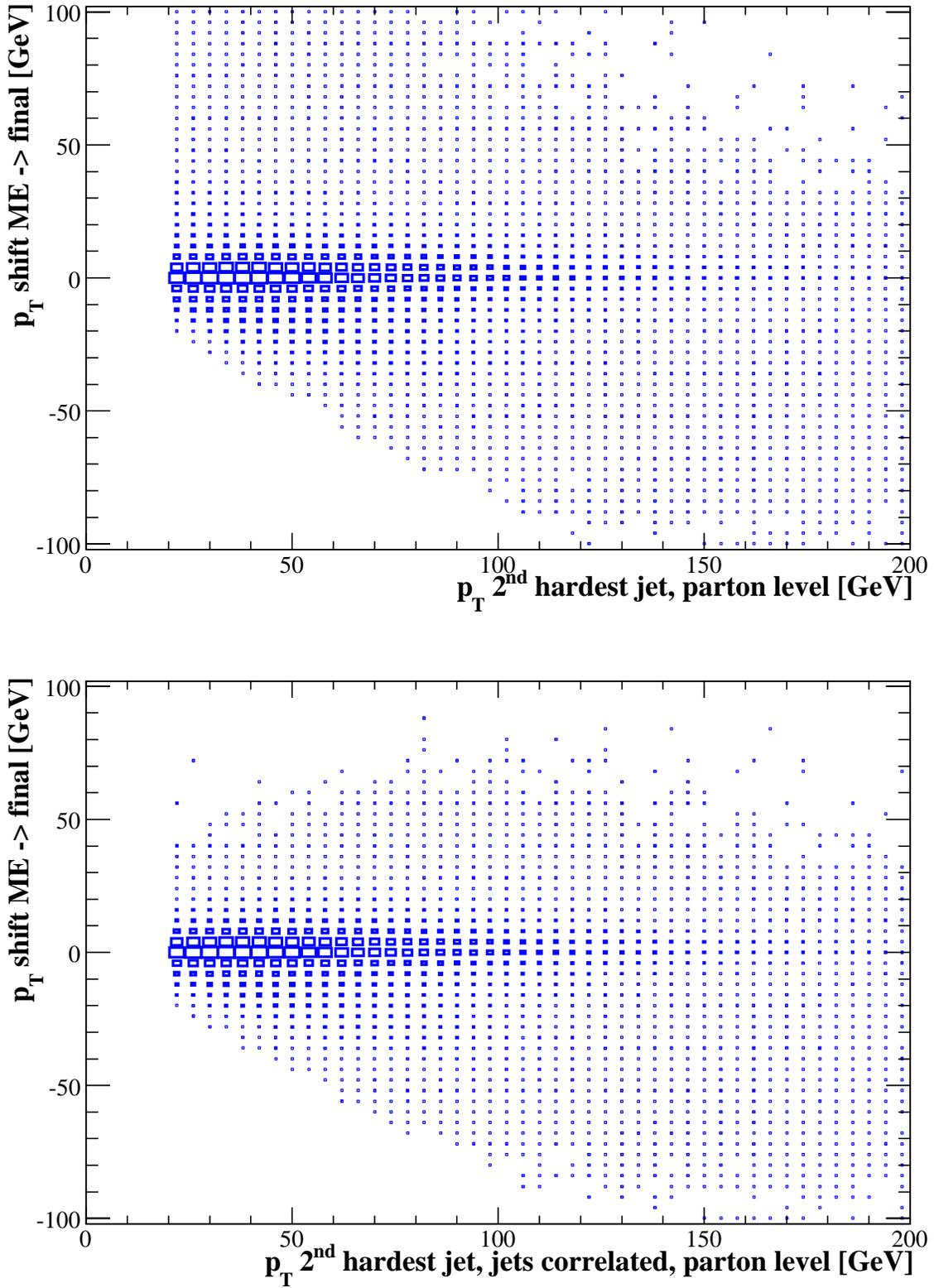


Figure 7.11: p_T shifts ME to final state, relative to p_T at ME level, MPI model, all jets (above) and correlated jets (below).

7.3 p_T distributions for correlated events

To see the impact of the correlation between the tagging jets at different levels on the VBF observables, a comparison was made for the transverse momentum distributions at the final state. This again was done for both of the used underlying event models. The overall shape of the distributions for the two tagging jets were not changed much, the correlated distribution seemed just to be scaled down. The overall behavior and especially the position of the peaks was unchanged, Fig. 7.12, 7.13 for the UA5 parametrization, Fig. 7.15, 7.16 for the MPI underlying event model.

The distribution for the transverse momentum difference between the tagging jets however was changed for the low Δp_T region, Fig. 7.14 respectively Fig. 7.17. Here the distribution for the correlated events has a dip for $\Delta p_T = 0$ GeV and then increases to peak around 5 GeV whereas the distribution for all events peaks at $\Delta p_T = 0$ GeV and decreases monotonously. Both curves run parallel for $\Delta p_T \geq 20$ GeV, so the effect occurs only for events with a small p_T difference between the tagging jets.

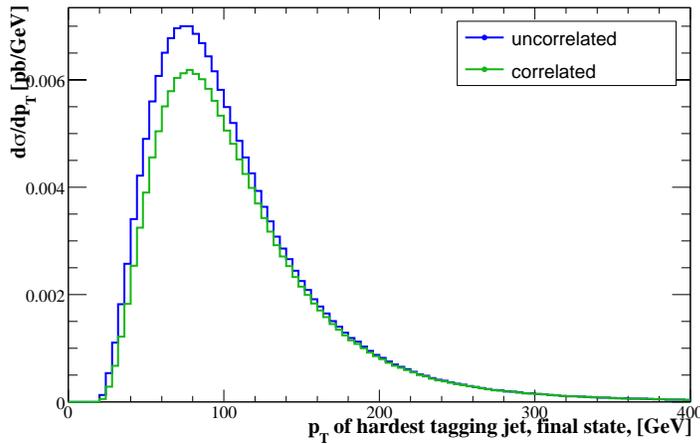


Figure 7.12: p_T distribution of the hardest tagging jet, final state, UA5 parametrization, all and correlated events.

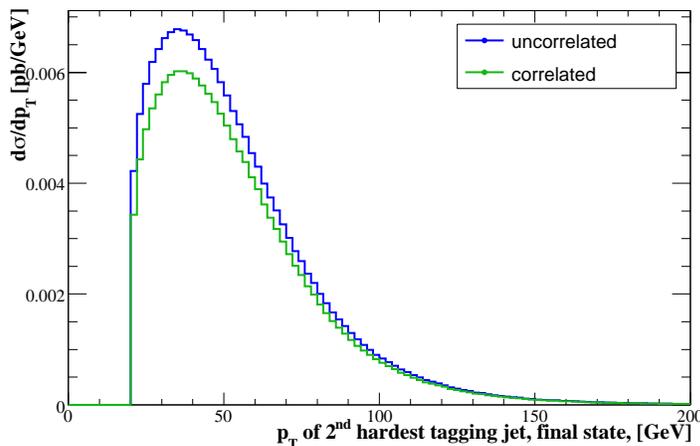


Figure 7.13: p_T distribution of the 2nd hardest tagging jet, final state, UA5 parametrization, all and correlated events.

