

Studies of granularity of a hadronic calorimeter for tens-of-TeV jets at a 100 TeV pp collider

S.V. Chekanov^a, A.V. Kotwal^{b,c}, J. Proudfoot^a, S. Sen^b, N.V. Tran^c, S.-S. Yu^e,
Chih-Hsiang Yeh^e

^a *HEP Division, Argonne National Laboratory, 9700 S. Cass Avenue, Argonne, IL 60439, USA.*

^b *Department of Physics, Duke University, USA*

^c *Fermi National Accelerator Laboratory*

^d *Department of Physics, Michigan State University, 220 Trowbridge Road, East Lansing, MI 48824*

^e *Department of Physics, National Central University, Chung-Li, Taoyuan City 32001, Taiwan*

Abstract

Texts

Keywords: multi-TeV physics, pp collider, future hadron colliders, FCC, SppC

1. Introduction

Particle collisions at energies beyond those attained at the LHC will lead to many challenges for detector technologies. Future experiments, such as high-energy LHC (HE-LHC), future circular pp colliders of the European initiative, FCC-hh [?] and the Chinese initiative, SppC [?] will be required to measure high-momentum bosons (W , Z , H) and top quarks with strongly collimated decay products that form jets. Studies of jet substructure can help identify such particles.

The reconstruction of jet substructure variables for collimated jets with transverse momentum above 10 TeV require an appropriate detector design. The most important for reconstruction of such jets are tracking and calorimeter. Recently, a number of studies [?] have been discussed using various fast simulation tools, such as Delphes [?], in which momenta of particles are smeared to mimic detector response.

A major step towards the usage of full Geant4 simulation to verify the granularity requirements for calorimeters was made in [?]. The studies included in this paper have illustrated a significant impact of granularity of electromagnetic (ECAL) and hadronic (HCAL) calorimeters on the shape of hadronic showers calculated using calorimeter hits for two particles separated by some angle. It was concluded that high granularity is essential in resolving two close-by particles for energies above 100 GeV.

This paper makes another step in understanding understanding of this problem in terms of high-level physics quantities typically used in physics analyses. Similar to the studies presented in [?], this paper is based on full Geant4 simulation with realistic jet reconstruction.

Email addresses: chekanov@anl.gov (S.V. Chekanov), ashutosh.kotwal@duke.edu (A.V. Kotwal), proudfoot@anl.gov (J. Proudfoot), sourav.sen@duke.edu (S. Sen), ntran@fnal.gov (N.V. Tran), syu@cern.ch (S.-S. Yu), jwzuzelski18@gmail.com (Chih-Hsiang Yeh)

23 2. Simulation of detector response and event reconstruction

24 The description of the detector and software used for this paper is discussed in
 25 [?]. We use the SiFCC detector geometry with a software package that represents a
 26 versatile environment for simulations of detector performance, testing new technology
 27 options, event reconstruction techniques for future 100 TeV colliders.

28 The GEANT4 (version 10.3) [?] simulation of calorimeter response was comple-
 29 mented with the full reconstruction of calorimeter clusters formed by the Pandora
 30 algorithm [? ?]. Calorimeter clusters were built from calorimeter hits in the ECAL
 31 and HCAL after applying the corresponding sampling fractions. No other corrections
 32 are applied. Hadronic jets were reconstructed with the FASTJET package [?] using
 33 the anti- k_T algorithm [?] with a distance parameter of 0.5.

34 In the following discussion, we use the simulations of a heavy Z' boson, a hypo-
 35 thetical gauge boson that arises from extensions of the electroweak symmetry of the
 36 Standard Model. The Z' bosons were simulated with the masses, $M = 5, 10, 20$ and
 37 40 TeV. The lowest value represents a typical mass that is within the reach of the LHC
 38 experiments. The value 40 TeV represents the physics reach of 100 TeV colliders. The
 39 Z' particles are forced to decay to two light-flavor jets ($q\bar{q}$), W^+W^- or $t\bar{t}$, where W
 40 and t decay hadronically. In all such scenarios, two highly boosted jets are produced,
 41 which are typically back-to-back in the laboratory frame. Typical transverse momenta
 42 of such jets are $\simeq M/2$. The main difference between considered decay types lays in
 43 different jet substructure. In the case of the $q\bar{q}$ decays, jets do not have any internal
 44 structure. In the case of W^+W^- , each jet originates from W , thus it has two subjects
 45 because of the decay $W \rightarrow q\bar{q}$. In the case of hadronic top decays, jets have three
 46 subjects due to the decay $t \rightarrow W^+ b \rightarrow q\bar{q}b$. The signal events were generated using the
 47 PYTHIA8generator with the default settings, ignoring interference with SM processes.
 48 The event samples used in this paper are available from the HepSim database [?].

49 3. Studies of jet properties

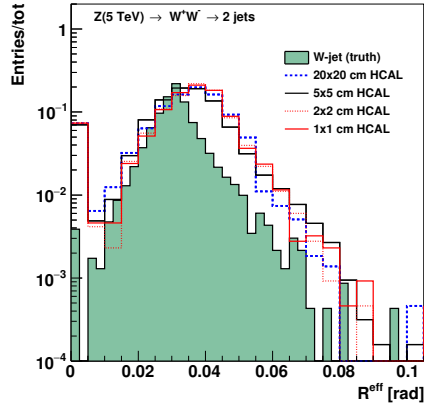
50 First let us consider several variables that represent jet substructure using different
 51 types of calorimeter granularity. The question we want to answer is how close the re-
 52 constructed jet substructure variables to the input "truth" value that are reconstructed
 53 using input particles directly from the PYTHIA8generator.

54 The effective radius is the average of the energy weighted radial distance in $\eta - \phi$
 55 space of jet constituents. Recently, it has been studied for multi-TeV jets in Ref.[?].

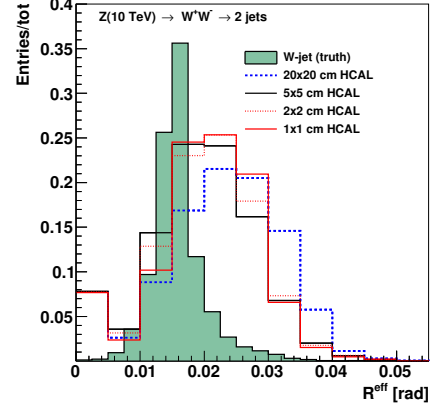
56 Let us study the effect of granularity on jet splitting scales. A jet k_T splitting scale [?
 57] is defined as a distance measure used to form jets by the k_T recombination algorithm
 58 [? ?]. This has been studied by ATLAS [?], and more recently in the context of 100
 59 TeV physics [?]. The distribution of the splitting scale $\sqrt{d_{12}} = \min(p_T^1, p_T^2) \times \delta R_{12}$ [?
 60] at the final stage of the k_T clustering, where two subjets are merged into the final
 61 one, is shown in Fig. 2.

