

# TRAVEL REPORT

Yu Hsiang Chang, 103282001

National Central University (NCU), Taoyuan City, Taiwan (R.O.C)

February 1 - July 31, 2017

## Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Search for Z' to ZH to llbb with 2016 CMS data</b>	<b>3</b>
2.1	MC/data comparison . . . . .	3
2.2	Double B Tagger Scale Factor . . . . .	6
2.3	Estimating background by ALPHABET method . . . . .	7
2.4	Signal contamination in the tau21 control sample . . . . .	10
2.5	The DY sample cross section correction K-factor . . . . .	11
2.6	Estimating background by BumpHunting method . . . . .	14
<b>3</b>	<b>High Granularity Calorimeter(HGCal)</b>	<b>15</b>
3.1	2016 HGCal paper . . . . .	15
3.2	2017 May Test Beam . . . . .	16
3.3	2017 July Test Beam . . . . .	21
<b>4</b>	<b>Service task for JMAR group</b>	<b>26</b>
<b>5</b>	<b>CMS trigger shift for 2017 data taking</b>	<b>26</b>

## List of Figures

1	The MC/data comparison of the muon channel in the Z $p_t$ (top left), the Z mass(top right), the fat jet $p_t$ (bottom left) and the fat jet Soft Drop mass(bottom right). . . . .	6
---	--	---

2	The MC/data comparison in double B-tagger after applying the scale factor . . . . .	8
3	In tau21 control sample, the Soft Drop mass distributions passing double B-tagger(left) and failing double B-tagger(right) . . . . .	9
4	In tau21 control sample, the pass/fail ratio in Soft Drop mass(left) and the comparison of failing double B-tagger, true passing double B-tagger and predicted double B-tagger in the ZH invariant mass(right) . . . . .	9
5	In search sample, the pass/fail ratio in Soft Drop mass(left) and the comparison of failing double B-tagger, true passing double B-tagger and predicted double B-tagger in the ZH invariant mass(right) . . . . .	10
6	The tau21 distribution of signal samples at different mass points. The pre-selection has been applied. . . . .	11
7	The tau21 distribution of signal at mass point = 1000 GeV(red), DY background(green), and the combined(purple). . . . .	12
8	The generator-level Z $p_t$ distributions of NLO sample, LO $H_T$ binned sample and the combined LO sample, before and after applying the k-factor. . . . .	13
9	Fit the sideband, which fails the double B-tagger and is in the Soft Drop mass window, in the ZH invariant mass . . . . .	15
10	The averaged total noises and the intrinsic noises of different cell types . . . . .	16
11	The PCB used in 2017, where 4 yellow frames indicate the SKIROC-CMS chips	17
12	The setup of 2017 May test beam. The yellow framed texts indicate the scintillators, absorber and the hexaboard (module containing the Silicon sensor,PCB ...etc). The beam direction is indicated by a white arrow. . . . .	18
13	An event display of run 106 (positron beam at 150 GeV), event 64, time sample 5 . . . . .	19

14	High gain ADCs distribution of chip=1/channel=36(from the center of hex-aboard where beam hits)/SCA=0/time sample=0 or 3. The good run 111-122 (around 11000 events) with electron beam at 200 GeV is used. The pion contamination could have possibility to cause MIP in time sample 3, and the time sample 0 shows the noise-like behavior by contrast. To reduce the significant common mode noise, chip=1/channel=58(from position not hit by beam)/SCA=0/time sample=0 or 3 have been used as reference (to be subtracted). A peak is clearly visible around 60 ADC for the MIP. . . . .	20
15	A cartoon shows how the capacitors sample the pulse shape, which is the charges received Silicon sensor in different time(left) and the relationship of SCA and the time sample. . . . .	21
16	Among 9 time samples of an event, using time sample 3,4,5,6 to fit and find the time of arrival and the time of maximum . . . . .	22
17	ADCs distribution in histogram of all time sample (left) and profile of time sample (right) of a good run 93 . . . . .	22
18	ADCs distribution in histogram of all time sample (left) and profile of time sample (right) of a bad run 53 . . . . .	23
19	Three modules glued in one copper cooling plate . . . . .	24
20	The setup of 2017 May test beam, which contain the ECAL and FH and AHCAL from left to right (beam direction) . . . . .	24
21	The movable position of 4 FH modules . . . . .	25
22	An event display of 300 GeV pion beam. The FH part only shows the hex-abords on the center. The 4-the and the 5-th hexaboard have high noise in the edge where the HDMI cable connects . . . . .	25
23	AHCAL data-taking monitoring . . . . .	25

24	the distribution of the charged EM energy(top left), the neutral EM energy(top right), the charged hadron energy(bottom left), and the charged EM energy(bottom right). The y-axes are in log scale . . . . .	27
25	the RECO-level v.s the generator-level 2D distribution of the charged EM energy(top left), the neutral EM energy(top right), the charged hadron energy(bottom left), and the charged EM energy(bottom right). . . . .	27
26	The averaged genjet energy variable fraction v.s $\eta$ . The Sum of all energy fractions equals to unity by the definition. . . . .	28
27	The level-1 trigger subsystem in 2017 . . . . .	29
28	The seats of the CMS control room . . . . .	29
29	Photo of trigger shifter's seat. Two PCs with 4 screen per PC for the DQM shifter(left) and trigger shifter(right) . . . . .	30
30	The trigger L1-page interface . . . . .	30

## List of Tables

1	dataset used for 2016 . . . . .	4
2	MC sample used 2016, the “*” are for different mass points or $H_T$ values . . . . .	4
3	event selection, the “*” are for different version. The data filter and noise cleaning haven't been applied here. The “same” means the same cut muon is applied on the muon channel as on the electron channel. . . . .	5
4	Double B-tagger's scale factor . . . . .	7
5	signal contamination and maximum difference . . . . .	11
6	the derived k-factor . . . . .	13

# 1 Introduction

In February 1 to July 31 of 2017, I was send by NCU to stay at the European Organization for Nuclear Research (CERN) to do my PhD studies, which is near the Geneva Swizterland and located in the boundary between France and Swizterland.

During these time, I worked for my analysis “Search for Z’ to ZH to llbb with 2016 CMS data” and participant the CMS future upgrade project “High Granularity Calorimeter(HGCal)”. Beside these two tasks, I was also assigned a task for JMAR group as the service work and took shift of CMS trigger shift for 2017 data taking.

This traval report is the summary of these four activities.

## 2 Search for Z’ to ZH to llbb with 2016 CMS data

Continuing the work in later 2016, I used 2016 CMS data to search the new particle Z’, which decayed to a Z boson and a Higgs, and then the Z decayed to two leptons (electrons or muons pair) and the Higgs decayed to two b-quarks.

The dataset used is listed in the Table 1, and the total luminosity =  $34.14 \text{ fb}^{-1}$ . The simulated Monte Carlo(MC) used is listed in the Table 2. The event selection is summarized in the Table 3. One can find more detail in [1].

The following subsections describe my progresses in six months.

### 2.1 MC/data comparison

The MC/data comparison was performed to make sure our understanding of physics is correct.

A pre-selection, which was the same to event selection but without the double B-tagger cut and revised the Soft Drop mass to be  $<65$  or  $>135$  (the  $65 < \text{Soft Drop mass} < 105$  was also blinded for W/Z boson region), was applied for MC/data comparison.

Table 1: dataset used for 2016

<b>dateset name</b>
/SingleMuon/Run2016B-23Sep2016-v3/MINIAOD
/SingleMuon/Run2016C-23Sep2016-v3/MINIAOD
/SingleMuon/Run2016D-23Sep2016-v3/MINIAOD
/SingleMuon/Run2016E-23Sep2016-v3/MINIAOD
/SingleMuon/Run2016F-23Sep2016-v3/MINIAOD
/SingleMuon/Run2016G-23Sep2016-v3/MINIAOD
/SingleMuon/Run2016H-PromptReco-v2/MINIAOD
/SingleMuon/Run2016H-PromptReco-v3/MINIAOD
/SingleElectron/Run2016B-23Sep2016-v3/MINIAOD
/SingleElectron/Run2016C-23Sep2016-v3/MINIAOD
/SingleElectron/Run2016D-23Sep2016-v3/MINIAOD
/SingleElectron/Run2016E-23Sep2016-v3/MINIAOD
/SingleElectron/Run2016F-23Sep2016-v3/MINIAOD
/SingleElectron/Run2016G-23Sep2016-v3/MINIAOD
/SingleElectron/Run2016H-PromptReco-v2/MINIAOD
/SingleElectron/Run2016H-PromptReco-v3/MINIAOD

Table 2: MC sample used 2016, the “\*” are for different mass points or  $H_T$  values

<b>MC sample name</b>
ZprimeToZhToZlephbb_narrow_M-*_13TeV-madgraph
DYJetsToLL_M-50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8
DYJetsToLL_M-50_HT-*to*_TuneCUETP8M1_13TeV-madgraphMLM-pythia8
TT_TuneCUETP8M2T4_13TeV-powheg-pythia8
ZZ_TuneCUETP8M1_13TeV-pythia8
WZ_TuneCUETP8M1_13TeV-pythia8
WW_TuneCUETP8M1_13TeV-pythia8
ZH_HToBB_ZToLL_M125_13TeV_amcatnloFXFX_madspin_pythia8
ZH_HToBB_ZToLL_M125_13TeV_powheg_pythia8

