

TRABAJO FIN DE GRADO
FACULTAD DE CIENCIAS

Estudio del proceso $t\bar{t}H$ en el experimento CMS del acelerador LHC (CERN) a 13 TeV

Autor

Víctor Rodríguez Bouza

Tutores

Francisco Javier Cuevas Maestro

Pietro Vischia

Curso 2016-2017



Universidad de Oviedo
Universidá d'Uviéu
University of Oviedo

Resumen

El LHC ha ofrecido desde su puesta en funcionamiento un marco incomparable para el estudio de la física de altas energías. Hoy en día se alcanzan energías en centro de masas de $\sqrt{s} = 13 \text{ TeV}$, que podrían dar lugar a descubrimientos físicos más allá del modelo estándar. En el futuro las mejoras del LHC le permitirán alcanzar una mayor luminosidad instantánea con la que observar procesos de los que hoy registramos pocos eventos, como el $t\bar{t}H$. Este se caracteriza por una gran variedad de estados finales, y permite medir directamente la constante de acoplo del Higgs al *quark top*, comparándola con la predicción del modelo estándar y poniendo así límites para física más allá del modelo estándar.

Hemos realizado análisis multileptónico de este proceso, empleando datos obtenidos del experimento CMS del LHC durante 2016 (35.9 fb^{-1}) y basándonos en un estudio previo realizado por la colaboración de dicho detector ([1]). Los resultados obtenidos no permiten realizar una medida de la sección eficaz del proceso, pero hemos puesto límites superiores en el valor de su *signal strength* ($\mu = \frac{\sigma_{\text{exp.}}}{\sigma_{\text{teo.}}}$), que son $\mu_{\text{obs.}} = 3.21$ y $\mu_{\text{esp.}} = 2.04^{+0.85}_{-0.59}$. Estos valores no nos permiten decir que hemos observado el proceso, si bien son resultados aceptables teniendo en cuenta el marco experimental de baja luminosidad.

Además, hemos observado una mejoría en la relación $\frac{\text{señal}}{\text{fondo}}$, así como en el indicador $\frac{\text{señal}}{\sqrt{\text{fdo.} + (\Delta \text{fdo})^2}}$ de un 0.1 % y de un 1.2 % respectivamente, al hacer una identificación de los leptones tau que decaen hadrónicamente más exigente. Esto permite una mejor observación del proceso y abre las puertas a mejoras en la selección de eventos.

Se ha estudiado la variable **Lepton MVA** usada en la identificación de leptones, obtenida de un BDT entrenado contra fondos de *fakes*. Hemos realizado una comparación sobre distintos criterios de selección en dicha variable, comprobando un aumento del recuento de eventos con *fakes* de hasta el 508 % de los del análisis estándar. Cambiando la identificación de leptones por otras sin **Lepton MVA**, llegamos a alcanzar el 907 % respecto a los que teníamos, dejando patente la eficacia de esta variable frente a este fondo.

Tangencialmente, se han analizado las diferencias entre simulaciones del proceso $t\bar{t} + \text{jets}$ de dos generadores: **aMC@NLO** y **Powheg**. Se han observado disparidades que se deben al recuento muy bajo de sucesos de la simulación sin pesar. Concluimos que la elección del generador en nuestro análisis podría influir en

los resultados del mismo y en algunas distribuciones de variables, sobre todo si se extrema la selección.

Índice general

Resumen	III
Introducción	1
1. El modelo estándar de física de partículas	3
1.1. Introducción histórica	3
1.1.1. Los atomistas (s. V-IV a.C.)	3
1.1.2. El concepto moderno de átomo y el «zoo» de partículas (s. XIX – XX)	5
1.1.3. La física de partículas con aceleradores y el modelo estándar (s. XX – actualidad)	10
1.2. El modelo estándar en la actualidad	13
1.3. Los límites del modelo estándar	16
1.4. El proceso $t\bar{t}H$	21
1.4.1. Descripción general	21
1.4.2. El canal multileptónico	22
1.4.3. Por qué estudiar $t\bar{t}H$	23
1.4.4. Situación actual de la investigación	23
2. La instalación LHC y el experimento CMS	25
2.1. El LHC	25
2.1.1. Descripción general	25
2.1.2. Funcionamiento	27
2.1.2.1. Encendido	27
2.1.2.2. Componentes esenciales	28
2.1.2.3. Luminosidad, <i>pile-up</i> y <i>trigger</i>	29
2.2. El detector CMS	32
2.2.1. Descripción general	32
2.2.2. Detector de trazas (<i>tracker</i>)	34
2.2.3. Calorímetro electromagnético (ECAL)	35
2.2.4. Calorímetro hadrónico (HCAL)	36
2.2.5. Solenoide superconductor	37
2.2.6. Sistema de muones	37

3. Descripción del trabajo experimental	39
3.1. Simulación de procesos	40
3.2. Reconstrucción de eventos e identificación de objetos	41
3.2.1. Vértice primario	42
3.2.2. Muones	43
3.2.3. Electrones	47
3.2.4. Hadrones: <i>jets</i>	49
3.2.5. Taus	50
3.2.6. Energía transversal faltante y otras cantidades características del proceso	51
3.3. Cálculo de secciones eficaces y límites superiores	52
4. Implementación computacional	55
4.1. Datos y muestras de Monte Carlo	56
4.2. <i>Triggers</i>	58
4.3. Identificación de objetos	61
4.3.1. Identificación de muones y electrones	61
4.3.2. Identificación de leptones tau	62
4.3.3. Identificación de <i>jets</i>	62
4.4. Los factores de escala (<i>scale factors</i> , SF)	63
4.5. Selección de eventos	65
4.5.1. Preselección	65
4.5.2. 2ISS	66
4.5.3. 3l	66
4.5.4. 4l	67
5. Resultados experimentales	69
5.1. Análisis del proceso $t\bar{t}H$	70
5.2. Estudio de los criterios de identificación de leptones	74
5.2.1. Comparación con otros análisis	74
5.2.2. Comparación con distintos criterios de selección de la variable Lepton MVA	77
5.3. Otras comparaciones y estudios	82
5.3.1. Cambios en la selección de leptones tau	82
5.3.2. Comparación de generadores aMC@NLO y Powheg para el proceso $t\bar{t} + jets$	82
6. Conclusiones	87
7. Bibliografía	91

Introducción

El primer objetivo de este trabajo de física de partículas experimental es la elaboración de un análisis funcional del proceso físico $t\bar{t}H$ empleando para ello datos recogidos en el experimento CMS del LHC durante el año 2016. Este estudio se corresponde con un análisis de alta luminosidad, puesto que se trata de un proceso que bajo las condiciones actuales del experimento de donde obtenemos los datos, no podemos observar (y si con unas condiciones de más luminosidad). A continuación, observaremos el papel que tiene la variable **Lepton MVA** (descrita más adelante en el texto) en la identificación de leptones. Este texto constituye una memoria de los esfuerzos dedicados a estos objetivos, así como a otros estudios tangenciales o secundarios.

Los conocimientos empleados guardan relación con las asignaturas «Física Nuclear y de Partículas Elementales», «Física de Altas Energías y Aceleradores», «Métodos Numéricos y sus Aplicaciones a la Física» e «Introducción a la Física Computacional». Otras asignaturas del Grado en Física también juegan un papel, aunque menor.

La estructura de este texto es la siguiente. El primer capítulo (precedido por un resumen del trabajo al completo y esta introducción) muestra una aproximación histórica a la física de partículas actual, que se puede resumir *grosso modo* en el modelo estándar. Se ofrece una descripción somera y cualitativa del mismo, a la que sigue una visión breve del proceso físico que estudiaremos: $t\bar{t}H$.

El segundo capítulo está dedicado a la descripción del experimento CMS y el acelerador de partículas LHC del CERN. Se ofrece una visión resumida de ambos, haciendo hincapié en cómo se recogen los datos y cómo se detectan las señales de los objetos físicos de los eventos. Se puede encontrar una relación del proceso conjunto de trabajo en física de partículas entre toda la comunidad científica del CMS y del LHC, a través de los sistemas de gestión y análisis de datos.

La descripción del proceso de reconstrucción de los objetos físicos se enmarca en el tercer capítulo. También ofrecemos una resumida explicación del proceso de identificación de estos objetos.

La elaboración del análisis se condensa en el cuarto capítulo. En ese epígrafe se ofrece una relación de

las muestras de datos y simulaciones usadas, así como las distintas partes (*triggers*, identificación de objetos, factores de escala y selección) del análisis estándar. Si bien pequeñas partes fueron cambiadas en los distintos estudios hechos, el grueso se mantiene siempre constante.

El siguiente capítulo recoge todos los resultados experimentales. En su primera sección explicamos los frutos del análisis que hemos hecho, entre los que están el límite a la sección eficaz del proceso. En la segunda describimos los estudios sobre el papel de la variable **Lepton MVA** en la identificación de los leptones, acabando después con aquellos estudios o resultados tangenciales a los objetivos que nos impusimos para este trabajo.

Finalmente, el último capítulo contiene las conclusiones de todo el trabajo, a las que sigue la bibliografía.

1 El modelo estándar de física de partículas

En un agujero en el suelo, vivía un hobbit. No un agujero húmedo, sucio, repugnante, con restos de gusanos y olor a fango, ni tampoco un agujero seco, desnudo y arenoso, sin nada en que sentarse o que comer: era un agujero-hobbit, y eso significa comodidad.

(*El Hobbit*, J.R.R. TOLKIEN)

HOY en día el modelo estándar de física de partículas, el marco teórico físico en el que la ciencia de altas energías se desenvuelve, está asentado y está considerado uno de los pilares de la concepción física del universo. Sin embargo, el descubrimiento e investigación de todas las leyes físicas que lo componen, no fue una tarea mundana.

En este capítulo, presentamos un resumen de la aproximación histórica al modelo estándar, seguida de una descripción cualitativa y somera del mismo. Finalizamos con una descripción del proceso que será objeto de estudio: el proceso $t\bar{t}H$.

1.1. Introducción histórica

1.1.1. Los atomistas (s. V-IV a.C.)

La curiosidad humana fue seguramente lo que llevó al primer filósofo de la historia, Tales de Mileto (circa 624 - 546 a.C.), a dar una explicación *racional* de distintas experiencias o realidades que conoció en vida. El paso de las explicaciones mitológicas (provenientes del *mythos*¹) a las basadas en argumentos lógicos, o racionales (*logos*) dio como fruto el nacimiento de la filosofía y, con ella, de la ciencia. Sería complicado tildar a las «filosofía» y «ciencia» de entonces de filosofía y ciencia actuales, puesto que no se circunscribían a las definiciones de ahora: el desarrollo de ambas disciplinas estuvo entrelazado desde el comienzo, y en aquellos tiempos hablar de la una o la otra, equivalía a lo mismo.

La intención de esos primeros pensadores, denominados **presocráticos**, consistía en buscar una esencia o principio (el *arkhé*²) a partir del cual poder explicar la totalidad, la naturaleza: el universo (la *phy-*

¹Del griego antiguo *μῦθος*: cuento, transcripción.

²*Idem*, de *ἀρχή*: «fuente», «principio», «origen».

*sis*³). Entre ellos, había unos que entendían el *arkhé* como *estoikheion*, que significa «elemento». Según estos presocráticos (quienes reciben el nombre de *pluralistas*) los elementos, el *arkhé*, seguían siendo la «esencia» de las cosas, aunque de forma distinta, pues las cosas estaban hechas de elementos, estaban hechas de *arkhé*⁴.

Se considera a Leucipo de Mileto (circa s. V a.C., contemporáneo de Sócrates) el fundador de la escuela atomista, aunque sin duda el más representativo de la corriente sería Demócrito de Abdera (circa 460 - 370 a. C., contemporáneo de Platón). Los **atomistas** consideraban que el universo estaba hecho de un número infinito de átomos⁵ que son eternos, cualitativamente iguales (i.e., que se distinguen solo por su forma, por su apariencia) y que se desplazan en un vacío eterno.

No es posible considerar esta como la primera concepción científica de partículas elementales, o de una explicación científica de los constituyentes de la materia, puesto que no se enmarcan dentro de lo que hoy entendemos por «ciencia». Sin embargo, parece procedente hacer un pequeño comentario de una de las primeras explicaciones *racionales* sobre los constituyentes de la materia, y que se acerca (salvando las distancias) más que otras de su tiempo a la visión que tenemos hoy en día. Habrá que esperar hasta el advenimiento de la ciencia como hoy la entendemos, cosa que sucederá en la Revolución Científica que tuvo lugar con el Renacimiento. Bajo este nuevo paradigma será en el siglo XIX donde encontraremos las primeras ideas científicas respecto al tema. El periodo de tiempo anterior al Renacimiento y posterior a los filósofos presocráticos atomistas, correspondiente en esencia a la Edad Media, estaba dominado en Europa occidental y central por las ideas aristotélicas, defendidas por la Iglesia, quien en esencia marcó el pensamiento filosófico. Las enseñanzas de Leucipo y sus sucesores no tuvieron cabida en el dogma eclesiástico.