62 3.1. Jet subjetteness

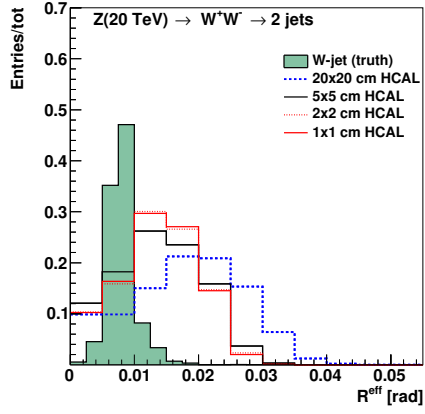
63 We recall that N -subjetteness [? ?], τ_N , of jets has been proposed as a class of
 64 variables with which to study the decay products of a heavy particle inside jets. τ_N is



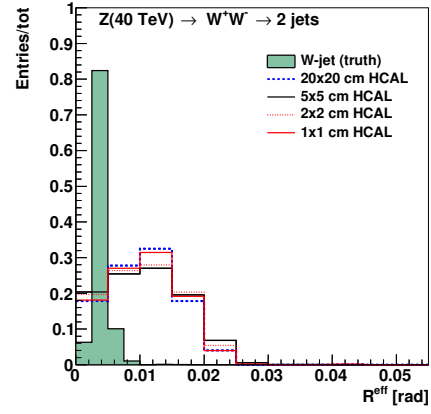
(a) 5 TeV



(b) 10 TeV

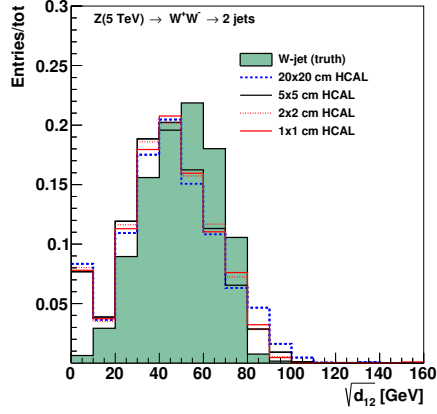


(c) 20 TeV

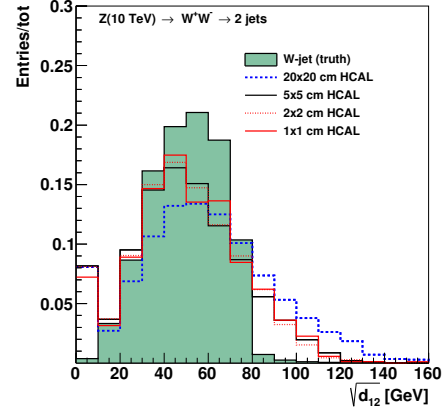


(d) 40 TeV

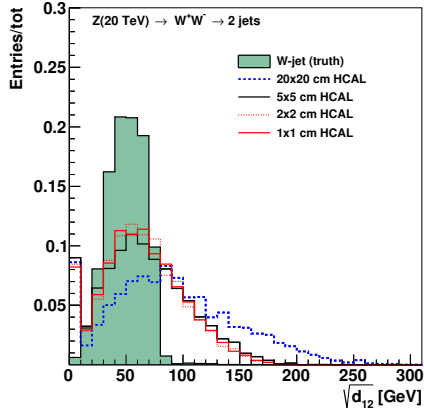
Figure 1: Jet effective radius for different jet transverse moment and HCAL granularity.



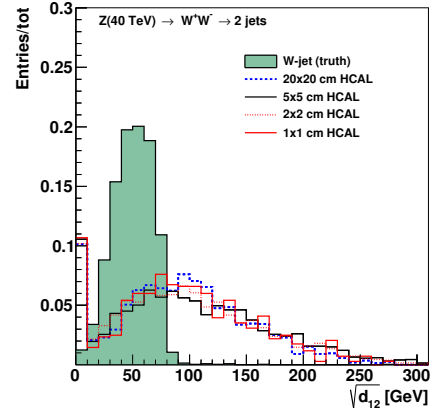
(a) 5 TeV



(b) 10 TeV



(c) 20 TeV



(d) 40 TeV

Figure 2: Jet splitting scale for different jet transverse moment and HCAL granularity.

a measure of the degree to which a jet can be considered as being composed of N k_T -subjets [?]. The variable τ_{32} , defined as the ratio of the N -subjettiness variables τ_3/τ_2 , is particularly sensitive to hadronically-decaying top-quark initiated jets. The variable, $\tau_{21} \equiv \tau_2/\tau_1$ can be used to reject background from W/Z decays. These variables do not strongly correlate with jet mass and can provide an independent check for the presence of top quarks. The jet substructure variables were obtained by re-running the k_T algorithm over the jet constituents of anti- k_T jets.

4. Study of detector performance with soft drop mass

In this section, we use the jet mass computed with a specific algorithm, soft drop declustering, to study the performance of detector with various detector cell sizes and center-of-mass (c.m.) energies.

4.1. The technique of soft drop declustering

The soft drop declustering [?] is a grooming method that removes soft wide-angle radiation from a jet. The constituents of a jet j_0 are first reclustered using the Cambridge-Aachen (C/A) algorithm [? ?]. Then, the jet j_0 is broken into two subjets j_1 and j_2 by undoing the last stage of C/A clustering. If the subjets pass the following soft drop condition, jet j_0 is the final soft-drop jet. Otherwise, the algorithm redefines j_0 to be the subset with larger p_T (among j_1 and j_2) and iterates the procedure.

$$\frac{\min(p_{T1}, p_{T2})}{p_{T1} + p_{T2}} > z_{\text{cut}} \left(\frac{\Delta R_{12}}{R_0} \right)^\beta, \quad (1)$$

where p_{T1} and p_{T2} are the transverse momenta of the two subjets, z_{cut} is soft drop threshold, ΔR_{12} is the distance between the two subjets in the η - ϕ plane, R_0 is the characteristic radius of the original jet, and β is the angular exponent.

In our study, we compare the performance of future detector when setting $\beta = 0$ versus when setting $\beta = 2$. For $\beta = 0$, the soft drop condition depends only on the z_{cut} . For $\beta = 2$, the condition depends on the angular distance between the two subjets and z_{cut} and the algorithm becomes infrared and collinear safe.

