Chapter 7

Reconstruction and Identification

The signals collected from the different parts of the detector are due to the particles produced by the \( pp \) collisions. The direct information from the detector is in the raw form and is stored on magnetic tapes in the form of analog and digital signals from the various sub-systems. This raw information cannot describe the kinematics of these particles. Hence, one needs to process this information into a form suitable for physics analyses by converting the raw signal into information about the various final state objects in the event. This process of turning the raw data into the kinematic parameters of the particles is known as reconstruction. In this chapter, we explain the reconstruction process in DØ which is carried out by a software package called DØRECO.

7.1 DØRECO

The DØRECO package is a set of software algorithms which are used to perform particle and jet reconstruction. It performs three major tasks. The very first step is hit finding, which involves the unpacking of the raw data and the conversion of counts into hits of definite energy and spatial location. The signals from each sense wire of the tracking chambers are converted into the spatial location of hits and the signals from each cell in the calorimeter are converted into energy deposits. The second step involves tracking and clustering of these hits. The hits which are close spatially are grouped together to form clusters in the calorimeter and the tracks in the tracking chambers. The final step, particle identification, is the step during which the tracking and calorimetric information is combined to reconstruct jets and identify electron/photon and muon candidates. The criteria applied by DØRECO in selecting these candidates are quite loose, and substantial rejection of spurious electrons and muons is gained by further offline processing. For more details about the working of DØRECO package refer to [28].
7.2 Event Vertex

The interaction point plays a major role in getting the kinematics correct. After interaction, the transverse energy and momentum of a particle is calculated as, $E_T = E \sin \theta$ and $p_T = p \sin \theta$, where $\theta$ is its polar angle in the lab frame. So, we must know $\theta$ for an emerging particle. The only information which is available for a particle being detected in the calorimeter is the location where it deposits energy. In order to get an angle from this direction, another point is needed along its trajectory. This point is fixed by finding the location of the hard $p\bar{p}$ collision from which this particle has emerged and is called the event vertex or interaction vertex. As the calorimeter determines the energy and the position of the particles where they hit, the position of the interaction vertex from which the particles originate is needed to find the direction of the particle. The vertex position is reconstructed with the following procedure:

- The drift chamber hits are fitted to reconstruct a track in the $(r - \phi)$ plane.
- A reconstructed $(r - z)$ track is associated with the reconstructed $(r - \phi)$ track.
- The intersection of these reconstructed space tracks with the $z$ axis form a distribution in $z$. The estimated $z$ position of the vertex is the mean of a gaussian fitted to this $z$ intercept distribution. In the case of a $z$ distribution with more than one peak and therefore multiple vertices, the vertex with the maximum number of tracks is considered as the primary vertex. The resolution of the vertex’s $z$ component is about 1-2 cm.

7.3 Electrons

Electrons are identified as localized deposits of energy in the electromagnetic calorimeter with a track in the central detector pointing back to the interaction vertex.

7.3.1 Reconstruction

To identify electron, the reconstruction involves following steps:

- A cluster is constructed by starting with the highest $E_T$ electromagnetic tower and adding nearby towers to the cluster with $E_T$ above a certain threshold using a ‘nearest neighbor’ algorithm [44]. The process repeats until no towers neighboring the cluster satisfy the energy requirement. A new cluster is then begun from the highest $E_T$ tower not previously assigned to cluster.

- A cluster is required to have at least 90% of the energy in the electromagnetic calorimeter and 40% of its energy must be contained in a single tower.
• The centroid of the cluster is calculated using the cells in the third electromagnetic layer.
• If there is a track in the central detector within a solid angle of $\Delta \eta = \pm 0.1$ and $\Delta \phi = \pm 0.1$, pointing from the interaction vertex to the cluster, the cluster will be identified as an electron otherwise it is considered as a photon.

The energy resolution of electrons can be expressed by the relation:

$$\frac{(\sigma/E)^2}{E} = C^2 + \frac{S^2}{E} + \frac{N^2}{E^2}$$

(7.1)

where $E$ is the mean energy of the incident electron, $C$ is a constant term due to calibration errors, $S$ is the error due to statistical fluctuations, $N$ corresponds to a noise term due to the contribution of electronic and radioactivity of the absorber (uranium) in the calorimeter. The energy scale of the electromagnetic calorimeter has been calibrated with reference to the mass of the $Z$ in $Z \rightarrow e^+ e^-$ events, which has been measured very accurately by the LEP experiments [45]. The measured electron energies are scaled up so that the mass peak in $Z \rightarrow e^+ e^-$ matches the LEP measurement [46]. This correction is about 5% in the CC and $(1 - 2)$% in EC.

