



UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
DEPARTAMENTO DE INFORMÁTICA
VALPARAÍSO – CHILE



ALGORITMO DE RECONSTRUCCIÓN PARA LA IDENTIFICACIÓN Y SEGREGACIÓN DE PARTÍCULAS DE ALTAS ENERGÍAS EN UN DETECTOR PRESHOWER.

Tesis presentada como requerimiento parcial
para optar al grado académico de

MAGISTER EN CIENCIAS DE LA INGENIERÍA
INFORMÁTICA

y al título profesional de
INGENIERO CIVIL EN INFORMÁTICA

por

Juan Guillermo Pavez Sepúlveda

OCTUBRE 2017

UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA
DEPARTAMENTO DE INFORMÁTICA
VALPARAÍSO – CHILE

TITULO DE LA TESIS:

ALGORITMO DE RECONSTRUCCIÓN PARA LA IDENTIFICACIÓN Y SEGREGACIÓN DE PARTÍCULAS DE ALTAS ENERGÍAS EN UN DETECTOR PRESHOWER.

AUTOR:

JUAN GUILLERMO PAVEZ SEPÚLVEDA

Tesis presentada como requerimiento parcial para optar al grado académico de **Magister en Ciencias de la Ingeniería Informática** y al título profesional de **Ingeniero Civil en Informática** de la Universidad Técnica Federico Santa María.

Prof. Dr. Héctor Allende

Profesor Guía

Prof. Dr. Hayk Hakobyan

Profesor Cotutor

Prof. Dr. Carlos Valle

Profesor Correferente

Prof. Dr. Max Chacón

Profesor Externo

Prof. Dr. Marcelo Mendoza

Presidente Comisión

OCTUBRE 2017.

Valparaíso, Chile.

Agradecimientos

Agradezco a mi familia y a mi novia por su apoyo durante el desarrollo de mi carrera y constante preocupación por el avance de esta tesis. Además, me gustaría agradecer al Dr. Hayk Hakobyan por su apoyo en el desarrollo de este trabajo y su guía especialmente en los temas relacionados con la física de altas energías. También me gustaría agradecer a mi profesor guía Dr. Héctor Allende por su apoyo en el desarrollo de esta tesis y al Dr. Carlos Valle por sus valiosos comentarios y contribuciones al escrito de esta tesis. Agradezco también al Centro Científico y Tecnológico de Valparaíso (CCTVal) por el financiamiento parcial de este trabajo a través del proyecto basal FB0821.

Valparaso, Chile

Juan Guillermo Pavez Sepúlveda

Resumen

La detección de partículas es un trabajo muy importante en campos como la física de altas energías, la medicina y la astronomía. El proceso de detección consiste en identificar las partículas producidas en distintos fenómenos físicos, como por ejemplo, las partículas producidas en los grandes colisionadores. El proceso de detección, sin embargo, es complicado y depende de muchos factores. Un caso especialmente difícil es la detección de partículas muy cercanas provenientes, por ejemplo, del decaimiento de piones neutros. Para solucionar este problema el centro Científico y Tecnológico de Valparaíso ha propuesto un detector del tipo calorímetro preshower, diseñado especialmente para tener una alta resolución en la identificación de partículas cercanas. La solución de este problema tiene varias aplicaciones prácticas: En los colisionadores electron-ion puede ayudar al estudio de los gluones, el material que une toda la materia del universo; En la medicina nuclear puede ayudar a mejorar las imágenes producidas en las tomografías por emisión de positrones, vitales en el tratamiento del cancer.

Para procesar los datos producidos por el detector se utiliza un algoritmo de reconstrucción. Este algoritmo es el encargado de convertir los datos brutos en información útil de la partícula detectada, como su posición de incidencia, su energía o su ángulo.

La contribución principal de este trabajo es el diseño e implementación de un algoritmo de reconstrucción completo para el detector preshower. Para esto, las diferentes opciones de la literatura son estudiadas y nuevas soluciones son propuestas para el trabajo de reconstrucción. Además, se construye una simulación computacional del detector y se estudia, mediante simulaciones, el desempeño del algoritmo de reconstrucción. Los resultados obtenidos muestran la capacidad del detector preshower para identificar partículas cercanas, superando otras propuestas de la literatura. Además, basándose en este análisis, se proponen posibles mejoras al detector, las cuales pueden ser implementadas en próximas iteraciones del diseño del detector.

Palabras Claves: Algoritmo de reconstrucción, Física de altas energías.

Abstract

Particles detection is an important task in fields such as high energy physics, medicine and astronomy. The detection process consists on identifying the particles produced in different physical phenomena, for instance, the particles produced in big colliders. The detection process, however, is complex and depends on the detector physics and the particles type. Specially hard is the case in which two particles arrive very close to each other, for example, from a neutral pion decay. In order to solve this problem researchers from the Science and Technology Center of Valparaíso have proposed a preshower calorimeter detector. This detector is specially designed to have high resolution to identify particles close to each other. The solution to this problem has many practical applications: In electron-ion colliders it can help in the study of gluons, the material binding all matter; In nuclear medicine it can help to improve the images produced in positron emission tomography, very important in cancer treatment.

In order to process the data produced by the detector, a reconstruction algorithm is used. The reconstruction algorithm converts the raw data obtained by the detector in useful information about the particle, such as incident position, energy or angle.

The main contribution of this work is the design and implementation of a full reconstruction algorithm for the preshower detector. For this, different options in the literature are studied and new solutions are proposed for the reconstruction in the preshower. Moreover, we build a computer simulation of the detector and using these simulations we study the performance of the reconstruction algorithm. The obtained results show the ability of the preshower detector to identify close particles, outperforming the results from other proposals found in the literature. Also, based on this analysis we propose some improvements to the preshower design. Those improvements can be implemented in new iterations of the detector design.

Keywords: Reconstruction algorithm, High energy physics.

Indice de Contenidos

Agradecimientos	III
Resumen	IV
Abstract	V
Indice de Contenidos	VI
Lista de Abreviaciones	VII
1. Introducción	1
1.1. Motivación	1
1.1.1. Detector Calorímetro	3
1.2. Contexto	5
2. Detectores de partículas	8
2.1. Física de los detectores calorímetros	8
2.1.1. Lluvias electromagnéticas	9
2.1.2. Materiales de construcción	11
2.1.3. Cristales centelleadores	11
2.1.4. Fotomultiplicadores	12
3. Detector Preshower	14
3.1. Diseño	14
3.1.1. Matriz de cristales	15

3.1.2. Sistema de lectura	17
4. Estado del Arte	20
4.1. Clústering de partículas	21
4.1.1. Ventana deslizante	22
4.1.2. Clústering topológico	24
4.1.3. Clústering Fuzzy	25
4.2. Identificación de máximos	28
4.2.1. Algoritmo de búsqueda de peaks	29
4.3. Separación de lluvias solapadas	31
4.4. Rechazo de partículas solapadas	33
4.5. Reconstrucción de posición	35
5. Algoritmo de Reconstrucción para el Detector Preshower	37
5.1. Descripción General	37
5.2. Propuesta y experimentos	38
5.2.1. Simulaciones	38
5.2.2. Algoritmo de reconstrucción	42
6. Resultados	55
6.1. Análisis de resultados	55
6.1.1. Identificación de partículas cercanas	56
6.1.2. Reconstrucción de posición	56
6.1.3. Rechazo de partículas solapadas con un sólo máximo	59
6.1.4. Reconstrucción de piones neutros	62
7. Conclusiones y Trabajo Futuro	64
Apéndices	67
A. Conceptos Básicos	68
A.1. Paso de partículas por la materia	68
A.1.1. Sección eficaz y recorrido libre	68

A.1.2. Interacción de partículas con la materia	71
A.2. Tipos de detectores	75
B. Resultados adicionales	77
B.1. Reconstrucción de posición	77
B.2. Reconstrucción en el eje Y	81
B.3. Reconstrucción del conteo de fotoelectrones	81
Bibliografía	86

Lista de Abreviaciones

- ATLAS** Detector parte del LHC (A Toroidal LHC Apparatus). 22
- BDT** Ensamblado de Árboles de Decisión (Boosted Decision Trees). 59
- BGO** Material utilizado en la construcción de cristales (Bismuth Germanium Oxide). 4
- BNL** Laboratorio Nacional de Brookhaven (Brookhaven National Laboratory). 5
- CALICE** Calorímetro para el Colisionador Lineal Internacional (Calorimeter for Linear Collider Experiment). 27
- CCTVal** Centro Científico y Tecnológico de Valparaíso. 6
- CERN** Organización Europea para la Investigación Nuclear. 4
- CMS** Detector Solenoide Compacto de Muones (Compact Muon Solenoid). 4
- ECAL** Calorímetro electromagnético parte de CMS (Electromagnetic Calorimeter). 4
- EIC** Colisionador Electrón-Ion (Electron Ion Collider). 5
- GeV** Giga Electronvoltios. 31
- HCAL** Calorímetro hadrónico parte de CMS (Hadron Calorimeter). 4
- HEP** Física de Alta Energía (High Energy Physics). 4
- JLab** Laboratorio Jefferson (Jefferson Lab). 5

- LYSO** Material para la construcción de cristales (Lutetium Yttrium Orthosilicate). 4
- MeV** Mega Electronvoltios. 28
- MLP** Perceptrón Multi Capas (Multilayer Perceptron). 59
- MPPC** Dispositivo de conteo de fotones (Multi-Pixel Photo Counter). 17
- rms** Media Cuadrática (Root Mean Square). 58
- SVM** Máquinas de Soporte Vectorial (Support Vector Machines). 59
- TMVA** Framework para Análisis Multivariado (Toolkit for Multivariate Data Analysis). 59

Capítulo 1

Introducción

1.1. Motivación

Grandes avances han sido logrados durante los últimos años en el campo de la física de partículas. Gran parte de estos adelantos han sido permitidos gracias al desarrollo y perfeccionamiento de complicados colisionadores de partículas que permiten a los científicos detectar y analizar millones de datos provenientes de colisiones entre partículas. Los grandes colisionadores permiten observar el universo en su escala más microscópica, es decir, analizar las partículas que componen todo lo que nos rodea. En conjunto con los colisionadores, es necesario construir potentes detectores que permitan medir y analizar el resultado de las colisiones. Los detectores de partículas son responsables de proveer los datos en bruto para el posterior análisis que permitirá realizar conclusiones acerca de los eventos observados.

Aunque los grandes colisionadores han acelerado considerablemente el desarrollo de los detectores de partículas, los detectores no sólo tienen uso en el estudio de las partículas producidas en estos. Su uso es fundamental en la tarea de percepción remota. Esta tarea es común en varios campos, como la astronomía, donde se estudian las partículas producidas a millones de años luz, con la finalidad de entender los procesos que las produjeron; o en la medicina, donde se desea obtener imágenes de los tejidos y procesos internos del cuerpo mediante la detección de las partículas emitidas, por ejemplo, en la tomografía por emisión de positrones (PET).

Básicamente, dado un evento que produce partículas secundarias de diferentes masas en varios ángulos y con distintos momentos, como se muestra en la Figura 1.1. El rol del

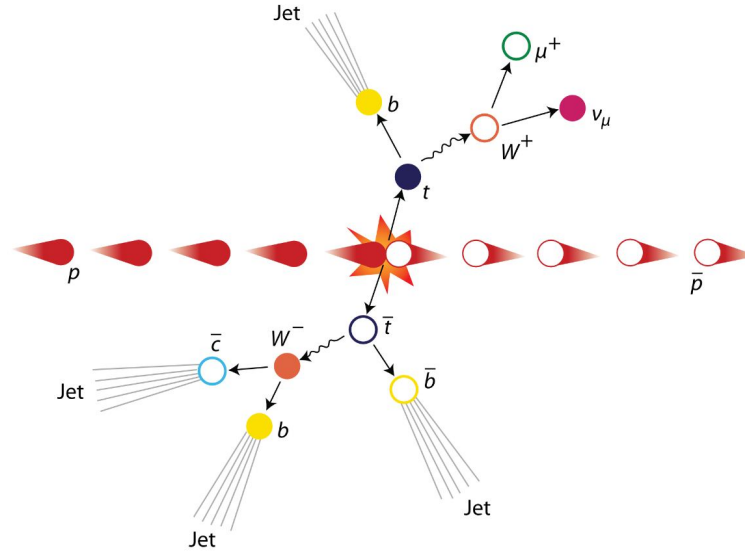


Figura 1.1: Colisión entre partículas y partículas secundarias producidas.

detector será medir variables como el tiempo de interacción t , el momento p y la masa M de estas partículas.

El proceso de detección no es un trabajo simple, primero se debe comprender que para cada tipo de partículas y para cada variable a analizar existen distintos tipos de detectores con diferentes propiedades de detección. Estos detectores a su vez pueden variar mucho entre sí, basándose en principios físicos muy distintos. Debido a que en los colisionadores las partículas producidas pueden ser de una gran cantidad de tipos, en diferentes rangos de energías, un detector en un acelerador puede estar compuesto por diversos sub-detectores que deben trabajar en conjunto para obtener una imagen completa del evento. Además de las complicaciones inherentes al proceso de detección, se debe considerar toda la electrónica asociada al trabajo de lectura de las señales. Esto involucra el tratamiento electrónico de las señales, la digitalización, el preprocesamiento en computadoras, entre otros.

Añadido a todo esto se debe observar que los colisionadores pueden producir una cantidad gigantesca de datos, colisionadores como el LHC ubicado en Suiza produce alrededor de 600 millones de colisiones por segundo (en datos se producirán 25 GB/s en la actual corrida del detector). Esto agrega la necesidad de que todo el proceso de detección debe ser muy rápido y eficiente, o en otras palabras, la detección debe ser muy escalable.

Si bien el proceso de detección entrega datos en bruto de las partículas detectadas, es necesario procesar estos datos computacionalmente con la finalidad de obtener información

útil para ser utilizada en el análisis. De esto último se encarga el procedimiento conocido como reconstrucción. Durante este paso los datos en bruto entregados por el detector son procesados utilizando métodos estadísticos y modelos teóricos en física con la finalidad de inferir el proceso mediante el cual los datos finales fueron producidos. Es decir, identificar las partículas que dieron origen a los datos observados y sus propiedades. Al igual que en los detectores, existe una gran cantidad de opciones para el proceso de reconstrucción, dependiendo del detector del que provienen los datos, el tipo de partícula a ser identificada, el rango de energía en que se opera, etcétera.

La contribución principal de este trabajo es la propuesta e implementación de un algoritmo de reconstrucción completo para el detector preshower, que nace de la propuesta de un grupo de investigadores del Centro Científico y Tecnológico de Valparaíso (CCTVAL). Debido al diseño innovador del detector, muchos de los componentes del algoritmo de reconstrucción son una contribución nueva, mientras que otros componentes deben ser adaptados de la literatura para ser utilizados en conjunto con la singular composición del detector. El detector preshower es un tipo de detector calorímetro encargado de localizar partículas muy cercanas entre sí. Como ya se mencionó, existe una gran gama de detectores, los cuales deben trabajar en conjunto para obtener una imagen completa del evento físico. A continuación se procederá a entregar una muy breve introducción a los detectores del tipo calorímetro. En el Apéndice A.2 se describen otros tipos de detectores a modo de información complementaria.

1.1.1. Detector Calorímetro

Los detectores del tipo calorímetros son muy útiles para medir la energía total de una o más partículas (tanto neutras como cargadas). Pueden ser de dos tipos: calorímetros electromagnéticos o calorímetros hadrónicos. Los calorímetros electromagnéticos son capaces de medir la energía de las partículas incidentes mediante la producción de lluvias electromagnéticas, compuestas de electrones, positrones y fotones y producidas mediante procesos de creación de pares y radiación de frenado (a altas energías) [1]. Los calorímetros hadrónicos, en cambio, miden la energía de la partícula usando la generación de lluvias hadrónicas producidas mediante interacciones del tipo fuerte. Ambos pueden ser homogéneos o heterogéneos (o de muestreo). Un calorímetro de muestreo intercala capas de materiales

sensibles, que son capaces de medir la energía de las lluvias y materiales de producción de las lluvias. Un ejemplo de este tipo de calorímetros es el *Hadron Calorimeter (HCAL)* del experimento CMS (*Compact Muon Solenoid*) en CERN. El calorímetro homogéneo, por otro lado, está compuesto en su totalidad por materiales sensibles. Estos materiales sensibles pueden ser cristales centelleadores (BGO, LYSO), cristales de vidrio de plomo, argón líquido, entre otros.

Este trabajo en particular utilizará un calorímetro de tipo electromagnético, por lo que se ahondará más en este tipo de calorímetros en la próxima sección.

En un calorímetro electromagnético la resolución espacial, es decir, la capacidad de identificar partículas muy cercanas entre sí, es un factor clave a considerar. El problema de identificar partículas muy cercanas entre sí es universal en la física de altas energías (o HEP, por las siglas de *High Energy Physics*) y es de interés particular en los colisionadores Electrón-Ión debido a la alta proporción de fotones, electrones y positrones en la colisión. Un caso muy común en que se presenta este problema es en el decaimiento de un pión neutro en dos fotones muy cercanos entre sí. Si la resolución no es suficiente es muy difícil diferenciar dos fotones generados por este tipo de decaimiento de un sólo fotón de alta energía.

Un ejemplo concreto es el detector ECAL (por *Electromagnetic Calorimeter*) del colisionador CMS en CERN, que tiene como propósito identificar el decaimiento del bosón de Higgs en dos fotones. Para hacer esto es necesario identificar los diferentes eventos background que se comportan similar, pero no corresponden al bosón de Higgs. Uno de estos eventos es el decaimiento de un pión neutro en dos fotones cercanos entre sí. El detector ECAL está conformado por una matriz de cristales del tipo $PbWO_4$ (tungsteno de plomo) como se muestra en la Figura 1.2. El tamaño transversal de cada cristal define la resolución espacial del detector. Hacia los extremos del detector el ángulo entre los fotones producidos por el decaimiento de un pión neutro es mucho menor y la resolución espacial del detector ECAL no es suficiente, como se muestra mediante simulaciones en la Figura 1.3b. Con la finalidad de solucionar este problema se utiliza un detector de alta resolución y sensible a los fotones incidentes que se coloca delante del detector calorímetro, conocido como detector preshower. El detector preshower es capaz de distinguir fotones muy cercanos entre sí y de esta manera rechazar eventos background provenientes del decaimiento del pion neutro. El detector preshower de CMS esta compuesto por capas de plomo, que producen las lluvias

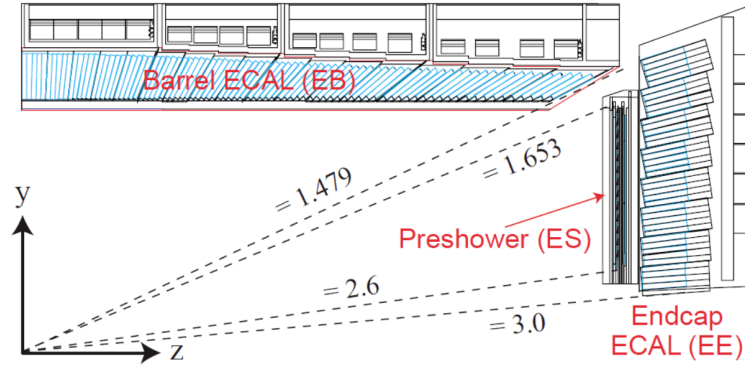
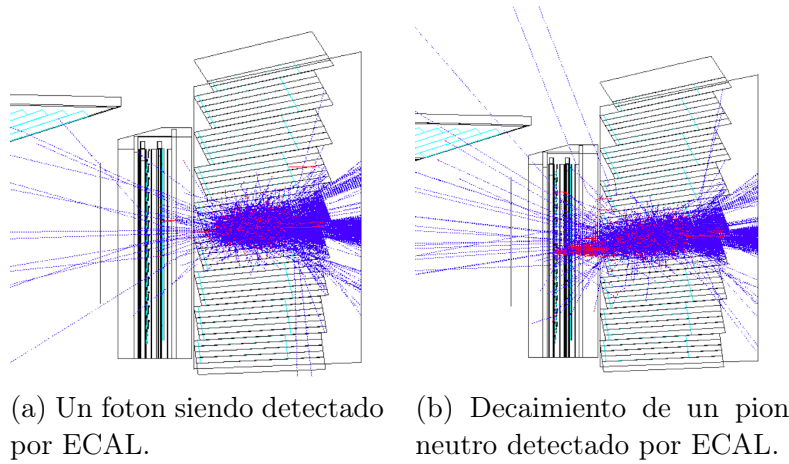


Figura 1.2: Diseño de detector ECAL en CMS.



(a) Un fotón siendo detectado por ECAL. (b) Decaimiento de un pión neutro detectado por ECAL.

Figura 1.3: Simulaciones detector ECAL (CERN).

electromagnéticas y capas sensibles de silicio.

1.2. Contexto

Como ya se mencionó, el problema de identificación de piones neutrales es de particular interés para cualquier colisionador Electrón-Ión. En este contexto, los laboratorio BNL (*Brookhaven National Laboratory*) y JLab (*Thomas Jefferson Laboratory*), en conjunto con la Oficina de Física Nuclear del Departamento de Energía de Estados Unidos convocaron propuestas en torno a un programa R&D en busca de satisfacer los requerimientos de un futuro colisionador EIC (*Electron-Ion Collider*). En torno a esta convocatoria, fue aceptada la propuesta por parte de investigadores del Centro Científico y Tecnológico de Valparaíso

(CCTVal) [2] para construir un detector preshower para calorímetros. El detector podría ser utilizado en experimentos como CMS en CERN o IC de CLAS en Jefferson Lab.

Si bien existen calorímetros preshower basados en materiales como silicio, plomo o tungsteno, lo innovador de la propuesta del CCTVal es la exploración de un calorímetro preshower basado en una matriz de cristales, lo que permite estudiar una alternativa con costos menores y características muy diferentes a la tecnología actualmente en uso. Los detectores basados en cristales se han desarrollado con gran éxito en campos como la medicina nuclear, para la visualización mediante radiación de tejidos internos (usando tomógrafos de emisión de positrones). Cristales destelladores del tipo LYSO (Lutetium Yttrium Silicon Oxide) han logrado un gran protagonismo debido a que ofrecen cualidades de producción de lluvia excepcionales, además que la industria de producción industrial de estos cristales está bien desarrollada. Estas ventajas hacen que la exploración de un diseño basado en cristales para la detección de alta resolución espacial sea muy interesante. El detector preshower debe cumplir varias exigencias:

- La capacidad de identificar dos partículas muy cercanas entre sí dependerá del tamaño transversal de cada cristal y de la sección longitudinal de estos. Ambos factores pueden ser optimizados para lograr una resolución que permita identificar la mayoría de los eventos problemáticos. Si la distancia entre dos fotones es menor que la mitad del tamaño transversal de cada cristal, será muy difícil identificar ambos fotones correctamente, por lo que es recomendable utilizar cristales de tamaño transversal menor.
- La superficie no sensible del detector preshower debe ser minimizada, con la finalidad de evitar la pérdida de energía debido a interacción con materiales de lectura u otros instrumentos. Para esto se propone la utilización de un sistema especial de lectura ubicado en el frente del detector.

En lo que continúa del trabajo se procederá a describir de manera general la composición y funcionamiento del detector preshower, para luego continuar con la presentación del algoritmo de reconstrucción y finalmente estudiar la efectividad del algoritmo mediante experimentos en datos simulados. En la Sección 2 se introducirán algunos conceptos fundamentales en física de detectores, vitales para comprender de manera general el

funcionamiento del detector preshower. Luego, en la Sección 3, se explicará de manera general la composición y funcionamiento del detector preshower. En la Sección 4 se analizarán las diferentes alternativas presentes en la literatura para la reconstrucción de partículas en calorímetros. Se continuará en la Sección 5 con la presentación del algoritmo de reconstrucción propuesto. En la Sección 6, experimentos en datos simulados serán expuestos. Finalmente, en la Sección 7, se concluirá sobre los resultados.

Capítulo 2

Detectores de partículas

En esta sección se analizarán los conceptos físicos relacionados a la detección en calorímetros. Para un tratamiento de los conceptos físicos básicos relacionados a la interacción general de partículas con la materia el lector se puede dirigir al Apéndice A.

2.1. Física de los detectores calorímetros

Si bien la física involucrada en el funcionamiento de los detectores es amplia y variada aquí nos concentraremos en los detectores del tipo calorímetro, que son el eje principal de este trabajo. En lo que sigue de esta sección se realizará una descripción de los procesos que ocurren en un detector calorímetro de manera general con la finalidad de entregar una base teórica para comprender las decisiones de diseño en la construcción del detector preshower. También, como ya se mencionó, un calorímetro puede ser del tipo electromagnético o hadrónico según el tipo de partícula que detecte. Aquí nos concentraremos en los elementos constitutivos de un calorímetro del tipo electromagnético y su funcionamiento.

Un calorímetro se conforma de un material de producción de lluvias electromagnéticas y un material de lectura de estas lluvias. A continuación se describirán los procesos que originan las lluvias electromagnéticas.

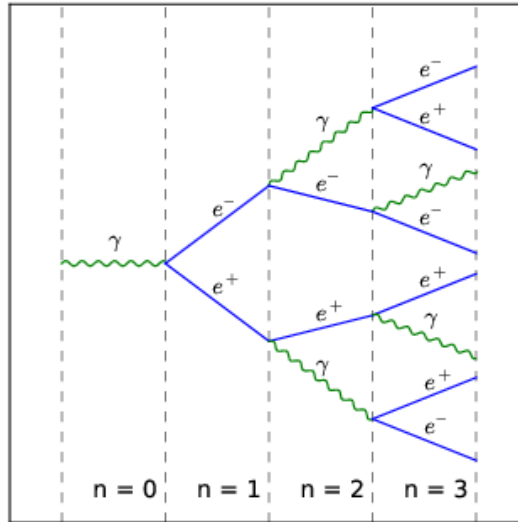


Figura 2.1: Diagrama de una lluvia electromagnética, descrita por el modelo Heitler.

2.1.1. Lluvias electromagnéticas

La generación de lluvias electromagnéticas es un proceso producido por el efecto combinado de la producción de pares y la emisión por radiación de frenado (o bremsstrahlung). A altas energías un fotón se convierte en un par e^+e^- , estos a su vez emiten fotones mediante el proceso de radiación de frenado, estos fotones producen pares e^+e^- y el proceso continúa intercalando ambos procesos como se muestra en la Figura 2.1, hasta alcanzar la energía crítica. El proceso es descrito de manera aproximada por el modelo de Heitler.

En el modelo de Heitler, en promedio un fotón se convertirá en un par electrón-positrón en una longitud de radiación X_0 . Entonces un fotón de E_0 que entra en el calorímetro producirá alrededor de 2^t partículas al cabo de t longitudes de radiación. Es decir,

$$N(t) = 2^t.$$

Además, dado a que en cada paso la energía de cada par será un medio de la partícula que lo produce, entonces

$$E(t) = \frac{E_0}{2^t},$$

o, trabajando esta ecuación, la energía media de las partículas a la profundidad t será,

$$t(E') = \frac{\ln(E_0/E')}{\ln 2}.$$

Al alcanzar la energía crítica E_c , la lluvia se detiene. En esta energía los pares e^+e^- empiezan a perder su energía por colisiones atómicas en vez que por radiación de frenado. Asumiendo que la cascada (o lluvia) se detiene abruptamente a esta energía tenemos que el máximo número de partículas se alcanza a

$$t_{max} = \frac{\ln(E_0/E_c)}{\ln 2},$$

donde el número de partículas es,

$$N_{max} = e^{t_{max} \ln 2} = E_0/E_c,$$

y la suma de recorrido de las partículas es,

$$L = X_0 \int_0^{t_{max}} N(t) dt \approx X_0 \frac{E_0/E_c}{\ln 2}.$$

Es decir, midiendo N_{max} , L o una cantidad proporcional es posible conocer el valor de energía inicial de la partícula incidente. Se debe notar eso si, que el modelo es solo una aproximación simple de la lluvia, es posible obtener aproximaciones más cercanas mediante métodos de Monte Carlo (es imposible obtener una forma analítica).