4.2. Analysis method

We employ the following method to quantify the detector performance and find out the cell size that gives the best separation power to distinguish signal from background. For each configuration of detector and c.m. energy, we draw the receiver operating characteristic (ROC) curves in which the x-axis is the signal efficiency (ϵ_{sig}) and y-axis is the inverse of background efficiency ($1/\epsilon_{\text{bkg}}$). In order to scan the efficiencies of soft drop mass cuts, we vary the mass window as follows. We first look for the median bin i_{med} ¹ of the soft drop mass histogram from simulated signal events. Taking the right boundary of bin i_{med} as the center of mass window x_{center} , we start increasing the width of mass window symmetrically on the left and on the right of x_{center} , in steps of

¹The integral from bin 0 to bin i_{med} ($i_{\text{med}} - 1$) should be greater (less) than half of the total number of events. Note, the bin width is 5 GeV.

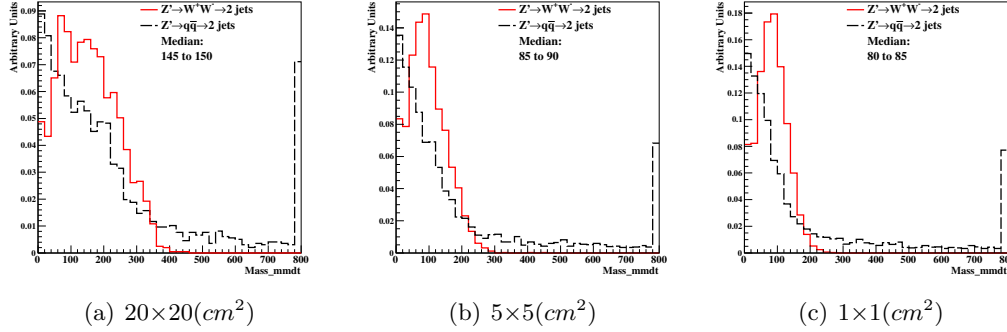


Figure 3: Distributions of soft drop mass for $\beta=0$, with 20 TeV c.m. energies and three different detector cell sizes: 20×20 , 5×5 , and 1×1 (cm^2). The signal (background) process is $Z'\rightarrow WW$ ($Z'\rightarrow q\bar{q}$).

5 GeV, i.e. the narrowest mass window is $[x_{\text{center}} - 5, x_{\text{center}} + 5]$. If one side reaches the boundary of the mass histogram, we only increase the width on the other side, also in steps of 5 GeV. For each mass window, there will be corresponding ϵ_{sig} and ϵ_{bkg} , which gives a point in the ROC curves.

4.3. Results and conclusion

Figures 3, 5, 7, and 9 present the distributions of soft drop mass for $\beta = 0$ and $\beta = 2$ with different c.m. energies and detector cell sizes; the signals considered are $Z'\rightarrow WW$ and $Z'\rightarrow t\bar{t}$. In Figs. 4, 6, 8, and 10, ROC curves from different detector cell sizes are compared for each c.m. energy, respectively.

Figures 4 and 6 show that for $\beta = 0$ the smallest detector cell size, $1 \text{ cm} \times 1 \text{ cm}$, has the best separation power at $\sqrt{s} = 5, 10$, and 20 TeV when the signal is $Z'\rightarrow WW$ and at $\sqrt{s} = 10$ and 20 TeV when the signal is $Z'\rightarrow t\bar{t}$. On the contrary, Figs. 8 and 10 show that for $\beta = 2$ the smallest detector cell size does not have improvements in the separation power with respect to those with larger cell sizes. In fact, the performances of the three cell sizes are similar. In addition, sometimes bigger detector cell sizes, $5 \text{ cm} \times 5 \text{ cm}$ or $20 \text{ cm} \times 20 \text{ cm}$ have the best separation power.

We also find compared to $\beta = 2$, soft drop mass with $\beta = 0$ has better performance for distinguishing signal from background. Therefore, we will apply requirements on this variable when studying the other jet substructure variables.

5. Studies of signal and background separation using jet substructure variables

In this section, we study different jet substructure variables and compare their ability to separate signal from background with different detector sizes and c.m. energy using Mann-Whitney U test and ROC curves.

By definition of Mann Whitney U test, if U value is close to 0.5, it means two distributions have similar compositions, and we can not distinguish them very well. On the

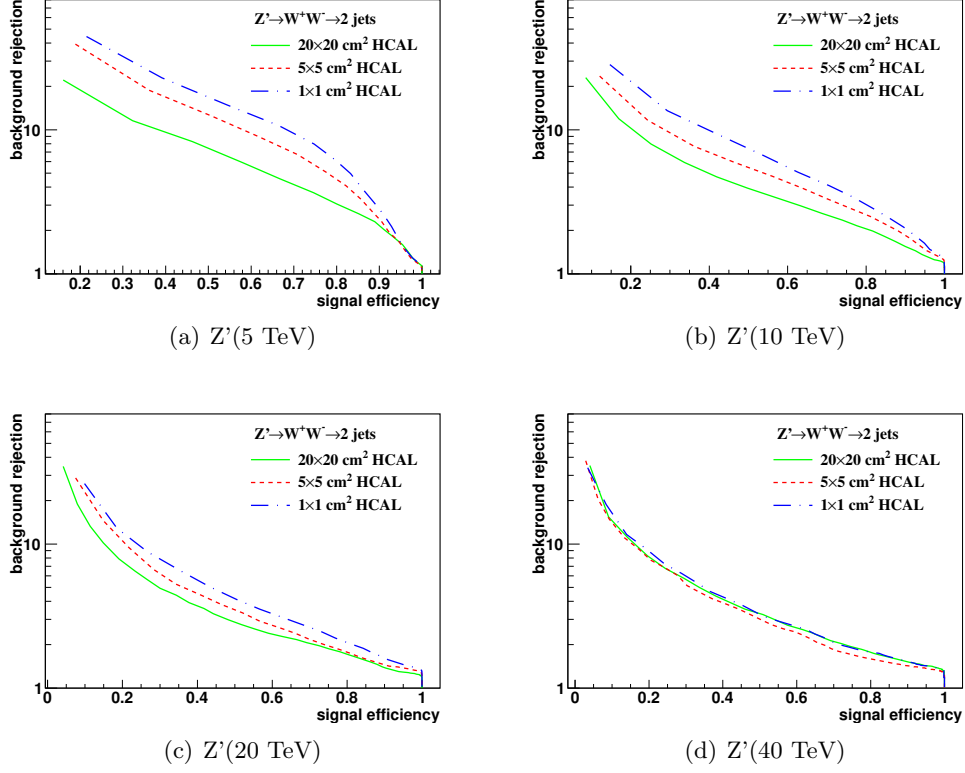


Figure 4: The ROC curves of soft drop mass selection for $\beta=0$ with 5, 10, 20, 40 TeV c.m. energies. Three different detector cell sizes are compared: 20×20 , 5×5 , and $1 \times 1 \text{ (cm}^2\text{)}$. The signal (background) process is $Z' \rightarrow WW$ ($Z' \rightarrow q\bar{q}$).

