



República Bolivariana de Venezuela
Facultad de Ciencias
Departamento de Física
Centro de Física Fundamental

Reproducción computacional, optimización y discusión del análisis del experimento ATLAS sobre la producción de ZZ en colisiones protón-protón a $\sqrt{s} = 13$ TeV usando datos abiertos

Por:

Br. Mildred Alexandra Arias Yanez

Tutor: Dr. Arturo Sanchez

Tutor Académico: Dr. Nelson Pantoja

Trabajo de grado presentado como requisito parcial para optar al grado académico de
Licenciado en Física

Marzo, 2022

ACTA DE TRABAJO ESPECIAL DE GRADO

En Mérida al primer día del mes de febrero del año dos mil veintidós, a las 2:30 p.m., se reunieron los Profesores NELSON PANTOJA, ALBERTO PATIÑO y MISAELO ROSALES, miembros del Jurado designado por el Departamento de Física y aprobado por el Consejo de la Facultad de Ciencias y el tutor externo Dr. ARTURO SÁNCHEZ (INAIT.AI, LAUSANNE, SUIZA), para revisar el Trabajo Especial de Grado que lleva como título:

“Reproducción computacional, optimización y discusión del análisis del experimento ATLAS sobre la producción de ZZ en colisiones protón-protón a $\sqrt{s} = 13 \text{ TeV}$ usando datos abiertos.”.

Presentado por el Bachiller MILDRED ALEXANDRA ARIAS YANEZ, titular de la Cédula de Identidad N° V- 26.209.311, para optar por el título de:

LICENCIADO EN FÍSICA

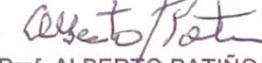
En la Facultad de Ciencias de la Universidad de Los Andes. Acto seguido, se procedió a oír la exposición sobre el tema, arriba mencionado, que realizó el bachiller MILDRED ALEXANDRA ARIAS YANEZ. Después del correspondiente interrogatorio, el Jurado procedió a deliberar sobre el contenido de la misma.

Finalmente, el Jurado lo declaró **APROBADO**, con la calificación de **VEINTE (20) PUNTOS**. Recomendado para su publicación.

En fe de lo cual firman:


Prof. NELSON PANTOJA
TUTOR ACADÉMICO


Prof. MISAELO ROSALES
JURADO


Prof. ALBERTO PATIÑO
JURADO



COORDINACIÓN DOCENTE

Departamento de Física

Núcleo "Pedro Rincón Gutiérrez", Edificio "A", La Hechicera, Mérida 5101 - República Bolivariana de Venezuela
Teléfono: (58 - 274) 240 1318 / 240 1319 / 240 1320 / 240 1321 - Fax: 240 1318 - Web: www.ciens.ula.ve _ Corre_cdfisica@ula.ve

If I could remember the names of all these particles, I'd be a botanist. .
Enrico Fermi

Agradecimientos

Primero que todo quisiera agradecer a quien ha sido mi ejemplo de mujer y de persona en general a seguir, quien me ha enseñado como echar para adelante y luchar en la vida sin importar nada, quien me ha enseñado sus valores, su sentido de responsabilidad y quien todos los días del mundo sin importar la situación me da todo el amor del mundo de la forma más dulce y sincera. Te amo mamá.

A mi otra mamá, Marisol que a pesar de la distancia siempre estuvo allí enviando sus buenas energías y ayudándome cuando mas lo necesite.

Quisiera agradecer a mis amigos durante la carrera Beatriz y Mario, quienes fueron de apoyo, ante todo, me ayudaron a crecer durante estos años, y compartimos muchísimos momentos que jamás olvidare, gracias simplemente por haber estado allí.

A mis tutores: Arturo, gracias primero por haberme motivado a seguir mis sueños, por esa primera conversación sobre la importancia de la ciencia y a su vez de la física. Por haberme enseñado el mejor camino a seguir en mi vida, por todo el tiempo dedicado, sé que no ha sido fácil, pero ha sido una experiencia que, aunque pudiera no cambiaría. Al profesor Nelson Pantoja, gracias por todo lo que me ha enseñado, de física y de la vida, gracias por todas las horas de conversaciones acerca de muchísimos temas de cualquier tópico que hemos tenido, gracias por abrirme las puertas de su oficina y tratar de ayudarme cuando estuve en sus manos, es un ejemplo a seguir tanto en lo profesional como en lo personal profesor.

Quisiera agradecer especialmente al grupo de CEVALE2VE, quien presta una gran ayuda para muchos estudiantes de conocer el mundo de la física de partículas, es algo asombroso lo que hacen, cuando la gente ama lo que hace, se nota.

En especial gracias a la Facultad de Ciencias de la ULA, la cual desde el primer día en que entre me enamoro y me enseñó a crecer a nivel personal y profesional.

Quisiera agradecer también a algunos profesores que me inspiraron para poder estar en el camino que estoy ahorita: Al profesor Nelson Pantoja nuevamente, que tal vez no lo sepa, pero fue mi primer modelo a seguir en la escuela, y alguien que admire muchísimo y todavía lo sigo haciendo. Al profe Alberto Patiño, por aquellas conversaciones motivacionales e inspiradoras para seguir adelante y seguir enamorándome más de la física. Al profe Cayetano, que, aunque no fue profe de la ULA sino de la USB fue un placer conocerlo y aprender mucho sobre él.

¡Gracias!

Reproducción computacional y discusión del análisis del experimento ATLAS sobre la producción de ZZ en colisiones protón-protón a $\sqrt{s} = 13$ TeV.

Br. Mildred Alexandra Arias Yanez

Proyecto de Grado — Licenciatura en Física, 118 páginas

Resumen

En este trabajo se genera la reproducción computacional, además de la optimización y discusión de uno de los análisis del experimento ATLAS sobre la producción de dos bosones Z (ZZ) en colisiones protón-protón a una energía de centro de masa de 13 TeV en el LHC en el CERN, utilizando datos abiertos publicados por la Colaboración ATLAS en el CERN. En una primera fase se procede a analizar la reproducción computacional utilizando notebooks de Jupyter, y el framework de análisis ROOT, creado en el CERN. Una vez analizados dicho conjunto de datos abiertos, se discuten los resultados obtenidos de la reproducción computacional, para finalmente optimizar los resultados en el canal de desintegración $ZZ \rightarrow 4l$ a $\sqrt{s} = 13$ TeV usando diferentes técnicas estadísticas de ajustes de datos. Se estimaron los parámetros de la Función de Densidad de Probabilidad (PDF) ajustada, mediante el método de mínimos cuadrados, utilizando el software de minimización proveído por ROOT, llamado TMinuit. Con éste, se obtuvo el valor mínimo de la cantidad relativa al Mínimo Cuadrado o χ^2 , que fue utilizado como prueba estadística de la bondad de ajuste, t . La hipótesis nula planteada es aceptada ya que la distribución de Landau utilizada describe nuestros datos con un nivel de confianza del 95 %. La forma de χ^2 en función del valor más probable (MPV) sigue un perfil parabólico, cuyo valor mínimo coincide con el valor más probable del modelo de Landau planteado, el cual es de 213 GeV. Los métodos estadísticos de mínimos cuadrados y de máxima verosimilitud (o log-likelihood), implementados en ROOT para histogramas de estadística baja, se compararon, dejándose ver que el método de log-likelihood tiene la ventaja de tratar correctamente los contenedores de los histogramas con estadísticas bajas, ya que el parámetro de MPV está sesgado cuando se trabaja con el método de χ^2 con baja estadística. El producto final de este trabajo de tesis consiste no solo en este documento, sino en una serie de notebooks de Jupyter almacenamos en un repositorio para ser utilizados posteriormente por quien desee reproducir y mejorar lo realizado, continuando así con la premisa de acceso abierto y reproducibilidad planteada por el proyecto de ATLAS Open Data sobre el cual se basó toda esta propuesta.

Palabras clave: Reproducción computacional, experimento ATLAS, producción de ZZ, colisiones protón-protón a $\sqrt{s} = 13$ TeV.

Índice General

| | |
|---|-----|
| Agradecimientos | I |
| Resumen | II |
| Índice de Tablas | VI |
| Índice de Figuras | VII |
| Introducción | 1 |
| 1. Física de Partículas | 3 |
| 1.1. Partículas Fundamentales | 3 |
| 1.2. Construcción del Modelo Estándar | 5 |
| 1.2.1. Electrodinámica Cuántica | 6 |
| 1.2.2. Teoría Electrodébil | 6 |
| 1.2.3. Ruptura Espontánea de Simetría y Mecanismo de Higgs | 7 |
| 1.2.4. La Cromodinámica Cuántica | 8 |
| 1.3. Interacciones en Colisiones Protón-Protón | 9 |
| 2. Producción ZZ | 11 |
| 2.1. Producción ZZ en el Colisionador de Hadrones | 11 |
| 2.1.1. Modos de Desintegración de ZZ | 12 |
| 2.1.2. Definición de Sección Transversal | 12 |
| 2.2. Acoplamientos Anómalos de Triple Calibre | 13 |
| 2.2.1. Origen de los Acoplamientos Anómalos de Triple Calibre Neutros | 14 |
| 2.3. Resultados Experimentales Previos | 15 |
| 3. CERN, LHC y el Experimento ATLAS | 19 |
| 3.1. El Gran Colisionador de Hadrones | 19 |
| 3.2. Detectores de Partículas: El experimento ATLAS | 23 |
| 3.2.1. Sistema de Coordenadas en ATLAS | 24 |
| 3.2.2. Sistema de Imanes en ATLAS | 25 |
| 3.2.3. Detector Interno | 26 |
| 3.2.4. Calorímetros | 27 |
| 3.2.5. Espectrómetro de Muones | 29 |
| 3.3. El Trigger de ATLAS | 31 |
| 3.4. Generadores de Monte Carlo en la Simulación de Eventos en ATLAS | 32 |
| 3.5. Simulación del Detector | 33 |

| | |
|--|-----------|
| 4. Reconstrucción de Objetos Físicos en ATLAS | 34 |
| 4.1. Calidad de los datos | 34 |
| 4.2. Reconstrucción de Trazas por el ID | 35 |
| 4.3. Hallazgo de Vértices | 37 |
| 4.4. Reconstrucción e Identificación de Electrones | 38 |
| 4.4.1. Reconstrucción de Electrones | 38 |
| 4.4.2. Identificación de Electrones | 39 |
| 4.5. Muones | 43 |
| 4.5.1. Reconstrucción e Identificación de Muones | 43 |
| 5. Selección de Objetos y Eventos | 45 |
| 5.1. Selección de Electrones | 45 |
| 5.1.1. Selección de Electrones Centrales | 45 |
| 5.2. Selección de Muones | 47 |
| 5.2.1. Selección de Muones Centrales | 47 |
| 5.2.2. Selección de Muón Directo | 47 |
| 5.2.3. Selección de Muones Etiquetado por Calorímetro | 49 |
| 5.3. Selección de Trigger | 49 |
| 5.4. Selección de Eventos ZZ | 49 |
| 5.4.1. Definiciones de candidatos Z | 49 |
| 5.4.2. Requisitos de Selección de Eventos | 49 |
| 5.5. Incertidumbres Sistemáticas | 51 |
| 6. Herramientas, Datos y Análisis Computacional para la Producción del Dibosón ZZ | 52 |
| 6.1. ATLAS Open Data | 52 |
| 6.1.1. Datos Reales y Simulados | 52 |
| 6.2. Herramientas Computacionales | 54 |
| 6.2.1. ROOT | 54 |
| 6.2.2. Jupyter Notebooks | 55 |
| 6.2.3. SWAN | 55 |
| 6.2.4. Git y GitHub | 56 |
| 6.3. Cómo hacer ciencia Reproducible | 57 |
| 6.3.1. Organización de Ficheros | 58 |
| 6.4. Descripción de la Reproducción Computacional para la Producción del Dibosón ZZ | 60 |
| 7. Optimización y Resultados | 64 |
| 7.1. Descripción de la P.D.F. utilizada para dicho ajuste en ROOT | 64 |
| 7.2. El Método de Ajuste en ROOT | 65 |
| 7.2.1. El Método “TH1::Fit” | 65 |
| 7.3. Optimización | 67 |
| 7.3.1. Prueba de Bondad de Ajuste | 67 |
| 7.3.2. Cálculo del Mínimo χ^2 a través de TMinuit (Curva de χ^2) | 70 |
| 7.4. Comparación entre dichos Métodos Estadísticos | 70 |
| 7.5. Análisis de las Distribuciones Cinemáticas generadas en la Reproducción | 74 |
| 8. Conclusiones y Discusión | 79 |
| A. Resumen de Conceptos Teóricos | 82 |

| | |
|--|------------|
| B. Glosario | 85 |
| C. Marco Estadístico | 88 |
| C.1. Funciones de Densidad de Probabilidad | 88 |
| C.1.1. Caracterización de la Forma de una Función de Densidad de Probabilidad | 89 |
| C.1.2. La Distribución de Landau | 90 |
| C.2. Método de Mínimos Cuadrados | 90 |
| C.3. Método de Máxima Verosimilitud | 91 |
| C.4. Máxima Verosimilitud con Datos Agrupados | 92 |
| C.5. Pruebas Estadísticas | 93 |
| C.5.1. Pruebas de Hipótesis | 93 |
| C.5.2. Pruebas de Bondad de Ajuste | 94 |
| C.6. Intervalos de Confianza | 94 |
| C.6.1. Intervalos de Confianza utilizando la Función de Verosimilitud o χ^2 | 96 |
| D. ROOT: Marco de Análisis de Datos | 99 |
| D.1. Descripción de las PDFs en ROOT | 99 |
| D.2. El Método de Ajuste en ROOT | 99 |
| D.2.1. El Método “TH1::Fit” | 100 |
| D.2.2. El Panel de Ajuste | 101 |
| D.3. Clase TMinuit | 101 |
| E. Abreviaturas | 103 |
| Referencias | 106 |

Índice de Tablas

| | |
|--|----|
| 1.1. Propiedades de las partículas elementales del modelo estándar: Carga, masa(eV), espín y color. | 5 |
| 1.2. Interacciones Fundamentales | 6 |
| 3.1. Parámetros operativos del LHC. Se comparan los parámetros nominales de diseño [60] y los utilizados en la operación de 2011 y en la operación de 2012 [61] [62]. | 22 |
| 5.1. Requisitos de selección de electrones. | 46 |
| 5.2. Requisitos de selección de muones. | 48 |
| 6.1. Conjunto de Datos [111]. | 53 |
| C.1. Los valores del nivel de confianza para diferentes valores del cuantil de la gaussiana estándar Φ^{-1} : Para los intervalos centrales (izquierda) el cuantil $\Phi^{-1}(1 - \gamma/2)$ y el nivel de confianza $(1 - \gamma)$; para intervalos unilaterales (derecha) el cuantil $\Phi^{-1}(1 - \alpha)$ y el nivel de confianza $(1 - \alpha)$ [172]. | 97 |

Índice de Figuras

| | |
|---|----|
| 1.1. Tipos de hadrones. A la izquierda mesones conformados por un par quark-antiquark. A la derecha bariones constituidos por tres quarks. | 4 |
| 1.2. Diagrama de Feynman del intercambio de un fotón mediante la interacción entre dos electrones.[19] | 7 |
| 1.3. Ilustración de una colisión protón-protón [25]. | 10 |
| 2.1. Diagramas de Feynman de menor orden para la producción de ZZ. Los diagramas de (a) canal t y (b) canal u contribuyen a la sección transversal de producción de ZZ, mientras que el diagrama del (c) canal s no está presente en el SM, ya que contiene un vértice neutral ZZZ o ZZ γ . En (d), (e) y (f) se muestran ejemplos de contribuciones de un bucle a la producción de ZZ mediante pares de gluones [30]. | 12 |
| 2.2. Decaimiento de Z en un Lepton-Antilepton. | 13 |
| 2.3. Producción de un vértice ZZV a través de un bucle de fermión pesado [18]. | 15 |
| 2.4. Mediciones de la sección transversal de producción ZZ en colisiones e^+e^- en fun- ción de la energía del centro de masa \sqrt{s} . Las medidas son una combinación de medidas de los cuatro experimentos LEP [47]. | 16 |
| 2.5. Curvas de probabilidad logarítmica negativas para acoplamientos de nTGCs de los cuatro experimentos LEP (bandas coloreadas) y su combinación (banda negra). No se aplica ningún factor de forma [42]. | 17 |
| 3.1. Gráfico de la luminosidad integrada obtenida en función del tiempo para 2011 - 2018 de operación con haces estables de protones de alta energía [59]. | 20 |
| 3.2. Luminosidad integrada en función del tiempo (para distintos años) que fue pro- porcionada por el LHC (verde) y que registró ATLAS(amarillo) [59]. | 21 |
| 3.3. (a) Cadena de aceleración de las partículas antes de entrar al LHC [63]. (b) Ex- perimentos del LHC [64]. | 22 |
| 3.4. Esquema del detector ATLAS y los subdetectores que lo componen [65]. | 23 |
| 3.5. Esquema de la detección de las partículas producidas en una colisión de protones mientras pasan a través de varios subdetectores de ATLAS [66]. | 24 |
| 3.6. Sistema de coordenadas del experimento ATLAS. Donde θ es el ángulo polar y describe el ángulo de una partícula con respecto al eje del haz. ϕ es el ángulo azimutal, este ángulo se mide alrededor del eje z o la dirección del haz. | 25 |
| 3.7. El sistema magnético del detector ATLAS: el cilindro interior es el solenoide su- perconductor, mientras que las partes externas son las bobinas del toroide [67]. | 26 |
| 3.8. Estructura del Detector interno (ID) [71]. | 27 |

| | |
|--|----|
| 3.9. Detector Interno de ATLAS. Se muestran los tres subdetectores (el Detector de píxeles, el Semiconductor Tracker y el Transition Radiation Tracker). La trayectoria de las partículas cruza sucesivamente 3 capas de píxeles de silicio, cuatro capas de micro bandas de silicio y aproximadamente 36 tubos del detector de radiación de transición [72]. | 28 |
| 3.10. Diagrama del calorímetro de argón líquido ATLAS, que muestra la estructura del acordeón y la diferente granularidad en las diferentes capas. Diagrama tomado de [67]. | 29 |
| 3.11. Estructura del Calorímetro Electromagnético y Hadrónico [73]. | 30 |
| 3.12. Se puede observar 3 capas concéntricas, cada capa consiste en cámaras MDT (Monitored Drift Tube). Las capas medias y externas además están equipadas con cámaras RPC (Resistive Plate Chambers) [74]. Las cámaras MDT proveen medidas de precisión para determinar el momentum y se encuentran en la región $ \eta < 2$. Las cámaras RPC proveen el trigger para el sistema central de barril [75]. | 30 |
| 3.13. Estructura principal del sistema de disparo ATLAS: se compone de tres niveles, cada uno de los cuales mejora la medición de los niveles anteriores combinando también información de diferentes subdetectores [67]. | 32 |
| | |
| 4.1. La trayectoria puede ser paramétrizada en un punto dado por medio de su posición, momento transverso y carga. La posición en cualquier otro punto puede ser calculado si el campo magnético y el material del detector es conocido. Tomado de [67] | 36 |
| 4.2. Eficiencia de identificación de electrones en 2011 (marcadores abiertos) y 2012 (marcadores sólidos) en función del número de vértices primarios reconstruidos en el evento. Los círculos azules muestran la eficiencia para la selección “Loose ++”, los triángulos rojos para “Medium ++” y los cuadrados verdes para “Tight ++”. Figura de [105]. | 41 |
| 4.3. Eficiencia de los requisitos de identificación de “Loose ++” en función de la energía transversal del clúster. Los puntos sólidos indican mediciones basadas en datos, mientras que los puntos abiertos indican predicciones de Monte Carlo. Los diferentes marcadores indican el método utilizado para medir la eficiencia: los resultados de la etiqueta $Z \rightarrow ee$ y la sonda se muestran como círculos negros, los resultados de $J/\psi \rightarrow ee$ como cuadrados azules y los resultados de $W \rightarrow e\nu$ como triángulos rojos. Figura de [106]. | 42 |
| 4.4. Eficiencia de reconstrucción de muones en 2011 en función de (a) la pseudorapidez y (b) el momento transversal del muón para los muones reconstruidos como Combinados o Etiquetados por Segmentos utilizando el algoritmo STACO. Los puntos negros sólidos muestran la eficiencia observada en los datos y los círculos rojos abiertos muestran la eficiencia predicha por la simulación de Monte Carlo. En la figura (a) también se muestran las eficiencias de los muones marcados con calorímetro para $ \eta < 0,1$ (triángulos azules sólidos para datos y triángulos verdes abiertos para Monte Carlo). Tomado de [108]. | 44 |
| | |
| 5.1. Corte de preselección para disparador de Electrones/Muones [109]. | 50 |
| 5.2. Cada par de leptones debe tener una masa invariante en el rango de $66 < m_{ll} < 116$ GeV [109]. | 51 |

| | |
|---|----|
| 6.1. Esquema de un flujo de trabajo reproducible. En primer lugar, los datos se recogen según un protocolo bien diseñado, se documentan con metadatos, se someten a un control de calidad (idealmente de manera automática, esto es, mediante funciones de código), y se almacenan en un repositorio de datos en la nube. Después procedemos al análisis, siempre utilizando “scripts” para manipular los datos, y creando funciones que pueden almacenarse en un paquete (para facilitar su documentación y posterior reutilización). El análisis propiamente dicho se haría mediante documentos en Jupyter Notebooks que integran texto, código y resultados (tablas y figuras). Estos documentos pueden convertirse fácilmente en presentaciones, páginas web, o artículos científicos plenamente reproducibles. | 53 |
| 6.2. ATLAS Open Data provee documentación, datos reales y simulados, software y máquinas virtuales para poder recrear análisis en HEP y aprender sobre ellos. Además de incentivar a comprender más sobre programación computacional y estadística [112]. | 54 |
| 6.3. Herramienta para el procesamiento de datos: ROOT [114]. | 55 |
| 6.4. Interfaz de Jupyter Notebook [116]. | 56 |
| 6.5. Interfaz de SWAN [117]. | 57 |
| 6.6. Interfaz de GitHub [120]. | 57 |
| 6.7. Herramientas para el Análisis de datos. | 58 |
| 6.8. Esquema de la Reproducción Computacional para el análisis de la producción ZZ. | 61 |
| 6.9. Directarios creados “Output_ZZDiBosonAnalysis”[127] e “histograms”[128], para almacenar la salida del código de dicho análisis y los gráficos de salida. | 62 |
| 6.10. a) Script Bash “run.sh” [134]. b) Llamado del Script Bash “run.sh” en el notebook [134]. | 63 |
| 7.1. Distribución de Landau con parámetros μ y c [142]. | 65 |
| 7.2. Panel de Ajuste en ROOT [144]. | 66 |
| 7.3. Histograma de la Masa del sistema de cuatro leptones [148]. | 68 |
| 7.4. Zonas donde se acepta o se rechaza la hipótesis nula, dependiendo del valor de χ^2_{prueba} | 68 |
| 7.5. Ajuste de la distribución de landau a los datos de la masa invariante del sistema de cuatro leptones [148]. | 69 |
| 7.6. a), b) χ^2 en función de MPV (el valor más probable. b) Interpolación parabólica en el mínimo. Donde χ^2 aumenta en 1 unidad para $\Delta MPV \sim \pm 5,5$ [148]. | 71 |
| 7.7. Distribución de landau con 96 eventos [150]. | 72 |
| 7.8. Comparación entre dichos Métodos Estadísticos con baja estadística [150]. | 74 |
| 7.9. Comparación entre los Métodos Estadísticos con alta estadística [150]. | 75 |
| 7.10. a) Las distribuciones de los leptones para el PDG ID muestran que se están simulando en este caso electrones y muones [153] y b) Leptones de carga opuesta [154]. | 75 |
| 7.11. Comparación entre datos y predicción MC obtenidas de la reproducción del marco de análisis en el canal de decaimiento $ZZ \rightarrow 4l$. a) Ángulo azimutal φ [155] y b) Pseudorapidez η de cada lepton seleccionado [156]. | 76 |
| 7.12. Comparación entre datos y predicción MC obtenidas de la reproducción del marco de análisis en el canal de decaimiento $ZZ \rightarrow 4l$. a) Momento transverso p_T [157] y b) Energía E de cada lepton seleccionado [158]. | 77 |

| | |
|---|----|
| 7.13. Distribuciones de las masas invariantes del (a) primero [159] y (b) segundo bosón Z reconstruido [160]. Los puntos representan datos experimentales. Los histogramas llenos muestran la predicción de diferentes simulaciones de MC. Los paneles inferiores de cada figura muestra la relación entre los puntos de datos y el histograma apilado. | 78 |
| 7.14. a) Momento transversal [161] y b) Masa invariante del sistema de cuatro leptones [162] en eventos seleccionados, reconstruida en la reproducción computacional. | 78 |
| B.1. Representación gráfica de un track con varios contactos. Tomado de [97]. | 86 |
| C.1. Función de Densidad de Probabilidad (PDF) para una variable x. Aparte de las condiciones de normalización las PDFs pueden exhibir uno o varios máximos locales o mínimos locales [171]. | 89 |
| C.2. Percentiles de la distribución χ^2 [149] | 95 |
| C.3. Representación gráfica del valor-p. La coordenada vertical es la densidad de probabilidad para cada resultado, calculado bajo la hipótesis nula. El valor-p es el área bajo la curva más adelante del punto de datos observados. Tomado de [173] | 95 |
| C.4. El χ^2/n , para n grados de libertad. Las curvas muestran en función de n el χ^2/n que corresponde a un valor de p dado [138]. | 96 |
| C.5. El p.d.f. Gaussiano estándar $\varphi(x)$ que muestra la relación entre los cuantiles Φ^{-1} y el nivel de confianza para (a) un intervalo de confianza central y (b) un intervalo de confianza unilateral [172]. | 98 |

Introducción

La reproducibilidad es un pilar fundamental del método científico. Es por ello que la motivación del presente trabajo es concebir la reproducción computacional, y a su vez la optimización del análisis original y una discusión de uno de los análisis del experimento ATLAS sobre la producción de bosones ZZ en colisiones protón-protón a una energía de centro de masa de 13 TeV en el LHC, usando datos abiertos. La colaboración ATLAS dió a conocer estos datos en el 2016 y corresponde a una luminosidad integrada de 10 fb^{-1} . Además, publicó un conjunto de herramientas que proporcionan interfaces interactivas de código abierto para la reproducción de dicho análisis. El marco de análisis de la producción ZZ , está escrito en $C++$ e interconectado con el framework de análisis llamado ROOT; disponible públicamente en la plataforma GitHub. Este código implementa los protocolos necesarios para leer los conjuntos de datos, hacer una selección inicial de los objetos físicos y el análisis, además de escribir histogramas y graficar resultados.

Una de las herramientas más interactivas para reproducir análisis son los notebooks de Jupyter, o Jupyter Notebooks, también conocido como cuaderno computacional. Es de código abierto y se usa para combinar código, resultados computacionales, texto explicativo y recursos multimedia en un solo documento. Por otro lado, se utilizó la interfaz de SWAN (servicio de análisis basado en la web proporcionado por el CERN) que está construida sobre Jupyter, lo que permite escribir y ejecutar estos cuadernos utilizando solo el navegador web. Todas estas herramientas fomentaron la reproducibilidad computacional al simplificar la reutilización del código, y como consecuencia de esto se propuso una optimización a partir de los resultados obtenidos de la reproducción computacional del análisis de la producción ZZ .

El interés por el estudio de la producción de pares de bosones Z en colisiones protón-protón (pp) en el Gran Colisionador de Hadrones (LHC) radica en que es una prueba del sector electrodébil del Modelo Estándar (SM) de la física de partículas a las escalas de energía más altas disponibles. Asimismo, la producción de dibosones no resonantes es también un proceso de fondo importante en las búsquedas y estudios del bosón de Higgs. Finalmente, es un antecedente en la búsqueda de nueva física que produzcan, igualmente, pares de bosones Z con una masa invariante alta, es decir del orden de varios cientos de GeV .

El objetivo principal de este trabajo es generar una reproducción computacional, idear e implementar una optimización al estudio publicado por el equipo de ATLAS Open Data y la correspondiente discusión del análisis sobre la producción de ZZ en colisiones protón-protón a $\sqrt{s} = 13\text{ TeV}$ en el LHC. Para ello se propuso analizar, discutir y optimizar los resultados obtenidos de la reproducción computacional utilizando Jupyter Notebooks con kernel ROOT $C++$. En el presente informe se ofrece en primer lugar una introducción al Modelo Estándar, con especial énfasis en la teoría electrodébil. Además, en el capítulo 2 se ofrece una introducción más detallada a la producción de pares de bosones Z y resultados experimentales previos. En el capítulo 3 se hace hincapié en qué consiste el LHC y el experimento ATLAS, incluyendo detalles de cada uno de sus subdetectores.

En el capítulo 4, se brinda una descripción general de los algoritmos y software utilizados para la identificación y reconstrucción de los objetos físicos como lo son los electrones y muones, siempre en el contexto de ATLAS y sus datos abiertos.

Luego, en el capítulo 5 se describen los requisitos de selección de electrones, muones y otras propiedades de los eventos para seleccionar aquellos catalogados de interés en la producción de ZZ . Después de describir los requisitos de selección, se discuten las eficiencias, el cómo se aceptan y se evalúa la eficiencia de los eventos seleccionados en comparación a los esperados. Finalmente, se consideran las incertidumbres sistemáticas asociadas con la reconstrucción de dichos eventos.

En el capítulo 6 se describen las herramientas y el análisis computacional para la producción de ZZ . Es decir, se presentan los elementos más importantes de un flujo de trabajo reproducible, dando a conocer las principales herramientas computacionales disponibles para el manejo de datos, análisis de datos mediante código, control de versiones, la organización de los archivos, y el manejo de las dependencias de software externo.

En el capítulo 7, se hace énfasis en la optimización y los resultados obtenidos, para lo cual se describe la función de densidad de probabilidad (PDF), definiendo un examen estadístico de bondad de ajuste, t , el método de mínimos cuadrados, cuyo valor reflejó el nivel de acuerdo entre los datos y la hipótesis dada. Además se hace una comparación del método de mínimos cuadrados y de máxima verosimilitud, implementados en ROOT para histogramas de estadística baja. Finalmente, se hace el análisis de las distribuciones cinemáticas generadas en la reproducción.

Los procedimientos y resultados presentados contribuyen a posteriores ejemplos didácticos de análisis sobre el estudio de la producción ZZ , además de las técnicas y herramientas usadas para llevar a cabo el análisis.

Capítulo 1

Física de Partículas

La física de partículas estudia los bloques fundamentales de la naturaleza también llamado partículas fundamentales y sus interacciones (fuerzas), una mejor comprensión de cómo estas partículas y tres de las fuerzas están relacionadas entre sí se encapsula en el modelo estándar. Matemáticamente, el SM (por sus siglas en inglés, Model Standard) [5] es una teoría consistente con la mecánica cuántica y la relatividad especial, describe tres de las cuatro interacciones conocidas en la Naturaleza: Las interacciones electromagnéticas, fuerte y débil, basándose en las ideas de unificación y simetrías [1], dicho modelo se utiliza actualmente para estudiar la física de altas energías. En este capítulo se da una visión general de las interacciones y partículas fundamentales.

1.1. Partículas Fundamentales

Toda la materia está compuesta de dos familias de partículas fundamentales: Fermiones y Bosones.

Están clasificadas por su espín, ya sea:

- Fermiones: Se caracterizan por tener un espín semi-entero, obedecen el principio de exclusión de Pauli [2] y la estadística de Fermi-Dirac [3]. Estas partículas se presentan en dos grupos llamados quarks y leptones, cada grupo consta de seis partículas que están relacionadas en pares o generaciones.

Los 6 fermiones elementales llamados leptones, interactúan solamente a través de la interacción electromagnética (si están cargados eléctricamente) y débil [4]. Los seis leptones están arreglados en tres generaciones, la primera generación lo conforman las partículas más ligeras y estables y las más pesadas y menos estables conforman la tercera generación. La primera, el electrón (e) y el neutrino electrónico (ν_e), la segunda, el muón (μ^-) y su neutrino muónico (ν_μ) y finalmente en la tercera está el tauón (τ^-) y el neutrino tauónico (ν_τ).

Por otro lado, se conoce la existencia de 6 quarks, el cual poseen 6 grados de libertad llamados “sabores” [5]. Un sabor de quark puede cambiar a otro sabor a través de las interacciones débiles mediadas por los bosones débiles cargados W ; además los quarks tienen otro grado de libertad llamado “color”. Las interacciones entre quarks debidas a la carga de “color”, la cual no es nada más que la interacción fuerte, es mediada por los gluones,

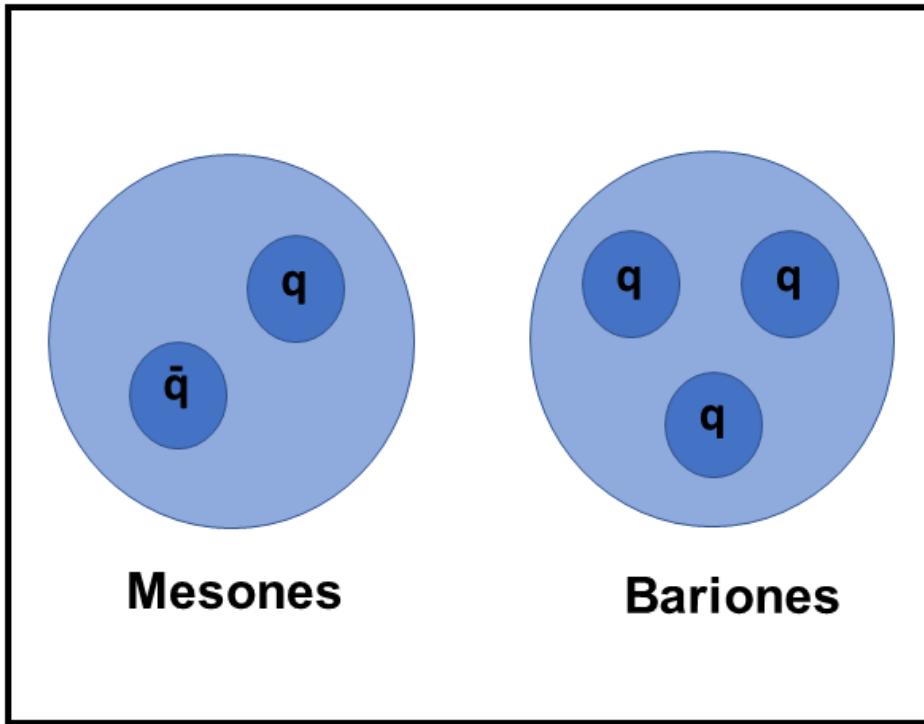


Figura 1.1: Tipos de hadrones. A la izquierda mesones conformados por un par quark-antiquark. A la derecha bariónes constituidos por tres quarks.

pero los quarks también interactúan a través de la fuerza electromagnética. Los quarks siempre están confinados en sistemas compuestos que se extienden sobre distancias de alrededor 1 fm. Los seis tipos o sabores de quarks están emparejados en tres generaciones, el up (u) y down (d) en la primera, seguidos por el charm (c) y strange (s), y después el top (t) y bottom (b). En la naturaleza los quarks se pueden agrupar a causa del fenómeno de confinamiento del color, formando bariónes, compuestos por tres quarks, y los mesones, formados por un par de quarks (quark-antiquark).

- **Bosones:** Se caracterizan por tener un espín entero, obedecen la estadística de Bose-Einstein [6] y su intercambio con los fermiones describe las interacciones fundamentales. Cada fuerza fundamental tiene su correspondiente Bosón. El fotón es el mediador de la interacción electromagnética, los bosones W^+ , W^- y Z^0 son los mediadores de la interacción débil, los ocho gluones mediadores de la interacción fuerte [7] y por último, el bosón de Higgs, responsable del mecanismo de Higgs, por el cual las partículas adquieren su masa, es una partícula de espín cero, sin carga eléctrica e inestable que fue introducida en el modelo para realizar la ruptura espontánea de simetría [8]; completando la lista de partículas elementales del Modelo Estándar.

Tabla 1.1: Propiedades de las partículas elementales del modelo estándar: Carga, masa(eV), espín y color.

| Partículas Elementales del Modelo Estándar | | | | | | | |
|--|----------------|----------------------|----------------|---------------------------|---------|----------|-------|
| | Generación | Nombre | Símbolo | Masa (eV/c ²) | Espín | Carga(e) | Color |
| Quark | 1 ^a | Up | u | ~3M | 1/2 | 2/3 | R/G/B |
| | | Down | d | ~6M | 1/2 | -1/3 | R/G/B |
| | 2 ^a | Charm | c | ~1.3G | 1/2 | 2/3 | R/G/B |
| | | Strange | s | ~100M | 1/2 | -1/3 | R/G/B |
| | 3 ^a | Top | t | 171G | 1/2 | 2/3 | R/G/B |
| | | Bottom | b | ~4.2G | 1/2 | -1/3 | R/G/B |
| Leptones | 1 ^a | Electrón | e ⁻ | 0.511M | 1/2 | -1 | - |
| | | Neutrino Electrónico | ν_e | <2 | 1/2 | 0 | - |
| | 2 ^a | Muón | μ^- | 106M | 1/2 | -1 | - |
| | | Neutrino Muónico | ν_μ | <2 | 1/2 | 0 | - |
| | 3 ^a | Tau | τ^- | 1.78G | 1/2 | -1 | - |
| | | Neutrino tauónico | ν_τ | <2 | 1/2 | 0 | - |
| Bosones | Fotón | γ | 0 | 1 | 0 | - | - |
| | W | W^\pm | 80.4G | 1 | ± 1 | - | - |
| | Z | Z^0 | 91.2G | 1 | 0 | - | - |
| | Gluón | g | 0 | 1 | 0 | R/G/B | - |
| | Higgs | H | 125G | 0 | 0 | - | - |

1.2. Construcción del Modelo Estándar

El modelo estándar de física de partículas es una teoría cuántica de campos consistente, renormalizable y libre de anomalías. Su construcción ha sido guiada por los principios de simetría, cuya matemática se fundamenta en la teoría de grupos [9]. La conexión profunda entre las simetrías y la física es reflejada en el teorema de Noether [10].

El desarrollo del Modelo Estándar o SM (del inglés Standard Model), comenzó con la teoría cuántica de campos de las interacciones electromagnéticas, llamada electrodinámica cuántica. Dicha teoría fue formulada por Dirac en 1920, y siguió hasta 1965 con los aportes de Tomonaga, Schwinger y Feynman que les valió el premio Nobel [11], esta teoría consiste en la interacción entre fotones y partículas cargadas de tipo fermiónico.

Posteriormente se desarrolló el modelo de las interacciones electrodébiles en los años 60 por Sheldon Glashow, Abdus Salam y Steven Weinberg (GWS) [12], dicho modelo es la formulación que permite unificar las interacciones débiles y electromagnéticas, además predijo la masa que deberían tener las partículas encargadas de la interacción débil y también la existencia de tres bosones, dos de ellos cargados (W^\pm) y uno neutro (Z^0). A su vez, en los años 70 fue propuesta la teoría cromodinámica cuántica por David Gross, David Politzer y Frank Wilczek [13] que describe una de las fuerzas fundamentales, la interacción fuerte. Finalmente la teoría electrodébil y la cromodinámica cuántica se unieron formando así el Modelo Estándar, que es hasta el momento la teoría más acertada para explicar las partículas elementales. En la tabla 1.2 se muestran las cuatro fuerzas fundamentales de la naturaleza y sus mediadores.

A continuación se explica con mas detalle cada una de estas teorías.

Tabla 1.2: Interacciones Fundamentales

| Interacción | Fuerza | Teoría | Mediador |
|------------------|------------|-------------------|----------|
| Fuerte | 10 | Cromodinámica | Gluón |
| Electromagnética | 10^{-2} | Electrodinámica | Fotón |
| Débil | 10^{-13} | Flavordynamics | W y Z |
| Gravitatoria | 10^{-42} | Gravedad Cuántica | Gravitón |

1.2.1. Electrodinámica Cuántica

La Electrodinámica Cuántica o QED (del inglés Quantum Electrodynamics) es una de las teorías físicas más precisas pero también más complicadas [5]. QED describe los fenómenos que implican las partículas eléctricamente cargadas que obran recíprocamente por medio de la fuerza electromagnética. Una forma básica de tal interacción se representa por una serie de diagramas de Feynman, como se muestra en la Figura 1.2. La QED se aplica a todos los fenómenos electromagnéticos asociados con las partículas fundamentales cargadas, como los electrones y positrones, y los fenómenos asociados tales como, la producción de pares, la aniquilación electrón-positrón, la dispersión de Compton, etc [14].

1.2.2. Teoría Electrodébil

En la década de 1960, Lee Glashow [15], Abdus Salam [16] y Steven Weinberg [17] propusieron la unificación de las fuerzas electromagnética y débil en una sola teoría que más tarde se conoció como la teoría electrodébil [12], dicha teoría requería la existencia de cuatro partículas portadoras sin masa, dos cargadas eléctricamente y dos neutras (γ , W^\pm , Z^0) para mediar en la interacción electrodébil unificada, pero el corto alcance de la fuerza débil indicaba, sin embargo, que es transportada por partículas masivas. Esto implicaba que la simetría subyacente de la teoría estaba “rota” por algún mecanismo que daba masa a las partículas intercambiadas en interacciones débiles pero no a los fotones intercambiados en interacciones electromagnéticas [8].

Sin embargo esta teoría es una teoría de gauge, y se requiere que sea invariante bajo transformaciones de gauge locales, es decir, que el Lagrangiano sea invariante bajo un conjunto de transformaciones continuas cuyos parámetros tengan una dependencia espacio-temporal [12]. Entonces requerir invariancia bajo una transformación de gauge local conduce a la aparición de campos de gauge vectoriales con bosones de gauge asociados.

Dicha teoría electrodébil unificada tiene un grupo de calibre:

$$U(1)_Y \times SU(2) \quad (1.1)$$

Los bosones gauge son un singlete B_μ sin masa asociado con el grupo $U(1)_Y$ y un triplete sin masa $W_\mu = (W_\mu^1, W_\mu^2, W_\mu^3)$ asociado con el grupo $SU(2)$. La cantidad conservada por el grupo $U(1)_Y$ es la hipercarga Y, y $SU(2)$ está asociada a la conservación del isospín débil [18]. Además, se observa que las interacciones débiles violan la conservación de la paridad; esto conduce a diferentes interacciones para los componentes zurdos y diestros de una partícula.

El lagrangiano de la teoría electrodébil es [5]:

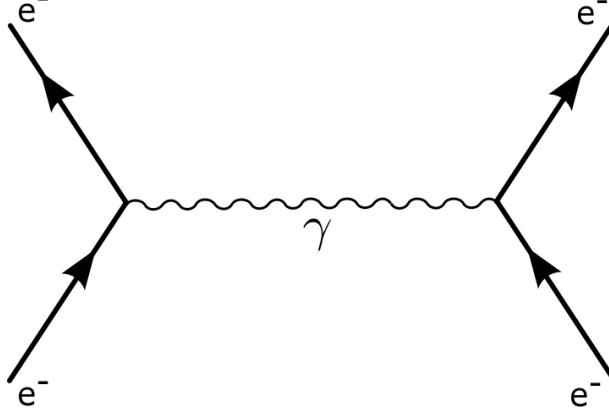


Figura 1.2: Diagrama de Feynman del intercambio de un fotón mediante la interacción entre dos electrones.[19]

$$\begin{aligned}
 L_{EW} = & -\frac{1}{4}B_{\mu\nu}B^{\mu\nu} - \frac{1}{4}F_{\mu\nu}^aF^{a\mu\nu} + \\
 & i\bar{l}_L^T\gamma^\mu D_\mu l_L + i\bar{e}_R^T\gamma^\mu D_\mu e_R + i\bar{\nu}_R^T\gamma^\mu D_\mu \nu_R + \\
 & i\bar{q}_L^T\gamma^\mu D_\mu q_L + i\bar{u}_R^T\gamma^\mu D_\mu u_R + i\bar{d}_R^T\gamma^\mu D_\mu d_R \quad (1.2)
 \end{aligned}$$

La primera línea proporciona términos cinemáticos para los campos de calibre, donde $B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$ es la intensidad del campo de hipercarga, y $F_{\mu\nu}^a = \partial_\mu W_\nu^a - \partial_\nu W_\mu^a - g f^{abc} W_\mu^b W_\nu^c$ es la intensidad de campo débil SU(2) [18]. El último término de la segunda expresión da lugar a acoplamientos triples y cuárticos de los bosones W^a . La segunda y tercera líneas de la ecuación 1.2 describen la cinemática de los fermiones en la teoría [5].

1.2.3. Ruptura Espontánea de Simetría y Mecanismo de Higgs

Los bosones gauge y los fermiones en el lagrangiano electrodébil dados en la ecuación 1.2 no tienen masa. Esto es claramente un problema, ya que se observa que los fermiones son masivos y, lo que es más preocupante, se observa que los bosones W y Z son extremadamente masivos. Desafortunadamente, los términos de masa explícitos para los bosones de gauge no pueden agregarse a la teoría sin violar la invariancia de gauge local, y no es posible incluir términos de masa explícitos para los fermiones ya que esto mezcla los componentes de la mano izquierda y derecha, que han sido asignados a diferentes multipletes SU(2) [20]. Para dar masa a los bosones gauge, la simetría debe romperse de alguna manera, la forma más sencilla de hacerlo sería simplemente agregar términos de masa para los bosones gauge a mano, sin embargo, esto conduciría a una teoría no renormalizable que contendría un número infinito de divergencias [20]. Entonces, a los bosones gauge se les da masa por ruptura espontánea de la simetría en el mecanismo de Higgs, introducido por Peter Higgs [20] en 1964, e independientemente por Robert Brout y Francoise Englert [21], y Gerald Guralnik, CR Hagen y Tom Kibble [22,23], mediante la incorporación de un doblete SU(2) de campos escalares complejos (el campo de Higgs), con un potencial de la forma [18]:

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \quad (1.3)$$

Si $\mu^2 < 0$, entonces el campo tiene un mínimo en $\Phi^\dagger \Phi = -\frac{1}{2} \mu^2/\lambda$, y por lo tanto tiene un valor de expectativa de vacío distinto de cero [18]. La invariancia de calibre del Lagrangiano se conserva, sin embargo, el estado de vacío ya no es invariante bajo las transformaciones de calibre. Los bosones gauge adquieren masa por interacciones con el campo y se encuentra que [18]:

$$M_W = \frac{1}{2}gv, M_Z = \frac{gv}{2\cos\theta_W} \quad (1.4)$$

Donde $v = \mu/\sqrt{\lambda}$ es el valor esperado de vacío [18]. También se pueden introducir términos que describan las interacciones de los fermiones con este campo, dando masas a los fermiones sin mezclar los componentes izquierdo y derecho. Esta interacción se denomina “interacción de Yukawa” [20]. Las fuerzas de los acoplamientos vienen dadas por los acoplamientos de Yukawa, que pueden ser diferentes para cada fermión y no están fijados por la teoría, requiriendo determinación experimental [20]. La introducción del campo de Higgs predice la existencia de una partícula fundamental, un bosón escalar masivo conocido como Bosón de Higgs con masa $m_H = \sqrt{2}|\mu|$ [18].

1.2.4. La Cromodinámica Cuántica

La Cromodinámica Cuántica o QCD (del inglés Quantum Chromodynamics) es una teoría cuántica de campos que describe las interacciones fuertes. Es una teoría de calibre perteneciente al grupo SU(3) [5]. La carga asociada con SU(3) es la carga de color, llamada así porque las tres cargas se denominan rojo, azul y verde [24]. La teoría se construye de manera similar a la teoría electrodébil, requiriendo invariancia bajo transformaciones de gauge locales generadas por el grupo [5]. Esto se logra reemplazando la derivada en la ecuación de Dirac con una derivada covariante, lo que requiere la introducción de campos de bosones gauge [24]. Como hay ocho generadores de SU(3), se introduce un octeto de campos gauge, que son los ocho gluones [24].