³ *Idem*, de *Φύσις*: «naturaleza», aunque entendida como «totalidad», como «universo», no como la traducción literal.

⁴ Sin entrar en detalles y en general, tales «elementos» respetaban las características del ser de Parménides (i.e., unicidad, permanencia, inmutabilidad, indivisibilidad y limitación) y la motivación tras estos elementos era poder explicar la diversidad del mundo y el cambio, la transformación.

⁵ Del griego *ἄτομον*: «indivisible», «no cortado», «sin partes».

1.1.2. El concepto moderno de átomo y el «zoo» de partículas (s. XIX – XX)

Desde la química se hicieron avances durante el siglo XIX que allanaron el camino de los modelos atómicos. John Dalton (1766 - 1844) sería quien «resucitara» el concepto de átomo. Primero, en un *paper* publicado en 1805, aunque leído en 1803, donde afirmaba ([2, p. 286]):

*Why does not water admit its bulk of every kind of gas alike? This question I have duly considered, and though I am not able to satisfy myself completely I am nearly persuaded that the circumstance depends on the weight and number of the ultimate particles of the several gases: [...]*⁶

La tabla periódica de Mendeleev (1869), así como desarrollos químicos anteriores a Dalton (ley de conservación de las masas de Lavoisier, ec. de los gases ideales de Gay-Lussac, etcétera) permitieron que la concepción atómica de la materia, (i.e.: esta está hecha de unas partículas elementales⁷ llamadas átomos) se viera elevada al nivel de teoría científica. Los químicos habían medido ya masas atómicas, aunque el concepto de «número atómico» no había sido desarrollado aún. Así mismo, se conocían los rayos catódicos y se habían medido ya espectros atómicos.

A finales de siglo se dieron distintos avances en áreas distintas de la física (y química). Comenzó esta avalancha W. C. Röntgen en 1895, con el hallazgo de los rayos X, seguido de Henri Becquerel al año posterior: sus investigaciones le llevaron a observar un oscurecimiento en láminas fotográficas debido a la radiación del uranio. Así mismo, Marie Curie y su marido Pierre descubrieron, además de realizar otros aportes a la física y a la química, dos nuevos elementos: el polonio y el radio.

Uno de los desarrollos más importantes fue el descubrimiento por el británico J. J. Thomson (1856 - 1940) del que todavía hoy permanece como partícula elemental (i.e., una de las últimas partes en las que la materia se puede dividir): el **electrón**⁸ (1897), sirviéndose del estudio de los rayos catódicos. Tras el descubrimiento, Thomson desarrolló su propio modelo atómico, que se basaba en muchos electrones que «revoloteaban» alrededor de carga positiva. La materia no cargada tendría un equilibrio entre ambas que las anularía.

Si bien en la segunda mitad del siglo existía la percepción entre los científicos de que la física estaba «acabada», todos estos descubrimientos sirvieron para eliminar de raíz tal idea. Sin embargo, el final de siglo (1896 - 1900) depararía aún la catastrofe ultravioleta y el advenimiento de la física cuántica, que

⁶ «¿Por qué el agua no entiende su volumen como el de cualquier tipo de gas? He considerado debidamente esta pregunta, y aunque no me satisfaga completamente mi respuesta, estoy casi convencido de que este [el volumen] depende del peso y el número de las partículas últimas de los distintos gases: [...]».

⁷ Cuando hablamos de partículas elementales nos referimos a entes **indivisibles**: es decir, a los componentes últimos de la materia.

⁸ Del griego *ήλεκτρον*: «ámbar».

revolucionaría los cimientos de la ciencia.

Ya en el siglo XX, las investigaciones sobre radioactividad que Becquerel, los Curie, etc. habían comenzado décadas atrás fueron continuadas por otros. Entre ellos destacaría el neozelandés Ernest Rutherford (1871 - 1937). Investigando en la Universidad de Mánchester, llevó a cabo el famoso experimento de la lámina de oro (1909), lo que le permitió establecer su modelo atómico (en el que ya estaría el núcleo, de carga positiva) al ser incompatibles los resultados experimentales con el modelo de Thomson. Rutherford bautizó como **protón**⁹ al núcleo del hidrógeno.

El avance de Rutherford allanaría el terreno para el danés Niels Bohr, quien desarrolló su propio modelo atómico (1913) con un formalismo teórico mayor que sus predecesores. Concretamente, era capaz de derivar la fórmula de Balmer para el espectro de hidrógeno y la constante de Rydberg. Más tarde, en 1914, Moseley aplicó técnicas de difracción de rayos X (en las que habían trabajado von Laue, Friedrich y Knipping) para comprender finalmente el concepto de número atómico, el cual daba particularmente la carga positiva del núcleo de los átomos. con él, se pudo establecer un orden en la tabla periódica, rellenándola con los elementos conocidos si era posible; se entendió que en las posiciones había «huecos», faltaban elementos no descubiertos aún. Además, la comunidad científica tardó poco en darse cuenta que la diferencia en masas atómicas respecto a la del hidrógeno parecía variar en cantidades prácticamente enteras, además de que había átomos que eran químicamente inseparables pero con distinta masa atómica: los isótopos.

Durante las dos primeras décadas del siglo, y en paralelo con lo relatado anteriormente, Albert Einstein desarrolló y publicó sus teorías especial y general de la relatividad, y la física cuántica siguió su desarrollo. Particularmente, durante los años 20: la mecánica cuántica (QM, del inglés *Quantum Mechanics*) se construyó con el trabajo de muchos físicos, quienes se enfrentaron a notorios problemas no solo matemáticos, sino también filosóficos. Resulta imperativo citar la labor de Schrödinger, Heisenberg, Bohr o Dirac, entre otros. Al final, se constituyó en una teoría científica que era capaz de describir con una serie de ecuaciones diferenciales el átomo de hidrógeno, así como los de otros elementos (aunque estos no de manera *exacta*).

Los avances en la física cuántica revelaron las carencias de los modelos atómicos, que fueron progresivamente siendo sustituidos por otros que respetaban los nuevos marcos teóricos. Tras el modelo de Bohr llegaría el de Sommerfeld (1916) y más tarde el de Schrödinger (1924). Este era ya cuántico, pero no relativista y habría que esperar a que Dirac desarrollase la ecuación que lleva su nombre (1928) para llegar a la concepción actual (relativista y cuántica) que tenemos en la actualidad del átomo. Pero para

⁹ *Idem*, de *πρῶτον*: «primero».

completar el rompecabezas y llegar a esa concepción faltaban algunas «piezas».

Con el desarrollo paulatino de la mecánica cuántica, la concepción de los átomos hasta entonces se fue destruyendo, debido a que se hacía imposible reconciliar los progresivos avances teóricos con los átomos conformados de protones y electrones exclusivamente. El propio Rutherford en 1920 ya había postulado la existencia del neutrón, aunque la comunidad científica consideraba, y siguió considerando por algún tiempo, que los «núcleos» tenían A protones y $A - Z$ electrones. Sería *Sir* James Chadwick (1892 - 1974) quien probara la existencia del mismo. Para ello, se sirvió del trabajo que Walter Bothe y Herbert Becker (y más tarde Irène Curie y Frédéric Joliot) habían hecho estudiando el proceso químico siguiente. En él, dos átomos de helio (He_2^{410}) y berilio (Be_4^9) dan lugar a uno de carbono (C_6^{12}) y un neutrón (n_0^1), aunque por aquel entonces no se sabía que los neutrones eran de hecho quienes participaban en ella: del proceso se conocía que emitía una radiación de algún tipo.



Curie y Joliot, quienes probaron que la radiación emitida era capaz de arrancar protones al hacerla incidir sobre parafina, entendieron que la misma eran rayos gamma (es decir, fotones).

Chadwick publicó en 1932 los resultados de su análisis del proceso, mostrando que considerando la velocidad a la que salían eyectados los protones ($\approx \frac{1}{10}c$), el fotón que hubiera hecho el *scattering* sobre ellos, tendría que tener una energía muy grande ($\approx 50 \text{ MeV}$) para los rayos gamma que normalmente emiten los núcleos ($\approx \text{MeV}$). Además, de hacer incidir la radiación sobre átomos de nitrógeno, ni siquiera con fotones tan energéticos se explicaría la energía que tendrían los átomos eyectados. Sin embargo, estos problemas desaparecían si se asumía que la radiación provenía de otra partícula de masa similar o igual a la del protón, pero neutra: el **neutrón**.

Uno de los procesos en los que participa este constituyente atómico es la conocida como desintegración beta. Cuando se empezó a estudiar este proceso, se analizaba como aquel en el que un neutrón (n) se desintegraba en un protón (p^+) y un electrón (e^-), lo cual podemos escribirlo como sigue.

$$n \longrightarrow p^+ + e^- \quad (1.2)$$

Sin embargo, a la hora de examinar espectros de los electrones (de radiación beta) originados en el proceso, se encontraba que la conservación de la energía no se cumplía. El austriaco Wolfgang Pauli (1900 - 1958) postuló la existencia en 1930 de otra partícula, el **neutrino**, (ν) que sería muy ligera, neutra y que interaccionaba poco con el resto de la materia, de modo que la diferencia de energía

¹⁰Con la notación X_Z^A , donde X es la representación del elemento en la tabla periódica y Z y A son el número atómico y el másico respectivamente.

encontrada en el proceso pudiera asignársele a él. Esto posibilitó un modelo cuantitativo del proceso, aunque no probaba su existencia. Habría que esperar a 1956 para que esto sucediera, de la mano de Cowan, Reines, Harrison y Kruse. El proceso correcto que se da es el siguiente.

$$n \longrightarrow p^+ + e^- + \nu_e \quad (1.3)$$

Otro austríaco, Víctor Hess, realizó una serie de experimentos entre 1911 y 1912 en los que medió, usando un globo aerostático, la radiación ionizante que existe en la atmósfera. Robert Andrews Millikan, en 1925, llamó a esta radiación «rayos cósmicos» y confirmó la conclusión de Hess, según la cual esta radiación venía del espacio exterior. Para demostrarlo ([3]), midió la ionización que se producía por estos rayos en dos lagos ubicados a 1500 y 3600 m sobre el nivel del mar. El resultado probó que se daban las mismas ionizaciones por minuto en el lago situado a 1500 m que en el lago superior pero a dos metros de profundidad. Por lo que los casi dos mil kilómetros de aire de diferencia absorbían necesariamente parte de la radiación, y esta debía de venir, por tanto, del espacio.

Una de las características de estos rayos es que sus energías son superiores a las que se podían obtener con fuentes radioactivas o rayos X, por lo que interesaron eventualmente a muchos físicos de partículas. El físico soviético Dimitry Skobelzyn sería pionero en usar una cámara de niebla¹¹ para observar rayos cósmicos.

El mero hecho de agregar un campo magnético a las cámaras de niebla posibilitaba la diferenciación de la carga de las partículas (cuando iban a la velocidad necesaria para que la curvatura fuera apreciable). Y esto fue indispensable para que C. D. Anderson descubriera en 1932 el **positrón**. Para ser exactos, él concluyó en su artículo que su descubrimiento era una nueva partícula de carga positiva, con carga en valor absoluto menor que dos veces la del electrón y una masa mucho menor que la del protón. El ingrediente que faltaba para que encajara en el rompecabezas con el resto de partículas lo había dado Dirac algunos años antes, cuando desarrolló su ecuación homónima, postulando la existencia de las **antipartículas**: partículas idénticas en masa a las conocidas, pero de carga eléctrica opuesta. El positrón sería identificado entonces como la antipartícula del electrón.

Los experimentos con cámaras de niebla continuaron, y en 1937 Anderson y S. H. Neddermeyer hicieron mediciones de pérdidas de energía de las partículas al atravesarlas, pues todas las partículas que atraviesan una, pierden energía. Existía una modelización de esta pérdida, la teoría de Bethe-Heitler, según la cual la energía perdida sería proporcional a la energía incidente. Sin embargo, Anderson y Neddermeyer comprobaron que había un grupo de partículas que se comportaba como Bethe-Heitler predecía, y otro que no. Además de que los que no la seguían no podían asociarse con fenómenos de

¹¹Contínente relleno de vapor sobresaturado que permite ver la traza de radiación ionizante.

cascadas electromagnéticas¹².

