

USING JET SHAPES IN THE SEARCH FOR TeV-SCALE
PARTICLES AT THE ATLAS DETECTOR*

by

CRAIG LEVY**
NORTHEASTERN UNIVERSITY
BOSTON, MA

under supervision of

SERGEI CHEKANOV
HEP DIVISION
ARGONNE NATIONAL LABORATORY

*Work performed at Argonne National Laboratory, a contract Laboratory of the United States Department of Energy.

**Participant in the Summer 2010 Student Program. This program is coordinated by the Division of Educational Programs.

Abstract

Using Jet Shapes in the Search for TeV-Scale Particles at the ATLAS Detector. CRAIG LEVY (Northeastern University, Boston MA, 02115) SERGEI CHEKANOV (Argonne National Laboratory, Argonne IL, 60439).

The ATLAS detector of the LHC proton-proton collider (CERN, Switzerland) is currently collecting data at record collision energies ($\sqrt{s} = 7$ TeV) and is searching for new physical phenomena at the most fundamental levels of matter. Many theories beyond the Standard Model of particles (SM) predict particles with invariant mass on the order of TeV/c^2 , much heavier than the heaviest currently observed elementary particle. A general feature of many such massive particles would be a primary decay channel through top-pairs; that is, the particle decays into a top/anti-top quark pair. The final state of such a decay is three light quarks from each top quark for a total of six light quarks which, in the fully hadronic decay channel, hadronize into six jets. However, in the case of the decay of a TeV-scale particle, the six jets will be highly Lorentz-boosted and collinear, resulting in overlap at the detector making it impossible to resolve the jets as independent objects; the jets will appear as two “composite” jets, apparently indistinguishable from “mono-jets” that occur due to less exotic processes. We propose a new method using jet-shape as a means of identifying overlapping “multi-jets” from the QCD (Quantum-ChromoDynamic) background “mono-jets”. The proposed method involves treating jets as conic sections (ellipses), allowing for the definition of several shape-variables. These variables include several ellipse eccentricities. Monte-Carlo (Pythia) truth studies were performed using over 20 identified shape-variables. The truth studies used a hypothetical Z' with 2- and 3- TeV mass as the signal event. Event cuts were made based on up to 8 shape-variables in order to improve signal-to-background ratio. It was found that shape-variable cuts coupled with jet-mass cuts can improve signal-to-background ratio by a factor of ~ 177 while maintaining low signal rejection (rejection factor < 4), with even larger signal-background improvements possible at the expense of the signal.

Introduction

The LHC (Large Hadron Collider), located at CERN (Geneva, Switzerland), is the largest particle accelerator in the world, currently operating at the record center-of-mass energy of 7 TeV (half the designed peak energy of 14 TeV) and luminosity of $1.44 \cdot 10^{30} \text{ cm}^{-2}\text{s}^{-1}$. The ATLAS (A Toroidal Lhc ApparatuS) detector is the largest detector of the LHC. The ATLAS collaboration is involved in the search for, and analysis of, a multitude of new physical phenomena that will likely become visible at this new energy scale, including the search for signatures of physics beyond the Standard Model (SM)- the currently accepted set of laws governing quarks, leptons, and their interactions. Despite the success of SM, the model is widely known to be incomplete (for reasons which are beyond the scope of this paper); numerous theories have been proposed in an attempt to complete it. Many models which attempt to complete SM, including theories of Extra Dimensions^[i], Supersymmetry^[ii] and others, predict new particles with mass on the order of TeV/c^2 , much larger than the top-quark mass ($\sim 173 \text{ GeV}/c^2$). Detecting such massive particles will be important to the validation or dismissal of these theories.

Identifying signatures of TeV-scale particles presents a problem, however, because the primary decay channel of any generic neutral particle with TeV-scale mass is through top-pair production. The fully-hadronic decay channel is

$$X \rightarrow t\bar{t} \rightarrow (W+b)(W-\bar{b}) \rightarrow (q\bar{q}b)(q\bar{q}\bar{b})$$

where X is a generic TeV-scale particle and q is a light quark (u,d,c,s). Each of the six quarks then hadronizes into a particle-shower known as a 'jet', producing a final-state of 6 jets at the detector. Fully hadronic top-pair decays have been observed at the Tevatron (Fermilab, IL). However, in the case of an initial TeV-scale particle X , the light decay products (jets) will be highly boosted and collinear in the lab-frame, resulting in jet-overlap at the detector which will make resolving the final state as multiple jets impossible; only one “multi-jet” from each top-quark will be visible, as opposed to three “mono-jets” (see **fig. 1**). It is thus necessary to find some way of separating signal “multi-jets” from background “QCD-jets” which occur due to less exotic hadronization processes.

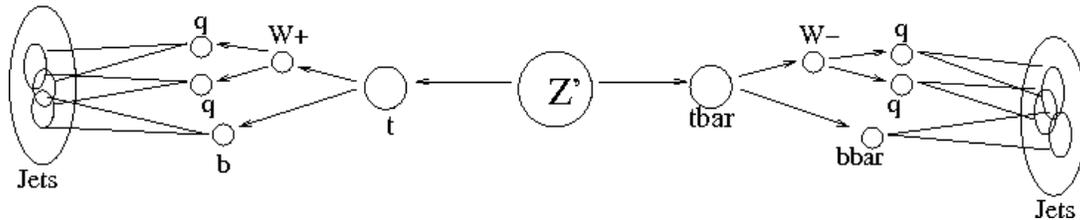


Fig.1 Schematic of a massive Z' decaying through top-pair production to a pair of multi-jets

Some suggested methods for improving signal to background ratio (SBR) include Principle-Component-Analysis (PCA) and Neural Networks^{[iii][iv][v]}. We propose a new method which uses

global jet-shape as means of improving SBR. This method is more intuitive than other approaches; it uses familiar geometrical variables rather than more abstract dimensional-reduction or cluster-analysis. We outline the proposed method and its implementation, and show the results of its application to Monte-Carlo truth-studies.