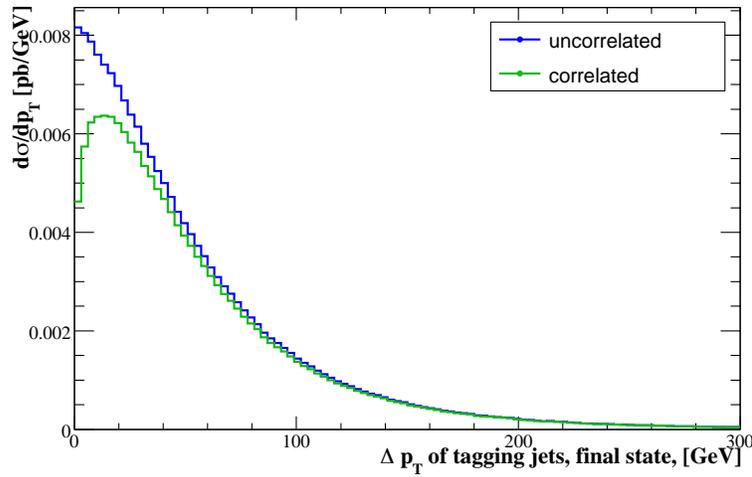


Figure 7.14: p_T difference between the two tagging jet, final state, UA5 parametrization, all and correlated events.

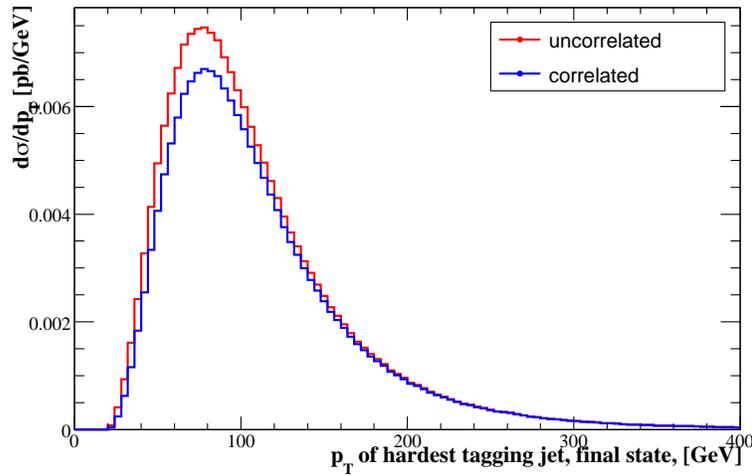


Figure 7.15: p_T distribution of the hardest tagging jet, final state, MPI model, all and correlated events.

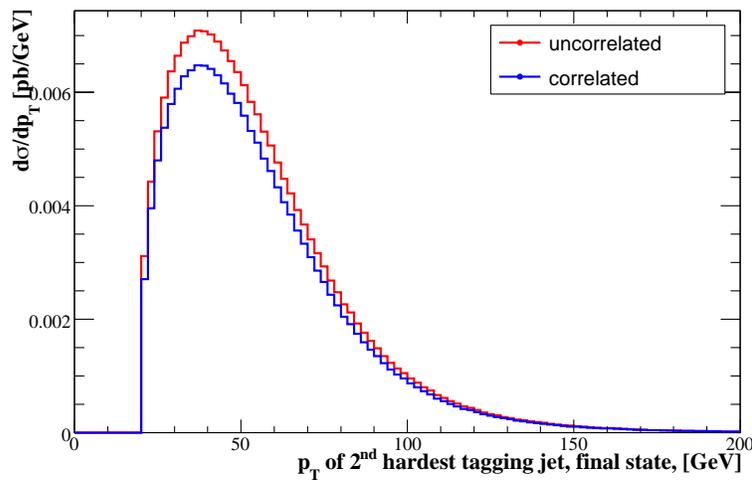


Figure 7.16: p_T distribution of the 2nd hardest tagging jet, final state, MPI model, all and correlated events.

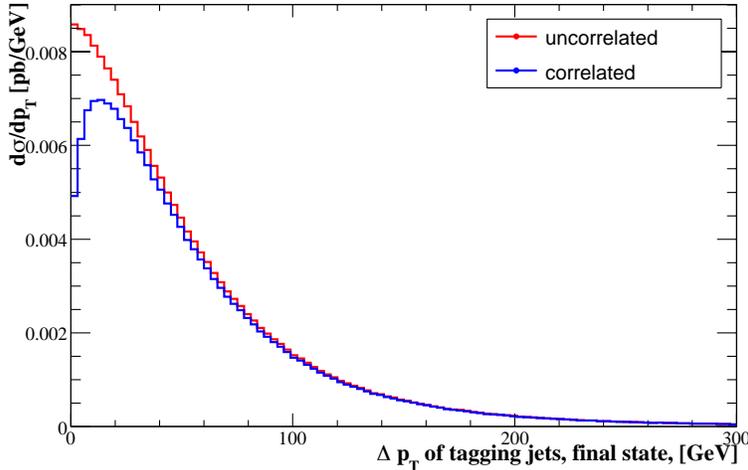


Figure 7.17: p_T difference between the two tagging jet, final state, MPI model, all and correlated events.

To see the difference more clearly the ratio of the distributions at matrix element level and at shower level was examined. Here a large difference between the two models can be seen.

For the UA5 parametrization, Fig. 7.18, 1st histogram, the hardest jet can be seen to differ only in the low p_T region, here the uncorrelated sample shows a peak at the lowest bin above the $p_{T,\min}$ cut, then falls slightly below one and increases slowly with the increase of transverse momentum to a value slightly above one. The correlated sample starts much lower, around 0.6 and then increases slowly towards the uncorrelated sample which it reaches at $p_T \approx 60$ GeV.

This is quite different to the case of the MPI model, Fig. 7.19, 1st histogram. Here the correlated and uncorrelated distributions are very similar. Both distributions have a sharp peak at the minimum $p_T = 20$ GeV, the uncorrelated starting at 1.6, the correlated at 1.2. Both distributions fall fast after the peak, then meet at $p_T \approx 30$ GeV. From here both distributions decrease slowly until they reach 1 at $p_T \approx 70$ GeV, from where they remain unchanged.

This behavior can be explained by regarding the p_T distributions in comparison, as was done in chapter 4, figure 4.3, it can be seen that the position of the peak shifts a little bit to a higher value between the matrix element level and the final state for the MPI model. So it is not surprising that the ratio for this model goes up for low transverse momentum and down for larger values.

For the second tagging jet the behavior of the ratio is similar to that from the hardest tagging jet. For the UA5 parametrization, Fig. 7.18, 2nd histogram, the uncorrelated sample shows a peak at $p_{T,\min} = 20$ GeV and a small dip behind it and then reaches the value 1, at which it stays. The correlated sample shows less activity for low p_T and grows towards the uncorrelated sample around $p_T = 35$ GeV, but then exceeds the uncorrelated sample and reaches a value of 1.075, where it stays constant. This hints at a slight shift in the peak position, but from the small deviations around one it is clear that this is a small effect.

