Determination of the parameters in the unintegrated Gluon Density Function

by analyzing azimuthal correlations in dijet events

Jessica Frank
DESY Summer Student Program 2008

Group: H1
Supervisor: Albert Knutsson

1. Introduction

The goal of my project is to determine some of the parameters in the unintegrated Gluon Density Function (uGDF). Therefore, I´m looking at processes in electron-proton scattering where the scattered quark emits a gluon, which splits up into a quark-antiquark-pair, these quarks then form the jets which are considered (if there are more than two jets, the two with the highest transverse momentum are picked). The uGDF depends on the transverse momentum of the gluon; the transverse momentum is related to the azimuthal angle between the two jets, which is measured. My fit of the uGDF uses these dijet data in combination with an interpolating fitting method.
Parton Density Functions like the uGDF are important for analyzing the structure of the proton and for studying some properties of Quantum Cromodynamics (QCD), like for example the energy dependence of the strong coupling constant .

2. Theoretical and experimental background

2.1. HERA and the H1 detector

HERA, which belongs to DESY in Hamburg, is a circular electron-proton collider which was operating until 2007. It can be used like a huge microscope to have a look inside the proton. HERA hosted four experiments: H1, ZEUS, HERMES and HERA-B.
H1 mainly analyses the structure of the proton and QCD. The H1 detector is shown below.

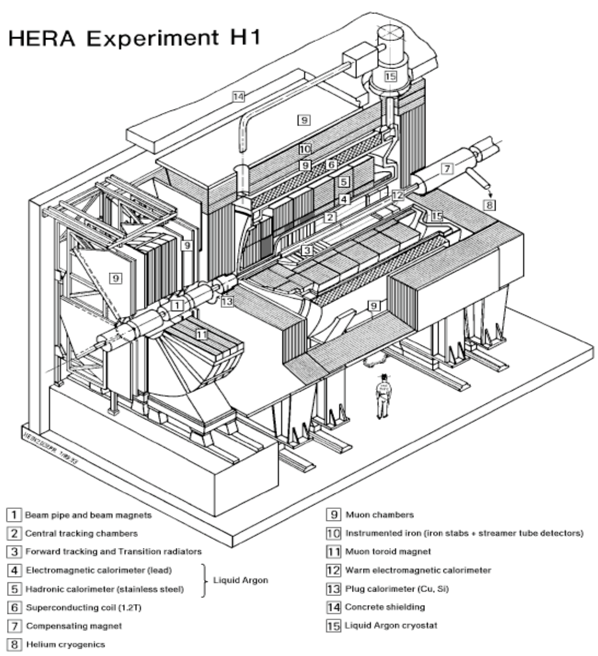


Figure 1: H1 detector

2.2. Deep inelastic electron-proton scattering (DIS)

In DIS events, like at HERA, an electron and a proton collide head-on. In the simplest case, the electron interacts with one of the partons (which means quarks and gluons) of the proton by exchanging a photon (or sometimes a Z^0 or a $W^{\pm}$, but these cases can be neglected in the studied kinematic region). After that, the scattered parton flies into a different direction as the proton remnant. The scattered parton emits a cascade of gluons and other quarks, which finally form hadrons, due to the fact that no single colored particle can exist. The result are the so-called jets. There are also more complicated possibilities of a scattering event: For example, the parton can first emit a gluon, which splits up into a quark-antiquark-pair, from which one quark then interacts with the photon.
Here is an example of a scattering event:

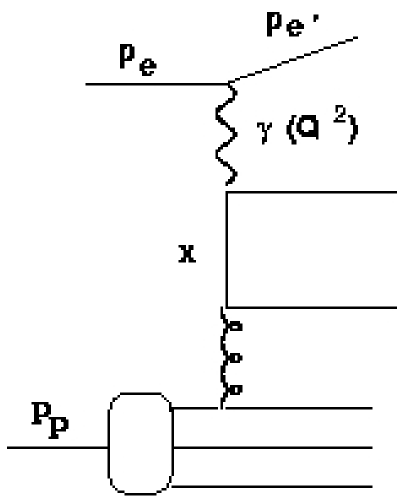


Figure 2: DIS event

To describe such events, one needs some variables like:

$$Q^2 = -q^2 \quad \text{with} \quad q = p_e - p_{e'}$$

$$x_{Bj} = \frac{Q^2}{2q\,p_p}$$

(p: four-momentum, index e: electron, index e': scattered electron, index p: proton)

Q² can be interpreted as the resolution with which one can look inside the proton; x is the fraction of the proton momentum which is carried by the scattering particle. In the lowest order reaction, where the electron interacts directly with a quark from the inside of the proton by exchanging a photon, x is the same than x_{Bj}.

