

CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS DEL  
INSTITUTO POLITÉCNICO NACIONAL

UNIDAD ZACATENCO  
DEPARTAMENTO DE FÍSICA

Estudio de técnicas multivariadas para la búsqueda de decaimientos  
 $B^0 \rightarrow D^-\pi^+\mu^+\mu^+$  en el experimento Belle II

**Tesis que presenta**

Alejandro De Yta Hernández

**para obtener el Grado de**

**Maestro en Ciencias**

**en la Especialidad de**

**Física**

**Director de Tesis:**

Dr. Eduard de la Cruz Burelo

Ciudad de México

Agosto 2017

*A mis padres, Rosa y Aureliano,  
a mis hermanos, Daniel y Jesús,  
con cariño y gratitud.*



# Agradecimientos

Mi más sincero agradecimiento es para mis padres, Rosa y Aureliano, que siempre han estado presentes y me han apoyado incondicionalmente. Gracias a ustedes aprendí a seguir adelante. A mis hermanos, Daniel y Jesús, que me hacen la vida más llevadera.

A mi asesor de tesis, Dr. Eduard de la Cruz Burelo, por guiarme con paciencia en la realización de este trabajo.

A mis profesores y sinodales, Dr. Iván Heredia de la Cruz y Dr. Pablo Roig Garcés, que tomaron parte de su tiempo para revisar y comentar este trabajo.

A todos mis profesores, que pacientemente compartieron sus conocimientos conmigo.

A Mary, por su total apoyo y comprensión.

A todos mis amigos y compañeros de generación, por su amistad, compañerismo y todas las experiencias compartidas.

A Michel E. Hernández Villanueva, por su valiosa ayuda.

A mi amigo Sam, por su amistad de todos estos años.

Al CINVESTAV, por sus atenciones durante este tiempo.

A CONACYT, por el apoyo (económico) que me brindo durante la realización de esta maestría, gracias.



# Resumen

En este trabajo se aplican y comparan tres diferentes métodos de clasificación multivariable basados en aprendizaje automatizado en el estudio del canal de decaimiento  $B^0 \rightarrow D^- \pi^- \mu^+ \mu^+$  en el experimento Belle II, localizado en la B-factory SuperKEKB. Este canal de decaimiento es de gran interés pues viola número leptónico total  $L$  en dos unidades,  $\Delta L = 2$ . Hasta la fecha no se ha encontrado experimentalmente, sin embargo de ser medido sería una prueba directa de que los neutrinos son fermiones de Majorana. La no observación de decaimientos con  $\Delta L = 2$  puede imponer constricciones importantes a los parámetros de masa y mezcla de un neutrino pesado de Majorana. Para realizar este estudio se generó una muestra de Monte Carlo de la cadena de interés mediante el uso del framework Basf2.



# Abstract

In this work, we apply and compare three different multivariate classification methods based on machine learning techniques in order to study the decay channel  $B^0 \rightarrow D^- \pi^- \mu^+ \mu^+$  in the Belle II experiment, located in the B-factory SuperKEKB. In this channel the total lepton number  $L$  is violated by two units,  $\Delta L = 2$ . Up to now, there is no experimental evidence of this process. However, if measured, it would be a direct proof of the Majorana nature of the neutrinos. The non-observation of this kind of processes can be turned out into important constraints on the Majorana neutrinos masses and mixings parameters. To carry out this study, a Monte Carlo sample of the decay channel of interest was generated using the Basf2 framework.





# Índice general

<b>1. Introducción</b>	<b>1</b>
<b>2. Neutrinos</b>	<b>5</b>
2.1. Decaimiento beta . . . . .	6
2.2. Neutrinos en el modelo estándar . . . . .	7
2.3. Neutrinos de Majorana . . . . .	8
2.4. Número leptónico . . . . .	8
2.5. Violación de número leptónico . . . . .	9
<b>3. El experimento Belle II</b>	<b>11</b>
3.1. Las B-factories . . . . .	11
3.2. Luminosidad . . . . .	14
3.3. SuperKEKB y el detector Belle II . . . . .	15
3.3.1. El colisionador SuperKEKB . . . . .	16
3.3.2. El detector Belle II . . . . .	17
<b>4. Computación y software</b>	<b>25</b>
4.1. Modelo computacional . . . . .	25
4.2. Modelo de análisis . . . . .	28
4.3. Producción de Monte-Carlo . . . . .	29
4.4. Simulación . . . . .	29
4.4.1. Generador de eventos . . . . .	30

4.4.2. Simulación del detector . . . . .	31
4.4.3. Digitalización . . . . .	32
4.5. Análisis . . . . .	33
<b>5. Técnicas multivariadas</b>	<b>35</b>
5.1. Árboles de decisión . . . . .	36
5.1.1. Construcción de un árbol . . . . .	36
5.1.2. Reglas de detención . . . . .	39
5.1.3. Árboles de decisión impulsados (BDT) en el paquete TMVA . . . . .	39
5.2. Redes neuronales . . . . .	40
5.2.1. Redes neuronales en el paquete TMVA . . . . .	41
5.2.2. Perceptrones multicapa . . . . .	41
5.3. Método de estimación por verosimilitud. . . . .	43
<b>6. Análisis</b>	<b>45</b>
6.1. Generación de la muestra de Monte Carlo . . . . .	45
6.2. Reconstrucción . . . . .	46
6.2.1. Implementación del paquete TMVA . . . . .	48
<b>Conclusiones</b>	<b>57</b>
<b>Apéndice A</b>	<b>59</b>
<b>Apéndice B</b>	<b>61</b>
<b>Bibliografía</b>	<b>63</b>

# Índice de figuras

2.1. Partículas elementales del modelo estándar. . . . .	7
3.1. Ambiente con alta multiplicidad observado en colisiones $pp$ en el experimento CMS [7]. . . . .	12
3.2. Ambiente con baja multiplicidad observado en colisiones $e^+e^-$ en el experimento Belle [8]. . . . .	13
3.3. Las B-factories han constreñido asombrosamente los parámetros del triángulo unitario. . . . .	14
3.4. SuperKEKB. . . . .	16
3.5. Detector Belle II. . . . .	18
3.6. Vista esquemática del arreglo geométrico de los sensores que forman el PXD. . . . .	19
3.7. Detector de silicio (SVD) mostrando sus 4 capas. . . . .	20
3.8. Distribución de los tubos que forman la CDC, en Belle II y en Belle, respectivamente. . . . .	21
3.9. Vista lateral del KLM, localizado fuera del ECL y del solenoide. Las líneas en gris muestran la aceptación angular polar nominal de Belle II. . . . .	23
3.10. Vista superior del detector Belle II. . . . .	24
3.11. Composición técnica del detector Belle II. . . . .	24
4.1. Un programa en BASF2 está formado por un conjunto de módulos que son ejecutados de forma lineal. . . . .	26

4.2.	Concepto del modelo computacional de Belle II. . . . .	27
4.3.	Diagrama de categorías por nivel de Geant4. . . . .	32
5.1.	Esquema de un árbol de decisión. . . . .	36
5.2.	Esquema de un arreglo multicapa con una capa oculta. . . . .	42
6.1.	Cortes en la variable de momento para identificar a las partículas de estado final candidatas a provenir del decaimiento $D^- \rightarrow K^+\pi^-\pi^-$ . . . . .	50
6.2.	Cortes en la variable de momento para identificar a las partículas de estado final candidatas a provenir del decaimiento directamente del mesón $B^0$ . . . . .	51
6.3.	Cortes en la variable PID para identificar a las partículas de estado final candidatas a provenir del decaimiento $D^- \rightarrow K^+\pi^-\pi^-$ . . . . .	52
6.4.	Cortes en la variable PID para identificar a las partículas de estado final candidatas a provenir del decaimiento directamente del mesón $B^0$ . . . . .	53
6.5.	Histograma de la masa invariante del mesón $D^-$ después de aplicar cortes en el momento y en la variable PID. . . . .	54
6.6.	Histograma de la masa invariante del mesón $B^0$ después de aplicar cortes en el momento y en la variable PID. . . . .	54
6.7.	Significancia como función de la variable $\alpha$ para los tres métodos implementados. . . . .	55
6.8.	Ajustes realizados aplicando cortes en la variable $\alpha$ asignada a los eventos por cada uno de los tres métodos implementados. . . . .	56

# Capítulo 1

## Introducción

El único proceso conocido en las primeras décadas de 1900 en el que aparecía la interacción débil era el decaimiento nuclear  $\beta$ . En 1914 Chadwick descubrió que el espectro de energía de los electrones liberados en el decaimiento  $\beta$  era continuo. Por muchos años este hecho se interpretó como un resultado de la pérdida de energía del electrón en el detector. Gracias a los experimentos realizados en 1927 por Ellis y Wooster se estableció que la energía promedio del electrón producido en decaimientos  $\beta$  es significativamente menor que la energía total liberada, con lo que demostraron que el espectro continuo del decaimiento  $\beta$  no puede explicarse con la energía perdida de los electrones en el detector. En otras palabras, parecía que se violaba la conservación de la energía. La idea de la existencia del neutrino fue propuesta por W. Pauli en una carta<sup>1</sup> enviada a un grupo de científicos en 1930. La existencia del neutrino, una partícula neutra con poca masa, de espín 1/2 y longitud de penetración grande (mucho más grande que la longitud de penetración del fotón), que sea emitida en el decaimiento  $\beta$  junto con el electrón, podía salvar la ley fundamental de la conservación de la energía.

Actualmente, sabemos que en la naturaleza existen 12 fermiones fundamentales:

---

<sup>1</sup>Ver Apéndice A

seis quarks ( $u, d, c, s, t, b$ ), tres leptones cargados ( $e, \mu, \tau$ ) y tres neutrinos ( $\nu_e, \nu_\mu, \nu_\tau$ ). En la lagrangiana de las interacciones electrodébiles, los neutrinos entran en el mismo nivel que los quarks y los leptones cargados. Apesar de esta similaridad los neutrinos son partículas muy especiales. Existen dos diferencias básicas entre los neutrinos y otros fermiones fundamentales:

- Hasta ahora, a todas las energías disponibles la sección eficaz de interacción de neutrinos con materia es muchos ordenes de magnitud menor que la sección eficaz de interacción electromagnética de los leptones cargados con la materia.
- La masa de un neutrino es varios ordenes de magnitud menor que la masa de un quark o un leptón cargado.

La medición de las masas de los neutrinos es un problema difícil y desafiante, que aún no se resuelve. Sin embargo, la observación del fenómeno de oscilación de los neutrinos permitió determinar la diferencia de los cuadrados de las masas de dos neutrinos. De los datos obtenidos de las oscilaciones de neutrinos y de experimentos de decaimientos  $\beta$  ha sido posible concluir que:

- Las masas de los neutrinos son diferentes de cero.
- Las masas de los neutrinos son menores a  $2 eV$ .

La explicación más común que se da para justificar que las masas de los neutrinos sean tan pequeñas se basa en la suposición de que se viola el número leptónico total. Si esta suposición es correcta, los neutrinos con masas definidas en realidad son fermiones neutros de Majorana. La observación de decaimientos  $\beta$  doble sin neutrinos sería una prueba de que los neutrinos son fermiones de Majorana.

La teoría que unifica las interacciones electromagnética y débil (Modelo Estándar) describe todos los datos experimentales que existen. Sin embargo, la existencia de la materia oscura y problemas internos del modelo estándar (tal como el problema de

las jerarquías) sugieren que existen teorías más allá del modelo estándar. Por otra parte, en este modelo los neutrinos son partículas estrictamente sin masa, sin carga eléctrica y sin color, por lo que únicamente interactúan de forma débil; sin embargo, ahora sabemos que en realidad los neutrinos son partículas con masa, lo que nos pone, nuevamente, en la necesidad de buscar física más allá del modelo estándar [1]. La posible existencia de neutrinos de Majorana puede examinarse mediante el estudio de procesos donde el número leptónico total  $L$  se viola en dos unidades,  $\Delta L = 2$ . Se estima que la razón de decaimiento de procesos con  $\Delta L = 2$  es bastante pequeña, por lo que se necesita una muestra de eventos bastante grande para la búsqueda de este tipo de procesos [2].

De particular interés para este trabajo es el canal de decaimiento de un mesón neutro  $B^0$  a cuatro cuerpos:  $B^0 \rightarrow D^- \pi^- \mu^+ \mu^+$ . Este proceso es una alternativa que se tiene para observar violación de número leptónico en dos unidades. Este canal no se ha observado experimentalmente; sin embargo vale la pena estudiarlo para saber si con la nueva generación de B-factories, en particular en el detector Belle II acoplado a SuperKEKB, será posible detectar este proceso. Un estudio previo de este canal ya se ha revisado, ver [11]. La ventaja de las B-factories es que están diseñadas para producir mesones  $B$  en grandes cantidades, por lo cual resulta ser la herramienta ideal para la búsqueda que deseamos realizar.

Esta tesis, se organiza de la siguiente manera.

- En el capítulo dos se revisa de forma breve el descubrimiento y el lugar que ocupan los neutrinos en el modelo estándar de física de partículas. Así como las consecuencias de que los neutrinos sean fermiones neutros de Majorana.
- En el capítulo tres, se comentan de forma breve la importancia de las B-factories. En particular, se resumen algunos aspectos importantes del acelerador de  $e^+e^-$  SuperKEKB y del experimento Belle II: sus mejoras, y los sub-



detectores principales que lo componen.

- En el capítulo cuatro, se da una descripción breve del modelo computacional y del framework que se emplea para realizar simulación y análisis en el experimento Belle II.
- En el capítulo cinco, se comentan de forma breve los métodos de clasificación multivariable que utilizamos para el estudio del canal de interés.
- Finalmente, en el capítulo seis se explica la metodología que se utilizó para realizar la simulación de Monte Carlo del canal  $B^0 \rightarrow D^- \pi^- \mu^+ \mu^+$  y el análisis de datos mediante la aplicación de diferentes técnicas de clasificación multivariable.

# Capítulo 2

## Neutrinos

La interacción débil se descubrió por primera vez en decaimientos beta ( $\beta$ ) en 1896 por Becquerel. En decaimientos alfa ( $\alpha$ ) y gama ( $\gamma$ ) se descubrió, simultáneamente, que las partículas emitidas tenían un espectro de energía monocromático, lo cual estaba de acuerdo con la diferencia de energías entre el núcleo padre e hija. Sin embargo, en los decaimientos  $\beta$ , la energía de los rayos  $\beta$  emitidos era continuo y no había señal de otras partículas. A saber, la ley de la conservación de la energía parecía violarse. Bohr estaba dispuesto a abandonar la ley de conservación de la energía, pero en 1930 Pauli propuso que una partícula (ahora llamada neutrino) hasta entonces desconocida era la responsable de esa pérdida de energía. En 1934, tomando en cuenta la hipótesis del neutrino de Pauli, Fermi propuso la primera versión de la teoría de interacción débil para explicar el fenómeno [3].

Un decaimiento  $\beta$  es un proceso en el cual un neutrón en el núcleo se convierte en un protón y al mismo tiempo emite un par electrón-antineutrino (del electrón):



Ahora sabemos que los neutrinos son partículas elementales, eléctricamente neutras, de espín 1/2. La probabilidad de que estas partículas interactúen con la materia es

despreciable, por lo cual simplemente pasan a través de ella. Hasta hace algunos años, se creía que eran partículas sin masa, como los fotones, pero en 1998 se descubrió que no es así [4].

## 2.1. Decaimiento beta

La teoría del decaimiento  $\beta$  fue propuesta por Fermi. Asumiendo la existencia del neutrino, formuló la teoría para la emisión de un electrón y un antineutrino usando un método similar al que se usa para estudiar la emisión de radiación por un átomo excitado. En la emisión de radiación por un átomo, la densidad hamiltoniana de interacción está formada por el producto escalar del vector de corriente electromagnética del electrón,  $\bar{\psi}(x)\gamma_\mu\psi(x)$ , y el potencial vectorial de campo electromagnético,  $A^\mu(x)$ ,

$$H(x) = e\bar{\psi}(x)\gamma_\mu\psi(x)A^\mu(x). \quad (2.2)$$

Por analogía, Fermi supuso que en el decaimiento  $\beta$ , la densidad hamiltoniana de interacción es el producto escalar de la corriente vectorial debida al sistema protón-neutrón con una corriente vectorial similar debida al sistema electrón-neutrino:

$$H(x) = G_F\bar{\psi}_p(x)\gamma_\mu\psi_n(x)\bar{\psi}_e(x)\gamma^\mu\psi_\nu(x). \quad (2.3)$$

En la emisión de radiación, la constante de acoplamiento es la carga eléctrica del electrón,  $e$ ; en el decaimiento  $\beta$ , la correspondiente constante de acoplamiento es  $G_F$ , llamada constante de acoplamiento débil de Fermi, y tiene dimensiones de energía  $\cdot$  volumen.