Aunque estas partículas «díscolas» ya habían sido detectadas por otros científicos, no había acuerdo en saber qué eran. Anderson y Neddermeyer pudieron descartar por sus mediciones que ninguno de los dos tipos de partículas eran ni protones ni electrones, pues su masa no se correspondía con la de estos. Además de estos experimentos, Street y Stevenson hicieron otros (1937) que corroboraban y precisaban más estos resultados, poniendo cotas en cuanto a masa de una de ellas, la cual estaría entre la del electrón y la del protón. Este sería el **muon** (nombre que recibiría más tarde), μ .

La investigación sobre los muones y las otras de partículas continuó, puesto que quedaba por saberse la naturaleza de estas. Relevante para la resolución fue la detección de desintegraciones de estas en muones. Sería un grupo de investigadores de Bristol, Lattes, Occhialini y Powell, quienes aclararan todo: las dos partículas medidas se trataban del ya citado muon y otra, denominado **pion** (ambos nombres dados por estos tres investigadores de Bristol), π . Y las desintegraciones observados se podían entender como piones decayendo en muones y neutrinos. También se observó que los piones aparecían con cargas de distinto signo, π^+ , π^- .

La teoría cuántica relativista de Dirac que habíamos mencionado (denominada más tarde teoría cuántica de campos o QFT, de *Quantum Field Theory*) dio más tarde lugar a la interpretación del electromagnetismo y la electrodinámica de forma también relativista y cuántica, lo cual se terminó llamando electrodinámica cuántica, QED (de *Quantum Electrodynamics*). Se terminó de edificar a finales de los años cuarenta estando entre sus más destacados impulsores Richard Feynman, Hans Bethe, Sin-Itiro Tomonaga, Julian Schwinger y Freeman Dyson.

Si bien QED explicaba las interacciones electromagnéticas de una forma cuántica y relativista, no se tenía lo equivalente para las otras tres fuerzas fundamentales: fuerte, débil y gravitatoria. Ignorando la gravitatoria, por sus peculiares características, durante las décadas siguientes se construyeron teorías tomando como modelo QED que explicaban las interacciones fuerte y débil. La teorización que se hizo para explicar las interacciones fuertes recibió el nombre de cromodinámica cuántica (QCD, de *Quantum Chromodynamics*) y fue desarrollada en la década de 1960. Se basa en la existencia de una serie de partículas *elementales*, los **quarks** (postulados por Murray Gell-Mann y George Zweig en 1963),

¹²Proceso físico consistente en una sucesión de desintegraciones de fotones en pares electrón-positrón que radian fotones que a su vez se desintegran en otros pares electrón-positrón. Este fenómeno se da, por ejemplo, cuando hay electrones que sufren radiación de frenado (*Bremsstrahlung*) al atravesar la materia. Cuando estos pasan cerca de un núcleo atómico (de carga positiva), se sienten atraídos por la fuerza electromagnética hacia él, por consiguiente, acelerados y necesariamente, sufren *Bremsstrahlung*, emitiendo fotones que decaen en pares electrón-positrón, y así sucesivamente. Las cascadas se detienen cuando las partículas no tienen suficiente energía para continuar el proceso: cuando un fotón no tiene la suficiente energía como para desintegrarse en un electrón y un positrón.

con tres sabores distintos, carga eléctrica no nula y una carga propia, llamada de color¹³. Estas partículas conforman todos los hadrones, y la interacción fuerte es mediada por unos bosones llamados **gluones**.

Los descubrimientos de muchos de los que hoy conocemos como **hadrones** (i.e., partículas conformadas de *quarks*) tienen lugar en los años cuarenta y cincuenta. En ese periodo de tiempo, los distintos **mesones** (formados por una pareja *quark* - *antiquark*) y **bariones** (formados por tres *quarks* de color distinto) fueron descubriéndose progresivamente, utilizando instrumentos como cámaras de burbujas y cámaras de chispas¹⁴.

La mera observación de partículas a través de instrumentos como los ahora citados ha dado lugar a un grandísimo avance en la física y en la comprensión científica del universo, en general. Sin embargo, nos hace dependientes de los rayos cósmicos o de las fuentes de radiación de partículas cuando deseemos realizar experimentos. Además, avances como el descubrimiento de partículas antes desconocidas, en general más masivas que las anteriormente descubiertas, requieren depender de radiación bastante energética ($\gtrsim$ GeV), lo cual no es tan común como pudiera ser deseable a efectos prácticos.

Debido a ello, muchos de los avances científicos en física de partículas durante el siglo XX (y después) se han llevado a cabo en aceleradores de partículas: máquinas que permiten dotar a las estas de una energía determinada a través de su aceleración, para después proceder a su estudio. Por su importancia histórica, además de porque la fuente de datos de este trabajo es un experimento ubicado en un acelerador de partículas, les dedicamos el siguiente subapartado.

1.1.3. La física de partículas con aceleradores y el modelo estándar (s. XX – actualidad)

Los aceleradores de partículas existían desde los años treinta, pero estaban muy limitados en energía ($\simeq$ 15 MeV). Sería a finales de los años cuarenta cuando los primeros sincrociclotrones aparecieron (y más tarde, los sincrotrones). Con ellos, se descubrió el pion de carga neutra, π^0 , y los nuevos regímenes de energías que poco a poco se alcanzaron abrirían la puerta a nuevos descubrimientos de partículas (como el antiprotón, entre muchos otros) y sustentos experimentales para los avances teóricos de lo que en el futuro constituirá el modelo estándar de física de partículas. Veamos brevemente en qué consisten los principales exponentes: ciclotrones, sincrociclotrones y sincrotrones.

¹³ Véase la sección siguiente, 1.2.

¹⁴ Ambos dos son instrumentos de detección de partículas. La cámara de burbujas es similar en concepto a la cámara de niebla, aunque la de chispas es más diferente.

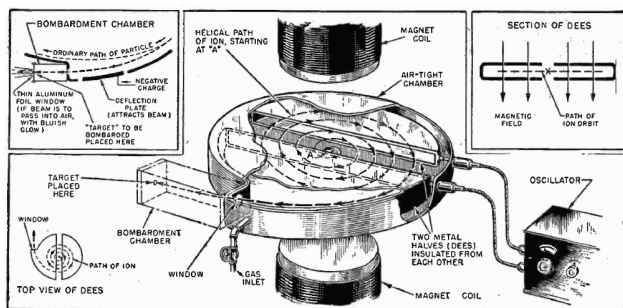


Figura 1.1: [4] Esquema de un ciclotrón.

Los sincrociclotrones son mejoras de otros aceleradores: los **ciclotrones**. Estos permiten acelerar iones haciendo uso de campos magnéticos y electromagnéticos. Como se aprecia en la figura 1.1, un ciclotrón está formado de dos semicírculos o «des» que se someten a una señal eléctrica alterna de alta frecuencia. Los iones, que entran desde el centro del ciclotrón, se encuentran embebidos

en un campo magnético que atraviesa las des de forma perpendicular, de modo que se ven forzados a tener una trayectoria circular. Sin embargo, como tienen carga y existe una diferencia de potencial entre las dos des (por la señal alterna eléctrica), los iones se aceleran, adquiriendo una trayectoria en espiral. Los sincrociclotrones constituyen una evolución, donde se mejora la frecuencia de la señal alterna eléctrica para compensar los efectos relativistas que ocurren cuando los iones se mueven a velocidades cercanas a la de la luz.

Los **sincrotrones**, por el contrario, son aceleradores de partículas en los cuales se consigue que estas se aceleren y mantengan una trayectoria cerrada. Esto se consigue con el uso de campos magnéticos y/o eléctricos que deben estar *sincronizados* (de ahí el nombre) con la frecuencia a la que temporalmente pasan las partículas, que según son aceleradas, varía. Uno de los ejemplos más reconocibles de sincrotrón es el Gran Acelerador de Hadrones, o LHC (*Large Hadron Collider*), del que hablaremos más en detalle en la sección 2.1.

Gracias a los experimentos en 1968 en el acelerador lineal de Stanford (SLAC) se comprobó la existencia de **partones** (i. e., *quarks*, *antiquarks* y gluones), que son los constituyentes de los protones, los neutrones, y todos los hadrones. Concretamente, se demostró la existencia de *quarks up* y *down*, así como indirectamente del *strange*. Además, los esfuerzos colaborativos iniciados con el PS (*Proton Synchrotron*) del CERN y seguidos por Fermilab y otras colaboraciones internacionales permitieron medir las funciones de estructura de distintas partículas.

También en SLAC se dio a conocer el leptón tau, τ en 1975 de la mano de Martin Lewis Perl, justo cuando el año pasado, 1974, habían descubierto el *quark charm*. Si bien inicialmente QCD contemplaba solo tres *quarks* (*up*, *down* y *strange*) con los años se fueron añadiendo más para que encajara, por ejemplo, con la descripción de la interacción débil (como es el caso del *quark charm*). Para terminar los seis *quarks* que hoy en día el modelo estándar contempla habrá que esperar hasta 1977, cuando se descubriría el *bottom* y mucho más, hasta 1995, para el *top* (ambos en Fermilab).

A su vez, fue con ayuda de los aceleradores de partículas que se probó la existencia de más de un neutrino. En 1962 Lederman, Schwartz y Steinberg demostraron la existencia del neutrino muónico y en los albores del tercer milenio se probaría en Fermilab que también había neutrinos tauónicos. La existencia de estos últimos ya había sido predicha, pues los resultados de medición de la masa del bosón Z en el LEP del CERN eran compatibles con la existencia de tres neutrinos. Al acabar el año 2000 los neutrinos encontrados eran tres: electrónico (ν_e), muónico (ν_μ) y tauínico (ν_τ).

Si los gluones transmitían la fuerza fuerte, y los fotones transmiten la electromagnética, ¿qué bosones transmiten la fuerza débil? El marco teórico que sirvió para describirla comenzó siendo la teoría de Fermi de la interacción débil, que intentaba hacer una analogía con QED: «un QED para la fuerza débil». No podía, debido a las diferencias entre la fuerza EM y la débil, ser igual, y entre otras cosas se vio necesaria la existencia de un bosón muy masivo (el bosón W) que la transmitiera. Por desgracia, tanto la idea original de Fermi, como modificaciones o mejoras de la misma, encontrarían problemas. En 1954 los primeros conatos de solución aparecieron de la mano de C. N. Yang y R. Mills, que fueron desarrollados y mejorados por (entre otros) Peter Higgs, Glashow, Weinberg o Salam.

Esta nueva teoría era una *generalización* de QED y la teoría de Fermi de la int. débil que las aglutinaba en una sola: la teoría electrodébil. Esta permitiría explicar las interacciones de ambas fuerzas (electromagnética y débil) a través de tres bosones: uno no masivo (el **fotón**, (γ)) y otros dos *muy* masivos (los bosones **Z** y **W**). Ambos bosones con masa serían descubiertos en 1983 con los datos obtenidos en los experimentos UA1 y UA2 del acelerador $S\bar{p}\bar{p}S$ en el CERN.

Fue entonces cuando comenzó a hablarse de **modelo estándar de física de partículas**, o modelo estándar (*Standard Model of Particle Physics*, *Standard Model*, o simplemente «SM»), aunque se hiciera hincapié sobre todo en la unificación entre electromagnetismo y fuerza débil. Podría decirse que el marco teórico electro-débil, junto con los alcances en ese momento de QCD era lo que empezó a denominarse SM. Hoy en día el concepto de modelo estándar ha cambiado, pues se han ido confirmando experimentalmente distintas partículas que pertenecen a él, y se ha justificado la masa de todas las partículas con el descubrimiento del bosón de Higgs (la última incorporación al mismo, en 2013).

El modelo estándar constituye la explicación científica más precisa sobre los constituyentes de la materia que ha existido nunca, tanto considerando los conceptos iniciales (pre-científicas¹⁵) de Leucipo, como ya desde el modelo (científico) atómico de Dalton, además de que ha sido refrendada experimentalmente en numerosas ocasiones. Explicaremos con más detalle qué se considera por modelo estándar de la física de partículas en la siguiente sección.

¹⁵Entendiendo por «ciencia» la concepción moderna surgida de la Revolución Científica.

