

# **Detector Traveling from Now to the Future: Study of the Calorimeter Performance for Future Colliders with the High Granularity Calorimeter in CMS and the Full Simulation based on FCC-hh**

Chih-Hsiang Yeh

Department of Physics and Center for High Energy and High Field Physics, National  
Central University, Chung-Li, Taiwan

## **Abstract**

The discovery of the standard-model-like Higgs boson at Large Hadron Collider(LHC) in 2012 opened a new era as well as unexplored windows for particle physics. Since then, people have been working hard on upgrading the LHC to higher luminosities and center-of-mass energy, in order to study the properties of Higgs boson and to discover new particles predicted by the physics beyond the standard model. Some people even proposed the idea of building new colliders, such as future circular collider(FCC), etc. locating adjacent to the LHC. My study mainly focuses on the topics that are related to the performance of detectors. The endcap part of the compact muon solenoid(CMS) at LHC will be replaced with a silicon-based high granularity calorimeter (HGCAL) owing to the issues of ascending pileup and insufficient radiation tolerance. First, I did the studies on the performance for HGCAL using data from test beam in CMS and the hexaboard of HGCAL electronics in National Taiwan University. Second, I present the performance of very-high-energy jets at FCC using the GEANT4 full simulation. We expect to help the future facilities to design a high resolution detector, and could have a large potential to which may shed light on new particles.



## Contents

1	Introduction . . . . .	2
1.1	Motivation and Expectation . . . . .	2
1.2	Summary for research topics . . . . .	2
1.3	Summary for Milestone Achieved within the project time . . . . .	3
2	Cross-talk studies with the silicon-sensor-unconnected hexaboard in NTU . . . . .	3
2.1	Experiment apparatus and Event reconstruction . . . . .	4
2.2	Events and Methods for analysis . . . . .	4
2.3	Results and Conclusion . . . . .	6
3	Pion-rejection studies with a CMS HGCALE test-beam prototype ECAL .	7
3.1	Introduction for HGCALE . . . . .	10
3.2	Introduction of the optimized cuts and new variable with June test-beam MC . . . . .	12
3.3	The results of application with October test-beam MC . . . . .	17
3.4	Conclusion . . . . .	18
4	Studies of granularity of a hadronic calorimeter for tens-of-TeV jets at a 100 TeV $pp$ collider . . . . .	18
4.1	Simulation of detector response and event reconstruction . . . . .	19
4.2	Study of detector performance with soft drop mass . . . . .	21
4.3	Studies of detector performance with jet substructure variables- N-subjettiness . . . . .	23
4.4	Studies of detector performance with jet substructure variables- Energy correlation function . . . . .	28
4.5	Conclusion . . . . .	29
5	Conclusion for all studies . . . . .	30
6	Acknowledgement . . . . .	30

# 1 Introduction

## 1.1 Motivation and Expectation

After identifying the Higgs boson at LHC in 2012 [1] [2], people are eager to explore new particles and the physics beyond the standard model [3], such as dark matter candidates [4], heavy Higgs [5]... etc. and so on. The CMS [6] and ATLAS [7] collaborations are also upgrading the detectors to improve the functions to cope with the new challenging conditions in the future colliders. For the future operation, one of the most important issues of detector performance is: How can we boost the efficiency of the distinguishability of signal from the background?

In the upcoming High Luminosity- LHC (HL-LHC) [8] era, the pileup will significantly increase compared with the condition at which the LHC is operating currently [9]. Therefore, first of all, improving the efficiency of identifying particles will be our top priority. Otherwise, we are not able to reject the unwanted contaminant particles very well. My first topic with HGCAL[10] in CMS focuses on the particle identification of electrons and pions. It is expected that there will be more new physics related to electrons as final states [11]. Nonetheless, the misidentification of pions as electrons will certainly be one of our primary issues, and the removal of these mistagged pions is the goal of our study. The outcome of the optimized cuts and their application on the cut-based analysis will benefit the signal/background studies that are very sensitive to pions [12]. Second, I study the cross-talk phenomena on the PCB boards(called "hexaboard") which are attached with the sensor used in the HGCAL, and try to find out the correlation between the injection pulse strength and cross-talk. We expect to understand the noise created in the electronics and reduce them in the future.

Some have proposed colliders with the higher center of mass (C.M.) energy for the next generation [13], such as Future Circular Collider (FCC) [14], Circular Electron-Positron Collider (CEPC) [15], Super proton-proton Collider (SppC) [16] etc. When those colliders operate, other crucial problems with higher C.M. energy will arise. Under these circumstances, the jets from the segmentation will be very boosted [17]. In this case, we need to explore other ways to look into the structure of boosted jets. Otherwise, we can't separate the signal from the background very well. My topic with FCC detector focuses on the boosted conditions, using different jet substructures with various configurations of the FCC detector to see whether the smallest detector configuration will give the best separation power to distinguish signal from the background.

## 1.2 Summary for research topics

The research activities of Chih-Hsiang Yeh from July 1, 2018 to February 28, 2019 include the following:

- Research with electronics of HGCAL : Cross-talk studies with the silicon-sensor-unconnected hexaboard in NTU
- Research with HGCAL in CMS : Pion-rejection studies with a CMS HGCAL test-beam prototype electromagnetic calorimeter(ECAL)
- Research with simulation FCC detector : Studies of granularity of a hadronic

calorimeter for tens-of-TeV jets at a 100 TeV  $pp$  collider

### 1.3 Summary for Milestone Achieved within the project time

- My first journal paper: "Studies of granularity of a hadronic calorimeter for tens-of-TeV jets at a 100 TeV  $pp$  collider "was submitted to arxiv [18] and the Journal of instrumentation(JINST). Now the paper is being reviewed by the referees of JINST.
- My first conference paper: "Jet Substructure Variables with the SiFCC Detector at 100 TeV" was submitted to arxiv, which was published by the Journal of Proceeding of Science(PoS) [19].
- The poster with FCC detector studies was presented in the International Conference of High Energy Physics in 2018(ICHEP2018).
- Two posters with HGCal studies and FCC detector studies were presented in the Taiwan Physics Society(TPS) annual meeting in 2019.

## 2 Cross-talk studies with the silicon-sensor-unconnected hexaboard in NTU

Cross-talk [20] is the one of the basic phenomena that exist in electronics, which a signal is transmitted on to one circuit or channel of a transmission system that creates an unwanted effect in another circuit or channel. For instance, there are two types of cross-talk that could happen:

- Correlated cross-talk  
The capacitors in the circuit are too closed to each other. Other capacitors might be affected if one of the adjacent capacitor is charging.
- Anti-correlated cross-talk  
The problem of power insufficient could exist in the circuit when one of the channels does not have enough energy, which would lead to the energy transfer from one another channel, creating also the phenomenon of cross-talk.

Thus, either of the circumstances described above would give rise to incorrect signal to the electronics. The issues need to be analyzed further more, or there will be significant impact when applying the results to reconstruct the events using in High energy physics(HEP) experiment.

I did this study with Prof. Stathes Paganis, Prof. Rong-Shayang Lu, master student Chia-Hong Chein from NTU, and Prof. Shin-Shan Eiko Yu from NCU. The contributions of Chih-Hsiang Yeh to this study includes the following:

- Study the correlation between the injection pulse strength and the cross-talk.

I will describe the detail as follows.

## 2.1 Experiment apparatus and Event reconstruction

The electronics we used are composed of three parts, the module (version 2), RPI Hexa-board and Raspberry Pi 3(RPI). The component of them are as follows:

- For the module(Fig.1 Top left), it also be called "HexaBoard", due to its hexagonal-like shape. The main function of it is to collect analog signals and convert them into the digital format with the four SKIROC2cms CHIPS [21] on it. It can be bonded to silicon sensor, such as the HGCAL, to collect real charge from particle to the sensors. The electrical pulse can be manually injected to simulate the condition when the sensors are attached.
- For the RPI Hexa(Fig.1 Top right), it bridges the signal between FPGAs and Raspberry Pi (RPI). When the signal comes out from the four CHIPS of hexa-board, initially, they will be stored to the secondary "slave" FPGA on the HexaBoard. Then after finishing collecting the data, the "slave" FPGA will transfer the data to MAX10 FPGA("master" FPGA) on the RPI Hexa.
- For Raspberry Pi 3(RPI) (Fig.1 Top right), acting like the small computer, can communicate with the hexaBoard. The main power for it is to receive the MAX10 data and decode the raw data to the formal data which can be applied in analyzing directly. Until now, we finish reconstructing the events for one run.

Every SKIROC2cms CHIP on the HexaBoard corresponds to 64 channels, while each channel has its own readout circuit(Pre-amplifier, shapers,etc.). In each event(run), it records 30 numbers to reconstruct the event for every channel. Total 30 numbers are given by 13 ADC counts in both highgain(HG) and lowgain(LG) plus 2 TOTgain(Time Over Threshold) and 2 TOAs(Time Of Arrival). The 13 ADC counts of HG and LG come from 13 SCA(Switched Capacitor Array) units in the circuit, which sample the input signal every 25ns<sup>1</sup>. They can be employed to define the hardware noise and differences between each capacitor. At the same time, we records the time stamp with the pulse trigger<sup>2</sup> for each event. Time stamp labels the whole pulse in order, so we can use the time stamp to mark the pulse location, specially for the peak of pulse.

## 2.2 Events and Methods for analysis

In order to simplify the study, the sensors are taken out from the HexaBoard. In this studies, we used two types of run to do the researches through the whole process:

- Pedestal run: The runs without injecting the electrical pulses are referred to as the pedestal runs.
- Charged run: The runs with injecting the electrical pulse to certain channels

<sup>1</sup>There exist thirteen different capacitors with the number SCA0, SCA1....SCA12, and they will record one ADC value in order every 25ns for the each channel of CHIPS. When going to SCA12 and finishing recording, it will return to SCA0. The order of SCA numbers are not related to the order of the electrical pulse we give, but the time stamps are.

<sup>2</sup>In real case, it records the starting SCA and the end of SCA with roll position array when the pulse trigger is on and off, and rearrange the SCAs with this array as the time stamp, so it can be specified as the pulse information.

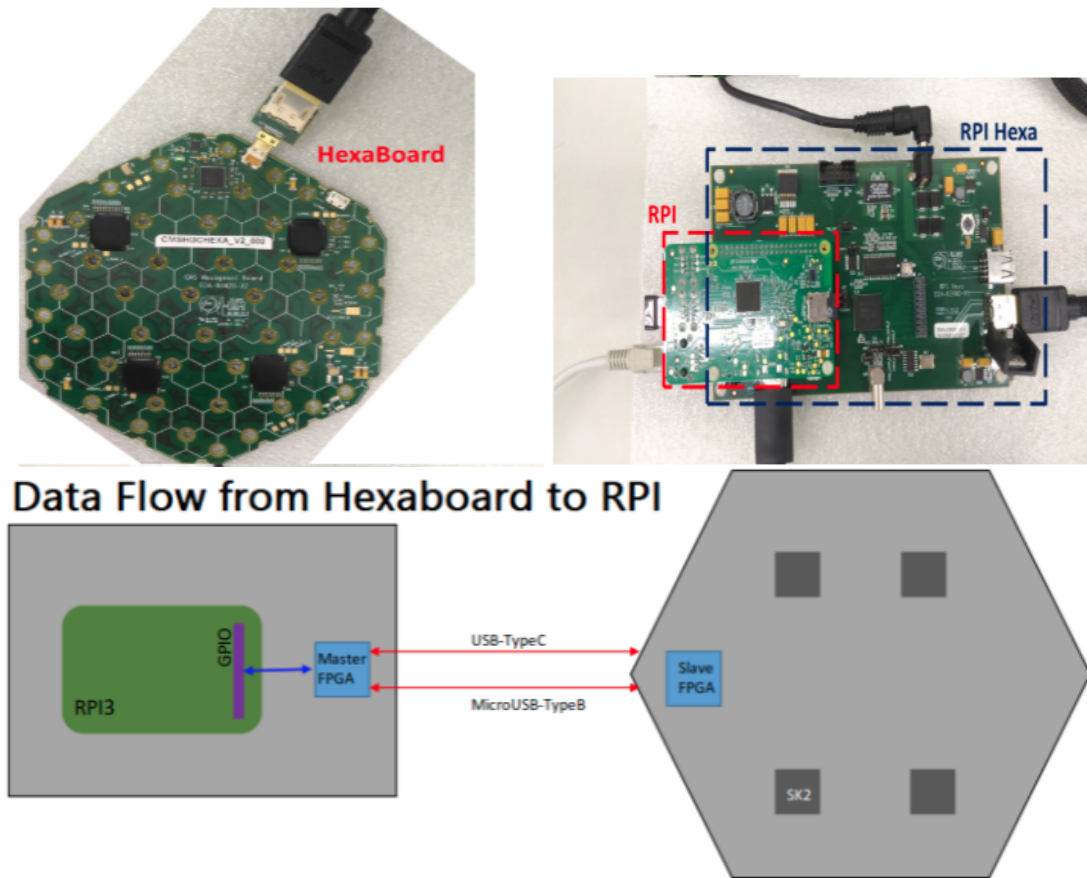


Figure 1: (Top left) The Hexaboard which was applied in the study. On the board, there are four SKIROC2cms CHIPS (black hexagonal) used to receive the electrical pulse which is given by the cable, and give out the signal to the Hexa RPI. (Top right) The Hexa RPI and RPI which is applied to transform the data and communicate with Hexaboard. (Down) The schematic plot illustrates the transfer of the data flow from Hexaboard to Hexa RPI and RPI.

are referred to as the charged run.

Notice that only 32 channels of the Hexaboard are simulated in the sensor-attached scenario due to the fact that only 32 channels are applied in reality. In the Fig.2 is the channel map that we used in the study.

The following three cases which were used in our studies are shown in the Fig.3.

