Energy flow in the forward region of the CMS Detector

Samantha Dooling

under the supervision of

Mira Kraemer and Albert Knutsson

September 6, 2010

Abstract

A nice summer at DESY in Hamburg
Summer Student Program 2010, DESY

The content of this report concerns the energy flow in the forward region of pseudorapidity $3.15 < |\eta| < 4.9$ of the CMS Detector. My task was to simulate pp scattering processes with the Monte Carlo generator Pythia6.4 and compare the energy flow in dijet data to the Monte Carlo predictions. The Monte Carlo simulations are performed for three different centre-of-mass energies $\sqrt{s} = 900, 2360, 7000$GeV. The energy flow with $\sqrt{s} = 7000$GeV was measured with two hadronic forward calorimeters in the region $2.9 < |\eta| < 5.2$ at the CMS. The predictions of the Monte Carlo generator are compared to the data.

Contents

1 Introduction

In this report the Monte Carlo (MC) events for dijet samples in the central region of the pseudo rapidity are generated and the energy flow in the forward region is represented.
The energy of charged and neutral stable particles is summed in bins of the pseudo rapidity $\eta = -log(tan(\frac{\theta}{2}))$. Where θ is the polar angle of the pp scattering process.
The energy flow is defined as:

$$\frac{1}{N_{events}} \frac{d \sum E_i}{d\eta}$$

Where E_i is the energy of the particle.
The forward region of the pseudo rapidity covers the range of $3.15 < |\eta| < 4.9$. The energy flow is measured at the CMS Detector with two hadronic forward calorimeters (HF) at a centre-of-mass energy of 7000GeV. For the dijet samples the energy flow is simulated for three different $\sqrt{s}$ and compared with each other and to the data.

The energy flow in the forward region is directly sensitive to the amount of parton radiation and the multiparton interaction. The amount of parton radiation in the forward region is expected to be larger than in the central region. This can be tested by using the energy flow in the central region.
A comparison between the energy flow in the forward region (events from minumum bias samples have zero or little partonic interactions) to the energy flow in the central region (events from dijet samples, have one or more hard partonic interactions) gives results for the physics of the underlying events.
If one has knowledge about the energy flow, one can discriminate between the different models of the multiparton interaction.

Due to all these important conclusions one can draw from the energy flow, it is from great interest to simulate it. Using Pythia6.4 MC event generator to make predictions of the energy flow for dijet samples one can gain some knowledge about the QCD. My main task during my stay at DESY was to write an analysis routine, which calculates the energy flow in the forward region of the CMS Detector of dijet samples.
In this report I will present my results for three different beam energies. The MC simulations made at $\sqrt{s} = 7000$GeV are in addition compared to corrected CMS data.

2 Physic background

In pp scattering two partons of the protons interact by radiating gluons, which can e.g. produce a $q\bar{q}$ pair or radiate more gluons (figure 1).

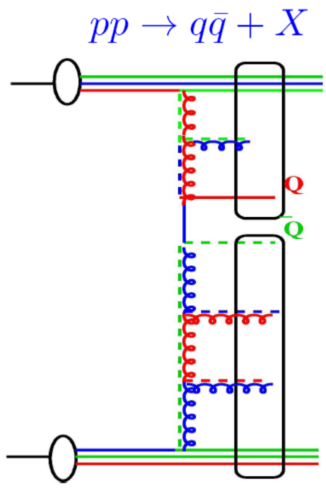


Figure 1: pp scattering process (http://www.desy.de/f/students/lectures2010/jung.pdf)

In this figure the pseudo rapidity decreases from the bottom to the top. In the centre, where the matrix element of the $q\bar{q}$ pair occurs, η has a value of zero. Due to the property of confinement the partons take part in the fragmentation process and jets emerge. A main issue of particle physics is the correct discribtion of the fragmentation and the influences of the reactions of the renmants, the so called multiparton radiation.

The parton distribution functions discribe the partonic content of the proton. For the QCD evolution of the parton densitites there exist several approaches, like DGLAP, BFKL and CCFM.
The different approximations distinguish in the following way:
DGLAP uses ordering of the virtuality of the propagation of the partons; BFKL has no ordering in k_T; CCFM has no ordering in k_T but in the angles of the radiation.
The Monte Carlo generator Pythia6.4 utilizes DGLAP.

The major task in my Summer Student work contains the calculation of the energy flow in the forward region only from dijet events. A jet is a collimated flow of particles, we are only interested in two jets. A dijet event is defined as two jets, which emerge in the central region and are back-to-back. The energy flow in the forward region is then calculated.

3 The CMS Detector

At the LHC (Large Hadron Collider) in Geneva particles are accelerated up to centre-of-mass energies of 7TeV. With the CMS Detector (Cessy, France) the resulting particles of a proton proton collision are measured. The particles which emerge directly after the collision, are in general unstable and decay after a short time into a cascade of long living particles. These particles are measured in the different sections of the CMS Detector.
The structure of the detector can be seen in figure 2. CMS stands for Compact Muon Soleniod. It has a size of 21m and a diameter of 15m. With a weight of 12500t the CMS is the heaviest particle detector, that has ever been built at an accelerator.

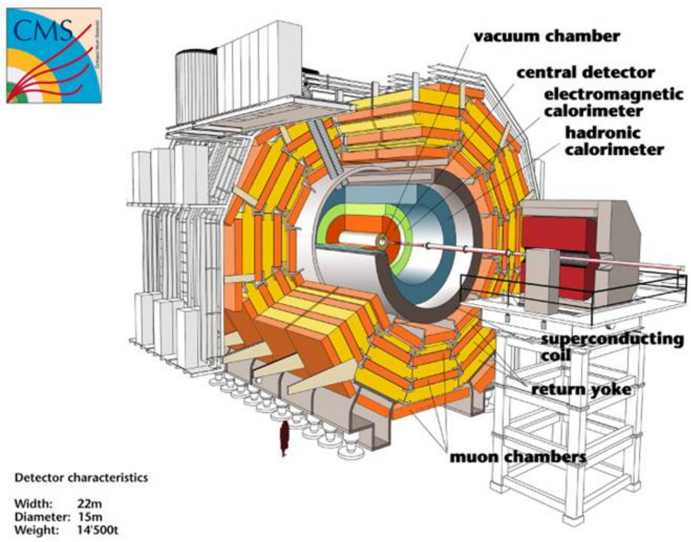


Figure 2: The CMS Detector (http://www-hep.colorado.edu/experimental/)

The three main aims of the CMS Detector is to discover the Higgs particle, search for the super symmetry (SUSY) and the measurement of collisions of heavy ions.
The structure of the detector is bowl-shaped, the components disposed around each other in several layers. In the centre the silizium-pixel detector is situated, to measure the position of charged particles. Thereafter the calorime-

ters are located, which measure the energy of the particles. On the outside margin are the muon chambers, to verificate the muons. All the elements are surrounded by a strong magnetic field, which forces the protons on a curved track, to select the particles due to their energy and charge.

