# The development of missing transverse momentum reconstruction with the ATLAS detector using the PUfit algorithm in $pp$ collisions at 13 TeV

by

Zhelun Li
B.Sc., University of Victoria, 2017

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Physics and Astronomy

The development of missing transverse momentum reconstruction with the ATLAS detector using the PUfit algorithm in $pp$ collisions at 13 TeV

by

Zhelun Li
B.Sc., University of Victoria, 2017

Supervisory Committee

---

Dr. R. Kowalewski, Supervisor
(Department of Physics and Astronomy)

---

Dr. R.Keeler, Departmental Member
(Department of Physics and Astronomy)

**Supervisory Committee**

---

Dr. R. Kowalewski, Supervisor
(Department of Physics and Astronomy)

---

Dr. R.Keeler, Departmental Member
(Department of Physics and Astronomy)

## ABSTRACT

Many interesting physical processes produce non-interacting particles that could only be measured using the missing transverse momentum. The increase of the proton beam intensity in the Large Hadron Collider (LHC) provides sensitivity to rare physics processes while inevitably increasing the number of simultaneous proton collisions in each event. The missing transverse momentum (MET) is a variable of great interest, defined as the negative sum of the transverse momentum of all visible particles. The precision of the MET determination deteriorates as the complexity of the recorded data escalates. Given the current complexity of data analysis, a new algorithm is developed to effectively determine the MET. Several well-understood physics processes were used to test the effectiveness of the newly designed algorithm. The performance of the new algorithm is also compared to that of the standard algorithm used in the ATLAS experiment.

# Contents

# List of Tables

# List of Figures

ACKNOWLEDGEMENTS

I would like to thank:

**all my great office mates** for the good memories.

**my Supervisor Dr Kowalewski and researcher Dr Hamano Kenji,** for mentoring, support, encouragement, and patience.

**All researchers affiliated with CERN/UVic/TRIUMF that I worked with,** for teaching me technical details patiently.

*A scientific man ought to have no wishes, no affections — a mere heart of stone.*
Charles Darwin,
Letter to T.H. Huxley, 9 July 1857

# DEDICATION

To Xiaoyun Kong

# Chapter 1

# Introduction

The **missing transverse momentum** ($E_T^{\text{miss}}$ ), which is sometimes referred to as the MET, is a crucial variable to study elusive particles produced in the large hadron collider. Many interesting physics processes involve invisible particles that escape the detection system. Consequently, the measured momenta of all particles produced in the collision will not be conserved because of the missing energy carried by elusive particles. $E_T^{\text{miss}}$ is sensitive to the momentum imbalance of a given collision, which allows us to determine the momentum carried by invisible particles without detecting them directly.

The **Large Hadron Collider** (LHC) collides bunches of high-energy protons, and many proton-proton scattering interactions occur in each bunch crossing. To produce more rare physics processes of interest, the LHC is being upgraded to offer a higher beam intensity to generate more collisions in each bunch crossing. As the LHC elevates the beam intensity to an unprecedented level, the increased number of collisions poses a challenge to the determination of $E_T^{\text{miss}}$ . To compute $E_T^{\text{miss}}$ one needs to sum the transverse momenta of all particles produced in the same collision. However, our ability to associate particles with their corresponding collisions is limited. This results in a significant deterioration in $E_T^{\text{miss}}$ resolution as the number of collisions increases. The PUfit algorithm is developed to effectively determine $E_T^{\text{miss}}$ given the large number of collisions. Specifically, this thesis is focused on the soft term which is a component of the $E_T^{\text{miss}}$ made from low momentum particles. The PUfit algorithm determines a new type of soft term using physics constraints to achieve higher resolution.

The rest of this chapter will be dedicated to introducing interesting physics processes that produce $E_T^{\text{miss}}$ . The experimental determination of $E_T^{\text{miss}}$ will be discussed

in Chapter 2. The ATLAS detector and the large hadron collider will be introduced in Chapter 3. Chapter 4 mainly focuses on the selection of reconstructed physics objects, which is the standard selection criteria developed by the ATLAS group. My original work on the PUfit algorithm starts from Chapter 5, in which the PUfit algorithm is presented with various input alternatives. Results of the implementation of the PUfit algorithm is presented and discussed in Chapter 6. Other alternative versions of the PUfit implementation will be briefly discussed in Chapter 7. Finally, the performance of PUfit is summarized in Chapter 8. A glossary of terms is included in Appendix A to help remind the readers of the nomenclatures used in the thesis.

## 1.1 The Standard Model

The building blocks of all matter are called fundamental particles which interact with each other through four known forces (interactions): Gravity, Strong force, Weak force and Electromagnetic force. The understanding of these fundamental particles and most interactions are encoded in a theory called the **Standard Model**. The Standard Model is a theory developed in the 1970s to explain three of the four forces, excluding gravity. It has been tested by multiple experiments over the years while agreeing with experimental results to an unprecedented level.

In the framework of the Standard Model, all matter in the universe can be decomposed into two types of fundamental particles: **quarks** and **leptons**. There are six different particles in each of the two types mentioned above. The six quarks are divided into three generations each with two quarks, with the two lightest quarks comprising the first generation. The heavier and less stable quarks are then grouped into the second and third generation. Similarly, leptons are also paired into three different generations. Electron, muon and tau particles are each in one of the generations with their corresponding neutrinos. Although electrons, muons, and tau particles have electric charges and noticeable mass, neutrinos are neutral particles with very small masses, which only interact through the weak force.

Three out of the four forces, excluding the gravitational force, are described by exchanges of carrier particles. These particles are called **gauge bosons**. Energy is transmitted between particles by exchanging the corresponding gauge boson of the interaction. The strong force is carried by gluons, and the electromagnetic force is carried by photons. There are three gauge bosons that carry the weak force: $W^{\pm}$ and $Z$. A table of the fundamental particles is shown in Table 1.1.

| Particle Type | Name | Letter | Generation | Spin | charge | Mass |
|---|---|---|---|---|---|---|
| Quarks | up | $u$ | I | $\frac{1}{2}$ | $\frac{2}{3}$ | 2.05 $MeV/c^2$ |
| | down | $d$ | I | | $-\frac{1}{3}$ | 4.7 $MeV/c^2$ |
| | charm | $c$ | II | | $\frac{2}{3}$ | 1.28 $GeV/c^2$ |
| | strange | $s$ | II | | $-\frac{1}{3}$ | 96 $MeV/c^2$ |
| | top | $t$ | III | | $\frac{2}{3}$ | 173.1 $GeV/c^2$ |
| | bottom | $b$ | III | | $-\frac{1}{3}$ | 4.18 $GeV/c^2$ |
| Leptons | e-neutrino | $\nu_e$ | I | $\frac{1}{2}$ | 0 | $< 2.2eV/c^2$ |
| | electron | $e$ | I | | $-1$ | 0.511 $MeV/c^2$ |
| | $\mu$-neutrino | $\nu_\mu$ | II | | 0 | $< 0.17$ $MeV/c^2$ |
| | muon | $\mu$ | II | | $-1$ | 105.66 $MeV/c^2$ |
| | $\tau$-neutrino | $\nu_\tau$ | III | | 0 | $< 18.2$ $MeV/c^2$ |
| | tau | $\tau$ | III | | $-1$ | 1.7768 $GeV/c^2$ |
| Gauge bosons | gluons | $g$ | — | 1 | 0 | 0 |
| | photon | $\gamma$ | — | | 0 | 0 |
| | Z boson | $Z$ | — | | 0 | 91.19 $GeV/c^2$ |
| | W bosons | $W$ | — | | $\pm 1$ | 80.39 $GeV/c^2$ |
| Scalar boson | higgs | $H$ | — | 0 | 0 | 124.97 $GeV/c^2$ |

Table 1.1: Fundamental particles in the Standard Model

Despite the success that the Standard Model has had over the years, it is still considered to be incomplete since the model does not include gravity. On top of that, the Standard Model does not include dark matter which only interacts gravitationally. Dark matter could be made of elusive particles that have not been discovered yet. Many hypothetical particles, in particular those that are candidates for Dark Matter, are predicted to interact weakly with SM particles. Therefore, they would only deposit a negligible amount of energy in the detector if they are produced by collisions in the LHC. However, they are still subject to conservation laws. Therefore, it is fruitful to infer the existence of such particles based on $E_T^{\text{miss}}$ which indicates the amount of momentum imbalance in the detector

## 1.2 Sources of missing transverse momentum

### 1.2.1 Processes within the Standard Model

In the Standard Model, the neutrino is a weakly interacting particle that often leaves no signal in detectors. According to the latest results shown in Ref.[1], the mass of neutrino $\nu_{\mathrm{e}}$ can not exceed 2.05 eV with a 95% confidence level. Since neutrinos interact weakly with other SM particles, they are invisible to most detectors. Displays of collision events involving neutrinos are shown in Figure 2.1.

Processes that procedure neutrinos in the Standard Model are major sources of the genuine $E_{\mathrm{T}}^{\mathrm{miss}}$ in events. Two major decays that involve neutrinos are W/Z bosons decaying into leptons and neutrinos. For example, in the decay $W^- \rightarrow e^-\bar{\nu}_e$ or $W \rightarrow \mu\nu_\mu$ , the charged leptonic signal will be captured by the calorimeter or the muon spectrometer while neutrinos escape from detection, leaving an imbalance of momentum in the event to be captured by $E_T^{\mathrm{miss}}$ .

### 1.2.2 Dark Matter

There is an enormous amount of astronomical evidence that suggests the existence of hypothetical **dark matter** which only interact gravitationally based on observations. Fritz Zwicky proposed the idea of dark matter to explain the observation that the velocity dispersion of distant galaxies are larger than the estimation based on luminous matter [2]. If dark matter is made of particles which interact with standard model particles, we should be able to produce them in the large hadron collider.

One of the most widely known models for dark matter is **weakly interacting massive particles** (WIMPs). WIMPs are particles in the mass range of $\mathrm{m}_{\mathbf{weak}} \sim 10$ GeV/c - TeV/c that only interact through the weak and gravitational interaction [3]. Since WIMPs are assumed to couple with the weak sector of the Standard Model, they could be produced in *pp* collisions in the LHC, and the coupling strength could be experimentally determined.

The calorimeter and tracker are not able to capture signals of WIMPs due to their elusive nature, but other Standard Model particles that emerge from the same process will be detected. Therefore, the existence of dark matter produced in *pp* collisions can only be inferred from the missing momentum in collision events.

### 1.2.3 Supersymmetry

**Supersymmetry** (SUSY) is an extension of the current Standard Model, and it is seen as a possible solution to many physics problems. SUSY offers explanations to many profound physics problems that we face today. It could be able to solve the hierarchy problem of the Standard Model to explain the large observed difference between the strength of the weak force and gravity, while providing dark matter candidates [4]. In SUSY, the **lightest stable supersymmetric particle** (LSP) is a name for the lightest SUSY particle that is stable if the R-parity is conserved. Instead of having conservation of the baryon and lepton number, the SUSY model proposes the conservation of the R-parity which is taken to be $+1$ for all SM particles and $-1$ for all SUSY particles. This prevents the LSP from decaying into SM particles. When other supersymmetric particles decay into LSP, the production of Standard Model particles accompanying the LSP would be detected. Therefore, $E_T^{\text{miss}}$ is a crucial variable to study SUSY decays given the non-interacting nature of the LSP.

# Chapter 2

# Missing transverse momentum

The proton beam consists of bunches of protons spaced 25ns apart along the beam axis. When two bunches of protons collide in the ATLAS detector, many proton-proton collisions take place in that bunch-crossing. The reconstructed interaction point of a recorded collision is called a **primary vertex**. In each proton-proton collision, it is the individual constituents of protons (quarks and gluons) that collide with each other. Since protons have internal structure, the momenta carried by the colliding constituents within the system varies in each collision. The **hard scatter** (HS) is defined as the hardest collision in the event, meaning that the colliding constituents carry large momenta. The "hardness" of a collision can be thought of as a way to measure how violent the collision is. The hard scatter is normally associated with the **hard scatter vertex** which is defined as the vertex with the largest sum of transverse momenta ($p_T$) of tracks associated with it. The other primary vertices are assumed to be produced by **pile-up interactions** which are collisions with constituents carrying less momenta. The particles inferred from detector signals in the data reduction procedure are referred to as physics objects or, for short, objects. All objects are classified into hard scatter objects and pile-up objects based on the type of collisions in which they are produced.

The **missing transverse momentum** $\vec{E}_T^{\text{miss}}$ is a 2-dimensional vector in the transverse plane defined as the negative sum of transverse momenta of all hard scatter particles. The magnitude of $\vec{E}_T^{\text{miss}}$ is referred to as $E_T^{\text{miss}}$ . The conservation law requires momentum of the two proton system to be conserved, resulting in a zero $E_T^{\text{miss}}$ . Non-zero $E_T^{\text{miss}}$ suggests the possible production of particles that interact weakly with the detector. A value incompatible with zero could be from the production of either standard model (SM) neutrinos or new particles beyond the SM that escape

the ATLAS detector without being detected[5].

## 2.1 Overview of the $E_T^{\text{miss}}$ determination

The goal of $E_T^{\text{miss}}$ determination is to quantify the momentum balance in the transverse plane by considering the momenta of all particles produced in the hard scatter. When the $E_T^{\text{miss}}$ is determined by the offline algorithm, two types of objects are used as inputs. The first is referred to as **hard objects** which are a collection of reconstructed electrons, photons, muons, $\tau$-leptons and also jets. A jet is a spray of hadronic particles which are loosely contained within a conical region. To be considered as a hard object, the reconstructed particle is required to pass a $p_{\text{T}}$ threshold and other selection criteria depending on the type of particle. Therefore, hard objects have relatively higher momentum, and they are assumed to be produced by the hardest interaction. The second contribution to $E_{\text{T}}^{\text{miss}}$ is called the **soft term**, which is comprised of hard scatter particles with low momenta that do not pass the hard object selection. The detail of how the soft term is computed is given in Section 4.2.

ATLAS uses a dedicated reconstruction procedure for each kind of particle as well as for jets, creating a particle or jet hypothesis based on the physical nature of the detector signals[5]. Reconstructed particles are fed into the offline algorithm to reconstruct the hard scatter $E_{\text{T}}^{\text{miss}}$. For each reconstructed object, the 2-dimensional transverse momentum vector $\vec{p}_T$ is determined by projecting the 3-dimensional momentum $p$ onto the transverse plane:

$$\begin{aligned} p_x &= p \sin\theta \cos\phi \\ p_y &= p \sin\theta \sin\phi \\ \vec{p}_T &= (p_x, p_y) \end{aligned} \tag{2.1}$$

The origin of the ATLAS coordinate system is taken to be the nominal interaction point while the z-axis is defined to be the direction along the beamline. Then the transverse plane can be easily defined as the plane that is perpendicular to the beamline with the positive x-axis pointing towards the center of the LHC ring and the positive y-axis point upward to the surface of the earth. The azimuthal angle $\phi$ is then defined as the angle in the transverse plane that measures relative to the x-axis. The polar angle $\theta$ is measured relative to the z-axis.

In practice, ATLAS uses a slightly different co-ordinate system in which $\theta$ is replaced by a variable called pseudorapidity. Pseudorapidity, $\eta$, is particularly useful

because differences in $\eta$ are Lorentz invariant under longitudinal boosts along the beam axis. $\eta$ is defined by:

$$\eta = -\ln\tan(\frac{\theta}{2}) \tag{2.2}$$

The $\vec{E}_T^{\text{miss}}$ of the hard scatter interaction is defined as the negative vectorial sum of the transverse momentum of hard objects and soft signals:

$$\begin{aligned}
\vec{E}_{\text{T}}^{\text{miss}} &= (E_{\text{x}}^{\text{miss}}, E_{\text{y}}^{\text{miss}}) \\
&= -\sum_{i\in\{\text{hard objects}\}} \vec{p}_{\text{T,i}} - \sum_{j\in\{\text{soft signals}\}} \vec{p}_{\text{T,j}} \\
&= -\sum_{\text{electrons}} \vec{p}_{\text{T}}^{\,e} - \sum_{\text{photons}} \vec{p}_{\text{T}}^{\,\gamma} - \sum_{\tau-\text{leptons}} \vec{p}_{\text{T}}^{\,\tau_{\text{had}}} - \sum_{\text{muons}} \vec{p}_{\text{T}}^{\,\mu} - \sum_{\text{HS jets}} \vec{p}_{\text{T}}^{\,\text{jet}} - \sum_{j\in\{\text{soft signals}\}} \vec{p}_{\text{T,j}}
\end{aligned} \tag{2.3}$$

Then it is easy to construct other widely used variables based on Equation 2.3:

$$\begin{aligned}
E_{\text{T}}^{\text{miss}} &= |\vec{E}_{\text{T}}^{\text{miss}}| \\
&= \sqrt{(E_{\text{x}}^{\text{miss}})^2 + (E_{\text{y}}^{\text{miss}})^2}
\end{aligned} \tag{2.4}$$

$$\phi^{\text{miss}} = \tan^{-1}\left(\frac{E_{\text{y}}^{\text{miss}}}{E_{\text{x}}^{\text{miss}}}\right) \tag{2.5}$$

The $E_T^{\text{miss}}$ has a non-negative value by definition. $\phi^{\text{miss}}$ is the azimuthal angle that represents the direction. The $E_T^{\text{miss}}$ is determined in both the trigger and the offline computation.

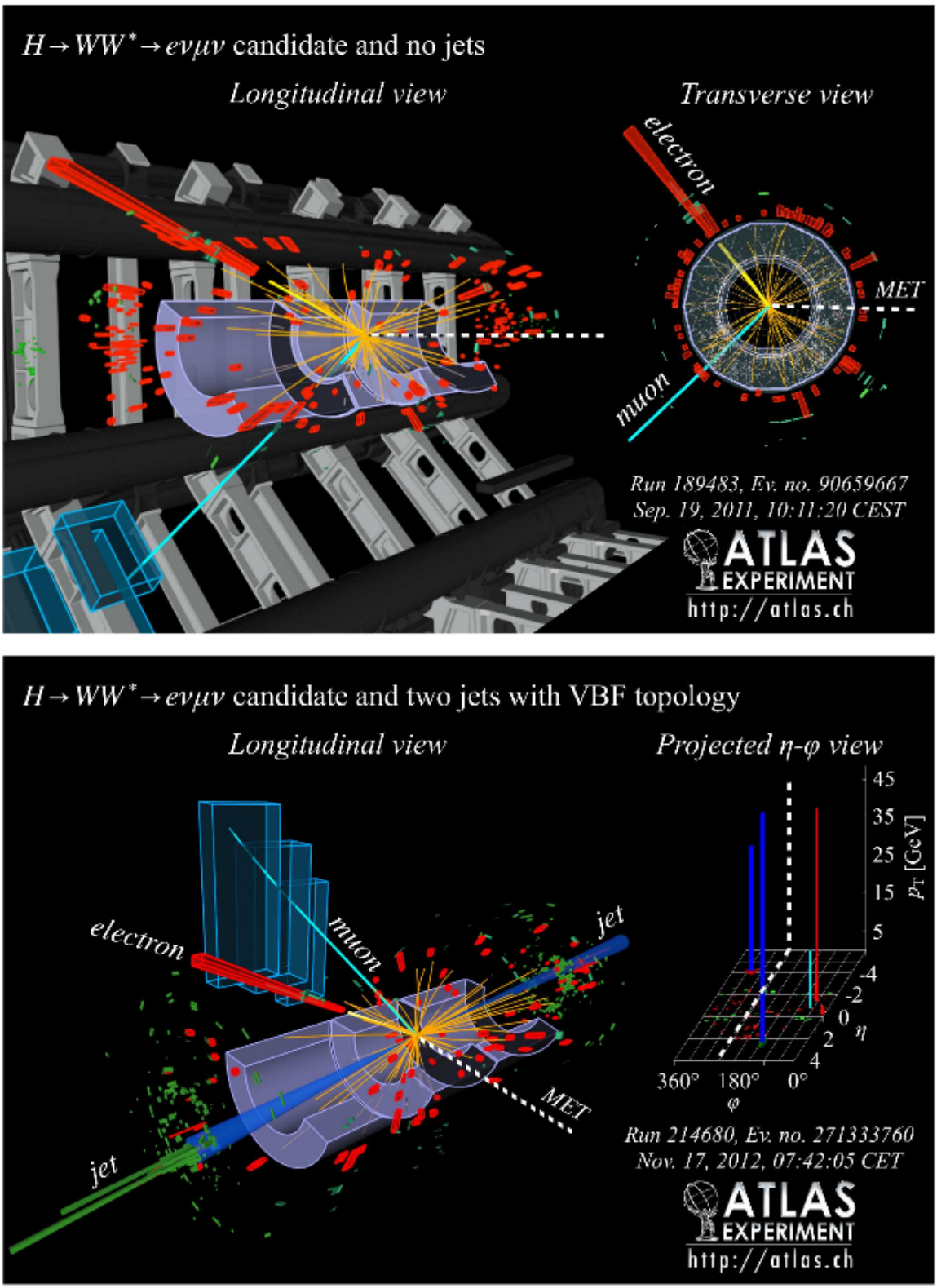


Figure 2.1: Event displays of H $\to$ WW* $\to$ e$\nu\mu\nu$ candidates[6]. The dotted white line represents the MET determined. Blue, light blue and red lines represent jets,electrons and muons respectively

## 2.2 Pile-up interactions

Additional proton-proton collisions in the same bunch crossing produce soft particles with low energies that accompany the energy deposit of hard scatter particles. This is referred to as in-time pile-up[5]. Other than the in-time pile-up, the calorimeter is also sensitive to residual signals from the previous and subsequent bunch crossings (out-of-time pile-up), as discussed in Ref.[7]. The resolution of missing transverse momentum is dependent on pile-up, since larger recorded energies result in larger measurement fluctuations leading to an apparent imbalance (fake $E_T^{\text{miss}}$ ). A number of strategies for mitigating the impact of pile-up on $E_T^{\text{miss}}$ have been developed; one of them is the subject of this thesis.

The **instantaneous luminosity** is defined as the ratio of the measured interaction rate divided by the cross-section, which is sensitive to the total number of interactions produced. It can be calculated using the revolution frequency $f$, number of bunches $n_b$, number of particles in bunches $N$ and the beam widths $\sigma_{x,y}$:

$$\mathcal{L} = \frac{n_b N^2 f}{4\pi\sigma_x\sigma_y} \tag{2.6}$$

The **integrated luminosity** is the integral of the instantaneous luminosity with respect to time. In order to increase the potential for physics discoveries, the integrated luminosity of the Large Hadron Collider was increased to a much higher level in 2017. A comparison of 2017 ATLAS online integrated luminosity with that of consecutive years is shown in Figure 2.2. The integrated luminosity recorded by ATLAS during 2017 reached a new record at the time, which is around 50 $\text{fb}^{-1}$. This means that the instantaneous luminosity in 2017 is significantly higher than previous years. With higher instantaneous luminosity, the amount of pile-up interactions also increases since the cross section of inelastic collisions is significantly higher than hard scatter collisions as shown in Figure 2.3. A higher dose of pile-up interactions challenges our ability to differentiate hard scatter signals from pile-up ones.

The number of interactions per bunch crossing is a Poisson-distributed quantity with a mean value denoted by the variable $\mu$. This value changes as the beam intensity varies and the number of interactions is a convolution over Poisson distributions characterized by an average mean denoted by $\langle\mu\rangle$. This average mean $\langle\mu\rangle$ for the 2017 ATLAS run is shown in Figure 2.4. The number of reconstructed primary vertices is shown in Figure 2.5 with the reconstruction efficiency being around 60%.

In Chapter 6 where $E_T^{\text{miss}}$ determination is investigated with respect to different pile-up conditions, the number of reconstructed primary vertices $N_{pv}$ is used to reflect the level of pile-up activity.

An easy way to investigate the missing transverse momentum in the event is to sum all energy deposits in the transverse plane. Theoretically, this is the same as summing up energies only from hard scatter interaction since pile-up interactions barely produce any real missing momentum. However, due to resolution issues, a minor deviation from zero $E_T^{\text{miss}}$ is present at each pile-up vertex. Therefore this method is less effective when the total number of vertices is high. It will be more effective to select only hard scatter objects than summing up all the energy deposits.

Charged pile-up objects can be effectively identified by exploiting their vertex-associations. By imposing the track selection cut discussed in Chapter 4.2.1, one can easily pick out all tracks from the hard scatter vertex. However, since the trajectory can only be determined within a limited resolution, the resolution is still degraded due to the presence of pile-up. For charged particles outside the tracker's acceptance region and neutral signals which could not be detected by the tracking system, pile-up discrimination is significantly harder to achieve.

## 2.3 $E_T^{\text{miss}}$ proxy

### 2.3.1 Motivation

Since detectors can not capture neutrinos and other invisible particles, results from MET algorithms cannot be compared and fine-tuned without the knowledge of the actual momentum of invisible particles. One way to deal with this is to test algorithms on Monte Carlo simulations in which the momentum of non-interacting particles can be accessed. However, if one wants to study $E_T^{\text{miss}}$ using data obtained from the ATLAS detector, it is desirable to come with a proxy for non-interacting particles based on a well understood physics process.

The idea is to treat an object that is visible to the detector as invisible particles and purposely ignore the associated energy deposits. Neglecting existing signals would surely make the sum of momentum be non-zero if there is no other source of $E_T^{\text{miss}}$ in that event. Then the $E_T^{\text{miss}}$ computed based on this treatment could be compared to the measured momentum values of the neglected particle to test the effectiveness of the MET algorithms.

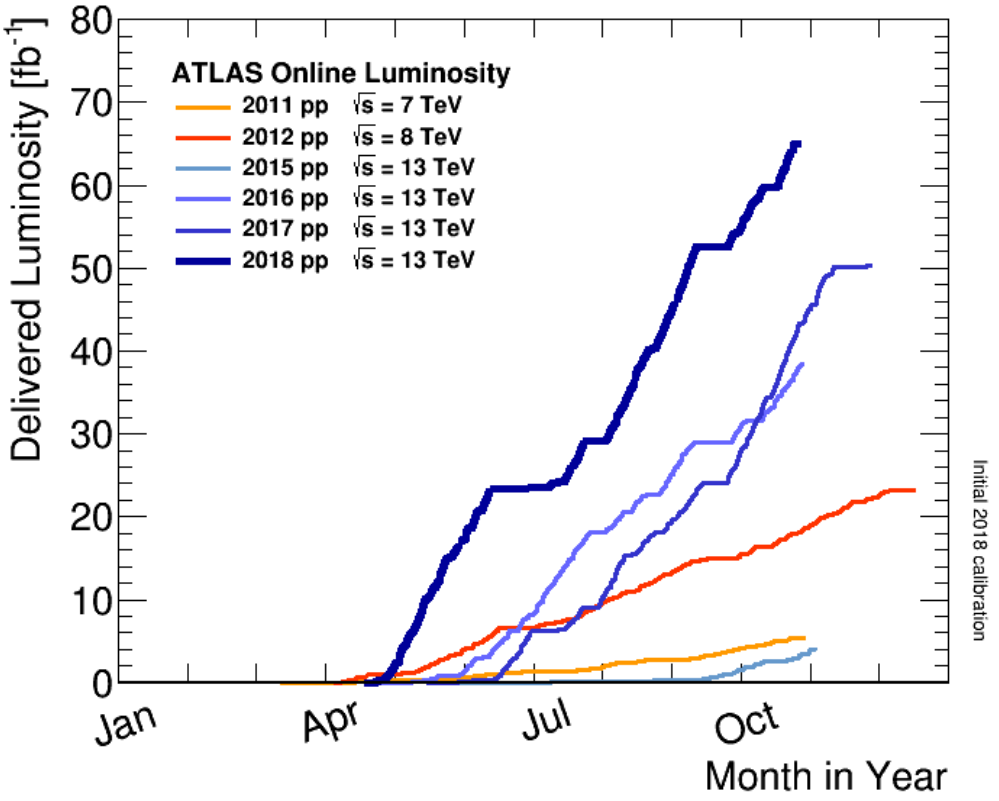


Figure 2.2: Cumulative luminosity recorded by the ATLAS detector[8].

A suitable process that can be used as an imaginary $E_T^{\text{miss}}$ source is the decay of the Z boson into two muons ($Z^0 \rightarrow \mu^+\mu^-$). For most calorimeters, the muon is a **minimum ionizing particle** (MIP) which leaves only a small fraction of its energy in the calorimeter. The stopping power of positive muon in copper as a function of energy and momentum is shown in Figure 2.6. We see that the stopping power barely changes within the momentum range of 100 MeV/c to 100 GeV/c, which is a typical range of muon momentum produced at the LHC.

A di-muon decay is then the perfect imaginary source of missing transverse momentum. The measurement of muon momentum is based on the combined tracking system and the muon spectrometer. It is not necessary to perform any modification on calorimeter signals since muons do not leave much energy in the calorimeter. After all recorded objects are reconstructed, one can reconstruct the Z boson by calculating the invariant mass of the di-muon system. Then the Z boson would serve as an imaginary $E_T^{\text{miss}}$ after neglecting signals from the muon spectrometer.

When algorithms are being tested, the estimated $E_T^{\text{miss}}$ can be compared with the

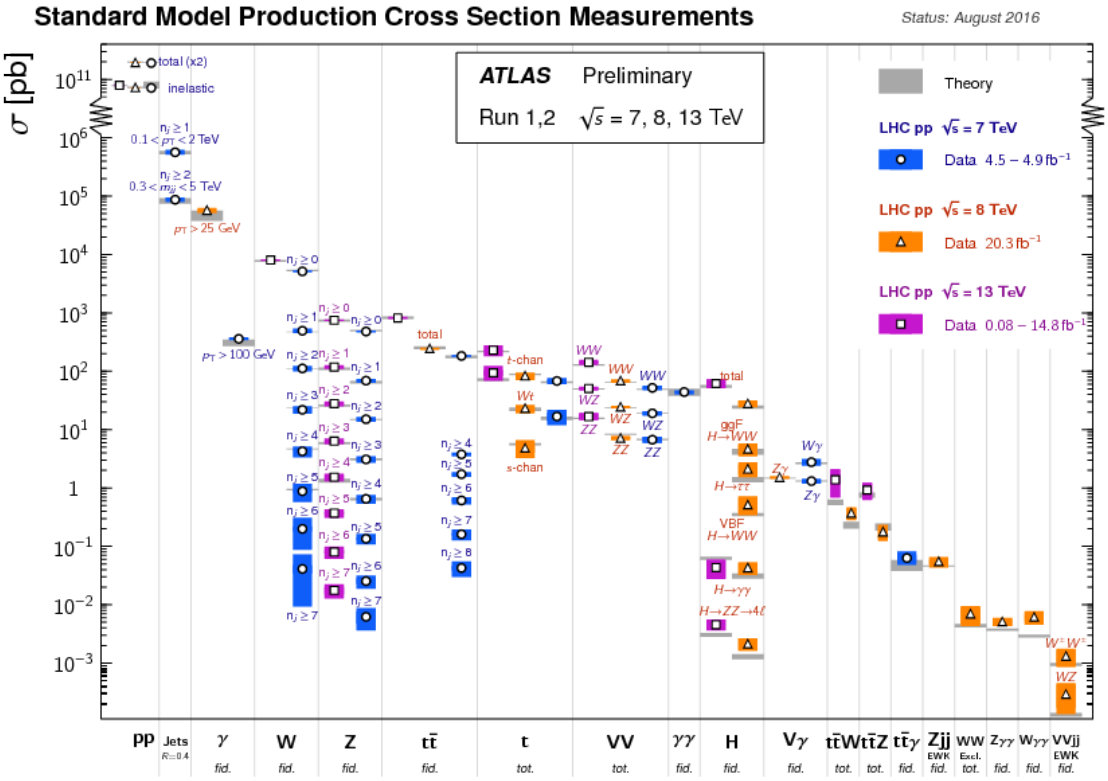


Figure 2.3: Comparison of theoretical and experimentally determined cross-sections.[8].

transverse momentum of the reconstructed Z boson. It also allows us to compare the estimation of the $E_T^{\text{miss}}$ scale with the magnitude of the Z boson's momentum.

### 2.3.2 $Z^0 \to \mu^+\mu^-$ selection

By treating the Z boson as an imaginary $E_T^{\text{miss}}$ source, one could easily test the resolution and scale of the PUfit algorithm in data. Monte Carlo samples are also used to test the precision of the Z boson momentum reconstruction. Datasets used are listed below.

- **Data**:

  **Center of mass energy**: 13 TeV

  **Total number of Events in the dataset**: 4238896

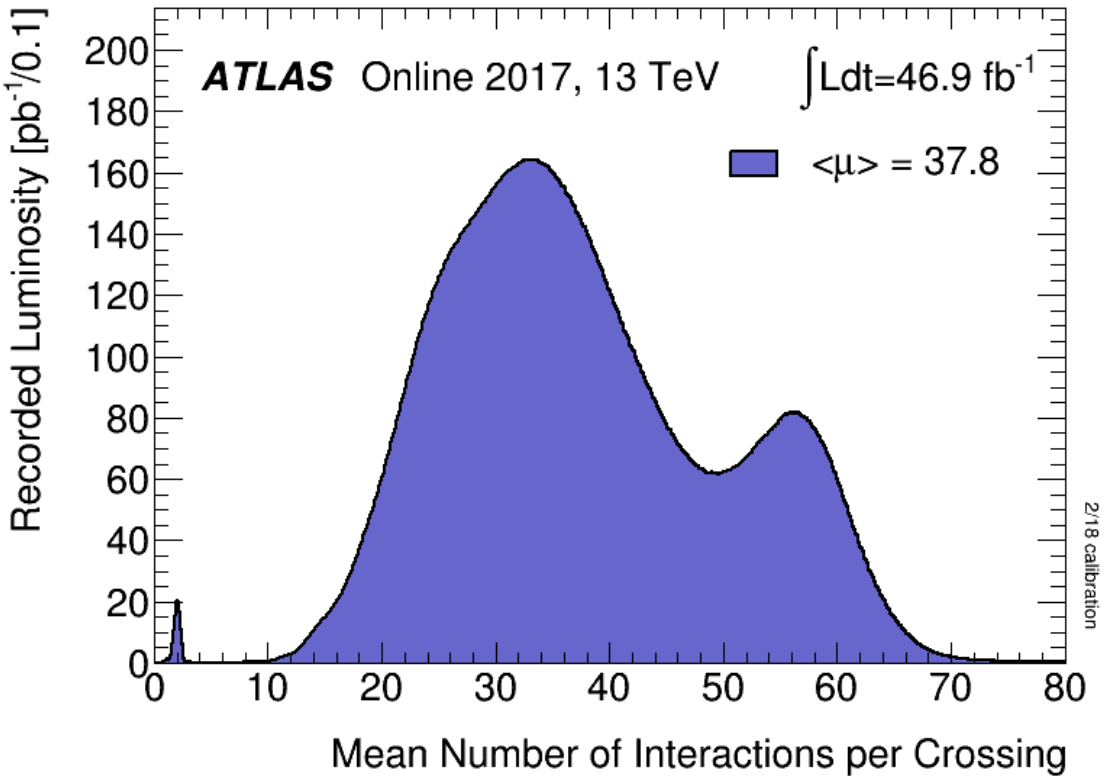


Figure 2.4: Mean number of interactions over the 2017 ATLAS run [8].

