Particle Identification tools and performance in the SHiP Experiment

Behzad Hosseini\textsuperscript{1} and Walter Bonivento\textsuperscript{1}

\textsuperscript{1}\textit{INFN Cagliari}

Abstract

This note describes in detail the implementation and performance of the PID algorithm in the FairSHiP simulation.
1 Introduction

Particle IDentification (PID) is a crucial issue to most particle physics experiments. This note reviews PID techniques and strategies used by the SHiP experiment. PID is provided by ECAL, HCAL and Muon sub-detectors.

This note describes in detail the procedure used to provide the PID for this experiment. General design of PID sub-detectors will be briefly explained in section 2. In section 3 the PID methods and strategies will be discussed. The results will be summarised in section 4.

2 Overview of SHiP's PID Sub-detectors

2.1 ECAL

The electromagnetic calorimeter is one of the PID sub-detectors to provide electron, photon and pion identification at the offline level. It is using the advantage of the shashlik technique, in which it contains a sample of scintillator-lead structure. The main goals to use this technique in the SHiP electromagnetic calorimeter are the fast response time, good energy resolution and relatively low cost-to-performance ratio. Recent developments in shashlik technology have shown that electromagnetic shower energies can be measured with a resolution of \( \sigma_E/E > 6\% \sqrt{E} \) (which E in GeV).

The calorimeter has outward dimensions of around \( 5 \times 10^2 \text{ m}^2 \) and contains 2876 modules arranged in 84 rows and 42 columns, for a total of 11504 readout channels.

2.2 HCAL

Same as the ECAL, the shashlik technology is the technology that used for hadronic calorimeter due to its similarity to the ECAL technology and ease of the MC implementation. The hadronic calorimeter is a system consisting of 2 read-out stations and is composed of several sampling modules. The HCAL detector has been placed behind ECAL with the same acceptance. Each module is composed by several layers of converting material (15 mm thickness of iron) alternated with active material (made of 5 mm of polystyrene-based scintillator tiles).

The main goals of the hadron calorimeter are to provide the pion identification and to provide pion/muon discrimination especially for low momentum particles \( (p < 5 \text{ GeV/c}) \). The longitudinal space taken by the whole calorimeter system will be about 3.3 m.

2.3 Muon Detector

Four stations of active layers interleaved by three muon filters construct the muon detector which is placed downstream of the calorimeter systems. The muon detector, in conjunction with the electromagnetic and hadronic calorimeters, is designed primarily to identify muons and it is 6 m wide and 12 m high. The multiple scattering of muons in the muon filters and the material of the calorimeter system drives the granularity of the system. Preliminary
simulation studies show that a readout granularity of 5-10 cm in the transverse directions is adequate for the interesting momentum range. The whole PID sub-detector layout is shown in Figure 1.

3 Particle Identification Criteria and Selections

In each experiment, the various elementary particles give different characteristic signatures in the specific and separate detectors. Charged particles leave particle hits due to ionization in their related detectors. Electrons produce Bremsstrahlung radiation when passing through matter and they are stable particles and have low mass ($m_e = 0.51$ MeV). Muons act like heavier versions of the electron, with mass 105.7 MeV. Since the muon mass is large, its Bremsstrahlung radiation is small, and as a lepton it does not feel the strong interaction.

In order to determine the particle identification in SHiP experiment, several especial selections based on the individual signals of each PID sub-detector are needed. After performing the related tracks in each event, each track will be extrapolated to all PID sub-detectors to be analysed by the information giving by all sub-detectors. The response of each sub-detector can be expressed in terms of its raw signal. The PID efficiency is been defined as the proportion of particles of a given species that are identified correctly by the PID selections. In the following the selection of various variables for electron/hadron and hadron/muon separation will be briefly discussed.
3.1 Electron-Hadron Identification

The important variables to distinguish between electrons and hadrons are given by electromagnetic calorimeter and they are as following:

- \( \Delta x \) and \( \Delta y \); which they are the difference in positions for both \( x \) and \( y \) projections of the extrapolated tracks and the reconstructed clusters in the ECAL detector (see Figure 2).

- \( E/P \); which is the Energy of clusters that are reconstructed by ECAL detector over the momentum of each track.

The thresholds of these variables depend on the momentum of particles. A dip study for different range of energies has been studied in this work. To calculate the right \( \Delta x \) and \( \Delta y \) in ECAL, each particle track is extrapolated to the \( Z \) position of the shower maximum.

