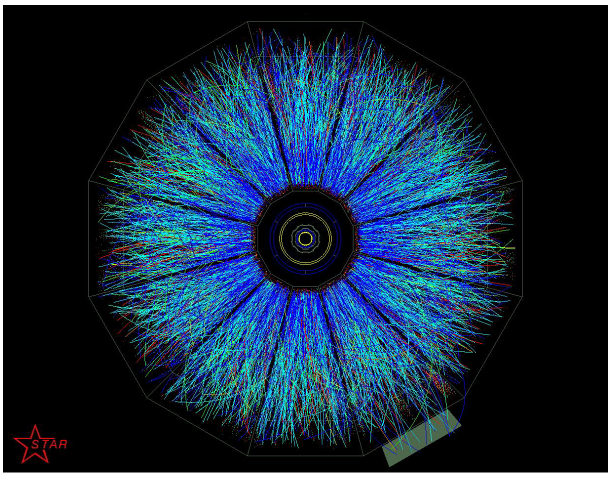
Study of Signal-to-Background Ratio of Reconstructed Neutral Pions from p+p Collisions at $\sqrt{s} = 200$ GeV

Jory Sonneveld
Institute of Subatomic Physics
Faculty of Physics and Astronomy
Utrecht University

July 24, 2009



Supervisor: Prof. Dr. Thomas Peitzmann
Cosupervisor: Dr. André Mischke

Front page image[1]

Particles streaming out from a head-on collision between gold nuclei, as seen by the STAR detector.

[1]Figure taken from http://www.lbl.gov/Science-Articles/Archive/RHIC-first.html.

Acknowledgments

I would like to express my gratitude to my supervisor Professor Dr. Thomas Peitzmann and my cosupervisor Dr. Andre Mischke for their inspiration and patient explanations on the many topics of particle physics that were new to me. I have learned much from them while they enthusiastically told me about this area of research and have found their corrections very helpful. I also want to thank Ermes Braidot for the time he spent explaining me the physics of and the programming for this analysis. Without his help, I would not have known the physical meaning of my data and some of my programs would still not have worked. In addition, I would like to thank Raoul, Cristian, Marek, Marco, and Marta for stopping by to help, as well as Wicher for his endless patience when helping me achieve what I want from a computer.

Abstract

One way to examine the theory of Colour Glass Condensate is to compare data analyses of proton-proton-collisions to proton-nucleus collisions. In this paper, proton-proton-collisions are analyzed to contribute to the baseline measurement for the Color Glass Condensate. Neutral pions are reconstructed from pairs of photon candidates. To optimize the pion reconstruction, a systematic study of the signal-to-background ratio in the invariant mass distributions as a function of transverse momentum is performed.

Contents

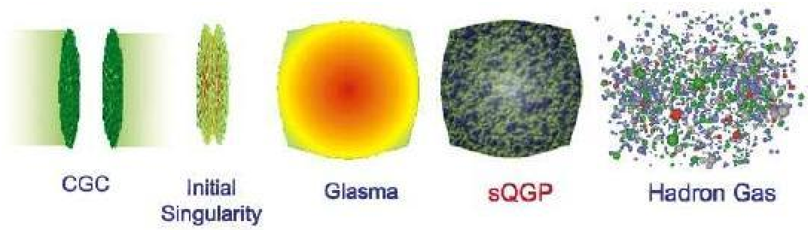


Figure 1: Evolution of a heavy ion collision. Figure taken from [9].

1 Introduction

1.1 Color Glass Condensate

The Color Glass Condensate is a state of matter expected to appear in dense strongly interacting systems such as collisions of hadrons or nuclei at high energies. Its production is much more likely in heavier systems, e.g. in p+Au or even Au+Au reaction systems. However, to interpret these, first a thorough understanding of the elementary p+p system that is studied here must be obtained. All these collisions of high energy particles occur at relativistic speeds. Because of Lorentz contraction, a nucleus of an atom appears contracted in the center-of-momentum of the collision. Furthermore, the high-energy collision probes the abundant low momentum gluons in the nuclei, which are virtual particles in a nucleus at rest. The result is a highly dense matter of quarks and gluons each carrying color charge, which behaves like a nearly solid fluid similar to glass: Color Glass Condensate [1]. This substance moves at a relativistic speed, so that the natural time evolution of gluons is time dilated and the substance behaves like a liquid [9]. This resembles ordinarily glass, which is liquid at longer time scales but solid at shorter time scales. The theory of Quark-Gluon Plasma describes the product of collisions of the highly dense Lorentz-contracted particles; the theory of Color Glass Condensate describes these dense particles themselves. How matter evolves after a heavy ion collision is displayed in figure 1. The early physical conditions of collisions of heavy ions resemble those of the early universe [9]. Statistics of millions of measurements taken after collisions are used to study what might have happened just after what is called “The Big Bang” in cosmology.

1.2 Neutral Pions

In the STAR experiment, among other particles, protons were collided with protons (p+p). Data from these collisions are used in this analysis. Other collisions are, for example, deuteron-gold nuclei collisions (d + Au). When two quarks, one from, for instance, a deuteron, and one from a gold nucleus, collide, they separate with such large energies that new quarks are formed. This is because quarks always come in pairs of two or three [7]. These quarks then form new hadrons, such as π mesons, kaons, muons, or protons. In this way the quark that has scattered will fragment into a jet of hadrons [7], which together carry the original momentum of the quark.[2] The two scattered quarks have to be balanced in transverse momentum – this is reflected in the back-to-back emission of the two jets. In more than 90 percent of the cases, one of the hadrons will be a pion ($\pi^{\pm}$, π^0). Here, neutral pions (π^0) are studied as "representatives" of the jets. With a probability of 99 percent, neutral pions decay into two photons after a mere 8.4×10^{-17} seconds with a decay energy (mass) of 135.0MeV/c^2. The photons are then measured by a detector; it is this information that will be used to reconstruct neutral pions.

In order to say more about the Color Glass Condensate, proton-proton collisions will be compared with collisions involving nuclei. For both collisions, 'measurements' of photons will be used. Energy from the collided particles is measured by two detectors (the Barrel Electromagnetic Calorimeter, or BEMC, and the Forward Meson Spectrometer, the FMS; see section 2). In this analysis, measurements from the BEMC will be used. Some of the signals measured by this detector are taken as energies of photon candidates. Pairs of those photon candidates are potential neutral pions. This hypothesis can be tested by studying the distribution of the invariant mass of the pairs. True neutral pions will contribute to a peak at the neutral pion mass; this peak will be called the signal. 'False combinations' of photons that together did not originate from the same pion will also arise; these form a combinatorial background of uncorrelated photon pairs in the invariant mass distribution. From studies of the signal-to-background ratio, the likelihood that a reconstructed neutral pion really was a neutral pion can be estimated. Ratios of about one and a half are sufficient to use these reconstructed pions for further calculations.

1.3 Beyond this Analysis

A study similar to the one described above can be performed on different combinations of reconstructed neutral pions to reconstruct the original collision. Instead of a mass distribution, now a plot of the relative azimuthal angle distribution is made from all possible combinations of the earlier reconstructed pions. The azimuthal angle is the difference between the angles of two neutral pions departing after a collision. The angles are measured in the transverse plane, or the plane perpendicular to the original beam momentum. A peak is likely to be seen around 180 degrees for proton-proton collisions. For proton-gold nucleus collisions, the peak is expected to be

[2] In relativistic mechanics, only the laws of conservation of energy and momentum apply, not the law of conservation of mass [8].

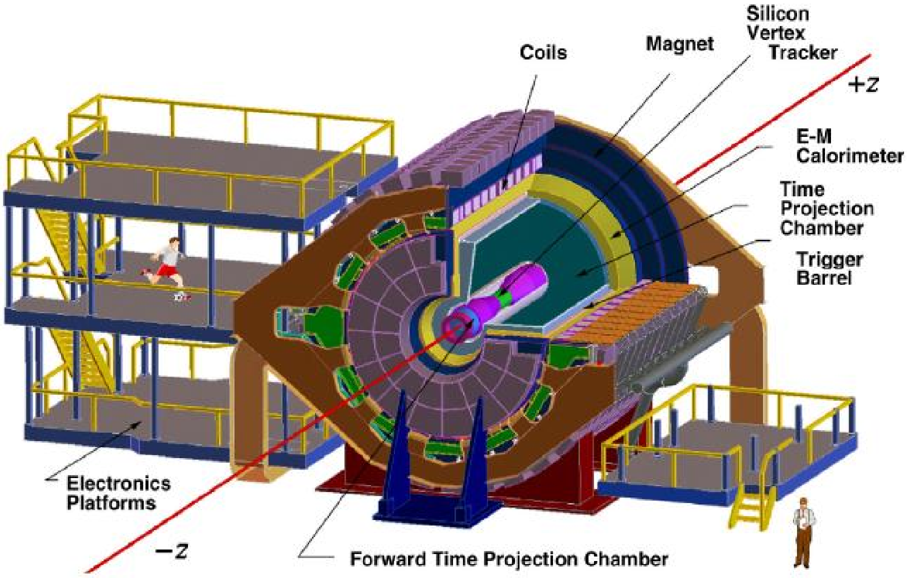


Figure 2: The STAR detector. Figure taken from [2].

different. Different theories predict different distributions of the azimuthal angle. If, for example, the peak of deuteron-gold collision angles is lower than that of proton-proton collision angles, one can expect the theory predicting a similar peak to be correct. The interpretation of such an analysis, however, relies on the probability that the particles used to reconstruct the collisions really were neutral pions, and thus on the signal-to-background ratio in the invariant mass distribution.

2 The STAR Detector

The STAR detector, which stands for Solenoidal Tracker at RHIC, is one of the four experiments at the Relativistic Heavy Ion Collider (RHIC) in Brookhaven National Laboratory. Several detectors are placed inside a magnetic field. The specific detector used for this analysis is the Barrel Electro-Magnetic Calorimeter (BEMC) (see figure 2). Inside the BEMC, the Time Projection Chamber is situated; this central tracking device of the STAR detector can track charged particles [6]. However, for technical reasons the data from the TPC are not used in this analysis.

2.1 The Barrel Electromagnetic Calorimeter

The barrel electromagnetic calorimeter, or BEMC, is a combination of lead and plastic scintillator layers that measures electrogmagnetic energy. A scintillator is a material that has the property of luminescence when excited by ionizing radiation. STAR uses the barrel electromagnetic calorimeter to study high transverse momentum processes such as jets or direct photons, and to provide a higher maximum range momentum for photons, electrons, neutral pions, and η mesons [3]. The transverse momentum, or p_{T}, is the momentum perpendicular to the axis along which the beam enters the detector. The BEMC is divided into two half barrels, each with a length of 293 cm, an inner radius of 223 cm, and an outer radius of 263 cm. Each barrel of the BEMC is divided into 60 modules of 23 cm wide and 6 degrees of azimuth (the angle in the azimuthal plane, which is perpendicular to the axis along which the beam comes in). Each module is segmented into 40 tower cells of lead-scintillator stacks [6] (see figure 5). They cover a pseudorapidity $-1 < \eta < 0$ and $0 < \eta < 1$, respectively [6]. Pseudorapidity is a measurement of the polar angle θ, or the azimuth:

$$\eta = \ln \tan \frac{\theta}{2}. \tag{1}$$

The pseudorapidity is the rapidity y in the limit where the mass, which is insignificantly small here, goes to zero:

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z}, \tag{2}$$

where $E = \sqrt{m^2 + p^2}$ is the energy of the particle including its mass. In figure 3 one can see the range of η and the azimuthal angle ϕ of the BEMC. In figure 4 the variables x, y, and z are plotted in a 3D graph; looping over data from many collisions it yields the shape of the detector.

In the towers, the signal of electromagnetic particles (photons, for instance) is measured. The total amplitude of the signal in the towers yields the energy (or absolute value of the momentum) of candidate photons. The spatial position of the signal can be used to calculate the direction of the candidate photon momentum.