Table 3: event selection, the “\*” are for different version. The data filter and noise cleaning haven’t been applied here. The “same” means the same cut muon is applied on the muon channel as on the electron channel.

	electron channel	muon channel
data filter	-	-
trigger	HLT_Ele105_CaloIdVT_GsfTrkIdT_v*	HLT_Mu45_eta2p1_v*
nVtx	$\geq 1$	same
<b>lepton</b>		
ID	two loose	one HighPt and one Custom tracker
Iso	already in ID	trackerIso<0.1
$p_t$	$>135, 20$	$>55, 20$
$\eta$	$ \eta  < 2.4$ , no $1.44 <  \eta  < 1.566$	$ \eta  < 2.4$
<b>Z boson</b>		
$p_t$	$>200$	same
mass	$70 < \text{mass} < 200$	same
<b>Higgs AK8 jet</b>		
ID	one loose	same
lepton cleaning	no jet overlapped in $\Delta R < 0.8$	same
$p_t$	$>200$	same
$\eta$	$ \eta  < 2.4$	same
Soft Drop mass	$105 < \text{SDmass} < 135$	same
b-tagging	Double B-tagger $>0.3$	same
noise cleaning	-	-

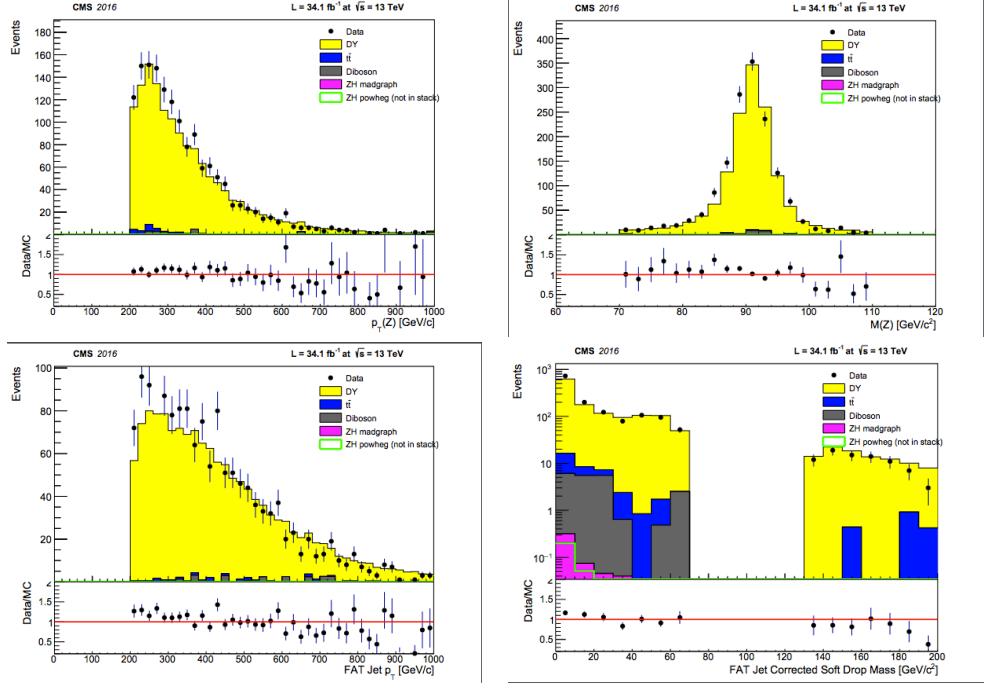


Figure 1: The MC/data comparison of the muon channel in the  $Z$   $p_t$ (top left), the  $Z$  mass(top right), the fat jet  $p_t$ (bottom left) and the fat jet Soft Drop mass(bottom right).

The results of the  $Z$  boson and the Higgs jet's transverse momentum ( $p_t$ ) and the mass in the muon channel are showed in Figure 1. The results in the electron channel are simialr to the muon channel.

Although the MC/data comparisons were not perfect, but since the final background estimation (see Sec. 2.3 ) did not rely on MC, the MC was still acceptable to use in later studies.

## 2.2 Double B Tagger Scale Factor

The double B-tagger is an algorithm used to identify the fat jet containing two b-quarks [2]. For every comparison between data and MC in the certain cut, the MC must be applied this cut's scale factor(SF) to cover the difference in efficency to data. This subsection describes the applying double B-tagger's SF to MC/data comparison.

Table 4: Double B-tagger's scale factor

signal topology		0-250	250-350	350-430	430-840	>840
$p_t$ bin	SF	0.96	0.96	1.00	1.01	1.01
background topology		0-250	250-350	350-430	430-700	>700
$p_t$ bin	SF	1.044	1.044	1.074	1.119	1.119

In general, the formula of the event weight could be more complicated when the event has more than one jets passing double B-tagger, but since this analysis only has one jet, each event just times the SF as the event weight of double B-tagger.

There are two SFs depending on the jet passing double B-tagger belonged to the signal topology ( $H \rightarrow bb$ ) or background topology ( $t\bar{t}$ ) [3]. The values of two kinds of SFs are listed in Table 4.

In this analysis, a fat jet tagged as b-quark and it has at least one subjet tagged as b-quark in the MC truth was considered as the signal topology, while a fat jet tagged as other (c-quark or u,d,s-quark or gluon) was considered as the background topology.

The MC/data comparison after applying the pre-selection, the double B-tagger cut and its SF is showed in Figure 2.

### 2.3 Estimating background by ALPHABET method

The ALPHABET method is to use the events in data failing the double B-tagger to predict the event passing double B-tagger in Soft Drop mass window ( $105 < SDmass < 135$ ). In the assumption that the double B-tagger and the Soft Drop mass have no correlation, one can fit the pass/fail double B-tagger event ratio in the left and right sidebands of the Soft Drop mass window then extrapolate the fit function to the mass window. Therefore, with the events failing double B-tagger and the pass/fail ratio description in the mass window, one can predict the events passing the double B-tagger and in the Soft Drop mass window, which is our final signal region(SR).

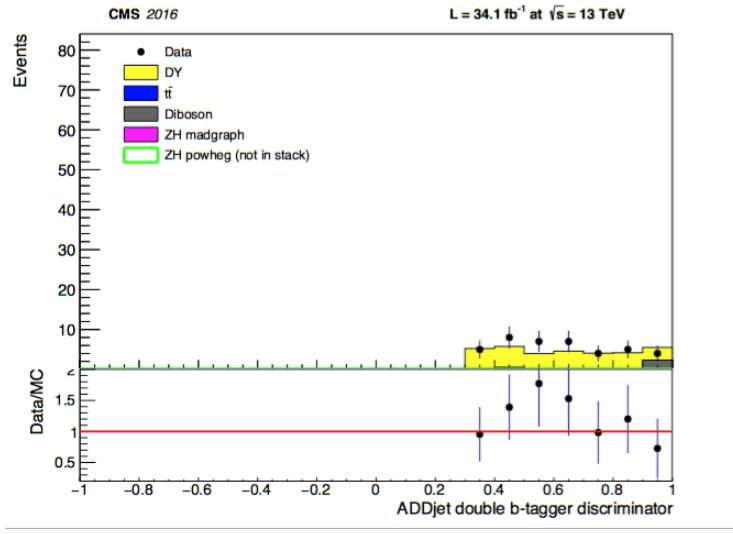


Figure 2: The MC/data comparison in double B-tagger after applying the scale factor

To increase the statistics, the electron channel and the muon channel were combined together.

This method could be validated by performing the same procedure in a control sample in data by requesting  $\text{tau21} > 0.4$ , which was assumed with negligible signal events, and then opening events passing double B-tagger and in the mass window to compare the true events and the predicted events. The passing and failing events in Soft Drop mass distribution is shown in Figure 3. The pass/fail ratio in the Soft Drop mass and the predicted and the true Zh invariant mass distribution are shown in Figure 4. The fit function for pass/fail ratio is the second order polynoimal (see Eqn. 1).

$$y = p_0 \cdot x^2 + p_1 \cdot x + p_2 \quad (1)$$

The difference between the true and the predicted events in the  $\text{tau21}$  control sample is around 9%, which is not big so there is potentials to use this method.

The pass/fail ratio in the Soft Drop mass and the predicted Zh invariant mass in the search sample (the data containing the real signal region) are shown in Figure 5. The events in signal region are blinded here.