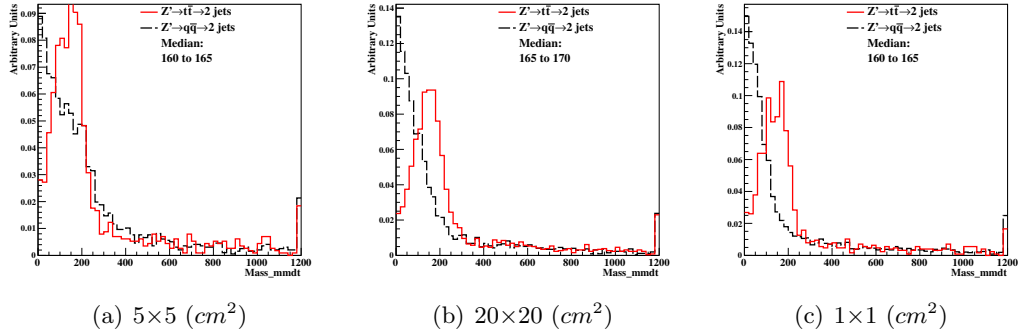


Figure 5: Distributions of soft drop mass for $\beta=0$, with 20 TeV c.m. energies and three different detector cell sizes: 20×20 , 5×5 , and $1 \times 1 \text{ (cm}^2\text{)}$. The signal (background) process is $Z' \rightarrow t\bar{t}$ ($Z' \rightarrow q\bar{q}$).

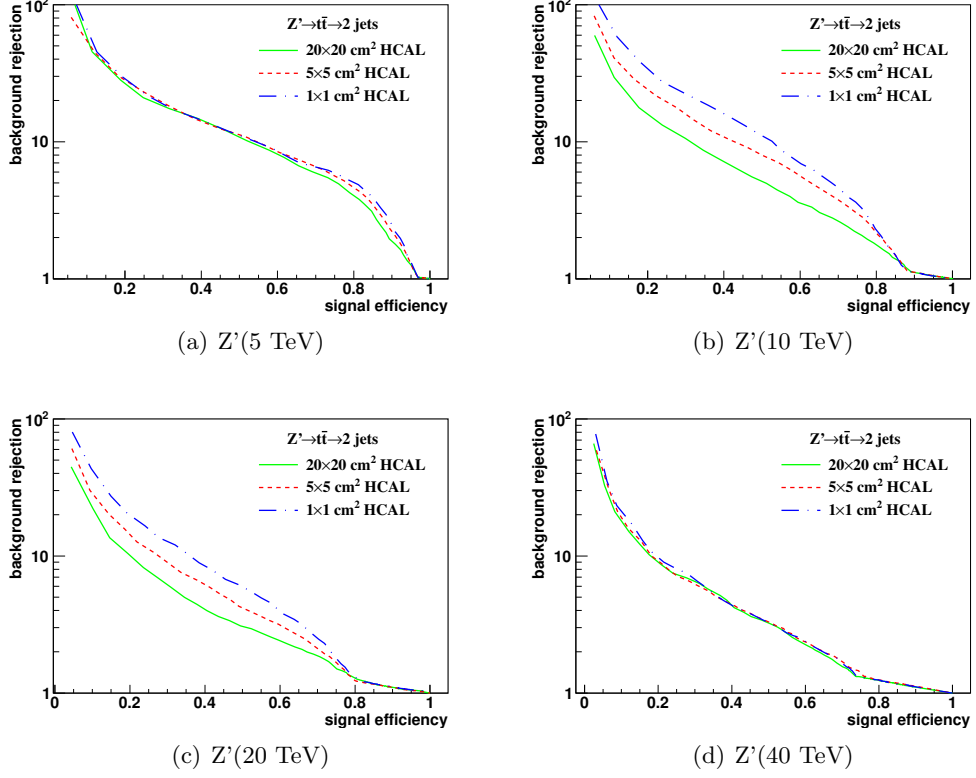


Figure 6: The ROC curves of soft drop mass selection for $\beta=0$ with 5,10, 20, 40 TeV c.m. energies. Three different detector cell sizes are compared: 20×20 , 5×5 , and 1×1 (cm^2). The signal (background) process is $Z' \rightarrow t\bar{t}$ ($Z' \rightarrow q\bar{q}$).

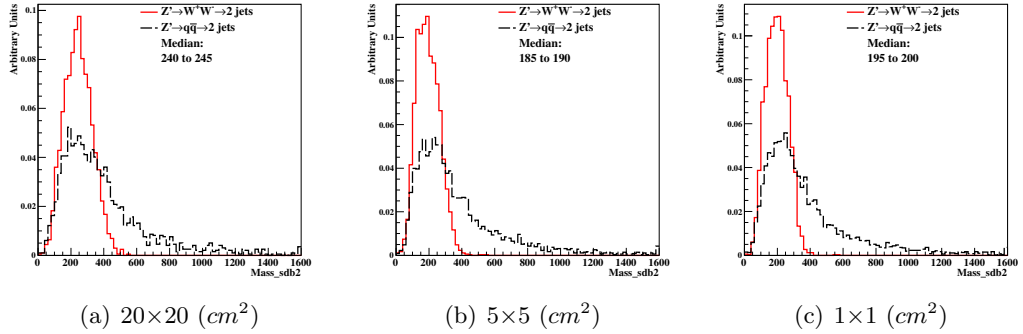


Figure 7: Distributions of soft drop mass for $\beta=2$, with 20 TeV c.m. energies and three different detector cell sizes: 20×20 , 5×5 , and 1×1 (cm^2). The signal (background) process is $Z' \rightarrow WW$ ($Z' \rightarrow q\bar{q}$).

