Event Shape Variables in Semileptonic Top Decays

Final Report - Summer Students 2007 CMS group

Antonia Mey B.Sc.

14th, September 2007

SUPERVISOR: PROF. DR. MNICH

Abstract

A Monte Carlo generator study was conducted by looking at the $lvb\bar{b}q\bar{q}$ (semileptonic) decay mode of a top quark pair, in hindsight of experiments at the LHC. Signal data were produced using Alpgen, Herwig, Mc@nlo and Pythia generators. Event shape variables were used to help identify a top quark pair signal from QCD and W+Jets background data, created by Alpgen generator. Modified Fox Wolfram Moments, sphericity, aplanarity, circularity, centrality are applied to the data by selecting events with four or five jets and a transverse momentum cut $p_t > 15$ GeV. Signal data was found to be generator independent as expected. Looking at event shape variables centrality appears to be the only useful variable for semileptonic top events to distinguish between signal data and background.

1 Introduction

The Compact Muon Solenoid is one of the new detectors for the Large Hadron Collider (LHC) going into operation in 2008. At DESY there is a small group working on the CMS project, in which I was involved as a summer student. The group at DESY is particularly interested in top quark physics in respect to the new CMS detector. The particular work that was undertaken was in foresight of the PhD thesis by Mr A. Flossdorf, who is carrying out an investigation into different Monte Carlo simulators. Monte Carlo simulations are used to produce top event signatures, as there are no experimental data available yet, which will hopefully improve the analysis of real detector data once the LHC is operational. In particular the generator study was used for delving into the analysis of event shapes of the semileptonic top decay channel. The report will give some brief background information on the CMS detector as well as the top quark pair production in hadronic collisions. Then an introduction to event shape variables will be given with the analysis of the results obtained.

2 Background Information

2.1 Compact Muon Solenoid - CMS

CMS is one of the four experiments at the LHC which is being build. It is a general puspose detector for hadron collisions of up to 14 TeV. The detector is represented in the graphic below: The main

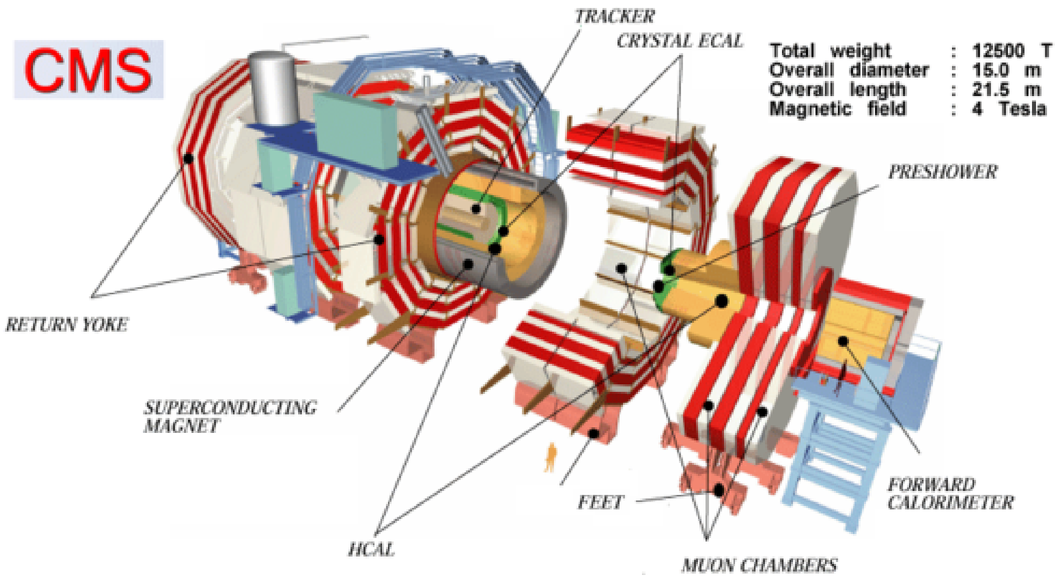

Figure 1: 3D drawing of CMS detector [1]

parts of this detector are: the tracker, an electromagnetic and hadronic calorimeter, a superconducting solenoid and a muonchamber. Both tracker and calorimeters fit inside the solenoid producing a strong magnetic field of 4 T. This design lays the ground for detecting particles such as the Higgs boson, which have not been experimentally found up till now. It will also be used for various other experiments such as looking at the physics of top quarks, which is the object of this report.