In this work, the MC simulations of the energy flow in the forward region of the detector are performed and compared to the data. The resulting data are obtained from two hadronic forward calorimeters (HF), which cover a pseudo rapidity range of $2.9 < |\eta| < 5.2$. The distance to the interaction point on each side is 11.2 m.

4 Monte Carlo Software

One main challenge of experimental and theoretical particle physics, is to compare the detailed measurement with the theoretical predictions. One solution is the simulation of data with the MC method. It refers to any procedure that makes use of random numbers and uses probability statistics to solve the problems. In high energy physics MC methods are used to simulate high order reactions and multiparticle final states, for which analytical calculations are currently not possible.
The MC samples are generated by Pythia6.4 This is a generator for events in high-energy physics. The focus lays on multiparticle interaction in collisions between elementary particles. In this analysis, the pp collison is used. To evalute the generated data we use Rivet-1.2.1. Certainly Pythia6.4 is written in Fortran 72, but Rivet-1.2.1 is programmed in C++. Therefore we additionally need an interface, we use AGILe-1.1.6.

AGILe stands for 'A Generator Interface Library (& executable)'. It provides a standard interface for Fortran generator codes. AGILe handeles the steering of the generators and also saves the output in the Hep MC form, which is used by Rivet. It runs the generator and provides a text file including all the MC events.
Rivet stands for 'Robust/Indipendent Validation of Experiment and Theory'. It is a toolkit to analyze and validate MC simulations. Rivet is efficient to compute observables, hence it provides a set of experimental analyses that are needed for a generator to do the sansity check. Moreover it supplies a suitabel interface to include own analyses.
The structure of a Rivet program is seperated into three functions: init(), analyze(), finalize(). The function init() contains all declarations that have

to be done once, e.g. variables and histograms. The function analyze() is accessed for every event, here the histograms are filled with our samples. The function finalize() is called once after the event generation and used for e.g. normalization of the histograms.

There are many free parametes in the MC generator that cannot be determined theoretically, but one has to tune them. The MC simulator Pythia is used in different tunes for the parameters of the simulation of the underlying event. Steering files correspond to the different tunes of the parameters. First the steering file is filled with two commands: MSEL = 2, CKIN(10) = 5. The first option makes sure all subprocesses are included. The second option just makes the event generation more effective by putting a p_T-cut on the hard scattering. All other parameters in Pythia not listed here, are at their default values. The predictions from this MC run are marked in my histograms as 'with MPI'. MPI stands for multiparticle interaction, which is turned on in this tune. Including the command MSTP(81)=0, has the consequence that the generator is runned without MPI. This MC run is labeled with 'without MPI'. Moreover we use the tunes: D6T, DW, PROQ20 and Perugia_0 (P0), which is using a new multiple interaction model implemented in Pythia.
The results for the different tunes are compared to each other and to the data.

5 Analysis

The major task in my Summer Student work contains the calculation of the energy flow in the forward region only from dijet events in the central region. I wrote a Rivet Analysis covering this issue. In the following section I will present the structure of my code and the main calculations.
The resulting dijet sample has to fullfill the following conditions: the pseudo rapidity of the hardest jets in the event has to be in between $|\eta| < 2.5$ and the difference of the azimuth angles of the two jets has to fulfill $||\Delta\phi(j_1, j_2)| - \pi| < 1.0$.
We are generating the samples for three different centre-of-mass energies and for each a different p_T cut on the jets is required. For $\sqrt{s} = 900$GeV and $\sqrt{s} = 2360$GeV the jets have to have a transverse momentum higher than $p_T > 8$GeV and for $\sqrt{s} = 7000$GeV a $p_T > 20$GeV is claimed. The cut has to be done in order to have well defined jets in the detector and a high p_T scale in the central region.

The structure of my Rivet Analysis split into the three parts init(), analize() and finalize(). The first function init() contains the booking of the histograms. Additional to the forward-energy-flow histogram I also set histograms concerning the number of particles in the two hardest jets, aswell as the number of jets, we obtain from the events. Furthermore the knowledge of the η distribution and the $\Delta\phi$ distribution is from great interest.
In the function init() the projections FinalState and FastJets are implemented. The jets are reconstructed by means of the anti-k_T-jet algorithm, with the jet radius R = 0.5, with stable particles as input.

In the following I will detail the construction of the main part of my analysis the subroutine analyze(). In the beginning the beam is set to a proton proton beam: setBeams(PROTON,PROTON) in the Constructor.
The analyze() function contains all the important cuts and calculations.

First the generated event is skipped, if it is empty and no resulting particles are emerged. After that I ask for the number of jets. For a dijet sample we need two jets, so if the number of jets in the event is below two, I can skip the event aswell, otherwise the program continues.
The next step concerns the two hardest jets in the event. That means, I have to find the jet with the highest p_T and the jet with the second highest p_T. Subsequently I ask for the centre-of-mass energy and implement the p_T-cut. For an energy of $\sqrt{s} = 900$ and 2360GeV we require the transverse momentum of the jets to be higher than 8GeV and for $\sqrt{s} = 7000$GeV the transverse momentum has to be $p_T > 20$GeV.
After that we conduct the η-cut for the two hardest jets, to examine they are in the central region, $|\eta| < 2.5$. Finally we ask for the back-to-back condition $||\Delta\phi(j_1, j_2)| - \pi| < 1.0$.
If all these conditions are fullfilled compute the energy flow in the forward region. This is done for all stable particles with $E > 4$GeV which correspond to the noise threshold cut in the data analysis.
Within the pseudo rapidity range of $3.15 < |\eta| < 4.5$ I loop over all particles and sum up all the energies in equidistant bins of η. The number of bins is 5 therefore the binwidth is 0.35. Furthermore I have to divide the energy sum by the bin width and norm the content with the number of events. Furthermore I have to divide by 2 because the energy flow is summed in bins of the absolute value of $|\eta|$.
This normalization is the content of the finalize() function.

