



Universidad de Oviedo

FACULTAD DE CIENCIAS - GRADO EN FÍSICA

TRABAJO FIN DE GRADO

**Producción asociada del bosón de Higgs y quarks
top en canales multileptónicos en el detector CMS
del LHC**

Alejandro Soto Rodríguez

Tutores:

Isidro González Caballero

Sergio Sánchez Cruz

Curso 2019-2020

RESUMEN

Producción asociada del bosón de Higgs y quarks top en canales multileptónicos en el detector CMS del LHC

En este trabajo se presentan los resultados de la búsqueda de la producción asociada del bosón de Higgs y *quarks top* en canales multileptónicos. Los datos analizados se corresponden con una luminosidad integrada de 137 fb^{-1} . Estos datos han sido recogidos por el detector CMS durante los años 2016, 2017 y 2018 en colisiones protón-protón a una energía en centro de masas de 13 TeV. Se ha obtenido un valor experimental para la fuerza de señal μ , definida como el cociente entre la sección eficaz experimental y teórica del proceso $t\bar{t}H$. El resultado ha sido $\mu^{obs} = 2.6 \pm_{0.9}^{0.9}$, correspondiente a una medida conjunta de los tres años en estados finales con dos leptones del mismo signo. Este resultado es compatible con el valor predicho por el Modelo Estándar en menos de dos desviaciones estándar. Por último, se ha comprobado que el volumen de datos recogidos hasta ahora por CMS no permite la observación de la resonancia de masa del bosón de Higgs en el canal 4ℓ de $t\bar{t}H$.

Índice

	Página
Resumen	I
Introducción	1
1 El Modelo Estándar y extensiones	3
1.1 El Modelo Estándar	3
1.2 Física más allá del Modelo Estándar	6
1.3 El proceso $t\bar{t}H$	8
2 Dispositivo Experimental	11
2.1 LHC	11
2.1.1 Luminosidad y <i>pile-up</i>	13
2.1.2 Componentes principales del LHC	14
2.2 El detector CMS	16
2.2.1 Sistema de Coordenadas	17
2.2.2 El detector de trazas (tracker)	18
2.2.3 Calorímetro Electromagnético	20
2.2.4 Calorímetro Hadrónico	21
2.2.5 Solenoide superconductor	22
2.2.6 Cámaras de muones	22
2.2.7 Trigger	24
3 Reconstrucción y simulación de sucesos	27
3.1 Reconstrucción de sucesos	27
3.1.1 Vértice principal	29
3.1.2 Muones	29
3.1.3 Electrones	31
3.1.4 <i>Jets</i>	32
3.1.5 Taus	34
3.1.6 Momento transverso faltante	34

3.1.7	Masa invariante	35
3.2	Simulación de sucesos	36
3.2.1	Generación	36
3.2.2	Simulación y digitalización	38
3.3	Herramientas de software	39
4	Estudio del proceso $t\bar{t}H$	41
4.1	Señal y fondo	41
4.2	Estados finales objeto de estudio	42
4.3	Muestras de datos y simulaciones de Montecarlo	44
4.3.1	Fondos irreducibles	44
4.3.2	Fondos reducibles	45
4.4	Selección de sucesos	47
4.4.1	Canal $2\ell ss$	47
4.4.2	Canal $H \rightarrow ZZ^* \rightarrow 4\ell$	51
4.5	Comparación de generadores de Montecarlo para el proceso $t\bar{t}$	53
4.6	Fuentes de incertidumbre	55
4.6.1	Incertidumbres estadísticas	55
4.6.2	Incertidumbres sistemáticas	56
5	Resultados Experimentales	59
5.1	Canal $2\ell ss$	59
5.1.1	Región de señal	59
5.1.2	Fuerza de señal	64
5.2	Medida de la masa del bosón de Higgs en el canal de $t\bar{t}H$: $H \rightarrow ZZ^* \rightarrow 4\ell$	68
6	Conclusiones	71
	Bibliografía	78

Lista de Figuras

1.1	Partículas fundamentales del Modelo Estándar de Física de Partículas y sus propiedades.	4
1.2	Interacciones entre partículas del Modelo Estándar.	5
1.3	Esquema de las unificaciones de teorías en física.	7
1.4	Diagrama de Feynman LO para la producción asociada de dos <i>quarks top</i> con un bosón de Higgs.	8
1.5	Constantes de acoplo del bosón de Higgs determinadas experimentalmente por la colaboración CMS.	9
1.6	Canales de desintegración del bosón de Higgs y sus correspondientes fracciones de desintegración.	9
2.1	Esquema del complejo de aceleradores del CERN.	11
2.2	Estructura de dos anillos del LHC junto a los cuatro detectores principales: CMS, ATLAS, LHCb y ALICE.	12
2.3	Promedio del número de interacciones por cada cruce paquetes de protones para los años 2015 a 2018.	14
2.4	Una de las cavidades de radiofrecuencia del LHC.	15
2.5	Sección transversal de un dipolo superconductor del LHC	16
2.6	Estructura del detector CMS.	17
2.7	Sistema de coordenadas utilizado en CMS.	18
2.8	Vértices reconstruidos (78 en total) en un suceso de alto <i>pile-up</i> . Las líneas verdes se corresponden con las trazas de las partículas cargadas y los puntos amarillos con el vértice de interacción.	19
2.9	Fotografía del detector de trazas de CMS.	19
2.10	Esquema del calorímetro electromagnético de CMS.	21
2.11	Foto de una de las tapas del calorímetro hadrónico.	21
2.12	Solenoide superconductor de CMS.	22
2.13	Disposición del sistema de detección de muones de CMS.	23
3.1	Corte transversal de CMS junto a las trayectorias de diversas partículas dentro del detector.	28

3.2	Esquema de las correcciones aplicadas a la energía de los <i>jets</i> tanto para datos como para las muestras de Montecarlo.	32
3.3	Esquema de un suceso con dos <i>jets</i> ligeros (en gris) y un <i>b-jet</i> (en azul).	33
3.4	Esquema de los pasos en los que se divide la generación de un suceso.	37
4.1	Un ejemplo de diagrama de Feynman de la producción de $t\bar{t}H$ en el LHC y su posterior desintegración en un estado final con dos leptones del mismo signo y un τ_h	42
4.2	Secciones eficaces experimentales para algunos de los procesos más relevantes del Modelo Estándar publicadas por la colaboración CMS.	44
4.3	Diagramas de Feynman para los procesos de fondo $t\bar{t}W$ y $t\bar{t}Z$ (de izquierda a derecha).	45
4.4	Diagramas de Feynman para los procesos de fondo WZ y ZZ (de izquierda a derecha).	45
4.5	Diagramas de Feynman para los procesos de fondo $t\bar{t}$, Drell-Yan, $Z + \gamma$, WW , $W + jets$ y <i>single top</i> (de izquierda a derecha y de arriba a abajo).	46
4.6	Representación de E_T^{Miss} frente a H_T^{Miss} . La línea recta negra se corresponde con el criterio de selección $L_D > 30$ GeV. Los sucesos por encima de la línea pasan la selección y los que están por debajo no. De esta forma, se elimina la contribución del fondo de Drell-Yan.	50
4.7	Distribución del número de <i>jets</i> y <i>b-jets</i> . Los criterios de selección aplicados a los sucesos mostrados en esta figura son los definidos para la región de señal exceptuando las restricciones en el número de <i>jets</i> . La señal $t\bar{t}H$ se representa en color rojo.	50
4.8	a) Distribución del número de <i>jets</i> identificados con $p_T \geq 25$ GeV. b) Distribución del número de <i>b-jets</i> identificados con $p_T \geq 25$ GeV y pertenecientes a la categoría <i>medium</i> . c) Distribución del número de <i>b-jets</i> identificados con $p_T \geq 25$ GeV y pertenecientes a la categoría <i>loose</i> . En todas las gráficas se ha utilizado una escala logarítmica en el eje de ordenadas y como criterios de selección se han aplicado los definidos en la tabla 4.3 pero sin incluir las restricciones en el número de <i>jets</i> y <i>b-jets</i>	52
4.9	Distribución del número de <i>jets</i> para las muestras de Montecarlo NLO (<i>Powheg</i>) y LO (<i>Madgraph</i>) de $t\bar{t}$	54
4.10	Distribución del número de <i>b-jets</i> en la categoría <i>medium</i> para las muestras de Montecarlo NLO (<i>Powheg</i>) y LO (<i>Madgraph</i>) de $t\bar{t}$	54

4.11	Distribución del momento transverso del <i>jet</i> más energético del suceso para las muestras de Montecarlo NLO (<i>Powheg</i>) y LO (<i>Madgraph</i>) de $t\bar{t}$.	54
4.12	Distribución del momento transverso del leptón más energético del suceso para las muestras de Montecarlo NLO (<i>Powheg</i>) y LO (<i>Madgraph</i>) de $t\bar{t}$.	55
5.1	Distribución del momento transverso del leptón más energético del suceso en la región de señal.	60
5.2	Distribución del momento transverso del segundo leptón más energético del suceso en la región de señal.	61
5.3	Sabores de los leptones <i>tight</i> en la región de señal.	61
5.4	Distribución de la pseudorapidez η del leptón más energético del suceso en la región de señal.	62
5.5	Distribución del ángulo ϕ del leptón más energético del suceso en la región de señal.	62
5.6	Distribución del momento transverso faltante en la región de señal.	63
5.7	Número de sucesos totales seleccionados en la región de señal.	63
5.8	Impacto de las fuentes de incertidumbre en el cálculo de la fuerza de señal.	67
5.9	Masa invariante de los 4 leptones seleccionados. A la derecha se ha hecho un discretizado más fino para poder observar el pico de masa del bosón de Higgs en torno a 125 GeV.	68
5.10	Diagrama de Feynman de la producción de un bosón de Higgs mediante fusión de dos gluones.	69

Lista de Tablas

4.1	Secciones eficaces de los procesos más importantes en este análisis y los generadores utilizados en el proceso de simulación de cada uno.	46
4.2	Selección de sucesos en la categoría $2\ell ss$	49
4.3	Selección de sucesos en la categoría 4ℓ	53
4.4	Sucesos totales seleccionados de $t\bar{t}$ para la muestra NLO (<i>Powheg</i>) y LO (<i>Madgraph</i>). Las incertidumbres que se incluyen en la tabla son la combinación de la incertidumbre estadística y sistemática.	55

5.1	Sucesos seleccionados de los procesos considerados en el análisis para los años 2016, 2017 y 2018. Las incertidumbres que se incluyen en la tabla son la combinación de la incertidumbre estadística y sistemática.	64
5.2	Sucesos totales seleccionados en el canal 4ℓ . Las incertidumbres que se incluyen en la tabla son la combinación de la incertidumbre estadística y sistemática.	69

Introducción

En este trabajo se presenta el estudio del proceso de producción de dos *quarks top* asociados a un bosón de Higgs ($t\bar{t}H$) en colisiones protón-protón realizadas en el LHC. La medida de la tasa de producción de este proceso permite obtener un valor preciso e independiente del modelo de y_t , la constante de acoplo de Yukawa del bosón de Higgs al *quark top*. Esta constante es de gran interés y, su valor, podría dar pistas a favor o en contra de la existencia de física más allá del Modelo Estándar. Además, dada la gran masa del *quark top*, mucho mayor a la de cualquier otro fermión del Modelo Estándar, puede indicar que juega un papel clave en el mecanismo de ruptura espontánea de simetría. Mecanismo por el cual los bosones W^\pm y Z adquieren masa.

El estudio se ha realizado utilizando los datos recogidos por el detector CMS durante los años 2016, 2017 y 2018 en colisiones protón-protón a 13 TeV de energía en centro de masas. Debido a que tanto el *quark top* como el bosón de Higgs son partículas inestables y se desintegran antes de alcanzar los detectores, se divide el estudio en categorías en función del estado final. En este análisis se han escogido estados finales con varios leptones, denominados multileptónicos.

Este trabajo se desarrolla dentro del marco de las asignaturas: Física Nuclear y de Partículas Elementales, Física de Altas Energías y Aceleradores y Método Numéricos Aplicados a la Física, impartidas dentro del Grado en Física de la Universidad de Oviedo.

El trabajo consta de seis capítulos estructurados de la siguiente forma. El capítulo 1 constituye una introducción al marco teórico vigente actualmente en física de altas energías, el Modelo Estándar de Física de Partículas, así como algunas de sus posibles extensiones.

En el capítulo 2 se expone el funcionamiento del gran colisionador de hadrones, LHC, y de CMS, uno de sus detectores. De ellos se han obtenido los datos analizados en este trabajo y, por lo tanto, comprender su funcionamiento será importante para interpretar los resultados.

El capítulo 3 está dedicado a la reconstrucción y simulación de sucesos. En él se explican los algoritmos y técnicas que se utilizan para extraer la información física

de las colisiones. Por ejemplo, las partículas producidas y sus propiedades. Se expone también el papel que desempeñan las simulaciones y cómo son los pasos para realizarlas.

Los capítulos 4 y 5 constituyen la parte central de este trabajo. En el capítulo 4 se definen los criterios de selección de sucesos empleados en el análisis, así como los procesos de fondo más importantes. Se introducen también las fuentes de incertidumbre consideradas explicando su origen y cálculo. El capítulo 5 recoge los resultados experimentales derivados del estudio del proceso $t\bar{t}H$ y su interpretación.

Por último, el capítulo 6 presenta las conclusiones finales del trabajo.

Capítulo 1

El Modelo Estándar y extensiones

1.1. El Modelo Estándar

El Modelo Estándar (*Standard Model*, SM) es una teoría física que describe las partículas y las interacciones fundamentales de la naturaleza. Surge de la combinación de la Teoría Electrodébil y la Cromodinámica Cuántica que aúnan en un único marco teórico las interacciones electromagnética, débil y fuerte. La figura 1.1 muestra las partículas que componen el Modelo Estándar. Consta de dos bloques fundamentales: los **fermiones** (*spin* semientero) y los **bosones** (*spin* entero). Los fermiones a su vez se subdividen en dos grupos: los **leptones** y los **quarks**. Estos últimos son los constituyentes fundamentales de la materia y las interacciones entre ellos se producen mediante el intercambio de bosones que son los mediadores de la interacción. Cada interacción lleva asociada una carga que tendrán las partículas que sufran esa interacción. La carga eléctrica está asociada a la interacción electromagnética, la carga débil a la interacción débil y la carga de color a la interacción fuerte. Existen tres tipos de carga de color denominados: rojo, verde y azul. La intensidad de cada interacción se cuantifica a través de su constante de acoplamiento. Estas constantes juegan un papel muy importante dentro de la dinámica de los sistemas, ya que frecuentemente, se establecen jerarquías de aproximación basadas en potencias de estas constantes. La notación utilizada en este trabajo será: α para la interacción electromagnética, α_S para la interacción fuerte y α_W para la interacción débil.

La familia de los *quarks* se divide en tres generaciones ordenadas de menor a mayor masa. La primera generación la componen los *quarks up* (**u**) y *down* (**d**), la segunda generación los *quarks charm* (**c**) y *strange* (**s**) y la tercera los *quarks top* (**t**) y *bottom* (**b**). Los *quarks* u, c y t tienen carga eléctrica $\frac{2}{3}$ (en unidades naturales de carga) y los *quarks* d, s y b, $-\frac{1}{3}$. Además de carga eléctrica, también tienen carga de color y carga

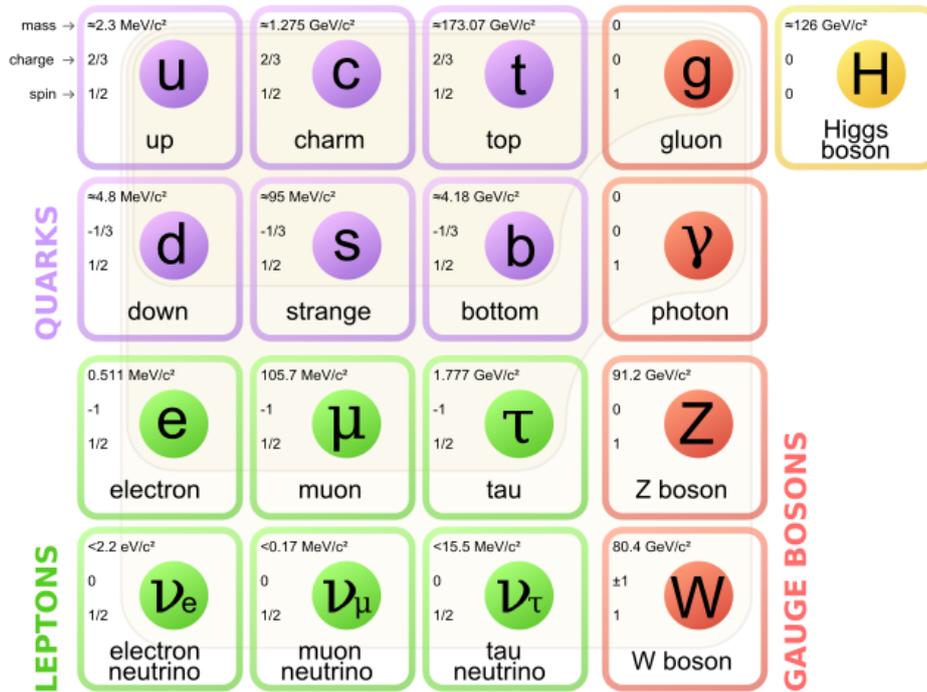


Figura 1.1: Partículas fundamentales del Modelo Estándar de Física de Partículas y sus propiedades [1].

débil.

Los leptones también se subdividen, como los *quarks*, en tres generaciones, siendo la primera la conformada 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 carga débil, ninguno tiene carga de color y entre ellos, el electrón, el muon y el tau tienen carga eléctrica de valor 1. Cabe destacar que entre generaciones las propiedades son esencialmente las mismas, solo cambia la masa.

A las partículas de la figura 1.1 hay que añadir la **antipartícula** de cada una de ellas. Una antipartícula tiene las mismas propiedades que su respectiva partícula excepto que su carga eléctrica es opuesta. La existencia de las antipartículas fue predicha por Paul Dirac en 1928 [2] al interpretar las soluciones de energía negativa de su ecuación de onda como un nuevo tipo de partícula.

Por otro lado, existen cinco tipos de bosones en el Modelo Estándar: el **fotón** (γ) mediador de la interacción electromagnética, los **gluones**¹ (g) mediadores de la interacción fuerte, los bosones W^\pm y Z^0 mediadores de la interacción débil y el **bosón de Higgs** (H^0), que mediante el mecanismo de Higgs dota de masa a las partículas. Un esquema de las interacciones descritas en el modelo estándar se muestra en la figura

¹Existen ocho tipos de gluones diferentes dentro del Modelo Estándar tanto como combinaciones de parejas color-anticolor se pueden hacer.

1.2. Las líneas azules unen los fermiones con los bosones que median en la interacción.

Una de las propiedades más importantes de la interacción fuerte es el confinamiento de color. Este fenómeno es debido a que los gluones también son portadores de la carga de color y por lo tanto son capaces de interactuar consigo mismos. El efecto es que tanto los gluones como los *quarks* no se pueden observar aislados en la naturaleza y siempre se encuentran formando sistemas compuestos denominados **hadrones**. Los hadrones son sistemas ligados por interacción fuerte y de carga de color nula. Existen dos clases de hadrones observados en la naturaleza: los **bariones**, formados por tres *quarks* y los **mesones**, formados por una pareja *quark-antiquark*. En una colisión de altas energías cuando se producen partículas con carga de color como *quarks* o gluones, rápidamente se hadronizan. Este hadrón, generalmente inestable, se desintegra en otras partículas que pueden a su vez volver a producir hadrones. Se inicia así una cascada de partículas que en física de altas energías se denomina *jet*.

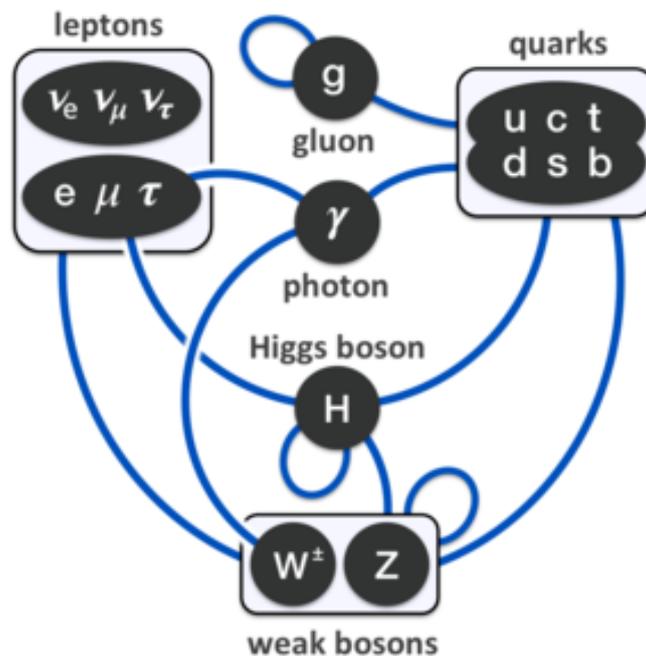


Figura 1.2: Interacciones entre partículas del Modelo Estándar [1].

En su formulación original, el Modelo Estándar predecía partículas sin masa. Para solucionar esto, Robert Brout, François Englert y Peter Higgs propusieron un mecanismo conocido como el mecanismo de Brout-Englert-Higgs que otorga masa a las partículas que interactúan con el campo de Higgs. La intensidad de la interacción del campo de Higgs con los fermiones se cuantifica a través de las constantes de acoplamiento de Yukawa y_f . El campo de Higgs, como todo campo fundamental, tiene una partícula

asociada llamada bosón de Higgs de masa 125 GeV^2 [3]. Tras la propuesta de esta partícula en 1964, su búsqueda fue uno de los objetivos principales de la comunidad científica hasta que finalmente fue descubierta en 2012 en los detectores CMS [4] y ATLAS [5] del CERN. Este descubrimiento fue premiado el 8 de octubre de 2013 con el Premio Nobel de Física a Peter Higgs y François Englert. Ese mismo año recibieron también el Premio Príncipe de Asturias de Investigación Científica y Técnica junto con el CERN.

1.2. Física más allá del Modelo Estándar

Aunque el SM es la teoría de física de partículas más precisa y robusta que se tiene hasta la fecha, existen una serie de fenómenos que no puede explicar. Muchos de los experimentos que se realizan en física de partículas están enfocados a medir parámetros del Modelo Estándar y a comprobar si las predicciones que hace son correctas. Pero otros de ellos buscan encontrar nuevos fenómenos no descritos dentro del SM. Sin embargo, todos los experimentos realizados hasta la fecha concuerdan con las predicciones del SM y no se tiene ninguna forma de motivar una nueva teoría. Algunas de las búsquedas que se están realizando actualmente van desde encontrar una partícula candidata a formar la materia oscura, hasta la búsqueda de dimensiones extra. Por otro lado, se han desarrollado teorías que buscan ampliar el Modelo Estándar incluyendo, por ejemplo, la interacción gravitatoria. Las teorías que aspiran a convertirse en una extensión del Modelo Estándar más conocidas son las teorías de supersimetría o teorías **SUSY**, la **teoría de cuerdas** y las teorías de **gravedad cuántica** entre otras.

La figura 1.3 muestra un esquema de cómo a lo largo de los años se han podido unificar teorías que describían interacciones diferentes en una sola. En este esquema se muestra también la situación actual y el marco donde se desarrolla el SM. Se piensa entonces, que al igual que antes se había conseguido, ahora también se tienen que poder unificar las interacciones restantes en una teoría única, la Teoría del Todo. Es por ello que muchos de los esfuerzos de los físicos teóricos actuales se dirigen hacia la creación de esta teoría.

Algunos de los fenómenos que el SM no es capaz de explicar son:

- **Interacción Gravitatoria:** Esta interacción no se encuentra incluida en el SM y la teoría que la describe (Teoría de la Relatividad General, Einstein 1915) utiliza un formalismo completamente diferente. Mientras que en el Modelo Estándar

²Todas las magnitudes en este trabajo se expresan en unidades naturales. Este sistema toma $\hbar = c = 1$ y, como consecuencia, tanto la masa como el momento tienen unidades de energía.

