

TRABAJO FIN DE GRADO

Estudio del acoplamiento de Yukawa entre el quark top y el bosón de Higgs en procesos $t\bar{t}H$ en el detector CMS del LHC del CERN



Universidad de Oviedo

FACULTAD DE CIENCIAS

GRADO EN FÍSICA

Covadonga Galán Fernández

Tutores:

Javier Fernández Menéndez

Clara Ramón Álvarez

Julio 2022

RESUMEN

En este trabajo se presentan los resultados del estudio del acoplamiento de Yukawa entre el *quark top* y el bosón de Higgs en procesos $t\bar{t}H$. Los datos analizados han sido recogidos por el detector CMS durante el *Run 2* en los años 2016, 2017 y 2018 en colisiones protón-protón con una energía de centro de masas de 13 TeV y una luminosidad integrada total de 137 fb^{-1} .

La intensidad del acoplamiento de Yukawa entre el bosón de Higgs y distintos fermiones y bosones aumenta con la masa de la partícula con la que se acopla. El bosón de Higgs y el *quark top* son las partículas más masivas del modelo estándar, por lo que el acoplamiento de Yukawa entre ellas es del orden de la unidad. Una desviación de la predicción dada por el modelo estándar podría dar indicios de una nueva física. Por esto, es de especial interés estudiar el acoplamiento entre el bosón de Higgs y el *quark top*.

En el modelo estándar el bosón de Higgs es escalar, y por tanto, par bajo inversiones de carga y paridad (CP). En este estudio se mide la estructura CP del acoplamiento entre el bosón de Higgs y el *quark top*. El estudio se realiza utilizando estados finales con dos leptones del mismo signo. Para la presentación de los resultados se utiliza el parámetro f_{CP} que relaciona las partes par e impar del acoplamiento de Yukawa entre el *quark top* y el bosón de Higgs. En este estudio se ha optimizado la clasificación de sucesos para obtener la medida más precisa de este parámetro. En los tres casos estudiados de variación de *bines* para la discriminación de la señal y el fondo en el nodo de $t\bar{t}H$, se ha obtenido un valor observado de 0.0 ± 0.7 para 13 *bines*, 0.0 ± 0.4 para 7 *bines* y 0.0 ± 0.8 para 5 *bines*. Estos valores son compatibles con el valor $f_{CP} = 0$ esperado por el modelo estándar, por lo que se comprueba que el acoplamiento de Yukawa entre el *quark top* y el bosón de Higgs es par.

Índice general

Resumen	I
Introducción	1
1. El modelo estándar	3
1.1. El modelo estándar	3
1.2. Más allá del modelo estándar	7
1.3. El proceso ttH	8
2. Dispositivo experimental	11
2.1. LHC y mecanismo de aceleración	11
2.1.1. Luminosidad y <i>pile-up</i>	13
2.1.2. Principales componentes del LHC.	15
2.2. El detector CMS	19
2.2.1. Sistema de coordenadas	20
2.2.2. El detector de trazas (<i>tracker</i>)	21
2.2.3. Calorímetro Electromagnético	22
2.2.4. Calorímetro Hadrónico	24
2.2.5. Solenoide superconductor	26
2.2.6. Cámaras de muones	27
2.2.7. <i>Trigger</i>	29
3. Reconstrucción y simulación de sucesos	30
3.1. Reconstrucción e identificación de sucesos	30
3.2. Generación y simulación de sucesos	35
3.2.1. Generación	36

3.2.2. Simulación	38
3.2.3. Herramientas del <i>software</i>	38
4. Estudio del proceso $t\bar{t}H$	39
4.1. Señal y fondo	39
4.2. Estados finales	40
4.3. Estimación de los fondos	41
4.3.1. Fondos irreducibles	41
4.3.2. Fondos reducibles	42
4.4. Selección de sucesos	45
4.4.1. Canal $2lss + 0\tau_H$	45
4.5. Fuentes de incertidumbre	50
4.5.1. Incertidumbres estadísticas	50
4.5.2. Incertidumbres sistemáticas	51
5. Resultados experimentales	53
5.1. Región de señal	53
5.1.1. Regiones de control	56
5.2. Discriminación CP	56
5.3. Clasificación	59
5.3.1. Nodo de $t\bar{t}H$	61
5.3.2. Variación del número de <i>bines</i>	62
5.4. Acuerdo <i>fakes</i> de Montecarlo vs <i>fakes</i> estimados de los datos	63
5.5. Ajuste de máxima verosimilitud y extracción del parámetro f_{CP}	67
5.5.1. Ajuste de máxima verosimilitud	67
5.5.2. Resultados	69
Conclusiones	72
A. Comparación de f_{CP} variando el número de <i>bines</i>	74
Glosario	78
Bibliografía	84

Índice de figuras

1.1. Partículas fundamentales del modelo estándar de la física de partículas. En color morado se representan los <i>quarks</i> , en color verde los leptones, en color rojo los bosones y en amarillo el bosón de Higgs. [3]	4
1.2. Interacciones entre las partículas del modelo estándar. [3]	6
1.3. Diagrama de Feynman de la producción de dos <i>quarks top</i> con un bosón de Higgs a primer orden en teoría de perturbaciones. [52]	8
1.4. Fracciones de desintegración del Higgs. [24]	9
1.5. Acoplamientos del bosón de Higgs a distintos fermiones y bosones. [20]	10
2.1. Complejo de aceleradores del CERN.[44]	11
2.2. Detalle del plan HL-LHC. [29]	14
2.3. Cavidades de radiofrecuencia [22]	16
2.4. Sección transversal de un dipolo del LHC [36]	16
2.5. Estructura del detector CMS y sus componentes principales. [48] . . .	19
2.6. Sistema de coordenadas en CMS. [9]	20
2.7. Detector de trazas de CMS. [32]	22
2.8. Esquema de la estructura del calorímetro electromagnético. [18]	23
2.9. Corte longitudinal del calorímetro electromagnético. [41]	23
2.10. Una de las tapas del calorímetro hadrónico. [57]	25
2.11. Corte longitudinal del calorímetro hadrónico que muestra la localización de los calorímetros HB, HE, HO y HF. [34]	25
2.12. Solenoide superconductor de CMS. [49]	26
2.13. Sistema de detección de muones de CMS. [17]	27
3.1. Corte transversal de CMS. [13]	31

3.2. Producción de <i>jets</i> . [30]	33
3.3. Generación de una colisión protón-protón. En rojo se representa la colisión principal, rodeada de las cascadas de partones. En color morado se ve el remanente de la colisión, en verde claro la hadronización y en verde oscuro la desintegración de los hadrones. [39]	37
4.1. Diagrama de Feynman de la producción $t\bar{t}H$ y su desintegración a dos leptones del mismo signo.	40
4.2. Diagramas de Feynman de los procesos $t\bar{t}W$ y $t\bar{t}Z$	41
4.3. Diagramas de Feynman de los procesos WZ y ZZ	42
4.4. Diagramas de Feynman de los procesos $t\bar{t}$, <i>Single top</i> , Drell-Yan y W^+W^-	43
4.5. Representación del momento transversal de los dos leptones más energéticos y de la masa invariante del par de leptones en el caso de aplicar criterios de selección en el momento y la carga eléctrica.	46
4.6. Representación del momento transversal del primer leptón y el número de <i>jets</i> y <i>b-tag</i> en el caso de aplicar todos los criterios previamente descritos. La señal $t\bar{t}H$ está representada en color rojo.	47
4.7. Representación del número de <i>jets</i> y número de <i>b-tag</i> para cada una de las combinaciones dadas en el cuadro anterior. Los sucesos esperados para cada contribución vienen apilados en distintos colores. La señal $t\bar{t}H$ está representada en color rojo.	48
4.8. Representación del número del momento transversal del primer leptón tras aplicar todos los criterios de selección.	50
5.1. Representación de la distribución del momento transversal del leptón más energético y del segundo leptón más energético para el conjunto de los tres años estudiados en este análisis.	54
5.2. Representación del número de <i>jets</i> y <i>b-tag medium</i> para el conjunto de los tres años estudiados en este análisis.	55
5.3. Observables físicos usados para la discriminación de escenarios en el canal $2l_{ss} + 0\tau_H$	58
5.4. Discriminante como combinación de diferentes observables físicos discriminantes.	59

5.5. Distribuciones de los fondos más importantes del análisis usando las simulaciones de Montecarlo.	60
5.6. Variación en el número de <i>bines</i> usando los <i>fakes</i> de las simulaciones de Montecarlo.	62
5.7. Representación de la distribución del momento transverso del leptón más energético y del segundo leptón más energético usando la contribución del fondo <i>fakes</i> estimado a partir de los datos	63
5.8. Representación del número de <i>jets</i> y <i>b-tag</i> usando la contribución del fondo <i>fakes</i> estimado a partir de los datos.	64
5.9. Separación de la clasificación de las redes neuronales en el canal $2l_{SS}+0\tau_H$ usando usando la contribución del fondo <i>fakes</i> estimado a partir de los datos	66
5.10. Variación en el número de <i>bines</i> para el nodo de $t\bar{t}H$ usando la contribución del fondo <i>fakes</i> estimados a partir de los datos.	67
5.11. Representaciones del parámetro f_{CP} tras el ajuste de máxima verosimilitud para los distintos <i>bines</i> considerados en cada región de CP en el nodo de $t\bar{t}H$	70
A.1. Representaciones del parámetro f_{CP} tras el ajuste de máxima verosimilitud para todas las variaciones de <i>bines</i> consideradas en cada región de CP en el nodo de $t\bar{t}H$	75

Índice de cuadros

4.1. Secciones eficaces de los fondos más importantes en este análisis y los generadores usados para simularlo. El generador de <i>Pythia</i> se combina con los anteriores para simular los procesos. [13]	44
4.2. Número de <i>jets</i> en cada <i>bin</i>	48
4.3. Proporción señal-fondo en cada <i>bin</i>	49
4.4. Selección de sucesos en el canal <i>2lss</i>	49
5.1. Sucesos considerados en el análisis para el conjunto de los años 2016, 2017 y 2018 usando las simulaciones de Montecarlo.	55
5.2. Sucesos considerados en el análisis para el conjunto de los años 2016, 2017 y 2018 usando los <i>fakes</i> estimados a partir de los datos.	64
5.3. Valores de f_{CP} obtenidos tras el ajuste para un intervalo de confianza (<i>Confidence Level</i>) del 68%. Se indican los resultados obtenidos para los valores esperados y los valores observados.	69
A.1. Valores de f_{CP} obtenidos tras el ajuste para todos los <i>bines</i> considerados en cada región de CP para un intervalo de confianza (<i>Confidence Level</i>) del 68%. Se indican los resultados obtenidos para las incertidumbres de los valores esperados y los valores observados.	74

Introducción

En este trabajo se estudiará la producción de un bosón de Higgs en asociación con un par de *quarks top* anti-*top* ($t\bar{t}H$) en colisiones protón-protón en el detector CMS del LHC del CERN. Dicho proceso, se puede usar en particular, para estudiar el acoplamiento entre el *quark top* y el bosón de Higgs.

El modelo estándar predice la existencia de un bosón que es responsable de dar masa al resto de partículas. Este bosón se conoce como bosón de Higgs y, en el modelo estándar, es escalar. Es decir, es par bajo inversiones de carga-paridad (CP). Esta partícula fue descubierta en los detectores CMS y ATLAS en el año 2012. Cualquier desviación de los acoplamientos entre las partículas del modelo estándar y el bosón de Higgs daría lugar a signos de nueva física.

Se presenta una medida directa de la estructura CP del acoplamiento de Yukawa, y_t , entre el *quark top* y el bosón de Higgs. Se han usado los datos recogidos por el detector CMS en colisiones protón-protón con una energía en centro de masas de 13 TeV durante los años 2016, 2017 y 2018 con una luminosidad integrada total de 137 fb^{-1} . Debido a las desintegraciones de los *quarks top* y el bosón de Higgs, hay varios canales de estudio según los estados finales que se produzcan. En este caso se estudiará el canal que contiene dos leptones con la misma energía en el estado final.

Este trabajo se relaciona con las siguientes asignaturas del Grado en Física de la Universidad de Oviedo: Física Nuclear y de Partículas Elementales, Métodos Numéricos Aplicados a la Física, Técnicas Experimentales y Física de Altas Energías y Aceleradores.

El trabajo se divide en 5 capítulos. El primer capítulo consiste en una introducción al modelo estándar descrito en física de partículas. En el segundo capítulo se desarrolla la estructura del gran colisionador de hadrones (LHC) y del detector del cual se han obtenido los datos para este análisis, el detector CMS. En el capítulo 3 se explica de manera introductoria los algoritmos usados para la reconstrucción de los sucesos, así como los pasos para realizar la simulación de los mismos. En el capítulo 4 se explica el proceso físico de estudio. Se definen los criterios de selección aplicados en el análisis y los fondos más importantes presentes en la región seleccionada. En el quinto capítulo se exponen los resultados experimentales obtenidos del estudio del acoplamiento de Yukawa. Por último, se recogerán las conclusiones del trabajo.

Capítulo 1

El modelo estándar

1.1. El modelo estándar

El modelo estándar (ME) es en física de partículas la mejor teoría que se tiene actualmente para describir la estructura fundamental de la materia, las partículas elementales y las interacciones fundamentales. Surge a partir de la unificación de la Teoría Electrodébil ¹ y la Cromodinámica Cuántica (QCD) ². El modelo estándar describe la materia ordinaria a partir de dos bloques fundamentales de partículas divididos según su spin: los fermiones (spin semientero) y los bosones (spin entero). Los fermiones se dividen a su vez en leptones y *quarks*. Las interacciones entre los fermiones se producen mediante el intercambio de los bosones mediadores de la interacción. Las partículas que sufran cierta interacción, tendrán una carga asociada a esta. La carga débil está asociada a la interacción débil, la carga eléctrica a la interacción electromagnética y la carga de color está asociada a la interacción fuerte. Existen tres tipos de carga de color, rojo, verde y azul. La intensidad de la interacción se mide a través de la constante de acoplamiento correspondiente, cuya representación será α_{em} para la interacción electromagnética, α_s para la interacción fuerte y α_W para la interacción débil.

La clasificación de las partículas se hace de acuerdo a su masa, su spin y su carga eléctrica. Según esta clasificación los fermiones se agrupan en tres familias o genera-

¹Teoría física desarrollada por Sheldon Lee Glashow, Abdus Salam y Steven Weinberg que unifica la interacción débil y el electromagnetismo.

²Teoría propuesta por David Politzer, Frank Wilczek y David Gross que describe la interacción fuerte entre quarks. Fue propuesta para entender la estructura de los bariones y los mesones.

ciones ordenadas de menor a mayor masa. La familia de los *quarks* se divide en tres generaciones. La primera está compuesta de los *quarks up* (u) y *down* (d), la segunda de los *quarks charm* (c) y *strange* (s) y la tercera de los *quarks top* (t) y *bottom* (b). Todos los *quarks* tienen spin $1/2$ y carga eléctrica fraccionaria ³. Los *quarks* u, c y t tienen carga eléctrica $\frac{2}{3}e$ y los *quarks* d, s y b, $-\frac{1}{3}e$. A parte de la carga eléctrica, poseen carga débil y carga de color. Mientras que los leptones existen como partículas libres, con los *quarks* no ocurre lo mismo. Una de las características más importantes de estas partículas es que no se han observado en libertad, están confinados. Este fenómeno es debido a que las partículas responsables de la interacción fuerte, los gluones, también poseen carga de color por lo que interaccionan consigo mismos. Esto da lugar a que los *quarks* estén confinados formando hadrones. Los hadrones son partículas no elementales formados por *quarks* unidos mediante la interacción fuerte. Dependiendo del número de *quarks* podemos clasificarlos en dos clases: los mesones formados por dos *quarks* (un *quark* y un *antiquark*) y los bariones formados por tres *quarks*.

mass →	=2.3 MeV/c ²	=1.275 GeV/c ²	=173.07 GeV/c ²	0	=126 GeV/c ²
charge →	2/3	2/3	2/3	0	0
spin →	1/2	1/2	1/2	1	0
	u up	c charm	t top	g gluon	H Higgs boson
	d down	s strange	b bottom	γ photon	
	e electron	μ muon	τ tau	Z Z boson	
	ν_e electron neutrino	ν_μ muon neutrino	ν_τ tau neutrino	W W boson	

Figura 1.1: Partículas fundamentales del modelo estándar de la física de partículas. En color morado se representan los *quarks*, en color verde los leptones, en color rojo los bosones y en amarillo el bosón de Higgs. [3]

³La carga eléctrica, denotada con el símbolo e , tiene un valor de $1.602176634 \cdot 10^{-19}$ C [38]

Para el caso de los leptones, también tenemos una división en tres generaciones. La primera está formada por el electrón (e^-) y el neutrino electrónico (ν_e), la segunda por el muon (μ^-) y el neutrino muónico (ν_μ) y la tercera por el tau (τ^-) y el neutrino tauónico (ν_τ). Todos los leptones tienen spin $1/2$, el electrón, el muon y el tau tienen carga eléctrica $-e$, no tienen carga de color por lo que no sufren la interacción fuerte, pero si tienen carga débil, sufren la interacción débil. Los neutrinos no tienen carga eléctrica ni carga fuerte, pero si carga débil.

Una clasificación de las partículas que forman el modelo estándar se puede ver en la figura 1.1. En ella se tiene en color morado los *quarks* y en verde los leptones (el conjunto son los fermiones) y en rojo, los bosones mediadores. En amarillo a la derecha se encuentra el bosón de Higgs. Además de las partículas anteriores, tenemos que añadir las antipartículas de cada una de ellas. Una antipartícula posee las mismas características que su partícula correspondiente salvo que su carga eléctrica es de signo opuesto. Si una partícula es diferente a su antipartícula, se llaman partículas de Dirac, sin embargo si una partícula es idéntica a su antipartícula, tenemos una partícula de Majorana.

El otro grupo de partículas del modelo estándar son los bosones, caracterizados por tener spin entero y no cumplir el principio de exclusión de Pauli ⁴. En el modelo estándar existen 5 tipos de bosones: el fotón (γ) responsable de la interacción electromagnética, los gluones (g) responsables de la interacción fuerte, los bosones W^\pm y Z^0 responsables de la interacción débil y el bosón de Higgs (H^0), que mediante el mecanismo de Higgs da lugar a la masa de las partículas. Todos tienen spin entero 1, menos el bosón de Higgs que tiene spin 0 (se denomina bosón escalar). Tanto el fotón como el gluón no tienen ni masa ni carga eléctrica, sin embargo el gluón posee carga de color, por lo que es la partícula mediadora de la interacción fuerte pero a la vez la sufre. Los bosones W^\pm y Z^0 son partículas masivas responsables de la interacción débil, por el hecho de ser tan masivas, el rango de la interacción débil es muy pequeño. La principal diferencia entre estos bosones es que los bosones W^\pm dan lugar a un cambio de sabor de las partículas.

⁴El principio de exclusión de Pauli establece que dos fermiones no pueden estar en el mismo estado cuántico, no pueden tener todos los números cuánticos iguales.

El bosón de Higgs se descubrió en el año 2012, en un momento en el cual era la partícula que faltaba para completar el marco teórico del modelo estándar. El mecanismo de Brout-Englert-Higgs fue propuesto por Robert Brout, François Englert y Peter Higgs y es el proceso que otorga masa a las partículas elementales que interactúan con el campo de Higgs. La intensidad de la interacción del bosón de Higgs con los fermiones se mide a través de las constantes de acoplamiento de Yukawa, y_f . El campo de Higgs tiene asociada una partícula llamada bosón de Higgs de masa $125.25 \pm 0.17 \text{ GeV}$ ⁵ [38] en unidades naturales ⁶. Esta partícula se descubrió en el año 2012 en los detectores CMS [11] y ATLAS [1] del CERN. En la figura 1.2 podemos ver los fermiones unidos a los bosones que median en las interacciones entre las partículas del modelo estándar.

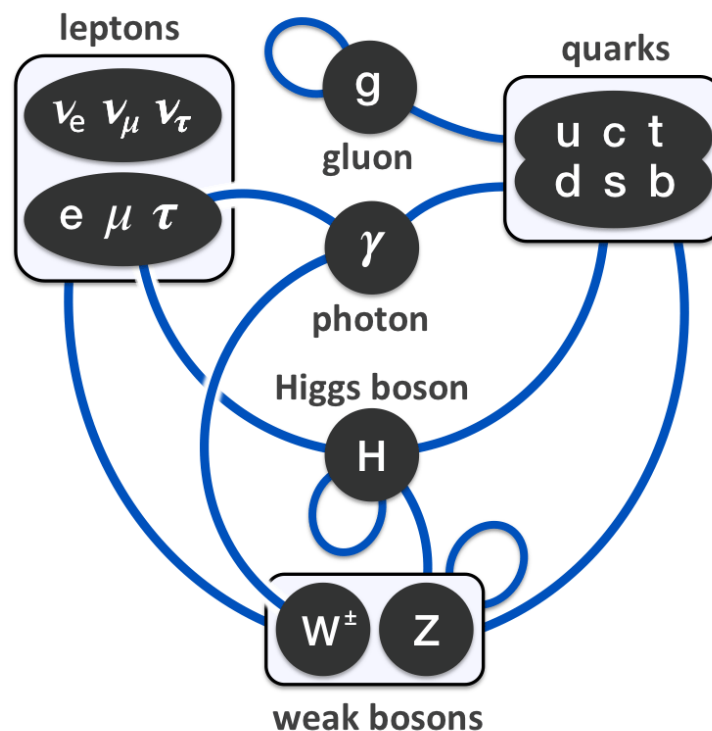


Figura 1.2: Interacciones entre las partículas del modelo estándar. [3]

⁵El electronvoltio (eV) es una unidad de energía cuyo valor es $1.602176634 \cdot 10^{-19} \text{ J}$ [38]. Representa la energía de un electrón al moverse de un punto a otro cuando la diferencia de potencial es de 1 voltio. En física de altas energías, esta unidad es muy pequeña por lo que se usan múltiplos como el megaelectronvoltio MeV, gigaelectronvoltio GeV o teraelectronvoltio TeV.

⁶Sistema de unidades que toma $\hbar = c = 1$ y en el cual tanto la masa como el momento tienen unidades de energía.

1.2. Más allá del modelo estándar

Durante décadas se han determinado los parámetros del modelo estándar con más y más precisión, buscando desviaciones de los valores teóricos que puedan indicar la presencia de nueva física. Aunque el modelo estándar es la teoría más precisa en la actualidad, tiene ciertas limitaciones. La búsqueda de nueva física mediante experimentos en el LHC del CERN ha sido motivada para responder a preguntas que abarcan desde la composición de la materia oscura hasta dimensiones extra. [42]

Algunos de los fenómenos para los cuales el modelo estándar no proporciona una explicación adecuada son los siguientes: [56]

- **Materia oscura y energía oscura:** No se conoce la composición de la materia oscura. Se descubrió a partir de las curvas de rotación de las galaxias, ya que con la materia visible no se describía el comportamiento observado. La materia oscura no interactúa con la materia bariónica ⁷, pues invisible a la luz y otras formas de radiación electromagnética. La materia oscura representa un 27% del Universo mientras que la energía oscura constituye un 68% del universo. La energía oscura se distribuye de manera uniforme por todo el Universo, lo que tiende a acelerar su expansión. [53]
- **Asimetría de la materia-antimateria:** La mayoría del Universo se encuentra compuesto principalmente por materia. Suponiendo que en el Big Bang se creó la misma cantidad de materia que de antimateria, dentro del modelo estándar no hay ninguna explicación que indique la prevalencia de la materia frente a la antimateria.
- **Interacción gravitatoria** El modelo estándar no proporciona una explicación para la interacción gravitatoria. Es incompatible con la Teoría de la Relatividad General de Einstein, en la cual la fuerza gravitatoria es producida por la deformación del espacio-tiempo. En el modelo estándar las interacciones son mediadas por partículas, en el caso de la interacción gravitatoria, se ha postulado una partícula llamada *gravitrón* la cual no ha sido aún descubierta.

⁷La materia bariónica es toda forma de materia que está formada por bariones y leptones, toda la materia visible.

El proceso $t\bar{t}H$ consiste en la producción de dos *quarks top* con un bosón de Higgs, como se puede ver en la figura 1.3. Por un lado, el *quark top* se desintegra casi de manera exclusiva en un *quark b* y un bosón W . El bosón W a su vez se desintegra en una pareja *quark-antiquark* o en una pareja leptón-neutrino.

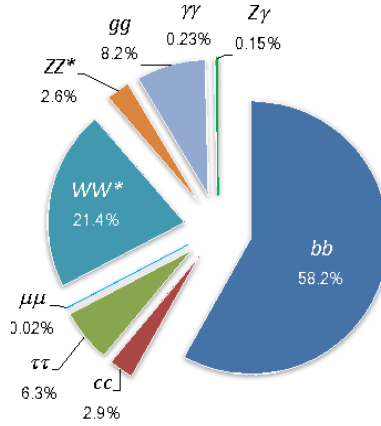


Figura 1.4: Fracciones de desintegración del Higgs. [24]

Por otro lado, el bosón de Higgs presenta un gran número de modos de desintegración, de las cuales las mayoritarias son las desintegraciones a $b\bar{b}$, W^+W^- , $\tau^+\tau^-$ y Z^0Z^0 . En la figura 1.4 se pueden ver las distintas fracciones de desintegración del Higgs. Por esto, el proceso $t\bar{t}H$ da lugar a una gran variedad de estados finales que permiten estudiar diferentes propiedades del bosón de Higgs.