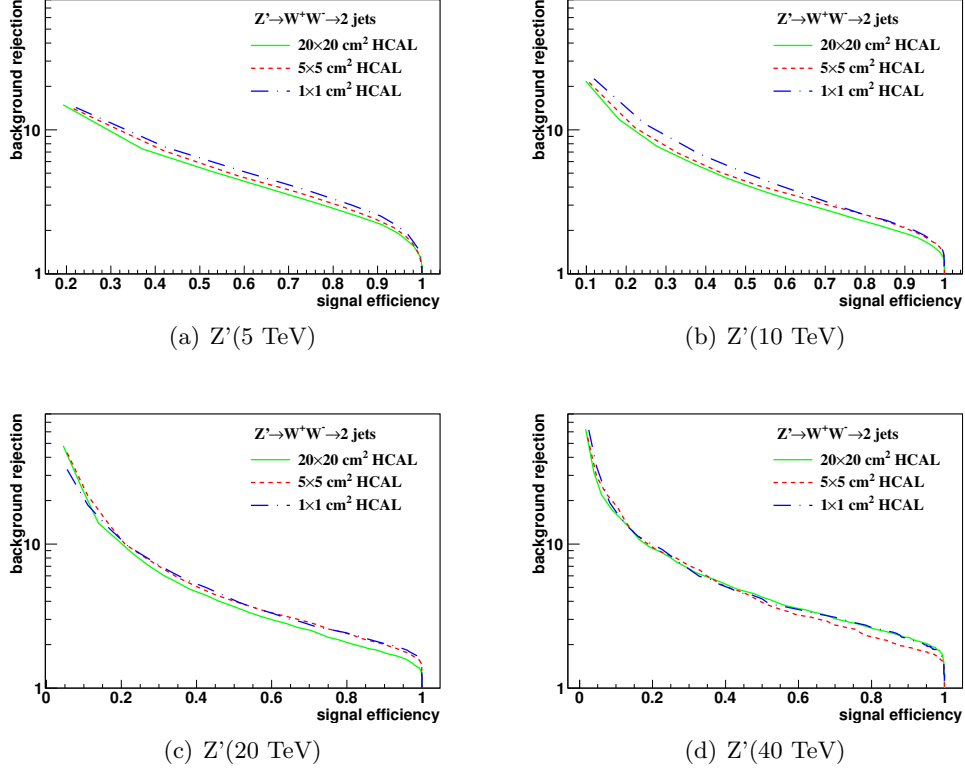


Figure 8: The ROC curves of soft drop mass selection for $\beta=2$ with 5, 10, 20, 40 TeV c.m. energies. Three different detector cell sizes are compared: 20×20 , 5×5 , and 1×1 (cm^2). The signal (background) process is $Z' \rightarrow WW$ ($Z' \rightarrow q\bar{q}$).

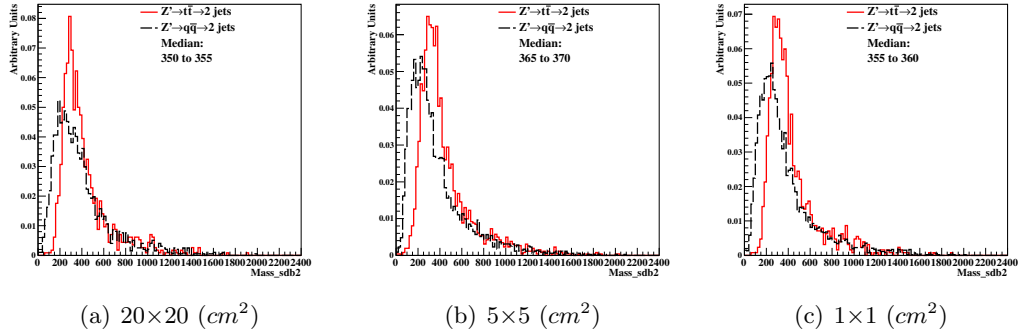
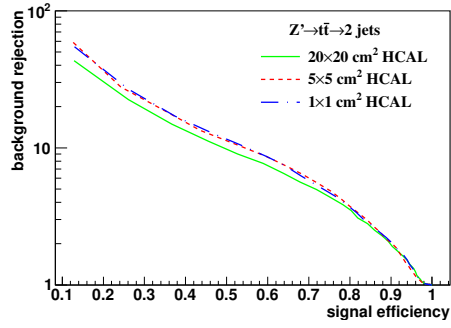
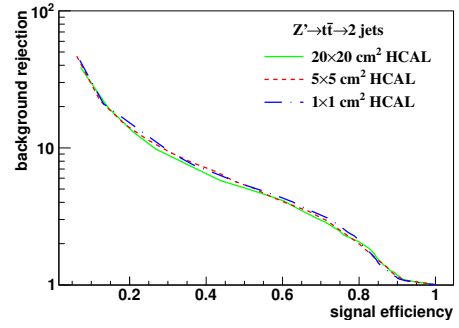


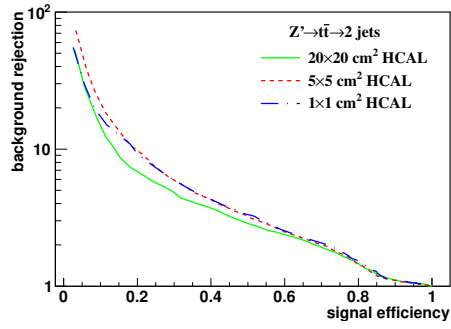
Figure 9: Distributions of soft drop mass for $\beta=2$, with 20 TeV c.m. energies and three different detector cell sizes: 20×20 , 5×5 , and 1×1 (cm^2). The signal (background) process is $Z' \rightarrow t\bar{t}$ ($Z' \rightarrow q\bar{q}$).



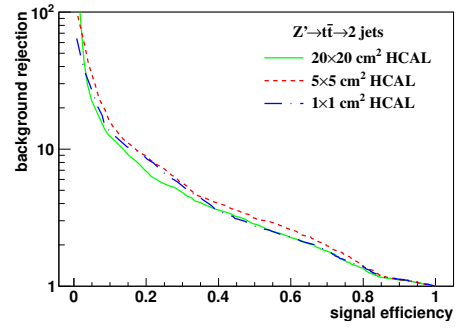
(a) $Z' (5 \text{ TeV})$



(b) $Z' (10 \text{ TeV})$



(c) $Z' (20 \text{ TeV})$



(d) $Z' (40 \text{ TeV})$

Figure 10: The ROC curves of soft drop mass selection for $\beta=2$ with 5, 10, 20, 40 TeV c.m. energies. Three different detector cell sizes are compared: 20×20 , 5×5 , and $1 \times 1 \text{ (cm}^2\text{)}$. The signal (background) process is $Z' \rightarrow t\bar{t}$ ($Z' \rightarrow q\bar{q}$).

other hand, if U value of two distributions are close to 0, it means both compositions of both distribution are much different from each other.

5.1. *N-subjettiness*

N-subjettiness[??] is the detection technique of jet substructure that is used to identify boosted hadronically-decaying objects under the high c.m. conditions. We use τ variable to distinguish the number of subjet in a fatjet to separate signal and background with different detector sizes and c.m. energy.