El lagrangiano para QCD viene dado por [18]:

$$L_{QCD} = -\frac{1}{4}G_{\mu\nu}^a G_a^{\mu\nu} + \bar{q}_a(i\gamma^\mu D_\mu)_{ab} q_b \quad (1.5)$$

donde el índice a corre sobre la carga de color. La derivada covariante se define como:

$$D_\mu = \partial_\mu + ig_s T^a G_\mu^a \quad (1.6)$$

donde T^a son los generadores de SU(3) y $G_{\mu\nu}^a$ es el tensor de intensidad de campo de color:

$$G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a - g_s f^{abc} G_\mu^b G_\nu^c \quad (1.7)$$

Los f^{abc} son las funciones de estructura del grupo SU(3) [18]. Como en el caso de los bosones W en la interacción débil, el último término de la Ecuación 1.6 da lugar a interacciones triples y cuarticas entre los gluones. g_s es la constante de acoplamiento fuerte, generalmente expresada como α_S , dada por [18]:

$$\alpha_S = \frac{g_s}{4\pi} \quad (1.8)$$

Dado que los gluones no tienen masa, QCD contiene divergencias UV que conducen a resultados infinitos en los cálculos de sección transversal [18]. Esto es claramente poco físico, por lo que se debe utilizar un procedimiento de renormalización para cancelar las divergencias, por ejemplo, introduciendo una masa de gluones o estableciendo una escala de corte de UV para regularizar

la divergencia [18]. Las divergencias son luego absorbidas por una redefinición de los parámetros “desnudos” de la teoría a los parámetros físicamente observables [18]. Los parámetros renormalizados como α_S dependen de la escala de renormalización μ_R , la escala a la que se realiza esta resta. Un resultado de este procedimiento es que α_S se vuelve dependiente del momento al cuadrado de las partículas externas que entran en un vértice, Q^2 [18]:

$$\alpha_S(Q^2) = \alpha_S(\mu_R^2) - \alpha_S(\mu_R^2)\beta_0 \ln \frac{Q^2}{\mu_R^2} \quad (1.9)$$

El parámetro β_0 es positivo, por lo que el acoplamiento efectivo disminuye a medida que aumenta la escala de momento Q^2 , un fenómeno que se conoce como el funcionamiento del acoplamiento [18]. Para escalas de momento por debajo de ~ 2 GeV, α_S se vuelve grande y no es posible realizar cálculos perturbativos en QCD. En este régimen no perturbativo se necesitan modelos fenomenológicos [18].

El funcionamiento del acoplamiento da lugar a dos fenómenos adicionales: el confinamiento y la libertad asintótica [18]. Dado que α_S aumenta a medida que Q^2 disminuye, la fuerza fuerte entre dos quarks aumenta a medida que aumenta la distancia entre ellos, lo que lleva a una unión cada vez más fuerte a medida que se separan. Con una separación suficientemente grande, la “cuerda” de color que conecta el quark se rompe y se produce un nuevo par de quarks. Un resultado de esto es que los quarks y gluones solo pueden observarse como estados ligados incoloros (hadrones) y nunca como cargas de color “desnudas” [25]. Por el contrario, a grandes escalas de momento (como en una colisión de partículas de alta energía) o distancias cortas (como dentro de un hadrón), el fuerte acoplamiento entre los quarks se vuelve muy pequeño y se comportan más como partículas libres.

1.3. Interacciones en Colisiones Protón-Protón

En la Figura 1.3 se muestra una ilustración de una colisión protón-protón en un colisionador de hadrones. En las colisiones protón-protón, no son los protones los que interactúan, sino sus constituyentes de quarks y gluones [18]. La interacción entre dos partones entrantes (quarks o gluones) de los dos protones iniciales se conoce como dispersión dura, que dará lugar a dos o más partículas salientes duras [18]. Las partículas entrantes y salientes emitirán más radiación QCD conocida como radiación de estado inicial (por sus siglas en inglés, ISR) y radiación de estado final (por sus siglas en inglés, FSR) en forma de gluones, que a su vez pueden dividirse en más partones [18]. Una vez que los quarks y gluones salientes se alejan lo suficiente entre sí, el acoplamiento entre ellos se vuelve fuerte y se forman en hadrones de color neutro en un proceso conocido como hadronización [5]. Las partículas cargadas entrantes y salientes también emitirán radiación electromagnética (QED) en forma de fotones.

La interacción entre las partículas entrantes deja atrás otros partones de la ruptura del protón original. También puede haber otras colisiones secundarias entre estos partones residuales, denominadas interacciones múltiples [5]. Las partículas de ISR y FSR, la ruptura del protón y las interacciones múltiples se denominan colectivamente el evento subyacente (por sus siglas en inglés, UE) [25].

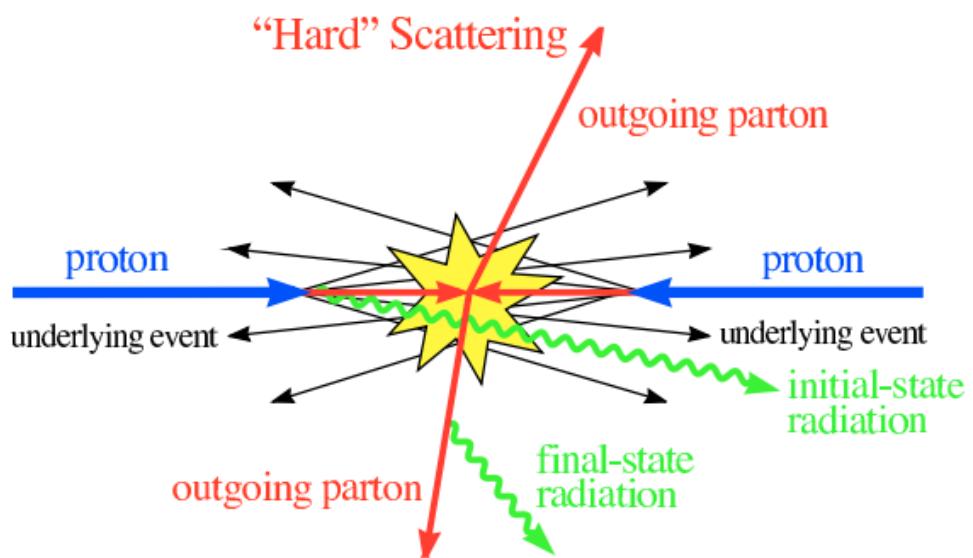


Figura 1.3: Ilustración de una colisión protón-protón [25].

Capítulo 2

Producción ZZ

En este capítulo se introduce una descripción del ZZ Diboson del Modelo Estándar y los esfuerzos para encontrarlo. Comenzando por el estudio de la producción de pares de bosones Z en colisiones de partículas (la denominada producción de dibosón ZZ) es de gran interés ya que proporciona una prueba de precisión del Modelo Estándar y una oportunidad única para sondear la estructura del sector electrodébil [26]. Los acoplamientos de bosones de calibre triple neutro ZZZ y $ZZ\gamma$ (por sus siglas en inglés nTGCs) son cero a nivel de árbol en el modelo estándar, y existen solo en el nivel de $O(10^{-4})$ en las correcciones de un lazo [27]. Sin embargo, los tamaños de estos nTGCs se mejoran en muchos modelos de nueva física. Además, la medición de estos acoplamientos proporciona así una prueba de la estructura del sector electrodébil del Modelo Estándar [26].

La producción de ZZ no resonante es también el trasfondo irreducible de las desintegraciones $H \rightarrow ZZ$, uno de los canales clave en la física del bosón de Higgs en el LHC [26]. Las desintegraciones $H \rightarrow ZZ$ fueron un canal de búsqueda clave para el descubrimiento del bosón de higgs con una masa cercana a 125 GeV, contribuyendo con un significado local de 3.6σ al significado local general de 6.0σ (los otros canales contribuyentes fueron $H \rightarrow YY$ y $H \rightarrow WW$) [28]. La buena comprensión de la producción de ZZ no resonante estudiada en esta tesis fue esencial para este descubrimiento y sigue siendo importante para estudiar las propiedades del nuevo bosón.

2.1. Producción ZZ en el Colisionador de Hadrones

El estudio de la producción de pares de bosones Z en interacciones protón-protón (pp) en el Gran Colisionador de Hadrones (LHC) con una energía de centro de masa de $\sqrt{s} = 13$ TeV, la producción ZZ está dominada por interacciones quark-antiquark ($q\bar{q}$), con una contribución de $O(10\%)$ de las interacciones gluón-gluón (gg) inducidas por bucle [29]. La producción del ZZ del modelo estándar puede proceder a través de un propagador de bosones de Higgs, aunque esta contribución se suprime en la región donde ambos bosones Z se producen en la cáscara [26]. Además, en la medición se incluyen pequeñas contribuciones de la producción de tribosones con dos bosones Z en descomposición leptónica y un tercer bosón débil en descomposición hadrónica y contribuciones de la dispersión del partón doble [26].

Los procesos de fusión gluón-gluón también contribuirán mediante diagramas de caja de quark, como se muestra en la Figura 2.1 [30]. Aunque estos son suprimidos por un factor de α_s^2 , debido al alto contenido de gluones del protón a las energías del LHC, contribuyen con una fracción considerable de la sección transversal de producción total de ZZ, aproximadamente con un 5-10 % [31], dependiendo de la energía y el rango de masa de los bosones Z.

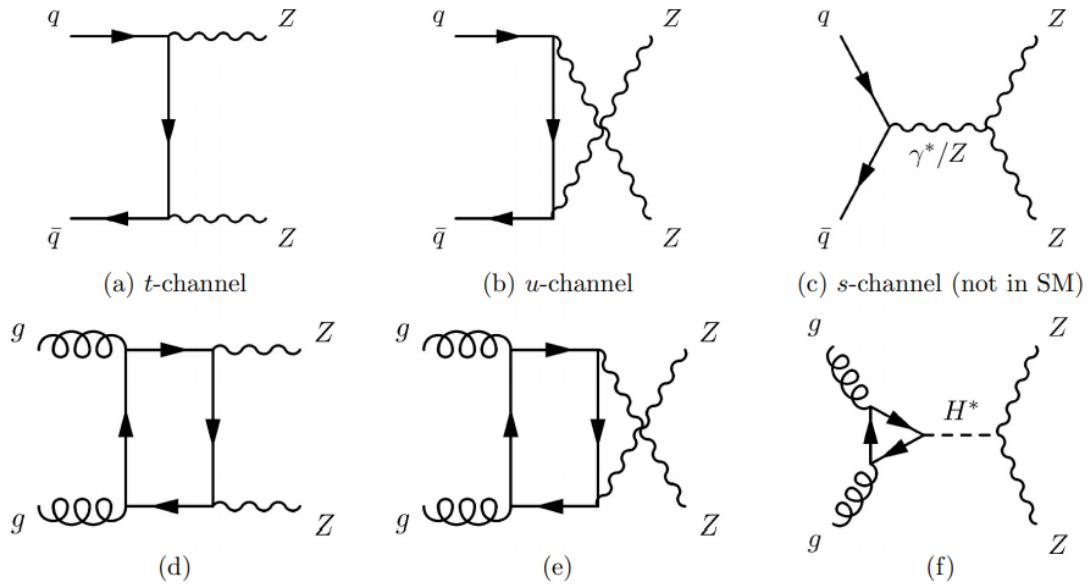


Figura 2.1: Diagramas de Feynman de menor orden para la producción de ZZ. Los diagramas de (a) canal t y (b) canal u contribuyen a la sección transversal de producción de ZZ, mientras que el diagrama del (c) canal s no está presente en el SM, ya que contiene un vértice neutral ZZZ o ZZ γ . En (d), (e) y (f) se muestran ejemplos de contribuciones de un bucle a la producción de ZZ mediante pares de gluones [30].

2.1.1. Modos de Desintegración de ZZ

Debido a que el bosón Z es neutro, la suma de las cargas de sus productos deben ser cero, así que debe decaer en un par partícula/antipartícula [18]. Los bosones Z pueden decaer en un par quark-antiquark, un par neutrino-antineutrino ó un par de leptones con carga opuesta [18]. Las fracciones de ramificación para cada uno de los estados finales son bien conocidas [32], y son 69,9 % para $q\bar{q}$, 20,0 % para $\nu\bar{\nu}$ y 10,1 % para l^-l^+ , en total son 24 posibilidades de desintegración, pero sólo 21 de ellas son visibles. En las desintegraciones ZZ, cada bosón se desintegra de forma independiente, por lo que la fracción de ramificación para un estado final dado es el producto de las fracciones de ramificación de los dos bosones Z. Las medidas de esta tesis se basan todas en medidas de $ZZ \rightarrow l^-l^+l^-l^+$, donde $l = e, \mu$, dando tres estados finales $e^-e^+e^-e^+, \mu^-\mu^+\mu^-\mu^+$ y $e^-e^+\mu^-\mu^+$. Las fracciones de ramificación a estos estados finales, obtenidas multiplicando las fracciones de ramificación de [32], son las siguientes:

$$B(ZZ \rightarrow e^-e^+e^-e^+) = 0,113 \pm 0,008 \% \quad (2.1)$$

$$B(ZZ \rightarrow \mu^-\mu^+\mu^-\mu^+) = 0,113 \pm 0,014 \% \quad (2.2)$$

$$B(ZZ \rightarrow e^-e^+\mu^-\mu^+) = 0,226 \pm 0,016 \% \quad (2.3)$$

2.1.2. Definición de Sección Transversal

La sección transversal de producción de ZZ no resonante en colisiones protón-protón a 13 TeV de energía del centro de masa se midió utilizando 10 fb^{-1} de datos registrados con el detector

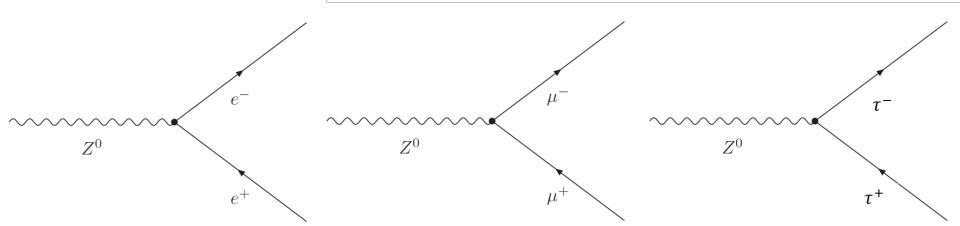


Figura 2.2: Decaimiento de Z en un Lepton-Antilepton.

ATLAS en el Gran Colisionador de Hadrones [33]. Los candidatos a bosones Z considerados decaen a un par de electrones o muones y se requiere que ambos bosones tengan una masa en el rango $66 < m_Z < 116$ GeV [33]. Las mediciones también se presentan en un espacio de fase restringido, denominado volumen fiducial, que corresponde estrechamente a los requisitos de selección experimental descritos en el Capítulo 5. La sección transversal fiducial correspondiente tiene incertidumbres teóricas menores que la sección transversal total, ya que al extrapolar de la sección transversal medida experimentalmente sección transversal fiducial a la sección transversal total surgen incertidumbres adicionales debido a incertidumbres en la PDF y las escalas de factorización y renormalización [18].

Las secciones transversales fiduciales se miden usando desintegraciones donde ambos bosones Z se desintegran en electrones o muones, denotados $ZZ \rightarrow l^-l^+l^-l^+$ [29]. Luego se extrapolan a la sección transversal total, corrigiendo la aceptación geométrica del volumen fiducial y las fracciones de ramificación a leptones [29]. El espacio de fase fiducial, que está diseñado para reflejar la aceptación del detector ATLAS (descrito a continuación), se define para eventos simulados aplicando los siguientes criterios a los objetos de nivel de partículas en estado final

La sección transversal fiducial $ZZ \rightarrow l^-l^+l^-l^+$ on-shell (ZZ) se define como [33]:

- $(Z/\gamma^*)(Z/\gamma^*) \rightarrow l^-l^+l^-l^+, l = e, \mu;$
- Disparador de un solo electrón o de un solo muón satisfecho;
- Exactamente cuatro leptones (electrones o muones) con $p_T > 20$ GeV, al menos uno de ellos debe tener $p_T > 25$ GeV;
- Se forman dos pares de leptones (e^+e^- o $\mu^+\mu^-$) dando lugar a tres canales: $4e, 4\mu$ y $2e2\mu$;
- Cada par de leptones SFOS debe tener una masa invariante dentro de un rango de $66 < m_Z < 116$ GeV;
- En los canales $4e, 4\mu$, donde hay dos formas posibles de formar pares de leptones, se elige la combinación que minimiza $|m_{ll,1} - m_Z| + |m_{ll,2} - m_Z|$ donde $m_{ll,1}$ y $m_{ll,2}$ son las masas invariantes de los dos pares de leptones.

2.2. Acoplamientos Anómalos de Triple Calibre

Los vértices ZZZ y $ZZ\gamma$ (colectivamente ZZV) están prohibidos a nivel de árbol en el Modelo Estándar, y surgen solo al nivel de $O(10^{-4})$ en las correcciones de un bucle [27]. Sin embargo, tales acoplamientos pueden surgir como resultado de contribuciones de Nueva Física (NF). Los

acoplamientos de triple calibre neutros anómalos (en lo sucesivo denominados por sus siglas en inglés nTGCs) se introducen por medio de una estructura lagrangiana eficaz [30]. La suposición básica es que hay algún NF más allá del Modelo Estándar a una escala masiva Λ , mucho más allá del alcance de los experimentos actuales [30]. Por tanto, las nuevas partículas resultantes no son directamente observables, y el único efecto observable de NF es interacciones anómalas de partículas del SM. Dado que en los acoplamientos ZZV siempre hay al menos dos partículas idénticas, las estadísticas de Bose prohíben los vértices ZZV con todas las partículas en el caparazón; por lo tanto, al menos uno de los bosones debe estar fuera de la cáscara [33]. La forma más general de los acoplamientos anómalos, asumiendo la invariancia de calibre de Lorentz y $U(1)_{EM}$, da dos acoplamientos independientes para cada uno de los vértices ZZV, o cuatro parámetros en total [34]. Dos de estos, denominados f_4^Z y f_4^γ , dan lugar a interacciones que violan CP, mientras que los dos restantes, f_5^Z y f_5^γ , dan lugar a interacciones que conservan CP [34]. A priori no existe relación entre estos acoplamientos.

La forma más general del lagrangiano para la teoría efectiva es [34]:

$$L = \frac{e}{m_Z^2} [f_4^V (\partial_\mu V^{\mu\beta}) Z_\alpha (\partial^\alpha Z_\beta) + f_5^V (\partial^\mu V_{\mu\alpha}) (\tilde{Z}^{\alpha\beta} Z_\beta)] \quad (2.4)$$

Donde $V = Z, A$ para Z o tensor de fotones, $\tilde{Z}_{\mu\nu} = \frac{1}{2} \epsilon_{\mu\nu\rho\sigma} Z^{\rho\sigma}$ y $V_{\mu\nu} = \partial_\mu V_\nu - \partial_\nu V_\mu$.

Dado que la desviación de la predicción SM crece rápidamente con el centro partónico de energía de masa $\sqrt{\hat{s}}$, y eventualmente puede violar la unitaridad, es convencional aplicar un factor de forma a los acoplamientos básicos dados anteriormente, para mantener la unitaridad en $\sqrt{\hat{s}}$ alta [34]. El factor de forma se aplica como [34]:

$$f_i^V(\hat{s}) = \frac{f_{i,0}^V}{(1 + \hat{s}/\Lambda)^n} \quad (2.5)$$

Donde $i = 4, 5, V = Z, \gamma$ y Λ se toma como la escala en la que la nueva física se vuelve directamente observable. n da la potencia del factor de forma y generalmente se considera $n = 3$ [35]. Para los límites derivados en un \hat{s} fijo, es trivial convertir entre parámetros de acoplamiento derivados con diferentes opciones de Λ y n en diferentes valores de \hat{s} [35]. Desafortunadamente, en los colisionadores de hadrones los límites surgen de la integración de un rango de \hat{s} , por lo que no es posible comparar los límites derivados en diferentes Λ , n o para diferentes rangos de \hat{s} [18]. Por lo tanto, los límites a menudo también se derivan sin un factor de forma (o de manera equivalente, tomando $\Lambda = \infty$) para permitir la comparación entre los resultados de diferentes experimentos [18].

2.2.1. Origen de los Acoplamientos Anómalos de Triple Calibre Neutros

El modelo más simple para generar acoplamientos anómalos de triple calibre neutros (por sus siglas en inglés, nTGCs) es de bucles de fermiones pesados que surgen de una nueva generación de fermiones; esto se muestra en la Figura 2.3 [18]. Nuevos fermiones pesados surgen en muchos modelos más allá del SM, como SUSY [18]. Dichos bucles solo pueden generar los acoplamientos f_5^V que conservan CP; Se necesitan procesos de orden superior para generar acoplamientos que violen CP f_4^V [18]. Un fermión pesado con acoplamientos del modelo estándar a Z y γ daría acoplamientos [36]:

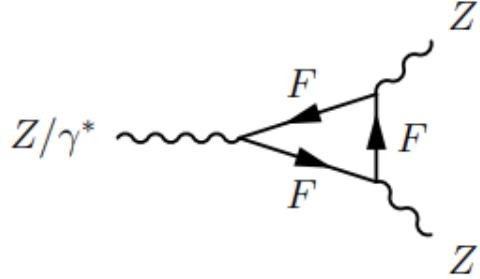


Figura 2.3: Producción de un vértice ZZV a través de un bucle de fermión pesado [18].

$$f_5^V = \frac{\alpha m_Z^2}{4\pi M_F^2} \quad (2.6)$$

Donde M_F^2 es la masa de fermiones pesados [36]. En un escenario experimentalmente favorable, si una partícula de la familia tuviera una masa mucho más baja que el resto de la familia, el acoplamiento se restablecería con una fuerza similar a la dada en la Ecuación 2.6 [36]. Esto todavía está suprimido por el factor de bucle $\alpha/4\pi$, y produce acoplamientos anómalos de orden 10^{-3} para M_F en el rango de 100 GeV [36]. Las contribuciones de los bucles de partículas SM generan acoplamientos al nivel de $O(10^{-4})$. Un ejemplo concreto de un nuevo modelo de física que da lugar a nTGCs anómalas es la supersimetría [36].

2.3. Resultados Experimentales Previos

La producción del Diboson ZZ se observó por primera vez en colisiones e^+e^- en LEP en 1997 cuando el centro de energía de masa del colisionador alcanzó por primera vez 183 GeV , el umbral para producir dos bosones Z en la capa [42]. El experimento L3 publicó la primera observación y medición de la sección transversal de la producción de ZZ en el caparazón [37], basada en $55,3 pb^{-1}$ de datos recopilados a una energía de centro de masa promedio de $182,7 \text{ GeV}$. Todos los canales de desintegración visibles se utilizaron en la medición. Si bien no se observaron eventos en el estado final $l^-l^+l^-l^+$, se observaron un total de 63 en los otros estados finales visibles [37]. La mayoría, (47) de estos estaban en el canal hadrónico, que sufría de antecedentes altos de $e^+e^- \rightarrow qq\gamma$ y $e^+e^- \rightarrow WW$ [37]. En este canal se utilizó un método de red neuronal para distinguir los eventos de señal del fondo. Se utilizó un ajuste de probabilidad logarítmica de la salida de la red neuronal y los espectros de masas observados en los otros canales para combinar los canales, dando una sección transversal de $\sigma_{ZZ} = 0.30^{+0.22}_{-0.16}{}^{+0.07}_{-0.03} \text{ pb}$, en muy buen acuerdo con la predicción del modelo estándar [37].

Las mediciones de la sección transversal de producción de ZZ en colisiones e^+e^- se realizaron posteriormente mediante los cuatro experimentos LEP en energías de centro de masa entre 183 GeV y 209 GeV [38-41]. Estos se han combinado en cada centro de energía de masa utilizando una técnica de minimización de χ^2 , teniendo en cuenta las correlaciones entre las incertidumbres sistemáticas [42]. Las secciones transversales combinadas resultantes en función de la energía del centro de masa, que se muestran en la Figura 2.4, concuerdan bien con las predicciones teóricas.

Cada uno de los experimentos LEP también estableció límites en los acoplamientos anómalos de triple vía f_i^V de ZZZ y $ZZ\gamma$ [38]. Los límites se establecieron utilizando la sección transversal ZZ total medida, así como distribuciones cinemáticas sensibles a aTGCs. ALEPH [41], DELPHI

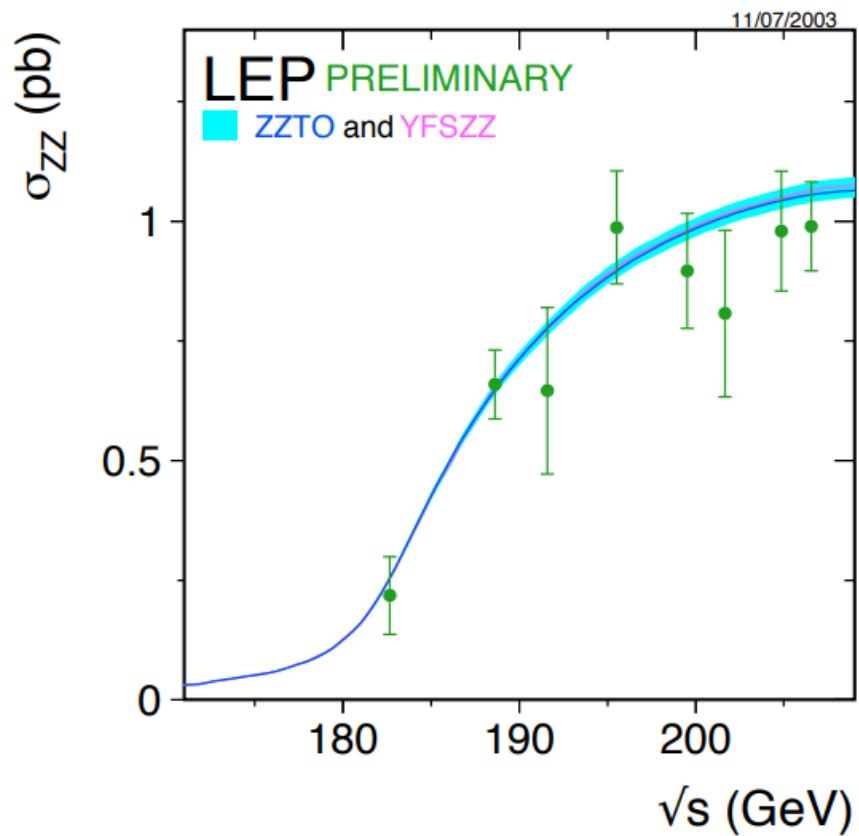


Figura 2.4: Mediciones de la sección transversal de producción ZZ en colisiones e^+e^- en función de la energía del centro de masa \sqrt{s} . Las medidas son una combinación de medidas de los cuatro experimentos LEP [47].

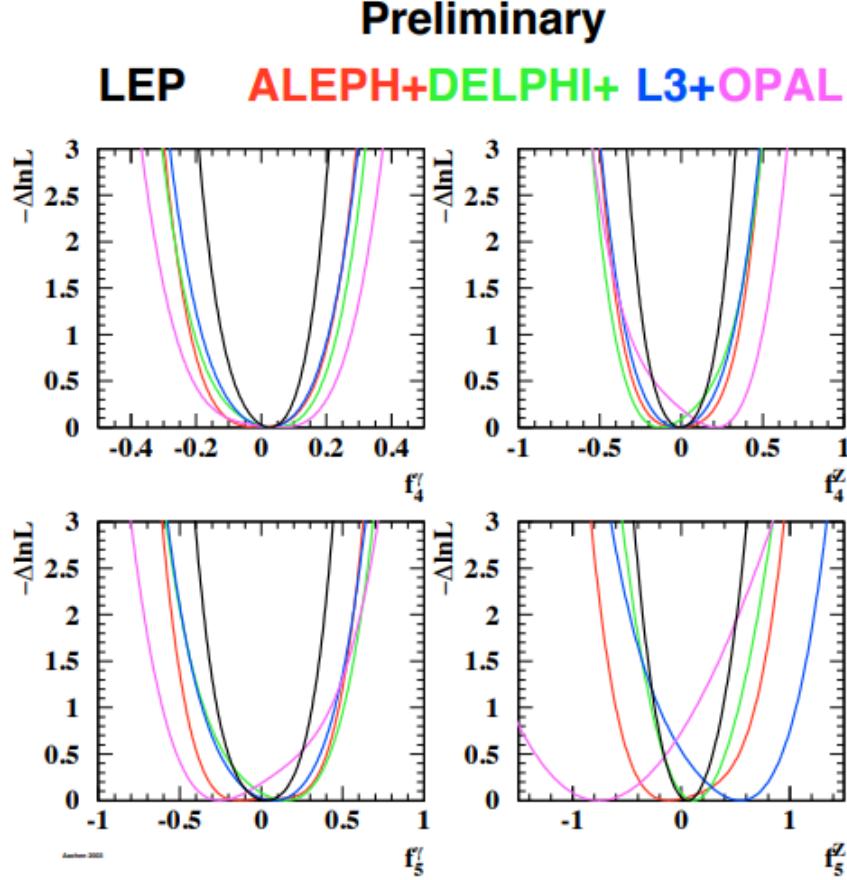


Figura 2.5: Curvas de probabilidad logarítmica negativas para acoplamientos de nTGCs de los cuatro experimentos LEP (bandas coloreadas) y su combinación (banda negra). No se aplica ningún factor de forma [42].

[43] y OPAL [44] utilizaron la distribución $\cos(\theta_Z)$, donde θ_Z es el ángulo de producción del bosón Z con respecto al eje del haz. L3 [40] utilizó ajustes para diferentes variables cinemáticas en cada canal de desintegración. Cada experimento primero combinó sus límites a través de los diferentes canales de desintegración y las energías del centro de masa, y proporcionó la probabilidad logarítmica negativa en función del parámetro a combinar. Los límites se establecieron variando cada parámetro de forma independiente, manteniendo los otros parámetros en su valor de modelo estándar de cero [40-44]. También se realizaron dos ajustes de parámetros. No se observaron desviaciones del modelo estándar. Las distribuciones de probabilidad logarítmica negativa de cada uno de los experimentos, así como la probabilidad logarítmica negativa combinada, se muestran en la Figura 2.5.

Se han realizado mediciones de la sección transversal de producción de ZZ en colisiones de pp en un centro de energía de masa de $\sqrt{s} = 1,96 \text{ TeV}$ en el Tevatron mediante los experimentos D0 y CDF [45-47]. D0 midió la sección transversal ZZ en 6.4 fb^{-1} de luminosidad integrada usando el canal $ZZ \rightarrow l^-l^+l^-l^+$ [45] y en 8.6 fb^{-1} usando el canal $ZZ \rightarrow l^-l^+\nu\bar{\nu}$ [46]. En el canal $ZZ \rightarrow l^-l^+l^-l^+$ se observaron 10 eventos con el requisito de que ambos pares de dileptones tengan una masa superior a 30 GeV, con un fondo esperado de $0,37 \pm 0,13$ eventos [45]. Combinando los dos canales, la sección transversal medida para la producción de ZZ con el requisito de

$60 < m_Z < 120\text{GeV}$ fue $\sigma(pp \rightarrow ZZ) = 1,44^{+0,31}_{-0,28}(stat)^{+0,17}_{-0,19}(syst)\text{pb}$, de acuerdo con la predicción del Modelo Estándar de $\sigma(pp \rightarrow ZZ) = 1,3 \pm 0,1\text{pb}$ [45-46]. Las distribuciones cinemáticas observadas también estaban en buen acuerdo con las predicciones del Modelo Estándar.

CDF [47] midió la sección transversal ZZ usando los canales $ZZ \rightarrow l^-l^+\nu\bar{\nu}$ y $ZZ \rightarrow l^-l^+l^-l^+$ en 6fb^{-1} de luminosidad integrada. En el canal $ZZ \rightarrow l^-l^+l^-l^+$, se requirió que los eventos tuvieran dos pares de leptones del mismo sabor de signo opuesto, con un par requerido para tener una masa invariante dentro de 25 GeV de la masa Z , y el otro en el rango $40 < m_{l^-l^+} < 140\text{ GeV}$. Se observaron 14 eventos de este tipo, con un fondo esperado de $0,26^{+0,53}_{-0,15}$ eventos [47]. La sección transversal medida, combinando los dos canales y corrigiendo a la aproximación de ancho cero, fue $(p\bar{p} \rightarrow ZZ) = 1,64^{+0,44}_{-0,38}(stat + syst)\text{pb}$, nuevamente de acuerdo con la predicción del Modelo Estándar de $(p\bar{p} \rightarrow ZZ) = 1,4 \pm 0,1\text{pb}$ [47].

D0 estableció límites en nTGCs usando $ZZ \rightarrow l^-l^+l^-l^+$ decaimientos en 1 fb^{-1} de datos [48], requiriendo que ambos pares de dileptones tuvieran una masa invariante $> 50(70)\text{ GeV}$ para pares de dimuones (di-electrones). No se observaron eventos que pasaran esta selección. Los límites se establecieron comparando el número observado de eventos con los rendimientos pronosticados en muestras simuladas con diferentes valores de los acoplamientos anómalos.

El experimento CMS ha medido la sección transversal de producción de ZZ en colisiones de pp a 7 TeV [49] utilizando desintegraciones $ZZ \rightarrow l^-l^+l^-l^+$. En un conjunto de datos de 5 fb^{-1} , se seleccionaron eventos con dos pares de dileptones del mismo sabor de signo opuesto, con el requisito de que el primer par de dileptones debe estar compuesto de electrones o muones y tener una masa de $60 < m_{l^-l^+} < 120\text{ GeV}$, y el segundo par estará compuesto por electrones, muones o taus [49]. En el caso de que el segundo par sea e^+e^- o $\mu^+\mu^-$ se aplica el mismo corte de masa que para el primer par. En el caso de $\tau^+\tau^-$, la masa visible se requiere para satisfacer $30 < m_{\tau\tau}^{vis} < 80\text{ GeV}$ [49]. En los canales $4e, 4\mu$ y $2e2\mu$ se observaron 54 eventos, con una predicción del Modelo Estándar de $54,5 \pm 0,3(stat) \pm 4,8(syst)$ (incluido un componente de fondo pequeño) [49]. En los canales $2l2\tau$, se observaron 11 eventos, con una predicción del modelo estándar de $11,7 \pm 0,8(stat) \pm 1,0(syst)$ (incluidos 4,4 eventos de procesos de fondo) [49]. Combinando todos los estados finales, la sección transversal medida, que requiere que ambos bosones Z estén en el rango $60 < m_Z < 120\text{ GeV}$, resultó ser $\sigma(pp \rightarrow ZZ) = 6,24^{+0,86}_{-0,80}(stat)^{+0,41}_{-0,32}(syst) \pm 0,14(lumi)\text{pb}$, en excelente acuerdo con la predicción del Modelo Estándar de $6,3 \pm 0,4\text{pb}$. Se establecieron límites en las nTGCs utilizando la distribución de masa de cuatro leptones observada para los canales $4e, 4\mu$ y $2e2\mu$ combinados [49].

CMS también publicó recientemente una medición de la sección transversal de producción de ZZ en $\sqrt{s} = 8\text{ TeV}$ en un conjunto de datos de $5,3\text{fb}^{-1}$ [50], utilizando un análisis similar a la medición de $\sqrt{s} = 7\text{ TeV}$. Midieron que la sección transversal era $\sigma(pp \rightarrow ZZ) = 8,4 \pm 1,0(stat) \pm 0,7(syst) \pm 0,4(lumi)\text{pb}$, consistente con la predicción del Modelo Estándar de $7,7 \pm 0,4\text{pb}$ [50].

Capítulo 3

CERN, LHC y el Experimento ATLAS

Las partículas y la radiación solo se pueden detectar a través de sus interacciones con la materia. Existen interacciones específicas para las partículas cargadas y neutras. Se puede decir que cada proceso de interacción se puede utilizar como base para el desarrollo de detectores. Uno de los primeros detectores de partículas fue el contador de Geiger, que solo detectaba partículas alfa, fue inventado por el físico alemán Hans Geiger y Ernest Rutherford en 1908 [51].

En 1954 se fundó la Organización Europea para la Investigación Nuclear conocida como CERN, el cual es uno de los más grandes centros de investigación científica de la actualidad, cuenta con 22 estados miembros y posee el laboratorio más grande de física de partículas del mundo, ubicado en la frontera Franco-Suiza en las cercanías de Ginebra [52]. El CERN cuenta con El Gran Colisionador de Hadrones (LHC, por sus siglas en inglés) es el mayor acelerador de partículas del mundo. Los instrumentos utilizados en el CERN son aceleradores y detectores de partículas especialmente diseñados. Los aceleradores inyectan energía a los haces de partículas antes de que los haces choquen entre sí. Los detectores, como ATLAS, observan y registran los resultados de estas colisiones [53].

En este experimento, los físicos hacen colisionar entre sí partículas subatómicas principalmente protones a velocidades cercanas a la velocidad de la luz, el proceso proporciona pistas sobre cómo interactúan las partículas e información sobre las leyes fundamentales de la naturaleza. En varios puntos del LHC se ubican grandes detectores (ATLAS, CMS, LHCb y ALICE) y tres (3) pequeños (LHCf, MoEDAL y TOTEM), donde estos registran las partículas resultantes de las colisiones para estudiar los elementos que componen la materia de la que está hecha el Universo y sus interacciones [53].

3.1. El Gran Colisionador de Hadrones

El Gran Colisionador de Hadrones (por sus siglas en inglés: LHC) es el colisionador de partículas más grande y poderoso del mundo, la instalación experimental más compleja jamás construida. LHC se creó en el CERN entre 1998 y el 2008 en colaboración con 10 mil científicos e ingenieros de más de 100 países, asociados a cientos de universidades y laboratorios [54].

El LHC es un anillo de 27 kilómetros de circunferencia ubicado a 100 metros bajo tierra. Su primera puesta marcha (Run I) fue en marzo del 2010, con una energía de 3,5 TeV por haz [56]. ATLAS alcanzó una luminosidad integrada de $23.22 f_b^{-1}$ con una energía de 4 TeV para el año

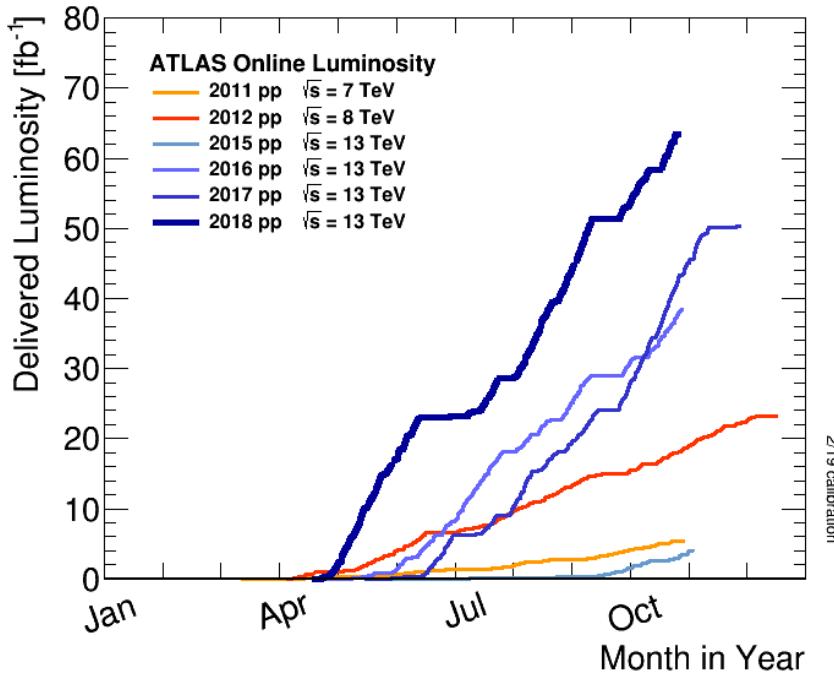


Figura 3.1: Gráfico de la luminosidad integrada obtenida en función del tiempo para 2011 - 2018 de operación con haces estables de protones de alta energía [59].

2012 [56]. En 2013 el LHC fue apagado para llevar a cabo una serie de mejoras programadas y para el año 2015 fue encendido otra vez (Run 2) en la cual consiguió una luminosidad integrada en ATLAS de 4.21fb^{-1} con una energía del centro de masa de 13 TeV [55]. En 2016 alcanzó una luminosidad integrada de 39.8fb^{-1} , y para el 2018 de 63.3fb^{-1} [55].

El colisionador cuenta con más de 50 tipos de electroimanes superconductores que hacen girar los haces de partículas alcanzando velocidades cercanas a las de la luz, estos deben refrigerarse a una temperatura inferior a la del espacio exterior [56]. Los haces están compuestos principalmente por protones, pero también se llevan a cabo estudios con haces de iones pesados, a menor energía. Cada uno consiste en alrededor de 2808 cúmulos o "bunches", mas o menos con 100 billones de protones, estrechamente empaquetados para asegurar una alta luminosidad en los puntos de interacción, y por tanto maximizar el número de colisiones [56].

Los dipolos principales (alrededor de 1232, de 15m de longitud) generan un campo nominal de 8,33T para inducir la trayectoria circular de los haces de partículas, mientras 392 cuadripolos (de $(5 - 7)\text{m}$ de largo, generadores de campos de alrededor de 6T) se encargan de enfocarlos [56]. También existen sextupolos, octupolos y decapolos empleados junto a los dipolos para la corrección de pequeñas imperfecciones en los extremos del campo magnético [57].

El proceso de acelerado de las partículas se da en escalas a través de anillos más pequeños integrados al LHC (ver figura 3.3 (a)). Núcleos de hidrógeno son primeramente insertados en un acelerador lineal (LINAC 2), de donde surgen con una energía de 50 MeV y pasan al Proton Synchrotron Booster (PSB). Allí adquieren una energía de 1,4 GeV para entrar al PS (Proton Synchrotron) y SPS (Super Proton Synchrotron), sincrotrones que los aceleran hasta los 26 y 450 GeV respectivamente [58].

Por último son inyectados al anillo principal, donde los bunches son acumulados y acelerados

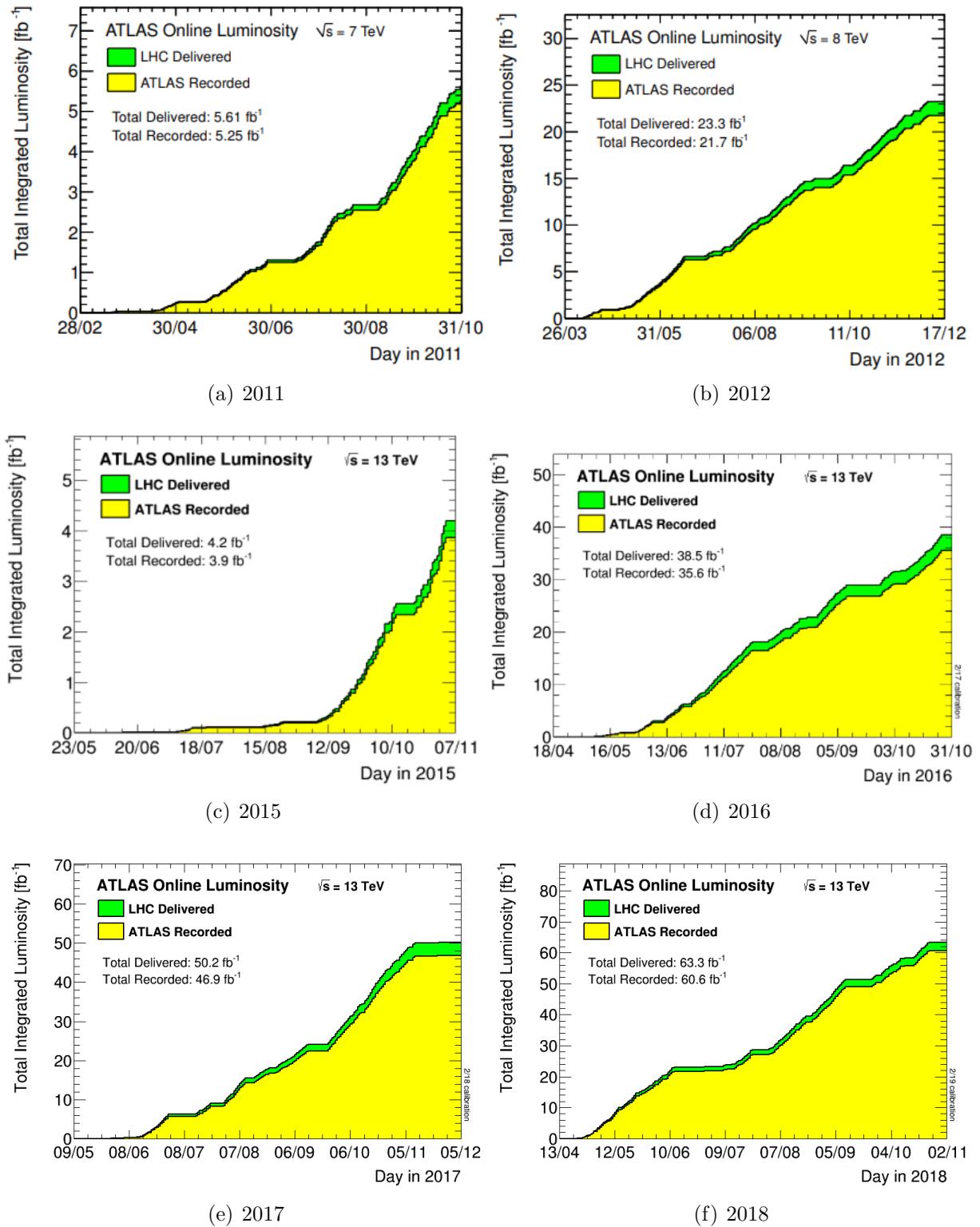


Figura 3.2: Luminosidad integrada en función del tiempo (para distintos años) que fue proporcionada por el LHC (verde) y que registró ATLAS(amarillo) [59].

Tabla 3.1: Parámetros operativos del LHC. Se comparan los parámetros nominales de diseño [60] y los utilizados en la operación de 2011 y en la operación de 2012 [61] [62].

| Parámetro | Nominal | Operación 2011 | Operación 2012 |
|----------------------------|-----------------------|----------------------|----------------------|
| Energía de Protones | 7 TeV | 3.5 TeV | 4 TeV |
| N_b | 1.15×10^{11} | 1.5×10^{11} | 1.6×10^{11} |
| n_b | 2808 | 1380 | 1380 |
| Espacio entre racimos (ns) | 25 | 50 | 50 |
| $\beta^*(m)$ | 0.55 | 1.0 | 0.6 |
| $\epsilon_n(\mu m)$ | 3.75 | 1.9 - 2.3 | 1.7 - 3.0 |
| Peak $L (cm^{-2}s^{-1})$ | 1.0×10^{34} | 3.6×10^{33} | 7.7×10^{33} |

en un período de 20 min hasta la energía nominal. Los haces permanecen en circulación durante 4 a 5 horas, mientras las colisiones son inducidas en 4 puntos del acelerador (ATLAS, CMS, LHCb y ALICE). En 4 de los puntos de interacción se sitúan los experimentos del LHC (ver figura 3.3 (b)), que se describen a continuación:

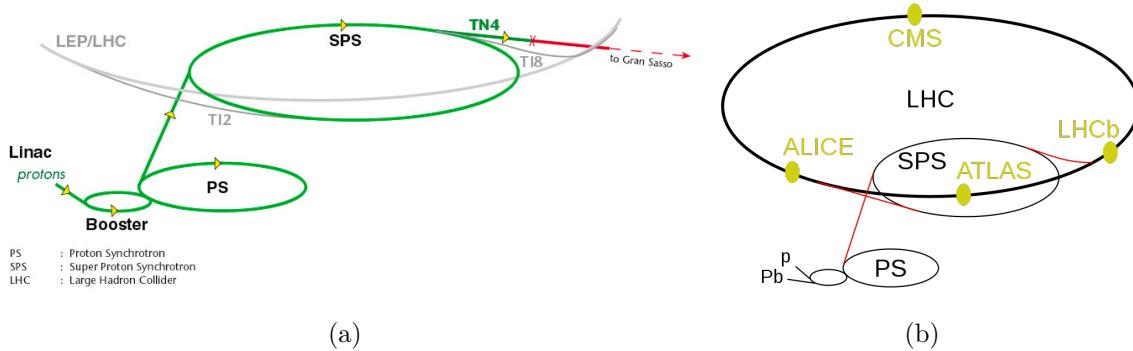


Figura 3.3: (a) Cadena de aceleración de las partículas antes de entrar al LHC [63]. (b) Experimentos del LHC [64].