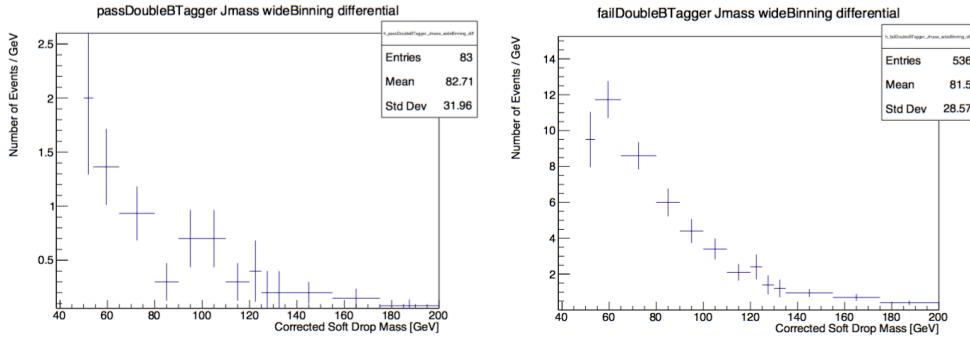


Figure 3: In tau21 control sample, the Soft Drop mass distributions passing double B-tagger(left) and failing double B-tagger(right)

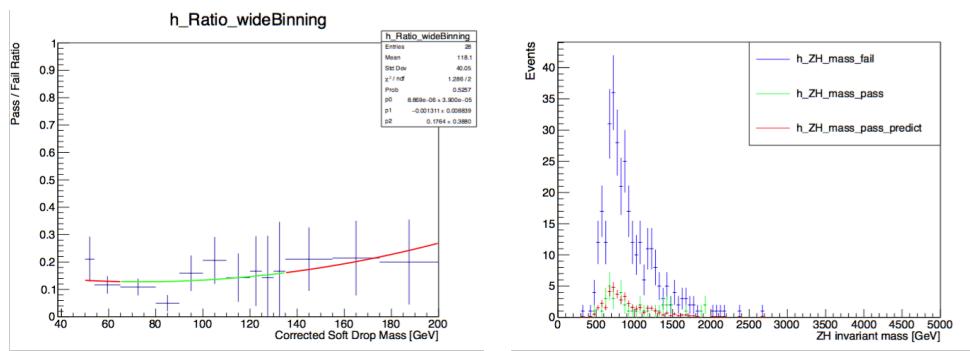


Figure 4: In tau21 control sample, the pass/fail ratio in Soft Drop mass(left) and the comparison of failing double B-tagger, true passing double B-tagger and predicted double B-tagger in the ZH invariant mass(right)

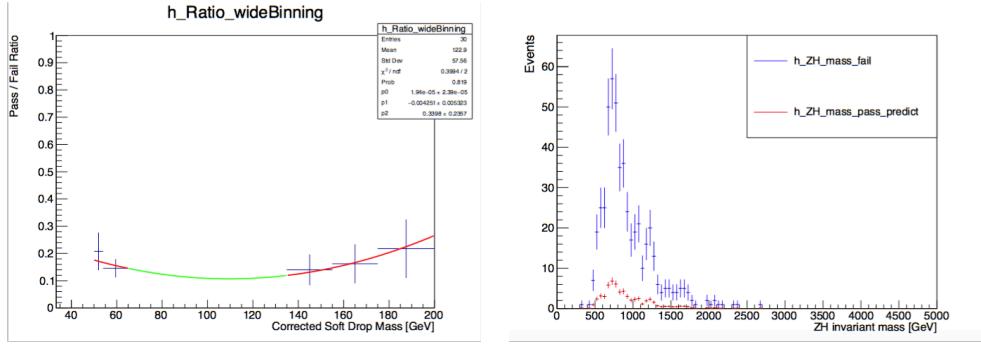


Figure 5: In search sample, the pass/fail ratio in Soft Drop mass(left) and the comparison of failing double B-tagger, true passing double B-tagger and predicted double B-tagger in the ZH invariant mass(right)

## 2.4 Signal contamination in the tau21 control sample

In previous subsection the validation in the tau21 control sample is shown, which is in the assumption that negligible or no signal events passing double B-tagger and in the mass window. So this subsection is for answering that if there is signal contamination then how it affects the validation.

The tau21 distributions in the muon channel of the signal sample at different mass points are shown in Figure 6. Each signal sample is scaled to data luminosity =  $34.1 \text{ fb}^{-1}$ .

The signal sample at mass = 1000 GeV is taken as the benchmark to compare with the backgrounds (dominated by DY sample), because it has more events after scaling. The tau21 distribution of the signal, the DY and the combined are shown in Figure 7. The signal fraction in  $\tau_{21} > 0.4$  is 7.6% when the signal sample at mass = 1000GeV exists. And in this case the presence of 7.6% signal fraction could change the difference between predicted background and true background (data events subtracts the signal contamination) from original 9% to 17.9%, as shown in Eqn. 2. The N is the data events passing double B-tagger and in mass window, and the calculation is in the case that predict events >original true events (the opposite case does not change the maximum difference).

Table 5: signal contamination and maximum difference

mass point	1000	1200	1400	1600	1800	2000	2500	3000	3500	4000
signal contanmination	7.8%	7.6%	5.2%	3.9%	2.2%	1.5%	1%	<1%	<1%	<1%
maximum difference	17.9%	14%	13%	11%	10%	9%	9%	9%	9%	9%

tau21 of signal MC weighted to lumi=34.1/fb

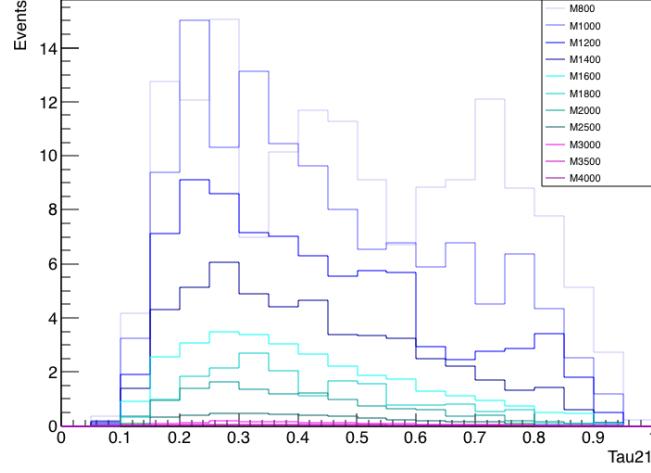


Figure 6: The tau21 distribution of signal samples at different mass points. The pre-selection has been applied.

$$\text{maximum difference} = \frac{(N \times 1.09 - N \times 0.924)}{N \times 0.924} = 17.9\% \quad (2)$$

If the signal does exist, the maximum differences in the tau21 control sample for validation are listed in Table 5.

## 2.5 The DY sample cross section correction K-factor

In this analysis, four DY  $H_T$ -binned samples (100-200, 200-400, 400-600, 600-Infinity, the 600-Infinity is combined by other higher  $H_T$ -binned samples) are used for simulation studies, where the  $H_T$  is the scalar sum of hadron  $p_t$  in the LHE level. However, the cross section of  $H_T$ -binned samples is in the leading order (LO) only. To describe the DY

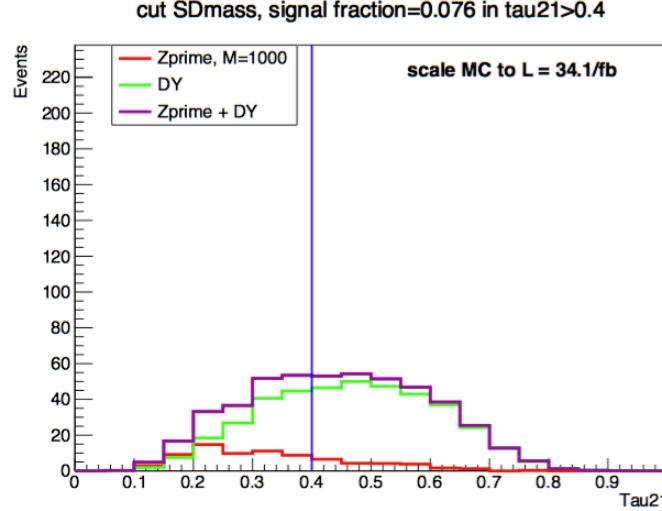


Figure 7: The tau21 distribution of signal at mass point = 1000 GeV(red), DY background(green), and the combined(purple).

process more precisely to the next leading order(NLO), one need to derive the correction k-factors for each DY  $H_T$ -binned samples's cross sections.

To derive the k-factors, one need to compare the NLO DY sample and the sum of each LO DY  $H_T$ -binned samples in the generator-level Z boson  $p_t$ (gen-Z  $p_t$ ) distribution when all samples are scaled to the same luminosity, and this is indeed equivalent to solve four linear equations with four unknowns as shown in Eqn. 3. The  $k_i$  is the k-factor for each DY  $H_T$ -binned sample, and the subscripts  $i=1,2,3,4$  correspond to four LO  $H_T$  binned samples 100-200, 200-400, 400-600, 600-Infinity respectively. The coefficients  $x^j$  correspond to the normalization of four samples (a,b,c,d) and NLO DY sample (e), and the superscripts  $j=1,2,3,4$  correspond to four generator-level Z boson  $p_t$  ranges 100-200, 200-400, 400-600, 600-Infinity respectively.

$$\begin{aligned}
 k_1 \cdot a^1 + k_2 \cdot b^1 + k_3 \cdot c^1 + k_4 \cdot d^1 &= e^1 \\
 k_1 \cdot a^2 + k_2 \cdot b^2 + k_3 \cdot c^2 + k_4 \cdot d^2 &= e^2 \\
 k_1 \cdot a^3 + k_2 \cdot b^3 + k_3 \cdot c^3 + k_4 \cdot d^3 &= e^3 \\
 k_1 \cdot a^4 + k_2 \cdot b^4 + k_3 \cdot c^4 + k_4 \cdot d^4 &= e^4
 \end{aligned} \tag{3}$$