Después de muchos años de trabajo detallado se determinó que los neutrinos tienen helicidad izquierda  $-1$ , mientras que los antineutrinos tienen helicidad derecha  $+1$ , donde la helicidad es la proyección del espín en la dirección del momento. Expe-

rimentalmente, la dirección de movimiento del neutrino se determina de la conservación del momento, si se conoce el momento del electrón y el momento de retroceso dado al núcleo. [5]

## 2.2. Neutrinos en el modelo estándar

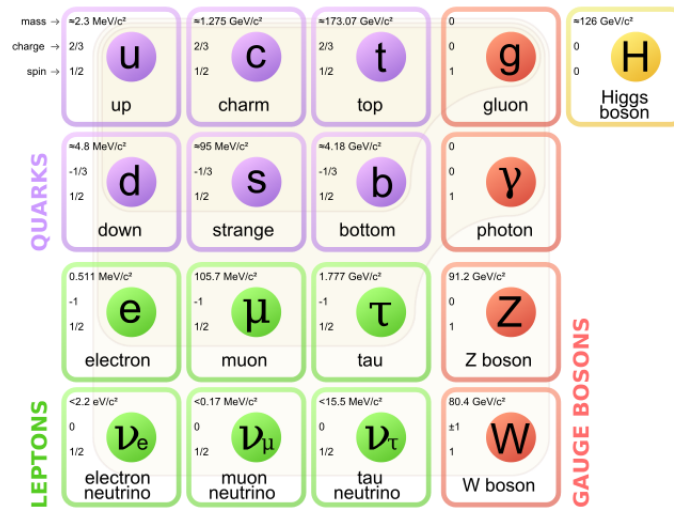


Figura 2.1: Partículas elementales del modelo estándar.

En el modelo estándar de física de partículas, los leptones aparecen en tres familias; cada una está formada por un leptón cargado:  $e^-$ ,  $\mu^-$ ,  $\tau^-$ , con carga eléctrica igual a la carga del electrón, y su correspondiente neutrino:  $\nu_e$ ,  $\nu_\mu$ ,  $\nu_\tau$ . De estas partículas, los neutrinos son las partículas menos conocidas. Lo que sabemos es que son partículas eléctricamente neutras, de espín  $1/2$  y que únicamente interactúan vía la interacción débil. En el modelo estándar los neutrinos son partículas con masa estrictamente cero, no tienen carga electromagnética ni de color, solamente carga débil. Bajo la conservación de la simetría CPT (conjugación de carga, paridad e inversión temporal, respectivamente) para cualquier fermión izquierdo existe una antipartícula derecha con carga opuesta. Pero no necesariamente existen partículas en estados derechos. Precisamente esto ocurre con los neutrinos en el modelo estándar. Cuando se postuló el modelo estándar, las masas de los neutrinos eran compatibles

con cero, por lo que los neutrinos fueron postulados como fermiones de Weyl: la partícula izquierda era el neutrino y la partícula derecha el antineutrino. [6]

### 2.3. Neutrinos de Majorana

En la teoría de Dirac del electrón, el electrón y el positrón son partículas diferentes, una es la antipartícula de la otra. Podemos distinguirlas por el signo de su carga. Por otro lado, existen bosones neutros, tal como el mesón  $\pi^0$ , para los cuales su antipartícula es la misma partícula. La pregunta natural es: ¿existe una teoría para fermiones neutros en la cual la antipartícula es la misma partícula? En efecto, tal teoría es posible, y fue propuesta por Majorana en 1937.

La siguiente pregunta que uno se puede plantear es: ¿los neutrinos son fermiones de Dirac o fermiones de Majorana? Para fermiones masivos, se pueden construir proyecciones quirales con quiralidad izquierda y quiralidad derecha, como número cuántico. En el modelo estándar de interacciones electrodébiles, sólo los neutrinos con quiralidad izquierda  $\nu_L$  y los antineutrinos con quiralidad derecha  $\bar{\nu}_R$  interactúan con los campos de norma. Las componentes con quiralidad opuesta,  $\nu_R$  y  $\bar{\nu}_L$ , son estériles, en el sentido de que no participan en las interacciones débiles. Actualmente, parece que el único experimento factible para determinar si la naturaleza de los neutrinos es de Majorana es el decaimiento beta doble sin neutrinos. El proceso de decaimiento beta doble sin neutrinos, es un proceso que puede ocurrir sólo si se viola la conservación de número leptónico en dos unidades. Un término de masa que viole número leptónico en dos unidades se conoce como "término de masa de Majorana".

### 2.4. Número leptónico

El número leptónico es un número cuántico asociado con los neutrinos y sus correspondientes leptones cargados. Al neutrino del electrón,  $\nu_e$ , y al electrón,  $e^-$ , se les

asigna un número leptónico (de la familia del electrón)  $l_e = +1$ ; y sus antipartículas  $\bar{\nu}_e$  y  $e^+$  tienen el valor opuesto  $l_e = -1$ . Similarmente, para las otras dos familias de leptones se asignan números leptónicos:  $l_\mu$ , para la familia del muon, y  $l_\tau$ , para la familia del tau. Al resto de partículas que no son leptones se les asigna un número leptónico igual a cero. En cualquier tipo de reacción, aparentemente se conserva el número leptónico por familia, independientemente de las otras familias y el número leptónico total. Por ejemplo, en el decaimiento  $\mu^+ \rightarrow e^+ + \nu_e + \bar{\nu}_\mu$ , tenemos  $l_\mu = -1$ , y  $l_e = 0$  de ambos lados [5].

## 2.5. Violación de número leptónico

Ahora que se ha acumulado suficiente evidencia experimental de las oscilaciones de los neutrinos, es un hecho que estos son partículas masivas. Esta es una de las motivaciones más fuertes para considerar nuevos escenarios de física más allá del modelo estándar (explicar el mecanismo de generación de sus masas). Además, la naturaleza de los neutrinos (partículas de Dirac o Majorana) aún es un problema abierto en la física de partículas.

Se sabe que si la naturaleza de los neutrinos es de Majorana, sus efectos deben de manifestarse a través de la observación de decaimientos y fenómenos de producción donde el número leptónico total se viola en dos unidades,  $\Delta L = 2$ . Una señal clara de procesos con  $\Delta L = 2$  implica la producción de dileptones del mismo signo en el estado final. El proceso más atractivo para observar este tipo de violación de número leptónico es el decaimiento nuclear beta doble sin neutrinos ( $0\nu\beta\beta$ ), el cual es considerado como la forma más factible de observarlo. El hecho de observar este decaimiento nuclear confirmaría la existencia de procesos con violación de número leptónico, implicando que los neutrinos son partículas de Majorana. Hasta ahora, no se ha observado experimentalmente el decaimiento nuclear  $0\nu\beta\beta$ .

En la literatura se han propuesto diferentes procesos con  $\Delta L = 2$  a altas y bajas energías como alternativa para probar la naturaleza de Majorana de los neutrinos.

Entre varias posibilidades, una fuente importante de procesos con LNV (violación de número leptónico) está dada por el intercambio de un neutrino de Majorana con una masa intermedia del orden de la escala de masa de los sabores pesados. Los efectos de tales neutrinos con masas de alrededor de  $100 \text{ MeV}$  y hasta algunos pocos  $\text{GeV}$  se han estudiado ampliamente en decaimientos de mesones  $B$  y  $D$  a cuatro cuerpos con  $\Delta L = 2$  como:  $\bar{B}^0 \rightarrow D^+ \pi^+ l^- l^-$ ,  $\bar{D}^0 \rightarrow (\pi^+ \pi^+, K^+ \pi^+) \mu^- \mu^-$ , y  $B^- \rightarrow D^0 \pi^+ \mu^- \mu^-$ , y adicionalmente, como procesos complementarios, decaimientos de leptones tau a cuatro cuerpos:  $\tau^- \rightarrow \pi^+ \nu_\mu l^- l^-$ . Colaboraciones experimentales tales como BABAR, Belle, y LHCb han reportado importantes esfuerzos experimentales en la búsqueda de decaimientos con  $\Delta L = 2$ , pero hasta ahora no se ha encontrado evidencia de tales procesos. La no observación de decaimientos con  $\Delta L = 2$  puede imponer restricciones significantes a los parámetros (masa y mezcla) de un neutrino pesado de Majorana [2].

# Capítulo 3

## El experimento Belle II

El experimento Belle II es una actualización del experimento Belle, acoplado al colisionador circular asimétrico de electrón-positrón SuperKEKB ubicado en Tsukuba, Japon, y provee una nueva forma de estudiar decaimientos de mesones  $B$ , mesones charm y leptones  $\tau$ . Con esta actualización se espera recolectar el equivalente a 50 veces más datos que su predecesor, Belle, lo cual permitirá hacer medidas de alta precisión, buscar decaimientos raros y nuevas partículas.

### 3.1. Las B-factories

Una B-factory es un experimento acoplado a un colisionador de partículas para producir y analizar un gran número de mesones  $B$ , partículas que contienen un quark bottom. Uno de los roles de las B-factories es mostrar la diferencia entre mesones  $B$  y su antipartícula para tratar de entender por qué el universo está dominado por materia. Una B-factory está formada, grosso modo, por tres partes: un colisionador de partículas que produce un gran número de mesones  $B$ ; detectores para observar los decaimientos, y una colaboración internacional.

Una de las principales ventajas de las B-factories se relaciona con el hecho de que el estado inicial es conocido, pues únicamente hay dos mesones  $B$  en el estado



final del decaimiento de un  $\Upsilon(4S)$ , sin partículas adicionales. Así, la reconstrucción de uno de los mesones  $B$  construye el cuadrivector y el sabor del otro. Esta característica comúnmente se usa en el método de "tagging", donde el primer mesón  $B$  puede reconstruirse completamente en modos hadrónicos, o parcialmente en modos semileptónicos.

Otra característica importante, se relaciona con la limpieza del ambiente experimental, el bajo nivel de ruido facilita la reconstrucción de canales con  $\gamma$ ,  $\pi^0$ ,  $\rho$  y  $\eta$  en el estado final. Las bajas multiplicidad en el tracker y ocupación del detector permiten una alta eficiencia de reconstrucción para  $B$ ,  $D$  y  $\tau$ , junto con un filtro de activación (trigger) bastante bajo, este hecho se ejemplifica en las figuras (3.1) y (3.2). Esto reduce las correcciones y errores sistemáticos en varios tipos de mediciones, en particular en los análisis de Dalitz plot.

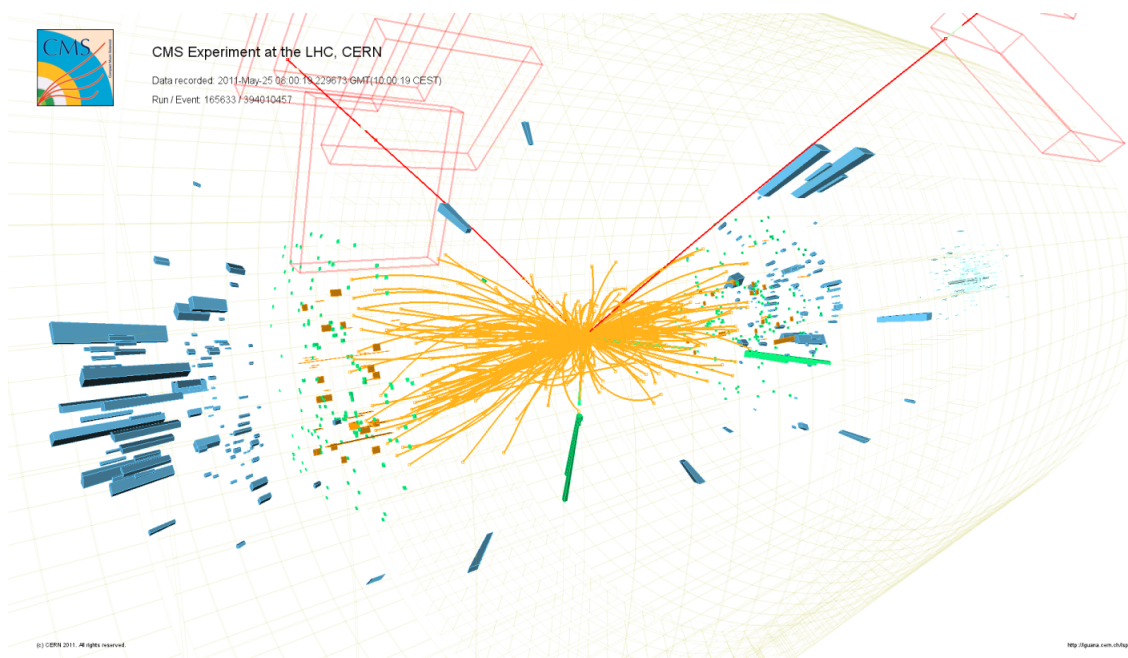


Figura 3.1: Ambiente con alta multiplicidad observado en colisiones  $pp$  en el experimento CMS [7].

Una característica excepcional de las B-factories es la posibilidad de hacer un

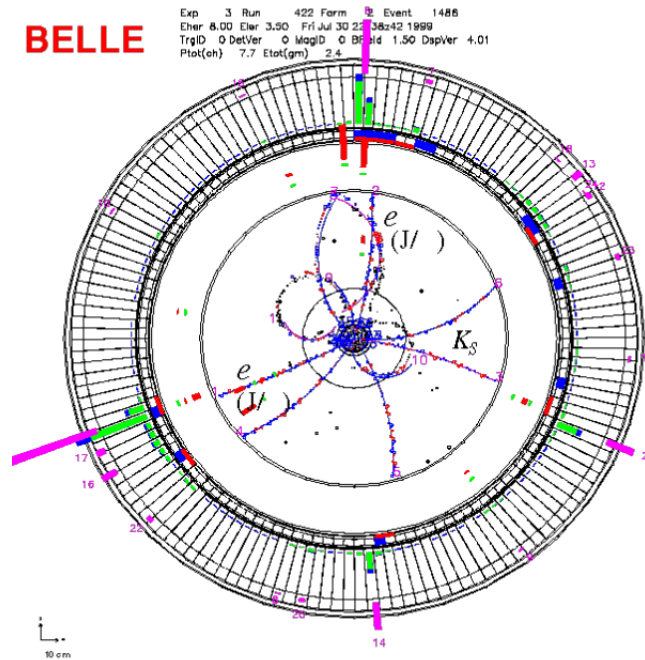
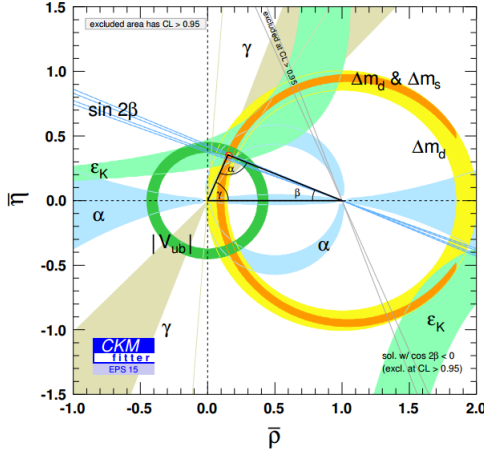


Figura 3.2: Ambiente con baja multiplicidad observado en colisiones  $e^+e^-$  en el experimento Belle [8].

escaneo de energía, donde la energía del haz puede ajustarse a diferentes valores correspondientes al estado específico  $\Upsilon$ . Así, aparte de la muestra de datos de la resonancia  $\Upsilon(4S)$ , también se pueden guardar muestras en las resonancias  $\Upsilon(1S)$ ,  $\Upsilon(2S)$ ,  $\Upsilon(3S)$  y  $\Upsilon(5S)$ .

Las características discutidas han resultado en una operación exitosa de B-factories durante la década pasada: Belle en KEKB en KEK, y BaBar en PEP2 en SLAC. Entre sus muchos logros, uno de los principales resultados fue el descubrimiento de violación CP en decaimientos de  $B$ , en particular, violación CP dependiente del tiempo (TCPV) en  $B^0 \rightarrow J/\psi K^0$ ,  $B^0 \rightarrow K^{*0}\gamma$ , etc., y violación directa CP (DCPV) en  $B^0 \rightarrow \pi\pi$  o decaimientos  $B^0 \rightarrow K\pi$ . Este tipo de estudios también contribuyen a la determinación de los lados y ángulos del Triángulo Unitario, figura (3.3), el cual es uno de los test de mayor precisión del Modelo Estándar. Además, se han estudiado muchos decaimientos raros de  $B$ , incluyendo transiciones altamente suprimidas

como  $b \rightarrow d\gamma$  y  $b \rightarrow sl^+l^-$ , cuya existencia se estableció en Belle. También hay aspectos interesantes en el sector charm. A saber, se han encontrado mezcla  $D^0 - \overline{D}^0$  y nuevos estados interesantes, algunos inesperados desde el punto de vista teórico. Algunos ejemplos son estados tipo charmonium  $X(3872)$ , y  $Z(4430)^+$ , los cuales fueron observados en Belle por encima del umbral del charm [9], [10].



