Bachelor Thesis

Identification of isolated photons at high energies at H1

Author: Arno Gadola, agadola@physik.uzh.ch

Supervisor: Dr. Katharina Müller, kmueller@physik.uzh.ch

Fall 2007

Contents

1 Abstract

The identification of isolated photons in electron proton collisions at HERA is hindered by the background of photons coming from the decay of neutral mesons as π^0, η etc. The existing methods to distinguish isolated and photons from decay work well for transverse energies in the range of 5 to 10 GeV. They use parameters like shape and dimension of the electromagnetic shower. In this bachelor thesis old and new separation parameters are analyzed on their separation power for transverse energies up to 15 GeV.

The first step includes almost all known shower shape variables. A rough behavioral analysis of the variables over the whole transverse energy range of 5 to 15 GeV leads to a first exclusion of very low separation power variables. The remaining variables are introduced to an analysis tool, which trains preselected methods, like Likelihood, with simulated data sets. The simulated data sets are split in sets of single photon and double photon (from π^0 decays) events which correspond to the signal of isolated and background photons. With the so trained methods, real data are analyzed with a second analysis tool. Also the single and double photon data sets are analyzed with the second tool to become an idea of the quality of the trained methods.

Due to lack of statistics in the real data set, only the simulated data give a feedback of the separation and method training quality for the higher transverse energy region. Nevertheless, satisfying separation power is achieved in most of the kinematic region without spending a lot of time on fine tuning on the training of the methods.

The results show, that the transverse radius of the shower has the most significant separation power. Likewise, a good separation power shows up for the variables of the first layer of the electromagnetic calorimeter as i.e. the cluster energy of only the first layer or the energy of the hottest cell in the first layer.

2 Theory

2.1 Physics processes at HERA

At HERA electrons or positrons are collided with protons. The pointlike electron or positron scatters off the proton via the exchange of a spin one gauge boson of the electroweak interaction. In neutral current (NC) processes a photon (γ) or and a Z^0 boson[1] is exchanged, whereas in charged current (CC) the exchanged boson is a $W^\pm$. In these scattering processes the proton usually dissolves. Figure 2.1 shows these processes.
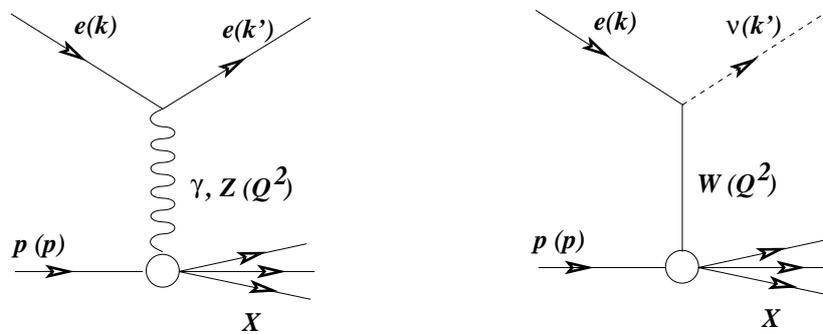
Figure 2.1: Diagrams of electron and proton scattering via photon and Z^0 exchange for neutral current interactions and $W^\pm$ exchange for charged current interactions. The four-momentum vectors of the particles are given in parentheses ([8], p.3).

In neutral current processes ($e^\pm p \overset{\gamma, Z^0}{\to} e^\pm X$) the final state yields an electron and eventually a broken-up proton (X) whereas in charged current processes ($e^\pm p \overset{W^\pm}{\to} \nu^\pm X$) the final state yields a neutrino (e$^-$, W$^-$) or antineutrino (e$^+$, W$^+$) and the eventually broken-up proton (X). For this thesis only neutral current interaction and photon exchange are of interest.

2.2 DIS (Deep Inelastic Scattering) and photo-production

At HERA the scattering processes are divided into two classes depending on the virtuality Q^2 of the exchanged photon: deep inelastic scattering ($Q^2 > 0$) and photo-production ($Q^2 \approx 0$). Table 2.1 shows the important differences.

---

[1]The Z^0 is only exchanged at higher photon virtuality Q^2 values. See i.e. [9].

	DIS	Photo-production
Q^2	> 0	≈ 0
θ	$< 2\pi$	$\approx 2\pi$
Effective coupling	α_s	$\alpha = 1/137$
Scattered e^-	Detectable	Emerges through beam pipe. About 30% are detected in a special detector close to the beam pipe.
Exchanged γ	No fluctuation	Fluctuation to quarks, gluons or to vector mesons possible*.
γ production	Prompt photons and bremsstrahlung	Basically prompt photons

*The concept of hadronic structure of photons was introduced by the Vector-Dominance-Model (VDM).

Table 2.1: DIS process vs. photo-production. $Q^2 = -q^2$ denotes the negative four-momentum transfer squared carried by the photon and θ is the scattering angle of the electron with respect to the direction of the incoming proton.

There are two subclasses in photo-production. In the direct process the exchanged photon couples directly to the quarks of the proton (fig. 2.2). In the resolved process the photon fluctuates into a hadronic state and a parton of the photon enters into the interaction (fig. 2.3). In this case there is a photon remnant (similar to the remnant of the proton) which is not participating in the interaction. The outgoing particles of the hard scattering process are normally quarks and gluons. In the strong color field they undergo a process of fragmentation and hadronize to so called jets. The jets and the proton remnant form the broken-up proton X seen in 2.1. Information about the parton of the hard subprocess is therefore only indirectly accessible via the reconstruction of the jet kinematics. In a process which is suppressed[2] by 1/137, photons are emitted instead of gluons. These so called prompt photons that are investigated in this thesis give a direct information on the internal scattering process and the structure of the proton or the photon.
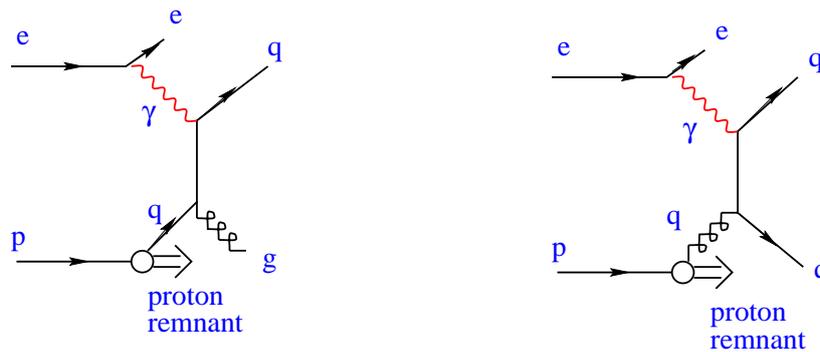
Figure 2.2: Feynman graphs of direct di-jet photo-production events ([8], p.10).

---

[2]Processes without prompt photons go with $\alpha_s = 1$, the coupling constant of the strong interaction, whereas processes with prompt photons only go with $\alpha = 1/137$, the coupling constant of the electroweak interaction and are therefore suppressed by a factor of ≈ 100.
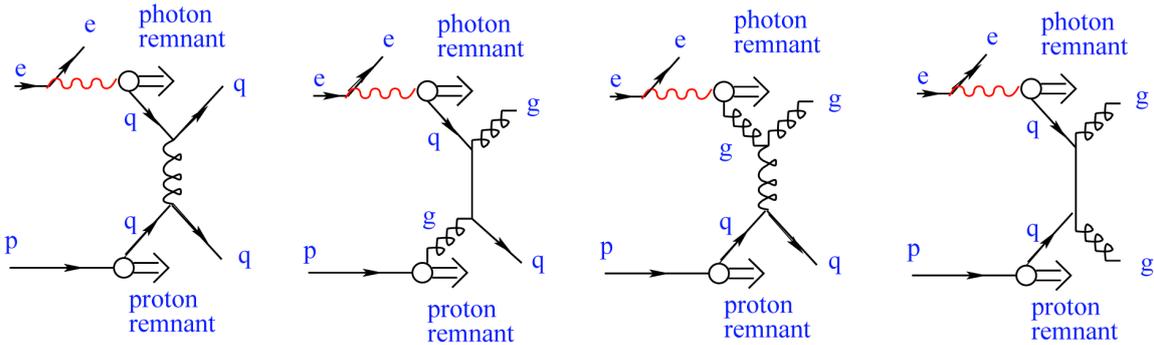
Figure 2.3: Feynman graphs of resolved di-jet photo-production events ([8], p.10).

2.3 Prompt photons

There are two processes of prompt photon production. The direct prompt photon process (fig. 2.4) which gives insight into the parton structure of the proton and the resolved prompt photon process (fig. 2.5) that gives also an insight to the parton content of the exchanged photon. A measurement of these rare processes is hence directly sensitive to the gluon and quark content of the exchanged photon and the parton structure of the proton.

The main experimental difficulty in the prompt photon measurement is the large background from photons of the decay of neutral mesons, especially from π^0 and η. These mesons originate from the fragmentation of the quarks and gluons at high energies. The opening angle ϑ of the decay photons of a meson is directly related to the meson energy ($E \propto \vartheta$, the angle between the two decay products) so the decay photons have a very small opening angle at high energies and are reconstructed in only a single electromagnetic cluster (see 5.1).
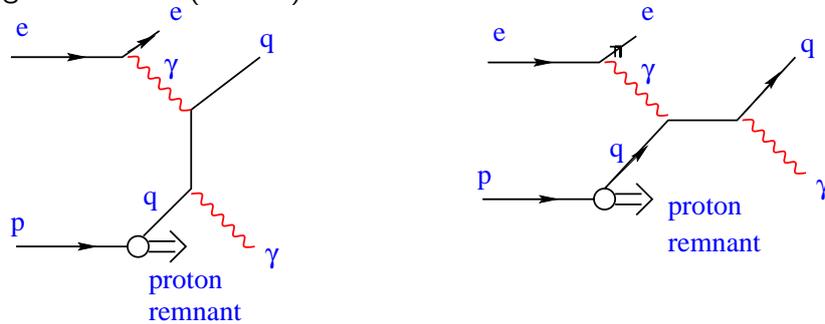
Figure 2.4: Example of direct prompt photon processes at leading order. The photon couples directly to a parton from the proton ([8], p.12).
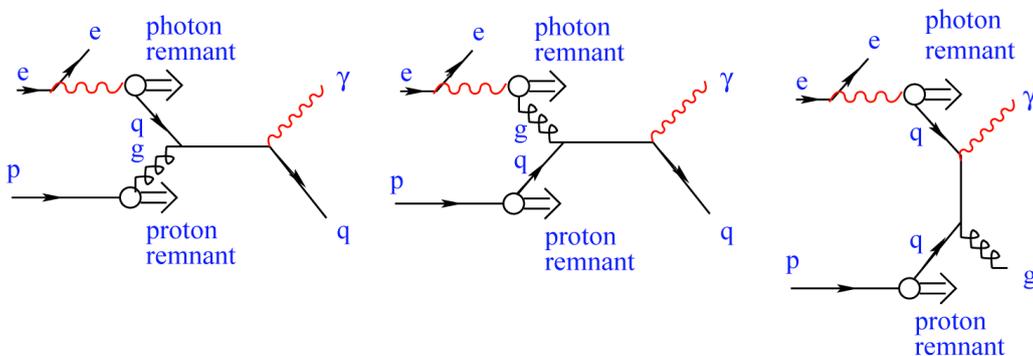
Figure 2.5: Examples of resolved prompt photon processes at leading order. The photon is resolved and a parton from the photon couples to a parton from the proton ([8], p.12).