Shape-Variables Method

In this method, all constituents of a reconstructed jet are mapped onto the eta-phi phase space. Each jet constituent (hadron) is defined by its position in eta (pseudorapidity) and phi (azimuthal angle) with respect to the beam-line and interaction-point, as well as by its energy. In this way, each particle is represented by a point (eta & phi) and a weight (energy), making the shape effectively 3-dimensional. If it is assumed that the jet in this phase space is a conic section (roughly elliptical), then we can define several shape-variables, including major axis length, minor axis length, ellipse eccentricity, and others (to be discussed below). The first task is thus to define the axes and lengths of the ellipse. This is not a trivial process, as a jet is not a true ellipse but a composite object of discrete points.

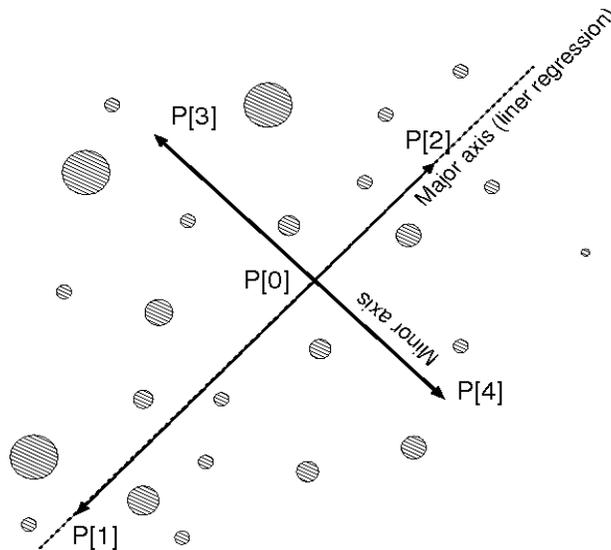


Fig. 2 Sketch of an approximately elliptical composite object. Each point is a constituent, with the size of each constituent representing its weight. The major-axis line is an unweighted linear regression, the minor-axis line is by definition perpendicular to the major axis, through the geometric mean.

Geometric Mean & Linear Regression

First, the geometric mean (i.e. without considering the weights) of all constituents is defined (P[0] in Fig.2). The geometric as opposed to weighted mean was chosen to produce less homogeneous shape-variables (i.e. variables more sensitive to differing geometry). Once the geometric mean is identified, unweighted linear regression is performed to define the major axis-line (see **fig. 2**). Unweighted linear regression was chosen over weighted linear regression for the same reason as for the geometric mean. Next, the minor axis-line is defined to be perpendicular to the major axis and passing through the geometric mean.

With the axis-lines of the ellipse defined, the next step is to define the axis-length. We identified two main classes of length-definition: the quadrant method and the non-quadrant method (NQ). We will discuss each of these methods respectively.

Quadrant Method

In the quadrant-method, the eta-phi space is first divided into four quadrants centered at the ellipse geometric center, each of which corresponds to one of the ellipse semi-axes. This is done by taking the major and minor axis lines (from linear regression) and rotating them by 45°, putting each semi-axis in one of the quadrants (see **fig. 3**). The length of each semi-axis is defined by finding the weighted center of each quadrant; that is, all constituent points are separated by the quadrant in which they lie and the weighted mean of each quadrant is found independently, without consideration to points in other quadrants. The length of the semi-axis is thus the length between the global geometric center and the quadrant center.

Non-Quadrant Method (NQ)

In NQ, rather than dividing the eta-phi space into four exclusive quadrants, it is divided into two sets of semi-planes; the major axis-line defines two semi-planes (the part above and the part below), as does the minor axis-line (see **fig. 3**). In this way, each point is in two of four semi-planes rather than a single exclusive quadrant. The weighted means above and below the major axis-line are the weighted centers defining the lengths of the minor semi-axes, while the means above and below the minor axis-line define the lengths of the major semi-axes.

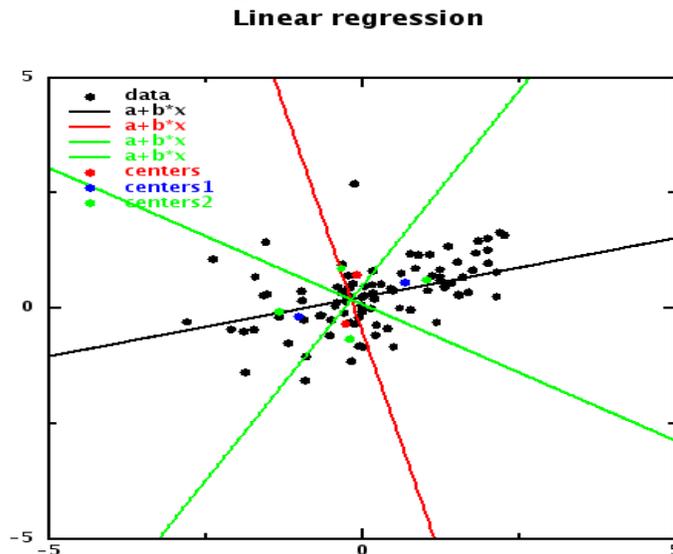


Fig. 3. The display of a java-based “calculator” developed for interactive debugging of the shape-variables program, where the weighted points are from three overlapping gaussian distributions, approximating a calorimeter’s jet display. The black/red lines are the major/minor axis-lines, while the (diagonal) green lines divide the quadrants. Here, the green points represent the weighted quadrant-centers (quadrant method). The blue points represent weighted centers above/below the minor axis-line (i.e. major semi-axes) and the red points represent weighted centers above/below the major axis-line (i.e. minor semi-axes), (NQ method). For “method 2”, each of these centers is projected onto the corresponding axis-line.