### 7.3.2 Identification

The procedure explained above favors efficiency rather than rejection. To discriminate electrons from other objects, we need to apply some selection cuts. One can further apply cuts according to specific analysis being carried out. The variables used are:

1. **Electromagnetic Energy Fraction**
   The electromagnetic energy fraction of a cluster is the fraction of its energy which is contained in electromagnetic calorimetry. For electrons, the electromagnetic calorimeter contains almost all of the energy, while charged hadrons will deposit only small fraction of their energy. Thus the electromagnetic energy fraction of a cluster serves as a powerful discriminant against charged hadrons.

2. **Isolation Fraction** ($f_{\text{isol}}$)
   The electron coming from $W$ should not be too close in space to other objects. Therefore the fraction of the energy outside a certain well defined cone should be small. DØ defines an isolation factor as:

   $$f_{\text{isol}} = \frac{E_{\text{Total}}(0.4) - E_{EM}(0.2)}{E_{EM}(0.2)}$$

   (7.2)

   where $E_{\text{Total}}(0.4)$ is the total energy deposited in all calorimeter cells within a cone of radius $R < 0.4$ ($R = \sqrt{\Delta \eta^2 + \Delta \phi^2}$) around the electron direction.
$E_{EM}(0.2)$ is the energy deposited in the EM calorimeter within a cone of radius $R < 0.2$. The present analysis requires $f_{	ext{had}} < 0.1$ which enhances electron identification.

3. Shower Shape
Shower shape is one of the primary and powerful tools used for selecting the electrons and rejecting the background. Electromagnetic showers can also be characterized by the fraction of the clusters energy deposited in each layer of the calorimeter. These are correlated and are also dependent on the incident electron energy. If an electron deposits large amount of energy in the first layer then it will deposit relatively small fraction of energy in the subsequent layers and vice-versa. To obtain the discrimination against hadrons and particles jets, we can exploit these correlations. Based on test beam studies and MC simulations of electrons with energies between 10 and 150 GeV, a 41 variable covariance matrix ($H$-matrix $\chi^2$) has been constructed given as:

$$\chi^2 = \sum_{i,j=1}^{41} (x_i - \bar{x}_i)H_{ij}(x_j - \bar{x}_j)$$

(7.3)

where $x_i$ are the variables which define the shape. A total of 41 variables are used. These are

- The fraction of the total energy contained in the first, second and fourth layers of the electromagnetic calorimeter.
- The fraction of the total energy contained in each cell of a $6 \times 6$ array in the third electromagnetic layer.
- The logarithm of the total energy of the cluster.
- The $z$ component of the primary vertex

A cut of $\chi^2 < 100$ is usually applied for better background rejection.

4. Track Match Significance ($S$)
A significant source to background for electrons is photons, either produced directly by the decay of $\pi^0$ or $\eta$ mesons. The photons do not leave tracks in the central detector, but a track might appear if a charged particle is nearby. One can enhance the electron identification by reducing the chance of reconstructing a track for photons due to nearby charged particles. In order to quantify the accuracy with which a track points to a calorimeter cluster, $S$ is defined as [47]:

$$S = \sqrt{(\frac{\Delta \phi}{\sigma_{\Delta \phi}})^2 + (\frac{\Delta z}{\sigma_{\Delta z}})^2}$$

(7.4)

where $\Delta \phi$ and $\Delta z$ are the azimuthal and the $z$ axis mismatch, respectively and $\sigma_{\Delta \phi}$ and $\sigma_{\Delta z}$ are the corresponding measurement resolutions. In the case
of EC track matching, the $z$ is replaced by the radial distance $r$.

5. TRD Efficiency ($\epsilon_t$)

The TRD response $\epsilon_t$ is defined as:

$$\epsilon_t(E) = \frac{\int_{E}^{\infty} \frac{dN}{dE'}(E')dE'}{\int_{0}^{\infty} \frac{dN}{dE'}(E')dE'}$$

where $E$ is the total energy recorded in the TRD minus that recorded in the layer with the largest signal and $\frac{dN}{dE'}$ is the energy spectrum from a sample of $W \rightarrow ev$ events. Since $\epsilon$ decreases as $E$ increases, hadrons will tend to have values near unity while the distribution from electrons is roughly uniform over the allowed range of zero to one.