El desarrollo longitudinal de la lluvia se puede cuantificar utilizando el radio de Moliere. A la profundidad máxima, el valor medio del ángulo de apertura de las partículas puede ser aproximado como

$$\langle \theta \rangle = \frac{21.2 MeV}{E_c},$$

entonces, el radio de Moliere se define como

$$\begin{aligned}\rho_M &= \langle \theta \rangle X_0 \\ &= \frac{21.2MeV}{E_c} X_0,\end{aligned}\tag{2.1.1}$$

que puede ser comprendido como la distancia transversal que una partícula en la energía crítica viaja al atravesar una longitud de radiación. Además, el 90 % de la energía de la lluvia esta contenida en alrededor de un ρ_M . Esta cantidad es muy importante en el diseño del calorímetro cuando la resolución espacial es un factor clave a considerar. Un menor radio de Moliere significa una mejor resolución espacial y por lo tanto una mejor capacidad para separar cascadas solapadas.

2.1.2. Materiales de construcción

Si bien, un calorímetro electromagnético puede estar construido de diferentes tipos de materiales (líquidos y sólidos, orgánicos e inorgánicos), en este trabajo nos concentraremos en la construcción de calorímetros homogéneos con cristales centelleadores inorgánicos (BGO, LYSO, LSO). Dos son los principales componentes de este tipo de calorímetros: el material centelleador y el sistema de lectura. Ambos se procederán a describir a continuación.

2.1.3. Cristales centelleadores

Los cristales centelleadores son al mismo tiempo materiales activos (generadores de lluvias electromagnéticas) y sensibles (mediante la generación de luz por centelleo). Estos son capaces de generar luminiscencia al desexcitarse los átomos que son excitados por el paso de partículas cargadas. En específico, la captura de electrones por impurezas (o centros de activación) del cristal produce emisión de luz al desexcitarse. Existen dos tipos de emisión, la fluorescencia, de tipo muy rápido ($\tau \sim 10^{-8}seg.$) y la fosforescencia, que es muy lenta ($\tau_d \sim \mu s$). Para los detectores se prefieren materiales con tiempos de desexcitación muy rápidos. La respuesta del centelleador es lineal a la energía depositada y se busca una alta eficiencia, rápida respuesta, con transparencia a la radiación fluorescente y un espectro de

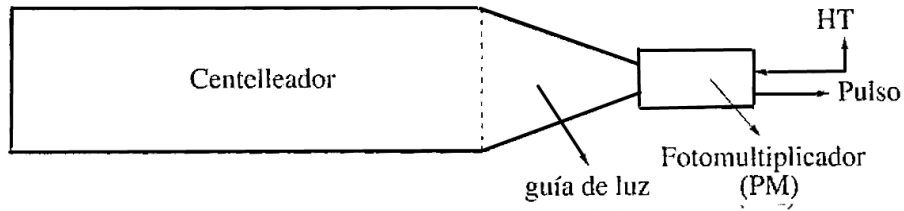


Figura 2.2: Diagrama simple de un centelleador conectado a un fotomultiplicador.

emisión similar al de los fotomultiplicadores utilizados en la lectura. Un valor importante para medir la calidad de un centelleador es su respuesta luminosa. Esta se define como la energía necesaria para crear un fotón de centelleo, o

$$\eta = \frac{\Delta E}{N_{\gamma}}.$$

Este valor se puede entender como la eficiencia del detector para producir una señal medible.

Los fotones producidos por el centelleador son comúnmente traspasados al fotomultiplicador que convierte los pulsos de luz en corriente eléctrica. La luz es trasladada al fotomultiplicador por guías de luz que pueden, a su vez, cambiar el rango de frecuencia de la luz emitida por el centelleador. Un diagrama se muestra en la Figura 2.2.

2.1.4. Fotomultiplicadores

Los fotomultiplicadores son capaces de convertir pulsos de luz en pulsos de corriente eléctrica medibles. Se componen del fotocátodo, dínodo y ánodo. El fotocátodo emite electrones en respuesta a los fotones incidentes mediante efecto fotoeléctrico. El material que lo compone debe ser un semiconductor con gran probabilidad de efecto fotoeléctrico. Comúnmente, los electrones producidos por el efecto fotoeléctrico se refieren como fotoelectrones γ_e . La eficiencia cuántica se puede definir como

$$\eta(\lambda) = \frac{N_{\gamma_e}}{N_{\gamma_{inc}}},$$

donde N_{γ_e} y $N_{\gamma_{inc}}$ es el número de fotoelectrones producidos y el número de fotones incidentes respectivamente. Además, λ es la longitud de onda de los fotones incidentes.

Los dínodos son los encargados de multiplicar los fotoelectrones producidos en el cátodo.

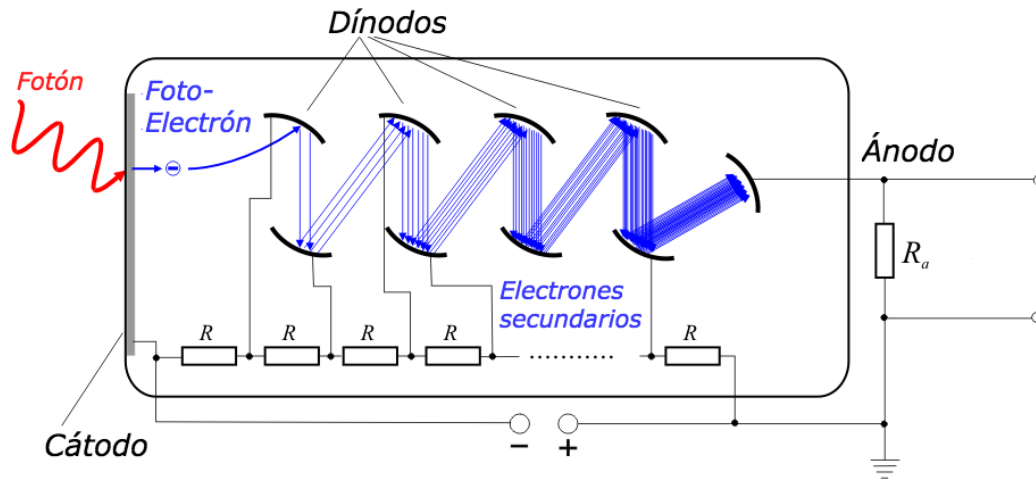


Figura 2.3: Fotomultiplicador en funcionamiento.

En estos, los fotoelectrones son acelerados en presencia de un campo eléctrico y dirigidos a los dínodos, dispuestos según la Figura 2.3, los que a su vez producen la emisión secundaria de más electrones.

El ánodo recoge la avalancha final de electrones producidos por los dínodos y produce un pulso eléctrico, normalmente un voltaje en función del tiempo $V(t)$. La ganancia del fotomultiplicador se puede definir como la amplificación en la cadena de dínodos. Si cada dínodo produce δ electrones secundarios, entonces la ganancia será

$$G = \delta^n = (KV_d)^n,$$

donde V_d es el potencial entre dínodos, n el número de dínodos y K es una constante. Se debe notar además que debido a que el funcionamiento de los centelleadores suele ser muy rápido, el fotomultiplicador también debe tener una resolución temporal muy alta.

Habiendo realizado un resumen muy general de los conceptos principales a tener en cuenta en el diseño de un calorímetro, en la próxima sección presentaremos el calorímetro preshower, su estructura general y las bases de su funcionamiento. La sección siguiente constituirá una visión general del diseño del preshower, sólo lo necesario para comprender a cabalidad el algoritmo de reconstrucción propuesto para el detector preshower.

Capítulo 3

Detector Preshower

En esta sección se procederá a describir de manera general, el diseño y funcionamiento del detector preshower. La finalidad del detector preshower es detectar partículas (neutras y cargadas) muy cercanas entre sí, por lo que el diseño está pensado en pos de esta finalidad. Un primer acercamiento por parte del equipo experimental de física nuclear y de altas energías de la UTFSM corresponde a un prototipo simple para probar el funcionamiento del detector antes de entrar en una fase de producción. A continuación, se presentará el diseño general del prototipo de preshower y su funcionamiento.

3.1. Diseño

El detector preshower es un detector calorímetro electromagnético del tipo homogéneo. Como se describió en la Sección 2, un detector calorímetro se conforma principalmente de un material de producción de lluvias (pueden ser cristales orgánicos e inorgánicos, líquido, gas, etcétera) y un sistema de amplificación y lectura de las señales producidas (por ejemplo, un sistema de fotomultiplicadores). En el prototipo de preshower estas dos secciones principales son la matriz de cristales, que producen las lluvias electromagnéticas y el sistema de lectura que convierte esta lluvia en una señal medible. Un esquema a alto nivel del diseño del detector se observa en la Figura 3.1.

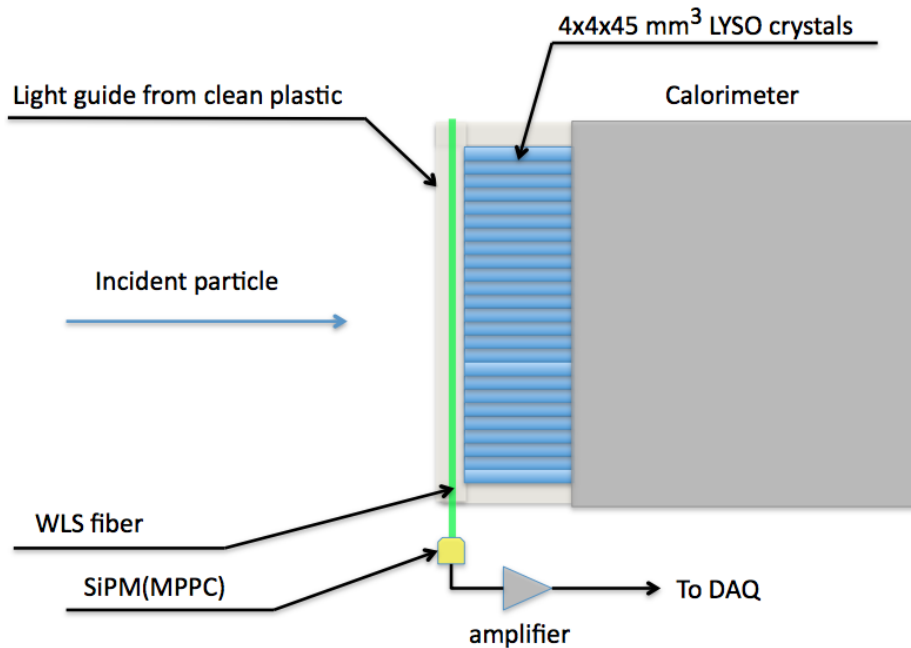


Figura 3.1: Prototipo de preshower, en este se puede observar la matriz de cristales en color azul, el sistema de lectura en color verde (fibras) y amarillo (fotomultiplicadores) y el material de cobertura en gris.

3.1.1. Matriz de cristales

La matriz de cristales se conforma de 25x25 (625 en total) cristales centelleadores del tipo LYSO ($Lu_{1.8}Y_{0.2}SiO_5$). Como fue descrito en la sección previa, los materiales centelleadores son capaces de producir luz mediante el proceso de luminiscencia, cuando partículas cargadas interactúan con el material. Los cristales LYSO ofrecen ciertas características que los hacen una muy buena elección para la construcción del preshower, estos presentan un tipo de luminiscencia muy veloz e intensa, lo que permite reducir los tamaños del calorímetro y al mismo tiempo producir una señal medible por el sistema de lectura. Otros materiales comúnmente utilizados como cristales centelleadores son BGO ($Bi_4Ge_3O_{12}$) y PWO ($PbWO_4$). Valores característicos de cada uno de estos materiales se muestran en el Cuadro 3.1, donde se observa claramente la superioridad en luminosidad (fotones producidos por energía incidente) de los cristales LYSO sobre las otras opciones.

Otro valor importante a considerar es la longitud de atenuación del cristal. Si bien el fabricante del cristal reporta un valor de λ de 12 [mm], otros estudios reportan valores muchos

Material	Densidad [gr/cm^3]	Luminosidad [ph/MeV]	Radio Moliere [cm]	Longitud atenuación [cm]
BGO	7,13	8.200	2,23	1,20
PWO	8,2	200	2,00	0,89
LYSO	7,1	32.000	2,07	1,14

Cuadro 3.1: Algunos valores para varios tipos de materiales centelleadores.

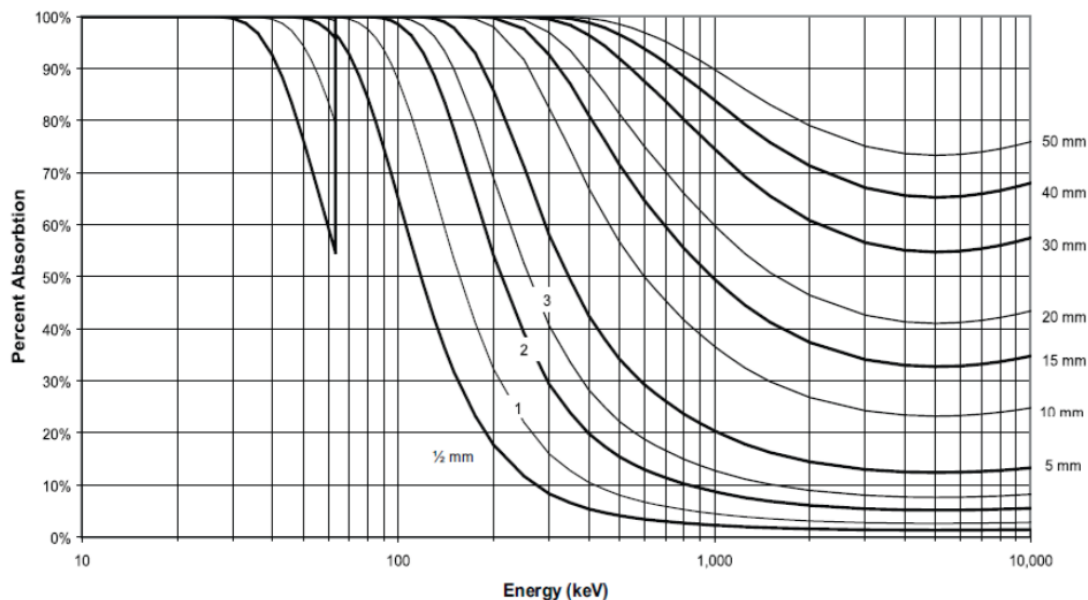


Figura 3.2: Porcentaje de absorción dada la energía para diversos largos de cristal LYSO.

mayores [3, 4]. Recordemos que la longitud de atenuación se relaciona con la probabilidad de absorber o no una partícula en el material, por lo que mayores longitudes de radiación implicarán cristales de mayor tamaño con la finalidad de aumentar la probabilidad de interacción con los cristales (de otra forma las partículas pueden pasar sin interactuar y por lo tanto sin ser detectadas). Las probabilidades de interacción de una partícula gamma entregadas por el fabricante, dependiendo de la energía y para varios largos del cristal, se muestran en la Figura 3.2. Para un largo de 45 [mm] y un fotón de 511 [keV] la probabilidad de absorción es aproximadamente 95 %, mientras que bajando a un largo de 4 [mm] la probabilidad disminuye a un valor cercano a 4 %. Esto demuestra la importancia de conocer correctamente la longitud de atenuación con la finalidad de seleccionar correctamente el tamaño del cristal usado.



Figura 3.3: Disposición de la matriz de fibras ópticas.

Finalmente, se utilizan cristales de tamaño $4 \times 4 \times 45$ [mm³], lo que (según las especificaciones del fabricante) supone un diseño de 4 longitudes de absorción.

Es importante aislar ópticamente los cristales para evitar la salida de luz de cada cristal. La luz viaja en el cristal confinada por el material reflectante hasta alcanzar los extremos del cristal donde es absorbida y re-emitida por fibras ópticas. Además, experimentos han demostrado que es posible optimizar la longitud de atenuación con diferentes materiales reflectantes [5]. Los experimentos realizados para seleccionar el material reflectante se pueden encontrar en [6].

3.1.2. Sistema de lectura

La luz producida por los cristales centelleadores es recolectada por una matriz de fibras ópticas que la llevan hasta el sistema de lectura. La fibra óptica absorbe la luz emitida y la re-emite en un diferente espectro de frecuencia, lo que facilita la lectura. La matriz de fibras se muestra en la Figura 3.3, esta es sostenida por una placa plástica tratada de manera especial, de forma que permite el paso de luz por la cara que da a los cristales y al sistema de lectura, mientras que refleja la luz por las caras opuestas.

La disposición del sistema de lectura en el detector preshower es poco común. Los lectores se encuentran en dos de los bordes perpendiculares de la matriz como se muestra en la Figura 3.4. Debido a esto, la señal observada consistirá en dos vectores de datos unidimensionales que luego deben ser procesados de manera de obtener la posición incidente en el plano. La luz emitida por los cristales centelleadores es guiada por las fibras ópticas a dos arreglos (uno para cada borde), de 25 (50 en total) fotomultiplicadores de silicio (MPPC

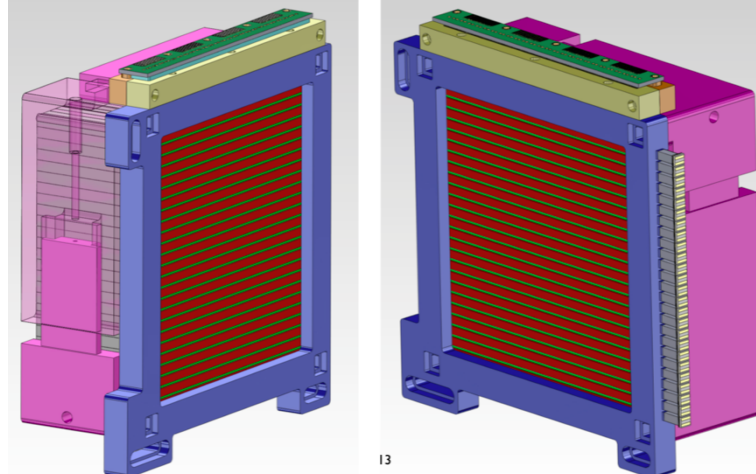


Figura 3.4: Prototipo de detector preshower y disposición del sistema de lectura en ambos bordes. En uno de los bordes se muestra la disposición de fotomultiplicadores y en el otro la salida de las guías de luz.

por *Multi-Pixel Photon Counter*).

El funcionamiento de los fotomultiplicadores fue descrito en la Sección 2.1.4. Estos convierten la luz de entrada en pulsos de corriente medibles mediante la amplificación de los fotoelectrones en los diodos. En específico, los MPPC utilizados en el detector preshower poseen 1440 fotodiodos de avalancha, cada uno con una ganancia de 2.75×10^5 . Finalmente, el pulso es digitalizado utilizando módulos de conversión de carga y conversores análogo-digital (ADC). Los datos entregados en formato digital son almacenados en un buffer de salida que puede ser leído con posterioridad. Los buffer guardan los valores de carga [pC] de cada uno de los fotomultiplicadores, obteniendo finalmente dos arreglos de 25 valores (25 para cada eje) de carga entre 0-400[pC] ($1 \text{ coulumb} = 6.25 \times 10^8 \text{ electrones}$). El funcionamiento del sistema de lectura es mucho más complejo que lo explicado aquí, sin embargo, no es estrictamente necesario comprender a cabalidad el mecanismo de lectura para explicar el algoritmo de reconstrucción. Un tratamiento más completo del sistema de lectura puede ser encontrado en [6]. Importante es notar el formato de salida del sistema de lectura y las posibles complicaciones que este puede significar para la reconstrucción de las partículas incidentes. En las próximas secciones se procederá a ahondar más en este tema.

En la siguiente sección se analizarán las propuestas de la literatura para cada uno de los pasos del algoritmo de reconstrucción. Si bien, muchas de estas propuestas se realizan en el contexto de detectores con capacidad para medir la posición en el plano, es necesario adaptar

su funcionamiento al caso uni-dimensional que tratamos en este trabajo. Al conocimiento del autor de esta propuesta, no existe trabajo en algoritmos de reconstrucción para una arquitectura similar a la del detector preshower, por lo que mucho de lo propuesto en la Sección 5 será trabajo original, pero tomando ideas de los algoritmos de reconstrucción propuestos para otros detectores.

Capítulo 4

Estado del Arte

El algoritmo de reconstrucción es la propuesta principal de este trabajo, por lo que en esta sección se estudiarán las propuestas de la literatura para cada una de las partes de un algoritmo de reconstrucción. Un algoritmo de reconstrucción se compone de varios sub-procesos, que en conjunto permiten reconstruir la energía y posición de una o más partículas detectadas. Si bien existen varias posibilidades en el proceso de reconstrucción, tales como la reconstrucción del ángulo de incidencia y la reconstrucción de trayectoria [7], las principales tareas que interesan en este trabajo son: el agrupamiento de los datos, la identificación de máximos en los clústeres, la identificación y separación (cuando es posible) de partículas solapadas, el rechazo de partículas solapadas cuando no pueden ser separadas y la reconstrucción de la posición y energía de cada una de estas partículas.

Para completar estas tareas un algoritmo de reconstrucción suele componerse de las siguientes sub-secciones:

- Un algoritmo de clústering, encargado de identificar y separar cada una de las partículas detectadas por el detector. La entrada para este algoritmo depende del sistema de lectura, ya que es el primer paso en el proceso de reconstrucción. La entrada puede corresponder a mediciones de energía o conteo de fotoelectrones. Estos valores pueden ser mediciones bidimensionales (x,y) o unidimensionales (como en el caso del detector preshower). Comúnmente, los valores de energía o conteos de fotoelectrones deben ser agrupados de manera de representar partículas incidentes. Además, es deseable eliminar posibles fuentes de ruido en las lecturas.

- En una segunda fase se deben identificar los máximos de cada uno de estos clústeres, varios máximos indican que más de una partícula corresponden a ese clúster.
- En el caso de identificar varios máximos en el mismo clúster es necesario separar la cantidad de fotoelectrones correspondientes a cada una de estas partículas. El algoritmo para la separación de partículas solapadas se conoce como algoritmo de *unfolding*¹.
- Si sólo un máximo es identificado, aún es posible que la lluvia provenga de dos o más partículas muy cercanas entre sí. Es deseable que el algoritmo sea capaz de identificar estos casos, con la finalidad de no reconstruir la posición erróneamente, como si fuese una sola partícula. A esto se le llama algoritmo de rechazo y es encargado de obviar las lluvias de este tipo.
- Luego de identificar la energía correspondiente a cada partícula para cada uno de los clúster detectados (y no rechazados en el paso anterior), es necesario reconstruir la posición incidente de la partícula basándose en los valores detectados por el sistema de lectura. El algoritmo de reconstrucción de posición se encarga de estimar la posición de la partícula incidente utilizando la energía depositada y la posición de las celdas de cada clúster. La energía de la partícula puede ser estimada luego integrando los valores de cada clúster.

A continuación se analizará el estado del arte para cada uno de estos sub-procesos por separado.

4.1. Clustering de partículas

La entrada al algoritmo es la medición de cada uno de los lectores del sistema. La estructura de esta entrada depende de la configuración del sistema de lectura, este puede ser capaz de medir la energía (o fotoelectrones) depositada en cada celda o sólo medir una proyección en cada eje de la energía depositada (como en el caso del preshower). La entrada depende del sistema de lectura, pero comúnmente corresponde a un conjunto de valores de energía o fotoelectrones que deben ser agrupados de manera de representar partículas

¹Se debe notar aquí, para evitar confusiones, que el nombre algoritmo de *unfolding* es usado en varios campos con significado diferente, la definición utilizada aquí es la correspondiente a [8].

incidentes. Las partículas incidentes usualmente depositan su energía en varias celdas o cristales del calorímetro. La finalidad de los algoritmos presentados aquí es agrupar estas celdas. El algoritmo debe también eliminar posibles fuentes de ruido en las lecturas en caso de ser posible. El ruido en los calorímetros proviene de dos fuentes principales. Primero, se introduce ruido por distintos procesos en la electrónica de lectura. Segundo, existe ruido conocido como de apilamiento (*pile-up*). Este tipo de ruido proviene de interacciones extras que ocurren al mismo tiempo que la partícula atraviesa el detector.

Varios métodos han sido propuestos en la literatura para formar los clústeres en calorímetros basados en cristales. Entre los algoritmos propuestos se pueden encontrar métodos de construcción incremental de los clústeres que funcionan muy bien en la práctica. También es posible encontrar algoritmos más avanzados que utilizan métodos probabilísticos para construir los clústeres.

En lo que continúa analizaremos algunos de los algoritmos más utilizados en el trabajo de clústering en detectores con geometrías relativamente similares al detector preshower.

4.1.1. Ventana deslizante

Este algoritmo es propuesto por el experimento ATLAS y es utilizado en el proceso de reconstrucción del calorímetro electromagnético y el calorímetro hadrónico [9]. Los clústeres pueden utilizar información combinada de ambos calorímetros (esto es útil para reconstruir jets e identificación de leptones Tau) o utilizar sólo la información del calorímetro electromagnético (útil para la reconstrucción de electrones y fotones).

El detector calorímetro en ATLAS se compone tanto de secciones longitudinales como secciones transversales, las secciones transversales se extienden en el espacio $\eta - \sigma$, y existen cuatro secciones longitudinales en el caso del calorímetro electromagnético: *middle*, *strips*, *pre-sampler* y *back*.

El algoritmo está constituido por tres pasos principales: construcción de torres, búsqueda de pre-clústeres y llenado de clústeres.

Construcción de torres

El espacio transversal $\eta - \sigma$ es segmentado en $N_\eta \times N_\phi$ celdas de tamaño $\Delta\eta \times \Delta\phi$ formando torres longitudinales de celdas donde la energía depositada es sumada. Los valores N_η y N_ϕ

son definidos de manera de contener la mayoría de la lluvia electromagnética producida por el calorímetro y depende, entre otras cosas, del material del detector. Se puede dar el caso en que una celda en alguna de las secciones longitudinales pertenezca a varias torres, en tales casos la energía de las celdas es repartida en proporción al área de la celda en cada torre.

Búsqueda de clústeres

Se define una ventana de tamaño fijo $N_\eta^{window} \times N_\phi^{window}$ en unidades del tamaño de torre $\Delta\eta \times \Delta\phi$. La ventana es movida a lo largo de las torres construidas y clústeres son formados en los máximos locales de energía acumulada en las torres. La energía de la torre es definida como la suma de energías en todas las celdas que forman parte de la torre. Además, se requiere que la energía de la torre supere un umbral E_T^{thresh} .

La posición en el eje $\eta - \phi$ de los clústeres obtenidos en este paso es calculada utilizando el centro de masa con pesos iguales a la energía en cada una de las celdas incluidas en las torres. Es decir, sean $\{T_1, T_2, \dots, T_C\}$ las torres incluidas en el clúster. Luego, sean $\{x_1^i, x_2^i, \dots, x_{k_i}^i\}$ las posiciones de las celdas incluidas en la torre T_i y $\{w_1^i, w_2^i, \dots, w_{k_i}^i\}$ las energías de las mismas celdas. Entonces la posición se calcula como

$$X = \frac{\sum_{i=1}^C \sum_{j=1}^{k_i} w_j^i x_j^i}{\sum_{i=1}^C \sum_{j=1}^{k_i} w_j^i}.$$

Es posible también definir tamaños menores para las ventanas utilizadas al calcular la posición, es decir utilizar $N_\eta^{pos} \times N_\phi^{pos}$, esto con la finalidad de minimizar el ruido introducido en el calculo de posición.

