PHOTOPRODUCTION OF EVENTS WITH RAPIDITY GAPS BETWEEN JETS WITH ZEUS AT HERA

by

PATRICK D. RYAN

A dissertation submitted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY
(PHYSICS)

at the

UNIVERSITY OF WISCONSIN – MADISON

2006

Abstract

Cross sections for the photoproduction of dijet events, where the two jets with the highest transverse energy are separated by a large gap in pseudorapidity, have been studied with the ZEUS detector using an integrated luminosity of 38.6 pb^{-1}. Rapidity-gap events are defined in terms of the energy flow between the jets, such that the total summed transverse energy in this region is less than some value $E_T^{\rm CUT}$. The data show a clear excess above the predictions of standard photoproduction models. Models which include color-singlet exchange are able to describe the data.

Acknowledgements

I would first like to thanks my advisors, Wesley Smith and Don Reeder, for providing leadership and support throughout my career as a graduate student.

I would also like to thank the other members of the Wisconsin group: Matthew Jaworski, Alexandre Savin, Dorian Kçira, Sabine Lammers, Liang Li, Adam Everett, Michele Sumstine, Tom Danielson, Eric Brownson, and Homer Wolfe.

Thanks to Claire Gwenlan (and her 388 page thesis), who worked with me on this analysis.

Thanks to my friends in Wisconsin: Charles Tahan, Ryan Miller, and Attila Mihalyi.

Thanks to Simon Albino, Bob Oliver, and Freddy Poirier, whose involvement in the DPC was responsible for many enjoyable beach parties and other activities.

Thanks to Alessandro Montanari, who showed me the proper way to cook frozen pizzas.

Thanks to the many other people in Hamburg who have made the last 4 years so enjoyable: Devis, Marcos, Nastja, Dave, Matt, Paul, Elke, Billy, Ricardo, Riccardo, Nabil, Gabi, Jessica, and Jean-Paul.

Thanks to the many places in Hamburg which have made the last 4 years so enjoyable: Nouar, Le Fonque, Fundbureau, Waagenbau, Kir, Ex-Sparr, Nachtlager, Bad Taste, Thomas Read, Golden Pudel, Cobra, Molly Malone, Datscha, Feuerstein, Aurel, Blaues Haus, Reh, Titanic, and Fabrik.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The goal of physics is to describe and explain the behavior of the universe. Particle physics is the discipline concerned with discovering the most basic constituents of matter and describing how these constituents interact with one another.

1.1 A Brief History of Particle Physics

The earliest attempts at particle physics are generally attributed to the Greeks, who set the foundations of modern principles such as conservation of matter and atomic theory. Anaxagoras (500-428 B.C.) postulated that the creation and destruction of matter is due to a rearranging of indivisible particles and Empedocles (484-424 B.C.) developed the theory that the indivisible particles belonged to one of four elements; air, earth, fire, or water. Democritus hypothesized that the universe consists of empty space and indivisible particles called atoms.(460-370 B.C.) Aristotle (384-322 B.C.) investigated a wide range of subjects and formed the groundwork for the methods of scientific study that would last for over 1000 years.

It wasn't until Copernicus (1473-1543) proposed his heliocentric model of the universe that the philosophical methods of Aristotle were superseded by empirical

study. Galileo Galilei (1564-1642) made many advances in the areas of astronomy and mechanics, and replaced many long held assumptions with rigorously developed theories. This paved the way for Isaac Newton (1642-1727), who developed the mechanical laws of motion and the law of gravity. In 1869, Dmitri Mendeleev ordered all known elements by their atomic mass into the Periodic Table. This implied an underlying structure of matter and accurately predicted the mass of those elements missing from the table. The electric and magnetic forces were united in 1837 by James Clerk Maxwell to form his electromagnetic theory.

The era of modern particle physics began in 1898 with Joseph Thompson's discovery of the electron. Thompson incorporated these electrons into a model of the atom in which he analogously called the electrons negatively charged plums in a positively charged pudding. In 1911, Ernest Rutherford scattered α-particles off gold atoms and demonstrated that most of the atom's mass is concentrated in a very small volume called the nucleus. Rutherford later found the first evidence of the proton, the name which he gave the hydrogen nucleus, while scattering α-particles off nitrogen gas. In 1905 and 1911 Einstein devolved the theories of special and general relativity, and soon after the major advances in quantum theory were made by de Broglie, Pauli, Schrödinger, Heisenberg, etc. In 1931 James Chadwick discovered the neutron, the third, and final, basic constituent of matter.

In 1934 Enrico Fermi developed a theory of β decay, which was the first to explicitly include neutrinos. In the following years, the muon, pion, and K^+ were observed, and 1953 marked the beginning of the "particle explosion" in which many more new particles were discovered. Schwinger, Bludman, and Glashow separately developed a theory of weak interactions between 1957 and 1959 and the first experimental evidence of neutrinos was found in 1962.

In 1964 Murray Gell-Mann and George Zweig developed a model of particle structure based upon quarks and in 1965 the quark property of color was proposed. Weinberg and Salam unified the electromagnetic and weak interaction into the electroweak interaction in 1967. An experiment at the Stanford Linear Accelerator Center in 1968, which scattered electrons off protons, provided the first experimental evidence of quarks. The theory of Quantum Chromodynamics, a quantum field theory of the strong force which describes the interactions of quarks and gluons, was formulated in 1973. The J/ψ particle, consisting of a charm and anti-charm quark was discovered independently at SLAC and Brookhaven in 1974, the tau lepton was discovered at SLAC in 1976, and the bottom quark and anti-quark were discovered at Fermilab in 1977. The first experimental evidence of the gluon was observed at the DESY laboratory in 1979 and in 1983 the $W^{\pm}$ and Z^0, force-carrying particles in the weak interaction, were seen at CERN. After many years of searching, the top quark was finally discovered at Fermilab in 1995.

1.2 The Standard Model

The **Standard Model** is currently the most complete theory of fundamental particles and their interactions. It includes the strong, weak, and electromagnetic forces and the three types of elementary particles; leptons, quarks, and gauge bosons, but not the force of gravity.

Quarks are spin-$\frac{1}{2}$ particles of fractional charge and color. There are six types (flavors) of quarks which can be divided into three generations, each consisting of two

Quarks			**Leptons**		
Flavor	**Mass (GeV/c^2)**	**Charge**	**Flavor**	**Mass (GeV/c^2)**	**Charge**
u	0.003	+2/3	ν_e	$< 1 \times 10^{-8}$	0
d	0.006	-1/3	e	5.11×10^{-3}	-1
c	1.3	+2/3	ν_μ	< 0.0002	0
s	0.1	-1/3	μ	0.106	-1
t	175	+2/3	ν_τ	< 0.02	0
b	4.3	-1/3	τ	1.7771	-1

Table 1.1: The quarks and leptons are grouped into three generations, with each generation consisting of two particles each. The only difference among generations are the mass and flavor of the particles. The basic unit of charge, -1, is defined as the charge of the electron. Each particle has an anti-particle (not shown) which has the same mass but opposite charge.

quarks, and differing from other generations only in flavor and mass. Quarks carry a color of either red, green, or blue, and combine to form colorless particles called **hadrons**. **Baryons** are hadrons consisting of three quarks of different colors (red, blue, and green) or three anti-quarks of different colors (anti-red, anti-blue, and anti-green). **Mesons** are hadrons consisting of a quark and anti-quark combination with colors of red and an anti-red, blue and anti-blue, or green and anti-green. Only these colorless combinations of quarks are observed[1].

There are six **leptons**, also spin-$\frac{1}{2}$ particles, which, in contrast to quarks, can be observed independently. The electron, muon, and tau all carry a negative charge of one unit, and their differences are due only to their different masses and flavors. Each also has a corresponding neutrino which carries no electric charge and has a very small mass. The quarks and leptons and their basic properties are listed in Table 1.2.

---

[1] The exception is the top quark, which is so short lived that it decays before forming a colorless meson.

A force between two particles can be explained by the exchange of a third particle between the two interacting particles. The exchange particles are said to carry, or “mediate” the force and constitute the third, and final, category of particles in the standard model. These particles are of integer spin and are called **gauge bosons** since the Lagrangian of each set of mediating bosons is invariant under a gauge transformation.

The **photon** is a massless particle with no electric charge that mediates the electromagnetic interaction. This interaction is responsible for binding the electron to the nucleus in an atom and for most everyday forces, such as friction and the force holding a table together. The weak interaction is mediated by the W^+ or W^- particles which have a charge of plus or minus one unit respectively, or the electrically neutral Z^0. The weak interaction is responsible for the decay of certain particles, such as the neutron, and allows quarks and leptons to change flavor when they decay. **Flavor** is a quantum number associated with elementary particles, and a particle’s flavor is the name by which it is usually referred. For example, up, down, strange, etc. for quarks and electron, electron neutrino, muon, muon neutrino, etc. for leptons. The theories of the electromagnetic (**Quantum Electrodynamics** or QED) and weak interactions have been merged into the into a single theory of the electroweak interaction, which supersedes and encompasses all aspects of both independent theories.

The **gluon**, also massless and electrically neutral, carries a color and anti-color arranged in one of eight possibilities (for example red and anti-green or blue and anti-red) and mediates the strong interaction. The strong interaction, described by the theory of **Quantum Chromodynamics** (QCD), binds quarks together using gluons and residual effects from this interaction hold the protons and neutrons together in the nucleus. The quarks and gluons are together referred to as **partons**.

Boson	**Mass**	**Charge**	**Color**	**Interaction**
γ (photon)	0	0	No	Electromagnetic
$W^\pm$	80.4	±1	No	Weak
Z^0	91.187	0	No	Weak
g (gluon)	0	0	Yes	Strong

Table 1.2: The gauge bosons. The basic unit of charge, -1, is defined as the charge of the electron.

The Higgs boson, which has yet to be experimentally observed, is the final piece of the Standard Model and is required to induce spontaneous symmetry breaking of the gauge groups and for the existence of mass. The gauge bosons and their basic properties are listed in Table 1.2.

Although the Standard Model has been very successful in explaining experimental results, it is not accepted as a fundamental theory of particle physics. Gravity is not incorporated into this theory, neutrino oscillations are not explained, and there are many free parameters, such as particle mass and the fundamental unit of charge, which can not be predicted and instead must be determined experimentally. One of the major goals of physics is to join all four forces into a Grand Unified Theory (GUT).

Chapter 2

Quantum Chromodynamics and Lepton-Hadron Physics

2.1 Quantum Chromodynamics

Quantum Chromodynamics (QCD) is a non-Abelian gauge theory of the group SU(3) which describes the strong interaction between quarks and gluons. A **gauge theory** requires symmetry transformations of the Lagrangian to be be valid locally (applicable in a certain region of space-time and not affecting other regions of space-time) in addition to globally (identical at every point in space-time). The color charge of the quarks and gluons is the local symmetry in QCD. **Non-Abelian** means that group operations are not commutative, but in practice this means that in QCD the gluons (gauge bosons) can interact with other gluons. The group SU(3) is that in which the matrices are special (determinant is 1), unitary (the matrix multiplied by its transpose conjugate is the identity matrix), and of dimension 3.

2.1.1 Asymptotic Freedom and Confinement

The two most distinct properties of QCD are asymptotic freedom and confinement. **Confinement** specifies that free quarks are unobservable and that only colorless objects, such as the proton or pion, can be seen in the laboratory. Although analytically unproven[1], the property is generally attributed to the idea that the potential energy increases infinitely as quarks are pulled infinitely far apart. It is therefore more advantageous, from an energy perspective, to create quark-anti-quark pairs from the vacuum to form colorless objects when the quarks are being pulled apart.

To understand asymptotic freedom, it is useful to draw an analogy to QED, an Abelian theory in which the photons do not interact with each other. Imagine that one wants to measure the charge of an electron in a vacuum. This electron, however, can radiate photons which then split into electron-positron pairs. The positrons of the electron-positron pairs are attracted to the original electron and the electrons are repelled (see Fig. 2.1). This creates an effect called screening in which the measured effective charge of the electron is reduced as the distance between the observer and the electron increases. The analogous screening effect happens for color charge in QCD when a quark emits gluons which then split into quark-anti-quark pairs, as shown in the diagram on the left side of Fig. 2.2. However, because of the self-coupling of gluons in QCD, the gluons can also radiate additional gluons, as shown in the diagram on the right side of Fig. 2.2. Since color charges like to be surrounded by charges of a similar color the latter effect dominates and thus the strength of the color charge decreases the closer one comes to the original quark, a property called **asymptotic freedom**.

---

[1] Confinement has been proven numerically using lattice techniques.

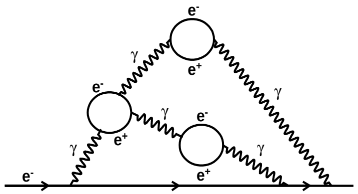


Figure 2.1: Electric charge screening in QED. The initial electron radiates a photon which splits into electron-positron pairs. The positrons are attracted to the original electron the electrons are repelled. This shielding reduces the measured charge of the original electron.

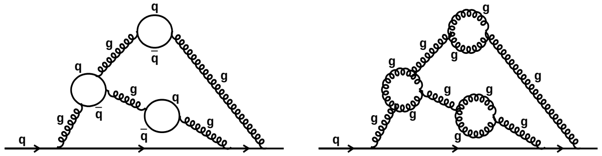


Figure 2.2: Color charge screening in QCD. As shown in the figure on the left, the initial gluon can radiate a gluon which splits quark-anti-quark pairs in analogy to screening in QED. In addition, as shown in the figure on the right, the gluons may also split into gluon-gluon pairs.

2.1.2 Running Coupling

The strength of the electroweak interaction is determined by the electroweak coupling constant, α_{EW}, which has a value of approximately 1/137 at zero energy. The strength of the strong interaction is analogously governed by the strong coupling constant, α_s. As described above, the strength of the strong interaction, and therefore α_s, is described by asymptotic freedom. When color charges are close together (at high energies), the value of α_s is small, but when the charges are far apart (at low energies), the value of α_s is large. Mathematically, this behavior is described by the expression

$$\alpha_s\left(Q^2\right) = \frac{12\pi}{(11n - 2f)\ln\left(Q^2/\Lambda^2\right)}, \tag{2.1}$$

where Q^2 is the square of the four momentum transfered between the two interacting particles (energy of the probe in a scattering experiment), n is the number of colors, f is the number of quarks, and Λ is an experimentally determined parameter whose exact value is difficult to determine but is generally accepted to be on the order of a few hundred MeV. In the standard model, $n = 3$ and $f = 6$, which gives a positive α_s. Therefore, α_s decreases as the energy increases and; as $Q^2 \rightarrow \infty$, $\alpha_s \rightarrow 0$, exhibits asymptotic freedom. Diagrams of the behavior of α_{EW} and α_s as a function of probing energy are shown in Fig. 2.3.

2.1.3 The Order in α_s

A generic example of the perturbative expansion of a quantity, in this case A, in terms of α_s is:

$$A = A_0 + A_1\alpha_s + A_2\alpha_s^2 + A_3\alpha_s^3 + ... \tag{2.2}$$

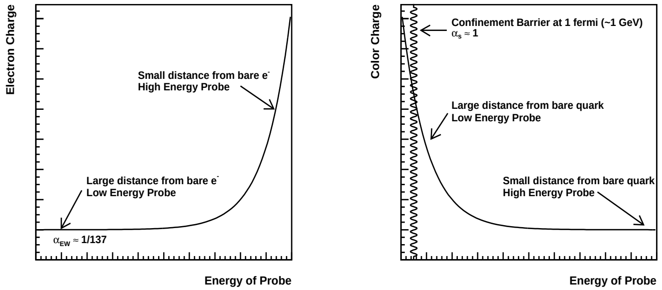


Figure 2.3: The QED electric charge (left figure) and QCD color charge (right figure) as a function of energy transfered between two interacting particles. As the energy of the probe increases the QED coupling constant, α_{EW} increases and the QCD coupling constant, α_s decreases. This decrease is termed asymptotic freedom. Diagrams based on those in [1].

The term "order" refers to the power of the α_s term in the expansion and corresponds to the number of QCD vertices in the Feynman Diagram describing the process. Each vertex increases the order by a term in α_s, so that a diagram calculated to $\mathcal{O}(\alpha_s^1)$ has one α_s vertex, a diagram of $\mathcal{O}(\alpha_s^2)$ has two α_s vertices, etc.

Leading Order (LO) refers to the lowest power of α_s in the expansion of A which contributes to the process being studied. For example, α_s^1 is the leading order for a process with one final state parton and α_s^2 is the leading order for a process with two final state partons. **Next to Leading Order** (NLO) refers to the power of α_s in the expansion of A which is one higher than that of the leading order term. This extra term in the expansion corresponds to an extra parton in the final state represents the radiation of a gluon. **Next to Next to Leading Order** (NNLO) refers to the power of α_s in the expansion of A which is two higher than that of the leading order term. This term corresponds to two extra partons in the final state and

represents the radiation of two gluons. NNLO process can also be a result of one of the partons temporarily fluctuating into a $q\bar{q}$ or gg pair via an internal loop[2]. These loops cause divergences in the calculations which must be dealt with using a mathematical technique called renormalization. Fig. 2.4 shows examples of LO, NLO, and NNLO processes.

2.1.4 Perturbative QCD

Hard interactions are those in which $Q^2 \gg \Lambda^2$, and therefore $\alpha_s \ll 1$. The small value of α_s allows for the convergence at higher orders in a perturbative expansion in terms of α_s. The method of calculating approximate analytic solutions using a perturbative expansion is called **Perturbative QCD** (pQCD). In **soft interactions**, where $Q^2 \rightarrow \Lambda^2$ and $\alpha_s \rightarrow 1$, the expansion in terms of α_s does not converge, and pQCD is not applicable. Instead, one must rely on other means of calculation, such as the numerical methods of lattice QCD.

2.2 Introduction to Lepton-Proton Collisions

In the most general case of a lepton-proton (ep) collision, the incoming lepton scatters from the proton via a gauge-boson into an outgoing lepton and a fragmented proton in its hadronic final state. A generic form of this interaction is pictured in Fig. 2.5, where k is the momentum of the incoming lepton, k' is the momentum of the outgoing lepton, and p is the momentum of the incoming proton. The exchange boson, a photon for electromagnetic interactions or a Z^0 or $W^{\pm}$ for weak interactions, carries a momentum of $q = k - k'$ and the hadronic final state of the proton (X) has a momentum of $p + q$.

---

[2] If there is no NLO process due to gluon radiation in the interaction, then pair production may be referred to as NLO.

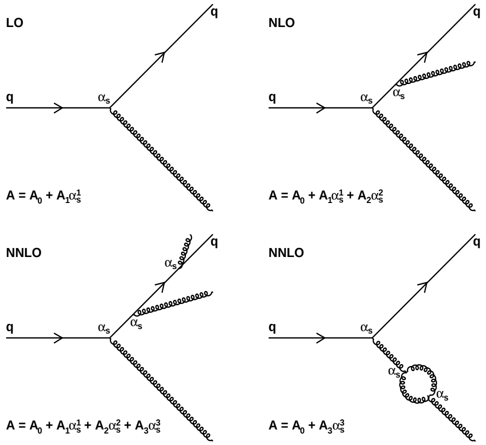


Figure 2.4: Diagrams of generic LO, NLO, and NNLO processes in which there is one initial particle which undergoes one split. The upper left diagram has one parton in the final state and is LO, the upper right diagram has two particles in the final state and is NLO, the lower left diagram has three partons in the final state and is NNLO, and the lower right diagram, which includes an internal loop, is also NNLO.

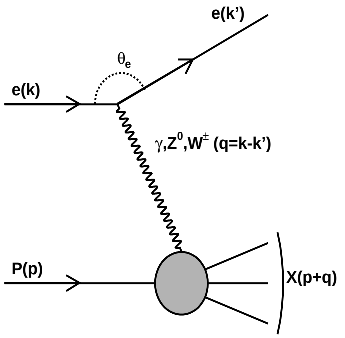


Figure 2.5: A schematic diagram of lepton-proton scattering. The four-momenta of the particles and systems of particles are given in parentheses.

The scattering angle between the incoming and outgoing leptons is given by θ_e.

The lepton-proton interaction can be completely described using the momenta of the incoming and outgoing leptons and the incoming proton. Knowledge of the hadronic final state of the proton is not needed since its properties can be determined from those of the other three particles and the application of four-momentum conservation.

Several Lorentz invariant quantities are used to describe the interactions. The center-of-mass energy of the *ep* system is given by

$$s = (k + P)^2 \approx 4E_eE_p, \tag{2.3}$$

where E_e is the energy of the incoming lepton and E_p is the energy of the incoming proton, The square of the four-momentum of the exchange boson is given by

$$Q^2 = -q^2 = (k - k')^2, \tag{2.4}$$

which, if the leptons are assumed to be massless, can be expressed as

$$Q^2 \approx 2k \cdot k' \approx 2E_e E_{e'}(1 - \cos\theta_e), \tag{2.5}$$

where E_e and $E_{e'}$ are the energies of the incoming and outgoing leptons. The negative sign in Equation 2.4 ensures that Q^2 is a positive quantity since the exchange boson in ep collisions is space-like with $q^2 < 0$[3]. Two more Lorentz invariant quantities are

$$x = \frac{Q^2}{2p \cdot q}, \tag{2.6}$$

the fraction of the proton's momentum carried by the quark struck by the lepton, and

$$y = \frac{p \cdot q}{p \cdot k}, \tag{2.7}$$

which, in the rest frame of the proton, reduces to

$$y = \frac{E_e - E_{e'}}{E_e} = 1 - \frac{E_{e'}}{E_e}, \tag{2.8}$$

where E_e and $E_{e'}$ are the energies of the incoming and outgoing leptons. This can be interpreted as the fraction of the lepton energy transfered to the proton, or the inelasticity of the interaction. The center of mass energy of the exchange boson-proton system is given by

$$W = (q + p)^2. \tag{2.9}$$

The above quantities are related by

$$Q^2 = sxy \tag{2.10}$$

and

$$W = sy - Q^2 + m_p^2(1 - y). \tag{2.11}$$

---

[3] In other collisions, such as $p\bar{p}$ and e^+e^-, the exchange boson can be time-like, with $q^2 > 0$.

Due to the quantum mechanical uncertainty principle, it is not required that the exchange boson obey the laws of conservation of momentum and energy as long as it exists for a time $\Delta t < \hbar/\Delta E$. An exchange boson which doesn't have the "correct" energy or momentum is termed a **virtual boson**, and Q^2 is referred to as the virtuality of the interaction. The virtuality can range from $0 < Q^2 < s$, with greater virtualities corresponding to smaller wavelengths and therefore a greater resolving power, of the the exchange boson. The wavelength, λ, is related to the three-momentum of the exchange boson by the relation $\lambda = 1/|\boldsymbol{q}|$, where $\boldsymbol{q} = \boldsymbol{k} - \boldsymbol{k}'$. The three-momentum is related to the four-momentum by

$$|\boldsymbol{q}| = \left[q_0^2 + Q^2\right]^{1/2} = \left[(E_e - E_{e'})^2 + Q^2\right]^{1/2}, \tag{2.12}$$

In the proton rest frame,

$$x = \frac{Q^2}{2p \cdot q} = \frac{Q^2}{2m_p(E_e - E_{e'})}, \tag{2.13}$$

where m_p is the mass of the proton. Substituting Equation 2.13 into Equation 2.12 gives

$$|\boldsymbol{q}| = \left[\left(\frac{Q^2}{2m_p x}\right)^2 + Q^2\right]^{1/2} = \left(\frac{Q^2}{2m_p x}\right)\left[1 + \frac{(2m_p x)^2}{Q^2}\right]^{1/2} \approx \left(\frac{Q^2}{2m_p x}\right) \tag{2.14}$$

for $Q^2 \gg (2m_p x)^2$. Therefore, the wavelength of the virtual boson is given by

$$\lambda \approx \frac{2m_p x}{Q^2}. \tag{2.15}$$

Interactions with large Q^2 ($Q^2 >> 1\,\mathrm{GeV}^2$), termed "hard", are therefore able to resolve smaller distances than interactions with small Q^2 ($Q^2 << 1\,\mathrm{GeV}^2$), which are termed "soft". Processes where $Q^2 \approx 0$ are classified as **photoproduction** and, if

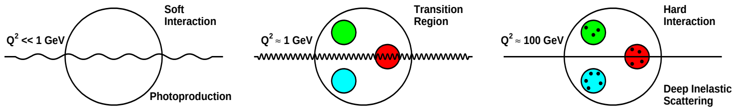


Figure 2.6: As Q^2 increases, the virtual exchange boson is able to probe smaller and smaller distances. In photoproduction ($Q^2 \approx 0$), only the gross structure of the proton is visible, for $Q^2 \approx 1$, the quarks are resolved, and for large Q^2 (DIS) the quarks are resolved into other quarks and gluons.

there is no other hard scale present, can only resolve the gross details of the photon. Process with $Q^2 \approx 1$ can resolve the quark structure of the proton, and those with larger Q^2, categorized as **Deep Inelastic Scattering** (DIS), can resolve the sea quarks and gluons. Resolution as a function of Q^2 is pictured in Fig. 2.6.

2.3 The ep Cross Section

2.3.1 Introduction to the Cross Section

The cross section is a quantity which describes the likelihood that a certain process will occur. Imagine scattering a particle of infinitesimal radius off a spherical target of radius r. The incoming particle "sees" only the cross sectional area πr^2 of the target particle, and therefore, the larger the cross section of the target particle, the greater the chance the incoming particle has to strike it.

In particle physics, the cross section, σ, for interactions such as $AB \to CD$ can be expressed as

$$\sigma\ [L^2] = \left(\frac{\text{Transition Rate } [1/T]}{\text{Initial Flux } [N/(L^2T)]}\right)(\text{Number of Final States } [N]) \tag{2.16}$$

where the dimensions, L for *Length*, T for *Time*, and N for *Number of Particles*, are

enclosed in square brackets. The *Transition Rate* multiplied by the *Number of Final States* represents the number of scatters per unit time and the *Initial Flux*, which represents the number of incoming particles per unit time per unit area, ensures that the cross section is independent of the luminosities of the colliding particle beams. It can therefore be seen that the cross section is the effective area over which particles A and B interact to produce C and D. The differential cross section can also be written as

$$d\sigma \ = \left(\frac{\text{Transition Rate}}{\text{Initial Flux}}\right) dQ, \tag{2.17}$$

where dQ is the differential of the number of final states.

2.3.2 The Deep Inelastic Scattering Cross Section

The general form for the Neutral Current (NC) DIS cross section interaction ($ep \to eX$) can be written as

$$d\sigma \sim L^e_{\mu\nu} W^{\mu\nu}, \tag{2.18}$$

where $L^e_{\mu\nu}$ is the leptonic tensor and $W^{\mu\nu}$ is the hardonic tensor. The leptonic tensor, which describes the connection among the incoming lepton, the gauge boson, and the outgoing lepton, can be written as

$$L^e_{\mu\nu} = 2\left[k'_\mu k_\nu + k'_\nu k_\mu - (k' \cdot k - m_e^2) g_{\mu\nu}\right] \tag{2.19}$$

where m_e is the mass of the lepton and $g_{\mu\nu}$ is the metric tensor. The hadronic tensor describes the connection among the incoming proton, the gauge boson, and the outgoing hadronic final states. It is is constructed from $g_{\mu\nu}$ and the four-momenta of the proton, p, and photon, q, and written as

$$\begin{aligned} W^{\mu\nu} = &- W_1 g^{\mu\nu} + \frac{W_2}{m_p^2} p^\mu p^\nu + i\epsilon^{\mu\nu\alpha\beta} p_\alpha q_\beta W_3 \\ &+ \frac{W_4}{m_p^2} q^\mu q^\nu + \frac{W_5}{m_p^2} \left(p^\mu q^\nu + q^\mu p^\nu\right) + \frac{W_6}{m_p^2} \left(p^\mu q^\nu - q^\mu p^\nu\right). \end{aligned} \tag{2.20}$$

The unknown information about the coupling between the gauge boson and the proton is absorbed into the W_i terms, called structure functions, which are process dependent and are related to the distribution of spatial charge within the proton. The antisymmetric W_6 term can be omitted since it will vanish anyway after contraction with the symmetric $L^e_{\mu\nu}$. Conservation of current at the hadronic vertex requires that $q_\mu W^{\mu\nu} = q_\nu W^{\mu\nu} = 0$, so that

$$W_5 = -\frac{p \cdot q}{q^2} W_2 \tag{2.21}$$

and

$$W_4 = \left(\frac{p \cdot q}{q^2}\right)^2 W_2 + \frac{m_p^2}{q^2} W_1. \tag{2.22}$$

This enables one to write

$$W_{\mu\nu} = W_1 \left(-g^{\mu\nu} + \frac{q^\mu q^\nu}{q^2}\right) + W_2 \frac{1}{m_p^2} \left(p^\mu - \frac{p \cdot q}{q^2} q^\mu\right) \left(p^\nu - \frac{p \cdot q}{q^2} q^\nu\right). \tag{2.23}$$

After contracting the leptonic and hadronic tensors and replacing the structure functions by their more commonly used forms given by

$$F_1(x, Q^2) = m_p W_1(x, Q^2) \tag{2.24}$$

$$F_2(x, Q^2) = \frac{Q^2}{2m_p x} W_2(x, Q^2) \tag{2.25}$$

$$F_3(x, Q^2) = \frac{Q^2}{2m_p x} W_3(x, Q^2) \tag{2.26}$$

$$\tag{2.27}$$

one obtains the neutral current double differential cross section (see, for example [1]):

$$\frac{d^2\sigma^{e\pm p}}{dxdQ^2} = \frac{4\pi\alpha_{EM}^2}{xQ^4} \left[y^2 x F_1(x, Q^2) + (1-y) F_2(x, Q^2) \mp y \left(1 - \frac{y}{2}\right) x F_3(x, Q^2)\right], \tag{2.28}$$

where the $-$ in $\pm$ is for leptons and the $+$ for the corresponding anti-lepton. Defining the longitudinally polarized structure function

$$F_L(x, Q^2) \equiv F_2 - 2xF_1 \tag{2.29}$$

and substituting it into Equation 2.28 one obtains

$$\frac{d^2\sigma^{e\pm p}}{dxdQ^2} = \left[Y_+ F_2(x, Q^2) - y^2 F_L(x, Q^2) \mp Y_- x F_3(x, Q^2)\right] \tag{2.30}$$

where $Y_\pm = 1 \pm (1-y)^2$. F_2 is the contribution to the cross section from γ and Z^0 exchange. F_3 is the parity violating contribution arising from Z^0 exchange and $\gamma - Z^0$ interference and it's contribution is very small small for $Q^2 \ll M_Z^2$. F_L is the contribution from the longitudinally polarized photon.

2.3.3 The Photon-Proton Cross Section

The role of the lepton beam in lepton-proton scattering is simply to provide a source of virtual photons with which to probe the proton. It is therefore illustrative to factor the lepton-proton cross section into terms corresponding to the photon-proton (γp) cross section and a flux factor giving the probability that the lepton beam will emit a photon of a certain energy $E_\gamma = yE$ and Q^2. For virtual photons with $Q^2 > 0$ the photon has both longitudinal and transverse polarizations and the total photon-proton cross section can be written as a sum of these terms:

$$\sigma^{\gamma p}_{\mathrm{Tot}} = \sigma^{\gamma p}_{\mathrm{T}} + \sigma^{\gamma p}_{\mathrm{L}}. \tag{2.31}$$

The longitudinal and transverse cross sections are related to the structure functions by

$$2xF_1 = \frac{K}{4\pi\alpha_{EW}}\frac{Q^2}{\nu}\sigma_{\mathrm{T}}^{\gamma p} \tag{2.32}$$

$$F_2 = \frac{K}{4\pi\alpha_{EW}}\frac{Q^2\nu}{Q^2+\nu^2}\left(\sigma_{\mathrm{T}}^{\gamma p}+\sigma_{\mathrm{L}}^{\gamma p}\right) \tag{2.33}$$

where K, the virtual photon flux, and $\nu \equiv \frac{p\cdot q}{m_p}$, are related by the Hand convention [2]:

$$K = \nu - \frac{Q^2}{2m_p}. \tag{2.34}$$

The lepton-proton cross section can then be written as

$$\frac{d^2\sigma^{ep}}{dydQ^2} = \frac{\alpha_{EW}}{2\pi}\frac{1}{Q^2}\left[\left(\frac{1+(1-y)^2}{y} - 2\frac{(1-y)}{y}\frac{Q^2_{min}}{Q^2}\right)\sigma_{\mathrm{T}}^{\gamma p} + 2\frac{(1-y)}{y}\sigma_{\mathrm{L}}^{\gamma p}\right] \tag{2.35}$$

where $Q^2_{min} = m_e^2\left(\frac{y^2}{1-y}\right)$ is the lower bound on Q^2.

2.3.4 The Photon-Proton Cross Section in the Photoproduction Regime

In the photoproduction regime ($Q^2 \sim 0$) the photons are quasi-real. Since the lepton-proton cross section is proportional to $\frac{1}{Q^4}$, photoproduction dominates over the higher Q^2 DIS processes in the total cross section. Real photons have no longitudinal component, so the photon-proton cross section in the photoproduction regime is given by Equation 2.35 with $\sigma_{\mathrm{L}}^{\gamma p} = 0$:

$$\frac{d^2\sigma^{ep}}{dydQ^2} = \frac{\alpha_{EW}}{2\pi}\frac{1}{Q^2}\left[\left(\frac{1+(1-y)^2}{y} - 2\frac{(1-y)}{y}\frac{Q^2_{min}}{Q^2}\right)\sigma_{\mathrm{T}}^{\gamma p}\right]. \tag{2.36}$$

Integrating over Q^2, one obtains the cross section as a function of only y

$$\frac{d\sigma^{ep}}{dy} = \frac{\alpha_{EW}}{2\pi}\left[\frac{1+(1-y)^2}{y}\ln\left(\frac{Q^2_{max}}{Q^2_{min}}\right) - \frac{2(1-y)}{y}\left(1-\frac{Q^2_{max}}{Q^2_{min}}\right)\right] \tag{2.37}$$

where Q^2_{min} and Q^2_{max} define the Q^2 range. This expression is equivalent to the virtual photon flux.

2.3.5 Factorization

It is not possible to calculate the entire lepton-proton scattering process using only perturbative methods. The incoming proton and photon, when resolved into hadrons, each contain a distribution of partons which partake in soft (long range with large α_s) interactions separate from the hard (short range with small α_s) scatter. Through a calculational method called **factorization**, it is possible to separate hard and soft terms. The soft terms are absorbed into the **Parton Distribution Functions** (PDFs), which are based on experimentally determined quantities, are process independent, and give the probability of finding a parton with a certain x and Q^2. Factorization introduces an additional scale, μ_F, which sets a limit below which pQCD is not valid. For LO processes, the structure functions and PDFs are related by

$$F_1(x,Q^2) = \frac{1}{2}\sum_i e_i^2 f_i(x,Q^2) \tag{2.38}$$

$$F_2(x,Q^2) = \sum_i e_i^2 x f_i(x,Q^2), \tag{2.39}$$

where the sum is over the partons in the hadron, e_i is the parton charge, and $f_i(x,Q^2)$ is the PDF. The PDFs for NLO processes are dependent on the renormalization scheme, which handles the divergent integrals present in calculations of these processes and introduces another scale, μ_R. There are several renormalization schemes, the most important being the Minimal Subtraction ($\overline{\mathrm{MS}}$) and Deep Inelastic Scattering (DIS) Schemes.

The hard term separated by factorization has a perturbatively calculable dependence on Q^2, called **parton evolution**. The parton evolution equations extrapolate the distributions provided by the PDFs at a certain x_0 and Q_0^2 over a large range of x

and Q^2.

2.4 Photoproduction

Photoproduction is the kinematic region in which $Q^2 \sim 0$. The lepton scattering angle is very small and the exchanged photon has a low virtuality, which means that the photon cannot resolve the internal structure of the proton as it does in DIS. Instead of being viewed as a process in which a photon probes the proton, photoproduction can be interpreted as a process in which hard partons from the proton probe the photon.

The low Q^2 of the exchange photon also means that it cannot be used as a hard scale for perturbative calculations. If present, a large momentum exchange between a parton in the proton and the photon, or a parton in the photon, is used as the hard scale. This large momentum exchange is indirectly observed through the high transverse energy (E_T) of outgoing partons, and in practice the E_T of these particles is used as the hard scale.

There are three different regions of photoproduction, as shown in Fig. 2.7, which are dependent on both Q^2 and the hard scale of the interaction. The time the photon has to fluctuate into a hadronic object varies approximately as $t_\gamma \sim \frac{E_\gamma}{Q^2}$. For very low Q^2, less than a certain scale k_0^2, the photon has enough time to fluctuate into a bound state having the same quantum numbers as the photon. These bound states are vector mesons and the process is described by the Vector Meson Dominance Model (VDM). The photon in this kinematic regime is also termed **VDM**. At larger Q^2, when the photon virtuality is smaller than the hard scale, the photon has time to fluctuate into an unbound $q\overline{q}$ pair. The photon is described as **anomalous** in this region. If the photon virtuality is larger than the hard scale, the photon does not have time to fluctuate into an hadronic system and interacts as a point-like object with the proton. The photon in this region is called **direct** or **bare**. Diagrams of each type of photon are shown in Fig. 2.8. There are no exact boundaries between the regions and transitions from one to another are smooth.

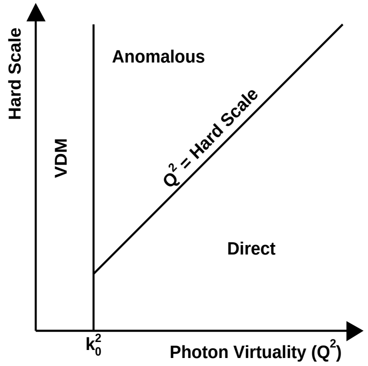


Figure 2.7: Photon structure in photoproduction for different regions of Q^2-hard scale phase space. Q^2 is the virtuality of the photon and k_0 is defined as the scale in which the transition between the VDM and direct/anomalous regions occur. The transitions between different regions are actually smooth.

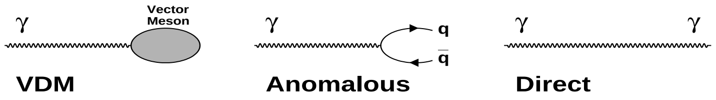


Figure 2.8: Diagrams of VDM (left), anomalous (middle), and direct (right) photons in photoproduction.

2.4.1 The VDM Region

For very low virtualities, the photon has enough time to fluctuate into an $q\overline{q}$ pair surrounded by soft gluons and form an hadronic bound state with the same quantum numbers as the photon. The bound states correspond to light vector mesons ($V = \rho^0, \omega, \phi, J/\psi \ldots$) and are explained by the non-perturbative Vector Meson Dominance Model (VDM) [3, 4], which describes the bound state as a superposition of the mesons. Higher mass vector mesons have been included in the General Vector Meson Dominance Model (GVDM) [5].

The lifetime of the fluctuation into a vector meson can be derived from the Uncertainty Principle, $\Delta t_V \approx \frac{\hbar}{\Delta E}$, where $\Delta E = E_V - E_\gamma = \sqrt{\boldsymbol{p}_V^2 + m_V^2} - E_\gamma$. Substituting $Q^2 = -q^2 = \boldsymbol{q}^2 - E_\gamma^2 = \boldsymbol{p_V}^2 - E_\gamma^2$, where $\boldsymbol{p_V} = \boldsymbol{q}$ because of three-momentum conservation, one obtains

$$\Delta E = \sqrt{E_\gamma^2 + Q^2 + m_V^2} - E_\gamma. \tag{2.40}$$

Assuming a very small Q^2 and that the mass of the vector meson is much smaller than its energy so that $E_\gamma \gg Q^2 + m_V^2$, the square root term can be expanded giving

$$\Delta E \approx \frac{Q^2 + m_V^2}{2E_\gamma}. \tag{2.41}$$

Therefore, the lifetime of the fluctuation is:

$$t_V \approx \frac{2\hbar E_\gamma}{Q^2 + m_V^2} \tag{2.42}$$

It can be seen that the lifetime varies as $\frac{1}{Q^2}$ and fluctuations into vector mesons are therefore highly suppressed for large Q^2. However, in the center-of-mass frame of the proton, $x = \frac{Q^2}{m_p} E_\gamma$ and $E_\gamma = \frac{p \cdot q}{m_p}$ which means that the lifetime also varies as $t_V \sim \frac{1}{x}$.

This means that VDM processes can be relevant at high Q^2 as long as x is small. In this situation, high energy partons from the proton probing the vector meson provide a hard scale and enable the application of pQCD.

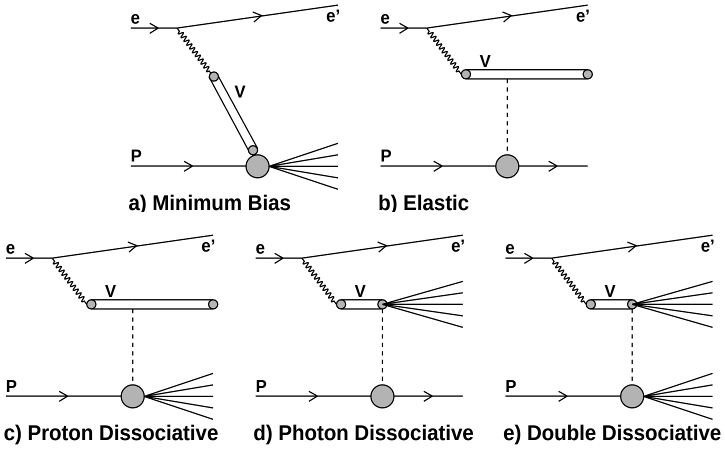
The dominant VDM processes are shown in Fig. 2.9. In a Minimum Bias event, depicted in Fig. 2.9a, the vector meson and proton interact in such a way that it is impossible to distinguish which final state particles came from the proton and which from the vector meson. In an Elastic process, shown in Fig. 2.9b, the interacting particles do not disassociate and an intact vector meson and proton are present in the final state. In Proton Dissociation, shown in Fig. 2.9c, the vector meson remains intact but the proton dissociates, and in Photon Dissociation, shown in Fig. 2.9d, the vector meson dissociates but the proton remains intact. In Double Dissociation, shown in Fig. 2.9e, both the vector meson and proton dissociate.

2.4.2 The Anomalous Region

In the anomalous region, Q^2 is too large to allow the photon time to fluctuate into a hadronic bound state, but small enough in relation to the hard scale that the photon can fluctuate into an unbound $q\overline{q}$ pair. One of the quarks, or a gluon which is radiated by one of the quarks, can interact with a parton in the proton. Large transverse momentum at the $\gamma \to q\overline{q}$ vertex means that, in principle, the anomalous photon can be calculated using pQCD. The six LO anomalous processes are shown in Fig. 2.10. Processes related by crossing symmetry[4] can also occur but are not shown.

---

[4] Crossing symmetry states if an interaction such as $A + B \to C + D$ can occur then so can $A + \overline{D} \to C + \overline{B}$, where particle change to their anti-particles when moved to the other side of the arrow. Therefore, $qg \to gq$ is actually the same as $q\overline{q} \to gg$ even though the appear different in the laboratory.



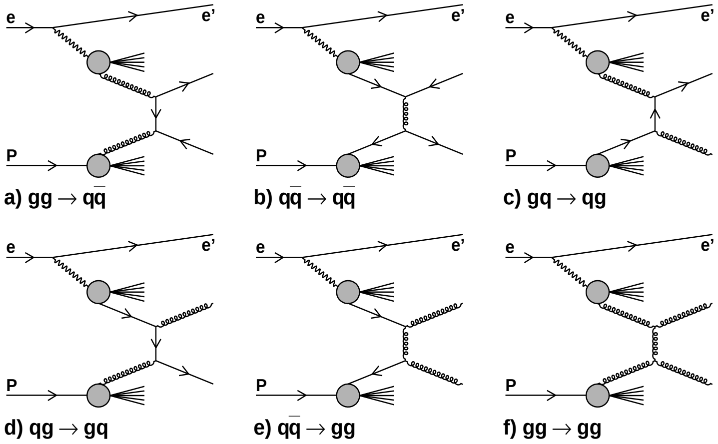
a) Minimum Bias b) Elastic

c) Proton Dissociative d) Photon Dissociative e) Double Dissociative

Figure 2.9: Diagrams of the vector meson photoproductive processes described by GVDM.

2.4.3 The Direct Region

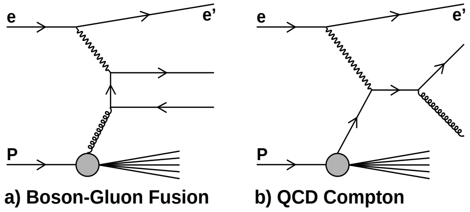
In the direct region, Q^2 is too large to allow the photon time to fluctuate into a bound state and too large in relation to the hard scale to allow the photon to fluctuate into an unbound $q\overline{q}$ pair. Therefore, the photon behaves like a point-like object and transfers all of its energy to the partons in the proton. The two possible LO direct interactions are shown in Fig. 2.11. In boson-gluon fusion, depicted in Fig. 2.11a, a gluon from the proton splits into an $q\overline{q}$ pair and one of these quarks interacts directly with the photon. In a QCD Compton process, shown in Fig. 2.11b, a quark from the proton interacts directly with the photon and then radiates a gluon.



a) $gg \rightarrow q\overline{q}$ b) $q\overline{q} \rightarrow q\overline{q}$ c) $gq \rightarrow qg$

d) $qg \rightarrow gq$ e) $q\overline{q} \rightarrow gg$ f) $gg \rightarrow gg$

Figure 2.10: The LO anomalous photoproduction processes. Note that the processes related by crossing symmetry occur, but are not shown.



a) Boson-Gluon Fusion b) QCD Compton

Figure 2.11: The LO direct photoproduction processes. On the left is Boson-Gluon Fusion and on the right is QCD Compton.

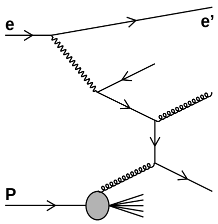


Figure 2.12: An event which could be interpreted as NLO Direct or LO Resolved

.

2.4.4 The Photoproduction Cross Section

The total photon-proton cross section can be written as $\sigma^{\gamma p}_{\text{Tot}} = \sigma^{\gamma p}_{\text{VDM}} + \sigma^{\gamma p}_{\text{Anomalous}} + \sigma^{\gamma p}_{\text{Direct}}$. The VDM and anomalous components can be combined into a single resolved photon component to give

$$\sigma^{\gamma p}_{\text{Tot}} = \sigma^{\gamma p}_{\text{Direct}} + \sigma^{\gamma p}_{\text{Resolved}}. \tag{2.43}$$

Whether a process is classified as direct or resolved is determined by the amount of the photon's momentum involved in the interaction with the partons in the proton. The fraction of the photon's momentum which is involved in the interaction is given by the variable x_γ, so that for direct interactions $x_\gamma = 1$ and for resolved interaction $x_\gamma < 1$. In practice, the boundary between direct and resolved is not clearly defined and is usually taken to be $x_\gamma = 0.75$. The terms direct and resolved are only valid at LO since an ambiguity arises between a NLO direct and a LO resolved event. This is demonstrated in Fig. 2.12, which shows an event that could be interpreted as either. Also, only the sum of direct and resolved processes are physically meaningful.

2.5 Diffraction

2.5.1 Introduction to Diffraction

There are three classifications of interactions in hadron-hadron collisions based upon the characteristics of their final states. In **elastic** interactions, both incoming hadrons remain intact and no additional final state particles are produced. In **diffractive** interactions, the momentum transfer between the incoming hadrons is small, but one (single dissociation) or both (double disassociation) hadrons dissociate into a multi-particle final state. The final state particles preserve the quantum numbers of the initial state, and therefore the exchange particle has the quantum numbers of the vacuum. Diffractive events contribute significantly (25-40%) to the total cross section for hadron-hadron interactions. The remaining events, classified as **inelastic**, are characterized by an exchange of quantum numbers and the dissociation of the final states.

Although diffraction has no precise definition, there are two descriptions which are generally used. The first is that a diffractive event is one in which the final state particles preserve the quantum numbers of the associated initial state particles. The second, attributed to Bjorken [6], is that "A process is diffractive if and only if there is a large rapidity gap in the produced particle phase space which is not exponentially suppressed". Rapidity is an angular quantity related to the polar angle and described in Section 2.5.2.

Since the momentum transfer in diffraction is small, there is no hard scale present at the vertices between the incoming and outgoing hadrons. If there is no other hard scale present, the process is classified as **soft diffraction** and non-perturbative methods, such as Regge theory, must be used to describe the interaction. However, if another hard scale is present, for example a large enough Q^2, pQCD can be used to describe the interaction and the process is classified as **hard diffraction**.

2.5.2 Rapidity and Pseudorapidity

It is possible to write the four-momentum of a particle in terms of its transverse momentum, p_{T}, rapidity, y, and azimuthal angle, ϕ, as

$$p^\mu = (E, p_x, p_y, p_z) = (m_T \cosh y, p_T \sin\phi, p_T \cos\phi, m_T \sinh y), \tag{2.44}$$

where $m_T = \sqrt{p_T^2 + m^2}$ and $p_T = \sqrt{p_x^2 + p_y^2}$. The rapidity[5] is then given by the expression

$$y = \frac{1}{2} \ln\left(\frac{E + p_z}{E - p_z}\right). \tag{2.45}$$

The advantages of this notation are that the distribution of final state particles as a function of y is fairly uniform for inelastic interactions and that differences in rapidity (but not rapidity itself) are boost invariant for a boost in the z (longitudinal) direction. **Boost invariant** means that a quantity is unchanged when it undergoes a Lorentz transformation. The accessible range in y is determined by the energy of the colliding particles and the rest mass of all particles participating in the interaction.

Pseudorapidity[6] is defined as the rapidity in the limit that $m \to 0$. In this limit (and using the variable η instead of y),

$$p_T \approx E_T = E \sin\theta \tag{2.46}$$

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right), \tag{2.47}$$

---

[5] The full derivation of rapidity is given in Appendix A.1.1.

[6] The full derivation of pseudorapidity is given in Appendix A.1.2.

where $\theta = \frac{E_T}{p_z}$. The variables (E_T, η, ϕ) associated with pseudorapidity, unlike the variables (p_T, y, ϕ) associated with rapidity, are those measured in experiments. Therefore, pseudorapidity is the more commonly used quantity.

2.5.3 Kinematics of Diffraction

The kinematic variables for diffraction in *ep* collisions are introduced in Fig. 2.13. A lepton radiates a photon, which interacts with a Pomeron emitted by the proton. The **Pomeron** ($\mathbb{P}$) carries the quantum numbers of the vacuum and is used to characterize the propagator function in diffractive interactions where the initial and final states have the same quantum numbers. The lack of color flow results in the hadronic final state and the final state proton being separated by a gap in rapidity. This is contrary to inelastic interactions in which the particle distribution is constant as a function of rapidity as a result of a propagator with color charge radiating colored particles.

The usual DIS variables apply in diffraction:

$$Q^2 = -q^2 = -(k - k')^2 \tag{2.48}$$

$$W^2 = (q + p)^2 \tag{2.49}$$

$$x = \frac{Q^2}{2p \cdot q} \tag{2.50}$$

$$y = \frac{p \cdot q}{p \cdot k}, \tag{2.51}$$

and in addition, there are variables specific to diffractive processes,

$$t = (p - p')^2 \tag{2.52}$$

$$x_{\mathbb{P}} = \frac{q \cdot (p - p')}{q \cdot p} = \frac{M_X^2 + Q^2 - t}{W^2 + Q^2 - m_p} \tag{2.53}$$

$$\beta = \frac{Q^2}{2q \cdot (P - P')} = \frac{x}{x_{\mathbb{P}}} = \frac{Q^2}{Q^2 + M_X^2 - t}. \tag{2.54}$$

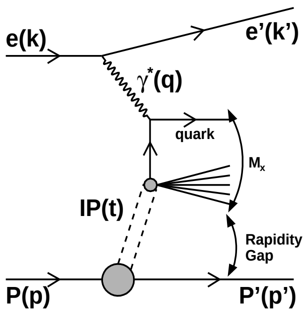


Figure 2.13: Diagram of diffractive kinematics in *ep* collisions.