- Pedestal run is used to evaluate the pedestals and noise for each channel. The histogram of the mean value of ADC-counts (Fig.3 Top left) is defined as the pedestal value from each channel, and it is the hardware-dependent(SCA-dependent) value.
- Charged run is used to record the responses with the histogram of ADC-counts values(Fig.3 Top right) from every channel with the electrical pulse injecting into the fixed-channel, and is applied to simulate the sensor-unconnected condition.
- "Pedestal-subtracted" value, the most important value for our study, is used

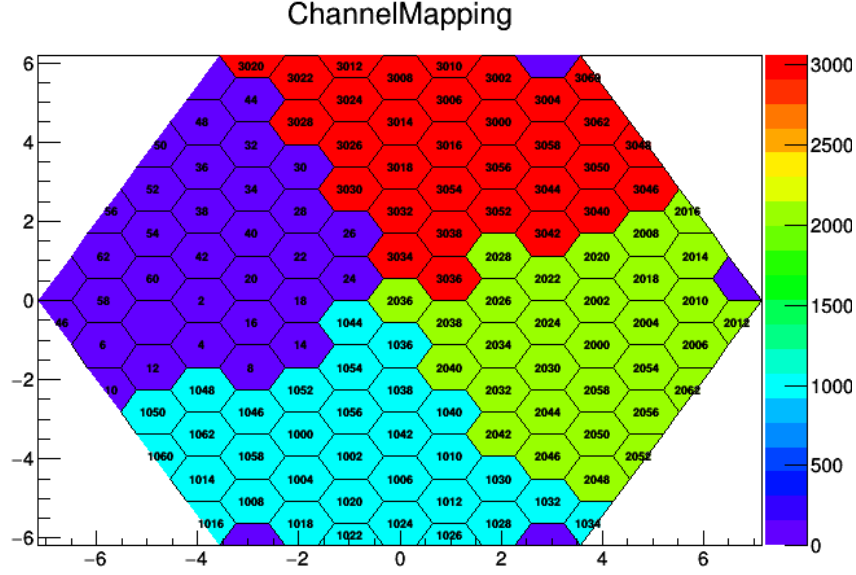


Figure 2: The channel map which was applied in the research, there are four regions with the different color represent the different SKIROC2cms CHiPs. A digit in thousands is the order of the CHiPs, and a digit in tens and ones represent the channels.

to subtract the "reference". The mean value of the ADC-counts for it is turned out after "Charged run ADC-counts" cut-off "the mean of ADC-counts from Pedestal run" with same SCA number. The error (width of gaussian) of the histograms of the already-subtracted values(Fig.3 Down) means the noise in the channels including the electronics noise, sensor noise(if it is installed, not in our case. ), etc.

Note, in the study, we always used SCAs of HG to do. So as to simplify the case, we fixed the injected-channel to number 20 for all CHiPs and saw all cases.

## 2.3 Results and Conclusion

We desired to explore the correlation between injection pulse strength and the cross-talk. In real physics, because particles own their energies individually in the collider, we could detect the different pulse strength in reality. At there, we wanted to see whether the different pulse strength will produce a different amount of the cross-talk, and expected to quantify them as well. In the Fig.4, they show that the mean and error values from Pedestal-subtracted value which is depicted above for different DAQ(strength unit for pulse) of injection pulse. The results for them are quite obviously showing that the cross-talk in two cases:

- Anti-correlated cross-talk  
DAC from 0 to 1000: No matter the closer or farther the channels, the slopes



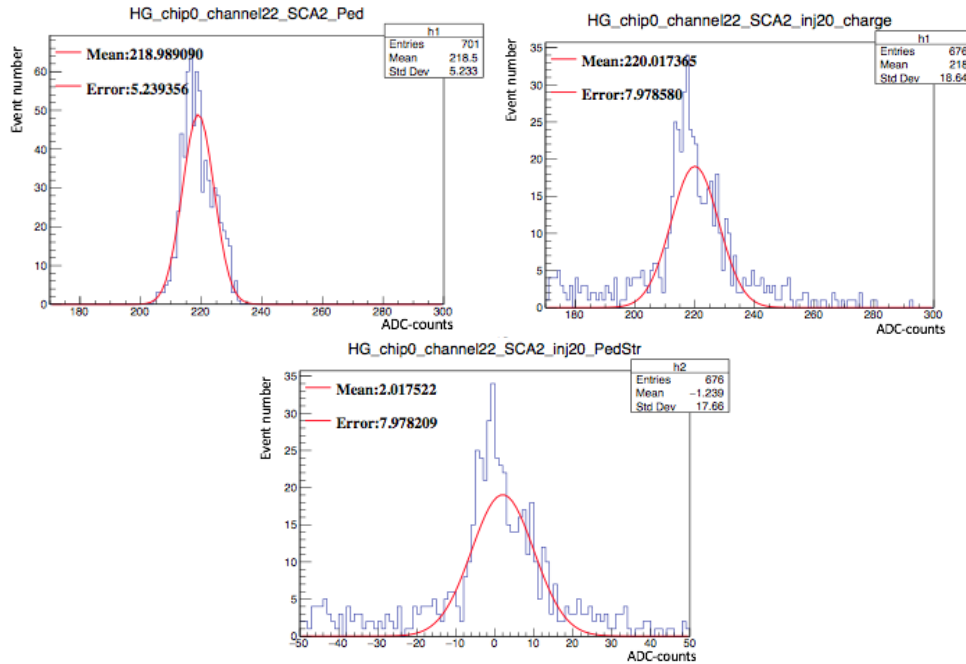


Figure 3: These figures perform the histograms that using in the study. All this cases are shown with fixing at SCA=2 and channel=22. (Top left ) The pedestal run histograms which is used to calculate the pedestal and noise. Fitting the line is to get the mean of ADC-counts. (Top right) The charged run for example. (Down) The pedestal-subtracted case is shown. Some of fitting cases are used to get the value of mean and error, taking to the calculation we apply in study.

are negative for all of them. That means the channels, except the injection channels, are taken out the energy by injection channels.

- Correlated cross-talk

DAC after 1000: When the channels are closed to the injection channels, the slopes are bigger compared with the channels which are farther from the injection channels. That means, the injection channels leak out more energy to the closed channels compared with the farther channels.

For the conclusion, the study indicates that we can't throw away the concerned issue of the noise of cross-talk, unfortunately, which was explored to exist in the Hexaboard and could affect the resolution of reconstructing the events. We used the ADC-counts from the Pedestal-subtracted values and let the cross-talk quantified properly. We expect that this study can help to quantify the cross-talk in the future and probe the way to get rid of it.

### 3 Pion-rejection studies with a CMS HGCAL test-beam prototype ECAL

In general, the electrons deposit most of the energy in ECAL by the processes of pair production and bremsstrahlung. Different from them, charged-pions store most of the energy in Hadronic Calorimeter(HCAL) through the process of segmentation and

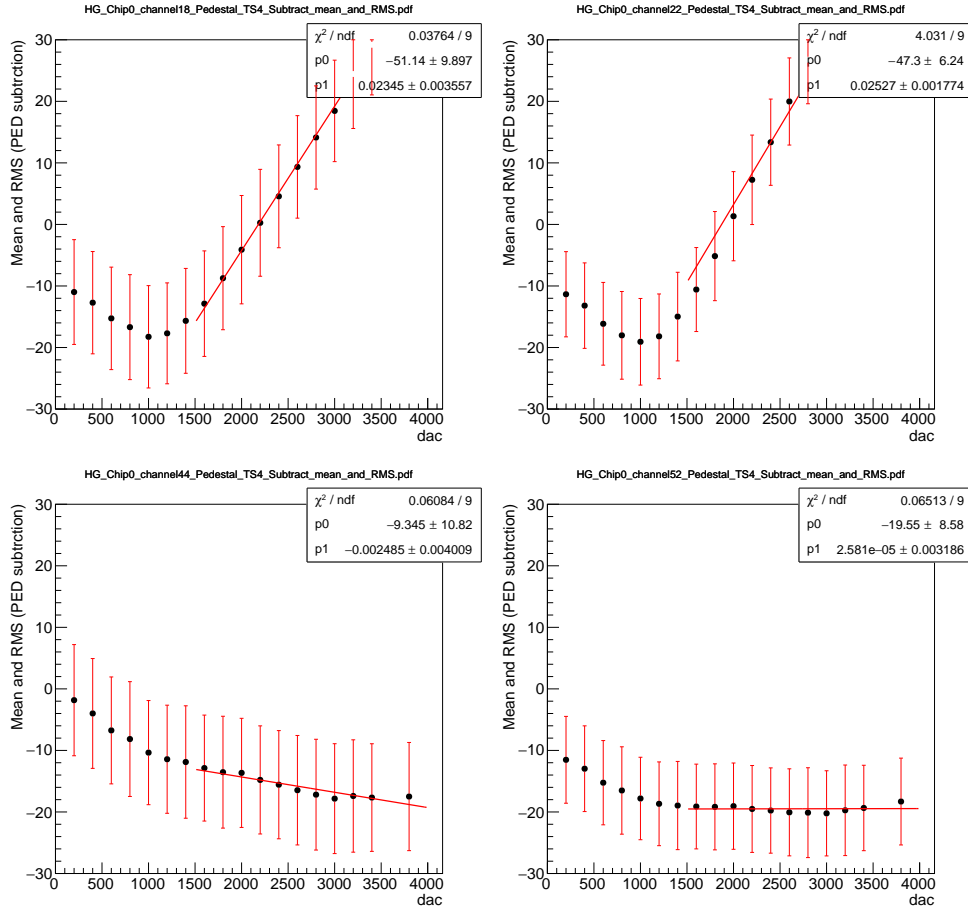


Figure 4: These figures present the preliminary results for the correlation between the different DAQ input and Pedestal-subtracted mean and error. The Top left and right pictures perform the channel of 18 and 22, which are closed to the injection channel. The down left and right pictures show channel of 44 and 52, which are farther from the injection channel.

hadronization. Unfortunately, in some special cases, those charged-pions could kick out the secondary particles, i.e. electrons and happen the fully Electromagnetic(EM) showering, along with leaving most of the energy in ECAL, violating with our general case. Even worse, those charged-pions which pass through the detector originally could get disappear momentarily in ECAL, and cause the problem with identifying them. If we can't explore the means to tag those pions and reject them, it will be very terrible. Because they will be misidentified as the electrons, and those "fake electrons" could contaminate with the real electrons. Furthermore, it could influence the analysis which is sensitive to the electron/pion distinguishability. We expect to handle this issue and help the physics analysis with this type of background in the future.

Since the significant challenges for radiation tolerance and event pileup on the detectors of the CMS experiment at LHC, they will replace the current Endcap calorimeter with a Si-pad HGCAL, a new generation state-of-the-art calorimeter, which can perform 3D imaging of the shower as well as provide 30ps timing. To test the performance of HGCAL, including calibration, particle identification, etc. before it will be

installed in CMS Phase II upgrade, we conduct the "test-beam"(Fig.5), which is held to incident the test particles, such like electron, pion and muon into HGCAL prototype regularly, and analyze its performance from the information of the hits come out from the detector. At the same time, The Monto Carlo(MC) samples are provided with the same prototype of the detectors to let the simulation of the experiment studied as well. We applied the test-beam MC samples which were done in CMS on June and October 2018 to do this study. The group of test-beam divided the studies into various topics, the topic of mine was allocated into the section "electron/pion identification".



Figure 5: The figure show the test beam experiment setup on October. The orange and red lines box mark the electromagnetic compartment(CE-E) and hadronic compartment (CE-H) part of the detector including the module(in the metal box) and readout electronics.

Charged-pion tagging and rejection are one of the strong points of HGCAL. Traditionally the most powerful discriminants for pion rejection are lateral shower containment and longitudinal energy leakage in the back of the ECAL. In this study, we investigate the capability of HGCAL to tag pions that yield the fully EM-showering in the ECAL and they passed the transverse containment and leakage from the back requirements. For our expectation, we want to find the best cuts and introduced the new variables to distinguish "fake electrons" from "real electrons".

I did this study with Prof. Stathes Paganis, Prof. Rong-Shayang Lu, master student Chia-Hong Chein from NTU, Dr. Shilpi Jain from University of Minnesota and Prof. Shin-Shan Eiko Yu from NCU. The contributions of Chih-Hsiang Yeh to this study include the following:

- Introduce the optimized cuts for finding the "fake electron" from pions, and the new variable with the longitudinal segmentation using the test-beam

MC samples from June.

- Apply the found optimized cuts and new variable in test-beam MC samples from October to see the electron efficiencies and pion rejections.

### 3.1 Introduction for HGICAL

During the Run 1 (2010-2012) [22] in LHC operated at  $\sqrt{s} = 7$  TeV, delivering  $\approx 6 \text{ fb}^{-1}$ , and at  $\sqrt{s} = 8$  TeV in 2012, delivering  $\approx 23 \text{ fb}^{-1}$ . The most significant physics results from this period was the discovery of the Higgs boson, and along with the Nobel prize for it in 2013. After Run 1, Run 2 [23] started in 2015 at a C.M of  $\sqrt{s} = 13$  TeV and the instantaneous luminosity, reaching to the value for  $1.7 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ . In the Figure.6, it presents the summary plot for the luminosity from Run1 to Run2. Surprisingly, it operated the surpass value compared with the original design. More studies with this unpredictable operation, such as Higgs boson, SM processes, BSM, will be carried out.

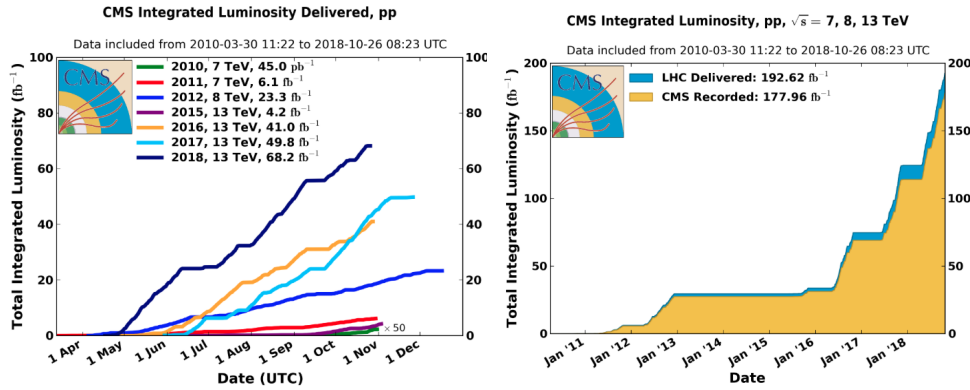


Figure 6: The figures show the summary plot of luminosity from the Run1(2010-2012) to Run2(2012 to 2018) in LHC. The left is the luminosity for the individual years, and the right is the totally integrated luminosity from 2010 to 2018.