(a) Triángulo unitario

UT 2014	Belle II
$\alpha$ 4° (WA)	1°
$\beta$ 0.8° (WA)	0.2°
$\gamma$ 8.5° (WA)	1-1.5°
	14° (Belle)

(b) Valores que se espera alcanzar en Belle II con  $50 ab^{-1}$ 

Figura 3.3: Las B-factories han constreñido asombrosamente los parámetros del triángulo unitario.

## 3.2. Luminosidad

La luminosidad instantánea se define como la cantidad que da una taza de eventos cuando se multiplica por la sección eficaz

$$N(\text{taza de eventos}/s) = \sigma L. \quad (3.1)$$

Las unidades de  $L$  son (en el sistema cgs)  $cm^{-2}s^{-1}$ . La sección eficaz  $\sigma$  está determinada por la naturaleza y no es controlable. Por lo tanto, el número de eventos es directamente proporcional a la luminosidad.

Cuando existen varias fuentes de ruido, uno tiene que filtrarlas usualmente aplicando cortes en algunas de las variables cinemáticas, tales como el momento, ángulos, localizaciones espaciales, etc. Sin embargo, no es posible hacer un filtro ideal que sea capaz de quitar todos los eventos no deseados y dejar pasar únicamente la señal buscada. En la mayoría de los casos, la señal de interés también se reduce. Cuando el ruido es abundante, los eventos tienen que pasar por muchos filtros y la señal que sobrevive depende críticamente de la luminosidad.

A partir de la luminosidad instantánea se define la luminosidad integrada

$$L_{int} = \int L dt, \quad (3.2)$$

de donde podemos conocer el número total de eventos buscados en un colisionador de partículas

$$N = \sigma L_{int}. \quad (3.3)$$

Las unidades de  $L_{int}$  son (en el sistema cgs)  $cm^{-2}$ . Dadas las magnitudes típicas de la luminosidad frecuentemente se utiliza el barn ( $b$ ) que, por definición, tiene un valor numérico de  $1b = 10^{-24} cm^2$ .

### 3.3. SuperKEKB y el detector Belle II

Las B-factories han tenido un periodo operacional exitoso al recolectar una muestra total de datos equivalente a  $1.5 ab^{-1}$  de luminosidad integrada, que corresponde a  $1.2 \times 10^9$  pares de mesones  $B$ . En particular, la B-factory Belle acoplada a KEKB alcanzó la luminosidad instantánea más grande del mundo,  $2 \times 10^{34} cm^{-2} s^{-1}$ , acumulando datos equivalentes a  $1 ab^{-1}$  de luminosidad integrada. Las colisiones se realizaron a una energía del centro de masa igual a la masa de  $\Upsilon(nS)$ , principalmente a  $m_{\Upsilon(4S)} = 10.58 GeV$ , almacenando datos equivalentes a  $710 fb^{-1}$ , correspondiente a 772 millones de pares de mesones B [10] [11].

Teniendo en cuenta todos los grandes logros que las B-factories han tenido en la década pasada, es clara la necesidad de una nueva generación de B-factories. Una nueva generación de (super) B-factories hará posible realizar mediciones de alta precisión en casi todos los elementos de la matriz  $CKM$  mediante decaimientos de mesones  $B$  y  $D$ . También son clave en el test de modelos más allá del Modelo Estándar al medir y/o buscar mezclas, CPV, violación de sabor y número leptónico (LFV y LNV, respectivamente), y decaimientos raros de partículas  $B$ ,  $D$ , y  $\tau$ . Además, es posible la búsqueda directa de partículas ligeras tales como neutrinos estériles, fotones oscuros, bosones de Higgs ligeros y axiones [10].

### 3.3.1. El colisionador SuperKEKB

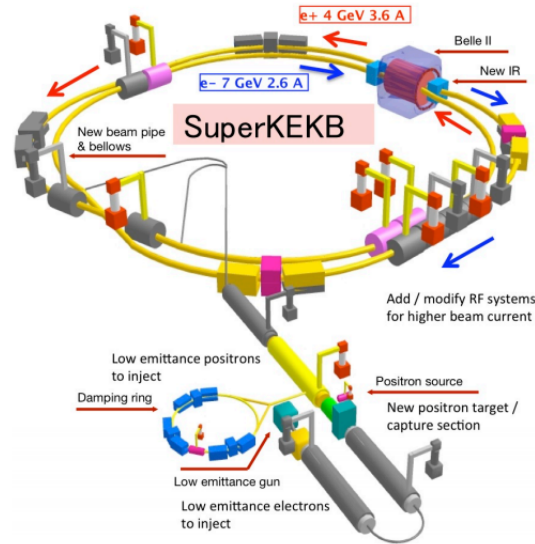


Figura 3.4: SuperKEKB.

El colisionador SuperKEKB, figura(3.4), está localizado en KEK (The High Energy Accelerator Research Organization) en Tsukuba, Japón, y es la actualización más grande hecha al acelerador KEKB. Consiste de dos anillos de 3  $Km$  equipados con sistemas de radio frecuencia (RF) los cuales aceleran a los haces de electrones y positrones a 7 y 4  $GeV$ , respectivamente, y los hacen colisionar en el centro del

detector Belle II. Respecto a KEKEB, SuperKEKEB tiene dipolos más grandes y redes de magnetos rediseñados para reducir la emitancia, e incluye sistemas de radio frecuencia adicionales o modificados para obtener corrientes de haz mayores ( $\times 2$ ). SuperKEKB reduce el tamaño del haz ( $\times 20$ ) usando un esquema de nano-haz, lo cual se logra al usar los nuevos magnetos cuadrupolares de enfoque final cerca de la región de interacción, además se incrementa la luminosidad usando cavidades de canchero. EL colisionador actualizado permitirá una luminosidad integrada de  $50 \text{ ab}^{-1}$ , y alcanzará una luminosidad máxima de  $8 \times 10^{35} \text{ cm}^{-2} \text{ s}^{-1}$ , cerca de 50 y 40 veces lo que alcanzó KEKB, respectivamente, con lo cual se tendrá cerca de  $10^{10}$  pares de mesones  $B$  o pares de leptones  $\tau$  por año [10], [12].

Colisionador	PEP-II	KEKB	SuperKEKB
Detector	BaBar	Belle	Belle II
Inicia toma de datos	1998	1998	2018
Fin de toma de datos	2008	2010	–
Energía (GeV): $e^-$	9	8	7
$e^+$	3.1	3.5	4
Luminosidad máxima ( $\text{cm}^{-2} \text{s}^{-1}$ )	$12.069 \times 10^{33}$	$2.1 \times 10^{34}$	$8 \times 10^{35}$

Cuadro 3.1: Parámetros de las B-factories mencionadas [13], [14].

### 3.3.2. El detector Belle II

Construir el detector Belle II, figura (3.5), implica actualizar o reemplazar todos los subdetectores del detector Belle, para aprovechar los avances en la tecnología de los detectores y asegurar que el nuevo detector pueda operar adecuadamente en el ambiente de luminosidad incrementada que SuperKEKB proveerá. Por ejemplo, un tracker más grande mejorará el parámetro de impacto y la resolución de vértices secundarios; también, incrementará la eficiencia en la detección de  $K_S^0$  y piones. Una descripción completa del detector Belle II se puede hallar en otras fuentes [15]; algunos puntos importantes se mencionan aquí.

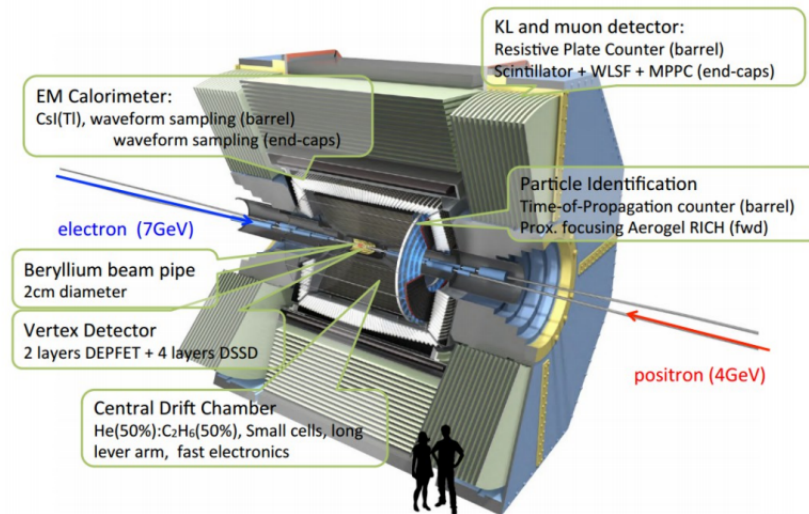


Figura 3.5: Detector Belle II.

### Región de interacción (IR)

El haz de electrones con energía de  $7\text{ GeV}$ , que circula en el anillo de alta energía (HER), y el haz de positrones con energía de  $4\text{ GeV}$ , que circula en sentido contrario en el anillo de baja energía (LER), chocan en un punto de interacción (IP) con un ángulo de  $83\text{ mrad}$ . En SuperKEKB la luminosidad esperada es 40 veces mayor que la alcanzada en KEKB, esto gracias a las altas corrientes de haz y la pequeña región de interacción de los haces. Debido a esto, el haz induce ruido en el detector Belle II. Para asegurar que el detector opere adecuadamente a tan alta luminosidad, el diseño de una región de interacción basada en estimaciones del ruido producido por el haz es importante. Detalles y propuestas se pueden encontrar en [15].

### Detector de pixeles (PXD)

A la alta luminosidad prevista para SuperKEKB, los detectores cercanos al haz son bombardeados por partículas a razones de colisión bastante altas, por lo que los detectores convencionales de silicio no son óptimos para realizar reconstrucciones. Para esto se colocó un detector de pixeles (PXD) que forma parte del sistema de tracking, este está formado por placas de  $50\ \mu\text{m}$  de grosor que forman dos cilindros

colineales al haz, figura (3.6), no requiere de un sistema de enfriamiento gracias a que consume poco energía.

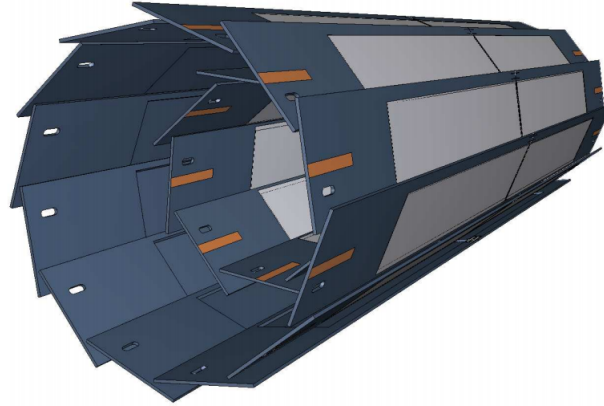


Figura 3.6: Vista esquemática del arreglo geométrico de los sensores que forman el PXD.

### Detector de silicio (SVD)

El trabajo principal del detector de silicio (SVD), figura (3.7), de Belle II, junto con el detector de pixeles (PXD) y la cámara central de acumulación (CDC), es medir los vértices del decaimiento de un par de B's para la medición de la asimetría  $CP$  inducida por la mezcla. Además, este detector capta información de los vértices en otros canales de decaimiento que involucran decaimientos de mesones  $D$  y leptones  $\tau$ . El diseño del detector de silicio (SVD) hereda las buenas características del detector de vértices de Belle: baja masa, alta precisión, inmunidad a ruido de fondo, tolerancia a radiación y estabilidad mecánica a largo plazo.

Algunas de sus características son las siguientes:

- Cubre completamente la región de aceptación definida por  $17^\circ < \theta < 150^\circ$ . El volumen que ocupa el SVD está limitado por los radios interno y externo que son, respectivamente,  $38\text{ mm}$  y  $140\text{ mm}$ , que a su vez están determinados por los radios del PXD y la CDC.



- Provee de datos para extrapolar los tracks reconstruidos en la CDC al PXD con gran eficiencia; esta es la estrategia que usamos para determinar un vertice.
- Junto con el PXD, el SVD es capaz de reconstruir tracks de bajo  $p_t$ , hasta pocas decenas de  $MeV/c$ , que regularmente no llegan a la CDC. Esto es de particular importancia para la eficiencia de reconstrucción de las partículas hijas del mesón  $D^*$  que etiqueta el sabor del mesón  $B$  padre.

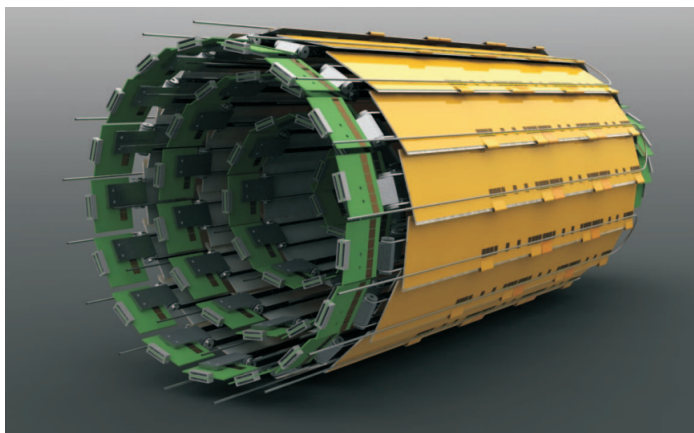


Figura 3.7: Detector de silicio (SVD) mostrando sus 4 capas.

### Cámara central de acumulación (CDC)

En el detector Belle II, la cámara central de acumulación (CDC) juega tres papeles importantes. El primero, reconstruye tracks cargados y mide de forma precisa su momento. Segundo, provee información de la identificación de partículas usando mediciones de las pérdidas de energía. Tracks de bajo momento, los cuales no alcanzan el identificador de partículas, pueden identificarse usando solamente la CDC. Finalmente, provee de señales de disparo (trigger) eficientes y fiables para partículas cargadas.

Esta cámara está formada por 14366 tubos que contienen helio ( $He$ ) y etano ( $C_2H_6$ ) en proporciones iguales, a través de estos se coloca un alambre que se somete a altos voltajes. La CDC detecta partículas mediante la ionización del gas.

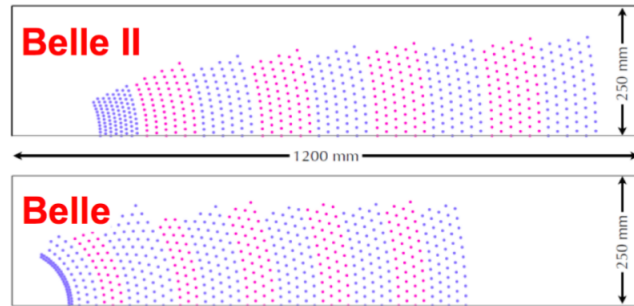


Figura 3.8: Distribución de los tubos que forman la CDC, en Belle II y en Belle, respectivamente.

### Sistema de identificación de partículas (PID)

En Belle II hay dos contadores, cuyo funcionamiento está basado en radiación Cherenkov, en las regiones del barril central y capas laterales (forward region) que permiten la identificación de partículas. El contador TOP localizado en la región del barril consiste de 16 módulos localizados en la separación que hay entre la CDC y el calorímetro. Un módulo TOP está constituido por dos barras de cuarzo con un arreglo de fotodetectores, un espejo y un prisma de extensión que recolecta la radiación Cherenkov producida por las partículas cargadas en los bordes de las barras. Este sistema discrimina principalmente entre piones y kaones. En las paredes laterales transversales al haz se utiliza un sistema de aerogel multicapa con diferentes índices de refracción, lo que además permite, en conjunto con el resto del detector, discriminar entre piones, muones y electrones con momento por debajo de  $1 \text{ GeV}$ .

### Calorímetro (ECL)

Dado que una tercera parte de los productos de decaimientos de mesones  $B$  son piones neutros,  $\pi^0$ , u otras partículas neutras que producen fotones en el amplio rango de energías que va de  $20 \text{ MeV}$  a  $4 \text{ GeV}$ , un calorímetro de alta resolución es fundamental para el detector Belle II. Por sus buenas características, se usó Yoduro de cesio,  $\text{CsI}$ , como material centelleador para la construcción del calorímetro. Las

principales tareas de un calorímetro son :

- Detección de fotones con gran eficiencia.
- Determinación precisa de la energía de los fotones.
- Identificación de electrones.
- Generación de una señal de disparo (trigger) apropiada.

El ECL instalado en Belle II se encuentra dentro de un solenoide superconductor que produce un campo magnético de  $1.5 T$ . El ECL usa cristal de  $CsI$  en el barril y  $CsI$  en las capas finales, y está equipado con electrónica más rápida respecto a la de Belle. Dadas sus características geométricas el ECL cubre la región angular  $12.4^\circ < \theta < 155.1^\circ$  a partir del punto de interacción.

### Detector de $K_L^0$ y $\mu$ (KLM)

El detector de kaones y muones (KLM), figura (3.9), consiste de un arreglo alter-nante de placas de acero de  $4.7 cm$  de grosor y elementos detectores activos colocados fuera del solenoide superconductor en forma octagonal.