Table 6: the derived k-factor				
H <sub>T</sub> -binned sample	100-200	200-400	400-600	600-Inf
k-factor	1.416	1.426	1.410	1.039

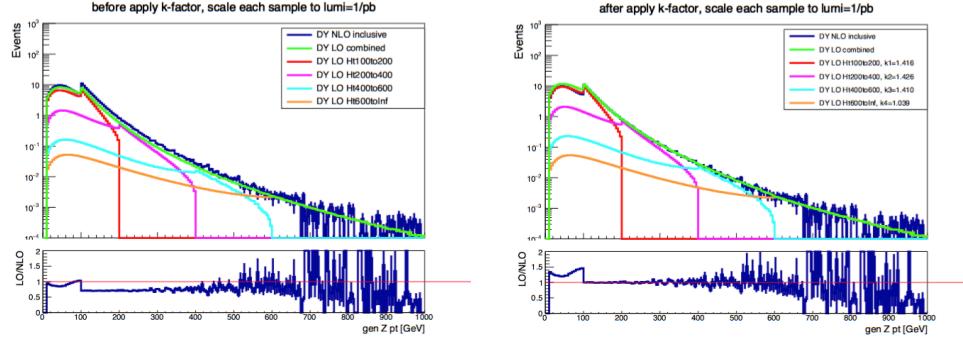


Figure 8: The generator-level Z  $p_t$  distributions of NLO sample, LO  $H_T$  binned sample and the combined LO sample, before and after applying the k-factor.

Since the gen-Z  $p_t$  will drop very rapid in the right edge of the  $H_T$ , the equations can be reduced to the easier one in Eqn.4 (the bottom-left parts of the off-diagonal line are zero).

$$\begin{aligned}
 k_1 \cdot a^1 + k_2 \cdot b^1 + k_3 \cdot c^1 + k_4 \cdot d^1 &= e^1 \\
 k_1 \cdot 0 + k_2 \cdot b^2 + k_3 \cdot c^2 + k_4 \cdot d^2 &= e^2 \\
 k_1 \cdot 0 + k_2 \cdot 0 + k_3 \cdot c^3 + k_4 \cdot d^3 &= e^3 \\
 k_1 \cdot 0 + k_2 \cdot 0 + k_3 \cdot 0 + k_4 \cdot d^4 &= e^4
 \end{aligned} \tag{4}$$

The derived k-factors are listed in Table 6, and the gen-Z  $p_t$  distributions of before and of after applying k-factors are shown in Figure 8. To be consistent with the LO DY  $H_T$ -binned sample in low  $H_T$  region, the  $H_T > 100$  has been applied in NLO DY sample.

## 2.6 Estimating background by BumpHunting method

Due to the insufficient statistics, we decided to use another method “BumpHunting” to describe the shape of background in the signal region, whereas the normalization is obtained by ALPHABET still.

In the BumpHunting method, the background shape of the  $Zh$  invariant mass spectrum in the signal region is assumed to be similar to the shape in the sideband (passing double B-tagger and in mass window) but allow to have small bump (it’s the reason for the name “BumpHunting”) for the possibly presence of signal. Therefore, one can fit the sideband first with a function and then use the similar functional form (with initial parameters from the first fit ) to fit the signal region or do the simultaneous fit in both the sideband and the signal region. The Eqn. 5 is the fit function for the sideband, and the Eqn.6 is the same function times the averaged value of the pass/fail ratio in the signal region obtained from the ALPHABET method, which is used to describe the background in the signal region originally. But for allowing a small bump in the background shape in the signal region, one linear term “ $a$ ” is adding to the Eqn. 6 to become Eqn.7 for the functional form used in the fit of the signal region finally.

$$y = f(x) \quad (5)$$

$$y = R_{\frac{p}{f}} \cdot f(x) \quad (6)$$

$$y = (1 + a \cdot x) \cdot R_{\frac{p}{f}} \cdot f(x) \quad (7)$$

The fit in the sideband region is shown in Figure 9, and the function form of fit is the levelled exponential (Eqn. 8).

$$y = e^{\frac{-p_1 x}{1 + p_1 \cdot p_2 \cdot x}} \quad (8)$$

This work haven’t finished and is going on still.

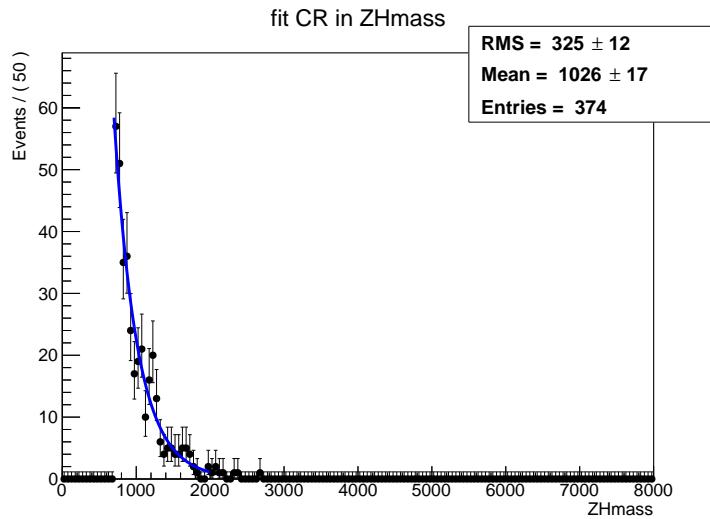


Figure 9: Fit the sideband, which fails the double B-tagger and is in the Soft Drop mass window, in the ZH invariant mass

### 3 High Granularity Calorimeter(HGCal)

In 2017, I continued to join the CMS phase II endcap calorimeter upgrade project “High Granularity Calorimeter(HGCal)” [4]. First I worked for the paper, which was the summary of our 2016 HGCal activity, then I participated two 2017 Test Beam at CERN, May and July Test beam (mainly to take shifts).

#### 3.1 2016 HGCal paper

For the 2016 HGCal paper, which was a detector note numbered as DN-17-011 [5], with Sandhya and Kai-Yu, who were my partners in 2016 for pedestal and noise study, we summarized our results and wrote our section in paper. Beside what we already had, we added a plot of the intrinsic noise to the paper.

For run with a lots of events on a single channel, the total noise is estimated by taking the RMS of the ADC counts distribution, and in the presence of the common mode noise, which are the noises with same amount happened in the channels of the same

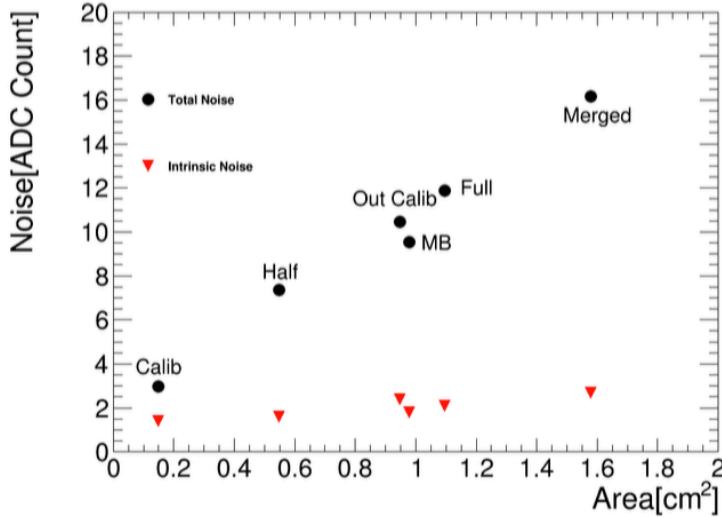


Figure 10: The averaged total noises and the intrinsic noises of different cell types

conditions in an events, one can define the intrinsic noise(IN noise) in Eqn9, where the  $ADCs\_i$  is the ADC counts of i-th channels substracted the pedestal value in an event, and the  $N$  is the number of channels with the same conditions in an event. The  $(-1)^i$  is used to cancel the common mode noise by each pair of channels in an event, then finally the bracket is taken the RMS as the IN noise.

$$IN\ noise = \left[ \frac{\sum_{i=1}^N (-1)^i \times ADCs_i}{\sqrt{N}} \right]_{RMS} \quad (9)$$

The intrinsic noise and the averaged total noise in different cell types are shown in Figure 10. The intrinsic noise is quite small compared to the total noise, which means the dominated noise is the common mode noise.

### 3.2 2017 May Test Beam

During 8<sup>th</sup> to 15<sup>th</sup> May 2017, one test beam was host in CERN's North Area beam facility (EHNA1, building 887, the H2 beamline). The purpose is to test the new ASIC chip (SKIROC-CMS, see Figure 11), new PCB, and new DAQ system (EUDAQ).