**Average expected number of interactions per beam crossing** ($\mu$): 20.9 - 58.4

- **Monte Carlo**:

  **Center of mass energy**: 13 TeV

  **Total number of Events in the dataset**: 436000

  **Pile-up condition:** same as LHC running conditions

The two muons produced by the Z boson need to have opposite electric charge due to the conservation of charge in the decay process. The leading muon is required to have a $p_T$ greater than 25 GeV, while the requirement on the sub-leading muon is $p_T > 20$ GeV. The identification of muon is mainly achieved by applying quality requirements which assess how likely the object is a muon based on variables such as the $\chi^2$ value of the fit and the $p_T$ difference between inner detector and muon spectrometer measurements[10]. The isolation criterion requires the muon to be isolated

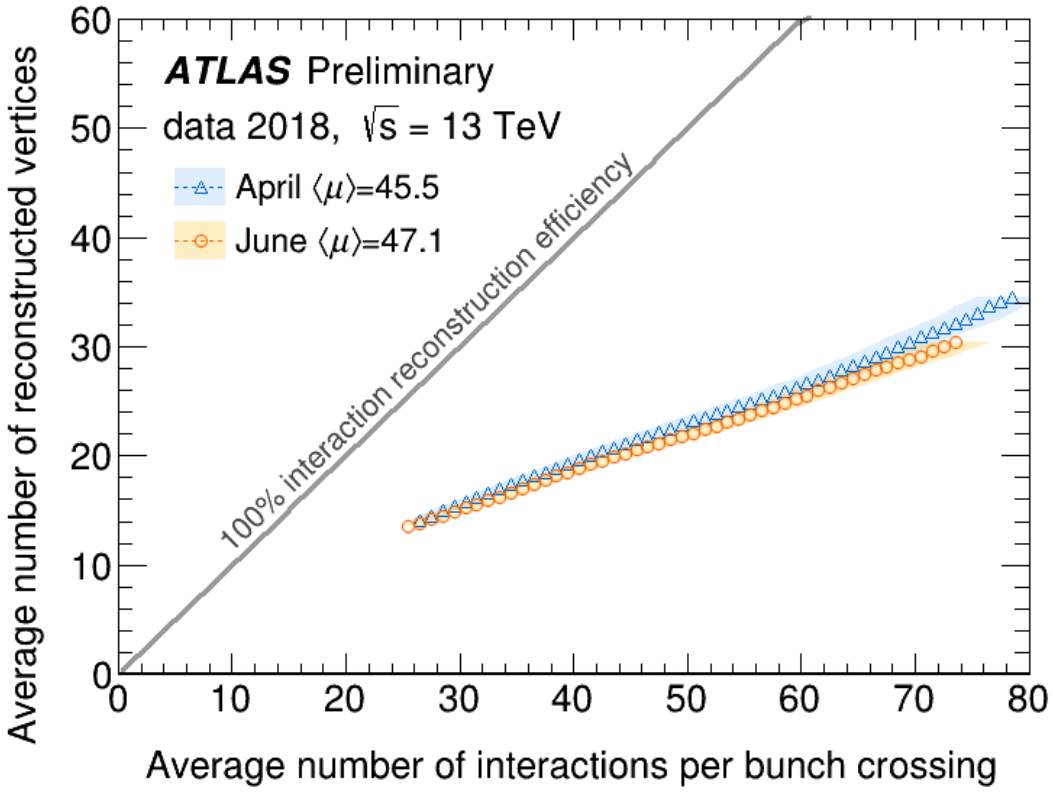


Figure 2.5: Number of reconstructed vertices as a function of the average number of interactions per bunch crossing.[9]

from other objects such as jets. Both muons are required to be a medium quality muon within $|\eta| < 2.5$ while passing isolation requirements[10].

From the two muons found in the events, one can construct the invariant mass of the Z-boson by using their four momenta. First the momentum $p_i^Z$ $(i = x, y, z)$ of the the Z boson along axis-$i$ is given by:

$$p_i^Z = p_i^{\mu_{\text{lead}}} + p_i^{\mu_{\text{sub}}} \tag{2.7}$$

where $p_i^{\mu_{\text{lead}}}$ and $p_i^{\mu_{\text{sub}}}$ are momenta of the leading and sub-leading muons along axis-$i$. The momentum $p_T^\mu$ is defined to be the transverse momentum of the muon. Then the energy and invariant mass of the Z boson are given by:

$$\begin{aligned} E_Z &= E_\mu^{lead} + E_\mu^{sub} \\ m_Z &= \sqrt{(E_Z)^2 - \sum_i^3 (p_i^Z)^2} \end{aligned} \tag{2.8}$$

In order for an event to be considered as a $Z^0 \to \mu^+\mu^-$ event, the invariant mass

calculated must fall within the Z-mass window given below:

- $p_T^{\mu_{lead}} > 25\text{GeV}$
- $p_T^{\mu_{sub}} > 20\text{GeV}$
- Charge ($\mu_{lead}$) $\cdot$ Charge ($\mu_{sub}$)$=-1$
- Medium quality of both muons [10].
- Muons must be isolated from other objects according to gradient isolation criterion in Ref.[10].
- $|m_Z - 91.1876\text{ GeV}| \leq 25\text{GeV}$

Selections listed above would select $Z^0 \rightarrow \mu^+\mu^-$ events, but it is convenient to impose another selection: $N_{jets} > 0$. It will be discussed in Chapter 5 that the current version of PUfit only makes a correction when there exists hard scatter jets. Therefore it is efficient to test the algorithm in those events in which corrections are non-zero. The distribution of jet multiplicity and the momentum of the Z boson in $Z^0 \rightarrow \mu^+\mu^-$ events are shown in Figures 2.7 and 2.8 respectively.

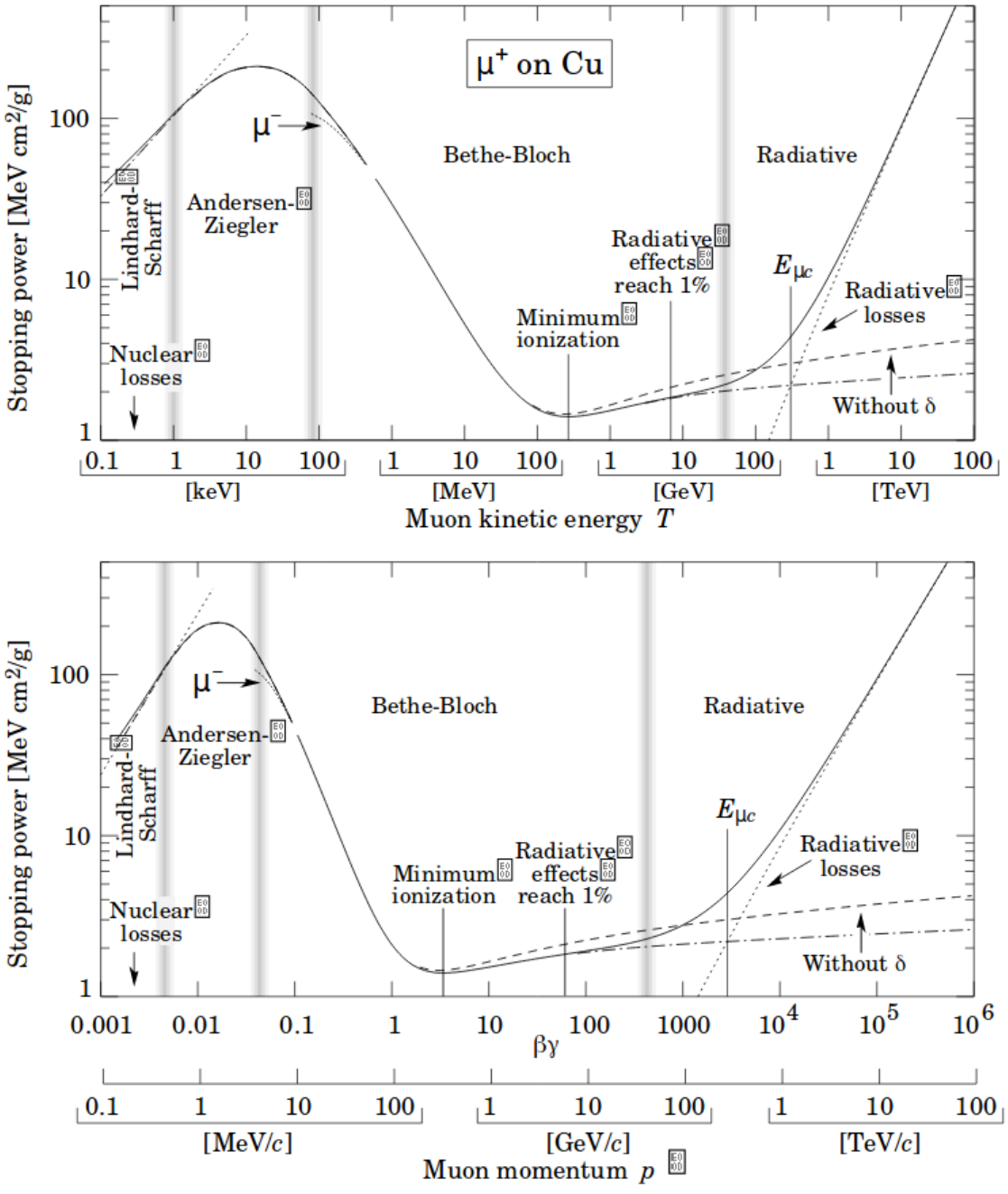


Figure 2.6: The stopping power of a positively charged muon in copper plotted against its kinetic energy and momentum (velocity). It is evident that there exist a plateau region in which the muon experience minimum ionization [11].

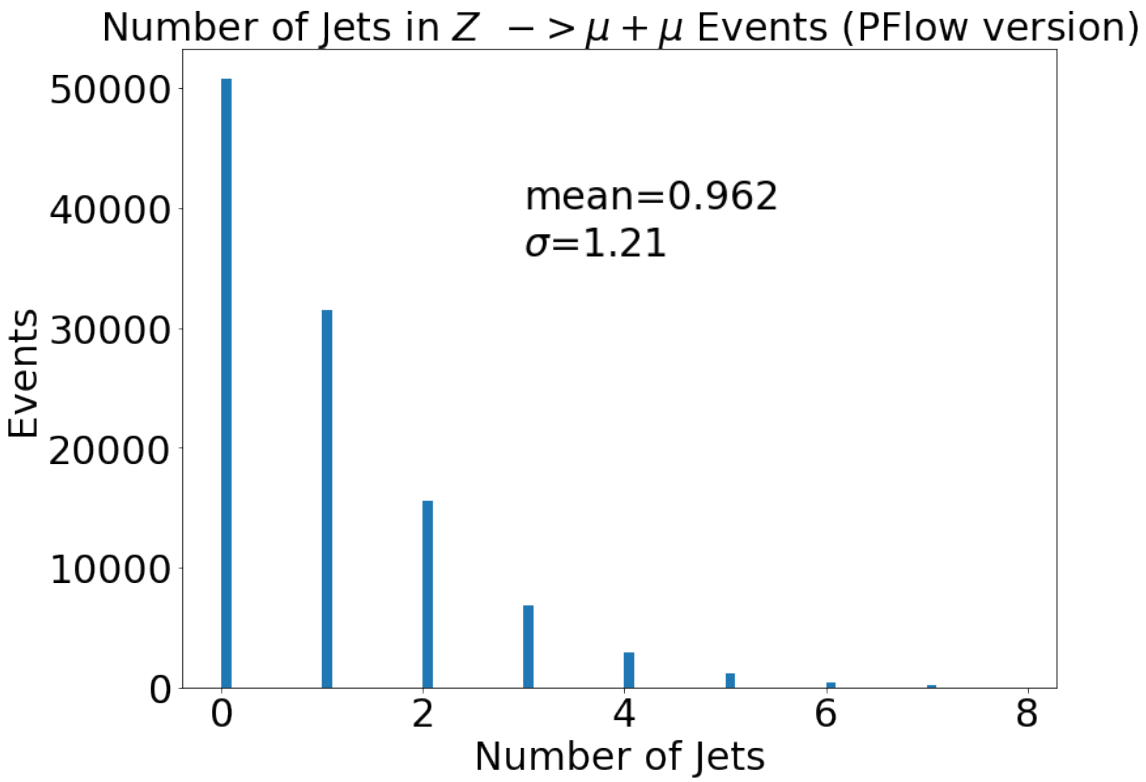


Figure 2.7: Number of jets in $Z^0 \rightarrow \mu^+\mu^-$ Events.

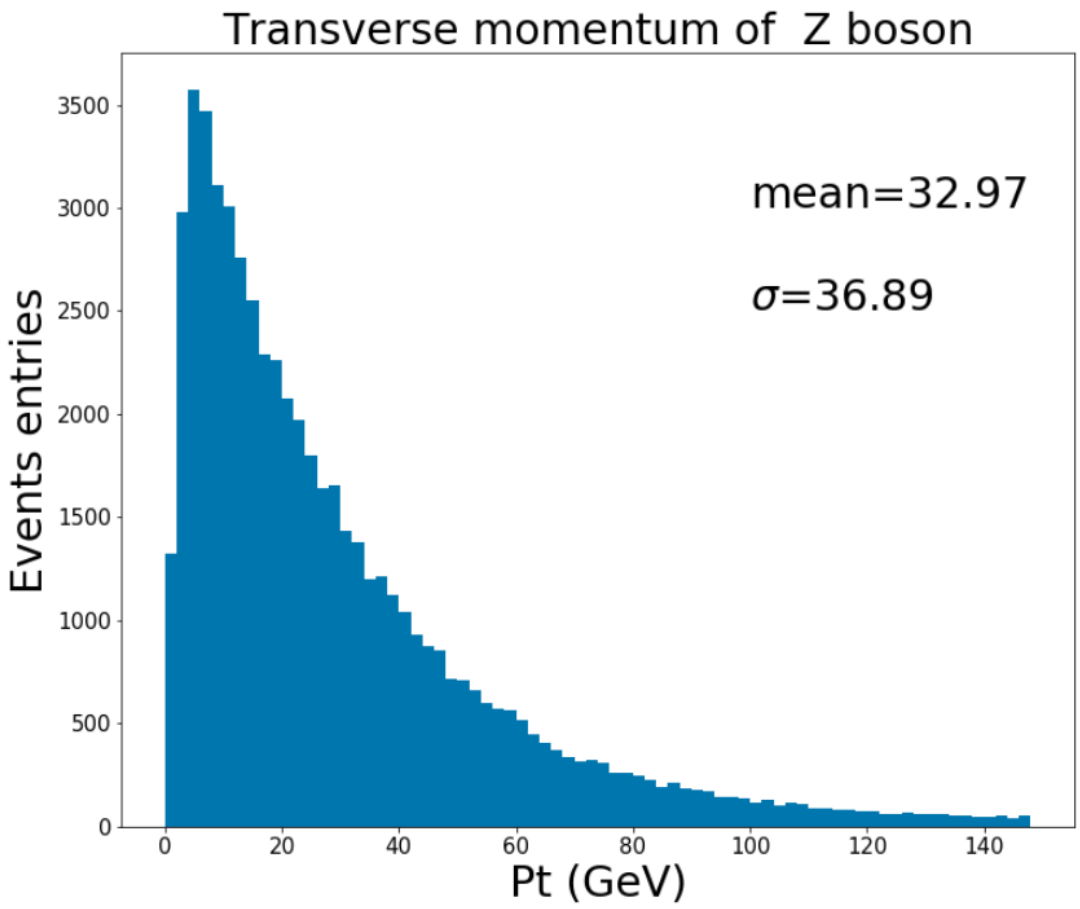


Figure 2.8: Distribution of transverse momentum of Z Boson.

# Chapter 3

# The ATLAS detector

## 3.1 LHC

The **Large Hadron Collider** (LHC) is the world's largest and most powerful accelerator for particle physics experiments. It boosts protons to travel in a 27 kilometers ring at a speed near the speed of light. After the upgrade in 2015, the LHC began its Run 2 phase and achieved an unprecedented center of mass energy of 13 TeV. The interior of the beam tube is kept at ultrahigh vacuum, allowing high energy particles to travel without undesired interactions. The two beams of protons are set to travel in opposite directions, and collide at four interaction locations along the ring to be measured by four different detectors. The ATLAS detector is one of the four detectors that probes interesting physics processes both within and beyond the standard model.

The proton beam in the LHC is made of groups of protons called **bunches**. There are $1.15 \cdot 10^{11}$ protons in each bunch and 2556 bunches in each beam. Inside the collider, proton bunches are guided by a magnetic field produced by superconducting electromagnets, enforcing protons to move in the pre-set trajectory. Bunches are collided every 25 ns to produce collisions. The action of colliding two bunches of protons is referred to as bunch crossing. Bunches of particles will then collide in four interactions points, each equipped with a detector to measure different physics interactions.

## 3.2 The ATLAS detector

### 3.2.1 Overview

The ATLAS detector is cylindrical detector with forward-backward symmetry. It offers nearly full coverage so that most particles produced would enter the detector regardless of their direction of motion[5]. It was designed to test new physics models while offering high precision measurements of QCD, electro-weak and other physical processes. In order to obtain high momentum resolution for physical analysis while focusing on rare events from interesting physical processes, the ATLAS detector is designed to meet the following requirements as described in Reference [12]:

- Due to the experimental conditions at the LHC, the detectors require fast, radiation-hard electronics and sensor elements. In addition, high detector granularity is needed to handle the particle fluxes and to reduce the influence of overlapping events.

- Large acceptance in pseudorapidity with almost full azimuthal angle coverage is required.

- Good charged-particle momentum resolution and reconstruction efficiency in the inner tracker is essential. For offline tagging of $\tau$-leptons and b-jets, vertex detectors close to the interaction region are required to observe secondary vertices.

- Very good electromagnetic (EM) calorimetry for electron and photon identification and measurements, complemented by full-coverage hadronic calorimetry for accurate jet and missing transverse energy measurements, are essential requirements, as these measurements form the basis of many of the studies mentioned above.

- Good muon identification and momentum resolution over a wide range of momenta and the ability to determine unambiguously the charge of high $p_T$ muons are fundamental requirements.

- Highly efficient triggering on low transverse-momentum objects with sufficient background rejection, is a prerequisite to achieve an acceptable trigger rate for most physics processes of interest.

A cut-away view of the ATLAS detector is shown in Figure 3.1. It is clearly shown that inner detector resides in the innermost part of the ATLAS detector, surrounded by calorimeters. The muon chamber is the outermost system with the wheels being 44 meters apart.

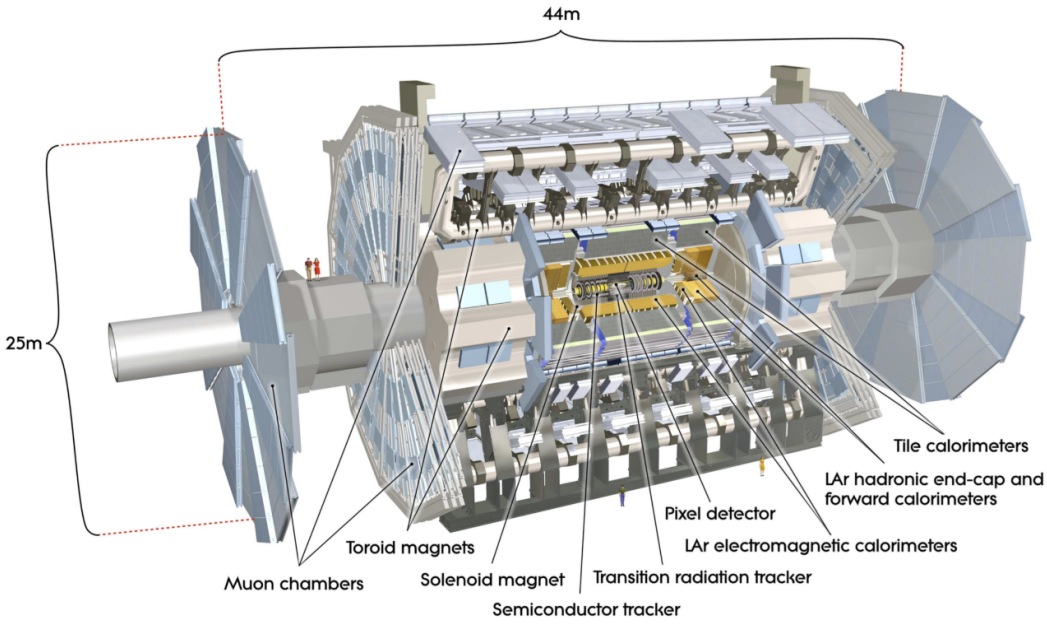


Figure 3.1: Cut-away view of the ATLAS detector[12].

### 3.2.2 Inner detector

After *pp* collisions take place inside the ATLAS detector, the inner detector is the first part of ATLAS that receives signals of particles produced in the collision. The inner detector is immersed in the strong magnetic field of 2T generated by the solenoid magnet[13]. Since electrically charged particles experience forces which bend their trajectories, the momentum and charge of the incoming charged particle can be determined by measuring the radius of curvature and also the direction of bending. By fitting the trajectory from detector signals, one can also determine the origin of the track along the beam axis. Using this information, it is then possible to associate tracks with vertices that produced them, offering better pile-up discrimination for charged particles. Details of offline selection and vertex association are given in Chapter 4. The transverse momentum resolution of the inner detector is designed to

be[12]:

$$\frac{\sigma_{p_T}}{p_T} = 0.05\% \times p_T(GeV) \oplus 1\% \tag{3.1}$$

There are three internal components of the inner detector: the silicon pixel detector that covers the radial range of 50.5 mm to 150 mm, the semiconductor tracker (SCT) that covers from 299 mm to 560 mm and finally the outermost transition radiation tracker (TRT) with a radial coverage of 563 mm to 1066 mm as shown in Figure 3.2.

The pixel detector is the innermost part of the inner detector, made of three barrel layers and three disks in each endcap. The barrel has 1744 modules with 46080 read-out channels per module while the three disks have 1456 modules. Each pixel is built to have an area of $50 \times 400 \mu\text{m}^2$

Semiconductor trackers are made of silicon that overall span 40 $m^2$ area. They are distributed over four cylindrical barrel layers and nine endcap disks on each side of the detector.

The transition radiation tracker (TRT) measures the transition radiation which is a form of radiation emitted when particles go through boundaries of two media with different refractive indices. The TRT is made of 350000 read-out channels, with a volume of $12m^3$. The detection is based on straw tubes, which are essentially drift tubes that measure electron ionizations produced by the traversing particle. Each straw tube has a diameter of 4 mm and contains a 0.03 mm diameter gold-plated tungsten wire in the center. There are 50000 straws in the barrel, each 144 cm long. Also, in endcaps, there are a total number of 250000 straws, each 39 cm long. The TRT also provides additional information on the type of the incoming particle, since particles of different types emit transition radiation with different characteristics.

Primary vertices are found by grouping reconstructed tracks. The parameters of vertex candidates are then determined in a fit using tracks associated with the vertex. The z resolution of the vertex, i.e, the resolution of position along the beam axis is shown in Figure 3.3. It is obvious that the resolution is much better for those vertices with large numbers of tracks.

### 3.2.3 Calorimeter

The calorimeter measures the energy of an incoming particle by absorbing the shower of secondary particles it produces as it interacts with the detector material. A cut-away view of the ATLAS calorimeter system is shown in Figure 3.4. The calorimeter

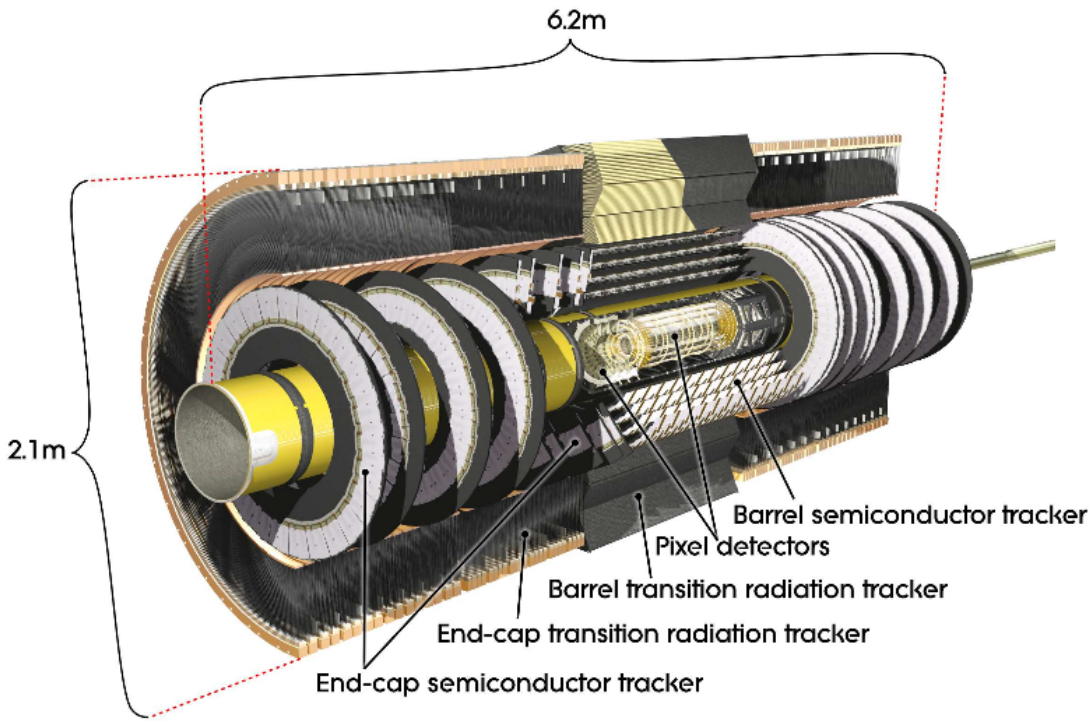


Figure 3.2: A cut-away view of the ATLAS inner detector [14].

system covers the range of $|\eta| < 4.9$ with different components being designed to measure showers produced by electro-magnetic or strong interaction processes [16].

The LAr electromagnetic calorimeter is the innermost detector in the entire calorimeter system. It is divided into the barrel ($|\eta| < 1.475$) and two end-caps ($1.375 < |\eta| < 3.2$). The barrel has a thickness of more than 22 radiation lengths so that the calorimeter would be able to contain high energy particle showers while preventing most particles from entering the muon system.

The barrel calorimeter is made of two half-barrels connected together with only a small gap of 4 mm at $z = 0$ [16]. There are two co-axial wheels in each end-cap calorimeter. The outer wheels cover the region $1.375 < |\eta| < 2.5$ while the inner ones cover $2.5 < |\eta| < 3.2$ [16].

The hadronic calorimeter consists of a tile barrel calorimeter, two LAr endcap calorimeters (HEC) and two LAr forward calorimeters. The tile calorimeter is placed directly outside the EM calorimeter and covers the range of $|\eta| < 1.0$. The two extended barrels cover the range of $0.8 < |\eta| < 1.7$. These are sampling calorimeters

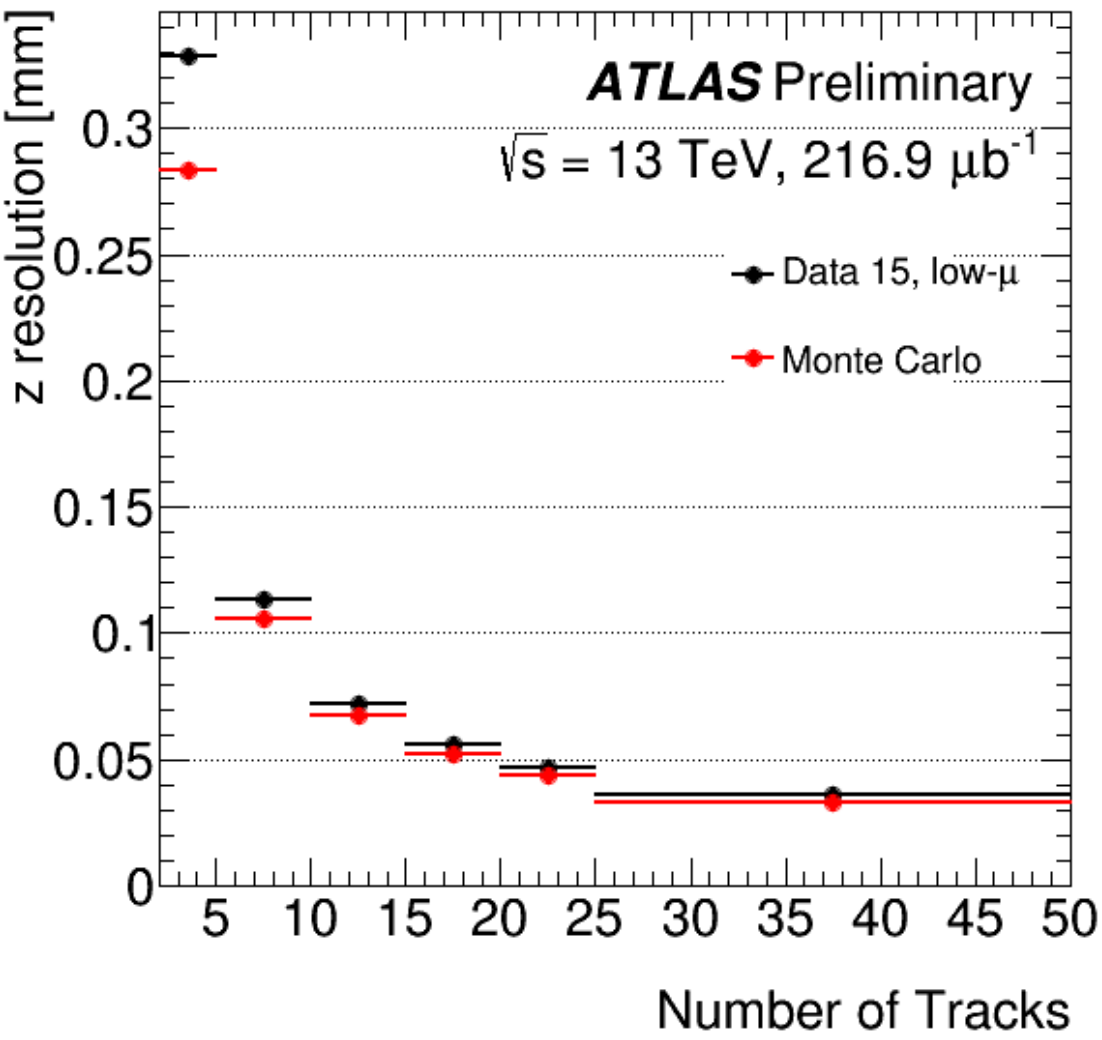


Figure 3.3: z resolution of vertex reconstruction [15]

with steel being the absorber [16]. The HEC covers up to $|\eta| = 3.2$, and it is located directly behind the electromagnetic endcap calorimeters. The forward calorimeter covers the high $\eta$ region and reduce the background in the muon system.

### 3.2.4 Topoclustering algorithm

In order to select desired significant signals, topologically connected calorimeter cells are grouped into clusters called **topoclusters**. Topoclusters are constructed in areas where the energy deposit is significantly higher than the expected noise fluctuation.

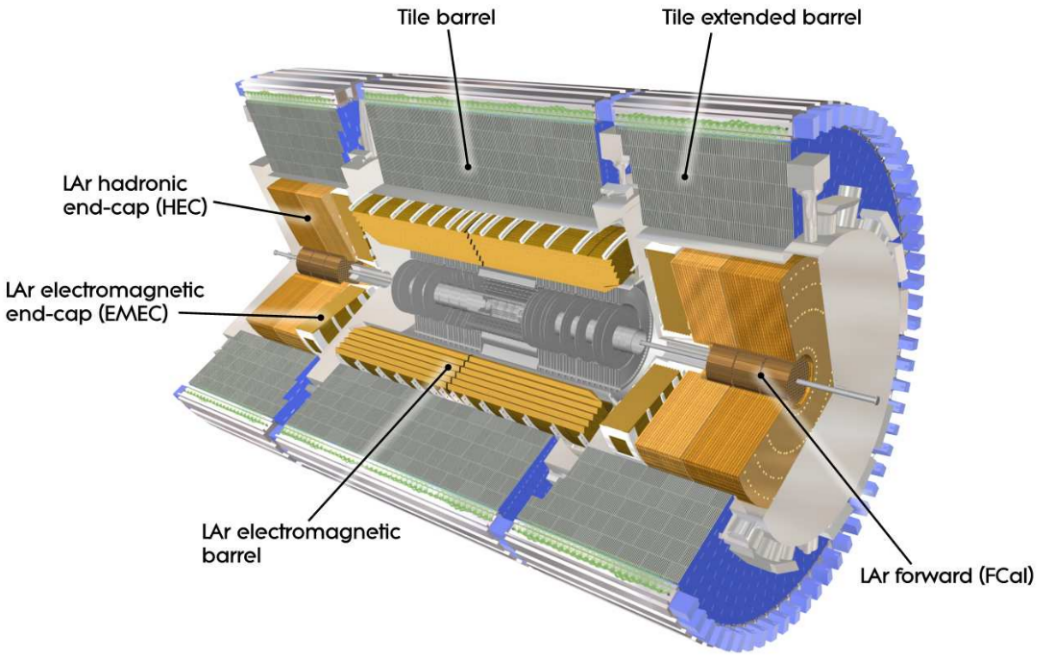


Figure 3.4: A cut-away view of the ATLAS calorimeter system [16].

However, the energy of a topocluster could consist of energy deposited by different particles from multiple vertices[17]. To construct topoclusters, the crucial variable to look at is the cell signal significance $\zeta_{cell}^{EM}$ which is defined to be the ratio of the signal $E_{signal}^{EM}$ in a given cell to the average noise $\sigma_{noise,cell}^{EM}$. The significance is given by:

$$\zeta_{cell}^{EM} = \frac{E_{signal}^{EM}}{\sigma_{noise,cell}^{EM}} \tag{3.2}$$

The process of building topoclusters is then described by the following procedures:

1. First, seed cells are determined by placing a cut of $|\zeta_{cell}^{EM}| > 4$. Each seed cell is called a proto-cluster.