- ALICE (por sus siglas en inglés, A Large Ion Collider Experiment): Dedicado al estudio de iones pesados relativistas, para explorar el potencial físico de las interacciones nucleón-nucleón a energías del LHC. Su objetivo es la física de la interacción fuerte a altas densidades, buscando recrear el plasma quark-gluón (QGP), verificar el estado de desconfinamiento y la reproducción de condiciones presentes instantes luego del Big Bang [58].
- LHCb (por sus siglas en inglés, Large Hadron Collider beauty): El experimento cuenta con un detector especializado en el estudio de las mediciones de los parámetros de violación CP y la física del quark b [66].
- CMS (por sus siglas en inglés, Compact Muon Solenoid): Es un detector multipropósito diseñado para el estudio de una amplia variedad de física que van desde el estudio del modelo estándar (incluyendo el bosón de Higgs) a la búsqueda de otras dimensiones y

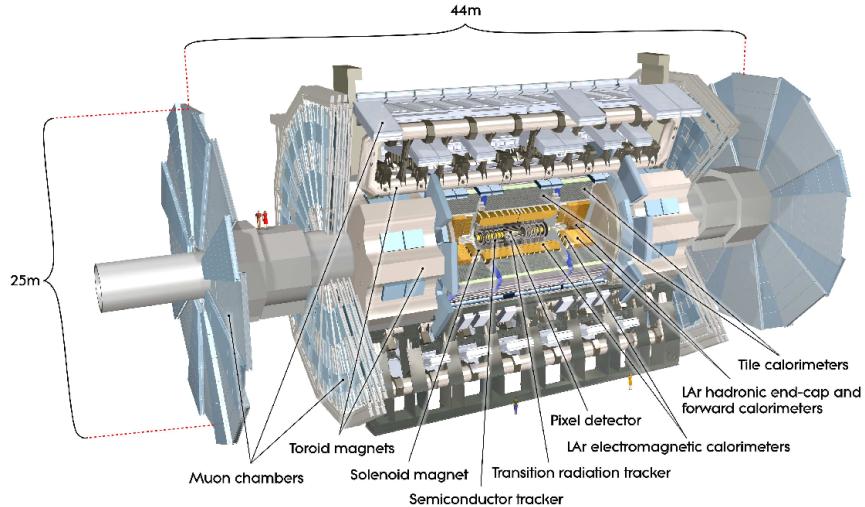


Figura 3.4: Esquema del detector ATLAS y los subdetectores que lo componen [65].

partículas que podrían conformar la materia oscura. Aunque tiene los mismos objetivos científicos que el experimento ATLAS utiliza diferentes soluciones técnicas y un diseño de sistema magnético diferente. Una sección será dedicada al detector ATLAS más adelante [66].

- MoEDAL (por sus siglas en inglés, Monopole and Exotics Detector At the Large Hadron Collider): La principal motivación de MoEDAL es la búsqueda de partículas súper exóticas como el monopolo magnético (dyon) y otras partículas masivas altamente ionizantes y estables (SMP) [66].
- LHCf (por sus siglas en inglés, Large Hadron Collider forward): Este experimento utiliza partículas creadas en el interior del LHC como fuente para simular los rayos cósmicos en condiciones de laboratorio [64].
- TOTEM (por sus siglas en inglés, TOTal cross section, Elastic scattering and diffraction dissociation Measurement at LHC): Para medir la sección eficaz total en las colisiones protón-protón, dispersión elástica y procesos de difracción en el LHC, se ubica a cada lado de CMS, a unos 200m del punto de interacción nominal de este experimento [66].

3.2. Detectores de Partículas: El experimento ATLAS

ATLAS es un detector multipropósito construido para el estudio de colisiones protón-protón ($p - p$) y de iones pesados ($A - A$) en el LHC [65]. Estudia un rango amplio de física, desde parámetros del Modelo Estándar hasta nueva física, pasando por la búsqueda del Bosón de Higgs, materia oscura, Supersimetría (SUSY) y dimensiones adicionales. Un diagrama del detector es mostrado en la figura 3.5 [65].

Haces de partículas del LHC colisionan en el centro del detector ATLAS creando nuevas partículas, las cuales salen del punto de colisión hacia todas las direcciones. Diferentes subdetectores están puestos en capas alrededor de donde ocurre la colisión, pudiendo así medir los caminos tomados, momentum, y energía de las partículas, pudiéndose identificar individualmente cuando

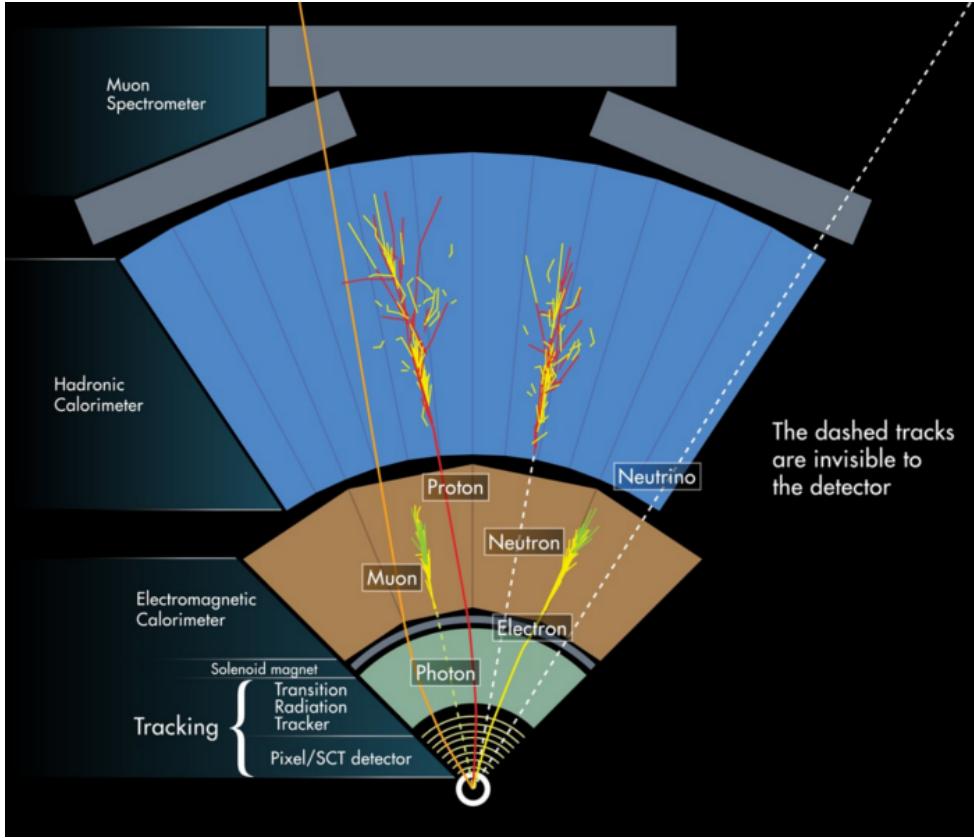


Figura 3.5: Esquema de la detección de las partículas producidas en una colisión de protones mientras pasan a través de varios subdetectores de ATLAS [66].

van desde la capa más interna hasta la más externa del detector, como se muestra en la figura 3.5 [66].

ATLAS fué diseñado para cumplir con una serie de requisitos. Entre estos destacan la resistencia de sus componentes a las condiciones experimentales en el LHC sin pérdida de fiabilidad ni rendimiento, sensores y sistemas electrónicos y de trigger (selección de eventos) lo suficientemente rápidos y eficientes, gran cobertura azimutal y polar y alta precisión tanto en medidas energéticas para la identificación y reconstrucción de fotones, electrones y jets, como para la lectura de trayectorias, vértices y momentos transversos [56]. De esta forma, el detector se compone de diferentes secciones, bastante especializadas en sus funciones, que pueden presentarse en 5 grupos principales. Estos se describen brevemente a continuación.

3.2.1. Sistema de Coordenadas en ATLAS

El sistema de coordenadas ATLAS es un sistema de coordenadas cartesiano diestro, con el punto de colisión nominal como origen (Figura 3.6), el eje z está a lo largo de la línea del haz y el plano $x - y$ es el plano perpendicular a la línea del haz [67]. El eje x apunta al centro del anillo LHC, mientras que el eje y va hacia arriba. El ángulo azimutal ϕ se define alrededor del eje del haz, mientras que el ángulo polar θ describe el ángulo desde el eje positivo z hasta el haz [67]. La variable θ no es invariante bajo aumentos a lo largo del eje z , por lo que en lugar del ángulo θ se usa la *pseudorapidez* η [67]:

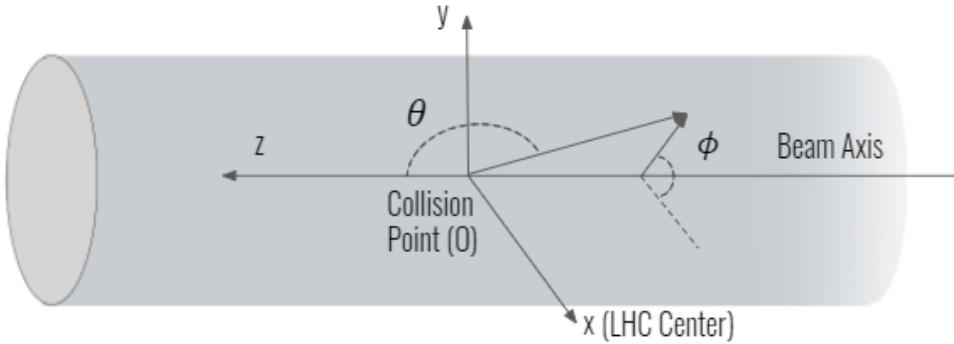


Figura 3.6: Sistema de coordenadas del experimento ATLAS. Donde θ es el ángulo polar y describe el ángulo de una partícula con respecto al eje del haz. ϕ es el ángulo azimutal, este ángulo se mide alrededor del eje z o la dirección del haz.

$$\eta = -\ln[\tan(\frac{\theta}{2})] \quad (3.1)$$

Dado que en un colisionador hadrónico las partículas en colisión reales son los partones dentro de los protones, podemos decir que el centro real de energía de masa es desconocido en cada colisión: $\hat{s} = x_1 \cdot x_2 \cdot s$, donde \hat{s} es la energía de colisión efectiva, x_1 y x_2 son las fracciones de momento transportadas por los dos partones en colisión y s es la energía de colisión de los dos protones [67]. Debido a esto, se desconoce la cantidad de movimiento total a lo largo del eje de la viga antes de la colisión, mientras que se sabe que la cantidad de movimiento total en el plano transversal (es decir, el plano x - y) es cero y, por lo tanto, solo podemos aplicar las leyes de conservación de la cantidad de movimiento y la energía en el plano transversal (porque sabemos cuál es el impulso total inicial).

Por esta razón, se consideran las cantidades transversales, y se denotarán con el subguion “T” (por ejemplo, p_T significa momento transversal, que es la proyección del momento en el plano $x - y$) [18].

3.2.2. Sistema de Imanes en ATLAS

El detector ATLAS está equipado con dos sistemas magnéticos: un solenoide superconductor [68], que proporciona un campo magnético al sistema de seguimiento interno, y un sistema de imanes toroidales superconductores con núcleo de aire [69, 70] ubicados en la parte exterior del detector como se muestra en la Figura 3.7. El solenoide cubre la región central del detector, proporciona un campo magnético uniforme de aproximadamente 2T a lo largo de las pistas de flexión del eje z de las partículas en el plano transversal para permitir que el sistema de seguimiento interno mida su momento transversal [68]. El solenoide está ubicado entre el detector interno y el calorímetro electromagnético.

El toroide es una de las peculiaridades del detector ATLAS: se encuentra fuera del sistema calorimétrico que cubre la región $|\eta| < 3$ (considerando todas sus subpartes), y proporciona un campo magnético cuyas intensidades máximas son 3.9T en la región central del detector y 4.1T en la región directa [69,70]. El objetivo de tal toroide es tener un brazo de palanca grande para mejorar la medición del momento transversal del muón, y está construido “en el aire” para minimizar la dispersión múltiple de muones dentro del detector.

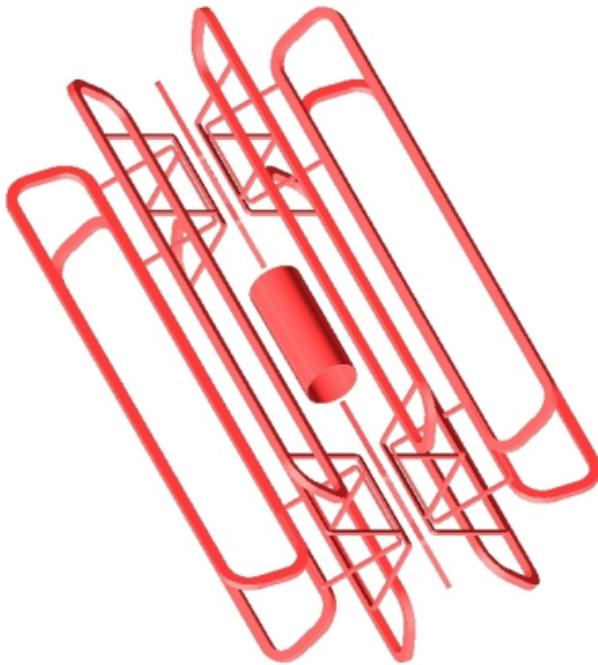


Figura 3.7: El sistema magnético del detector ATLAS: el cilindro interior es el solenoide superconductor, mientras que las partes externas son las bobinas del toroide [67].

El sistema magnético doble ATLAS ha sido diseñado para proporcionar dos mediciones independientes del momento transversal del muón en el detector interno y en el espectrómetro de muones, asegurando así una buena resolución del momento muónico desde unos pocos GeV hasta la escala TeV.

3.2.3. Detector Interno

El detector interno (por sus siglas en inglés: ID) es el primero en ver los productos de decaimientos de las colisiones, este mide la dirección, el momento, los vértices y la carga de partículas producidas en cada colisión protón-protón [56]. Consiste en tres sistemas diferentes de sensores, todos inmersos en un campo magnético de 2T paralelo al eje del haz, cubriendo la región $|\eta| < 2,5$ [18]. El detector interno está construido para brindar una buena reconstrucción de las trayectorias de las partículas, momentos y vértices. Los tres sensores del detector interno son: Detector de píxeles, rastreador de semiconductores (por sus siglas en inglés, SCT: Semiconductor Tracker) y rastreador de radiación de transición (por sus siglas en inglés, TRT: Transition Radiation Tracker) [71].

En la figura 3.8, se muestra la estructura del detector interno, el cual está compuesto por tres subdetectores cilíndricos concéntricos, que son [71]:

- El Detector de Píxeles (Pixel Detector): Es el más cercano al punto de interacción y se compone de tres capas de píxeles de silicio. Además, proporciona una medición de seguimiento de alta precisión
- El Detector de Semiconductores (SCT, Semi-conductor Tracker): Es el segundo detector de alta precisión del detector interno ATLAS. Se compone de ocho capas de tiras de silicio.

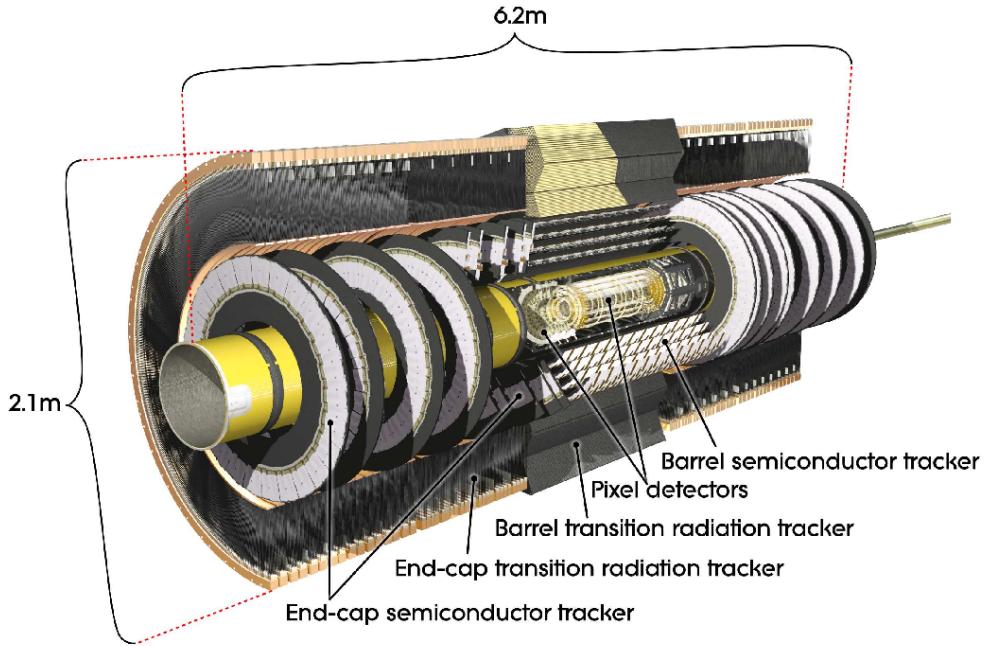


Figura 3.8: Estructura del Detector interno (ID) [71].

El detector de píxeles y el rastreador de semiconductores juntos proporcionan un promedio de ocho hits de alta precisión por pista.

- Detector de Radiación de Transición (TRT: Transition Radiation Tracker): Los dos primeros utilizan tecnología de semiconductores de silicio y este último hace uso de la ionización de distintos gases (X_e , CO_2 , O_2). A pesar de la baja resolución en comparación con los rastreadores de silicio y la falta de una medición en la dirección z, los golpes en el TRT contribuyen significativamente al reconocimiento de patrones y la resolución del momento debido a la gran cantidad de mediciones y a la mayor longitud de la pista medida.

3.2.4. Calorímetros

En los calorímetros, las partículas cargadas van a interactuar en procesos como la ionización y Bremsstrahlung, los fotones por otro lado van a interactuar con la materia por medio de la dispersión de Compton, efecto fotoeléctrico y producción de pares [73]. Los hadrones por medio de interacciones nucleares, y los neutrinos solamente interactúan de forma débil, no van a interactuar con la materia.

Los calorímetros en ATLAS tienen una forma cilíndrica centrada alrededor del punto de interacción con su eje en el eje z. La principal función de los calorímetros es medir la energía que depositan las partículas cuando pasan a través del detector así como la dirección y posición (η, ϕ) de éstas [73]. Por lo general, está diseñado para detener la totalidad o “absorber la energía” de la mayoría de las partículas provenientes de una colisión. Su estructura consiste en capas de material de alta densidad “pasivo” o “absorbente”, por ejemplo, plomo intercalado con capas de un medio “activo” como vidrio de plomo sólido o argón líquido [56].

Los calorímetros ATLAS se dividen en dos subsistemas distintos, el calorímetro electromagnético y el calorímetro hadrónico, y se representan en la Figura 3.11.

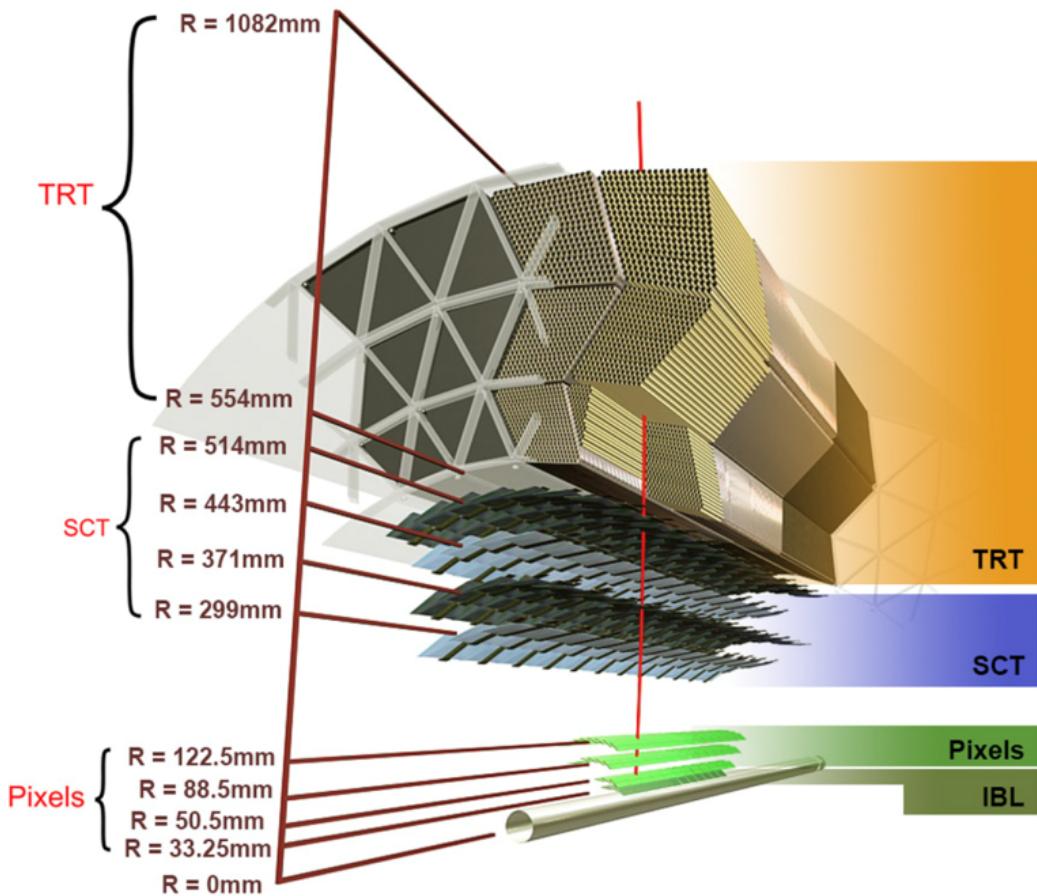


Figura 3.9: Detector Interno de ATLAS. Se muestran los tres subdetectores (el Detector de píxeles, el Semiconductor Tracker y el Transition Radiation Tracker). La trayectoria de las partículas cruza sucesivamente 3 capas de píxeles de silicio, cuatro capas de micro bandas de silicio y aproximadamente 36 tubos del detector de radiación de transición [72].

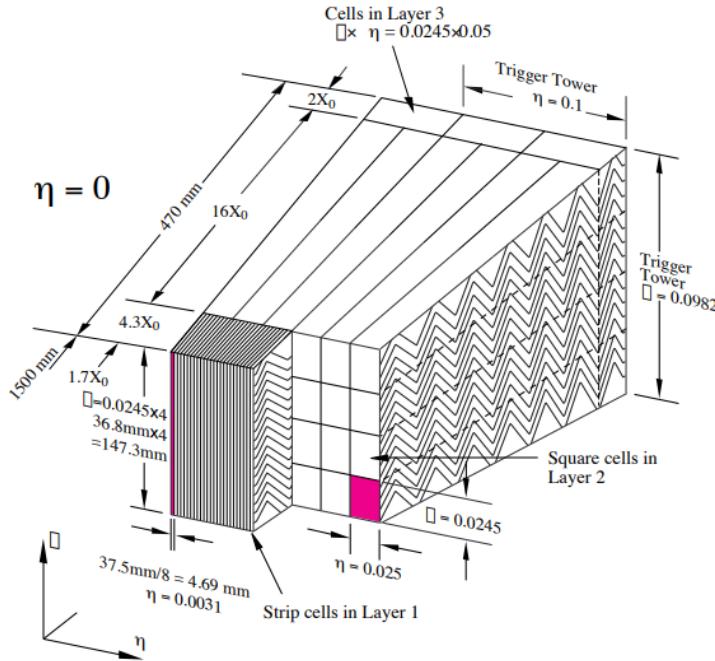


Figura 3.10: Diagrama del calorímetro de argón líquido ATLAS, que muestra la estructura del acordeón y la diferente granularidad en las diferentes capas. Diagrama tomado de [67].

El Calorímetro Electromagnético (EM por su nombre en inglés Electromagnetic Calorimeter): Su estructura es de acordeón como se observa en la figura 3.10, la cual permite medir la energía de electrones (positrones) y fotones a medida que interactúan con la materia. Los materiales que absorben esa energía son plomo y acero inoxidable, con argón líquido como material de muestreo, y se requiere un criostato alrededor del calorímetro EM para mantenerlo lo suficientemente fresco. Además, cubre la región $|\eta| < 3.2$ [67].

Los Calorímetros Hadrónicos (HTC por su nombre en inglés Hadronic Tile Calorimeter): Miden la energía de los hadrones (partículas que contienen quarks, como los protones y los neutrones) a medida que interactúan con los núcleos atómicos, el material que absorbe esa energía es el acero, con baldosas centelleantes que muestrean la energía depositada, además cubre la región de pseudorapidez de $|\eta| < 4.5$. Estos calorímetros pueden detener la mayoría de las partículas conocidas, excepto muones y neutrinos [67].

3.2.5. Espectrómetro de Muones

Los muones son partículas que generalmente pasan a través del detector interno y los calorímetros sin ser detectados. El espectrómetro de muones (MS por su nombre en inglés Muon Spectrometer), está ubicado en la parte más externa del detector ATLAS, su diseño está basado en la deflexión magnética de las trazas de muones en el gran toroide magnético superconductor de núcleo de aire, instrumentado con cámaras de trigger y de alta precisión, distribuidas separadamente, para la reconstrucción de trazas, así mismo poder medir el momento de los muones y las partículas altamente penetrantes en la región $|\eta| < 2.7$ [73]. El espectrómetro consta de tres partes; un campo magnético provisto por tres imanes toroidales, un conjunto de 1200 cámaras con alta precisión para la reconstrucción especializada de las trayectorias de este tipo de partículas

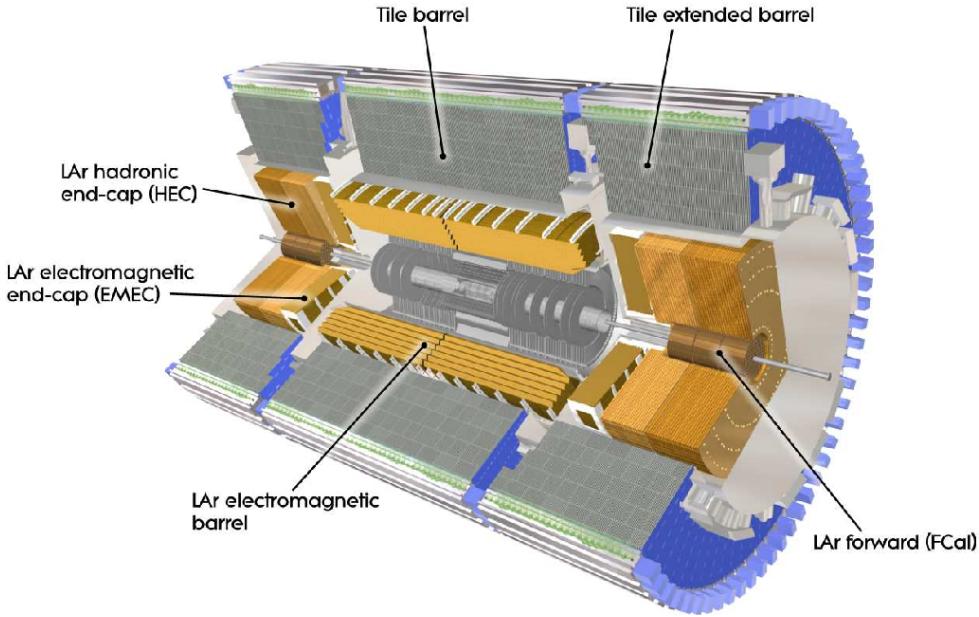


Figura 3.11: Estructura del Calorímetro Electromagnético y Hadrónico [73].

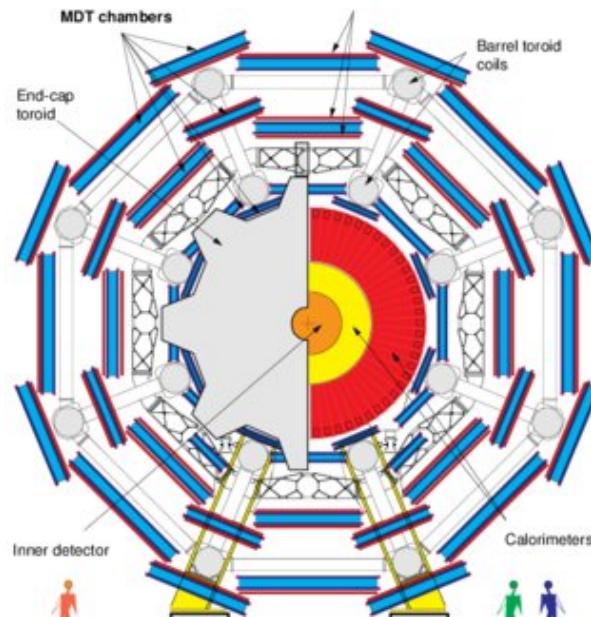


Figura 3.12: Se puede observar 3 capas concéntricas, cada capa consiste en cámaras MDT (Monitored Drift Tube). Las capas medias y externas además están equipadas con cámaras RPC (Resistive Plate Chambers) [74]. Las cámaras MDT proveen medidas de precisión para determinar el momentum y se encuentran en la región $|\eta| < 2$. Las cámaras RPC proveen el trigger para el sistema central de barril [75].

y un conjunto de cámaras de disparo con una precisa resolución de tiempo [73]. A su vez, este detector constituye un sistema de trigger para muones. Ver la figura 3.12.

3.3. El Trigger de ATLAS

El trigger de ATLAS está diseñado para inspeccionar rápidamente los eventos detectados por los subdetectores del mismo, su principal objetivo es seleccionar eficientemente aquellos eventos candidatos a involucrar física nueva e interesante, descartando los asociados al ruido o background [58]. Así como también elegir entre guardar o desechar los eventos, después de ser comparadas sus características principales con un conjunto de valores establecidos en el trigger. De esta forma si el trigger decide descartar un evento, este no se graba y se elimina permanentemente [58]. Para poder seleccionar los eventos de interés para su posterior almacenamiento, sin perder información relevante, se utilizan tres niveles de trigger, los cuales van a ser explicados a continuación [67].

La selección se refina en varios niveles, introduciendo también nuevos criterios de selección y combinando la información de diferentes subdetectores [67].

- Trigger Nivel - 1 (L1): El primer nivel está completamente basado en circuitos electrónicos especializados (Hardware) y utiliza únicamente el subconjunto de datos procedentes de los calorímetros y el espectrómetro de muones. Este nivel está diseñado para tomar una decisión sobre el evento en $2,5\mu s$ y una tasa de aceptación máxima de 100 kHz. Los resultados que se obtienen después de esos $2\mu s$ forman una lista llamada Regiones de Interés (RoI), estas son regiones del plano - del detector, en el que se han identificado actividades de interés y objetos causantes de la selección del evento. El trigger L1 busca p_T altos de muones, electrones, fotones, jets, entre otros.
- Trigger Nivel - 2 (L2): Este nivel está apoyado por el uso de redes de software para continuar el filtrado de los eventos que pasaron el primer nivel (L1) y toma como entrada las Regiones de Interés (RoI) proporcionados por el mismo refinando así la medición de estas regiones. Además, los datos de los diferentes subdetectores se combinan para obtener una mejor reconstrucción / identificación del objeto usando información ahora de todos los detectores, incluyendo el ID, el cual no fue usado en el L1. Un ejemplo de esto es que toma las trazas obtenidas del ID y MS para combinarlas y obtener así información sobre los muones, también la información del ID y los calorímetros son combinados para distinguir entre electrones y fotones. Este nivel está diseñado para tomar una decisión sobre el evento en 10 ms reduce la tasa de trigger a 3kHz.
- Filtro de Eventos (EF): Este nivel está completamente basado en software. En esta etapa, se realiza una reconstrucción completa del detector, donde la medición no se limita a los RoI, y los algoritmos ejecutados en el EF son en su mayoría algoritmos de reconstrucción offline, los que serán finalmente almacenados para su posterior análisis. La decisión del evento se toma en 1s y la tasa de trigger se reduce a 400Hz.

Las tasas de eventos ya han alcanzado cientos de millones de colisiones por segundo, lo que significa que los físicos deben filtrar decenas de petabytes de datos por año [56]. Y, a medida que mejora la resolución de los detectores, se necesita un software cada vez mejor para el procesamiento y el filtrado en tiempo real de los eventos más prometedores, produciendo aún más datos. La dimensión promedio de un evento en ATLAS es de alrededor de 1,6 MB por evento. Los dos triggers basados en software fueron fusionados en uno solo, el High Level Trigger (HLT), el efecto de esto es la redundancia y duplicación de código así como son disminuidos los procesos de recolección de data [67].

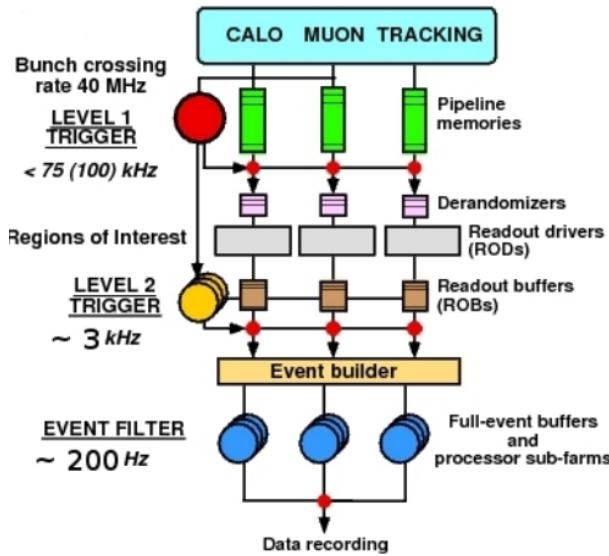


Figura 3.13: Estructura principal del sistema de disparo ATLAS: se compone de tres niveles, cada uno de los cuales mejora la medición de los niveles anteriores combinando también información de diferentes subdetectores [67].

3.4. Generadores de Monte Carlo en la Simulación de Eventos en ATLAS

Los eventos simulados son un ingrediente crucial para la medición de la física de partículas por varias razones. Por supuesto, es importante poder comparar cantidades medidas experimentalmente, como secciones transversales y distribuciones cinemáticas, con predicciones de la teoría. Además, los eventos simulados son importantes para calibrar la respuesta del detector y para estimar las eficiencias de selección con el fin de traducir los eventos observados a una cantidad física como una sección transversal.

La cadena de simulación ATLAS, que simula una amplia gama de procesos físicos de señal y de fondo que se esperan en el LHC. Dicha cadena presenta cuatro pasos bien definidos, que son [67]:

- Generación de Eventos: Corresponde a la fase de generación de eventos de colisión protón-protón (pp). Se encarga de la producción y descomposición de partículas en un proceso determinado. Hay varios generadores de eventos disponibles.
- Simulación de Detector: Es la simulación de interacciones entre las partículas generadas y el detector.
- Digitalización: Corresponde a la simulación de la lectura del detector, o mejor dicho, la conversión de la energía depositada en el detector en tiempos, corrientes y voltajes para la lectura electrónica. El formato de salida de la simulación es idéntico al formato de salida del detector real.
- Reconstrucción: En este paso se aplica un conjunto de algoritmos de reconstrucción de objetos. Estos algoritmos se aplican tanto a la simulación como a los datos reales exactamente de la misma manera. Esta fase se describe mejor en el Capítulo 4.

Las distribuciones cinemáticas para un proceso dado pueden diferir entre diferentes generadores Monte Carlo. Dependiendo del problema en estudio, un generador puede ser más adecuado que otros. Para el análisis de producción del DiBoson ZZ, la muestra de señal principal se genera con el método y marco POWHEG [76-78], con un generador de eventos de dibosón [79,80] utilizado para simular el proceso de producción de ZZ en el orden siguiente al líder (NLO).

La simulación de la lluvia de partón, del evento subyacente y de la hadronización se realiza con PYTHIA8 [81,82] utilizando el conjunto AZNLO de parámetros ajustados (tune) [83]. SHERPA [84-86] se utiliza para generar una muestra con el proceso iniciado por $q\bar{q}$ simulado en NLO para ZZ más cero o un chorro adicional y en el orden principal (LO) para dos o tres chorros adicionales, así como una muestra con el proceso iniciado por gg inducido por bucle simulado en LO con cero o un chorro adicional. Estos se utilizan para incluir la producción iniciada por gg inducida por bucle, que no se incluye en la muestra POWHEG + PYTHIA8, así como para estimar, mediante la comparación de las diversas muestras, una incertidumbre sistemática debido a la elección del generador de eventos [80-81]. Las funciones de distribución de partones (PDF) de CT10 NLO [87] se utilizan en la generación de eventos para todas las muestras anteriores. Se generan muestras adicionales para estimar la contribución de los eventos de fondo. Los eventos de Triboson se simulan con SHERPA, usando PDF CT10, y los eventos ttZ se simulan con MADGRAPH [88] interconectado con PYTHIA8 usando los PDF NNPDF 2.3 LO [89] y la melodía A14 [90].

En todas las muestras de MC (por sus siglas en inglés, Monte Carlo), las interacciones de pp adicionales que ocurren en el mismo cruce de racimos que la producción de ZZ, o en las cercanas, se simulan con PYTHIA8 con MSTW 2008 LO PDFs [91] y la sintonía A2 [92].

3.5. Simulación del Detector

La simulación de la respuesta del detector a varios procesos físicos es un ingrediente vital para realizar una medición física. Es importante para comprender y calibrar la respuesta del detector a diferentes firmas, para la estimación de aceptaciones y eficiencias, y también para poder comparar resultados experimentales con modelos teóricos. Se utiliza un simulador de detector personalizado, que utiliza el juego de herramientas Geant4 [93], para simular interacciones entre partículas y materia en el detector ATLAS. Toma como entrada la salida de los generadores Monte-Carlo. El simulador luego propaga las partículas a través del detector ATLAS, usando Geant4 para simular su interacción con el material del detector. También se simula la respuesta de los componentes del detector a las partículas producidas, y las señales de salida simuladas se producen en el mismo formato que la salida real del detector [18]. A continuación, se reconstruyen de la misma forma que para los datos reales (como se describe en el Capítulo 4 (reconstrucción de partículas)).

Capítulo 4

Reconstrucción de Objetos Físicos en ATLAS

En este capítulo se describen los algoritmos de software utilizados para reconstruir e identificar diferentes partículas. El primer paso en la reconstrucción de partículas implica reconstruir pistas en el detector interno y el espectrómetro de muones, identificar vértices de interacción e identificar grupos de depósitos de energía en los sistemas calorimétricos. Luego, estos se combinan para reconstruir partículas como electrones, muones, chorros, fotones y leptones tau, así como para medir las propiedades del evento, como la energía transversal faltante. Dado que la activación de electrones y muones es un componente crucial de las mediciones descritas en esta tesis, esto también se describe con más detalle aquí.

4.1. Calidad de los datos

El monitoreo de la calidad de los datos es un aspecto importante de cualquier experimento de física de altas energías o HEP por sus siglas en inglés High Energy Physics [18]. Debido a que los detectores en el LHC en general son equipos extremadamente sofisticados, se necesita una respuesta en línea de los datos que se están recogiendo y guardando, para evitar tener posteriormente datos de mala calidad, así como para mantener una buena línea de base, es decir una primera medición buena de todos los indicadores [94], para el análisis fuera de línea, cuando se inicie la reconstrucción [67].

Un paso importante para la búsqueda de nueva física es poder ser capaz de obtener un buen conjunto de datos de las colisiones. Ésta es la razón por la cual los cuatro experimentos del LHC mejoraron algunos de los sistemas de sus subdetectores, considerando las grandes cantidades de datos que se obtienen actualmente.

Esto hace que sea de suma importancia diseñar algoritmos y software para controlar la calidad de los datos grabados por cada experimento [18].

Para poder hacer seguimiento a las lecturas de los millones de canales de lectura y de las cambiantes condiciones del detector es importante el book-keeping [67], que es la asociación entre el sistema de adquisición de datos de ATLAS y las bases de datos que interactúan con él. La información sobre voltajes, torres calientes, canales muertos, temperaturas, intensidades de campo magnético y muchas otras variables capturan el estado actual del detector. Esta información es monitorizada mediante controles automatizados diseñados para detectar irregularidades y signos de problemas [67].

Cada uno de los detectores que conforman ATLAS tiene su propio sistema de monitoreo para encontrar problemas en tiempo real y dar una evaluación de su rendimiento. Se monitorean también los objetos reconstruidos: electrones, muones, jets, energía transversa perdida, entre otras cantidades, ya que estos objetos son sensibles al rendimiento del detector, y para poder ser reconstruidos se necesita información de diferentes subdetectores [67]. Los eventos en donde alguno de los subdetectores relevantes de ATLAS no este operando correctamente no van a poder ser usados para un posteriores análisis físicos.

En ATLAS, cada sistema está encargado de establecer su propia de calidad de datos e indicadores de integridad para cada Bloque de Luminosidad (LB). Esta información puede ser usada para crear una lista de LB la cual se puede usar para diferentes análisis, y esta lista tiene el nombre de Good Runs List [56].

En el párrafo anterior se introdujeron dos términos nuevos que requieren un poco más de explicación:

La Good Runs List (GRL) es una forma de seleccionar muestras de datos para análisis físicos. Hay varias formas en que los datos obtenidos no satisfagan las condiciones impuestas por la GRL, por ejemplo que el LHC no este en un modo de haz estable, que subdetectores estén apagados, que haya mucho ruido, etc. En su forma más simple, la GRL es un archivo .xml en el cual se especifica los conjuntos de buenos LB, sobre los cuales los datos deberían correr. Puede contener también listas negras, para excluir de forma explícita algunas secciones. Cada análisis tiene un GRL [95].

Un Bloque de Luminosidad (LB) es la unidad de tiempo para la toma de datos, y tiene una duración $\sim 2\text{min}$.

4.2. Reconstrucción de Trazas por el ID

Las partículas que atraviesan el detector interno viajan en una trayectoria aproximadamente helicoidal bajo la influencia del campo magnético, dejando impactos en los diversos componentes del detector que atraviesan [58]. Para utilizar estos impactos para identificar y medir partículas, es necesario reconstruir las pistas de partículas a partir de estos impactos, en un proceso conocido como seguimiento [18]. Una reconstrucción completa de la pista es necesario para la determinación de la masa invariante de las partículas y sus tiempos de vida media. En las energías de colisión y los niveles de choque en el LHC, normalmente habrá cientos de impactos en el detector interno. El algoritmo de seguimiento debe poder asociar correctamente los impactos con las pistas, así como reconstruir los parámetros de la pista. Además de interactuar con los elementos activos del detector, las partículas también interactuarán con el material muerto en el detector interno, lo que conducirá a dispersiones múltiples, pérdida de energía de ionización y, especialmente para los electrones, pérdida de energía de radiación de bremsstrahlung [18]. En [96] se ofrece una descripción detallada del seguimiento de ATLAS.

En ATLAS las trayectorias de las partículas son parametrizadas con cinco observables, obteniendo un vector en 5-dim $\tau = (d_0, z_0, \phi, \theta_0, q/p_T)$, estos parámetros son definidos desde el punto más cercano de la trayectoria de la partícula al punto en el cual ambos haces interactúan, éste no es necesariamente el centro geométrico del detector [67]. Se enumeran a continuación, y se puede ver en la figura (4.1) que representa cada parámetro [67]:

- d_0 : Parámetro de impacto radial

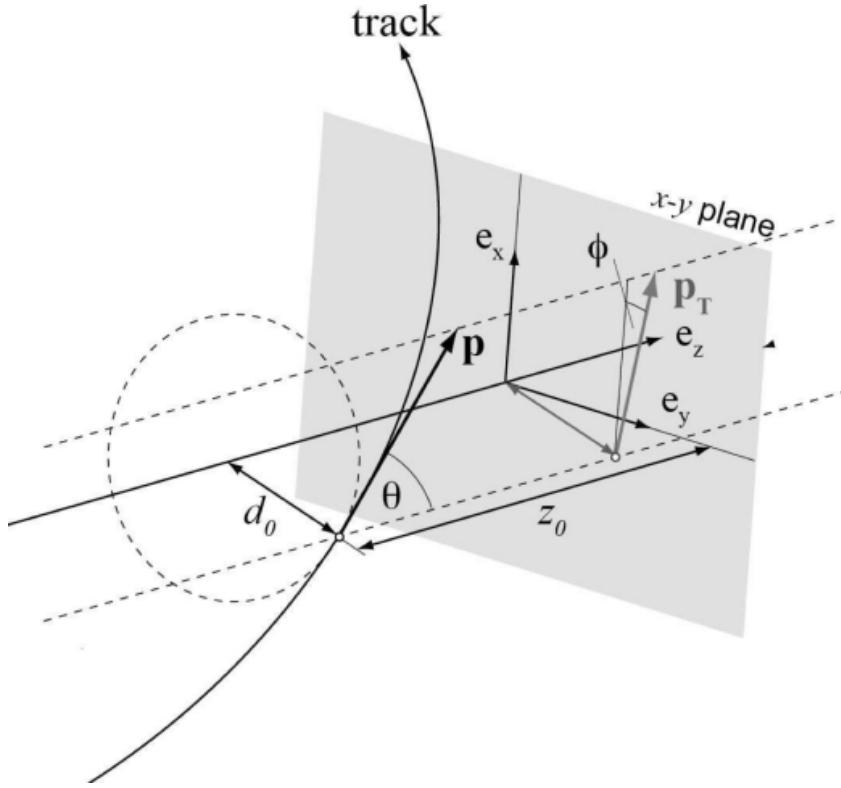


Figura 4.1: La trayectoria puede ser paramétrizada en un punto dado por medio de su posición, momento transverso y carga. La posición en cualquier otro punto puede ser calculado si el campo magnético y el material del detector es conocido. Tomado de [67]

- z_0 : Parámetro de impacto longitudinal
- ϕ : Ángulo azimutal
- θ_0 : Ángulo polar
- q/p_T : Carga dividida entre el momento transverso

La reconstrucción de la pista utiliza algoritmos de reconocimiento de patrones locales y globales para identificar mediciones que provienen de la misma partícula cargada. Las mediciones encontradas son utilizadas en el ajuste de las pistas para estimar los parámetros de estas mismas. La reconstrucción de las pistas en el ID consta de varias etapas que se enumeran a continuación [97]:

- Identificación del Track: Asignación del ID a las pistas encontradas como posibles candidatas. Se utilizan algoritmos de reconocimiento de patrones para poder asignar ciertos puntos o una serie de contactos (Ver Apéndice B) a una pista.
- Ajuste de la pista: Una vez que se identifican una serie de puntos que vienen del track se le hace un ajuste para determinar la trayectoria que sigue. Se determina en esta etapa los parámetros del track y sus errores. Son procesados de tal manera que se eliminan los candidatos a tracks de los contactos aleatorios, que se le conocen como falsos o duplicados, los cuales pueden ser identificados por medio de medidas que comparten con otros candidatos a track.