3 HERA and the H1 detector
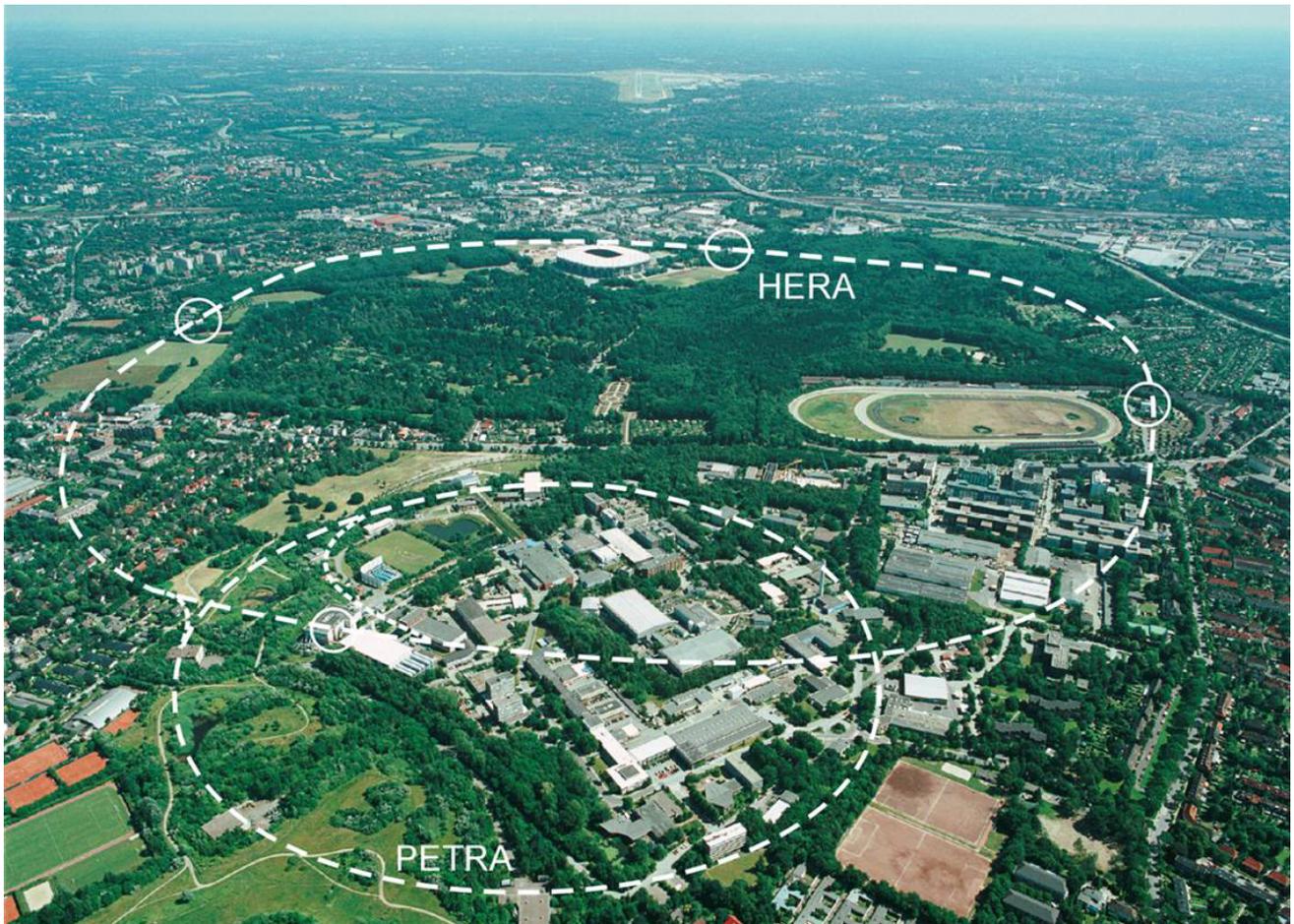
Figure 3.1: A bird's eye view of the DESY site and the surroundings [1].

The two accelerators HERA and PETRA at DESY (Deutsches Elektronen-Synchrotron) are shown in figure 3.1 as dashed lines. HERA, with its circumference of 6.3 km is the biggest accelerator at DESY and it is housed in a tunnel with an inner diameter of 5.2 m which is situated about 10-20 m underground. In HERA, protons with a maximum energy of 920 GeV collide with electrons or positrons with an energy of 27.5 GeV thereby providing a way to study the inner structure of protons.

The accelerator PETRA, with its circumference of 2.3 km, serves as a pre-accelerator of protons, electrons and positrons for HERA. Moreover, PETRA provides synchrotron radiation (X-rays) for research at the HASYLAB which is situated on the DESY site. The DESY site lies essentially within the PETRA ring.

One of the four detectors at HERA is the H1 detector. The H1 detector was designed to provide clear identification and precise measurement of electrons, muons and penetrating neutral particles together with a good performance in the measurement of jets with high particle densities. The characteristics of the particles detected with the H1 detector after an ep-collision provides access to the following areas of current research:
- Structure functions of the proton, and of the diffractive exchange.
- Parton distributions of real and virtual photons.
- Tests of QCD in jet, photon and heavy-quark production.
- Tests of the electro-weak theory.
- Study of non-perturbative phenomena of strong interactions.
- Diffractive scattering and the mechanism of colour neutralization.
- Searches for new particles and new interactions.

This thesis could give a contribution especially to the first two topics.

3.1 Setup of the H1 detector

In the following, the detector is briefly discussed. A more detailed description of the H1 detector can be found in [11].

The origin of the H1 coordinate system (fig. 3.2) is the nominal *ep* interaction point (vertex), with the direction of the proton beam defining the positive z-axis (forward direction). Transverse momenta are measured in the x - y plane. Polar (θ) and azimuthal (ϕ) angles are measured with respect to this reference system. The pseudorapidity is defined to be $\eta = -ln(tan(\theta/2))$.
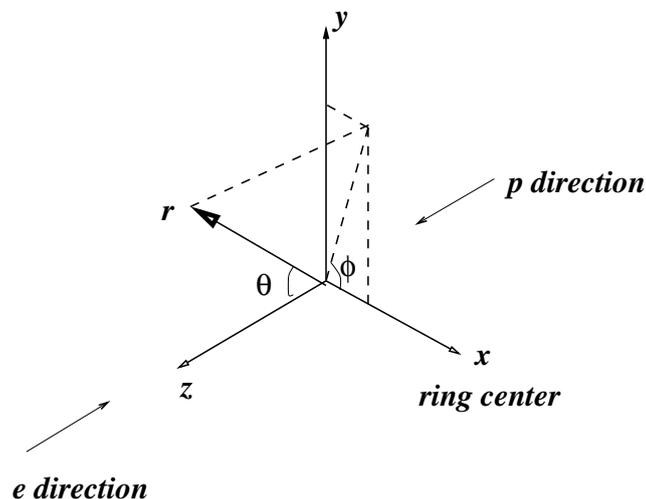
Figure 3.2: The definition of the Cartesian coordinate system (x, y, z) with respect to the *e* and *p* beam directions. The corresponding spherical coordinate system (r,θ,ϕ) is also shown ([8], p.27).

In the central region ($25°< \theta <165°$) the interaction region is surrounded by the central tracking system, which consists of a silicon vertex detector, drift chambers and multi-wire proportional chambers, all located within a solenoidal magnetic field of 1.16 T. The forward tracking detector ($7°< \theta <25°$)

and the backward drift chamber (BDC, operated in 1999-2000) or backward proportional chambers (BPC, for 2003-2007) measure tracks of charged particles at smaller and larger θ than the central tracker, respectively. In each event the interaction vertex is reconstructed from the measured charged tracks. The polar angle of the reconstructed particles is measured with respect to this vertex.

The liquid argon (LAr) sampling calorimeter which is also located within the magnet, surrounds the tracking chambers. It covers the angle $4° < \theta < 154°$ and will be described in detail in the next chapter.

In the backward region $153° < \theta < 178°$, the particle energies are measured by a lead-scintillating fibre spaghetti calorimeter (SpaCal).

The luminosity is determined from the rate of the Bethe-Heitler process $ep \rightarrow ep\gamma$, measured using a photon detector located close to the beam pipe at $z = $ -103 m and a small angle calorimeter at $z = $ -33.9 m.
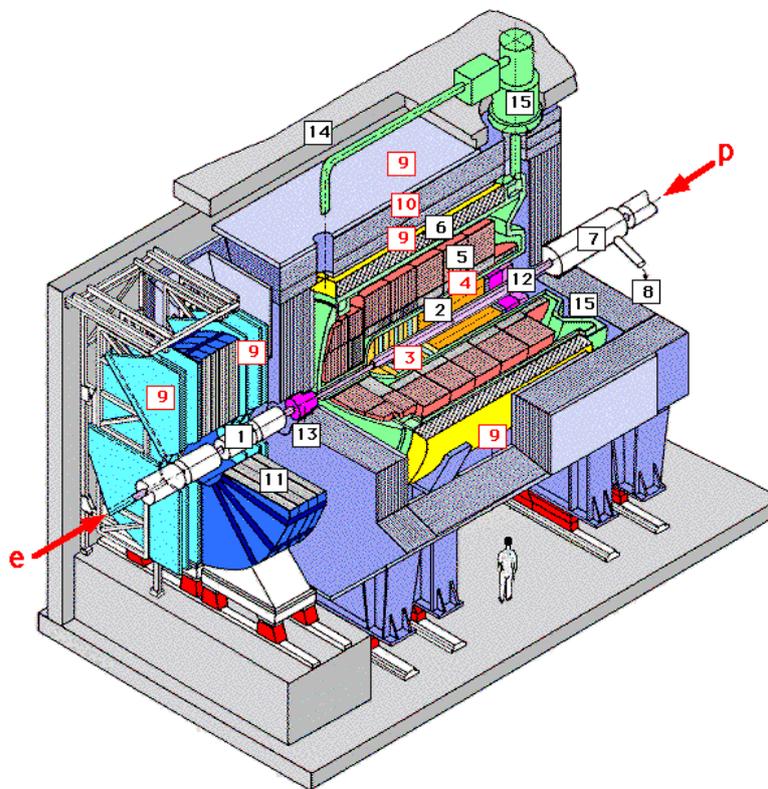
Figure 3.3: An isometric view of the H1 detector with its major detector components [2].

1. Beam pipe and beam magnets
2. Central tracking chambers
3. Forward tracking and Transition radiators
4. Electromagnetic Calorimeter (lead, liquidargon)
5. Hadronic Calorimeter (stainless steel, liquidargon)
6. Superconducting coil (1.2T)
7. Compensating magnet
8. Helium cryogenics
9. Muon cambers
10. Return yoke (instrumented iron with streamer tubes)
11. Myon-Toroid-Magnet
12. Warm calorimeter (Spacal)
13. Plug calorimeter
14. Concrete screen
15. Liquid Argon cryostat.

For this thesis only events with a photon detected in the electromagnetic, liquid argon (LAr) calorimeter (see pic. 3.3 No. 4) of the detector are of importance.

3.2 The LAr calorimeter

The liquid argon (LAr) sampling calorimeter surrounds the tracking chambers. It has an asymmetric polar angle coverage of $4° < \theta < 154°$ and full azimuthal acceptance. It consists of an inner electromagnetic section with lead absorbers and an outer hadronic section with steel absorbers (see fig. 3.4a). The calorimeter is divided into eight wheels along the beam axis, each of them segmented in ϕ into eight modules (fig. 3.4b), separated by small regions of inactive material. The electromagnetic and the hadronic sections are highly segmented in the transverse and the longitudinal direction with about 44'000 cells in total. The granularity is larger in the electromagnetic part and increasing in both sections in the forward part. For particles coming from the *ep* interaction region, the laterally projected cell size in the electromagnetic part varies between 5×5 cm^2 in the forward and at most 7×13 cm^2 in the central region. The longitudinal segmentation in the different wheels varies from three (central) to four (forward) layers in the electromagnetic and from four to six in the hadronic section. The first electromagnetic layer has a thickness of about 3 to 6 radiation lengths for particles coming from the interaction region.
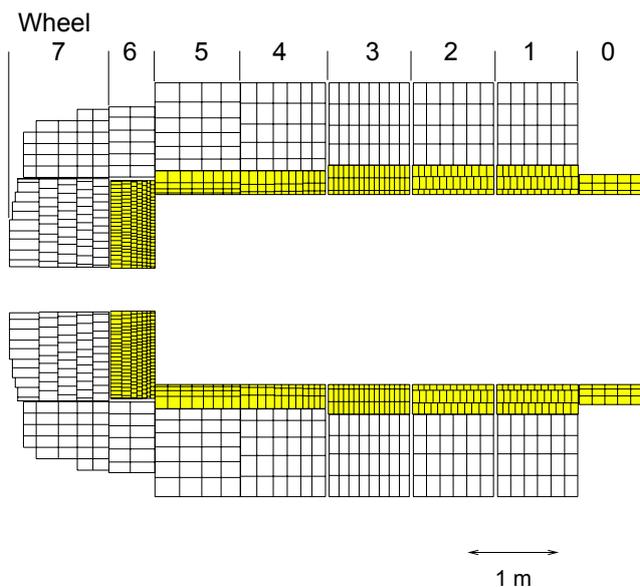
(a) Wheel and cell structure of the LAr e.m. (yellow) and hadronic calorimeter ([8], p.31).
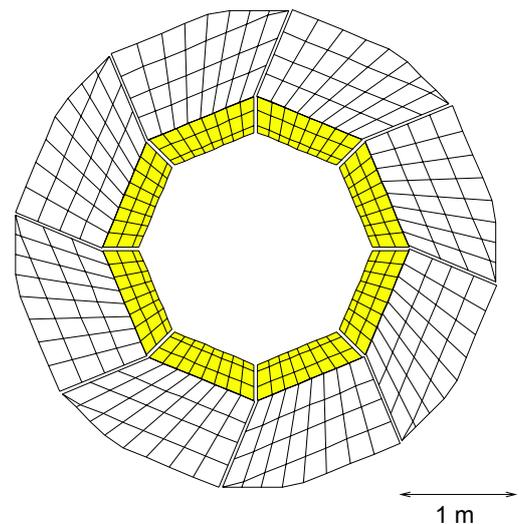
(b) Octant and cell structure of the CB1 wheel of the LAr e.m (yellow) and hadronic calorimeter ([8], p.31).