En el modelo estándar, las masas de los fermiones se explican introduciendo las interacciones de Yukawa, y_f , entre el bosón de Higgs y los fermiones. Los fermiones se acoplan al bosón de Higgs con una intensidad y_f proporcional a la masa m_f de estos $y_f \approx m_f/v$, donde $v \approx 246$ GeV es el valor esperado del campo de Higgs en el vacío. [52] Las medidas de las desintegraciones del bosón de Higgs a fermiones corresponden con las predicciones hechas por el modelo estándar dentro de sus incertidumbres. Debido a la alta masa del *quark top*, el bosón de Higgs no se desintegra a un par *top anti-top*. No obstante, se puede estudiar el acoplamiento entre el Higgs y el *quark top* de forma directa usando el proceso $t\bar{t}H$. Como se puede ver en la figura 1.5, se representa la intensidad del acoplamiento de Yukawa frente a la masa de distintos fermiones y bosones. Esta relación es proporcional, a medida que aumenta el acoplamiento con el bosón de Higgs de una partícula, su masa aumenta. Observando la comparación con la

predicción hecha por el modelo estándar, desviaciones del ME podrían dar indicios de una nueva física. En particular, dado que el *quark top* es el fermión más masivo del ME con una masa de 173.34 ± 0.76 GeV [4], su acoplamiento de Yukawa es del orden de la unidad. Esto hace que sea de especial interés el estudio del acoplamiento entre el Higgs y el *quark top*.

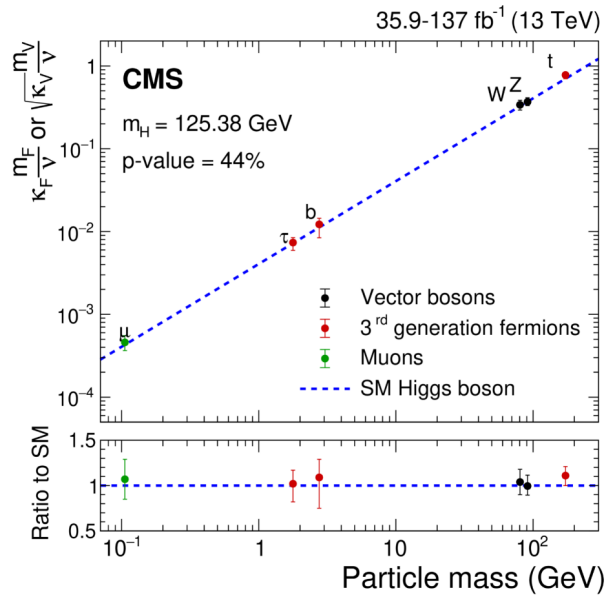


Figura 1.5: Acoplamientos del bosón de Higgs a distintos fermiones y bosones. [20]

En este trabajo se escogerá para el estudio de este proceso, el canal con dos leptones, donde los estados finales contienen al menos dos leptones, haciendo referencia a los leptones como electrones y muones, dejando de lado a los taus ya que tienen una vida media muy corta y no pueden ser detectados directamente, pero sí a través de sus productos de desintegración. Seleccionando estos estados finales, se está eligiendo sucesos en los que el Higgs se ha desintegrado a W^+W^- , que es el segundo canal con una fracción de desintegración mayor, tras el canal $b\bar{b}$. La ventaja con respecto al canal $b\bar{b}$ es que los leptones nos dan una señal clara, más fácil de distinguir de otras colisiones dado que las colisiones en el LHC son hadrónicas. La sección eficaz del proceso $t\bar{t}H$ es de 0.507 pb, mientras que la sección eficaz del proceso $t\bar{t}$ es de 833 pb, lo que da una idea de la dificultad del estudio del proceso $t\bar{t}H$.

Capítulo 2

Dispositivo experimental

2.1. LHC y mecanismo de aceleración

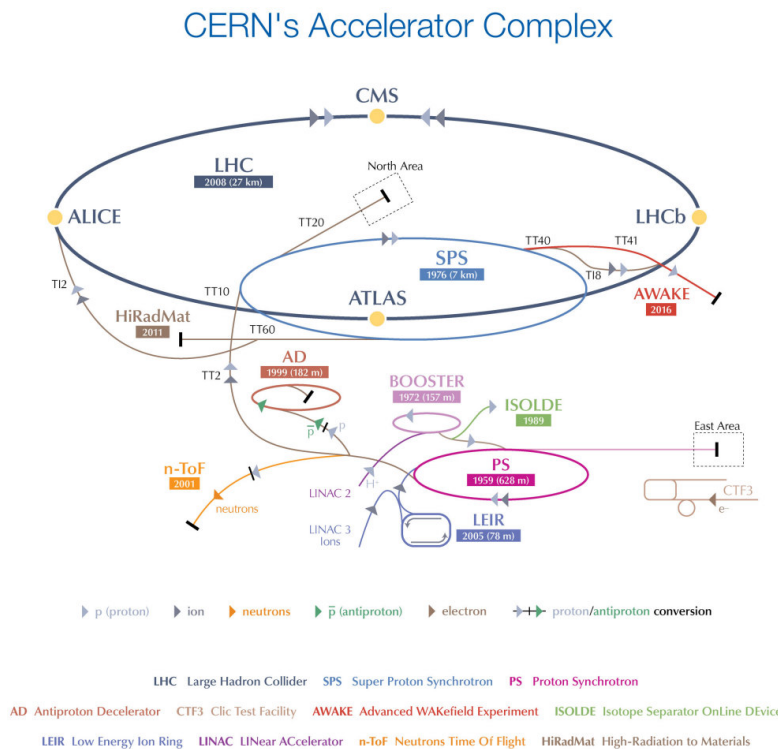


Figura 2.1: Complejo de aceleradores del CERN.[44]

El Large Hadron Collider (LHC) es el acelerador de partículas circular más grande y de mayor energía del mundo. Este acelerador fue construido entre 1998 y 2008 por el CERN gracias a la colaboración de miles de científicos y universidades de todo el

mundo. Está construido en el túnel ya existente del Large Electron-Positron (LEP), en funcionamiento desde 1989 hasta el 2000, de 26.7 kilómetros de circunferencia situado a 100 metros de profundidad cerca de la ciudad de Ginebra, en la frontera entre Francia y Suiza.

Por el LHC se hacen viajar haces de partículas para producir colisiones protón-protón, plomo-plomo o protón-plomo. Estos haces se dividen y se aceleran de manera secuencial a través de diferentes aceleradores antes de ser inyectados en el LHC. El complejo acelerador del CERN está compuesto por una sucesión de máquinas que aceleran partículas a energías cada vez más altas antes de inyectarlo a la siguiente máquina en la secuencia.

Proceso de aceleración

El proceso de aceleración empieza en el Acelerador lineal 4 (LINAC 4). El LINAC 4 está diseñado para impulsar los iones de hidrógeno negativos (átomos de hidrógeno con un electrón adicional) a altas energías para prepararlos para ingresar al Proton Synchrotron Booster (PSB). En este acelerador circular, se despojan a los iones de sus dos electrones, acelerando los protones resultantes hasta 1.4 GeV para ser inyectados en el Proton Synchrotron (PS). Este opera hasta los 26 GeV, inyectando los protones al Super Proton Synchrotron (SPS) donde se aceleran hasta los 450 GeV. Finalmente los protones se transfieren a los dos tubos de haz del LHC. Los protones no son las únicas partículas aceleradas en el LHC, los iones de plomo para el LHC parten de una fuente de plomo vaporizado y entran en Acelerador lineal 3 (LINAC 3) antes de ser recogidos y acelerados en el anillo de iones de baja energía, Low Energy Iron Ring (LEIR). Luego siguen el mismo camino que los protones. [48] Se puede observar el proceso descrito anteriormente en la figura 2.1.

Los dos tubos del LHC son dos anillos paralelos en ultra alto vacío, donde los haces viajan en direcciones opuestas a energías de 13 TeV y velocidades cercanas a la luz. A lo largo de los anillos del LHC existen varios puntos en los cuales se hacen colisionar los protones, en ellos se colocan los detectores: ATLAS, ALICE, LHCb, CMS.

2.1.1. Luminosidad y *pile-up*

Luminosidad

Uno de los principales parámetros en un acelerador es la luminosidad instantánea (\mathcal{L}). Es una medida del número de colisiones que pueden producirse por unidad de área y tiempo en un detector. Cuanto mayor sea \mathcal{L} , mayor es el número de colisiones en un experimento. Se puede calcular \mathcal{L} de la siguiente manera, ecuación (2.1):

$$\mathcal{L} = \frac{N_1 N_2 f N_b}{4\pi\sigma_x\sigma_y} \quad (2.1)$$

donde f es la frecuencia de cruce entre paquetes de protones (*bunches*), N_1 y N_2 son el número de partículas en cada *bunch*, N_b es el número de *bunches* en el haz y σ_x y σ_y es la desviación típica de la función de distribución gaussiana que determina la forma del *bunch*. Se corresponde con el tamaño de la sección del *bunch* en el punto de interacción (del orden de micras en este caso). [47]

Si se integra esta cantidad en un periodo de tiempo, se obtiene la luminosidad integrada. Se mide en unidades inversas de sección eficaz ¹, es decir, unidades inversas de área y se suele expresar en barns, b^{-1} , ($1b = 10^{-28}m^2$). Es por lo tanto, una medida de la cantidad de datos obtenidos de colisiones en un periodo de tiempo T, ecuación (2.2).

$$L = \int_0^T \mathcal{L}(t') dt' \quad (2.2)$$

Dado un proceso cuya sección eficaz (σ) es conocida, se puede estimar el número de sucesos esperados como $N = \sigma \cdot L$.

Experimento de alta luminosidad (HL-LHC)

El CERN ha adoptado la Estrategia Europea para la Física de Partículas, que tiene como alta prioridad la plena explotación del LHC. Esto ha dado lugar al proyecto LHC de Alta Luminosidad (HL-LHC), que consiste en un conjunto de actualizaciones sobre

¹Es el área efectiva en una colisión. Indica la probabilidad de que un proceso tenga lugar, se mide en unidades de área (m^2)

el LHC y todos sus detectores. Debido a retrasos, está previsto que la instalación final de HL-LHC empiece una vez finalice el *Run 3*, a partir de 2026, donde el LHC operará con su energía en centro de masas de $\sqrt{s} = 14$ TeV, pero con una luminosidad de entre 5 y 7 veces la luminosidad diseñada para el LHC. Esto permitirá la recopilación de muchas más datos de los que se han obtenido en el *Run 1*, *Run 2* y *Run 3*. Cada *Run* es cada periodo de toma de datos del LHC. [48] En la figura 2.2 se pueden ver los distintos periodos de toma de datos del LHC, la energía en centro de masas a la que opera y la previsión del HL-LHC.

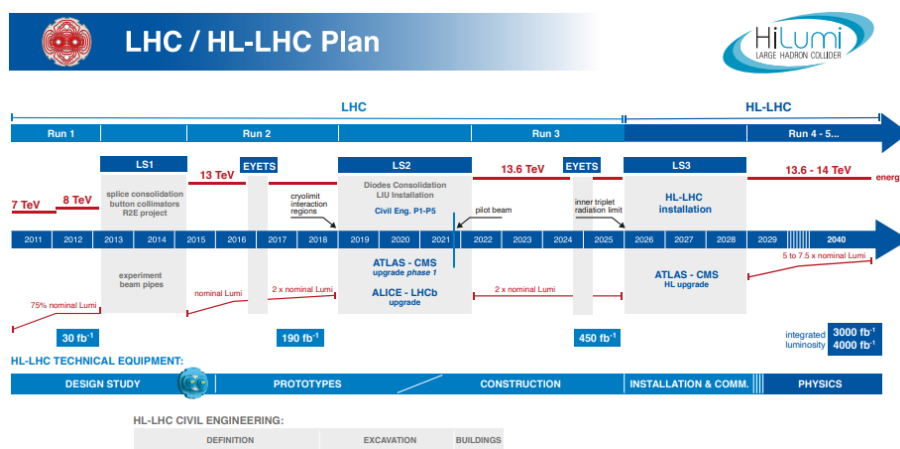


Figura 2.2: Detalle del plan HL-LHC. [29]

En este trabajo se realizará el análisis de los datos recogidos por el detector CMS durante el *Run-2* en los años 2016, 2017 y 2018, con una luminosidad integrada de 36.3, 41.4 y 59.7 fb^{-1} , respectivamente. El valor promedio del número de colisiones simultáneas (también denominado *pile-up*, ver siguiente sección) en el año 2018 fue de 32.

Pile-up

En las colisiones protón-protón en el LHC, uno de los retos a superar a la hora de analizar los datos, es el apilamiento de datos, *pile-up*. Cuando chocan dos paquetes de protones, se produce más de una interacción protón-protón y en el punto de colisión se observarán muchas de estas colisiones en un periodo de tiempo muy pequeño. Es debido a la alta luminosidad instantánea del LHC y da lugar a muchas partículas que puedan aparentar venir del proceso principal.

2.1.2. Principales componentes del LHC.

La energía en centro de masas máxima a la que ha operado el LHC en su último periodo de toma de datos (*Run 2*) ha sido de $\sqrt{s} = 13$ TeV, lo que corresponde a 6.5 TeV por haz. En esta sección se describirán los componentes principales del acelerador que han permitido llegar a esta energía en centro de masas. Una información más detallada del funcionamiento y los componentes del LHC se puede encontrar en [7].

El LHC se ha construido utilizando infraestructuras ya existentes para reducir su coste. Esto ha dado lugar a un gran número de restricciones técnicas. La máxima energía que se puede alcanzar en una máquina circular depende del radio que se curvan las partículas por los dipolos y la máxima intensidad de campo magnético que se alcanza. Como el radio de curvatura de las partículas está determinado por la geometría del túnel, el campo debería ser lo más alto posible. El campo magnético que se requiere para conseguir una energía de 6.5 TeV es de 8.33 T aproximadamente. Esto dio lugar a un nuevo límite a la hora de diseñar los imanes superconductores y sus sistemas de refrigeración asociados.

La siguiente restricción fue el diámetro del túnel, 3.8 m. Para que el imán alcance el estado superconductor ha de conservarse frío y para ello debe insertarse un recipiente de vacío llamado criostato. Debido al pequeño tamaño transversal del túnel, los dos anillos por los que circulan los haces están separados por 19 cm dentro de un yugo común y un criostato.

Cavidades de radiofrecuencia

Para acelerar las partículas, los aceleradores disponen de unas cámaras metálicas que contienen un campo electromagnético, llamadas cavidades de radiofrecuencia (RF). Cuando entran las partículas cargadas en el seno de este campo, reciben un impulso que las acelera. El campo en una cavidad está hecho para oscilar a una frecuencia dada. Esto hace que los protones con energías diferentes a la deseada sean acelerados o frenados para homogeneizar el haz. Cuando el haz ha alcanzado la energía requerida, un protón con la energía correcta no se acelerará. De esta manera, el haz de partículas se clasifica en los paquetes de protones llamados *bunches*.



Figura 2.3: Cavidades de radiofrecuencia [22]

En el LHC se usan 16 cavidades de radiofrecuencia que hacen que las partículas pasen de 450 GeV a 6.5 TeV. En la figura 2.3 se pueden ver las cavidades de RF del LHC. La energía máxima se alcanza en unos 20 minutos cuando los haces han pasado unas 10 millones de veces por las cavidades. Estas cavidades se encuentran en cuatro refrigeradores llamados criomódulos, lo que les permite estar a temperaturas muy bajas y llegar al estado superconductor ². En el LHC, cada cavidad de RF está sintonizada para oscilar a 400 MHz. En el corazón del LHC se encuentra el sistema de imanes superconductores y su sistema criogénico asociado. [22]

Imanes superconductores

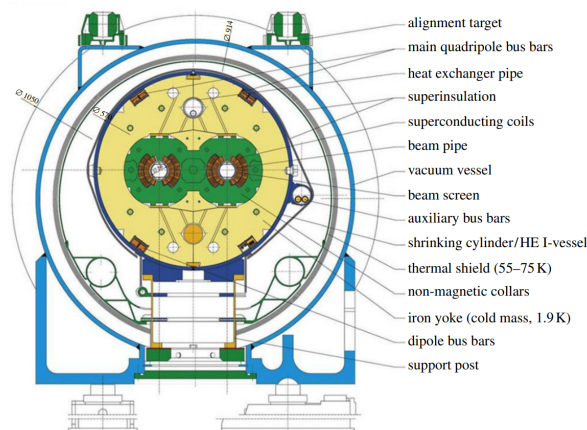


Figura 2.4: Sección transversal de un dipolo del LHC [36]

²La superconductividad es la capacidad que tienen algunos materiales de conducir la corriente eléctrica sin resistencia ni pérdidas de energía. En un material superconductor la resistencia cae bruscamente cuando se enfría por debajo de su temperatura crítica.

Otro componente importante del LHC son los imanes superconductores. Para que la trayectoria de los haces sea circular, las tuberías están rodeadas de una gran sistema de imanes, operando en un estado superconductor, que generan un campo magnético intenso y desvían la trayectoria de las partículas. La parte principal del sistema consta de 1232 imanes dipolares en total que miden 15 metros de largo, pesan unas 35 toneladas y operan produciendo un campo constante de 8.33 T. El campo creado por estos imanes se dirige en la dirección vertical para que la fuerza de Lorentz ³ curve el haz en la dirección deseada. En la figura 2.4 podemos ver la estructura de la sección transversal de estos imanes dipolares.

En los extremos de los imanes dipolares, debido a efectos de borde, se crean imperfecciones que hacen que el campo creado no sea homogéneo. Por ello, en los bordes de los imanes dipolares, hay imanes cuadrupolares para enfocar los haces colocados de manera simétrica en torno al tubo por el que circula el haz, y miles de imanes sextupolares y octupolares adicionales para corregir las pequeñas imperfecciones en el campo magnético. Los imanes dipolares están fabricados por niobio-titanio (NbTi) y necesitan una temperatura de -271.3 °C para obtener el rendimiento deseado, por ello, para conseguir esta temperatura se necesita un sistema criogénico asociado. [31]

Sistema criogénico

El uso de materiales superconductores es la mejor manera de evitar el sobrecalentamiento en las bobinas ya que los cables deben mantenerse a bajas temperaturas para alcanzar un estado superconductor. Para ello, los imanes deben enfriarse a -271.3 °C (1.9 K), una temperatura inferior a la del espacio exterior (2.7 K). Para enfriar los imanes, gran parte del acelerador está conectado a un sistema de distribución de nitrógeno y helio líquidos. Todo el proceso de refrigeración tarda semanas en completarse y consta de tres etapas diferentes. En la primera, el helio se enfría a 80 K en unos intercambiadores de calor de los refrigeradores utilizando 10000 toneladas de nitrógeno líquido y luego a 4.5 K en las turbinas refrigeradoras. En la segunda etapa se inyecta en las masas frías de los imanes y en la tercera, se enfría a una temperatura de 1.9 K mediante las

³Una partícula cargada en movimiento en el seno de un campo magnético, experimentará la acción de una fuerza magnética que recibe el nombre de fuerza de Lorentz. Esta fuerza vendrá dada por la expresión $F = q(\vec{v} \times \vec{B})$.

unidades de refrigeración. En total, el sistema criogénico enfría unas 36000 toneladas de masas frías magnéticas. El helio es un excelente refrigerante para estabilizar los sistemas superconductores del LHC. Esto hace que sea el mayor sistema criogénico del mundo, uno de los lugares mas fríos de la tierra. [25]

Sistema de vacío

Para evitar las colisiones de los haces con las moléculas de gas residual del tubo del haz, es fundamental estar en ultra-alto vacío para preservar la criogenia. Se tiene ultra-alto vacío a unos 10^{-7} mbar ⁴. El LHC tiene 3 sistemas de vacío: un vacío de aislamiento para los imanes enfriados por el sistema criogénico, un vacío de aislamiento para la distribución del helio y un vacío del tubo del haz. En el sistema criogénico y en el sistema de distribución del helio, el vacío actúa como un aislante térmico para reducir la cantidad de calor que se filtra desde el ambiente a las partes del sistema. El ultra-alto vacío se mantiene mediante el bombeo criogénico de 9000 metros cúbicos de gas. Se requieren algo menos de dos semanas de bombeo para bajar las presiones por debajo de 10^{-10} mbar. [21]

Imanes de inserción

Cuando los haces de partículas entran en los detectores, los imanes de inserción ejercen fuerzas que comprimen los paquetes de protones que circulan por el tubo para concentrar los protones que vienen de direcciones opuestas. Después del choque de los haces en los detectores, los haces se separan de nuevo por imanes dipolares. Cuando llega el momento de deshacerse de las partículas, se desvían del LHC a lo largo de una línea recta y un imán reduce la intensidad del haz antes de que choque con un bloque de hormigón y compuesto de grafito para su parada final. Los imanes de inserción también son responsables de la limpieza del haz, lo que garantiza que las partículas parásitas no entren en contacto con los componentes más sensibles del LHC.

⁴El milibar (mbar) es una unidad de presión equivalente a una milésima parte del bar. Un milibar equivale a un hectopascal (hPa). La presión atmosférica a nivel del mar es de 1013.25 hPa.

2.2. El detector CMS

El detector *Compact Muon Solenoid* (CMS) es un detector de propósito general que se encuentra en uno de los cuatro puntos de colisión del LHC y está diseñado para observar cualquier nuevo fenómeno físico que pueda revelar el LHC, desde el estudio del modelo estándar, hasta la búsqueda de materia oscura.

El detector tiene forma cilíndrica y está colocado alrededor del tubo donde viajan los haces y se producen las colisiones. Sus dimensiones son las siguientes: mide 21 metros de largo, tiene unos 15 metros de diámetro y pesa unas 14000 toneladas. Su diseño detecta los muones de manera precisa y tiene el solenoide más potente jamás creado, capaz de generar un campo de 3.9 T. Estas características le dan el nombre de *Compact Muon Solenoid*.

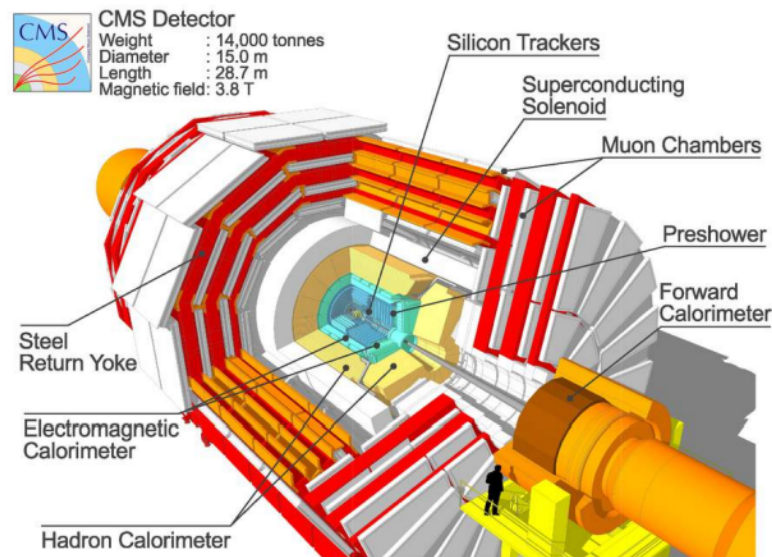


Figura 2.5: Estructura del detector CMS y sus componentes principales. [48]

Como se puede ver en la figura 2.5, está formado por varias capas concéntricas, en la que cada una detecta uno o varios tipos de partículas que se pueden producir en la colisión. Las distintas capas son: el detector de trazas (*tracker*), el calorímetro electromagnético (ECAL), el calorímetro hadrónico (HCAL), el solenoide superconductor y por último las cámaras de muones. A continuación se describirá de manera breve cada uno de estos componentes. Se puede encontrar una descripción completa del detector en [5].

2.2.1. Sistema de coordenadas

El sistema de coordenadas tomado por CMS tiene su origen en el punto de colisión de los haces. Dada la simetría cilíndrica del detector, lo más natural es utilizar coordenadas cilíndricas. El eje X va hacia el centro del LHC, el eje Y hacia la superficie y el eje Z en la dirección del túnel del LHC. El plano XY es el plano transversal y se define de manera en que el eje X apunte hacia dentro de la circunferencia del LHC. En la figura 2.6 se pueden ver las coordenadas descritas.

