



Universiteit Utrecht

Opleiding Natuur- en Sterrenkunde

Study of event shape observables and jet
production in pp collisions generated by PYTHIA 8
at $\sqrt{s} = 14$ TeV

BACHELOR THESIS

J.N. Schoppink

Supervisors:

MSc. M.H.P.A. SAS
National Institute for Subatomic Physics (Nikhef)

Prof. Dr. T. PEITZMANN
Institute for Subatomic Physics (SAP)

January 15, 2020

Abstract

In this thesis we discuss event shapes of simulated pp collisions as generated by PYTHIA 8, a program for high-energy particle collision simulations developed by Lund University, based on the Monte-Carlo method. We analyze two sets of simulated events (one minimum bias, one with hard only hard QCD processes at $\hat{p}_{T,\min} = 20\text{GeV}$) by calculating a triplet of event shapes: transverse sphericity, sphericity and 3-dimensional sphericity, the correlations between them, and how they perform for different cuts in multiplicity and pseudorapidity η . Subsequently, we will use FastJet3 to find jets in these events and analyze how the existence of one or multiple jets affects performance of the different event shape observables. We conclude that for events with hard QCD processes, all event shape observables are reasonable indicators of jets in an event. Transverse sphericity and sphericity struggle with events with large spread in η , which 3-dimensional sphericity does account for. For minimum bias events, none of the event shape observables are trustworthy indicators of jets in an event, nor the other way around, unless all low-multiplicity events are omitted from the dataset, in which case the performance of all event shape observables is improved significantly.

Contents

1	Introduction	3
2	Theory	4
2.1	Standard Model	4
2.1.1	Bosons	4
2.1.2	Fermions	4
2.2	QCD and Jets	5
2.3	Event shapes	5
2.3.1	Transverse Sphericity	6
2.3.2	Transverse Spherocity	7
2.3.3	3-dimensional sphericity	7
3	Analysis	9
3.1	Datasets	9
3.2	Method	9
4	Results	10
4.1	Transverse sphericity and Multiplicity	10
4.2	Sphericity, Spherocity and 3-dimensional Sphericity	10
4.3	Jets	14
4.3.1	Jet selection on events with hard QCD	14
4.3.2	Pseudorapidity dependence of particles in events with hard QCD	18
4.3.3	Jet selection on minimum bias events	20
5	Conclusion	23
6	Discussion and Outlook	24

1 Introduction

Of the four known fundamental forces that govern our universe, the strongest one, properly called the strong nuclear force, is one which, unlike gravity and the electromagnetic force, most people never realize is such a key part of the structure of the world around us. This is largely because it is relevant at distances of around 1 femtometre ($1 \cdot 10^{-15}$ metres), or about the size of a single proton. It is the force which keeps all matter together, keeps hadrons from falling apart into their constituent quarks, and lets protons and neutrons form nuclei with each other. The strong nuclear force, often referred to as the colour force, is so strong in fact, that if you were to pull apart two quarks (which carry a so-called colour charge) hard enough, eventually the strong force will resist so fiercely that it will become energetically more efficient to simply create a new quark-antiquark pair out of "thin air", such that the original quarks both get a new partner to form a new colour neutral pair with, this effect is called colour confinement.

However, there does exist at least one state of matter so hot and dense that even the colour confinement of the strong force is no match for it and has to let quarks and gluons roam free, the quark-gluon plasma or QGP. The universe is thought to have consisted of QGP in the very first few milliseconds after it formed, before expanding so fast that it quickly cooled down to a temperature where colour confinement took over. This QGP is such an interesting state of matter, it is now the subject of a great part of physical research around the globe.

To this day, the only known way for us to recreate QGP is in heavy-ion collisions in hadron colliders like the LHC or Tevatron. In these experiments, heavy ions, consisting of many protons and neutrons collide at energies high enough to let a tiny droplet of QGP form in the middle of the collision. This droplet only exists for a small time before rapidly expanding and cooling down. However, new research shows that high-multiplicity pp (proton-proton) collisions show properties reminiscent of those observed in heavy-ion events with QGP, suggesting that QGP may even be formed in the relatively small pp collisions. [6]

The most accurate description of the strong interaction between gluons and quarks to date is a theory called Quantum Chromo-Dynamics (QCD). An interesting effect predicted by QCD is the formation of jets, a closely packed, high-momentum stream of particles observed in some collisions as a result of hard scattering process. A jet is produced when a quark or gluon carrying colour charge is ejected with enough energy to create new particles with colour charge in its vicinity in order to keep the overall colour charge neutral. This process may repeat several times and the result is a collection of hadrons all heading in roughly the same direction, called a jet.

A promising way to analyze and learn more about the formation of these jets is to look at event shapes. In this thesis we will cover a triplet of them, transverse sphericity, transverse spherocity, or simply spherocity, and 3-dimensional sphericity. All three of these observables share the property that they approach a value of 1 for sphere-like events and 0 for perfectly pencil-like events. It is often thought that events which score low on these observables, must therefore be a dijet (pencil-like) event. However, there are some downsides to using this logic. Transverse sphericity, for example, only selects pencil-like events in which the jets are back-to-back in the transverse plane, and does not take the pseudorapidity of a particle (its momentum in the direction of the incoming beams) into account. In this thesis we will look at some of the imperfections for each of the observables and discuss what their underlying cause is and how these observables compare to each other.

2 Theory

2.1 Standard Model

The standard model is a gauge quantum field theory which describes all different particles we currently know to make up our universe. It divides these particles into three subgroups, which we will shortly discuss in this section.

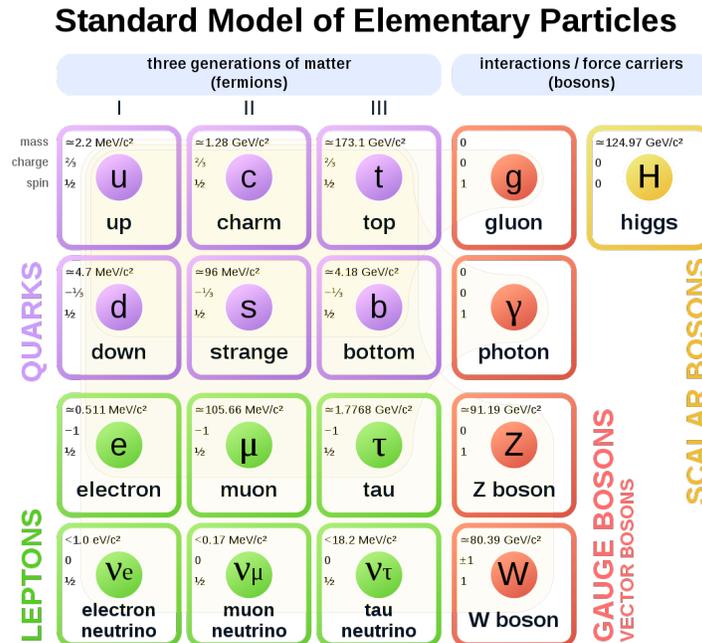


Figure 1: A visual representation of all particles in the standard model, their mass, charge, and spin.

[7]

2.1.1 Bosons

Bosons are particles with integer spin, the most fundamental ones are often referred to as the force carriers, although the higgs-boson is -as the name suggests- also a boson, but not a force carrier. It is called a scalar boson, referring to the fact that the higgs-field is a scalar field. The other four bosons described by the standard model are vector bosons, and these are the force-carriers. They are the mediators of three out of the four fundamental forces. The photon for the electromagnetic force, the gluon for the strong nuclear force and the W- and Z-bosons for the weak nuclear force. Note that the gravitational force is, to this day, not included in the standard model. This, alongside a handful of other deficiencies of the standard model, prevent it from being the first "theory of everything".[1]

2.1.2 Fermions

Fermions are the matter-like particles of the standard model, they are particles with half-integer spin. More formally, fermions are particles which follow fermi-dirac statistics, as opposed to bosons, which follow bose-einstein statistics. Fermions themselves can be split into two subgroups: quarks and leptons. Quarks are the massive particles which make up most of ordinary matter, like protons and neutrons. The leptons are six lighter particles divided up into three generations, electrons, muons, taus, and their corresponding neutrino's. Composite particles consisting of three quarks, called baryons are also fermions themselves.

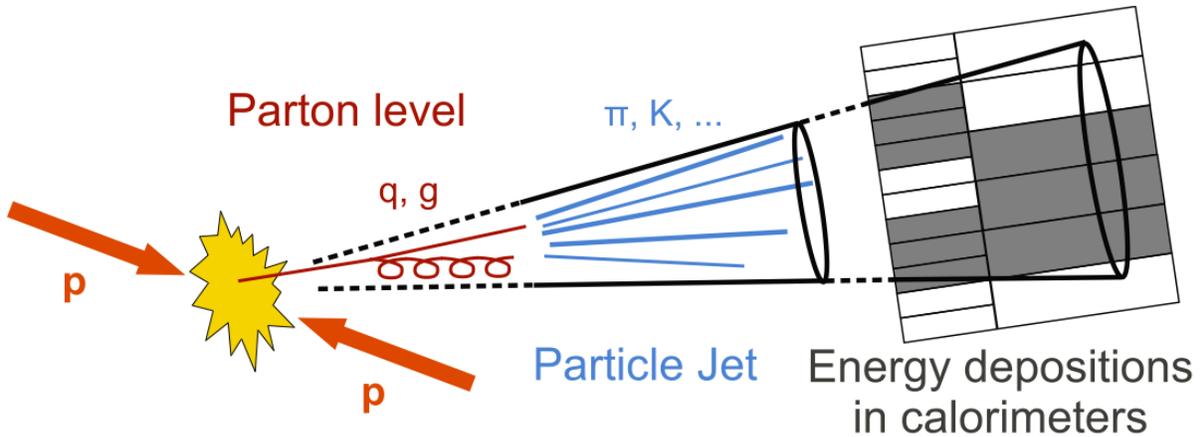


Figure 2: Illustration of jet formation in a proton-proton collision.
[8]

2.2 QCD and Jets

Quantum Chromodynamics is a theory which describes the behaviour of colour charge, a property of particles like gluons and quarks which makes them interact via the strong nuclear interaction. Colour charge is to the strong interaction what electrical charge is to the electromagnetic interaction. It differs in the fact that colour charge comes in three types, as opposed to the two observed in electrical charge (+ and -). The components of colour charge are appropriately named red, green and blue. One other major difference here is that, where electrical charge simply flips for antiparticles, in QCD each colour charge has its own anticolour, bringing the total to six unique colour charges: red, green, blue, antired, antigreen, and antiblue.