The situation is a little different for ratios in the MPI model, Fig. 7.19, 2nd

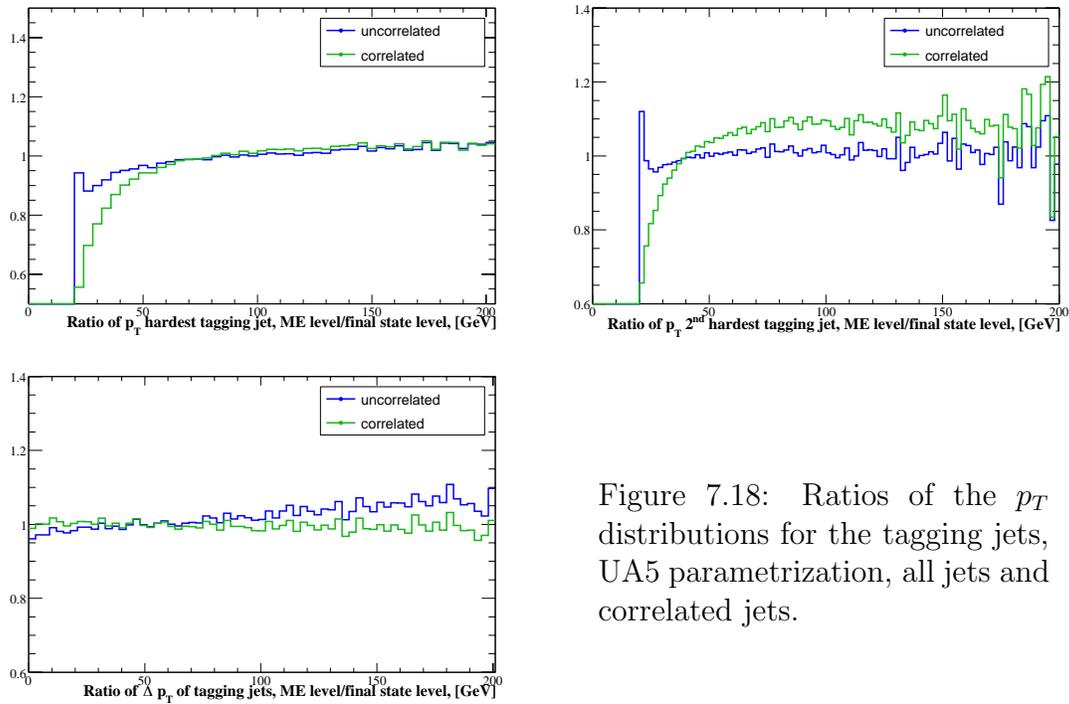


Figure 7.18: Ratios of the p_T distributions for the tagging jets, UA5 parametrization, all jets and correlated jets.

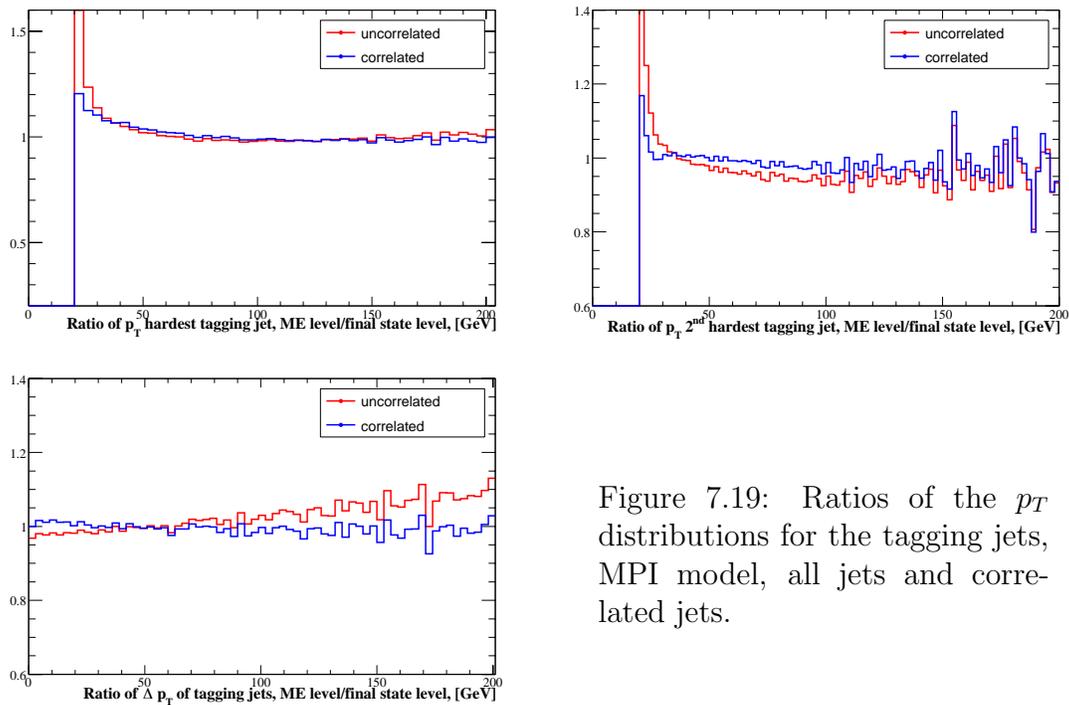


Figure 7.19: Ratios of the p_T distributions for the tagging jets, MPI model, all jets and correlated jets.

histogram. Here, as for the hardest tagging jet, a peak at the minimum p_T can be seen, from which both distributions drop rapidly. But instead of reaching a constant ratio at 1, both distributions exhibit a permanent decrease, the uncorrelated sample being below the correlated. This can be explained by the bigger difference between

the correlated sample and the overall sample for the second hardest jet that can as well be seen in the slight shift in the peak position of the p_T distribution, Fig. 7.16.

For the p_T difference Δp_T the ratios for the correlated samples in both underlying event models considered are very exactly at one, showing that for the correlated events the transverse momentum difference stays the same at matrix element level and the final state. Fig. 7.18, 7.19, 3rd histogram, respectively.

The ratios for the uncorrelated samples both start below one, at 0.95 for the UA5 parametrization and 0.97 for the MPI model. From there they both increase linearly until they reach one at $\Delta p_T = 50$ GeV. Both ratios keep rising slightly up to 1.07 and 1.1 at $\Delta p_T = 200$ GeV. So the uncorrelated distributions show that for the set of all events the p_T difference grows a little bit during the event simulation, which is to be expected since the jets radiate partons or acquire additional particles from the underlying event independent of each other.

Chapter 8

Conclusion

In this work the influence of full event simulations on the properties of vector boson fusion observables was examined. This was done in order to gain insight into the possible differences between the results from matrix element calculations and the final states in experiments. Since predictions for hadronic particles cannot be made from exact calculations, approximations and phenomenological models for the various effects influencing the final state of the events of interest have to be made.

A review of the properties of the signal of interest and the models and techniques used in event generators was made in chapter 2. Then the course undertaken in the analysis was described along with the programs used in chapter 3.

The following four chapters showed the results of the investigations.

First, the influence of the full event simulation on the VBF signal and the cut efficiency was examined, chapter 4.

The vector boson fusion observables turned out to be very stable under the various steps in the simulation. The cut efficiency became lower, which means that some of the events passing the vector boson fusion cuts at matrix element level got distorted, for example by emission of hard partons, so that the two hardest jets did not correspond to the two jets from the VBF process any more. The cuts imposed on the events however turned out to be quite sensible, rejecting these events and thus conserving the signal.

The cut efficiency for the final state went down by about eleven percentage points compared to the cut efficiency at the hard process in case of the UA5 parametrization, and approximately six percentage points lower for the MPI model.

The major impact on the VBF signal turned out to stem from the underlying event model. Here, the only significant changes could be found. The analysis was made using two underlying event models, a model based on the UA5 parametrization of experimental data and a multiparton interaction model, MPI, which includes additional interactions between constituents of the remnant. For the MPI model the peaks in the most of the VBF signal distributions were shifted, but even these effects were small.

The different possible methods to define the tagging jets, as either the two jets with hardest transverse momentum or hardest transverse energy do not give different

results. The small impact of the different orderings can be understood from the low jet masses, between ten and twenty GeV. This is small compared to the transverse momentum of the tagging jets, which peaks around 90 GeV for the hardest and 40 GeV for the second hardest tagging jet, so the difference between p_T and E_T is negligible.

In addition the jet activity created by the full event simulation was examined. This was done in two parts. First, in chapter 5 the properties of the third hardest jet were investigated.

The third hardest jet, which is the hardest jet that does not stem from the process at matrix element level was studied in detail. A comparison with the distributions for the third jet after the shower and in the final states for the two underlying event models used was made. In addition the distributions for the third jet in a next-to-leading order calculation for the same process were examined to see the difference between the results from the shower approximation and the full calculation. In conclusion it can be found that the overall distribution of the third hardest jets is close to the tagging jets, the UA5 parametrization hardly has an impact on the distribution obtained from the final state of the shower step. The MPI model, however, showed to be different, it could be seen that in this model a quite large amount of activity is present that is uncorrelated with the hard process. The comparison between the third hardest jet after the shower and the NLO calculation shows that the shower tends to place the third hardest jet outside the rapidity gap formed by the two tagging jets, whereas the third hardest jet in the NLO matrix element calculation is distributed more evenly between the outside and inside region.