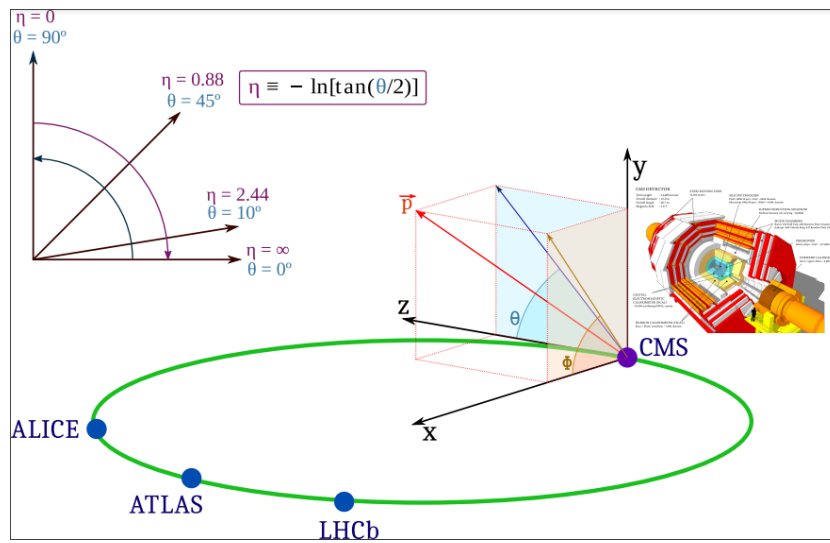


Figura 2.6: Sistema de coordenadas en CMS. [9]

La coordenada radial se denota por r . El ángulo azimutal $\phi \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ se mide respecto al eje X en el plano XY. El ángulo polar $\theta \in [0, \pi)$ se mide respecto al eje Z en el plano y a partir de él se define la pseudorapidez η , con valores entre $(-\infty, \infty)$. La pseudorapidez, ecuación (2.3), describe el ángulo con el que sale la partícula después de la colisión respecto al eje del haz. Esta magnitud es importante porque es invariante bajo las transformaciones de Lorentz, la medición no depende del marco teórico de referencia elegido.

$$\eta = -\ln \left[\operatorname{tg} \left(\frac{\theta}{2} \right) \right] \quad (2.3)$$

Se usa η en lugar del ángulo polar porque los detectores cilíndricos al ser concéntricos limitan los ángulos polares en los que se puede obtener una reconstrucción. Partículas

con $\eta = 0$ estarán en el plano transversal, mientras que partículas con una pseudorapidez alta estarán más próximas al tubo del haz. El momento de una partícula viene dado por el producto entre la masa y la velocidad, $\vec{p} = m \cdot v = (p_x, p_y, p_z)$. En lugar del módulo del momento, se utiliza su componente transversal \vec{p}_T en el plano XY, cuyo módulo vendrá dado por

$$p_T = \sqrt{p_x^2 + p_y^2} \quad (2.4)$$

2.2.2. El detector de trazas (*tracker*)

El detector de trazas es la parte más interna de CMS. Gracias a él, se pueden identificar las trayectorias de las partículas con carga que lo atraviesan, y permite determinar su momento y el vértice principal del que procedían. Además, el *tracker* permite a su vez reconstruir vértices de partículas cuya vida media es lo suficientemente grande como para viajar una distancia apreciable por el detector, como es el caso de los *quarks b*. El proceso de identificación de estos *quarks* se conoce como *b-tagging*, el cual se explicará más adelante. El *tracker* debe registrar las trayectorias de las partículas con precisión pero perturbando la partícula lo menos posible, esto lo hace tomando mediciones de posición tan precisas que las trayectorias se pueden reconstruir de manera fiable utilizando solo unos pocos puntos de medición. [32]

El *tracker* está situado en torno al punto de choque de los haces, por lo que al ser una de las capas más internas de CMS, recibe el mayor volumen de partículas. Tiene una longitud de 5.8 metros, un diámetro de 2.5 metros y está compuesto por un detector de píxeles de silicio y de tiras de silicio, porque es un material resistente a la radiación para que no se sobrecaliente por la energía de las partículas que recibe. [34]

El *tracker* está compuesto por:

- **Píxeles de silicio:** Los píxeles de silicio son la parte del *tracker* más cercana al punto de colisión. Cubren un rango de pseudorapidez $|\eta| < 2.5$ y está formado por 65 millones de píxeles, lo que permite determinar la trayectoria de las partículas con precisión. [34] Cuando una partícula cargada pasa a través del detector, tiene

energía suficiente para arrancar los electrones de los átomos de silicio, que son recogidos por los píxeles, creando una señal eléctrica que se amplifica. Se estima que pasan unas 10 millones de partículas por segundo y por centímetro cuadrado.

[32]

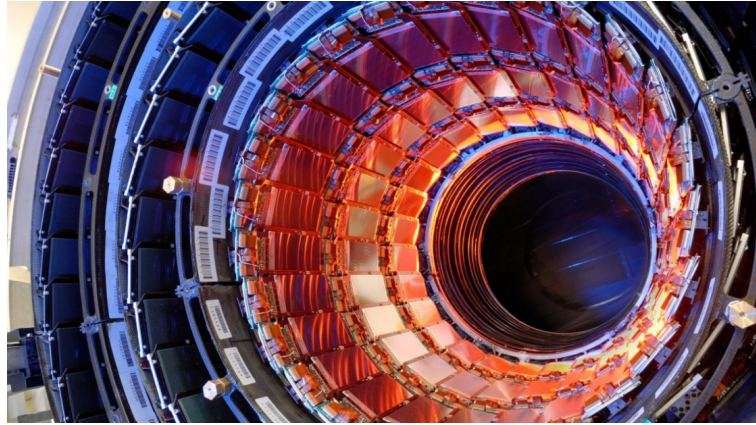


Figura 2.7: Detector de trazas de CMS. [32]

- **Tiras de silicio:** Después de los píxeles, las partículas pasan a través de diez capas de detectores de tiras de silicio. Esta parte del rastreador contiene 15200 módulos altamente sensibles con un total de 10 millones de tiras detectoras leídas por 80000 chips microelectrónicos. Las tiras de silicio funcionan de la misma manera que los píxeles: la energía de la partícula cargada que pasa, es suficiente para arrancar los electrones del material, creando una señal eléctrica que es amplificada, lo que nos permite reconstruir su camino. [32]

Los píxeles son mucho más pequeños y se ponen cerca del punto de interacción para optimizar la resolución (unas pocas micras) mientras que las tiras sirven para medir la trayectoria en un rango mayor de distancias pero con menor granularidad (decenas o casi centenas de micras).

2.2.3. Calorímetro Electromagnético

El calorímetro electromagnético (ECAL) es un subdetector de CMS compuesto de 80000 cristales de tungsteno de plomo ($PbWO_4$), que se encarga de medir la energía de los fotones y los electrones. Este material emite unos destellos rápidos y cortos cuando es excitado por una radiación ionizante. Cuando las partículas pasan a través

de los cristales, se produce una cantidad de luz proporcional a la energía de estas, que es recogida por unos fotodetectores y se traduce a una señal eléctrica que lleva la información de la partícula. El uso de estos cristales permitió diseñar un calorímetro compacto dentro del solenoide que es rápido y resistente a la radiación.[5]

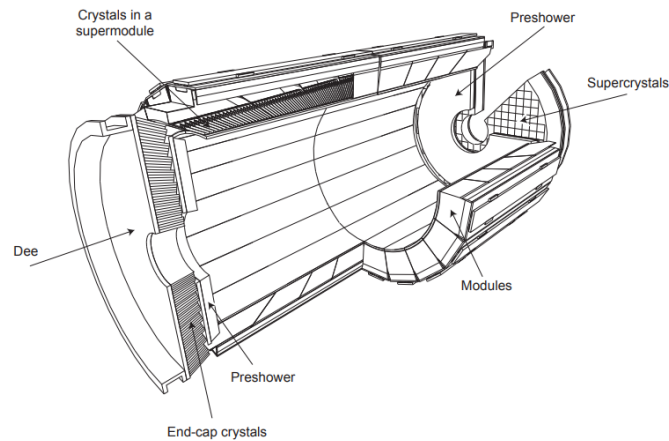


Figura 2.8: Esquema de la estructura del calorímetro electromagnético. [18]

Como se ve en la figura 2.8 el ECAL, al igual que CMS, es cilíndrico. Además, se divide en dos partes, una parte con forma de "barril", *Barrel Part (EB)*, y las "tapas finales", *endcaps (EE)*. La parte cilíndrica con forma de barril está formado por 61200 cristales agrupados en 36 supermódulos con 1700 cristales cada uno. Las tapas están formadas por 15000 cristales y sellan el barril en cada extremo. Para una mayor precisión, el ECAL contiene un detector "precascada" (*preshower*), delante de las tapas. Esto permite diferenciar fotones de alta energía con pares de fotones de baja energía. [27]

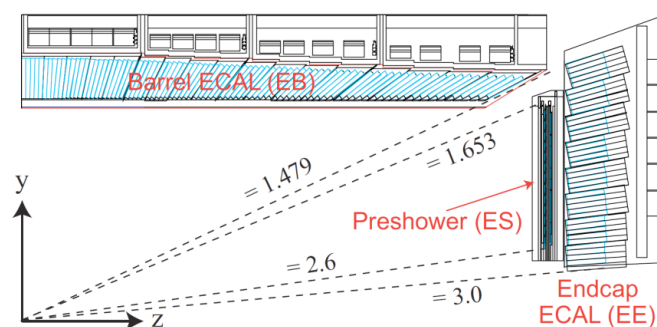


Figura 2.9: Corte longitudinal del calorímetro electromagnético. [41]

En la figura 2.9, se puede ver un corte longitudinal del calorímetro electromagnético. La *Barrel Part* cubre una pseudorapidez de $|\eta| < 1.48$ y las *Endcaps* abarcan un rango de la pseudorapidez entre $1.65 < |\eta| < 3.0$. En el rango de pseudorapidez $1.48 < |\eta| < 1.65$, la eficiencia de detección de electrones es baja.

2.2.4. Calorímetro Hadrónico

Después del calorímetro electromagnético se encuentra el calorímetro hadrónico, (HCAL). Su función es medir la energía de los hadrones, partículas hechas de *quarks* y gluones, como pueden ser protones, neutrones, piones o kaones. También ofrece una medida indirecta de la presencia de partículas no cargadas que no interactúan con los detectores, como son los neutrinos, a partir de la energía transversa faltante (MET). Este detector absorbe la gran mayoría de las partículas, pudiendo escapar de él solamente neutrinos y muones. Como puede haber partículas que no interaccionen con el detector, el calorímetro hadrónico ha de ser lo más hermético posible, de tal manera que recoja cada partícula producida en las colisiones. Así, si se detectan partículas en un lado del detector, pero por el otro lado no, hay un desequilibrio de energía y momento. Esto no significa que la energía y el momento no se conserven, sino que hay partículas que no se pueden detectar. [28]

El calorímetro hadrónico está formado por capas de materiales absorbentes de latón o acero intercaladas con placas de plástico centelleador que producen un pulso de luz rápido cuando una partícula los atraviesa. Unas fibras ópticas dentro del centelleador recogen esa luz y las señales ópticas obtenidas, se van sumando formando "torres", generando la trayectoria de la partícula además de medir su energía. Estas señales ópticas se convierten en señales electrónicas mediante unos fotodiodos híbridos (HPDs). [10]

EL HCAL se divide en varias secciones. Tiene una sección en forma de barril, *Hadron Barrel* (HB), a continuación del calorímetro electromagnético; y un calorímetro exterior, *Hadron Outer* (HO), que se llama así porque está colocado fuera del solenoide magnético para detectar la energía de partículas que no haya detectado el HB. Tiene dos tapas en los laterales, *Hadron Endcaps* (HE), figura 2.10, y dos calorímetros frontales, *Hadron*

Foward (HF) que están al final del tubo de CMS con el fin de detectar las partículas producidas en la colisión con un ángulo muy pequeño respecto a la dirección de los haces. Esta zona recibirá la mayor parte de la energía contenida en la colisión, por tanto, deberá estar hecha de materiales muy resistentes a la radiación y usar diferentes materiales en el resto del calorímetro. Tanto el barril como las tapas experimentan campo magnético de casi 4 T del solenoide CMS y, por lo tanto, tienen que estar hechos de un material no magnético. [54]



Figura 2.10: Una de las tapas del calorímetro hadrónico. [57]

En la figura 2.11 se ve un corte longitudinal del calorímetro hadrónico. Las líneas discontinuas corresponden a un valor fijo de η . El HB y el HO cubren una pseudorapidez de $|\eta| < 1.4$, el HE de $1.3 < |\eta| < 3.0$ y HF $3.0 < |\eta| < 5.2$. De esta manera todo el espacio está cubierto por lo que es difícil que alguna partícula no sea detectada.



Figura 2.11: Corte longitudinal del calorímetro hadrónico que muestra la localización de los calorímetros HB, HE, HO y HF. [34]

2.2.5. Solenoide superconductor

El solenoide superconductor es uno de los componentes más importantes del detector, genera el campo magnético que curva la trayectoria de las partículas que surgen de la colisión en CMS. Esto permite calcular el signo de las partículas cargadas, así como su momento. Cuanto mayor sea su momento, menos será curvada por el campo magnético. En la figura 2.12 se puede ver el solenoide superconductor de CMS. [23]

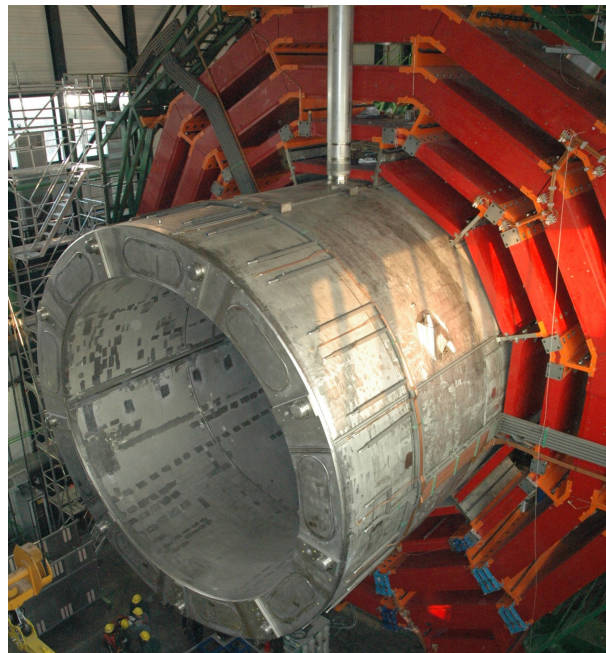


Figura 2.12: Solenoide superconductor de CMS. [49]

El solenoide está formado por bobinas de alambre que producen un campo magnético de 3.9 T cuando las atraviesa una corriente eléctrica. Cuantas más bobinas tenga, mayor será el campo magnético producido, y se obtendrán resultados más precisos. Es un solenoide superconductor, lo que significa que permite el paso de la corriente eléctrica sin ofrecer ninguna resistencia. Para ello se necesita enfriar el solenoide hasta 4.65 K. [23] Dentro de él, se encuentran el *tracker*, el calorímetro electromagnético y el calorímetro hadrónico; fuera se encuentran las cámaras de muones. El solenoide está incrustado en un armazón de hierro con el fin de guiar las líneas de campo magnético y el sentido de este cambia según se mida dentro o fuera del solenoide. El solenoide de CMS es el imán superconductor más potente que se ha construido hasta la fecha.

2.2.6. Cámaras de muones

Como su nombre sugiere, *Compact Muon Solenoid*, la detección de muones es una de las principales tareas de CMS. Los muones son leptones cargados más pesados que el electrón y son una de las principales firmas del bosón de Higgs en su desintegración a cuatro muones ($H \rightarrow ZZ \rightarrow 4\mu$).

Una vez pasado el solenoide, se encuentran las cámaras de muones, forman el detector más externo y de mayor tamaño de CMS. Se encargan de detectar y medir el momento de los muones que proceden de las colisiones en CMS. A este detector solamente llegan muones y partículas no cargadas como son los neutrinos. Por la naturaleza de los muones, estos son capaces de atravesar los calorímetros y las capas de hierro sin interactuar con el resto de los detectores. Por esto, las cámaras de muones se colocan en la parte más externa del detector. [26]

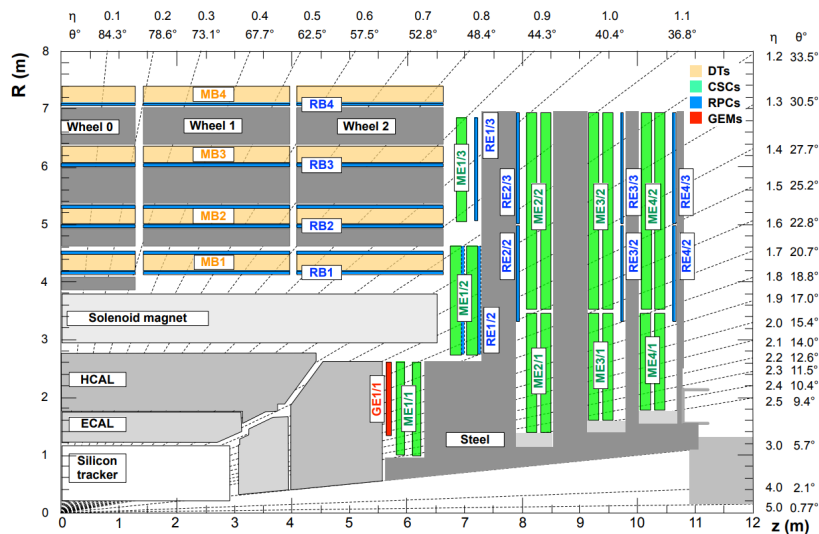


Figura 2.13: Sistema de detección de muones de CMS. [17]

Como las cámaras de muones están fuera del solenoide, las líneas de campo magnético curvarán la trayectoria de los muones en sentido contrario. Al igual que el resto de detectores, las cámaras de muones tienen forma cilíndrica y se dividen en un "barril" y dos "tapas". En la figura 2.13 se puede ver el sistema de detección de muones de CMS. Se utilizan tres tipos de subdetectores para reconstruir la trayectoria de los muones.

En cada uno de estos componentes se tiene un volumen de gas, cuando un muon o una partícula cargada los atraviesa, arranca los electrones de los átomos del gas generando una señal eléctrica. [26] [45] [34]

- **Tubos de deriva**, *Drift Tubes* (DF): Los tubos de deriva miden la posición de los muones en el "barril". Está formado por 5 ruedas, cada una dividida en 12 sectores y se extiende hasta $|\eta| < 1.2$. Cuando un muon o una partícula cargada atraviesa uno de estos tubos, en los cuales se tiene un cable dentro de un volumen de gas, tendrá la energía suficiente para arrancar un electrón de los átomos del gas. El cable registra una señal que da las coordenadas de la posición del muon.
- **Cámaras de tiras catódicas**, *Cathode Strip Chambers* (CSC): Las cámaras de tiras catódicas se encuentran en las tapas y se extienden en $0.9 < |\eta| < 2.4$. En esta zona el campo magnético no es uniforme y la tasa de partículas es alta. Cada CSC tiene forma trapezoidal, está formada por 6 huecos de gas y contienen cables de "ánodo" cruzados con tiras de "cátodo". Debido a que las tiras y los cables son perpendiculares, se obtienen dos coordenadas de posición para cada partícula que pasa. Cada módulo CSC contiene seis capas que lo hacen capaz de identificar con precisión los muones.
- **Cámaras de placas resistivas**, *Resistive Plate Chambers* (RPC): Las cámaras de placas resistivas se encuentran tanto en la zona del barril como en las tapas. Cubren una pseudorapidez de $|\eta| < 1.6$. Cuando un muon pasa se genera una señal que da lugar a un impulso que luego es utilizado por un disparador para tomar decisiones inmediatas sobre si vale la pena mantener los datos. Las RPC combinan una buena resolución espacial con una resolución de tiempo de solo un nanosegundo. Esto permite complementar las medidas del resto de cámaras y alcanzar una mayor precisión en la determinación de la trayectoria de los muones.

En total hay 1400 cámaras de muones: 250 tubos de deriva y 540 cámaras de tiras catódicas rastrean las posiciones de las partículas y proporcionan un disparador, mientras que 610 cámaras de placas resistivas forman un sistema de disparo, que rápidamente decide si mantener los datos de muones adquiridos o no. El sistema completo de detección de muones abarca hasta $|\eta| < 2.4$. [26]

2.2.7. *Trigger*

Dentro de CMS tienen lugar millones de colisiones protón-protón cada segundo, por lo que no hay manera de almacenar tal cantidad de datos. Además, no todos los procesos son relevantes para los análisis, la mayoría tienen baja energía y no pueden dar lugar a partículas masivas. Por lo tanto, se necesita de un *trigger* que sea capaz de seleccionar y almacenar los sucesos interesantes para su posterior análisis. Para no confundir las partículas de dos sucesos diferentes, los detectores deben tener una muy buena resolución y las señales de los millones de canales electrónicos deben estar sincronizadas para poder identificar las del mismo suceso.

Para seleccionar sucesos de interés, el *trigger* utiliza dos niveles. El Nivel 1 (L1), un proceso rápido y automático que busca signos de física interesante como por ejemplo, partículas de mucha energía. Selecciona sucesos que contienen objetos candidatos, reduciendo los millones que sucesos que tienen lugar cada segundo. El siguiente nivel es el Nivel Alto (HLT). El HLT sincroniza la información de las diferentes partes del detector para recrear todo el suceso. Los ordenadores revisan la información realizando pruebas complejas para buscar firmas específicas como por ejemplo, haciendo coincidir los depósitos de energía de las cámaras de muones con trazas. En general, en el *Run 2* se seleccionaron al final 1000 sucesos por segundo y el resto se descartaron. [33]

Capítulo 3

Reconstrucción y simulación de sucesos

3.1. Reconstrucción e identificación de sucesos

En el capítulo anterior se ha visto cómo se aceleran las partículas en el LHC y las colisiones en los distintos detectores. Estas colisiones dan lugar a un gran número de partículas, que pueden ser estables o inestables, que al pasar por los subdetectores dejan depósitos de energía, los cuales se pueden usar para identificar las partículas. Se llama reconstrucción a pasar de las señales eléctricas a todo aquello compatible con un objeto físico.

En la figura 3.1 se puede ver un corte transversal de CMS. Empezando por la región de colisión, las partículas entran primero en el *tracker*, donde las partículas cargadas son las únicas que dejan señal y describen una trayectoria curva por el imán. Los electrones se curvan al pasar por el *tracker* y depositan su energía en el ECAL, mientras que los hadrones cargados lo harán en el HCAL. Las partículas que no tienen carga como es el caso de los fotones, se detectan en el ECAL y los hadrones neutros en el HCAL. Por último los muones son detectados en las cámaras de muones, en la parte más externa de CMS. Por lo tanto, dependiendo del tipo de partícula y sus propiedades, interaccionará con las distintas partes del detector, pudiendo de esta manera identificarla.

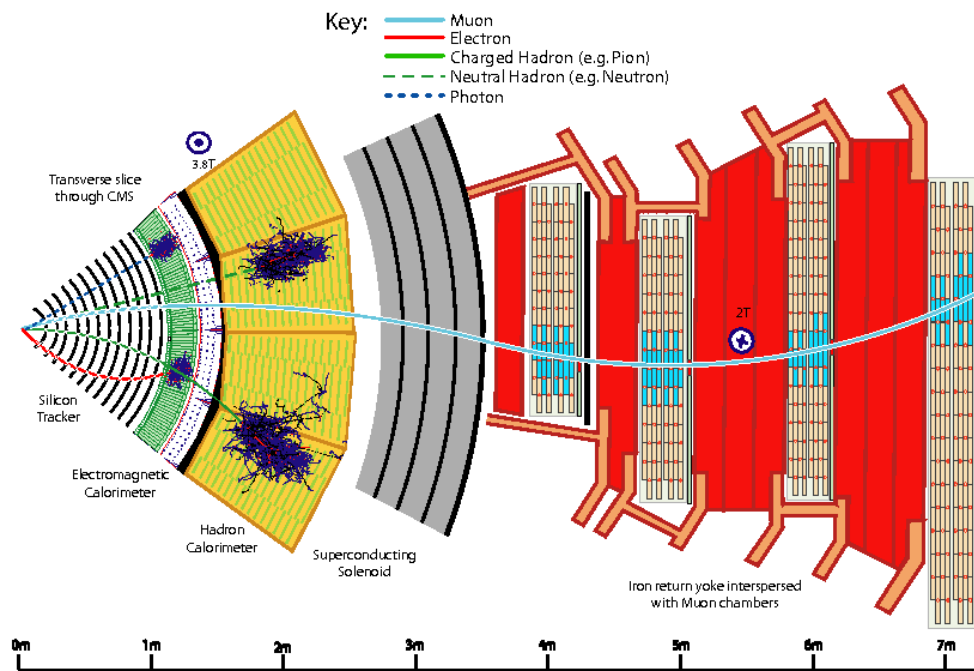


Figura 3.1: Corte transversal de CMS. [13]

El algoritmo *Particle Flow* (PF) reconstruye e identifica cada partícula mediante la combinación de la información dada por cada subdetector. A partir de las señales que deja la partícula en su paso por el detector, se unen las señales y deposiciones de energía mediante un algoritmo iterativo, partiendo de una mayor precisión y acabando con un criterio para reconstruir el mayor número de trazas posibles.

El algoritmo trabaja con cinco tipos de partículas: electrón, muon, fotón, hadrón cargado y hadrón neutro. Para la identificación de las partículas se piden ciertos criterios de calidad como pueden ser criterios angulares o criterios en la compatibilidad de la traza. El PF analiza cada suceso de manera global, siguiendo una secuencia y excluyendo los sucesos ya reconstruidos. Primero reconstruye los objetos más sencillos, los muones mediante la unión de las trazas en el *tracker* con las deposiciones de energía en las cámaras de muones. A continuación, pasa a los electrones y hadrones cargados con los depósitos de energía correspondientes en el ECAL y el HCAL. Una vez están reconstruidos los objetos, el algoritmo los identifica verificando lo que han reconstruido y obteniendo el momento de la partícula correspondiente. Dependiendo del tipo de análisis que se quiera realizar prevalecerán los criterios que nos darán mayor eficiencia frente a la pureza de la señal o viceversa. [11]

Vértice primario

El vértice primario es el punto en el cual se produce la colisión protón-protón. Se reconstruye extrapolando las trayectorias de las partículas cargadas hasta que converjan en un único punto. Debido a la presencia de *pile-up*, el vértice primario de interacción protón-protón debe elegirse entre los varios candidatos a vértices que se reconstruyen en cada evento de colisión. El vértice con el mayor valor de la suma del momento transversal se toma como el vértice primario de la colisión. [13]

Muones

Los muones son más sencillos de reconstruir, usando la información del *tracker* y las cámaras de muones. Los electrones y los muones en los eventos de señal se espera que estén aislados. Estos leptones aislados se distinguen usando la suma del momento transversal sobre las partículas cargadas, hadrones neutros, y fotones que son reconstruidos en un cono estrecho centrado en la dirección del leptón, que dependerá de la pseudorapidez y el ángulo azimutal. Cuanto mayor sea el cono, más aislado estará el leptón, y más se reducirá el efecto de *pile-up*. Una vez se reconstruyen los muones, las trazas correspondientes no se tienen en consideración para la reconstrucción de las demás partículas. Según la eficiencia o la pureza de la reconstrucción se pueden distinguir tres categorías de muones, *loose*, *medium* o *tight*, siendo la categoría *tight* la más restrictiva. [51]

Electrones

Los electrones son más complicados de reconstruir que los muones. Para ello se utiliza la información del *tracker* y los depósitos de energía del ECAL. La reconstrucción de los electrones está orientada a caracterizar electrones energéticos y aislados. A la hora de reconstruir sus trayectorias hay que tener en cuenta que los electrones son partículas cargadas y pierden energía emitiendo radiación por *Bremsstrahlung*¹. Estas trayectorias se reconstruyen mediante un algoritmo que busca depósitos de energía en el ECAL y se usan métodos de análisis multivariante para distinguir los electrones de los hadrones cargados. [51]

¹La radiación de frenado o *Bremsstrahlung* es una radiación producida por la desaceleración de una partícula cargada.