Figure 3.4: Structure of the LAr electromagnetic (yellow) and hadronic calorimeter. (a) Longitudinal sectional drawing, (b) profile of a wheel.

The main reason for choosing a liquid argon technique for the detector are good stability and ease of calibration, fine granularity for e$/\pi$ separation and energy flow measurements as well as homogeneity of response. Sophisticated electronics converts the deposited energy of an event in a detector cell into an electric signal for further treatment. The demand on the electronics is enormous because large energies may be deposited at short time intervals (96ns at Hera) into the detectors with large

capacities and long collection time, and the information has to be stored until the arrival of the trigger signal ($\approx 2.5\mu s$). But also long time stability behavior is a big issue due to minimization of statistical errors.

After collecting all data from the detector, a computer farm performs different processes on the data as for example noise reduction, event reconstruction, corrections of dead material (derived by Monte Carlo simulations) and clustering before saving and archiving the data. All saved data are available to users.

3.3 Clustering

The clustering of neighboring cells is done in the event reconstruction. All cells passing the cell level reconstruction are subject to clustering. The algorithms used are tuned such that the cells containing energy depositions from an electromagnetic shower (more about e.m. showers see [6], p. 42) initiated by a photon or electron are most probably merged into one cluster (fig. 3.5). Hadronic showers with their larger spatial fluctuations are in general split into several clusters.
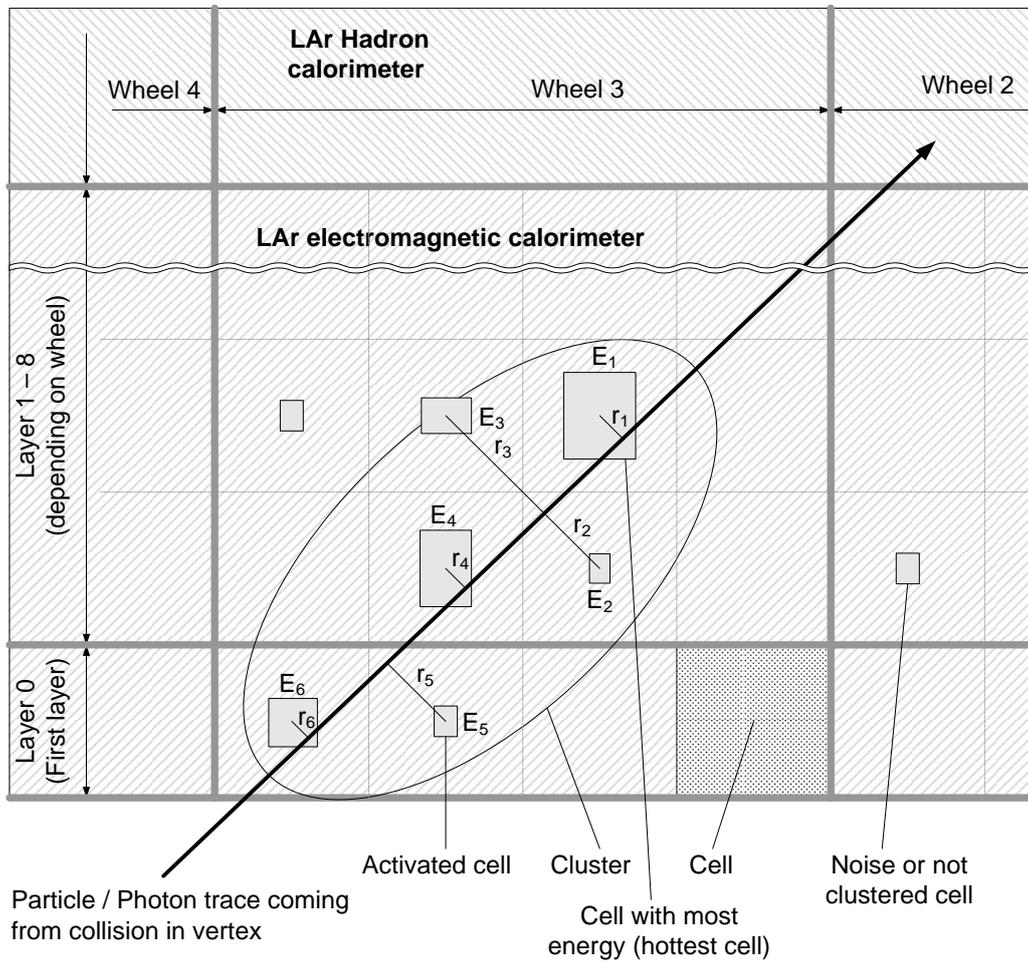
Figure 3.5: Schematic drawing of the calorimeter response to a photon event. The figure shows a two dimensional view ($r - z$ plane) of the detector. The deposited energy magnitude E_i in a cell is shown by the area of rectangles. One can also see, that a cluster does not need to contain all activated cells. Some cells are too far away from the hottest cell and or have not enough energy deposed.

4 Monte Carlo simulations

4.1 Data generation

To generate shower data for the analysis, events needed to be simulated and reconstructed with the H1 Monte Carlo (MC) simulation software H1SIM (Version: H1SIMREC90722) and the new version of the integrated module H1FAST (see [2] for further details) by Katharina Müller. To simplify the data analysis and to avoid the complex and problematic selection of events and photon clusters, only events with exactly one generated particle (so called single particle events) are taken for the simulation. The simulation of single particle events is easy and fast and allows the production of events in a large number. For signal events, single photons with a flat distribution in polar angle θ, azimuthal angle ϕ and transverse energy E_T for all E_T and all wheels are used. The background consists only of pions, who provide the biggest part of the real background and suite therefore perfectly for this analysis. The pions have the same flat distribution as the single photons. For an overall analysis also multi-photon events from hadronic decays of other particles as i.e. η, ρ, K and (anti)neutrons should be taken in account (see fig. 4.1). The simulated data with their corresponding E_T and wheel numbers are saved in files, which are after some compilation processed by the $clustertool.C$ software (see section 7.1). In the $clustertool.C$, event data can be analyzed or separated identified by applying cuts on E_T and wheel number for further analysis.
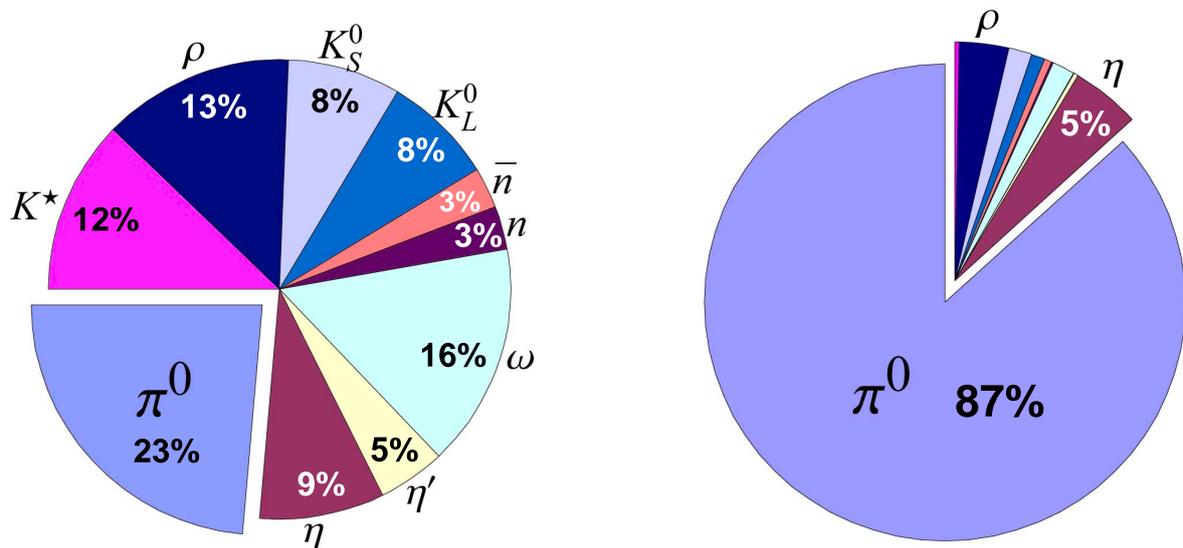
Figure 4.1: Relative contribution of the relevant neutral hadrons to the background of photon-like clusters before and after the standard cluster shape selection ([6], p.35).

4.2 Data variables

The following table 4.1 lists the most important and later on used variables from the MC data.

$E_{Cluster}$	Cluster energy	$E_{HottestCell}$	Hottest cell energy
$E_{ClusterL1}$	Cluster energy layer 1	$E_{HottestCellL1}$	Hottest cell energy layer 1
E_T	Transverse energy	$E_{HotCore}$	Hot core energy
E_{TL1}	Transverse energy layer 1	$E_{HotCoreL1}$	Hot core energy layer 1
θ	Polar angle	$Wheelno$	Wheel number

Table 4.1: MC data variables.

4.2.1 Cluster energy E$_{\text{Cluster}}$

The cluster energy is provided by the H1 data class H1Cluster and is a summation of all activated cell energies in the defined cluster.

4.2.2 Cluster energy in first e.m. LAr layer E$_{\text{ClusterL1}}$

The cluster energy in the first LAr layer is a summation of all activated cell energies of the first electromagnetic LAr layer in the defined cluster. This energy should be slightly higher for multi-photon events due to a higher conversion probability of the photons.

4.2.3 Transverse energy E$_{\text{T}}$

The transverse cluster energy is the projection of the cluster energy perpendicular to the beam axis and defined with the polar angle θ as

$$E_T = E_{Cl}sin\theta \tag{4.1}$$

4.2.4 Hottest cell energy E$_{\text{HottestCell}}$

The energy of the hottest cell in the cluster is provided directly by the H1 data class H1Cluster. The hottest cell contains the most deposited energy. Single photon events yield larger values than multi-photon events.

For comparison of different events, the energy fraction of the hottest cell and the cluster is taken as variable (see chapter 6.1).

4.2.5 Hottest cell energy in first layer E$_{\text{HottestCellL1}}$

The energy of the hottest cell in the first layer in the cluster is determined through the analysis of all activated cell energies in the first e.m. LAr layer. Since the conversion of multiple photons is more probable than for single photons, the hottest cell in the first layer ($E_{HotCellL1}$) or better its fraction with the cluster energy is a good separation variable.

For comparison of different events, the energy fraction of the hottest cell in first e.m. layer and the cluster is taken as variable (see chapter 6.1).

4.2.6 Hot core energy $E_{HotCore}$

This variable is provided directly by the H1 data class H1Cluster. Starting from the most energetic cell (hottest cell) the energy of the adjacent activated cells are added (4 cells for wheel 1 to 3, 8 cells for wheel 4 and 5 and 12 cells for wheel 6 and 7) such that the energy in the hot core is maximized. Single photon events yield larger values than multi-photon events.
For comparison of different events, the energy fraction of the hot core and the cluster is taken as variable (see chapter 6.1).