After this, the influence of the additional jets on the VBF signal, in particular on the rapidity gap was studied in chapter 6. To preserve the signal, a central jet veto is introduced, rejecting events with hard jet activity between the tagging jets.

To investigate the activity between the two tagging jets in more detail and to see whether the rapidity gap found at matrix element level can also be seen at the final state, the central jets are analyzed. This was done finding the hardest central jet, i.e. the hardest jet lying in the rapidity gap. Since the behavior of the central jets was expected to be dominated by the underlying event, this was compared for the two different models used and a sample without a sophisticated underlying event model.

The two different underlying event models compared did not show a large deviation regarding the rapidity of the central jets, even the y^* distributions showing the rapidity of the central jet relative to the mean value of the rapidity of the tagging jets was quite similar. The third jet in the sample without underlying event showed a behavior similar to the third jet stemming from the shower.

The transverse momentum of the central jets showed clearly the different regions of activity for the possible underlying event models. Both models showed an excess to the distribution without underlying event in the low p_T region, the MPI model in a higher and wider part than the UA5 parametrization. Above $p_{T,\text{central}} = 25$ GeV the three samples run almost identically, reaching the result for the matrix element calculation.

From this the efficiency for a central jet veto was deduced, that is the fraction of vector boson fusion events passing a veto. Such a *central jet veto* rejects events

where a central jet with transverse momentum above $p_{T,\min}$ is present. The veto efficiency depending on $p_{T,\min}$ was calculated. It turned out to be quite good for a reasonable choice of $p_{T,\min}$, for low values the discrepancies between the models were rather large, from 90% in case of the case without model down to 65% for the MPI model at $p_{T,\min} = 10$ GeV. For $p_{T,\min} \approx 20$ GeV however, the veto efficiency is between 95% and 97%, and since the transverse momentum distributions for $p_T \geq 25$ GeV are identical, the veto efficiency for all three models is the same here as well.

To see the influence of the simulation on the tagging jets in more detail, a direct comparison between the tagging jets at the different steps in the simulation was made in chapter 7.

In order to study this, the changes in transverse momentum between the tagging jets at matrix element level and after the shower respectively the full event simulation were examined and the stability of the direction of the jets tested.

The result to this was that the hardest tagging jet mostly is not affected much by the simulation, especially its direction stays stable and only a small fraction suffers a big change in transverse momentum. This holds for the change to shower level as well as to the final state.

For the second tagging jet this was a little bit different. While the overall change in transverse momentum was not big either, a significantly higher amount of jets seemed to undergo a larger change in direction. This could best be explained by the possibility that the second hardest jet after the simulation was not the original second hardest jet any more. This can happen if during the simulation a hard jet arises which then is taken to be a potential tagging jet. This mixing of jets can be seen by the jets seemingly acquiring a huge amount of transverse momentum during the simulation, most of which are rejected by the correlation condition.

A comparison of distributions for correlated and all events showed hardly a different behavior in the observables, only slight changes could be perceived.

The results in this work show that the Vector Boson Fusion signal is not affected much by a full event simulation. The VBF cuts devised for this process turned out to be very sensible and reject the events that would distort the signal. The cut efficiency showed that only for a comparably small number of events the effects from the full simulation are big enough to change the signal significantly.

The additional hadronic activity not stemming from the original VBF process turned out to be no threat to the signal. The rapidity gap as distinguishing feature of the signal was fully visible after the complete simulation, activity falling between the tagging jets mostly turned out to be of low transverse momentum. The veto efficiency for a central jet veto depends on the underlying event model used only for small values of $p_{T,\min}$. For values greater than 25 GeV it gets dominated by jets stemming from the parton shower, which are close to the tagging jets, so the veto efficiency turned out to be very good for reasonable values of $p_{T,\min} \approx 20$ GeV.

In this analysis detector effects have not been taken into account, so no loss or smearing for the signal was considered. This of course is an important field for further studies, since the impact of these effects cannot be estimated a priori.

In conclusion vector boson fusion shows to be a very promising Higgs discovery process indeed, since the distinguishing features are quite stable under the effects of parton shower, hadronization and underlying event. The VBF cuts and the central jet veto offer a very good method to extract this signal.

Appendix A

Sample Les Houches File

```
<LesHouchesEvents version="1.0">
<!--
File generated with VBFNLO - parton level MC program (LO)
pp -> H jj -> mu+ mu- jj

Higgs + 2 jets production in vector boson fusion with Higgs

decay into mu+ mu-.
Process is implemented at LO and NLO QCD.
Anomalous coupling parameters can be set in "anom_HVV.dat".
## Number of Events : 22702397
-->
<init>
  2212  2212  7.000000E+03  7.000000E+03  0  0  10042  10042  2  1
  2.035589E-04  3.448832E-08  1.741728E-07  102
</init>
<event>
  7  102  1.058494E-08  6.360184E+01  7.554144E-03  1.376638E-01
    3  -1  0  0  501  0  0.0000000000E+00  0.0000000000E+00  7.7127008479E+02
    7.7127008479E+02  0.0000000000E+00  0.0E+00  9.0E+00
   -1  -1  0  0  0  502  0.0000000000E+00  0.0000000000E+00  -4.7510403211E+02
    4.7510403211E+02  0.0000000000E+00  0.0E+00  9.0E+00
    4  1  1  2  501  0  4.0431488591E+01  2.4829448551E+00  6.6093340603E+02
    6.6217357052E+02  0.0000000000E+00  0.0E+00  9.0E+00
   -2  1  1  2  0  502  -5.1764506989E+01  -3.2480996279E+01  -4.3649404954E+02
    4.4075121621E+02  0.0000000000E+00  0.0E+00  9.0E+00
   25  2  1  2  0  0  1.1333018398E+01  2.9998051424E+01  7.1726696188E+01
    1.4344933017E+02  1.2001946086E+02  0.0E+00  9.0E+00
  -13  1  5  0  0  0  -3.3939222322E+01  7.0199312873E+00  -1.8970000138E+01
    3.9509633665E+01  0.0000000000E+00  0.0E+00  -1.0E+00
   13  1  5  0  0  0  4.5272240721E+01  2.2978120137E+01  9.0696696326E+01
    1.0393969650E+02  0.0000000000E+00  0.0E+00  1.0E+00
</event>
```

```

<event>
7 102 1.041277E-08 1.325916E+02 7.554144E-03 1.225421E-01
  2 -1 0 0 501 0 0.0000000000E+00 0.0000000000E+00 2.3633886684E+02
    2.3633886684E+02 0.0000000000E+00 0.0E+00 9.0E+00
  3 -1 0 0 502 0 0.0000000000E+00 0.0000000000E+00 -1.2194214820E+03
    1.2194214820E+03 0.0000000000E+00 0.0E+00 9.0E+00
  1 1 1 2 501 0 7.9959166375E+00 5.8182704321E+01 2.1657098735E+02
    2.2439285712E+02 0.0000000000E+00 0.0E+00 9.0E+00
  4 1 1 2 502 0 4.4074545182E+01 -1.0014716842E+02 -8.2679905569E+02
    8.3400761349E+02 0.0000000000E+00 0.0E+00 9.0E+00
25 2 1 2 0 0 -5.2070461820E+01 4.1964464095E+01 -3.7285454682E+02
    3.9735987823E+02 1.2000837677E+02 0.0E+00 9.0E+00
-13 1 5 0 0 0 -8.5991775320E+01 5.2115226233E+00 -2.0953690576E+02
    2.2655564496E+02 0.0000000000E+00 0.0E+00 1.0E+00
 13 1 5 0 0 0 3.3921313501E+01 3.6752941472E+01 -1.6331764107E+02
    1.7080423326E+02 0.0000000000E+00 0.0E+00 -1.0E+00
</event>
:
</LesHouchesEvents>

```

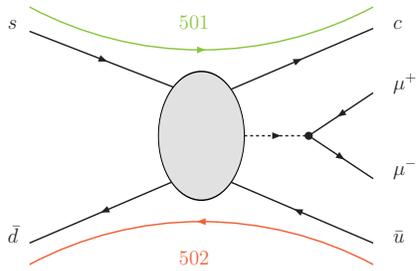


Figure A.1: Schematic description of first event. The blob marks the VBF process producing the Higgs.