Jets

Los jets son las firmas experimentales de los *quarks* y los gluones producidos en procesos de altas energías en las colisiones protón-protón. Como los *quarks* y los gluones tienen carga de color, no pueden existir en libertad, por lo tanto, se hadronizan dando lugar a un chorro de partículas llamado *jet*. La excepción es el *quark top*, cuya vida media es tan corta que se desintegra antes de hadronizarse. Todos estos jets de partículas se propagan a través del detector y dejan depósitos de energía en los calorímetros. Estas señales se combinan usando algoritmos para formar un *jet* reconstruido. Los jets se reconstruyen por la acumulación de objetos *particle-flow* mediante el algoritmo *anti-kt*. *Anti-kt* es un algoritmo genérico de reconstrucción de jets usado en otros experimentos también, puede tomar de entrada objetos PF o trazas y depósitos de energía simplemente. [8]

Se utilizan objetos reconstruidos por el *particle flow* y se agrupan siguiendo una forma cónica usando un parámetro de distancia $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.4$. La energía de los jets reconstruidos se corrige por efectos de *pile-up* y se calibra como una función del momento transverso y la pseudorapidez. [13]

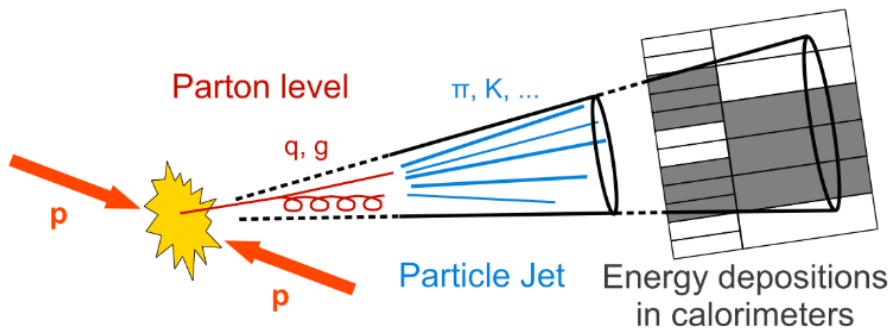


Figura 3.2: Producción de jets. [30]

Jets etiquetados como provenientes de un *quark b*

Los jets que provienen de la desintegración de un *quark b* son particulares. Los hadrones con un *quark b* en su interior tienen una vida media más larga, y se desintegran dando lugar a un vértice secundario dentro del detector. Debido a la precisión del detector, es posible identificar los jets que vengan de la desintegración de un *quark b*, se denotarán como *b-tag*. Los *b-tag* se identifican mediante algoritmos de *b-tagging*, uno de ellos es el algoritmo DEEPJET, [15] que utiliza características relacionadas con la

vida media de los hadrones y masa de los *jets* comparando con *jets* ligeros de gluones y *quarks* para poder identificarlos. [13] Se basan en técnicas de análisis multivariante y redes neuronales cuyas entradas son los *jets* reconstruidos. Los *b-tag* se pueden clasificar en tres categorías *loose*, *medium*, *tight* que nos indican la probabilidad creciente de que el *jet* identificado venga de un *quark b*.

Taus

Reconstruir e identificar los taus se hace de manera diferente a los electrones y muones, ya que por su corta vida media se desintegran antes de hadronizarse. Con una masa de 1.777 GeV, el leptón tau es lo suficientemente pesado como para desintegrarse en hadrones. Un tercio de las veces, se desintegra en un muon o en un electrón con dos neutrinos, y se reconstruyen e identifican con las técnicas para ellos. En el resto de los casos, los taus se desintegran hadrónicamente en una combinación de mesones y un neutrino. Los taus que se desintegran hadrónicamente se denotan con τ_H y se reconstruyen e identifican con el algoritmo *Hadrons-Plus-Strips* (HPS) [19] La reconstrucción de los taus se hace a partir de combinaciones de partículas neutras y cargadas reconstruidas por el algoritmo PF que son compatibles con las desintegraciones de los taus. [46]

Momento transverso faltante

Partículas como los neutrinos no interactúan con los elementos del detector. No obstante, su presencia se infiere de la conservación de la energía y el momento total. En el plano transversal de una colisión protón-protón el momento transversal debería ser nulo. Si es distinto de 0, debe haber partículas que no han podido ser detectadas. El momento transversal faltante o MET se define como el vector negativo del momento sobre la suma de todas las partículas reconstruidas por el algoritmo PF.

$$E_T^{Miss} = - \left| \sum_{i=1}^{Num_{particulas}} \vec{p}_T \right| \quad (3.1)$$

Esta información se complementa con el observable H_T^{Miss} , definido como la magnitud de la suma vectorial de leptones, τ_H y *jets*.

$$H_T^{Miss} = \left| \sum_{leptones} \vec{p}_T^l + \sum_{\tau_H} \vec{p}_T^{\tau} + \sum_{jets} \vec{p}_T^j \right| \quad (3.2)$$

La resolución de H_T^{Miss} es peor que la resolución de E_T^{Miss} pero la ventaja del observable es que los leptones, τ_H y los *jets* mayoritariamente se originan en la interacción *hard-scattering*, que se detallará más adelante, y rara vez de las interacción de *pile-up*, lo que hace que el observable sea menos sensible a las variaciones de las condiciones de *pile-up*. [13]

Masa invariante

Se llama masa invariante a la masa en reposo de una partícula. En unidades naturales se puede calcular a partir de la energía de la partícula, E , y de su momento lineal, \vec{p} , mediante la ecuación (3.3). Se llama invariante porque en cualquier sistema de referencia, el valor es el mismo.

$$m^2 = E^2 - |\vec{p}|^2 \quad (3.3)$$

En los experimentos de colisión de partículas, se estudian las partículas que se generan en la desintegración de otra partícula. A partir de la energía y el momento de las partículas resultantes se puede obtener la masa de la partícula inicial. En este caso, se está considerando un sistema de N partículas en el cual se define la masa invariante en unidades naturales, M , del sistema como

$$M^2 = p^2 = \left(\sum_{i=1}^N E_i \right)^2 - \left(\sum_{i=1}^N \vec{p}_i \right)^2 \quad (3.4)$$

donde E_i es la energía de la partícula i -ésima y \vec{p}_i el momento. [47]

3.2. Generación y simulación de sucesos

Tras explicar cómo se reconstruyen e identifican las partículas en CMS, se pasa a la simulación de los sucesos. La simulación de los sucesos nos permite comparar los datos

experimentales con los teóricos para comprobar las teorías propuestas. Para ello se usan las simulaciones de Montecarlo, generando procesos que reproducen los que ocurren en el LHC.

El proceso de simulación de los sucesos tiene varias etapas. En la primera etapa, que recibe el nombre de generación, se simula la colisión y los procesos que ocurren en esta. A continuación, llega a la simulación, que consiste en simular el paso de las partículas previamente generadas por el detector. Y, por último, en la digitalización, a partir de la simulación de las partículas con los subdetectores, se simulan las señales electrónicas. El proceso de digitalización realizado sobre las simulaciones, se utiliza también sobre los datos experimentales para obtener una reproducción fiable de los datos reales.

3.2.1. Generación

Existen diversos programas de generación de sucesos con técnicas de Montecarlo, algunos ejemplos son *MadGraph* [40], *Powheg* [55] o *aMC@NLO* [43], que permite obtener información a segundo orden de teoría de perturbaciones, Next to Leading Order (NLO) y *Pythia* [12] que da información a primer orden, Leading Order (LO). Todos ellos se combinan para reconstruir el resto de interacciones.

Antes de explicar la simulación, se va a introducir de manera breve los procesos que tienen lugar en una colisión protón-protón. En la colisión no choca un solo protón con otro, en el haz hay un elevado número de protones que van a colisionar dando lugar a distintas partículas que forman parte del *pile-up*. Por lo tanto, la colisión da lugar a diferentes procesos. En la figura 3.3 se puede ver las distintas etapas en las cuales se divide la generación de un suceso.

El proceso principal o interacción principal, se denomina *hardscattering* y se da entre los componentes de cada protón, los denominados partones. Se simula a partir de la distribución de densidad de partones (PDF) que nos da la probabilidad de encontrar partones (*quarks* y gluones) en un hadrón como función de la fracción x del momento del protón llevado por el partón.

Después, se tienen las cascadas de partones, *parton-showers*, al tener colisiones hadrónicas, se emiten *quarks* y gluones en forma de radiación. Estos procesos se simulan mediante *Pythia*.

Por otro lado, tenemos el remanente de la colisión, *underlying event*, es la interacción del resto del protón que no tuvo *hardscattering*. Son procesos muy poco energéticos en comparación con la colisión principal, por lo que se identifican de manera sencilla. Protones del mismo haz que produce el *hardscattering*, dan lugar al *pile-up*, son un añadido y empezaron a ser relevantes en el LHC, por eso no aparece en el diagrama 3.3.

Tras la *parton shower*, tenemos la **hadronización** que forma los hadrones a partir de los quarks, del cual se encarga *Pythia*. Por último, se tiene la **desintegración**. Los hadrones que han sido producidos en el punto anterior, pueden desintegrarse diversas partículas. Todos estos procesos dan lugar a numerosas desintegraciones que es necesario simular por lo que tendrá un alto coste computacional.

Una vez se tienen generadas las partículas y los procesos, se realiza una simulación en la cual se obtengan las interacciones de las partículas con el detector.

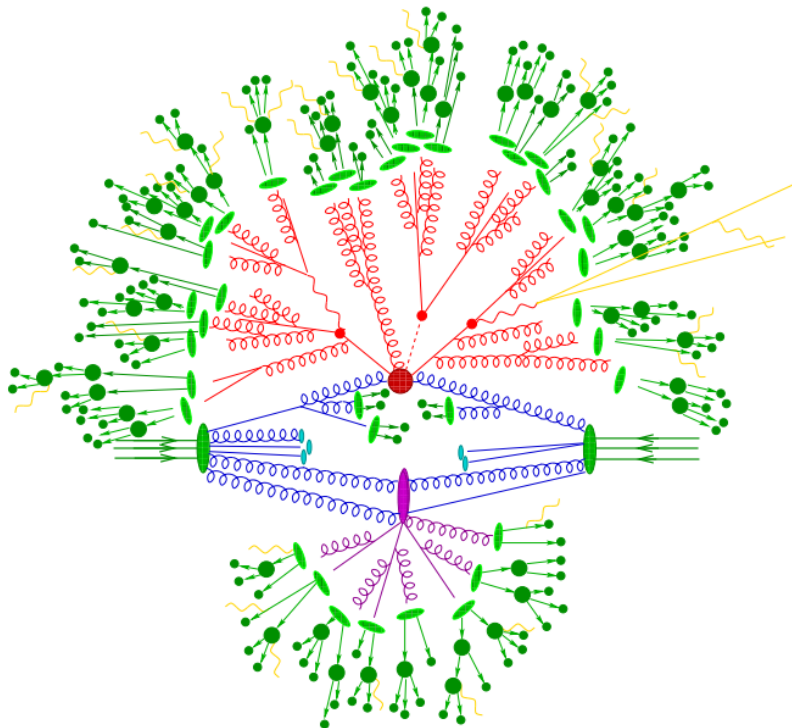


Figura 3.3: Generación de una colisión protón-protón. En rojo se representa la colisión principal, rodeada de las cascadas de partones. En color morado se ve el remanente de la colisión, en verde claro la hadronización y en verde oscuro la desintegración de los hadrones. [39]

3.2.2. Simulación

Una vez que se tiene generado el suceso, se simula el paso de cada una de las partículas producidas por los distintos subdetectores de CMS mediante el *software Geant4* [2]. A partir de los datos con las secciones eficaces de los procesos, los regímenes de energía, la geometría del detector y todas sus características, simula la trayectoria las partículas y su interacción con cada subdetector y genera las deposiciones de energía y las señales electrónicas. El *Geant4* fue creado a partir de la ingeniería de *software* y la tecnología orientada a objetos y se implementó en el lenguaje de programación *C++*. Se utiliza en aplicaciones en física de partículas, física nuclear, diseño de aceleradores, ingeniería espacial y física médica.

3.2.3. Herramientas del *software*

Para la realización del análisis en este trabajo se han utilizado los lenguajes *Python* y *C++* usando ROOT [6]. ROOT es un conjunto de herramientas programadas en *C++*, que permiten trabajar con grandes cantidades de datos, así como realizar análisis estadísticos y visualizaciones. El trabajo se ha desarrollado con el paquete *CMG Tools* dentro del marco de trabajo (*framework*) *CMSSW* [16], que ha sido desarrollado por la colaboración CMS. Así mismo, para la descarga de versiones y desarrollo del código se ha utilizado *Git* y *Github* [37]. Para realizar los ajuste mediante el método de máxima verosimilitud se utilizará la herramienta *Combine* [35].

Capítulo 4

Estudio del proceso $t\bar{t}H$

El objetivo de este trabajo es el estudio del acoplamiento entre el *quark top* y el bosón de Higgs en procesos $t\bar{t}H$ en estados finales con dos leptones del mismo signo. Los datos utilizados para este análisis han sido recogidos por el detector CMS del LHC en colisiones protón-protón a 13 TeV en centro de masas durante los años 2016, 2017 y 2018, con una luminosidad integrada de 36.3, 41.4 y 59.7 fb^{-1} respectivamente, con una luminosidad total de 137 fb^{-1} . Estos datos se comparan con las simulaciones de Montecarlo de la señal y de los fondos.

4.1. Señal y fondo

En colisiones de partículas en experimentos de altas energías solo se conocen los estados finales que surgen de la colisión de partículas. Un solo proceso puede dar lugar a varios estados finales iguales o similares siendo muy difícil identificarlos, por lo que es complicado identificar el proceso que ha tenido lugar en la colisión simplemente analizando las partículas presentes en el estado final.

A la hora de estudiar un proceso, es muy importante distinguir dos conceptos, la señal y el fondo. La señal son los procesos a estudiar en el análisis y los fondos, el resto de procesos que dan lugar a estados finales similares a la señal. Aplicando criterios de selección, estos fondos se pueden reducir con el fin de discriminar la cantidad de sucesos de señal, del fondo, obteniendo una región en la cual predomina la señal, llamada región señal.

4.2. Estados finales

En función del número de leptones (considerando electrones y muones) y τ_H en el estado final, se pueden distinguir diferentes canales multileptónicos. En este trabajo se considerarán estados finales con dos leptones que tengan la misma carga eléctrica y ningún tau, denotado por $2lss + 0\tau_H$. En la figura 4.1 se ve un diagrama de Feynman del proceso $t\bar{t}H$ y su desintegración en un estado final con dos leptones del mismo signo.

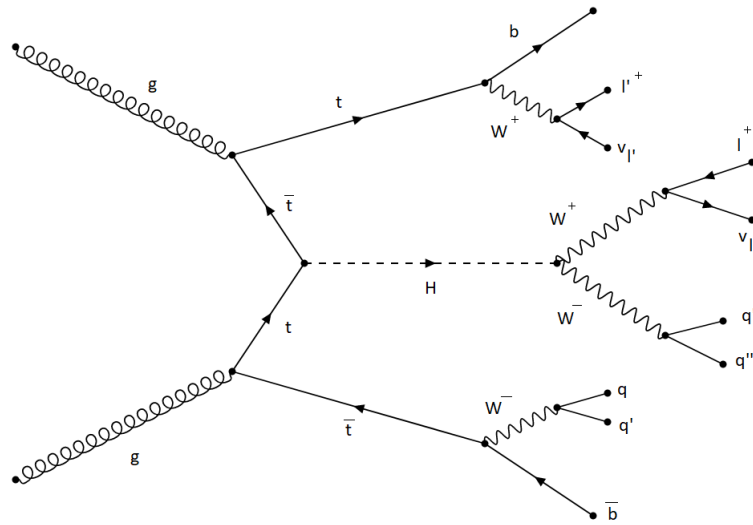


Figura 4.1: Diagrama de Feynman de la producción $t\bar{t}H$ y su desintegración a dos leptones del mismo signo.

Los *quarks top* se desintegran antes de hadronizarse, aproximadamente el 100% de las veces, hacia un *quark b* y un bosón W , el cual se desintegra a un leptón y un neutrino o a *quarks*. Por otro lado, el canal $2lss$ corresponde en su gran mayoría al canal de desintegración donde el bosón de Higgs se desintegra a dos bosones W . Para llegar a tener los dos leptones del mismo signo, o se desintegran los dos bosones W^+ que provienen de la desintegración del Higgs y la desintegración leptónica del top, respectivamente o los dos bosones W^- que provienen de la desintegración del Higgs y el anti-top, respectivamente. Los bosones W que no se desintegren a leptones, lo harán a *quarks* dando lugar a *jets* de partículas. Por lo que la cadena de desintegración completa vendrá dada por $t\bar{t}H \rightarrow 2b + 2W^+ + 2W^- \rightarrow 2b(b\text{-jets}) + 2l(2ss) + 2\nu(E_T^{Miss}) + 4q(\text{jets})$.

El proceso $t\bar{t}H$ tiene una sección eficaz muy baja por lo que es difícil de observar y se encontrará eclipsado por los fondos del análisis, por lo que es necesario hacer

una selección cuidadosa para ser capaces de eliminar los fondos con una sección eficaz mayor y que dificultarán la medida. Usando estados finales que contienen leptones, se pueden discriminar muchos de los fondos que aparecen en el colisionador. Estos fondos se pueden clasificar en fondos reducibles o irreducibles. Los fondos irreducibles son aquellos procesos que producen estados finales muy similares a la señal, suelen estar simulados por métodos de Montecarlo. Se consideran fondos reducibles cuando se tiene al menos un electrón o un muon de tal modo que pasan criterios estrictos de selección pero su origen no es un bosón W . Los fondos reducibles se pueden estimar a partir de los datos o de las simulaciones de Montecarlo.

4.3. Estimación de los fondos

Las contribuciones a los fondos de la región señal del análisis definida por los criterios de selección vienen de varias fuentes.

4.3.1. Fondos irreducibles

Los fondos irreducibles hacen referencia a procesos que producen estados finales muy similares a la señal. Las contribuciones dominantes en estos fondos son debidas a la producción de *quarks top* en asociación con bosones W o Z , $t\bar{t}W$ y $t\bar{t}Z$.

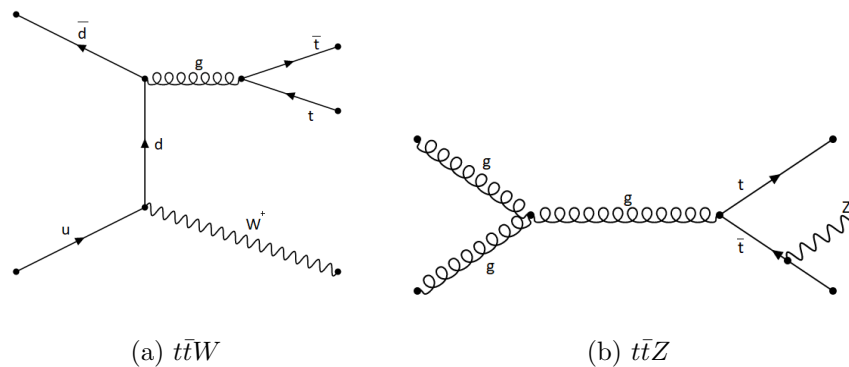


Figura 4.2: Diagramas de Feynman de los procesos $t\bar{t}W$ y $t\bar{t}Z$.

Estos procesos contribuyen al canal $2lss$ con un leptón que viene de la desintegración del *quark top* y otro que viene o del bosón W o del bosón Z . El bosón Z se desintegra a dos leptones de distinto signo, por lo que en principio no contribuiría a estados finales con dos leptones, pero sí se tendría en el caso de que se midiera mal la

carga de uno de los leptones, en ese caso los W del top se desintegrarían a $quarks$. Otra opción es que no se reconstruya uno de los leptones del Z de modo que se tenga un leptón con una carga que proviene del Z y otro (con carga opuesta) que provenga de uno de los W . Para el caso del proceso $t\bar{t}Z$, se puede reducir su contribución aplicando una selección en los sucesos que tengan en el estado final parejas de leptones con masa invariante cercana a la masa del bosón Z . [50]

La producción de WZ y ZZ en sucesos con uno o más jets es otro fondo relevante. Tienen una sección eficaz muy grande en comparación con el proceso $t\bar{t}H$, pero su contribución es menos que la de $t\bar{t}W$ y $t\bar{t}Z$. Esto es debido a que estos procesos se espera que den lugar a un número de $jets$ mucho menor que la señal, y ninguno procedentes de un $quark$ b , por lo que al exigir en los criterios de selección un número de $jets$ alto y al menos uno procedente de un $quark$ b , se pueden reducir estos fondos.

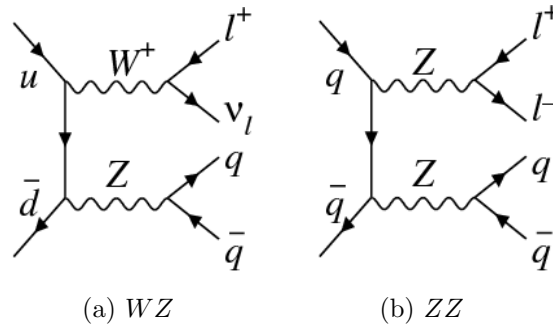


Figura 4.3: Diagramas de Feynman de los procesos WZ y ZZ .

También están en los fondos muestras de tHq , tHW , $t\bar{t}WW$ y las contribuciones más pequeñas que son debidas a procesos con secciones eficaces bajas como la producción de tres bosones o de $t\bar{t}t\bar{t}$, que en su conjunto no son despreciables. Las contribuciones de fondo irreducibles se modelan mediante simulaciones de Montecarlo. [13]

4.3.2. Fondos reducibles

Los fondos reducibles son procesos con una sección eficaz varios órdenes de magnitud más alta que la de $t\bar{t}H$ y que contienen leptones mal identificados, de modo que entran en la selección aplicada. En los fondos reducibles se encuentran procesos como $t\bar{t}$, Drell-

Yan ¹, *single top*, W^+W^- , $W+jets$ y $Z + \gamma$.

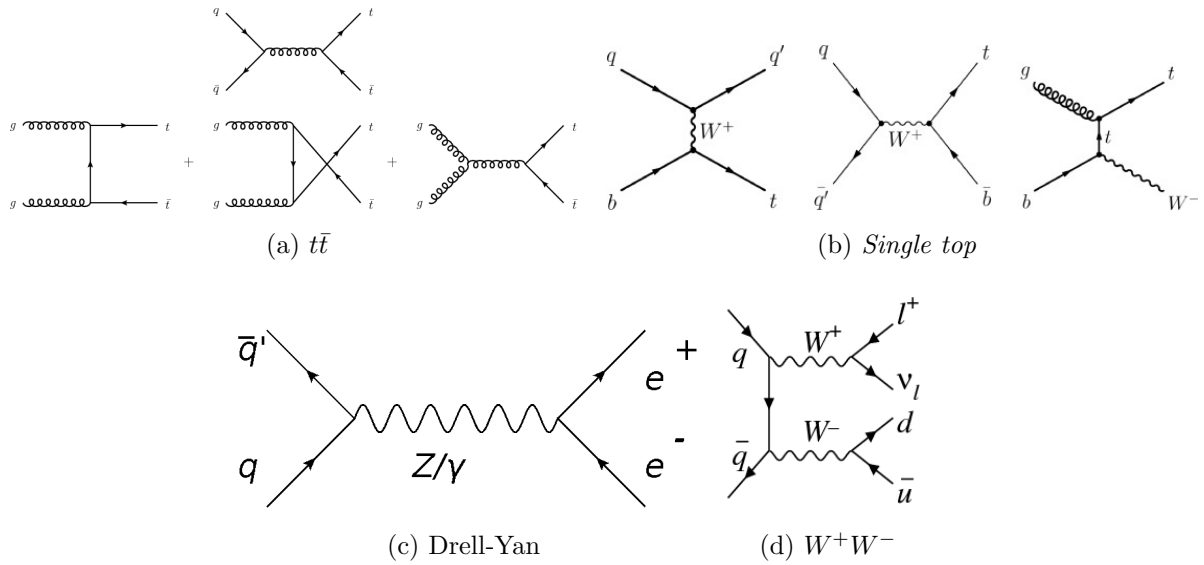


Figura 4.4: Diagramas de Feynman de los procesos $t\bar{t}$, *Single top*, Drell-Yan y W^+W^- .

La sección eficaz para la señal y para los fondos más relevantes, así como los generadores usados para las simulaciones se observa en el cuadro 4.1. En el canal $2lss + 0\tau_H$ se pueden distinguir tres fuentes de fondos reducibles: [13]

- La mala identificación de los leptones (*fakes*). Los leptones mal identificados se refieren a sucesos en los que al menos un electrón o muon reconstruido es causado por la identificación errónea de un leptón o hadrón reconstruido, que surge de la identificación errónea de un *quark* o *jets* de gluón. La principal contribución a este fondo proviene de la producción de $t\bar{t} + jets$, lo que refleja la gran sección transversal de este proceso de fondo.
- Fotones que sufren el proceso $\gamma \rightarrow e^+e^-$ (*conversiones*) en la cual un electrón o positrón se lleva la mayor parte de la energía, mientras que el otro es de baja energía y se reconstruye de manera errónea.
- La medida incorrecta de la carga de un leptón (*flips*). Consiste en sucesos en los que se mide mal la carga de un leptón reconstruido. La principal contribución al fondo *flips* proviene de los sucesos $t\bar{t} + jets$ en los que ambos *top quarks* se desintegran.

¹Proceso que ocurre en la dispersión de hadrones de alta energía. Se da cuando un *quark* de un hadrón y un anti-*quark* de otro hadrón se aniquilan dando lugar a un fotón virtual o un bosón Z que decae en un par de leptones de carga opuesta.