This colour charge is the major driving component behind the formation of jets. QCD prevents any free roaming objects from having a net colour charge. All composite particles must have neutral colour charge. Baryons, consisting of three quarks, must therefore have one of each basic colour charge (or anticolour for an antibaryon) to be colour neutral. Mesons are particles which consist of a quark-antiquark ($q\bar{q}$) pair and these must therefore have any colour charge and its corresponding anticolour charge to be colour neutral.

Jets are created in hard scattering QCD processes. These are processes with high momentum transfer. This causes partons (quarks or gluons) to be ejected with high momenta. At low energies, these partons would simply be pulled back in by the strong interaction, because they carry colour charge and can therefore not be let go on their own. What happens in this case however is that the energy that the parton carries is high enough to convert some of that energy into new mass, creating a new quark-antiquark pair such that the ejected parton can keep travelling in the same direction, while remaining colour neutral due to its new companion. This results in groups of particles with relatively high momentum, travelling in the same direction, called a jet.

2.3 Event shapes

The main focus of this thesis will be on event shapes. These are observables which try to separate spherical-like events from pencil-like events by taking on values between 0 and 1, where 1 indicates a perfectly spherical event, meaning that the transverse momentum p_T , i.e. the momentum perpendicular to the axis of the incoming proton beams, is uniformly spread in all directions, and 0 indicates a perfectly pencil-like event, where all of the momentum is directed in two back-to-back, or a single, narrow stream of particles. Because of the cylindrical nature of detectors like LHC, it is useful to express the position of particles

in cylindrical coordinates. PYTHIA 8 provides a particles position in ϕ and η , where ϕ is the azimuthal angle, which indicates the direction in the xy -plane, and η is defined as

$$\eta = -\ln \left[\tan \left(\frac{\theta}{2} \right) \right], \quad (1)$$

where θ is the angle in radians between the momentum vector of the particle of interest and the positive direction of the incoming proton beams.

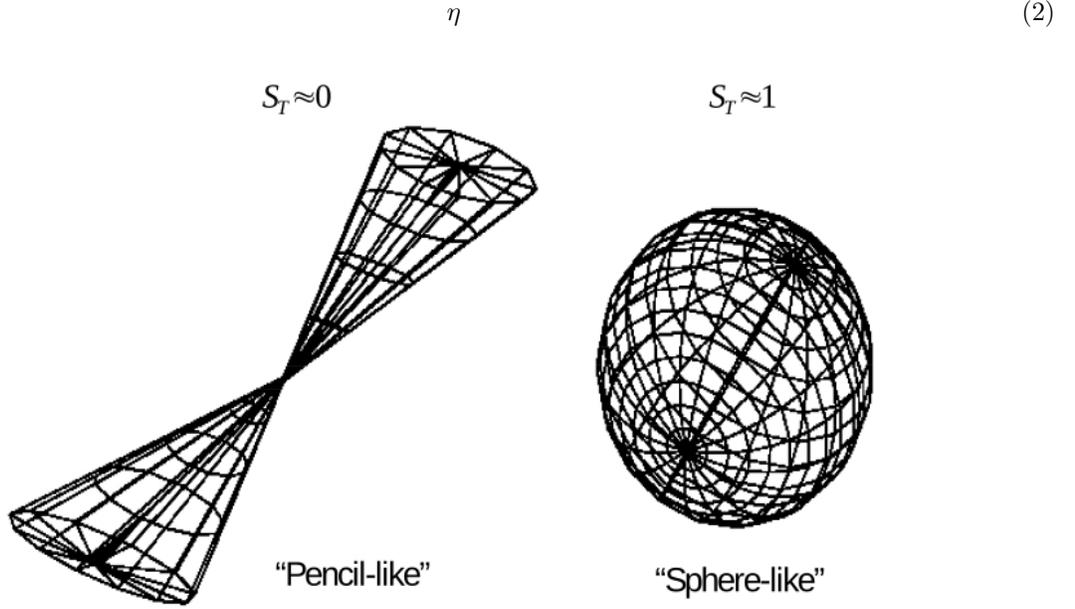


Figure 3: A visualization of a pencil-like event and a sphere-like event, with corresponding values of transverse sphericity.

2.3.1 Transverse Sphericity

The first event shape we will cover is also the most commonly used one, transverse sphericity. To calculate transverse sphericity we must first calculate the eigenvalues of the transverse momentum tensor

$$\mathbf{S}_{xy} = \frac{1}{\sum_i p_{T_i}} \sum_i \begin{pmatrix} p_{x_i}^2 & p_{x_i} p_{y_i} \\ p_{x_i} p_{y_i} & p_{y_i}^2 \end{pmatrix}. \quad (3)$$

However, the fact that this matrix depends on the square of the transverse momentum makes transverse sphericity a collinear unsafe quantity. This means that if we were to split a single particle track with $p_T = \alpha$ into two separate tracks, each with $p_T = \frac{\alpha}{2}$, the resulting contribution to the total would be only half that of the single particle.[2] To circumvent this phenomenon, we need to use the linearized transverse momentum tensor

$$\mathbf{S}_{xy}^L = \frac{1}{\sum_i p_{T_i}} \sum_i \frac{1}{p_{T_i}} \begin{pmatrix} p_{x_i}^2 & p_{x_i} p_{y_i} \\ p_{x_i} p_{y_i} & p_{y_i}^2 \end{pmatrix}. \quad (4)$$

If λ_1 and λ_2 are the eigenvalues of this matrix, and $\lambda_1 > \lambda_2$, then the transverse sphericity is defined as

$$S_T = \frac{2\lambda_2}{\lambda_1 + \lambda_2}. \quad (5)$$

2.3.2 Transverse Sphericity

Transverse sphericity, in the remainder of this thesis simply called sphericity, is a different way of separating sphere-like events from jet-like events, i.e. it measures the same property of events as transverse sphericity does, however the manner in which it does this is quite different.

The transverse sphericity of an event is defined as [3]

$$S_0 = \frac{\pi^2}{4} \min_{\vec{n}_T=(n_x, n_y, 0)} \left(\frac{\sum_i |\vec{p}_{Ti} \times \vec{n}_T|}{\sum_i |\vec{p}_{Ti}|} \right), \quad (6)$$

where p_{T_i} is the transverse momentum of a single particle and \vec{n}_T is a unit vector (a vector of length 1), called the transverse thrust axis, which minimizes the expression. Sphericity is linearly dependent on momentum and therefore collinear safe. To prove this we consider N particles with momenta p_{T_i} , two of which (arbitrary, so let us take particle numbers 1 and 2) are in the same direction. Then, by definition of the modulus of a cross product,

$$S_0 = \frac{\pi^2}{4} \min_{\vec{n}_T=(n_x, n_y, 0)} \left(\frac{|\vec{p}_{T1}| |\vec{n}_T| \sin(\theta_1)}{\sum_i |\vec{p}_{Ti}|} + \frac{|\vec{p}_{T2}| |\vec{n}_T| \sin(\theta_2)}{\sum_i |\vec{p}_{Ti}|} + \frac{\sum_{i=3}^N |\vec{p}_{Ti}| |\vec{n}_T| \sin(\theta_i)}{\sum_i |\vec{p}_{Ti}|} \right), \quad (7)$$

where θ_i is the angle between particle i and the unit vector \vec{n} . Then,

$$S_0 = \frac{\pi^2}{4} \min_{\vec{n}_T=(1, \theta_n, 0)} \left(\frac{|\vec{p}_{T1}| \sin(\theta_1)}{\sum_i |\vec{p}_{Ti}|} + \frac{|\vec{p}_{T2}| \sin(\theta_2)}{\sum_i |\vec{p}_{Ti}|} + \frac{\sum_{i=3}^N |\vec{p}_{Ti}| \sin(\theta_i)}{\sum_i |\vec{p}_{Ti}|} \right). \quad (8)$$

But particles 1 and 2 have the same direction, $\theta_1 = \theta_2$, therefore,

$$S_0 = \frac{\pi^2}{4} \min_{\vec{n}_T=(1, \theta_n, 0)} \left(\frac{\sin(\theta_1) |\vec{p}_{T1+2}|}{\sum_i |\vec{p}_{Ti}|} + \frac{\sum_{i=3}^N |\vec{p}_{Ti}| \sin(\theta_i)}{\sum_i |\vec{p}_{Ti}|} \right). \quad (9)$$

From this final expression we conclude that particles 1 and 2 behave like a single particle with the sum of the momenta of particles 1 and 2 $p_{T1+2} = p_{T1} + p_{T2}$. Thus, we have proven that this definition of sphericity is collinear safe.