2. Then neighbouring cells of a given proto-cluster are merged into the proto-cluster if they satisfy $|\zeta_{cell}^{EM}| > 2$

Since the absolute value $|\zeta_{cell}^{EM}|$ is used in the procedure stated above, cells with negative signals are also included in the construction of topoclusters. It is shown in Figure 3.5 that the signal development takes much longer than the 25 ns interval between consecutive bunch crossings. It takes around 450 ns for the signal to decrease

linearly to 0. In order to shorten the signal development time, the pulse is shaped to have a sharp peak and then quickly decreasing to negative values so that the integral of amplitude over time is 0. The negative tail shown in Figure 3.5 balances the residual charges that are still being collected. However, this still has impact on consecutive bunch crossings, and the residual energy is referred to as out-of-time pile-up. Negative signal cells are usually products of out-of-time pile-up and electronic noise. Due to the existence of noise fluctuations in cell energies, the positive noise brought by the cells could form a bias in the overall distribution of cell energies. The bias introduced by only using positive cells can be partially suppressed by allowing negative cells to enter the algorithm[17].

There are two different energy scales for topocluster, the **EM-scale** and the **LCW** (local hadronic cell weighting) scale. The scale is determined by the LCW calibration which corrects energy deposits based on the topocluster shape: either hadronic (HAD) or electromagnetic (EM). The energy used in the formation procedure above ($E^{EM}_{signal}$) is in the EM-scale.

### 3.2.5 Muon Spectrometer

The muon spectrometer is the outermost detection system in the ATLAS detector. It is based on the magnetic deflection of muon tracks in the large superconducting air-core toroid magnets, instrumented with separate trigger and high-precision tracking chambers.[16]

Muon tracks are bent by the magnetic field supplied by the large barrel toroid over the range $|\eta| < 1.4$. Two endcap magnets are installed at the end of the barrel toroid to provide magnetic bending for $|\eta|$ between 1.6 and 2.7.

As shown in Figure 3.6, monitored drifting tubes (MDT) are implemented in the barrel and endcap region up to $|\eta| = 2.7$ to offer precision tracking. For large $\eta$ regions where $2.0 < |\eta| < 2.7$, cathode strip chambers are also used for precision tracking. Triggering is also separated into two regions. For $|\eta| < 1.05$, resistive plate chambers are used for triggering while thin gap chambers are used for regions of $1.05 < |\eta| < 2.7$.

### 3.2.6 Trigger

Only a small subset of the proton collisions produced by LHC correspond to interesting physics processes. Since the computing power and the bandwidth for digitizing

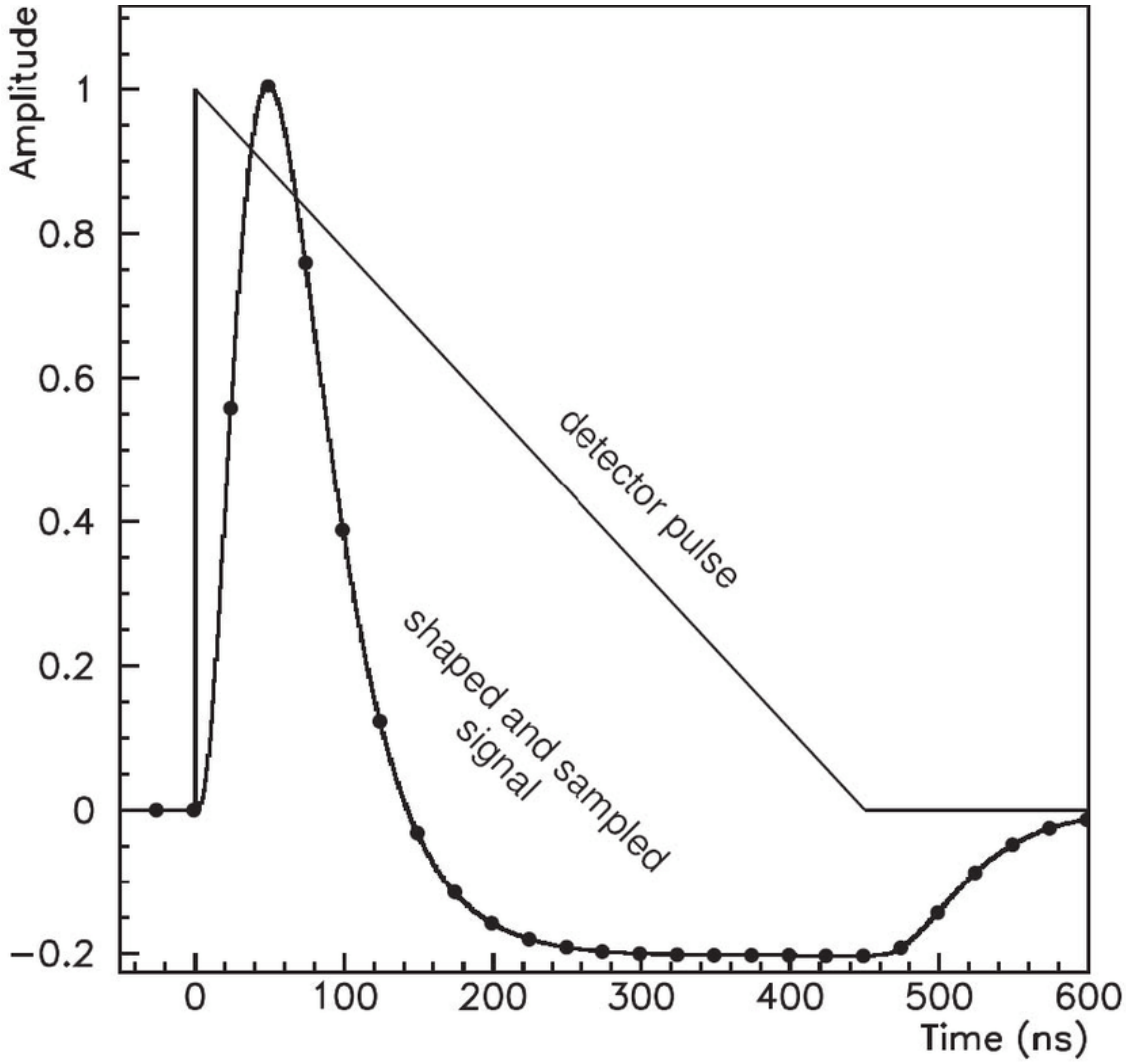


Figure 3.5: Shapes of the LAr calorimeter pulse[18].

are limited, it is vital to select the events of interest for the analysis instead of using all produced events. The decision of whether to keep an event is made by the two-level trigger system which drops the event rate from 40 MHz to 1 kHz[16].

The level-1 (L1) trigger determines the Regions-of-Interests (RoIs) using information from the calorimeter and the muon spectrometer[19]. The decision is made within 25 $\mu s$, and the event rate is dropped from 40 MHz to 100 kHz. If a event is selected by the L1 trigger, it will be subsequently sent to the high level trigger (HLT) to be analyzed by a more sophisticated trigger algorithm. The HLT further reduces the event rate from 100 kHz to about 1 kHz. The decision of event selection in the HLT is made within 200 ms[19]. Events selected by the HLT are then analyzed by

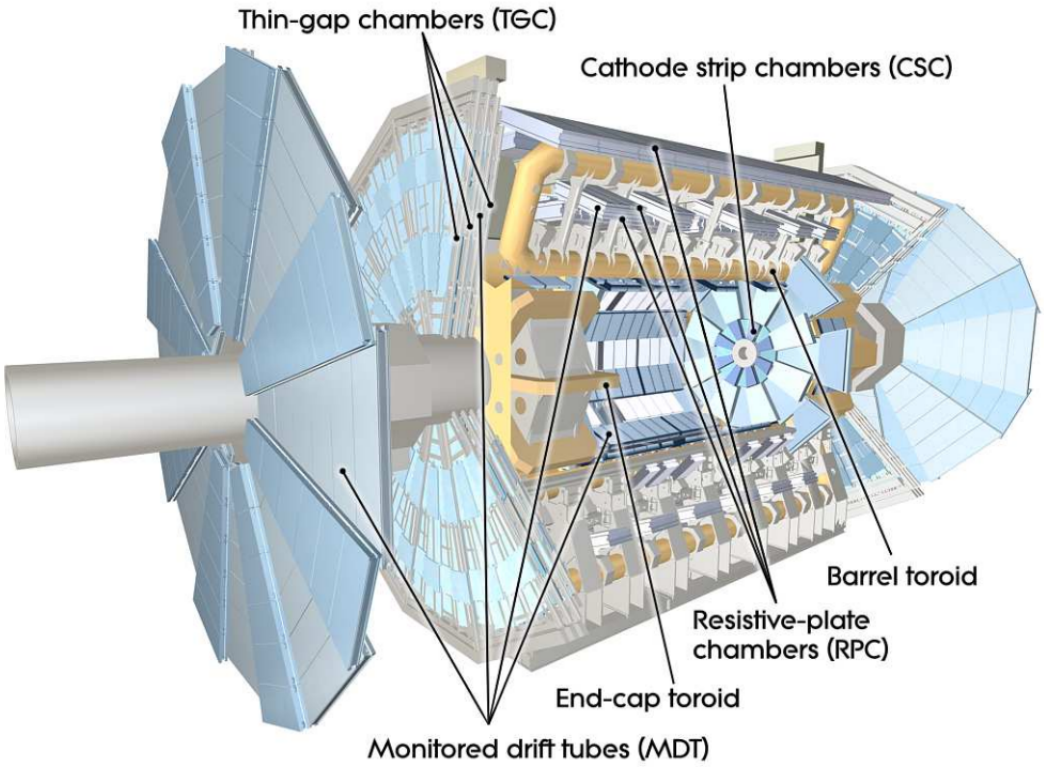


Figure 3.6: A cut-away view of the muon system. [16].

the offline algorithm for physics objects reconstruction.

# Chapter 4

# Offline MET reconstruction

Particles of different types are reconstructed based on signals recorded by a series of detectors listed in Chapter 3. Hard scatter reconstructed objects are selected by imposing momentum cuts, isolation requirements, and other criteria. The final $E_T^{\text{miss}}$ is determined by the offline MetMaker algorithm which reconstructs $E_T^{\text{miss}}$ based on hard scatter objects and a soft term described below.

## 4.1 Object selection

As shown in Equation 2.3, the inputs to the $E_T^{\text{miss}}$ calculation are hard scatter objects and soft signals. In typical event topologies used in this thesis ($Z^0 \rightarrow \mu^+\mu^-$ , $W \rightarrow \mu\nu_\mu$ , $t\bar{t}$ ), the main particles of interest are muons, electrons, jets and soft particles( $\pi^{\pm}$, $K^{\pm}$, $p$, $n$ and photons). It is fairly rare to find other particles in these samples, and they do not contribute significantly to the $E_T^{\text{miss}}$ .

### 4.1.1 Muon

Signals from the muon spectrometer and the inner detector are used to reconstruct muons within $|\eta| < 2.5$. Muons that are outside the tracker acceptance region but within $|\eta| < 2.7$ are reconstructed by fits using information from the muon spectrometer [5]. Apart from what is listed above, muon candidates also have to pass a $p_T$ threshold of 10 GeV while having a medium reconstruction quality. For those muon candidates within the acceptance region of the inner detector, their tracks must be associated with the hard scatter vertex. The track association will be discussed in Section 4.2.1.

### 4.1.2 Electron

Reconstructed electrons are selected based on their calorimeter shower shapes and the matching between their clusters and tracks[20]. Electrons need to pass the medium reconstruction quality selection to be considered as candidates. The calibration of electrons is described in Ref.[20]. All electron candidates passing medium quality selection are required to have $p_T > 10$ GeV. In the transition region between central and endcap calorimeters($1.37 < |\eta| < 1.52$), the resolution is too poor for the electron reconstruction. In order to avoid the transition region, electrons are additionally required to have $|\eta| < 1.37$ or $1.52 < |\eta| < 2.47$. Energy deposits by electrons within the range of $1.37 < |\eta| < 1.52$ could be picked up by the jet finding algorithm and thus included in the jet collection if it meets the selection criteria mentioned in Section 4.1.3 [5].

### 4.1.3 Jet

Jets are collimated sets of color neutral hadrons arising from high momentum quarks or gluons. Jets are reconstructed from electromagnetic scale topoclusters which are discussed in Ref.[17]. To form jets, a radius parameter of 0.4 is used in the anti-kt algorithm [21]. The EM-scale jets would then be calibrated according to steps discussed in Ref.[22]. A schematic diagram of EM-scale jet calibration is shown in Figure 4.1.

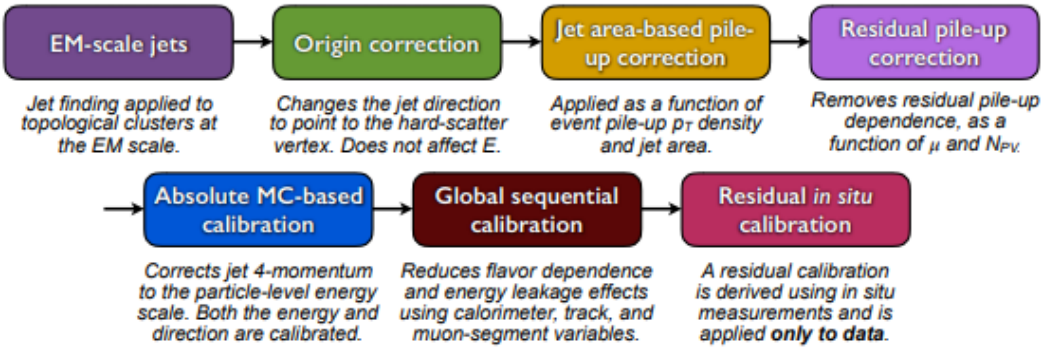


Figure 4.1: Calibration stages for EM-scale jets [22].

Calibrated jets will be considered for further selection involving kinematic thresholds and the Jet Vertex Tagger (JVT). JVT takes all charged tracks used in the jet

reconstruction and uses their vertex associations to output a number ranging from 0 (pile-up like) to 1 (hard-scatter like). The JVT algorithm is discussed in Ref.[23]

The selection of hard scatter jets is based on requirements on $p_T$ and JVT that vary as a function of $\eta$. Jets that meet any of the following criteria are selected.

- $2.4 < |\eta| < 4.5$

  $p_T > 30$ GeV

- $|\eta| < 2.4$

  $p_T > 50$ GeV

- $|\eta| < 2.4$

  20 GeV $< p_T < 50$ GeV

  JVT $> 0.64$

Apart from the selection criteria listed above, there are also selections based on reconstruction quality and leptonic overlap. These technical details are discussed in Ref.[5].

## 4.2 Soft term

There are two main types of soft terms to be used to reconstruct the $E_T^{\text{miss}}$ according to Equation 2.3. The first is called the **track soft term**(TST) which is based on selected hard scatter tracks. TST is the standard soft term to be used in offline algorithms. Alternatively, the **cluster soft term**(CST), a soft term based on topoclusters, could be used instead to achieve better resolution in some cases.

### 4.2.1 Track and vertex selection

Track selection is not only crucial to the TST but also a vital part of hard object reconstruction discussed in the last section. To correctly use tracking information, one must determine whether the track is well reconstructed while associating the track with a vertex in the event. Vertex association allows us to focus on tracks that belong to the hard scatter vertex and thus enhance the performance of $E_T^{\text{miss}}$ reconstruction.

First of all, all tracks and vertices must pass quality requirements to be considered as inputs to reconstruction. Tracks need to have their $p_T$ to be greater than 400

MeV and $|\eta| < 2.5$ while passing quality requirements discussed in Ref.[24]. The **transverse impact parameter**($d_0$) is defined to be the shortest distance from the trajectory to the nearby vertex in the transverse plane and the corresponding $z$ value is the longitudinal impact parameter $z_0$ . Each vertex candidate must have at least two tracks that satisfy $|d_0| < 1.5$mm and $|z_0 sin(\theta)| < 1.5$mm to that candidate[5]. Then by applying the impact parameter cuts mentioned above, one can easily match tracks with the hard scatter vertex which is defined in Chapter 2.1.

### 4.2.2 Track soft term

The definition of the track soft term(TST) is intuitively easy to understand. It merely uses all tracks associated with the hard scatter vertex but not used in the reconstruction of any hard object. In order to select unused tracks to avoid double counting, tracks that meet the following criteria are removed as described in Ref.[5], where $\Delta$R is defined as the distance in the $\eta-\phi$ plane at the surface of the calorimeter: $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ .

- ID tracks with $\Delta R$(track,electron/photon cluster)$< 0.05$;
- ID tracks with $\Delta R$ (track,$\tau$-lepton) $< 0.2$;
- ID tracks associated with muons;
- ID tracks ghost-associated with jets. Ghost association is a technique that determines whether a track belongs to a jet[22].

Hard scatter tracks that are associated with jets rejected due to their overlap with other physics objects are used in the calculation of the soft term ($E_T^{miss,soft}$) [5]. Moreover, tracks from jets that are likely to come from pile-up, as tagged by the JVT procedure which assesses how pile-up like an object is, are included in the soft term as well [5].

The TST is insensitive to pile-up since it only uses tracks associated with the hard scatter vertex as inputs. However, since tracks do not measure neutral energy deposit, soft neutral particles from the hard scatter vertex are neglected. Also, charged particles with $|\eta| > 2.4$ are outside the tracker's acceptance region and thus not included in the TST. The loss of resolution from ignoring neutral particles from the HS vertex is compensated by the gain in excluding particles from pileup interactions once the number of simultaneous collisions $\mu$ is around 15 or higher.

### 4.2.3 Cluster soft term

The cluster soft term (CST) is sometimes also referred to as the calorimeter soft term. It uses topoclusters recorded in the calorimeter to provide a soft term that includes signals of neutral particles.

The CST is reconstructed by using topoclusters with positive energy in the LCW scale. Clusters in hard objects are not considered in the CST. Topoclusters formed at the same locations of hard objects would not contribute to CST even if those signals are not directly used in the reconstruction of the hard objects.

The CST provides neutral signals from the hard scatter vertex that are neglected by the TST. But at the same time, since it sums up clusters produced by all vertices, fluctuations in the measurement of the large amount of pileup energy used in the CST makes it sensitive to pile-up and it performs worse than TST when pile-up is high.

# Chapter 5

# PUfit algorithm

In this chapter, the PUfit algorithm is introduced with mathematical derivations. The constraints behind the algorithm will be formulated, and the solution to the fit will also be derived. The PUfit algorithm determines the MET by using our knowledge on the pile-up energy distribution. A simpler version of the PUfit algorithm was deployed as the default MET trigger in 2017 and 2018.

## 5.1 PUfit Theory

As shown in earlier chapters, both TST and CST have their own disadvantages as a soft term. The TST does not include neutral and large-$\eta$ signals, while the CST is sensitive to pile-up activities. In PUfit, a new soft term is introduced to capture large-$\eta$ and neutral soft signals without the severe deterioration due to pile-up increase. The new soft term is based on a $\chi^2$ fit that uses pile-up distribution constraints.

The $\vec{E}_T^{\text{miss}}$ can be determined by the negative vector sum of the transverse momentum ($\vec{p}_T$) of hard scatter (HS) objects. In order to avoid pileup (PU) contributions, we select jets and leptons that satisfy certain selection cuts on $p_T$, JVT, etc. Hard scatter objects are labeled in blue color as shown below while pile-up objects are labeled in red. Therefore contribution to $E_T^{\text{miss}}$ from selected HS objects is given by the following equation.

$$-\sum_{k}^{J} \vec{p}_{T_k} - \sum_{k}^{L} \vec{p}_{T_k}$$

where $J$ and $L$ are the number of selected jets and leptons. However, there are also soft HS objects with lower momenta. We account for these soft contributions by

introducing the Soft Term (ST). The ST has contributions from charged and neutral objects and is given by

$$\vec{E}_T^{\,\mathrm{ST}} = -\sum_k^C \vec{E}_{T_k} - \sum_k^N \vec{E}_{T_k} \tag{5.1}$$

where $C$ and $N$ are collections of charged and neutral HS signals that are not used in the reconstruction of any HS objects. With these definitions, the $\vec{E}_T^{\mathrm{miss}}$ associated with the HS interaction is

$$\vec{E}_T^{\mathrm{miss}} = -\sum_k^J \vec{E}_{T_k} - \sum_k^L \vec{E}_{T_k} + \vec{E}^{\,\mathrm{ST}} \tag{5.2}$$

### 5.1.1 TST/CST

Currently, we use tracks to determine the soft contribution from HS. We select tracks which

- do not belong to HS jets and leptons
- belong to the primary vertex

These tracks make up the Track Soft Term (TST)

$$\vec{E}^{\,\mathrm{TST}} = -\sum_k^{T(|\eta|<2.4)} \vec{E}_{T_k} \tag{5.3}$$

where $T$ is the number of HS tracks which do not belong to HS jets and leptons. As denoted above, this covers only $|\eta| < 2.4$ region, where we have the tracking system. The TST does not include neutral HS contributions and does not cover forward regions ($|\eta| > 2.4$).

Then the offline MET (TST) is defined as:

$$\begin{aligned} \vec{E}_T^{\mathrm{miss}} &= -\sum_k^J \vec{E}_{T_k} - \sum_k^L \vec{E}_{T_k} + \vec{E}^{\,\mathrm{TST}} \\ &= -\sum_k^J \vec{E}_{T_k} - \sum_k^L \vec{E}_{T_k} - \sum_k^{T(|\eta|<2.4)} \vec{E}_{T_k} \end{aligned} \tag{5.4}$$

Alternatively, the **cluster soft term** (CST) can be used. The CST is made of calorimeter clusters outside jet objects which includes both charged and neutral contribution to the $E_T^{\mathrm{miss}}$ . However, the resolution of CST $E_T^{\mathrm{miss}}$ deteriorates as the amount of pile-up becomes large. The CST is defined by the following equation:

$$\vec{E}^{\,\mathrm{CST}} = -\sum_{k}^{clus\notin J} \vec{E}_{T_k} \tag{5.5}$$

where $clus \notin J$ denotes the set of clusters outside jets.

### 5.1.2 The PUfit soft term(PST) and the Pileup-imbalance Adjustment Term (PAT)

The objects from pileup interactions are expected to be isotropic and to have no contribution to $\vec{E}_T^{\mathrm{miss}}$. However the object detection and reconstruction of the detector is not perfect and pileup interactions produce fake $\vec{E}_T^{\mathrm{miss}}$. This is what causes the cluster soft term (CST) to be sensitive to pileup. To adjust for the detector-response related imbalance of the PU contribution, we use a fit (PUfit) based on topoclusters. The PUfit algorithm computes the pileup-imbalance adjustment term (PAT) as a correction to the existing TST. Then a new soft term called PUfit Soft Term (PST) is given by the sum of the TST and PAT:

$$\vec{E}_T^{\mathrm{PST}} = \vec{E}_T^{\mathrm{TST}} + \vec{E}_T^{\mathrm{PAT}} \tag{5.6}$$

To construct the PAT, one must determine the characteristics of the pile-up distribution in the event. In particular, the PUfit algorithm requires the determination of expected pile-up energy under HS jets to construct the PAT. In order to calculate the average pile-up energy density in the calorimeter, the transverse energy $E_T$ of all topoclusters outside HS jets is summed into an $\eta - \phi$ grid of 112 equal-sized towers. Towers that overlapped with HS jets will not be used in the determination of the average pile-up density $\langle\rho\rangle$. Let the number of HS jets be denoted by $n_J$ and the number of towers overlapping with HS jets be $m$. The average PU energy density of towers is given by:

$$\langle\rho\rangle = \frac{1}{N-m}\sum_{j=1}^{N-m}\frac{E_{T_j}^{\mathrm{tower}} - p_{T_j}^{\mathrm{HS}}}{A_j} \tag{5.7}$$

where $E_{T_j}^{\mathrm{tower}}$ is the overall transverse energy deposit inside the tower, $p_{T_j}^{\mathrm{HS}}$ is the sum of hard scatter track $p_T$ in the given tower, and $A_j$ is the area of the $j$-th tower. Note that since we are only interested in the pile-up energy density, the HS energy must be removed. While we cannot distinguish hard scatter and PU activity for neutral clusters, we remove the HS activity for charged particles by subtracting the HS track

$p_T$ from the tower energy. In $Z^0 \to \mu^+\mu^-$ events, there are not many leptonic signals other than muons, and these muons do not leave much energy in the calorimeter. So the subtraction in Equation 5.7 only needs to consider HS soft tracks which are HS tracks that have not been used in the reconstruction of any HS object.

Then the average PU energy under the $k$-th HS jet can be approximated by $< \rho > A_k$, where $A_k$ is the area of the $k$-th jet. Since we are using jets with a radius of 0.4, the area $A_k$ is then taken to be $\pi \cdot 0.4^2$. The variance is given by

$$V_k^J = \frac{A_k^2}{N-m} \sum_{j=1}^{N-m} \left( \frac{E_{T_j}^{\text{tower}} - p_{T_j}^{\text{HS}} - \langle\rho\rangle A_j}{A_j} \right)^2 \tag{5.8}$$

The expected PU energy under the area of the $k$-th jet is expected to be $\langle\rho\rangle A_k$ if the PU transverse energy is isotropic and balanced. The pile-up energy is isotropic on average, but this is not necessarily true in a given event. Using a fit, we can determine the PU energy value for the $k$-th Jet, $\mathcal{E}_k$, needed to balance the overall PU contributions in an event. We can adjust for the pileup imbalance by taking the difference

$$-\sum_{k=1}^{n_J} (\mathcal{E}_k - \langle\rho\rangle A_k) \frac{\vec{E}_{T_k}^{\text{Jet}}}{E_{T_k}^{\text{Jet}}} \tag{5.9}$$

This is the fake PU contribution to $\vec{E}_T^{\text{miss}}$ due to detector effects. Since this can not be accounted for by any HS jets, leptons or the TST, it needs to be subtracted from the final $\vec{E}_T^{\text{miss}}$ as an additional term to reduce fake contributions from PU. Then, the Pileup-imbalance Adjustment Term (PAT), $E_T^{\text{PAT}}$, is given by the negative of the above term

$$\vec{E}_T^{\text{PAT}} = \sum_{k=1}^{n_J} (\mathcal{E}_k - \langle\rho\rangle A_k) \frac{\vec{E}_{T_k}^{\text{Jet}}}{E_{T_k}^{\text{Jet}}} \tag{5.10}$$

In the PUfit algorithm, only jets are used as inputs as shown in Equation 5.10. Signals of leptons only deposit energy over a small area in the calorimeter, meaning that the amount of pile-up energy buried under that area is negligible. Therefore, including the leptons will not make a huge difference to the fit.

### 5.1.3 Pile-up Contribution

One of the constraints of PUfit is that pile-up vertices produce no real MET. Namely, the sum of all PU energies in the calorimeter should be zero within the resolution. The contributions to the sums of clusters and tracks are specified in the following two equations. The pile-up objects are labeled by the red colour, and $C\&N$ are charged

and neutral pile-up objects that are outside HS jets.

$$\begin{aligned}\sum_{j=1}^{\text{clus}\notin\text{J}} \vec{E}_{T_j} &= \sum_{j}^{C} \vec{E}_{T_j} + \sum_{j}^{N} \vec{E}_{T_j} + \sum_{j}^{N} \vec{E}_{T_k} + \sum_{j}^{C} \vec{E}_{T_k} \\ \sum_{j=1}^{\text{track}\notin\text{J}} \vec{p}_{T_j}^{\text{HS}} &= \sum_{j}^{T(|\eta|<2.4)} \vec{E}_{T_j} = \sum_{j}^{C(|\eta|<2.4)} \vec{E}_{T_j}\end{aligned} \tag{5.11}$$

In the second equation above, we are assuming that the momentum measured by the tracking system is the same as the measurement from the calorimeter.

One of the global constraints in the PUfit algorithm is that the sum of transverse energy vectors of pile-up objects will be zero, i.e, no real $E_T^{\text{miss}}$ is produced in pile-up interactions. This global constraint can be formulated mathematically by summing the expected pile-up energy under jets $\mathcal{E}$ with all PU energies outside jets. For PU energies outside jets, we first add up all clusters and then subtract HS track $p_T$ to eliminate identifiable HS components.

$$\sum_{j=1}^{\text{clus}\notin\text{J}} \vec{E}_{T_j} - \sum_{j=1}^{\text{HS track}\notin\text{J}} \vec{p}_{T_j} + \sum_{k}^{n_J} \vec{\mathcal{E}}_k = 0 \tag{5.12}$$

This is equivalent to :

$$\sum_{j}^{C(|\eta|>2.4)} \vec{E}_{T_j} + \sum_{j}^{N} \vec{E}_{T_j} + \sum_{j}^{N} \vec{E}_{T_j} + \sum_{j}^{C} \vec{E}_{T_j} + \sum_{k}^{n_J} \vec{\mathcal{E}}_k = 0 \tag{5.13}$$

Notice that after the subtraction, soft neutral and charged($|\eta| > 2.4$) HS signals will still be present since they can not be successfully identified by the current algorithm. These signals will cause the calorimeter to be unbalanced and thus mandating $\mathcal{E}$ to cover the deficit after the fit. Therefore, the direction and magnitude information of these un-identified soft signals would be captured by fit parameters $\mathcal{E}$.

In PUfit, the strength of the constraint is determined by the variance associated with it in the covariance matrix. In the global constraint shown in Equation 5.12, each cluster in the sum contributes to the overall variance of the constraint according to its energy. The variance on the measured transverse energy $E_{T_j}$ is known to be proportional to $|E_{T_j}|$

$$\sigma_j^2 = r^2|E_{T_j}| + r_0^2 \tag{5.14}$$

where $r$ is the resolution scale and $r_0$ is the resolution floor to avoid computational problem when $\sigma_j$ being zero. We take

$$\begin{aligned} r &= 0.5 \quad \text{GeV}^{1/2} \\ r_0 &= 0.05 \quad \text{GeV} \end{aligned} \tag{5.15}$$

These parameters are not fully optimized but they are good enough to test the algorithm. The momenta of the soft HS tracks are well measured and uncertainties on $p_{T_j}^{\mathrm{HS}}$ are negligible compared to the $\sigma_j$.

### 5.1.4 Fit

The PUfit algorithm uses the following two constraints:

- Pileup vertices should not produce any real $\vec{E}_T^{miss}$.
- Pileup energies under HS jets ($\mathcal{E}_k$) are close to the average pileup ($<\rho> A_k$)

We construct a $\chi^2$ function:

$$\chi^2(\mathcal{E}_1, \ldots, \mathcal{E}_{n_J}) = \Delta^T V^{-1} \Delta \tag{5.16}$$

where the variables $\mathcal{E}_k$ are the fitted parameters corresponding to $n_J$ HS jets, determined by minimizing the $\chi^2$ function. The $(2 + n_J)$-dimensional vector $\Delta$ is given by

$$\Delta = \begin{pmatrix} \sum_{j=1}^{\text{clus}\notin\text{J}} E_{T_j} \cos\phi_j - \sum_{j=1}^{\text{HS track}\notin\text{J}} p_{T_j} \cos\phi_j + \sum_{k=1}^{n_J} \mathcal{E}_k \cos\phi_k \\ \sum_{j=1}^{\text{clus}\notin\text{J}} E_{T_j} \sin\phi_j - \sum_{j=1}^{\text{HS track}\notin\text{J}} p_{T_j} \sin\phi_j + \sum_{k=1}^{n_J} \mathcal{E}_k \sin\phi_k \\ \mathcal{E}_1 - \langle\rho\rangle A_1 \\ \vdots \\ \mathcal{E}_{n_J} - \langle\rho\rangle A_{n_J} \end{pmatrix} \tag{5.17}$$

and $V$ is the associated covariance matrix. The angles $\phi_j$ ($\phi_k$) are the azimuthal directions of their corresponding clusters. The first two lines impose constraints corresponding to Equation 5.12. Each remaining line corresponds to the constraint of $\mathcal{E}_k$ being close to their expected values discussed above. The average energy is given by the product of the jet area $A_k$ and the average $E_T$ density, $\langle\rho\rangle$.

The full covariance is

$$V = \begin{pmatrix} V_{11} & V_{12} & 0 & 0 & \dots & 0 \\ V_{21} & V_{22} & 0 & 0 & \dots & 0 \\ 0 & 0 & V^{\mathrm{J}} & 0 & \dots & 0 \\ 0 & 0 & 0 & V^{\mathrm{J}} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & V^{\mathrm{J}} \end{pmatrix} \tag{5.18}$$

where $V^{\mathrm{J}}$ is defined in Equation 5.8 and the upper $2 \times 2$ submatrix is given by

$$\begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{\mathrm{clus}\notin\mathrm{J}} \sigma_j^2 \cos^2\phi_j & \sum_{j=1}^{\mathrm{clus}\notin\mathrm{J}} \sigma_j^2 \cos\phi_j \sin\phi_j \\ \sum_{j=1}^{\mathrm{clus}\notin\mathrm{J}} \sigma_j^2 \cos\phi_j \sin\phi_j & \sum_{j=1}^{\mathrm{clus}\notin\mathrm{J}} \sigma_j^2 \sin^2\phi_j \end{pmatrix} \tag{5.19}$$

Minimizing this $\chi^2$ function with respect to the $\mathcal{E}_k$ provides $\mathcal{E}_k$ to be used in Equation 5.10 and the PST from Equation 5.6 can be computed accordingly.

### 5.1.5 Double subtraction of pile-up energy

It is vital that the PAT term has the form given in Equation 5.10 instead of only having $\vec{E}_T^{\mathrm{PAT}} = \sum_{k=1}^{n_J} \mathcal{E}_k \frac{\vec{E}_{T_k}^{\mathrm{Jet}}}{E_{T_k}^{\mathrm{Jet}}}$. The reason for this is that offline jets had already gone through sophisticated calibration with pile-up subtraction. If one merely subtracts PU energy $\mathcal{E}$, then it means that PU energies under jets would have been subtracted twice, in both calibration and PUfit. To avoid this, one uses the definition of PAT in Equation 5.10.

The problem of double subtraction of pile-up originates from the definition of $\mathcal{E}$. It is defined to represent the pile-up energy under HS jets, which is a feature inherited from the trigger version of PUfit. The original PUfit algorithm was designed for the trigger system, in which no pile-up subtraction is applied. Therefore the correction in the trigger PUfit can be defined easily as $\vec{E}_T^{\mathrm{PAT}} = \sum_{k=1}^{n_J} \mathcal{E}_k \frac{\vec{E}_{T_k}^{\mathrm{Jet}}}{E_{T_k}^{\mathrm{Jet}}}$.

Another approach, perhaps a more intuitive one, could be used to avoid the double subtraction of pile-up. This new approach will be derived in detail, and it will be shown to be equivalent to the approach described in earlier sections.