2.2 Semi Leptonic $t\bar{t}$ decay

The top quark was discovered at Fermilab in March 1995 [2]with a mass of around 174.2 GeV. This discovery lead to a whole new area of physics. Top physics will play an important role at LHC when operational. The hadron collider will produce a large number of top events, due to the high centre of mass energy of the collision. For example gluons will interact and create a top antitop pair. Such an event is represented in the diagram below: It can be seen that, when two hadrons collide, gluons

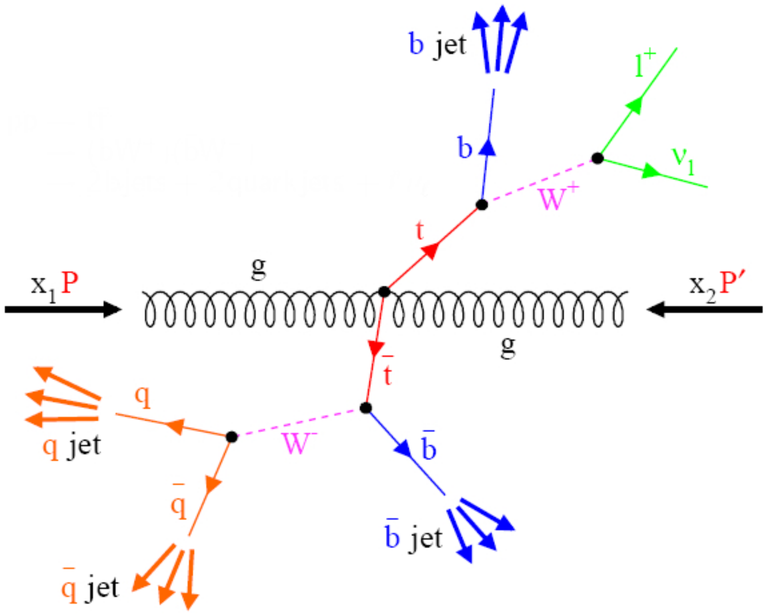

Figure 2: semileptonic top decay topology [3]

carrying a certain momentum fraction of the colliding protons, can create a $t\bar{t}$ pair. The top quarks then decay into a $b\bar{b}$ and two W bosons. Depending on the decay mode of the W bosons three different decay channels for the top decay can be identified: Fully hadronic, semileptonic and dileptonic. The semileptonic decay channel is the one studied here. In this case one of the W bosons decays into a lepton and the corresponding antineutrino and the other one into a light $q\bar{q}$ pair. Both b quark and anti b quark will produce all sort of final state particles bundled into a jet with a certain cone opening angle. The quark antiquark pair will produce further two jets. The percentage of semileptonic top decays lies at about 44%. However, generally when looking at the semileptonic decay only the first two lepton generations are of interest, that is the production of an electron or muon and the τ will be neglected. This then leaves a decay probability of 30% for the semileptonic decay channel. One characteristic of the event can be observed due to the creation of the bottom and the distance it travels before its decay occurs.

The question that arises is for what purpose are top events being analysed? There are two main reasons for it, one is of experimental nature. The highly scientific machines will enable even more precise measurements of quantities such as spin, mass and charge of the quark. The other motivation is of theoretical nature. The top quark is a very curious particle in the Standard Model and is closely linked to the Higgs boson. So far it has only been detected at the Tevatron therefore the repetition of the detection of such an event is desired. However it will be hard to distinguish a top event from numerous background signals. One way of screening for top signals is looking at so called event shape variables.Therefore Monte Carlo generators are used for event simulations to gain an insight into shape of such an event to make detection more likely once the experiment is running at the LHC.

3 Event Shape Variables

Event shape variables are an ideal way to help to distinguish between the actual signal of a detector event and the background. So far this can only be done with Monte Carlo generated data, however these can give a certain idea which event shape variable will be suitable to help distinguish a semileptonic top event from background data. A Monte Carlo generator study was carried out, by looking at different event shape variables for the four jets produced in a semileptonic top event. The variables under investigation were: Fox Wolfram Moments, sphericity, centrality, circulariy and aplanarity. First Fox Wolfram Moments will be discussed.