2.3.3 3-dimensional sphericity

3-dimensional sphericity is an extension of transverse sphericity to include the z-component of the particles momentum (momentum in the direction of the incoming beams of protons). We begin by taking the sphericity tensor

$$\mathbf{S}^{\alpha\beta} = \frac{\sum_i p_{\alpha_i} p_{\beta_i}}{\sum_i |p_i|^2}, \quad (10)$$

where $\alpha, \beta = 1, 2, 3$ represent the x, y and z components. Diagonalization of this tensor will lead to three eigenvalues $\lambda_1 > \lambda_2 > \lambda_3$, where $\lambda_1 + \lambda_2 + \lambda_3 = 1$. The 3-dimensional sphericity is then defined as

$$S_{3D} = \frac{3}{2} (\lambda_2 + \lambda_3), \quad (11)$$

leading to a value of 1 in events where the momentum is spread out equally amongst all three cartesian axes and a value of 0 when all momentum is focused in a single, or two back-to-back jets. However, this expression is, again, collinear unsafe. In this case we can circumvent this by rewriting the expression as [9]

$$\mathbf{S}^{\alpha\beta} = \frac{\sum_i |p_i|^{2-r} p_{\alpha_i} p_{\beta_i}}{\sum_i |p_i|^r}, \quad (12)$$

where, for all analyses in this thesis, we have set $r = 1$. Note that Eq. 10 is simply the case where $r = 2$. This resulting expression (with $r = 1$) now very much resembles the expression we used for the

transverse sphericity, but it is now a three-dimensional matrix. We can then use the eigenvalues of this expression combined with Eq. 11 to calculate the 3-dimensional sphericity.

3 Analysis

3.1 Datasets

All events studied in this thesis are generated by PYTHIA 8.2, a monte-carlo event generator created by Lund University for simulating high-energy particle collisions.

PYTHIA 8 takes an input of two incoming particles, their center-of-mass energy, and a range of options, like generating only events with hard QCD processes or minimum bias. The output it delivers consists of a list of particles, both final state and mother particles. For each of these particles the following information is given:

- The particle ID according to the monte-carlo particle numbering scheme. [4]
- The particle status, indicating how a particle was produced, i.e. where in the program execution it was inserted into the event record and why. This also indicates whether the particle is a final state particle. [5]
- Amount of mother- and daughter-particles
- The colour code, referring to the particles colour charge, 0 if the particle is colour neutral.
- The particles mass, energy and momentum-vector.

For more information on PYTHIA and its components, see [5].

All of this information is then loaded into ROOT 6, the C++ based industry standard data analysis program, developed by CERN. Using ROOT we can make selections in the data and visualize it. ROOT also enables us to analyze possible jets in the data by using a jetfinder such as FastJet 3 as a plugin to ROOT.

We have generated two datasets with PYTHIA, one with hard QCD at $\hat{p}_{T,\min} = 20\text{GeV}$, which means that in every event there is at least one hard QCD interaction with a momentum transfer of 20GeV. The other dataset is run with minimum bias, i.e. $\hat{p}_{T,\min} = 0$, meaning we will generate significantly less jets in this dataset, however, it is a better representation of real data gathered at hadron colliders like the LHC.

3.2 Method

For this thesis we have used PYTHIA 8 to generate two sets of proton-proton collisions at $\sqrt{s} = 14\text{TeV}$. One set of 20.000 events with minimum bias, and one set of 40.000 events with hard QCD, $\hat{p}_{T,\min} = 20\text{ GeV}$.

For all analyses in this thesis we have made the following particle selections. Only particles with a transverse momentum of $p_T > 0.5\text{ GeV}$ are selected. Of those, only the final-state particles are selected, i.e. photons, charged pions, charged kaons, electrons, protons and any corresponding antiparticles. Furthermore, we only select particles with pseudorapidity $|\eta| < 1$. Then for all remaining events we only accept events with at least 3 final state particles, because events with only 1 or 2 (back-to-back) particles will by definition have 0 sphericity.

All jets in this thesis have been found using FastJet 3 [10]. PYTHIA 8 does provide three jetfinders of its own, however FastJet 3 is more complete and therefore more suitable for more complex studies of jets, hence our decision for using it. All jets are found using the anti-kt algorithm with $R = 0.4$ and minimum jet- p_T of either 10 GeV or 15 GeV which will be specifically indicated for each figure or calculation.

4 Results

In this section we will present the results obtained from analyzing the pp collisions simulated by PYTHIA 8. The first subsection will be on the effect of multiplicity cuts on the value of transverse sphericity. The second will be on the correlations between the different event shapes and in the final part we will present what happens when we select only events with a certain amount of jets.

4.1 Transverse sphericity and Multiplicity

It is important to realize that multiplicity, i.e. the number of final state particles produced in a collision, has an effect on event shapes and can therefore not be ignored. In figure 4 the transverse sphericity distribution is shown for several cuts in multiplicity. It can be clearly seen that events with a low multiplicity (3 - 10 particles) have a lower average sphericity (Mean = 0.4888 σ = 0.2356), than high multiplicity (>30 particles) events (Mean= 0.6282, σ = 0.1857). This is logical because the chance of all particles having momentum in the same direction decreases as the number of particles increases. This result will become very relevant in section 4.3.3.

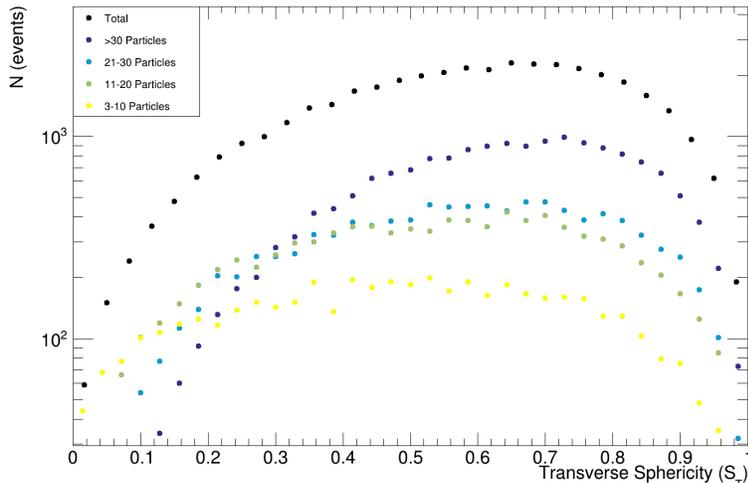


Figure 4: The transverse sphericity distributions of simulated pp collisions for several cuts in multiplicity. The events are simulated by PYTHIA 8 at $\sqrt{s} = 14$ TeV, with Hard QCD, $p_{Tmin} = 20 GeV$. Multiplicity cuts made are $3 \leq N \leq 10$, $10 < N \leq 20$, $20 < N \leq 30$, and $N > 30$, where N is multiplicity.

4.2 Sphericity, Spherocity and 3-dimensional Sphericity

In figure 5 we see the total distributions for transverse sphericity, spherocity, and 3-dimensional sphericity. It is interesting to see how the distributions for transverse sphericity and 3-dimensional sphericity look very similar in shape and mean value, whereas spherocity has a significantly lower mean and a different shape overall. However, as can be seen in figure 6, it is actually transverse sphericity and spherocity which show the cleanest correlation.

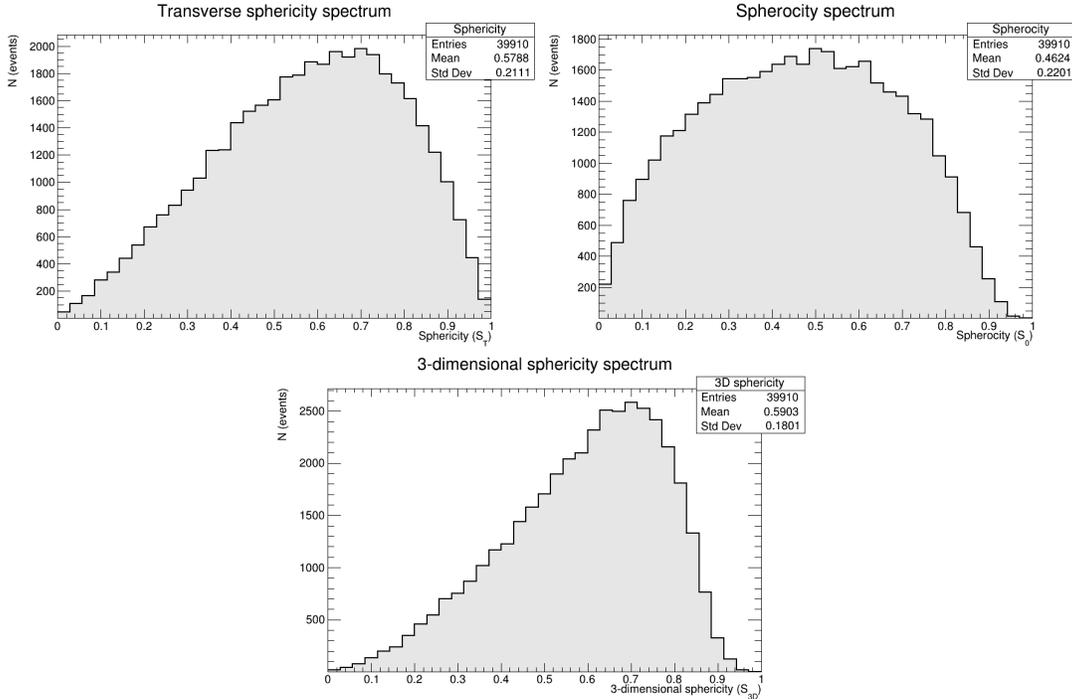


Figure 5: The total distribution for transverse sphericity, sphericity, and 3-dimensional sphericity of 40,000 hard QCD events ($\hat{p}_{T,\min} = 20$ GeV) generated by PYTHIA 8. Only final state particles with $|\eta| < 1$ and $p_t > 0.5$ GeV are accepted. Only events with more than two final state particles are considered.