The invariant mass squared of the hadron final state is given by $M_X^2 = (q + p - p')^2$. The square of the momentum difference between the initial and final state proton is given by t. The fraction of the proton's momentum carried by the Pomeron is given by $x_{\mathbb{P}}$ and the fraction of the Pomeron's momentum involved in the interaction with the photon is given by β. The fraction of the proton's momentum carried by the quark in the proton involved in the interaction is related to these quantities by

$$x = x_{\mathbb{P}}\beta. \tag{2.55}$$

2.5.4 Properties of Diffraction

The Diffractive Peak

The diffractive cross section has a pronounced forward peak and falls off rapidly when moving away from the forward direction. Kinematically, this corresponds to an exponential decrease in the cross section as t becomes more negative. The cross section

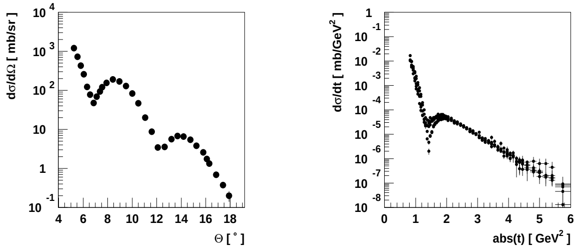


Figure 2.14: Cross sections for diffractive processes. The plot on the left is for proton-calcium scattering as a function of θ and the plot on the right is for proton-proton scattering as a function of t.

can be empirically parameterized as

$$\frac{d\sigma}{dt} = \left(\frac{d\sigma}{dt}\right)_{t=0} e^{bt} \approx \left(\frac{d\sigma}{dt}\right)_{t=0} \left[1 - b(p\theta)^2\right] \tag{2.56}$$

with the slope parameter $b = R^2/4$, where R is the radius of interaction and generally taken to be about 1 fm, and t is negative. The sharpness of the forward peak increases slowly with $\sqrt{s}$, a property known as **shrinkage**.

Two examples of diffractive cross sections can be seen in Fig. 2.14. The plot on the left [7], clearly exhibiting the diffractive peak and secondary maximums, is for proton-calcium scattering as a function of θ and the plot on the right [8] is proton-proton scattering as a function of t. The shapes of these distributions are similar to those observed in optical diffraction, and this similarity is the origin of the application of the term diffraction to particle physics. In optical diffraction, light passes through a circular aperture and emerges with a distribution of intensity described by

$$\frac{I}{I_0} = \left[\frac{\sin(\frac{kR}{2}\sin\theta)}{\frac{kR}{2}\sin\theta}\right]^2 \approx 1 - \frac{R^2}{4}(k\theta)^2, \tag{2.57}$$

where k is the wave number and R is the radius of the aperture. A sketch of this

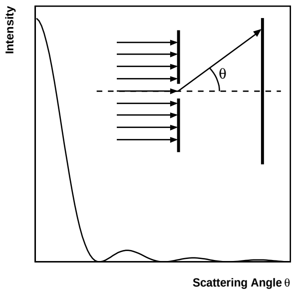


Figure 2.15: Diffraction pattern of light passing through a circular aperture. The similarity in shape with the cross section of a certain class of hadron-hadron collisions is the origin of the term diffraction.

pattern is shown in Fig. 2.15.

The Rapidity Gap

Because of the large forward peak, the diffractive final state is separated in rapidity from the proton. This separation in rapidity can be easily estimated for a single diffractive γp process in which the proton does not dissociate and $t = 0$. In the γp center of mass system, the outgoing proton and hadronic system X move in opposite directions with longitudinal momentum $p_z = W/2$. Therefore, the rapidities of the systems are given by

$$y_P \approx \frac{1}{2}\ln\left(\frac{E_P + p_{z_P}}{E_P - p_{z_P}}\right) = \frac{1}{2}\ln\frac{W^2}{m_p^2} \tag{2.58}$$

$$y_X \approx \frac{1}{2}\ln\left(\frac{E_X + p_{z_X}}{E_X - p_{z_X}}\right) = \frac{1}{2}\ln\frac{W^2}{M_X^2}, \tag{2.59}$$

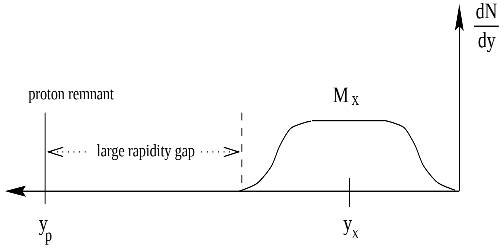


Figure 2.16: Sketch of the cross section as a function of rapidity for a diffractive event. There is a large rapidity gap between the hadronic system (M_X) and the proton remnant.

where $E = \sqrt{p^2 + m^2} = \frac{W}{2}\sqrt{1 + \left(\frac{2m}{W}\right)^2}$ has been expanded. The rapidity gap between the proton and hadronic system is then given by

$$\Delta y = y_P - y_X \approx \ln\left(\frac{W^2}{m_p M_X}\right). \tag{2.60}$$

For the typical values of $W = 200\,\mathrm{GeV}$ and $M_X = 20\,\mathrm{GeV}$, there is a rapidity gap $\Delta y \approx 7.7$. However, since the hadronic system X dissociates, some of these hadrons will spray into the gap and decrease it by an amount $\Delta y \sim \ln M_X \approx 3$. The proton and hadronic system will therefore be separated by a rapidity gap larger than 4 units. For the general case of the interaction, there are also soft hadron-hadron interactions which decrease the rapidity gap, but a gap of $\Delta y > 2$ is still observed. This is pictured in Fig. 2.16, which shows a sketch of the cross section as a function of rapidity for a diffractive event. There is a large rapidity gap between the hadronic system and the proton remnant.

2.5.5 Regge Theory

The momentum transfer at the proton vertex (t) in diffractive processes is small and cannot be used as a hard scale. Perturbative methods are not valid in this regime, and the Regge theory of the strong interaction is used instead of QCD to describe the diffractive hadron-hadron cross section. **Regge theory** was developed in the 1960's, before QCD, with the aim of describing the asymptotic behavior of the scattering amplitude of hadron-hadron collisions in the limit that the center-of-mass energy (s) becomes large.

The Hadron-Hadron Cross Section

The cross section of a two-body elastic interaction, $A + B \rightarrow C + D$, in the s-channel with $s > 0$ and $t < 0$, as shown on the left side of Fig. 2.17, can be written as

$$\frac{d\sigma}{dt} = \frac{1}{64p^2 s}|A(s,t)|^2 \tag{2.61}$$

where p is the momentum of the proton and $A(s,t)$ is the scattering amplitude. If $A(s,t)$ is an analytic function of s and t, the crossed channel of the interaction, as shown on the right side of Fig. 2.17, with $t > 0$ and $s < 0$, allows the scattering amplitude to be written as an expansion of partial waves

$$A(s,t) = \sum_{l=0}^{\infty}(2l+1)f_l(t)P_l\cos(\theta), \tag{2.62}$$

where $P_l\cos(\theta)$ is the Legendre polynomial of order l, θ is the scattering angle between particles A and C, $f_l(t)$ is the l^{th} partial wave amplitude, and $t = (p_A + p_{\overline{C}})^2$. Using the Sommerfeld-Watson [9] transformation, the partial wave amplitude can be written as a contour integral in the complex plane with poles

$$f_l(t) \sim \frac{\beta(t)}{l - \alpha(t)}. \tag{2.63}$$

A **Regge pole**, with angular momentum $l = \alpha(t)$, is identified as a particle resonance, and $\beta(t)$ is the residue function which describes the coupling of the Regge pole particle

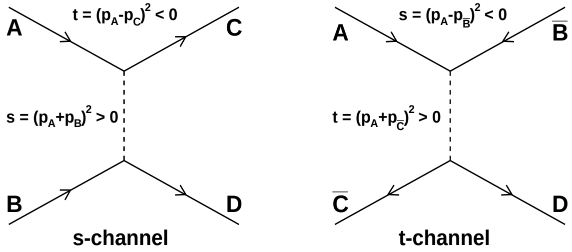


Figure 2.17: Diagram of two body scattering in the s and t channels. Note that $s > 0$ and $t < 0$ in the s channel and $s < 0$ and $t > 0$ in the t channel.

to external particles. Going back to the s-channel, where $s > 0$ and $t < 0$, the scattering amplitude in the limit that $s \rightarrow \infty$ and t is constant can be written as

$$A(s,t) \sim \beta(t)\left(\frac{s}{s_0}\right)^{\alpha(t)}, \tag{2.64}$$

where $s_0 \approx 1\,\mathrm{GeV}$ is the hadronic mass scale.

Unlike in QCD, where individual particles are exchanged in an interaction, trajectories of particles are exchanged in Regge Theory. A **Regge Trajectory** describes group of particles with the same quantum numbers but different spin. The trajectories can be written as an expansion of the angular momentum $\alpha(t)$:

$$\alpha(t) = \alpha_0 + \alpha' t, \tag{2.65}$$

where α_0 is the intercept and α' is the slope. An example of a Regge trajectory for the ρ, ω, f, and a mesons is shown in Fig. 2.18. This trajectory has the parameters $\alpha(0) = 0.55$ and $\alpha' = 0.86\,\mathrm{GeV}^2$ and is extrapolated to negative t based upon the results from $\pi^- p \rightarrow \pi^0 n$ scattering [10].

Substituting the scattering amplitude $A(s,t) \sim s^{\alpha_0 + \alpha' t}$ into Equation 2.61, where

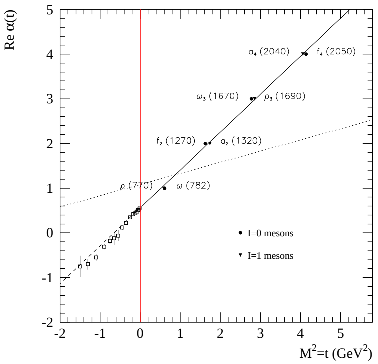


Figure 2.18: The solid line shows a fit to Equation 2.65 for the ρ, ω, f, and a mesons. The open squares in the region $t < 0$ represent data from $\pi^- p \to \pi^0 n$ scattering [10] and the dashed line is an extension of the fit to these points. The dotted line is the trajectory corresponding to Pomeron exchange.

the high energy approximation $s \approx 4p^2$ has been used, one obtains

$$\frac{d\sigma}{dt} = \frac{1}{16\pi s^2}|A(s,t)|^2 \sim \frac{1}{s^2}\left(\frac{s}{s_0}\right)^{2\alpha_0 + 2\alpha' t}, \tag{2.66}$$

which can be rewritten as

$$\frac{d\sigma}{dt} \sim \frac{1}{s_0^2}\left(\frac{s}{s_0}\right)^{2(\alpha_0 - 1)} e^{2\alpha' \ln(s/s_0) t}. \tag{2.67}$$

In the forward direction $t = 0$, and moving away from the forward direction causes t to grow negatively, which causes $\frac{d\sigma}{dt}$ to decrease exponentially. This explains the forward peak in diffractive scattering. Comparing Equation 2.67 to Equation 2.56 shows that

$$b = b_0 + 2\alpha' \ln\left(\frac{s}{s_0}\right). \tag{2.68}$$

Since t is negative, the exponential falloff of the diffractive peak decreases with increasing s [11]. This explains the property of shrinkage.

The Need for a Pomeron

The **optical theorem** relates the imaginary part of the forward scattering amplitude to the total (elastic + diffractive +inelastic) cross section

$$\sigma_{Tot}(s) \sim \frac{1}{s}\mathrm{Im}A(s, t = 0) \sim s^{\alpha_0 - 1}, \tag{2.69}$$

where α_0 is the intercept of a Regge trajectory. Experimentally, it has been shown that the total cross section decreases until $\sqrt{s} \approx 10\,\mathrm{GeV}$ and then rises slowly with increasing[7] $\sqrt{s}$. However, all Regge trajectories fit to known particles have an intercept $\alpha_0 < 1$ and therefore predict a cross section which can only decrease as a function of s. In order to describe the rise in the cross section at large s it was proposed by Gribov [13] to introduce a particle called a **Pomeron** ($\mathbb{P}$), in honor of Pomeranchuk, with $\alpha_0 > 1$. The Pomeron trajectory is shown as the dotted line in Fig. 2.18 with the parameters $\alpha_0 = 1.08$ and $\alpha' = 0.25\,\mathrm{GeV}^{-2}$.

The total cross section can be parameterized into contributions from the Pomeron and contributions from all other Regge trajectories represented by one average trajectory called a **reggeon** ($\mathbb{R}$)

$$\sigma_{Tot}(s) = As^{(\alpha_{\mathbb{R}} - 1)} + Bs^{(\alpha_{\mathbb{P}} - 1)}. \tag{2.70}$$

While the Pomeron represents the exchange of vacuum quantum numbers (0 electric charge, no color, and isospin 0), the reggeon represents the exchange of non-vacuum quantum numbers. Donnachie and Landschoff [14] have fit Equation 2.70 to pp and

---

[7] The rise is limited by the **Froissart Bound** [12], which states $\sigma_{Tot}(s) \leq \frac{\pi}{m_\pi^2}(\ln s)^2$.

$p\bar{p}$ data, as shown in Fig. 2.19 and obtained the intercepts

$$\alpha_{0,\mathbb{R}} = 0.4525$$

$$\alpha_{0,\mathbb{P}} = 1.0808.$$

The Pomeron couples equally to particles and anti-particles because it has the quantum numbers of the vacuum, and therefore the coefficient A in Equation 2.70 is the same for pp and $p\bar{p}$ collisions. The reggeon couples differently to particles and anti-particles, and is reflected in the difference in the value of the coefficient B for each case.

Data from $\pi^{\pm}p$ and γp total cross sections have also been fit to Equation 2.70 and also show the need for a Pomeron. Instead of being viewed as an actual particle, the Pomeron serves as a construct to characterize interactions with the properties needed to describe various sets of data. This suggests that the Pomeron is a universal object.

2.5.6 Hard Diffraction

Hard diffractive events are those in which there is a large rapidity gap in the hadronic final state and at least one hard scale which allows the application of pQCD. Examples of this hard scale are a large Q^2 in diffractive DIS events, as shown in Fig. 2.13, and a large momentum transfer across the rapidity gap. There exists an interplay between hard and soft physics in hard diffractive events, and their separation is necessary to explore QCD at both a quantitative and qualitative level.

The Pomeron in QCD

Regge theory gives no insight into the structure of the Pomeron. In QCD, the structure of the Pomeron is still unknown, but there are several models which describe a particle

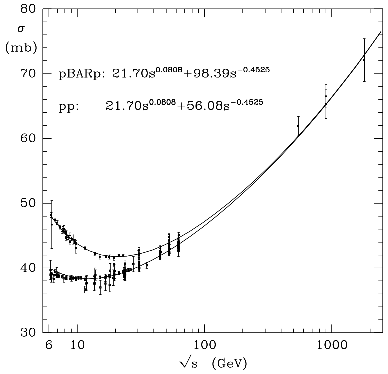


Figure 2.19: A compilation of data showing the total pp and $p\bar{p}$ cross sections as a function of $\sqrt{s}$. The data samples for each cross sections are fit to $\sigma_{Tot}(s) = As^{(\alpha_{\mathbb{R}}-1)} + Bs^{(\alpha_{\mathbb{P}}-1)}$. The fit coefficients are shown in the text in the plot.

with the same properties as the Pomeron predicted by Regge theory. These range from a simple two gluon exchange [15, 16, 17] to more complicated models involving a multi-gluon ladder whose properties depend on the nature of the interaction. This points to a non-universal Pomeron in QCD.

The **BFKL**[8] **Pomeron** [18, 19, 20] is a sum of all gluon ladder diagrams and has an intercept $\alpha_0 \approx 1.3 - 1.5$ [21]. The gluon ladder in the BFKL Pomeron consists of a special radiation pattern in which the gluons are ordered strongly in x but not ordered in k_T.

---

[8] Also known as a "hard Pomeron" or "perturbative Pomeron"

2.5.7 Hard Diffraction in Photoproduction

Two hard diffractive hadron-hadron interactions, which can be observed in lepton-proton photoproduction, have been suggested to study the nature of the Pomeron. Since $Q^2 \approx 0$ in photoproduction, it can not be used as the hard scale. The first process, shown on the left in Fig. 2.20, and proposed by Ingelman and Schlein [22], is characterized by an outgoing parton with high transverse momentum separated in rapidity from one of the outgoing original hadrons. In this process, t is small and the hard scale is in the exchange between the Pomeron and the quark from the photon. The second process, shown on the right in Fig. 2.20 and proposed by Bjorken [23], is characterized by two final state partons separated by a rapidity gap. In this process, the diffractive exchange takes place between the two outgoing partons, but t is large and can be used as the hard scale. Since t and the transverse momentum of the outgoing hadrons, p_T, are related by

$$|t| \geq p_T^2, \tag{2.71}$$

p_T of the outgoing hadrons can be used as a selection criteria for these events. This relation is proven in Appendix A.2.

This analysis will concentrate on the study of hard diffractive photoproduction events with large t and the rapidity gap between two final state hadrons with high transverse momentum. These events are sometimes called **gaps between jets**[9] and the analysis of these events is sometimes referred to as the study of **interjet energy flow**.

---

[9] A jet is composed of final state hadrons observed in a detector and will be described in subsequent chapters.

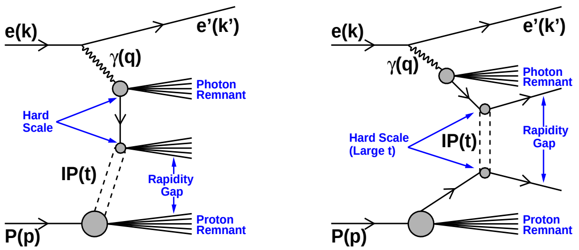


Figure 2.20: Diagrams of hard diffractive photoproduction. The diagram on the left has the rapidity gap between the proton remnant and a final state hadron. The diagram on the right has the rapidity gap between two jets of high transverse momentum.

2.6 Gaps Between Jets in Hard Diffractive Photoproduction

2.6.1 Color Connection and Rapidity Gaps

Scattering in high energy hadronic collisions is dominated by QCD processes involving quarks and gluons. These particles carry color and radiate color charge as they accelerate. Similar to an accelerated electric charge radiating a photon, an accelerated color charge will radiate soft gluons. When a color-carrying particle is exchanged, the acceleration occurs as it travels through an angular path. Two partons connected by a quark or gluon are said to be color-connected.

It is possible that color-connected outgoing partons are separated by a large distance in rapidity, and the only way rapidity gaps can appear between them is by random fluctuations in particle multiplicity. This behavior, pictured on the left side of Fig. 2.21, can be behavior described by

$$\frac{dN}{dy} = -PN, \tag{2.72}$$

where N is the number of events with a rapidity gap and P is the probability that a particle is radiated into a rapidity interval dy. The probability can be assumed to be constant for events described by QCD. Solving for N, one obtains

$$N = N_0 e^{-P\Delta y}, \tag{2.73}$$

which shows an exponential decrease in the number of events with a rapidity gap as the size of the rapidity gap increases.

A **color-singlet** propagator, which has no color, does not radiate color charge into the rapidity gap. The outgoing hadrons on either side of the gap are not color-connected to each other, but are instead color-connected to the remnants closest in rapidity. The lack of radiation from the progagator means that the number of events with rapidity gaps due to color-singlet exchange are constant as a function of rapidity.

Color-singlet exchange in hard diffractive photoproduction is shown on the right side of Fig. 2.21. In this situation, the color-singlet propagator may be a Pomeron or an electro-weak gauge boson. However, the contribution from electro-weak exchange is expected to be small and not contribute significantly to the cross section.

2.6.2 The Gap Fraction

The **gap fraction**, f, provides a convenient method of measuring the rate of color singlet exchange in hard diffractive photoproduction. The gap fraction is defined as

$$f \equiv \frac{\sigma_{\mathrm{Gap}}}{\sigma}. \tag{2.74}$$

The inclusive cross section, σ, is the cross section for all events containing two final state hadrons with high transverse energy separated in rapidity, regardless of the amount of transverse energy of the particles between them. The gap cross section, σ_{Gap}, is the cross section for all events containing two final state hadrons with high

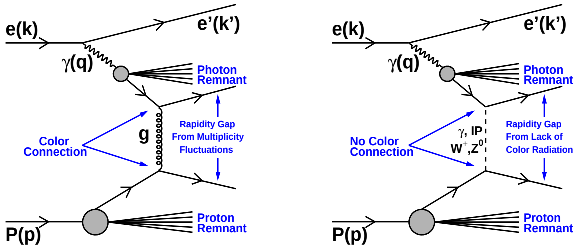


Figure 2.21: Diagrams of processes which produce gaps between jets in photoproduction. The diagram on the left shows color non-singlet exchange between the final state hadrons. If there is an empty gap, it is caused by fluctuations in particle multiplicity. The frequency of these gaps decreases exponentially as a function of rapidity. The diagram on the right shows color singlet exchange. The empty gap in these events are due to lack of color radiation in the gap between the outgoing hadrons.

transverse energy separated in rapidity and with only a few GeV of transverse energy between them. σ_{Gap} includes contributions from both color non-singlet and color singlet exchange.

$$\sigma_{\mathrm{Gap}} = \sigma_{\mathrm{Gap}}^{\mathrm{singlet}} + \sigma_{\mathrm{Gap}}^{\mathrm{non-singlet}}, \tag{2.75}$$

where the contribution from non-singlet exchange is due to random fluctuations in particle fluctuations and the contribution from color singlet exchange is due to the Pomeron propagator. The ratio of σ_{Gap} to σ is an ideal quantity to measure from both a theoretical and experimental viewpoint. Many uncertainties in the measurement of the cross sections and dependencies on the kinematics of hadronization should cancel in the ratio.

The color non-singlet contribution to the gap fraction is expected to fall off exponentially with increasing rapidity separation between the two high-energy outgoing partons, as explained in Section 2.6.1, but the color singlet contribution is expected to

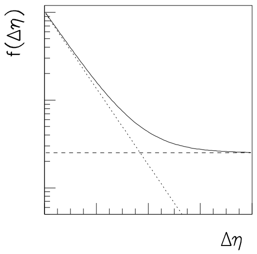


Figure 2.22: A prediction of the $f(\Delta\eta)$. The dotted line shows the exponential decrease of the color non-singlet f as a function of $\Delta\eta$, the dashed line shows the constant contribution of the color singlet contribution, and the solid line shows the convolution of both contributions.

be constant as a function of rapidity separation. The combination of these contributions is a convolution of both individual behaviors. These behaviors are illustrated in Fig. 2.22, which shows predictions for the gap fraction as a function of pseudorapidity.

2.6.3 Gap Survival

The measurement of rapidity gaps is complicated by the possibility of their destruction by other processes. One method of destruction is from soft gluon radiation during hadronization which can spill into the gap. A gap may also be destroyed by the **underlying event**, which is any interaction taking place in addition to the hard scatter. In particular, **multi-parton interactions** (MPI), the scattering of partons in the proton and photon remnants, can destroy the gap since they tend to be color mediated processes. These effects are non-perturbative and, although very difficult to model, can be accounted for by the **gap survival probability**, $\mathcal{S}$. The measurable gap fraction, $f(\Delta\eta)$, can then be related to the gap fraction produced at parton level,

$\hat{f}(\Delta\eta)$, by

$$f(\Delta\eta) = \mathcal{S}\hat{f}(\Delta\eta). \tag{2.76}$$

It is possible to estimate $\mathcal{S}$ by comparing simulated events at the hadron and parton levels.

At lowest order, QCD color singlet exchange is due to a Pomeron composed of two gluons. However, it has been shown in NLO calculations [24] that in this situation radiation from soft gluons is suppressed in the central rapidity region. In fact, the radiation pattern of a composite gluon object is the same as for an exchanged photon.

Chapter 3

HERA and the ZEUS Experiment

3.1 The DESY Laboratory

The Deutsches Elektronen-Synchrotron (DESY) was established in 1959 with the signing of a State Treaty by the mayor of Hamburg and the German Minister of Nuclear Energy. Its name was taken from the first accelerator, an electron synchrotron which was constructed starting in 1960 and called "DESY". DESY is a publicly funded national research center based in Hamburg with a second site in Zeuthen. The purpose of the laboratory is to conduct basic research in the natural sciences with an emphasis on the development, construction, and operation of accelerator facilities, particle physics research, and research using photons.

Today, DESY contains many particle accelerators, with the smaller ones being used to pre-accelerate particles before they are passed on to the larger ones. In increasing size, the most significant accelerators are the liner accelerator (LINAC), DESY, the Positron Elektron Tandem Ring Anlage (PETRA), the Doppel Ring Speicher (DORIS), and the Hadron Elektron Ring Anlage (HERA). Particle physics research is currently centered around HERA and the other rings are sources of synchrotron radiation for research with photons, in addition to preaccelerators for HERA. The Hamburg Synchrotron Radiation Laboratory (HASYLAB) is the DESY installation in which the synchrotron radiation research is coordinated and performed. A theory institute and research programs for the European Free Electron Laser (XFEL) and International Liner Collider (ILC) are also based at DESY. An aerial view of DESY is shown in Fig. 3.1, with the underground location of the HERA and PETRA accelerators marked by the white dashed line and the location of the detectors shown by the small white circles.

Figure 3.1: An aerial view of DESY. The white dashed lines show the underground locations of PETRA and HERA and the small white circles show the location of the detectors.

There are approximately 1550 members of the DESY staff, including 365 permanent scientists. In addition, there are approximately 90 undergraduate students, 450 PhD students, and 240 postdoctoral scientists working at DESY. There are also approximately 3000 scientists from 33 countries using the DESY facilities. The budget for DESY is 160 million Euros per year with the German Federal Government financing 90% of the cost and the states of Hamburg and Brandenburg sharing the remaining 10%.

3.2 The HERA Collider

HERA is the world's first and only lepton-proton collider. It was constructed between between May 1984 and November 1990 at a cost of 700 Million Euros.

3.2.1 The HERA Design

HERA is 6.336 kilometers in circumference and varies in depth between 10 and 25 meters. There are two separate rings in HERA, one for the acceleration of protons and the other for the acceleration of leptons. The rings are not completely circular, but instead consist of four straight sections and four arcs of 90°. At the center of each straight section is a particle detector. H1 is located on the northern (north is towards the left in Fig. 3.1) side of HERA, ZEUS on located the southern side, HERMES on the eastern side, and HERA-B on the western side. H1 and ZEUS are general purpose detectors that study interactions produced by the collisions of lepton and proton beams, HERMES[1] studies the interactions of leptons with a polarized gas-jet target, and HERA-B[2] studies the interaction between particles in the proton beam halo and a fixed wire grid target with the goal of measuring CP violation in the $b\bar{b}$ system.

---

[1] HERMES started taking data in 1995

[2] HERA-B started taking data in 1999 stopped in February 2003

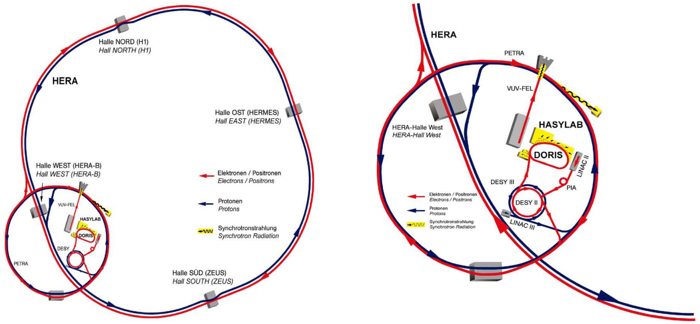


Figure 3.2: The figure on the left shows HERA and its injection system, which is comprised of several smaller accelerators. The figure on the right is a closeup view of the injection system.

3.2.2 The HERA Injection System

Protons and electrons are injected separately into HERA and the injection process consists of several different stages in various accelerators. The injection system is shown in Fig. 3.2.

Proton injection begins with the acceleration of H^- ions to 50 MeV in the H^- LINAC. The ions are passed through a thin foil to strip their electrons and then transfered to the DESY III storage ring, where they are accelerated to 7.5 GeV. The protons in the DESY III ring are grouped into 11 bunches, each consisting of approximately 10^{11} protons, with a 96 ns separation between bunches that is identical to the bunch separation in HERA. The protons are then passed to the PETRA II storage ring and accelerated to 40 GeV. PETRA II can accommodate 70 bunches and also has a 96 ns bunch spacing. The protons are then transfered to HERA where they are accelerated up to 920 GeV by a series of radio frequency (RF) cavities. HERA can hold a maximum of 210 bunches, although not all spaces available for bunches are filled. The proton beam is guided and focused by superconducting magnets which are cooled with liquid helium and create a field of 4.65 Tesla.

Lepton injection begins with the production of either electrons or positrons. Electrons are collected from a hot metal filament and positrons are obtained from e^+e^- pair production caused by bremsstrahlung radiation as an electron passes through a sheet of tungsten. The leptons are grouped into bunches of approximately 3.5×10^{10} and accelerated to 450 MeV by the LINAC II. The leptons are then passed to DESY II where they are accelerated to 7 GeV, and then to PETRA II where they are accelerated further to 14 GeV. PETRA can hold 70 lepton bunches with a 96 ns separation between bunches. The leptons are next transfered to HERA where they are accelerated to 27.52 GeV by a series RF cavities. The RF cavities must continuously supply energy to the leptons in order to compensate for the energy lost through synchrotron radiation. The HERA lepton ring can hold a maximum of 210 bunches, although not all spaces available for bunches are filled. The lepton beam is guided and focused by conventional magnets which produce a field of 0.165 Tesla.

The lepton and proton beams circulate in opposite directions in separate rings which are kept at a vacuum pressure of 3×10^{-11} Torr. Although HERA can potentially have 210 colliding bunches in the lepton and positron beams, with a 96 ns spacing between bunches, not all bunch positions are filled during normal operation. The lepton beam has approximately 18-20 bunches with no counterpart in the proton beam and the proton beam has approximately 3 to 6 bunches with no counterpart in the lepton beam. These unpaired bunches are called **pilot bunches** and are used to

estimate the **beam-gas background**, which is produced from interactions between the beams and residual gas particles in the beam-pipes. Each beam also has approximately 15 empty bunches in a row, a gap which provides enough time to energize the magnets used to deflect the beams when they are dumped.

3.2.3 Luminosity Delivered by HERA

HERA can accelerate both electrons and positrons and has switched between the two over the years. Running with electrons occurred between 1992-94, 1998-99, and 2004-5, while running with positrons occurred between 1995-97, 1999-2000, and 2002-4. The advantage of a positron beam is that its lifetime is longer than an electron beam, which partly due the repulsion between positively charged ions remaining in the beam-pipe and the positrons. In contrast, the electrons attract these positive ions, and pull them into the path of the beam, which increases the interaction rate between the electrons and ions and therefore decreases the lifetime of the electron beam. While this effect was large in the early running at HERA I, modifications to the accelerator have mitigated the differences between electron and positron running.

Between 1992 and 1997, HERA collided 27.52 GeV leptons with 820 GeV protons, giving a center-of-mass energy $\sqrt{s} \approx \sqrt{4E_eE_p} \approx 300\,\mathrm{GeV}$. Since 1998, HERA has been colliding 27.52 GeV leptons with 920 GeV protons, giving a center-of-mass energy $\sqrt{s} \approx 320\,\mathrm{GeV}$.

Between 2000 and 2002 HERA underwent a **luminosity** (the number of collisions per unit area per unit time) upgrade primarily by decreasing the cross sectional area of the colliding beams. In the period before the upgrade, called HERA I, the luminosity reached a peak value of approximately $2 \times 10^{31}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$. In the period

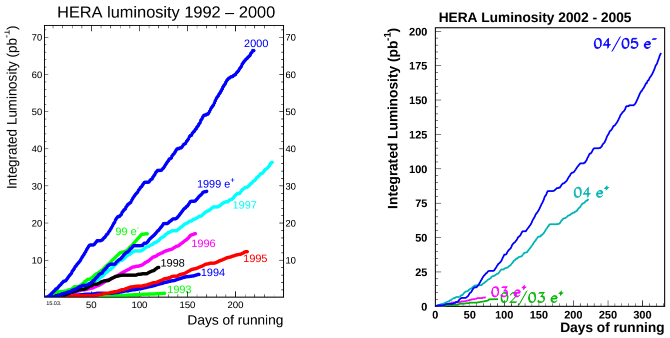


Figure 3.3: The integrated luminosity delivered for each year of running. The figure on the left is for HERA I, which ran between 1992 and 1999. The figure on the right is for HERA II, which has run from 2002. Note the different vertical scales on the figures.

since the upgrade, called HERA II, the luminosity has more than doubled, reaching approximately $4.5 \times 10^{31}\ \mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ in 2005. The integral of the luminosity with respect to time, or **integrated luminosity**, is shown in Fig. 3.3 for HERA I and HERA II, and listed in Table 3.1, for each year of running.

3.3 Detecting Particles

3.3.1 Passage of Particles Through Matter

As particles move through matter, they interact with the matter via the electromagnetic, strong, or weak force and deposit energy. It is possible to detect particles by measuring this deposited energy. The interaction which takes place, and therefore the amount and pattern of energy deposited, depends on the particle moving through the matter.

Luminosity (pb^{-1})					
		HERA Delivered		**ZEUS Physics**	
Running Period	**Year**	e^-p	e^+p	e^-p	e^+p
HERA I	1993	1.09		0.54	
	1994	1.08	5.11	0.28	3.02
	1995		12.31		6.62
	1996		17.16		10.77
	1997		36.35		27.85
	1998	8.08		4.60	
	1999	17.12	28.54	12.08	19.66
	2000		66.41		46.22
HERA II	2002		5.20		0.97
	2003		6.53		2.08
	2004		77.94		38.51
	2005	204.80		152.26	

Table 3.1: The integrated luminosity delivered by HERA I and HERA II, and gated (recorded for physics) by ZEUS, for each year of running.

High energy electrons (and positrons) predominantly lose their energy through **bremsstrahlung**, or braking radiation, which is produced as a result of the electron's deceleration as it is deflected by atomic nuclei. High energy photons lose most of their energy from production of e^+e^- pairs which then undergo bremsstrahlung. High energy electrons and photons can produce an **electromagnetic shower** by successive bremsstrahlung and pair production.

Hadronic particles, such as the proton, neutron, and pion, can produce new hadronic particles via the strong interaction. These new particles can in turn produce more new particles, leading to an **hadronic shower**, which are broader and deeper than electromagnetic showers. Hadronic showers also have an electromagnetic component since photons are produced in the decay of π^0 particles and hadronic particles can interact via the electromagnetic processes of ionization and atomic excitation.

Muons interact mainly via ionization and atomic excitation. Their interaction rate is much lower than electrons and they produce an electromagnetic shower much less often. Neutrinos interact via the weak interaction and have an extremely low interaction rate, which means they are almost never observed unless a detector of very large mass is explicitly designed to look for them.

3.3.2 A Generic Particle Detector

A particle detector is comprised of sub-detectors which can be classified as either tracking chambers or calorimeters. A **tracking chamber** relies on ionization to measure the path of a charged particle as it travels through a magnetic field. The cyclotron formula, $p_T = qBr$, can then be applied to calculate the transverse momentum. The direction of curvature of the particle trajectory determines whether it is positively or negatively charged.

A **calorimeter** functions by placing enough active material in front of the particle to cause the development of electromagnetic or hadronic showers (or a mixture of the two) and using absorbing material to detect the showers. The size of the shower in the active material is proportional to the energy of the original particle, and since the lateral and longitudinal distributions of energy are different for hadronic and electron showers, they can be used for particle identification. If the absorbing material in the calorimeter is segmented, a position measurement is also possible. A perfect calorimeter absorbs all of a particle's energy.

A cross sectional view of a generic particle detector is shown in Fig. 3.4. The tracking detectors are closest to the interaction region and the calorimeters surround the tracking detectors. Only charged particles are measured in the tracking detector.

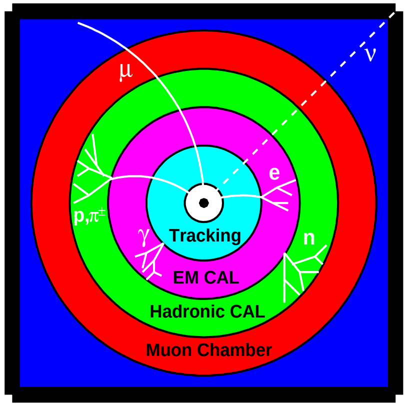


Figure 3.4: A cross sectional view of a generic particle detector and the passage of various particles through this detector.

Surround the tracking detector is the electromagnetic calorimeter. Electrons and positrons, characterized by the short and narrow showers caused by bremsstrahlung, deposit most of the energy here, along with photons, which pair produce and then undergo bremsstrahlung. Charged hadrons, such as the proton and $\pi^{\pm}$, leave a track resulting from electromagnetic interactions in the electromagnetic calorimeter and then produce a hadronic shower in the hadronic calorimeter, although it is possible that hadronic showers start in the electromagnetic calorimeter. Neutral hadrons, such as the neutron, begin to shower in the hadronic calorimeter. Muons leave a path of ionizing radiation in all sections of the detector but do not shower, and neutrinos pass through the entire detector without depositing any energy.

3.4 The ZEUS Experiment

The ZEUS experiment involves over 450 physicists, students, technicians, and staff from 52 institutes in 12 countries. Construction was completed and the first collisions were observed in 1992.

3.4.1 Introduction to the ZEUS Detector

The ZEUS detector is located approximately 30 m underground in the south hall of HERA. It has the dimensions $12\,\mathrm{m} \times 10\,\mathrm{m} \times 19\,\mathrm{m}$, weighs 3600 tonnes, and is hermetic with the exception of the beam-pipe openings. It is a general purpose detector used to study a wide range of physics by measuring the energies, direction of travel, and types of particles produced in lepton-proton collisions. A schematic drawing of ZEUS is shown in Fig. 3.5, with protons traveling from left to right and leptons traveling from right to left. The **forward direction** is defined as the direction in which the proton travels and the direction in which the lepton travels is called the **rear direction**. The ZEUS detector is asymmetric, with larger, and deeper, sub-detector components in the forward direction. This is because the asymmetric energy in HERA interactions causes the majority of the final state particles to travel in the forward direction.

The ZEUS detector is comprised of many sub-detectors which can be categorized as either tracking chambers, which measure trajectories and momenta of particles, or calorimeters, which measure the energy of particles. Concentrically surrounding the beam-pipe is the silicon Micro Vertex Detector[3] (MVD), which provides information about vertex locations. Surrounding the MVD are the inner tracking detectors; the Central Tracking Detector (CTD), the Rear Tracking Detector (RTD), and the Forward Tracking Detector (FTD). A solenoid with a magnetic field of 1.43 Tesla surrounds the CTD and allows the measurement of charged particle momenta. A uranium-scintillator calorimeter (UCAL) surrounds the inner tracking detectors and measures the energy of showers produced by particles as they pass through the uranium. The UCAL is divided into 3 regions, the Forward Calorimeter (FCAL), the Barrel Calorimeter (BCAL), and the Rear Calorimeter (RCAL). There is a Hadron-Electron Separator (HES) located between the layers of the RCAL which can distinguish between hadronic and electromagnetic showers. An iron yoke, which compensates for the field produced by the solenoid, and the Backing Calorimeter (BAC), which measures the energy of showers extending out the back of the UCAL, are sandwiched together and surround the UCAL. The Forward, Barrel, and Rear Muon Chambers (FMUON, BMUON, and RMUON) surround the BAC. They are drift chambers which detect minimum ionizing particles (muons) not yet already absorbed. There are also Forward, Barrel, and Rear Muon Identification chambers (FMUI, BMUI, and RMUI) located between the UCAL and iron yoke.

---

[3] The original vertex detector (VTX) was removed in 1996 because it didn't function properly and the MVD was installed in 2002. From 1996 to 2002, there was no vertex detector and the CTD was the innermost detector.

The ZEUS Blue Book [25] contains detailed information about all the ZEUS components. Only a brief description of the most relevant detector components will be given here.

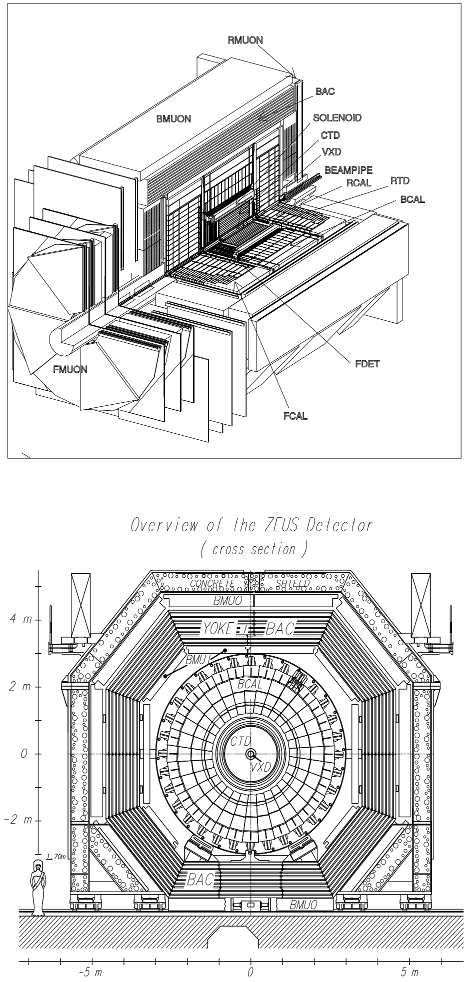


Figure 3.5: A 3D cutout (top) and a 2D x-y cross sectional (bottom) view of the ZEUS detector.

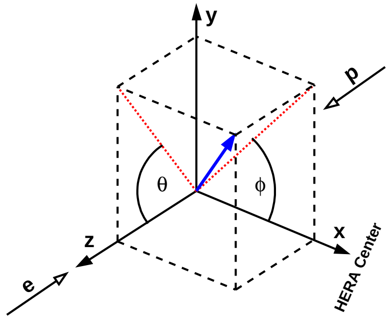


Figure 3.6: The ZEUS coordinate system.

3.4.2 The ZEUS Coordinate System

ZEUS uses a cylindrical polar coordinate system with its origin at the nominal interaction point (IP). The IP coincides with the geometrical center of the experiment[4]. The positive z direction, or the forward direction, is defined as the direction in which the proton travels. The positive x direction is defined to point from the IP to the center of the HERA ring, and in order to have a right-handed coordinate system, the positive y direction is chosen to be the vertical direction. The polar angle, θ, is the angle with respect to the z-axis and the azimuthal angle, ϕ, is the angle with respect to the x-axis. A diagram of the ZEUS coordinate system is illustrated in Fig. 3.6.

3.4.3 The Central Tracking Detector (CTD)

The CTD [26] provides a precision measurement of transverse momentum and vertex location. It is a cylindrical drift chamber with an active length of 2.05 m, an active

---

[4] The IP varies from event to event and an average value determined from many events is actually used.

inner radius of 0.182 m, and an active outer radius of 0.794 m. The polar angle coverage is $15^\circ < \theta < 164^\circ$, which corresponds to the pseudorapidity coverage of $-1.96 < \eta < 2.04$.

The CTD structure is visible in the $x-y$ cross sectional view shown in Fig. 3.7. It is organized into 16 azimuthal sectors and nine concentric superlayers. The superlayers are numbered starting at the innermost layer, and each superlayer consists of eight radial layers. Each superlayer also contains between 32 and 96 drift cells oriented at 45° to the radial direction. Each cell contains 38 field wires, which maintain a constant electric field, and eight sense wires. The wires in the axial (odd numbered) superlayers run parallel to the z-axis and the wires in the stereo (even numbered) superlayers are rotated by $\pm 5^\circ$ with respect to the z-axis. This allows a measurement of both radial and z positions. The CTD is filled with a mixture of Argon (83.4%), CO_2 (2.8%), and ethane (13.8%) gases and is in a 1.43 Tesla magnetic field produced by the solenoid.

When a charged particle passes through the CTD it ionizes the gas molecules. Electrons travel towards the positive sense wires and positive ions travel towards the negative field wires. The electrons are accelerated in the electric field and liberate other electrons from gas molecules through further ionization. The resulting electron shower produces a current pulse in the sense wire which is recorded by read-out electronics. The spatial resolution for a track which has passed all superlayers is approximately 180 μm in the $r-\phi$ plane and 2 mm in the z direction. The track transverse momentum resolution is $\sigma(p_T)/p_T = \sqrt{(0.005 p_T)^2 + (0.0016)^2}$, where p_T is in GeV. The vertex position is accurate to about 4 mm in z and 1 mm in x and y.

The inner three axial superlayers (numbers 1, 3, and 5) are equipped with a z-by-timing system, which provides a quick estimate of the track's z-coordinate. This information is obtained from the difference in arrival times of a pulse at either end of the detector. Although the resolution is large (≈ 4 cm) the speed with which the information is obtained is ideal for triggering (see Section 3.5).

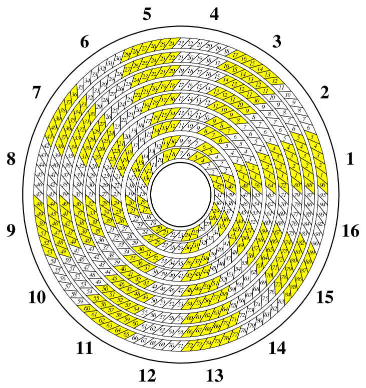


Figure 3.7: An $x-y$ view of the CTD . There are nine concentric superlayers, numbered 1-9 starting at the innermost layer. There are 16 sectors, labeled by the outer numbers. The individual cells are numbered in each superlayer starting in sector 1.

3.4.4 The Calorimeter System

In order to contain an event and ensure an accurate measurement of its energy, position, and time, the following properties are required from the calorimeter; hermeticity, good energy and position resolutions, the ability to separate hadrons from isolated electrons and electrons mixed with the hadronic final state, and the ability to calibrate the absolute energy scale and individual calorimeter sections with respect to each other to a level of 1%. In addition, a fast readout to accommodate the high interaction rate, good timing resolution in order to reject out of time backgrounds, and radiation tolerance to ensure longevity of the detector are desired. The ZEUS answer to these needs is a segmented, compensating, sampling, uranium-scintillator sandwich calorimeter with a photomultiplier tube readout. This device, abbreviated UCAL[27], is shown in Fig. 3.8.

Layers of 3.3 mm thick depleted uranium[5] absorber plates encased in thin stainless steel are **sandwiched** with layers of 2.6 mm thick scintillator. Particle showers are induced in the uranium and detected via the scintillators. Since most of the particle energy is absorbed by the uranium and not passed to the scintillator for detection, the calorimeter is termed **sampling**.

The thickness of the uranium and scintillator were chosen so that the UCAL is **compensating**, which means that it responds equally to hadrons and electrons of the same energy. Electromagnetic showers produce more photons than hadronic showers for initial particles of the same energy, and this effect can be compensated for by uranium absorption of neutrons produced in an hadronic interaction and the subsequent production of photons. his property is important when reconstructing hadronic final states, which have an unknown mixture of hadronic and electromagnetic components. The ZEUS calorimeter has a response of $e/h = 1 \pm 0.05$ when the initial particles have the same energy.

Another advantages of uranium is its very high Z value, which means that particle showers form after traversing a short distance. This allows for a compact calorimeter. Also, its constant uniform radioactivity provides an ideal means of calibration.

---

[5] The exact composition is 98.1% U^{238}, 1.7% Nb, and 0.2% U^{235}.

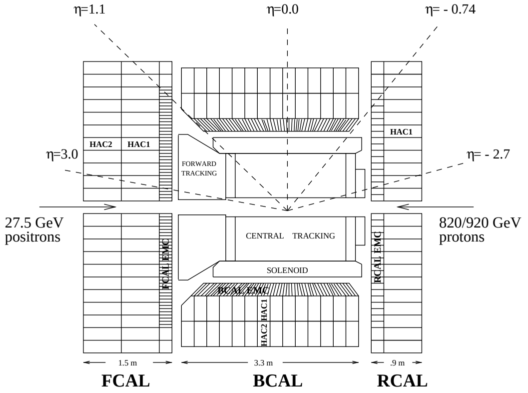


Figure 3.8: An $x - y$ view of the UCAL. Several angles in pseudoraditity are shown which correspond approximately to the boundaries of the different calorimeter sections.

The UCAL is divided into three regions; the Forward Calorimeter (FCAL), the Barrel Calorimeter (BCAL), and the Rear Calorimeter (RCAL). The segmentation of the calorimeter depends upon the region. Each region is divided into modules, which are vertical slices in the FCAL and RCAL and longitudinal slices in the BCAL. The modules are divided into towers and the towers are further divided into cells. An FCAL tower has a vertical stack of four electromagnetic calorimeter (EMC) cells in front of two hadronic calorimeter (HAC) cells. An RCAL tower has a vertical stack of two EMC cells in front of one HAC cell. The BCAL towers, which are projective in ϕ, consist of horizontally stacked EMC cells followed by two HAC cells. A BCAL tower is illustrated in Fig. 3.9.

All EMC cells are one absorption length (λ) long. An **absorption length** is the mean distance a particle must travel before it undergoes an inelastic collision.

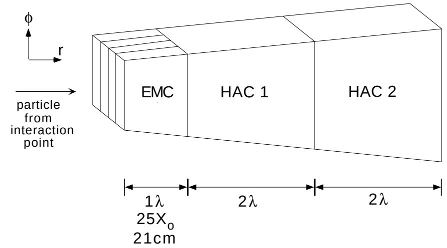


Figure 3.9: Diagram of a BCAL tower. There are 4 EMC cells stacked horizontally and behind them two adjacent HAC cells. The BCAL towers, unlike those in the FCAL and RCAL, are projective.

In uranium, one absorption length is equal to 25 radiation lengths (X_0), where the **radiation length** is defined as the mean distance an electron travels before losing all but $1/e$ of its energy. The HAC cells in the FCAL and RCAL are each 3λ in length and are 2λ in length in the BCAL. The asymmetric arrangement of cells, visible in Fig. 3.8, is to compensate for the difference in lepton and proton beam energies. A greater number of, and more energetic, final state particles travel in the forward direction, which necessitates a more finely segmented and deeper detector in this region. The characteristics of each section of the UCAL are summarized in Table 3.2.

Particles traveling through the scintillators produce light and this light is transported via a wavelength shifter to a photomultiplier tube (PMT). Each calorimeter cell is attached to two PMTs, giving a total of approximately 12,000 PMTs for the entire UCAL. The PMT digitizes the light signal it receives from the wavelength shifter and passes the information to the readout electronics. There are several advantages of having two PMTs per cell, the most important being that a more uniform detector response is obtained by averaging the two sides of a cell and the failure of a single PMT does not result in an dead cell.

	FCAL	**BCAL**	**RCAL**
Angular Coverage (θ)	2.2° → 39.9°	36.7° → 129.1°	128.1° → 176.5°
Angular Coverage (η)	$3.95 \rightarrow 1.01$	$1.10 \rightarrow -0.74$	$-0.72 \rightarrow -3.49$
Number of Cells	2172	2592	1668
Number of Modules	24	32	24
Towers/Module	11-23	14	11-23
Depth (m)	1.5	1.07	0.84
Absorption Length (λ)	7.1	5.1	4.0
Radiation Length (X_0)	181.0	129.0	103.0
EMC cell size (cm)	5 × 20	5 × 20	10 × 20
HAC cell size (cm)	20 × 20	20 × 20	20 × 20

Table 3.2: Properties of the UCAL listed by section.

The single particle energy resolution of the calorimeter, as measured in a test beam [28], is $\sigma(E) = 18\%\sqrt{E}$ for purely electromagnetic showers and $\sigma(E) = 35\%\sqrt{E}$ for purely hadronic showers, where E is in units of GeV. The timing resolution of a calorimeter cell is $\sigma(t) = \left[\left(\frac{1.5}{\sqrt{E}}\right)^2 + (0.5\,\mathrm{ns})^2\right]^{1/2}$, where again E is in GeV.

There are several other components which, at some time during the operation of the experiment, have been part of the calorimeter system. These include the small-angle rear tracking detector (SRTD), the beam-pipe calorimeter (BNC), the proton remnant tagger (PRT), the forward neutron calorimeter (FNC), the forward and rear presamplers (PRES), the barrel presampler (BPRES), the forward plug calorimeter (FPC), the backing calorimeter (BAC), and the hadron-electron separator (HES).

3.4.5 The Veto Wall and C5 Counter

The Veto Wall, located at $z = -7.5\,\mathrm{m}$, is an $8\,\mathrm{m} \times 7.6\,\mathrm{m} \times 0.86\,\mathrm{m}$ iron slab covered on both sides with scintillators. There is a $95\,\mathrm{cm} \times 95\,\mathrm{cm}$ hole to accommodate the beam-pipe and the beam magnets. The Veto Wall shields the detector from particles in the proton beam halo, provides information used to reject events originating from halo particles that pass the veto wall, and provides timing information used to reject beam-gas interactions.