Definition of Variables

A list of the identified shape-variables and their definitions is presented below. Each variable may be either quadrant or non-quadrant, and may be of the standard method or “method 2”, which describes the way in which the lengths were calculated. In the standard method, lengths are calculated directly between the centers; the major length is the distance between the two major-semi-axis centers (in **Fig. 3** either the blue dots or the green dots on either side of the red line). In method 2, the centers are first orthogonally projected onto the corresponding axis-line, and the length is the distance between these projections.

- Major length - distance between major semi-axis centers (P[1] and P[2])
- Minor length - distance between minor semi-axis centers (P[3] and P[4])
- Eccentricity (ECC) - $1 - (\text{minor length} / \text{major length})$. Range [0,1].
~ a measure of the degree to which the ellipse fails to be circular. $Ecc=0$ is a perfect circle, while $ecc=1$ is an infinitely elongated object (line).
- Semi-major 1(major length 1) – shorter of two major semi-axes. Distance between P[0] and P[2].
- Semi-major 2(major length 2) - longer of two major semi-axes. Distance between P[0] and P[1].
- Major ECC - $1 - (\text{semi-major1} / \text{semi-major 2})$. Range[0,1].
~ a measure of the degree to which the ellipse is 'skewed' to one side of the minor axis-line. A large value signifies a large difference between lengths of the major semi-axes.
- Semi-minor 1(minor length 1) - shorter of two minor semi-axes. Distance between P[0] and P[4].
- Semi-minor 2(minor length 2) - longer of two minor semi-axes. Distance between P[0] and P[3].
- Minor ECC - $1 - (\text{semi-minor1} / \text{semi-minor2})$. Range [0,1].
~ a measure of the degree to which the ellipse is 'skewed' to one side of the major axis-line.
- Absolute length – distance between the two points furthest from the geometric center, when the points are projected onto the major axis.
- Absolute width – distance between the two points furthest from the geometric center, when the points are projected onto the minor axis.

With the exception of absolute length & width, all the above variables have as many as four manifestations, with quadrant vs. non-quadrant and standard method vs. method 2 possible in any combination (i.e. quadrant-standard, quadrant-method 2, NQ-standard, NQ-method2). Eccentricity is a well-known value describing conic-sections. We are, however, to our knowledge the first to identify the major and minor eccentricities of an ellipse, which are unlikely to differ from unity except when attempting to define a classical ellipse out of a cluster of discrete points where weighting-effects become apparent.

Monte-Carlo Truth Studies

The ability of the more-than 20 variables described in the previous section to improve SBR was tested by performing truth-studies with Pythia^[viii]. This Monte-Carlo program was used to produce 500,000 events each for 2 TeV and 3 TeV signal Z' particles decaying to a top-pair, as well as 500,000 general QCD background events (including top-quarks) producing jets. Jets were reconstructed using the anti- K_T jet clustering algorithm^[vi], with cone size $R=0.6$. A C++ code was developed to calculate and return 24 variables for the first and second leading p_T jets from each event (only jets with $p_T \geq 500$ GeV were accepted). The results of this code as well as a list of the eta, phi, and energy of all jet-constituents could be read into a file, allowing for further debugging by placing reconstructed jets of interest into the java “calculator” (**fig. 3**) for direct observation.

Figure 4 shows histogram plots for all 24 variables of the leading p_T (transverse momentum) jet. The first plot is invariant jet-mass. The two signal plots peak at the top quark mass of ~ 173 GeV, while the QCD-background has a broader profile, peaking much lower at ~ 40 GeV. The second plot is a kinematic (rather than shape) variable F_{\max} , defined as $F_{\max} = E_{\text{CONSTIT-MAX}} / E_{\text{TOT}}$, where $E_{\text{CONSTIT-MAX}}$ is the energy of the single most energetic constituent in the jet and E_{TOT} is the total energy of all constituents in the jet. The remaining plots are shape variables.

Before applying any cuts, it is important to know the correlations between variables. If two variables have a large correlation coefficient (~ 1), then making a cut on one variable is essentially equivalent to making a cut on the other. CERN's ROOT framework has provisions for calculating such inter-variable correlation coefficients (with a range $[-1,1]$, where 0 is entirely uncorrelated, 1 is entirely correlated, and -1 is entirely anti-correlated), which was used to produce a “correlation matrix” for the 24 variables. An abridged version of this matrix is presented as **Table 3** at the end of this paper. This version includes those variables which were used as cut variables due to their ability to discriminate between signal and background.

Most of the shape-variables have clear separation between signal and background events, however jet-mass has by-far the most separation, both in terms of peak-position and distribution-width. This is expected, as most QCD jets are not products of top-quarks. For this reason it was decided to first make a jet-mass cut and then study the subsequent plots. **Figure 5** shows the same 24 variables after a mass cut from 140 GeV to 250 GeV (all jets with invariant mass below 140 GeV or above 250 GeV are rejected). The resulting QCD distributions of most variables are significantly different from the corresponding distributions prior to the mass cut, as can be seen by comparing figures 4 and 5.