La identificación de muones inicia con la reconstrucción de una partícula cargada en la CDC. Cada track se extrapola hacia fuera de la CDC identificandola como pión. Si un muón impacta en la RCP se generará un fotón que es recuperado por fibra óptica y transmitida a un fotodetector. Mediante reconstrucción se puede decidir si el track detectado en la CDC es un muón. Análogamente para kaones.

### Solenoide superconductor

Un solenoide superconductor proporciona un campo magnético de  $1.5 T$  en un volumen cilíndrico de  $3.4 m$  de diámetro y  $4.4 m$  de largo. La bobina está rodeada por una estructura multicapa formada por placas de acero y calorímetros.

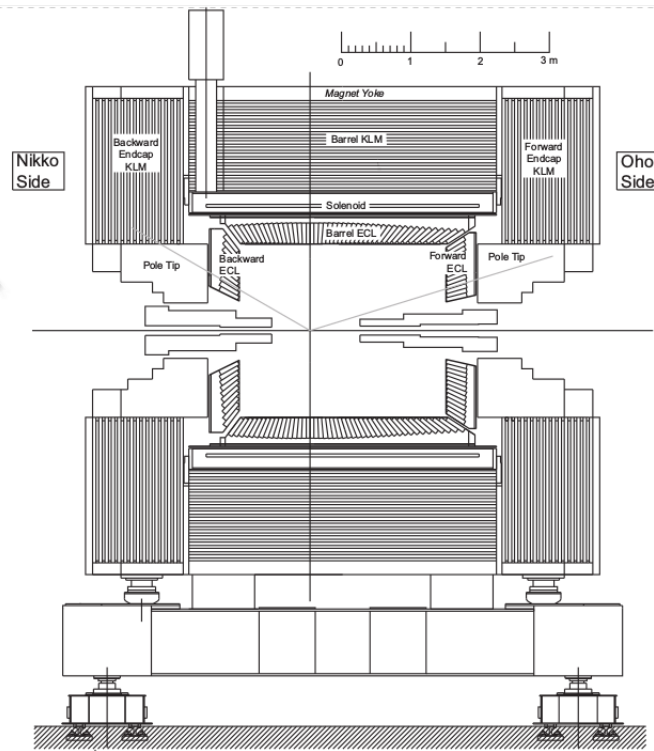


Figura 3.9: Vista lateral del KLM, localizado fuera del ECL y del solenoide. Las líneas en gris muestran la aceptación angular polar nominal de Belle II.

### Sistema de trigger y adquisición de datos

Se espera que la dispersión Bhabha ( $e^+e^- \rightarrow e^+e^-$ ) y procesos  $\gamma\gamma$ , también como la dispersión intrahaz (Touschek) sean las principales fuentes de ruido en Belle II. Para trabajar con física de muchos eventos en poco tiempo, se implementará un sistema de trigger de dos etapas basado en el hardware (L1) y el software (HLT) para reducir la frecuencia a menos de  $10\text{KHz}$ , guardando únicamente las colisiones de interés. El sistema de adquisición de datos (DAQ) recolecta la información de los subdetectores y produce eventos de unos  $200\text{KB}$ , de lo cual se tiene una rapidez de almacenamiento de  $1.5\text{GB/s}$ , aproximadamente. Belle II usa un modelo computacional Grid basado en varios niveles, llamados Tiers, para distribuir el almacenamiento de datos sin procesar, datos procesados, producción de simulación Monte-Carlo, etc., en diversos centros de datos en más de 17 países. Detalles más específicos pueden encontrarse en [15].

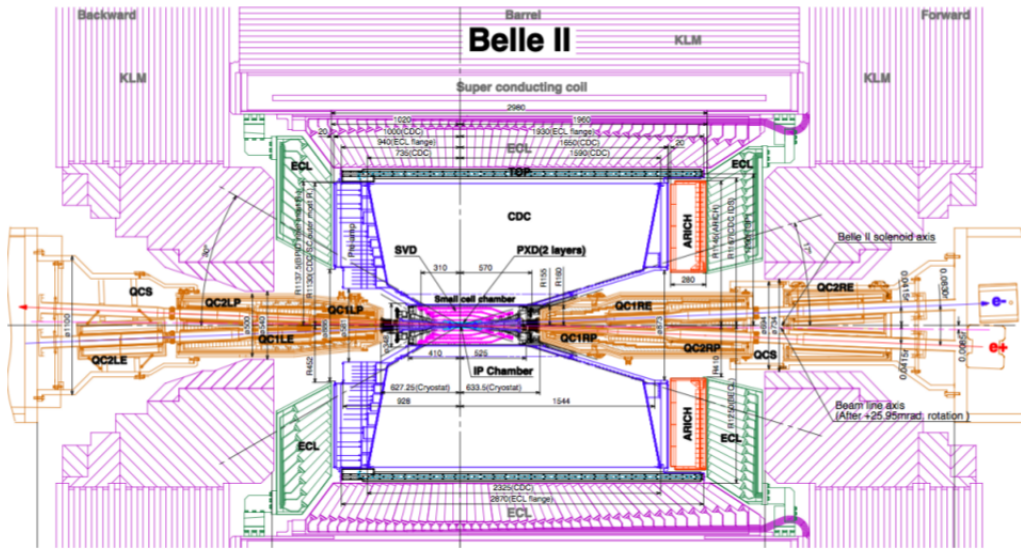


Figura 3.10: Vista superior del detector Belle II.

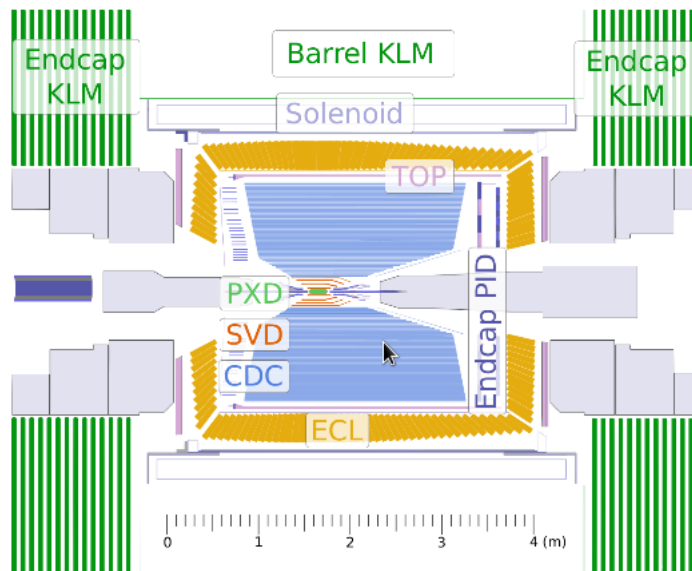


Figura 3.11: Composición técnica del detector Belle II.

# Capítulo 4

## Computación y software

La luminosidad del colisionador de  $e^+e^-$ , SuperKEKB, aumentará en un factor de 40, lo cual permitirá crear una muestra de datos 50 veces mayor comparada con la muestra de Belle. Sin embargo, el ruido y la tasa de eventos activados también se incrementará al menos por un factor de 10. El software de Belle II está diseñado para adaptarse a estos retos y funcionar en la grid, en la nube, y en recursos locales alrededor del mundo. Se utilizan varios paquetes de software externos para mejorar la interfaz del usuario. El sistema de software, BASF2 (Belle Analisis Framework 2), está estructurado como un marco construido con módulos de carga dinámica y la posibilidad de realizar procesos en paralelo. La estructura de BASF2 está basada en una cadena de módulos que se ejecutan linealmente mientras interactúan con un almacén de datos, figura (4.1). El sistema está escrito en  $C++$  con scripts guía en *Python*. También se desarrolló una librería para simular completamente el detector basada en Geant4 [16].

### 4.1. Modelo computacional

El modelo computacional de Belle II asume la siguiente estructura basada en tecnologías de grid [15]:

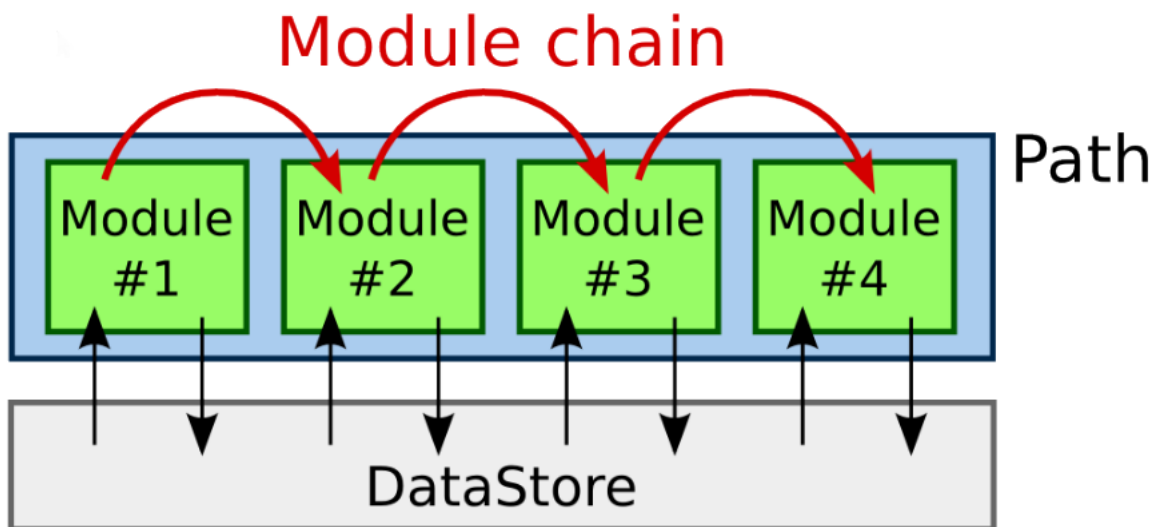


Figura 4.1: Un programa en BASF2 está formado por un conjunto de módulos que son ejecutados de forma lineal.

- El centro principal se encuentra en KEK y es responsable del almacenamiento de los datos sin procesar, y procesados. Se espera que el centro principal opere 24/7, está diseñado para ser extremadamente confiable.
- Los sitios grid proporcionan recursos informáticos geográficamente cercanos al usuario. Sumado al rol de centro principal, KEK tendrá el rol de sitio grid. Los sitios grid proporcionan recursos para la administración de la producción y almacenamiento de muestras genéricas de MC. También permiten al usuario crear ntuples a partir de datos y muestras de Monte-Carlo existentes.
- Los análisis basados en ntuples usualmente se realizan en los recursos locales, ya sea computadoras personales o clusters institucionales.

El framework de análisis de datos es la base del software de Belle II. Varios tipos de componentes del software están escritos como "módulos" para el framework, y son agregados al framework bajo demanda. Los programas en Belle II, tales como la reconstrucción de eventos, y los códigos de los usuarios, se escriben combinando un conjunto de módulos del framework. El framework lee un evento guardado de un

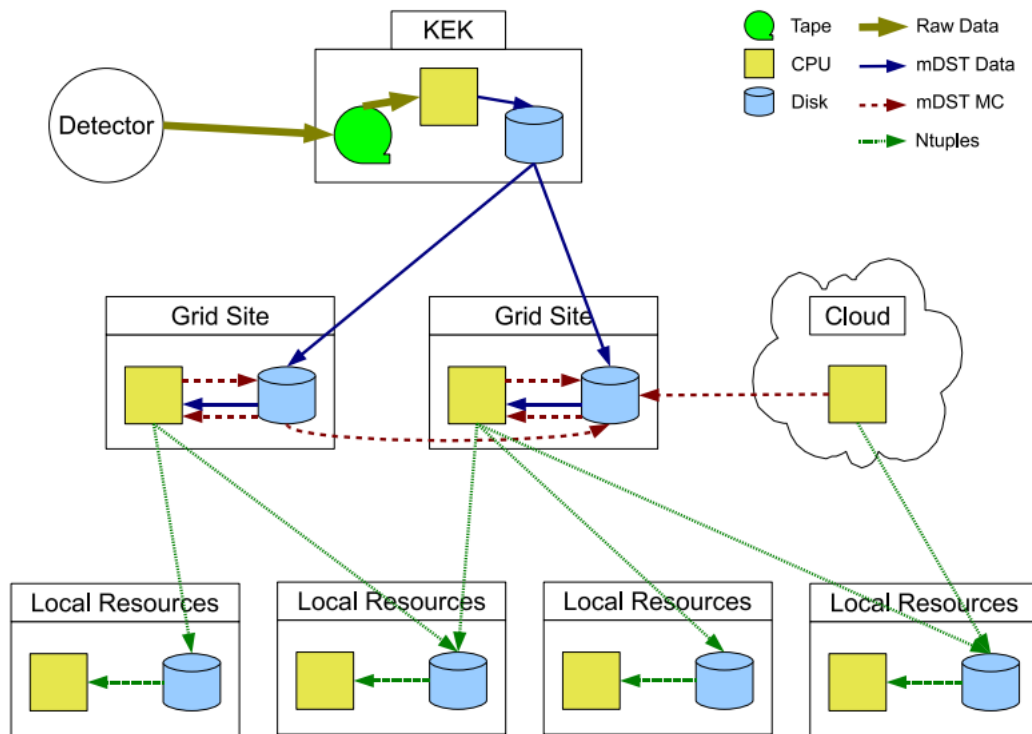


Figura 4.2: Concepto del modelo computacional de Belle II.

archivo, entonces lo procesa ejecutando una cadena de módulos, y guarda el evento resultante en un archivo nuevo. El almacenamiento de los histogramas, los ntuples, y el resto de información solicitada también es gestionado por el framework.

El framework se divide en varios paquetes que reflejan diferentes aspectos del detector. Esos paquetes contienen objetos y módulos. Los módulos realizan tareas específicas sobre los objetos. Algunos de los paquetes son:

- Analisis.
- Adquisicion de datos (DAQ).
- ir, cdc, ecl, ekln, pxd, svd, top, vxd (módulos asociados a subdetectores).
- Reconstrucción.
- Simulacion (MC).



cuyas tareas pueden ser la escritura y/o lectura de datos, la simulación, la reconstrucción, el análisis, entre muchas otras. Adicionalmente, el framework incluye paquetes de terceros, entre ellos se encuentran: ROOT, genfit2, EvtGen, GEANT4, por mencionar algunos.

## 4.2. Modelo de análisis

Una parte importante en los análisis de la física, y el procesamiento de los datos obtenidos, es el procesamiento de archivos seleccionados por una parte del código de análisis escrito por el usuario. A esta tarea le conocemos como proyecto. Para obtener resultados físicos correctos, es esencial que cada archivo de entrada sea procesado exactamente una vez. Si algunos archivos se pierden o se procesan dos veces, los resultados que obtenemos del análisis muy probablemente sean erróneos. Por lo tanto, tenemos que diseñar un sistema que minimize el riesgo de fallar con los requisitos.

Dado que el número de archivos pertenecientes al proyecto usualmente es grande, para procesarlos utilizamos varias copias ejecutadas en paralelo con diferentes archivos de entrada. Desde el punto de vista del sistema, todas estas copias paralelas son independientes y no están relacionadas. Sin embargo, desde el punto de vista del usuario, todas ellas pertenecen al mismo proyecto. Entonces, cuando el usuario quiere iniciar un proyecto, llamamos al cliente y proporcionamos la siguiente información:

- Una cadena de caracteres que contenga los metadatos para identificar los archivos de entrada.
- Una tarball que contiene el código del análisis.
- El nombre del script dentro de la tarball que debe ejecutarse.
- La versión del software de Belle II que debe ser usado.

- El número de copias deseadas, o alternativamente el número deseado de archivos de entrada por copia.

Todas las copias se mandan a un sitio, o también pueden repartirse en varios sitios. Opcionalmente, el usuario puede proporcionar una lista de sus sitios preferidos.

### 4.3. Producción de Monte-Carlo

La simulación de Monte-Carlo (MC) es el aspecto más intensivo en términos de CPU y el menos intensivo en términos de datos de los requisitos computacionales de Belle II. Este es un método numérico que se usa para evaluar expresiones que no se pueden resolver analíticamente, o cuya solución analítica necesita muchos recursos computacionales. Se basa en la generación de números aleatorios y su principal uso es en la evaluación de integrales. El usuario solicita producción de Monte Carlo de un modelo físico en particular, un rango en el número de corridas, número de eventos, etc. El cliente del proyecto traduce esta petición a una colección física de recursos y manda el número de copias solicitadas. En el modelo estadístico, a cada copia se le asignan parámetros fijos para la simulación (experimento y número de corrida, número de eventos por corrida). En el modelo dinámico, las copias reciben esta información al tiempo de la corrida desde el servidor del proyecto. Un requisito adicional para la producción de Monte-Carlo es que se proporcionen los Random overlay data files adecuados. Estos son datos reales recolectados junto con eventos físicos pero almacenados con un trigger aleatorio y sin software de filtrado. Esos datos fueron usados exitosamente para simular ruido debido al haz para el experimento Belle; se planea usar esta técnica en Belle II.

### 4.4. Simulación

La simulación de muestras de Monte-Carlo de un modelo de decaimiento se realiza en tres pasos: generación de eventos, simulación del detector, y digitalización.