There are a series of collimators, labeled C1 to C6 counting forward to rear, positioned around the beam-pipe. Attached to the C5 collimator, and located at $z = -3.15\,\text{m}$, is the C5 counter. This is a scintillation device which provides information used to synchronize the HERA and ZEUS clocks, measure synchrotron radiation, and veto beam-gas interactions.

3.4.6 The Luminosity System

In its most simplistic form, an experimental cross section is obtained by dividing the number of detected events by the integrated luminosity: $\sigma = N/\mathcal{L}$. A precise measurement of the luminosity is therefore needed in order to accurately normalize cross section.

The luminosity at ZEUS is determined from a cross section measurement of the Bethe-Heitler process [29], $e + p \rightarrow e + p + \gamma$, in which lepton-proton bremsstrahlung produces a lepton and photon with very small scattering angles. The Bethe-Heitler cross section, integrated over final state scattering angles, is given by

$$\frac{d\sigma_{BH}}{dE_\gamma} = 4\alpha_{EW} r_e^2 \frac{E_e'}{E_\gamma E_e}\left(\frac{E_e}{E_e'} + \frac{E_e'}{E_e} - \frac{2}{3}\right)\left(\ln\frac{4E_pE_eE_e'}{m_pm_eE_\gamma} - \frac{1}{2}\right), \tag{3.1}$$

where E_e and E_e' are the energies of the incoming and outgoing lepton energies, E_γ is the energy of the photon, α_{EW} is the fine structure constant, r_e is the classical radius of the electron, m_p is the proton mass, and m_e is the electron mass. This cross section

is well known for a fixed photon scattering angle and photon energy. Therefore, the measurement of the rate of photons, N_γ, at a fixed angle and energy can be used to obtain the luminosity: $\mathcal{L} = N_\gamma/\sigma_{BH}$.

The ZEUS luminosity systems differed for HERA I and HERA II. The HERA I luminosity system, as pictured in Fig. 3.10, used a lead-scintillator photon calorimeter, located at $z = -107$m, to determine the luminosity. Photons produced in the bremsstrahlung interaction travel down the proton beam-pipe unbent by the magnetic field of the beam-line magnets. They exit through a thin beryllium-copper window located at $z = -82\,\text{m}$ and enter the photon calorimeter. A carbon-lead filter was placed in front of the photon calorimeter to shield it from low energy (0.01 MeV) photons produced in synchrotron radiation. The energy resolution of the photon calorimeter [30], as measured in a test beam, is $\sigma(E) = 18\%\sqrt{E}$, and the resolution of the photon calorimeter plus shielding is $\sigma(E) = 25\%\sqrt{E}$, where E is in units of GeV. Layers of scintillators installed on the surface of the photon calorimeter can measure the impact point of the photon with a resolution of 2 mm in both the horizontal and vertical directions. The photon calorimeter can also measure the electron beam tilt and photons produced in initial state radiation.

In the HERA II luminosity system, two separate detectors are used for photon detection. A photon is detected by only one of the two detectors, and this allows for two separate luminosity measurements. The first is an upgraded version of the HERA I photon detector and the second measures the fraction of photons that pair produce as they travel through the beryllium-copper exit window in the beam-pipe. The e^+e^- pairs are separated by a dipole magnet and deflected into one of two separate

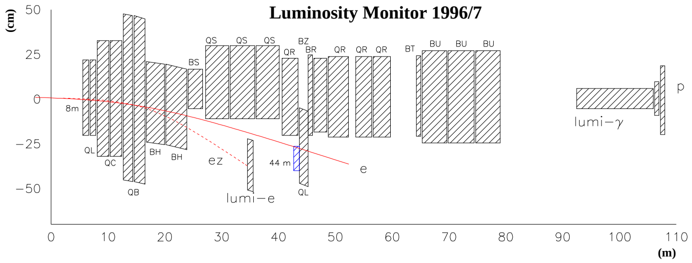


Figure 3.10: The ZEUS luminosity system for 1996-97.

calorimeters (the former BPC).

The luminosity system also includes lepton detectors ("taggers") which measure leptons with a small scattering angle that are bent by the magnetic field of the beam-pipe. These detectors are used for purposes such as calibrating the energy scale of the photon calorimeter, since $E_e = E_\gamma + E'_e$, and tagging photoproduction events in which the lepton scatters with a very small angle.

3.5 The ZEUS Trigger and Data Acquisition Systems

The HERA beams collide every 96 ns, which corresponds to a rate of 10.4 MHz. It would be impossible to record information for every event since the ZEUS Data Acquisition System (DAQ) is able to process only a few events per second. Also, only a limited amount of storage exists to record the approximately 0.3 MByte of information describing each event. These limitations are overcome by the ZEUS three-level trigger system, which rejects background and selects only the subset of events resulting from *ep* collisions which are considered interesting. Less than 10 events are passed from the trigger to the DAQ for recording out of the 10 million beam crossings occurring every second.

Events not resulting from an *ep* collision are termed **background** and are primarily due to beam-gas events, cosmic rays, and synchrotron radiation. Beam-gas events are by far the largest source of background. They occur when the lepton or proton beams interact inside, or in the vicinity of the detector, with gas particles or material in the beam-pipe wall. Through pion decay, these events can also produce **halo muons**, which are muons traveling in the exterior regions of the proton beam. Cosmic rays originating in outer-space or the upper atmosphere are another source of background. A 1 m thick layer of concrete and 30 m of earth provide substantial shielding, but some cosmic rays are still able to deposit energy in the detector. Synchrotron radiation is caused by the bending of the lepton beam.

Interesting events are those which contribute to a physics measurement. In general, these events are required to satisfy certain kinematic cuts and are able to be described by one of the broad categories of particle physics interactions, such as DIS, photoproduction, or diffraction.

A schematic diagram of the ZEUS trigger and DAQ system is shown in Fig. 3.11.

3.5.1 The First Level Trigger

The First Level Trigger (FLT) is built from custom hardware and is the fastest but least pure of the triggers. Each component has its own local FLT and decisions whether to accept or reject an event are based upon global and regional energy sums, limited tracking information, and primitive electron finding.

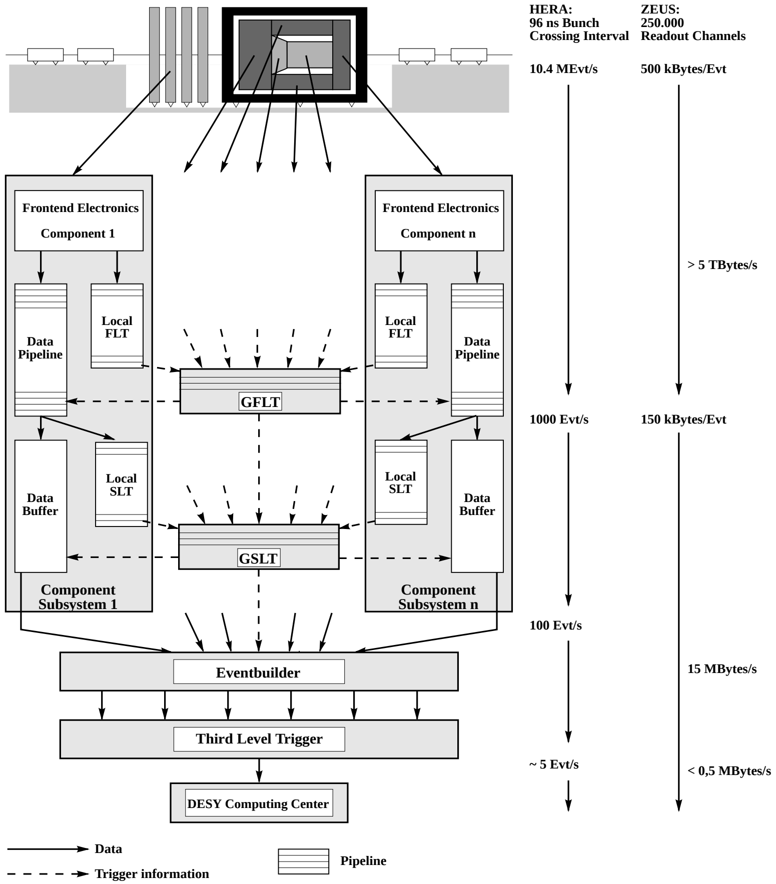


Figure 3.11: The ZEUS trigger and DAQ systems.

In each component, a small fraction of the signal exiting the front end electronics passes to the local FLT while the remaining fraction of the signal passes into a 4.4 μs[6] data pipeline. This pipeline is necessary in order to store the data while the local FLT produces information about the event. Approximately 2 μs after the interaction, each local FLT sends its information to the Global First Level Trigger (GFLT), which, based upon the information from each local FLT, makes the decision whether to accept or reject the event. The GFLT then sends this decision back to the individual components. This entire process takes a maximum of 4.4 μs.

The information in the data pipelines of the components is overwritten if the event is rejected and passed to the Second Level Trigger (SLT) for that component if the event is accepted. The FLT reduces the beam crossing rate of 10.4 MHz to a rate of about 1 kHz. Approximately 98% of the events passed from the FLT to the SLT are background.

3.5.2 The Second Level Trigger

The SLT is a software trigger running on a programmable transputer network [31]. Due to the longer time available to make a decision, approximately 6 ms, more sophisticated algorithms than those of the FLT can be run on a much larger subset of the data. Selection at the SLT level is based upon vertex and tracking information, calculations of E_T and $E - p_z$, and calorimeter timing cuts. The timing cuts veto events not originating at the nominal vertex by comparing the impact time of outgoing particles in different sections of the calorimeter. As depicted in Fig. 3.12, events originating at the nominal vertex strike opposite sides of the calorimeter at the same time while

---

[6] This time corresponds to exactly 46 bunch crossings.

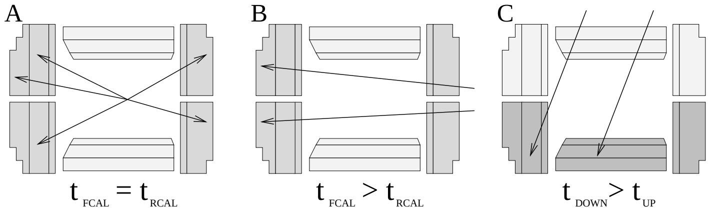


Figure 3.12: The timing of various events in the ZEUS detector. Figure A shows a lepton-proton collision which occurs at the nominal interaction point, Figure B shows a typical beam-gas event, and Figure C shows a typical cosmic muon event.

background events originating upstream do not.

Analogous to the FLT, each component has a local SLT. The data is stored in a memory buffer while the local SLT produces information. The local SLTs pass their information to the Global Second Level Trigger (GSLT), which then makes its decision and passes the result back to the local SLTs. The SLT reduces the rate from about 1 kHz to 100 Hz.

3.5.3 The Event Builder

Events which are accepted by the GSLT are passed to the Event Builder (EVB). For each event, the EVB receives information simultaneously from the multiple SLT memory buffers and reorganizes the information so that one memory buffer contains all the information corresponding to only one event. The EVB can build up to 75 events in parallel and has buffers which can hold up to 72 additional events.

3.5.4 The Third Level Trigger

The Third Level Trigger (TLT) is a software trigger running a computer farm. The TLT takes approximately 300 ms to make a decision based upon the full event information. The TLT uses the same code that used in the offline reconstruction (although some algorithms are limited by processing time) to reconstruct the kinematic quantities of the event, and then it applies selection criteria based upon these quantities. The selected events are then written to tapes stored in the DESY computing center. The total time between a collision and the decision by the TLT to accept an event is approximately 0.3 s.

3.5.5 Offline Reconstruction

Events passing the trigger selection are next processed by the offline reconstruction software. This software applies algorithms which are too slow to to run during the online TLT selection, or need constants determined by running over a large data set. Each event is then labeled by Data Storage Type (DST) bits according to which TLT cuts it passes. The DST bits are used as an easy way to select only certain categories of physics events from the storage tapes for analysis. Offline reconstruction is further discussed in Chapter 5.

Chapter 4

Event Simulation

Experimental measurements are validated by comparison to theoretical predictions. It is extremely difficult, if not impossible, to perform theoretical analytic calculations to infinite precision, and instead predictions are produced using numerical methods.

4.1 Applications of Monte Carlo Simulations

Monte Carlo methods are a class of computational algorithms which are stochastic in nature, meaning that the next state of the environment is not fully determined from the previous state. Randomly chosen numbers[1] in one phase space are mapped into another phase space according to a weight determined by the process being modeled. This technique well suits quantum calculations, which are probabilistic in nature, and is therefore extremely applicable in particle physics. The primary uses of Monte Carlo simulations are:

- To provide a theoretical prediction within a larger framework, such as the Standard Model, with which to compare experimental results;

---

[1] More accurately, pseudo-random numbers

- To provide a prediction of the type of events produced in an interaction and an estimate of their rates;
- To optimize the design of a detector;
- To aid in the development of analysis strategies so that signal-to-background conditions are optimized;
- To correct data for detector and radiative effects in order to extract a parton or hadron level result.

4.2 Stages of Monte Carlo Event Simulation

Event simulation is very complicated and it is divided into several separate stages, as shown in Fig. 4.1, in order to make the procedure more manageable. The incoming particles and their corresponding PDFs, which describe the partons at a certain x and Q^2, are provided as input, and the dependence of the parton properties as a function of x and Q^2 are calculated using parton evolution equations (see Section 2.3.5). The hard scatter is then calculated and the initial and final state QCD radiation is simulated perturbatively. The process in which colored partons form colorless hadrons is simulated using a non-perturbative hadronization model. Finally, the hadrons are passed through a detector simulation. Each of these stages is described in greater detail below.

4.2.1 Incoming Hadrons

As described in Section 2.3.5, it is possible to separate the hard and soft terms in scattering processes using factorization. The soft terms are absorbed into the PDFs,

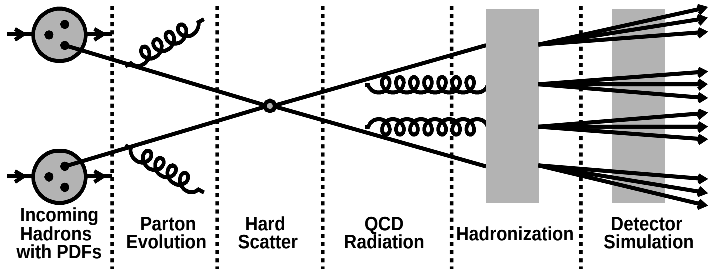


Figure 4.1: A schematic diagram of the stages involved in the simulation of particle physics events.

which are experimentally determined quantities describing the probability of finding a parton in a hadron at a certain x and Q^2. The basic strategy for determining PDFs is to calculate an initial distribution $\gamma(x_0, Q_0^2)$, parameterize the hadron's quark and gluon distributions at Q_0^2, and perform a global fit to data using equations which predict the evolution of the distributions as a function of x and Q^2 [32, 33, 34, 35].

Commonly used proton PDFs are the Glück Reya Vogt (GRV) [36], the Coordinated Theoretical-Experimental Project on QCD (CTEQ) [37], and Martin Roberts Stirling Thorne (MRST) [38] parameterizations. Commonly used photon PDFs are the GRV [39, 40], Schuler and Sjöstrand (SaS) [41], and Watanabe Hagiwara Izubchi Tanakra (WHIT) [42] parameterizations.

4.2.2 Parton Evolution

The parton evolution equations extrapolate the distributions provided by the PDFs at a certain x_0 and Q_0^2 over a large range of x and Q^2. The general form of the parton evolution equations require a summation over all leading double logarithms in

---

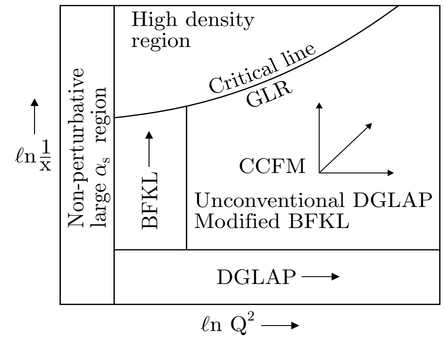


Figure 4.2: Expected regions of validity for different parton evolution equations.

$\ln(Q^2) \cdot \ln(1/x)$. DGLAP [32, 33, 34, 35], the most commonly used evolution equation, re-sums only over the single logarithms in $\alpha_s \ln(Q^2)$, neglecting the $1/x$ terms, and is expected to be valid in regions where x is not too small ($x > 10^{-2}$). The partons in the cascade are strongly ordered in k_T and ordered in x, where the parton interacting with the photon has the highest k_T and lowest x. The BFKL evolution equation [43, 20] resums over single logarithms in $\alpha_s \ln(1/x)$ and is therefore valid in regions of low x ($x < 0.1$) and Q^2 not too large ($Q^2 < 100\,\mathrm{GeV}$). The partons in the cascade are ordered in x in the same way as those in DGLAP, but are not ordered in k_T. The CCFM equation [44, 45] resums over both $\alpha_s \ln(Q^2)$ and $\alpha_s \ln(1/x)$ and should therefore be valid in large regions of x ($10^{-5} < x < 1$) and Q^2 ($1 < Q^2 < 1000$). The expected regions of validity for the different parton evolutions equations are summarized in Fig. 4.2.

4.2.3 The Hard Scatter

The hard scatter involves two incoming beam particles, or their constituents, which interact to form one or more outgoing particles. The interaction is calculated exactly to a certain fixed order using pQCD.

4.2.4 QCD Radiation

Perturbative methods are used to simulate initial state and final state QCD radiation. In e^+e^- and ep interactions, photon emissions may be significant, but the large value of the strong coupling constant, α_s, and the presence of triple gluon vertices leads to predominance of QCD quark and gluon emissions. The amount of emissions is bound by the momentum transfer, Q, occurring in the hard scatter. After the simulation of QCD radiation, the event consists of colored partons and is said to be simulated at the "parton level".

A common approach used to simulate QCD radiation is the Matrix Element and Parton Shower (MEPS) method, which is the combination of two separate perturbative calculations. In the matrix-element method, Feynman diagrams are calculated to a certain order. Although exact to the calculated order, it is increasingly difficult to calculate matrix elements at each higher order. The parton shower method relies on arbitrary branchings, such as $q \rightarrow qg$ and $g \rightarrow q\bar{q}$, of one parton into two or more partons. Approximations obtained from simplification of the kinematics allow simulations to higher orders, but predictions obtained from this method are only valid in certain phase spaces.

Heavy particles such as top quarks, electroweak gauge bosons, and Higgs bosons, which decay on time scales smaller than those of QCD radiation can also initiate

partons showers.

4.2.5 Hadronization

The colored partons resulting from the simulation of QCD radiation are not directly comparable with experimental results because confinement prevents the existence of colored partons in nature. The quarks and gluons are combined into colorless hadrons through hadronization.

Hadronization cannot be calculated from first principles due to a lack of complete understanding of the processes and perturbative methods cannot be used because the interactions have a low momentum transfer (large α_s). Therefore, non-perturbative phenomenological models are relied upon. The three main classes of hadronization models are String Hadronization, Independent Hadronization, and Cluster Hadronization. They all describe the hadronization process by iterative branchings, $\mathrm{X} \rightarrow \mathrm{Hadron} + \mathrm{X}_{\mathrm{Remainder}}$, where probabilistic rules govern the particles produced at each branch.

Hadronization Models

The Cluster Hadronization model is based upon the property of preconfinement [46, 47], which implies that that the effects of a colored particle, for example a red quark, are dependent upon the location of its partner particle, in this case an anti-red quark. The gluons produced from the simulation of QCD radiation are split into pairs of quarks and anti-quarks, or diquarks and anti-diquarks. Each quark or diquark can be connected by color to an anti-quark or anti-diquark with which it can form a color-singlet cluster. The clusters have a distribution of mass and spatial size peaked at small values because of confinement. Each cluster is fragmented into two hadrons or, if it is too light to fragment, is taken to represent the lightest single hadron corresponding to the flavor of the quarks in the cluster.

The String Hadronization Model [48], also known as the Lund String Model, is based upon preconfinement and assumes that confinement is linear. It is easiest to understand this model by imagining the production of a $q\bar{q}$ pair in which the quarks are connected by a one dimensional string with a string constant $\kappa \approx 1\,\mathrm{GeV}/\,\mathrm{fm}$ corresponding to its "mass density". As the $q\bar{q}$ pair move apart, the potential energy stored in the string increases linearly. A some distance, it is more advantageous for the string to split and combine with a $q'\bar{q}'$ vacuum fluctuation to produce $q\bar{q}'$ and $q'\bar{q}$ pairs. These pairs further divide if the mass of their connecting strings is large enough, and the divisions continue until only on-shell hadrons remain with a small piece of string connecting the quark and anti-quark. In the simplest method, baryons are produced by replacing the anti-quark by a diquark at the end of a string and having the string break as a result of either quark-anti-quark or diquark-anti-diquark fluctuations. Gluons are treated as "kinks" on the string which carry energy and momentum.

The Independent Hadronization model popularized by Field and Feynman [49] is not based upon preconfinement since it does not incorporate long range interactions between partons. It assumes that the hadronization of a system of partons can be calculated by summing the hadronization processes of each individual parton, which is independent of the hadronization processes of all other partons. For example, a quark of flavor q_1 combines with a quark of flavor $\bar{q}_2$ from the created pair $q_2\bar{q}_2$ to form $q_1\bar{q}_2$, q_2 combines with $\bar{q}_3$ from the pair $q_3\bar{q}_3$, and so on until the quarks are

below a certain energy. The hadronization of $q_1\bar{q}_2$ is independent of the hadronization of $q_2\bar{q}_3$, and so on. There is no standard method for handling gluons in Independent Hadronization, but it is common to split the gluon into a $q\bar{q}$ pair or treat it like a quark with a randomly chosen flavor. Independent Hadronization has not matched the success of the newer models based upon preconfinement and is no longer widely used.

The Underlying Event

The underlying event includes all hadronic activity in addition to that produced by the hard scatter. In hadron-hadron interactions the underlying event can be caused by **Multi-Parton Interactions (MPI)**, which result from interactions between the beam remnants. In *ep* photoproduction the MPIs are simulated only for resolved interactions, which, unlike direct interactions, contain both proton and photon remnants. A diagram of MPIs in *ep* resolved photoproduction is shown in Fig. 4.3.

MPIs are calculated using phenomenological models with several assumptions and input parameters. They make a direct impact on the hadronic final state by increasing the energy, and in rare cases the MPIs have enough energy to produce additional hadronic final states.

Many unstable hadrons are produced during hadronization and their decay to stable hadrons is calculated using tables of branching ratios and decay modes. The stable hadrons, which do not decay in a reasonable amount of time, are said to be simulated at the "hadron level".

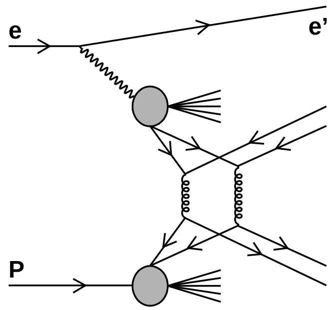


Figure 4.3: Multi-parton interactions in *ep* collisions.

4.2.6 Detector and Trigger Simulation

Event generators used in the ZEUS collaboration are wrapped in the AMADEUS software package, which organizes the information describing the hadrons produced in the hadronization process into data structures using the ADAMO[2] [50] management system. The database structure is the same as that used to store data events and in the format required as input by the ZEUS detector and trigger simulations.

Hadrons produced in the hadronization process are passed through a full simulation of the ZEUS detector and trigger. MOZART[3] uses the GEANT [51] package to describe the geometry and material of the ZEUS subcomponents and simulates the passage of particles through these subcomponents as well as through dead material. Events are then passed through CZAR[4], which simulates the ZEUS trigger logic.

---

[2] Aleph Data Model

[3] Monte Carlo for Zeus Analysis, Reconstruction, and Trigger

[4] Complete ZGANA Analysis Routine, where ZGANA stands for ZG313 Analysis

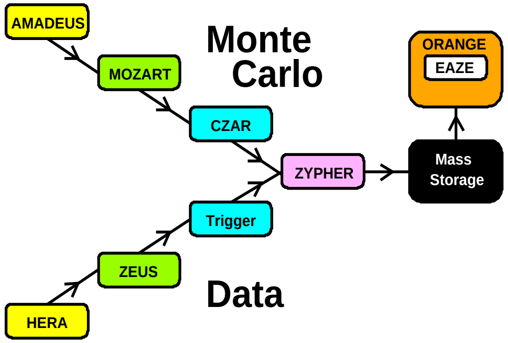


Figure 4.4: Diagram showing the processing of ZEUS data and simulated events.

Events at this stage are referred to as “detector level”.

The simulated events are stored in the same format as the data, which allows the same reconstruction methods to be applied to both samples. Events are reconstructed using the ZEPHYR[5] package, as described in Chapter 5, and then written to tape, where they are then available for offline analysis. EAZE[6] is the standard ZEUS offline analysis package and produces an Ntuple containing the event information. ORANGE[7] [52] is a software library that is compiled with all EAZE jobs. It provides a standardized set of commonly used analysis routines with the goal of making ZEUS analyses more reliable and reproducible. Analysis at the EAZE and ORANGE level is described in Chapter 6. A diagram summarizing the processing of ZEUS data and simulated events is shown in Fig. 4.4.

---

[5] ZEUS Physics Reconstruction

[6] Easy Analysis of ZEUS Events

[7] Overlying Routine for Analysis Ntuple Generation

4.3 Monte Carlos used to Simulate Photoproduction Events

The Pythia[8] [53] and Herwig[9] [54] Monte Carlos are general purpose event generators with an emphasis on multi-particle productions in e^+e^-, ep, and pp collisions, and both have been shown to provide valid predictions for photoproduction data. Both use LO matrix elements to calculate the hard shower, but Pythia attempts to model the non-perturbative processes as closely as possible, while Herwig attempts to use the simplest universal model available. As a result, Pythia has many input parameters, while Herwig has relatively few. Direct, resolved, and color singlet samples are generated separately in both Monte Carlos.

4.3.1 PYTHIA

PYTHIA uses the Weiszäcker-Williams [55] approximation to generate the spectrum of photons emitted from the lepton. The factorization scale (see Section 2.3.5) is set to the mean transverse mass of the two outgoing partons in the hard scatter,

$$\mu_F^2 = m_T^2 = \frac{1}{2}\left(m_1^2 + p_{\mathrm{T1}} + m_2^2 + p_{\mathrm{T2}}\right), \tag{4.1}$$

where m_1^2 and p_{T1} are the mass and transverse momentum one of the outgoing particles and m_2^2 and p_{T2} are the mass and transverse momentum of the other outgoing particle. In Pythia, it is possible to adjust the $p_{\mathrm{T}}^{\mathrm{min}}$ of the hard interaction and the $p_{\mathrm{T}}^{\mathrm{min}}$ of the MPIs separately.

The QCD radiation is calculated according to the MEPS method. In Pythia, each parton is characterized by some virtuality scale, Q^2, which is associated with the mass of the branching partons by the relation $m^2 = E^2 - \boldsymbol{p}^2$. Partons initiating final state radiation are time-like ($m^2 > 0$) and the virtuality of partons produced in each successive branch is decreased until a cutoff Q_0^2 is reached and all partons are on-shell ($m^2 \approx 0$). Initial state radiation is calculated backwards in time from the hard scatter to the partons in the beam hadrons. The partons initiating initial state radiation are space-like ($m_2 > 0$), and once again the virtuality of the partons produced in each successive branch is decreased until the particles are on-shell.

The String Hadronization model is used to calculate the hadronization in Pythia and there are two main models of MPIs available. The first is based upon a simple model in which partonic showers in an event are ordered in a decreasing series in p_{T}. All hadronic collisions are considered to be equivalent and all parton-parton interactions independent of each other. The rate of partonic interactions is then generated according to a Poisson distribution. The second model assumes that each hadronic collision depends upon a varying impact parameter, which in turn depends upon the distribution of the partons within the hadron. The rate of interactions increase based upon the amount of overlap of the impact parameters.

Pythia does not contain a simulation of color-singlet exchange in hard interactions. However, an event topology similar to that of color-singlet exchange can be simulated by high-t photon exchange for quark-quark scattering in LO resolved processes since the photon does not radiate color charge. It is important to note that these events are used only to produce events with a similar topology as those produced by Pomeron exchange and are not a source of events with rapidity gaps in hard diffractive photoproduction.

---

[8] Named after the priestess of the Oracle in Delphi.

[9] Hadron Emissions Reactions With Interfering Gluons

4.3.2 HERWIG

Herwig uses the Equivalent Photon Approximation [56] to generate the spectrum of photons emitted from the lepton. The factorization scale of the hard scatter is given by

$$\mu_F^2 = \frac{2stu}{s^2 + t^2 + u^2}, \tag{4.2}$$

where s, t, and u are the Mandelstam variables. The resulting cross section is divergent in the limit that $p_{\mathrm{T}}^{\mathrm{min}} \to 0$ and therefore a cut on the minimum transverse momentum of the produced partons must be applied. Herwig only allows the adjustment of $p_{\mathrm{T}}^{\mathrm{min}}$ in the hard interaction.

The QCD radiation is calculated according to the MEPS method. The partons are ordered angularly, and at each branching the angle between the two emitted partons is smaller than the angle in the previous branching. The showering stops when the angle is less than a certain cutoff value. The calculation of both initial and final state radiation starts from the partons involved in the hard scatter, with initial state radiation being evolved backwardly.

Herwig relies on the Cluster Hadronization model in the hadronization process and is interfaced with the Jimmy [57, 58] library to simulate MPIs. Jimmy uses an eikonal model [59], which combines the assumption of a Poisson distribution at each fixed impact parameter with an impact dependent overlap function. The model is independent of the value of $p_{\mathrm{T}}^{\mathrm{min}}$.

Herwig implements BFKL color-singlet exchange according to the model of Mueller and Tang [60]. The hard-Pomeron intercept, $1 + \omega_0$, is related to the strong coupling, α_s, along the gluon ladder in the BFKL parton evolution by $\omega_0 = \alpha_s C_A \left[4 \ln (2/\pi)\right]$.

Chapter 5

Event Reconstruction

The reconstruction process combines several forms of raw information from the detector components, such as pulse heights in photo-multiplier tubes and drift times in tracking chambers, into physics objects. Reconstruction determines the kinematic properties of each particle involved in an interaction, which enables the characterization of the event as a whole.

5.1 Reconstruction in the ZEUS Environment

Data stored on tape after the ZEUS trigger selection is processed offline by the ZEPHYR software package, which performs reconstruction based upon information from the individual detector components. EAZE level routines then use the reconstructed information to form objects corresponding to the particles produced in the *ep* collision and to determine the variables which describe these objects.

5.2 Track and Vertex Reconstruction

The signals recorded by the tracking chambers provide the information necessary to reconstruct charged particle tracks and to determine the location of the interaction

vertex. ZEUS uses the VCTRACK [61, 62] routine to perform track reconstruction and primary and secondary vertex reconstruction.

5.2.1 Track Reconstruction

VCTRACK can be run in regular mode, which used information from the CTD, FTD, RTD, and SRTD, or CTD-only mode, which uses only information from the CTD. The regular mode has the advantage of an increased angular acceptance, but the CTD, which has an angular acceptance of $-15^\circ < \theta < 164^\circ$, is the best understood of all the tracking detectors, and therefore the CTD-only mode is used in this analysis.

Track finding begins with a seed consisting of three hits in an axial superlayer in the outer part of the CTD. The seed is then extrapolated towards the inner superlayers using a pattern recognition algorithm, and the trajectory parameters are updated as more CTD hits with increasing precision are accumulated. The longest tracks are found first and the shorter ones next, making sure not to include segments of the longer tracks in the shorter ones. Tracks with too many shared hits are removed. A second iteration of the pattern recognition algorithm is run to find tracks in the outer superlayers of the CTD which do not extend into the inner superlayers.

Each track candidate is fit to a five parameter Helix model [62]. The fit begins with the innermost hits, and during the course of the fit some hits are swapped between tracks or discarded. The five parameter helix model is illustrated in Fig. 5.1.

5.2.2 Vertex Reconstruction

The fitted tracks are used as input for the vertex finding. A loop over all tracks is performed and the determination of which tracks originate from the primary vertex

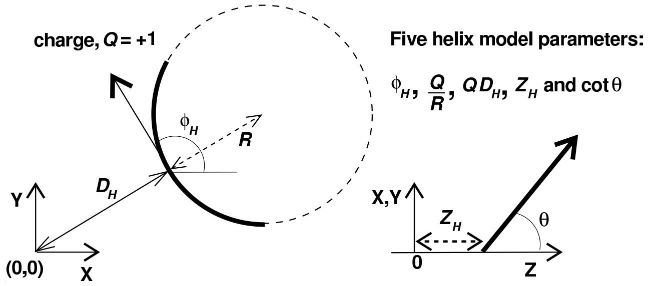


Figure 5.1: An illustration of the five parameter helix model used in track fitting.

and which originate from the secondary vertices is achieved using a χ^2 fit.

5.3 Calorimeter Reconstruction

The calorimeter reconstruction relies on the position of the calorimeter cells, the magnitude of the signal pulses from the two PMTs associated with each cell, and the timing difference between the PMT pulses. First, calorimeter noise is suppressed, then corrections are applied to the cells, and finally, the calorimeter cells are grouped into clusters.

If a PMT is known to be broken, the energy of the calorimeter cell is set to double the energy of the working PMT and the imbalance is set to 0. Imbalance is defined as the fractional energy difference between the two PMTs of a cell, $I_{cell} \equiv \left|\frac{E_{left}-E_{right}}{E_{left}+E_{right}}\right|$.

5.3.1 Calorimeter Noise Suppression

The various sources of calorimeter noise are accounted for by the Noise96s [63] routine. EMC cells with $E < 60\,\text{MeV}$ and HAC cells with $E < 110\,\text{MeV}$ are removed, as are

isolated EMC cells with $E < 80\,\text{MeV}$ and isolated HAC cells with $E < 140\,\text{MeV}$. This removes most of the noise from the depleted uranium, which is concentrated at low energies. Cells with an energy imbalance are removed if $I_{cell} < 0.49E + 0.03$, which accounts for cells in which one of the PMT bases produces a small spark. Hot cells are defined as those producing an unusually large or frequent signal and are often caused by a hardware failure. The hot cells are stored in a list and are used in a data correction routine at the EAZE level.

5.3.2 Calorimeter Energy Scale Corrections

The simulation of the calorimeter response is not accurate enough to provide an exact description of the data [64, 65]. Two classes of events, both in which the scattered lepton energy can be predicted from the rest of the event, were used to determine the constants for a recalibration that achieved equality in the energy response of the data and MC over the detector. [66]. For events in which the scattered lepton was very close to the RCAL beam-pipe, the kinematic peak method [66] was used. These events have a very low y, and therefore, lepton energy distributions which peak near the beam energy. The energy distributions for both the data and MC were fit to a Gaussian and the ratio of the peak positions was used as the recalibration factor. For events in which the scattered lepton was not close to the RCAL beam-pipe, the double-angle method (see Section 5.5.3) was used. In this method, the energy of the scattered lepton is determined only from the angle of the scattered lepton and the angle of the hadronic system. The average difference in the calorimeter measured energy and the double angle measured energy of the scattered lepton was used as the calibration constant. This method has the advantage of including a wider range of lepton energies since it takes into account cells far from the RCAL beam-pipe, but has the disadvantage of poor hadronic angle resolution for low lepton energies.

The RCALCORR [67] routine applies different recalibration constants to the calorimeter cells based upon the section of the CAL in which the cell is located. The cells in FCAL are not scaled at all, the cells in BCAL are scaled by 1.05, and the cells in RCAL are scaled individually by a factor between 1.00 and 1.025. If no information from the SRTD and PRES is available, all cells are scaled by these factors. Otherwise, the SRTD and PRES information is used to recalibrate the energy deposited by the scattered lepton.

5.3.3 Island Formation

Energy deposited in the CAL is usually spread over several adjacent calorimeter cells and the cells having energy deposits most likely belonging to a single particle are merged. This process is based on the idea of islands of energy surrounded by areas lacking energy, and uses either a tower (see Section 3.4.4) or cell as the smallest geometrical unit. As illustrated in Fig. 5.2, each cell is uniquely associated with its highest energy neighbor. If the cell being considered is itself the cell with the highest energy, it becomes the "peak" of the island. Island finding is first applied to the individual sections of the CAL and is then joined across sections.

The cell-islands can be combined using a probability function based upon their angular separation to form three dimensional objects called cone-islands. The position of a cone-island is determined by the logarithmic center of mass of the shower. Cone-islands are used in this analysis for lepton finding (see Section 5.4) and reconstruction of the hadronic final state (see Section 5.6).

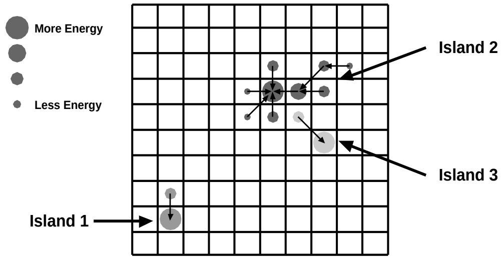


Figure 5.2: An illustration of the island finding algorithm used in the CAL reconstruction.

5.4 Lepton Reconstruction

A scattered lepton deposits almost all of its energy in the EMC, and this deposit, along with a corresponding track in the CTD, provides the strongest signature for lepton identification. However, the larger angular coverage of the CAL means that particles do not have to first pass through the CTD before interacting in the CAL, and therefore initial lepton finding is only done with the CAL.

Sinistra95 [68], a neural network trained on NC DIS data and MC, is used for lepton finding. The longitudinal and transverse energy distributions of the islands, which characterize the origin of the particle shower, are the input to Sinistra95. The output is a list of lepton candidates, each with probability between 0 and 1, signifying the likelihood that the energy deposit was of electromagnetic origin. Sinistra95 is 80% efficient at finding leptons if the energy of the electromagnetic deposit is greater than 10 GeV and the probability is greater than 0.9 [66].

In this analysis, the scattered lepton travels down the beam-pipe. Therefore, events in which Sinistra95 found an lepton were excluded from the data sample in order to remove NC DIS background (see Section 6.3.1).

5.5 Reconstruction of Kinematic Variables

In order to fully describe a general lepton-proton collision, each four-momentum component of the scattered lepton and hadronic system must be known. The number of free variables can be reduced from eight to four by imposing energy and momentum conservation, and further to three by fixing the mass of the lepton. These are usually chosen to be the Lorentz invariant quantities; x, y, and Q^2 (see Section 2.2). By fixing the energy of the incoming lepton, it is possible to provide a full description with only two variables, most often chosen to be y and Q^2. Different reconstruction methods allow the determination of y and Q^2 from only the lepton energy and angle, from only the hadronic system energy and angle, and from only the angles of the both lepton and hadronic systems. The choice of method depends on the kinematic region, the interaction type, and the reference frame being studied.

5.5.1 The Electron Method

The electron method uses only information from the scattered electron to calculate y and Q^2;

$$y_{el} = 1 - \frac{E'_e}{2E_e}\left(1 - \cos\theta_e\right) \tag{5.1}$$

$$Q^2_{el} = 2E_e E'_e \left(1 + \cos\theta_e\right), \tag{5.2}$$

where E_e is the energy of the incoming lepton, E'_e is the energy of the scattered lepton, and θ_e is the angle of the scattered electron. This method assumes that E_e and E'_e are the actual energies involved in the hard scatter and is therefore sensitive to initial and final state radiation.

5.5.2 The Jacquet-Blondel Method

The Jacquet-Blondel method uses only information from the hadronic system to calculate y and Q^2;

$$y_{\mathrm{JB}} = \frac{(E - p_z)_{\mathrm{Had}}}{2E_e} = \frac{\sum_i \left(E_i - p_{z,i}\right)}{2E_e} \tag{5.3}$$

$$Q^2 = \frac{p^2_{\mathrm{T,Had}}}{1 - y_{\mathrm{JB}}} = \frac{\left(\sum_i p_{x,i}\right)^2 + \left(\sum_i p_{y,i}\right)^2}{1 - y_{\mathrm{JB}}}, \tag{5.4}$$

where the sum is performed over all CAL cells except those associated with the scattered lepton. The quantities p_{T} and $(E - p_z)_{\mathrm{Had}}$ are close to 0 for remnant particles emerging at a small angle and traveling down the beam-pipe, which ensures that the calculations are not biased by these undetectable particles. This method is sensitive to initial state but not final state radiation since it depends on E_e but not E'_e.

For the incoming lepton and proton,

$$E - p_z = (E_e + E_p) - (p_{z,e} + p_{z,p}) \approx 2E_e = 55.04\,\mathrm{GeV}, \tag{5.5}$$

where the lepton is traveling in the negative z direction, $E_p \approx p_{z,p}$ and $E_e \approx -p_{z,e}$, and $E_e \approx 27.52\,\mathrm{GeV}$. This quantity is conserved in the collision and therefore the measured $E - p_z$ of the final state should be approximately 55 GeV if all particles are contained in the detector. However, this is not the case in photoproduction, where the lepton escapes undetected down the beam-pipe. Requiring y_{JB} to be below a certain maximum value can exclude NC DIS events and therefore help select a photoproduction sample. This is especially useful for rejecting events in which scattered lepton was not identified by SINISTRA95 and therefore misclassified as photoproduction. Requiring y_{JB} to be above a certain minimum value excludes beam-gas events, which may have a different $E - p_z$ than the lepton proton collision.

5.5.3 The Double Angle Method

The double angle method relies on the angle of the scattered lepton and hadronic system;

$$y_{da} = \frac{\sin\theta\,(1-\cos\gamma_{\mathrm{Had}})}{\sin\gamma_{\mathrm{Had}} + \sin\theta - \sin(\gamma_{\mathrm{Had}}+\theta)} \tag{5.6}$$

$$Q^2_{da} = 4E_e^2 \frac{\sin\gamma_{\mathrm{Had}}\,(1+\cos\theta)}{\sin\gamma_{\mathrm{Had}} + \sin\theta - \sin(\gamma_{\mathrm{Had}}+\theta)}. \tag{5.7}$$

The advantage of using only angles is that angular resolution is usually better than energy resolution, which leads to a more precise measurement in some kinematic regions.

5.6 ZEUS Unidentified Flow Objects (ZUFOs)

Combining tracking and calorimeter information significantly improves the reconstruction of the hadronic final state. The tracking system has a better angular resolution and a better energy resolution at low energy. Also, the calorimeter system is sensitive to dead material between it and the tracking system and low energy particles may not reach the calorimeter but are still measured by the tracker. The hadronic objects resulting from a combination of CAL and CTD information are called **ZEUS Unidentified Flow Objects** (ZUFOs) [69] within the ZEUS collaboration and **Energy Flow Objects** (EFOs) in ZEUS publications.

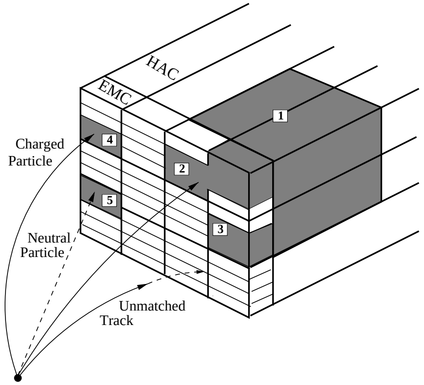


Figure 5.3: Neighboring calorimeter cells are clustered to form cell-islands. One HAC cell-island (1) and four EMC cell-islands (2,3,4,5) are depicted. The HAC cell-island 1 is joined with the EMC cell-islands 2 and 3 to form a cone-island. The cell-islands and cone-islands are then associated with tracks to form ZUFOs.

5.6.1 Track and Island Matching

ZUFOs are formed by extrapolating “good” charged tracks in the CTD to the inner surface of the CAL and associating them with cone-islands. Good tracks are those which have traversed as least four layers of the CTD and have a transverse momentum in the range[1] $0.1 < p_{\mathrm{T}} < 20\,\mathrm{GeV}$. A track “matches” an island if the distance of closest approach between the extrapolated track and the island is less than 20 cm or if the track lies within the radius of the island.

The result of the matching process determines whether to assign the track energy or calorimeter energy to the ZUFO. Tracks with no matching cone-islands are attributed to charged particles and the energy of the ZUFO is determined by assuming

---

[1] The maximum p_{T} is increased to 25 GeV if the track passed through more than seven superlayers.

the particle is a pion. Cone-islands with no matching tracks are attributed to neutral particles and the energy measured by the calorimeter is used as the ZUFO energy. Cone-islands associated with more than three tracks are assigned the calorimeter energy. For cone-islands associated with less than three tracks, the tracking information is used if $E_{cal}/p_{trk} < 1.0 + 1.2 \cdot \sigma\left(E_{cal}/p_{trk}\right)$ and $\sigma(p_{trk})/p_{trk} < \sigma(E_{cal})/E_{cal}$. The first requirement ensures that the energy deposit in the calorimeter is due to the associated track alone and the second requirement ensures that the momentum resolution of the track is better than the energy resolution of the associated calorimeter object.

5.6.2 Corrections

Backsplash consists of small energy deposits at large polar angles in the calorimeter which bias the hadronic angle measurement. It originates from neutral particles which escape from the front of the CAL during a shower and traverse the detector, or showering in dead material. It can also be caused by noisy calorimeter cells, cosmic rays, and beam-gas. Backsplash is removed by comparing the average and maximum hadronic angles. The backsplash correction mainly benefits DIS events and is therefore not applied in this analysis.

It is also possible to make hadronic energy corrections to the ZUFOs in order to compensate for energy loses in dead material. The correction factors are determined separately for data and MC using high Q^2 NC DIS events and only applied to those ZUFOs which are assigned calorimeter energy. ZUFOs which undergo hadronic corrections are called **corrected ZUFOs**. Corrected ZUFOs are not used in this analysis.

5.7 Jets Algorithms

As described earlier, partons produced in a high-energy collision initiate a parton shower through QCD radiation and the partons in the shower then combine via hadronization to form colorless hadrons. Since the high-energy collision is on the order of several GeV and the QCD radiation is on the order of 1 GeV, the partons in the shower and the subsequent hadrons are collimated in the direction of the original parton produced in the collision. A group of collimated particles is called a **jet**. The relation between partons and jets allows the determination of hadron-level information from the properties of the jets.

5.7.1 Requirements of Jet Algorithms

A jet is not fundamental to QCD, but it is possible to form jets from various objects such as partons, hadrons, calorimeter cells, and ZUFOS, as long as the properties of a jet reflect the properties of the constituents from which it was formed. A jet should be easy to construct from the hadronic final state and easy to calculate in pQCD. Algorithms used to construct the jets should satisfy the following:

- **Collinear safety** The jet finding should be insensitive to collinear radiation, which means that it should treat a single particle the same as an any number of particles with the same total momentum as that single particle. Divergences in theoretical calculations disappear only when no distinction is made between a single particle with energy E and N collinear particles whose energies sum to E and whose contributions are integrated over. Experimentally, this means that the jet algorithm should be independent of the detector granularity. In other words, the measured energy should be the same whether the energy deposit is contained in one calorimeter cell or spread over many cells.

- **Infrared safety** The jet finding should be insensitive to infrared radiation, which means that the result should not change if an arbitrary number of infinitely soft partons with arbitrary directions are added. This avoids soft divergences in perturbative calculations. Experimentally, this means that the measured energy should not be effected by very low energy particles.

5.7.2 The Cone Algorithm

The cone algorithm has historically been used to find jets originating from pp collisions. All objects with an E_{T} greater than a certain value are selected as seeds for jets. The seed is defined as the center of a cone of radius R_{cone} in $\eta - \phi$ space and all objects with

$$R^2 = \left[(\eta_i - \eta)^2 + (\phi_i - \phi)^2\right] < R_{cone} \tag{5.8}$$

are combined with the seed to form a jet candidate according to the Snowmass Accord [70],

$$E_{\mathrm{T}}^{\mathrm{Jet}} = \sum_i E_{\mathrm{T},i} \tag{5.9}$$

$$\eta^{\mathrm{Jet}} = \frac{1}{E_{\mathrm{T}}^{\mathrm{Jet}}} \sum_i E_{\mathrm{T},i}\eta_i \tag{5.10}$$

$$\phi^{\mathrm{Jet}} = \frac{1}{E_{\mathrm{T}}^{\mathrm{Jet}}} \sum_i E_{\mathrm{T},i}\phi_i. \tag{5.11}$$

The summation is repeated with the jet candidates as the center of the cone. This process repeats until the center of the cones are the same as the center of the jet candidates or a maximum number of iterations is reached.

Cone algorithms are collinear safe but not infrared safe. The presence of soft radiation between between cones can lead to a merging of jets which would not occur in its absence. The treatment of overlapping jets is dependent upon the exact implementation of the algorithm and the lack of a standardized treatment of overlapping jets also presents a problem because two cone algorithms with the same R_{cone} can produce different results. However, the use of seeds minimizes the CPU time required to find the jets.

In this analysis, the EUCELL version of the cone algorithm is used in the TLT. Cells are required to have $E_{\mathrm{T}} > 1\,\mathrm{GeV}$ to be considered as seeds and $R_{cone} = 1$. The EUCELL algorithm is designed so that the energy in the overlap region is associated with the highest E_{T} jet.

5.7.3 The Cluster Algorithm

The cluster algorithm has historically been used in e^+e^- experiments but has been modified for use in ep experiments. The variations of this algorithm differ in the method used to measure the distance between two particles and the way in which particles are merged. The k_{T} cluster algorithm, used in this analysis, defines the distance from a particle, i, to the beam-pipe in momentum space as

$$d_i = E_{\mathrm{T},i}^2, \tag{5.12}$$

and the distance between two particles i and j as

$$d_{i,j} = \min\left(\mathrm{E}_{\mathrm{T},i}^2, \mathrm{E}_{\mathrm{T},j}^2\right)\left[(\eta_{\mathrm{i}} - \eta_{\mathrm{j}})^2 + (\phi_{\mathrm{i}} - \phi_{\mathrm{j}})^2\right], \tag{5.13}$$

as illustrated in Fig. 5.4.

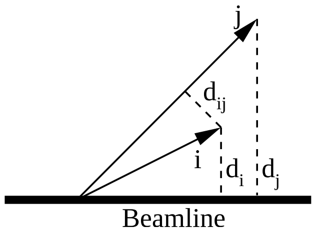


Figure 5.4: An illustration of variables used in the k_T jet finding algorithm. Distances are in momentum space.

For every particle, d_i is calculated, and for every pair of particles, $d_{i,j}$ is calculated. If the smallest of all the d_i and $d_{i,j}$ quantities is a $d_{i,j}$, the objects i and j are combined into a new object k according to

$$E_{\mathrm{T},k}^2 = E_{\mathrm{T},i}^2 + E_{\mathrm{T},j}^2 \tag{5.14}$$

$$\eta_k = \frac{E_{\mathrm{T},i}\eta_i + E_{\mathrm{T},j}\eta_j}{E_{\mathrm{T},i} + E_{\mathrm{T},j}} \tag{5.15}$$

$$\phi_k = \frac{E_{\mathrm{T},i}\phi_i + E_{\mathrm{T},j}\phi_j}{E_{\mathrm{T},i} + E_{\mathrm{T},j}}. \tag{5.16}$$

If the smallest of all the d_i and $d_{i,j}$ quantities is a d_i, the object is classified as a jet candidate and no longer merged. The process is repeated for the remaining particles until there are none left. Jet candidates with E_T greater than a certain predefined value are considered jets.

Cluster algorithms are collinear and infrared safe. In addition, there is no problem with overlapping jets since every particle is unambiguously assigned to a single jet and there is no dependence on a seed choice.

In this analysis, the k_T algorithm is used to find jets from the ZUFOs in the laboratory frame.

5.8 Reconstruction of Jet Based Variables

5.8.1 Reconstruction of E_T and η

The transverse energy and pseudorapidity of a jet with four-momentum (E, p_x, p_y, p_z) are determined by

$$E_\mathrm{T} = \sqrt{p_x^2 + p_y^2} \tag{5.17}$$

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right), \tag{5.18}$$

where $\theta = \tan^{-1}\left(\frac{E_\mathrm{T}}{p_z}\right)$.

5.8.2 Reconstruction of x_γ and x_p

At leading order, the fraction of photon and proton momentum involved in the hard scatter can be calculated according to

$$x_\gamma = \frac{\sum_i E_{\mathrm{T},i} e^{-\eta_i}}{2E_\gamma} \tag{5.19}$$

$$x_p = \frac{\sum_i E_{\mathrm{T},i} e^{\eta_i}}{2E_p}, \tag{5.20}$$

where the sum is over the outgoing partons. $E_{\mathrm{T}i}$ and η_i are the transverse momentum and pseudorapidity of parton i in the laboratory frame. A derivation of these quantities is given in Appendix A.3.