For the future in Run3(2023)[24], LHC intended to accumulate around  $300 \text{ fb}^{-1}$ . After the third long shutdown (LS3), they prepare to design the value to the instantaneous luminosity at  $5 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$  with the target of integrating around  $3000 \text{ fb}^{-1}$  by the time of the mid-2030s, and it will go into the HL-LHC era. The corresponding mean number of collisions (pileup) per bunch crossing will be 140. In addition, the LHC has the ability to deliver 50% higher values for both the instantaneous and integrated luminosities.

In this case, It will bring out the two significant issues in the HL-LHC era: (1) radiation damage and (2) event pileup on CMS detectors, especially for calorimetry in the forward region. As part of its HL-LHC upgrade program, the CMS Collaboration is proposing to build a HGICAL to replace the existing endcap calorimeters. There are many requirements for the HGICAL upgrade to let the good performance preserved after working the period of time. Write some of them as follows:

- This detector must be "radiation tolerance", otherwise, it will lead to the unrecoverable condition, same as the circumstance of scintillator now in CMS. To solve this issue, the active layer with the silicon sensor [25] will be applied in the HGICAL.



- This detector needs to be designed with the fine granularity for lateral and longitudinal segmentation:
  - For the lateral part, it can help with separating two showers and can observe the narrow jets.
  - For the longitudinal part, it can help us to probe the longitudinal development of showers, providing good energy resolution, and also can know more about the patterns in the physical processes.

Both of them can remove the "pileup event" if they are great enough to distinguish the processes of different particles.

The HGCAL prototype consists of an electromagnetic compartment (CE-E) followed by a hadronic compartment (CE-H) with Forward region(FH) and Backward region(BH) in the Fig.7. For the CE-E, it consists of 28 sampling layers with a depth of approximately  $26 X_0$  and  $1.7 \lambda$ . The element of active layer for this part is a hexagonal silicon sensor, which is sandwiched between layer of WCu (75%, 25%) baseplate and a printed circuit board(PCB), which is used to study my cross-talk studies in NTU, that carries the front-end electronics to form a silicon module in the Fig.8 top right. These modules are tiled on a Cu cooling plate, which together with the two WCu baseplates form one absorber layer. The alternate absorber layer is formed by lead planes clad with stainless steel (SS) sheets that are placed on module-cooling plate sandwich. In the Fig.8 top left shows one sampling layer of the CE-E for example.

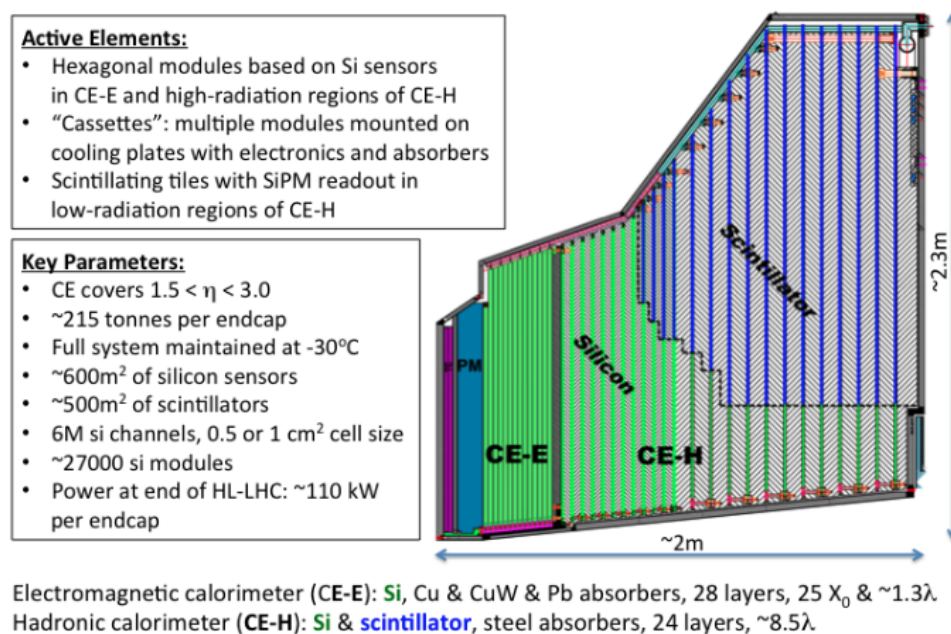


Figure 7: Summary for the basic design and coverage of the CMS high-granularity Endcap Calorimeter (CE) including CE-E and CE-H

For the hadronic compartment of HGCAL, they consist of 12 planes of thick SS plates followed by another 12 SS planes with different thickness. Between these absorber plates sit silicon modules(show one sampling layer in Fig.8 down left) in most regions with high-radiation and silicon modules mixed with scintillator tileboards(show one sampling layer Fig.8 down right) in part of low-radiation regions mounted on a copper

cooling plates to form the wide cassettes. These cassettes are similar to those in the EM compartment, but include sensors on only one side of the cooling plate(See Fig.8 down left), which are formed as a separate mechanical structure. This leads to a total interaction length with number  $10.7\lambda$ . All layers are read out for use in energy measurement, but only alternate layers in CE-E and all in CE-H, are used for producing L1 trigger primitives [26].

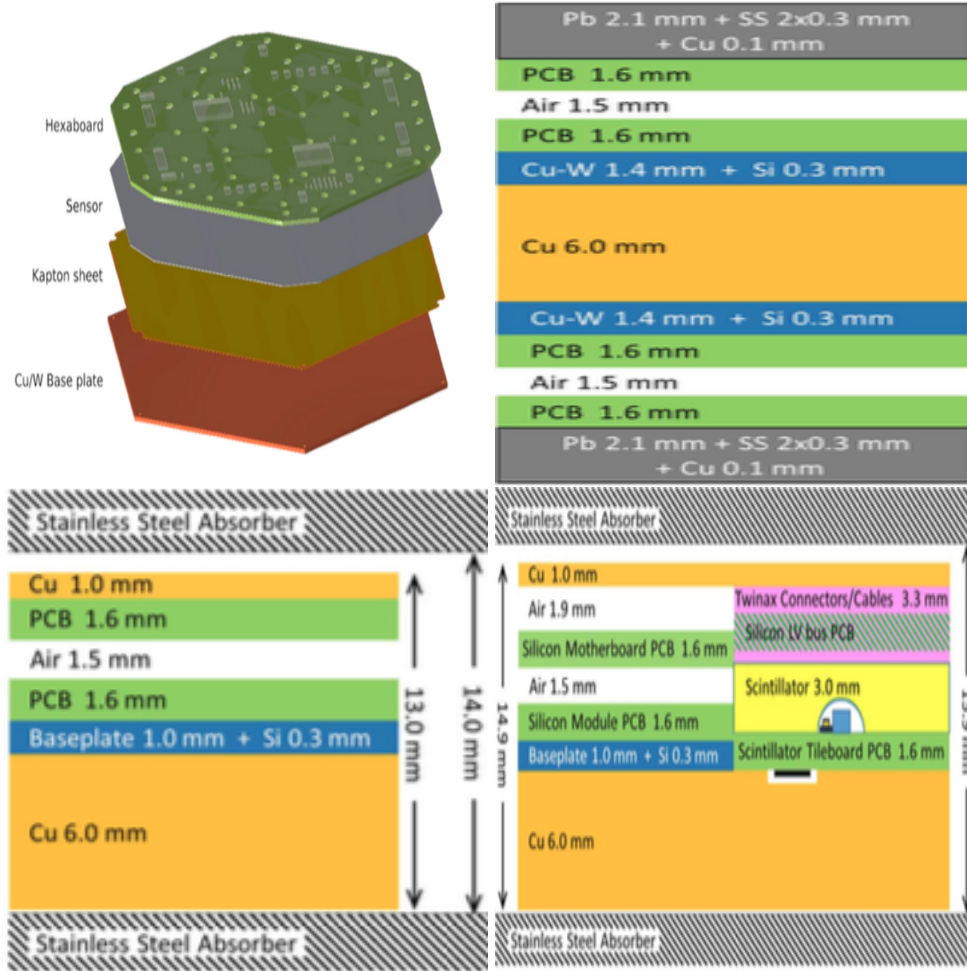


Figure 8: The figures show the different material arrangement in the different part of the detector. The top left shows the silicon module used in the CE-E and CE-H. The top right is the one sampling layer of the CE-E. The down left is the one sampling layer of the CE-H which sits on the high-radiation region, oppositely, the downright is the one sampling layer of the CE-H which is based on the place with low-radiation.

In the test-beam, we employed the material based on the prototype of HGCal written above, but for the detail of its, each component will use the CMS phase 2 design. I will write the results of the test-beam in 2018 in June and October.

### 3.2 Introduction of the optimized cuts and new variable with June test-beam MC

For the purpose of tagging the "electron-like"(e-like) pion, at the starting, we must quest the properties of the electron, including shower shape, energy deposition, etc.

in the detector. For the next step, we apply all of the properties of the electron in runs of pion to pick up those "bad pion". I will describe the detail for the cuts and the selections as follows. For notes, in this studies, we used the sample with 50GeV electron and pion MC which were done in June test-beam of 2018.

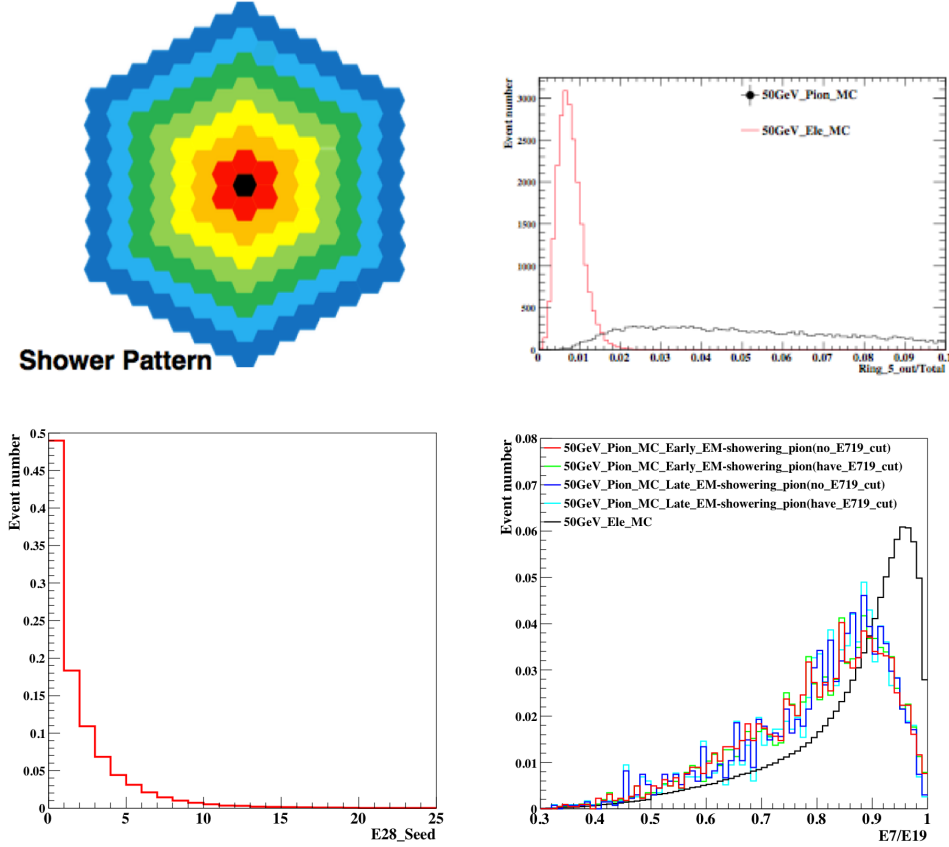


Figure 9: These figures present the cuts we used in the studies. Top left is the shower pattern on the transverse plane in the detector. Different rings mean the different sizes of the clustering on it. For example, the 5-Rings means the fifth circle calculated from the black pad to the shadow green. The top right plot is the energy distribution of  $\frac{\text{energy out of the 5 Rings}}{\text{total energy}}$ . The down left is the last layer energy with electrons. The down right is the cut with  $\frac{\text{energy in 2 Rings}}{\text{energy in 3 Rings}}$ , we plot every layer  $\frac{\text{energy in 2 Rings}}{\text{energy in 3 Rings}}$  value with and without the  $\frac{E_7}{E_{19}}$  cut

The first selection, because electron and pion have very different shower shape(width) on the transverse plane in ECAL, we studied on them at the starting point. From the theory, it tells us the electron has the narrower shower shape than pion because of the mechanism of interaction, we explored the cut with the different number of rings on the transverse plane of the detector. In the Fig.9 top left shows the distribution plot of the ring. Due to the plot which we draw in Fig.9 top right, it shows that the energy of electrons showering is collected out of the 5 Rings with at most 1% compared with total energy in ECAL, we employed it for being the cuts. Although, it is the tight cut because it also discarded some of the electrons as well, but we wanted to make sure that we cut-off the bad pion in reality. We called this cut as "5-Rings cut".

The second selection, because we wanted to pick up the "fully EM-showering" event, we must need to tag the events which deposit most of the energy in ECAL. Normally, we use HCAL to cut-off the events which leave much energy out of ECAL, but without having the information of HCAL energy in this month test-beam, we only could use the "poor man's solution"-"finding the last layer energy of electrons" and using it to be our cut. In the Fig.9 down left is the plot which was applied to search out the cut-off criteria in the last layer energy with the events which passed 5-Rings cut. Eventually, we used the last layer energy smaller than 20MIPs to tag the fully showering events, we called this cut as "The last layer energy cut".

The third selection, it is the cut always employed to tag the EM-showering in pion traditionally, and is related to the shower shape also, similar to the first selection. In every run, we applied the cut with " $\frac{\text{energy in two Rings}}{\text{energy in three Rings}} = \frac{E_7}{E_{19}}$ " need to be at least 0.75 at the shower depth maximum layer. Shower depth(unit: $X_0$ ) is the energy-weighted in the detector, it means the center-of-energy in the shower. In the Fig.9 downright shows the events with the values of  $\frac{E_7}{E_{19}}$  of all layers. Compare between with and without the cut, most of the events with cut have bigger value than without the cut, this means we cut-off some of the border shower shape events in pion runs. Both of Early EM-showering and Late EM-showering in the plot are two kinds of the e-like pion, will describe the detail later.