1.2. El modelo estándar en la actualidad

El **modelo estándar de física de partículas**, más conocido por **modelo estándar** (o por las siglas **SM** del inglés *Standard Model*) es un modelo¹⁶ científico basado en un conjunto de teorías científicas que describe todos los constituyentes últimos de la materia descubiertos y las interacciones fuerte, débil y electromagnética entre ellos. Ha sido comprobado experimentalmente en muy numerosas ocasiones y de forma cada vez más precisa. En este texto no hablaremos de forma rigurosa sobre él, puesto que ello conllevaría demasiado y necesitaría conocimientos (QFT) no impartidos en el grado: nos limitaremos a dar una descripción más *cualitativa*.

Todo el modelo estándar está sustentado sobre la teoría cuántica de campos (relativista y cuántica), donde se apoyan la teoría electrodébil y la cromodinámica cuántica, que son la esencia del SM junto al mecanismo de Higgs. Este último permite explicar las masas de todas las partículas elementales que la tengan.

Según el modelo estándar, los componentes últimos de la materia común se denominan **fermiones**, y las interacciones entre sí que sufren se explican como intercambios de otras partículas mediadoras, llamados **bosones**. El bosón de Higgs se incluye porque es necesario para dotar de masa a aquellas partículas que la tengan (como él mismo). En la figura 1.2 se aprecian los fermiones y los bosones con un solo gluon, el par de bosones W y ninguna antipartícula fermiónica. El conjunto total de partículas es mayor, puesto que hay que considerar todas las antipartículas de los fermiones y los ocho gluones (se explicará más adelante). Recordemos que la antipartícula de una partícula es otra similar a la original pero que tiene carga eléctrica opuesta.

¹⁶Existe cierta controversia sobre si el «modelo» estándar es una «teoría» científica o un «modelo» (¿científico?). Podría decirse que una **teoría científica** es una colección de enunciados matemáticos deducidos de un conjunto finito de axiomas que describen un aspecto del universo y que han sido refrendados experimentalmente (p. ej., la mecánica cuántica o la teoría de la relatividad general).

Por otra parte, un **modelo** podría definirse como un caso (muy) particular de teoría científica. Por ejemplo, para explicar el movimiento transversal de una cuerda atada por sus extremos podemos preparar unas ecuaciones o enunciados (como hicimos en varias asignaturas durante el curso) que lo explique; cosa que podríamos hacer recurriendo a la ecuación de ondas, p. ej.. Tal modelo sirve para explicar un aspecto *muy* concreto y en un rango de energías *muy* delimitado del sistema físico que es una cuerda, y claramente sería posible hacer una explicación mucho más elaborada y *general* de tal sistema (¿de qué está hecha la cuerda? ¿Cómo afectan las cuatro fuerzas fundamentales a sus componentes? ¿Se mueven estos internamente?) que en virtud a su «nivel de generalidad» podría llegar a tildarse de *teoría* científica. Como vemos, los bordes entre «teoría» y «modelo» no son ni mucho menos nítidos.

En el **caso del SM** existe controversia, pues aunque sea una descripción de las interacciones débil, fuerte y electromagnética entre todas las partículas elementales experimentalmente descubiertas, hechos como que no explica el $\approx 95\%$ de la masa-energía del universo, su dependencia de «numerosos» parámetros o que no se incluye la interacción gravitatoria, entre otros, hacen que parte de la comunidad científica considere al modelo estándar «solo» un *modelo*. Sin embargo, otra parte piensa que el nivel de generalidad es suficiente para tildarlo de *teoría* científica.

Aunque quien firma este texto se inclina por *modelo* (basándose en unas definiciones que, si bien muy generales, también él ha dado), en parte por la dependencia de parámetros, parece necesaria esta aclaración. De este modo, el lector podrá comprender la diversidad de opiniones de la comunidad científica.

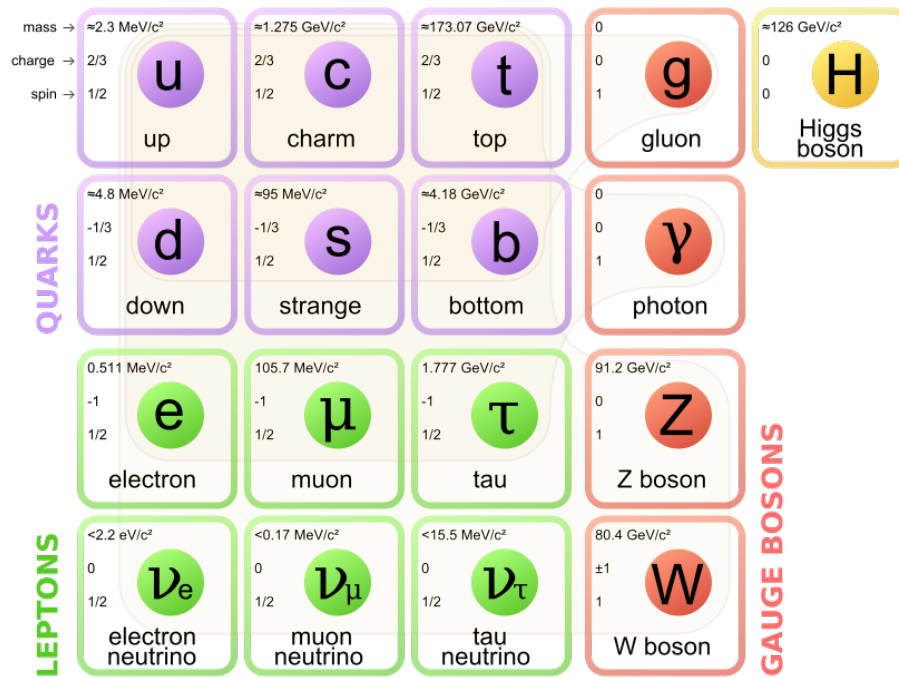


Figura 1.2: [5] Partículas elementales que describe el modelo estándar. En morado se han pintado los *quarks*, en verde los leptones (que, en conjunto suman los fermiones) y en rojo, los bosones mediadores (i.e., vectoriales, de espín la unidad). En amarillo a la derecha se encuentra la última incorporación al SM: el bosón de Higgs.

Más formalmente, los fermiones son aquellas partículas definidas por tener espín semientero. Estas siguen la estadística de Fermi-Dirac y su función de onda es antisimétrica respecto a cambios de posición, mientras que los bosones se definen como poseedores de espín entero, función de onda simétrica respecto a cambios de posición y además siguen la estadística de Bose-Einstein.

Como hemos dicho, la forma de interaccionar entre sí dos fermiones es a través del intercambio de un bosón, pero para que se dé este, es necesario que los fermiones en concreto posean la carga asociada a la interacción que están sufriendo. De acuerdo con el modelo estándar, cada fuerza lleva asociada una carga: EM - eléctrica, fuerte - color, débil - débil. Por ello, si dos partículas quieren sufrir una interacción electromagnética, p. ej., deberán tener carga eléctrica no nula.

Los bosones encargados de mediar la **interacción débil** son el bosón Z (neutro) y los bosones $W^\pm$. Se diferencian del resto de bosones mediadores en tanto que son especialmente masivos (motivo por el cual se tardaron tanto en descubrir experimentalmente), lo cual también explica el rango limitado de la fuerza débil. Todos los fermiones poseen carga débil, así que pueden sufrir la fuerza débil mediada por estos bosones (o, dicho de otra forma, todos los fermiones *se acoplan* a ellos).

El caso de la **interacción fuerte** es más complicado. Es responsabilidad del gluon, g , mediar esta fuerza. Este bosón es neutro eléctricamente, no tiene carga débil y posee, a diferencia de sus com-

pañeros mediadores, la carga asociada a la interacción que gobierna: la carga de color. Esta diferencia complica mucho la explicación de estas interacciones (QCD), en contraposición con la teoría electrodébil que explica las otras dos fuerzas. Por ejemplo, propiedades exclusivas de las partículas que sufren esta fuerza como el confinamiento del color se asocian con este hecho. Tampoco otros mediadores pueden interactuar consigo mismos, cosa de que sí son capaces los gluones.

Los fermiones que tienen carga de color y, por lo tanto, sufren la fuerza fuerte, son seis y se denominan **quarks**. También tienen carga débil y carga eléctrica; son el *quark* arriba, u (**up**), abajo, d (**down**), encanto, c (**charm**), extraño, s (**strange**), fondo, b (**bottom**) y cima, t (**top**). Los otros seis fermiones, *sin* carga de color, se dicen **leptones**. Esta propiedad¹⁷ puede tomar tres *valores*: rojo, verde o azul (para los *antiquarks*: antirrojo, antiazul o antiverde). La aritmética de esta carga de color se diseñó para ajustarse a las predicciones de QCD, según la cual habría estados sin carga de color, que se explican cuando se suman los tres «colores» (y también, cualquier suma color-anticolor). Los gluones, que poseen como carga un par color-anticolor, tienen por consiguiente ocho posibilidades de carga de color.

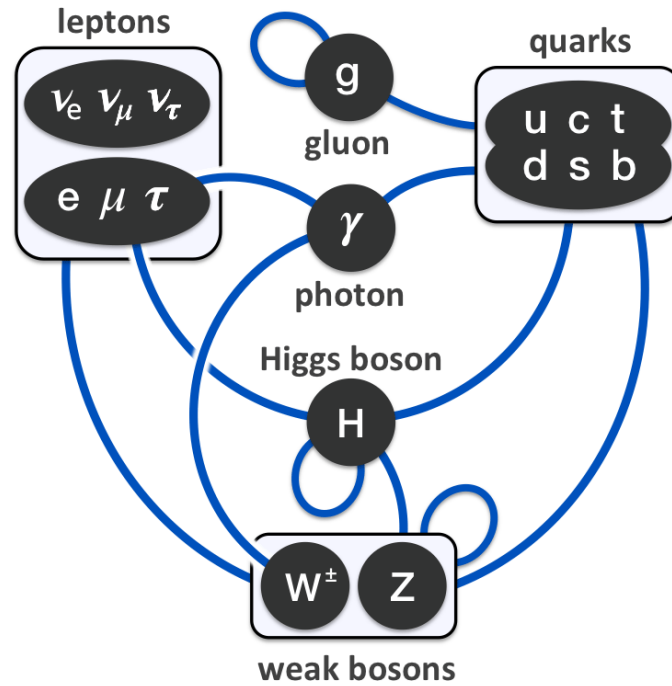


Figura 1.3: [6] Esquema de las partículas del modelo estándar y de los acoplamientos entre ellas.

La **interacción electromagnética** es mediada por el fotón, γ , neutro (eléctricamente) y de masa nula (lo que explica el alcance infinito de la interacción EM). Todos los *quarks*, la mitad de los leptones y los bosones $W^\pm$ poseen carga eléctrica y son susceptibles por ende de sufrir esta interacción. Los leptones

¹⁷No se debe entender por el nombre (carga «de color») que esta característica está relacionada de algún modo con el color visual de los objetos: es mera notación.

que *no* sufren la fuerza EM son los **neutrinos**, ν_e , ν_μ y ν_τ : solo tienen carga débil¹⁸ (lo cual explica en parte que sean tan difíciles de detectar). Los otros leptones, por el contrario, tienen carga débil y eléctrica (negativa para partículas y positiva para sus antipartículas) y son el **electrón**, e , el **muon**, μ y el **tau** τ .

Finalmente, el bosón de Higgs interacciona con todas aquellas partículas con masa (incluyéndose a sí mismo) para darles masa a través del **mecanismo de Higgs**. La figura 1.3 permite ver de forma más clara todos los fermiones y bosones, así como los mediadores de fuerzas y cómo se acoplan entre sí. Cabe destacar que según la formulación actual del SM, los neutrinos *no* tienen masa.

Precisamente, en función a la masa que les da este mecanismo, se puede encontrar otra clasificación por generaciones que se puede resumir en las tres primeras columnas de la figura 1.2. Estas generaciones de leptones y *quarks* son tres y son crecientes en masa. En la primera generación de leptones estarían el electrón y el neutrino electrónico, y en la de *quarks* el *up* y *down*. En la segunda, el muon y su neutrino y los *quarks charm* y *strange*, y en la última el tau con su neutrino y los últimos *quarks*: *bottom* y *top*. Los **hadrones** son agrupaciones de *quarks* que pueden ser **mesones** (dos *quarks*) o **bariones** (tres). Ejemplos de estos últimos son el protón y el neutrón, ambos hechos con *quarks* de la primera generación. Teniendo en cuenta que el electrón también está en la misma, podemos decir que toda la materia usual está constituida por elementos de ella.