In figure 6a we can see the response matrices for all combinations of sphericity, sphericity and 3-dimensional sphericity and in figure 6b their corresponding profiles (Average value and root mean square for "y" in each bin of "x"). It is clear that transverse sphericity and sphericity are nicely correlated, where the value of transverse sphericity is a bit higher for almost all events, which is to be expected from looking at the total distributions of the event shapes and their mean values in figure 5. It also has less extreme outliers, in contrast to the two histograms which include 3-dimensional sphericity, where a small number of events can be seen with high 3-dimensional sphericity, but low transverse sphericity, and the other way around. The momentum distribution in (η, ϕ) -space of a few examples of these extraordinary events are displayed in figures 7 and 8. The following paragraphs will be dedicated to analyzing these events, by looking at their structure, and deriving from their structure what unique properties they have that gives rise to this behaviour. Please note that these are only a handful of events from a set of 40,000, therefore, they are by no means common, but we will plot and analyze them in the hopes of learning more about the behaviour of our event shape observables in extreme cases.

It is clear from figures 7 and 8, which show the p_T distribution of single events, that events with only a few particles, which have a relative angle of either 0 or π in ϕ , but have an angle of approximately $\frac{\pi}{2}$ in η result in relatively high 3-dimensional sphericity but an extremely low transverse sphericity. This is logical, because transverse sphericity does not take their difference in η into account. It interprets these events as two almost perfectly back-to-back streams/jets of particles in the transverse plane. 3-dimensional sphericity however, does consider this angle in η and therefore interprets these events as being much more sphere-like than they appear to be in the transverse plane.

For the events with high transverse sphericity and low 3-dimensional sphericity it can be seen that they are characterized by a homogeneous spread of momentum in ϕ , with a distance of approximately $\phi = \frac{2\pi}{3}$ between the (groups of) particles, making them appear very sphere-like in the transverse momentum plane. However, when we take into account their z-component, we see that they lie in a plane. This

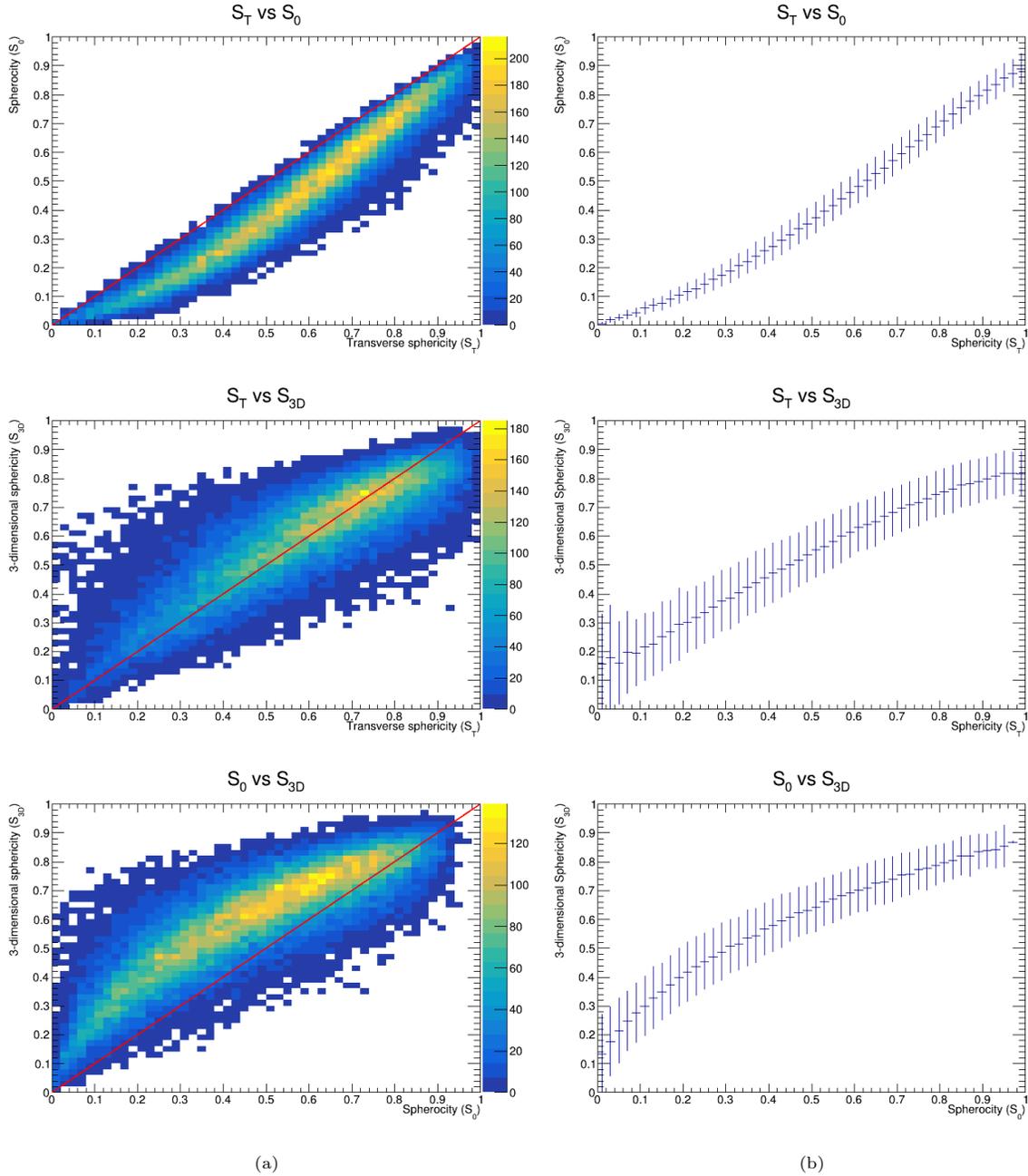


Figure 6: 2D histograms (a) and their TProfiles (b) for all combinations of transverse sphericity, sphericity and 3-dimensional sphericity. The red line is simply the diagonal line and only included for visual clarity. All events are pp collisions at $\sqrt{s} = 14$ TeV simulated by PYTHIA 8 with hard QCD, $\hat{p}_{T,\min} = 20$ GeV.

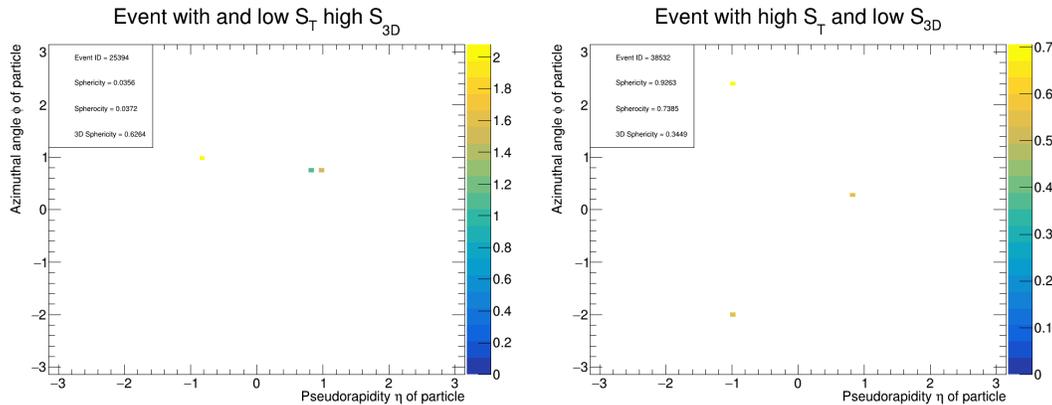


Figure 7: Two event displays of events with unexpectedly high differences in S_T and S_{3D} . Events are pp collisions at $\sqrt{s} = 14$ TeV simulated by PYTHIA 8 with hard QCD, $\hat{p}_{T,\min} = 20$ GeV. Each particle is weighted with its p_T , indicated by the colour scale.

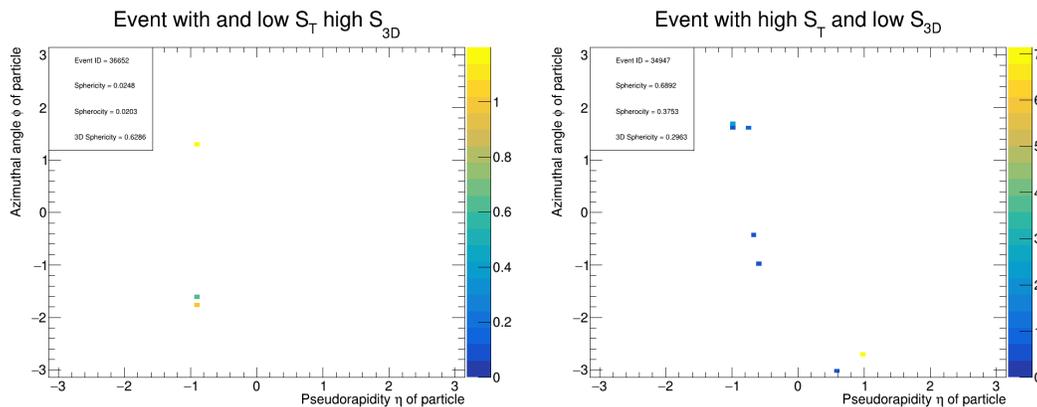


Figure 8: Two event displays of events with unexpectedly high differences in S_T and S_{3D} . Events are pp collisions at $\sqrt{s} = 14$ TeV simulated by PYTHIA 8 with hard QCD, $\hat{p}_{T,\min} = 20$ GeV. Each particle is weighted with its p_T , indicated by the colour scale.

means that one component in our 3-dimensional space is completely void of momentum. This leads to the relatively low score in 3-dimensional sphericity.

For comparison, in figure 9 two much more ordinary events are displayed; one event with low S_{3D} and low S_T (left), and one with high S_{3D} and high S_T (right). We can see that events with low values on all observables are very jet-like events. Events with high values on all observables have a very homogeneous spread of the total p_T over the entire space, as we would expect.

We conclude this part on the outliers in the response matrices of 6a by noting that the events with high discrepancies in S_T and S_{3D} are almost exclusively low-multiplicity events. This increases the chance of the few particles that are present to have momenta at just the right angles to cause these unexpected observable values. It is safe to say that these outliers are simply a result of statistics and are therefore to be expected in large sample sizes.