The fourth selection, because we applied the cuts with first to third selections, and we observed the shower depth of those passing events in the Fig.10, we found there existed two types of electron-like pion with the cut described in the next paragraph:

- Early EM-showering: The events have a similar shower depth with the electrons in ECAL, which means the energy distributions for these events are similar to the electrons. We can't distinguish them from electrons very well.
- Late EM-showering: The events have very distinct shower depth compared with the electrons in ECAL, owing to those events put most of the energy in the backward ECAL. As a result of it, we can use some variables to distinguish this kind of pion from electrons.

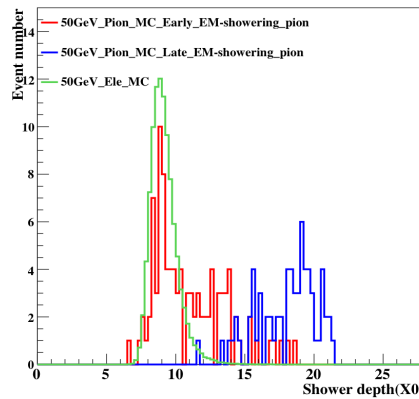


Figure 10: This figure shows the shower depth for both Early EM-showering and Late EM-showering.



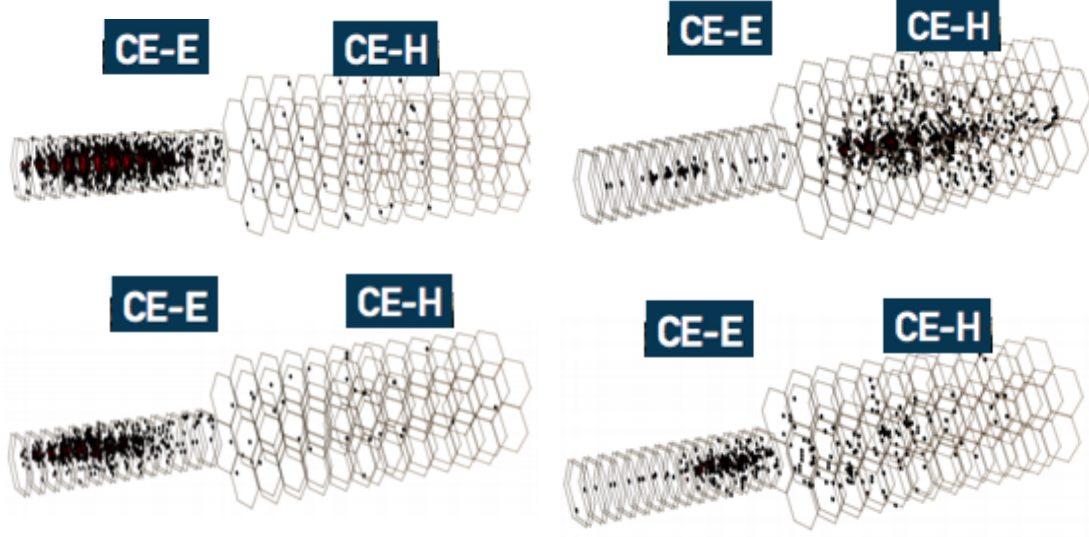


Figure 11: These figures show the event display with the cases which are the used to probe in the studies. (Top left) is the normal electron, most of the energy put in the CE-E, oppositely, (Top right) is the normal pion, which put most of the energy in CE-H. The both cases below the normal particles are the "e-like pion" cases, which we target to distinguish them from the normal electron. The left plot is the case for "Early EM-showering" and the right plot is the case for "Late EM-showering".

We present some events display for the normal electron, pion showering, Early EM-showering and Late EM-showering in Fig.11. For the fifth selection, we required a consecutive number of hits in CE-E starting from the first layer and checking until the 14 Layer. We recorded the first layer which is found with more than 3MIPs, and pre-selection of it is at least 2 consecutive layers with more than energy > 3MIPs individually. If the layer of start showering was from the first to fifth layer, we called it "Early EM-showering", on the other hand, if it started at sixth to later layers, we called it "Late EM-showering", and all of this caused by the various interaction point. For note, because this was the convenient way to let us see the dynamic of the shower in ECAL approximately, we used this cut to distinguish them. In the real case, we don't do this cut, just use one variable to distinguish "both" of them from electrons. I will show the results later.

Tight cut types	value
5-Rings cut	> 0.99
The last layer energy	< 20MIPs
$\frac{E_7}{E_{19}}$	> 0.75

Table 1: E10/Etotal Tight cuts selection used in June test-beam

So far, the one strong point of the difference is seen between electron and e-like pion is the longitudinal segmentation of shower energy deposited in the ECAL. Because some of them put energy in the different section of succeeding layers in ECAL, especially for comparing the electrons with Late EM-showering. In this case, we thought about the variables with "longitudinal segmentation" with ECAL layers. We developed the

discrimination variable as following:

$$E_{10} = \frac{E_{\text{first 10 layers energy}}}{E_{\text{total energy}}} \quad (1)$$

Intuitively, because the electrons put most of the energy in the front of the detector, they will have a bigger value for  $E_{10}$ . Oppositely, for the e-like pions, because they put all of the energy in the different sections due to the different interaction point, there exist many values for  $E_{10}$  from all events. In the Fig.12, we can see that we can use  $E_{10}$  to reject some of e-like pions out of the electron region.

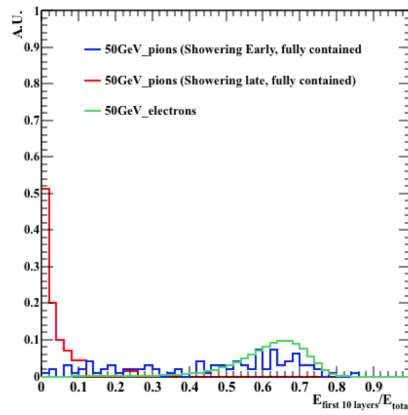


Figure 12: The  $E_{10}$  value for electrons, Early EM-showering and Late EM-showering.

In my study, we need to find the critical point of  $E_{10}$  to be the threshold, because we wanted to reserve the electrons as many as possible, and cut-off the bad pions for our best. We used the quantities about the "Electron Efficiency" and "Background Rejection" as our standard in the following formula, to see whether this cut is good:

$$\text{Electron Efficiency} = \frac{\text{Pass electrons}}{\text{Total electrons}} \quad (2)$$

$$\text{Background Rejection} = \frac{\text{Total e-like pions}}{\text{Pass e-like pions}} \quad (3)$$

Our goal was to cut-off the most e-like pion with the highest background rejection (pass the least e-like pions) in (2), but reserved the most electrons with the highest electron efficiency. (pass the most electrons) in (3).

First, we tried to use the critical point at 0.2, 0.3 and 0.4, and observed whether it could give us the best results. In the Table.2, it tells us that the background rejection is good (more than half of e-like pions are rejected), but we hoped that our electron efficiency can achieve to 99.8% or more. On the other hand, we still wanted to keep the electrons alive after the higher critical point to reserve most of the important researches on them.

In order to improve the issue we told before, we tried to use the other points near 0.2 and see whether we can get great results without discarding more electrons. In the Table.3, it said that near the point 0.2, it can reserve most of the electrons near 99.9%

E10/Etotal cut	Background Rejection	Electrons Efficiency
>0.2	2.15+/-0.18(stat)	99.85%+/-0.01%(stat)
>0.3	2.53+/-0.24(stat)	99.30%+/-0.01%(stat)
>0.4	2.74+/-0.28(stat)	97.32%+/-0.01%(stat)

Table 2: E10 cuts of Electrons Efficiency and Background Rejection comparison with point 0.2,0.3 and 0.4

E10/Etotal cut	Background Rejection	Electrons Efficiency
>0.16	2.02+/-0.15(stat)	99.93%+/-0.01%(stat)
>0.18	2.07+/-0.16(stat)	99.89%+/-0.01%(stat)
>0.20	2.15+/-0.17(stat)	99.85%+/-0.01%(stat)

Table 3: E10 cuts of Electrons Efficiency and Background Rejection comparison with point 0.16,0.18 and 0.20

and can get the background rejection with the expectation value near 2 (Half of the rejection).

For the conclusion of this month test-beam results, we found that Background Rejection can be near 2 without killing many electrons with 99.9% electron efficiency. For our expectation, we wanted to apply this cut in the October test-beam, to see whether we can see similar results.

### 3.3 The results of application with October test-beam MC

In this month test-beam, because we had the information of HCAL, we applied the cut no more than 0.4% energy fraction in the HCAL energy to replace the poor man's solution. And also, we altered some parameters of the cuts to get the best results. In the Table.4, the summary table for the cuts is shown.

Tight cut types	value
5-Rings cut	> 0.99
$\frac{\text{HCAL total energy}}{\text{ECAL total energy}}$	< 0.004
$\frac{E_7}{E_{19}}$	> 0.85

Table 4: Tight cuts selection used in October test-beam

In June test-beam, we only made the plot of  $E_{10}$  cut after the tight-cut selections. With this month data, we wanted to study the power of the cuts from tight-cut selections to  $E_{10}$  cut, it means we wanted to study the total number of how many e-like pions will be rejected after all cuts included in. We had two steps rejection, and we used the same definitions with the equation (2) and (3).

- Tight cuts rejection power:  
This is the first step rejection study. Because we need to find out the e-like pions with the cut in the Table.4 first, and then we can use our segmentation variable to reject them for the next step. For this step, we found that the number of background rejection is 3600. Most of the pions were rejected by those cuts.

- $E_{10}$  cuts rejection power:

This is the second step rejection study. After we found out the e-like pions, we applied the discrimination variable which we had invited in June test-beam,  $E_{10}$  cut, to distinguish them from the real electrons. In this step, we used the critical point at 0.18, 0.20 and 0.22, to see whether at the critical point near 0.2 will give us the good results. In the Table.5 summarize the results for them. Although this time didn't have better results compared with the June test-beam, we still got the number about 1.5.

E10/Etotal cut	Electron Efficiency	Background Rejection
$> 0.18$	$99.56\% \pm 0.01\%(\text{syst}) \pm 0.20\%(\text{stat})$	$1.52 \pm 0.27(\text{stat})$
$> 0.20$	$99.05\% \pm 0.01\%(\text{syst}) \pm 0.20\%(\text{stat})$	$1.53 \pm 0.26(\text{stat})$
$> 0.22$	$98.91\% \pm 0.01\%(\text{syst}) \pm 0.20\%(\text{stat})$	$1.56 \pm 0.24(\text{stat})$

Table 5:  $E_{10}$  cuts electrons efficiency and Background rejection after the tight cuts

In the end, totally we can get the total rejection power of cuts which was calculated by the number 3600 with the Tight cuts rejection power and the number 1.5 with the  $E_{10}$  cuts rejection power. After times them together, we can get the number  $3600 \times 1.5 = 5400$ . In A Toroidal LHC ApparatuS(ATLAS), the expected value for it is between 5000 to 10000 [27], and in our case, we arrive the value for the expectation.

### 3.4 Conclusion

In this section, we obtained the cuts and new variable  $E_{10}$  to distinguish the e-like pions from real electrons with the June test-beam MC samples, and we applied in the October test-beam MC samples, excitedly, we got the expected results compared with the previous studies in ATLAS. Those cuts and variable we used in this study have never been applied in CMS analysis before, and we expect that we can help to reduce the background which is sensitive to electron/pion identification with HGCal in the future.

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

After the discovery of the standard-model-like Higgs boson at 2012[1], we have established a milestone in HEP field. We used the LHC at CERN to solve an important puzzle in the SM of particle physics, Higgs boson, and it gave us the answer why the elementary particles retain their mass[28] from the early universe to now. We expect more particles predicted by BSM can be found in the future. A lot of these new particles may be heavier than 13 TeV[29][30], which is operating now in LHC. Therefore, particle physicists plan to build a very high energy collider with  $\sqrt{s} = 100$  TeV in the future, such as Future circular  $pp$  colliders [31] such as the European initiatives, FCC-ee [32], FCC-hh [33], high-energy LHC (HE-LHC) [34], and the Chinese initiative, SppC [16].

In the very-high-energy era, the C.M. energy will be increased compared with now in LHC, and the instantaneous luminosity will be also expected to increase. In this

case, more interesting events will occur within a shorter amount of time, but incidentally, it will accompany with the unexpected background. We need to figure out some methods to pick out the signals which we are interested in, and filter the unwanted huge background events extremely. The silicon future circular collider (SiFCC) detector, which is based on the silicon detector (SiD)[35] designed for international linear collider (ILC)[36][37], was our simulated detector prepared for the high energy collider for the future era in Ref.[38]. We used this detector to simulate the condition under very high C.M., and applied the different configurations of detector with some methods such as jet substructure variables, to see the detector performance of distinguishing the signal from background. I will describe the detail later.

Traditionally, we used scintillators as the active layers of the ECAL and HCAL to measure the energy of the showering of particles which are induced by the absorbers (with material such as lead, brass, or tungsten, etc.). But when the radiation dosage is much higher than before, scintillators can deteriorate easily. This is the same circumstance which CMS is bumped into now. In addition, because of time consuming for making the special shape for fitting the detector, it is not easy to sub-divide scintillators into smaller areas. And also, we need to have a readout with high position resolution to match them. In this case, we want to use the new design- "high granularity" detector to solve the resolution problems, i.e., we used the different configurations of detector to record the precise position and energies. For the radiation problem, because the silicon detector was known to be radiation tolerant.[39], we could use it in the future collider, same as HGCAL which was presented in my previous studies.

I did this study with Prof. Shin-Shan Eiko Yu from NCU, Prof. Ashutosh Kotwal and student Sourvan Sen from Duke University, Dr. Sergei Chekanov and Dr. Proudfoot James from Argonne National Laboratory(ANL), and Nhan Viet Tran from Fermi National Accelerator Laboratory. The contributions of Chih-Hsiang Yeh to this study includes the following:

- Study of detector performance with soft drop mass.
- Studies of signal and background separation using the different jet substructure variables.

I will describe the detail as following.