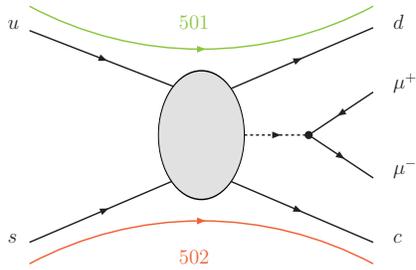


Figure A.2: Schematic description of second event. The blob marks the VBF process producing the Higgs.

Bibliography

- [1] M. E. Peskin and D. V. Schroeder, “An Introduction To Quantum Field Theory,” *Reading, USA: Addison-Wesley (1995) 842 p*
- [2] L. H. Ryder, “Quantum Field Theory,” *Cambridge, Uk: Univ. Pr. (1985) 443p*
- [3] D. Graudenz, M. Spira and P. M. Zerwas, “QCD corrections to Higgs boson production at proton proton colliders,” *Phys. Rev. Lett.* **70** (1993) 1372.
M. Spira, A. Djouadi, D. Graudenz and P. M. Zerwas, “Higgs boson production at the LHC,” *Nucl. Phys. B* **453** (1995) 17 [arXiv:hep-ph/9504378].
- [4] T. Han, G. Valencia and S. Willenbrock, “Structure function approach to vector boson scattering in p p collisions,” *Phys. Rev. Lett.* **69** (1992) 3274 [arXiv:hep-ph/9206246].
- [5] S. L. Glashow, D. V. Nanopoulos and A. Yildiz, “Associated Production Of Higgs Bosons And Z Particles,” *Phys. Rev. D* **18** (1978) 1724.
- [6] S. Dawson, C. Jackson, L. H. Orr, L. Reina and D. Wackerth, “Associated Higgs production with top quarks at the Large Hadron Collider: NLO QCD corrections,” *Phys. Rev. D* **68** (2003) 034022 [arXiv:hep-ph/0305087].
P. Hafliger and M. Spira, “Associated Higgs boson production with heavy quarks in e+ e- collisions: SUSY-QCD corrections,” *Nucl. Phys. B* **719** (2005) 35 [arXiv:hep-ph/0501164].
- [7] T. Figy, C. Oleari and D. Zeppenfeld, “Next-to-leading order jet distributions for Higgs boson production via weak-boson fusion,” *Phys. Rev. D* **68** (2003) 073005 [arXiv:hep-ph/0306109].
C. Oleari and D. Zeppenfeld, “Next-to-leading order QCD corrections to W and Z production via vector-boson fusion,” *Phys. Rev. D* **69** (2004) 093004 [arXiv:hep-ph/0310156].
B. Jager, C. Oleari and D. Zeppenfeld, “Next-to-leading order QCD corrections to W+ W- production via vector-boson fusion,” *JHEP* **0607** (2006) 015 [arXiv:hep-ph/0603177].
- [8] H. Chehime and D. Zeppenfeld, “‘Rapidity gaps’ A New tool to find the Higgs?,” C93-02-21;
A. Duff and D. Zeppenfeld, “Heavy Higgs boson production in association with

- three jets at hadron supercolliders,” *Phys. Rev. D* **50** (1994) 3204 [arXiv:hep-ph/9312357].
- [9] D. Green, “Vector boson fusion Higgs production at the LHC: Mass variables,” [arXiv:hep-ex/0501027].
D. Green, “Kinematics in vector boson fusion,” [arXiv:hep-ph/0603022].
- [10] R. K. Ellis, W. J. Stirling and B. R. Webber, “QCD and collider physics,” *Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.* **8** (1996) 1.
- [11] G. Dissertori, I. G. Knowles and M. Schmelling, “High energy experiments and theory,” *Oxford, UK: Clarendon (2003) 538 p*
- [12] V. N. Gribov and L. N. Lipatov, “Deep Inelastic E P Scattering In Perturbation Theory,” *Sov. J. Nucl. Phys.* **15** (1972) 438 [*Yad. Fiz.* **15** (1972) 781].
G. Altarelli and G. Parisi, “Asymptotic Freedom In Parton Language,” *Nucl. Phys. B* **126** (1977) 298.
Y. L. Dokshitzer, *Sov. Phys. JETP* **46** (1977) 641 [*Zh. Eksp. Teor. Fiz.* **73** (1977) 1216].
- [13] R. Odorico, “Exclusive Calculations For QCD Jets In A Monte Carlo Approach,” *Nucl. Phys. B* **172** (1980) 157.
A. Bassetto, M. Ciafaloni and G. Marchesini, “Jet Structure And Infrared Sensitive Quantities In Perturbative QCD,” *Phys. Rept.* **100** (1983) 201.
Y. L. Dokshitzer, G. Marchesini and B. R. Webber, “Dispersive Approach to Power-Behaved Contributions in QCD Hard Processes,” *Nucl. Phys. B* **469** (1996) 93 [arXiv:hep-ph/9512336].
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. B* **632** (2002) 343 [arXiv:hep-ph/0108069].
- [14] G. Marchesini and B. R. Webber, “Simulation Of QCD Jets Including Soft Gluon Interference,” *Nucl. Phys. B* **238** (1984) 1.
G. Marchesini and B. R. Webber, “Monte Carlo Simulation of General Hard Processes with Coherent QCD Radiation,” *Nucl. Phys. B* **310** (1988) 461.
S. Catani, B. R. Webber and G. Marchesini, “QCD coherent branching and semiinclusive processes at large x,” *Nucl. Phys. B* **349** (1991) 635.
- [15] S. Catani, S. Dittmaier and Z. Trocsanyi, “One-loop singular behaviour of QCD and SUSY QCD amplitudes with massive partons,” *Phys. Lett. B* **500** (2001) 149 [arXiv:hep-ph/0011222].
- [16] M. Bengtsson and P. M. Zerwas, “Four Jet Events in e+ e- Annihilation: Testing the Three Gluon Vertex,” *Phys. Lett. B* **208** (1988) 306.
- [17] J. Frenkel and J. C. Taylor, “Exponentiation Of Leading Infrared Divergences In Massless Yang-Mills Theories,” *Nucl. Phys. B* **116** (1976) 185.