Proceso	σ (fb)	Generador	Proceso	σ (fb)	Generador
$t\bar{t}H$	507	<i>Powheg</i>	$t\bar{t} + jets$	$8.33 \cdot 10^5$	<i>Powheg</i>
tHq	74.3	<i>Madgraph</i>	Drell-Yan	$6.23 \cdot 10^7$	<i>aMC@NLO</i>
tHW	15.2	<i>Madgraph</i>	WW	$1.19 \cdot 10^5$	<i>Powheg</i>
$t\bar{t}Z$	839	<i>aMC@NLO</i>	WZ	$4.50 \cdot 10^4$	<i>Powheg</i>
$t\bar{t}W$	650	<i>aMC@NLO</i>	ZZ	$1.6 \cdot 10^4$	<i>Powheg</i>
$t\bar{t}WW$	6.98	<i>Madgraph</i>			

Cuadro 4.1: Secciones eficaces de los fondos más importantes en este análisis y los generadores usados para simularlo. El generador de *Pythia* se combina con los anteriores para simular los procesos. [13]

El proceso $t\bar{t}$ es el que da lugar a una mayor contribución al fondo reducible de los leptones *non-prompt* (no vienen de un bosón W) en el canal $2lss + 0\tau_H$. Puede contribuir al fondo de los *flips*, cuando los dos *quarks top* se desintegran a dos leptones y se mide mal la carga de uno de ellos, y al fondo de *conversions* cuando se produce este proceso con fotones ($t\bar{t}\gamma$). Este proceso tiene una sección eficaz alta ($833 pb$) por lo que en comparación con el proceso de señal, da lugar a una contribución importante a los sucesos totales.

El proceso Drell-Yan (DY) tiene lugar cuando dos *quarks* se aniquilan dando lugar a un fotón o a un bosón Z que se desintegra a leptones. Este proceso tiene una sección eficaz alta ($6230 pb$) y puede ser que se identifiquen mal algunas de las partículas producidas, o que se mida mal la carga de un leptón. El proceso DY tiene mayor contribución en la región con un menor número de *jets*, por lo que se puede reducir aplicando ciertos criterios de selección.

Los tres tipos de fondo reducible se excluyen mutuamente al dar preferencia al *fakes* sobre los *flips* y *conversions* y al dar preferencia a los *flips* sobre los *conversions* cuando un suceso se clasifica para más de un tipo de fondo reducible. Los fondos *fakes* y *flips* se determinan a partir de datos. La estimación de este número de sucesos a partir de los datos se realiza en una muestra ortogonal a la región de señal donde se realiza la medida. Posteriormente, se extrapola para poder aplicar la estimación de la región de señal sin inducir sesgos. Sin embargo, el fondo de *conversions* se modela utilizando la simulación de Montecarlo. [13]

4.4. Selección de sucesos

Según el estado final buscado para el análisis y los fondos presentes en el proceso, se aplican a los datos y las simulaciones unos criterios de selección para discriminar los sucesos que corresponden a la señal, de los fondos. En este análisis se realizará una selección de sucesos con el fin de detectar las señales de los procesos $t\bar{t}H$ y tH en estados finales multileptónicos, en concreto en el canal $2lss + 0\tau_H$.

4.4.1. Canal $2lss + 0\tau_H$

Este canal está formado por estados finales que tengan dos leptones del mismo signo, es decir, que tenga la misma carga eléctrica. Es uno de los canales más sensibles en las medidas de estos procesos, y contiene suficientes sucesos de $t\bar{t}H$ para ser sensible a las propiedades cinemáticas que son necesarias para buscar una violación CP, que se detallará en la sección 5.2. Para obtener el estado final buscado se aplican una serie de criterios de selección en la señal y el fondo.

A continuación se detallará la elección de los criterios de selección usados en este análisis, tomando como referencia las distribuciones para los observables físicos del año 2016. Una vez establecidos los criterios de selección, se usarán los mismos para la región señal establecida a partir de la combinación de los datos de los tres años estudiados en el análisis. Inicialmente se aplicaron los siguientes criterios de selección:

- Para el *trigger* de uno y dos leptones, el momento transversal del leptón más energético tiene que ser mayor que 25 GeV y el momento transversal del leptón menos energético, mayor que 15 GeV.
- Solo puede haber dos leptones con requisitos de calidad altos en su selección. Según el estado final que se quiere obtener, se necesitan dos leptones *prompt* (vienen de un bosón W).
- El signo de la carga eléctrica ha de ser el mismo, esto se impone con la condición que el producto de sus cargas sea mayor que cero. Esta condición reduce de manera considerable el fondo de $t\bar{t} + jets$. El canal $2lss$ contiene sucesos en los cuales un leptón se origina de la desintegración del bosón de Higgs, mientras que el otro

viene de la desintegración de un *quark top*. Imponer que los dos leptones tengan la misma carga reduce el rendimiento de la señal a la mitad, pero aumenta la relación señal-fondo en gran medida al eliminar el fondo que surge de la producción de $t\bar{t} + jets$ con desintegraciones dileptónicas de los *quarks top*. Por lo que la relación señal-fondo es más favorable para los sucesos con pares de leptones del mismo signo.

En la figura 4.5 se representan el números de sucesos frente algunas variables como el momento del leptón más energético, primer leptón (*leading lepton*) y del segundo leptón más energético (*subleading lepton*), así como la masa invariante de los dos leptones tras haber hecho la primera selección. Las contribuciones esperadas de cada proceso se encuentran apiladas en diferentes colores.

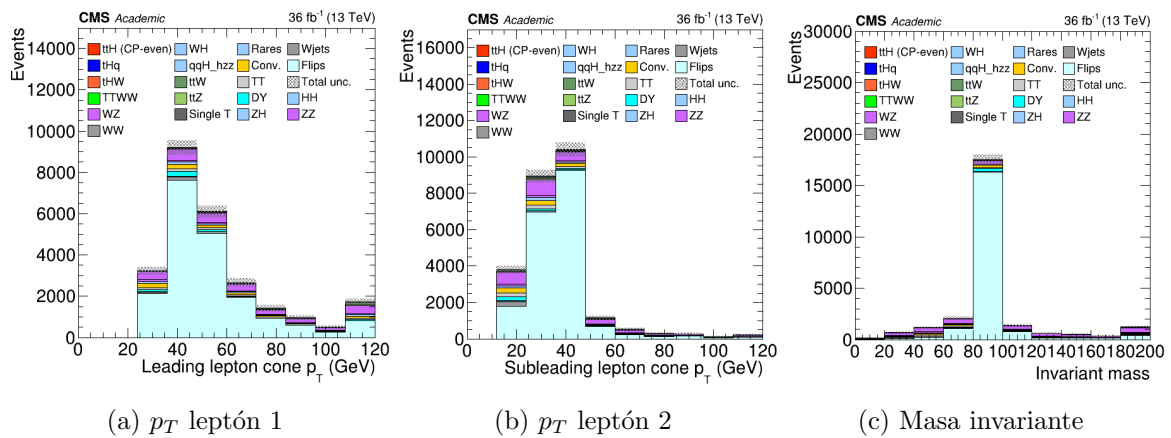


Figura 4.5: Representación del momento transverso de los dos leptones más energéticos y de la masa invariante del par de leptones en el caso de aplicar criterios de selección en el momento y la carga eléctrica.

Tras aplicar los criterios de selección detallados anteriormente, no se detecta la señal del proceso $t\bar{t}H$, pero sí se observa una gran contribución de diversos fondos, predominando los *flips*. Para reducir los fondos y lograr obtener la señal deseada, se añadieron los siguientes criterios:

- Se introducen criterios de calidad en la medida de la carga de los leptones lo que hace reducir el fondo de procesos que dan lugar a leptones de carga opuesta donde la carga de uno de ellos se ha determinado incorrectamente. Las parejas de signo contrario se producen más a menudo, como ejemplo, en las desintegraciones de partículas neutras como el bosón Z .

- Para eliminar las contribuciones de los fondos que surgen de $t\bar{t}Z$, tZ , WZ y DY , se introduce un veto sobre el pico del bosón Z . Imponiendo que $|m_{ll} - m_Z| > 10$ donde $m_Z = 91.2$ GeV es la masa del bosón Z y m_{ll} es la masa invariante del par de leptones de carga opuesta, se vetan los procesos en los cuales la masa invariante del par de leptones dista menos de 10 GeV de la masa del bosón Z . Así se reduce la contribución de los procesos que contienen el bosón Z , ya sea de señal o de la desintegración del bosón de Higgs a dos bosones Z .

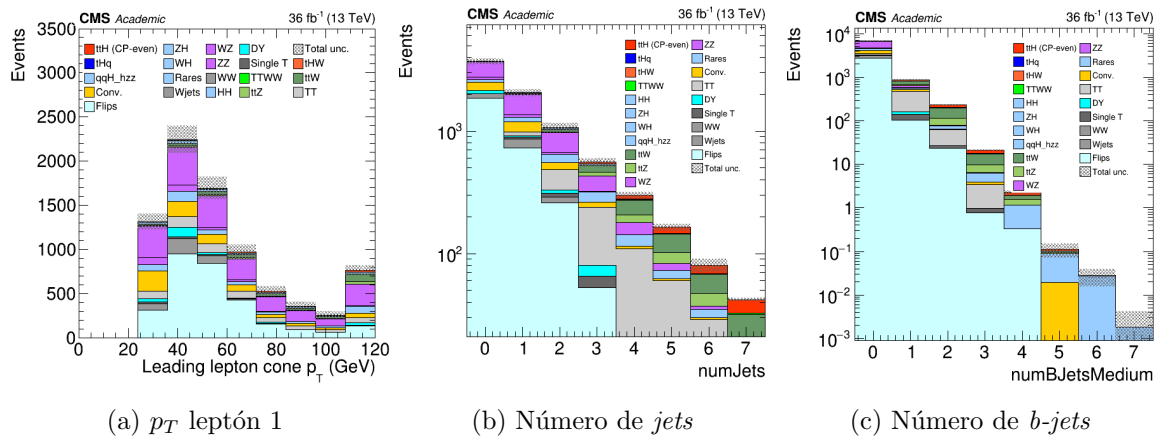


Figura 4.6: Representación del momento transversal del primer leptón y el número de *jets* y *b-tag* en el caso de aplicar todos los criterios previamente descritos. La señal $t\bar{t}H$ está representada en color rojo.

En la figura 4.6 se representan los números de sucesos frente al momento transversal del leptón más energético, el número de *jets* y el número de *b-tag*. Las contribuciones esperadas para cada suceso se encuentran apiladas en distintos colores, la contribución en rojo corresponde a la señal de $t\bar{t}H$. Comparando la representación del leptón más energético con los primeros criterios aplicados, figura 4.5a, con la figura 4.6a, se ve una disminución considerable en el número de sucesos debido a los criterios aplicados hasta este punto. El fondo que más se ha reducido ha sido el fondo de los *flips*, ya que al introducir un criterio de calidad a la hora de medir la carga de los leptones, se reduce el fondo en el cual la carga de uno de los leptones se ha medido incorrectamente. Como se puede ver en la figura 4.6b, y se comprobará más adelante, la señal puebla regiones con alto número de *jets* y *b-tag*. Esto indica que será necesario establecer un criterio de selección en el número de *jets* en el estado final, ya que la proporción de la señal respecto al fondo es muy pequeña. La proporción señal-fondo es $83.04 \pm 0.83 / 7431.29 \pm 263.94$, por lo que aún hay demasiado fondo en los datos.

A continuación, se determinará el criterio de selección que se aplicará al número de *jets*. Para ello, se determinarán todas las combinaciones posibles de *jets* y de *b-tag*, dadas en el cuadro 4.2. Con esta clasificación, se obtiene la figura 4.7, donde están representados en el eje de ordenadas el números de sucesos que contienen *jets*, en escala logarítmica. Por otro lado, en el eje de abscisas están representados el número de sucesos que tienen la combinación de *jets* y *b-tag* descrita en el cuadro 4.2.

<i>Bines</i>	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
<i>jets</i>	0	1	1	2	2	2	3	3	3	4	4	4	5	5	5	>5
<i>b-jets</i>	0	0	1	0	1	2	0	1	2	0	1	2	0	1	2	2

Cuadro 4.2: Número de *jets* en cada *bin*.

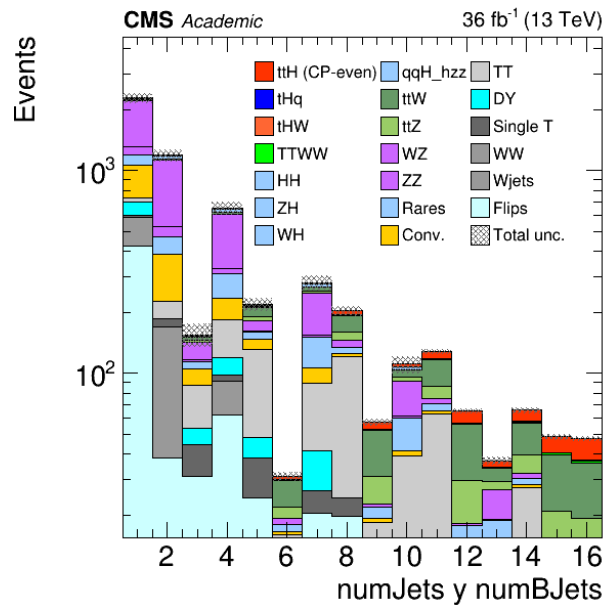


Figura 4.7: Representación del número de *jets* y número de *b-tag* para cada una de las combinaciones dadas en el cuadro anterior. Los sucesos esperados para cada contribución vienen apilados en distintos colores. La señal $t\bar{t}H$ está representada en color rojo.

Se puede ver como a medida que aumenta el número de *jets*, se reducen los fondos *flips*, *Drell-Yan*, *conversions* y $t\bar{t}W$. Por lo tanto, queda establecido que el proceso $t\bar{t}H$ se aprecia con un número de *jets* elevado y al menos un *b-tag*. En el cuadro 4.3 se ve la proporción señal-fondo en cada *bin*. Esta empieza a ser apreciable a partir del *bin* número 11, lo que corresponde con tener 4 o más *jets* y al menos 1 *jet* procedente de

la hadronización de un *quark* b . Por lo tanto, este será el criterio de selección que se aplicará en el número de *jets*.

<i>Bines</i>	8	9	10	11	12	13	14	15	16
Señal	7.46	4.40	3.17	9.06	8.01	2.58	7.83	7.88	10.14
Fondo	170.79	51.01	105.83	107.14	54.03	34.84	56.72	37.27	34.66
Porporción	0.043	0.08	0.029	0.084	0.148	0.074	0.138	0.211	0.2925

Cuadro 4.3: Proporción señal-fondo en cada *bin*.

Se añadirán los siguientes criterios de selección:

- El suceso ha de tener 4 o más *jets* con momento transverso superior a 25 GeV, $p_T \geq 25$ GeV, y al menos uno de ellos debe proceder de la hadronización de un *quark* b .
- Se vetan sucesos que tengan parejas de leptones cuya masa invariante sea menor que 12 GeV, ya que están mal modelizados por las simulaciones de Montecarlo.
- El número de leptones τ_H identificados tiene que ser 0. De esta manera se tiene el canal $2lss + 0\tau_H$. Esto permite eliminar del análisis cualquier suceso que tenga un τ_H , que en gran parte vienen de la desintegración del Higgs a dos τ_H .

El cuadro 4.4 recoge la selección completa que se aplicará en la realización de este análisis.

Criterio	$2lss + 0\tau_H$
p_T de los leptones	≥ 25 GeV ≥ 15 GeV
Identificación de leptones	Sólo 2 leptones <i>tight</i>
Criterio sobre la carga	Criterios de calidad + mismo signo
<i>Jets</i>	≥ 4
<i>b-tag</i>	≥ 1
τ_H	0 τ_H
Masa invariante de los leptones	$ m_Z - m_{ee} \geq 10$ GeV $m_{ll} \geq 12$ GeV

Cuadro 4.4: Selección de sucesos en el canal $2lss$.

Una vez se han aplicado todos los criterios de selección, representando el número de sucesos frente al momento transverso del primer leptón, se obtiene la figura 4.8. En ella las contribuciones esperadas de cada proceso se encuentran apiladas en diferentes colores. Se ve como se han eliminado gran parte de los fondos y se puede apreciar la señal $t\bar{t}H$ con claridad en color rojo. Esta será la selección completa que se aplicará para el desarrollo del análisis.

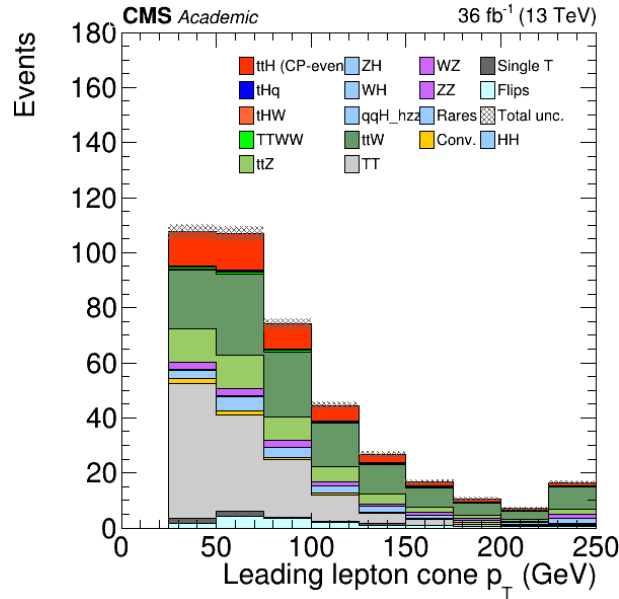


Figura 4.8: Representación del número del momento transverso del primer leptón tras aplicar todos los criterios de selección.

4.5. Fuentes de incertidumbre

En esta sección se detallarán las diferentes fuentes de incertidumbre consideradas en este trabajo. Al llevar a cabo una medida, es necesario tener en cuenta las limitaciones de nuestro sistema experimental así como de los cálculos teóricos usados. Estas se tienen en cuenta como incertidumbres que nos darán un rango en el que nuestra medida es válida. Las incertidumbres se clasifican en estadísticas y sistemáticas.

4.5.1. Incertidumbres estadísticas

Las incertidumbres estadísticas asociadas a un experimento surgen del hecho de que cada medida tiene fluctuaciones aleatorias por la naturaleza estocástica de los procesos. Los procesos simulados mediante los métodos de Montecarlo también tienen asociada

una incertidumbre estadística, que será menor cuanto mayor sea el número de procesos. Si una variable sigue una distribución de Poisson, la incertidumbre estadística asociada a ella será \sqrt{N} , donde N es el número de sucesos observados que pasan la selección.

4.5.2. Incertidumbres sistemáticas

Las fuentes de incertidumbre sistemáticas son debidas al método o instrumento usado. Las fuentes experimentales comprenden las incertidumbres en las mediciones auxiliares, que se realizan para validar o si es necesario, para corregir el tratamiento de los distintos objetos usados en el análisis que se detallan más adelante. La estimación de los fondos dada por la simulación están afectados por incertidumbres teóricas, surgen principalmente de la falta de correcciones de orden superior a las expansiones perturbativas empleadas para el cálculo de las secciones eficaces y de las incertidumbres asociadas a las PDF. Las fuentes de incertidumbre sistemáticas se modelizan como parámetros de confusión en el ajuste de máxima verosimilitud. [13]

- **Eficiencia del *trigger*:** La eficiencia del *trigger* tiene una incertidumbre asociada, basada en la presencia de uno, dos o tres leptones (electrones o muones) y se miden como función del momento transversal, basada en la detección de un leptón con un p_t mínimo. Tiene una incertidumbre en el rango de 1 % y 2 %.
- **Identificación de los leptones:** Los algoritmos de reconstrucción e identificación de leptones permiten identificar electrones y muones de forma precisa. No obstante, no son infalibles, pues están sujetos a las limitaciones del detector: hay ciertas regiones que no están instrumentadas, los detectores tienen una cierta resolución espacial y temporal, etc. todo esto da lugar a una cierta pérdida de la eficiencia que se debe tener en cuenta, por lo tanto, tienen unas incertidumbres sistemáticas asociadas. Estas incertidumbres dependen del momento del leptón, de la pseudorapidez y de lo fuerte que sean los criterios de calidad aplicados sobre el leptón.
- **Energía de los *jets*:** La detección e identificación de *jets* lleva asociada una incertidumbre en la medida de su energía dependiendo del momento del *jet* y de η . El efecto resultante en la señal y el fondo esperado se calcula variando las

energía de los *jets* en los sucesos simulados con sus incertidumbres, recalculando todos los observables cinemáticos y aplicando de nuevo los criterios de selección. El efecto de las incertidumbres en la resolución de energía de los *jets* se evalúa de una manera similar, pero es más pequeña que el efecto de las incertidumbres en la escala de la energía de los *jets*.

- **Eficiencia de *b*-tagging:** La eficiencia de etiquetar los *quarks b* tiene una incertidumbre que depende del momento del *jet* y de la pseudorapidez. Los algoritmos no identifican todos los *jets* procedentes de un *quarks b* con una eficiencia total, pueden confundir los *jets* procedentes de un *quarks b* con *jets* ligeros y viceversa.
- **Luminosidad:** La luminosidad es otra fuente de incertidumbre en las medidas. Esta incertidumbre está asociada a la medida de la luminosidad del LHC que recoge el experimento CMS y proporciona como dato. Según el año de la toma de datos esta incertidumbre es diferente. Para el año 2016, la incertidumbre asociada a la luminosidad integrada recogida por CMS es de un 2.5 %. [14] La luminosidad integrada para los años 2016, 2017 y 2018 tienen un 1.2 – 2.5 % de incertidumbre individual, mientras que la incertidumbre total para los tres años es de 1.6 %.
- ***Pile-up*:** Las incertidumbres asociadas a los procesos de *pile-up* se calcula variando el número de colisiones inelásticas entre protones que se impone en los sucesos simulados. Así se puede estimar como el *pile-up* afecta a las medidas de diferentes procesos. El efecto en la señal $t\bar{t}H$ y tH y de las contribuciones de fondo modelizando con las simulaciones de Montecarlo es de un 1 %.
- **Simulaciones de Montecarlo:** La modelización de la señal y en fondo a partir de las simulaciones de Montecarlo también lleva una incertidumbre asociada. Esto es debido a que la simulación se hace a partir del cálculo de la teoría de perturbaciones y no se pueden tener en cuenta todos los órdenes, se suelen hacer las aproximaciones a LO o NLO.
- Las incertidumbres en las secciones eficaces son relevantes para el propósito de medir la producción con respecto a lo esperado del modelo estándar, y en el caso de los procesos $t\bar{t}H$ y tH , es relevante para marcar los límites en los acoplamientos del bosón de Higgs al *quark top*.

Capítulo 5

Resultados experimentales

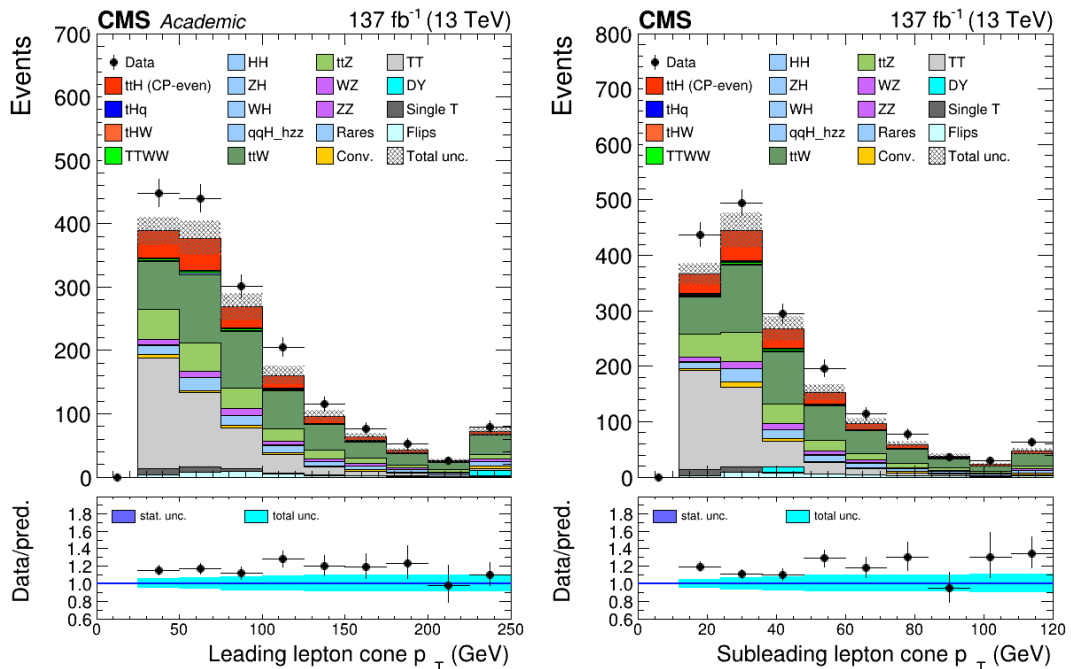
En este capítulo se recogerán los resultados del estudio del acoplamiento de Yukawa entre el *quark top* y el bosón de Higgs en el canal $2lss + 0\tau_H$. Los criterios de selección aplicados en este canal, se han definido en el capítulo anterior y están recogidos en el cuadro 4.4. Primero se determinarán las distribuciones de los observables físicos definidos en la región señal. Se introducirá la discriminación de los escenarios CP, así como las variables que se usan para discriminar las contribuciones de los escenarios. A continuación, se realizará una clasificación de los sucesos con el fin de separar aún más la señal del fondo y, también, para ser capaz de distinguir los dos escenarios CP que se buscan. Por último, se extraerá el parámetro f_{CP} que permite estudiar el acoplamiento de Yukawa entre el *quark top* y el bosón de Higgs.