## 4.1 Simulation of detector response and event reconstruction

The description of the detector and software used for our study is discussed in [38]. We used the geometry of SiFCC detector with the software package that represents a various environment for simulations of detector performance. It was used to prepare for new technology options, event reconstruction techniques for future 100 TeV colliders. Figure 13 is the SiFCC detector simulated by the GEANT4 package[40] and it was the detector studied in Ref.[38]. The baseline detector discussed in [38] uses a steel-scintillator hadronic calorimeter with a transverse cell size of  $5 \times 5 \text{ cm}^2$ , which corresponds to  $\Delta\eta \times \Delta\phi = 0.022 \times 0.022$ . In addition, because we wanted to test the detector

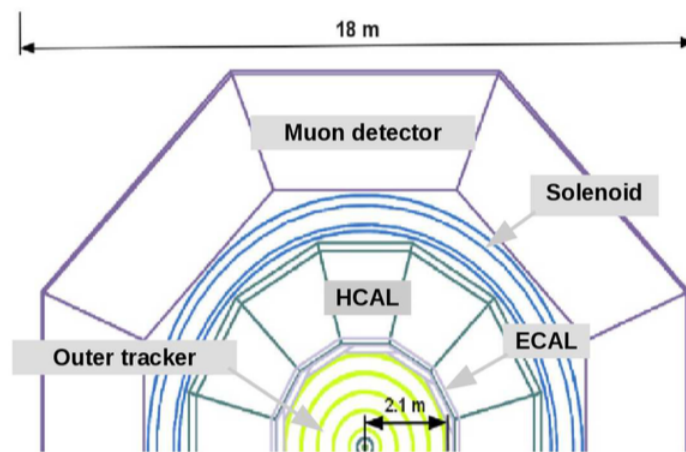


Figure 13: The azimuthal cross section of the SiFCC detector.

Barrel	Technology	pitch/cell	radii (cm)	z  size (cm)
Vertex detector	silicon pixels/5 layers	25 $\mu\text{m}$	1.3 - 6.3	38
Outer tracker	silicon strips/5 layers	50 $\mu\text{m}$	39 - 209	921
ECAL	silicon pixels+Tungsten	2 $\times$ 2 cm	210 - 230	976
HCAL	scintillator+steel	5 $\times$ 5 cm	230 - 470	980
Solenoid	5 T (inner), -0.6 T (outer)	-	480 - 560	976
Muon detector	RPC+steel	3 $\times$ 3 cm	570 - 903	1400

Figure 14: The barrel structure in the SiFCC, including sizes and materials.

performance of granularity effect with hadronic processes of signal and background, several geometry variations were considered, such as  $20 \times 20 \text{ cm}^2$  and  $1 \times 1 \text{ cm}^2$ , which correspond to  $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$  and  $0.0043 \times 0.0043$ , respectively.

SiFCC detector which was simulated in our study contains five main parts: tracker, ECAL, HCAL, solenoid, and muon chambers. When particles are produced in the collider, they will go through the detector and leave signals to be recorded. First, the tracker can record the trajectory of a charged particle because it can leave the energy in and bended by the solenoid, we can obtain the transverse momentum of this charged particle. Electrons and photons will leave energy in ECAL, and hadronic particles, such as neutrons, protons, etc., will deposit energy in HCAL mainly. Muons will go through the whole detector and could be detected by muon chamber. Figure 14 lists the material and the configuration for the sub-detectors in SiFCC. Note, since for the time being we have more experimental results coming from the scintillator-based HCAL in CMS and ATLAS, we have been using this type of HCAL as a first study. After finishing the first study, we will move to the study of HCAL with silicon sensors as active layers later, same as HGCAL material. In our studies, we used the simulations of a heavy  $Z'$  boson, a hypothetical gauge boson that arises from extensions of the electroweak symmetry of the SM. The  $Z'$  bosons were simulated with the masses,  $M = 5, 10, 20$

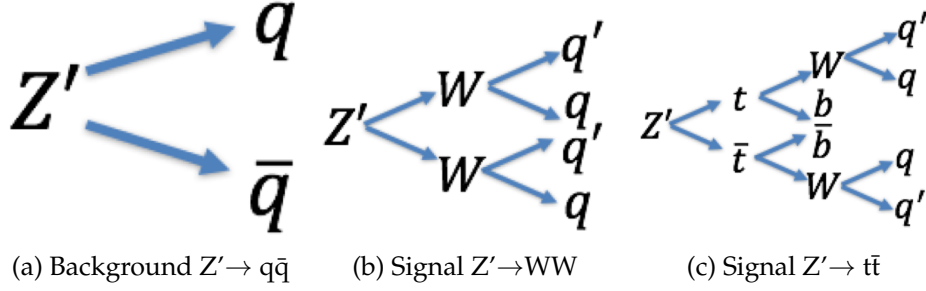


Figure 15: The signals and background we study.

and 40 TeV. The lowest value 5 TeV represents a mass that is the achieved C.M energy of the LHC experiments. The value 40 TeV represents the physics for 100 TeV colliders. The  $Z'$  particles are forced to decay to two light-flavor jets ( $q\bar{q}$ ) [41] with one-prong jet as background,  $W^+W^-$  [42] or  $t\bar{t}$  [43] as signal, where  $W(\rightarrow q\bar{q})$  with two-prong jets, and  $t(\rightarrow W^+ b \rightarrow q\bar{q}b)$  with three-prong jets decay hadronically. The processes for signal and background are shown in Fig.15. In all such scenarios, two highly boosted jets were produced, which were typically back-to-back in the laboratory frame. The main difference between considered decay types lays in different jet substructure. The events of signals were generated using the generator with the default settings, ignoring interference with SM processes. The event samples which were used in this study are available from the HepSim database [44] now.

## 4.2 Study of detector performance with soft drop mass

In this section, we used the jet mass which was computed with soft drop declustering, and goal was to study the performance of detector with three types of detector cell sizes and C.M. energies.

### 4.2.1 The technique of soft drop declustering

The soft drop declustering [45] is a grooming method which is used to remove the soft wide-angle radiation from a jet. The particles of a jet  $j_0$  are first reclustered by using the Cambridge-Aachen (C/A) algorithm [46, 47]. Second, the jet  $j_0$  is broken into two subjets  $j_1$  and  $j_2$  by returning back to the last stage of C/A clustering. If the subjets pass the soft drop condition in the formula, jet  $j_0$  will be the final soft-drop jet. Otherwise, the algorithm will replace the  $j_0$  with the subjet which has the larger  $p_T$  (among  $j_1$  and  $j_2$ ). Then, iterate the procedure until all of the jet are the final soft-drop jets.

$$\frac{\min(p_{T1}, p_{T2})}{p_{T1} + p_{T2}} > z_{\text{cut}} \left( \frac{\Delta R_{12}}{R_0} \right)^\beta, \quad (4)$$

where  $p_{T1}$  and  $p_{T2}$  are the transverse momenta of the two subjets,  $z_{\text{cut}}$  is soft drop threshold,  $\Delta R_{12}$  is the distance between the two subjets in the  $\eta$ - $\phi$  plane,  $R_0$  is the characteristic radius of the original jet, and  $\beta$  is the angular exponent.

In our study, we compared the performance of future detector when setting  $\beta = 0$  versus when setting  $\beta = 2$ . For  $\beta = 0$  [48, 49], the soft drop condition depends only on the

$z_{\text{cut}}$  and is angle-independent. For  $\beta = 2$  [50], the condition is angle-dependent, which depends on the angular distance between the two subjects and  $z_{\text{cut}}$  and the algorithm becomes infrared and collinear safe. Both of them have different sensitivities to large angle radiation.

#### 4.2.2 Analysis method

We applied the following method to quantify the detector performance and found out the cell size that could give us the best separation power to distinguish signal from background. For each configuration of detector and C.M. energy, we drew the receiver operating characteristic (ROC) curves in which the x-axis is the signal efficiency ( $\epsilon_{\text{sig}}$ ) and y-axis was is inverse of background efficiency ( $1/\epsilon_{\text{bkg}}$ ).

For scanning the efficiencies of soft drop mass cuts, we used many mass window that is described as follows. We first look for the median bin of the soft drop mass histogram in the simulated signal events. Taking the right boundary of bin  $i_{\text{med}}$  as the center of mass window  $x_{\text{center}}$ . Second, we start increasing the width of mass window symmetrically on the left and right of  $x_{\text{center}}$  i.e. the narrowest mass window is about 10 GeV (Left and right are 5 GeV width). If one side reaches the boundary of the mass histogram, for the rest part of the window, we only expand the width on the other side, also in steps of 5 GeV. For each mass window, there would be corresponding  $\epsilon_{\text{sig}}$  and  $\epsilon_{\text{bkg}}$ , which give a point in the ROC curves.

#### 4.2.3 Results and conclusion

Figures 16, 18, 20, and 22 show 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  $Z' \rightarrow WW$  of two-prong jets and  $Z' \rightarrow t\bar{t}$  of three-prong jets. In Figs. 17, 19, 21, and 23, ROC curves from different detector cell sizes are compared for each C.M. energy, respectively.

Figures 17 and 19 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 when the signal is  $Z' \rightarrow WW$  and at  $\sqrt{s} = 10$  and when the signal is  $Z' \rightarrow t\bar{t}$ . On the contrary, Figs. 21 and 23 show that for  $\beta = 2$  the smallest detector cell size does not have any improvement in the separation power correspond to those 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.

Note that the separation between ROC curves depends on the physics variable and on the boost of the top quarks or the  $W$  bosons. For example, the similarity between the ROC curves shown in Fig. 19(a) is due to the insufficient boost of the top quarks. On the other hand, Fig. 19(d) does not show a difference between the ROC curves because the boost is too high. For both of the cases, they indicate the situation that for the different detector cell sizes, if they are not small or big enough to distinguish the different angle of fatjets, it can lead to producing the same separation power for all cell sizes.

We also found that when we compared between  $\beta = 0$  and  $\beta = 2$ ,  $\beta = 0$  had better performance for distinguishing signal from background. Therefore, we will employ



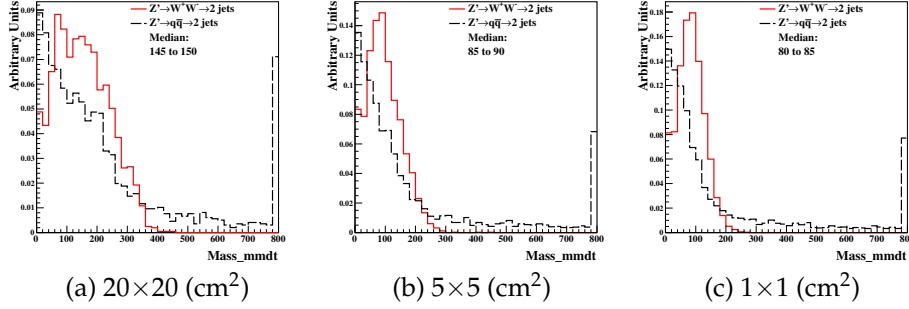


Figure 16: Distributions of soft drop mass for  $\beta=0$ , with 20 TeV C.M. energies and three different detector cell sizes:  $20 \times 20$ ,  $5 \times 5$ , and  $1 \times 1$  ( $\text{cm}^2$ ). The signal (background) process is  $Z' \rightarrow WW$  ( $Z' \rightarrow q\bar{q}$ ).

requirements on this variable when studying the other jet substructure variables in the next studies.

### 4.3 Studies of detector performance with jet substructure variables- N-subjettiness

In this section, we studied the one kind of jet substructure variable-N-subjettiness and compared their ability to separate signal from background with different detector sizes and C.M. energy by using ROC curves.

#### 4.3.1 The technic of N-subjettiness

The formula and the technique are as following:

$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min\{\Delta R_{1,k}, \Delta R_{2,k}, \dots, \Delta R_{N,k}\} \quad (5)$$

$$d_0 = \sum_k p_{T,k} R_0 \quad (6)$$

$k$  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.

To use this technique, first, Anti-kt(AK) algorithm [51] with radius( $R_0=0.4$  in formula 6) is used to reconstruct jets in the detector. Second, after reconstructing the jets, exclusive  $k_T$  algorithm [52] is used to find the jet axis in a fatjet, and applied in calculating the radius from constituent particles to jet axis. Third, start applying the formula (5), looping all constituent particles in a fatjet. Finally, when finishing running all particles, it will give out  $\tau_N$ , where  $N$  is positive integer.

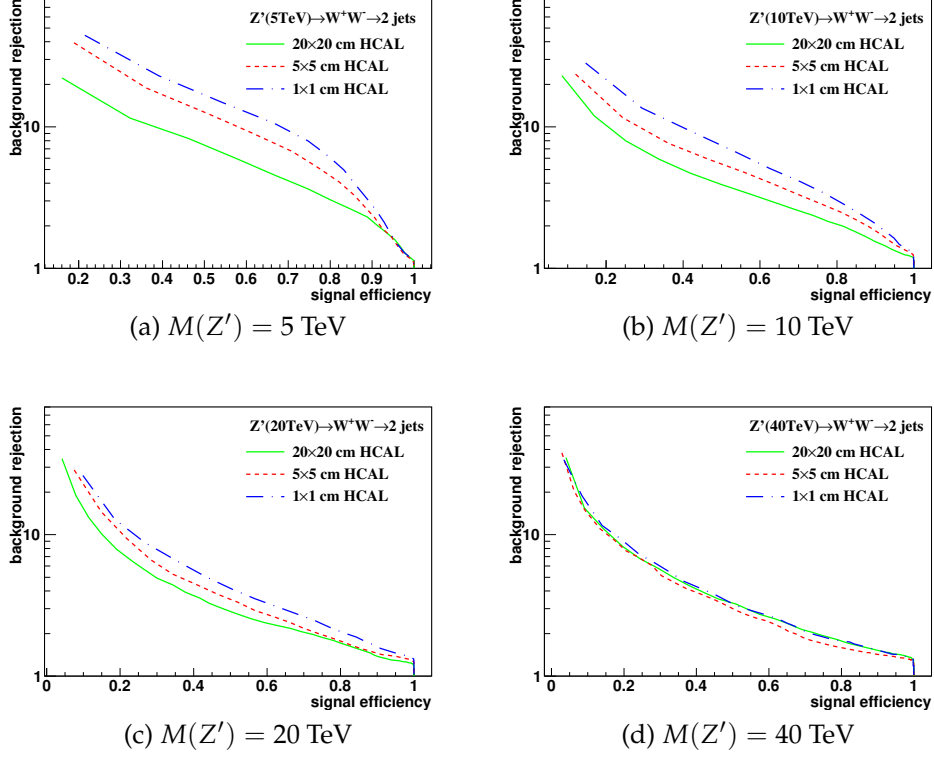


Figure 17: The ROC curves of soft drop mass selection for  $\beta=0$  with 5, 10, 20, and 40 TeV C.M. energies. Three different detector cell sizes are compared:  $20 \times 20$ ,  $5 \times 5$ , and  $1 \times 1$  ( $\text{cm}^2$ ). The signal (background) process is  $Z' \rightarrow WW$  ( $Z' \rightarrow q\bar{q}$ ).