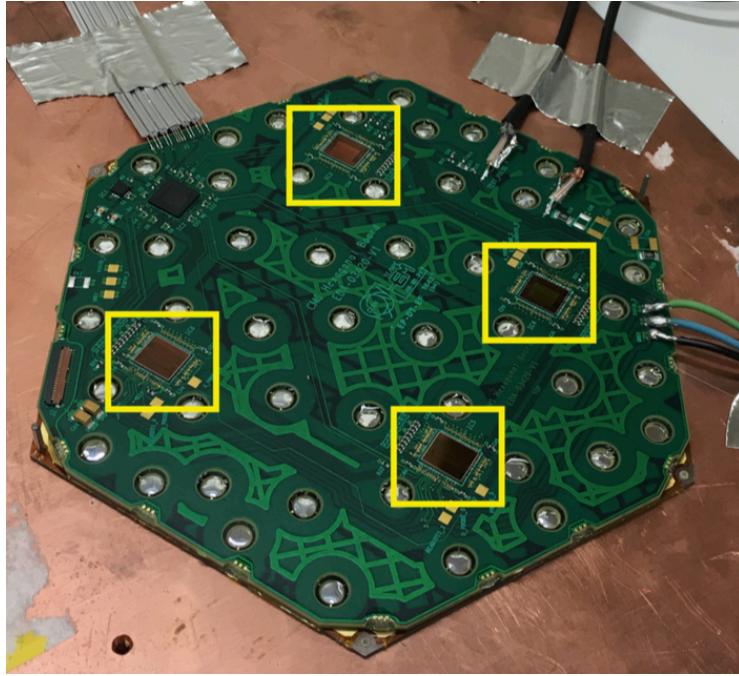


Figure 11: The PCB used in 2017, where 4 yellow frames indicate the SKIROC-CMS chips

The test beam setup is shown in Figure 12. One module was tested for the EM endcap calorimeter(EE) in this test. The data of several beam energies (positron at 20,32,50,100,150,200,250 GeV [6] were taken. The electron signal could be seen as an event display shown in Figure 13. A non-showering single particle(MIP)’s signal due to the Pion contamination was also observed as shown in Figure 14.

One feature of May test beam (and July test beam) in 2017 is the ASIC chip “SKIROC-CMS” can save thirteen time samples, whereas the ASIC chip used in 2016 saved one time sample only, so this provides us more information to analyze the data (but also more complicated).

In each channel of the SKIROC-CMS chip there are 13 switched capacitor array(SCA) to save the charges received from the Silicon sensor in turns of the period of 25 ns per SCA. Once the 13-th SCA has been charged the amount in that 25 ns period, then the first SCA will be discharged (erase original information) and charged again for the next 25 ns period. This cycle will keep going until the trigger signal arrives, whose delay time is

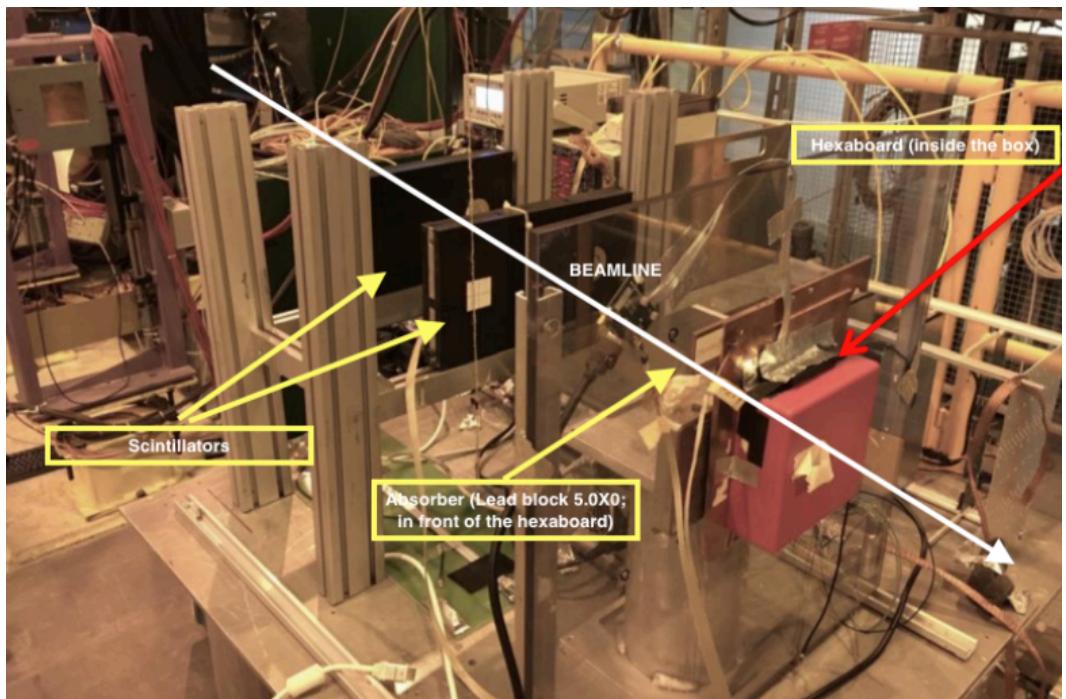


Figure 12: The setup of 2017 May test beam. The yellow framed texts indicate the scintillators, absorber and the hexaboard (module containing the Silicon sensor,PCB ...etc). The beam direction is indicated by a white arrow.

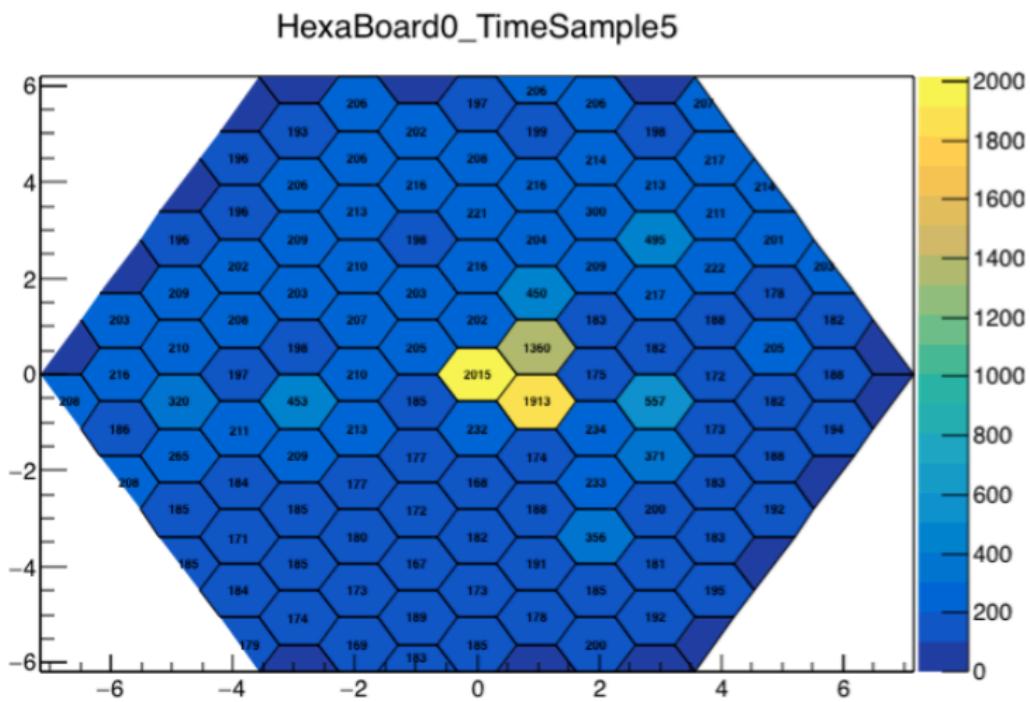


Figure 13: An event display of run 106 (positron beam at 150 GeV), event 64, time sample 5

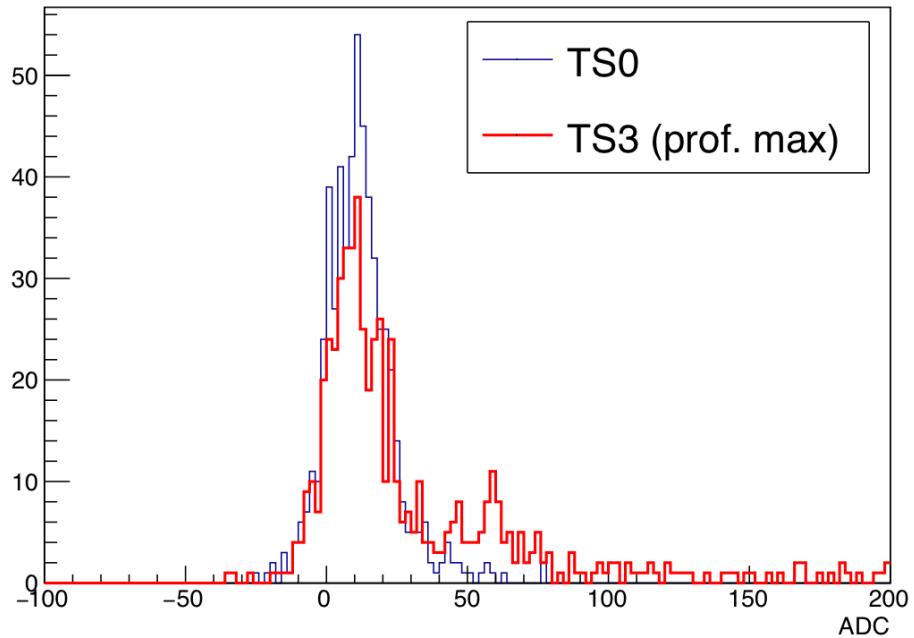


Figure 14: High gain ADCs distribution of chip=1/channel=36(from the center of hexaboard where beam hits)/SCA=0/time sample=0 or 3. The good run 111-122 (around 11000 events) with electron beam at 200 GeV was used. The pion contamination could have possibility to cause MIP in time sample 3, and the time sample 0 shown the noise-like behavior by contrast. To reduce the significant common mode noise, chip=1/channel=58(from position not hit by beam)/SCA=0/time sample=0 or 3 had been used as the reference (to be subtracted). A peak was clearly visible around 60 ADC for the MIP.