Finalmente, se definen como clústeres duplicados aquellos que tengan posiciones centrales dentro de $\Delta\eta_{dupl} \times \Delta\phi_{dupl}$ (en unidades de torres). En tales casos sólo el clúster con mayor cantidad de energía integrada es conservado.

Este algoritmo es utilizado en el detector ATLAS para la reconstrucción de electrones y fotones [10, 11] y la reconstrucción de leptones Tau [12] (aunque en versiones actuales del algoritmo esto se cambió a clústering topológico).

4.1.2. Clustering topológico

Otro método de clustering utilizado en el experimento ATLAS es el clúster topológico [9, 13]. Este algoritmo es más flexible que el algoritmo de clustering de ventana deslizante presentado en la sección anterior ya que permite construir clústeres de tamaño variable. El algoritmo consiste de tres pasos principales: búsqueda de semillas, búsqueda de vecinos, construcción de clústeres.

Búsqueda de semillas

El algoritmo comienza con la búsqueda de semillas para construir los clústeres. Para esto se identifican las celdas con un valor de energía mayor a un umbral t_{seed} definido de antemano. La definición de t_{seed} se hace en base a valores esperados de ruido, el cual depende de las características electrónicas y físicas del detector.

Búsqueda de vecinos

La lista de semillas es ordenada en orden descendente basándose en el valor de energía en cada celda. Se marcan todas las semillas como no usadas. Entonces, se itera por cada semilla no usada de manera descendente. En cada paso se crea un proto-clúster con la semilla actual como único integrante. Luego, se agregan las celdas vecinas a la semilla que superen un umbral t_{cell} y que no estén marcadas como usadas. Cada vez que se agrega una celda a un proto-clúster se marca como usada. Si la celda es adyacente también a otro proto-clúster entonces ambos son combinados. Si la celda agregada, además, supera un umbral $t_{neighbor}$, entonces la celda es utilizada como semilla para expandir el proto-clúster utilizando el mismo mecanismo. Esto último asegura que las colas de las lluvias electromagnéticas no sean descartadas pero al mismo tiempo asegura que no se agregue ruido electrónico y de apilamiento. El proceso continua hasta marcar todas las semillas como usadas.

Construcción de clústeres

Como paso final, los proto-clústeres que quedan del paso anterior (algunos fueron combinados) son ordenados de mayor a menor (respecto a la energía total del clúster u otro valor equivalente) y convertidos en clústeres. Adicionalmente se pueden eliminar los que

tengan energía total E_t menor a un umbral t_E , con la finalidad de evitar clústeres contruidos de ruido.

Notar que es posible que los clústeres obtenidos con el algoritmo de clústering topológico correspondan a más de una partícula incidente, debido al solapamiento de lluvias. Por lo anterior, es necesario agregar un paso de división de clústeres, esto se tratará en la Sección 4.3.

El algoritmo de clúster topológico ha sido utilizado extendidamente en la reconstrucción de jets y MTE (*missing transverse momentum*) en los calorímetros electromagnético y hadrónico de ATLAS [14, 15, 16], mostrando una gran eficiencia en la reconstrucción [13].

4.1.3. Clústering Fuzzy

Otro método exitosamente aplicado en la literatura es el algoritmo de clústering fuzzy c-means. El algoritmo fuzzy c-means es una extensión del algoritmo k-means, que es una técnica de clústering que aprende de manera no supervisada a encontrar k clústeres en los datos. Para el trabajo de reconstrucción de partículas en detectores calorímetros, esta técnica (y en general cualquier técnica de clústering duro) no es adecuada ya que asigna cada dato a un sólo clúster, lo que no permite lluvias solapadas, que como ya se observó son muy comunes. Los métodos fuzzy, en cambio, permiten que cada punto pertenezca en menor o mayor grado a diferentes clústeres. Estos métodos combinan los algoritmos de clústering básicos con la teoría fuzzy [17] que permite agregar conceptos de imprecisión e incerteza a los métodos duros. Existen estudios de la aplicación de este algoritmo para la reconstrucción de partículas en calorímetros [18, 19], sin embargo, al conocimiento de este autor no hay aplicaciones reales (en detectores) de este algoritmo.

En [18] el algoritmo fuzzy c-means y una extensión de este, el algoritmo fuzzy c-means dinámico, son aplicados en la reconstrucción de datos simulados de un detector calorímetro de muestreo. En aquel trabajo se demuestra la utilidad de los algoritmos fuzzy en la tarea de reconstrucción debido a la capacidad de estos métodos de manejar las lluvias solapadas automáticamente.

Fuzzy c-means comienza con un conjunto de datos $X = (x_1, x_2, \dots, x_N)$ con $x_i \in \mathbb{R}^d$, para estos datos busca el conjunto de clústeres con centros $V = (v_1, v_2, \dots, v_C)$ con $v_i \in \mathbb{R}^d$ que

minimicen la función de costo

$$J_m(U, V; X) = \sum_{k=1}^N \sum_{i=1}^C (u_{ik})^m \|x_k - v_i\|^2,$$

donde U consiste de los valores u_{ik} que corresponden a los grados de pertenencia del punto k al clúster i , $\|x\| = \sqrt{x^T x}$ es la norma de producto interno y el valor $m \in [1, \infty[$ (factor fuzzy) define el grado de fucificación del algoritmo, donde m cercano a 1 corresponde a un simple k-means. Imponiendo restricciones

$$\begin{aligned} \sum_{k=1}^N u_{ik} &> 0, i \in \{1, \dots, C\}, \\ \sum_{i=1}^C u_{ik} &= 1, k \in \{1, \dots, N\}. \end{aligned}$$

Es posible resolver el problema iterando el siguiente algoritmo

$$u_{ik} = \left[\sum_{j=1}^C \left(\frac{D_{ik}}{D_{jk}} \right)^{\frac{2}{m-1}} \right]^{-1},$$

con $D_{ik} = \|x_k - v_i\| > 0$ para todo i y k . Y

$$v_i = \frac{\sum_{k=1}^n (u_{ik})^m x_k}{\sum_{k=1}^n (u_{ik})^m}.$$

El algoritmo procede alternando las estimaciones de V y U hasta alcanzar un número definido de iteraciones o alcanzar un umbral de error ϵ definido de ante mano. Es necesario definir con anterioridad el número de clústeres C . Para decidir entre distintas opciones de C se suele utilizar un índice de validación. El índice Xie-Beni se define como

$$v_{XB}(U, V : X) = \frac{\sum_{i=1}^C \sum_{k=1}^n u_{ik}^2 \|x_k - v_i\|^2}{n(\text{mín}_{i \neq j} \{\|v_i - v_j\|\})},$$

el cual al ser minimizado busca maximizar la separación entre clústeres mientras que se minimiza la separación entre los puntos en cada clúster.

Además del algoritmo fuzzy c-means básico, en [18] se presenta una modificación que

funciona de manera online. Esto último permite al método adaptarse al patrón de datos de manera dinámica. El algoritmo funciona de la siguiente manera:

1. Con una cantidad de datos iniciales se definen los parámetros μ (umbral de pertenencia), ϵ (umbral de error) y los valores C_{min} y C_{max} del algoritmo. Se construyen C_{min} clústeres de manera aleatoria y se calculan los grados de pertenencia para los puntos iniciales.
2. Por cada dato que llega se calcula su pertenencia a todos los clústeres. Si la pertenencia máxima es mayor que el umbral μ , entonces los clústeres y pertenencias son actualizadas utilizando el algoritmo de fuzzy c-means. En caso contrario (cuando la máxima pertenencia es menor a μ), se procede al paso 3.
3. Dado que el número actual de clústeres es C , el valor del índice de validación para C se compara con los índices de validación para $L = C - 2$ a $L = C + 2$ clústeres. Los clústeres para cada uno de estos valores son construidos utilizando valores guardados con anterioridad. Si ya se habían construido L clústeres entonces los valores de clústeres son actualizados utilizando el algoritmo de fuzzy c-means para luego evaluar el índice de validación. Si no se han construido L clústeres con anterioridad pero sí $L - 1$, entonces se agrega un nuevo clúster y se actualizan los valores con el algoritmo estándar. Finalmente se selecciona el número de clústeres con el mejor valor para el índice de validación.

El algoritmo permite construir los clústeres de manera incremental y decidir automáticamente el valor para C (aunque sí se debe definir un mínimo y máximo).

En [18] ambos algoritmos son aplicados al trabajo de identificación de clústeres en un calorímetro de muestreo compuesto de tungsteno y silicio (este tipo de calorímetros son utilizados en el experimento CALICE [20]). Los algoritmos son usados en la reconstrucción de fotones que decaen de piones neutros. El algoritmo de fuzzy c-means no funciona bien en caso de observar más de dos clústeres debido a que subvalora el número de clústeres (principalmente cuando estos clústeres se presentan en forma de grupos), en cambio el algoritmo de fuzzy c-means dinámico es capaz de inferir el número correcto de clústeres siendo más robusto a estos casos.

El algoritmo dinámico muestra en general buena performance en el trabajo de reconstrucción de clústeres para fotones, sin embargo es considerablemente más lento que los algoritmos de ventana deslizante y clústering topológico. Otro problema de este algoritmo es que al no considerar la energía de cada celda, se está perdiendo información que puede ser importante para el proceso de separación de lluvias solapadas.

4.2. Identificación de máximos

Luego de obtener los clústeres es necesario estimar los puntos de incidencia de las partículas y separar lluvias solapadas. Para esto es necesario conocer los máximos locales de los clústeres. Si observamos más de un máximo local en un sólo clúster podría indicar el solapamiento de lluvias. Si bien, existen algoritmos relativamente simples para encontrar los máximos locales, estos métodos no consideran ni el ruido presente en los valores ni las fluctuaciones estadísticas. Es posible que muchos de los máximos locales sean producidos por ruido, esos casos deben ser identificados y eliminados con el fin de evitar la identificación de partículas falsas.

Un algoritmo simple para la identificación de máximos locales es presentado en [8]. En este, se definen dos umbrales: $cut2$ y $cut3$. Un máximo local es identificado cuando el valor de energía de la celda es mayor que $cut2$ y además mayor que todas las celdas inmediatamente vecinas por un valor $cut3$. Los valores para ambos umbrales son definidos con dependencia de las características del detector, sin embargo, los autores mencionan que la búsqueda de máximos locales no es muy sensible a estos valores. Por lo anterior definen los umbrales fijos como 100 MeV y 60 MeV. El algoritmo presenta varios problemas, por ejemplo, consideremos el caso en que dos celdas vecinas tengan energías altas muy parecidas, en tales casos la definición anterior no identificaría un máximo local. Si bien, este caso puede ser poco probable cuando los cristales tienen un área relativamente grande, debido a que la mayoría de la energía se deposita en el cristal de incidencia, es un problema muy común al utilizar cristales de menor diámetro (como en el preshower). En este caso, muchas veces la energía se reparte entre dos o más cristales. Otro problema es la definición de los valores para los umbrales, en un detector con cristales amplios se puede asumir cierta estabilidad en la probabilidad de interacción, debido a que la longitud del cristal supera por mucho la longitud de radiación

del material. Sin embargo, para cristales de menor tamaño la probabilidad de interacción es mucho menor y los casos en que las partículas sólo depositan una pequeña parte de su energía, o nada, son más comunes. Esto hace que la definición de estos umbrales sea muy complicada y dependiente de cada evento.

4.2.1. Algoritmo de búsqueda de peaks

Un algoritmo más elaborado para la identificación de máximos, que también es más robusto respecto al ruido y no necesita la definición a priori de umbrales (aunque si necesita de otros parámetros) es presentado en [21]. Este algoritmo asume que los peaks pueden ser descritos por funciones normales y el background puede ser aproximado por funciones lineales por intervalos. En estos casos se puede aproximar el valor en la celda x como

$$N(x) = G(x) + B + Cx,$$

donde $G(x) \sim \mathcal{N}(x, \sigma^2)$ y B y C son constantes. El método usa la segunda derivada de $N(x)$ para encontrar los peak, esto debido a que para la segunda derivada $N''(x)$ el background desaparece y sólo se observan valores $N''(x) \neq 0$ en donde se presenta un peak.

Para cada valor de x considerar los valores discretos de conteo N_i . Estos se definen como

$$N_i = A \exp[-(i - i_0)^2 / (2\sigma^2)] + B + Cx. \quad (4.2.1)$$

Donde la segunda derivada se reemplaza por la segunda diferencia

$$S_i = N_{i+1} - 2N_i + N_{i-1}. \quad (4.2.2)$$

Un problema es que los valores de N_i fluctúan alrededor de su valor esperado según una desviación estándar F_i . Entonces si S_{i_0} es comparable a la desviación estándar alrededor del peak, no se puede considerar como diferente a cero, produciendo un peak falso. La desviación estándar de S_i puede ser calculada como

$$F_i = (N_{i+1} + 4N_i + N_{i-1})^{1/2}.$$

Una forma de reducir la desviación estándar de S_i es suavizando la función promediando los valores vecinos, es decir

$$S_i(w) = \sum_{j=i-m}^{i+m} S_j,$$

con $w = 2m + 1$. Considerando el caso especial en que $B = C = 0$ y definiendo A_{min} como la intensidad del peak en la eq. 4.2.1 para la cual $S_{i_0} = F_{i_0}$, entonces $S_{i_0} \gg F_{i_0}$ sólo si $A \gg A_{min}$. Por lo que se desea minimizar A_{min} . Luego de la primera iteración de suavización es posible volver a suavizar la función suavizada, y continuar así, por lo que se puede definir la función suavizada generalizada como

$$S_i(z, w) = \underbrace{\sum_{j=i-m}^{i+m} \dots \sum_{h=l-m}^{l+m} S_h}_z, \quad (4.2.3)$$

donde se desea encontrar los parámetros z y w dado que se minimice el valor A_{min} . En [21] se demuestra que los valores que minimizan esta cantidad son

$$z = 4, \quad w = 0.6\Gamma,$$

donde Γ es el valor de fwhm (*full width at half maximum*) del peak, que es equivalente a

$$\Gamma = 2.355\sigma.$$

Estos valores son definidos analíticamente de manera de maximizar la capacidad del algoritmo para identificar peak individuales y peaks dobles. Un peak doble es un caso en que dos peak son observados solapados con una distancia entre máximos de $S/2$.

Un peak es encontrado en el caso de observar valores $S_i > fF_i$ donde f es un factor de confianza. En caso de que el background no sea lineal por secciones, entonces se pueden considerar condiciones extras para definir un peak, como se explica en [21].

Una versión actualizada del algoritmo para casos multidimensionales es presentado en [22]. El algoritmo sigue la misma idea buscando peaks en el espacio m-dimensional.

4.3. Separación de lluvias solapadas

Dos peaks en un clúster probablemente indican el solapamiento de dos o más lluvias electromagnéticas, en estos casos es necesario separar las celdas que corresponden a cada una de las lluvias. Además, los valores en las celdas solapadas deben ser repartidos entre las lluvias.

Esto se realiza utilizando un algoritmo para la separación de partículas solapadas (o algoritmo de *unfolding*). Este algoritmo es capaz de distribuir la energía entre las partículas de manera de obtener dos o más clústeres de energías separados que luego serán utilizados en el calculo de posición o energía. Comúnmente se utiliza la distribución de energía lateral de la lluvia de fotones en los cristales del detector, estos valores están relacionados con el radio de Moliere explicado en la Sección 2.1.1. En [23] se propone un algoritmo de unfolding que permite separar el conteo de cada celda entre las lluvias correspondientes a cada partícula.

Para separar las lluvias el algoritmo hace uso del perfil lateral del detector. Este define en que forma la energía depositada en las celdas se distribuye a medida que la distancia entre la celda y la posición incidente de la partícula crece. Una buena forma de obtener este perfil es simular la geometría del detector y obtener mediante simulaciones datos que permitan ajustar una función que defina el comportamiento lateral de la lluvia. La función ajustada depende de la energía de la lluvia electromagnética (en GeV) y la distancia de la celda a la posición incidente. La función entrega la cantidad de energía esperada en una celda a distancia Δr de la posición incidente. En [23] se propone la función

$$\begin{aligned} f(\Delta r, E) &= A \text{máx}(\exp(-\Delta r^2/0.33), d \exp(\Delta r/s)), \\ s &= 0.254 + 0.013E^{0.7}, \\ d &= 1.67 - 0.374 \ln(E), \end{aligned}$$

donde A es un valor de normalización.

La curva parametrizada y la función obtenida se usa en el calculo de energía para cada celda utilizando la ecuación 4.3.1, donde A_i es la energía depositada en cada celda, N es el número de máximos, k representa cada clúster a separar y va entre 1 y N , r_{ij} es la distancia entre la celda i y la posición del máximo j , E_j es la energía del máximo j y f es la función

obtenida por la parametrización.

$$A_i^k = A_i \cdot \frac{f(r_{ik}, E_k)}{\sum_{j=1}^N f(r_{ij}, E_j)}. \quad (4.3.1)$$

Con estos datos el algoritmo se resume en los siguientes pasos:

1. El algoritmo comienza con una estimación de las energías y posiciones de las lluvias correspondientes a cada máximo, para esto se utilizan las celdas correspondientes a los máximos locales en el clúster. Es decir, la posición de la lluvia es dada por el centro de estas celdas y sus energías por la energía depositada en estas celdas.
2. Luego, la energía en cada celda es dividida en $A_i^1, A_i^2, \dots, A_i^N$ las energías correspondientes a cada uno de los máximos, utilizando la ecuación 4.3.1. Esto permite calcular nuevas energías y posiciones para cada una de las lluvias correspondientes a los máximos locales. Estos nuevos valores son utilizadas como entrada para una nueva iteración del algoritmo.
3. La iteración continua sólo si al menos alguna de las coordenadas de cualquier lluvia cambia en más de 0.01 celdas en cualquier eje y termina si se alcanzan 20 iteraciones, en cuyo caso se considera que el algoritmo no converge. Comúnmente no se necesitan más de 4-5 iteraciones [8].

En Algoritmo 1 se describe el algoritmo de unfolding completo.

Algoritmo 1 Algoritmo de unfolding

- 1: Se tiene como entrada un arreglo con las posiciones y conteos de las celdas en el clúster y un arreglo con las posiciones del o los máximos en el clúster.
 - 2: **if** número de máximos > 1 **then**
 - 3: Inicia los valores de cada lluvia como la posición y energía de las celdas con los máximos.
 - 4: **while** Número iteraciones < 20 AND ΔX (cambio de posición de la lluvia) > 0.01 **do**
 - 5: Calcular nuevos valores de energía para cada celda.
 - 6: Calcular la nueva posición y energía de cada una de las lluvias.
-

4.4. Rechazo de partículas solapadas

Puede ocurrir el caso que dos o más partículas den origen a un clúster de sólo un máximo. Esto ocurre cuando las partículas vienen muy cercanas entre sí. Considerar erróneamente estos casos como una sola partícula producirá reconstrucciones falsas de partículas, afectando el resultado final del algoritmo. Si bien, es imposible reconstruir la posición de las partículas en este caso, si es factible rechazar estos casos para evitar falsos positivos. El algoritmo de rechazo utiliza información de la forma de la lluvia para clasificar entre lluvias producidas por una o por más de una partícula. Intuitivamente se podría esperar que las lluvias producidas por dos o más partículas sean más anchas, o tengan algún tipo de asimetría.

Un método común para rechazar partículas solapadas hace uso de cortes simples sobre variables de sumario extraídas de los valores de las celdas en el clúster [23, 8]. En [23] los autores proponen usar un corte en el segundo momento centrado de la lluvia, o dispersión, definido como:

$$D_x = \frac{\sum E_i x_i^2}{\sum E_i} - \left(\frac{\sum E_i x_i}{\sum E_i} \right)^2. \quad (4.4.1)$$

O equivalentemente D_y para el eje y . La dispersión permite medir el ancho de la lluvia, que es útil para identificar dos o más partículas. Los autores observan que la eficiencia del corte depende de la posición incidente de la partícula respecto al centro de la celda. Debido a esto, un corte parabólico en D_x es propuesto. Para esto se define el valor

$$D_x^{corr} = D_x - D_x^{min}, \quad (4.4.2)$$

$$D_x^{min} = (\bar{x} - x_R)(\bar{x} - x_L), \quad (4.4.3)$$

donde \bar{x} es el primer momento del clúster y x_R y x_L son los ejes derecho e izquierdo de la celda central. Los autores muestran que un corte lineal en la distribución uni-dimensional de D_x^{corr} es equivalente a un corte parabólico en la distribución bi-dimensional de D_x versus \bar{x} . Usando esto, es posible obtener cortes que no dependen de la posición de incidencia de la

partícula. El corte final en el plano X-Y se define en base a los valores

$$D = \text{máx}(D_x^{corr}, D_y^{corr}), \quad (4.4.4)$$

$$D^{min} = \text{máx}(D_x^{min}, D_y^{min}). \quad (4.4.5)$$

Se observa que este corte es dependiente de la energía de la partícula incidente (esto es debido a que la forma de la lluvia depende de la energía de manera logarítmica).

En [8] los autores presentan un método alternativo basado en un análisis mas detallado de la forma de la lluvia. Para esto se mide el perfil de la lluvia mediante el calculo de (λ_1, λ_2) , que son los valores propios del tensor 2×2 de esfericidad, definido como

$$S_{xx} = \sum_i E_i x_i^2, \quad (4.4.6)$$

$$S_{xy} = S_{yx} = \sum_i E_i x_i y_i, \quad (4.4.7)$$

$$S_{yy} = \sum_i E_i y_i^2. \quad (4.4.8)$$

El corte se realiza en el factor de esfericidad L :

$$L = \frac{|\lambda_1 - \lambda_2|}{\lambda_1 + \lambda_2}. \quad (4.4.9)$$

Luego de definir las variables sobre las que se realizará el corte, es necesario seleccionar el valor del corte. Para esto se suelen utilizar simulaciones. Las simulaciones permiten saber exactamente el tipo de partícula incidente, por lo que es posible definir el corte óptimo en términos de eficiencia en la detección de señal y en el rechazo de background. Una manera simple de seleccionar este valor es mediante inspección gráfica, decidiendo, cuando es posible, el valor del corte observando gráficos de la variable de selección para datos correspondientes a la señal y al background. Es deseable, sin embargo, definir un método automatizado para seleccionar el valor del corte. En [24] se propone un método de selección automatizada de cortes. El método define como criterio de optimización la maximización del rechazo de background, r_B , a dada eficiencia de señal, e_S , y esto se hace para todo el rango de valores de eficiencia. Para seleccionar el valor del corte (o valores, en caso de usar más de una variable para realizar el corte) se utiliza un método de muestreo de Monte Carlo. Cada valor

muestreado corresponde a un punto en el espacio de (r_B, e_s) . Para facilitar el trabajo, el espacio e_s es dividido en intervalos y para cada intervalo se retiene la muestra que maximice r_B . Otros métodos más avanzados de optimización, como algoritmos genéticos, pueden ser utilizados en la optimización del valor del corte.

Los métodos de rechazo basados en cortes son muy populares en física de partículas debido a su simpleza en la implementación y su transparencia a la hora de entender el proceso de decisión. Sin embargo, el uso de algoritmos de clasificación (o métodos de selección multivariados, como se conocen en física de partículas) permite construir reglas de decisión más complejas, a costo de perder transparencia en el proceso de decisión. En [25, 26] se propone el uso de clasificadores para el rechazo de partículas en calorímetros (aunque cabe notar que las geometrías de los calorímetros presentados en ambos trabajos difieren considerablemente de la del detector presentado en este trabajo, por ejemplo, ambos tienen segmentación longitudinal). En ambos trabajos se comienza extrayendo variables que representen la forma de la lluvia, como por ejemplo: El número de celdas en las que se detectó energía, la profundidad máxima de la lluvia en el eje Z, el ancho de la lluvia, la fracción de energía depositada en alguna de las capas longitudinales, entre otras variables. Luego, las variables (o características) extraídas son utilizadas como entrada a un algoritmo de clasificación que se encargará de discriminar entre señal o background. En [25] se propone el uso de *boosted decision trees*, mientras que en [26] se hace uso de *support vector machines* y *multilayer perceptron*.

4.5. Reconstrucción de posición

La reconstrucción de posición es un método mediante el cual la posición de la partícula incidente es estimada utilizando la energía depositada y la posición de las celdas del clúster. Un método simple propuesto por [27] es usar el centro de gravedad como se muestra en la Eq. 4.5.1.

$$X_{cg} = \frac{\sum_i w_i x_i}{\sum_i w_i}. \quad (4.5.1)$$

Donde x_i representa el centro de la celda i y w_i corresponde al peso asignado a la celda, comúnmente la energía, pulso o conteo correspondiente a la celda. El denominador corresponde, entonces, a la energía total del clúster. Como se menciona en [28] este método

tiende a presentar un sesgo sistemático hacia la posición del centro de una celda, lo que genera la conocida correlación de forma S entre la posición reconstruida y la posición incidente.

La ecuación lineal 4.5.1, además, no considera el decaimiento exponencial de las lluvias. Para resolver esto, una posible solución es corregir los errores sistemáticos en los valores ajustando la siguiente función

$$x'_{cg} = x_{cg} + b \operatorname{arcsinh} \left(\frac{X_{cg}}{\Delta} \sinh \delta \right),$$

donde el parámetro b se relaciona con la forma exponencial de la lluvia, Δ es la la mitad del ancho de la celda y $\delta = \Delta/b$.

En [28] se propone una corrección del sesgo basada en la distribución de la respuesta lateral de los cristales, mediante un método iterativo que reajusta los puntos basándose en cuan parecida es la energía obtenida mediante esta función con la energía observada. Mientras que en [29], la corrección necesaria es ajustada mediante el método de mínimos cuadrados a los valores observados.

Otro método, más simple y efectivo y que también usa el centro de gravedad de la ecuación 4.5.1 pero con diferentes pesos es presentado en [30]. Este método calcula los pesos como:

$$w_i = \max \left\{ 0, \left[w_0 + \ln \left(\frac{E_i}{E_t} \right) \right] \right\}, \quad (4.5.2)$$

$$E_t = \sum_i E_i.$$

Donde w_0 es un parámetro adimensional que se debe ajustar y depende del tamaño y tipo del cristal. Este impone un umbral a la menor energía depositada en una celda que es aceptada para la reconstrucción. Valores mayores de w_0 llevan a pesos más uniformes mientras que menores valores favorecen celdas con bajas energías depositadas. El método funciona muy bien y el logaritmo se encarga automáticamente del decaimiento exponencial en el perfil lateral de la lluvia [30]. Si bien, el parámetro w_0 es dependiente de la energía incidente de la partícula, esta dependencia es pequeña y un valor fijo de w_0 se puede usar con una pequeña pérdida en la resolución de posición [30]. Este algoritmo en específico ha sido usado extendidamente en la literatura [31, 32, 33].

Capítulo 5

Algoritmo de Reconstrucción para el Detector Preshower

En esta sección se presentará el algoritmo de reconstrucción para el detector preshower. La entrada para el algoritmo de reconstrucción corresponde a la salida del sistema de lectura del detector preshower. Esta salida se conforma por el conteo de fotoelectrones producidos por cada uno de los MPPC. Los MPPC, a su vez, están dispuestos en dos bordes perpendiculares de la matriz de cristales, por lo que la salida corresponde a dos vectores uni-dimensionales de tamaño 25 (que llamaremos eje X y eje Y). El conteo de fotoelectrones es proporcional a la energía depositada por las partículas en el detector.