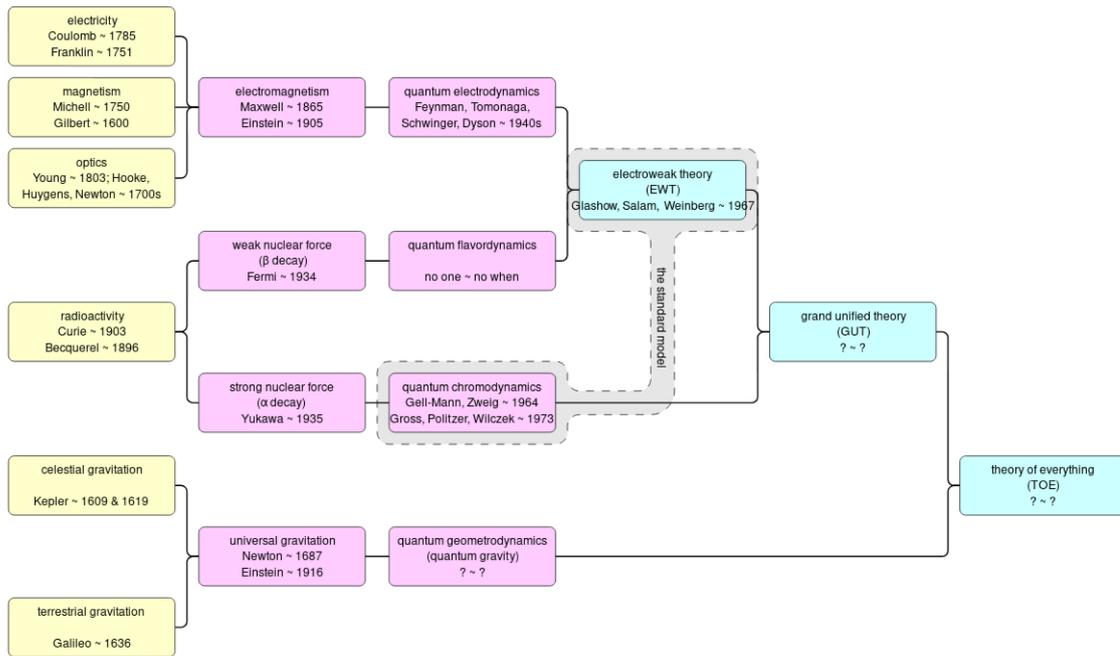


Figura 1.3: Esquema de las unificaciones de teorías en física [6].

una interacción es llevada a cabo por el intercambio de partículas, en la Teoría de la Relatividad es la deformación del espacio tiempo la que produce la fuerza gravitatoria. Se ha teorizado sobre la existencia de una partícula que se ha denominado *gravitón*, que sería la mediadora de esta interacción. Sin embargo, dicha partícula aún no ha sido descubierta.

- **Materia Oscura:** Se desconoce la composición de la materia oscura. Fue descubierta gracias a la influencia gravitatoria que ejercía sobre las galaxias, cuya curva de rotación no era explicable si solo se tenía en cuenta la materia visible. Actualmente se estima que un 26% [7] de la materia del Universo es materia oscura, y de momento solo se sabe que interacciona gravitacionalmente.
- **Energía Oscura:** Se sabe que compone el 69% [7] de la materia existente en el Universo y su naturaleza no está clara. Surge como explicación a la expansión acelerada del Universo y se representa dentro de las ecuaciones de Einstein como un término denominado constante cosmológica.
- **Asimetría materia-antimateria:** Es un hecho evidente que el Universo se encuentra compuesto principalmente por materia. Dentro del SM no existe ninguna explicación (suponiendo que en el Big Bang se crearon en cantidades iguales) acerca de la prevalencia de materia frente a antimateria. Se busca por ello, alguna propiedad que las diferencie y que permita explicar esta anomalía.

1.3. El proceso $t\bar{t}H$

Gracias a los datos recogidos por los detectores del LHC a diferentes energías se han podido determinar varias propiedades del bosón de Higgs y su interacción con algunas de las partículas del SM. Como comentamos en la sección 2.1, la interacción del resto de partículas del SM con el campo de Higgs las dota de la masa que se observa en los experimentos. Las constantes de acoplo y_f del Higgs con los fermiones son proporcionales a la masa de los mismos ($y_f = m_f/v$ con $v \approx 246$ GeV, el valor esperado del campo de Higgs en el vacío [8]), y expresan la intensidad de la interacción entre el fermión correspondiente y el bosón de Higgs. Muchos de estos valores han sido comprobados experimentalmente, sin embargo, hay algunos de ellos que por falta de datos no se han podido determinar con precisión. La figura 1.5 recoge los resultados experimentales de la determinación de estas constantes para algunas partículas del Modelo Estándar. Entre ellas, el acoplamiento del Higgs con el *quark top* y_t es de gran interés por varias razones. El *quark top* es la partícula con más masa del Modelo Estándar (173 GeV [3]). Esto puede indicar que juega un papel especial en el mecanismo de ruptura espontánea de simetría electrodébil. Además, es la única partícula que tiene un acoplamiento al Higgs cercano a 1. El interés en el proceso $t\bar{t}H$ surge gracias a que permite obtener una medida precisa de y_t , que a su vez nos daría pistas a favor o en contra de contribuciones de Física más allá del Modelo Estándar.

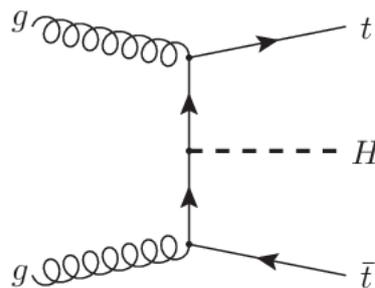


Figura 1.4: Diagrama de Feynman LO para la producción asociada de dos *quarks top* con un bosón de Higgs.

El proceso $t\bar{t}H$ consiste en la producción asociada de dos *quarks top* con un bosón de Higgs como ilustra la figura 1.4. El *quark top* se desintegra casi el 100% de las veces en un *quark b* y un bosón W. Éste a su vez se puede desintegrar en una pareja leptón-neutrino o en una *quark-antiquark*. Por otro lado, el bosón de Higgs presenta mayoritariamente las siguientes desintegraciones: $H \rightarrow W^+W^-, Z^0Z^0, \tau^+\tau^-, b\bar{b}, \gamma\gamma, gg, c\bar{c}$. La figura 1.6 muestra los canales de desintegración

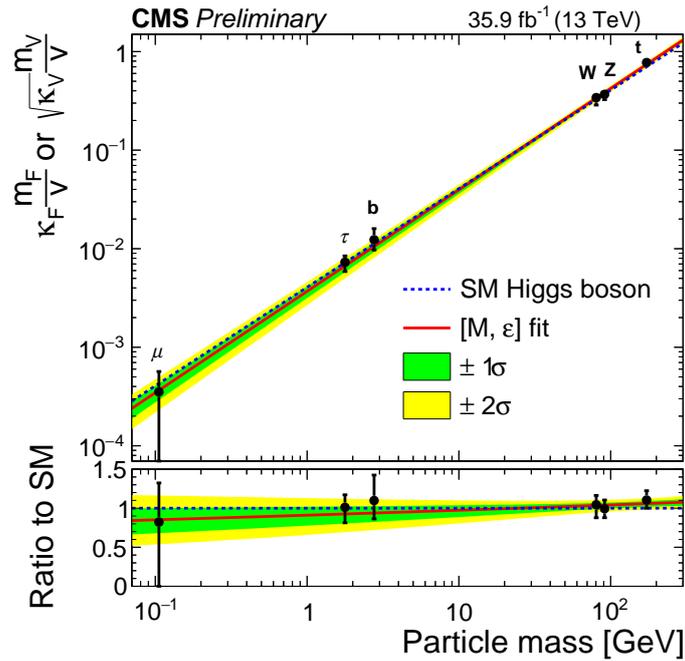


Figura 1.5: Constantes de acoplo del bosón de Higgs determinadas experimentalmente por la colaboración CMS [9].

del Higgs junto a su fracción de desintegración. Todas estas posibilidades hace que existan muchos estados finales que pueden ser objeto de estudio. El canal multileptónico, escogido en este trabajo, abarca mayoritariamente las desintegraciones del bosón de Higgs a W^+W^- , Z^0Z^0 y $\tau^+\tau^-$. Donde los estados finales contienen un cierto número de leptones. A lo largo de este trabajo se hará referencia a los leptones como únicamente electrones y muones excluyendo de esta categoría a los τ , ya que su vida media es muy corta (290 fs [3]) y no pueden ser detectados. Los otros canales de desintegración del bosón de Higgs también son objeto de estudio en otros análisis y permiten medir otras propiedades diferentes del bosón.

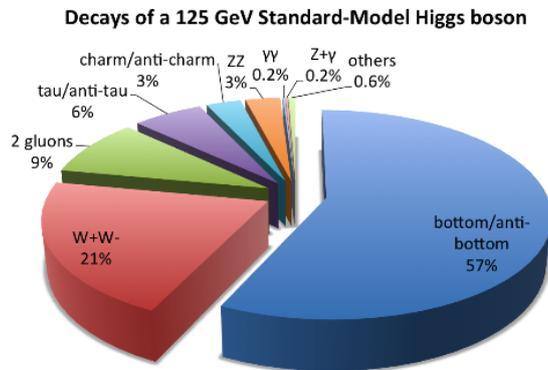


Figura 1.6: Canales de desintegración del bosón de Higgs y sus correspondientes fracciones de desintegración [10].

Capítulo 2

Dispositivo Experimental

2.1. LHC

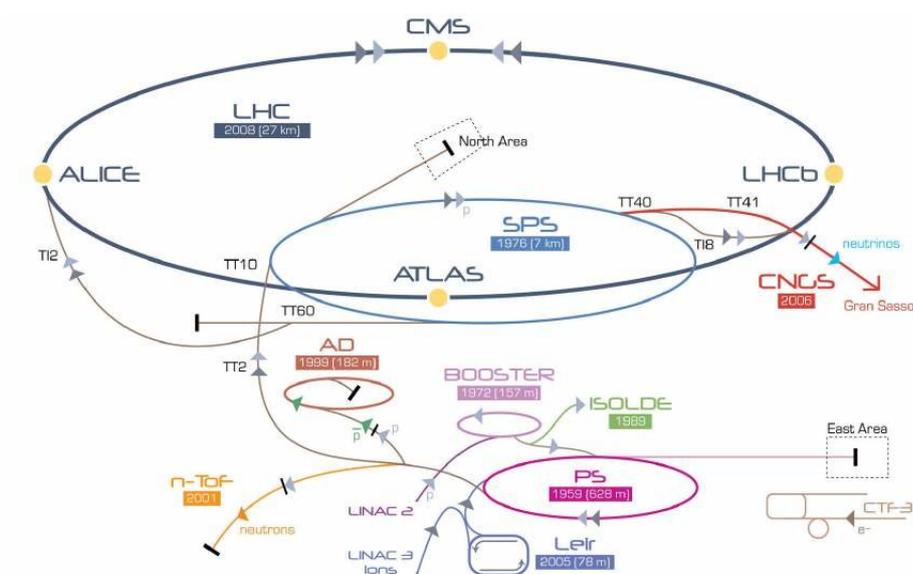


Figura 2.1: Esquema del complejo de aceleradores del CERN [11].

El LHC (*Large Hadron Collider*) es un colisionador de hadrones circular. Está instalado en el túnel de 26.7 km que fue construido entre 1984 y 1989 para el colisionador LEP (*Large Electron-Positron collider*) ya desmontado¹. Este túnel se encuentra a una profundidad entre 45 m y 170 m por debajo de la superficie. Por el LHC pueden circular haces de protones o de iones pesados (principalmente plomo) para producir colisiones protón-protón, plomo-plomo o plomo-protón. Los haces de partículas iniciales son divididos en dos anillos paralelos por donde viajan los haces

¹LEP fue un colisionador de electrones y positrones que estuvo en funcionamiento desde 1989 hasta el 2000 en el CERN.

en sentidos opuestos. En este trabajo se estudiarán únicamente datos tomados en colisiones protón-protón, por ello, profundizaremos a partir de ahora en su proceso de aceleración. Antes de inyectarse los protones en el LHC, son acelerados secuencialmente en una cadena de cuatro aceleradores: el acelerador lineal Linac 2, PSB, el Sincrotrón de Protones (PS) y el Super Sincrotrón de Protones (SPS). Estos aceleradores cumplen también la función de dividir el haz en paquetes o *bunches* de aproximadamente 10^{11} protones separados 25 ns . La figura 2.1 muestra el complejo de aceleradores construido en el CERN, y la figura 2.2 representa en mayor detalle el LHC, su estructura de dos anillos y los cuatro detectores donde se hacen colisionar los protones: CMS, ATLAS, ALICE y LHCb.

La energía en centro de masas (\sqrt{s}) que alcanza el LHC actualmente es 13 TeV. Sin embargo, se espera aumentar en un futuro a 14 TeV, energía para la que fue diseñado. Para mantener las partículas girando dentro de la trayectoria circular en el LHC, se necesitan fuertes campos magnéticos que son generados por imanes superconductores. Estos imanes pueden alcanzar intensidades de 8.3 T del campo magnético. Para que funcionen correctamente deben ser enfriados a una temperatura de 1.9 K y ser aislados del exterior lo máximo posible. Esto se consigue situándolos en cavidades de vacío llamadas criostatos.

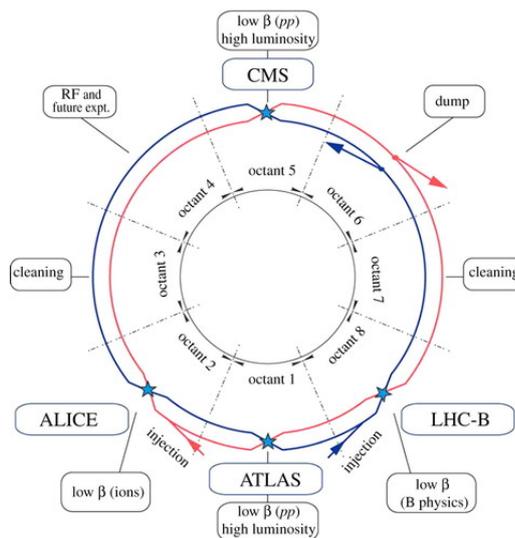


Figura 2.2: Estructura de dos anillos del LHC junto a los cuatro detectores principales: CMS, ATLAS, LHCb y ALICE [12].

2.1.1. Luminosidad y *pile-up*

Luminosidad

Uno de los parámetros más importantes a la hora de construir un acelerador es la luminosidad instantánea \mathcal{L} . La luminosidad instantánea es una medida del número de colisiones que pueden producirse en un detector por unidad de área y de tiempo. Cuanto mayor sea \mathcal{L} , mayor será el número de colisiones que podrán producirse. Actualmente el LHC opera a una luminosidad instantánea de $2 \cdot 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$. Si asumimos paquetes de protones (*bunches*) distribuidos espacialmente según una función de densidad de probabilidad gaussiana y colisionando frontalmente a velocidades cercanas a la de la luz, podemos obtener una expresión analítica aproximada para la luminosidad (ec 2.1).

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

Donde:

- f : frecuencia de cruce entre los *bunches*.
- N_i : número de partículas por *bunch*.
- σ_i : 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).
- N_b : número de *bunches* en el haz.

Otra magnitud importante que caracteriza el desempeño de un acelerador es la luminosidad integrada L (ec 2.2), que está relacionada con el número total de colisiones observadas en un periodo de tiempo T . Tiene unidades del inverso de área y se suele expresar en b^{-1} ($1b = 10^{-28} \text{ m}^2$). La luminosidad también relaciona el número total de sucesos esperados N de un proceso con sección eficaz σ , y se calcula: $N = L \cdot \sigma$. En física de partículas la sección eficaz es una medida de la probabilidad de que un proceso ocurra en una colisión y tiene unidades de área. Como comparación, la sección eficaz inelástica total de una colisión protón-protón es del orden de 80 mb [13] y la del proceso $t\bar{t}H$ estudiado en este trabajo $0.50 \pm_{0.05}^{0.04} \text{ pb}$ [14] (ambos a 13 TeV en centro de masas). Es decir, solo 1 de cada $1.6 \cdot 10^{11}$ colisiones inelásticas da lugar al proceso $t\bar{t}H$.

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

En este trabajo se analizarán datos recogidos por el detector CMS en el *Run-2* del LHC durante los años 2016, 2017 y 2018, que registraron una luminosidad integrada de 35.9 fb^{-1} , 41.4 fb^{-1} y 59.7 fb^{-1} respectivamente.

Pile-up

Cada vez que se cruzan dos paquetes de protones más de una interacción protón-protón tiene lugar. Los detectores colocados en el punto de interacción observan inevitablemente los resultados de muchas de estas colisiones en un lapso de tiempo muy corto. Esto afecta a las medidas realizadas sobre los procesos de interés, ya que aparecen muchas más partículas que pueden confundirse con las procedentes del proceso principal. Este fenómeno se conoce como *pile-up*, y es consecuencia de la alta luminosidad instantánea del LHC.

En los datos analizados en este trabajo se observa un promedio de 27 interacciones por cada cruce de paquetes en el año 2016. Para 2017 y 2018 esta cifra aumenta a 38 y 37 respectivamente (figura 2.3).

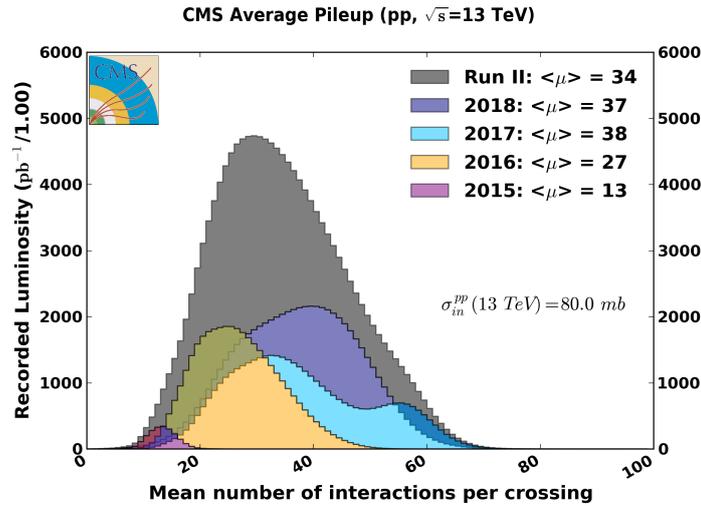


Figura 2.3: Promedio del número de interacciones por cada cruce de paquetes de protones para los años 2015 a 2018 [15].

2.1.2. Componentes principales del LHC

La energía máxima en centro de masas a la que ha operado el LHC ha sido de 13 TeV en su último periodo de toma de datos (*Run-2*). Esto corresponde a 6.5 TeV por protón, una energía jamás alcanzada anteriormente y que ha supuesto todo un reto tecnológico conseguirlo. En esta sección describiremos los componentes principales del acelerador que han permitido alcanzar una energía en centro de masas tan grande. Una

descripción detallada del funcionamiento y componentes del LHC se puede encontrar en [16].

Las **cavidades de radiofrecuencia** son el dispositivo utilizado para acelerar las partículas (protones e iones de plomo) en el LHC. En concreto, el LHC cuenta con 16 de estas que elevan la energía de 450 GeV, proporcionada por los aceleradores previos a los protones, hasta los 6.5 TeV. Estas cavidades están agrupadas de 4 en 4 y requieren temperaturas muy bajas para funcionar como superconductores, por ello, se sitúan dentro de 4 refrigeradores cilíndricos llamados *criomódulos*. Para alcanzar la máxima energía solo necesitan alrededor de 20 minutos, tiempo en el cuál los protones son capaces de dar alrededor de 10 millones de vueltas. Las cavidades de radiofrecuencia se ajustan de tal manera que los campos electromagnéticos que generan oscilen a 400 MHz y sincronizados con el paso del haz. Esto provoca que protones con energía superiores o inferiores a la deseada sean frenados y acelerados respectivamente, homogeneizando la energía del haz. La figura 2.4 muestra una de las cavidades de radiofrecuencia instaladas en el LHC.

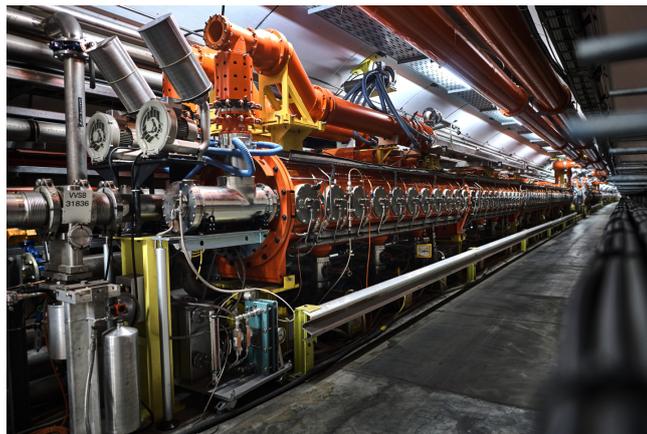


Figura 2.4: Una de las cavidades de radiofrecuencia del LHC [17].

Otros componentes básicos y cruciales para el funcionamiento del LHC son los imanes superconductores. Existen varios tipos (más de 50), cada uno con una función determinada como puede ser mantener la trayectoria circular o corregir la forma de los *bunches*. Esta forma puede verse afectada por la repulsión entre los protones que componen el *bunch*.

Los imanes que cumplen la función de curvar el haz son los **imanes dipolares** y hay 1232 en total. Cada uno de ellos mide 15 m, pesa 35 toneladas y es capaz de alcanzar 8.3 T [18]. La figura 2.5 muestra la estructura interna de dichos imanes. El campo que crean estos imanes se dirige en la dirección vertical para que la fuerza de Lorentz curve el haz en la dirección deseada. Sin embargo, el campo que crean no es completamente homogéneo, y presenta imperfecciones sobre todo en los extremos del

imán. Para solucionar esto, se utilizan imanes sextupolares, octopolares y decapolares.

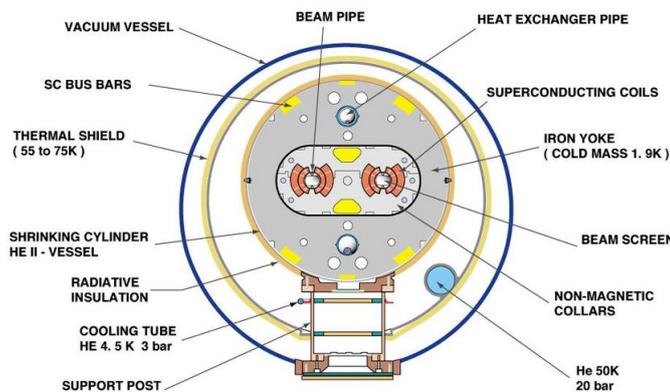


Figura 2.5: Sección transversal de un dipolo superconductor del LHC [19].

La otra función mencionada es la de enfocar el haz. Este aspecto es importante ya que mejora sustancialmente el número de colisiones que ocurren cuando se cruzan los haces. Los imanes que realizan esta función son los **imanes cuadrupolares**; constan de cuatro polos simétricamente colocados en torno al tubo por el que circula el haz, y pueden ejercer fuerzas de compresión sobre los paquetes de protones que circulan por el tubo. Se utilizan también cerca de los detectores donde ocurren las colisiones para concentrar los protones provenientes de direcciones opuestas.

Después de cada cruce solo unos pocos protones habrán interactuado. Existen de nuevo otro conjunto de imanes que permiten focalizar y reutilizar el haz. Cuando no se quiera hacer esto se puede desechar el haz desviándolo y haciéndolo colisionar con un enorme bloque de grafito.

2.2. El detector CMS

El detector CMS (*Compact Muon Solenoid*) es un detector de propósito general diseñado para detectar una gran variedad de procesos producidos a las escalas de energía del LHC. Sus funciones van desde el estudio del Modelo Estándar hasta la búsqueda de materia oscura o dimensiones extra.

CMS es un detector de forma cilíndrica situado en torno al tubo del haz donde se hacen colisionar los haces de protones previamente acelerados en el LHC. Está compuesto de varios subdetectores, cada uno de ellos especializado en la detección de uno o varios tipos de partículas que se pueden producir en la colisión. Una representación del detector se muestra en la figura 2.6. El detector se divide en dos regiones: la zona del barril rodeando al solenoide superconductor y las tapas (o en inglés

endcaps) en la base del cilindro que conforma CMS. En su conjunto el detector pesa 14000 toneladas, mide 15 m de alto y 21 m de largo y posee el solenoide superconductor más potente del mundo, capaz de generar un campo magnético de 3.8 T [20].

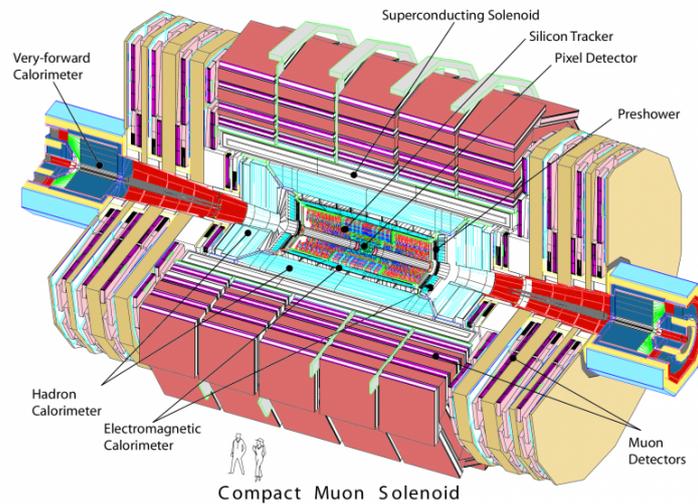


Figura 2.6: Estructura del detector CMS [21].

Los principales componentes de CMS son: el detector de trazas o *tracker*, el calorímetro electromagnético (ECAL), el calorímetro hadrónico (HCAL), el solenoide superconductor y las cámaras de muones. Debido a la importancia que tienen vamos a dedicar un apartado para describir brevemente cada uno de ellos. Lo primero que vamos a definir es el sistema de coordenadas con respecto al cuál se expresarán las medidas.

Se puede encontrar una descripción completa del detector CMS en [22].

2.2.1. Sistema de Coordenadas

El sistema de coordenadas utilizado en CMS tiene el origen situado en el punto de colisión de los haces. La elección más natural, dada la simetría cilíndrica del detector, es utilizar coordenadas cilíndricas. La orientación de los ejes se tomará de forma que el eje z sea tangente a la circunferencia del acelerador y apunte hacia las montañas de Jura en Suiza. El plano XY se denomina plano transverso y se define de forma que el eje x apunte hacia el centro de la circunferencia del LHC.