2.3. Parton Density Functions (PDFs) and the cross-section

A fundamental value for the analysis of DIS events is the cross-section, which is given by the following formula:

$$\frac{d\sigma}{dx_{Bj}dQ^2} = \frac{4\pi\alpha^2}{x_{Bj}Q^4}((1-y)F_2(x_{Bj},Q^2) - y^2 x_{Bj}F_1(x_{Bj},Q^2))$$

α is the electromagnetic coupling constant, y is the inelasticity variable of the event, F_1 and F_2 are the structure functions of the proton, which depend on the PDFs. For example, F_2 is given by:

$$F_2(x_{Bj},Q^2) = x_{Bj}\sum_i c_i f_i(x_{Bj},Q^2)$$

The sum runs over all kinds of partons, c_i is the charge of a parton i and $f_i(x_{Bj},Q^2)$ is the corresponding PDF.
The PDF can be interpreted as the probability of finding a parton i carrying a fraction x of the proton momentum. As one can see, the PDFs play an important role in the description of scattering events and the proton structure.

2.4. Evolution Equations

Evolution Equations describe how a parton splits into two partons, one of them also splits and so on. These processes form parton ladders as shown below:

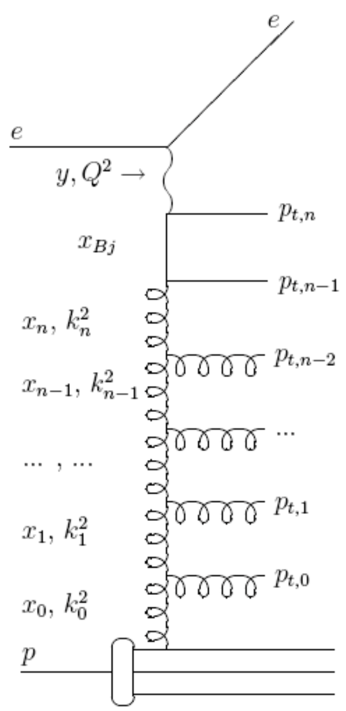


Figure 3: parton ladder

I will not go into the theoretical details, but the practical use of the Evolution Equations is that one can get rid of the Q^2 dependence in the PDFs. Once you know a PDF at one value of Q^2, you can use an Evolution Equation to determine it on every other Q^2.
One example of an Evolution Equation is the DGLAP- scheme:

$$\frac{d}{d\ln(Q^2)} f_i(x,Q^2) = \frac{\alpha_s Q^2}{2\pi} \sum_j \int_x^1 \frac{dx'}{x'} f_i(x',Q^2) P_{ij}\left(\frac{x}{x'}\right)$$

where α_s is the strong coupling constant and P_{ij} gives the probability for a parton i to radiate a parton j.

2.5. The unintegrated Gluon Density Function

The formula of the starting distribution of the uGDF is:

$$xG(x,k_T) = N\,x^{-B}\,(1-x)^C \exp\left(\frac{-(k_T-\mu)^2}{2\sigma^2}\right)$$

Unintegrated here means that it depends on the transverse momentum k_T of the gluon. N is the normalization, B (or sometimes called B_G) gives the behavior of the function at small values of x, C gives the large x behavior, μ and σ determine the shape of the k_T distribution.

The parameters N, B, C, μ and σ can't be calculated theoretically, they need to be fitted experimentally. In the case of my project, C is kept fixed at 4 because the data is assumed not to be sensitive to C, since the measurement was performed at low x. There were several attempts to determine these parameters. One recent result from Albert Knutsson was (see [3]):

N= 0.47, B= 0.08, σ= 0.5, μ= 3.0

3. The data

The uGDF is here fitted to dijet data which was collected by H1 in the years 1999 and 2000. DIS events with $5\ GeV^2 < Q^2 < 100\ GeV^2$ and at least two jets with a transverse energy larger than 5 GeV are taken. For such events, the cross-section is measured as a function of $\Delta\phi$.

The measurement contains scattering events like the one shown in figure 2. There, a gluon splits up into a quark and an antiquark, which produce two jets, one from the quark and the other from the antiquark (involving the interaction with the gluon). Because of the conservation of momentum, the azimuthal angle between these two jets depends on the transverse momentum k_T of the gluon.

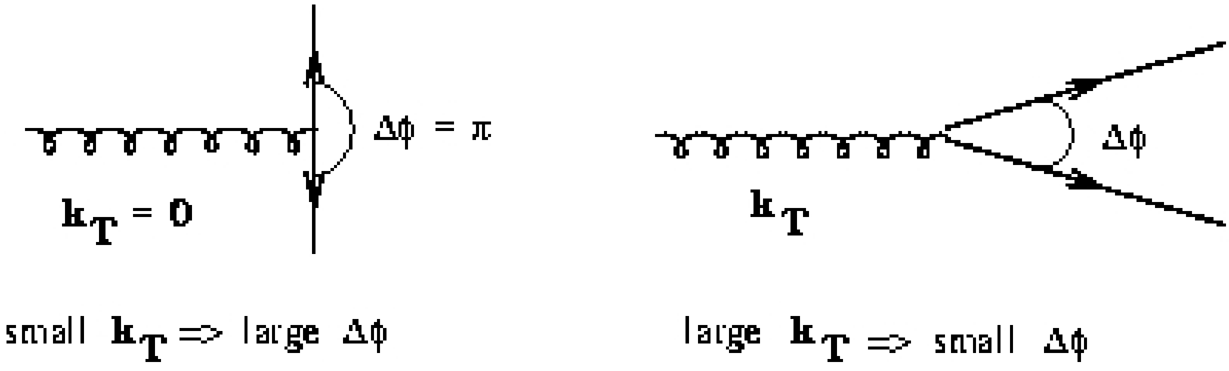


Figure 4: Dependence of $\Delta\phi$ on k_T

Therefore, the measurement of $\frac{d\sigma}{d\Delta\phi}$ is sensitive to the k_T distribution (here, σ is the cross-section, not the parameter in the uGDF).