Since the partons are not directly measured in an experiment, a sum over jets can be used to define the quantities $x_\gamma^{\mathrm{OBS}}$and x_p^{OBS}, which are correlated to x_γ and x_p

at the hadron level,

$$x_\gamma^{\mathrm{OBS}} = \frac{\sum_i E_{\mathrm{T},i} e^{-\eta_i}}{2E_\gamma} \tag{5.21}$$

$$x_p^{\mathrm{OBS}} = \frac{\sum_i E_{\mathrm{T},i} e^{\eta_i}}{2E_p}. \tag{5.22}$$

These definitions are valid to all orders in perturbation theory and can also be applied at the detector level if y_{JB} and the jet based $E_{\mathrm{T}i}$ and η_i are used in the calculation.

At LO, $x_\gamma = 1$ for direct processes and falls somewhere between 0 and 1 for resolved processes, but beyond LO the direct and resolved terms are ambiguous (see Section 2.4.4). Parton showers and higher order hadronization, as well as detector resolution, cause $x_\gamma^{\mathrm{OBS}}$to have values other than exactly 1 for direct processes. Events with $x_\gamma^{\mathrm{OBS}} > 0.75$ are more likely to be direct than resolved and are termed **direct enhanced**. Events with $x_\gamma^{\mathrm{OBS}} < 0.75$ are more likely to be resolved than direct and are termed **resolved enhanced**.

Chapter 6

Event Selection

The data used for this analysis was collected during the ZEUS 1996-97 running period. A clean dijet photoproduction sample was selected by applying both online and offline cuts to data taken with stable beam and detector conditions. Further offline cuts were applied in order to select the rapidity gap sample studied in the analysis.

6.1 Run Selection

The EVTAKE routine was used at the EAZE level (see Section 4.2.6) to select data runs taken under sufficiently good beam and detector conditions [71]. It was required that the LUMI monitor was operational, the solenoid was turned on, the CTD (see Section 3.4.3) had high voltage and no large dead regions, and the UCAL (see Section 3.4.4) did not have too many dead channels.

During the 1996-97 running period (see Section 3.2.3), HERA delivered $53.5\,\mathrm{pb}^{-1}$ of luminosity, of which $38.6\,\mathrm{pb}^{-1}$ was recorded by ZEUS after the EVTAKE selection.

6.2 Online Event Selection (Trigger)

The online event selection was performed by the ZEUS three level trigger (see Section 3.5). Each level of the trigger, with decreasing speed and increasing strictness, applied selection criteria to remove background and select a phase space characteristic of photoproduction. Various selection criteria at each trigger level are grouped together to form **trigger slots**, which correspond to one bit in a 32-bit word of trigger information. The trigger slots at the FLT are numbered 0 to 63, while the trigger slots at the SLT and TLT level are arranged according to the ZEUS analysis working groups. For example, HPP SLT 1, is the first SLT slot assigned to the Hard Photoproduction working group[1].

When taken together, a series of trigger slots at the FLT, SLT, and TLT levels required to accept an event is called a **trigger chain**. As detailed below, the trigger chain for this analysis is FLT 42, HPP SLT 1, and HPP TLT 14. DST Bit 77, which corresponds to HPP TLT 14, was applied at the EAZE level, but the FLT and SLT requirements were selected in the analysis code specific to this analysis. Detector level Monte Carlo (see Section 4.2.6) includes a trigger simulation by CZAR and was required to pass the same trigger chain as the data.

6.2.1 FLT Slot 42

In FLT Slot 42, the GFLT selects events based upon global and regional energy sums provided by the CFLT, simple tracking information provided by the CTD FLT, and vetos provided by various other components. An accepted event must meet the following

---

[1] The names of the ZEUS working groups has evolved over the years but the names of the trigger slots have not.

criteria:

- Calorimeter Energy Requirement[2]
  - Total UCAL Energy > 14.68 GeV;
  - **OR** Total EMC Energy > 10.068 GeV;
  - **OR** BCAL EMC Energy > 3.404 GeV;
  - **OR** RCAL EMC Energy > 2.032 GeV.
- At least one "good track" in the CTD, where a good track is defined as having a z-position of $-50\,\text{cm} < z < 80\,\text{cm}$ in the first superlayer of the CTD.
- **Pass Vetos**
  - Information from ep collisions reaches the C5 counter (see Section 3.4.5) in 3 ns and therefore events outside a 6 ns window are classified as beam-gas and rejected;
  - The SRTD (see Section 3.4.4) also rejects events which have a timing characteristic of beam-gas;
  - The Veto Wall rejects events in which there is a coincidence between the sides of the wall facing towards and facing away from the detector. This coincidence is characteristic of halo muon events and beam gas events, which, unlike ep collisions, can deposit energy on the side of the wall facing away from the detector.

---

[2] Excluding energy in the 3 innermost rings of the FCAL and the innermost ring of the RCAL.

6.2.2 SLT HPP Slot 1

SLT HPP Slot 1 is designed to select photoproduction events containing jets with large transverse energies. It requires that an event satisfy all of the following:

- Pass FLT Slot 40, 41, 42, or 43;
- A vertex reconstructed from the CTD with $|Z_{vtx}| < 60\,\text{cm}$ **OR** no reconstructed vertex[3];
- At least 1 vertex track;
- $\sum E - p_z > 8\,\text{GeV}$, where the sum is over all calorimeter cells. This rejects beam-gas events which deposit energy only in the forward regions;
- $\sum E_{\text{T}}\,(-1ir) > 8\,\text{GeV}$, where the sum is over all calorimeter cells except those in the first inner ring around the FCAL beam-pipe. This selects events with energy deposits in the UCAL characteristic of jets with high transverse momentum.
- $\sum E - p_z > 12\,\text{GeV}$ **OR** $\sum p_z/E < 0.95$. This further rejects beam-gas events while preserving photoproduction events.

In addition, global timing cuts are applied to all SLT slots. The readout systems of the ZEUS sub-components are synchronized to the HERA clock so that events occurring at the nominal interaction point result in a measured event time of 0. Timing cuts therefore reject background originating somewhere other than in the expected interaction point. If the timing information exists for an event, events must satisfy the following:

---

[3] In the case of no reconstructed vertex, Z_{vtx}, is set to 0.

- $t_{\mathrm{RCAL}} > -8\,\mathrm{ns}$. Rejects upstream beam-gas events, which normally reach the RCAL approximately 10 ns before *ep* events from the nominal vertex.
- $t_{\mathrm{FCAL}} - t_{\mathrm{RCAL}} < 8\,\mathrm{ns}$. Rejects upstream beam-gas events which first pass through the RCAL and then the FCAL.
- $t_{\mathrm{top}} - t_{\mathrm{down}} > -10\,\mathrm{ns}$. Rejects cosmic ray events, which pass from the top through the bottom of the detector in approximately 12 ns.

6.2.3 TLT HPP Slot 14

HPP TLT 14 uses the full event information to select dijet photoproduction events by requiring that each event satisfy the following:

- Pass HPP SLT Slots 1, 2, or 3, **OR** pass the SLT Special Slot;
- Two or more jets, each with $E_{\mathrm{T}} \geq 4\,\mathrm{GeV}$ and $\eta < 2.5$. Jet finding is performed using the EUTLT cone algorithm, which is a modified version EUCELL (see Section 5.7.2).

In addition, the following global TLT requirements must be satisfied:

- The event has an reconstructed vertex with $|Z_{vtx}| < 60\,\mathrm{cm}$. Even though the SLT and TLT vertex requirements are the same, the TLT requirement is more refined because more tracking information is available at the TLT.
- Less than 6 “Bad Tracks”. A bad track is associated with beam-gas and does *not* satisfy one of the following:

  - Number of degrees of freedom > 20, where the number of degrees of freedom is the number of CTD hits minus the number of parameters in the track fit;
  - $p_T \geq 2\,\mathrm{GeV}$;
  - $-3.13 < \eta < 1.75$;
  - Number of hits in the CTD axial superlayers > 5;
  - Number of hits in the CTD stereo superlayers > 5;
  - Distance of closest approach to the vertex in $z \leq 75\,\mathrm{cm}$.
- $5.0\,\mathrm{GeV} < E - p_z < 75.0\,\mathrm{GeV}$ **AND** $p_z/E < 1.0$.
- $E_{\mathrm{T}}^{\mathrm{Cone}} < 5.0\,\mathrm{GeV}$, where $E_{\mathrm{T}}^{\mathrm{Cone}}$ is the sum of all transverse energy in the calorimeter except in a 10° cone around the FCAL beam-pipe.
- Timing information not characteristic of beam-gas and cosmic ray backgrounds
  - $t_{\mathrm{RCAL}} > -6\,\mathrm{ns}$;
  - $t_{\mathrm{FCAL}} < 8\,\mathrm{ns}$;
  - $t_{\mathrm{FCAL}} - t_{\mathrm{RCAL}} < 6\,\mathrm{ns}$;
  - $t_{\mathrm{Global}} < 8\,\mathrm{ns}$, where t_{Global} is an average time calculated using energy weighted means from all calorimeter cells above a threshold of 200 MeV.

6.3 Offline Event Selection

The data is stored on tape after the online event selection and is reconstructed according to the methods described in Chapter 5. More accurate cuts are then applied to the fully reconstructed events at the EAZE level and in the code specific to this analysis in order to further reduce the background, select the kinematic range, and choose events from this kinematic range specific to this analysis.

6.3.1 Selection of a Clean Photoproduction Sample

A clean photoproduction sample was selected by removing background not rejected by the trigger and removing NC and CC DIS events. The following selection criteria were applied offline:

- the longitudinal position of the reconstructed vertex was required to be in the range $-40\,\mathrm{cm} < Z_{vtx} < 40\,\mathrm{cm}$ in order to reduce contributions from beam-gas interactions, cosmic-ray showers, and beam-halo muons;
- events with a scattered positron in the UCAL having a Sinistra (see Section 5.4) probability greater than 0.9 and $y_e < 0.85$ and $E'_e > 5\,\mathrm{GeV}$, where E'_e is the energy of the scattered positron, were rejected. This cut reduced contamination from NC DIS events in the region where efficiency for the scattered positron to be detected approached 100%;
- events were required to have $0.2 < y_{\mathrm{JB}} < 0.75$. The upper cut on y_{JB} further reduced contamination from neutral current DIS events which were not removed by the cut on y_e and the lower cut removed beam-gas events, as described in Section 5.5.2;
- events were required to have a relative transverse momentum $\frac{\not{p}_{\mathrm{T}}}{\sqrt{E_{\mathrm{T}}}} < 2\,\mathrm{GeV}^{1/2}$, where $\not{p}_T$ (called "missing p_{T}") and E_{T} are calculated by summing the transverse momentum and energy in the UCAL. This cut reduces the contribution of CC DIS events and further reduces the contribution from cosmic-ray showers. The incoming lepton and proton beams have negligible transverse components and in NC DIS, where the lepton is detected, and in photoproduction, where the lepton travels at a small angle undetected down the beam-pipe, the final measured $\not{p}_T$ should be close to zero. However, in CC DIS, the undetected neutrino must be balanced by particles having a transverse momentum component, and therefore the magnitude of $\frac{\not{p}_{\mathrm{T}}}{\sqrt{E_{\mathrm{T}}}}$ is substantial. The energy resolution of the calorimeter, which is proportional to $\sqrt{E_{\mathrm{T}}}$, is accounted for by scaling $\not{p}_{\mathrm{T}}$ by $1/\sqrt{E_{\mathrm{T}}}$.

The cuts on y_e and y_{JB} reduced the contribution of DIS events to less than 0.5% and restricted the photon virtuality to a range of $Q^2 < 1\,\mathrm{GeV}^2$ with a median value of $Q^2 \sim 10^{-3}\,\mathrm{GeV}^2$ [72]. There were 3.1×10^6 data events in the clean photoproduction sample.

6.3.2 Selection of Inclusive Dijet Photoproduction Events

Jets were reconstructed from the ZUFOs using the k_T algorithm [73] in the longitudinally invariant inclusive mode [74] and ordered in E_{T}. A sample of high E_{T} dijet events well separated in pseudorapidity was selected by requiring the following:

- $E_{\mathrm{T}}^{jet1} \geq 5.1\,\mathrm{GeV}$ and $E_{\mathrm{T}}^{jet2} \geq 4.25\,\mathrm{GeV}$, where E_T^{jet1} and E_T^{jet2} are the transverse energy of the jets with the highest E_{T} satisfying the η requirement listed in the next bullet. The high cuts on the jet E_{T} ensures that processes can be calculated using pQCD and their asymmetric values reduce the region of phase space where NLO QCD calculations suffer from incomplete cancellations of real and virtual contributions.

- $-2.4 < \eta^{jet1,2} < 2.4$, where η^{jet1} and η^{jet2} are the pseudorapidities of the two jets with the highest E_{T} satisfying this requirement. The most forward jet of the pair is called the **leading jet** and the most backward jet of pair is called the **trailing jet**. This cut ensure that the jets were well constructed in the UCAL.

- $|\overline{\eta}| \equiv \frac{1}{2}|\,(\eta^{jet1} + \eta^{jet2})\,| < 0.75$, to ensure that a large fraction of photon energy has participated in the hard interaction. This enables a better MC simulation and less uncertainty from the photon PDFs, decreases the likelihood of multi-parton interactions by reducing the probability of large non-perturbative corrections, and ensured that events are more uniformly distributed in η-space.

- $2.5 < \Delta\eta < 4$, where $\Delta\eta \equiv |\eta^{jet1} - \eta^{jet2}|$ is the absolute difference in pseudorapidity between the leading and trailing jets, so that the jets are well separated in pseudorapidity.

Events passing the above cuts are termed **inclusive events**. There are $56,211$ inclusive events in the data sample used in this analysis.

6.3.3 Selection of Gap Events

The total transverse energy of all jets lying in the pseudorapidity region between the leading and trailing jets is defined as $E_{\mathrm{T}}^{\mathrm{GAP}}$

$$E_{\mathrm{T}}^{\mathrm{GAP}} \equiv \sum_{i>2} E_{\mathrm{T}}^{jeti} \ , \ \eta^{jet}_{trailing} < \eta^{jeti} < \eta^{jet}_{leading}. \tag{6.1}$$

A **gap event** is defined as an event in which $E_{\mathrm{T}}^{\mathrm{GAP}}$ is less than an $E_{\mathrm{T}}^{\mathrm{CUT}}$ value. $E_{\mathrm{T}}^{\mathrm{GAP}}$ must be small enough compared to the E_{T} of the leading and trailing jets in order for the event to retain gap-like properties. The number of gap events for the four different $E_{\mathrm{T}}^{\mathrm{CUT}}$ values, $E_{\mathrm{T}}^{\mathrm{CUT}} = 0.6, 1.2, 1.8, 2.4\,\mathrm{GeV}$, used in this analysis are listed in Table 6.1.

$E_{\mathrm{T}}^{\mathrm{CUT}}$	**Number of Gap Events**
0.6	1500
1.2	3554
1.8	6265
2.4	9432

Table 6.1: The number of gap events in the data sample for four different values of $E_{\mathrm{T}}^{\mathrm{CUT}}$.

A gap event definition based on E_{T} has both theoretical and experimental advantages over the definition based on particle multiplicity in the gap [75]. From a theoretical standpoint, a definition based on E_{T} is both collinear and infrared safe. Also, since the E_{T} is calculated from k_{T} jet clusters instead of individual particles, corrections for hadronization effects are reduced. Although color-singlet exchange produces no QCD radiation in a gap event, hadronic activity resulting from the color connections between the leading jet and proton remnant and the trailing jet and photon remnant “leak” into the gap region. The definition based on E_{T} determined from the k_{T} cluster algorithm is more likely than multiciplicity definitions to combine the hadronic activity at the edges of leading and trailing jets with jets lying outside the rapidity gap.

Experimentally, the use of a gap definition based on E_{T} determined from a jet algorithm allows the pseudorapidity range of the gap to span between the centers of the leading and trailing jets, while the multiplicity definitions only allows for a pseudorapidity range spanning between the inner edges of the jets. The increased gap range expands the region in $\eta - \phi$ space in which rapidity gap events may be observed

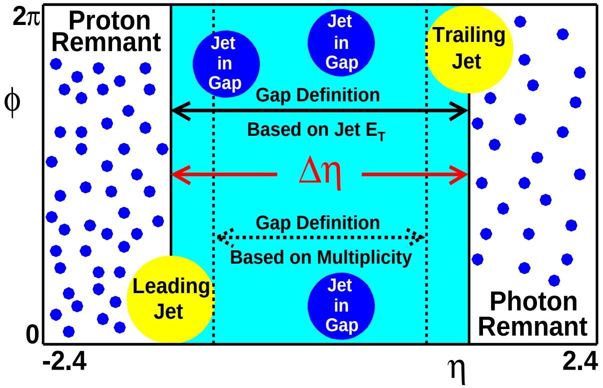


Figure 6.1: The topology of a rapidity gap event in photoproduction at HERA. Gap definitions based on Jet E_T and multiplicity are shown. For the definition based on Jet E_T, $\Delta\eta$ is defined as the distance between the centers of the leading and trailing jets. $E_{\rm T}^{\rm GAP}$ is calculated by summing the transverse energy of all jets in the $\eta - \phi$ space, indicated by the shading, between the leading and trailing jets.

and therefore increases the statistics of the gap event sample. Another advantage of this definition is that jet energies are easier to calibrate than particle energies.

The topology of a gap event with the gap defined as the pseudorapidity range between the centers of the leading and trailing jets is illustrated in Fig. 6.1. A typical gap event, as depicted in the ZEUS Event Display, is shown in Fig. 6.2.

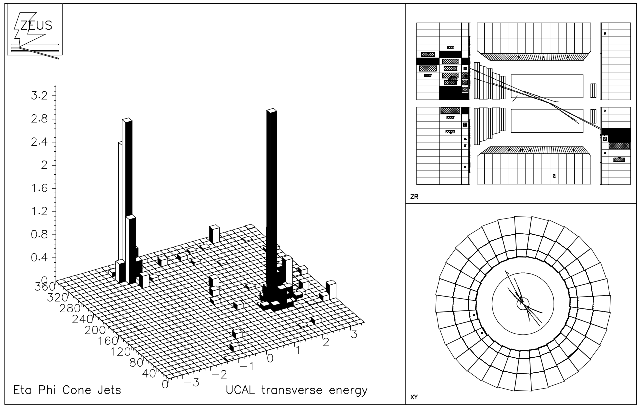


Figure 6.2: A rapidity gap event in the ZEUS Event Display.

Chapter 7

Modeling of Photoproduction Data

An accurate detector level Monte Carlo description of the data sample is necessary to obtain a cross section measurement at the hadron level by correcting the data for detector effects and to have reasonable systematic uncertainties on this cross section measurement. Pythia version 6.1 and Herwig version 6.1 (see Section 4.3) interfaced through AMADEUS (see Section 4.2.6) were used to simulate hard photoproduction events. Direct, resolved, and color-singlet MC samples were generated separately and passed through a full simulation of the ZEUS detector and trigger (see Section 4.2.6).

7.1 Tuning the Monte Carlo

Despite the all the successes of QCD, the production of the hadronic final state is still difficult to calculate and it is necessary to use general purpose MC generators, such as Pythia and Herwig, to obtain a detailed description. The interpretation of measurements at past, present, and future colliders relies on an accurate tuning of the free input parameters of the MC. However, determining which MC models are consistent with which experimental results is difficult since many different variables are measured with a variety of colliding beams and in various regions of kinematic phase space. The JetWeb [76] facility provides a means to compare a broad range of kinematic data with various models through a web-based interface.

Monte Carlo			**PDFs**		**Default Parameters**			
MC	**Process**	**MPI**	**PDF(p)**	**PDF(γ)**	$\mathbf{p}_{\mathrm{T}}^{\mathrm{min1}}$	$\mathbf{p}_{\mathrm{T}}^{\mathrm{min2}}$	$\mathbf{P}_{\mathrm{SUE}}$	$\mathbf{R}_p$
Pyt	Direct	-	GRV 94 LO	GRV G LO	2.0	-	-	-
	Resolved	Yes	GRV 94 LO	GRV G LO	2.0	1.5	-	-
	High-t γ	No	GRV 94 LO	GRV G LO	2.0	-	-	-
Her	Direct	-	GRV 94 LO	WHIT-G 2	1.8	-	-	1.0
	Resolved	Yes	GRV 94 LO	WHIT-G 2	1.8	-	1.0	1.0
	BFKL	Yes	GRV 94 LO	WHIT-G 2	1.8	-	1.0	1.0

Table 7.1: The standard free parameters used as input for the generation of Pythia and Herwig. The minimum transverse momentum of the hard interaction is set by $p_{\mathrm{T}}^{\mathrm{min1}}$ and the minimum transverse momentum of the MPIs is set by $p_{\mathrm{T}}^{\mathrm{min2}}$, where $p_{\mathrm{T}}^{\mathrm{min1}}$ and $p_{\mathrm{T}}^{\mathrm{min2}}$ are listed in GeV. In Pythia 6.1, both parameters are adjustable, while in Herwig 6.1 only $p_{\mathrm{T}}^{\mathrm{min1}}$ can be changed. $\mathrm{P}_{\mathrm{SUE}}$ is the probability of the Soft Underlying Event and R_p is the square of the factor by which the proton radius is reduced.

The parameters determined by JetWeb's global fits to various collider data were used as the starting point for tuning the MC to the rapidity gap data in this analysis. For Pythia 6.1, the transverse momentum of the hardest interaction ($p_{\mathrm{T}}^{\mathrm{min1}}$) and transverse momentum of all secondary interactions ($p_{\mathrm{T}}^{\mathrm{min2}}$) were tuned to give the best agreement between the MC and hadron level data (see Section 8.1) for the $E_{\mathrm{T}}^{\mathrm{GAP}}$ distribution. For Herwig 6.1, only $p_{\mathrm{T}}^{\mathrm{min1}}$ was adjustable. The MC used in the tuning was generated using HZTOOL [77]. The standard free parameters are listed in Table 7.1 and the tuned parameters used in this analysis are listed in Table 7.2.

All further mention of Pythia and Herwig refer to the tuned samples of these MCs.

Monte Carlo			PDFs		Tuned Parameters			
MC	**Process**	**MPI**	**PDF(p)**	**PDF(γ)**	$\mathbf{p_T^{min1}}$	$\mathbf{p_T^{min2}}$	$\mathbf{P_{SUE}}$	$\mathbf{R_p}$
PYT	Direct	-	CTEQ 5L	SaS-G 2D	1.9	-	-	-
	Resolved	Yes	CTEQ 5L	SaS-G 2D	1.9	1.7	-	-
	High-t γ	No	CTEQ 5L	SaS-G 2D	1.9	-	-	-
HER	Direct	-	CTEQ 5L	SaS-G 2D	2.7	-	-	3.0
	Resolved	Yes	CTEQ 5L	SaS-G 2D	2.7	-	0.03	3.0
	BFKL	Yes	CTEQ 5L	SaS-G 2D	2.7	-	0.03	3.0

Table 7.2: The tuned free parameters used as input for the generation of PYTHIA and HERWIG. The minimum transverse momentum of the hard interaction is set by p_T^{min1} and the minimum transverse momentum of the MPIs is set by p_T^{min2}, where p_T^{min1} and p_T^{min2} are listed in GeV. In PYTHIA 6.1, both parameters are adjustable, while in HERWIG 6.1 only p_T^{min1} can be changed. P_{SUE} is the probability of the Soft Underlying Event and R_p is the square of the factor by which the proton radius is reduced.

7.2 The Monte Carlo Sample

The separately generated direct, resolved, and color-singlet detector level MC samples were combined into one final MC sample to which the data was compared.

7.2.1 Mixing the Direct and Resolved MC

The direct and resolved samples were mixed according to fractions obtained from a χ^2 minimization (see Appendix B) of the inclusive data and the direct and resolved MC x_γ^{OBS} distributions. The combination of direct and resolved MC formed a color-non-singlet MC sample. The fit fractions for the direct and resolved samples are listed in Table 7.3 and shown in Fig. 7.3 for both PYTHIA and HERWIG. These mixing ratios were used in all subsequent detector level comparisons between the data and MC.

It can be seen in Fig. 7.3 that the data distribution peaks at approximately $x_\gamma^{OBS} = 0.75$, which means that the hadron involved in the hard interaction usually carries a large fraction of the photon's momentum. However, the shaded histogram, which shows the contribution of direct interactions, indicates that resolved interactions do make a significant contribution, even at high x_γ^{OBS}. The dashed line at $x_\gamma^{OBS} = 0.75$ divides the sample into direct and resolved enhanced regions.

MC	Direct	Resolved
PYTHIA	28%	72%
HERWIG	42%	58%

Table 7.3: The fractions of detector level direct and resolved MC giving the best fit to the data for the inclusive x_γ^{OBS} distribution.

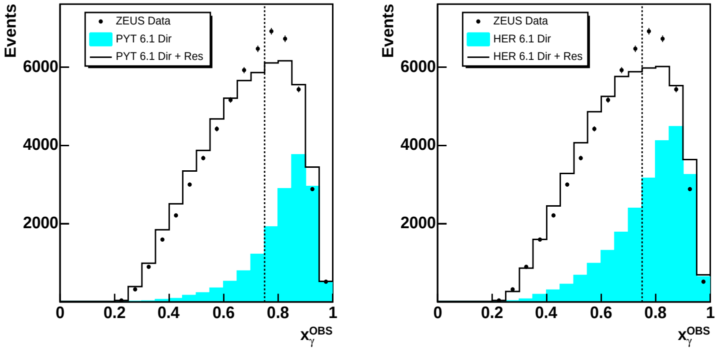


Figure 7.1: The fractions of detector level direct and resolved MC giving the best fit to the data for the inclusive x_γ^{OBS} distribution. The black points show the ZEUS data, the shaded histogram indicates the contribution of direct photoproduction, and the solid line indicates the combination of direct and resolved photoproduction. The dashed line at $x_\gamma^{OBS} = 0.75$ divides the sample into direct and resolved enhanced regions. The fit using PYTHIA is shown in the left plot and the fit using HERWIG is shown in the right plot.

7.2.2 Mixing the Color-Non-Singlet and Color-Singlet MC

Color-singlet MC was added to the color-non-singlet MC sample obtained by mixing the direct and resolved MC samples. Color-singlet events are simulated in HERWIG

	$E_{\mathrm{Tot}}^{\mathrm{GAP}}$ for $E_{\mathrm{T}}^{\mathrm{GAP}} < 1.5\,\mathrm{GeV}$		**Inclusive**	
MC	**Non-CS**	**CS**	**Non-CS**	**CS**
PYTHIA	74%	26%	96%	4%
HERWIG	77%	23%	94%	6%

Table 7.4: The fractions of detector level color-non-singlet and color-singlet MC giving the best fit to the data for the $E_{\mathrm{Tot}}^{\mathrm{GAP}}$ distribution for events in which $E_{\mathrm{T}}^{\mathrm{GAP}} < 1.5\,\mathrm{GeV}$. The fit result in the region that $E_{\mathrm{T}}^{\mathrm{GAP}} < 1.5\,\mathrm{GeV}$ and the corresponding fractions for the inclusive sample are listed.

by BFKL Pomeron exchange and in PYTHIA by high-t photon exchange. The high-t photon exchange is only used to compare the data and MC predictions and is not expected to describe the mechanism of color-singlet exchange (See Section 4.3).

The fractions of color-non-singlet and color-singlet MC were determined by a χ^2 minimization using the total gap energy distribution, $E_{\mathrm{Tot}}^{\mathrm{GAP}}$, for events in which $E_{\mathrm{T}}^{\mathrm{GAP}} < 1.5\,\mathrm{GeV}$. The color-non-singlet and color-singlet MC samples were fit to the data to form a mixed color-non-singlet and color-singlet sample. The region $E_{\mathrm{T}}^{\mathrm{GAP}} < 1.5\,\mathrm{GeV}$ was chosen to perform the fit because color-singlet exchange is expected to dominate when $E_{\mathrm{T}}^{\mathrm{GAP}}$ is small. The fit was performed using $E_{\mathrm{Tot}}^{\mathrm{GAP}}$ because this quantity is well described in the low $E_{\mathrm{T}}^{\mathrm{GAP}}$ region.

Table 7.4 lists the results of this fit and the corresponding fractions of color-singlet and color-non-singlet MC for the inclusive samples of both PYTHIA and HERWIG. The results of the fit are also shown in Fig 7.2, where it can be seen that the mixture of the color-non-singlet and color-singlet samples describes the data better than the color-non-singlet sample alone.

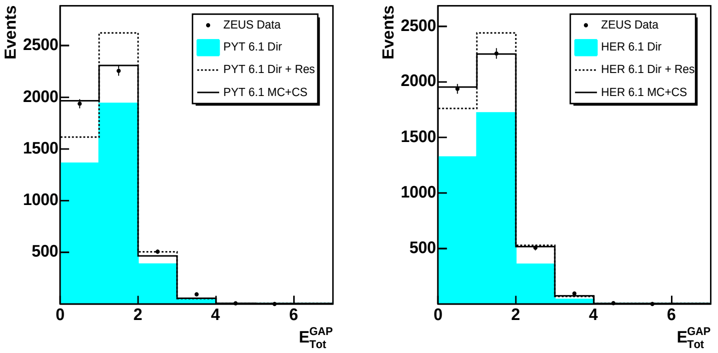


Figure 7.2: The fractions of detector level color-non-singlet and color-singlet MC giving the best fit to the data for the $E_{\mathrm{Tot}}^{\mathrm{GAP}}$ distribution in the region $E_{\mathrm{T}}^{\mathrm{GAP}} < 1.5\,\mathrm{GeV}$. The black points show the ZEUS data, the shaded histogram indicates the contribution of direct photoproduction, the dashed line indicates the combination of direct and resolved photoproduction, and the solid line indicates the mixed sample resulting from the combination of the direct and resolved samples with the color-singlet sample. The fit using PYTHIA is shown in the left plot and the fit using HERWIG is shown in the right plot.

7.3 The Monte Carlo Description of the Data

The direct photoproduction sample, the color-non-singlet sample, and the mixed color-non-singlet and color-singlet sample were compared to the data. In each plot, the data is signified by the points, the direct sample is signified by the shaded histogram, the color-non-singlet sample by the dashed line, and the mixed sample by the solid line.

The vertical dashed lines depict the location of the cuts in the offline selection. If the plotted variable was used as a selection criteria in the offline cuts described in Chapter 5, all cuts, except for that on the plotted variable were applied[1]. The MC was area normalized to the data in the region selected by the offline cuts (between the lines showing the cut locations).

An accurate description of these variables is necessary to have confidence that the kinematic phase space is well described and to minimize the systematic error on the cross section measurement.

7.3.1 The Monte Carlo Description of Inclusive Events

The inclusive sample of events has no requirement on the energy between the dijets separated in pseudorapidity and therefore is expected to be described by the the color-non-singlet MC alone.

PYTHIA

Figure 7.3 shows the the highest E_{T} jet (E_{T}^{1}), the second highest E_{T} jet (E_{T}^{2}), the leading jet (η^{1}), the trailing jet (η^{2}), the separation between the leading and trailing jets ($\Delta\eta$), and the average η of the leading and trailing jets ($|\overline{\eta}|$). The MC matches the data well in all plots and there is regular behavior at the cut boundaries.

Figure 7.4 shows the fractional missing transverse momentum ($\frac{\not{p}_{\mathrm{T}}}{\sqrt{E_{\mathrm{T}}}}$), the longitudinal vertex (Z_{vtx}), y reconstructed using the Jacquet-Blondel method (y_{JB}), $E_{\mathrm{T}}^{\mathrm{GAP}}$, and $x_{\gamma}^{\mathrm{OBS}}$. $\frac{\not{p}_{\mathrm{T}}}{\sqrt{E_{\mathrm{T}}}}$ is well described for $\frac{\not{p}_{\mathrm{T}}}{\sqrt{E_{\mathrm{T}}}} > 2\,\mathrm{GeV}^{1/2}$, but for $\frac{\not{p}_{\mathrm{T}}}{\sqrt{E_{\mathrm{T}}}} < 2\,\mathrm{GeV}^{1/2}$, the MC underestimates the data. This disagreement is most likely due to the contamination of CC DIS events. Z_{vtx} is well described, as is y_{JB}, except for in the region $y_{\mathrm{JB}} > 0.85$. This disagreement, approximately 60% at $y_{\mathrm{JB}} = 0.85$, is most likely due to contamination of NC DIS events in the data which were not removed by the cleaning cut based on electron finding. $E_{\mathrm{T}}^{\mathrm{GAP}}$ is fairly well described, except for the lowest few bins, although when the MC is normalized to the data in the low $E_{\mathrm{T}}^{\mathrm{GAP}}$ region alone, there is improved agreement. The $x_{\gamma}^{\mathrm{OBS}}$ plot is the same as that shown in Fig. 7.1 and was included to show that the addition of color-singlet exchange MC does not improve the agreement at the inclusive level. The small fraction of color-singlet MC at the inclusive level, as determined by the fit to $E_{\mathrm{Tot}}^{\mathrm{GAP}}$ in the region that $E_{\mathrm{T}}^{\mathrm{GAP}} < 1.5\,\mathrm{GeV}$ (see Table 7.4), does not significantly affect the agreement in any of the plots and confirms the prediction that the fraction of color-singlet exchange processes in the inclusive sample is small.

HERWIG

Figure 7.5 shows that the MC agrees with the data for the quantities E_{T}, η, $\Delta\eta$, and $|\overline{\eta}|$. E_{T}^{1} is better simulated with Herwig than Pythia and the other variables are simulated equally well by Herwig and Pythia.

It can be seen in Fig. 7.5 that the Herwig gives a better description of $\frac{\not{p}_{\mathrm{T}}}{\sqrt{E_{\mathrm{T}}}}$ than Pythia, especially in the region $\frac{\not{p}_{\mathrm{T}}}{\sqrt{E_{\mathrm{T}}}} > 2\,\mathrm{GeV}^{1/2}$. The descriptions by Herwig

---

[1] Except for E_{T} and η, which were not possible to decouple because the event sample requires the two highest jets in E_{T} which satisfy the η requirements and not just the two highest jets in E_{T}.

and PYTHIA are similar for Z_{vtx}, $y_{\rm JB}$, $E_{\rm T}^{\rm GAP}$, and $x_\gamma^{\rm OBS}$, with disagreement for large $y_{\rm JB}$ and small $E_{\rm T}^{\rm GAP}$ existing in both.

Once again, an insignificant fraction of color-singlet exchange MC, as determined by the fit to $E_{\rm Tot}^{\rm GAP}$ in the region that $E_{\rm T}^{\rm GAP} < 1.5\,{\rm GeV}$ (see Table 7.4), was needed to improve agreement with the data.

7.3.2 The Monte Carlo Description of Gap Events

The gap sample is defined by requiring the transverse energy in the pseudorapidity region between the leading and trailing jets, $E_{\rm T}^{\rm GAP}$, to be less than a certain value $E_{\rm T}^{\rm CUT}$. This sample is expected to contain a significant fraction of color-singlet processes, and therefore a substantial amount of color-singlet MC is expected to be required to match the data. Comparisons between data and MC for $E_{\rm T}^{\rm CUT} = 1.2\,{\rm GeV}$ are shown.

PYTHIA

Figure 7.7 shows $E_{\rm T}^1$, $E_{\rm T}^2$, η^1, η^2, $\Delta\eta$, and $|\overline{\eta}|$ for $E_{\rm T}^{\rm GAP} < 1.2\,{\rm GeV}$. The MC matches the data well in all plots and there is regular behavior at the cut boundaries. The addition of the color-singlet MC significantly improves the agreement for η^1, η^2, $\Delta\eta$, and $|\overline{\eta}|$.

Figure 7.8 shows $\frac{\not{p}_{\rm T}}{\sqrt{E_{\rm T}}}$, Z_{vtx}, $y_{\rm JB}$, $E_{\rm T}^{\rm GAP}$, and $x_\gamma^{\rm OBS}$ for $E_{\rm T}^{\rm GAP} < 1.2\,{\rm GeV}$. The level of agreement for $\frac{\not{p}_{\rm T}}{\sqrt{E_{\rm T}}}$ and Z_{vtx} is similar to that in the inclusive sample, with the region $\frac{\not{p}_{\rm T}}{\sqrt{E_{\rm T}}} > 2\,{\rm GeV}^{1/2}$ once again showing disagreement due to CD DIS contamination. There is a more pronounced disagreement at large $y_{\rm JB}$ than the inclusive sample, and once again this can be attributed to NC DIS contamination. The $E_{\rm T}^{\rm GAP}$ distribution for the gap sample is described, unlike in the inclusive sample (see Fig. 7.4), where the MC was normalized to the data over the entire range of $E_{\rm T}^{\rm GAP}$ and disagrees and with the data in the lowest few $E_{\rm T}^{\rm GAP}$ bins. The $x_\gamma^{\rm OBS}$ distribution is also well described by the mixed MC sample.

HERWIG

Figure 7.9 shows $E_{\rm T}^1$, $E_{\rm T}^2$, η^1, η^2, $\Delta\eta$, and $|\overline{\eta}|$ for $E_{\rm T}^{\rm GAP} < 1.2\,{\rm GeV}$. HERWIG describes all variables well and has approximately the same level of agreement with the data as PYTHIA. The addition of the color-singlet MC improves the agreement for η^1, η^2, $\Delta\eta$, and $|\overline{\eta}|$.

Figure 7.10 shows $\frac{\not{p}_{\rm T}}{\sqrt{E_{\rm T}}}$, Z_{vtx}, $y_{\rm JB}$, $E_{\rm T}^{\rm GAP}$, and $x_\gamma^{\rm OBS}$ for $E_{\rm T}^{\rm GAP} < 1.2\,{\rm GeV}$. The level of agreement with HERWIG is comparable to that with PYTHIA, except for $\frac{\not{p}_{\rm T}}{\sqrt{E_{\rm T}}}$, which HERWIG describes better for $\frac{\not{p}_{\rm T}}{\sqrt{E_{\rm T}}} > 2\,{\rm GeV}^{1/2}$. The addition of the color-singlet MC improves the agreement with the data, especially for $E_{\rm T}^{\rm GAP}$ and $x_\gamma^{\rm OBS}$.

7.3.3 The Monte Carlo Description of NC DIS Contamination

In order to ensure that the data rejected because a $y_{\rm JB}$ value was greater than 0.75 originated from NC DIS, the data was compared to a NC DIS sample of MC generated using ARIADNE [78]. ARIADNE relies on the Color Dipole Model [79, 80, 81, 82] to simulate QCD radiation and the String Model to simulate hadronization, and is commonly used in the simulation of DIS events. The Color Dipole Model assumes that all initial and final state radiation occurs from color dipoles formed between the struck quark and proton remnant. The String Hadronization Model is described in Section 4.2.5.

Comparisons between the data and ARIADNE are shown in Fig. 7.11 for the inclusive sample and each of the four different gap samples. The MC are area normalized to the data between $y_{\rm JB} = 0.85$ (indicated by the vertical dashed line) and $y_{\rm JB} = 1.0$, the region where DIS events are expected to dominate over photoproduction events. The MC describes the data well in this region, and this confirms the prediction that the excess in the data over PYTHIA and HERWIG seen at high y_{JB} is indeed due to DIS contamination and that the requirement of $y_{\mathrm{JB}} < 0.75$ removes most of this contamination.

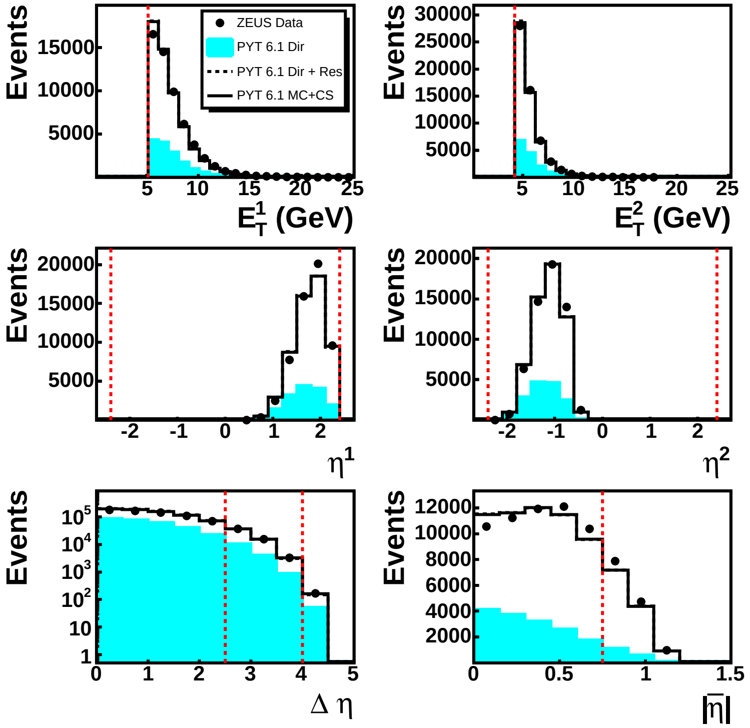


Figure 7.3: Description of the inclusive data sample by PYTHIA for the variables E_{T}, η, $\Delta\eta$, and $|\overline{\eta}|$. The ZEUS data is depicted by the points, the direct MC sample by the shaded histogram, the color-non-singlet MC sample by the dashed line, and the mixed MC sample by the solid line. The dashed vertical lines show the cuts applied in the offline selection. The MC was area normalized to the data in the region selected by the offline cuts.

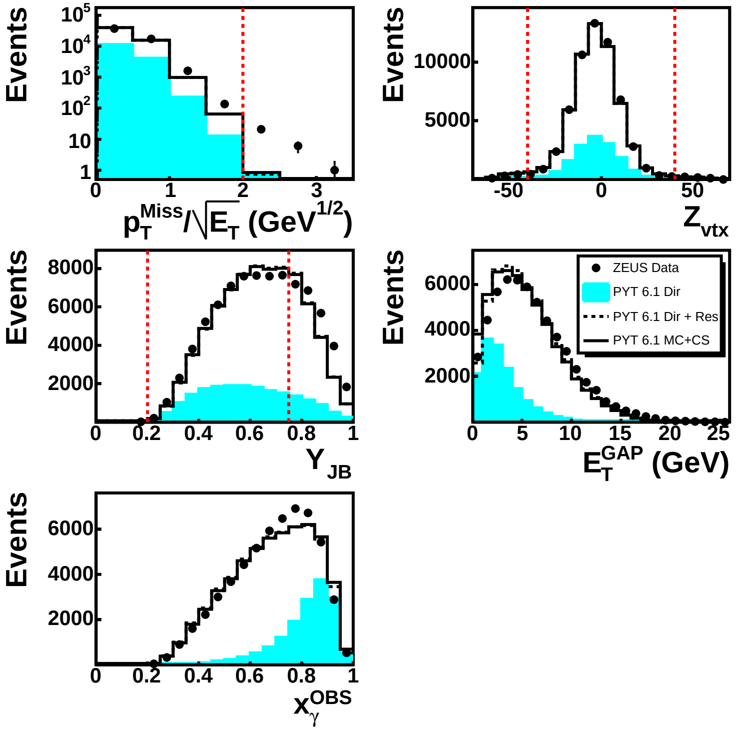


Figure 7.4: Description of the inclusive data sample by PYTHIA for the variables $\frac{\not{p}_T}{\sqrt{E_T}}$, Z_{vtx}, y_{JB}, E_T^{GAP}, and x_γ^{OBS}. The ZEUS data is depicted by the points, the direct MC sample by the shaded histogram, the color-non-singlet MC sample by the dashed line, and the mixed MC sample by the solid line. The dashed vertical lines show the cuts applied in the offline selection. The MC was area normalized to the data in the region selected by the offline cuts.

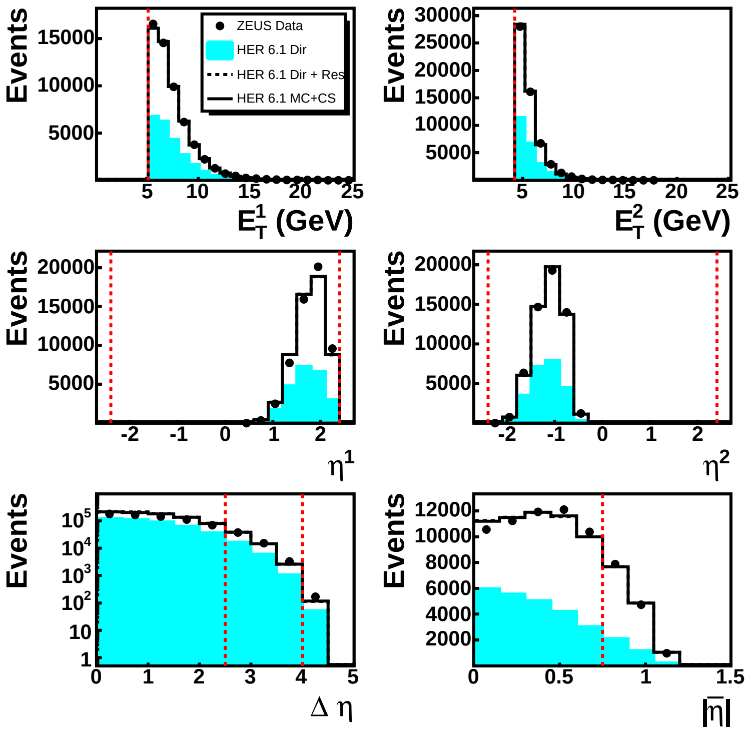


Figure 7.5: Description of the inclusive data sample by HERWIG for the variables E_T, η, $\Delta\eta$, and $|\overline{\eta}|$. The ZEUS data is depicted by the points, the direct MC sample by the shaded histogram, the color-non-singlet MC sample by the dashed line, and the mixed MC sample by the solid line. The dashed vertical lines show the cuts applied in the offline selection. The MC was area normalized to the data in the region selected by the offline cuts.

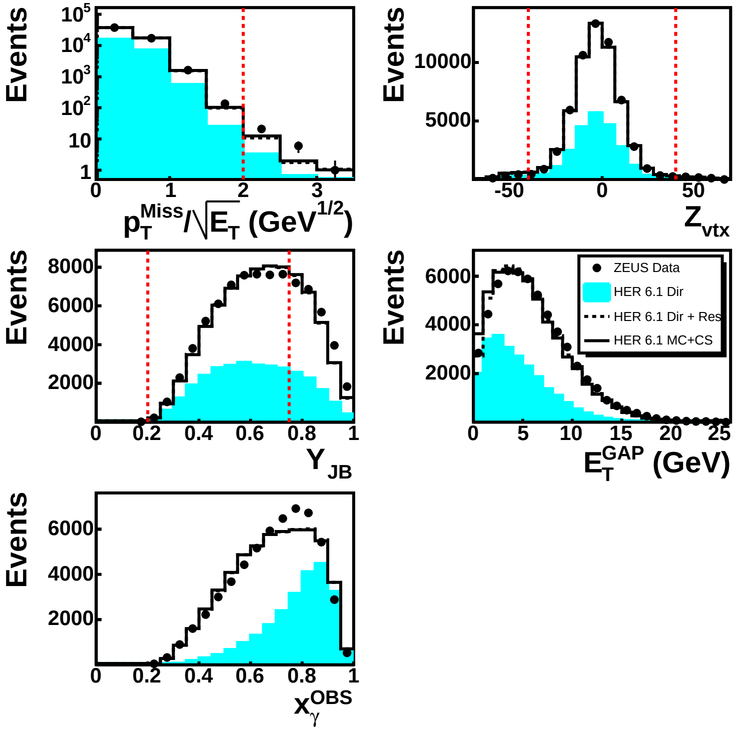


Figure 7.6: Description of the inclusive data sample by HERWIG for the variables $\frac{\not{p}_T}{\sqrt{E_T}}$, Z_{vtx}, y_{JB}, E_T^{GAP}, and x_γ^{OBS}. The ZEUS data is depicted by the points, the direct MC sample by the shaded histogram, the color-non-singlet MC sample by the dashed line, and the mixed MC sample by the solid line. The dashed vertical lines show the cuts applied in the offline selection. The MC was area normalized to the data in the region selected by the offline cuts.

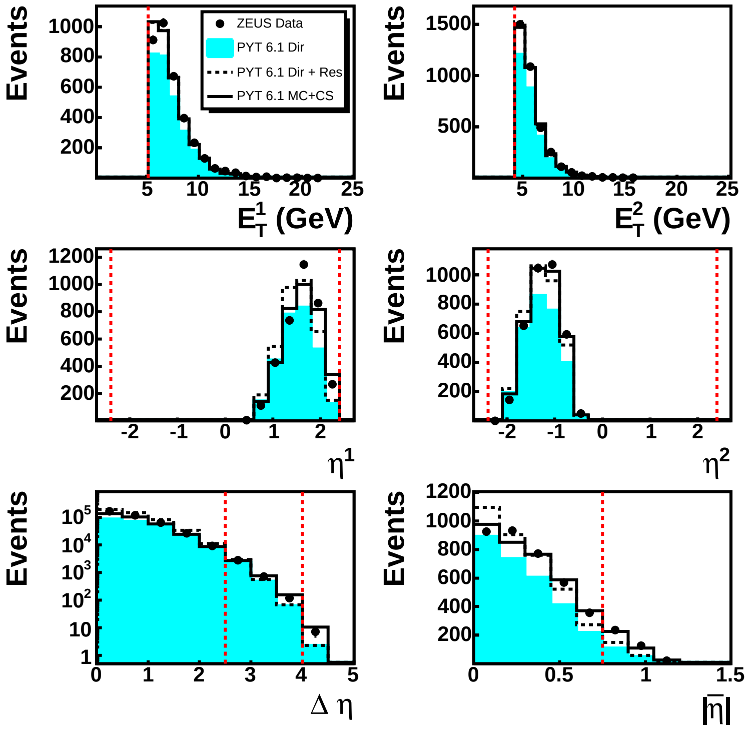


Figure 7.7: Description of the gap data sample, for $E_T^{CUT} = 1.2\,\text{GeV}$, by PYTHIA for the variables E_T, η, $\Delta\eta$, and $|\overline{\eta}|$. The ZEUS data is depicted by the points, the direct MC sample by the shaded histogram, the color-non-singlet MC sample by the dashed line, and the mixed MC sample by the solid line. The dashed vertical lines show the cuts applied in the offline selection. The MC was area normalized to the data in the region selected by the offline cuts.

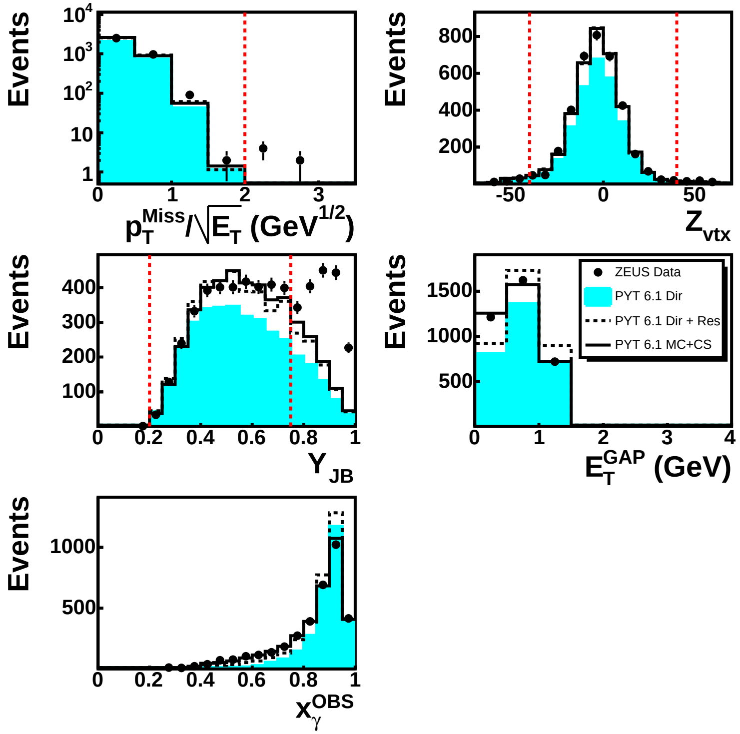


Figure 7.8: Description of the gap data sample, for $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.2\,\mathrm{GeV}$, by PYTHIA for the variables $\frac{\not{p}_{\mathrm{T}}}{\sqrt{E_{\mathrm{T}}}}$, Z_{vtx}, y_{JB}, $E_{\mathrm{T}}^{\mathrm{GAP}}$, and $x_{\gamma}^{\mathrm{OBS}}$. The ZEUS data is depicted by the points, the direct MC sample by the shaded histogram, the color-non-singlet MC sample by the dashed line, and the mixed MC sample by the solid line. The dashed vertical lines show the cuts applied in the offline selection. The MC was area normalized to the data in the region selected by the offline cuts.