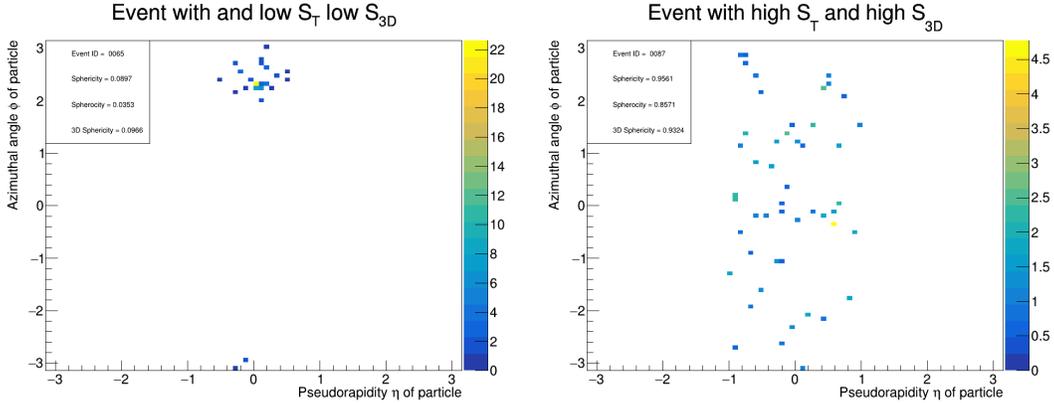


Figure 9: Two events with roughly equal values for S_T and S_{3D} . Events are pp collisions at $\sqrt{s} = 14$ TeV simulated by PYTHIA 8 with hard QCD, $\hat{p}_{T,\min} = 20$ GeV. Each particle is weighted with its p_T , indicated by the colour scale.

4.3 Jets

4.3.1 Jet selection on events with hard QCD

In this section and section 4.3.3, we will discuss the effect that the presence or absence of jets has on the different event shape observables. All event shape observables described in this thesis are created and used as tools to separate pencil-like events from sphere-like events. It is therefore not unreasonable to think that, considering conservation of momentum, events with two jets should be pencil-like and therefore have low values on all our observables. However, as we can see in figure 10, this is not always the case. Although events with one or two jets are definitely more likely to be pencil-like, there still exists a relatively large amount of events containing two jets which are very sphere-like.

To put this into numbers, in table 1 we can see, for every observable and events sorted by their amount of jets, the fraction of those events for which the value of the observable is lower (more pencil-like) than that observable's mean value, giving an indication of how good these observables are at separating events with different amounts of jets from each other. If our observables perform well, we expect to see higher values in the rows "1 Jet" and "2 Jets" and lower values in the rows "0 Jets" and "3 Jets". We conclude that 3-dimensional sphericity is best at finding events with jets, as we expect events with 1 or 2 jets to have below average values on the event shape observables, and 3-dimensional sphericity gives the highest fraction of events with below average 3-dimensional sphericity for events with 1 or 2 jets. Keep in mind that the dataset used in this subsection consists of only events with hard QCD processes with $\hat{p}_{T,\min} = 20$ GeV. For events with soft QCD (minimum bias events), the table looks very different and will be discussed in section 4.3.3.

Fraction of events with observable value lower than its mean			
	Sphericity	Spherocity	3D-Sphericity
0 Jets	0.3234	0.3356	0.2985
1 Jet	0.6318	0.6835	0.7095
2 Jets	0.6270	0.6349	0.6247
3 Jets	0.5688	0.5792	0.5480

Table 1: Fractions of events with observable value lower than that observable's mean for events with hard QCD ($\hat{p}_{T,\min} = 20$ GeV). Only considering events with multiplicity $N \geq 3$. Minimum jet- $p_T = 10$ GeV. Mean values used are $\langle S_T \rangle = 0.5788$, $\langle S_0 \rangle = 0.4624$, $\langle S_{3D} \rangle = 0.5903$.

In figure 10 we see the effect of jet-selection on the event shapes visualized in six diagrams. Note

that they are all logarithmic in the vertical axis. It can be seen that for all event shapes the selection of events with zero jets has an average far above the total average. For selections of one and two jets, the opposite is the case. What is interesting to see here is that, as we may also have concluded from table 1, for all event shape observables, the average value for a selection of events with a single jet is lower than that of events with two jets. Combined with our knowledge of momentum conservation, this suggests that there are a lot of events in which we only see two jets, but they are still not pencil-like because there may be a third jet, of which we are not aware, which causes the other jets to be at an angle instead of back-to-back. There may also be cases in which two jets are out of our range of measurement and we only observe a single jet, which is then interpreted as a highly pencil-like event, contributing to the low mean in single-jet events. As we will see in section 4.3.2, by taking the widely accepted pseudorapidity range of $|\eta| < 1$ we only measure approximately 30 percent of all jets in the events.

Another way to visualize this is to take the response matrices from figure 6a and subject them to jet-selection. Which is shown in figure 11. Here we see the response matrices for all combinations of event shape observables but the events are sorted by their amount of jets (with minimum jet- p_T of 10 GeV). The shift in average for all event shapes observables for events with one jet as opposed to events with zero jets can be clearly seen. However the difference between 1-jet and 2-jet events is marginal.

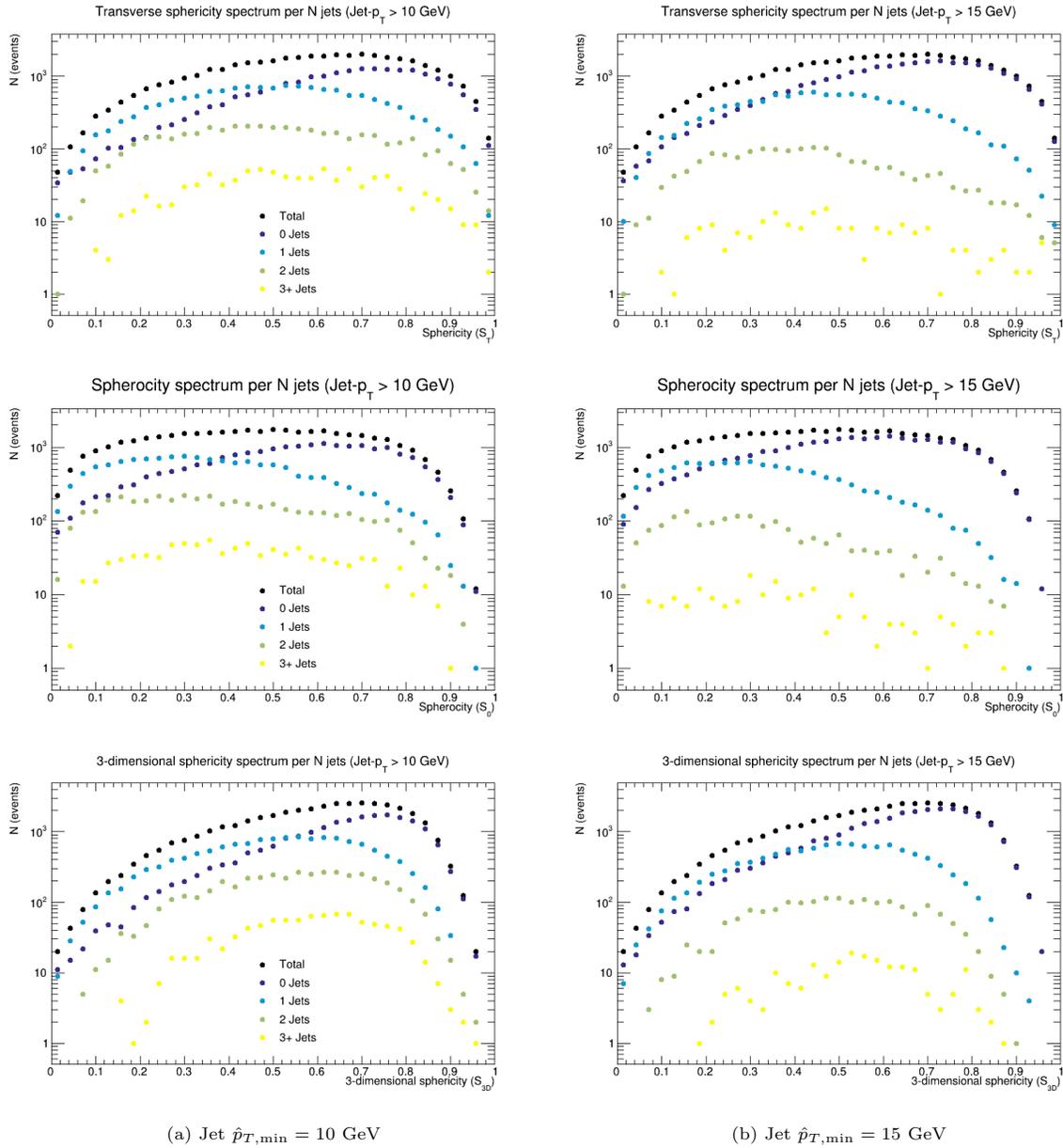


Figure 10: Sphericity, spherocity and 3-dimensional sphericity distributions for 40,000 simulated pp collisions generated by PYTHIA 8, sorted by amount of jets. Jets found with FastJet3, $\hat{p}_{T,\min} = 10$ GeV on the left, $\hat{p}_{T,\min} = 15$ GeV on the right. Legend is applicable to both figures.