El problema es que para obtener información útil para el análisis del experimento necesitamos identificar las partículas, medir su posición y obtener su energía, sin embargo, sólo poseemos como salida del detector dos histogramas de conteo de fotoelectrones. Es por esto que se necesita un algoritmo de reconstrucción que se debe encargar de analizar los datos brutos obtenidos por el detector y obtener información útil para el posterior análisis.

5.1. Descripción General

Como ya se mencionó en la Sección 4, un algoritmo de reconstrucción debe realizar varias tareas, como la agrupación de los datos, la identificación de máximos en los clústeres, la identificación y separación de partículas solapadas, el rechazo de partículas solapadas con un sólo máximo y la reconstrucción de la posición y energía de las partículas. Una característica

que debe ser considerada específicamente para el detector preshower es la siguiente:

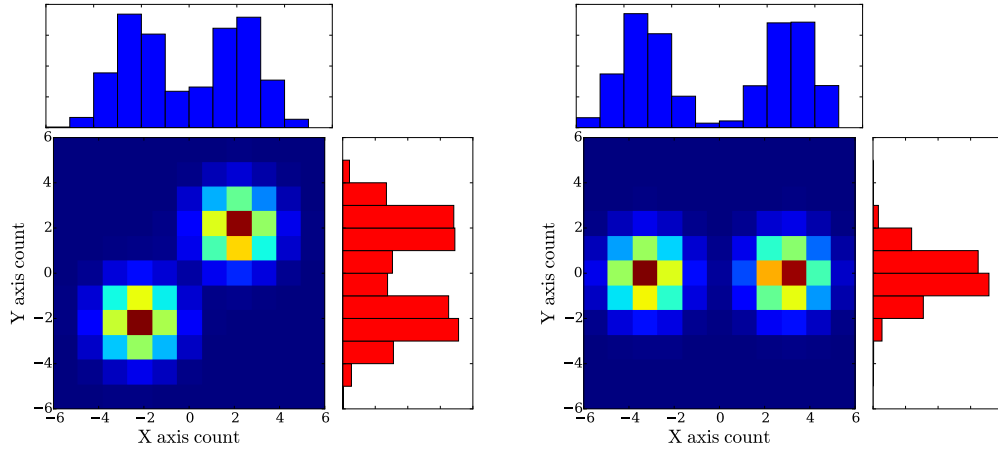
Es común que los detectores sean capaces de medir valores de energía (o un valor proporcional) de manera bi-dimensional, es decir, obtener una posición (X,Y) de la partícula incidente, como por ejemplo en [8]. También un detector puede medir valores de manera tridimensional, obteniendo una posición (X,Y,Z) donde Z corresponde a la posición longitudinal, esto, en casos en que el detector este también segmentado longitudinalmente [31]. Sin embargo, una particularidad del detector preshower es que sólo es capaz de detectar datos uni-dimensionales de valores acumulados en los ejes transversales X e Y . Dado que el sistema de lectura sólo entrega la información de la cantidad acumulada de fotoelectrones en cada uno de los ejes transversales, la medición de la energía depositada en cada cristal no esta disponible, lo que genera ciertas ambigüedades a la hora de reconstruir dos o más partículas. Considerar por ejemplo que dos clústeres son observados en el eje X y en el eje Y . Pensar que se reconstruye la posición de estos clúster con un método cualquiera, obteniéndose posiciones (x_1, x_2) en el eje X y (y_1, y_2) en el eje Y . Entonces, tenemos 2 distintas formas de reconstruir la posición en el plano, $\{(x_1, y_1), (x_2, y_2)\}$ o $\{(x_1, y_2), (x_2, y_1)\}$. El problema se complica mucho más cuando se observan más de dos clústeres en los ejes, debido a que pueden ocurrir combinaciones como $\{(x_1, y_2), (x_2, y_2)\}$ en casos en que ambas partículas se observen como un sólo clúster en uno de los ejes, debido al solapamiento de las lluvias. En la Figura 5.1 se observan ambos casos mencionados. Entonces, se debe agregar un paso encargado de la combinación de clústeres entre ambos ejes y la resolución de ambigüedades que pudiesen aparecer al combinar los clústeres.

5.2. Propuesta y experimentos

En esta sección se presentan cada una de las partes del algoritmo de reconstrucción propuesto para el detector preshower. Además, se usan simulaciones para analizar cada uno de los procesos y seleccionar los parámetros necesarios para el funcionamiento del algoritmo.

5.2.1. Simulaciones

Para estudiar el algoritmo de reconstrucción propuesto se utilizan simulaciones. El detector ha sido simulado utilizando GEANT 4 [34]. El prototipo simulado corresponde a un

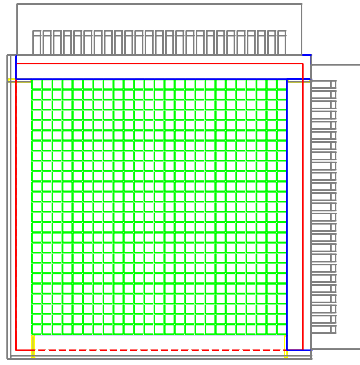


(a) Existe un problema de ambigüedad en la reconstrucción de posiciones. Dos opciones son igual de factibles. (b) Dos partículas se observan como una en uno de los ejes debido al solapamiento de las lluvias.

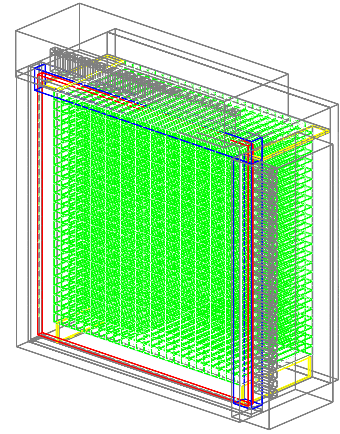
Figura 5.1: Diagramas que muestran algunos de los problemas que presenta la particular disposición del sistema de lectura del detector preshower. Debido a que sólo se conocen los valores acumulados de conteo en los ejes X e Y, la reconstrucción de la posición en el plano presenta algunas ambigüedades que se deben solucionar.

arreglo de 25x25 cristales LYSO. Cada cristal tiene un tamaño de 4x4x45 [mm³]. Los cristales están cubiertos por una pintura reflectora en la cara trasera para evitar fuga de fotones por la cara no sensible. En la cara frontal de la matriz de cristales una placa de acrílico de ancho igual a 1.5 [mm] es posicionada. Esta placa está cubierta por un material reflector en su cara frontal para evitar la fuga de fotones. En la placa se posicionan 25 fibras horizontales y 25 fibras verticales de 1 [mm] de diámetro que pasan por el centro de cada cristal. Estas fibras también están cubiertas por materiales reflectantes en las esquinas no sensitivas para evitar la fuga de fotones. Los MPPC son reemplazados por materiales sensibles que permiten medir los fotones que llegan al sistema de lectura. Toda la geometría está cubierta por placas de aluminio de 2.5 [mm] de espesor con uniones de cobre. Una simulación de la geometría completa del detector es mostrada en la Figura 5.2.

En los experimentos la geometría ha sido simplificada con la finalidad de acelerar el lento proceso de simulación en los experimentos para la reconstrucción. Muchos de los elementos exteriores del detector (como la cubierta protectora) no fueron considerados. Estos elementos se utilizan principalmente para evitar la influencia de factores externos que no están presentes en las simulaciones. Además, la matriz de fibras para la lectura y el sistema de lectura es



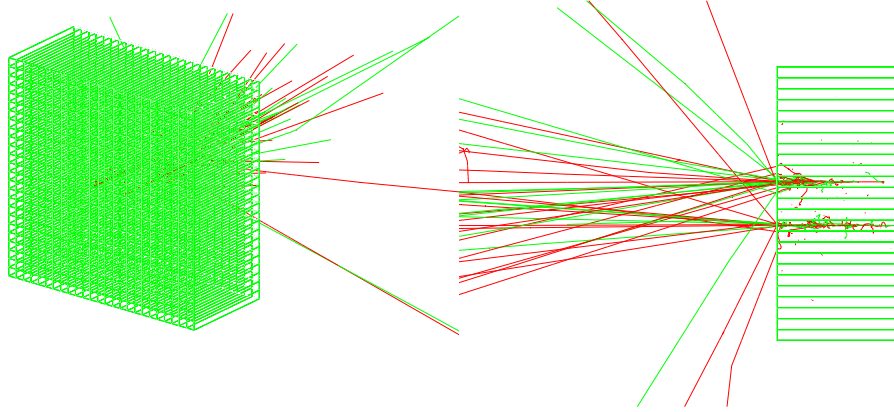
(a) Cara frontal.



(b) Perspectiva rotada.

Figura 5.2: Geometría del detector preshower completa simulada. Mostrado en verde se observa la matriz de cristales, en plomo el material cobertor de la matriz y el sistema de lectura, en rojo se muestra la placa de acrílico en la que se montan las fibras de luz (no mostradas para facilitar la visualización de los componentes).

removido para evitar el lento trabajo de simulación de los procesos en las fibras. En su lugar la respuesta obtenida por el panel de lectura es aproximada usando un modelo estadístico basado en mediciones reales que permite obtener la cantidad aproximada de fotones que llegan a cada MPPC en base a la energía medida en cada cristal. El modelo fue propuesto por el equipo de físicos involucrados en este trabajo. Para obtener el modelo estadístico utilizado para simular la salida de cada MPPC se asume que la cantidad de fotoelectrones que llegan a cada MPPC es una función de Poisson de la energía total que llega al MPPC. La energía que llega al MPPC desde cada cristal se modela como una distribución normal centrada en la posición del cristal en el eje correspondiente al MPPC normalizada según el valor de energía depositada en el cristal. En términos matemáticos, los valores p_{xl} y p_{yk} corresponden a la salida de la fibra l y k en el eje x e y respectivamente. Estos valores se calculan en base a las energías E_{ij} para cada cristal, donde (i, j) corresponde a la posición



(a) Perspectiva rotada.

(b) Cara superior.

Figura 5.3: Geometría simplificada del detector preshower, donde sólo se simula la matriz de cristales del detector. En la imagen se muestra la matriz en verde y la interacción de dos partículas γ incidentes de 5 [GeV]. Las trayectorias de los electrones son mostradas en rojo mientras que las trayectorias de positrones son mostradas en verde.

del cristal. Primero se calculan los valores u_{ij}^l, u_{ij}^k como:

$$\begin{aligned}
 u_{ij}^l &= 20 \cdot E_{ij} \cdot \left[\frac{1}{2\sqrt{2} \cdot \pi \cdot 4} \exp\left(\frac{-(l-i)^2}{2 \cdot 4}\right) \right], \\
 u_{ij}^k &= 15 \cdot E_{ij} \cdot \left[\frac{1}{2\sqrt{2} \cdot \pi \cdot 6.25} \exp\left(\frac{-(k-j)^2}{2 \cdot 6.25}\right) \right],
 \end{aligned} \tag{5.2.1}$$

entonces, p_x^l y p_y^k las salidas de las fibras l y k en los eje x e y son calculados como

$$\begin{aligned}
 p_x^l &= Poiss\left(\sum_{ij} u_{ij}^l\right), \\
 p_y^k &= Poiss\left(\sum_{ij} u_{ij}^k\right).
 \end{aligned}$$

Estas salidas corresponden al conteo de fotoelectrones detectado por el MPPC l o k . En la Figura 5.3 se muestra la versión simplificada del detector en conjunto a la interacción de dos partículas γ incidentes.

Los experimentos fueron realizados en el rango 0.5-15 [GeV]. Se consideró un valor de

ruido de 2.25 [MeV], este valor fue seleccionado en base a la eficacia del algoritmo de clústering para identificar clústeres correctamente en el rango de energía mencionado anteriormente (este valor es mucho menor que los valores utilizados en la literatura [8], debido al tamaño considerablemente menor de los cristales para el detector preshower). Los experimentos fueron realizados utilizando partículas γ , lanzadas distribuidas de manera uniforme en el cuadrado central de 80x80 [cm] de la matriz de cristales. Esto último para evitar casos especiales en los cuales las partículas inciden muy cerca de los límites. En estos casos cierta energía es perdida provocando algunos problemas en la reconstrucción. Aunque la simulación ha sido simplificada para acelerar el proceso, la simulación de cada evento es aún una tarea relativamente lenta. Los experimentos presentados en esta sección fueron realizados con muestras de 1000 a 5000 eventos (dependiendo del experimento), lo que observamos es un buen compromiso entre tamaño de muestra y tiempo de ejecución.

5.2.2. Algoritmo de reconstrucción

En esta sección se presenta el algoritmo de reconstrucción propuesto para el detector preshower y se estudian sus propiedades mediante simulaciones.

El algoritmo completo de reconstrucción consiste de los siguientes procesos

1. Construcción de clústeres.
2. Identificación de máximos en los clústeres.
3. Separación de lluvias solapadas.
4. Rechazo de partículas solapadas con un sólo máximo.
5. Reconstrucción de posición y energía.
6. Resolución de ambigüedades.

En la Figura 5.4 se presenta un diagrama de flujo que resume el proceso completo de reconstrucción.

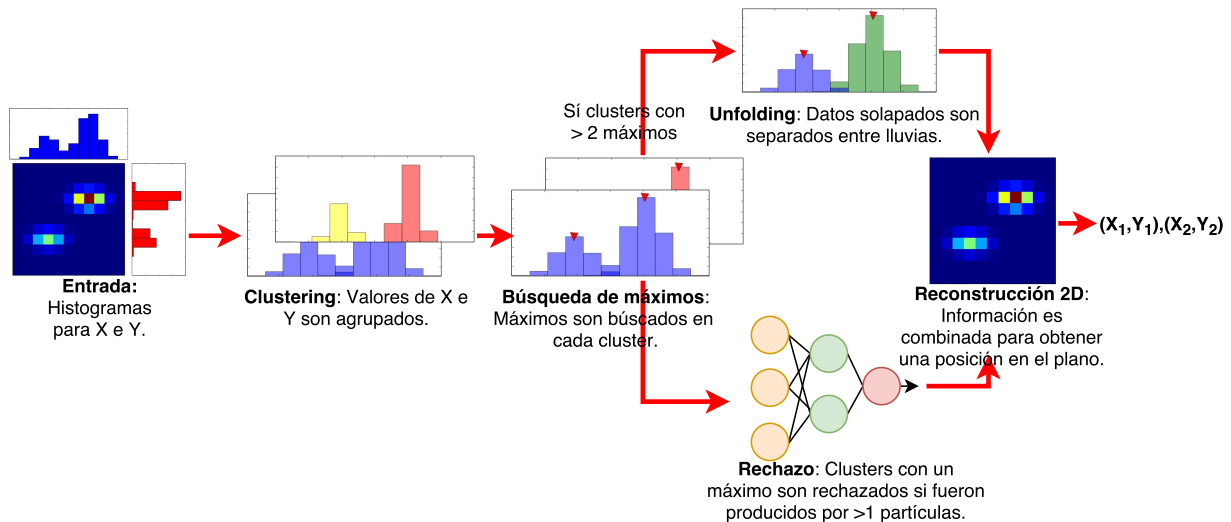


Figura 5.4: Diagrama del proceso de reconstrucción completo.

Algoritmo de clústering

El algoritmo de clústering es el primer paso del algoritmo de reconstrucción. Para el detector preshower se decidió utilizar una adaptación del algoritmo topológico presentado en la Sección 4.1.2. El algoritmo de clústering topológico es un algoritmo muy utilizado en la práctica y permite la construcción de clústeres de tamaños variables, a diferencia del algoritmo de ventana deslizante. Esto último es importante para el detector preshower, debido a que la profundidad menor de los cristales provoca una gran variabilidad en la cantidad de energía depositada en cada evento (esto a causa de que muchas de las partículas interactúan en la sección final del cristal, depositando sólo una parte de su energía). Otras opciones, como el algoritmo de clústering fuzzy, fueron rechazadas debido a la poca o nula utilización en experimentos reales de estos algoritmos y debido a que no utilizan toda la información disponible en el preshower. En el caso del algoritmo de clústering fuzzy, no se hace uso de la información disponible en la distribución lateral de la matriz de cristales.

El algoritmo de clústering topológico implementado para el detector preshower se resume en el Algoritmo 2.

Los valores $cut1$ y $cut2$ corresponden al umbral de ruido y a un umbral útil para identificar máximos respectivamente. Estos valores fueron elegidos observando la eficacia del algoritmo de clústering para identificar partículas simuladas en el umbral 0.5-15 [GeV]. Los valores usados son $cut1 = 2.25$ y $cut2 = 3.75$.

Algoritmo 2 Algoritmo de clústering

```
1: Como entrada se tiene dos arreglos de valores para cada celda y para cada eje (X e Y).
2: procedure ADYACENTES(celda, clúster)
3:   for Cada vecino de la celda do
4:     Añadir vecino a clúster
5:     Marcar celda de vecino como usado
6:     Llamar Adyacentes(vecino, clúster)
7: Inicializar arreglo clústeres vacío
8: for Cada eje do
9:   Inicializar clústeres[eje]
10:  El arreglo de conteos es llenado con los valores mayores a cut1 en el eje
    correspondiente.
11:  El arreglo de conteos es ordenado de mayor a menor y todos los valores son
    inicializados como no usados.
12:  for Cada valor en el arreglo de conteos do
13:    if Valor es mayor a cut2 y el valor no esta marcado como usado then
14:      Crear nuevo clúster
15:      Añadir posición como semilla de un clúster, marcar celda como usada.
16:      Llamar Adyacentes(semilla,clúster)
17:      Añadir clúster a clústeres[eje]
18: Retornar clústeres
```

Notar que el algoritmo no es exactamente igual al algoritmo topológico presentado en la Sección 4.1.2. De este se ha quitado, primero, la recombinación de clústeres. Esto debido a que el algoritmo de unfolding se encarga de la separación de lluvias adyacentes y solapadas. Segundo, se ha quitado el umbral utilizado en la decisión de utilizar el vecino como semilla o no, en cambio todos los vecinos son utilizados como semilla. Esto último debido a que las colas de la lluvia son muy importantes para lograr una buena reconstrucción en el detector preshower. Teniendo los clústeres identificados, el algoritmo continúa con la búsqueda de máximos locales en cada uno de los clústeres por separado.

Búsqueda de máximos

El siguiente paso del algoritmo es la búsqueda de máximos en cada uno de los clústeres encontrados. El análisis posterior depende de la cantidad de máximos encontrados en cada clúster. En caso que se encuentre sólo un máximo, entonces se considera que el clúster ha sido originado por una sola lluvia o puede ser originado por más de una lluvia pero imposibles de separar. Si el clúster posee dos o más máximos, entonces se considera una superposición de lluvias electromagnéticas y el algoritmo de unfolding es aplicado con la finalidad de separar las lluvias.

Es posible identificar los máximos asignando reglas simples que funcionen en la mayoría de los casos. En [8] se revisan los valores mayores a un umbral y si el valor de esta celda es además mayor a todos sus vecinos por una cantidad definida con anterioridad, entonces la celda es identificada como máximo local. Sin embargo, para el detector preshower el problema es considerablemente más complicado. Debido a la relativamente baja probabilidad de interacción en el detector preshower, la cantidad de energía depositada en cada evento es muy variable (a diferencia de un detector con una profundidad de varias longitudes de radiación, donde la mayoría de las partículas interactúan). Por esto, la definición de umbrales a priori para identificar los máximos es complicada. Además, producto del diámetro menor de los cristales en el preshower, se dan muchos casos como el mostrado en la Figura 5.5 (varios cristales comparten gran parte de la energía, imposibilitando la identificación de máximos locales), en los que sería imposible identificar los máximos utilizando las reglas simples mencionadas con anterioridad.

En base a lo anterior, se ha decidido utilizar el algoritmo presentado en la Sección 4.2.1.

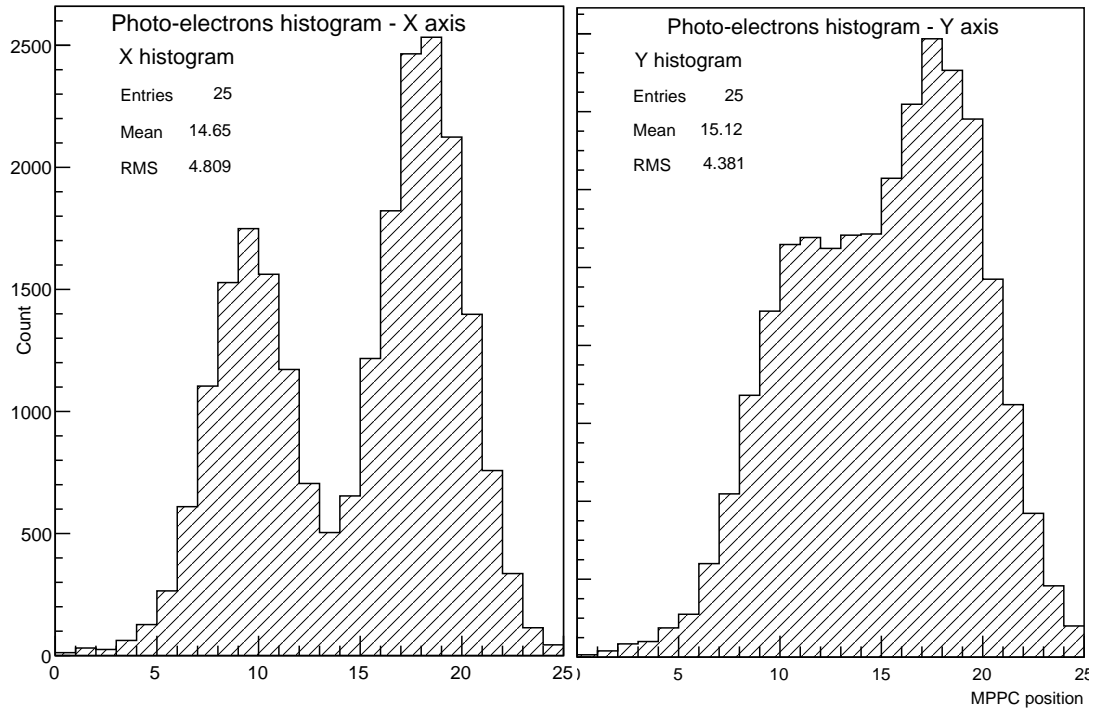


Figura 5.5: Histograma de fotoelectrones para cada uno de los ejes para un evento de dos gamas incidentes de 5 [GeV]. Se observa que la partícula incidente alrededor de la posición 11 del eje Y es muy difícil de identificar como un máximo, ya que no produce un máximo local fácilmente observable.

La versión utilizada del algoritmo es la del paquete ROOT [35], en la clase *TSpectrum*. Si bien el algoritmo se basa principalmente en el algoritmo presentado en [21] y su versión para varias dimensiones presentada en [22], se agregan algunos pasos extra. Primero, se agrega un paso extra de suavizado, basado en utilizar la distribución estacionaria de una cadena de markov, que asigna mayor probabilidades a las celdas en que se encuentran los peaks, como se menciona en [36]. Este método suaviza y al mismo tiempo intensifica los peaks de la distribución original. Además, se agrega un filtro para los peaks encontrados, que sólo retorna los peaks que tienen una amplitud mayor a $umbral \times 100\%$ del mayor peak. El valor de umbral y el σ utilizado en el algoritmo presentado en 4.2.1 son definidos como argumentos de la función, para el detector preshower se definieron valores de $\sigma = 2$ y $umbral = 0.05$ (estos son los valores por defecto, no se encontró mayor diferencia al variar estos valores). En la Figura 5.6 se muestra el algoritmo de búsqueda de peaks funcionando para un caso

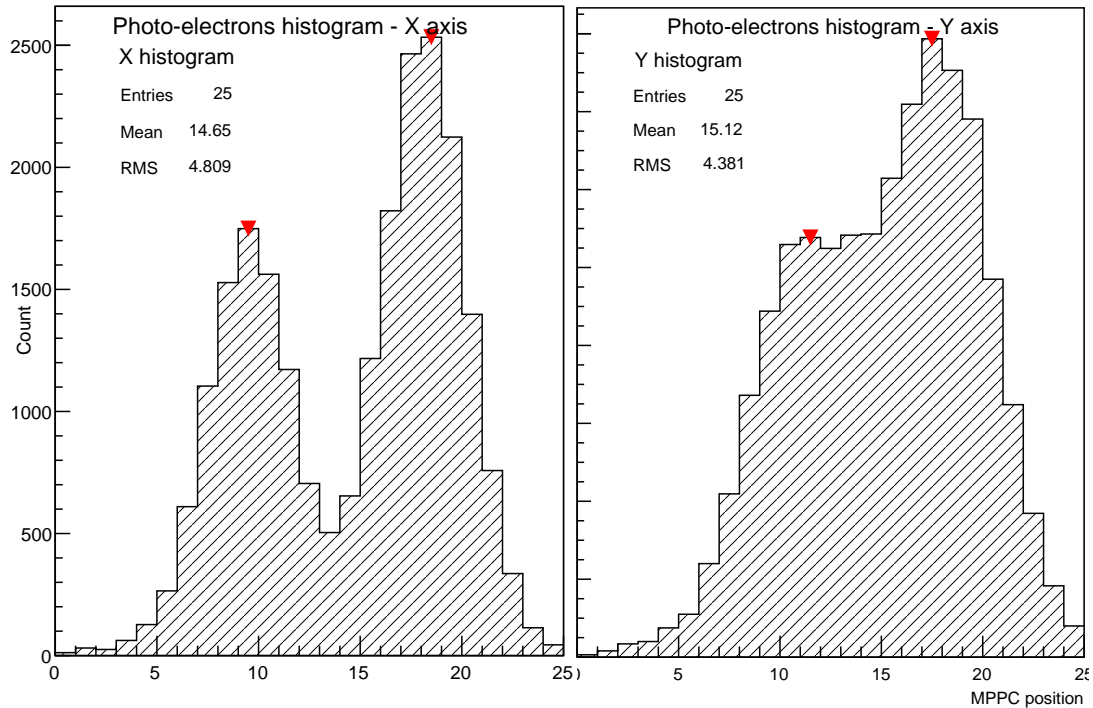


Figura 5.6: Histograma de fotoelectrones para cada uno de los ejes para un evento de dos gamas incidentes de 5 [GeV]. El algoritmo de búsqueda de máximos identifica todas las partículas incidentes en cada eje de manera exitosa, incluso para el caso difícil del eje Y.

difícil del detector preshower.