- Prueba de las Hipótesis a Tracks: Se verifica la calidad de los candidatos a tracks y superposición con otros candidatos.

En ATLAS, los pasos de búsqueda de pistas y de ajuste se combinan. Los impactos en el detector de píxeles y el SCT se transforman en puntos espaciales tridimensionales. Justo después, el algoritmo de adentro hacia afuera [98] entra en acción. Consiste en sumar los puntos del espacio tridimensional uno a uno alejándose del punto de interacción para formar un “camino”. Utiliza un filtro combinatorio de Kalman [99]. El algoritmo de filtro de Kalman agrega puntos espaciales tridimensionales de forma iterativa y se ajusta simultáneamente al candidato de pista. Se aplican recortes en la calidad del ajuste para eliminar pistas de mala calidad y evitar superposiciones con otras pistas candidatas. Las pistas seleccionadas se amplían en el TRT y se ajustan de nuevo para obtener los valores finales de los parámetros de la pista [97].

La pérdida de energía media en el material del detector, la dispersión múltiple, el efecto Bremsstrahlung y los cambios en el campo magnético a lo largo de la trayectoria de la pista se tienen en cuenta durante el proceso de ajuste de la pista [18]. El algoritmo de adentro hacia afuera proporciona la mejor eficiencia de reconstrucción de partículas cargadas primarias producidas directamente en una colisión de pp o de desintegraciones o interacciones de partículas con una vida útil corta ($< 3 \times 10^{-11} s$) [98].

Para reconstruir mejor las partículas cargadas secundarias, producidas en la interacción de las primarias (con una vida útil $> 3 \times 10^{-11} s$), o candidatos de conversión, se aplica un algoritmo de búsqueda de pistas adicional, llamado de afuera hacia adentro [98]. El proceso de búsqueda de pistas comienza con segmentos TRT no utilizados por el algoritmo de adentro hacia afuera. Luego se extienden al SCT y al detector de píxeles [98].

4.3. Hallazgo de Vértices

Uno de los principales objetivos de un detector de tracking es poder medir o estimar cuáles son los vértices de interacción [97]. La idea geométrica básica de un vértice es encontrar un punto en donde dos o más partículas son originadas. Los vértices ayudan entre varias cosas a los siguientes puntos: permiten identificar partículas, estimar la resolución o precisión de la posición de las partículas, estudiar o medir la cantidad de material que se tienen en las diferentes capas de los detectores, entre otros [97].

En un evento pueden estar presentes dos tipos diferentes de vértices [18]:

- Los vértices primarios (VP) que corresponden al punto de colisión de las partículas del haz. Se caracterizan por tener muchas pistas-partícula asociadas, por lo tanto una alta multiplicidad de pistas. En un evento seleccionado generalmente hay un VP de dispersión dura, mientras que el resto están asociados, por convenciones, a interacciones acumuladas.
- Los vértices secundarios corresponden a la desintegración de partículas de vida corta, que se desintegran a una distancia medible de los VP. La multiplicidad de pistas para los vértices secundarios es menor.

En general, la reconstrucción de los VP se realiza en tres pasos [100]:

- Identificación del vértice: Asignación de pistas reconstruidas a candidatos VP.

- Ajuste de vértice: Reconstrucción de la posición del VP, cálculo de su matriz de error, estimación de la calidad de ajuste y reajuste opcional de los parámetros de la pista asociada para restringirlos para que se originen en el VP correspondiente y no en el punto del haz.
- Prueba de hipótesis de vértice: Verificar la calidad de los candidatos a vértices y sus solapamientos con los otros candidatos.

Normalmente estas dos primeras etapas son indistinguibles, ambos ajustes después de la identificación, e identificación por el ajuste, son métodos implementados por ATLAS. Los tracks incompatibles con los candidatos de VP, que estén separados más de 7σ del vértice, son usados para buscar por nuevos VP [18]. Este procedimiento es repetido hasta que no queden tracks sin ser asociadas a un vértice o no se puedan encontrar vértices adicionales. A los vértices primarios se les requiere que tengan al menos dos tracks asociados. Los VP con mayor suma de p_T^2 de sus tracks, son escogidos como hard-scatter VP [67], lo cual es una etiqueta que se le coloca para diferenciar su mayor momento y se usa para posteriores cortes en los análisis de datos.

4.4. Reconstrucción e Identificación de Electrones

4.4.1. Reconstrucción de Electrones

La firma distintiva de un electrón es una pista curva en el detector interno y una lluvia estrecha en el calorímetro EM. Los fotones a diferencia de los electrones, se reconocen gracias a tener solamente la cascada producida en el calorímetro EM. Los electrones pueden ser producidos en interacciones de hard-scatter, pero también en jets y conversiones de fotones [105]. Se aplican varios criterios para la reconstrucción e identificación de electrones para determinar si un dado depósito de energía y track asociado fueron producidos en una interacción de hard-scatter o no [105]. Los electrones en la región central ($|\eta| < 2,47$) son reconstruidos usando algoritmos que combinan información del calorímetro EM y de ID, otros algoritmos reconstruyen electrones en la región directa utilizando solo información del calorímetro EM [105].

La reconstrucción de electrones en ATLAS en la región central se realiza mediante el algoritmo de ventana deslizante [101]. Busca “cluster” en la capa media del calorímetro EM con un total de $E_T > 2,5 \text{ GeV}$. La ventana utilizada para definir los “clusters” tiene un tamaño de 35 en unidades de celda de capa media ($\Delta\eta \times \Delta\phi = 0,025 \times 0,025$) [101]. Si los “clusters” sembrados coinciden con un par de pistas que se originan en un vértice de conversión de fotones reconstruido, se etiqueta como fotón convertido. De lo contrario, si el algoritmo puede hacer coincidir una pista del ID con el “cluster” sembrado, se etiqueta como candidato a electrón [105].

El emparejamiento se realiza en una ventana $\Delta\eta \times \Delta\phi$ de $0,05 \times 0,10$, para tener en cuenta las pérdidas bremsstrahlung [105]. Se requiere que el impulso de la pista sea compatible con la energía del grupo. En caso de que varias pistas coincidan con el grupo EM, se prefieren las pistas con impactos en los detectores de silicio y se elige la más cercana en ΔR [18]. Además, la información del TRT se puede utilizar para mejorar la separación de los candidatos a electrones de los piones. Los “clusters” finales se construyen alrededor de los “clusters” sembrados emparejados con una pista, al incluir todas las células de diferentes capas de calorímetro EM ubicadas dentro de un rectángulo centrado en la posición de la semilla [101]. El tamaño del rectángulo depende de la posición en el calorímetro de los “clusters” sembrados (barril o tapón). La energía del grupo se calibra a la escala de energía EM, que se derivó de correcciones basadas en MC para tener en

cuenta la pérdida de energía en material pasivo, mediciones de haz de prueba y mediciones de decaimientos $Z \rightarrow ee$ para la calibración final [102,103].

4.4.2. Identificación de Electrones

Los electrones se pueden distinguir de los hadrones ya que las lluvias EM depositan la mayor parte de su energía en la segunda capa del calorímetro EM [18]. El ancho de las lluvias de electrones es más estrecho que el de los hadrones [56]. La relación entre la energía transversal reconstruida en la primera capa del calorímetro hadrónico y la energía transversal reconstruida en el calorímetro EM, conocida como fuga hadrónica o R_{had} , es menor para los electrones que para los hadrones [105]. Para identificar electrones rápidos, se utiliza una identificación basada en cortes. Se reducen las variables relacionadas con la forma de la ducha electromagnética en el calorímetro, la calidad de la pista del detector interior, la coincidencia pista-calorímetro y la información de identificación de partículas del TRT [105].

En ATLAS, se utilizan tres conjuntos de cortes de referencia, denominados “Loose ++”, “Medium ++” y “Tight ++”, diseñados para proporcionar un rechazo de fondo progresivamente mayor, a costa de la eficiencia de la señal [104]. Los rechazos de chorro esperados (de la simulación) de los tres puntos son 500, 5000 y 50000 respectivamente [104].

Requisitos “Loose ++”

Tanto en 2011 como en 2012, la selección “Loose ++” realiza recortes en las variables de forma de ducha en la primera y segunda capas del calorímetro EM, fugas en el calorímetro hadrónico, calidad de pista en los detectores de silicio y coincidencia de grupos de pistas sueltas. Las variables cortadas son las siguientes [67]:

- **Formas de Ducha:** Se realizan cortes en las siguientes variables de forma de ducha que distinguen entre las duchas electromagnéticas que se originan a partir de electrones o fotones y las duchas hadrónicas que se originan a partir de partículas en chorros:
 - Fuga hadrónica, R_{had} , la relación entre E_T en la primera capa del calorímetro hadrónico y la E_T del grupo EM. En el rango $0,8 < |\eta| < 1,37$ se utiliza la E_T del calorímetro hadrónico completo para compensar la grieta entre el barril y el barril extendido del calorímetro de baldosas.
 - R_η , la relación de la energía en la capa media del calorímetro en 3×7 celdas a la energía en 7×7 celdas centradas en la posición del grupo de electrones.
 - $w_{\eta 2}$, el ancho lateral de la ducha en la segunda capa del EM calorímetro.
 - E_{ratio} , la relación de la diferencia de energía entre el depósito de energía más grande y el segundo más grande en la primera capa del calorímetro EM con la suma de sus energías.
 - w_{stot} , el ancho total de la ducha en la primera capa del calorímetro EM.
- **Silicon Hits:** Al menos 7 hits en los detectores de silicio, de los cuales al menos uno debe estar en el detector de píxeles. Esto asegura una buena calidad de pista y rechaza fondos de conversiones o decaimientos como $\pi^0 \rightarrow e^+e^-$.

- **Coincidencia Pista-Grupo:** Se aplica una coincidencia flexible en η , que requiere $\Delta\eta < 0,015$; esto asegura que la pista y el grupo se originan a partir de la misma partícula física y rechaza los fondos de las falsificaciones combinatorias.

Las variables de forma de ducha R_η y R_{had} son particularmente susceptibles de acumularse, ya que toman muestras de una gran área del calorímetro [18]. Por lo tanto, los recortes en estas variables se relajaron en 2012 con respecto a 2011, reduciendo el poder de rechazo de “Loose ++” en aproximadamente un 20 % [18].

Requisitos “Medium ++”

Todos los cortes “Loose ++” deben aprobarse y, además [18]:

- **Shower Shapes:** Los cortes en forma de ducha realizados en “Loose ++” (R_η , R_{had} , $w_{\eta 2}$, E_{ratio} , w_{stot}) se hacen más ajustados. Para 2012, los cortes en R_η y R_{had} se relajaron con respecto a 2011, y se hicieron al mismo valor que para “Loose ++”, mientras que los cortes en $w_{\eta 2}$, E_{ratio} y w_{stot} se endurecieron con respecto a 2011.
- **Coincidencia Track-Cluster:** Se aplica una coincidencia más estricta en η , lo que requiere $\Delta\eta < 0,005$.
- **Parámetro de Impacto:** Requiere que la trayectoria del electrón tenga un parámetro de impacto transversal $|d_0| < 5mm$; esto rechaza los fondos de los electrones que se originan en la desintegración de los hadrones en los chorros.
- **Golpes de Silicio:** Se imponen requisitos más estrictos sobre los golpes en los detectores de silicio. Se requiere que haya al menos un impacto en la capa b para $|\eta| < 2,01$ ($|\eta| < 2,37$ en 2012). En 2011, se requiere al menos 1 acierto de Pixel para $|\eta| < 2,01$ y al menos dos para $|\eta| > 2,01$. Para 2012, se requieren dos aciertos de Pixel en todos los contenedores
- **Fracción en la tercera capa del calorímetro f_3 :** Para 2012 se agregó un corte en la fracción de la energía de la ducha depositada en la tercera capa del calorímetro EM para compensar el aflojamiento de los cortes en la primera capa del calorímetro. Este corte solo se aplica para $E_T < 80 GeV$, ya que la profundidad de la ducha EM y, por lo tanto, la fuga hacia la tercera capa aumenta con la energía.
- **Impactos de umbral alto de TRT** Se establece un requisito impreciso sobre la fracción de impactos de umbral alto (HT) de los fotones de radiación de transición en el detector de TRT (consulte la Sección 3.2.3)

Requisitos “Tight ++”

Todos los cortes de “Medium ++” deben aprobarse y, además [18]:

- **Formas de Ducha:** Los cortes en las variables de forma de ducha se realizan a valores iguales o más ajustados a los de “Medium ++”. En cuanto a “Medium ++”, los recortes en R_η y R_{had} se aflojaron en 2012 con respecto a 2011, mientras que se endurecieron los recortes en $w_{\eta 2}$, E_{ratio} y w_{stot} .

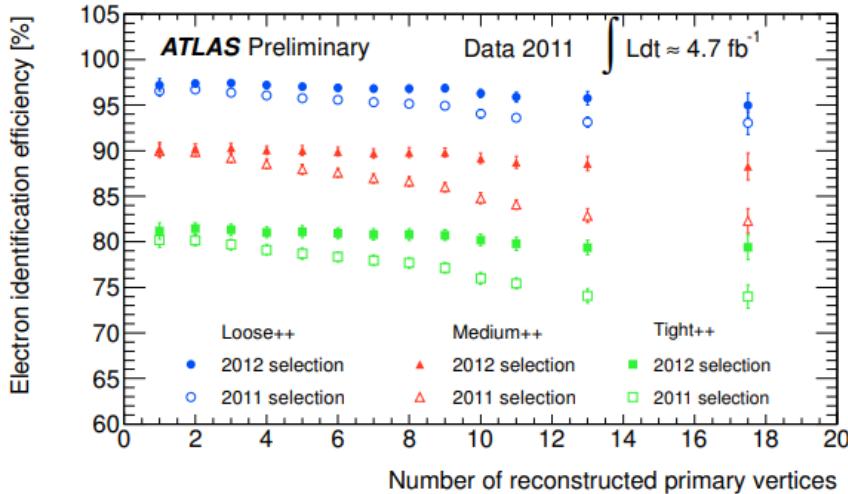


Figura 4.2: Eficiencia de identificación de electrones en 2011 (marcadores abiertos) y 2012 (marcadores sólidos) en función del número de vértices primarios reconstruidos en el evento. Los círculos azules muestran la eficiencia para la selección “Loose ++”, los triángulos rojos para “Medium ++” y los cuadrados verdes para “Tight ++”. Figura de [105].

- **Emparejamiento de Pista-Grupo:** se agrega un emparejamiento de clúster en ϕ , lo que requiere $\Delta\phi < 0,02$, y se realizan recortes en la relación entre la energía del clúster y el impulso de la pista, E/p .
- **Parámetro de Impacto:** El corte del parámetro de impacto transversal se aprieta a $|d_0| < 1mm$.
- **Impactos de Silicio:** Se imponen requisitos más estrictos sobre los impactos en los detectores de silicio, que requieren que haya al menos un impacto en la capa b para todo η y, en 2012, al menos 2 impactos en el detector de píxeles para todo η .
- **Rechazo de conversión:** Los candidatos que coinciden con las conversiones de fotones reconstruidos son rechazados.

Eficiencias de Identificación de Electrones

La Figura 4.2 muestra la eficiencia de identificación de electrones en 2011 y 2012, medida de manera similar a la eficiencia de reconstrucción descrita anteriormente, en función del número de vértices primarios reconstruidos en el evento. Usando los requisitos de identificación de 2011, la eficiencia puede caer más del 5 % en eventos con 18 vértices primarios reconstruidos con respecto a la eficiencia en eventos con un solo vértice primario. La figura 4.3 muestra la eficiencia de identificación de “Loose ++” en función de E_T , utilizando los requisitos de 2011. Se observan diferencias al nivel de un pequeño porcentaje entre la eficiencia medida en los datos y la eficiencia medida en la simulación de Monte Carlo. Esto se atribuye principalmente a un modelo erróneo de las variables de forma de la ducha en el Monte Carlo. Los factores de escala, parametrizados en función de η y E_T , se aplican al Monte Carlo para corregir la eficiencia de reconstrucción a la observada en los datos [18].

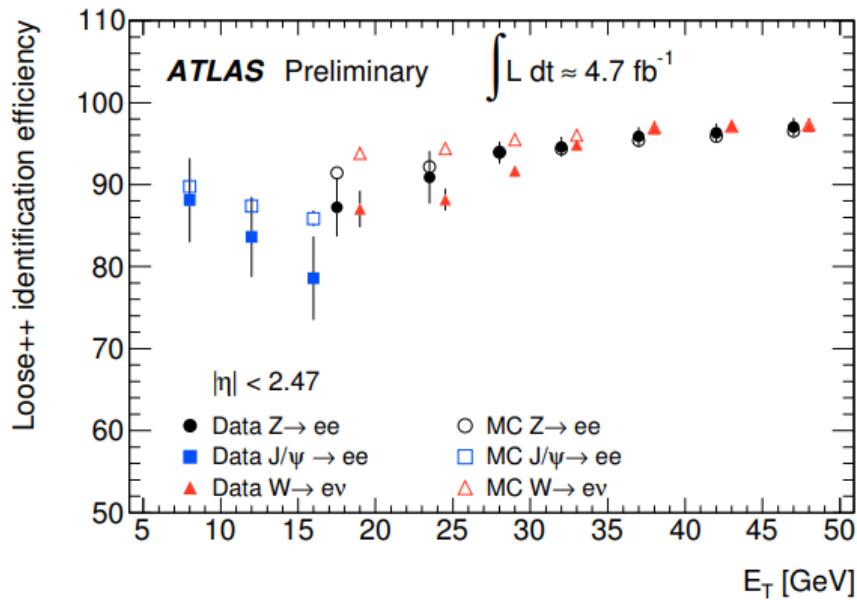


Figura 4.3: Eficiencia de los requisitos de identificación de “Loose ++” en función de la energía transversal del clúster. Los puntos sólidos indican mediciones basadas en datos, mientras que los puntos abiertos indican predicciones de Monte Carlo. Los diferentes marcadores indican el método utilizado para medir la eficiencia: los resultados de la etiqueta $Z \rightarrow ee$ y la sonda se muestran como círculos negros, los resultados de $J/\psi \rightarrow ee$ como cuadrados azules y los resultados de $W \rightarrow e\nu$ como triángulos rojos. Figura de [106].

4.5. Muones

4.5.1. Reconstrucción e Identificación de Muones

La reconstrucción de muones se basa en información procedente de mediciones precisas en el espectrómetro de muones y el detector interno [106]. Su objetivo es reconstruir muones con alta eficiencia en un amplio espectro de p_T (desde pocos GeV hasta la escala TeV). En ATLAS se utilizan cuatro categorías de muones, cada una optimizada para una necesidad diferente [106]:

- **Combinado:** Combinación de una pista MS (por sus siglas en inglés, muon spectrometer) con una pista ID. En general limitado por la aceptación de la identificación, $|\eta| < 2,5$, pero debido a la naturaleza complicada del campo magnético, es posible formar muones combinados más allá de este.
- **Segmento Etiquetado:** Combinación de una pista de identificación con un segmento de pista MS. Un segmento de vía MS es un segmento de vía recta reconstruido en una única estación MS (como se define en la Sección 3.2.5) donde el segmento no formó una vía MS completa. Los parámetros de la pista del muón reconstruido se toman únicamente de la pista de identificación.
- **Stand-Alone:** Una reconstrucción de la pista de muones basada únicamente en mediciones de MS. Posible con la plena aceptación del EM, $|\eta| < 2,7$.
- **Etiquetado por Calorímetro:** las pistas de identificación se etiquetan como que se originan a partir de muones al compararlas con depósitos de calorímetro consistentes con un muón ionizante mínimo. No se utiliza información de MS.

Dado que los muones combinados tienen una pista completamente reconstruida tanto en el ID como en el MS, son el tipo de muón preferido y tendrán la mejor resolución de parámetros de pista. El ID proporciona la mejor medición del impulso en momentos bajos a intermedios, mientras que el MS proporciona la mejor medición en un p_T más alto (aproximadamente para $p_T > 100 \text{ GeV}$) [106]. La combinación con una pista de ID mejora la resolución del impulso en el rango $4 < p_T < 100 \text{ GeV}$ [106]. Los muones marcados con segmentos son útiles para recuperar la eficiencia a p_T bajo, donde los muones solo pueden alcanzar la capa interna de la cámara de muones y en regiones de aceptación limitada del detector. Los muones autónomos son útiles para extender la cobertura más allá de la cobertura del ID. Los muones marcados con calorímetro sufren grandes tasas de falsificación de chorros y electrones, pero se pueden usar para recuperar la aceptación en $|\eta| < 0,1$ donde hay espacios en las cámaras de muones para proporcionar espacio para los servicios del ID, el imán de solenoide y los calorímetros [106].

Hay dos cadenas de reconstrucción de muones paralelas en uso en ATLAS, STACO y MUID [107]. Cada uno utiliza algoritmos de búsqueda de pistas ligeramente diferentes y enfoca la combinación de pistas ID y MS de diferentes maneras. MUID realiza un reacondicionamiento global de aciertos en el MS y el ID, mientras que STACO intenta una combinación estadística de las dos mediciones de la pista, ponderando las contribuciones relativas según sus matrices de covarianza [107]. Se ha descubierto que las dos cadenas ofrecen un rendimiento similar.

Eficiencias de reconstrucción

La Figura 4.4 muestra la eficiencia de reconstrucción observada para los muones reconstruidos como Combinados o Etiquetados por Segmentos en los datos de 2011, medidos usando una

técnica de etiqueta y sonda en decaimientos del bosón Z [108], así como la eficiencia predicha por la simulación de Monte Carlo. Se considera que la eficiencia disminuye significativamente para $|\eta| < 0,1$ por las razones explicadas anteriormente. También se muestra la eficiencia de los muones marcados con calorímetro, y se ve que recuperan eficazmente la pérdida de eficiencia en $|\eta| < 0,1$ [108]. Se considera que la eficacia de la reconstrucción de muones es casi constante en función de p_T . Tanto en p_T como en η se observa una buena concordancia entre la simulación de Monte Carlo y los datos, pero al igual que con los electrones, se aplican factores de escala al Monte Carlo para reproducir la eficiencia observada en los datos. La tasa de identificación errónea de la carga de muones es insignificante [108].

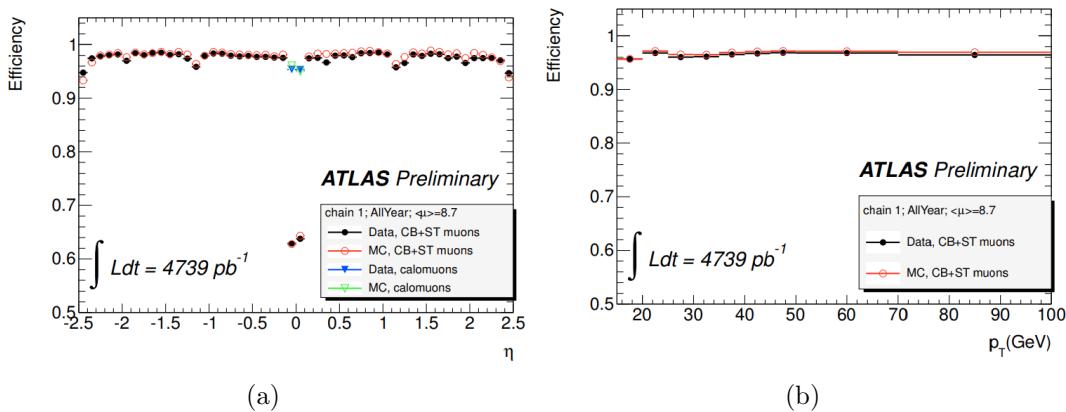


Figura 4.4: Eficiencia de reconstrucción de muones en 2011 en función de (a) la pseudorapidez y (b) el momento transversal del muón para los muones reconstruidos como Combinados o Etiquetados por Segmentos utilizando el algoritmo STACO. Los puntos negros sólidos muestran la eficiencia observada en los datos y los círculos rojos abiertos muestran la eficiencia predicha por la simulación de Monte Carlo. En la figura (a) también se muestran las eficiencias de los muones marcados con calorímetro para $|\eta| < 0,1$ (triángulos azules sólidos para datos y triángulos verdes abiertos para Monte Carlo). Tomado de [108].

Capítulo 5

Selección de Objetos y Eventos

En este capítulo se describen los requisitos de selección de electrones, muones y niveles de eventos utilizados para seleccionar eventos ZZ. Después de describir los requisitos de selección, se discuten las eficiencias y aceptación de la selección y se dan los rendimientos de eventos esperados. Finalmente, se consideran las incertidumbres sistemáticas asociadas con la aceptación de la reconstrucción.

5.1. Selección de Electrones

La reconstrucción e identificación de electrones se describió en la Sección 4.4. Los electrones centrales (con $|\eta| < 2,47$) se utilizan para dicho análisis. Se imponen requisitos de selección adicionales a los descritos en la Sección 4.4 para seleccionar los electrones que probablemente se hayan originado a partir de la desintegración del bosón Z y para rechazar los fondos. Los requisitos de selección de electrones se resumen en la Tabla 5.1 y se describen con más detalle a continuación.

5.1.1. Selección de Electrones Centrales

Electrones centrales, con $|\eta_{cluster}| < 2,47$, se reconstruyen utilizando el algoritmo de ventana deslizante como se describe en la Sección 4.4.1. Posteriormente se utiliza el algoritmo que reajusta las pistas utilizando un filtro de suma gaussiana (GSF) [18] para tener en cuenta el efecto de bremsstrahlung en el detector interno. Se requieren electrones centrales para pasar los requisitos de identificación de “Loose ++”, descritos en el capítulo anterior.

Para los candidatos a electrones con 4 o más impactos en la capa de silicio (SCT y Pixel), la energía del electrón se toma de la medición del grupo, y η y ϕ se toman de la medición de la pista (este requisito se satisface automáticamente para los electrones que pasan los requisitos de identificación “Loose ++”) [105]. Para los candidatos a electrones con menos de 4 impactos de silicio, todos los parámetros de los electrones se toman directamente del grupo [105]. En ambos casos, el grupo η y ϕ se utilizan para el requisito de η y para la eliminación de superposiciones. Usando la energía y la dirección definidas de esta manera, se requiere que los candidatos a electrones tengan $E_T > 7 \text{ GeV}$, donde E_T se define como $E \sin(\theta) = E / \cosh(\eta)$ [33].

Los electrones se rechazan si caen en una región de η , ϕ espacio consistente con la presencia de un pequeño número de placas frontales del calorímetro LAr muerto en la primera o segunda capa de muestreo [105].

Tabla 5.1: Requisitos de selección de electrones.

| Requisito | 13TeV |
|------------------------------|---|
| Algoritmo | Estándar (con reacondicionamiento GSF) |
| Calidad | Buena calidad de datos |
| Selección ID | “Loose ++” |
| η | $ \eta_{cluster} < 2,47$ |
| E_T | $E_T > 7\text{GeV}$ |
| z_o | $ z_o \sin(\theta) < 0,5\text{mm}$ |
| $d_o/\sigma(d_o)$ | $ d_o/\sigma(d_o) < 5$ |
| Aislamiento de Pistas | $\sum p_T(\Delta R < 0,2)/p_T < 0,15$ |
| Aislamiento del Calorímetro | $\sum E_T^{\text{calo}}(\Delta R < 0,2)/p_T < 0,3$ |
| Eliminación de superposición | Eliminar el electrón (e) con E_T mas bajo en $\Delta R < 0,1$ |

Para asegurar que los candidatos provienen del vértice primario, se aplican requisitos al parámetro de impacto longitudinal z_o de la pista de electrones con respecto al vértice primario. Para datos de 13 TeV, el requisito es $|z_o \sin(\theta)| < 0,5\text{ mm}$ [33]. Se utiliza el parámetro de impacto insesgado, obtenido reajustando el vértice sin la pista en cuestión, luego calculando el parámetro de impacto con respecto a este vértice reajustado, eliminando así el tirón de la pista del ajuste del vértice. El parámetro de impacto transversal d_o debe tener un significado (d_o dividido por el error en su medición, $d_o/\sigma(d_o)$) de magnitud menor que 5 [33]. Esto ayuda a reducir los fondos, particularmente desde la desintegración de piones y kaones en chorros a electrones, ya que estos decaimientos ocurren más lejos del punto de interacción dando parámetros de impacto más grandes [29].

Los requisitos de aislamiento se aplican para reducir los fondos de chorros identificados erróneamente como electrones, o de electrones de desintegraciones en chorros [105]. Las pistas de tales objetos normalmente estarán rodeadas por muchas otras pistas, y estarán rodeadas por grandes depósitos de energía en el calorímetro de otras partículas en el chorro. Se imponen requisitos de aislamiento de pistas, que requieren que la suma del p_T de las pistas que rodean la pista del electrón en un cono de $\Delta R < 0,2$ sea inferior al 15 % del p_T del electrón [33]. Las pistas incluidas en el cálculo deben tener $p_T > 0,4\text{ GeV}$, tener parámetros de impacto consistentes con el mismo vértice que el electrón y tener al menos 9 impactos de silicio, lo que garantiza una buena calidad y pureza de la pista [105]. El requisito del parámetro de impacto significa que esta variable es insensible al pileup ya que se excluyen la mayoría de las pistas de otros vértices (pileup) [18].

Además, se impone un requisito de aislamiento del calorímetro. Esto se define como la suma de las energías de las células calorimétricas en un cono de $\Delta R < 0,2$ alrededor del baricentro del grupo de electrones, excluyendo una cuadrícula de células de 5×7 en el centro del cono (que se supone que se deben al electrón) [105]. El requisito es que la suma de dichas energías sea inferior al 30 % de la E_T de electrones [105]. Esta variable es particularmente sensible al pileup, ya que los eventos pileup tienden a depositar energía adicional de forma isotrópica en todo el detector, aumentando así la energía en el cono de aislamiento [18]. Si dos electrones seleccionados se superponen dentro de $\Delta R < 0,1$, el electrón de E_T inferior se elimina, aunque en el caso de un electrón central que se superpone con un electrón delantero, este último se elimina ya que el electrón central tiene prioridad [105].

5.2. Selección de Muones

La reconstrucción de muones se describió en la Sección 4.5. En este análisis se utilizan tres categorías distintas de muones: Muones “centrales” con $|\eta| < 2,5$, y muones “hacia adelante (directo)” con $2,5 < |\eta| < 2,7$, ambos reconstruidos con el algoritmo STACO [18,33]; y muones marcados por calorímetro con $|\eta| < 0,1$, reconstruido sin utilizar el espectrómetro de muones [33]. Los requisitos de selección de muones se resumen en la Tabla 5.2 y se describen con más detalle a continuación.

5.2.1. Selección de Muones Centrales

Los muones centrales pueden estar combinados (con una pista de ID y MS completa) y tener $p_T > 7 \text{ GeV}$ y $|\eta| < 2,47$ [29].

En cuanto a los electrones, se imponen requisitos de aislamiento de pistas y calorímetros para rechazar los fondos. En el caso de los muones, esto se debe principalmente a la fuerte desintegración del sabor en los chorros y la “penetración” de los hadrones en el espectrómetro de muones [29]. El aislamiento de la pista se define de la misma manera que para los electrones. La variable de aislamiento del calorímetro se define de manera similar a la de los electrones, sin embargo, el tamaño del depósito de energía del “núcleo” atribuido al muón es más pequeño que en el caso del aislamiento de electrones, y varía según el componente del calorímetro dependiendo del muón esperado [107].

Para asegurar que los candidatos provienen del vértice primario, la magnitud del parámetro de impacto longitudinal con respecto al vértice primario, $|z_o \sin(\theta)| < 0,5 \text{ mm}$ [33]. La magnitud del significado del parámetro de impacto transversal, $|d_o/\sigma(d_o)|$ se requiere que sea inferior a 3 [33]. El requisito de $|d_o/\sigma(d_o)|$ para los muones es más estricto que para los electrones, ya que debido a que los muones no emiten bremsstrahlung en la misma medida que los electrones, se mejora el ajuste de la pista y, por lo tanto, la resolución del parámetro de impacto, por lo que se pueden establecer requisitos más estrictos, manteniendo una alta eficiencia de señal [29].

Las pistas internas del detector asociadas con los muones centrales deben tener un número mínimo de impactos en cada subdetector de silicio para garantizar una buena calidad de la pista [107]. Para todas las condiciones de impacto de silicio, los sensores muertos cuentan como impactos observados, no como agujeros [107].

5.2.2. Selección de Muón Directo

Se requiere que los muones en la región $2,5 < |\eta| < 2,7$ tengan una pista de espectrómetro de muones completa con aciertos en cada estación del espectrómetro y con $p_T > 10 \text{ GeV}$ [29]. Aunque el detector interno solo se extiende hasta $|\eta| = 2,5$, la naturaleza complicada del campo magnético hace posible que algunos muones hasta $|\eta| < 2,6$ para combinarse con una pista de ID, formando muones combinados [107]. Dado que no todos los muones directos tienen pistas, no se aplica un requisito de aislamiento de pistas.

Para los muones directos con una pista de ID, los cortes de calidad se aplican al número de impactos en los detectores de silicio, similares a los cortes para los muones centrales pero ligeramente más flojos [107]. No se aplica ningún corte al número de aciertos de TRT, ya que las huellas de muones directos generalmente están fuera de la aceptación de TRT [106]. Se requieren

Tabla 5.2: Requisitos de selección de muones.

| Requisito | 13TeV |
|--|--|
| Muones centrales | |
| Algoritmo | STACO |
| Tipo | Combinado o Etiquetado por Segmento |
| η | $ \eta < 2,5$ |
| p_T | $p_T > 7GeV$ |
| Calidad de la pista de identificación | |
| - Pixel Hits | ≥ 1 |
| - SCT Hits | ≥ 5 |
| - TRT | <p>if $0,1 < \eta < 1,9$: Require $\eta_{hits} + \eta_{outliers} > 5$ y $\eta_{outliers}/(\eta_{hits} + \eta_{outliers}) < 0,9$</p> <p>if $\eta > 1,9$ o $\eta < 0,1$: if $\eta_{hits} + \eta_{outliers} > 5$ require $\eta_{outliers}/(\eta_{hits} + \eta_{outliers}) < 0,9$</p> |
| z_o | $ z_o \sin(\theta) < 0,5mm$ |
| $d_o/\sigma(d_o)$ | $ d_o/\sigma(d_o) < 3$ |
| Aislamiento de Pistas | $\sum p_T(\Delta R < 0,2)/p_T < 0,15$ |
| Aislamiento del Calorímetro | $\sum E_T^{calo}(\Delta R < 0,2)/p_T < 0,3$ |
| Eliminación de superposición | Eliminar el electrón (e) con E_T mas bajo en $\Delta R < 0,1$ |
| Muón Directo | |
| Algoritmo | STACO |
| Tipo | Combinado o Etiquetado por Segmento |
| η | $2,5 < \eta < 2,7$ |
| p_T | $p_T > 10GeV$ |
| Calidad de la pista de identificación | |
| - Pixel Hits | ≥ 1 |
| - SCT Hits | ≥ 3 |
| Aislamiento del Calorímetro | $\sum E_T^{calo}(\Delta R < 0,2)/p_T < 0,15$ |
| z_o | $ z_o \sin(\theta) < 0,5mm$ |
| $d_o/\sigma(d_o)$ | $ d_o/\sigma(d_o) < 3$ |
| Muones Etiquetado por Calorímetro | |
| Tipo | Calorímetro etiquetado |
| η | $ \eta < 0,1$ |
| p_T | $p_T > 20GeV$ |
| Eliminación de superposición | Eliminar si se superpone con un muón central en $\Delta R < 0,1$ |

muones directos con una pista de ID para satisfacer los mismos requisitos de z_o y $d_o/\sigma(d_o)$ que los muones centrales [29].

5.2.3. Selección de Muones Etiquetado por Calorímetro

Los muones marcados con calorímetro se utilizan para recuperar la eficiencia en $|\eta| < 0,1$ donde hay un espacio en el espectrómetro de muones [107]. Debido al mayor nivel de falsificaciones de chorros y electrones a un p_T bajo, y la dificultad de medir la eficiencia de identificación a un p_T bajo, el requisito de p_T para los muones marcados con calorímetro es $p_T > 20 \text{ GeV}$ [33]. Deben pasar el corte o los requisitos de identificación de muones marcados con calorímetro basados en probabilidad descritos en la Sección 4.5.1, y no se seleccionan si se superponen con un muón central seleccionado dentro de $\Delta R < 0,1$. Los muones marcados con calorímetro deben satisfacer los mismos requisitos de pista y aislamiento calorimétrico, calidad de pista ID, z_o y $d_o/\sigma(d_o)$ que los muones centrales [107]. Para el análisis en estudio los eventos se ignoran si más de un muón seleccionado está etiquetado por calorímetro o es independiente [33].

5.3. Selección de Trigger

Los eventos candidatos se preseleccionan mediante un disparador de un solo muón o dielectrón [33]. Al menos uno de los leptones seleccionados en los eventos activados debe coincidir con el objeto que provocó el disparo del activador con $\Delta R < 0,1$ [33]. Dichos leptones se conocen como leptones emparejados por disparadores. Se requiere que los electrones acoplados por disparador en los datos de 13 TeV tengan un $p_T > 20 \text{ GeV}$ y $|\eta| < 2,47$ [29]. Se requiere que al menos un muón en el canal de 4μ debe satisfacer $|\eta| < 2,4$ lo que corresponde a la aceptación del disparador del muón. Los otros muones deben satisfacer $|\eta| < 2,7$ con $p_T > 20 \text{ GeV}$ [29].

5.4. Selección de Eventos ZZ

5.4.1. Definiciones de candidatos Z

Al referirse a los pares de di-leptones candidatos Z en la selección de dibosones ZZ, se utilizan las siguientes definiciones [18]:

- **Líder y sublíder:** El candidato Z con el p_T más grande de los dos se denomina Z líder; el candidato Z con el p_T más bajo se conoce como el sublíder Z.
- **Primario y secundario:** El candidato Z más cercano a la masa del bosón Z (de [32]) se conoce como el Z primario; el otro candidato se denomina Z secundario.

5.4.2. Requisitos de Selección de Eventos

Los eventos ZZ candidatos se seleccionan imponiendo los siguientes requisitos [33]:

- **Calidad de los datos:** Se requiere que los eventos pasen una “Lista de ejecución correcta” para eliminar los eventos que afectaron la calidad de los datos.
- **Disparador:** El evento debe pasar un disparador de un solo muón o dielectrón con un p_T dado como se describe en la Sección 5.3.

```

//Preselection cut for electron/muon trigger
if(trigE || trigM)
{
    //

    // Preselection of good leptons
    int goodlep_index[4];
    int goodlep_n = 0;
    int lep_index = 0;

    for(unsigned int i=0; i<lep_n; i++)
    {

        TLorentzVector leptemp; leptemp.setPtEtaPhiE(lep_pt->at(i)/1000., lep_eta->at(i), lep_phi->at(i), lep_E->at(i)/1000.);

        // Lepton is Tight
        if( lep_isTightID->at(i) )
        {
            // Lepton is isolated and with at least 20 GeV
            if( lep_pt->at(i) >2000. && ( (lep_ptcone30->at(i)/lep_pt->at(i)) < 0.15 ) && ( (lep_etcone20->at(i) / lep_pt->at(i)) < 0.15 ) )
            {
                if ( lep_type->at(i)==11 && TMath::Abs(lep_eta->at(i)) < 2.47 ) {
                    if( TMath::Abs(lep_tracksd0pvunbiased->at(i))/lep_tracksigd0pvunbiased->at(i) < 5 && TMath::Abs(lep_z0->at(i)*TMath::Sin(leptemp.Theta())) < 0.5 ) {
                        goodlep_n = goodlep_n + 1;
                        goodlep_index[lep_index] = i;
                        lep_index++;
                    }
                }
                // muon selection
                if ( lep_type->at(i) == 13 && TMath::Abs(lep_eta->at(i)) < 2.5 ) {
                    if( TMath::Abs(lep_tracksd0pvunbiased->at(i))/lep_tracksigd0pvunbiased->at(i) < 3 && TMath::Abs(lep_z0->at(i)*TMath::Sin(leptemp.Theta())) < 0.5 ) {
                        goodlep_n = goodlep_n + 1;
                        goodlep_index[lep_index] = i;
                        lep_index++;
                    }
                }
            }
        }
    }
}

```

Figura 5.1: Corte de preselección para disparador de Electrones/Muones [109].

- **Cuatro Leptones:** El evento debe tener exactamente cuatro electrones o muones que superen los requisitos de selección descritos en las secciones 5.1 y 5.2. El requisito de exactamente cuatro leptones, en lugar de al menos cuatro leptones, simplifica enormemente la estimación de fondo.
- **Separación de Leptones:** Se requiere que los leptones estén separados espacialmente por $\Delta R(l, l) > 0.2$. Este requisito se satisface automáticamente después de aplicar los requisitos de aislamiento de leptones (aplicados a todos los leptones).
- **Coincidencia del Disparador:** Al menos un leptón seleccionado debe coincidir con el objeto que provocó el disparo del disparador dentro de $\Delta R < 0.1$, y pasar los mismos requisitos cinemáticos y de identificación aplicados en el disparador, como se describe en la Sección 5.3.
- **Formación de Cuatrillizos:** Debe haber dos pares de leptones del mismo sabor (por sus siglas en inglés, SF) y de signo opuesto (por sus siglas en inglés, OS), esto da lugar a tres canales de señal: $4e$, 4μ y $2e2\mu$. En los eventos $eeee$ y $\mu\mu\mu\mu$ hay dos formas posibles de emparejar los cuatro leptones en pares OS. El emparejamiento que minimiza la cantidad $|m_{12} - m_Z| + |m_{34} - m_Z|$ se elige, donde m_{12} , m_{34} son las masas invariantes de los dos pares de leptones en un cierto emparejamiento y m_Z es la masa del bosón Z, tomado de [33].
- **Candidato Z “Primario”:** El candidato Z primario debe satisfacer el corte de masa $66 < m_{12} < 116 \text{ GeV}$.
- **Candidato Z ”Secundario”:** El candidato Z secundario debe satisfacer el corte de masa $66 < m_{34} < 116 \text{ GeV}$

```

// final selection, each lepton pair must have an invariant mass in the range 66-116 GeV, sum of differences < 50 GeV
if ( (sum_types == 44 || sum_types == 52 || sum_types == 48) && sum_ZZ_fin < 50 )
{

TLorentzVector FourLepSystem = TLorentzVector();
FourLepSystem = Lepton_1 + Lepton_2 + Lepton_3 + Lepton_4;
float FourLepSystem_pt = FourLepSystem.Pt()/1000.;
float FourLepSystem_y = FourLepSystem.Rapidity();
float FourLepSystem_m = FourLepSystem.M()/1000.;

```

Figura 5.2: Cada par de leptones debe tener una masa invariante en el rango de $66 < m_{ll} < 116 \text{ GeV}$ [109].

5.5. Incertidumbres Sistemáticas

Las incertidumbres surgen debido a la determinación de las eficiencias de identificación y reconstrucción de electrones y muones, las escalas y resoluciones de energía de electrones y muones, y la eficiencia del disparador, así como las incertidumbres teóricas asociadas con la simulación de Monte Carlo. Las incertidumbres surgen debido a los métodos utilizados para restar el fondo, los posibles sesgos de la metodología, la definición del electrón etiqueta y las diferencias entre los resultados de diferentes desintegraciones, así como debido a las estadísticas limitadas en los datos y las muestras de Monte Carlo [29].

Capítulo 6

Herramientas, Datos y Análisis Computacional para la Producción del Dibosón ZZ

La reproducibilidad es un pilar fundamental del método científico. La trazabilidad y reproducibilidad de los resultados son sin embargo condiciones inherentes a la ciencia de calidad, y un requisito cada vez más frecuente por parte de revistas y organismos financiadores de la investigación. Los estudios científicos reproducibles incluyen código informático capaz de recrear todos los resultados a partir de los datos originales. De esta manera el proceso de análisis queda perfectamente registrado, se reduce drásticamente el riesgo de errores, y se facilita la reutilización de código para análisis posteriores. A continuación presentamos los elementos más importantes de un flujo de trabajo reproducible (ver figura 6.1) e introducimos las principales herramientas disponibles para el manejo de datos, análisis de datos mediante código, control de versiones, la organización de los archivos, y el manejo de las dependencias de software externo.

6.1. ATLAS Open Data

La mayoría de los estudios científicos no son reproducibles: resulta muy difícil, si no imposible, trazar todo el proceso de análisis y obtención de resultados a partir de un conjunto de datos, incluso tratándose de los mismos investigadores. Entonces, en busca del camino hacia la reproducibilidad, el portal de ATLAS Open Data [110] provee una guía para poder visualizar datos provenientes del experimento, descargarlos y utilizarlos. Además proporciona software interactivos de código abierto para que se puedieran hacer análisis de física de altas energías [33].

6.1.1. Datos Reales y Simulados

En 2020, la colaboración ATLAS publicó un nuevo conjunto de datos de colisión de pp con fines educativos. Los datos fueron recogidos por el detector ATLAS en el LHC a 13 TeV durante el año 2016 y corresponden a una luminosidad integrada de $10fb^{-1}$ [33]. Este conjunto de datos va acompañado de eventos simulados que describen tanto varios procesos del modelo estándar como la producción de señales hipotéticas más allá del modelo estándar [110]. Las muestras de simulación de Monte Carlo (MC) describen varios procesos del modelo estándar (SM) y más allá del modelo estándar (BSM), que son utilizados para modelar las distribuciones esperadas de

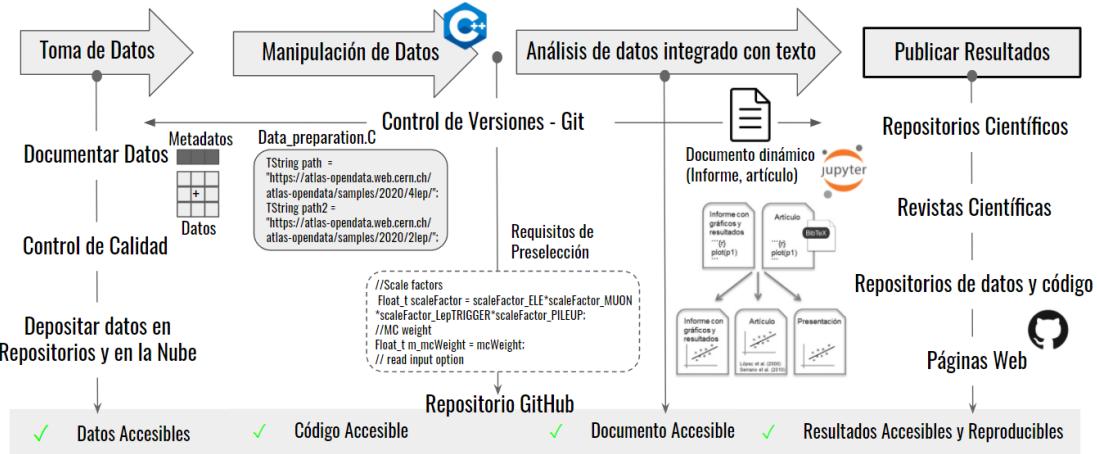


Figura 6.1: Esquema de un flujo de trabajo reproducible. En primer lugar, los datos se recogen según un protocolo bien diseñado, se documentan con metadatos, se someten a un control de calidad (idealmente de manera automática, esto es, mediante funciones de código), y se almacenan en un repositorio de datos en la nube. Después procederemos al análisis, siempre utilizando “scripts” para manipular los datos, y creando funciones que pueden almacenarse en un paquete (para facilitar su documentación y posterior reutilización). El análisis propiamente dicho se haría mediante documentos en Jupyter Notebooks que integran texto, código y resultados (tablas y figuras). Estos documentos pueden convertirse fácilmente en presentaciones, páginas web, o artículos científicos plenamente reproducibles.

diferentes procesos de señal y de fondo. Todas las muestras simuladas son procesadas a través de los mismos algoritmos de reconstrucción y cadena de análisis que los datos y se someten a una preselección de eventos “loose” para reducir el tiempo de procesamiento.