4. The method

In the fitting procedure, in general, one compares the data to predictions given by a monte- carlo generator (I used CASCADE). The predictions are made for several values of the parameters. To see which value of the parameters gives the best predictions, one looks at the value of χ^2, which is defined as:

$$\chi^2 = \sum_i \frac{(Y_i^{MC} - Y_i^{exp})^2}{(\delta Y_i^{MC})^2 + (\delta Y_i^{exp})^2}$$

(Y: observable, MC: monte-carlo prediction, exp: experimental data, δY_i: error of Y_i) The lower χ^2, the better is the prediction. So one searches for the minimum of χ^2 to get the best description of the data.

In my fitting method, the idea is the following procedure: First I make a rough χ^2-scan for each parameter, while the other parameters are kept fixed at the expected values (see 2.5). That means that I choose some values in a rather large region and search for the minimum of χ^2. After the scans, I choose some values around these minima, which altogether build the grid. The grid is the combination of all $N_N * N_B * N_\sigma * N_\mu$ points (N_i: number of the chosen values for parameter i).

Then I use a program to fit a polynomial to this grid:

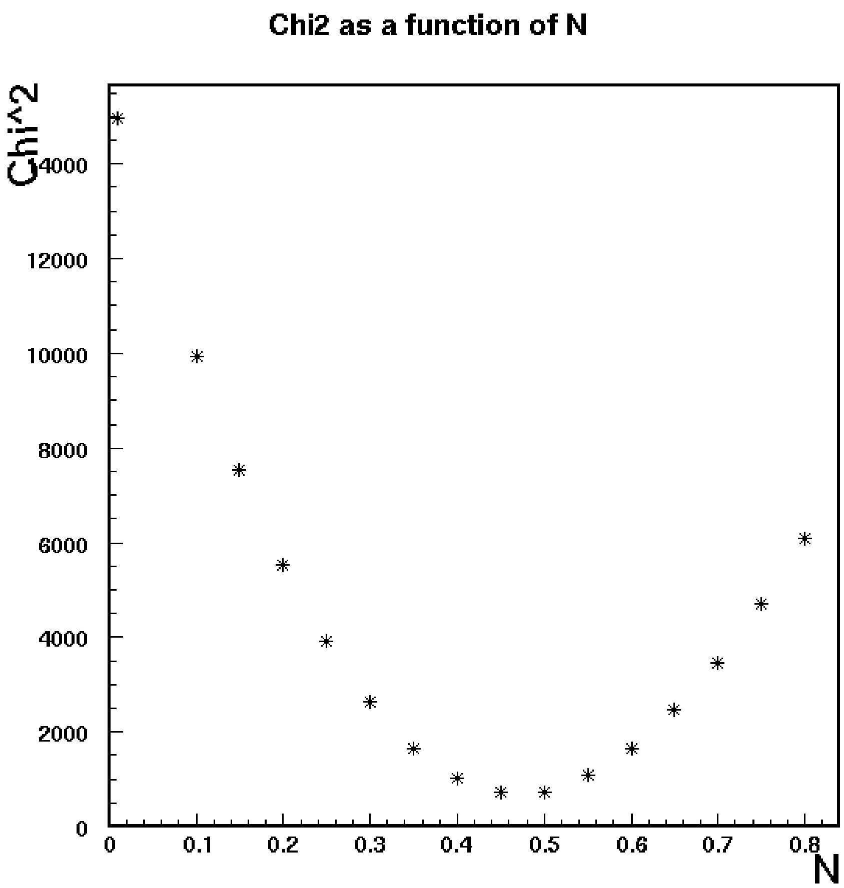
$$y(p_i) = A + \sum_{i=1}^{n} B_i p_i + \sum_{i=1}^{n} C_i {p_i}^2 + \sum_{i=1}^{n} \sum_{j=i+1}^{n} D_{ij} p_i p_j + \ \dots$$