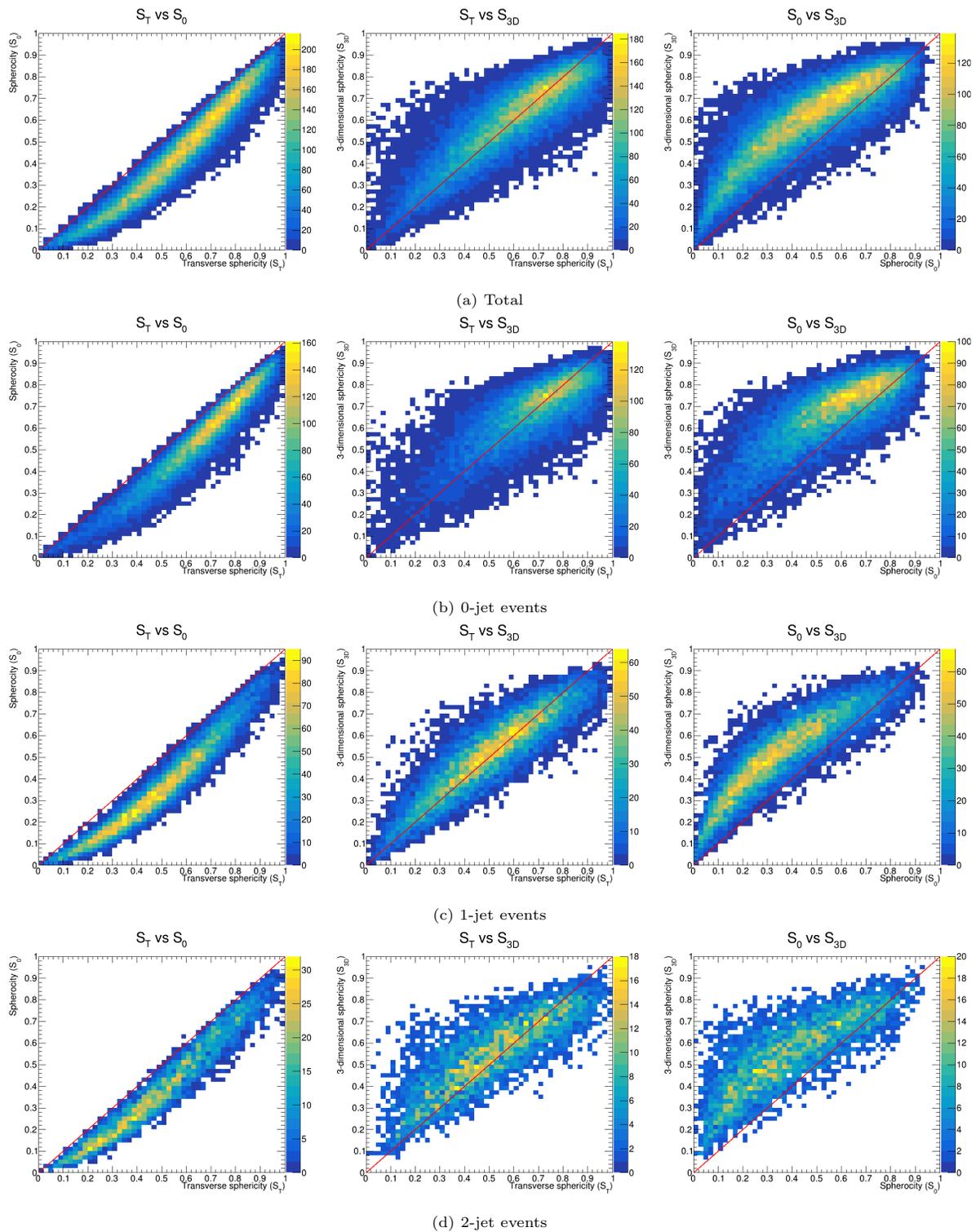


Figure 11: 2D histograms for all combinations of transverse sphericity, sphericity and 3-dimensional sphericity. Events are selected by their amount of jets with minimum $\text{jet-}p_T$ of 10 GeV. The red line is simply the diagonal line and only included for visual clarity. Events are pp collisions at $\sqrt{s} = 14$ TeV generated by PYTHIA 8 with hard QCD, $\hat{p}_{T,\text{min}} = 20$ GeV.

4.3.2 Pseudorapidity dependence of particles in events with hard QCD

When looking at previous results, we have reason to believe we may be missing a lot of jets that are outside of our pseudorapidity selection of $|\eta| < 1$. It may be interesting to find out just how many jets we are missing. But first let us look at one more interesting phenomenon which points us in this direction. We will, recreate the response matrices of figure 6a, but this time we will also accept particles with very high pseudorapidity, say $|\eta| < 5$. The resulting response matrices are shown in figure 12. Now, we can see that transverse sphericity and sphericity, both of which only depend on transverse momentum are not very much affected by this. If anything, their average increases a little. 3-dimensional sphericity, on the other hand, is displaying some exceptional behaviour. The average value of 3-dimensional sphericity drops to an extremely low value. This behaviour could be explained by many high- p_T jets with angles close to the incoming proton beams, which incites us to further investigate this possibility.

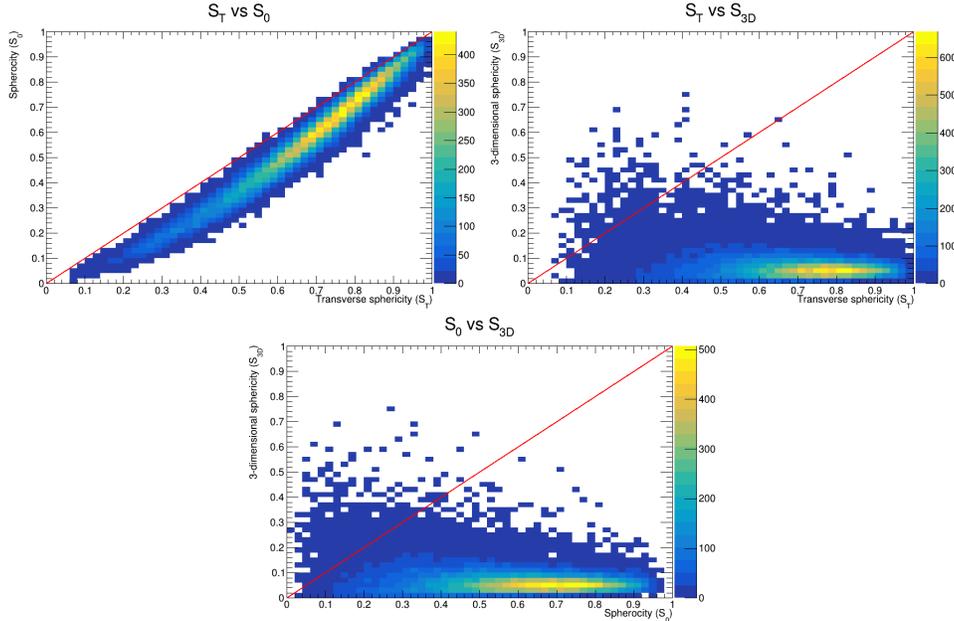


Figure 12: 2D histograms for all combinations of transverse sphericity, sphericity and 3-dimensional sphericity, for events with pseudorapidity $|\eta| < 5$. The red line is simply the diagonal line and only included for visual clarity. Events are pp collisions at $\sqrt{s} = 14$ TeV simulated by PYTHIA 8 with hard QCD, $\hat{p}_{T,\min} = 20$ GeV.

In figure 13 we see the distributions of high p_T jets ($p_T > 20$ GeV) over η and ϕ . The distribution in terms of η seems to be fairly logical. However, pseudorapidity is not linear in θ , meaning that the spherical surface enclosed within the interval $0 < \eta < 1$ is much smaller than, for example the area enclosed within the interval $2 < \eta < 3$. To overcome this counterintuitive idea we write θ in terms of η as

$$\theta = \arctan e^{-\eta}. \quad (13)$$

Using this expression we can transform figure 13a to figure 13b. Note that the vertical axis in figure 13b is logarithmic. In this figure we can see how uneven the distribution of high p_T jets really is. For almost the entirety of this thesis we have selected only events with pseudorapidity of $|\eta| < 1$, which corresponds to a spherical surface of

$$\int_0^{2\pi} d\phi \int_{\arctan(e^{-1})}^{\arctan(e^1)} \sin \theta d\theta \approx 9.57, \quad (14)$$

where we assumed radius $r = 1$. Whereas the surface of the entire sphere is, of course, 4π . The fraction of the total spherical surface covered by taking $|\eta| < 1$ is then approximately

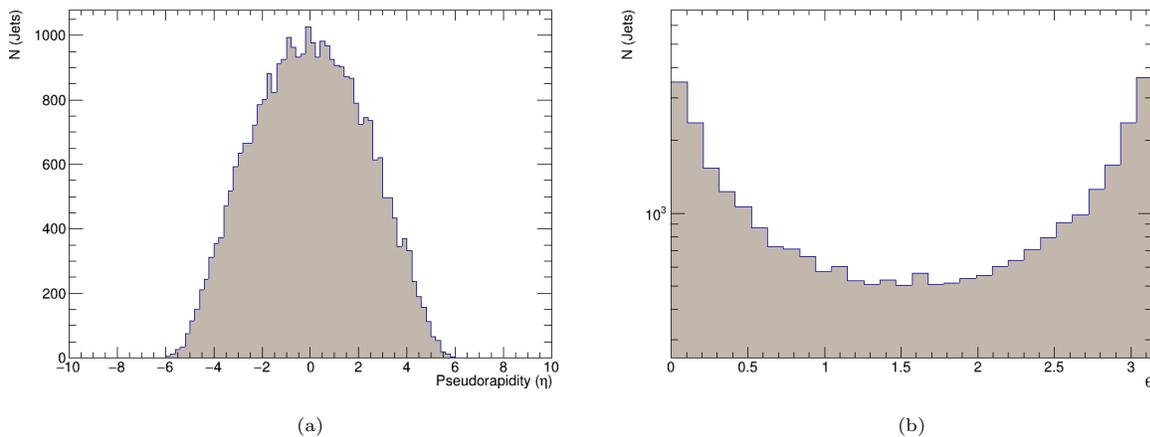


Figure 13: The distribution of jets with $p_T > 20\text{GeV}$ in η (a) and θ (b)

$$\frac{9.57}{4\pi} \approx 0.76. \quad (15)$$

From the data of figure 13 it can be calculated that approximately 30 percent of high p_T jets lies in this area. This would mean that approximately 70 percent of high p_T jets lie in an area which is less than 25 percent of the surface of the sphere, leading to a much higher jet-density in the circular area straddling the incoming beam axis. This might explain the unexpected behaviour of 3-dimensional sphericity in large pseudorapidity cuts, observed in figure 12.