5.1.1. *The technic of N-subjettiness*

The formula and the technique are as following:

$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min\{\Delta R_{1,k}, \Delta R_{2,k}, \dots, \Delta R_{N,k}\} \quad (2)$$

$$d_0 = \sum_k p_{T,k} R_0 \quad (3)$$

k runs over all constituent particles in the given jets (fatjet), $p_{T,k}$ are their transverse momentum, $\Delta R_{J,k} = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ is the distance between the constituent particles k and the candidate subjet J on the $\eta - \phi$ plane. R_0 is the characteristic jet radius used in Anti-kt(AK) jet algorithm at starting. d_0 is the normalization factor.

1. First, Anti-kt(AK) algorithm is used to reconstruct jets
2. Second, after reconstructing the AK4 jets, exclusive k_T algorithm[??] is used in finding the jet axis in a fatjet.
3. Third, start running formula [2] and loop all constituent particles in a fatjet.
4. Finally, when finishing running all particles, it will give out τ_N , where N is positive integer.

If a fatjet has N subjet(s)[??], its τ_N is smaller than the τ_N of the fatjet with different number of subjets. For example, if we compare the τ_2 of one-subjet fatjet and two-subjets one, two-subjets fatjet has smaller τ_2 than one-subjet one. On the other hand, one-subjet fatjet has smaller τ_1 than two-subjets one. In the end, we can use the ratio of τ_2 and τ_1 (τ_{21}) to distinguish fatjet with one-subjet case and two subjets case. τ_{21} is used to discriminate the fatjet shape, and it can be modified with different number of subjet.

In our study, we use τ_{21} and τ_{32} in distinguishing two-subjets fatjet and three-subjets fatjet from one-subjet fatjet individually. We want to use this two ratio values to distinguish signal from background.

5.1.2. Analysis method

First, we select the events in mass window by using SD with $\beta = 0$ and 75% signal efficiency. Then, we find the highest ratio bin to be our seed bin. Next, we compare the left and right of ratio bin, and add the higher bin to be our width. Finally, We can use this width to draw the ROC curves.

2

5.1.3. The results and conclusion

In the figure [13][15], they show the histograms of τ_{21} and τ_{32} after selecting the events. In all figures, they also include the Mann-Whitney U value in ir the legend.

As a result of figure [14][16], they perform the ROC curves of τ_{21} and τ_{32} with different detector cell sizes and c.m. energy. The smallest detector cell (1×1) doesn't have the best separation power to distinguish signal from background. Some of them have the best separation power with the bigger cell size (5×5 and 20×20).

In Figure [17](a)(b), they show the summary plots of τ_{21} and τ_{32} with the rawhit cut with 0.5GeV using Mann Whitney U test. In τ_{21} , 5TeV has better separation power when detector sizes get smaller. When energy increases, there is no improvement in the smallest detector cell size (1×1). In τ_{32} , the case is similar to τ_{21} . Even worse, with some c.m. energies, the bigger detector sizes (5×5 and 20×20) have better separation power than the smallest detector sizes (1×1).

5.2. Studies of signal and background separation using jet substructure variable: Energy correlation function

Energy correlation function (ECF) [??] is another kind of detection technique of jet substructure that is used to distinguish the number of subjets in a fatjet under high c.m. energy conditions. This method only uses the momenta of particles and the angles between them without additional algorithm.

5.2.1. The technic of energy correlation function

The basic ECF formula is as following:

$$ECF(N, \beta) = \sum_{i_1 < i_2 < \dots < i_N \in J} \left(\prod_{a=1}^N E_{ia} \right) \left(\prod_{b=1}^{N-1} \prod_{c=b+1}^N \theta_{i_b i_c} \right)^\beta \quad (4)$$

In the formula 4, the sum loop all particles in the jet J , E are the energy of particles, and θ are the angles between the particles.

We apply two approximation. First, because under the high energy limitation $p \gg m$, $E \approx p$. Second, we use Radius R between particles naturally, so our ECF formula (4) can be modified to the formula:

$$ECF(N, \beta) = \sum_{i_1 < i_2 < \dots < i_N \in J} \left(\prod_{a=1}^N P_{ia} \right) \left(\prod_{b=1}^{N-1} \prod_{c=b+1}^N R_{i_b i_c} \right)^\beta \quad (5)$$

From the modified ECF formula (5), in order to use the dimensionless observation to determine whether the number of subjets in system, parameter τ_N is defined as:

$$\tau_N^{(\beta)} \equiv \frac{ECF(N+1, \beta)}{ECF(N, \beta)} \quad (6)$$

The idea of formula (6) is from N-subjetness, because the behavior of it is very similar to N-subjetness as reference [??]. In general, if the system has N subjets, $ECF(N+1, \beta)$ should be significantly smaller than $ECF(N, \beta)$, so we can use this advantage to distinguish different number of subjets. Finally, because it is suggested to used τ_{21} , τ_{32} [??] to distinguish two-subjets fatjet and three-subjets fatjet from one-subjet fatjet, in the ECF, it also defines the ratio of τ there, and define the energy correlation double ratio that is used in our study:

$$C_N^{(\beta)} \equiv \frac{\tau_N^{(\beta)}}{\tau_{N-1}^{(\beta)}} = \frac{ECF(N-1, \beta)ECF(N+1, \beta)}{ECF(N, \beta)^2} \quad (7)$$

We set $N=2$ and $\beta = 1$ (C_2^1) to distinguish two-subjets fatjet from one-subjet fatjet.

5.2.2. Analysis method

First, we select the events in mass window by using SD with $\beta = 0$ and 75% signal efficiency. Then, we find the highest ratio bin to be our seed bin. Next, we compare the left and right of ratio bin, and add the higher bin to be our width. Finally, We can use this width to draw the ROC curves.

5.2.3. The results and conclusion

In the figure [11], they show the histograms of τ_{21} and τ_{32} after selecting the events. In all figures, they also include the Mann-Whitney U value in their legend

As a result of figure [12], they perform the ROC curves of C_2^1 with different detector cell sizes and c.m. energy. The smallest detector cell (1×1) doesn't have the best separation power to distinguish signal from background. In addition, in some cases such like (a), the biggest one (20×20) has the best distinguish power under the same c.m. energy.

In Figure [17](c), it shows the summary plots with the rawhit cut with 0.5GeV using Mann Whitney U test. For conclusion, all separation power aren't improved by smallest cell size (1×1).