1.3. Los límites del modelo estándar

Pese a que el modelo estándar ha conseguido tremendo éxito en sus comprobaciones experimentales durante sus cuatro décadas de existencia, no es capaz de explicar todo aquello que cabría esperar de una teoría científica sobre la composición de la materia y las interacciones fundamentales. Esto no debería ser motivo de sorpresa, puesto que desde que empezamos a hablar del SM una de las interacciones fundamentales, la gravitatoria, se ha obviado. A continuación sigue una pequeña lista de los principales aspectos que el modelo estándar no es capaz de explicar ([7]).

Interacción gravitatoria Sin duda, la gran ausente. La gravedad y su efecto sobre las partículas no se explica en el modelo estándar y esto *per se* ya constituye una carencia, pues todas las partículas con masa la sufren. De forma análoga a con QED se intentó hacer una «QED para la gravedad», i.e., una gravedad cuántica y relativista. Sin embargo, esto no fue posible, debido a la no renormalizabilidad que se encontraron los físicos a la hora de intentarlo; y ello fue el principal problema para incluirla en una descripción conjunta junto a la teoría electrodébil y a QCD.

Masa de los neutrinos El modelo estándar no otorga, en principio, masa a los neutrinos. Sin embargo,

¹⁸I.e., tienen carga eléctrica y de color nulas.

se ha comprobado experimentalmente que, *de facto*, su masa no es nula. Además, se ha explicado cómo la consiguen a través de la llamada oscilación de neutrinos, que también ha encontrado justificación experimental. El SM no encuentra cabida a estas masas.

Materia oscura Los estudios cosmológicos y astrofísicos de la segunda mitad del siglo XX de las curvas de rotación de las galaxias acabaron conduciendo a la necesidad de delimitar un $\approx 20\%$ de la masa-energía del universo para asignársela a la «materia oscura». Se sabe que es materia que no, o casi no interacciona electromagnéticamente, aunque sí gravitatoriamente. El problema que tiene el modelo estándar es que ninguna partícula elemental (o no elemental) se puede corresponder con esa descripción.

Energía oscura La cosmología avanzó mucho durante el siglo XX, tomando como herramienta principal la teoría de la relatividad general. Las pruebas experimentales de la segunda mitad de ese siglo llevadas a cabo con explosiones de supernovas permitieron asegurar que la mejor forma de modelizar el universo a gran escala pasa por darse cuenta de que un $\approx 75\%$ de la masa-energía del universo no es ni materia «usual» (bariónica, constituye el $\approx 5\%$) ni oscura. Este modelo se denomina modelo estándar de cosmología (en clara analogía al SM) o modelo Λ_{CDM} y el nombre que le da a ese 75% es, en un alarde de originalidad, «energía oscura».

La explicación que se le da a la energía oscura es la de la densidad de energía del estado de vacío del universo. Y he aquí la discrepancia del modelo estándar de física de partículas con el de cosmología. Es posible, usando el SM, calcular esa energía de vacío. Sin embargo, a la hora de comparar el resultado del SM con los datos experimentalmente justificados del modelo Λ_{CDM} se comprueba una diferencia **ingente** entre los órdenes de magnitud de los dos valores. En función a las hipótesis que se hagan para el cálculo desde el SM o algunas de sus mejoras, la diferencia está entre los cuarenta o *más de cien* órdenes de magnitud. Esto se ha conocido, por motivos evidentes, como la «peor predicción experimental jamás hecha de la física».

Abundancia de la materia frente a la antimateria El motivo de que la evidente abundancia respecto a la antimateria de la materia hoy en día se puede explicar desde el SM requiriendo violaciones de la simetría CP (i.e., que la acción de un operador que intercambie el signo de las coordenadas espaciales y la carga de una partícula no sea simétrica). Y, dentro del modelo estándar existen fuentes de violación de la simetría CP, pero no las suficientes para explicar la diferencia entre materia y antimateria que se estima que debería haber habido instantes después del Big Bang. Hay extensiones al modelo estándar que incluyen más fuentes de violación de esta simetría que podrían ser la solución.

Las tres generaciones Hemos visto que se pueden agrupar las partículas que el modelo estándar describe en tres familias o generaciones¹⁹, pero, ¿por qué ha de haber tres específicamente? ¿Por qué

¹⁹Los experimentos

no más? El SM no nos dice por qué.

Origen de los valores de los parámetros del SM Entre las masas de las partículas, constantes de acoplamiento, y otros valores, el modelo estándar depende de unos 19 parámetros que han de ser medidos experimentalmente e introducidos en el marco teórico para que el SM funcione. A una parte de la comunidad científica le parece que, en pos de la búsqueda de simplicidad sería adecuado tratar de reducir el número de variables que hay que introducir «a mano» en el modelo, sin olvidarnos de que el SM tampoco explica por qué todos esos parámetros tienen los valores que se miden.

El problema de la jerarquía Existe una gran diferencia en órdenes de magnitud entre las fuerzas de la interacción débil y de la interacción gravitatoria: la primera es muchísimo más fuerte (25 órdenes de magnitud) que la segunda. Sin embargo, si se parte de la base del SM, la constante de Fermi (la constante de acoplamiento de la int. débil) debería ser *similar* a la gravitatoria, (i.e., la constante gravitatoria de Newton, G). Sin embargo, esto no es así y la explicación que se puede dar desde el modelo estándar sería un ajuste **muy** fino de las correcciones cuánticas en el cálculo de estas constantes. Esta solución, que parece ser demasiado *ad hoc*, sugiere otras explicaciones para describir estas diferencias de magnitudes, ya allende el modelo estándar.

Debido a todas estas problemáticas, la comunidad científica ha ido proponiendo distintas formas de, bien mejorar el modelo estándar, o bien sustituirlo por una explicación nueva que solvete, si no todas, parte de las susodichas. Los nuevos marcos que ahondan en una mejoría o superación del modelo estándar se llaman «física más allá del modelo estándar», o BSM (del inglés *Beyond the Standard Model*). Apuntamos brevemente las principales corrientes.

Teorías de supersimetría Esta apuesta se basa en añadir otra simetría al modelo estándar: entre bosones y fermiones. De este modo, todo fermión tendría una partícula *simétrica* que sería un bosón, y análogamente todo bosón podría relacionarse con un fermión. Los nuevos bosones tienen el mismo nombre que sus compañeros fermiones pero añadiéndoles una «s» delante (squark, sleptón, selectrón, sneutrino...), mientras que los fermiones supersimétricos tienen el mismo nombre que sus bosones asociados del modelo estándar pero añadiendo el sufijo «-ino» al mismo (i.e., fotino, gluino, wino, higgsino, etcétera), como se puede apreciar en la figura 1.4.

Según estas teorías de supersimetría, más comúnmente **teorías SUSY** (del inglés *SUPERSymmetry*), de no estar rota tal simetría, haría que existieran partículas exactamente análogas a las actuales, salvo por el espín (p. ej., el fotino sería un fermión sin masa de espín $\frac{1}{2}$). Esto claramente no es así, puesto que ya habrían sido detectadas. Por ello la simetría debe haber sufrido una rotura espontánea de la simetría y la masa de las partículas supersimétricas debe ser mayor que la de sus análogos del SM.

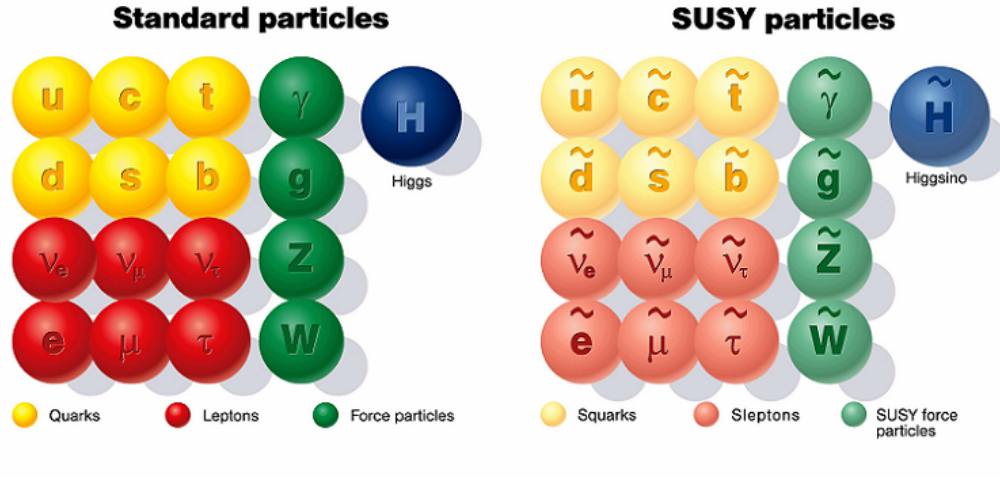


Figura 1.4: [8] Relación de las partículas supersimétricas junto a las análogas del modelo estándar.

Las teorías SUSY arreglan varios problemas de la lista anterior. Entre ellos, ofrece una solución al problema de la jerarquía, además de algún candidato a materia oscura. También permite la reunificación de las constantes de acoplamiento de las tres interacciones del SM a escalas de energía más elevadas. Esto último da soporte a las interpretaciones de muchos científicos según las que, en busca de mayor simplicidad y generalidad, sería posible reunificar estas tres fuerzas en una sola a energías mayores (ver fig. 1.5). Estas teorías se dicen «teorías de gran unificación» o GUT (del inglés *Great Unified Theory*), y el problema bajo el modelo estándar es que parece que la evolución esperada de esas constantes de acoplamiento con la energía no permitiría tal reunificación. Sin embargo, modelos como SUSY abren la puerta a que así sea.

Estos motivos hacen que la comunidad de física de partículas hoy en día dedique buena parte de sus esfuerzos a buscar partículas SUSY, o alguna prueba experimental que sustente estas hipótesis. En pos de ello, se han creado muy numerosos modelos SUSY, que se diferencian esencialmente en las hipótesis de las que se parte. Si bien el modelo estándar tiene unos 19 parámetros libres, los modelos SUSY más generales (MSSM, *Minimal SuperSymmetry Model*) alcanzan los 124 parámetros a introducir.

Teoría de cuerdas Es esta una apuesta distinta a SUSY, puesto que no aspira solo a mejorar el modelo estándar, sino a construir un nuevo marco teórico que supere al SM y lo englobe dentro de sí, incluyendo a la gravedad también. Con raíces en la mitad del siglo XX, la teoría de cuerdas parte de la base de que las partículas (cero-dimensionales) que el SM tilda de elementales, no son tales, sino que los componentes últimos de la materia serían *cuerdas* unidimensionales que se pueden «conectar» con *d-branas* de dos o más dimensiones. Las partículas del modelo estándar se podrían recuperar como un determinado modo de vibración de estas cuerdas, así como la descripción de la relatividad general en el límite de bajas energías.

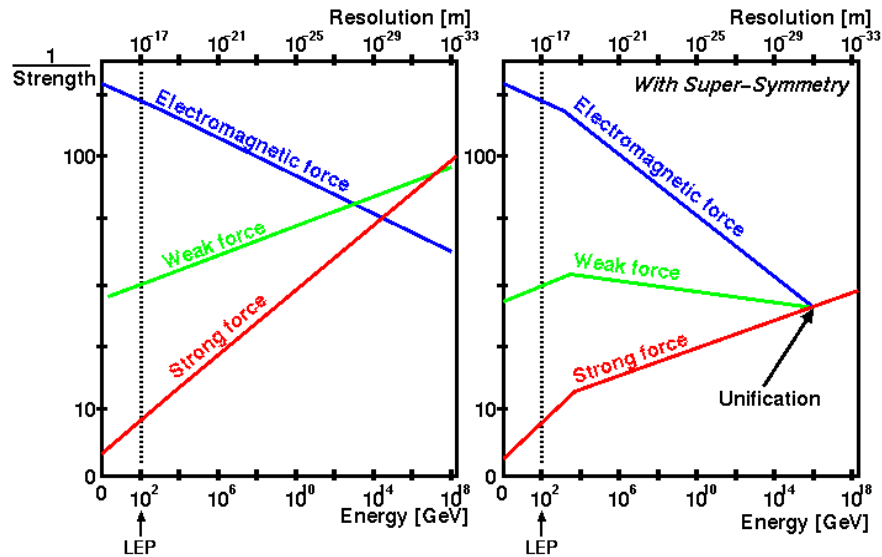


Figura 1.5: [9] Diagrama que muestra la evolución con la energía del inverso de las constantes de acoplo de las interacciones electromagnética, fuerte y débil. A la derecha se muestra la comparación con una teoría SUSY.