6 Results

The energy flow for events having a hard scale defined by dijets with $E_{T,jet} >$ 8 GeV for $\sqrt{s} = 900, 2360$GeV and $E_{T,jet} > 20$ GeV for $\sqrt{s} = 7000$GeV in $|\eta| < 2.5$ has been computed in the forward region.
In figure 3 the energy flow for the centre-of-mass energy $\sqrt{s} = 900$GeV is shown for all the different tunes. I generated 2.5 Mio events with Pythia6.4 with the steering file 'with MPI' and 'without MPI'. For the other four tunes I generated 100000 events.
In figure 4 one can see the energy flow for $\sqrt{s} = 2360$GeV for all the tunes and in figure 5 the predictions of the generator are shown with $\sqrt{s} = 7000$GeV and are compared to the data from the HF calorimeters. The error bars correspond to statistical errors. The systematic uncertainty of the measurement effects mainly from the HF calorimeters. Additional distributions such as $\Delta\phi$ and η of the dijets are shown in the appendix.

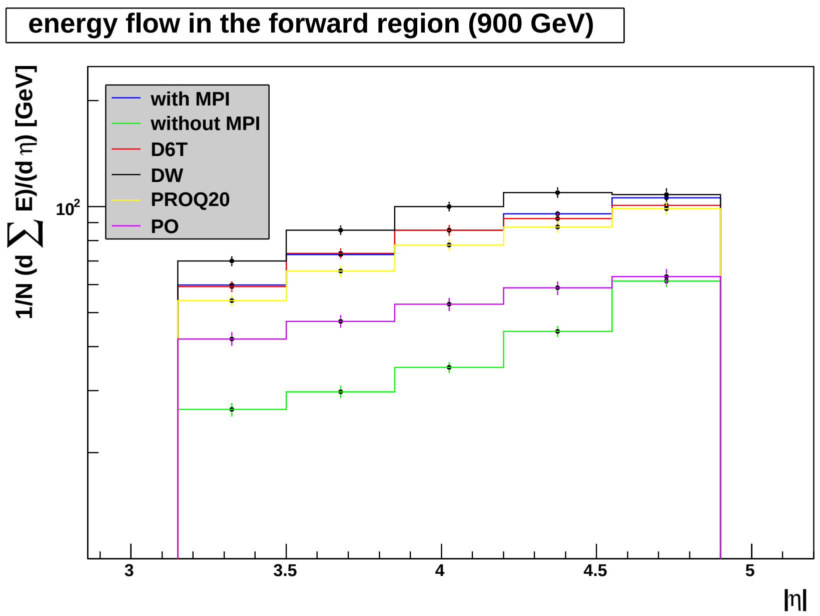


Figure 3: Energy flow in the forward region for the centre-of-mass energy 900GeV generated with Pythia6.4 for different tunes

In figure 6 the results of the CMS Collaboration published 2010/07/16 (Contact: cms-pag-conveners-fwd@cern.ch) are shown. The energy flow in dijet samples for the three different centre-of-mass energies are shown. The uncorrected data are shown as points. The error bars correspond to statistical errors. The shaded bands represent the systematic uncertainties of the measurements. I included this histogram to compare it with my results. Aswell the Pythia generator with the tunes D6T, PROQ20, P0, and DW are used.

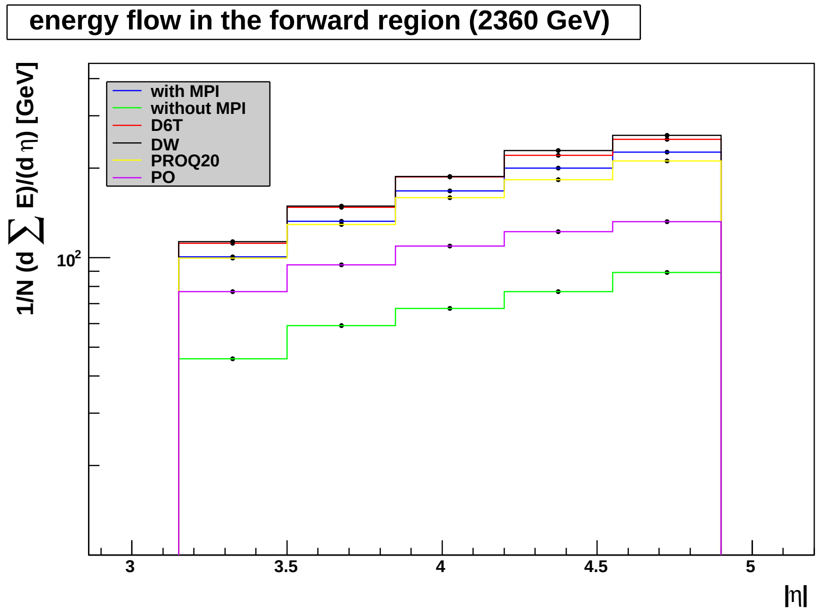
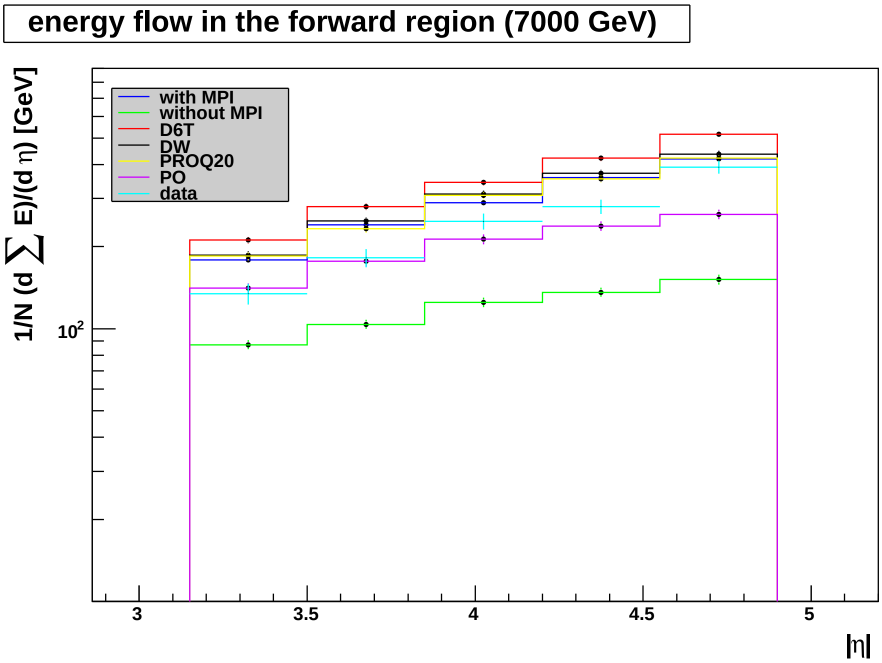