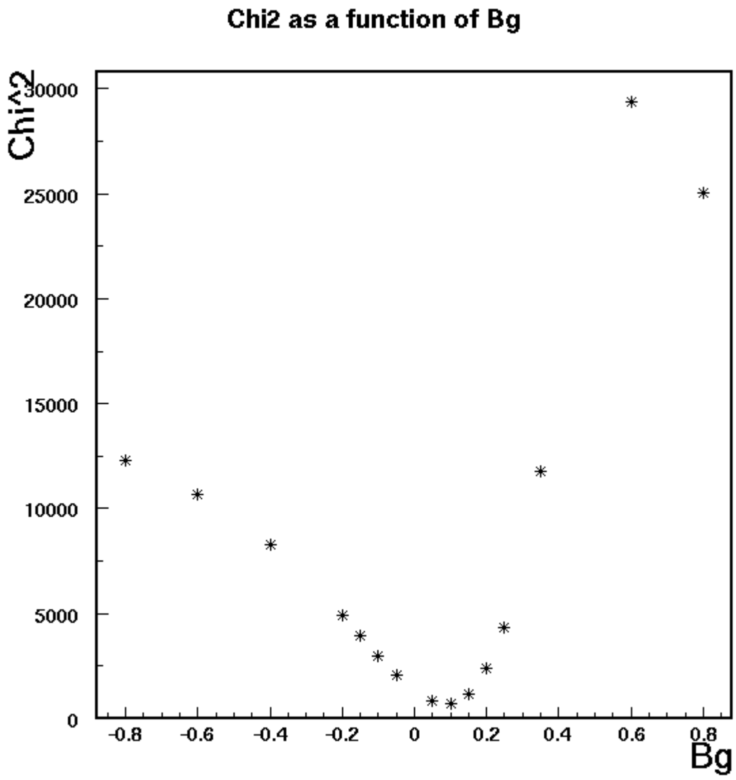
(p_i: parameters (in this case N, B, σ and μ), n is the number of parameters (here: 4) , A, $B_i, C_i, D_{ij}, \dots$ are fitted by the program). In the last step, the parameters p_i are determined by fitting all the polynomials to data simultaneously.

5. Results

First, I was making a rough χ^2-scan for each parameter, while the other parameters were kept fixed at the expected values from the result I mentioned above (N= 0.47, B= 0.08, σ= 0.5, μ= 3.0).

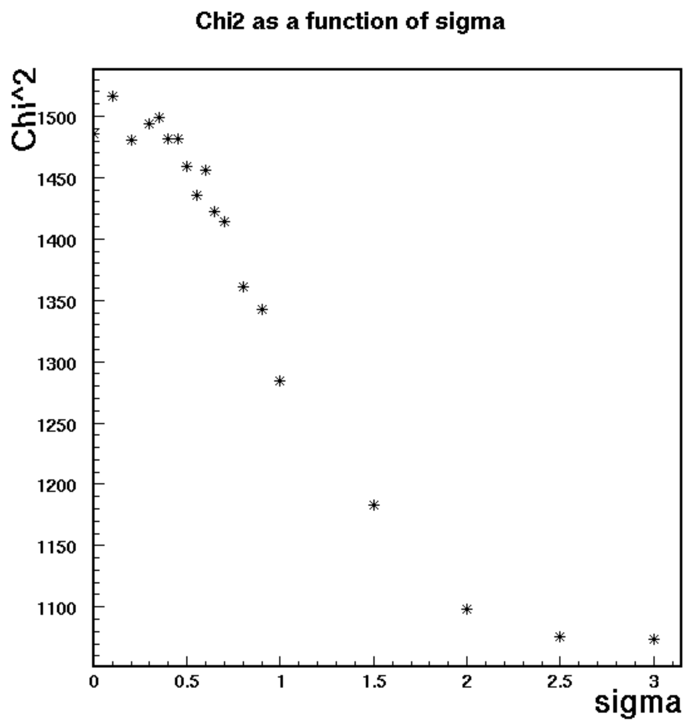
These are the resulting scans for N and B:

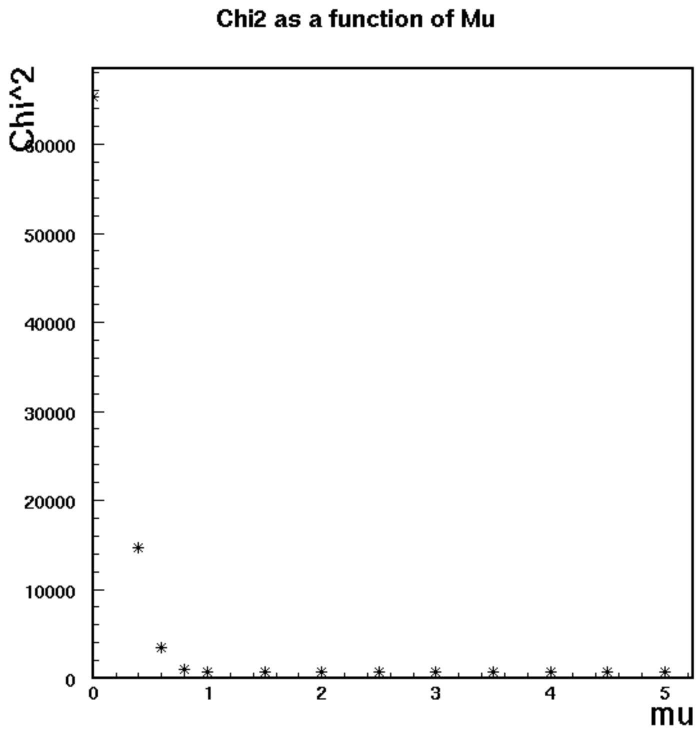




In both scans, one can see a clear minimum, which is at the expected value (for N, it's around 0.47 and for B, it's around 0.08). But these minima don't have to be the same than my final results (and as one will see, they aren't), because the other parameters are fixed at values based on another result and all the parameters are correlated.