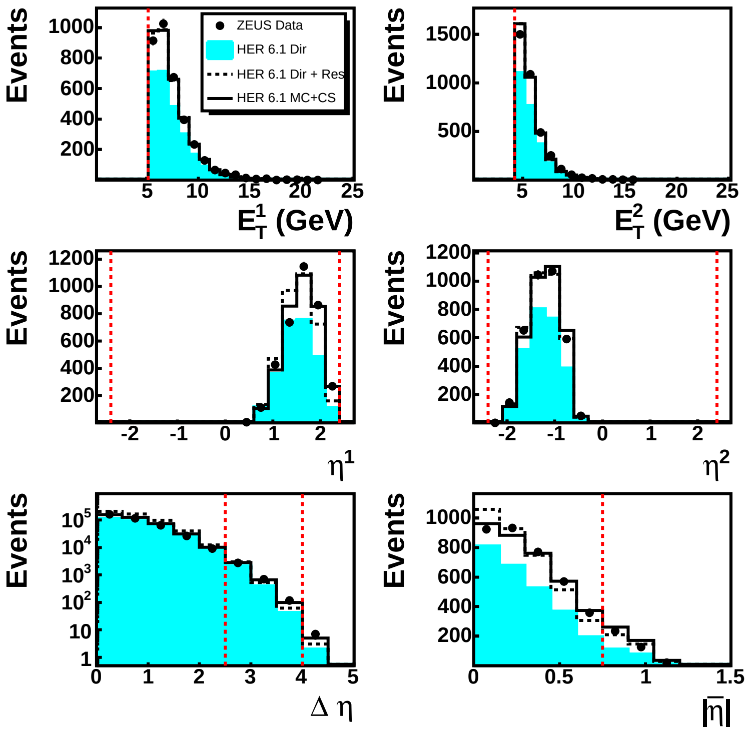


Figure 7.9: Description of the gap data sample, for $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.2\,\mathrm{GeV}$, by HERWIG for the variables E_{T}, η, $\Delta\eta$, and $|\overline{\eta}|$. The ZEUS data is depicted by the points, the direct MC sample by the shaded histogram, the color-non-singlet MC sample by the dashed line, and the mixed MC sample by the solid line. The dashed vertical lines show the cuts applied in the offline selection. The MC was area normalized to the data in the region selected by the offline cuts.

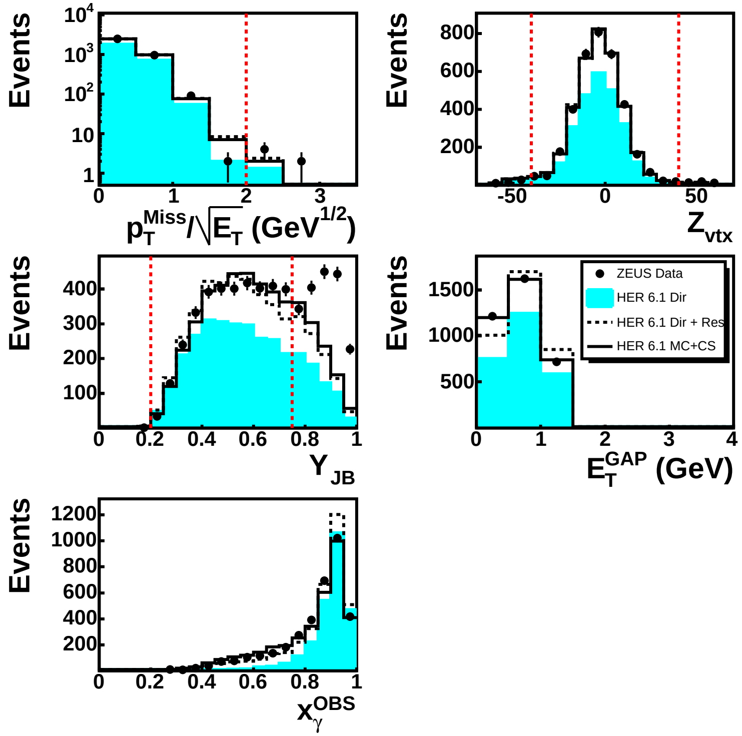


Figure 7.10: Description of the gap data sample, for $E_{\rm T}^{\rm CUT} = 1.2\,{\rm GeV}$, by Herwig for the variables $\frac{\not{p}_{\rm T}}{\sqrt{E_{\rm T}}}$, Z_{vtx}, $y_{\rm JB}$, $E_{\rm T}^{\rm GAP}$, and $x_\gamma^{\rm OBS}$. The ZEUS data is depicted by the points, the direct MC sample by the shaded histogram, the color-non-singlet MC sample by the dashed line, and the mixed MC sample by the solid line. The dashed vertical lines show the cuts applied in the offline selection. The MC was area normalized to the data in the region selected by the offline cuts.

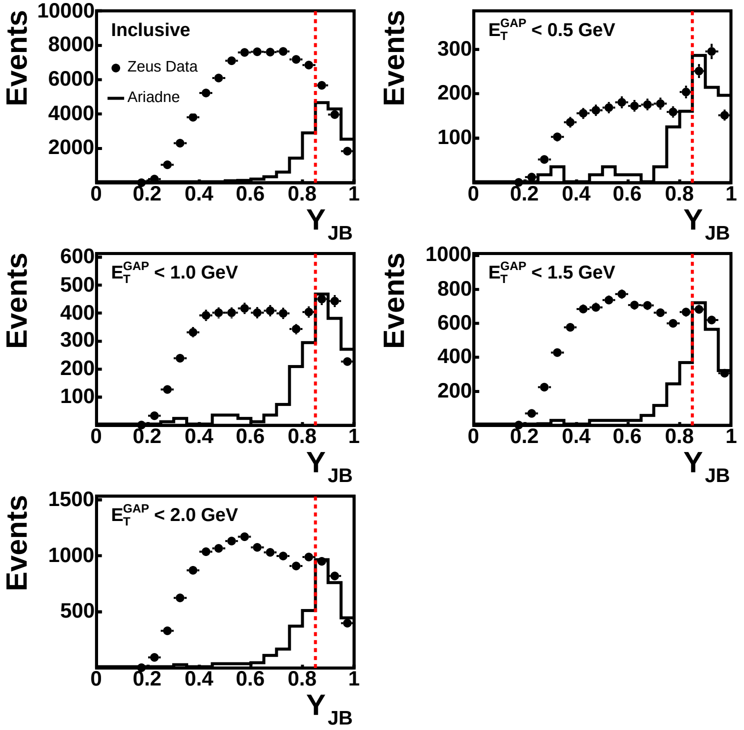


Figure 7.11: Comparisons of inclusive and gap data to Ariadne MC for the variable $y_{\rm JB}$. The ZEUS data is depicted by the points and the MC by the solid line. The dashed vertical line at $y_{\rm JB} = 0.85$ indicates the boundary between the photoproduction ($y_{\rm JB} < 0.85$) and DIS ($y_{\rm JB} > 0.85$) enhanced regions.

Chapter 8

Analysis Method

The inclusive and gap data samples, selected according to the criteria defined in Chapter 6, were corrected for detector acceptance to give differential cross sections and gap fractions at the hadron level. A detailed study of the sources contributing to the systematic uncertainties of the measurements was performed.

8.1 Correcting to the Hadron Level

Due to the resolution of the detector, trigger inefficiencies, and imperfect reconstruction, observed data does not correspond exactly with the processes occurring at the "truth" level[1]. Detector level measurements must therefore be corrected to this level in order make comparisons with corrected data from other experiments and theoretical predictions.

---

[1] In this analysis, the hadron level is the "truth" level. The "truth" level is where the measurement is compared to theoretical predictions.

8.1.1 Unfolding

A measured distribution, $m(\mathbf{x})$, of dimension N_m and a true distribution, $t(\boldsymbol{y})$ of dimension N_t are related by

$$m(\mathbf{x}) = \int R(\mathbf{x}|\boldsymbol{y}) t(\boldsymbol{y}) d^{N_t}\boldsymbol{y}, \tag{8.1}$$

where $R(\mathbf{x}|\boldsymbol{y})$ is a response function of the detector used to make the measurement. The determination of $t(\boldsymbol{y})$ from a known $m(\mathbf{x})$ and $R(\mathbf{x}|\boldsymbol{y})$ is a process called **unfolding**.

Multi-dimensional unfolding, where distributions are dependent on several variables, is computationally complex and requires high statistics. If the variables are essentially independent, the process can be simplified to one dimension, which reduces Equation 8.1 to

$$m(x) = \int R(x|y) t(y) dy. \tag{8.2}$$

In the case that the distributions m and t are represented as histograms, Equation 8.2 can be expressed as

$$\boldsymbol{m} = R\boldsymbol{t}, \tag{8.3}$$

where $\boldsymbol{m}$ is a vector of dimension N_m, $\boldsymbol{t}$ is a vector of dimension N_t, and R is a tensor of dimension $N_m \times N_t$. Equivalently, the elements of $\boldsymbol{m}$ can be expressed as

$$m_i = \sum_{j=1}^{N_t} R_{ij} t_j \quad i = 1, 2, \ldots, N_m. \tag{8.4}$$

An element R_{ij} in the response matrix can be interpreted as the probability that an event was measured in bin i divided by the probability that an event has a true value in bin j. The true level events, $\boldsymbol{t}$, can be determined from R^{-1} and $\boldsymbol{m}$.

8.1.2 The Bin-by-Bin Method

In the case that the bins of the measured distribution and true distribution are the same in size and number, the bin-by-bin method of unfolding data can be applied. Assuming the effects of detector resolution between variables are negligible, the off-diagonal terms in R disappear. Defining the correction factor as $\mathcal{C}_i \equiv \frac{1}{R_{ij}}\delta_{ij}$, true level events are then determined by

$$t_i = C_i m_i. \tag{8.5}$$

In practice, the correction factor is determined from the ratio of hadron to detector level Monte Carlo,

$$\mathcal{C}_i^{MC} = \frac{t_i^{MC}}{m_i^{MC}}, \tag{8.6}$$

for bins $i = 1, 2, \ldots, N$. This correction factor is then used to correct the data to the hadron level,

$$t_i^{Data} = \mathcal{C}_i^{MC} m_i^{Data}. \tag{8.7}$$

8.1.3 The Hadron-Level Cross Section

The cross section describes the likelihood that a certain process will occur. Equation 2.16 can be written for measured quantities as

$$\sigma^m = \frac{N^{Data}}{\mathcal{L}^{int}}, \tag{8.8}$$

where N^{Data} is the total number events in the data sample and $\mathcal{L}^{int} = \int \mathcal{L} dt$ is the integrated luminosity of the data sample. Since experimental results are often presented in the form of a histogram binned in a certain variable, x, the derivative of the cross section with respect x is often presented. For each bin, i, in the histogram, the differential cross section can be written as

$$\frac{d\sigma_i^m(x)}{dx} = \frac{N_i^{Data}(x)}{\mathcal{L}^{int}\Delta_i x}, \tag{8.9}$$

where Δ_i is the width of bin i. The differential cross section at the hadron level is obtained by correcting the data according to Equation 8.7,

$$\frac{d\sigma_i^t(x)}{dx} = \frac{N_i^{Data}(x)\mathcal{C}_i^{MC}(x)}{\mathcal{L}^{int}\Delta_i x}. \tag{8.10}$$

8.2 Definition of the Hadron Level Cross Section

In order to calculate the differential cross section at the hadron level according Equation 8.10, the correction factor, $\mathcal{C}^{MC}$, must be determined from the ratio of generator level to detector level MC. The detector level MC sample was selected by applying the same cuts as those applied offline to the data (see Section 6.3). The kinematic range of the generator level MC sample was selected by requiring

- $Q^2 < 1\,\text{GeV}$, which selects events in the photoproduction regime;
- $0.2 < y < 0.75$, which corresponds to a photon-proton center-of-mass energy in the range $134 < W < 260\,\text{GeV}$.

The inclusive sample of the generator level MC was selected by requiring, for the same reasons as listed in Section 6.3.2 for the detector level requirements, that

- $E_{\mathrm{T}}^{jet1} \geq 6\,\text{GeV}$ and $E_{\mathrm{T}}^{jet2} \geq 5\,\text{GeV}$;
- $-2.4 < \eta^{jet1,2} < 2.4$;
- $|\overline{\eta}| \leq 0.75$;

- $2.5 < \Delta\eta < 4$.

Four gap samples, as defined in Section 6.3.3, of generator level MC were selected by requiring

- $E_{\mathrm{T}}^{\mathrm{CUT}} = 0.5, 1.0, 1.5, \text{and } 2.0\,\mathrm{GeV}$.

The detector level jet E_{T} requirements (see Section 6.3.2) of $E_{\mathrm{T}}^{jet1} \geq 5.1\,\mathrm{GeV}$ and $E_{\mathrm{T}}^{jet2} \geq 4.24\,\mathrm{GeV}$ were determined by lowering the cuts used to define the hadron level cross section at the generator level by 15%. Similarly, the detector level $E_{\mathrm{T}}^{\mathrm{CUT}}$ values were raised[2] by 20%. These percentages were chosen partly to account for the pull on the jet E_{T} and partially to increase the efficiency of the selection (see Section 8.4) [83]. The usual procedures of correcting jet E_{T}, either a ratio of generator to detector level jet E_{T} or modifying ZUFO energy during reconstruction (see Section 5.6.2), could not be used in this analysis because the low E_{T} of the jets used to calculate $E_{\mathrm{T}}^{\mathrm{GAP}}$ limited their accuracy. A study of the reconstruction of $E_{\mathrm{T}}^{\mathrm{GAP}}$ performed using calorimeter cells, cone islands, uncorrected ZUFOs and corrected ZUFOs found that uncorrected ZUFOs provided the best resolution of true and reconstructed values. They also provided the best correlation between reconstructed and true levels over the entire $E_{\mathrm{T}}^{\mathrm{GAP}}$ range, most importantly at low $E_{\mathrm{T}}^{\mathrm{GAP}}$ [83]

In this analysis, differential inclusive and gap cross sections as a function of $E_{\mathrm{T}}^{\mathrm{GAP}}$, $\Delta\eta$, $x_{\gamma}^{\mathrm{OBS}}$, and W were measured and used to calculate gap fractions. The $E_{\mathrm{T}}^{\mathrm{GAP}}$ cross section was measured between 0 and 12 GeV because color singlet exchange dominates at low $E_{\mathrm{T}}^{\mathrm{GAP}}$ and the statistics decrease rapidly for $E_{\mathrm{T}}^{\mathrm{GAP}}$ above 12 GeV. The $\Delta\eta$ cross section was measured between 2.5 and 4.0 units of pseudorapidity. The lower limit was chosen to be slightly larger than the minimum separation of 2 units in pseudorapidity allowable by jets with a radius of 1 unit in pseudorapidity and the upper limit was constrained by the acceptance of the calorimeter. The $x_{\gamma}^{\mathrm{OBS}}$ cross section was measured from 0 to 1, which corresponds to the full range of the fraction of the photon's energy participating in the hard interaction. The W cross section was measured between 150 and 260 GeV, a range which was selected to correspond to the restricted W range of $134 < W < 260\,\mathrm{GeV}$ imposed by the kinematic requirement $0.2 < y < 0.75$.

---

[2] Therefore lowering the amount $E_{\mathrm{T}}^{\mathrm{GAP}}$ needed to define the gap

8.3 Corrections, Correlations, and Resolutions

It is important that the cross section variables show a strong correlation between their reconstructed and true values in order to ensure that the detector is not causing a bias in the measurement and the data is corrected accurately to the hadron level.

8.3.1 Corrections

It is possible to improve the correlation between reconstructed and true values by scaling the reconstructed level MC (as well as the data). The $E_{\mathrm{T}}^{\mathrm{GAP}}$ cross-section variable was corrected simply dividing the reconstructed MC (and data) by 1.2. This amount corresponds to the ratio between true and reconstructed jet energies, as described in Section 8.2. The W cross-section variable was corrected by scaling the reconstructed values by an amount so that the correlation between the true and reconstructed values have a slope of 1 as illustrated in Fig. 8.1. The upper-left figure shows a scatter plot of true and reconstructed values for uncorrected W. The dashed line with a slope of 1 shows the ideal correlation. The scatter plot was converted into a profile histogram and fit with a straight line, as shown in the upper-right figure. In order to align the fit of the profile histogram with the ideal correlation, a new reconstructed value, W_{new}, was determined by

$$W_{new} = 1.21W - 34.0. \tag{8.11}$$

A scatter plot and profile histogram using the corrected W values are shown in the lower-left and lower-right figures respectively. The correction, based upon the fit to the profile histogram, required to obtain correlation for PYTHIA was the same as that for HERWIG within errors.

All following plots of $E_{\mathrm{T}}^{\mathrm{GAP}}$ and W include the corrections described above. It was not necessary correct $\Delta\eta$ and $x_{\gamma}^{\mathrm{OBS}}$ since their correlation was already adequate.

8.3.2 Correlations

Correlations between the true and reconstructed values for all cross section variables are shown in Fig. 8.2 for PYTHIA and Fig. 8.3 for HERWIG. A strong correlation is observed for all variables.

8.3.3 Resolutions

The correlations were quantified by plotting the resolutions and fitting them to a Gaussian distribution, as shown in Fig. 8.4 for PYTHIA and Fig. 8.5 for HERWIG. The fit box in the upper right corner of each plot shows the fit parameters. The mean of the Gaussian corresponds to the pull and the sigma corresponds to the resolution of the distribution. The pull is between 0.05 and 0.08 for $E_{\mathrm{T}}^{\mathrm{GAP}}$ and between 0.001 and 0.01 for $\Delta\eta$, $x_{\gamma}^{\mathrm{OBS}}$, and W. The resolutions for $\Delta\eta$, $x_{\gamma}^{\mathrm{OBS}}$, and W are between 3 and 8%, while for $E_{\mathrm{T}}^{\mathrm{GAP}}$ they are approximately 30%. The poor $E_{\mathrm{T}}^{\mathrm{GAP}}$ resolution can be attributed to the low E_{T} jets in the gap which are not reconstructed as well as jets

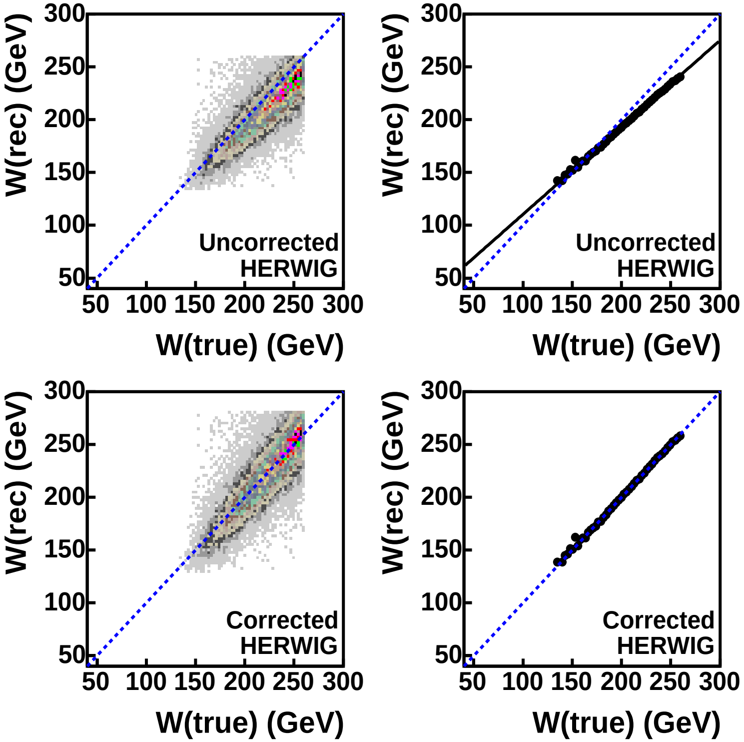


Figure 8.1: Correction of W calculated using HERWIG. The upper-left plot shows the uncorrected correlation between true and reconstructed values as a scatter plot. The points in the upper-right show the uncorrected correlation as a profile histogram and the solid line is a fit to these points. The lower-left plots shows a scatter plot using the corrected reconstructed values. The lower-right plot shows a profile histogram of the correlation using corrected reconstructed values. The dashed line in each plot, having a slope of 1, shows the ideal correlation between true and reconstructed values.

Variable	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5
$E_{\mathrm{T}}^{\mathrm{GAP}}(\,\mathrm{GeV})$	0.0 - 0.5	0.5 - 1.5	1.5 - 3.5	3.5 - 7.0	7.0 - 12.0
$\Delta\eta$	2.5 - 2.8	2.8 - 3.1	3.1 - 3.5	3.5 - 4.0	-
$x_{\gamma}^{\mathrm{OBS}}$	0.0 - 0.5	0.5 - 0.75	0.75 - 0.9	0.9 - 1.0	-
$W(\,\mathrm{GeV})$	150 - 180	180 - 210	210 - 240	240 - 260	-

Table 8.1: The binning of the cross section variables $E_{\mathrm{T}}^{\mathrm{GAP}}$, $\Delta\eta$, $x_{\gamma}^{\mathrm{OBS}}$, and W.

with higher E_{T}.

The bin widths of the cross section variables must be larger than the resolution of the variables in each bin so that bin migrations do not significantly affect the correction of the cross section to the hadron level. **Bin migrations** occur when events are generated and reconstructed in different bins and their magnitude is sensitive to hadronic corrections applied to the data. These correction routines are usually applied to the ZUFOs during reconstruction (see Section 5.6.2) but in this analysis were implemented by applying different jet energy cuts at the hadron and detector levels (see Section 8.2).

The strong correlation between the true and reconstructed values over the entire cross section ranges permits the resolution in each bin to be calculated from the resolutions for the entire variable range shown in Fig. 8.4 and Fig. 8.5. Bin widths were chosen to be significantly larger than the approximate resolution in that bin. For example, the resolution of the entire $E_{\mathrm{T}}^{\mathrm{GAP}}$ is 0.30 and therefore at $E_{\mathrm{T}}^{\mathrm{GAP}} = 0.5\,\mathrm{GeV}$, the minimum bin width should be $0.30 \times 0.5\,\mathrm{GeV} = 0.15\,\mathrm{GeV}$. A larger bin width of 0.5 GeV was actually chosen for the cross section measurement. The binning of each cross section variable is summarized in Table 8.1.

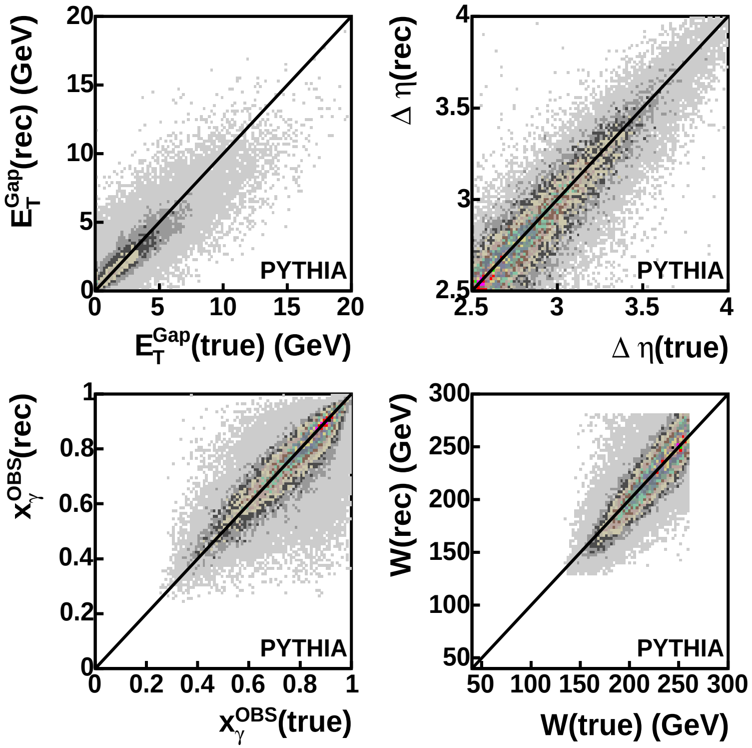


Figure 8.2: Correlation between true and reconstructed values of the cross section variables $E_{\mathrm{T}}^{\mathrm{GAP}}$, $\Delta\eta$, $x_{\gamma}^{\mathrm{OBS}}$, and W from Pythia.

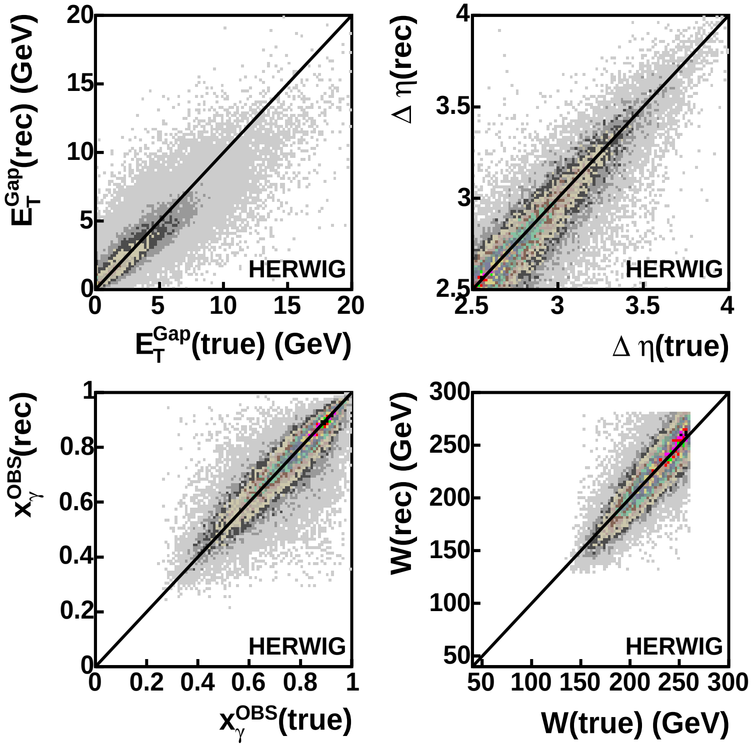


Figure 8.3: Correlation between true and reconstructed values of the cross section variables $E_{\mathrm{T}}^{\mathrm{GAP}}$, $\Delta\eta$, x_γ^{OBS}, and W from HERWIG.

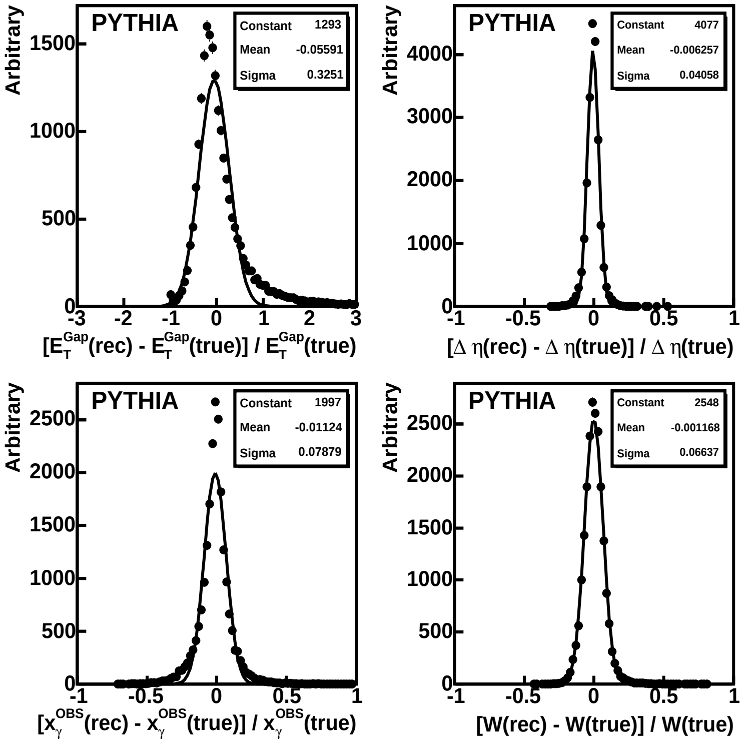


Figure 8.4: The pull and resolution of the cross section variables $E_{\mathrm{T}}^{\mathrm{GAP}}$, $\Delta\eta$, x_γ^{OBS}, and W from PYTHIA. The curve shows a fit to a Gaussian distribution. The fit parameters are in the box in the upper right corner of the plot.

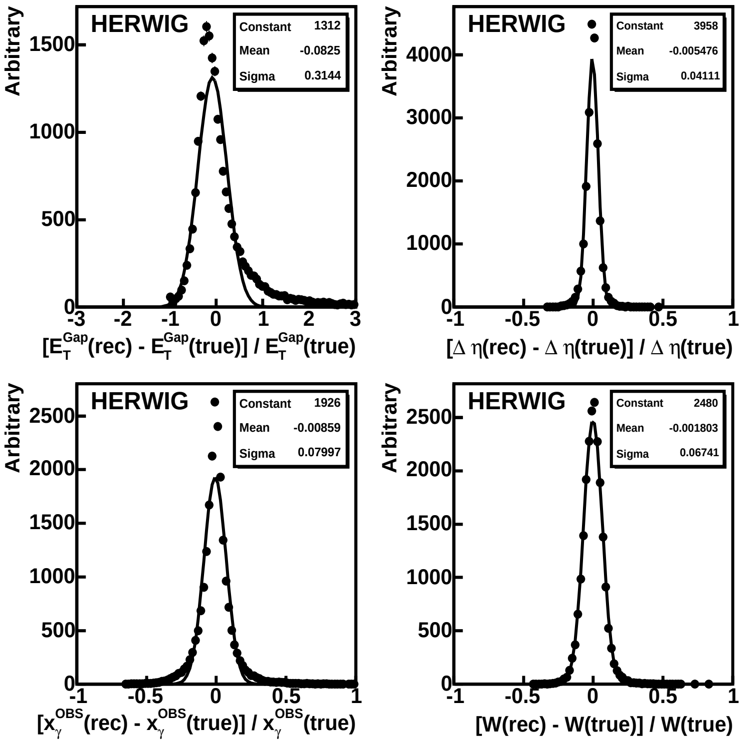


Figure 8.5: The pull and resolution of the cross section variables $E_{\mathrm{T}}^{\mathrm{GAP}}$, $\Delta\eta$, $x_{\gamma}^{\mathrm{OBS}}$, and W from HERWIG. The curve shows a fit to a Gaussian distribution. The fit parameters are in the box in the upper right corner of the plot.

8.4 Purities, Efficiencies, Stabilities, and Correction Factors

Since the bin-by-bin method for correcting data to the hadron level depends only on the number of MC events generated in bin i and the number of MC events reconstructed in bin i, it is useful to quantitatively determine the effects of bin migrations.

The purity and efficiency of a certain bin, i, in a MC sample are defined as

$$\mathcal{P}_i \equiv \frac{b_i}{r_i} \tag{8.12}$$

$$\mathcal{E}_i \equiv \frac{b_i}{t_i} \tag{8.13}$$

where r_i is the number of events reconstructed in bin i, t_i is the number of events at the true level, or generated, in bin i, and b_i is the number of events *both* generated and reconstructed in the same bin i. The purity gives the fraction of events generated in bin i which remained in bin i after reconstruction and the efficiency gives the fraction of events reconstructed in bin i which were also generated in bin i. The errors on the purity and efficiency are given by[3].

$$\delta\mathcal{P}_i = \left[\frac{(1-\mathcal{P}_i)\mathcal{P}_i}{r_i}\right]^{1/2} \tag{8.14}$$

$$\delta\mathcal{E}_i = \left[\frac{(1-\mathcal{E}_i)\mathcal{E}_i}{t_i}\right]^{1/2}. \tag{8.15}$$

The stability (also known as smearing) provides information about how stable a distribution is with respect to bin migrations and is defined as

$$\mathcal{S}_i \equiv \mathcal{P}_i\frac{r_i}{b} = \frac{b_i}{b}, \tag{8.16}$$

---

[3] The errors in this section are derived in Appendix C.

where b is the number of events generated in bin i that are reconstructed in any bin of the final sample. The error on the stability is given by

$$\delta\mathcal{S}_i = \left[\frac{(1-\mathcal{S}_i)\mathcal{S}_i}{b_i}\right]^{1/2}. \tag{8.17}$$

The correction factor in bin i is related to the purity and efficiency in bin i by

$$\mathcal{C}_i = \frac{r_i}{t_i} = \frac{\mathcal{P}_i}{\mathcal{E}_i} \tag{8.18}$$

and has an error

$$\delta\mathcal{C}_i = \left[\frac{t_i}{r_i^3}\left(t_i + r_i - 2b_i\right)\right]. \tag{8.19}$$

.

The purity, efficiency, stability, and correction factor for $\Delta\eta$ are shown in Fig. 8.6 for Pythia and Fig. 8.7 for Herwig. The direct, resolved, and color-singlet contributions are mixed according to the procedure described in Chapter 7. The solid points show the inclusive sample and the open squares show the gap sample defined by $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.0\,\mathrm{GeV}$(true) and $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.2\,\mathrm{GeV}$(reconstructed).

The purity for Pythia (Fig. 8.6) is approximately 0.32 across the entire $\Delta\eta$ range, except for the last bin where it is slightly lower, for both the inclusive and gap samples. The efficiency decreases slightly across the $\Delta\eta$ range for the inclusive sample, from approximately 0.27 to 0.20, and is approximately constant at 0.21 for the gap sample. The stability is approximately 0.85 in the lowest $\Delta\eta$ bin and approximately 0.70 in the other bins for the inclusive sample and slightly higher for the gap sample. The correction factor is approximately 1.2 for the inclusive sample and 1.5 for the gap sample in all bins.

The purity for Herwig (Fig. 8.7) is slightly lower that that of Pythia, with a value of approximately 0.3 across the entire $\Delta\eta$ range, except for the last bin where it is

slightly lower, for both the inclusive and gap samples. The efficiency for Herwig also shows a decline across $\Delta\eta$, with the inclusive sample having approximately the same value as Pythia and the gap sample having a slightly lower value than Pythia. The stability for Herwig has similar values and trends as that for Pythia. The correction factor for the inclusive sample for both Pythia and Herwig is also similar in value and trend, while the correction factor for the gap sample is larger at high $\Delta\eta$ for Herwig.

The purities, efficiencies, stabilities, and correction factors for the other cross section variables show a similar behavior to those for $\Delta\eta$. These quantities are shown in Appendix D.1 for all cross section variables.

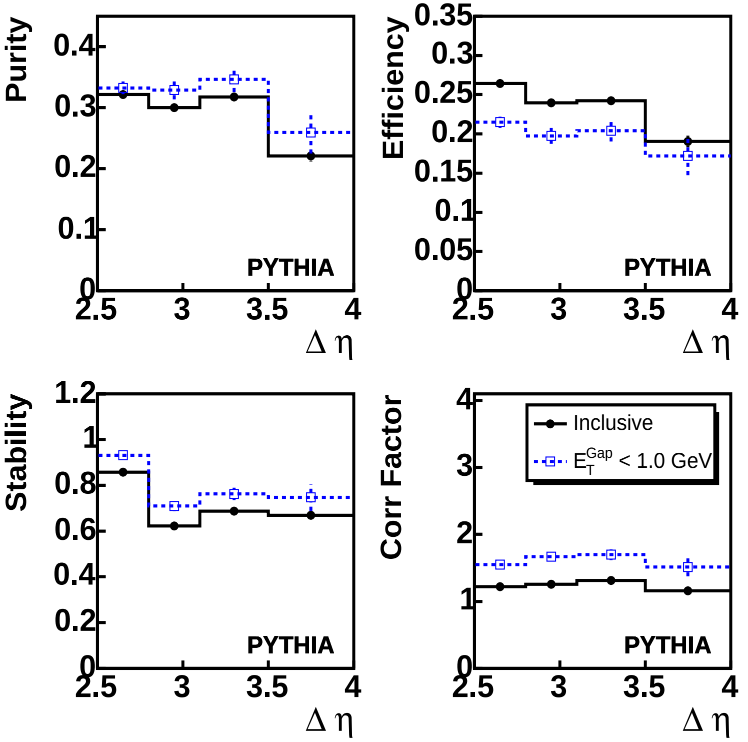


Figure 8.6: The purity, efficiency, stability, and correction factor of $\Delta\eta$ calculated from Pythia. The solid points show the inclusive sample and the open squares show the gap sample defined by $E_T^{\mathrm{CUT}} = 1.0\,\mathrm{GeV}$(true) and $E_T^{\mathrm{CUT}} = 1.2\,\mathrm{GeV}$(reconstructed).

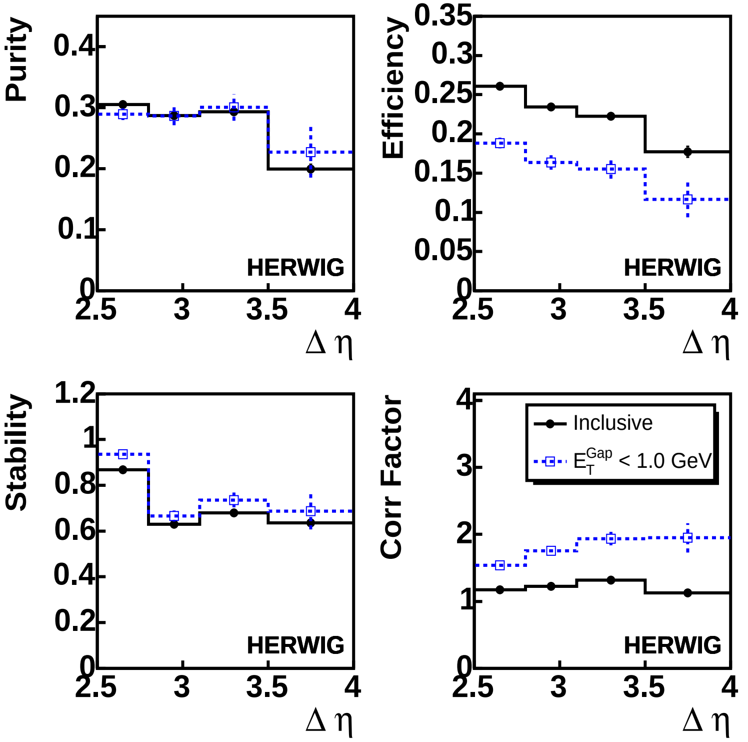


Figure 8.7: The purity, efficiency, stability, and correction factor of $\Delta\eta$ calculated from Herwig. The solid points show the inclusive sample and the open squares show the gap sample defined by $E_T^{\mathrm{CUT}} = 1.0\,\mathrm{GeV}$(true) and $E_T^{\mathrm{CUT}} = 1.2\,\mathrm{GeV}$(reconstructed).

8.5 Cumulative Efficiencies

The efficiencies at each stage of the event selection were also studied. In this case, the efficiency was defined as the number of true level events passing the true level cuts *and* passing a certain selection after reconstruction to the number of true level events passing the true level cuts. For example, the FLT efficiency in bin i is

$$\mathcal{E}_i = \frac{t_i \cap r_i^{FLT}}{t_i}, \tag{8.20}$$

where t_i is the number of events at the true level which passed the true level cuts (see Section 8.2) and r_i^{FLT} is the number of events at the reconstructed level which passed the FLT selection. The true and reconstructed events are not required to lie in the same bin, unlike the efficiency defined in Section 8.4. The efficiencies at each level of selection were calculated independently of other levels of selection. For example, events used to measure the SLT efficiency were not required to pass the FLT.

The FLT, SLT, TLT, offline, and total efficiencies for $\Delta\eta$ are shown in Fig. 8.8 for Pythia and Fig. 8.9 for Herwig. The direct, resolved, and color-singlet contributions are mixed according to the procedure described in Chapter 7. The solid points show the inclusive sample and the open squares show the gap sample defined by $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.0\,\mathrm{GeV}$(true) and $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.2\,\mathrm{GeV}$(reconstructed).

For both Pythia and Herwig, the FLT and SLT efficiencies rise from approximately 85% to almost 100% across the $\Delta\eta$ range for both the inclusive and gap samples. The TLT efficiency rises from approximately 65% to 70% for Pythia and Herwig for both the inclusive and gap samples. The offline efficiencies are also similar for both MCs, with the inclusive sample varying between approximately 40% and 50% and the gap samples varying between approximately 20% and 30%. The total efficiencies are also similar for the inclusive samples, varying between approximately 30% and 40%, and the gap sample, varying between approximately 20% and 35%. These efficiencies are slightly higher than those shown in Section 8.4 because it was not required that the events be generated and reconstructed in the same bin.

A very small decrease in the efficiency occurs at the FLT and SLT, and a moderate decrease occurring at the TLT. The largest decrease in efficiency by far occurs during the offline selection. This is because the trigger selection removes mostly background events while the stricter offline cuts remove a greater proportion of physics events. The lower efficiency of the gap sample compared to the inclusive sample is due to the large resolution of $E_{\mathrm{T}}^{\mathrm{GAP}}$. The other cumulative efficiencies of the other cross section variables show similar behavior to those of $\Delta\eta$ and can be seen in Appendix D.2.

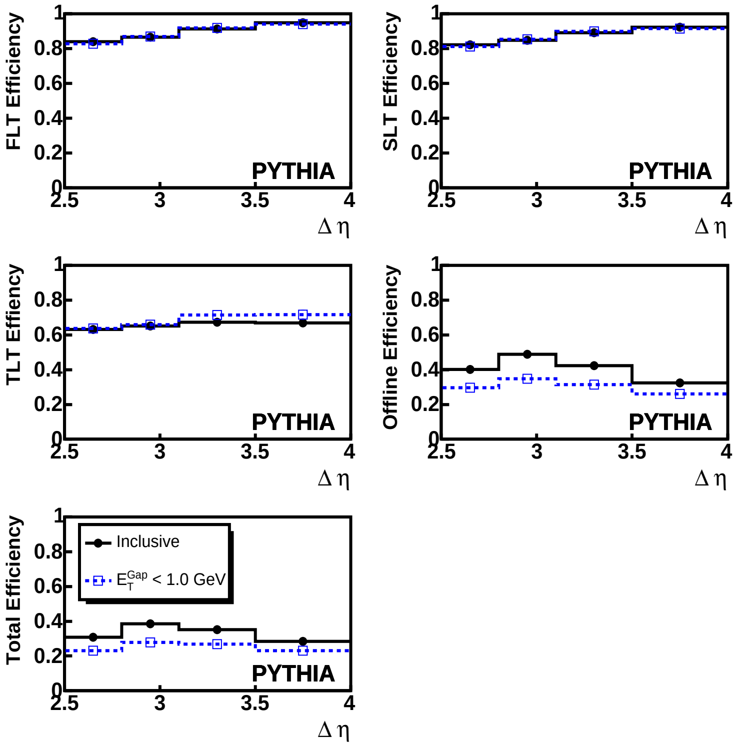


Figure 8.8: The cumulative efficiencies at each stage of event selection for $\Delta\eta$ calculated from Pythia. The solid points show the inclusive sample and the open squares show the gap sample defined by $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.0\,\mathrm{GeV}$(true) and $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.2\,\mathrm{GeV}$(reconstructed).

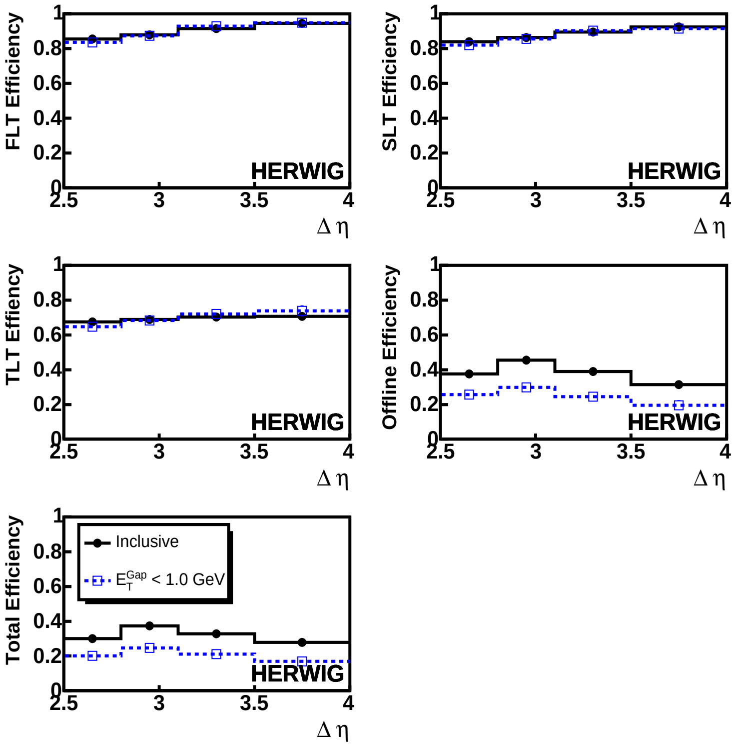


Figure 8.9: The cumulative efficiencies at each stage of event selection for $\Delta\eta$ calculated from Herwig. The solid points show the inclusive sample and the open squares show the gap sample defined by $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.0\,\mathrm{GeV}$(true) and $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.2\,\mathrm{GeV}$(reconstructed).

8.6 The Uncorrected Results

The uncorrected results for the data and MC at the detector level are presented for the cross section variables $E_{\mathrm{T}}^{\mathrm{GAP}}$, $\Delta\eta$, $x_{\gamma}^{\mathrm{OBS}}$, and W in Figs 8.10 to 8.17. In addition to providing a detector level description of the variables used in the offline selection (see Section 7.3), the MC must also provide a detector level description of the cross section variables. This ensures that the kinematic phase space is well understood and that the correction procedure produces accurate hadron level distributions. Also, the uncorrected results can provide information, albeit with detector effects included, about the underlying physical processes.

The uncorrected differential cross sections are calculated from Equation 8.10 with the luminosity and correction factor set to one,

$$\frac{d\sigma_i^{Uncor}(x)}{dx} = \frac{N_i^{Data}(x)}{\Delta_i x}. \tag{8.21}$$

The ZEUS data in each plot is shown as black points, the direct MC contribution as the shaded area, the mixture of direct and resolved MC samples as the dashed line, and the mixture of direct, resolved, and color-singlet MC samples as the solid line. The MC was mixed according to the procedure described in Chapter 7 and was area normalized to the data for each distribution.

The uncorrected differential cross sections as a function of $E_{\mathrm{T}}^{\mathrm{GAP}}$ (Figs. 8.11 and 8.11) show a reasonable agreement between the data and MC. The addition of color-singlet MC improves agreement only in the lowest and the third bin. The level of agreement between the data and the color-non-singlet and color-singlet samples is

approximately the same for both Pythia and Herwig. The uncorrected differential cross sections as a function of $\Delta\eta$ (Figs. 8.13 and 8.13) show excellent agreement between the data and MC for both Pythia and Herwig. The addition of the color-singlet MC does not significantly improve the agreement with the data for the inclusive sample, but does for the gap samples. Similar behavior is seen for the differential cross sections as a function of $x_{\gamma}^{\mathrm{OBS}}$ and W (Figs. 8.14 to 8.17).

The good level of agreement between the data and MC in the uncorrected differential cross sections provides confidence in the accuracy of the corrected results. Also, the need for color-singlet MC to match the gap samples confirms that its addition to the MC sample used in the unfolding process was necessary and predicts that color-singlet MC will also be necessary to obtain agreement with the data at the hadron level.

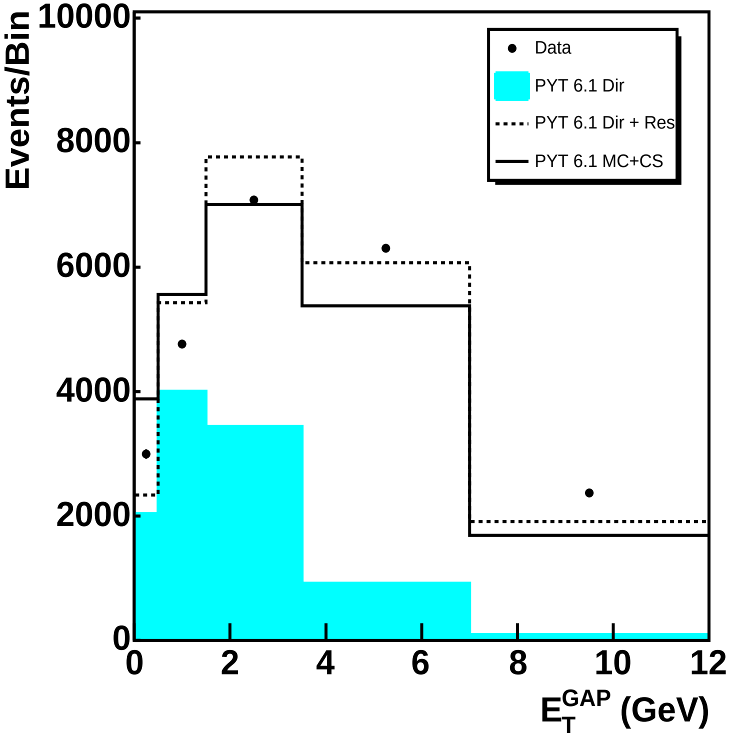


Figure 8.10: The uncorrected $E_{\mathrm{T}}^{\mathrm{GAP}}$ distribution compared to PYTHIA. The ZEUS data is depicted by the points, the direct MC sample by the shaded histogram, the color-non-singlet MC sample by the dashed line, and the mixed sample by the solid line.

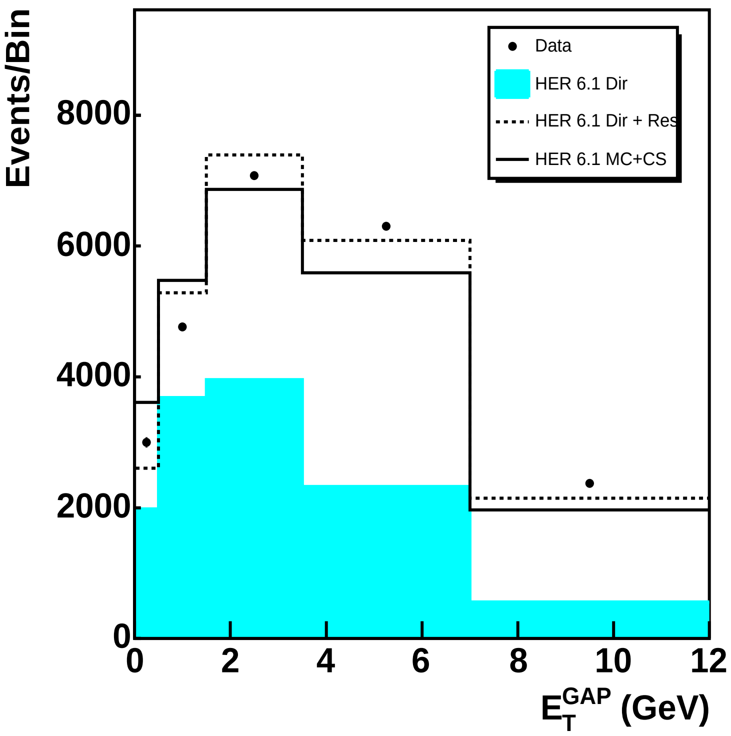


Figure 8.11: The uncorrected $E_{\mathrm{T}}^{\mathrm{GAP}}$ distributions compared to HERWIG. The ZEUS data is depicted by the points, the direct MC sample by the shaded histogram, the color-non-singlet MC sample by the dashed line, and the mixed sample by the solid line.

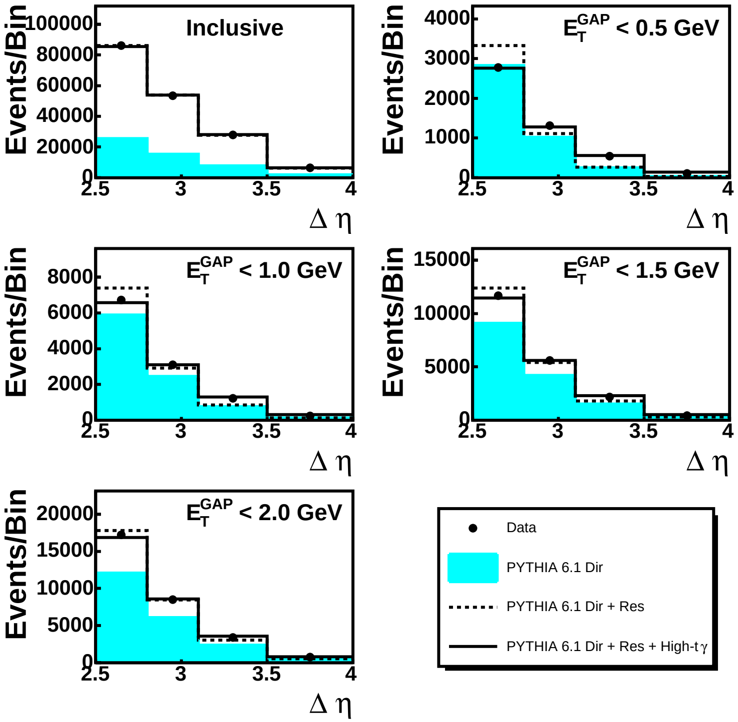


Figure 8.12: The uncorrected $\Delta\eta$ distributions compared to PYTHIA. The ZEUS data is depicted by the points, the direct MC sample by the shaded histogram, the color-non-singlet MC sample by the dashed line, and the mixed sample by the solid line.