#### 4.4.1. Generador de eventos

El framework BASF2 usa el módulo EvtGen para generar eventos del modelo físico de interés para el usuario. EvtGen es un generador de Monte-Carlo de eventos físicos diseñado para la simulación de la física de decaimientos de mesones  $B$ . EvtGen proporciona un framework para manipular decaimientos secuenciales complejos. Implementa muchos modelos de decaimiento incluyendo modelos semileptónicos, decaimientos hadrónicos, decaimientos con violación de CP, distribuciones angulares y decaimientos de Dalitz.

BASF2 usa EvtGen para generar los mesones  $B$  que se producen de las resonancias  $\Upsilon(nS)$ . El algoritmo de un decaimiento en EvtGen es el siguiente:

- La entrada es la partícula padre con (PDG) ID y 4-momento conocidos.
- Selecciona el modo de decaimiento basado en los modos de decaimiento permitidos para partículas específicas. En nuestro caso la entrada completa está basada en un archivo DECAY.DEC. Las partículas que no se listan en el archivo de entrada DECAY.DEC se suponen estables.
- La cadena de decaimiento continúa hasta que todas las partículas que restan son partículas estables.
- Determina la masa de cada partícula.
- Determina la cinemática del decaimiento. Antes de determinar la cinemática, la cadena de decaimiento debe fijarse, incluyendo la masa de las partículas [17].

A partir del último punto se puede extraer un archivo con extensión de archivo root para hacer un análisis directo o conservarlos para su posterior uso en el módulo de simulación.

Cabe mencionar que es posible introducir archivos con extensión DEC que sobreescriben los decaimientos, esto con el fin de modificar las razones de decaimiento o buscar decaimientos no listados en el PDG.

#### 4.4.2. Simulación del detector

Otro de los módulos que usa BASF2 es Geant4. Geant4 se orienta a objetos y es una herramienta para simular el paso de partículas a través de materia. Incluye un amplio rango de funcionalidades entre las cuales están simular la trayectorias, la geometría del detector, modelos físicos e impactos.

Las partes fundamentales de la simulación del paso de partículas a través de materia son:

- Geometría y materiales.
- Interacción de las partículas con la materia.
- Manejos de eventos y trayectorias de las partículas en el material.
- Manejo de las señales dejadas en el detector.
- Visualización.
- Interfaz de usuario.

Esta herramienta permite al usuario crear un modelo geométrico con un gran número de componentes de diferentes formas y materiales, y definir elementos que registren la información necesaria para simular la respuesta del detector (digitalización). Adicionalmente, el usuario puede usar una interfaz gráfica y visualizar la geometría y las trayectorias. Geant4 tiene implementada una amplia variedad de procesos físicos para modelar el comportamiento de partículas y sus interacciones con la materia [18].

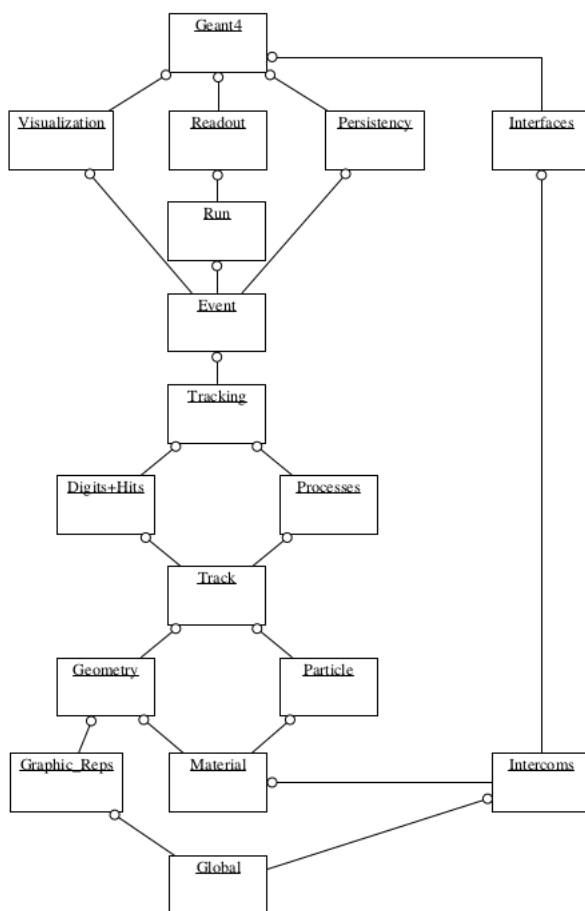


Figura 4.3: Diagrama de categorías por nivel de Geant4.

### 4.4.3. Digitalización

Los impactos en el detector simulados por Geant4 se convierten en señales que el detector mide via un proceso de digitalización. El resultado que se obtiene del proceso de digitalización corresponde a los datos reales sin analizar. Para simular los procesos debemos aprovechar toda la información que tenemos de los materiales con los que se construyeron los detectores y de la electrónica que se usa. Expertos de cada componente del detector colaboran para crear las librerías, en  $C++$ , que lleva acabo el proceso de digitalización [19].

## 4.5. Análisis

La rapidez de almacenamiento que se utilizará en Belle II es de  $1.5 \text{ GB/s}$ , aproximadamente, de lo cual es claro que la cantidad de datos que se recolectaran es enorme. Desde hace ya 20 años los físicos de partículas utilizan una herramienta de licencia abierta llamada ROOT que fue desarrollada en cooperación entre el CERN y Fermilab para manipular una gran cantidad de datos de una forma bastante eficiente. Como framework, ROOT está diseñado específicamente para realizar análisis de datos a gran escala. El uso de esta herramienta en física de partículas es fundamental pues permite manipular cantidades enormes de datos de manera fácil y versátil. ROOT está escrito en  $C++$ , y está orientado a objetos. En cuanto al almacenamiento, ROOT guarda los datos de una forma muy eficiente: utiliza una base de datos jerárquica orientada a objetos. Esta base de datos es independiente de la máquina y está altamente comprimida. ROOT también tiene herramientas para interactuar de manera eficiente con los datos. Se construyó para ser capaz de crear histogramas multidimensionales de alta calidad, ajustar curvas, modelar y simular. Todas estas herramientas están diseñadas para manejar grandes volúmenes de datos [21] [20].



# Capítulo 5

## Técnicas multivariadas

En física de altas energías usualmente buscamos señales pequeñas de algún proceso de interés en un conjunto de datos bastante grande, por lo que es esencial extraer el máximo de información de los datos disponibles. Por lo anterior, los métodos de clasificación multivariada, basados en técnicas de máquinas de aprendizaje, son una herramienta fundamental para realizar análisis. Esto significa, grosso modo, el reemplazo de suposiciones y aproximaciones por poder computacional como método para obtener información característica acerca de un conjunto de datos. Entre los métodos que se han desarrollado se encuentran los árboles de decisión (decision trees), redes neuronales (neural nets), estimación basada en el uso de verosimilitudes (projective likelihood estimation), curvas de aprendizaje (learning curves), entre otras.

En particular, es de nuestro interés utilizar el paquete TMVA (Toolkit for Multivariate Analysis) el cual proporciona un ambiente de aprendizaje automatizado integrado a ROOT para el procesamiento y evaluación de forma paralela de técnicas de clasificación y regresión multivariada. El paquete TMVA está diseñado específicamente para cubrir las necesidades de la física de altas energías; sin embargo, su uso no se limita a eso.



## 5.1. Árboles de decisión

Un árbol de decisión es un árbol binario clasificador estructurado de forma similar al que se muestra en la figura 5.1. Se toman decisiones repetidas de izquierda a derecha en una variable a la vez y concluye cuando el criterio de optimización impuesto se cumple. De esta forma, el espacio fase se divide en en varias regiones que eventualmente son clasificadas como señal o ruido, dependiendo de la mayoría de eventos de entrenamiento que terminen en el nodo final.

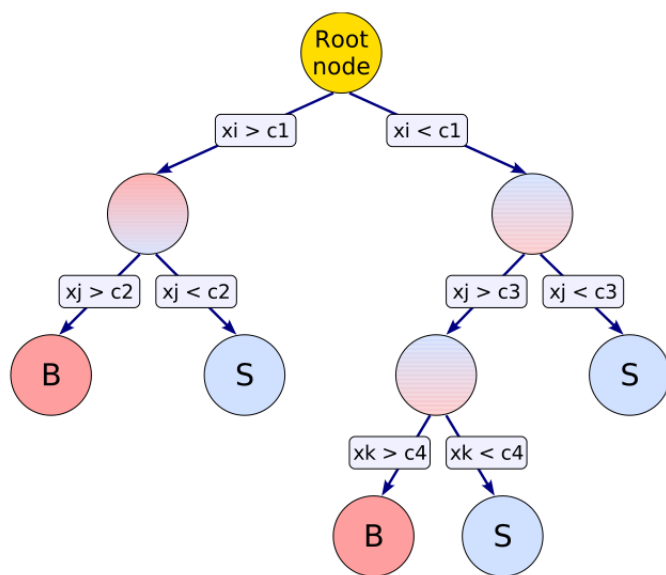


Figura 5.1: Esquema de un árbol de decisión.

### 5.1.1. Construcción de un árbol

En la construcción de un árbol de decisión, tomamos una muestra con  $N$  eventos  $x_n$ , cada uno de los cuales está descrito por  $D$  variables de entrada. En el nodo raíz del árbol,  $O$ , se coloca toda la muestra. En seguida se parte el nodo raíz en dos subconjuntos: uno izquierdo,  $L$ , y uno derecho,  $R$ ; esto mediante la asignación de cada uno de los eventos a uno y sólo uno de los subconjuntos. Se aplica el mismo proceso de partición a cada uno de los nuevos nodos. El proceso de partición de los

nodos continúa hasta que el criterio de optimización, impuesto, se satisfaga.

La pregunta que surge es: ¿cuál es un buen criterio de partición? Antes de tratar de responder consideremos la siguiente notación: sea  $t_O$  un evento en el nodo raíz,  $t_L$  y  $t_R$  eventos en los nodos izquierdo y derecho, respectivamente. También,  $N_0$ ,  $N_L$ , y  $N_R$ , el número de eventos en el los nodos raíz, izquierdo y derecho, respectivamente. La probabilidad de que cierto evento esté en uno de los nodos está dada por:  $P(t) = N_t/N$  con  $t \in \{O, L, R\}$ . De aquí en adelante  $N$  es el número total de eventos en la muestra que usamos para entrenar el árbol. Es claro que el nodo raíz contiene todos los eventos, y la probabilidad de que uno de ellos esté en tal nodo es uno.

Ahora, consideremos el caso en el que existen dos clases de eventos en la muestra, la clase  $A$  y la clase  $B$  (por ejemplo señal y ruido). Si el nodo raíz contiene  $N_A$  eventos de la clase  $A$  y  $N_B$  eventos de la clase  $B$ , la probabilidad de que uno de los eventos en el nodo raíz sea de la clase  $A$  es:  $P(A|t) = N_A/N_t$ , análogamente para los eventos de clase  $B$ . Claramente la suma de estas probabilidades debe ser igual a uno. Un nodo del árbol predice la clase con mayor probabilidad posterior, y su error en el entrenamiento es  $\epsilon(t) = \min_{Y \in \{A, B\}} P(Y|t)$ .

Si un nodo del árbol contiene sólo eventos de una clase, se le llama *puro* y su error de entrenamiento es cero. Al realizar un análisis lo ideal sería tener árboles con nodos puros, pues la separación sería totalmente confiable. Sin embargo, en ocasiones esto no es posible debido a que la distribuciones de las clases pueden estar superpuestas. Para seleccionar un buen criterio de partición, se requiere de una medida de la impureza.

Se define la *impureza de un nodo* como una función de las probabilidades de las clases:

$$i(t) = \phi(P(A|t), P(B|t)), \quad (5.1)$$

podemos poner cotas a los valores de esta función,  $0 \leq \phi(p, q) \leq 1/2$ . La impureza es máxima cuando las dos clases están mezcladas en igual proporción y mínima cuando un nodo es puro. Esperamos que la medida de la impureza de una mezcla con dos

clases distintas sea simétrica. De lo anterior, es posible pedir que  $\phi(1, 0) = \phi(0, 1) = 0$  y  $\phi(1/2, 1/2) = 1/2$ .

Una buena condición para llevar a cabo la separación debe minimizar la impureza del nodo. Una separación binaria produce dos nodos, por lo que deberíamos minimizar el promedio de las impurezas. Los nodos pueden diferir en tamaño, así que para promediar podemos usar las probabilidades de los nodos como peso para sus impurezas.

Se define la impureza pesada del nodo como,

$$I(t) = P(t)i(t), \quad (5.2)$$

entonces, la impureza promedio después de la separación se puede escribir como  $I(t_L) + I(t_R)$ . Finalmente, se define la ganancia en la impureza,

$$\Delta I = I(t_0) - I(t_L) - I(t_R). \quad (5.3)$$

La mejor regla de separación puede hallarse maximizando la ganancia en la impureza  $\Delta I$  sobre todas las posibles separaciones para todas las variables. Si no hay divisiones con ganancia positiva, el nodo no puede dividirse y es un nodo *terminal*.

La pieza faltante es dar la forma específica de la función de impureza  $\phi$ . Una elección simple es una función cuadrática llamada *Gini diversity index*:

$$\phi(p, q) = 1 - p^2 - q^2. \quad (5.4)$$

Otra elección usual es la llamada *entropía cruzada*:

$$\phi(p, q) = -\frac{p \log_2 p + q \log_2 q}{2}. \quad (5.5)$$

En la práctica, usualmente elegir una regla de detención es más importante que elegir una medida de la impureza.

### 5.1.2. Reglas de detención

Una lista de criterios de detención encontrados en la literatura se muestra en seguida. Un nodo de un árbol no puede dividirse si [22] :

- El nodo es puro, esto es, todos los eventos pertenecen a una de las clases.
- El tamaño del nodo es menor que el tamaño mínimo permitido del padre nodo.
- La profundidad máxima del árbol se ha alcanzado.
- El número máximo de nodos terminales se ha alcanzado.
- La máxima ganancia en la impureza es menor a cierto umbral.

por tamaño del nodo se entiende número de eventos en tal nodo.

### 5.1.3. Árboles de decisión impulsados (BDT) en el paquete TMVA

Un árbol de decisión impulsado (BDT) es una generalización de un árbol de decisión a varios árboles que forman un bosque [23]. En el paquete *TMVA* integrado en el framework de análisis *ROOT* podemos utilizar la herramienta BDT mediante el comando:

```
factory → BookMethod(Types :: kBDT, "BDT", "⟨options⟩");
```

en donde el primer argumento está predefinido, el siguiente es un identificador dado por el usuario, y finalmente las opciones de configuración que se pueden consultar en [23].

Al proceso mediante el cual se define un criterio para llevar a cabo la separación de una muestra se le conoce como entrenamiento (construcción o crecimiento), ya hemos hablado de esto con anterioridad. En el uso del paquete *TMVA*, un nodo deja de separarse una vez que se obtiene el número mínimo de eventos que el usuario especifica como opción en la configuración del BDT (*nEventsMin*). En un BDT, los

nodos finales se clasifican en señal o ruido según la mayoría de eventos pertenezca a una clase u otra.

A la hora de implementar un BDT es posible configurar un criterio de separación (`SeparationType`). Se han realizado pruebas en los siguientes criterios de separación, definidos en el paquete, y no se han encontrado diferencias significantes [23]:

- Gini Index (usada por defecto).
- Entropía cruzada.
- Significancia estadística.
- Error en la clasificación.
- Error cuadrático medio.

En teoría se puede continuar haciendo cortes hasta alcanzar nodos finales con pura señal/ruido, sin embargo el árbol estaría sobreentrenado. Para evitar esto debemos podar el árbol. El proceso de podar tiene como fin quitar nodos con insignificancia estadística y con esto evitar árboles sobreentrenados. Sin embargo, aunque ya hay algoritmos para podar implementados en TMVA, estos ya son obsoletos y es preferible no hacer uso de ellos. Debido a la simplicidad del método y los buenos resultados que se obtienen, los árboles de decisión son una de las herramientas más usadas cuando se estudia un problema de clasificación.

## 5.2. Redes neuronales

La implementación de las redes neuronales (NN) nació como un intento de modelar la inteligencia biológica. La aplicación de redes neuronales a problemas de clasificación ha tenido éxito, por ejemplo en física de altas energías fue de los primeros métodos de aprendizaje automatizado empleados para identificar partículas,

reconstruir eventos y optimizar análisis de datos. Una red neuronal la podemos entender como la simulación de un conjunto de neuronas conectadas, las cuales pueden responder a cierta información de entrada. En otros términos, una red neuronal se puede entender como un mapeo del espacio de las variables de entrada  $x_1, x_2, \dots, x_n$  en el espacio de las variables de salida  $y_1, y_2, \dots, y_m$ . Este mapeo puede no ser lineal si al menos una de las neuronas tiene una respuesta no lineal.