Algoritmo de unfolding

El método a utilizar para la separación de lluvias solapadas es el presentado en la Sección 4.3. El método se basa en el conocimiento de la distribución lateral de la lluvia en los cristales, por lo que primero esta debe ser calculada y parametrizada para ser usada en la separación de la energía depositada en celdas solapadas. La distribución lateral puede ser calculada mediante simulaciones. Para esto se simularon 2500 eventos, cada uno correspondiente a un gamma de 10 [GeV] lanzado en las dos celdas centrales. Entonces es graficado el conteo de fotoelectrones en una celda versus la distancia entre el centro de la celda y la posición incidente. La función obtenida se muestra en la Figura 5.7. Sobre estos

Lateral Distribution of Lyso crystal

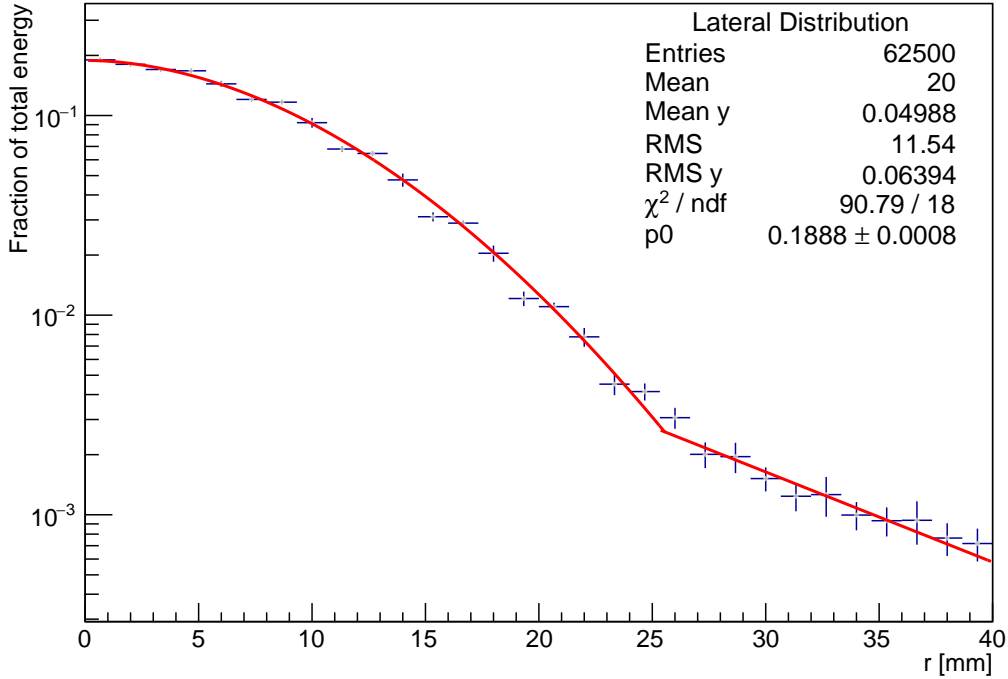


Figura 5.7: Gráfico del conteo de fotoelectrones para cada una de las celdas respecto a su distancia a la posición incidente de la partícula (en uno de los ejes). Los valores son calculados para un gamma de 10 [GeV] en incidencia normal. En rojo se muestra la parametrización de la curva utilizando la ecuación 5.2.2.

datos se ajusta la función 5.2.2 que provee una parametrización sobre los datos.

$$f(r, E) = \begin{cases} A \cdot \exp(-0.09 \cdot r^{1.89}) & r < 26 \\ A' \cdot \exp(-0.1 \cdot r) & r \geq 26 \end{cases}, \quad (5.2.2)$$

donde r es la distancia de la celda a la posición incidente en [cm] y A es una constante de normalización que corresponde al conteo estimado de fotoelectrones de la lluvia. Las simulaciones entregan curvas laterales muy similares para diferentes gamas con diferentes energías, por lo que la dependencia en E de la función esta sólo contenida en el parámetro de normalización. El ajuste se realiza utilizando el paquete *Minuit* del framework *Root* [35]. Los errores en cada uno de los bins en el histograma corresponden al error estándar en la media σ/\sqrt{n} .

El algoritmo de separación de lluvias comienza con una estimación del conteo de

fotoelectrones y de la posición de las lluvias correspondientes a cada máximo. Para esto se utilizan los valores de conteo y posición central de las celdas correspondientes a los máximos locales. Entonces el algoritmo prosigue de acuerdo a lo explicado en la Sección 4.3, adaptado para el caso unidimensional, con la diferencia de que el algoritmo continúa si cualquiera de las posiciones estimada para los clústeres en uno de los ejes cambia en más de 0.01 celdas. En caso que el algoritmo no converja luego de 20 iteraciones el clúster es descartado.

Rechazo de partículas solapadas con un sólo máximo

En la Sección 4.4 se presentaron dos métodos para el rechazo de partículas solapadas con un sólo máximo. Si bien, los métodos basados en cortes son simples y transparentes, los métodos basados en clasificadores permiten modelar reglas de decisión más complejas. Además de esto, un gran problema del método basado en cortes presentado en la Sección 4.4 es que es dependiente de la energía incidente de la partícula. Una forma de solucionar esto es utilizando un algoritmo de clasificación multivariado capaz de variar suavemente la superficie de decisión a medida que la energía observada de la partícula incidente cambia.

Finalmente, se decide utilizar un método de clasificación multivariado con la finalidad de: 1) Proveer cortes no lineales más complejos que permitan mejorar el desempeño del algoritmo respecto a la eficiencia en la señal y el rechazo de background. 2) Obtener un desempeño que sea invariante a la energía de la partícula incidente. Esto último mediante la inclusión de la energía medida de la partícula como característica de entrada. Ideas similares han sido propuestas en [37, 38].

Como entrada para el algoritmo de clasificación se seleccionó un conjunto de características capaces de resumir la forma de la lluvia. Es de esperar, como ya se mencionó, que la forma de la lluvia cambie cuando dos o más partículas sean las que dieron origen a la lluvia. En base a esto, las características utilizadas son:

- El centro de gravedad del clúster, definido como en la ecuación 4.5.1 y con pesos definidos en la ecuación 4.5.2.
- La varianza con pesos del clúster. Definida como

$$\sigma_x^2 = \frac{\sum_i w_i (x_i - \bar{x})^2}{\sum_i w_i}, \quad (5.2.3)$$

y con pesos definidos en la ecuación 4.5.2.

- El coeficiente de asimetría de Fisher del clúster, definido como

$$\gamma_x = \frac{1}{\sigma_x^3} \frac{\sum_i w_i (x_i - \bar{x})^3}{\sum_i w_i}. \quad (5.2.4)$$

- La altura normalizada del clúster, definida como $E_{max} / \sum_i E_i$. Es decir, la energía normalizada de la celda con el máximo.
- La dispersión corregida, definida en la ecuación 4.4.2.

Cada una de las características es calculada para el eje X e Y, obteniendo un total de 14 variables como entrada para el algoritmo de clasificación, además de la energía de la partícula incidente.

Reconstrucción de posición y energía

Como se mencionó en la Sección 4.5, la forma más simple de reconstruir la posición de una partícula incidente es usando el centro de gravedad, expuesto en la ecuación 4.5.1. Donde w_i en la suma con pesos corresponde al conteo de fotoelectrones en cada una de las celdas. Se sabe que esta ecuación produce errores sistemáticos dependiendo de la posición dentro del cristal en la que incide la partícula, esto debido a la subvaloración de las colas exponenciales en la ecuación. Para observar si esto ocurre para el detector preshower se gráfica la posición incidente x_{inc} versus la posición reconstruida x_{rec} . Esto se realiza para una partícula gamma de 10 [GeV] lanzada en el cristal central de la matriz. Como se observa en la Figura 5.9 a), para el detector preshower existe un error sistemático en la reconstrucción utilizando el centro de gravedad, la forma S de la curva muestra que las posiciones son reconstruidas erróneamente en las posiciones lejos de los límites o lejos del centro del cristal. Esto último puede ser explicado debido a la cantidad de energía que se filtra al resto de los cristales: En el centro la mayoría de la energía esta en el cristal de incidencia, por lo que la posición puede ser reconstruida correctamente, por otro lado, en los límites, la mayoría de la energía esta en los cristales aledaños, permitiendo una buena reconstrucción en aquellos cristales también.

En la Sección 4.5 se presentó un método simple para solucionar el problema que presenta la ecuación del centro de gravedad. Este método se basa en la ecuación 4.5.2, que se encarga

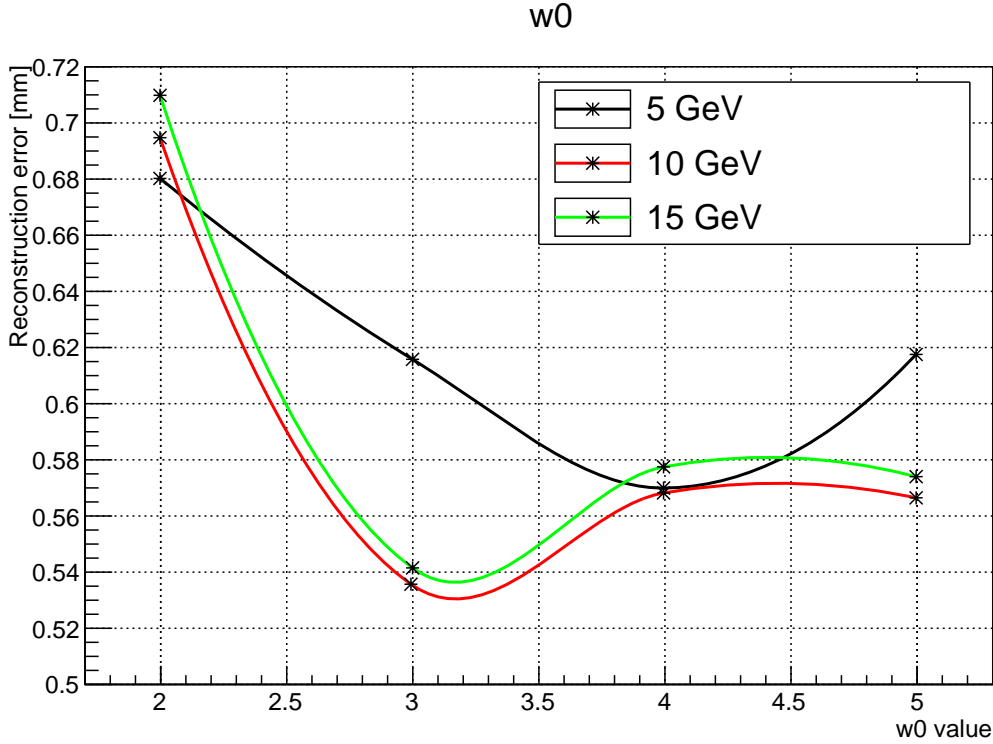
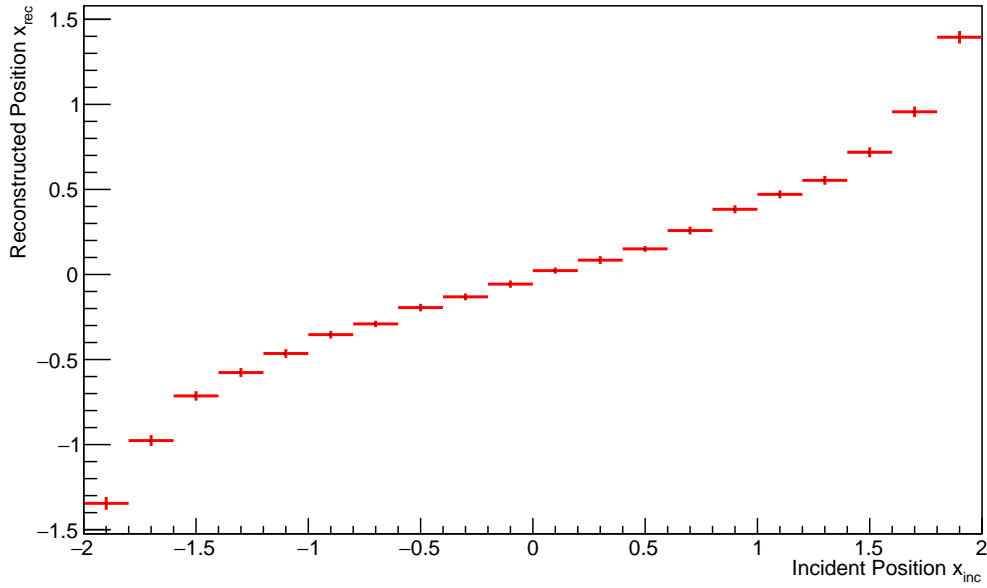


Figura 5.8: Error de reconstrucción versus valor de w_0 en 4.5.2 para el método de reconstrucción de posición

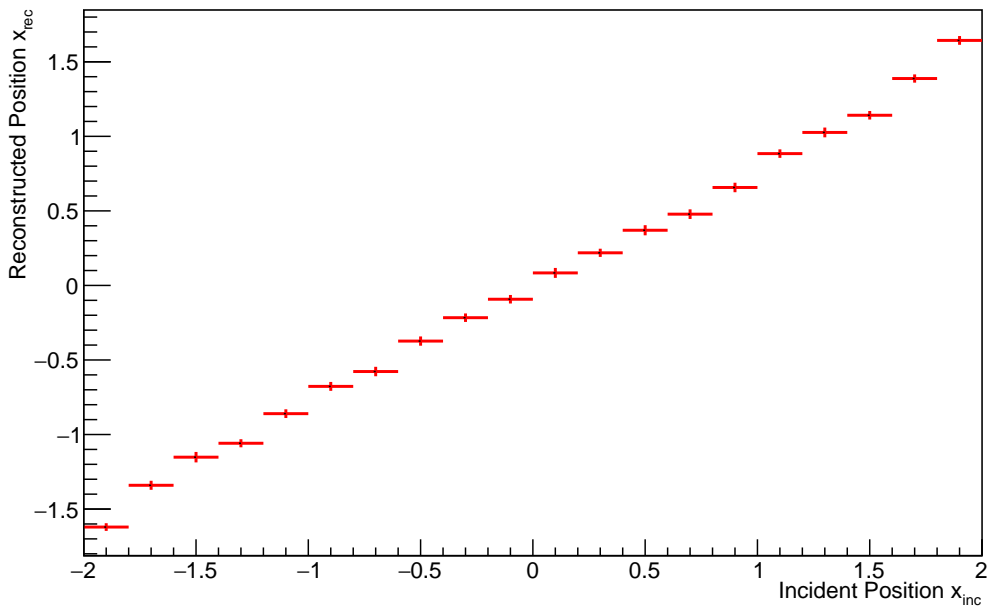
automáticamente de las colas con forma exponencial de la lluvia electromagnética. Para utilizar esta ecuación es necesario definir el valor w_0 . Recordemos que el valor de w_0 controla la importancia de las colas de la lluvia en el calculo de posición. Para definir este valor se utilizaron simulaciones. Se midió el error medio de reconstrucción $\langle |x_{rec} - x_{inc}| \rangle$ para diferentes valores de w_0 para gammas incidentes de manera normal y uniformemente distribuidos en el rectángulo central de 40x40 [mm]. Se midieron los valores para gammas de 5, 10 y 15 [GeV]. En la Figura 5.8 se muestran los resultados, en base a esto se decidió utilizar un $w_0 = 3$ que presenta la menor perdida de resolución a diferentes energías. En la Figura 5.9 b) se muestra la misma curva S obtenida anteriormente, que gráfica la posición reconstruida versus la posición incidente, pero utilizando esta vez la ecuación 4.5.2. Se observa que los errores sistemáticos ahora son mucho menores, solucionando el problema que presenta el centro de gravedad.

Systematics errors on position reconstruction



(a) Posición reconstruida utilizando el método de centro de gravedad.

Systematics errors on position reconstruction



(b) Posición reconstruida utilizando la ecuación 4.5.2.

Figura 5.9: Gráfico de posición incidente x_{inc} versus posición reconstruida x_{rec} (para uno de los ejes). En el gráfico se muestran los valores medios para la posición reconstruida en varias posiciones de incidencia. El gráfico permite observar los errores sistemáticos que se presentan debido a las diferentes posiciones de incidencia dentro de un mismo cristal.

Resolución de ambigüedades

Una gran diferencia entre el sistema de lectura del detector Preshower y los sistemas de lecturas comunes de otros detectores es que la información del detector preshower es obtenida por dos arreglos de MPPC ubicados en los extremos derecho y superior de la matriz de cristales. Cada uno de estos arreglos de lectores sólo recibe una proyección de la energía en el plano unidimensional, mientras que la posición debe ser reconstruida en el plano bidimensional. Esto provoca ciertas ambigüedades que deben ser resueltas si se quiere obtener la posición incidente de la partícula. Tal como se observa en la Figura 5.1, este sistema provoca situaciones en las que es difícil combinar las posiciones incidentes unidimensionales para reconstruir la posición bidimensional. Si, por ejemplo, en uno de los ejes se observa sólo una partícula pero en el otro se observan dos, significa que la partícula aparece solapada en uno de los ejes y la energía correspondiente a cada partícula puede ser identificada en base a la reconstrucción del eje en que se observan ambas partículas. Por otro lado, considerar que en ambos ejes se observan dos partículas, entonces no es claro que posición en uno de los ejes corresponde a que posición en el otro eje, es decir, sólo considerando la información de posición, es imposible decir si es que la reconstrucción debiese ser (x_1, y_1) y (x_2, y_2) o (x_1, y_2) y (x_2, y_1) .

El problema se complica más aún cuando más de dos partículas inciden, ya que en tales casos las combinaciones posibles crecen factorialmente con el número de partículas observadas en los ejes.

Ahora, es posible solucionar este problema considerando el conteo de fotoelectrones reconstruidos para cada partícula en cada eje. Debido a que el valor medido en cada eje es una proyección de la suma de conteos de los cristales por los que pasa la fibra que llega al lector, entonces, clústeres con energías similares (no iguales, ya que existe algo de pérdida en la propagación) en distintos ejes corresponderán a la misma partícula. Para esto debemos asumir, primero, que las pérdidas en la propagación de fotones de los cristales a las fibras es despreciable con respecto a la energía depositada por la partícula y segundo que dos partículas no depositarán exactamente la misma energía. Debido a que la interacción de la partícula con el material en el detector preshower es un proceso probabilístico, es muy poco probable que dos partículas interactúen en la misma posición y por consiguiente depositen la misma energía. Basado en esto es posible resolver las ambigüedades en la reconstrucción

mediante la unión de clústeres con energías parecidas en los distintos ejes.

En este trabajo proponemos un algoritmo para resolver ambigüedades cuando dos partículas inciden en el detector. Este algoritmo es resumido en Algoritmo 3. Dos partículas inciden solas en el detector cuando un pión neutro decae en dos fotones. El foco de este trabajo y del detector preshower es identificar estos casos. Sin embargo, se puede dar el caso de que más de dos partículas lleguen al detector. En estos casos, más de dos clústeres pueden ser identificados en alguno de los ejes, lo que complica mucho más el problema debido a que varias combinaciones entre una, dos o más de dos partículas son posibles. Cuando el número de partículas es muy grande, es muy costoso computacionalmente considerar las posibles combinaciones para resolver las ambigüedades, por lo que un algoritmo heurístico que permita encontrar la combinación más probable basado en el conteo de fotoelectrones de cada clúster sería una buena opción a considerar. Para esto se necesita estudiar y proponer un nuevo algoritmo para la tarea de resolución de ambigüedades, lo que queda fuera del alcance de este trabajo.

Algoritmo 3 Resolución de ambigüedades

```

1:  $X, E_x \leftarrow \text{Reconstruction1D}(X)$ 
2:  $Y, E_y \leftarrow \text{Reconstruction1D}(Y)$ 
3: if  $\text{len}(X) > 1$  and  $\text{len}(Y) > 1$  then
4:    $D_0 = |E_x[0] - E_y[0]|$ 
5:    $D_1 = |E_x[0] - E_y[1]|$ 
6:   if  $D_0 < D_1$  then
7:      $P_0 = (X[0], Y[0])$ 
8:      $P_1 = (X[1], Y[1])$ 
9:      $E_0, E_1 = E_x[0], E_x[1]$ 
10:  else
11:     $P_0 = (X[0], Y[1])$ 
12:     $P_1 = (X[1], Y[0])$ 
13:     $E_0, E_1 = E_x[0], E_x[1]$ 
14: else if ( $\text{len}(X) > 1$  and  $\text{len}(Y) == 1$ ) then ▷ Similar para el caso opuesto.
15:    $P_0 = (X[0], Y[0])$ 
16:    $P_1 = (X[1], Y[0])$ 
17:    $E_0, E_1 = E_x[0], E_x[1]$ 
18: else if ( $\text{len}(X) == 1$  and  $\text{len}(Y) == 1$ ) then
19:    $P_0 = (X[0], Y[0])$ 
20:    $E_0, E_1 = E_x[0], E_x[1]$ 

```

Capítulo 6

Resultados

6.1. Análisis de resultados

En esta sección se presentarán los resultados del algoritmo propuesto. Se debe notar que el propósito principal del detector preshower es identificar la posición de partículas muy cercanas entre sí, que decaen de piones neutros. Debido a esto los resultados presentados en esta sección se concentrarán en evaluar la capacidad del detector para la separación de partículas cercanas y su capacidad para reconstruir su posición. Resultados relacionados al calculo del conteo de fotoelectrones pueden ser encontrados en el Apéndice B.

El foco principal del detector preshower es detectar partículas γ muy cercanas entre sí, que comúnmente provienen del decaimiento de un pión neutro. Para lograr esto especial cuidado se debe tomar en los siguientes puntos:

- Identificación de lluvias electromagnéticas provenientes de partículas gammas individuales y reconstrucción de posición de incidencia de estas partículas.
- Reconocimiento y separación de partículas γ solapadas, especialmente para el caso de dos partículas muy cercanas (el caso del pión neutro).
- Rechazo de partículas solapadas cuando estas se presentan como una lluvia con un sólo máximo.

Los resultados presentados en esta sección se obtuvieron mediante simulaciones en el rango 1-15 [GeV]. Se utilizaron 5000 pares de partículas gammas lanzadas en posiciones de

incidencia uniformemente distribuidas en el cuadrado de 40x40 [mm] central de la matriz de cristales y con ángulo de incidencia perpendicular a la cara frontal del detector preshower. No se presentan experimentos con ángulos de incidencia no perpendicular y estos casos se analizarán con más profundidad en las conclusiones.

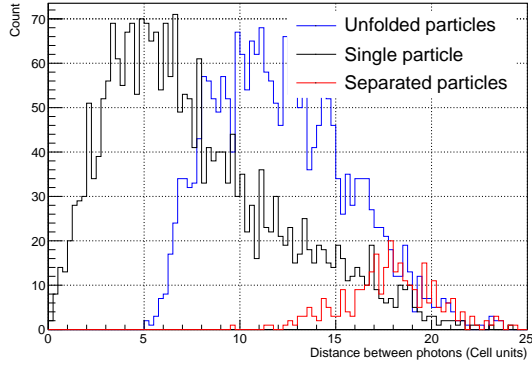
6.1.1. Identificación de partículas cercanas

La Figura 6.1 muestra la distancia (entre las partículas) a la que dos partículas pueden ser identificadas por el detector, usando diferentes combinaciones de partículas con distintas energías. Se observa que clústeres con un sólo máximo (cuando dos partículas son imposibles de detectar) son observados principalmente a distancias pequeñas entre las dos partículas (≤ 5 unidades de celda, es decir ≤ 20 [mm]). Por otro lado, clústeres con dos máximos que deben ser separados mediante el procedimiento de unfolding se observan entre las 5 y 20 unidades de celda (25-100 [mm]). Los clústeres totalmente separados se observan en poca cantidad y principalmente sobre la distancia de 15 unidades de celdas (75 [mm]). También se observa que existe un límite claro para la resolución de dos partículas del detector preshower, el que es 5 unidades de celda (o 20 [mm]).

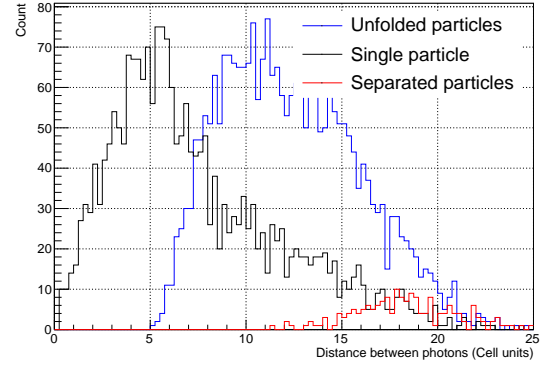
Podemos comparar estos resultados a los del detector PHOS, presentados en [8]. Los autores reportan una distancia mínima para la detección de dos partículas de 26 [mm], es decir el detector preshower permite una ganancia en resolución de alrededor de 23 %.

6.1.2. Reconstrucción de posición

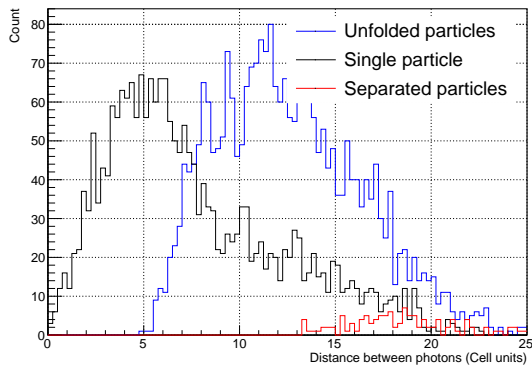
Para estudiar la eficacia del algoritmo para reconstruir la posición de las partículas incidentes se observó el error absoluto de reconstrucción entre la posición reconstruida por el algoritmo de reconstrucción y la posición de incidencia de la partícula ($< |X_{rec} - X_{inc}| >$). Esto para pares de partículas de 10 [GeV] + 1 [GeV], 10 [GeV] + 5 [GeV], 10 [GeV] + 10 [GeV], 10 [GeV] + 15 [GeV]. Se estudió el error de reconstrucción en los casos en que el algoritmo de separación de partículas solapadas deba ser aplicado y en los casos en que ambas partículas formen dos clústeres separados (el caso de un sólo máximo no fue considerado). En la Figura 6.2 se muestran histogramas para el error de reconstrucción para cada uno de los casos mencionados anteriormente. Mientras que en el Cuadro 6.1 se resumen los valores para cada uno de los casos.



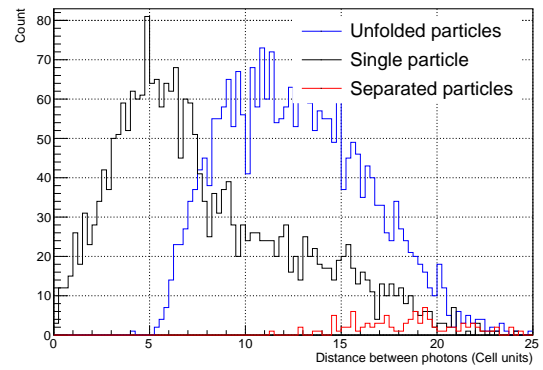
(a) 10 + 1 [GeV].



(b) 10 + 5 [GeV].



(c) 10 + 10 [GeV].



(d) 10 + 15 [GeV].

Figura 6.1: Número de clústeres identificables como una función de la distancia entre las partículas incidentes. Se analizaron 4 casos (a) 10 [GeV] + 1 [GeV] γ , (b) 10 [GeV] + 5 [GeV] γ , (c) 10 [GeV] + 10 [GeV] γ , (d) 10 [GeV] + 15 [GeV] γ . Para cada uno de los casos se muestran los clústeres con sólo un máximo, con dos máximos y los clústeres totalmente separados.

Energías [GeV]		Resolución posición [mm]		
		Unfolded	Separados	Total
10 + 1	10	1.416 ± 1.102	1.075 ± 0.7435	1.3777 ± 1.0731
	1	1.789 ± 1.45	1.025 ± 0.6956	1.7032 ± 1.4067
10 + 5	10	1.562 ± 1.189	1.183 ± 1.122	1.54 ± 1.1885
	5	1.683 ± 1.26	1.144 ± 0.7378	1.6517 ± 1.2421
10 + 10	10	1.571 ± 1.21	1.088 ± 0.8137	1.5532 ± 1.2011
10 + 15	10	1.623 ± 1.302	1.413 ± 1.012	1.6157 ± 1.2936
	15	1.556 ± 1.184	1.109 ± 0.815	1.5406 ± 1.1760

Cuadro 6.1: Error de reconstrucción para partículas γ solapadas y totalmente separadas. Se presentan resultados para partículas de 1, 5, 10 y 15 [GeV].