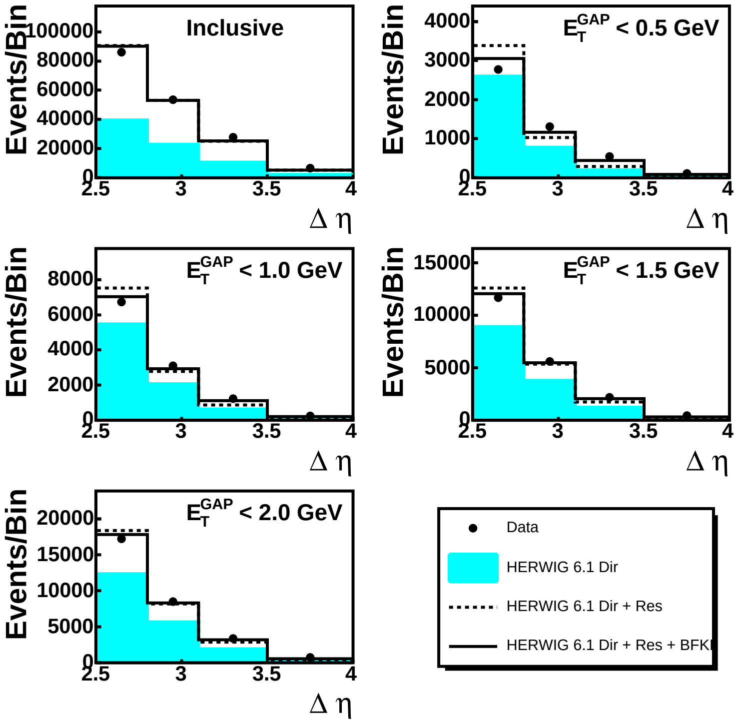


Figure 8.13: The uncorrected $\Delta\eta$ distributions compared to HERWIG. The ZEUS data is depicted by the points, the direct MC sample by the shaded histogram, the color-non-singlet MC sample by the dashed line, and the mixed sample by the solid line.

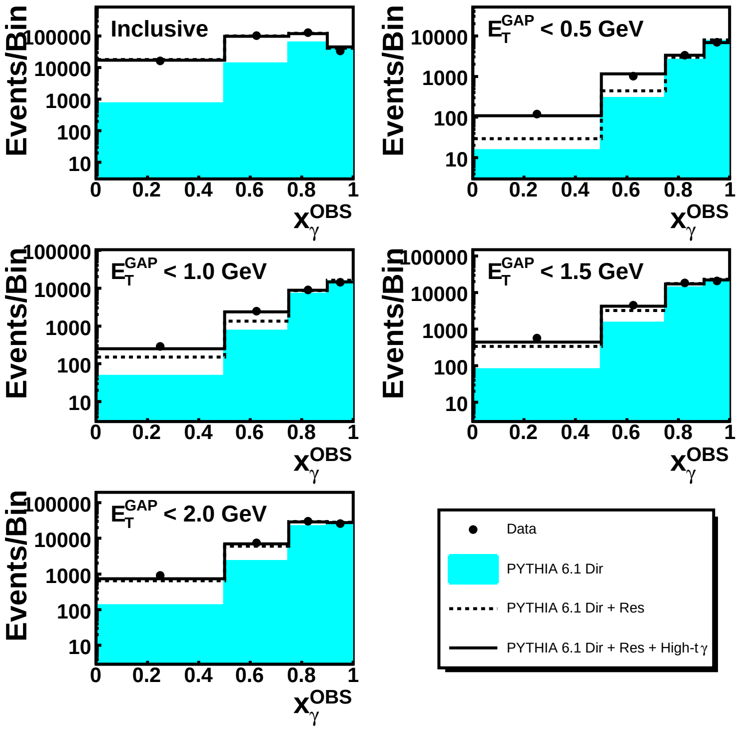


Figure 8.14: The uncorrected x_γ^{OBS} distributions compared to PYTHIA. The ZEUS data is depicted by the points, the direct MC sample by the shaded histogram, the color-non-singlet MC sample by the dashed line, and the mixed sample by the solid line.

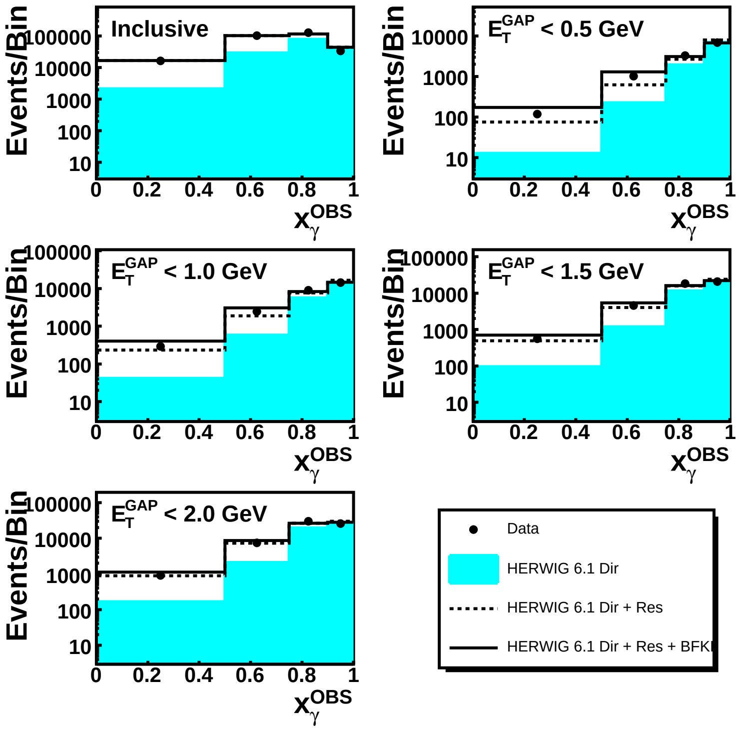


Figure 8.15: The uncorrected x_γ^{OBS} distributions compared to HERWIG. The ZEUS data is depicted by the points, the direct MC sample by the shaded histogram, the color-non-singlet MC sample by the dashed line, and the mixed sample by the solid line.

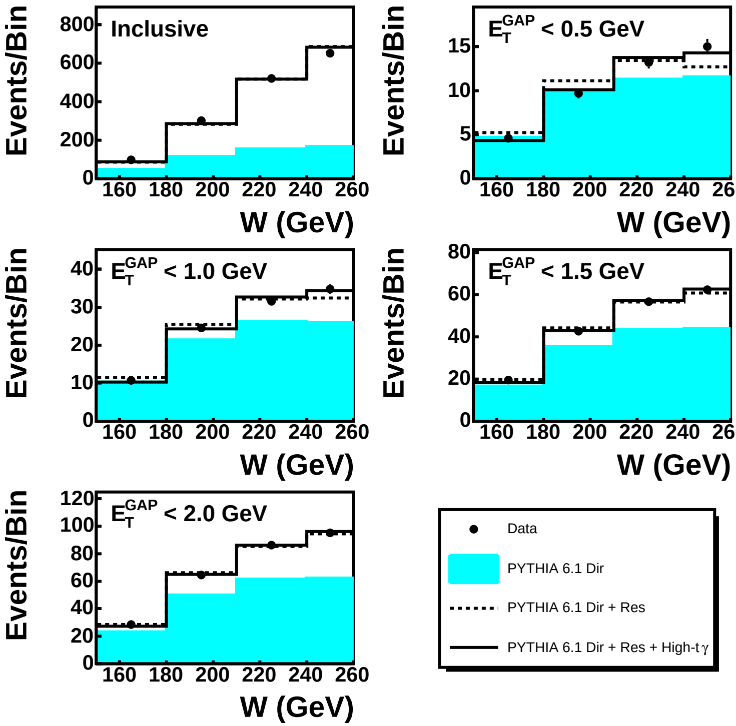


Figure 8.16: The uncorrected W distributions compared to Pythia. The ZEUS data is depicted by the points, the direct MC sample by the shaded histogram, the color-non-singlet MC sample by the dashed line, and the mixed sample by the solid line.

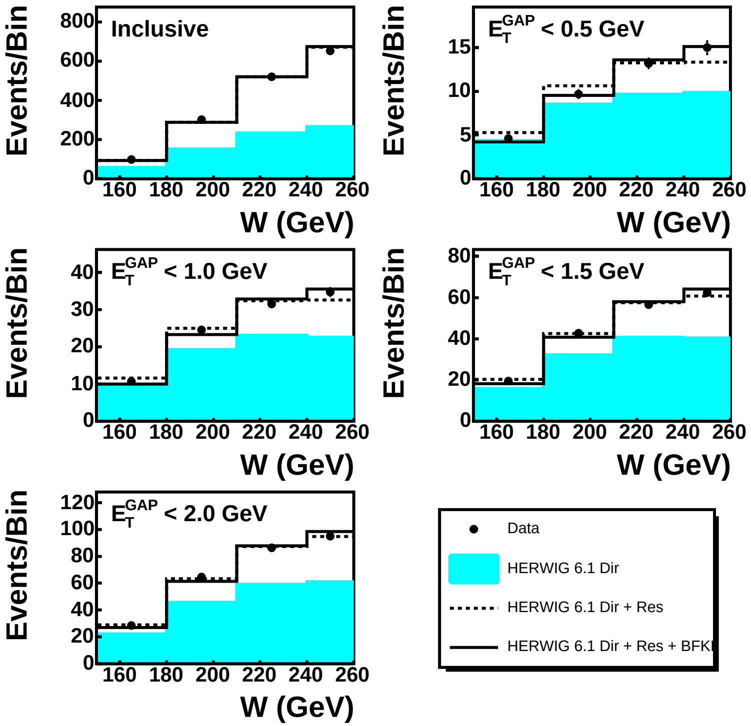


Figure 8.17: The uncorrected W distributions compared to Herwig. The ZEUS data is depicted by the points, the direct MC sample by the shaded histogram, the color-non-singlet MC sample by the dashed line, and the mixed sample by the solid line.

8.7 Systematic Uncertainties

Systematic uncertainties arise from the chosen method of measurement and bias the measured value in a certain direction. Unlike statistical uncertainties, systematic uncertainties cannot be minimized simply by increasing the size of the data sample and do not have a fixed method of calculation. The usual procedure used to estimate systematic uncertainty is to first make the measurement using the nominal method and then to make the measurement again after altering one parameter in this method. The difference between the results from nominal and altered method is then added in quadrature to the statistical uncertainty and systematic uncertainties resulting from other alterations. Exactly which parameters in the measurement methods are altered in order to quantify the systematic uncertainties is not fixed and depends on the specific measurement. The major sources of systematic uncertainties in this analysis are discussed below.

Since HERWIG provides a slightly better description of the kinematic variables than PYTHIA, it was used to estimate the systematic uncertainty due to the modification of the kinematic cuts. HERWIG was also used to estimate the uncertainty due to the calorimeter energy scale and amount of color-singlet MC used in the unfolding in order to maintain consistency.

8.7.1 Luminosity Measurement

The uncertainty in the ZEUS gated luminosity measurement for the 1996-97 data taking period is ±1.6% [71]. Since the incorporation of this systematic uncertainty causes a change of only about 2% in the cross sections and since it cancels in the gap fraction, it has been ignored.

8.7.2 Reconstruction Method

The analysis described in this thesis used uncorrected ZUFOs for the reconstruction of the hadronic final state. It has been shown [83] that they provide the best reconstruction of low energy particles and isolated clusters. The use of cells and corrected ZUFOs [83] has also been studied to ensure the stability of the reconstruction method. The level of difference in cross sections calculated using uncorrected ZUFOs and cells is approximately 5%. The choice of reconstruction method has not been included in the systematic uncertainties since ZUFOs, which rely on tracking information in addition to calorimeter information, generally provide a more accurate reconstruction than cells and corrected ZUFOs do not properly reconstruct the low-energy particles in the gap.

8.7.3 Model Dependence

The data was corrected with both PYTHIA and HERWIG in order to evaluate the systematic uncertainty originating from the hadronization process. The model dependence was the major contribution to the systematic uncertainty, about 10%, and in order to minimize this uncertainty, the average correction factor of PYTHIA and HERWIG was used (see Section 8.4). The difference between the correction factors of PYTHIA and HERWIG was included in the systematic uncertainty.

8.7.4 Calorimeter Energy Scale

The calorimeter energy scale was assigned an uncertainty of ±3% [84]. This systematic was implemented by raising and lowering the jet energy by 3%, and affected the jet E_{T}, $\frac{\not{p}_{\mathrm{T}}}{\sqrt{E_{\mathrm{T}}}}$, and y_{JB}. Since this uncertainty is correlated between bins of the cross section, it is displayed as a band overlayed on the data points and not added in quadrature with the other errors. The calorimeter energy scale uncertainty was estimated using Herwig.

8.7.5 Color-singlet MC in the Unfolding

The amount of color-singlet exchange MC used in the unfolding was varied by ±25%. This resulted in a variation of approximately 1% in the cross section. The uncertainty due to the amount of color-singlet MC used in unfolding was estimated using Herwig.

8.7.6 Kinematic Selection

The uncertainty due to the imperfect understanding and simulation of detector effects in the MC was estimated by varying the kinematic cuts used in the offline event selection (see Section 6.3) for the data and detector level Monte Carlo sample, while keeping the cuts on the true level MC fixed (see Section 8.2), and then recalculating the differential cross section.

Rewriting Equation 8.10 as

$$\frac{d\sigma}{dx} = \frac{N_t^{MC}}{\mathcal{L}^{int}\Delta x}\left(\frac{N^{Data}}{N_r^{MC}}\right), \tag{8.22}$$

where N^{Data} is the number of data events, N_t^{MC} is the number of MC events at the true level, and N_r^{MC} is the number of data events at the reconstructed level, it can be seen that variation in the differential cross section in the estimation of systematic uncertainty depends solely upon the variation of the ratio N^{Data}/N_r^{MC}. If the data is well described by the detector level MC, especially at the boundaries of the cuts, the systematic uncertainty will be very small. It is therefore essential that the MC provides a good description of the data (see Chapter 7).

Cuts on $\frac{\not{p}_{\mathrm{T}}}{\sqrt{E_{\mathrm{T}}}}$, Z_{vtx}, and y_e were only applied in the offline selection to the data and to the detector level MC and had no corresponding cut at the true level. The systematic uncertainty for these cuts was estimated by varying the cut by an amount corresponding to the resolution of the variables.

The other kinematic variables had corresponding cuts at the true level and these cuts were varied by an amount corresponding to the standard deviation (σ) of a Gaussian fit to their resolutions. The resolutions of these kinematic variables is the variance of the reconstructed value with respect to the true value and is analogous to those shown for the cross section variables in Section 8.3. The variation of the kinematic by σ accounts for bin migrations between the true and detector levels.

The sample of Herwig used to calculate the resolutions was a combination of direct and resolved MC mixed according to the generated luminosity. Resolutions calculated using Pythia differed by a few percent from those calculated using Herwig for all variables used in the kinematic selection except $E_{\mathrm{T}}^{\mathrm{GAP}}$, where the difference was 17%. The variation of the cuts on $E_{\mathrm{T}}^{\mathrm{GAP}}$, y_{JB}, and E_{T} gave the largest contributions to the systematic uncertainty. Depending upon the variable measured, the change in the cross section ranged from a few percent to approximately 30% in regions where the statistical significance was extremely low.

8.7.7 Plots of Systematics

The event selection variations used to calculate the systematic uncertainties is summarized in Table 8.2. The cross sections and gap fractions resulting from the variation used in the study of the systematic uncertainties are displayed in Figs. 8.18 to 8.21 for each cross section variable. In each bin of each plot, the first (left-most) point shows the nominal cross section value and each successive point shows the cross section calculated by varying one parameter in the measurement. The correspondence between the order of the points and the variation of parameters is summarized in Table 8.3. The error bars on the points show the statistical uncertainty on the measurement resulting from the propagation of the data and MC statistics. The fractional uncertainty of the MC for the cross sections and gap fractions varies between between 1 and 6% depending on the bin. In the bins with the least statistics (for example, the highest bin in $\Delta\eta$), the error on the MC is approximately 10% of the total error, statistical plus systematic, on the data. A dashed horizontal line is drawn through the highest and lowest systematic variation in each bin. In most bins, the calorimeter energy scale uncertainty causes the largest systematic uncertainty. Of the kinematic variables, the variation of y_{JB} and $E_{\mathrm{T}}^{\mathrm{GAP}}$cause the largest systematic uncertainty. This was predicted from the disagreement between the data and MC near the cut of $y_{\mathrm{JB}} = 0.75$ in Figs 7.8 and 7.10 and the large resolution of $E_{\mathrm{T}}^{\mathrm{GAP}}$.

All systematic uncertainties in each bin above the nominal point and all systematic uncertainties in each bin below the nominal point, except that of the calorimeter energy scale, were added in quadrature to give a total upper and lower total systematic uncertainty for each bin. These uncertainties were then added in quadrature to the

systematic uncertainty due to model dependence and the statistical uncertainty.

Nominal Cuts	σ	**+ Cuts**	**− Cuts**	± %
KINEMATIC CUTS				
Offline Cuts *with no* Corresponding True Level Cuts				
$\not{p}_T/E_{\mathrm{T}} < 2.0\,\mathrm{GeV}^{1/2}$	–	$\not{p}_T/E_{\mathrm{T}} < 2.2$	$\not{p}_T/E_{\mathrm{T}} < 1.8$	10.00
$y_e < 0.85$	–	$y_e < 0.90$	$y_e < 0.85$	5.88
$Z_{\mathrm{vtx}} < 40$ cm	–	$Z_{\mathrm{vtx}} < 50$	$Z_{\mathrm{vtx}} < 30$	25.00
Offline Cuts *with* Corresponding True Level Cuts				
$E_{\mathrm{T}}^1 > 5.10\,\mathrm{GeV}$	0.126	$E_{\mathrm{T}}^1 > 5.744$	$E_{\mathrm{T}}^1 > 4.456$	12.64
$E_{\mathrm{T}}^2 > 4.25\,\mathrm{GeV}$	0.126	$E_{\mathrm{T}}^2 > 4.784$	$E_{\mathrm{T}}^2 > 3.716$	12.56
$-2.40 < \eta^{1,2} < 2.40$	0.047	$-2.34 < \eta^{1,2} < 2.46$	$-2.46 < \eta^{1,2} < 2.34$	2.37
$\lvert\overline{\eta}\rvert > 0.75$	0.158	$\lvert\overline{\eta}\rvert > 0.809$	$\lvert\overline{\eta}\rvert > 0.691$	7.92
$2.50 < \Delta\eta < 4.00$	0.04	$2.550 < \Delta\eta < 4.080$	$2.450 < \Delta\eta < 3.920$	2.01
$0.20 < y_{\mathrm{JB}} < 0.75$	0.11	$0.211 < y_{\mathrm{JB}} < 0.789$	$0.189 < y_{\mathrm{JB}} < 0.711$	5.25
$E_{\mathrm{T}}^{\mathrm{GAP1}} < 0.60\,\mathrm{GeV}$	0.367	$E_{\mathrm{T}}^{\mathrm{GAP1}} < 0.820$	$E_{\mathrm{T}}^{\mathrm{GAP1}} < 0.380$	36.74
$E_{\mathrm{T}}^{\mathrm{GAP2}} < 1.20\,\mathrm{GeV}$	0.367	$E_{\mathrm{T}}^{\mathrm{GAP2}} < 1.641$	$E_{\mathrm{T}}^{\mathrm{GAP2}} < 0.759$	36.74
$E_{\mathrm{T}}^{\mathrm{GAP3}} < 1.80\,\mathrm{GeV}$	0.367	$E_{\mathrm{T}}^{\mathrm{GAP3}} < 2.461$	$E_{\mathrm{T}}^{\mathrm{GAP3}} < 1.139$	36.74
$E_{\mathrm{T}}^{\mathrm{GAP4}} < 2.40\,\mathrm{GeV}$	0.367	$E_{\mathrm{T}}^{\mathrm{GAP4}} < 3.282$	$E_{\mathrm{T}}^{\mathrm{GAP4}} < 1.518$	36.74
MODEL DEPENDENCE				
CS = 5.7%	–	7.1%	4.3%	25
CALORIMETER ENERGY SCALE UNCERTAINTY				
E_{T}	–	E_{T}+3%	E_{T}−3%	3.00
p_{T}	–	p_{T}+3%	p_{T}−3%	3.00
y_{JB}	–	y_{JB}+3%	y_{JB}−3%	3.00

Table 8.2: The modifications of kinematic cuts, amount of color-singlet used in the unfolding, and calorimeter energy scale, determined from HERWIG, used to estimate the systematic uncertainties. The nominal cuts on the variables, resolution of the variable, the varied cuts used in the systematic estimation, and the percent difference between the nominal and varied cut value are listed.

Number	**Variation**	**Number**	**Variation**
1	Nominal		
2	$E_{\mathrm{T}}+$	3	$E_{\mathrm{T}}-$
4	$\eta+$	5	$\eta-$
6	$\Delta\eta+$	7	$\Delta\eta-$
8	$\lvert\overline{\eta}\rvert+$	9	$\lvert\overline{\eta}\rvert-$
10	$y_{\mathrm{JB}}+$	11	$y_{\mathrm{JB}}-$
12	$\frac{\not{p}_{\mathrm{T}}}{\sqrt{E_{\mathrm{T}}}}+$	13	$\frac{\not{p}_{\mathrm{T}}}{\sqrt{E_{\mathrm{T}}}}-$
14	y_e+	15	y_e-
16	$Z_{vtx}+$	17	$Z_{vtx}-$
18	$E_{\mathrm{T}}^{\mathrm{GAP}}+$	19	$E_{\mathrm{T}}^{\mathrm{GAP}}-$
20	%CS+	21	%CS−
22	CAL Ene+	23	CAL Ene−

Table 8.3: The list of variations used in the measurement of systematic uncertainties corresponding to the points in the bins of the systematic plots. The first point in each bin corresponds to the nominal value, the second point corresponds to raising the cut on E_{T}, the third point corresponds to lowering the cut on E_{T}, and so on.

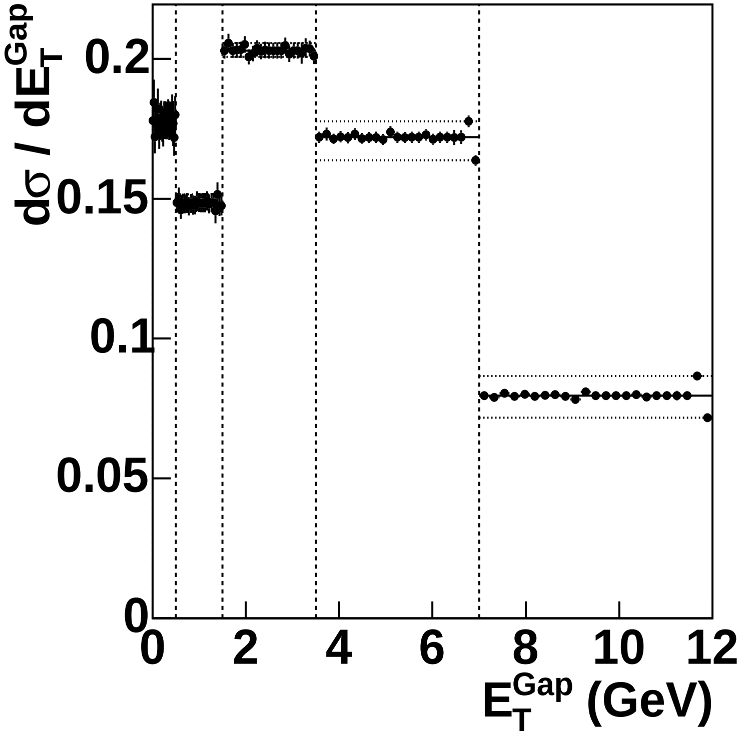


Figure 8.18: The inclusive cross section as a function of E_T^{GAP} plotted for each variation used in the calculation of the systematic uncertainty. The first (left-most) point in each bin shows the nominal cross section and each successive point shows the cross section calculated after applying a variation to one parameter in the measurement. The error bars show the statistical error on each point. The dashed horizontal lines correspond to the the highest and lowest systematic variation in each bin.

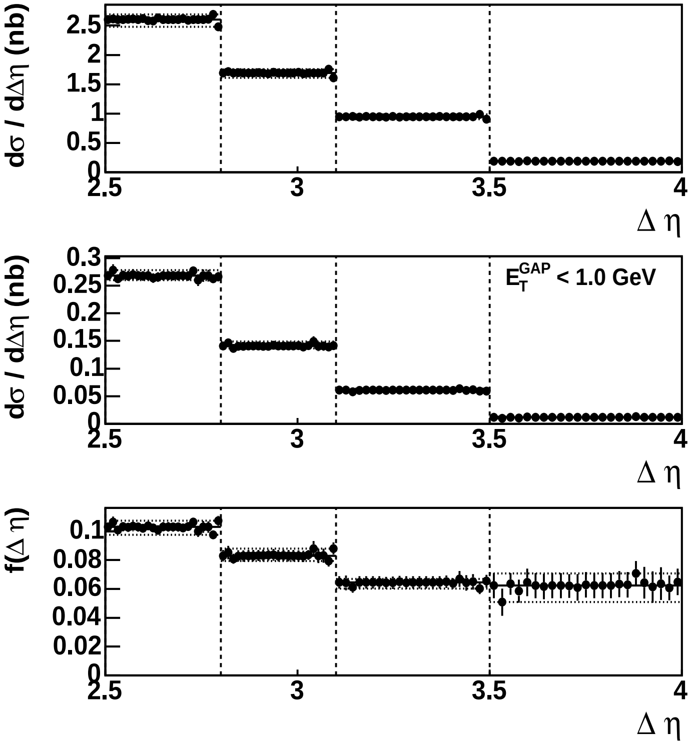


Figure 8.19: The inclusive cross section, gap cross section, and gap fraction as a function of $\Delta\eta$ plotted for each variation used in the study of the systematic uncertainty. The first (left-most) point in each bin shows the nominal cross section and each successive point shows the cross section calculated after applying a variation to one parameter in the measurement. The error bars show the statistical error on each point. The dashed horizontal lines correspond to the the highest and lowest systematic variation in each bin.

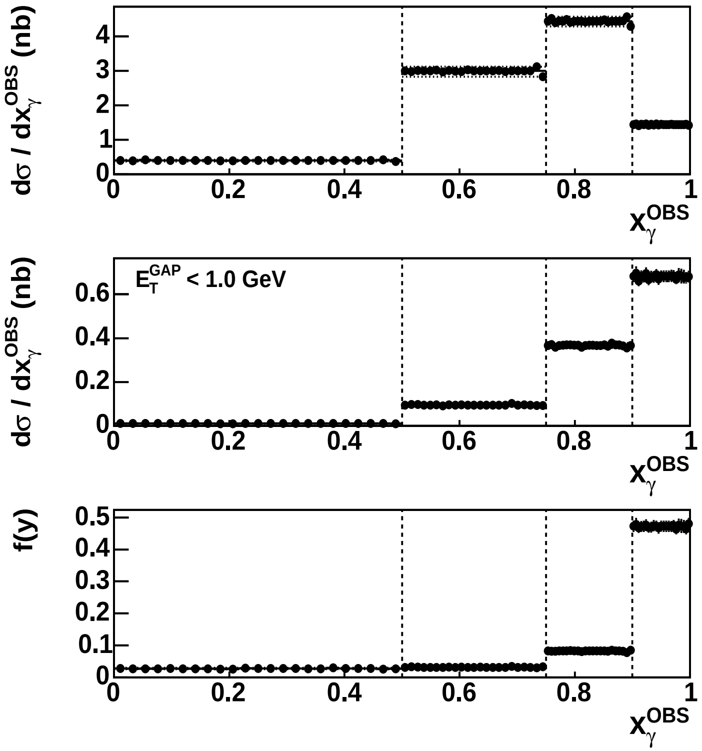


Figure 8.20: The inclusive cross section, gap cross section, and gap fraction as a function of x_γ^{OBS} plotted for each variation used in the study of the systematic uncertainty. The first (left-most) point in each bin shows the nominal cross section and each successive point shows the cross section calculated after applying a variation to one parameter in the measurement. The error bars show the statistical error on each point. The dashed horizontal lines correspond to the the highest and lowest systematic variation in each bin.

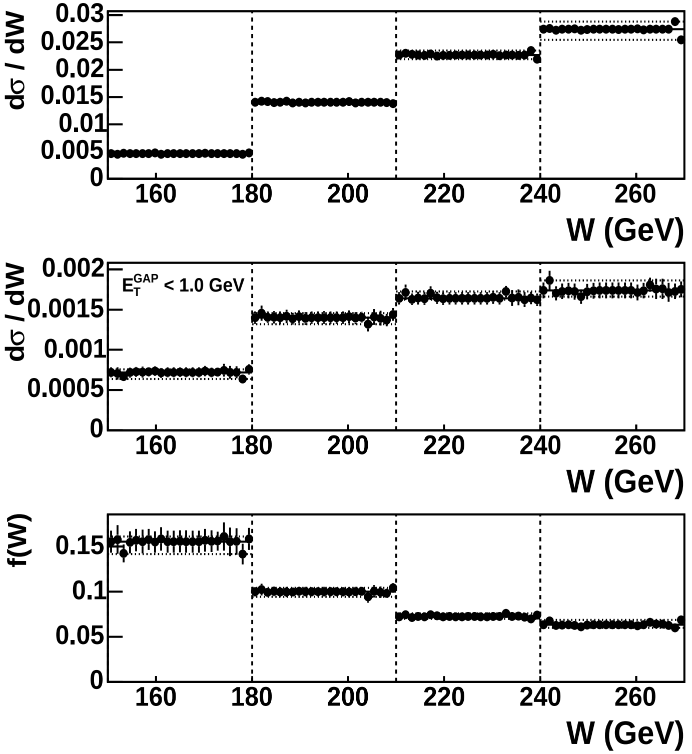


Figure 8.21: The inclusive cross section, gap cross section, and gap fraction as a function of W plotted for each variation used in the study of the systematic uncertainty. The first (left-most) point in each bin shows the nominal cross section and each successive point shows the cross section calculated after applying a variation to one parameter in the measurement. The error bars show the statistical error on each point. The dashed horizontal lines correspond to the the highest and lowest systematic variation in each bin.

Chapter 9

Results

The data was corrected, as described in Section 8.1, with the average correction factors of the Pythia and Herwig MC samples detailed in Section 7.2. The inclusive dijet cross section as a function of $E_{\rm T}^{\rm GAP}$ is presented in Fig. 9.1. At low $E_{\rm T}^{\rm GAP}$ values, where the color-singlet contribution should be most pronounced, the data demonstrates a clear rise towards the small $E_{\rm T}^{\rm GAP}$ values. In order to estimate the amount of the color singlet, the direct and resolved components of each MC were mixed according to their cross sections, as predicted by the MC, to give the color-non-singlet MC sample. The color-non-singlet and color-singlet MC samples were then fitted to the data according to

$$F = P_1 \frac{d\sigma^{CS}}{dE_{\rm T}^{\rm GAP}} + P_2 \frac{d\sigma^{NCS}}{dE_{\rm T}^{\rm GAP}},$$

where P_1 and P_2 are the free parameters of the fit. The best fit to the data resulted in $P_1 = 1.31 \pm 0.01$ and $P_2 = 327 \pm 20$ for Pythia and $P_1 = 1.93 \pm 0.01$ and $P_2 = 1.02 \pm 0.13$ for Herwig. The large color-singlet scale factor, P_2, for Pythia is due to the use of the high-t γ exchange model to simulate the rapidity gap topology. These scaling parameters were used in this analysis when comparing data to the MC predictions. The amount of the color singlet contribution to the total cross section, estimated by integrating the MC predictions over the entire $E_{\rm T}^{\rm GAP}$ range, was found to be in the region $2-3\%$ for both Pythia and Herwig.

The inclusive dijet cross section, the gap cross section, and the gap fraction, as a function of the separation, $\Delta\eta$, of the two leading jets, are presented in Fig. 9.2 for $E_{\rm T}^{\rm CUT} = 1\,{\rm GeV}$. Both cross sections and the gap fractions decrease as a function of $\Delta\eta$. In the inclusive cross section, both MC models with and without color-singlet exchange describe the data equally well, but for the gap cross section, the MC models without color-singlet exchange fall below the data while the MC models with color-singlet exchange describe the data. The contribution of color-singlet exchange to the total gap fraction increases as the dijet separation increases from 2.5 to 4 units in pseudorapidity.

Figure 9.3 shows the gap fraction as a function of the dijet separation, $\Delta\eta$, for the four values of $E_T^{\rm CUT} = 0.5, 1.0, 1.5$ and $2\,{\rm GeV}$. The data first fall and then level out as $\Delta\eta$ increases for all values of $E_T^{\rm CUT}$ except $E_T^{\rm CUT} = 0.5$, where the data is almost constant with $\Delta\eta$. The behavior at $E_T^{\rm CUT} = 0.5$ is due to the fraction color-non-singlet processes being small compared to the color-singlet processes in all regions of $\Delta\eta$. The predictions of Pythia and Herwig without color-singlet exchange lie below the data over the entire $\Delta\eta$ range. With the addition of the color-singlet contribution both MCs describe the data well.

In order to estimate the color-singlet contribution to the gap fraction, the ZEUS data were compared to the MC predictions as shown in Fig. 9.3. Depending on the MC used in the comparison, the color-singlet contribution can be estimated to be at the level of 1 to 4%, increasing slightly with increasing $\Delta\eta$. These numbers are well in agreement with the value estimated using the total cross section.

The data are consistent with the previously published ZEUS results [85], shown in Fig. 9.12, where the gap definition was based on multiplicity of the objects in the gap and the gap fraction was measured to be $0.11 \pm 0.02\ ^{+0.01}_{-0.02}$ for $3.5 < \Delta\eta < 4$. The new results presented here, with much better statistical and systematic errors, agree with the previous H1 measurement [86], where the gap was defined using the transverse energy in the gap as in the current analysis, but with slightly different kinematic cuts. The H1 measurement is shown in Fig. 9.13. In order to compare the ZEUS and H1 measurements, the ZEUS data was scaled by the ratio of hadron level HERWIG MC generated with the ZEUS event selection and scaled to fit the ZEUS data to hadron level HERWIG generated using the H1 event selection and scaled by the same amount as the ZEUS data. It can be seen in Fig. 9.14 that the ZEUS and H1 data agree within errors and that the current ZEUS measurement has much smaller statistical and systematic errors than the H1 measurement. The gap fractions also exhibit a smaller decrease as $\Delta\eta$ increases, which is due to the removal of the $|\overline{\eta}|$ cut.

The cross sections and gap fractions were also measured as a function of x_γ^{OBS} for comparisons with other experiments and pp measurements, since at low values of x_γ^{OBS} the photon appears as an hadronic object. These results are presented in Figs. 9.4 and 9.5 for four different values of $E_{\mathrm{T}}^{\mathrm{CUT}}$. The gap fraction decreases as a function of x_γ^{OBS} and the data are reasonably well described by both MC models. In the x_γ^{OBS} region below 0.75 HERWIG predicts larger cross sections than PYTHIA. Although the data have sufficiently small errors, the difference in the model predictions preclude an accurate determination of the color-singlet contribution to the gap fraction and its behavior as a function of x_γ^{OBS}.

The same is also valid for the W (the photon-proton center of mass energy) dependence, which allows for comparison with experiments at different energies. The W dependence of the cross sections and gap fractions are presented in Figs. 9.6 and 9.7. Both the cross sections and the gap fractions are described by the MC. The gap fraction falls with increasing W and the color-singlet contribution can be estimated to be at the level of approximately 2 to 4 %.

In order to compare with pp measurements the $\Delta\eta$ and W behavior was investigated in the resolved enhanced region ($x_\gamma^{\mathrm{OBS}} < 0.75$). Figure 9.8 shows the cross sections as a function of $\Delta\eta$ in resolved enhanced region for $E_{\mathrm{T}}^{\mathrm{GAP}} < 1\,\mathrm{GeV}$. The MC models do not describe the data well. Both MCs fail to describe the inclusive cross section and PYTHIA does not describe the gap cross section either. In the ratio of gap to inclusive cross sections, performed to calculate the gap fraction, the effects cancel out and the gap fraction as a function of $\Delta\eta$ is still reasonably well described. Figure 9.9 shows the gap fractions as a function of $\Delta\eta$ in the resolved enhanced region for the four different $E_{\mathrm{T}}^{\mathrm{CUT}}$ values. For $E_{\mathrm{T}}^{\mathrm{GAP}} < 0.5\,\mathrm{GeV}$ and $E_{\mathrm{T}}^{\mathrm{GAP}} < 1.0\,\mathrm{GeV}$, both MC models predict almost no contribution to the gap fractions from the non-color-singlet component at high values of $\Delta\eta$. These points are therefore less biased by the model predictions and can be used to estimate the color-singlet contribution in the resolved enhanced sample. The fraction of events with a gap lies in the range 1-2% with large uncertainties mainly due to unfolding using different MC models. These values, within errors, agree well with the measurement for the total x_γ^{OBS} region.

The W behavior in the resolved enhanced sample is presented in Figs. 9.10

and 9.11. The inclusive cross section is not well described at high values of W, although Herwig describes the gap cross section better than Pythia. Both MC models predict that the gap fraction should fall as W increases, but the data show an almost flat behavior.

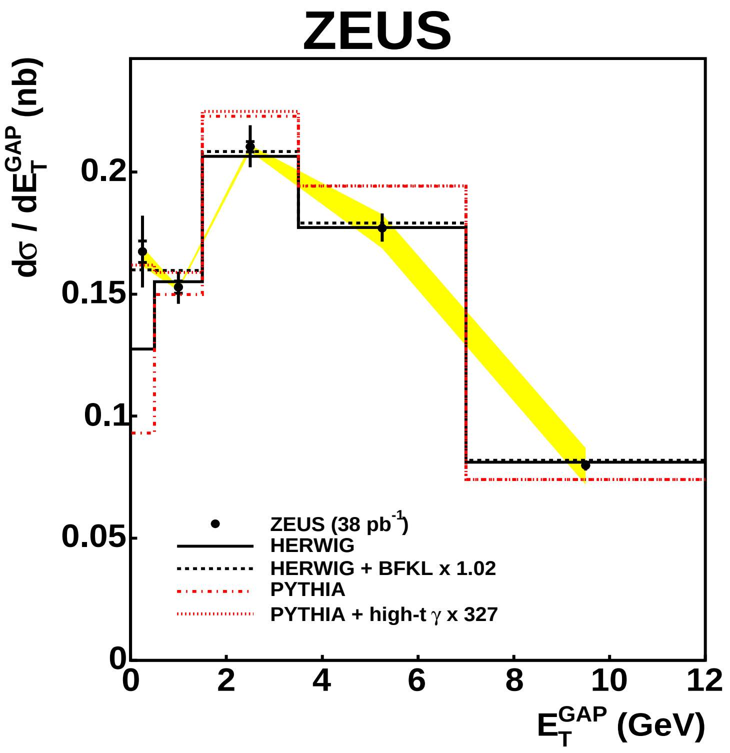


Figure 9.1: The inclusive dijet cross section differential in $E_T^{\rm GAP}$ The black circles represent the ZEUS data, with the inner error bars representing the statistical errors and the outer error bars representing the statistical and systematic uncertainties added in quadrature. The solid black line shows the prediction of Herwig and the black dashed line shows the prediction of Herwig plus BFKL Pomeron exchange. The dot-dashed line shows the prediction of Pythia and the dotted line shows the prediction of Pythia plus high-t photon exchange. The band shows the calorimeter energy scale uncertainty.

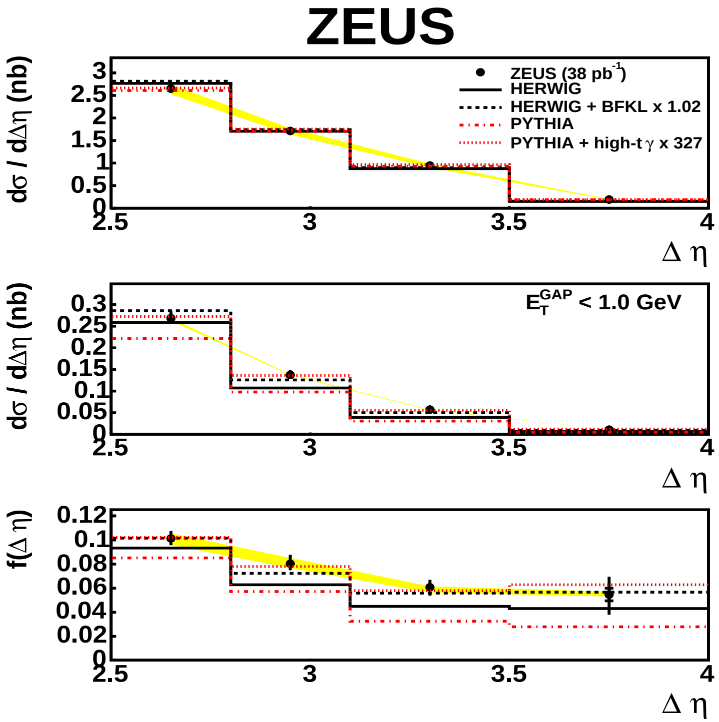


Figure 9.2: The top plot is the inclusive dijet cross section differential in $\Delta\eta$, the middle plot is gap dijet cross section differential in $\Delta\eta$, and the bottom plot is the gap fraction in $\Delta\eta$. The black circles represent the ZEUS data, with the inner error bars representing the statistical errors and the outer error bars representing the statistical and systematic uncertainties added in quadrature. The solid black line shows the prediction of HERWIG and the black dashed line shows the prediction of HERWIG plus BFKL Pomeron exchange. The dot-dashed line shows the prediction of PYTHIA and the dotted line shows the prediction of PYTHIA plus high-t photon exchange. The band shows the calorimeter energy scale uncertainty.

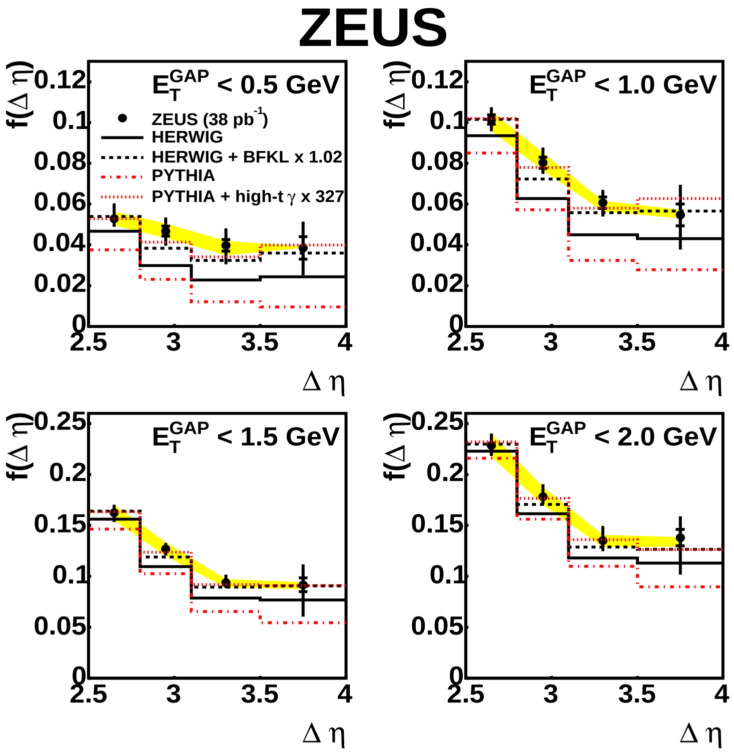


Figure 9.3: The gap fraction, f, in $\Delta\eta$. The black circles represent the ZEUS data, with the inner error bars representing the statistical errors and the outer error bars representing the statistical and systematic uncertainties added in quadrature. The solid black line shows the prediction of HERWIG and the black dashed line shows the prediction of HERWIG plus BFKL Pomeron exchange. The dot-dashed line shows the prediction of PYTHIA and the dotted line shows the prediction of PYTHIA plus high-t photon exchange. The band shows the calorimeter energy scale uncertainty.

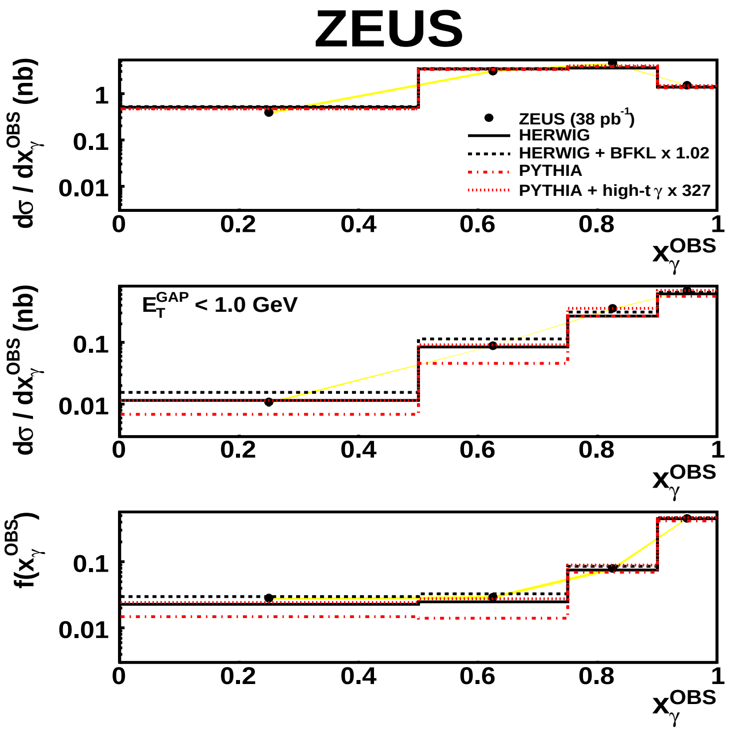


Figure 9.4: The top plot is the inclusive dijet cross section differential in $x_\gamma^{\rm OBS}$, the middle plot is gap dijet cross section differential in $x_\gamma^{\rm OBS}$, and the bottom plot is the gap fraction in $x_\gamma^{\rm OBS}$. The black circles represent the ZEUS data, with the inner error bars representing the statistical errors and the outer error bars representing the statistical and systematic uncertainties added in quadrature. The solid black line shows the prediction of Herwig and the black dashed line shows the prediction of Herwig plus BFKL Pomeron exchange. The dot-dashed line shows the prediction of Pythia and the dotted line shows the prediction of Pythia plus high-t photon exchange. The band shows the calorimeter energy scale uncertainty.

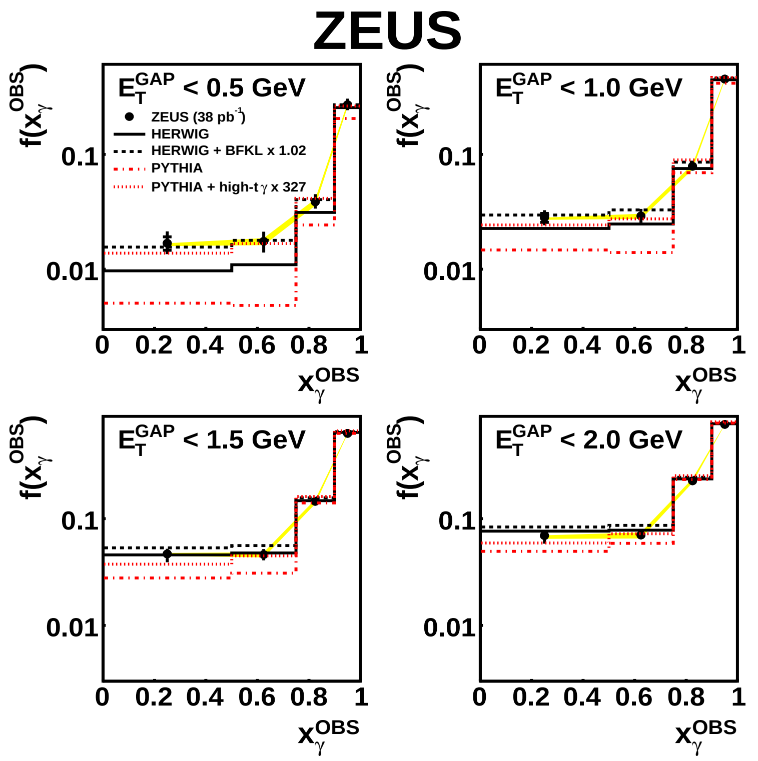


Figure 9.5: The gap fraction, f, in $x_\gamma^{\rm OBS}$. The black circles represent the ZEUS data, with the inner error bars representing the statistical errors and the outer error bars representing the statistical and systematic uncertainties added in quadrature. The solid black line shows the prediction of Herwig and the black dashed line shows the prediction of Herwig plus BFKL Pomeron exchange. The dot-dashed line shows the prediction of Pythia and the dotted line shows the prediction of Pythia plus high-t photon exchange. The band shows the calorimeter energy scale uncertainty.

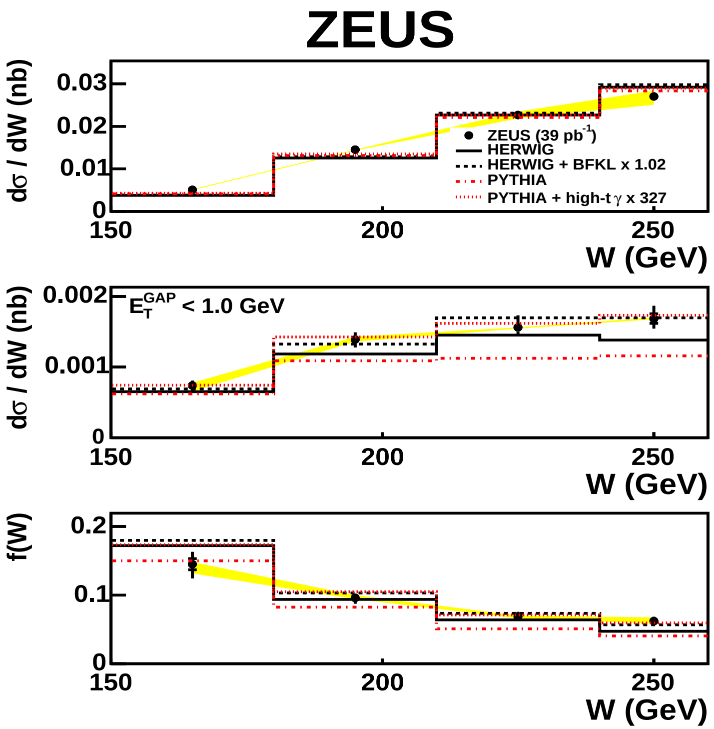


Figure 9.6: The top plot is the inclusive dijet cross section differential in W, the middle plot is gap dijet cross section differential in W, and the bottom plot is the gap fraction in W. The black circles represent the ZEUS data, with the inner error bars representing the statistical errors and the outer error bars representing the statistical and systematic uncertainties added in quadrature. The solid black line shows the prediction of HERWIG and the black dashed line shows the prediction of HERWIG plus BFKL Pomeron exchange. The dot-dashed line shows the prediction of PYTHIA and the dotted line shows the prediction of PYTHIA plus high-t photon exchange. The band shows the calorimeter energy scale uncertainty.