3 Data Set

For this analysis, proton-proton collisions from the 2008 data taking are used. The energy of the center of mass of the collisions was $\sqrt{s} = 200$ GeV. Eleven to fifteen hundred files, each of around 5000 events, were used; this means that a total of 5.5 to 7.5 million events were used for this analysis. A

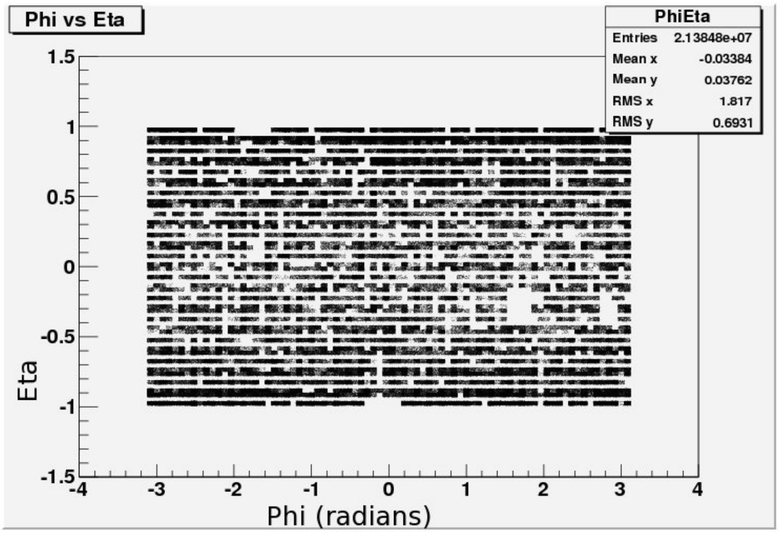


Figure 3: Hit distribution in the ϕ-η plane of the BEMC.

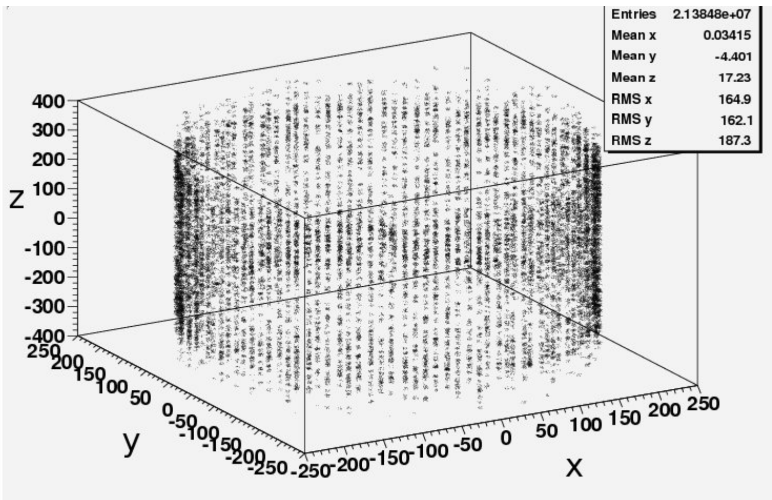


Figure 4: The x-,y-, and z-coordinates of many events, which altogether from the shape of the BEMC detector.

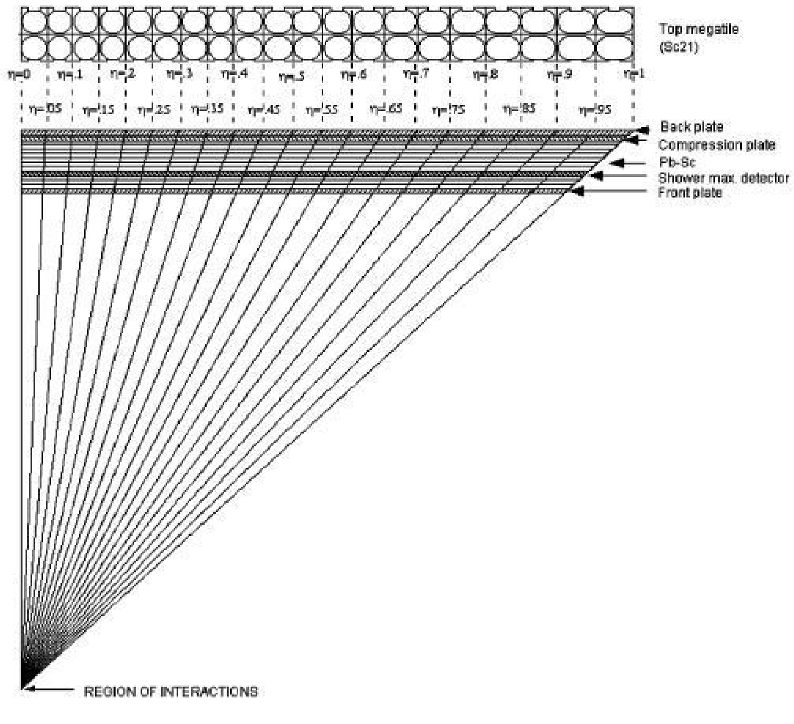


Figure 5: Side view of one BEMC module. Figure taken from [3].

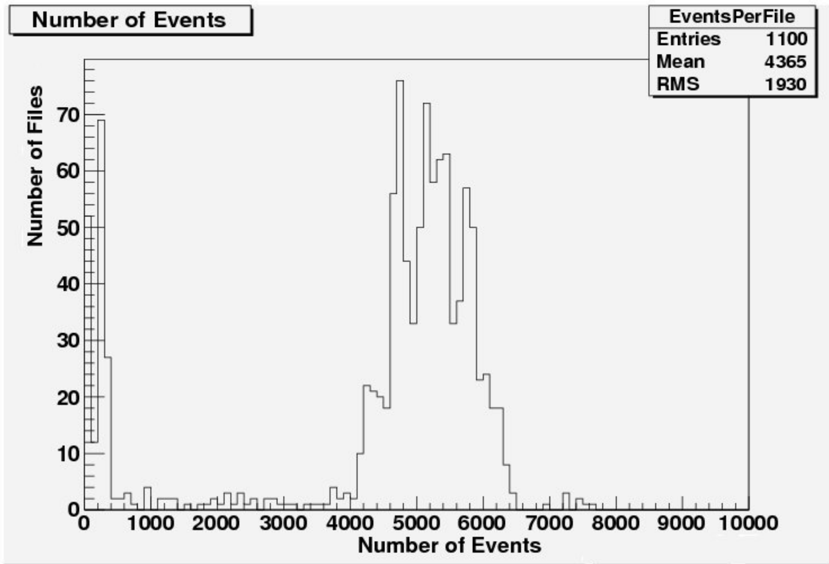


Figure 6: Distribution of the number of events stored in a file.

distribution of the number of entries in the files, which indicate the number of events in a file, can be seen in figure 6. It can be seen that a number of files contain only few events; these are likely to be corrupted files, which were rejected in this analysis.

All candidate photons in an event are combined to pairs of which the invariant mass is calculated. The invariant mass distributions show a peak around the mass of a neutral pion. In order to find candidate photons that can be used for reconstructing neutral pions, deposits of energy must be measured in the modules of the BEMC. A cluster finding algorithm is used to categorize certain clusters of energy as photons [6]. It starts with listing all energy deposits with a minimum energy E_{seed}. Starting with the highest energy measured in a tower on the list, the cluster finder then adds all adjacent hits with a minimum energy E_{add} to the cluster, and removes it from the list. This clustering of one candidate photon stops when no more adjacent hits are found, or when a maximum cluster size N_{max} is reached. It must be noted that the cluster finder does not combine hits from adjacent modules. These modules are, however, separated by a distance of 12 mm, making it very unlikely that two hits from adjacent modules came from a single photon [6].

This analysis uses data that originate from a slightly adjusted cluster finding algorithm. Because the towers are very large, it is possible that two photons end up in one tower and be counted as one photon. For these data, all towers surrounding a peak (a photon) were excluded from being taken for a photon, so that any one of two photons that would land in neighboring towers is not counted. In other words, a minimum distance of photons was made so that the cluster finder was more likely to find real photons. The minimum add-threshold E_{add} is set larger than the peak-threshold E_{seed}, so any energy of a photon ending up in a neighboring tower is not used. If more cells are added (E_{add}) to E_{seed}, more information will be available, but there will also be more noise. Here, none of E_{add} were added to E_{seed}; less emphasis is laid on the statistics (number of events) and more emphasis is laid on the guarantee of dealing with real neutral pions.

4 Analysis

To analyze the data, the analysis and graphic program ROOT has been used [4]. Below follows a description of the process of reconstructing neutral pions. After that, quality cuts are discussed. Then, the fitting procedure is given, whereafter the calculation of the signal-to-background ratio will explained. Finally, the statistic and systematic errors in this process are examined.

4.1 Reconstruction of Neutral Pions

As mentioned before, a neutral pion decays into two photons:

$$\pi^0 \rightarrow \gamma\gamma, \qquad \qquad \mathrm{BR} = 99\%. \tag{3}$$

After the clustering and listing photons, it remains unknown which photons together originated from an actual neutral pion. To reconstruct a pion, combinations of all photons are made.[3] This results in a graph with a peak around the mass of the neutral pion and a combinatorial background of photon pairs that did not originate from one single neutral pion. The invariant mass of photon combinations is calculated according to:

$$M_{\text{pion}} = \sqrt{E_{\text{pion}}^2 - (p_{x,1} + p_{x,2})^2 - (p_{y,1} + p_{y,2})^2 - (p_{z,1} + p_{z,2})^2} \tag{4}$$

where E is the energy, and p is the momentum of one of the two photons. The energy of the pion is just the sum of the measured energies of the

[3]For looping over all files, a program for making a list and opening files one by one by Michael B. Anderson (dating from August 19, 2007) has been used. See http://www.hep.wisc.edu/~mbanderson/public/ROOT_Macros/PlotFromFiles-OLD.C (01/04/09)

individual candidate photons:

$$E_{\text{pion}} = E_1 + E_2. \tag{5}$$

Then η, ϕ, and the transverse momentum p_T can be calculated for each candidate pion:

$$\begin{aligned}
p_{\text{T}} &= \sqrt{p_x^2 + p_y^2} && (6)\\
\eta &= \frac{1}{2} \ln \frac{E + p_z}{E - p_z} && (7)\\
\phi &= \arctan \frac{y}{x}. && (8)
\end{aligned}$$

This is possible using the following variables, in terms of photon one and photon two:

$$\begin{aligned}
x_{\text{pion}} &= \frac{E_1 x_1 + E_2 x_2}{E_1 + E_2} && (9)\\
y_{\text{pion}} &= \frac{E_1 y_1 + E_2 y_2}{E_1 + E_2} && (10)\\
z_{\text{pion}} &= \frac{E_1 z_1 + E_2 z_2}{E_1 + E_2} && (11)\\
R_{\text{pion}} &= \sqrt{x_{pion}^2 + y_{pion}^2 + z_{pion}^2} && (12)\\
p_{x,\text{pion}} &= E_{\text{pion}} \frac{x_{\text{pion}}}{R_{\text{pion}}} && (13)\\
p_{y,\text{pion}} &= E_{\text{pion}} \frac{y_{\text{pion}}}{R_{\text{pion}}} && (14)\\
p_{z,\text{pion}} &= E_{\text{pion}} \frac{z_{\text{pion}}}{R_{\text{pion}}}. && (15)\\
& && (16)
\end{aligned}$$

With these equations it is now possible to impose cuts on the distribution, leaving out the values of certain parameters (such as η, Δ (the asymmetry, see section 4.2), and p_T). Some values are cut out because they are not relevant in this study (for example, very small values of transverse momentum cannot have originated from neutral pions). Cuts are also used to study the change in the signal-to-background ratio as a function of changing parameters.