En el paquete TMVA se pueden hallar tres implementaciones de redes neuronales. Sin embargo, las tres implementaciones están basadas en técnicas de perceptrones multicapa.

### 5.2.1. Redes neuronales en el paquete TMVA

Como ya se ha mencionado con anterioridad, en el paquete TMVA podemos encontrar tres implementaciones de redes neuronales.

Red neuronal de Clermont-Ferrand, esta se puede llamar mediante el comando:

```
factory → BookMethod(Types :: kCFMlpANN, "CF_ANN", "⟨options⟩");
```

Otra red neuronal que se puede emplear es la de ROOT, esta es una interfaz de la clase *TMultiLayerPerceptron* y se puede llamar mediante el comando:

```
factory → BookMethod(Types :: kTMlpANN, "TMlp_ANN", "⟨options⟩");
```

Finalmente, la red neuronal MLP se llama mediante el comando:

```
factory → BookMethod(Types :: kMLP, "MLP_ANN", "⟨options⟩");
```

Las opciones de configuración de estas tres opciones se pueden consultar en [23].

### 5.2.2. Perceptrones multicapa

Un perceptrón es una forma de modelar una neuróna, que tiene la tarea de seleccionar entre dos clases  $C_1$  y  $C_2$ , en base a la información de entrada [22]. La técnica de perceptrones multicapa consiste en ordenar perceptrones (neuronas) en capas, cada perceptrón puede conectarse directamente con perceptrones de la siguiente capa. La forma en que una red neuronal se comporta está determinada por la configuración

de la neurona, los pesos dados a las conexiones neurona-neurona, y la respuesta que tienen las neuronas a la información de entrada, dada por la función de respuesta  $\rho$ . En un arreglo multicapa, la primera capa es la capa de entrada (input layer), la última capa es la capa de salida (output layer), y las capas intermedias se conocen como capas ocultas. En el caso de un problema de clasificación con  $n$  variables de entrada la capa de entrada consta de  $n$  neuronas las cuales reciben los  $n$  valores,  $x_1, x_2, \dots, x_n$  y la capa de salida tiene una neurona que recibe a la variable de salida, el estimador de la red neuronal,  $y_{ANN}$ . Lo anterior se ejemplifica en la imagen 5.2. En la imagen 5.2,  $y_m^l$  es la etiqueta de la neurona que se encuentra en la capa  $l$ , en

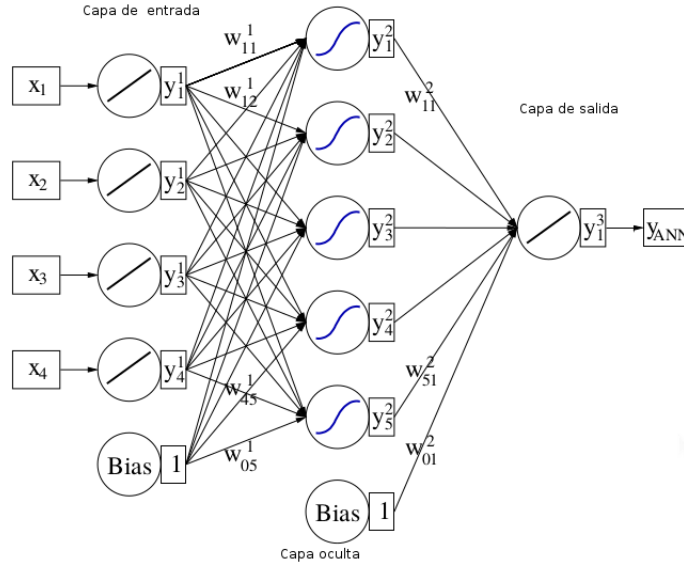


Figura 5.2: Esquema de un arreglo multicapa con una capa oculta.

la posición  $m$ ;  $w_{mn}^l$  etiqueta al peso correspondiente a la conexión de la neurona  $y_m^l$  con la  $y_n^{l+1}$ . En el paquete *TMVA* es posible configurar el número de capas ocultas y el número de neuronas en esas capas mediante la opción *HiddenLayers*.

Usualmente, la función de respuesta de una neurona puede separarse en una función de sinapsis:  $\kappa : \mathbb{R}^n \rightarrow \mathbb{R}$ , y una función de activación de la neurona:  $\alpha : \mathbb{R} \rightarrow \mathbb{R}$ . De tal forma que podemos escribir la función de respuesta como la composición:  $\rho = \alpha \circ \kappa$ . Las formas que pueden tener las funciones  $\kappa$  y  $\alpha$  pueden consultarse en [23].

Se han realizado pruebas y según los resultados obtenidos, la red *MLP* y la implementada en *ROOT* tienen un desempeño bastante similar, teniendo la red *MLP* ventaja en cuanto a rapidez. Por otra parte, la red de Clermont-Ferrand no es eficiente con los problemas de clasificación.

### 5.3. Método de estimación por verosimilitud.

Otro método que resulta de utilidad cuando se busca la presencia de una señal dentro de un conjunto de datos es el uso del método llamado razón de verosimilitud (likelihood ratios).

Podemos contruir un modelo basado en densidades de probabilidad que reproduzca las variables de entrada para la señal y para el ruido, a este método se le conoce como método de máxima verosimilitud. La verosimilitud de que un evento sea señal se obtiene multiplicando las densidades de probabilidad de todas las variables de entrada, que se suponen independientes, y normalizando este producto con la suma de las verosimilitudes de la señal y ruido. Esto es, para el evento  $i$ , se define la razón de verosimilitud como:

$$y_L(i) = \frac{L_S(i)}{L_S(i) + L_B(i)}, \quad (5.6)$$

donde

$$L_{S(B)}(i) = \prod_{k=1}^n p_{S(B)k}(x_k(i)), \quad (5.7)$$

con  $p_{S(B)k}$  la distribución de probabilidad de señal (ruido) para la  $k$ -ésima variable de entrada  $x_k$ . Debemos tener en mente que las distribuciones de probabilidad deben normalizarse:

$$\int_{-\infty}^{\infty} p_{S(B)k}(x_k) dx_k = 1, \quad \forall k. \quad (5.8)$$

La forma de las distribuciones de probabilidad se hallan de forma empírica a partir del conjunto de datos que se usan para el entrenamiento de este método.



Entre las ventajas que tiene este método está el hecho de que el entrenamiento y la aplicación de este método de clasificación son bastante rápidos lo cual lo hace un método ideal para el estudio de un conjunto de datos grande.

En el paquete *TMVA* podemos implementar este método de clasificación mediante el comando:

```
factory → BookMethod(Types :: kLikelihood, "Likelihood", "⟨options⟩");
```

Las opciones de configuración de este clasificador pueden encontrar en [23].

# Capítulo 6

## Análisis

Para realizar este trabajo se utilizó una muestra de 2.25 millones de eventos del canal de decaimiento  $B^0 \rightarrow D^- \pi^- \mu^+ \mu^+$  generada por el método de Monte Carlo mediante el uso del framework *Basf2*. La versión que se utilizó es la *release* – 00 – 08 – 00 instalada en el centro de computo *JAGUAR* en el departamento de Física del CINVESTAV, que corre las versiones 6.06/08 y 3.5.2 de *ROOT* y *Python*, respectivamente.

La muestra mencionada la dividimos en dos partes, una de medio millón la cual utilizaremos para entrenar y probar los métodos de clasificación multivariable que aquí emplearemos; la segunda muestra la utilizamos como nuestra muestra de datos independientes a analizar.

### 6.1. Generación de la muestra de Monte Carlo

Anteriormente ya se describió la forma en la que se realizan la generación de eventos y la simulación del detector. Sin embargo, es importante mencionar que debido a que el canal de decaimiento  $B^0 \rightarrow D^- \pi^- \mu^+ \mu^+$  no se ha observado experimentalmente debemos escribir explícitamente un archivo DEC en donde se especifique de forma clara cual es la cadena de decaimiento que se quiere generar, así como el modelo de

decaimiento que se desea usar<sup>1</sup>.

## 6.2. Reconstrucción

Teniendo la muestra de eventos generados podemos pasar a realizar la simulación de la cadena de decaimiento dentro del detector para así obtener la cinemática de las partículas de estado final y, posteriormente, continuar con la reconstrucción. Una vez más el framework *Basf2* nos ayuda a realizar este trabajo. Queremos reconstruir el canal:  $B^0 \rightarrow D^- \pi^- \mu^+ \mu^+$ ; para esto, primero buscamos reconstruir el meson  $D^-$  mediante el canal  $D^- \rightarrow K^+ \pi^- \pi^-$ . Hecho lo anterior buscamos identificar un vertice en el que coincidan un mesón  $D^-$ , un mesón  $\pi^-$  y dos antimuones  $\mu^+$ , con lo que se completa la cadena de decaimiento que buscamos. La ventana de masa que utilizamos para el mesón  $B^0$  es de  $5.2 \text{ GeV}$  a  $5.4 \text{ GeV}$ .

Una vez reconstruido el decaimiento podemos, por ejemplo, hacer un histograma de la masa invariante del meson  $D^-$  y comparar el valor central que obtenemos con el valor reportado por el PDG. Sin embargo, la reconstrucción no es perfecta pues el framework tomará en cuenta diferentes posibilidades que no corresponden al canal reconstruido correctamente y que se pueden confundir con este. Para solucionar este inconveniente podemos apoyarnos en el hecho de que en este punto estamos usando una muestra que fue generada mediante el método de Monte Carlo y que tiene la ventaja de que nos permite saber si cierto evento fue reconstruido correctamente. Apoyados en lo anterior, podemos buscar cortes en ciertas variables cinemáticas que nos permitan separar la señal de un evento correctamente reconstruido del ruido combinatorio que tenemos en nuestra reconstrucción.

Una de las variables cinemáticas que podemos usar para quitar parte del ruido es la magnitud del momento. Lo que aquí hicimos fue hacer dos histogramas del momento superpuestos para cada una de las partículas de esto final. Uno de los histogramas corresponde al momento de la partícula reconstruida sin poner restricción

---

<sup>1</sup>Ver Apéndice B

alguna, y el otro histograma se realizó bajo la condición de que la reconstrucción del canal sea correcta. Con ésto podemos ver a partir de que valor del momento para la partícula en cuestión podemos asumir que una buena parte de los datos corresponde a señal y no a ruido. Por ejemplo, de la figura (6.1) (a), observamos que si pedimos que el momento asociado al kaón que proviene del mesón  $D^-$  sea mayor que  $0.35 GeV$  estamos quitando eventos que no fueron correctamente reconstruidos; sin embargo aún no estamos quitando todo el ruido; de forma análoga lo hacemos para el resto de las partículas que buscamos en el estado final. Los valores para los cortes que utilizamos en la variable de momento se muestran en los histogramas de las figuras (6.1) y (6.2). Estos histogramas están normalizados, lo que nos permite ver de forma gráfica en que valor podemos hacer un corte de tal forma que podamos distinguir entre señal y ruido. Evidentemente, al realizar un corte estamos perdiendo eventos que son señal; también nos estamos quedando con eventos que son parte del ruido producido en la reconstrucción.

De forma análoga, buscamos cortes en una segunda variable, a saber, utilizamos la probabilidad de identificación de cada partícula, PID. Esta variable nos da la probabilidad de confundir una partícula en el estado final<sup>2</sup> con otra partícula diferente, por ejemplo, la probabilidad de que el detector confunda un electrón con un muón . En este caso, comparamos la probabilidad de identificar correctamente un kaón sin ninguna restricción con la probabilidad de identificar el mismo kaón bajo la condición de que la reconstrucción del canal sea correcta. Lo mismo realizamos para el resto de partículas de estado final, que en nuestro caso son: kaones, piones, y antimuones. Los valores para los cortes que utilizamos en la variable PID se muestran en los histogramas de las figuras (6.3) y (6.4).

Utilizando estos cortes en la muestra generada de 1.75 millones de eventos, obtenemos los histogramas para las masa invariantes de los mesones  $D^-$  y  $B^0$ , (6.5) y (6.6), respectivamente.

---

<sup>2</sup>Entre las partículas de estado final que el detector puede identificar están: electrones, kaones, piones, muones y protones, y sus correspondientes antipartículas.

Podemos observar en las figuras (6.5) y (6.6) que aún después de realizar cortes no hemos conseguido limpiar completamente las señales asociadas a los mesones  $D^-$  y  $B^0$ . Nos vemos en la necesidad de limpiar aún más nuestra muestra de datos. Por esta razón requerimos utilizar técnicas de análisis multivariable.

### 6.2.1. Implementación del paquete TMVA

Ya describimos brevemente las características del paquete *TMVA* (Toolkit for Multivariate Data Analysis) de *ROOT*, a saber, este nos provee de un ambiente de máquinas de aprendizaje integrado a *ROOT* para la ejecución y evaluación en paralelo de técnicas de clasificación multivariable. En este trabajo se emplearon tres métodos: árboles de decisión estimulados (de ahora en adelante BDT), redes neuronales artificiales (NN) y estimación basada en el uso de verosimilitudes (projective likelihood estimation).

Para entrenar y probar estos métodos de clasificación usamos una muestra de medio millón de eventos. Las variables que utilizamos para el entrenamiento de estos clasificadores son:

- Los PID de cada una de las partículas finales, 6 en total.
- Los momentos transversales de las 6 partículas finales.
- Los momentos transversales de los mesones  $D^-$  y  $B^0$ .

Para discriminar entre señal y ruido utilizamos el código del PDG asignado al mesón  $B^0$ . Una vez concluido el entrenamiento, el resultado es un conjunto de archivos que contienen los pesos asociados a cada uno de los métodos. Estos pesos los utilizamos para analizar nuestra muestra de datos independientes de 1.75 millones de eventos.

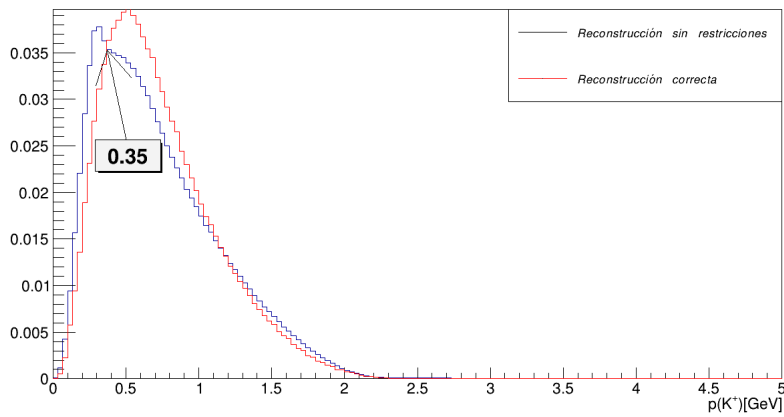
Después del entrenamiento, y usando los pesos hallados en el paso anterior cada uno de los métodos de clasificación asigna un valor, aquí llamado  $\alpha$ , a cada evento de la muestra. Para el método BDT tenemos  $\alpha \in [-1, 1]$ ; valores tendiendo a  $-1$  se

asocian a eventos de ruido y valores tendiendo a 1 se asocia con eventos de señal. Para el método NN tenemos  $\alpha \in [0, 1]$ ; valores cercanos a 0 se asocian a eventos de ruido y valores tendiendo a 1 se asocia con eventos de señal. Análogamente al método NN, para el método de verosimilitud tenemos  $\alpha \in [0, 1]$ ; valores cercanos a 0 se asocian a eventos de ruido y valores tendiendo a 1 se asocia con eventos de señal.

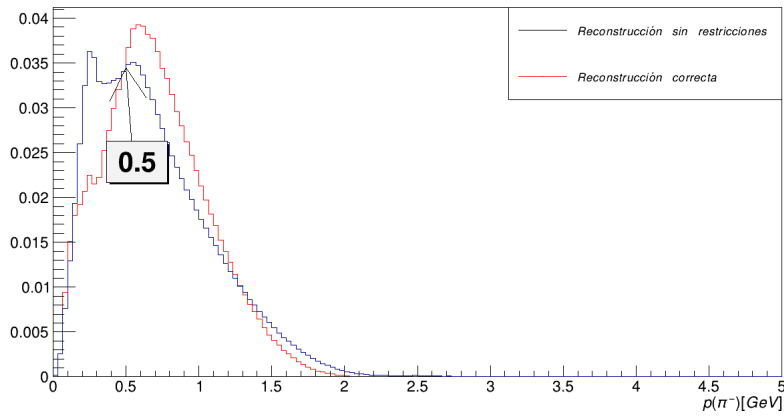
Ahora, tenemos una variable más asignada a cada evento por cada método y es de nuestro interes optimizar esta variable de forma que podamos separar la mayor cantidad posible de eventos de señal. Lo que en este trabajo se realizó es ver el comportamiento de la significancia  $S = N_S/\sqrt{N_S + N_B}$  como función de la nueva variable, donde  $N_S$  es el número de eventos de señal y  $N_B$  es el número de eventos de ruido. Así pues, el comportamiento encontrado se muestra en las figura 6.7.