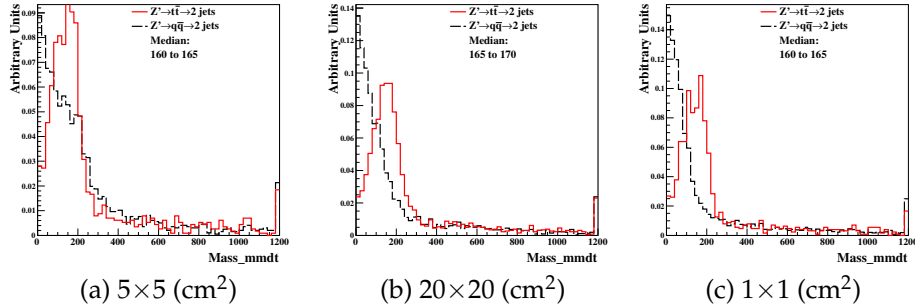


Figure 18: Distributions of soft drop mass for  $\beta=0$ , with 20 TeV C.M. energies and three different detector cell sizes:  $20 \times 20$ ,  $5 \times 5$ , and  $1 \times 1$  ( $\text{cm}^2$ ). The signal (background) process is  $Z' \rightarrow t\bar{t}$  ( $Z' \rightarrow q\bar{q}$ ).

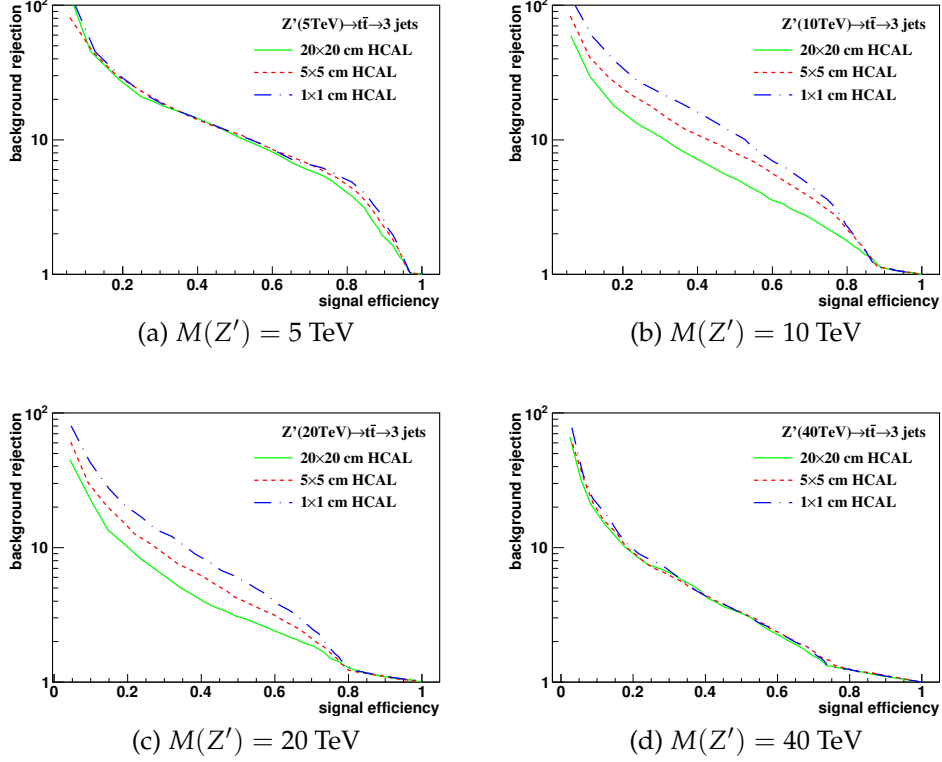


Figure 19: The ROC curves of soft drop mass selection for  $\beta=0$  with 5, 10, 20, and 40 TeV C.M. energies. Three different detector cell sizes are compared:  $20 \times 20$ ,  $5 \times 5$ , and  $1 \times 1$  ( $\text{cm}^2$ ). The signal (background) process is  $Z' \rightarrow t\bar{t}$  ( $Z' \rightarrow q\bar{q}$ ).

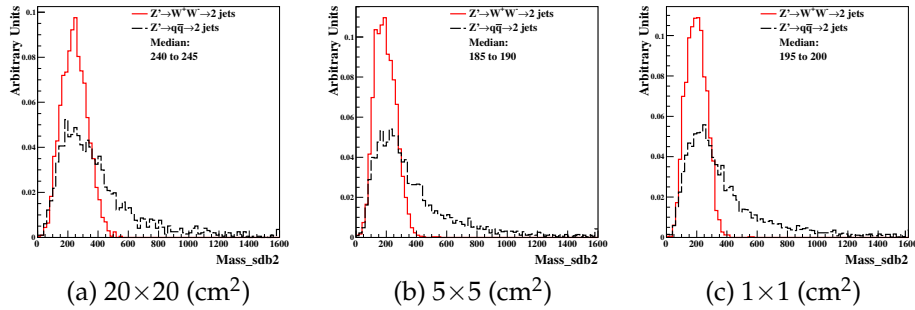


Figure 20: Distributions of soft drop mass for  $\beta=2$ , with 20 TeV C.M. energies and three different detector cell sizes:  $20 \times 20$ ,  $5 \times 5$ , and  $1 \times 1$  ( $\text{cm}^2$ ). The signal (background) process is  $Z' \rightarrow WW$  ( $Z' \rightarrow q\bar{q}$ ).

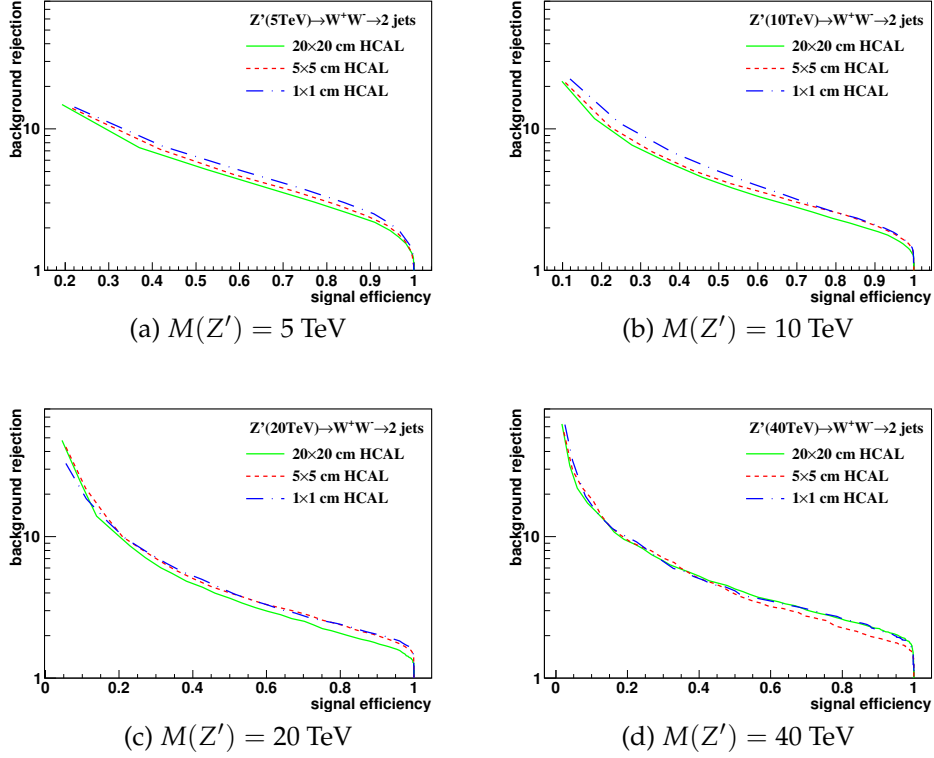


Figure 21: The ROC curves of soft drop mass selection for  $\beta=2$  with 5, 10, 20, and 40 TeV C.M. energies. Three different detector cell sizes are compared:  $20 \times 20$ ,  $5 \times 5$ , and  $1 \times 1$  ( $\text{cm}^2$ ). The signal (background) process is  $Z' \rightarrow WW$  ( $Z' \rightarrow q\bar{q}$ ).

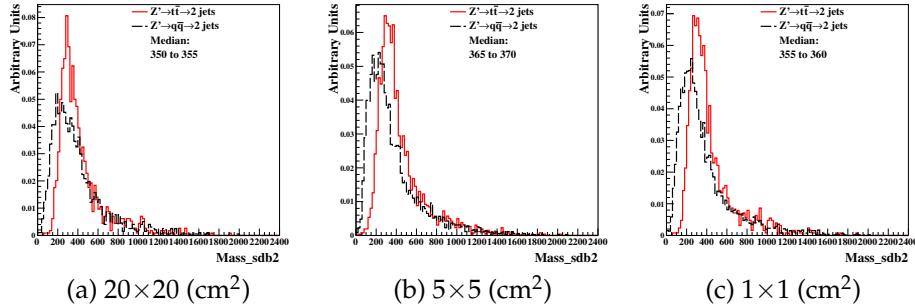


Figure 22: Distributions of soft drop mass for  $\beta=2$ , with 20 TeV C.M. energies and three different detector cell sizes:  $20 \times 20$ ,  $5 \times 5$ , and  $1 \times 1$  ( $\text{cm}^2$ ). The signal (background) process is  $Z' \rightarrow t\bar{t}$  ( $Z' \rightarrow q\bar{q}$ ).

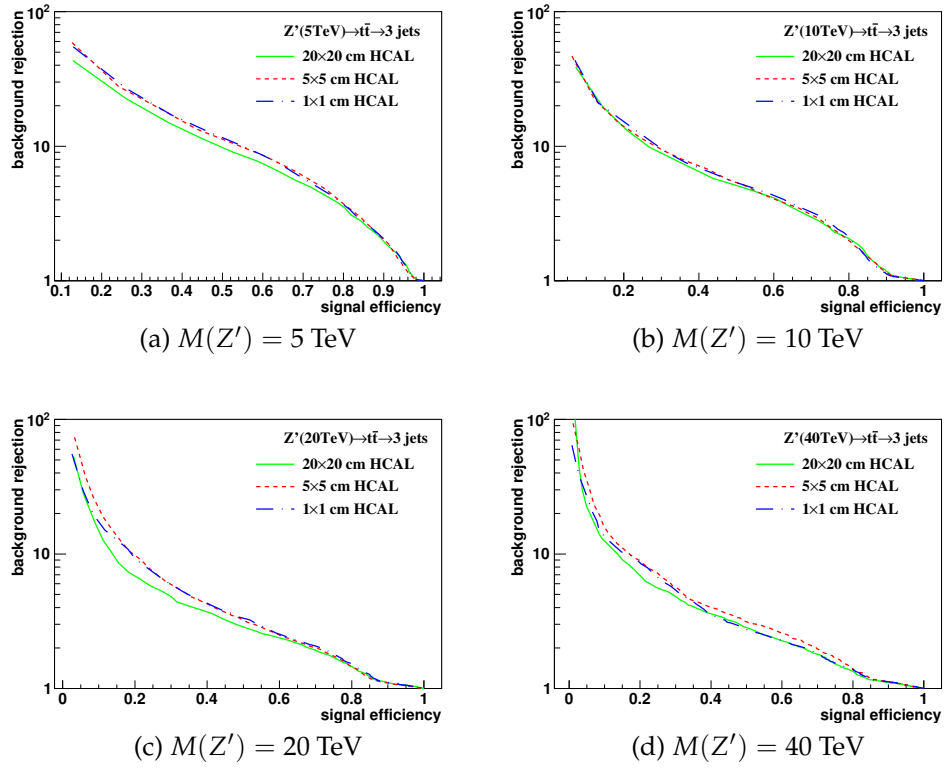


Figure 23: The ROC curves of soft drop mass selection for  $\beta=2$  with 5, 10, 20, and 40 TeV C.M. energies. Three different detector cell sizes are compared:  $20 \times 20$ ,  $5 \times 5$ , and  $1 \times 1$  ( $\text{cm}^2$ ). The signal (background) process is  $Z' \rightarrow t\bar{t}$  ( $Z' \rightarrow q\bar{q}$ ).

If a fatjet has exactly  $N$  subjet(s), its  $\tau_N$  is smaller than the  $\tau_{N-1}$ . It can return the truth that how many subjet(s) there is(are) in a fatjet. In our study, we used  $\tau_{21}=\tau_2/\tau_1$  and  $\tau_{32}=\tau_3/\tau_2$  in distinguishing two-subjets fatjet and three-subjets fatjet of signal processes from one-subjet fatjet of background process individually.

#### 4.3.2 Analysis method

Similar with the soft drop mass analysis, we need to apply some methods to quantify the separation power for different jet substructure variables with different configurations of detector, and figure out the best configuration for distinguishing signal from background. We still applied the ROC curves to be our comparison tools in this study.

Before scanning the efficiencies, we first did some pre-selection tasks. Inspired by the paper [53], it pointed out that applying the mass cut would give us the best separation power, so first we selected the optimized mass window. We did the prerequestion that in the signal events, compared the adjacent bins from the largest number of events iteratively. Until We included the 75% mass events, we stopped and used the events which were included.

Second, inspired by the Pearson lemma, it told us that if we used the ratio of the two histograms to be the window selection for the ROC curves, it can give us the best solution, so we applied this method in our study. We used  $\frac{\text{Signal histogram}}{\text{Background histogram}}$  of ratio histogram to be our window to draw the ROC curves. Compare the adjacent bins from the bin of most event in ratio histogram, and extended the width with higher side until all bins were included in.