Uno de los iconos de esta teoría es que recurre a más dimensiones (en el menor de los casos a unas nueve o diez dimensiones espaciales) para construirse. Las ventajas de la teoría de cuerdas es que incluye dentro de sí una gravedad que es relativista (i.e., relatividad general) y también cuántica. La teoría de cuerdas también incluye dentro de sí la simetría de fermiones y bosones de las teorías SUSY, y sería una buena candidata, no solo a GUT (i.e., a una explicación unificada de las interacciones débil, fuerte y EM) sino a «teoría del todo» o TOE (del ing. *Theory Of Everything*).

Otras Existen más ampliaciones del SM aparte de SUSY y teorías de cuerdas, como otras vías para cuantizar la gravedad que podríamos mencionar aquí: es el caso de las teorías efectivas de campos (EFT de *Effective Field Theory*) o las teorías de gravedad cuántica de bucles (*loop quantum gravity theories*); o también los axiones: partículas elementales teorizadas para solventar la violación clara de la simetría CP en QCD. Hay muchas otras alternativas para completar el modelo estándar o superarlo, pero no entraremos en detalles por no ser este el fin de este texto.

Como vemos, el modelo estándar constituye la frontera de la física en lo que se refiere a un modelo científico experimentalmente probado en física de partículas y de altas energías. Es un marco teórico que ha traído muchos avances en la comprensión de la estructura interna de la materia y ha avanzado en la aglutinación y progresiva unión de las teorías científicas (QED, QCD, interacción débil) todo ello sustentándose en unas contundentes pruebas experimentales. Hoy en día los esfuerzos de los físicos de partículas tratan de ampliar este modelo para suplir sus carencias, por una parte poniendo a prueba estos nuevos marcos teóricos y, por otra, comprobando que las predicciones del modelo estándar se siguen cumpliendo según avanzan los umbrales de energía disponibles para los experimentos.

El proceso de física de partículas que se estudia en este texto, conocido como proceso $t\bar{t}H$, se enmarca en principio en un análisis del segundo tipo. En la siguiente sección explicaremos en más detalle en qué consiste y la motivación del análisis del mismo.

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

1.4.1. Descripción general

El proceso físico conocido sencillamente por $t\bar{t}H$ parte de dos gluones que decaen a dos pares respectivos de *quarks top-antitop*, donde un *quark top* proveniente de uno de los gluones, y uno *antitop* proveniente de otro de los gluones interaccionan para dar lugar a un bosón de Higgs, H . Esto viene recogido en el diagrama de Feynmann de la figura 1.6.

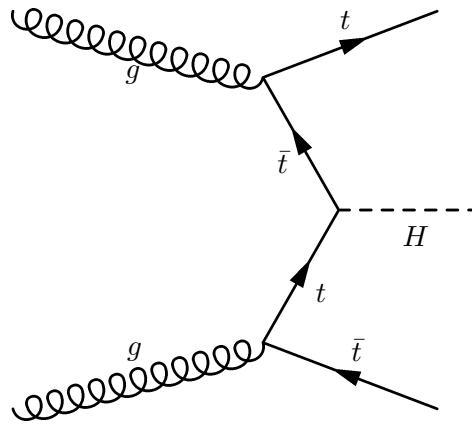


Figura 1.6: Diagrama de Feynmann representando el proceso $t\bar{t}H$.

Debido a las distintas formas de desintegración del bosón de Higgs, a las que se añade las del par t y $\bar{t}$ ²⁰, el número de estados finales es muy grande. Dentro de la colaboración del detector CMS (de donde se han obtenido los datos) del CERN se ha decidido estudiar este proceso en varios canales ([11]), en función a la desintegración del bosón de Higgs.

- $t\bar{t}H$ con $H \rightarrow \gamma\gamma$.
- $t\bar{t}H$ con $H \rightarrow b\bar{b}$.
- $t\bar{t}H$ con $H \rightarrow \tau\bar{\tau}$.²¹
- $t\bar{t}H$ multileptónico.

Cada canal tiene sus peculiaridades, en virtud a los estados finales que puede alcanzar, aunque no entraremos en detalles de los mismos, salvo del último: pues es ese el canal donde se han hecho análisis en este trabajo.

²⁰Sobre todo, las de los bosones $W^\pm$ a los que, junto a un *quark b*, el *top* decae en un 93.2% ([10])

²¹En rigor, se estudia $H \rightarrow \tau\bar{\tau}$ donde alguno de los dos τ decae hadrónicamente.

1.4.2. El canal multileptónico

El **canal multileptónico** está formado por los estados finales en los que se encuentran dos o más leptones (de ahí su nombre), de los cuales alguno procede de la desintegración semileptónica (a través de un bosón $W^\pm$) de uno de los dos *quarks top* que siempre hay en el proceso. Ningún estado final multileptónico pertenecen a los otros canales (pues estos entre sí son «ortogonales», en el sentido de que no comparten estados finales). Si bien comparativamente con (p. ej.) los estados finales del canal $t\bar{t}H$ con $H \rightarrow b\bar{b}$, las fracciones de desintegración (*branching ratios*, BR) son bastante bajas²², la ventaja del análisis multileptónico de este proceso radica en que:

- Necesariamente los eventos que seleccionemos estarán primordialmente definidos por muones y electrones²³ y dependerán menos de otros observables, como *jets*, que se miden con menor precisión que los primeros.
- Como requisito para seleccionar eventos con dos leptones, se exige que los dos leptones tengan carga del mismo signo, cosa en que en el proceso $t\bar{t}H$ sí se puede dar (ver fig 1.7). Este criterio de selección es muy restrictivo (debido a que no todos los procesos tienen estados finales así) y permite eliminar muchos otros procesos físicos similares que no nos interesan (que son el fondo de nuestros datos).

Pese a estas ventajas, no es ni mucho menos uno de los procesos que más abundan en las colisiones p-p del LHC actualmente. Su sección eficaz es de $\sigma \simeq 0.5 \text{ pb}$ ([12]), pequeña, en comparación con la del proceso $t\bar{t}$ ($\sim 8 \cdot 10^2 \text{ pb}$), y diminuta, en comparación con la del proceso $b\bar{b}$ ($\sim 10^{11} \text{ pb}$), por poner algún ejemplo; un gráfico comparativo de las secciones eficaces a distintas energías en el acelerador LHC se muestra en la figura 1.8. Esto hace que tengamos que buscar precisamente una señal lo más clara y nítida posible, y que sea esencial entender los fondos (esto es, procesos con estados finales iguales o similares a los nuestros que afectan a nuestro análisis).

Los principales fondos de $t\bar{t}H$ multileptónico son los procesos $t\bar{t}$, $t\bar{t}W$ y $t\bar{t}Z$. Los segundo y tercer fondos se consideran irreducibles. Esto se debe a que son procesos tan similares (con un par $t\bar{t}$ y otro bosón) que comparten estados finales con $t\bar{t}H$, y por tanto nuestros criterios de selección se ajustarían también a algunos de sus estados finales. En cambio, el primero sí es reducible: es uno de los fondos más importantes debido a que se detectan muchas veces falsos pares de leptones con misma carga (leptones *fake*). Esto se puede deber a una medición mala de la carga de las partículas, y debido a cómo son los estados finales de $t\bar{t}$ (que pueden dar pares de leptones con carga diferente), así como a la mayor sección eficaz de este proceso, esto ocurre con mucha frecuencia.

²²Como comparación, se estima que un bosón de Higgs ($m_H = 125 \text{ GeV}$) decae en un 60 % a $b\bar{b}$ ([10]), y del 40 % restante hay aún que sustraer los estados finales que no entran en el canal multileptónico.

²³Es común referirse por «leptones» en jerga de física de partículas experimental solo a muones y electrones, debido a las desintegraciones del leptón tau, que hacen su identificación más compleja que en los casos de sus otros dos compañeros.

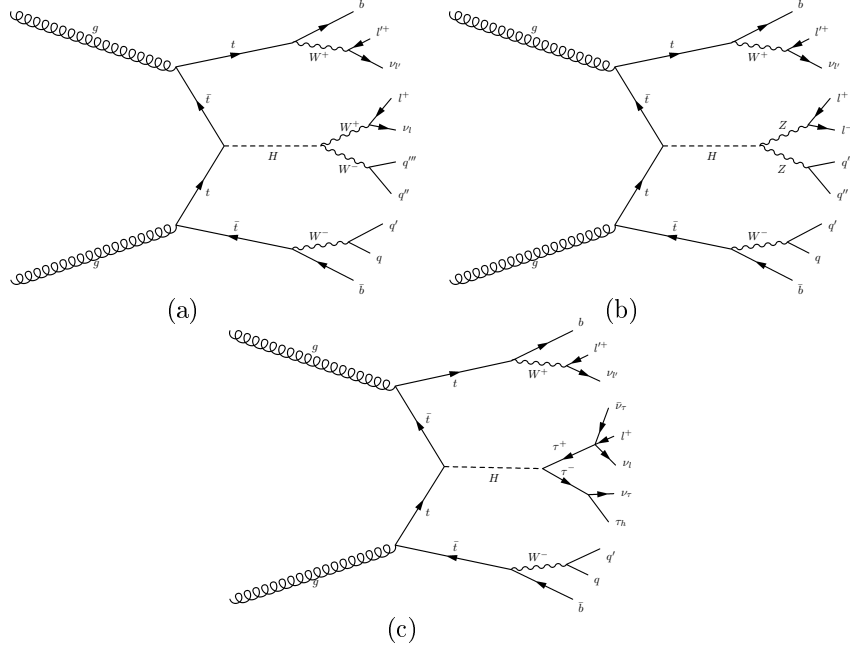


Figura 1.7: Diagramas de Feynman representando un proceso $t\bar{t}H$ con distintos estados finales. Tanto los diagramas (a) como (c) muestran estados finales con el indicador de dos leptones con igual carga, aunque solo el (a) entraría dentro del análisis multileptónico (puesto que el (c) tiene un leptón tau decayendo hadrónicamente). El diagrama (b) muestra un estado final con tres leptones que también pertenecería al análisis multileptónico, de acuerdo con la clasificación antes hecha.

1.4.3. Por qué estudiar $t\bar{t}H$

El proceso $t\bar{t}H$ ofrece una forma **directa** de medir la constante de acoplo de Yukawa entre el *quark top* y el bosón de Higgs, la cual tiene importancia en tanto en cuanto el *top* es el más masivo de todos los *quarks* (y a través del mecanismo de Higgs, le debe su masa al bosón homónimo). Esta constante aparece en las ecuaciones que describen el proceso, concretamente, en lo relativo a la interacción en la que el par $t\bar{t}$ resultante de la desintegración de los dos gluones da pie a un bosón de Higgs.

Así mismo, y como este proceso se estudia en el límite de energías conocido en física de partículas en aceleradores (13 TeV), toda desviación de las predicciones del modelo estándar podría ser un indicador de «nueva física». Teniendo en cuenta las faltas del modelo estándar, y las esperanzas de que a escalas mayores de energía se pudieran hallar indicios de un sustituto o mejoría de este, siempre resulta útil comprobar la física esperada del SM cuando se alcanzan estas nuevas cotas de energía.

1.4.4. Situación actual de la investigación

El estudio del proceso $t\bar{t}H$ se enmarca dentro de la física del Higgs. Esta partícula fue descubierta con una significancia mayor que 5σ sobre todo gracias a la información de los canales $H \rightarrow \gamma\gamma$ y $H \rightarrow ZZ$ (y en menor medida $H \rightarrow WW$) en el año 2012, y de manera independiente por las colaboraciones de los experimentos CMS y ATLAS. La confirmación de la existencia del bosón de Higgs confirmó la justificación de las masas a través de la interacción de Yukawa de todas las partículas del

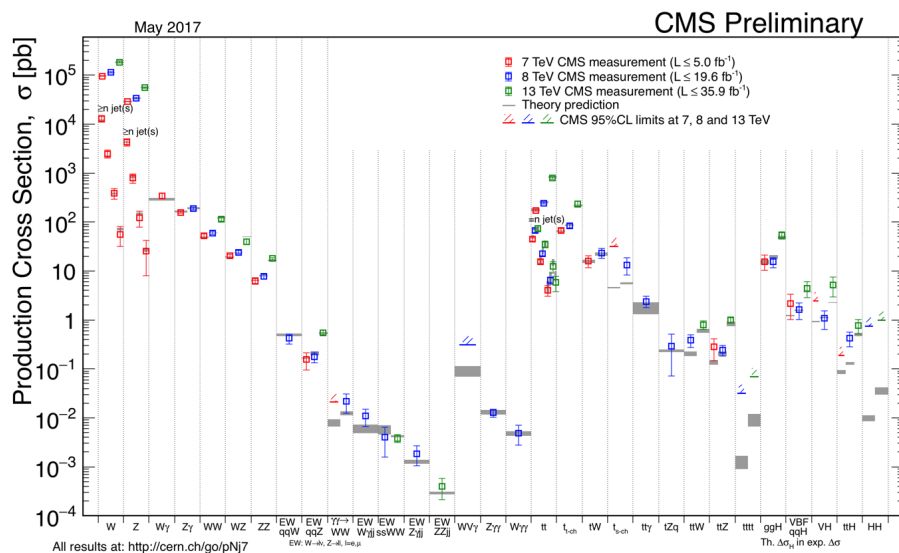


Figura 1.8: [13] Gráfica que muestra distintas secciones eficaces de numerosos procesos a distintas energías en centro de masas en el acelerador de partículas LHC del CERN.

modelo estándar. Su masa se ha medido como $m_H = 125.09 \pm 0.24 \text{ GeV}$ ([10]).