#### The intuitive approach

The most natural method to avoid double subtraction is to stop letting $\mathcal{E}$ representing the PU energy but only targeting the leftover soft signals. It would be tempting to

revise the $\Delta$ function defined in Equation 5.17 so that the expectation value of $\mathcal{E}$ is set to zero starting from the third row. At the same time, one also needs to add in $E_{J,k}^{PU}$, the PU subtracted from jets in the offline calibration, to the first two rows.

After such revisions, the new $\Delta$ function and correction could be defined as the following:

$$\Delta^{new} = \begin{pmatrix} \sum_{i=1}^{O} E_{T_i} \cos\phi_i + \sum_{k=1}^{n_J} \mathcal{E}_k^{new} \cos\phi_k + \sum_{k=1}^{n_J} E_{J,k}^{PU} \cos\phi_k \\ \sum_{i=1}^{O} E_{T_i} \sin\phi_i + \sum_{k=1}^{n_J} \mathcal{E}_k^{new} \sin\phi_k + \sum_{k=1}^{n_J} E_{J,k}^{PU} \sin\phi_k \\ \mathcal{E}_1^{new} \\ \vdots \\ \mathcal{E}_{n_J}^{new} \end{pmatrix} \tag{5.20}$$

where $E_{J,k}^{PU}$ is the amount of PU energy subtracted in the offline calibration and the sum over O is defined as:

$$\sum_{j=1}^{O} = \sum_{j=1}^{\text{clus}\notin\text{J}} - \sum_{j=1}^{\text{HS track}\notin\text{J}} \tag{5.21}$$

One can see that in $\Delta^{new}$, expectation values of $\mathcal{E}$ are set to be zero starting in the third row and $E_{J,k}^{PU}$ is added into the global pile-up constraint. So we will still be solving for $\mathcal{E}$ to balance the calorimeter energy deposit while requiring the correction magnitude $\mathcal{E}$ to be small, i.e., the expectation of $\mathcal{E}$ being zero.

It is obvious that this definition of $\Delta^{new}$ avoids the double subtraction problem when it is used in offline framework. The correction term can then be directly borrowed from trigger PUfit:

$$\vec{E}_T^{\text{PAT}} = \sum_{k=1}^{n_J} \mathcal{E}_k \frac{\vec{E}_{T_k}^{\text{Jet}}}{E_{T_k}^{\text{Jet}}} \tag{5.22}$$

**Theorem of Equivalence**

It is crucial for readers to understand that the PUfit algorithm was thoroughly debugged and optimized when it was tested in the trigger system, with the $\Delta$ function defined in the form of Equation 5.17. Although the definition of $\Delta^{new}$ in Ch.5.1.5 suits the need of implementing PUfit in the offline framework, this new definition

would significantly alter the solution matrix. To avoid a large number of changes in the existing code, one might wonder if there is an easier way to cope with the double subtraction problem that does not involve a completely new $\Delta$ function.

The answer is yes. There is an easier way to deal with this and it is described in the following theorem:

**Theorem 5.1.1** (Theorem of equivalence)**.** *The outcome of the fit using $\Delta$ is equivalent to that using $\Delta^{new}$ if $\langle\rho\rangle A_k = E^{PU}_{J,k}$ and the correction term $\vec{E}^{PAT} = \sum_{k=1}^{J} \mathcal{E}_k \frac{\vec{E}^{jet}_{T_k}}{E^{jet}_{T_k}}$ is modified to be $\sum_{k=1}^{J}(\mathcal{E}_k - \langle\rho\rangle A_k)\frac{\vec{E}^{jet}_{T_k}}{E^{jet}_{T_k}}$*

The theorem itself is not hard to understand intuitively. Since there is a double subtraction problem, one must eliminate the $E^{PU}_{J,k}$ part from the final correction in PUfit if $\Delta$ remains unchanged. If $\langle\rho\rangle A_k$ is taken to be close to $E^{PU}_{J,k}$, then it is straight forward to subtract it in the PAT.

However it is still important that theorem 5.1.1 can be proven to be true. So let us first consider the assumption that $\langle\rho\rangle A_k = E^{PU}_{J,k}$. Then substituting this into Equation 5.20 yields:

$$\Delta^{new} = \begin{pmatrix} \sum_{i=1}^{O} E_{T_i}\cos\phi_i + \sum_{k=1}^{n_J}\mathcal{E}^{new}_k \cos\phi_k + \sum_{k=1}^{n_J}\langle\rho\rangle A\cos\phi_k \\ \sum_{i=1}^{O} E_{T_i}\sin\phi_i + \sum_{k=1}^{n_J}\mathcal{E}^{new}_k \sin\phi_k + \sum_{k=1}^{n_J}\langle\rho\rangle A\sin\phi_k \\ \mathcal{E}^{new}_1 \\ \vdots \\ \mathcal{E}^{new}_{n_J} \end{pmatrix} \tag{5.23}$$

Then one can collect terms by defining new variable $\mathcal{E}'$:

$$\mathcal{E}'_k = \mathcal{E}^{new}_k + E^{PU}_{J,k} = \mathcal{E}^{new}_k + \langle\rho\rangle A_k \tag{5.24}$$

Substituting $\mathcal{E}'$ into Equation 5.23, we have:

$$\Delta^{new} = \begin{pmatrix} \sum_{i=1}^{O} E_{T_i} \cos\phi_i + \sum_{k=1}^{n_J} \mathcal{E}'_k \cos\phi_k \\ \sum_{i=1}^{O} E_{T_i} \sin\phi_i + \sum_{k=1}^{n_J} \mathcal{E}'_k \sin\phi_k \\ \mathcal{E}'_1 - \langle\rho\rangle A_1 \\ \vdots \\ \mathcal{E}'_{n_J} - \langle\rho\rangle A_{n_J} \end{pmatrix} \tag{5.25}$$

Now we have transformed $\Delta^{new}$ into the same form as $\Delta$ defined in Equation 5.17 without adding or deleting terms but merely collecting terms with new nomenclatures. Since $\sum_{j=1}^{O} = \sum_{j=1}^{\text{clus}\notin\text{J}} - \sum_{j=1}^{\text{HS track}\notin\text{J}}$ , we see that the two $\Delta$ functions are the same.

Therefore, we can set up the following relationship after acknowledging that $\Delta$ is equivalent to $\Delta^{new}$:

$$\mathcal{E'}_k = \mathcal{E}_k^{new} + \langle\rho\rangle A_k = \mathcal{E}_k^{new} + E_{J,k}^{PU} = \mathcal{E}_k \tag{5.26}$$

This gives:

$$\mathcal{E}_k^{new} = \mathcal{E}_k - E_{J,k}^{PU} = \mathcal{E}_k - \langle\rho\rangle A_k \tag{5.27}$$

According to Equation 5.27, if we replace $\mathcal{E}_k$ in the PAT term by $\mathcal{E}_k - \langle\rho\rangle A_k$, then the outcome of the fit should be equivalent to that of $\Delta^{new}$ given that $\langle\rho\rangle A_k = E_{J,k}^{PU}$. This is why we have to use the definition in Equation 5.10. Ideally, one would use $E_{J,k}^{PU}$ itself instead of using $\langle\rho\rangle A_k$ which is computed inside PUfit. However, it is hard to retrieve information like $E_{J,k}^{PU}$ and one can assume $\langle\rho\rangle A_k \approx E_{J,k}^{PU}$ since the offline calibration also mainly uses an area-density approach to determine pile-up.

## 5.2 PUfit with Particle Flow

In PUfit, we are concerned with the imbalance in the energy deposit inside the calorimeter. The inputs to the algorithm described above are topoclusters and inner detector tracks. In order to sum up all pile-up energies, hard scatter tracks $\vec{p}_T$ are subtracted manually as in Equation 5.12. This means that we are assuming the energy deposited in the calorimeter by charged particles is equal to the energy measured by the trackers. However, this assumption is not entirely accurate, and such energy subtractions can be greatly improved by using an algorithm called Particle Flow. Furthermore, particle flow allows the calorimeter energy deposits within $|\eta| < 2.4$ to

be classified as either from charged particles or from neutral particles.

Particle Flow is developed to incorporate both tracking and calorimeter signals in object reconstruction since tracking signals have a better resolution at lower momentum, while the calorimeter signals are more precise when momentum is high, a combination of both systems would be the optimal choice. In order to avoid double counting the same signal in the tracker and the calorimeter during reconstruction, the PFlow (Particle Flow) algorithm identifies energy deposited by charged particles in topoclusters. After estimating the energy deposit of tracks, PFlow does a much more sophisticated energy subtraction based on the location of the track in the calorimeter. The PFlow algorithm is optimized to subtract only one particle's energy deposit in the calorimeter without affecting the energy deposited by other particles[25].

It is obvious that the PFlow algorithm suits the exact need of PUfit. With PFlow, one can perform the subtraction in Equation 5.12 in a more sophisticated fashion and use tracking signals for all charge particles with $|\eta| < 2.4$ for better resolution. PFlow categorizes all objects into two types: charged PFO (PFlow Objects) and neutral PFO. All tracks within $|\eta| < 2.4$ are classified as charged PFO and all topoclusters after the energy subtractions are classified as neutral PFO. The following equation gives the correspondence between charged and neutral PFOs and the notation used previously for topoclusters in Equation 5.11.

$$\begin{aligned}\sum_{j}^{\mathrm{PFO_N}\notin J} \vec{E}_{T_j} &= \sum_{k}^{C(|\eta|>2.4)} \vec{E}_{T_j} + \sum_{k}^{N} \vec{E}_{T_j} + \sum_{k}^{N} \vec{E}_{T_k} + \sum_{k}^{C(|\eta|>2.4)} \vec{E}_{T_k} \\ \sum_{j}^{\mathrm{PFO_{C,PU}}\notin J} &= \sum_{k}^{C(|\eta|<2.4)} \vec{E}_{T_k}\end{aligned} \tag{5.28}$$

Then Equation 5.12 can be re-written with $PFO_N$ (neutral PFO) and $PFO_C$(charged PFO):

$$\sum_{j}^{\mathrm{PFO_N}\notin J} \vec{E}_{T_j} + \sum_{j}^{\mathrm{PFO_{C,PU}}\notin J} \vec{E}_{T_j} + \sum_{k}^{n_J} \vec{\mathcal{E}}_k = 0 \tag{5.29}$$

where $PFO_{C,PU}$ represents charged PFO that are linked to pile-up tracks.

Therefore, the input to PUfit was changed to $PFO_N$ and $PFO_{C,PU}$. They are summed into 112 towers as before, with the rest of fit being the same. The only modification in the formulation is that we changed the first two rows of Equation

5.17 according to Equation 5.29. The PFlow version of $\Delta$ is given by Equation 5.30

$$\Delta = \begin{pmatrix} \sum_{j}^{\mathrm{PFO_N}\notin J} \vec{E}_{T_j}\cos\phi_j + \sum_{j}^{\mathrm{PFO_{C,PU}}\notin J} \vec{E}_{T_j}\cos\phi_j + \sum_{k=1}^{n_J}\mathcal{E}_k\cos\phi_k \\ \sum_{j}^{\mathrm{PFO_N}\notin J} \vec{E}_{T_j}\sin\phi_j + \sum_{j}^{\mathrm{PFO_{C,PU}}\notin J} \vec{E}_{T_j}\sin\phi_j + \sum_{k=1}^{n_J}\mathcal{E}_k\sin\phi_k \\ \mathcal{E}_1 - \langle\rho\rangle A_1 \\ \vdots \\ \mathcal{E}_{n_J} - \langle\rho\rangle A_{n_J} \end{pmatrix} \tag{5.30}$$

The covariance sub-matrix shown in Equation 5.19 would then be changed according to the substitution: $\sum_{j=1}^{\mathrm{clus}\notin \mathrm{J}} \rightarrow \sum_{j}^{\mathrm{PFO_N}} + \sum_{j}^{\mathrm{PFO_{C,PU}}\notin J}$.

## 5.3 Analytical solution

In order to compute the correction, we must first derive the solution to the $\chi^2$ fit to obtain $\mathcal{E}$. The solution will be derived from Equation 5.16 and Equation 5.17. The solution to the PFlow version of PUfit would be similar as what is shown below with the only difference being the inputs.

Let us first define some new variables in Table 5.1 to simplify the derivation, where $\sum_j^O$ is defined in Equation 5.21.

| new variable | meaning |
|---|---|
| $d$ | $V_{11}V_{22} - V_{12}V_{21}$ |
| $E_{x_j}$ | $E_{T_j}\cos\phi_j$ |
| $E_{y_j}$ | $E_{T_j}\sin\phi_j$ |
| $\Delta_x$ | $\sum_j^O E_{x_j} + \sum_k^{n_J}\mathcal{E}_k\cos\phi_k$ |
| $\Delta_y$ | $\sum_j^O E_{y_j} + \sum_k^{n_J}\mathcal{E}_k\sin\phi_k$ |
| $\Delta_k$ | $\mathcal{E}_k - \langle\rho\rangle A_k$ |
| $\alpha_k$ | $\frac{V_{22}\cos\phi_k - V_{12}\sin\phi_k}{d}\sum_j^O E_{x_j} + \frac{V_{11}\sin\phi_k - V_{12}\cos\phi_k}{d}\sum_j^O E_{y_j}$ |
| $\beta_{kj}$ | $\frac{V_{22}\cos\phi_k - V_{12}\sin\phi_k}{d}\cos\phi_j + \frac{V_{11}\sin\phi_k - V_{12}\cos\phi_k}{d}\sin\phi_j$ |

Table 5.1: New variables

Let us first expand $\chi^2$ using variables defined above. Since we are treating every jet in the same way, the variance $V_k^J$ on each $\mathcal{E}$ is constant and denoted as $V^J$

$$\begin{aligned}\chi^2 &= \Delta^T V^{-1} \Delta \\ &= \frac{V_{22}}{d}\Delta_x^2 - 2\frac{V_{12}}{d}\Delta_x\Delta_y + \frac{V_{11}}{d}\Delta_y^2 + \sum_k^{n_J}\frac{\Delta_k^2}{V^J}\end{aligned} \tag{5.31}$$

Then the solution that minimized $\chi^2$ is obtained by solving:

$$\frac{\partial \chi^2}{\partial \mathcal{E}_k} = 0 \tag{5.32}$$

Derivatives with respect to $\mathcal{E}_k$ are :

$$\begin{aligned}\frac{\partial \Delta_x}{\partial \mathcal{E}_k} &= \cos\phi_k \\ \frac{\partial \Delta_y}{\partial \mathcal{E}_k} &= \sin\phi_k \\ \frac{\partial \Delta_j}{\partial \mathcal{E}_k} &= \delta_k^j\end{aligned} \tag{5.33}$$

Then we substitute these derivatives into Equation 5.31 and Equation 5.32:

$$\begin{aligned}\frac{1}{2}\frac{\partial \chi^2}{\partial \mathcal{E}_k} &= \frac{V_{22}}{d}\Delta_x\frac{\partial \Delta_x}{\partial \mathcal{E}_k} - \frac{V_{12}}{d}\Delta_x\frac{\partial \Delta_y}{\partial \mathcal{E}_k} - \frac{V_{12}}{d}\Delta_y\frac{\partial \Delta_x}{\partial \mathcal{E}_k} + \frac{V_{11}}{d}\Delta_y\frac{\partial \Delta_y}{\partial \mathcal{E}_k} + \sum_k^{n_J}\frac{\Delta_k}{V^J}\frac{\partial \Delta_k}{\partial \mathcal{E}_k} \\ &= \frac{\Delta_x}{d}(V_{22}\cos\phi_k - V_{12}\sin\phi_k) + \frac{\Delta_y}{d}(V_{11}\sin\phi_k - V_{12}\cos\phi_k) + \frac{\Delta_k}{V^J} \\ &= (\alpha_k - \frac{\langle\rho\rangle A_k}{V^J}) + \beta_{11}\mathcal{E}_1 + \beta_{12}\mathcal{E}_2 + ... + \beta_{k,n_J}\mathcal{E}_{n_J} + \frac{1}{V^J}\mathcal{E}_k \\ &= 0\end{aligned} \tag{5.34}$$

So for $\frac{\partial \chi^2}{\partial \mathcal{E}_1} = 0$, we have:

$$\left(\alpha_1 - \frac{\langle\rho\rangle A_1}{V^J}\right) + \left(\beta_{11} + \frac{1}{V^J}\right)\mathcal{E}_1 + \beta_{12}\mathcal{E}_2 + ... + \beta_{k,n_J}\mathcal{E}_{n_J} = 0 \tag{5.35}$$

Similarly, we have the following for $\frac{\partial \chi^2}{\partial \mathcal{E}_2} = 0$:

$$\left(\alpha_2 - \frac{\langle\rho\rangle A_2}{V^J}\right) + \beta_{21}\mathcal{E}_1 + \left(\beta_{22} + \frac{1}{V^J}\right)\mathcal{E}_2 + ... + \beta_{k,n_J}\mathcal{E}_{n_J} = 0 \tag{5.36}$$

We can deduce the Equation for all other $\mathcal{E}$'s and then express them in the matrix form:

$$\begin{pmatrix} \beta_{11} + \frac{1}{V^J} & \beta_{12} & \dots & \beta_{1n_J} \\ \beta_{21} & \beta_{22} + \frac{1}{V^J} & \dots & \beta_{2,n_J} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{n_J,1} & \beta_{n_J,2} & \dots & \beta_{n_J,n_J} + \frac{1}{V^J} \end{pmatrix} \begin{pmatrix} \mathcal{E}_1 \\ \mathcal{E}_2 \\ \vdots \\ \mathcal{E}_{n_J} \end{pmatrix} = \begin{pmatrix} \frac{\langle\rho\rangle A_1}{V^J} - \alpha_1 \\ \frac{\langle\rho\rangle A_2}{V^J} - \alpha_2 \\ \vdots \\ \frac{\langle\rho\rangle A_{n_J}}{V^J} - \alpha_{n_J} \end{pmatrix} \tag{5.37}$$

Then the solution can be easily determined by inverting a $n_J \times n_J$ matrix.

# Chapter 6

# Results

Both the topocluster and PFlow version of PUfit are introduced in Chapter 5, and they will both be tested in this chapter. The procedures used to compare different algorithms are developed and applied to both versions of PUfit. The algorithm is mainly tested in $Z^0 \rightarrow \mu^+\mu^-$ samples with the Z boson used as an imaginary MET source as discussed in section 2.3.2. The two main variables used in the analysis are the resolution and the scale difference.

In these selected $Z^0 \rightarrow \mu^+\mu^-$ events, the transverse momentum of the Z boson is compared to the estimates of MET produced by different algorithms. To obtain the resolution along a given axis for the MET, one takes the difference between components of MET and transverse momentum of the Z boson ($p_T$ (Z)) along the given axis to get a distribution. The standard deviation of the obtained distribution is used as resolution. The resolution along the x-axis is defined as:

$$\text{Resolution } Ex = \sigma_x = \boldsymbol{\sigma}\left[(\vec{E}_{\mathrm{T}}^{\mathrm{miss}} - \vec{p}_{\mathrm{T}}^{\mathrm{Z}}) \cdot \hat{x}\right] \tag{6.1}$$

where the $\boldsymbol{\sigma}$ operator represents the action of taking the standard deviation of the given distribution and $\vec{p}_{\mathrm{T}}^{\mathrm{Z}}$ is the transverse momentum vector of the Z boson. Since the resolution is defined by the standard deviation, the uncertainty of the resolution is then:

$$\boldsymbol{\sigma}(\sigma_x) = \frac{\sigma_x}{\sqrt{2N-2}} \tag{6.2}$$

where $N$ is the number of events used in the calculation.

Aside from resolution, it is also fruitful to look at the scale of the MET compared to that of the Z transverse momentum. One could simply take the difference between magnitudes to check the scale difference but it is often useful to look at the **parallel**

**scale difference**(PSD). The parallel scale difference is defined as:

$$\begin{aligned} \text{PSD} &= \vec{E}_{\text{T}}^{\text{miss}} \cdot \hat{p}_{\text{T}}^{\text{Z}} - |\vec{p}_{\text{T}}^{\text{Z}}| \\ &= |\vec{E}_{\text{T}}^{\text{miss}}| \cdot \cos\theta - |\vec{p}_{\text{T}}^{\text{Z}}| \end{aligned} \tag{6.3}$$

where $\hat{p}_T^Z$ is the unit vector in the direction of the Z transverse momentum and $\theta$ is the angle between $\vec{E}_{\text{T}}^{\text{miss}}$ and $\hat{p}_T^Z$. Then $\cos\theta$ can be easily computed using

$$\cos\theta = \frac{\vec{E}_{\text{T}}^{\text{miss}} \cdot \vec{p}_{\text{T}}^{\text{Z}}}{|\vec{E}_{\text{T}}^{\text{miss}}| \cdot |\vec{p}_{\text{T}}^{\text{Z}}|} \tag{6.4}$$

## 6.1 Topocluster based PUfit

For topocluster based PUfit, the $\Delta$ function used is shown in Equation 5.17 in which all inputs are either from topoclusters or ID tracks. After computing the solutions of $\mathcal{E}_k$ using Equation 5.37, the PAT is obtained by Equation 5.10 and then substituted into Equation 5.6 to get the PST.

The magnitudes of PAT and PST are compared with TST in Figure 6.1. All events have non-zero PAT since they are required to have at least one jet in the selection criterion. It is evident that the mean value of the new soft term PST is much higher than TST after adding the correction (PAT) from PUfit. The tail of PST appears to be much higher than TST. This is an indication that PUfit offered significant corrections in many events. However, CST is still the one with the largest mean and variance. This is simply due to the fact that each pile-up vertex contributes its own variance in the calculation of CST. Occasionally, extremely imbalanced calorimeter energy deposit would yield PAT of extremely large magnitude. In order to focus on events without outliers, a cut is placed at 100 GeV for the PAT.

Components of TST MET and PST MET are shown in Figure 6.2 and 6.3 respectively. It is evident that the jet term is much larger compared to the soft terms, and the tail of MET is mostly dominated by the jet term. Since $\text{Z}^0 \to \mu^+\mu^-$ is mostly dominated by single jets events as shown in Figure 2.7, these peaks are results of the jet selection criteria discussed in Chapter 4.1.3. The first peak has a sharp edge at exactly 20 GeV which is the $p_T$ cut of jet selections in the central region. The second peak is at 30 GeV which corresponds to the cut in the forward region.

Components of CST MET are shown in Figure 6.4. It is clear that the CST is of comparable magnitude to the jet term. As a result, the CST MET is of higher mean and also has a higher tail.

All three MET are plotted in Figure 6.5 to be compared with the magnitude of Z

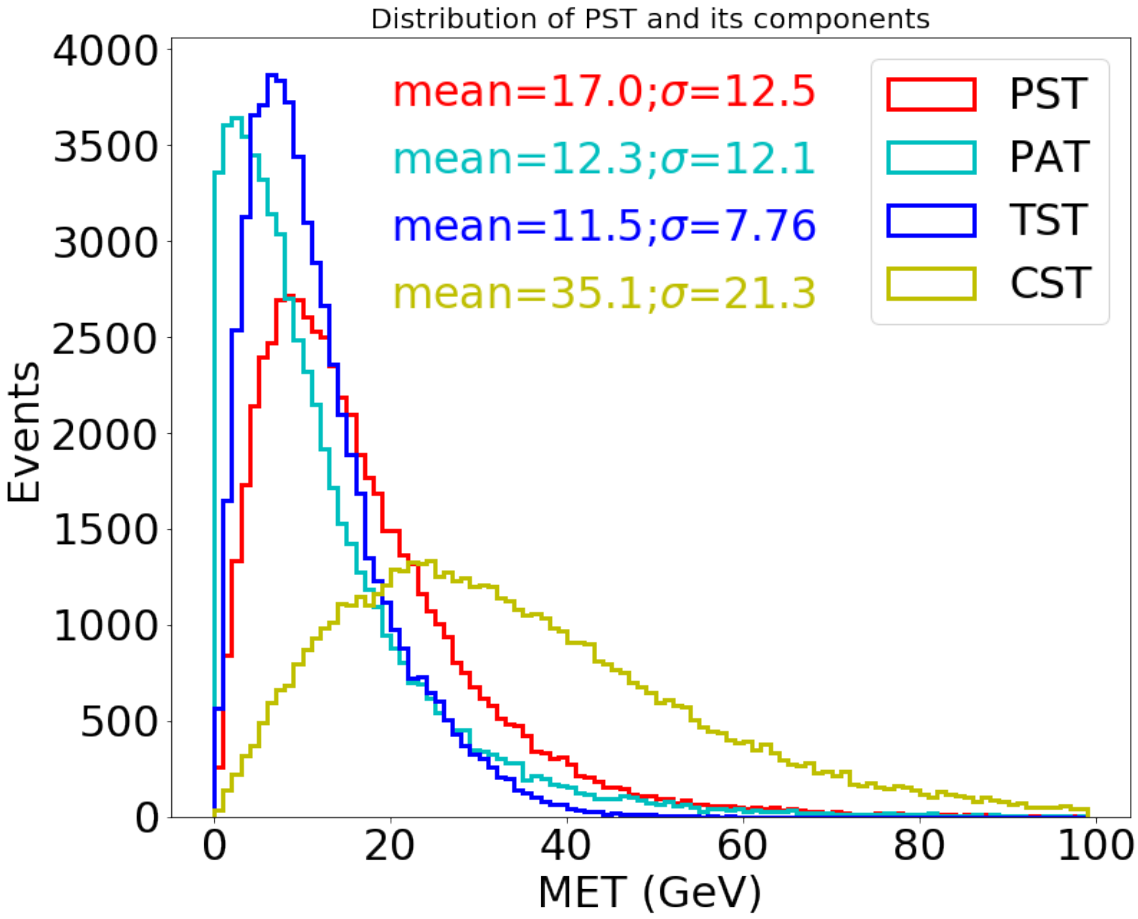


Figure 6.1: Distribution of different soft terms and topocluster based PUfit corrections in $Z^0 \to \mu^+\mu^-$ events.

boson's $p_T$. All three MET appear to have larger magnitudes compared to Z $p_T$ due to the resolution of MET reconstruction. The MET reconstruction involves more objects and it is subject to a larger resolution influence. Therefore the MET distribution could be thought of as the distribution of the 2-dimensional Z $p_T$ vector convolved with gaussian noise, and thus resulting in a larger magnitude. The distribution of PST MET magnitude is similar to that of the TST MET, especially at higher energies. This is expected since the two only differ in the soft term while the high energy tail is dominated by the jet term. The major difference between PST MET and TST MET is in the region of 0 Gev to 20 GeV, in which the PST MET has a higher overall number of entries. This resulted in a lower average for the PST MET. Since the scale of Z $p_T$ is quite low compared to MET, the CST MET distribution showed the worst results among all three algorithms with a significantly higher mean value. CST MET

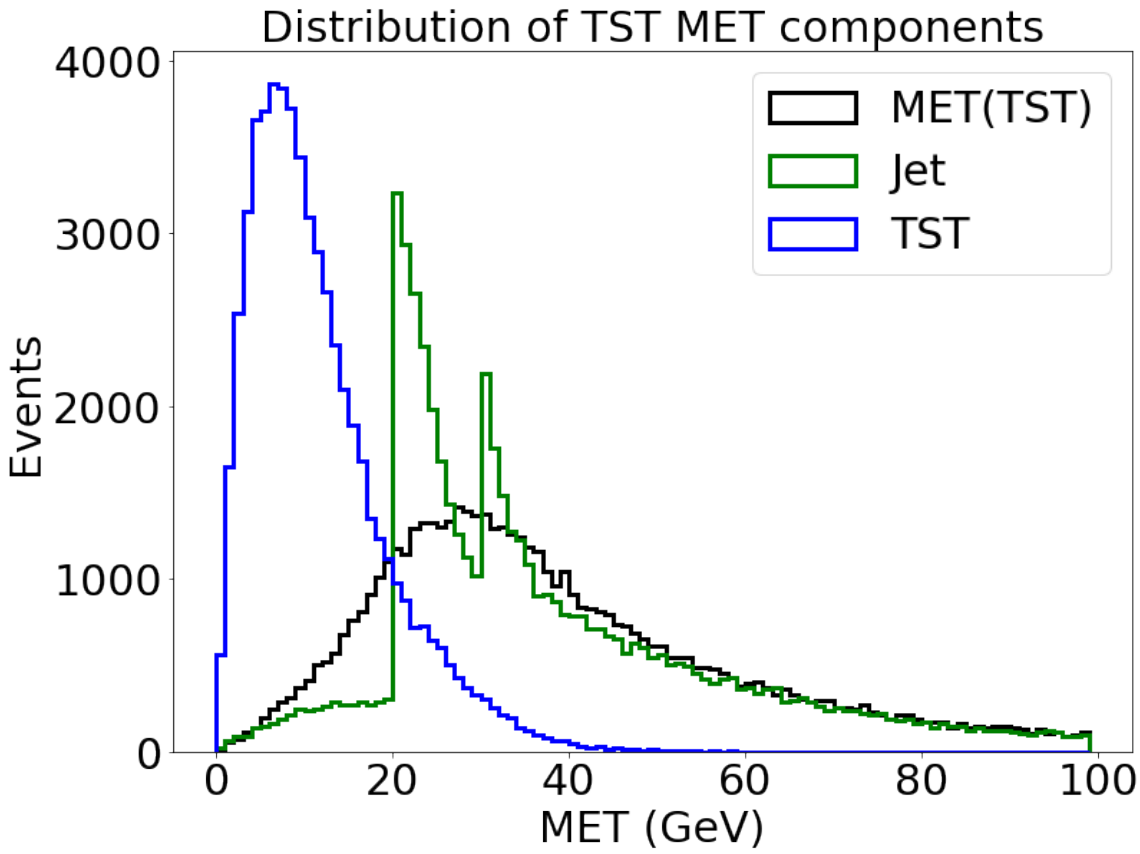


Figure 6.2: Distribution of TST MET components in $Z^0 \rightarrow \mu^+\mu^-$ events.

is similar to the other two MET in the low energy region (0 GeV to 20 GeV), but it quickly begins to differ after 20 GeV due to the large magnitude of CST. Since CST has a tail that is comparable to the jet term as shown in Figure 6.4, the tail of CST MET is much higher than that of the other two. This results in a worse resolution which will be shown later.

The resolution against Z $p_T$ is shown in Figure 6.6 and the Z $p_T$ distribution is shown in Figure 6.5. It is clear that the resolution decreased at first as the Z $p_T$ increases, and increases back up when Z $p_T$ is high. This behavior is explained by two separate factors. First, the reconstructed MET could be viewed as the convolution of the $\vec{Z}_{p_T}$ and noises along each axis in the transverse plane. This means that the mean value of the distribution is enlarged for the reconstructed MET distribution. Therefore, it is very rare for the reconstructed MET to be under 20 GeV while that is common for the Z $p_T$ . This explains the worse resolution of MET in the low Z $p_T$ region. The other factor is the increase in jet multiplicity as Z $p_T$ increases. A larger

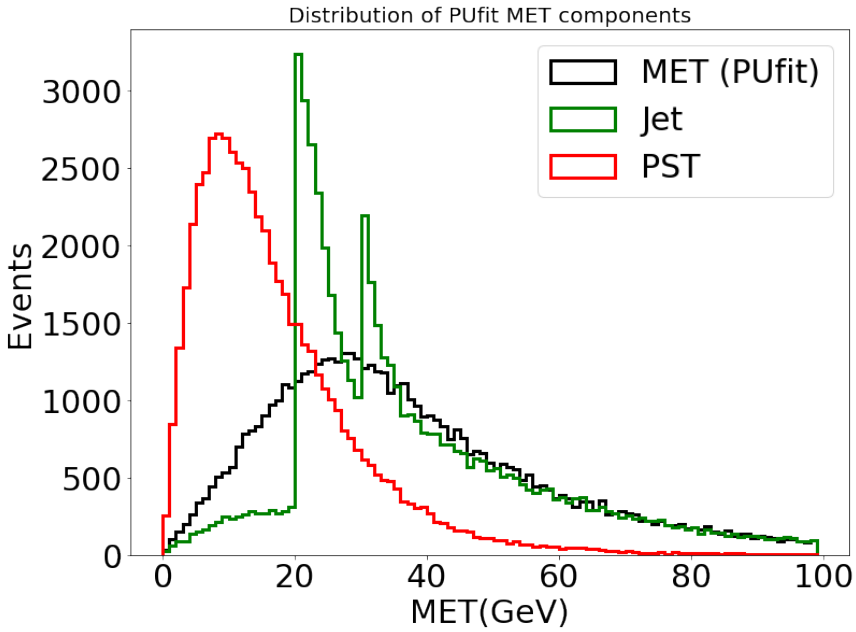


Figure 6.3: Distribution of topocluster based PUfit MET components in $Z^0 \to \mu^+\mu^-$ events.

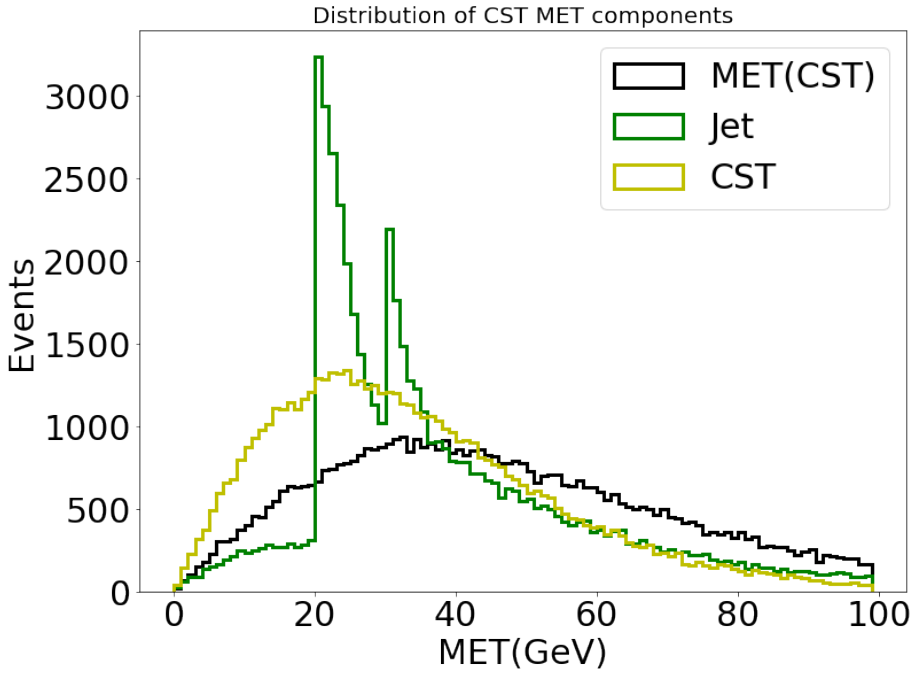


Figure 6.4: Distribution of CST MET components in $Z^0 \to \mu^+\mu^-$ events.

jet multiplicity results in the worse resolution in the high Z $p_T$ region.