Además, la colaboración ATLAS introdujo 12 análisis, los cuales están agrupados de la siguiente manera: 4 análisis de alta estadística, 3 análisis de baja estadística, 3 análisis del Bosón de Higgs y 2 análisis más allá del modelo estándar (BSM) [33]. Nuestro análisis de estudio es el de baja estadística. Estos análisis ilustran las limitaciones estadísticas del conjunto de datos publicado dada la baja sección transversal de producción de los procesos raros, donde las variaciones entre los datos y la predicción de MC se atribuyen a fluctuaciones estadísticas considerables [33].

Tabla 6.1: Conjunto de Datos [111].

| Descripción | Nombre | Peso |
|--|--------|-------|
| Eventos seleccionados con al menos dos leptones (e ó μ) | 2lep | 24Gb |
| Eventos seleccionados con al menos cuatro leptones (e ó μ) | 4lep | 427Mb |

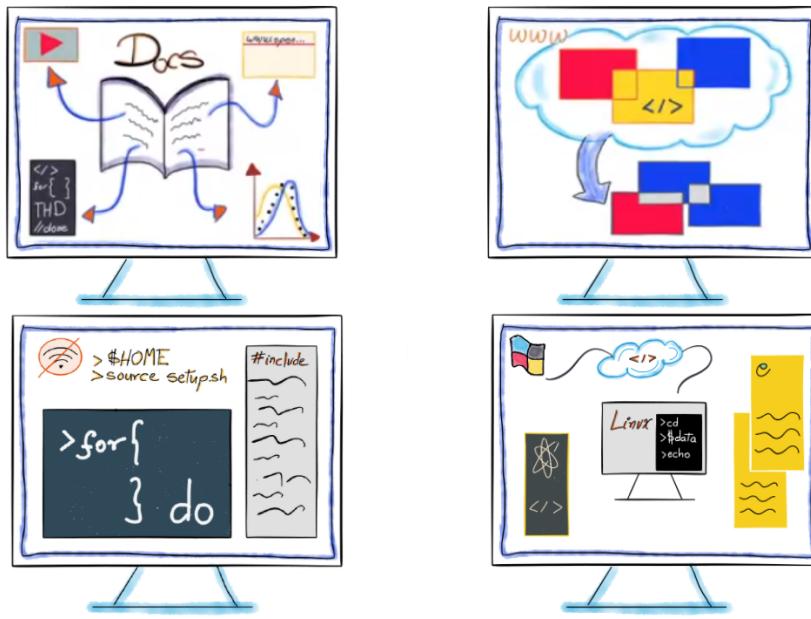


Figura 6.2: ATLAS Open Data provee documentación, datos reales y simulados, software y máquinas virtuales para poder recrear análisis en HEP y aprender sobre ellos. Además de incentivar a comprender más sobre programación computacional y estadística [112].

6.2. Herramientas Computacionales

Adoptar un flujo de trabajo reproducible requiere un esfuerzo inicial importante. Por eso es necesario familiarizarse con diversas herramientas (bases de datos, programación, sistemas de control de versiones) que proporcionen interfaces interactivas sencillas y fáciles de usar para el análisis de datos [113]. La forma tradicional para ejecutar dicho análisis de datos es a través de las siguientes herramientas descritas a continuación.

6.2.1. ROOT

Una de las herramientas mas importantes para el procesamiento de datos es ROOT (Software científico desarrollado por el CERN) [114]. Dicha herramienta provee de todas las funciones necesarias para poder trabajar con el procesamiento de cantidades grandes de datos, análisis estadísticos, visualización y almacenamiento. Está escrito principalmente en C++ pero tiene integrado otros lenguajes como Python y R.

ROOT permite las siguientes actividades:

- **Guardar datos:** Se puede guardar los datos (y cualquier objeto C++) en un formato binario comprimido en un archivo ROOT. El formato del objeto también se guarda en el mismo archivo: los archivos ROOT son autodescriptivos. Incluso en el caso de que los archivos fuente que describen el modelo de datos no estén disponibles, la información contenida en un archivo ROOT es siempre legible. ROOT proporciona una estructura de datos, el árbol, que es extremadamente potente para acceder rápidamente a grandes cantidades de datos, órdenes de magnitud más rápido que acceder a un archivo normal.



Figura 6.3: Herramienta para el procesamiento de datos: ROOT [114].

- **Acceso a los datos:** Se puede acceder a los datos guardados en uno o varios archivos ROOT desde la PC y desde la web, éste último método es el usado durante este trabajo de tesis.
- **Extraer Datos:** Se proporcionan potentes herramientas matemáticas y estadísticas para operar sobre los datos.
- **Publicar Resultados:** Los resultados pueden ser presentados en forma de histogramas, diagramas de dispersión, funciones de ajuste (fitting), entre muchas otras.

6.2.2. Jupyter Notebooks

Una de las herramientas más interactivas de código abierto para reproducir análisis es Jupyter Notebook, también conocida como cuaderno computacional, son esencialmente cuadernos de laboratorio para computación científica, se usa para combinar código de software, resultados computacionales, texto explicativo y recursos multimedia en un solo documento [115].

Jupyter Notebook tiene dos componentes. Se ingresa código de programación o texto en celdas rectangulares en una página web frontal. El navegador luego pasa ese código a un “kernel” de back-end, que ejecuta el código y devuelve los resultados. Esta ejecución del código se realiza mediante la comunicación con un núcleo (Kernel) de cálculo, es importante destacar que los núcleos Jupyter, que admiten docenas de lenguajes de programación, no necesitan residir en la computadora del usuario. Además, los cuadernos computacionales también se pueden ejecutar en la nube. En la figura 6.4 se muestra la interfaz de Jupyter Notebook.

6.2.3. SWAN

Por varios años, la física de altas energías ha estado enfrentando retos sin precedentes en el almacenamiento, procesamiento y análisis de datos. Entre las diversas opciones y soluciones que son exploradas por estas comunidades, hay una tendencia notable hacia los “web-based” análisis interactivos, donde los usuarios interactúan con un servicio en línea por medio de un navegador web. Este “Software as a Service” (SaaS) provee modelos que permiten a los usuarios concentrarse

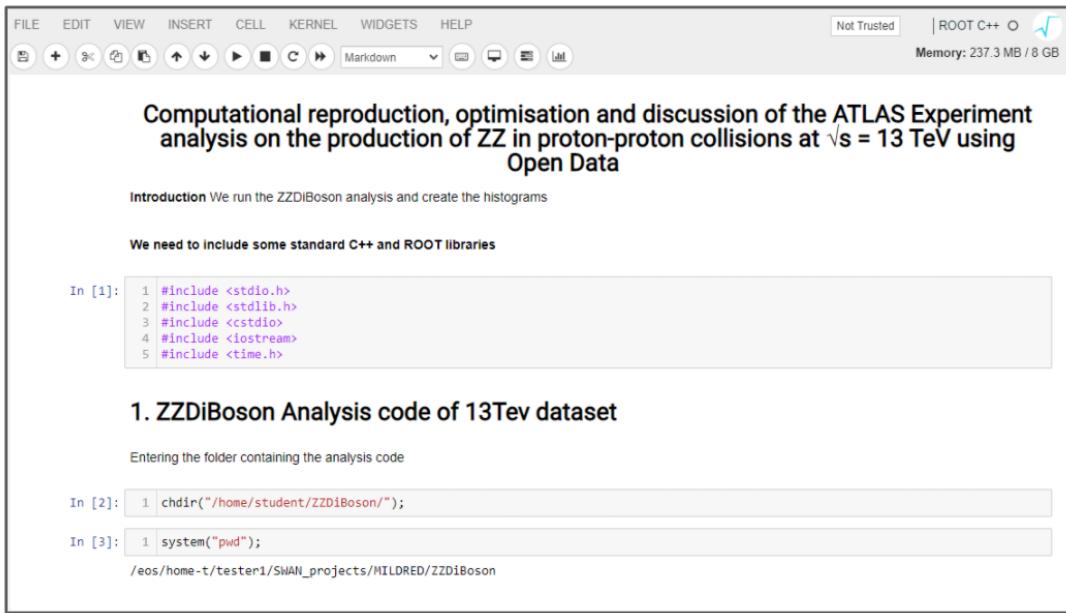


Figura 6.4: Interfaz de Jupyter Notebook [116].

en la solución del problema en cuestión en vez de la instalación, demás ajustes y problemas operacionales. SWAN (Servicio de análisis basado en web) es una plataforma para realizar análisis de datos interactivos en la nube [117]. SWAN permite a los usuarios escribir y ejecutar sus análisis de datos con solo un navegador web, aprovechando la interfaz de portátil Jupyter ampliamente adoptada. El código de usuario, las ejecuciones y los datos viven íntegramente en la nube. SWAN facilita la producción y el intercambio de resultados y código científico, el acceso a software científico, la producción de tutoriales y demostraciones y la conservación de análisis. Además, también es una herramienta poderosa para el análisis de datos no científicos y no está ligado a un lenguaje de programación en particular. En la actualidad ofrece la posibilidad de escribir notebooks en tres lenguajes: C++, Python y R. Un elemento innovador es que gracias a ROOT [114], estos lenguajes pueden ser mezclados sin problema, tal que pueden ser creados notebooks con múltiples lenguajes.

SWAN representa otro paso hacia la ciencia reproducible.

6.2.4. Git y GitHub

Otra herramienta importante es GitHub [118] el cual es un sitio web y un servicio en la nube que permite gestionar proyectos y controlar versiones de código. El sistema de control de versiones usado es Git [119], quien ayuda a gestionar nuestros archivos de proyecto. Una de las cosas principales que hace Git y también la razón principal por la que existe es para mantener un registro de toda la historia de las cosas en las que se esta trabajando. Esto es especialmente útil para los desarrolladores porque cuando se esta trabajando en un proyecto primero se construye una versión básica del mismo y luego se trata de mejorarlo añadiendo nuevas características o simplemente experimentando con las cosas.

Todo este proceso de experimentar con nuevas características es increíblemente propenso a errores y es posible que desee volver a su código original. Aquí es donde el control de versiones entra en juego, ya que realiza un seguimiento automático de cada minuto de cambio en el proyecto

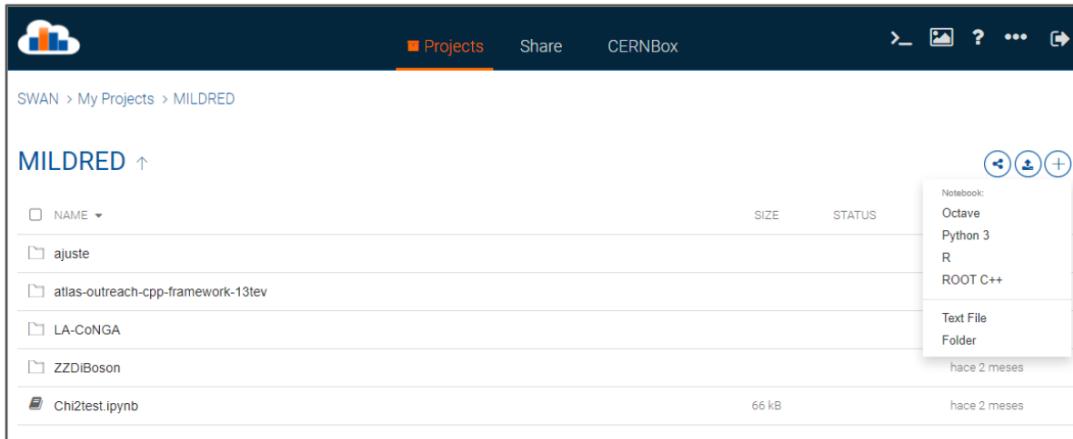


Figura 6.5: Interfaz de SWAN [117].

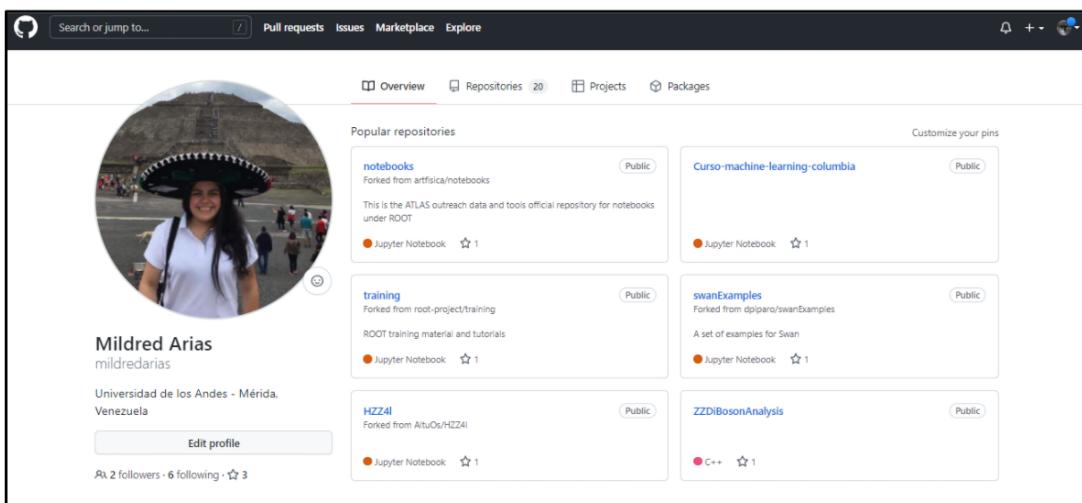


Figura 6.6: Interfaz de GitHub [120].

y permite volver a una versión anterior sin importar cuántas veces se haya cambiado los archivos. Otra cosa asombrosa que Git permite hacer es que las personas trabajen juntas en el mismo proyecto al mismo tiempo sin perturbar los archivos de los demás. La colaboración es aún más fácil con Git, los miembros de un equipo pueden trabajar en diferentes características y fusionar fácilmente los cambios.

Todas estas herramientas fomentan la reproducibilidad computacional al simplificar la reutilización del código.

6.3. Cómo hacer ciencia Reproducible

Para que un estudio sea reproducible, debemos seguir una serie de criterios comenzando por el proceso de recolección y manejo de datos, por lo tanto es muy importante garantizar la calidad de este proceso. Las muestras publicadas por ATLAS a 13 TeV se proporcionan en un formato de datos simplificado, lo que reduce el contenido de información del formato de análisis de datos



Figura 6.7: Herramientas para el Análisis de datos.

original utilizado dentro de la Colaboración ATLAS. El formato resultante es una tupla ROOT con más de 80 ramas [121]. Varios objetos físicos reconstruidos como los electrones, muones para nuestro análisis están contenidos dentro de los datos abiertos de 13 TeV ATLAS, y sus requisitos de selección se detallaron en el capítulo anterior.

Además, varios criterios de calidad de los datos garantizan que el detector funciona correctamente y los eventos se rechazan si contienen chorros reconstruidos asociados con depósitos de energía que pueden surgir de problemas de hardware, eventos de haz o lluvias de rayos cósmicos.

6.3.1. Organización de Ficheros

Mantener un sistema consistente de organizar todos los archivos relacionados con un proyecto es otro punto importante para garantizar su reproducibilidad. Cuando no se tiene ningún criterio los archivos se acumulan desordenadamente y resulta muy difícil manejar los distintos componentes del proyecto bien sea datos, código, figuras, texto, entre otros. Esto no sólo dificulta la comprensión y reutilización en el futuro, sino que también favorece la aparición de errores incontrolados. Hay muchas maneras de organizar un proyecto, pero existen algunos principios básicos que se deben tomar en cuenta que son [113]:

- Todos los ficheros relacionados con el proyecto deben estar dentro del mismo directorio.
- Existen subdirectorios independientes para los datos, código, figuras, resultados y manuscrito.
- Las funciones se definen en ficheros independientes del código que ejecuta el análisis.
- En el directorio raíz hay un fichero llamado README que describe el proyecto y sus componentes.

Teniendo esto en mente veamos la estructura que tiene nuestro fichero para el análisis de datos de la producción ZZDiBoson en el marco *C ++* [122]. Este marco consta de dos partes principales [122]:

- La parte de análisis, ubicada dentro del directorio “Análisis” en el cual se realiza la selección del objeto particular y almacena los histogramas de salida. Dicho directorio contiene los siguientes archivos:
 - Código principal del análisis (ZZDiBosonAnalysis.C) en el cual se realiza toda la selección de eventos, descrita en el capítulo 5 y el almacenamiento de los histogramas de salida.
 - También tenemos el encabezado principal del análisis (ZZDiBosonAnalysis.h) en donde se define los histogramas y se da acceso a las variables almacenadas en las muestras de entrada.
 - Además tenemos el encabezado de los histogramas (ZZDiBosonAnalysisHistograms.h) en el cual se define el nombre de los histogramas de salida.
 - Por otro lado el código de control principal del análisis (main_ZZDiBosonAnalysis.C) que controla qué muestras de entrada se van a utilizar y su ubicación.
 - Además se creó el directorio (Output_ZZDiBosonAnalysis), donde se almacenará la salida del código de dicho análisis (un archivo con histogramas por cada muestra de entrada). Si este directorio de salida no existe, el código fallará.
 - Por último en esta carpeta encontramos un script bash (run.sh), el cual ejecuta el análisis de forma interactiva a través de un shell de Linux.
- Por otro lado la parte de trazado, ubicada dentro del directorio “Plotting” en el cual se realizan los gráficos finales de datos/predicción. Cada subcarpeta del directorio Plotting contiene los siguientes archivos:
 - Directorio auxiliar (archivos de entrada): Dentro de él se encuentra el archivo de control con nombre Files_base.txt, cuyas muestras de entrada exactamente se utilizarán para este análisis en particular, su sección transversal y suma de pesos.
 - Directorio de ayuda (list_histos): Dentro de él se encuentra los nombres de los histogramas para cada uno de los análisis a estudiar el nombre del archivo para nuestro caso es HistoList_ZZDiBosonAnalysis.txt en nuestro caso y para los otros Histolist_Name Analysis.txt.
 - Por otro lado creamos el directorio “histograms” el cual contendrá los gráficos de salida.
 - El código principal de trazado se encuentra en (Plotting.cxx) que controla automáticamente qué hacer para cada uno de los 12 análisis.
 - El encabezado principal de trazado que se encuentra en (Plotting.h), en donde se define todas las piezas necesarias para el código de trazado.
 - Por último encontramos un script bash (plotme.sh): El cual ayuda a ejecutar el código de trazado de forma interactiva.

6.4. Descripción de la Reproducción Computacional para la Producción del Dibosón ZZ

Como se mencionó anteriormente la reproducibilidad va desde trabajos totalmente irreproducibles (que sólo contienen el texto, tablas y figuras finales) a estudios perfectamente reproducibles donde se integran todos estos directorios mencionados anteriormente en la estructura del fichero, además de la documentación, texto, control de versiones y datos que permite regenerar fácilmente el resultado final a partir de los datos originales. Para llevar acabo la reproducción del análisis utilizando Jupyter notebook con el kernel “ROOT C ++”, se crea un notebook dentro de la carpeta de análisis, llamado “ATLAS-Open-Data-cpp-13-TeV-framework-ZZDiBosonAnalysis.ipynb” [123].

1. En primer lugar se carga las librerías necesarias para dicho análisis, las cuales son:
 - *< stdio.h >*: Es el archivo de cabecera que contiene las definiciones de las macros, las constantes, las declaraciones de funciones de la biblioteca estándar del lenguaje de programación C para hacer operaciones, estándar, de entrada y salida, así como la definición de tipos necesarias para dichas operaciones. Por motivos de compatibilidad, el lenguaje de programación C++ (derivado de C) también tiene su propia implementación de estas funciones, que son declaradas con el archivo de cabecera *< cstdio >* [124].
 - *< stdlib.h >*: Es el archivo de cabecera de la biblioteca estándar de propósito general del lenguaje de programación C. Contiene los prototipos de funciones de C para gestión de memoria dinámica, control de procesos y otras. Es compatible con C++ [125].
 - *< iostream >*: Es un componente de la biblioteca estándar (STL) del lenguaje de programación C++ que es utilizado para operaciones de entrada/salida. Su nombre es un acrónimo de Input/Output Stream [126].
2. Luego se crea los directorios de “OutputZZDiBosonAnalysis” e “histograms” que son los que almacenan la salida del código de dicho análisis y los gráficos de salida. En la figura 6.9 se puede ver la creación de dichos directorios [127-128].
 - Otra forma más interactiva de hacer este proceso es crear dichos directorios con un “script bash”. Entonces, se ejecutó el “script bash” “welcome.sh” [129], en el cual fue llamado dentro del notebook mediante la función “system()” que se encuentra dentro de la librería “*< stdlib >*”.
3. Una vez creado estos directorios, llamamos las cabeceras “ZZDiBosonAnalysis.h” [130] y “ZZDiBosonAnalysisHistograms.h” [131] en el código principal del análisis “ZZDiBosonAnalysis.c” [132], cuyas cabeceras definen los histogramas, da acceso a las variables almacenadas en las muestras de entrada y el nombre de los histogramas de salida.
4. Luego se llamó a ROOT que se encargó de analizar todas las muestras de entrada que están definidas en “main_NNAnalysis.C” [133], y así se fue llenando el directorio “OutputZZDiBosonAnalysis” [127].

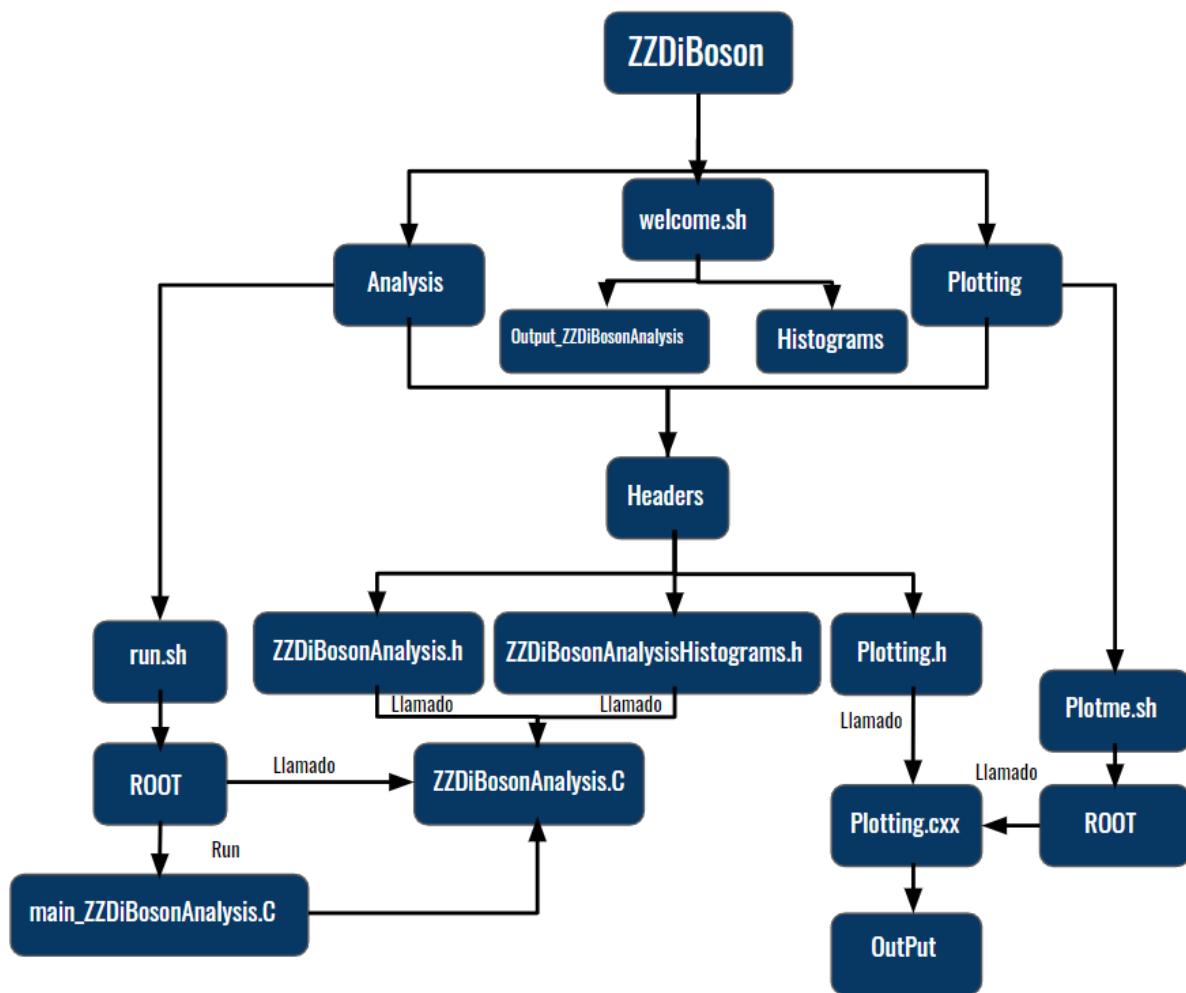


Figura 6.8: Esquema de la Reproducción Computacional para el análisis de la producción ZZ.

1. ZZDiBoson Analysis code of 13Tev dataset

Entering the folder containing the analysis code

```
In [2]: chdir("/home/student/ZZDiBoson/");

In [3]: system("pwd");
/eos/home-t/tester1/SWAN_projects/MILDRED/ZZDiBoson
```

Create the output folder

Welcome to ATLAS 13 TeV Open Data C++ framework!!!

Creating all output directories and cleaning created files

```
In [4]: system("mkdir -p Output_ZZDiBosonAnalysis");
```

Creating the directory for the Plotting code: Plotting/histograms

```
In [5]: system("mkdir -p Plotting/histograms");
```

Figura 6.9: Directorios creados “Output_ZZDiBosonAnalysis” [127] e “histograms” [128], para almacenar la salida del código de dicho análisis y los gráficos de salida.

- Todo este proceso se ejecutó llamando el script bash “run.sh” [134], el cual llamó la data que se iba a utilizar y mediante ROOT se analizaron todas estas muestras. Este script bash se llama dentro del notebook mediante la función “system()”. Para que esto se ejecutara de manera correcta dentro del notebook, se hizo una adaptación al “run.sh” el cual se seleccionó trabajar con toda la data tanto real como la simulada, como se ve en la figura 6.10a.
5. Por otra parte se llamó la cabecera de “plotting.h” [135] en el código principal de trazado “Plotting.cxx” [136], en donde se definieron las piezas necesarias para el trazado. Luego se ubicó el directorio “Output_ZZDiBosonAnalysis” [127] y se llamó a ROOT para que comenzará el trazado.
 - Para que esto se ejecutara de manera correcta dentro del notebook, se hizo una adaptación al “plotme.sh” [137], para ubicar el directorio “Output_ZZDiBosonAnalysis” [127], en donde se iba almacenar la salida del código de dicho análisis, luego se llamó ROOT y este comenzó el trazado. Cuyos gráficos de salida se almacenaron dentro del directorio de histograms. Cuyos gráficos de salida se almacenaron dentro del directorio de histograms.

En el siguiente capítulo se muestra los gráficos de salida de dicha reproducción.

```

1 #!/bin/bash
2 ## small script to run the analysis
3 analysis="main_ZZDiBosonAnalysis"
4
5 ##OPTION
6 echo Which option should I run?
7 echo Options are:
8 echo 0 = run all data and MC one after another
9 echo 1 = run data only
10 echo 2,3 = run MC samples only
11 #read varname
12 #echo Option is $varname
13 #option=$varname
14
15 echo Should I use PROOF? \will make things faster\
16 echo Options are:
17 echo 0 = NO
18 echo 1 = YES
19 #read proofvarname
20 #echo PROOF option is $proofvarname
21 #parallel=$proofvarname
22
23
24 ## execute and run ROOT
25 echo "starting ROOT"
26 ##
27 root -l -b << EOF
28 .L $analysis.C+
29 $analysis(1,0)
30 EOF
31 ##
32 echo "End of ROOT execution"

```

3. Begin analysis

Which option should I run?

Options are:

0 = run all data and MC one after another
 1 = run data only
 2,3 = run MC samples only

Should I use PROOF? (will make things faster)

Options are:

0 = NO
 1 = YES

Now it's time to RUN!

In [8]: `system("chmod 755 run.sh"); //permission`

In [9]: `system("./run.sh");` ←

Which option should I run?
 Options are:
 0 = run all data and MC one after another
 1 = run data only
 2,3 = run MC samples only
 Should I use PROOF? (will make things faster)
 Options are:
 0 = NO
 1 = YES
 starting ROOT
 Info in <TUnixSystem::ACLiC>: creating shared lib
 +++ Starting PROOF-Lite with 16 workers +++

(a) (b)

Figura 6.10: a)Script Bash “run.sh” [134]. b)Llamado del Script Bash “run.sh” en el notebook [134].

Capítulo 7

Optimización y Resultados

En física de partículas para hacer análisis de datos es necesario tomar y procesar datos, formular hipótesis así como estimar observables físicos. A menudo, se desea cuantificar el nivel de acuerdo entre los datos y una hipótesis. Esto se puede hacer definiendo un estadístico de bondad de ajuste, una función de los datos cuyo valor refleja el nivel de acuerdo entre los datos y una hipótesis dada (Ver Apéndice C). En nuestro caso, se plantea una hipótesis, el cual hace referencia a una función de densidad de probabilidad (por sus siglas en inglés, P.D.F.s), a causa de buscar un buen acuerdo entre los datos y la hipótesis dada. Por otro lado, nos interesa utilizar dichos datos para hacer inferencias sobre un modelo probabilístico, por ejemplo, para evaluar la validez del modelo o para determinar los valores de sus parámetros, en ambos casos se usa los métodos de ajuste de mínimos cuadrados y de máxima verosimilitud (Ver la sección C.2 y C.3 del Apéndice C.) [138]. Por otro lado, en este capítulo se presentarán los resultados de dicha reproducción. A continuación se dará una descripción de la P.D.F. utilizada para dicho ajuste en ROOT.

7.1. Descripción de la P.D.F. utilizada para dicho ajuste en ROOT

La clase de ROOT, que es “ROOT :: Math” [139], proporciona la mayoría de las funciones especiales y las funciones estadísticas, como lo son las funciones de densidad de probabilidad, la cual se usa para el ajuste de datos [140].

Las función de densidad de probabilidad usada para buscar un buen acuerdo entre los datos y las hipótesis dadas, es:

- **La Distribución de Landau:** Representa una distribución estadística definida y sustentada en el conjunto de números reales \mathbb{R} y parametrizada por un número real μ (llamado “parámetro de ubicación” y corresponde aproximadamente al valor más probable (MPV)) y por un número real c (llamado “parámetro de escala”, que corresponde con la desviación estándar (RMS)) [141], para mayor información ver apéndice C.

La función de densidad de probabilidad, es [141]:

$$p(x; \mu, c) = \frac{1}{\pi c} \int_0^{\infty} e^{-t} \cos(t(\frac{x-\mu}{c}) + \frac{2t}{\pi} \log(\frac{t}{c})) dt \quad (7.1)$$

donde $c \in (0, \infty)$ y $\mu \in (-\infty, \infty)$

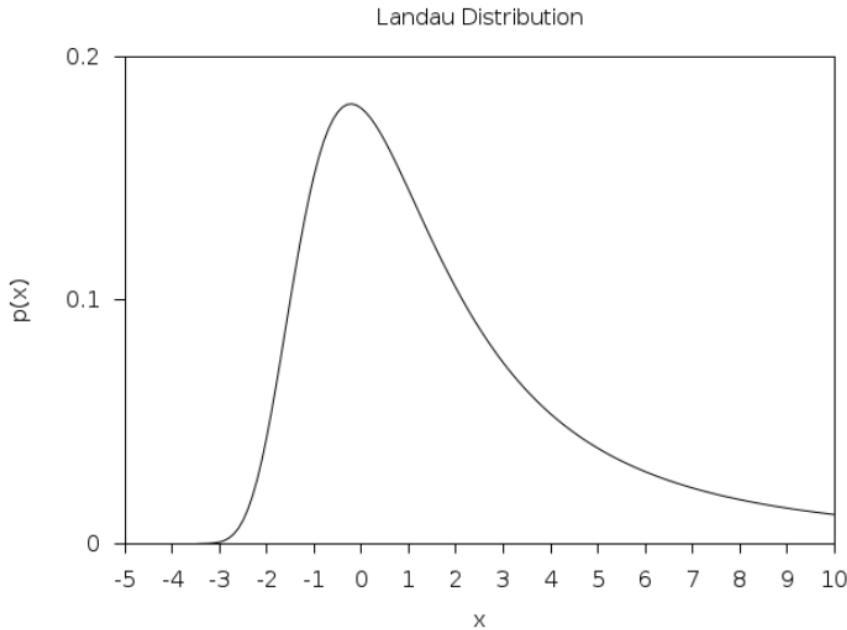


Figura 7.1: Distribución de Landau con parámetros μ y c [142].

Se observa que la forma original de $p(x)$ se obtiene para $\mu = 0$ y $c = \frac{\pi}{2}$, que es la definición de la distribución que utiliza ROOT para el ajuste, cuya función ha sido adaptada de la rutina CERNLIB G110 denlan [143].

La estadística juega un papel vital en la ciencia, ya que permite cuantificar los conocimientos, las incertidumbres y comunicar los resultados de los experimentos. La estadística provee herramientas para analizar los datos, extraer información, lidiar con los errores experimentales, cuantificar los grados de compatibilidad entre las medidas y las predicciones teóricas, y reportar los resultados partiendo de las mediciones. Es por ello que se usan métodos de ajustes para cuantificar el nivel de acuerdo entre los datos y la hipótesis dada, además de determinar los valores de los parámetros de la P.D.F [138]. En la siguiente sección se describe el método de ajuste en ROOT.

7.2. El Método de Ajuste en ROOT

Para ajustar un conjunto de datos, necesitamos un modelo para describir nuestros datos, por ejemplo, una función de densidad de probabilidad que describe nuestros datos observados, además de determinar los parámetros que mejor se adapten a los datos. En ROOT para ajustar un histograma, se puede usar el panel de ajuste a través del menú contextual (Ver figura 7.2), o se puede usar el método “TH1::Fit” [144], ver apéndice D para más información.

A continuación se describe el método “TH1::Fit”.

7.2.1. El Método “TH1::Fit”

Se presenta la implementación del método “TH1::Fit”, cuya explicación de los parámetros se extiende en el apéndice D:

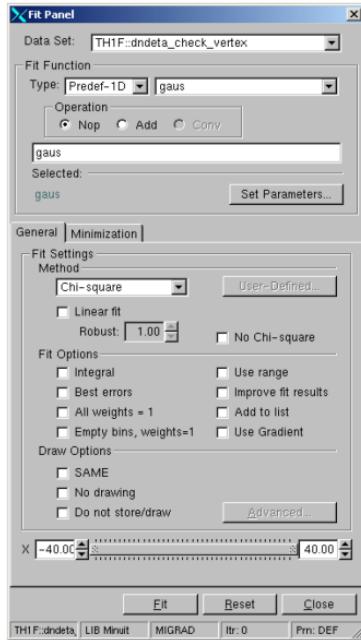


Figura 7.2: Panel de Ajuste en ROOT [144].

```
1 | TFitResultPtr Fit(TF1 *function, Option_t *option, Option_t *
    goption, Axis_t xxmin, Axis_t xxmax);
```

- *function: Este nombre puede ser uno de los nombres de función predefinidos de ROOT o una función definida por el usuario. En nuestro caso la función de “landau” ya esta predefinida en ROOT con mean y sigma. Esta función se ha adaptado de la rutina CERNLIB G110 denlan(ver TMath::Landau) [144]
- *option: La opción de ajuste utilizada es [144]:
 - “L”: Utiliza el método de probabilidad logarítmica (el método predeterminado es chi-cuadrado, “ χ^2 ”). Para usar cuando el histograma representa recuentos.

En el caso de histogramas ROOT (clases TH1C , TH1S , TH1F , TH1D), se define la función de ajuste Minuit como H1FitChisquare (“ χ^2 ”) ó H1FitLikelihood (“L”) según las opciones seleccionadas [145]. Por ejemplo si se elige H1FitChisquare, este calcula el χ^2 entre la función de ajuste del usuario (gaussiana, landau, definida por el usuario, etc.) y los datos para los valores dados de los parámetros [145]. Es tarea de MINUIT encontrar aquellos valores de los parámetros que dan el valor más bajo de χ^2 [145]. Una descripción mas detallada del software de minimización MINUIT se puede ver en el Apéndice D.

- *goption: El tercer parámetro es la opción de gráficos que es la misma que en el “TH1::Draw” [144].
- xxmin, xxmax: Rango sobre el que aplicar el ajuste [144].

De forma predeterminada, el objeto de función ajustado se agrega al histograma y se dibuja en el panel actual.

7.3. Optimización

Los datos recogidos por el detector ATLAS en el LHC a 13 TeV durante el año 2016 [33], fueron usados para hacer inferencias sobre un modelo probabilístico y determinar los valores de sus parámetros. Esto se pudo hacer definiendo un estadístico de bondad de ajuste, t [138] (Ver apéndice C, sección C.5.2.), el cual es una función de los datos cuyo valor refleja el nivel de acuerdo entre los datos y la hipótesis dada. En nuestro caso, la hipótesis planteada, hizo referencia a la función de densidad de probabilidad (por sus siglas en inglés, P.D.F.s) de landau (Ver apéndice C). Por otro lado, se estimó los parámetros de la p.d.f ajustada, mediante el método de mínimos cuadrados, cuyo software de minimización utilizado en ROOT fue TMinuit (Ver Apéndice D) [145], en el cual se obtuvo el valor mínimo de la cantidad χ^2 (Ver apéndice C, sección C.4.), que fue utilizado como estadístico de bondad de ajuste, t [138].

7.3.1. Prueba de Bondad de Ajuste

La aplicación de la prueba de bondad de ajuste requiere:

- Que los datos estén agrupados en categorías o clases [146]. Lo cual fue necesario construir un histograma que contenía los datos de la variable m_{4l} , a través de ROOT usando el kernel C++. Como se ve en la figura 7.3.

El estadístico de bondad de ajuste, usado para buscar un nivel de acuerdo entre los datos y la hipótesis planteada fue el método de mínimos cuadrados [147], el cual la función a minimizar es la suma de los cuadrados de las desviaciones entre los valores medidos y las predicciones del modelo que contiene parámetros variables, cuya ecuación viene dada por (Ver apéndice C.2),

$$\chi^2(\theta) = -2 \ln L(\theta) + ctte = \sum_{i=1}^N \frac{(y_i - f(x_i; \theta))^2}{\sigma_i^2} \quad (7.2)$$

donde $y(i)$ es el contenido del contenedor para cada contenedor i , $x(i)$ es el centro del contenedor, σ_i^2 es el error del contenedor (sqrt(y(i)) para un histograma no ponderado), $f(x_i; \theta)$ es la función modelo y θ son los parámetros [147]. A continuación se presenta el código para llevar a cabo el estadístico de bondad de ajuste, este también se detalla en el repositorio “mildredarias/ZZDiBoson/Chi2_minuit.ipynb” en Github [148].

```

1 TF1 *fit1=new TF1("fit1","[0]*TMath::Landau(x,[1],[2])",120,500);
2 //Definicion de la funcion a estudiar
3 fit1->SetParameters(m_four_lep->GetMaximum(),m_four_lep->GetMean()
4 ,m_four_lep->GetRMS());
5 //Conjunto de parametros
6 fit1->SetParNames("const","MPV","RMS");
7 //Nombre de los parametros
8 m_four_lep->Fit("fit1","","",120,500);
9 //Ajuste
10 m_four_lep->Draw("E1");
11 c2->Draw();
12 //Se pinta la funcion ajustada

```

La hipótesis nula H_0 , en el test de bondad de ajuste es:

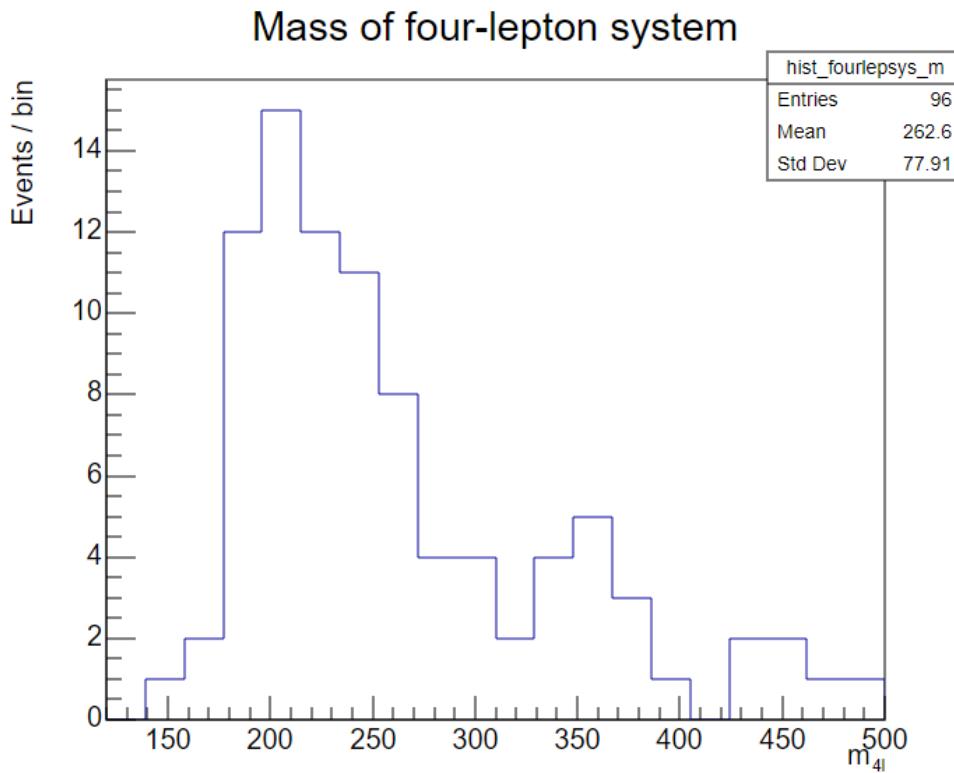


Figura 7.3: Histograma de la Masa del sistema de cuatro leptones [148].

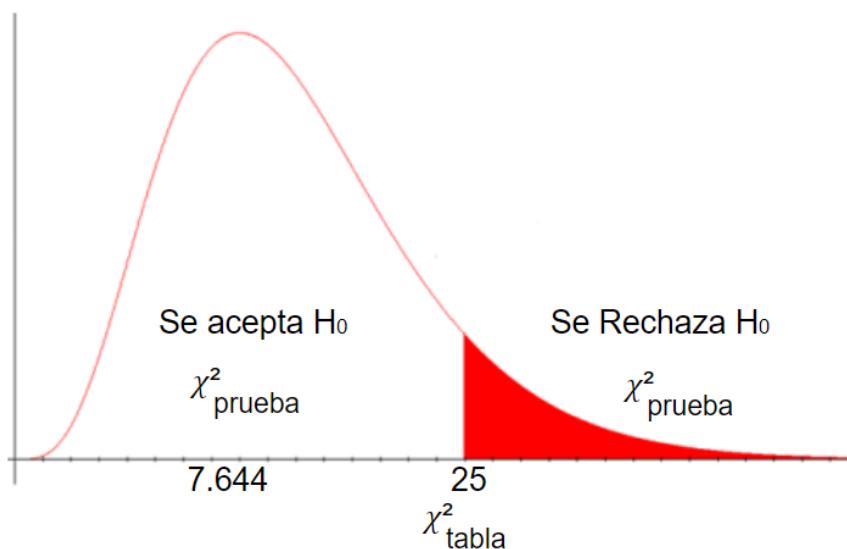


Figura 7.4: Zonas donde se acepta o se rechaza la hipótesis nula, dependiendo del valor de χ^2_{prueba} .

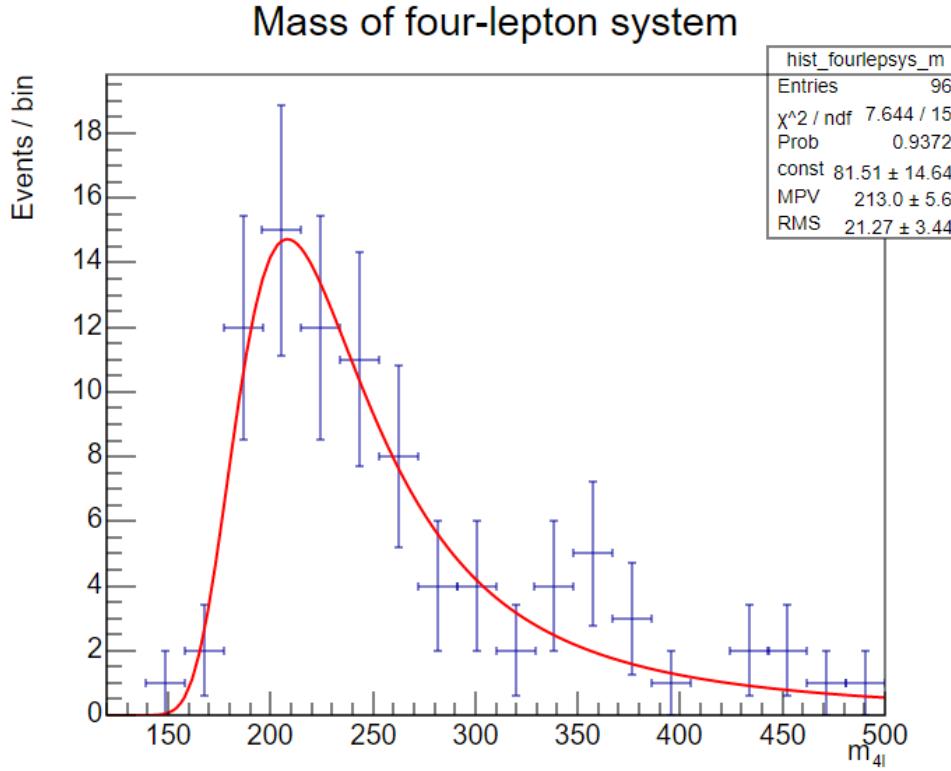


Figura 7.5: Ajuste de la distribución de landau a los datos de la masa invariante del sistema de cuatro leptones [148].

- H_0 : La variable m_{4l} sigue la distribución de landau la cual esta parametrizada (μ, c) .

Donde dicha hipótesis fue evaluada para un nivel de significancia de 5 % (denotado por α : alfa) definido en el apéndice C (comúnmente se consideran valores de alfa de 1 %, 5 % o 10 %), que es equivalente a un nivel de confianza de un 95 % [146]. Adicionalmente se pudo obtener el valor p (o p-value) el cual es el valor “Prob” que proporciona ROOT, como la probabilidad de obtener un χ^2 dado un número de grados de libertad n [138]. El valor p también depende del nivel de significancia establecido (Ver apéndice C). Por otro lado, se obtuvieron los valores de los parámetros de dicha p.d.f mediante el método de mínimos cuadrados, los cuales son “const” (valor para la normalización), MPV (Valor más Probable) y RMS (desviación estándar), cuyos valores se observan en la figura 7.5.