3.2 Hadron-Muon Identification

Hadrons mostly stop in hadron calorimeter, however muons passing through the HCAL and reach muon detector. The interested variables to distinguish between hadrons and muons within the muon detector are:

- \( \Delta x \) and \( \Delta y \); same as the ECAL detector, \( \Delta x \) and \( \Delta y \) are the difference in positions of the extrapolated tracks and the muon hits within the muon detector (see Figure 3).
• **Depth**: as muon detector contains four active layers, this gives the possibility of calculating the depth of particle inside the muon detector.

• **Number of hits around the extrapolated track**: which is the number of muon hits related to muon particles.

Low energy muons (\( P < 5 \text{ GeV} \)) might not reach the muon detector. Therefore, for those particles we also check the information given by HCAL detector in order to see if we have a mip for the muon or not.

### 4 The Structure of PID Code

Taking into account all the information giving by PID sub-detectors, there is the opportunity to identify different type of particles in SHiP experiment. A step by step code - PID code - has been written to distinguish electron, hadron and muon particles. The structure of the code is designed in a way that for each event first the program access to the track container and checks all the tracks of each event to identify them.

First, each track inside the track container will be extrapolated to the ECAL detector and later by distracting the information given by ECAL detector the code challenging to check if the particle is an electron or not.

Depending on the respond of the code, different labelling has been introduced for each particle. These labels in the case of electrons is as following:

• **ElectronID = 1**: The particle is an electron.
- **ElectronID = 0;** It is not an electron.
- **ElectronID = -1;** The information given by the ECAL detector is not enough to discuss if the particle is an electron.
- **ElectronID = -2;** The particle is outside of the ECAL detector.
- **ElectronID = -3;** The track does not satisfied ”FitConverged” or ”ndf” cuts.

If the particle is not an electron, the next step will be to check if the particle is a muon. For that, same as electron, the first step is to extrapolate the track to Muon detector. Later, by using the information giving by Muon detector, the PID code decide if the particle is a muon or not. As mentioned in Section 3.2, low momentum particles (P < 5 GeV) are double checked by checking their HCAL response.

Labels for muons and hadrons are same as electrons and are as following:
- **HadronID (MuonID) = 1;** The particle is a hadron (muon).
- **HadronID (MuonID) = 0;** It is not a hadron (muon).
- **HadronID (MuonID) = -1;** The information given by the ECAL (muon) detector is not enough to discuss if the particle is a hadron (muon).
- **HadronID (MuonID) = -2;** The particle is outside of the ECAL (muon) detector.
- **HadronID (MuonID) = -3;** The track does not satisfied ”FitConverged” or ”ndf” cuts.

The summary of the structure of PID code has been illustrated in the diagram shown in Figure 4.

5 Implementing PID in FairShip

The PID has been implemented to the FairShip software in away that after running the ShipReco, PID container will be created automatically. This container stores the information of
- **TrackID;** which is correspond to the number of tracks within each event,
- **ParticleID;** that gives the information related to the identification of particles of each event.

A view of PID container is represented in Figure 5.

In order to do any physics analysis regarding the HNL events, the data analysis experts can easily extract the PID information in the ShipAna software. This overcomes by looping on the PID container and connecting to each track by its trackID. After selecting the interesting particles for specific analysis (for example muons), one can connect the FitTracks container to extract all the information related to that track, see Figure 6.
6 Results

To calculate the efficiency of the PID code, 30000 events are generated with Pythia8 generator in the FairShip software each of them having a mass of 1 GeV. These events include the following channels: \(\mu\mu\), \(\mu\mu\nu\), \(e\nu\), \(e^+\nu\), \(\pi\pi\), \(\mu\pi\), \(\pi e\).

A basic set of kinematic and topological criteria are applied to select the right events. The cuts are as following:

- events with only one HNL candidate.
- all vertices and tracks be within the fiducial volume.
- track n.d.f > 25
- \(\chi^2/n.d.f < 5\)
- DOCA < 1 cm
- \(P_{\text{daughter}} > 1\) GeV
Cut on IP
no particle out of the acceptance of the PID detectors

Cut on the IP is different for 2 body and 3 body events. For 2 body events (such as $\mu\mu$, $ee$, ...) IP cut should be less then 10 while for 3 body events (like $\mu\mu\nu$, $ee\nu$, ...) this cut needs to be less then 250.

The result of this work is shown on Figures 7 and 8 for 2 and 3 body events respectively. The first row of the table shown in Figure 7 shows the reconstructed events in SHiP while the most left column of the table illustrates the Monte Carlo events generated by FairShip software. The blue columns represents the 2 body events and the red column is shown to identify the immigrated events to 3 body events. However, the two red columns on
Table shown in Figures 7 and 8 shows the immigrated events to 2 body events. The percentage of events on the diagonal part of the table shows the efficiency of PID software for each channel. The outliers are due to systemical or physical affects that will be briefly explained in the following.