Resolutions of all three MET algorithms against pile-up activities are shown in

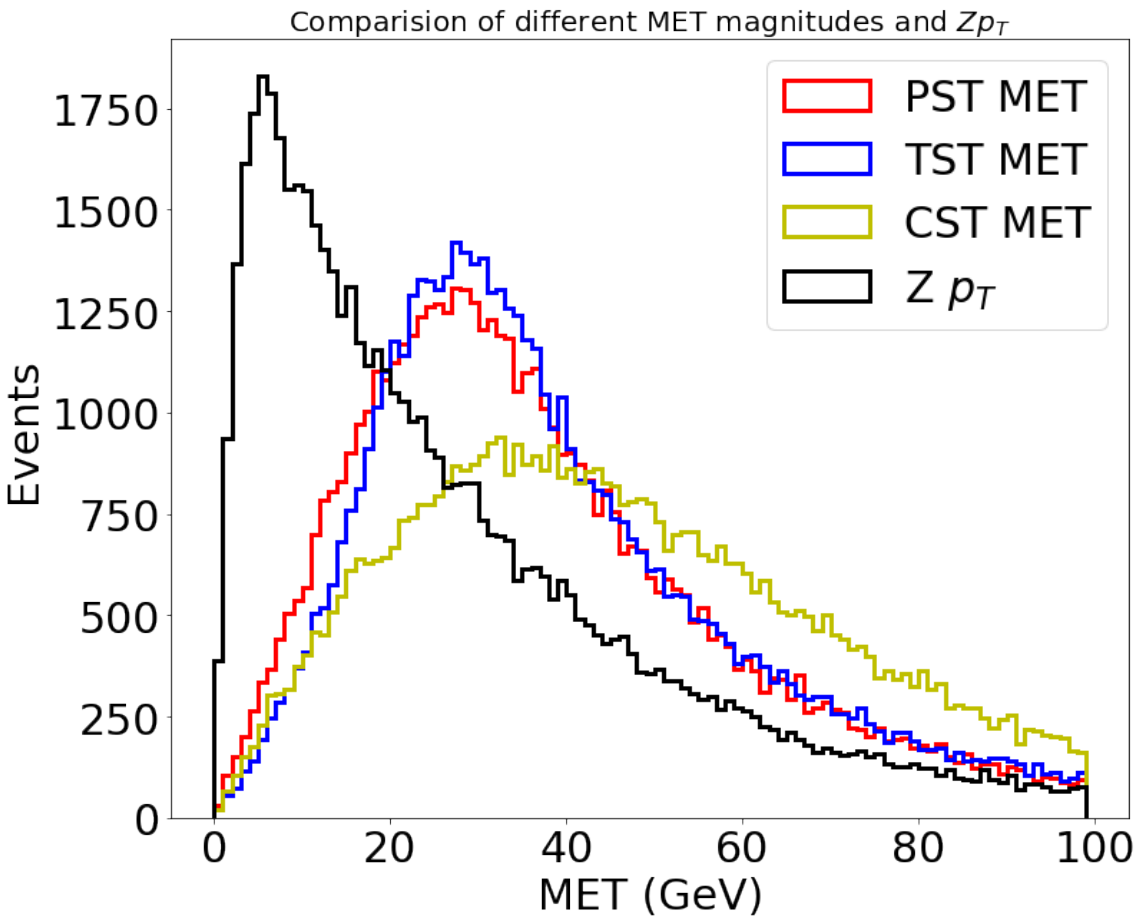


Figure 6.5: Comparison of magnitude of different MET in $Z^0 \rightarrow \mu^+\mu^-$ events.

Figure 6.7. The x-axis of the plot is the number of primary vertices in the events, which measures the amount of pile-up. Since these three MET measures only differ in their soft terms, they are labeled by the soft terms used. The resolution and its standard deviation along the x-axis are defined in Equations 6.1 and 6.2. From Figure 6.7, one can clearly see that the Ex resolution of PST MET is similar to that of the TST MET throughout the entire domain of $N_{pv}$ which ranges from 10 to 39. The resolution of CST MET starts at around 20 GeV when the pile-up is low. This is about the same resolution compared to the other two MET, but the resolution of CST MET quickly deteriorates as pile-up increases. This is expected since we have seen that the scale of CST MET is significantly larger than that of the Z $p_T$. Since each pile-up vertex brings its own variance into the calculation, it makes sense that the resolution of CST MET gradually gets worse as $N_{pv}$ increases.

Despite adding a new PAT term to the soft term, PST showed no clear improve-

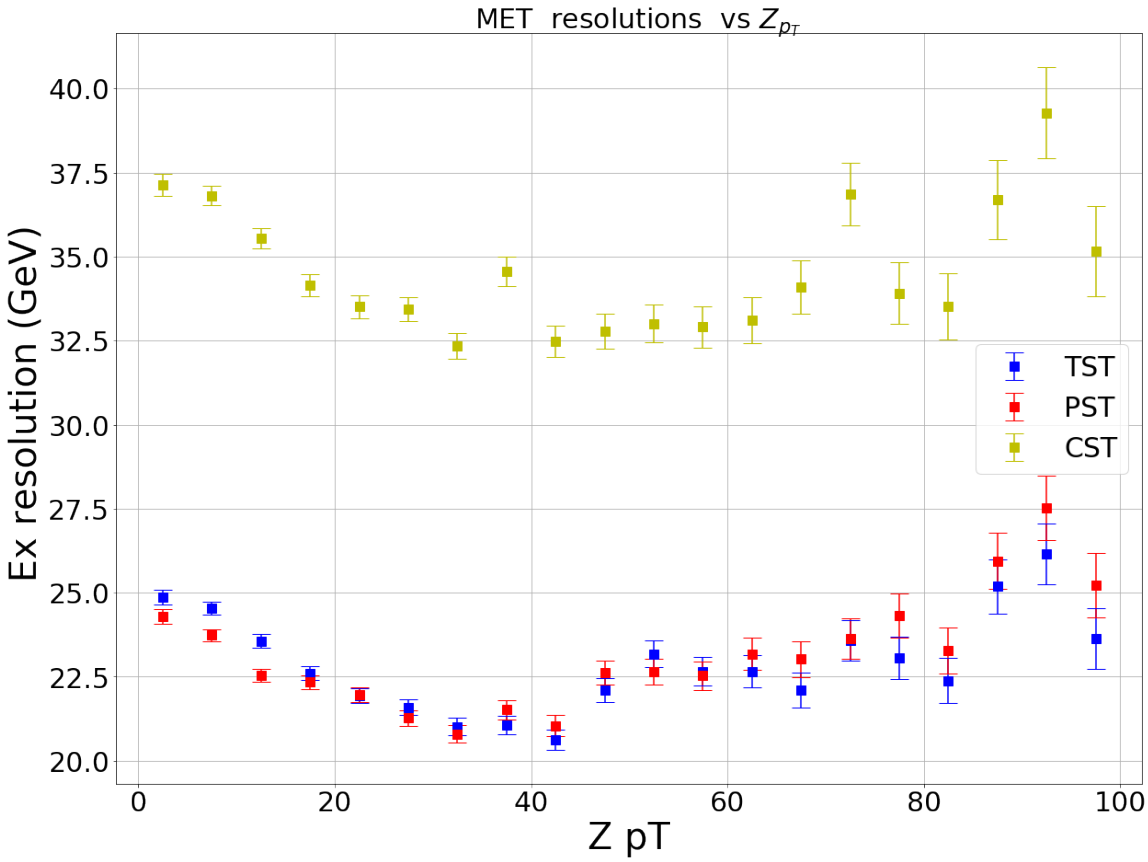


Figure 6.6: Comparison of magnitude of different MET in $Z^0 \rightarrow \mu^+\mu^-$ events.

ment in Ex resolution. One could point out that PST showed slight improvements in Figure 6.7 when $N_{pv}$ is between 15 and 25. But these improvements are mostly within 1 GeV along the x-axis. Then one could argue the difference in the soft term is not reflected in the resolution but on the scale of the MET. The angular difference between MET & Z $p_T$ and the parallel scale difference can be computed using Equation 6.3 and Equation 6.4. These two variables are shown in Figure 6.8 and Figure 6.9 respectively. The x-axis in Figure 6.8 and Figure 6.9 is Z $p_T$ instead of $N_{pv}$. This is because the main difference between PST and TST is shown to be in the magnitude of MET as in Figure 6.1. By plotting these scale variables against Z $p_T$, one could examine scale results with different truth scale (Z $p_T$ magnitudes).

In Figure 6.8, the angular difference between MET and $p_T$ (Z) is shown to be similar for TST and PST based MET. The Z $p_T$ distribution is shown in Figure 6.5. The CST based MET is significantly worse, as expected. As $p_T$ (Z) increases to be above 80 GeV, the TST and PST MET are mostly dominated by jets so that the

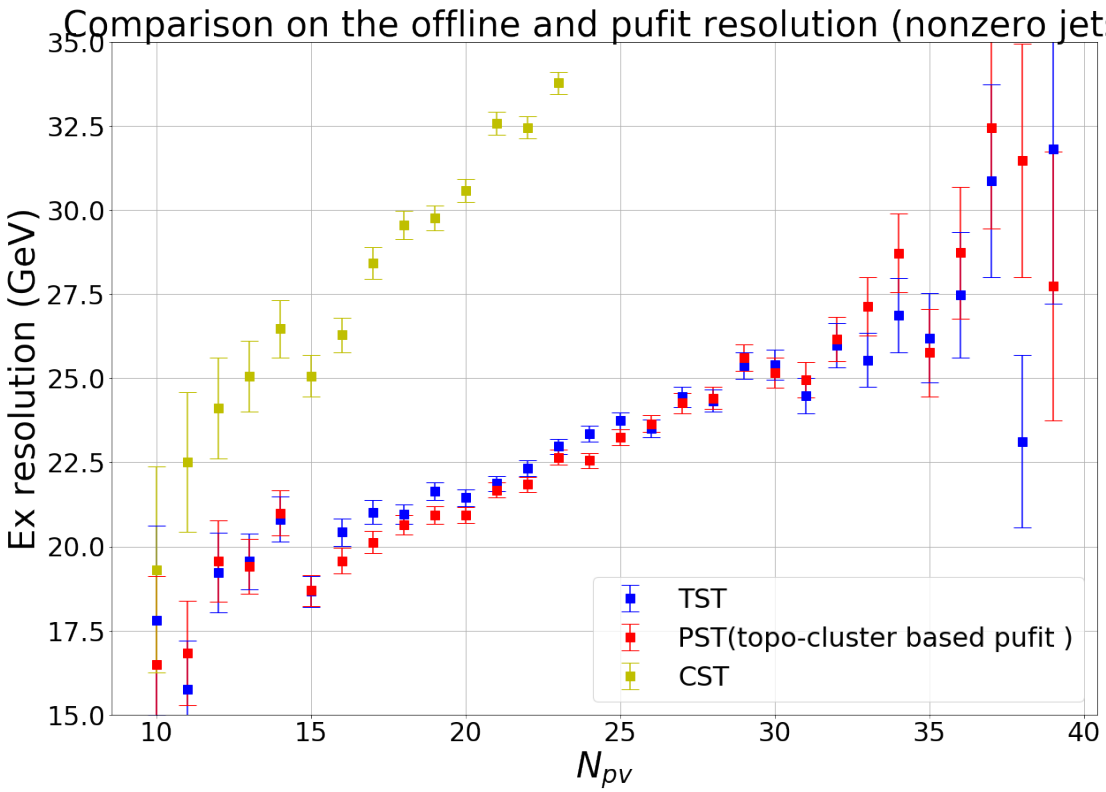


Figure 6.7: Comparison of Ex Resolution in $Z^0 \rightarrow \mu^+\mu^-$ events. The topocluster based PST MET is shown in red. The x-axis of the plot ($N_{pv}$) is the number of primary vertices which measures the amount of pile-up.

angular difference becomes similar at larger $p_T$ (Z).

In Figure 6.9, the parallel scale difference for TST MET is higher than that of the PST MET for all Z $p_T$ intervals, especially when Z $p_T$ is high. This is consistent with Figure 6.5 which indicates that the magnitude of PST MET is indeed lower on average. Note that the CST MET has the lowest scale, which can be explained by the larger angular difference shown in Figure 6.8. Since the angle is on average larger, there are more events in which the CST MET is in the opposite direction of the Z $p_T$ and thus could result in a negative scale difference.

## 6.2 PFlow PUfit

PFlow based PUfit is constructed from PFlow objects using the $\Delta$ function defined in Equation 5.30. In addition to the magnitude and resolution analysis, the angular correlation between each pair of components is also investigated for PFlow based

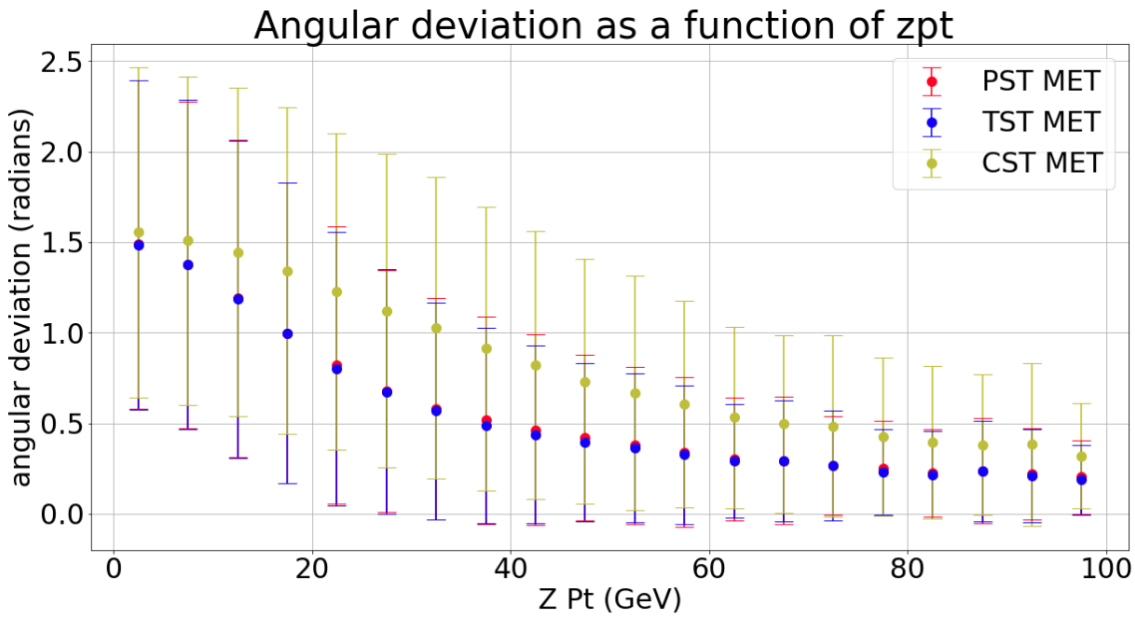


Figure 6.8: Angular difference between MET and Z $p_T$ in $Z^0 \rightarrow \mu^+\mu^-$ events. topocluster based PST MET is shown.

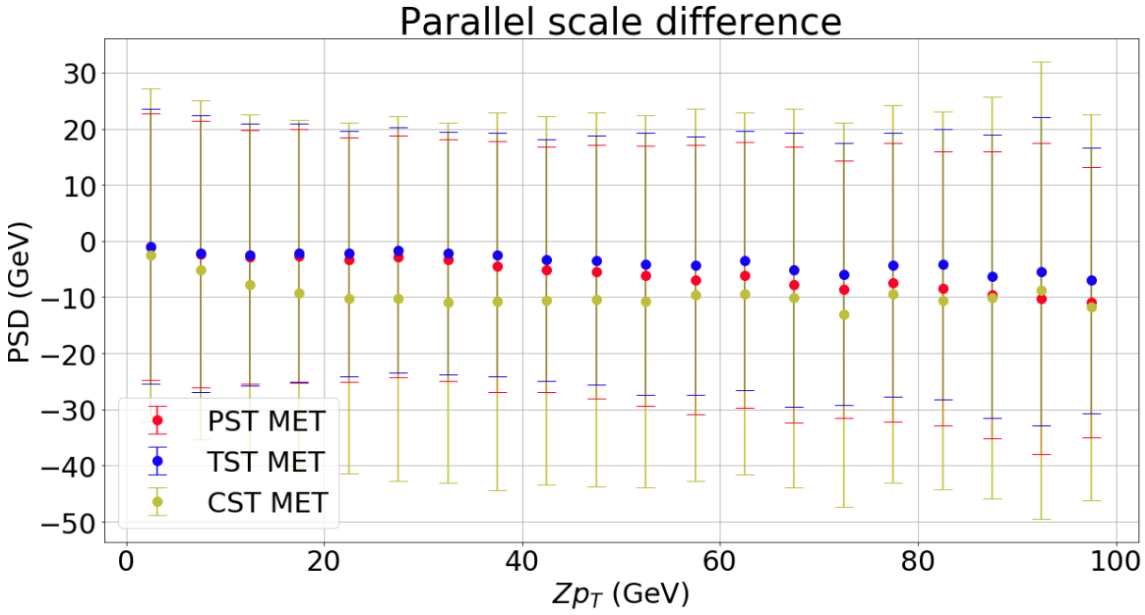


Figure 6.9: Parallel scale difference for all three MET in $Z^0 \rightarrow \mu^+\mu^-$ events. topocluster based PST MET is shown. The error-bar is the standard deviation of the PSD distribution in each $p_T$ slice

PUfit.

### 6.2.1 Magnitudes and resolution

The PFlow algorithm provides its own jet collection found by using PFlow objects. So the number of jets in each event could be slightly different from that of EM-topo jets which are used in topocluster based PUfit. The new distribution is shown in Figure 6.10. Note that zero jet events are not used in the following results as discussed in Chapter 2.3.2.

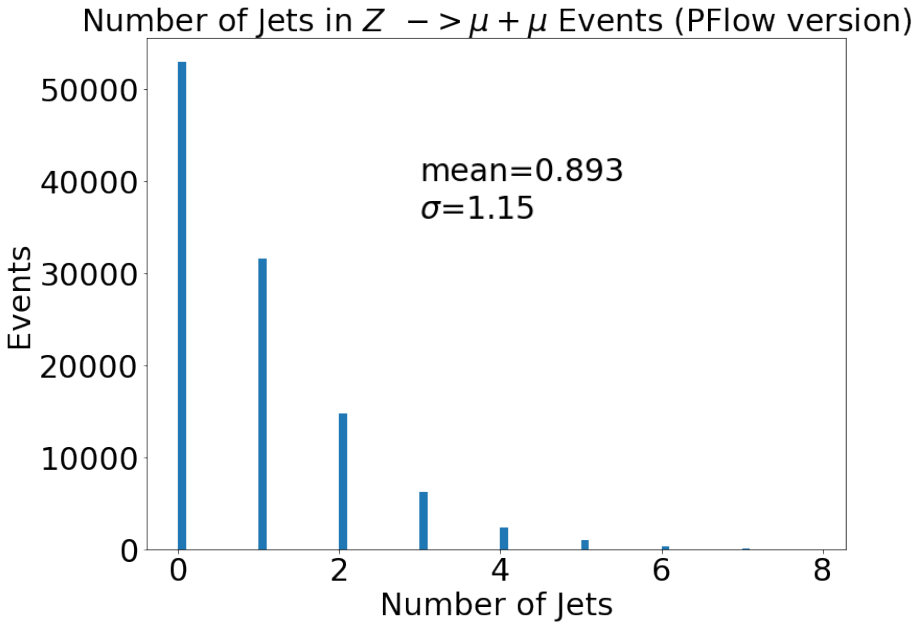


Figure 6.10: Distribution of the number of PFlow jets in $Z^0 \rightarrow \mu^+\mu^-$ events.

Magnitudes of soft terms are shown in Figure 6.11. One can see that by using PFlow inputs, all three MET varied in their magnitude.

By comparing Figure 6.11 to Figure 6.1, it is obvious that the magnitude of CST dropped drastically. In the PFlow version, CST is of comparable magnitude to other soft terms. The mean value of CST dropped 12.5 GeV to around 22.6 GeV while the standard deviation also dropped to 13.8 GeV. The impact of this new CST on the resolution will be shown later. TST appears to be more concentrated, and the mean value and standard deviation both decreased compared to previous results. However, the mean value and standard deviation of PAT did not decrease like the other two but increased instead. The PST, which is the vectorial combination of PAT and TST, stayed at the same mean value with a slightly increased standard deviation. A direct comparison of EM topocluster based MET, and PFlow MET is shown in Figure 6.12.

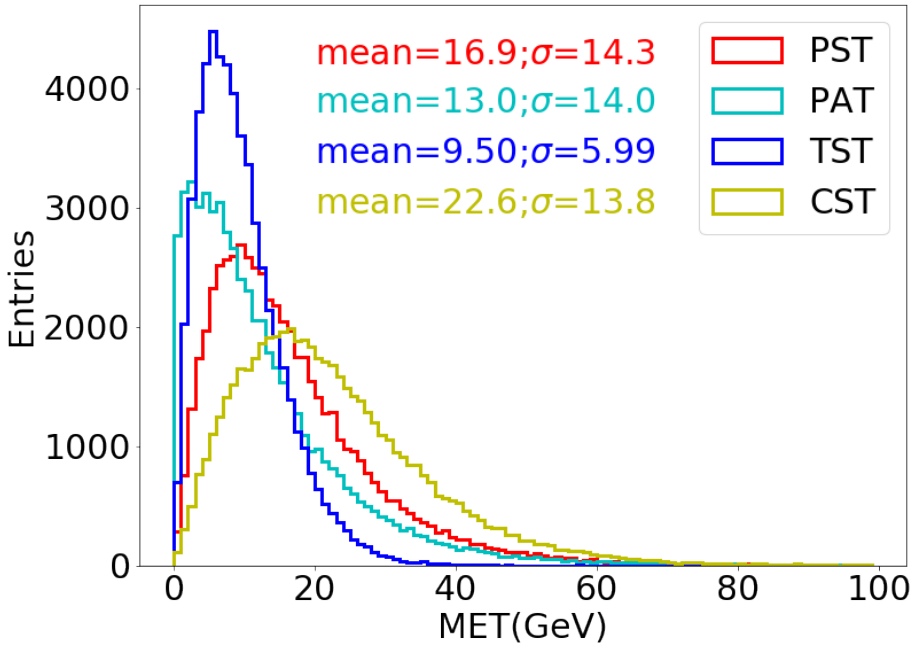


Figure 6.11: Distribution of PFlow based soft terms in $Z^0 \to \mu^+\mu^-$ events.

In general, the PAT correction made by PUfit is much higher than the previously obtained results. However, the overall structure of the MET is not changed as shown in Figure 6.13. The composition is similar to what was shown in Figure 6.3, only differing slightly in individual terms.

Resolution of PFlow based MET is shown in Figure 6.14. PST MET still shows similar results compared to TST MET. Improvements in the resolution are still mostly in the $N_{pv}$ range of 15 GeV to 25 GeV, and improve the resolution by at most 1 GeV.

All three PFlow based MET are compared to those produced by topocluster inputs in Figure 6.15. The three PFLow based MET showed improvements in the entire domain of $N_{pv}$. CST benefits the most from adopting PFlow, the resolution dropped to the same level of other MET. This is to be expected since the magnitude of CST is shown to be more reasonable in Figure 6.13. TST and PST with PFlow inputs showed slight improvements of around 1 GeV in most cases while the relative difference between the two with the same inputs has not changed much.

In Figure 6.16, the resolution of all three MET measures are shown in two different cases. As expected, the resolution in events with more than one jet is much worse than that in events with only one jet due to the increased energy deposition in these events.

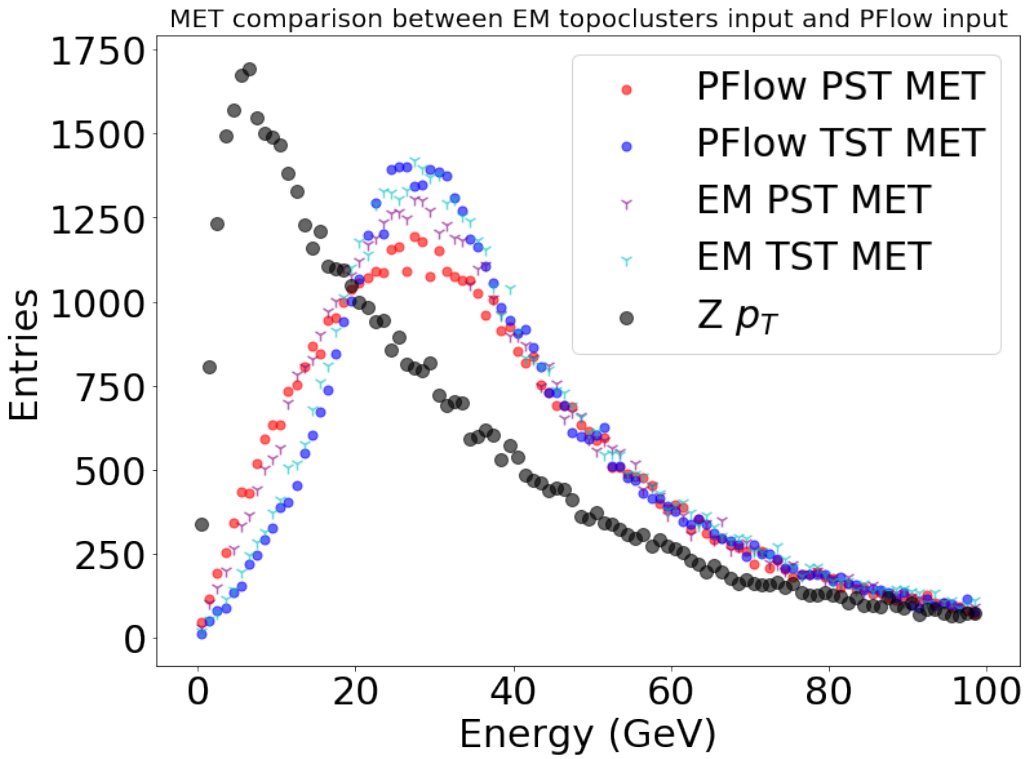


Figure 6.12: Comparison between EM-topo and PFlow based MET distribution in $Z^0 \to \mu^+\mu^-$ events.

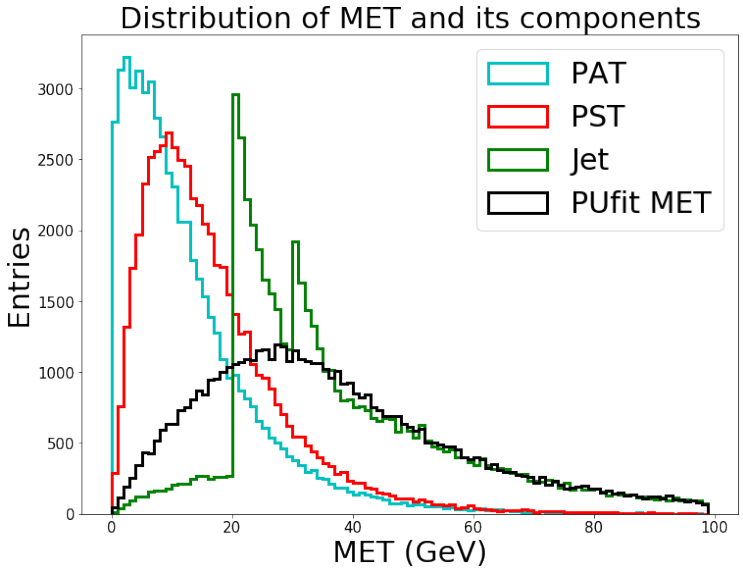


Figure 6.13: Distribution of different components of PFlow MET in $Z^0 \to \mu^+\mu^-$ events.

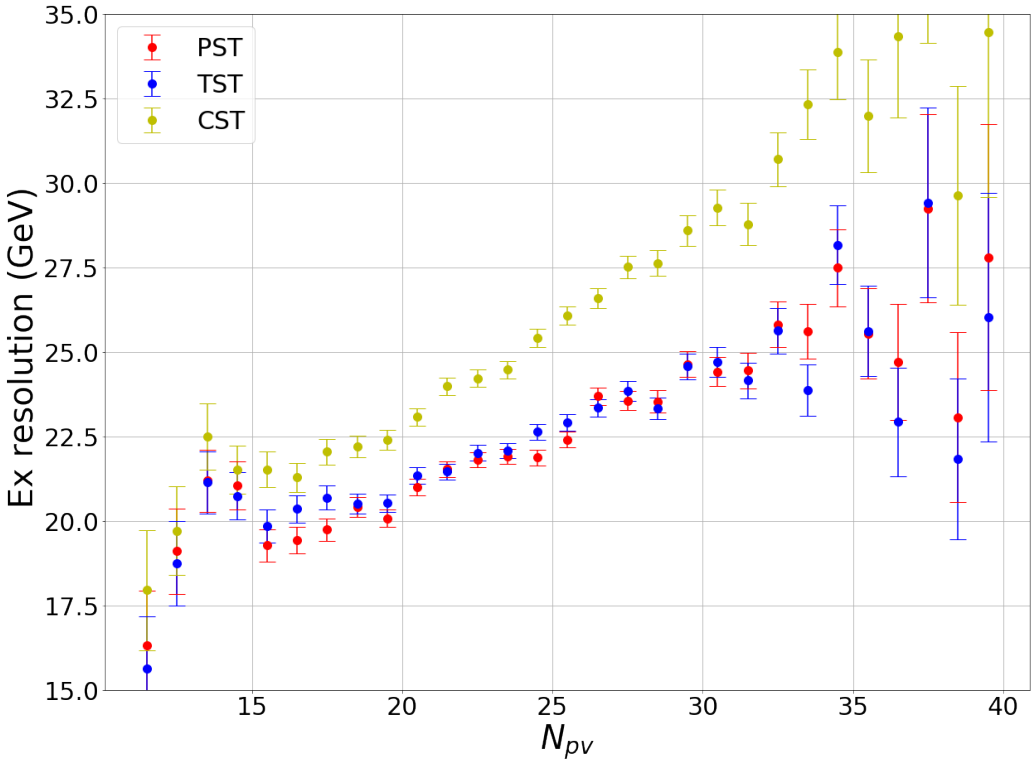


Figure 6.14: PFlow based TST/CST/PST MET Ex resolution against $N_{pv}$ in $Z^0 \rightarrow \mu^+\mu^-$ events.

In Figure 6.17, resolutions of all three MET measures are plotted against $p_T$ (Z) instead of $N_{pv}$. The resolution of all three algorithms is the best when Z $p_T$ is about 40 GeV, similar to what is previously shown in Figure 6.6.

The angular deviation and parallel scale difference are shown in Figure 6.18 and Figure 6.19. The angular difference did not change very much except for the better angular resolution of CST MET. This also true for the parallel scale difference. The PST MET is shown to have the lowest scale in the PFlow version of PUfit.

### 6.2.2 Angular correlation between MET components

Angles between different MET components could be calculated similarly to what is defined in Equation 6.4. By observing the distribution of angular differences between components, one could investigate correlations between these components. All angular differences between MET components are shown in Figure 6.20, C.1 and C.2.