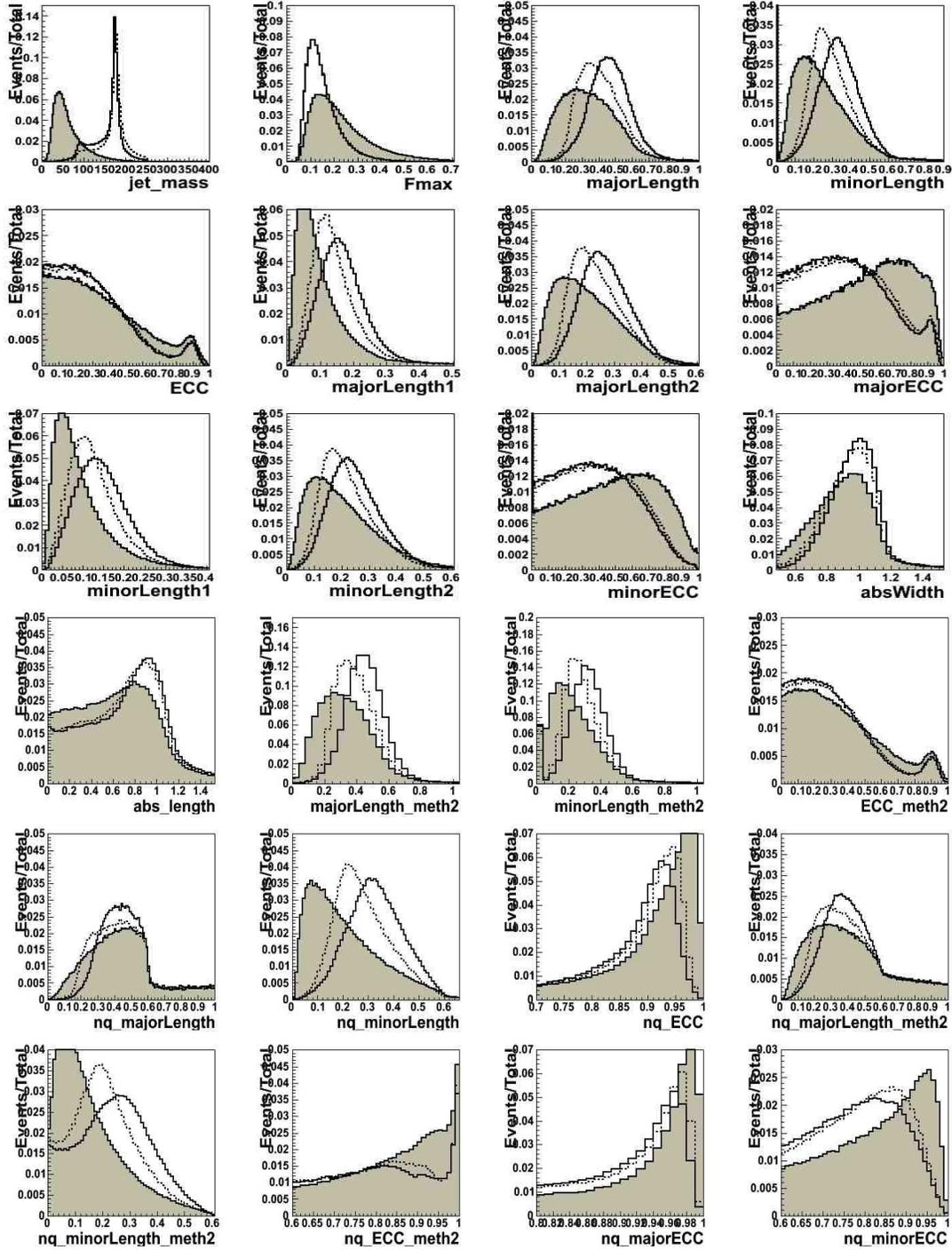


Fig. 4. Histogram plots for leading- P_T jet (**Jet 1**). The filled histogram is QCD background, the solid line is the 2 TeV resonance, and the dotted line the 3 TeV. Jet invariant-mass, Fmax, and 22 shape-variables shown. In these plots, a prefix “nq_” denotes non-quadrant method, while no prefix denotes quadrant-method, and a suffix “_meth2” denotes method 2, while no suffix denotes standard method. The vertical axis is the unitless fraction of events per total number of events. The horizontal axis for any length variable (major length, minor length, etc.) is in distance units in eta-phi space. The horizontal axis for any of the eccentricities, as well as Fmax, is unitless.