It should be noted that all of the above can also be said for the total of all particles in events with hard QCD interactions. The total of all particles shows a similar distribution in η and therefore also in θ . In figure 14 we can see the amount of high- p_T jets per the amount of particles detected in each bin of θ and we can see that it fluctuates but stays roughly the same over the entire space. Therefore we can conclude from this section that the amount of high- p_T jets detected per the amount of particles is consistent throughout the space. However, the amount of particles (and therefore jets) is heavily centered around the beam axis.

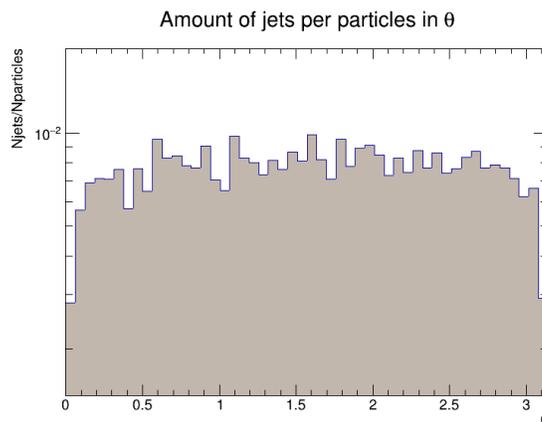


Figure 14: The amount of jets with $p_T > 20\text{GeV}$ per the amount of particles in theta.

4.3.3 Jet selection on minimum bias events

We have recreated the table we made for hard QCD events (table 1) for events with minimum bias, which can be seen in table 2, where we should note that there were only two events with three jets in this dataset, which explains the zeroes in the bottom row. In table 2 we see that for minimum bias events, all three of our event shape observables perform either very bad or not at all when it comes to separating events with jets from events with no jets. This is unexpected, because one would say that even when there are very few jets in the total dataset, the events which do have one or two jets should still be pencil-like. However, as we can see, this is not the case. Let us now take a closer look at some of these events to see what their structure is and what may cause this ineffectiveness of the event shape observables.

Fraction of events with observable value lower than its mean			
	Sphericity	Sphericity	3D-Sphericity
0 Jets	0.4621	0.5017	0.4553
1 Jet	0.4754	0.5130	0.4173
2 Jets	0.4286	0.5	0.3929
3 Jets	0	0	0

Table 2: Fractions of events with observable value lower than that observable’s overall mean for events with soft QCD (minimum bias). Only considering events with multiplicity $N \geq 3$. Minimum jet- $p_T = 10$ GeV. Mean values used are $\langle S_T \rangle = 0.541$, $\langle S_0 \rangle = 0.4198$, $\langle S_{3D} \rangle = 0.5487$.

One might argue from table 2 that sphericity, even though its performance is still weak to non-existent, is our least weak observable when it comes to separating jetty events from events without jets for minimum bias events. Let us therefore look at two events with two jets and high sphericity. See figure 15. It appears that these are two events with such a high multiplicity that the whether or not the jets that our jetfinder finds in these events are actually jets or not, which is sometimes questionable, they are overshadowed by the many other particles heading off in other directions.

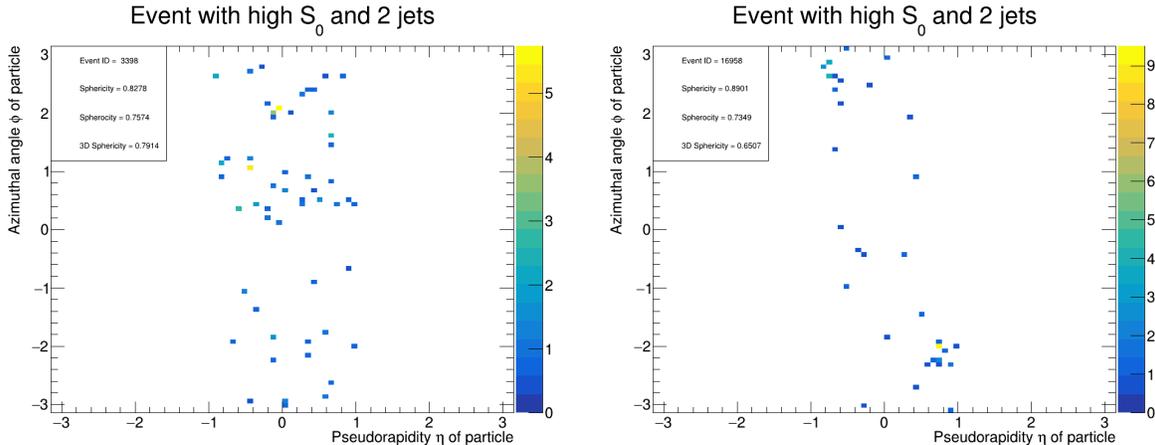


Figure 15: Two event displays of minimum bias events with high sphericity and multiple jets. Events are pp collisions at $\sqrt{s} = 14$ TeV simulated by PYTHIA 8. Each particle is weighted with its p_T , indicated by the colour scale

Now there are of course also minimum bias events with two jets that also have the expected low sphericity, two of which are shown in figure 16. These events are more reminiscent of what we would expect from an event with two jets, namely two compact clusters of particles with an angle in ϕ between them of approximately $\Delta\phi = \frac{\pi}{2}$. However, as can be concluded from the ‘Sphericity’ column of table 2, events such as the ones displayed in figure 15 are more common than the ones displayed in figure 16. More

specifically, we calculated from our data that events with one or more jets have an average of 33 final state particles, compared to an average of 10 particles over all minimum bias events. The production of jets is therefore strongly correlated to multiplicity. (For comparison, in the data with hard QCD events, the average multiplicity is approximately 39 for events with jets, and 29 for events without jets. Note that we have omitted all events with 1 or 2 particles in both calculations). This has as a result that for minimum bias events, event shape observables such as transverse sphericity, sphericity and 3-dimensional sphericity, are not good indicators of whether an event has jets or not. Also, the appearance of jets in a minimum bias event says little to nothing about the value of that event for the event shape observables. If anything, events with two jets are likely high multiplicity events are therefore sphere-like, as we have concluded in section 4.1.

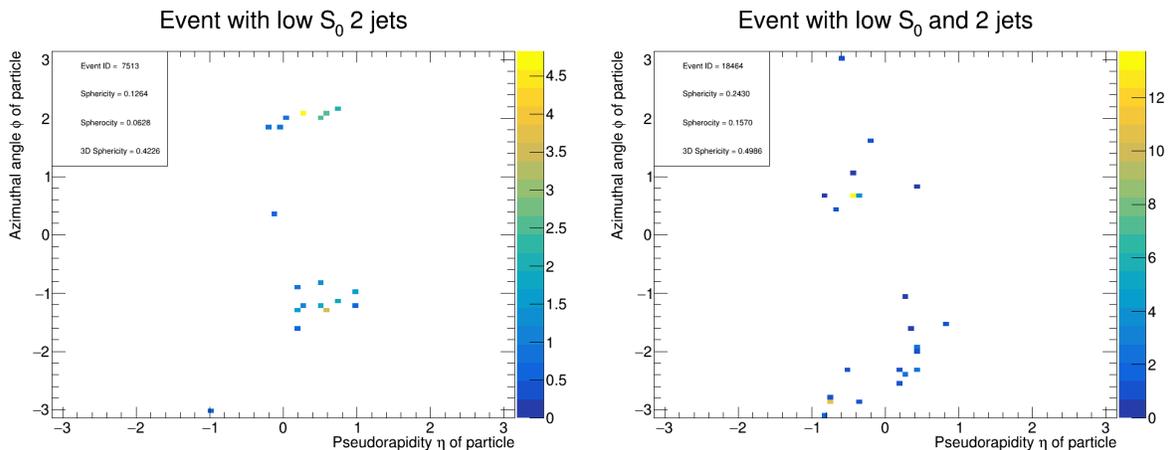


Figure 16: Two event displays of minimum bias events with low sphericity and multiple jets. Events are pp collisions at $\sqrt{s} = 14$ TeV simulated by PYTHIA 8. Each particle is weighted with its p_T , indicated by the colour scale

But what could cause this strong correlation between the amount of jets and multiplicity which is so much stronger in minimum bias events than in events with hard QCD? Let us look at the total multiplicity distributions for both datasets and the distributions for events which have jets. As we can see in figure 17, our dataset with minimum bias events contains very many events with low multiplicity. Approximately half of the events in this dataset has a multiplicity of 5 or lower. Now when we judge our event shape observables on how well they separate events with jets from events without jets, we look at how much the average shifts to the right for events without jets (which we expect to have high sphericity, etc.) and how much the average shifts to the left for events with one or two jets (which we expect to have low sphericity, etc). But, because of the large amount of low multiplicity events without jets in minimum bias data, the overall average sphericity for events without jets gets lowered significantly, and the sphericity for events with jets is relatively high. This counteracts the aforementioned effect which we use to measure the effectiveness of our event shape observables to such an extent, that they give little to no useful information.

With this in mind we have remade table 2 once again in table 3, but this time we have dismissed all events with multiplicity of 18 or lower. Because, beyond this point the multiplicity distribution for minimum bias somewhat resembles that of hard QCD, see figure 17. We can see that with this cut in multiplicity, all of our event shape observables perform excellently. All of our event shapes indicate that more than 85 percent of events with jets score below average on our event shape observables, which is exactly what we would expect. Note: There are only 17 events with two jets in this dataset, one of which has above average value on all event shapes, leading to the recurring value of 0.9412 in the row "2 Jets".