In the first plot in Figure 6.20, the PAT showed a strong correlation with the jet term. The angular differences between the two components are mostly 0 or $\pi$. This

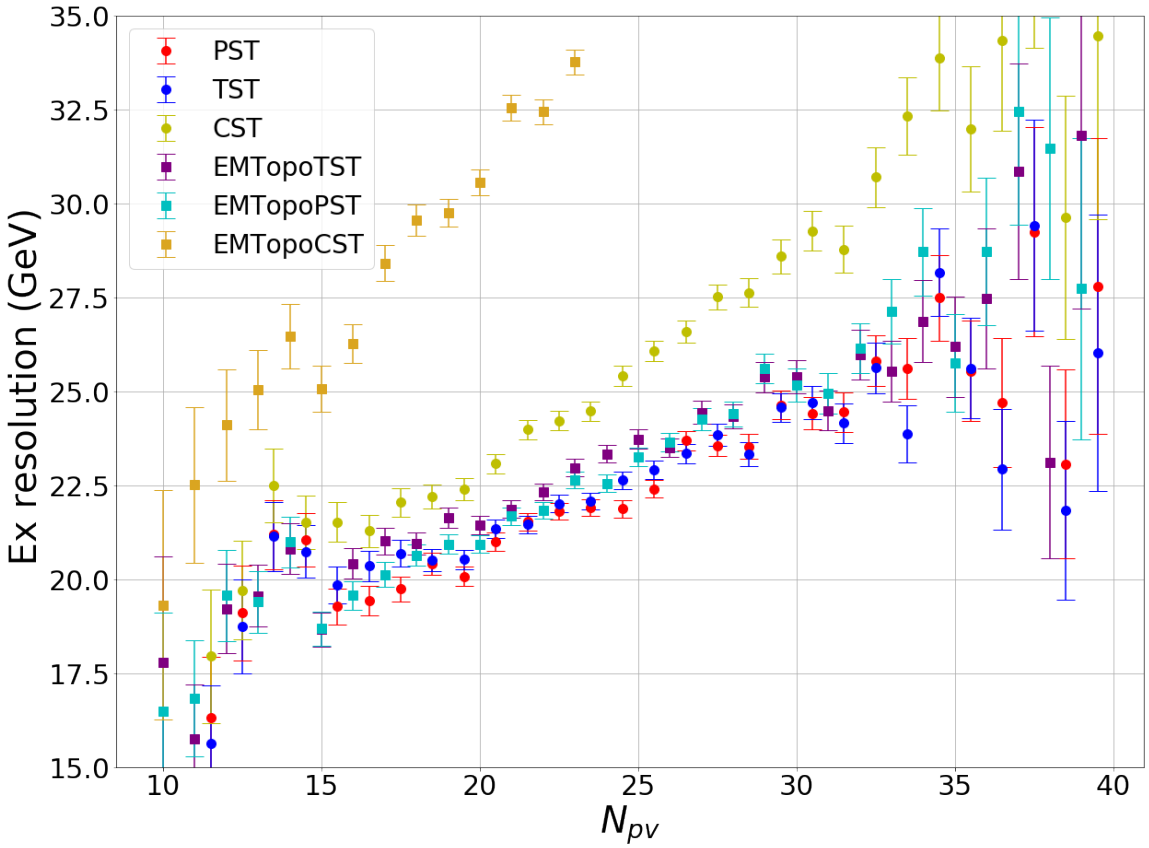


Figure 6.15: Comparison of topocluster based MET and PFlow based MET resolution in $Z^0 \to \mu^+\mu^-$ events.

indicates that PAT is mostly parallel to the jet term. This is consistent with how PUfit is designed. In PUfit, $\mathcal{E}$ parameters are only assigned to hard scatter Jets and the construction of PAT in Equation 5.10 uses the direction of HS jets in the $\eta - \phi$ plane. This means that if there is only one HS jet in the event, the PAT will be in either the same or the opposite direction of that jet. This is exactly what is being shown in the first plot in Figure 6.20. Since we have only selected events with non-zero HS jets, the dominating event topology would be single HS jet. This is shown in Figure 2.7. If one sums all events that have PAT parallel to the jet term and compares the overall number to single jet events entries in Figure 2.7, one can see that those two numbers agree with each other. This indicates that the PUfit algorithm is indeed executed as designed. There are still a large number of events between 0 and $\pi$ in the first plot in Figure 6.20. These are events with multiple HS jets in which PAT is allowed to be not parallel to the jet term. This is because that the PAT direction is a vector that is obtained by summing directions over multiple jets. Another interesting

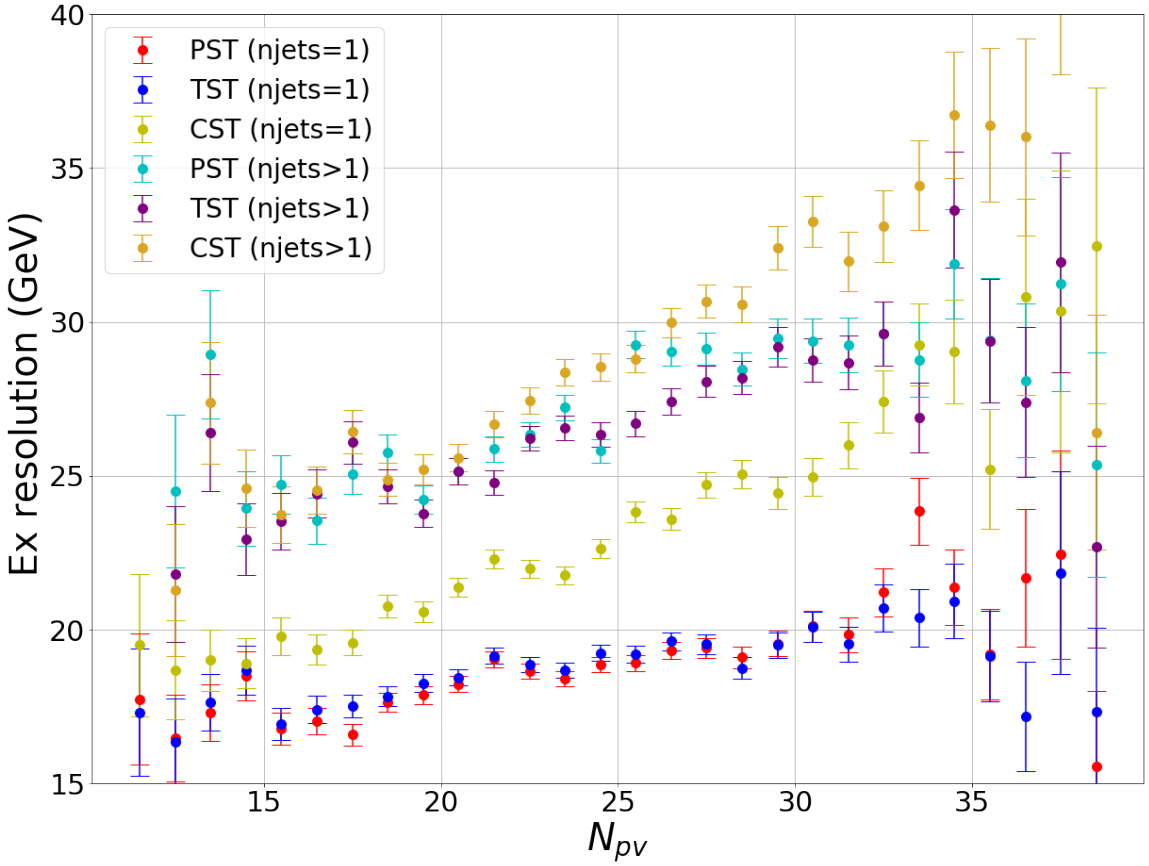


Figure 6.16: PFlow based Resolution in $Z^0 \rightarrow \mu^+\mu^-$ events with different jet multiplicity.

aspect of the first plot is that there are more events in which PAT is in the opposite direction of the jet term than in the same direction. In the second plot in Figure 6.20, we see that the jet term is mostly positively correlated with MET which makes sense since the jet term is the most dominating term in the MET composition as shown in Figure 6.13. Then by having more events in which PAT is opposite to the jet term, the overall scale of MET would be lowered by the PAT. This agrees with what is shown in Figure 6.19 that the PST MET has a lower scale compared to the TST MET.

As shown in the 4th plot in Figure 6.20, TST showed limited correlation with the jet term. However, there is a slightly increasing trend of increasing entries as the angle gets larger. PST showed correlation with the jet term since entries are high at both ends, but it also has quite a lot of uncorrelated entries in angles between 0 and $\pi$. This is simply because that the PST is defined to be the vectorial sum of PAT and

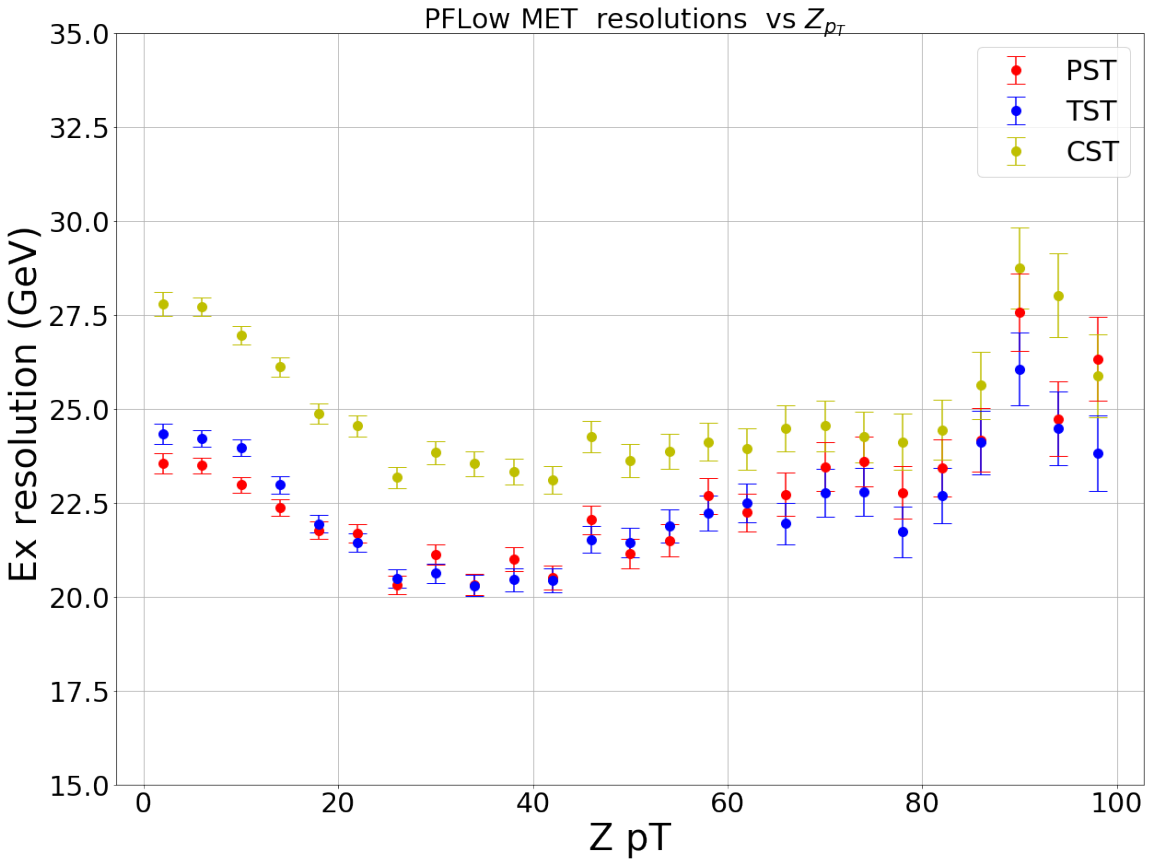


Figure 6.17: PFlow based Resolution vs Z$p_T$ in $Z^0 \to \mu^+\mu^-$ events.

TST, so it is expected to inherit correlation characteristics from both PAT and TST.

## 6.3 $Z^0 \to \mu^+\mu^-$ Monte Carlo

To check results shown in section 6.2, $Z^0 \to \mu^+\mu^-$ MC samples are used for the same analysis. The PFlow algorithm was applied in the same fashion, and the estimated MET is then compared to the $p_T$ of the truth Z boson retrieved from the generated particles container.

The reconstructed $p_T$ (Z) is calculated in the same way mentioned before, and it is compared to the truth $p_T$ (Z) in Figure 6.21. One can see that the reconstructed $p_T$ (Z) using recorded detector signals agrees well with the truth $p_T$ . The MET response is also similar to what we had previously using data. The PST MET showed smaller mean compared to the TST MET. The difference between the two MET is slightly larger than what is shown in Figure 6.12. Since the scale of PST MET is slightly

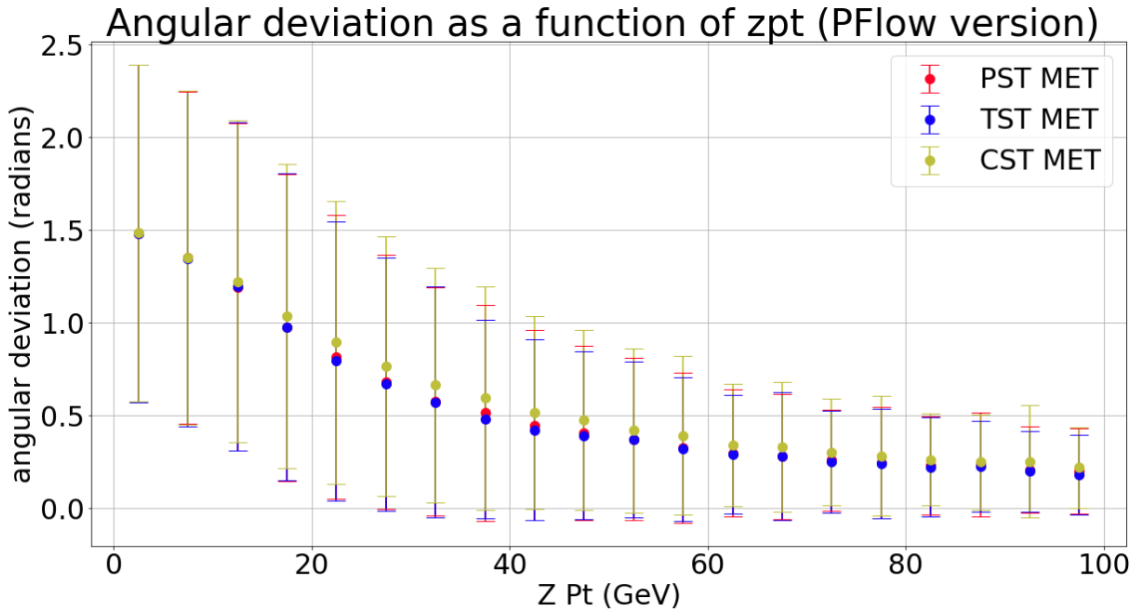


Figure 6.18: Comparisons of angular deviation for PFlow based MET in $Z^0 \rightarrow \mu^+\mu^-$ events.

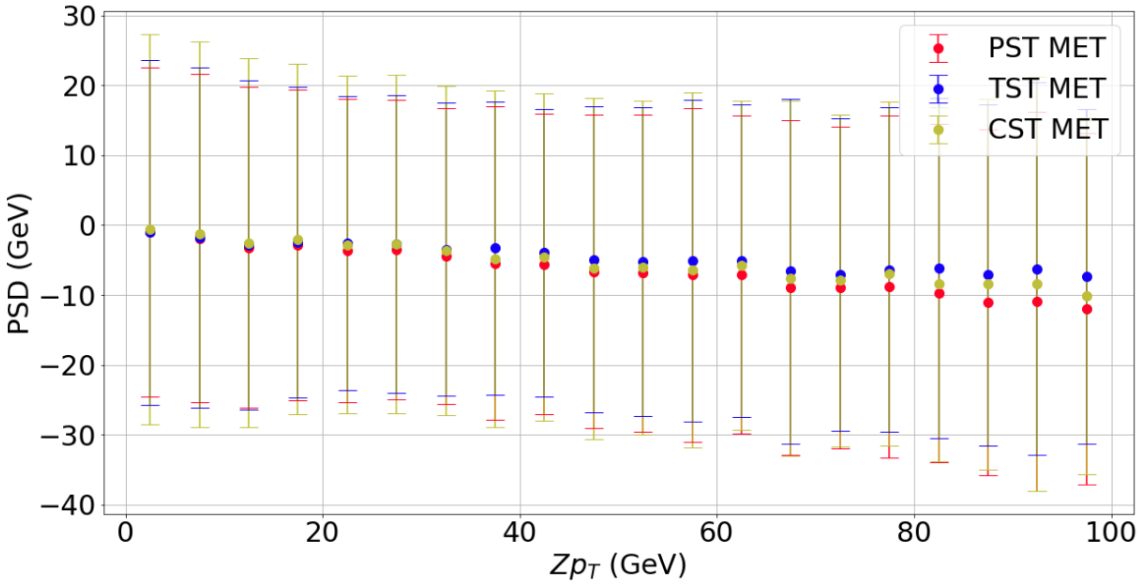


Figure 6.19: Comparisons of parallel scale difference for PFlow based MET in $Z^0 \rightarrow \mu^+\mu^-$ events.

lower compared to that of the TST MET and thus closer to the truth $p_T$ (Z), PST MET is expected to perform well in terms of resolution which is shown in Figure 6.22. Unlike previous resolution results, the PST MET showed consistent improvement of

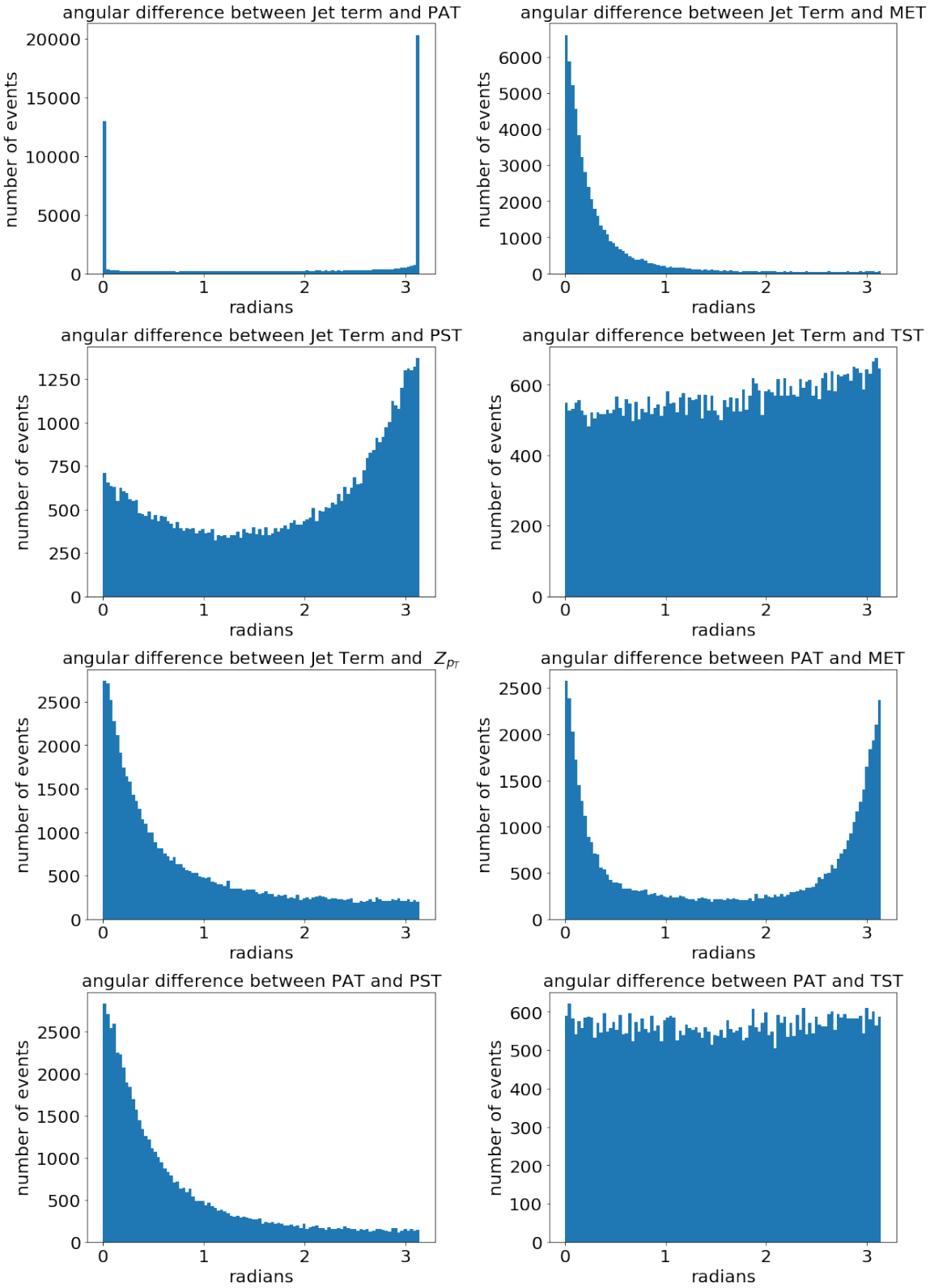


Figure 6.20: Angular difference between each MET term in $Z^0 \rightarrow \mu^+\mu^-$ events. (1) (PFlow version)

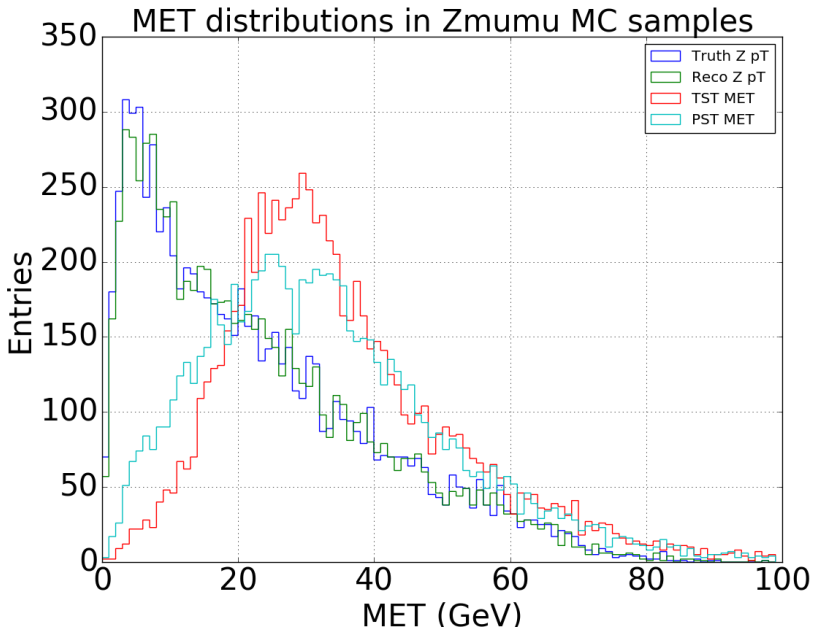


Figure 6.21: Magnitudes of PFlow-based MET distributions in $Z^0 \rightarrow \mu^+\mu^-$ MC samples

around 1 GeV. The parallel scale difference is shown in Figure 6.23 and it is slightly different compared to the PSD plot shown previously. The mean values of PSD in MC samples are shown to be mostly negative and stayed nearly constant in intermediate $p_T$ ranges.

## 6.4 W boson decay

The PUfit algorithm is tested in the $Z^0 \rightarrow \mu^+\mu^-$ samples so far. Alternatively, it could also be tested in different event topologies. An interesting topology to look at is $W \rightarrow \mu\nu_\mu$ in which the neutrino signal is the main MET source. Unlike in $Z^0 \rightarrow \mu^+\mu^-$ , muon signals are no longer considered to be MET sources, i.e. the measured muon $p_T$ is used as a hard object in the MET calculation. The $W \rightarrow \mu\nu_\mu$ samples used for the analysis are:

**Center of mass energy:** 13 TeV
**Total number of Events in the dataset:** 20000
**Pile-up condition:** Same as the LHC running condition.

The transverse momentum of neutrino could be estimated by the offline or PUfit

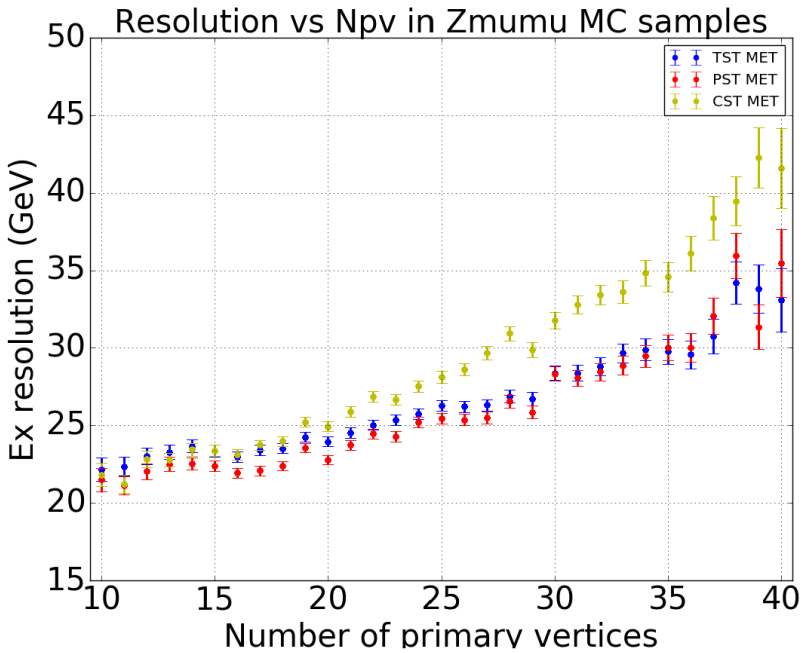


Figure 6.22: resolution of PFlow-based MET in $Z^0 \to \mu^+\mu^-$ MC samples

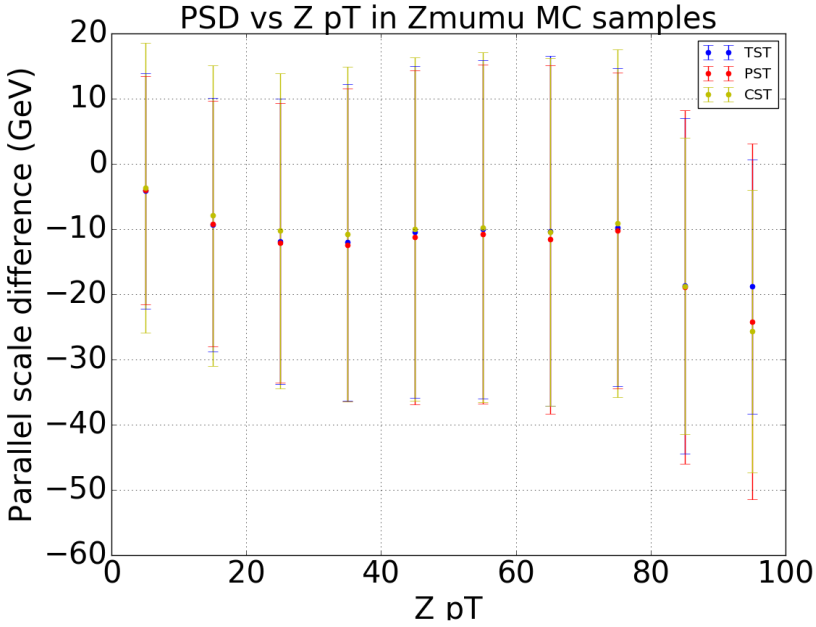


Figure 6.23: PSD of PFlow-based MET in $Z^0 \to \mu^+\mu^-$ MC samples

MET and consequently used in the reconstruction of the W boson. The PFlow version of PUfit is tested in $W \to \mu\nu_\mu$ in which $p_T$ (W) $< 70$ GeV. The first thing one needs to look at is the number of jets in each event. Similar to the previous selection, we

will be only looking at events with HS jets since they are the only ones with non-zero PAT corrections. There are also many events in the dataset that have truth MET very close to zero. These events are filtered out by a cut of 5 GeV on the truth MET. Since we are mainly interested in $W \to \mu\nu_\mu$ events, all events are required to have exactly one selected muon. The distribution of numbers of HS jets in these events are shown in Figure 6.24. Notice that only events with non-zero jets are entering this histogram and the overall distribution is similar to what $Z^0 \to \mu^+\mu^-$ offers. The distribution of truth neutrino momentum is shown in Figure 6.25.

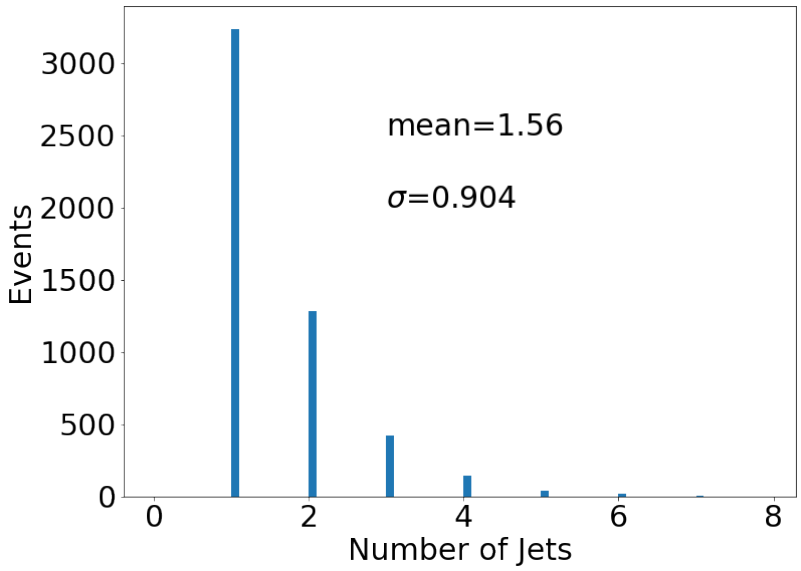


Figure 6.24: PFlow Jet multiplicity in events with at least one HS jet in $W \to \mu\nu_\mu$ events.

Magnitudes of different soft terms are shown in Figure 6.26. One can see that the general form of distributions is similar to what is shown in Figure 6.11. The distribution of MET is shown in Figure 6.25. All three MET have similar distributions that are of much higher mean value than that of the truth neutrino $p_T$ , which is mainly due to the resolution on the reconstructed MET.

An useful variable in $W \to \mu\nu_\mu$ events is the transverse mass of the W boson $m_W$, defined in Equation 6.5:

$$m_T^2 = 2E_{T,\mu}E_{T,\nu_\mu}\,(1 - \cos\theta). \tag{6.5}$$

where $E_{T,\mu}$ and $E_{T,\nu_{mu}}$ are the transverse momentum of the muon and neutrino produced by the decaying W boson. The angle $\theta$ is the angle between the muon and the neutrino in the transverse plane.

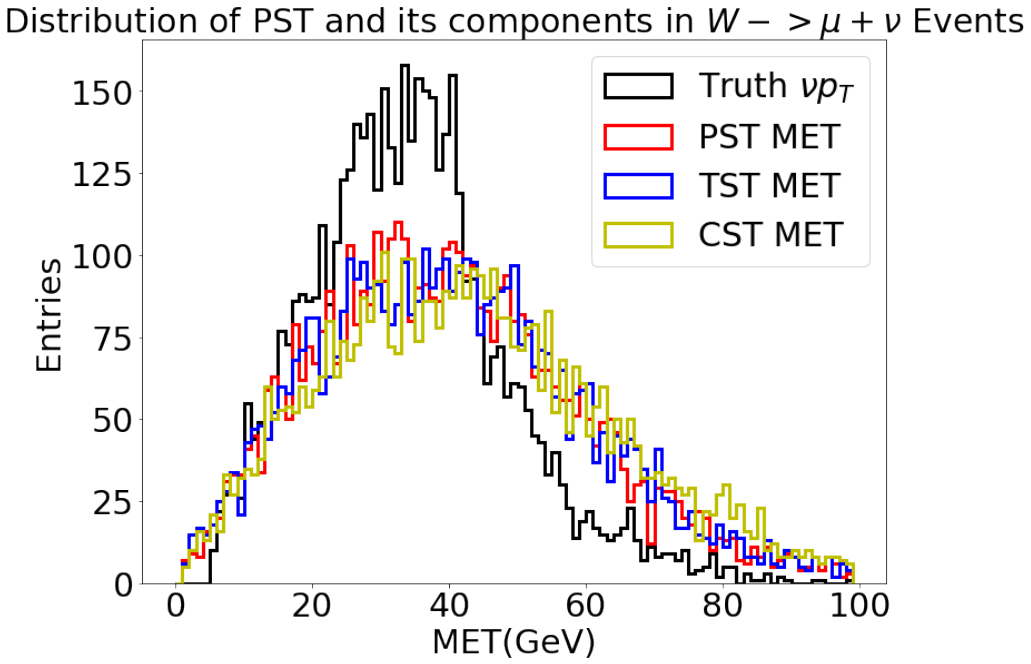


Figure 6.25: Magnitudes of PFlow-based MET in $W \to \mu\nu_\mu$ events

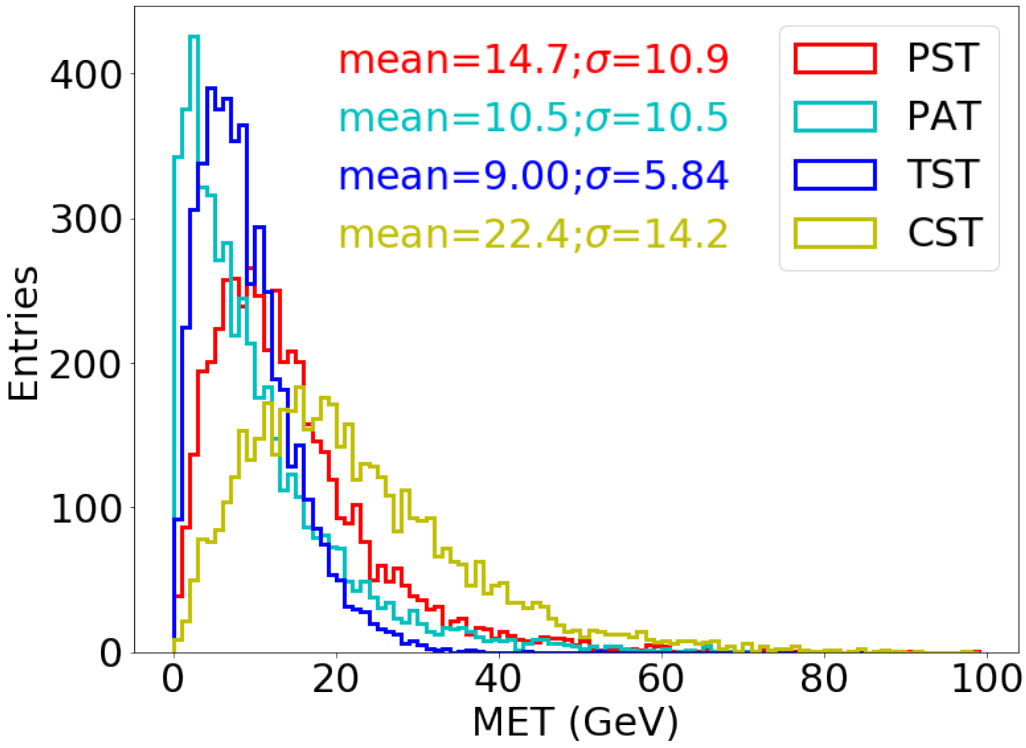


Figure 6.26: Magnitudes of PFlow-based soft terms in $W \to \mu\nu_\mu$ events

Different MET calculations are used to reconstruct the W boson's transverse mass with the recorded muon signal. The truth transverse mass is also computed using truth MET and truth muon information and plotted in Figure 6.27. The reconstructions using truth neutrinos and recorded muons are similar to those using full truth information. Therefore, it is clear that the MET resolution is the dominant factor in the transverse mass resolution.

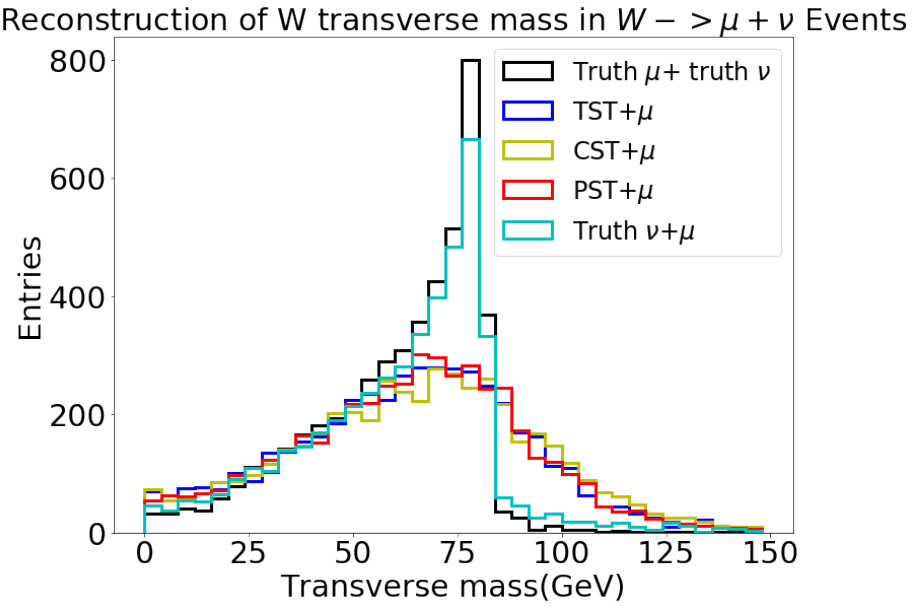


Figure 6.27: Distribution of W boson transverse mass in $W \to \mu\nu_\mu$ events.(PFlow-based)

The resolution of all three MET measures are plotted against $N_{pv}$ in Figure 6.28. The difference in resolution between the CST MET and the other two is about 2 GeV as $N_{pv}$ increases up to 25. TST MET and PST MET are shown to have similar resolutions with minor differences that are statistically insignificant.

