

Diss. ETH No. 18652 and
DESY-THESIS-2009-047



MEASUREMENT OF BEAUTY
PHOTOPRODUCTION AT THRESHOLD
USING DI-ELECTRON EVENTS
WITH THE H1 DETECTOR AT HERA

A dissertation submitted to the
SWISS FEDERAL INSTITUTE OF TECHNOLOGY
ZURICH
for the degree of
DOCTOR OF SCIENCE

presented by
MICHEL DAVID SAUTER

Dipl. Phys. ETH
born on June 27, 1980
citizen of Sulgen (TG)

accepted on the recommendation of
Prof. Dr. Ralph Eichler, examiner and
Prof. Dr. Christoph Grab, co-examiner and
Prof. Dr. André Schöning, co-examiner

2009

Abstract

The cross section of $b\bar{b}$ photoproduction in ep collisions has been measured with the H1 detector at HERA. Events containing b -quarks were identified through detection of two low momentum electrons in the final state. Semileptonic decays $b\bar{b} \rightarrow eeX$ were exploited in the kinematic range of the photon virtuality $Q^2 < 1 \text{ GeV}^2$, the inelasticity $0.2 < y < 0.8$ and the pseudorapidity of the b -quarks $|\eta(b), \eta(\bar{b})| < 2$. The measured differential b -quark production cross section as a function of the transverse b -quark momentum extends the previously experimentally accessible phase space towards the b -quark production threshold. The results are compared to other b -quark cross section measurements, as well as to leading-order and next-to-leading-order QCD predictions. The extension to lower b -quark momenta became possible with a dedicated low momentum electron trigger in the data period 2007, which combines track (Fast Track Trigger) and calorimeter information (Jet Trigger), and by mastering the experimental challenges of low p_T -electron identification.

Zusammenfassung

Der Wirkungsquerschnitt von $b\bar{b}$ -Photoproduktion wurde in ep Kollisionen mit Hilfe des H1 Detektors bei HERA gemessen. b -Quark Ereignisse wurden über die Detektion von zwei Elektronen bei niedrigem Impuls im Endzustand identifiziert, mittels des semileptonischen Zerfallskanal $b\bar{b} \rightarrow eeX$ und im kinematischen Bereich der Photonvirtualität $Q^2 < 1 \text{ GeV}^2$, der Photoninelastizität $0.2 < y < 0.8$ und der Pseudorapitität der b -Quarks $|\eta(b), \eta(\bar{b})| < 2$. Der Differentielle b -Quark Produktionswirkungsquerschnitt wurde als Funktion des transversalen b -Quark Impulses gemessen und erweitert den experimentell zugänglichen Phasenraum früherer Messungen zur b -Quark Produktionsschwelle. Die Resultate werden sowohl mit anderen b -Quarkmessungen, als auch mit leading-order und next-to-leading-order QCD Vorhersagen verglichen.

Die Erweiterung zu kleineren transversalen Impulsen des b -Quarks erfolgte in der Datenperiode 2007 mit speziellen Niederimpuls Elektron Triggern, die Spurinformaton (Fast Track Trigger) mit Kalorimeterinformation (Jet Trigger) kombinieren und dank der Kontrolle der experimentell schwierigen Identifikation von Elektronen bei niedrigen transversalen Impulsen.

Contents

1	Introduction	1
2	Theoretical Framework	3
2.1	Quantum Chromodynamics	3
2.2	Electron Proton Scattering	5
2.3	The Naive Quark Parton Model	6
2.4	Divergences in Perturbative Quantum Chromodynamics	7
2.4.1	Ultraviolet Divergences and Renormalization	7
2.4.2	Infrared Divergences and the Factorization Theorem	8
2.5	Factorization for Structure Functions	8
2.6	Evolution Equations	8
2.7	Heavy Quark Photoproduction in ep Scattering	9
2.7.1	Fragmentation	11
2.7.2	Weizsäcker-Williams distribution	13
2.8	NLO Calculations	13
3	The H1 Detector at HERA	15
3.1	The HERA Accelerator at DESY	15
3.2	The H1 Detector at HERA	16
3.2.1	Central Tracking System	19
3.2.2	The Liquid Argon Calorimeter	19
3.2.3	Luminosity Measurement	20
3.2.4	The H1 Trigger System	20
3.2.5	Trigger Upgrades	22
4	Aim and Concept of the Measurement	25
4.1	Motivation and Aim of the Measurement	25
4.2	Concepts of Previous Beauty Measurements at HERA	25
4.3	Concept of this Measurement	26
4.3.1	Conclusions on the Concept of this Measurement	29
5	Monte Carlo Simulations	31
5.1	Simulation of Events	31
5.1.1	Event Generator	31
5.1.2	Leading Order Monte Carlo Event Generators: Pythia and Cascade	31
5.1.3	Simulation of the Detector Response	32

5.2	Monte Carlo Samples	32
6	Electron Identification	35
6.1	Method	35
6.2	Acceptance and Preselection Cuts	36
6.2.1	Track Selection	36
6.2.2	LAr Acceptance Cuts	37
6.2.3	Preselection Cuts	37
6.3	Input Variables to the Artificial Neural Network	37
6.4	Artificial Neural Network Structure and Training	39
6.5	Control Distributions of the NN Input Variables	45
6.6	Result	45
6.7	Other Measurements based on the Electron Identification presented	45
7	Event Selection	49
7.1	Trigger Selection	49
7.1.1	A FTT-JT based Electron Trigger for Beauty Events	49
7.1.2	Combination of Different Subtriggers and Correction for Prescale Factors	50
7.1.3	Event Weights Correcting for Prescale Factors	52
7.2	Subtriggers Selection	54
7.3	Trigger Efficiency for Beauty Events	54
7.4	Offline Selection	56
7.4.1	Run Selection	56
7.4.2	Rejection of Background Events	57
7.5	Selection of Di-Electron Events in Photoproduction	58
7.5.1	Preselected Di-Electron Sample in Photoproduction	58
7.5.2	Final Di-Electron Sample	58
7.6	Control Distributions for the Di-Electron Preselection Sample	59
8	Cross Section Measurement	69
8.1	Measurement of the Mean Transverse Momentum of the Beauty Quarks	69
8.1.1	An Estimator for the Mean Transverse Energy of Beauty Quarks based on the Thrust Axis in the Transverse Plane	69
8.2	Unfolding the Differential Beauty Cross Section	73
8.2.1	Definition of the Response Matrix	73
8.2.2	Filling of the Response Matrix	78
8.2.3	Determination of the Regularization Parameter	78
8.3	Final Distributions	81
8.4	Systematic Uncertainties and Bias Correction	81
8.5	Results and Discussion	84
8.5.1	Determination of the Signal and Background Contributions	84
8.5.2	Differential Cross Section as Function of the Mean Trans- verse Momentum of the Beauty Quarks	84
8.5.3	NLO pQCD Calculation	90
8.6	Comparison with Previous Measurements at HERA	94

9	Conclusions and Outlook	97
A	Unfolding	99
A.1	Formulation of the Unfolding Problem	99
A.2	Determination of the Response Matrix	101
A.3	Regularized Unfolding	101
A.3.1	Choice of the Regularization Parameter	104
A.3.2	Interpretation of the Regularization Parameter	104
A.3.3	Computer Programs for Unfolding	105
B	Dead Material Correction	107
C	Calibration of the Jet Trigger Simulation	111
D	The FTT L3 Trigger	117
D.1	Introduction to the FTT L3 System	117
D.2	L3 Physics Algorithms	117
D.2.1	D* and J/ψ Triggers	118
D.2.2	Muon Trigger	118
D.2.3	Electron Trigger	118
D.3	L3 Trigger Simulation	123
E	Example for the Combination of Subtriggers	127
E.1	Introduction	127
E.2	Uncorrelated Subtriggers	128
E.3	Correlated Subtriggers	128
E.3.1	Method I - List	129
E.3.2	Method II - Formula	129
	List of Figures	133
	List of Tables	135
	List of Acronyms	137
	Bibliography	139
	Acknowledgments	145
	Curriculum Vitae	147

Chapter 1

Introduction

High energy colliders are the strongest microscopes mankind ever constructed, and permit us to probe the structure of matter at the shortest distances. The purpose of these experiments is to explore questions about the fundamental structure and the fundamental forces of nature related to the composition and development of the universe directly after the Big Bang.

To perform these experiments billions of particles are packed into tiny bunches which are then accelerated with the help of electromagnetic fields to almost the speed of light. The accelerated particle bunches form a beam which is often bent with the strongest available magnets in storage rings of several kilometers circumference. Two beams are then sent in opposite directions and focused in specific points on each other. At these points particles can collide, where each collision produces a shower of new particles. By measuring these showers in detectors that surround the collision points, one can study the footprints of the initial particle interaction.

Particle physics experiments have a tendency to operate at, and also to push the technological frontier. Behind each component of such an accelerator and such a particle detector an enormous technological development and innovation is hidden. To approach these challenges scientists form collaborations involving hundreds of people.

The Standard Model of particle physics serves since approximately 30 years as the theory to describe matter. It reduces the wide spectrum of particles composing matter to a set of twelve elementary fermions (particles with half-integral spin). This set can be subdivided into three families of quarks and leptons. These fermions interact within the Standard Model through three forces, by exchanging a boson (particle with integral spin). Quarks predominantly interplay among each other via the strong force by exchanging a gluon, whilst the coupling to the leptons is described by the electroweak force. The electroweak force splits at low energies into the electromagnetic force, carried by a massless photon, and the weak force, which is mediated by the exchange of a massive Z or W boson. All the predictions derived from the Standard Model have been experimentally measured to a general high accuracy except for one, the Higgs boson, which is a crucial piece of the electroweak unification and the masses of the Z and W bosons. Nevertheless, due to the lack of the incorporation of gravity in the Standard Model and the wide range of parameters, it is believed that an even more fundamental, yet unknown theory hides behind the Standard Model. Questions related to the measurement of the remaining Higgs boson and to the measurement of physics beyond the Standard Model are addressed at the new LHC collider at the research center CERN, which is just to start its regular operation these days.

This work deals with beauty quark production at threshold at the electron proton collider HERA at the research center DESY. Why should one study heavy quark production at the

eve of a next generation particle collider, the LHC? The LHC collides protons with protons. Protons are composed of three quarks that predominantly interact via the strong force, by exchanging a gluon. Quantum Chromodynamics (QCD) is the theory that describes the strong force and nowadays is regarded as one of the cornerstones of the Standard Model. In order to exploit physics at the LHC one needs a quantitative and precise understanding of QCD.

Unlike to the carriers of the electroweak force, gluons can also couple to each other. This self-coupling leads to a scale dependence of the strong coupling α_s , that becomes weaker at higher scales, respectively smaller distances (asymptotic freedom).

The appropriate calculation techniques to derive predictions in QCD are perturbative expansions in α_s , which are generally denominated as perturbative Quantum Chromo Dynamics (pQCD). PQCD calculations only converge if at least one hard scale is present.

The production of heavy quarks in ep collisions at HERA is an important testing ground for pQCD, since among others the large b-quark mass provides such a hard scale. In particular the beauty production cross section at HERA as function of the transverse momentum of the beauty quark was predicted by pQCD calculations quite accurately. Frixione et al. comment their calculation in a paper [1] by "We can regard [our calculation] as reliable prediction of QCD for the p_T spectrum of b hadrons at HERA. The comparison of this prediction with the data would be extremely useful in the light of the status of the comparison between theory and data...".

The first beauty production measurements at HERA have shown a discrepancy between the measured data and the prediction. In particular in photoproduction and towards low values of the transverse momentum of the beauty quarks $p_T(b)$, the prediction underestimated the data.

This thesis deals with the measurement of beauty photoproduction as function of $p_T(b)$. The measurement presented extends the phase space of all previous measurements at HERA in $p_T(b)$ and includes the beauty production threshold. At the beginning of this thesis new sophisticated beauty triggers were developed that allowed the data collection within only three months, notably just before the shut down of the HERA collider. Afterwards the measured data was analyzed with techniques developed with a special focus on not restricting the $p_T(b)$ spectrum.

After some remarks on basic theoretical aspects of HERA physics, and a short introduction of the HERA collider and the H1 experiment, the thesis concentrates on the differential beauty cross section measurement. The thesis is completed by a detailed Appendix covering some side topics that are not necessarily needed for the understanding of the measurement. But each chapter in the Appendix certainly contributes to the overall picture of the work done in the context of this thesis, and might also contain some interesting ideas and concepts to other analyses.

Chapter 2

Theoretical Framework

2.1 Quantum Chromodynamics

Quantum Chromodynamics (QCD) [3–6] is the field theory of the strong force, the fundamental force describing the interaction of quarks and gluons. Quarks have three 'color' charges (r=red, g=green and b=blue) and interact by exchanging virtual gluons. Gluons are massless spin=1 bosons and are the 'force carriers' of the strong interaction, similar to the photons in Quantum Electrodynamics (QED) [7, 8]. Like QED, QCD is a field theory, but with eight colored charged gluons instead of a single chargeless photon. Due to their charge gluons are able to interact with each other, and therefore are also allowed to split up: $g \rightarrow gg$. The experimental consequences of this behavior are dramatic:

The electromagnetic force acts on long distances. An electron placed at the origin is measurable with a test charge placed at infinity. The closer the test charge is brought to the origin, the higher will be the charge measured. The reason for this phenomenon is that the electron constantly emits virtual photons, which may split-up into electron positron pairs. The electron is surrounded by a cloud of polarized charges that screen it's bare charge at long distances. This charge-screening effect is directly reflected in the scale dependence of the fine structure constant α_{em} (see Figure 2.1(a)).

In QCD color-charged quarks also constantly emit virtual gluons that are allowed to split up. QCD would be a carbon copy of QED if not the additional gluon splitting configurations contributed. It turns out, that the gluon self-interaction diagrams dominate, which reverses the QED result and leads to an anti-screening effect referred to as 'asymptotic freedom'. The strong coupling constant α_s is at high momentum scales (short distances)¹ small and increases towards small momentum scales (large distances). It becomes large at momentum scales on the order of $\Lambda_{QCD} \approx 0.2 \text{ GeV}$, which corresponds to distances on the order of $\approx 1 \text{ fm}$. Color charged, strongly interacting partons are at small momenta bound to color-neutral hadrons (designated as 'confinement') and become only at higher momentum scales effectively free (see Figure 2.1(b)).

The property that α_s becomes small at high scales turns QCD into a quantitative calculable theory. In the scheme of perturbative QCD (pQCD) particle scattering cross sections are calculated in power series of α_s . Naturally, the smaller α_s becomes the more reliable are the corresponding cross section calculations. The ideal hard scale at which this is the case and at which the cross sections are evaluated, depends on the involved experimental scales. In electron proton interactions these are the transverse momentum of the studied quark, the

¹According to the Heisenberg Uncertainty Principle, short (long) distances correspond to high (low) momenta, and short time (long) scales correspond to high (low) energy scales.

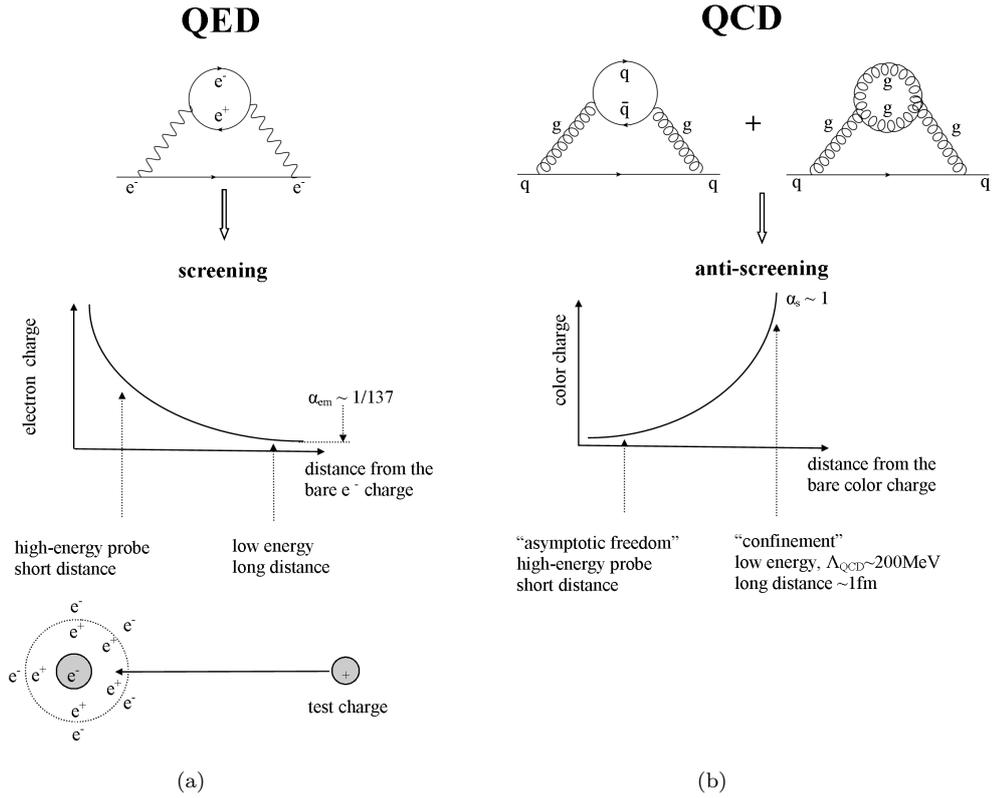


Figure 2.1: Coupling Constant in QED and QCD

Quantum Electrodynamics (QED) and Quantum Chromodynamics (QCD) are compared in the figure. In QED (a) electrons constantly emit photons that may split-up into e^+e^- -pairs. This results in charge screening and the electromagnetic coupling constant α_{em} becomes small at long distances. In QCD (b) the additional gluon-splitting diagrams lead to an anti-screening effect and a strong coupling constant that increases towards long distances. (Adapted from [2]).

photon virtuality and in the case of heavy quarks the quark mass. A QCD calculation to all orders in α_s does not depend on the choice of the hard scale. In practice however, a dependency arises due to the truncation of the non available higher orders of the perturbative power series.

Measurements of electron proton cross sections provide a good test for pQCD calculations. In this thesis beauty quark production in ep interactions is measured and compared to next to leading order (NLO) pQCD predictions. The hard scale in these calculations is given by the heavy beauty quark mass and the transverse momentum of the beauty quark. In the following the discussion concentrates on the theoretical basics of ep scattering.

2.2 Electron Proton Scattering

The generic scattering of a high energetic electron (or positron) off a proton is described by the exchange of a single virtual gauge boson, as shown in the Feynman diagram in Figure 2.2. At HERA the scattering is predominately mediated by the exchange of a virtual photon. However the exchange of Z^0 and W^\pm bosons is possible as well. W^\pm boson exchange is referred to Charged Current, whereas Z^0 and γ exchange are summarized as Neutral Current. In Charged Current processes the outgoing lepton is a neutrino, which leads to experimental events signatures with missing transverse momentum.

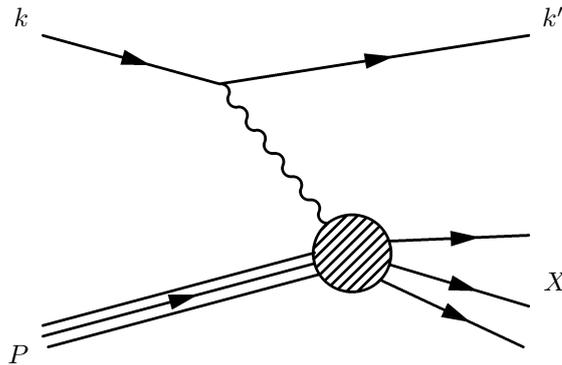


Figure 2.2: Lepton Proton Scattering
Feynman diagram of lepton proton scattering.

Using the notation k , k' for the four-momenta of the electron and P for the proton respectively, then the standard Lorentz invariant variables describing the kinematics of the interaction are defined as:

- The center of mass energy squared

$$s = (k + P)^2 \approx 4 \cdot E_e E_P \quad , \quad (2.1)$$

which is determined by the energy of the colliding beams. With the electron beam having $E_e = 27.6 \text{ GeV}$ and the proton beam $E_P = 920 \text{ GeV}$, the center of mass energy at HERA was 318 GeV .

- The negative squared four momentum transfer,

$$Q^2 = -q^2 = -(k - k')^2 \quad , \quad (2.2)$$

which describes the virtuality of the exchanged boson. If the exchanged boson is a quasi real photon ($Q^2 \lesssim \Lambda_{QCD}^2$) the process is referred to as photoproduction (γp), while high Q^2 processes are designated as Deep Inelastic Scattering (DIS). In the H1 experiment DIS events are distinguished from photoproduction by the detection of the scattered beam electron. In the case of photoproduction the beam electron escapes below the experimental acceptance of the backward calorimeter ($Q^2 > 4 \text{ GeV}^2$) in the beam pipe.

- The Bjorken scaling variable

$$x = \frac{Q^2}{2P \cdot q} \quad (0 < x < 1) \quad . \quad (2.3)$$

In the quark parton model (see section 2.3) x corresponds to the fraction of the proton momentum carried by the struck quark.

- The inelasticity

$$y = \frac{P \cdot q}{P \cdot k} \quad (0 < y < 1) \quad , \quad (2.4)$$

which can be interpreted as the fraction of the incoming electron energy carried by the mediating photon in the photon proton rest frame.

- The energy of the photon-proton center of mass system

$$W_{\gamma P}^2 = (P + q)^2 \quad . \quad (2.5)$$

These variables are not independent, but related via the equations:

$$Q^2 = sxy \quad (2.6)$$

$$W_{\gamma P}^2 = ys - Q^2 \quad (2.7)$$

Hence for fixed beam energies (respectively s) the kinematics of the inclusive scattering process can be described by any set of two independent variables out of Q^2 , x , y and $W_{\gamma P}$.

2.3 The Naive Quark Parton Model

In the single photon approximation² the differential NC cross section is commonly written as [9]:

$$\frac{d^2\sigma_{NC}}{dx dQ^2} = \frac{2 \cdot \pi\alpha}{xQ^4} [(1 + (1 - y)^2)F_2(x, Q^2) - y^2 F_L(x, Q^2)] \quad . \quad (2.8)$$

$F_i(x, Q^2)$ are the so-called Structure Functions, that parameterize the structure of the proton, seen by the virtual photons.

In the naive Quark Parton Model (QPM) [10] the proton is described by the composition of point-like spin 1/2 quarks. Quarks in the proton are elastically scattered on photons emitted from the incoming electron. A necessary condition for such a picture to make sense is that the interaction time between the photon and the struck quark is much shorter than the interaction time among the quarks. The internal proton structure is expressed in terms of parton density functions $f_q(x)$, which represent the probability to find a quark q in the proton, carrying the proton momentum fraction between x and $x + dx$. The structure functions

²In the single photon approximation the Z boson exchange, which is strongly suppressed at low Q^2 values, is neglected.

of equation (2.8) are related to the parton density functions of the naive QPM as

$$F_2(x) = \sum_{q,\bar{q}} e_q^2 x f_q(x) \quad , \quad (2.9)$$

$$F_L(x) = 0 \quad . \quad (2.10)$$

The sum in equation (2.9) is computed over the quark flavors in the proton and e_q are the quark charges. The following two aspects are important in the naive QPM:

- The structure function F_2 only depends on the Bjorken scaling variable x .
- The structure function $F_L(x, Q^2)$, which is related to the absorption of longitudinally polarized virtual photons, is zero due to the spin 1/2 nature of quarks (Callan-Gross relation).

Although the naive QPM described the data of the early ep scattering experiments quite well, things are not that simple. Higher order virtual states, in which quark antiquark pairs of any flavor occur, are not included in the model. In addition, gluon radiation and gluon splitting are necessary to describe scaling violation, i.e. the Q^2 dependence of F_2 [11]. The corresponding additional diagrams however, also lead to divergences in the cross section calculations.

2.4 Divergences in Perturbative Quantum Chromodynamics

The cross section calculations in fixed order pQCD suffer from two types of divergences, the ultraviolet divergences and the infrared divergences [12, 13].

2.4.1 Ultraviolet Divergences and Renormalization

The ultraviolet divergences are associated with virtual loop corrections leading to divergent integrals. In theoretical calculations ultra violet divergences are compensated by replacing the divergent integrals in a systematic way through finite expressions. This technical procedure is called renormalization, which is performed at a fixed scale, the renormalization scale μ_R . The renormalization scale μ_R is not an intrinsic parameter of QCD, even though μ_R is required for cross section calculations. Therefore any physical observable measured at an experimental scale Q cannot depend on the choice of μ_R .

The independence of a physical observable on μ_R can mathematically be used to derive the renormalization group equation [13]. The solution of it describes the relation of the strong coupling constant evaluated at the renormalization scale μ_R^2 and the strong coupling constant evaluated at the experimental scale Q^2 . The relation between $\alpha_s(\mu_R^2)$ and $\alpha_s(Q^2)$ is given in lowest order by

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu_R^2)}{1 + \alpha_s(\mu_R^2) \cdot b \cdot \ln\left(\frac{Q^2}{\mu_R^2}\right)} \quad b = \frac{33 - 2 \cdot n_f}{12 \cdot \pi} \quad , \quad (2.11)$$

with n_f being the number of active flavors in the loop correction. Obviously $\alpha_s(Q^2)$ depends on logarithm $\ln\left(\frac{Q^2}{\mu_R^2}\right)$. It decreases for large scales Q (asymptotic freedom) and becomes large for small scales Q (confinement).

2.4.2 Infrared Divergences and the Factorization Theorem

Infrared divergences arise due to real gluon emissions. They either occur, if the gluon is radiated in the direction of the outgoing parton (collinear divergence), or by the emission of a low energetic soft gluons (soft divergence). Both divergences are related to long-distance physics. In pQCD almost all cross sections are infrared divergent. Nevertheless there exists also infrared safe quantities, i.e. quantities that only depend on the short-distance QCD-physics and not on the long-distance dynamics. The separation of the long- from the short-distance physics is based on the factorization theorem.

According to the factorization theorem [14–16], the description of a scattering process can be written as the product of two independent components: a) the hard interaction, related to the short-distance physics process, and b) the long-distance effects that include the infrared divergences. a) is calculable in pQCD, whereas b) is non-perturbative and needs to be determined by experiment. It is assumed that these long-distance effects are universal, i.e. process independent.

2.5 Factorization for Structure Functions

The field theoretical realization of the quark parton model is based on the factorization theorem, which is essential to separate the short-distance physics (hard subprocess) from the long-distance physics (soft subprocesses). The generalization of equation (2.9) is [12]

$$F_2(x, Q^2) = \sum_{i=q,\bar{q},g} \int_0^1 \frac{d\xi}{\xi} C_2(x/\xi, Q^2/\mu_R, \mu_F^2/\mu_R^2, \alpha_s(\mu_R^2)) \cdot f_i(\xi, \mu_F^2, \mu_R^2) \quad . \quad (2.12)$$

The hard-scattering function $C_2(x/\xi, Q^2/\mu_R, \mu_F^2/\mu_R^2, \alpha_s(\mu_R^2))$ contains all information about the hard subprocess (the short-distance physics), and can be calculated in perturbative Quantum Chromodynamics (pQCD), by power series expansions of the strong coupling constant α_s . The distributions $f_i(\xi, \mu_F^2, \mu_R^2)$ are the parton distribution functions PDFs and describe the probability to find a certain parton of type i with a longitudinal momentum ξP in the proton. These distributions cannot be calculated from first principles, since they receive contributions from the long-distance (non perturbative) part of the strong interaction, but have to be measured. The factorization scale μ_F is the scale which separates the long- from the short-distance physics. Thus a parton emitted with a transverse momentum less than μ_F is absorbed in the PDF. A parton emitted at large transverse momentum is part of the short-distance cross section.

However, the details on this separation also depend on the factorization scheme. The commonly used factorization schemes are the modified minimal subtraction factorization scheme (\overline{MS} -scheme) and DIS factorization scheme (see also [12]). The second scale of equation (2.12), μ_R is the renormalization scale, already discussed in section 2.4.1.

2.6 Evolution Equations

As mentioned above, the PDFs cannot be derived from pQCD calculations. However, what can be calculated perturbatively is how they depend on the factorization scale μ_F . Once the PDFs are known at a starting scale μ_0 , the slopes of the PDFs - and in particular their scaling violation behavior - can be translated to any other scale. This transition is described by the QCD parton evolution equations. The scaling violation behavior is incorporated in the theory by the allowance of gluon radiation and gluon splitting, which results in a gluon ladder as shown in Figure 2.3. However, the incorporation of gluon emissions in the theory

also introduces large logarithms that have to be resummed. Several approaches exist to organize the resummation of the partons contributing to the gluon ladder, which result in different evolution equation models.

In the DGLAP evolution equation [17–20] the partons of the gluon ladder are ordered according to their transverse momenta, $k_{T,i} > k_{T,i-1} > \dots > k_{T_0}$, which results in sums over $\alpha_s \log(Q^2)$.

In the BFKL ansatz [21, 22] a contrary approach is followed by ordering the partons of the gluon ladder by their fractional longitudinal momenta: $x_i > x_{i-1} > \dots > x_0$. This results in sums over $\alpha_s \log(1/x)$. Depending on which set of logarithms is dominant, the DGLAP and the BFKL equation are valid in a different region of phase space (x, Q^2) , as illustrated in Figure 2.4. The DGLAP evolution equation is expected to be valid for large Q^2 and not too small x values, while the BFKL equation is more adequate for moderate Q^2 and low x values.

The CCFM evolution equation [23–26] performs an evolution of PDFs to x and Q^2 by an angular ordering of the gluons, until a cut-off angle that can be related to the hard scale of the problem. This approach however is more complicated and involves unintegrated parton density functions $A(x, k_T^2, Q^2)$, which specify the probability of finding a gluon with the longitudinal momentum fraction x and transverse momentum k_T . The traditional integrated gluon density depending on the longitudinal momentum fraction x only, can be recovered by integration over the transverse momenta k_T :

$$xg(x, Q^2) = \int^{Q^2} dk_T^2 A(x, k_T^2, Q^2) \quad (2.13)$$

Since in the CCFM approach the gluon densities depend on the transverse momenta, it is often referred to as k_T -factorization, while in the case of DGLAP one speaks of collinear factorization.

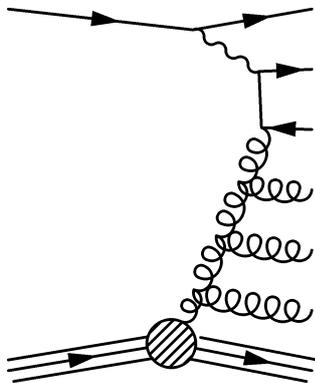


Figure 2.3: Gluon Ladder

Gluon emissions in the parton evolution, the so called gluon ladder.

2.7 Heavy Quark Photoproduction in ep Scattering

Electron proton scattering at HERA produced quarks with the flavors up, down, strange, charm and beauty, while top quarks were kinematically not accessible. The three first quarks are designated as light quarks, because their mass is small compared to Λ_{QCD} , while charm

and beauty quarks are called heavy, $m_{c,b} > \Lambda_{QCD}$. In the following the discussion is focused on heavy quark photoproduction ($Q^2 \lesssim \Lambda_{QCD}^2$).

In photoproduction a quasi real photon, emitted from the incoming electron, interacts with the incoming proton. Heavy quarks can be produced in several processes, for which the main Feynman diagrams are listed in Figure 2.5. The dominant contribution for beauty production is the boson-gluon-fusion (BGF), in which a photon emitted from the electron fuses with a gluon from the proton to a beauty quark pair. This process is referred to direct, or point-like. A second class of processes are the so called resolved processes, in which the photon fluctuates into a hadronic state before undergoing a hard collision, Figures 2.5(b)-(e). Resolved processes are further subdivided into hadron-like, Figure 2.5(b), and heavy flavor excited processes, Figures 2.5(c)-(e), which are strongly suppressed in beauty photoproduction. Therefore, a differential photon-proton cross section can be written as the sum of a point-like and a hadronic-contribution [1]

$$d\sigma^{\gamma P}(P_\gamma, P_P) = d\sigma_{point}^{\gamma P}(P_\gamma, P_P) + d\sigma_{hadr}^{\gamma P}(P_\gamma, P_P) \quad , \quad (2.14)$$

with P_P and P_γ being the four-momenta of the incoming proton and photon. Thanks to the factorization theorem [14–16] (and section 2.4.2) the equation (2.14) factorizes into different contributions [1, 12]:

$$d\sigma_{point}^{\gamma P}(P_\gamma, P_P) \propto \sum_{h,j,k} \int dx \frac{dz}{z} f_j^P(x, \mu_F) \cdot d\hat{\sigma}_{\gamma j \rightarrow kX}(P_\gamma, xP_P, \alpha_s(\mu_R), \mu_R, \mu_F) \cdot D_k^h(z, \mu_F) \quad (2.15)$$

$$d\sigma_{hadr}^{\gamma P}(P_\gamma, P_P) \propto \sum_{h,i,j,k} \int dx dy \frac{dz}{z} f_i^\gamma(y, \mu_F) \cdot f_j^P(x, \mu_F) \cdot d\hat{\sigma}_{ij \rightarrow kX}(yP_\gamma, xP_P, \alpha_s(\mu_R), \mu_R, \mu_F) \cdot D_k^h(z, \mu_F) \quad (2.16)$$

The different contributions in the equations (2.15) and (2.16) are:

- The proton PDF $f_j^P(x, \mu_F)$ and the photon PDF $f_i^\gamma(y, \mu_F)$. They determine the probability to find the parton j (i) in the proton (photon).
- $d\hat{\sigma}_{\gamma j \rightarrow kX}(P_\gamma, xP_P, \alpha_s(\mu_R), \mu_R, \mu_F)$ and $d\hat{\sigma}_{ij \rightarrow kX}(P_\gamma, xP_P, \alpha_s(\mu_R), \mu_R, \mu_F)$ denote the Born cross section of the hard collision $i + j \rightarrow k + X$ and $\gamma + j \rightarrow k + X$ (with i, j and k being partons). Apart from coefficients, these are the matrix elements calculable

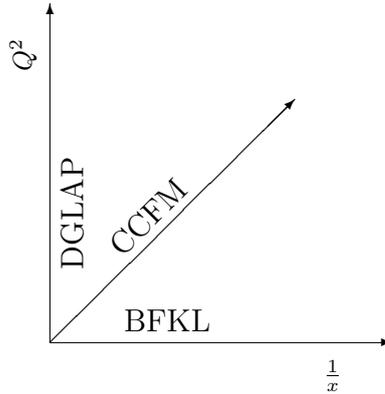


Figure 2.4: Schematic Overview on the DGLAP, BFKL and CCFM Evolution Equation
The direction of evolutions are drawn in the $1/x - Q^2$ plane.

in pQCD. μ_F is the factorization scale that separates the short- from the long-distance physics and μ_R the renormalization scale.

- The fragmentation function $D_k^h(z, \mu_F)$ describes the probability to hadronize a parton k into a colorless hadron h . z is the energy fraction of the parent quark that is transferred to the newly created hadron (see section 2.7.1 for more details).

The factorization theorem is applied twice in this form of the $d\sigma^{\gamma P}$ and $d\sigma_{hadr}^{\gamma P}$ parametrization: Infrared divergences related to soft gluons in the proton and photon are absorbed in the proton and photon PDFs, whereas the 'final state' infrared divergences related to the transition of partons to hadrons are absorbed in the fragmentation function.

The perturbative calculation of the matrix elements of the hard processes mentioned above leads to complications due to the presence of several hard scales. These are the transverse momentum of the beauty quark p_T , the beauty quark mass³ m_b and in the case of DIS also Q^2 .

Several perturbative calculations exist, that vary in the way the multiple scales are taken into account, and in the region of phase space that they are expected to give a reliable result. The following competing approximations are most commonly applied [27]:

- In the massless scheme the heavy quarks are treated as massless. For scales μ_F larger than the heavy quark mass, they are treated as active partons in the proton, respectively the resolved photon [28]. These calculations are called massless or 'Zero Mass Variable Flavor Number Scheme' (ZMVFNS). In photoproduction the massless scheme is expected to be valid for transverse momenta of the heavy quark that are much higher than the beauty (charm) quark mass, $p_T \gg m_{b(c)}$.
- In the massive or 'Fixed Flavor Number Scheme' (FFNS) the beauty and charm quarks are treated as massive, and their mass is taken into account as a parameter. This approach is expected to describe the data in the regions of phase space, where the transverse momentum of the beauty quark is of the same order as the beauty mass. Charm and beauty quarks are not treated as constituents of the proton, but generated perturbatively only.

2.7.1 Fragmentation

The fragmentation formalism describes how a quark from the hard interaction becomes a colorless hadron. For quarks leaving the hard interaction the strong coupling constant rises at long distances, spoiling the perturbative calculations. The effect can be modeled by the convolution of the partonic cross sections with a fragmentation function. Different phenomenological parameterizations for heavy quark fragmentation exist, among the popular ones is the Peterson fragmentation function [29],

$$D_Q^H(z) = \frac{N}{z[1 - (1/z) - \epsilon_Q/(1-z)]^2} \quad , \quad (2.17)$$

which describes the probability that a heavy quark Q fragments into a hadron H . z is the energy fraction of the parent quark that is transferred to the newly created hadron, ϵ_Q a phenomenological parameter and N a normalization constant.

Heavy quark fragmentation functions peak at high z values, since by attaching a light anti-quark \bar{q} to a heavy quark Q (or vice versa), the heavy quark decelerates only slightly. Thus Q and $Q\bar{q}$ carry almost the same energy. Whereas light quark fragmentation functions peak at lower z values and behave more like $D_q^h(z) \sim z^{-1} \cdot (1-z)^2$, i.e. the peak position can be

³Since the mass of the heavy quarks is above Λ_{QCD} it can serve as the hard scale for perturbative calculations, whereas this is not the case for the light quarks.

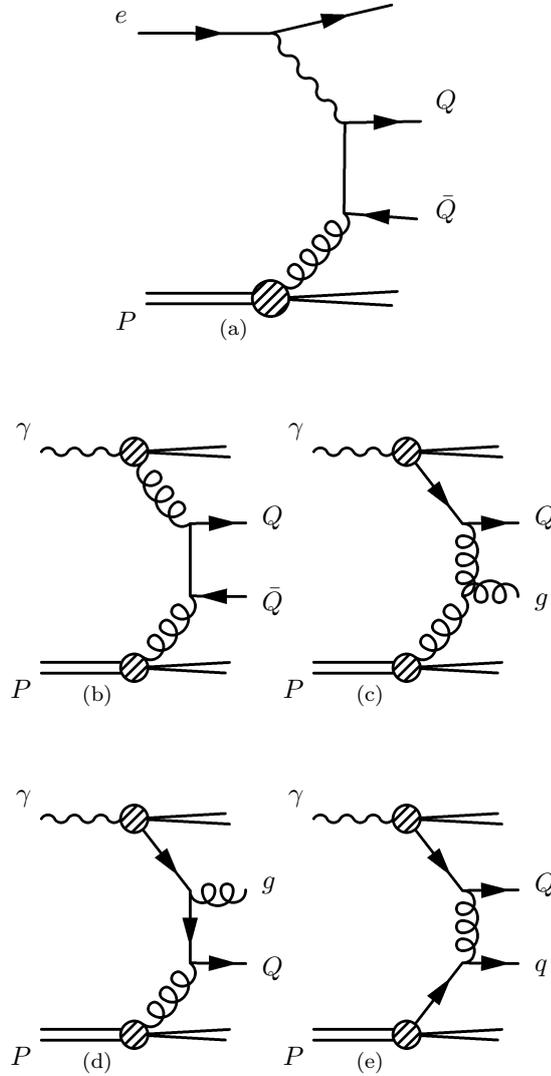


Figure 2.5: Heavy Quark Photoproduction Processes in LO QCD

Direct process (a) boson gluon fusion (dominant process for b-quark production in γp), resolved processes (b)-(e), whereas (b) is designated as hadron-like and (c)-(e) as heavy flavor excitation (highly suppressed for beauty production).

adjusted with the parameter ϵ_Q .

Other fragmentation function parameterizations are the Kartvelishvili [30] and the Bowler [31] fragmentation functions.

2.7.2 Weizsäcker-Williams distribution

In order to calculate the $q\bar{q}$ cross section in ep-collisions, the photoproduction cross section given in the equations (2.14-2.15) must be convoluted with the incoming photon flux $f_\gamma^{(e)}$. The incoming electron radiates quasi-real ($Q^2 < 1 \text{ GeV}^2$) photons with an energy fraction⁴ y according to the Weizsäcker-Williams approximation [32, 33].

Photoproduction events are experimentally defined by means of anti-tagging scattered beam electrons: all those events in which the electron is scattered at an angle larger than ϑ_c are rejected. The form of the Weizsäcker-Williams approximation in the presence of an anti-tag condition is given by [34]:

$$f_\gamma^{(e)}(y) = \frac{\alpha_{em}}{2\pi} \left\{ 2(1-y) \left[\frac{m_e^2 y}{E^2(1-y)^2 \vartheta_c^2 + m_e^2 y^2} - \frac{1}{y} \right] + \frac{1 + (1-y)^2}{y} \log \frac{E^2(1-y)^2 \vartheta_c^2 + m_e^2 y^2}{m_e^2 y} \right\} \quad (2.18)$$

2.8 NLO Calculations

Perturbative calculations of heavy quark production in electron-proton collisions are available up to NLO in α_s .

In chapter 8 the measured cross section of $b\bar{b}$ photoproduction will be compared to the NLO pQCD prediction computed with the program FMNR [35]. FMNR evaluates double differential parton-level cross sections for heavy flavor photoproduction in the fixed order massive scheme. It is a cross section evaluation program which generates weighed partonic events, and should not be considered as a Monte Carlo event generator (see section 5.1.2), which provide LO event-by-event predictions for the complete hadronic final state. It is conceived as a tool to determine the NLO prediction for heavy quark cross sections, when complicated kinematical cuts are imposed on the final partons.

The pQCD calculation implemented in FMNR is based on the ansatz of the equations (2.14), (2.15) and (2.16), i.e. the cross section is split into a photon-hadron cross section (direct contribution) and a contribution in which the photon is formally treated as a hadron (resolved contribution). Neither, the direct nor the resolved components are separately independent in NLO, only the sum of both contributions has a physical interpretation. The matrix elements of the equations (2.15) and (2.16) are expanded to the order $\mathcal{O}(\alpha_{em}\alpha_s^2)$ and $\mathcal{O}(\alpha_s^3)$ respectively [36]:

$$d\hat{\sigma}_{\gamma j}(P_\gamma, xP_P, \alpha_s(\mu_R), \mu_R, \mu_F) = \alpha_{em}\alpha_s(\mu_R)d\hat{\sigma}_{\gamma j}^{(0)}(P_\gamma, xP_P) + \alpha_{em}\alpha_s^2(\mu_R)d\hat{\sigma}_{\gamma j}^{(1)}(P_\gamma, xP_P, \mu_R, \mu_F) \quad (2.19)$$

$$d\hat{\sigma}_{ij}(yP_\gamma, xP_P, \alpha_s(\mu_R), \mu_R, \mu_F) = \alpha_s^2(\mu_R)d\hat{\sigma}_{ij}^{(0)}(yP_\gamma, xP_P) + \alpha_s^3(\mu_R)d\hat{\sigma}_{ij}^{(1)}(yP_\gamma, xP_P, \mu_R, \mu_F) \quad (2.20)$$

The partonic subprocesses relevant for the computation of the direct matrix elements $d\hat{\sigma}_{\gamma j}^{(0)}$, $d\hat{\sigma}_{\gamma j}^{(1)}$ are [36]

⁴The variable y is interpreted as the longitudinal momentum fraction of the incoming electron carried by the photon.

$$\begin{aligned}
 \gamma + g &\rightarrow Q + \bar{Q} \\
 \gamma + g &\rightarrow Q + \bar{Q} + g \\
 \gamma + q &\rightarrow Q + \bar{Q} + q \quad ,
 \end{aligned}$$

whereas the resolved matrix elements $d\hat{\sigma}_{ij}^{(0)}$, $d\hat{\sigma}_{ij}^{(1)}$ are based on the partonic subprocesses [37]

$$\begin{aligned}
 g + g &\rightarrow Q + \bar{Q} & g + g &\rightarrow Q + \bar{Q} + g \\
 q + \bar{q} &\rightarrow Q + \bar{Q} & q + \bar{q} &\rightarrow Q + \bar{Q} + g \\
 g + q &\rightarrow Q + \bar{Q} + q & g + \bar{q} &\rightarrow Q + \bar{Q} + \bar{q} \quad .
 \end{aligned}$$

Q and \bar{Q} denote the produced heavy quark pair, q a light quark, γ a photon and g a gluon.

The result of FMNR depends on the chosen PDF parameterizing the proton and the photon structure, the values for the heavy quark masses and the chosen renormalization and factorization scale. In FMNR calculations usually the renormalization and factorization scale are chosen to be equal $\mu_R = \mu_F = \mu_0$, with $\mu_0 = \sqrt{p_T^2 + m_Q^2}/2$, where p_T is the average transverse momentum of the heavy quark pair $Q\bar{Q}$, and m_Q the heavy quark mass. To evaluate the corresponding uncertainty on the result, the scale μ_0 and the heavy quark mass m_Q have to be varied within a certain window around the nominal settings. The uncertainty due to the PDF parameterization is estimated by deriving the result with different PDF sets, however is usually not the dominant contribution.

Chapter 3

The H1 Detector at HERA

3.1 The HERA Accelerator at DESY

The HERA (Hadron-Elektron Ring Anlage) collider at DESY (Deutsches Elektronen Synchrotron) in Hamburg consists of an electron¹ and a proton accelerator, built in a 6.3 km long ring tunnel. HERA operated from 1992 until 2007. Figure 3.1 shows a schematic view of HERA and its preaccelerators. Protons were accelerated counterclockwise to an energy of $E_p \simeq 920$ GeV, electrons clockwise to an energy of $E_e \simeq 27.6$ GeV. In the middle of the four straight parts of the accelerator, where the beams were brought to collision, four experiments were situated. Inside the multipurpose experiments H1 and ZEUS the two beams collided at a center of mass energy of $\sqrt{s} \simeq 318$ GeV, with $s = 4E_e E_p$. In HERMES the polarized electron beam collided with a polarized gas target to investigate the spin structure of the nucleons. HERA B was already dismantled in 2001, its purpose was to study CP violation in nucleon-proton interactions by colliding the proton beam with fixed target wires. HERA had 220 buckets, each separated by 96 ns. 180 of these buckets were filled with particle bunches, each containing on the order of 10^{10} particles². The electron bunches had a length of ~ 8 mm and were much shorter than the proton bunches that had an elongation of ~ 30 cm.

After a first data taking period from 1992-2000 (HERA I) the HERA accelerator was shut-down and underwent an upgrade program to increase the luminosity of the accelerator from $1.4 \cdot 10^{31} \text{ cm}^{-2} \text{ s}^{-1}$ to $7.5 \cdot 10^{31} \text{ cm}^{-2} \text{ s}^{-1}$ [38]. Among other things this involved the installation of additional focusing magnets inside the experiments H1 and ZEUS. After the successful modifications HERA operated from 2004 until summer 2007 (HERA II). For the last three months the energy of the proton beam was reduced to measure the structure function F_L [39] (see equation 2.8). Afterwards the data taking was stopped and the remaining three experiments dismantled.

¹The electron accelerator could be operated with electrons or positrons. The term 'electron' is used in the following to denote both lepton types.

²The beam profiles also contained side bunches, so called satellite bunches, that occurred with a time offset of $\sim \pm 4.8$ ns to the main bunch. Satellite bunches arose due to the compression of the main bunches. In addition there were pilot bunches that had no counterpart in the other beam to collide. Pilot bunches were important to study the background of non ep-interactions.