#### 4.3.3 Results and conclusion

In the figure 26, 28, they show the histograms of  $\tau_{21}$  and  $\tau_{32}$  with  $\sqrt{s} = 20$  TeV after selecting the events as an example. As a result of figure 27, 29, they perform the ROC curves of  $\tau_{21}$  and  $\tau_{32}$  with different detector cell sizes and C.M. energy. The smallest detector cell size( $1 \times 1$ ) 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 ( $5 \times 5$  and  $20 \times 20$ ).

### 4.4 Studies of detector performance with jet substructure variables- Energy correlation function

Energy correlation function (ECF) [54] 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.

#### 4.4.1 The technic of energy correlation function

The basic ECF formula is as following:

$$ECF(N, \beta) = \sum_{i_1 < i_2 < \dots < i_N \in J} \left( \prod_{a=1}^N P_{ia} \right) \left( \prod_{b=1}^{N-1} \prod_{c=b+1}^N R_{i_b i_c} \right)^\beta \quad (7)$$

In the formula 7, the sum loop all particles in the jet  $J$  by using the information of  $P_{ia}$ , which are the momenta of all particles, and  $R_{i_b i_c}$ , which are the angles between each particles in the  $y$ - $\phi$  plane. in order to use the dimensionless observation to determine whether the number of subjets in system, parameter  $\tau_N$  is defined as:

$$\tau_N^{(\beta)} \equiv \frac{ECF(N+1, \beta)}{ECF(N, \beta)} \quad (8)$$

The idea of formula (8) is from N-subjetness, because the behavior of it is very similar to N-subjetness as reference [54]. 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 use  $\tau_{21} = \tau_2 / \tau_1$  and  $\tau_{32} = \tau_3 / \tau_2$  to distinguish two-subjets fatjet and three-subjets fatjet from one-subjet fatjet in the previous studies, it is also defined as the ratio of  $\tau$  there, which is called "Double ratio-ECF" that is used in our study:

$$C_N^{(\beta)} \equiv \frac{\tau_N^{(\beta)}}{\tau_{N-1}^{(\beta)}} = \frac{ECF(N-1, \beta) ECF(N+1, \beta)}{ECF(N, \beta)^2} \quad (9)$$

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

For the analysis method, it is the same as the part 4.3.2.

#### 4.4.2 The results and conclusion

In the figure [24], they show the histograms of  $C_2^1$  with  $\sqrt{s} = 20$  TeV after selecting the events as an example. For the figure [25], they present the ROC curves of  $C_2^1$  with different detector cell sizes and C.M. energy. The smallest detector cell size ( $1 \times 1$ ) doesn't have the best separation power to distinguish signal from background. In addition, in some cases such like (a), the biggest one ( $20 \times 20$ ) has the best distinguish power under the same C.M. energy.

### 4.5 Conclusion

The studies presented in this previous sections show that the reconstruction of jet substructure variables for future colliders will benefit from small cell sizes of the hadronic calorimeters. This conclusion is obtained using the simulation of calorimeter response combined with reconstruction of calorimeter clusters and the Rawhits information used as inputs for jet reconstruction. Hadronic calorimeters that used the cell sizes

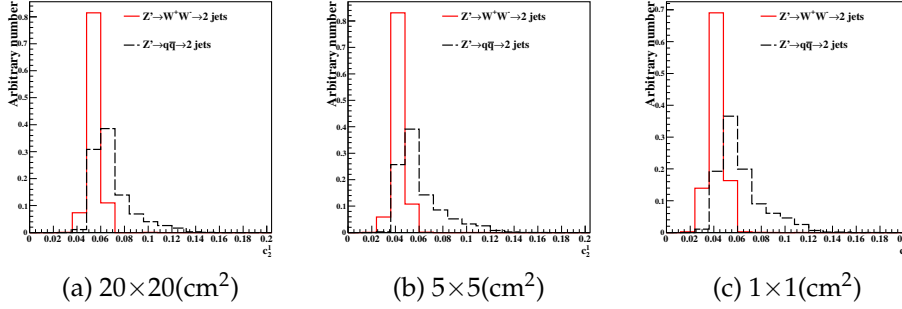


Figure 24: Distributions of  $C_2^1$  with 20 TeV C.M. energy and three different detector cell sizes:  $20 \times 20$ ,  $5 \times 5$ , and  $1 \times 1$  ( $\text{cm}^2$ ). The signal (background) process is  $Z' \rightarrow WW$  ( $Z' \rightarrow q\bar{q}$ )

of  $20 \times 20 \text{ cm}^2$  ( $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ ) have the worst performance for almost every substructure variable which is considered in our analysis, for the resonance mass at  $\sqrt{s} = 5 \text{ TeV}$  to  $\sqrt{s} = 20 \text{ TeV}$ . Such cell sizes are similar to those used for the ATLAS and CMS detectors at the LHC. In our studies, we obtained the performance of a hadronic calorimeter with  $\Delta\eta \times \Delta\phi = 0.022 \times 0.022$  ( $5 \times 5 \text{ cm}^2$  cell size) is, in most cases, better than for a detector with  $0.087 \times 0.087$  cells.

Thus this study confirms the HCAL geometry of the SiFCC detector [38], with the  $\Delta\eta \times \Delta\phi = 0.022 \times 0.022$  HCAL cells. It also confirms the HCAL design of the baseline FCC-hh [55, 56] detector with  $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$  HCAL cells.

## 5 Conclusion for all studies

I have done the "detector traveling" from now to the future within the project time. For now, at the start, I studied the cross-talk on the no sensor-based Hexaboard which will be used in HGAL. We explored the phenomenon and saw the correlated cross-talk and the anti-correlated cross-talk. Later, we did the project with the HGAL test-beam results of MC samples to do the e-like pions tagging and rejected them by the cuts and variable we developed. We saw the expected results in the October test-beam MC samples. Last but not least, we came to the future project with the SiFCC detector. By using the different configurations of the detector with the different jet substructure variables and soft-drop mass, to see whether the high granularity could give us the best results to distinguish the signal from background in the SiFCC detector. For our surprise, the smallest configuration is not the best, and we think that it need to be optimized.

## 6 Acknowledgement

First, thanks for my supervisors, Prof. Shin-Shan Eiko Yu from NCU, Prof. Ashtusoh from Duke University, Dr. Sergei from Argon National Lab(ANL) , Dr. Nhan from Fermilab, and Prof. Stathes Paganis, Prof. Rong-Shayang Lu from NTU, without them,



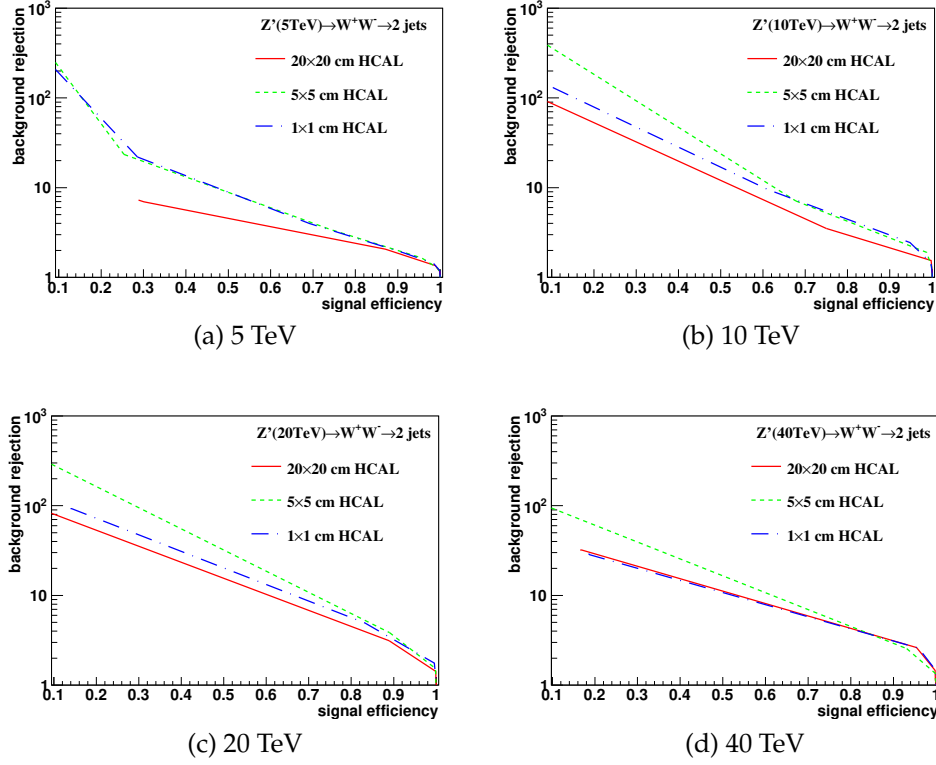


Figure 25: Signal efficiency versus background rejection rate using c2b1. The energies of collision at (a)5, (b)10, (c)20, (d)40 TeV are shown here. In each picture, the three ROC curves correspond to different detector sizes.

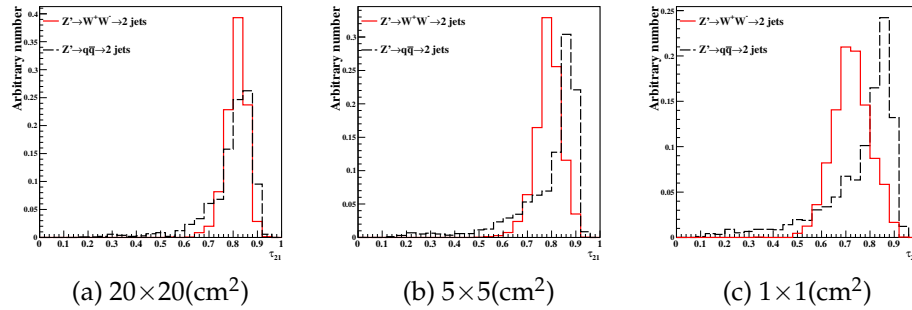


Figure 26: Distributions of Distributions of  $\tau_{21}$  with 20 TeV C.M. energy and three different detector cell sizes:  $20 \times 20$ ,  $5 \times 5$ , and  $1 \times 1$  ( $\text{cm}^2$ ). The signal (background) process is  $Z' \rightarrow WW$  ( $Z' \rightarrow q\bar{q}$ )

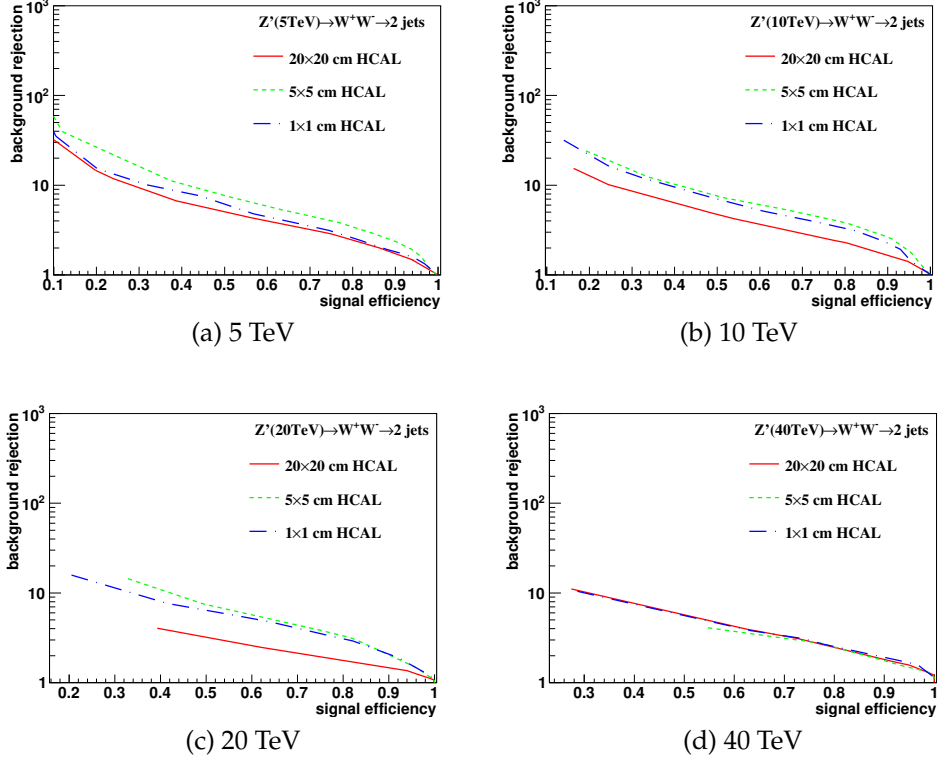


Figure 27: Signal efficiency versus background rejection rate using  $\tau_{21}$ . The energies of collision at (a)5, (b)10, (c)20, (d)40TeV are shown here. In each picture, the three ROC curves correspond to different detector sizes.