**Systemical misidentification:** The misidentification of particles caused by systematic affects are sorted in 3 items:

- **Muon particles that identified as pions:** this happens when the energy of the muons are very low (around 1-2 GeV), therefore the information giving by muon
detector is not enough to classify them as muon particles. If the opening angle of these muons with the other particles in the same events are small, the chance of using the information of HCAL detector will fail too.

- **Electron particles that identified as pions:** Sometimes the energy reconstruction of the reconstructed cluster in ECAL detector fails for very high energy electrons. Since the information from the muon and ECAL detector of this electrons are missed, they recognised as pions.

- **Electron particles that identified as muons:** This misidentification is very similar to those electrons which are misidentified as pions. The difference here is when there is a muon in the event and the opening angle between the muon and electron is very small. In this case, after the extrapolation of the tracks to muon detector, electron uses the same muon hit in the muon detector. Since the reconstructed cluster energy in ECAL is also failed, the program mistakenly identify this electrons as muons. This kind of misidentification is under the control and in new upgraded versions of PID code will be solved.

**Physical misidentification:** Physical misidentifications are classified in 2 categories:

- **Pions decaying in muons:** charged pions can decay into a muon and a muon neutrino or antineutrino. When pions decay before the track reconstruction, there is no chance to recognise them as pions.

\[
\pi^+ \rightarrow \mu^+ + \nu_\mu \\
\pi^- \rightarrow \mu^- + \bar{\nu}_\mu
\]

- **Pions that do the charge exchange:** interactions in which a charged pion does not produce a \( \pi^\pm \) can be a pion charge exchange interactions:

\[
\pi^- + p \rightarrow \pi^0 + N \\
\pi^+ + \bar{p} \rightarrow \pi^0 + N
\]

this cause a misidentification for pion particles.

Efficiency of PID code is calculated for lower mass events as well. For that, 30000 events with a mass around 400 MeV are generated for this study. The result for 2 and 3 body events are reported in Figures 9 and 10.
Figure 9: PID efficiency for 2 body events having mass around 400 MeV. Most upper row represent the reconstructed events in SHiP experiment and the first column from left shows the Monte Carlo events. The red column stands for the immigration of 2body to 3body events. The upper number in each cell shows the number of reconstructed events over the Monte Carlo events for each channel; the lower number is the efficiency of particle identification for each channel.

<table>
<thead>
<tr>
<th>GEN</th>
<th>REC</th>
<th>µ-µ</th>
<th>e-e</th>
<th>π-π</th>
<th>µ-π</th>
<th>π-e</th>
<th>µ-e</th>
</tr>
</thead>
<tbody>
<tr>
<td>µ-µ</td>
<td>2 body</td>
<td>287/291</td>
<td>266/267</td>
<td>268/297</td>
<td>23/296</td>
<td>12/236</td>
<td><strong>98.63%</strong></td>
</tr>
<tr>
<td>e-e</td>
<td>2 body</td>
<td>4/291</td>
<td>1/267</td>
<td>5/297</td>
<td>20/297</td>
<td>221/236</td>
<td></td>
</tr>
<tr>
<td>π-π</td>
<td>2 body</td>
<td>268/297</td>
<td>266/267</td>
<td>259/296</td>
<td>23/296</td>
<td></td>
<td></td>
</tr>
<tr>
<td>µ-π</td>
<td>2 body</td>
<td>5/297</td>
<td>20/297</td>
<td>1/296</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>π-e</td>
<td>2 body</td>
<td>221/236</td>
<td>23/296</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 10: PID efficiency for 3 body events having mass around 400 MeV. Most upper row represent the reconstructed; and the first column from left shows the Monte Carlo events. The red column stands for the immigration of 3body to 2body events. The upper number in each cell shows the number of reconstructed events over the Monte Carlo events for each channel; the lower number is the efficiency of particle identification for each channel.

<table>
<thead>
<tr>
<th>GEN</th>
<th>REC</th>
<th>µ-µ</th>
<th>e-e</th>
<th>µ-µ</th>
<th>µ-π</th>
<th>π-e</th>
</tr>
</thead>
<tbody>
<tr>
<td>µ-µ</td>
<td>3 body</td>
<td>312/317</td>
<td>230/231</td>
<td>12/240</td>
<td>5%</td>
<td>223/240</td>
</tr>
<tr>
<td>e-e</td>
<td>3 body</td>
<td>5/317</td>
<td>1/231</td>
<td>223/240</td>
<td>5/240</td>
<td></td>
</tr>
</tbody>
</table>