3.2 The H1 Detector at HERA

The H1 detector was a classical multi purpose detector designed for the measurement of electron-proton (ep) interactions. Figure 3.2 presents a technical drawing of the detector after the luminosity upgrade. The electron and proton beams were focused with the help of magnetic lenses and collided in the center of the detector. The nominal interaction point is also the origin of the H1 coordinate system that has its z-axis along the proton beam, its x-axis points from the origin to the middle of the accelerator ring and the y-axis points upwards. The polar angle ϑ is defined between the trajectory and the z-axis and the azimuthal angle φ in the xy-plane, whereas $\varphi = 0^\circ$ corresponds to the x-axis.

The H1 detector was arranged like a cylindrical onion, different layers of subdetector stopped and measured the different particles produced, to reconstruct a picture of events at the point of the collisions: The tracks of charged decay products were measured in the silicon detectors enclosing the interaction point (IP) to identify decay vertices from long lived particles, multi wire proportional chambers and the drift chambers for the reconstruction of tracks. The tracking detectors were enclosed by calorimeters that measured the energy of electrons, photons and hadrons. In the central and forward region the calorimeter consisted of a liquid argon calorimeter (LAr), that was split in an electromagnetic part and a hadronic part. In the backward part the (SpaCal) calorimeter was optimized to detect the scattered beam electron and was composed of scintillating fibers as the active material and lead as absorber. The tracking and calorimeter detectors were surrounded by a superconducting coil, flooding the inner detectors with a magnetic fields of 1.15 T. The charged decay products bent in the magnetic field on a helix, which allowed the determination of their transverse momenta from the curvature of the track. Muons were detected in the muon detectors implemented in the return yoke of the solenoid.

Table 3.1 contains a list of the main detector components of H1. In the following only the components most relevant to this analysis are briefly explained further. A more detailed description of the detector is found in [40] and [41].

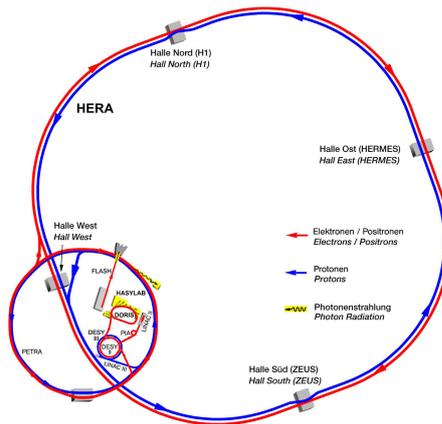


Figure 3.1: The HERA Collider

Schematic view of the HERA collider and its preaccelerators at DESY.

Number	Name	Abbreviation
1	Nominal I nteraction P oint	IP
2	C entral S ilicon T racker	CST
3	B ackward S ilicon T racker	BST
4	F orward S ilicon T racker	FST
5	C entral I nnner P roportional chamber	CIP
6	C entral O uter Z -chamber	COZ
7	Inner C entral J et C hamber	CJC1
8	Outer C entral J et C hamber	CJC2
9	F orward T racker	FT
10	B ackward P roportional C hamber	BPC
11	Liquid argon vessel	
12	Electromagnetic L iquid A rgon calorimeter	LAr elm.
13	Hadronic L iquid A rgon calorimeter	LAr had.
14	Liquid argon cryostat	
15	Electromagnetic S paghetti C alorimeter	SpaCal elm.
16	Hadronic S paghetti C alorimeter	SpaCal had.
17	P lug calorimeter	Plug
18	Superconducting solenoid	
19	Return yoke including C entral M uon D etector	CMD
20	F orward M uon D etector	FMD
21	Focusing magnets	GO / GG
22	Concrete shielding	

Table 3.1: Legend to Figure 3.2

The Table contains the main subdetectors of the H1, in the configuration after the luminosity upgrade. The indicated numbers refer to the corresponding labels in Figure 3.2.

3.2.1 Central Tracking System

H1's central tracking system consisted of silicon detectors (CST, FST, BST), measuring displaced vertices, the Central Inner Proportional chamber (CIP), whose signals were mainly used for triggering, and two concentric Central Jet Chambers (CJC1 and CJC2), that were intersected by the Central Outer z-Chamber (COZ). The CJC measured tracks precisely in $r\varphi$ but with a poorer resolution in the z -direction. The purpose of the COZ was to determine the exact z -position. The main tracking detector of H1 was the CJC, some aspects of it are described in the following.

The CJC consisted of two separated drift chambers, split in an inner (CJC1) and an outer (CJC2) ring. Both chambers had an active length of 2.2 m. CJC1 encircled the CIP at a starting radius of 20.3 cm and ended at 42.6 cm. The inner (and outer) radius of CJC2 were 53 cm (84.4 cm). Charged particles passing through the chambers ionized a special gas composition filled into the chambers. Anode sense wires were span in parallel to the beam line with two adjacent cathode planes (also made out of wires) shaping the drift field. The volume between two cathode planes formed a cell. CJC1 (CJC2) was split in 30 (60) cells that were tilted with regard to the radial direction, such that in presence of the magnetic field, the ionization electrons drifted approximately perpendicular to stiff, high momentum tracks originating from the center.

To determine the position of hits in the transverse plane, the drift time of the ionization charge to the signal wires was measured. The combination of this measurement with the precise knowledge of the wire positions, the drift velocity and the magnetic field determined the position in $r\varphi$. The position along the wire (z coordinate) was determined from the ratio of the charges read out from both wire ends. The single hit resolution of the CJC was about $160\ \mu\text{m}$ in $r\varphi$ and between 2 cm and 8 cm in z [42].

Analogue signals of the CJC were the input to the Fast Track Trigger (presented in section 3.2.5), that derived trigger decisions for the first three trigger levels of the H1 trigger system.

3.2.2 The Liquid Argon Calorimeter

The liquid argon calorimeter (LAr) [43] covered the central and forward regions of the detector. It contained two sections, an inner electromagnetic section (ECAL) optimized for the detection of electrons and photons, and an outer hadronic section that measured the energy deposition of hadrons (HCAL).

The LAr calorimeter was a non-compensating³ sampling calorimeter, composed of alternating absorber layers and liquid argon filled gaps. The absorber in the ECAL were lead plates, whereas in the HCAL the absorber was made out of steel plates. The thickness of the ECAL was 20-30 radiation lengths, that of the HCAL corresponded to 4.7-7 interaction lengths. The calorimeter contained eight sections along the z -axis, called wheels or barrels. Each wheel had its own name starting from backward barrel (BB), to three central barrels (CB1-CB3), followed by two forward barrels (FB1, FB2), and the inner and outer forward barrels (IF, OF) (see Figure 3.3). From the CB1 up to the FB2 the wheels had an eightfold geometry in the transverse plane. Inactive regions occurred in the gaps between the wheels and between two octants, and are referred to z - respectively φ -cracks. The orientation of the absorber plates depended on the z -position of the wheel, and was chosen such that the angle of incidence of a particle from the interaction point was not smaller than 45 degrees. The LAr calorimeter was highly-segmented into 45000 read-out cells as shown in Figure 3.3.

³The average ratio between signals from electromagnetic and hadronic particles of the same incident energy is for non-compensating calorimeters larger than 1. The response for electromagnetic particles is higher.

The basic granularity followed from the criteria to separate electromagnetic and hadronic particle showers in the transverse and longitudinal plane, while also optimizing the capacities of the single cells. The entire calorimeter was hermetically enclosed by a cryostat, to cool the argon down to its liquid phase.

The resolution in the electromagnetic part obtained in test beam measurements [44, 45] was $\sigma/E \approx 11\% \sqrt{E/\text{GeV}} \oplus 1\%$, and in the hadronic part $\sigma/E \approx 50\% \sqrt{E/\text{GeV}} \oplus 2\%$.

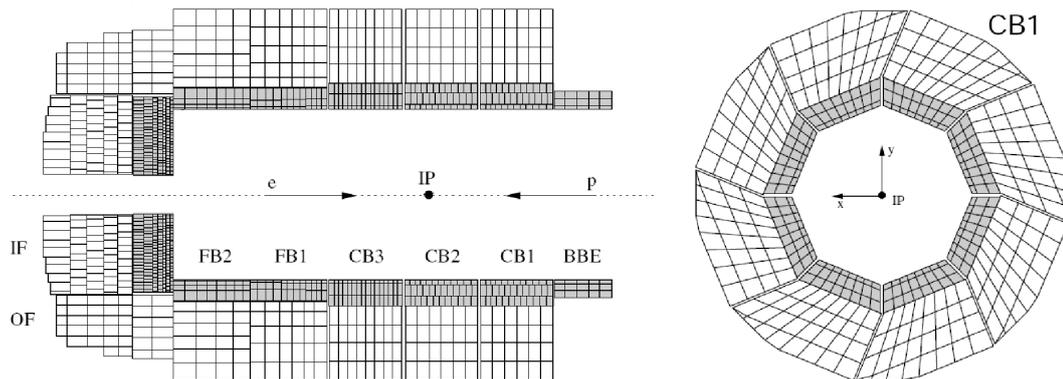


Figure 3.3: Segmentation of the LAr Calorimeter

Longitudinal (left) and transversal (right) cut through the Liquid Argon calorimeter illustrating the fine granularity of the readout-cells. The shaded area indicates the electromagnetic section of the calorimeter.

3.2.3 Luminosity Measurement

H1 measured the luminosity delivered by HERA by counting the rate of the Bethe-Heitler Process [46], $ep \rightarrow ep\gamma$. The Bethe-Heitler Process is theoretically well understood, and within quantum electrodynamics (QED) its cross section is precisely calculable.

Bethe-Heitler events were measured by the detection of the photons in the Photondetector situated at $z = -101.8$ m [47], [48]. The photon detector consisted of a sampling calorimeter composed of quartz-fiber and tungsten as absorber. The scattered electrons were not hitting the photon detector, but were deflected by the magnetic field of HERA.

The precision of the luminosity measurement was on the order of 2.5%, except for the last running period of the experiment during which problems occurred and the uncertainty on the luminosity determination raised to 5%. The observed problems are yet not fully understood, and are presently being investigated, as well as an alternative luminosity measurement via Compton scattering events.

3.2.4 The H1 Trigger System

The bunch crossing rate at HERA was 10.4 MHz, while the interesting ep-interaction rate was on the order of 1 kHz. During the H1 detector readout, that lasted on average 1.4 ms, the detector collected the information generated by a previous event and became insensitive to new events. The insensitive time of the detector is called dead-time. The readout frequency

was restricted to ~ 50 Hz, in order to keep the dead-time below 10%. Events were only written to tape and saved permanently with a rate of ~ 10 – 20 Hz.

The rate reduction by a factor of one million, from ~ 10 MHz to ~ 10 – 20 Hz, was performed by the H1 trigger system. The H1 trigger system was implemented as a four level trigger system that selected physics events of interest, while successively reducing background and events of low interest. On each trigger level the input rate was reduced further, while the available time to derive the trigger decision rose, and therefore allowed more complex event analyses.

Trigger Level 1, L1 The first trigger level was implemented dead time free, i.e. the data taking was not stopped during the time the L1 decision was formed. This was achieved by storing the data in temporary buffers (pipelines) during the L1 latency time of $2.3 \mu\text{s}$ (which corresponds to 24 bunch crossings (BC) ~ 24 BC). Within this time the subsystems built 256 trigger elements (TE), that were combined in the central trigger logic (CTL) to 128 subtrigger elements. If at least one subtrigger element condition was fulfilled, the pipelines were stopped by the CTL and the dead time started to accumulate. The rates of the L1 subtriggers depended substantially on the luminosity of the accelerator and on the background conditions. However, in order not to overstrain the subsequent trigger levels and the data acquisition with too high event rates, and on the other hand in order to fully exploit the available bandwidth, the L1 rate had to stay in a narrow band. This was achieved by adjusting the L1 rates with prescale factors. Prescale factors are integer numbers that were computed on a regular basis⁴ according to a predefined strategy. A subtrigger with a positive L1 condition got a raw trigger bit assigned. If the subtrigger had an assigned prescale factor N that was larger than one, it only triggered the event every N^{th} time it fired. It then got in addition to the raw trigger bit an actual trigger bit. A period with stable prescale conditions was called a run. The prescale counter from which the actual trigger bits were determined started at zero at run start. If the L1 conditions of two subtriggers were different, shortly after the run had started the two subtriggers became independent. However, two subtriggers with equal L1 conditions were correlated, even if the prescale factors were different. (Compare with section 7.1.2 where this discussion becomes relevant for the combination of subtriggers.)

Trigger Level 2, L2 After a positive L1 trigger decision the second trigger level started. The time available to obtain the L2 trigger decision was $20 \mu\text{s}$. The L2 decision was derived by combining L1 information in the Topological Trigger [49] and the Neural Network Trigger [50, 51]. The third system contributing was the Fast Track Trigger (FTT), providing high resolved track information (compare with section 3.2.5). The three systems sent 96 TE to the CTL, that were again combined to subtrigger elements. Typically a L2 subtrigger element validated a L1 subtrigger element. Many L1 subtriggers however, had no L2 condition and were validated by default. If at least one L1 subtrigger was validated on L2, the readout of the detector started. On the other hand, if a L1 decision was rejected by the CTL, the readout pipelines were restarted and the data taking continued.

Trigger Level 3, L3 Events with a positive L2 trigger decision were partially reconstructed on the third trigger level within about $100 \mu\text{s}$. If for a subtrigger with a L3 condition no verification was derived within the mentioned time window, the complete readout of the event was aborted, the pipelines cleared and the data taking restarted.

Trigger Level 4, L4 After a completed detector readout the data events were passed to a PC computer farm and reconstructed with the H1 reconstruction program (hlrec).

⁴The prescale factors were adjusted by an hour at latest.

Based on the completely reconstructed events software filters (L4 trigger) further reduced the rate, before the remaining events of interest were permanently saved for the offline analyses. For monitor purposes a small part of the events classified with a negative L4 trigger decision were stored as well, together with a corresponding event weight.

3.2.5 Trigger Upgrades

With the HERA luminosity upgrade the contributions to the trigger rate of ep-physics increased by roughly a factor 5. Of more importance for the trigger system was, that also the background rates increased substantially. To cope with this new situation, H1 decided to upgrade the existing trigger system with new subtriggers providing higher selectivity. The main upgrade projects were the CIP2000 trigger [52–55], the Fast Track Trigger (FTT) [56–61] and the Jet Trigger (JT) [62,63]. CIP2000 was a proportional chamber that was mainly used to reject non-ep background. The basic idea of the JT was to search for localized energy depositions in the LAr calorimeter. The concept of the FTT L1 and L2 was to reconstruct on the first two trigger levels highly resolved charged particle tracks in the CJs. On the third trigger level the FTT partially reconstructed final states, based on the tracks measured on the FTT L2 and on information of other systems (in particular also of the JT). The triggers used in this analysis are FTT L3 triggers that combine on the third trigger level the rich track information of the FTT L2 with the energy depositions measured by the JT on L1. These triggers and aspects of FTT L3 are discussed in Appendix D. The basic essentials of the FTT L1 and L2 and the JT are briefly summarized in the following.

The Fast Track Trigger on the First Two Trigger Levels

On the first two trigger levels the Fast Track Trigger (FTT) derived trigger decisions, based on FTT tracks in the two Jet Chambers CJC1 and CJC2. The FTT was capable to recognize up to 48 tracks. During the $2.3\ \mu\text{s}$ latency of the first trigger level the FTT reconstructed the curvature $\kappa = 1/p_T$ and the azimuthal angle φ of tracks. During the $20\ \mu\text{s}$ latency of the second trigger level three dimensional tracks were reconstructed, almost reaching offline resolution.

Out of 56 wire layers of the CJC only 12 wire layers grouped in four trigger layers delivered the input to the FTT (see Figure 3.4). The analogue signals from both ends of in total 450 wires of the CJs were continuously digitized with 80 MHz Fast Analogue Digital Converters (FADC).

Wire hits were identified by a Qt-algorithm and afterwards filled into shift registers operating at 80 MHz. To reduce the bandwidth at L1 an effective sampling rate of only 20 MHz was used, while the L2 processing was based on the full information. For each trigger layer the stored hit pattern were compared on L1 to 3072 precalculated characteristic masks, each containing the hypothetical hit pattern of a track originating from the vertex. These characteristic masks were precomputed by a simulation running on a standard PC. The output of a recognized pattern was the curvature κ and the azimuthal angle φ of track segments located on the corresponding trigger layers. In order to suppress falsely recognized patterns, the track segments were combined to tracks. Tracks were identified, if at least two out of potentially four track segments on different trigger layers had similar values in κ and φ .

Five trigger elements (TE) were sent to the CTL for the number of tracks above a certain p_T threshold, with the lowest being $p_T = 100\ \text{MeV}$. In addition TE existed for topological criteria, and the sum over the charge of all charged particles in the event.

The longer latency on the second trigger level allowed a more accurate track segment search, using the higher resolved hit patterns of the 80 MHz shift registers. Track segments on dif-

ferent trigger layers were again combined to tracks by requiring similar κ and φ values. Afterwards the parameters were refined by two fits. The segment coordinates in the transverse $r\varphi$ -plane were fitted by a circle containing the primary vertex. For each track segment the z positions were measured as well, and fitted in the longitudinal rz -plane by a straight line passing through the z -vertex determined by the FTT. The result of these fits were precisely measured tracks in the variables κ , φ and ϑ . The resolution of the FTT L2 tracks reached almost offline resolution: $\sigma_{1/p_T} = 2.2\%/ \text{GeV}$, $\sigma_\varphi = 2.4 \text{ mrad}$ and $\sigma_\vartheta = 50 \text{ mrad}$ [61]. The L2 trigger elements essentially verified the L1 trigger elements, i.e. TE for the number of tracks above a certain p_T -threshold and for topological criteria. In addition to the L1 TE, also TE for more sophisticated applications, like the invariant mass of two-body decays and the z -vertex, existed [42].

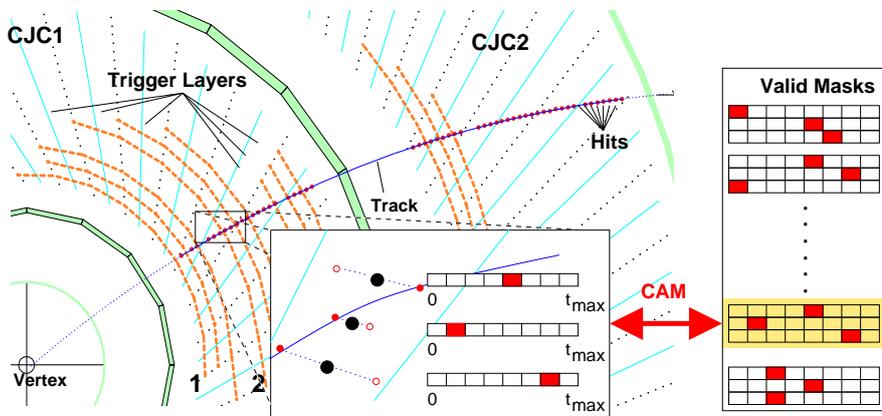


Figure 3.4: Illustration to the FTT Track Finding.

Radial view of the H1 central tracker, illustrating the principles of the FTT: A charged particle is passing from the interaction point (vertex) through CJC1 and CJC2 and ionizing the tracker gas. The analogue signals of selected wires (grouped on four trigger layers) were the input to the FTT. The digitized signal hits were filled into shift registers, and compared to precalculated valid masks to identify track segments. (The comparison was done by a Content Addressable Memory (CAM)). (The Figure is taken from [42].)

The Jet Trigger

The basic idea of the Jet Trigger (JT) was to search for localized energy depositions in the LAr calorimeter within a typical cone radius of $\Delta r = \sqrt{(\Delta\eta)^2 + (\Delta\varphi)^2} \approx 1$, with η being the pseudo rapidity $\eta = \ln(\tan(\vartheta/2))$. The JT derived trigger decisions on the first trigger level, and sent its results also to FTT L3.

The input to the JT were signals of the fine-granularity trigger towers (TT), also available in the LAr calorimeter trigger [64]. A TT was a group of LAr cells that have an angular coverage of typically 0.2 radians. Within the JT the energy of a TT was formed from their analogue pulses that were sent to the JT from the electromagnetic and the hadronic section of the LAr calorimeter. These pulses were first digitized, before thresholds and weights were applied. Afterwards the two sections were added together and mapped to the so called input

towers (IT). The mapping differed for the forward and the backward region. In the forward region the finest possible granularity of the TTs was preserved, whereas in the backward region several TTs were combined, as illustrated in Figure 3.5.

With the help of highly sophisticated and time optimized algorithms, nearest neighbor ITs around a local transverse energy maximum were summed up, to define a local energy deposition. The measured energy depositions were ordered by energy and the 16 highest energy depositions stored in a list. The list contained for every energy deposition the transverse energy and its position in the $\vartheta\varphi$ -plane.

This list was on the one hand side placed at the disposal for the FTT L3, and on the other hand side used internally to form L1 subtrigger elements.

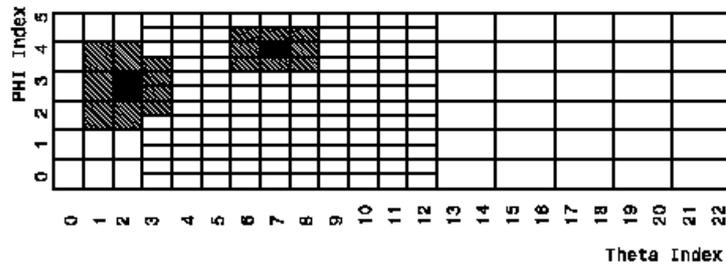


Figure 3.5: Granularity of the JT

Granularity of the JT in the $\vartheta\varphi$ -plane shown for 3 contiguous of 8 φ -octants. The ϑ index 0 (22) corresponds to the most forward (backward) region in the calorimeter. The shaded areas indicate the summing of the nearest neighbors around the central (dark shade) input tower [63].

Chapter 4

Aim and Concept of the Measurement

4.1 Motivation and Aim of the Measurement

Perturbative Quantum Chromodynamics (pQCD) based theoretical predictions of beauty production in electron proton collisions have shown a tendency to underestimate previous measurements of beauty production at HERA. The discrepancies occurred especially at small scales, namely at small values of the photon momentum transfer Q^2 and at small values of the transverse momentum of the beauty quark¹ $p_T(b)$.

The aim of this analysis is to measure the differential beauty cross section at HERA in photoproduction ($Q^2 < 1 \text{ GeV}^2$) as function of the mean transverse momentum of the beauty quark²,

$$\frac{d\sigma}{d\langle p_T(b) \rangle} . \quad (4.1)$$

The kinematical phase space to which the result of this measurement is extrapolated as described in chapter 8, is chosen similar³ to the kinematical phase space of the common $\frac{d\sigma}{d\langle p_T(b) \rangle}$ summary plot⁴ combining different beauty measurements at HERA, see [65] and the references therein. The corresponding analysis cuts are listed in Table 4.1. Special attention is given to include the production threshold in the measurement and to extend the experimental visible phase space of the previous measurements down to $p_T(b) \approx 0 \text{ GeV}$.

4.2 Concepts of Previous Beauty Measurements at HERA

In ep collisions at HERA beauty quarks are mainly produced in pairs via the fusion of a quasi-real photon emitted by the incoming electron and a gluon of the proton. Traditionally, beauty measurements at HERA tagged the beauty quarks with the help of one or two reconstructed jets and one or two additional beauty decay leptons. A second class of analyses tag the beauty quarks via the reconstruction of two muons or a muon and a D^* , utilizing

¹The transverse momentum is defined as the momentum in the plane transverse to the beam axis.

² $\langle p_T(b) \rangle$ is the average transverse momentum of the produced $b\bar{b}$ -pair, $\langle p_T(b) \rangle = \sqrt{(p_{T,b}^2 + p_{T,\bar{b}}^2)}/2$.

³The only difference is a slightly different cut on the pseudo rapidity of the beauty quarks, and is discussed in section 8.6.

⁴The mentioned summary plot is shown in Figure 8.16, including the results of this measurement.

the semileptonic decay channel $b \rightarrow lX$ and the decay channel $b \rightarrow D^*X$, respectively. Jet algorithms are a suitable tool for $p_T(b)$ values of 6–7 GeV or above, but for lower values the convergence of jet algorithms becomes problematic due to threshold effects. Often threshold effects are also present with the lepton tag analyses. Muons are preferably identified in the outermost part of the H1 and ZEUS detectors, which involves a cut-off on the transverse momentum of the reconstructed muons of $p_T(\mu) \sim 2$ GeV and thereby also on the transverse momentum of the beauty quarks $p_T(b)$.

4.3 Concept of this Measurement

In this analysis a new concept is chosen: The differential beauty cross section is measured from electron pairs only, exploiting the decays $b\bar{b} \rightarrow eeX$. Jet algorithms are avoided in order not to introduce related cut-off effects⁵. Electrons from the decay channel $b\bar{b} \rightarrow eeX$ either result directly from a semileptonic beauty quark decay, or indirectly from a beauty quark decaying into a charm quark, which then decays semileptonically into an electron. The Feynman Diagrams for the different possibilities are shown in Figures 4.1(a)-(d).

Electrons have the advantage over muons that the reconstruction cut-off on its transverse momentum is lower. The relation of the transverse momentum of the beauty quark $p_T(b)$ and the transverse momentum of the decay lepton $p_T(l)$ is illustrated in the Figures 4.2 and 4.3, obtained from a simulation⁶. All plots are based on the decay channel $b\bar{b} \rightarrow eeX$ in the phase space defined in Table 4.1. Figure 4.2 (a) and (b) present the corresponding differential cross section from the direct and resolved (hatched) contribution (see chapter 2.7) as function of $\langle p_T(b) \rangle$ and $\langle p_T(e) \rangle$, the mean transverse momentum of the decay electrons⁷. Figure 4.3 shows as function of $p_T(b)$ the reconstruction acceptance for different cuts on the lower value of $p_T(e)$ of both electrons. Exemplary the cut on $p_T(e)$ of this analysis is indicated in Figure 4.2 (b) by the black line, which corresponds to the dashed plotted acceptance of Figure 4.3. The Figures clearly show that lowering the cut on the transverse electron momentum $p_T(e)$ improves not only the absolute value of the acceptance, but also makes the low $p_T(b)$ phase space experimentally accessible.

kinematical cuts
$0.2 < y < 0.8$
$Q^2 < 1 \text{ GeV}^2$
$ \eta(b), \eta(\bar{b}) < 2$

Table 4.1: Kinematical Cuts

The kinematical cuts on the inelasticity y , the photon momentum transfer Q^2 and the rapidity of the beauty quarks $\eta(b), \eta(\bar{b})$. These cuts define the kinematical phase space to which the final result is extrapolated.

⁵A similar approach has been followed by the ZEUS collaboration in a recent measurement, utilizing the decay channel $b\bar{b} \rightarrow \mu\mu X$ [66].

⁶The used simulation program is Pythia, discussed in chapter 5.1.2.

⁷Both beauty decay electrons are filled into the histogram, weighed by 0.5.

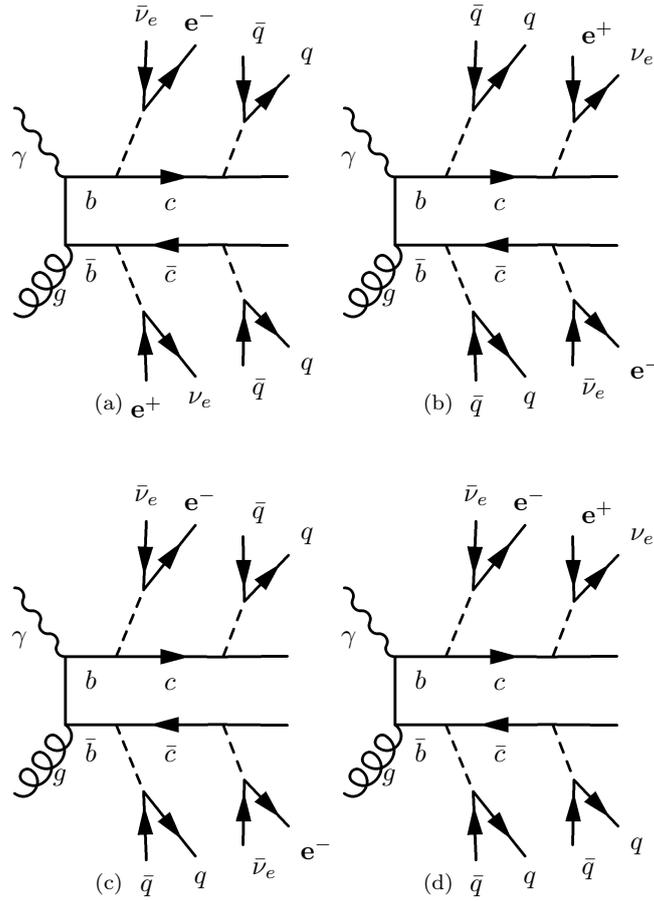


Figure 4.1: Feynman Diagrams for $b\bar{b} \rightarrow eeX$ Decays.

Feynman diagrams for $b\bar{b}$ pairs produced in boson gluon fusion and decaying in two electrons. The electron pair can either originate from the $b\bar{b}$ pair (a), the $c\bar{c}$ pair (b), one b and one c quark with the c quark having the same (c) or different (d) b quark mother as the first electron.

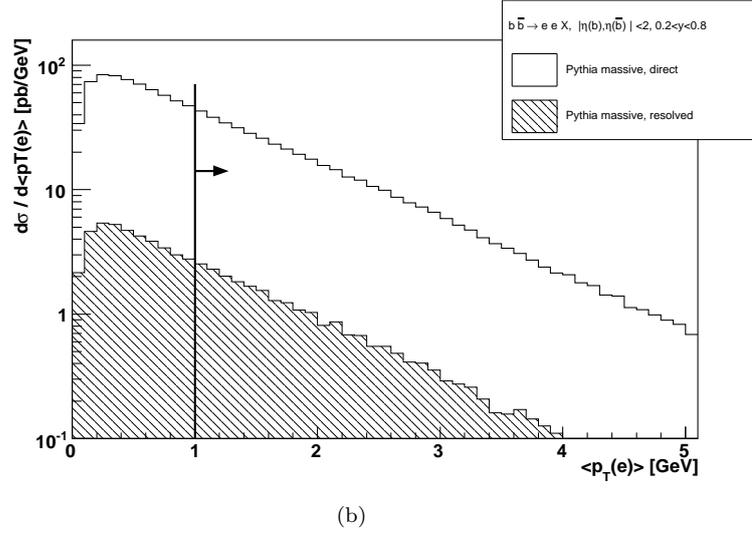
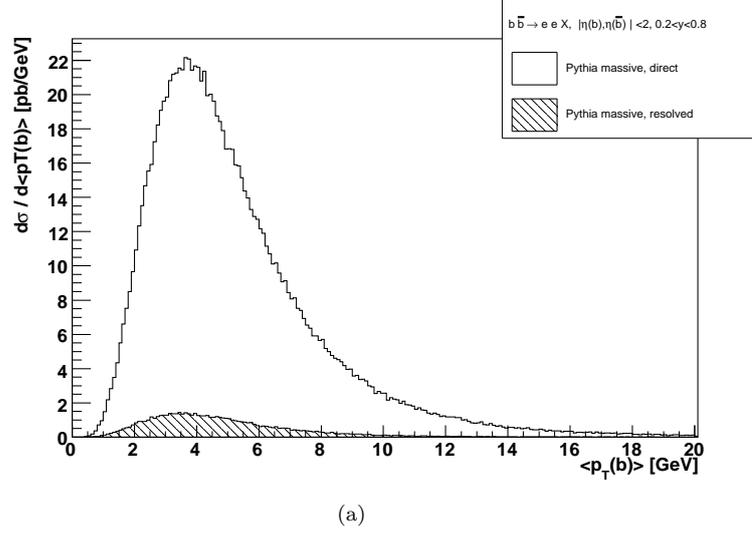


Figure 4.2: Simulated Differential Beauty Cross Sections

Differential beauty cross sections for the kinematical phase space of this analysis (see Table 4.1) obtained from simulation (Pythia), as function of the mean transverse momentum of the beauty quarks (a) and the mean transverse momentum of the two electrons (b). The vertical black line indicates the acceptance cut of the electron identification. The open part of the histograms shows the contribution from the direct matrix element, while the shaded part shows the contributions from the resolved matrix elements, cf. section 2.7.

4.3.1 Conclusions on the Concept of this Measurement

The advantage of using low- p_T -electrons is that beauty pairs become measurable at production threshold, i.e. down at $\langle p_T(b) \rangle \approx 0$ GeV. Consequently the aim to lower the $p_T(e)$ threshold on trigger level (chapter 7.1.1 and D.2.3) and for the offline electron identification (chapter 6) was followed within the whole thesis. The disadvantage is that due to the avoidance of jet algorithms the kinematical properties of the beauty quarks are only accessible indirectly.

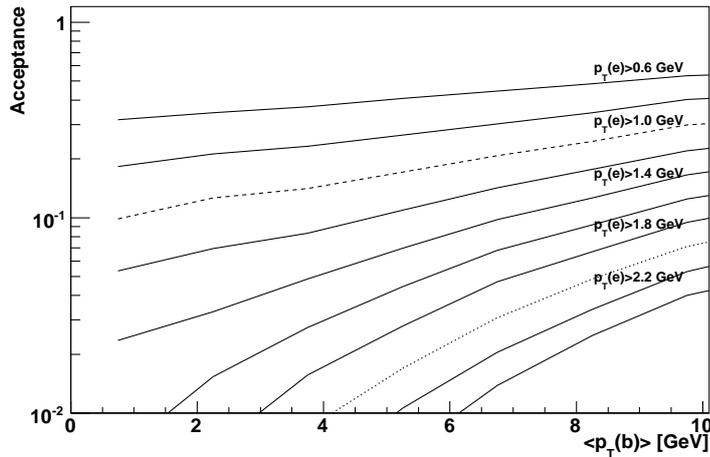


Figure 4.3: $p_T(e)$ -Acceptance

Acceptance for $ep \rightarrow b\bar{b} \rightarrow llX$ events determined from simulation (Pythia) in the kinematical phase space of this analysis (see Table 4.1), as function of the mean transverse momentum of the beauty quarks. The leptons also may originate from a c -quark having a b -quark mother (see Figure 4.1). Events containing only charm ($ep \rightarrow c\bar{c} \rightarrow llX$, $ep \rightarrow J/\psi X \rightarrow llX$) are treated as background (see chapter 8).

Shown are different lines for different cuts on the $p_T(e)$ of both electrons. In this analysis electrons are reconstructed above $p_T(e) > 1$ GeV, indicated in the plot by the dashed line. Muons analyses usually have a lower cut-off on the transverse momentum of the muon of $\sim p_T(\mu) > 2$ GeV (dotted line).

Chapter 5

Monte Carlo Simulations

5.1 Simulation of Events

The simulation of events in the H1 experiment is split into two main parts: In the first part a so called event generator program is executed that contains the physics under study. The second part simulates the detector response to the final state particles of these generated events.

5.1.1 Event Generator

An event generator is a Monte Carlo program that simulates on a event-by-event basis the interaction of high energetic incoming particles. The output of an event generator are for every event a set of four momenta of all simulated final state particles, as well as information concerning their production.

Based on the fact that the full process can be approximated by factorization into individual problems, these individual processes are calculated separately. The starting point of an event generator are matrix elements of the initial hard interaction, usually calculated in a perturbative ansatz of the strong (QCD) and electroweak force. For the simulation of ep interactions additional input distributions are required, modeling the proton and photon structure. From these distributions and the matrix elements under study four vectors of partons are simulated on the basis of random number generators. Initial and final state parton showers are added to simulate QCD radiation down to a cut-off scale. Below this cut-off scale the colored partons are transferred into colorless hadrons by the fragmentation processes. Decays of the short lived particles are performed as well, which gives rise to decay leptons and displaced secondary vertices.

5.1.2 Leading Order Monte Carlo Event Generators: Pythia and Cascade

For this analysis event samples are generated with the event generators Pythia [67,68] and Cascade [69,70]. Both event generators are based on matrix elements in leading order (LO) in α_s . However, the perturbative QCD ansatz implemented in Pythia and Cascade is different. Pythia uses a collinear factorization approach of the parton density functions, based on the DGLAP evolution equation [17–20], i.e. the partons are restricted to be on the mass-shell. In Cascade, on the other hand, the k_T -factorization approach is implemented, based on the CCFM evolution equation [23–26]. The matrix elements are calculated for partons being off-

mass-shell and are not integrated over k_T , which implies that the parton densities depend on the transverse momentum of partons, in contrast to the collinear factorization, as discussed in chapter 2.

5.1.3 Simulation of the Detector Response

For every event the event generator produces a set of four vectors of all stable final state particles. Yet, these quantities are difficult to compare to measured data due to the missing simulation of the detector response. The complete detector response to the stable particles of the generated event, the trigger simulation and the event reconstruction are added by the consecutive execution of three additional programs:

- The interaction of the generated stable particles with the detector and the subsequent detector response is simulated with the program `h1sim`, based on the GEANT 3 package [71]. Detector effects due to geometrical acceptances and inefficiencies, and the production of secondary particles are taken into account in this simulation step.
- The output of `h1sim` is fed into the H1 reconstruction program `h1rec`. This program is identically used to reconstruct measured data events.
- Due to technical reasons the trigger simulation is only partly implemented in `h1sim`. The complete trigger simulation is resimulated and the missing parts of it are added by the program `fttemu` [72, 73]¹.

The computing time used by these programs can be considerable, where the main time consumption is due to `h1sim`. Even with the computing power available today, it is often necessary to restrict the input to the detector simulation. This is usually done by cuts that confine the event generator output to the events in the experimentally accessible phase space.

5.2 Monte Carlo Samples

The different Monte Carlo Samples used in this analysis are presented in the following, and are summarized in Table 5.1. Cuts on the output of the event generator to restrict the experimentally visible phase space were necessary for two background samples, in order to reduce the large number of events in the sample. However, no such cuts are applied on the event generator output of the signal samples.

The following **signal** MC samples were produced:

Beauty Samples Two Monte Carlo samples were generated simulating beauty production, one based on the event generator Pythia and the other one on Cascade.

Pythia is run for these samples in the massive mode [68], i.e. using massive heavy quarks in the matrix elements (cf. section 2.7). Samples for the direct and the resolved contributions were generated, which are added luminosity weighted.

For Cascade the distinction of a resolved and direct contribution is not necessary.

The measured differential cross section will be compared in shape and normalization to the predictions of these two generators in chapter 8.

The following **background** MC samples were produced:

J/ψ Sample To estimate the $J/\psi \rightarrow ee$ background, an event sample of inelastic $J/\psi \rightarrow X$ has been simulated with the Cascade event generator. These events are not well

¹The trigger simulation of the third trigger level and the calibration of the Jet Trigger simulation, and its implementation into the H1 software framework were part of this thesis. Both subjects are summarized in the Appendix.

described by the Pythia event generator [75], therefore Pythia is not considered for this decay channel.

Charm Sample The open charm background is estimated from a sample generated with Cascade. A generated event was only passed to the detector response simulation, if at least two charged particles have a transverse momentum above 0.7 GeV. This cut is slightly looser than a corresponding L2 trigger condition.

Inclusive Sample The background due to misidentified electrons is estimated from a flavor inclusive Monte Carlo sample, generated with Pythia in the massless mode. Similar as for the charm sample, the output of the event generator was restricted to the experimental accessible phase space. The following cuts, slightly looser than corresponding trigger conditions, were applied on the transverse momentum ordered charged particles: $p_{T,1} > 1.1$ GeV and $p_{T,2} > 0.7$ GeV. Even after these cuts the output of the event generator for an integrated luminosity of $\mathcal{L} = 25.7 \text{ pb}^{-1}$ was more than 69 million events large, which resulted in a detector response simulation running for several weeks.

In order also to generate low p_T -physics processes, the steering parameter for \hat{p}_\perp was set from the H1 standard of $\hat{p}_\perp = 4$ GeV to the minimum, as recommended for such samples in the Pythia manual [68]. For hard 2-2 processes \hat{p}_\perp is the transverse momentum between the two outgoing partons, defined in the rest frame of the interaction. A cut-off on it restricts the generator to the high p_T -jet production, which is not suitable for this analysis (cf. with the discussion in chapter 4).

The drawback of this setting is that in the generated sample the resolved contribution considerably overshoots the direct contribution, by roughly a factor of five². For the unfolding of the differential beauty cross section (chapter 8) the resolved contribution therefore is scaled, i.e. the relative normalization of the two contributions is adjusted. It is important to mention that, first of all the measurement does not depend on the overall normalization of this MC sample, and second the differences in the shape of the

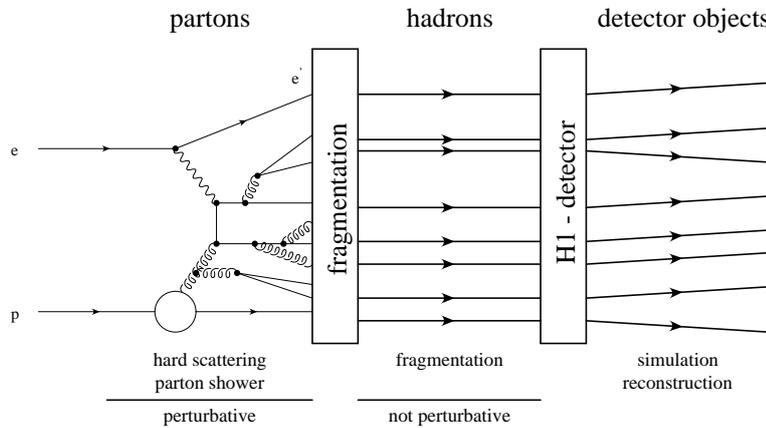


Figure 5.1: Schematic Illustration of the Event Simulation in the H1 Experiment

The simulation of the hard interaction and the parton showering is based on matrix elements calculated in perturbative QCD, QED and EW. The resulting partons are afterwards hadronised into colorless hadrons by non perturbative fragmentation functions. Finally the response of the H1 detector is simulated. (Figure is taken from [74].)

²Similar problems were reported in the past with similar MC samples [76].

used distributions due to the relative normalization are propagated to the systematic uncertainty of the final cross section (section 8.4).

The reasons contributing to this unpleasant behavior of the event generator are the followings: Several processes contribute to the total cross section, if Pythia is run in the flavor inclusive mode. Some matrix elements of considered processes diverge in the massless pQCD approach like $d\sigma/d\hat{p}_\perp \approx 1/\hat{p}_\perp^4$ for $\hat{p}_\perp \rightarrow 0$. This characteristic have in particular LO matrix elements of cross sections dominated by the t-channel³ and contributing to the resolved ep cross section [68, 77]. Within Pythia these divergences are taken into account by firstly applying process dependent cut-offs⁴, and secondly by modeling the mixture of the direct and resolved cross section by phase space dependent dumping factors [68, 78]. In the Pythia manual it is argued that this mixing is nontrivial and that with the default settings one obtains it "hopefully [!] consistently, rather than building it up by combining separate runs". The default mixture model might be suitable for the cross section in the entire phase space, however the interplay with the applied generator cuts is unclear. Therefore, the latter procedure was followed by adjusting the relative normalization of the resolved and direct contributions, as described above.

Sample	Generator	Gen. Cuts	PDF, uPDF	# Ev.[M]	\mathcal{L} [pb^{-1}]
beauty	Pythia, massive, dir.	no	CTEQ 6L / SAS 2D	15.0	2809.9
	Pythia, massive, res.	no	CTEQ 6L / SAS 2D	3.0	5080.8
	Pythia, massive, res.	no	CTEQ 6L / SAS 2D	0.1	5487.4
	Cascade	no	A0	12.5	2472.1
J/ψ	Cascade	no	A0	2.4	492.1
charm	Cascade	yes	A0	2.0	500.0
inclusive	Pythia, massless mode	yes	CTEQ 6L / SAS 2D	69.0	25.7

Table 5.1: Overview of the MC Samples

The table gives an overview on the different Monte Carlo samples considered in this analysis. The parton distribution functions (PDF) used in Pythia are documented for the proton in [79] and for the photon in [80]. Cascade is based on unintegrated parton distribution functions (uPDF) [81].

³These processes are: $qq' \rightarrow qq'$, $q\bar{q} \rightarrow q'\bar{q}'$, $q\bar{q} \rightarrow gg$, $qg \rightarrow qg$ and $gg \rightarrow q\bar{q}$.

⁴Pythia sets for the divergent processes the \hat{p}_\perp -cut-off to $\hat{p}_\perp = 1$ GeV, if the corresponding steering parameter is set to the minimum and nothing else is specified (as recommended in the manual).

Chapter 6

Electron Identification

Electrons in H1 are identified by their signature in the electromagnetic part of the calorimeter, which was specifically optimized to detect scattered high energetic electrons. The H1 standard and well established electron identification algorithm works for electrons with transverse momenta above 5 GeV, a threshold too high for this analysis, as discussed in chapter 4. Other methods to identify also electrons with lower energies exist, but are found not to be efficient and not sufficiently well described by the simulation for this analysis.

Therefore, a new electron identification for electrons with $p_T > 0.5$ GeV was developed within this thesis, and is presented in the following. An important element of this new electron identification, extending the work of [82], is a method to correct the measured energy in the calorimeter for energy losses in the dead material in front of it. This dead material depends on the energy and the position of the energy deposition in the calorimeter and is described in Appendix B.

6.1 Method

The main background contribution to an electron identification in H1 results from pions. The electron identification is based on the combination of track and calorimeter information in an Artificial Neural Network (NN) that is trained to discriminate between electrons (signal) and pions (background). Special attention is given to the following two aspects: First, to efficiently separate electrons with $p_T > 0.5$ GeV from background, and second to develop an algorithm with an identical output for data and simulation. To ensure this, all input variables to the NN are checked to have the same characteristics for the data and the simulation. Within this context a new elaborate method to correct effects due to dead material was developed, which is presented in Appendix B. Input variables to the NN for which the data is not sufficiently described by the simulation are discarded. This is in particular the case for the specific energy loss dE/dx , with which specially electrons at transverse momenta of 1 GeV or less could be efficiently discriminated from background.¹

One can regard the NN as a mapping from the multi-dimensional space of input variables onto a single output variable, designated as MLP ². MLP is a quality criterion that separates electrons ($MLP = 1$) from background ($MLP = -1$), e.g. the NN reduces the multidimensional input to a one-dimensional signal-versus-background discrimination problem.

¹The calibration and implementation of dE/dx in the simulation is foreseen for a future event reconstruction version [83].

² MLP refers to the name of the NN algorithm, cf. with section 6.4.

6.2 Acceptance and Preselection Cuts

6.2.1 Track Selection

The electron identification is based on particles for which a well reconstructed track exists. In H1, tracks are reconstructed from the hit information measured in the central- (CTD) and the forward tracker (FTD). Depending on the tracking detector in which hits are found and combined to tracks, three different track types are distinguished within the H1 software. 'Central tracks' are measured only in the CTD, 'forward tracks' only in the FTD and 'combined tracks' in both systems. For each of these track types a standard selection exists that ensures the necessary reliability of the measured track parameters. All three selections only take tracks into account that are fitted to the primary or a secondary vertex.

In addition, there are minimum requirements on the transverse momentum, the polar angle and the length of the tracks. The different selection cuts used for the analysis are summarized in Table 6.1. For a single track one can have several track type hypotheses. If this is the case, the track types are compared and the most preferred track type is chosen. If two track types are equal, the track hypothesis with the most preferred vertex type is taken [84]. Within the H1 software framework, all tracks that fulfill these selection conditions are saved after the type comparison and designated as 'H1PartSelTrack' tracks.