4.2 Quality Cuts

To minimize the combinatorial background, or, in other words, to maximize the signal-to-background ratio, quality cuts can be made on the data that are used. One example is cutting out candidate photons with low transverse momentum; for these momenta it is very likely that particles other than

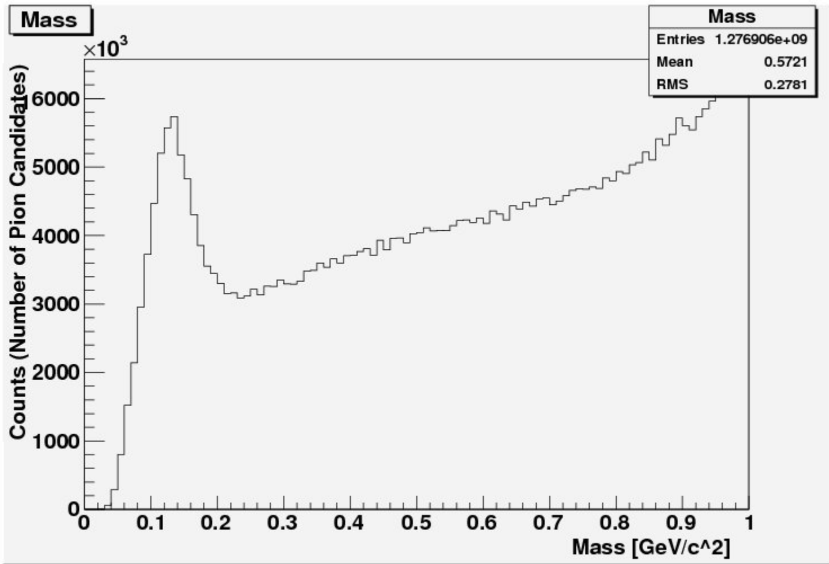


Figure 7: Invariant mass of photon pairs. No cuts were applied.

photons contribute to the photon candidate sample. Simultaneously, the asymmetry can be cut. The energy asymmetry Δ is defined as follows:

$$\Delta = \left| \frac{E_1 - E_2}{E_1 + E_2} \right| \tag{17}$$

The figures 7 and 8 an example of how the combinatorial background is reduced is shown. In figure 7 no cuts were applied. In figure 8 only data where

$$p_{\mathrm{T}}^{\pi^0} \geq 1.5 \quad \mathrm{GeV}/c \tag{18}$$

$$\Delta \leq 0.7 \tag{19}$$

were used. There are more entries for $p_{\mathrm{T}}^{\pi^0}$ greater than 1 than for $p_{\mathrm{T}}^{\pi^0}$ greater than 1.5; in figure 8, a sharper peak around the mass of a pion is visible, and the background is reduced. Altering the upper bound for the asymmetry does not change much. In this analysis, an asymmetry upper bound of 0.7 is taken for all invariant mass distributions.

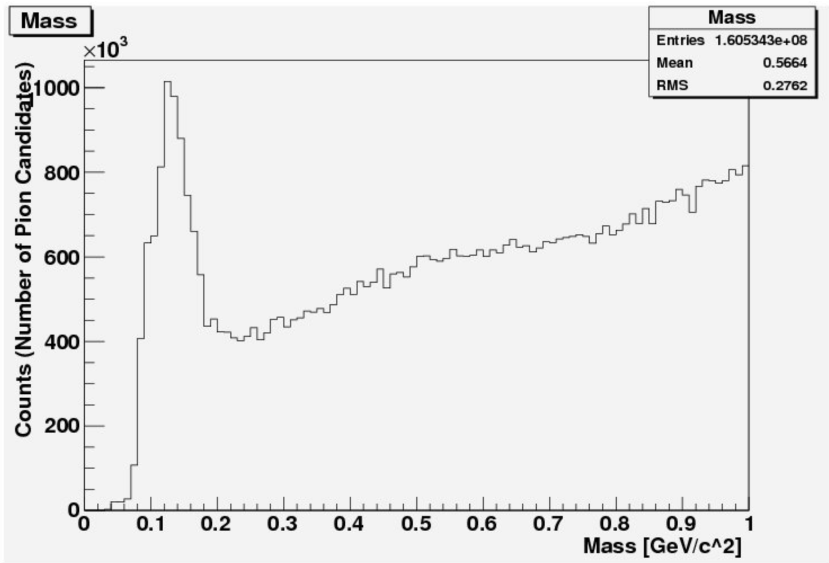


Figure 8: Invariant mass distribution of photon pairs for a transverse momentum greater than 1.5 GeV/c^2, and an asymmetry less than 0.7.

The rapidity η, which is a measure of the momentum along the z-axis [10], has been cut for all photons. It is defined as follows:

$$\eta = \frac{1}{2} \ln \frac{E + p_z}{E - p_z}. \tag{20}$$

The rapidity, which is 0 at the center of the BEMC, and 1 at the ends, is constrained to the values $[-0.9, 0.9]$ because measurements from the edges of the BEMC are found to be less reliable [6]. The value of η can be extracted directly from the data set.

Another possibility to eliminate particles that are not photons is to cut out measured electrically charged particles, since photons do not carry an electric charge. However, it could be possible that a photon would accidentally overlap with charged particle; then it will be erroneously discarded. In the data for this analysis, such cuts have not been applied.

4.3 Fitting the Invariant Mass Distributions

Plots of invariant mass distributions are fitted with functions to make a reasonable estimate of the combinatorial background. This will facilitate obtaining the signal-to-background ratio. First, the variables used from the measurements are elaborated; after, multiple fit functions will be discussed.

4.3.1 Parameters

Various parameters of the measured photons were used in this analysis. The most important ones include the rapidity η (see equation 20) and the azimuthal angle ϕ, which is, like η, taken from the data set. The inner radius of the detector R_{EMC} is taken to be 231.23 cm.

The position and momentum variables are defined as follows:

$$x = R_{\mathrm{EMC}} \cdot \cos\phi \tag{21}$$

$$y = R_{\mathrm{EMC}} \cdot \sin\phi \tag{22}$$

$$z = R_{\mathrm{EMC}} \cdot \sinh\eta - z_{vert} \tag{23}$$

$$p_x = E\frac{x}{R} \tag{24}$$

$$p_y = E\frac{y}{R} \tag{25}$$

$$p_z = E\frac{z}{R} \tag{26}$$

$$p_{\mathrm{T}} = \sqrt{p_x^2 + p_y^2} \tag{27}$$

where $R = \sqrt{x^2 + y^2 + z^2}$, z_{vert} is measured by the detector, and E is the energy of the measured photon.

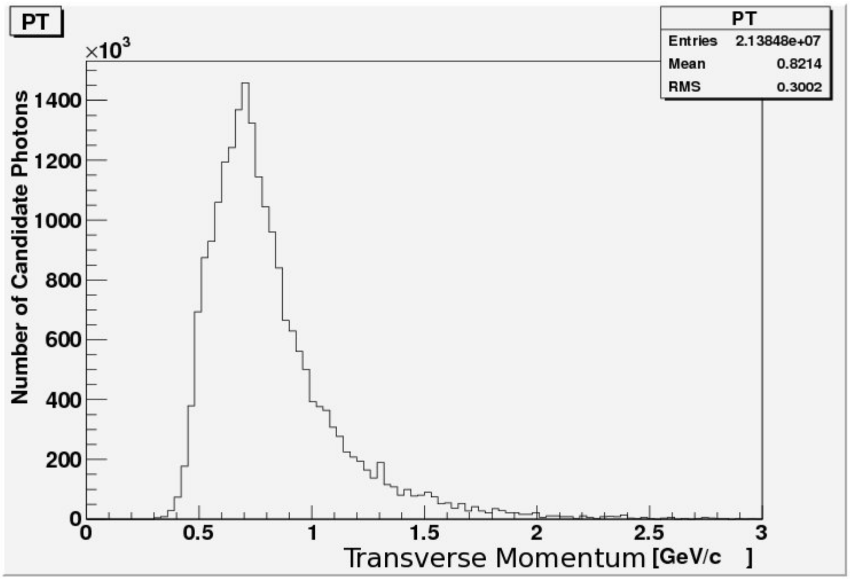


Figure 9: The transverse momentum distribution of individual photons.

With these variables, graphs can be made using single photons as follows. In figure 9, the transverse momentum is plotted of only photons. In figure 10, the energy distribution of these photons is shown. The sharp rise in the distribution near 0.7 GeV is a result of the cluster finding algorithm that did not count low energies.

For the reconstructed pions, the transverse momentum is cut into several intervals. In these intervals, the invariant mass distributions are fitted to study the signal-to-background ratio. The p_T distribution in graph 11 shows that reasonable intervals would be 0-1, 1-1.5, 1.5-2, and 2-∞ GeV/c. The lower p_{T} intervals are not split into even smaller intervals; this is because they will, as can be seen later in this analysis, yield a lower signal-to-background ratio. In addition, the measurements of lower transverse momentum are not likely to come from neutral pions and measurements of energies below 700 MeV are not considered reliable.

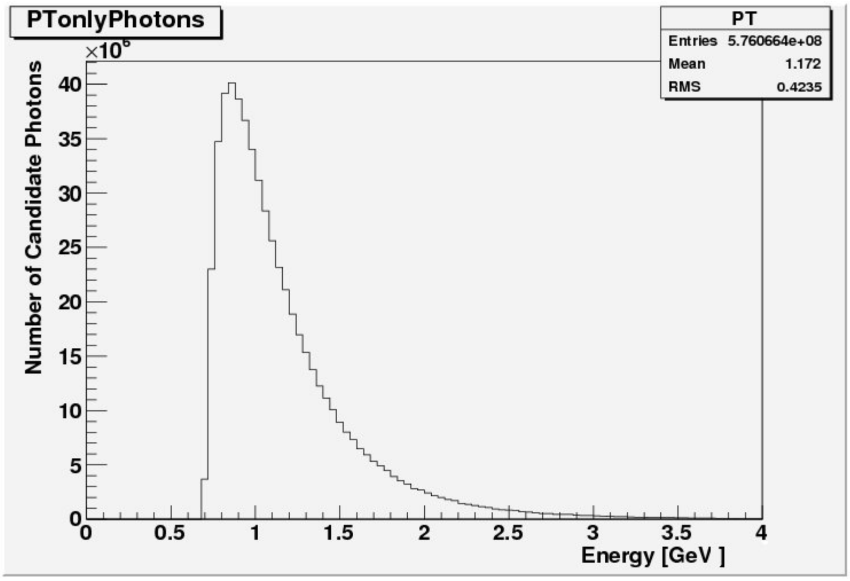


Figure 10: The energy distribution of individual photons. The sharp rise at 0.7 MeV is a result of excluding low energies in the cluster finding algorithm.

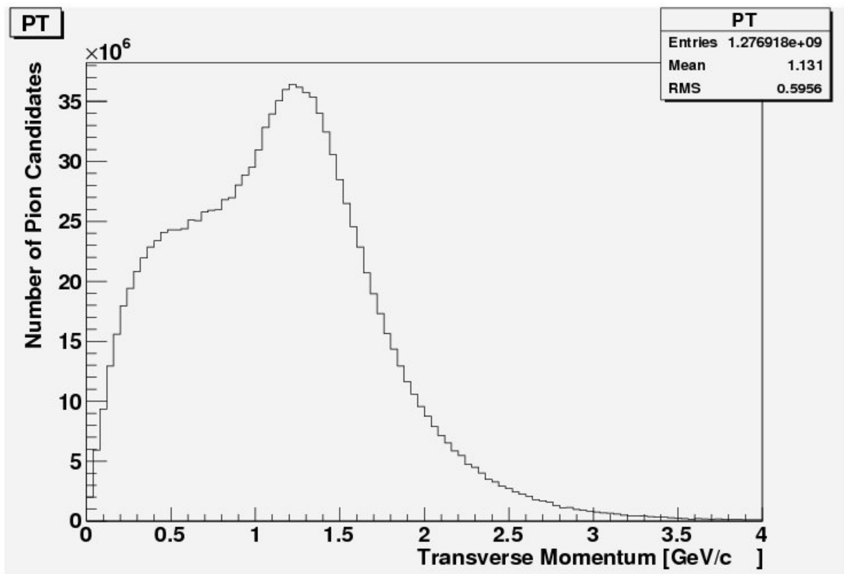


Figure 11: The transverse momentum distribution of candidate pions.

4.3.2 Fitting Procedure

Functions matching the data distribution can help estimate the combinatorial background in the invariant mass distributions. With integrals of these functions, an estimate of the signal-to-background ratio can obtained.