3.1 Fox Wolfram Moments

This variable was devised in 1979 by Fox and Wolfram and were originally used to characterise the final shapes of electron proton annihilations. [4] They are useful as they provide a set of rotationally invariant observables Hl. When looking at these variables masses have to be neglected. The Fox Wolfram Moments H0, H1, H2 etc, can be defined by:

$$Hl = \sum_{i,j} \frac{\mid p_i \mid\mid p_j \mid}{E_{vis}^2} Pl(cos\theta_{ij}) \tag{1}$$

For the purpose of the generator study the original Fox wolfram moments were a little altered. Instead of looking at the total momentum of the event a selection has been made. Only events containing four or five jets was used. The reason is apparent as four is the expected number of jets for a semileptonic top event topology. Depending on the jets algorithm used five jets can still occur as a top event. The other constraint made was that all jets have to highly energetic with a p_t of greater than 15 GeV. P_t in this case represents the momentum fourvector of the jets, also p_i and p_j in the equation. The produced neutrino and lepton in the event will not be taken into account hence E_{vis} represents the visible energy of the jets of the event. These adjustions to the examination of the Fox Wolfram moments is modeled after a paper by Fields et al. [4] Pl in the equation for Hl are the Legendre Polynomials. These are given by:

$$\begin{aligned} P_0(x) &= 1 \\ P_1(x) &= x \end{aligned}$$

$$
\begin{aligned}
P_2(x) &= \frac{1}{2}(3x^2 - 1) \\
P_3(x) &= \frac{1}{2}(5x^2 - 3x) \\
P_4(x) &= \frac{1}{8}(35x^4 - 30x^2 + 3)
\end{aligned}
$$

In the case of the Fox Wolfram Moments x is given by θ_{ij}. This will mean that H1 will be zero if the momentum of the event is balanced. [5] The generator study in particular was looking at four different event generators for the signal data, which are Alpgen, Herwig, Mc@nlo and Pythia. Only the jets of the events were of importance and the cut of $p_t > 15GeV$ will prevent irrelevant low energy jets to be included. Fox Wolfram Moments have been observed for numerous Polynomials. At first signal plots using all four different generators for H0, H1, H2 and H4 were generated. These can be found below: It is easily observed, that the shape of the curves suing different generators is very similar. Not only