Figure 4: Energy flow in the forward region for the centre-of-mass energy 2360GeV generated with Pythia6.4 for different tunes

In table 2 - 7 in the appendix the predictions of the MC dijet samples are collected and compared to the data. The results are outlined for each bin and with the statistical error.



2

Figure 5: Energy flow in the forward region for the centre-of-mass energy 7000GeV generated with Pythia6.4 for different tunes and compared to the data

7 Conclusions

The energy flow is measured at detector level with the CMS hadronic forward calorimeters in the region $2.9 < |\eta| < 5.2$. The data was measured at a centre-of-mass energy of 7000GeV and is compared to the MC dijet event samples. Furthermore I simulated the energy flow for three different $\sqrt{s}$ to compare the flow regarding increasing centre-of-mass energy.

One can see in figures 3 - 5 and from the table 2 - 7 that the energy flow is increasing with increasing $\sqrt{s}$. The amount of events provide sufficient and convincing results due the small errors about 1-3%.

The energy flow has the tendency to increase with the pseudo rapidity. One can see from the MC predictions that the increase per each bin is almost the same. In comparision to the data the MC predictions reveal the same behaviour of the energy flow in the different bins concerning the increasing trend of $|\eta|$.

The percentage derivation of the measurement from the MC values differ for each Pyhia tune. The reason for using different tunes is to find the best predictions for the measurements. In table 1 the percentage derivatives of

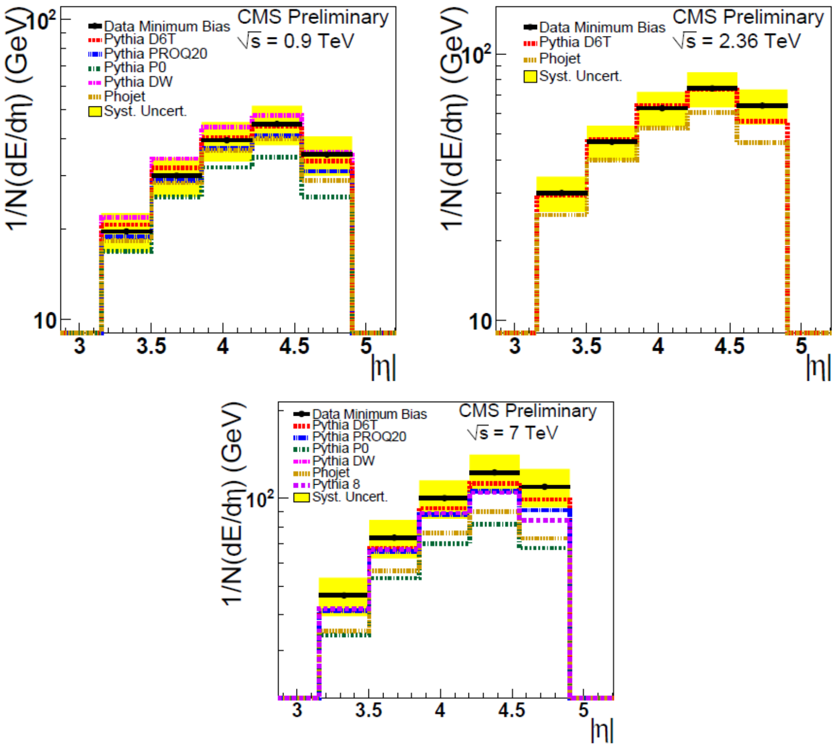


Figure 6: Energy flow in the forward region at $\sqrt{s} = 900$ GeV (left), $\sqrt{s} = 2360$GeV (right) and $\sqrt{s} = 7000$GeV (bottom). The uncorrected data are shown as points.

the data from the different tunes for $\sqrt{s} = 7000$GeV are shown for each bin and each tune.
None of the Pythia tunes discribes the corrected data within the error bars of the MC simulations. One can see a high sensitivity to the choice of the parmaeters. The tune where the MPI is turned of does not discribe the measured data at all. The percentage derivative are in between 54% - 158 %. This is not a satisfying result for a MC generaor. The D6T tune also has a high derivation from the data (24% - 36%). The first two bins are well predicted by Perugia-0, but the other three bins differ about 16% - 49% from the data.
My conclusion is that the MC simulations cannot discribe the the energy flow in all aspects. They predict the increasing of the energy flow for increasing pseudo rapidity, but the values do not coincide with the data. Hence we really have to learn more about what is inside of the data.
The big discrepancy between the data and the reaults of the MC generation

without MPI shows that this interaction has a great contribution to the energy flow in the forward region.

percentage derivation of the data at $\sqrt{s} = 7000\text{GeV}$						
$\|\eta\|$	Pythia tunes					
	'with MPI'	'without MPI'	D6T	DW	P0	PROQ20
3.15 - 3.50	24.74	53.75	36.39	27.88	4.71	27.13
3.50 - 3.85	24.15	75.37	34.96	26.75	3.02	21.77
3.85 - 4.20	14.59	98.08	28.10	20.49	16.29	19.63
4.20 - 4.55	21.77	106.01	33.61	24.61	17.85	20.85
4.55 - 4.90	6.76	158.24	24.31	10.28	48.84	7.20

Table 1: percentage derivation of the measurement from the MC values

In comparision to the uncorrected results of the CMS Collaboration (figure 6), one can see that the energy flow is increasing for the first four bins, but in the fith bin the energy flow gets less. This is a detector effect and is not seen in the corrected data I used.
As well as our results, the predictions from the MC generator do not coincide with the data.