The electron discrimination variable developed in this analysis is calculable for H1PartSelTrack tracks with a transverse momentum of more than 0.5 GeV.

central tracks (CTD)	
transverse momentum	$p_T > 70 \text{ MeV}$
radial track length	$R_{length} \geq 10 \text{ cm}$ for $\vartheta \leq 150^\circ$ $R_{length} \geq 5 \text{ cm}$ for $\vartheta > 150^\circ$
radius of first hit	$R_{start} \leq 50 \text{ cm}$
dca'	$ dca' \leq 2 \text{ cm}$
combined tracks (CTD & FTD)	
transverse momentum	$p_T > 120 \text{ MeV}$
polar angle	$10^\circ < \vartheta < 30^\circ$
radius of first hit	$R_{start} \leq 50 \text{ cm}$
dca'	$ dca' \leq 5 \text{ cm}$
forward tracks (FTD)	
momentum	$p > 0.5 \text{ GeV}$
polar angle	$6^\circ < \vartheta < 25^\circ$
radius of first hit	$R_{start} \leq 10 \text{ cm}$
number of modules	$N_{plan} \geq 2$

Table 6.1: Track Selection Cuts

Track selection cuts (H1oo software analysis framework, release 3.3.10). The different track types are combined to H1PartSelTracks, avoiding double counting.

6.2.2 LAr Acceptance Cuts

For H1PartSelTrack tracks the whole trajectory within the detector is reconstructed using a broken line fit [85]. In particular the point where the track hits the calorimeter can be determined. This entry point is required to be between the BBE and the FB2 (compare with Figure 3.3), which restricts the polar angle acceptance of the electron finding to $\sim 20^\circ < \vartheta < \sim 140^\circ$. Furthermore, entry points close to an inactive region of the calorimeter are excluded. Eight such regions are located between the calorimeter octants and designated as φ -cracks. Entry points within $\pm 1^\circ$ of a φ -crack are excluded. In addition the regions between the calorimeter wheels are also (partially) inactive and are named as z-cracks. The inactive gap depends on the position along the z-axis, or rather the angle of incidence of a trajectory pointing from the interaction point to the z-crack. The inactivity is mostly pronounced for the central z-crack between CB2 and CB3, which has a corresponding incidence angle of $\sim 90^\circ$. Tracks that are extrapolated into the region of ± 1.5 cm around this z-crack are excluded as well.

6.2.3 Preselection Cuts

To ensure a minimal energy deposition measured in the calorimeter at the position hit by the corresponding track, two further preselections are applied on the calorimeter cells behind the track, whereas cells situated within a cylinder of 30 cm radius and with the axis along the track at its entry point are considered (cf. with Figure 6.1):

- The energy deposition behind the track has to be larger than 100 MeV.
- The energy has to be deposited on more than one electromagnetic cell layer³. This cut reduces the pion background, but is completely transparent for electrons.

6.3 Input Variables to the Artificial Neural Network

The Artificial Neural Network is trained with seven input variables that all require a measured track. Two input variables furnish the NN with the position and the energy scale of the absorption shower. The other five variables all contain information about the shower development in the calorimeter.

The first variable is the z-position of the point, where the track hits the calorimeter. Its intention is to provide some information to the NN, which reflects that the calorimeter differs along the z-axis. The second variable is the logarithm of the transverse momentum of the track. Particle showers typically show a scaling behavior depending on the logarithm of the particle energy.

The remaining five variables all discriminate between signal and background. They are based on calorimeter cells that are situated within a cylinder of 30 cm around the extrapolated track trajectory, as illustrated in Figure 6.1. All measured energies are corrected for energy losses in the dead material in front of the calorimeter, according to the dead material correction developed within this thesis and which is described in Appendix B. The definition and intention of these input variables is summarized in Table 6.2 and described in details in [82].

³The calorimeter cells are arranged in layers; compare with Figure 3.3.

designation	definition	description
z-impact point	z_{imp}	z position of the track impact point on the calorimeter. This variable incorporates differences of the calorimeter along the z-axis into the NN.
Energy scale	$\ln(p)$	Logarithm of the track momentum. Scale variable for the energy of the incoming particle; particle showers scale logarithmically with the energy.
E/p	$E/p = E_{em,calo}/p_{track}$	Energy measured in the electromagnetic part of the calorimeter, divided by the track momentum. Due to the non compensating calorimeter this ratio is 1 for electrons and has lower values for pions.
Isolation	$I = \frac{E_{inner}}{E_{outer}}$	Fraction of the energy in an inner cylinder around the track with radius $R_i = 15\text{cm}$ divided by the fraction of energy in an outer cylinder around the track with $R_i = 30\text{cm}$. The lateral shower development is smaller for electromagnetic than for hadronic showers.
S_{rad}	$S_{rad} = \frac{\sum_{cells} \sqrt{E_{cell} \cdot dist^2}}{\sum_{cells} \sqrt{E_{cell}}}$	Sum over the energy weighted distance $dist$ between the calorimeter cell and the cylinder axis, $dist = \frac{ \vec{r}_{cell} \times \vec{p}_{atcalo} }{ \vec{p}_{atcalo} }$. \vec{p}_{atcalo} is the track momentum at the calorimeter, \vec{r}_{cell} the position of the calorimeter cell, compare with Figure 6.1. S_{rad} is a measure for the radial shower distribution.
S_{len}	$S_{len} = \frac{\sum_{cells} layer \cdot E_{cell}}{\sum_{cells} E_{cell}}$	Sum over the energy weighted cell layer number $layer$. S_{len} is a measure for the longitudinal shower distribution. The cells of the LAr calorimeter are arranged in layers (cf. with Figure 3.3). The layer numbering starts from the innermost layer and increases towards each subsequent layer.
Energy fraction	$E_f = \frac{E_{em}}{E_{em} + E_{had}}$	The fraction of the energy deposited in the electromagnetic part of the calorimeter divided by the energy deposited in the electromagnetic and the hadronic part. Electrons are mostly absorbed in the electromagnetic section, whereas this is not the case for pions.

Table 6.2: Input Variables to the NN

The Table lists the input variables to the Artificial Neural Network (NN). The first two input variables provide the NN with the position and the energy scale of the particle shower. The other variables are used to discriminate electromagnetic from hadronic particle showers. All energy depositions are calculated from calorimeter cells that are within a cylinder of 30 cm around the track (see Figure 6.1).

6.4 Artificial Neural Network Structure and Training

The algorithm of the Artificial Neural Network is designed as multi layer perceptron and is implemented in the RooT based software package Toolkit for Multivariate Analysis (TMVA). In TMVA the algorithm used is denoted as 'TMultiLayerPerceptron'. Documentation about the algorithm and TMVA can be found in [86] and the references therein. The NN architecture for this analysis is composed of 7 input layers, 2 intermediate layers and the neuron response function is chosen for best discrimination as 'tanh'.

The NN was trained from so called Single Particle Monte Carlo samples⁴; the training samples used are listed in Table 6.3.

	particle	p_T range [GeV]	number of events
signal	$e+$	0-15	200000
	$e-$	0-15	200000
background	$\pi+$	0-15	200000
	$\pi-$	0-15	200000

Table 6.3: Training Sample

The Artificial Neural Network was trained from the Single Particle Monte Carlo samples that are listed in the Table.

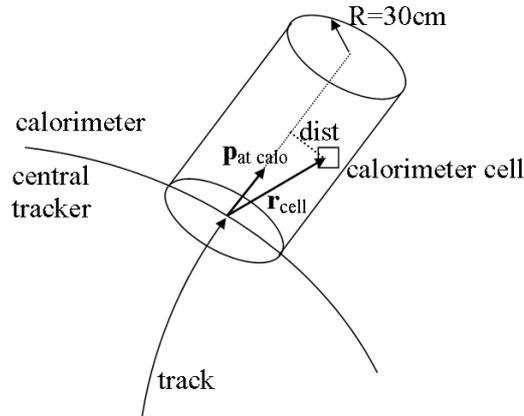


Figure 6.1: Illustration of the Calorimeter Volume Considered

Schematic illustration of the calorimeter volume considered for the calculation of the NN input variables (defined in Table 6.2): The track of a particle produced in the center of the detector is extrapolated into the calorimeter. Calorimeter cells within a cylinder around the track with a radius of 30 cm are used to discriminate between electromagnetic and hadronic absorption showers.

⁴In a Single Particle Monte Carlo the detector response due to a single particle in the detector is simulated.

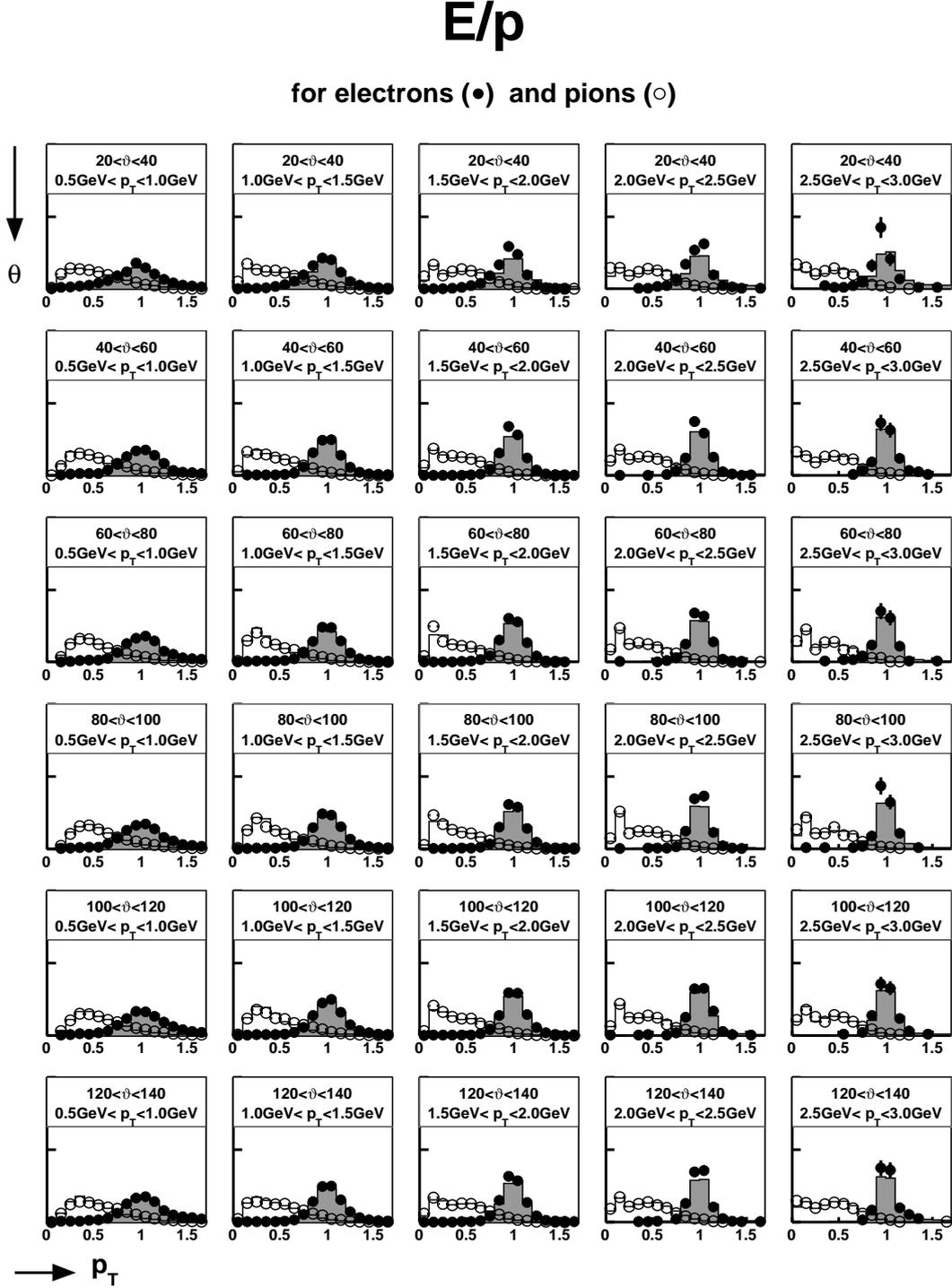


Figure 6.2: NN Input Variable E/p

Distributions for the NN input variable E/p as defined in Table 6.2. Data (points) is compared to simulation (histogram) for electrons (shaded histogram, black points) and pions (white histogram, open points), whereas all distributions are normalized to unity. For the data samples pions are obtained from a $\rho \rightarrow \pi\pi$ enriched sample, and electrons from $J\psi \rightarrow ee$ enriched events. The individual subfigures correspond to different phase space regions in ϑ and p_T , indicated in the subfigure captions and by the two arrows at the side.

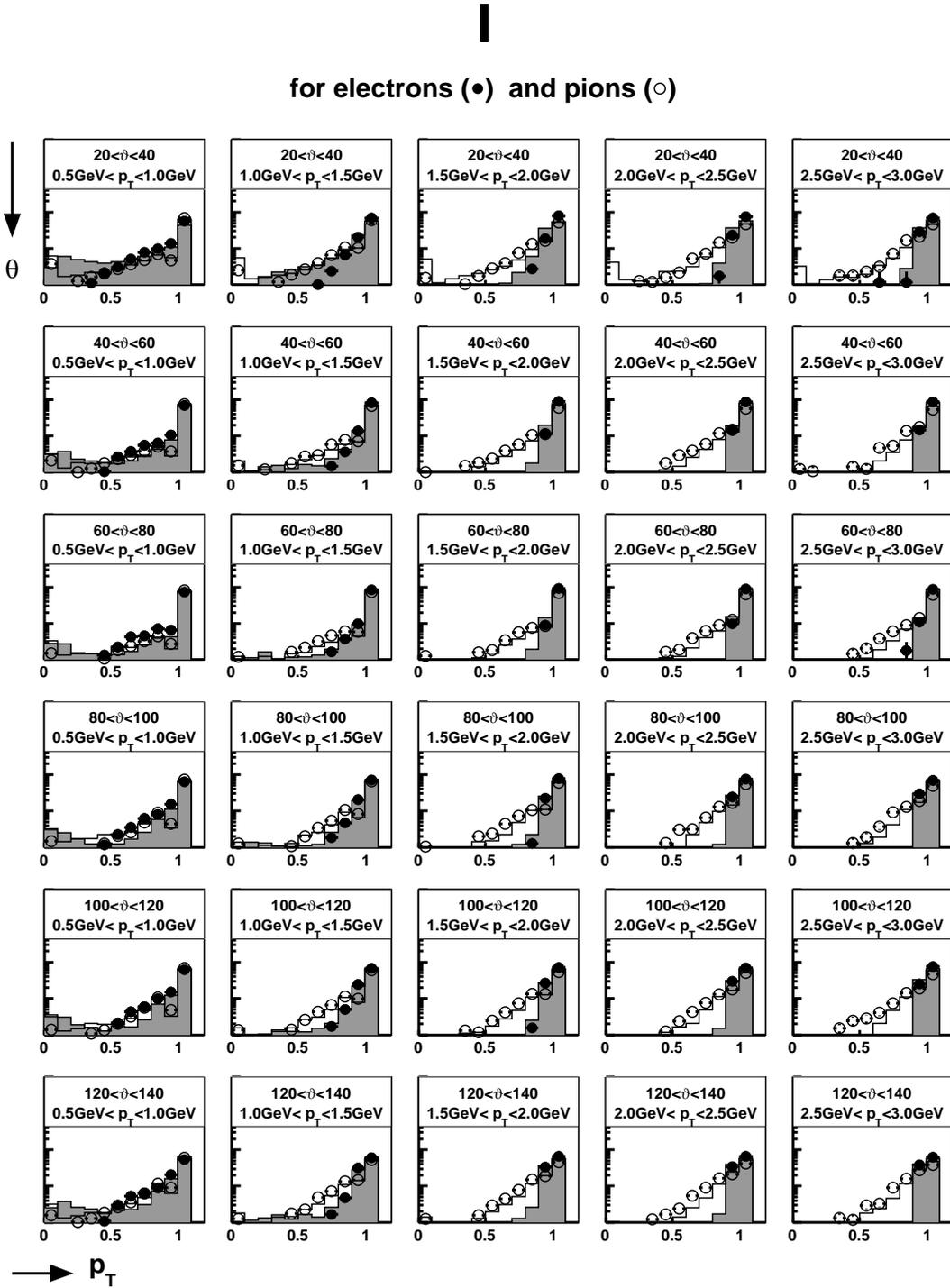


Figure 6.3: NN Input Variable Isolation
Distributions for the NN input variable Isolation as defined in Table 6.2, cf. with the caption of Figure 6.2.

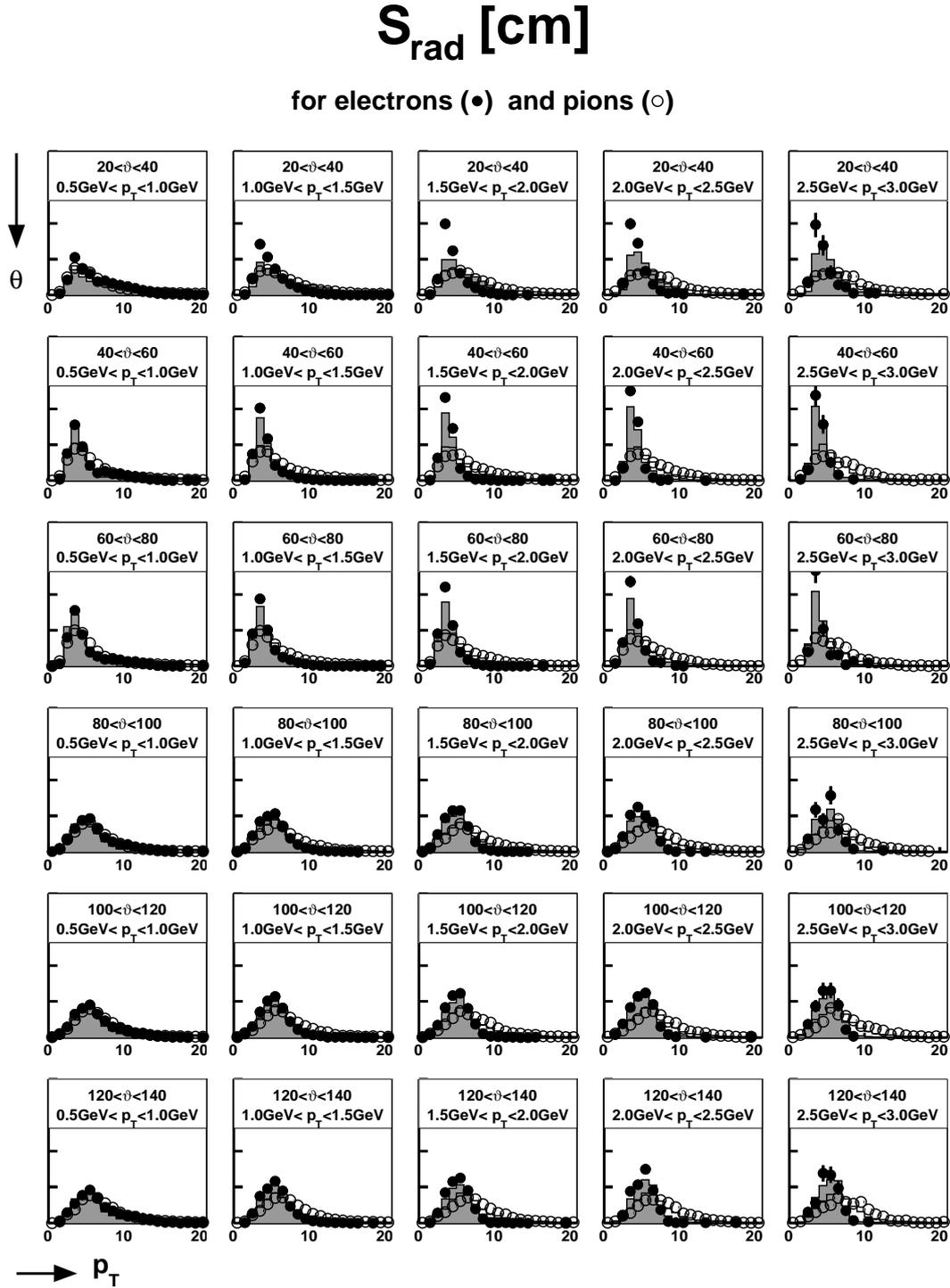


Figure 6.4: NN Input Variable S_{rad}
 Distributions for the NN input variable S_{rad} as defined in Table 6.2,
 cf. with the caption of Figure 6.2.

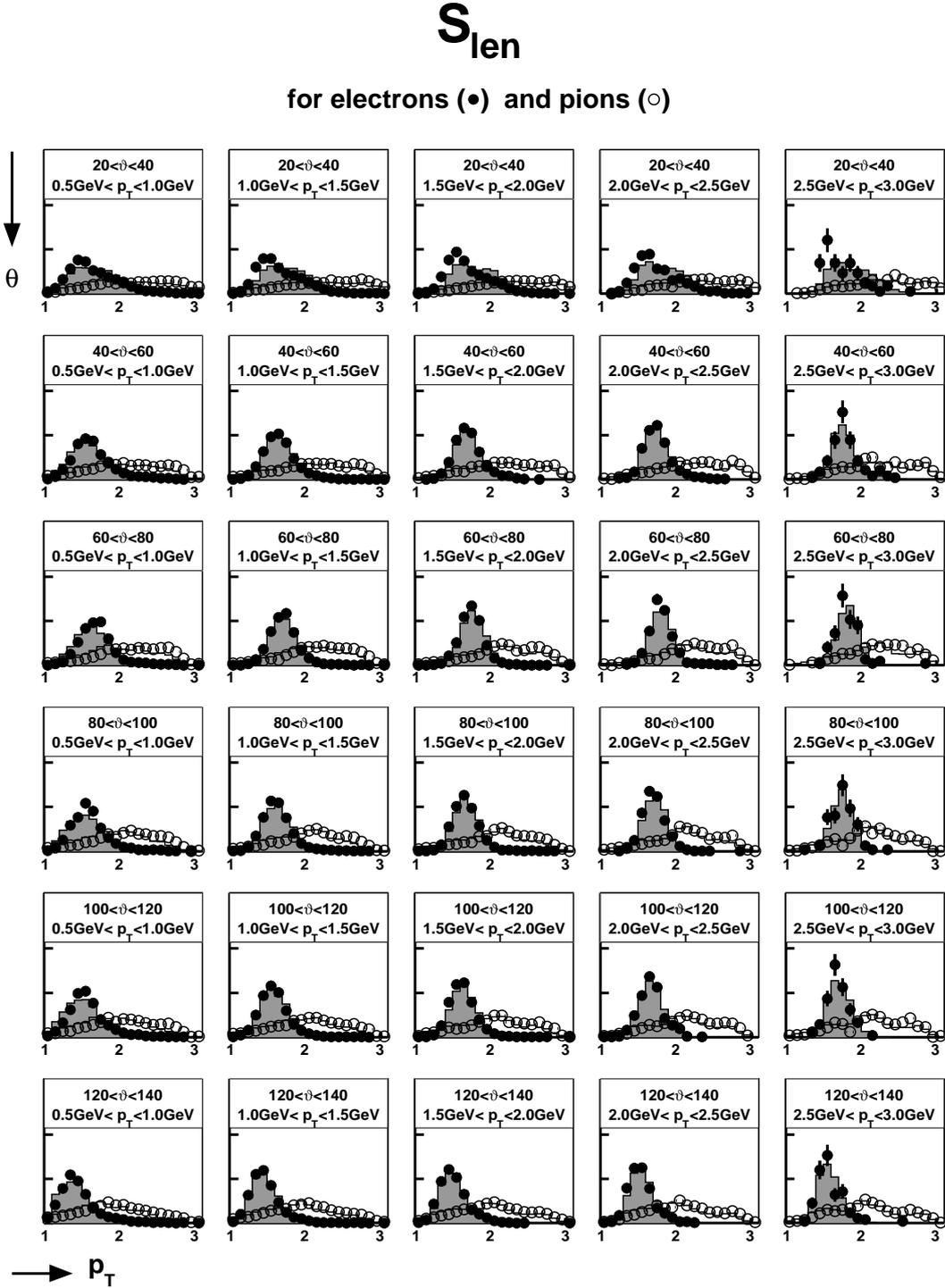


Figure 6.5: NN Input Variable S_{len}
Distributions for the NN input variable S_{len} as defined in Table 6.2,
cf. with the caption of Figure 6.2.

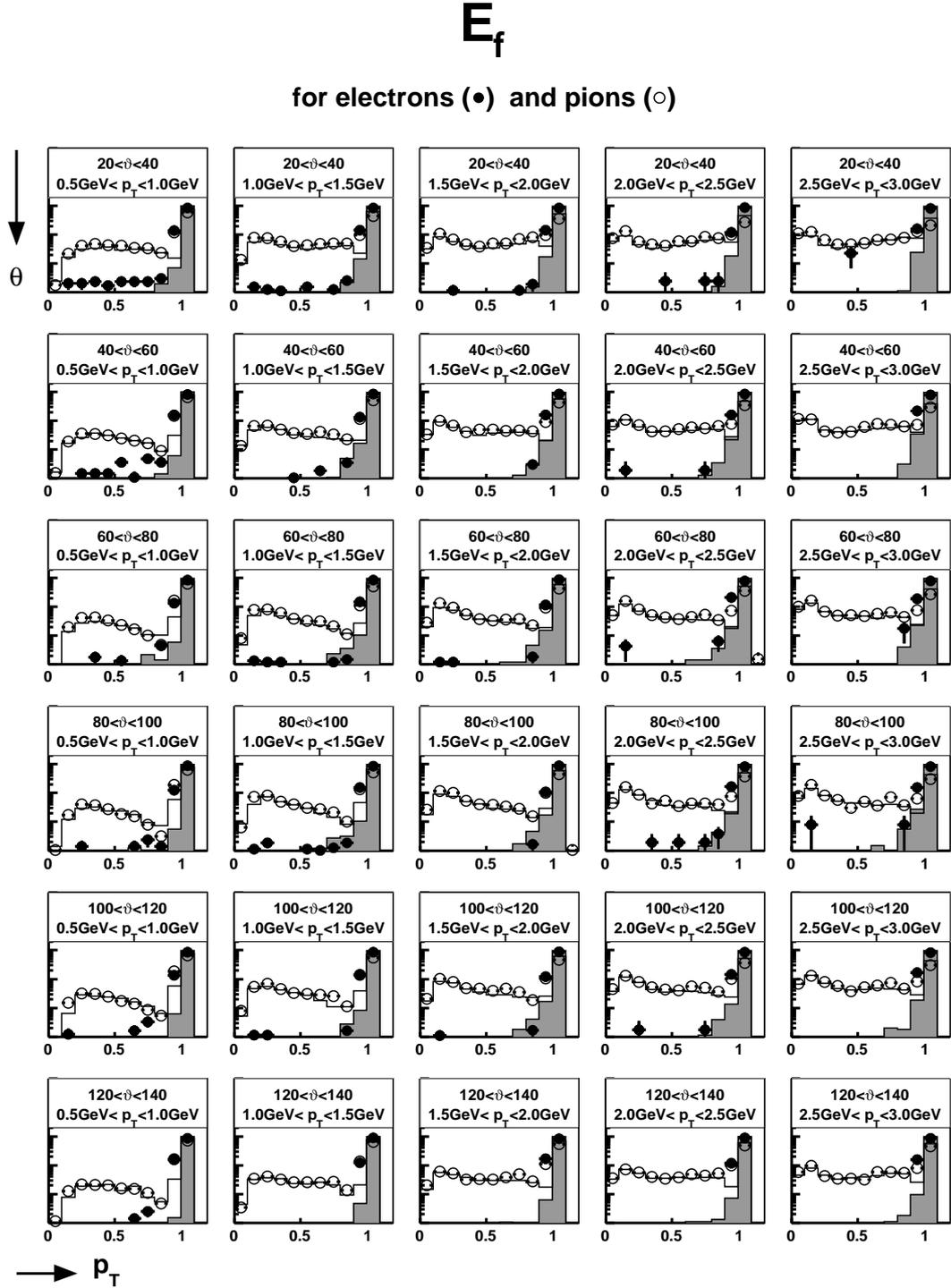


Figure 6.6: NN Input Variable E_f
 Distributions for the NN input variable E_f as defined in Table 6.2,
 cf. with the caption of Figure 6.2.

6.5 Control Distributions of the NN Input Variables

The five NN input variables that discriminate between signal and background showers are compared in Figure 6.2-6.6 between data and simulation, for signal and for background. The plots are based on two Single Particle Monte Carlo samples containing electrons (signal), respectively pions (background). For the data samples pions are obtained from a $\rho \rightarrow \pi\pi$ enriched sample, and electrons from $J/\psi \rightarrow ee$ enriched events. The plots presented are filled from these data samples such that one obtains single track distributions.

Each Figure contains six times five subfigures, whereas each subfigure covers a different kinematical phase space in $p_T(e)$ and $\vartheta(e)$, indicated in the subfigure captions. Electrons in data (simulation) are presented as black points (shaded histogram), pions as open points (white histogram). All NN input variables clearly separate electrons from pions in the whole kinematical phase space, whereas the signal-versus-background discrimination improves towards higher $p_T(e)$ values, or more precisely the more energy is deposited in the calorimeter. Moreover, the simulation describes the data for signal and background and all NN input variables reasonably well, which is obtained by the dead material correction presented in Appendix B.

6.6 Result

The NN output variable MLP is compared between data and simulation in Figure 6.8, based on the same samples as discussed above. Pions peak at $MLP = -1$, while electrons peak at $MLP = 1$. The two distributions are clearly separated from each other and the data is described by the simulation in the whole kinematical phase space.

Also at low transverse momenta MLP allows an almost optimal distinction between signal and background. This is demonstrated in Figure 6.7: The background (pion) rejection is plotted against the signal (electron) efficiency for different cuts on MLP and for electrons (pions) having a transverse momentum smaller than 3 GeV. At 95 % signal efficiency 95 % of the background is rejected.

6.7 Other Measurements based on the Electron Identification presented

The electron identification algorithm presented in this chapter was implemented within the H1 analysis framework H1oo, and that way made available to the H1 collaboration. It was exploited in a recent analysis studying inelastic $J/\psi \rightarrow eeX$ in DIS [75]. Figure 6.9 shows the e^+e^- invariant mass distribution for electrons selected by this electron identification algorithm: Clearly visible is the J/ψ peak above the background. In particular at low z values⁵, where the background significantly contributes, the new electron identification improves the peak reconstruction with respect to other electron identification algorithms used in the corresponding analysis.

Furthermore, the analysis team measuring the F_L structure function at high Q^2 [87] has also shown some interest in the method. Moreover, in a recent diploma thesis [88] elastic $J/\psi \rightarrow ee$ events were measured at low $W_{\gamma P}$.

⁵The Lorentz invariant elasticity z is defined as

$$z = \frac{\sum_{J/\psi} E - P_z}{\sum_{HFS} E - P_z} .$$

In the proton rest frame, z is the fractional energy of the virtual photon transferred to the J/ψ meson.

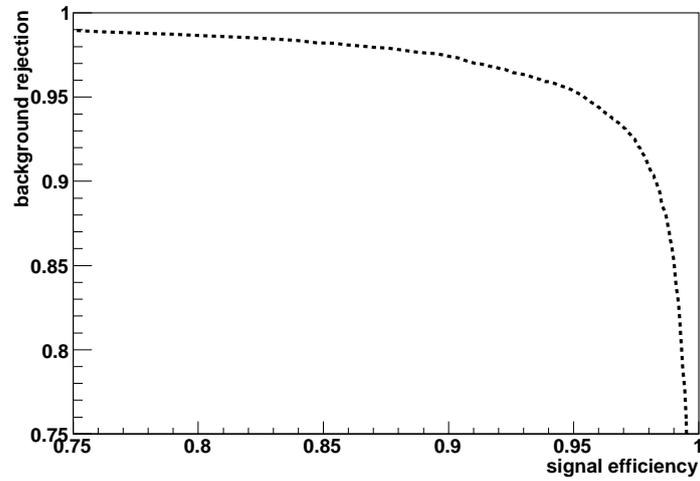


Figure 6.7: Signal Efficiency vs. Background Rejection
Signal (electrons) efficiency versus background (pions) rejections in the range $0.5 \text{ GeV} < p_T < 3 \text{ GeV}$ as obtained from a single particle Monte Carlo.

MLP

for electrons (•) and pions (○)

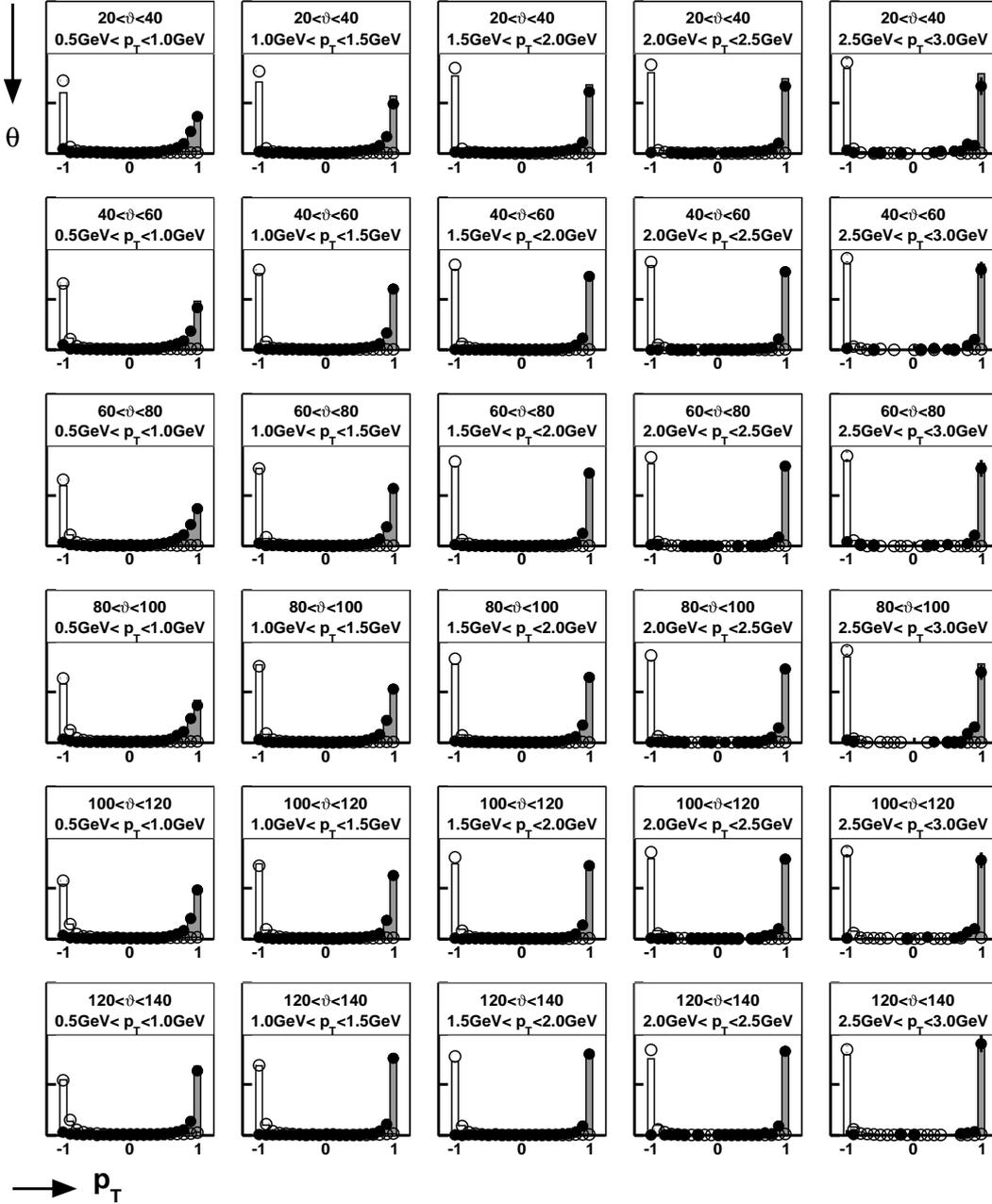


Figure 6.8: NN Output Variable MLP

Distributions for the NN output variable MLP , cf. with the caption of Figure 6.2.

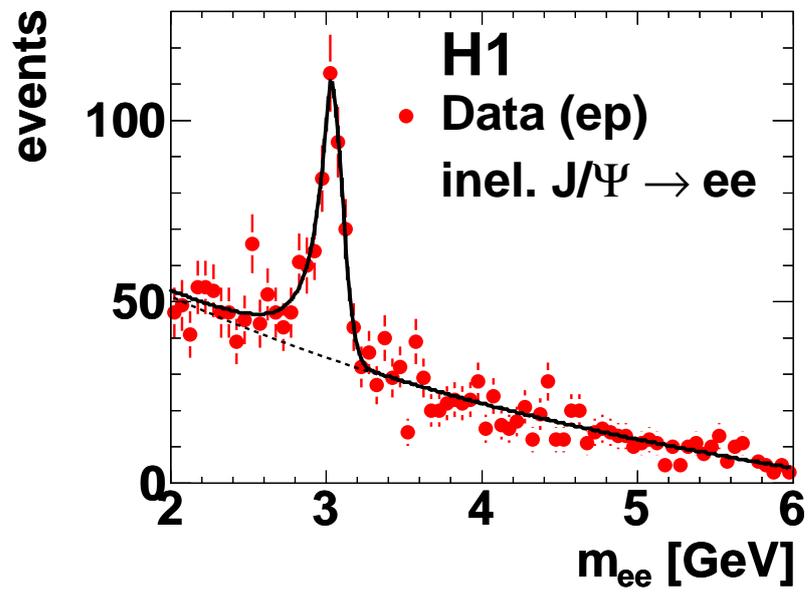


Figure 6.9: Inelastic J/ψ Peak in DIS
 J/ψ peak of an analysis studying inelastic J/ψ production in DIS, applying the electron identification developed in this thesis. Clearly visible is the J/ψ peak above the background [75].

Chapter 7

Event Selection

The goal of this analysis is the measurement of beauty photoproduction using the decay channel $b\bar{b} \rightarrow eeX$. The experimental challenge to identify this decay signature is the discrimination between electrons and the huge background consisting predominantly of hadrons. The chapter starts with a section discussing the online selection of beauty events by triggering on electrons, and a method combining the data sets of different subtriggers. The next sections treat the offline selection of photoproduced events containing two electrons in the final state, followed by a section verifying various issues discussed in the sections before, with a set of control distributions.

The offline selection is split in different selection steps, each presented in a subsection. The full set of selection criteria is summarized in Table 7.4.

7.1 Trigger Selection

7.1.1 A FTT-JT based Electron Trigger for Beauty Events

The combination of the rich track information determined by the Fast Track Trigger on the second trigger level (FTT L2) with the energy depositions in the calorimeter measured by the Jet Trigger (JT) allowed an efficient identification of electrons from b-quark decays, while strongly suppressing background. The concept of the FTT-JT based electron subtrigger was to correlate FTT tracks to the energy measurement of the JT on the third trigger level (L3). Both systems together provided particle information about momentum, energy and the position in the $\vartheta\varphi$ -plane.

Three L3 electron subtriggers (s29, s32 and s34) with equal level 1 and similar level 2 trigger conditions (L1, L2), but different conditions on the third trigger level have been commissioned. An overview of the composition of the main subtrigger elements is given in Table 7.1.

The L1 condition was mainly composed of subtrigger elements provided by the CIP trigger and the FTT. These conditions were designed to reject background and to select events of physical interest at an early stage. The purpose of the CIP trigger elements was to select events with tracks pointing to the nominal interaction point: The trigger element 'CIP_mul' corresponded to the ratio between the number of central tracks $N(\text{ctr})$ (measured in the CIP), and the number of the backward $N(\text{bwd})$ and forward tracks $N(\text{fwd})$ with respect to the beam line. (Events containing many tracks originating from the backward (forward) region were usually non-ep background.) The trigger element 'CIP_mul' represented the track multiplicity measured by the CIP. Only events with at least 10 CIP tracks were selected.

More information about CIP trigger elements can be found in [52]. The purpose of the other L1 trigger elements is to enrich ep-events with a high track multiplicity. These elements all represent the number of tracks (measured by the FTT L1) which are above a certain threshold on the transverse track momentum.

On the second trigger level, events had to contain at least two (respectively three¹) tracks (measured by the FTT L2) with a transverse momentum above 800 MeV. In addition an online reconstructed vertex was required, that was determined by extrapolating the FTT L2 tracks to their origin (compare with [42]).

As already mentioned above, electrons were identified on the third trigger level by combining FTT L2 tracks with energy depositions in the LAr calorimeter, measured by the JT. The FTT tracks were topologically linked to the JT energy depositions in the variables φ and ϑ . The main focus was on discriminating electrons from hadrons, which was achieved by a cut on the ratio between the transverse energy measured by the JT $E_{T,JT}$ and the transverse track momentum measured by the FTT $p_{T,FTT}$. Due to the non compensating LAr calorimeter, more energy was observed for electrons than for hadrons having the same momentum. The main differences between the three commissioned electron subtriggers are the p_T -thresholds and the number of online reconstructed electron candidates. Furthermore, the commissioned subtriggers differed in the E_T/p_T -cut and the topological match between the track and the energy deposition. These technical details are described in section D.2.3 and reference [62].

7.1.2 Combination of Different Subtriggers and Correction for Prescale Factors

The three FTT-JT based electron subtriggers listed in Table 7.1 covered an overlapping kinematic phase space. The data set of each trigger corresponds to a different integrated luminosity due to different prescale factors. The technique at H1 to prescale subtriggers is discussed in section 3.2.4. The mean prescale factors and the integrated and prescale corrected luminosity of each electron subtrigger are listed in Table 7.2. A method to solve the complexity of such a situation, i.e. to combine correlated subtriggers each having a different integrated luminosity, is described in [89] and [90], and is summarized in the following.

The crucial point is a uncommon interpretation of the trigger system: The entire kinematic phase space was in principle split in different, partially overlapping phase space regions by the different subtriggers. The conceptual idea of overlapping phase space regions is illustrated in Figure 7.1. Depending on the physical interest in a certain phase space region, different band widths were assigned to the corresponding subtrigger. If necessary, these band widths were adjusted by prescale factors.

The standard procedure in an analysis is to select a certain trigger (respectively its phase space) and to correct the luminosity for the mean prescale factor. An equivalent method is to weight the events with the mean prescale factor:

$$\sigma = \frac{N_{Events}}{\varepsilon \cdot \mathcal{L} \cdot \frac{1}{\langle pf \rangle}} = \frac{N_{Events} \cdot w}{\varepsilon \cdot \mathcal{L}}, \quad w = \langle pf \rangle \quad (7.1)$$

σ is the cross section, ε the corresponding efficiency, \mathcal{L} the integrated (uncorrected) luminosity, $\langle pf \rangle$ the mean prescale factor and w the corresponding weight.

When combining several subtriggers, the same procedure can be applied. In exclusive phase space regions covered by a single subtrigger, the weights are again the mean prescale factors. For the overlapping regions the determination of the weights is more complicated, and is explained below. Once the weights for every phase space region are known, a unique weight

¹Depending on the subtrigger, cf. Table 7.1.

Subtrigger	Trigger Element	Trigger Element Definition
LEVEL1		
s29, s32, s34	CIP_sig > 2 & CIP_mul4 & FTT_mul.Ta > 5 & FTT_mul.Tc > 2 & FTT_mul.Td > 1 & FTT_mul.Te > 0	N(ctr) > (N(bwd) + N(fwd)) N(ctr) + N(bwd) + N(fwd) > 10 #L1tracks > 5, with p _T > 100 MeV #L1tracks > 2, with p _T > 400 MeV #L1tracks > 1, with p _T > 900 MeV #L1tracks > 0, with p _T > 1.8 GeV
LEVEL2		
s29	& FTT_mul.Ta >= 6 & FTT_mul.Te >= 3 & FTT_zvtx_hist >= 2	#L2tracks >= 6, with p _T > 100 MeV #L2tracks >= 3, with p _T > 800 MeV reconstructed FTT zVtx
s32, s34	& FTT_mul.Te >= 2 & FTT_zvtx_hist >= 2	#L2tracks >= 2, with p _T > 800 MeV reconstructed FTT zVtx
LEVEL3		
s29	medium-p _T , single tag	1 electron, with p _T > 1.5 GeV
s32	low-p _T , double tag	2 electrons, with p _T > 1.2 GeV
s34	high-p _T , single tag	1 electron, with p _T > 2.0 GeV

Table 7.1: Definitions of FTT-JT Based Electron Subtriggers

The trigger definition on L1, L2 and L3 of the FTT-JT based electron subtriggers, s29, s32, s34 are listed in the table. For each trigger level the composition of trigger elements, and their constraint is given in the middle-column. The right column contains a brief definition of the trigger elements (compare with the text and Table D.1 in Appendix D).

subtrigger	name	$\langle \text{pf} \rangle$	$\mathcal{L} / \langle \text{pf} \rangle$ [pb ⁻¹]
s32	low-p _T , double tag	1.95	24.4
s29	medium-p _T , single tag	3.73	12.7
s34	high-p _T , single tag	1.45	32.8

Table 7.2: Integrated Luminosity of the L3 Electron Subtriggers

For the data period considered in this analysis the mean prescale factors $\langle \text{pf} \rangle$ and the prescale corrected integrated luminosity $\mathcal{L} / \langle \text{pf} \rangle$ are listed for the the JT- FTT based L3 electron subtriggers, s29, s32, s34.

can be assigned to every data event, depending on its phase space categorization by the trigger. This categorization was recorded and saved in the raw trigger bits of the event.

The weighted data distributions are the prescale corrected distributions, i.e. the distributions that were measured with an 'OR'-combination of the considered subtriggers and none of them downscaled by a prescale factor.

Note, that in such a subtrigger combination one can also take subtriggers into account for which one does not have a decent experimental understanding, since the weights are only determined from the prescale factors and the data period. The weights are independent from the trigger efficiency. It is preferable to exclude the subtriggers with a non understood trigger efficiency after the combination, since in the overlapping regions the weights are reduced. In particular for subtriggers that are well understood, but were highly prescaled, this is an interesting option.

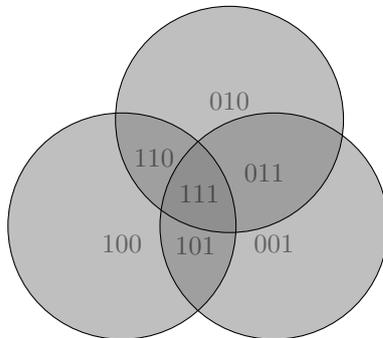


Figure 7.1: Overlapping Subtriggers

Sketch illustrating the overlapping phase space of three different subtriggers, each represented by a circle. The phase space region of an event is stored in its raw trigger bit pattern, represented here by the bit pattern.

7.1.3 Event Weights Correcting for Prescale Factors

As mentioned above, phase space regions are formed by the overlap of different subtriggers and the phase space of an event is stored in its raw trigger bit pattern, RawTb (see Figure 7.1).

The weight for run k in the phase space region RawTb is given according to [89] by

$$w_{k,uncorr.}(\text{RawTb}) = \frac{1}{1 - \prod_{i=\text{subtrig} \in \text{RawTb}} \left(1 - \frac{1}{\text{pf}_{i,k}}\right)}, \quad (7.2)$$

where $\text{pf}_{i,k}$ is the prescale factor of subtrigger i in run k . The product runs over all subtriggers considered that overlap.

In equation (7.2) it is assumed that the downscaling of the different overlapping subtriggers is uncorrelated. In the H1 trigger system the downscaling of a subtrigger was solved such, that if the corresponding L1 counter modulo the assigned prescale factor was zero, the event

was passed to the second trigger level. The L1 counter started at 0 at run start and counted the number of times the L1 condition of the subtrigger was fulfilled. Thus the assumption of uncorrelated downscaling of the subtriggers is true for long runs and subtriggers with independent L1 conditions.