Respecto al proceso que nos atañe, pese a que su estudio se ha dado a lo largo de varios años, día de hoy no se ha observado todavía. Por «observar el proceso $t\bar{t}H$ » nos referimos a que ningún análisis ha conseguido, estudiando los datos obtenidos en experimentos (en nuestro caso, el CMS en el acelerador LHC (CERN)), calcular la sección eficaz experimental del proceso con suficiente nitidez. Los esfuerzos se han orientado, pues, a la búsqueda de límites superiores en la fuerza de señal (*signal strength*) que han ido mejorando según se hacían experimentos con energía en centro de masa mayor²⁴ y mayor luminosidad.

Los últimos resultados del experimento CMS ([1]) de marzo de este año apuntan a un límite asintótico superior con un intervalo de confianza del 95 % en *signal strength* de $\mu = 2.5$ (observado) y $\mu = 0.8^{+0.3}_{-0.2}$ (esperado, incertidumbres de 1σ). Estos valores se corresponden con una significancia observada de 3.4σ y una esperada de 2.4σ , de donde no se deduce la «observación del proceso». ²⁵ En el futuro, se espera que con mayor cantidad de datos se pueda llegar a observar el proceso y a obtener resultados de él con mayor precisión.

²⁴La sección eficaz de $t\bar{t}H$ aumentó un factor 4 al hacer el paso de $\sqrt{s} = 8$ TeV a 13 TeV, mientras que los fondos escalaron por un factor 3, lo que permitió mejorar la relación entre señal y fondo.

²⁵En la sección 3.3 de resultados detallaremos los significados de estas magnitudes y cómo se obtienen.

2 La instalación LHC y el experimento CMS

El poder que manejamos aquí es inconmensurable.

(El teniente comandante Orson Krennic, en *Rogue One: una historia de Star Wars*, DISNEY/LUCASARTS)

EL 16 de diciembre de 1994 ([14]) el Consejo del CERN aprobó consensuadamente la construcción de un colisionador de partículas protón-protón (también plomo-protón y plomo-plomo) en el túnel del viejo LEP. El colisionador, llamado en inglés *Large Hadron Collider* (Gran Colisionador de Hadrones) o simplemente LHC, prometía conseguir energías en centro de masas de $\sqrt{s} = 14 \text{ TeV}$, nunca vistas hasta la fecha. Obtener datos experimentales bajo tales condiciones permitirían avances significativos en el campo de la física de partículas, como el posible descubrimiento del bosón de Higgs. El año siguiente, se aprobó un presupuesto para su construcción del orden de $\sim 3000 \text{ M€}$. Las obras terminaron una década después, concretamente en 2008, cuando se dieron por finalizadas.

Todos los datos que han sido empleados en este análisis fueron recogidos en el experimento CMS que está ubicado en el LHC, de ahí que dediquemos este capítulo a describir, sin entrar en detalle, tanto el colisionador como el detector.

2.1. El LHC

2.1.1. Descripción general

([15], [16]) El Gran Colisionador de Hadrones tiene forma circular, con cerca de 27 km de longitud y está ubicado cerca de la ciudad de Ginebra, entre Francia y Suiza: justo donde el CERN tiene sus instalaciones y se encuentra levemente inclinado (respecto al suelo, no respecto a la tangencial respecto al centro del planeta). Si bien de media se encuentra a unos 100 m bajo la superficie, esta profundidad oscila en función a qué parte del anillo nos refiramos. El hecho de estar soterrado proporciona protección frente a parte de la radiación que sufriría de estar al aire libre.

La localización no es casual, puesto que se encuentra dentro del mismo túnel donde ya estaba el antiguo colisionador electrón-positrón LEP (*Large Electron Positron Collider*) y, el hecho de estar en las cercanías del CERN le permite utilizar el conjunto de aceleradores anteriores que allí hay (véanse el SPS,

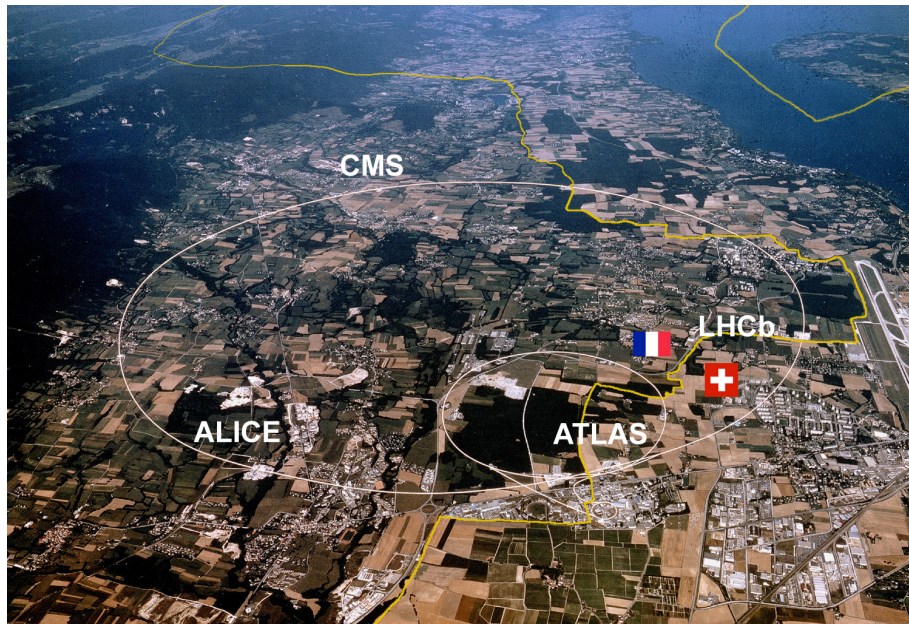


Figura 2.1: [17] Mapa de la región cercana a Ginebra sobre el que se ha superpuesto un diagrama del LHC.

o el PS) para inyectar los cationes que se quieran colisionar con más energía que sin ellos. Los cationes inyectados viajan a través de dos tubos en sentidos opuestos, y estos se cruzan en cuatro localizaciones exactas (véase 2.1), donde se dan las colisiones entre cationes. Y donde, también, se ubican los cuatro experimentos principales: CMS (*Compact Muon Solenoid*), ATLAS (*A Thoroidal LHC Apparatus*), ALICE (*A Large Ion Collider Experiment*) y LHCb (*LHC-beauty*). Los dos primeros (CMS y ATLAS) tienen carácter «general» y permiten estudiar una gran variedad de procesos físicos, mientras que LHCb y ALICE se dedican a análisis más específicos (lo cual se muestra, de hecho, en su estructura).

Como ya avanzamos, está diseñado para (eventualmente) trabajar con energías en centro de masas de unos 14 TeV en colisión protón-protón (7 TeV cada uno), pese a que en sus comienzos no se empezase con tanta. Se puso en marcha por primera vez el 10 de septiembre de 2008, aunque la alegría no llegó mucho. El 19 de septiembre, nueve días después, hubo un problema con los imanes superconductores que mantienen los iones en la trayectoria circular: un fallo eléctrico provocó la pérdida del gas helio (que se usa para enfriar los dipolos) y que acabó afectando a 57 imanes. Durante el resto de 2008 y gran parte de 2009 se realizaron estudios sobre el incidente y se reparó la instalación.

Sería el 20 de noviembre de 2009 cuando los haces de partículas volviesen al LHC, dando comienzo al primer *run*. Diez días después la energía de los mismos se aumentó hasta los 1.18 TeV (por haz), superando el récord previamente establecido por el Tevatron de Fermilab ([18]) que estaba en 0.98 TeV. Tras esto, se fue aumentando la energía por haz hasta llegar a 3.5 TeV el 30 de marzo de 2010. Se hizo una pausa el 4 de noviembre de 2010, tras el cual se introdujeron núcleos de plomo durante un mes en el LHC. Más tarde, en febrero de 2011 se retomó la actividad p-p, y tras otra parada, en 2012 se

aumentó la energía por haz hasta los 4 TeV. Sería ese año cuando se anunciase el descubrimiento de una partícula «compatible» con el bosón de Higgs, uno de los más importantes hitos del LHC. El primer *run* se completaría el 13 de febrero de 2013.

Desde ese momento el LHC se sometió a mejoras que le permitieron reiniciar su actividad (dando comienzo al segundo *run*) el 5 de abril de 2015 con 6.5 TeV de energía por haz de hadrones, permitiendo energías en centro de masas de 13 TeV. Este *run* es en el que nos encontramos todavía, y durará hasta diciembre de 2018.

2.1.2. Funcionamiento

2.1.2.1. Encendido

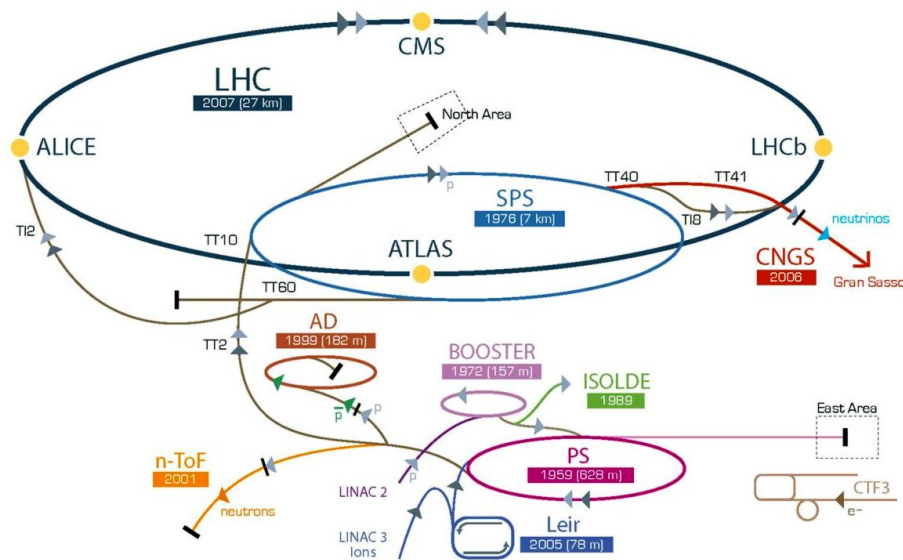


Figura 2.2: [19] Diagrama de los distintos aceleradores que sirven como inyectores de haces de hadrones al LHC.

El inicio de la actividad en el LHC comienza obteniendo ora núcleos de hidrógeno de una botella común de gas H_2 , ora núcleos de plomo de una muestra muy pura de este metal. Los protones (que son los que nos interesan) son inyectados en el acelerador lineal LINAC 2 que los introduce en el sincrotrón PSB¹ con 50 MeV. Allí se aumenta la energía del haz hasta los 1.4 GeV, momento en el cual pasan al PS². Este sincrotrón acelera los hadrones hasta los 26 GeV, cuando son transferidos al SPS³. Finalmente se introducen en los dos tubos del LHC en sentidos opuestos tras haber alcanzado los 450 GeV: será en el propio LHC cuando se aceleren hasta la energía final, que hoy en día es de 6.5 TeV por haz (ver fig. 2.2). Para el caso de los núcleos de plomo, la cadencia es similar, aunque partiendo del LINAC 3 y sin pasar por el PS: estos haces podrán llegar a ser acelerados a 1.38 TeV.