5.1. Región de señal

La región de señal recoge las distribuciones de los observables físicos definidos según los criterios de selección previamente descritos en el cuadro 4.4. A continuación se representarán las distribuciones de algunos de los observables más importantes para la región señal establecida para el conjunto de los años estudiados en este análisis: 2016, 2017 y 2018.

En la figura 5.1, se representan el número de sucesos frente al momento transverso del primer y segundo leptón más energéticos, *leading lepton* y *subleading lepton*, respectivamente para el conjunto de los años estudiados en el análisis. En la figura 5.2 se representan el número de sucesos frente al número de *jets* y *b-tag* etiquetados como *medium*.

En las gráficas superiores se tiene en el eje de ordenadas el número de eventos y en el eje de abscisas el observable correspondiente que se ha representado. En distintos colores se tienen apiladas las contribuciones esperadas cada proceso. Los puntos indican el número de sucesos de datos tomados por el detector CMS junto con su incertidumbre estadística representada con una barra. La señal del proceso $t\bar{t}H$ se representa en color rojo. En las gráficas inferiores se representa el cociente entre los datos experimentales y la predicción simulada para cada caso, además también se tiene la incertidumbre estadística y la total, siendo la total la suma de la estadística más la sistemática. Se ve como los fondos que mayor contribución tienen son $t\bar{t}W$ y $t\bar{t}$.



(a) Momento transverso del *leading lepton* (b) Momento transverso del *subleading lepton*

Figura 5.1: Representación de la distribución del momento transverso del leptón más energético y del segundo leptón más energético para el conjunto de los tres años estudiados en este análisis.

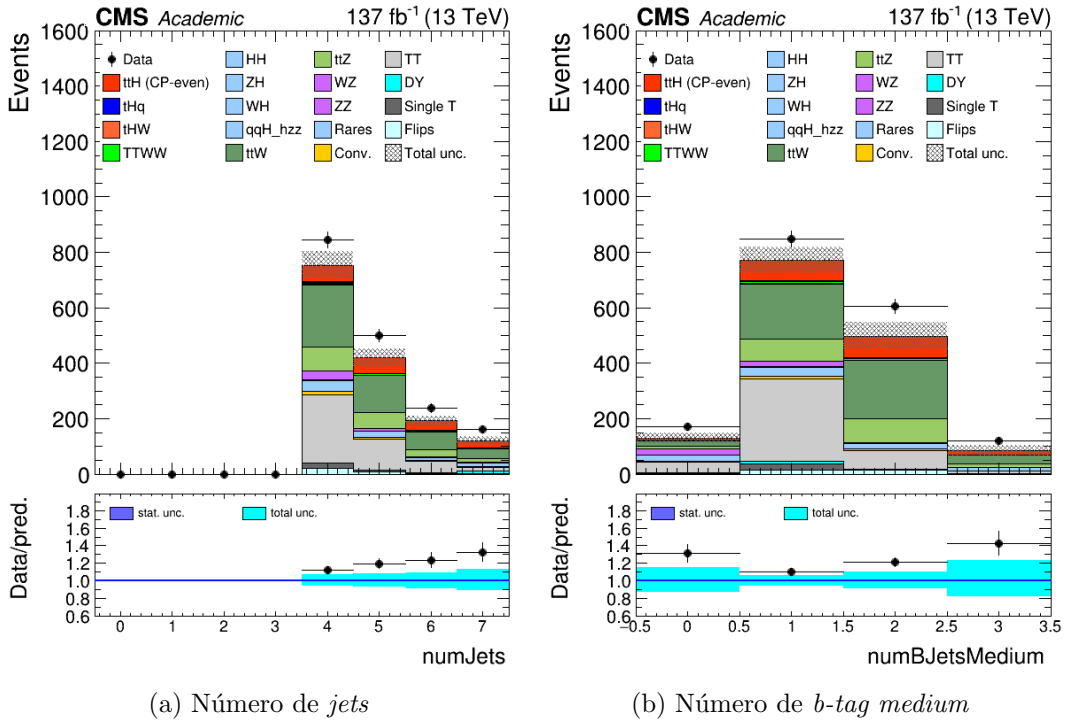


Figura 5.2: Representación del número de *jets* y *b-tag medium* para el conjunto de los tres años estudiados en este análisis.

En el cuadro 5.1 se pueden ver el número de sucesos de señal y de los distintos fondos, así como sus incertidumbres para el canal considerado usando las simulaciones de Montecarlo.

Proceso	2016+2017+2018	Proceso	2016+2017+2018
$t\bar{t}H$	172.03 ± 3.74	tHq	4.74 ± 0.25
tHW	4.26 ± 0.29	$t\bar{t}WW$	13.07 ± 0.38
WH	7.21 ± 3.71	ZH	1.02 ± 0.90
$t\bar{t}W$	459.44 ± 5.91	WZ	50.7 ± 5.16
$t\bar{t}Z$	182.78 ± 8.27	ZZ	2.68 ± 1.11
DY	8.91 ± 8.91	$t\bar{t}$	412.32 ± 23.14
<i>Single t</i>	26.94 ± 9.4	<i>Flips</i>	35.14 ± 5.97
<i>Conversions</i>	26.82 ± 8.70	<i>Rares</i>	85.70 ± 7.50
Fondos	1321.91 ± 38.67	Datos	1742

Cuadro 5.1: Sucesos considerados en el análisis para el conjunto de los años 2016, 2017 y 2018 usando las simulaciones de Montecarlo.

5.1.1. Regiones de control

Para determinar la normalización de los procesos $t\bar{t}Z$, $t\bar{t}W$, WZ y ZZ se usarán las regiones de control enriquecidas en estos procesos. Se definen dos regiones de control, $3l - CR$ y $4l - CR$, usando regiones con tres y cuatro leptones en el estado final, respectivamente. Para ello se usa una selección que hace que estas regiones estén enriquecidas en los procesos $t\bar{t}Z$ y WZ para la región de 3 leptones y ZZ para la región de 4 leptones. Estas regiones se incluyen en los ajustes para dar una estimación de estos fondos. La región de control para el fondo de $t\bar{t}W$ se construye a partir del canal $2lss + 0\tau_H$ usando la clasificación en nodos dada por las redes neuronales que se explica más adelante en la sección 5.3.

A pesar de la selección descrita anteriormente, las regiones todavía están dominadas por los fondos, siendo los que más contribución tienen, $t\bar{t}W$ y $t\bar{t}$. Una vez se tiene definida la región de señal, se usarán técnicas de análisis multivariante, en particular, redes neuronales entrenadas para lograr separar la contribución de los diversos procesos de señal y de fondo. Estas redes neuronales proporcionarán como salida, varios nodos que estiman la probabilidad de que cada suceso corresponda a $t\bar{t}H$, tHq , $t\bar{t}W$ u otro fondo. Para el nodo de $t\bar{t}H$ dado por la salida de las redes neuronales, se aplicarán técnicas de análisis multivariante que discriminan los escenarios CP par y CP impar. Los discriminantes usados se determinarán observando como varían las propiedades cinemáticas según el escenario CP usado.

5.2. Discriminación CP

En el modelo estándar, el bosón de Higgs es par bajo la inversión carga-paridad (CP), lo que quiere decir que cambiando de signo a la carga e invirtiendo la paridad, se obtiene el mismo signo del escalar. Determinar la estructura CP de los acoplamientos del bosón es relevante, la presencia de un término impar en CP en el lagrangiano¹ del bosón de Higgs sería una indicación directa de presencia de nueva física. Las colaboraciones ATLAS y CMS han estudiado los acoplamientos de bosones en búsqueda

¹Operador a partir del cual se puede describir un sistema físico. A partir de él se puede obtener la evolución temporal o las leyes dinámicas del sistema.

de violación CP. Todos estos estudios fueron compatibles con un lagrangiano del bosón de Higgs puramente par, siguiendo de esta manera las predicciones del modelo estándar.

El acoplamiento de Yukawa del bosón de Higgs a los fermiones ha sido determinado con una gran incertidumbre durante el *Run 1*. A lo largo, del *Run 2*, el acoplamiento de Yukawa entre el *quark top* y el bosón de Higgs, y_t , dividido por el valor esperado por el modelo estándar, descrito como $\kappa_t = y_t/y_y^{SM}$, ha sido determinado por la colaboración CMS en el rango $-0.9 < \kappa_t < -0.7$ ó $0.7 < \kappa_t < 11$ con un 95 % de nivel de confianza. La producción del bosón de Higgs en asociación con dos *quarks top* o un *quark top* da acceso a la magnitud y el signo de y_t . [50]

El lagrangiano de la interacción fermión-Higgs se puede escribir como una superposición de una fase CP par e impar:

$$\mathcal{L} = \mathcal{L}_{CP-even} + \mathcal{L}_{CP-odd} \quad (5.1)$$

donde cualquier desviación de los valores esperados por el modelo estándar de los acoplamientos significaría una violación de la simetría CP y se podría describirse como un fenómeno más allá del modelo estándar. El lagrangiano de $t\bar{t}H$ puede parametrizarse como: [50]

$$\mathcal{L}_{t\bar{t}H} = \frac{m_t}{v} \bar{\phi}_t (\kappa_t + i\gamma_5 \tilde{\kappa}_t \phi_t) H \quad (5.2)$$

Aquí, m_t es la masa del *quark top* y v es el valor esperado del vacío en el campo de Higgs. El parámetro κ_t representa la parte par del acoplamiento entre el bosón de Higgs y el *quark top* y $\tilde{\kappa}_t$, la parte impar. En el Modelo Estándar $\kappa_t = 1$ y $\tilde{\kappa}_t = 0$ y en el caso $\kappa_t = 0$ y $\tilde{\kappa}_t = 1$ se estaría en la hipótesis de un escenario completamente impar. Para cualquier otros valores de κ_t y $\tilde{\kappa}_t$ se estará en un escenario mezcla.

La separación de las contribuciones par e impar de la señal y de los fondos se realiza mediante técnicas de aprendizaje automático. En particular, los árboles de clasificación (BDTs) se emplean para mejorar la separación de los escenarios CP par e impar tanto puros como mixtos de la señal de $t\bar{t}H$. Para ello, se usan el 50 % de las muestras para entrenar los BDTs, el 40 % de *test* y el 10 % como validación durante el entrenamiento.

En la figura 5.3 se representan algunos de los observables físicos que se usan como variables de entrada por el BDT para discriminar entre los distintos escenarios para la señal. En color rojo se muestra la contribución de $t\bar{t}H$ del escenario par y en color azul, la contribución del escenario impar. Una vez representadas todos los observables se ve que las que más diferencia tienen entre los escenarios CP par e impar son aquellas que describen la cinemática del proceso. Estos observables físicos son los que se usan para entrenar el BDT que separará las contribuciones de los procesos de señal y fondo según las contribuciones de los escenarios CP par y CP impar.

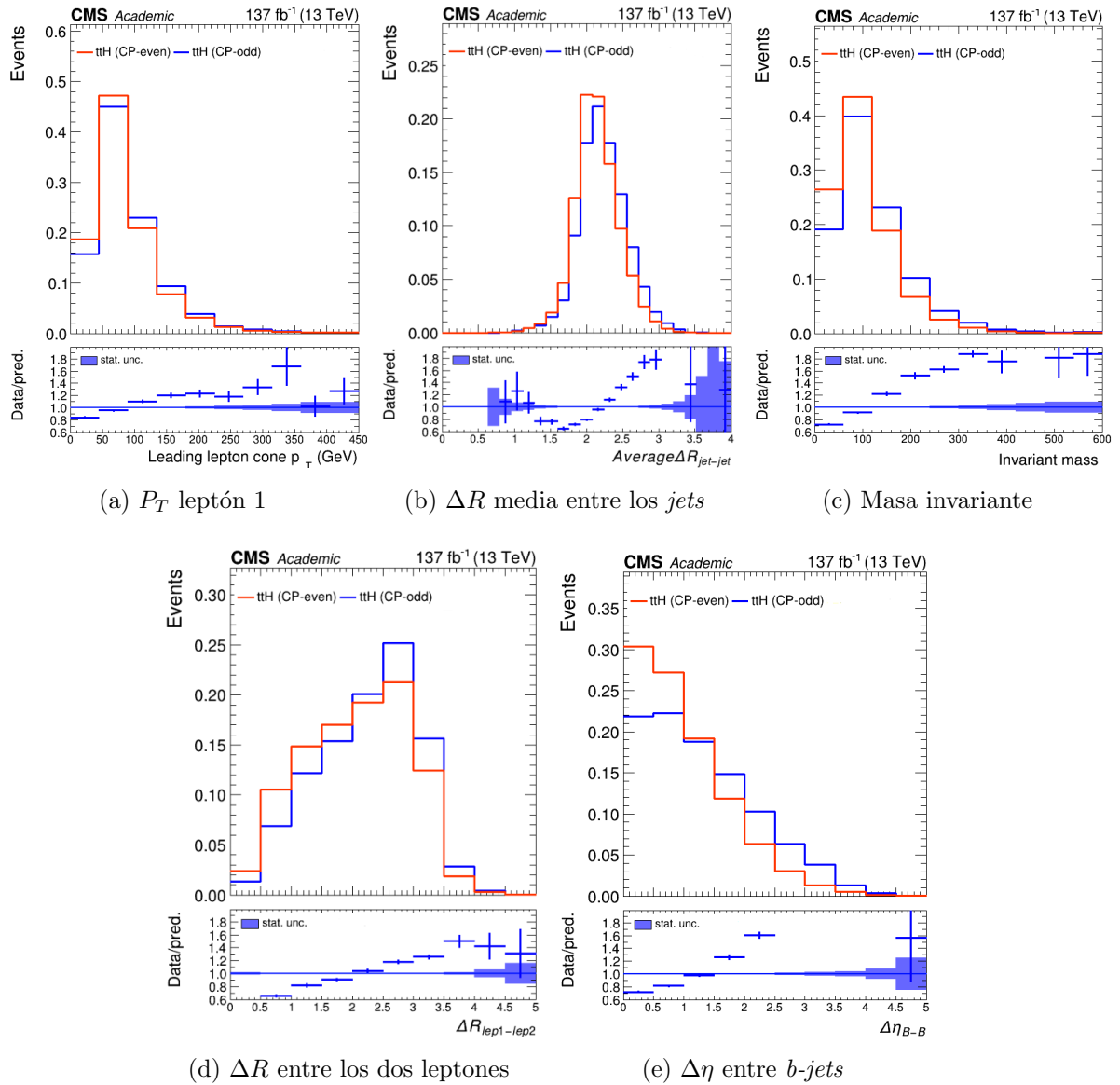


Figura 5.3: Observables físicos usados para la discriminación de escenarios en el canal $2lss + 0\tau_H$.

A partir de todos los observables físicos que discriminan los procesos, se obtiene el discriminante de los escenarios CP par e impar como combinación de estos observables. En la figura 5.4, el discriminante de los escenarios se construye con todos los observables que entrena los BDTs y da como salida la información sobre si el proceso es más parecido a un escenario CP par o impar.

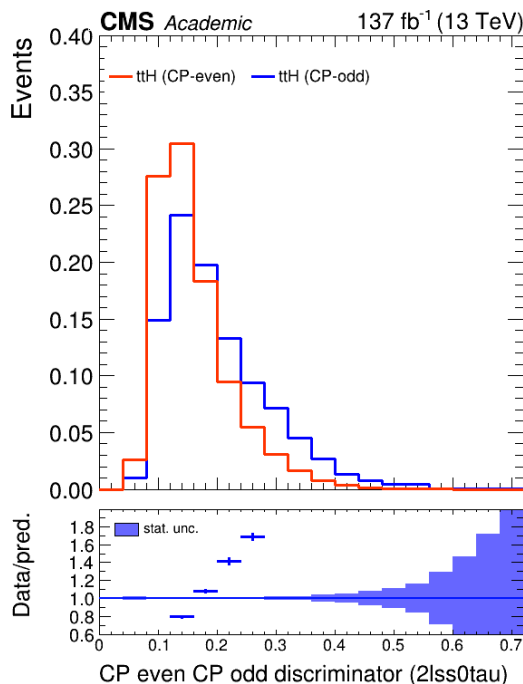


Figura 5.4: Discriminante como combinación de diferentes observables físicos discriminantes.

5.3. Clasificación

Se usan redes neuronales entrenadas para discriminar los procesos de señal y fondo. Estas redes neuronales darán como salida distintos nodos que indican la probabilidad de que cada suceso corresponda a $t\bar{t}H$, tHq a $t\bar{t}W$ u otro fondo. Por lo tanto, los sucesos se clasifican de acuerdo al nodo más probable.

Mediante el uso de las redes neuronales, se separarán los procesos en distintos nodos y se realizará una comparación con los resultados experimentales. La separación según los distintos nodos se puede ver en la figura 5.5. En ella se tienen los cuatro nodos obtenidos de la clasificación de las redes neuronales: $t\bar{t}H$, $t\bar{t}W$, tHq y el resto de fondos.

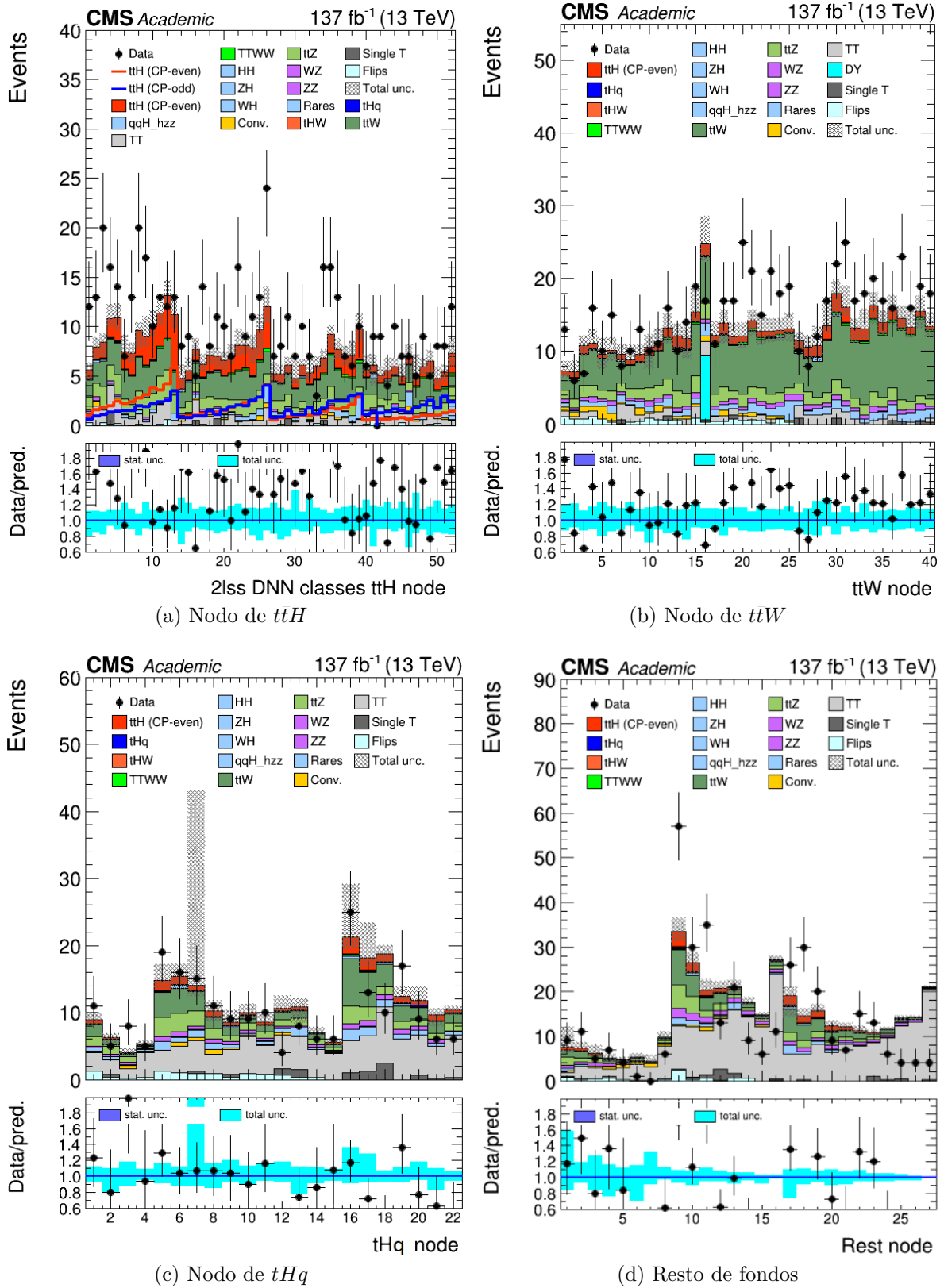


Figura 5.5: Distribuciones de los fondos más importantes del análisis usando las simulaciones de Montecarlo.

En cada uno de los nodos domina un proceso diferente. En el nodo de $t\bar{t}W$, figura 5.5b, se ve como la mayor contribución viene del color verde, que representan los sucesos esperados de $t\bar{t}W$ según la simulación de Montecarlo. En el caso del nodo de tHq , figura 5.5c, la mayor parte de la contribución es debido a leptones *non-prompt*. Como tHq tiene una sección eficaz tan baja y dada la selección aplicada, casi no hay sucesos de tHq . La clasificación del resto de fondos, figura 5.5d, está también dominada principalmente por leptones etiquetados como *non-prompt*.

5.3.1. Nodo de $t\bar{t}H$

La clasificación para el nodo de $t\bar{t}H$ se puede ver en la figura 5.5a. En ella, la línea de color rojo indica la contribución del escenario CP par y la línea azul la contribución del escenario CP impar, ambas normalizadas.

En el nodo de $t\bar{t}H$ se tiene una discriminación de los sucesos según las hipótesis de los escenarios CP, obteniendo de esta manera 4 regiones diferentes. El discriminante usado se detalla en la sección 5.2. Sucesos que sean más parecidos a aquellos descritos por el modelo estándar (escenario CP par) se situarán en la región izquierda, mientras que, sucesos que sean más parecidos a los descritos por un escenario CP impar se situarán en la región derecha. En las regiones centrales se sitúan sucesos que corresponden a estados mezcla de escenarios. Esta separación en regiones de CP se ha hecho a partir del discriminante 5.4, y permite explotar las diferencias cinemáticas del proceso dependiendo del escenario. Por otro lado, en cada región de CP, se hace una discriminación de la señal respecto al fondo obteniendo en este caso 13 *bines* (intervalos).

Hay que tener en cuenta que en cada nodo se tendrán las contribuciones de distintos procesos, ya que las redes neuronales usadas pueden fallar a la hora de separar los procesos. Además, los sucesos están categorizados según el sabor de los leptones ($ee/\mu\mu/e\mu$) porque no se mide con la misma eficiencia muones que electrones, excepto en el nodo de $t\bar{t}H$ donde no se aplica una clasificación en función del sabor de los leptones por la poca estadística que se tiene. En el nodo de $t\bar{t}H$, como se ha explicado, los sucesos se dividen en cuatro regiones dependiendo del valor que tomen en el discriminante de CP. En los nodos $t\bar{t}W$, tHq y el resto de fondos no se aplica una discriminación en CP porque no

son lo suficientemente sensibles, ni si quiera en el nodo de tHq , donde aunque se tenga una interacción entre el *quark top* y el bosón de Higgs, no permite ver las diferencias en la cinemática porque no se tienen suficientes sucesos.

Para estudiar si se puede mejorar la clasificación de sucesos, reduciendo las fluctuaciones estadísticas sin perder sensibilidad, se reducirán el número de *bines* en cada región de CP para el nodo de $t\bar{t}H$.

5.3.2. Variación del número de *bines*

A continuación, se repetirá el mismo procedimiento clasificando los procesos, pero modificando el número de *bines* usados en la discriminación de la señal y el fondo, para estudiar la estadística. Una reducción en el *binning* implica un mayor número de procesos por cada *bin*. Se representará solo la variación de *bines* en el nodo de $t\bar{t}H$, ya que es aquel en el cual se tiene la clasificación según las hipótesis de los escenarios CP par e impar. Se reducirán de 13 a 7 y 5 *bines* por cada una de las cuatro regiones de CP, la representación se muestra en la figura 5.6.