4.2.7 Hot core energy in first layer $E_{HotCoreL1}$

This variable is provided directly by the H1 data class H1Cluster. Starting from the most energetic cell (hottest cell) in the first e.m. layer adjacent cells are added (4 cells for wheel 1 to 3, 8 cells for wheel 4 and 5 and 12 cells for wheel 6 and 7) such that the energy in the hot core is maximized.
For comparison of different events, the energy fraction of the hot core in first e.m. layer and the cluster is taken as variable (see chapter 6.1).

5 Photon signal

5.1 Single versus multiple photons

As mentioned in chapter 2.3 single photons have to be identified and distinguished from a background from decay of hadrons to multiple photons. The signal and background samples display slightly different shapes in the detector (fig. 5.1) depending on their energy and on the wheel, they are detected. These shapes can be described by so called shape variables (see chapter 6.1). The separation of these two different event samples at higher energies (10 to 15 GeV) with a set of shape variables is the subject of this thesis.
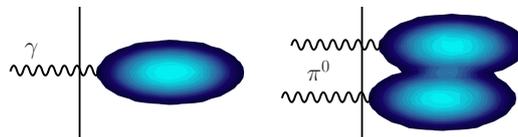
Figure 5.1: Schematic representation of the different shower profiles of clusters induced by single photons (left) and multiple photons (right). The black line indicates the calorimeter surface ([6], p.80).

5.2 Example of a detection

The following figure 5.2 shows a typical event with an isolated photon in the H1 detector. The total collision energy is $\sqrt{s} = 318$ GeV. At the left detector side ① a part of the proton remnant is detected. The biggest part of the remnant escapes through the beam pipe. On the right side the scattered electron ② is detected in the SpaCal. ③ shows the reconstructed track of the jet of the parton which takes part in the scattering. The photon leaves no track in the tracking chambers but has a narrow energy deposit in the LAr calorimeter ④. No track is allowed to point to the cluster ⑤ in order to reject charged particles. Furthermore the photon is required to be isolated which significantly reduces the background from neutral hadrons and their decay products. Usually the isolation criteria is based on the energy around the photon candidate (for more details see [6] and [13]). ⑥ shows the vertex (collision point) reconstructed from the tracks.
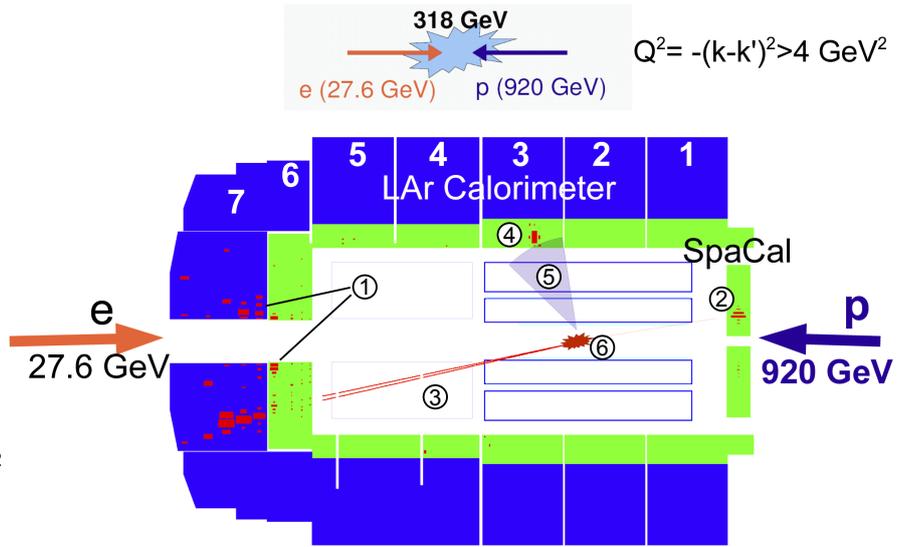
Figure 5.2: Sectional drawing of the H1 detector [10].

6 Methods evaluation for event separation

As shown schematically in figure 5.1, single photons and multiple photons have different cluster shapes. The multi-photon cluster is typically less compact, transversely wider and more asymmetric. The shower is likely to start closer to the calorimeter surface, as the probability of conversion increases with the number of incident photons. But with increasing energy the multi-photon event gets similar to the single photon event due to a decreasing opening angle so the variables describing the cluster shapes look pretty alike.

To become an idea of how the different shower variables look for a single photon or pion (double photons, respectively) event, the most popular variables are plotted and compared. Table 6.1 lists all these variables. For a more detailed understanding of the variables they are plotted separately

R_T	Transverse cluster radius
R_L	Longitudinal cluster dimension
$Kurtosis\ K$	Weighted energy distribution
$D_{HottestCells}$	Distance between the two hottest cells
$E_{HottestFrac}$	Fraction of $E_{HottestCell}$ and $E_{Cluster}$
$E_{HottestL1Frac}$	Fraction of $E_{HottestCellL1}$ and $E_{Cluster}$
$E_{HotCoreFrac}$	Fraction of $E_{HotCore}$ and $E_{Cluster}$
$E_{HotCoreL1Frac}$	Fraction of $E_{HotCoreL1}$ and $E_{Cluster}$

Table 6.1: Examined shape variables before optimizing them to the significant ones.

for each wheel and in transverse energy steps of 2 GeV. Figures 7.1 to 7.6 show the variables for all wheels at a transverse energy of 13-15 GeV. The variables θ and $Wheelno$ (see table 4.1) provided a verification of the cuts on the data. They are associated as following:

Wheel	Polar angle θ
1	140° - 150°
2	120° - 140°
3	80° - 120°
4	45° - 80°
5	27° - 47°
6	17° - 30°

Table 6.2: Covert polar angle θ of each detector wheel. Note that there are no exact boundaries since θ is smeared.

Out of these first plots the most significant and promising variables were picked out for further treatment. The next chapter describes in more detail the most important shower variables. A detailed description of the individual energy variables can be found in subsections 4.2.1 to 4.2.7.

6.1 Shower variables

The description of the different shower (cluster) variables refer to figure 3.5. Some variables are calculated with so called central transverse moments. The moments are defined as

$$\mu_{T,k} = \left\langle |\vec{r_T} - \langle \vec{r_T} \rangle|^k \right\rangle \tag{6.1}$$

where $\vec{r_T}$ is the transverse projection of the cell vector to the particles trajectory and the energy weighted average of the transverse cell positions is

$$\langle \vec{r_T} \rangle = \frac{\sum_i E_i \vec{r_{T,i}}}{\sum_i E_i}, \quad i: \; all \; cells \; of \; the \; cluster. \tag{6.2}$$

The first and until now most promising variable is the transverse radius of the cluster (see also [6]).

6.1.1 Transverse radius R$_T$

The transverse radius is defined as the square root of the second central transverse moment

$$R_T = \sqrt{\mu_{T,2}} \tag{6.3}$$

and describes in principal the transverse dimension of the cluster with respect to the particle trajectory. Tight showers induced by single photons have small values of R_T. Due to the limited transverse radius for single as well as for multi-photon events the radius is cut at 6 cm, means that all events with a larger R_T are ignored.

6.1.2 Longitudinal dimension R$_L$

The longitudinal dimension of the shower is defined similar to the transverse radius.

$$R_L = \sqrt{\mu_{L,2}} \tag{6.4}$$

The longitudinal radii of the activated cells r_L are calculated with respect to the center of gravity (COG) of the cluster. The longitudinal dimension is, as figure 5.1 shows, slightly larger for a multi-photon event.
This variable was not yet used in shower shape analysis but makes a good figure in the first plotting so it could as well be a good candidate for shower separation.

6.1.3 Kurtosis K

The kurtosis defines how strongly the energy distribution of the cluster is peaked. For a single photon a gaussian distribution with $K = 0$ is expected. For a double photon event (i.e. $\pi^0 \rightarrow \gamma\gamma$) two superposed gaussian with $K > 0$ are expected (fig. 6.1). The transverse kurtosis is defined as

$$K_T = \frac{\mu_{T,4}}{(\mu_{T,2})^2} - 3 \tag{6.5}$$

where the -3 is just an offset on the x-axis. This variable has however no separation power at higher energies and is therefore skipped for the analysis.
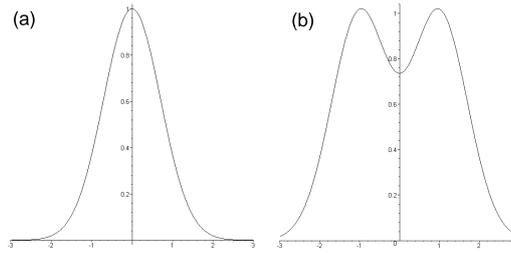
Figure 6.1: Kurtosis for: (a) single photon event with $K = 0$ and for (b) double photon event with $K > 0$.

6.1.4 Hottest cell energy fraction $E_{HottestFrac}$

The hottest cell fraction is defined as:

$$E_{HottestFrac} = \frac{E_{HottestCell}}{E_{Cluster}} \qquad (6.6)$$

It is sensitive to the compactness of the cluster.

6.1.5 Hottest cell energy fraction in first e.m. layer $E_{HottestL1Frac}$

Same as $E_{HottestFrac}$ but instat of $E_{HottestCell}$ the energy of the hottest cell in the first electromagnetic LAr layer is taken.

6.1.6 Hot core energy fraction $E_{HotCoreFrac}$

The hot core energy fraction is defined as:

$$E_{HotCoreFrac} = \frac{E_{HotCore}}{E_{Cluster}} \qquad (6.7)$$

The fraction is sensitive to the compactness of the cluster. This variable is not that significant and is therefore skipped for further analysis.

6.1.7 Hot core energy fraction in first e.m. layer $E_{HotCoreL1Frac}$

Same as $E_{HotCoreFrac}$ but instat of $E_{HotCore}$ the energy of the hot core in the first electromagnetic LAr layer is taken.
This variable is also not that significant and is therefore skipped for further analysis.

6.2 Variables selection

The following table 6.3 gives an overview of the shower shape variables which will be used further and the expected (or predicted by the first plots, see fig. 7.1 to 7.6) attitudes at higher energies as well as the expected separation quality.
As shown later in chapter 7, the separation qualities of the selected variables at higher energies differ slightly from the expectations or do not even separate.

Variable		$\gamma : \gamma\gamma$ **event**	**Separation quality**
R_T	Transverse cluster radius	Smaller	$+++$
R_L	Longitudinal cluster dimension	Slightly smaller	$++$
$E_{HottestFrac}$	Fraction of $E_{HottestCell}$ and $E_{Cluster}$	Smaller	$++$
$E_{HottestL1Frac}$	Fraction of $E_{HottestCellL1}$ and $E_{Cluster}$	Much smaller	$+++$

Table 6.3: Selected shower shape variables and their predicted properties.

6.3 Multivariate data analysis tool TMVA

After choosing the best or promising shower shape variables the multivariate data analysis tool TMVA [5] of ROOT[1] imposed to be used for further analysis of simulated single particle (γ and $\pi^0 \to \gamma\gamma$ events) and real detector data events (kindly provided by Krzysztof Nowak).

6.3.1 TMVA tools

The TMVA tool provides a very easy and fast way for data analysis with several methods like Fisher, Likelihood, Neural Network Analysis etc. The analysis happens in two steps with the C^{++} scripts $TMVAnalysis.C$ and $TMVApplication.C$. Both scripts with explanations and examples can be found on the web [5].

TMVAnalysis tool

The TMVAnalysis tool trains the selected methods (described in A.1) with separated signal (single photon events) and background (double photon events) data. The tool produces for each method and for each provided shower shape variable a weighting of signal and background. These weightings will later be used on real data by the TMVApplication tool to build a discriminator and hence separate single or isolated photon events from background.

To compare the signal-efficiency and background-rejection performance of the different methods, the TMVAnalysis tool provides some tabulated benchmark values as well as smooth efficiency versus background rejection curves (see fig. 7.7).

After playing a bit with the tool and getting experience, the methods Likelihood, Likelihood PCA, Fisher, FDA_MT, MLP and BDT are considered to be most significant.