Se observa una pequeña dependencia entre el error de reconstrucción y la energía de la partícula incidente. Esto es de esperar ya que las partículas de menor energía pueden ser cubiertas por las lluvias de las partículas de energía mayor, dificultando su reconstrucción. Esto es claro en el caso 10+1 donde el error para la partícula de 1 [GeV] es claramente mayor que para la partícula de 10 [GeV]. Sin embargo, las diferencias entre los errores de reconstrucción para el rango 1-15 [GeV] son relativamente pequeñas, lo que muestra la dependencia menor del algoritmo de reconstrucción de posición con la energía de la partícula incidente. El error de reconstrucción es similar para el caso de clústeres separados y clústeres solapados donde el algoritmo de unfolding debe ser utilizado, lo que demuestra la capacidad del algoritmo de unfolding para separar partículas cercanas de manera correcta. Existe una pequeña diferencia entre ambos errores, lo que es de esperar ya que el algoritmo de unfolding agrega un paso extra de estimación, además que el cubrimiento mutuo de las lluvias electromagnéticas dificulta la reconstrucción de posición.

Es posible comparar los resultados en la reconstrucción de posición a los reportados para PHOS en [8]. Ahí se reporta en términos de la media cuadrática para el error, o rms, definida como $(\sqrt{\frac{1}{n} \sum_i (x_{rec} - x_{inc})^2})$. Para partículas gammas de 1 y 10 [GeV] con ángulo de incidencia de 0 grados se reporta un error cuadrático medio de 3.4 [mm] y 1 [mm]. Para el detector preshower, en cambio, el rms para partículas de 1 y 10 [GeV] es de 0.9641 [mm] y 0.9198 [mm] respectivamente. Aunque para el caso de 10 [GeV] los errores son parecidos, para 1 [GeV] el detector preshower tiene un error cuadrático medio mucho menor (equivalente al 28.3% del error en PHOS).

Para un análisis más detallado de los errores de reconstrucción para las distintas

situaciones que se presentan en el algoritmo de reconstrucción el lector puede dirigirse al Apéndice B.

Los errores en Y son presentados en el Apéndice B. El error para el eje Y es mayor que para el eje X , la razón de esto se estudiará con mayor profundidad en las conclusiones.

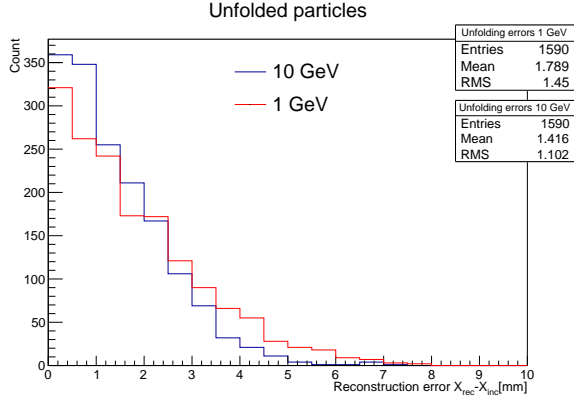
6.1.3. Rechazo de partículas solapadas con un sólo máximo

A continuación se estudia el algoritmo de rechazo para partículas con distancias ≤ 20 [mm], que es donde el algoritmo de reconstrucción falla en la identificación de dos partículas. Para esto, se simularon partículas individuales y pares de partículas (80 mil en total). Se compararon varios algoritmos de clasificación con el método basado en cortes en la variable de dispersión presentado en la Sección 4.4. Los algoritmos de clasificación utilizados fueron: SVM (*support vector machines*), BDT (*boosted decision trees*) y MLP (*multilayer perceptron*).

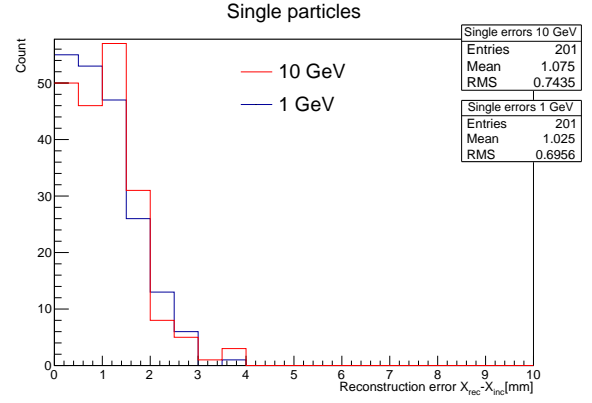
Para cada uno de los algoritmos de clasificación, las variables fueron normalizadas al rango $[-1, 1]$. Además se incluyó la energía total para cada clasificador. En el caso de la SVM se usó un kernel *Gaussiano* con regularización de $C = 1.0$. Para el BDT se usó el algoritmo de aprendizaje *AdaBoost* [39] con 200 árboles de decisión. Además, se usó una profundidad máxima de 3 y un factor de aprendizaje de 0.5. El MLP fue entrenado con *stochastic gradient descent* con un factor de aprendizaje de 0.02, regularización ℓ_2 y durante 600 iteraciones. La arquitectura utilizada consistió de dos capas ocultas de tamaño $15 - 5$, con activaciones del tipo *tanh* en las capas ocultas y activación *sigmoid* en la salida. Los parámetros fueron seleccionados mediante búsqueda de grilla. El 50 % del conjunto de datos se usó en el entrenamiento, dejando el resto para test.

En el caso de los cortes, los valores de los cortes fueron seleccionados automáticamente con la finalidad de maximizar el rechazo de background dado diferentes valores de la eficiencia en la señal. Para esto se utilizó el método de Monte Carlo, tal como se explicó en la Sección 4.4. Los cortes fueron realizados en las variables de dispersión y dispersión corregida (ecuación 4.4.2). Se observó que los resultados del método empeoraban al agregar más variables. Todos los métodos fueron implementados utilizando TMVA [24].

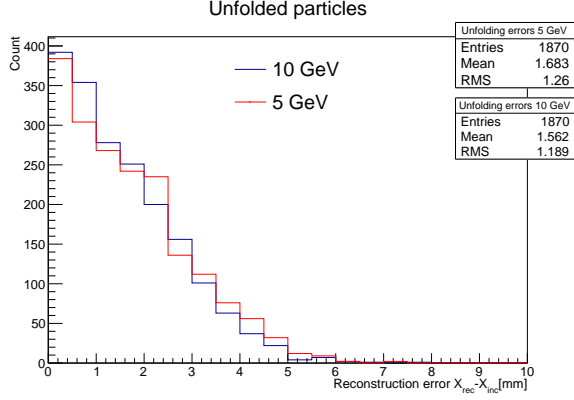
Las curvas ROC para cada uno de los métodos son mostradas en la Figura 6.3. Se puede observar que los mejores resultados son obtenidos por el *multilayer perceptron*.



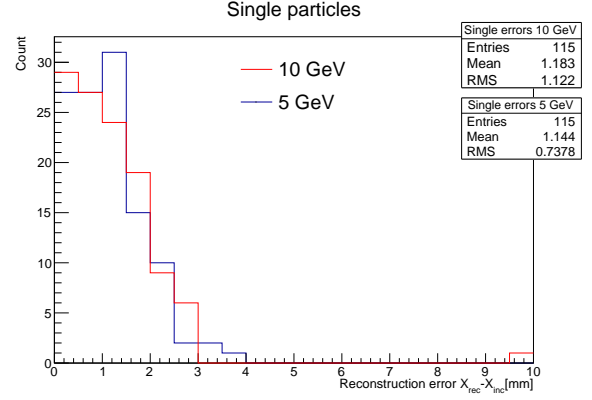
(a) 10 + 1 [GeV].



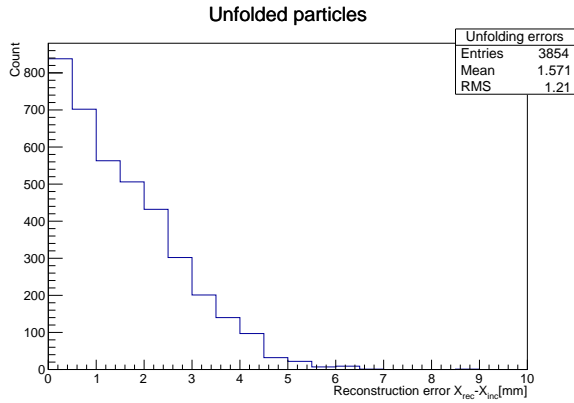
(b) 10 + 1 [GeV].



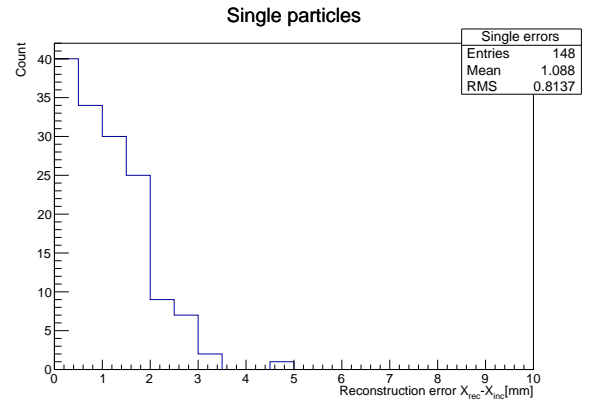
(c) 10 + 5 [GeV].



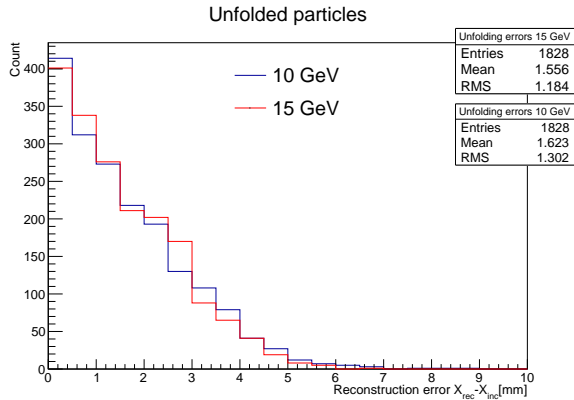
(d) 10 + 5 [GeV].



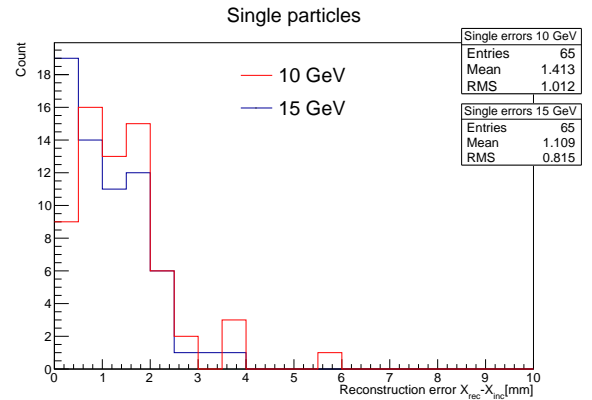
(e) 10 + 10 [GeV].



(f) 10 + 10 [GeV].



(g) 10 + 15 [GeV].



(h) 10 + 15 [GeV].

Figura 6.2: Histogramas del error de reconstrucción $X_{rec} - X_{inc}$ para varias combinaciones de partículas γ . En cada gráfico se muestra el error para cada una de las partículas por separado (en el caso de 10+10 [GeV] sólo se muestra el error para 10 [GeV]).

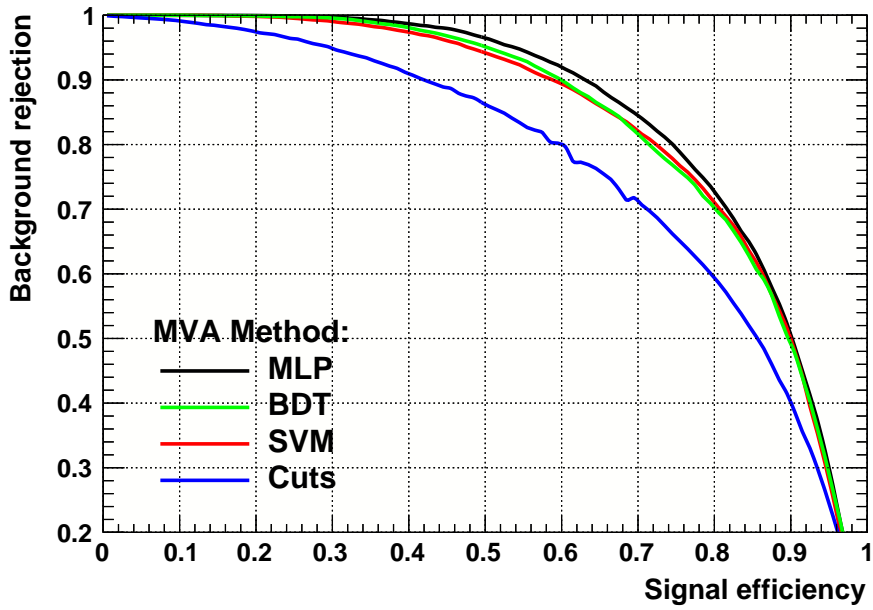


Figura 6.3: Curvar ROC para cada uno de los métodos.

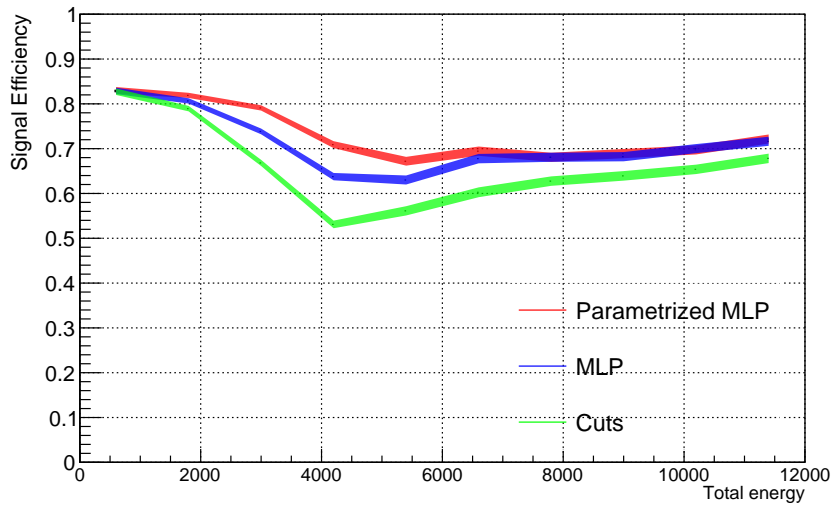


Figura 6.4: Eficiencia en la señal para: a) Una red neuronal con la energía total como variable de entrada (parametrizada), b) Una red neuronal sin la energía como variable de entrada. c) El método basado en cortes. La comparación fue realizada para diferentes rangos de energía.

Para continuar, se estudia la dependencia de los clasificadores a la energía de la partícula incidente. Esto para comprobar la hipótesis de que al usar la energía total como variable de entrada se logra obtener un clasificador más resistente al efecto de distintos rangos de energía. Para esto, se comparó el *multilayer perceptron* al método basado en cortes y a un *multilayer perceptron* que no incluye la energía como variable de entrada. Se midió la eficiencia en la señal para cada método y para diferentes rangos de energías. Notar, sin embargo, que el entrenamiento fue realizado en el rango completo de energías. Los resultados son mostrados en la Figura 6.4. Mientras que el método basado en cortes varía considerablemente con la energía, el desempeño del clasificador es mucho más estable, en especial cuando la energía es usada como variable de entrada.

6.1.4. Reconstrucción de piones neutros

Para estudiar la capacidad del detector preshower para detectar piones neutros decayendo en dos gammas muy cercanos entre sí se utilizaron simulaciones de piones lanzados al detector de manera perpendicular. Sólo los piones que decaen y para los cuales ambos gammas llegan al detector fueron considerados. Los piones fueron lanzados desde una distancia de 250 [cm] a la cara frontal del detector. Se utilizó un rango de energía de 10-60 [GeV]. En la Figura 6.5 se observa la fracción de partículas gammas producidas por piones neutros que fueron detectadas como dos partículas diferentes. Se observa que a mayor energía es más difícil identificar el decaimiento del pión neutro como dos partículas separadas. Esto se debe a que a mayor energía, menor el ángulo de apertura entre los gammas, lo que hace más difícil la identificación. Incluso para energías bajas no se identifican la totalidad de piones, esto se produce porque a menor energía existe mayor probabilidad de que se produzca un gamma de baja energía que no sea identificado por el detector preshower.

Pi0 decay reconstruction

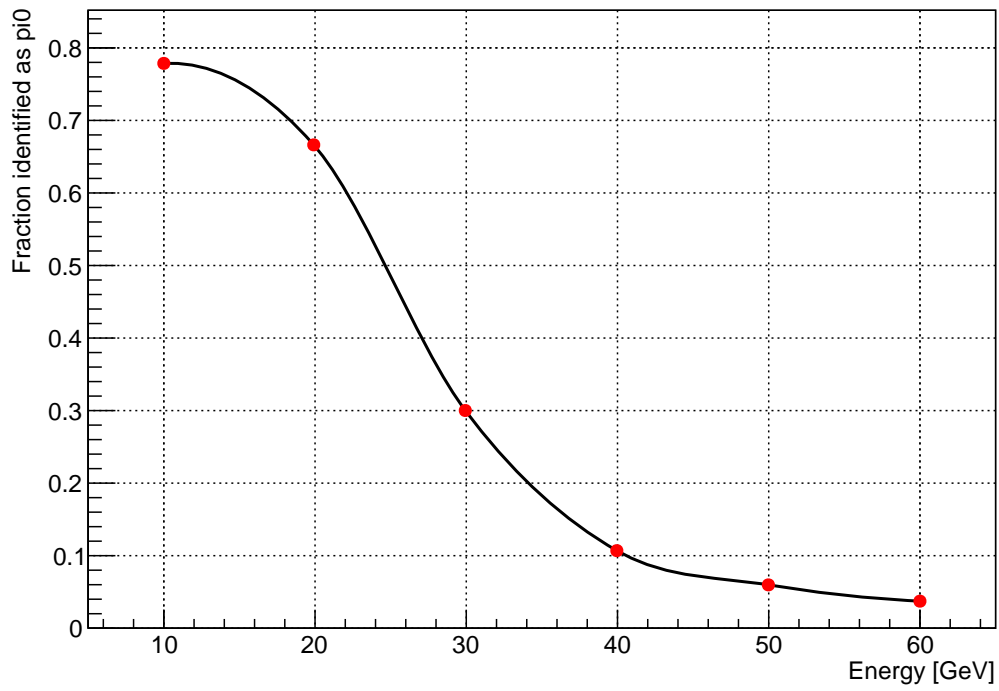


Figura 6.5: Fracción de π^0 decayendo en dos gammas que golpean el detector y que son identificados como dos gammas por el algoritmo de reconstrucción. Se presentan resultados para gammas en el rango de energía 10-60 [GeV].

Capítulo 7

Conclusiones y Trabajo Futuro

En este trabajo hemos propuesto un algoritmo de reconstrucción completo para el detector preshower siendo construido por investigadores del Centro Científico y Tecnológico de Valparaíso (CCTVal). El algoritmo es capaz de identificar y reconstruir la posición de partículas muy cercanas entre sí. De esta manera soluciona un problema de gran importancia para el trabajo de detección de partículas en la física de altas energías. El algoritmo es, además, capaz de identificar piones neutros decayendo en dos gammas muy cercanos entre sí. De este trabajo se pueden concluir algunos puntos de vital interés para próximas iteraciones en el diseño del detector preshower.

- El detector preshower es capaz de identificar partículas muy cercanas entre sí, mejor que otras propuestas similares de la literatura. Los resultados muestran que el detector preshower permite una ganancia en resolución espacial de alrededor del 23% en comparación con el detector PHOS [8]. Esto ya es una ganancia considerable y puede ser mejorada aún más en próximas iteraciones del diseño del detector.
- En términos de error de reconstrucción el detector preshower obtiene resultados muy competitivos al ser comparado con los resultados obtenidos en [8]. Esto se observa especialmente para partículas de baja energía, por ejemplo para partículas de 1 [GeV] el detector preshower tiene un error cuadrático medio equivalente al 28.3% del error del detector PHOS. Esto es debido a que el diseño de la matriz de cristales del detector preshower provee una mejor resolución para la reconstrucción espacial, incluso cuando la energía de la partícula incidente es menor.

- Se propone el uso de un algoritmo de clasificación para el rechazo de dos o más partículas observadas como una sola. Se concluye que el método supera al método basado en cortes que es comúnmente utilizado en física de partículas. Además, mediante la inclusión de la energía total como variable de entrada hemos disminuido la dependencia del algoritmo en la energía de la partícula incidente.
- El error de reconstrucción es mayor en el eje Y que en el eje X. Esto se entiende principalmente por dos razones. Según el modelo de la ecuación 5.2.2 que considera los efectos de la electrónica y diseño físico del detector, la varianza de la propagación de la energía en el eje Y es mayor que en el eje X. Lo anterior es debido a factores electrónicos y físicos fuera del alcance del algoritmo de reconstrucción (por ejemplo, diferencias en el aislamiento de las fibras que transportan los fotones). El efecto de ensanchamiento de las lluvias se puede observar claramente en la Figura 5.6. Esto dificulta la reconstrucción de posición produciendo errores mayores en el eje Y.

Otra razón de esta diferencia es el método de resolución de ambigüedades. Según el funcionamiento actual del algoritmo, cuando dos partículas son observadas en ambos ejes se resuelve el problema de ambigüedad mediante el uso de la diferencia de conteo de fotones entre las lluvias detectadas en cada eje. Como se mencionó al presentar el método, este algoritmo no funciona en el caso que ambas partículas depositen una cantidad de energía similar. Aunque estos casos son poco probables, al ocurrir producen una mala elección de posiciones en el eje Y (el algoritmo funciona emparejando posiciones en el eje Y a las posiciones reconstruidas en el eje X) lo que a su vez produce errores de reconstrucción muy mayores a la media (para un mayor análisis de este problema el lector se puede dirigir al Apéndice B, Figura B.3). Para mitigar este problema se puede utilizar información extra para emparejar las lluvias electromagnéticas, por ejemplo, utilizar variables derivadas de la forma de la lluvia.

- El detector preshower por sí sólo no puede manejar la situación en las que el ángulo de incidencia es no perpendicular a la cara frontal del detector. Para lograr esto otros tipos de detectores implementan métodos como la segmentación transversal o la medición de los tiempos en los que la partícula atraviesa el detector. Diferentes ángulos de incidencia afectan los resultados del preshower. Por ejemplo, el error

medio para partículas individuales de 10 [GeV] es de 0.7402 [mm] cuando el ángulo de incidencia es de 0 grados. En cambio, si el ángulo es 6 grados el error medio pasa a ser 3.573 [mm]. Si bien el detector preshower no es capaz de solucionar este problema por sí solo, otras secciones podrían medir el ángulo de incidencia y mediante estas mediciones implementar correcciones de posición para los valores medidos por el detector preshower.

En base a esto se propone como trabajo futuro:

- En base a los experimentos realizados en este trabajo, se observó que un importante factor que afecta la capacidad del algoritmo para reconstruir la posición de la partícula incidente de una manera precisa (especialmente en el eje Y) es el ensanchamiento de las lluvias debido a factores electrónicos de la detección. Este factor afecta, además, la capacidad de identificar partículas incidentes cercanas. Este efecto se refleja en el modelo de propagación utilizado para simular las mediciones del MPPC de la ecuación 5.2.2. En base a esto, una de las conclusiones más importantes de este trabajo es: Las próximas iteraciones del diseño del detector preshower debiesen intentar mejorar estos valores, lo que producirá una mejora directa en la capacidad para identificar partículas cercanas y en la precisión de la reconstrucción del detector preshower. Para evitar el ensanchamiento de las lluvias se puede, por ejemplo, considerar diferentes materiales reflectantes en los cristales o fibras, con la finalidad de evitar la fuga de fotones.
- Problemas como la pérdida de precisión para ángulos diferentes a cero pueden ser solucionados mediante el trabajo conjunto del detector preshower con el resto de la cadena de detección. Esto sólo puede ser considerado en un contexto más amplio, teniendo un conocimiento más preciso del uso final del detector preshower. En la etapa actual de producción no se conoce la constitución final de la cadena de detección, por lo que este tipo de trabajo sólo se puede hacer en base a supuestos. Debido a lo anterior esto se deja como trabajo futuro.

Apéndices

Apéndice A

Conceptos Básicos

En esta sección se presentará una breve descripción de los conceptos físicos relevantes para comprender de una manera general el diseño y funcionamiento del detector Preshower. La finalidad de esta sección es entregar al lector, con conocimientos básicos en física moderna, los conceptos principales que le permitirán comprender de manera general las decisiones de diseño tanto en la construcción del detector, pero principalmente, en la propuesta de algoritmo de reconstrucción. Para mayor información y un tratamiento más técnico de los temas el lector se puede dirigir a [40, 41, 42].

A.1. Paso de partículas por la materia

El estudio de las interacciones de las partículas con la materia constituye la base del estudio de los detectores. Un detector detectará una partícula mediante las interacciones que esta tiene con los materiales del detector, por lo que el estudio de las principales interacciones de las partículas con la materia es indispensable para entender el funcionamiento de los detectores.

A.1.1. Sección eficaz y recorrido libre

La sección eficaz es un concepto fundamental en la física de partículas. Este se relaciona con la probabilidad de interacción de un partícula con un blanco. Considerar un haz de partículas que incide en un blanco de modo que son dispersadas (en parte) en un ángulo

θ en dirección de un detector, siendo el blanco de área A menor que la superficie del haz. Considerar que el flujo de partículas en el haz es $F[cm^{-2}s^{-1}]$ (número de partículas por unidad de área y por unidad de tiempo). Además considerar que el blanco X es muy delgado (δx), de manera que la probabilidad de que las partículas del blanco se posicionen una delante de otra es baja. También, el blanco tiene una densidad de N partículas, es decir que el haz incidente verá $N\delta x$ partículas. Sea, N_s el número promedio de partículas dispersadas en el ángulo $d\Omega$ del detector, entonces se puede definir la sección eficaz diferencial como:

$$\frac{d\sigma}{d\Omega}(E, \Omega) = \frac{1}{F} \frac{dN_s}{d\Omega},$$

que es la fracción promedio de partículas dispersadas hacia $d\Omega$ por unidad de tiempo y por unidad de flujo. La sección eficaz total se define como

$$\sigma(E) = \int d\Omega \frac{d\sigma}{d\Omega}.$$

Ahora, debido a que el área del blanco es menor que el haz, el número de partículas que pueden interactuar del haz es $N_{inc} = FA$ y el número promedio de partículas dispersadas en $d\Omega$ es:

$$N_s(\Omega) = FAN\delta x \frac{d\sigma}{d\Omega},$$

entonces, el número de partículas dispersadas en todos los ángulos es

$$\begin{aligned} N_{tot} &= FAN\delta x\sigma \\ &= N_{inc}N\delta x\sigma, \end{aligned} \tag{A.1.1}$$

de esta manera se puede definir la probabilidad de interacción de una partícula en un material de espesor δx como

$$\frac{N_{tot}}{N_{inc}} = n\sigma\delta x, \tag{A.1.2}$$

es decir que la sección eficaz es proporcional a la probabilidad de interacción del haz incidente con el blanco. El concepto de sección eficaz es fundamental y constituirá el valor principal para evaluar las posibles interacciones presentes en el detector. Se puede, también, calcular la *distancia media recorrida* por una partícula en un material de cualquier grosor x . Considerar

$P(x)$ como la probabilidad de que una partícula no tenga interacción en una distancia x de material y $w dx$ como la probabilidad de tener una interacción entre x y dx . La probabilidad de no tener una interacción entre x y $x + dx$ es

$$\begin{aligned} P(x + dx) &= P(x)(1 - w dx), \\ P(x) + \frac{dP(x)}{dx} dx &= P(x) - P(x)w dx, \\ dP(x) &= -wP(x) dx. \end{aligned}$$

Resolviendo la ecuación diferencial anterior y requiriendo que $P(0) = 1$ se obtiene que

$$P(x) = \exp(-wx).$$

Finalmente, es posible calcular la distancia media recorrida por una partícula en un material como

$$\begin{aligned} \lambda &= \frac{\int x P(x) dx}{\int P(x) dx} \\ &= \frac{1}{w} \\ &= \frac{1}{N\sigma}. \end{aligned}$$

Usando la ecuación A.1.2 y el hecho de que la probabilidad de interacción en un blanco delgado puede ser expandida como

$$\begin{aligned} P_{int}(\delta x) &= 1 - \exp(-wx), \\ &= 1 - \left(1 - \frac{\delta x}{\lambda} + \dots\right), \\ &\simeq \frac{\delta x}{\lambda}. \end{aligned}$$

Entonces, la probabilidad de supervivencia queda como

$$P(x) = e^{-x/\lambda}.$$

El valor λ también se conoce como longitud de atenuación y se puede entender como la distancia a la cual la probabilidad de que la partícula no sea absorbida cae en $1/e$. La longitud de atenuación es un valor importante para evaluar y comprar diferentes materiales utilizados como productores de lluvia en un detector calorímetro.