The MC generator tunes which dicsribe the corrected data for dijet events best are 'with MPI' and PROQ20. The percentage derivation of the MC values from the measured data are in between 7% - 25% respectivly 7% - 27%. In both tunes the data values for the energy flow is below the MC prediction.

8 Literature

- CMS Collaboration, *CMS Physics Analysis Summary: Measurement of the energy flow at pseudorapidity at the LHC at* $\sqrt{s} = 900, 2360$ *and 7000 GeV*, Contact: cms-pag-conveners-fwd@cern.ch

- H1 Collaboration, *Measurements of transverse energy flow in deep inelastic-scattering at HERA*, arXiv:hep-ex/9907027

- Lectures given by H. Jung, *Monte Carlo Simulations*, http://www.desy.de/f/students/lectures2010/jung.pdf, 08.08.2010

9 Appendix

'with MPI'				
$\|\eta\|$	$\sqrt{s}$ [GeV]			
	900	2360	7000	Data 7000
3.15 - 3.50	59.84 ± 0.95	100.55 ± 0.72	178.51 ± 3.03	134.35 ± 11.84
3.50 - 3.85	73.11 ± 1.17	132.64 ± 0.95	239.82 ± 4.07	181.91 ± 14.00
3.85 - 4.20	85.78 ± 1.37	167.76 ± 1.20	289.72 ± 4.92	247.45 ± 16.29
4.20 - 4.55	95.35 ± 1.52	200.10 ± 1.43	358.14 ± 6.08	280.18 ± 16.94
4.55 - 4.90	105.83 ± 1.69	226.50 ± 1.62	419.45 ± 7.12	391.08 ± 21.40

Table 2: Pythia6.4 MC predictions in the tune 'with MPI' for the forward energy flow in the η bins and data measured from the HF calorimeters on detector level

'without MPI'				
$\|\eta\|$	$\sqrt{s}$ [GeV]			
	900	2360	7000	Data 7000
3.15 - 3.50	26.52 ± 0.57	45.73 ± 0.46	87.38 ± 1.77	134.35 ± 11.84
3.50 - 3.85	29.80 ± 0.64	59.14 ± 0.60	103.73 ± 2.10	181.91 ± 14.00
3.85 - 4.20	34.93 ± 0.75	67.47 ± 0.68	124.92 ± 2.53	247.45 ± 16.29
4.20 - 4.55	44.24 ± 0.95	76.94 ± 0.77	136.00 ± 2.75	280.18 ± 16.94
4.55 - 4.90	61.48 ± 1.32	89.13 ± 0.90	151.44± 3.07	391.08 ± 21.40

Table 3: Pythia6.4 MC predictions in the tune 'without MPI' for the forward energy flow in the η bins and data measured from the HF calorimeters on detector level

'D6T'				
$\|\eta\|$	$\sqrt{s}$ [GeV]			
	900	2360	7000	Data 7000
3.15 - 3.50	59.23 ± 1.72	111.68 ± 1.33	211.22 ± 6.01	134.35 ± 11.84
3.50 - 3.85	73.72 ± 2.14	147.70 ± 1.76	279.68 ± 7.96	181.91 ± 14.00
3.85 - 4.20	85.62 ± 2.49	186.82 ± 2.22	344.16 ± 9.80	247.45 ± 16.29
4.20 - 4.55	92.58 ± 2.69	221.10 ± 2.63	422.03 ± 12.01	280.18 ± 16.94
4.55 - 4.90	100.80 ± 2.93	249.59 ± 2.97	516.70 ± 14.71	391.08 ± 21.40

Table 4: Pythia6.4 MC predictions in the tune 'D6T' for the forward energy flow in the η bins and data measured from the HF calorimeters on detector level

'DW'				
$\|\eta\|$	$\sqrt{s}$ [GeV]			
	900	2360	7000	Data 7000
3.15 - 3.50	70.03 ± 1.93	113.26 ± 1.37	186.28 ± 5.73	134.35 ± 11.84
3.50 - 3.85	85.80 ± 2.36	149.18 ± 1.80	248.34 ± 7.64	181.91 ± 14.00
3.85 - 4.20	100.10 ± 2.76	187.67 ± 2.26	311.23 ± 9.57	247.45 ± 16.29
4.20 - 4.55	109.68 ± 3.02	229.38 ± 2.77	371.63 ± 11.43	280.18 ± 16.94
4.55 - 4.90	108.29 ± 2.98	257.93 ± 3.11	435.91 ± 13.41	391.08 ± 21.40

Table 5: Pythia6.4 MC predictions in the tune 'DW' for the forward energy flow in the η bins and data measured from the HF calorimeters on detector level

'Perugia-0'				
$\|\eta\|$	$\sqrt{s}$ [GeV]			
	900	2360	7000	Data 7000
3.15 - 3.50	42.11 ± 1.21	76.89 ± 1.05	140.99 ± 4.372	134.35 ± 11.84
3.50 - 3.85	47.25 ± 1.36	94.61 ± 1.29	176.57 ± 5.48	181.91 ± 14.00
3.85 - 4.20	52.80 ± 1.51	109.28 ± 1.49	212.78 ± 6.60	247.45 ± 16.29
4.20 - 4.55	58.79 ± 1.69	122.31 ± 1.67	237.75 ± 7.37	280.18 ± 16.94
4.55 - 4.90	63.35 ± 1.82	132.08 ± 1.81	262.75 ± 8.15	391.08 ± 21.40