El ángulo azimutal $\phi \in [0, 2\pi)$, se mide con respecto al eje x en el plano XY , y la coordenada radial en este plano se denota por r . Por último, el ángulo polar $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, se mide con respecto al eje z en el plano YZ y, a partir de él, se define la pseudorapidez η (ec. 2.3). La utilidad de esta magnitud en lugar del ángulo θ radica en el hecho de que es un invariante Lorentz para partículas sin masa y, bajo aproximaciones, también

para partículas ultrarrelativistas.

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

De esta forma, partículas con mucha pseudorapidez estarán más próximas al tubo del haz y, aquellas con pseudorapidez 0, estarán contenidas en el plano transverso. La figura 2.7 ilustra las coordenadas que acabamos de definir. Utilizando estas coordenadas se define el momento transverso \vec{p}_T , como la proyección del momento lineal sobre el plano XY. En este trabajo denotaremos por p_T al módulo del momento transverso:

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

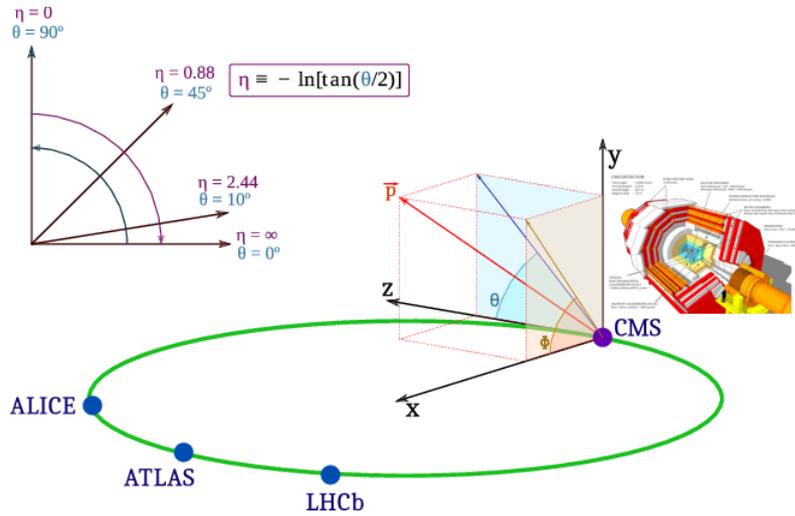


Figura 2.7: Sistema de coordenadas utilizado en CMS [23].

2.2.2. El detector de trazas (tracker)

El detector de trazas (figura 2.9) es el subdetector más interno de CMS. Permite reconstruir las trayectorias de las partículas cargadas que lo atraviesan y, a partir de ellas, inferir su momento y el vértice principal del que procedían. El vértice principal es el punto del que parten las trazas y en el que se ha producido la colisión. Un efecto del *pile-up* es que aumenta el número de señales registradas por el *tracker* y el número de trayectorias posibles también. Este efecto se puede observar en la figura 2.8, donde se ha podido reconstruir el vértice de la interacción de 78 colisiones de *pile-up*. Es necesario poder distinguir la procedencia de cada traza para poder eliminar el efecto del *pile-up* sobre la colisión de interés (no todas las colisiones producen sucesos interesantes desde el punto de vista físico). La alta granularidad y eficiencia del *tracker* permite

distinguir entre ellas, alcanzando una precisión en la determinación de las trayectorias del orden de $10 \mu m$ [24]. Además, altera poco la trayectoria inicial de las partículas que lo atraviesan. Esto permite que medidas posteriores de energía y posición en los calorímetros sean suficientemente precisas.

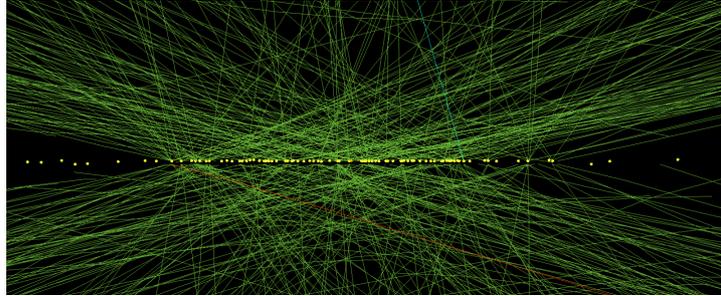


Figura 2.8: Vértices reconstruidos (78 en total) en un suceso de alto *pile-up* [25]. Las líneas verdes se corresponden con las trazas de las partículas cargadas y los puntos amarillos con el vértice de interacción.

También permite reconstruir vértices secundarios de partículas cuya vida media es suficiente como para viajar una distancia resoluble por el detector, este es el caso de los *quarks b*. El proceso de identificación de estos *quarks b* es conocido como *b-tagging* y debido a su importancia será explicado en detalle más adelante. El detector de trazas necesita de un campo magnético que curve la trayectoria de las partículas para, a partir de su curvatura, inferir el momento que tenían (partículas de bajo momento se curvarán más).

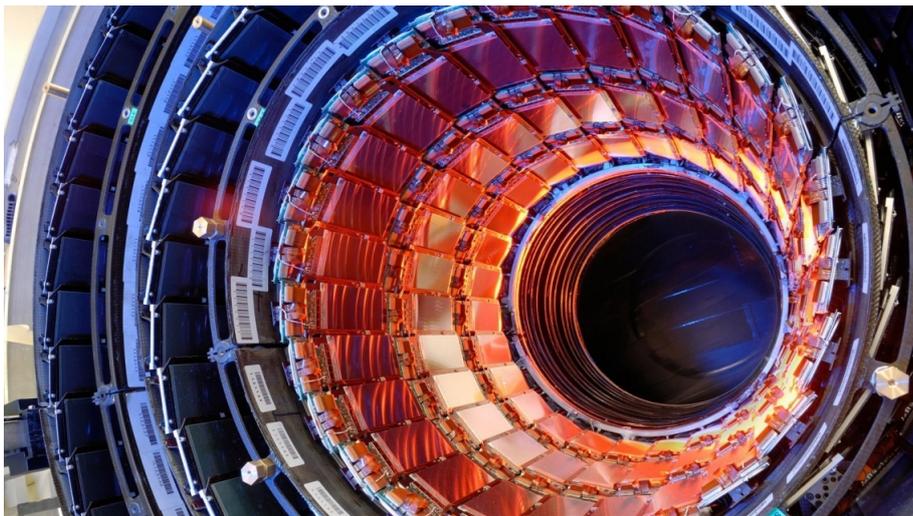


Figura 2.9: Fotografía del detector de trazas de CMS [24].

El detector de trazas está compuesto a su vez de dos partes: los **píxeles de silicio** que se encuentran en la parte más interna del detector y las **tiras de silicio**.

El detector de píxeles cuenta con 124 millones de canales en total y con 6000

conexiones por centímetro cuadrado. Se divide en 4 capas (antes de 2017 el detector de píxeles únicamente contaba con 3 capas) cilíndricas a 2.9 cm, 6.8 cm, 10.9 cm y 16.0 cm de distancia al tubo del haz, y tres anillos externos e internos que extienden la región del *tracker* hasta $|\eta| = 2.5$ [26]. Cuando una partícula cargada lo atraviesa, ioniza a su paso los átomos de silicio más cercanos. Estos electrones libres son recogidos y amplificados para obtener una señal eléctrica del paso de la partícula. Se estima que atraviesan este detector 10 millones de partículas por segundo y por centímetro cuadrado.

Las tiras de silicio se encuentran en la parte más externa del *tracker*. Se organizan en 10 capas cilíndricas concéntricas hasta un radio máximo de 130 cm. Cuatro de estas capas forman la región interna del barril y las otras seis la región externa. Ambas regiones se cierran con dos tapas en las bases de los cilindros.

2.2.3. Calorímetro Electromagnético

El calorímetro electromagnético (ECAL) es el subdetector de CMS encargado de medir la energía de los fotones y electrones. Su diseño y características están fuertemente orientados al descubrimiento del bosón de Higgs. Una forma de detectarlo es a través del pico de resonancia de masa. La anchura de este pico está condicionada por la resolución en energía del ECAL, que se diseñó para que sea la mayor posible. Las características del ECAL son las siguientes [27]:

- Excelente resolución en energía, posición y ángulo hasta un $|\eta| < 3.0$.
- Hermético, compacto y de alta granularidad.
- Respuesta rápida de 25 ns.
- Posibilidad de medir energías desde los pocos GeV hasta los TeV.
- Resistente a la alta radiación producida en las colisiones.

El ECAL está formado por cerca de 80000 cristales de tungstenato de plomo ($PbWO_4$), un material muy denso que emite destellos rápidos y cortos cuando es excitado por radiación ionizante. Dichos destellos son amplificados por un fotomultiplicador y traducidos a una señal eléctrica que lleva información de la energía de la partícula necesaria para un posterior análisis. Cada cristal pesa 1.5 kg y se reparten entre la parte del barril y las tapas del detector en grupos denominados *supermódulos*. Además de estos, el ECAL cuenta con detectores de “precascada” que le permite resolver con mayor precisión pares de fotones de baja energía muy cercanos que pudieran interpretarse como un único fotón de alta energía. Un esquema del ECAL se muestra en la figura 2.10.

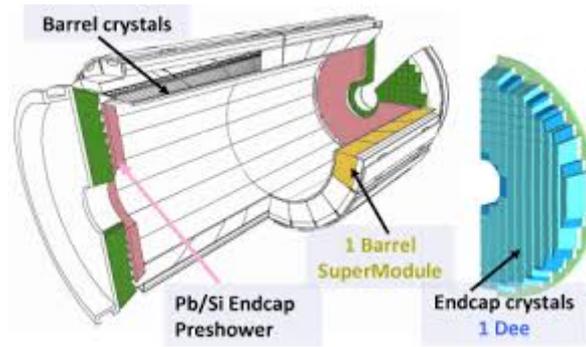


Figura 2.10: Esquema del calorímetro electromagnético de CMS [27].

2.2.4. Calorímetro Hadrónico

Continuando hacia la parte más externa del detector se encuentra el calorímetro hadrónico o HCAL (figura 2.11). Su función es medir la energía de los hadrones producidos en la colisión, tanto cargados como neutros. Además, se diseñó con el objetivo de que fuera lo más hermético posible para que solo partículas poco interaccionantes como los neutrinos o muones pudieran atravesarlo. Esto permite determinar con buena resolución el balance de momento total en el suceso. El HCAL se organiza en capas alternadas de material absorbente y centelleador situadas de tal forma que no existan huecos por los que pudieran escapar las partículas. Se subdivide al igual que el ECAL en una sección de barril que cubre hasta $|\eta| < 1.2$ y dos tapas que extienden el calorímetro hasta $|\eta| < 3$. Cuenta con dos calorímetros *forward* situados al final de CMS para capturar partículas que emergen con muy alta pseudorapidez. Esta última sección utiliza materiales algo diferentes al resto debido a que la radiación que debe soportar es mayor y cubre hasta $|\eta| < 5.2$ [28]. El barril y las tapas cuentan con 5000 canales de medida y el calorímetro *forward* con 1700.

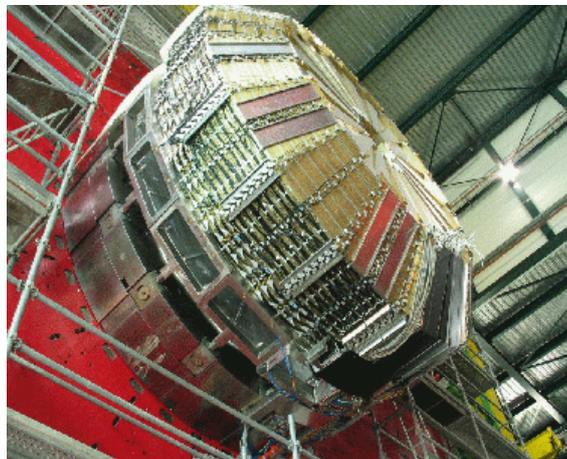


Figura 2.11: Foto de una de las tapas del calorímetro hadrónico [29].

2.2.5. Solenoide superconductor

El solenoide superconductor (figura 2.12) es el dispositivo encargado de generar el campo magnético que curva la trayectoria de las partículas resultantes de la colisión en CMS. Curvar la trayectoria de las partículas da información acerca del momento y carga eléctrica de las mismas, de hecho, a mayor campo magnético más precisas podrán ser estas medidas.

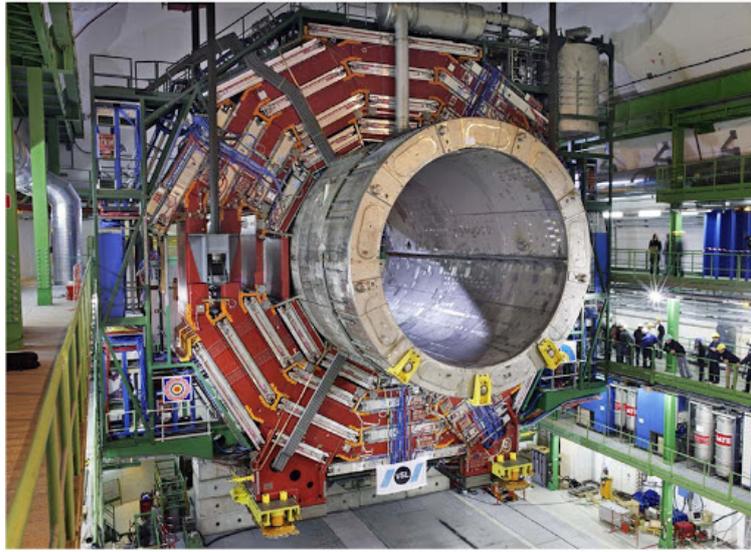


Figura 2.12: Solenoide superconductor de CMS [30].

El solenoide está formado por bobinas de fibras superconductoras que conducen una intensidad de corriente de 18500 A sin sufrir resistencia eléctrica. Para lograr este efecto es necesario enfriar el imán hasta los 4.65 K , gracias a ello, el solenoide es capaz de producir un campo magnético de 3.8 T [31]. Dentro del solenoide se encuentran los calorímetros y el *tracker* y, fuera, las cámaras de muones. El sentido del campo magnético cambia según se mida fuera o dentro del solenoide. Para confinar el campo a solo la región del detector se utiliza un armazón de hierro que sirve como guía del campo magnético. Con todo esto el imán de CMS se ha convertido en el imán superconductor más grande construido hasta la fecha.

2.2.6. Cámaras de muones

Las cámaras de muones forman el subdetector más externo de CMS y su objetivo es detectar y medir el momento de los muones procedentes de las colisiones. Se encuentran fuera del solenoide superconductor y son en tamaño el subdetector más grande de CMS. La trayectoria de los muones se reconstruye utilizando las señales o *hits* que dejan a su paso por las diferentes estaciones de muones con las que cuenta CMS. Estas estaciones

se encuentran acopladas al retorno del imán. Con la información recogida en el *tracker* y en las cámaras de muones, se puede reconstruir la trayectoria completa de los muones a lo largo del detector e inferir de ella tanto la carga como el momento de los mismos. Las cámaras de muones también se utilizan como *trigger* ya que permiten identificar rápidamente sucesos que contienen muones.

La figura 2.13 muestra un cuadrante del sistema de muones de CMS. Cuenta con tres subsistemas de detectores gaseosos basados en diferentes tecnologías: los tubos de deriva (DTs) situados en el barril, las cámaras de tiras catódicas (CSCs) situadas en las tapas y las cámaras de placas resistivas (RPCs) situadas tanto en el barril como en las tapas. En total CMS cuenta con 250 DTs, 540 CSCs y 610 RPCs [32].

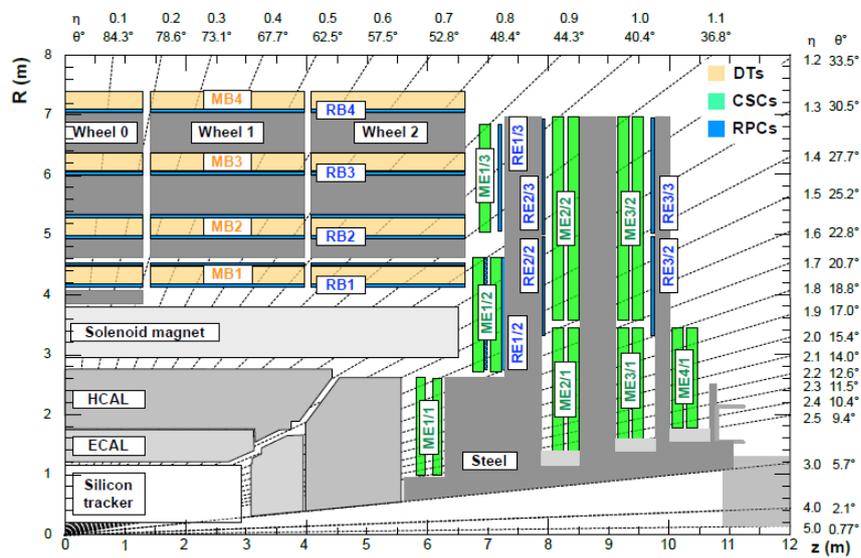


Figura 2.13: Disposición del sistema de detección de muones de CMS[21].

El sistema de DTs se divide a lo largo del eje z en 5 ruedas (*wheels* en la figura 2.13), a lo largo del ángulo ϕ en 12 sectores y cada sector cuenta con 4 detectores de deriva (DT). Cada DT está formado por grupos de pequeñas celdas llenas de una mezcla de gases de Ar (85 %) y CO₂ (15 %) sometidos a una diferencia de potencial del orden 1800 V [21]. Por último, las cámaras de DTs se subdividen en tres o dos *superlayers* (dependiendo del caso) y cada *superlayer* en 4 *layers*. El sistema de DTs se localiza en el barril de CMS y se extienden hasta $|\eta| < 1.2$.

Las CSCs se situán entre $0.9 < |\eta| < 2.4$ en la región de las tapas. Están diseñadas para soportar una mayor radiación que el resto, ya que se situán en una región donde el fondo de neutrones es mayor. Estas cámaras están montadas en cuatro discos perpendiculares al tubo del haz.

Por último, las RPCs se montan junto a las DTs y CSCs hasta un $|\eta| < 1.9$ [21]. Su resolución espacial es menor que la de las DTs y CSCs pero su resolución temporal

es mayor. Esto permite complementar las medidas del resto de cámaras y alcanzar una mayor precisión en la determinación de la trayectoria de los muones.

En su conjunto, el sistema de muones cubre toda la región del barril y se extiende hasta un $|\eta| < 2.4$ [21].

2.2.7. Trigger

Uno de los principales desafíos que tiene la alta luminosidad del LHC es el volumen de datos que genera por segundo. El LHC produce colisiones a una frecuencia de 40 MHz pero la tecnología actual solamente permite guardar datos a una frecuencia de 1 kHz [21]. No todos los sucesos que se producen son interesantes desde el punto de vista físico. Muchos de ellos, son procesos de baja energía donde no se espera que se puedan producir partículas tan masivas como el bosón de Higgs. Por ello, el diseño del sistema de *trigger* está enfocado a identificar cuáles de estas colisiones son interesantes y deben ser guardadas y cuáles pueden ser desechadas. El correcto funcionamiento del *trigger* es crucial, todo suceso que no es guardado es perdido para siempre.

La información de la que dispone el *trigger* es muy poca ya que, a este nivel, aún no se han podido reconstruir todas las trayectorias de las partículas ni parámetros como el momento o la energía. Las primeras decisiones que toma se basan en la observación de firmas (estados finales) con presencia de muones o altas deposiciones de energía en los calorímetros. El *trigger* de CMS está dividido en dos niveles: el Nivel-1 (L1) basado en *hardware*, con una frecuencia de salida de 100 kHz y el Nivel Alto (HLT), basado en *software*.

El *trigger* L1 utiliza información procedente de las cámaras de muones y de los calorímetros dejando para el siguiente nivel la información del *tracker*. Este nivel se subdivide a su vez en *trigger* local, regional y global. En él, se hace una primera reconstrucción de objetos como electrones, muones o *jets* con estimaciones de su momento y energía. Basado en esto el Nivel-1 decide que suceso debe ser descartado y cuál puede pasar al HLT.

El HLT puede acceder a información procedente de todo el detector y llevar a cabo una reconstrucción similar a la que se hace posteriormente en el análisis *offline*², pero hecho de una manera más rápida y menos precisa. El HLT se subdivide a su vez en el Nivel-2 (L2) y Nivel-3 (L3), que son capaces de reconstruir las trayectorias de las partículas y los vértices en un tiempo de 100 ms por suceso. Finalmente, los sucesos seleccionados son clasificados en diferentes rutas denominadas *HLT paths*, cada una

²En este contexto el término *offline* hace referencia al análisis y reconstrucción hechos a posteriori de la colisión, en contraposición al análisis *online* llevado a cabo por el *trigger* instantes después de la colisión.

con unas características específicas del suceso en cuestión.

Capítulo 3

Reconstrucción y simulación de sucesos

3.1. Reconstrucción de sucesos

En el capítulo anterior se expuso el procedimiento de aceleración de los protones en el LHC y su posterior colisión en los puntos de detección. Estas colisiones dan lugar a numerosos tipos de partículas (algunas estables y otras inestables) que son detectadas por su respectivo subdetector dejando algún tipo de señal a su paso como depósitos de energía o destellos de luz. El detector CMS fue construido para efectuar esta tarea de la manera más óptima posible. Gracias al potente imán superconductor que tiene instalado puede medir muy eficientemente partículas cargadas hasta una pseudorapidez de $|\eta| < 2.5$ [22]. Ahora bien, esta información no es suficiente para realizar un análisis físico de las colisiones que se han producido. Lo ideal sería contar con una forma de poder identificar cada depósito de energía con una partícula determinada y asociar esa partícula a una colisión en concreto. De esta forma, se contaría con un conjunto de datos cuyo significado físico sería mucho más evidente y se podría estudiar las interacciones que han tenido lugar y el estado final que han producido. De esta tarea se encarga el algoritmo *particle-flow* [33], que recopila información de todos los subdetectores y la aglutina obteniendo candidatos a objetos físicos como electrones, muones o *jets* y sus respectivos observables físicos como su momento transversal y energía.

Para entender como funciona este algoritmo vamos a fijarnos en la figura 3.1. En ella se representan las trayectorias de diferentes tipos de partículas a lo largo del detector. Como vemos, las partículas cargadas presentan una trayectoria curva y son detectadas por el *tracker*. Las cámaras de muones situadas en la parte más externa de CMS

detectan el paso de los muones. Los electrones depositan su energía en el ECAL, y los hadrones cargados en el HCAL. Por otro lado, los fotones solo son detectados en el ECAL y los hadrones neutros en el HCAL. Como vemos, cada partícula en función de sus propiedades interacciona de una forma u otra con el detector, esto lo podemos utilizar para identificarlas. Un ejemplo sería la reconstrucción de un muon, donde debemos unir la traza que ha dejado en el *tracker* con los *hits* en las cámaras de muones.

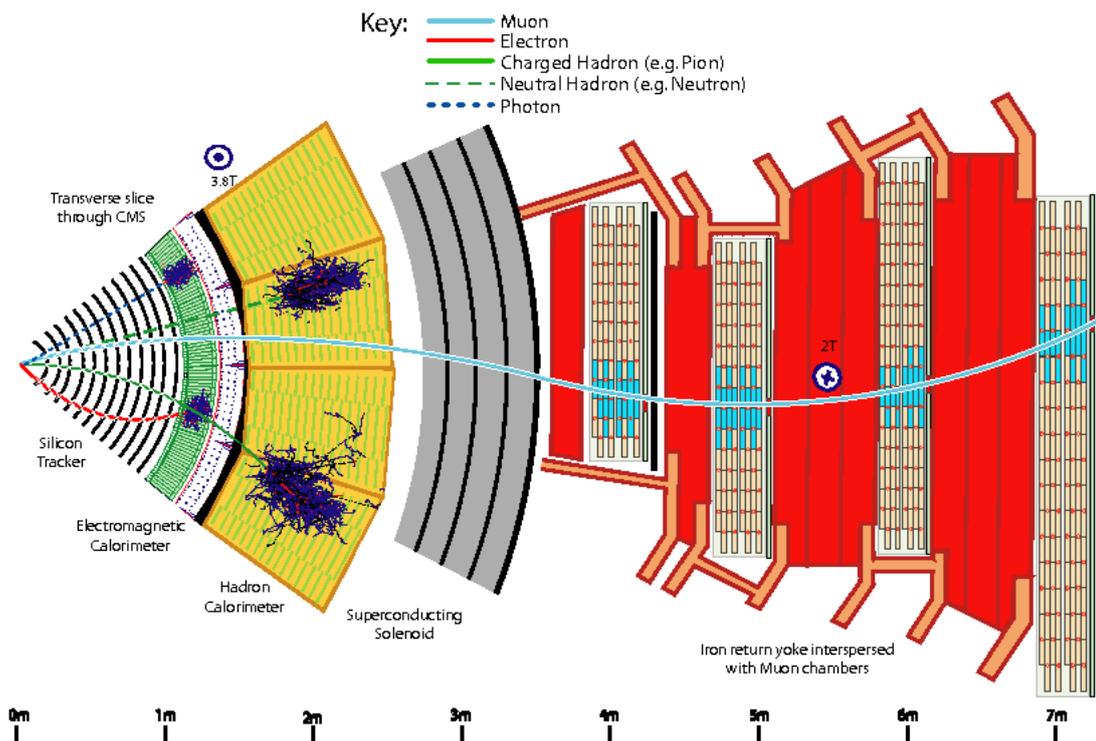


Figura 3.1: Corte transversal de CMS junto a las trayectorias de diversas partículas dentro del detector [33].