En la siguiente sección se analizarán las interacciones de partículas con la materia útiles para comprender el funcionamiento del detector preshower.

A.1.2. Interacción de partículas con la materia

Las partículas pueden interactuar de muchas maneras con la materia, muchas de estas interacciones están muy estudiadas y sus secciones eficaces son conocidas. En este trabajo nos concentraremos en las interacciones sufridas por fotones y electrones que pudiesen tener influencia en el funcionamiento y posterior reconstrucción del detector preshower. El detector preshower es un detector del tipo calorímetro que produce lluvias electromagnéticas (esto se procederá a explicar en mayor detalle más adelante), por lo que son las interacciones involucradas en esto las que interesaran en este trabajo.

Interacción de partículas con átomos

En este caso la partícula incidente interactúa directamente mediante colisiones con el átomo. Se pueden dar dos casos, una colisión con los electrones del átomo, de manera que el átomo queda ionizado y la partícula incidente pierde energía, o una colisión elástica con el núcleo del átomo, donde la partícula incidente cambia de dirección y pierde un poco de energía (despreciable). La pérdida de energía por colisión con electrones es modelada por la fórmula *Bethe-Bloch* mientras que las colisiones con los núcleos, en la que es mas probable que la partícula incidente cambie de dirección, es modelada por la fórmula de *Rutherford*.

Interacción de electrones y positrones

Para electrones y positrones otra interacción del tipo electromagnética, conocida como radiación de frenado (o *bremsstrahlung*), afecta a la partícula incidente (además de las

interacciones por colisión), por lo que la pérdida energía total será

$$\frac{dE}{dx} = \left(\frac{dE}{dx} \right)_{col} + \left(\frac{dE}{dx} \right)_{rad}.$$

La radiación de frenado (o *bremssstrahlung*) sucede en presencia de un campo eléctrico intenso. En este los electrones son frenados y emiten fotones, perdiendo energía. Es la principal causa de pérdida de energía a altas energía. El electron incidente con energía E_i emite un fotón de energía $E_\gamma = E_i - E_f$ donde E_f es la energía final del electrón. Es posible demostrar que la pérdida de energía por la radiación de frenado es linealmente proporcional con E_i . Además, a altas energías la sección eficaz de la radiación de frenado es casi constante y es posible definir la *longitud de radiación* X_0 como

$$\frac{1}{X_0} = n_\alpha \sigma_{rad},$$

donde n_α es el número de átomos por cm^3 en el material y σ_{rad} la sección eficaz de la radiación de frenado. Es posible demostrar que

$$E = E_0 e^{-x/X_0},$$

es decir, el valor X_0 representa la distancia media recorrida en la que el electrón ha disminuido a $1/e$ de su valor de energía inicial mediante el proceso de bremsstrahlung. La longitud de radiación depende del material y será un valor importante a considerar al diseñar un detector.

El positrón, al igual que el electrón, pierde energía por ionización y radiación de frenado, sin embargo, a muy baja velocidad puede formar un *positronio* con el electrón, un partícula muy parecida a un átomo de hidrógeno pero con estados menos espaciados. El positronio se aniquila con sección eficaz σ_{an} produciendo dos fotones de $0.511MeV$,

$$e^+ + e^- \rightarrow \gamma + \gamma,$$

aunque también se pueden producir tres gammas finales (dependiendo del estado del positronio).

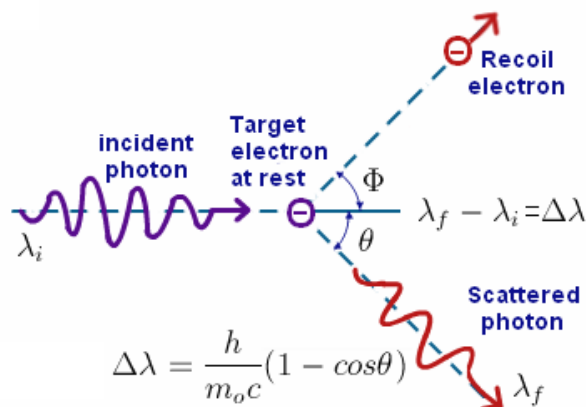


Figura A.1: Diagrama del efecto Compton entre un fotón y un electrón.

Interacción de partículas gamma

Los fotones pueden interactuar de varias maneras con la materia, estas interacciones pueden producir partículas cargadas que permitirán identificar al fotón. Las interacciones de fotones con la materia pueden ser por efecto fotoeléctrico, efecto Compton, creación de pares, difusión Rayleigh con un electrón del material y absorción fotonuclear. Los dos últimos procesos son poco frecuentes en el rango de energía importante para este trabajo por lo que no se analizarán.

El efecto fotoeléctrico ocurre cuando un fotón es absorbido por un electrón del átomo, obteniendo este último la energía del fotón incidente, si esta energía supera la función de trabajo del material, el electrón es liberado del átomo. Los electrones liberados de esta manera son llamados fotoelectrones, para diferenciarlos de los electrones incidentes. La energía del electrón liberado es $E_e = E_\gamma - E_b$, donde E_b es la función de trabajo y E_γ es la energía del fotón que es igual a $E_\gamma = h\nu$.

Por otro lado, el efecto de Compton se trata de una difusión elástica del fotón con un electrón que actúa como si estuviese libre, $\gamma + e_- \rightarrow \gamma' + e'_-$. El fotón sale en un ángulo θ y con una nueva longitud de onda λ' , dada por $\lambda' - \lambda = \frac{h}{mc}(1 - \cos\theta)$. Un diagrama de la interacción se muestra en la Figura A.1.

La creación de pares, proceso mediante el cual un fotón produce un electrón y un positrón, $\gamma \rightarrow e^+ + e^-$, sucede en presencia de un campo eléctrico intenso. Para que esto ocurra es necesario que la energía del fotón cumpla que $E_\gamma \geq 1.022MeV$. El ángulo de emisión del par e^\pm respecto al fotón a alta energía sigue $\theta \sim \frac{mc^2}{E_\pm}$, por lo que a mayor energía menor el

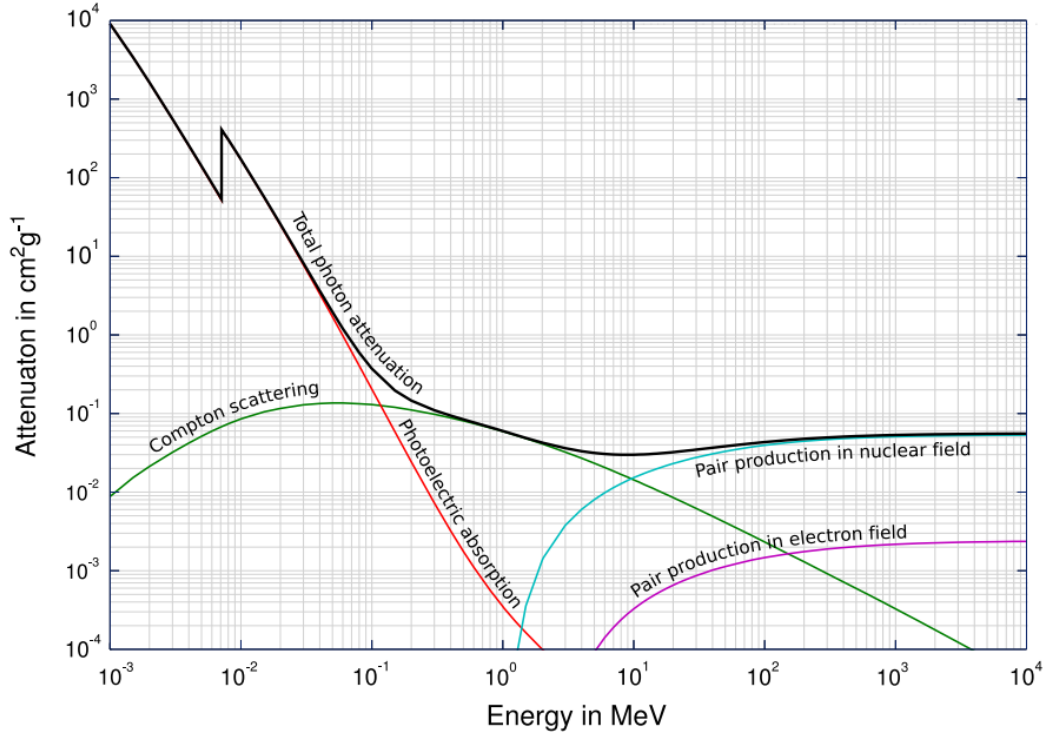


Figura A.2: Longitud de atenuación para diversos procesos de un fotón.

ángulo de emisión.

A altas energías la creación de pares es el proceso dominante, con sección eficaz σ_{par} constante. Al igual que para el electrón se puede introducir la longitud de conversión (atenuación) X_c como

$$\frac{1}{X_c} = n_\alpha \sigma_{par} = \frac{9}{7} X_0.$$

Además, se cumple que la intensidad del haz de fotones se comporta en función de X_c como

$$I = I_0 e^{-x/X_c}.$$

En la Figura A.2 se muestran diferentes longitudes de atenuación dada la energía del fotón incidente, para cada uno de los procesos analizados. La figura muestra longitudes de atenuación comunes que deben ser adaptadas dependiendo del material. En esta se observa claramente como la producción de pares predomina a altas energías, mientras que en la región de bajas energías el proceso predominante es el efecto fotoeléctrico.

A.2. Tipos de detectores

Como una primera clasificación los detectores pueden agruparse en dos grupos, detectores de partículas cargadas y detectores de partículas neutras. Mientras que las partículas cargadas pueden ser detectadas directamente mediante diversos métodos, la única forma de detectar partículas neutras es hacerlas interactuar en el detector con la finalidad de que produzcan partículas cargadas secundarias. Luego es posible reconstruir las partículas neutras mediante la información de las partículas secundarias.

Los detectores de partículas neutras se pueden clasificar según la partícula a detectar:

- Los detectores de fotones (γ) usan las interacciones electromagnéticas de los fotones con la materia, es decir el efecto fotoeléctrico, efecto Compton y la creación de pares. Aquí son útiles los detectores calorímetros que serán analizados más adelante.
- La detección de neutrinos (ν) utiliza la interacción débil de estos con la materia, debido a que esta es la única fuerza que los afecta. Dado que la probabilidad de interacción de los neutrinos es muy baja el proceso de detección es considerablemente más complicado.
- Los detectores de neutrones (n) hacen uso de la interacción fuerte de estos con la materia. Un ejemplo es hacer uso de la colisión elástica entre un protón y un neutrón, en esta el protón adquiere energía y se puede utilizar esto para detectar el neutrón.

Los detectores de partículas cargadas, por otro lado, pueden ser clasificados según el proceso que utilizan para medir las propiedades de las partículas con carga.

- Los detectores electrónicos miden los electrones liberados por ionización. Estos electrones acelerados en un potencial producen a su vez más ionización generando una avalancha que produce una señal medible. El medio de ionización puede ser un gas (por ejemplo, en los detectores Geiger-Müller o en los contadores proporcionales) o puede ser un cristal semiconductor donde la avalancha es generada por la creación de pares e^- -hueco.
- En la detección por trazas se utiliza un medio en el cual la trayectoria de la partícula se hace visible (cámaras de niebla, de burbujas, de emulsiones, de chispas, entre otras). En presencia de un campo eléctrico \vec{B} la trayectoria de la partícula es curvada en función de su momento, lo que permite estimarlo.

- Los contadores por centelleo y de Cherenkov detectan la emisión luminosa de las partículas que son excitadas por las partículas incidentes. Estos fotones pueden ser medidos utilizando el efecto fotoeléctrico y amplificadores del pulso eléctrico (fotomultiplicadores).

Apéndice B

Resultados adicionales

En esta sección se presentarán resultados adicionales del funcionamiento del algoritmo de reconstrucción.

B.1. Reconstrucción de posición

Como una forma de comprender mejor el funcionamiento del algoritmo de reconstrucción, podemos analizar los errores de reconstrucción para cada una de las posibilidades que se dan al reconstruir dos partículas:

1. Sólo una partícula es identificada de manera errónea.
2. Un clúster es identificado en uno de los ejes y en el otro un clúster con dos máximos.
3. Dos clústeres son identificados en un eje y en el otro un clúster con un máximo.
4. Dos clústeres separados son identificados en ambos ejes.
5. Dos clústeres en un eje y en el otro un clúster con dos máximos.
6. Un clúster con dos máximos en ambos ejes.

El algoritmo de unfolding debe ser aplicado en los casos 2, 5 y 6. Mientras que la resolución de ambigüedades utilizando el conteo de fotoelectrones debe ser aplicada en los casos 3 y 6.

En la Figura B.1 y B.2 se presentan histogramas para el error de reconstrucción para cada uno de estos casos y para las combinaciones de partículas de $10+1$ [GeV] y $10+10$

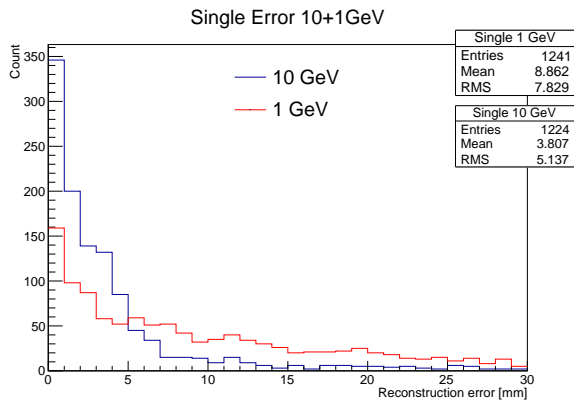
[GeV]. En el Cuadro B.1 se presentan los valores resumidos para el error de reconstrucción para las combinaciones de partículas 10+1, 10+5, 10+10 y 10+15 [GeV] y para cada uno de los casos enumerados anteriormente.

Energías [GeV]	Resolución posición [mm]						
	Caso 1	Caso 2	Caso 3	Caso 4	Caso 5	Caso 6	
10 + 1	10	3.807 ±	2.1 ±	1.876 ±	1.382 ±	1.632 ±	2.253 ±
		5.137	2.523	2.66	0.7241	1.135	1.081
	1	8.862 ±	3.603 ±	3.032 ±	1.035 ±	1.68 ±	3.093 ±
		7.829	4.685	4.503	0.4122	1.357	1.394
10 + 5	10	5.224 ±	2.538 ±	2.105 ±	1.18 ±	2.051 ±	2.407 ±
		6.268	3.169	2.943	0.6835	1.692	1.103
	5	6.936 ±	2.996 ±	3.187 ±	0.6828 ±	1.7 ±	2.67 ±
		7.14	3.798	4.803	0.33	1.043	1.129
10 + 10	10	6.397 ±	2.617 ±	2.773 ±	1.132 ±	2.039 ±	2.526 ±
		6.915	3.588	4.963	0.7415	1.408	1.101
10 + 15	10	5.929 ±	2.819 ±	3.153 ±	0.7519 ±	1.795 ±	2.665 ±
		6.585	3.709	4.502	0.0	0.9704	1.424
	15	5.656 ±	2.628 ±	2.145 ±	0.6303 ±	1.665 ±	2.449 ±
		6.585	3.471	3.18	0.0	0.9577	1.088

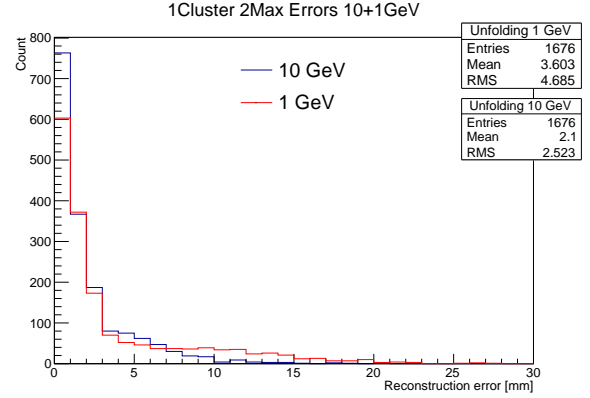
Cuadro B.1: Error de reconstrucción para partículas γ para cada uno de los casos del algoritmo. Se presentan resultados para partículas de 1, 5, 10 y 15 [GeV].

Algunas conclusiones que se pueden obtener de estos resultados:

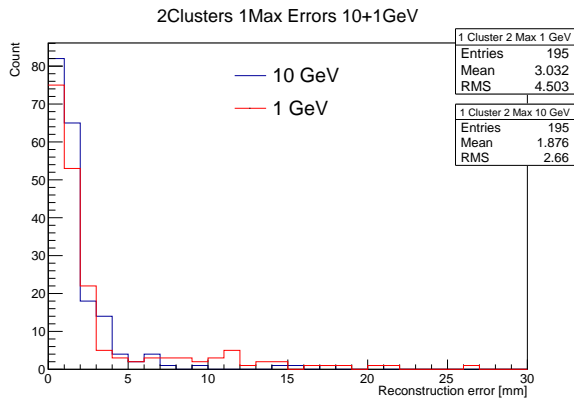
- Los casos en que se observan dos clústeres separados son poco probables en comparación con el resto. Esto debido al ancho de las lluvias observadas por cada uno de los ejes (ensanchamiento propiciado por los efectos electrónicos tomados en cuenta en el modelo de la ecuación 5.2.2).
- El error es mayor en los casos en que se observa sólo un clúster o un máximo en alguno de los ejes. Esto es de esperar ya que en caso de observarse un máximo la posición reconstruida es una posición intermedia entre ambas partículas, lo que aumenta el error de reconstrucción.
- Aunque existen algunas diferencias entre los errores para cada uno de los casos, estas son mas bien pequeñas, lo que demuestra que el algoritmo de reconstrucción funciona bien para cada uno de los posibles casos.



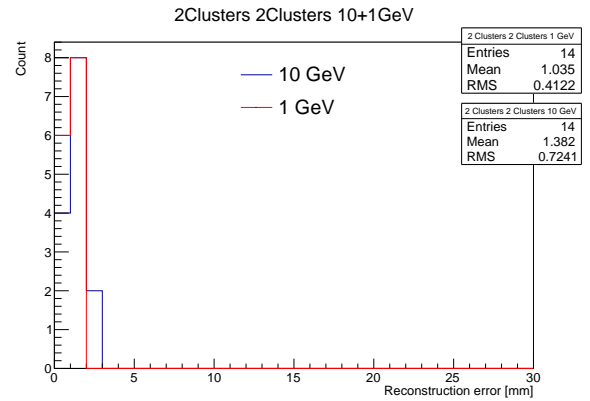
(a) Una partícula identificada.



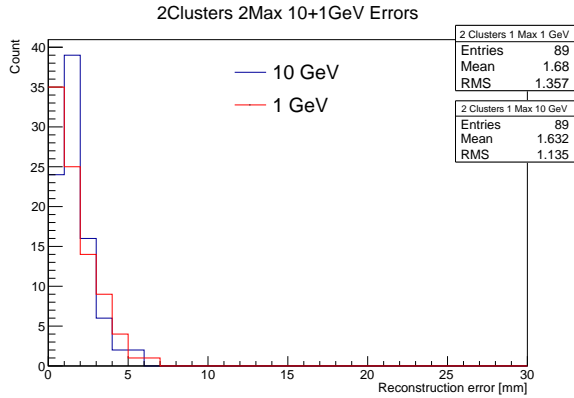
(b) Un clúster en un eje y un clúster con dos máximos en el otro.



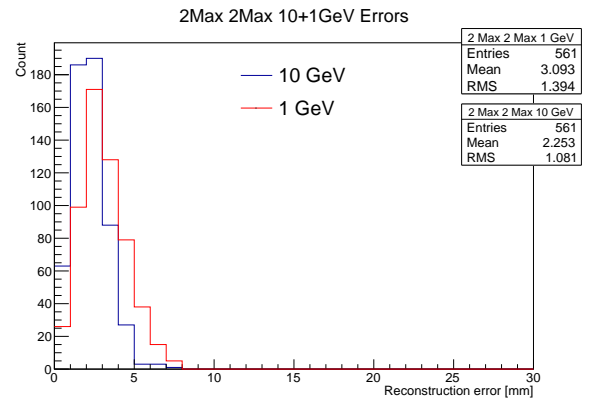
(c) Dos clústeres en un eje y un clúster con un máximo en el otro.



(d) Dos clústeres en ambos ejes.

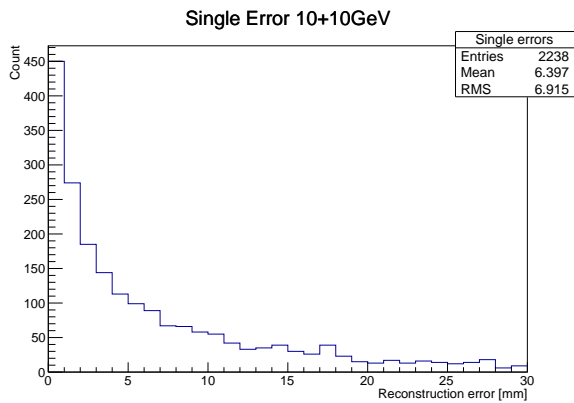


(e) Dos clústeres en un eje y un clúster con dos máximos en el otro.

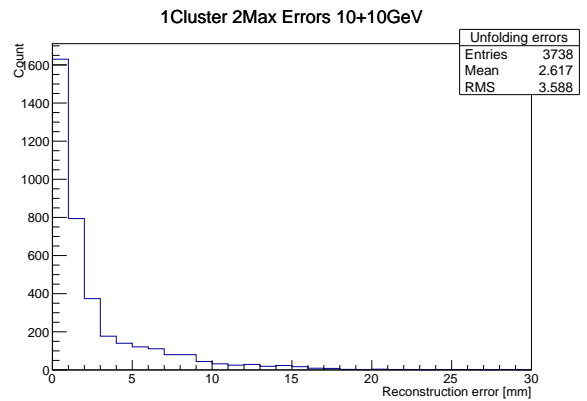


(f) Un clúster con dos máximos en ambos ejes.

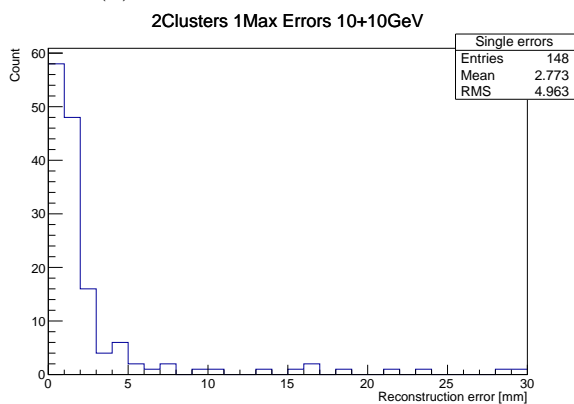
Figura B.1: Histogramas del error de reconstrucción $X_{rec} - X_{inc}$ para combinaciones de partículas γ de 10 y 1 [GeV] y diferentes casos del algoritmo de reconstrucción. En cada gráfico se muestra el error para cada una de las partículas por separado.



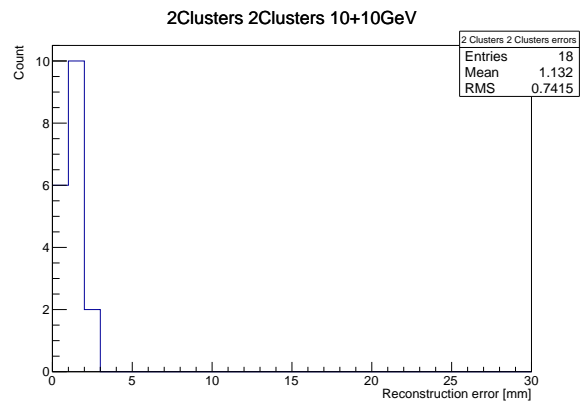
(a) Una partícula identificada.



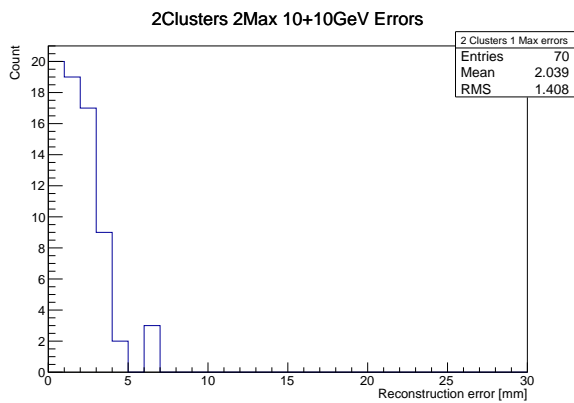
(b) Un clúster en un eje y un clúster con dos máximos en el otro.



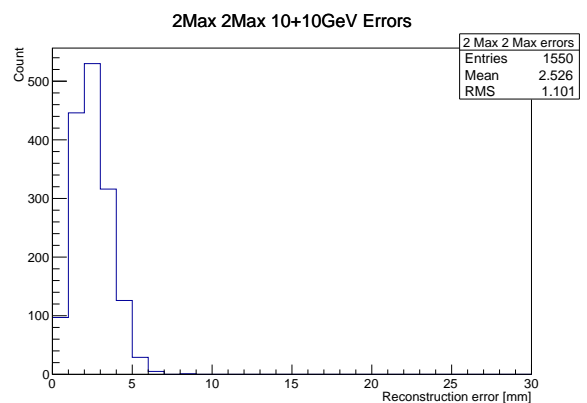
(c) Dos clústeres en un eje y un clúster con un máximo en el otro.



(d) Dos clústeres en ambos ejes.



(e) Dos clústeres en un eje y un clúster con dos máximos en el otro.



(f) Un clúster con dos máximos en ambos ejes.

Figura B.2: Histogramas del error de reconstrucción $X_{rec} - X_{inc}$ para dos partículas γ de 10 [GeV] y diferentes casos del algoritmo de reconstrucción.

B.2. Reconstrucción en el eje Y

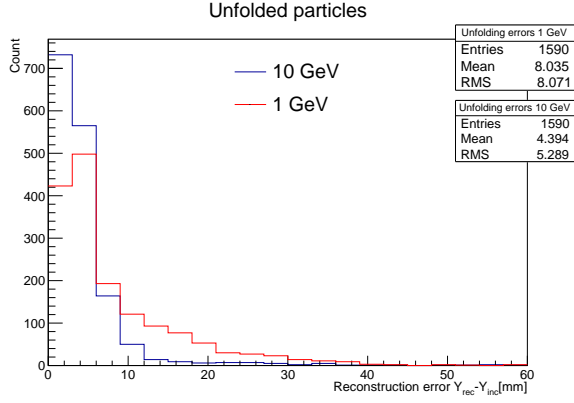
En la Figura B.3 se presentan los errores de reconstrucción para el eje Y para las combinaciones de partículas gamma de 10+1, 10+5, 10+10 y 10+15 [GeV]. Los errores son presentados para los caso en que las partículas deban ser separadas mediante el algoritmo de unfolding y para cuando las partículas se presentan en dos clústeres totalmente separados.