Subtriggers with the same L1 condition, as it was for example the case with the three FTT-JT based electron triggers, are correlated and have to be treated differently. The weights are then given by [91]:

$$w_{k,corr.}(\text{RawTb}) = \left(\sum_{s=1}^{N_{pf,RawTb}} \left(\sum_{i=1}^{N_{C(\mathbb{A}_{pf}(k,RawTb),s)}} \frac{(-1)^{s+1}}{LCM(C(\mathbb{A}_{pf}(k,RawTb),s,i))} \right) \right)^{-1} \quad (7.3)$$

The quite complicated notation of equation (7.3) is explained at the bottom of this page², an example is given in the Appendix E. For the combination of correlated and uncorrelated subtriggers, first the weight for the correlated subtriggers has to be determined with equation (7.3). The combined correlated subtriggers can then be interpreted as an uncorrelated subtrigger with respect to others, and plugged into equation (7.2).

In the case of prescale factors varying from run to run (as it usually was the case at H1), events with the same subtrigger bit pattern are assigned different weights for different runs. This inappropriate variation of the weights can be smoothed out by averaging the weights over the run range:

$$w(\text{RawTb}) = \frac{\sum_{k=1}^{N_{runs}} \mathcal{L}_k}{\sum_{k=1}^{N_{runs}} \mathcal{L}_k \cdot \frac{1}{w_k(\text{RawTb})}} \quad (7.4)$$

The sum is performed over all runs being considered and the individual factors $\frac{1}{w_k(\text{RawTb})}$ are weighted by the integrated run luminosity \mathcal{L}_k .

s29	s32	s34	weight
1	0	0	3.73478
0	1	0	1.94555
1	1	0	1.84973
0	0	1	1.45037
1	0	1	1.28153
0	1	1	1.22642
1	1	1	1.20758

Table 7.3: Event Weights to Correct for Prescale Factors

Event weights applied in this analysis, as function of the raw trigger bit pattern (first three columns) of the subtriggers that triggered the data set. The weights in the rows with only one trigger bit set are equal to the mean prescale factor of the corresponding subtrigger.

² $\mathbb{A}_{pf}(k, RawTb)$ is the set of prescale factors in run k of all subtriggers that overlap in the phase space region given by $RawTb$. $N_{pf,RawTb}$ is the number of subtriggers that overlap. $C(\mathbb{A}, s)$ are the combinations of size s of the set \mathbb{A} , i.e. the un-ordered collection of distinct elements. $C(\mathbb{A}, s, i)$ is the i^{th} combination. $N_{C(\mathbb{A}, s)}$ is the number of combinations of size s . $N_{\mathbb{A}}$ = number of elements in \mathbb{A} . $LCM(\mathbb{A})$ is the least common multiple of the numbers in the set \mathbb{A} .

7.2 Subtriggers Selection

The data set analyzed herein was triggered by the L3-electron subtriggers s29, s32, s34. An actual trigger bit of at least one of these triggers is required. The event weights applied to the data to correct for the different prescale factors of the three subtriggers are listed in Table 7.3.

In addition to the weights mentioned above, event weights correcting the L4 trigger decision are applied, as discussed in section 3.2.4. The L4 weights for all three subtriggers are almost one, equivalent to an almost transparent level four filter.

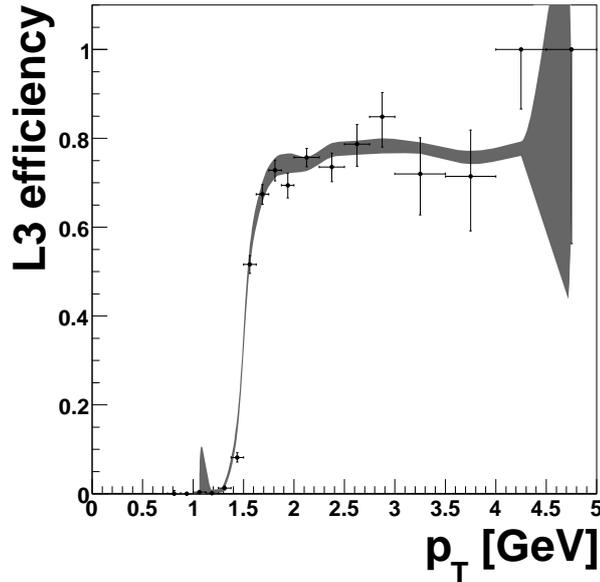


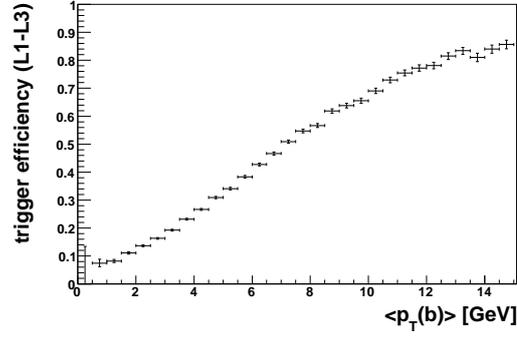
Figure 7.2: L3 Single Electron Efficiency for s29

The L3 trigger efficiency for single electrons (of subtrigger s29) is shown as function of p_T of the electron, for data (points) and for simulation (dark band), (cf. with plots in Figure C.4).

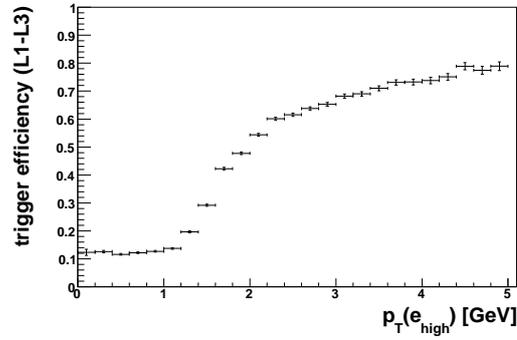
7.3 Trigger Efficiency for Beauty Events

The trigger efficiency for beauty events is determined from the Monte Carlo trigger simulation, for two reasons: First, the trigger efficiency for a $b\bar{b}$ pair decaying into two electrons is different than for background events not containing two electrons. Unfortunately, it is impossible to obtain a pure signal sample with high enough statistics to determine the signal efficiency directly in data. Second, for the low momentum phase space region covered by the L3 electron subtriggers no independent subtrigger exists that could be used as monitor trigger to determine the overall L1-L3 trigger efficiency. For possible candidates all kind of (partially unknown) correlations would have to be taken into account.

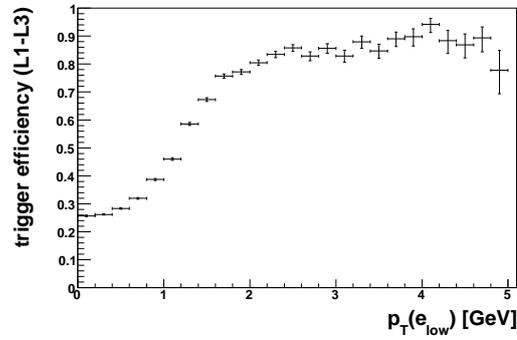
The determined L1-L3 trigger efficiency for signal events relies on the overall performance of the trigger simulation. However, trigger efficiencies obtained from data are in this and in other analyses compared to the simulated trigger efficiencies on the different trigger levels, and are found to be reasonably well described. The L1 and L2 trigger simulation of the FTT



(a)



(b)



(c)

Figure 7.3: L1-L3 Combined Trigger Efficiency for Beauty Events

The L1-L3 trigger efficiency of the combined L3 electron triggers (s29, s32, s34) to select beauty events decaying semileptonically into two electrons, $b\bar{b} \rightarrow eeX$, with $|\eta(b), \eta(\bar{b})| < 2$, $0.2 < y < 0.8$ and $Q^2 < 1 \text{ GeV}^2$, determined from the simulation. Plot (a) shows the efficiency as function of the mean transverse momentum of the $b\bar{b}$ quarks. Plot (b) as function of the transverse momentum of the higher p_T electron, and plot (c) as function of the transverse momentum of the lower p_T electron.

was extensively studied in [88, 92–95]. In [88] the combined L1-L2 efficiency of the FTT to trigger on electrons from elastic $J/\psi \rightarrow ee$ decays is compared between data and simulation as function of various variables. The performance of the L3 simulation is presented in the Appendix D.3 (see also [92] and [73]). The least tested point in the whole simulation chain is the calibration of the JT simulation. The control of this part of the simulation is demonstrated exemplarily in the Figure 7.2 and discussed in detail in Appendix C.

Figure 7.2 shows the L3 single electron trigger efficiencies for the subtriggers s29 as function of the transverse electron momentum. Data is compared to simulation. The electrons in data are obtained from elastic $J/\psi \rightarrow ee$ events, that are restricted in the phase space such, that the single electron efficiency can be tested. The simulation is obtained from Single Particle Monte Carlo simulations. Clearly visible is the well described p_T -threshold behavior of the subtrigger at $p_T(e) = 1.5 \text{ GeV}$, and the flat plateau above of it at $\sim 75\%$.

Additional control distributions for the L3 single electron efficiency are shown in Figure C.4 in the Appendix. A set of control plots comparing the data with simulation are presented at the end of this chapter in section 7.6. The plots shown in this section rely on a good overall description of the trigger simulation and can be regarded as a final cross-check for the trigger simulation.

Figure 7.3 presents the L1-L3 efficiency of the combined L3 electron subtriggers (s29, s32, s34) to trigger beauty events decaying into two electrons, in the acceptance window $b\bar{b} \rightarrow eeX$ $|\eta(b), \eta(\bar{b})| < 2$, $0.2 < y < 0.8$ and $Q^2 < 1 \text{ GeV}^2$. The efficiency is determined from the simulation. The trigger efficiency rises as function of the transverse beauty quark momentum and as function of the transverse decay electron momentum up to $\sim 90\%$. This efficiency increase is determined by the $p_T(b)$ acceptance (Figures 4.2 and 4.3) and the $p_T(e)$ thresholds of the three L3 electron subtriggers (cf. Table 7.1).

7.4 Offline Selection

7.4.1 Run Selection

The data set considered in this analysis was measured from January 2007 until end of March 2007, when HERA collided positrons with protons at a center of mass energy of 318 GeV. The beginning of this period is given by the final commissioning of the Jet Trigger. At the end of this period the center of mass energy of HERA was lowered, before HERA and H1 came to their final shut down in summer 2007.

In order to exclude runs measured during abnormal operation of the detector, only runs passing the following quality criteria are selected.

- A run has to be flagged as 'good' or 'medium', i.e. the main detector components were operational, the run conditions were stable and no serious problem has occurred.
- The high voltage and the read-out of all for the analysis essential systems were at nominal settings. These systems are: CJC1, CJC2, LAR, TOF, VETO, LUMI, FTT, CaloTrig³.
- The integrated luminosity per run has to be larger than 0.1 nb^{-1} .

The data fulfilling these selection criteria (cf. I., II. in Table 7.4) correspond to a total integrated luminosity of $\mathcal{L} = 47.6 \text{ pb}^{-1}$.

³The JT read-out is connected to the calorimeter trigger read-out.

7.4.2 Rejection of Background Events

Non-ep background events happened by the interaction of the beams with residual gas in the beam pipe, or by the collision of satellite bunches with the nominal bunches. In both cases the interactions occurred anywhere along the beam pipe, and are not in time coincidence with the normal bunch crossing (BC) time in H1. These events are rejected by requiring the reconstructed z -vertex to be close to the nominal vertex position ($|zvtx| < 30$ cm), and the event occurring within the correct time window. Two event timing variables are determined in the offline reconstruction using information from the CJC and the LAr calorimeter. The CJC timing $T_{0,CJC}$ is defined as the time when most charged particles cross the chamber. It is measured relative to the nominal bunch crossing in units of CJC ticks (500 CJC ticks = 96 ns). The LAr calorimeter timing $T_{0,LAr}$ is derived from the time structure of the energy depositions measured in units of bunch crossing (BC). The constraints on the event timing are $410 \text{ ticks} < T_{0,CJC} < 510 \text{ ticks}$ and $-0.54 \text{ BC} < T_{0,LAr} < 0.54 \text{ BC}$ (cf. IV. in Table 7.4). These are standard values and are adopted from [96].

Since the z -vertex position depends a lot on the properties of the injected beam bunches during the HERA run operation, it is rather difficult to simulate within the detector response simulation. Hence, a good agreement in the z -vertex description by the MC should not be expected. Therefore the z -vertex position of each MC sample has been reweighted to describe the data.

Run Selection Cuts	
I.	good, medium runs
II.	HV of relevant systems on
Trigger Selection & Background Rejection	
III.	subtrigger: s29 s32 s34
IV.	$ zvtx < 30 \text{ cm}$ $410 \text{ ticks} < T_{0,CJC} < 510 \text{ ticks}$ $-0.54 \text{ BC} < T_{0,LAr} < 0.54 \text{ BC}$
Di-electron Preselection in Photoproduction	
V.	2 electron candidates with: $MLP > 0.8$ $1 \text{ GeV} < p_T(e) < 5 \text{ GeV}$ $20^\circ < \theta(e) < 140^\circ$
VI.	no identified scattered beam electron, cf. text
VII.	$\sum(E - P_z) < 40 \text{ GeV}$
Final Di-electron Selection	
VIII.	$MLP > 0.9125$
IX.	$m_{ee} > 0.75 \text{ GeV}$

Table 7.4: Overview on the Selection Cuts

An overview on the offline selection cuts is given in the table. Each selection step is discussed in the text in a corresponding section.

7.5 Selection of Di-Electron Events in Photoproduction

7.5.1 Preselected Di-Electron Sample in Photoproduction

Electrons are identified by means of an Artificial Neural Network (NN) that combines calorimeter and tracker information (as discussed in chapter 6). The output of this NN is a quality criterion for electrons and designated as MLP . The value varies between $MLP = -1$ for background and $MLP = 1$ for electrons.

In a first selection a di-electron sample is obtained by requiring at least two electron candidates ($MLP > 0.8$), each with a transverse momentum in the window $1 \text{ GeV} < p_T(e) < 5 \text{ GeV}$ and a polar angle $20^\circ < \vartheta(e) < 140^\circ$, whereas $p_T(e)$ and $\vartheta(e)$ are measured by the electron track. On one hand, below 1 GeV the electron identification becomes more difficult due to a low energy deposition in the calorimeter. On the other hand, the trigger becomes inefficient below 1 GeV , as can be seen from Figure 7.3b and 7.3c. The upper boundary of 5 GeV is a trade off between the selection of electrons from semileptonic beauty decays, and a first rejection of beam electrons from high Q^2 DIS events. Figure 4.2b shows that the electron spectrum of semileptonic beauty decays drops exponentially. On the other hand, the scattered beam electrons usually have transverse momenta far above 5 GeV .

To select so called untagged photoproduction events it is necessary to reject scattered beam electrons $e_{scat.}$ from DIS events in the main calorimeters. With the help of the standard electron identification algorithms of the H1 analysis software the scattered beam electrons are searched for in the LAr and SpaCal calorimeters. These algorithms apply a lower cut on the transverse momentum of the scattered beam electron candidate of 5 GeV , $p_T(e_{scat.}) > 5 \text{ GeV}$. All events with an identified scattered beam electron candidate in the final state are excluded.

In addition, DIS events can be rejected by a cut on the longitudinal energy flow of the hadronic final state (HFS)⁴, $\sum(E - P_z)$. The value $\sum(E - P_z)$ peaks for DIS events at the double electron beam energy, i.e. $\sim 55 \text{ GeV}$. For photoproduction events this is not the case, because the beam electron escapes into the beam pipe. In order to suppress the DIS background further, the longitudinal energy flow of the HFS is restricted to $\sum(E - P_z) < 40 \text{ GeV}$. The described selection cuts are summarized in Table 7.4 (cuts V.-VII.) and define the 'Pre-selection Sample' used for various control plots at the end of this chapter.

7.5.2 Final Di-Electron Sample

The cross section is extracted from spectra obtained from di-electron candidates. In the case of more than two electron candidates in the event, all pair combinations are formed. Due to the combinatorics it is important to suppress misidentified electron candidates as much as possible, in order not to spoil the spectra built from the electron pairs.

Misidentified electron candidates are reduced further by restricting the electron identification criterion for both electron candidates to $MLP > 0.915$. The efficiency of this cut to select $b\bar{b} \rightarrow eeX$ events with both electrons in the acceptance of the electron identification is 0.63, thus 0.79 per electron.

In addition, the invariant mass of the electron pair is required to be above $m_{ee} > 0.75 \text{ GeV}$. The invariant electron pair mass m_{ee} is derived from the 4-momentum sum of the two electron candidates $m_{ee}^2 = (P_{e,1} + P_{e,2})^2$, whereas the electron 4-momenta $P_{e,i}$ are reconstructed from the track momenta \mathbf{p}_i and the hypothesis of measuring an electron, $P_{e,i} = ((m_e^2 + \mathbf{p}_i^2), \mathbf{p}_i)$. First, this cut rejects real electron pairs from photon-conversions.

⁴For photoproduction events the kinematic variables of the hadronic final state have to be computed from the sum over all hadronic final state particles. In the strict sense the notation $\sum(E - P_z)$ refers to $\sum_{i \in \text{HFS}} (E_i - P_{z,i})$.

Second, it rejects fake pairs, where both candidates of the pair hit the LAr calorimeter side by side⁵, and have therefore overlapping energy depositions in the calorimeter. Most of this kind of background arises if the NN is spoiled by overlapping particle showers in the calorimeter⁶.

These cuts define the final di-electron sample containing 3069 electron pairs, from which the beauty cross section is extracted. The whole selection is summarized in Table 7.4. Control distributions for this final di-electron selection are discussed in chapter 8 in the Figures 8.8 and 8.9.

7.6 Control Distributions for the Di-Electron Preselection Sample

A set of control distributions is shown in the following in the sequence of the cut flow of the di-electron preselection (cuts V.-VII. in Table 7.4). The presented control plots demonstrate the rejection of DIS events, the combination of the subtriggers and the performance of the trigger simulation.

Data is compared with the flavor inclusive Monte Carlo sample (presented in section 5.2) based on the Pythia simulation in the massless mode. The relative normalization of the resolved and the direct contribution in the simulation is adjusted by scaling the resolved contribution by a factor 0.3, chosen for best agreement between simulation and data (cf. with the corresponding discussion in section 5.2).

The plot structure in the Figures 7.4-7.7 is chosen such, that the following aspects of the selection can be studied:

- The **rejection of DIS events**: Each row of the Figures has an additional cut in the sequence of the preselection cuts, listed in Table 7.4 (i.e. cut V.-VII.). As can be seen from the peak in $\sum(E - P_z)$ (plots in Figure 7.4), DIS events are rejected with each additional cut. Since the simulation is for photoproduction events, agreement between data and the simulation is only expected on the last row.
- The **combination and prescale correction of the subtriggers**: The different columns in the Figures each correspond to different subtrigger selections. In the first three columns the selection is restricted to one of the three electron subtriggers s29, s32 and s34. The data is not weighted, but the Monte Carlo is scaled to the prescale corrected luminosity. In the last column these subtriggers are combined by an 'OR' condition. In this column the data is weighted to correct for the prescale factors, as discussed in section 7.1.2 and the Monte Carlo is scaled to the nominal luminosity.
- The **trigger simulation**: The different thresholds that are visible in the p_T spectra for the three highest p_T tracks (Figures 7.5-7.7) are determined by the trigger.

From the clear suppressions of the $\sum(E - P_z)$ peak by rejecting DIS events (cut VI.) and the additional cut on the variable $\sum(E - P_z)$ (cut VII.), the remaining background from DIS events can be considered as negligible. This statement was also cross checked with a standard DIS Monte Carlo, less than 10^{-4} of all simulated DIS events were not rejected.

On the last rows the main characteristics of the shapes of the spectra are described well by the Monte Carlo simulation. The fact that these shapes are completely determined by the

⁵ The invariant mass of the electron pair m_{ee} is small for electron candidates having a small intermediary angle $\Delta\phi$.

⁶ For instance, consider the NN input variable E/p , which compares the energy measured in the calorimeter behind the track with the track momentum. It peaks at 1 for electrons and has lower values for pions. Overlapping pion showers shift the energy E deposited behind the track towards higher values, while this is not the case for the momentum p .

trigger, can be regarded as a cross check that the trigger simulation is under control. On the last row in each column the normalization between data and Monte Carlo agrees reasonably well, after the downscaling of the resolved contribution and normalizing each column by its corresponding luminosity. The fact that agreement is also present in the last column, in which all three subtriggers are combined and in which the data is weighted, verifies the prescale correction procedure for overlapping subtriggers (described in section 7.1.2). In the Figures 7.8-7.10 further control distributions are shown for the p_T distribution of the highest and second highest p_T electron candidate and the polar angle distribution ϑ of all tracks. Clearly also in these distributions agreement between data and simulation is obtained after the di-electron preselection.

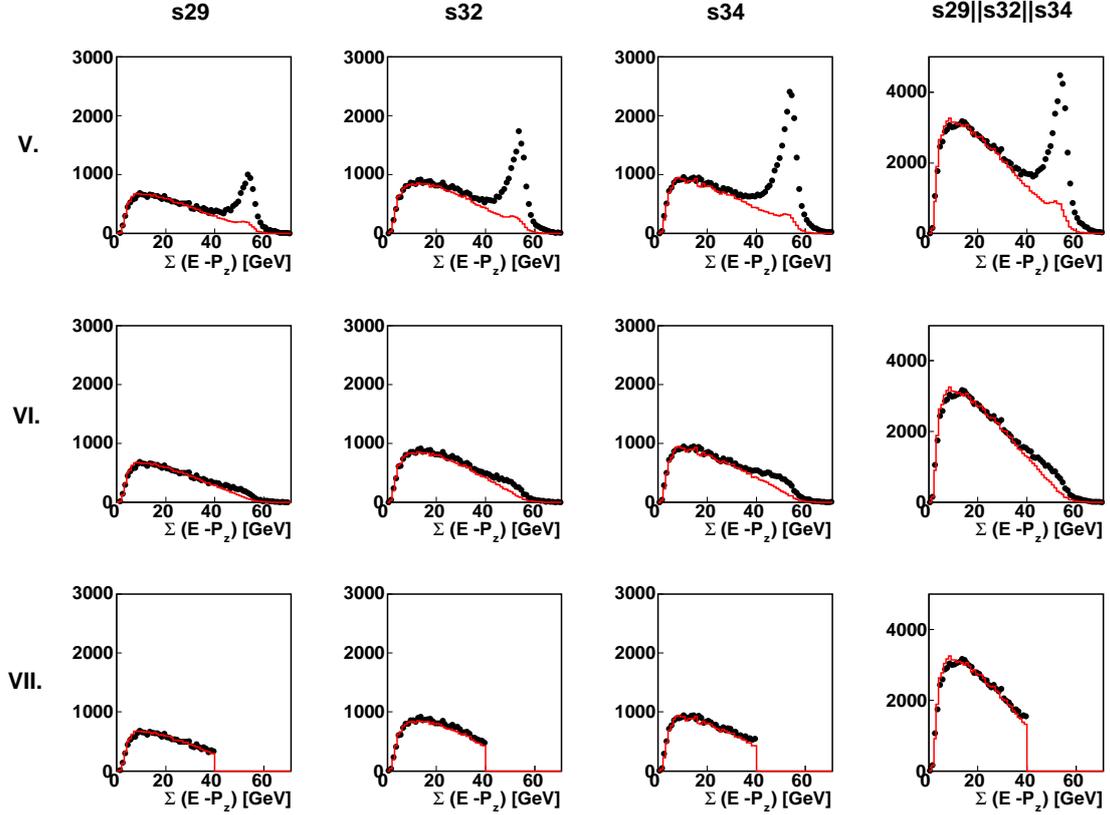


Figure 7.4: Control Distributions for $\Sigma(E - P_z)$

Distributions as function of the longitudinal energy flow $\Sigma(E - P_z)$ for data (points) and simulation (histogram). Each row contains an additional cut listed in Table 7.4 (cut V.-VII.). The columns show the three individual electron subtriggers s29, s32, s34 and the 'OR'-combination of all (as discussed in section 7.1.2). DIS events populate the histograms in the peak at ~ 55 GeV and are rejected by every additional cut. Since the simulation (mostly) contains γp events, agreement between data and simulation can be expected on the last row.

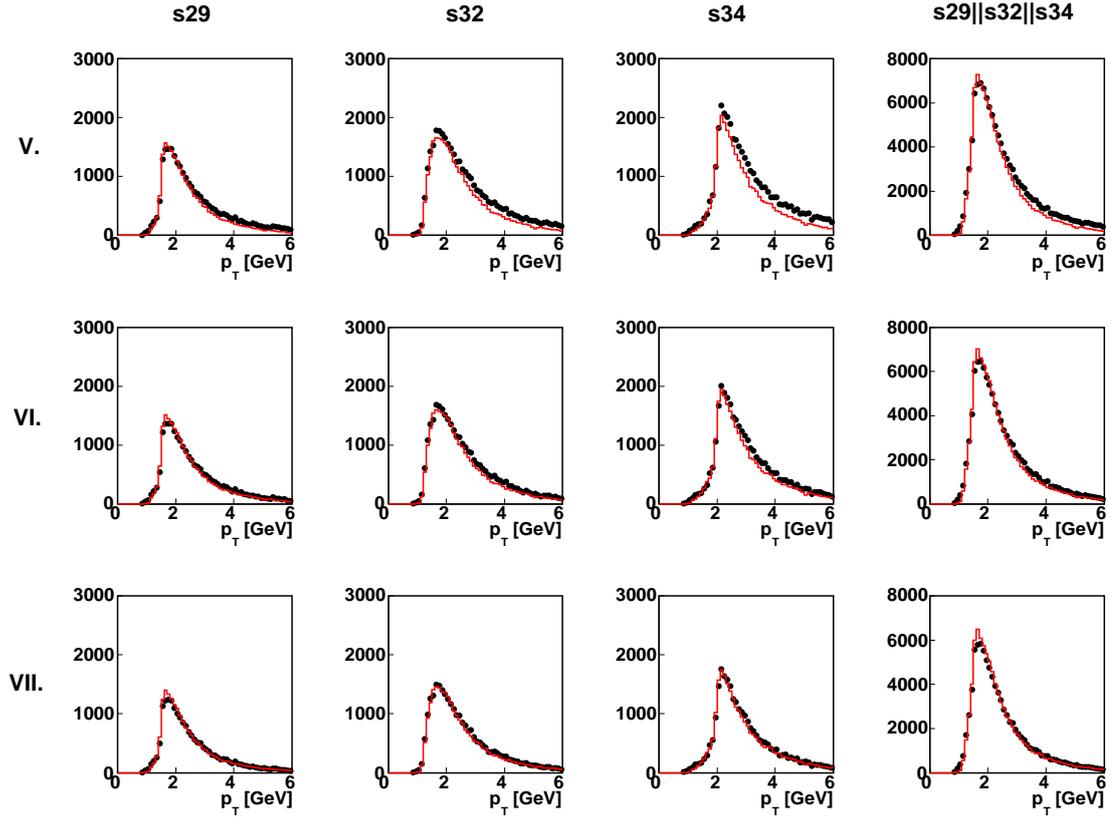


Figure 7.5: Control Distributions for Highest p_T -Track
Distributions as function of the the transverse momentum of the track having the highest transverse momentum for data (points) and simulation (histogram). Compare with the caption of Figure 7.4 for the structure of the subfigures. Agreement between data and simulation can be expected on the last row.

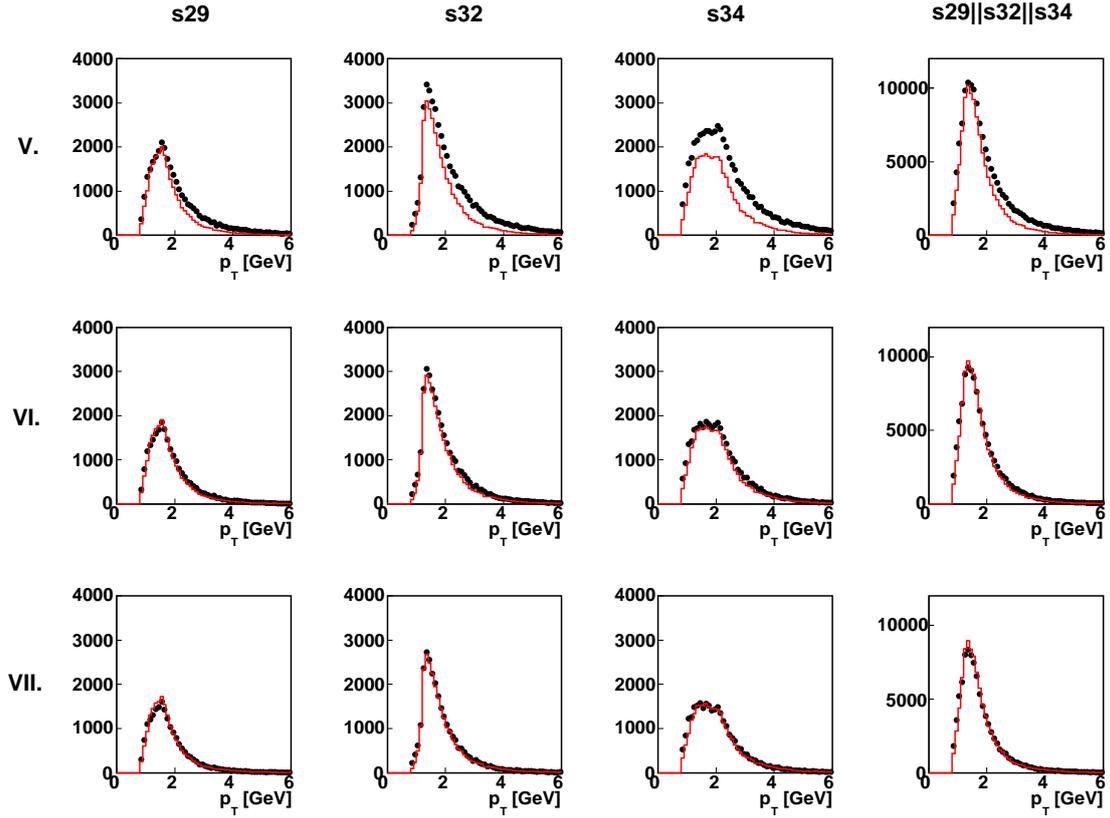


Figure 7.6: Control Distributions for 2nd Highest p_T -Track
Distributions as function of the transverse momentum of the track having the second highest transverse momentum for data (points) and simulation (histogram). Compare with the caption of Figure 7.4 for the structure of the subfigures. Agreement between data and simulation can be expected on the last row.

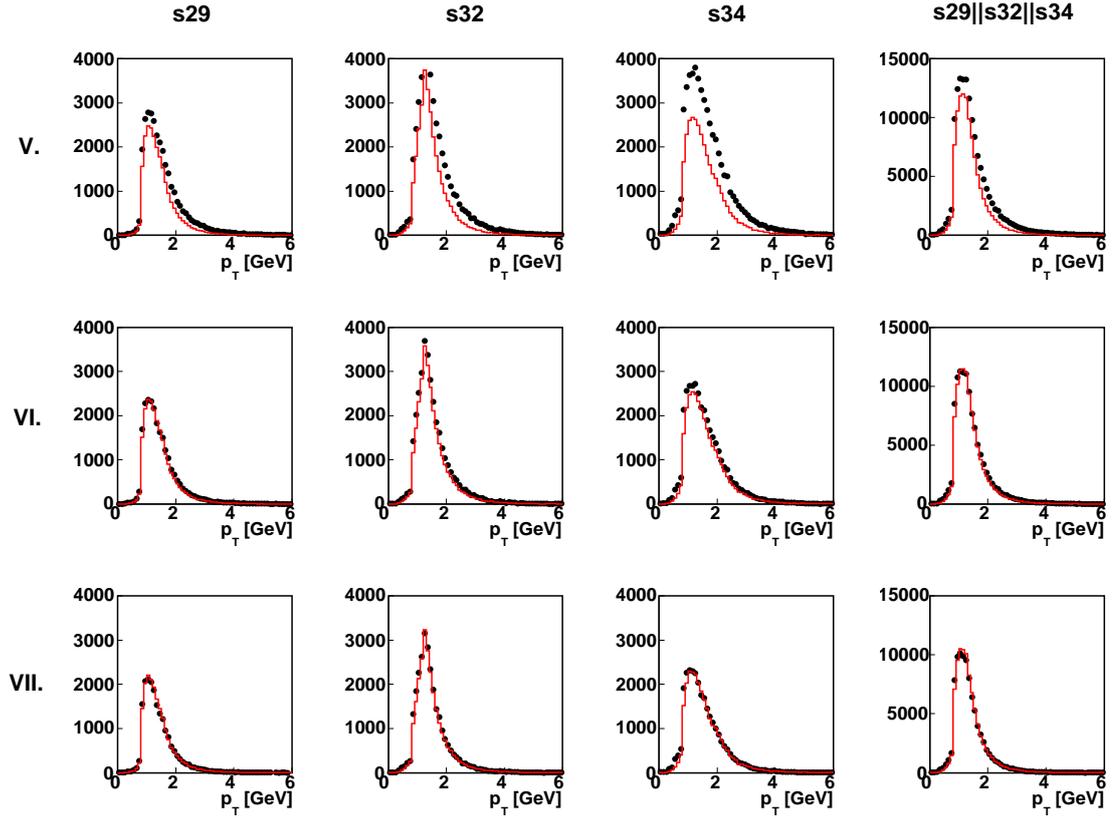


Figure 7.7: Control Distributions for 3rd Highest p_T -Track
 Distributions as function of the transverse momentum of the track having the third highest transverse momentum for data (points) and simulation (histogram). Compare with the caption of Figure 7.4 for the structure of the subfigures. Agreement between data and simulation can be expected on the last row.

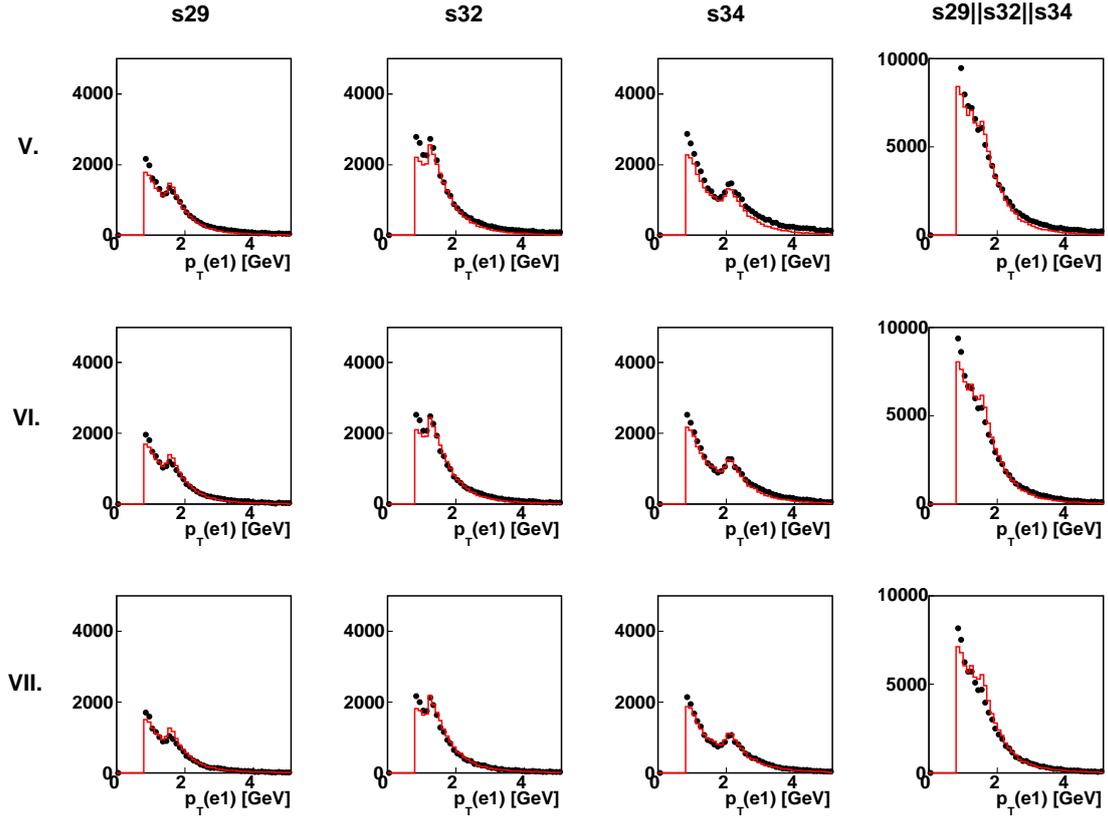


Figure 7.8: Control Distributions for Highest p_T -Electron Track Distributions as function of the transverse momentum of the electron candidate $MLP > 0.8$ having the highest transverse momentum for data (points) and simulation (histogram). Compare with the caption of Figure 7.4 for the structure of the subfigures. Agreement between data and simulation can be expected on the last row.

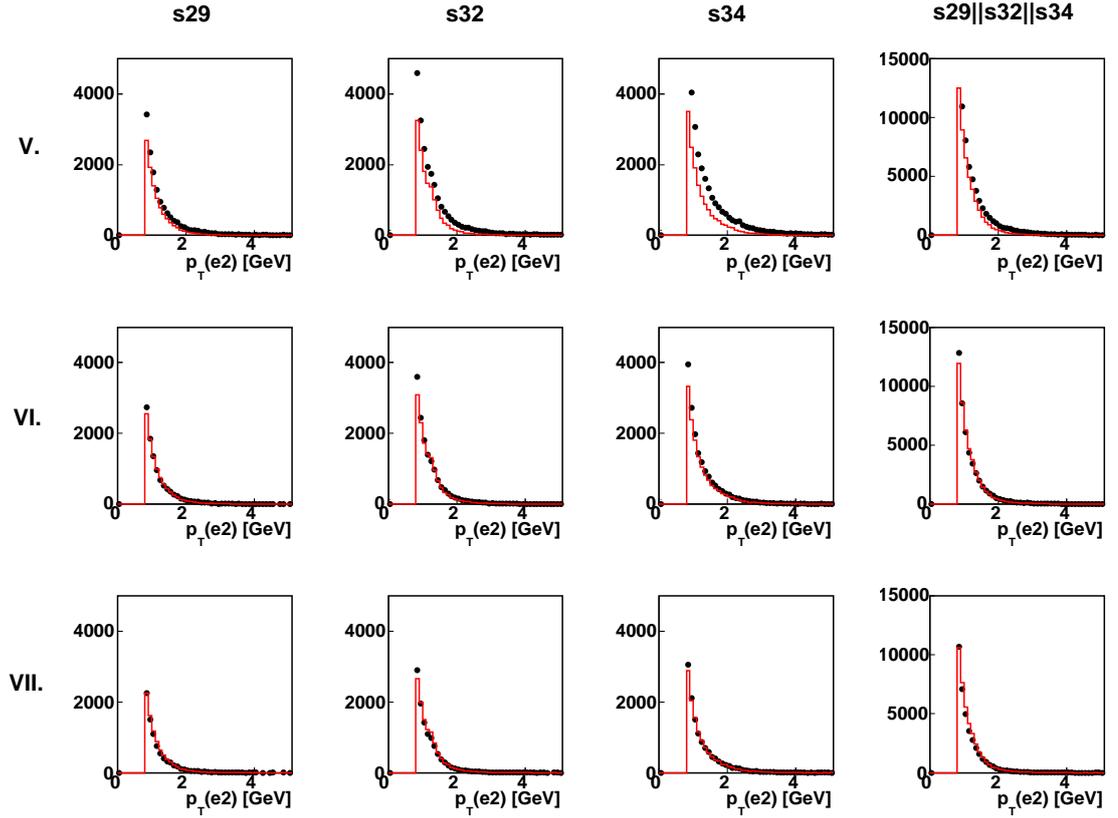


Figure 7.9: Control Distributions for 2nd Highest p_T -Electron Track Distributions as function of the transverse momentum of the electron candidate $MLP > 0.8$ having the second highest transverse momentum for data (points) and simulation (histogram). Compare with the caption of Figure 7.4 for the structure of the subfigures. Agreement between data and simulation can be expected on the last row.

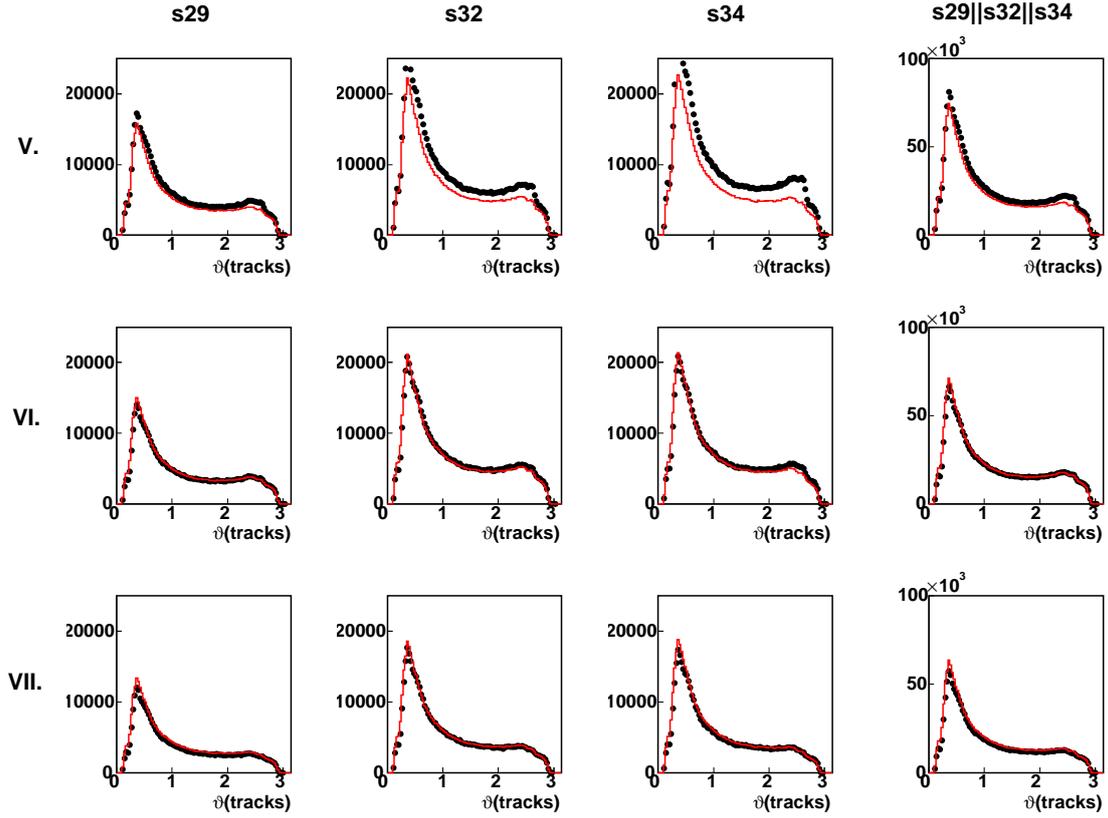


Figure 7.10: Control Distributions for all Tracks
 Distributions as function of the polar angle ϑ of all tracks for data (points) and simulation (histogram). Compare with the caption of Figure 7.4 for the structure of the subfigures. Agreement between data and simulation can be expected on the last row.

Chapter 8

Cross Section Measurement

The determination of the differential beauty production cross section in terms of transverse b-quark momenta is described in this chapter. It is based on the unfolding of $b\bar{b} \rightarrow eeX$ events (see chapter 7 for the event selection). The purpose of unfolding is twofold, on the one hand side it separates the signal from the remaining background, on the other hand side it deconvolutes the p_T -dependence of the differential cross section.

The chapter starts with a section about the measurement of the transverse b-quark momentum, followed by various aspects concerning unfolding and ends with the presentation of the measured differential beauty cross section. If the reader is not familiar with basics of unfolding, it might be preferable to first read Appendix A, where an introduction to unfolding is given.

8.1 Measurement of the Mean Transverse Momentum of the Beauty Quarks

The measurement of $\langle p_T(b) \rangle = \sqrt{(p_{T,b}^2 + p_{T,\bar{b}}^2)/2}$ within the unfolding procedure is based on an estimator for the mean transverse beauty mass of the $b\bar{b}$ pair $\langle m_T(b) \rangle = \sqrt{m_b^2 + \langle p_T(b) \rangle^2}$ (see [97] for the definition of m_T). This estimator is calculated by using the thrust axis in the transverse plane and will be discussed in section 8.1.1. Compared to the usual practice of measuring the transverse momentum of quarks with the help of jet algorithms, this approach has no lower $p_T(b)$ -cut-off, and therefore includes the beauty production threshold (cf. with the discussion in chapter 4).

8.1.1 An Estimator for the Mean Transverse Energy of Beauty Quarks based on the Thrust Axis in the Transverse Plane

Originally the thrust axis was developed for physics at colliders with symmetrical beam energies. For these experiments the laboratory frame is identical to the center of mass frame. The decay particles arrange symmetrically in the detector, due to momentum conservation. Thrust is a measure of the alignment of the particles within an event along a common axis. The lower the thrust, the more spherical is the event. The higher the thrust, the more jet-like is the event. The thrust axis is a vector along which the maximum alignment is found.

Since at HERA the colliding electrons and protons had different energies, the thrust axis

used in the following definition is constraint to the transverse plane with respect to the incoming beams [98]. The axis of the vector \vec{a} in

$$T = \max(\vec{a}) \left(\frac{\sum_{i \in HFS} |\vec{a} \cdot \vec{p}_{T,i}|}{\sum_{i \in HFS} |\vec{p}_{T,i}|} \right) \quad (8.1)$$

is the thrust axis in the transverse plane. It is chosen such that in the transverse plane the projections of all particles of the hadronic final state (HFS) onto the axis are most balanced. The procedure is illustrated in Figure 8.1. Subfigure (a) illustrates the determination of the the thrust axis. The plane that is perpendicular to the thrust axis and contains the beam axis, splits the event in two hemispheres: The first hemisphere, labeled as upper hemisphere in the figure, lies on the positive direction of the vector \vec{a} ; the other on the negative side ¹ of \vec{a} .

The motivation for this subdivision of the event is, that in ep interactions beauty quarks are dominantly produced in $b\bar{b}$ pairs, which emerge back-to-back, i.e. that such events have to be balanced in the transverse plane. Therefore all hadronic final state particles in one hemisphere are related to the first beauty quark, and the particles in the other hemisphere to the second one. However, the hadronic final state also contains particles of the so called proton remnant, leaving the interaction in the positive direction of the detector. This is illustrated in Subfigure (b) by a boson-gluon-graph like diagram, with an incoming proton from the right-hand side and a incoming electron from the left-hand side producing a $b\bar{b}$ pair. PR refers to the proton remnant leaving the interaction in the flight direction of the proton beam.

Simulation studies show, that it is favorable to split the event in three parts: One in the forward direction associated with the proton remnant and two hemispheres associated with the two beauty quarks. A good choice is to connect all HFS particles reconstructed below a polar angle of 15 degrees ($\vartheta < 15^\circ$) with the proton remnant. All other HFS particles are associated with one of the two quarks, depending on their hemisphere.

The two quark hemispheres allow the definition of a measurable quantity $m_{T,est}(b)$,

$$m_{T,est}(b) = ((|\vec{p}_{T,upper\ h.}| + |\vec{p}_{T,lower\ h.}|)/2 + p1) * p2 \quad , \quad (8.2)$$

which correlates with the non-directly measurable mean transverse mass of the $b\bar{b}$ quark pair,

$$\langle m_T(b) \rangle = \sqrt{m_b^2 + \langle p_T(b) \rangle^2} \quad . \quad (8.3)$$

The vectors $\vec{p}_{T,upper\ h.}$ ($lower\ h.$) are the transverse momenta of the summed 4-momenta of the HFS particles in the corresponding hemispheres

$$\vec{p}_{T,upper\ h.} \text{ (} lower\ h.\text{)} = \sum_{\substack{HFS_{upper\ h.} \\ (HFS_{lower\ h.})}} \vec{p}_{T,i} \quad . \quad (8.4)$$

$p1$ and $p2$ are two parameters, fixed to $p1 = 1.0$ GeV and $p2 = 1.3$, such that the correlation is maximized. The correlation of the two quantities is demonstrated in Figure 8.2, based on the Beauty simulation (Pythia massive, cf. section 5.2).