In order to fit a function, the Fit() class in ROOT is used [4].[4] This function uses default (1) or user input start parameters to find the best parameters to fit a given input function on a given histogram. After fitting, a covariance matrix, the number of degrees of freedom, parameters and parameter errors, χ^2, and other properties can be requested from the input function. The verb 'fit' used below indicates the use of the function Fit().[5]

The procedure to find a good fit for the background function is carried out as follows: Two separate functions are chosen for the signal and combinatorial background. These are fitted separately; a Gaussian function is fitted on the peak, and the background function is fitted on a histogram excluding the peak region. The boundaries of this peak interval are determined using the width (σ) and mean (invariant mass) parameters of the Gaussian, which was first fitted on the peak. The peak region X_{peak} is then

$$(m - 3\sigma) < X_{\text{peak}} < (m + 3\sigma), \tag{28}$$

with m the mean of the Gaussian.

At this point in the process, two functions have been fitted: A Gaussian on the peak, and another function on the background. The Gaussian function $g(x)$ is defined as:

$$g(x) = c \cdot \mathrm{e}^{-\frac{(x-m)^2}{2\sigma^2}}, \tag{29}$$

where c, m, and σ are the parameters of the Gaussian: a constant, the mean (in this case the average candidate pion mass), and the width of the peak, respectively.

Now, the background fit will be improved using a third function, which is the sum of the first two functions. This function, here called the global fit function, will be fit onto the entire histogram using the parameters from the first two fits:

$$f_{globalfit}(x) = g(x) + f_{background}(x). \tag{30}$$

Subsequently, the parameters from the global fit function are used to draw (not fit) the background function again; it is this function that is used to estimate the background and the signal-to-background ratio. The

[4]"Fitting Demo" for fitting signal and background of Rene Brun has been used to create fit programs for ROOT in C++. See http://root.cern.ch/root/html/tutorials/fit/FittingDemo.C.html (01/04/09).

[5]See also http://root.cern.ch/root/html/TH1.html#TH1:Fit .

Gaussian function for the peak can be drawn with the new parameters to see how the global fit worked, and whether there were any significant errors in this fit.

Background functions are fitted on the graph on an interval from 0 to 0.7 GeV/c^2. Any greater values are insignificant for this analysis as particles with a mass larger than 0.25 GeV/c^2 are not neutral pions. For a third order polynomial fit, however, a larger interval reduced χ^2. The difference between the third order polynomial fits using different intervals can be seen in figure 12.

Fits were tried with various functions, such as first-, second-, and third-order polynomials (where p_i indicate the parameters):

$$
\begin{aligned}
f_{\mathrm{pol1}}(x) &= \mathrm{p}_0 + \mathrm{p}_1 x && (31)\\
f_{\mathrm{pol2}}(x) &= \mathrm{p}_0 + \mathrm{p}_1 x + \mathrm{p}_2 x^2 && (32)\\
f_{\mathrm{pol3}}(x) &= \mathrm{p}_0 + \mathrm{p}_1 x + \mathrm{p}_2 x^2 + \mathrm{p}_3 x^3, && (33)
\end{aligned}
$$

a square-root function:

$$f_{\mathrm{sqrt}} = \mathrm{p}_0 + \mathrm{p}_1 \sqrt{x}, \qquad (34)$$

and first- and second-order logarithmic functions:

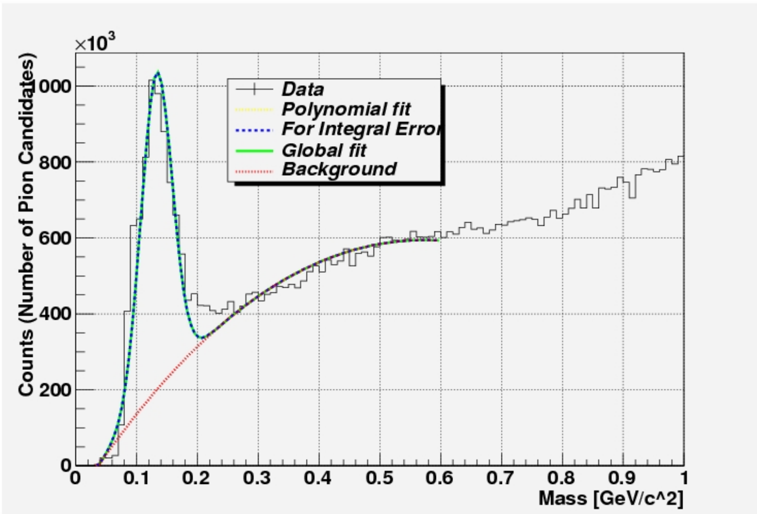
$$
\begin{aligned}
f_{\mathrm{log1}} &= \mathrm{p}_0 + \mathrm{p}_1 \log(x - \mathrm{p}_2) && (35)\\
f_{\mathrm{log2}} &= \mathrm{p}_0 + \mathrm{p}_1 \log(x - \mathrm{p}_2) + \mathrm{p}_3 \log^2(x - \mathrm{p}_2). && (36)
\end{aligned}
$$

From the graphs in figures 13, 14 and 15, it is evident that f_{sqrt}, f_{pol2}, and f_{log1} do not yield good fits. Second order logarithmic and third order polynomial functions proved to fit the combinatorial background better, as can be seen in figures 16 and 17. As a result, they have been used to study the signal-to-background ratio of the mass distributions using various cuts.

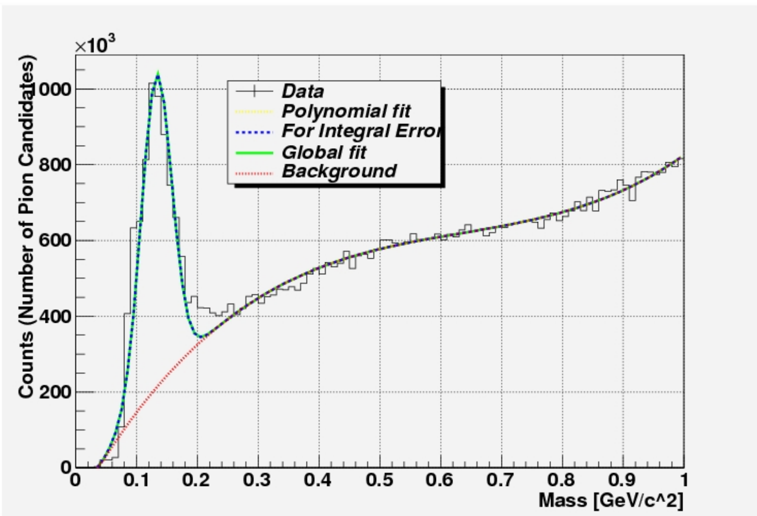
4.4 The Signal-To-Background Ratio

The signal-to-background ratio, or the ratio of the 'true' neutral pions to the 'false' combinations of photons, will be explained further in this section. This ratio is defined as follows:

$$\frac{\text{signal}}{\text{background}} = \frac{S}{B} = \frac{T - B}{B} = \frac{T}{B} - 1 \equiv R, \qquad (37)$$



(a) Fitting the combinatorial background with a third order polynomial using a shorter interval of 0 to 0.6 GeV/c^2.



(b) Fitting the combinatorial background with a third order polynomial using a longer interval of 0 to 1 GeV/c^2.

Figure 12: The different fits of a third order polynomial function yielded by intervals of two different lengths. The function “For Integral Error” is used later to find the error on the integral of the background function (see section 4.5.1).

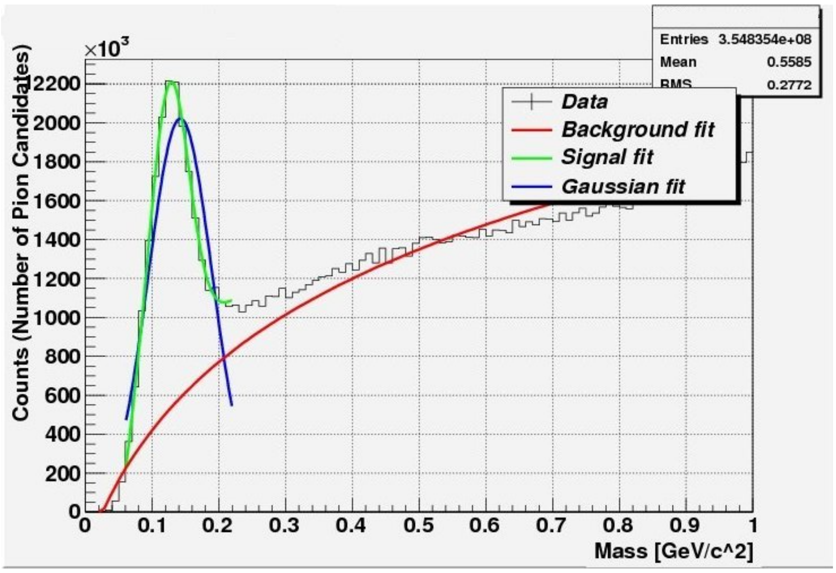


Figure 13: A histogram of reconstructed pion masses fitted with a first order logarithmic function.

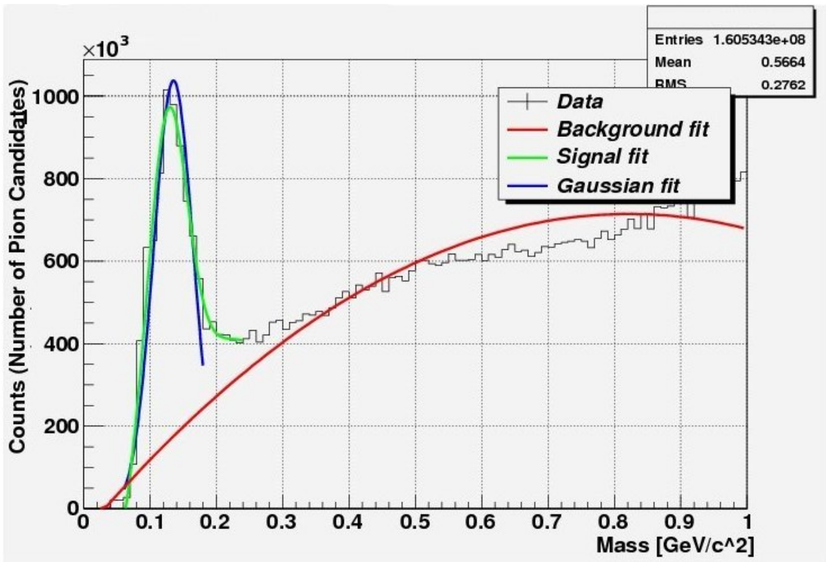


Figure 14: A histogram of reconstructed pion masses fitted with a second order polynomial function.

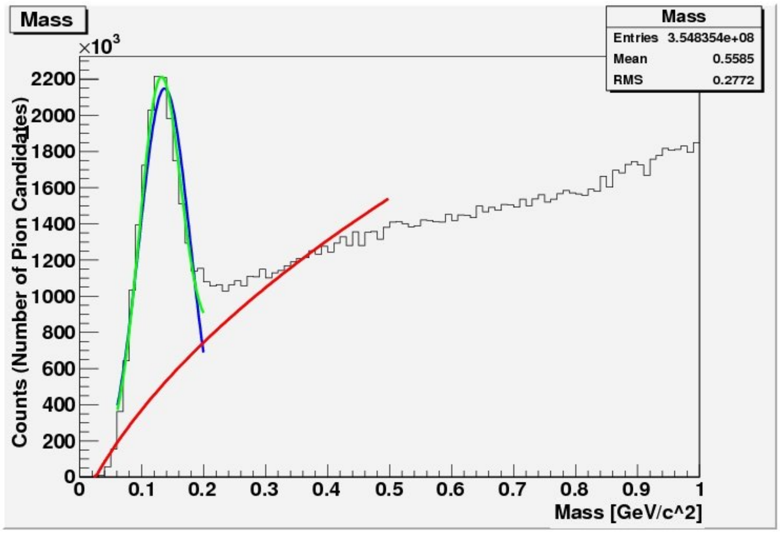


Figure 15: A histogram of reconstructed pion masses fitted with a square root function.