Gráficamente observamos que para el método BDT la significancia tiene un punto máximo en algún punto tal que  $\alpha \in (0.075, 0.125)$ . De forma análoga, para el método NN la significancia tiene un máximo en algún punto tal que  $\alpha \in (0.775, 0.825)$ . Finalmente, para el método de verosimilitud, la significancia tiene un máximo en una vecindad de  $\alpha = 0.925$ .

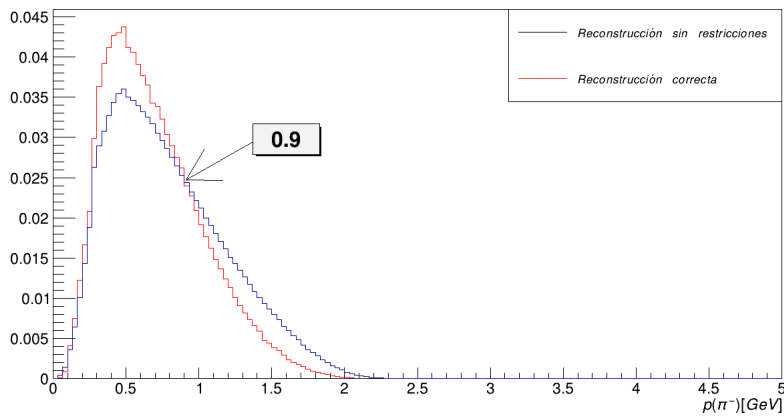
Con esta información extra podemos realizar un corte más en la variable  $\alpha$  limitandola a los intervalos que mencionamos en el parrafo anterior, con lo cual logramos limpiar más la señal de los eventos que buscamos. Una vez realizados los cortes en las variables de momento, PID y  $\alpha$  guardamos en un archivo nuevo los eventos que han pasado los cortes y realizamos un ajuste, por cada método, a la masa invariante de los candidatos a mesones  $B^0$  para lo cual hacemos uso de RooFit. Este ajuste nos permite estimar el número de eventos de señal y ruido. En los ajustes realizados se utilizó un modelo de masa gaussiano para la señal y un modelo lineal para el ruido combinatorio. Los ajustes mencionados se muestran en la figura (6.8). En las leyendas,  $N_s$  es el número de eventos de señal estimados y  $N_b$  el número de eventos de ruido,  $a0$  es el coeficiente del polinomio usado en el modelo de ruido,  $Mean$  es el valor central obtenido para la masa del  $B^0$ , y  $RMS$  es el valor de la desviación estándar.



(a) Histograma del momento asociado al kaón proveniente de un mesón  $D^-$ .

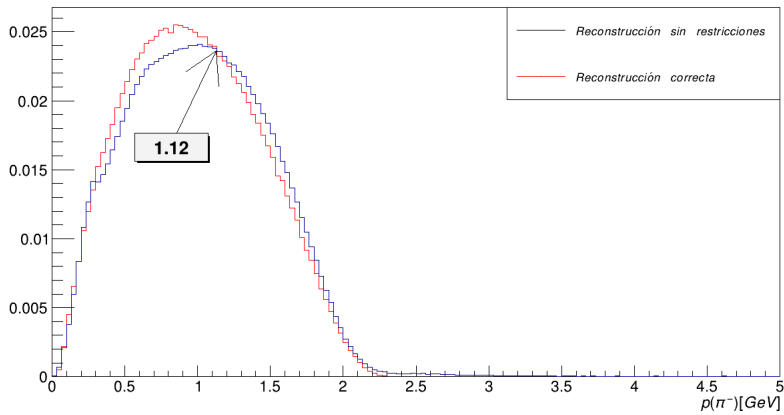
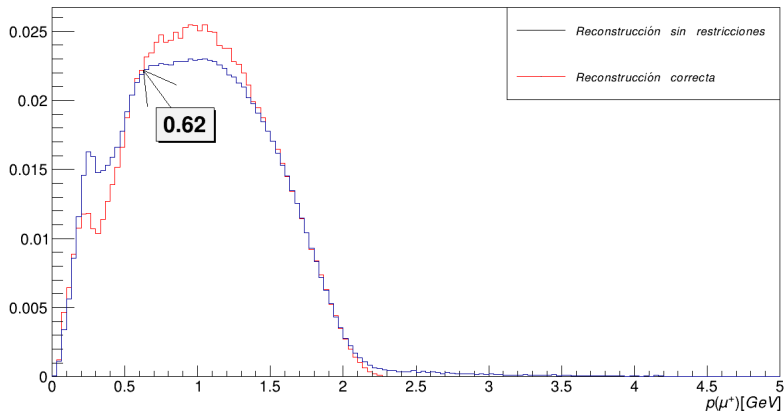
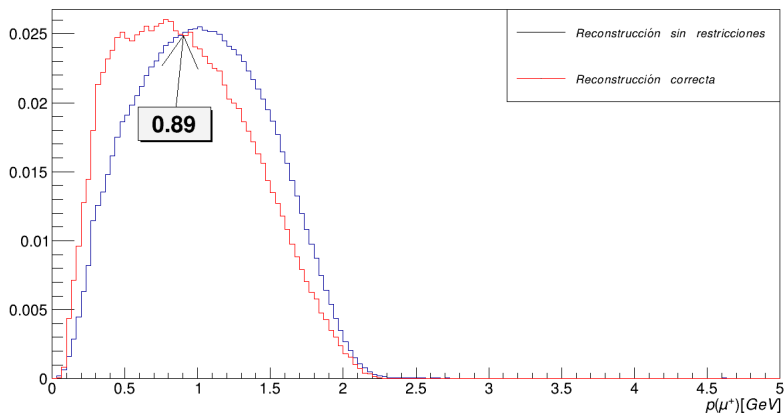


(b) Histograma del momento asociado al primer pión proveniente de un mesón  $D^-$ .

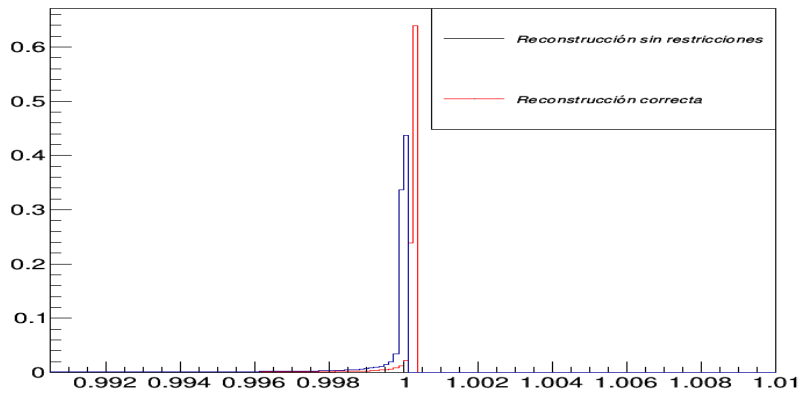


(c) Histograma del momento asociado al segundo pión proveniente de un mesón  $D^-$ .

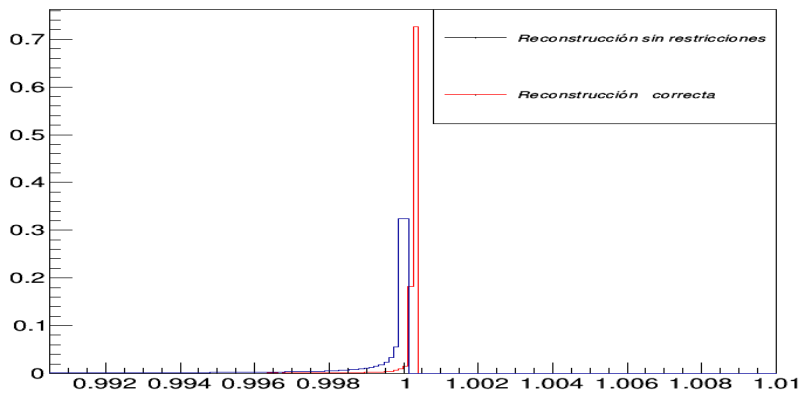
Figura 6.1: Cortes en la variable de momento para identificar a las partículas de estado final candidatas a provenir del decaimiento  $D^- \rightarrow K^+ \pi^- \pi^-$ .

(a) Histograma del momento asociado al pión proveniente de un mesón  $B^0$ .(b) Histograma del momento asociado al primer antimuón proveniente de un mesón  $B^0$ .(c) Histograma del momento asociado al segundo antimuón proveniente de un mesón  $B^0$ .Figura 6.2: Cortes en la variable de momento para identificar a las partículas de estado final candidatas a provenir del decaimiento directamente del mesón  $B^0$ .

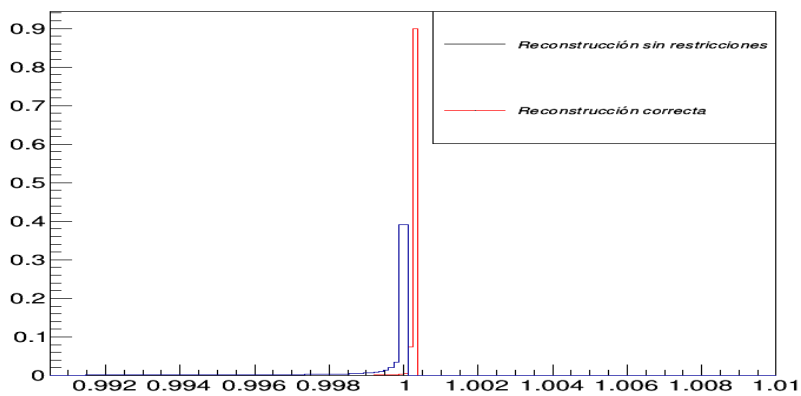




(a) Histograma de la variable PID asociada al kaón proveniente de un mesón  $D^-$ .

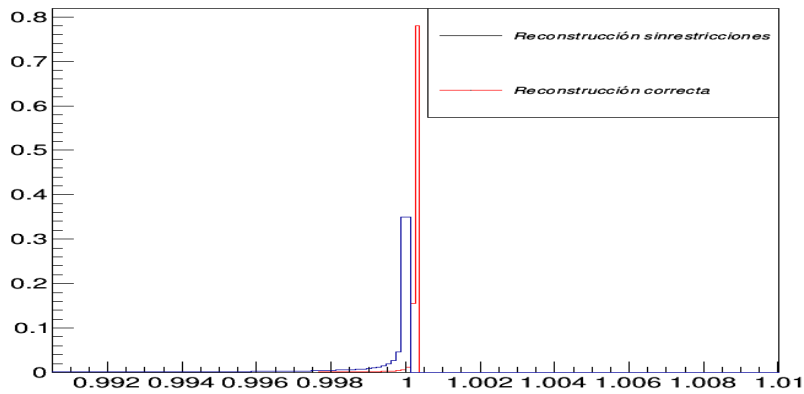


(b) Histograma de la variable PID asociada al primer pión proveniente de un mesón  $D^-$ .

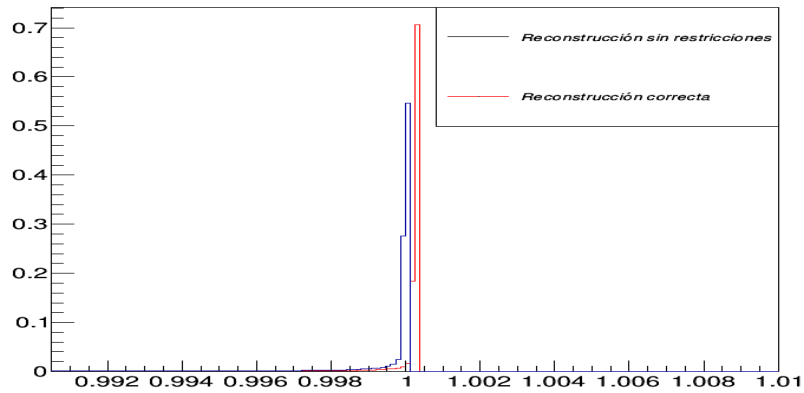


(c) Histograma de la variable PID asociada al segundo pión proveniente de un mesón  $D^-$ .

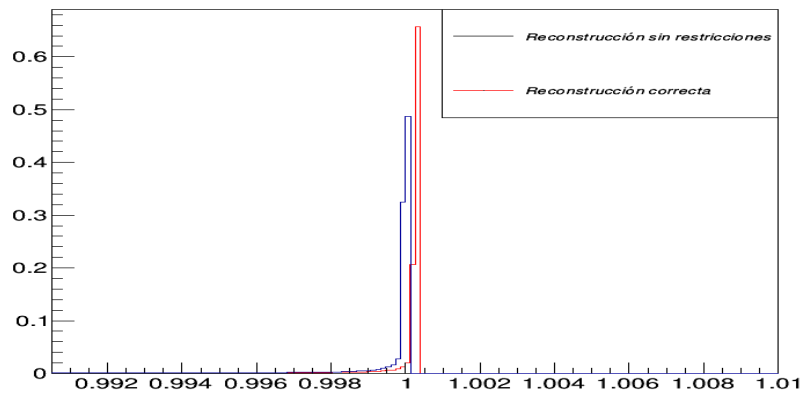
Figura 6.3: Cortes en la variable PID para identificar a las partículas de estado final candidatas a provenir del decaimiento  $D^- \rightarrow K^+ \pi^- \pi^-$ .



(a) Histograma de la variable PID asociada al pión proveniente de un mesón  $B^0$ .



(b) Histograma de la variable PID asociada al primer antimuón proveniente de un mesón  $B^0$ .



(c) Histograma de la variable PID asociada al segundo antimuón proveniente de un mesón  $B^0$ .

Figura 6.4: Cortes en la variable PID para identificar a las partículas de estado final candidatas a provenir del decaimiento directamente del mesón  $B^0$ .

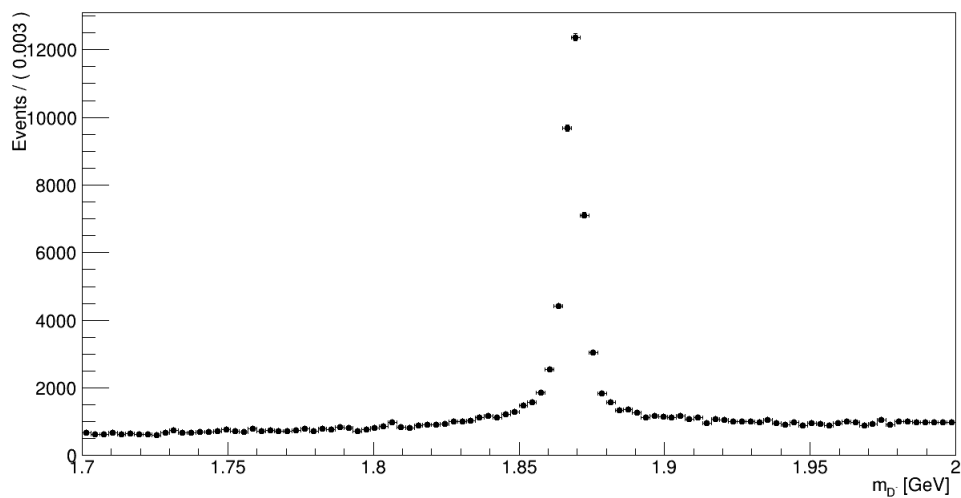


Figura 6.5: Histograma de la masa invariante del mesón  $D^-$  después de aplicar cortes en el momento y en la variable PID.

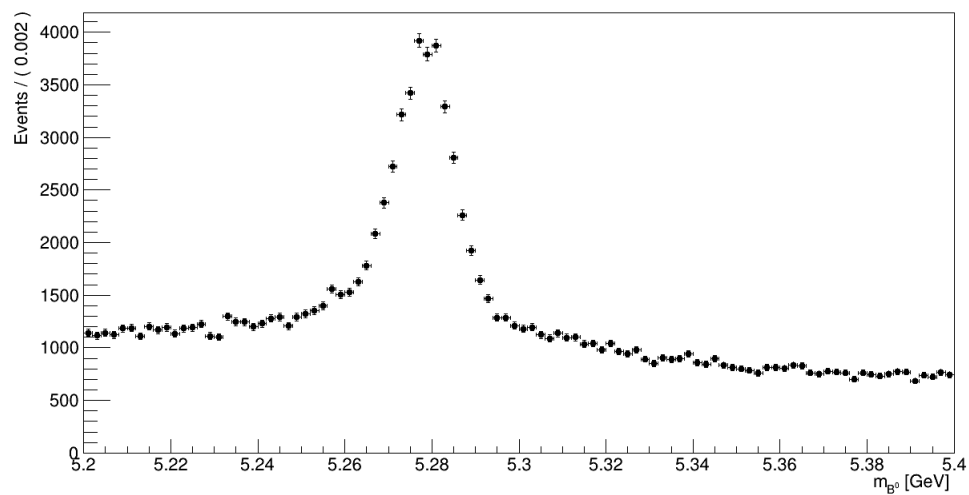


Figura 6.6: Histograma de la masa invariante del mesón  $B^0$  después de aplicar cortes en el momento y en la variable PID.