Acknowledgements

This research was performed using resources provided by the Open Science Grid, which is supported by the National Science Foundation and the U.S. Department of Energy's Office of Science. We gratefully acknowledge the computing resources provided on Blues, a high-performance computing cluster operated by the Laboratory Computing Resource Center at Argonne National Laboratory. Argonne National Laboratory's work

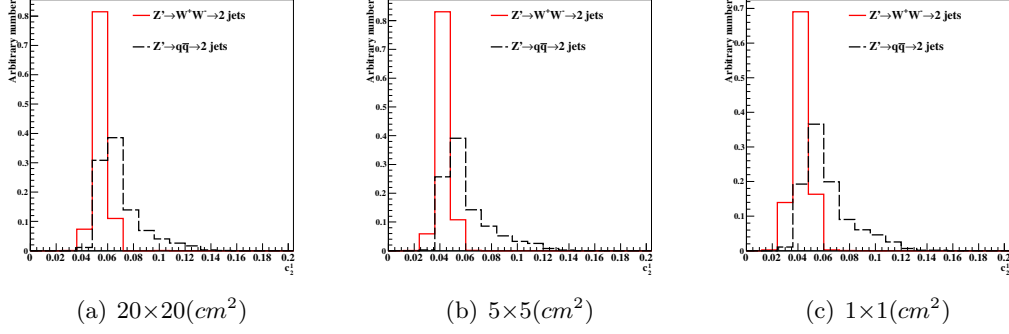


Figure 11: Distributions of Mann-Whitney value U in 20 TeV energy collision for c2b1 in different detector sizes. Cell Size in 20×20 , 5×5 , and 1×1 ($\text{cm} \times \text{cm}$) are shown here.

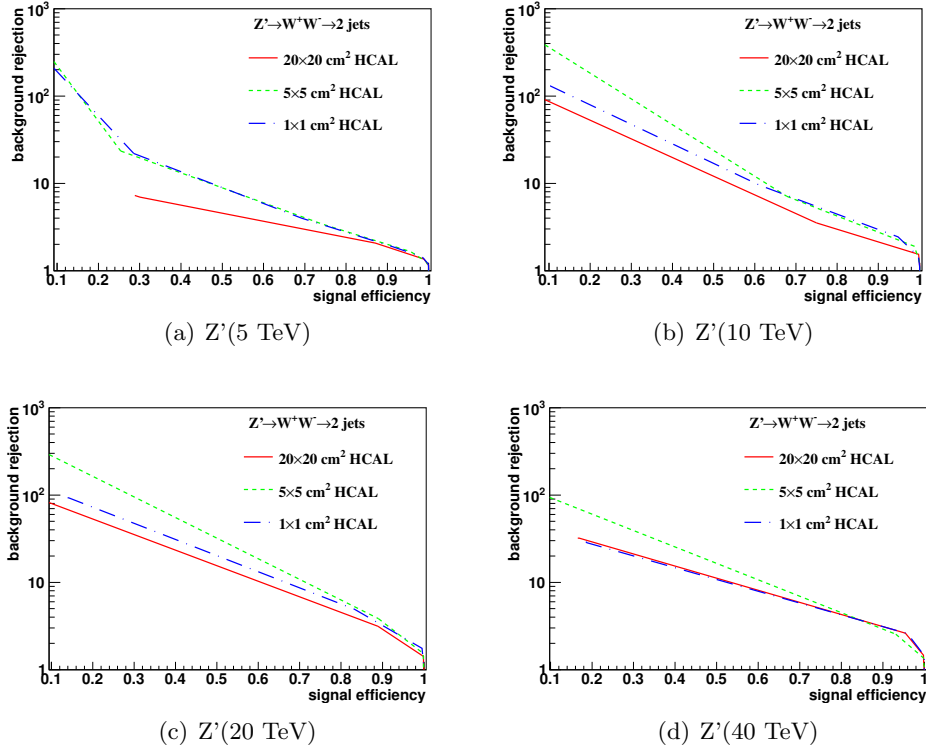


Figure 12: Signal efficiency versus background rejection rate using c2b1. The energies of collision at (a) 5, (b) 10, (c) 20, (d) 40 TeV are shown here. In each picture, the three ROC curves correspond to different detector sizes.

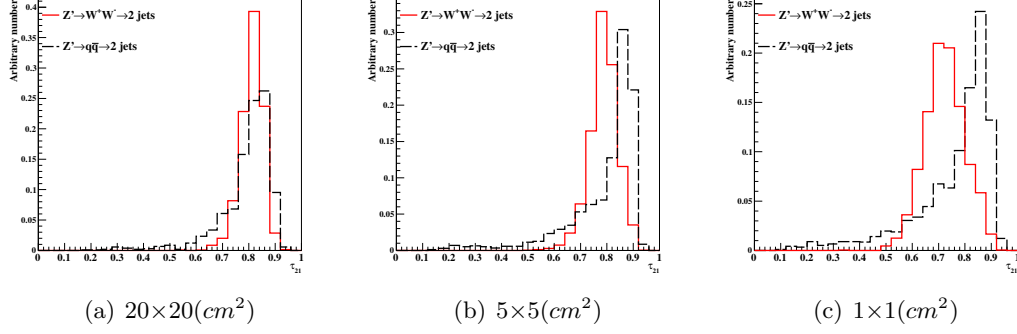


Figure 13: Distributions of Mann-Whitney value U in 20 TeV energy collision for τ_{21} in different detector sizes. Cell Size in 20×20 , 5×5 , and $1 \times 1 (cm \times cm)$ are shown here.