A description of these methods can be found in the appendix A.1 and at [5].

TMVApplication tool

With the TMVApplication tool and the weighting files from the training with the TMVAnalysis tool real data can be analysed on the provided shower shape variables. For a better or more precise interpretation of the results, the tool runs separately over signal (single photons), background (double photons) and real data. The first two runs are used to check the quality of the trained methods.

---

[1]The ROOT system provides a set of object oriented (OO) frameworks with all the functionality needed to handle and analyse large amounts of data in a very efficient way. For more details see [4].

7 Results

7.1 Input variables

The figures 7.1 to 7.6 show as an example for E_T=13-15 GeV the shape variables fed into the TMVAnalysis tool for training. These plots (and the corresponding data files) were created with a separate tool ($clustertool.C$ by A. Gadola, K. Müller, C. Schmitz), that reads the H1 data (real or MC data), calculates the different shape variables as described in chapter 6.1 and fills the histograms and data files used with the TMVAnalysis tool. With these histograms a first qualitative estimation of the separation power of the different shower shape variables can be made. The data used are MC simulations for single and double photon events created by K. Müller.

The longitudinal dimension Cl_Rl shows very few separation power in these histograms. Nevertheless, this variable has an influence in the overall analysis as a comparison of the TMVApplication output with and without the variable showed. For the transverse radius Cl_Rt the higher tail of the background on the right side of the histograms in figure 7.1 and 7.2 show as expected, that double photon events yield larger transverse radii which gives a contribution to the separation power of this variable. The separation of signal and background with this variable becomes better as approaching the forward detector region (wheels 4 to 6). Especially for wheel 6, a good separation of signal and background is possible since the background distribution extends to larger radii. For the energy fraction in first e.m. layer $Cl\_EHottestCellL1$ the higher bin content for the signal at small energy fractions for all wheels but wheel 6 gives a good separation power. Multi-photon events have more probability to convert in the first e.m. layer and depose therefore more energy there. For wheel 6 this variable has hardly any separation power. The reason is, that most of the photons convert before they reach the calorimeter in the forward region because there is a significant amount of dead material from readout electronics, cables and the endwalls of the tracking detectors. In opposition to the small energy deposition for single photons in the first e.m. layer, the variable $Cl\_EHottestCell$ shows a higher energy deposition for signal than for background. Even if the separation of signal and background in this variable is not that significant for wheel 1 and 2, it is for the other wheels, especially for wheel 6, where the before mentioned converted photons are detected.
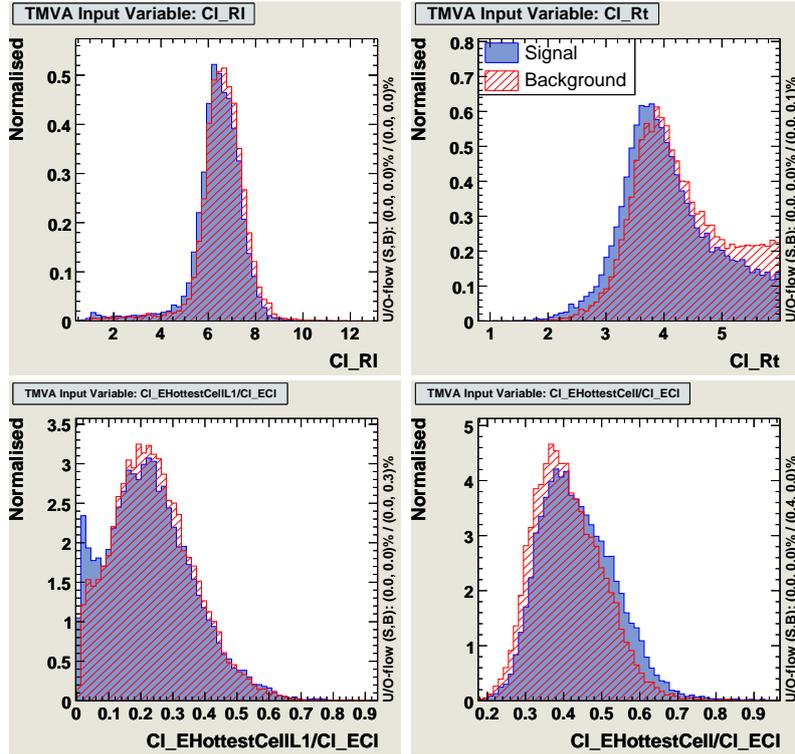
Figure 7.1: Shower shape variables of the MC files for single and double photon events at E_T=13-15 GeV for wheel 1.
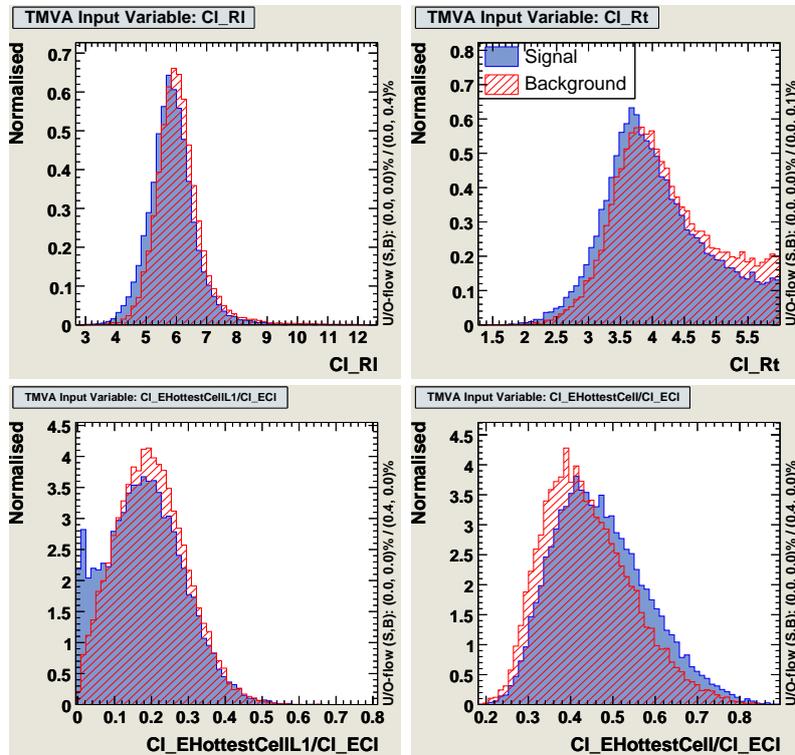
Figure 7.2: Shower shape variables of the MC files for single and double photon events at E_T=13-15 GeV for wheel 2.
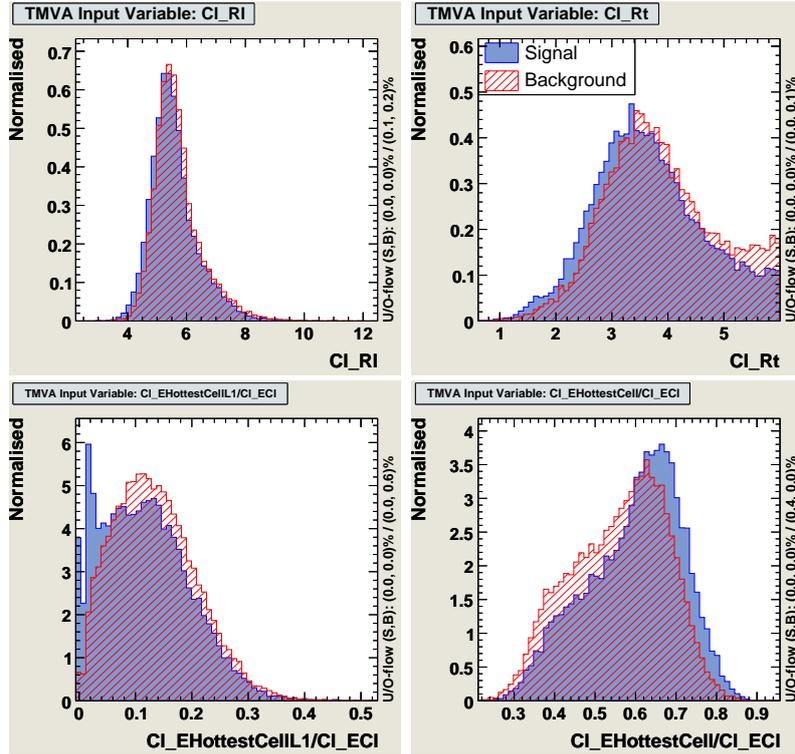
Figure 7.3: Shower shape variables of the MC files for single and double photon events at E_T=13-15 GeV for wheel 3.
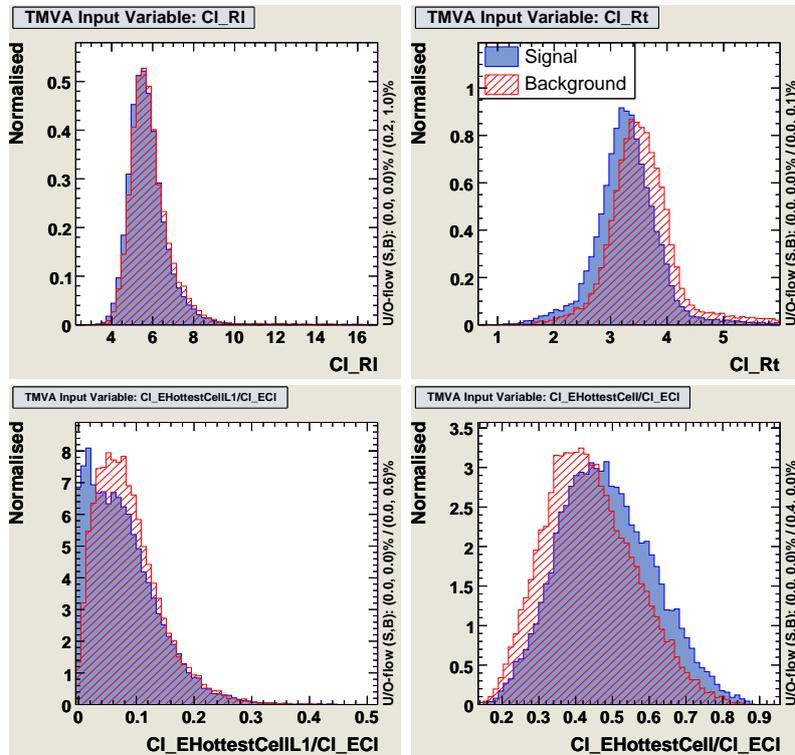
Figure 7.4: Shower shape variables of the MC files for single and double photon events at E_T=13-15 GeV for wheel 4.
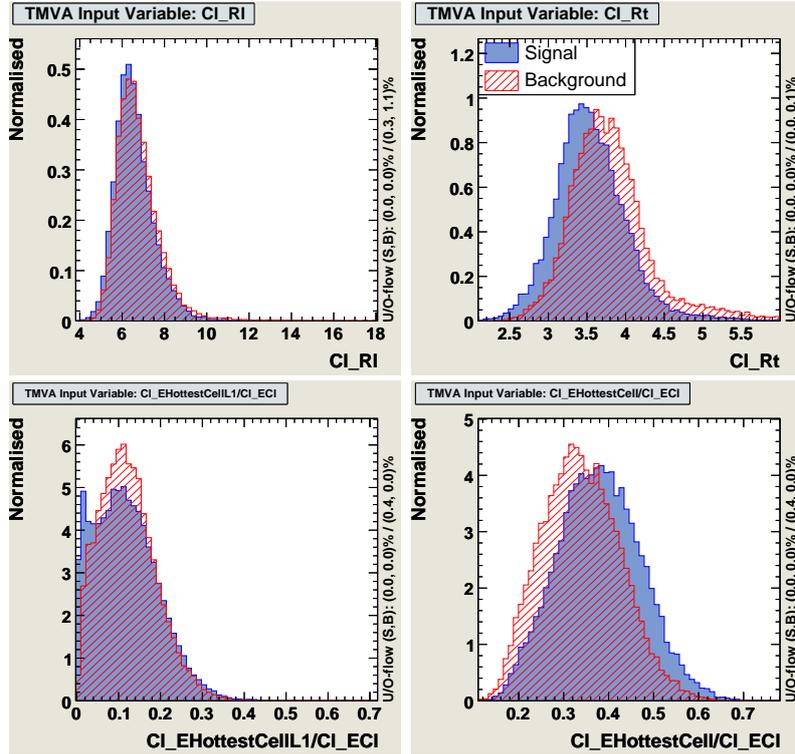
Figure 7.5: Shower shape variables of the MC files for single and double photon events at E_T=13-15 GeV for wheel 5.
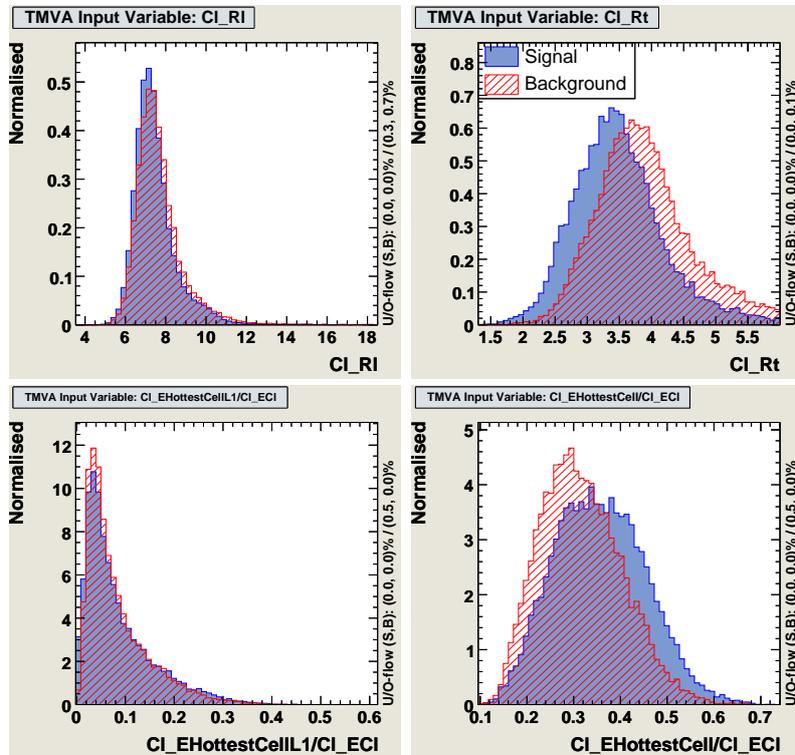
Figure 7.6: Shower shape variables of the MC files for single and double photon events at E_T=13-15 GeV for wheel 6.