For σ and μ, the scans are the following:

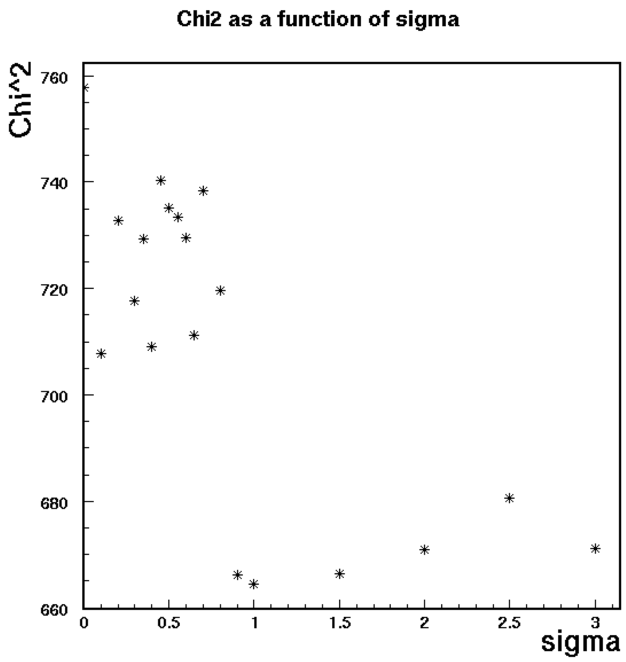




In both scans, there isn't a minimum, but χ^2 is getting smaller and smaller for larger values of σ and μ.

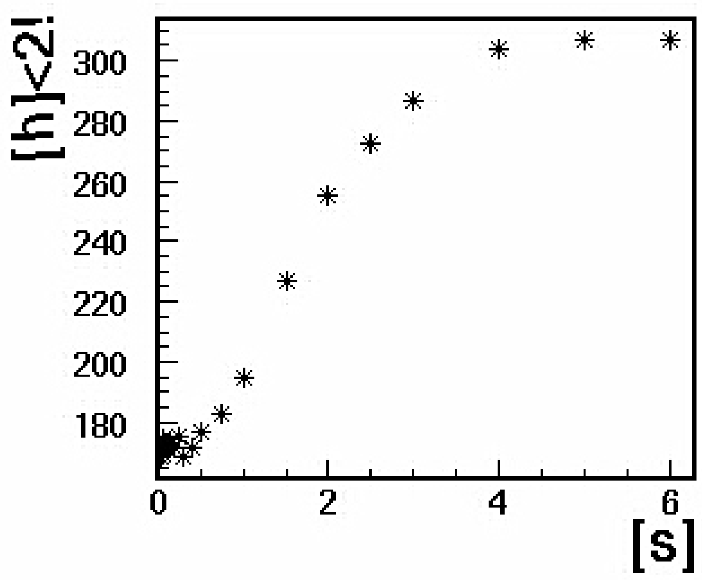
One thing I especially realized while making the scan for σ is the importance of a p_T-cut. For all my following results, I also used a p_T- cut (at $p_T > 3$). This means, that the monte- carlo generator only calculates events where the transverse momentum of the jets is larger than $p_T = 3$. Therefore, one gets more statistics and can avoid errors due to statistical fluctuations.

This was my χ^2-scan for σ without a p_T- cut:



Here one can see a lot of statistical fluctuations, especially at low values of σ. So one can't get any information from such a scan.

My result of the σ-scan is surprising, because it is quite different than the one Albert received. This is Alberts σ-scan:



Here, χ^2 is getting larger for larger values of σ, so it goes in the opposite direction than in my scan, and it has a minimum around 0.5.

After these scans, I was building up and analyzing the grid. I started with few points in 2 dimensions (just N and B, the other parameters were kept fixed at one value), analyzed this grid, added some new points in the interesting regions and also added some points in sigma later to get a 3 dimensional grid. Building and analyzing the grids always took at lot of time, so unfortunately I didn´t have the time to build up a 4 dimensional grid to analyze mu also and to add more points in sigma.

My final grid was the following:

N	0.04	0.1	0.2	0.36	0.47	0.54	0.7	
B	- 0.2	- 0.05	0.08	0.15	0.3	0.45	0.6	0.75
$\boldsymbol{\sigma}$	1.0	2.0	3.0					
$\boldsymbol{\mu}$	3.0							

By analyzing this grid, the results weren't consistent: When I analyzed all different regions of x_{Bj} together, I got another result than when I analyzed all these regions separately. The result should be the same; I don´t know the exact reason for this inconsistency.