- [18] R. D. Field and R. P. Feynman, “Quark Elastic Scattering As A Source Of High Transverse Momentum Mesons,” *Phys. Rev. D* **15** (1977) 2590.
R. D. Field and R. P. Feynman, “A parametrization of the properties of quark jets,” *Nucl. Phys. B* **136** (1978) 1.
R. P. Feynman, R. D. Field and G. C. Fox, “Quantum-chromodynamic approach for the large-transverse-momentum production of particles and jets,” *Phys. Rev. D* **18** (1978) 3320.
- [19] T. Sjostrand and M. Bengtsson, “The Lund Monte Carlo For Jet Fragmentation And E+ E- Physics. Jetset Version 6.3: An Update,” *Comput. Phys. Commun.* **43** (1987) 367.
- [20] X. Artru, “Classical String Phenomenology. 1. How Strings Work,” *Phys. Rept.* **97** (1983) 147.
- [21] T. Sjostrand,
“Phenomenological Studies On Jet Fragmentation,” RX-981-LUND;B. Andersson, G. Gustafson and T. Sjostrand, “Baryon Production In Jet Fragmentation And Upsilon Decay,” *Phys. Scripta* **32** (1985) 574.
E. Norrbin and T. Sjostrand, “Production and hadronization of heavy quarks,” *Eur. Phys. J. C* **17** (2000) 137 [arXiv:hep-ph/0005110].
- [22] R. D. Field and S. Wolfram, “A QCD Model For E+ E- Annihilation,” *Nucl. Phys. B* **213** (1983) 65.
S. Wolfram, “Jet Development In Leading Log QCD,” CALT-68-740;
T. D. Gottschalk, “A Simple Phenomenological Model For Hadron Production From Low Mass Clusters,” *Nucl. Phys. B* **239** (1984) 325.
T. D. Gottschalk, “An Improved Description Of Hadronization In The QCD Cluster Model For E+ E- Annihilation,” *Nucl. Phys. B* **239** (1984) 349.
- [23] B. R. Webber, “A QCD Model For Jet Fragmentation Including Soft Gluon Interference,” *Nucl. Phys. B* **238** (1984) 492.
- [24] G. J. Alner *et al.* [UA5 Collaboration], “UA5: A general study of proton-antiproton physics at $s^{*1/2} = 546\text{-GeV}$,” *Phys. Rept.* **154** (1987) 247.
G. J. Alner *et al.* [UA5 Collaboration], “The UA5 High-Energy anti-p p Simulation Program,” *Nucl. Phys. B* **291** (1987) 445.
- [25] J. M. Butterworth, J. R. Forshaw and M. H. Seymour, “Multiparton interactions in photoproduction at HERA,” *Z. Phys. C* **72** (1996) 637 [arXiv:hep-ph/9601371].
- [26] R. Kleiss and W. J. Stirling, “Massive multiplicities and Monte Carlo,” *Nucl. Phys. B* **385** (1992) 413.
- [27] E. P. Venugopal and B. R. Holstein,
Phys. Rev. D **57** (1998) 4397 [arXiv:hep-ph/9710382].
Phys. Scripta **T99** (2002) 55 [arXiv:hep-ph/0112150].

- [28] G. C. Blazey *et al.*, “Run II jet physics,” [arXiv:hep-ex/0005012].
- [29] G. Sterman and S. Weinberg, “Jets From Quantum Chromodynamics,” *Phys. Rev. Lett.* **39** (1977) 1436.
- [30] S. Catani, Y. L. Dokshitzer, M. Olsson, G. Turnock and B. R. Webber, “New clustering algorithm for multi - jet cross-sections in e+ e- *Phys. Lett. B* **269** (1991) 432.
S. D. Ellis and D. E. Soper, “Successive combination jet algorithm for hadron collisions,” *Phys. Rev. D* **48** (1993) 3160 [arXiv:hep-ph/9305266].
S. Catani, Y. L. Dokshitzer, M. H. Seymour and B. R. Webber, “Longitudinally Invariant K(T) Clustering Algorithms For Hadron Hadron Collisions,” *Nucl. Phys. B* **406** (1993) 187.
- [31] J. M. Butterworth, J. P. Couchman, B. E. Cox and B. M. Waugh, “KtJet: A C++ implementation of the K(T) clustering algorithm,” *Comput. Phys. Commun.* **153** (2003) 85 [arXiv:hep-ph/0210022].
- [32] Y. L. Dokshitzer, G. D. Leder, S. Moretti and B. R. Webber, “Better jet clustering algorithms,” *JHEP* **9708** (1997) 001 [arXiv:hep-ph/9707323].
M. Wobisch and T. Wengler, “Hadronization corrections to jet cross sections in deep-inelastic scattering,” [arXiv:hep-ph/9907280].
- [33] M. Cacciari and G. P. Salam, *Phys. Lett. B* **641** (2006) 57 [arXiv:hep-ph/0512210].
M. Cacciari, “FastJet: Dispelling the N**3 myth for the k(t) jet-finder,” [arXiv:hep-ph/0607071].
- [34] G. P. Salam and G. Soyez, “A practical Seedless Infrared-Safe Cone jet algorithm,” *JHEP* **0705** (2007) 086 [arXiv:0704.0292 [hep-ph]].
- [35] H. Murayama, I. Watanabe and K. Hagiwara, “HELAS: HELicity amplitude subroutines for Feynman diagram evaluations,” KEK-91-11;
- [36] F. Maltoni and T. Stelzer, “MadEvent: Automatic event generation with MadGraph,” *JHEP* **0302** (2003) 027 [arXiv:hep-ph/0208156].
- [37] G. P. Lepage, “Vegas: An Adaptive Multidimensional Integration Program,” CLNS-80/447;
- [38] S. Kretzer, H. L. Lai, F. I. Olness and W. K. Tung, “CTEQ6 parton distributions with heavy quark mass effects,” *Phys. Rev. D* **69** (2004) 114005 [arXiv:hep-ph/0307022].
- [39] M. R. Whalley, D. Bourilkov and R. C. Group, “The Les Houches accord PDFs (LHAPDF) and LHAGLUE,” [arXiv:hep-ph/0508110].
<http://hepforge.cedar.ac.uk/lhapdf/>

- [40] E. Boos *et al.*, “Generic user process interface for event generators,” [arXiv:hep-ph/0109068].
- [41] J. Alwall *et al.*, “A standard format for Les Houches event files,” *Comput. Phys. Commun.* **176** (2007) 300 [arXiv:hep-ph/0609017].
- [42] <http://www.w3.org/XML/>
<http://www.w3.org/XML/1999/XML-in-10-points.html.en/>
<http://www.xml.org/>
- [43] S. Gieseke, A. Ribon, M. H. Seymour, P. Stephens and B. Webber, “Herwig++ 1.0: An event generator for e+ e- annihilation,” *JHEP* **0402** (2004) 005 [arXiv:hep-ph/0311208].
S. Gieseke *et al.*, “Herwig++ 2.0 release note,” [arXiv:hep-ph/0609306].
- [44] HERWIG 6.5, G. Corcella, I.G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M.H. Seymour and B.R. Webber, *JHEP* 0101 (2001) 010 [hep-ph/0011363]; [hep-ph/0210213]
- [45] S. Gieseke, P. Stephens and B. Webber, “New formalism for QCD parton showers,” *JHEP* **0312** (2003) 045 [arXiv:hep-ph/0310083].
- [46] P. Richardson, “Spin correlations in Monte Carlo simulations,” *JHEP* **0111** (2001) 029 [arXiv:hep-ph/0110108].
- [47] L. Lönnblad, *Comput. Phys. Commun.* **118** (1999) 213.
M. Bertini, L. Lonnblad and T. Sjostrand, “Pythia version 7-0.0: A proof-of-concept version,” *Comput. Phys. Commun.* **134** (2001) 365 [arXiv:hep-ph/0006152].
- [48] T. Plehn, D. L. Rainwater and D. Zeppenfeld, “Determining the structure of Higgs couplings at the LHC,” *Phys. Rev. Lett.* **88** (2002) 051801 [arXiv:hep-ph/0105325].
V. Hankele, G. Klamke and D. Zeppenfeld, “Higgs + 2 jets as a probe for CP properties,” [arXiv:hep-ph/0605117].
V. Hankele, G. Klamke, D. Zeppenfeld and T. Figy, “Anomalous Higgs boson couplings in vector boson fusion at the CERN LHC,” *Phys. Rev. D* **74** (2006) 095001 [arXiv:hep-ph/0609075].