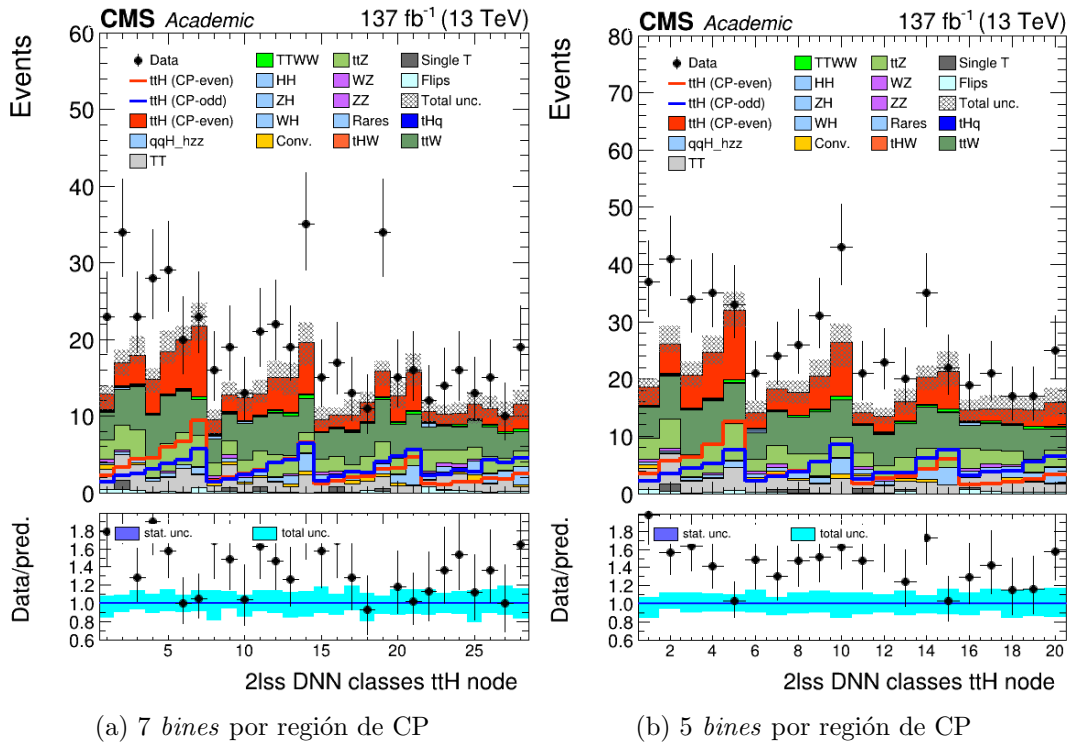
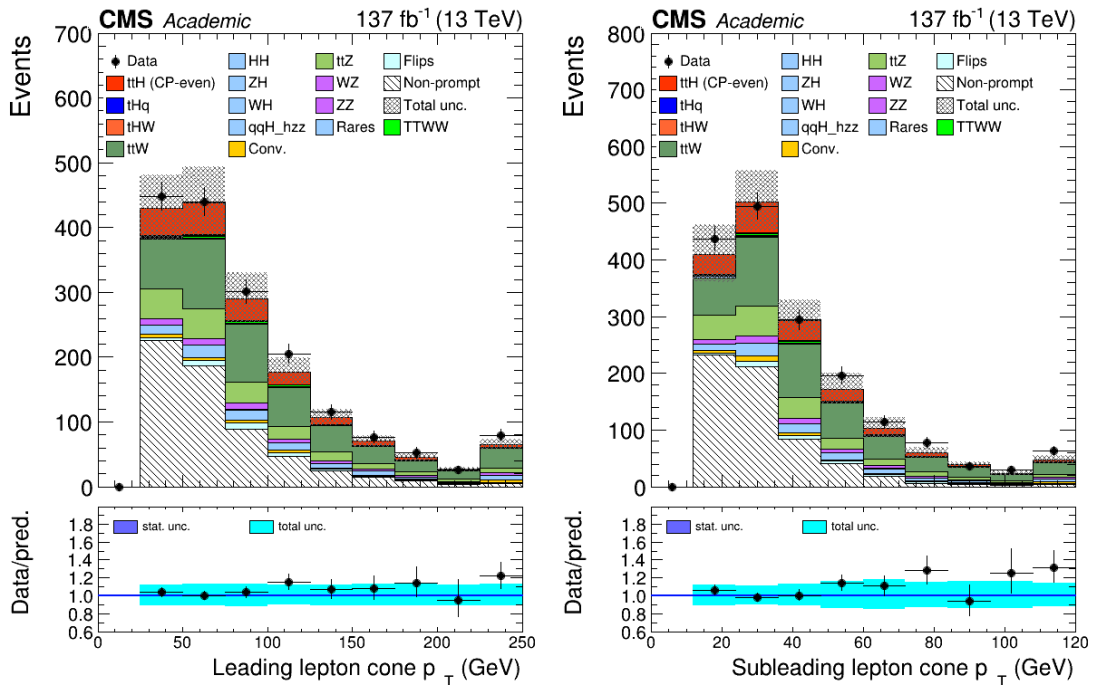


Figura 5.6: Variación en el número de *bines* usando los *fakes* de las simulaciones de Montecarlo.

5.4. Acuerdo *fakes* de Montecarlo vs *fakes* estimados de los datos

En las figuras 5.1 y 5.2 se han representado el momento del primer y segundo leptón más energéticos, así como el número de *jets* y *b-tag* para la región señal, donde se tenían apilados en colores las contribuciones esperadas de cada proceso obtenido mediante la simulación de métodos de Montecarlo, y con un punto los datos experimentales. Se ve que hay un desacuerdo ente los datos experimentales y los procesos simulados con Montecarlo. Para tratar de reproducir los datos con un mejor acuerdo, se sustituye el fondo de *fakes* simulado, por el fondo de *fakes* estimado a partir de los datos experimentales. De esta manera, la representación de los observables físicos vistos en la sección 5.1 utilizando los *fakes* estimados a partir de los datos se pueden ver en las figuras 5.7 y 5.8. En la figura 5.7 se representa el momento transverso del primer y segundo leptón más energéticos, mientras que en la figura 5.8 se tienen el número de *jets* y *b-tag* etiquetados como *medium*.



(a) Momento transverso del *leading lepton* (b) Momento transverso del *subleading lepton*

Figura 5.7: Representación de la distribución del momento transverso del leptón más energético y del segundo leptón más energético usando la contribución del fondo *fakes* estimado a partir de los datos

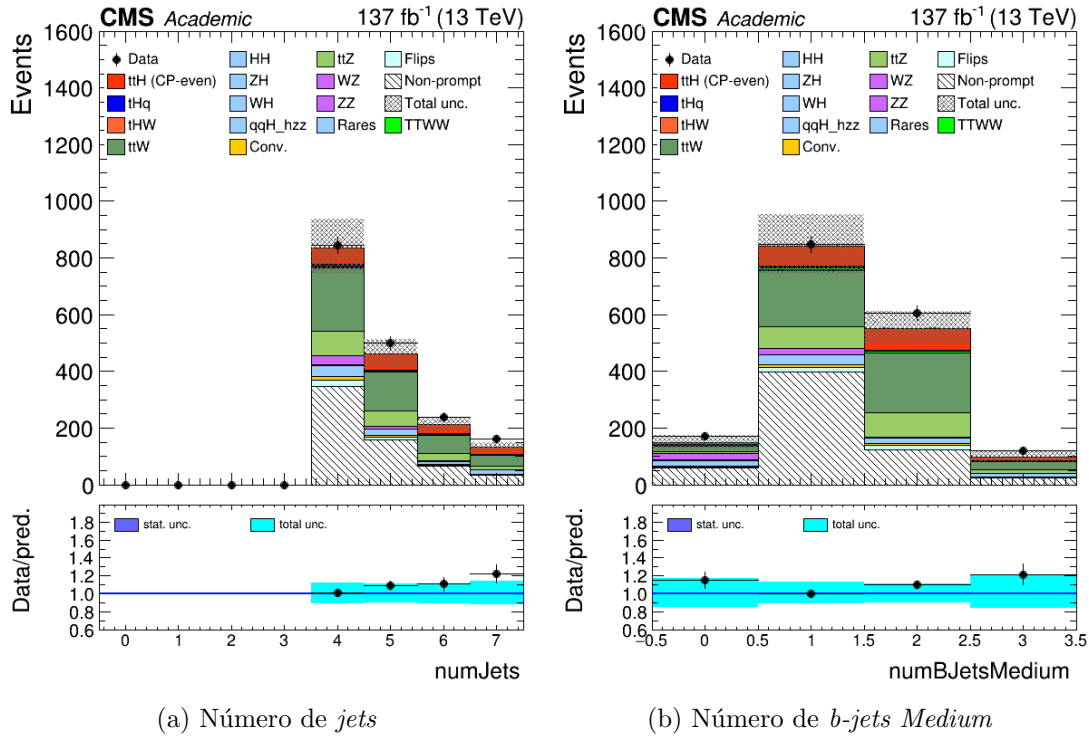


Figura 5.8: Representación del número de *jets* y *b-tag* usando la contribución del fondo *fakes* estimado a partir de los datos.

En el cuadro 5.2 se pueden ver el número de sucesos de señal y de los distintos fondos, así como sus incertidumbres en el canal $2lss + 0\tau_H$ considerando los tres años estudiados en el análisis para el caso de utilizar los *fakes* estimados a partir de los datos.

Proceso	2016+2017+2018	Proceso	2016+2017+2018
$t\bar{t}H$	172.03 ± 3.74	tHq	4.74 ± 0.25
tHW	4.26 ± 0.29	$t\bar{t}WW$	13.07 ± 0.38
WH	7.21 ± 3.71	ZH	1.02 ± 0.90
$t\bar{t}W$	459.44 ± 5.91	WZ	50.7 ± 5.16
$t\bar{t}Z$	182.78 ± 8.27	ZZ	2.68 ± 1.11
<i>Non-prompt</i>	605.17 ± 24.24	<i>Flips</i>	35.14 ± 5.97
<i>Conversions</i>	26.82 ± 8.7	<i>Rares</i>	85.70 ± 7.50
Fondos	1478.94 ± 30.7	Datos	1742

Cuadro 5.2: Sucesos considerados en el análisis para el conjunto de los años 2016, 2017 y 2018 usando los *fakes* estimados a partir de los datos.

Comparando la tabla 5.1 con la tabla 5.2, se ve como en el caso de la simulación con los *fakes* estimados de los datos, la suma de los fondos es ligeramente mayor que la suma de los fondos utilizando los *fakes* simulados. Esto explica por qué con los *fakes* estimados a partir de los datos se obtiene un mejor acuerdo entre la simulación y los datos experimentales. El proceso $t\bar{t}$ es el que da lugar a una mayor contribución al fondo de los leptones *non-prompt*, sin embargo, la suma de los procesos DY, *Single t* y $t\bar{t}$ de la tabla 5.1 no da como resultado el número de sucesos de la categoría *non-prompt*. Esto indica que, al añadir los *fakes* de los datos, se están añadiendo contribuciones de procesos que las simulaciones de Montecarlo no son capaces de reproducir.

El desacuerdo entre la simulación y los datos es debida a que el fondo de los *fakes* es un fondo instrumental que no se es capaz de modelar de forma correcta con la simulación. Para hacerlo se necesita generar cantidades más grandes de sucesos de Montecarlo, pero esto tiene un gran coste computacional.

De esta manera, la clasificación en los distintos nodos hecha por las redes neuronales, se puede ver en la figura 5.9. Se representan en colores y apilados las contribuciones esperadas para los procesos y la señal, que se representa en color rojo. Los puntos indican el datos experimentales junto con su incertidumbre estadística.

En particular, se ve que efectivamente, el nodo en el que peor acuerdo hay es el dedicado al resto de fondos, figura 5.9d, donde hay más proporción de fondo de los *fakes*. Por ello, de ahora en adelante, se utilizará para seguir con el estudio del análisis, los *fakes* estimados a partir de los datos.

Al igual que antes, se hace una reducción de 13 a 7 y 5 *bines* en el discriminante de los fondos y la señal, la variación de los *bines* se representa en la figura 5.10. Con esta estimación se obtiene un mejor acuerdo entre los datos y la simulación, por lo que para los siguientes pasos de este análisis se continuará usando los *fakes* estimados a partir de los datos. Estas discrepancias eran esperadas pues, en variables sencillas como son el momento transversal o el número de *jets*, se observaron discrepancias, lo que hace que la red neuronal, que las usa como entrada, vaya a reflejar diferencias similares.

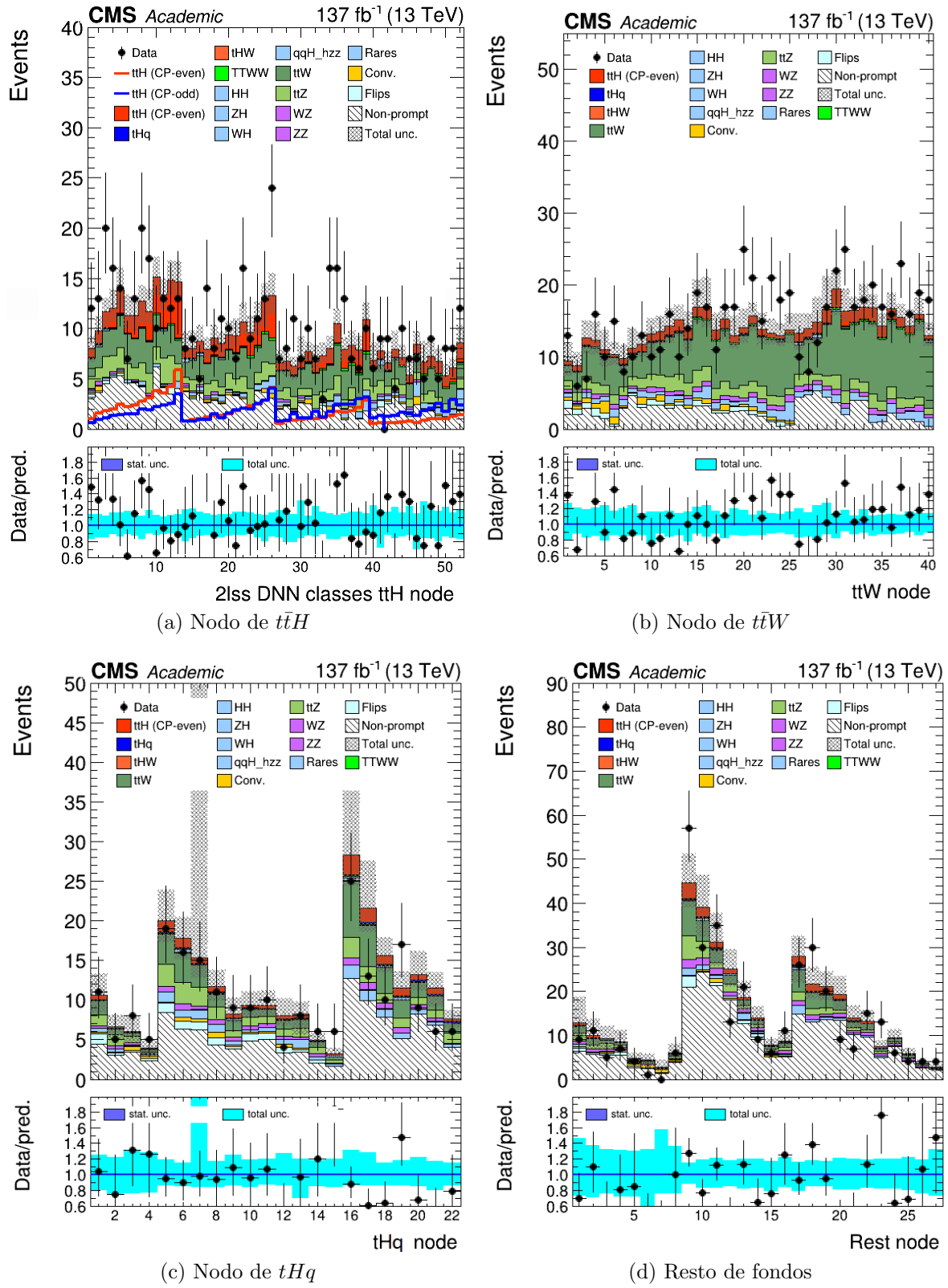


Figura 5.9: Separación de la clasificación de las redes neuronales en el canal $2lss + 0\tau_H$ usando usando la contribución del fondo *fakes* estimado a partir de los datos

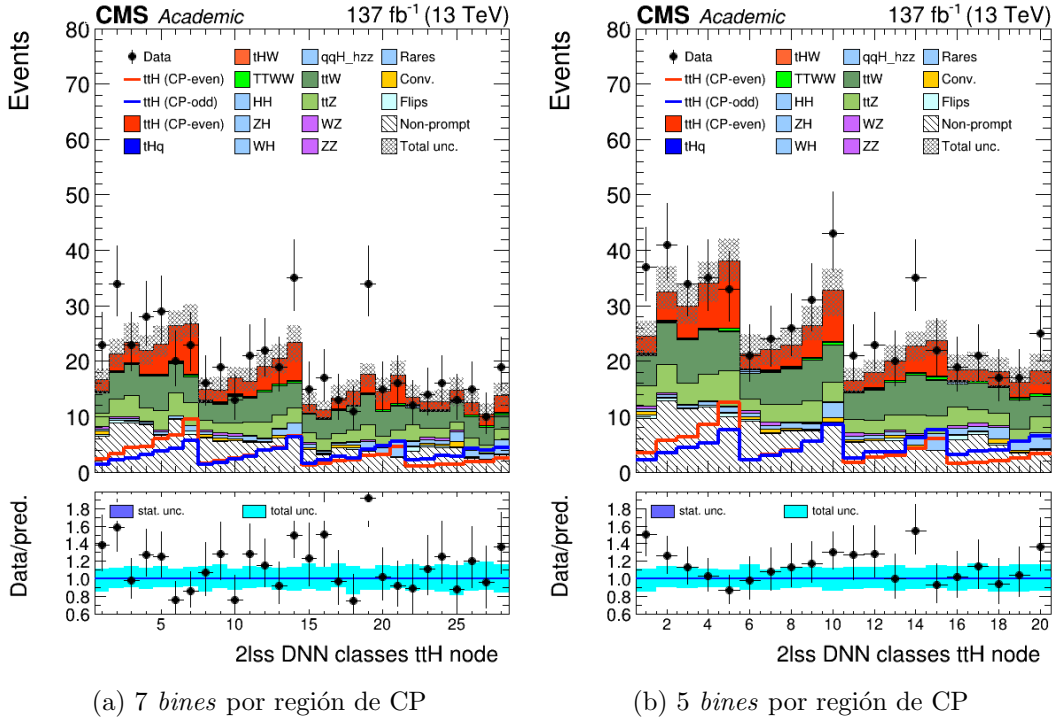


Figura 5.10: Variación en el número de *bin*s para el nodo de $t\bar{t}H$ usando la contribución del fondo *fakes* estimados a partir de los datos.

5.5. Ajuste de máxima verosimilitud y extracción del parámetro f_{CP}

Se parametriza el proceso $t\bar{t}H$ mediante los parámetros κ_t y $\tilde{\kappa}_t$, que son respectivamente, la parte par e impar del acoplamiento entre el bosón de Higgs y el *quark top*. Para obtener la representación final se utilizará el parámetro f_{CP} , que relaciona κ_t y $\tilde{\kappa}_t$ y permite medir la estructura de las contribuciones de los escenarios CP par e impar mediante la ecuación (5.3), donde los acoplamientos del Higgs a las demás partículas se restringen a su predicción por el modelo estándar. [50]

$$f_{CP}^{t\bar{t}H} = \frac{|\tilde{\kappa}_t|^2}{|\kappa_t|^2 + |\tilde{\kappa}_t|^2} \text{sign} \left(\frac{\tilde{\kappa}_t}{\kappa_t} \right) \quad (5.3)$$

5.5.1. Ajuste de máxima verosimilitud

Para la obtención de f_{CP} se usará un ajuste de máxima verosimilitud (ML) en las distribuciones de los observables en las regiones de señal para el canal $2lss + 0\tau_H$ y las

regiones de control $3l - CR$ y $4l - CR$. Las regiones han sido separadas en distintas categorías que separan los procesos de señal de los de fondo y, además para el proceso $t\bar{t}H$, separan sucesos según el escenario CP más probable. Esta clasificación ha sido detallada en la sección 5.2.

El ajuste de máxima verosimilitud es un método estadístico que se usa para ajustar un modelo y poder estimar sus parámetros. En este caso, denominarán parámetros de interés (POI) a κ_t y $\tilde{\kappa}_t$. El ajuste se hace sobre una serie de observables que discriminan correctamente la señal del fondo. Se hace el ajuste teniendo en cuenta no solo el valor de cada proceso, sino también las incertidumbres, que se modelan en este ajuste como parámetros de confusión y vienen denotados por el símbolo θ . Se quiere aplicar este procedimiento a una muestra de datos con N observaciones. Los sucesos (x_1, x_2, \dots, x_N) vienen dados por una función de distribución desconocida y la función de verosimilitud se denotará con el símbolo \mathcal{L} , que viene dada por la expresión: [13]

$$\mathcal{L}(x_1, x_2, \dots, x_N | \kappa_t, \tilde{\kappa}_t, \theta) = \prod_{i=1}^N P(x_i; \nu_i(\kappa_t, \tilde{\kappa}_t, \theta)) \prod_{j=1}^m p_j(\theta_j) \quad (5.4)$$

donde el x_i se refiere al número de sucesos observados en cada *bin*, $\nu_i(\kappa_t, \tilde{\kappa}_t, \theta)$ es el número de eventos esperados en ese *bin* y el factor $P(n_i | \kappa_t, \tilde{\kappa}_t, \theta)$ representa la probabilidad de observar n_i eventos en un *bin* dado. La distribución p_j se espera que modele los parámetros de confusión. En cada *bin* el número de sucesos que se observan sigue una distribución de probabilidad Poissoniana debido a la naturaleza estocástica de las colisiones y se conoce la medida porque se espera siempre la predicción de la señal y el fondo. La probabilidad $P(n_i | \kappa_t, \tilde{\kappa}_t, \theta)$ vendrá entonces dada por:

$$P(n_i | \kappa_t, \tilde{\kappa}_t, \theta) = e^{-\nu_i(\kappa_t, \tilde{\kappa}_t, \theta)} \frac{(\nu_i(\kappa_t, \tilde{\kappa}_t, \theta))^x}{x!} \quad (5.5)$$

donde x es el número de sucesos observados y ν_i es el número de sucesos esperados. Las incertidumbres sistemáticas que afectan a la normalización se modelan mediante una distribución normal logarítmica, mientras que las incertidumbres de forma se modelizan por una distribución gaussiana. [13]

El objetivo es encontrar los parámetros del modelo que maximicen la función de verosimilitud para los datos dados. La forma usual de resolver este problema es resolver la ecuación de verosimilitud: [13]

$$\hat{\kappa}_t, \hat{\tilde{\kappa}}_t, \hat{\theta} = \operatorname{argmax} \left[\mathcal{L} \left(x_1, x_2, \dots, x_N; \kappa_t, \tilde{\kappa}_t, \theta \right) \right] \quad (5.6)$$

En los casos más sencillos se puede resolver (5.6) de manera analítica, pero normalmente se resolverá de forma numérica. Para la resolución de la ecuación en este trabajo se utiliza la herramienta *Combine*, que busca minimizar la función $-2 \ln \mathcal{L}$.

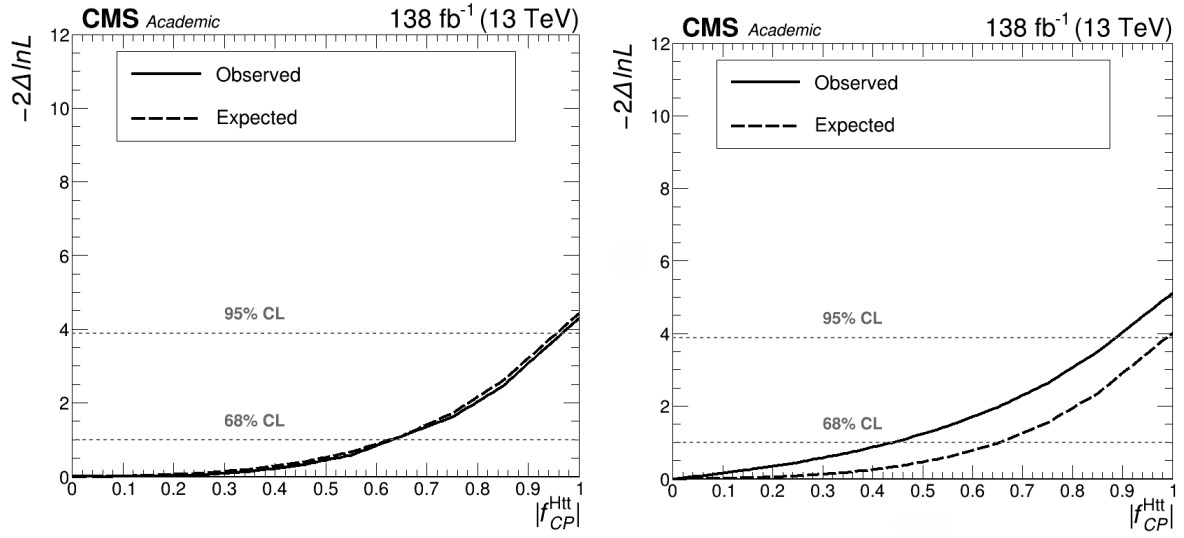
5.5.2. Resultados

Para la obtención de los resultados, se utilizará el ajuste de máxima verosimilitud descrito en la sección 5.5.1 para minimizar la función $-2 \ln \mathcal{L}$ y obtener así la mejor estimación del parámetro f_{CP} junto con su incertidumbre. Primero se realizará el ajuste sin mirar los datos, utilizando los valores esperados, y posteriormente otro en el cual se añaden los datos experimentales.

Ajustando las distribuciones correspondientes para el distinto *binning* determinado en la sección 5.3, se obtienen las representaciones de la figura 5.11. En ellas se tiene en el eje de ordenadas el logaritmo negativo de la función de máxima verosimilitud y en el eje de abscisas el valor absoluto del parámetro f_{CP} obtenido a partir de la ecuación (5.3), mediante los valores obtenidos para κ_t y $\tilde{\kappa}_t$ por el ML. La línea discontinua representa los valores esperados de f_{CP} y la línea continua indica el ajuste en el cual se han usado los datos experimentales. Los resultados obtenidos se pueden ver en el cuadro 5.3.

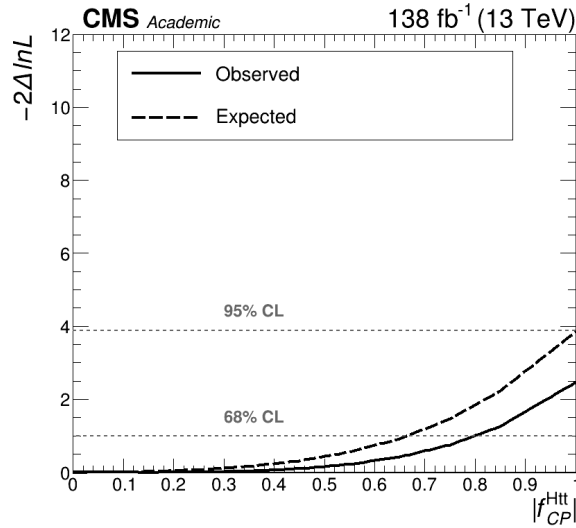
Valor de f_{CP} para un CL del 68 %		
<i>binning</i>	Esperado	Observado
13	0.0 ± 0.6	0.0 ± 0.7
7	0.0 ± 0.7	0.0 ± 0.4
5	0.0 ± 0.7	0.0 ± 0.8

Cuadro 5.3: Valores de f_{CP} obtenidos tras el ajuste para un intervalo de confianza (*Confidence Level*) del 68 %. Se indican los resultados obtenidos para los valores esperados y los valores observados.