¹ 'Lower' and 'upper', respectively 'positive' and 'negative' are only used as names here, and do not have a profounder meaning. The definition of the sign of \vec{a} in equation 8.1 is ambiguous.

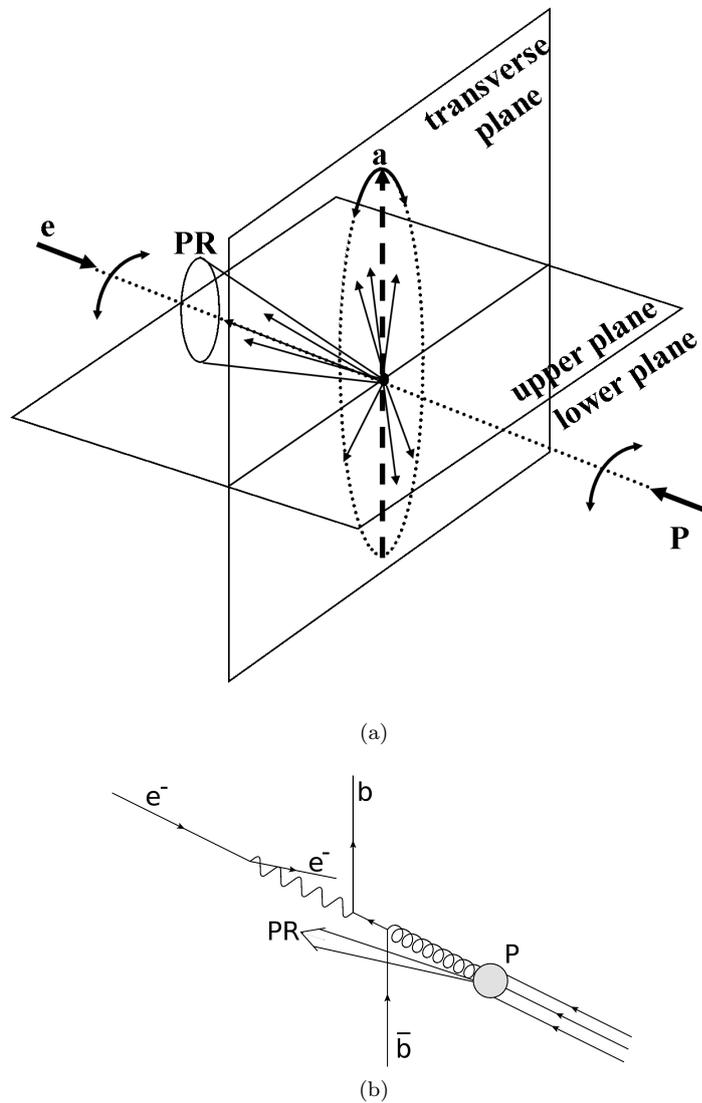


Figure 8.1: Illustration of the Thrust Axis in the Transverse Plane
 Figure (a) schematically illustrates the determination of the thrust axis in the transverse plane. The transverse thrust axis is the axis in the transverse plane along which the maximum alignment is found. The dashed line in the figure sketches the thrust axis in the transverse plane **a** and the arrows the decay products. Decay products in the forward direction often are related to the proton remnant indicated by PR. The thrust axis subdivides the event into two hemispheres, each containing the decay products of a beauty quark.
 The Figure (b) is a boson-gluon like diagram that illustrates the topology of beauty production in ep-collisions at HERA.

$m_{T,est}(b)$ [GeV]	0	5.5	7.5	9.0	10.5	14.0	-
$\langle m_T(b) \rangle$ [GeV]	4.75	6.11	7.08	8.30	9.93	12.84	20.0

Table 8.1: Binning

Chosen binning for the reconstructed transverse beauty mass $m_{T,est}(b)$ and the mean generated transverse beauty mass $\langle m_T(b) \rangle$.

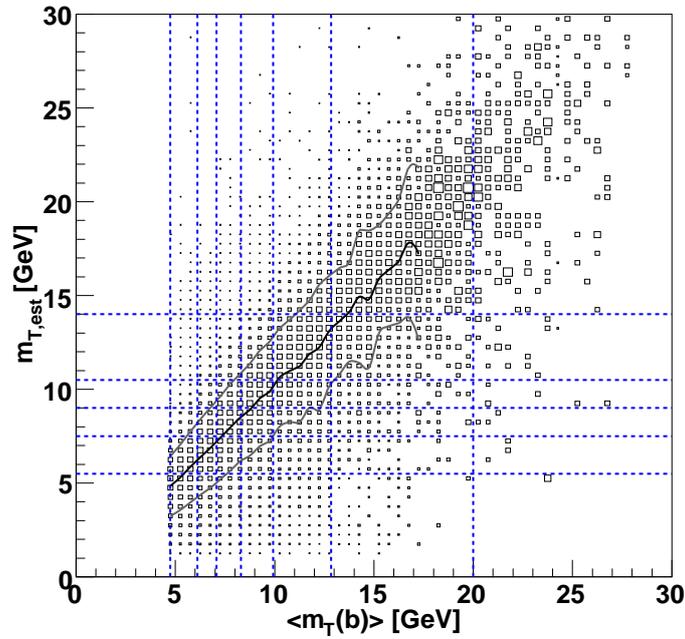


Figure 8.2: Reconstructed vs. Generated Transverse Beauty Mass. The reconstructed transverse beauty mass $m_{T,est}(b)$ is correlated with the mean generated transverse beauty mass $\langle m_T(b) \rangle$. The line on the diagonal indicates the mean of $m_{T,est}(b)$ in each slice of $\langle m_T(b) \rangle$, and the two lines in parallel the one- σ error band. The dashed lines show the binning used for the unfolding (section 8.2). The numerical values of the chosen binning are listed in Table 8.1.

8.2 Unfolding the Differential Beauty Cross Section

A regularized unfolding procedure is used to relate the measured mean transverse beauty mass $m_{T,est}(b)$ and the measured spectra of electron pairs with the mean transverse momentum of the beauty quarks. In parallel it also determines the fractions of background contributions resulting from $J/\psi \rightarrow ee$, $c\bar{c} \rightarrow ee$ and misidentified electron candidates.

The idea of unfolding is to correct a measured distribution for detector effects and background. The measured distribution \mathbf{y} is related to the true distributions \mathbf{x} on parton level, via the matrix equation

$$\mathbf{y} = \mathbf{A}\mathbf{x} \quad , \quad (8.5)$$

where the matrix \mathbf{A} is designated as response matrix and \mathbf{x} and \mathbf{y} as true and reconstructed vector². The response matrix \mathbf{A} has to be determined from the simulation.

Unfolding is to construct an estimator $\hat{\mathbf{x}}$ for the true distribution \mathbf{x} . The naive approach of choosing a quadratic response matrix and simply inverting it, i.e. $\mathbf{x} = \mathbf{A}^{-1}\mathbf{y}$, often leads to a catastrophic highly fluctuating result, since \mathbf{A} and \mathbf{y} are subject to statistical fluctuations [99]. Problems in this regard are overcome by prior assumptions concerning smoothness.

In regularized unfolding³ the estimator $\hat{\mathbf{x}}$ is chosen such that the $\chi^2(\hat{\mathbf{x}})$ given by

$$\chi^2(\hat{\mathbf{x}}, \tau, \mu) = \chi_A^2(\hat{\mathbf{x}}) + \tau \cdot \chi_L^2(\hat{\mathbf{x}}) + \mu \cdot \chi_N^2(\hat{\mathbf{x}}) \quad (8.6)$$

$$\chi_A^2(\hat{\mathbf{x}}) = 1/2 (\mathbf{y} - \mathbf{A}\hat{\mathbf{x}})^T V^{-1} (\mathbf{y} - \mathbf{A}\hat{\mathbf{x}}) \quad (8.7)$$

$$\chi_L^2(\hat{\mathbf{x}}) = \hat{\mathbf{x}}^T L \hat{\mathbf{x}} \quad (8.8)$$

$$\chi_N^2(\hat{\mathbf{x}}) = \left(n_{obs} - \sum_{j=1}^N (\mathbf{A}\hat{\mathbf{x}})_j \right)^2 \quad , \quad (8.9)$$

is minimized. $V_{i,j} = cov(y_i, y_j)$ is the covariance matrix of the data, L the regularization matrix, n_{obs} the number of observed events and τ, μ Lagrange multipliers (the variable τ is often referred to as regularization parameter).

$\chi_A^2(\hat{\mathbf{x}})$ is a standard χ^2 definition that minimizes the deviation of $\mathbf{A}\hat{\mathbf{x}}$ from the measured vector \mathbf{y} . $\tau \cdot \chi_L^2(\hat{\mathbf{x}})$ imposes smoothness constraints, and $\mu \cdot \chi_N^2(\hat{\mathbf{x}})$ ensures the total normalization. In this analysis μ is chosen such that $|n_{obs} - \sum_{i=1}^N (\mathbf{A}\hat{\mathbf{x}})_i| / n_{obs} < 10^{-3}$.

Appendix A gives an overview on the concept of regularized unfolding. The equations (8.6)-(8.9) and some arguments and methods used within this chapter are discussed therein in a broader context.

The discussion in this chapter is in the following restricted to a) the determination of the response matrix \mathbf{A} of this analysis and b) the extraction of the differential beauty cross section $\frac{d\sigma}{d\langle p_T(b) \rangle}$.

8.2.1 Definition of the Response Matrix

Figure 8.3 shows a sketch of the response histogram, corresponding to the response matrix \mathbf{A} . The detailed structure of it, as well as the structure of the vectors \mathbf{y} and \mathbf{x} are explained below. The construction of the estimator $\hat{\mathbf{x}}$ for the true distributions \mathbf{x} incorporates prior assumptions on the composition of the measured spectrum and on the smoothness of the deconvoluted distributions.

²All background contributions considered are incorporated in the response matrix \mathbf{A} . The vector \mathbf{b} of equation (A.9) is therefore set to 0. Cf. with the corresponding discussion in section A.1.

³See Appendix A and the references therein.

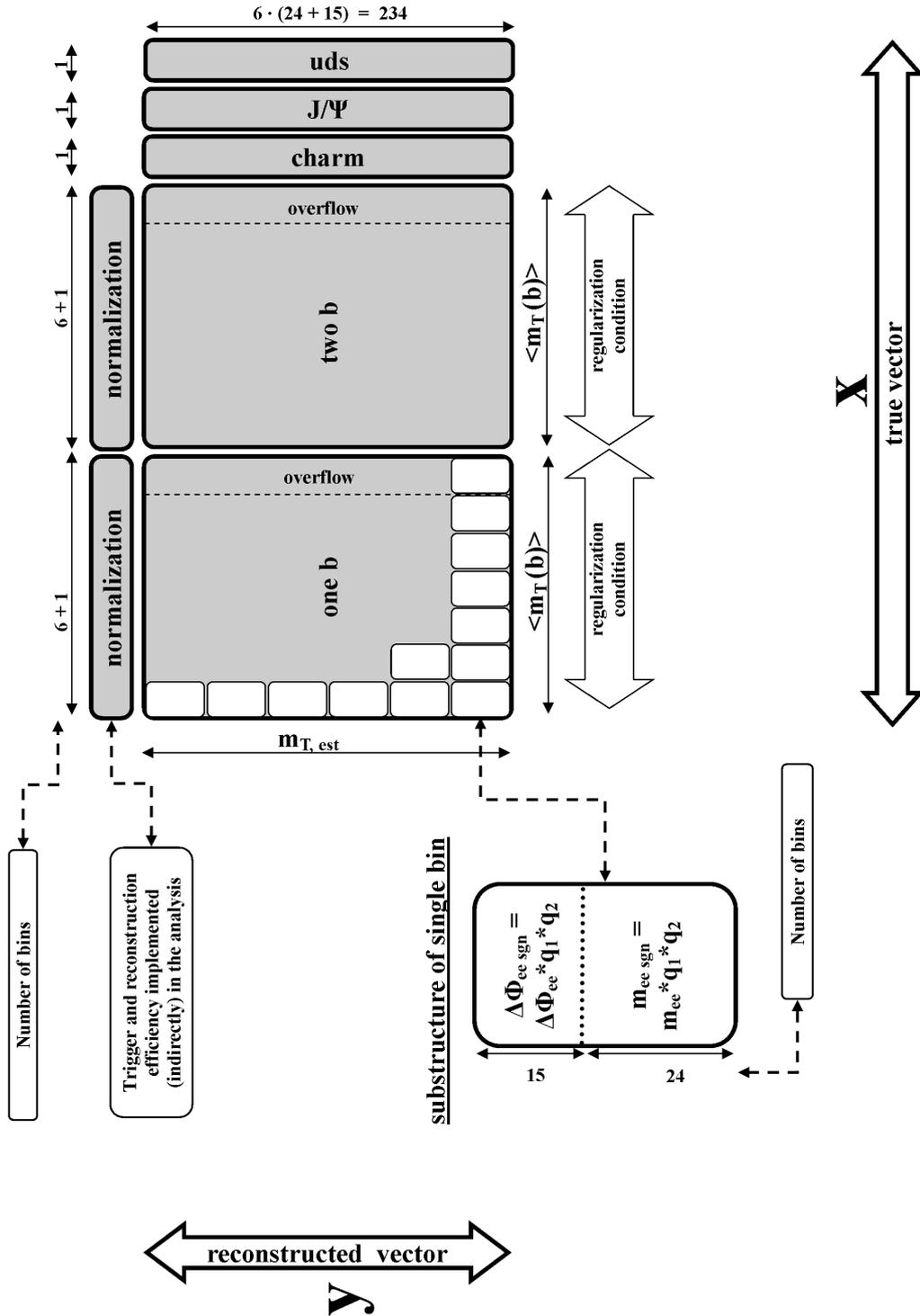


Figure 8.3: Sketch of the Response Histogram

The structure of the response histogram is sketched, used to determine the response matrix \mathbf{A} of the unfolding problem $\mathbf{y} = \mathbf{A}\mathbf{x}$. Compare with section 8.2.1 for the details.

Composition of the Measured Spectrum and Structure of the True Vector \mathbf{x}

It is assumed that the measured spectrum of di-electron candidates is a composition of the following five true distributions:

signal Events containing a $b\bar{b}$ pair and at least two tagged electron candidates. These events are further subdivided into:

one b Events where the electron candidates originate from the same beauty quark.

two b Events where one electron candidate originates from the b quark and the other electron candidate from the \bar{b} quark.

background All events with at least two tagged electron candidates, but containing no beauty quarks are treated as background. These events are classified into:

J/ψ Events containing J/ψ decaying into two electron candidates, $J/\psi \rightarrow e^+e^-X$.

charm Events where the electron candidates originate from charm quarks, but not from J/ψ . These are mainly events from semileptonic charm decays $c\bar{c} \rightarrow e^+e^-X$.

uds Events only containing light quarks (u,d,s). These events are mainly tagged due to two misidentified electron candidates.

The vector \mathbf{x} consists of two signal contributions **one b** and **two b** with 6+1 bins each and three backgrounds (1 bin each) for the contributions **charm**, **J/ψ** and **uds** background⁴:

$$\mathbf{x} = (\mathbf{x}_{\text{one b}}^{(6+1)}, \mathbf{x}_{\text{two b}}^{(6+1)}, \mathbf{x}_{\text{charm}}^{(1)}, \mathbf{x}_{J/\psi}^{(1)}, \mathbf{x}_{\text{uds}}^{(1)}) . \quad (8.10)$$

The signal contributions are binned in the mean transverse beauty mass $\langle m_T(b) \rangle$, such that from the unfolded result (i.e. the vector \mathbf{x}) differential cross sections in six bins of $\langle p_T(b) \rangle$ can be obtained. The seventh bin of $\langle m_T \rangle$ is used as an overflow bin for events with $\langle m_T(b) \rangle > 20$ GeV (see Table 8.1) and is not used for the final result.

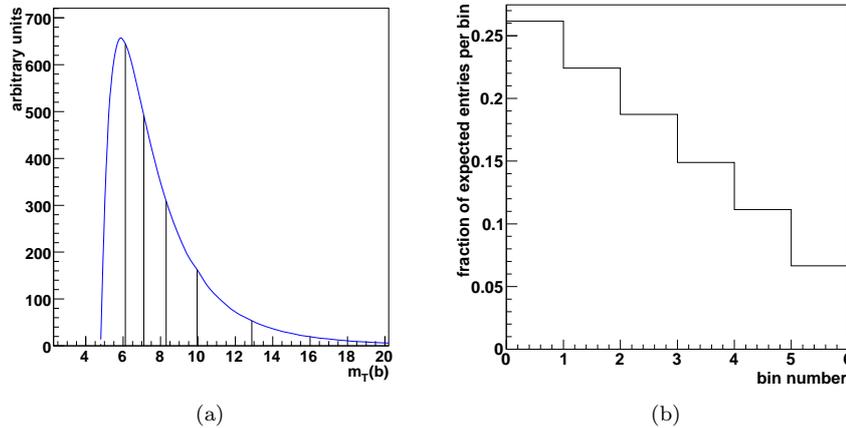


Figure 8.4: Binning of $\langle m_T(b) \rangle$

The left plot shows the predicted transverse beauty mass. The vertical lines show the binning used for the unfolding, which is chosen such, that the number of expected events per bin follow a straight line (right plot).

⁴Notation: the superscript number in parenthesis refers to the number of bins.

Smoothness Constraints

Smoothness constraints for the unfolding are applied on the signal bins of \mathbf{x} : The second derivative for the bins corresponding to **one b** and the second derivative for the bins corresponding to **two b** are each minimized (cf. equation (A.17) and (A.18) and Figure 8.3). To ensure this constraint, a non equidistant binning in $\langle m_T(b) \rangle$ is chosen such, that the expected entries in the corresponding histograms follow a straight line. The procedure is illustrated in Figure 8.4. The obtained binning is drawn in Figure 8.2 (vertical dashed lines) and indicated in the Table 8.1.

This smoothness constraints reduce the fluctuations of the result, but have no implications on the scale and the slope of the obtained distributions.

Structure of the Reconstructed Vector \mathbf{y}

The measured observables to construct the vector \mathbf{y} are the estimator for the transverse beauty mass $m_{T,est}$ (cf. equation (8.2)), the signed invariant mass of the two electron candidates $m_{e_1,e_2,sgn} = m_{e_1,e_2} * (q_{e_1} \cdot q_{e_2})$ (i.e. the invariant two electron mass multiplied by the charge of each electron candidate) and the signed angle between the two electron candidates in the transverse plane $\Delta\phi_{e_1,e_2,sgn} = \Delta\phi_{e_1,e_2} * (q_{e_1} \cdot q_{e_2})$.

The latter two variables both discriminate between the contributions **one b**, **two b**, **charm**, **J/ψ** and **uds**. For the variable $\Delta\phi_{e_1,e_2,sgn}$ this is illustrated in Figure 8.5. For the variable $m_{e_1,e_2,sgn}$ the charge product of the two electron candidates classifies the events as discussed in Figure 8.5. In addition m_{e_1,e_2} discriminates electron signal from background, since electrons from true b-quarks produce the largest invariant mass m_{e_1,e_2} . In the case of J/ψ events m_{e_1,e_2} peaks at the J/ψ mass.

The precise structure of the vector \mathbf{y} are six bins of $m_{T,est}$, further subdivided into 24 bins of $m_{e_1,e_2,sgn}$ in the range ± 10.4 GeV and into 15 bins of $\Delta\phi_{e_1,e_2,sgn}$ in the range $\pm\pi$ (cf. with Figure 8.3):

$$\mathbf{y} = \left(\underbrace{\left(m_{e_1,e_2,sgn}^{(24)}, \Delta\phi_{e_1,e_2,sgn}^{(15)} \right)}_{m_{T,est}^{(1)}}, \underbrace{\left(m_{e_1,e_2,sgn}^{(24)}, \Delta\phi_{e_1,e_2,sgn}^{(15)} \right)}_{m_{T,est}^{(2)}}, \dots, \underbrace{\left(m_{e_1,e_2,sgn}^{(24)}, \Delta\phi_{e_1,e_2,sgn}^{(15)} \right)}_{m_{T,est}^{(6)}} \right) \quad (8.11)$$

Structure of the Response Matrix \mathbf{A}

The response matrix \mathbf{A} is composed of 17×234 bins, corresponding to the $6 \cdot (24 + 15) = 234$ bins of \mathbf{y} and the $7 + 7 + 1 + 1 + 1 = 17$ bins of \mathbf{x} . It is composed of 5 submatrices each for a considered contribution **one b**, **two b**, **charm**, **J/ψ** and **uds** :

$$\mathbf{A} = (\mathbf{A}^{\text{one b}}, \mathbf{A}^{\text{two b}}, \mathbf{A}^{\text{charm}}, \mathbf{A}^{\text{J/}\psi}, \mathbf{A}^{\text{uds}}) \quad (8.12)$$

The estimator for the transverse beauty mass $m_{T,est}$ is used for the unfolding of $\langle m_T(b) \rangle$, which afterwards can be directly converted to $\langle p_T(b) \rangle$ for the extraction of $\frac{d\sigma}{d\langle p_T(b) \rangle}$. The chosen substructure in \mathbf{y} produces different submatrices $\mathbf{A}^{\text{one b}}$, $\mathbf{A}^{\text{two b}}$, $\mathbf{A}^{\text{charm}}$, $\mathbf{A}^{\text{J/}\psi}$, \mathbf{A}^{uds} , which ensures the separation of the different signal and background contributions during the unfolding.

In addition, the chosen substructure of \mathbf{y} also contributes to the unfolding of the transverse beauty mass $\langle m_T(b) \rangle$, since $m_{e_1,e_2,sgn}$ and $\Delta\phi_{e_1,e_2,sgn}$ also depend on $\langle m_T(b) \rangle$.

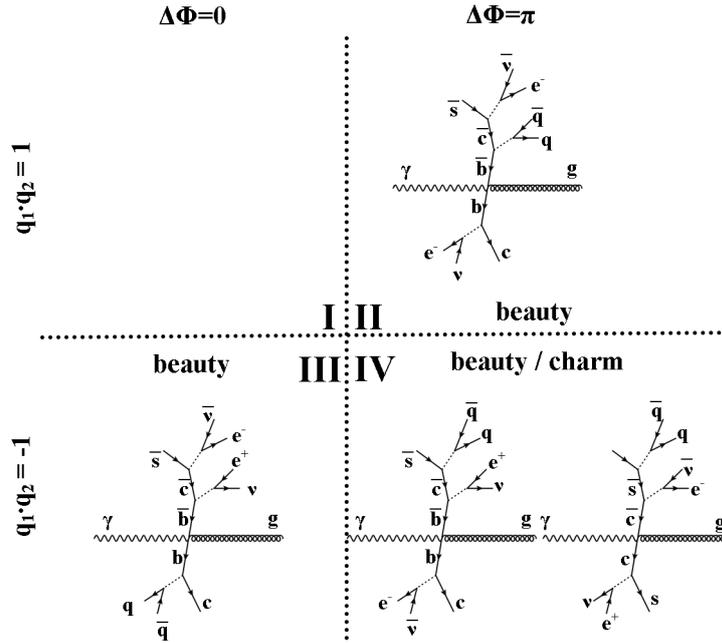


Figure 8.5: Illustration of the Variable $\Delta\phi_{e_1,e_2,sgn}$

The Feynman diagrams of the Figure show charm respectively beauty events with two electrons in the final state. The diagrams are arranged according to the charge product of the two electrons $q_{e_1} \cdot q_{e_2}$ and the angle between the electrons $\Delta\phi$. Charm events populate only in sector IV, beauty events in the sectors II, III and IV. Events with misidentified electrons can occur in all sectors.

8.2.2 Filling of the Response Matrix

The response matrix \mathbf{A} is computed using a combination of the Pythia (in the massive and massless mode) and the Cascade event generator (the used event samples are summarized in section 5.2). All correction factors due to limited acceptances and detector inefficiencies are based on the detector response and the trigger simulation, discussed in section 5.1.3. The systematic uncertainties due to differences between the models and due to the trigger simulation are estimated and propagated to the final result (cf. with the discussion in section 8.4).

The simulated events filled into the reconstructed part of the response matrix \mathbf{A} all have to pass the final di-electron selection, discussed in section 7.5.2 and listed in Table 7.4. The elements of \mathbf{A} are determined in the following manner:

- the submatrices $\mathbf{A}^{\text{one } \mathbf{b}}$ and $\mathbf{A}^{\text{two } \mathbf{b}}$ are computed from the two Beauty simulation samples generated with Pythia in the massive mode⁵ and Cascade. The two samples are added luminosity-normalized.
- the submatrix $\mathbf{A}^{\text{charm}}$ is extracted from the Cascade simulation, because for charm Cascade is known to describe the shape of the differential charm cross section as function of the transverse momentum of the charm quark quite well [92].
- the submatrix $\mathbf{A}^{\text{J}/\psi}$ is based on the Cascade simulation, since it is known from a previous measurement that Pythia does not describe the data in this decay channel very well [94].
- the submatrix \mathbf{A}^{uds} is determined from the Pythia simulation in the massless mode. As described in section 7.6, the resolved part of the MC is scaled by a factor 0.3 (see also the sections 5.2 and 8.4).

The normalization bins are computed according to equation (A.11). They normalize the y columns of the response matrix such, that the result is extrapolated to the kinematic cuts listed in the Table 4.1.

Table 8.2 summarizes the composition of the MC samples, and in Figure 8.6 the computed response histogram is presented. Clearly visible in the response histogram are along the x -axis the different contributions of the generated vector \mathbf{x} and along the y -axis the substructure of the reconstructed vector \mathbf{y} . Further recognizable is for the two signal contributions **one \mathbf{b}** and **two \mathbf{b}** the correlation given by $m_{T,est}$ and $\langle m_T(b) \rangle$ and the dependence of $m_{e_1, e_2, sgn}$ and $\Delta\phi_{e_1, e_2, sgn}$ on $\langle m_T(b) \rangle$, where both effects ensure the unfolding of $\langle m_T(b) \rangle$.

8.2.3 Determination of the Regularization Parameter

The regularization parameter τ (see equation 8.6) is chosen such that the correlations between the unfolded bins are minimized. This is achieved by minimizing the average of the global correlation coefficients $\langle \rho \rangle$, as explained in the Appendix A.3.1.

Figures 8.7(a) and (c) illustrate the correlation between the \mathbf{x} bins and motivate the choice of the regularization parameter τ . The choice of the regularization parameter τ influences the number of degrees of freedom (NDF) of the unfolding problem. For $\tau = 0$ the effective NDF is 17, corresponding to the 17 bins of the vector \mathbf{x} , that are successively reduced the higher the chosen value for τ is, or rather the more the smoothness constraints are applied to the unfolding problem (see Figures 8.7(b)).

⁵The Pythia sample is composed of a direct and a resolved contribution that are added luminosity weighted.

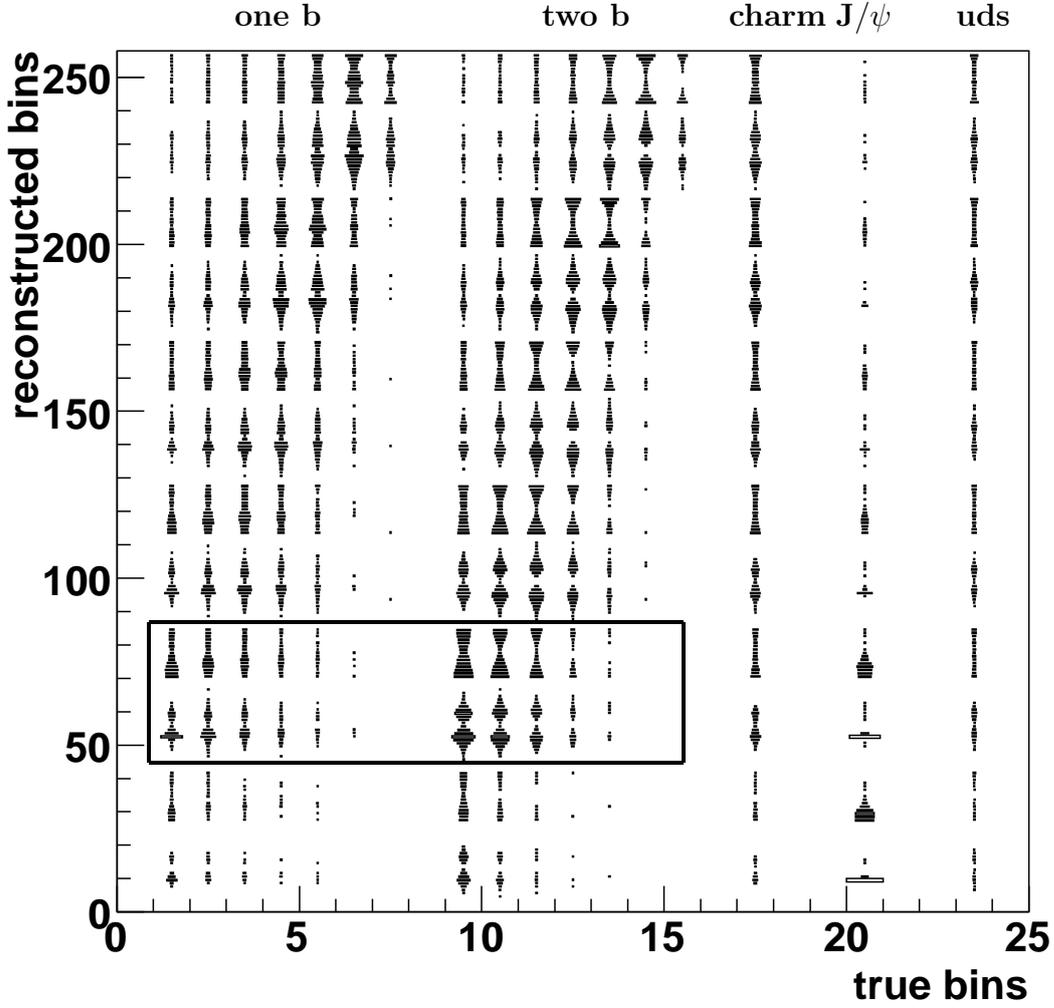


Figure 8.6: Response Histogram

The plot shows the response histogram used for the unfolding. The units are bin numbers. The different contributions of the response matrix are separated by empty bins for a better visibility (the number of filled bins is 17×234). The normalization bins are suppressed. The size of the bin reflects the number of entries in it. Recognizable in the box: the dependence of $m_{e_1, e_2, sgn}$ and $\Delta\phi_{e_1, e_2, sgn}$ on $\langle m_T(b) \rangle$. Compare also with Figure 8.3 for the structure of the histogram.

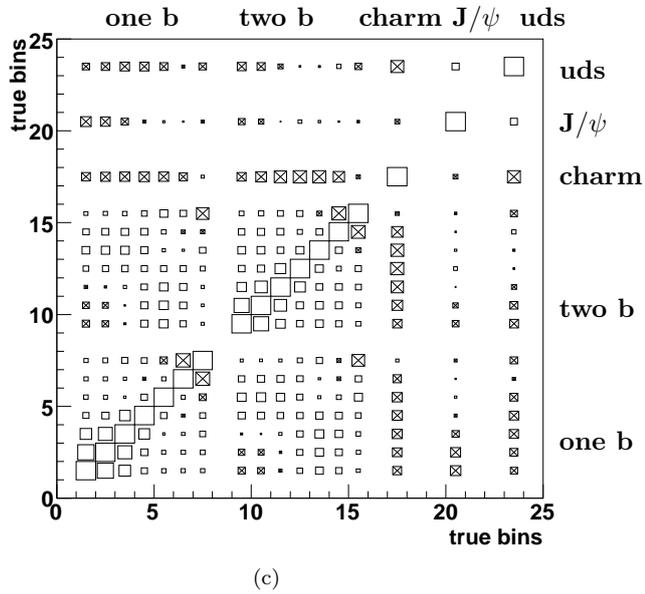
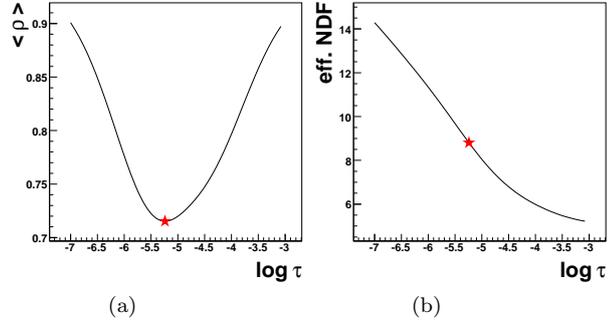


Figure 8.7: Control Distributions for the Unfolding

Figure (a) shows the average of the global correlation coefficient $\langle \rho \rangle$ as function of $\log(\tau)$, with τ being the regularization parameter. τ is chosen in the minimum of $\langle \rho \rangle$ (red star), where the correlation among the bins of the vector \mathbf{x} (true bins) are minimized (see Figure 8.3). Figure (b) shows the number of effective degrees of freedom of the whole unfolding problem as function of $\log(\tau)$. The higher the value of τ is chosen, the more the whole unfolding problem is constraint to the theory prediction. The red star corresponds to the determined value of τ . In Figure (c) the correlations between the bins of the vector \mathbf{x} are shown. The size of a box reflects the bin value and a cross in the box a negative sign.

8.3 Final Distributions

In this section properties of relevant observables are studied and compared to the Monte Carlo simulation samples from which the response matrix is computed. The decomposition into the different signal and background contributions is obtained from the unfolded vector $\hat{\mathbf{x}}$. Since in the applied unfolding procedure the normalization is fixed, only the shape of the obtained distributions are comparable. In Figures 8.8(a)-(d) distributions for the electron candidates are presented. The main characteristics of the distribution $m_{e_1, e_2, sgn}$, $\Delta\phi_{e_1, e_2, sgn}$, $p_T(e)$ and $\vartheta(e)$ are well described by the Monte Carlo simulation.

The rise of the electron candidates in the forward region (Figure 8.8(d)) occurs in the transition region between the central and the forward barrel of the calorimeter. It indicates a slight difference in the efficiency of the electron identification in the two parts of the detector. However, this difference is described by the simulation. All corrections referring to this effect are therefore taken into account by the unfolding procedure.

Further control distributions are presented in Figures 8.9(a)-(e): The longitudinal energy flow $\sum(E - P_z)$, the z -Vertex distribution and the p_T -distributions for the three highest p_T -tracks.

Agreement between the data and the Monte Carlo simulation is obtained in these distributions. It should be noticed that serious differences in the detector simulation relevant to this measurement would be visible in these distributions. Differences in the trigger simulation would occur in the track p_T -distributions.

8.4 Systematic Uncertainties and Bias Correction

The systematic uncertainties of the cross section measurement are determined by evaluating variations of the response matrix \mathbf{A} due to certain systematic effects. The resulting differences in the response matrix \mathbf{A} are translated into the systematic error of the measurement by means of error propagation of every single element of \mathbf{A} . The resulting cross section error should be interpreted as the uncertainty, when varying the input distribution within its 1σ -error band.

The effects being considered are summarized in Table 8.3 and are discussed in the following:

Electron Identification

In order to determine the variation of the response matrix \mathbf{A} due to the systematic uncertainty arising from the electron identification by means of the NN, the 1σ -uncertainty on the NN output variable MLP has to be estimated.

The uncertainty on MLP is determined by comparing an almost clean electron sample in data obtained from elastic $J/\psi \rightarrow ee$ events, with simulated electrons from an elastic $J/\psi \rightarrow ee$ Monte Carlo simulation. Electrons are identified in the data sample by means of dE/dx . The resulting invariant mass spectrum calculated from the two electron candidates is shown in Figure 8.10(a) (as points) together with the corresponding Monte Carlo sample (histogram). The two distributions are normalized to the peak height. In order to further purify the data sample only electron candidates with an invariant mass in the window $2.9 \text{ GeV} < m_{ee} < 3.2 \text{ GeV}$ are selected, which corresponds to the gray colored band in the plot drawn around the J/ψ -mass. Figure 8.10(b) shows the output of the electron identification variable MLP for these events. Over the full range between $MLP = -1$ and $MLP = 1$ the data (points) is reasonably well described by the simulation (histogram). Entries at $MLP = -1$ are mostly events that are close to a φ - or z -crack in the LAr calorimeter, where the deposited particle shower leaks into an inactive part of the LAr

calorimeter. Figure 8.10(c) is an enlargement of Figure 8.10(b) around the MLP selection value of the analysis, $MLP > 0.9125$. A small shift ΔMLP between the two distributions becomes visible that can be quantified by the following χ^2 definition:

$$\chi^2(MLP, \Delta MLP) = \sum_{\text{bin } i=i_0 \pm 30} \frac{(\text{hist}_i^{\text{data}} - \text{hist}_{i'}^{\text{MC}})^2}{(\sigma_i^{\text{data}})^2} \quad (8.13)$$

$$\Delta MLP = (i - i') \cdot \text{binwidth}$$

$$MLP = \text{bin center of bin } i_0$$

$$\text{hist}_i^{\text{data/MC}} = \text{bin content of bin } i \text{ of data (MC) histogram}$$

The sum is performed over 60 bins around a central bin with the bin number i_0 and with a bin center value MLP. $\text{hist}_i^{\text{data}}$ ($\text{hist}_{i'}^{\text{MC}}$) is the bin content of the data (MC) histogram in bin i (i'). The Monte Carlo histogram is shifted with respect to the data histogram by $i' - i$ bins, respectively ΔMLP . Figure 8.10(d) shows χ^2 as function of ΔMLP , with χ^2 computed around $MLP = 0.9125$ and fitted with a polynomial of degree 4. From the minimum of the polynomial fit the bias between the data and the MC histogram is determined, from the increase of χ^2 by 1 the one sigma uncertainty. Figure 8.10 (e) shows the minimum of the χ^2 (solid line) and the one sigma uncertainty band (dotted line) as function of different MLP values.

In order to incorporate the resulting bias in the analysis, the cut value on the MLP-discriminator is shifted accordingly for all Monte Carlo samples. Around this new value the MLP discriminator is varied within the one sigma uncertainty, to evaluate the corresponding systematic uncertainty on the response matrix \mathbf{A} . The arrows drawn in Figure 8.10 (e) indicate the bias correction and the systematical uncertainty on MLP at the cut value of the analysis, $MLP = 0.9125$. The maximal relative difference on each entry of \mathbf{A} between the default response matrix and the newly evaluated ones is determined and taken as the relative uncertainty on the response matrix \mathbf{A} due to the electron identification.

The mean relative systematical error per bin of the final cross section due to the uncertainty on the electron identification is determined to be **6%**.

Trigger Efficiency

The Level 1 and Level 2 trigger simulation of the FTT were studied in many other analyses [73, 75, 88, 92, 93, 95]. The discrepancy between data and the Monte Carlo simulation are on the percent level and are not considered in the following. The Level 3 Trigger simulation is based on exactly the same program code as was loaded in the L3 Trigger Hardware, differences due to this part of the trigger simulation are therefore also negligible (see Appendix D.3). Even though a lot of time and effort was spent within this thesis to calibrate the Jet Trigger simulation, it is known not to be perfect. The biggest differences in the trigger simulation from the data are expected due to the calibration of the JT simulation. The calibration constants for the JT simulation were determined from data, and were identically applied in the trigger simulation of all MC samples (see Appendix C).

The approach chosen to determine the systematic uncertainty of the measurement due the calibration of the JT simulation is similar to the one discussed in the last section. In a first step the uncertainty of the calibration constants is estimated, which is afterwards translated via the uncertainty of the response matrix \mathbf{A} to the corresponding uncertainty of the measurement.

To quantify the uncertainty of these calibration constants, the single electron L3 trigger efficiency obtained from $J/\psi \rightarrow ee$ events is compared to the single electron L3 trigger efficiency obtained from Single Particle Monte Carlo samples. For the MC samples, the L3 trigger

efficiency is computed for different variations of the calibration constants, by globally scaling the default calibration constants up and down. In Figure 8.11 the single electron L3 trigger efficiency obtained from data is compared to the single electron L3 efficiency obtained from MC, as function of the scaling factor on the calibration constants. Figure (a) and (c) present the L3 trigger efficiencies for the single tag medium p_T electron trigger s29, whereas the Figures (b) and (d) are the corresponding plots for the single tag high p_T electron trigger s34. The precision for the later plots is reduced, because less data events are above the trigger threshold. In the two upper plots ((a) and (b)) the comparison of all data points above the high p_T trigger threshold is performed by a standard χ^2 definition, in the lower plots the mean trigger efficiencies above the threshold are compared directly to each other.

The uncertainty on the calibration constants was estimated to be on the order of $\pm 10\%$, indicated by the straight line in the two lower plots. Based on this, the trigger efficiency of all MC samples is resimulated twice, with all calibration constants scaled up and down by 10%. From the resimulated MC samples the response matrix \mathbf{A} is recomputed for both variations. For every entry in \mathbf{A} the maximal relative difference between the default response matrix to the re-simulated ones is determined, and propagated to the data points of the measurement. The mean relative systematical error per bin of the final cross section due to the uncertainty on the trigger efficiency is determined to be **12%**.

Model Dependence - Beauty

In the signal part (**one b, two b**), the response matrix is computed by the luminosity weighted sum of the Cascade and Pythia massive Beauty Monte Carlo samples, as discussed in section 8.2.2. The uncertainty due to the two models is determined by two alternative response matrices, each based on one of the Monte Carlo samples. From these two samples the relative maximal difference in shape with respect to the default response matrix is computed for each entry of the matrix, and taken as the relative uncertainty of the response matrix.

The mean relative systematical error per bin of the final cross section due to the beauty model is determined to be **4%**.

Model Dependence - Charm

The uncertainty for the charm contribution is evaluated from the relative difference in shape between the (default) Cascade charm Monte Carlo sample and the charm contribution in the Pythia massless Monte Carlo sample.

The mean relative systematical error per bin of the final cross section due to the charm model is determined to be **14%**.

Model Dependence - Light Flavors

The bins in the response matrix corresponding to the light flavors are determined from the Pythia massless Monte Carlo, where the resolved part is downscaled by a factor 0.3, as discussed in section 8.2.2 and 5.2. The differences in the shape of the distributions between the default settings of Pythia, and the one obtained by the downscaled resolved contribution, are treated as systematic uncertainty.

The mean relative error per bin due to the uncertainty of the relative fraction between the resolved and direct contribution in the inclusive MC sample is **2%**.

Tracking Efficiency

The uncertainty on the track finding efficiency is on the level of 2% per track. The unfolding is based on the spectra of two electron candidates. The tracking efficiency therefore translates into an uncertainty of 4%, adding the systematical uncertainty linearly.

Luminosity

For the data period analyzed the luminosity is known to an accuracy of 5%. This uncertainty is directly added to the total error of the determined cross section and not translated into an uncertainty of the response matrix \mathbf{A} . The luminosity enters the analysis after the unfolding, cf. with equation (8.14).

8.5 Results and Discussion

8.5.1 Determination of the Signal and Background Contributions

The fraction of signal and background in the 3069 selected electron pairs in data (cf. section 7.5.2), can be determined from the result of the unfolding procedure, i.e. by combining the corresponding elements of the vector $\hat{\mathbf{x}}$. Table 8.4 lists the determined number of beauty, charm, J/ψ and uds events and the resulting fractions. The indicated errors are split into correlated and uncorrelated errors, whereas the error matrix E of the unfolding was split into a correlated and an uncorrelated error matrix $E = E_{uncorr} + E_{corr}$ [100], before combining the associated elements.

The correlation between the different considered contributions is given in Table 8.5. Especially charm is strongly anti-correlated to beauty and therefore almost not separable from the signal. However, the resulting charm contribution overshoots the prediction of the used simulation (Cascade) by a factor 1.35, which is in agreement with [92]. The J/ψ and uds backgrounds also agree with the used predictions, whereas J/ψ overestimates the simulation by a factor 1.6 and uds is underestimated by a factor 0.7.

8.5.2 Differential Cross Section as Function of the Mean Transverse Momentum of the Beauty Quarks

The extracted number of events for each signal bin determined by the unfolding procedure $N_{sig,i} = \hat{\mathbf{x}}_i$, is translated into the average differential cross section in each bin by

$$\frac{d\sigma}{d\langle p_T(b) \rangle} = \frac{N_{sig,i}}{\mathcal{L} \cdot \Delta p_{T,i}(b)} \quad , \quad (8.14)$$

where \mathcal{L} is the luminosity and $\Delta p_{T,i}(b)$ the bin width. The unfolded differential cross sections $\frac{d\sigma}{d\langle p_T(b) \rangle}$ for $b\bar{b} \rightarrow eeX$ in the visible range $0.2 < y < 0.8$, $Q^2 < 1 \text{ GeV}^2$, $|\eta(b), \eta(\bar{b})| < 2$ is presented in Figure 8.12 for the two signal contributions, where the two electron candidates result from the same b quark **one b** (a), and where one electron results from the b quark and the second one from the \bar{b} quark **two b** (b). The inner error bars indicate the uncorrelated error, the full error bars contain in addition the correlated errors added in quadrature. Also shown are the predictions calculated with the Pythia and Cascade simulation that were used for the unfolding.

Together twelve bins are shown. On the other hand the estimator \hat{x} is composed of seventeen bins, including the two overflow bins and the three bins of the background contributions

Contribution in A	Generator	Filled in A	Sys. Uncert.
one beauty	Pythia massive	combined	✓
	Cascade		✓
two beauty	Pythia massive	combined	✓
	Cascade		✓
charm	Cascade	✓	-
	Pythia, massless	-	✓
J/ψ	Cascade	✓	-
uds	Pythia massless	✓	✓

*Table 8.2: MC Samples Used to Determine the Response Matrix **A***
The table lists for each contribution of the response histogram, the MC samples considered for the default computation. The last column indicates the MC samples from which the systematic uncertainty on the default was determined (cf. section 8.4).

Effect	Error
Electron Identification	6%
Trigger Efficiency	12%
Model dependence - Beauty	4%
Model dependence - Charm	14%
Model dependence - Light flavors	2%
Tracking Efficiency	4%
Luminosity	5%

Table 8.3: Summary of Systematic Errors
The relative uncertainty of the systematic errors considered.

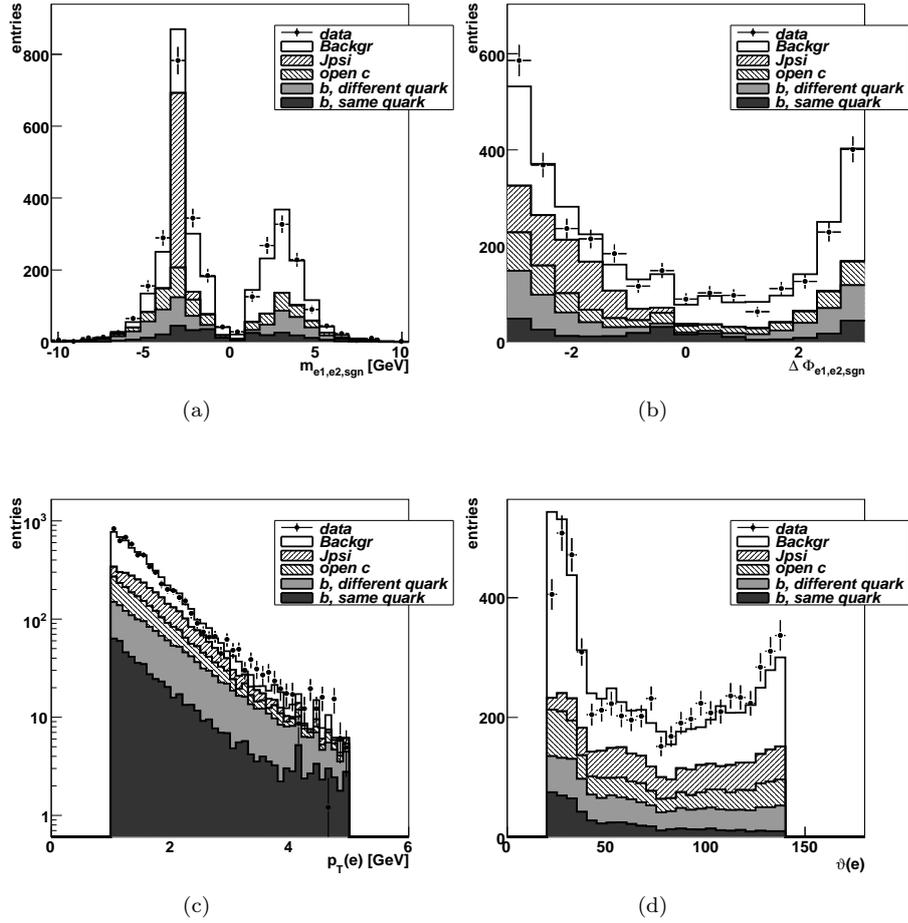


Figure 8.8: Final Distributions

Distributions for the electron candidates are shown. Data is compared to the prediction of the Monte Carlo simulation. The fraction of the different Monte Carlo contributions is determined by unfolding. Figure (a) and Figure (b) are input variables to the response histogram. (In Figure (c) and (d) two electrons are filled per electron pair.)