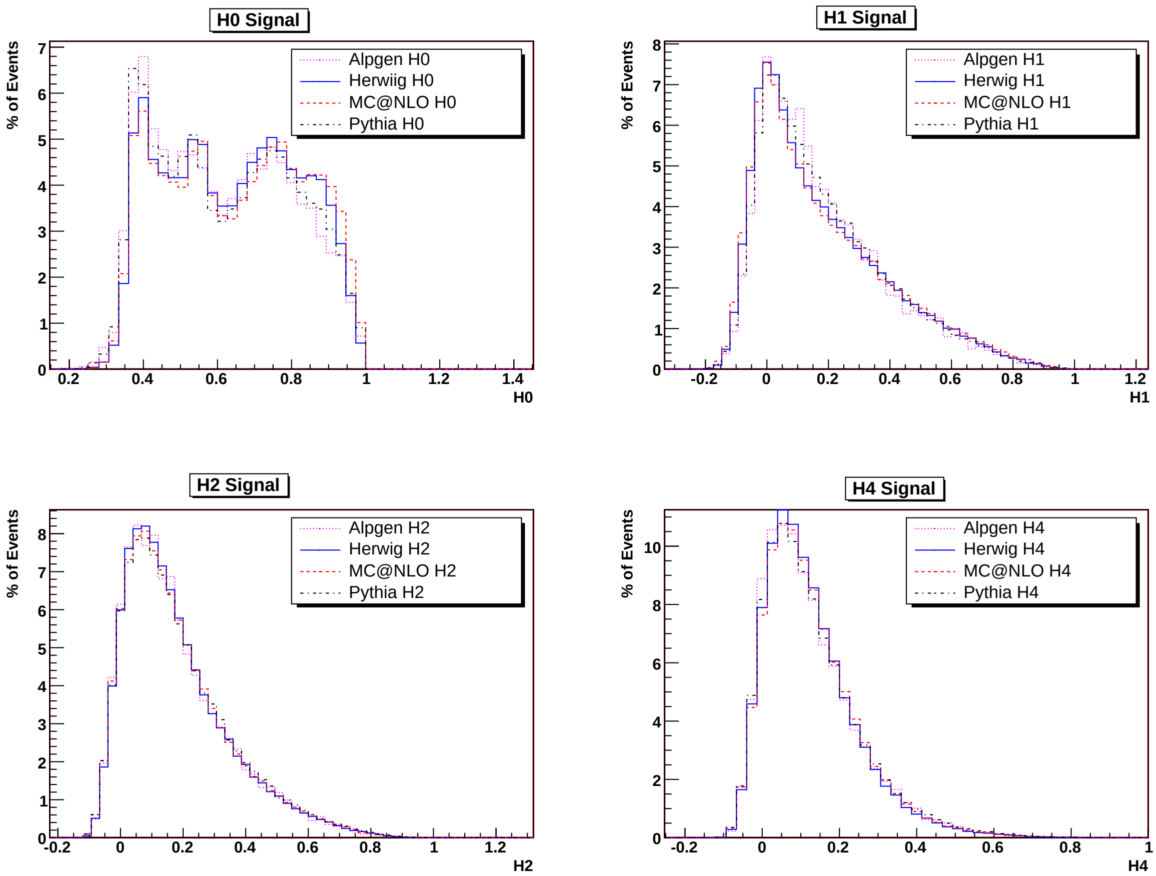

Figure 3: Fox Wolfram Moments Signal

Fox Wolfram Moments but all event shape variables have been normed between [0,1]. On this account the y-axis has been assigned a general percentage of events to make plots comparable with each other. Therefore the comparative study of different event generators is possible. Hence for the observed Fox

Wolfram moments it can be said that the event shape is generator independent. Furthermore plots using background data were created, for the same Hl Fox Wolfram Moments. These represent QCD and W+Jets background data created using the Alpgen generator as well as the signal produced with Alpgen. Again the four graphs can be found below:

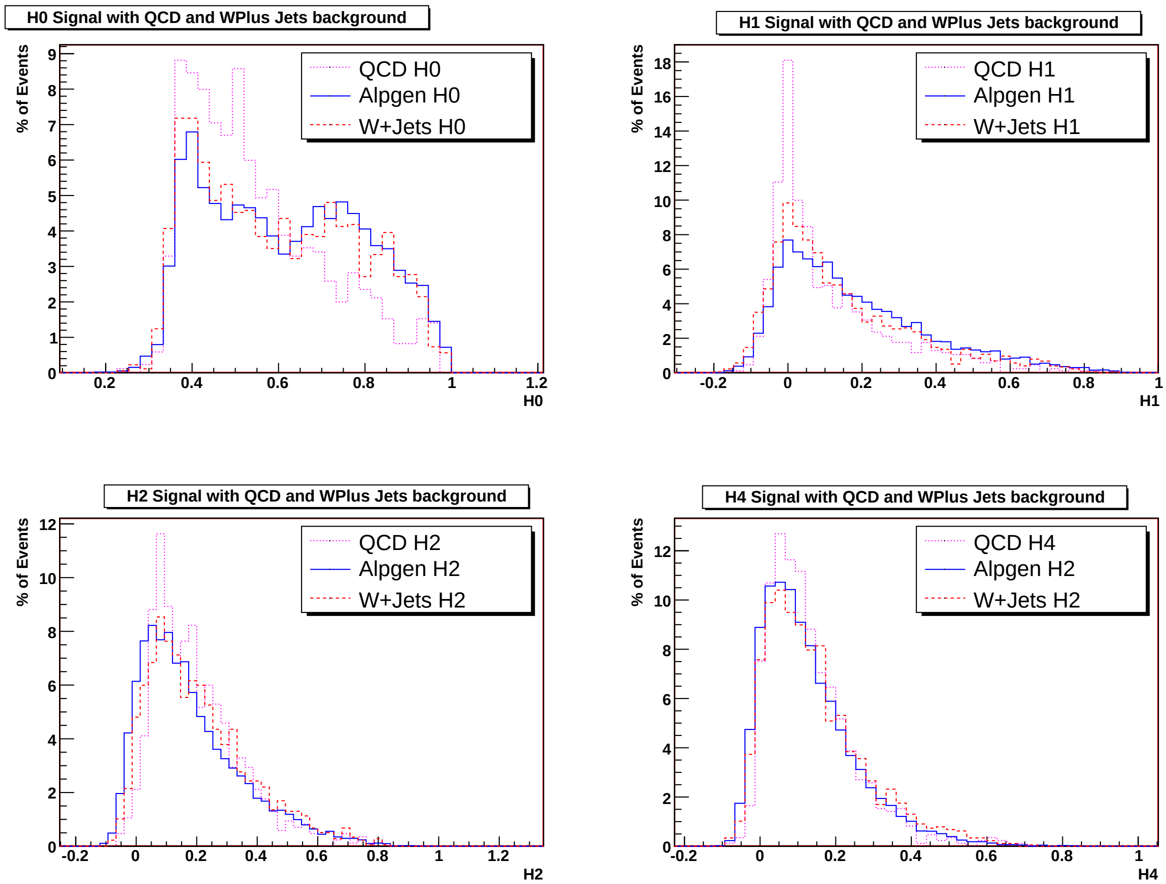

Figure 4: Fox Wolfram Moments with Background signal

All for plots show similar shapes for the event shape variable, hence it can be concluded that Fox Wolfram is not sufficiently useful to separate signal and background data as desired.

3.2 Sphericity

The next event shape variable observed is sphericity. Sphericity can be defined as a measure of the summed transverse moment squared with respect to the event axis [5]. By defining a sphericity tensor Sphericity can be deduced.
Sphericity Tensor:

$$S^{\alpha\beta} = \frac{\sum_i p_i^\alpha p_i^\beta}{\sum_i \mid p_i \mid^2}, \tag{2}$$

with $\alpha, \beta = 1, 2, 3$ a matrix using x, y and z components of the momentum vector of the jets can be constructed. Finding the normalised eigenvalues of these matices the actual sphericity can be constructed. It will be found that to the eigenvalues applies the following to satisfy the normalisation condition: $\lambda_1 \geq \lambda_2 \geq \lambda_3$ and $\lambda_1 + \lambda_2 + \lambda_3 = 1$. Sphericity can be found as a linear combination of these:

$$S = \frac{3}{2}(\lambda_2 + \lambda_3) \tag{3}$$

This formula was then used to evaluate the sphericity of signal data and background data. This can be summarised in the two plots below:

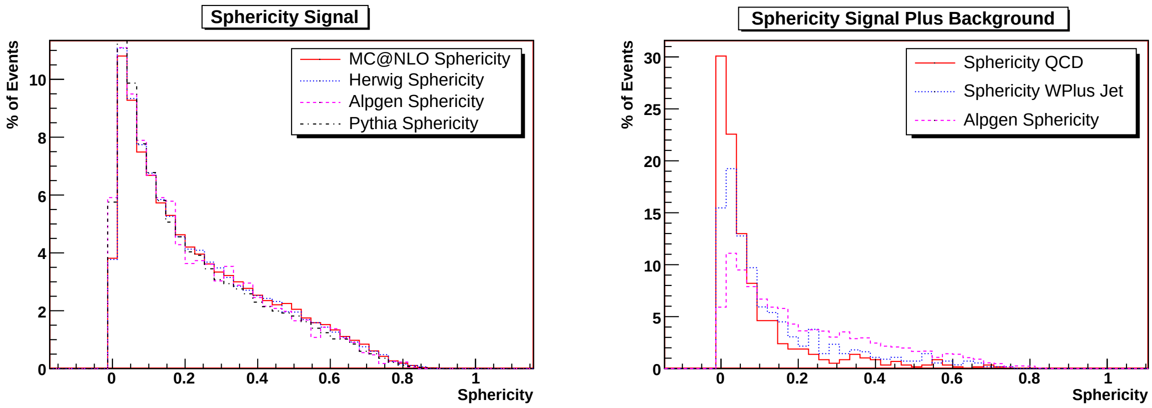

Figure 5: Sphericity

It can be seen, that the signal data is very similar no matter which generator is used for the data production, which was to be expected. Furthermore both QCD and W+Jets background Sphericity behave in a similar way as the sphericity of the signal. The shape of all three graphs in the plot on the right are very similar and therefore will make it hard to use sphericity as an event shape variable to distinguish between background and the actual signal.