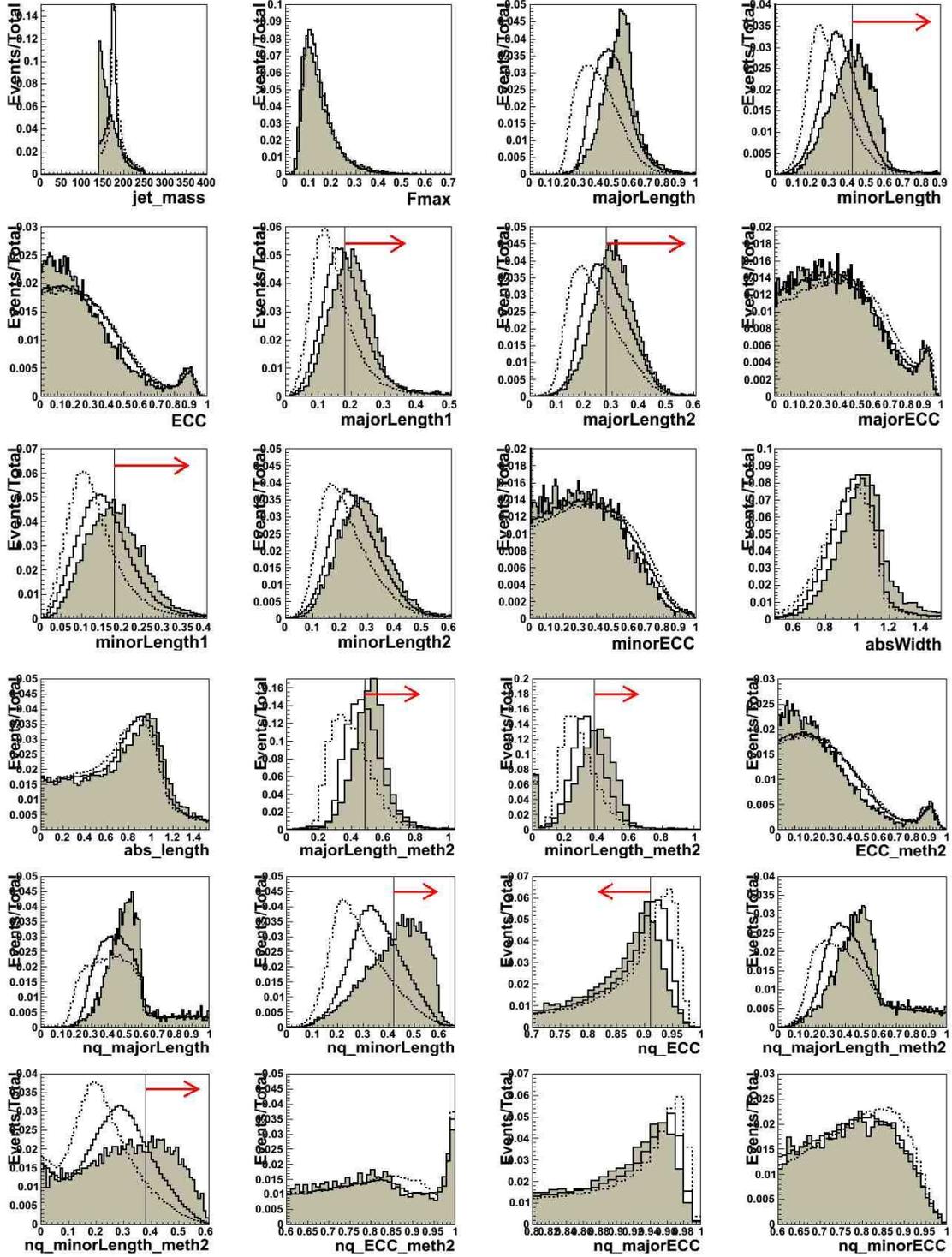


Fig. 5. Histogram plots for leading- p_T jet (**Jet 1**) after placing a cut on jet-mass. Only jets with mass between 140 GeV and 250 GeV are accepted. Filled histograms are QCD background, solid-line is for a 2 TeV resonance, dashed-line for 3 TeV. Note the dramatic change in the shape of many of the QCD plots relative to figure 4. Vertical lines show the position of the cuts listed in table 1. Red arrows show the section being rejected.