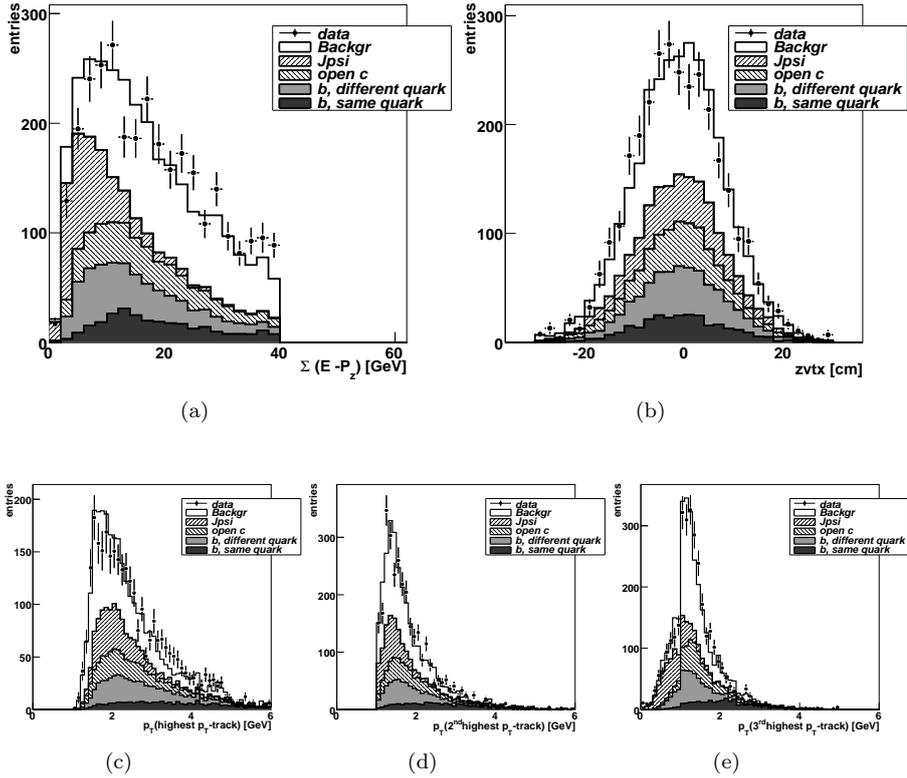


Figure 8.9: Final Distributions

Distributions for event quantities are shown. Data is compared to the prediction of the Monte Carlo simulation. The fraction of the different Monte Carlo contributions is determined by unfolding. In Figure (a) and (b) the longitudinal energy flow $\sum(E - P_z)$ and the z-Vertex distribution are shown. Figures (c)-(e) are the transverse momentum distributions for the three highest p_T tracks. These distributions are sensitive to the trigger simulation.

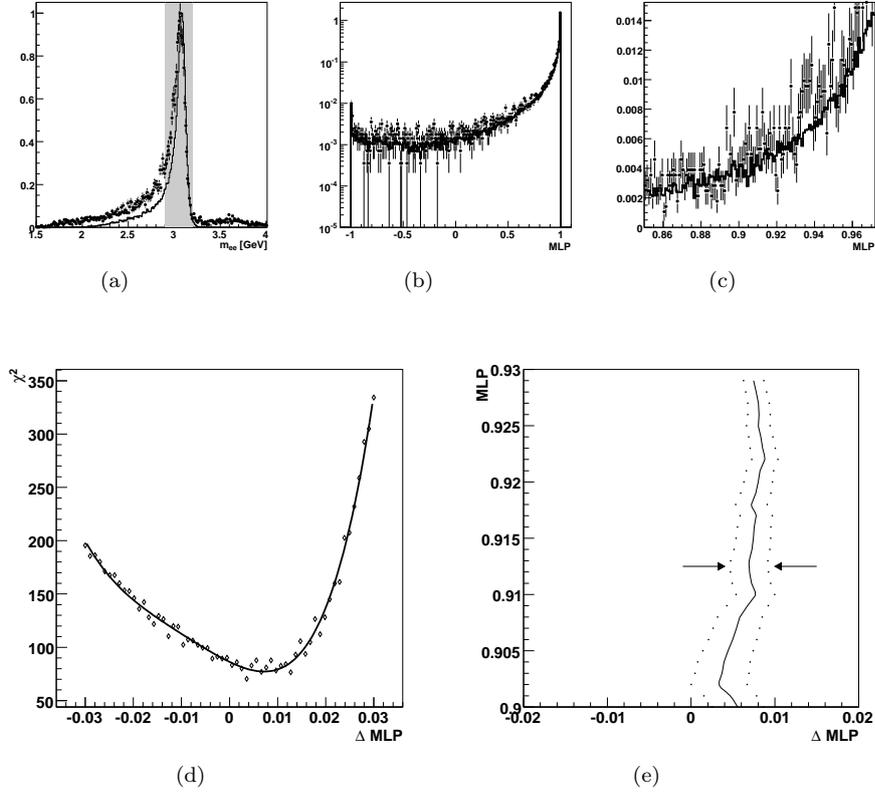


Figure 8.10: Electron Identification Systematics

The Figure contains plots that were used for the determination of the 1σ -uncertainties of the electron identification. In all Figures a clean electron data sample obtained from $J/\psi \rightarrow ee$ events is compared to electrons from a corresponding Monte Carlo sample. Figure (a) shows the invariant mass of the J/ψ sample. For all further plots the sample is restricted to the (shaded) peak region. Figure (b) compares data with MC for the output of the Neural Network variable MLP . Figure (c) is an enlargement of Figure (b) around the electron selection value ($MLP > 0.9125$) of the analysis. The simulation in general describes the data, however it is slightly shifted. This shift can be quantified via the χ^2 definition of equation (8.13). Figure (d) presents the χ^2 computed around $MLP = 0.9125$ as function of $\Delta MLP = MLP_{MC} - MLP_{data}$. The minimum of the χ^2 distribution and the one- σ uncertainty band are shown in Figure (e) for different values of MLP . The two arrows indicate the bias correction and the systematic uncertainty on the NN variable applied in the analysis.

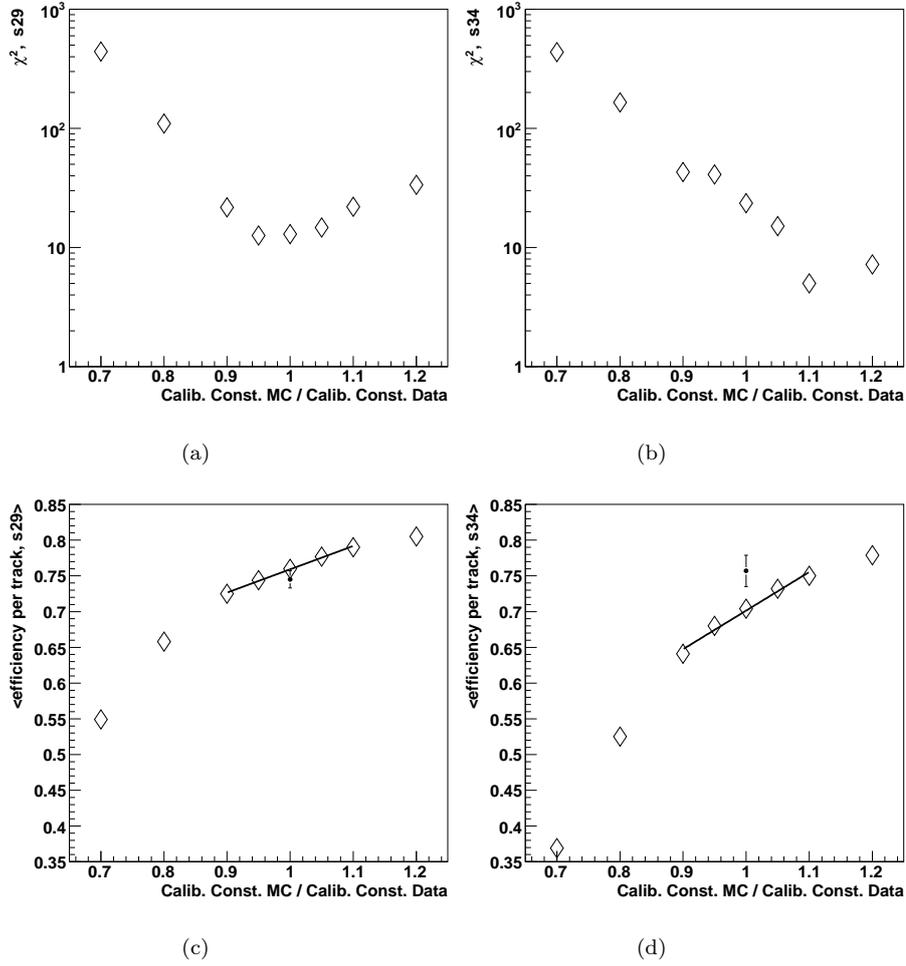


Figure 8.11: JT Calibration Systematics

In the plots (a) and (b) the single electron trigger efficiency obtained from data is compared to one determined from simulation via a standard χ^2 definition for the medium p_T (s29) and high p_T (s34) electron trigger. In the plots (c) and (d) the single track trigger efficiency obtained from simulation (diamonds), is compared to one obtained from data (points). The two straight lines indicate the region in which the calibration constants were varied for the determination of the uncertainty of the trigger efficiency.

charm, J/ψ and uds . However, the smoothness constraints introduced by the regularized unfolding reduces the effective number of degrees of freedom to the order of ten, as illustrated in Figure 8.7(b). It is therefore reasonable to combine bins and to reduce the number of represented bins.

Figure 8.14 shows the differential beauty cross section in photoproduction $d\sigma / \langle p_T(b) \rangle$ in the kinematic range specified above. The distribution is obtained by combining the two signal distributions **one b** and **two b**, and by scaling the obtained distribution for the effective branching fraction⁶ of 6.2% for a $b\bar{b}$ decaying into at least two electrons. In $\langle p_T(b) \rangle$ the bin centers of the data points are corrected according to [101]. The correction is based on the $\langle p_T(b) \rangle$ distributions of the Pythia massive⁷ simulation. The data is compared to theoretical predictions based on the LO simulation Pythia massive (dotted line) and Cascade (dashed line), and a fixed order NLO pQCD calculation (yellow shaded band) implemented in the computer program FMNR (see section 8.5.3). The lower plot shows the ratio of the measured cross section and the calculated NLO pQCD cross section, $R = d\sigma_{measured} / d\sigma_{NLO\ pQCD}$.

The shape of the presented distribution agrees well with the theoretical predictions, but is underestimated by a factor ~ 1.7 on average. However, within errors data and theory are compatible.

The uncertainty of the result is dominated by the correlated error which mostly is determined by the uncertainty on the charm background determination: This can be seen in Figure 8.13 which displays the measured correlated and uncorrelated relative errors on the three background contributions **charm, J/ψ and uds** . The correlated error on the charm contribution is on the order of 50%. On the other hand charm is highly anti-correlated with both beauty signal contributions **one b, two b**, as can be identified in the correlation plot of Figure 8.7(c). The separation of the charm and beauty is difficult, which is reflected in the high correlated error of these two contributions.

The individual data points and the associated uncertainties of the differential beauty cross section in photoproduction $d\sigma / \langle p_T(b) \rangle$ presented in Figure 8.14 and discussed in this section are listed in Table 8.6.

8.5.3 NLO pQCD Calculation

The differential beauty cross section of Figure 8.14 is compared to the NLO pQCD prediction obtained from the program FMNR [35, 102] (see section 2.8). The differential cross section was evaluated in the double differential mode and in the kinematic phase space of this measurement, listed in the upper part of Table 4.1. The differential cross section is determined as function of the mean transverse momentum of the produced beauty pair,

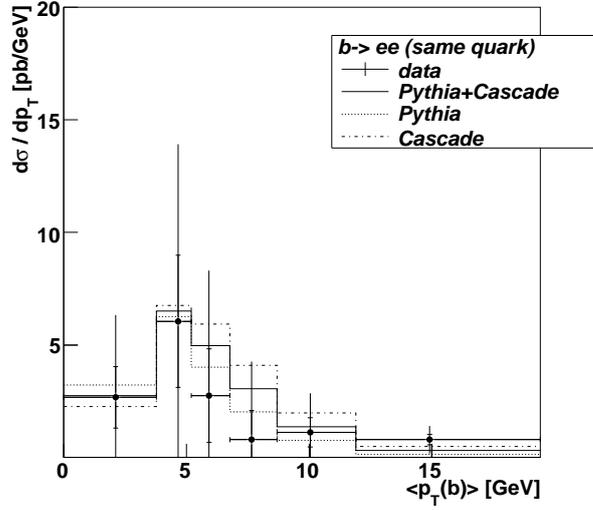
$$\langle p_T(b) \rangle = \sqrt{(p_{T,b}^2 + p_{T,\bar{b}}^2)/2}.$$

The used parton density functions are CTEQ5M [103] for the proton and GRVO-HO [104] for the photon. The renormalization μ_R and the factorization μ_F scales were parameterized to be equal, $\mu_R = \mu_F = \mu_0$, $\mu_0 = \sqrt{p_T^2 + m_b^2}/2$, where p_T is the average transverse momentum of the b quarks, and $m_b = 4.75$ GeV the beauty mass. The theoretical uncertainty, reflected in the yellow band of the Figure, were evaluated by varying the scale⁸ in the window $\mu_0/2 < \mu_0 < 2\mu_0$.

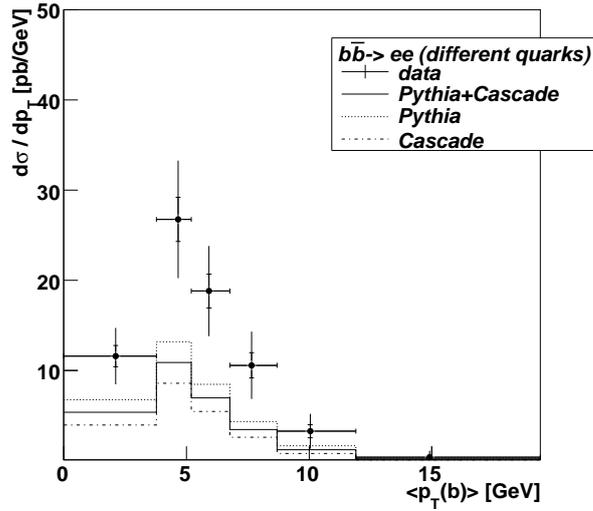
⁶The effective branching fraction $b\bar{b} \rightarrow eeX$ (see Figure 4.1) is determined from the Pythia simulation, and is in agreement with the number quoted in [66] and [97].

⁷The correction was also determined with Cascade, the differences between the two models are negligible.

⁸The theoretical uncertainty due to a variation of the beauty mass in the window $4.5 \text{ GeV} < m_b < 5 \text{ GeV}$ results at fixed μ_0 in almost the same uncertainty.



(a)



(b)

Figure 8.12: Differential Beauty Cross Sections

Differential beauty cross section obtained via the unfolding procedure for electrons having the same beauty quark mother (a), and electrons having a different beauty quark mother (b) in the kinematic range $0.2 < y < 0.8$, $Q^2 < 1 \text{ GeV}^2$, $|\eta(b), \eta(\bar{b})| < 2$. The inner error bars indicate the uncorrelated error, the full error bars contain in addition the correlated errors added in quadrature. Note that the effective number of degrees of freedom in these plots is smaller than the number of represented points (compare with the discussion in the text).

contribution	# events	uncorr. err.		corr. err.		fraction
		uncorr. err.	corr. err.	uncorr. err.	corr. err.	
beauty	773	45.3	221	0.252	0.015	0.072
charm	442	92.1	228	0.144	0.030	0.074
J/ψ	514	11.4	28	0.167	0.004	0.009
uds	1340	63	156	0.437	0.021	0.051

Table 8.4: *Unfolding Results*

The number of events and the corresponding fractions determined with the unfolding procedure are listed for the signal and background contributions.

	beauty	charm	J/ψ	uds
beauty	1	-0.68	-0.23	-0.04
charm	-0.68	1	-0.02	-0.49
J/ψ	-0.23	-0.02	1	0.33
uds	-0.04	-0.49	0.33	1

Table 8.5: *Correlation Matrix*

The correlations between the signal and background contributions considered in the unfolding procedure are listed.

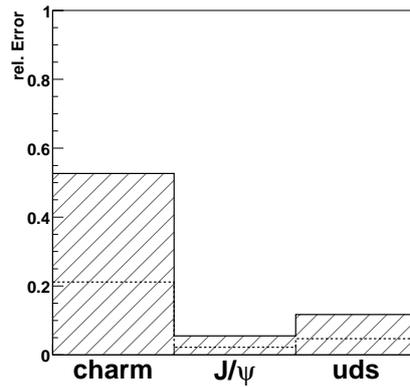


Figure 8.13: *Relative Error of the Background*

Relative correlated and uncorrelated error of the background contributions **charm**, **J/ψ** and **uds**. The correlated error is represented by the hatched histogram, the uncorrelated error by the dotted lines.

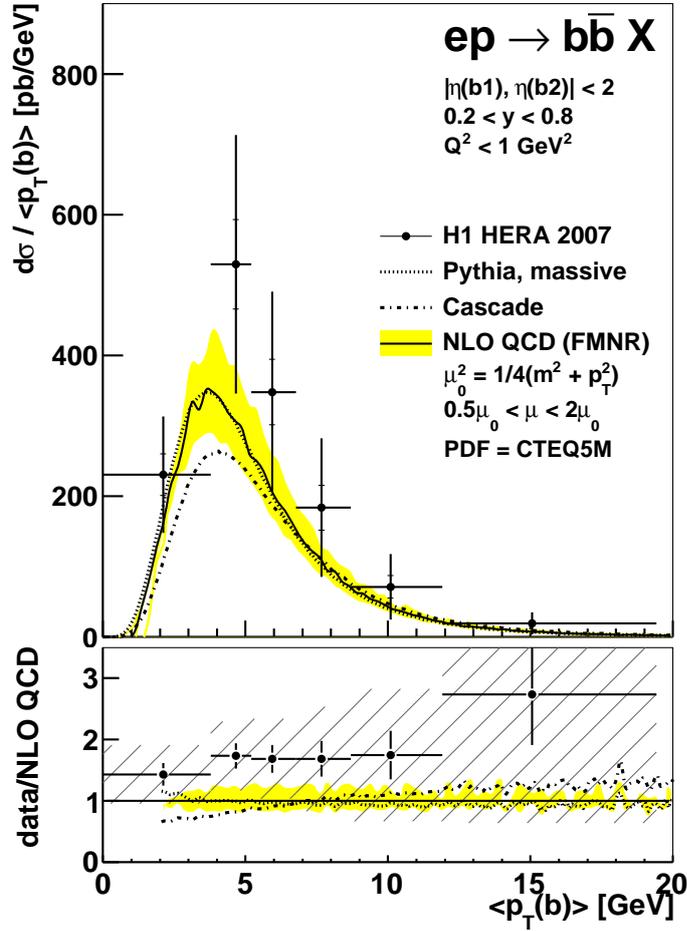


Figure 8.14: Beauty Photoproduction Cross Section

Upper plot: The measured differential beauty cross section $d\sigma/d\langle p_T(b) \rangle$ is shown as function of the mean transverse momentum of the beauty quarks $\langle p_T(b) \rangle$. The data is represented as points, where the inner vertical errors given are the uncorrelated errors of the measurement and the outer error bars represent the total error, for which the uncorrelated and the correlated error are added in quadrature. The bin centers of the indicated data points are corrected in $\langle p_T(b) \rangle$ according to [101] (based on the Pythia simulation). The horizontal errors indicate the bin width in $\langle p_T(b) \rangle$ of each data point. Also shown are the predictions of the LO calculation of Pythia (dotted line) and Cascade (dashed line) and the FMNR NLO QCD calculation (band).

Lower plot: Ratio of the measured cross section and the calculated NLO pQCD cross section, $R = d\sigma_{measured}/d\sigma_{NLO\ pQCD}$. The error bars of the data points are the uncorrelated errors, whereas the correlated error is indicated by the hatched area. The yellow shaded band is the relative error of the NLO calculation. Below $\langle p_T(b) \rangle = 2\text{ GeV}$ the predictions are not shown due to numerical instabilities in the ratio calculation.

8.6 Comparison with Previous Measurements at HERA

The differential beauty cross section $d\sigma/dp_T(b)(ep \rightarrow eb X)$ of this beauty measurement is compared to previous beauty measurements at HERA in Figure 8.16. The data points of this measurement are presented as (green) points, with the inner errors being the uncorrelated error and the outer errors the total error. The yellow shaded band is the NLO pQCD calculation⁹ and the solid line a calculation based on the k_T -factorization approach [105]. The phase space restrictions of the summary plot differ in terms of η -range from the phase space definition of this measurement (Figure 8.14). Correction factors result from modifying the cut on the rapidity from $(|\eta(b)| < 2)$ AND $(|\eta(\bar{b})| < 2)$ to $(|\eta(b)| < 2)$ OR $(|\eta(\bar{b})| < 2)$. The correction factors applied were determined from the Pythia massive and Cascade MC simulation and are presented in Figure 8.15. The mean value of the two models was taken and the differences treated as an additional systematical error.

The presented data points of this analysis extend the previous measurements towards lower transverse momenta of the beauty quark. In the overlapping region this measurement agrees with the other measurements within errors. It also agrees with the NLO pQCD prediction within errors, even though the data exhibit a larger beauty production than is predicted. This is also the case for all other analyses extracting the beauty cross section from double tagged correlations and without using jet algorithms, like the $\mu\mu$ and the $D^*\mu$ measurements (see Figure 8.16). On the other hand, the jet based analyses agree better with the predictions and achieve a higher accuracy, but they are restricted to the high $p_T(b)$ regime.

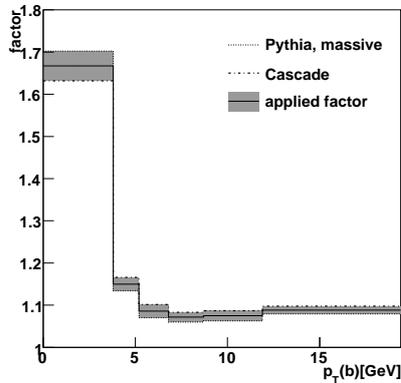


Figure 8.15: η Correction Factors

Correction factors that result from the transition of the phase space by modifying the cut on the rapidity from $(|\eta(b)| < 2)$ AND $(|\eta(\bar{b})| < 2)$ to $(|\eta(b)| < 2)$ OR $(|\eta(\bar{b})| < 2)$. This phase space transition is necessary in order to compare the data points of this measurements (Figure 8.14) with the other beauty measurements at HERA (Figure 8.16). The correction factors are determined from the Pythia and Cascade simulation, the mean value is applied and the difference between the two models treated as an additional systematic error.

⁹The theoretical uncertainty shown results from a simultaneous variation of the beauty mass and the scale in the ranges $4.5 \text{ GeV} < m_b < 5 \text{ GeV}$ and $\mu_0/2 < \mu_0 < 2\mu_0$. See also [65, 66].

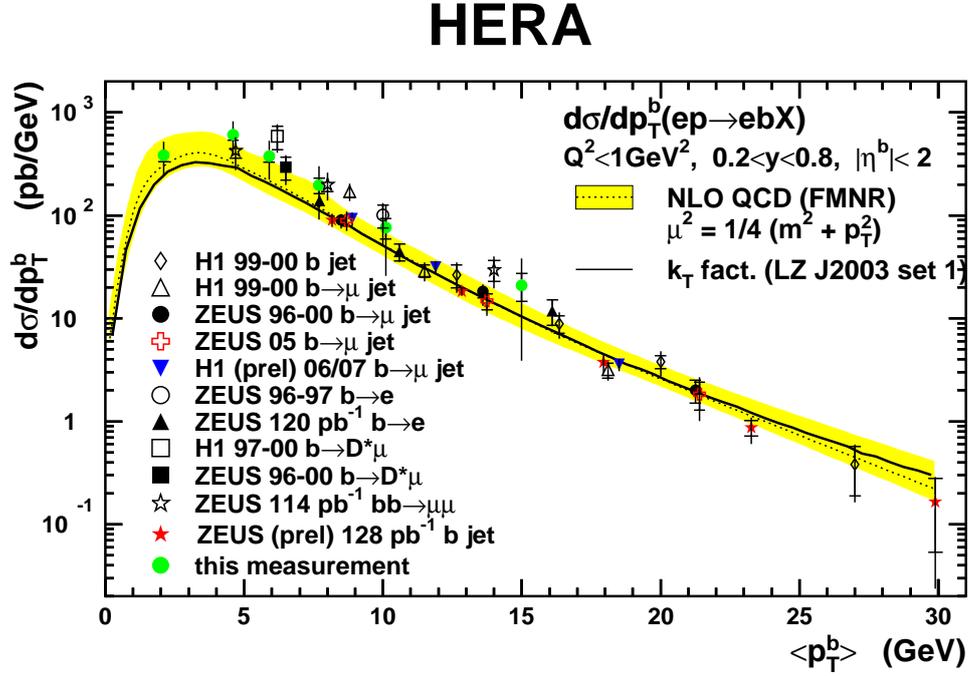


Figure 8.16: Beauty Photoproduction at HERA.

The differential beauty cross section $d\sigma/dp_T(b)$ of this analysis is compared to previous measurements at HERA [106], as well as the predictions based on FMNR NLO QCD calculation (band) and a calculation based on the k_T -factorization approach (line). The inner error bars on the data points of this measurement indicate the uncorrelated error, the full error bars contain in addition the correlated error added in quadrature.

The phase space of this Figure differs in terms of η -range from the Figure 8.14. The correction factors applied are presented in Figure 8.15).

$p_T(b)$ range [GeV]	$\langle p_T(b) \rangle$ [GeV]	$d\sigma / \langle p_T(b) \rangle$	uncorr. err. [pb/GeV]	corr. err.
0 3.8	2.1	230	29	74
3.8 5.2	4.6	530	62	163
5.2 6.8	5.9	348	45	128
6.8 8.7	7.7	184	31	89
8.7 11.9	10.1	70	16	42
11.9 19.4	15.0	19	5.7	14

Table 8.6: Measured Differential Beauty Cross Section Values

Measured differential cross section values for beauty in photon production as a function of the mean transverse momentum of the beauty quark $\langle p_T(b) \rangle$ in the kinematic range $0.2 < y < 0.8$, $Q^2 < 1 \text{ GeV}^2$, $|\eta(b), \eta(\bar{b})| < 2$. The bin range, bin center corrected [101] transverse momentum of the beauty quark $\langle p_T(b) \rangle$, the bin averaged cross section values and the uncorrelated and correlated uncertainties are listed. The data points listed are plotted in the Figure 8.14.

Chapter 9

Conclusions and Outlook

Beauty photoproduction was measured as function of the mean transverse momentum of the beauty quark $\langle p_T(b) \rangle$, with a special focus on the low $\langle p_T(b) \rangle$ regime, using the H1 detector at the HERA collider. This measurement extends the previously experimental accessible phase space towards the beauty production threshold, which became possible by exploiting for the first time the semielectronic decay channel $ep \rightarrow b\bar{b} \rightarrow eeX$ and because a low cutoff was achieved online and offline, by mastering the experimental challenges of low p_T electron identification.

A new path was followed to extract the differential beauty cross section only from the measured di-electron spectrum by means of a regularized unfolding procedure.

The measurement presented is limited by the correlated error related to the uncertainty on the determination of the open charm background contribution. Compared to the NLO pQCD prediction the obtained result is compatible within the (still high) errors, even though the prediction shows a tendency to underestimate the data. In the overlapping region this measurement agrees well with other beauty production measurements at HERA. In particular this is the case for all other double tag measurements that similarly to this measurement extract the beauty signal by means of correlations between beauty decay products, either $\mu\mu$ or μD^* . The jet based analyses in general obtain a higher accuracy and agree better with the theoretical prediction, however are only sensitive at higher transverse momenta of the beauty quark.

This measurement has a high potential to gain in precision due to two recent developments within the H1 collaboration. First, the new detector response software of the H1 experiment will be capable to simulate also the differential energy loss dE/dx . In particular at low transverse electron momenta, this will allow to further reduce the background, and might allow to push the electron reconstruction threshold to even lower values. Second, H1 will publish a new D^* measurement in photoproduction [92]. The inclusion of this result in the unfolding, will allow to control the charm background and therefore to reduce the corresponding error on the beauty cross section.

Appendix A

Unfolding

A.1 Formulation of the Unfolding Problem

A standard task in high energy experiments is the extraction of a signal distribution $f^{sig}(x^{sig})$ of a quantity x^{sig} . With an ideal detector this could be done by simply histogramming x^{sig} , whereas with a real detector the measured quantity y is modified with respect to the true quantity x^{sig} due to limited acceptances, reconstruction efficiencies and resolution effects. In addition, it is usually impossible to extract a pure signal, as the measured quantity y can be diluted by a background quantity x^b ¹.

Mathematically the distribution $g(y)$ of the measured quantity y can be related to the true underlying distribution of the signal $f^{sig}(x^{sig})$ and the background distribution $f^b(x^b)$ by the equation

$$g(y) = \int A^{sig}(x^{sig}, y)f(x^{sig})dx^{sig} + \int A^b(x^b, y)f(x^b)dx^b. \quad (\text{A.1})$$

All detector effects enter this equation via the so called response functions A^{sig} for the signal and A^b the background distribution. All inefficiencies due to the detector response, and the consequential correction factors, are taken via A^{sig} and A^b into account. For a given signal/background value $x_0^{sig/b}$, $A^{sig/b}(y, x_0^{sig/b})$ describes the response of the detector in the variable y . The distribution $g(y)$ of the measured quantity y is obtained by folding the true² distributions of the signal $f(x^{sig})$ and the background $f(x^b)$ with the detector response A^{sig} and A^b , respectively.

In the case of multiple signal (N_{sig}) and background (N_b) contributions equation (A.1) has to be generalized to

$$g(y) = \sum_i^{N_{sig}} \int A^{sig_i}(x^{sig_i}, y)f(x^{sig_i})dx^{sig_i} + \sum_i^{N_b} \int A^{b_i}(x^{b_i}, y)f(x^{b_i})dx^{b_i} \quad , \quad (\text{A.2})$$

where x^{sig_i} (x^{b_i}) is the i^{th} signal (background) contribution and $A^{sig_i}(x^{sig_i}, y)$ ($A^{b_i}(x^{b_i}, y)$) the associated response function. For the numerical solution of the equation (A.2), it is

¹ In the analysis presented in this thesis x^{sig} is the mean transverse momentum of the beauty quarks $x^{sig} = \langle p_T(b) \rangle$, and y is derived from the measured electron pairs and $m_{T,est}(b)$ determined by the thrust axis method. The background contributions are electron pairs from semileptonic charm and J/ψ decays as well as from misidentified electrons.

²In the context of unfolding, the term 'true' is used for the distributions that are not folded with the detector response function.

convenient to linearize the problem and to represent all distributions by histograms, and the response functions by matrices, such that the integral equation reduces to a matrix equation:

$$\mathbf{y} = \sum_i^{N_{sig}} \mathbf{A}^{sig_i} \mathbf{x}^{sig_i} + \sum_i^{N_b} \mathbf{A}^{b_i} \mathbf{x}^{b_i} \quad (\text{A.3})$$

The vector $\mathbf{y} = (y_1, y_2, \dots, y_m)$ represents the histogram of N observed events distributed with $g(y)$, where y_j is the number of measured events in bin j and m the number of bins in the histogram.

The relationship between the histogram \mathbf{y} and the distribution $g(y)$ has to be interpreted in the standard way, i.e.

$$y_i = n_{obs} \cdot \int_{\substack{\text{bin width} \\ \text{of bin } i}} g(y) dy \quad , \quad (\text{A.4})$$

$$n_{obs} = \sum y_i = \text{Number of observed events} \quad . \quad (\text{A.5})$$

Likewise the vectors $\mathbf{x}^{sig/b_i} = (x_1^{sig/b_i}, x_2^{sig/b_i}, \dots, x_{n^{sig/b_i}}^{sig/b_i})$ represents the histograms of the i^{th} signal- ($f(x^{sig_i})$), respectively the i^{th} background distribution ($f(x^{b_i})$) having n^{sig_i} and n^{b_i} bins. Note that the number of bins $m, n^{sig/b_1}, n^{sig/b_2}, \dots, n^{sig/b_N}$ for the different histograms do not have to be equal. The element (k, l) of the response matrix \mathbf{A}^{sig/b_i} is obtained by integrating the corresponding response function over the l^{th} bin of the histogram \mathbf{y} and k^{th} bin of the histogram \mathbf{x}^{sig/b_i} :

$$\left(\mathbf{A}^{sig/b_i}\right)_{k,l} = \int_{\text{bin } k} dx^{sig/b_i} \int_{\text{bin } l} A^{sig/b_i}(x^{sig/b_i}, y) dy. \quad (\text{A.6})$$

In practice, however, the response matrices are not computed by integrating a function, but from appropriate signal and background Monte Carlo samples, for which the detector response has been simulated. A possible procedure is described in section A.2.

Only contributions to the reconstructed vector \mathbf{y} that have a different detector response function are separable with unfolding. Contributions with equal or similar detector response functions are correlated. It is preferable to split the second sum in equation (A.3) in a set of separable (uncorrelated) and a set of non-separable (correlated) backgrounds, such that the correlated part can be subtracted later from \mathbf{y} .

$$\mathbf{y} = \sum_i^{N_{sig}} \mathbf{A}^{sig_i} \mathbf{x}^{sig_i} + \sum_i^{N_b^{sep.}} \mathbf{A}^{b_i} \mathbf{x}^{b_i} + \sum_{i > N_b^{sep.}}^{N_b} \mathbf{b}^i \quad (\text{A.7})$$

$$\mathbf{b}^i = \mathbf{A}^{b_i} \mathbf{x}^{b_i} \quad (\text{A.8})$$

For similar response functions, a higher discrimination may be achieved by combining several reconstructed variables, and by applying substructures to the reconstructed vector \mathbf{y} . (Compare with Figure 8.3 and section 8.2.1 for an example.)

In order to simplify the notation in the following, the different separable response matrices \mathbf{A}^{sig/b_i} are combined to one single response matrix \mathbf{A} , the different separable (uncorrelated) vectors \mathbf{x}^{sig/b_i} to one single vector \mathbf{x} , and the different non-separable (correlated)

vectors $\mathbf{x}^{\text{sig}/b_i}$ to a single vector \mathbf{b} , by rewriting equation (A.7):

$$\begin{aligned}
\mathbf{y} &= \sum_i^{N_{\text{sig}}} \mathbf{A}^{\text{sig}} \mathbf{x}^{\text{sig}} + \sum_i^{N_b^{\text{sep.}}} \mathbf{A}^{b_i} \mathbf{x}^{b_i} + \sum_{i > N_b^{\text{sep.}}}^{N_b} \mathbf{b}^i \\
&= \underbrace{(\mathbf{A}^{\text{sig}1}, \dots, \mathbf{A}^{\text{sig}N_{\text{sig}}}, \mathbf{A}^{b1}, \dots, \mathbf{A}^{bN_b})}_{\mathbf{A}} \underbrace{(\mathbf{x}^{\text{sig}1}, \dots, \mathbf{x}^{\text{sig}N_{\text{sig}}}, \mathbf{x}^{b1}, \dots, \mathbf{x}^{bN_b})^T}_{\mathbf{x}^T} + \underbrace{\sum_{i > N_b^{\text{sep.}}}^{N_b} \mathbf{b}^i}_{\mathbf{b}} \\
&\Rightarrow \mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b} \tag{A.9}
\end{aligned}$$

To illustrate the idea, a sketch of a response histogram, corresponding to a response matrix \mathbf{A} composed of two signal and three background contributions is presented in Figure A.1.

A.2 Determination of the Response Matrix

The response matrix $\mathbf{A} = (\mathbf{A}^{\text{sig}1}, \dots, \mathbf{A}^{\text{sig}N_{\text{sig}}}, \mathbf{A}^{b1}, \dots, \mathbf{A}^{bN_b})$, taking several signal and background contributions into account, can be determined from Monte Carlo simulations by first defining a response histogram³ $\tilde{A} = (\tilde{A}^{\text{sig}1}, \dots, \tilde{A}^{\text{sig}N_{\text{sig}}}, \tilde{A}^{b1}, \dots, \tilde{A}^{bN_b})$ reflecting the structure of \mathbf{A} . The concept is illustrated in Figure A.1 for two signal contributions and three background contributions. Every signal (background) submatrix $\mathbf{A}^{\text{sig}i}$ (\mathbf{A}^{b_i}) is represented by a corresponding subhistogram $\tilde{A}^{\text{sig}i}$ (\tilde{A}^{b_i}). The number of bins of \tilde{A} corresponds along the x-axis to the number of elements of \mathbf{A} , whereas along the y-axis an additional bin is added for normalization purposes. The response histogram is filled by the following procedure: Every simulated event of the i^{th} signal (background) contribution is filled in the corresponding subhistogram $\tilde{A}_{k,l}^{\text{sig}/b_i}$, where the first index k is given by the bin number of x^{sig/b_i} in the histogram $\mathbf{x}^{\text{sig}/b_i}$ and the second index l by the bin number of y in the histogram \mathbf{y} . The normalization bins of the response histogram are filled such that the sum over all bins in the corresponding column k of $\tilde{A}_{k,l}$ are equal to the number of events in the true histogram \mathbf{x} in bin k :

$$x_k = \sum_{l \neq \text{norm. bin}} \tilde{A}_{k,l} + \tilde{A}_{k,l=\text{norm. bin}} \tag{A.10}$$

$$\Rightarrow \tilde{A}_{k,l=\text{norm. bin}} = x_k - \sum_{l \neq \text{norm. bin}} \tilde{A}_{k,l} \tag{A.11}$$

After a subsequent columnwise normalization of the response histogram \tilde{A} , the response matrix \mathbf{A} is given by the corresponding bin values of \tilde{A} .

A.3 Regularized Unfolding

In order to determine the true vector \mathbf{x} from the measured vector \mathbf{y} , one could try to invert the response matrix \mathbf{A} of equation (A.9), such that

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{y} - \mathbf{b} \quad . \tag{A.12}$$

Even if the response matrix \mathbf{A} is quadratic and the inverse of it exists, the so obtained \mathbf{x} can result in an unacceptable solution with widely fluctuating bins. The reason for this

³Response matrices are printed in bold (\mathbf{A}) and the response histograms are labeled with a tilde (\tilde{A}).

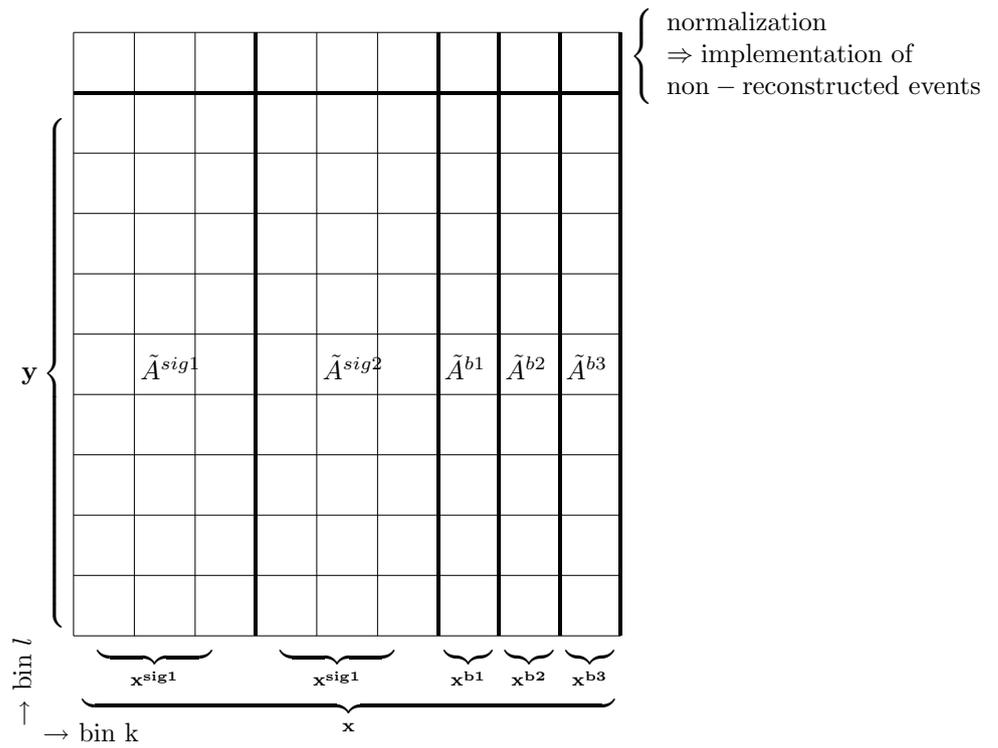


Figure A.1: Response Histogram

Sketch of a response histogram with nine input and nine output bins. The input vector \mathbf{x} is composed of two signal and three background contributions indicated by $\mathbf{x}^{\text{sig/b}}$. Compare with equation (A.9) and the text above.

behavior are statistical fluctuations in the determination of the response matrix, noise in the measured vector \mathbf{y} and model assumptions. The art of unfolding is to construct a biased estimator $\hat{\mathbf{x}}$ for \mathbf{x} such that the bias is small, if some prior assumptions are true [99]. Besides the assumption on the signal and background models, the standard approach is to apply an assumption on the smoothness of the result. This method is called regularization (see [99, 111, 112]).

One possibility⁴ to construct such an estimator, is to choose $\hat{\mathbf{x}}$ such that the $\chi^2(\hat{\mathbf{x}})$ given by

$$\chi^2(\hat{\mathbf{x}}, \tau, \mu) := \chi_A^2(\hat{\mathbf{x}}) + \tau \cdot \chi_L^2(\hat{\mathbf{x}}) + \mu \cdot \chi_N^2(\hat{\mathbf{x}}) \quad (\text{A.13})$$

$$\chi_A^2(\hat{\mathbf{x}}) := 1/2 (\mathbf{y} - \mathbf{b} - \mathbf{A}\hat{\mathbf{x}})^T V^{-1} (\mathbf{y} - \mathbf{b} - \mathbf{A}\hat{\mathbf{x}}) \quad (\text{A.14})$$

$$\chi_L^2(\hat{\mathbf{x}}) := \hat{\mathbf{x}}^T L \hat{\mathbf{x}} \quad (\text{A.15})$$

$$\chi_N^2(\hat{\mathbf{x}}) := \left(n_{obs} - \sum_{j=1}^m (\mathbf{A}\hat{\mathbf{x}})_j \right)^2, \quad (\text{A.16})$$

is minimized for fixed τ and μ , whereas τ often is denoted as regularization parameter. $V = cov(y_i, y_j)$ is the covariance matrix of the data, L the regularization matrix, n_{obs} the number of observed events.

Equation (A.13) is in principle nothing more than a Lagrange minimization problem of $\chi_A^2(\hat{\mathbf{x}})$ with the two side conditions given by $\chi_L^2(\hat{\mathbf{x}})$ and $\chi_N^2(\hat{\mathbf{x}})$. However, the values of the Lagrange multipliers τ and μ are determined differently than in standard Lagrange minimization problems (see below). The individual χ^2 -terms of equation (A.13) have the following interpretation:

- $\chi_A^2(\hat{\mathbf{x}})$ is the standard χ^2 definition and minimizes the deviation of $\mathbf{A}\hat{\mathbf{x}}$ from the measured vector $\mathbf{y} - \mathbf{b}$.
- $\chi_L^2(\hat{\mathbf{x}})$ is a measure for the smoothness of the result, with

$$L = \begin{bmatrix} L1 & 0 & \dots & 0 \\ 0 & L2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & LN \end{bmatrix}, \quad (\text{A.17})$$

where L_i correspond to the i^{th} signal/background contribution of equation (A.7). The smoothness conditions are only applied to the individual contributions i with the L given in equation (A.17). In the literature different choices of L_i are described, such that the size or the n -th derivative of $\hat{\mathbf{x}}$ is minimized. The best choice of L_i depends on the problem studied. In the analysis presented in this thesis the second derivative of the signal contributions of $\hat{\mathbf{x}}$ was minimized and no regularization conditions were applied to the background contributions. Since $\hat{\mathbf{x}}_i'' \propto \hat{\mathbf{x}}_{i-1} - 2 \cdot \hat{\mathbf{x}} + \hat{\mathbf{x}}_{i+1}$, L_i is then given by

$$L_i = \begin{cases} \begin{bmatrix} 1 & -1 & 0 & 0 & \dots & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & \dots & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & \dots & 0 & -1 & 1 \end{bmatrix} & \text{regularization} \\ \begin{bmatrix} 0 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 0 \end{bmatrix} & \text{no regularization} \end{cases}, \quad (\text{A.18})$$

⁴In [99] an overview on several other approaches is given.

- $\chi_N^2(\hat{\mathbf{x}})$ ensures the normalization of the obtained result, i.e. that the total number of estimated events is equal to the number of actually observed events [99].

The parameters τ and μ determine how the two assumptions (smoothness and normalization) influence the undisturbed $\chi_A^2(\hat{\mathbf{x}})$. There is a certain arbitrariness of how these parameters are determined, especially for the choice of the regularization parameter τ . A possible strategy is presented in section A.3.1.

For given parameters τ and μ the estimator $\hat{\mathbf{x}}$ can be determined by setting the derivatives of $\chi^2(\hat{\mathbf{x}})$ with respect to \hat{x}_i to zero [114]:

$$\nabla\chi^2(\hat{\mathbf{x}}) = 0 \quad (\text{A.19})$$

$$\Rightarrow \hat{\mathbf{x}} = (A^T V^{-1} A + \tau L)^{-1} A^T V^{-1} \hat{\mathbf{y}} \quad (\text{A.20})$$

$$+ \underbrace{\mu \cdot \left(n_{obs} - \sum_{j=1}^N (\mathbf{A}\hat{\mathbf{x}})_j \right)}_{=: \alpha(\hat{\mathbf{x}})} \cdot \left(\sum_i A_{i,j} \right)$$

For τ and μ equal zero, equation (A.20) reduces to $\mathbf{x} = \mathbf{A}^{-1}\mathbf{y}$, the solution without side conditions. For $\mu \neq 0$ equation (A.20) is not a solution for $\hat{\mathbf{x}}$ since $\alpha = \alpha(\hat{\mathbf{x}})$. In a numerical implementation however, μ can be replaced by α and treated as a parameter.

$$\Rightarrow \hat{\mathbf{x}} = (A^T V^{-1} A + \tau L)^{-1} A^T V^{-1} \hat{\mathbf{y}} + \alpha \cdot \left(\sum_i A_{i,j} \right) \quad (\text{A.21})$$

The parameter α then has to be adjusted such that the difference between the number of estimated events and the number actually observed events is smaller than a certain limit, i.e. $\left| n_{obs} - \sum_{i=1}^N (\mathbf{A}\hat{\mathbf{x}})_i \right| < limit$.