Table 6: Pythia6.4 MC predictions in the tune 'Perugia-0' for the forward energy flow in the η bins and data measured from the HF calorimeters on detector level

'PROQ20'				
$\|\eta\|$	$\sqrt{s}$ [GeV]			
	900	2360	7000	Data 7000
3.15 - 3.50	54.14 ± 1.49	99.78 ± 1.19	184.36 ± 5.44	134.35 ± 11.84
3.50 - 3.85	65.59 ± 1.80	129.21 ± 1.55	232.53 ± 6.87	181.91 ± 14.00
3.85 - 4.20	77.87 ± 2.14	159.052 ± 1.90	307.88 ± 9.09	247.45 ± 16.29
4.20 - 4.55	87.61 ± 2.41	183.11 ± 2.19	353.98 ± 10.45	280.18 ± 16.94
4.55 - 4.90	98.82 ± 2.71	211.21 ± 2.53	421.40 ± 12.44	391.08 ± 21.40

Table 7: Pythia6.4 MC predictions in the tune 'PROQ20' for the forward energy flow in the η bins and data measured from the HF calorimeters on detector level

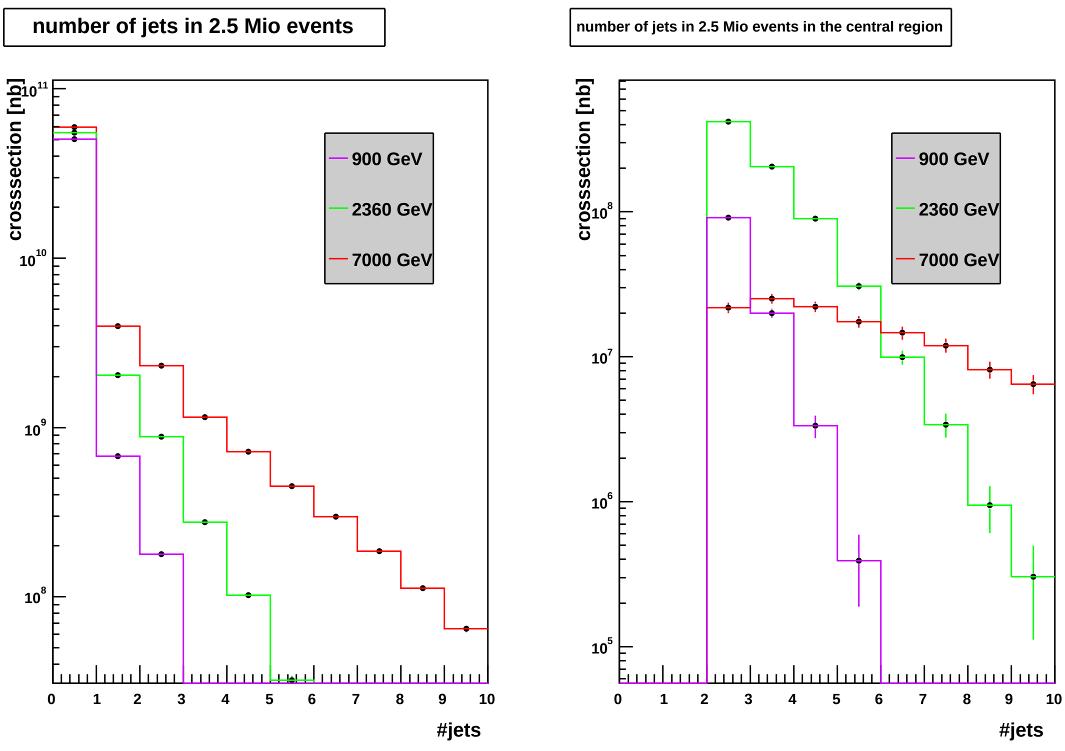


Figure 7: Left: Number of jets in the generated event for three different $\sqrt{s}$; Right: Number of jets in the central region for three different $\sqrt{s}$

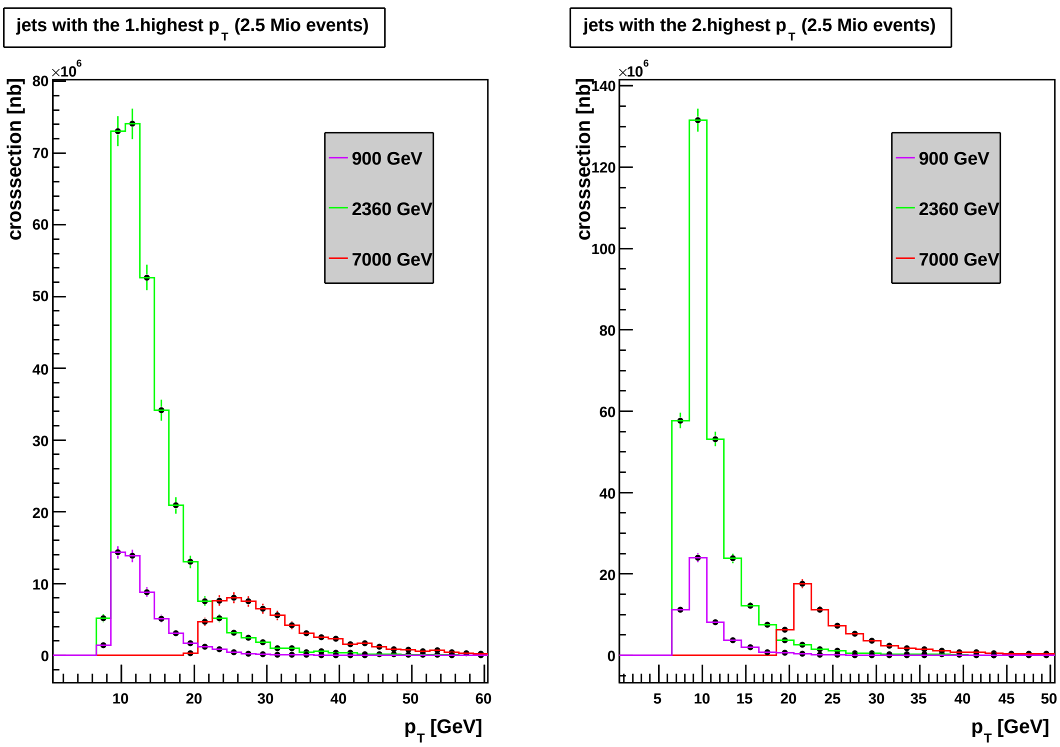


Figure 8: Left: Jets with the highest transverse momentum for three different $\sqrt{s}$; Right: Jets with the second highest transverse momentum for three different $\sqrt{s}$. The cross section of $\sqrt{s} = 7000\text{GeV}$ is less because of the transverse momentum cut at 20 GeV.