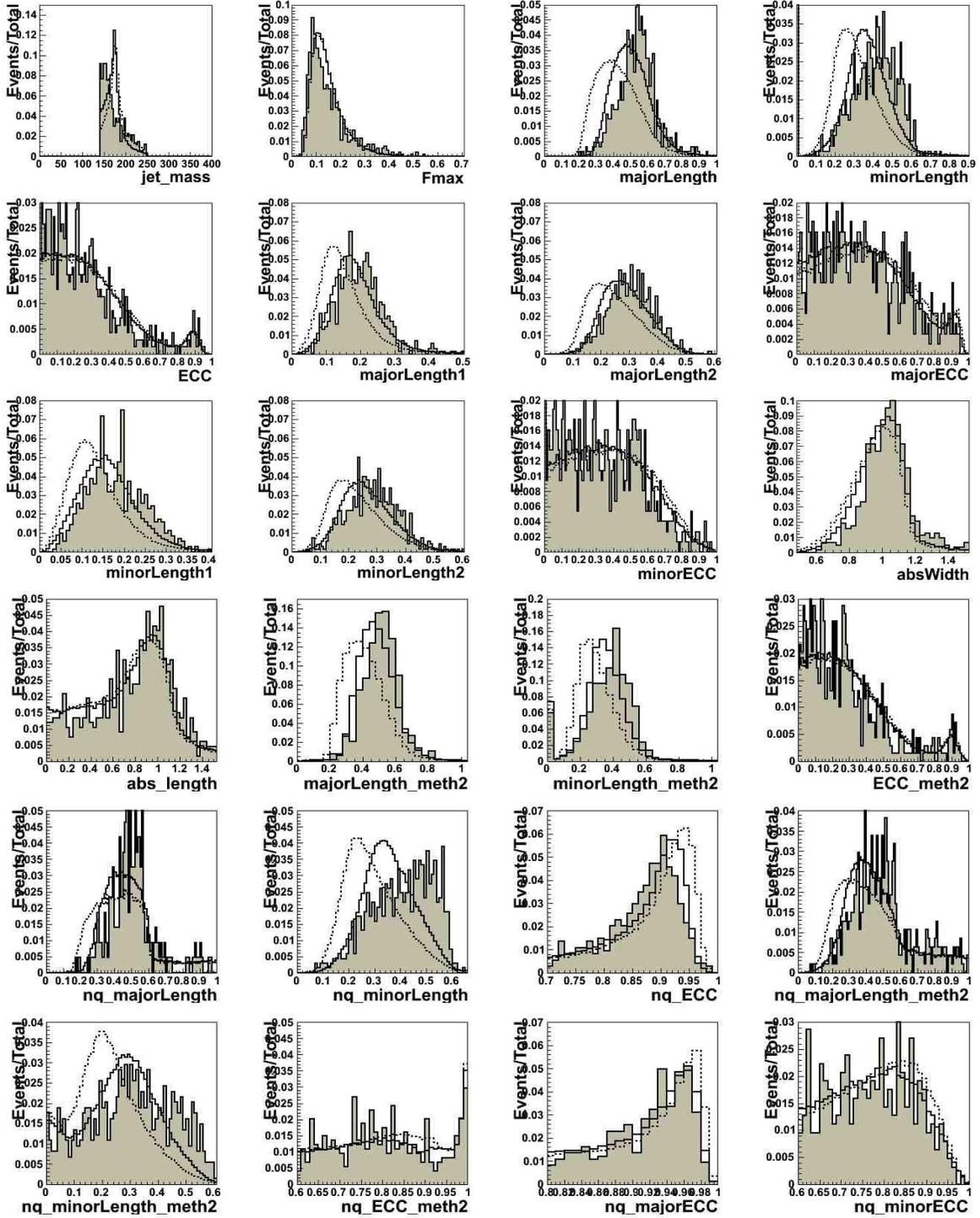


Fig. 6. Histogram plots for **JET 2** (second leading P_T jet) after mass cut at 140 GeV. Filled histograms are background, solid and dashed lines are 2 TeV and 3 TeV signals, respectively. Signal plots are very similar to Jet 1 plots, as jets from top quarks will have nearly the same P_T . By contrast, the background appears much noisier than Jet 1 background, as the second-leading P_T QCD jet may be much softer than the leading QCD jet, causing lower jet-acceptance.

The shape-variable distributions after the mass-cut were studied for their suitability as signal/background discriminators, with each jet (leading and second-leading p_T) being independently considered. For each jet, 8 variables were chosen for their low correlations and relatively high rejection-power. A variable is considered a good rejection-variable if there exists a range on the plot (fig. 5) where the signal has a significantly higher fraction of entries than the background. The variables and the cuts placed on them are listed in **table 1**. The different methods of applying cuts and their results are discussed in the Results section.

Variable	Cut (jet rejected if...)
nq_ECC	< 0.91
Major length 1 (jet1 only)	> 0.18
Major length 2	> 0.28
Minor length 1	> 0.18
Minor length meth2	> 0.38 (jet 1) > 0.42 (jet 2)
nq_minor length	> 0.42
Minor length	> 0.42
nq_minor length_meth2	> 0.38 (jet 1) > 0.35 (jet 2)
Major length meth2 (jet 2 only)	> 0.48

Table 1. Shape variables & cuts placed. In cases where the difference between jet 1 and jet 2 cuts differed by 0.02 or less, only the jet 1 cut is listed.

Results & Discussion

In determining how to improve SBR, “rejection” and “relative rejection” factors were used. The rejection factor of a cut is found by dividing the accepted entries prior to the cut by the accepted entries after the cut; larger rejection factors indicate many rejected events. For each cut, a rejection factor is assigned to the background as well as both signals (2- and 3- TeV). The relative rejection of a cut is the rejection factor of the background divided by the rejection factor of the signal. Large relative rejections indicate large improvement of SBR. It is therefore desirable to make cuts which produce very low signal rejection but very high background rejection. It is necessary to maintain low signal rejection, even if relative rejection could be improved by producing a larger signal rejection, as cross-sections for any new massive particle are extremely small, and any signal rejection which is too large will leave a statistically insufficient signal sample. For all results below, the rejection factors given are for the jet1-jet2 invariant mass range 1.5-2.5 TeV for the 2 TeV signal and the range 2.5-3.5 TeV for the 3 TeV signal.

The single most important rejection variable is the jet invariant-mass (as is clear from **fig. 4**);

Changing the mass-cut impacts the rejection factor greatly. **Table 2** outlines the effects of the cuts. A lower bound of 120 GeV on the leading jet and 100 GeV on the second-leading jet produces a relative rejection factor of ~ 30 , while hardening the cut to 150 GeV for jet1 and 130 GeV for jet2 produced a relative rejection of 84.4 for 2 TeV signal and 74.9 for 3 TeV signal. It is clear from the table that small changes to the mass-cut produce large differences in background rejection, but very small differences in signal rejection.

Several different methods for shape-variable cuts were used. In many cases, it was found that making hard 8-variable cuts was not as effective as soft 8-variable cuts. This is to say that requiring a jet to pass all 8 cuts was in some cases ineffective compared with requiring a jet to pass any 4 of 8 cuts, or 5 of 8 cuts, etc. Shape-variable cuts, in comparison to mass cuts, had a relatively small effect individually, but coupled with the mass-cuts provided an important boost to the relative rejection. Relative rejection from mass cuts at 150 GeV and 130 GeV (for jets 1 and 2, respectively) was improved from ~ 80 to as much as 135 when shape-variable cuts were included. For 140 GeV mass cuts on both jets, inclusion of variable-cuts improved the relative rejection from ~ 100 to as much as 177 (in the case of the 3 TeV signal). This boost could be important in differentiating QCD top-pair events (from energetic gluons) from signal top-pair events.

a) 2-TeV Relative Rejection

Mass Cut	Just Mass	≥ 4 variables	≥ 5 variables	≥ 6 variables	≥ 7 variables
Jet1 120 GeV Jet2 100 GeV	46.1 / 1.4 =32.9	78.8 / 2.2 = 35.8	91.9 / 2.6 = 35.3	112.0 / 3.5 =32.0	158.3 / 5.4 =29.3
Jet1 150 GeV Jet2 130 GeV	177.3 / 2.1 =84.4	467.5 / 4.3 =108.7	444.2 / 4.0 =111.1	584.4 / 5.5 =106.3	925.3 / 9.0 =102.8
Jet1 140 GeV Jet2 140 GeV	185.8 / 2.2 =84.5	407.5 / 3.5 =116.4	545.0 / 4.2 =129.8	746.5 / 6.0 =124.4	1388.0 / 9.9 =140.2

b) 3-TeV Relative Rejection

Mass Cut	Just Mass	≥ 4 variables	≥ 5 variables	≥ 6 variables	≥ 7 variables
Jet1 120 GeV Jet2 100 GeV	40 / 1.4 =28.6	63.0 / 1.7 =37.1	71.7 / 1.9 =37.7	82.4 / 2.2 =37.5	104.0 / 2.8 =37.1
Jet1 150 GeV Jet2 130 GeV	127.3 / 1.7 =74.9	284.7 / 2.7 =118.6	260.9 / 2.3 =113.4	319.5 / 2.7 =118.3	460.5 / 3.4 =135.4
Jet1 140 GeV Jet2 140 GeV	133.3 / 1.7 =78.4	233.7 / 2.2 =106.2	292.7 / 2.3 =127.3	372.8 / 2.7 =138.1	602.2 / 3.4 =177.1