A.3.1 Choice of the Regularization Parameter

The choice of the regularization parameter τ is in principle a trade off between the bias and the variance of the estimator $\hat{\mathbf{x}}$ [99]. Several methods to determine the best value for the regularization parameter τ are discussed in the literature, good references are for example [99, 111, 115].

The method followed in this thesis is characterized in [111] as “... seems to be the best method”. The basic idea of this approach is to minimize the correlation coefficients of $\hat{\mathbf{x}}$, i.e. to choose τ such that the covariance matrix of $\hat{\mathbf{x}}$, V_X becomes mostly diagonal. In order to quantify the correlations, the global correlation coefficient ρ_i is defined:

$$\rho_i = \sqrt{1 - ((V_x)_{ii} * (V_x^{-1})_{ii})^{-1}} \quad (\text{A.22})$$

The global correlation coefficient measures the total correlation between the element i of $\hat{\mathbf{x}}$ and all other elements. Its value is between zero and one. For the present choice of τ , the mean of all n global correlation coefficients $\langle \rho \rangle$ is sampled over a wide range of τ -values, and the τ -value with the smallest mean value $\langle \rho \rangle$ is taken.

A.3.2 Interpretation of the Regularization Parameter

The choice of τ influences the effective number of degrees of freedom (NDF) of the estimator $\hat{\mathbf{x}}$ (cf. equation (A.13)). For no side conditions $\tau = 0$ the number of degrees of freedom are

equal to the number of bins of $\hat{\mathbf{x}}$. For $\tau \neq 0$, the side condition induced by the regularization reduces the effective NDFs⁵.

In [111, 112, 114] regularization is interpreted in this way. By a set of linear matrix multiplications the vector \mathbf{x} is transformed to a vector $\mathbf{c} = \mathbf{C} \cdot \mathbf{B} \cdot \mathbf{E} \cdot \mathbf{x}$, such that the transformed estimator becomes

$$\hat{\mathbf{c}} = (\mathbb{1} + \tau \mathbf{D})^{-1} \hat{\mathbf{c}}_0 \quad (\text{A.23})$$

\mathbf{D} is a diagonal matrix, and $\hat{\mathbf{c}}_0$ the unregularized solution. In this representation \mathbf{x} is expressed as the superposition of Eigenvectors, each weighted with a filter factor

$$f_i = \frac{1}{1 + \tau \lambda_i}, \quad (\text{A.24})$$

where λ_i are the Eigenvalues of the matrix \mathbf{D} . The filter factors f_i depend on the regularization parameter τ . For $\tau = 0$, all filter factors are equal to one, $f_i = 1$. With increasing regularization parameters, $\tau > 0$, the filter factors with large Eigenvalues λ_i rapidly decrease towards zero and suppress the non significant Eigenvector contributions. The interpretation is motivated by a Fourier analysis, in which the high frequencies of the spectrum are damped. The sum of all filter factors can be interpreted as the effective NDFs:

$$NDF(\tau) = \sum_{i=0}^N \frac{1}{1 + \tau \lambda_i} \quad (\text{A.25})$$

For $\tau = 0$ the effective NDFs are equal to the number of bins of the histogram \mathbf{x} , $NDF(\tau = 0) = N$. For $\tau > 0$ filtering sets in, $NDF(\tau)$ then only counts the number of those Eigenvectors not being suppressed.

If the normalization of the unfolding result is fixed, i.e. if the parameter μ in equation (A.13) is chosen > 0 , the NDF of the estimator $\hat{\mathbf{x}}$ is reduced further by 1.

A.3.3 Computer Programs for Unfolding

The computer program used in this thesis for regularized unfolding is a modified version of the TUnfold [118] class, developed within H1 and implemented in the RooT software package. The existing code of TUnfold was extended by the side condition on the number of reconstructed events (equation A.16). Besides the unfolded results the Unfold package provides the full error matrix of the unfolded points and provides methods to determine the systematical errors. The latter is done by simple error propagation of a systematical uncertainty on the response matrix to the final result.

⁵The argumentation is motivated by the similar and well know discussion on NDFs in least square fits with constraints: $NDF_{LSQ} = N_{bin} - \#\text{constraints}$.

Appendix B

Dead Material Correction

The development of the electron identification in this thesis (chapter 6) necessitated a method to correct for energy losses in dead material in front of the calorimeter. The dead material correction presented, depends on the energy deposition in the calorimeter and the amount and the material in front of it. With high statistical samples the method can be used to determine the amount and type of dead material as function of its position. The basic ideas of the correction derived are summarized in this section.

If a certain amount of dead material is in front of the calorimeter, the shower will already start to develop before it enters into the calorimeter and will therefore deposit a part of its energy in the dead material. This energy can not be measured in the calorimeter. However, if one has a rough knowledge of the amount and the type of dead material in front of the calorimeter one can correct for it.

The starting point for the correction derived is a parameterization for the mean longitudinal profile of the energy deposition in an electromagnetic cascade, as shown in Figure B.1a. In [97] and [117] the following parameterization is described to give a reasonable description of the profile:

$$\frac{dE}{dt} = E_0 b \cdot \left(\frac{(bt)^{a-1} e^{-bt}}{\Gamma(a)} \right) \quad (\text{B.1})$$

E_0 is the total energy in the shower, a and b are parameters related to the material in which the cascade showers, t is the depth in the material measured in radiation lengths from the starting point of the cascade and $\Gamma(a)$ is the gamma function, defined as $\Gamma(a, x) = \int_x^\infty t^{a-1} e^{-t} dt$, $\Gamma(a) = \Gamma(a, 0)$.

Assuming that equation (B.1) is also valid for the dead material and taking X_0 as the depth of the dead material in radiation lengths, the amount of energy lost E_{lost} in it can be estimated by integration:

$$E_{lost} = \int_0^{X_0} \left(\frac{dE}{dt} \right) dt = E_0 (1 - \Gamma(a, bX_0) / \Gamma(a)) \quad (\text{B.2})$$

It is assumed in the following that the remaining energy contained in the shower $E_{dep} = E_0 - E_{lost}$ is completely deposited in the calorimeter, and that the measured energy $E_{measured}$ is related to the deposited energy by a linear calibration constant C , i.e $E_{dep} = C \cdot E_{measured}$. By combining these equations with equation (B.2) one gets:

$$E_0 = E_{measured} \cdot C \cdot \frac{\Gamma(a)}{\Gamma(a, bX_0)} \quad (\text{B.3})$$

The parameters a and b are related with each other through the following expression (derived from equations¹ published in [97]):

$$a \approx b \cdot \left[\log(E) - \log\left(\frac{0.8}{Z + 1.2}\right) - 0.5 \right] + 1 \quad (\text{B.4})$$

E is the energy of the shower in GeV and Z the atomic number of the dead material. b is a material dependent parameter describing the shape of the profile, and is usually of the order 0.5 (compare with Figure B.1b).

Within H1 the measured energy is already of the same order as the corrected energy due to a rough calibration, e.g. $E_{measured} \simeq E_0$. With this and equation (B.4), equation (B.3) can be rewritten as:

$$E_0 = E_{measured} \cdot C \cdot f(X_0, b, Z, E_{measured}) \quad (\text{B.5})$$

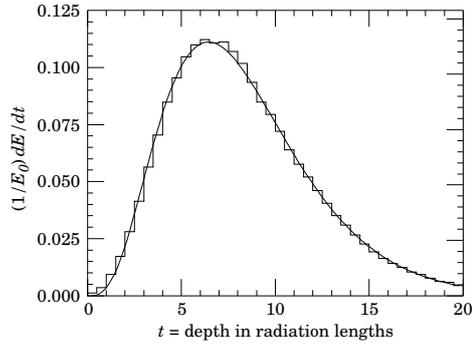
$$f(X_0, b, Z, E_{measured}) = \frac{\Gamma\left(b \cdot \left[\log(E_{measured}) - \log\left(\frac{0.8}{Z + 1.2}\right) - 0.5 \right] + 1\right)}{\Gamma\left(b \cdot \left[\log(E_{measured}) - \log\left(\frac{0.8}{Z + 1.2}\right) - 0.5 \right] + 1, bX_0\right)} \quad (\text{B.6})$$

Once the parameters X_0 , b and Z are fixed, the function $f(X_0, b, Z, E_{measured})$ is an energy dependent function to correct for energy losses due to dead material in front of the calorimeter. C is a linear calibration constant.

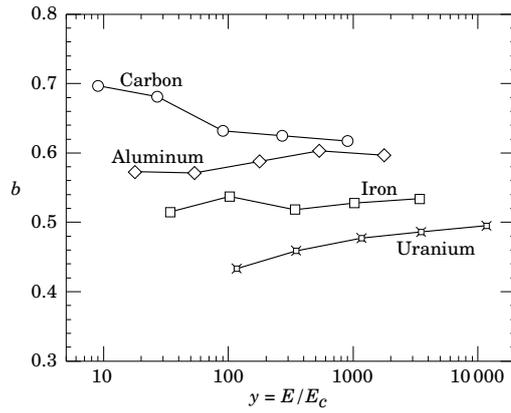
All input variables to the Artificial Neural Network for the electron identification (discussed in chapter 6) are corrected for energy losses in the dead material, according to the scheme described here. The constants X_0 , b and Z were determined for every LAr cell on the lowest layer of the calorimeter by means of fits, based on electron showers from Single Particle Monte Carlo samples. In principle this could also be done with data, but no electron sample with high enough statistics exists to determine all three constants simultaneously with the full granularity. However, $J/\psi \rightarrow ee$ data samples were used to cross check the results obtained from the Single Particle Monte Carlo samples. Typical values for the parameters were: $Z \sim 9$, $b \sim 0.8$ and $X_0 \sim 1 - 2$, depending on the cell position. The amount of dead material is around one radiation length for the central barrel and increases towards the forward barrel to 2 radiation lengths. These constants are equally used for data and Monte Carlo, whereas the calibration parameters C were readjusted for data with $J/\psi \rightarrow ee$ samples.

The effect of this energy correction is demonstrated in the Figures B.2(a)-(b). Both plots are based on a Single Particle electron Monte Carlo sample, simulating the detector response to a single electron. As an example for electrons hitting the CB3, the reconstructed energy relative to the generated energy E/E_{gen} as function of the generated energy E_{gen} is shown before the dead material correction (a) and after the dead material correction (b).

¹The equations published are:
 $\frac{a-1}{b} = \log\left(\frac{E}{E_c}\right) - 0.5$ and $E_c = \frac{0.8 \text{ GeV}}{Z+1.2}$



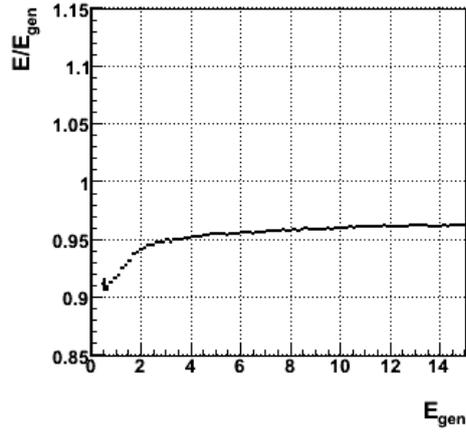
(a)



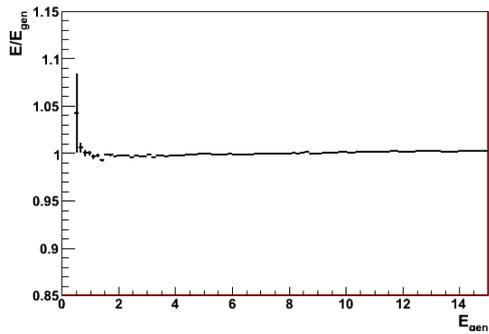
(b)

Figure B.1: Electromagnetic Showers

In (a) the longitudinal profile of an electron induced cascade is shown. The histogram shows the fractional energy deposition per radiation length, and the curve is a gamma-function fit to the distribution. (b) shows values of the scale factor b for energy deposition profiles for a variety of elements and for incident electrons with 1 – 100 GeV. (Both Figures are taken from [97], (a) is slightly adapted).



(a)



(b)

Figure B.2: Performance of the Dead Material Correction

For a Single Particle electron MC sample the fraction of the reconstructed energy over the generated energy, as function of the generated energy is shown in (a) before the dead material correction, and in (b) after the dead material correction.

Appendix C

Calibration of the Jet Trigger Simulation

The input information to the FTT L3 trigger system (cf. section D) relevant for the online identification of electrons were tracks measured in the CJC by the Fast Track Trigger on L2 (cf. section 3.2.5), and energy depositions in the LAr calorimeter measured by the Jet Trigger (cf. section 3.2.5). This information¹ was transferred to the FTT L3 in the form of two lists. The first one contained for each FTT track 3 floating point numbers, reflecting the 3 track coordinates $1/p_T$, φ and ϑ . The second one contained $16 \cdot 3$ integers, reflecting the energy and the spatial position coded in φ and ϑ of the 16 highest energetic energy depositions in the LAr calorimeter.

The L3 trigger simulation (cf. section D.3) determines from these two lists the L3 trigger elements². The two input lists to the L3 simulation can be either the ones that were measured online by the FTT-L2 and the JT, or the ones derived from the FTT-L2 and the JT simulations. The discussion concentrates on the JT trigger simulation in the following, while the FTT-L2 simulation is documented in [72].

The input to the JT simulation are energy depositions in the so called trigger towers (TT). These are cell units of the LAr calorimeter also used in the H1 LAr trigger, as for instance discussed in [113]. The simulation of the TT is implemented in the standard detector response software, i.e. h1simrec. From the energy depositions in the TT, the JT simulation determines among other things the input to the L3 simulation, thus the list of the 16 highest energetic energy depositions. As for the L3 simulation, the input to the JT simulation can either be the one measured, or the one determined from the simulation. A sketch of the simulation structure discussed is presented in Figure C.2.

Every simulation unit can use as input data or the simulation output of the previous systems. This allows the control of each simulation unit by resimulating the data and comparing the simulated response to the actual response of the system. In particular this can be done for the JT simulation.

As an example Figure C.1 shows for a certain TT in the CB1 region the simulated JT energy correlated to the actually measured JT energy, before the JT simulation was calibrated. The plot is obtained from a high Q^2 data sample, by comparing the 16 highest energetic energy depositions in the simulated JT output list with the corresponding measured list.

¹FTT L3 also received input information from the central muon system (CMD), besides FTT L2 and the JT.

² The algorithms implemented in the L3 simulation are exactly the same as were loaded in the hardware. Compare with the discussion in section D.3.

If the two lists have an overlapping entry in the two spatial integers (i.e. the integers representing φ and ϑ), the third integer (containing the energy measurement) is filled into the scatter plot. Clearly one can identify two correlations, that are related to the hadronic and the electromagnetic energy scale of the TT.³ Under the assumption of having either an electromagnetic or a hadronic energy deposition the sample is split into two, and each subsample is fitted with a straight line:

- The initial separation is based on the diagonal. Afterwards the straight line in-between the two fitted straight lines defines the split-up. This procedure is iterated until it converges.
- The fit minimizes the χ^2 given by $\chi^2 = \sum_i (E_{\text{measured}} - c \cdot E_{\text{simulated}})^2$.

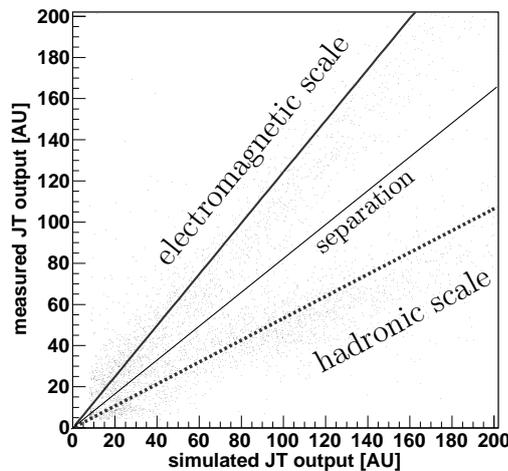


Figure C.1: Measured vs. Simulated JT Energy, before the Calibration

The figure compares the measured with the simulated JT energy, before the JT simulation was calibrated. The energy deposition is restricted to a single TT of the central barrel. The energy is given in JT internal units, 1 unit roughly corresponds to 0.1 GeV. The plot is obtained from a high Q^2 data sample, which is also the input to the simulation. Clearly two correlations can be identified that can be related to the electromagnetic and the hadronic energy scales. From the straight lines fitted into the two correlations, calibration constants for the hadronic and the electromagnetic energy scale are derived for each TT.

These plots were repeated for every TT separately. From the slopes of the two straight lines a first approximation for calibration constants of the hadronic and the electromagnetic energy scale were derived for every TT. Based on these constants the hadronic and the electromagnetic energy of the TT are scaled linearly before passing them to the JT simulation. From this estimate the calibration constants are further adjusted using a $J/\psi \rightarrow ee$ data sample. During this adjustment the electromagnetic and the hadronic calibration constants are varied simultaneously. This second adjustment was also primarily determined by a linear correlation fit as described above, however some 'eye tuning' (on the order of less than

³By repeating this plot with a pure electron sample, only the upper correlation remains.

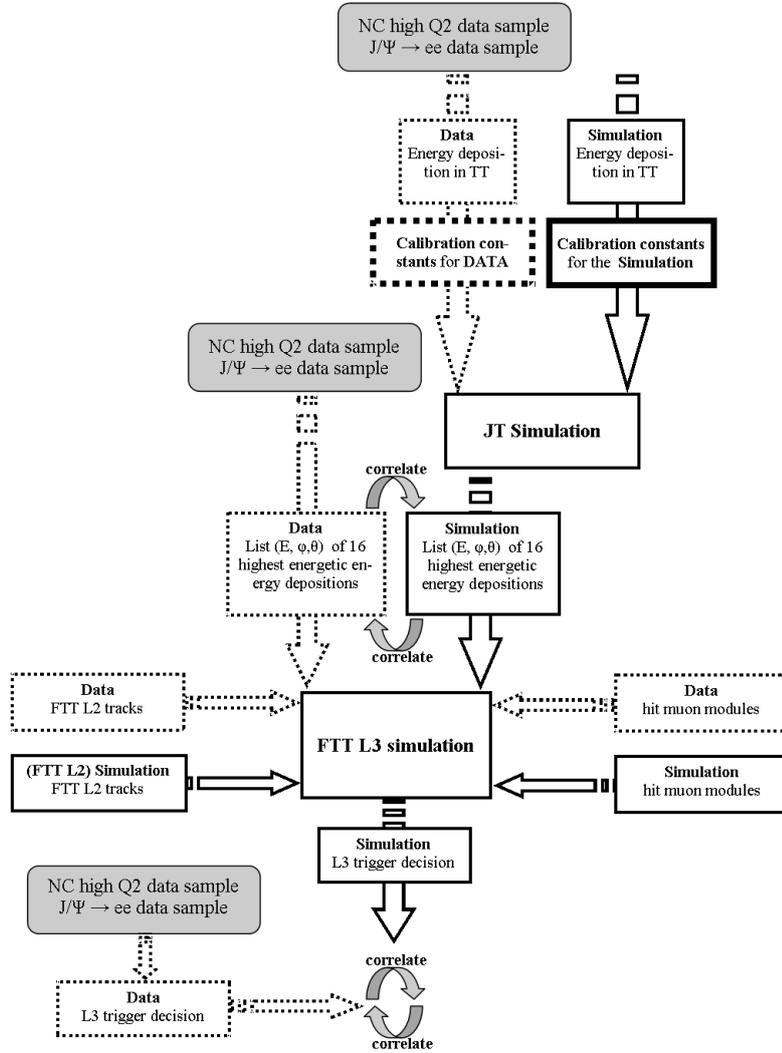


Figure C.2: Sketch of the Calibration Procedure for the JT Simulation
 Schematic representation of the L3 simulation code structure and of the chosen procedure for the calibration of the JT simulation. Every simulation unit (JT simulation, FTT L3 simulation) can switch its input between data (dashed arrows) and simulation (solid arrows). The JT calibration constants were obtained from a NC high Q^2 and a $J/\psi \rightarrow ee$ data sample, by correlating the simulation output with the measured data of the corresponding level.

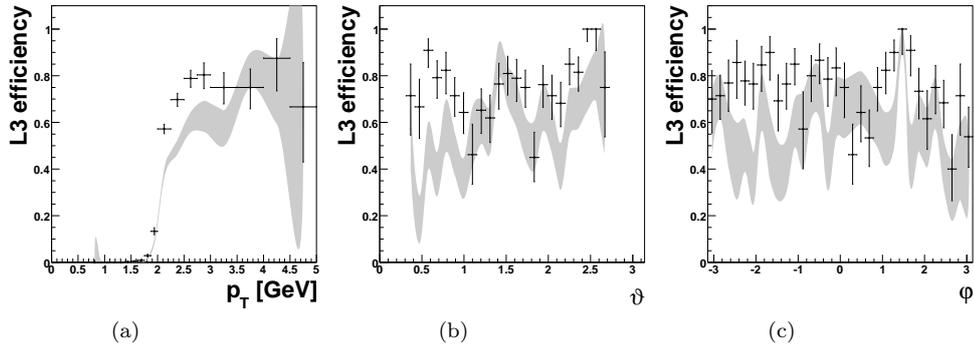


Figure C.3: Control Plot for the L3 Trigger Efficiency, Uncalibrated JT Simulation. The L3 trigger efficiency for single electrons is shown as function of p_T , φ and ϑ of the electron, for the high p_T single tag electron trigger (s34) and for the uncalibrated JT simulation. The measured trigger efficiency (points) is compared to the simulated trigger efficiency (bright band). Input to the JT simulation is the same data. The data is a $J/\psi \rightarrow ee$ sample that is kinematically restricted in a way that the single electron efficiency can be determined. For the φ and ϑ plots the electron is required to be above the p_T -threshold.

10% on the input constants) was necessary. This first estimate for the calibration constants basically fixes the relative ratio between the electromagnetic and the hadronic scale. In the second adjustment the trigger efficiency is tuned with electrons that are in the same kinematic p_T regime, as the electrons from semileptonic beauty decays, i.e. the regime relevant for this analysis.

The derived calibration constants are stored in two BOS banks (AJC1 and AJC2) and saved in the H1 database. AJC1 stores for each TT the electromagnetic and the hadronic calibration constant. AJC2 allows relative shifts of these constants to be applied for Monte Carlo, i.e. the constants in the Monte Carlo used are $ACJ1 * ACJ2$. For this analysis all constants in AJC2 are set to 1, and are varied for the systematics between 0.9 and 1.1.

The effect of this calibration is presented in the Figures C.3 and C.4. Figure C.3 compares for the high- p_T L3 electron subtrigger (s34) the L3 trigger efficiency for single electrons in data with the L3 trigger efficiency obtained from the simulation, before the calibration. The L3 trigger efficiency is presented as function of p_T , φ and ϑ . The points are the trigger efficiency determined from data⁴ directly, the bright gray band is the same trigger efficiency, based on the same data sample, but now determined from the trigger simulation. Obviously the simulation does not describe the measured L3 electron trigger response at all.

In Figure C.4 the same plots are shown, now for all three L3 electron subtriggers. In addition, the single track efficiency obtained from a Single Particle electron Monte Carlos is shown (dark band) for the single tag electron triggers (s29 and s34)⁵. After the calibration of the JT simulation the simulated single electron efficiencies describes the main characteristics of the L3 electron subtriggers in the track variables p_T , φ and ϑ

⁴The data sample was obtained from $J\psi \rightarrow ee$ events. The kinematic phase space is constrained such that one electron is below the p_T -trigger threshold, which allows the determination of the single electron efficiency. For φ and ϑ plots the second electron is required to be above the p_T -threshold.

⁵Since for the double tag trigger at least two electrons need to be reconstructed, these plots were unfortunately not possible with a Single Track Monte Carlo Sample.

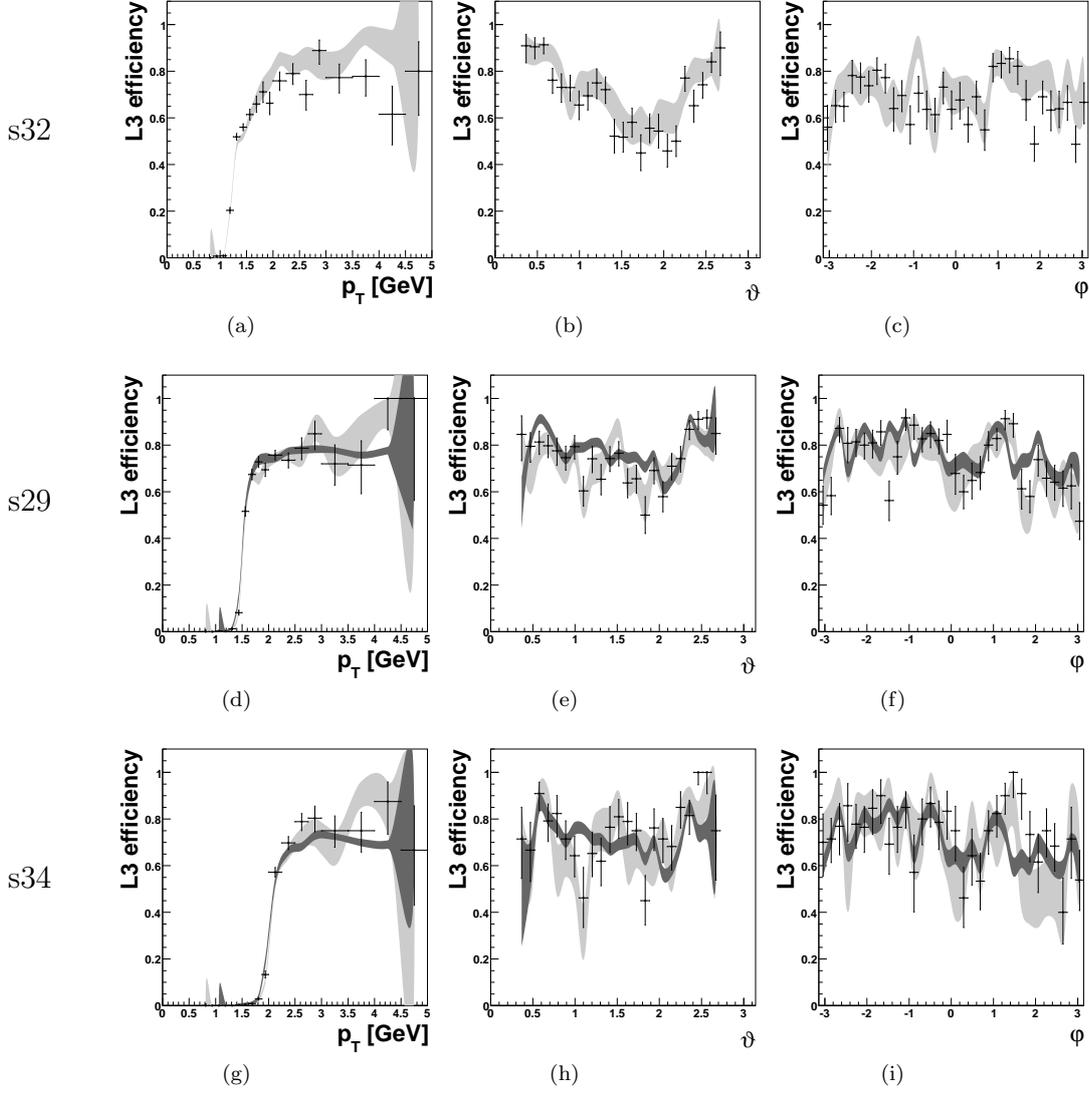


Figure C.4: Control Plot for the L3 Trigger Efficiency, Calibrated JT Simulation.

The L3 trigger efficiency for single electrons is shown as function of p_T , φ and ϑ of the electron, for the three L3 electron subtriggers and for the calibrated JT simulation. The first row presents the trigger efficiency for the low- p_T double tag trigger (s32) (plots (a), (b), (c)), the second row for the medium- p_T single tag subtrigger (s29) (plots (d), (e), (f)) and the third row for the high- p_T single tag subtrigger (s34) (plots (g), (h), (i)). The measured trigger efficiency (points) is compared to the simulated trigger efficiency, based on the data (bright band). The input to the JT simulation is the data. In addition, the L3 trigger efficiency determined from a single particle electron Monte Carlo is given for the single tag triggers (dark band). The data is a $J/\psi \rightarrow ee$ sample that is kinematically restricted (for each subtrigger separately) in a way that the single electron efficiency can be determined. For the φ and ϑ plots the electron is required to be above the p_T -threshold.

Appendix D

The FTT L3 Trigger

D.1 Introduction to the FTT L3 System

The intention of the third trigger level in H1 was to abort the detector readout of non interesting events. The only system that contributed was the third trigger level of the Fast Track Trigger (FTT L3) [61]. In order to profit in terms of a reduced detector dead time, the L3 decision had to be derived faster than the detector readout time of 1–2 ms. During the allowed latency of $\sim 100 \mu\text{s}$, the FTT L3 system performed a partial event reconstruction, based on fitted tracks in the CJC measured by the FTT L2 (see section 3.2.5), energy deposition in the LAr calorimeter measured by the Jet Trigger (JT, see section 3.2.5) and a coarse pattern of the muon modules containing a hit. The FTT L3 triggers were designed to trigger heavy flavor events in photoproduction, by reconstructing decay signatures of heavy quarks. The particles reconstructed on FTT L3 were namely, D^* s, inelastic J/ψ s, electrons, and muons. Each of these particles was reconstructed by a different physics selection algorithm discussed below.

In order to cope with the stringent time constraints to derive the L3 trigger decision, the L3 online event reconstruction was performed on a computer farm, processing in parallel the different physics selection algorithms. The L3 programs are written in the real-time operating system VxWorks and based on the programming language C. The hardware of the FTT L3 finally implemented, consisted of five commercial standard VME PowerPCs, with a 32-bit processor clocked at 450 MHz and with 32 MB memory. The FTT L3 computer farm was directly connected with the FTT L2 system via a fast channel link (throughput of 5 Gbit/s) from where all input data was received.

In the scope of this thesis the reconstruction of muons and electrons¹ was developed and implemented into FTT L3 together with the simulation of the FTT L3 system. These three topics will be presented in this Appendix. The hardware of FTT L3, the data preparation and the performance of the FTT L3 are discussed in a broader context in [73], the read-out of the FTT is documented in [108] and the online monitoring of the FTT (besides lots of other FTT L1 and L2 related information) in [42].

D.2 L3 Physics Algorithms

The reconstruction of the different heavy quark signatures on the third trigger level was based on different selection algorithms. In order to derive the L3 decision as fast as possible,

¹The initial work on electron triggers was done in a diploma thesis [62].

these selection algorithms are all time optimized by the avoidance of double calculations, the use of Taylor expansions and look-up tables. The D^* and J/ψ selection algorithms are purely track based, while the electron and muon selection algorithm combine L2 track information with L1 trigger information from the calorimeter (respectively JT) and the muon system.

D.2.1 D^* and J/ψ Triggers

The D^* and J/ψ selection algorithms are both based on the reconstruction of the invariant mass from the fitted FTT L2 tracks. The D^* mesons are identified in the so called golden decay channel ($D^{*\pm} \rightarrow D^0 \pi_{slow}^\pm \rightarrow K^\pm \pi^\pm \pi_{slow}^\pm$), whereas the J/ψ algorithms exploit semileptonic J/ψ -decays. No particle identification are applied in these algorithms, thus the full combinatorics has to be taken into account. More details can be found in [73].

D.2.2 Muon Trigger

Muons were identified on FTT L3 by linking FTT L2 tracks to muon modules with a muon signature measured on L1. Via a data-bus the muon system (CMD) transmitted to the FTT the hit pattern of its 64 modules. Under the hypothesis of a muon coming from the ep-interaction point, FTT L3 determined for each track the possible triggered muon modules using precalculated look-up tables. The principle is illustrated in Figure D.1. If one module with a positive L1 trigger decision was in coincidence with a looked up module the event was accepted and otherwise rejected.

The FTT muon identification was applied as an additional trigger condition on 2 out of 3 H1 standard muon triggers. By suppressing background from cosmic muons and from problematic beam background, this opened bandwidth as can be seen from Figure D.2. The rate of the muon trigger is shown as function of different runs, before and after the switch on of the L3 condition (indicated by the arrow). Clearly visible: the L3 condition roughly halved the trigger rate. No efficiency losses due to the L3 muon veto were reported by analyses based on these triggers [94], [95]. In addition, a new trigger setup was developed for previously excluded regions in the forward direction of the muon system that suffered from beam background, using the muon identification of the FTT.

D.2.3 Electron Trigger

The L3 electron trigger was developed to allow an efficient online identification of electrons from semileptonic b-quark decays having low transverse momenta, while rejecting background events most effectively [62]. The concept of the L3 electron trigger is based on the idea of matching the FTT L2 track to the energy deposition (cluster) in the LAr calorimeter measured by the JT. Both systems provided particle information about the position in the $\vartheta\varphi$ -plane, the track momentum and the energy deposition, respectively.

The track cluster matching was done by allocating to every JT cluster in the calorimeter all FTT tracks that lay within a quadratic pyramid around the axis determined by the cluster: $\Delta\varphi = |\varphi_{FTT} - \varphi_{JetTrigger}| < \varphi_{cut}$ and, $\Delta\vartheta = |\vartheta_{FTT} - \vartheta_{JetTrigger}| < \vartheta_{cut}$, compare with Figure D.3 (a). Since the FTT track variables were measured at the production vertex, the track cluster match in $\Delta\varphi$ smeared out for strongly curved (low momentum) tracks. Therefore the FTT φ -measurement was optionally corrected for the track curvature, as illustrated in Figure D.3 (b).

For the purpose of particle identification already on trigger level the energy and momentum

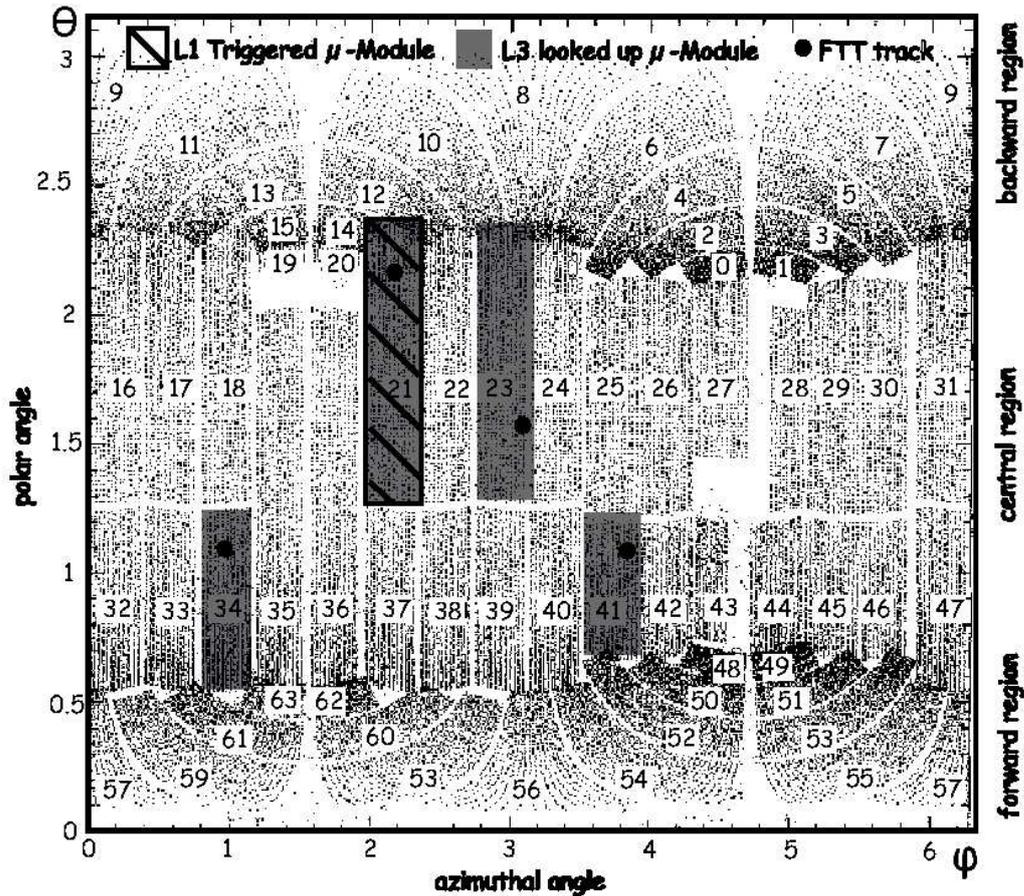


Figure D.1: L3 Muon Trigger Schematic

Hit map of the central muon system as function of the azimuthal and polar angle. The numbers label the 64 different CMD modules. The basic ideas of the muon identification on FTT L3 are illustrated: The points refer to FTT tracks extrapolated to the CMD under the hypothesis of a muon from an ep-interaction. The shaded modules indicate the corresponding linked modules. One L1 triggered module is represented hatched. If one module with a positive L1 trigger decision is in coincidence with a FTT-track linked module the event was accepted.

measurements were correlated. The main focus was on discriminating electrons and pions. In the non compensating LAr calorimeter, the fraction of the detectable energy is smaller for hadrons than for electrons. Therefore, the ratio between the transverse energy measured by the JT and the transverse momentum of the FTT track, $E_{T,JT}/p_{T,FTT}$, allows the distinction of electrons and hadrons: Figure D.4 shows the online measured $E_{T,JT}/p_{T,FTT}$ ratio for a pion and an electron sample obtained from $\rho \rightarrow \pi\pi$ and $J/\psi \rightarrow ee$ decays. The peak position of the two distributions is clearly separated. Electrons are selected on trigger level by a lower cut on $E_{T,JT}/p_{T,FTT}$. An upper $E_{T,JT}/p_{T,FTT}$ -cut is not introduced in order to allow also non isolated electrons to pass the trigger selection.

Three subtrigger elements were implemented with different settings on the p_T -thresholds, the track cluster match, the $E_{T,JT}/p_{T,FTT}$ -cut and the number of online reconstructed electrons. The different trigger setups are listed in Table D.1.

The single electron L3 trigger efficiencies of the different setups are shown in Figure D.5 as function of p_T and determined from $J/\psi \rightarrow ee$ events. The different implemented p_T -thresholds are well recognizable by the sharp rise of the efficiencies, that above the p_T thresholds become flat. The level of the flat plateau is determined by the $E_{T,JT}/p_{T,FTT}$ -cut and the track cluster match condition.

These FTT-JT based subtriggers covered very efficiently a unique phase space, not covered by any other subtrigger. They were not only used in the analysis presented in the first part of this thesis, but also for the measurement of the Proton Structure Function F_L at high Q^2 [87].

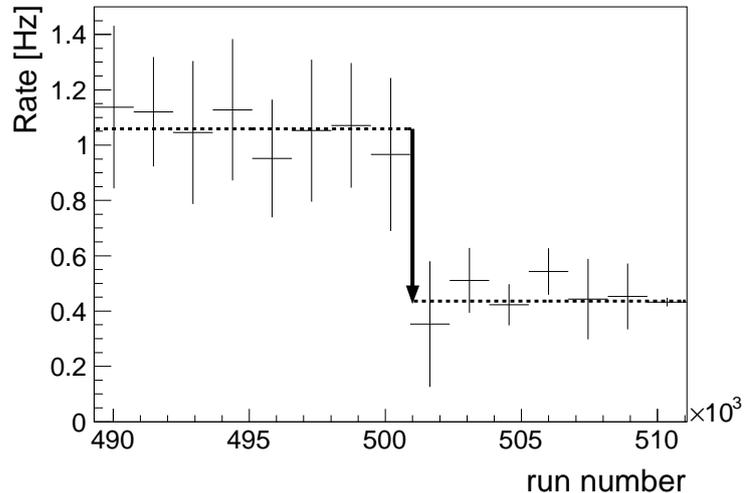


Figure D.2: Trigger Rate of Muon Trigger

The rate of the muon trigger is shown as function of different runs, before and after the switch on of the L3 condition, indicated by the arrow [116].

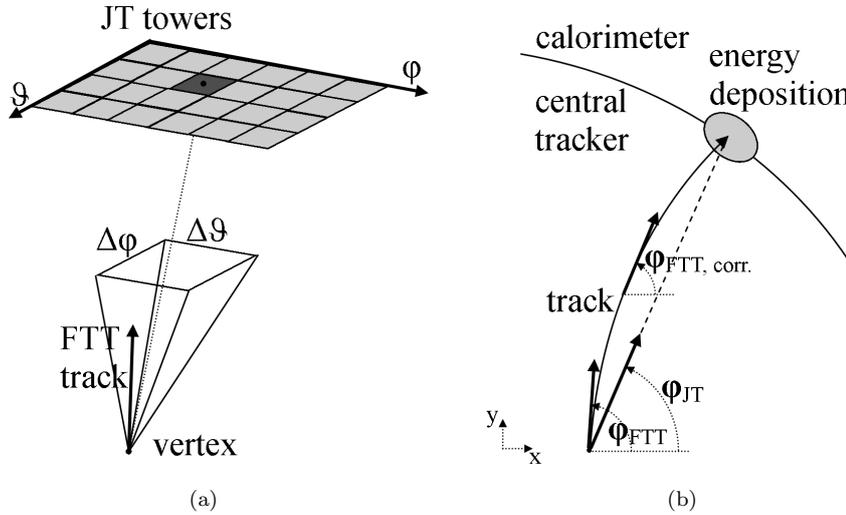


Figure D.3: Track Cluster Match

Left: The track cluster matching is done by allocating to every FTT track a JT cluster in an acceptance window determined by $\Delta\varphi$, $\Delta\vartheta$. Right: The FTT φ -measurement was (optionally) corrected for the track curvature in the transverse plane.

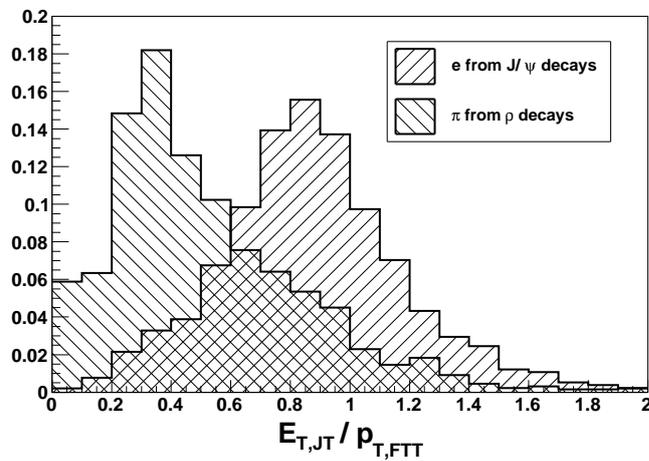


Figure D.4: Online Measured $E_{T, JT}/p_{T, FTT}$ Ratio.

Online measured ratio between the transverse energy deposition in the calorimeter measured by the JT $E_{T, JT}$, and the transverse momentum of the corresponding FTT track $p_{T, FTT}$, for pions and for electrons. A lower cut on $E_{T, JT}/p_{T, FTT}$ allowed the separation of electrons and pions already on trigger level.

Subtrigger	Trigger Element	Curv. Corr.	$\Delta\varphi$ [rad]	$\Delta\vartheta$ [rad]	$\frac{E_{T,JT}}{p_{T,FTT}}$	#electrons	$p_{T,FTT}$ [GeV]
s32	low- p_T , double tag	-	0.3	0.25	0.3	2	1.2
s29	medium- p_T , single tag	✓	0.15	0.2	0.5	1	1.5
s34	high- p_T , single tag	-	0.2	0.2	0.6	1	2.0

Table D.1: Setup of Different Trigger Elements

The settings for the different implemented trigger elements are listed in the table. Note that the curvature correction was only applied on one subtrigger element, for the others a larger acceptance window in $\Delta\varphi$ was chosen.

The first column refers to the subtrigger in which the trigger elements were applied, compare with Table 7.1.

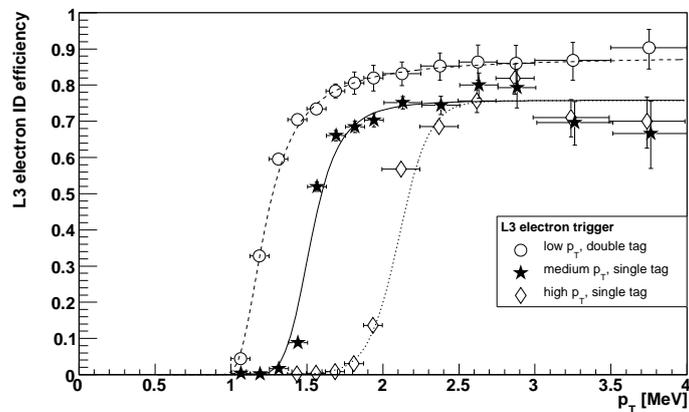


Figure D.5: L3 Electron Trigger Efficiency

L3 trigger efficiency for single electrons as function of the transverse momentum of the electron and for the different settings of the trigger, as listed in Table D.1. The sharp rise corresponds to different p_T -thresholds at 1.2 GeV, 1.5 GeV and 2.0 GeV. The electrons are obtained from data using $J/\psi \rightarrow e^+e^-$ decays. (Also published in [61]).

D.3 L3 Trigger Simulation

The simulation program for the FTT is called FTTEMU and is written in the programming language C. FTTEMU is a hardware like program, which means that the structure and algorithms of the program try to mirror the hardware as close as possible. This is also the case for the three trigger levels of the FTT. The parts of the simulation corresponding to the first two trigger levels are documented in [72]. The implementation of FTTEMU into the global H1 simulation scheme is described in detail in [73]. The trigger simulation of the third trigger level was developed in the scope of this thesis. This part of the program is sketched in the following.

The input information to the L3 system consisted of tracks measured in the CJC by the FTT L2 system, energy depositions in the LAr calorimeter, measured by the JT, and a 64-bit word encoding the muon modules with a hit. Once this information was filled into the computer memory of the PowerPCs of the L3 computer farm, dedicated selection algorithms searched for interesting physic events and derived trigger decisions.² The program loaded into the PowerPCs was written in the programming language C, which allowed a fast commissioning of the system, and a regular refinement of the selection algorithms.

The L3 trigger simulation starts from exactly the same input³ as the L3 system, its output are simulated trigger decisions. Since FTTEMU is also written in the programming language C, the L3 trigger simulation could be based to a great extent on the same code, as was used in the program loaded into the L3 computer farm. However, the realization of this involved the following non trivial tasks:

1. The PowerPCs utilized in the L3 computer farm filled and addressed their computer memory in an optimized, non-standard way.
2. The code of the online selection algorithms contains a lot of (for the simulation) redundant functions related to the read-out, internal communication and communication with the central trigger.
3. The commissioning of the FTT L3 system, and the refinement of the selection algorithms, resulted in several code versions, which were loaded into the hardware at different running periods. The L3 simulation should implement all versions simultaneously and, depending on the run number of the simulated event, automatically call the correct version.⁴

The memory access of the L3 simulation is implemented in the standard way. The first point only influences the program in so far as the sequence in which the memory is filled might differ⁵ between the simulation and the online code.

The other two points are solved with the C preprocessor (cpp). Cpp is a program that preprocesses c-code before the compilation. The trick applied is, that the behavior of cpp is steerable by define-statements saved in a macro. The concept is schematically illustrated in Figure D.6. The different code versions (sketched in the middle of the Figure) can on one hand be compiled to individual programs flagged with a version number. These programs were then loaded into the L3 trigger system and the code version saved in the database. On

²Depending on the time consumption each PowerPC was loaded with one or two selection algorithms.

³The input can either be the one simulated by the L2 simulations, or the one measured by the L2 trigger systems.

⁴ Every run measured in H1 got flagged with a different (run) number. Before the generated Monte Carlo events are passed through the detector response simulation, every event gets a run number assigned, depending on the running period of interest.

Since the program version loaded into the L3 system was saved on the H1 database for every run, the run numbers assigned to the Monte Carlo events are used to steer the L3 simulation executing the correct code version.