3.3 Aplanarity

After sphericity aplanarity was the next event shape variable under investigation. Aplanarity is used to look at different event topologies. It will help distinguish spherical from planar and linear events. Its mathematical definition is 1.5 times the smallest eigenvalue of the momentum tensor(earlier introduced as the sphericity tensor). [6]

Aplanarity:

$$A = \frac{3}{2}\lambda_1 \tag{4}$$

For the aplanarity, again two plots were constructed one containing signal data of all four Monte Carlo generators and the second one plotting the Alpgen signal data against a W+Jets and QCD background. These two graphs display extremely well the fact that signal data is almost identical using different generators. It can also be seen that aplanarity as an event shape variable is not efficient in determining a top event by separating signal and background data, as QCD as well as W+Jets show a very similar pattern to the Alpgen signal.

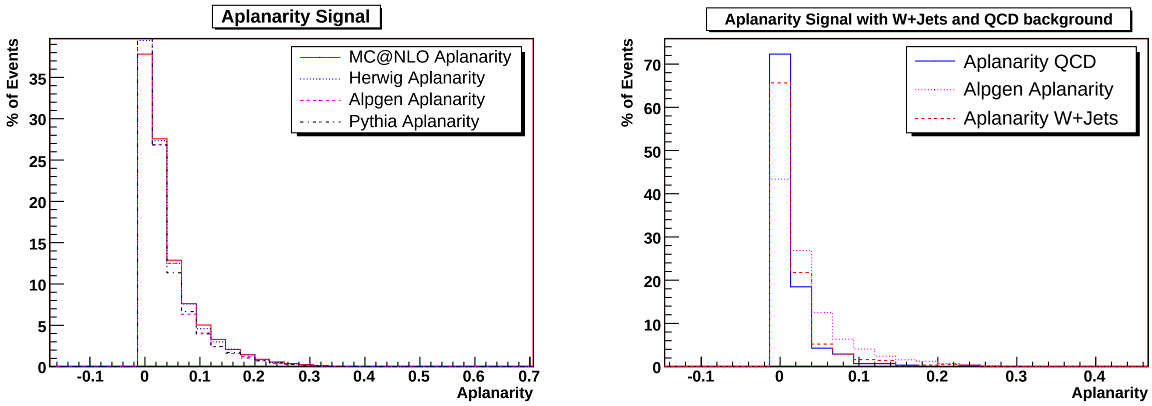

Figure 6: Aplanarity

3.4 Circularity

The next event shape variable used for the generator study was circularity(C). Circularity is a measure of energy spread after the collision. It can again be deduced from the momentum tensor. [7]
Circularity is given by:

$$C = 2\frac{min(\lambda_1, \lambda_2)}{\lambda_1 + \lambda_2} \tag{5}$$

This will again allow the production of two plots one representing the four generator signals and the other one the Alpgen signal with the QCD and W+Jets background. As expected the different signal

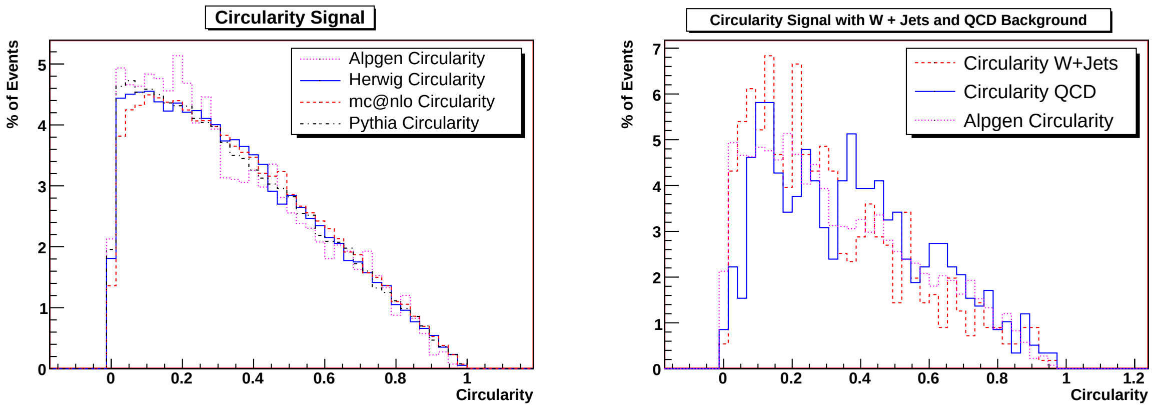

Figure 7: Circularity

data are very similar to eachother. The same is observed with the background data for QCD and W+jets making circularity an inapt event shape variable for the analysis.