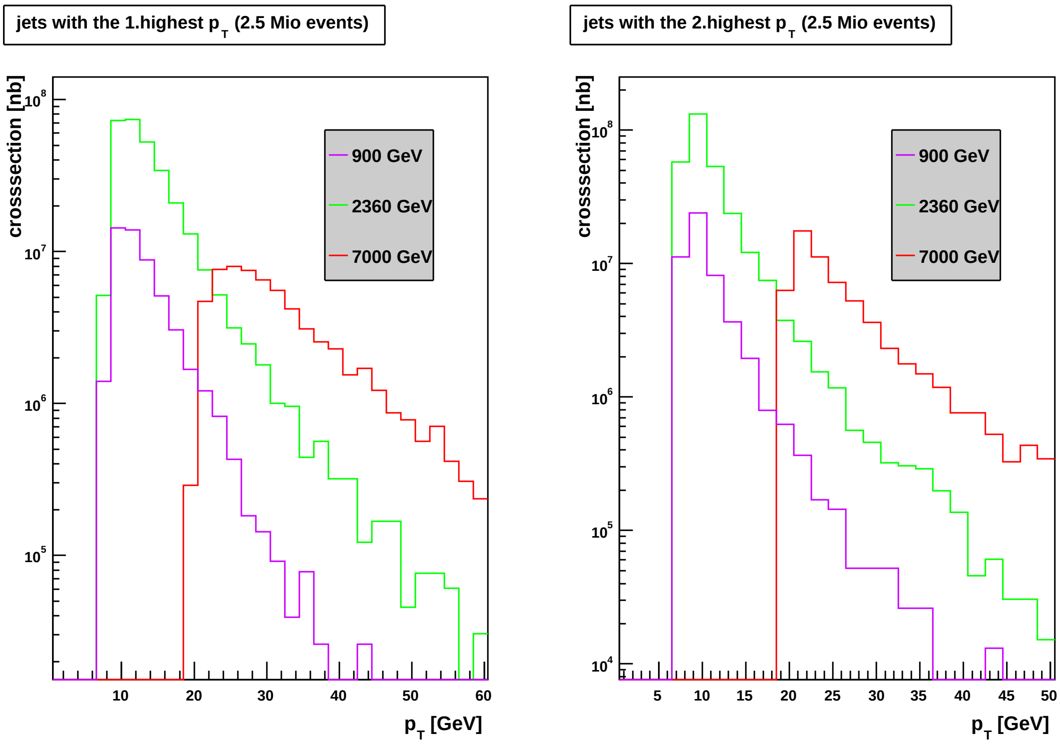


Figure 9: Left: logarithmic scale of jets with the highest transverse momentum for three different $\sqrt{s}$; Right: logarithmic scale of jets with the second highest transverse momentum for three different $\sqrt{s}$. In this figure one notices the declension for higher p_T

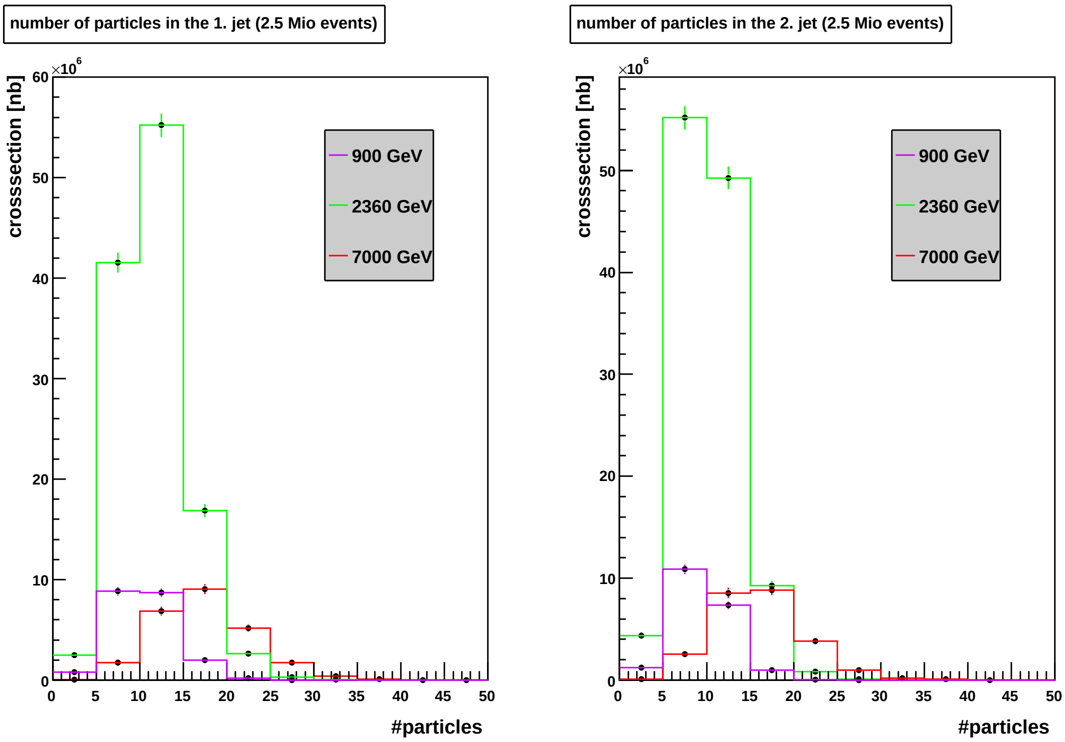


Figure 10: Left: Number of particles in the hardest jets for three different $\sqrt{s}$; Right: Number of particles in the 2. hardest jets for three different $\sqrt{s}$. One can see from this histogram that the distribution for $\sqrt{s} = 7000\text{GeV}$ is broader compared to the two other energies, due to the different p_T cuts.