6. Likelihood Ratio ($L_4, L_5$)

In order to obtain a better background rejection the DØ electron identification algorithm uses a 4 or 5 variable likelihood function. The 4 variable likelihood, $L_4$, uses the combinations of $f_{EM}$, $\chi^2$, $\sigma_{trk}$ and $dE/dx$ and the 5 variable likelihood, $L_5$, uses the 4 variable from $L_4$ and $\epsilon_t$.

7.4 Muons

Muons interact weakly with the matter and have long enough lifetime to pass through all of the detector material without decaying. These are identified as tracks in the muon drift chambers which point back at the interaction vertex and the reconstruction of muon tracks is similar to the reconstruction of tracks in the central detector. Similar to the central detector reconstruction, muon reconstruction involves three steps: hit sorting, track finding and global fitting. The first two steps make use of the information from the muon system only whereas the last step uses the information from the full DØ detector. They are reconstructed and identified using the hits and timing information from the muon spectrometers. Since a muon deposits a little of its energy in the calorimeter as it passes through, the resulting minimum ionizing trace can also be used for muon identification. The momentum of the muon is measured from the bend in the track produced by the magnetic field of the muon spectrometer.

7.4.1 Reconstruction

The muon reconstruction is similar to that of CDC track reconstruction except the differences in algorithm which take account of the geometry of the muon system. DØRECO uses timing information to determine the position of the hits in all the planes (A, B and C) of the muon system. For the B and C chambers, hits in four out of six possible planes are required and for the chamber A, this requirement is two out of four. The hits from the planes A, B and C chambers are then used to
form tracks. Because the B and C chambers are outside the magnetic toroid and the A chamber is inside, the tracking is done separately before and after the magnet. The segments are then matched, and a measurement of the momentum is made by measuring how much the track bends. Lastly, a global fit is performed using the tracks in the muon chamber, the interaction vertex, the energy profile in the calorimeter, and the track from the CDC/FDC. Additional corrections are made for the effects of multiple scattering in the calorimeter and the iron toroid and for the expected energy loss in the calorimeter.

7.4.2 Identification

The two major backgrounds to muon are from cosmic rays and leakage out of the backs of hadronic showers. To reduce these backgrounds, several variables are used to identify good muons which are as:

1. **Muon Track Quality (IFW4)**
   IFW4 represents the quality of a track fit. Track with perfect fits have IFW4 of 0 and those with one failure have IFW4 of 1. The IFW4 cut is a very powerful tool to reduce the cosmic rays and the fake track backgrounds constructed from random hits.

2. **Isolation (ΔR cut)**
   The analysis requires that the distance in R between the muon and the nearest jet be $\Delta R(\mu, jet) > 0.5$ and for muon and nearest electron/photon to be $\Delta R(\mu, e/\gamma) > 0.25$.

3. **Muon Track in the Calorimeter (MTC)**
   As muon passes through the calorimeter, it deposits energy through ionization, and these energy traces are used in the track fit. The fraction of all possible hadronic calorimeter layers which had energy deposits large enough to be included in the fit is recorded (MTC), along with the fraction of energy deposited in the outermost possible layer. Both of these quantities are useful in rejecting muon tracks formed from random noise in the muon system.

4. **Impact Parameters (IP)**
   Two impact parameter cuts are used to require that the muon tracks point towards the interaction vertex, and thus reject cosmic ray backgrounds.
   The non-bend impact parameter is defined by projecting the muon track into the $xy$ plane (it does not get bent in this plane), extrapolating the track formed by the B and C layers towards the center of the detector, and calculating the impact parameter between this extrapolated track and the interaction vertex. The impact parameter is required to be less than 40 cm.
   The bend-view impact parameter is calculated by projecting the track into the plane in which the muon bends and calculating the impact parameter of this projection. This parameter is required to be less than 25 cm.
7.5 Jets

Quarks other than top and gluons which are produced in the \( p\bar{p} \) collisions hadronize into colorless particles. When a quark or gluon leaves the site of a hard scattering, it can not remain free but instead hadronizes or fragments into a collection of colorless hadronic particles. This collection will typically lie in a cone around the direction of motion of the original parton and will show up in a calorimeter as a cluster of energy. This is called a jet.