## 6.5 $t\bar{t}$ Monte Carlo samples

The PUfit algorithm has been primarily tested in $Z^0 \to \mu^+\mu^-$ samples for the moment. The PUfit algorithm is mainly dependent on the number of hard scatter jets which has a similar distribution for both $Z^0 \to \mu^+\mu^-$ and $W \to \mu\nu_\mu$ events as shown in Figures 2.7 and 6.24. It is worthwhile to check the performance of PUfit in event topologies that are significantly different from $Z^0 \to \mu^+\mu^-$ , especially those with

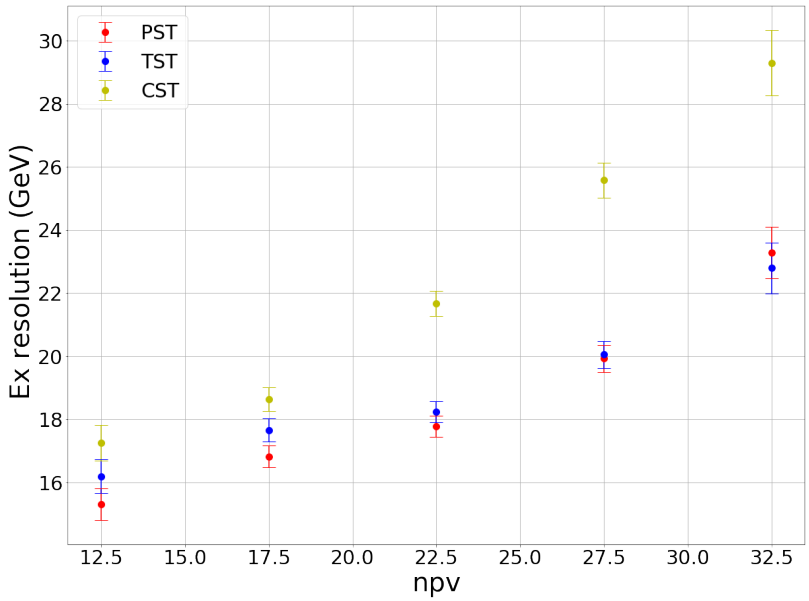


Figure 6.28: Resolution of PFlow-based offline MET and PUfit MET vs $N_{pv}$ in $W \to \mu\nu_\mu$ events.

larger jet multiplicities. Given this motivation, the $t\bar{t}$ sample is a good choice. The $t\bar{t}$ MC sample used for the analysis is:

**Center of mass energy:** 13 TeV
**Total number of Events in the dataset:**159696000
**Number of events used:**36281
**Pile-up condition:** Same as the LHC running condition.

The hard object multiplicities in $t\bar{t}$ events are shown in Figure 6.29. The jet multiplicity in $t\bar{t}$ samples is much higher than that in $Z^0 \to \mu^+\mu^-$ or $W \to \mu\nu_\mu$ samples. As a result, PUfit will have more fit parameters $\mathcal{E}$ in $t\bar{t}$ events. Note that some $t\bar{t}$ events contain high-$p_T$ electrons or muons. Recall that we used the PFLow algorithm for energy substraction as discussed in Chapter 6.2. The PFlow algorithm used in PUfit does not include corrections for leptons since it is designed for use with hadrons. Muon signals can be neglected in the PUfit algorithm since they do not leave a significant amount of energy in the calorimeter to affect the fit. However, electrons deposit non-negligible energy and their energy will not be subtracted by the PFlow algorithm. It is shown in Figure 6.30 that electrons have many entries that contribute significant energies. In order to deal with this, all events with electrons are neglected and do not enter the following plots. Distributions of soft terms and

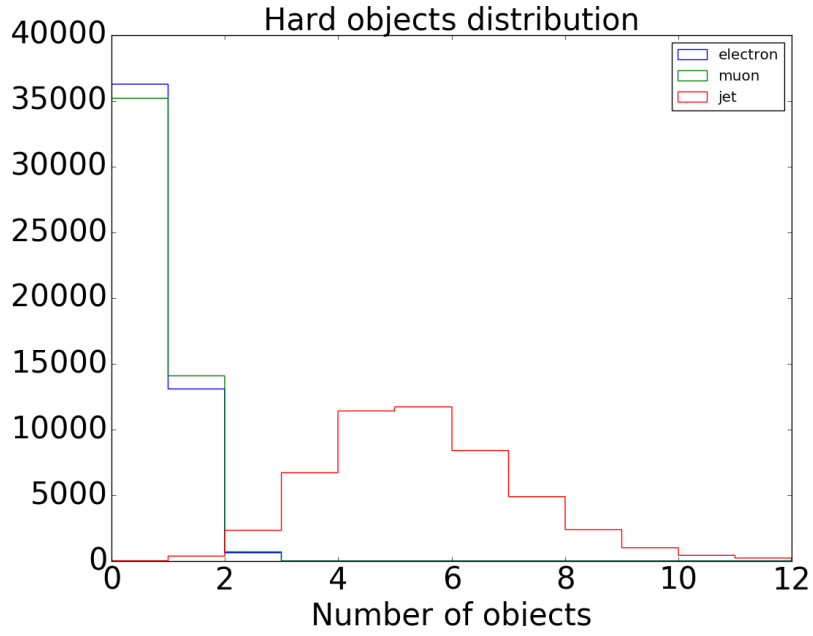


Figure 6.29: Hard object multiplicity in $t\bar{t}$ events. (PFlow version)

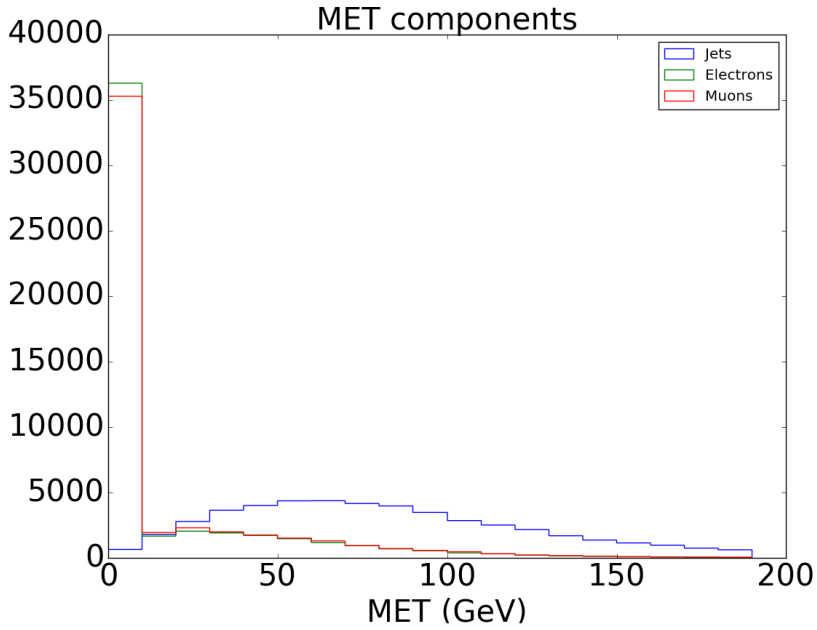


Figure 6.30: PFlow-based MET components in $t\bar{t}$ events.

MET are shown in Figure 6.31 and Figure 6.32 respectively. The PST is shown to be on the same scale as CST after adding the correction term PAT. But the scale of soft terms are much lower than the overall MET so the scale of the MET appears to not

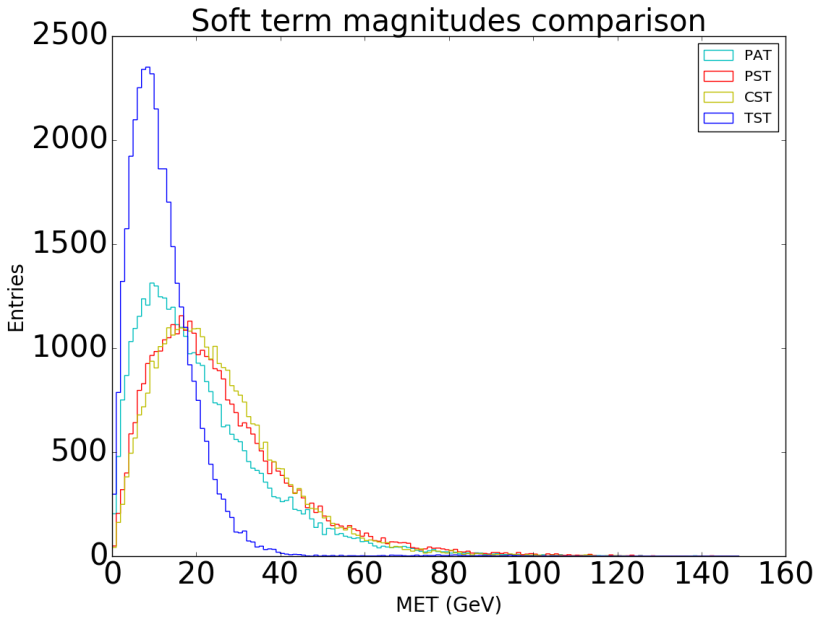


Figure 6.31: Distributions of PFlow-based soft terms in $t\bar{t}$ events.

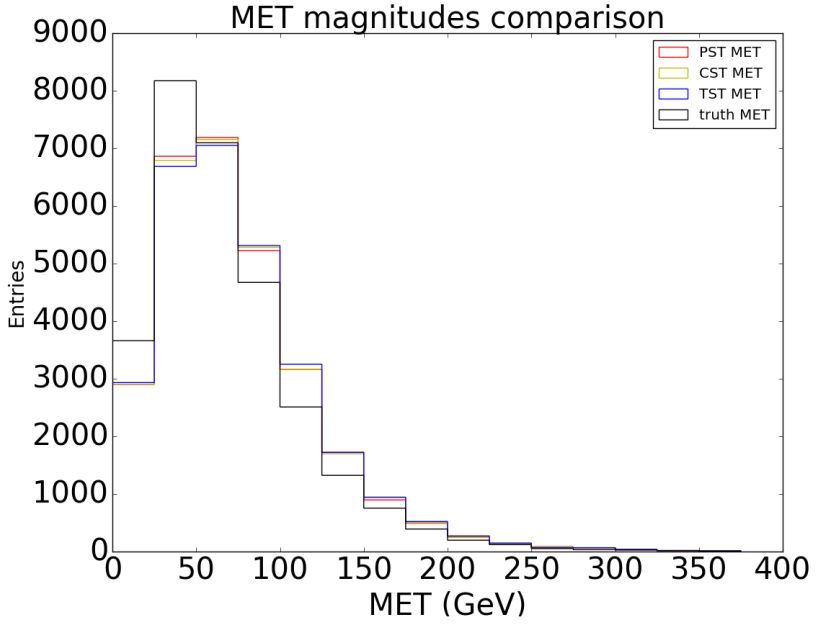


Figure 6.32: PFlow-based MET distributions in $t\bar{t}$ events.

be heavily influenced by the soft term. The large magnitude of PST is not expected based on the previous results in other event topologies. It is worthwhile to check parameters in the fit to see if the fit is working properly. The covariance terms and jet variance are shown in Figure 6.33. Both $V_{11}$ and $V_{22}$ are defined with the second

power of a trigonometric function as shown in Equation 5.19, so they turn out to be non-negative. $V_{12}$ is defined as the multiplication of sine and cosine function which allows it to be negative. It is clear that the two diagonal covariance terms are about the same magnitudes, as expected.

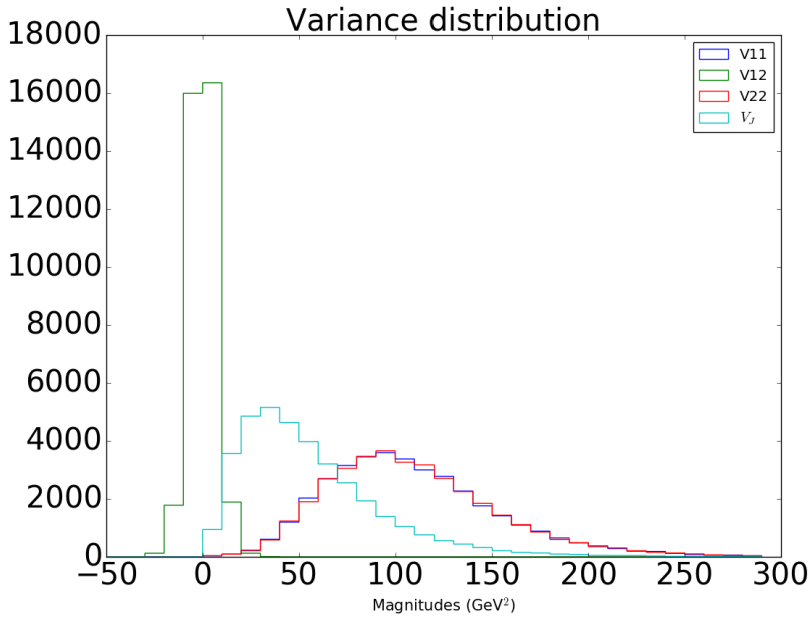


Figure 6.33: Covariance matrix components in $t\bar{t}$ events

Resolutions of different algorithms against $N_{pv}$ and truth MET in $t\bar{t}$ events are shown in Figures 6.34 and 6.35. One can see that the CST MET is performing much better than its resolution performance in $Z^0 \rightarrow \mu^+\mu^-$ events. This could be due to the increased hard scatter activities in $t\bar{t}$ events. As the proportion of hard scatter energy outside HS jets increases, the resolution gain of using CST could be maintained up to a higher $N_{pv}$ value. The resolution of PST MET is actually closer to the CST MET in $t\bar{t}$ events as shown in Figure 6.34. Also, the PSD scale plot is shown in Figure 6.36 with similar responses from all three algorithms.

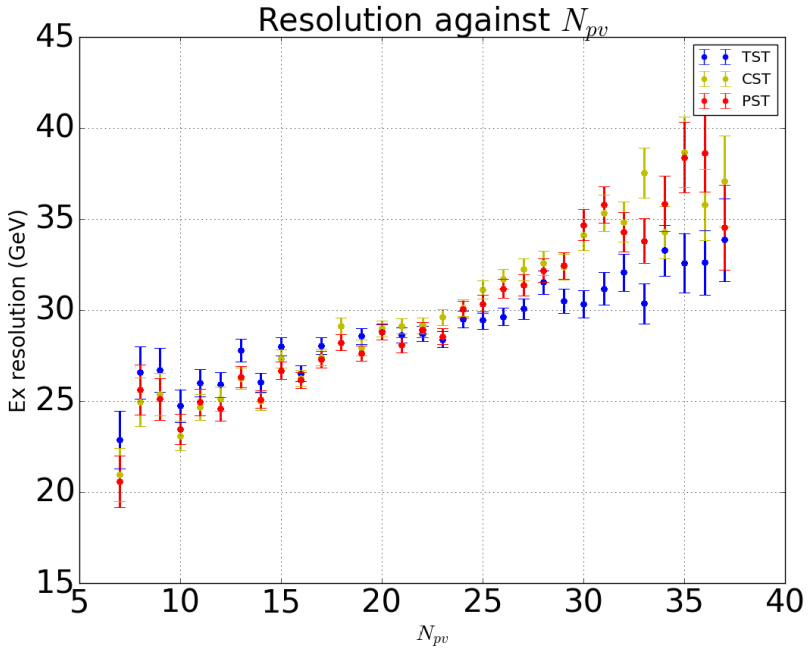


Figure 6.34: Resolution comparisons in $t\bar{t}$ events. (PFlow version)

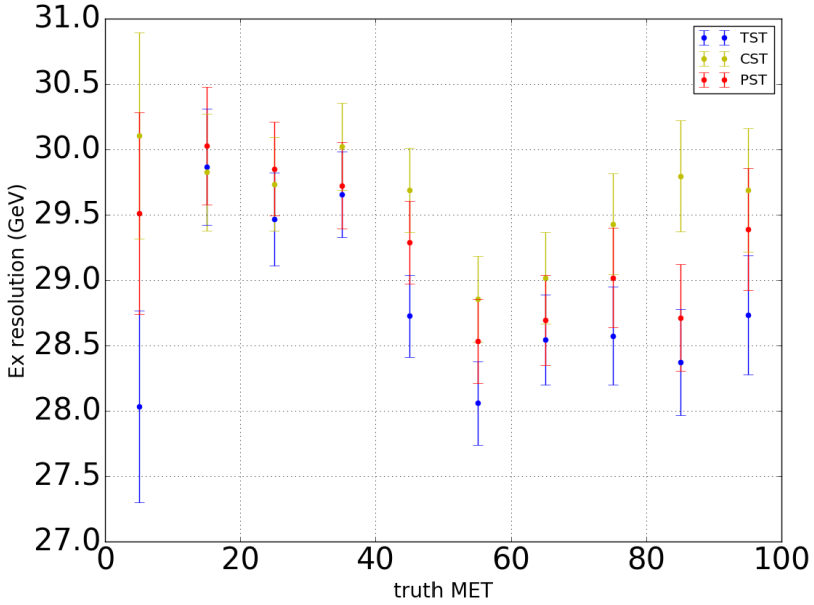


Figure 6.35: Resolution of PFlow-based MET versus truth MET in $t\bar{t}$ events

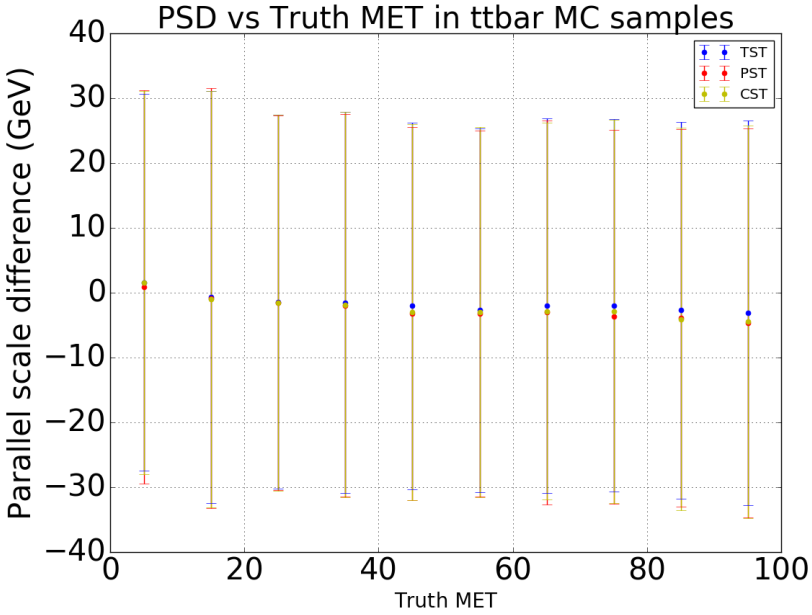


Figure 6.36: PSD vs truth MET in $t\bar{t}$ events. (PFlow version)

# Chapter 7

# Alternative PUfit designs

Other than the two PUfit versions described in Chapter 5, several other versions with new features were also designed, and preliminary test results were obtained. These new features will be briefly introduced and discussed in this section. Most of these versions still need revisions and validations to investigate the potential gain of adopting these new features.

## 7.1 Scale factor

The scale factor is a dynamically determined variable which scales the track soft term. The assumption is that neutral soft activity is highly correlated with charged soft objects in terms of the position in the $\eta - \phi$ plane.

Let the energy ratio between neutral and charged soft objects in a given event be denoted by $f$. Given the TST which describes charged soft objects, the scale factor $\omega$ is inserted to account for neural soft signals.

$$\vec{E}_T^{ST} = (1+f)\vec{E}_T^{TST} = \omega\vec{E}_T^{TST} \tag{7.1}$$

This omega factor is designed to include all soft object signals except for those outside the track's acceptance region since these are not in the original TST. Then it is easy to merge this with the PUfit algorithm described in Chapter 5 so that $\omega$ could also be determined based on constraints in the fit. Specifically, the global constraint would include omega as shown below.

$$\sum_j^{\text{clus}} \vec{E}_{T_j} - \omega \sum_j^{C} \vec{p}_{T_j}^{\text{HS}} + \sum_k^{j} \vec{\mathcal{E}}_k = 0 \tag{7.2}$$

Similarly, one could add the constraint that the expectation value of $\omega$ is $\omega_0 = 1.5$

given by the isospin symmetry. The $\chi^2$ is then defined as the following.

$$\Delta = \begin{pmatrix} (\sum_{j}^{\text{clus}} \vec{E}_{T_j} \cos\phi_j - \omega \sum_{j}^{C} \vec{p}^{\,\text{HS}}_{T_j} \cos\phi_j) + \sum_{k=1}^{m} \mathcal{E}_k \cos\phi_k \\ (\sum_{j}^{\text{clus}} \vec{E}_{T_j} \sin\phi_j - \omega \sum_{j}^{C} \vec{p}^{\,\text{HS}}_{T_j} \sin\phi_j) + \sum_{k=1}^{m} \mathcal{E}_k \sin\phi_k \\ \langle\rho\rangle A_1 - \mathcal{E}_1 \\ \vdots \\ \langle\rho\rangle A_J - \mathcal{E}_J \\ \omega_0 - \omega \end{pmatrix} \tag{7.3}$$

Since the amount of energy deposited by charged particles could be 0 due to statistical fluctuation while it is half of the neutral on average, the standard deviation of $\omega$ is about 0.5 and the variance is 0.25. The covariance matrix will also add one more entry in the diagonal with the magnitude of $V_\omega = 0.25$.

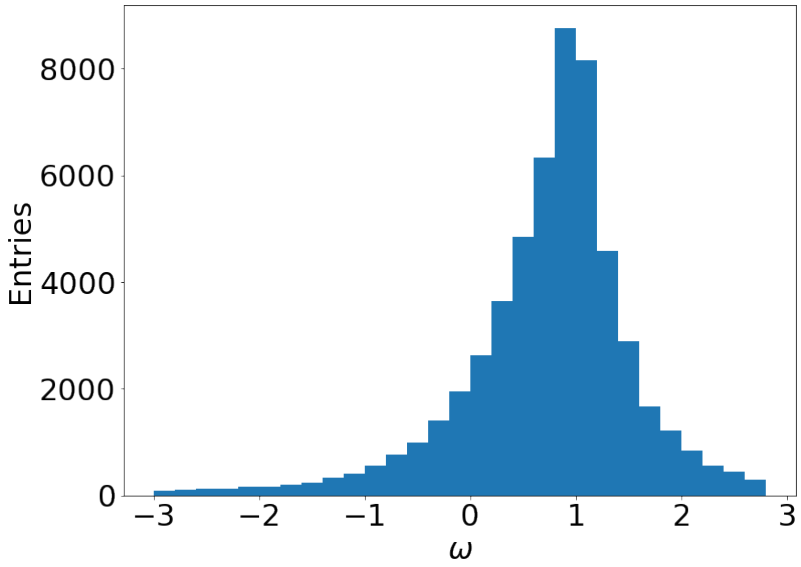


Figure 7.1: Distribution of $\omega$ determined by the fit

However, the implementation of $\omega$ did not achieve any noticeable improvement in terms of resolution. The reason is that the isospin expectation is regarding the charged to neutral ratio of the scalar energy deposition, while the soft term is a vector sum whose magnitude is a small fraction of the scalar sum due to cancellations. The distribution of $\omega$ is shown in 7.1. It is clear that the distribution has values that differ substantially from 1.5 which means our assumption is questionable.

It would be ideal to investigate the truth soft energy distribution in Monte Carlo samples so that the $\omega$ implementation could be fine-tuned. This involves using the Digi-truth algorithm which outputs the energy composition of a given cluster so that one could easily determine the percentage of HS energy in that cluster. Digi-truth is not yet available, but it would be a good starting point for work in the future.

## 7.2 $\eta$ re-scaling

The overall pile-up energy deposit is not uniform in $\eta$, so the expected pile-up energy deposit would be dependent on that $\eta$ values of the given jet. A pre-defined function $f(\eta)$ that encodes $\eta$ dependence is used to describe the shape of PU energy deposit in a given event. Since the overall PU energy scale varies event by event, $f(\eta)$ is scaled as described below to obtain optimal results.

The $(N-m)$ non-overlapping towers are further grouped based on their $\eta$ values, and the average PU deposits are calculated for each $\eta$ slice. Then we compare each tower's PU energy with the expected PU energy to obtain the normalization factor G:

$$G = \frac{1}{N-m}\sum_{j}^{N-m}\frac{E_{T_j}^{\text{tower}} - \omega p_{T_j}^{\text{HS}}}{f(\eta_j)} \tag{7.4}$$

Then $Gf(\eta)\frac{A_{jet}}{A_{tower}}$ would describe the PU deposition in the event, and the estimated PU energy under HS jets can be computed accordingly based their $\eta$ values.

The variance of the estimated PU energy is computed from averaging the variance of each tower that does not overlap with jets with respect to its corresponding $Gf(\eta)$. Note that the variance also has $\eta$ dependence which is encoded in the predefined function $f_\eta^V$, so the same procedure can be used again to describe the variance analytically. The normalization factor $G^V$ is determined by:

$$G^V = \frac{1}{N-m}\sum_{j}^{N-m}\frac{\left(E_{T_i}^{\text{tower}} - \omega p_{T_i}^{\text{HS}} - Gf(\eta_i)\right)^2}{f^V(\eta_j)} \tag{7.5}$$

So for a jet at $\eta_k$, the expected PU energy and its variance is given by:

$$\begin{aligned} E_k^{PU} &= Gf(\eta_k)\frac{A_{jet}}{A_{tower}} \\ V_k &= G^V f^V(\eta_k)\frac{A_{jet}}{A_{tower}} \end{aligned} \tag{7.6}$$

Using the fit, we can determine the PU value for this jet, $\mathcal{E}_k$, to balance the overall PU contributions in a event. Details of the numerical solution are shown in

the appendix D. We can adjust the pileup imbalance term by taking the difference

$$-\sum_{k=1}^{J}\left(\mathcal{E}_k - Gf(\eta_k)\frac{A_{jet}}{A_{tower}}\right) \tag{7.7}$$

This is the fake PU contribution to $\vec{E}_T^{\text{miss}}$. Since this cannot be accounted for by any HS jets, leptons or the soft term, it needs to be subtracted from the final $\vec{E}_T^{\text{miss}}$ as an extra term to avoid the fake contribution from PU. Then the Pileup-imbalance Adjustment Term (PAT), $E_T^{\text{PAT}}$, is given by the negative of the above term

$$\vec{E}_T^{\text{PAT}} = \sum_{k=1}^{J}\left(\mathcal{E}_k - Gf(\eta_k)\frac{A_{jet}}{A_{tower}}\right) \tag{7.8}$$

The shape of pile-up energy deposit in $\eta$ is shown in Figure 7.2. The average energy deposit in a given $\eta$ slice is computed in all events, and the figure displays the overall accumulated energy. The energy deposit drops sharply as $\eta$ increases up to around 2. This sudden drop is the result of several factors: non-linearity of $\eta$, topocluster sizes and also the overlap of the barrel and endcap calorimeters.

Including a $\eta$-dependence feature should improve resolution since it provides a more accurate description of the pile-up energy. But as Figure 7.3 shows, the resolution deteriorates significantly after the $\eta$ - dependence is implemented. This was also the case when $\eta$ dependence was being implemented in the trigger version previous to this study. A possible reason for this behavior is that even though the expected shape is considered, the energy deposit in a given event fluctuates severely so that the scaling breaks down. Also, a flat eta-dependence function will not emphasize the measurements at large $\eta$ where the detector response is less precise.

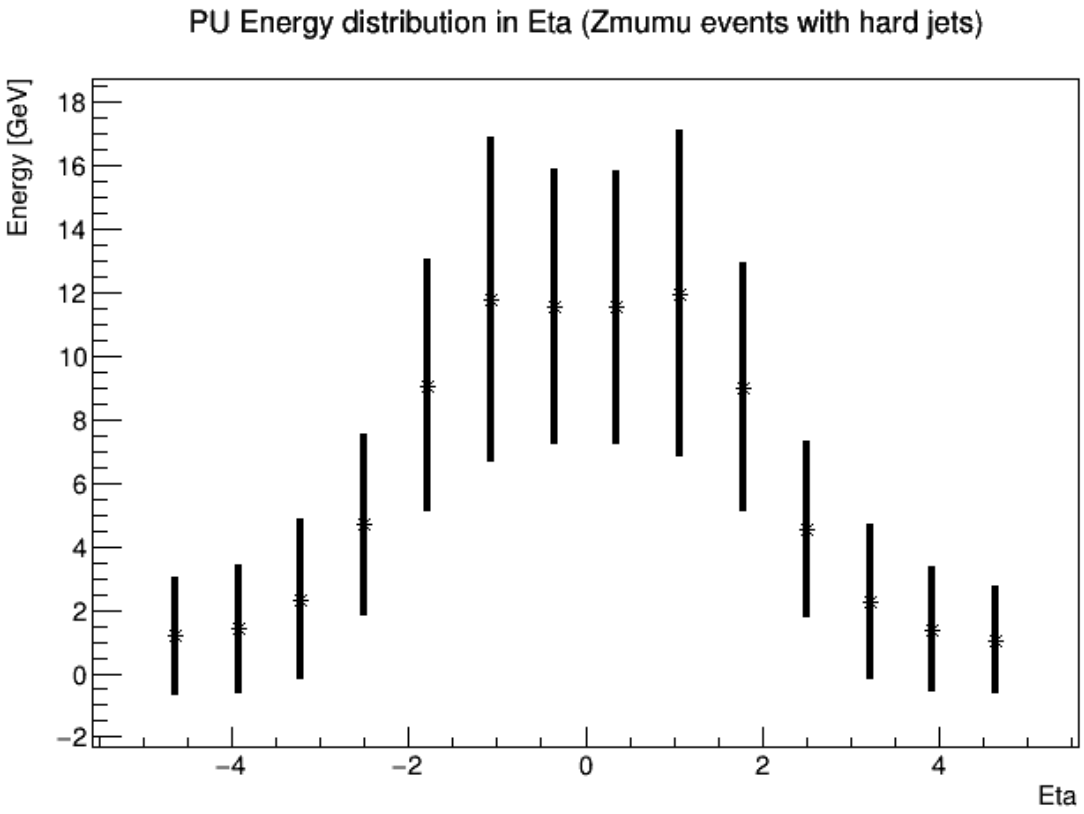


Figure 7.2: Distribution of PU energy deposit in $\eta$. The distribution is obtained by accumulating entries from all events.

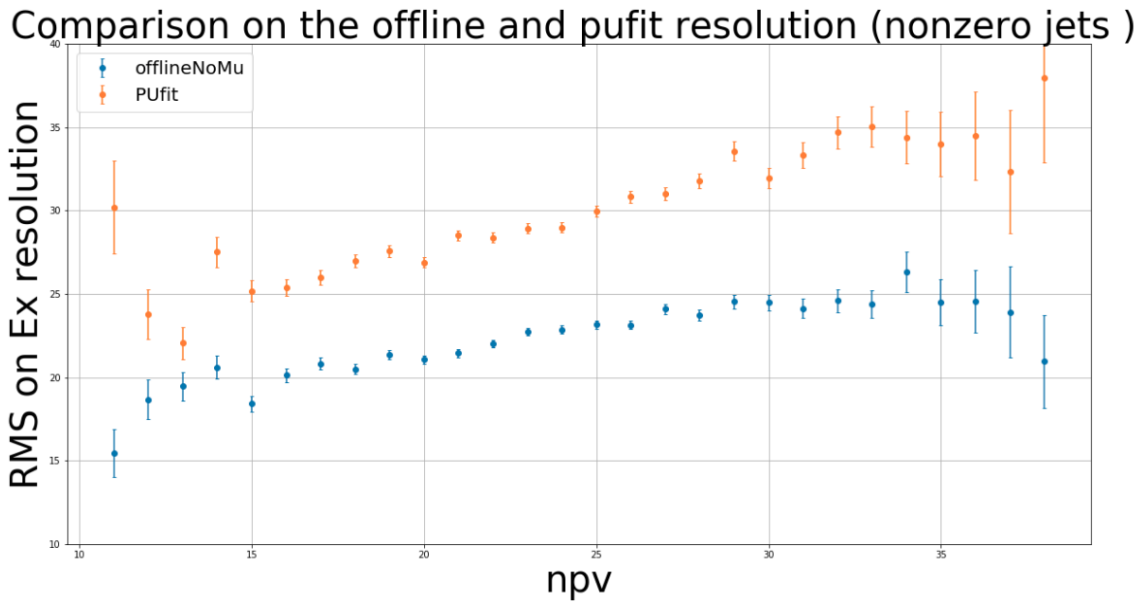


Figure 7.3: Resolution against $N_{pv}$ of eta-dependent PUfit in $Z^0 \rightarrow \mu^+\mu^-$ events

# Chapter 8

# Conclusion

The performance of the PUfit algorithm in the offline framework is tested and discussed in this thesis. Different versions of the PUfit algorithm with various input types were put to test, and their resolutions are compared. Overall, three event topologies were used in the testing. $Z^0 \rightarrow \mu^+\mu^-$ is the main topology with the Z boson being the imaginary $E_T^{\text{miss}}$ source. Monte Carlo samples of $t\bar{t}$ and $W \rightarrow \mu\nu_\mu$ are also used.

The magnitude of the PAT correction term offered by the PUfit algorithm is shown to be of similar scale to the track soft term in the offline $E_T^{\text{miss}}$ . However, the overall scale of PST $E_T^{\text{miss}}$ does not increase on average after vectorially adding the PAT. The scale is lowered in most datasets, meaning that the PAT is mostly negatively correlated with the offline $E_T^{\text{miss}}$ .

The two different offline $E_T^{\text{miss}}$ with TST and CST are both compared to the PST $E_T^{\text{miss}}$ determined by the PUfit. PST $E_T^{\text{miss}}$ with topocluster inputs showed limited resolution improvement and a slightly lower scale in the region of $15 < \text{N}_{pv} < 25$ in $Z^0 \rightarrow \mu^+\mu^-$ events. The resolution of the PFlow PST $E_T^{\text{miss}}$ exhibit similar resolution behavior compared to the topocluster one. In the $Z^0 \rightarrow \mu^+\mu^-$ Monte Carlo sample, the PST $E_T^{\text{miss}}$ showed consistent improvement of about 1 GeV over most of the $\text{N}_{pv}$ domain. However, no such clear improvements were obtained in $W \rightarrow \mu\nu_\mu$ and $t\bar{t}$ samples.

The PUfit algorithm adds neutral energy to the soft term without the severe pile-up dependence observed in CST. But the PUfit soft term does not yield significant improvements compared to the TST, and its behavior in the $t\bar{t}$ sample needs further investigation. The PUfit algorithm is expected to offer improvements in events with relatively lower truth MET. In these events, the MET would not be dominated by

the jet term and thus leaving room for the corrections on the soft term to make a larger impact.