⁵For instance the tracks might occur in a different sequence.

the other hand all code versions can be combined and compiled into the L3 trigger simulation. The statements in the preprocessor macro are defined such, that the cpp output of the online compilation slightly differs from the cpp output of the simulation. In particular for the compilation of the L3 simulation cpp is used to filter the redundant code pieces. In addition, the function name of every selection algorithm is extended by the version number during the preprossion step, such that all algorithms can be implemented simultaneously in the L3 simulation.

Performance of the L3 Trigger Simulation

Figure D.7 presents the performance of the L3 simulation. The plot is derived from data, the input to the simulation is the measured output of the L2 systems. The points are the measured L3 trigger bits of each selection algorithm, while the simulated L3 trigger bits are represented in the histogram. The simulated L3 trigger decision agrees with the measured one, as can be expected from the program design. The only effect that is not reflected in the L3 trigger simulation are time constraints present in the L3 trigger: In rare cases [73] the online computing time consumption of a selection algorithm was more than the allowed latency time of the L3 system ($100\ \mu\text{s}$). In these cases the L3 simulation might find a positive trigger decision, while the L3 system did not finish its calculations. However, the effect turned out to be completely negligible.

Besides the general benefits for this and other analyses of having a reliable trigger simulation, the L3 trigger simulation accelerated the commissioning of the selection algorithms for the FTT L3 system. It allowed an offline cut optimization to obtain the anticipated event rates, while not loosing in trigger efficiency and without frequent changes of the online trigger settings [110].

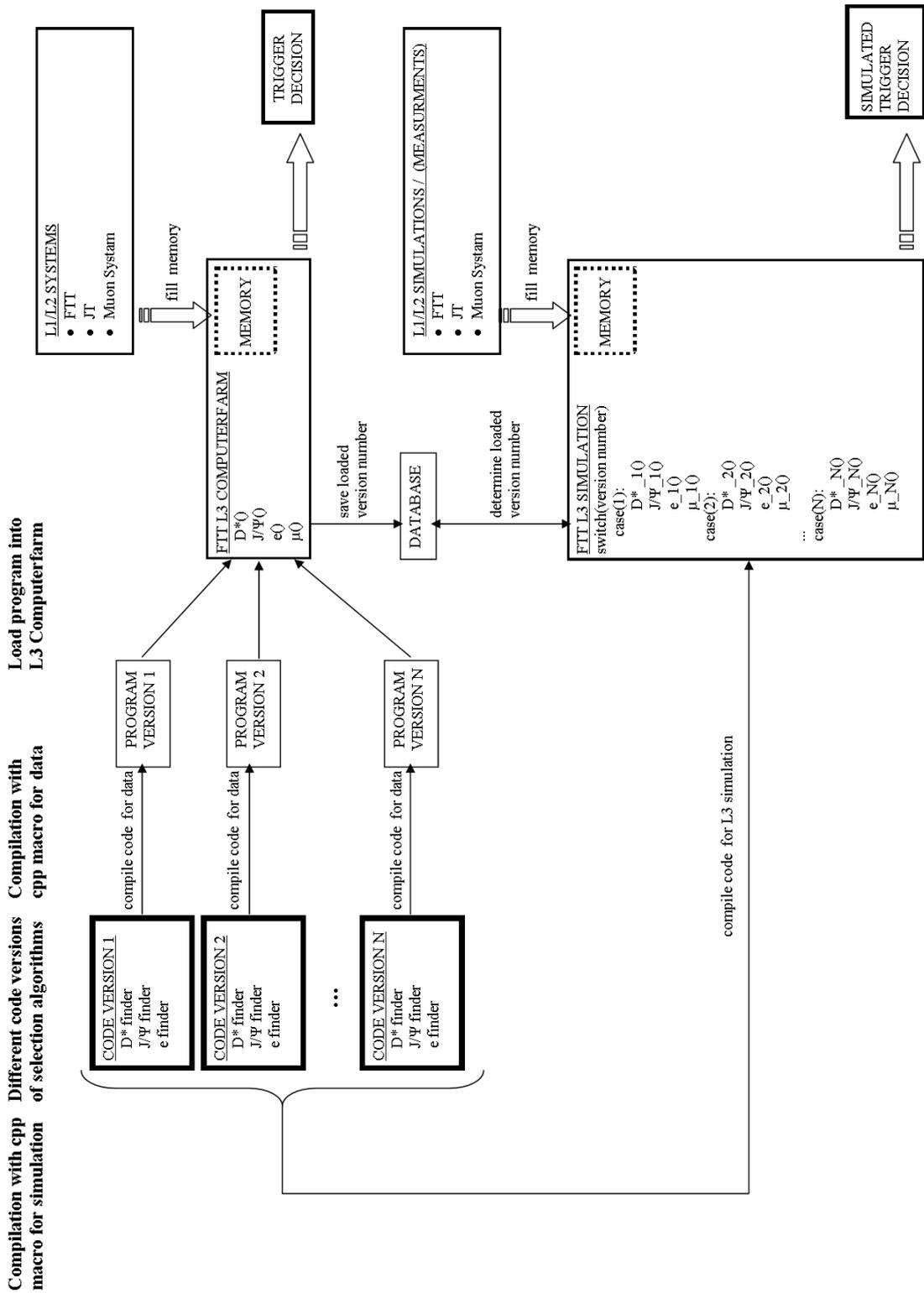


Figure D.6: Sketch Illustrating the Preprocessing of the L3 Program Code. Schematic representation of the online code implementation into the L3 trigger simulation with the help of the c-preprocessor cpp.

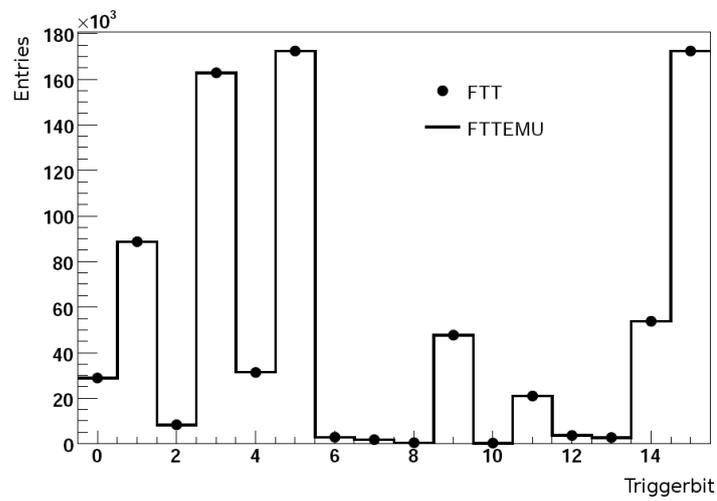


Figure D.7: L3-Triggerbits Online and Offline

The L3-simulation is run on data: the input to the simulation is the measured output of the L2-systems. Each triggerbit corresponds to a different L3 physics selection algorithm. (The plot is adapted from [110]).

Appendix E

Example for the Combination of Subtriggers

E.1 Introduction

This chapter illustrates the calculation of event weights correcting for prescale factors as discussed in section 7.1.2 .

As an example the weights for three subtriggers with the prescale factors 3, 4 and 6 are calculated for each overlapping phase space region, once for uncorrelated subtriggers and once for correlated subtriggers. With correlated subtriggers, the correlations between the subtriggers that occurred in the H1 triggering scheme due to the same L1 trigger condition is meant (compare with the discussion in the sections 3.2.4 and 7.1.2).

As a model illustrating the difference between correlated and uncorrelated subtriggers (of this example) one could think of three spinning cogwheels with 3, 4 and 6 cogs. If the 3rd, 4th or 6th tooth points upwards, it should open the corresponding trigger (see Figures E.1). Correlated triggers are represented by interleaved cogwheels, while uncorrelated triggers are freely turning cogwheels.

Before calculating the details, it is useful to remind the following points:

- The event weight has to be determined for each overlapping phase space region separately (see Figure 7.1).
- The overlapping phase space regions are defined by the subtriggers considered, and can be addressed by their raw trigger bits (RawTb), cf. with the discussion in section 7.1.2.
- The actual opening of the subtrigger is determined by the assigned prescale factor. If a subtrigger is allowed to open the trigger by the CTD subtrigger counter (see section 3.2.4), the term 'the subtrigger **fires**' is used in the following.
- The event weight w is the prescale factor pf of the corresponding phase space region (equation (7.1)):

$$w = pf \tag{E.1}$$

- The probability for an event to fulfill the prescale condition (i.e. to fire) is the inverse of the event weight:

$$P_{fire} = w^{-1} = pf^{-1} \tag{E.2}$$

E.2 Uncorrelated Subtriggers

The probability that one of the considered subtriggers (cogwheels) fires $P_{fire,uncorr}$ is for uncorrelated subtriggers calculable as follows:

If the probability that none of them fires is $P_{no\ fire}$, P_{fire} can be rewritten as

$$P_{fire} = 1 - P_{no\ fire} \quad . \quad (E.3)$$

Since the probability that the subtrigger i does not fire is $1 - 1/pf_i$, $P_{no\ fire,uncorr}$ (RawTb) can be expressed as

$$P_{no\ fire,uncorr}(\text{RawTb}) = \prod_{i=\text{subtrig} \in \text{RawTb}} \left(1 - \frac{1}{pf_i}\right) \quad , \quad (E.4)$$

where the product is performed over all subtriggers that overlap in the phase space considered. By combining equation (E.3) and (E.4) one gets

$$P_{fire,uncorr}(\text{RawTb}) = 1 - \prod_{i=\text{subtrig} \in \text{RawTb}} \left(1 - \frac{1}{pf_i}\right) \quad , \quad (E.5)$$

which in combination with equation (E.2) is equation (7.2).

Table E.1 lists the corresponding weights of this example, i.e. for three uncorrelated subtriggers with prescale factors 3, 4 and 6. Each overlapping phase space region is listed in a different row.

E.3 Correlated Subtriggers

The probability that at least one of the considered subtriggers (cogwheels) fires $P_{fire,corr}$ can for correlated subtriggers be determined by either writing a long list, or by applying equation (7.3). Both methods are demonstrated for the example considered. The list-method is simpler to understand, however can become very expensive in terms of (computer) computation time.

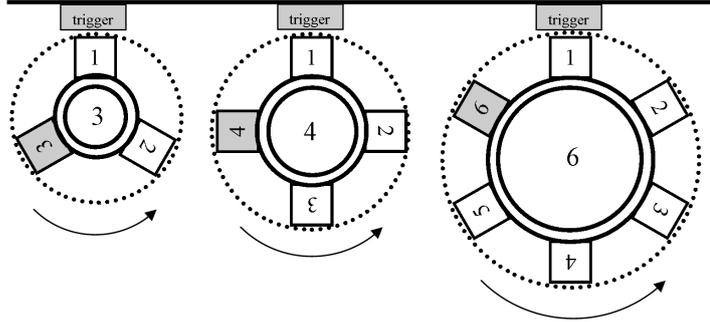


Figure E.1: Cogwheel Model

Model illustrating the three subtriggers with prescale factors 3, 4 and 6: Each subtrigger is represented by a turning cogwheel. If the 3rd, 4th or 6th (shaded) tooth points upwards, the trigger fires. The cogwheels of correlated subtriggers are interleaved, while cogwheels of uncorrelated subtriggers are spinning independently.

E.3.1 Method I - List

The method is best explained by the cogwheel model: $P_{fire,corr}$ is determined by turning the cogwheels corresponding to the overlapping phase space considered, and by simply counting the number of times that at least one of them opens the trigger.

For the three subtriggers considered in this example and for all overlapping phase space regions this is done in a systematic way in the Table E.2: The overlapping phase space regions are indicated in the first three columns by their bit pattern. The numerical series below 'Tooth-Number Sequence' correspond to the tooth-number sequence on the three cogwheels (see Figure E.1). The first row belongs to the cogwheel 3, followed by the cogwheels 4 and 6. The tooth-numbers that open the trigger are printed in bold.

$P_{fire,corr}$ is determined for every overlapping phase space region, by first selecting the corresponding cogwheels. They are then turned until they are again in the starting position. Every time a new tooth points upwards and fires a 'x' is noted, otherwise a '-'. The number of entries in these series of 'x' and '-' corresponds to the least common multiple (LCM) of the prescale factors considered, so in the case of $\{3, 6\}$, $LCM(\{3, 6\}) = 6$ and in the case of $\{3, 4, 6\}$, $LCM(\{3, 4, 6\}) = 12$, etc. .

The probability $P_{fire,corr}$ is determined by the number of 'x' in the line divided by the corresponding LCM, $P_{fire,corr} = \frac{\#\text{'x'}}{LCM}$.

E.3.2 Method II - Formula

The expression given in equation (7.3) does exactly the same but in a more direct way. According to equation (7.3) $P_{fire,corr}$ is

$$P_{fire,corr}(RawTb) = \sum_{s=1}^{N_{pf,RawTb}} \left(\sum_{i=1}^{N_{C(\mathbb{A}_{pf}(RawTb),s)}} \frac{(-1)^{s+1}}{LCM(C(\mathbb{A}_{pf}(RawTb),s,i))} \right) . \quad (E.6)$$

The notation is as follows:

- $\mathbb{A}_{pf}(RawTb)$ is the set of prescale factors of all subtriggers that overlap in the phase space region given by $RawTb$.

pf / cogwheel			$P_{fire,uncorr}$	weight
3	4	6		
1	0	0	$\frac{1}{3}$	3
0	1	0	$\frac{1}{4}$	4
0	0	1	$\frac{1}{6}$	6
1	1	0	$1 - (1 - \frac{1}{3})(1 - \frac{1}{4}) = \frac{1}{2}$	2
1	0	1	$1 - (1 - \frac{1}{3})(1 - \frac{1}{6}) = \frac{9}{4}$	$\frac{4}{9}$
0	1	1	$1 - (1 - \frac{1}{4})(1 - \frac{1}{6}) = \frac{8}{3}$	$\frac{3}{8}$
1	1	1	$1 - (1 - \frac{1}{3})(1 - \frac{1}{4})(1 - \frac{1}{6}) = \frac{7}{12}$	$\frac{12}{7}$

Table E.1: Event Weights of Uncorrelated Subtriggers

The event weights correcting for prescale factors of three uncorrelated subtriggers with the prescale factors 3, 4 and 6 are listed. Each row corresponds to a different overlapping phase space region encoded in the bit pattern of the first three columns.

- $N_{pf,RawTb}$ is the number of subtriggers that overlap.
- $C(\mathbb{A}, s)$ are the combinations of size s of the set \mathbb{A} , i.e. the un-ordered collection of distinct elements. $C(\mathbb{A}, s, i)$ is the i^{th} combination of size s . $N_{C(\mathbb{A}, s)}$ is the number of combinations of size s .
For example, given the set of letters $\mathbb{L} = \{a, b, c\}$ the combinations of size 2 are $\{a, b\}$, $\{a, c\}$, and $\{b, c\}$. This are 3 combinations of size 2, $N_{C(\mathbb{L}, 2)} = 3$, and the 3rd element of size 2 is $C(\mathbb{L}, 2, 3) = \{b, c\}$.
- $LCM(\mathbb{A})$ is the least common multiple of the numbers in the set \mathbb{A} .

Writing equation (E.6) explicitly out for three subtriggers with prescale factors a , b and c and for the phase space where all three subtriggers overlap ($RawTb = 111$), one gets

$$\begin{aligned}
 P_{fire,corr}(RawTb = 111) &= \underbrace{\frac{(-1)^{1+1}}{LCM(\{a\})} + \frac{(-1)^{1+1}}{LCM(\{b\})} + \frac{(-1)^{1+1}}{LCM(\{c\})}}_{s=1 \quad N_{C(\{a,b,c\},1)}=3} \\
 &+ \underbrace{\frac{(-1)^{2+1}}{LCM(\{a,b\})} + \frac{(-1)^{2+1}}{LCM(\{b,c\})} + \frac{(-1)^{2+1}}{LCM(\{a,c\})}}_{s=2 \quad N_{C(\{a,b,c\},2)}=3} \\
 &+ \underbrace{\frac{(-1)^{3+1}}{LCM(\{a,b,c\})}}_{s=3 \quad N_{C(\{a,b,c\},3)}=1}, \quad (E.7)
 \end{aligned}$$

where the curly brace enclose the inner sum of equation (E.6).

The event weights for the three subtriggers considered in this example and for all overlapping phase space regions are calculated in Table E.3 with equation (E.6).

pf / cogwheel			Tooth-Number Sequence												$P_{fire,corr}$ $= \frac{\#x'}{LCM}$	weight																	
3	4	6	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3				
1	0	0	-	-	x																											$\frac{1}{3}$	3
0	1	0	-	-	-	x																										$\frac{1}{4}$	4
0	0	1	-	-	-	-	-	x																								$\frac{1}{6}$	6
1	1	0	-	-	x	x	-	x	-	x	x	-	-	x																		$\frac{6}{12}$	2
1	0	1	-	-	x	-	-	x																								$\frac{2}{6}$	3
0	1	1	-	-	-	x	-	x	-	x	-	-	-	x																		$\frac{6}{12}$	3
1	1	1	-	-	x	x	-	x	-	x	x	-	-	x																		$\frac{6}{12}$	2

Table E.2: Event Weights of Correlated Subtriggers - Method I
The event weights correcting for prescale factors of three correlated subtriggers with the prescale factors 3, 4 and 6 are listed. Each row corresponds to a different overlapping phase space region encoded in the bit pattern of the first three columns. The columns below 'Cog-Number Sequence' and the method to calculate the event weights are discussed in the text.

pf / cogwheel			$P_{fire,corr}$, equation (E.6)	weight
3	4	6		
1	0	0	$\left(\frac{1}{LCM(3)}\right) = \frac{1}{3}$	3
0	1	0	$\left(\frac{1}{LCM(4)}\right) = \frac{1}{4}$	4
0	0	1	$\left(\frac{1}{LCM(6)}\right) = \frac{1}{6}$	6
1	1	0	$\left(\frac{1}{LCM(3)}\right) + \left(\frac{1}{LCM(4)}\right) + \left(\frac{-1}{LCM(\{3,4\})}\right) = \frac{1}{3} + \frac{1}{4} - \frac{1}{12} = \frac{1}{2}$	2
1	0	1	$\left(\frac{1}{LCM(3)}\right) + \left(\frac{1}{LCM(6)}\right) + \left(\frac{-1}{LCM(\{3,6\})}\right) = \frac{1}{3} + \frac{1}{6} - \frac{1}{6} = \frac{1}{3}$	3
0	1	1	$\left(\frac{1}{LCM(4)}\right) + \left(\frac{1}{LCM(6)}\right) + \left(\frac{-1}{LCM(\{4,6\})}\right) = \frac{1}{4} + \frac{1}{6} - \frac{1}{12} = \frac{1}{3}$	3
1	1	1	$\left(\frac{1}{LCM(3)}\right) + \left(\frac{1}{LCM(4)}\right) + \left(\frac{1}{LCM(6)}\right) + \left(\frac{-1}{LCM(\{3,6\})} + \frac{-1}{LCM(\{3,4\})} + \frac{-1}{LCM(\{4,6\})}\right) + \left(\frac{1}{LCM(\{3,4,6\})}\right)$ $= \frac{1}{3} + \frac{1}{4} + \frac{1}{6} - \frac{1}{12} - \frac{1}{12} - \frac{1}{6} + \frac{1}{12} = \frac{1}{2}$	2

Table E.3: Event Weights of Correlated Subtriggers - Method II
The event weights correcting for prescale factors of three correlated subtriggers with the prescale factors 3, 4 and 6 are listed. Each row corresponds to a different overlapping phase space region encoded in the bit pattern of the first three columns. The event weights are calculated with equation (E.6).

List of Figures

2.1	Coupling Constant in QED and QCD	4
2.2	Lepton Proton Scattering	5
2.3	Gluon Ladder	9
2.4	Schematic Overview on the DGLAP, BFKL and CCFM Evolution Equation	10
2.5	Heavy Quark Photoproduction Processes in LO QCD	12
3.1	The HERA Collider	16
3.2	The H1 Detector	17
3.3	Segmentation of the LAr Calorimeter	20
3.4	Illustration to the FTT Track Finding.	23
3.5	Granularity of the JT	24
4.1	Feynman Diagrams for $b\bar{b} \rightarrow eeX$ Decays.	27
4.2	Simulated Differential Beauty Cross Sections	28
4.3	$p_T(e)$ -Acceptance	29
5.1	Schematic Illustration of the Event Simulation in the H1 Experiment	33
6.1	Illustration of the Calorimeter Volume Considered	39
6.2	NN Input Variable E/p	40
6.3	NN Input Variable Isolation	41
6.4	NN Input Variable S_{rad}	42
6.5	NN Input Variable S_{len}	43
6.6	NN Input Variable E_f	44
6.7	Signal Efficiency vs. Background Rejection	46
6.8	NN Output Variable MLP	47
6.9	Inelastic J/ψ Peak in DIS	48
7.1	Overlapping Subtriggers	52
7.2	L3 Single Electron Efficiency for s29	54
7.3	L1-L3 Combined Trigger Efficiency for Beauty Events	55
7.4	Control Distributions for $\sum(E - P_z)$	61
7.5	Control Distributions for Highest p_T -Track	62
7.6	Control Distributions for 2 nd Highest p_T -Track	63
7.7	Control Distributions for 3 rd Highest p_T -Track	64
7.8	Control Distributions for Highest p_T -Electron Track	65
7.9	Control Distributions for 2 nd Highest p_T -Electron Track	66
7.10	Control Distributions for all Tracks	67

LIST OF FIGURES

8.1	Illustration of the Thrust Axis in the Transverse Plane	71
8.2	Reconstructed vs. Generated Transverse Beauty Mass.	72
8.3	Sketch of the Response Histogram	74
8.4	Binning of $\langle m_T(b) \rangle$	75
8.5	Illustration of the Variable $\Delta\phi_{e_1, e_2, sgn}$	77
8.6	Response Histogram	79
8.7	Control Distributions for the Unfolding	80
8.8	Final Distributions	86
8.9	Final Distributions	87
8.10	Electron Identification Systematics	88
8.11	JT Calibration Systematics	89
8.12	Differential Beauty Cross Sections	91
8.13	Relative Error of the Background	92
8.14	Beauty Photoproduction Cross Section	93
8.15	η Correction Factors	94
8.16	Beauty Photoproduction at HERA.	95
A.1	Response Histogram	102
B.1	Electromagnetic Showers	109
B.2	Performance of the Dead Material Correction	110
C.1	Measured vs. Simulated JT Energy, before the Calibration	112
C.2	Sketch of the Calibration Procedure for the JT Simulation	113
C.3	Control Plot for the L3 Trigger Efficiency, Uncalibrated JT Simulation.	114
C.4	Control Plot for the L3 Trigger Efficiency, Calibrated JT Simulation.	115
D.1	L3 Muon Trigger Schematic	119
D.2	Trigger Rate of Muon Trigger	120
D.3	Track Cluster Match	121
D.4	Online Measured $E_{T, JT}/p_{T, FTT}$ Ratio.	121
D.5	L3 Electron Trigger Efficiency	122
D.6	Sketch Illustrating the Preprocessing of the L3 Program Code.	125
D.7	L3-Triggerbits Online and Offline	126
E.1	Cogwheel Model	128

List of Tables

3.1	Legend to Figure 3.2	18
4.1	Kinematical Cuts	26
5.1	Overview of the MC Samples	34
6.1	Track Selection Cuts	36
6.2	Input Variables to the NN	38
6.3	Training Sample	39
7.1	Definitions of FTT-JT Based Electron Subtriggers	51
7.2	Integrated Luminosity of the L3 Electron Subtriggers	51
7.3	Event Weights to Correct for Prescale Factors	53
7.4	Overview on the Selection Cuts	57
8.1	Binning	72
8.2	MC Samples Used to Determine the Response Matrix \mathbf{A}	85
8.3	Summary of Systematic Errors	85
8.4	Unfolding Results	92
8.5	Correlation Matrix	92
8.6	Measured Differential Beauty Cross Section Values	96
D.1	Setup of Different Trigger Elements	122
E.1	Event Weights of Uncorrelated Subtriggers	129
E.2	Event Weights of Correlated Subtriggers - Method I	130
E.3	Event Weights of Correlated Subtriggers - Method II	131

List of Acronyms

B

BC	Bunch Crossing.
BFKL	Balitsky, Fadin, Kuraev, Lipatov.
BOS	Bank Object System.
BPC	Backward Proportional Chamber.
BR	Branching Ratio.
BSM	Beyond the Standard Model.
BST	Backward Silicon Tracker.

C

CC	Charged Current.
CCFM	Ciafaloni, Catani, Fiorani, Marchesini.
CDAQ	Central Data Acquisition.
CIP	Central Inner Proportional Chamber.
CJC	Central Jet Chamber.
CMD	Central Muon Detector.
COP	Central Outer Proportional Chamber.
COZ	Central Outer z Chamber.
CPLD	Complex Programmable Logic Device.
Cpp	C Preprocessor.
CST	Central Silicon Tracker.
CTL	Central Trigger Logic.

D

DAQ	Data Acquisition.
DESY	Deutsches Elektronen-Synchrotron (German electron synchrotron).
DGL	Dokshitzer, Gribov, Lipatov, Altarelli, Parisi.
DIS	Deep Inelastic ep Scattering.
DST	Detector Summary Tape.

E

ep	electron-proton.
-----------	------------------

F

FMD	Forward Muon Detector.
FMNR	Frixione, Mangano, Nason, Ridol.
FPGA	Field Programmable Gate Array.
FST	Forward Silicon Tracker.
FTD	Forward Tracking Detector.
FTT	Fast Track Trigger.

H

HERA	Hadron Elektron Ring Anlage (Hadron electron ring facility).
HFS	Hadronic Final State.

I

IP	Internet Protocol.
IT	Input Tower.

J

JT	Jet Trigger.
-----------	--------------

L

L1	Trigger Level 1.
L2	Trigger Level 2.
L2NN	Level 2 Neural Network Trigger.
L2TT	Level 2 Topological Trigger.
L3	Trigger Level 3.
L45	Trigger Level 4/5.
LAr	Liquid Argon Calorimeter.
LCM	Least Common Multiple.
LO	Leading Order.

M

MC	Monte Carlo.
-----------	--------------

LIST OF ACRONYMS

N

- NC** Neutral Current.
NLO Next to Leading Order.
NN (Artificial) Neural Network.

P

- PC** Personal Computer.
PDF Parton Distribution Function.
pQCD Perturbative QCD.
PQZP Parallel Quickbus Zero-suppressed Processor.
PRT Proton Remnant Tagger.

Q

- QCD** Quantum Chromo-Dynamics.
QED Quantum Electro-Dynamics.
Qt Charge - Time.

S

- SM** Service Module.
SpaCal Spaghetti Calorimeter.
ST Sub-Trigger.

T

- TE** Trigger Element.
TMVA Toolkit for Multivariate Data Analysis.
TT Trigger Tower.

U

- uPDF** uninintegrated Parton Distribution Function.

Bibliography

- [1] S. Frixione, M. L. Mangano, P. Nason and G. Ridolfi, “*Heavy-Quark Production*”, *Adv.Ser.Direct.High Energy Phys.*, **15** 609–706, 1998.
- [2] A. D. Martin F. Halzen, *Quarks and leptons*. John Wiley & Sons, 1984.
- [3] D. J. Gross and F. Wilczek, “*Asymptotically Free Gauge Theories. 1*”, *Phys. Rev.*, **D8** 3633–3652, 1973.
- [4] H. D. Politzer, “*Reliable Perturbative Results for Strong Interactions?*”, *Phys. Rev. Lett.*, **30** 1346–1349, 1973.
- [5] D. J. Gross and F. Wilczek, “*Ultraviolet Behavior of Non-Abelian Gauge Theories*”, *Phys. Rev. Lett.*, **30** 1343–1346, 1973.
- [6] H. D. Politzer, “*Asymptotic Freedom: An Approach to Strong Interactions*”, *Phys. Rept.*, **14** 129–180, 1974.
- [7] R.P. Feynman, *Quantum Electrodynamics*. W.A. Benjamin, Inc. , New York, 1961.
- [8] S. S. Schweber, “*QED and the men who made it: Dyson, Feynman, Schwinger, and Tomonaga*”, Princeton, USA: Univ. Pr. (1994) 732 p.
- [9] C. Adloff et al., [H1], “*Measurement and QCD analysis of neutral and charged current cross sections at HERA*”, *Eur. Phys. J.*, **C30** 1–32, 2003.
- [10] R. P. Feynman, “*Photon-hadron interactions*”, Reading 1972, 282p.
- [11] H1 Collaboration, “*A Precision Measurement of the Inclusive ep Scattering Cross Section at HERA*”, 2009.
- [12] G. Sterman et al., “*Handbook of perturbative QCD*”, *Rev. Mod. Phys.*, **67**(1) 157–248, Jan 1995.
- [13] W. Stirling K. Ellis and B. Webber, *QCD and Collider Physics*. Cambridge Monographs, 2003.
- [14] R. K. Ellis, H. Georgi, M. Machacek, H. D. Politzer and G. G. Ross, “*Perturbation Theory and the Parton Model in QCD*”, *Nucl. Phys.*, **B152** 285, 1979.
- [15] J. C. Collins, D. E. Soper and G. Sterman, “*Factorization of Hard Processes in QCD*”, *Adv. Ser. Direct. High Energy Phys.*, **5** 1–91, 1988.
- [16] J. C. Collins and D. E. Soper, “*The Theorems of Perturbative QCD*”, *Ann. Rev. Nucl. Part. Sci.*, **37** 383–409, 1987.
- [17] V. N. Gribov and L. N. Lipatov, “ *$e^+ e^-$ pair annihilation and deep inelastic $e p$ scattering in perturbation theory*”, *Sov. J. Nucl. Phys.*, **15** 675–684, 1972.
- [18] V. N. Gribov and L. N. Lipatov, “*Deep inelastic $e p$ scattering in perturbation theory*”, *Sov. J. Nucl. Phys.*, **15** 438–450, 1972.
- [19] G. Altarelli and G. Parisi, “*Asymptotic Freedom in Parton Language*”, *Nucl. Phys.*, **B126** 298, 1977.
- [20] Y. L. Dokshitzer, “*Calculation of the Structure Functions for Deep Inelastic Scattering and $e^+ e^-$ Annihilation by Perturbation Theory in Quantum Chromodynamics*”, *Sov. Phys. JETP*, **46** 641–653, 1977.

- [21] I. I. Balitsky and L. N. Lipatov, “*The Pomeron Singularity in Quantum Chromodynamics*”, Sov. J. Nucl. Phys., **28** 822–829, 1978.
- [22] E. A. Kuraev, L. N. Lipatov and Victor S. Fadin, “*The Pomeron Singularity in Nonabelian Gauge Theories*”, Sov. Phys. JETP, **45** 199–204, 1977.
- [23] M. Ciafaloni, “*Coherence Effects in Initial Jets at Small q^2/s* ”, Nucl. Phys., **B296** 49, 1988.
- [24] S. Catani, F. Fiorani and G. Marchesini, “*QCD Coherence in Initial State Radiation*”, Phys. Lett., **B234** 339, 1990.
- [25] S. Catani, F. Fiorani and G. Marchesini, “*Small x Behavior of Initial State Radiation in Perturbative QCD*”, Nucl. Phys., **B336** 18, 1990.
- [26] G. Marchesini, “*QCD coherence in the structure function and associated distributions at small x* ”, Nucl. Phys., **B445** 49–80, 1995.
- [27] R. S. Thorne and W. K. Tung, “*PQCD Formulations with Heavy Quark Masses and Global Analysis*”, 2009.
- [28] O. Behnke, “*Production of Charm and Beauty Quarks at HERA*”, Ruprecht-Karls- Univ. Heidelberg, 2005, Habilitation.
- [29] C. Peterson, D. Schlatter, I. Schmitt and Peter M. Zerwas, “*Scaling Violations in Inclusive $e^+ e^-$ Annihilation Spectra*”, Phys. Rev., **D27** 105, 1983.
- [30] V. G. Kartvelishvili, A. K. Likhoded and V. A. Petrov, “*On the Fragmentation Functions of Heavy Quarks Into Hadrons*”, Phys. Lett., **B78** 615, 1978.
- [31] M. G. Bowler, “ *$e^+ e^-$ Production of Heavy Quarks in the String Model*”, Zeit. Phys., **C11** 169, 1981.
- [32] C. F. von Weizsäcker, “*Radiation emitted in collisions of very fast electrons*”, Z. Phys., **88** 612–625, 1934.
- [33] E. J. Williams, “*Nature of the high-energy particles of penetrating radiation and status of ionization and radiation formulae*”, Phys. Rev., **45** 729–730, 1934.
- [34] S. Frixione, M. L. Mangano, P. Nason and G. Ridolfi, “*Improving the Weizsäcker-Williams Approximation in Electron-Proton Collisions*”, Phys.Lett. B, **319** 339–345, 1993.
- [35] S. Frixione, M. L. Mangano, P. Nason and G. Ridolfi, “*Total Cross Sections for Heavy Flavour Production at HERA*”, Phys. Lett., **B348** 633–645, 1995.
- [36] S. Frixione, M. L. Mangano, P. Nason and G. Ridolfi, “*Heavy quark correlations in photon - hadron collisions*”, Nucl. Phys., **B412** 225–259, 1994.
- [37] M. L. Mangano, P. Nason and G. Ridolfi, “*Heavy quark correlations in hadron collisions at next-to- leading order*”, Nucl. Phys., **B373** 295–345, 1992.
- [38] M. Seidel and H. DESY, “*Luminosity upgrade of HERA*”, **1**, 1999.
- [39] F. D. Aaron et al., [H1], “*Measurement of the Proton Structure Function F_L at Low x* ”, Phys. Lett., **B665** 139–146, 2008.
- [40] I. Abt et al., [H1], “*The H1 detector at HERA*”, Nucl. Instrum. Meth., **A386** 310–347, 1997.
- [41] I. Abt et al., [H1], “*The Tracking, calorimeter and muon detectors of the H1 experiment at HERA*”, Nucl. Instrum. Meth., **A386** 348–396, 1997.
- [42] N. Berger, *Measurement of Diffractive Phi Meson Photoproduction at HERA with the H1 Fast Track Trigger*, PhD thesis, ETH Zurich, 2007.
- [43] B. Andrieu et al., [H1 Calorimeter Group], “*The H1 liquid argon calorimeter system*”, Nucl. Instrum. Meth., **A336** 460–498, 1993.
- [44] B. Andrieu et al., [H1 Calorimeter Group], “*Beam tests and calibration of the H1 liquid argon calorimeter with electrons*”, Nucl. Instrum. Meth., **A350** 57–72, 1994.

-
- [45] B. Andrieu et al., [H1 Calorimeter Group], “*Results from pion calibration runs for the H1 liquid argon calorimeter and comparisons with simulations*”, Nucl. Instrum. Meth., **A336** 499–509, 1993.
- [46] H. Bethe and W. Heitler, “*On the Stopping of fast particles and on the creation of positive electrons*”, Proc. Roy. Soc. Lond., **A146** 83–112, 1934.
- [47] H1 collaboration, “*Proposal for an Upgrade of the H1 Luminosity System and its Associated Electronics for HERA 2000*”, Proposal submitted to the DESY PRC 98/05, 1998.
- [48] <http://www.h1.desy.de/h1/www/h1det/lumi/>.
- [49] L. Goerlich et al. Cracow, “*Strategy Studies for the H1 Topological L2-Trigger (L2TT)*”, **H1-IN-508(01/1997)**, 1997.
- [50] H. Kolanoski et al., “*Concept and design of the fast h1 second level trigger using artificial neural networks, Prepared for International Conference on Computing in High-energy Physics (CHEP 95)*”, Prepared for International Conference on Computing in High-energy Physics (CHEP 95), 1995.
- [51] J. Mock et al., “*Artificial neural networks as a second-level trigger at the H1 experiment: Performance analysis and results*”, Prepared for 4th International Workshop on Software Engineering and Artificial Intelligence for High-energy and Nuclear Physics (AIHENP 95), Pisa, Italy, 1995.
- [52] M. C. Urban, *The new CIP2k z-Vertex Trigger for the H1 Experiment at HERA*, PhD thesis, Univ. Zurich, 2004.
- [53] M. Cuje et al., “*H1 high luminosity upgrade 2000 cip and level 1 vertex trigger*”, DESY-PRC 98/02 and H1 internal note H1-01/98-535, 1998.
- [54] D. Müller and F. Sefkow, “*Improving the z-Vertex Trigger*”, H1 internal note H1-04/98-539, 1998.
- [55] J. Becker, “*H1 CIP-Vertex-Trigger upgrade - Triggerhardware-Simulation*”, Internal note Physikalisches Institut, Ruprecht-Karls-Universität Heidelberg, 1998.
- [56] J. Naumann et al., “*A fast high-resolution track trigger for the H1 experiment*”, **48(4)** 1276–1281, Aug. 2001.
- [57] D. Meer, D. Müller, J. Müller, A. Schöning and Ch. Wissing, “*A Multifunctional Processing Board for the Fast Track Trigger of the H1 Experiment*”, IEEE Trans.Nucl.Sci., **49** 357–361, 2002.
- [58] A. Schöning, [H1], “*A fast track trigger for the H1 collaboration*”, Nucl. Instrum. Meth., **A518** 542–543, 2004.
- [59] N. Berger et al., First results from the first level of the h1 fast track trigger, In *IEEE Nuclear Science Symposium Conference Record*, volume 3, pages 1976–1979, 16–22 Oct. 2004.
- [60] A. Schöning, [H1], “*The Fast Track Trigger at the H1 experiment design concepts and algorithms*”, Nucl. Instrum. Meth., **A566** 130–132, 2006.
- [61] A. W. Jung et al., First results from the third level of the h1 fast track trigger, In *Proc. 15th IEEE-NPSS Real-Time Conference*, pages 1–7, April 29 2007–May 4 2007.
- [62] L. Caminada, “*Implementation of a Trigger for the Decay $b \rightarrow eX$ on the Third Trigger Level at the H1 Experiment*”, 2008, Diploma thesis.
- [63] H1 Collaboration, “*Proposal to Upgrade the LAr Calorimeter Trigger: The Jet Trigger*”, DESY-PRC-99-02.
- [64] M. Nedden, B. Reisert and T. Schöner, “*H1 Liquid Argon Trigger: Overview, Simulation and Performance*”, H1-IN-592(04/2001), 2001.

- [65] A. Geiser, “*Heavy flavour production, QCD, and the quark-photon coupling*”, Nucl. Phys. Proc. Suppl., **184** 189–195, 2008.
- [66] S. Chekanov et al., [ZEUS], “*Measurement of beauty production from dimuon events at HERA*”, JHEP, **02** 032, 2009.
- [67] T. Sjostrand, “*The PYTHIA and JETSET programs*”, Prepared for Particles & Fields 92: 7th Meeting of the Division of Particles Fields of the APS (DPF 92), Batavia, Illinois, 10-14 Nov 1992.
- [68] T. Sjostrand, L. Lonnblad and S. Mrenna, “*PYTHIA 6.2: Physics and manual*”, 2001.
- [69] H. Jung, “*The CCFM Monte Carlo generator CASCADE*”, Comput. Phys. Commun., **143** 100–111, 2002.
- [70] H. Jung and G. P. Salam, “*Hadronic final state predictions from CCFM: The hadron-level Monte Carlo generator CASCADE*”, Eur. Phys. J., **C19** 351–360, 2001.
- [71] CERN Program Library Long Writeup W5013, “*GEANT - Detector Description and Simulation Tool*”.
- [72] C. Wissing, *Entwicklung eines Simulationsprogramms und Implementierung schneller Spurfindalgorithmen fuer den neuen H1-Driftkammertrigger*, PhD thesis, Univ. Dortmund, 2003.
- [73] A. Jung, *Measurement of the $D^* \pm$ Meson Cross Section and Extraction of the Charm Contribution ($F_2^c(x, Q^2)$), to the Proton Structure in Deep Inelastic ep Scattering with the H1 Detector at HERA*, PhD thesis, Ruprecht-Karls-Univ. Heidelberg, 2009.
- [74] S. Caron, *Jets in Photoproduction at HERA*, PhD thesis, RWTH Aachen, 2002, PhD thesis RWTH Aachen, DESY-THESIS-2002-035.
- [75] F. D. Aaron et al., “*Inelastic Production of J/ψ mesons in Photoproduction and Deep Inelastic Scattering at HERA*”, to be submitted to Eur. Phys. J. C.
- [76] Ch. Grab, private communications.
- [77] L. M. Magro, *Multiple Parton Interactions in Photoproduction at HERA/H1*, PhD thesis, Univ. of Hamburg, DESY-THESIS-2009-007.
- [78] G. A. Schuler and T. Sjostrand, “*Towards a complete description of high-energy photoproduction*”, Nucl. Phys., **B407** 539–605, 1993.
- [79] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. Nadolsky and W. K. Tung, “*New Generation of Parton Distributions with Uncertainties from Global QCD Analysis*”, JHEP, **0207:012,2002**, JHEP0207:012,2002.
- [80] G. A. Schuler and T. Sjostrand, “*Parton Distributions of the Virtual Photon*”, Phys.Lett. B, **376** 193–200, 1996.
- [81] Hannes Jung, “*Un-integrated PDFs in CCFM*”, 2007.
- [82] R. Hälg, “*Electron Identification in Heavy Quark Decays in the H1 Experiment*”, 2007, Diploma thesis.
- [83] E. Hennekemper, *thesis in preparation*, PhD thesis.
- [84] P. Laycock, “*H1 TrkFinder*”, www-h1.desy.de/icas/oop/finders/.
- [85] V. Blobel, “*A new fast track-fit algorithm based on broken lines*”, Prepared for PHYS-TATO5: Statistical Problems in Particle Physics, Astrophysics and Cosmology, Oxford, England, United Kingdom, 12-15 Sep 2005.
- [86] A. Hoecker et al., “*TMVA - Toolkit for Multivariate Data Analysis*”, PoSACAT:, **040,2007**, PoSACAT:040,2007.
- [87] V. Chekelian, “*Direct FL measurement at high Q^2 at HERA*”, Proc. of XVI Int. Workshop on Deep-Inelastic Scattering and Related Topics, London, England, 2008.
- [88] F. Huber, “*Elastic J/ψ Production at low Q^2 at HERA*”, 2009, Diploma thesis.

-
- [89] S. Egli et al., “Calculating Event Weights in Case of Downscaling on Trigger Levels 1-4”, H1 Note-04/97-517.
- [90] V. Lendermann et al., “Combining Triggers in HEP Data Analysis”, Nucl. Instrum. Meth., **A604** 707–718, 2009.
- [91] A. Schöning, S. Schmitt and M. Sauter, unpublished.
- [92] K. Urban, *Measurement of Inclusive and DiJet D^* Meson Photoproduction at the H1 Experiment at HERA*, PhD thesis, Ruprecht-Karls-Univ., 2009.
- [93] D. Sunar, *Measurement of $K^{*\pm}(892)$ Production in Deep Inelastic ep Scattering with the H1 Detector at HERA*, PhD thesis, 2009, University Antwerpen.
- [94] M. Steder, *Measurement of Inelastic Charmonium Production at HERA*, PhD thesis, Univ. Hamburg, 2008.
- [95] M. Krämer, *Measurement of Charm and Beauty Cross Sections in Photoproduction using Events with Muons and Dijets at HERA*, PhD thesis, Univ. Hamburg, 2009.
- [96] T. Zimmermann, *Measurement of Charm in Charged Current at HERA*, PhD thesis, ETH Zurich, 2008.
- [97] C. Amsler et al. (Particle Data Group), Physics Letters, **B667** 1, 2008.
- [98] F. D. Aaron et al., [H1], “Study of Charm Fragmentation into $D^{*\pm}$ Mesons in Deep-Inelastic Scattering at HERA”, Eur. Phys. J., **C59** 589–606, 2009.
- [99] G. Cowan, *Statistical Data Analysis*. Oxford University Press, 1998.
- [100] B. List, “Decorrelation Of Correlated Data Points”, unpublished, 2008.
- [101] G. D. Lafferty and T. R. Wyatt, “Where to stick your data points: The treatment of measurements within wide bins”, Nucl. Instrum. Meth., **A355** 541–547, 1995.
- [102] C. Grab, “PQCD calculations based on FMNR”.
- [103] H. L. Lai et al., “Global QCD Analysis of Parton Structure of the Nucleon: CTEQ5 Parton Distributions”, Eur.Phys.J.C, **12:375-392,2000**, Eur.Phys.J.C12:375-392,2000.
- [104] M. Glück, E. Reya and A. Vogt, “Dynamical Parton Distributions Revisited”, Eur.Phys.J. C, **5** 461–470, 1998.
- [105] A. V. Lipatov and N. P. Zotov, “Beauty photoproduction at HERA: kt -factorization versus experimental data”, Phys.Rev.D, **73:114018,2006**, Phys.Rev.D73:114018,2006.
- [106] A. Geiser, “Private communications”.
- [107] C. Wissing et al., Performance of the h1 fast track trigger - operation and commissioning results, In *Proc. 14th IEEE-NPSS Real Time Conference*, page 4pp., 10–10 June 2005.
- [108] M.-O. Bönig, *Messung des D^* -Meson-Produktionsquerschnitts in tiefinelastischer Streuung mit dem H1-Experiment*, PhD thesis, Univ. Dortmund, 2007.
- [109] A. Baird et al., “A Fast High Resolution Track Trigger for the H1 Experiment”, IEEE Trans.Nucl.Sci., **48** 1276–1285, 2001.
- [110] D. Benekenstein, “Effizienzstudie und Optimierung der dritten Stufe des schnellen Spurtriggers bei H1”, 2007, Diploma thesis.
- [111] V. Blobel, “Unfolding for HEP experiments”, DESY Computing Seminar, 23rd June 2008.
- [112] V. Blobel, “An unfolding method for high energy physics experiments”, Proc. Advanced 510 Statistical Techniques in Particle Physics, 2002.
- [113] T. Carli et al., “Performance of the H1 Lar Trigger in 1994”, H1 internal report H1-07/95-445, 1995.

BIBLIOGRAPHY

- [114] G. Cowan, “*Note on Blobel’s Unfolding Method*”, <http://www.desy.de/~blobel/unfold.html>.
- [115] P. C. Hansen, “*The L-Curve and its Use in the Numerical Treatment of Inverse Problems*”, pages 119–142, 2000.
- [116] A. Jung, Private communications.
- [117] E. Longo and I. Sestili, “*Monte Carlo Calculation of Photon Initiated Electromagnetic Showers in Lead Glass*”, Nucl. Instrum. Meth., **128** 283, 1975.
- [118] S. Schmitt, TUnfold (Unfolding class for RooT).
- [119] C. Schmitz, *Isolated Photon Production in Deep-Inelastic Scattering at HERA*, PhD thesis, Univ. Zurich, 2007.

Acknowledgements

It is a great pleasure to express my deepest gratitude to all those who assisted, encouraged and supported me during the last four years to bring this thesis to a happy end. I am obliged to my teachers and colleagues, who all shared their knowledge, insight and enthusiasm in innumerable fruitful discussion with me, and among whom many became my friends. In a more personal way, a special thank goes to my family for their interest, understanding and love. *Thank you all!*

Curriculum vitae

MICHEL DAVID SAUTER

Born June 27 1980
in Bern, Switzerland
Citizen of Sulgen (TG),
Switzerland

EDUCATION

2005 – 2009	Doctoral studies at ETH Zürich, research at the H1 experiment at DESY, Hamburg
2005	Diploma in Physics at ETH Zürich, diploma thesis <i>“Investigation of the background in an experiment to search for the invisible decay of orthopositronium”</i> in the group of Prof. A. Rubbia, at CERN Geneva
2000 – 2005	Studies of physics at ETH Zürich
1999	Eidgenössische Matura Typus C
1996 - 1999	Gymnasium in Oberwil
1987 - 1996	Primary and Secondary School in Bottmingen and Binningen