7.5.1 Reconstruction

There are several algorithms which could be used for jet reconstruction. The most common algorithm used in the \( p\bar{p} \) environment is the “cone algorithm” in which a jet is considered as the energy inside of a cone with a fixed size in \((\eta, \phi)\) space. This definition was used by UA1 [48, 49], and CDF and is also used by most DØ analyses. A jet is required to satisfy some minimum \( E_T \) threshold before it could be considered in analyses so as one of the event’s objects. This threshold is 8 GeV and is needed in order to suppress random noise fluctuation which can produce small energy clusters. The main steps in the reconstruction process are:

**Pre Clustering:** The transverse energy is calculated for all of the calorimeter towers which are then sorted in order of decreasing \( E_T \) to form a “seed” cluster. Beginning with the highest \( E_T \) tower, clusters are formed by adding the towers within a radius \( R \) of the highest energy tower. The process is repeated for the remaining calorimeter towers.

**Cone Clustering:** The centroid of each cluster is calculated by performing an \( E_T \) weighted sum of the tower \((\eta, \phi)\) positions. Then the whole process is iterated using the jet centers as cluster seeds until the position of the cluster converges.

**Merging and Splitting:** Once the cone clustering is completed, some cells may have been assigned to more than one jet. If two jets share cells, the fraction of the total energy which is shared between them is examined. If the fraction is greater then 50%, the two jets are merged together and the jet axis is recalculated from the centroid of the cells in the merged jet. Otherwise the jets are split and the shared cell is assigned to the closest jet. At this stage all jets with transverse energy \( E_T > 8 \) GeV are retained for further analysis.

7.5.2 Identification

In order to remove any fake jets produced by calorimeter or MR noise, DØ has developed a set of quality cuts based on the jet characteristics. These are cuts on the jet ElectroMagnetic Fraction (EMF) which is used to distinguish between electrons/photons and jets, the Hot Cell Energy Fraction (HCF) which helps reduce calorimeter noise and Coarse Hadronic Energy Fraction (CHF) which helps to remove activity caused by the MR.
The various kinematic quantities defining a jet are:

\[ E_T = \sqrt{E_x^2 + E_y^2} \]  

where \( E_x, E_y \) are the sums of the components of the individual cell energies:

\[ E_x = \sum_i E_{xi} \]  
\[ E_y = \sum_i E_{yi} \]

(7.6)

and

\[ \phi = \tan^{-1} \left( \frac{E_y}{E_x} \right) \]  
\[ \theta = \cos^{-1} \left( \frac{E_z}{\sqrt{E_x^2 + E_y^2 + E_z^2}} \right) \]

(7.9)

(7.10)

\[ \phi = \tan^{-1} \left( \frac{E_y}{E_x} \right) \]

7.6 Missing Transverse Energy

According to conservation of transverse momenta and the fact that the colliding proton and antiproton have nearly opposite momenta, it follows that the sum of the transverse momenta of the particles produced by \( pp \) collisions should be zero. Neutrinos do not interact in the detector and if the total transverse momentum is significantly different from zero, the difference is attributed to neutrinos. And this imbalance in the transverse energy is known as “missing \( E_T \)” and is denoted as \( \not{E}_T \).

In order to calculate the transverse energy of the neutrinos, a vector \( \not{E}_T \) is assigned to each calorimeter cell, including the Intercryostat Detector (ICD), whose magnitude is the measured energy in the cell and it points from the interaction vertex to the center of the cell. The calorimeter missing energy is defined as:

\[ \not{E}_T^{\text{cal}} = - \sum_i \not{E}_i \]  

(7.11)

and its magnitude is given as:

\[ \not{E}_T^{\text{cal}} = \sqrt{(\not{E}_x^{\text{cal}})^2 + (\not{E}_y^{\text{cal}})^2} \]

(7.12)

The muons deposit small portion of their energy in the calorimeter, so the transverse momenta of all good muon tracks should be subtracted from \( \not{E}_T^{\text{cal}} \) to get the total transverse missing energy. Therefore,

\[ (\not{E}_T)_x = (\not{E}_T^{\text{cal}})_x - \sum_i \mu_i \]

(7.13)
\[(\mathbf{p}_T)_y = (\mathbf{p}_{T,\text{cal}})_y - \sum_i p_{i_y}^e\]  

\[\mathbf{p}_T = \sqrt{(\mathbf{p}_{T,\text{x}})^2 + (\mathbf{p}_{T,\text{y}})^2}\]  

(7.14)  

(7.15)

Since all the objects in the calorimeter contribute to \(\mathbf{p}_{T,\text{cal}}\), any mismeasurement in the energy of these objects would cause a mismeasurement in \(\mathbf{p}_{T,\text{cal}}\). Therefore whenever corrections are applied to the calorimeter objects like electron and jets, the corresponding correction must be applied to \(\mathbf{p}_{T,\text{cal}}\) [50].