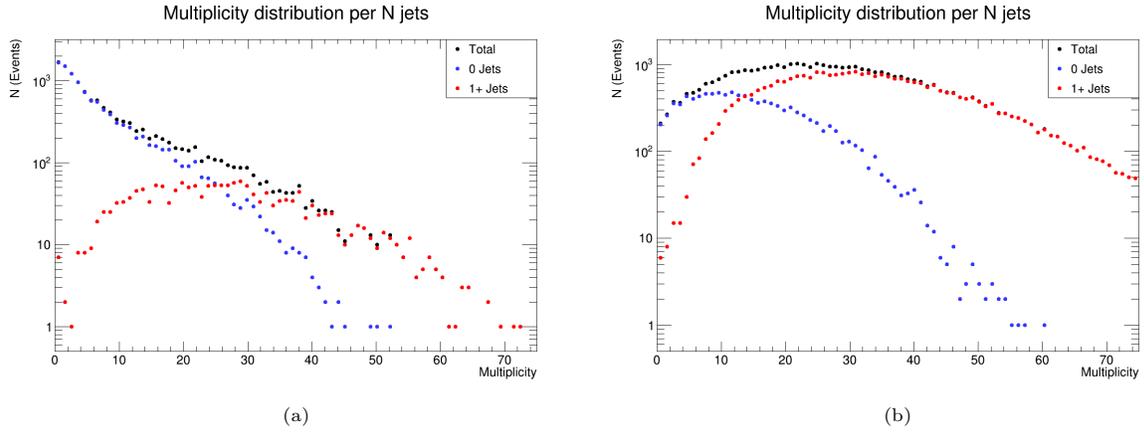


Figure 17: The multiplicity distributions per amount of jets with $p_T > 5\text{GeV}$ for min. bias events (a) and events with hard QCD (b)

Fraction of events with observable value lower than its mean			
	Sphericity	Sphericity	3D-Sphericity
0 Jets	0.4498	0.4634	0.4323
1 Jet	0.8566	0.875	0.8640
2 Jets	0.9412	0.9412	0.9412
3 Jets	0	0	0

Table 3: Fractions of events with observable value lower than that observable's overall mean for events with soft QCD (minimum bias). Only considering events with multiplicity $N > 18$. Minimum jet- $p_T = 10\text{ GeV}$. Mean values used are $\langle S_T \rangle = 0.6966$, $\langle S_0 \rangle = 0.5938$, $\langle S_{3D} \rangle = 0.701$.

5 Conclusion

In this thesis we have analyzed the behaviour of a triplet of event shape observables: transverse sphericity, transverse spherocity and 3-dimensional sphericity. We have analyzed how they perform when we use them to find jets in an event, but also how they correlate with each other upon selecting only events with or without jets. We have also tried to find and report the strong and weak points of each event shape observable. Furthermore, we have looked at the effect of multiplicity cuts on transverse sphericity, the effect of jet selection on multiplicity, and the distribution of particles and jets in η .

First of all it can be concluded, like it has been in many other instances, that the average transverse sphericity increases with multiplicity.

For the events generated with hard QCD processes, all three event shapes are reasonable, although far from decisive, indicators of the presence of jets in an event. 3-dimensional sphericity seems to be best for separating events with jets from events without jets.

For events generated with minimum bias, we have discovered a strong correlation between the formation of jets and the multiplicity of an event. Namely, the larger the amount of jets in an event, the more likely an event is to have high multiplicity. Upon combining this result with our earlier conclusion that high multiplicity leads to high values on the event shape observables, we run into a complication. When we use event shape observables to determine whether or not an event has jets, we assume that events with (especially 1 or 2) jets have low values on our event shape observables, because we have proven this to be the case in our section on hard QCD processes. However, the strong correlation between multiplicity and the amount of jets for minimum bias event works in the opposite direction. It states that events with jets are most likely high multiplicity events and are therefore sphere-like. These two conflicting correlations render all of our event shape observables ineffective when using them to predict jets in minimum bias events.

It is possible overcome this effect by omitting all low-multiplicity events (1-15, for example). However, in doing this we throw away approximately 70 percent of all events. It is debatable whether this is the right approach, or that maybe another way can be found to improve performance of event shape observables in minimum bias data. Nonetheless, we have shown that the large amount of low-multiplicity events has a large effect on the performance of our event shape observables.

The greatest weakness of transverse sphericity and spherocity in being able to judge how pencil-like an event is lies in their inability to take into account a particle's pseudorapidity, causing them to wrongly interpret some events where 2 particles or jets are at right angles to each other as extremely pencil-like. This results in some events where the transverse observables; sphericity and spherocity, which correlate nicely, give very different results than 3-dimensional sphericity.

Finally, we conclude that in the area within $|\eta| < 1$, relatively few particles are detected. From our calculations, in the hard QCD data generated by PYTHIA 8, most of the particles, and therefore also jets, appear in the circular areas straddling the incoming beam axis. Now, our choice for $|\eta| < 1$ is not random, it is approximately equal to the area in which detectors like ALICE can detect massive particles. Because of this technological restriction, the high particle-density around the incoming beam axis is something which would be extremely hard to measure in real life experiments. We are therefore unable to say at this point whether this is a flaw (or feature) of PYTHIA 8, or an accurate representation of reality, because we are unable to verify.

6 Discussion and Outlook

Transverse sphericity is an observable that is widely used within the high-energy physics community. Our analysis of how transverse sphericity and other event shape observables perform for certain datasets, or cuts in jets or pseudorapidity, may help some scientists or students better understand its behaviour and why it and other event shape observables are very useful in some situations and not as much in others and perhaps even what the underlying cause is for their good or bad performance.

One of the shortcomings of this thesis is the amount of data used. We have used datasets of 20.000 events with minimum bias and 40.000 events with hard QCD. This is large enough for most of the analyses done in this thesis, but, for example in the jet selections for minimum bias events, we ended up with only 16 events with two jets, which is not very much. It may therefore be interesting to redo some of the analyses with larger datasets, say of the order 10^6 events.

The results presented in this thesis encourage more research to be done into this subject. Further research could focus on finding other ways to let event shape observables perform better in minimum bias data without having to throw away most of the events. Also, we expect that some of the jets found in minimum bias data may simply be groups of particles heading in the same direction, not originating from a hard QCD interaction and therefore not what we would define as a jet. One could research if this is true, and if so, for how many of the detected jets this is the case. A good start would be to experiment with changes in the jetfinder parameters or jetfinding algorithm.

Furthermore, we encourage anyone interested in doing further research into this subject to look deeper into our finding that for events generated with hard QCD most of the particles and jets move roughly along the direction of the incoming beam axis. It would be interesting to see if the same effect is observed in events with minimum bias, and if maybe this effect is of a mathematical nature, for example due to the non-linearity of η in θ . Because this is so hard to check in real experiment, it may also be a viable option to contact the developers of PYTHIA 8 to see if they are aware of this and if so, what causes this effect.

References

- [1] Prof. Sean Carroll. *Dark Matter, Dark Energy: The Dark Side of the Universe* California Institute of Technology
- [2] Andrea Banfi, Gavin P. Salam and Giulia Zanderighi *Phenomenology of event shapes at hadron colliders*. Institute for Theoretical Physics, ETH Zurich, 8093 Zurich, Switzerland.
- [3] Arvind Khuntia, Sushanta Tripathy, Ashish Bisht, and Raghunath Sahoo *Event Shape Engineering and Multiplicity dependent Study of Identified Particle Production in proton+proton Collisions at $\sqrt{S} = 13\text{TeV}$ using PYTHIA* Discipline of Physics, School of Basic Sciences, Indian Institute of Technology Indore, Simrol, Indore 453552, India November 13, 2018
- [4] L. Garren (Fermilab), I.G. Knowles (Edinburgh U.), S. Navas (U. Granada), P. Richardson (Durham U.), T. Sjöstrand (Lund U.), and T. Trippe (LBNL). *MONTE CARLO PARTICLE NUMBERING SCHEME*
- [5] Torbjörn Sjöstrand, Stephen Mrenna, Peter Skands *A Brief Introduction to PYTHIA 8.1* CERN/PH, CH-1211 Geneva 23, Switzerland. Department of Theoretical Physics, Lund University, Sölvegatan 14A, SE-223 62 Lund, Sweden. Fermi National Accelerator Laboratory, Batavia, IL 60510, USA.
- [6] B. Jacak, *Stony Brook University*, P. Steinberg *Brookhaven National Laboratory Creating the perfect liquid in heavy-ion collisions*
- [7] MissMJ, *Standard Model of Elementary Particles*, Own work by uploader, PBS NOVA, Fermilab, Office of Science, United States Department of Energy, Particle Data Group. https://commons.wikimedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg.
- [8] H. Kirschenmann, CERN. *Jets at CMS and the determination of their energy scale*
- [9] Torbjörn Sjöstrand, Stephen Mrenna, Peter Skands *PYTHIA 6.4 Physics and Manual* CERN/PH, CH-1211 Geneva 23, Switzerland. Department of Theoretical Physics, Lund University, Sölvegatan 14A, SE-223 62 Lund, Sweden. Fermi National Accelerator Laboratory, Batavia, IL 60510, USA.
- [10] Matteo Cacciari, Gavin P. Salam and Gregory Soyez. *FastJet user manual* LPTHE, UPMC Univ. Paris and CNRS UMR 7589, Paris, France Université Paris Diderot, Paris, France 3CERN, Physics Department, Theory Unit, Geneva, Switzerland Department of Physics, Princeton University, Princeton, NJ 08544, USA 5 Institut de Physique Théorique, CEA Saclay, France.