En la figura 7.5 se puede observar que el valor de χ^2 es 7.644 con 15 grados de libertad y con una probabilidad de 0.9372. Al comparar el valor del estadístico de prueba $\chi^2(\chi^2_{prueba}) = 7,644$ con el valor de χ^2 esperado (χ^2_{tabla}), se puede apreciar en la tabla de percentiles de la distribución de χ^2 [149] que el valor esperado es 25, por lo tanto 7.644 se encuentra a la izquierda de 25 y la probabilidad de valores mayores a él es superior al nivel de significancia $\alpha = 0,05$. Además, el valor p es mayor que el nivel de significancia escogido, entonces no hay evidencia estadística que permita rechazar dicha hipótesis nula, esto significa que los datos se ajustan a la distribución de landau.

7.3.2. Cálculo del Mínimo χ^2 a través de TMinuit (Curva de χ^2)

Una vez que se realizó la prueba de bondad de ajuste, se deseó ver la curva de χ^2 , y para ello se empleo el software de minimización “TMinuit” (Ver Apéndice D) a través de ROOT, en la cual se efectuó un nuevo ajuste (Ver línea 12 del código presentado a continuación) y además se fijaron parámetros al azar, en donde se fueron calculando los valores de χ^2 .

El parámetro que fue variando fue el MPV (valor más probable), como se observa en la línea 10 del código presentado a continuación. Esto quiere decir que se calculó el χ^2 cuando el MPV variaba entre un valor mínimo y un valor máximo, los cuales fueron 208-216 GeV (Ver línea 1 del código). Este código se detalla en el repositorio “mildredarias/ZZDiBoson/Chi2_minuit.ipynb” [148].

```

1 double mpvMin=200, mpvMax=220;
2 double rmsMin=5, rmsMax=30;
3 double stepMPV = (mpvMax-mpvMin)/100.;
4 double stepRms = (rmsMax-rmsMin)/100.;

5
6 TH1D * histo = new TH1D("histo" , "histo" , 20 , 120 , 500 ) ;
7 TH1D * theMap = new TH1D("theMap" , "theMap;MPV;Chi2" , 100 , mpvMin ,
8 mpvMax);

9 for (double i= mpvMin; i< mpvMax ; i+= stepMPV){
10   fit1 -> FixParameter(1,i);
11   fit1 -> SetParameters(70,i,20);
12   m_four_lep-> Fit("fit1","","","",110,600);
13   cout<<fit1->GetChisquare ()<< endl;
14   theMap->SetBinContent(theMap->FindBin(i),fit1->GetChisquare ());
15 }
```

La forma de χ^2 en función del valor más probable (MPV) sigue un perfil parabólico, cuyo valor mínimo coincide con el valor más probable del modelo, el cual es 213 GeV, ver figura 7.6a, este valor se calculo en el repositorio “mildredarias/ZZDiBoson/Chi2_minuit.ipynb” en Github [148]. Además, el error estadístico es el intervalo que hace que el χ^2 aumente en 1 unidad. Cuyo intervalo de confianza central del 68,3 % viene dado por los valores de θ en los que χ^2 aumenta una unidad de su valor mínimo (Ver apéndice C, sección C.6.1).

Así que, se realizó una interpolación parabólica en el mínimo y se vio cual fue el error al hacer el ajuste y la diferencia en el eje vertical de un χ^2 , ver figura 7.6b. Por inspección alrededor del mínimo se observó que el χ^2 aumenta en 1 unidad para $\Delta MPV \sim \pm 5,5$, cuyo error estadístico corresponde con el error del parámetro MPV del modelo.

7.4. Comparación entre dichos Métodos Estadísticos

Para evaluar la fortaleza de ambos métodos, el método de mínimos cuadrados y de máxima verosimilitud, implementado en ROOT, para histogramas de estadística baja, se realizó una serie de procedimientos:

- Se tomó la distribucion de landau, el cual es el modelo.

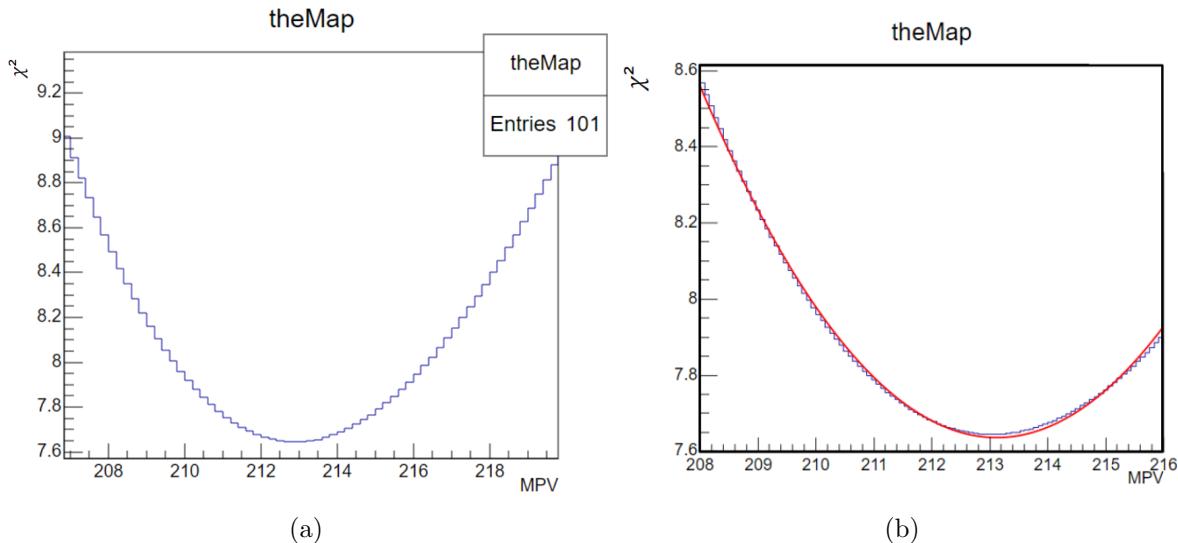


Figura 7.6: a), b) χ^2 en función de MPV (el valor más probable. b) Interpolación parabólica en el mínimo. Donde χ^2 aumenta en 1 unidad para $\Delta MPV \sim \pm 5.5$ [148].

- Se generó datos con ese modelo y se efectuó una realización aleatoria con 96 eventos y con $\text{MPV} = 200$ y $\sigma = 20$, ver figura 7.7.

```

1 int nEvents = 96
2 TF1 * func = new TF1("func" , "[0]*TMath::Landau(x,[1],[2])"
   ,100,500);
3
4 func->SetParameters(nEvents,200,20);
5 func->SetParNames("N","MPV","Sigma");
6
7 TH1D * histo = new TH1D("histo" , "histo" , 20 , 120 , 500 );
8
9 histo->FillRandom("func" , nEvents);
10 histo->Fit("func" , "" , "" , 120 , 500 );

```

- Luego, se tomó 140 veces los 96 eventos.
 - Se hizo un ajuste de χ^2 individual a cada conjunto de 96 eventos y para cada uno de esos ajustes, se obtuvo una estimación del parámetro MPV.

```
1 TH1D * Mean_square_least = new TH1D("Mean_square-least", "Mean_square-least", 205,120,500);
2 int valor = 0 ;
3
4 //140 realizaciones
5 for( int index = 0 ; index < 140 ; ++index){
6
```

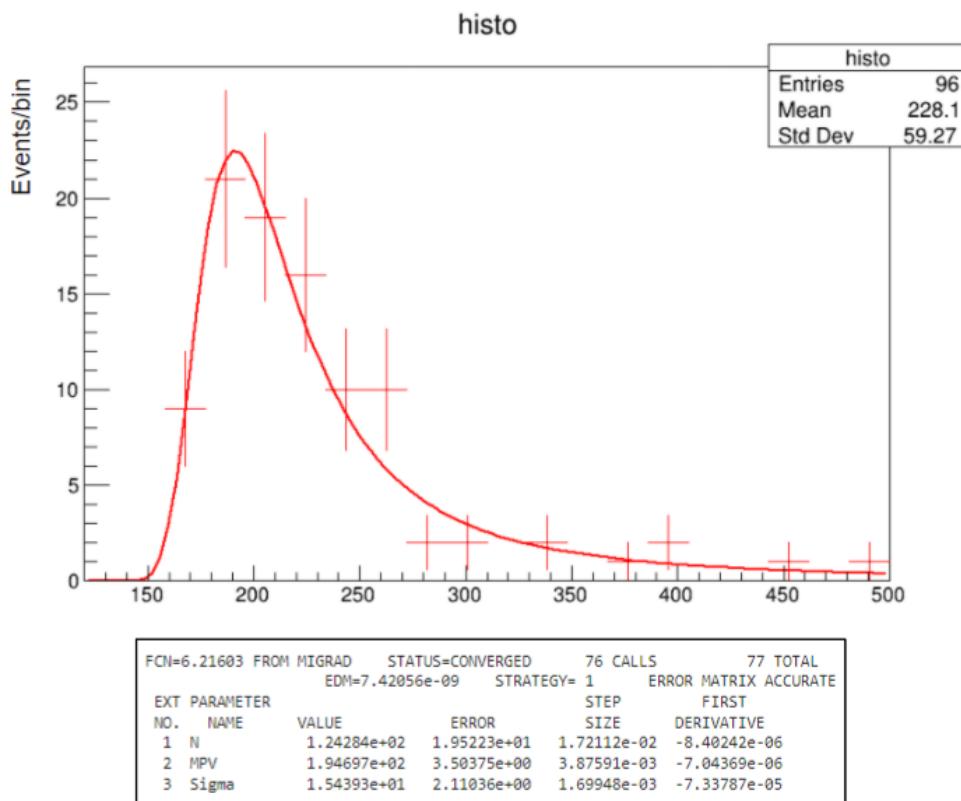


Figura 7.7: Distribución de landau con 96 eventos [150].

```

7 //Se generan numeros aleatorios que sigan una distribucion
8 //poissoniana
9
10 valor = gRandom->Poisson(nEvents);
11 func->SetParameters(nEvents ,200 ,20);
12
13 //Reseteo el histograma
14 histo->Reset();
15
16 //Usamos FillRandom para llenar aleatoriamente el
17 //histograma usando el contenido de la funcion TF1 que es
18 //landau en este caso.
19 //Se creo y se lleno un histograma (valor) veces con una
20 //distribucion landau con MPV 200 y sigma 20:
21 histo->FillRandom("func" , valor);
22
23 //Se ajusto el histograma con la misma funcion predefinida
24 //, en este caso es la distribucion de landau, usando:
25 histo->Fit("func" , "" , "" , 120 , 500 );
26
27 //Se lleno un histograma con los valores del parametro 1(
28 //MPV) de la funcion ajustada landau, para ver
29 //si el valor central es el valor real.
30 Mean_square_least->Fill(func->GetParameter(1));
31
32 }
33 Mean_square_least->Draw();

```

- Por ultimo se hizo un ajuste de log-likelihood individual a cada conjunto de 96 eventos, esto para poder comparar la fortaleza de ambos métodos. Para cada uno de esos ajustes, se obtuvo una también una estimación del MPV.

```

1 TH1D * Mean_likelihood = new TH1D("Mean_likelihood" , "Mean_likelihood" , 205 ,120 ,500) ;
2 int valor = 0 ;
3
4 //140 realizaciones
5 for( int index = 0 ; index < 140 ; ++index){
6     valor = gRandom->Poisson(nEvents);
7     func->SetParameters(nEvents ,200 ,20);
8
9     histo->Reset();
10
11     histo->FillRandom("func" , valor);
12
13     //Ajusto mi funcion con likelihood(minimiza la funcion de
14     //menos log-likelihood)

```

```

14     histo->Fit("func" , "QL" , "" , 120 , 500 );
15
16     //lleno mi histograma con el parametro 1 de la funcion
17     //ajustada
18     Mean_likelihood->Fill( func->GetParameter(1));
19
20 Mean_likelihood->Draw();

```

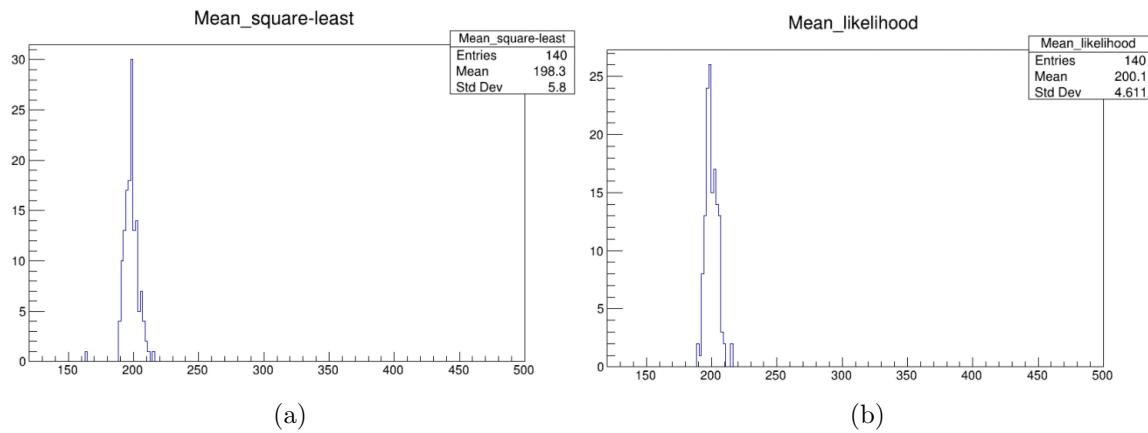


Figura 7.8: Comparación entre dichos Métodos Estadísticos con baja estadística [150].

Se puede ver que el método de log-likelihood tiene la ventaja de tratar correctamente los contenedores con estadísticas bajas, ya que el parámetro de MPV esta sesgado cuando se trabaja con el método de χ^2 con baja estadística, esto quiere decir que el valor central no es en el valor real del modelo, estadísticamente hablando la distribución esta desviada, ademas el RMS de χ^2 es notablemente mayor que el de log-Likelihood, ver figura 7.8. En caso de estadísticas altas / bin, la distribución del contenido de bin se convierte en una distribución normal y ambos métodos como log-likelihood y χ^2 dan el mismo resultado [151], ver figura 7.9. Normalmente, los recuentos de un histograma siguen una distribución de Poisson [151].

Además, cuando se usa el método χ^2 , existe el problema de los contenedores con cero entradas, en ROOT, estos contenedores se excluyen del ajuste, pero en realidad contienen información estadística valiosa [138]. En cambio, se consideran en el método de log-likelihood. El método log-likelihood, es el método recomendado en caso de estadísticas de bin bajas, donde el método χ^2 da resultados incorrectos, en particular cuando hay varios bins vacíos [147, 151].

7.5. Análisis de las Distribuciones Cinemáticas generadas en la Reproducción

Se analiza las distribuciones cinemáticas generadas en la reproducción que pasaron la selección de eventos (capítulo 5) para la producción ZZ en el canal de decaimiento 4l. A continuación, se compara los datos y la predicción de MC para varias distribuciones en la selección $ZZ \rightarrow 4l$.

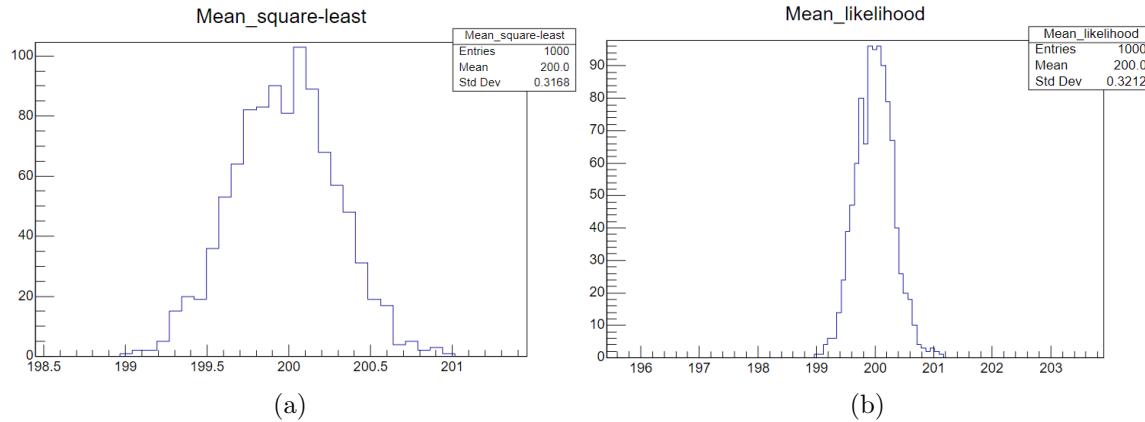


Figura 7.9: Comparación entre los Métodos Estadísticos con alta estadística [150].

Las distribuciones de los leptones para el $|PDG\ ID|$ [152] muestran que se están simulando en este caso electrones y muones. Como se observa en la figura 7.10a. Por otro lado en la figura 7.10b se muestra que dichos leptones son de carga opuesta.

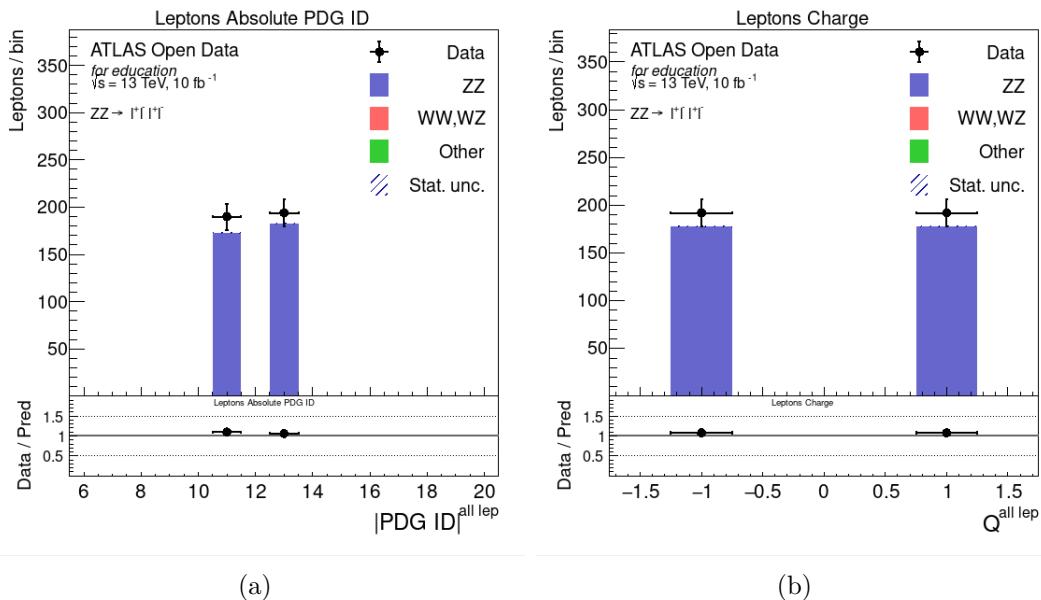


Figura 7.10: a) Las distribuciones de los leptones para el $|PDG ID|$ muestran que se están simulando en este caso electrones y muones [153] y b) Leptones de carga opuesta [154].

Como podemos ver, la forma de la figura 7.11a es debido a que no hay una dirección preferencial en los que van a salir esos leptones ya que pueden suceder los eventos a cualquier ángulo azimutal. Por otro lado en la figura 7.11b esta la pseudorapidez η cuya forma se debe a que los procesos duros (producto de las colisiones pp) van a aparecer de forma más frecuente ortogonales a la dirección de la colisión, por eso el pico en cero.

En la figura 7.12a se muestra el momento transverso de cada uno de los leptones, en el cual la gráfica tiene esa forma debido a la selección de eventos que se hizo cuando exigimos $p_T > 20 \text{ GeV}$ para los leptones (Ver capítulo 5). Por otro lado en la figura 7.12b observamos que la energía está

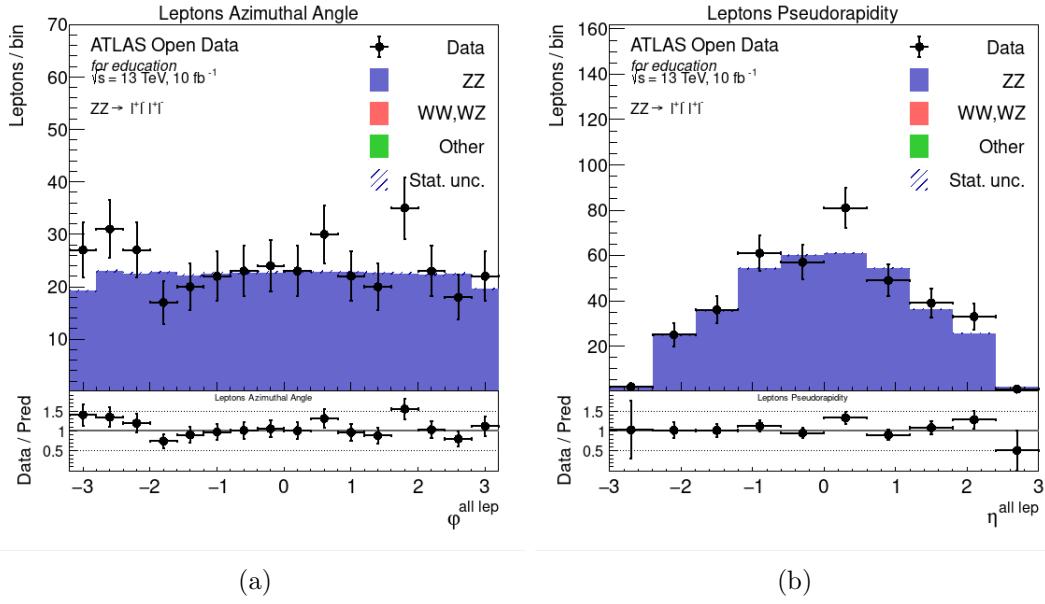


Figura 7.11: Comparación entre datos y predicción MC obtenidas de la reproducción del marco de análisis en el canal de decaimiento $ZZ \rightarrow 4l$. a) Ángulo azimutal φ [155] y b)Pseudorapidez η de cada leptón seleccionado [156].

altamente correlacionada con el p_T , se observa que la caída a baja energía se debe a la selección en p_T cuando se pidió leptones con $p_T > 20 \text{ GeV}$.

Hablamos de cada lepton individualmente pero realmente queremos identificar cual es el par que viene de cada Z, para ello se hizo pares en función de la masa invariante que se tiene. En la figura 7.13 podemos ver las distribuciones de las masas invariantes de los pares de bosones Z, el cual los datos concuerdan con las simulaciones. En la figura 7.13a se puede ver la distribución de la masa invariante del primer par. En la cual esta distribución es así debido a que este par tiene el lepton más energético, es conocido como el par de leptones primarios, en el que contribuirá más a la masa invariante del lepton que el de la figura 7.13b, que tiene el lepton menos energético y conocido como el par de leptones secundario (Ver la sección 5.4).

En la figura 7.14a se muestra que la mayoría de los eventos con 4 leptones tienen un bus pequeño y luego se va viendo cada vez una menor sección eficaz para eventos que salen con mayor momento transverso. Entonces esto quiere decir que la mayoría de los pares de bosones Z se generan con poco momento transverso. Por último en la figura 7.14b se puede observar que el histograma de la masa invariante del sistema de cuatro leptones no comienza en cero sino en 180 GeV y esto es debido al límite cinemático. Además como se tiene dos bosones Z que decaen en 4 leptones, la mínima masa que se tiene al final es la suma de las dos masas más la suma de los momentos de los leptones y esto último es lo que da la forma de dicho histograma.

En dichas distribuciones cinemáticas, se puede notar que:

- Los puntos representan datos experimentales.
 - Los histogramas llenos muestran la predicción de diferentes simulaciones de MC.
 - Las contribuciones se apilan.

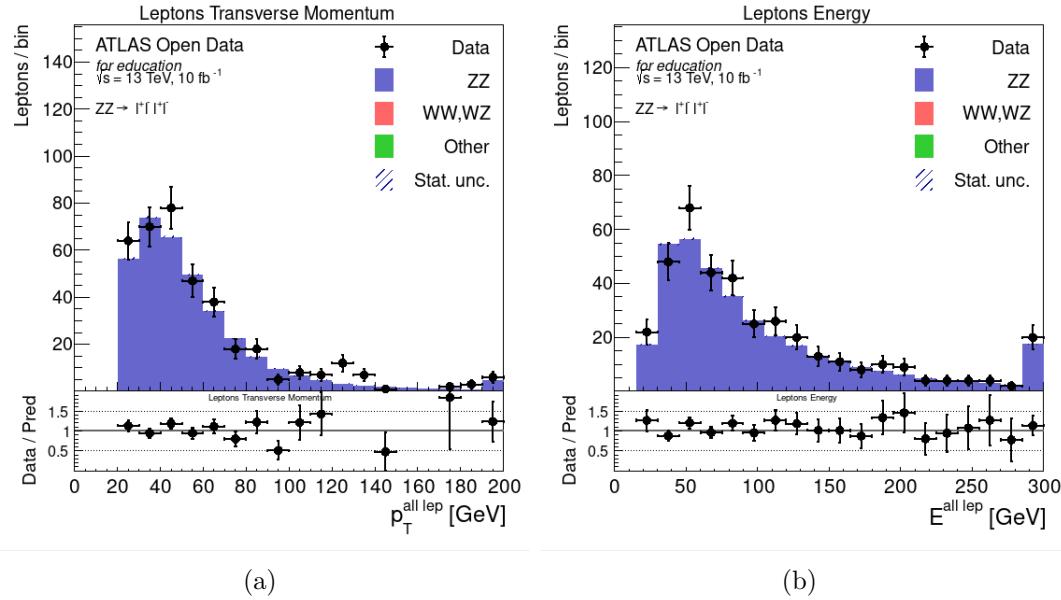


Figura 7.12: Comparación entre datos y predicción MC obtenidas de la reproducción del marco de análisis en el canal de decaimiento $ZZ \rightarrow 4l$. a) Momento transverso p_T [157] y b) Energía E de cada lepton seleccionado [158].

- La incertidumbre estadística estan representada por las barras de error en los puntos de datos.
 - El último contenedor de todas las figuras contiene el desbordamiento.
 - Los paneles inferiores de cada figura muestran la relación entre los puntos de datos y el histograma apilado.
 - Las contribuciones de fondo que son WW , WZ y otros son despreciables y no aparecen en los histogramas, ya que se ha procesado las muestras tanto de los WW y WZ y no han pasado, debido a la selección de eventos.
 - La masa invariante del bosón Z coincide con el valor tabulado [33].
 - En general se nota una buena concordancia entre los datos reales y simulados.
 - La cinemática de los leptones, muestra que los datos simulados están modelados correctamente.

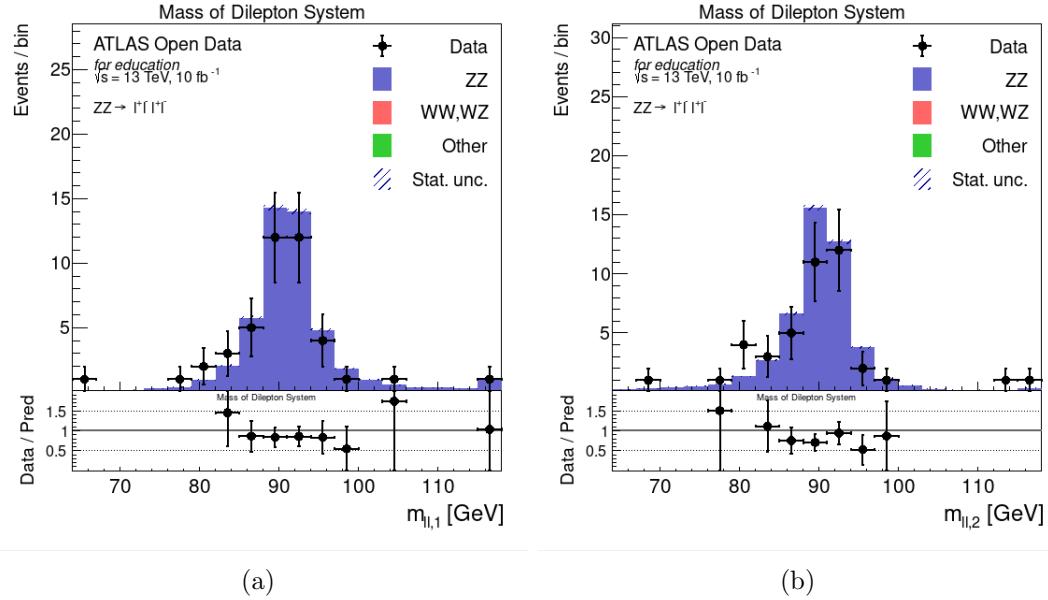


Figura 7.13: Distribuciones de las masas invariantes del (a) primero [159] y (b) segundo bosón Z reconstruido [160]. Los puntos representan datos experimentales. Los histogramas llenos muestran la predicción de diferentes simulaciones de MC. Los paneles inferiores de cada figura muestra la relación entre los puntos de datos y el histograma apilado.

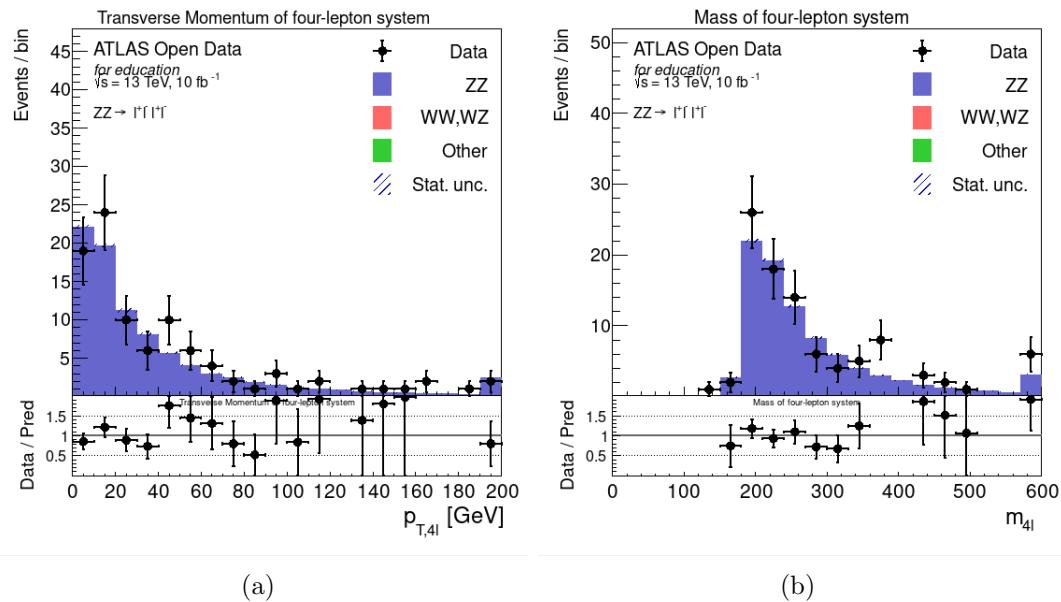


Figura 7.14: a) Momento transversal [161] y b) Masa invariante del sistema de cuatro leptones [162] en eventos seleccionados, reconstruida en la reproducción computacional.

Capítulo 8

Conclusiones y Discusión

Los estudios científicos reproducibles incluyen código informático capaz de recrear todos los resultados a partir de los datos originales. De esta manera el proceso de análisis queda registrado, se reduce drásticamente el riesgo de errores, y se facilita la reutilización de código para análisis posteriores. Otro componente importante es el acceso a tales recursos. En el caso de accesibilidad, el proyecto de acceso abierto ATLAS Open Data[110] consiste en la publicación de un conjunto de datos y las herramientas de software, proporcionando también interfaces interactivas para replicar procedimientos utilizados por investigadores de física de altas energías y permitir a los usuarios experimentar el análisis de datos de física de partículas en entornos educativos.

Para efectos del presente trabajo, se aprovecharon los datos compartidos y recopilados por el experimento ATLAS en 2016, los cuales se utilizaron para realizar mediciones de la producción de ZZ a $\sqrt{s} = 13 \text{ TeV}$ por el LHC, seleccionando eventos que fueran consistentes con la producción de bosones Z que se desintegran en electrones o muones. Con tales recursos se logró recrear el análisis de la producción de ZZ en el canal de decaimiento $ZZ \rightarrow 4l$ a $\sqrt{s} = 13 \text{ TeV}$ utilizando Jupyter Notebooks y ROOT. Está escrito principalmente en $C++$ pero tiene integrado otros lenguajes como Bash.

Las muestras publicadas por ATLAS a 13 TeV se proporcionan en un formato de datos simplificado, lo que reduce el contenido de información del formato de análisis de datos original utilizado dentro de la Colaboración ATLAS. El formato resultante es una tupla ROOT con más de 80 ramas. Varios objetos físicos reconstruidos como los electrones, muones para nuestro análisis están contenidos dentro de los datos abiertos de 13 TeV ATLAS, y sus requisitos de selección se detallaron en el capítulo 5. Además, varios criterios de calidad de los datos garantizan que el detector funciona correctamente y los eventos se rechazan si contienen chorros reconstruidos asociados con depósitos de energía que pueden surgir de problemas de hardware, eventos de haz o lluvias de rayos cósmicos.

Como se menciona, además de ROOT, Jupyter Notebooks se utilizaron como cuaderno de laboratorio en la nube para la computación necesaria, permitiendo combinar código, resultados computacionales, texto explicativo y recursos multimedia en un solo documento [115].

Sumado a ello, fue utilizada como herramienta informática para el análisis de los datos la plataforma SWAN (servicio de análisis basado en la web proporcionado por el CERN), la cual se usó para realizar los análisis de datos interactivos en la nube [117]. SWAN permitió escribir y ejecutar los análisis de datos a través de un navegador web, aprovechando la interfaz de Jupyter. El código de usuario, las ejecuciones y los datos viven íntegramente en la nube.

Otra herramienta importante es GitHub [118] el cual es un sitio web y un servicio en la nube que permite gestionar proyectos y controlar versiones de código. El sistema de control de versiones usado es Git [119], quien ayuda a gestionar los archivos del proyecto.

Todas estas herramientas fomentan la reproducibilidad computacional al simplificar la reutilización del código.

Asimismo, se logró la discusión de los resultados obtenidos de la reproducción computacional usando Jupyter Notebooks con el kernel ROOT *C++* para la producción de ZZ en el canal de desintegración $ZZ \rightarrow 4l$ a $\sqrt{s} = 13\text{ TeV}$, en el que además se generó un esquema de un flujo de trabajo reproducible (ver figura 6.1).

Este consistió, en primer lugar, el comprender cómo fue la construcción de los datos a utilizar, es decir, el cómo se da la manipulación previa de los datos según un protocolo previamente diseñado. Durante todo el procedimiento, los físicos generan metadatos necesarios para maximizar el entendimiento y la reproducibilidad de los análisis a publicar. Como es habitual en esta clase de análisis de física de partículas, los datos reales y simulados (es decir, el MonteCarlo, o MC) se sometieron a controles de calidad y se almacenaron en un repositorio de datos en la nube. Luego, se procedió al pre-análisis, siempre utilizando “scripts” para manipular los datos, y creando funciones que pudieran almacenarse en un paquete (software) para facilitar su documentación y posterior reutilización.

El análisis de este trabajo de tesis se hizo mediante documentos en Jupyter Notebooks que integran texto, código y resultados (tablas y figuras). Estos documentos pueden convertirse en presentaciones, páginas web, o artículos científicos plenamente reproducibles. Así pues, una gran parte del trabajo (que entre otras cosas, consiste en leer más de 2.1 millones de eventos de colisión protón-protón) reside en comprender las herramientas de análisis, en cómo traducir algoritmos a diferentes lenguajes y plataformas. En este último caso, de código en *C++* en macros a Jupyter notebooks: es a ello lo que llamamos reproducción computacional.

En nuestro análisis, las distribuciones cinemáticas generadas en la reproducción contienen los eventos que pasaron la selección de ZZ descrita en el capítulo 6. En donde los puntos representan los datos experimentales; los histogramas llenos muestran la predicción de diferentes simulaciones de MC de forma apilada. La incertidumbre estadística está representada por las barras de error en los puntos de datos; el último contenedor de todas las figuras contiene el desbordamiento (algo habitual en esta clase de análisis); los paneles inferiores de cada figura muestran la relación entre los puntos de datos y el histograma apilado, es decir, las contribuciones de MC. De estas contribuciones de fondo, WW, WZ y otros son despreciables y no aparecen en los histogramas, aunque si son tomados en cuenta en el código y no han pasado la selección de eventos. La masa invariante del bosón *Z* coincide con el valor tabulado en la literatura [33]. En general se nota una buena concordancia entre los datos reales y simulados. La cinemática de los leptones, muestra que los datos simulados están modelados correctamente.

Dicha discusión y puesta a punto de todo este material experimental y computacional permitió a su vez tener un ambiente en línea (SWAN) en donde poder optimizar los resultados obtenidos. Tal optimización es el valor agregado en términos de física y se diseñó con revisión previa múltiples métodos estadísticos que pudiesen permitir hacer un fit en las muestras de datos reales y conseguir, de ser posible, una función que los describiese. Tal resultado permitirá conseguir una posible función que ayude a otros análisis aprovechando esta función, o modelo, y evitar utilizar grandes cantidades de MC para sumar esta contribución ZZ a sus análisis; como es el caso de los múltiple análisis en la búsqueda del bosón de Higgs que decae en un par de

bosones Z , y en menor medida, en bosones W y fotones.

Parte de formular esta hipótesis de modelado bajo distintas funciones estadísticas es poder estimar observables físicos que permitan comparar con resultados previos y dar una valoración del buen ajuste a los datos (o la falta de éste). En este caso, se planteó una hipótesis, haciendo referencia a una función de densidad de probabilidad (por sus siglas en inglés, PDF) de Landau, con el fin de buscar un buen acuerdo entre los datos y la hipótesis dada. Interesándose en utilizar dichos datos para hacer inferencias sobre un modelo probabilístico, evaluar la validez del modelo y determinar los valores de sus parámetros, es decir, los de la función a conseguir. En todos los casos se usaron los métodos de ajuste de mínimos cuadrados y de máxima verosimilitud para tales exámenes y valoraciones.

Subsecuentemente, se estimaron los parámetros de la PDF ajustada, mediante el método de mínimos cuadrados, con el software de minimización TMinuit en ROOT, obteniendo el valor mínimo de la cantidad χ^2 , que fue utilizado como prueba estadística de la bondad de ajuste, t . La hipótesis planteada es aceptada ya que la distribución de Landau describe nuestros datos con un nivel de confianza del 95 %.

La forma de χ^2 en función del valor más probable (MPV) sigue un perfil parabólico, cuyo valor mínimo coincide con el valor más probable del modelo, el cual es 213 GeV . Pero dado que nuestra misión es describir este proceso, y no la búsqueda de alguna resonancia, tal valor no debe dar cabida a discusiones en términos de partículas “a observar”, sino simplemente como parte del trabajo de conseguir una buena función que describa nuestra distribución de masa invariante de los leptones reconstruidos, es decir, la “forma” de nuestro proceso ZZ .

El método de mínimos cuadrados y de máxima verosimilitud, implementados en ROOT para histogramas de estadística baja, se usaron para evaluar la fortaleza de ambos métodos estadísticos, pudiéndose observar que el método de log-likelihood tiene la ventaja de tratar correctamente los contenedores de los histogramas con estadísticas bajas (o estadística nula), a sabiendas que la literatura nos dice que el parámetro de MPV es sesgado cuando se trabaja con el método de χ^2 con baja estadística:

- Es decir, en caso de estadísticas altas/bin, la distribución del contenido de bin se convierte en una distribución normal y ambos métodos como log-likelihood y el método χ^2 dan el mismo resultado, ver figura 7.9.
- Normalmente, los recuentos de un histograma siguen una distribución de Poisson.
- Además, cuando se usa el método χ^2 , existe el problema de que en ROOT, contenedores sin entradas (vacíos) se excluyen del ajuste, pero en realidad contienen información estadística valiosa.
- Es por ello, que se recomienda el método log-likelihood en caso de estadísticas de bin bajas, donde el método χ^2 da resultados incorrectos, en particular cuando hay varios bins vacíos.

Así pues, como sumario, se concluye que se ha podido generar y almacenar una reproducción computacional del análisis del experimento ATLAS en la nube y bajo una plataforma distinta (Jupyter), para luego pasar a la búsqueda de una posible optimización y mejora del análisis de ATLAS Open Data sobre la producción de ZZ en colisiones protón-protón a $\sqrt{s} = 13 \text{ TeV}$. Tales objetivos fueron logrados con un buen nivel de satisfacción, incluidos el hecho de revisar, aprender, utilizar y documentar herramientas de análisis consideradas de vanguardia en lo que se refiere a esta clase de física y su divulgación a nivel universitario.

Apéndice A

Resumen de Conceptos Teóricos

1. Grupo: Estructura algebraica que consiste en un conjunto de elementos que contiene una operación que combina dos elementos para formar un tercero. Las condiciones necesarias y suficientes para que G sea un grupo son [163]:

- Si $g_1 g_2$ Son dos elementos cualesquiera del un grupo G, entonces $g_1 \times g_2 \in G$ (Producto cerrado).
- $(g_1 \times g_2) \times g_3 = g_1 \times (g_2 \times g_3)$ (El producto es asociativo).
- Existe una unidad (elemento neutro) e $\in G$ $g \times e = e \times g = g$.
- Existe un inverso. Para todo g existe un g^{-1}
 $g \times (g^{-1}) = (g^{-1}) \times g = e$

Si se cumple con estas 4 propiedades es un grupo. Si además el producto es conmutativo se dice que el grupo es Abeliano [163].

2. Grupo Abeliano: Si g_1 y g_2 son dos elementos cualesquiera de un grupo G y son tales que $g_1 \times g_2 = g_2 \times g_1$, se dice que los elementos del grupo commutan y llamamos Grupo Abeliano a G. Específicamente sea G el grupo unitario unidimensional llamado $U(1)$ que consiste en el conjunto de parámetros continuos de todas las fases de una función de onda, $U(\alpha) = e^{i\alpha}$ donde α es un parámetro escalar real [164]:

$$U(\alpha_1)U(\alpha_2) = e^{i\alpha_1}e^{i\alpha_2} = e^{i\alpha_2}e^{i\alpha_1} = U(\alpha_2)U(\alpha_1) \quad (\text{A.1})$$

3. Grupo no-Abeliano: Cuando en el grupo los elementos no commutan, se dice que es un Grupo no-Abeliano. Se considera el grupo de Lie $SU(2)$ como el conjunto de las matrices unitarias 2×2 con determinante igual a 1, tales que cualquiera de sus elementos pueden ser escritos de la forma [164]:

$$U(\alpha) = e^{i\alpha\vec{\sigma}} \quad (\text{A.2})$$

donde $\alpha \in \mathbb{R}^3$ es un parámetro arbitrario y $\vec{\sigma} = (\sigma_1 \sigma_2 \sigma_3)$ corresponde a un vector con las matrices de Pauli generadoras del grupo. En general cualquier grupo de Lie $SU(N)$ resulta ser no-Abeliano.

4. Grupo de Lie: Es un grupo en el cual sus elementos son expresados por medio de un conjunto de parámetros continuos con una ley de multiplicación que depende de la suavidad de los parámetros [165]. Un grupo de rotaciones es un grupo de Lie porque las rotaciones se expresan en términos de los ángulos de rotación que son números reales, y no son elementos discretos.
5. SU(3) : Grupo especial unitario de tercer orden. Es un grupo de Lie de dimensión 8. Es el grupo de simetría de teorías físicas fundamentales como la cromodinámica cuántica [163]. Por ser $n = 3$, se tienen $n^2 - 1 = 8$ generadores, los cuales son el análogo de las matrices de Pauli, pero son 8 matrices 3×3 llamadas matrices de Gell-Mann λ_a ($a = 1, 2, 3..,8$) [166].
6. SU(2): Grupo especial unitario de segundo orden. Es un conjunto de matrices unitarias 2×2 con determinante uno [166]. Para este caso $n = 2$ representa 3 componentes, lo que son 3 partículas de calibre.
7. U(1) : Consiste en el grupo de todos los números complejos con valor absoluto igual a 1, o sea un círculo unitario en el plano complejo [166],

$$\mathbb{T} = z \in \mathbb{C} : |z| = 1 \quad (\text{A.3})$$

Estas actúan en el plano como una rotación alrededor del origen.

8. Luminosidad Instantánea: Es uno de los parámetros más relevantes en un colisionador de partícula ya que caracteriza el desempeño de un acelerador de partículas, el cual es proporcional a la tasa de eventos dN/dt , definido como [58]:

$$\frac{N}{t} = L \times \sigma \quad (\text{A.4})$$

donde σ es la sección eficaz del proceso considerado. La luminosidad instantánea depende de forma general de

$$L = \frac{N_1 N_2 f k}{4\pi S_x S_y} \quad (\text{A.5})$$

Donde N_1 y N_2 son las partículas del haz que revolucionan a una frecuencia f . S_x y S_y corresponden al área de la sección transversal del haz y k es el número de pulsos del haz o bunches, que circulan. Para nuestro caso la luminosidad sería [58]:

$$L = \frac{N_p^2 f k}{4\pi R^2} \quad (\text{A.6})$$

Donde N_p es el número de protones por cada bunch y R es el radio medio de la distribución de protones en el plano ortogonal a la distribución del haz.

La luminosidad se expresa normalmente en unidades de $cm^{-2}s^{-1}$, y ésta se define como luminosidad instantánea.

9. Luminosidad Integrada: Se define como [58]:

$$L_{int} = \int L dt \quad (\text{A.7})$$

y se expresa normalmente en unidades de fb^{-1} . El barn es una unidad muy usada en la física nuclear y de altas energías, se denota como b, y es una unidad de área igual a $10^{-28} m^2 = 100 fm^2$, donde el prefijo femto es $f = 10^{-15}$. El femtobarn inverso (fb^{-1}) es una unidad usada típicamente para medir el número de eventos de colisiones de partículas por femtobarn de la sección transversal del objetivo, y es la unidad convencional, como ya se mencionó, para la luminosidad integrada.

Apéndice B

Glosario

1. Bunch-crossing: En el LHC los protones dentro de los dos haz se agrupan en bunches, los cuales son apretados para disminuir su volumen, tal que la probabilidad de que ocurra una colisión aumente. Para los datos que se están usando en este trabajo, los bunches se cruzan cada 50 ns, cada vez que los bunches se crucen (bunchcrossing) ocurren en promedio 30 colisiones [95].
2. Evento: Para HEP, son los datos resultantes por cada bunch-crossing [95].
3. Interacción Hard-scatter: Denominación dada al vértice primario que tenga el mayor valor de la suma del cuadrado de los momentos transversos, $\sum \vec{p}_T^2$, para los tracks [67].
4. Cluster: Es una idea geométrica de agrupamiento. Para este trabajo se hace referencia a cluster como los depósitos de energía que depositan las diferentes partículas al pasar por los calorímetros electromagnéticos y hadrónicos [167].
5. Gaussian-Sum Filter: Es un algoritmo usado en la reconstrucción de electrones en ATLAS con el fin de mejorar la estimación de los parámetros de los tracks de los electrones [168].
Un electrón puede perder cantidades significativas de su energía inicial debido al bremsstrahlung, al interactuar con el material que atraviesa. Debido a la masa pequeña de los electrones, las perdidas radioactivas pueden ser muy relevantes, ocasionando una alteración en la trayectoria del electrón cuando se propaga por el campo magnético. Este algoritmo está diseñado para tomar en cuenta esas pérdidas de energía debido al bremsstrahlung.
6. Tracking: Acto de medir la dirección y magnitud del momentum de las partículas cargadas. Las partículas que entran al tracker liberan parte de su energía en el dispositivo. El tracker debe ser finamente segmentado tal que se pueda ser capaz de reconstruir con una buena precisión por donde pasaron las partículas. Se puede ver representado como la línea de la figura B.1 [169].
7. Tracker: Dispositivo o detector usado para el tracking [95].
8. Track: Hélice hecha por las partículas cargadas dentro del tracker [95].
9. Contacto: En inglés conocido como hit, es la medida de varios puntos en un track. Cuando una partícula cargada pasa por el detector de tracking al interactuar con éste va a dejar una serie de marcas. Se puede ver representado por las X en la figura B.1 [97].