3.5 Centrality

The last one of the studied event shapes was centrality(C'). Centrality is not dependent on the momentum tensor as all of the event shape variables before.

Centrality

$$C' = \frac{\sum p_t}{\sum E_{vis}}, \tag{6}$$

where E_{vis} represents the visible energy of all the jets but not the lepton and neutrino of the event, as they are not being considered in the observations.

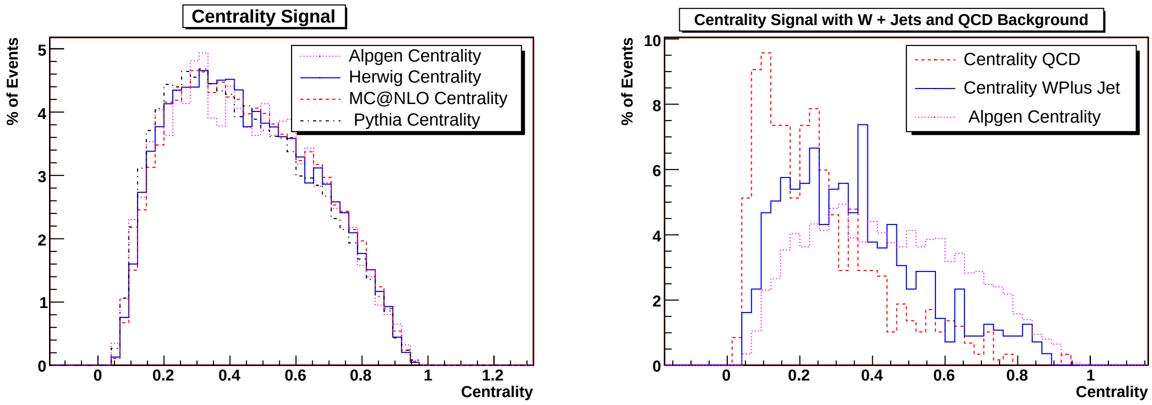

Figure 8: Centrality

This equation could be used to produce two further graphs showing the four centrality signals and the Alpgen signal with the two background signals. The Centrality plot for the signal represents again the expected similar shapes of the different generators. However there can be a difference detected between the background signal of both W+Jets and QCD in respect to the signal data. This means that circularity can be used, to identify a top event from background data.

4 Conclusion

The generator study of different event shapes has revealed some useful information. Using four different Monte Carlo event generators, Alpgen, Pythia, Herwig and Mc@nlo has shown that the production of signal data is indeed independent of the generator used. All studied event shapes showed similar plots for the signal therefore it can be happily said that the expected generator independence is indeed true. For the event shape variables themselves the observations were a little more unexpected. For the fully hadronic decay of the $t\bar{t}$ more than one event shape variable can be used for the identification of such an event from background signals. However it appears, that for the semileptonic decay channel only centrality shows very obvious differences in the signal pattern.

References

[1] http://cmsinfo.cern.ch/outreach/CMSdocuments/DetectorDrawings/DetectorDrawings.html last viewed 14.9.2007

[2] **'The discovery of the Top Quark'**, *Scientific America*, September 1997

[3] Lars Sonnenschein: **'The $t\bar{t}$ production in pp collisions at $\sqrt{s} = 14TeV$'**, Dissertation III. Physikalisches Institut RWTH Aachen Mai 2001

[4] R. D. Fields et al. *Physical Review D*, **5**, 9, *1996*

[5] Torbjoern Sjoerstrand, Stephen Mrenna, Peter Skands: **'Pythis 6.4 - Physics and Manual'**, March 2006

[6] Greenlee H.B., 'The D0 collaboration' **FERMILAB-Conf-93/363-E**

[7] hhtp://atlas.group.shef.ac.uk/physics/nuffield/davesmith.pdf last viewed 13.9.2007