La primera tarea que realiza el algoritmo *particle-flow* está enfocada a reconstruir las trayectorias de las partículas (trazas) dentro de cada subdetector. El proceso es similar en el *tracker* y las cámaras de muones. Se parte de las señales que ha dejado la partícula a su paso y mediante un algoritmo iterativo se van uniendo los puntos que pueden corresponder a la trayectoria de una misma partícula. En los primeros pasos solamente se reconstruyen las trazas que se han medido con mayor precisión y, en iteraciones siguientes, se relaja este criterio para reconstruir la mayor cantidad de trazas posibles. En esta fase también se reconstruyen los llamados *clusters* en los calorímetros. Esta operación consiste en agrupar las deposiciones de energía próximas que serían debidas a la interacción de una sola partícula con el detector.

Una vez se ha reconstruido la información de cada subdetector, el algoritmo procede

a relacionar la información de cada uno de ellos. Lo primero que se hace es reconstruir los objetos más sencillos de identificar, los muones. Para ello, se emplea un algoritmo de *enlace* que asocia una traza del *tracker* con otra de la cámara de muones. Una vez reconstruidos éstos, se pasa a los electrones y hadrones cargados usando la información del *tracker* y el ECAL o el HCAL. Por último, los *clusters* que queden en el ECAL se asocian con fotones y los que queden en el HCAL con hadrones neutros.

Tras reconstruir los candidatos a objetos físicos el algoritmo pasa a obtener los observables correspondientes. A partir de la trayectoria se puede obtener el momento de la partícula y, usando los *clusters*, también se puede conocer su energía.

3.1.1. Vértice principal

Se denomina vértice principal (PV) al punto del espacio donde se ha producido la colisión protón-protón. La manera que tiene el algoritmo *particle-flow* de identificarlo es extrapolando las trayectorias reconstruidas en los subdetectores hasta que estas confluyan en un único punto del tubo del haz. El proceso en realidad es algo más complicado ya que a las luminosidades que trabaja el LHC, el *pile-up* complica mucho la tarea ocasionando la aparición de vértices de interacción adicionales. Lo que se suele hacer después de extrapolar las trayectorias, es realizar un ajuste de máxima verosimilitud para estudiar la probabilidad de que esas trayectorias se originasen de un vértice común. Se toma como vértice principal aquel en el que la suma del p_T de las trazas asociadas es más alta. Un aspecto importante en el diseño de este algoritmo consiste en encontrar un equilibrio en la eficiencia de la reconstrucción de vértices muy cercanos. Tiene que ser suficiente como para distinguir los posibles casos de *pile-up*, pero no demasiado alta como para dividir una trayectoria de una partícula en dos diferentes.

3.1.2. Muones

En esta sección vamos a describir brevemente el procedimiento que se utiliza en CMS para la reconstrucción de muones. Como comentábamos en la introducción de este capítulo, los muones son reconstruidos utilizando la información del *tracker* y las cámaras de muones. Primero se procede con una reconstrucción local de la traza en cada subdetector, y luego se utilizan estas trazas como punto de entrada para la reconstrucción global de la trayectoria.

En el *tracker* las trazas se construyen utilizando un algoritmo iterativo que en cada paso elimina los *hits* ya utilizados y utiliza una lógica ligeramente diferente. Las

trazas reconstruidas en las cámaras de muones se denominan *standalone-muon tracks*, y comienzan utilizando información de segmentos de las CSCs o DTs generando una serie de semillas. Estas semillas se utilizan después para construir un candidato a muon usando un filtro de Kalman¹ que tiene en cuenta las señales de todos los subdetectores de muones.

El siguiente paso es asociar las trazas de muones en el *tracker* con las de las cámaras. Para ello, se comienza extrapolando las trayectorias del *tracker* con $p_T > 0.5$ GeV y momento total $p > 2.5$ GeV hacia las cámaras de muones. Si el resultado de la extrapolación coincide con alguna trayectoria reconstruida previamente en las cámaras de muones y si cumple una serie de criterios de consistencia, entonces ambas trazas son unificadas en una sola. Este procedimiento también se realiza al revés, donde la reconstrucción va desde fuera hasta dentro del detector. Cuando ambos resultados coinciden se unifican en uno solo. Los muones reconstruidos de esta forma se denominan *global muons*.

Posteriormente estos muones son utilizados por el algoritmo de *Particle-Flow*, donde además se les imponen algunos criterios de calidad adicionales y, utilizando el resto de información de la que dispone, se le aplican también diversos criterios de aislamiento. Esto último es crucial ya que permite distinguir muones *prompt* de los *non-prompt*. Se denomina muon *prompt* a aquellos muones que proceden de desintegraciones ocurridas en el vértice principal, generalmente de bosones masivos como el W o el Z . Suelen ser muones más aislados y energéticos que los *non-prompt*. Por otro lado, los muones *non-prompt* son los producidos en desintegraciones débiles de hadrones o incluso hadrones identificados incorrectamente como muones. Esto también es aplicable a electrones y se denominan en general leptones *prompt* o *non-prompt*.

El aislamiento se evalúa estimando el momento transverso total de las partículas emitidas a lo largo de la dirección del muon. Se calcula utilizando la ecuación 3.1, donde la suma se extiende a todos los hadrones cargados (h^\pm), fotones (γ) y hadrones neutros (h^0) situados dentro del cono $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$ alrededor del muon con momento transverso p_T . Con esto se establece un punto de trabajo que proporciona un criterio para decidir si el muon se considera aislado o no.

$$I = \frac{1}{p_T} \left(\sum_{h^\pm} p_T^{h^\pm} + \sum_{\gamma} p_T^{\gamma} + \sum_{h^0} p_T^{h^0} \right) \quad (3.1)$$

Tras la reconstrucción, se realiza un proceso de **identificación** para clasificar los

¹El filtro de Kalman es un algoritmo recursivo que sirve para estimar el estado de un sistema dinámico a partir de medidas sujetas a una cierta incertidumbre. Fue desarrollado por Rudolf E. Kalman y publicado en 1960 [34].

muones en diferentes categorías en función de los valores que toman ciertas propiedades. Se crean varias categorías con un doble objetivo. Primero, unificar los criterios de identificación en unos comunes para todos los estudios. Y segundo, permitir a cada análisis utilizar la categoría más adecuada en función de la pureza y eficiencia buscada. Algunos ejemplos de estas categorías serían las siguientes: *loose muon*, *medium muon* y *tight muon* (cuya definición precisa se puede encontrar en [35]). Los muones de tipo *loose* han sido definidos utilizando unos criterios de selección optimizados para identificar muones *prompt* y muones procedentes de desintegraciones débiles de hadrones. Mientras que los muones *tight* suprimen los muones que no se han creado en el vértice principal. Para ello, se exige que el muon haya dejado un número mayor de señales en el *tracker*.

El momento que llevaban los muones es una cantidad cinemática importante de los mismos que se estima a partir de las trayectorias reconstruidas. En función del momento transverso del muon, se utiliza o bien la traza global del muon en todo CMS, o bien únicamente la trayectoria dentro del *tracker*. Esto se hace porque aunque la precisión del *tracker* es mayor que la de las cámaras gracias a su alta granularidad, muones de alto momento transverso dejan prácticamente una línea recta como traza en el *tracker*, lo que dificulta la medida de su curvatura.

3.1.3. Electrones

La reconstrucción de electrones es más complicada que la de los muones. Los electrones son detectados por el *tracker* y por los depósitos de energía que dejan en el ECAL. Sin embargo, a lo largo de su trayectoria van perdiendo energía en forma de radiación *bremstrahlung* como consecuencia de su interacción con el detector lo que hace que su trayectoria pierda la forma helicoidal ideal. Esta trayectoria puede ser reconstruida utilizando un filtro de Kalman, procedimiento habitual para partículas cargadas. Pero en este caso, las grandes pérdidas de energía hacen que sea mejor utilizar otro método más específico.

Se comienza generando una serie de semillas, que consiste en escoger dos o tres *hits* en el *tracker* desde los cuáles se pueda comenzar a reconstruir la traza. Luego se genera la traza correspondiente a estas semillas y se extrapolan hasta el ECAL teniendo en cuenta las pérdidas de energía. El algoritmo de *particle-flow* se encarga en este punto de asociar cada traza con los depósitos de energía correspondientes. Para seleccionar electrones bien reconstruidos y evitar identificar incorrectamente electrones *non-prompt*, se aplican una serie de criterios de calidad a los observables reconstruidos del electrón. Como en el caso de los muones se exigen ciertos criterios de aislamiento para seleccionar electrones *prompt*. También se restringen las distancias angulares ($\Delta\phi$

y $\Delta\eta$) entre la traza extrapolada y el depósito del ECAL asignado.

Otra cantidad que se suele utilizar para discriminar entre electrones *prompt* y *non-prompt* es: $\frac{1}{E} - \frac{1}{p}$ donde E es la energía del electrón y p su momento. Este observable tiende a 0 para electrones *prompt*, mientras que los *non-prompt* presentan valores negativos.

3.1.4. *Jets*

En una colisión de altas energías cuando se producen partículas con carga de color como *quarks* o gluones lo que se observa en el detector es un chorro de partículas originadas por su hadronización. Esos chorros de partículas se les denomina *jets*. Identificar correctamente los *jets* producidos en un suceso permite distinguir entre diferentes procesos. En nuestro caso, el proceso ($t\bar{t}H$) puede dar lugar a estados finales con un número grande de *jets*.

La reconstrucción de *jets* se realiza a través del algoritmo *anti- κ_T* [36], y utiliza los objetos reconstruidos previamente por *particle-flow*. Estos objetos son agrupados en *jets* de forma cónica con una amplitud ΔR . La energía reconstruida de los *jets* debe ser corregida para tener en cuenta diversos factores como: *pile-up*, partículas del *underlying-event*², forma y calibración de los detectores. Estas correcciones se aplican tanto a los datos como a las simulaciones de Montecarlo aunque no tienen porqué ser las mismas. La figura 3.2 esquematiza este procedimiento.

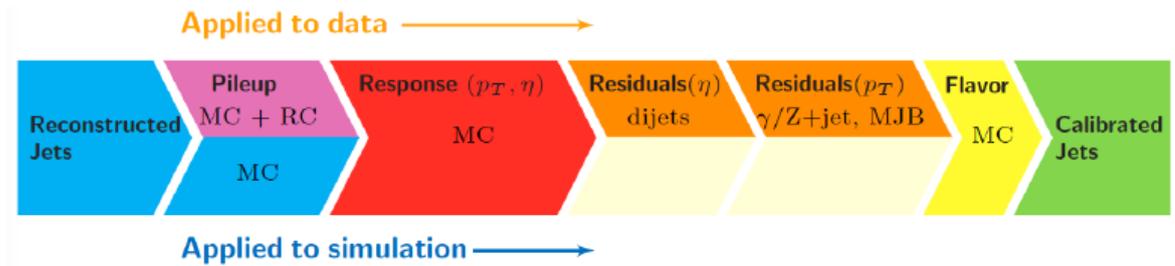


Figura 3.2: Esquema de las correcciones aplicadas a la energía de los *jets* tanto para datos como para las muestras de Montecarlo [21].

Una forma de caracterizar la actividad hadrónica en un suceso es a través del observable H_T :

$$H_T = \sum_i p_T^{jet}[i] \quad (3.2)$$

²Se entiende por *underlying-event* a los procesos (en general poco energéticos) debidos al resto de los constituyentes del protón que no participan en el proceso principal (*hard-scattering*) de la colisión.

Donde la suma se extiende a todos los *jets* del suceso y $p_T^{jet}[i]$ es el momento transversal de cada uno.

Identificación de *jets* *b*

Un tipo especial de *jets* son aquellos originados por la hadronización de un *quark* *b*. Tienen la característica particular de que forman hadrones (B^0 , B^\pm y B_s^0 entre otros) que son capaces de desplazarse una distancia resoluble del vértice principal antes de desintegrarse, formando un vértice secundario. Identificar correctamente estos *jets* permite realizar medidas de mayor precisión y es imprescindible en el estudio de todo proceso que involucre *quarks top*. La figura 3.3 muestra una representación de un suceso con dos *jets* ligeros y un *b-jet*.

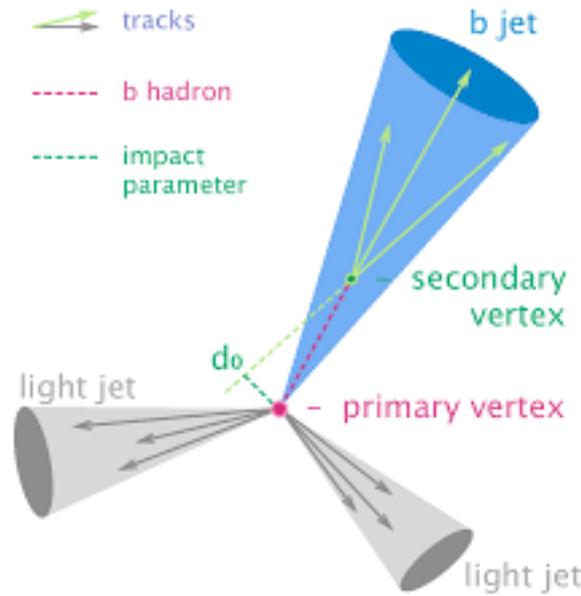


Figura 3.3: Esquema de un suceso con dos *jets* ligeros (en gris) y un *b-jet* (en azul).

Los algoritmos de *b-tagging* se apoyan en las características particulares de estos *jets* para identificarlos. Están basados en técnicas de análisis multivariable (MVA) que tienen en cuenta toda esta información a la vez y producen como salida un discriminante de *b-tagging*. Este discriminante toma valores de 0 a 1 e indica como de probable es que el *jet* observado haya sido generado por un *quark* *b*. Con él, se definen puntos de trabajo que buscan un equilibrio entre eficiencia de detección y probabilidad de identificar incorrectamente un *jet* ligero como *b-jet*. Al igual que en el caso de identificación de muones, se emplean las categorías *loose*, *medium* o *tight b-jet*, que se escogen para que la probabilidad de identificar un *jet* ligero como proveniente de un *quark* *b* sea 10 %, 1 % y 0.1 % respectivamente. Los principales algoritmos están basados en redes neuronales, como por ejemplo: CSVv2, DeepCSV o DeepJet.

3.1.5. Taus

La reconstrucción e identificación de los leptones τ es diferente a la llevada a cabo en los electrones y muones. Su corta vida media hace que sean sus productos de desintegración los que son detectados. Además, la masa del leptón τ es de 1.777 GeV [3], lo que le permite desintegrarse en hadrones. Una tercera parte de las veces se desintegra en un electrón o un muon y dos neutrinos. En este caso, el electrón y el muon pueden ser reconstruidos, pero los neutrinos escapan sin interactuar con el detector. En el resto de casos, se acaba desintegrando hadrónicamente dando lugar a un *jet* de partículas y un ν_τ . Los taus que se desintegran hadrónicamente se denotan por τ_h y se reconstruyen con el algoritmo HPS (*hadron-plus-strip*)[37]. La dificultad de esta tarea radica en ser capaz de distinguir entre *jets* que provienen de la desintegración de un τ y aquellos que no. Típicamente los *jets* provenientes de un τ_h son de menor multiplicidad, depositan su energía en el HCAL en una región más reducida que el resto y se suelen encontrar más aislados. Además, el hecho de que los leptones τ se desintegren siempre en estados con neutrinos dificulta la medida de sus propiedades ya que habrá siempre energía que no ha podido ser detectada.

La reconstrucción de τ_h utiliza los objetos reconstruidos por *particle-flow* (electrones, muones, fotones, hadrones cargados y hadrones neutros) para llevar a cabo la suya propia. El algoritmo HPS utiliza como datos los *jets* reconstruidos anteriormente y, estudiando sus constituyentes, establece si el *jet* reconstruido podría provenir de un τ_h o no.

3.1.6. Momento transverso faltante

Aunque el detector CMS fue diseñado para detectar el mayor número de partículas posibles resultantes de las colisiones hay un cierto tipo que se le escapan, los neutrinos. Al ser partículas neutras que solo interactúan débilmente, no dejan absolutamente ningún rastro en el detector, sin embargo, su presencia puede ser inferida aplicando las leyes de conservación del momento y la energía. En una colisión pp frontal en el LHC se espera que en el plano transversal el momento total sea 0 tanto para el estado inicial como para el final. De esta forma, si la suma de los momentos transversos de todas las partículas en el estado final es distinto de 0, debe haber una o más partículas que no han sido detectadas. Este momento no medido es asociado a los neutrinos y se representa por E_T^{Miss} . Se calcula según la ecuación 3.3 y se conoce como momento faltante en el plano transversal o MET. Este observable es muy importante en el estudio de Física más allá del Modelo Estándar y también juega un papel muy importante en

las medidas de bosones W y procesos con *quarks top*.

$$E_T^{Miss} = \left| \sum_{\text{partículas}} \vec{p}_T \right| \quad (3.3)$$

Idealmente en casos en los que no se hayan producido neutrinos este observable sería cero. Sin embargo, esto no es siempre así ya que la eficiencia del detector no es del 100% y se le pueden escapar otras partículas que no sean neutrinos. Además, el alto *pile-up* de las colisiones también puede influir en esta medida.

El observable H_T^{Miss} complementa la información del momento transverso faltante. Se define como la suma vectorial del momento transverso de los leptones, τ_h y *jets* presentes en el suceso [38].

$$H_T^{Miss} = \left| \sum_{\ell} \vec{p}_{T\ell} + \sum_{\tau_h} \vec{p}_{T\tau} + \sum_{jets} \vec{p}_{Tj} \right| \quad (3.4)$$

Aunque la resolución en este observable es peor que en E_T^{Miss} , las partículas que entran en la suma suelen provenir del *hard-scattering* y no de sucesos de *pile-up*. Esto hace a este observable menos sensible a las variaciones en las condiciones de *pile-up*.

3.1.7. Masa invariante

Se denomina masa invariante m a la energía en reposo de una partícula. Es un invariante Lorentz, es decir, su valor es el mismo en cualquier sistema de referencia. Utilizando el sistema de unidades naturales ($\hbar = c = 1$), se expresa:

$$m^2 = E^2 - |\vec{p}|^2 \quad (3.5)$$

Donde E y \vec{p} son la energía y el momento lineal de la partícula medidos desde cualquier sistema de referencia.

En el caso de un sistema de N partículas, se puede definir análogamente la masa invariante M del sistema:

$$M^2 = \left(\sum_{i=1}^N E_i \right)^2 - \left| \sum_{i=1}^N \vec{p}_i \right|^2 \quad (3.6)$$

Donde E_i y \vec{p}_i son la energía y el momento lineal de la partícula i -ésima medidos desde cualquier sistema de referencia. En este caso, la masa M coincide con la energía del sistema de partículas en el sistema de referencia del centro de masas.

Un caso importante de aplicación de esta magnitud es la reconstrucción de la masa

de una partícula que se ha desintegrado en varias partículas. Si se consigue detectar todas las partículas producto de la desintegración y, reconstruir su momento y energía, entonces se puede determinar la masa de la partícula que las generó a través de la ecuación 3.6. Esta es una de las técnicas que se utiliza para descubrir nuevas partículas y permitió descubrir el bosón de Higgs en 2012.

3.2. Simulación de sucesos

La simulación de sucesos o colisiones en CMS tiene un objetivo múltiple. Por un lado, proporciona una forma directa de comparación entre los datos experimentales y los modelos teóricos, lo que permite comprobar la validez de la teoría o, en caso contrario, encontrar desviaciones que puedan ser señal de nueva física. También se utilizan las simulaciones para realizar cálculos de eficiencias comparando datos con simulaciones en procesos bien caracterizados. La complejidad de los procesos que ocurren en el LHC impide que se puedan obtener resultados analíticos sencillos y, en su lugar, se recurre a simularlos mediante métodos de Montecarlo. Los métodos de Montecarlo tratan de reproducir los procesos que ocurren en el LHC de manera estocástica, generando experimentos “pseudo-aleatorios”.

El proceso de simulación consta de varias etapas. Primero se debe simular la colisión y los procesos que tienen lugar en la misma, este paso se conoce como **generación**. El segundo paso, denominado **simulación**, consiste en simular el paso de las partículas generadas a través del detector. Y por último, se simulan las señales electrónicas de los detectores a partir de las interacciones de las partículas con los materiales activos. Este último paso se conoce como **digitalización**.

El procedimiento de reconstrucción que se lleva a cabo sobre los datos simulados es el mismo que el que se realiza sobre los datos experimentales. Esto se hace para reproducir de la manera más fiable posible los datos reales.

3.2.1. Generación

En este paso, se generan todos los posibles diagramas de Feynman para el proceso considerado. Al existir un número infinito de ellos se generan los más relevantes, es decir, los que más contribuyen a la sección eficaz del proceso. En este sentido, se clasifican los diferentes generadores en función del orden en teoría de perturbaciones que son capaces de simular. La expansión normalmente se hace en potencias de α_S , la constante de acoplamiento de la interacción fuerte, aunque también se pueden considerar órdenes en EWK (electrodébil). La terminología utilizada es la siguiente:

LO (*Leading Order*) para el orden fundamental, NLO (*Next to Leading Order*) para el primer orden, NNLO (*Next to Next to Leading order*) para el segundo orden, etc. Existen distintos modelos para la generación de sucesos: *MadGraph* permite obtener información a LO y, *PowHeg* y *aMC@NLO* a NLO. Estas simulaciones permiten obtener valores teóricos de la sección eficaz total del proceso, pero también de la sección eficaz diferencial, partículas involucradas y sus distribuciones de energía y momento. La generación consta de varias etapas, un esquema de todas ellas se ilustra en la figura 3.4.

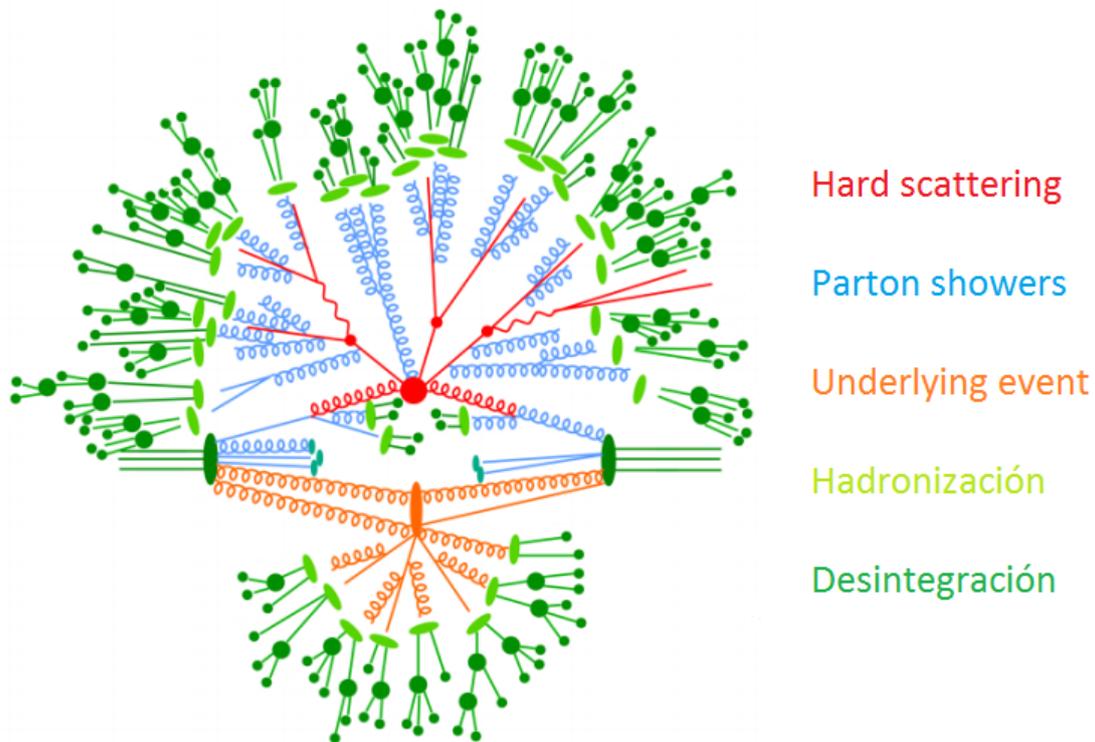


Figura 3.4: Esquema de los pasos en los que se divide la generación de un suceso.

En la primera etapa se simula el proceso principal (***hard-scattering***) a partir de las PDFs (*Parton Distribution Functions*) de los protones que van a colisionar. Es en esta etapa donde se generan los distintos diagramas de Feynman y los elementos de matriz. Las PDF dan la probabilidad de encontrar un partón (*quark* o *gluón*) en un hadrón como función de la fracción de momento x del protón que lleva el partón. Son determinadas experimentalmente a partir de experimentos de dispersión inelástica ya que la estructura del protón está por debajo de las escalas de energía en las que se pueden realizar cálculos perturbativos.

A este paso le siguen las cascadas de partones (***parton showers***) que simulan los procesos de emisión de gluones y *quarks* de las partículas con carga de color presentes en el suceso. Esto puede ocurrir antes de que se produzca el *hard-scatter* y se consideraría

radiación en el estado inicial (ISR) o después, y se consideraría radiación en el estado final (FSR). La simulación de estos procesos se lleva a cabo con el paquete de *software Pythia* [39].