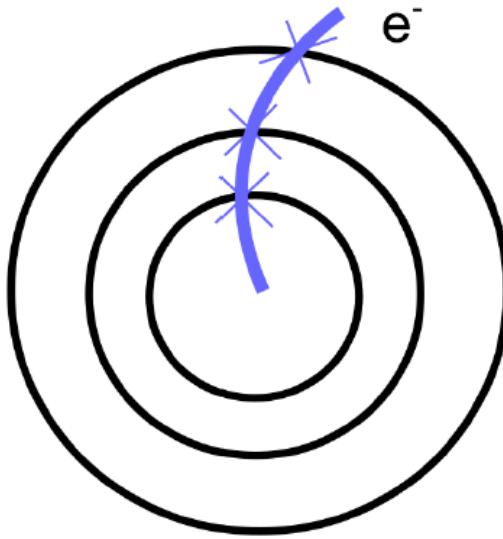


Figura B.1: Representación gráfica de un track con varios contactos. Tomado de [97].

10. Partícula Virtual: Una partícula virtual no es una partícula normal, se refiere a una perturbación en un campo que no es una partícula, nunca se va a encontrar por si sola, en cambio es causada por la presencia de otras partículas, y generalmente otros campos. El concepto viene de teoría de perturbaciones de QFT donde las interacciones entre partículas ordinarias son descritas en términos de intercambios de partículas virtuales.

Una partícula virtual no tiene necesariamente la misma masa que su correspondiente partícula real, de hecho las partículas virtuales pueden tener cualquier masa, esto viene de la relatividad especial donde la energía, momentum y masa de una partícula libre viene dada por $E^2 - \vec{p}^2 c^2 = m^2 c^4$ pero para una partícula virtual $E^2 - \vec{p}^2 c^2$ puede tomar cualquier valor, esto se interpreta muchas veces como que los procesos virtuales violan la conservación de la energía.

Como consecuencia del principio de incertidumbre en la mecánica cuántica, cualquier objeto o proceso que exista por un periodo limitado de tiempo no puede tener un valor preciso de energía, por esta razón las partículas virtuales, las cuales existen de forma temporal mientras estas son intercambiadas por otras partículas, no obedecen necesariamente la relación mass-shell, o sea se dice que están off shell porque no satisfacen las ecuaciones de energía momentum. Sin embargo, mientras más tiempo exista una partícula, más cercana está a cumplir la relación mass-shell. Una partícula virtual que existe por un tiempo largo arbitrario es simplemente una partícula ordinaria [5].

11. Background o fondo: Los datos contienen tanto eventos provenientes de la señal como eventos de background o fondo. Los eventos del fondo son aquellos eventos en los datos que no provienen de la señal de interés, normalmente pueden ser combinaciones de estados finales de partículas o procesos identificados de forma errónea, estos procesos son rechazados tal que el proceso físico deseado pueda ser observado [95].
12. Masa Transversa: En la física de partículas, la masa transversa es una cantidad definida

por su utilidad, ya que es invariante ante transformaciones de Lorentz a lo largo de la dirección z [5]. En unidades naturales es:

$$m_T^2 = m^2 + p_x^2 + p_y^2 \quad (\text{B.1})$$

donde la dirección z se encuentra a lo largo del haz, por lo tanto p_x , y p_y son los momentos transversos y m es la masa invariante. La ecuación que se utiliza para usar esta variable como discriminante, en los cortes aplicados a los eventos, con las variables proporcionadas por las n-tuplas es [5]:

$$m_T = \sqrt{2\vec{p}_T^l \vec{E}_T^{miss} [1 - \cos(\Delta\phi(l, \vec{E}_T^{miss}))]} \quad (\text{B.2})$$

donde \vec{p}_T^l es el momentum transverso del lepton, \vec{E}_T^{miss} es el momentum transverso perdido y $\Delta\phi(l, \vec{E}_T^{miss})$ es el ángulo que hay entre el momentum transverso perdido y el lepton.

13. Masa Invariante: Cuando se ven dos o mas tracks en el detector en un evento, originados del mismo punto, es decir con un vértice en común, puede que éstas pertenezcan a una misma partícula que decayó en esos tracks. Para poder verificar esta hipótesis se debe calcular la masa invariante de la partícula inicial e investigar si estos productos de decaimiento vienen de esta partícula, es decir, investigar si los productos pertenecen al decaimiento de la partícula que tiene una masa invariante en particular [5].

La masa invariante, es la masa en el sistema de reposo de la particula, está definida como [5]:

$$m \equiv \sqrt{-p_\mu p^\mu} \quad (\text{B.3})$$

$$p_\mu p^\mu = -E^2 + \vec{p}^2 \quad (\text{B.4})$$

Donde, $p = (E, \vec{p})$ es el 4-momentum.

La masa invariante es la norma del 4-momentum. Sustituyendo las componentes del 4-vector se llega a:

Para un sistema de partículas.

$$m^2 \equiv E^2 - ||\vec{p}||^2 \quad (\text{B.5})$$

$$m^2 \equiv (\sum E)^2 - ||\sum \vec{p}||^2 \quad (\text{B.6})$$

14. TLorenzVector: Es una clase general de cuadrivectores de ROOT, que puede ser usada ya sea para la descripción de la posición y el tiempo (x, y, z, t) o el momentum y la energía (p_x, p_y, p_z, E) [170].

Apéndice C

Marco Estadístico

En estadística, estamos interesados en usar una muestra dada de datos para hacer inferencias sobre un modelo probabilístico, para evaluar la validez del modelo o para determinar los valores de sus parámetros. Esto se puede hacer definiendo un estadístico de bondad de ajuste, el cual es una función de los datos cuyo valor refleja el nivel de acuerdo entre los datos y una hipótesis dada, que se describirá más adelante en este apéndice. En nuestro caso, se plantea una hipótesis, que hace referencia a una función de densidad de probabilidad (por sus siglas en inglés, P.D.F.s), a causa de buscar un buen acuerdo entre los datos y la hipótesis dada.

A continuación, se ofrece una descripción de las funciones de densidad de probabilidad, los métodos estadísticos usados, los cuales son el método de mínimos cuadrados y el método de máxima verosimilitud, pruebas estadísticas que son las pruebas de hipótesis y pruebas de bondad de ajuste e intervalos de confianza y valor de probabilidad (valor-p) [138].

C.1. Funciones de Densidad de Probabilidad

Un histograma es la mejor forma de poder representar gráficamente una cierta distribución de datos numéricos. En el eje horizontal está la cantidad o la variable que se está midiendo. En el eje vertical se mide cuántas veces se obtiene un mismo valor para la variable. Se define los Bins, como intervalos en los cuales están divididos los datos. El histograma cuenta las entradas para cada bin y dibuja una barra de ese tamaño. Tener una visualización así da una mejor impresión de cómo es la distribución [58].

La curva que se crea cuando se reduce el bin tomando un límite al continuo es denominada como $f(x)$ y es llamada función de densidad de probabilidad. La probabilidad de que una variable aleatoria continua x esté sobre un intervalo de valores se puede hallar al integrar debajo de la curva $f(x)$ [58].

Formalmente, la definición de función de densidad de probabilidad es una generalización al continuo de la distribución de probabilidad. Si se considera un espacio la medición x es un punto $\vec{x} = (x_1, \dots, x_n) \in \Omega \subseteq \mathbb{R}^n$, cada extracción aleatoria sería un experimento. El resultado es una medición de \vec{x} , lo cual es punto en el espacio Ω . Se puede asociar a cualquier punto \vec{x} una densidad de probabilidad $f(\vec{x}) = f(x_1, \dots, x_n)$ lo cual es un valor real mayor o igual que cero. La probabilidad de un evento A , donde $A \subseteq \Omega$, es decir la probabilidad que $\vec{x} \in A$, está dada por [171]:

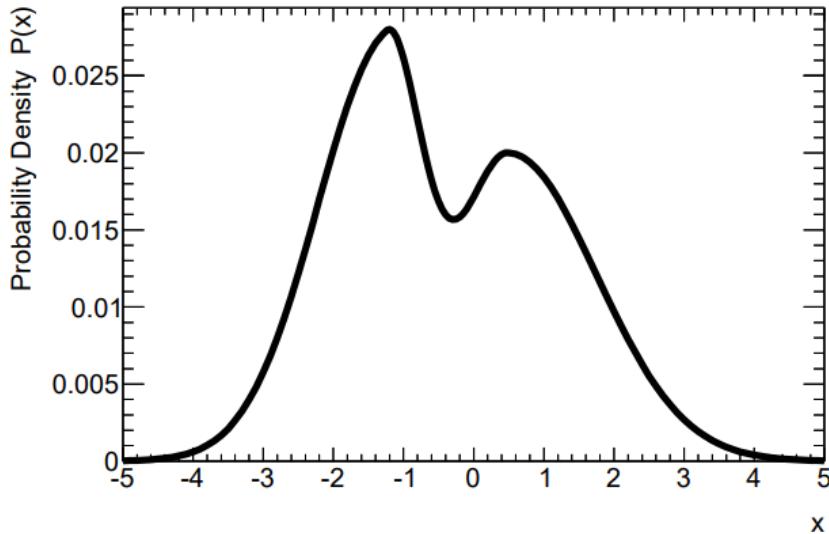


Figura C.1: Función de Densidad de Probabilidad (PDF) para una variable x . Aparte de las condiciones de normalización las PDFs pueden exhibir uno o varios máximos locales o mínimos locales [171].

$$P(A) = \int_A f(x_1, \dots, x_n) d^n x \quad (\text{C.1})$$

La condición de normalización para una distribución de probabilidad discreta ($\sum_{i=1}^n P(x_i = 1)$), puede ser generalizada para una función de distribución de probabilidad continua de la siguiente manera [171]:

$$\int_{\Omega} f(x_1, \dots, x_n) dx^n = 1 \quad (\text{C.2})$$

En una dimensión se puede reescribir como,

$$\int_{-\infty}^{+\infty} f(x_1, \dots, x_n) dx = 1 \quad (\text{C.3})$$

C.1.1. Caracterización de la Forma de una Función de Densidad de Probabilidad

- En la práctica, la verdadera dependencia funcional de una PDF es a menudo desconocida.
- La información sobre su forma solamente puede extraerse a partir de una muestra de talla finita (contiene N eventos).
- Si consideramos que la PDF es de tipo paramétrico, la caracterización de su forma es un procedimiento para estimar los valores numéricos de sus parámetros.
- No existe una receta única para la elección de los parámetros a ser estimados, con lo que el proceso es intrínsecamente incompleto.

- Se puede sin embargo confirmar en la práctica que el proceso de caracterización de forma es bastante poderoso, si los parámetros seleccionados proveen información útil y complementaria.

La función de densidad de probabilidad usada para buscar un buen acuerdo entre los datos y las hipótesis dadas, es la distribución de Landau la cual se describe a continuación.

C.1.2. La Distribución de Landau

Representa una distribución estadística definida y sustentada en el conjunto de números reales \mathbb{R} y parametrizada por un número real μ (llamado “parámetro de ubicación” y corresponde aproximadamente al valor más probable) y por un número real c (llamado “parámetro de escala”). En general, la función de densidad de probabilidad (PDF) de una distribución de Landau es unimodal con un solo “pico” (es decir, un máximo global), aunque su forma general está determinada por los valores de μ y c [141].

La función de densidad de probabilidad, tal como fue escrita originalmente por Landau, está definida por la integral compleja [141]:

$$p(x) = \frac{1}{2i\pi} \int_{a-i\infty}^{a+i\infty} e^{s \log(s) + xs} ds \quad (\text{C.4})$$

donde a es un número real positivo arbitrario, y \log se refiere al logaritmo natural.

Otra forma de escribir la función de densidad de probabilidad es [141],

$$p(x) = \frac{1}{\pi} \int_0^{\infty} e^{-t \log(t) - xt} \sin(\pi t) dt \quad (\text{C.5})$$

La familia completa de distribuciones de Landau se obtiene al extender la distribución original a una familia de distribuciones estables con parámetros de estabilidad $\alpha = 1$ y de asimetría $\beta = 1$ con la función característica [141]:

$$\phi(t; \mu, c) = \exp\left(it\mu - \frac{2ict}{\pi} \log|t| - c|t|\right) \quad (\text{C.6})$$

C.2. Método de Mínimos Cuadrados

Probablemente la aplicación más común de la minimización en la investigación científica es el ajuste por mínimos cuadrados ó LS (del inglés, Least Squares), donde la función a minimizar es la suma de los cuadrados de las desviaciones entre los valores medidos y las predicciones de un modelo que contiene parámetros variables. Dicho método coincide con el método de máxima verosimilitud cuando la densidad de probabilidad a priori es gaussiana [138].

Se considera un conjunto de N medidas independientes y_i en puntos conocidos x_i . Se supone que la medida y_i tiene una distribución gaussiana con media $f(x_i; \theta)$ y varianza conocida σ_i^2 . El objetivo es construir estimadores para los parámetros desconocidos θ . La función de verosimilitud contiene la suma de cuadrados [138],

$$\chi^2(\theta) = -2 \ln L(\theta) + ctte = \sum_{i=1}^N \frac{(y_i - f(x_i; \theta))^2}{\sigma_i^2} \quad (\text{C.7})$$

El conjunto de parámetros θ que maximizan L es el mismo que los que minimizan χ^2 . El mínimo de la ecuación (C.7) define los estimadores de mínimos cuadrados $\hat{\theta}$ para el caso más general donde los y_i no tienen distribución gaussiana siempre que sean independientes [138].

Al construir la cantidad $\chi^2(\theta)$, se requieren las varianzas. A menudo, estas cantidades no se conocen a priori y deben estimarse a partir de los datos; en nuestro caso el valor medido y_i representa un número contado de eventos en el contenedor del histograma [138]. Los y_i representan una variable de Poisson, para la cual la varianza es igual a la media, entonces se puede estimar la varianza a partir del valor predicho, $f(x_i; \theta)$, o del número observado en sí, y_i . En la primera opción, las variaciones se convierten en funciones de los parámetros ajustados, lo que puede dar lugar a dificultades de cálculo. La segunda opción puede no estar definida si y_i es cero, y en ambos casos para y_i pequeños, la varianza se estimará mal. Una alternativa para solucionar este dilema es utilizar el método de máxima verosimilitud [138].

C.3. Método de Máxima Verosimilitud

Las frecuencias medidas en un histograma ("barras" de histograma) están sujetas a fluctuaciones estadísticas que siguen la distribución de Poisson o Gaussiana. Ajustar histogramas con estadísticas bajas no es lo ideal con el método χ^2 , porque las fluctuaciones del intervalo no son gaussianas, por lo que el método de máxima verosimilitud, se construye asumiendo una función de densidad de probabilidad de Poisson para cada contenedor [151].

Desde un punto de vista teórico, el método general de estimación más importante es el método de máxima verosimilitud (en inglés, "Maximum Likelihood Method") [138]. Se tiene un conjunto de N cantidades x_i medidas independientemente y se quiere ver qué PDF va a representar mejor lo que se observa en el experimento, tal que se pueda extraer información luego. La unión de la PDF con los datos x_i es una función que depende de un conjunto de parámetros θ . La función de verosimilitud va a representar los datos y dar los parámetros θ que deberían ayudar a hacer que el modelo sea lo más cercano a los datos posibles, donde $\theta = (\theta_1, \dots, \theta_n)$ es un conjunto de n parámetros cuyos valores se desconocen [138]. El método de máxima verosimilitud toma los estimadores $\hat{\theta}$ como aquellos valores de θ que maximizan la función de verosimilitud [138],

$$L(\theta) = \prod_{i=1}^N f(x_i; \theta) \quad (\text{C.8})$$

La función se debe evaluar con los datos obtenidos y considerarla como una función de los parámetros. Para un valor de θ cercano a la realidad, se espera una alta probabilidad de que el modelo sea lo más cercano a los datos que se tienen. Si el parámetro θ está lejos del valor real, se tiene una probabilidad baja de observar lo que se midió. La función de verosimilitud no es una P.D.F. para los parámetros θ [171].

Por lo general, es más fácil trabajar con $\ln L$, y dado que ambos se maximizan para los mismos valores de parámetro θ , los estimadores de máxima verosimilitud (por sus siglas en inglés, ML: Maximum likelihood) se pueden encontrar resolviendo las ecuaciones de verosimilitud [138],

$$\frac{\partial \ln L}{\partial \theta_i} = 0, i = 1, \dots, n \quad (\text{C.9})$$

Los estimadores de máxima verosimilitud son importantes porque son aproximadamente insesgados. Por otro lado, el tamaño n de la muestra de datos x_1, \dots, x_n puede considerarse fijo

o se puede optar por tratarlo como una variable distribuida por Poisson; esta última opción se denomina “Ajuste de Probabilidad Extendida” [172].

En particular, el método de máxima verosimilitud generalmente no tiene problemas con contenedores con cero o pocos eventos, mientras que los otros métodos pierden eficiencia debido a la no normalidad de la distribución del número de eventos [147]. Estos inconvenientes suelen solucionarse llevando acabo el método de máxima verosimilitud con datos agrupados, el cual se presenta en la siguiente sección, y es utilizado por ROOT cuando se selecciona el método de ajuste “L”.

C.4. Máxima Verosimilitud con Datos Agrupados

Se considera un n_{tot} de observaciones de una variable aleatoria x distribuida de acuerdo con un p.d.f. $f(x; \theta)$ para el que nos gustaría estimar el parámetro desconocido $\theta = (\theta_1, \dots, \theta_m)$ [172]. Para muestras de datos muy grandes, la función de verosimilitud logarítmica se vuelve difícil de calcular, ya que se debe sumar $\log f(x; \theta)$ para cada valor x_i . En tales casos, en lugar de registrar el valor de cada medición, normalmente se hace un histograma, lo que arroja un cierto número de entradas $n = (n_1, \dots, n_N)$ en N bins. Los valores esperados $\nu = (\nu_1, \dots, \nu_N)$ del número de entradas están dados por [172],

$$\nu_i(\theta) = n_{tot} \int_{x^{min}}^{x^{max}} f(x; \theta) dx, \quad (\text{C.10})$$

Donde x^{min} y x^{max} son los límites de bin [172]. En muchos problemas se puede querer considerar el número total de entradas n_{tot} como una variable aleatoria de una distribución de Poisson con media ν_{tot} . Es decir, que la medición consiste en determinar primero n_{tot} a partir de una distribución de Poisson y luego distribuir n_{tot} observaciones de x en un histograma con N intervalos, dando $n = (n_1, \dots, n_N)$. El pdf conjunto para n_{tot} y n_1, \dots, n_N es el producto de una distribución de Poisson y una distribución multinomial [172],

$$f_{joint}(n; \nu) = \prod_{i=1}^N \frac{\nu_i^{n_i} e^{-\nu_i}}{n_i!} \quad (\text{C.11})$$

donde el número esperado de entradas en cada bin ν_i ahora depende de los parámetros θ y ν_{tot} ,

$$\nu_i(\nu_{tot}, \theta) = \nu_{tot} \int_{x^{min}}^{x^{max}} f(x; \theta) dx, \quad (\text{C.12})$$

A partir del p.d.f. conjunto (6.9) se ve que el problema es equivalente a tratar el número de entradas en cada intervalo como una variable aleatoria de Poisson independiente n_i con valor medio ν_i . Tomando el logaritmo del p.d.f. conjunto y descartando términos que no dependen de los parámetros, da como resultado [172],

$$\log L(\nu_{tot}, \theta) = -\nu_{tot} + \sum_{i=1}^N n_i \log \nu_i(\nu_{tot}, \theta) \quad (\text{C.13})$$

Esta es la función logarítmica de verosimilitud extendida, ahora para el caso de datos agrupados. Si no existe una relación funcional entre ν_{tot} y θ , entonces se obtiene $\nu_{tot} = n_{tot}$ y las estimaciones de $\hat{\theta}$ son las mismas que cuando no se incluyen los términos de poisson para n_{tot}

[172]. Si ν_{tot} se da en función de θ , entonces las varianzas de los estimadores $\hat{\theta}$ se reducen en general al incluir la información de n_{tot} [172].

C.5. Pruebas Estadísticas

Además de estimar parámetros, a menudo se desea evaluar la validez de ciertas afirmaciones sobre la distribución subyacente de los datos. Las pruebas de hipótesis proporcionan una regla para aceptar o rechazar hipótesis según el resultado de una medición [138]. En las pruebas de bondad de ajuste se da la probabilidad de obtener un nivel de incompatibilidad con una determinada hipótesis mayor o igual al nivel observado con los datos reales.

C.5.1. Pruebas de Hipótesis

Se considera un experimento cuyo resultado se caracteriza por un vector de datos \mathbf{x} [138]. Una hipótesis es un enunciado sobre la distribución de \mathbf{x} . Podría, por ejemplo, definir completamente el p.d.f. para los datos (una hipótesis simple) o podría especificar solo la forma funcional de la p.d.f., dejando abiertos los valores de uno o más parámetros (una hipótesis compuesta) [138]. Una prueba estadística es una regla que establece para qué valores de \mathbf{x} se debe rechazar una hipótesis dada (a menudo llamada hipótesis nula, H_0). Esto se hace definiendo una región del espacio \mathbf{x} llamada región crítica; si el resultado del experimento cae en esta región, se rechaza H_0 [138]. De manera equivalente, se puede decir que la hipótesis se acepta si \mathbf{x} se observa en la región de aceptación, es decir, el complemento de la región crítica. Aquí se entiende que “aceptado” significa simplemente que la prueba no rechazó H_0 .

Rechazar H_0 si es verdadera se denomina error de primer tipo [172]. La probabilidad de que esto ocurra se denomina nivel de significancia de la prueba, α , que a menudo se elige para que sea igual a algún valor preespecificado. También puede ocurrir que H_0 sea falsa y la hipótesis verdadera esté dada por alguna alternativa, H_1 [172]. Si se acepta H_0 en tal caso, esto se denomina error de segundo tipo. La probabilidad de que esto ocurra, β , depende de la hipótesis alternativa, por ejemplo, H_1 , y $1 - \beta$ se denomina potencia de la prueba para rechazar H_1 [172].

En la Física de Altas Energías, las componentes de \mathbf{x} podrían representar las propiedades medidas de los eventos candidatos, y la región de aceptación se define por los cortes que se imponen para seleccionar eventos de un determinado tipo deseado. Es decir, H_0 podría representar la hipótesis de la señal y varias alternativas, H_1 , H_2 , etc., podrían representar procesos de fondo [138].

A menudo, en lugar de usar la muestra completa de datos \mathbf{x} , es conveniente definir una estadística de prueba, t , que puede ser un solo número o, en cualquier caso, un vector con menos componentes que \mathbf{x} [138]. Cada hipótesis para la distribución de \mathbf{x} determinará una distribución para t , y la región de aceptación en el espacio \mathbf{x} corresponderá a un rango específico de valores de t [138].

En terminología de física de partículas, la probabilidad de aceptar la hipótesis de la señal, H_0 , es la eficiencia de selección, es decir, uno menos el nivel de significación [138].

A menudo, uno quiere cuantificar el nivel de concordancia entre los datos y una hipótesis sin referencia explícita a hipótesis alternativas, es por ello que se define en la siguiente sección las pruebas de bondad de ajuste.

C.5.2. Pruebas de Bondad de Ajuste

Para cuantificar el nivel de concordancia entre los datos y una hipótesis sin referencia explícita a hipótesis alternativas se hace definiendo una estadística de bondad de ajuste, t , que es una función de los datos cuyo valor refleja de alguna manera el nivel de acuerdo entre los datos y la hipótesis, un ejemplo de esto es una prueba de χ^2 en la cual requiere la comparación entre el χ^2_{prueba} (obtenido de los datos) con el χ^2_{tabla} (Ver figura C.2), si el valor estadístico de prueba (χ^2_{prueba}) es menor que el valor tabulado (χ^2_{tabla}), la hipótesis nula es aceptada, caso contrario, H_0 se rechaza [138]. Al estimar parámetros mediante el método de mínimos cuadrados, se obtiene el valor mínimo de la cantidad χ^2 , que puede utilizarse como estadístico de bondad de ajuste, t [138]. El usuario debe decidir qué valores del estadístico corresponden a mejores o peores niveles de acuerdo con la hipótesis en cuestión; para muchas estadísticas de bondad de ajuste existe una opción obvia [138].

La hipótesis en cuestión, digamos, H_0 , determinará la p.d.f. $g(t|H_0)$ para el estadístico [138]. La bondad de ajuste se cuantifica dando el valor p , definido como la probabilidad de encontrar t en la región de igual o menor compatibilidad con H_0 que el nivel de compatibilidad observado con los datos reales. Por ejemplo, si t se define de tal manera que los valores grandes corresponden a un acuerdo pobre con la hipótesis, entonces el valor p sería [138].

$$\int_{t_{obs}}^{\infty} g(t|H_0) dt, \quad (\text{C.14})$$

donde t_{obs} es el valor del estadístico obtenido en el experimento real. El valor p no debe confundirse con el nivel de significación de una prueba o el nivel de confianza de un intervalo de confianza (Sección D.6.), los cuales son constantes preespecificadas [138].

El valor- p es una probabilidad, la cual es el resultado de dicha prueba estadística. Es la probabilidad de un acuerdo "peor" que el observado [58]. Los valores de p pequeños corresponden a pruebas sólidas. Si el valor p está por debajo del límite predefinido, los resultados se designan como estadísticamente significativos [58]. Se puede ver representado en la figura C.3. La hipótesis propuesta puede ser rechazada si esta probabilidad es menor o igual a un valor límite α , que se escoge de forma arbitraria, y es referido como la significación estadística [58].

Solamente el valor- p no permite tener una declaración acerca de las hipótesis planteadas, es usado sobre todo cuando los resultados no son significativos. Por ello los límites de confianza contienen mayor información, la cual es presentada en la siguiente sección. Sin embargo, en el análisis final la definición de este límite es arbitrario, y el valor- p puede ser dado aunque no se haya seleccionado una significación estadística.

C.6. Intervalos de Confianza

Para informar errores estadísticos en una medición es a través de un intervalo de confianza, el cual fue desarrollado por primera vez por Neyman [172]. Por otro lado el término nivel de confianza se refiere a la probabilidad de cobertura de un intervalo de confianza.

Cuando el objetivo de un experimento es determinar un parámetro θ , el resultado suele expresarse citando, además de la estimación puntual, algún tipo de intervalo que refleje la precisión estadística de la medición. En el caso más simple, esto puede estar dado por el valor estimado del parámetro $\hat{\theta}$ más o menos una estimación de la desviación estándar de $\hat{\theta}$, $\sigma_{\hat{\theta}}$ [138]. Esto es válido, si el p.d.f. del estimador es gaussiano [138].

39

**PERCENTILE VALUES (χ^2_p)
FOR χ^2 (CHI-SQUARE)
DISTRIBUTION**

with n degrees of freedom (shaded area = p)

| n | $\chi^2_{.995}$ | $\chi^2_{.99}$ | $\chi^2_{.975}$ | $\chi^2_{.95}$ | $\chi^2_{.90}$ | $\chi^2_{.75}$ | $\chi^2_{.50}$ | $\chi^2_{.25}$ | $\chi^2_{.10}$ | $\chi^2_{.05}$ | $\chi^2_{.025}$ | $\chi^2_{.01}$ | $\chi^2_{.005}$ |
|-----|-----------------|----------------|-----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|-----------------|----------------|-----------------|
| 1 | 7.88 | 6.63 | 5.02 | 3.84 | 2.71 | 1.32 | .455 | .102 | .0158 | .0039 | .0010 | .0002 | .0000 |
| 2 | 10.6 | 9.21 | 7.38 | 5.99 | 4.61 | 2.77 | 1.39 | .575 | .211 | .103 | .0506 | .0201 | .0100 |
| 3 | 12.8 | 11.3 | 9.35 | 7.81 | 6.25 | 4.11 | 2.37 | 1.21 | .584 | .352 | .216 | .115 | .072 |
| 4 | 14.9 | 13.3 | 11.1 | 9.49 | 7.72 | 5.39 | 3.36 | 1.92 | 1.06 | .711 | .484 | .297 | .207 |
| 5 | 16.7 | 15.1 | 12.8 | 11.1 | 9.24 | 6.63 | 4.35 | 2.67 | 1.61 | 1.15 | .831 | .554 | .412 |
| 6 | 18.5 | 16.8 | 14.4 | 12.6 | 10.6 | 7.84 | 5.35 | 3.45 | 2.20 | 1.64 | 1.24 | .872 | .676 |
| 7 | 20.3 | 18.5 | 16.0 | 14.1 | 12.0 | 9.04 | 6.35 | 4.25 | 2.83 | 2.17 | 1.69 | 1.24 | .989 |
| 8 | 22.0 | 20.1 | 17.5 | 15.5 | 13.4 | 10.2 | 7.34 | 5.07 | 3.49 | 2.73 | 2.18 | 1.65 | 1.34 |
| 9 | 23.6 | 21.7 | 19.0 | 16.9 | 14.7 | 11.4 | 8.34 | 5.90 | 4.17 | 3.33 | 2.70 | 2.09 | 1.73 |
| 10 | 25.2 | 23.2 | 20.5 | 18.3 | 16.0 | 12.5 | 9.34 | 6.74 | 4.87 | 3.94 | 3.25 | 2.56 | 2.16 |
| 11 | 26.8 | 24.7 | 21.9 | 19.7 | 17.3 | 13.7 | 10.3 | 7.58 | 5.58 | 4.57 | 3.82 | 3.05 | 2.60 |
| 12 | 28.3 | 26.2 | 23.3 | 21.0 | 18.5 | 14.8 | 11.3 | 8.44 | 6.30 | 5.23 | 4.40 | 3.57 | 3.07 |
| 13 | 29.8 | 27.7 | 24.7 | 22.4 | 19.8 | 16.0 | 12.3 | 9.30 | 7.04 | 5.89 | 5.01 | 4.11 | 3.57 |
| 14 | 31.3 | 29.1 | 26.1 | 23.7 | 21.1 | 17.1 | 13.3 | 10.2 | 7.79 | 6.57 | 5.63 | 4.66 | 4.07 |
| 15 | 32.8 | 30.6 | 27.5 | 25.0 | 22.3 | 18.2 | 14.3 | 11.0 | 8.55 | 7.26 | 6.26 | 5.23 | 4.60 |
| 16 | 34.3 | 32.0 | 28.8 | 26.3 | 23.5 | 19.4 | 15.3 | 11.9 | 9.31 | 7.96 | 6.91 | 5.81 | 5.14 |
| 17 | 35.7 | 33.4 | 30.2 | 27.6 | 24.8 | 20.5 | 16.3 | 12.8 | 10.1 | 8.67 | 7.56 | 6.41 | 5.70 |
| 18 | 37.2 | 34.8 | 31.5 | 28.9 | 26.0 | 21.6 | 17.3 | 13.7 | 10.9 | 9.39 | 8.23 | 7.01 | 6.26 |
| 19 | 38.6 | 36.2 | 32.9 | 30.1 | 27.2 | 22.7 | 18.3 | 14.6 | 11.7 | 10.1 | 8.91 | 7.63 | 6.84 |

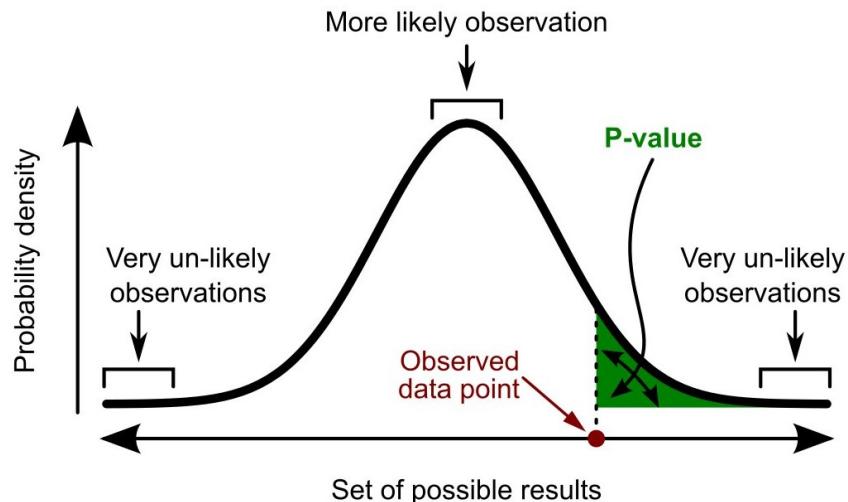
Figura C.2: Percentiles de la distribución χ^2 [149]

Figura C.3: Representación gráfica del valor-p. La coordenada vertical es la densidad de probabilidad para cada resultado, calculado bajo la hipótesis nula. El valor-p es el área bajo la curva más adelante del punto de datos observados. Tomado de [173]

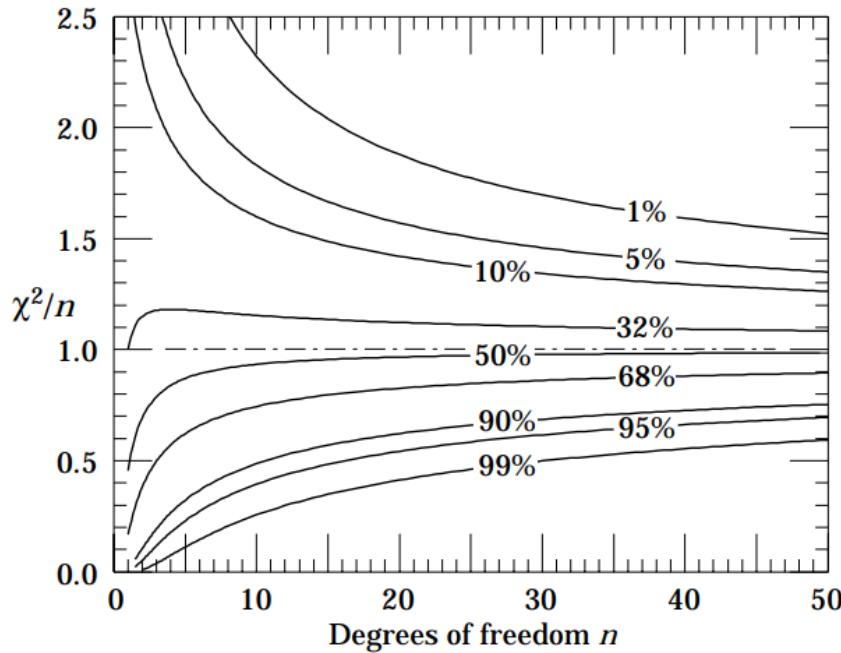


Figura C.4: El χ^2/n , para n grados de libertad. Las curvas muestran en función de n el χ^2/n que corresponde a un valor de p dado [138].

Como tal, la desviación estándar a menudo se informa como la incertidumbre estadística de una medición y se denomina error estándar.

C.6.1. Intervalos de Confianza utilizando la Función de Verosimilitud o χ^2

Incluso en el caso de un estimador no gaussiano, el intervalo de confianza se puede determinar con una técnica aproximada simple que utiliza la función de verosimilitud o, de manera equivalente, la función χ^2 donde se tiene $L = \exp(-\chi^2/2)$. Se considera primero un estimador de máxima verosimilitud $\hat{\theta}$ para un parámetro θ en el límite de muestra grande la función de verosimilitud en sí se vuelve gaussiana centrada en la estimación de ML $\hat{\theta}$ [172],

$$L(\theta) = L_{max} \exp\left(\frac{-(\theta - \hat{\theta})^2}{2\sigma_{\hat{\theta}}^2}\right) \quad (\text{C.15})$$

Una receta simple para estimar $\sigma_{\hat{\theta}}$, es que al cambiar el parámetro θ por N desviaciones estándar en la ecuación 6.24 de la sección 6.7 de la referencia [172], la función logarítmica de verosimilitud disminuye en $N^2/2$ desde su valor máximo,

$$\log L(\hat{\theta} \pm N\sigma_{\hat{\theta}}) = \log L_{max} - N^2/2 \quad (\text{C.16})$$

Para un estimador distribuido gaussiano $\hat{\theta}$, el intervalo de confianza central del 68,3% se puede construir a partir del estimador y su desviación estándar estimada $\hat{\sigma}_{\hat{\theta}}$ como $[a, b] = [\hat{\theta} - \hat{\sigma}_{\hat{\theta}}, \hat{\theta} + \hat{\sigma}_{\hat{\theta}}]$ [172]. El intervalo de confianza central del 68,3% viene dado por los valores de θ en los que la función logarítmica de verosimilitud disminuye a la mitad de su valor máximo (esto

supone, por supuesto, que $\hat{\theta}$ es el estimador ML y, por lo tanto, corresponde al de la función de verosimilitud) [172]. El resultado final de la medición de θ se informa simplemente como $\hat{\theta} \pm \sigma_{\hat{\theta}}$.

De hecho, se puede demostrar que incluso si la función de verosimilitud no es una función gaussiana de los parámetros, el intervalo de confianza central $[a, b] = [\hat{\theta} - c, \hat{\theta} + d]$ todavía se puede aproximar usando [172],

$$\log L(\hat{\theta}_{-c}^{+d}) = \log L_{max} - N^2/2 \quad (\text{C.17})$$

Donde $N = \Phi^{-1}(1 - \gamma/2)$ es el cuantil de la gaussiana estándar correspondiente al nivel de confianza deseado $1 - \gamma$ (por ejemplo, $N = 1$ para el intervalo de confianza central de 68,3%; ver tabla C.1). En el caso de ajustes por mínimos cuadrados con errores gaussianos, es decir, con $\log L = -\chi^2/2$, la prescripción se convierte en [172],

$$\chi^2(\hat{\theta}_{-c}^{+d}) = \chi^2_{min} + N^2 \quad (\text{C.18})$$

Las ecuaciones (C.16) y (C.18) representan uno de los métodos más utilizados para estimar las incertidumbres estadísticas. Se debe tener en cuenta que estas mismas se interpretan como una aproximación al intervalo de confianza clásico, es decir, un intervalo aleatorio construido para incluir el verdadero valor del parámetro con una probabilidad dada [172].

Tabla C.1: Los valores del nivel de confianza para diferentes valores del cuantil de la gaussiana estándar Φ^{-1} : Para los intervalos centrales (izquierda) el cuantil $\Phi^{-1}(1 - \gamma/2)$ y el nivel de confianza ($1 - \gamma$); para intervalos unilaterales (derecha) el cuantil $\Phi^{-1}(1 - \alpha)$ y el nivel de confianza ($1 - \alpha$) [172].

| $\Phi^{-1}(1 - \gamma/2)$ | $(1 - \gamma)$ | $\Phi^{-1}(1 - \alpha)$ | $(1 - \alpha)$ |
|---------------------------|----------------|-------------------------|----------------|
| 1 | 0.6827 | 1 | 0.8413 |
| 2 | 0.9544 | 2 | 0.9772 |
| 3 | 0.9973 | 3 | 0.9987 |

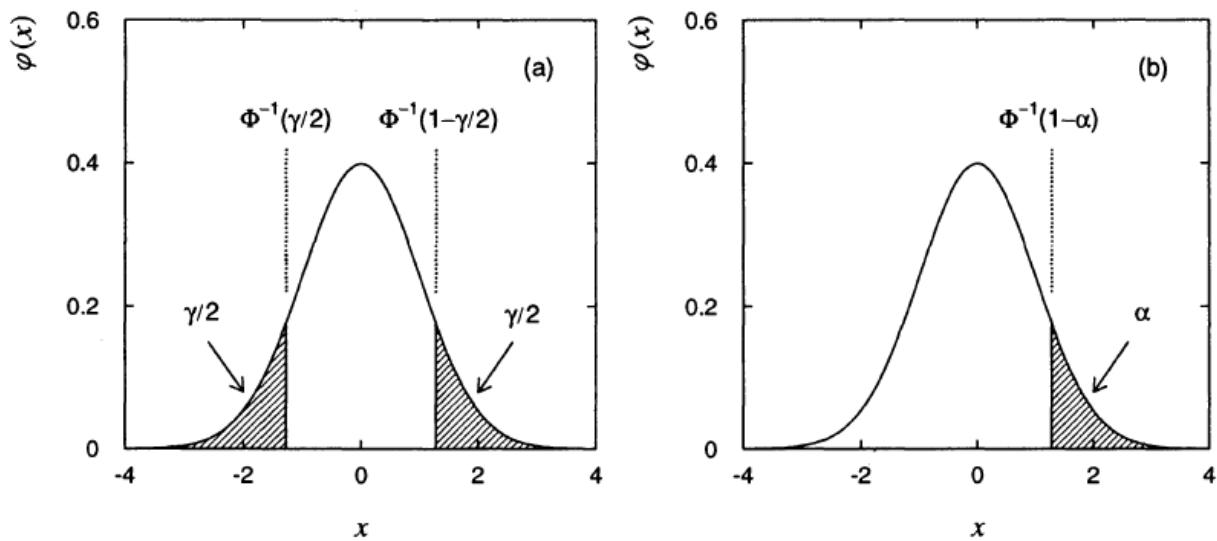


Figura C.5: El p.d.f. Gaussiano estándar $\varphi(x)$ que muestra la relación entre los cuantiles Φ^{-1} y el nivel de confianza para (a) un intervalo de confianza central y (b) un intervalo de confianza unilateral [172].

Apéndice D

ROOT: Marco de Análisis de Datos

D.1. Descripción de las PDFs en ROOT

En ROOT, la biblioteca “MathCore” [174] proporciona una colección de funciones, clases $C +$ y clases ROOT para computación numérica HEP. El “MathCore” es un conjunto mínimo autoconsistente de herramientas necesarias para la computación numérica básica. Lo siguiente está incluido en la biblioteca “MathCore” [174]:

- **Funciones Especiales**
- **Funciones Estadísticas:** Funciones utilizadas en estadística, como las funciones de densidad de probabilidad y las funciones de distribuciones acumulativas para distribuciones continuas y discretas.
- **Clases de Funciones e Interfaces**
- **Ajuste y Estimación de Parámetros**
- **Algoritmos Numéricos**

Además, la biblioteca “MathCore” contiene la siguiente clase de ROOT, que es “ROOT :: Math”[139], la cual proporciona la mayoría de las funciones especiales y las funciones estadísticas, como lo son las funciones de densidad de probabilidad, la cual se usa para el ajuste de datos.

Estas funciones de densidad de probabilidad, o simplemente densidad de una variable aleatoria continua describe la probabilidad relativa según la cual dicha variable aleatoria tomará determinado valor (Ver la sección D.1 del Apéndice D.).

D.2. El Método de Ajuste en ROOT

En ROOT para ajustar un histograma, se puede usar el panel de ajuste a través del menú contextual (Ver figura 7.2), o se puede usar el método “TH1::Fit” [144]. El panel de ajuste, es el mejor para la creación de prototipos, además el histograma debe dibujarse en un pad antes de invocar este método. El método “TH1::Fit” es más poderoso y se usa en scripts y programas [144].

El método de ajuste se implementa en ROOT para las clases de histogramas “TH1”, las clases de histogramas dispersos “THnSparse”, las clases de gráficos “TGraph”, “TGraph2D” y

“TMultiGraph” para ajustar una colección de gráficos con la misma función. A continuación se describe el método “TH1::Fit” y el Panel de Ajuste.

D.2.1. El Método “TH1::Fit”

Se presenta la implementación del método “TH1::Fit” [144] y una explicación de los parámetros:

```
1 | TFitResultPtr Fit(TF1 *function, Option_t *option, Option_t *
goption, Axis_t xxmin, Axis_t xxmax);
```

- *function: Es un puntero al objeto de la función ajustada (el modelo de ajuste). También se puede utilizar el nombre de la función. Este nombre puede ser uno de los nombres de función predefinidos de ROOT o una función definida por el usuario.
- *option: Es el segundo parámetro de la opción de ajuste. Aquí está la lista de opciones de ajuste:
 - “W”: Establece todos los pesos en 1 para contenedores no vacíos; ignora las barras de error.
 - “I”: Usa la integral de la función en el contenedor en lugar del valor en el centro del contenedor.
 - “L”: Utiliza el método de probabilidad logarítmica (el método predeterminado es chi-cuadrado, “ χ^2 ”). Para usar cuando el histograma representa recuentos.
 - “P”: Utiliza el método de chi-cuadrado de Pearson, además emplea los errores esperados en lugar del observado dado por “TH1::GetBinError” (caso predeterminado). En cambio, el error esperado se estima a partir de la raíz cuadrada del valor de la función bin.
 - “R”: Utiliza el rango especificado en el rango de funciones.
 - “B”: Se usa esta opción cuando se quiere fijar uno o más parámetros y la función de ajuste es una predefinida, como polN, expo, landau y gaus. Se tiene en cuenta que, en el caso de funciones predefinidas, se establecen algunos valores y límites iniciales predeterminados.
 - “Otras Opciones”
- *goption: El tercer parámetro es la opción de gráficos que es la misma que en el “TH1::Draw” [140].
- xxmin, xxmax: Los parámetros cuarto y quinto especifican el rango sobre el que aplicar el ajuste.

De forma predeterminada, el objeto de función ajustado se agrega al histograma y se dibuja en el panel actual.

Una clase de funciones mencionada en el método “TH1::Fit” es el objeto TF1 que es una función, el cual está definida entre un límite superior e inferior. La función puede ser una función simple basada en una expresión “TFormula” o una función de usuario precompilada. Además, la función puede tener parámetros asociados. A continuación, se hará una mejor descripción de dicha clase.

- **La Clase de Función “TF1”:** Aquí se muestra cómo crear la clase “TF1” que se usa para ajustar histogramas y gráficos. Se puede hacer ajustes con una función predefinida o con las funciones definidas por el usuario [140].

- **Ajustar con una Función Predefinida:** Para ajustar un histograma con una función predefinida, simplemente se pasa el nombre de la función al primer parámetro de “TH1::Fit”. Por ejemplo, en nuestro caso se ajusta al objeto de histograma “hist” con landau.

```
1 |     hist.Fit("landau");
```

La lista de funciones predefinidas que se pueden utilizar con el método Fit son las siguientes:

- “gaus” Función gaussiana con 3 parámetros.
- “expo” Un exponencial con 2 parámetros.
- “pol N” Un polinomio de grado N , donde N es un número entre 0 y 9.
- “chebyshev N” Un polinomio de Chebyshev de grado N, donde N es un número entre 0 y 9.
- “landau” Función de Landau con mean y sigma. Esta función se ha adaptado de la rutina CERNLIB G110 denlan(ver TMath::Landau) [140].
- “gausn” Forma normalizada de la función gaussiana con 3 parámetros.

- **Creación de Funciones Definidas por el Usuario:**

Se puede crear un objeto “TF1” y usarlo en la llamada “TH1::Fit”. El parámetro del método fit es el nombre del objeto “TF1”. Hay formas de crear un “TF1”.

- Usando una función que ha definido. Puede ser una función libre o un objeto funtor o una función miembro particular de una clase.
- Usando la expresión de C ++ usando x con un conjunto fijo de operadores y funciones definidas en TFormula.
- Otras formas.