Table 2. Relative rejection factors for different mass & shape-variable cuts. All rejections a) in the jet-jet invariant mass region 1.5-2.5 TeV, b) in the jet-jet invariant mass region 2.5-3.5 TeV. In calculation of relative rejection, the numerator is the background rejection factor, the denominator is the signal rejection factor. All jet-mass cuts include an upper bound of 250 GeV. For shape-variable cuts, a jet is accepted if $\geq n$ cuts are passed, where n is 4, 5, 6, or 7 out of 8 variables.

From **table 2**, it is clear that the larger Z' mass is easier to separate from the background than the smaller, as all signal rejection factors (denominators) in **table 2b** are less than 4, as opposed to the signal rejection factors **2a** which easily rise above 4. This is expected from the inspection of **figure 5**, in which 3 TeV signal peaks are generally more separated from background peaks than 2 TeV signal peaks. This observation proves to hold for particularly punitive cuts; A mass cut of 150 GeV on both jets and inclusion of shape-variable cuts for ≥ 7 variables gives a 3 TeV signal rejection of only 3.8, with a large relative rejection of 284.2. Meanwhile, the same cut produces a somewhat large 2 TeV signal rejection of 12.1 (for a relative rejection of 215.9). **Fig. 7** shows the Jet1-Jet2 invariant mass distributions of different cut methods for jet-mass cuts at 140 GeV. For the 3 TeV signal, most of the signal loss is outside the jet-jet invariant mass range of 2.5-3.5 TeV (the range in which rejections are calculated), contributing to low signal rejection.

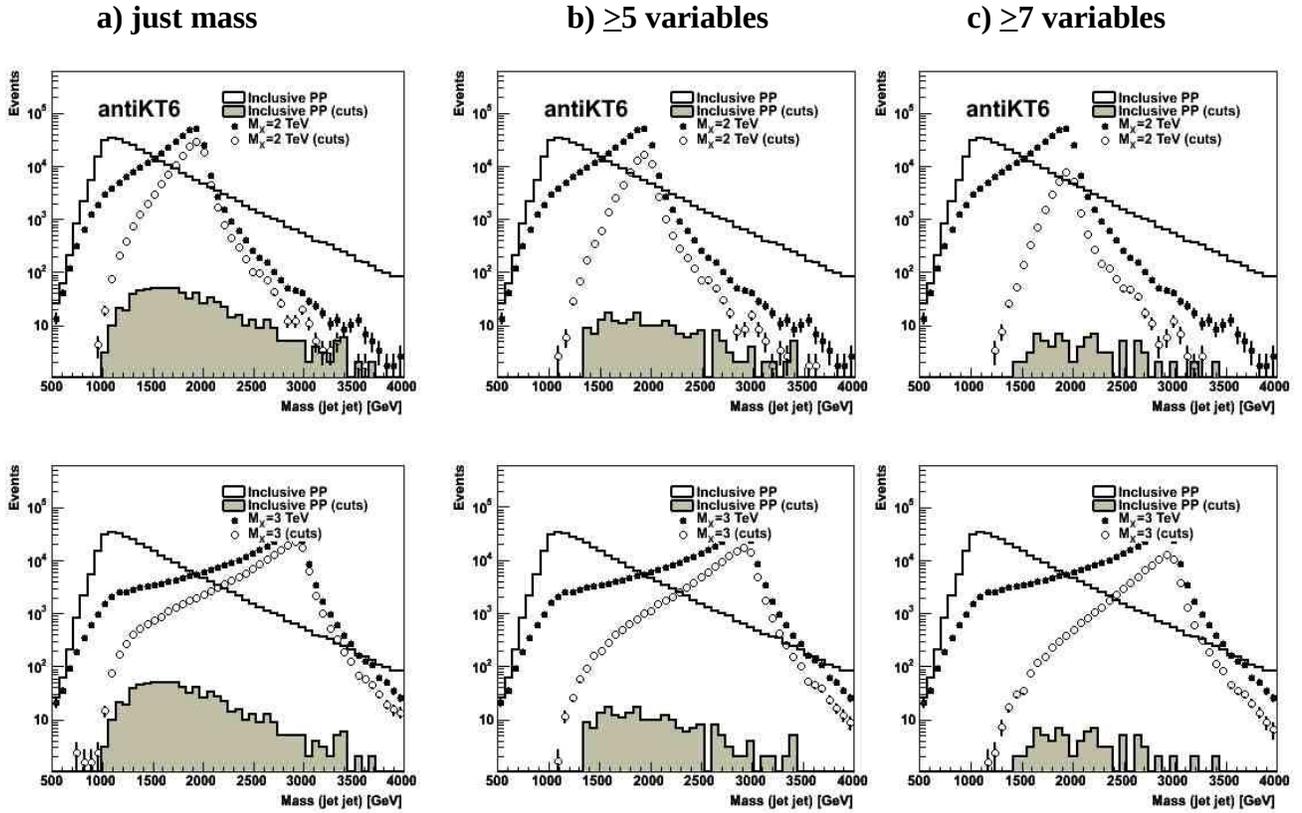


Fig. 7. Number of accepted events vs. Jet1-Jet2 invariant mass for three different rejection methods. a) only mass cut at 140 GeV, b) mass cut and ≥ 5 variables included, c) mass cut and ≥ 7 variables included. Line-histograms represent QCD background, points represent signal. For a, b, and c, the top plot is for a 2 TeV signal, the bottom plot is for a 3 TeV signal. Histograms for which the legend reads “(cuts)” represent events after applied cut. Rejection and relative rejection factors for all 6 plots can be read from **table 2**. Plots are normalized, in that 500,000 events were produced for both signal and background.