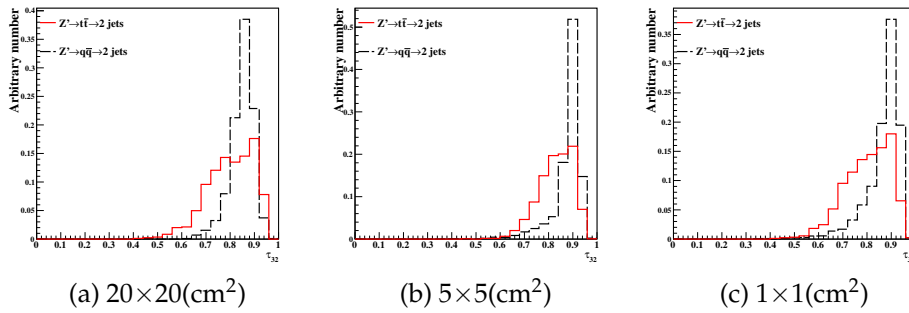


Figure 28: Distributions of  $\tau_{32}$  with 20 TeV C.M. energy and three different detector cell sizes: 20x20, 5x5, and 1x1 (cm<sup>2</sup>). The signal (background) process is  $Z' \rightarrow t\bar{t}$  ( $Z' \rightarrow q\bar{q}$ )

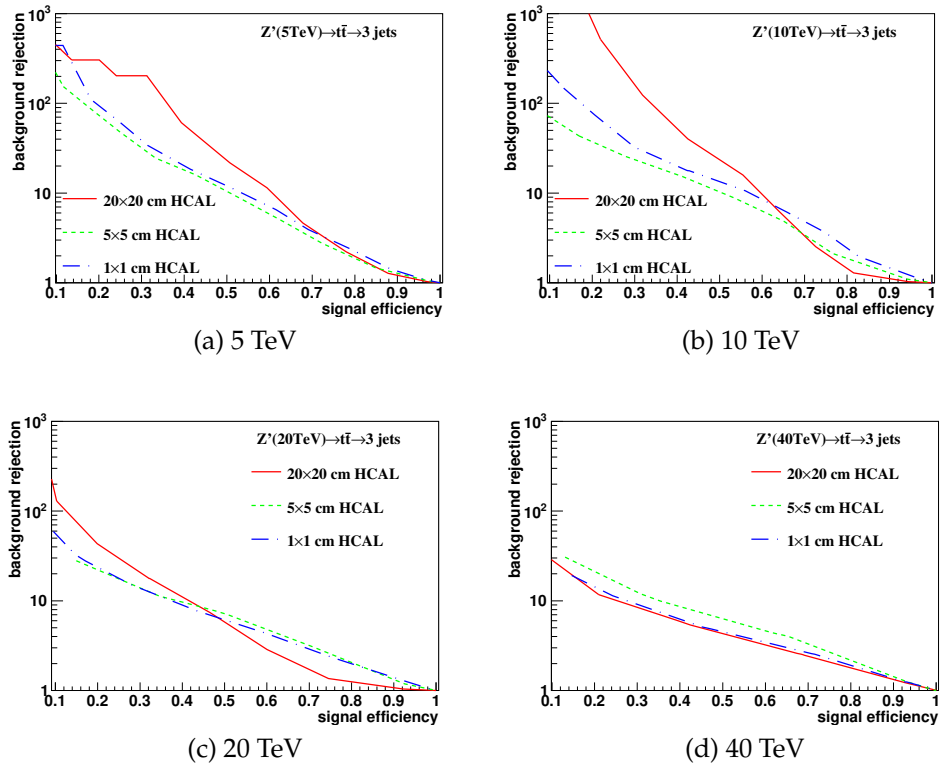


Figure 29: Signal efficiency versus background rejection rate using  $\tau_{32}$ . The energies of collision at (a)5, (b)10, (c)20, (d)40TeV are shown here. In each picture, the three ROC curves correspond to different detector sizes.

I can't do those good studies smoothly, and publish the results well. Also, I would like to thanks LHC-CMS, NCU and NTU to give me the resource, including the data, hardware and the server resource, to study my interested topics. And thanks for all people in NCUHEP, NTUHEP groups who have helped me. Last but not least, I'm glad to thanks Ministry of Science and Technology(MOST) to support me to do this studies.

## References

- [1] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett. B* **716** (2012) 30–61, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.
- [2] ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Phys. Lett. B* **716** (2012) 1–29, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.
- [3] CMS Collaboration, "Search for physics beyond the standard model in high-mass diphoton events from proton-proton collisions at  $\sqrt{s} = 13$  TeV", *Phys. Rev. D* **98** (2018), no. 9, 092001, doi:10.1103/PhysRevD.98.092001, arXiv:1809.00327.
- [4] D. Abercrombie et al., "Dark Matter Benchmark Models for Early LHC Run-2 Searches: Report of the ATLAS/CMS Dark Matter Forum", arXiv:1507.00966.
- [5] CMS Collaboration, M. Pelliccioni, "CMS High mass WW and ZZ Higgs search with the complete LHC Run1 statistics", in *Proceedings, 50th Rencontres de Moriond Electroweak Interactions and Unified Theories: La Thuile, Italy, March 14-21, 2015*, pp. 47–52. 2015. arXiv:1505.03831.
- [6] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.
- [7] ATLAS Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider", *JINST* **3** (2008) S08003, doi:10.1088/1748-0221/3/08/S08003.
- [8] G. Apollinari, O. Brning, T. Nakamoto, and L. Rossi, "High Luminosity Large Hadron Collider HL-LHC", *CERN Yellow Report* (2015), no. 5, 1–19, doi:10.5170/CERN-2015-005.1, arXiv:1705.08830.
- [9] L. Medina Medrano, G. Arduini, and R. Toms, "Studies on Luminous Region, Pile-up and Performance for HL-LHC Scenarios",.
- [10] C. Collaboration, "The Phase-2 Upgrade of the CMS Endcap Calorimeter", Technical Report CERN-LHCC-2017-023. CMS-TDR-019, CERN, Geneva, Nov, 2017. Technical Design Report of the endcap calorimeter for the Phase-2 upgrade of the CMS experiment, in view of the HL-LHC run.

- [11] CMS Collaboration, “Search for heavy resonances decaying into a vector boson and a Higgs boson in final states with charged leptons, neutrinos and b quarks at  $\sqrt{s} = 13$  TeV”, *JHEP* **11** (2018) 172, doi:10.1007/JHEP11(2018)172, arXiv:1807.02826.
- [12] M. Rimoldi, “Search for physics beyond the Standard Model in di-lepton final state with the ATLAS experiment”, 2018. Presented 20 Jul 2018.
- [13] V. Shiltsev, “On the Future High Energy Colliders”, in *Proceedings, Meeting of the APS Division of Particles and Fields (DPF 2015): Ann Arbor, Michigan, USA, 4-8 Aug 2015*. 2015. arXiv:1509.08369.
- [14] J. Ellis, “The Future of High-Energy Collider Physics”, in *38th International Symposium on Physics in Collision (PIC 2018) Bogot, Colombia, September 11-15, 2018*. 2018. arXiv:1810.11263.
- [15] CEPC Study Group Collaboration, “CEPC Conceptual Design Report: Volume 1 - Accelerator”, arXiv:1809.00285.
- [16] J. Tang et al., “Concept for a Future Super Proton-Proton Collider”, 2015.
- [17] E. Coleman et al., “The importance of calorimetry for highly-boosted jet substructure”, arXiv:1709.08705.
- [18] C. H. Yeh et al., “Studies of granularity of a hadronic calorimeter for tens-of-TeV jets at a 100 TeV  $pp$  collider”, arXiv:1901.11146.
- [19] C. H. Yeh et al., “Jet Substructure Variables with the SiFCC Detector at 100 TeV”, arXiv:1811.12805. [PoSICHEP2018,905(2018)].
- [20] and and and, “The study of crosstalk in high-speed circuit design”, in *2002 3rd International Symposium on Electromagnetic Compatibility*, pp. 617–620. May, 2002. doi:10.1109/ELMAGC.2002.1177507.
- [21] J. Borg et al., “SKIROC2.CMS an ASIC for testing CMS HGCALE”, *Journal of Instrumentation* **12** (feb, 2017) C02019–C02019, doi:10.1088/1748-0221/12/02/c02019.
- [22] M. Giovannozzi and F. V. der Veken, “Description of the luminosity evolution for the CERN LHC including dynamic aperture effects. Part II: application to Run1 data”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **908** (2018) 1 – 9, doi:https://doi.org/10.1016/j.nima.2018.08.019.
- [23] K. Bloom, “CMS software and computing for LHC Run 2”, *PoS ICHEP2016* (2016) 185, doi:10.22323/1.282.0185, arXiv:1611.03215.
- [24] HEP Software Foundation Collaboration, “A Roadmap for HEP Software and Computing R&D for the 2020s”, *Comput. Softw. Big Sci.* **3** (2019), no. 1, 7, doi:10.1007/s41781-018-0018-8, arXiv:1712.06982.

- [25] Z. LI, "RADIATION HARDNESS / TOLERANCE OF SI SENSORS / DETECTORS FOR NUCLEAR AND HIGH ENERGY PHYSICS EXPERIMENTS.",.
- [26] C. Collaboration, "The Phase-2 Upgrade of the CMS L1 Trigger Interim Technical Design Report", Technical Report CERN-LHCC-2017-013. CMS-TDR-017, CERN, Geneva, Sep, 2017. This is the CMS Interim TDR devoted to the upgrade of the CMS L1 trigger in view of the HL-LHC running, as approved by the LHCC.
- [27] ATLAS Collaboration Collaboration, "Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data", Technical Report ATLAS-CONF-2016-024, CERN, Geneva, Jun, 2016.
- [28] P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons", *Phys. Rev. Lett.* **13** (Oct, 1964) 508–509, doi:10.1103/PhysRevLett.13.508.
- [29] R. Contino et al., "Physics at a 100 TeV pp collider: Higgs and EW symmetry breaking studies ", arXiv:1606.09408. CERN-TH-2016-113.
- [30] M. L. Mangano et al., "Physics at a 100 TeV pp collider: Standard Model processes ", arXiv:1607.01831. CERN-TH-2016-112, FERMILAB-FN-1021-T.
- [31] M. Mangano et al., "Future Circular Collider : Vol. 1 Physics opportunities",.
- [32] M. Benedikt et al., "Future Circular Collider: Vol. 2 The Lepton Collider (FCC-ee)", Technical Report CERN-ACC-2018-0057, CERN, Geneva, Dec, 2018. Submitted for publication to Eur. Phys. J. ST.
- [33] M. Benedikt et al., "Future Circular Collider : Vol. 3 The Hadron Collider (FCC-hh)",.
- [34] F. Zimmermann et al., "Future Circular Collider : Vol. 4 The High-Energy LHC (HE-LHC)",.
- [35] H. Aihara et al., "SiD Letter of Intent", 2009. arXiv:0911.0006. Presented to ILC IDAG.
- [36] C. Adolphsen et al., "The International Linear Collider Technical Design Report - Volume 3. II: Accelerator Baseline Design", 2013. arXiv:1306.6328.
- [37] H. Abramowicz et al., "The International Linear Collider Technical Design Report - Volume 4: Detectors", arXiv:1306.6329.
- [38] S. V. Chekanov et al., "Initial performance studies of a general-purpose detector for multi-TeV physics at a 100 TeV pp collider", *JINST* **12** (2017), no. 06, P06009, doi:10.1088/1748-0221/12/06/P06009, arXiv:1612.07291.
- [39] Z. Li, "Radiation hardness / tolerance of Si sensors / detectors for nuclear and high energy physics experiments", in *Semiconductor pixel detectors for particles and X-rays. Proceedings, International Workshop, PIXEL2002, Carmel, USA, September 9-12, 2002*. 2002.

- [40] J. Allison et al., “Recent developments in Geant4”, *Nuclear Instruments and Methods in Physics Research A* **835** (2016) 186.
- [41] CMS Collaboration, “Search for narrow and broad dijet resonances in proton-proton collisions at  $\sqrt{s} = 13$  TeV and constraints on dark matter mediators and other new particles”, *JHEP* **08** (2018) 130, doi:10.1007/JHEP08(2018)130, arXiv:1806.00843.
- [42] CMS Collaboration, “Search for massive resonances decaying into  $WW$ ,  $WZ$ ,  $ZZ$ ,  $qW$ , and  $qZ$  with dijet final states at  $\sqrt{s} = 13$  TeV”, *Phys. Rev. D* **97** (2018), no. 7, 072006, doi:10.1103/PhysRevD.97.072006, arXiv:1708.05379.
- [43] CMS Collaboration, “Search for resonant  $t\bar{t}$  production in proton-proton collisions at  $\sqrt{s} = 13$  TeV”, *Submitted to: JHEP* (2018) arXiv:1810.05905.
- [44] S. Chekanov, “HepSim: a repository with predictions for high-energy physics experiments”, *Advances in High Energy Physics* **2015** (2015) 136093. Available as <http://atlaswww.hep.anl.gov/hepsim/>.
- [45] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, “Soft Drop”, *JHEP* **05** (2014) 146, doi:10.1007/JHEP05(2014)146, arXiv:1402.2657.
- [46] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, “Better jet clustering algorithms”, *JHEP* **08** (1997) 001, doi:10.1088/1126-6708/1997/08/001, arXiv:hep-ph/9707323.
- [47] M. Wobisch and T. Wengler, “Hadronization corrections to jet cross-sections in deep inelastic scattering”, in *Monte Carlo generators for HERA physics. Proceedings, Workshop, Hamburg, Germany, 1998-1999*. 1998. arXiv:hep-ph/9907280.
- [48] CMS Collaboration, “Jet algorithms performance in 13 TeV data”,.
- [49] A. Tripathy et al., “Jet Substructure Studies with CMS Open Data”, *Phys. Rev. D* **96** (2017), no. 7, 074003, doi:10.1103/PhysRevD.96.074003, arXiv:1704.05842.
- [50] ATLAS Collaboration, “Measurement of the Soft-Drop Jet Mass in pp Collisions at  $\sqrt{s} = 13$  TeV with the ATLAS Detector”, *Phys. Rev. Lett.* **121** (2018), no. 9, 092001, doi:10.1103/PhysRevLett.121.092001, arXiv:1711.08341.
- [51] M. Cacciari, G. P. Salam, and G. Soyez, “The anti-kt jet clustering algorithm”, *JHEP* **0804** (2008) 063, arXiv:0802.1189.
- [52] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, “Longitudinally-invariant  $k_{\perp}$ -clustering algorithms for hadron-hadron collisions”, *Nucl. Phys. B* **406** (February, 1993) 187–224.
- [53] F. A. Dreyer, L. Necib, G. Soyez, and J. Thaler, “Recursive Soft Drop”, *JHEP* **06** (2018) 093, doi:10.1007/JHEP06(2018)093, arXiv:1804.03657.

- 
- [54] A. J. Larkoski, G. P. Salam, and J. Thaler, “Energy Correlation Functions for Jet Substructure”, *JHEP* **06** (2013) 108, doi:10.1007/JHEP06(2013)108, arXiv:1305.0007.
  - [55] C. Neubüser, “Performance Studies and Requirements on the Calorimeters for a FCC-hh Experiment”, in *Proceedings of International Conference on Technology and Instrumentation in Particle Physics 2017*, Z.-A. Liu, ed., pp. 37–43. Springer Singapore, Singapore, 2018.
  - [56] J. Faltova, “Design and performance studies of a hadronic calorimeter for a FCC-hh experiment”, *Journal of Instrumentation* **13** (2018), no. 03, C03016.