Zusammenfassung

Das Standardmodell der Teilchenphysik beschreibt drei der vier bekannten fundamentalen Wechselwirkungen zwischen Elementarteilchen. Es handelt sich dabei um eine Quantenfeldtheorie, die auf einer $SU(3)_C \times SU(2)_L \times U(1)_Y$ Eichsymmetrie basiert. Kräfte zwischen Teilchen werden im Standardmodell durch die Eichfelder A_μ^a , W_μ^i und B_μ vermittelt, während Materie durch drei 'Generationen' von Quarks und Gluonen beschrieben wird. Experimentelle Beobachtungen haben eine hervorragende Übereinstimmung der Vorhersagen der Theorie mit den Ergebnissen gezeigt, womit das Standardmodell eine der am genauesten überprüften physikalischen Theorien ist.

Allerdings handelt es sich beim Standardmodell um keine alles beschreibende Theorie, zum Einen, weil die Gravitation nicht darin enthalten ist und zum Anderen weil eine große Menge an Parametern wie beispielsweise Massen und Kopplungskonstanten sich nicht aus der Theorie ergeben sondern zusätzlich eingeführt werden müssen.

Aus der Formulierung als Eichtheorie ergibt sich ein weiteres Problem. Die Feldquanten als Träger der schwachen Wechselwirkung müssten aus Gründen der Eichinvarianz masselos sein, wurden aber im Experiment als massiv gefunden. Um die Eichinvarianz dennoch zu gewährleisten wurde der Higgs-Sektor in die Theorie eingeführt. Dieser ist experimentell noch am wenigsten bestätigt, bisher konnte das Higgs-Boson nicht gefunden werden, es war lediglich möglich mithilfe der elektroschwachen Präzisionsdaten Schranken an die Higgsmasse zu ermitteln. Die Ergebnisse des LEP Beschleunigers deuten auf ein eher leichtes Higgs-Boson hin, mit einem Massenbereich zwischen 114 GeV und 166 GeV. Somit ist eine der Hauptaufgaben des momentan im Bau befindlichen Large Hadron Colliders (LHC) darin nach dem Higgs-Boson zu suchen und seine Eigenschaften zu bestimmen.

Die beste Methode zur Higgs-Suche hängt dabei entscheidend von der tatsächlichen Higgsmasse ab, da diese sich auf die möglichen Zerfallskanäle auswirkt. Der im durch die elektroschwachen Präzisionsdaten bevorzugten Massenbereich dominante Produktionsprozess ist die Gluon-Fusion. Dahinter folgt die Vektorboson-Fusion als zweitwichtigster Produktionsprozess, die zwar einen niedrigeren Wirkungsquerschnitt, dafür aber ein klareres Signal aufweist und außerdem die Möglichkeit bietet, die Kopplungen des Higgs-Bosons an Eichbosonen und Fermionen einzeln zu messen. Das Vektorboson-Fusions-Signal ist durch zwei harte Taggingjets gegeben, die im vorwärts- bzw. rückwärtsgerichteten Teil des Detektors liegen und eine große Rapiditätslücke aufweisen. Die Zerfallsprodukte des Higgsbosons fallen zwischen die beiden Jets, in der Mitte des Detektors.

Die Vorhersagen des Standardmodells beziehen sich auf Partonen, also Quarks

und Gluonen und nicht auf Hadronen. Nur letztere aber sind messbar, was auf die Tatsache zurückgeführt wird, dass sie *farbneutral* sind, während Partonen Farbladung tragen. Weiter besteht der Endzustand einer Berechnung auf Matrixelement-Ebene aus einigen wenigen Teilchen, der Endzustand eines Beschleunigerexperiments jedoch aus mehreren Hundert hauptsächlich hadronischer Teilchen.

Um also eine Vorhersage mit experimentellen Ergebnissen vergleichen zu können, muss der Übergang von Partonen im Matrixelement zu Hadronen im Endzustand simuliert werden.

Dies geschieht mit einer eigenen Klasse von Programmen, den *Eventgeneratoren*. Diese verbinden verschiedene Modelle und Konzepte um ausgehend von Partonen messbare hadronische Endzustände zu simulieren. Dazu werden perturbative Ansätze wie der *Partonschauer* und Modelle wie *Hadronisierung* und *Underlying Event* verwendet.

In dieser Arbeit wurde der Einfluss von vollständigen Eventsimulationen auf die Observablen für Vektorboson-Fusion simuliert. Dazu wurden zuerst, nach einer Einführung in den Higgs-Sektor und Vektorboson-Fusions-Reaktionen die Grundlagen der Quantenchromodynamik und der daraus abgeleiteten Observablen beschrieben. Darauf aufbauend folgt eine Beschreibung der Konzepte Partonschauer, Hadronisierung und Underlying Event. Da die Hadronen im Endzustand den auslaufenden Partonen im Matrixelement entsprechen, treten die erzeugten Teilchen in Bündeln, sog. Jets auf. Spezielle Algorithmen, die *Jet-Finder*, rekombinieren die Endzustandsteilchen in Jets, um infrarot und kollinear sichere Observablen zu erhalten. Eine Übersicht über die momentanen Entwicklungen im Bereich der Jet-Algorithmen folgt nach der Beschreibung der Eventgeneratoren.

Im Anschluss daran wird die eigentliche Analyse beschrieben, zusammen mit den verwendeten Eventgeneratoren `vbfnlo` und `Herwig++`.

Hierauf folgt die Diskussion der Ergebnisse. Zunächst wurden die Observablen für VBF Prozesse nach den verschiedenen Simulationsstufen verglichen. Dabei zeigte sich keine große Veränderung in den betrachteten Verteilungen. Die Cut-Effizienz, also die Anzahl der Ereignisse, die die Cut-Bedingungen erfüllen, nahm ab, wobei das Underlying Event einen deutlichen Einfluss hierauf auszuüben scheint. Die Abnahme beträgt elf Prozentpunkte zwischen dem Prozess auf Matrixelement-Niveau und dem Endzustand wenn das Underlying Event mithilfe der UA5-Paramterisierung simuliert wird und ungefähr sechs, wenn das Multiparton Interactions (MPI) Modell verwendet wird.

Der Rückgang der Cut-Effizienz lässt sich durch die Tatsache erklären, dass die Ereignisse im Laufe der Simulation so verändert werden, dass sich beispielsweise die Position oder Energie der Jets ändert und somit die kinematische Signatur des Ereignisses. Diese Ereignisse werden jedoch durch die Cuts aussortiert wodurch das eigentliche Signal erhalten bleibt.

Der größte Einfluss auf die Observablen scheint vom verwendeten Underlying Event Modell auszugehen, je nach verwendetem Modell änderte sich die Position der Peaks in den Verteilungen. Ein schwacher Effekt zeigte sich bei den Verteilungen mit dem MPI Modell, während die UA5 Paramterisierung kaum einen Unterschied zu der

Verteilung nach dem Schauer aufweist.

Weiter wurden die unterschiedlichen Möglichkeiten die Taggingjets zu definieren verglichen. Da die Taggingjets sehr harte Jets sind, wählt man die beiden härtesten Jets als mögliche Taggingjets und wendet die Cuts auf diese an. Dabei lassen sich die härtesten Jets wahlweise als die beiden Jets mit größtem Transversalimpuls oder die mit größter Transversal-Energie definieren. Da die einzelnen Jets aber nur eine kleine Masse $m^2 = E^2 - \vec{p}^2$ aufweisen, ist der Effekt der unterschiedlichen Definitionen minimal, Unterschiede für die Verteilungen wurden erst im Promillbereich gefunden.

Zudem wurde die Jetaktivität, die durch die volle Eventsimulation entstand untersucht.