The following table shows the results I obtained for all the different regions of x_{Bj} separately and for all combinations of them:

	N	**B**	$\boldsymbol{\sigma}$	$\boldsymbol{\chi^2}$**/ndf**
1	0.44 ± 0.09	0.21 ± 0.06	3 ± 1.65	0.41
2	0.28 ± 0.07	0.31 ± 0.09	3 ± 1.43	1.43
3	0.47 ± 0.1	0.05 ± 0.07	3 ± 1.56	0.7
1 and 2	0.27 ± 0.06	0.35 ± 0.08	2.45 ± 0.7	3.17
1 and 3	0.1 ± 0.02	0.70 ± 0.06	1 ± 0.19	11.1
2 and 3	0.15 ± 0.04	0.50 ± 0.1	1.95 ± 0.7	7.0
1, 2 and 3	0.11 ± 0.02	0.65 ± 0.07	1.03 ± 1.75	7.51

$$(1: 8 * 10^{-5} < x_{Bj} < 3 * 10^{-4},$$
$$2: 3 * 10^{-4} < x_{Bj} < 6 * 10^{-4},$$
$$3: 6 * 10^{-4} < x_{Bj} < 10^{-2})$$

So I had two candidates for the result, which means for the values of the parameters that describe my data best: One of them was the result for fitting all regions together, the other was the average of the results I obtained by fitting every region separately (average of the first three lines of the table). These two were:

	Result for fitting all regions of x_{Bj} together	Result for fitting all regions of x_{Bj} separately
N	0.11	0.4
B	0.65	0.2
σ	1.0	3.0
μ	3.0	3.0

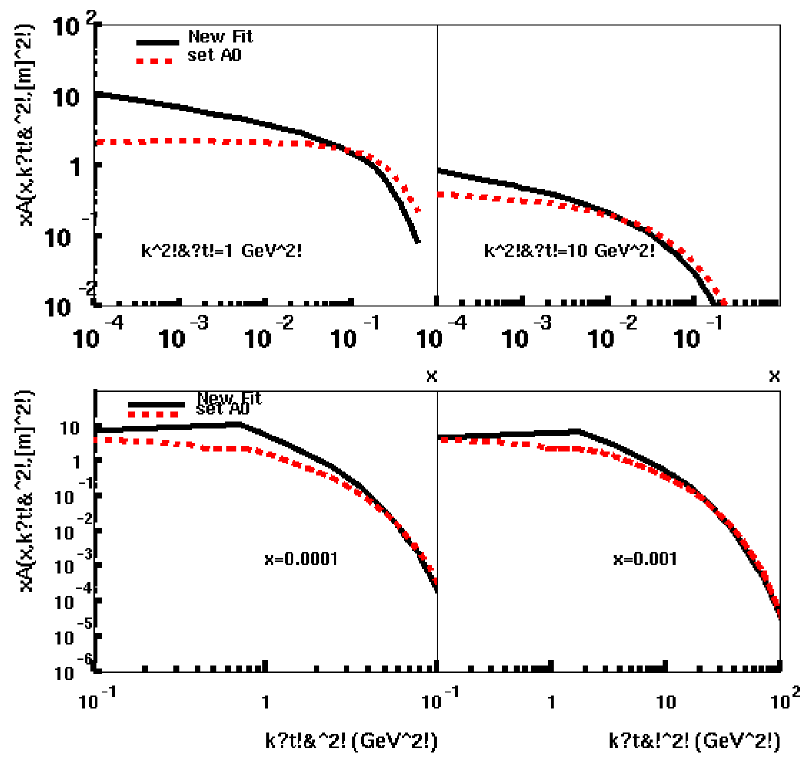
Then I calculated χ^2 for both of these candidates, and I found out that the second one describes the data better (χ^2 = 185.8) than the first one (χ^2= 347.7).

So my final result for the best fit of the parameters is:

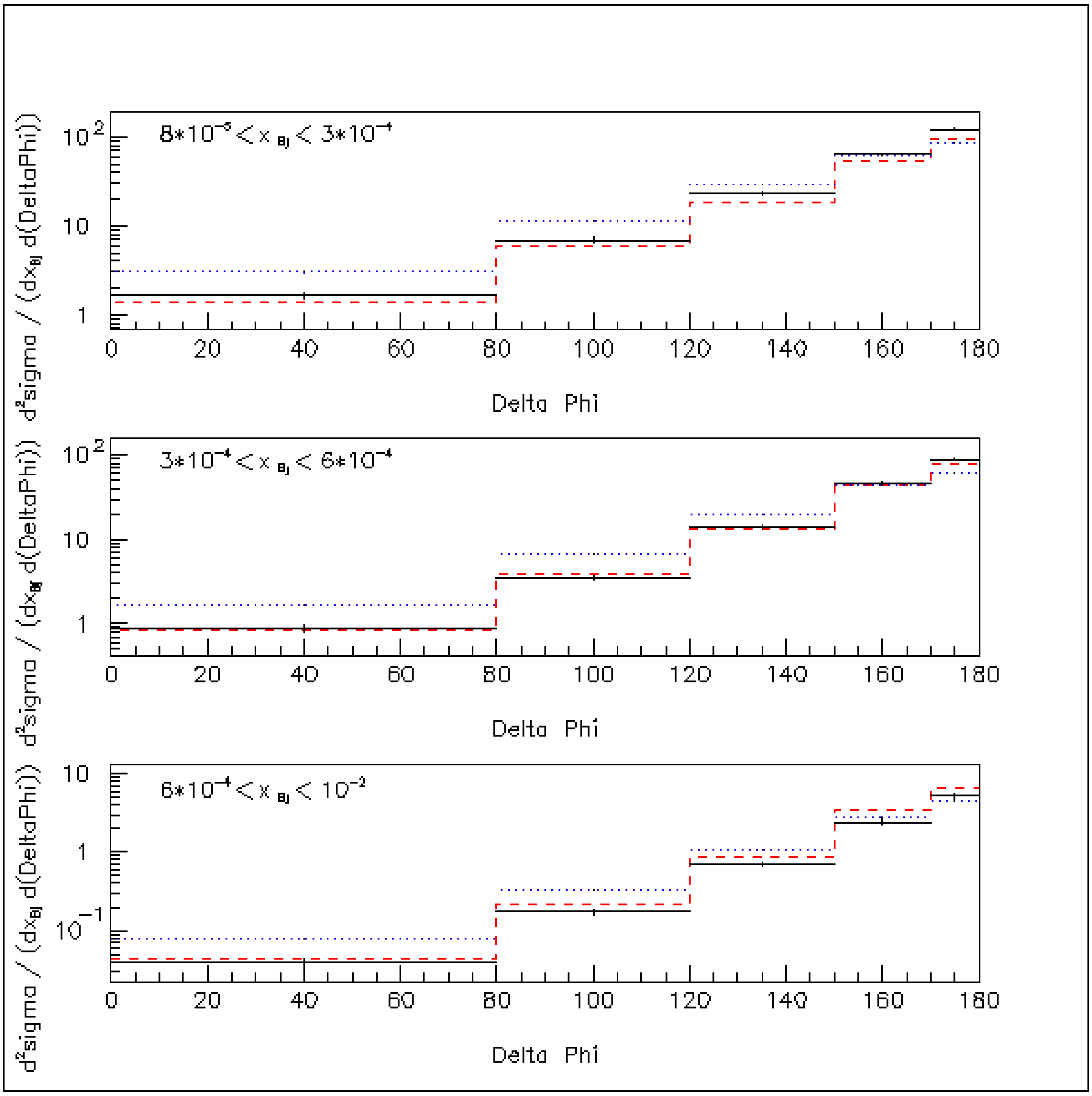
N = 0.4
B = 0.2
σ = 3.0
μ = 3.0

But these values aren´t very exact because N and B are an average, for sigma I didn’t have many points in the grid and my result for mu is just obtained from the chi^2 scan and not from a 4 dimensional grid.
For sigma and mu, one can also use larger values than 3.0, because one can see in the chi^2 scans that the curves flatten out for large values of sigma and mu.

With these values, the uGDF looks like this (new fit means my actual one):



I was also creating a graphical comparison of my fit with the data and with one older fit, which is called A0 (there, the values of the parameters are: B = 0.0, σ = 1.0, μ = 0.0, this is also used in the previous plot):