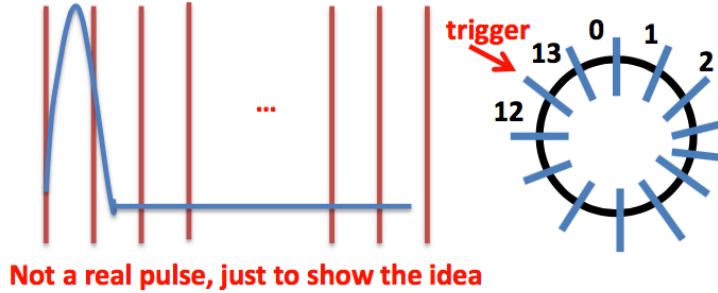


Figure 15: A cartoon shows how the capacitors sample the pulse shape, which is the charges received Silicon sensor in different time(left) and the relationship of SCA and the time sample.

around 200 ns (if  $>25 \times 13$  ns then we lost the event), then the whole 13 SCAs' information (the amount of charges) will be send out, and in this step the 13 SCAs will be called 13 time samples according to the coming order, where the two SCAs near the trigger become the 12th and 13th time samples as the cartoon shown in Figure 15.

Except the last 4 time samples' values, which do not reflect the charges received in the Silicon sensor due to the trigger signal, the values of other 9 time samples describ the pulse shape in an event as in Figure 16.

In the pedestal run in May test beam ( and aslo in July test beam), there are few good runs and a lot of bad run as the examples shown in Figure 17 and Figure 18. Since it's hard to find a stable pedestal run to represent the pedestal in all times, finally we decided to use the first SCA as the pedestal value to do the pedestal subtraction in every event.

### 3.3 2017 July Test Beam

During 12<sup>th</sup> to 19<sup>th</sup> July 2017, another test beam was host in the same place. The purpose is to measure the hadronic shower in the front hadron calorimeter(FH). Due to the nuclear interaction length (the mean free path of a nuclear interacting particle) is usually larger than the radiation length, the width of hadronic shower is also larger than the width of EM shower, therefore we bound three modules (instead of one in May) together in

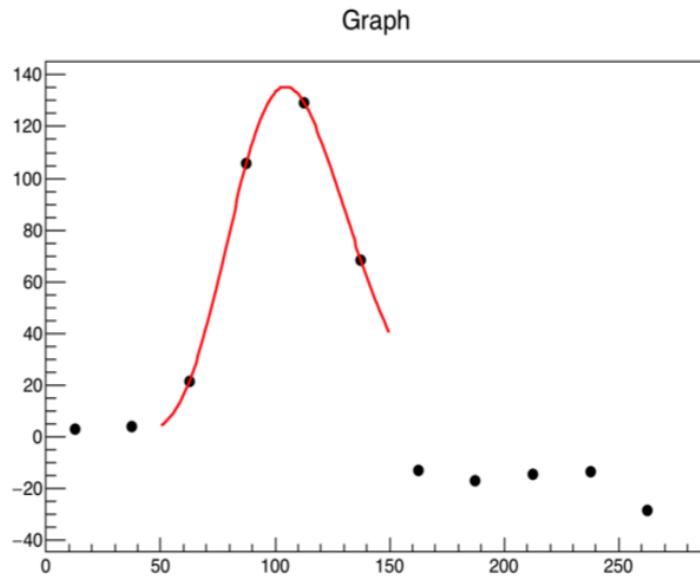


Figure 16: Among 9 time samples of an event, using time sample 3,4,5,6 to fit and find the time of arrival and the time of maximum

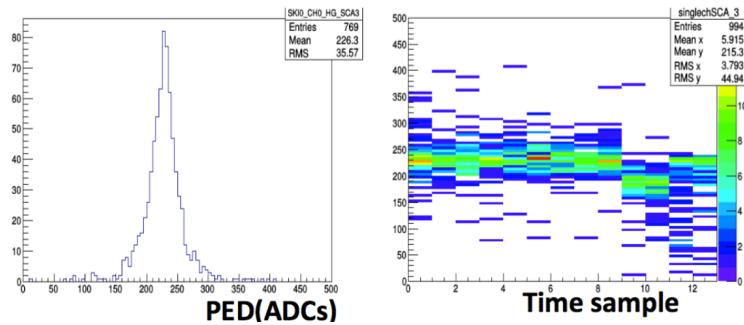


Figure 17: ADCs distribution in histogram of all time sample (left) and profile of time sample (right) of a good run 93

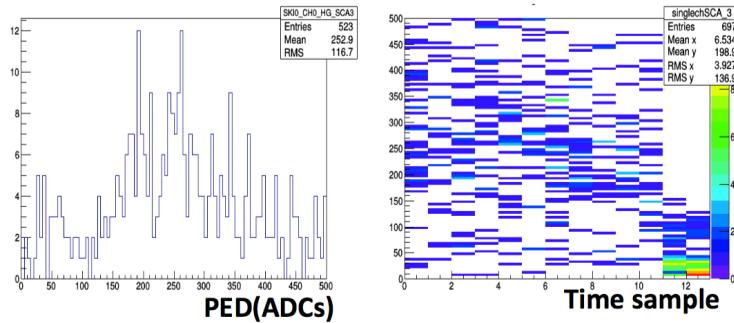


Figure 18: ADCs distribution in histogram of all time sample (left) and profile of time sample (right) of a bad run 53

the Copper cooling plate to cover the wide width of the hadronic shower (see Figure 15). However, the limited number of modules in hands in that time forced us to use one module per layer for two layers (1+1) to mimic the EE part and use one module per layer for two layers and three modules per layer for two layers (1+1+3+3) to mimic the FH part. In the downstream of FH, we borrowed the AHCAL modules from the CALICE experiment [7] to mimic the backing hadron calorimeter(BH). The setup of full July test beam and the side view of FH are shown in Figure 16 and Figure 17.

In the period of data-taking, we had ever difficult time for the low data-taking rate (this was solved after the IPbus protocol between the chip and EUDAQ PC was established) and the EUDAQ stuck, so only in the last 8 hours after we fixed all issues we could take more statistical data. The data of pion beam (at 100, 150, 200, 250 and 300 GeV), electron beam(at 150 and 250 GeV), and muon beam (at 150 GeV)are taken [8]. One event display of HCal and the monitoring of AHCAL data-taking are shown in Figure 18 and Figure 19.

After the booking period, we also kept our setup in H2 beamline and benifited from the experiments in the upstream of us in either using the residue muons or that they allow us to use the beam in their spare time to take more data.



Figure 19: Three modules glued in one copper cooling plate

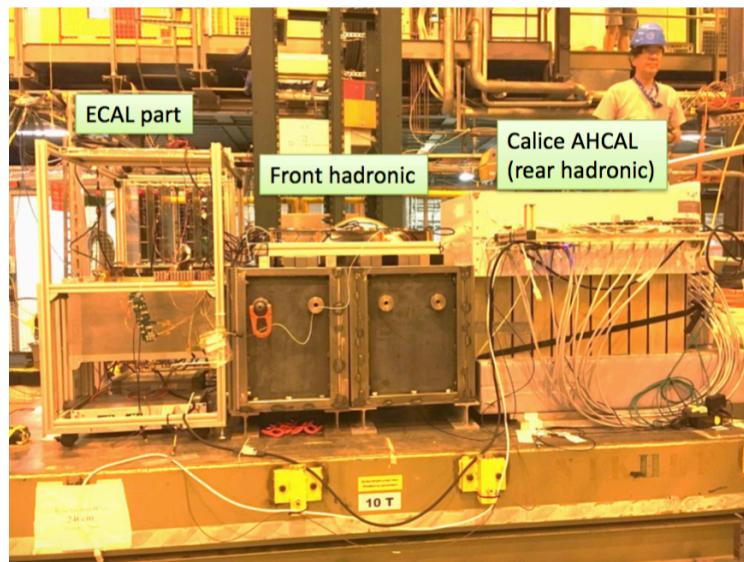


Figure 20: The setup of 2017 May test beam, which contain the ECAL and FH and AHCAL from left to right (beam direction)