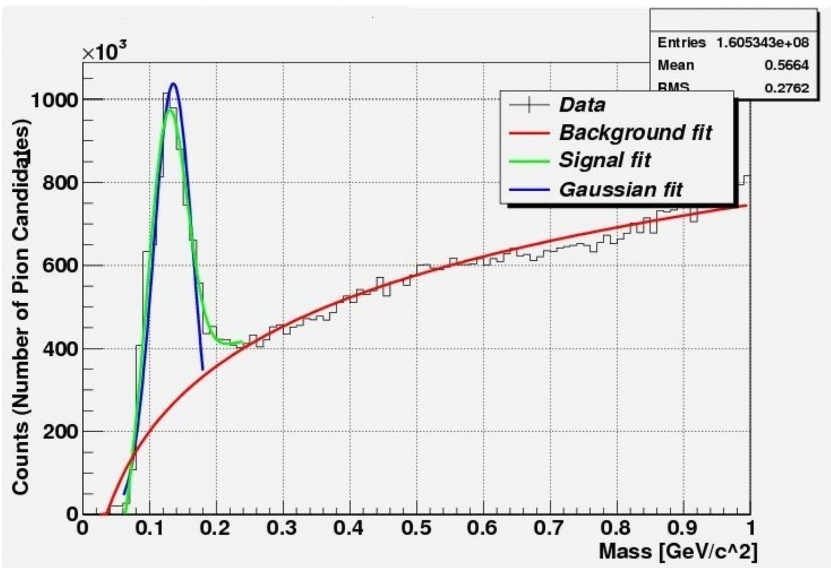


Figure 16: A histogram of reconstructed pion masses fitted with a second order logarithmic function.

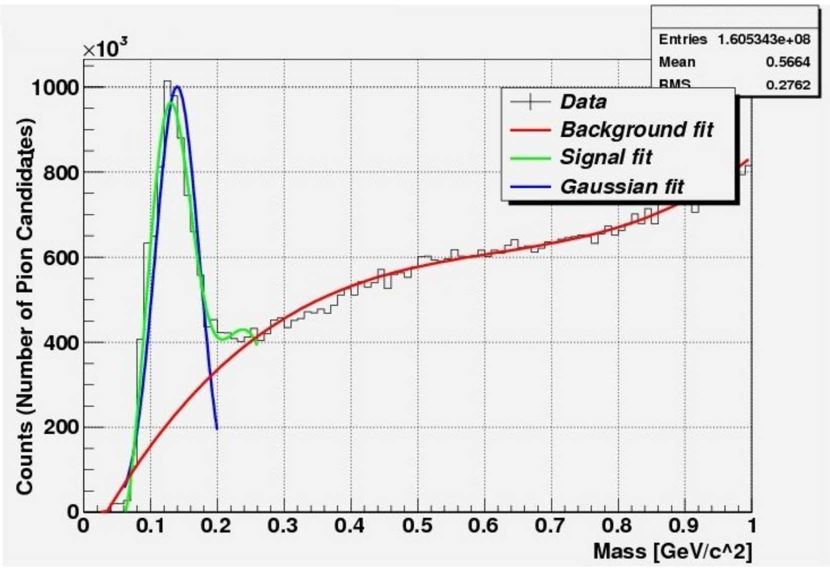


Figure 17: A histogram of reconstructed pion masses fitted with a third order polynomial function.

where S is the signal, and the total area in the peak region T and combinatorial background in the peak region B are defined as:

$$T = \sum_{i=b_{\text{low}}}^{b_{\text{high}}} C_i \tag{38}$$

$$B = \int_{m_{\text{pion}}-3\sigma}^{m_{\text{pion}}+3\sigma} f_{\text{background}}(x)dx, \tag{39}$$

where $f_{\text{background}}(x)$ is the function fitted on the combinatorial background, C_i are the contents of the i^{th} bin of the mass histogram, and b_{high} and b_{low} are the upper and lower bins located at the end points of the peak region (a factor 100 that is necessary to scale B to the same dimensions as T has been left out here for simplicity). The total of the bin contents T consists of the contents of the bins in the peak region; the background B is the integral taken (in the same region) over the function that fits the combinatorial background using the new parameters from the global fit. Here, the assumption is made that the combinatorial background in the region of the peak interval behaves more or less similar to the fitted background function. Now, the signal-to-background ratio is

$$R = \frac{\sum_{i=b_{\text{low}}}^{b_{\text{high}}} C_i}{\int_{m_{\text{pion}}-3\sigma}^{m_{\text{pion}}+3\sigma} f_{\text{background}}(x)dx} - 1 \tag{40}$$

An example of a fitted graph can be seen in figure 18. The combinatorial

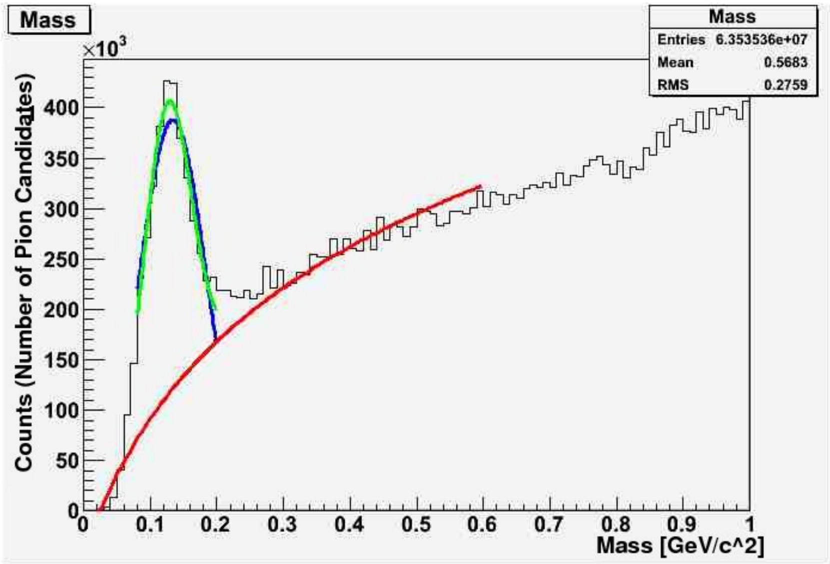


Figure 18: A histogram of reconstructed pion masses fitted with a second order logarithmic function. The signal-to-background ratio is 1.5.

background is fitted with a second order logarithmic function, and the peak with a Gaussian; the signal-to-background ratio is 1.5.

4.5 Statistical and Systematic Uncertainties

There are two kinds of errors that must be taken into account: random errors and systematic errors [11]. Random errors, or statistical errors, can be solved by repeating measurements. Using statistics, it is possible to get a reliable estimate of this kind of error. The program of ROOT can be used to obtain most of the statistical errors. Systematic errors must be estimated using other methods.

4.5.1 Statistical Errors

ROOT can give the statistical errors of the bins of a histogram. From these bin errors, ROOT can also calculate the parameter errors of fitted functions; these depend on the bin errors. Using the parameter errors, it is also possible to obtain the error of the integral of the background function. One can do this by calculating the integral and then solve for a parameter. An example is a first order polynomial:

$$
\begin{aligned}
f(x) &= p_0 + p_1 x & (41)\\
p_0' \equiv \int_a^b f(x)dx &= p_0(b-a) + \frac{p_1}{2}(b^2 - a^2) & (42)\\
p_0 &= \frac{p_0'}{(b-a)} - \frac{p_1}{2}(b+a), & (43)
\end{aligned}
$$

where b and a are the end points of the peak interval at $m \pm 3\sigma$. Now, the equation for the new parameter p_0 can be substituted into the original function $f(x)$:

$$
f(x) = \frac{p_0'}{b-a} - \frac{p_1}{2}(b+a) + p_1(x), \qquad (44)
$$

where the integral will be treated as a parameter. The background function $f(x)$ can be fitted again as part of the global fit function, whereafter the error of the parameter p_0' can be obtained from ROOT, and thus the error of the integral will be given. This procedure of calculating the integral error has been used for the third order polynomial and second order logarithmic functions that are utilized in this analysis. For the third order polynomial function:

$$
\begin{aligned}
f(x) &= p_0 + p_1 x + p_2 x^2 + p_3 x^3\\
p_0' \equiv \int_a^b f(x)dx &= p_0(b-a) + \frac{p_1}{2}(b^2 - a^2) + \frac{p_2}{3}(b^3 - a^3) +\\
&\quad + \frac{p_3}{4}(b^4 - a^4)\\
p_0 &= \frac{p_0'}{(b-a)} - \frac{p_1}{2}(b+a) - \frac{p_2}{3}(b^2 + ab + a^2) -\\
&\quad - \frac{p_3}{4}(b+a)(b^2+a^2)\\
f(x) &= \frac{p_0'}{(b-a)} + p_1(x - \frac{(b+a)}{2}) +\\
&\quad + p_2(x^2 - \frac{(b^2+ab+a^2)}{3}) +\\
&\quad + p_3(x^3 - \frac{(b+a)(b^2+a^2)}{4}).
\end{aligned}
$$

The second order logarithmic function is more complicated. The integrals of a first and second order logarithmic function are defined as follows:

$$
\begin{aligned}
\int \log x dx &= x \log x - x\\
\int (\log x)^2 dx &= x(\log x)^2 - 2x \log x + 2x,
\end{aligned}
$$

so that the integral of the logarithmic function is

$$
\begin{aligned}
f(x) &= p_0 + p_1 \log(x - p_2) + p_3(\log(x - p_2))^2 \\
\int_a^b p_1 \log(x - p_2)dx &= p_1[(-b + a - (a - p_2)\log(a - p_2) + (b - p_2)\log(b - p_2)] \\
\int_a^b p_3(\log(x - p_2))^2 dx &= p_3[2(b - a) - 2(b - p_2)\log(b - p_2) + \\
&\quad +2(a - p_2)\log(a - p_2) + (b - p_2)(\log(b - p_2))^2 - \\
&\quad -(a - p_2)(\log(a - p_2))^2)] \\
p_0' \equiv \int_a^b f(x)dx &= p_0(b - a) + p_1[-b + a - (a - p_2)\log(a - p_2) + \\
&\quad +(b - p_2)\log(b - p_2)] + \\
&\quad +p_3[2(b - a) - 2(b - p_2)\log(b - p_2) + \\
&\quad +2(a - p_2)\log(a - p_2) + (b - p_2)(\log(b - p_2))^2 - \\
&\quad -(a - p_2)(\log(a - p_2))^2].
\end{aligned}
$$

This gives the following when solving for p_0:

$$
\begin{aligned}
p_0 = & \frac{p_0'}{(b-a)} + p_1[-\frac{1}{b-a}(b - a + (a - p_2)\log(a - p_2) - \\
& -(b - p_2)\log(b - p_2))] + \\
& +p_3[-\frac{1}{b-a}(-2(b - a) + 2(b - p_2)\log(b - p_2) - \\
& -2(a - p_2)\log(a - p_2) - (b - p_2)(\log(b - p_2))^2 + \\
& +(a - p_2)(\log(a - p_2))^2)],
\end{aligned}
$$

and, when substituting the function p_0 into $f(x)$:

$$
\begin{aligned}
f(x) = & \frac{p_0'}{(b-a)} + p_1[\log(x - p_2) - \frac{1}{b-a}(b - a + \\
& +(a - p_2)\log(a - p_2) - (b - p_2)\log(b - p_2))] + \\
& +p_3[(\log(x - p_2))^2 - \frac{1}{b-a}(-2(b - a) + \\
& +2(b - p_2)\log(b - p_2) - 2(a - p_2)\log(a - p_2) - \\
& -(b - p_2)(\log(b - p_2))^2 + (a - p_2)(\log(a - p_2))^2)].
\end{aligned}
$$