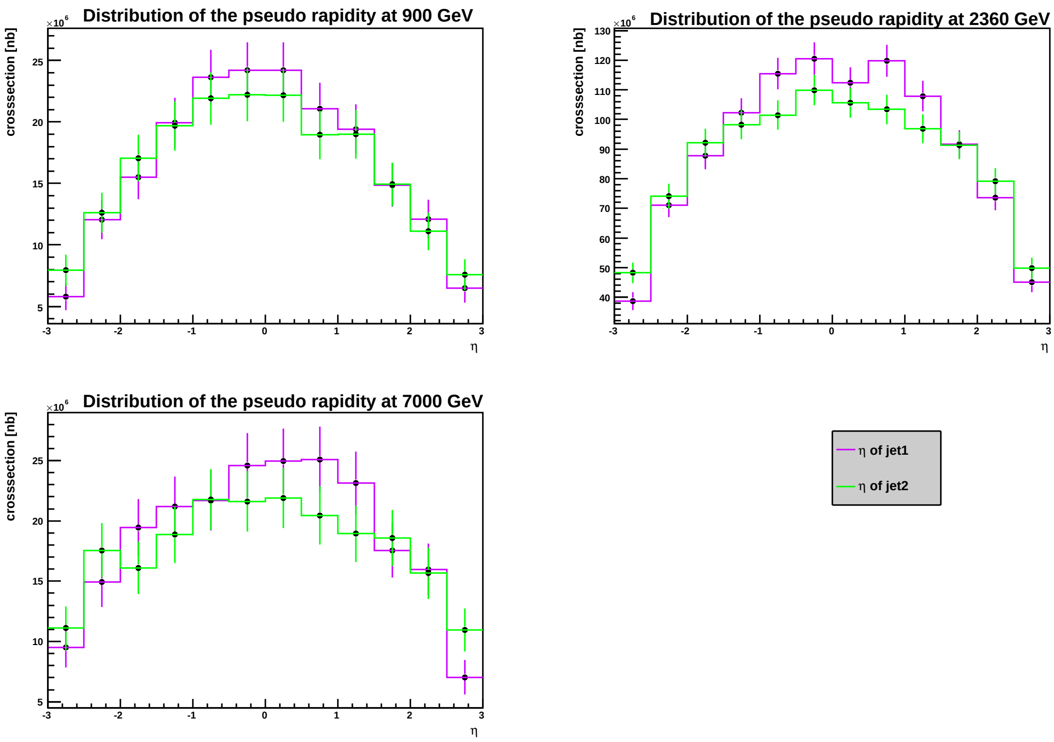


Figure 11: The distribution of the pseudo rapidity for three different centre-of-mass energies. In each figure are the pseudo rapidities for the two dijets are presented.

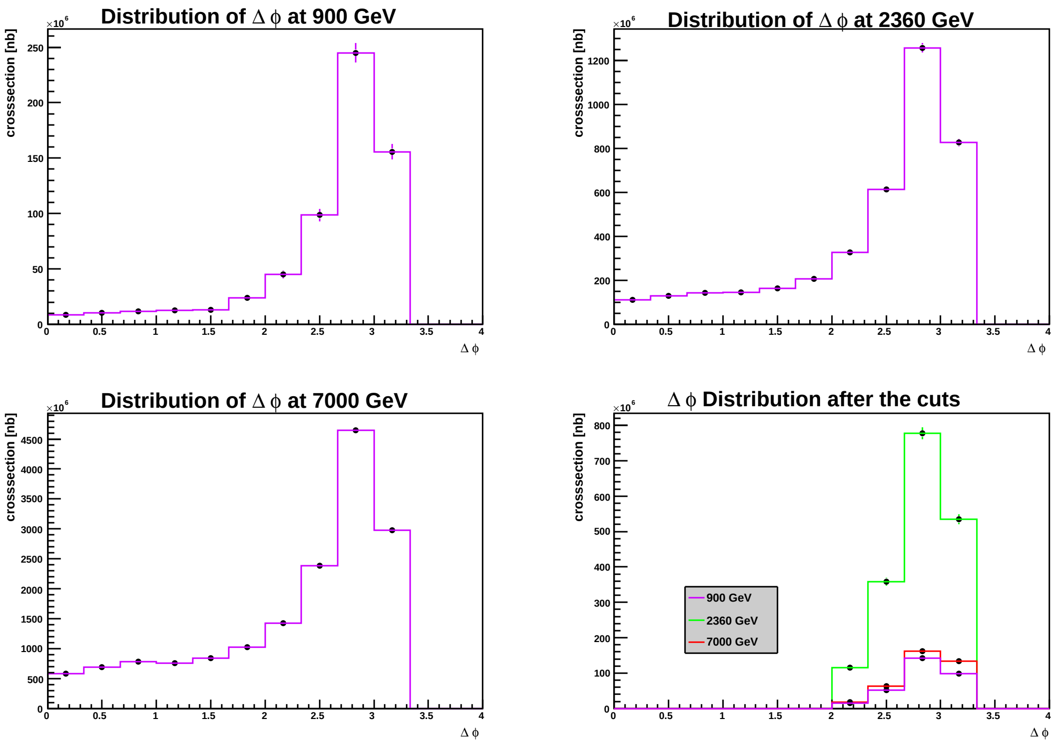


Figure 12: The distribution of $\Delta\phi$ of the dijets for three different centre-of-mass energies. The expected course of the distribution is well presented in the histograms. One can see that the cross section increases and diverges at $\Delta\phi = \pi$. In the figure on the bottom right side the distribution after the cuts is represented. The appearance of the hisrogram is not a precise peak at $\Delta\phi = \pi$ because of the allowed variation of 1.0.