Los partones que no participan en el *hard-scattering* también interaccionan y pueden ser detectados. Son procesos menos energéticos que el principal y se denominan globalmente como *underlying-event*. Su simulación se realiza también a través de *Pythia*.

Tras producirse la cascada de partones el siguiente paso que se realiza es la **hadronización** que transforma los *quarks* (estados de color) en hadrones (estados sin color). Este proceso ocurre cerca del régimen no perturbativo de QCD donde α_S es cercana a 1. El *software* encargado de esta parte es *Pythia*.

Por último, muchos de los hadrones producidos en el paso anterior tienen vidas medias muy cortas, de esta manera también se simulan los procesos de **desintegración** que sufren.

El efecto de los sucesos de *pile-up* también debe ser tenido en cuenta en la simulaciones. De esta tarea se encarga el *software Pythia*, que agrega este efecto a las simulaciones teniendo en cuenta el perfil de luminosidad de cada año.

Al final, el resultado que se obtiene de la generación es un conjunto de partículas de las que se conoce su momento y energía y que ofrecen una imagen completa del suceso en cuestión. Todo esto ocurre en una fracción de tiempo muy pequeña, tanto, que las partículas aún no han podido escapar del tubo del haz. El siguiente paso que se realiza es, por lo tanto, la simulación de la interacción de estas partículas con el detector.

3.2.2. Simulación y digitalización

Una vez generado el suceso se procede a simular la interacción de cada una de las partículas producidas, con todos los subdetectores de CMS. Existen dos formas de realizar esta tarea, a través de una simulación completa del detector (*Full-Sim*) o a través de una simulación más rápida (*Fast-Sim*).

La simulación completa tiene en cuenta todas las características del detector como: material, geometría, etc, y también el tipo de partícula con la que interacciona. El paquete de simulación *Geant4* [40] se encarga de esta tarea y produce como resultado las deposiciones de energía y señales electrónicas de cada subdetector. Estas simulaciones son muy precisas, pero requieren grandes recursos computacionales. Las señales electrónicas simuladas pueden ser utilizadas por los algoritmos de reconstrucción de manera análoga a las señales generadas en los datos reales.

La simulación rápida permite reducir el tiempo de computación empleado en

la simulación completa produciendo unos resultados igualmente precisos. Para ello, es necesario tomar ciertas simplificaciones y parametrizaciones del detector. Estos parámetros son ajustados a través de simulaciones completas con *Geant4* y sus resultados son validados con los obtenidos por simulación completa y también a través de datos experimentales. Al igual que en la simulación completa, estos resultados pueden ser utilizados por los mismo algoritmos de reconstrucción que los datos experimentales.

3.3. Herramientas de software

Todos los análisis realizados en este trabajo se han programado en los lenguajes *Python* y *C++* utilizando ROOT [41]. ROOT es un conjunto de herramientas programadas en *C++* y compatibles con *Python* que permiten, entre otras cosas, trabajar con grandes cantidades de datos, realizar análisis estadísticos y visualizaciones. ROOT es utilizado en prácticamente todos los experimentos de física de partículas y astropartículas, así como en numerosos experimentos de física médica o física nuclear. Este trabajo se ha desarrollado dentro del paquete *CMG Tools* integrado dentro del marco de trabajo (*framework*) *CMSSW* [42] desarrollado por la colaboración CMS. Para la gestión de versiones y desarrollo del código se ha utilizado *Git* y *Github*.

Capítulo 4

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

El objetivo de este trabajo es determinar la fuerza de señal $\mu = \frac{\sigma_{exp}}{\sigma_{SM}}$, definida como el cociente entre la sección eficaz experimental y la predicha por el Modelo Estándar, del proceso $t\bar{t}H$ en estados finales multileptónicos. Para ello, se utilizará un método estadístico para obtener el valor más probable de μ según los datos observados. Se analizará también las fuentes de incertidumbre, tanto estadísticas como sistemáticas, que más afectan a la medida y como se modelizan dentro del análisis.

4.1. Señal y fondo

En los experimentos de colisiones a altas energías de partículas lo único que se conoce son los estados finales de las colisiones. Estos estados finales se corresponden a las partículas estables que han podido alcanzar el detector. Sin embargo, no es posible conocer el proceso que ha tenido lugar en la colisión a través de las partículas presentes en el estado final. Esto es debido a que existen muchos procesos que producen estados finales iguales o muy parecidos resultando imposible diferenciarlos.

Dos conceptos importantes y a los que se hará referencia continuamente en este trabajo son la señal y el fondo. La señal es el proceso o procesos objeto de estudio en el análisis, en este caso $t\bar{t}H$. El fondo es el resto de procesos que producen estados finales similares a la señal. Estos fondos pueden ser reducidos aplicando criterios de selección sobre los observables adecuados que permitan discriminar la señal del fondo. Esto define una región en el espacio de fases enriquecida en el proceso de señal y denominada región de señal.

4.2. Estados finales objeto de estudio

Los canales multileptónicos se subdividen en función del número de leptones (electrones y muones en este contexto) y τ_h en el estado final. En este análisis vamos a considerar los siguientes dos casos:

- Estados finales con dos leptones con la misma carga eléctrica ($2\ell ss$).
- Estados finales con cuatro leptones (4ℓ).

La figura 4.1 representa un diagrama de Feynman del proceso $t\bar{t}H$ y su posterior desintegración hacia un estado final con dos leptones del mismo signo.

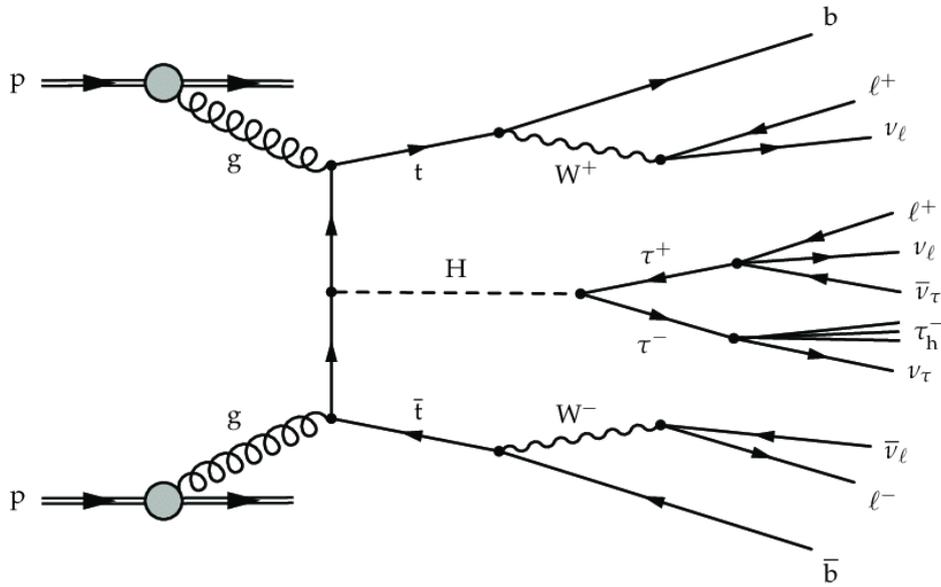


Figura 4.1: Un ejemplo de diagrama de Feynman de la producción de $t\bar{t}H$ en el LHC y su posterior desintegración en un estado final con dos leptones del mismo signo y un τ_h [38].

El canal $2\ell ss$ abarca mayoritariamente el canal de desintegración de $t\bar{t}H$ donde el bosón de Higgs se desintegra a dos bosones W . Por otro lado, los *quarks top* se desintegran antes de hadronizarse, y casi el 100% de las veces lo hacen hacia un *quark b* y un W . Tenemos por lo tanto la cadena: $t\bar{t}H \rightarrow 2b + 2W^+ + 2W^-$. También la desintegración $H \rightarrow \tau^+\tau^-$ puede producir dos leptones del mismo signo. Sin embargo, como se explicará más adelante, se van a aplicar criterios de selección que reducen esta contribución. Para tener dos leptones del mismo signo, o bien los dos bosones W^+ se desintegran en leptones, o bien lo hacen los W^- . En cualquier caso, los bosones W que no se desintegren a leptones se desintegrarán en *quarks* que darán lugar a *jets* de partículas. Resumiendo la cadena de desintegración completa: $t\bar{t}H \rightarrow 2b + 2W^+ + 2W^- \rightarrow 2b$ (*b-jets*) + 2ℓ ($2\ell ss$) + 2ν (E_T^{Miss}) + $4q$ (*jets*).

En el canal 4ℓ la desintegración objetivo del bosón de Higgs es $H \rightarrow ZZ^* \rightarrow 4\ell$. Este estado final donde los 4 leptones provienen de la desintegración del bosón de Higgs permite reconstruir su pico de masa a través de la masa invariante del sistema de 4 leptones. Los estados finales esperados son por lo tanto: $t\bar{t}H \rightarrow 2b + 2W + 2Z \rightarrow 2b$ (*b-jets*) + $4q$ (*jets*) + 4ℓ .

Una ventaja de realizar el análisis en estados finales que contienen leptones es que permiten discriminar muchos de los fondos presentes en un colisionador hadrónico. Esto es debido a que muchas de las colisiones que ocurren entre protones no dan lugar a leptones aislados y energéticos como los que buscamos en este caso. Además, los leptones son las partículas que con mayor precisión se miden en CMS. Sin embargo, la dificultad radica en que la sección eficaz esperada del proceso $t\bar{t}H$ es muy baja (0.5 pb aproximadamente) en comparación a otros procesos producidos en las colisiones. Esto se puede observar en la figura 4.2, en ella se representan las secciones eficaces experimentales de muchos procesos medidos en CMS a diferentes energías en centro de masas. Además, al restringirnos a regiones de señal con varios leptones, estamos reduciendo el espacio de fases a la mitad, puesto que los casos en los que el H se desintegra a $b\bar{b}$ no son seleccionados.

Por todas estas razones, este proceso es muy difícil de observar y se va a encontrar enmascarado por los procesos de fondo que van a contribuir mucho más que él. Consideramos dos tipos de fondos en nuestro análisis: los fondos irreducibles y los reducibles. Los fondos irreducibles hacen referencia a procesos que producen estados finales muy similares a la señal. Este tipo de fondos son simulados mediante métodos de Montecarlo. Por otro lado, los fondos reducibles son en general procesos claramente distinguibles de la señal a partir de una selección adecuada. Su sección eficaz suele ser mucho mayor y debido a errores en la medida pueden pasar la selección. En este análisis, un fondo se considera reducible si al menos un electrón o muon identificados dentro de la categoría *tight* son debidos a un leptón *non-prompt*, es decir, originado de la desintegración de un hadrón o un hadrón identificado incorrectamente [38]. Este tipo de fondos se les denomina *fakes*. En la región de $2\ell ss$ aparecen otras fuentes adicionales de fondos reducibles originados por la medida incorrecta de la carga de un leptón (*flips*), y también fondos originados por la producción de pares de *quarks top* junto con fotones. Los fotones pueden sufrir el proceso $\gamma \rightarrow e^+e^-$ y dar lugar a un fondo denominado *conversions*. Los fondos reducibles se pueden estimar a partir de los datos o simularlos mediante métodos de Montecarlo, como haremos en nuestro análisis.

que no tiene estados finales que contengan 4ℓ . En este canal es mayor la contribución de $t\bar{t}Z$ con dos leptones procedentes del bosón Z y los otros dos de los *quarks top*.

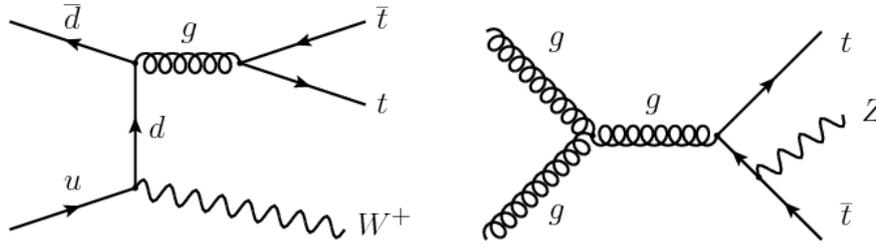


Figura 4.3: Diagramas de Feynman para los procesos de fondo $t\bar{t}W$ y $t\bar{t}Z$ (de izquierda a derecha).

Los procesos WZ y ZZ (figura 4.4) también se incluyen dentro del fondo irreducible. Estos procesos son de muy alta sección eficaz en comparación con $t\bar{t}H$ como se puede ver en la tabla 4.1. Sin embargo, su contribución es menor que la de $t\bar{t}W$ y $t\bar{t}Z$ ya que se espera que produzcan menor número de *jets* que la señal y, en concreto, no presentan *jets* procedentes de un *quark b*. Por lo tanto, exigiendo en los criterios de selección un número elevado de *jets* en el estado final y alguno de ellos etiquetado como procedente de un *quark b*, podemos reducir estos fondos.

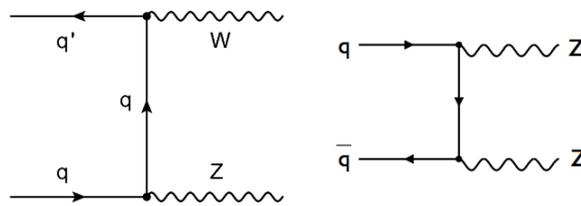


Figura 4.4: Diagramas de Feynman para los procesos de fondo WZ y ZZ (de izquierda a derecha).

También se incluyen muestras de tHq , tHW , $t\bar{t}WW$ y algunos procesos denominados raros como $t\bar{t}t\bar{t}$ que tienen una sección eficaz muy pequeña pero que en conjunto suponen una contribución no despreciable. El grupo de procesos raros se representa en las gráficas como una única categoría, pero es importante resaltar que dentro de ella se incluyen varios procesos.

4.3.2. Fondos reducibles

Como fondos reducibles se incluyen $t\bar{t}$, Drell-Yan, *single top*, WW , $W + jets$ y $Z + \gamma$. Los diagramas de Feynman de estos procesos se muestran en la figura 4.5. La

Proceso	σ (pb)	Generador
$t\bar{t}H$	0.507	<i>Powheg + Pythia</i>
$t\bar{t}W$	0.601	<i>aMC@NLO + Pythia</i>
$t\bar{t}Z$	0.839	<i>aMC@NLO + Pythia</i>
$t\bar{t}WW$	0.007	<i>Madgraph + Pythia</i>
WZ	45.0	<i>aMC@NLO + Pythia</i>
ZZ	16.9	<i>Powheg + Pythia</i>
tHq	0.074	<i>Madgraph + Pythia</i>
tHW	0.015	<i>Madgraph + Pythia</i>
$t\bar{t}$	833	<i>Powheg/Madgraph + Pythia</i>
Drell-Yan	6230	<i>aMC@NLO + Pythia</i>
WW	119	<i>Madgraph + Pythia</i>

Tabla 4.1: Secciones eficaces de los procesos más importantes en este análisis y los generadores utilizados en el proceso de simulación de cada uno.

contribución de los fondos reducibles es muy importante en el canal $2\ell ss$, sin embargo, en canal 4ℓ como se verá más adelante su contribución es mucho menor.

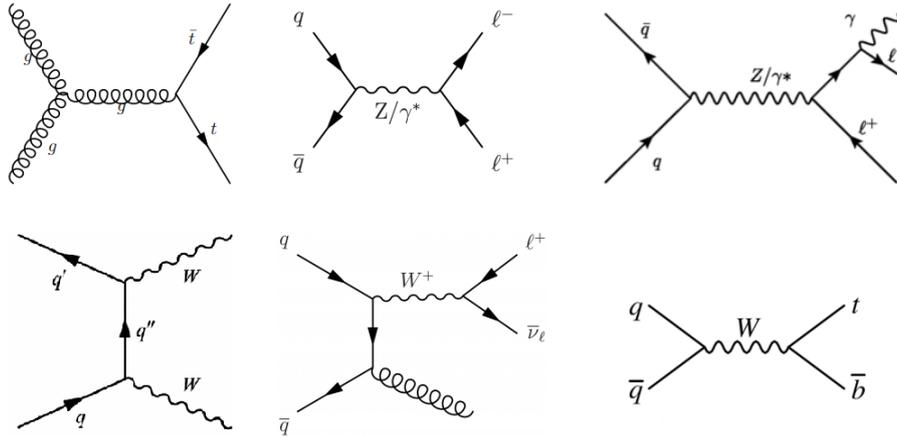


Figura 4.5: Diagramas de Feynman para los procesos de fondo $t\bar{t}$, Drell-Yan, $Z + \gamma$, WW , $W + jets$ y *single top* (de izquierda a derecha y de arriba a abajo).

El proceso $t\bar{t}$ es el que supone la mayor contribución al fondo reducible de leptones *non-prompt* en el canal $2\ell ss$. Aunque no tiene estados finales con dos leptones del mismo signo, puede ocurrir que algún hadrón procedente de un *jet* sea identificado incorrectamente como un leptón. También puede ocurrir que algún leptón producido en un *jet* sea identificado como un leptón procedente de una desintegración en el vértice principal. Esto es poco probable, pero como la sección eficaz de $t\bar{t}$ es tan alta (833 pb) en comparación con el proceso de señal, supone una contribución importante a los sucesos totales. Este proceso también puede contribuir al fondo de *flips*. Esto puede ocurrir cuando los dos *quarks top* se desintegran a dos leptones y la carga de uno de ellos se mide incorrectamente. Por último, también se incluye la contribución de este

proceso al fondo de *conversions* cuando se produce con fotones ($t\bar{t} + \gamma$).

En el proceso de Drell-Yan dos *quarks* se aniquilan para dar lugar a un fotón o bosón Z que posteriormente se desintegra a leptones. La sección eficaz de este proceso es muy alta (6230 pb) y, al igual que en el proceso de $t\bar{t}$, puede ocurrir que se identifique mal alguna de las partículas producidas en un *jet* dando lugar a un fondo de leptones *non-prompt*. También puede medirse mal la carga de un leptón y dar lugar a dos leptones del mismo signo y contribuir en el fondo de *flips*. La mayor contribución del fondo de Drell-Yan se localiza en la región de bajo número de *jets* y puede ser reducido eficientemente aplicando criterios de selección sobre estos observables.

De manera similar a los anteriores, los procesos *single top*, WW y $W + jets$ pueden contribuir en la región de señal como fondos de leptones *non-prompt* y $Z + \gamma$ en *conversions*.

Los fondos de *flips* y *conversions* se agrupan en una única categoría en las gráficas, ya que la contribución de cada proceso en particular no es relevante en este estudio. Por otro lado, los fondos de leptones *non-prompt* se separan en función del proceso ya que al final la mayor contribución proviene de $t\bar{t}$.

4.4. Selección de sucesos

Con el objetivo de discriminar los sucesos de señal de los fondos, se aplican una serie de criterios de selección. En este caso, al extender el estudio a dos estados finales posibles, se definen criterios de selección independientes para cada canal de desintegración.

4.4.1. Canal $2\ell ss$

Este canal se restringe a estados finales que contengan dos leptones del mismo signo, es decir, cuya carga eléctrica sea la misma. Los datos se han recogido con una combinación de *triggers* basados en la presencia de 1 o 2 leptones. En base al estado final buscado y a los procesos de fondo presentes en el experimento, se definen los criterios de selección. Los criterios aplicados en este caso han sido:

- Momento transversal del leptón más energético mayor que 25 GeV y momento transversal del segundo más energético mayor que 15 GeV. Estos criterios están basados en los requisitos del *trigger*.
- Únicamente puede haber dos y solo dos leptones identificados como *tight* en el suceso. Esta categoría está optimizada para identificar los leptones *prompt* del

suceso. En nuestro caso, únicamente esperamos dos según el estado final objetivo.

- El signo de la carga eléctrica de los dos leptones debe ser el mismo, esto reduce considerablemente el fondo de $t\bar{t} + jets$. Este criterio de selección limita notablemente los fondos del experimento, ya que los procesos dentro del SM que dan lugar a dos leptones del mismo signo son muy raros y además de baja sección eficaz.
- Se exige también ciertos criterios de calidad sobre la medida de la carga de los leptones. Esto se hace para reducir el fondo procedente de procesos que producen dos leptones de carga opuesta en los que la carga de uno de ellos se ha determinado incorrectamente. Al contrario que la producción de pares de leptones del mismo signo, las parejas de signo contrario se producen más a menudo, por ejemplo, en las desintegraciones de partículas neutras como el bosón Z.
- Sucesos que contengan parejas de leptones cuya masa invariante sea menor de 12 GeV son rechazados porque están mal modelizados por las simulaciones de Montecarlo [38].
- Veto sobre el pico del bosón Z para electrones: $|m_Z - m_{ee}| > 10$ GeV, donde $m_Z = 91.2$ GeV y m_{ee} la masa invariante del sistema de dos electrones. La probabilidad de medir la carga incorrectamente es mayor para electrones que para muones, por eso, se exige este criterio y el siguiente solo para los electrones. Esto permite rechazar sucesos que procedan de procesos como $Z + jets$ (Drell-Yan).
- Para sucesos con electrones se exige: $L_D = 0.6E_T^{Miss} + 0.4H_T^{Miss} > 30$ GeV. Este observable es una combinación lineal de E_T^{Miss} y H_T^{Miss} , que son magnitudes sensibles a partículas invisibles no detectadas. La resolución de E_T^{Miss} es mayor que la de H_T^{Miss} . Sin embargo, esta última es más robusta porque no es tan sensible a los procesos de baja energía y a las condiciones de *pile-up*. Este criterio se aplica únicamente sobre electrones porque la probabilidad de medir mal su carga es mayor. Se puede eliminar parte de la contribución de Drell Yan con el criterio de selección en L_D puesto que normalmente el momento transverso faltante es pequeño en comparación con $t\bar{t}H$. En la figura 4.6 se puede ver cómo este criterio separa los procesos de Drell-Yan de $t\bar{t}H$.
- $|m_Z - m_{Z1}| > 10$ GeV. El observable m_{Z1} sobre el que se aplica este criterio de selección es la masa invariante del par de leptones de carga opuesta y mismo sabor que dista menos de la masa del bosón Z. Con este criterio se reduce la contribución de todos los procesos que contienen bosones Z, tanto los fondos,

como la señal cuando se desintegra el bosón de Higgs a dos bosones Z . El caso $H \rightarrow ZZ$ es estudiado en el canal 4ℓ .

- El número de leptones τ identificados debe ser 0. Este criterio de selección permite excluir del análisis la categoría $2\ell ss + 1\tau_h$ que proviene mayoritariamente del modo de desintegración del Higgs: $H \rightarrow \tau^+\tau^-$. Se aplica porque en el caso de realizar una medida conjunta del proceso $t\bar{t}H$ en varias categorías, estas deben ser disjuntas. Es decir, los sucesos contenidos en una no deben pertenecer a la otra.
- El suceso debe contener 4 o más *jets* con $p_T \geq 25$ GeV y, al menos 1 de ellos debe proceder de la hadronización de un *quark* b . Los *jets* etiquetados como b , deben pertenecer a la categoría *medium*. Este criterio de selección se ha escogido en base a lo observado en la figura 4.7. En ella se representa en el eje de abscisas el número de *jets* y *b-jets* identificados en el suceso y, en el eje de ordenadas, el número de sucesos seleccionados que contienen ese número de *jets* en escala logarítmica. La gráfica inferior representa el cociente datos/predicción para cada categoría junto con la incertidumbre estadística y total (sistemática + estadística). En la gráfica superior se representa con colores el número de sucesos de cada proceso apilados y con un punto los datos experimentales. La banda rayada representa la incertidumbre total. Lo que se observa es que a mayor número de *jets* se reducen los fondos de WZ , $W + jets$, *conversions* y Drell-Yan. El proceso de señal $t\bar{t}H$ (en rojo en la gráfica) contribuye a alto número de *jets* y con al menos un *b-jet* identificado.

La tabla 4.2 recoge de manera más resumida los criterios de selección aplicados en este análisis.

Selección	$2\ell ss$
Desintegración de $t\bar{t}H$ objetivo	$t \rightarrow b + \ell\nu$ $t \rightarrow b + qq'$ $H \rightarrow WW \rightarrow \ell\nu + qq'$
Trigger	Un leptón o dos leptones
p_T leptones	$> 25/15$ GeV
Identificación de leptones	Únicamente 2 leptones <i>tight</i>
Criterios de carga	Criterios de calidad + mismo signo
Jets y b-tag	≥ 4 <i>jets</i> y ≥ 1 b-tag
MET	$L_D = 0.6E_T^{Miss} + 0.4H_T^{Miss} > 30$ GeV
Identificación de τ_h	$0\tau_h$
Masa invariante de los leptones	$m_{\ell\ell} > 12$ GeV $ m_Z - m_{Z1} > 10$ GeV $ m_Z - m_{ee} > 10$ GeV

Tabla 4.2: Selección de sucesos en la categoría $2\ell ss$

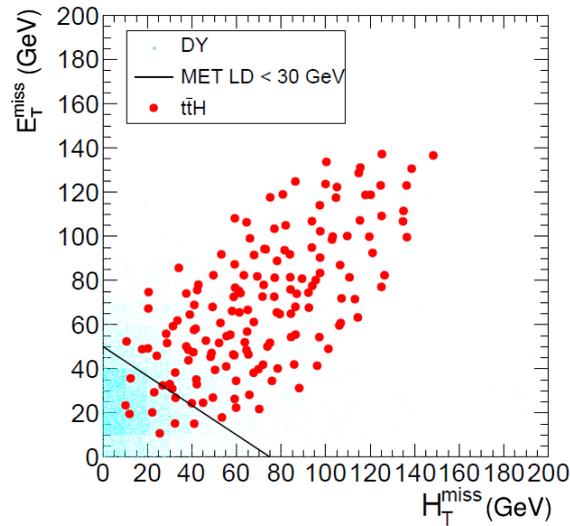


Figura 4.6: Representación de E_T^{Miss} frente a H_T^{Miss} . La línea recta negra se corresponde con el criterio de selección $L_D > 30$ GeV. Los sucesos por encima de la línea pasan la selección y los que están por debajo no. De esta forma, se elimina la contribución del fondo de Drell-Yan.