7.2 Methods efficiency

Figure 7.7 shows the background rejection versus signal efficiency for six different methods. All methods are remarkably comparable and have about 80% signal efficiency at 80% background rejection for a transverse energy E_T=5-7 GeV. For higher energies (E_T=13-15 GeV) signal efficiency reaches only about 37% at 80% background rejection but the fraction of background events decreases with energy. In the following the results are only presented for the Likelihood method due to its popularity and good understanding but also for its very good background rejection versus signal efficiency at higher energies.

(a) E_T=5-7 GeV   (b) E_T=13-15 GeV
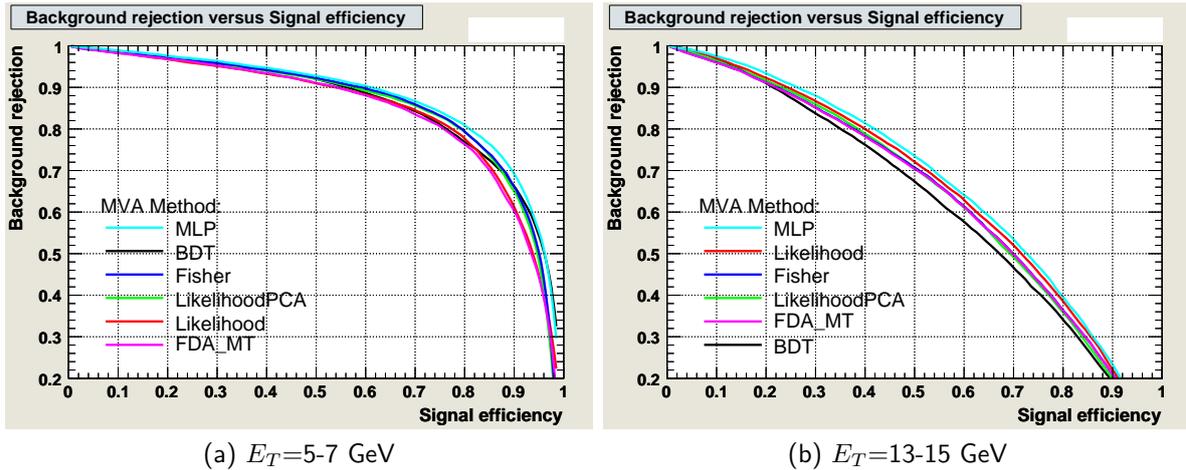
Figure 7.7: Background rejection versus signal-efficiency for selected TMVA methods for wheel 4 and E_T=5-7 GeV (left) and 13-15 GeV (right). Desired are high values for both parameters.

7.3 Variable correlations

The correlation of the chosen shower shape variables (fig. 7.8) show as expected only a rather small correlation between the transverse radius and the longitudinal dimension. All other variables seem to have very little correlation what will later on show up in a high quality discriminator.
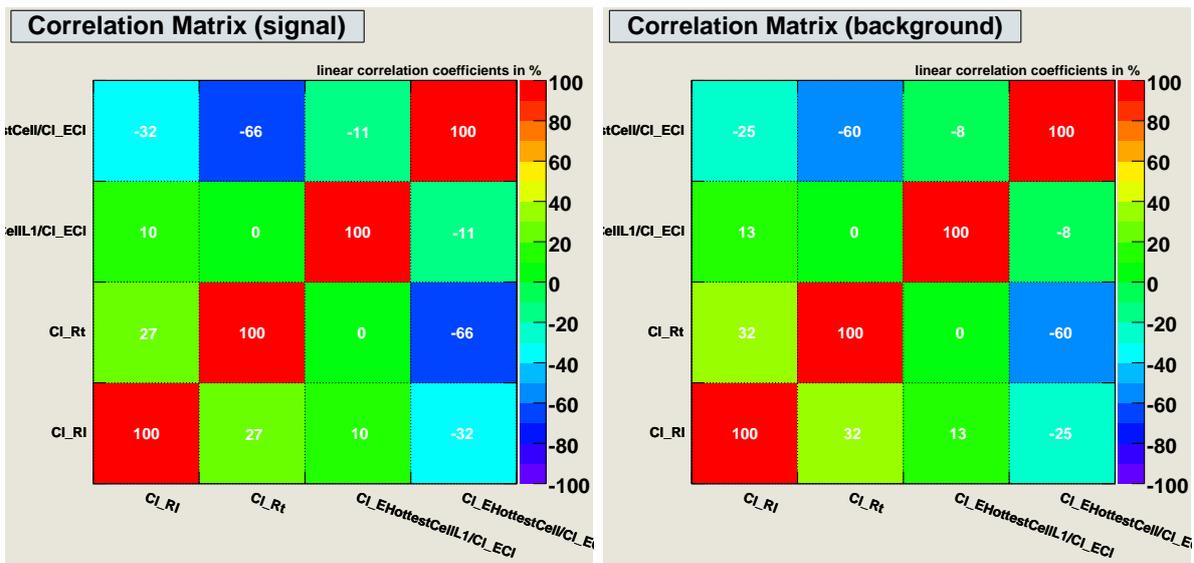
(a) Correlation matrix for signal.   (b) Correlation matrix for background.

Figure 7.8: Correlations between the different shower shape variables for wheel 4 and E_T=13-15 GeV.

7.4 Likelihood discriminator

The figures 7.10 and 7.11 show the Likelihood method with different binnings after passing the TM-VApplication tool. Unfortunately real data were available only for wheels 1 to 3.

The distributions show that there is no distinct separation possible between signal and background mostly because pions decay in two photons with very asymmetrical energies[1] so that these photons cannot be distinguished from single photons. Nevertheless background and single photons have a different shape of the discriminator distribution. This allows to fit the fraction of photon and background events.

The Likelihood histograms show a good separation power for all wheels and all energy intervals except for wheel 1 and 2. In these two wheels signal and background shape of the histograms are still slightly different, so that a separation with a smaller probability is possible. Especially for the energy region of 9 to 11 GeV, the signal separates not too badly in the right bins and has therefore a good probability for correct identification of single photons. The worse separation power of wheel 1 and 2 can be explained by their granularity. An outstanding separation can be found in the forward wheels 4, 5 and 6 for the energy interval 9 to 13 GeV. There, signal and background peak very nicely and the probability to identify single photons correctly is very high.

For a better conclusion, the used methods should be tuned better (mainly the number of neural knots of the neural network could give a big impact to the results but needs longer simulation time).

Real data are shown in figures 7.10a to 7.10i only for illustration. For a conclusion about how real data fit the MC data, signal and background data need to be weighted, summed and normalized. To become an idea how real data fit with the selected variables, figure 7.12 shows a separation with a good data set and well tuned methods which use very similar input variables as in this study (figure 7.9 shows schematically the analysis structure and the used variables). The plots show the likelihood distribution for five wheels and six bins in transverse energies (5-15 GeV). The data is described well in all bins by the sum of the signal for background and photons. The scaling factor for the background and the signal is determined by the fit. The results show that the method works also for high transverse energies where the separation power is poorer. These figures were kindly provided by Krzysztof Nowak [7].
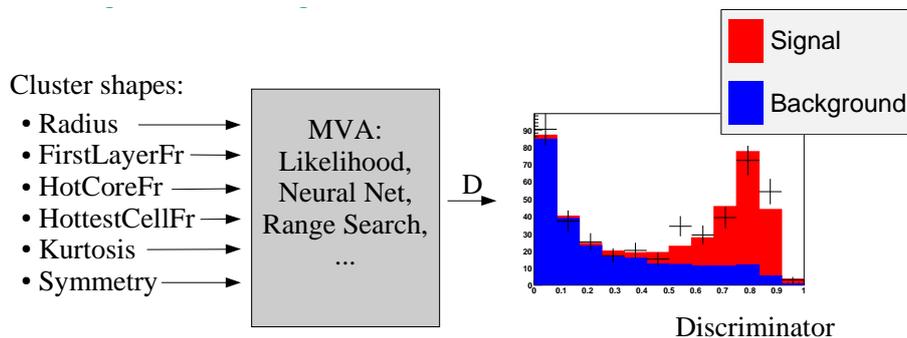
Figure 7.9: Schematic structure of the ongoing analysis of K. Nowak. Source [7].

---

[1] $E_{\gamma 1} \propto 1 + \beta cos(\theta_{CM})$
$E_{\gamma 2} \propto 1 - \beta cos(\theta_{CM})$

(a) Wheel 1, E_T=9-11 GeV.

(b) Wheel 1, E_T=11-13 GeV.

(c) Wheel 1, E_T=13-15 GeV.

(d) Wheel 2, E_T=9-11 GeV.

(e) Wheel 2, E_T=11-13 GeV.

(f) Wheel 2, E_T=13-15 GeV.

(g) Wheel 3, E_T=9-11 GeV.

(h) Wheel 3, E_T=11-13 GeV.

(i) Wheel 3, E_T=13-15 GeV.

(j) Wheel 4, E_T=9-11 GeV.

(k) Wheel 4, E_T=11-13 GeV.

(l) Wheel 4, E_T=13-15 GeV.

(m) Wheel 5, E_T=9-11 GeV.

(n) Wheel 5, E_T=11-13 GeV.

(o) Wheel 5, E_T=13-15 GeV.

(p) Wheel 6, E_T=9-11 GeV.

(q) Wheel 6, E_T=11-13 GeV.