These results represent a significant improvement over previous attempts [vii] to use jet-shapes. These attempts differed in that they used jet-shapes from Principal-Component-Analysis rather than global geometric jet-shapes. Reported results used cuts on jet-mass and 2 PCA shape-variables to achieve relative rejection factors of ~ 30 .

Conclusion

A method for classifying jets according to global jet-shape has been presented as a means of identifying TeV-scale particles in the fully hadronic top-pair decay channel. This method is geometrically intuitive compared with such methods as PCA and cluster-analysis. Pythia truth studies for a Z' boson show that this method can easily improve SBR by a factor of over 100 with low signal rejection, and improvement factors of well over 200 are possible, with better SBR improvement possible for more massive signal particles. The next steps in investigating this method will be conducting full-detector simulations (for the ATLAS detector) to study the effects of finite detector resolution, followed by analysis of actual ATLAS data as it becomes available.

Acknowledgments

Thanks very much to Dr. Sergei Chekanov of ANL for being incomparably helpful and knowledgeable, and for making the field of experimental high-energy physics comprehensible. Thanks to Dr. James Proudfoot for his help as well as his enthusiasm and encouragement. Thanks also to Dr. Rik Yoshida, the ANL-ASC, and to Argonne's SRP for making this work possible.

	mass	nq ECC	Major 1	Major 2	Minor 1	minor length	minor length meth2	minor length meth1	nq minor length meth2	major length meth2
Mass (2 TeV)	1									
(3 TeV)	1									
(QCD)	1									
nq ECC	-0.0465	1								
	-0.0412	1								
	-0.0524	1								
Major 1	0.0301	-0.1343	1							
	0.0096	-0.1777	1							
	0.0350	-0.1466	1							
Major 2	0.0033	-0.1218	0.6665	1						
	-0.0018	-0.1766	0.6533	1						
	0.0066	-0.1282	0.6505	1						
Minor 1	0.0136	-0.1736	0.8323	0.6672	1					
	0.0041	-0.2223	0.8281	0.6459	1					
	0.0175	-0.1792	0.8414	0.6431	1					
minor length	0.0215	-0.0548	-0.0978	0.1699	-0.1474	1				
	0.0012	-0.0756	-0.0421	0.2300	-0.1020	1				
	0.0538	-0.0574	-0.1235	0.1562	-0.1728	1				
minor length meth2	0.0680	0.0015	-0.3755	-0.1187	-0.4140	0.4670	1			
	0.0128	-0.0057	-0.2732	0.0168	-0.3301	0.4740	1			
	0.1036	-0.0087	-0.3605	-0.1099	-0.3989	0.4484	1			
nq minor length	0.0058	-0.3112	0.6666	0.6797	0.7284	0.1358	-0.1970	1		
	-0.0148	-0.3684	0.6615	0.6898	0.7271	0.1924	-0.0893	1		
	0.0356	-0.3231	0.6651	0.6712	0.7166	0.1246	-0.1816	1		
nq minor length meth2	0.0919	0.0786	-0.1279	-0.1650	-0.1315	0.0230	0.1139	-0.1630	1	
	0.0223	0.0396	-0.0763	-0.1053	-0.0843	0.0569	0.1743	-0.1025	1	
	0.1627	0.1156	-0.1440	-0.1933	-0.1461	0.0427	0.1093	-0.1647	1	
major length meth2	0.0092	0.0299	0.8008	0.9768	0.7606	0.1089	-0.1916	0.7238	0.8896	1
	0.0009	-0.0367	0.7873	0.9788	0.7400	0.1754	-0.0545	0.7309	0.8905	1
	0.0131	0.0363	0.7886	0.9716	0.7499	0.0911	-0.1689	0.7214	0.8678	1

Table 3. Correlation matrix for mass and 9 shape-variables. Correlations were calculated using the ROOT calculator. For each variable, the first line is for 2 TeV signal, the second line is for 3 TeV signal, and the third line is for background. The matrix is symmetric, so entries above the diagonal are not shown. While some coefficients are somewhat large, most are in an acceptable range (≤ 0.4).

- i Matsumoto, Sato, Senami, Yamanaka. “Productions of second Kaluza-Klein bosons in the minimal universal extra dimension model at LHC”. *Phys. Rev. D.* **80.** 056006. Sept. 2009.
- ii Kane, Kolda, Roszkowski, Wells. “Study of constrained minimal supersymmetry”. *Phys. Rev. D.* **49.** 6173-6210. June 1994.
- iii Almeida, Lee, Perez, Sung, Virzi. “Top quark jets at the LHC”. *Phys. Rev. D.* **79.** 074012. April 2009.
- iv Almeida, Lee, Perez, Stermann, Sung, Virzi. “Substructure of high- p_T jets at the LHC”. *Phys. Rev. D.* **79.** 074017. April 2009.
- v Ellis, Vermilion, Walsh. “Recombination algorithms and jet substructure: Pruning as a tool for heavy particle searches”. *Phys. Rev. D.* **81.** 094023. May 2010.
- vi Cacciari, Salam, Soyez. “The anti- k_r jet clustering algorithm”. *J. High Energy Physics.* **04.** 063. April 2008.
- vii Chekanov, Proudfoot. “Searches for TeV-scale particles at the LHC using jet shapes”. *Phys. Rev. D.* **81.** 114083. June 2010.
- viii Sjöstrand, Mrenna, Skands. “A Brief Introduction to PYTHIA 8.1”. CERN-LCGAPP-2007-04. October 2007.