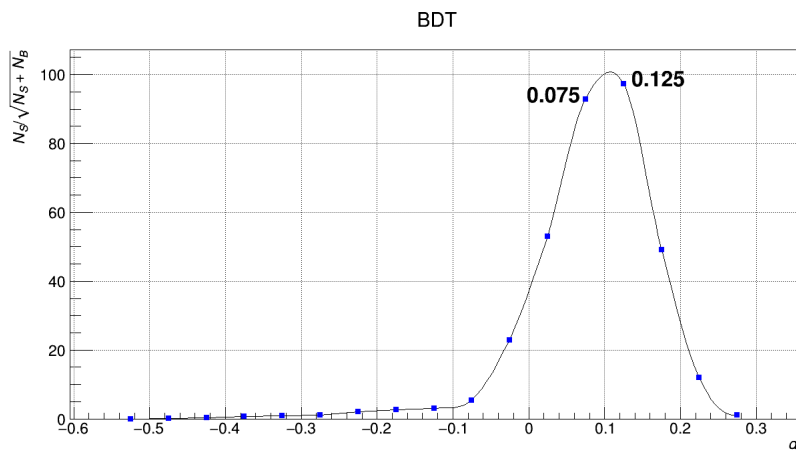
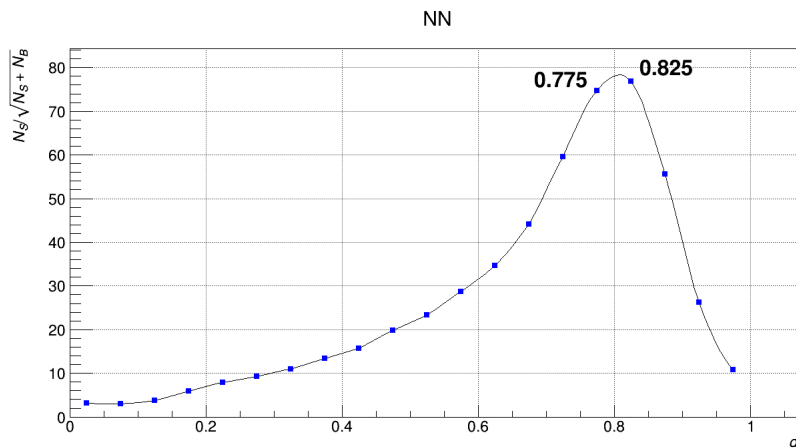
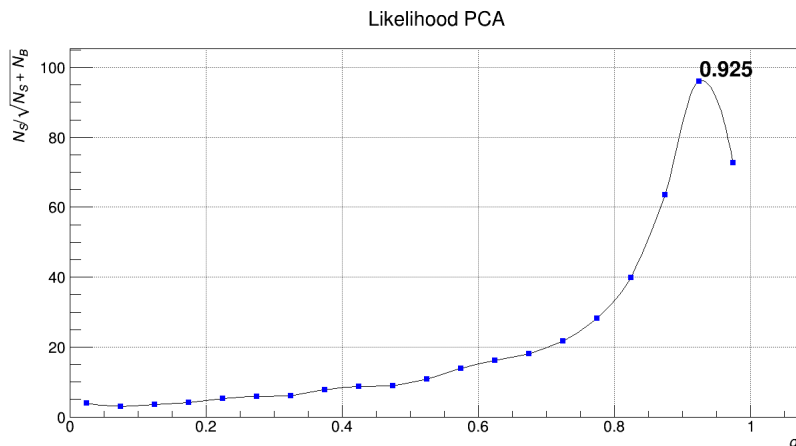
(a) Significancia como función de  $\alpha$  para el método BDT.(b) Significancia como función de  $\alpha$  para el método NN.(c) Significancia como función de  $\alpha$  para el método Likelihood PCA.

Figura 6.7: Significancia como función de la variable  $\alpha$  para los tres métodos implementados.

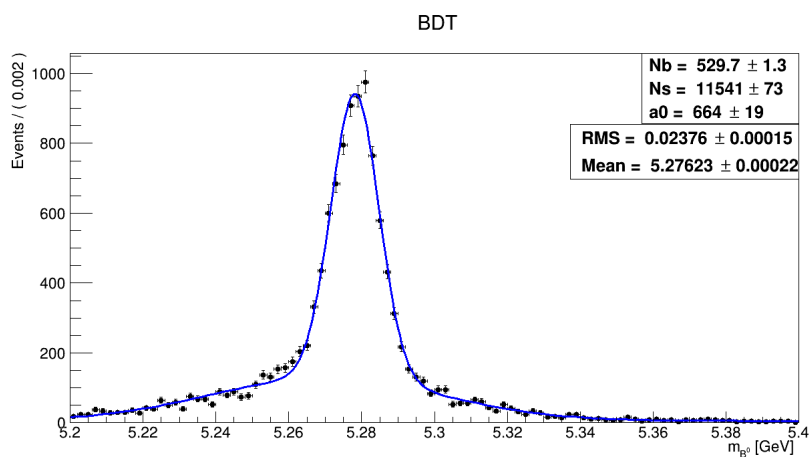
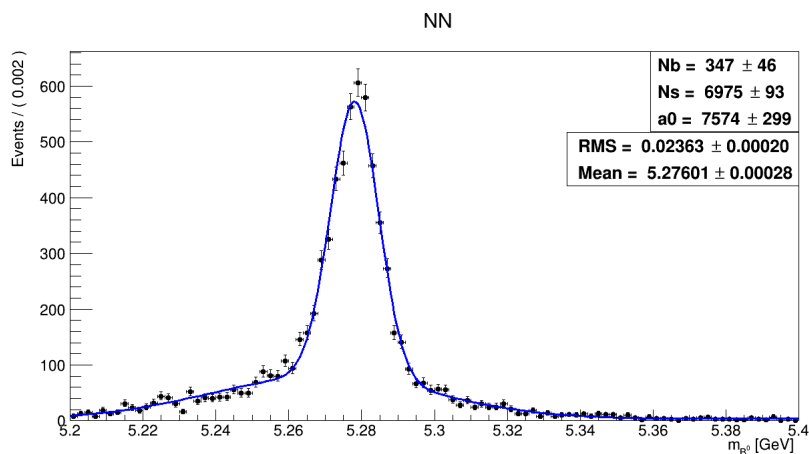
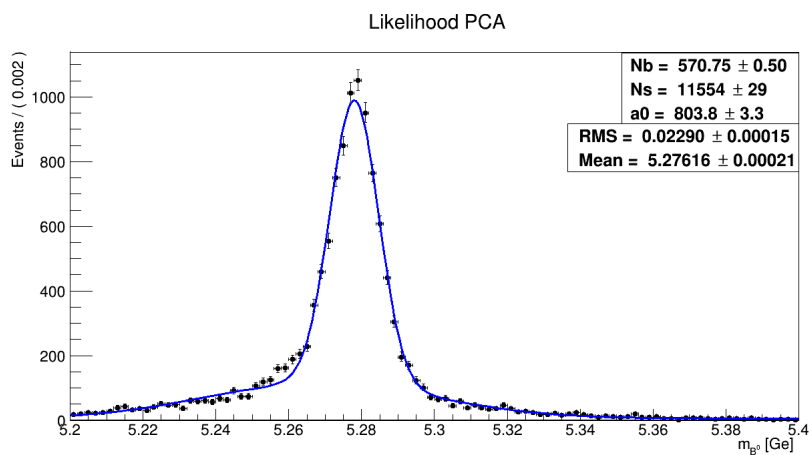
(a) Ajuste con cortes en la variable  $\alpha$  del método BDT.(b) Ajuste con cortes en la variable  $\alpha$  del método NN.(c) Ajuste con cortes en la variable  $\alpha$  del método Likelihood PCA.

Figura 6.8: Ajustes realizados aplicando cortes en la variable  $\alpha$  asignada a los eventos por cada uno de los tres métodos implementados.

# Conclusiones

En este trabajo implementamos tres diferentes clasificadores multivariable incluidos en el paquete TMVA integrado a ROOT en el estudio del canal de decaimiento  $B^0 \rightarrow D^- \pi^- \mu^+ \mu^+$ . La configuración utilizada en estos métodos de clasificación es la que tienen por defecto en el paquete TMVA. El uso de estas técnicas basadas en máquinas de aprendizaje tiene un papel fundamental en física de altas energías pues nos permite analizar datos recolectados por los aceleradores de partículas; al tiempo que el estudio de estos datos nos permite comprobar de forma precisa las teorías fundamentales existentes, así como descartar modelos de extensiones del modelo estándar y buscar física más allá del modelo estándar.

Al comparar los resultados de aplicar estos métodos a nuestra muestra de Monte Carlo observamos, de la figura (6.7), que los métodos que tienen mejor resultado a la hora de optimizar la significancia son el BDT y el Likelihood PCA, pues al ver el comportamiento de la significancia como función de la variable  $\alpha$  vemos que para estos dos métodos hay un máximo que es mayor al máximo alcanzado cuando se usan redes neuronales, NN. Por otro lado, también observamos que el comportamiento de la significancia como función de  $\alpha$  para el Likelihood PCA es más suave en el intervalo de definición que el comportamiento que hay en el BDT, esto es importante porque si nos interesa realizar estudios en una vecindad de un valor dado de la variable  $\alpha$  podríamos obtener resultados muy diferentes al variar un poco los extremos del intervalo de interés.

Después de aplicar los métodos de clasificación, realizamos un ajuste a la masa

invariante del mesón  $B^0$ . En el siguiente recuadro se resumen los valores resultantes de los parametros de los ajustes realizados a los eventos que pasaron los cortes en la variable  $\alpha$ .

Método	$N_s$	$N_b$	Media	Desviación
BDT	$11541 \pm 73$	$529.7 \pm 1.3$	$5.27623 \pm 0.00022$	$0.02376 \pm 0.00015$
NN	$6975 \pm 93$	$347 \pm 46$	$5.27601 \pm 0.00028$	$0.02363 \pm 0.00020$
Likelihood PCA	$11554 \pm 29$	$570.75 \pm 0.5$	$5.27616 \pm 0.00021$	$0.02290 \pm 0.00015$

Es posible continuar con este análisis, un punto que se debe mencionar aquí es que en la realización de este trabajo empleamos simulación de Monte Carlo en donde únicamente se generó el canal deseado, sin embargo actualmente existe simulación de Monte Carlo realizada en campañas por Belle II en donde se incluyen todos los posibles canales, medidos, en los que puede decaer un mesón  $B^0$ , así como ruido debido al haz de partículas. Utilizando estas dos muestras de datos podemos identificar otras fuentes de ruido y limpiar más nuestra señal. También podemos estudiar el uso de otras variables cinemáticas que nos permitan entrenar estos métodos de clasificación de forma que podamos quedarnos con la mayor cantidad posible de eventos de señal.

# Apéndice A

Pauli escribió una carta a Geiger y Meitner, quienes participaban en una reunión en Tübingen (Diciembre 4, 1930), para comunicarles acerca de su idea de la existencia de una nueva partícula. Pidiendo, además, que se les comunicara al resto de participantes acerca de su propuesta. La traducción al inglés de tal carta se muestra enseguida [1].

*Dear Radioactive Ladies and Gentlemen,*

*As the bearer of these lines, to whom I graciously ask you to listen, will explain to you in more detail, how because of the “wrong” statistics of the  $N$  and  $Li\beta$  nuclei and the continuous beta spectrum, I have hit upon a desperate remedy to save the “exchange theorem” of statistics and the law of conservation of energy. Namely, the possibility that there could exist in the nuclei electrically neutral particles, that I wish to call neutrons, which have spin  $1/2$  and obey the exclusion principle and which further differ from light quanta in that they do not travel with the velocity of light. The mass of the neutrons should be of the same order of magnitude as the electron mass and in any event not larger than  $0.01$  proton masses. The continuous beta spectrum would then become understandable by the assumption that in beta decay a neutron is emitted in addition to the electron such that the sum of the energies of the neutron and the electron is constant.*

*I agree that my remedy could seem incredible because one should have seen those neutrons very earlier if they really exist. But only the one who dare can win and the*



*difficult situation, due to the continuous structure of the beta spectrum, is lighted by a remark of my honored predecessor, Mr Debye, who told me recently in Brussels: "Oh, It's well better not to think to this at all, like new taxes". From now on, every solution to the issue must be discussed. Thus, dear radioactive people, look and judge. Unfortunately, I cannot appear in Tübingen personally since I am indispensable here in Zürich because of a ball on the night of 6/7 December. With my best regards to you, and also to Mr Back.*

*Your humble servant W. Pauli*

# Apéndice B

El archivo DEC que se utilizó para la generación de la muestra es el siguiente:

```
Alias MyB0 B0
Alias Myanti - B0 anti - B0
ChargeConj MyB0 Myanti - B0
Alias MyD+ D+
Alias MyD- D-
ChargeConj MyD+ MyD-
Decay Upsilon(4S)
1.00 MyB0 Myanti - B0 VSS_BMIX dm;
Enddecay
Decay MyB0
1.00 MyD- pi- mu+ mu+ PHSP;
Enddecay
CDecay Myanti - B0
Decay MyD-
1.00 K+ pi- pi- D_DALITZ;
Enddecay
CDecay MyD+
End
```

Para los detalles completos acerca de los modelos puede visitar [24].



# Bibliografía

- [1] Samoil M Bilenkij. Introduction to the physics of massive and mixed neutrinos. Berlin: Springer, 2011.
- [2] Diego Milanés, Néstor Quintero, and Carlos E. Vera. "Sensitivity to Majorana neutrinos in  $\Delta L = 2$  decays of B meson at LHCb." Physical Review D 93.9 (2016): n. pag. Web.
- [3] Yorikiyo Nagashima. Elementary particle physics. Weinheim, Germany: Wiley-VCH Verlag GmbH, 2010. Print.
- [4] Rajasekaran, G. "The Story of the Neutrino." [1606.08715] The Story of the Neutrino. N.p., 14 Oct. 2016. Web. 04 July 2017.
- [5] Sundaresan, Monsur K. Handbook of particle physics. Boca Raton: CRC Press, 2001. Print.
- [6] I. Gil-Botella. "Neutrino Physics." [1504.03551] Neutrino Physics. N.p., 14 Apr. 2015. Web. 04 July 2017.
- [7] McCauley, T., and L. Taylor. CMS collision events (7 TeV): candidate ZZ to  $2e+2\mu$ . CERN Document Server. N.p., 11 Aug. 2011. Web. 04 July 2017.
- [8] Belle Event Gallery. N.p., n.d. Web. 05 July 2017.
- [9] Wiechczynski, J. "Future Belle II experiment at the KEK laboratory." Belle II Document Server. N.p., n.d. Web. 05 July 2017.

- [10] Iván Heredia de la Cruz. "The Belle II experiment: fundamental physics at the flavor frontier." Belle II Document Server. N.p., n.d. Web. 05 July 2017.
- [11] Michel Enrique Hernández Villanueva, Eduard De La Cruz Burelo, and Gabriel López Castro. "Sensitivity study of the process  $B^0 \rightarrow D^- \pi^- \mu^+ \mu^+$  at the Belle II experiment." Belle II Document Server. Cinvestav, n.d. Web. 05 July 2017.
- [12] Matthew Barrett (on Behalf of the Belle II Collaboration). "The Belle II Experiment." Belle II Document Server. N.p., n.d. Web. 05 July 2017.
- [13] Accelerator Systems Division PEP-II: Asymmetric B Factory. SLAC. N.p., n.d. Web. 05 July 2017.
- [14] Belle II. N.p., n.d. Web. 05 July 2017.
- [15] T. Abe, et al. "Belle II Technical Design Report." [1011.0352] Belle II Technical Design Report. N.p., 01 Nov. 2010. Web. 05 July 2017.
- [16] Dorisyangsoo Kim. "The software library of the Belle II experiment." Nuclear and Particle Physics Proceedings 273-275 (2016): 957-62. Web.
- [17] "DESY Confluence." DESY Confluence. N.p., n.d. Web. 05 July 2017.
- [18] Agostinelli S., and J. Allison, et al. "Geant4—a simulation toolkit." Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 506.3 (2003): 250-303. Web.
- [19] "The software library of the coming Belle II experiment and its simulation package." 2013 IEEE Nuclear Science Symposium and Medical Imaging Conference (2013 NSS/MIC) (2013): n. pag. Web.
- [20] CERN Accelerating science. Big data takes ROOT — CERN. N.p., n.d. Web. 05 July 2017.

- [21] Kumar Ravi, and Arun Tripathi. ROOT: A Data Analysis and Data Mining Tool from CERN . N.p., n.d. Web. 5 July 2017.
- [22] Narsky, Ilya, and Frank C. Porter. Statistical analysis techniques in particle physics: fits, density estimation and supervised learning. Weinheim: Wiley-VCH, 2014. Print.
- [23] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, TMVA - Toolkit for Multivariate Data Analysis, PoS ACAT 040 (2007), arXiv:physics/0703039.
- [24] Welcome - EvtGen Project. EvtGen Development Team. Accessed August 01, 2017. <http://evtgen.warwick.ac.uk/docs/external/>.