(r) Wheel 6, E_T=13-15 GeV.
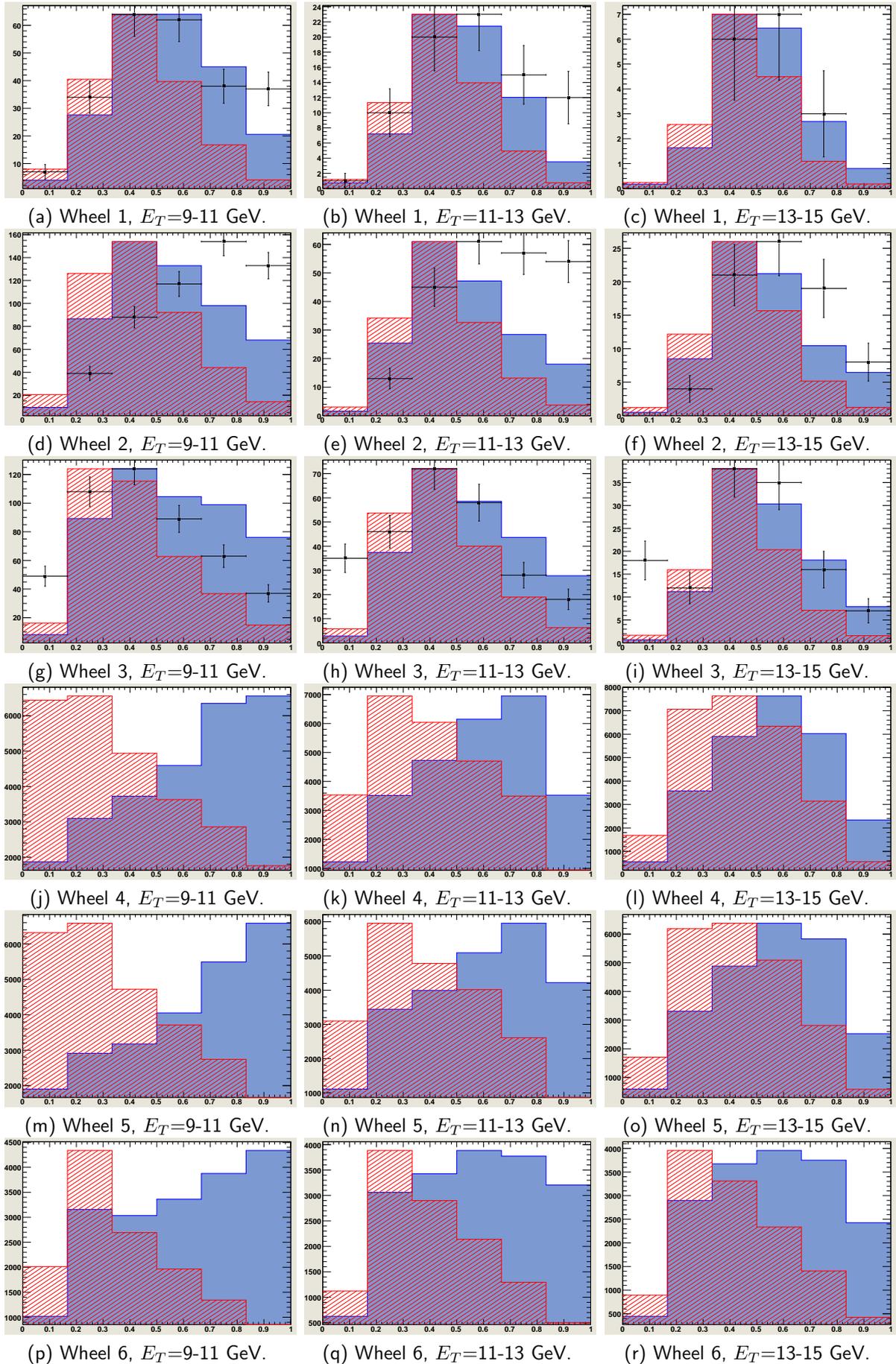
Figure 7.10: Likelihood discriminator (6 bins) of signal (full, blue), background (dashed, red) and data (wheel 1 to 3 only).

(a) Wheel 1, E_T=9-11 GeV.

(b) Wheel 1, E_T=11-13 GeV.

(c) Wheel 1, E_T=13-15 GeV.

(d) Wheel 2, E_T=9-11 GeV.

(e) Wheel 2, E_T=11-13 GeV.

(f) Wheel 2, E_T=13-15 GeV.

(g) Wheel 3, E_T=9-11 GeV.

(h) Wheel 3, E_T=11-13 GeV.

(i) Wheel 3, E_T=13-15 GeV.

(j) Wheel 4, E_T=9-11 GeV.

(k) Wheel 4, E_T=11-13 GeV.

(l) Wheel 4, E_T=13-15 GeV.

(m) Wheel 5, E_T=9-11 GeV.

(n) Wheel 5, E_T=11-13 GeV.

(o) Wheel 5, E_T=13-15 GeV.

(p) Wheel 6, E_T=9-11 GeV.

(q) Wheel 6, E_T=11-13 GeV.

(r) Wheel 6, E_T=13-15 GeV.
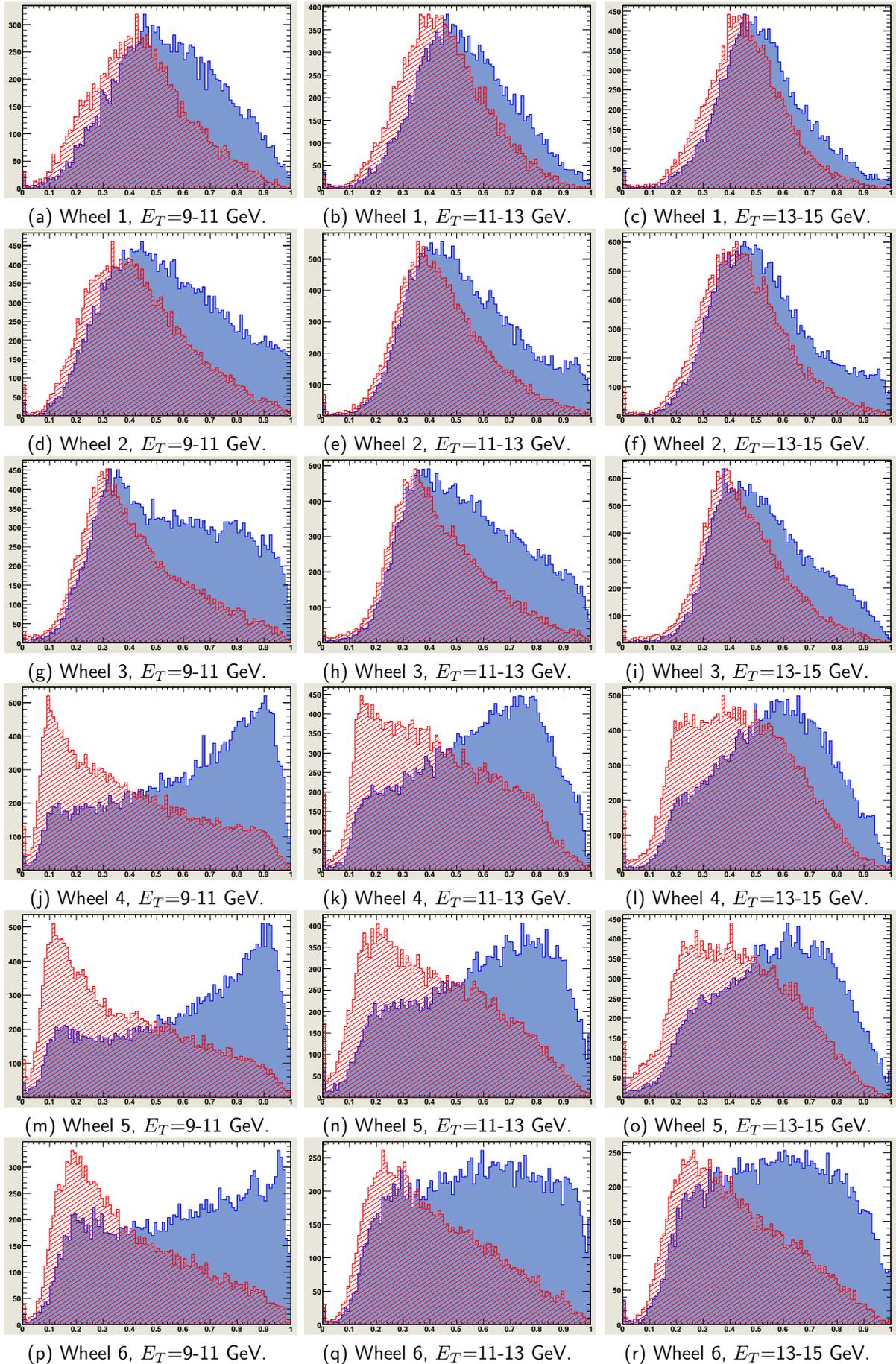
Figure 7.11: Likelihood discriminator (100 bins) of signal (full, blue) and background (dashed, red).
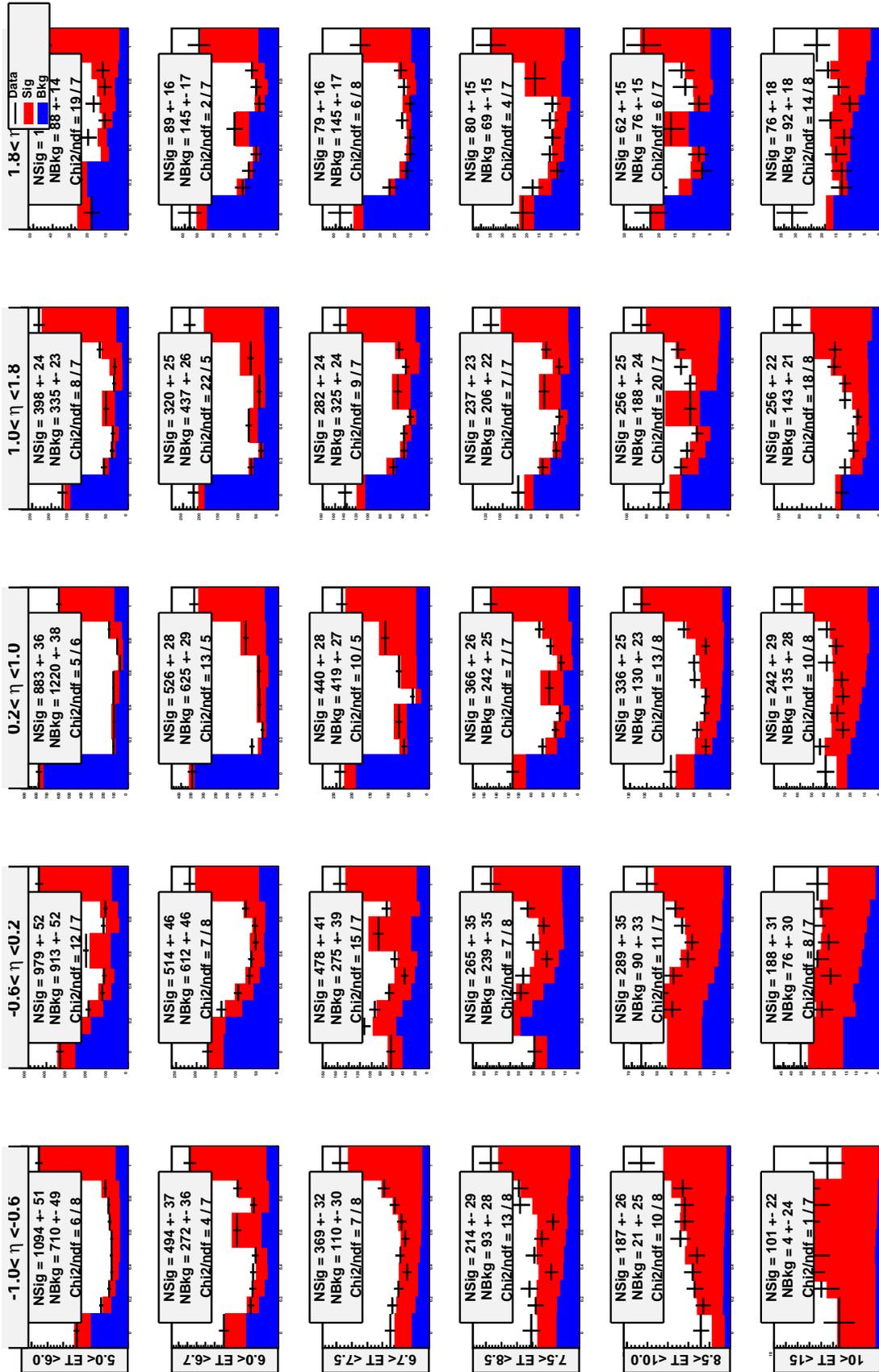
Figure 7.12: Simulated versus real data. Columns=Wheel, rows=E_T. Source [7].