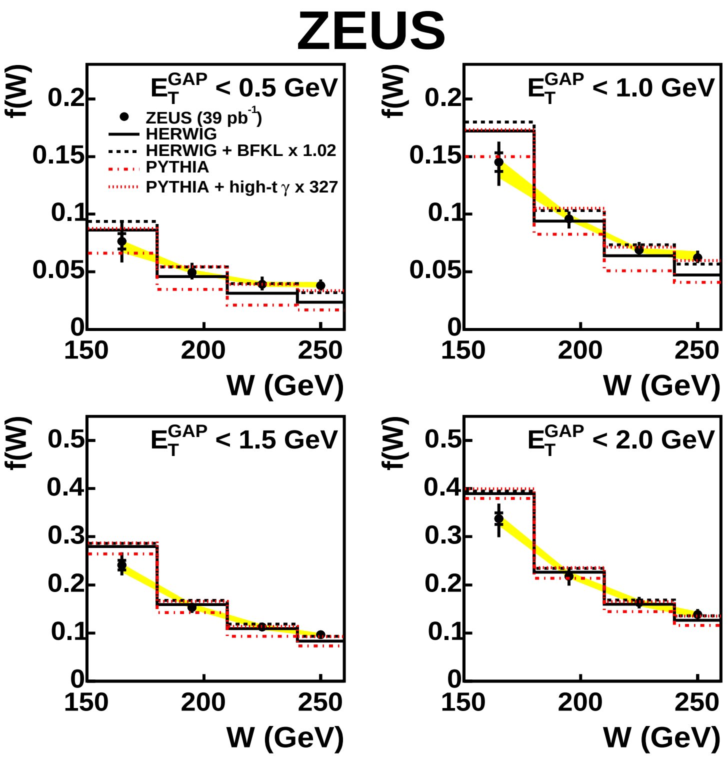


Figure 9.7: The gap fraction, f, in W. The black circles represent the ZEUS data, with the inner error bars representing the statistical errors and the outer error bars representing the statistical and systematic uncertainties added in quadrature. The solid black line shows the prediction of HERWIG and the black dashed line shows the prediction of HERWIG plus BFKL Pomeron exchange. The dot-dashed line shows the prediction of PYTHIA and the dotted line shows the prediction of PYTHIA plus high-t photon exchange. The band shows the calorimeter energy scale uncertainty.

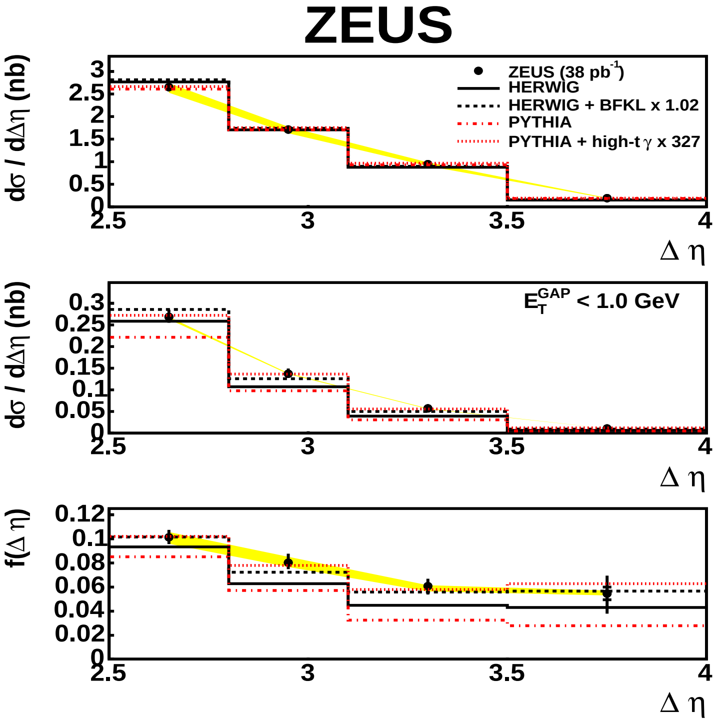


Figure 9.8: The top plot is the inclusive dijet cross section differential in $\Delta\eta$, the middle plot is gap dijet cross section differential in $\Delta\eta$, and the bottom plot is the gap fraction in $\Delta\eta$ for $x_\gamma^{\rm OBS} < 0.75$. The black circles represent the ZEUS data, with the inner error bars representing the statistical errors and the outer error bars representing the statistical and systematic uncertainties added in quadrature. The solid black line shows the prediction of Herwig and the black dashed line shows the prediction of Herwig plus BFKL Pomeron exchange. The dot-dashed line shows the prediction of Pythia and the dotted line shows the prediction of Pythia plus high-t photon exchange. The band shows the calorimeter energy scale uncertainty.

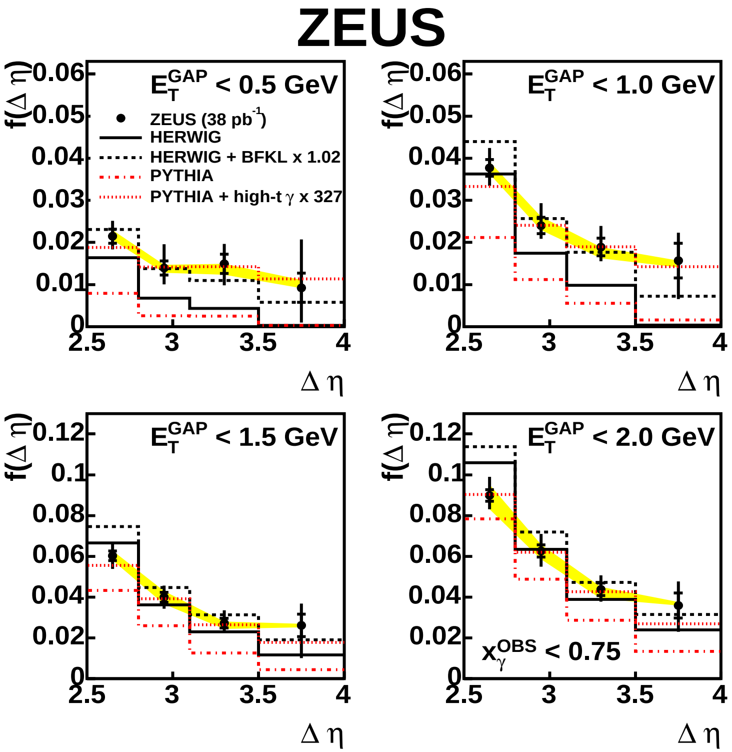


Figure 9.9: The gap fraction, f, in $\Delta\eta$ for $x_\gamma^{\rm OBS} < 0.75$. The black circles represent the ZEUS data, with the inner error bars representing the statistical errors and the outer error bars representing the statistical and systematic uncertainties added in quadrature. The solid black line shows the prediction of Herwig and the black dashed line shows the prediction of Herwig plus BFKL Pomeron exchange. The dot-dashed line shows the prediction of Pythia and the dotted line shows the prediction of Pythia plus high-t photon exchange. The band shows the calorimeter energy scale uncertainty.

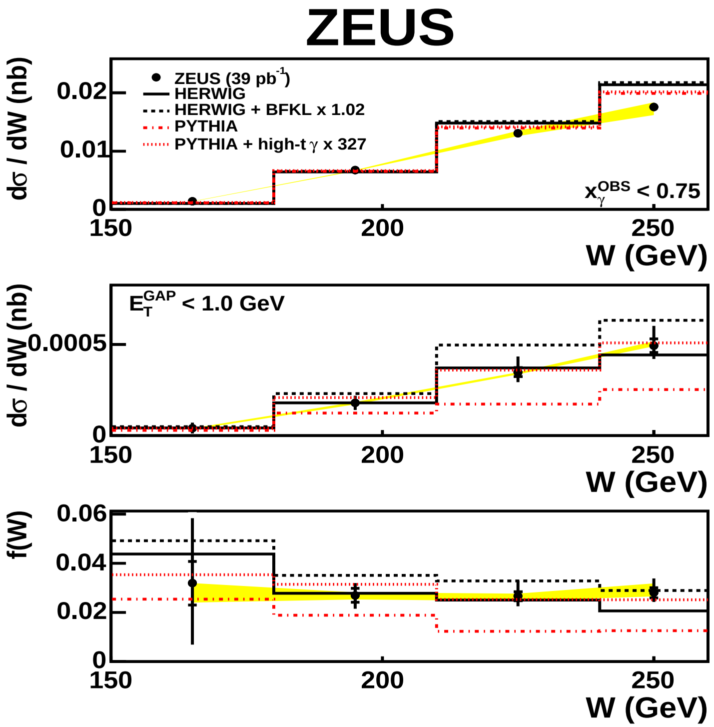


Figure 9.10: The top plot is the inclusive dijet cross section differential in W, the middle plot is gap dijet cross section differential in W, and the bottom plot is the gap fraction in W for $x_\gamma^{\mathrm{OBS}} < 0.75$. The black circles represent the ZEUS data, with the inner error bars representing the statistical errors and the outer error bars representing the statistical and systematic uncertainties added in quadrature. The solid black line shows the prediction of Herwig and the black dashed line shows the prediction of Herwig plus BFKL Pomeron exchange. The dot-dashed line shows the prediction of Pythia and the dotted line shows the prediction of Pythia plus high-t photon exchange. The band shows the calorimeter energy scale uncertainty.

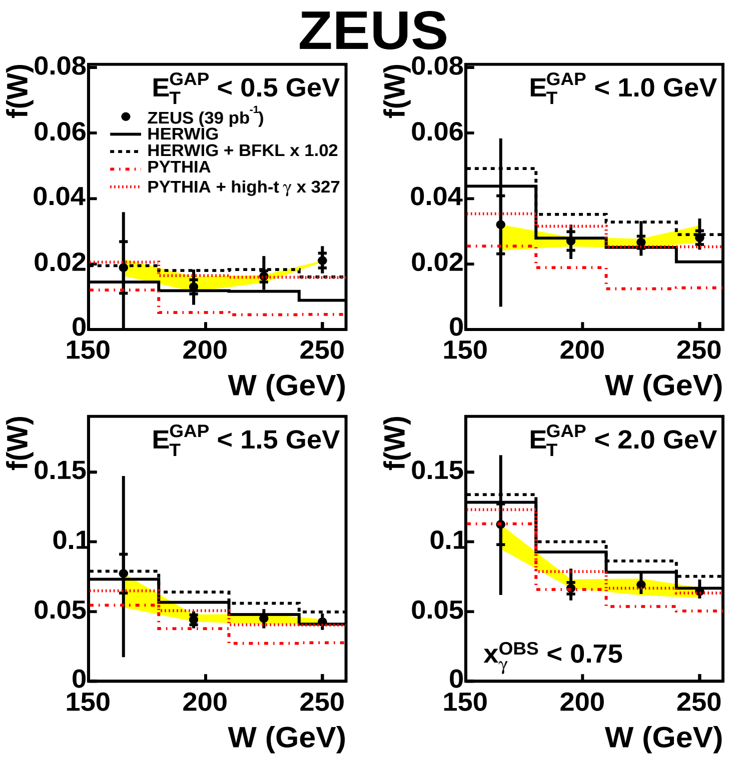


Figure 9.11: The gap fraction, f, in W for $x_\gamma^{\mathrm{OBS}} < 0.75$. The black circles represent the ZEUS data, with the inner error bars representing the statistical errors and the outer error bars representing the statistical and systematic uncertainties added in quadrature. The solid black line shows the prediction of Herwig and the black dashed line shows the prediction of Herwig plus BFKL Pomeron exchange. The dot-dashed line shows the prediction of Pythia and the dotted line shows the prediction of Pythia plus high-t photon exchange. The band shows the calorimeter energy scale uncertainty.

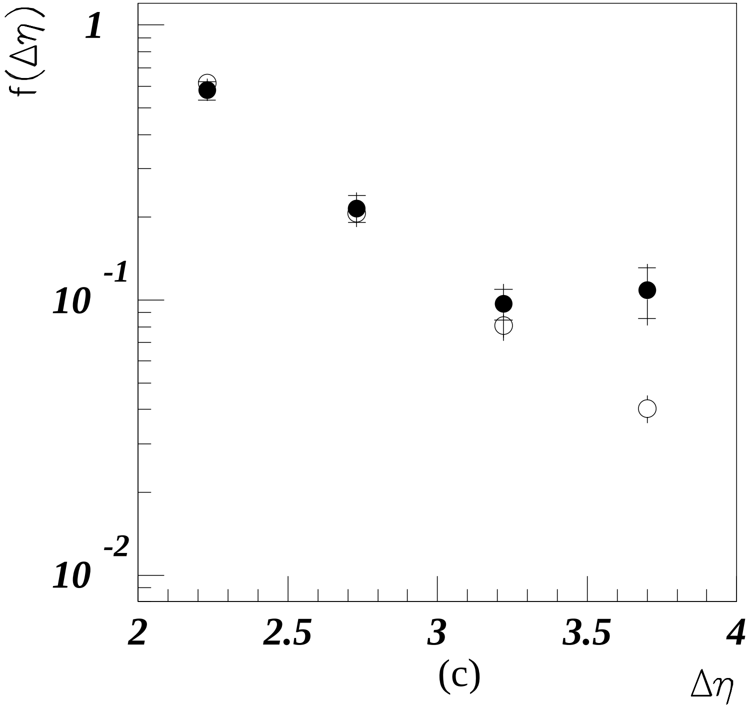


Figure 9.12: The gap fraction as a function of $\Delta\eta$ measured by ZEUS in 1995 [85]. The black circles represent the ZEUS data, with the inner error bars representing the statistical errors and the outer error bars representing the statistical and systematic uncertainties added in quadrature. The white circles show Pythia MC without color-singlet exchange. The gap is defined by multiplicity and not jet E_{T}.

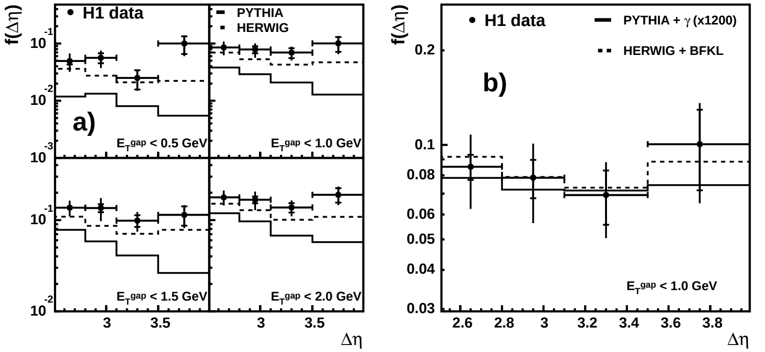


Figure 9.13: The gap fraction as a function of $\Delta\eta$ measured by H1. The black circles represent the H1 data, with the inner error bars representing the statistical errors and the outer error bars representing the statistical and systematic uncertainties added in quadrature. The solid black line shows the prediction of Pythia and the black dashed line shows the prediction of Herwig. In the four plots on the left, each having a different value of $E_{\mathrm{T}}^{\mathrm{CUT}}$, the MC does not include color-singlet exchange. In the plot on the right, where $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.0\,\mathrm{GeV}$, the MC includes color-singlet exchange.

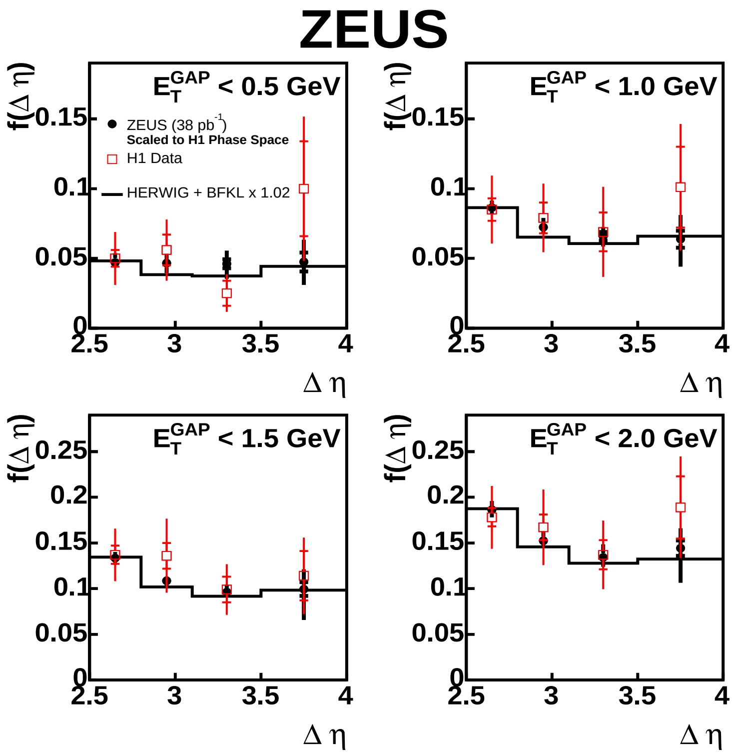


Figure 9.14: Comparisons between ZEUS and H1 gap fractions as a function of $\Delta\eta$. The black circles represent the ZEUS data scaled to the H1 results, the red open squares show the H1 data. The inner error bars on the data represent the statistical errors and the outer error bars represent the statistical and systematic uncertainties added in quadrature. The solid black line shows the prediction of HERWIG with color-singlet-exchange included.

Chapter 10

Conclusions

Cross sections of dijet photoproduction events, where the two jets with the highest transverse energy are separated by a large region in pseudorapidity and have very little transverse energy between them, show a clear excess when compared to the predictions of standard photoproduction MC models. The same models, with the inclusion of a color-singlet exchange sample on the level of $2-3\%$, are able to describe the data. The same amount of color-singlet exchange is observed by direct comparison of the data to the MC predictions for the gap fractions.

The main systematic uncertainties in the measurement come from using PYTHIA instead of HERWIG in unfolding of the cross sections. The difference in the model predictions preclude an exact determination of the color-singlet contribution and its behavior as a function of different kinematic variables such as x_γ^{OBS} or W.

The level of color-singlet exchange observed in the current analysis is consistent with the previously published ZEUS and H1 results and the level of color-singlet exchanged in the resolved region is consistent with results from the Tevatron.

Appendix A

Kinematics

A.1 Rapidity and Pseudorapidity

A.1.1 Derivation of Rapidity

It is possible to write the four-momentum of a particle as

$$p^{\mu} = (E, p_x, p_y, p_z) = (m_T \cosh y, p_T \sin\phi, p_T \cos\phi, m_T \sinh y), \tag{A.1}$$

where the transverse mass $m_T = \sqrt{p_T^2 + m^2}$ and the transverse momentum $p_T = \sqrt{p_x^2 + p_y^2}$. The z axis specifies the longitudinal direction, ϕ is the azimuthal angle with respect to the z axis, and y is the rapidity. The E and p_z components of Equation A.1 can be written as

$$E = \sqrt{p_T^2 + m^2}\cosh y = \sqrt{p_T^2 + m^2}\left(\frac{e^{-y} + e^{y}}{2}\right) \tag{A.2}$$

$$p_z = \sqrt{p_T^2 + m^2}\sinh y = \sqrt{p_T^2 + m^2}\left(\frac{e^{-y} - e^{y}}{2}\right). \tag{A.3}$$

Therefore,

$$\frac{E + p_z}{E - p_z} = e^{2y}. \tag{A.4}$$

Solving for the rapidity,

$$y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right). \tag{A.5}$$

A.1.2 Derivation of Pseudorapidity

In the limit that $m \to 0$, $E^2 = \boldsymbol{p^2}$. The four-momentum of a massless particle can be written as

$$p^{\mu} = (E, p_x, p_y, p_z) = (E, E\sin\theta\cos\phi, E\sin\theta\sin\phi, E\cos\phi). \tag{A.6}$$

Therefore,

$$E + p_z = E\left(1 + \cos\theta\right) \tag{A.7}$$

$$E - p_z = E\left(1 - \cos\theta\right). \tag{A.8}$$

In the massless limit, the rapidity from Equation A.5 becomes

$$\eta = \frac{1}{2}\ln\left(\frac{1 + \cos\theta}{1 - \cos\theta}\right), \tag{A.9}$$

where the variable η is now used instead of y. Using the trigonometric relations

$$\sin^2\frac{\psi}{2} = \frac{1}{2}\left(1 - \cos\psi\right) \tag{A.10}$$

$$\cos^2\frac{\psi}{2} = \frac{1}{2}\left(1 + \cos\psi\right) \tag{A.11}$$

and substituting $\tan^2\frac{\psi}{2} = \left(\frac{1-\cos\psi}{1+\cos\psi}\right)$ into Equation A.9, one obtains the expression for pseudorapidity

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right). \tag{A.12}$$

A.1.3 The Boost Invariance of Rapidity Differences

A particle boosted in the longitudinal (z) direction undergoes a Lorentz Transformation

$$\begin{pmatrix} E' \\ p'_x \\ p'_y \\ p'_z \end{pmatrix} = \begin{pmatrix} \gamma & 0 & 0 & -\beta\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta\gamma & 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} E \\ p_x \\ p_y \\ p_z \end{pmatrix} \tag{A.13}$$

where $\gamma \equiv \frac{1}{\sqrt{1-v^2/c^2}}$ and $\beta \equiv v/c$. Substituting

$$E' = \gamma E - \beta\gamma p_z \tag{A.14}$$

$$p'_z = -\beta\gamma E + \gamma p_z. \tag{A.15}$$

into Equation A.5, one obtains

$$y' = \frac{1}{2}\ln\left(\frac{E'+p'_z}{E'-p'_z}\right) \tag{A.16}$$

$$= \frac{1}{2}\ln\left(\frac{(E+p_z)(\gamma-\beta\gamma)}{(E-p_z)(\gamma+\beta\gamma)}\right) \tag{A.17}$$

$$= \frac{1}{2}\ln\left(\frac{E+p_z}{E-p_z}\right) + \frac{1}{2}\ln\left(\frac{1-\beta}{1+\beta}\right) \tag{A.18}$$

$$= y + \frac{1}{2}\ln\left(\frac{1-\beta}{1+\beta}\right). \tag{A.19}$$

Under a longitudinal boost, the rapidity changes only by the addition of a constant. Differences in rapidity are therefore invariant under a boost,

$$y'_2 - y'_1 = y_2 - y_1. \tag{A.20}$$

A.2 Relation between t and p_T in a hard process

For a process $A(a) + B(b) \rightarrow C(c) + D(d)$, where the quantities in parenthesis are the four-momenta of the particles, the Lorentz invariant Mandelstam variable t is defined

as $t \equiv (a-c)^2$. In the center-of-mass frame, where particles A and B have the same energy, and particle A is traveling in the $+z$ direction and particle B is traveling in the $-z$ direction,

$$a = (E, 0, 0, E) \tag{A.21}$$

$$b = (E, 0, 0, -E) \tag{A.22}$$

$$c = (E, p_x, p_y, E\cos\theta) \tag{A.23}$$

$$d = (E, -p_x, -p_y, -E\cos\theta) \tag{A.24}$$

where θ is the scattering angle and the mass is considered negligible. Then,

$$-t = p_T^2 + E^2(1-\cos\theta)^2, \tag{A.25}$$

and therefore

$$|t| \geq p_T^2. \tag{A.26}$$

A large transverse momentum of the outgoing partons signifies that t is a hard scale.

A.3 Derivation of x_γ and x_p

In an ep collision, the fraction of the photon's energy involved in the hard interaction is represented by x_γ and the fraction of the proton's energy involved in the hard interaction is represented by x_p.

From conservation of energy,

$$x_\gamma E_\gamma + x_p E_p = \sum_i E_i \tag{A.27}$$

$$= \sum_i \frac{E_i \sin\theta_i}{\sin\theta_i} \tag{A.28}$$

$$= \sum_i \frac{E_{\mathrm{T},i}}{\sin\theta_i} \tag{A.29}$$

where the sum is performed over the outgoing partons from the collision between the parton from the proton and the parton from the proton. From conservation of momentum,

$$x_\gamma |\boldsymbol{p}_\gamma| \cos(\pi - \theta) + x_p |\boldsymbol{p}_p| \cos\theta = \sum_i |\boldsymbol{p}_i| \cos\theta. \tag{A.30}$$

In the approximation that the photon and proton are collinear to the beam-line ($\theta = 0$) and in the massless limit ($\boldsymbol{p} \to E$),

$$\begin{aligned} x_\gamma E_\gamma - x_p E_p &= \sum_i E_i \cos\theta_i && \text{(A.31)} \\ &= \sum_i E_i \frac{\sin\theta_i}{\tan\theta_i} && \text{(A.32)} \\ &= \sum_i \frac{E_{\mathrm{T},i}}{\tan\theta_i} && \text{(A.33)} \end{aligned}$$

Adding Equations A.27 and A.31, one obtains

$$2x_\gamma E_\gamma = \sum_i E_{\mathrm{T},i} \left[\frac{1}{\sin\theta_i} - \frac{1}{\tan\theta_i}\right], \tag{A.34}$$

and subtracting them, one obtains

$$2x_p E_p = \sum_i E_{\mathrm{T},i} \left[\frac{1}{\sin\theta_i} + \frac{1}{\tan\theta_i}\right]. \tag{A.35}$$

Using the trigonometric relations A.10 and A.11,

$$\begin{aligned} \frac{1}{\sin\theta_i} - \frac{1}{\tan\theta_i} &= \frac{1 - \cos\theta_i}{\sin\theta_i} = \tan\frac{\theta_i}{2} && \text{(A.36)} \\ \frac{1}{\sin\theta_i} + \frac{1}{\tan\theta_i} &= \frac{1 + \cos\theta_i}{\sin\theta_i} = \left[\tan\frac{\theta_i}{2}\right]^2, && \text{(A.37)} \end{aligned}$$

and substituting $\eta = -\ln\left(\tan\frac{\theta}{2}\right)$ and $E_\gamma = E_e$, one obtains

$$\begin{aligned} x_\gamma &= \frac{\sum_i E_{\mathrm{T},i} e^{-\eta_i}}{2yE_e} && \text{(A.38)} \\ x_p &= \frac{\sum_i E_{\mathrm{T},i} e^{\eta_i}}{2E_p}. && \text{(A.39)} \end{aligned}$$

Appendix B

Method of Least Squares

The method of least squares finds a function best approximating a set of data points by minimizing the the square of the difference between the function and the data. The sum of the squares of the distance perpendicular to the y axis from a function f to a set of N data points can be written as

$$R^2 = \sum_{i=1}^{N} \left[z_i - f(x_i, a_1, a_2, \ldots, a_N, y_i, b_1, b_2, \ldots, b_N, \ldots) \right], \tag{B.1}$$

where z_i are the data data points and f is the function dependent on the variables x_i, y_i, …, with corresponding adjustable parameters a_i, b_i, …. The condition that R is a minimum is that

$$\frac{\partial R}{\partial a_i} = 0, \quad \frac{\partial R}{\partial b_i} = 0, \quad \ldots \tag{B.2}$$

for $i = 1, \ldots, N$.

B.1 Binned Data

If it is assumed that each bin in a histogram is independent of the other bin and that the bins have a Gaussian distribution around the model $f(x_i, \ldots, a_i, \ldots)$ with a known

standard deviation, σ_i, R has a χ^2 distribution and one can write

$$\chi^2 = \sum_{i=1}^{N} \left[\frac{z_i - f(x_i, \ldots a_1, \ldots)}{\sigma_i} \right]^2, \tag{B.3}$$

where the sum is over the N bins in the histogram. The method of least squares can be considered a χ^2 minimization in this case.

B.2 Combining Two Histograms to Match the Data

In the analysis described in this thesis, it was desired to combine two Monte Carlo histograms, each with a single weight, such that they gave the best fit to the data. In this case,

$$f = a x_i + b y_i, \tag{B.4}$$

and therefore,

$$\chi^2 = \sum_{i=1}^{N} \left[\frac{z_i - a x_i - b y_i}{\sigma_i} \right]^2. \tag{B.5}$$

χ^2 is minimized by setting $\frac{\partial \chi^2}{\partial a_i} = 0$ and $\frac{\partial \chi^2}{\partial b_i} = 0$ and solving for a and b. Doing so gives

$$a = \frac{\sum \frac{x_i y_i}{\sigma_i} \sum \frac{y_i z_i}{\sigma_i} - \sum \frac{x_i z_i}{\sigma_i} \sum \frac{y_i y_i}{\sigma_i}}{\sum \frac{x_i y_i}{\sigma_i} \sum \frac{x_i y_i}{\sigma_i} - \sum \frac{x_i x_i}{\sigma_i} \sum \frac{y_i y_i}{\sigma_i}} \tag{B.6}$$

$$b = \frac{\sum \frac{x_i x_i}{\sigma_i} \sum \frac{y_i z_i}{\sigma_i} - \sum \frac{x_i z_i}{\sigma_i} \sum \frac{x_i y_i}{\sigma_i}}{\sum \frac{x_i x_i}{\sigma_i} \sum \frac{y_i y_i}{\sigma_i} - \sum \frac{x_i y_i}{\sigma_i} \sum \frac{x_i y_i}{\sigma_i}}, \tag{B.7}$$

where the sums are performed from $i = 1 \ldots N$. The standard deviation is given by $\sigma_i = \delta z_i^2$, where δz_i is the error on the bin i.

Appendix C

Derivation of Errors

C.1 Errors on the Purity, Efficiency, and Correction Factor

The purity ($\mathcal{P}$), efficiency ($\mathcal{E}$), and correction factor ($\mathcal{C}$) are defined as[1]

$$\begin{aligned} \mathcal{P} &= \frac{b}{r} \\ \mathcal{E} &= \frac{b}{t} \\ \mathcal{C} &= \frac{r}{t}, \end{aligned} \tag{C.1}$$

where r is the number of events at the reconstructed, or detector, level, t is the number of events at the true, or generator, level, and b is the number of events *both* generated and reconstructed in the same bin. The quantities r, t, and b are correlated and must be written in such a way which uncorrelates them. This allows the application of error propagation formulas for independent and random quantities, which are less complex than those for correlated quantities. Defining

$$\begin{aligned} r &\equiv x + z \\ t &\equiv x + y \\ b &\equiv x, \end{aligned}$$

where x is the number of events generated and reconstructed in the same bin i, y is the number of events generated in bin i but not reconstructed in bin i, and z is the number of events not generated in bin i but reconstructed in bin i, one can write

$$\begin{aligned} \mathcal{P} &= \frac{x}{x+z} \\ \mathcal{E} &= \frac{x}{x+y} \\ \mathcal{C} &= \frac{x+y}{x+z}. \end{aligned} \tag{C.2}$$

The general for formula for the propagation of independent and random errors,

$$\delta q = \sqrt{\sum_{i=1}^{N} \left(\frac{\partial q}{\partial x_i} \delta x_i \right)^2}, \tag{C.3}$$

where $q = q(x_1, x_2, \ldots, x_N)$ may now be applied. It is assumed that the x_i follows a Poisson distribution. If the events are unweighted, as is the case in this thesis,

$$x_i = M_i \tag{C.4}$$

$$\delta x_i = \sqrt{M_i}, \tag{C.5}$$

---

[1] Note that these quantities are calculated separately for each bin i in an histogram but in this section are written without the subscript i.

where M_i is the number of events in bin i. The error on the purity is calculated by

$$\begin{aligned}\delta\mathcal{P} &= \left[\left(\frac{\partial\mathcal{P}}{\partial x}\delta x\right)^2 + \left(\frac{\partial\mathcal{P}}{\partial z}\delta z\right)^2\right]^{1/2} \\ &= \left[\left[\frac{1}{x+z} - \frac{x}{(x+z)^2}\right]^2 x + \left[\frac{x}{(x+z)^2}\right]^2 z\right]^{1/2} \\ &= \left[\frac{xz^2 + zx^2}{(x+z)^2}\right]^{1/2}.\end{aligned} \tag{C.6}$$

Replacing the original variables gives

$$\delta\mathcal{P} = \left[\frac{(1-\mathcal{P})\mathcal{P}}{r}\right]^{1/2}. \tag{C.7}$$

Similar calculations for $\mathcal{E}$ and $\mathcal{C}$ give

$$\delta\mathcal{E} = \left[\frac{(1-\mathcal{E})\mathcal{E}}{t}\right]^{1/2} \tag{C.8}$$

and

$$\delta\mathcal{C} = \left[\frac{t}{r^3}\left(t + r - 2b\right)\right]^{1/2}. \tag{C.9}$$

C.2 Error on the Gap Fraction

The gap fraction is defined as

$$f \equiv \frac{g}{i}, \tag{C.10}$$

where g is the number of events in the gap sample and i is the number of events in the inclusive sample. Substituting $i = g + n$, where n is the number of events which are *not* gap events,

$$f = \frac{g}{g+n} \tag{C.11}$$

Applying the formula for the propagation of uncorrelated errors (Equation C.3),

$$\begin{aligned}\delta f &= \sqrt{\left(\frac{\partial f}{\partial g}\delta g\right)^2 + \left(\frac{\partial f}{\partial n}\delta n\right)^2} \\ &= \left(\frac{g}{g+n}\right)\left(\frac{n}{g+n}\right)\sqrt{\left(\frac{\delta g}{g}\right)^2 + \left(\frac{\delta n}{n}\right)^2}.\end{aligned} \tag{C.12}$$

Substituting $f = g/(g+n)$ gives

$$\delta f = f(1-f)\sqrt{\left(\frac{\delta g}{g}\right)^2 + \left(\frac{\delta n}{n}\right)^2}. \tag{C.13}$$

Finally, replacing n with $i - g$ and δn with $\sqrt{\delta i^2 + \delta g^2}$ gives

$$\delta f = f(1-f)\left[\left(\frac{\delta g}{g}\right)^2 + \left(\frac{\sqrt{\delta i^2 + \delta g^2}}{i-g}\right)^2\right]^{1/2}. \tag{C.14}$$

Appendix D

Characteristics of the Monte Carlo Sample

D.1 Purities, Efficiencies, Stabilities, and Correction Factors

The purity, efficiency, stability, and correction factor are defined in Section 8.4. Each quantity is plotted for the cross section variables; $E_{\mathrm{T}}^{\mathrm{GAP}}$, $\Delta\eta$, $x_{\gamma}^{\mathrm{OBS}}$, and W. The direct, resolved, and color-singlet contributions are mixed according to the procedure described in Chapter 7. The solid points show the inclusive sample and the open squares show the gap sample defined by $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.0\,\mathrm{GeV}$(true) and $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.2\,\mathrm{GeV}$(reconstructed).

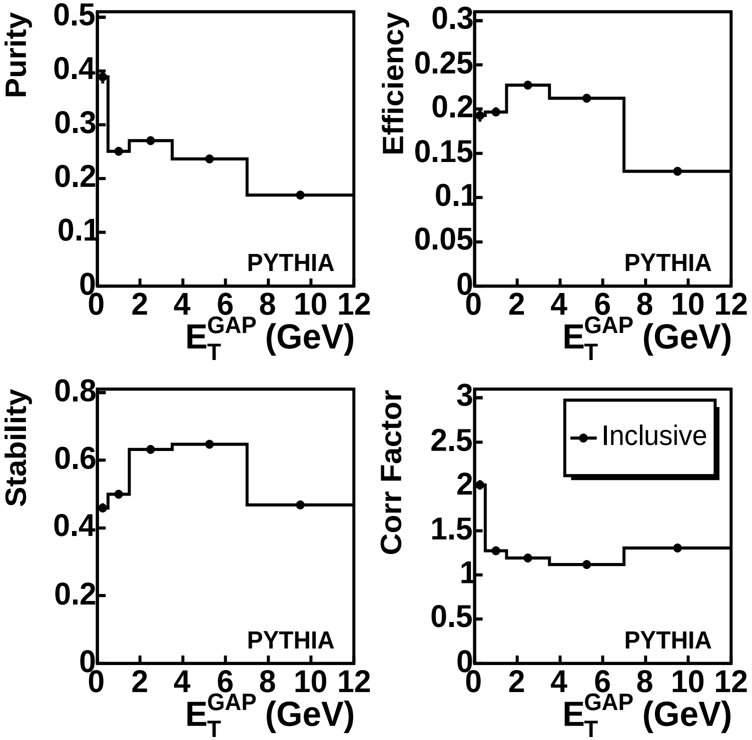


Figure D.1: The purity, efficiency, stability, and correction factor of $E_{\mathrm{T}}^{\mathrm{GAP}}$ calculated from Pythia. The solid points show the inclusive sample and the open squares show the gap sample defined by $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.0\,\mathrm{GeV}$(true) and $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.2\,\mathrm{GeV}$(reconstructed).

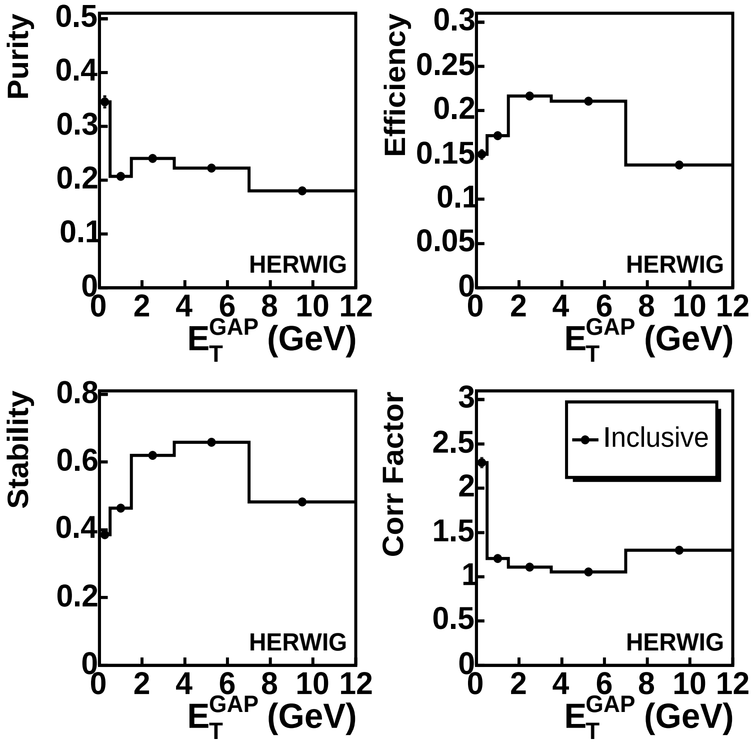


Figure D.2: The purity, efficiency, stability, and correction factor of E_T^{GAP} calculated from Herwig. The solid points show the inclusive sample and the open squares show the gap sample defined by $E_T^{CUT} = 1.0\,\text{GeV}$(true) and $E_T^{CUT} = 1.2\,\text{GeV}$(reconstructed).

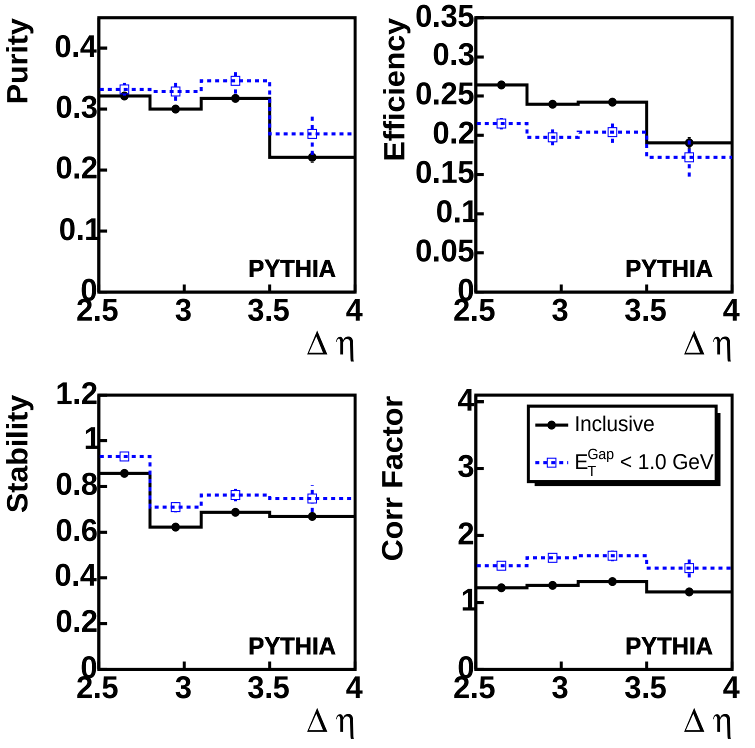


Figure D.3: The purity, efficiency, stability, and correction factor of $\Delta\eta$ calculated from Pythia. The solid points show the inclusive sample and the open squares show the gap sample defined by $E_T^{CUT} = 1.0\,\text{GeV}$(true) and $E_T^{CUT} = 1.2\,\text{GeV}$(reconstructed).

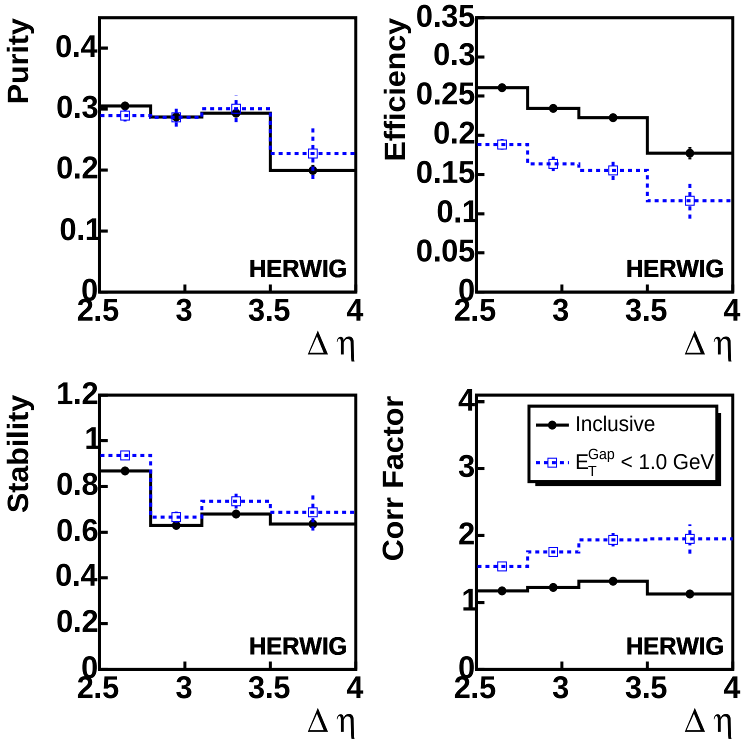


Figure D.4: The purity, efficiency, stability, and correction factor of $\Delta\eta$ calculated from Herwig. The solid points show the inclusive sample and the open squares show the gap sample defined by $E_T^{\rm CUT} = 1.0\,{\rm GeV}$(true) and $E_T^{\rm CUT} = 1.2\,{\rm GeV}$(reconstructed).

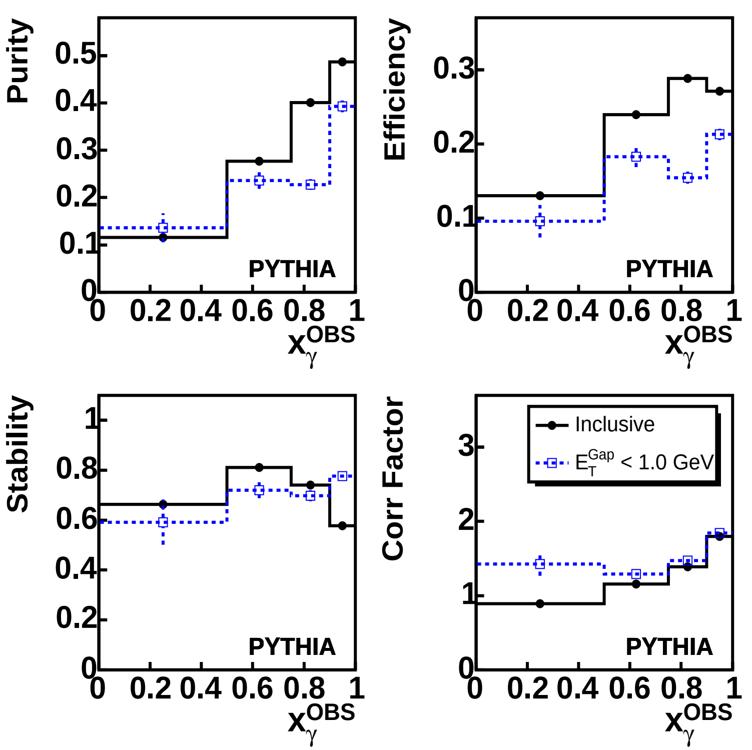


Figure D.5: The purity, efficiency, stability, and correction factor of $x_\gamma^{\rm OBS}$ calculated from Pythia. The solid points show the inclusive sample and the open squares show the gap sample defined by $E_T^{\rm CUT} = 1.0\,{\rm GeV}$(true) and $E_T^{\rm CUT} = 1.2\,{\rm GeV}$(reconstructed).

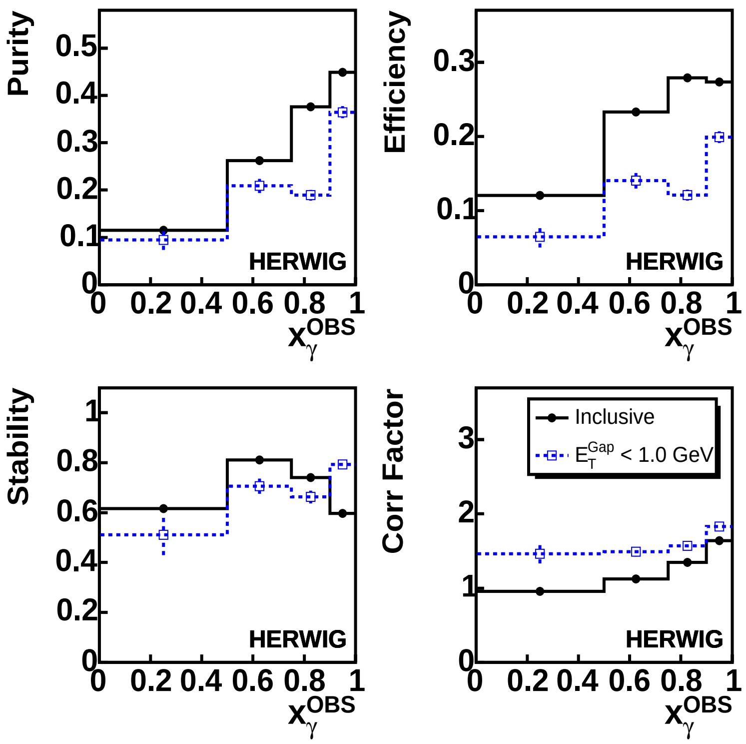


Figure D.6: The purity, efficiency, stability, and correction factor of $x_\gamma^{\rm OBS}$ calculated from Herwig. The solid points show the inclusive sample and the open squares show the gap sample defined by $E_{\rm T}^{\rm CUT} = 1.0\,{\rm GeV}$(true) and $E_{\rm T}^{\rm CUT} = 1.2\,{\rm GeV}$(reconstructed).

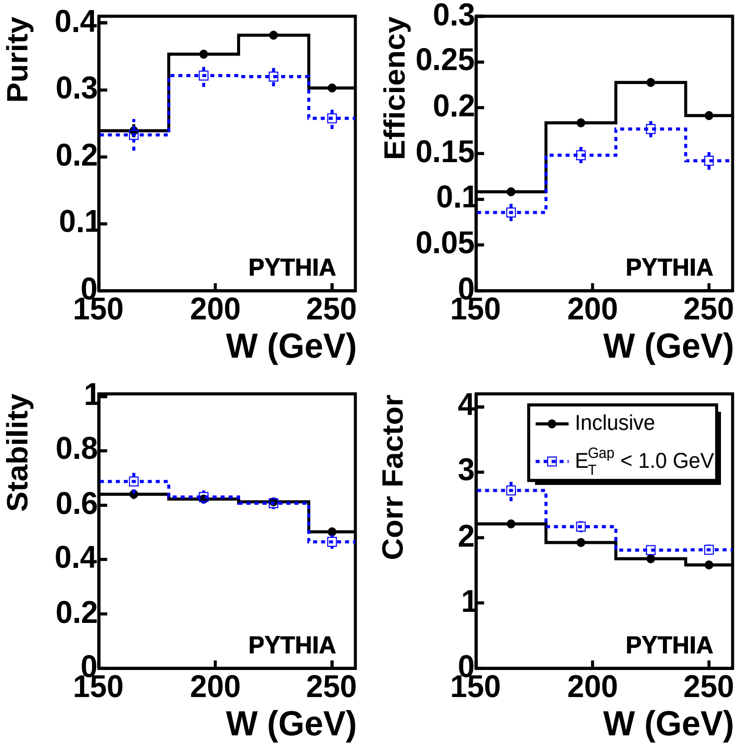


Figure D.7: The purity, efficiency, stability, and correction factor of W calculated from Pythia. The solid points show the inclusive sample and the open squares show the gap sample defined by $E_{\rm T}^{\rm CUT} = 1.0\,{\rm GeV}$(true) and $E_{\rm T}^{\rm CUT} = 1.2\,{\rm GeV}$(reconstructed).

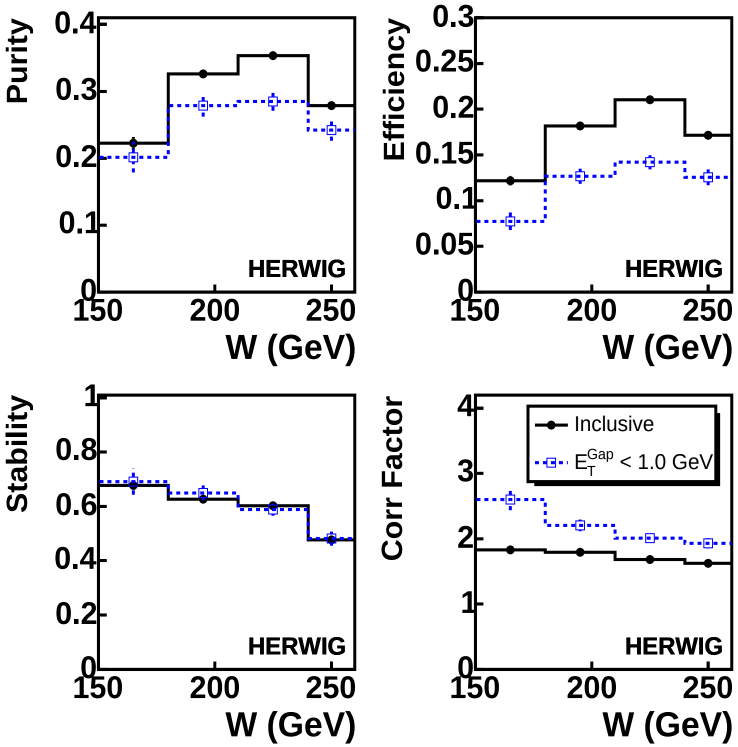


Figure D.8: The purity, efficiency, stability, and correction factor of W calculated from Herwig. The solid points show the inclusive sample and the open squares show the gap sample defined by $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.0\,\mathrm{GeV}$(true) and $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.2\,\mathrm{GeV}$(reconstructed).

D.2 Cumulative Efficiencies

The cumulative efficiencies are defined in Section 8.5. Each quantity is plotted for the cross section variables; $E_{\mathrm{T}}^{\mathrm{GAP}}$, $\Delta\eta$, $x_{\gamma}^{\mathrm{OBS}}$, and W. The direct, resolved, and color-singlet contributions are mixed according to the procedure described in Chapter 7. The solid points show the inclusive sample and the open squares show the gap sample defined by $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.0\,\mathrm{GeV}$(true) and $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.2\,\mathrm{GeV}$(reconstructed).

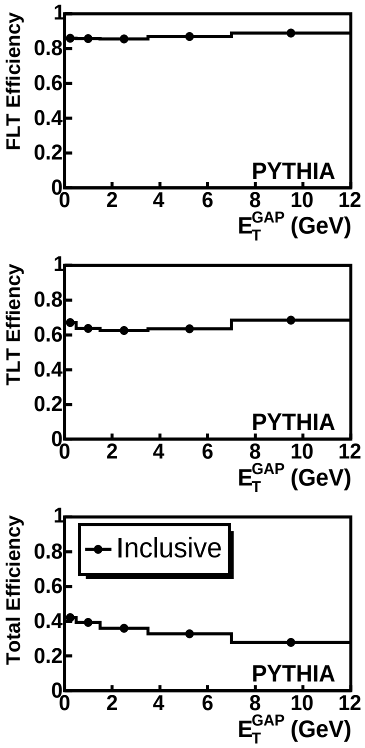


Figure D.9: The cumulative efficiencies at each stage of event selection for $E_{\mathrm{T}}^{\mathrm{GAP}}$ calculated from Pythia. The solid points show the inclusive sample and the open squares show the gap sample defined by $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.0\,\mathrm{GeV}$(true) and $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.2\,\mathrm{GeV}$(reconstructed).