D.2.2. El Panel de Ajuste

Esta interfaz está planificada para contener dos pestañas: “General” y “Minimización” [140]. La pestaña “General” proporciona elementos de interfaz de usuario para configurar la función de ajuste, el método de ajuste y diferentes opciones de ajuste, dibujo e impresión. La “pestaña Minimización” ofrece la opción de configurar el Minimizador para usar en el ajuste y sus opciones específicas (Ver figura 7.2).

D.3. Clase TMinuit:

MINUIT, ahora MINUIT2, es una biblioteca de software de minimización numérica escrita originalmente en el lenguaje de programación FORTRAN [175] por el físico del CERN Fred James en la década de 1970. Proporciona varios algoritmos de minimización que buscan valores de parámetros que minimizan una función definida por el usuario. Además de eso, puede calcular intervalos de confianza para los parámetros escaneando la función alrededor del mínimo.

Hay cinco algoritmos de minimización disponibles [176]. El algoritmo predeterminado recomendado MIGRAD es en general el mejor minimizador para casi todas las funciones. Es un método de métrica variable con búsqueda de línea inexacta, un esquema de actualización de métrica estable y controles de definición positiva. Su principal debilidad es que depende en gran medida del conocimiento de las primeras derivadas y falla miserablemente si son muy imprecisas [145].

Apéndice E

Abreviaturas

SM: Model Standard. Modelo Estándar.

GWS: Sheldon Glashow, Abdus Salam y Steven Weinberg.

QED: Quantum Electrodynamics. Electrodinámica Cuántica.

QCD: Quantum Cromodynamics. Cromodinámica Cuántica.

ISR: Initial State Radiation. Radiación de Estado Inicial.

FSR: Final State Radiation. Radiación de Estado Final.

UE: Underlying Event Evento Subyacente.

nTGCs: Neutral Triple Gauge Boson Couplings. Acoplamientos de Bosones de Calibre Triple Neutro.

LHC: Large Hadron Collider. Gran Colisionador de Hadrones.

gg: gluón-gluón.

$q\bar{q}$: Quark-antiquark.

ATLAS: A Toroidal LHC ApparatuS. Aparato Toroidal del LHC.

SFOS: Same Flavor Opposite Sign. Mismo sabor y signo opuesto.

PDF: Probability Distribution Function. Función de Distribución de Probabilidad.

NF: Nueva Física.

SUSY: Supersymmetry. Supersimetría.

CP: Charge Parity. Paridad de Carga.

LEP: Large Electron-Positron Collider. Colisionador de Electrones y Positrones.

CDF: Collider Detector at Fermilab. Detector Colisionador en Fermilab.

D0: DZero. Experimento D0.

CMS: Compact Muon Solenoid. Solenoide Compacto de Muones.

CERN: The European Organization for Nuclear Research. Organización Europea para la Investigación Nuclear.

LHCb : Large Hadron Collider Beauty. La belleza del Gran Colisionador de Hadrones.

ALICE: A Large Ion Collider Experiment. Un gran experimento de colisionador de iones.

LHCf: Large Hadron Collider forward. Gran Colisionador de Hadrones Adelante.

MoEDAL: Monopole and Exotics Detector At the Large Hadron Collider. Detector Monopolar y Exótico en el Gran Colisionador de Hadrones.

TÓTEM: TOTal Cross Section, Elastic Scattering and Diffraction Dissociation Measurement at LHC. Medida de la sección transversal TOTal, la dispersión elástica y la disociación por difracción en el LHC.

PS: ProtonSynchrotron. ProtónSincrotrón.

SPS: Super Proton Synchrotron. Supersincrotrón de Protones.

PSB: ProtonSynchrotron Booster. Impulsor de Sincrotrón de Protones.

ID: Internal Detector. Detector Interno.

SCT: Semiconductor Tracker. Rastreador de Semiconductores.

TRT: Transition Radiation Tracker. Rastreador de Radiación de Transición.

EM: Electromagnetic Calorimeter. Calorímetro Electromagnético.

HTC: Hadron calorimeters. Calorímetros Hadrónicos.

MS: Muon Spectrometer. Espectrómetro de muones.

MDT: Monitored Drift Tube. Tubo de deriva monitoreado.

RPC: Resistive Plate Chambers. Cámaras de Placas Resistivas.

RoI: Regions of Interest. Regiones de Interés.

EF: Event Filter. Filtro de Eventos.

HLT: High Level Trigger. Disparador de alto nivel.

MC: Monte-Carlo.

LINAC 2: Linear Accelerator 2. Acelerador lineal 2.

HEP: High Energy Physics. Física de Altas Energías.

LB: Bloque de Luminosidad. Luminosity Block.

GRL: Good Runs List. Lista de Buenas Carreras.

VP: Primary Vertices . Vértices Primarios.

GSF: Gaussian Sum Filter. Filtro de Suma Gaussiana.

BSM: Beyond the Standard Model. Más allá del modelo Estándar.

SaaS: Software as a Service. Software como Servicio.

Swan: Web Based Analysis Service. Servicio de Análisis basado en web.

P.D.F.s: Función de Densidad de Probabilidad. Probability Density Function.

MPV: Most Likely Value. Valor más Probable.

RMS: Root Mean Square. Media cuadrática.

LS: Least Squares. Mínimos Cuadrados.

ML: Maximum Likelihood. Máxima Verosimilitud.

Referencias

1. Graciela B. Gelmini. El Bosón de Higgs. Revista de CIENCIA E INVESTIGACIÓN - TOMO 64 N° 3. [Internet]. 2014 [Consultado 22 febrero 2021] Disponible en: <http://www.aargentinapciencias.org/2/images/RevistasCeI/tomo64-3/Paginas5-22desdeRevista64-3.pdf>
2. HyperPhysics traducido por M Olmo R Nave. Principio de Exclusión de Pauli. Universidad Estatal de Georgia [internet] 2000 [Consultado 20 Enero 2021] Disponible en: <http://hyperphysics.phy-astr.gsu.edu/hbasees/pauli.html>
3. Torres J. Tema 10: Estadística de Fermi-Dirac [internet] 2009 [Consultado 20 Febrero 2021] Disponible en: <http://www.ugr.es/~jtorres/tema10fe.pdf>
4. Rosenfeld, Léon. Nuclear Forces. Interscience Publishers, New York. pp. xvii. [internet] 1948 [Consultado 10 Febrero 2021] Disponible en: <https://archive.org/details/in.ernet.dli.2015.169077/page/n17/mode/2up>
5. Griffiths, David. Introduction to elementary particles. John Wiley and Sons Inc, [internet] 1987 [Consultado 12 enero 2021] Disponible en: <http://nuclphys.sinp.msu.ru/books/b/Griffiths.pdf>
6. Academic, Estadística de Bose-Einstein [internet] 2000-2021 [Consultado 8 enero 2021] Disponible en: <https://es-academic.com/dic.nsf/eswiki/456070>
7. Chen Ning Yang - Nobel Lecture: The Law of Parity Conservation and Other Symmetry Laws of Physics. [Internet] 1957 [Consultado 12 febrero 2021] Disponible en: <https://www.nobelprize.org/uploads/2018/06/yang-lecture.pdf>
8. Vargas, Miguel. Mecanismo de Higgs. [Internet] 2017 [Consultado 18 enero 2021] Disponible en: https://www.miguelangelvargascruz.com/mecanismodehiggs_blog_2222.html
9. F. J. Ynduráin. Elements of Group Theory [Internet] 2008 [Consultado 19 enero 2021] Disponible en: <https://arxiv.org/pdf/0710.0468.pdf>
10. Goldstein, Herbert. Classical Mechanics. Editorial Addison-Wesley; 2002
11. Wikipedia. Electrodinámica cuántica. [Internet] 2015 [Consultado 13 febrero 2021] Disponible en: http://i3campus.co/CONTENIDOS/wikipedia/content/a/electrodi%25c3%25a1mica_cu%25c3%25a1ntica.html
12. Halzen. Quarks and Leptons - An Introductory Course in Modern Particle Physics. Wiley; 1984.
13. A. Di Giacomo. Understanding Color Confinement. EPJ Web of Conferences 70,00019 [Internet] 2014 [Consultado 14 febrero 2021] Disponible en: https://www.epj-conferences.org/articles/epjconf/pdf/2014/07/epjconf_icfp2012_00019.pdf

14. HyperPhysics traducido por M Olmo R Nave. Electrodinámica Cuántica (QED). [Internet] 2000 [Consultado 15 febrero 2021] Disponible en:<http://131.96.55.77/hbasees/Forces/qed.html>
15. Wikipedia. Modelo electrodébil de Sheldon Lee Glashow. [Internet] 2006 [Consultado 16 febrero 2021] Disponible en: https://es.wikipedia.org/wiki/Modelo_electrod%C3%A9bil
16. Wikimedia Foundation. Interacción electrodébil de Abdus Salam. [Internet] 2021 [Consultado 26 febrero 2021] Disponible en: https://es.hrvwiki.net/wiki/Abdus_Salam
17. FreeJournal. Interacción electrodébil de Steven Weinberg [Internet] 2020 [Consultado 26 febrero 2021] Disponible en: <https://amp.es.freejournal.org/403921/1/steven-weinberg.html>
18. Nicholas, Charles E. Measurement of the ZZ production cross section and limits on anomalous neutral triple gauge couplings with the ATLAS detector. [Internet] 2013 [Consultado 21 febrero 2021] Disponible en: <https://cds.cern.ch/record/1563582/files/CERN-THESIS-2013-081.pdf?subformat=pdfaversion=1>
19. Wikipedia. Diagrama de Feynman. [Internet] 2008 [Consultado 18 febrero 2021] Disponible en: https://es.wikipedia.org/wiki/Electrodin%C3%A1mica_cu%C3%A1ntica/media/Archivo:Feynmandiagram.svg
20. Higgs, Peter W., Broken Symmetries and the Masses of Gauge Bosons. Physical Review Letters 13 (1964).
21. Englert, F., Brout, R., Broken Symmetry and the Mass of Gauge Vector Mesons. Physical Review Letters 13; 1964.
22. Hagen, C. R. G. S. Guralnik, T. W. B. Kibble, Global Conservation Laws and Massless Particles. Physical Review Letters 13; 1964.
23. Guralnik, Gerald S. The History of the Guralnik, Hagen and Kibble development of the Theory of Spontaneous Symmetry Breaking and Gauge Particles. International Journal of Modern Physics A24; 2009.
24. Martinez, Homero. Lecture IV: Introduction to the standard model; Abril 2016.
25. Campbell, J. M., Huston, J., and Stirling, W., Hard Interactions of Quarks and Gluons: A Primer for LHC Physics, Rept.Prog.Phys. 70. arXiv:hep-ph/061114889; 2007.
26. European Organisation For Nuclear Research (CERN). Measurement of the ZZ Production Cross Section in pp Collisions at $s = 13$ TeV with the ATLAS Detector. [Internet] 2021 [Consultado 18 julio 2021] Disponible en: <https://arxiv.org/pdf/1512.05314.pdf>
27. G. J. Gounaris, J. Layssac, and F. M. Renard, Off-shell structure of the anomalous Z and self-couplings, Phys. Rev. D62 (2000) 073012, arXiv:hep-ph/0005269
28. ATLAS Collaboration, Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC , Phys. Lett. B 716 (2012) 1, arXiv:1207.7214 [hep-ex].
29. G. Aad et al. Measurement of the ZZ Production Cross Section in pp Collisions at $s=13$ TeV with the ATLAS Detector. American Physical Society under the terms of the Creative Commons Attribution 3.0 License [Internet]. 2016 [Consultado 18 febrero 2021] Disponible en: <https://journals.aps.org/prl/pdf/10.1103/PhysRevLett.116.101801>

30. EUROPEAN ORGANISATION FOR NUCLEAR RESEARCH (CERN). Measurement of ZZ production in pp collisions at $\sqrt{s} = 7\text{TeV}$ and limits on anomalous ZZZ and ZZ γ couplings with the ATLAS detector arXiv:1211.6096v1 [Internet] 26 Nov 2012. [Consultado 25 febrero 2021] Disponible en: <https://arxiv.org/pdf/1211.6096.pdf>
31. Campbell, J. M., Ellis R. K., and Williams, C. Vector boson pair production at the LHC , JHEP 1107 018, arXiv:1105.0020 [Internet] 2011 [Consultado 15 febrero 2021] Disponible en: <https://arxiv.org/abs/1105.0020>
32. Beringer, J et al. (Particle Data Group), Review of Particle Physics, Phys. Rev. D 86 010001 [Internet] 2012 [Consultado 17 febrero 2021] Disponible en: <http://link.aps.org/doi/10.1103/PhysRevD.86.010001>
33. ATLAS PUB Note PUB-OTRC-2020-01. Review of the 13 TeV ATLAS Open Data release. [Internet] 28 Abril 2020. [Consultado 19 febrero 2021] Disponible en: <https://cds.cern.ch/record/2707171/files/ANA-OTRC-2019-01-PUB-updated.pdf>
34. G. Gounaris, J. Layssac, and F. Renard, Signatures of the anomalous Z and ZZ production at the lepton and hadron colliders, Phys.Rev. D61; 2000. 073013, arXiv:hep-ph/9910395.
35. Baur, U. and Rainwater, D. L., Probing neutral gauge boson self interactions in ZZ production at hadron colliders, Phys.Rev. D62 113011; 2000. arXiv:hep-ph/0008063
36. Gounaris, G., Layssac, J., and Renard, F., New and standard physics contributions to anomalous Z and gamma self couplings, Phys.Rev. D62; 2000. 073013, arXiv:hep-ph/0003143
37. The L3 Collaboration, Study of neutral-current four-fermion and ZZ production in e+e collisions at s= 183 GeV , Physics Letters B 450; 1999, no. 13, 281 – 293.
38. The OPAL Collaboration, Study of Z pair production and anomalous couplings in e + e collisions at s between 190 GeV and 209 GeV , Eur. Phys. J. C32; 2003. 303, arXiv:hep-ex/0310013.
39. The DELPHI Collaboration, ZZ production in e^+e^- interactions at $\sqrt{s} = 183\text{GeV}$ to 209GeV , Eur. Phys. J. C30; 2003. 447, arXiv:hep-ex/0307050.
40. The L3 Collaboration, Study of Z Boson Pair Production in e^+e^- collisions at LEP at $\sqrt{s} = 189\text{GeV}$, Phys. Lett. B465; 1999. 363, arXiv:hep-ex/9909043.
41. The ALEPH Collaboration, Measurement of Z-pair production in e + e collisions and constraints on anomalous neutral gauge couplings, Journal of High Energy Physics; no. 04, 124. [Internet] 2009 [Consultado 20 febrero 2021] Disponible en: <http://stacks.iop.org/1126-6708/2009/i=04/a=124>.
42. The LEP Collaborations ALEPH, DELPHI, L3, OPAL, and the LEP Electroweak Working Group, A Combination of preliminary electroweak measurements and constraints on the Standard Model. [Internet] 2006 [Consultado 20 febrero 2021] Disponible en: arXiv:hep-ex/0612034.
43. The DELPHI Collaboration, Study of Trilinear Gauge Boson Couplings, ZZZ, ZZv and Zvv., DELPHI-2001-097-CONF-525. [Internet] 2001 [Consultado 21 febrero 2021] Disponible en: <http://cds.cern.ch/record/1002930>.
44. The OPAL Collaboration, Z boson pair production in e^+e^- collisions at p (s) = 183-GeV and 189-GeV , Phys.Lett. B476 (2000) 256–272, arXiv:hep-ex/0003016.

45. D0 Collaboration, Measurement of the ZZ production cross section in $p\bar{p}$ collisions at $\sqrt{s} = 1,96TeV$, Phys.Rev. D84011103. [Internet] 2011 [Consultado 22 febrero 2021] Disponible en: arXiv:1104.3078
46. D0 Collaboration, A measurement of the WZ and ZZ production cross sections using leptonic final states in $8,6fb^{-1}$ of $p\bar{p}$ collisions, Phys.Rev. D85 112005. [Internet] 2012 [Consultado 27 febrero 2021] Disponible en: arXiv:1201.5652
47. CDF Collaboration, Measurement of ZZ production in leptonic final states at \sqrt{s} of $1,96TeV$ at CDF, Phys.Rev.Lett. 108 101801, [Internet] 2012 [Consultado 27 febrero 2021] Disponible en: arXiv:1112.2978.
48. D0 Collaboration, Search for ZZ and $Z\gamma^*$ production in $p\bar{p}$ collisions at $\sqrt{s} = 1,96TeV$ and limits on anomalous ZZZ and $ZZ\gamma^*$ couplings, Phys.Rev.Lett. 100 131801, [Internet] 2008 [Consultado 28 febrero 2021] Disponible en: arXiv:0712.0599
49. CMS Collaboration, Measurement of the ZZ production cross section and search for anomalous couplings in $2l2l$ final states in pp collisions at $\sqrt{s} = 7TeV$, JHEP 1301 063, [Internet] 2013 [Consultado 28 febrero 2021] Disponible en: arXiv:1211.4890.
50. CMS Collaboration, Measurement of W+W and ZZ production cross sections in pp collisions at $\sqrt{s} = 8TeV$, [Internet] 2013 [Consultado 28 febrero 2021] Disponible en: arXiv:1301.4698.
51. ihow.pro. ¿Cómo funcionan los contadores Geiger? [Internet] 2021 [Consultado 28 febrero 2021] Disponible en: <https://ihow.pro/es/p/como-funcionan-los-contadores-geiger/ry1MXtEMvpLjFuo0g2hKBg>
52. Consejo Europeo de Investigación Nuclear. CERN. [Internet] 2010 [Consultado 01 marzo 2021] Disponible en: <https://home.cern/>
53. Centro Nacional de Física de Partículas Astropartículas y Nuclear (CPAN) El Gran Colisionador de Hadrones (LHC). [Internet] 2021 [Consultado 28 febrero 2021] Disponible en: <https://www.i-cpan.es/es/content/el-gran-colisionador-de-hadrones-lhc>
54. Consejo Europeo de Investigación Nuclear. CERN. The Large Hadron Collider. [Internet] 2021 [Consultado 28 febrero 2021] Disponible en: <https://home.cern/science/accelerators/large-hadron-collider>
55. Accelerator Performance and Statistics. [Internet] 2014 - 2021 [Consultado 28 marzo 2021] Disponible en: <https://acc-stats.web.cern.ch/acc-stats/lhc/overview-panel>
56. Román, Julia E. P. Reconstrucción de Fotones con el Detector ATLAS en Preparación Para la Segunda Corrida del LHC. [Trabajo de Grado presentado como requisito parcial para optar al título de Licenciado en Física] Universidad de Los Andes; 2019.
57. Taking a closer look at LHC. Magnets Detectors I Taking a closer look at LHC. [Internet] 2013 - 2015 [Consultado 28 abril 2021] Disponible en: https://www.lhc-closer.es/taking_a_closer_look_at_lhc/0.magnets__detectors_i
58. Di Domenico F. María R. Reconstrucción de Masas Invariantes de Bosones del Modelo Estándar Usando Datos Públicos de Atlas Open Data. [Trabajo de Grado presentado como requisito parcial para optar al título de Licenciado en Física] Universidad Central de Venezuela; 2017.
59. Atlas Experiment – Publics Results. Pileup Interactions and Data Taking Efficiency. [Internet] 2016 [Consultado 12 marzo 2021] Disponible en:

- https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2#2016_pp_Collisions
60. Brning, O. S., Collier, P., Lebrun, P., et al., LHC Design Report, CERN-2004-003, [Internet] 2004 [Consultado 13 marzo 2021] Disponible en: <https://cdsweb.cern.ch/record/782076>.
61. Accelerator Performance and Statistics. [Internet] 2014 - 2021 [Consultado 28 marzo 2021] Disponible en: <https://lhcb-statistics.web.cern.ch/LHC-Statistics/>.
62. Fournier, D. Performance of the LHC, ATLAS and CMS in 2011 , EPJ Web Conf. 28 01003. [Internet] 2012 [Consultado 14 marzo 2021] Disponible en: arXiv:1201.4681.
63. ChronoZoom!. The SPS. [Internet] 2014 [Consultado 15 marzo 2021] Disponible en: <http://www.chronozoom.com/cern/f7ece365-70d2-4474-ad55-66839815a770/t29b28831-5999-46ae-a724-636d23eb1e99/t737a8002-403b-48f9-9410-ca8d784ec3c0/ee11219bb-d4f0-442a-b552-15faf9a5c863/f7ece365-70d2-4474-ad55-66839815a770@x=0y=0w=2.53375796178344h=1.050955414012739>
64. Wikipedia. LHCf experiment. [Internet] 2021 [Consultado 28 julio 2021] Disponible en: https://en.wikipedia.org/wiki/LHCf_experiment
65. ATLAS Experiment. The ATLAS Detector. [Internet] 2021 [Consultado 26 julio 2021] Disponible en: <https://atlas.cern/discover/detector>
66. ATLAS Experiment. Detectores de Partículas: El experimento ATLAS [Internet] 2021 [Consultado 27 julio 2021] Disponible en: <https://atlas.cern>
67. Sánchez P., Arturo. Search for Higgs-like resonances in the $ZZ \rightarrow l^\pm l^\pm q\bar{q}$ decay channel with the ATLAS experiment at LHC. [Tesis para optar al título Doctor en Filosofía] Universidad de Nápoles "Federico II", Departamento de Ciencias Físicas. Doctorado en Física Fundamental y Aplicada; 2013.
68. ATLAS Collaboration, CERN/LHCC 97-21; 1997.
69. ATLAS Collaboration, CERN/LHCC 97-19; 1997.
70. ATLAS Collaboration, CERN/LHCC 97-20; 1997.
71. ATLAS Experiment. It is the first part of ATLAS to see the decay products of the collisions. [Internet] 2021 [Consultado 21 junio 2021] Disponible en: <https://atlas.cern/discover/detector/inner-detector>
72. Studies Concerning the ATLAS IBL Calibration Architecture - Scientific Figure on ResearchGate. [Internet] 2021 [Consultado 8 marzo, 2021] Disponible en: https://www.researchgate.net/figure/The-inner-detector-of-ATLAS-with-the-pixel-detector-the-silicon-tracker-SCT-and-the_fig3_254469228
73. ATLAS Collaboration, The ATLAS Experiment at the CERN Large Hadron Collider Journal of Instrumentation 3 no. 08, S08003. [Internet] 2008 [Consultado 20 julio 2021] Disponible en: <http://stacks.iop.org/1748-0221/3/i=08/a=S08003>.
74. Gastón, Leonardo R. Reconstrucción de muones cósmicos en el espectrómetro de muones del experimento ATLAS. Tesis de Licenciatura en Ciencias Físicas Facultad de Ciencias Exactas y Naturales Universidad de Buenos Aires [Internet] 2007 [Consultado 23 mayo de 2021] Disponible en: <http://users.df.uba.ar/aia/Varios/Tesis.Seminario.Gaston.pdf>

75. J. Dubbert , M. Groh, O. Kortner , H. Kroha , J. von Loeben , R. Richter , J. Schmaler , H. von der Schmitt , O. Biebel , D. Merkl , F. Rauscher , A. Staude. Registro de la conferencia del Simposio de Ciencias Nucleares del IEEE de 2006. Integración, instalación y puesta en servicio de grandes cámaras de tubos de deriva supervisadas del espectrómetro de muones de barril ATLAS. [Internet] 2006 [Consultado 25 mayo de 2021] Disponible en: <https://www.semanticscholar.org/paper/Integration%2C-Installation%2C-and-Commissioning-of-of-Dubbert-Groh/35051ccc297e9c171a327db39b1415f58a737b1f/figure/2>
76. P. Nason, A new method for combining NLO QCD with shower Monte Carlo algorithms, *J. High Energy Phys.* 11 040; 2004.
77. S. Frixione, P. Nason, and C. Oleari, Matching NLO QCD computations with Parton Shower simulations: The POWHEG method, *J. High Energy Phys.* 11 070; 2007
78. S. Alioli, P. Nason, C. Oleari, and E. Re, A general framework for implementing NLO calculations in shower Monte Carlo programs: The POWHEG BOX, *J. High Energy Phys.* 06 043; 2010.
79. T. Melia, P. Nason, R. Röntsch, and G. Zanderighi, WW, WZ and ZZ production in the POWHEG-BOX, *J. High Energy Phys.* 11 078; 2011.
80. P. Nason and G. Zanderighi, WW, WZ and ZZ production in the POWHEG-BOX-V2, *Eur. Phys. J. C* 74, 2702; 2014.
81. T. Sjöstrand, S. Mrenna, and P. Skands, PYTHIA 6.4 physics and manual, *J. High Energy Phys.* 05 026; 2006.
82. T. Sjöstrand, S. Mrenna, and P. Skands, A brief introduction to PYTHIA 8.1, *Comput. Phys. Commun.* 178, 852 ;2008.
83. ATLAS Collaboration, Measurement of the Z= boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector, *J. High Energy Phys.* 09 145; 2014
84. T. Gleisberg, S. Höche, F. Krauss, M. Schönher, S. Schumann, F. Siegert, and J. Winter, Event generation with SHERPA 1.1, *J. High Energy Phys.* 02 007; 2009
85. S. Höche, F. Krauss, S. Schumann, and F. Siegert, QCD matrix elements and truncated showers, *J. High Energy Phys.* 05 053; 2009.
86. T. Gleisberg and S. Höche, Comix, a new matrix element generator, *J. High Energy Phys.* 12 039; 2008.
87. H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin, and C.-P. Yuan, New parton distributions for collider physics, *Phys. Rev. D* 82, 074024; 2010.
88. J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, MADGRAPH5: Going beyond, *J. High Energy Phys.* 06 128; 2011.
89. R. D. Ball et al., Parton distributions with LHC data, *Nucl. Phys.* B867, 244; 2013.
90. ATLAS Collaboration, Report No. ATL-PHYS-PUB-2014- 021. [Internet] 2014 [Consultado 9 mayo de 2021] Disponible en: <http://cdsweb.cern.ch/record/1966419>
91. A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, Parton distributions for the LHC, *Eur. Phys. J. C* 63, 189; 2009.
92. ATLAS Collaboration, Report No. ATL-PHYS-PUB-2012- 003. [Internet] 2012 [Consultado 10 mayo de 2021] Disponible en: <http://cds.cern.ch/record/1474107>.

93. S. Agostinelli, J. Allison, K. Amako, et al., Geant4 - a simulation toolkit, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 506 no. 3, 250 – 303; 2003
94. Wikipedia. Línea de base (investigación científica). [Internet] 2017 [Consultado 19 mayo de 2021] Disponible en: [https://es.wikipedia.org/wiki/Linea_de_base_\(investigacioncientifica\)](https://es.wikipedia.org/wiki/Linea_de_base_(investigacioncientifica))
95. ATLAS experiment. Software book- data and tools - documentation. [Internet] 2016 [Consultado 20 mayo de 2021] Disponible en: <https://cheatham1.gitbooks.io/openatlasdatatools/content/>
96. T. Cornelissen, M. Elsing, I. Gavrilenko, et al., The new ATLAS track reconstruction (NEWT), Journal of Physics: Conference Series 119 no. 3, 032014. [Internet] 2008 [Consultado 9 mayo de 2021] Disponible en: <http://stacks.iop.org/1742-6596/119/i=3/a=032014>
97. Camacho, Reina. Tracking and vertexing, Mayo 2016.
98. R. e. Kalman, A new approach to linear filtering and prediction problems. Journal of Basic Engineering 82 (1): 35-45. Retrieved 2008-05-03; 1960
99. ATLAS Collaboration. Performance of the ATLAS Inner Detector Track and Vertex Reconstruction in the High Pile-Up LHC Environment. Technical Report ATLAS-CONF-2012-042, CERN, Geneva; 2012.
100. V Lacuesta. Track and vertex reconstruction in the atlas experiment. Journal of Instrumentation, 8(02):C02035; 2013.
101. G. Aad et al. Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data, Eur.Phys.J., C72:1909; 2012
102. M. Aharrouche et al. Measurement of the response of the ATLAS liquid argon barrel calorimeter to electrons at the 2004 combined test-beam. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 614(3):400-432; 2010.
103. E. Abat, J. M. Abdallah, T. N. Addy, P. Adragna, M. Aharrouche, A. Ahmad, T. P. A. Akesson, M. Aleksa, C. Alexa, K. Anderson, et al. Combined performance studies for electrons at the 2004 ATLAS combined test-beam. Journal of Instrumentation, 5:11006; November 2010.
104. ATLAS Collaboration, Expected electron performance in the ATLAS experiment, ATL-PHYS-PUB-2011-006, [Internet] Apr, 2011. [Consultado 19 mayo de 2021] Disponible en: <https://cdsweb.cern.ch/record/1345327>.
105. ATLAS Collaboration, Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $s = 13$ TeV, Eur. Phys. J. C 79 639, (cit. on pp. 4, 5). [Internet] 2019 [Consultado 13 mayo de 2021] Disponible en: arXiv: 1902.04655
106. ATLAS Collaboration, Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $s = 13$ TeV, Eur. Phys. J. C 76 292, (cit. on pp. 4, 5, 12, 16, 35). [Internet] 2016 [Consultado 9 mayo de 2021] Disponible en: arXiv: 1603.05598
107. R. Nicolaïdou, L. Chevalier, S. Hassani, et al., Muon identification procedure for the ATLAS detector at the LHC using Muonboy reconstruction package and tests of its

- performance using cosmic rays and single beam data, Journal of Physics: Conference Series 219 no. 3, 032052. [Internet] 2010 [Consultado 14 mayo de 2021] Disponible en: <http://stacks.iop.org/1742-6596/219/i=3/a=032052>;
108. ATLAS Collaboration. Muon reconstruction efficiency in reprocessed 2010 LHC proton-proton collision data recorded with the ATLAS detector, ATLAS-CONF-2011-063, [Internet] Apr, 2011. [Consultado 15 mayo de 2021] Disponible en: <http://cdsweb.cern.ch/record/1345743>.
 109. Arias M. ZZDiBosonAnalysis.C at master mildredarias/atlas-outreach-cpp-framework-13tev. [Internet]. 2019. [Revisado Septiembre de 2021] Disponible en: <https://github.com/mildredarias/atlas-outreach-cpp-framework-13tev/blob/master/Analysis/ZZDiBosonAnalysis/ZZDiBosonAnalysis.C>
 110. ATLAS Open Data 13 TeV Documentation. The ATLAS Open Data 13 TeV docs. [Internet]. S.f. [Recuperado el 25 de mayo de 2021] Disponible en: <http://opendata.atlas.cern/release/2020/documentation/index.html>
 111. ATLAS Open Data 13 TeV Documentation. Datasets in ROOT format. [Internet]. 2019. [febrero a marzo de 2021] Disponible en: <http://opendata.atlas.cern/release/2020/documentation/datasets/files.html>
 112. Cheatham. ATLAS open data · Software Book [Internet]. Gitbooks.io. [Citado 2021 Mar 5]. Disponible en: <https://cheatham1.gitbooks.io/openatlasdatatools/content/>
 113. Rodríguez, F., Pérez, A.J. Bartomeus, I., Varela, S. Ciencia reproducible: qué, por qué, cómo. [Internet]. 2016. Ecosistemas. Revista Científica de Ecología y Medio Ambiente. ISSN 1697-2473 / Open access. Disponible en: www.revistaecosistemas.net
 114. ROOT team. ROOT: analyzing petabytes of data, scientifically [Internet]. ROOT. 2019 [Citado 5 Febrero 2021]. Disponible en: <https://root.cern>
 115. Project jupyter [Internet]. 2016. Jupyter.org. [Citado 4 Marzo 2021]. Disponible en: <https://jupyter.org/>
 116. Arias M. ATLAS-Open-Data-cpp-13-TeV-framework-ZZDiBosonAnalysis.ipynb at master · mildredarias/ZZDiBoson. [Internet]. 2016. Jupyter.org. [Actualizado 4 Septiembre 2021]. Disponible en: <https://github.com/mildredarias/ZZDiBoson/blob/master/Analysis/ATLAS-Open-Data-cpp-13-TeV-framework-ZZDiBosonAnalysis.ipynb>
 117. The swan service [Internet]. 2016. Cern.ch. [Revisado 25 Abril de 2021]. Disponible en: <https://swan.web.cern.ch/swan/>
 118. Where the world builds software. [Internet]. 2016. Cern.ch. [Revisado 25 Mayo de 2021]. Disponible en: <https://github.com/>
 119. Git [Internet]. Git-scm.com. [Citado Mayo de 2021]. Disponible en: <https://git-scm.com>
 120. Arias M. Mildredarias – Descripción General. [Internet]. 2021. Cern.ch. [Actualizado 25 Septiembre de 2021]. Disponible en: <https://github.com/mildredarias>
 121. Overview - ATLAS open data 13 TeV documentation [Internet]. Atlas.cern. [Citado febrero 2021]. Disponible en: <http://opendata.atlas.cern/release/2020/documentation/datasets/intro.html>
 122. Arias M. ZZDiBoson. [Internet]. 2021. Cern.ch. [Actualizado 20 Septiembre de 2021]. Disponible en: <https://github.com/mildredarias/ZZDiBoson>

123. Arias M. ATLAS-Open-Data-cpp-13-TeV-framework-ZZDiBosonAnalysis.ipynb at master . mildredarias/ZZDiBoson. [Internet]. 2021. Cern.ch. [Actualizado 20 Septiembre de 2021]. Disponible en: <https://github.com/mildredarias/ZZDiBoson/blob/master/Analysis/ATLAS-Open-Data-cpp-13-TeV-framework-ZZDiBosonAnalysis.ipynb>
124. C Library - [Internet]. Tutorialspoint.com. [Citado Abril de 2021]. Disponible en: https://www.tutorialspoint.com/c_standard_library/stdio_h.htm
125. C Library - [Internet]. Tutorialspoint.com. [Citado Abril 22 de 2021]. Disponible en: https://www.tutorialspoint.com/c_standard_library/stdlib_h.htm
126. iostream;[Internet]. Cplusplus.com. [Revisado marzo de 2021]. Disponible en: <https://wwwcplusplus.com/reference/iostream/>
127. Arias M. Analysis/Output_ZZDiBosonAnalysis at master . mildredarias/ZZDiBoson. [Internet]. 2019. [Revisado Agosto de 2021]. Disponible en: https://github.com/mildredarias/ZZDiBoson/tree/master/Analysis/Output_ZZDiBosonAnalysis
128. Arias M. Plotting/histograms at master . mildredarias/ZZDiBoson. [Internet]. 2019. [Revisado Septiembre de 2021]. Disponible en: <https://github.com/mildredarias/ZZDiBoson/tree/master/Plotting/histograms>
129. Arias M. welcome.sh at master . mildredarias/ZZDiBoson. [Internet]. 2019. [Revisado Septiembre de 2021]. Disponible en: <https://github.com/mildredarias/ZZDiBoson/blob/master/welcome.sh>
130. Arias M. ZZDiBosonAnalysis.h at master . mildredarias/ZZDiBoson. [Internet]. 2019. [Actualizado 11 Septiembre de 2021]. Disponible en: <https://github.com/mildredarias/ZZDiBoson/blob/master/Analysis /ZZDiBosonAnalysis.h>
131. Arias M. ZZDiBosonAnalysisHistograms.h at master . mildredarias/ZZDiBoson. [Internet]. 2019. [Actualizado 11 Septiembre de 2021]. Disponible en: <https://github.com/mildredarias/ZZDiBoson/blob/master/Analysis /ZZDiBosonAnalysisHistograms.h>
132. Arias M. ZZDiBosonAnalysis.C at master . mildredarias/ZZDiBoson.[Internet]. 2019. [Actualizado 11 Septiembre de 2021]. Disponible en: <https://github.com/mildredarias/ZZDiBoson/blob/master/Analysis /ZZDiBosonAnalysis.C>
133. Arias M. main_ZZDiBosonAnalysis.C at master . mildredarias/ZZDiBoson. [Internet]. 2019. [Actualizado 11 Septiembre de 2021]. Disponible en: https://github.com/mildredarias/ZZDiBoson/blob/master/Analysis/main _ZZDiBosonAnalysis.C
134. Arias M. run.sh at master . mildredarias/ZZDiBoson. [Internet]. 2019. [Actualizado 11 de Septiembre de 2021]. Disponible en: <https://github.com/mildredarias/ZZDiBoson/blob/master/Analysis/run.sh>
135. Arias M. Plotting.h at master . mildredarias/ZZDiBoson. [Internet]. 2019. [Actualizado 11 Septiembre de 2021]. Disponible en: <https://github.com/mildredarias/ZZDiBoson/blob/master/Plotting/Plotting.h>

136. Arias M. Plotting.cxx at master · mildredarias/ZZDiBoson. [Internet]. 2019. [Revisado del 10 al 11 de Septiembre de 2021]. Disponible en: <https://github.com/mildredarias/ZZDiBoson/blob/master/Plotting/Plotting.cxx>
137. Arias M. plotme.sh at master · mildredarias/ZZDiBoson. [Internet]. 2019. [Revisado del 10 al 11 de Septiembre de 2021]. Disponible en: <https://github.com/mildredarias/ZZDiBoson/blob/master/Plotting/plotme.sh>
138. James F. (CERN) Revisor en abril de 1998; Cousins R. (UCLA) revisor febrero de 2000; Cowan G. (RHUL) revisor en octubre 2001 y octubre de 2003. 32. STATISTICS [Internet]. Lbl.gov. 2019. [citado el 5 de marzo de 2021]. Disponible en: <https://pdg.lbl.gov/2004/reviews/statrpp.pdf>
139. ROOT: Probability density functions (PDF) [Internet]. Root.cern. 2019. [citado el 6 de marzo de 2021]. Disponible en: https://root.cern/doc/master/group_PdfFunc.html
140. Guide U. ROOTUsersGuide [Internet]. Cern.ch. 2018. [citado el 8 de junio de 2021]. Disponible en: <https://root.cern.ch/root/htmldoc/guides/users-guide/ROOTUsersGuide.html>
141. LandauDistribution—wolfram language documentation [Internet]. Wolfram.com. 2017. [citado el 4 de junio de 2021]. Disponible en: <https://reference.wolfram.com/language/ref/LandauDistribution.html>
142. Random Number Distributions — GSL 2.7 documentation [Internet]. Gnu.org. 2018. [citado el 5 de junio de 2021]. Disponible en: <https://www.gnu.org/software/gsl/doc/html/randist.html>
143. ROOT: Probability density functions (PDF) [Internet]. Root.cern. [citado el 6 de junio de 2021]. Disponible en: https://root.cern/doc/master/group_PdfFunc.html
144. Guide U. ROOTUsersGuide. The Fit Method. [Internet]. Cern.ch. 2018. [citado el 8 de junio de 2021]. Disponible en: <https://root.cern.ch/root/htmldoc/guides/users-guide/ROOTUsersGuide.htmlthe-fit-method>
145. ROOT: TMinuit class reference [Internet]. Cern.ch. 2018. [citado el 9 de junio de 2021]. Disponible en: <https://root.cern.ch/doc/master/classTMinuit.html>
146. Nicolas. Prueba de Bondad de Ajuste Chi Cuadrado [Internet]. Probabilidades y Estadística. 2017. [citado el 5 de julio de 2021]. Disponible en: https://webcache.googleusercontent.com/search?q=cache:Z2x_jfe5KPMJ:https://www.probabilidadesyestadistica.com/prueba-de-bondad-de-ajuste-chi-cuadrado/+cd=12hl=esct=clnkgl=cogoogle_abuse=GOOGLE_ABUSE_EXEMPTION%3DID%3Dded4cb62ee3aea32:TM%3D1646432067:C%3Dr:IP%3D190.90.39.154:S%3D3GjKrJmWNbM6O33Rkyvpoc%3B+path%3D/%3B+domain%3Dgoogle.com%3B+expires%3DSat,+05-Mar-2022+01:14:27+GMT
147. ROOT: TH1 class reference [Internet]. Cern.ch. 2017. [citado el 15 de junio de 2021]. Disponible en: <https://root.cern.ch/doc/master/classTH1.html>
148. Arias M. Chi2_minuit.ipynb at master · mildredarias/ZZDiBoson. [Internet]. github.com. 2017. [Actualizado el 20 de enero de 2022]. Disponible en: https://github.com/mildredarias/ZZDiBoson/blob/master/Chi2_minuit.ipynb
149. Ruiz, R. Estudio del Comportamiento a Fatiga por Flexión de La Planta de Calzado de Seguridad para Determinar su Tiempo de Vida Útil, en La Fábrica Carvicaucho de

- La Ciudad de Ambato. Universidad técnica de Ambato Facultad de Ingeniería Civil y Mecánica Carrera de Ingeniería Mecánica. [Internet]. Docplayer.es. 2013. [citado el 15 de julio de 2021]. Disponible en: <https://docplayer.es/98747433-Universidad-tecnica-de-ambato-facultad-de-ingeneria-civil-y-mecanica-carrera-de-ingeneria-mecanica.html>
150. Arias M. Least_square-likelihood.ipynb at master · mildredarias/ZZDiBoson [Internet]. 2021 [Actualizado el 19 de Enero de 2022]. Disponible en: <https://github.com/mildredarias/ZZDiBoson/commits/master>
151. Karimaki, V. Selected Topics on Data Analysis in Particle Physics Lecture Notes. [Internet]. Helsinki.fi. 1992. [citado el 15 de julio de 2021]. Disponible en: <https://www.mv.helsinki.fi/home/karimaki/CMHEP/Lectures/statLectures.pdf>
152. Garren L. (Fermilab), Knowles I.G. (U. de Edimburgo), Navas S. (U. de Granada), Richardson P. (U. de Durham), Sjostrand T. (Lund U.) y Trippe T. (LBNL), Revisores. 34. Monte Carlo particle numbering scheme [Internet]. Lbl.gov. 2006. [citado el 13 de julio de 2021]. Disponible en: <https://pdg.lbl.gov/2007/reviews/montecarlorpp.pdf>
153. Arias M. hist_fourleptID.png at master · mildredarias/ZZDiBoson [Internet]. 2021 [Actualizado el 9 de enero de 2022]. Disponible en: https://github.com/mildredarias/ZZDiBoson/blob/master/Plotting/histograms/hist_fourleptID.png
154. Arias M. hist_fourleptch.png at master · mildredarias/ZZDiBoson [Internet]. 2021 [Actualizado el 11 de septiembre de 2021]. Disponible en: https://github.com/mildredarias/ZZDiBoson/blob/master/Plotting/histograms/hist_fourleptch.png
155. Arias M. hist_fourleptphi.png at master · mildredarias/ZZDiBoson [Internet]. 2021 [actualizado el 11 de septiembre de 2021]. Disponible en: https://github.com/mildredarias/ZZDiBoson/blob/master/Plotting/histograms/hist_fourleptphi.png
156. Arias M. hist_fourlepteta.png at master · mildredarias/ZZDiBoson [Internet]. 2021 [actualizado el 11 de septiembre de 2021]. Disponible en: https://github.com/mildredarias/ZZDiBoson/blob/master/Plotting/histograms/hist_fourlepteta.png
157. Arias M. hist_fourleptpt.png at master · mildredarias/ZZDiBoson [Internet]. 2021 [actualizado el 11 de septiembre de 2021]. Disponible en: https://github.com/mildredarias/ZZDiBoson/blob/master/Plotting/histograms/hist_fourleptpt.png
158. Arias M. hist_fourleptE.png at master · mildredarias/ZZDiBoson [Internet]. 2021 [actualizado el 11 de septiembre de 2021]. Disponible en: https://github.com/mildredarias/ZZDiBoson/blob/master/Plotting/histograms/hist_fourleptE.png
159. Arias M. hist_mLL1.png at master · mildredarias/ZZDiBoson [Internet]. 2021 [citado el 9 de noviembre de 2021]. Disponible en: https://github.com/mildredarias/ZZDiBoson/blob/master/Plotting/histograms/hist_mLL1.png
160. Arias M. hist_mLL2.png at master · mildredarias/ZZDiBoson [Internet]. 2021 [revisado el 11 de septiembre de 2021]. Disponible en:

- https://github.com/mildredarias/ZZDiBoson/blob/master/Plotting/histograms/hist_mLL2.png
161. Arias M. hist_fourlepsys_pt.png at master · mildredarias/ZZDiBoson [Internet]. 2021 [revisado el 11 de septiembre de 2021]. Disponible en: https://github.com/mildredarias/ZZDiBoson/blob/master/Plotting/histograms/hist_fourlepsys_pt.png
 162. Arias M. hist_fourlepsys_m.png at master · mildredarias/ZZDiBoson [Internet]. 2021 [citado el 11 de Septiembre de 2021]. Disponible en: https://github.com/mildredarias/ZZDiBoson/blob/master/Plotting/histograms/hist_fourlepsys_m.png
 163. Wikipedia contributors. Group [Internet]. Wikipedia, The Free Encyclopedia. 2017. Disponible en: <https://en.wikipedia.org/w/index.php?title=Groupoldid=1074523208>
 164. Quintero N, Molina F. Una descripción sencilla de las Teorías Gauge. Tumbaga [Internet]. 2009 [citado el 5 de noviembre de 2021]; 1(4):19–29. Disponible en: <https://dialnet.unirioja.es/servlet/articulo?codigo=3632005>
 165. Howard Georgi. Lie algebras in particle physics. Frontiers in Physics. AddisonWesley Publishing Company, 2 edition, 1982.
 166. Abdelgabar M. I., Yakubu, A., Abdallah M. D. The Lagrangian Formulation and Gauge Theory of the Standard Model [Internet]. Scirp.org. 2021 [citado el 9 de noviembre de 2021]. Disponible en: https://www.scirp.org/pdf/oalibj_2021092915280295.pdf
 167. Wikipedia contributors. Cluster (physics) [Internet]. Wikipedia, The Free Encyclopedia. 2017. [citado el 5 de mayo de 2021]. Disponible en: [https://en.wikipedia.org/w/index.php?title=Cluster_\(physics\)&oldid=1070997097](https://en.wikipedia.org/w/index.php?title=Cluster_(physics)&oldid=1070997097)
 168. Server CD. The ATLAS Collaboration Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung [Internet]. 2012 [citado el 8 de mayo de 2021]. Disponible en: <https://cds.cern.ch/record/1449796?ln=es>
 169. Wikipedia.org. Tracking (particle physics). [Internet]. 2017 [citado el 10 de mayo de 2021]. Disponible en: [https://en.wikipedia.org/wiki/Tracking_\(particle_physics\),_2017](https://en.wikipedia.org/wiki/Tracking_(particle_physics),_2017).
 170. Root Reference Guide. ROOT: TLorentzVector class reference [Internet]. Cern.ch. 2017 [citado el 9 de noviembre de 2021]. Disponible en: <https://root.cern.ch/doc/master/classTLorentzVector.html>
 171. Ocariz J. Probability and statistics for particle physicists [Internet]. arXiv [physics.data-an]. 2014 [citado el 15 de abril de 2021]. Disponible en: <http://arxiv.org/abs/1405.3402>
 172. Cowan G. Statistical Data Analysis. 1era Edición. Oxford, England: Clarendon Press; 1998.
 173. Wikipedia contributors. p-value [Internet]. Wikipedia, The Free Encyclopedia. 2017. Disponible en: <https://en.wikipedia.org/w/index.php?title=P-value&oldid=1075362413>
 174. ROOT team. Mathematical libraries [Internet]. ROOT. 2018 [citado el 11 de septiembre de 2021]. Disponible en: <https://root.cern/manual/math/>
 175. Wikipedia contributors. MINUIT [Internet]. Wikipedia, The Free Encyclopedia. 2017. Disponible en: <https://en.wikipedia.org/w/index.php?title=MINUIT&oldid=1070916412>

176. James. F. MINUIT Tutorial Function Minimization Reprinted from the Proceedings of the 1972 CERN Computing and Data Processing School [Internet]. 2004 [citado el 11 de septiembre de 2021]. Disponible en: <http://seal.web.cern.ch/seal/documents/minuit/mntutorial.pdf>