En base a los resultados para el eje Y se puede observar que:

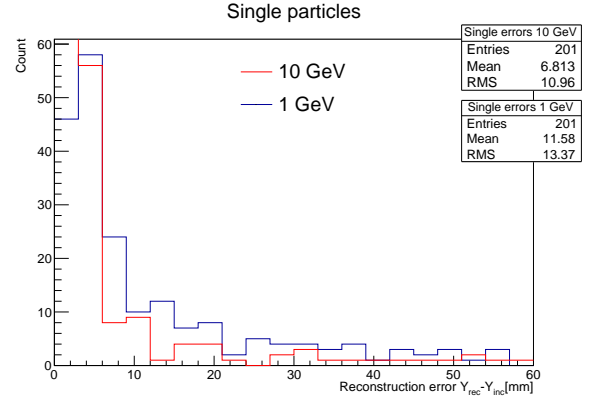
- Los errores son mayores para el eje Y que para el eje X. Esto se explica en parte por la mayor varianza del modelo presentado en 5.2.2. Esta varianza representa los distintos efectos electrónicos y físicos de la propagación de la lluvia por las fibras. Una mayor varianza en las distribuciones normales del modelo provoca lluvias más anchas y por lo tanto más difíciles de reconstruir. Además las lluvias se cubren entre ellas, haciendo parecer que sólo una partícula fue observada.
- Algunos peaks anormales son observados en las colas de la distribución de errores, en especial para el caso 10+10 [GeV]. Esto se debe al funcionamiento del algoritmo de resolución de ambigüedades. En los casos en que la energía depositada por ambas partículas sea parecida (caso más probable para la combinación 10+10) es posible que las posiciones sean emparejadas incorrectamente, produciendo un gran error en uno de los ejes. Debido al funcionamiento del algoritmo, el eje Y es el que se empareja a la reconstrucciones obtenidas en el eje X, por lo que se producen estos casos en los que el error es muy mayor a la media. Para solucionar esto es necesario trabajar más en el algoritmo de resolución de ambigüedades.

B.3. Reconstrucción del conteo de fotoelectrones

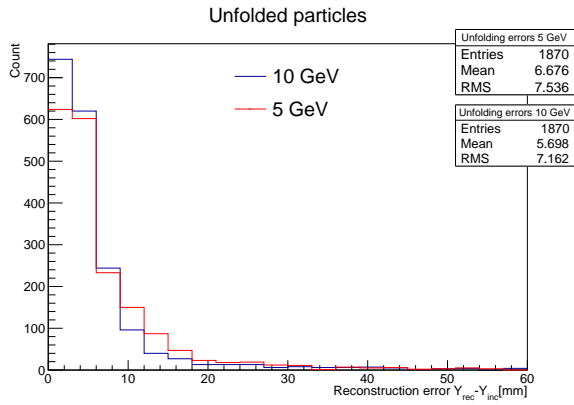
La reconstrucción de energía es un trabajo conjunto de toda la cadena de detección. El detector preshower mide el conteo de fotoelectrones que se comporta como una distribución de Poisson con media relacionada a la energía detectada. Esta medición es influenciada por otros factores, como la ganancia en el sistema de lectura, efectos ópticos en las fibras, pérdida de energía por los bordes de la matriz, entre otros. Es responsabilidad del detector preshower medir sólo parte de la energía de la partícula incidente, el resto de la energía será



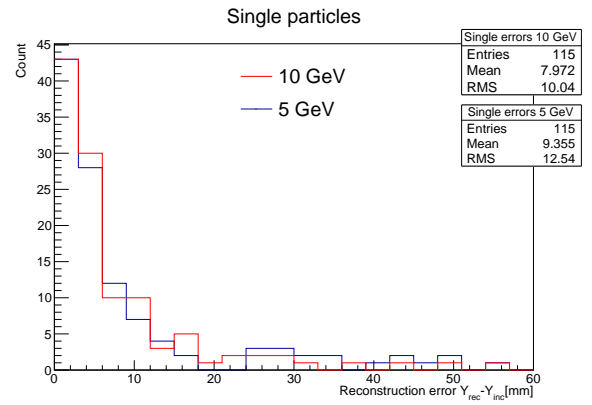
(a) 10 + 1 [GeV].



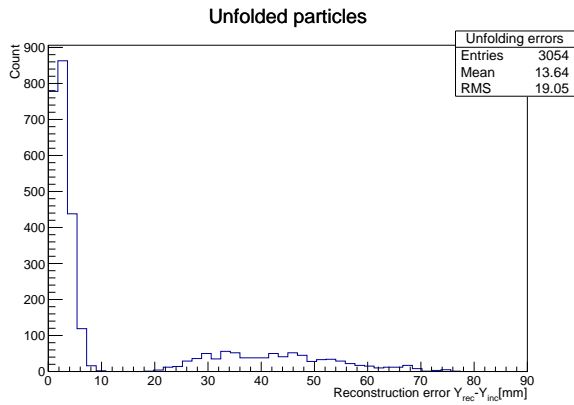
(b) 10 + 1 [GeV].



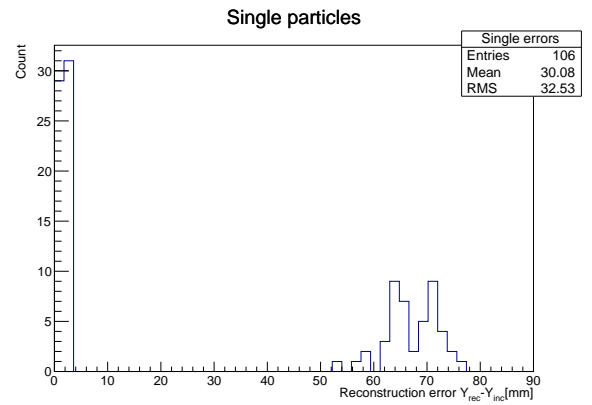
(c) 10 + 5 [GeV].



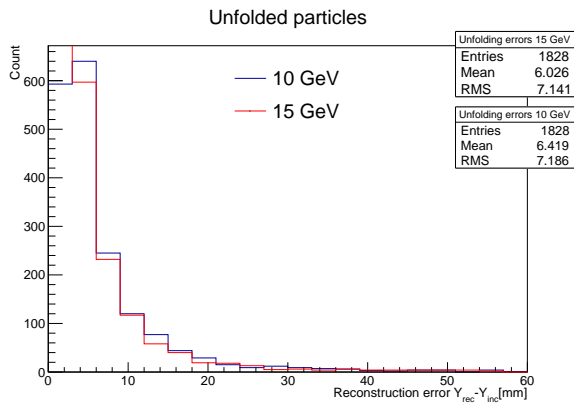
(d) 10 + 5 [GeV].



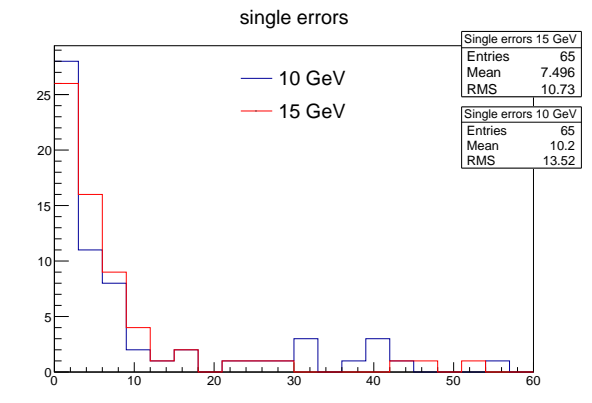
(e) 10 + 10 [GeV].



(f) 10 + 10 [GeV].



(g) 10 + 15 [GeV].



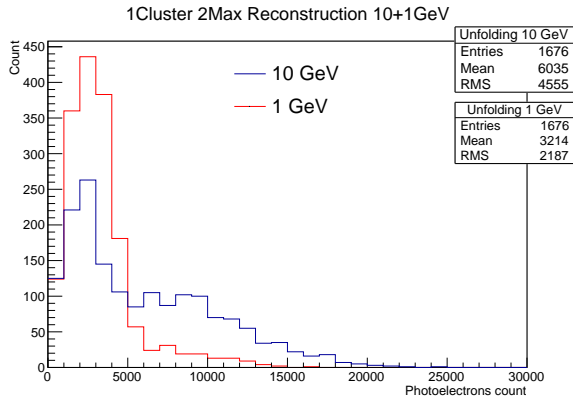
(h) 10 + 15 [GeV].

Figura B.3: Histogramas del error de reconstrucción $Y_{rec} - Y_{inc}$ para varias combinaciones de partículas γ . En cada gráfico se muestra el error para cada una de las partículas por separado (en el caso de 10+10 [GeV] sólo se muestra el error para 10 [GeV]).

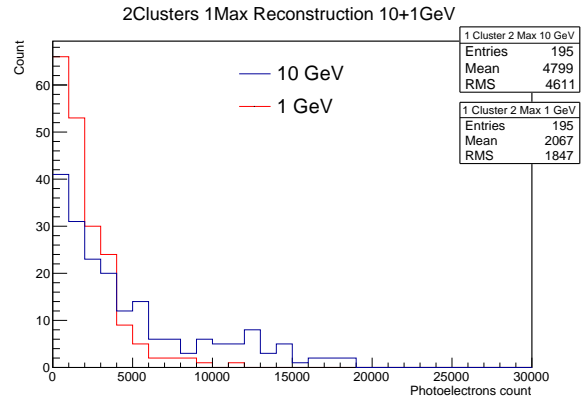
medida en detectores especializados en esta tarea. Además, es útil observar las distribuciones de fotoelectrones obtenidas por el algoritmo de reconstrucción para encontrar cualquier comportamiento anormal que pudiese estar indicando un mal funcionamiento del algoritmo. En la Figura B.4 y B.5 se observa el conteo de fotoelectrones para cada una de las partículas en las combinaciones 10+1 y 10+10 [GeV] y para cada uno de los posibles casos del algoritmo.

De los resultados de la reconstrucción de energía se observa que:

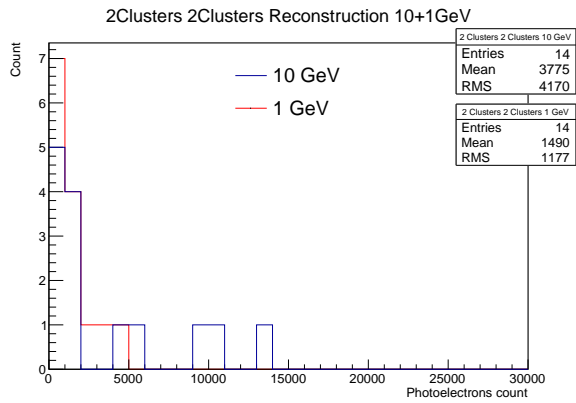
- Las partículas de baja energía depositan menos fotoelectrones que las partículas de mayor energía (observar el caso 10+1). Algo esperado y que indica que el algoritmo no está realizando una reconstrucción errónea del conteo de fotoelectrones.
- Para los casos en que el algoritmo de unfolding debe ser implementado existe una pequeña diferencia en la media del conteo de fotoelectrones respecto a cuando no se usa el algoritmo. Esto se debe a que al realizar la reconstrucción, las energías que están bajo el umbral de ruido y que son descartadas en el caso de clústeres separados, podrían aportar en el caso de clústeres solapados, ya que la suma del solapamiento de las lluvias supera el umbral de ruido.



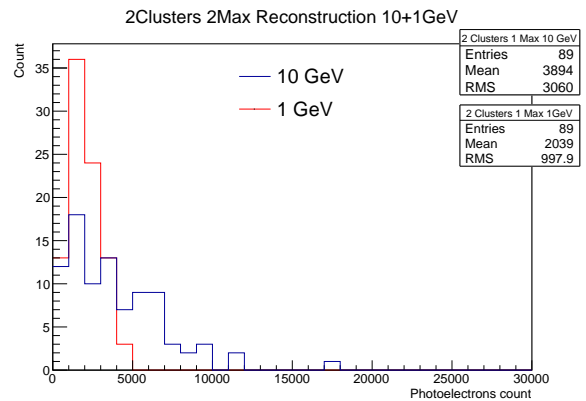
(a) Un clúster en un eje y un clúster con dos máximos en el otro.



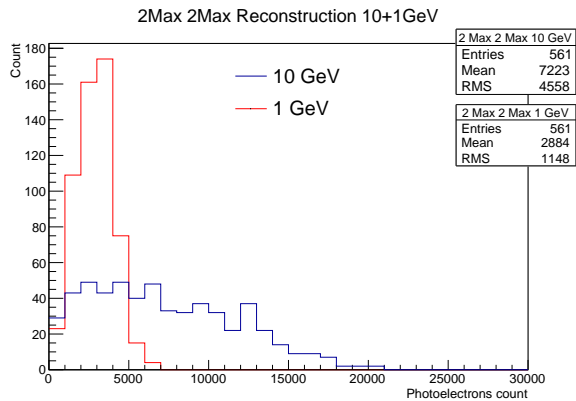
(b) Dos clústeres en un eje y un clúster con un máximo en el otro.



(c) Dos clústeres en ambos ejes.

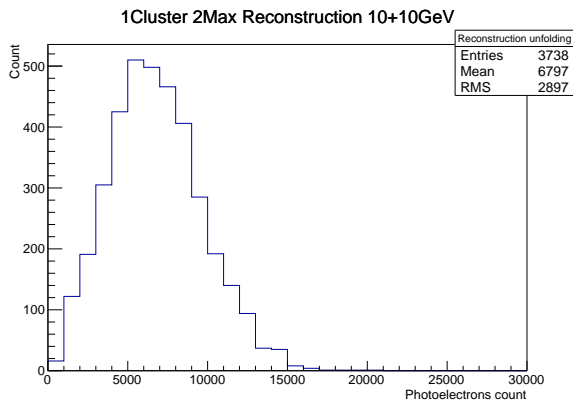


(d) Dos clústeres en un eje y un clúster con dos máximos en el otro.

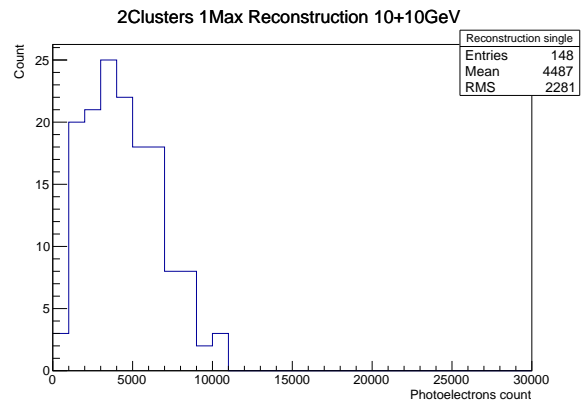


(e) Un clúster con dos máximos en ambos ejes.

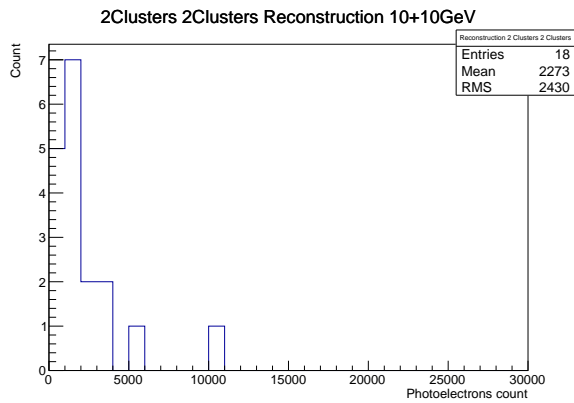
Figura B.4: Histogramas de reconstrucción del conteo de fotoelectrones para combinaciones de partículas γ de 10 y 1 [GeV] y diferentes casos del algoritmo de reconstrucción. En cada gráfico se muestra la reconstrucción para cada una de las partículas por separado.



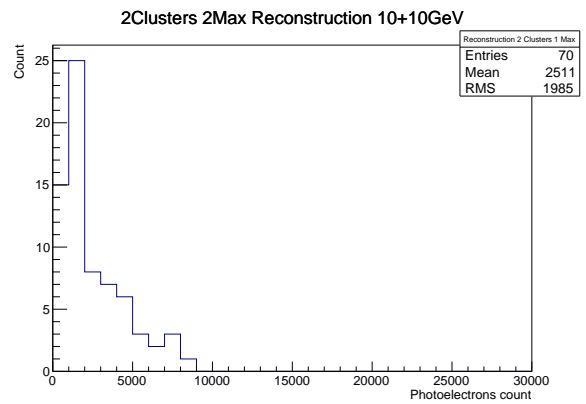
(a) Un clúster en un eje y un clúster con dos máximos en el otro.



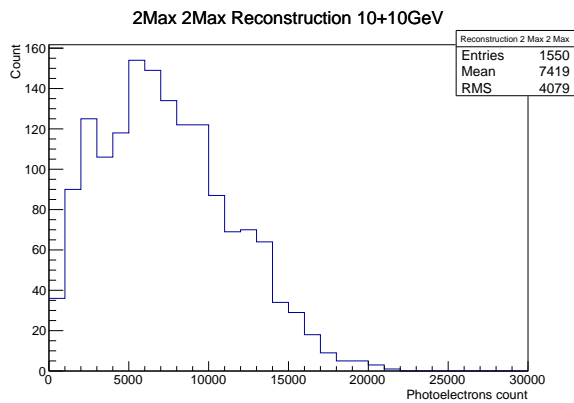
(b) Dos clústeres en un eje y un clúster con un máximo en el otro.



(c) Dos clústeres en ambos ejes.



(d) Dos clústeres en un eje y un clúster con dos máximos en el otro.



(e) Un clúster con dos máximos en ambos ejes.

Figura B.5: Histogramas de reconstrucción del conteo de fotoelectrones para combinaciones de partículas γ de 10 [GeV] y diferentes casos del algoritmo de reconstrucción.

Bibliografía

- [1] H. Davies, H. Bethe, and L. Maximon, “Theory of bremsstrahlung and pair production. ii. integral cross section for pair production,” *Physical review*, vol. 93, no. 4, p. 788, 1954.
- [2] S. Kuleshov and W. Brooks, “A pre-shower detector for forward electromagnetic calorimeters.” https://wiki.bnl.gov/conferences/images/b/bd/RD2012-13_Brooks_preshower_final.pdf, 2012.
- [3] T. A.-P. Collaboration, “Ax-pet - a demonstrator for an axial pet camera with wls strips and g-apd readout.” <https://twiki.cern.ch/twiki/bin/view/AXIALPET>, 2012.
- [4] M. Heller, “Ax-pet:a novel pet concept with g-apd readout.” <http://ndip.in2p3.fr/ndip11/AGENDA/AGENDA-by-DAY/Presentations/3Wednesday/AM/ID120-HELLER.pdf>, 2012.
- [5] J. Huber, W. Moses, M. Andreaco, and O. Petterson, “A lso scintillator array for a pet detector module with depth of interaction measurement,” in *Nuclear Science Symposium Conference Record, 2000 IEEE*, vol. 2, pp. 14–46, IEEE, 2000.
- [6] E. Zambrano, “Detector de pre-shower para calorímetros de lluvia electromagnética,” 2015. Universidad Técnica Federico Santa María, Departamento de Electrónica. Tesis de grado.
- [7] T. Speer, W. Adam, R. Frühwirth, A. Strandlie, T. Todorov, and M. Winkler, “Track reconstruction in the cms tracker,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 559, no. 1, pp. 143–147, 2006.
- [8] A. collaboration *et al.*, “Technical design report of the photon spectrometer (phos),” *CERN/LHCC*, vol. 99, no. 4, 1999.
- [9] W. Lampl, P. Loch, S. Menke, S. Rajagopalan, S. Laplace, G. Unal, H. Ma, S. Snyder, D. Lelas, and D. Rousseau, “Calorimeter clustering algorithms,” tech. rep., 2008.
- [10] G. Aad, B. Abbott, J. Abdallah, A. Abdelalim, A. Abdesselam, O. Abdinov, B. Abi, M. Abolins, H. Abramowicz, H. Abreu, *et al.*, “Electron performance measurements

- with the atlas detector using the 2010 lhc proton-proton collision data,” *The European Physical Journal C*, vol. 72, no. 3, pp. 1–46, 2012.
- [11] C. A.-B. S. Snyder et al, “Em calorimeter calibration and performance,” 2008.
- [12] A. Kalinowski, A. Collaboration, *et al.*, “Tau lepton reconstruction and identification with the atlas detector at the lhc,” *Nuclear Physics B-Proceedings Supplements*, vol. 189, pp. 305–310, 2009.
- [13] A. Collaboration *et al.*, “Topological cell clustering in the atlas calorimeters and its performance in lhc run 1,” *arXiv preprint arXiv:1603.02934*, 2016.
- [14] J. Pinfold, J. Soukup, J. Archambault, C. Cojocar, M. Khakzad, G. Oakham, M. Schram, M. Vincter, V. Datskov, V. Drobin, *et al.*, “Evaluation of the local hadronic calibration with combined beam-test data for the endcap and forward calorimeters of atlas in the pseudorapidity region $2.5 < |\eta| < 4.0$,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 693, pp. 74–97, 2012.
- [15] T. Barillari, A. Jantsch, P. Stavina, A. Kiryunin, T. Carli, K. Grahn, P. Giovannini, E. Rauter, G. Pospelov, P. Strízenec, *et al.*, “Local hadronic calibration,” tech. rep., 2008.
- [16] C. Cojocar, J. Pinfold, J. Soukup, M. Vincter, V. Datskov, A. Fedorov, S. Golubykh, N. Javadov, V. Kalinnikov, S. Kakurin, *et al.*, “Hadronic calibration of the atlas liquid argon end-cap calorimeter in the pseudorapidity region $1.6 < |\eta| < 1.8$ in beam tests,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 531, no. 3, pp. 481–514, 2004.
- [17] J. C. Bezdek, R. Ehrlich, and W. Full, “Fcm: The fuzzy c-means clustering algorithm,” *Computers & Geosciences*, vol. 10, no. 2, pp. 191–203, 1984.
- [18] R. P. Sandhir, S. Muhuri, and T. K. Nayak, “Dynamic fuzzy c-means (dfcm) clustering and its application to calorimetric data reconstruction in high-energy physics,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 681, pp. 34–43, 2012.
- [19] S. K. Pal, S. Chattopadhyay, and Y. Viyogi, “Application of fuzzy-based pattern recognition techniques for cluster finding in a preshower detector in high energy heavy ion experiments,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 626, pp. 105–113, 2011.
- [20] F. Simon, “Energy reconstruction of hadron showers in the calice calorimeters,” in *Nuclear Science Symposium Conference Record (NSS/MIC), 2009 IEEE*, pp. 2292–2295, IEEE, 2009.

- [21] M. Mariscotti, “A method for automatic identification of peaks in the presence of background and its application to spectrum analysis,” *Nuclear Instruments and Methods*, vol. 50, no. 2, pp. 309–320, 1967.
- [22] M. Morháč, J. Kliman, V. Matoušek, M. Veselský, and I. Turzo, “Identification of peaks in multidimensional coincidence γ -ray spectra,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 443, no. 1, pp. 108–125, 2000.
- [23] F. Berger, D. Bock, G. Clewing, L. Dragon, R. Glasow, G. Hölker, K. Kampert, T. Peitzmann, M. Purschke, B. Roters, *et al.*, “Particle identification in modular electromagnetic calorimeters,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 321, no. 1, pp. 152–164, 1992.
- [24] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, H. Voss, M. Backes, T. Carli, O. Cohen, A. Christov, *et al.*, “Tmva-toolkit for multivariate data analysis,” *arXiv preprint physics/0703039*, 2007.
- [25] F. Palma, C. Collaboration, *et al.*, “Simulation studies of the expected proton rejection capabilities of calet,” in *34th International Cosmic Ray Conference (ICRC2015)*, vol. 34, 2015.
- [26] R. Bellotti, M. Boezio, and F. Volpe, “A particle classification system for the pamela calorimeter,” *Astroparticle Physics*, vol. 22, no. 5, pp. 431–438, 2005.
- [27] G. Akopdjanov, A. Inyakin, V. Kachanov, R. Krasnokutsky, A. Lednev, Y. V. Mikhailov, Y. D. Prokoshkin, E. Razuvaev, and R. Shuvalov, “Determination of photon coordinates in a hodoscope cherenkov spectrometer,” *Nuclear Instruments and Methods*, vol. 140, no. 3, pp. 441–445, 1977.
- [28] Y. A. Bashmakov and M. S. Korbut, “Photon position and energy reconstruction in a cherenkov hodoscope,” in *AIP Conference Proceedings*, pp. 562–570, IOP INSTITUTE OF PHYSICS PUBLISHING LTD, 1998.
- [29] L. Bugge, “On the determination of shower central positions from lateral samplings,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 242, no. 2, pp. 228–236, 1986.
- [30] T. Awes, F. Obenshain, F. Plasil, S. Saini, S. Sorensen, and G. Young, “A simple method of shower localization and identification in laterally segmented calorimeters,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 311, no. 1, pp. 130–138, 1992.
- [31] J. Colas, L. Di Ciaccio, M. El Kacimi, O. Gaumer, M. Gouanère, D. Goujdami, R. Lafaye, C. Le Maner, L. Neukermans, P. Perrodo, *et al.*, “Position resolution and

- particle identification with the atlas em calorimeter,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 550, no. 1, pp. 96–115, 2005.
- [32] M. Amarian, G. Asryan, K. Beard, W. Brooks, V. Burkert, T. Carstens, A. Coleman, R. Demirchyan, Y. Efremenko, H. Egiyan, *et al.*, “The clas forward electromagnetic calorimeter,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 460, no. 2, pp. 239–265, 2001.
- [33] D. A. Bremer, “Measurements and simulations on position dependencies in the response of single pwo crystals and a prototype for the panda emc,” 2014.
- [34] S. Agostinelli, J. Allison, K. a. Amako, J. Apostolakis, H. Araujo, P. Arce, M. Asai, D. Axen, S. Banerjee, G. Barrand, *et al.*, “Geant4 - a simulation toolkit,” *Nuclear instruments and methods in physics research section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 506, no. 3, pp. 250–303, 2003.
- [35] I. Antcheva, M. Ballintijn, B. Bellenot, M. Biskup, R. Brun, N. Buncic, P. Canal, D. Casadei, O. Couet, V. Fine, *et al.*, “Root-a c++ framework for petabyte data storage, statistical analysis and visualization,” *Computer Physics Communications*, vol. 180, no. 12, pp. 2499–2512, 2009.
- [36] Z. Silagadze, “A new algorithm for automatic photopeak searches,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 376, no. 3, pp. 451–454, 1996.
- [37] K. Cranmer, J. Pavez, and G. Louppe, “Approximating likelihood ratios with calibrated discriminative classifiers,” *arXiv preprint arXiv:1506.02169*, 2015.
- [38] P. Baldi, K. Cranmer, T. Faucett, P. Sadowski, and D. Whiteson, “Parameterized machine learning for high-energy physics,” *arXiv preprint arXiv:1601.07913*, 2016.
- [39] Y. Freund and R. E. Schapire, “A desicion-theoretic generalization of on-line learning and an application to boosting,” in *European conference on computational learning theory*, pp. 23–37, Springer, 1995.
- [40] A. F. Soria and E. R. Martínez, *Física de partículas y de astropartículas*. Universitat de València, 2011.
- [41] W. R. Leo, *Techniques for nuclear and particle physics experiments: a how-to approach*. Springer Science & Business Media, 2012.
- [42] D. Green, *The physics of particle detectors*, vol. 12. Cambridge University Press, 2000.