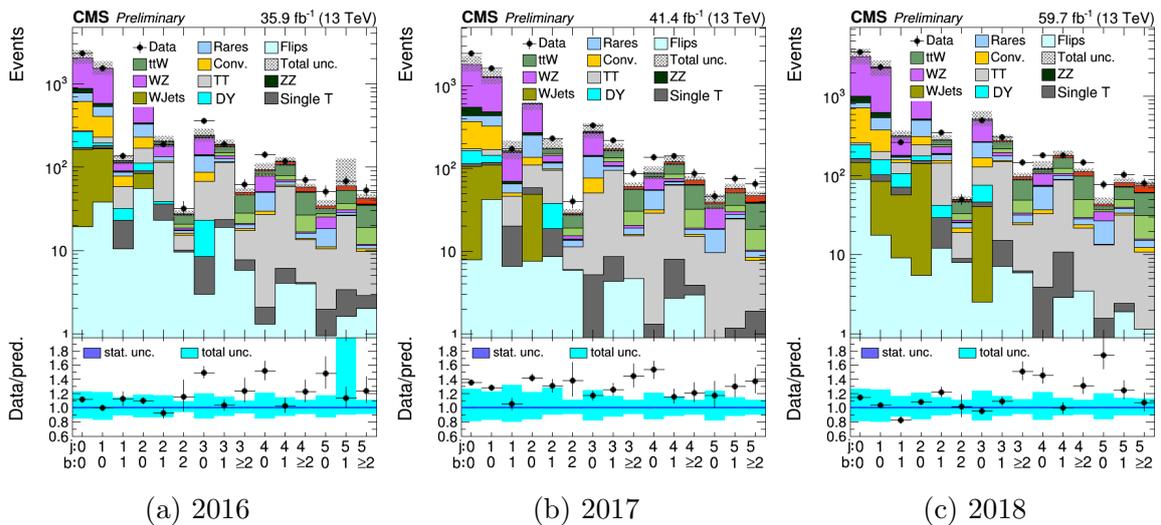


Figura 4.7: Distribución del número de *jets* y *b-jets*. Los criterios de selección aplicados a los sucesos mostrados en esta figura son los definidos para la región de señal exceptuando las restricciones en el número de *jets*. La señal ttH se representa en color rojo.

4.4.2. Canal $H \rightarrow ZZ^* \rightarrow 4\ell$

Los sucesos seleccionados en este canal son aquellos que contienen 4 leptones en el estado final. El interés de este canal es que permite estudiar el pico de masa del bosón de Higgs a partir de la masa invariante de los 4 leptones. Para la preselección de los datos se ha utilizado una combinación de *triggers* basados en la presencia de 1, 2 o 3 leptones. Teniendo en cuenta el canal de desintegración $t\bar{t}H \rightarrow 2b + 2W + 2Z$, se presentan dos posibilidades:

- Los 4 leptones proceden de la desintegración de los bosones Z, y los W dan lugar a 4 *jets*. En este caso se esperaría por lo tanto 6 *jets* de los cuáles dos provienen de un *quark* b y al no haber neutrinos no se espera momento transverso faltante. La cadena de desintegración completa: $t\bar{t}H \rightarrow 2b + 2W + 2Z \rightarrow 2b$ (b -*jets*) + $4q$ (*jets*) + 4ℓ . Este caso es el más interesante puesto que permite reconstruir la masa del Higgs a partir de la masa invariante del sistema de 4 leptones.
- Los 4 leptones proceden 2 de un bosón Z y los otros 2 de los W. El bosón Z restante puede desintegrarse en 2 neutrinos o en 2 *quarks*. Habría por lo tanto en el estado final momento transverso faltante y 4 o 2 *jets*. Queda por lo tanto la siguiente cadena de desintegración: $t\bar{t}H \rightarrow 2b + 2W + 2Z \rightarrow 2b$ (b -*jets*) + $4\ell + 2\nu/4\nu$ (E_T^{Miss}) + $2q/0q$ (*jets*).

Con estas consideraciones se han definido los siguientes criterios de selección:

- El suceso seleccionado debe tener al menos 4 leptones identificados. Los criterios de identificación son menos restrictivos en este caso que los exigidos en la categoría *tight* empleada en el canal $2\ell ss$. La razón es que el número de sucesos de leptones *non-prompt* esperado es pequeño y por lo tanto se pueden relajar los criterios de identificación. Además, esto permite aumentar el número de sucesos seleccionados totales.
- El momento transverso de los leptones debe satisfacer: $p_T^1 > 25$ GeV, $p_T^2 > 15$ GeV, $p_T^3 > 15$ GeV y $p_T^4 > 10$ GeV. El superíndice denota 1 para el leptón más energético en el suceso y 4 para el menos energético. Este criterio está basado en los requisitos del *trigger*.
- El número de leptones τ en el suceso debe ser 0. Este criterio se ha mantenido siguiendo los pasos definidos en el canal $2\ell ss$.
- Se rechazan sucesos que contengan parejas de leptones cuya masa invariante sea menor que 12 GeV por la misma razón que en la selección del canal $2\ell ss$.

- El número de *jets* con $p_T \geq 25$ GeV debe ser superior o igual a 3. Además, se exige que de ellos al menos 1 proceda de la hadronización de un *quark b* identificado con la categoría *medium* o al menos 2 identificados con la categoría *loose*. En esta ocasión se utilizan dos criterios de identificación en los *b-jets* para que la selección sea menos estricta y el número de sucesos seleccionados sea mayor. En la figura 4.8 se representan los observables: número de *jets* identificados en el suceso con $p_T \geq 25$ GeV, el número de *b-jets medium* y el número de *b-jets loose*. Aunque los sucesos de señal son pocos, se encuentran en la región con un número alto de *jets*, donde además, la cantidad de sucesos del fondo de *ZZ* es menor. También comprobamos que la contribución de los leptones *non-prompt* (MC_fakes en la figura 4.8) es muy inferior en la región de 4ℓ comparada con la del canal $2\ell ss$.

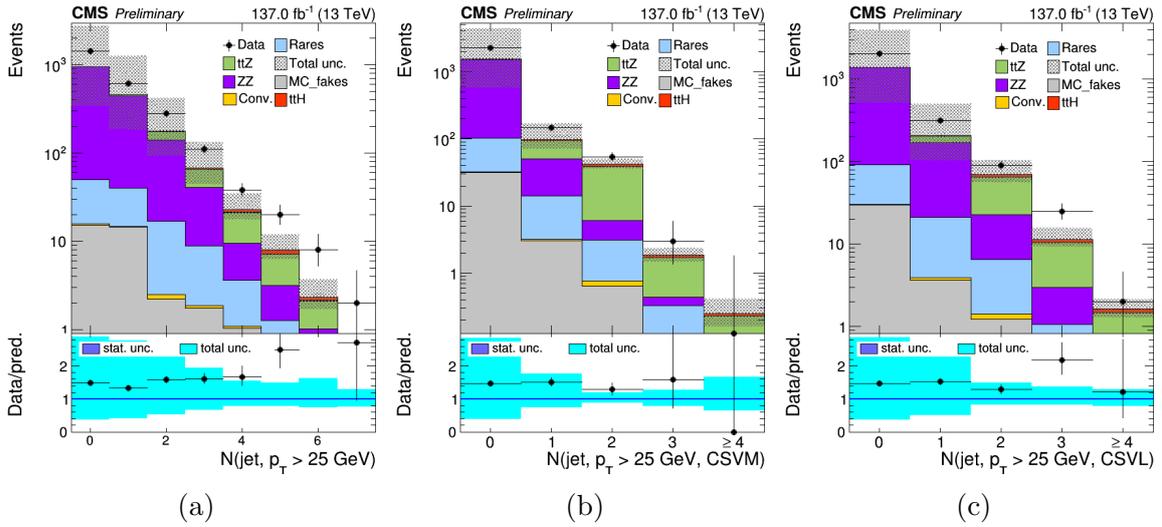


Figura 4.8: a) Distribución del número de *jets* identificados con $p_T \geq 25$ GeV. b) Distribución del número de *b-jets* identificados con $p_T \geq 25$ GeV y pertenecientes a la categoría *medium*. c) Distribución del número de *b-jets* identificados con $p_T \geq 25$ GeV y pertenecientes a la categoría *loose*.

En todas las gráficas se ha utilizado una escala logarítmica en el eje de ordenadas y como criterios de selección se han aplicado los definidos en la tabla 4.3 pero sin incluir las restricciones en el número de *jets* y *b-jets*.

La tabla 4.3 resume los criterios de selección exigidos.

Selección	4ℓ
Desintegración de $t\bar{t}H$ objetivo	$t \rightarrow b + qq'$ $t \rightarrow b + qq'$ $H \rightarrow ZZ \rightarrow 4\ell$
Trigger	Uno, dos o tres leptones
p_T leptones	$> 25/15/15/10$ GeV
Jets	≥ 3 jets
b-tag	≥ 1 medium b-tag o ≥ 2 loose b-tag
Masa invariante de los leptones	$m_{\ell\ell} > 12$ GeV

Tabla 4.3: Selección de sucesos en la categoría 4ℓ

4.5. Comparación de generadores de Montecarlo para el proceso $t\bar{t}$

Uno de los procesos que más contribuyen en la región de señal del canal $2\ell ss$ es $t\bar{t}$. Este proceso aparece en nuestro análisis como una contribución al fondo reducible, en concreto, fondos debidos a leptones *non-prompt*. En los estudios realizados por CMS y ATLAS sobre el proceso $t\bar{t}H$ estiman este tipo de fondos con técnicas basadas en datos. En este trabajo se plantea la posibilidad de utilizar en su lugar simulaciones de Montecarlo. En este caso, se han utilizado dos simulaciones del proceso $t\bar{t}$ a LO y NLO. Como generador a LO se ha escogido *Madgraph* y a NLO *Powheg*. Uno de los primeros pasos es comparar el desempeño de estos generadores entre ellos para comprobar si los resultados son compatibles. La razón es que los métodos basados en datos siempre se han optimizado utilizando muestras a LO. Al simular este fondo mediante simulaciones de Montecarlo es necesario utilizar la muestra más precisa posible y, por ello, se busca utilizar la simulación a NLO.

Los criterios de selección aplicados son los mismos que los utilizados en la categoría $2\ell ss$ (tabla 4.2), ya que el objetivo es comparar la estimación de sucesos de leptones *non-prompt* que produce cada generador. Los criterios en el número de *jets* no se han aplicado para poder comparar también esos observables.

Las figuras 4.9, 4.10, 4.11 y 4.12 representan los observables de número de *jets* en el suceso, número de *b-jets*, p_T del *jet* más energético y p_T del leptón más energético respectivamente. A diferencia del resto de gráficas presentadas en este trabajo, las de esta sección no se han construido apilando los sucesos de cada proceso. En su lugar, se representan superpuestos para poder comparar sus contribuciones.

La tabla 4.4 recoge el número de sucesos totales seleccionados de $t\bar{t}$ para las dos muestras estudiadas. Los sucesos simulados por cada generador son compatibles entre ellos para cada año dentro de los márgenes de incertidumbre. Para el resto del análisis realizado en este trabajo se utilizará las muestras simuladas a NLO dada la

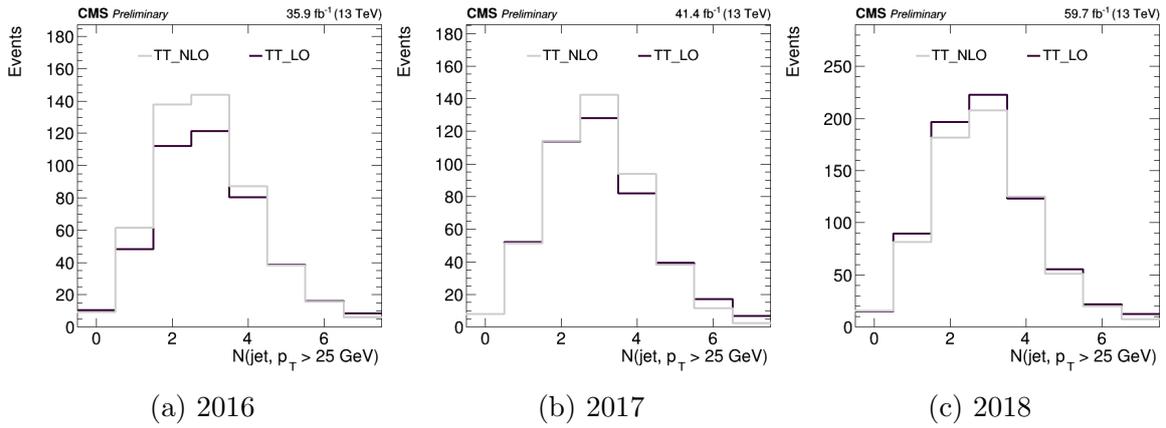


Figura 4.9: Distribución del número de *jets* para las muestras de Montecarlo NLO (*Powheg*) y LO (*Madgraph*) de $t\bar{t}$.

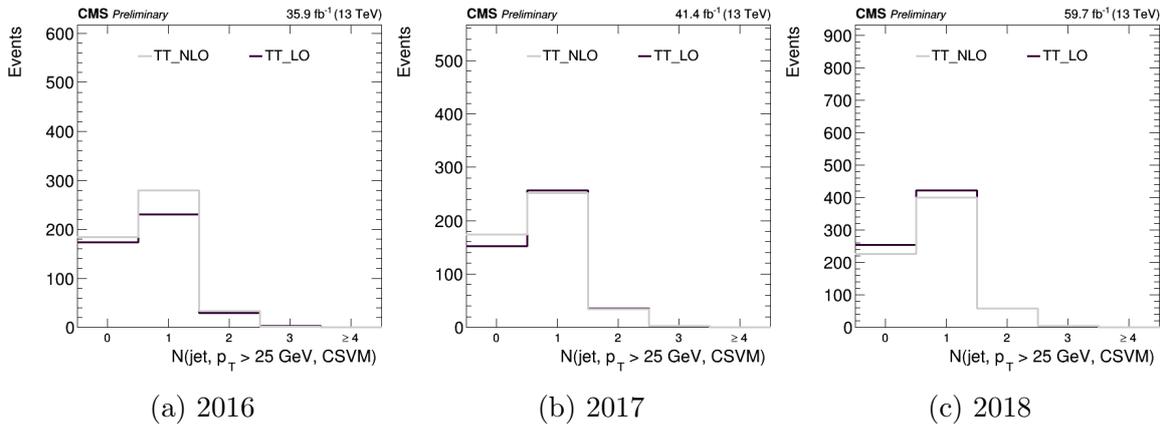


Figura 4.10: Distribución del número de *b-jets* en la categoría *medium* para las muestras de Montecarlo NLO (*Powheg*) y LO (*Madgraph*) de $t\bar{t}$.

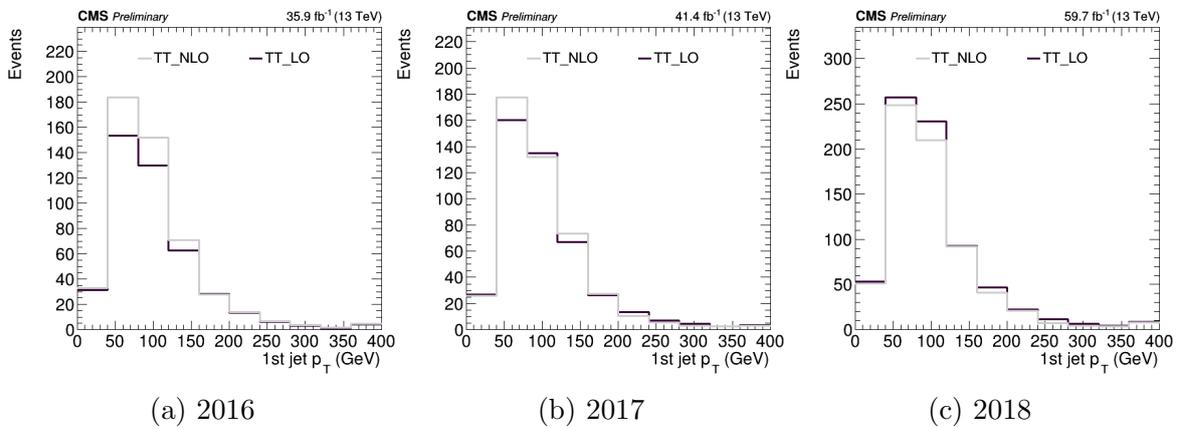


Figura 4.11: Distribución del momento transversal del *jet* más energético del suceso para las muestras de Montecarlo NLO (*Powheg*) y LO (*Madgraph*) de $t\bar{t}$.

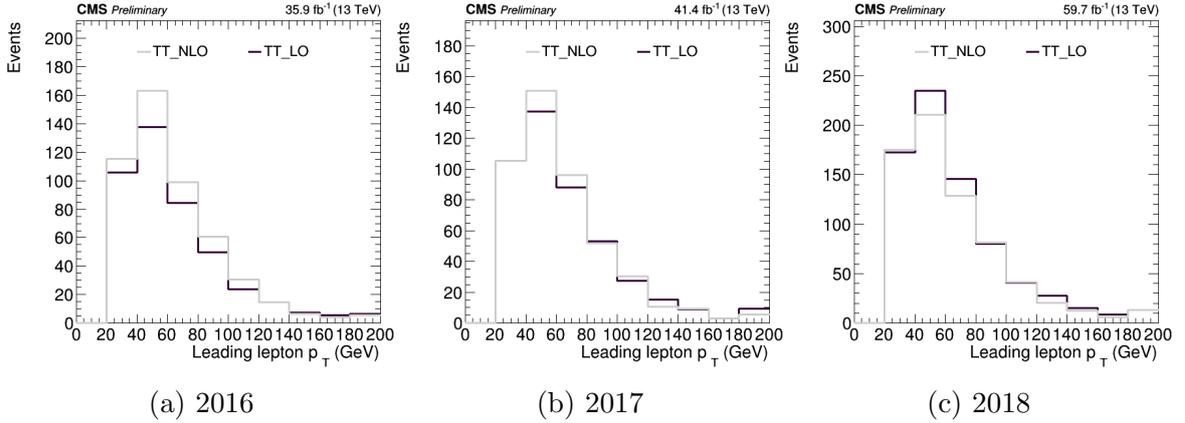


Figura 4.12: Distribución del momento transverso del leptón más energético del suceso para las muestras de Montecarlo NLO (*Powheg*) y LO (*Madgraph*) de $t\bar{t}$.

compatibilidad de los resultados comprobada entre los generadores.

Orden	2016	2017	2018
LO	435.91 ± 116.85	447.57 ± 120.02	738.16 ± 197.93
NLO	499.83 ± 133.96	462.36 ± 124.42	689.57 ± 184.91

Tabla 4.4: Sucesos totales seleccionados de $t\bar{t}$ para la muestra NLO (*Powheg*) y LO (*Madgraph*). Las incertidumbres que se incluyen en la tabla son la combinación de la incertidumbre estadística y sistemática.

4.6. Fuentes de incertidumbre

En esta sección vamos a explicar las diferentes fuentes de incertidumbre consideradas en el análisis. Es importante cuantificar las incertidumbres presentes en cualquier experimento en física. Al realizar una medida nunca se determina su valor exacto, en su lugar se proporciona un rango de valores entre los que el valor real se encuentra. En función de su naturaleza, las incertidumbres se dividen en estadísticas y sistemáticas.

4.6.1. Incertidumbres estadísticas

Las incertidumbres estadísticas surgen del hecho de que estamos realizando un experimento de conteo donde la naturaleza estocástica de los procesos considerados produce fluctuaciones aleatorias. Si N es una variable aleatoria que sigue una distribución de Poisson, entonces la incertidumbre estadística asociada es \sqrt{N} . En este caso N sería el número de sucesos observados que pasan la selección.

También, los procesos simulados mediante métodos de Montecarlo tienen una incertidumbre estadística asociada. Esta incertidumbre es menor cuanto mayor es el número de sucesos simulados. Por esta razón, se suelen simular muchos más sucesos de los que realmente se esperaría observar en las colisiones, aunque hay procesos en los que no es posible como por ejemplo Drell-Yan.

4.6.2. Incertidumbres sistemáticas

Las fuentes de incertidumbre sistemática surgen del desconocimiento del dispositivo experimental y no son originadas enteramente por efectos estocásticos. Estas pueden ser debidas a la eficiencia de los detectores, a la eficiencia de los algoritmos de identificación y reconstrucción y a las incertidumbres teóricas de la simulación de los fondos y señal.

Las incertidumbres consideradas en este análisis han sido las siguientes:

- **Eficiencia de leptones *non-prompt*:** El cálculo de la eficiencia de los sucesos de leptones *non-prompt* que pasan la selección tiene una incertidumbre. En este análisis se ha asignado una incertidumbre del 30 % (que es el orden visto en las medidas de la eficiencia) a las muestras de $t\bar{t}$ para tener en cuenta este efecto.
- **Modelado MC:** El modelado de los fondos y la señal mediante simulaciones de Montecarlo tienen una incertidumbre asociada. Esta incertidumbre proviene del hecho de que no se pueden tener en cuenta en la simulación todos los órdenes en teoría de perturbaciones sino que, normalmente, la aproximación se hace a LO o NLO. También se incluye aquí la incertidumbre asociada a las PDF (*Parton Density Functions*). Estas incertidumbres sistemáticas afectan a la sección eficaz de los procesos de Montecarlo. Para el proceso de señal $t\bar{t}H$ se tiene una incertidumbre de $\pm_{9.97}^{6.83}$ %. Y para los fondos irreducibles: $\pm_{12.18}^{13.51}$ % para $t\bar{t}W$ y $\pm_{10.18}^{17.70}$ % para $t\bar{t}Z$. Estas incertidumbres representan la suma cuadrática de la incertidumbre debida a las PDF y a los órdenes superiores en teoría de perturbaciones no tenidos en cuenta.
- **Luminosidad:** La luminosidad supone otra fuente de incertidumbre en las medidas. Es la incertidumbre asociada a la medida de la luminosidad integrada proporcionada por el LHC y recogida por CMS. Afecta a la normalización de los procesos y es diferente para cada año. Se divide a su vez en incertidumbres correlacionadas y no correlacionadas entre años. Las incertidumbres totales (sumando efectos correlacionados y no correlacionados) serían de un 2.5 %, 2.3 % y 2.5 % para 2016, 2017 y 2018 respectivamente.

- **Eficiencia de *Trigger*:** La eficiencia de *trigger* también tiene una incertidumbre asociada. Dentro de nuestro análisis, consideramos las incertidumbres asociadas a la eficiencia de los *triggers* basados en la presencia de pares de leptones, donde para cada año es diferente. Aproximadamente esta incertidumbre es del orden del 5 %.
- **Identificación y aislamiento de leptones:** La eficiencia en los algoritmos de identificación de leptones tienen asociadas una serie de incertidumbres sistemáticas. El origen es que los algoritmos no son exactos, y objetos que no son leptones pueden ser identificados como tal y viceversa. Estas incertidumbres dependen de los años, del p_T y η del leptón y de la categoría a la que pertenezcan: *loose* o *tight*.
- **Energía de los *jets*:** Las incertidumbres asociadas a la escala de energía de los *jets* se calculan variando las propiedades de los *jets* reconstruidos por su incertidumbre hacia arriba y hacia abajo y recalculando las cantidades cinemáticas relevantes. Hay dos tipos de incertidumbres, las asociadas a la escala de energía, y las asociadas a la resolución. A su vez, estas incertidumbres se dividen en varias componentes que tienen en cuenta diferentes aspectos de la escala de energía de los *jets*.
- **Eficiencia de *b-tagging*:** Los algoritmos empleados en el etiquetado de *quarks* b no identifican todos los *jets* procedentes de un *quark* b con una eficiencia del 100 %. A veces los confunden con *jets* ligeros y viceversa, *jets* ligeros pueden etiquetarse erróneamente como *b-jets*. Esta eficiencia se mide en simulaciones de Montecarlo, generalmente en muestras de $t\bar{t} + jets$ y en procesos con un alto número de *jets*. Al igual que las incertidumbres asociadas a la escala de energía de los *jets*, las incertidumbres en la eficiencia de *b-tagging* se dividen también en varias componentes. Tanto la eficiencia de *b-tagging* como la probabilidad de etiquetar incorrectamente un *jet* ligero como b , tienen una incertidumbre asociada.
- ***Pileup*:** Las condiciones de *pileup* pueden influir en las medidas de aislamiento de los leptones y la energía de los *jets*, y esto afectar en última instancia al número de sucesos de señal y fondo obtenidos de las simulaciones de Montecarlo. Se calcula variando el número de colisiones inelásticas por cada cruce de paquetes de protones en las simulaciones. De esta forma, se estima el efecto que tiene sobre la medida de diferentes procesos las condiciones de *pile-up*. El efecto es en todos los casos pequeño y del orden del 1 %.