In the future, it will be interesting to check if the large jet multiplicities will cause are associated with the performance degradation seen in the $t\bar{t}$ sample. Also, the $\eta$ dependence of pile-up energy distribution could be the reason why the gain in resolution is limited. A more sophisticated method of dealing with $\eta$ dependence needs to be proposed and tested. And finally, the PUfit algorithm in the trigger system had shown great improvement in events with extremely large pile-up activities($N_{pv} \approx 200$). It will be worthwhile to check the performance of PUfit in offline computation when the high pile-up simulation is available.

# Bibliography

[1] M Tanabashi et.al (Particle Data Group). Review of particle physics. *Phys. Rev. D*, 98:030001, Aug 2018.

[2] F. Zwicky. Die Rotverschiebung von extragalaktischen Nebeln. *Helv. Phys. Acta*, 6:110–127, 1933. [Gen. Rel. Grav.41,207(2009)].

[3] Jonathan L. Feng. Dark Matter Candidates from Particle Physics and Methods of Detection. *Ann. Rev. Astron. Astrophys.*, 48:495–545, 2010.

[4] Philip Bechtle, Tilman Plehn, and Christian Sander. Supersymmetry. In Thomas Schrner-Sadenius, editor, *The Large Hadron Collider: Harvest of Run 1*, pages 421–462. 2015.

[5] Morad Aaboud et al. Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13$ TeV. *Eur. Phys. J.*, C78(11):903, 2018.

[6] Georges Aad et al. Observation and measurement of Higgs boson decays to WW$^*$ with the ATLAS detector. *Phys. Rev.*, D92(1):012006, 2015.

[7] Georges Aad et al. Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1. *Eur. Phys. J.*, C77:490, 2017.

[8] ATLAS collaboration. *Luminosity Public Results Run2*. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2 Pileup_Interactions_and_Data_AN2.

[9] ATLAS collaboration. *Vetex performance in 2018 data*. `https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/IDTR-2018-006/`.

[10] Georges Aad et al. Muon reconstruction performance of the ATLAS detector in protonproton collision data at $\sqrt{s}$ =13 TeV. *Eur. Phys. J.*, C76(5):292, 2016.

[11] Donald E. Groom, Nikolai V. Mokhov, and Sergei I. Striganov. Muon stopping power and range tables 10-MeV to 100-TeV. *Atom. Data Nucl. Data Tabl.*, 78:183–356, 2001.

[12] G. Aad et al. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008.

[13] *ATLAS inner detector: Technical Design Report, 1.* Technical Design Report ATLAS. CERN, Geneva, 1997.

[14] Georges Aad et al. Operation and performance of the ATLAS semiconductor tracker. *JINST*, 9:P08009, 2014.

[15] ATLAS collaboration. *Vertex Reconstruction Performance of the ATLAS Detector at 13 TeV.* `https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2015-026/`.

[16] The ATLAS Collaboration. The ATLAS experiment at the CERN large hadron collider. *Journal of Instrumentation*, 3(08):S08003–S08003, aug 2008.

[17] Georges Aad et al. Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1. *Eur. Phys. J.*, C77:490, 2017.

[18] Heather Abreu et al. Performance of the electronic readout of the atlas liquid argon calorimeters. *Journal of Instrumentation*, 5, 09 2010.

[19] Aranzazu Ruiz-Martinez and ATLAS Collaboration. The Run-2 ATLAS Trigger System. Technical Report ATL-DAQ-PROC-2016-003, CERN, Geneva, Feb 2016.

[20] The ATLAS Collaboration. Electron and photon energy calibration with the atlas detector using lhc run 1 data. *The European Physical Journal C*, 74(10):3071, Oct 2014.

[21] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-$k_t$ jet clustering algorithm. *JHEP*, 04:063, 2008.

[22] M. Aaboud et al. Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *Phys. Rev.*, D96(7):072002, 2017.

[23] Georges Aad et al. Performance of pile-up mitigation techniques for jets in $pp$ collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector. *Eur. Phys. J.*, C76(11):581, 2016.

[24] J Lacey, A Pilkington, Will Leight, M Queitsch-Maitland, G Vardanyan, and F Ruehr. Data-driven determination of the energy scale and resolution of jets reconstructed in the ATLAS calorimeters using dijet and multijet events at $\sqrt{s} = 8\ TeV$. Technical Report ATLAS-COM-CONF-2014-068, CERN, Geneva, Aug 2014.

[25] Morad Aaboud et al. Jet reconstruction and performance using particle flow with the ATLAS Detector. *Eur. Phys. J.*, C77(7):466, 2017.

# Appendix A

# Glossary of terms

Many terminologies have been introduced in this thesis to discuss the performance of $E_T^{\text{miss}}$ algorithm. In order to help readers understand the nomenclature, a brief glossary of terms is given below.

- **Transverse momentum ($p_T$):**

  The transverse momentum is defined as the projection of the three-dimensional momentum in the transverse plane perpendicular to the beam axis in the ATLAS detector. $\vec{p}_T$ is a 2-dimensional vector in the transverse plane and $p_T$ is the magnitude of this 2-d vector.

- **Primary vertex:**

  A primary vertex is the location of a reconstructed proton-proton collision.

- **Hard scatter:**

  The hardest interaction of all recorded in a given event, meaning that it exhibits the largest momentum transfer within the two proton system. The primary vertex associated with hard scatter is referred to as the hard scatter vertex. The hard-scatter vertex is defined as the one with the largest sum of $p_T^2$ of tracks associated with it.

- **Pile up:**

  Pile-up interactions are low momentum transfer proton-proton collisions that accompany the hard scatter. All vertices other than the hard scatter vertex are assumed to be pile-up vertices. Objects produced in the pile-up vertex are referred to as pile-up objects.

- **Missing transverse momentum ($E_T^{\text{miss}}$ ; MET):**

  The missing transverse momentum, often referred to as $E_T^{\text{miss}}$ or the MET, is defined as the negative vectorial sum of $\vec{p}_T$ of all hard scatter objects. $\vec{E}_T^{\text{miss}}$ is the 2-dimensional vector in the transverse plane and $E_T^{\text{miss}}$ is the magnitude of that vector.

- **Hard objects:**

  Reconstructed objects that pass offline object selection. These objects are mostly of high transverse momentum and thus assumed to be originating from the hard scatter vertex.

- **Soft term:**

  The soft term is a component of the $E_T^{\text{miss}}$ . It is defined as the vectorial sum of $\vec{p}_T$ of objects produced in the hard scatter collision but not classified as hard objects.

- **Track:**

  Tracks are signals of charged particles recorded by the inner detector tracking system. They offer information on the trajectory and also a precise measurement of momentum, especially when the momentum is low. Given the trajectory, it is also possible to associate a track with one of the primary vertices.

- **Topocluster:** A cluster of topologically connected calorimeter cell signals. It represents the full or large fraction of a given particle's energy deposit in the calorimeter. Both charged and neutral particles are reconstructed using topoclusters. They offer better measurements of momentum than the tracker for high energy particles.

# Appendix B

# Datasets details

## $\mathbf{Z^0 \to \mu^+\mu^-}$ data:

- **Name**: data17_13TeV.00339435.physics_Main.merge.DAOD_ZMUMU.f904_m1831_f913_m1920
- **Run number**: 339435
- **Center of mass energy**: 13 TeV
- **Total number of Events in the dataset**: 4238896
- **Instantaneous luminosity range**: 5488 - 15330 $(10^{30} cm^{-2} s^{-1})$
- **average expected number of interactions per beam crossing**($\mu$): 20.9 - 58.4

## $\mathbf{Z^0 \to \mu^+\mu^-}$ Monte Carlo sample:

- **Name**: data17_13TeV.00339435.physics_Main.merge.DAOD_ZMUMU.f904_m1831_f913_m1920
- **Run number**: 339435
- **Center of mass energy**: 13 TeV
- **Total number of Events in the dataset**: 4238896
- **Instantaneous luminosity range**: 5488 - 15330 $(10^{30} cm^{-2} s^{-1})$

- **average expected number of interactions per beam crossing**($\mu$): 20.9 - 58.4

## $W \to \mu\nu_\mu$ Monte Carlo sample:

- **Name:** mc16_13TeV.363436.Sherpa_NNPDF30NNLO_Wmunu_Pt0_70 _CVetoBVeto.merge.AOD.e4715_e5984_s3126_r10724_r10726
- **Center of mass energy:** 13 TeV
- **Total number of Events in the dataset:** 20000
- **Pile-up condition:** Same as the LHC running condition.
- **Cross section**:19.984 nb

## $t\bar{t}$ Monte Carlo sample:

- **Name**: mc16_13TeV.410470.PhPy8EG_A14_ttbar_hdamp258p75_ nonallhad.merge.AOD.e6337_e5984_s3126_r10724_r10726
- **Center of mass energy:** 13 TeV
- **Total number of Events in the dataset:**159696000
- **Number of events used:**36281
- **Pile-up condition:** Same as the LHC running condition.
- **Cross section**:729.770 pb

# Appendix C

# Additional angular correlation plots

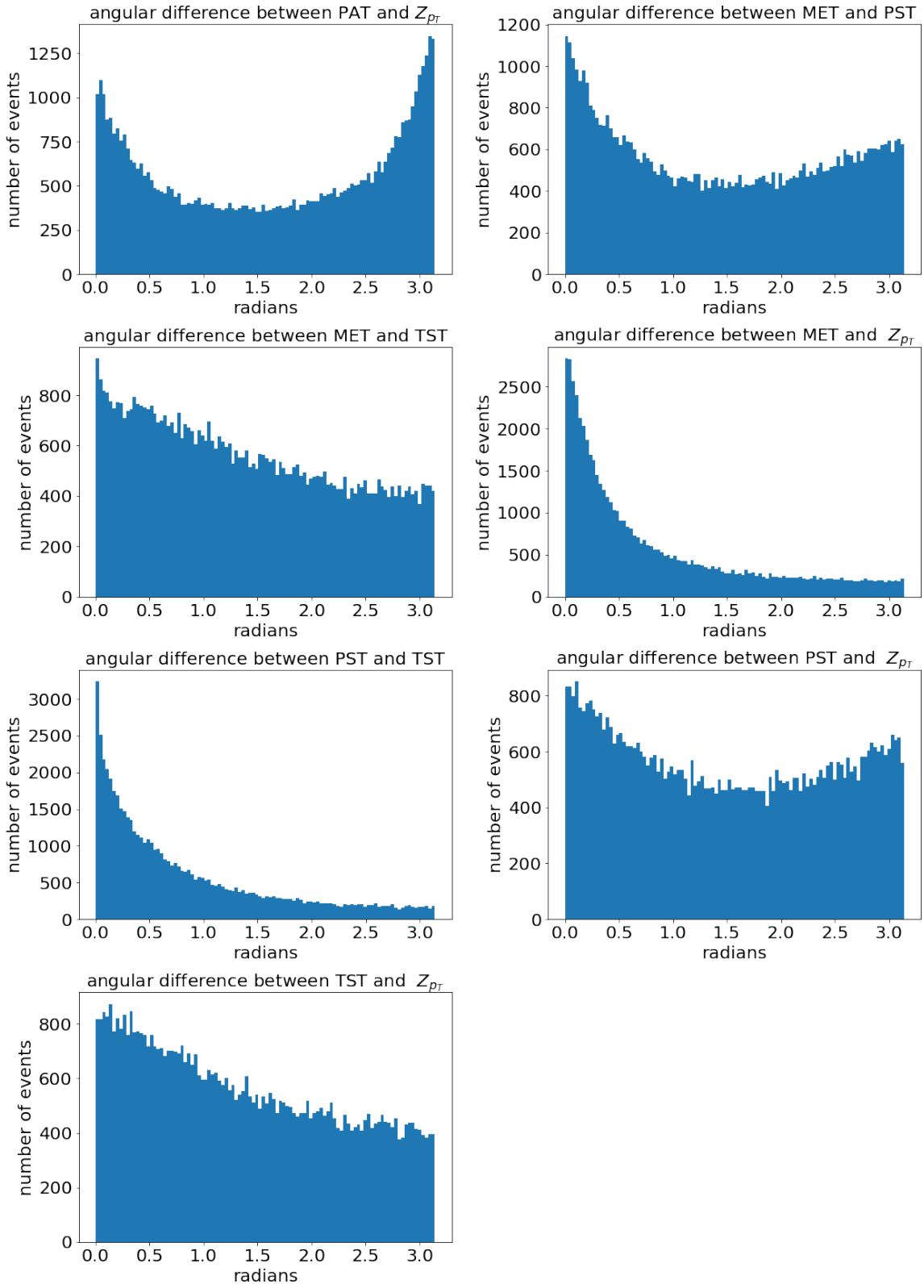


Figure C.1: Angular difference between each MET term in $Z^0 \rightarrow \mu^+\mu^-$ events. (2) (PFlow version)

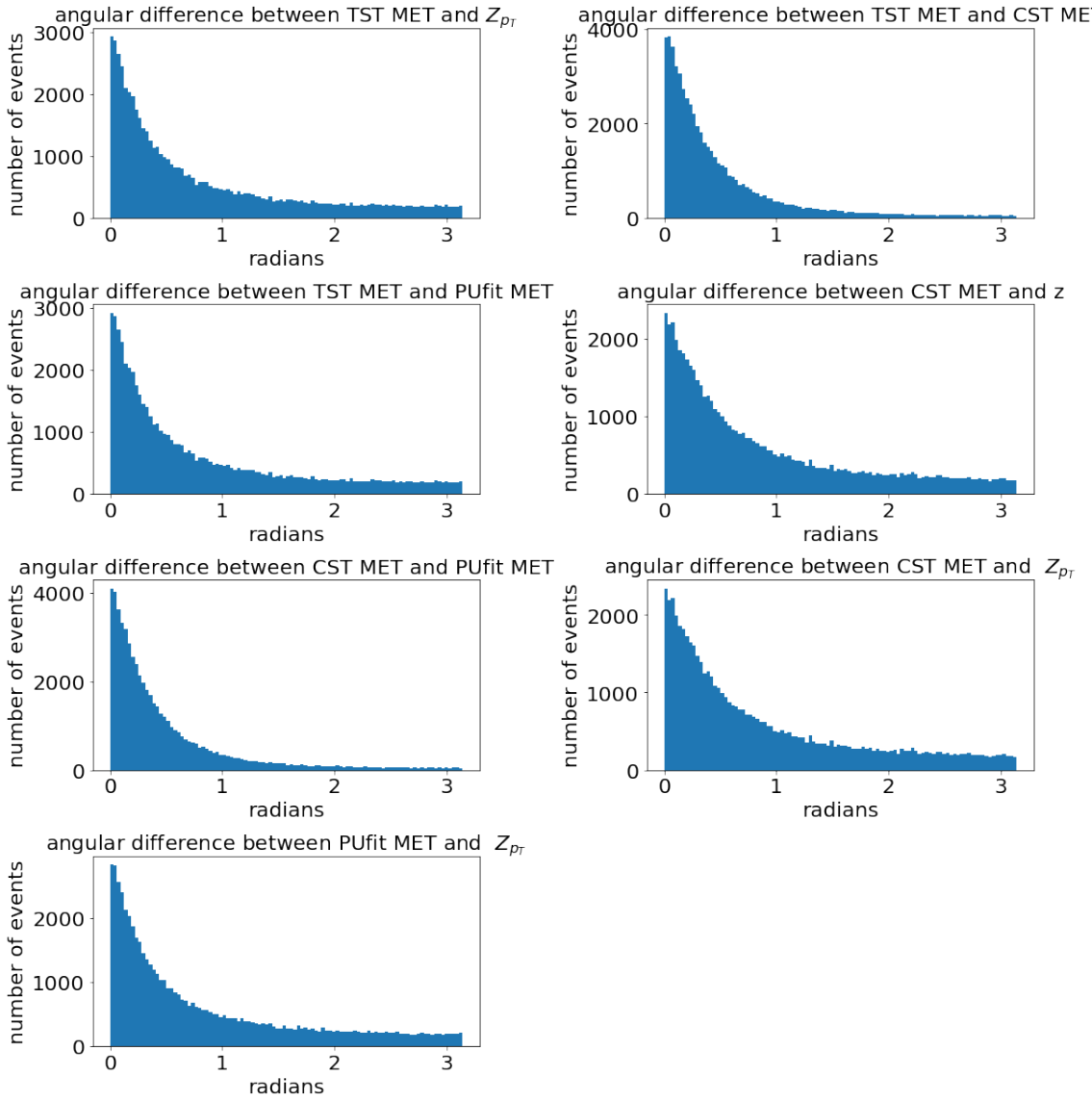


Figure C.2: Angular difference between each MET term in $Z^0 \rightarrow \mu^+\mu^-$ events. (3) (PFlow version)

# Appendix D

# $\eta$ - dependent PUfit solution

### D.0.1 Formulation

We construct a $\chi^2$ function:

$$\chi^2(\mathcal{E}_1, \ldots, \mathcal{E}_J, \omega) = \Delta^T V^{-1} \Delta \tag{D.1}$$

where the variables $\mathcal{E}_k$ are the fitted parameters corresponding to the $m$ high $p_T$ towers and $\omega$ is the HS soft contribution coefficient, determined by minimizing the $\chi^2$ function. The $(2 + J + 1)$-dimensional vector $\Delta$ is given by

$$\Delta = \begin{pmatrix} \sum_j^{\text{clus}} E_{T_j} \cos\phi_j - \omega \sum_j^{C} p_{T_j}^{\text{HS}} \cos\phi_j + \sum_k^{J} \mathcal{E}_k \cos\phi_k \\ \sum_j^{\text{clus}} E_{T_j} \sin\phi_j - \omega \sum_j^{C} p_{T_j}^{\text{HS}} \sin\phi_j + \sum_k^{J} \mathcal{E}_k \sin\phi_k \\ Gf(\eta_1) - \mathcal{E}_1 \\ \vdots \\ Gf(\eta_J) - \mathcal{E}_J \\ \omega_0 - \omega \end{pmatrix} \tag{D.2}$$

and $V$ is the associated covariance matrix. The angles $\phi_j$ ($\phi_k$) are the azimuthal directions of the corresponding low $p_T$ tower. The first two lines impose constraints corresponding to Equation 5.12. Each remaining line corresponds to the $\mathcal{E}_k$ and $\omega$ are close to their expected values discussed above. The average energy is given by the product of the jet area $A_k$ and the average $E_T$ density, $\langle\rho\rangle$.

The full covariance is

$$V = \begin{pmatrix} V_{11} & V_{12} & 0 & 0 & \dots & 0 & 0 \\ V_{21} & V_{22} & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & V_1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \ddots & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & V_J & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & V_\omega \end{pmatrix} \tag{D.3}$$

where the upper $2 \times 2$ submatrix is given by

$$\begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix} = \begin{pmatrix} \sum_j^O \sigma_j^2 \cos^2\phi_j & \sum_j^O \sigma_j^2 \cos\phi_j \sin\phi_j \\ \sum_j^O \sigma_j^2 \cos\phi_j \sin\phi_j & \sum_j^O \sigma_j^2 \sin^2\phi_j \end{pmatrix}, \tag{D.4}$$

where $\sum_j^O = \sum_j^{\text{clus}} - \sum_j^c$

Minimizing this $\chi^2$ function with respect to the $\mathcal{E}_k$ and $\omega$ provides the PST from Equation 5.6.

### D.0.2 Minimization

The minimization is done analytically rather than numerically using

$$\frac{\partial \chi^2}{\partial \mathcal{E}_k} = 0 \tag{D.5}$$

$$\frac{\partial \chi^2}{\partial \omega} = 0 \tag{D.6}$$

The inverse of the covariance matrix is

$$V^{-1} = \begin{pmatrix} V_{22}/d & -V_{12}/d & 0 & 0 & \dots & 0 & 0 \\ -V_{21}/d & V_{11}/d & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1/V_1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \ddots & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & 0 \\ 0 & 0 & 0 & 0 & \dots & 1/V_J & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 & 1/V_\omega \end{pmatrix} \tag{D.7}$$

where $d = V_{11}V_{22} - V_{12}V_{21}$ is the determinant of the 2 by 2 part. Therefore,

$$\chi^2 = \frac{V_{22}}{d}\Delta_1^2 + \frac{V_{11}}{d}\Delta_2^2 - 2\frac{V_{12}}{d}\Delta_1\Delta_2 + \sum_{k=1}^{J} \frac{\Delta_{E_k}^2}{V_k} + \frac{\Delta_\omega^2}{V_\omega} \tag{D.8}$$

where

$$\begin{aligned}
\Delta_1 &= \sum_j^{\text{clus}} E_{T_j}\cos\phi_j - \omega\sum_j^{C} p_{T_j}^{\text{HS}}\cos\phi_j + \sum_k^{J}\mathcal{E}_k\cos\phi_k \\
&= E_x^0 - \omega P_x^0 + \sum_{k=1}^{m}\mathcal{E}_k\cos\phi_k \\
\Delta_2 &= \sum_j^{\text{clus}} E_{T_j}\sin\phi_j - \omega\sum_j^{C} p_{T_j}^{\text{HS}}\sin\phi_j + \sum_k^{J}\mathcal{E}_k\sin\phi_k \\
&= E_y^0 - \omega P_y^0 + \sum_{k=1}^{m}\mathcal{E}_k\sin\phi_k \\
\Delta_{E_k} &= Gf(\eta_k)\frac{\text{A}_{\text{jet}}}{\text{A}_{\text{tower}}} - \mathcal{E}_k \\
\Delta_\omega &= \omega_0 - \omega
\end{aligned} \tag{D.9}$$

where

$$\begin{aligned}
E_x^0 &\equiv \sum_j^{\text{clus}} E_{T_j}\cos\phi_j \\
E_y^0 &\equiv \sum_j^{\text{clus}} E_{T_j}\sin\phi_j \\
P_x^0 &\equiv \sum_j^{C} p_{T_j}^{\text{HS}}\cos\phi_j \\
P_y^0 &\equiv \sum_j^{C} p_{T_j}^{\text{HS}}\sin\phi_j
\end{aligned} \tag{D.10}$$

Differentiation gives:

$$\begin{aligned}
\frac{\partial\chi^2}{\partial\mathcal{E}_i} = 2\frac{V_{22}}{d}\Delta_1\frac{\partial\Delta_1}{\partial\mathcal{E}_i} + 2\frac{V_{11}}{d}\Delta_2\frac{\partial\Delta_2}{\partial\mathcal{E}_i} - 2\frac{V_{12}}{d}\Delta_2\frac{\partial\Delta_1}{\partial\mathcal{E}_i} - 2\frac{V_{12}}{d}\Delta_1\frac{\partial\Delta_2}{\partial\mathcal{E}_i} \\
+ \sum_k \frac{2\Delta_{E_k}}{V_k^{\text{tower}}}\frac{\partial\Delta_{E_k}}{\partial\mathcal{E}_i}
\end{aligned} \tag{D.11}$$

and

$$\begin{aligned}
\frac{\partial\chi^2}{\partial\omega} = 2\frac{V_{22}}{d}\Delta_1\frac{\partial\Delta_1}{\partial\omega} + 2\frac{V_{11}}{d}\Delta_2\frac{\partial\Delta_2}{\partial\omega} - 2\frac{V_{12}}{d}\Delta_2\frac{\partial\Delta_1}{\partial\omega} - 2\frac{V_{12}}{d}\Delta_1\frac{\partial\Delta_2}{\partial\omega} \\
+ \sum_k \frac{2\Delta_{E_k}}{V_k}\frac{\partial\Delta_{E_k}}{\partial\omega} + \frac{2\Delta_\omega}{V_\omega}\frac{\partial\Delta_\omega}{\partial\omega}
\end{aligned} \tag{D.12}$$

By substituting $\Delta$s, the $\partial\chi^2/\partial\mathcal{E}_1 = 0$ gives

$$\begin{aligned}
&\frac{V_{22}\cos\phi_1 - V_{12}\sin\phi_1}{d}(E_x^0 - \omega P_x^0 + \mathcal{E}_1\cos\phi_1 + \mathcal{E}_2\cos\phi_2 + ... + \mathcal{E}_m\cos\phi_m) \\
+&\frac{V_{11}\sin\phi_1 - V_{12}\cos\phi_1}{d}(E_y^0 - \omega P_y^0 + \mathcal{E}_1\sin\phi_1 + \mathcal{E}_2\sin\phi_2 + ... + \mathcal{E}_m\sin\phi_m) \\
&-\frac{1}{V_1}\left[\psi_k^E - w\psi_k^P - \mathcal{E}_1\right] = 0
\end{aligned} \tag{D.13}$$

where

$$\begin{aligned}
\psi_k^P &= -\frac{\partial\Delta_{E_k}}{\partial\omega} = f(\eta_k)\frac{1}{N-m}\frac{\mathrm{A_{jet}}}{\mathrm{A_{tower}}}\sum_j^{N-m}\frac{p_{T_j}^{HS}}{f(\eta_j)} \\
\psi_k^E &= f(\eta_k)\frac{1}{N-m}\frac{\mathrm{A_{jet}}}{\mathrm{A_{tower}}}\sum_j^{N-m}\frac{E_{T_j}^{\mathrm{tower}}}{f(\eta_j)}
\end{aligned} \tag{D.14}$$

This gives

$$\alpha_1 + \left(\beta_{11} + \frac{1}{V_1}\right)\mathcal{E}_1 + \beta_{12}\mathcal{E}_2 + ... + \beta_{1J}\mathcal{E}_J - \gamma_1\omega = 0 \tag{D.15}$$

where

$$\begin{aligned}
\alpha_k &= \frac{V_{22}\cos\phi_k - V_{12}\sin\phi_k}{d}E_x^0 + \frac{V_{11}\sin\phi_k - V_{12}\cos\phi_k}{d}E_y^0 - \frac{\psi_k^E}{V_k} \\
\beta_{kj} &= \frac{V_{22}\cos\phi_k - V_{12}\sin\phi_k}{d}\cos\phi_j + \frac{V_{11}\sin\phi_k - V_{12}\cos\phi_k}{d}\sin\phi_j \\
\gamma_k &= \frac{V_{22}\cos\phi_k - V_{12}\sin\phi_k}{d}P_x^0 + \frac{V_{11}\sin\phi_k - V_{12}\cos\phi_k}{d}P_y^0 - \frac{\psi_k^P}{V_k}
\end{aligned} \tag{D.16}$$

for $k, j = 1, ..., m$. Similarly $\partial\chi^2/\partial\mathcal{E}_2 = 0$ gives

$$\alpha_2 + \beta_{21}\mathcal{E}_1 + \left(\beta_{22} + \frac{1}{V_2}\right)\mathcal{E}_2 + ... + \beta_{2m}\mathcal{E}_m - \gamma_2\omega = 0 \tag{D.17}$$

Similarly, $\partial\chi^2/\partial\omega = 0$ gives

$$\begin{aligned}
&\frac{V_{22}P_x^0 - V_{12}P_y^0}{d}(E_x^0 - \omega P_x^0 + \mathcal{E}_1\cos\phi_1 + \mathcal{E}_2\cos\phi_2 + ... + \mathcal{E}_J\cos\phi_J) \\
+&\frac{V_{11}P_y^0 - V_{12}P_x^0}{d}(E_y^0 - \omega P_y^0 + \mathcal{E}_1\sin\phi_1 + \mathcal{E}_2\sin\phi_2 + ... + \mathcal{E}_J\sin\phi_J) \\
&+\sum_k^J\frac{\psi_k^P}{V_k^{\mathrm{tower}}}(\psi_k^E - \omega\psi_k^P) - \sum_k^J\frac{\psi_k^P}{V_k}\mathcal{E}_k + \frac{\omega_0 - \omega}{V_\omega} = 0
\end{aligned} \tag{D.18}$$

This gives

$$\kappa + \lambda_1\mathcal{E}_1 + ... + \lambda_m\mathcal{E}_m - \mu\omega = 0 \tag{D.19}$$

where

$$\begin{aligned}
\kappa &= \frac{V_{22}P_x^0 - V_{12}P_y^0}{d}E_x^0 + \frac{V_{11}P_y^0 - V_{12}P_x^0}{d}E_y^0 + \sum_k^J \frac{\psi_k^P}{V_k^{\text{tower}}}\psi_k^E + \frac{\omega_0}{V_\omega} \\
\lambda_k &= \frac{V_{22}P_x^0 - V_{12}P_y^0}{d}\cos\phi_k + \frac{V_{11}P_y^0 - V_{12}P_x^0}{d}\sin\phi_k - \frac{\psi_k^P}{V_k} \\
\mu &= \frac{V_{22}P_x^0 - V_{12}P_y^0}{d}P_x^0 + \frac{V_{11}P_y^0 - V_{12}P_x^0}{d}P_y^0 + \sum_k^J \frac{(\psi_k^P)^2}{V_k^{\text{tower}}} + \frac{1}{V_\omega}
\end{aligned} \tag{D.20}$$

Combining all equations

$$\begin{pmatrix}
\beta_{11} + \frac{1}{V_{\text{tower}}} & \beta_{12} & \dots & \beta_{1m} & -\gamma_1 \\
\beta_{21} & \beta_{22} + \frac{1}{V_{\text{tower}}} & \dots & \beta_{2m} & -\gamma_2 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
\beta_{m1} & \beta_{m2} & \dots & \beta_{mm} + \frac{1}{V_{\text{tower}}} & -\gamma_m \\
\lambda_1 & \lambda_2 & \dots & \lambda_m & -\mu
\end{pmatrix}
\begin{pmatrix}
\mathcal{E}_1 \\ \mathcal{E}_2 \\ \vdots \\ \mathcal{E}_m \\ \omega
\end{pmatrix}
= -
\begin{pmatrix}
\alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \\ \kappa
\end{pmatrix} \tag{D.21}$$

This can be solved for $\mathcal{E}_1$, ..., $\mathcal{E}_m$ and $\omega$ by inverting the $(m+1)$ by $(m+1)$ matrix. Using this analytical equation, computationally expensive minimization is avoidable.

# Appendix E

# Calorimeter geometry and segmentation

| | Barrel | | End-cap | |
|---|---|---|---|---|
| **EM calorimeter** | | | | |
| Number of layers and $\|\eta\|$ coverage | | | | |
| Presampler | 1 | $\|\eta\| < 1.52$ | 1 | $1.5 < \|\eta\| < 1.8$ |
| Calorimeter | 3 | $\|\eta\| < 1.35$ | 2 | $1.375 < \|\eta\| < 1.5$ |
| | 2 | $1.35 < \|\eta\| < 1.475$ | 3 | $1.5 < \|\eta\| < 2.5$ |
| | | | 2 | $2.5 < \|\eta\| < 3.2$ |
| Granularity $\Delta\eta \times \Delta\phi$ versus $\|\eta\|$ | | | | |
| Presampler | $0.025 \times 0.1$ | $\|\eta\| < 1.52$ | $0.025 \times 0.1$ | $1.5 < \|\eta\| < 1.8$ |
| Calorimeter 1st layer | $0.025/8 \times 0.1$ | $\|\eta\| < 1.40$ | $0.050 \times 0.1$ | $1.375 < \|\eta\| < 1.425$ |
| | $0.025 \times 0.025$ | $1.40 < \|\eta\| < 1.475$ | $0.025 \times 0.1$ | $1.425 < \|\eta\| < 1.5$ |
| | | | $0.025/8 \times 0.1$ | $1.5 < \|\eta\| < 1.8$ |
| | | | $0.025/6 \times 0.1$ | $1.8 < \|\eta\| < 2.0$ |
| | | | $0.025/4 \times 0.1$ | $2.0 < \|\eta\| < 2.4$ |
| | | | $0.025 \times 0.1$ | $2.4 < \|\eta\| < 2.5$ |
| | | | $0.1 \times 0.1$ | $2.5 < \|\eta\| < 3.2$ |
| Calorimeter 2nd layer | $0.025 \times 0.025$ | $\|\eta\| < 1.40$ | $0.050 \times 0.025$ | $1.375 < \|\eta\| < 1.425$ |
| | $0.075 \times 0.025$ | $1.40 < \|\eta\| < 1.475$ | $0.025 \times 0.025$ | $1.425 < \|\eta\| < 2.5$ |
| | | | $0.1 \times 0.1$ | $2.5 < \|\eta\| < 3.2$ |
| Calorimeter 3rd layer | $0.050 \times 0.025$ | $\|\eta\| < 1.35$ | $0.050 \times 0.025$ | $1.5 < \|\eta\| < 2.5$ |
| Number of readout channels | | | | |
| Presampler | 7808 | | 1536 (both sides) | |
| Calorimeter | 101760 | | 62208 (both sides) | |
| **LAr hadronic end-cap** | | | | |
| $\|\eta\|$ coverage | | | $1.5 < \|\eta\| < 3.2$ | |
| Number of layers | | | 4 | |
| Granularity $\Delta\eta \times \Delta\phi$ | | | $0.1 \times 0.1$ | $1.5 < \|\eta\| < 2.5$ |
| | | | $0.2 \times 0.2$ | $2.5 < \|\eta\| < 3.2$ |
| Readout channels | | | 5632 (both sides) | |
| **LAr forward calorimeter** | | | | |
| $\|\eta\|$ coverage | | | $3.1 < \|\eta\| < 4.9$ | |
| Number of layers | | | 3 | |
| Granularity $\Delta x \times \Delta y$ (cm) | | | FCal1: $3.0 \times 2.6$ | $3.15 < \|\eta\| < 4.30$ |
| | | | FCal1: $\sim$ four times finer | $3.10 < \|\eta\| < 3.15$, |
| | | | | $4.30 < \|\eta\| < 4.83$ |
| | | | FCal2: $3.3 \times 4.2$ | $3.24 < \|\eta\| < 4.50$ |
| | | | FCal2: $\sim$ four times finer | $3.20 < \|\eta\| < 3.24$, |
| | | | | $4.50 < \|\eta\| < 4.81$ |
| | | | FCal3: $5.4 \times 4.7$ | $3.32 < \|\eta\| < 4.60$ |
| | | | FCal3: $\sim$ four times finer | $3.29 < \|\eta\| < 3.32$, |
| | | | | $4.60 < \|\eta\| < 4.75$ |
| Readout channels | | | 3524 (both sides) | |
| **Scintillator tile calorimeter** | | | | |
| | Barrel | | Extended barrel | |
| $\|\eta\|$ coverage | $\|\eta\| < 1.0$ | | $0.8 < \|\eta\| < 1.7$ | |
| Number of layers | 3 | | 3 | |
| Granularity $\Delta\eta \times \Delta\phi$ | $0.1 \times 0.1$ | | $0.1 \times 0.1$ | |
| Last layer | $0.2 \times 0.1$ | | $0.2 \times 0.1$ | |
| Readout channels | 5760 | | 4092 (both sides) | |

Figure E.1: Main parameters of the calorimeter system. [16].