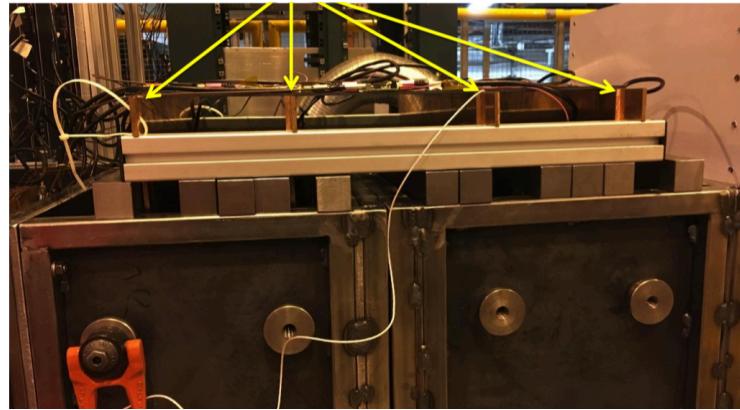


Figure 21: The movable position of 4 FH modules

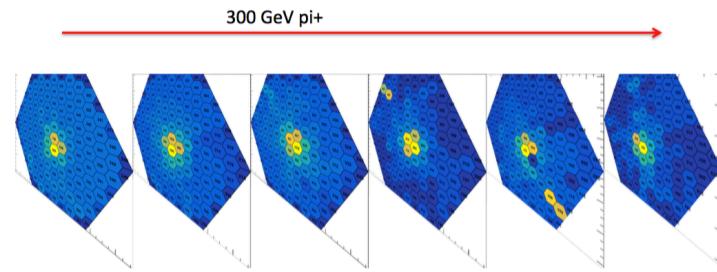


Figure 22: An event display of 300 GeV pion beam. The FH part only shows the hexaboard on the center. The 4-th and the 5-th hexaboard have high noise in the edge where the HDMI cable connects

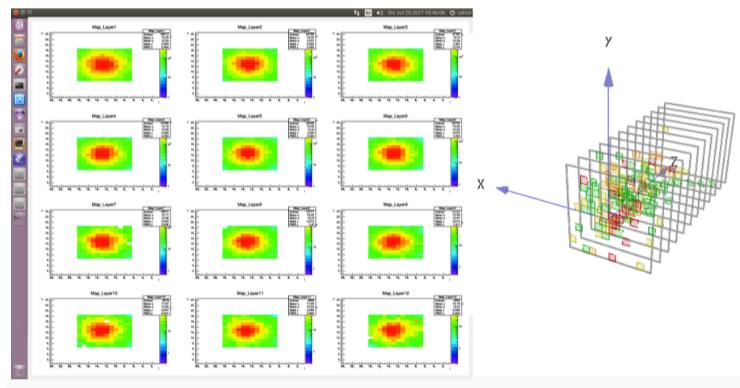


Figure 23: AHCAL data-taking monitoring

## 4 Service task for JMAR group

In June and Juky I worked for a task for the Jet MissingET Algorithm Reconstruction (JMAR) group as the service work, which was to adding few genjet energy fraction variables to the current CMSSW code, because they were missing in current code. The variables that I added were the charged EM energy, the neutral EM energy, the charged hadron energy and the neutral hadron energy, and I used the particle ID to classify each generator-level particles composed the genjet and then summed their energy to the corresponding energy variables. My github code can be found in [9].

Also I produced a private TTbar sample to validate the task. Figure 24 shows the four variables. Figure 25 shows the variables in the RECO-level v.s generator-level, where three variables look linear but the neutral hadron energy does not, which is mainly caused by the definition of the RECO-level neutral hadron energy in CMSSW code doesn't contain the electromagnetic calorimeter(ECAL) energy whereas the generator-level one does.

Figure 26 shows the averaged energy fraction in the TTbar sample. The lepton cleaning (remove the genjet overlapped with the leptons from the W boson in its cone size) had been implemented. The rest lepton fraction was caused by the lepton decayed from the hadronization of the b-quark.

This result has been represented in the JMAR meeting and the XPOG meeting. Now I am waiting for the pull request to merged my code to current CMSSW code.

## 5 CMS trigger shift for 2017 data taking

During March to June 2017, I took the trigger shifts of the Compact Muon Solenoid(CMS) experiment. The trigger system is used to reduce the rate of data-taking and only select the event that people are interest. In CMS, the trigger is composed by the level-1 trigger (L1) and the high-level trigger (HLT), where the L1 is the hardware-based trigger to reduce the rate <100 kHz by basic thresholds and the HLT is the computer-farm based trigger to subsequently reduce the rate <1 kHz by the algorithms. All L1 subsystems

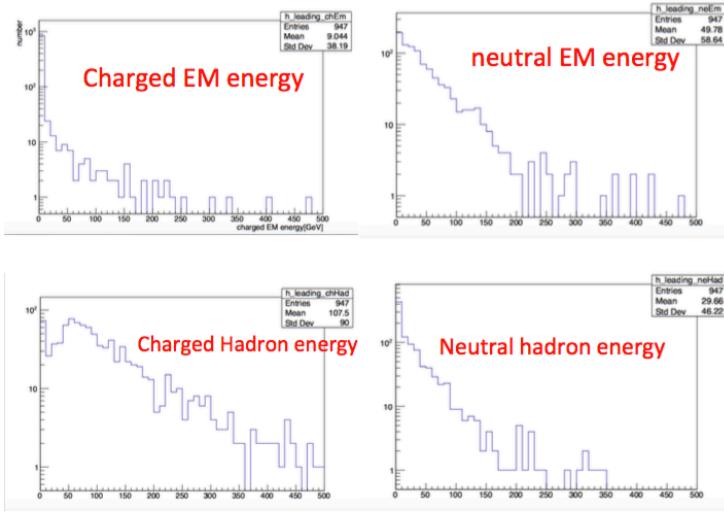


Figure 24: the distribution of the charged EM energy (top left), the neutral EM energy (top right), the charged hadron energy (bottom left), and the charged EM energy (bottom right). The y-axes are in log scale

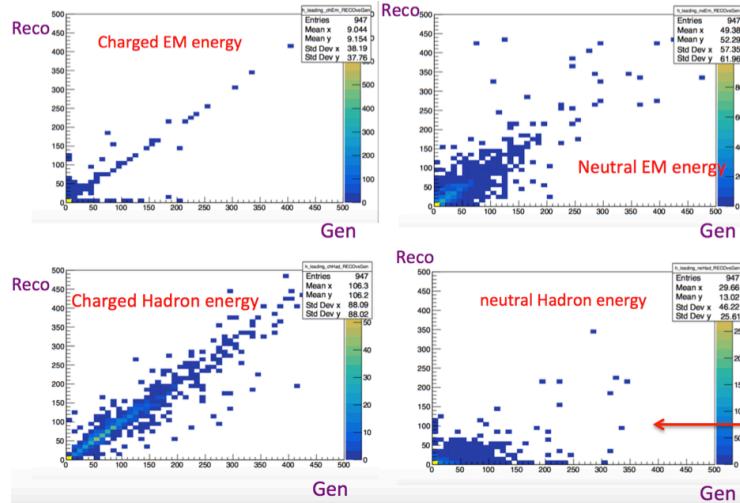


Figure 25: the RECO-level v.s the generator-level 2D distribution of the charged EM energy (top left), the neutral EM energy (top right), the charged hadron energy (bottom left), and the charged EM energy (bottom right).