Capítulo 5

Resultados Experimentales

En este capítulo se recogen los resultados de la búsqueda de la producción de $t\bar{t}H$ en los canales $2\ell ss$ y 4ℓ . Los criterios de selección aplicados para cada canal se han definido en el capítulo anterior y se resumen en las tablas 4.2 y 4.3. El capítulo se divide en dos secciones con objetivos diferentes. La primera sección, tiene como objetivo la determinación experimental de la fuerza de señal μ en el canal $2\ell ss$. La segunda sección por otro lado, busca estudiar la posibilidad de observación del pico de resonancia del bosón de Higgs en el canal 4ℓ . Esta medida ha sido realizada ya en otros modos de producción del bosón de Higgs diferentes al $t\bar{t}H$. Sin embargo, es interesante realizarla también aquí, sobre todo para comprobar la compatibilidad de los resultados.

5.1. Canal $2\ell ss$

5.1.1. Región de señal

Esta sección recoge las principales distribuciones de los observables físicos característicos en la región de señal definida según los criterios de selección recogidos en la tabla 4.2. La figura 5.1 representa el momento transversal del leptón más energético del suceso (*leading lepton*) para los años 2016, 2017 y 2018. En el eje de ordenadas se representa el número de sucesos seleccionados y en el eje de abscisas el momento transversal del leptón. La mayor contribución de los fondos irreducibles procede del proceso $t\bar{t}W$. Por otro lado, la mayor contribución de los fondos reducibles corresponde al proceso $t\bar{t}$ habiéndose reducido en gran medida el fondo procedente de los *flips* y *conversions*. El p_T del otro leptón (*sub-leading*) se representa en la figura 5.2. Los leptones presentes en el estado final pueden ser o bien dos electrones (ee), dos muones ($\mu\mu$) o un electrón y un muon ($e\mu$). La mayoría de los leptones seleccionados provienen

de desintegraciones débiles y, por lo tanto, la probabilidad de tener un electrón es la misma que la de tener un muon. Como consecuencia, la proporción esperada sería que el número de sucesos de $e\mu$ fuera igual al de ee y $\mu\mu$ juntos. Sin embargo, en la figura 5.3 se observa una situación diferente. La proporción en la que se observan se puede explicar teniendo en cuenta que los criterios de identificación de los electrones son mucho menos eficientes que los de los muones. Como consecuencia, el número de sucesos observados es menor. Además, sobre los electrones se aplican dos criterios de selección adicionales, luego es de esperar que el número de sucesos seleccionados con 2 electrones en el estado final sea menor. Esto se observa también en el número de sucesos de $e\mu$ que es similar al de $\mu\mu$. Las coordenadas angulares η y ϕ para el leptón más energético se representan en las figuras 5.4 y 5.5. Son distribuciones simétricas como cabría esperar por la cinemática del proceso.

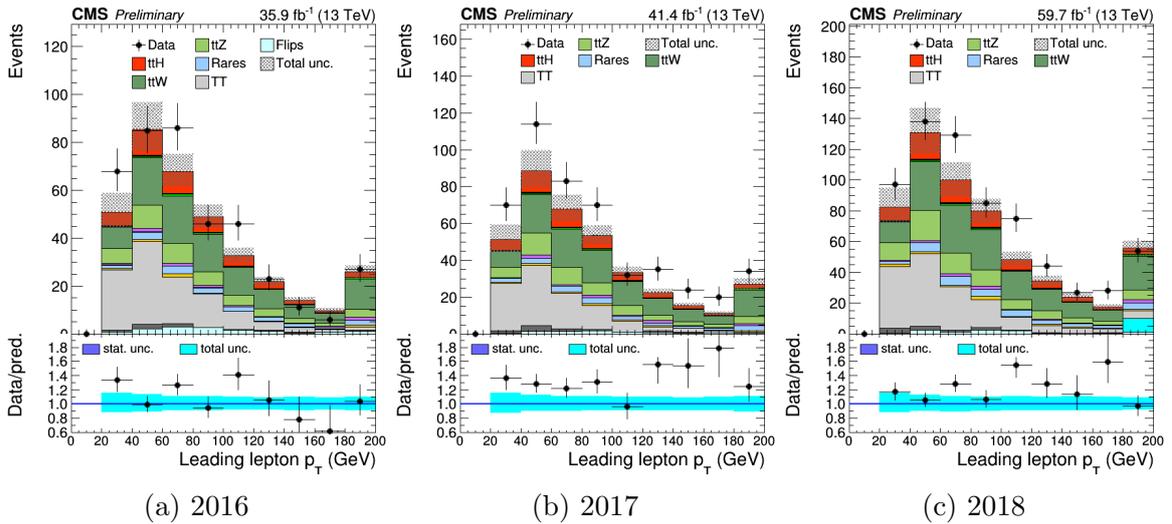


Figura 5.1: Distribución del momento transverso del leptón más energético del suceso en la región de señal.

Otro observable importante es el momento transverso faltante E_T^{Miss} . En el canal $2\ell ss$ como mínimo habrá dos neutrinos en el estado final que no podrán ser detectados. Este defecto de energía lo observaremos en esta magnitud (ver figura 5.6).

El número total de sucesos seleccionados por año en la región de señal se muestra en la figura 5.7. De entre todos ellos, 2016 es el que mejor relación datos/predicción presenta. El detector CMS ha sufrido modificaciones y envejecimiento de sus subdetectores durante el periodo 2016-2018 que se analiza en este trabajo. La continua exposición de sus piezas a la radiación emitida en la colisiones produce un desgaste que afecta a las mediciones. En concreto, en 2017 fallaron algunos módulos del detector de píxeles que fueron reparados en 2018. También las condiciones de *pile-up* son diferentes para los tres años. En 2016 el número promedio de interacciones en cada

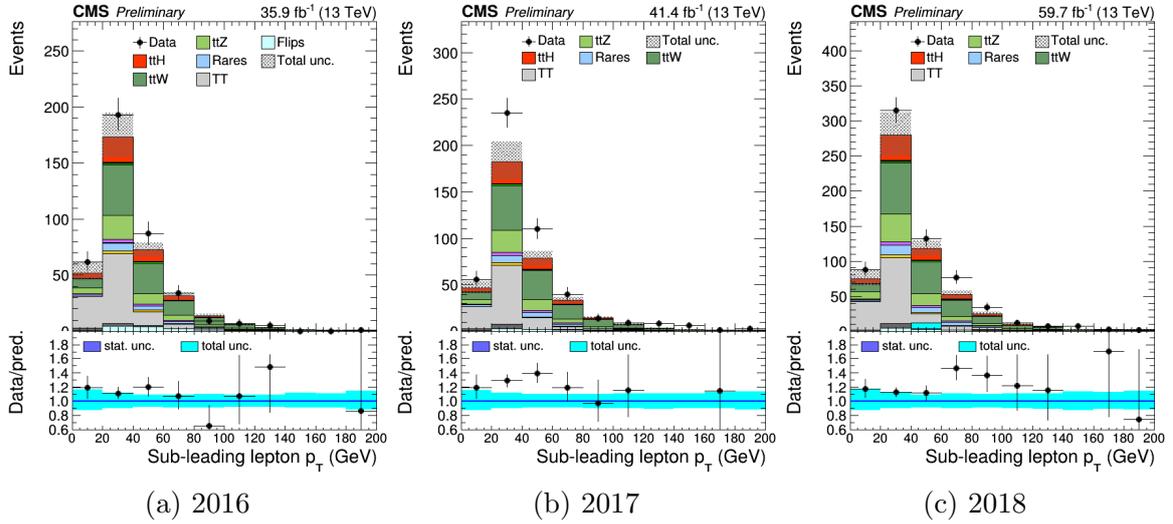


Figura 5.2: Distribución del momento transverso del segundo leptón más energético del suceso en la región de señal.

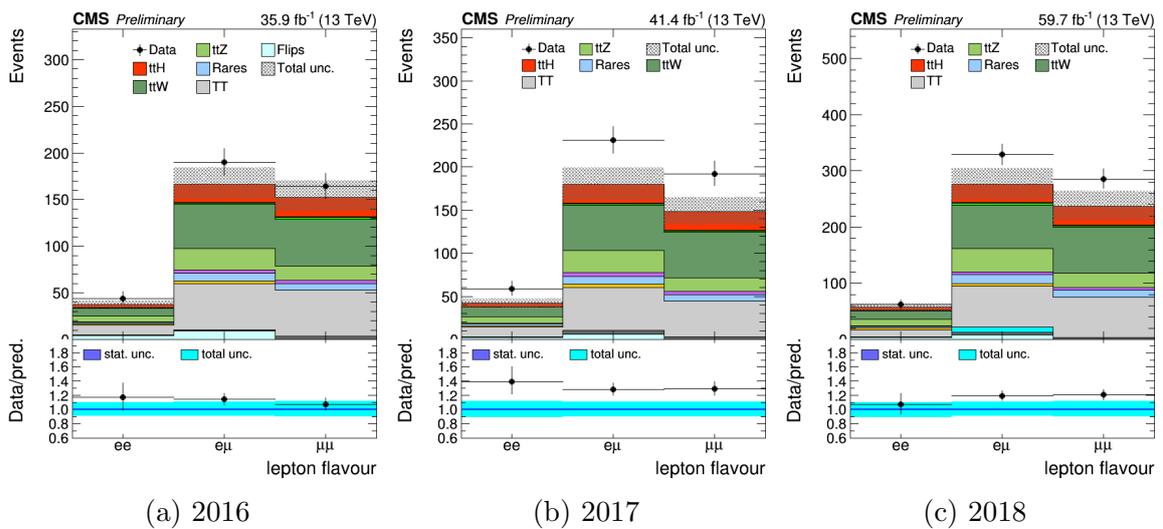


Figura 5.3: Sabores de los leptones *tight* en la región de señal.

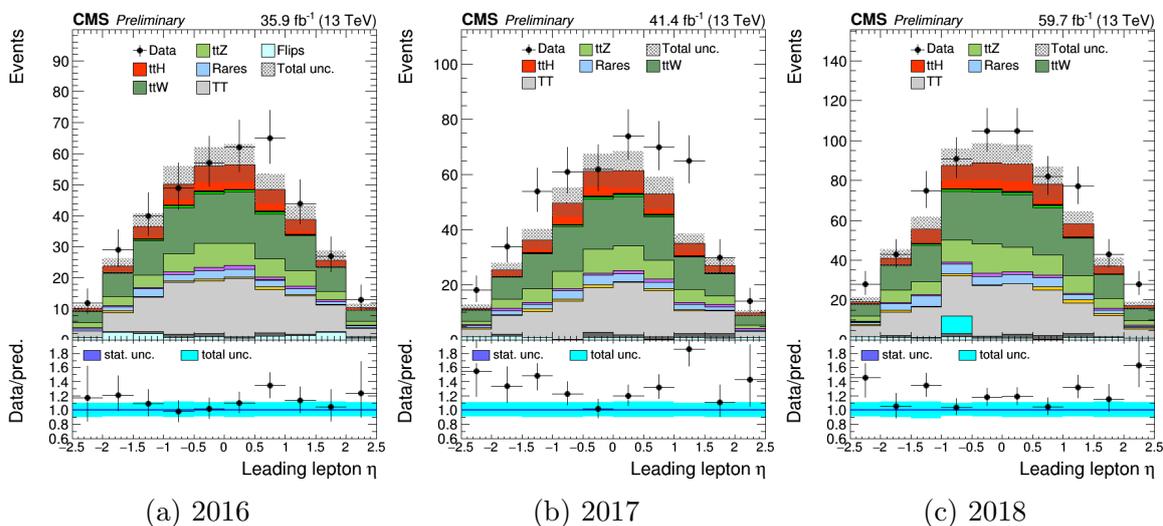


Figura 5.4: Distribución de la pseudorapidez η del leptón más energético del suceso en la región de señal.

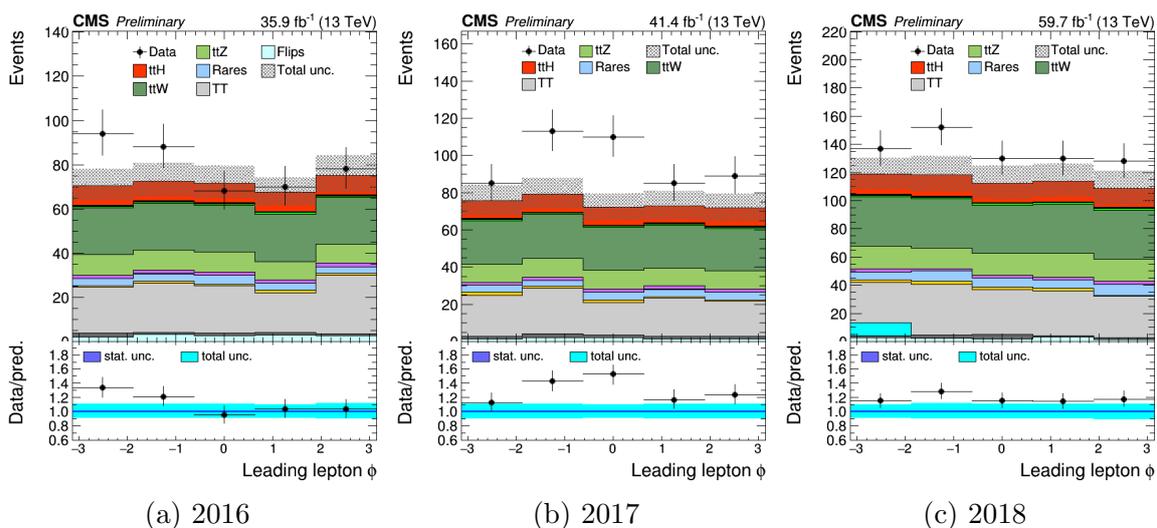


Figura 5.5: Distribución del ángulo ϕ del leptón más energético del suceso en la región de señal.

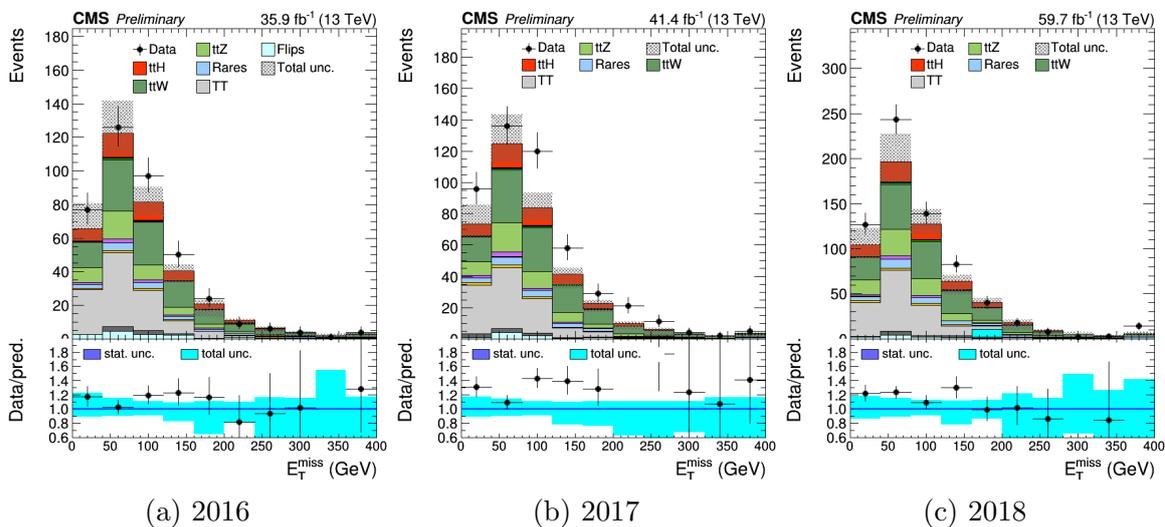


Figura 5.6: Distribución del momento transverso faltante en la región de señal.

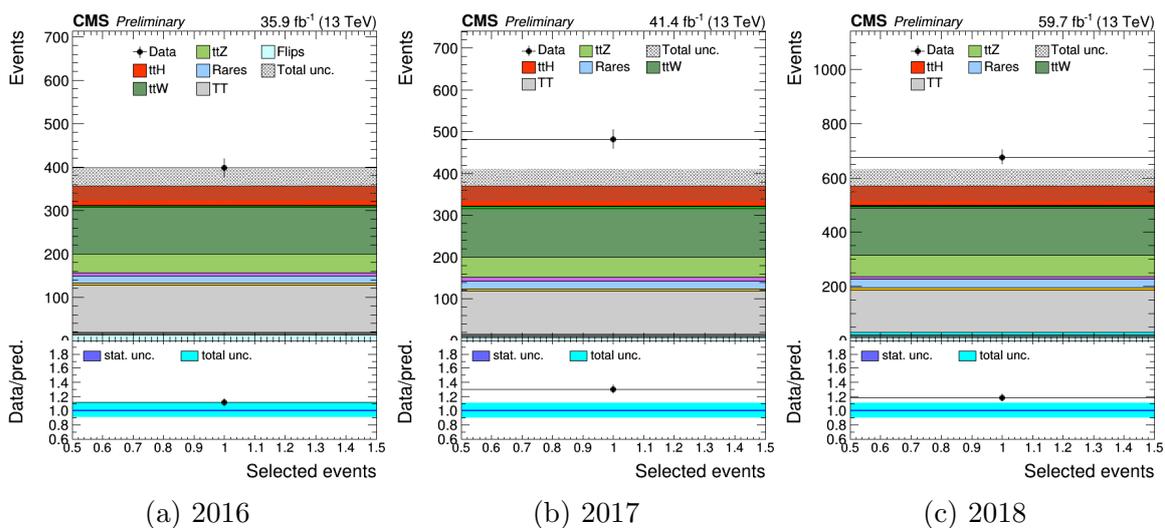


Figura 5.7: Número de sucesos totales seleccionados en la región de señal.

cruce de los paquetes de protones era de 27, mientras que en 2017 y 2018 era de 38 y 37 respectivamente (estos datos se pueden consultar en [15]). Estas podrían ser las razones que explicarían el peor acuerdo datos/predicción en 2017 comparado con 2016 y 2018.

La tabla 5.1 recoge el número de sucesos seleccionados con sus incertidumbres en el estudio del canal $2\ell ss$ para cada uno de los procesos mencionados en la presente memoria, así como los datos para los tres años considerados.

Proceso	2016	2017	2018
ttH	44.11 ± 4.68	47.54 ± 5.31	71.26 ± 8.15
tHq	1.24 ± 0.12	1.29 ± 0.15	1.96 ± 0.23
tHW	1.09 ± 0.12	1.07 ± 0.10	1.81 ± 0.19
ttWW	3.21 ± 0.37	3.61 ± 0.44	5.58 ± 0.67
ttW	106.36 ± 15.08	116.12 ± 16.85	173.76 ± 25.30
ttZ	44.25 ± 5.83	49.21 ± 6.75	79.27 ± 10.91
WZ	7.17 ± 3.31	8.19 ± 3.79	9.53 ± 4.81
ZZ	0.28 ± 0.47	0.31 ± 0.49	0.39 ± 0.63
Rares	16.15 ± 7.10	18.47 ± 8.31	32.09 ± 14.64
Conv.	4.43 ± 1.56	5.01 ± 1.69	7.66 ± 2.83
tt	109.04 ± 29.38	103.54 ± 28.33	157.35 ± 42.50
<i>Single t</i>	5.49 ± 1.35	7.70 ± 2.07	10.16 ± 2.51
<i>Flips</i>	13.49 ± 1.14	8.15 ± 0.98	10.82 ± 1.38
Fondo	312.22 ± 35.87	322.67 ± 37.19	499.31 ± 57.47
Datos	398	482	677

Tabla 5.1: Sucesos seleccionados de los procesos considerados en el análisis para los años 2016, 2017 y 2018. Las incertidumbres que se incluyen en la tabla son la combinación de la incertidumbre estadística y sistemática.

5.1.2. Fuerza de señal

Para estimar la fuerza de señal μ se ha utilizado un ajuste por el método de máxima verosimilitud. El ajuste se suele realizar sobre una serie de observables que discriminen correctamente la señal del fondo. En este trabajo se ha aplicado sobre las distribuciones de sucesos de la tabla 5.1.

Supongamos que se quiere aplicar este procedimiento a una muestra de datos representada en un histograma con un número de categorías N . Sea x_i el número de sucesos observados en la categoría i , se denomina función de verosimilitud asociada a las observaciones (x_1, \dots, x_N) , a la probabilidad conjunta de la muestra, y se denota por $L(x_1, \dots, x_N; \mu, \vec{\theta})$. Donde $\vec{\theta}$ son los denominados parámetros de confusión, utilizados para modelizar las fuentes de incertidumbre sistemática, y μ la fuerza de señal.

En el caso de tener un experimento de conteo, la función de distribución de probabilidad de cada categoría tendrá la forma de una Poissoniana:

$$P(x; \nu) = \frac{e^{-\nu} \nu^x}{x!} \quad (5.1)$$

Donde:

- x : es el número de sucesos observados.
- ν : es el número de sucesos esperados.

Si asumimos que cada categoría es independiente, podemos construir la función de verosimilitud como un producto de poissonianas $P(x_i; \nu_i(\vec{\theta}))$ para cada categoría [44]. El número de sucesos esperados en cada categoría ν_i depende de los parámetros $\vec{\theta}$. De esta forma la ecuación para la función de verosimilitud se puede escribir como:

$$L(x_1, \dots, x_N; \mu, \vec{\theta}) = \prod_{i=1}^N P(x_i; \nu_i(\vec{\theta})) \prod_{j=1}^m p_j(\theta_j) \quad (5.2)$$

Donde p_j son las distribuciones esperadas que modelan los parámetros de confusión. En este caso se ha utilizado la distribución normal logarítmica [45] para las incertidumbres sistemáticas que afectan a la normalización de los procesos, y la distribución gaussiana para las que afectan a la forma de las distribuciones.

Si el histograma únicamente contiene una categoría, la expresión 5.2 se simplifica:

$$L(x; \mu, \vec{\theta}) = P(x; \nu_{\vec{\theta}}) \prod_{j=1}^m p_j(\theta_j) = \frac{e^{-\nu_{\vec{\theta}}} \nu_{\vec{\theta}}^x}{x!} \prod_{j=1}^m p_j(\theta_j) = \frac{e^{-(\mu S_{\vec{\theta}} + B_{\vec{\theta}})} (\mu S_{\vec{\theta}} + B_{\vec{\theta}})^x}{x!} \prod_{j=1}^m p_j(\theta_j) \quad (5.3)$$

Donde $S_{\vec{\theta}}$ el número de sucesos de señal esperados y $B_{\vec{\theta}}$ el número de sucesos de fondo esperados obtenidos a través de las simulaciones de Montecarlo. La señal y el fondo se pueden dejar como parámetros libres que dependen de las fuentes de incertidumbre $\vec{\theta}$.

El objetivo es encontrar los parámetros del modelo μ y $\vec{\theta}$ que maximizan la función de verosimilitud para los datos observados. La forma usual de resolver este problema es resolver la ecuación de verosimilitud:

$$\hat{\mu}, \hat{\vec{\theta}} = \underset{\mu, \vec{\theta}}{\operatorname{argmax}} L(x_1, \dots, x_N; \mu, \vec{\theta}) \quad (5.4)$$

La resolución analítica de 5.4 solamente se puede llevar a cabo en los casos más sencillos, normalmente se resolverá de forma numérica. En este trabajo, la búsqueda

de los parámetros que maximicen la función de verosimilitud se hace numéricamente a través del *framework* CMSSW que utiliza el paquete de minimización MINUIT [46]. Ajustando a las distribuciones 5.7 obtenemos los siguientes valores de μ :

$$\mu_{2016}^{obs} = 1.9 \pm_{1.0}^{1.0} \quad \mu_{2017}^{obs} = 3.4 \pm_{1.0}^{1.1} \quad \mu_{2018}^{obs} = 2.5 \pm_{1.0}^{1.0} \quad (5.5)$$

Y haciendo el ajuste para los tres años a la vez:

$$\mu^{obs} = 2.6 \pm_{0.9}^{0.9} \quad (5.6)$$

De estos resultados μ_{2016}^{obs} , μ_{2018}^{obs} y μ^{obs} son compatibles con lo predicho por el Modelo Estándar ($\mu = 1$) en menos de dos desviaciones estándar. Con respecto a este resultado, el exceso observado en los datos comparado con lo predicho por las simulaciones puede explicarse teniendo en cuenta que se ha observado que el proceso $t\bar{t}W$ está mal modelado (este resultado se puede encontrar en [47]). En concreto, el número de sucesos simulados de $t\bar{t}W$ debería ser mayor para ajustarse a los datos experimentales. Esto afecta a la medida realizada en $t\bar{t}H$ observando una fuerza de señal superior a la esperada.

Las incertidumbres en la fuerza de señal 5.6 incluyen todas las contribuciones de las incertidumbres estadísticas y sistemáticas presentadas en el capítulo anterior. Se han considerado en total 108 fuentes de incertidumbre sistemática en el análisis. La figura 5.8 muestra, en la columna de la derecha, la contribución de cada fuente de incertidumbre por separado a la incertidumbre total de μ . La columna de la izquierda contiene los nombres de cada fuente de incertidumbre. Por último, en el centro se representa el valor del parámetro de confusión que mejor se ajusta a los datos. Si el valor es cercano a cero quiere decir que el valor esperado de la fuente de incertidumbre coincide con el valor que proporciona el ajuste de máxima verosimilitud a los datos observados.