The absolute error for the signal-to-background ratio is defined as:

$$\sigma_R = \sigma_{\frac{T}{B}}. \tag{45}$$

The relative error squared on $\frac{T}{B}$ is

$$\left(\frac{\sigma_{\frac{T}{B}}}{\frac{T}{B}}\right)^2 = \left(\frac{\sigma_T}{T}\right)^2 + \left(\frac{\sigma_B}{B}\right)^2, \tag{46}$$

and the statistical error on the signal-to-background ratio σ_R is

$$\sigma_R = \frac{T}{B}\sqrt{\left(\frac{\sigma_T}{T}\right)^2 + \left(\frac{\sigma_B}{B}\right)^2}, \tag{47}$$

where σ_T is

$$\sigma_T = \sqrt{T}, \tag{48}$$

the square root of the bin contents, and σ_B is obtained from ROOT by requesting the error on the integral parameter (see procedure above). The error on the signal is given by a quadratic sum [11]:

$$\sigma_S = \sqrt{\sigma_T^2 + \sigma_B^2}. \tag{49}$$

4.5.2 Systematic Errors

Systematic errors are errors of missing knowledge and where assumptions are made. Systematic errors cannot be revealed by statistics [11]. A systematic error in this analysis could originate from the use of a particular fit function that is actually not suitable, and gives rise to errors not shown by statistics. For this analysis, the deviation of the average of the two different fit functions could be calculated to get an idea of the systematic error of using these functions. The plots are fitted with a second order logarithmic and a third order polynomial function. The ideal combinatorial background is then taken to be the average, and the error will be the difference of the mean with the polynomial and logarithmic functions. For example, in the case of a transverse momentum between 1 and 1.5 GeV/c, the integral in the peak region of a second order logarithmic fit function is 3141630, while the integral in the peak region of a third order polynomial fit function is 3038080. The average would than be

$$\frac{\mathrm{B_{log2}} + \mathrm{B_{pol3}}}{2} = \frac{3141630 + 3038080}{2} = 3089855,$$

and an estimate of the systematic error of this average combinatorial background is then

$$\frac{\mathrm{B_{log2}} - \mathrm{B_{pol3}}}{2} = \frac{3141630 - 3038080}{2} = 51775.$$

As a result, the systematic error in this trial is about $(51814/3089855 =)$ $0.016769 \approx 1.7$ %. Using table 3 on page 31, similar calculations can be performed to find that the error for 1.5 GeV/c $< p_T <$ 2 GeV/c is

$$\frac{4650}{2300850} = 0.00202 \approx 0.2\%,$$

and for $p_T > 2$ GeV/c

$$\frac{3900}{452600} = 0.00862 \approx 0.9\%.$$

We can then assume the systematic error due to the function choice to be around

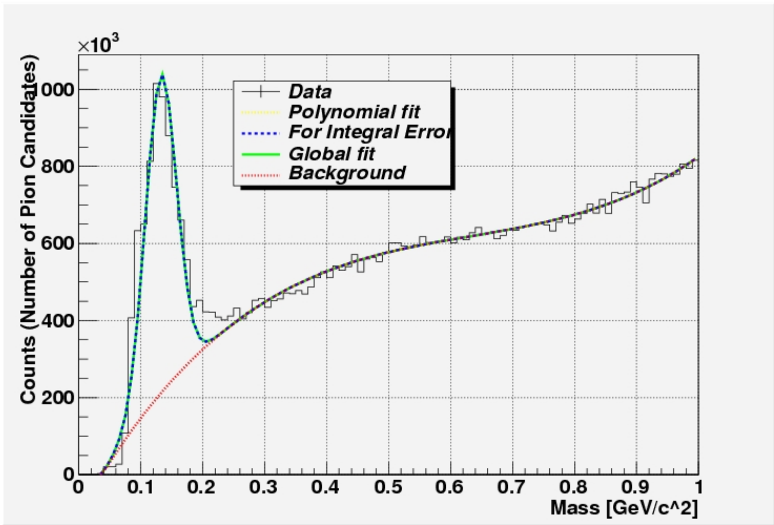
$$\frac{1.7 + 0.2 + 0.9}{3} \approx 0.9\%.$$

One assumption made in this analysis was that the size of the peak interval that was used in calculating the signal-to-background ratio was the interval of the mean $\pm 3\sigma$. In order to check the systematic error on the use of the length of 3σ, the peak interval was changed using values of 2.5σ and 3.5σ; these, however, changed little. See figure 19 for these trials. As a comparison, the same background function is given in figure 20 using the interval of 3σ assumed in this analysis. From these figures, it is visible that the systematic error of the width of the peak interval is negligible.

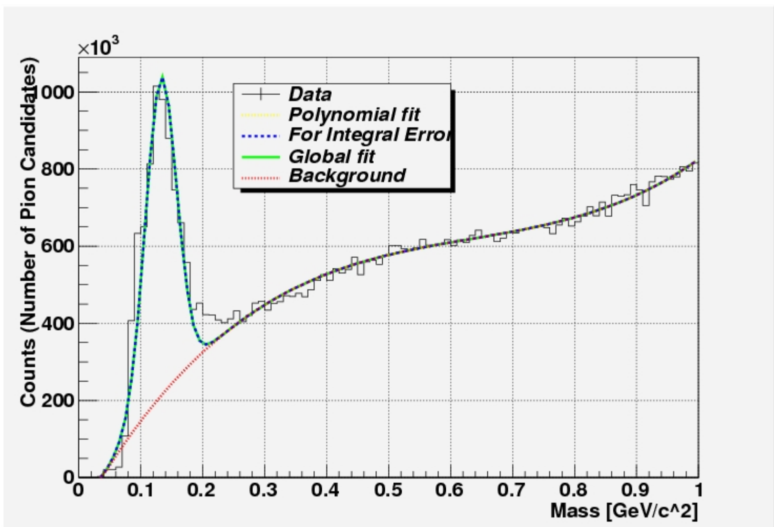
Another assumption made is that the shape of the combinatorial background in the peak region behaves like the function fitted on the rest of the combinatorial background. This behavior in the peak region is largely due to the shape of the detector. If there were only two small detectors, separated by a large distance from one another, then only very small opening angles (two photons leaving in nearly similar directions, towards one detector) and very large opening angles (the angle between two separating photons very large, so that they go in opposite directions to different detectors) will yield a peak. Then two peaks will appear around these masses (smaller for the small opening angle, larger for the large opening angle) in a histogram. The shape of a combinatorial background, then, varies with different detectors and with different particles. The BEMC detector is fairly smoothly shaped, and therefore, the combinatorial background can be considered smooth as well. One way to verify the assumption of the shape of this combinatorial background is event mixing, in which photons from different events are combined [6]. From this, a similarly shaped combinatorial background arises, which asserts our assumption of the combinatorial background shape.

5 Results

A summary of the fit results is given in tables below. These data are taken from graphs in which the third order polynomial and second order polynomial are fitted onto invariant mass distributions of reconstructed pions of different intervals of transverse momenta: p_T from 1 to 1.5 GeV/c (figure 21), p_T from 1.5 to 2 GeV/c (figure 22), p_T from 2 GeV/c and up (figure



(a) A third order polynomial fit on the combinatorial background using $\pm 3.5\sigma$



(b) A third order polynomial fit on the combinatorial background using $\pm 2.5\sigma$

Figure 19: Fits using different peak intervals to check the systematic error of the choice of the peak interval. The difference is very small, so this error is negligible.

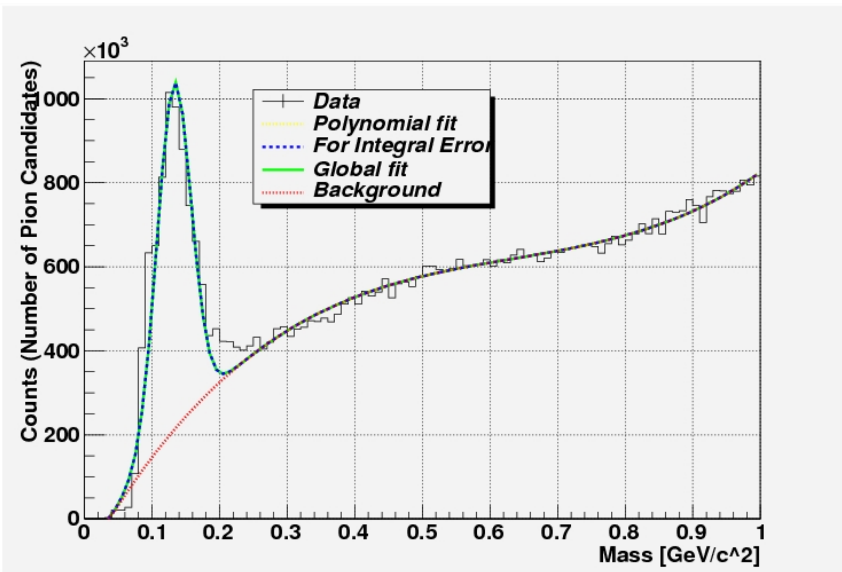


Figure 20: A third order polynomial fit on the combinatorial background using $\pm 3\sigma$.

23). From the results in the tables, new plots were made in which the mean parameter of the Gaussian (or mass, see figure 24), σ (width of the peak, see figure 25), and the signal-to-background ratio (see figure 26) were plotted against different intervals of transverse momentum.

5.1 The Mean and Standard Deviation

In table 1, the mean, or average mass of the reconstructed neutral pions, σ, χ^2, and number of degrees of freedom are given for different fit functions and different intervals of transverse momentum. The mean increases as the transverse momentum is increased (see figure 24). With ideal measurements, one would expect the mean to be constant and equal to the mass of a neutral pion, which is 134.9766 ± 0.0006 MeV [5]. This does not apply in this analysis. It is possibly a result of the type of cluster finding algorithm; however, in order to give a reliable explanation of this non-linearity in the measurements, more research is necessary.

The standard deviation, or σ, decreases as the transverse momentum is increased (see figure 25). This is as expected; as the transverse momentum increases, the energy increases, and the measurements of the calorimeter get a better resolution. The calorimeter has a better energy resolution for higher energies, as shown below.

The invariant mass of neutral pion can be written as a function of the energies and the opening angle of two candidate photons [12]:

$$m = \sqrt{2\mathrm{E}_1\mathrm{E}_2(1 - \cos\psi)}.$$

For the case when the two energies are approximately equal ($E_1 \approx E_2 \approx E$) one can easily estimate the square of the standard deviation of the mass as follows:

$$\sigma_m^2 = \sigma_\mathrm{E}^2 \left|\frac{\partial m}{\partial \mathrm{E}}\right|^2 + \sigma_\psi^2 \left|\frac{\partial m}{\partial \psi}\right|^2.$$

This approximation is reasonable because of the maximum cut on the energy asymmetry and the lower energy thresholds used in the cluster algorithm. For the relatively low p_T values studied here one can also assume that the contribution from the energy resolution is dominant, because the opening is large and we can neglect the relatively small uncertainties in the angle measurement. Now, ignoring the term of the opening angle and taking the energies to be the same,

$$\begin{aligned}
\sigma_m^2 &= \sigma_{\mathrm{E}}^2 \left(\frac{\partial m}{\partial \mathrm{E}}\right)^2 \\
&= \sigma_{\mathrm{E}}^2 \left(\frac{\partial \sqrt{2\mathrm{E}^2(1-\cos\psi)}}{\partial \mathrm{E}}\right)^2 \\
&= \sigma_{\mathrm{E}}^2 \left(\frac{\partial \left(\mathrm{E}\sqrt{2(1-\cos\psi)}\right)}{\partial \mathrm{E}}\right)^2 \\
&= \sigma_{\mathrm{E}}^2 \left(\frac{m}{\mathrm{E}}\right)^2 .
\end{aligned}$$

The standard deviation of the mass is then proportional to the standard deviation of the energy:

$$\left(\frac{\sigma_m}{m}\right)^2 = \left(\frac{\sigma_{\mathrm{E}}}{\mathrm{E}}\right)^2 .$$

The standard deviation of the energy is a known property of the detector: it is the so-called energy resolution of the calorimeter, which is estimated to be [6]

$$\frac{\sigma_{\mathrm{E}}}{\mathrm{E}} = \sqrt{\left(\frac{15\%}{\sqrt{\mathrm{E[GeV]}}}\right)^2 + (1.5\%)^2},$$

where the energy goes as a Poisson distribution due to energy fluctuations in the detector, and the percentages are constants (the second constant is one independent of the energy). It can now be seen that for higher energies, the calorimeter has a better energy resolution and the standard deviation of the mass (σ) decreases.