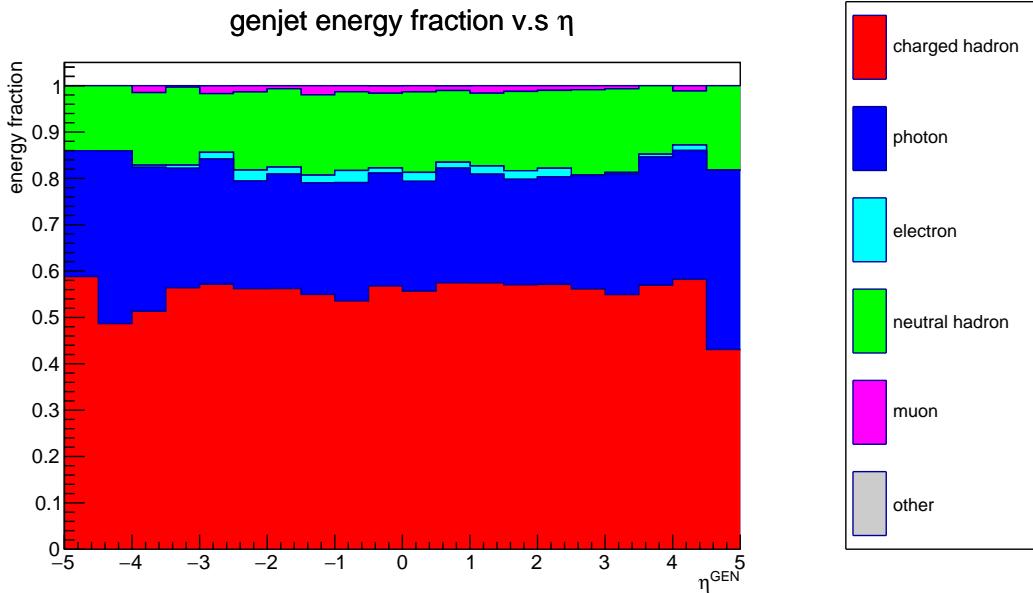
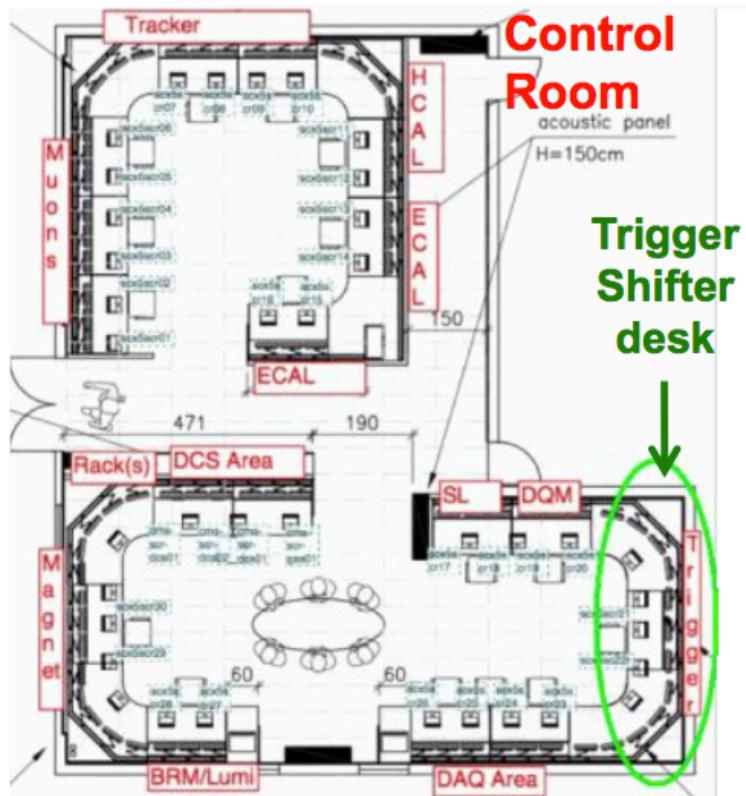
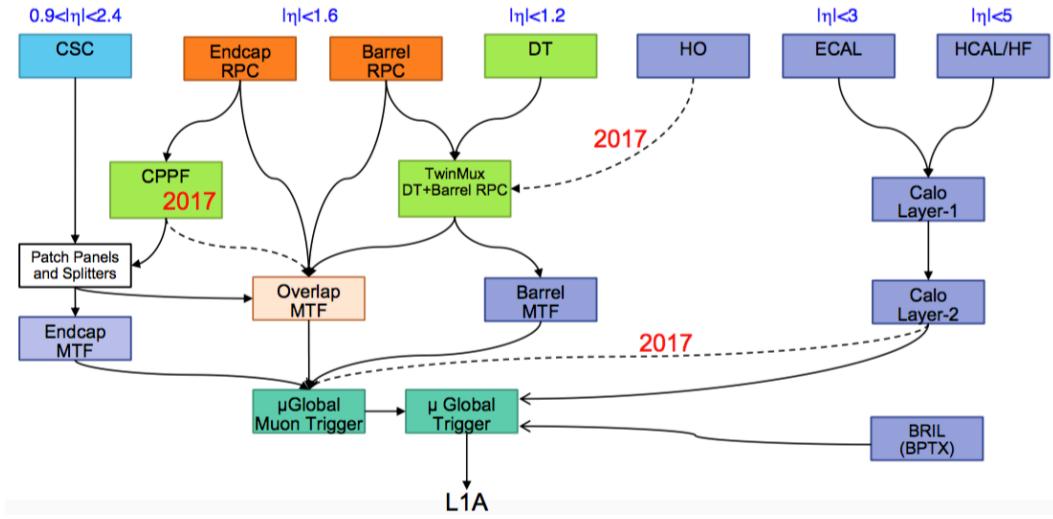


Figure 26: The averaged genjet energy variable fraction v.s  $\eta$ . The Sum of all energy fractions equals to unity by the definition.

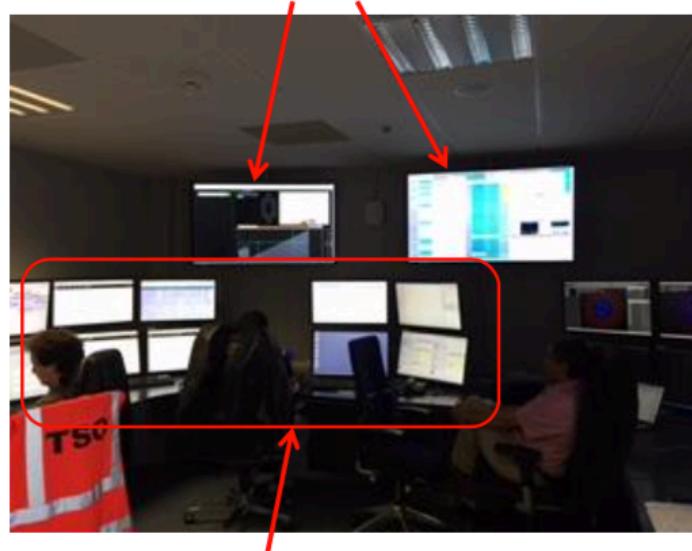
used in 2017 are shown in Figure 27.

The trigger shifter's role is to cooperate with other crews in the CMS control room (in building 3562 or SCX5) located at the LHC Point5 when the CMS is in operation. The trigger shifter need to make sure the triggers in use working well (function ok and has resonable rate), communicate with other crews when error happens, change the pre-scale column, which decide differnt pre-scale number for different trigger, to suppress total rate according to the current luminosity, and upload a shift report to the Elog in the end of the shift. The seats and photo of trigger shift are shown in Figure 28 and Figure 29, and the trigger L1-page, which is the one of the most frequently used interfaces, is shown in Figure 30.

In totally I took 15 shifts and get 19 credits for the NCU institue, and each shift took 8 hours (11-7,7-15 or 15-23).



## Run Information



**There are 2 PCs**

Figure 29: Photo of trigger shifter's seat. Two PCs with 4 screen per PC for the DQM shifter(left) and trigger shifter(right)

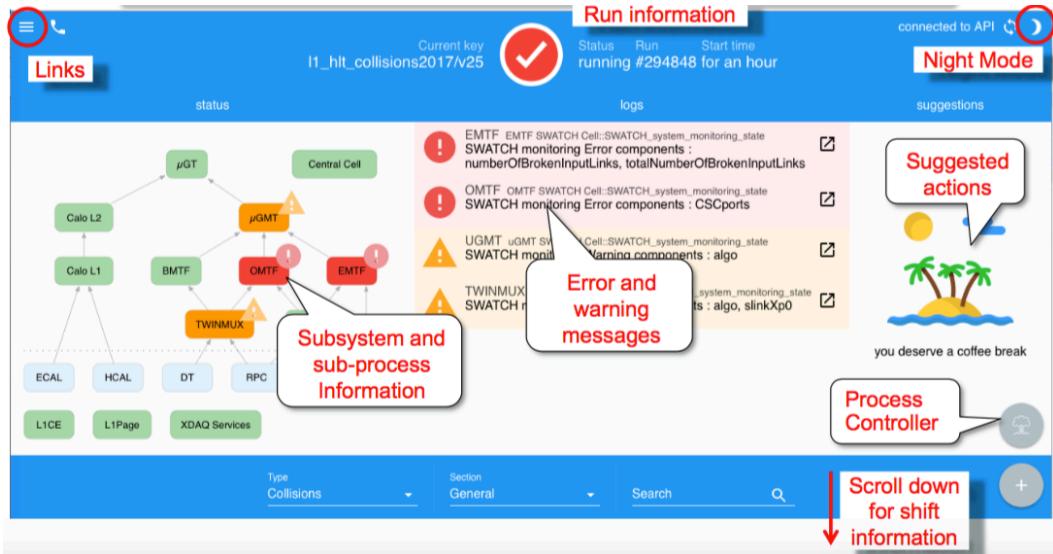


Figure 30: The trigger L1-page interface

## References

- [1] 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”, CERN, 2016.  
<https://arxiv.org/abs/1610.08066>
- [2] CMS Collaboration, “Identification of double-b quark jets in boosted event topologies”, CERN, 2016.  
<https://cds.cern.ch/record/2195743>
- [3] CMS internal twiki link, “Usage of b/c Tag Objects for 13 TeV Data in 2016 and 80X MC”.  
<https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation80XReReco>
- [4] CMS Collaboration, “CMS Forward Calorimeters Phase II Upgrade”, CERN, 2014.  
<https://cds.cern.ch/record/1969369>
- [5] CMS internal twiki link, “Paper Summarising 2016 Beam Tests at FNAL and CERN”, 2017.  
<https://twiki.cern.ch/twiki/bin/viewauth/CMS/BeamTestPaper2016>
- [6] Google doc for 2017 May test beam data.  
[https://docs.google.com/spreadsheets/d/1ju1r2Be2dpvIa7\\_QnCeufZ9UqPE1CeJrtBJivwTI68U/edit#gid=0](https://docs.google.com/spreadsheets/d/1ju1r2Be2dpvIa7_QnCeufZ9UqPE1CeJrtBJivwTI68U/edit#gid=0)
- [7] Home page of the CALICE experiment.  
<https://twiki.cern.ch/twiki/bin/view/CALICE/WebHome>
- [8] Google doc for 2017 July test beam data.  
[https://docs.google.com/spreadsheets/d/1ZUL7VM7EK651mL\\_FGNUkcBkE0tJtqkRVMfAqTUYqt7g/edit#gid=91437087](https://docs.google.com/spreadsheets/d/1ZUL7VM7EK651mL_FGNUkcBkE0tJtqkRVMfAqTUYqt7g/edit#gid=91437087)
- [9] My github repository forked for the JMAR task  
<https://github.com/yuchanggit/cmssw/blob/master/RecoJets/JetProducers/src/JetSpecific.cc>  
<https://github.com/yuchanggit/cmssw/blob/master/DataFormats/JetReco/interface/GenJet.h>