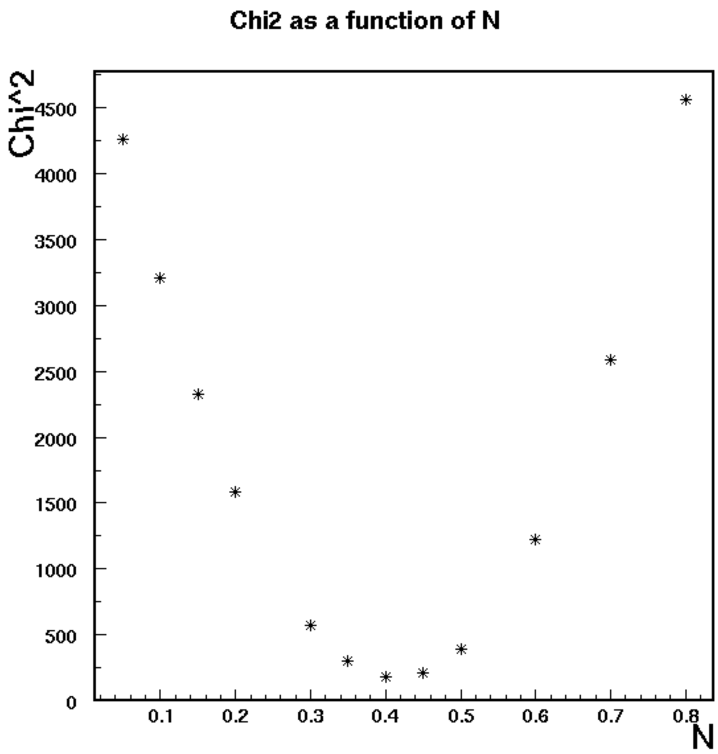


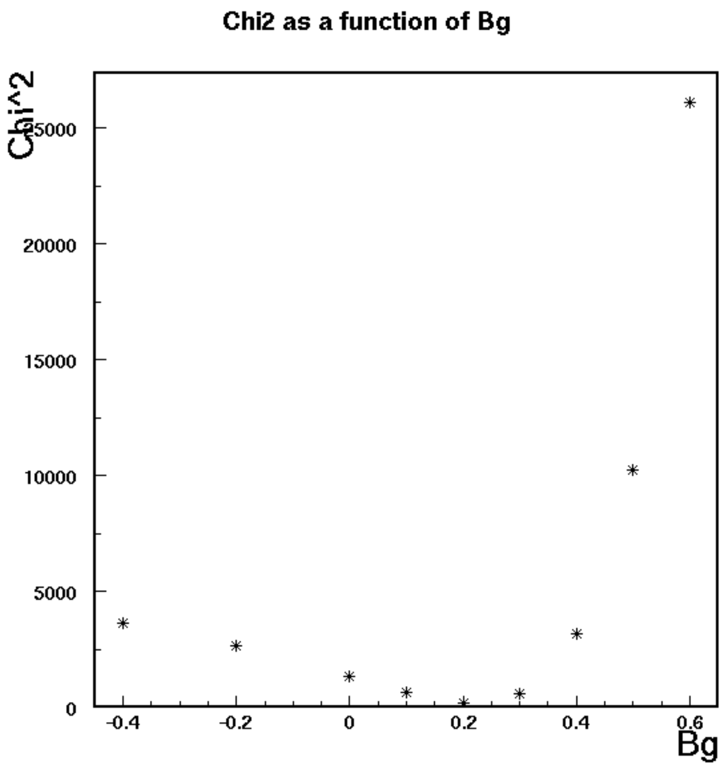
The black line is the data, the blue, dotted line is the older fit A0 and the red, dashed line is my new fit.
In this plot, one can see that my new fit is really an improvement, because is describes the data better than the old one in almost all the bins.

Finally I was making some final chi^2 scans for my new fit. In the ones shown above, I kept the other parameters fixed at the values N= 0.47, B= 0.08, σ= 0.5, μ= 3.0 while scanning one parameter, in the scans shown now I kept them fixed at

N= 0.4, B= 0.2, σ= 3.0, μ= 3.0, which is my new fit. (I didn't scan mu again because I didn't fit it in a 4 dimensional grid).

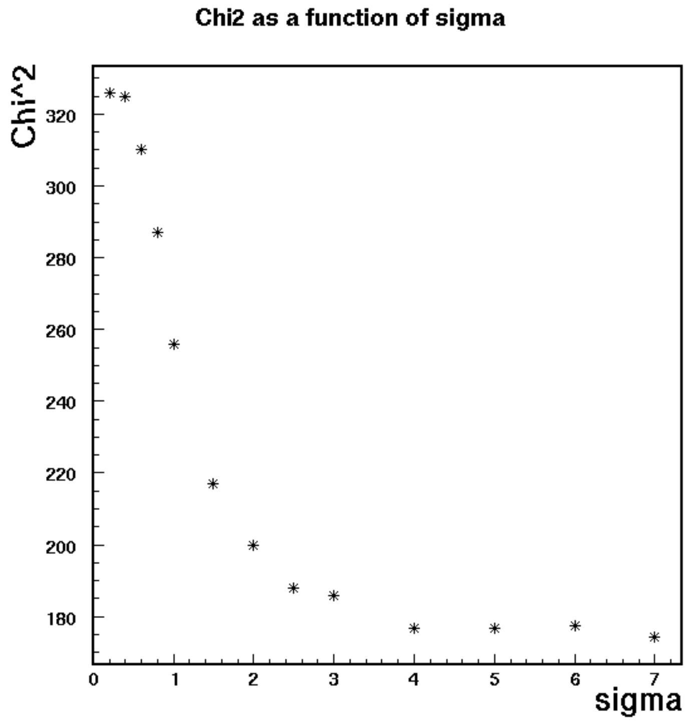
These are the scans for N and B:





One can see a clear minimum at N = 0.4 for the N-scan and at B = 0.2 for the B-scan. That confirms my obtained results, because it shows that my obtained values for N and B really give the lowest chi^2, which means the best description of the data.

For sigma, the scan is the following:



As in the other chi^2 scan for sigma, which I made before the grid, one can't see a minimum, but the curve flattens out for large values of sigma. So this scan is also consistent with the results.

6. Summary

My obtained fit for the parameters in the unintegrated Gluon Density Function is the following:

N = 0.4
B = 0.2
σ = 3.0 (or larger)
μ = 3.0 (or larger)

N, B and σ are obtained from analyzing a 3 dimensional grid, μ just comes from a rough chi^2 scan.

There were some inconsistencies during the fitting procedure, but these values describe the dijet-data they were fitted to well.

For further projects, it would be useful to take a larger grid for the fit, which contains also different values for μ, more points for σ and more points at higher values of B. Maybe that could solve the inconsistency and give a fit which is more exact.
It would also be interesting to analyze how well this fit can describe other data or for example the proton structure function F_2.

Compared to other results, where B was 0.08 (Alberts result) or even 0.0 (A0), my fit prefers a steeper rising and higher values of the uGDF if one goes to smaller and smaller values of x.

References

[1] Albert Knutsson: Forward jet production in ep-collisions at HERA (PhD thesis)

[2] Magnus Hansson: Azimuthal correlations in dijet events (PhD thesis)

[3] A. Bacchetta, H. Jung, A. Knutsson, K. Kutak: An approach to fast fits of the unintegrated gluon density (HERA-LHC workshop 2008)