La mayor contribución a la incertidumbre total de μ proviene de la incertidumbre en la eficiencia de leptones *non-prompt*. Esta incertidumbre se había fijado a un 30 % para el proceso $t\bar{t}$. Otras contribuciones grandes a la incertidumbre de μ provienen de fuentes teóricas como son el modelado mediante simulaciones de Montecarlo de $t\bar{t}W$ y $t\bar{t}H$. Las incertidumbres en la escala de energía de los *jets* y en la eficiencia de *b-tagging* están divididas en varias componentes y suponen también una contribución importante. En vista a los resultados, para mejorar la precisión en la medida de μ es necesario reducir la incertidumbre en la eficiencia de leptones *non-prompt*. Este proceso pasaría por mejorar las técnicas utilizadas para medir esta eficiencia y también por optimizar los criterios de identificación de leptones *tight* de forma que fueran más eficientes discriminando a los leptones *non-prompt*.

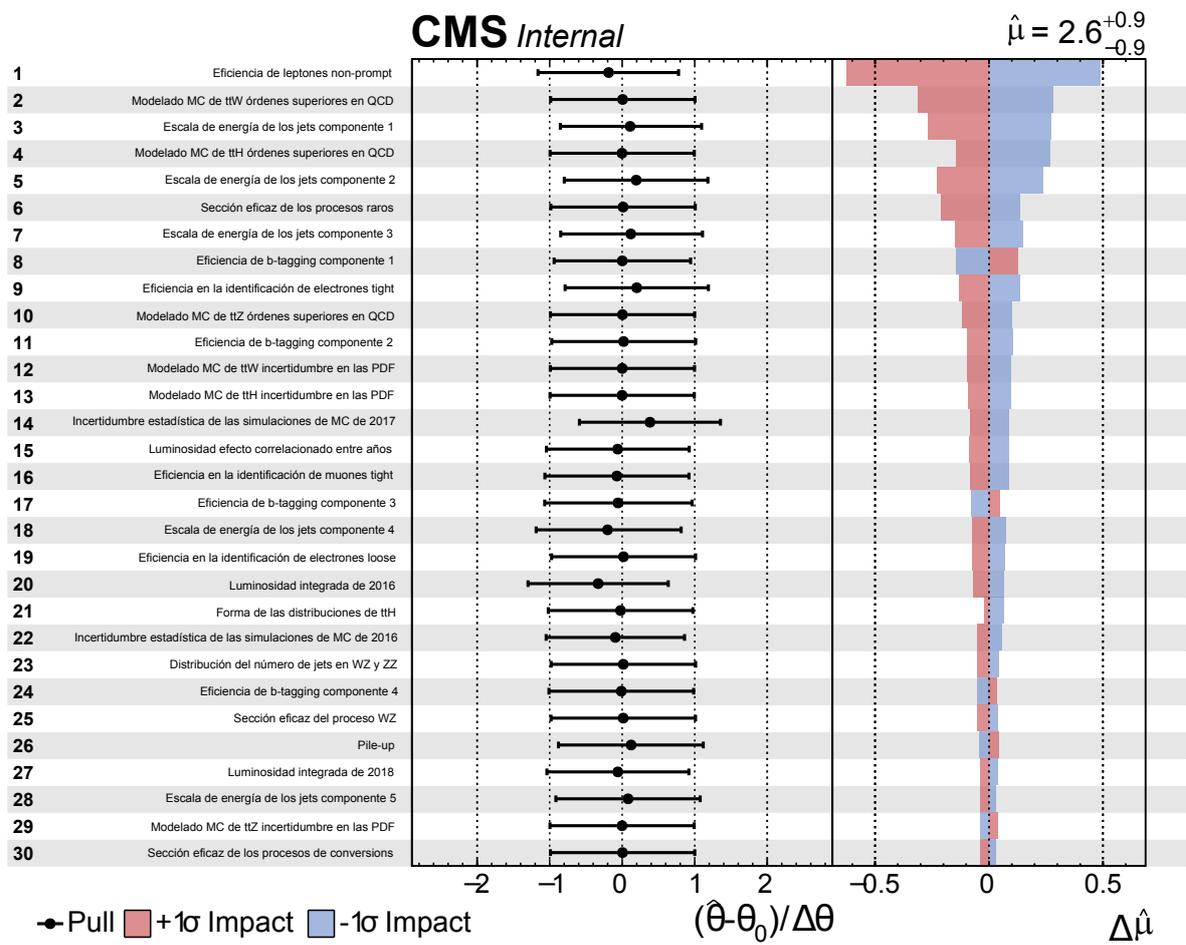


Figura 5.8: Impacto de las fuentes de incertidumbre en el cálculo de la fuerza de señal.

5.2. Medida de la masa del bosón de Higgs en el canal de $t\bar{t}H$: $H \rightarrow ZZ^* \rightarrow 4\ell$

El objetivo de esta sección es determinar la posibilidad de observación del pico de masa del bosón de Higgs en torno a 125 GeV. Para ello, hay que reconstruir la masa del bosón a partir de sus productos de desintegración, en este caso, leptones. Esto se puede hacer calculando la masa invariante de los cuatro leptones seleccionados en el suceso. Si estos leptones proceden de un bosón de Higgs, su masa invariante se corresponderá con $m_H = 125$ GeV.

La figura 5.9 representa la masa invariante de los 4 leptones seleccionados según los criterios de selección definidos en la sección 4.4.2. Los procesos de fondo que más contribuyen son $t\bar{t}Z$ y ZZ que son fondos irreducibles. La tabla 5.2 recoge los sucesos seleccionados en el canal 4ℓ para cada proceso. A diferencia de la región $2\ell ss$, la contribución de los leptones *non-prompt* es menor. La selección tan estricta de 4 leptones es la causa de que el número de sucesos totales seleccionados sea muy bajo. Por ello, como se observa en la figura 5.9 derecha, no existe suficientes datos que permitan afirmar la observación del proceso $t\bar{t}H$ en el canal seleccionado. De hecho, aunque la simulación de Montecarlo permite observar el pico de masa, no hay ningún suceso observado del mismo.

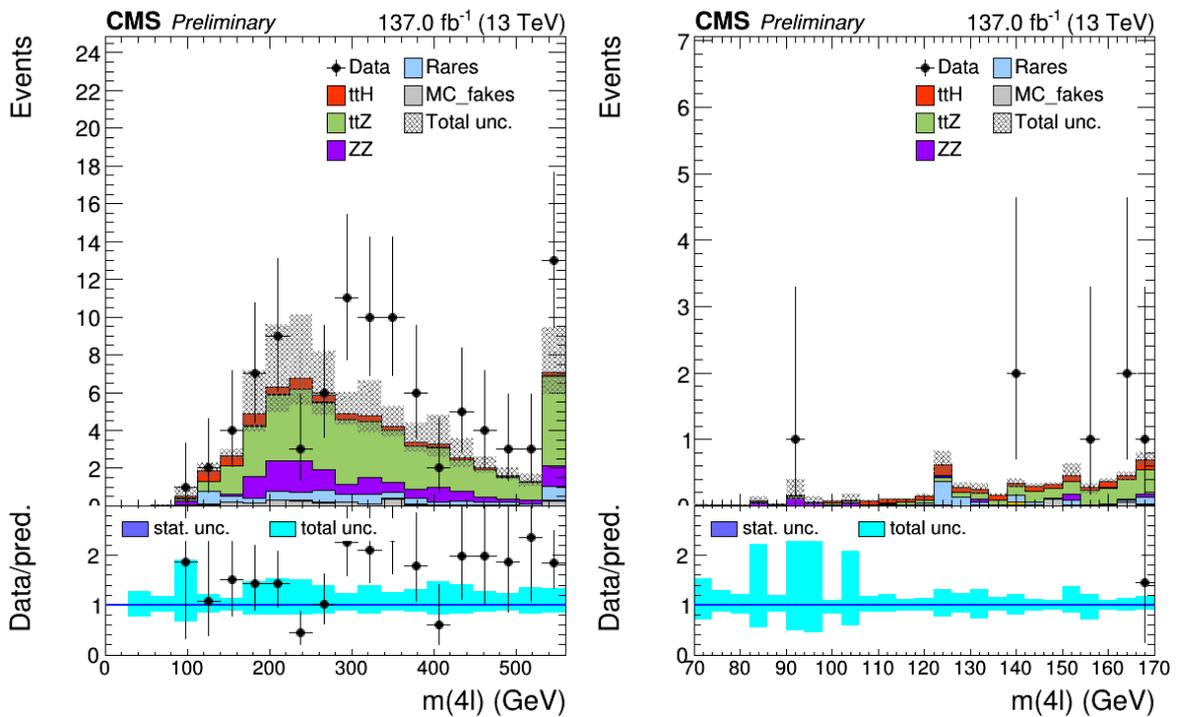


Figura 5.9: Masa invariante de los 4 leptones seleccionados. A la derecha se ha hecho un discretizado más fino para poder observar el pico de masa del bosón de Higgs en torno a 125 GeV.

Proceso	Sucesos
ttH	4.80 ± 0.55
tHq	0.04 ± 0.01
tHW	0.10 ± 0.01
ttWW	0.19 ± 0.03
ttW	0.03 ± 0.01
ttZ	39.14 ± 4.93
ZZ	11.22 ± 17.18
Rares	5.48 ± 2.39
Conv.	0.20 ± 0.08
MC_fakes	2.19 ± 0.34
Fondo	58.57 ± 18.25
Datos	99

Tabla 5.2: Sucesos totales seleccionados en el canal 4ℓ . Las incertidumbres que se incluyen en la tabla son la combinación de la incertidumbre estadística y sistemática.

Es interesante ver también que existe una contribución grande de los procesos raros en torno a 125 GeV. Esta contribución proviene en concreto del proceso de producción de bosones de Higgs mediante fusión de gluones ggH (figura 5.10). La sección eficaz de este proceso ($48.4 \pm_{4.4}^{3.6} pb$ [14]) es mucho mayor que la de $t\bar{t}H$ ($0.50 \pm_{0.05}^{0.04} pb$ [14]), sin embargo, la contribución en esta región es del mismo orden. Esto ocurre porque se está aplicando un criterio de selección que exige que en el estado final haya al menos 3 *jets* y algún *b-jet* identificado (criterio en la tabla 4.3). El proceso ggH en principio no tiene *jets* en su estado final ya que se están seleccionando sucesos con 4ℓ , es decir, el canal $ggH \rightarrow ZZ^* \rightarrow 4\ell$. Pero durante el proceso de *hard-scattering* las partículas con carga de color pueden radiar gluones o *quarks* que producirán los *jets*. Esto explica la contribución de ggH en esta región que es pequeña porque según se aumenta el número de *jets* en el suceso este se vuelve menos probable.

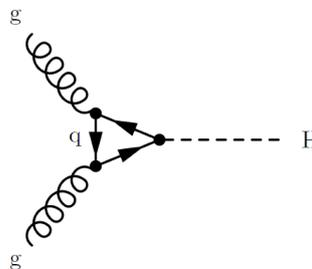


Figura 5.10: Diagrama de Feynman de la producción de un bosón de Higgs mediante fusión de dos gluones.

Una vez visto el resultado, se pueden plantear una serie de mejoras de cara a un estudio futuro posiblemente una vez haya comenzado el *Run-3*. Un posible punto de

mejora sería relajar los criterios de identificación de leptones. La definición de leptones utilizada en este trabajo es muy estricta y está muy optimizada para el canal $2\ell ss$ donde los fondos de leptones *non-prompt* son una importante contribución a los sucesos totales. En la región de 4ℓ al haber menos sucesos de leptones *non-prompt* sería posible emplear criterios menos restrictivos y aumentar así los sucesos observados. También es esperable que en el futuro, cuando la luminosidad integrada sea mayor, el proceso comience a ser observable.

Capítulo 6

Conclusiones

En este trabajo se han presentado los resultados de la búsqueda de la producción asociada de un bosón de Higgs con un par de *quarks top* en canales multileptónicos. Los estados finales seleccionados se han agrupado en dos categorías en función del número de leptones presentes en él. La categoría $2\ell ss$ para pares de leptones en el estado final con la misma carga eléctrica, y la categoría 4ℓ para estados finales con 4 leptones. Los datos analizados se corresponden con los tomados por CMS en los años 2016, 2017 y 2018, registrando una luminosidad integrada total de 137 fb^{-1} a una energía en centro de masas de 13 TeV en colisiones protón-protón.

Los resultados obtenidos en el canal $2\ell ss$ muestran un buen acuerdo datos/predicción en el año 2016 y 2018 comparado con 2017. Esta diferencia puede ser atribuible a los daños sufridos por el detector de píxeles y al mayor número de sucesos de alto *pile-up* que pueden afectar a las medidas. La determinación de la fuerza de señal $\mu = \frac{\sigma_{exp}}{\sigma_{SM}}$ se ha realizado utilizando un ajuste de máxima verosimilitud. El ajuste se ha realizado sobre las distribuciones 5.7 de sucesos totales, obteniéndose:

$$\mu^{obs} = 2.6 \pm_{0.9}^{0.9} \quad (6.1)$$

El valor observado de la fuerza de señal es compatible con lo predicho por el Modelo Estándar en menos de dos desviaciones estándar. La mayor contribución a la incertidumbre en la fuerza de señal proviene de la incertidumbre asociada a la eficiencia de los leptones *non-prompt* que pasan la selección. La estimación de estos sucesos ha sido simulada mediante métodos de Montecarlo utilizando generadores a LO y NLO que se han comparado comprobando que el número de sucesos que simulan es compatible dentro de los márgenes de incertidumbre.

Por otro lado, en la región de 4ℓ se ha comprobado que el volumen de datos recogidos hasta ahora por CMS no permite la observación del proceso $t\bar{t}H$ en el canal 4ℓ debido a

su bajísima sección eficaz. Por este motivo, no se ha podido medir la resonancia de masa del bosón de Higgs. Si bien se puede incrementar la eficiencia de la selección relajando los criterios exigidos a los leptones, será necesario acumular más datos para hacer esta medida factible. El *Run-3* del LHC cuyo comienzo está previsto para 2021-2022 prevé acumular un volumen de datos 3 veces superior al disponible actualmente.

Bibliografía

- [1] César Tomé. Del modelo estándar. *Cultura Científica*, Jun 2014. <https://culturacientifica.com/2014/06/17/del-modelo-estandar/>.
- [2] P.A.M Dirac. The Quantum Theory of the Electron. *Royal Society*, pages 610–624, 1928. <https://royalsocietypublishing.org/doi/pdf/10.1098/rspa.1928.0023>.
- [3] M. Tanabashi. Review of particle physics. *Phys. Rev. D*, 98:030001, Aug 2018. <https://link.aps.org/doi/10.1103/PhysRevD.98.030001>.
- [4] S. Chatrchyan, V. Khachatryan, A.M. Sirunyan, A. Tumasyan, W. Adam, E. Aguilo, et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B*, 716(1):30–61, Sep 2012. <http://dx.doi.org/10.1016/j.physletb.2012.08.021>.
- [5] G. Aad, T. Abajyan, et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716(1):1 – 29, 2012. <http://www.sciencedirect.com/science/article/pii/S037026931200857X>.
- [6] Glenn Elert. Beyond the Standard Model. *The Physics Hypertextbook*, 1998-2020. <https://physics.info/beyond/>.
- [7] Planck Collaboration. Planck 2015 results - I. Overview of products and scientific results. *A&A*, 594:A1, 2016. <https://doi.org/10.1051/0004-6361/201527101>.
- [8] A. M. Sirunyan, A. Tumasyan, et al. Observation of $t\bar{t}H$ Production. *Phys. Rev. Lett.*, 120:231801, Jun 2018. <https://link.aps.org/doi/10.1103/PhysRevLett.120.231801>.
- [9] CMS Collaboration. Combined measurements of the Higgs boson’s couplings at $\sqrt{s} = 13$ tev. http://cds.cern.ch/record/2308127/files/Figure_010-b.pdf.

- [10] Konrad Jende et al. El bosón de Higgs. *LHCInternationalMasterclasses*. https://atlas.physicsmasterclasses.org/es/zpath_hboson.htm.
- [11] J Quijada. Recent upgrades and results from the CMS experiment. *Journal of Physics: Conference Series*, 912:012005, Oct 2017. https://www.researchgate.net/publication/320698309_Recent_upgrades_and_results_from_the_CMS_experiment.
- [12] Lyndon Evans. The Large Hadron Collider. *The Royal Society*, Feb 2012. <https://royalsocietypublishing.org/doi/10.1098/rsta.2011.0453>.
- [13] R. Aaij, B. Adeva, and et al. Measurement of the inelastic pp cross-section at a centre-of-mass energy of 13 tev. *Journal of High Energy Physics*, 2018(6), Jun 2018. [http://dx.doi.org/10.1007/JHEP06\(2018\)100](http://dx.doi.org/10.1007/JHEP06(2018)100).
- [14] SM Higgs production cross sections at $\sqrt{s} = 13$ TeV (CERN Report 4). <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageAt13TeV>.
- [15] CMS Collaboration. CMS Luminosity - Public Results. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>.
- [16] Oliver Sim Brüning, Paul Collier, P Lebrun, Stephen Myers, Ranko Ostojic, John Poole, and Paul Proudlock. *LHC Design Report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2004. <https://cds.cern.ch/record/782076>.
- [17] Accelerating: Radiofrequency cavities. *CERN*. <https://home.cern/science/engineering/accelerating-radiofrequency-cavities>.
- [18] CERN. Pulling together: Superconducting electromagnets. <https://home.cern/science/engineering/pulling-together-superconducting-electromagnets>.
- [19] Stepan Kunc. *Study of the Magnetically Induced QED Birefringence of the Vacuum in experiment OSQAR*. PhD thesis, abr 2018. https://www.researchgate.net/figure/Cross-section-of-LHC-dipole-superconducting-magnet-62_fig14_324588058.
- [20] CMS Collaboration. Detector. <https://cms.cern/detector>.
- [21] Juan Rodrigo. *Estudio de procesos con presencia de pares de quarks top-antitop y energía transversa faltante en colisiones protón-protón con el experimento CMS en el Run 2 del LHC*. PhD thesis, may 2019.

- [22] G L Bayatian, S Chatrchyan, et al. *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software*. Technical Design Report CMS. CERN, Geneva, 2006. <https://cds.cern.ch/record/922757>.
- [23] Mario González. <http://www.hep.uniovi.es/wiki/>.
- [24] CMS Collaboration. Tracking. <http://cms.cern/detector/identifying-tracks>.
- [25] CMS Collaboration and A Holzner. 78 reconstructed vertices in event from high-pileup run 198609. <https://cds.cern.ch/record/1479324>, CMS Collection, Sep 2012.
- [26] Martin Lipinski. The Phase-1 Upgrade of the CMS Pixel Detector. Technical Report CMS-CR-2017-135. 06, CERN, Geneva, May 2017. <https://cds.cern.ch/record/2265423>.
- [27] Cristina Biino. The CMS electromagnetic calorimeter: overview, lessons learned during run 1 and future projections. *Journal of Physics: Conference Series*, 587:012001, feb 2015. <https://iopscience.iop.org/article/10.1088/1742-6596/587/1/012001>.
- [28] Francesca Cavallari. Performance of calorimeters at the LHC. *Journal of Physics: Conference Series*, 293:012001, apr 2011. <https://iopscience.iop.org/article/10.1088/1742-6596/293/1/012001>.
- [29] CMS Collaboration. HCAL, 2018. <https://cmshcal.web.cern.ch/>.
- [30] The superconducting solenoid for the CMS detector. CEA, 2013. http://irfu.cea.fr/dacm/en/Phoce/Vie_des_labos/Ast/ast_visu.php?id_ast=3011.
- [31] CMS Collaboration. Bending particles. <http://cms.cern/detector/bending-particles>.
- [32] CMS Collaboration. Detecting muons. <http://cms.cern/detector/detecting-muons>.
- [33] A.M. Sirunyan et al. Particle-flow reconstruction and global event description with the CMS detector. *JINST*, 12(CMS-PRF-14-001. CMS-PRF-14-001-004. 10):P10003. 82 p, Jun 2017. <https://cds.cern.ch/record/2270046>.
- [34] Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. 1960. <https://www.cs.unc.edu/~welch/kalman/media/pdf/Kalman1960.pdf>.

- [35] A.M Sirunyan et al. Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV. *JINST*, 13(06), 2018. <https://inspirehep.net/literature/1667449>.
- [36] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. The anti-ktjet clustering algorithm. *Journal of High Energy Physics*, 2008(04), 2008. <http://dx.doi.org/10.1088/1126-6708/2008/04/063>.
- [37] A. M Sirunyan et al. Performance of reconstruction and identification of τ leptons decaying to hadrons and ν_τ in pp collisions at $\sqrt{s}=13$ TeV. *Journal of Instrumentation*, 13(10):P10005–P10005, oct 2018. <https://doi.org/10.1088/2F1748-0221%2F13%2F10%2Fp10005>.
- [38] A. M. Sirunyan, A. Tumasyan, M. Dragicevic, and et al. Evidence for associated production of a Higgs boson with a top quark pair in final states with electrons, muons, and hadronically decaying τ leptons at $\sqrt{s} = 13$ TeV. *JHEP*, 08(arXiv:1803.05485. CMS-HIG-17-018-003):066. 50 p, Mar 2018. <https://cds.cern.ch/record/2308650>.
- [39] Torbjörn Sjöstrand, Stephen Mrenna, and Peter Skands. A brief introduction to pythia 8.1. *Computer Physics Communications*, 178(11):852–867, Jun 2008. <http://dx.doi.org/10.1016/j.cpc.2008.01.036>.
- [40] S. Agostinelli, J. Allison, et al. Geant4—a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250 – 303, 2003. <http://www.sciencedirect.com/science/article/pii/S0168900203013688>.
- [41] Rene Brun and Fons Rademakers. ROOT - An Object Oriented Data Analysis Framework. Proceedings AIHENP'96 Workshop, Lausanne, Sep. 1996, Nucl. Inst. & Meth. in Phys. Res. A 389 (1997) 81-86. See also <http://root.cern.ch/>.
- [42] CMS Collaboration. CMSSW Application Framework. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookCMSSWFramework>.
- [43] CMS Collaboration. Summaries of CMS cross section measurements. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsCombined>.
- [44] Luca Lista. Practical statistics for particle physicists. *arXiv preprint arXiv:1609.04150*, 2016. <https://arxiv.org/pdf/1609.04150.pdf>.

- [45] Eric W Weisstein. Log normal distribution. <https://mathworld.wolfram.com/LogNormalDistribution.html>.
- [46] Fred James and Matthias Winkler. MINUIT User's Guide. *InspireHEP*, 6 2004. <https://inspirehep.net/literature/1258345>.
- [47] Analysis of $t\bar{t}H$ and $t\bar{t}W$ production in multilepton final states with the ATLAS detector. Technical Report ATLAS-CONF-2019-045, CERN, Geneva, Oct 2019. <https://cds.cern.ch/record/2693930>.
- [48] Kyle Cranmer. Practical Statistics for the LHC. (arXiv:1503.07622):267–308. 41 p, Mar 2015. <https://cds.cern.ch/record/2004587>.
- [49] Povh Bogdan, Rith Klaus, Scholz Christoph, and Zetsche Frank. *Particles and Nuclei*. Springer, 6 edition, 2008. p (165-168).
- [50] Lyndon Evans and Philip Bryant. LHC machine. *Journal of Instrumentation*, 3(08):S08001–S08001, aug 2008. <https://iopscience.iop.org/article/10.1088/1748-0221/3/08/S08001>.
- [51] CMS Collaboration. Triggering and data acquisition. <https://cms.cern/detector/triggering-and-data-acquisition>.
- [52] CMS Collaboration. Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET. Technical Report CMS-PAS-PFT-09-001, CERN, Geneva, Apr 2009. <https://cds.cern.ch/record/1194487>.
- [53] CMS Collaboration. Description and performance of track and primary-vertex reconstruction with the CMS tracker. *Journal of Instrumentation*, 9(10):P10009–P10009, Oct 2014. <http://dx.doi.org/10.1088/1748-0221/9/10/P10009>.
- [54] CMS Collaboration. Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s}= 8$ TeV. *Journal of Instrumentation*, 10(06):P06005–P06005, jun 2015. <https://iopscience.iop.org/article/10.1088/1748-0221/10/06/P06005>.
- [55] CMS Collaboration. Performance of reconstruction and identification of tau leptons in their decays to hadrons and tau neutrino in LHC Run-2. Technical Report CMS-PAS-TAU-16-002, CERN, Geneva, 2016. <https://cds.cern.ch/record/2196972>.

-
- [56] Didar Dobur. Jets and Missing Transverse Energy Reconstruction with CMS. *PoS*, 2008LHC:039, 2014. <https://inspirehep.net/literature/817095>.
- [57] Luca Scodellaro. b tagging in ATLAS and CMS, 2017. <https://arxiv.org/abs/1709.01290>.
- [58] Artur Apresyan et al. Performance of MET reconstruction in CMS. *Journal of Physics: Conference Series*, 587:012006, feb 2015. <https://iopscience.iop.org/article/10.1088/1742-6596/587/1/012006>.
- [59] SM Higgs production cross sections at $\sqrt{s} = 13-14$ TeV (CERN Report 3). <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageAt1314TeV2014>.
- [60] CMS Collaboration. Simulating and Reconstructing events with Fast Simulation. <https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookFastSimulationArchive>.