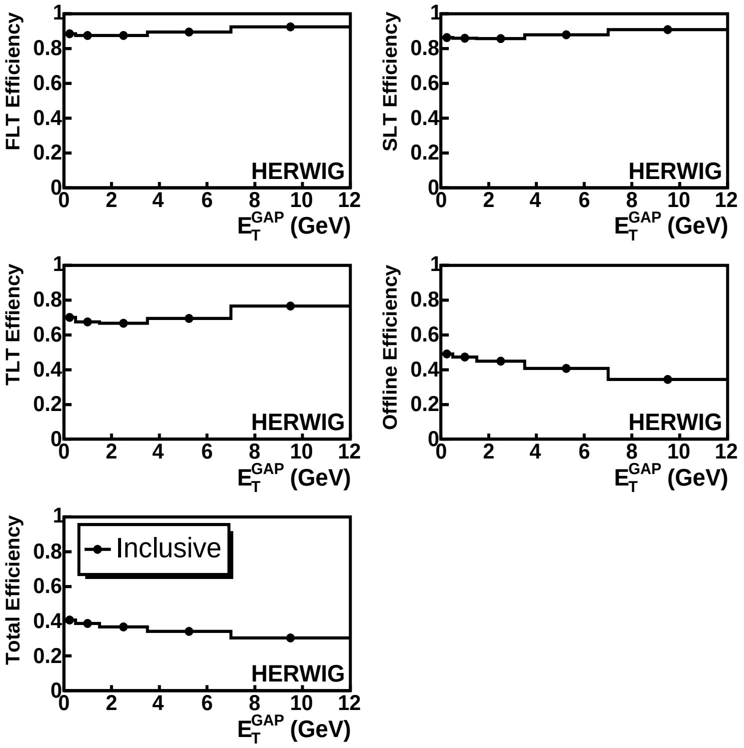


Figure D.10: The cumulative efficiencies at each stage of event selection for $E_{\mathrm{T}}^{\mathrm{GAP}}$ calculated from Herwig. The solid points show the inclusive sample and the open squares show the gap sample defined by $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.0\,\mathrm{GeV}$(true) and $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.2\,\mathrm{GeV}$(reconstructed).

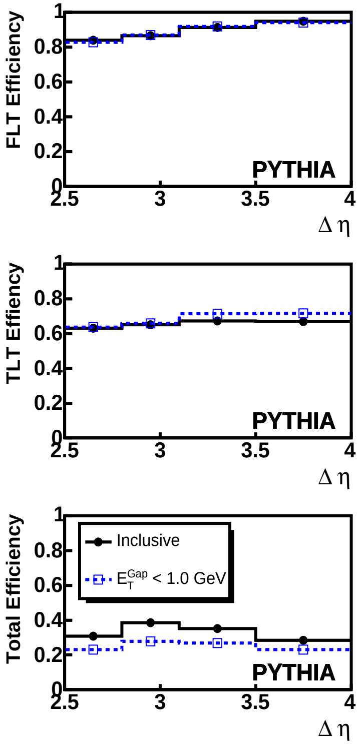


Figure D.11: The cumulative efficiencies at each stage of event selection for $\Delta\eta$ calculated from Pythia. The solid points show the inclusive sample and the open squares show the gap sample defined by $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.0\,\mathrm{GeV}$(true) and $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.2\,\mathrm{GeV}$(reconstructed).

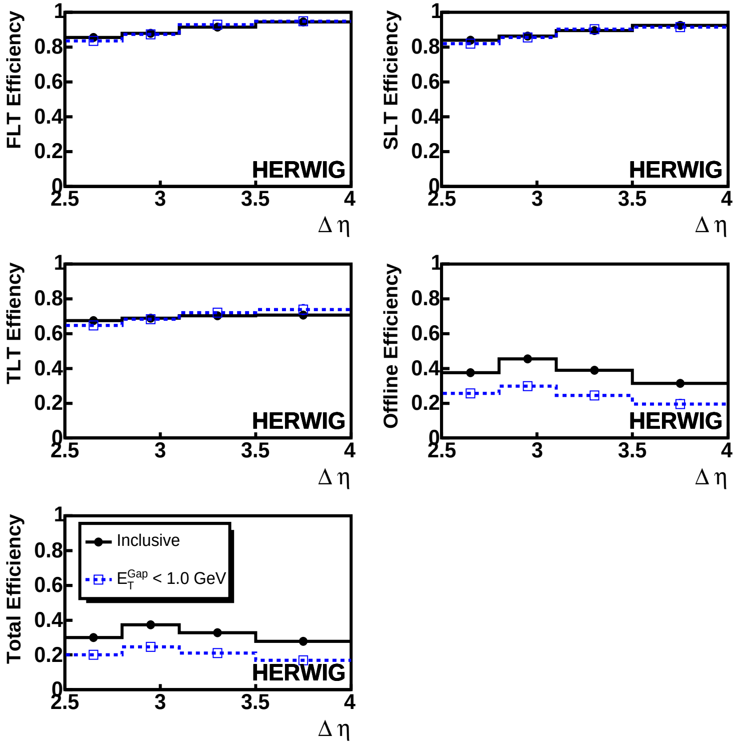


Figure D.12: The cumulative efficiencies at each stage of event selection for $\Delta\eta$ calculated from Herwig. The solid points show the inclusive sample and the open squares show the gap sample defined by $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.0\,\mathrm{GeV}$(true) and $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.2\,\mathrm{GeV}$(reconstructed).

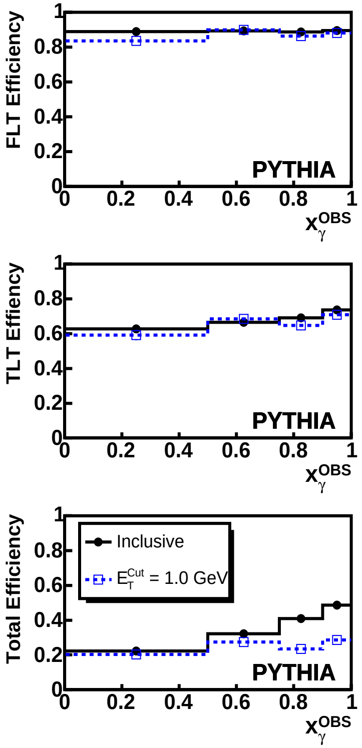


Figure D.13: The cumulative efficiencies at each stage of event selection for x_γ^{OBS} calculated from Pythia. The solid points show the inclusive sample and the open squares show the gap sample defined by $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.0\,\mathrm{GeV}(\mathrm{true})$ and $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.2\,\mathrm{GeV}(\mathrm{reconstructed})$.

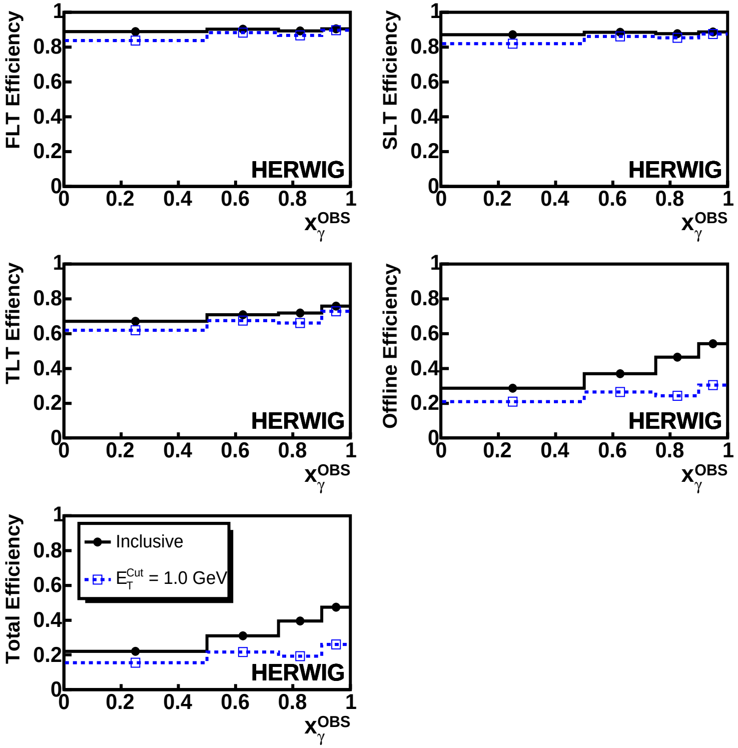


Figure D.14: The cumulative efficiencies at each stage of event selection for x_γ^{OBS} calculated from Herwig. The solid points show the inclusive sample and the open squares show the gap sample defined by $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.0\,\mathrm{GeV}(\mathrm{true})$ and $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.2\,\mathrm{GeV}(\mathrm{reconstructed})$.

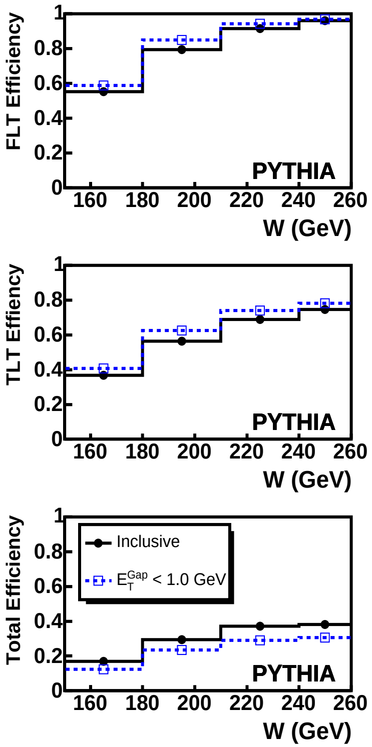


Figure D.15: The cumulative efficiencies at each stage of event selection for W calculated from Pythia. The solid points show the inclusive sample and the open squares show the gap sample defined by $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.0\,\mathrm{GeV}$(true) and $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.2\,\mathrm{GeV}$(reconstructed).

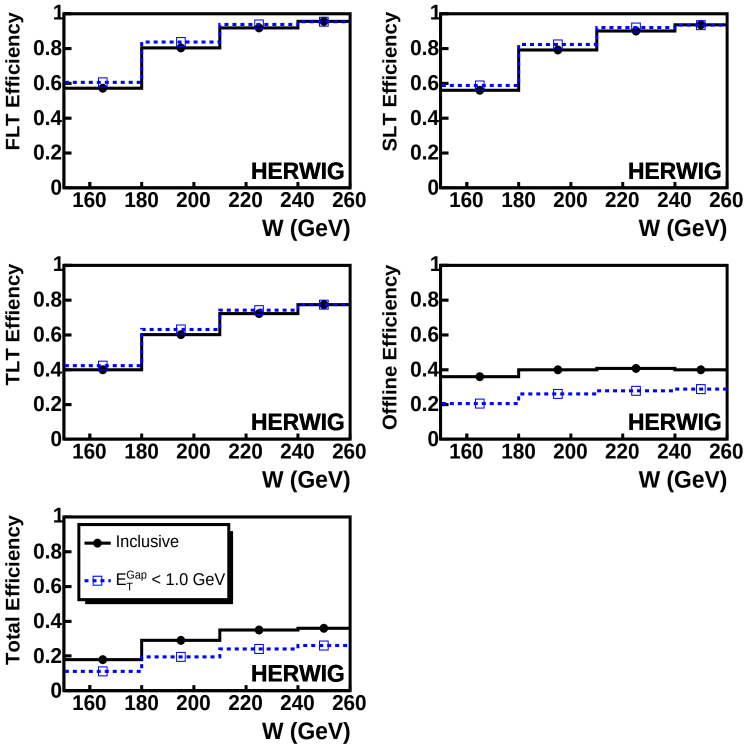


Figure D.16: The cumulative efficiencies at each stage of event selection for W calculated from Herwig. The solid points show the inclusive sample and the open squares show the gap sample defined by $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.0\,\mathrm{GeV}$(true) and $E_{\mathrm{T}}^{\mathrm{CUT}} = 1.2\,\mathrm{GeV}$(reconstructed).

Appendix E

Tables of Cross Sections and Gap Fractions

$E_{\mathrm{T}}^{\mathrm{GAP}}$ bin (GeV)	σ(nb)	± stat	± sys	± cal
0.0 − 0.5	0.167	± 0.004	$^{+0.014}_{-0.014}$	$^{+0.002}_{-0.006}$
0.5 − 1.5	0.153	± 0.002	$^{+0.006}_{-0.006}$	$^{+0.000}_{-0.001}$
1.5 − 3.5	0.210	± 0.002	$^{+0.009}_{-0.008}$	$^{+0.001}_{-0.002}$
3.5 − 7.0	0.177	± 0.001	$^{+0.006}_{-0.005}$	$^{+0.006}_{-0.008}$
7.0 − 12.0	0.080	± 0.001	$^{+0.002}_{-0.002}$	$^{+0.007}_{-0.008}$

Table E.1: The measured differential cross section $d\sigma/dE_{\mathrm{T}}^{\mathrm{GAP}}$ unfolded with the average correction factors of Pythia and Herwig for the inclusive sample of events. The statistical error, systematic errors, and calorimeter energy scale uncertainty on the measurement are also listed.

$\Delta\eta$ bin	σ(nb)	± stat	± sys	± cal
2.5 − 2.8	2.652	± 0.019	$^{+0.065}_{-0.062}$	$^{+0.085}_{-0.122}$
2.8 − 3.1	1.712	± 0.016	$^{+0.033}_{-0.027}$	$^{+0.061}_{-0.083}$
3.1 − 3.5	0.945	± 0.011	$^{+0.015}_{-0.015}$	$^{+0.039}_{-0.043}$
3.5 − 4.0	0.193	± 0.004	$^{+0.009}_{-0.007}$	$^{+0.007}_{-0.009}$

Table E.2: The measured differential cross section $d\sigma/d\Delta\eta$ unfolded with the average correction factors of Pythia and Herwig for the inclusive sample of events. The statistical error, systematic errors, and calorimeter energy scale uncertainty on the measurement are also listed.

$\Delta\eta$ bin	$E_{\mathrm{T}}^{\mathrm{CUT}}$ GeV	σ(nb)	± stat	± sys	± cal
2.5, 2.8	0.5	0.140	± 0.005	$^{+0.018}_{-0.008}$	$^{+0.002}_{-0.004}$
2.8, 3.1		0.080	± 0.004	$^{+0.009}_{-0.010}$	$^{+0.002}_{-0.003}$
3.1, 3.5		0.038	± 0.003	$^{+0.007}_{-0.008}$	$^{+0.000}_{-0.003}$
3.5, 4.0		0.007	± 0.001	$^{+0.002}_{-0.002}$	$^{+0.000}_{-0.000}$
2.5, 2.8	1.0	0.269	± 0.006	$^{+0.015}_{-0.013}$	$^{+0.000}_{-0.006}$
2.8, 3.1		0.137	± 0.005	$^{+0.011}_{-0.006}$	$^{+0.001}_{-0.001}$
3.1, 3.5		0.057	± 0.003	$^{+0.005}_{-0.006}$	$^{+0.000}_{-0.002}$
3.5, 4.0		0.011	± 0.001	$^{+0.003}_{-0.003}$	$^{+0.000}_{-0.000}$
2.5, 2.8	1.5	0.431	± 0.008	$^{+0.019}_{-0.022}$	$^{+0.000}_{-0.006}$
2.8, 3.1		0.217	± 0.006	$^{+0.006}_{-0.007}$	$^{+0.000}_{-0.004}$
3.1, 3.5		0.089	± 0.003	$^{+0.007}_{-0.004}$	$^{+0.000}_{-0.001}$
3.5, 4.0		0.018	± 0.001	$^{+0.004}_{-0.006}$	$^{+0.000}_{-0.000}$
2.5, 2.8	2.0	0.606	± 0.009	$^{+0.032}_{-0.027}$	$^{+0.001}_{-0.010}$
2.8, 3.1		0.305	± 0.007	$^{+0.019}_{-0.009}$	$^{+0.001}_{-0.003}$
3.1, 3.5		0.128	± 0.004	$^{+0.013}_{-0.009}$	$^{+0.000}_{-0.001}$
3.5, 4.0		0.027	± 0.002	$^{+0.004}_{-0.007}$	$^{+0.000}_{-0.001}$

Table E.3: The measured differential cross section $d\sigma/d\Delta\eta$ unfolded with the average correction factors of Pythia and Herwig for the gap sample of events. The statistical error, systematic errors, and calorimeter energy scale uncertainty on the measurement are also listed.

$\Delta\eta$ bin	$E_{\mathrm{T}}^{\mathrm{CUT}}$ GeV	f	± stat	± sys	± cal
$2.5, 2.8$	0.5	0.053	± 0.002	$^{+0.007}_{-0.004}$	$^{+0.003}_{-0.003}$
$2.8, 3.1$		0.047	± 0.002	$^{+0.006}_{-0.007}$	$^{+0.004}_{-0.003}$
$3.1, 3.5$		0.040	± 0.003	$^{+0.008}_{-0.009}$	$^{+0.002}_{-0.005}$
$3.5, 4.0$		0.038	± 0.005	$^{+0.012}_{-0.012}$	$^{+0.001}_{-0.000}$
$2.5, 2.8$	1.0	0.101	± 0.002	$^{+0.006}_{-0.005}$	$^{+0.004}_{-0.005}$
$2.8, 3.1$		0.080	± 0.003	$^{+0.007}_{-0.005}$	$^{+0.005}_{-0.004}$
$3.1, 3.5$		0.061	± 0.003	$^{+0.006}_{-0.006}$	$^{+0.001}_{-0.004}$
$3.5, 4.0$		0.055	± 0.005	$^{+0.014}_{-0.016}$	$^{+0.003}_{-0.002}$
$2.5, 2.8$	1.5	0.163	± 0.003	$^{+0.007}_{-0.009}$	$^{+0.008}_{-0.007}$
$2.8, 3.1$		0.127	± 0.003	$^{+0.005}_{-0.005}$	$^{+0.007}_{-0.007}$
$3.1, 3.5$		0.094	± 0.003	$^{+0.007}_{-0.005}$	$^{+0.003}_{-0.005}$
$3.5, 4.0$		0.092	± 0.007	$^{+0.019}_{-0.030}$	$^{+0.003}_{-0.004}$
$2.5, 2.8$	2.0	0.228	± 0.003	$^{+0.011}_{-0.010}$	$^{+0.012}_{-0.011}$
$2.8, 3.1$		0.178	± 0.004	$^{+0.012}_{-0.006}$	$^{+0.010}_{-0.008}$
$3.1, 3.5$		0.135	± 0.004	$^{+0.014}_{-0.010}$	$^{+0.006}_{-0.006}$
$3.5, 4.0$		0.138	± 0.008	$^{+0.019}_{-0.035}$	$^{+0.001}_{-0.009}$

Table E.4: The measured gap fraction $f(\Delta\eta)$ unfolded with the average correction factors of Pythia and Herwig. The statistical error, systematic errors, and calorimeter energy scale uncertainty on the measurement are also listed.

$x_{\gamma}^{\mathrm{OBS}}$ bin	σ(nb)	± stat	± sys	± cal
$0.00 - 0.50$	0.386	± 0.005	$^{+0.024}_{-0.021}$	$^{+0.017}_{-0.027}$
$0.50 - 0.75$	3.040	± 0.023	$^{+0.075}_{-0.069}$	$^{+0.127}_{-0.162}$
$0.75 - 0.90$	4.504	± 0.038	$^{+0.123}_{-0.098}$	$^{+0.128}_{-0.152}$
$0.90 - 1.00$	1.513	± 0.026	$^{+0.077}_{-0.084}$	$^{+0.010}_{-0.028}$

Table E.5: The measured differential cross section $d\sigma/dx_{\gamma}^{\mathrm{OBS}}$ unfolded with the average correction factors of Pythia and Herwig for the inclusive sample of events. The statistical error, systematic errors, and calorimeter energy scale uncertainty on the measurement are also listed.

$x_{\gamma}^{\mathrm{OBS}}$ bin	$E_{\mathrm{T}}^{\mathrm{CUT}}$ GeV	σ(nb)	± stat	± sys	± cal
$0.00, 0.50$	0.5	0.007	± 0.001	$^{+0.002}_{-0.001}$	$^{+0.000}_{-0.000}$
$0.50, 0.75$		0.053	± 0.003	$^{+0.011}_{-0.011}$	$^{+0.000}_{-0.001}$
$0.75, 0.90$		0.174	± 0.008	$^{+0.026}_{-0.020}$	$^{+0.004}_{-0.008}$
$0.90, 1.00$		0.410	± 0.016	$^{+0.043}_{-0.037}$	$^{+0.007}_{-0.015}$
$0.00, 0.50$	1.0	0.011	± 0.001	$^{+0.002}_{-0.001}$	$^{+0.000}_{-0.001}$
$0.50, 0.75$		0.089	± 0.004	$^{+0.013}_{-0.011}$	$^{+0.000}_{-0.001}$
$0.75, 0.90$		0.356	± 0.011	$^{+0.020}_{-0.021}$	$^{+0.000}_{-0.011}$
$0.90, 1.00$		0.685	± 0.019	$^{+0.028}_{-0.038}$	$^{+0.000}_{-0.007}$
$0.00, 0.50$	1.5	0.018	± 0.001	$^{+0.002}_{-0.003}$	$^{+0.000}_{-0.001}$
$0.50, 0.75$		0.140	± 0.005	$^{+0.017}_{-0.016}$	$^{+0.000}_{-0.002}$
$0.75, 0.90$		0.653	± 0.014	$^{+0.029}_{-0.043}$	$^{+0.003}_{-0.016}$
$0.90, 1.00$		0.953	± 0.021	$^{+0.045}_{-0.043}$	$^{+0.000}_{-0.004}$
$0.00, 0.50$	2.0	0.027	± 0.001	$^{+0.003}_{-0.004}$	$^{+0.000}_{-0.002}$
$0.50, 0.75$		0.214	± 0.006	$^{+0.024}_{-0.016}$	$^{+0.000}_{-0.007}$
$0.75, 0.90$		1.021	± 0.017	$^{+0.069}_{-0.056}$	$^{+0.009}_{-0.013}$
$0.90, 1.00$		1.153	± 0.023	$^{+0.051}_{-0.055}$	$^{+0.003}_{-0.009}$

Table E.6: The measured differential cross section $d\sigma/dx_{\gamma}^{\mathrm{OBS}}$ unfolded with the average correction factors of Pythia and Herwig for the gap sample of events. The statistical error, systematic errors, and calorimeter energy scale uncertainty on the measurement are also listed.

$x_\gamma^{\rm OBS}$ bin	$E_{\rm T}^{\rm CUT}$ GeV	f	± stat	± sys	± cal
0.00, 0.50	0.5	0.017	± 0.002	$^{+0.004}_{-0.002}$	$^{+0.000}_{-0.001}$
0.50, 0.75		0.018	± 0.001	$^{+0.004}_{-0.003}$	$^{+0.001}_{-0.001}$
0.75, 0.90		0.039	± 0.002	$^{+0.006}_{-0.005}$	$^{+0.002}_{-0.003}$
0.90, 1.00		0.272	± 0.010	$^{+0.033}_{-0.028}$	$^{+0.011}_{-0.012}$
0.00, 0.50	1.0	0.028	± 0.003	$^{+0.004}_{-0.003}$	$^{+0.000}_{-0.001}$
0.50, 0.75		0.029	± 0.001	$^{+0.004}_{-0.003}$	$^{+0.001}_{-0.002}$
0.75, 0.90		0.079	± 0.002	$^{+0.005}_{-0.005}$	$^{+0.003}_{-0.005}$
0.90, 1.00		0.454	± 0.012	$^{+0.024}_{-0.026}$	$^{+0.008}_{-0.008}$
0.00, 0.50	1.5	0.047	± 0.003	$^{+0.005}_{-0.007}$	$^{+0.001}_{-0.002}$
0.50, 0.75		0.046	± 0.001	$^{+0.006}_{-0.005}$	$^{+0.003}_{-0.003}$
0.75, 0.90		0.145	± 0.003	$^{+0.007}_{-0.010}$	$^{+0.006}_{-0.008}$
0.90, 1.00		0.630	± 0.015	$^{+0.028}_{-0.022}$	$^{+0.010}_{-0.007}$
0.00, 0.50	2.0	0.069	± 0.004	$^{+0.007}_{-0.010}$	$^{+0.001}_{-0.005}$
0.50, 0.75		0.070	± 0.002	$^{+0.008}_{-0.005}$	$^{+0.004}_{-0.005}$
0.75, 0.90		0.227	± 0.004	$^{+0.016}_{-0.013}$	$^{+0.010}_{-0.009}$
0.90, 1.00		0.763	± 0.018	$^{+0.023}_{-0.021}$	$^{+0.009}_{-0.003}$

Table E.7: The measured gap fraction $\mathrm{f}(x_\gamma^{\rm OBS})$ unfolded with the average correction factors of Pythia and Herwig. The statistical error, systematic errors, and calorimeter energy scale uncertainty on the measurement are also listed.

W bin (GeV)	σ(pb)	± stat	± sys	± cal
150.0 − 180.0	5.12	± 0.11	$^{+0.51}_{-0.53}$	$^{+0.16}_{-0.12}$
180.0 − 210.0	14.5	± 0.2	$^{+0.7}_{-0.6}$	$^{+0.0}_{-0.2}$
210.0 − 240.0	22.7	± 0.2	$^{+0.5}_{-0.3}$	$^{+0.8}_{-0.7}$
240.0 − 260.0	27.1	± 0.3	$^{+0.5}_{-0.5}$	$^{+1.4}_{-2.0}$

Table E.8: The measured differential cross section $d\sigma/dW$ unfolded with the average correction factors of Pythia and Herwig for the inclusive sample of events. The statistical error, systematic errors, and calorimeter energy scale uncertainty on the measurement are also listed.

W bin (GeV)	$E_{\rm T}^{\rm CUT}$ GeV	σ(pb)	± stat	± sys	± cal
150.0, 180.0	0.5	0.387	± 0.035	$^{+0.053}_{-0.064}$	$^{+0.017}_{-0.054}$
180.0, 210.0		0.717	± 0.043	$^{+0.109}_{-0.074}$	$^{+0.021}_{-0.018}$
210.0, 240.0		0.887	± 0.046	$^{+0.154}_{-0.112}$	$^{+0.017}_{-0.020}$
240.0, 260.0		1.027	± 0.062	$^{+0.147}_{-0.117}$	$^{+0.011}_{-0.000}$
150.0, 180.0	1.0	0.738	± 0.044	$^{+0.058}_{-0.071}$	$^{+0.038}_{-0.081}$
180.0, 210.0		1.391	± 0.054	$^{+0.082}_{-0.101}$	$^{+0.038}_{-0.030}$
210.0, 240.0		1.559	± 0.052	$^{+0.164}_{-0.101}$	$^{+0.003}_{-0.017}$
240.0, 260.0		1.686	± 0.067	$^{+0.170}_{-0.127}$	$^{+0.015}_{-0.010}$
150.0, 180.0	1.5	1.230	± 0.055	$^{+0.118}_{-0.113}$	$^{+0.053}_{-0.097}$
180.0, 210.0		2.219	± 0.065	$^{+0.122}_{-0.111}$	$^{+0.077}_{-0.065}$
210.0, 240.0		2.553	± 0.065	$^{+0.211}_{-0.170}$	$^{+0.037}_{-0.047}$
240.0, 260.0		2.622	± 0.080	$^{+0.184}_{-0.183}$	$^{+0.009}_{-0.108}$
150.0, 180.0	2.0	1.723	± 0.064	$^{+0.159}_{-0.191}$	$^{+0.098}_{-0.109}$
180.0, 210.0		3.166	± 0.077	$^{+0.216}_{-0.250}$	$^{+0.082}_{-0.077}$
210.0, 240.0		3.691	± 0.077	$^{+0.301}_{-0.253}$	$^{+0.033}_{-0.036}$
240.0, 260.0		3.754	± 0.093	$^{+0.306}_{-0.152}$	$^{+0.000}_{-0.118}$

Table E.9: The measured differential cross section $d\sigma/dW$ unfolded with the average correction factors of Pythia and Herwig for the gap sample of events. The statistical error, systematic errors, and calorimeter energy scale uncertainty on the measurement are also listed.

W bin (GeV)	$E_T^{\rm CUT}$ GeV	f	± stat	± sys	± cal
150.0, 180.0	0.5	0.077	± 0.007	$^{+0.017}_{-0.017}$	$^{+0.001}_{-0.010}$
180.0, 210.0		0.049	± 0.003	$^{+0.008}_{-0.005}$	$^{+0.003}_{-0.001}$
210.0, 240.0		0.039	± 0.002	$^{+0.006}_{-0.005}$	$^{+0.002}_{-0.002}$
240.0, 260.0		0.038	± 0.002	$^{+0.005}_{-0.004}$	$^{+0.005}_{-0.003}$
150.0, 180.0	1.0	0.145	± 0.008	$^{+0.016}_{-0.019}$	$^{+0.003}_{-0.013}$
180.0, 210.0		0.096	± 0.004	$^{+0.005}_{-0.008}$	$^{+0.005}_{-0.002}$
210.0, 240.0		0.069	± 0.002	$^{+0.007}_{-0.004}$	$^{+0.001}_{-0.002}$
240.0, 260.0		0.062	± 0.002	$^{+0.006}_{-0.004}$	$^{+0.005}_{-0.003}$
150.0, 180.0	1.5	0.241	± 0.010	$^{+0.025}_{-0.019}$	$^{+0.003}_{-0.015}$
180.0, 210.0		0.153	± 0.004	$^{+0.010}_{-0.010}$	$^{+0.008}_{-0.004}$
210.0, 240.0		0.113	± 0.003	$^{+0.008}_{-0.008}$	$^{+0.005}_{-0.006}$
240.0, 260.0		0.097	± 0.003	$^{+0.006}_{-0.006}$	$^{+0.003}_{-0.004}$
150.0, 180.0	2.0	0.339	± 0.012	$^{+0.029}_{-0.038}$	$^{+0.008}_{-0.016}$
180.0, 210.0		0.218	± 0.005	$^{+0.016}_{-0.019}$	$^{+0.010}_{-0.004}$
210.0, 240.0		0.163	± 0.003	$^{+0.012}_{-0.011}$	$^{+0.007}_{-0.007}$
240.0, 260.0		0.138	± 0.003	$^{+0.011}_{-0.004}$	$^{+0.006}_{-0.008}$

Table E.10: The measured gap fraction $f\,(W)$ unfolded with the average correction factors of Pythia and Herwig. The statistical error, systematic errors, and calorimeter energy scale uncertainty on the measurement are also listed.

$\Delta\eta$ bin	σ(pb)	± stat	± sys	± cal
$2.5 - 2.8$	1463	± 15	$^{+37}_{-31}$	$^{+61}_{-85}$
$2.8 - 3.1$	925.2	± 12.4	$^{+17.2}_{-14.0}$	$^{+41.6}_{-56.4}$
$3.1 - 3.5$	491.6	± 8.1	$^{+16.3}_{-13.5}$	$^{+24.3}_{-25.6}$
$3.5 - 4.0$	72.98	± 2.59	$^{+3.91}_{-3.49}$	$^{+0.55}_{-3.38}$

Table E.11: The measured differential cross section $d\sigma/d\Delta\eta$ in the region $x_\gamma^{\rm OBS} < 0.75$ unfolded with the average correction factors of Pythia and Herwig for the inclusive sample of events. The statistical error, systematic errors, and calorimeter energy scale uncertainty on the measurement are also listed.

$\Delta\eta$ bin	$E_T^{\rm CUT}$ GeV	σ(pb)	± stat	± sys	± cal
2.5, 2.8	0.5	31.4	± 2.5	$^{+4.9}_{-3.8}$	$^{+0.0}_{-0.0}$
2.8, 3.1		12.9	± 1.6	$^{+4.5}_{-2.6}$	$^{+0.0}_{-0.4}$
3.1, 3.5		7.34	± 1.15	$^{+2.09}_{-2.26}$	$^{+0.00}_{-0.91}$
3.5, 4.0		0.67	± 0.25	$^{+0.78}_{-0.54}$	$^{+0.09}_{-0.00}$
2.5, 2.8	1.0	55.1	± 3.0	$^{+6.1}_{-5.4}$	$^{+0.3}_{-1.4}$
2.8, 3.1		22.2	± 1.8	$^{+4.3}_{-1.4}$	$^{+0.4}_{-0.0}$
3.1, 3.5		9.30	± 1.03	$^{+2.22}_{-1.35}$	$^{+0.00}_{-0.93}$
3.5, 4.0		1.14	± 0.30	$^{+0.37}_{-0.59}$	$^{+0.00}_{-0.16}$
2.5, 2.8	1.5	88.1	± 3.6	$^{+8.2}_{-8.4}$	$^{+0.5}_{-1.9}$
2.8, 3.1		37.0	± 2.3	$^{+4.2}_{-4.1}$	$^{+0.2}_{-1.5}$
3.1, 3.5		13.4	± 1.2	$^{+2.8}_{-1.6}$	$^{+0.0}_{-0.8}$
3.5, 4.0		1.90	± 0.41	$^{+0.65}_{-1.07}$	$^{+0.00}_{-0.06}$
2.5, 2.8	2.0	131.5	± 4.3	$^{+12.5}_{-8.2}$	$^{+0.0}_{-4.7}$
2.8, 3.1		58.0	± 2.9	$^{+6.7}_{-6.0}$	$^{+0.0}_{-1.2}$
3.1, 3.5		21.5	± 1.5	$^{+2.8}_{-2.5}$	$^{+0.0}_{-1.3}$
3.5, 4.0		2.62	± 0.45	$^{+0.72}_{-0.80}$	$^{+0.14}_{-0.14}$

Table E.12: The measured differential cross section $d\sigma/d\Delta\eta$ in the region $x_\gamma^{\rm OBS} < 0.75$ unfolded with the average correction factors of Pythia and Herwig for the gap sample of events. The statistical error, systematic errors, and calorimeter energy scale uncertainty on the measurement are also listed.

$\Delta\eta$ bin	$E_{\mathrm{T}}^{\mathrm{CUT}}$ GeV	f	± stat	± sys	± cal
$2.5, 2.8$	0.5	0.021	± 0.002	$^{+0.003}_{-0.003}$	$^{+0.001}_{-0.001}$
$2.8, 3.1$		0.014	± 0.002	$^{+0.005}_{-0.004}$	$^{+0.001}_{-0.001}$
$3.1, 3.5$		0.015	± 0.002	$^{+0.004}_{-0.005}$	$^{+0.000}_{-0.003}$
$3.5, 4.0$		0.009	± 0.003	$^{+0.011}_{-0.007}$	$^{+0.002}_{-0.000}$
$2.5, 2.8$	1.0	0.038	± 0.002	$^{+0.004}_{-0.004}$	$^{+0.001}_{-0.001}$
$2.8, 3.1$		0.024	± 0.002	$^{+0.005}_{-0.003}$	$^{+0.002}_{-0.001}$
$3.1, 3.5$		0.019	± 0.002	$^{+0.005}_{-0.003}$	$^{+0.000}_{-0.003}$
$3.5, 4.0$		0.016	± 0.004	$^{+0.005}_{-0.008}$	$^{+0.000}_{-0.002}$
$2.5, 2.8$	1.5	0.060	± 0.002	$^{+0.006}_{-0.006}$	$^{+0.003}_{-0.002}$
$2.8, 3.1$		0.040	± 0.002	$^{+0.005}_{-0.005}$	$^{+0.003}_{-0.003}$
$3.1, 3.5$		0.027	± 0.002	$^{+0.006}_{-0.003}$	$^{+0.001}_{-0.003}$
$3.5, 4.0$		0.026	± 0.006	$^{+0.009}_{-0.015}$	$^{+0.001}_{-0.001}$
$2.5, 2.8$	2.0	0.090	± 0.003	$^{+0.009}_{-0.006}$	$^{+0.005}_{-0.007}$
$2.8, 3.1$		0.063	± 0.003	$^{+0.008}_{-0.007}$	$^{+0.003}_{-0.004}$
$3.1, 3.5$		0.044	± 0.003	$^{+0.006}_{-0.005}$	$^{+0.001}_{-0.005}$
$3.5, 4.0$		0.036	± 0.006	$^{+0.010}_{-0.011}$	$^{+0.002}_{-0.000}$

Table E.13: The measured gap fraction $f\,(\Delta\eta)$ in the region $x_{\gamma}^{\mathrm{OBS}} < 0.75$ unfolded with the average correction factors of Pythia and Herwig. The statistical error, systematic errors, and calorimeter energy scale uncertainty on the measurement are also listed.

W bin (GeV)	σ(pb)	± stat	± sys	± cal
$150.0 - 180.0$	1.37	± 0.06	$^{+0.19}_{-0.20}$	$^{+0.02}_{-0.02}$
$180.0 - 210.0$	6.72	± 0.13	$^{+0.48}_{-0.42}$	$^{+0.03}_{-0.14}$
$210.0 - 240.0$	13.05	± 0.16	$^{+0.29}_{-0.19}$	$^{+0.42}_{-0.47}$
$240.0 - 260.0$	17.55	± 0.22	$^{+0.58}_{-0.59}$	$^{+0.83}_{-1.34}$

Table E.14: The measured differential cross section $d\sigma/dW$ in the region $x_{\gamma}^{\mathrm{OBS}} < 0.75$ unfolded with the average correction factors of Pythia and Herwig for the inclusive sample of events. The statistical error, systematic errors, and calorimeter energy scale uncertainty on the measurement are also listed.

W bin (GeV)	$E_{\mathrm{T}}^{\mathrm{CUT}}$ GeV	σ(pb)	± stat	± sys	± cal
$150.0, 180.0$	0.5	0.025	± 0.010	$^{+0.017}_{-0.020}$	$^{+0.004}_{-0.004}$
$180.0, 210.0$		0.086	± 0.015	$^{+0.036}_{-0.036}$	$^{+0.017}_{-0.009}$
$210.0, 240.0$		0.211	± 0.023	$^{+0.080}_{-0.049}$	$^{+0.000}_{-0.017}$
$240.0, 260.0$		0.370	± 0.040	$^{+0.075}_{-0.065}$	$^{+0.009}_{-0.030}$
$150.0, 180.0$	1.0	0.042	± 0.011	$^{+0.028}_{-0.026}$	$^{+0.000}_{-0.009}$
$180.0, 210.0$		0.180	± 0.019	$^{+0.034}_{-0.034}$	$^{+0.004}_{-0.009}$
$210.0, 240.0$		0.348	± 0.025	$^{+0.083}_{-0.049}$	$^{+0.000}_{-0.013}$
$240.0, 260.0$		0.492	± 0.037	$^{+0.102}_{-0.062}$	$^{+0.024}_{-0.003}$
$150.0, 180.0$	1.5	0.101	± 0.019	$^{+0.070}_{-0.067}$	$^{+0.000}_{-0.027}$
$180.0, 210.0$		0.295	± 0.024	$^{+0.041}_{-0.038}$	$^{+0.017}_{-0.007}$
$210.0, 240.0$		0.588	± 0.031	$^{+0.086}_{-0.089}$	$^{+0.017}_{-0.035}$
$240.0, 260.0$		0.749	± 0.043	$^{+0.113}_{-0.109}$	$^{+0.007}_{-0.030}$
$150.0, 180.0$	2.0	0.151	± 0.020	$^{+0.043}_{-0.055}$	$^{+0.000}_{-0.024}$
$180.0, 210.0$		0.447	± 0.029	$^{+0.090}_{-0.050}$	$^{+0.031}_{-0.000}$
$210.0, 240.0$		0.903	± 0.039	$^{+0.104}_{-0.079}$	$^{+0.022}_{-0.069}$
$240.0, 260.0$		1.132	± 0.051	$^{+0.154}_{-0.097}$	$^{+0.000}_{-0.044}$

Table E.15: The measured differential cross section $d\sigma/dW$ in the region $x_{\gamma}^{\mathrm{OBS}} < 0.75$ unfolded with the average correction factors of Pythia and Herwig for the gap sample of events. The statistical error, systematic errors, and calorimeter energy scale uncertainty on the measurement are also listed.

W bin (GeV)	$E_{\mathrm{T}}^{\mathrm{CUT}}$ GeV	f	± stat	± sys	± cal
$150.0, 180.0$	0.5	0.019	± 0.008	$^{+0.015}_{-0.017}$	$^{+0.003}_{-0.003}$
$180.0, 210.0$		0.013	± 0.002	$^{+0.005}_{-0.005}$	$^{+0.003}_{-0.001}$
$210.0, 240.0$		0.016	± 0.002	$^{+0.006}_{-0.004}$	$^{+0.000}_{-0.002}$
$240.0, 260.0$		0.021	± 0.002	$^{+0.004}_{-0.003}$	$^{+0.000}_{-0.001}$
$150.0, 180.0$	1.0	0.032	± 0.009	$^{+0.025}_{-0.023}$	$^{+0.000}_{-0.008}$
$180.0, 210.0$		0.027	± 0.003	$^{+0.004}_{-0.005}$	$^{+0.001}_{-0.002}$
$210.0, 240.0$		0.027	± 0.002	$^{+0.006}_{-0.004}$	$^{+0.001}_{-0.002}$
$240.0, 260.0$		0.028	± 0.002	$^{+0.005}_{-0.003}$	$^{+0.004}_{-0.001}$
$150.0, 180.0$	1.5	0.077	± 0.014	$^{+0.068}_{-0.058}$	$^{+0.000}_{-0.024}$
$180.0, 210.0$		0.044	± 0.004	$^{+0.005}_{-0.005}$	$^{+0.004}_{-0.001}$
$210.0, 240.0$		0.045	± 0.002	$^{+0.006}_{-0.007}$	$^{+0.003}_{-0.004}$
$240.0, 260.0$		0.043	± 0.002	$^{+0.005}_{-0.005}$	$^{+0.002}_{-0.002}$
$150.0, 180.0$	2.0	0.113	± 0.015	$^{+0.048}_{-0.049}$	$^{+0.000}_{-0.018}$
$180.0, 210.0$		0.067	± 0.004	$^{+0.013}_{-0.007}$	$^{+0.007}_{-0.000}$
$210.0, 240.0$		0.069	± 0.003	$^{+0.007}_{-0.006}$	$^{+0.004}_{-0.008}$
$240.0, 260.0$		0.064	± 0.003	$^{+0.008}_{-0.004}$	$^{+0.003}_{-0.005}$

Table E.16: The measured gap fraction $f(W)$ in the region $x_{\gamma}^{\mathrm{OBS}} < 0.75$ unfolded with the average correction factors of PYTHIA and HERWIG. The statistical error, systematic errors, and calorimeter energy scale uncertainty on the measurement are also listed.

Bibliography

[1] F. Halzen and A.D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. John Wiley & Sons, Inc, 1984.

[2] L.N. Hand, Phys. Rev. **129**, 1834 (1963).

[3] J.J. Sakurai, Ann. Phys. **11**, 1 (1960).

[4] J.J. Sakurai, Phys. Rev. Lett. **22**, 981 (1969).

[5] J.J. Sakurai and D. Schildknecht, Phys. Lett. **B 40**, 121 (1972).

[6] J.D. Bjorken, *International Workshop on Deep Inelastic Scattering and Related Subjects, Eliat, Isreal*, A. Levy (ed.), p. 151. World Scientific, Singapore (1994).

[7] E. Nagy, Nucl. Phys **B 150**, 221 (1979).

[8] G.D. Alkhazov, Preprint LENI-79-541, 1979, available on `http://durpdg.dur.ac.uk/hepdata/reac.html`.

[9] A. Sommerfeld, *Partial Differential Equations in Physics*. Adademic Press, 1949.

[10] A.V. Barnes, Phys. Rev. Lett. **37**, 76 (1976).

[11] Particle Data Group, K. Hagiwara et al., Phys. Rev. **D 66**, 10001 (2002).

[12] M. Froissart, Phys. Rev. **123**, 1053 (1962).

[13] V.N. Gribov, JETP **41**, 667 (1961).

[14] A. Donnachie and P.V. Landshoff, Phys. Lett. **B 296**, 227 (1992).

[15] F.E. Low, Phys. Rev. **D 12**, 163 (1975).

[16] S. Nussinov, Phys. Rev. Lett. **34**, 1286 (1975).

[17] N.N. Nikolaev and B.G. Zakharov, Z. Phys. **C 49**, 607 (1991).

[18] L.N.. Lipatov, Sov. J. Nucl.Phys. **23**, 338 (1976).

[19] E.A. Kuraev, L.N. Lipatov and V.S. Fadin, JETP **44**, 443 (1976).

[20] Ya.Ya. Balitskiĭ and L.N. Lipatov, Sov. J. Nucl. Phys. **28**, 822 (1978).

[21] L.N.. Lipatov, JETP **63**, 904 (1986).

[22] G. Ingelman and P.E. Schlein, Phys. Lett. **B 152**, 256 (1985).

[23] J.D. Bjorken, Phys. Rev. **D 47**, 101 (1993).

[24] H.N. Chehime and D. Zeppenfeld, Preprint hep-ph/9401244, 1994.

[25] ZEUS Coll., U. Holm (ed.), *The ZEUS Detector.* Status Report (unpublished), DESY (1993), available on `http://www-zeus.desy.de/bluebook/bluebook.html`.

[26] N. Harnew et al., Nucl. Inst. Meth. **A 279**, 290 (1989);
B. Foster et al., Nucl. Phys. Proc. Suppl. **B 32**, 181 (1993);
B. Foster et al., Nucl. Inst. Meth. **A 338**, 254 (1994).

[27] M. Derrick et al., Nucl. Inst. Meth. **A 309**, 77 (1991);
A. Andresen et al., Nucl. Inst. Meth. **A 309**, 101 (1991);
A. Caldwell et al., Nucl. Inst. Meth. **A 321**, 356 (1992);
A. Bernstein et al., Nucl. Inst. Meth. **A 336**, 23 (1993).

[28] J. Crittenden et al., *Calibration of FCAL and RCAL Modules at CERN* (unpublished). ZEUS-90-111, internal ZEUS-note, 1990;
A. Caldwell et al., *BCAL Electron Studies Part I: MC and Test Beam Energy Studies* (unpublished). ZEUS-98-02, internal ZEUS-note, 1998.

[29] H. Bethe and W. Heitler, Proc. Roy. Soc. Lond. **A146**, 83 (1934).

[30] Z. Jakubowski et al., ZEUS Internal Note (Unpublished) **ZEUS-Note 90-108** (1990).

[31] H. Boterenbrood et al., Nucl. Inst. Meth. **A 332**, 263 (1993).

[32] V.N. Gribov and L.N. Lipatov, Sov. J. Nucl. Phys. **15**, 438 (1972).

[33] L.N. Lipatov, Sov. J. Nucl. Phys. **20**, 94 (1975).

[34] Yu.L. Dokshitzer, JETP **46**, 641 (1977).

[35] G. Altarelli and G. Parisi, Nucl. Phys. **B 126**, 298 (1977).

[36] M. Gluck, E. Reya, and A. Vogt, Z. Phys. **C 53**, 127 (1992).

[37] J. Botts et al., Phys. Lett. **B 304**, 159 (1993).

[38] A.D. Martin, W.J. Stirling, and R.G. Roberts, Phys. Rev. **D 50**, 6734 (1994).

[39] M. Gluck, E. Reya, and A. Vogt, Phys. Rev. **D 46**, 1973 (1992).

[40] M. Gluck, E. Reya, and A. Vogt, Phys. Rev. **D 45**, 3986 (1992).

[41] G.A. Schuler and T. Sjostrand, Z. Phys. **C 68**, 607 (1995).

[42] K. Hagiwara, M. Tanaka, I. Watanabe, and T. Izubuchi, Phys. Rev. **D 51**, 3197 (1995).

[43] E.A. Kuraev, L.N. Lipatov and V.S. Fadin, JETP **45**, 199 (1977).

[44] M. Ciafaloni, Nucl. Phys. **B 296**, 49 (1988).

[45] S. Catani, F. Fiorani and G. Marchesini, Nucl. Phys. **B 336**, 18 (1990).

[46] B. R. Webber, Nucl. Phys. **B 238**, 492 (1984).

[47] G. Marchesini, B.R. Webber, Nucl. Phys. **B 310**, 461 (1988).

[48] B. Andersson et al., Phys. Rep. **97**, 31 (1983).

[49] R.D. Field and R.P. Feynman, NP **B 136**, 1 (1978).

[50] S.M. Fisher and P. Palazzi, *ADAMO Programmers Manual – Version 3.2.* CERN ECP and RAL, available on `http://adamo.web.cern.ch/Adamo/programmers_manual/TOC_of_adamo.html`.

[51] *Geant v3.13 modules GMOLI.FPP & GMOLS.FPP*, available on `http://wwwinfo.cern.ch/adsoc/geant_html3/node247.html`.

[52] Available on `http://www-zeus.desy.de/ZEUS_ONLY/analysis/orange/index.html`.

[53] M. Bengtsson and T. Sjöstrand, Comp. Phys. Comm. **46**, 43 (1987).

[54] G. Marchesini et al., Comp. Phys. Comm. **67**, 465 (1992).

[55] C.F. von Weizsäcker, Z. Phys. **88**, 612 (1934). In German.

[56] E. Fermi, Z. Phys. **29**, 315 (1924).

[57] J.R. Forshaw and J.K. Storrow, Phys. Rev. **D 46**, 4955 (1992).

[58] J.M. Butterworth, J.R. Forshaw, and M.H. Seymour, Z. Phys. **C 72**, 637 (1996).

[59] G. Moliere, Naturforschung **A2**, 133 (1947).

[60] A.H. Mueller and W.K. Tang, Phys. Lett. **B 284**, 123 (1992).

[61] G.F. Hartner, *VCTRAK(3.07/04): Offline Output Information* (unpublished). ZEUS-97-064, internal ZEUS-note, 1997.

[62] G. F. Hartner, *VCTRAK Briefing: Program and Math* (unpublished). Zeus-98-058, internal ZEUS-note, 1998.

[63] A. Savin, *Study of Calorimeter Noise in the 1996 Data* (unpublished). ZEUS-98-007, internal ZEUS-note, 1998.

[64] C. Catterall and F. Pelucchi, *Comparison of CTD and Calorimeter Measurement of the DIS Positron Angle with High Q^2 Events* (unpublished). ZEUS-97-56, internal ZEUS-note, July 1997.

[65] C. Catterall and Y. Iga, *Improved Simulation of Hadronic Cascades in the Uranium Calorimeter* (unpublished). ZEUS-98-077, internal ZEUS-note, 1998.

[66] M. Wodarczyk, *Measurement of the F_2 Structure Function of the Proton at HERA from 1996 and 1997 ZEUS Data.* Ph.D. Thesis, University of Wisconsin, 1999.

[67] N. Tuning and M. Wodarczyk, *rcalcor.fpp, Calorimeter correction routine, PHANTOM Users Library.*

[68] H. Abramowicz, A. Caldwell and R. Sinkus, Nucl. Inst. Meth. **A 365**, 508 (1995).

[69] N. Tuning, *ZUFOs: Hadronic Final State Reconstruction with Calorimeter, Tracking and Backsplash Correction* (unpublished). ZEUS-Note-01-021, 2001.

[70] J.E. Huth et al., *Research Directions for the Decade. Proceedings of Summer Study on High Energy Physics, 1990*, E.L. Berger (ed.), p. 134. World Scientific (1992). Also in preprint FERMILAB-CONF-90-249-E.

[71] *ZEUS Lumi Web Page*, available on `http://www-zeus.desy.de/physics/lumi/`.

[72] S. Chekanov, et al., Eur. Phys. J. **C23**, 615 (2002).

[73] S.Catani et al., Nucl. Phys. **B406**, 187 (1993).

[74] S.D. Ellis and D.E. Soper, Phys. Rev. **D 48**, 3160 (1993).

[75] G. Oderda and G. Sterman, Phys. Rev. Lett. **81**, 3591 (1998).

[76] J.M. Butterworth and S.Butterworth, Comput. Phys. Commun. **153**, 164 (2003).

[77] J. Bromley, et al.. Prepared for Workshop on Future Physics at HERA (Preceded by meetings 25-26 Sep 1995 and 7-9 Feb 1996 at DESY), Hamburg, Germany, 30-31 May 1996.

[78] L. Lönnblad, Comp. Phys. Comm. **71**, 15 (1992).

[79] Y. Azimov, Y. Dokshitzer,V.A. Khoze, and S.I. Troian, Phys. Lett. **B165**, 147 (1985).

[80] G. Gustafson, Phys. Lett. **B175**, 453 (1986).

[81] G. Gustafson and U. Pettersson, Nucl. Phys. **B306**, 746 (1988).

[82] B. Andersson, G. Gustafson, L. Lonnblad, and U. Pettersson, Z. Phys. **C43**, 625 (1989).

[83] C. Gwenlan, *Jets and Energy Flow in Photoproduction using the ZEUS Detector at HERA.* Ph.D. Thesis, University College London, 2003.

[84] D. Chapin, S. Lammers, ZEUS Internal Note (Unpublished) **ZEUS-Note 2001-006** (2001).

[85] ZEUS Coll., M. Derrick et al., Phys. Lett. **B 369**, 55 (1996).

[86] C. Adloff, et al., Eur. Phys. J. **C24**, 517 (2002).