5.2 The Signal-To-Background Ratio

In table 3 the areas of the signal and the background on the graphs are given with their errors. The signal is calculated using the total contents of the bins in the peak region that are given in table 2. In table 4, the signal-to-background ratio and the significance are given. The significance α is defined as the ratio of the signal to the square root of the signal plus background:

$$\alpha = \frac{\mathrm{T}-\mathrm{B}}{\sqrt{\mathrm{T}}} = \frac{\mathrm{S}}{\sqrt{\mathrm{S}+\mathrm{B}}}. \tag{50}$$

The peak of the signal is clearly distinguishable on the graphs (see figures 21, 22, and 23), so the significance does not yield important information in this case.

fit method	p_T [GeV/c]	m [MeV/c^2]	σ [MeV/c^2]	χ^2	**NDF**
log 2nd order	1 - 1.5	124	30	3996.8	51
pol 3rd order	1 - 1.5	125	31	2486.42	91
log 2nd order	1.5 - 2	130	25	8173.34	50
pol 3rd order	1.5 - 2	130	25	4590.62	90
log 2nd order	> 2	142	22	3443.73	49
pol 3rd order	> 2	142	22	1954.78	89

Table 1: Summary of fit results using different fit functions and p_T intervals. The errors of the mean and sigma parameters are of the order 10^{-5} GeV/c^2 and thus negligible compared to the systematic error.

p_T [GeV/c]	T [GeV/c^2]
1 - 1.5	$(5483.4 \pm 2.3) \times 10^3$
1.5 - 2	$(5143.0 \pm 2.3) \times 10^3$
> 2	$(1366.2 \pm 1.2) \times 10^3$

Table 2: The bin contents in the peak region for different p_T intervals.

The signal-to-background ratio increases as the transverse momentum is increased (see figure 26). For higher transverse momentum, and thus higher energies, less particles are available and, as a consequence, less (false) combinations of photons can be made. Also, for lower energies more clusters are measured that are not actually photons. As a result, higher energies yield fewer false combinations and a more distinguished peak.

6 Discussion

With the plot of the varying signal-to-background ratios, a more educated estimate of the authenticity of the reconstructed neutral pions can be made. Further analysis can then be performed on the original collisions. Some uncertainties, however, must be taken into account and could be investigated more thoroughly.

To verify the assumption of the shape of the combinatorial background, the earlier described procedure of event mixing could be used. In this procedure, candidate photons from different events are combined. As these surely did not originate from the same pion, these combinations will yield a combinatorial background which serves as a test of our assumption of the smoothly shaped background.

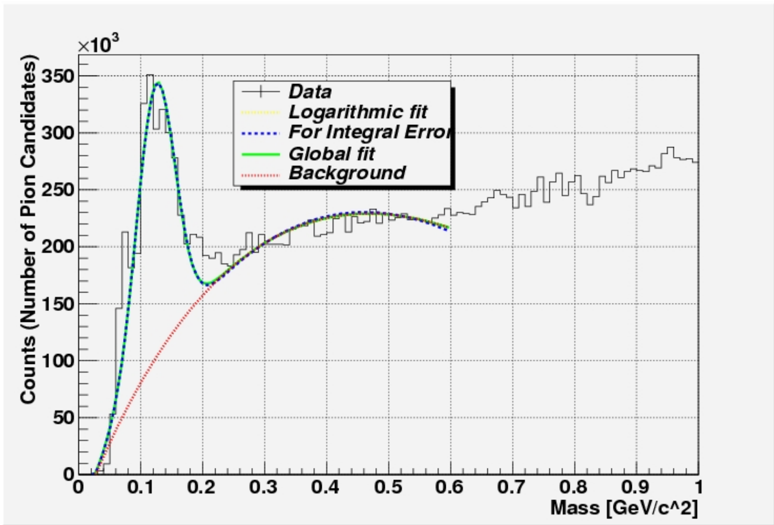
In the invariant mass distributions of the transverse momentum intervals 1 to 1.5 and 1.5 to 2 GeV/c, a second peak arises next to the peak around the

fit method	p_T [GeV/c]	S(= T − B) [GeV/c^2]	B [GeV/c^2]
log 2nd order	1 - 1.5	$(2341.8 \pm 3.1) \times 10^3$	$(3141.6 \pm 2.0) \times 10^3$
pol 3rd order	1 - 1.5	$(2445.3 \pm 3.0) \times 10^3$	$(3038.1 \pm 1.9) \times 10^3$
log 2nd order	1.5 - 2	$(2837.5 \pm 2.7) \times 10^3$	$(2305.5 \pm 1.4) \times 10^3$
pol 3rd order	1.5 - 2	$(2846.8 \pm 2.7) \times 10^3$	$(2296.2 \pm 1.4) \times 10^3$
log 2nd order	> 2	$(917.5 \pm 1.3) \times 10^3$	$(448.7 \pm 0.6) \times 10^3$
pol 3rd order	> 2	$(909.7 \pm 1.3) \times 10^3$	$(456.5 \pm 0.6) \times 10^3$

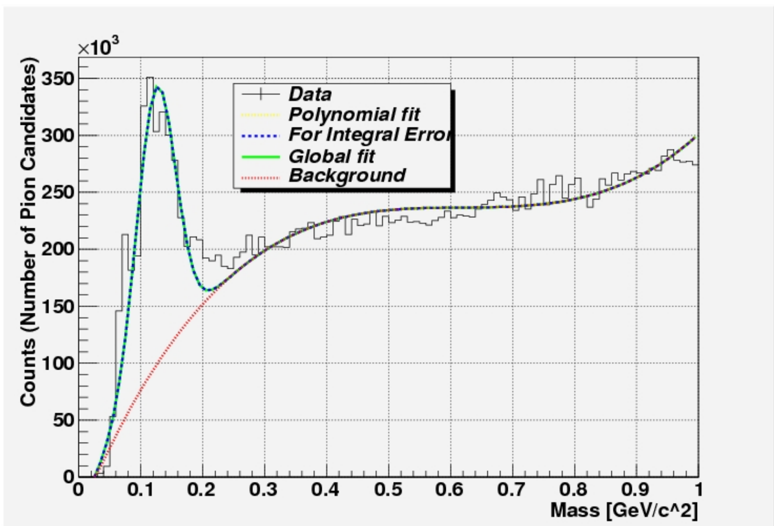
Table 3: The signal and background areas for different fit functions and p_T intervals. The area of the signal on the graph is calculated using the integral of the background function and bin contents of the histogram in the peak region.

fit method	p_T [GeV/c]	R	α
log 2nd order	1 - 1.5	0.745404 ± 0.00058	1000.05
pol 3rd order	1 - 1.5	0.804897 ± 0.00061	1044.27
log 2nd order	1.5 - 2	1.23074 ± 0.00092	1251.19
pol 3rd order	1.5 - 2	1.23979 ± 0.00093	1255.3
log 2nd order	> 2	2.04489 ± 0.00332	784.975
pol 3rd order	> 2	1.99295 ± 0.00295	778.314

Table 4: The signal-to-background ratio and the significance for different fit functions and p_T intervals. The error of the significance is not given, since each graph has a clearly distinguishable peak (see figures 21, 22, and 23).

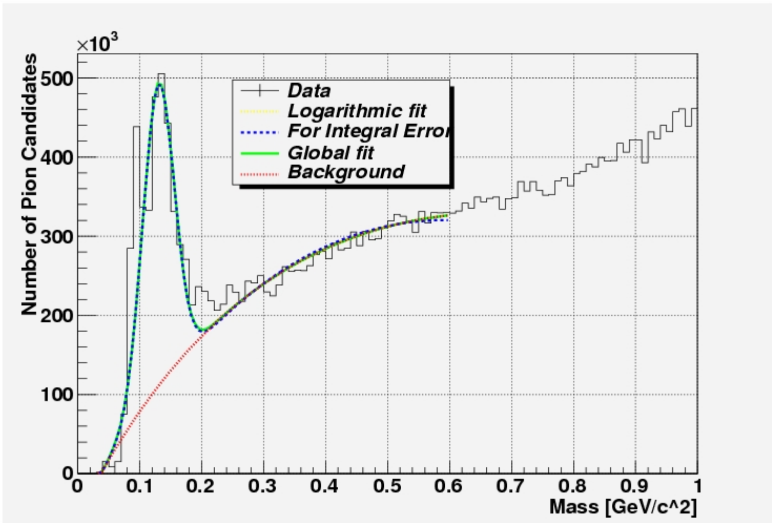


(a)

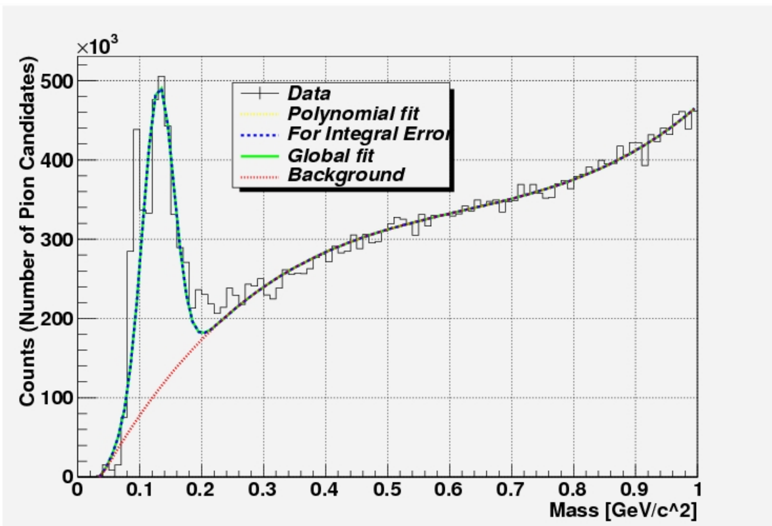


(b)

Figure 21: A mass distribution of particles of transverse momentum 1 to 1.5 GeV/c fitted with (a) a second order logarithmic function and (b) with third order polynomial function.

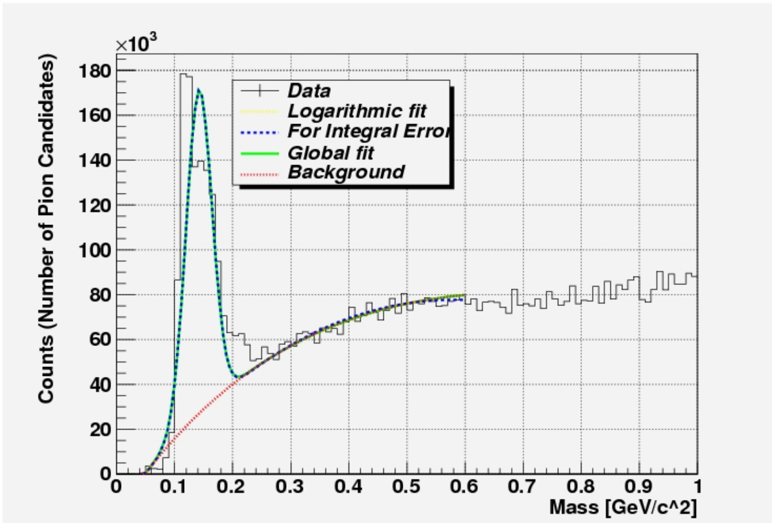


(a)

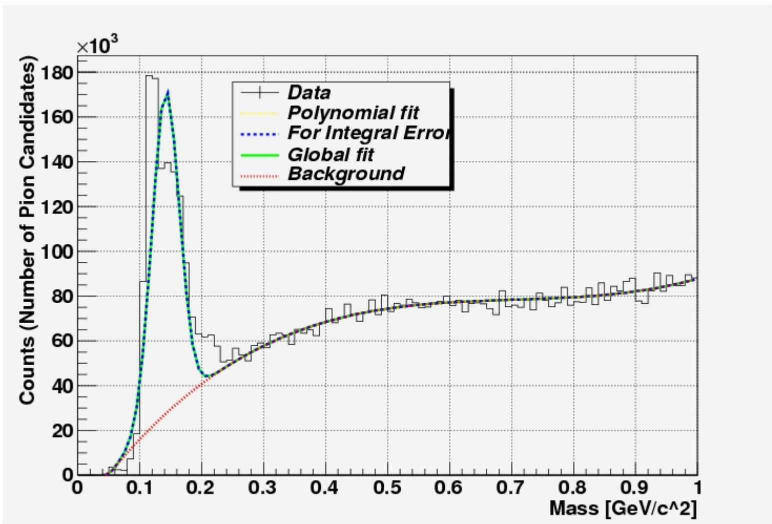


(b)

Figure 22: A mass distribution of particles of transverse momentum 1.5 to 2 GeV/c fitted with (a) a second order logarithmic function and (b) with third order polynomial function.