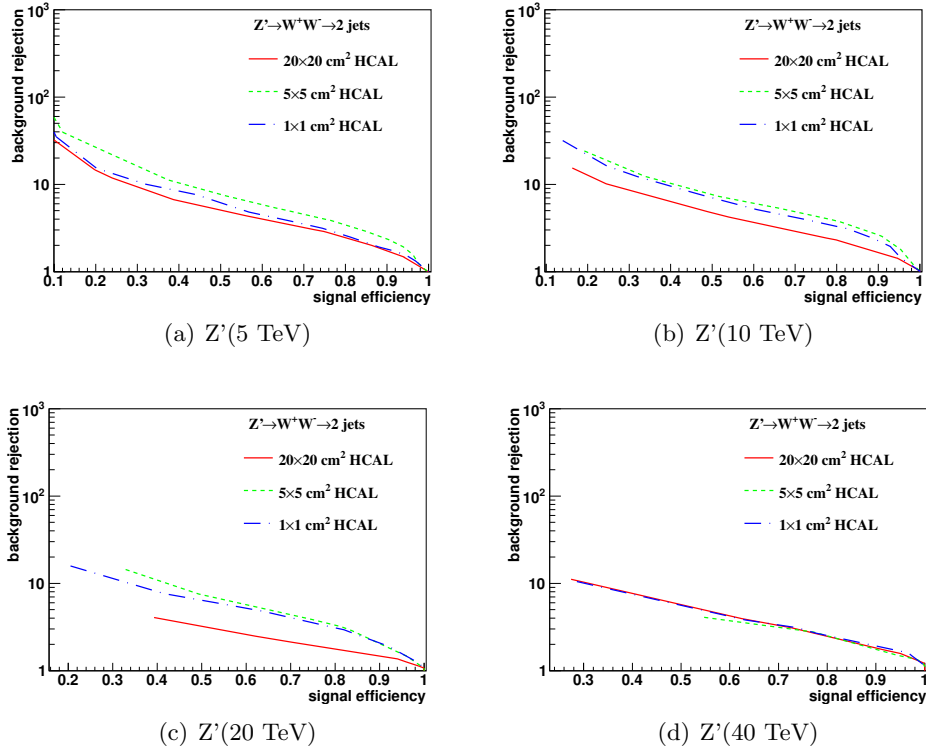


Figure 14: Signal efficiency versus background rejection rate using τ_{21} . The energies of collision at (a) 5, (b) 10, (c) 20, (d) 40 TeV are shown here. In each picture, the three ROC curves correspond to different detector sizes.

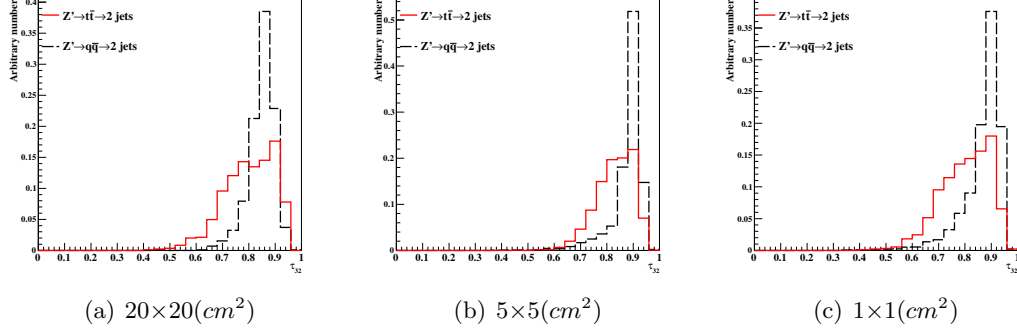


Figure 15: Distributions of Mann-Whitney value U in 20 TeV energy collision for τ_{32} in different detector sizes. Cell Size in 20×20 , 5×5 , and $1 \times 1 (cm \times cm)$ are shown here.

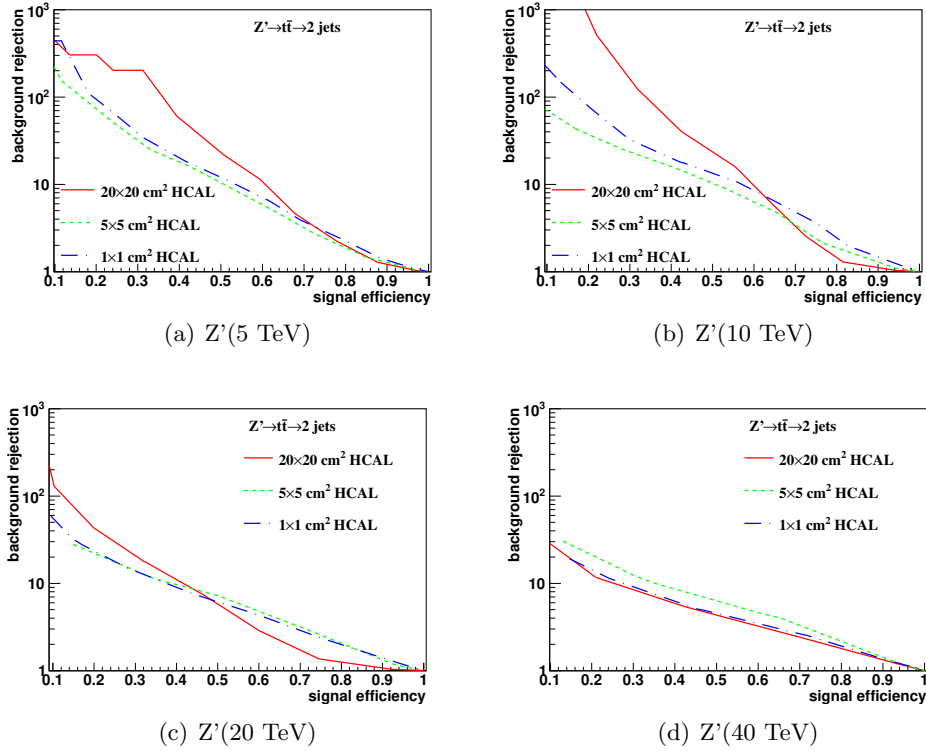


Figure 16: Signal efficiency versus background rejection rate using τ_{32} . The energies of collision at (a) 5, (b) 10, (c) 20, (d) 40 TeV are shown here. In each picture, the three ROC curves correspond to different detector sizes.

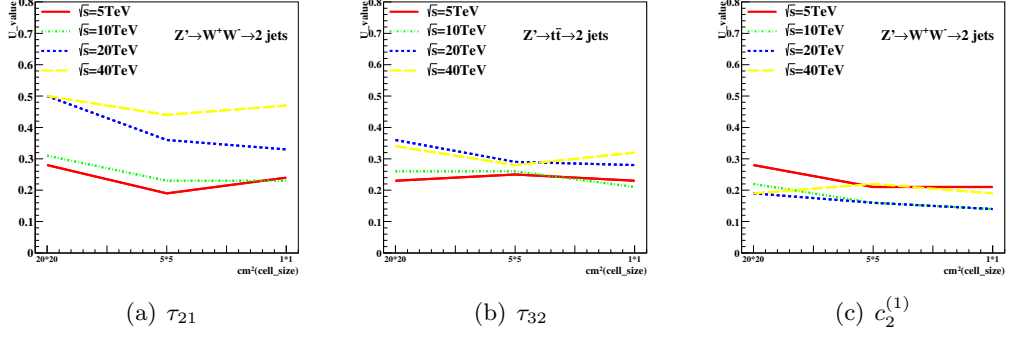


Figure 17: The Mann-Whitney U values for τ_{21}, τ_{32} and $c_2^{(1)}$ reconstructed from calorimeter hit at 05GeV cut with different collision energies correspond to different detector sizes in rawhit cut with 05GeV. The energies of collision at 5, 10, 20, 40, 20, 40TeV are shown in each figure.

237 was supported by the U.S. Department of Energy, Office of Science under contract DE-
 238 AC02-06CH11357. The Fermi National Accelerator Laboratory (Fermilab) is operated
 239 by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the
 240 United States Department of Energy.

241 **References**