(a) 13 *bines* por región de CP

(b) 7 *bines* por región de CP



(c) 5 *bines* por región de CP

Figura 5.11: Representaciones del parámetro f_{CP} tras el ajuste de máxima verosimilitud para los distintos *bines* considerados en cada región de CP en el nodo de $t\bar{t}H$.

Tras la realización del ajuste, se obtiene la figura 5.11. La curva es simétrica, y por esto se representa el valor absoluto. Esta curva tiene un mínimo en 0, que corresponde con el mejor valor de f_{CP} dado por el ajuste y que está de acuerdo con el modelo estándar. Para valores más altos de $-2\ln\mathcal{L}$, este valor de f_{CP} es menos compatible con los datos, por lo tanto indica que el escenario con $f_{CP} = 1$ está totalmente desfavorecido por la medida.

En las representaciones se puede ver como la reducción de 13 a 7 *bines*, figuras 5.11a y 5.11b, en la discriminación de la señal respecto al fondo para el nodo de $t\bar{t}H$, da lugar a una disminución en la incertidumbre. La reducción de *bines* aumenta el número de sucesos en cada *bin* por lo que se tiene una reducción de la incertidumbre estadística. Sin embargo, una mayor reducción no implica una mayor disminución en la incertidumbre, figuras 5.11b y 5.11c, ya que disminuir demasiado el número de *bines* hace que se pierda eficacia a la hora de discriminar los sucesos. Cabe destacar que para un número menor de *bines*, los valores observados siguen un comportamiento ligeramente distinto a los esperados.

Recordando que el parámetro f_{CP} viene dado por la ecuación (5.3), en la cual se indica la parte fraccionaria de la contribución del escenario impar, los resultados obtenidos en la figura 5.11 son compatibles con el valor $f_{CP} = 0$ esperado por el modelo estándar. Con esto, se puede comprobar que el acoplamiento de Yukawa entre el *quark top* y el bosón de Higgs es par, el bosón de Higgs se acopla al *quark top* según lo esperado por el modelo estándar.

Conclusiones

En este capítulo se detallarán las conclusiones de los resultados experimentales obtenidos en este análisis. Se han presentado los resultados del estudio del acoplamiento de Yukawa entre el *quark top* y el bosón de Higgs. Se ha estudiado el canal en el cual hay dos leptones del mismo signo en el estado final. Los datos analizados en este trabajo han sido recogidos por el detector CMS durante los años 2016, 2017 y 2018 con una energía en centro de masas de 13 TeV y con una luminosidad integrada de 137 fb^{-1} en colisiones protón-protón.

Se han utilizado simulaciones de Montecarlo para estimar los fondos y la señal y así compararlas con los datos experimentales. Para ello se establecieron ciertos criterios de selección para reducir los fondos presentes y obtener así una región en la cual predomine la señal. También se ha estudiado la estructura CP del acoplamiento entre el *quark top* y el bosón de Higgs. Para esto se han usado unas variables cinemáticas para entrenar un BDT que es capaz de distinguir entre los escenarios par e impar. Con el fin de obtener una región en la cual predomine el proceso $t\bar{t}H$ se han clasificado los sucesos usando redes neuronales para separar la señal de $t\bar{t}H$ y tH del fondo de $t\bar{t}W$ y el resto de fondos. Estas redes neuronales proporcionan varios nodos que clasifican los sucesos según cuál de las cuatro categorías anteriores es más probable. El nodo de $t\bar{t}H$ se divide a su vez en cuatro regiones según el discriminante de CP y en cada uno se obtiene una discriminación de la señal respecto al fondo. Para estudiar si se puede mejorar la clasificación de los sucesos, se hace una reducción de *bines* en la discriminación de la señal y el fondo para cada región de CP.

Comparando el acuerdo entre los datos experimentales y simulación se han visto discrepancias en las regiones enriquecidas en el fondo $t\bar{t}$, que es uno de los fondos principales. La contribución de este fondo se debe a la mala identificación de leptones y el desacuerdo proviene de un mal modelado de dicho proceso. Se comprueba que se puede estimar este fondo usando datos experimentales y obteniendo así un mejor modelado y acuerdo entre datos y predicción.

El objetivo principal del trabajo es el estudio del acoplamiento de Yukawa ente el *quark top* y el bosón de Higgs. Mediante un ajuste de máxima verosimilitud se minimizó la expresión $-2 \ln \mathcal{L}$ y se obtuvo el mejor valor del parámetro f_{CP} que relaciona κ_t y $\tilde{\kappa}_t$, que son respectivamente, la parte par e impar del acoplamiento entre el *quark top* y el bosón de Higgs. El ajuste obtenido presenta un mínimo en el valor 0, un valor compatible con el Modelo Estándar, lo que indica que el acoplamiento es par. Por otro lado se ve como para valores altos de la expresión $-2 \ln \mathcal{L}$, el escenario impar en el cual $f_{CP} = 1$, está desfavorecido por la medida. También se ha visto como una reducción en el número de *bines* en la discriminación de la señal y el fondo de 13 a 7 da lugar a una reducción en la incertidumbre de f_{CP} , pero una mayor reducción, de 7 a 5 *bines*, no implica una menor incertidumbre.

Esta medida está hecha en uno de los canales que contienen estados finales con leptones de igual carga. Este mismo estudio se puede realizar para más estados finales, lo que hará que haya más estadística y que el análisis sea más sensible. Este estudio sigue limitado por la estadística, por lo que sería necesaria una mayor cantidad de datos. El *Run-3* del LHC comenzó el 5 de julio de 2022, lo que dará lugar a una mayor cantidad de datos para analizar y mejorar la precisión de los estudios previamente hechos.

Apéndice A

Comparación de f_{CP} variando el número de *bines*

Una vez obtenidos los resultados de la sección 5.5.2, en la cual se ha visto como reduciendo el número de *bines* en cada región de CP a 7, se reduce la incertidumbre, pero disminuyendo aún más vuelve a aumentar. Se ha hecho el ajuste para dos variaciones más de *bines* y así observar la tendencia según la modificación hecha. Los valores de f_{CP} obtenidos se pueden ver en la tabla A.1 y las representaciones del ajuste, en la figura A.1. A medida que se reduce el número de *bines*, se reduce la incertidumbre, hasta alcanzar el valor más bajo con 7 *bines*, a partir de este número, una disminución en el número de *bines*, aumentará la incertidumbre del parámetro f_{CP} .

Valor de f_{CP} para un CL del 68 %		
<i>binning</i>	Esperado	Observado
13	0.0 ± 0.6	0.0 ± 0.7
8	0.0 ± 0.7	0.0 ± 0.6
7	0.0 ± 0.7	0.0 ± 0.4
6	0.0 ± 0.7	0.0 ± 0.5
5	0.0 ± 0.7	0.0 ± 0.8

Cuadro A.1: Valores de f_{CP} obtenidos tras el ajuste para todos los *bines* considerados en cada región de CP para un intervalo de confianza (*Confidence Level*) del 68 %. Se indican los resultados obtenidos para las incertidumbres de los valores esperados y los valores observados.

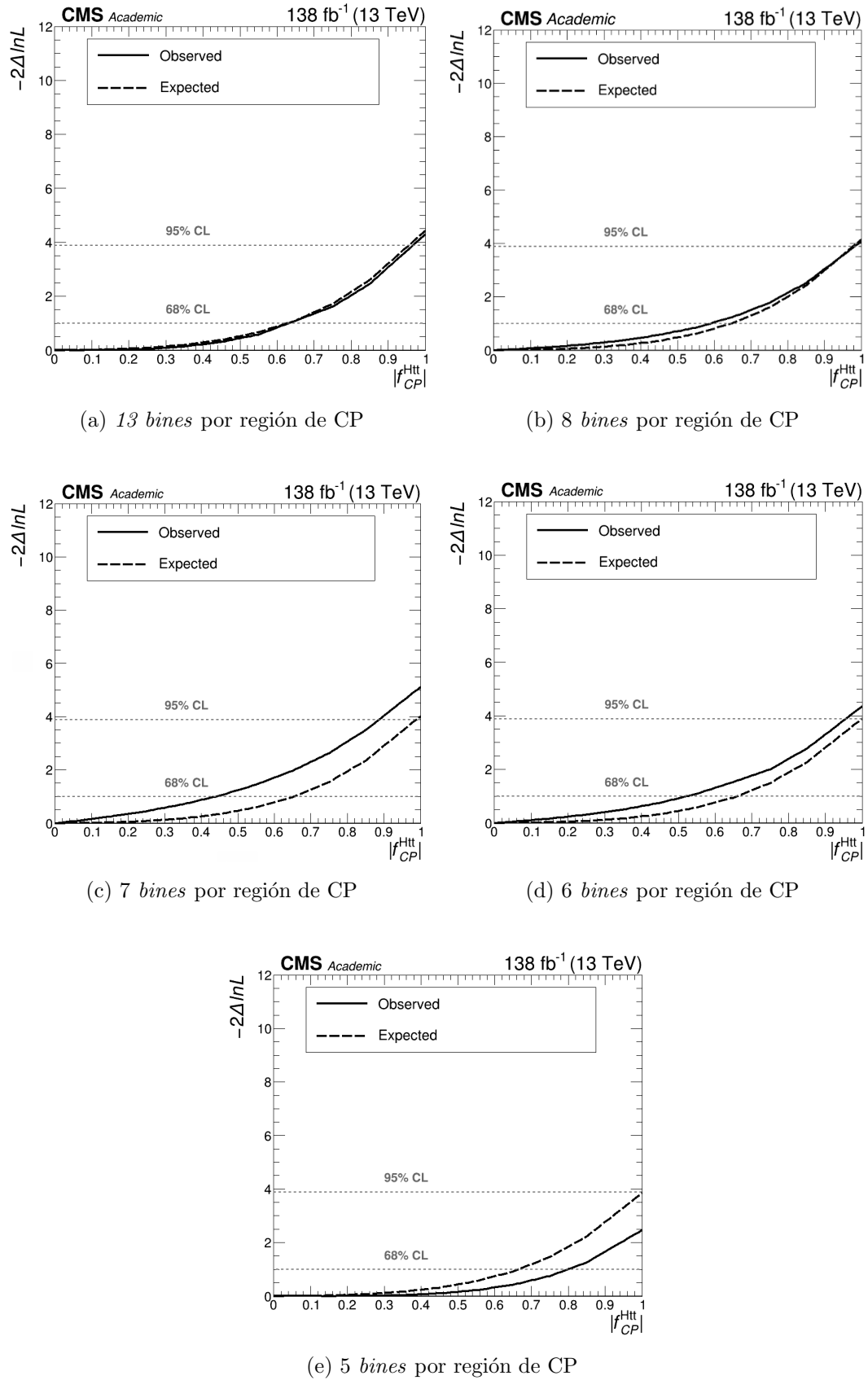


Figura A.1: Representaciones del parámetro f_{CP} tras el ajuste de máxima verosimilitud para todas las variaciones de *bines* consideradas en cada región de CP en el nodo de $t\bar{t}H$.

Siglas

ALICE A Large Ion Collider Experiment. 12

ATLAS A Toroidal LHC ApparatuS. 12

BDTs Boosted Decision Trees. 57

CERN Conseil Européen pour la Recherche Nucléaire. 11

CMS Compact Muon Solenoid. 12, 19

CP Charge-Parity. 56

CSC Cathode Strip Chambers. 28

DF Drift Tubes. 28

DY Drell-Yan. 44

ECAL Electromagnetic calorimeter. 19, 22

HCAL Hadronic calorimeter. 19, 24

HLT High Level Trigger. 29

HPDs Hybrid Photodiodes. 24

HPS Hadrons-Plus-Strips. 34

LEIR Low Energy Iron Ring. 12

LEP Large Electron-Positron. 12

-
- LHC** Large Hadron Collider. 11
- LHCb** Large Hadron Collider beauty. 12
- LINAC 3** Acelerador lineal 3. 12
- LINAC 4** Acelerador lineal 4. 12
- LO** Leading Order. 36
- ME** Modelo Estándar. 3
- MET** Missing Transverse Energy. 24
- ML** Most Likelihood. 67
- NLO** Next to Leading Order. 36
- PDF** Parton Distribution Function. 36
- PF** Particle Flow. 31
- POI** Parameter of interest. 68
- PS** Proton Synchrotron. 12
- PSB** Proton Synchrotron Booster. 12
- QCD** Quantum chromodynamics. 3
- RF** Radiofrequency. 15
- RPC** Resistive Plate Chambers. 28
- SPS** Super Proton Synchrotron. 12
- SUSY** Super Symmetry. 8

Bibliografía

- [1] G. Aad, T. Abajyan, B. Abbott y J. Abdallah. «Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC». En: *Physics Letters B* 716.1 (2012), págs. 1-29. ISSN: 0370-2693. DOI: <https://doi.org/10.1016/j.physletb.2012.08.020>. URL: <https://www.sciencedirect.com/science/article/pii/S037026931200857X>.
- [2] S. Agostinelli y col. «GEANT4—a simulation toolkit». En: *Nucl. Instrum. Meth. A* 506 (2003), págs. 250-303. DOI: [10.1016/S0168-9002\(03\)01368-8](https://doi.org/10.1016/S0168-9002(03)01368-8).
- [3] A.B. Arbuzov. *Quantum Field Theory and the Electroweak Standard Model*. 35 pages. Ene. de 2018. DOI: [10.23730/CYRSP-2017-004.1](https://doi.org/10.23730/CYRSP-2017-004.1). arXiv: [1801.05670](https://arxiv.org/abs/1801.05670). URL: <https://cds.cern.ch/record/2315477>.
- [4] The ATLAS, CDF, CMS y D0 Collaborations. *First combination of Tevatron and LHC measurements of the top-quark mass*. 2014. DOI: [10.48550/ARXIV.1403.4427](https://doi.org/10.48550/ARXIV.1403.4427). URL: <https://arxiv.org/abs/1403.4427>.
- [5] G L Bayatian y col. *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software*. Technical design report. CMS. There is an error on cover due to a technical problem for some items. Geneva: CERN, 2006. URL: <https://cds.cern.ch/record/922757>.
- [6] R. Brun y F. Rademakers. «ROOT: An object oriented data analysis framework». En: *Nucl. Instrum. Meth. A* 389 (1997). Ed. por M. Weren y D. Perret-Gallix, págs. 81-86. DOI: [10.1016/S0168-9002\(97\)00048-X](https://doi.org/10.1016/S0168-9002(97)00048-X).
- [7] Oliver Sim Brüning y col. *LHC Design Report*. CERN Yellow Reports: Monographs. Geneva: CERN, 2004. DOI: [10.5170/CERN-2004-003-V-1](https://doi.org/10.5170/CERN-2004-003-V-1). URL: <https://cds.cern.ch/record/782076>.

- [8] Matteo Cacciari, Gavin P Salam y Gregory Soyez. «The anti-ik/isubit/i/subject clustering algorithm». En: *Journal of High Energy Physics* 2008.04 (abr. de 2008), págs. 063-063. DOI: [10.1088/1126-6708/2008/04/063](https://doi.org/10.1088/1126-6708/2008/04/063). URL: <https://doi.org/10.1088/1126-6708/2008/04/063>.
- [9] Mario González Carpintero. *Trabajo Fin de Grado, Producción electrodébil de WZ (VBS) en el experimento CMS (LHC), Universidad de Oviedo, año 2019*.
- [10] Francesca Cavallari. «Performance of calorimeters at the LHC». En: *Journal of Physics: Conference Series* 293 (abr. de 2011), pág. 012001. DOI: [10.1088/1742-6596/293/1/012001](https://doi.org/10.1088/1742-6596/293/1/012001). URL: <https://doi.org/10.1088/1742-6596/293/1/012001>.
- [11] S. Chatrchyan y col. «Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC». En: *Physics Letters B* 716.1 (2012), págs. 30-61. ISSN: 0370-2693. DOI: <https://doi.org/10.1016/j.physletb.2012.08.021>. URL: <https://www.sciencedirect.com/science/article/pii/S0370269312008581>.
- [12] *Citing Pythia*. URL: <https://pythia.org/documentation/#citing-pythia>.
- [13] CMS Collaboration. «Measurement of the Higgs boson production rate in association with top quarks in final states with electrons, muons, and hadronically decaying tau leptons at $\sqrt{s} = 13$ TeV». En: (2021). DOI: [10.48550/ARXIV.2011.03652](https://arxiv.org/abs/2011.03652). URL: <https://arxiv.org/abs/2011.03652>.
- [14] *CMS Luminosity Measurements for the 2016 Data Taking Period*. Inf. téc. Geneva: CERN, 2017. URL: <http://cds.cern.ch/record/2257069>.
- [15] «CMS Phase 1 heavy flavour identification performance and developments». En: (mayo de 2017). URL: <http://cds.cern.ch/record/2263802>.
- [16] *CMSSW*. URL: <https://github.com/cms-sw/cmssw>.
- [17] A Colaleo, A Safonov, A Sharma y M Tytgat. *CMS Technical Design Report for the Muon Endcap GEM Upgrade*. Inf. téc. Jun. de 2015. URL: <https://cds.cern.ch/record/2021453>.

- [18] CMS Collaboration. «Performance and operation of the CMS electromagnetic calorimeter». En: *Journal of Instrumentation* 5.03 (mar. de 2010), T03010-T03010. DOI: [10.1088/1748-0221/5/03/t03010](https://doi.org/10.1088/1748-0221/5/03/t03010). URL: <https://doi.org/10.1088/1748-0221/5/03/t03010>.
- [19] CMS Collaboration. «Performance of q-lepton reconstruction and identification in CMS». En: *Journal of Instrumentation* 7.01 (ene. de 2012), P01001-P01001. DOI: [10.1088/1748-0221/7/01/p01001](https://doi.org/10.1088/1748-0221/7/01/p01001). URL: <https://doi.org/10.1088/1748-0221/7/01/p01001>.
- [20] CMS Collaboration. URL: http://cms-results.web.cern.ch/cms-results/public-results/publications/HIG-19-006/CMS-HIG-19-006_Figure_014-b.png.
- [21] CMS Collaboration. *A vacuum as empty as interstellar space*. URL: <https://home.cern/science/engineering/vacuum-empty-interstellar-space>.
- [22] CMS Collaboration. *Accelerating: Radiofrequency cavities*. URL: <https://home.cern/science/engineering/accelerating-radiofrequency-cavities>.
- [23] CMS Collaboration. *Bending Particles*. URL: <https://cms.cern/detector/bending-particles>.
- [24] CMS Collaboration. *Branching fractions predicts for an SM Higgs boson of mass 125 GeV*. URL: https://cds.cern.ch/record/2236645/files/Figures_HiggsProdAndBRs.png.
- [25] CMS Collaboration. *Cryogenics: Low temperatures, high performance*. URL: <https://home.cern/science/engineering/cryogenics-low-temperatures-high-performance>.
- [26] CMS Collaboration. *Detecting muons*. URL: <https://cms.cern/detector/detecting-muons>.
- [27] CMS Collaboration. *Energy of electrons and photons (ECAL)*. URL: <https://cms.cern/detector/measuring-energy/energy-electrons-and-photons-ecal>.
- [28] CMS Collaboration. *Energy of hadron (HCAL)*. URL: <https://cms.cern/detector/measuring-energy/energy-hadrons-hcal>.

- [29] CMS Collaboration. *High Luminosity LHC Project*. URL: <https://hilumilhc.web.cern.ch/content/hl-lhc-project>.
- [30] CMS Collaboration. *Jets at CMS and the determination of their energy scale*. URL: <https://cms.cern/news/jets-cms-and-determination-their-energy-scale>.
- [31] CMS Collaboration. *Pulling together: Superconducting electromagnets*. URL: <https://home.cern/science/engineering/pulling-together-superconducting-electromagnets>.
- [32] CMS Collaboration. *Tracking*. URL: <https://cms.cern/detector/identifying-tracks>.
- [33] CMS Collaboration. *Triggering and data acquisition*. URL: <https://cms.cern/detector/triggering-and-data-acquisition>.
- [34] The CMS Collaboration y col. «The CMS experiment at the CERN LHC». En: *Journal of Instrumentation* 3.08 (ago. de 2008), S08004-S08004. DOI: [10.1088/1748-0221/3/08/s08004](https://doi.org/10.1088/1748-0221/3/08/s08004). URL: <https://doi.org/10.1088/1748-0221/3/08/s08004>.
- [35] *Combine*. URL: <https://cms-analysis.github.io/HiggsAnalysis-CombinedLimit/>.
- [36] Lyndon Evans. «The Large Hadron Collider». En: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 370.1961 (2012), págs. 831-858. DOI: [10.1098/rsta.2011.0453](https://doi.org/10.1098/rsta.2011.0453). eprint: <https://royalsocietypublishing.org/doi/pdf/10.1098/rsta.2011.0453>. URL: <https://royalsocietypublishing.org/doi/abs/10.1098/rsta.2011.0453>.
- [37] *Github*. URL: <https://github.com/about>.
- [38] Particle Data Group. URL: <https://pdg.lbl.gov/>.
- [39] Stefan Höche. *Introduction to parton-shower event generators*. 2014. DOI: [10.48550/ARXIV.1411.4085](https://doi.org/10.48550/ARXIV.1411.4085). URL: <https://arxiv.org/abs/1411.4085>.
- [40] High Energy Physics Illinois. *The MadGraph5_aMC@NLO homepage*. URL: <http://madgraph.physics.illinois.edu/>.
- [41] Bora Isildak. *Measurement of the differential dijet production cross section in proton-proton collisions at $\sqrt{s} = 7$ tev*. 2013. arXiv: [1308.6064](https://arxiv.org/abs/1308.6064) [hep-ex].

- [42] Hyun Min Lee. «Lectures on physics beyond the Standard Model». En: *Journal of the Korean Physical Society* (mayo de 2021). ISSN: 1976-8524. DOI: [10.1007/s40042-021-00188-x](https://doi.org/10.1007/s40042-021-00188-x). URL: <http://dx.doi.org/10.1007/s40042-021-00188-x>.
- [43] *MadGraph5_aMC@NLO webpage*. URL: http://amcatnlo.web.cern.ch/amcatnlo/list_refs.htm.
- [44] Fabienne Marcastel. «CERN's Accelerator Complex. La chaîne des accélérateurs du CERN». En: (ago. de 2013). General Photo. URL: <https://cds.cern.ch/record/1621583>.
- [45] Pierluigi Paolucci. *The CMS Muon system*. Inf. téc. Geneva: CERN, abr. de 2005. DOI: [10.1142/9789812773678_0096](https://doi.org/10.1142/9789812773678_0096). URL: <https://cds.cern.ch/record/927394>.
- [46] *Performance of reconstruction and identification of tau leptons in their decays to hadrons and tau neutrino in LHC Run-2*. Inf. téc. Geneva: CERN, 2016. URL: <https://cds.cern.ch/record/2196972>.
- [47] Bogdan Povh y col. *Particles and Nuclei, An introduction to the Physical Concepts*. Seventh Edition. Springer, 2014. ISBN: 978-3-662-46320-8.
- [48] J Quijada. «Recent upgrades and results from the CMS experiment». En: *Journal of Physics: Conference Series* 912 (oct. de 2017), pág. 012005. DOI: [10.1088/1742-6596/912/1/012005](https://doi.org/10.1088/1742-6596/912/1/012005).
- [49] UC Davis Richard Breedon. «View through the CMS detector during the cooldown of the solenoid on February 2006». CMS Collection. Feb. de 2006. URL: <https://cds.cern.ch/record/930094>.
- [50] *Search for CP violation in $t\bar{t}H$ and tH production in multilepton channels at $\sqrt{s} = 13$ TeV*. Inf. téc. Geneva: CERN, 2022. URL: <https://cds.cern.ch/record/2803420>.
- [51] A. M. Sirunyan y col. «Evidence for associated production of a Higgs boson with a top quark pair in final states with electrons, muons, and hadronically decaying leptons at $\sqrt{s}=13$ TeV». En: *Journal of High Energy Physics* 2018.8

- (ago. de 2018). DOI: [10.1007/jhep08\(2018\)066](https://doi.org/10.1007/jhep08(2018)066). URL: <https://doi.org/10.48550/arXiv.1803.05485>.
- [52] A. M. Sirunyan y col. «Observation of $t\bar{t}H$ Production». En: *Phys. Rev. Lett.* 120 (23 jun. de 2018), pág. 231801. DOI: [10.1103/PhysRevLett.120.231801](https://doi.org/10.1103/PhysRevLett.120.231801). URL: <https://link.aps.org/doi/10.1103/PhysRevLett.120.231801>.
- [53] M. Tanabashi, K. Hagiwara, K. Hikasa y P. Nakamura. «Review of Particle Physics». En: *Phys. Rev. D* 98 (3 ago. de 2018), pág. 030001. DOI: [10.1103/PhysRevD.98.030001](https://doi.org/10.1103/PhysRevD.98.030001). URL: <https://link.aps.org/doi/10.1103/PhysRevD.98.030001>.
- [54] *The CMS hadron calorimeter project: Technical Design Report*. Technical design report. CMS. Geneva: CERN, 1997. URL: <https://cds.cern.ch/record/357153>.
- [55] *The Powheg Box*. URL: <https://powhegbox.mib.infn.it/#References>.
- [56] T. S. Virdee. «Beyond the standard model of particle physics». En: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 374.2075 (2016), pág. 20150259. DOI: [10.1098/rsta.2015.0259](https://doi.org/10.1098/rsta.2015.0259). eprint: <https://royalsocietypublishing.org/doi/pdf/10.1098/rsta.2015.0259>. URL: <https://royalsocietypublishing.org/doi/abs/10.1098/rsta.2015.0259>.
- [57] Tejinder S. Virdee. «Assembled HCAL endcap (HE) in November 2002». CMS Collection. Nov. de 2002. URL: <https://cds.cern.ch/record/947004>.