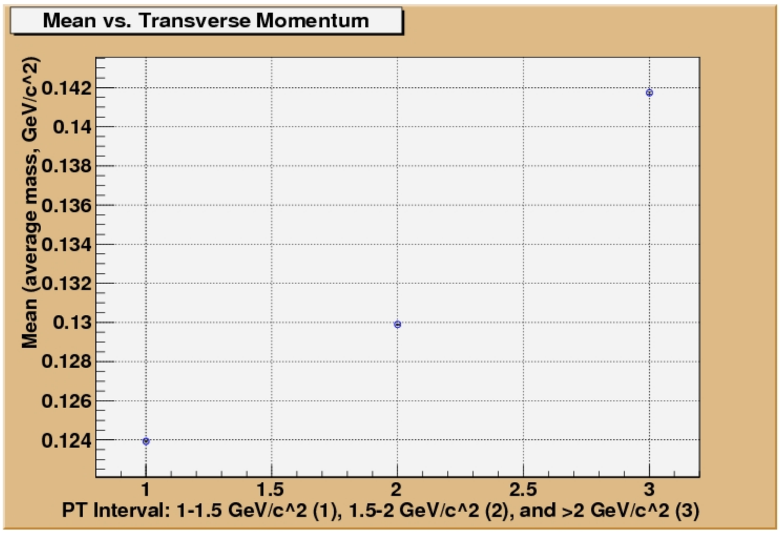


(a)

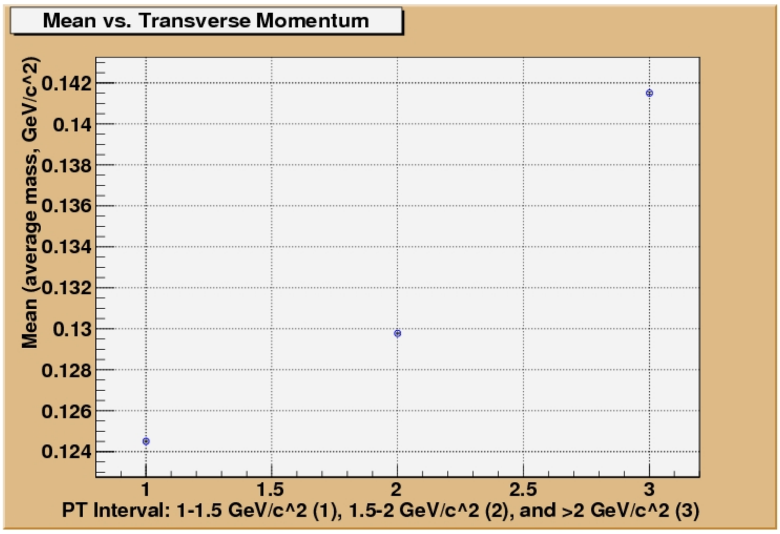


(b)

Figure 23: A mass distribution of particles of transverse momentum greater than or equal to 2 GeV/c fitted with (a) a second order logarithmic function and (b) with third order polynomial function.

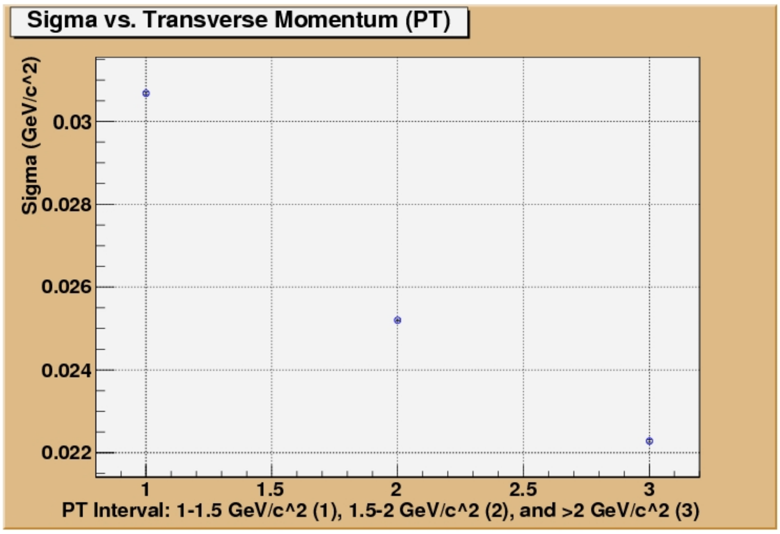


(a) Results for the mean of a fit with a second order logarithmic function.

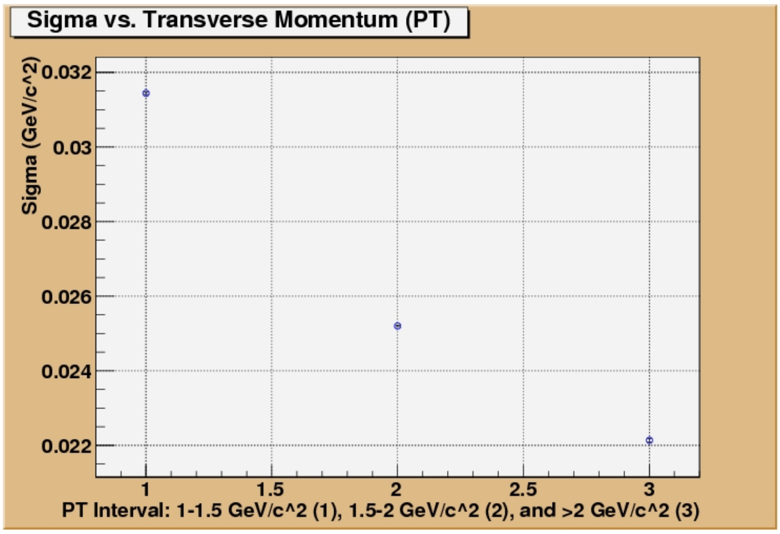


(b) Results for the mean of a fit with a third order polynomial function.

Figure 24: A plot with error bars of the average candidate neutral pion mass against the intervals of transverse momentum of 1 to 1.5 GeV/c (1 on the x-axis), 1.5 to 2 GeV/c (2), and 2 GeV/c and up (3). The errors are negligibly small.

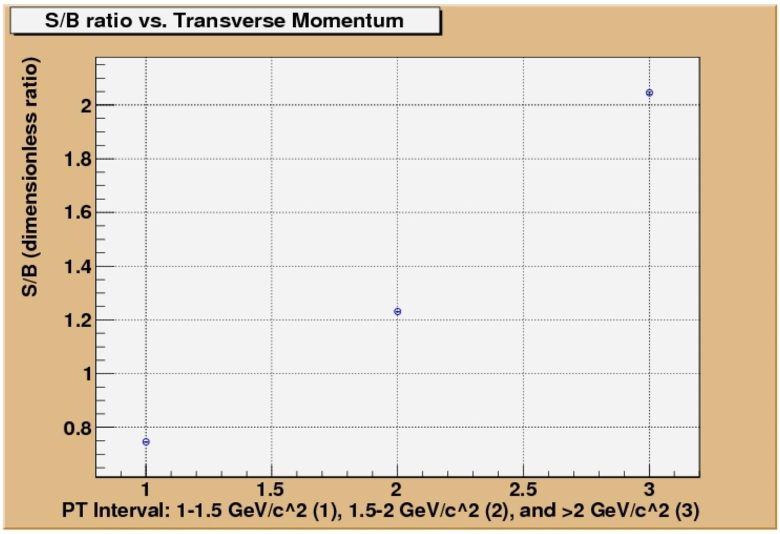


(a) Results for σ of a fit with a second order logarithmic function.

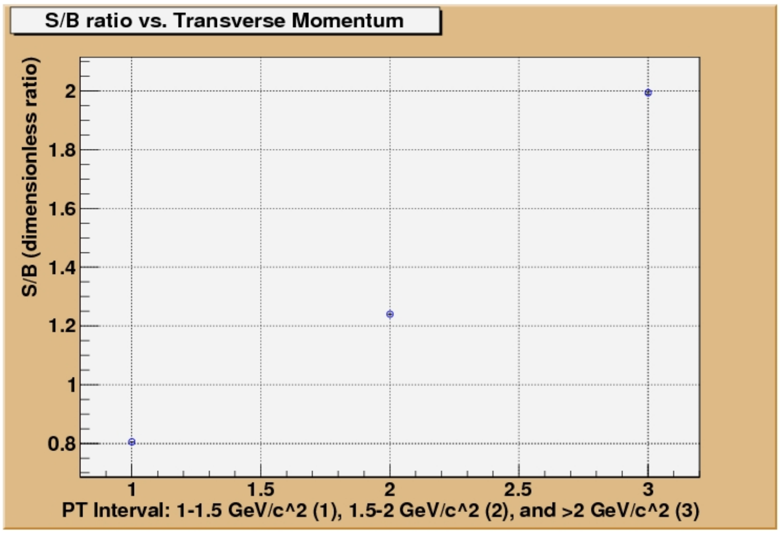


(b) Results for σ of a fit with a third order polynomial function.

Figure 25: A plot with error bars of the peak width σ against the intervals of transverse momentum of 1 to 1.5 GeV/c (1 on the x-axis), 1.5 GeV/c to 2 (2), and 2 GeV/c and up (3). The errors are negligibly small.



(a) Results for the signal-to-background ratio of a fit with a second order logarithmic function.



(b) Results for the signal-to-background ratio of a fit with a third order polynomial function.

Figure 26: A plot with error bars of the signal-to-background ratio against the intervals of transverse momentum of 1 to 1.5 GeV/c (1 on the x-axis), 1.5 to 2 GeV/c (2), and 2 GeV/c and up (3).

mass of the pion. This is possibly a result of the cluster finding algorithm. To examime the cause of this second peak, the peak could be analyzed by fitting a second Gaussian on it. This second Gaussian can then be treated as part of the combinatorial background.

As mentioned earlier (see section 4.5), a systematic error occurs when using specific fit functions on a combinatorial background. To obtain an even more precise estimate of this error, a third function, for example a fifth order polynomial, could be used to obtain an average of the three functions. The error of each function would then be the average deviation in the peak region from this average function.

Another systematic error arose in the plot of the mean, where the mean unexpectedly varied with the transverse momentum (see figure 24). More research can to be done to estimate this systematic error, which could also be a result of the cluster finding algorithm.

References

[1] Background on Color Glass Condensate, March 2009. http://www.bnl.gov/bnlweb/pubaf/pr/2003/colorglasscondensate-background.htm.

[2] K. H. Ackermann et al. STAR Detector Overview. *Nuclear instruments and methods in physics research.A, accelerators, spectrometers, detectors and associated equipment*, 499(2-3):624–632, 2003.

[3] M. Beddo et al. The STAR Barrel Electromagnetic Calorimeter. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 499(2-3):725 – 739, 2003. The Relativistic Heavy Ion Collider Project: RHIC and its Detectors.

[4] Rene Brun and Fons Rademakers. ROOT - an object oriented data analysis framework. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 389(1-2):81 – 86, 1997.

[5] S. Eidelman et al. Review of Particle Physics. *Physics Letters*, B592(1), 2004.

[6] Oleksandr Grebenyuk. *Neutral meson production in d+Au and p+p collisions at $\sqrt{s_{NN}}$ = 200 GeV in STAR.* PhD thesis, University of Utrecht, 2007.

[7] David Griffiths. *Introduction to Elementary Particles.* Wiley-VCH, second edition, 2008.

[8] L.D. Landau and E.M. Lifshitz. *Classical Theory of Fields.* Butterworth-Heinemann, fourth edition, 1980.

[9] Larry McLerran. The Color Glass Condensate and the Glasma: Two Lectures, 2008.

[10] Michael Kliemant; Raghunath Sahoo; Tim Schuster; Reinhard Stock. Global Properties of Nucleus-Nucleus Collisions. arXiv:0809.2482v1 [nucl-ex], September 2008.

[11] John R. Taylor. *An Introduction to Error Analysis.* University Science Books, second edition, 1982.

[12] Naomi van der Kolk. η Meson Analysis in Deuteron-Gold Collisions in STAR. Master's thesis, Universiteit Utrecht, the Netherlands, August 2006.