Dazu wurde zunächst der dritthärteste Jet analysiert, der der härteste und deshalb wohl auch dominanteste Jet ist, der nicht direkt aus der Matrixelementberechnung stammt. Hierzu wurde der dritthärteste Jet nach dem Schauer mit dem dritten Jet im Endzustand für die beiden untersuchten Underlying-Event-Modelle und dem dritten Jet in einer *next-to-leading order* QCD Berechnung des VBF Prozesses verglichen. Die dritten Jets nach dem Schauer konnten in der Mehrheit der Fälle in der Nähe der Taggingjets gefunden werden, wobei circa zwei Drittel der Jets außerhalb der Rapiditätslücke lag. Dies steht im Gegensatz zu den Ergebnissen der NLO Berechnung, bei der der dritte Jet zwar auch nahe bei den Taggingjets lag, aber ungefähr gleichhäufig zwischen und außerhalb dieser.

Für den Endzustand der Simulation zeigte sich wieder eine deutliche Abhängigkeit von dem verwendeten Underlying Event Modell. Während bei der UA5 Parametrisierung kaum eine Abweichung zum Schauer feststellbar ist, zeigen die Verteilungen für das MPI Modell stärkere Unterschiede, die zusätzlichen Wechselwirkungen in diesem Modell sorgen für weitere Jet-Aktivität, die nicht mit der des harten Prozesses korreliert ist.

Um diese weitere Aktivität genauer zu untersuchen und insbesondere den Einfluss auf das eigentliche Signal, wurde die Jet-Aktivität in der Zentralregion analysiert. Hierzu wurden bei allen Ereignissen, die den VBF-Cuts genügen sämtliche Jets, die zwischen den beiden Taggingjets in die Rapiditätslücke fallen nach Transversalimpuls sortiert und der härteste Jet analysiert.

Da erwartet wurde, dass die Jetaktivität in der Zentralregion hauptsächlich vom Underlying Event bestimmt wird, wurde hierfür ein Vergleich von mehreren Underlying-Event-Modellen durchgeführt, verglichen wurden dabei wie bisher die UA5-Parametrisierung, das MPI Modell und zusätzlich eine Simulation ohne ausgefeiltes UE-Modell, in der die Proton-Remnants einfach als normale Cluster, wie sie in der Hadronisierung auftreten aufgefasst werden.

Die Rapiditätsverteilungen für die beiden Underlying-Event-Modelle zeigten keine großen Abweichungen voneinander. Im Vergleich zu der Verteilung ohne Underlying-Event-Modell befanden sich die zentralen Jets weiter in der Mitte, während der Jet aus der Simulation ohne UE-Modell eher in der Nähe der Taggingjets lag, ähnlich den Verteilungen der dritthärtesten Jets nach dem Schauer. Ein deutlicher Unterschied zwischen den Modellen konnte in den Transversalimpulsverteilung des härtesten zentralen Jets gefunden werden, die UE-Modelle produzierten härtere Jets als die Sim-

ulation ohne ein solches Modell, wobei die UA5-Parametrisierung vermehrt Jets mit Transversalimpuls $p_T \approx 5$ GeV und das MPI Modell Jets mit $p_T \approx 15$ GeV erzeugt. Für Transversalimpulse größer als 25 GeV waren die drei Verteilungen nahezu identisch. Aus diesen Verteilungen wurde dann eine *Central Jet Veto* Effizienz ermittelt. Da Ereignisse mit harten Jets in der Zentralregion nicht die gewünschte VBF-Signal Signatur aufweisen, sollen sie verworfen werden. Also wurde aus der Verteilung der Transversalimpulse der zentralen Jets eine Verteilung erstellt, die den Anteil der Events anzeigt, die auch bei einem Veto mit gegebenem $p_{T,veto}$ noch akzeptiert werden würden. Es ergab sich eine sehr hohe Effizienz bei akzeptablem Wert für $p_{T,veto}$, für ein Veto mit $p_{T,veto} = 20$ GeV ergibt sich abhängig vom UE-Model eine Effizienz von ca. 95% bis 97%, für niedrigere Werte von $p_{T,veto}$ waren die Abweichungen größer, für $p_{T,veto} = 10$ GeV lag die Effizienz zwischen 65% für das MPI Modell und 90% bei dem Sample ohne Underlying Event.

Zuletzt wurde der Einfluss der Simulation auf die Taggingjets selbst untersucht. Hierzu wurden die Veränderungen im Transversalimpuls der Taggingjets und die Änderungen in der Richtung untersucht. Dies erfolgte für die Übergänge vom Matrixelementniveau zum Schauer und zum Endzustand der Simulation, wieder für beide verwendeten Underlying-Event-Modelle.

Als Ergebnis konnte festgehalten werden, dass der härteste Taggingjet im wesentlichen kaum durch die Simulation beeinflusst wird, für die meisten Ereignisse ändern sich weder der Transversalimpuls noch die Richtung des härtesten Jets, nur für eine kleine Anzahl der Ereignisse konnten deutliche Veränderungen festgestellt werden.

Der zweithärteste Taggingjet wies hier eine größere Aktivität auf. Für die Mehrheit der Ereignisse waren zwar Transversalimpuls und Richtung auch nur geringfügig verändert, jedoch konnte eine gewisse Menge an Ereignissen gefunden werden, bei denen der Transversalimpuls erheblich zunahm und die Richtung stark von der ursprünglichen abwich. Dies kann dadurch erklärt werden, dass in diesen Ereignissen, die sich zusätzlich durch eher weiche Taggingjets auszeichnen, in der Simulation ein weiterer harter Jet entsteht, der härter ist als der ursprüngliche zweite Jet. Damit ändern sich dann natürlich Richtung und Transversalimpuls erheblich.

Zum Abschluss wurde noch die Transversalimpulsverteilung der Taggingjets für Ereignisse, bei denen die Taggingjets ihre Richtung beibehalten und für alle Ereignisse verglichen. Dabei zeigte sich kein erheblicher Unterschied zwischen den beiden Verteilungen, die Ereignisse, bei denen die Jetsignatur stark verändert wurde, scheinen also keinen großen Einfluss auf die Observablen auszuüben.

Für diese Analyse konnten keine Detektoreffekte berücksichtigt werden, da hierzu eine volle Detektorsimulation notwendig wäre. Selbstverständlich spielen jedoch die hierbei auftretenden Effekte eine große Rolle, die in eine volle Analyse und beim Vergleich mit experimentellen Daten mit berücksichtigt werden müssen.

Danksagung

Zunächst möchte ich mich bei Herrn Prof. Dr. Zeppenfeld für die Möglichkeit bedanken, diese Diplomarbeit auf einem spannenden Themengebiet anfertigen zu können. Dankbar bin ich auch für die gute Betreuung und die Möglichkeit, auch neben dem eigentlichen Themengebiet viel über die aktuelle Teilchenphysik zu lernen.

Ebenso möchte ich mich bei Herrn Dr. Gieseke bedanken, der diese Arbeit mitbetreut hat und sich viel Zeit genommen hat, mir weiterzuhelfen und Rückmeldungen zu geben.

Bei Herrn Professor Dr. Klinkhamer bedanke ich mich für seine Bereitschaft, das Korreferat zu übernehmen.

Allen Institutsangehörigen danke ich für die freundliche und professionelle Atmosphäre und das Beantworten zahlreicher Fragen. Besonders möchte ich dabei Manuel Bähr und Simon Plätzer danken, die mir beim Umgang mit **Herwig++** viel geholfen haben. Bei meinen Mitdiplomanden und Bürokollegen Christoph Englert, Jan Germer und Matthias Werner bedanke ich mich für die vielen Diskussionen und das gute Miteinander in unserem Büro.

Weiter bedanke ich mich bei Leif Lönnblad, Peter Richardsson und Mike Seymour, die mir bei Problemen mit **ThePEG** und **Herwig++** weiterhalfen.

Simon Plätzer danke ich stellvertretend für die Systemadministratoren für die Wartung und Bereitstellung einer funktionierenden IT-Infrastruktur.

Schließlich möchte ich meinen Eltern danken, die mich während meines Studiums unterstützt haben.

Ebenso will ich mich bei meinen Freunden, die mich während meiner Diplomarbeitszeit begleitet haben bedanken.