8 Conclusions

The results show that good separation power can be reached for single and double photon events at high transverse energies for wheels in the forward region of the detector. It seems that the selected shower shape variables R_T, R_L, $E_{HottestFrac}$ and $E_{HottestL1Frac}$ are well chosen. For a better understanding of the influence of each variable, a continuing treatment of the topic is indispensable. A fine tuning of the methods by grouping different variables and applying them to the analysis tools in combination with larger data sets for signal, background and real data would surely increase the quality of separation power. Furthermore other analysis methods (especially the MLP method) should be taken into account for a more appropriate or more complete analysis at higher energies.

A next step would also be the calculation of a cross section that would give an even more precise statement about separation power of different combinations of shower shape variables linked to different analysis methods.

The results show with good credibility that future analysis should definitively make use of the energy range above 10 GeV and the longitudinal dimension of the shower.

9 Acknowledgments

A Appendix

A.1 Description of TMVA methods

A.1.1 Projective Likelihood (PDE Approach)

The method of maximum likelihood is among the most straightforward multivariate analyser approaches. We define the likelihood ratio, R, for an event by the ratio of the signal to the signal plus background likelihoods. The individual likelihoods are products of the corresponding probability densities of the discriminating input variables used. In practice, TMVA uses polynomial splines fitted to histograms, or unbinned Gaussian kernel density estimators, to estimate the probability density functions (PDF) obtained from the distributions of the training variables.
Likelihood responses are often strongly peaked at $0/1$. The booking option "TransformOutput" zooms into these peaks (with no change in the performance) using an inverse sigmoid transformation.

A.1.2 Fisher and Mahalanobis Discriminants

In the method of Fisher discriminants event selection is performed in a transformed variable space with zero linear correlations, by distinguishing the mean values of the signal and background distributions. The linear discriminant analysis determines an axis in the (correlated) hyperspace of the input variables such that, when projecting the output classes (signal and background) upon this axis, they are pushed as far as possible away from each other, while events of a same class are confined in a close vicinity. The linearity property of this method is reflected in the metric with which "far apart" and "close vicinity" are determined: the covariance matrix of the discriminant variable space.
The classification of the events in signal and background classes relies on the following characteristics (only): overall sample means for each input variable, class-specific sample means, and total covariance matrix. The covariance matrix can be decomposed into the sum of a within- and a between-class class matrix. They describe the dispersion of events relative to the means of their own class (within-class matrix), and relative to the overall sample means (between-class matrix). The Fisher coefficients are then given by the product of the difference vector of signal and background sample means and the inverse within-class matrix.

A.1.3 Function Discriminant Analysis (FDA)

The common goal of all TMVA discriminators is to determine an optimal separating function in the multivariate space represented by the input variables. The Fisher discriminant solves this analytically for the linear case, while artificial neural networks, support vector machines or boosted decision trees provide nonlinear approximations with – in principle – arbitrary precision if enough training statistics is available and the chosen architecture is flexible enough.
The function discriminant analysis (FDA) provides an intermediate solution to the problem with the

aim to solve relatively simple or partially nonlinear problems. The user provides the desired function with adjustable parameters via the configuration option string, and FDA fits the parameters to it, requiring the signal (background) function value to be as close as possible to 1 (0). Its advantage over the more involved and automatic nonlinear discriminators is the simplicity and transparency of the discrimination expression. A shortcoming is that FDA will underperform for involved problems with complicated, phase space dependent nonlinear correlations.

The FDA performance depends on the complexity and fidelity of the user-defined discriminator function. As a general rule, it should be able to reproduce the discrimination power of any linear discriminant analysis. To reach into the nonlinear domain, it is useful to inspect the correlation profiles of the input variables, and add quadratic and higher polynomial terms between variables as necessary. Comparison with more involved nonlinear classifiers can be used as a guide.

A.1.4 Artificial Neural Networks (Non-Linear Discriminant Analysis)

Three different ANN implementations are used in TMVA. The TMlpANN, implemented in ROOT, the Clermont-Ferrand ANN (CFMlpANN), which has been translated from FORTRAN, and a new ANN (MLP), which is very similar to the ROOT ANN, but can be trained significantly faster. All ANNs belong to the class of Multilayer Perceptrons (MLP), which are feed-forward networks according to the following propagation schema: The input layer contains as many neurons as input variables used
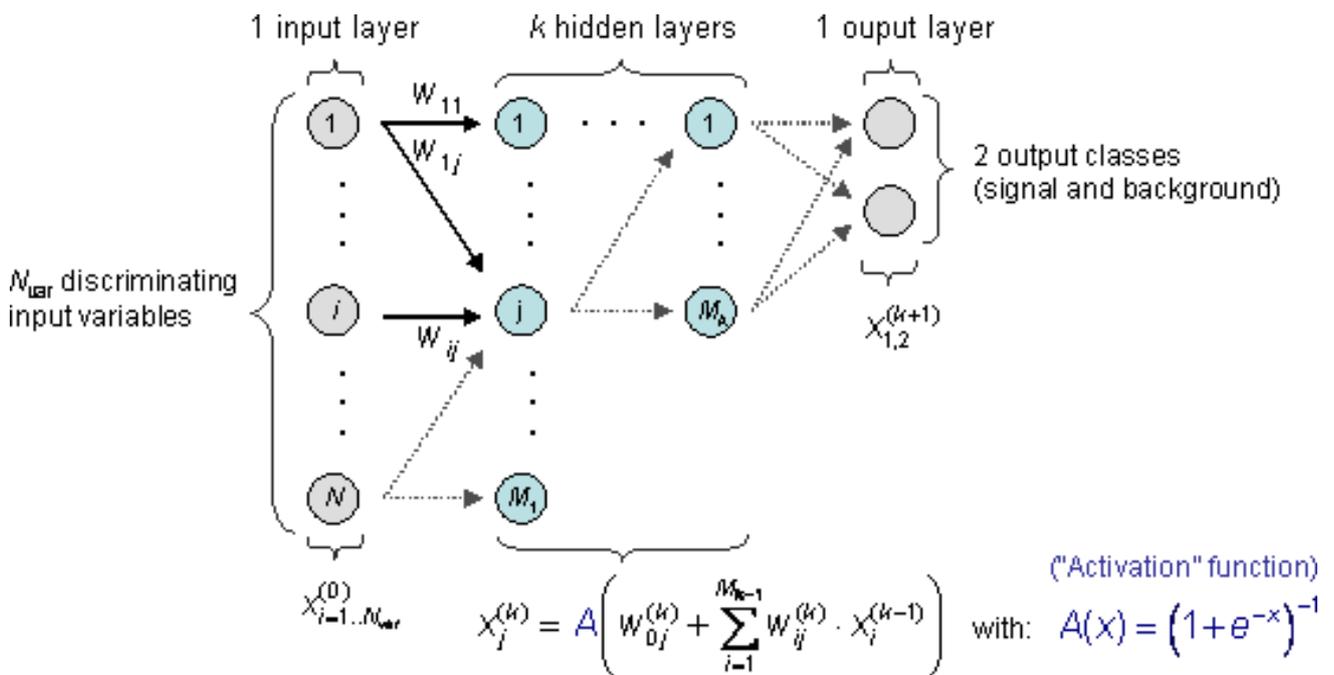
Figure A.1: [5].

in the MVA. The output layer contains a single neuron for the signal weight. In between the input and output layers are a variable number of k hidden layers with arbitrary numbers of neurons. (While the structure of the input and output layers is determined by the problem, the hidden layers can be configured by the user through the option string of the method booking.)

As indicated in the sketch, all neuron inputs to a layer are linear combinations of the neuron output of the previous layer. The transfer from input to output within a neuron is performed by means of an

"activation function". In general, the activation function of a neuron can be zero (deactivated), one (linear), or non-linear. The above example uses a sigmoid activation function. The transfer function of the output layer is usually linear. As a consequence: an ANN without hidden layer should give identical discrimination power as a linear discriminant analysis (Fisher). In case of one hidden layer, the ANN computes a linear combination of sigmoid.

A.1.5 Boosted Decision Trees

Boosted decision trees have been successfully used in High Energy Physics analysis for example by the MiniBooNE experiment (Yang-Roe-Zhu, physics/0508045). In Boosted Decision Trees, the selection is done on a majority vote on the result of several decision trees, which are all derived from the same training sample by supplying different event weights during the training.

Decision trees: successive decision nodes are used to categorize the events out of the sample as either signal or background. Each node uses only a single discriminating variable to decide if the event is signal-like ("goes right") or background-like ("goes left"). This forms a tree like structure with "baskets" at the end (leave nodes), and an event is classified as either signal or background according to whether the basket where it ends up has been classified signal or background during the training. Training of a decision tree is the process to define the "cut criteria" for each node. The training starts with the root node. Here one takes the full training event sample and selects the variable and corresponding cut value that gives the best separation between signal and background at this stage. Using this cut criterion, the sample is then divided into two subsamples, a signal-like (right) and a background-like (left) sample. Two new nodes are then created for each of the two sub-samples and they are constructed using the same mechanism as described for the root node. The devision is stopped once a certain node has reached either a minimum number of events, or a minimum or maximum signal purity. These leave nodes are then called "signal" or "background" if they contain more signal respective background events from the training sample.

Boosting: the idea behind the boosting is, that signal events from the training sample, that end up in a background node (and vice versa) are given a larger weight than events that are in the correct leave node. This results in a re-weighed training event sample, with which then a new decision tree can be developed. The boosting can be applied several times (typically 100-500 times) and one ends up with a set of decision trees (a forest).

Bagging: In this particular variant of the Boosted Decision Trees the boosting is not done on the basis of previous training results, but by a simple stochasitc re-sampling of the initial training event sample.

Analysis: applying an individual decision tree to a test event results in a classification of the event as either signal or background. For the boosted decision tree selection, an event is successively subjected to the whole set of decision trees and depending on how often it is classified as signal, a "likelihood" estimator is constructed for the event being signal or background. The value of this estimator is the one which is then used to select the events from an event sample, and the cut value on this estimator defines the efficiency and purity of the selection.

Bibliography

[1] http://m.desy.de/

[2] http://www-h1.desy.de/

[3] http://www-h1.desy.de/h1det/detpaper/contents.html

[4] http://root.cern.ch/

[5] http://tmva.sourceforge.net/#root_scripts

[6] Isolated Photon Production in Deep-Inelastic Scattering at HERA
Carsten Schmitz, Ph.D. Thesis Zurich University, 2007, in preparation, to be made available at
http://www-h1.desy.de/psfiles/theses/h1th-477.pdf

[7] Prompt photons in photoproduction - analysis update
Krzysztof Nowak, Talk at HaQ Meeting 05.02.2008
https://www-h1.desy.de/icgi-h1wiki/moin.cgi/HaqGroup/Meeting2008-02-05

[8] Prompt Photon Production at HERA
Rachid Lemrani-Alaoui, Ph.D. Thesis, Hamburg, 2003
http://www-h1.desy.de/publications/theses_list.html

[9] Kern- und Teilchenphysik
Claude Amsler, Publisher UTB, ISBN: 978-3-8252-2885-9, Chapter 16.7, p. 287ff

[10] Isolated Photons in DIS
Dr. Katharina Müller, University of Zurich, Presentation at DIS 17 April 2007, Munich

[11] I. Abt et al., Nucl. Instr. and Meth. **A 386** (1997) 310; ibid, 348.

The following papers were used for introduction into topic and for general information and are available on www.arxiv.org.

[12] Measurement of prompt photon cross sections in photoproduction at HERA
By H1 Collaboration (A. Aktas et al.). DESY-04-118, Jul 2004. 21pp.
Published in Eur.Phys.J.C38:437-445,2005.
hep-ex/0407018

[13] Measurement of Isolated Photon Production in Deep-Inelastic Scattering at HERA
F. D. Aaron *et al.* [H1 Collaboration].
hep-ex/0711.4578