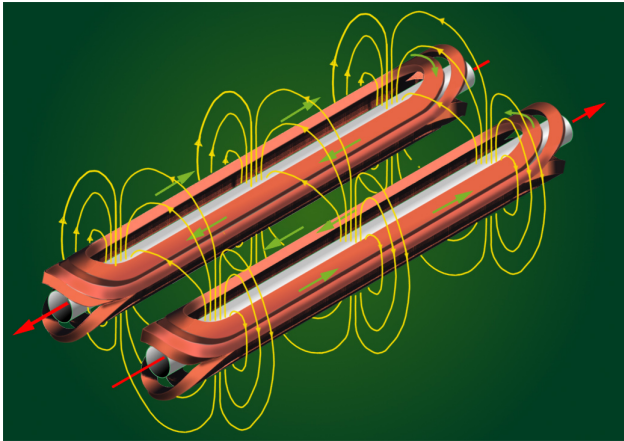
¹ Proton Synchrotron Booster.

² Proton Synchrotron.

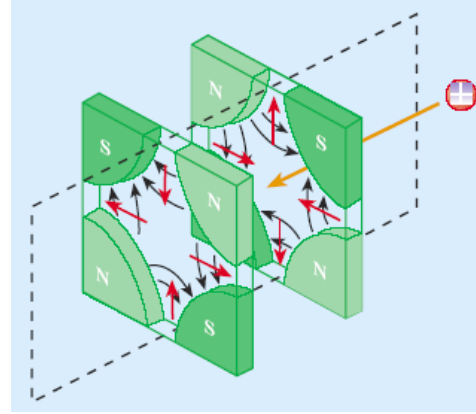
³ Super Proton Synchrotron.

Tras la puesta en funcionamiento del LHC, los llamados haces estables (*stable beams*) de protones están formados por grupos o *puñados* (*bunches*) de unos 10^7 hadrones. Actualmente, la separación temporal de cada *bunch* de partículas es de unos 25 ns.

2.1.2.2. Componentes esenciales



(a) [20] Representación del campo magnético que crean los dipolos superconductores en los dos tubos del LHC.



(b) [21] Representación del campo magnético que crean los cuadrupolos magnéticos y que sufren las partículas cargadas que viajan por los tubos del LHC.

Figura 2.3: Imágenes de componentes magnéticos del LHC.

Los hadrones se mantienen dentro del LHC a través de un sistema de millares de imanes distribuidos a lo largo de la circunferencia del colisionador, de los cuales los 1232 **dipolos superconductores** (fig. 2.3) son los más destacables. Son ellos los que se encargan de que las partículas no pierdan su trayectoria circular. Enfriados con helio líquido a unos 1.9 K, son capaces de operar en régimen semiconductor a 8.33 T (aunque podrían llegar a ofrecer 9 T).

Al igual que estos sistemas de dipolos, existen otros imanes, **multipolos magnéticos**, que se encargan de «concentrar» o «colimar» los haces de partículas. Esto permite que las dimensiones de cada *bunch* estén controladas dentro de unos límites, por lo que en los cuatro puntos de colisión habrá un mayor número de choques de partículas (fig. 2.3). Los motivos del ensanchamiento de estos *bunches* varían desde la interacción gravitatoria de los protones, las electromagnéticas entre distintos *bunches* etcétera. Son empleados particularmente cerca de los puntos de colisión de los haces, y pueden tener desde cuatro (cuadrupolos) hasta ocho imanes (octopolos) pasando por seis (sextupolos).

Habíamos dicho ya que los protones y núcleos de plomo no llegaban al LHC acelerados al máximo de lo que la instalación soporta, sino que dentro de la máquina debían todavía alcanzar mayores energías. La responsabilidad de tal menester recae sobre las ocho **cavidades de radiofrecuencia** que cada uno de

los dos tubos del LHC tienen. Estas, conformadas por imanes superconductores que trabajan a 4.5 K, inducen un campo eléctrico de corriente alterna en el tubo por donde viajan los hadrones de modo que fuerzan a las partículas que la atraviesan a *sincronizarse* con la frecuencia del campo, que es de 400 MHz. Si las partículas no están exactamente sincronizadas con ella, sufren una fuerza que bien las frenaría o las aceleraría para que así lo estuvieran. De este modo, se consiguen dos cosas: por un lado, acelerar los haces de partículas hasta la energía final (pues la frecuencia de los campos se cambia con el tiempo) y, por otro, concentrar y separar esos *bunches* de partículas espacialmente (lo cual repercute, como dijimos antes, en el número de colisiones final). Además, la sincronización es crucial para que los haces de cada tubo se crucen exactamente en el momento preciso en los cuatro detectores del colisionador.

2.1.2.3. Luminosidad, *pile-up* y *trigger*

La forma en la que se mide cuántas colisiones se dan en un acelerador es a través de la magnitud física llamada **luminosidad**. Esta se define como el número de eventos (colisiones) detectados por unidad de tiempo y de superficie, usándose como unidades comúnmente $[L] = \text{b}^{-1} \text{s}^{-1} = \text{Hz b}^{-1}$, o múltiplos de estas. La forma más sencilla (y semicualitativa) de definirla para nuestro acelerador es

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

donde N_1 y N_2 son el número de partículas de cada *bunch* que colisiona en uno de los detectores, f es la frecuencia con la que los *bunches* viajan por el LHC (en la actualidad, $40 \cdot 10^6$ Hz) y σ_x y σ_y son el ancho y alto del *bunch* en los puntos de colisión. Una definición más general (de donde se puede deducir la anterior expresión) sería

$$L := \frac{1}{\sigma} \frac{dN}{dt}, \quad (2.2)$$

donde σ es la sección eficaz de la colisión protón-protón y $\frac{dN}{dt}$ la rapidez con la que se produce una colisión.

Ahora mismo, la luminosidad que se da en el LHC es de $L = 10^{-2} \text{fb}^{-1} \text{s}^{-1}$. Una de las ventajas de esta magnitud es que podemos obtener una medida del número total de colisiones que han tenido lugar en el LHC en un periodo determinado de tiempo. Para ello se define la **luminosidad integrada** (en contraposición a la luminosidad, o luminosidad *instantánea* anterior) como

$$\mathcal{L} := \int_{t_0}^{t_1} L(t) dt, \quad (2.3)$$

esto es, simplemente, la acumulación de la luminosidad instantánea en virtud del tiempo.

La luminosidad *nominal* es la que el equipo que controla y maneja el LHC ofrece a los experimentos

que en él están. Estos, así mismo, tienen capacidad de medir la luminosidad que ellos observan, que se aproxima a la que ofrece el LHC, pero no es exactamente la misma debido esencialmente a los tiempos muertos en el que el detector no recoge información, lo cual ocurre constantemente. Esto se puede ver claramente en la gráfica 2.4, donde se observa la luminosidad integrada a lo largo del periodo de funcionamiento del LHC en el año 2016 que ha sido medida por CMS y la que el LHC ofrece (en colisiones p-p).

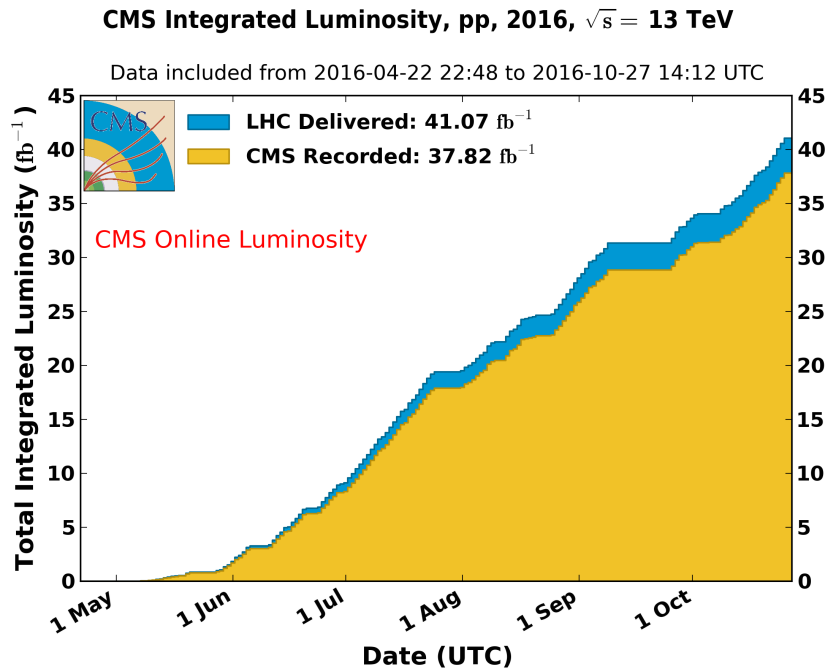


Figura 2.4: [22] Luminosidad integrada ofrecida por el equipo del LHC y registrada por CMS durante el periodo de trabajo con colisiones protón-protón en 2016.

La luminosidad que es capaz de ofrecer el LHC es tal que por cada segundo y cada centímetro cuadrado, se dan del orden de 10^{34} colisiones en un detector, produciéndose al mismo tiempo un sinfín de procesos físicos entre partículas de cada uno de los *bunches*, así como entre los constituyentes de estas. Los detectores, colocados en los sitios de colisión por un motivo trivial, cuando registran un evento observan inevitablemente los resultados de muchas otras colisiones en un lapso de tiempo muy corto. A esta «contaminación» que se da entre colisiones se la denomina *pile-up* («apilamiento»), PU.

Otra desafortunada consecuencia de la alta luminosidad del LHC es que es imposible que los detectores puedan registrar todas las colisiones que se dan. El motivo es que no existe ninguna tecnología que sea capaz de guardar tanta información en tan poco tiempo. Teniendo en cuenta que la frecuencia de colisiones es de 40 MHz, y que los datos de cada suceso ocupan entre 1 MB y 1.5 MB, la velocidad a la que debería guardarse la información serían 40 TBs^{-1} . Además, hay que tener en cuenta que no todos esos datos nos interesan: en general, buscamos información de uno, o algunos procesos físicos determinados.

Lo que se hace, inevitablemente, es seleccionar los sucesos cuya información será guardada. Con este fin,

existe un conjunto de herramientas tanto a nivel de *hardware* como de *software* que se *disparan* cuando se detecta un suceso que podría ser interesante y que luego, si cumple unas determinadas condiciones, es guardado. Estas herramientas y procesos se conocen como **trigger**⁴, y son específicos de cada detector. Sin embargo, todos siguen unas líneas generales: a continuación describimos, basándonos en el *trigger* de CMS, en qué consisten ([23]).

El primer nivel (L1, *Level 1*) del *trigger* consiste en una selección que se hace a nivel de *hardware*, de circuitos electrónicos y procesadores dedicados embebidos, con el fin de que sea lo más rápida posible. Se utilizan fragmentos sueltos de la información total de un evento, y la velocidad a la que sale la información de este nivel es del orden de 10^5 GB s^{-1} .

La segunda gran etapa (aunque algunos *triggers* añaden otras por el medio) se denomina *High Level Trigger* (HLT) y se caracteriza por estar basado en *hardware* y *software* convencional (comercial), y tiene lugar en servidores de computación en línea (lo que se denomina comúnmente una «granja en línea» de ordenadores). En este nivel del *trigger* se hacen selecciones más complejas, basadas en toda la información del evento, que reducen la velocidad de la información al orden de 10^2 MB s^{-1} , siendo ya asumible su almacenaje.

En esta última labor colabora la red (*grid*) de computación del CERN, la conocida como *LHC Computing Grid* o LCG (anteriormente, WLCG). Conformada por centros de computación distribuidos por todo el planeta, está dividida en varios segmentos o niveles (*tiers*). El *Tier-0* son los servidores de computación del CERN que se encuentran operativos 24 horas al día, todos los días de la semana. A este régimen de disponibilidad también están sujetos los centros de computación del *Tier-1*, distribuidos por todo el planeta, e interconectados por conexiones dedicadas de internet de muy alta capacidad. Este bloque se usa para computación y para salvaguarda de copias de seguridad de archivos. Finalmente, los *Tier-2* y *Tier-3* consisten en servidores de computación ubicados en departamentos de universidades o institutos de investigación, dedicados sobre todo a análisis específicos y accesibles por parte de la comunidad científica que colabora con el LHC, si bien también se hacen copias de seguridad en los *Tier-2* de ficheros.

El *run 2* del LHC continuará hasta finales de 2018, cuando se iniciará un proceso de mejora de la instalación de varios años conocido como HL-LHC (*High Luminosity LHC*, [24]). Este proyecto mejorará mucho la luminosidad que la máquina podrá ofrecer.

⁴Podría traducirse por «disparador», «activador», «desencadenador».

2.2. El detector CMS

2.2.1. Descripción general

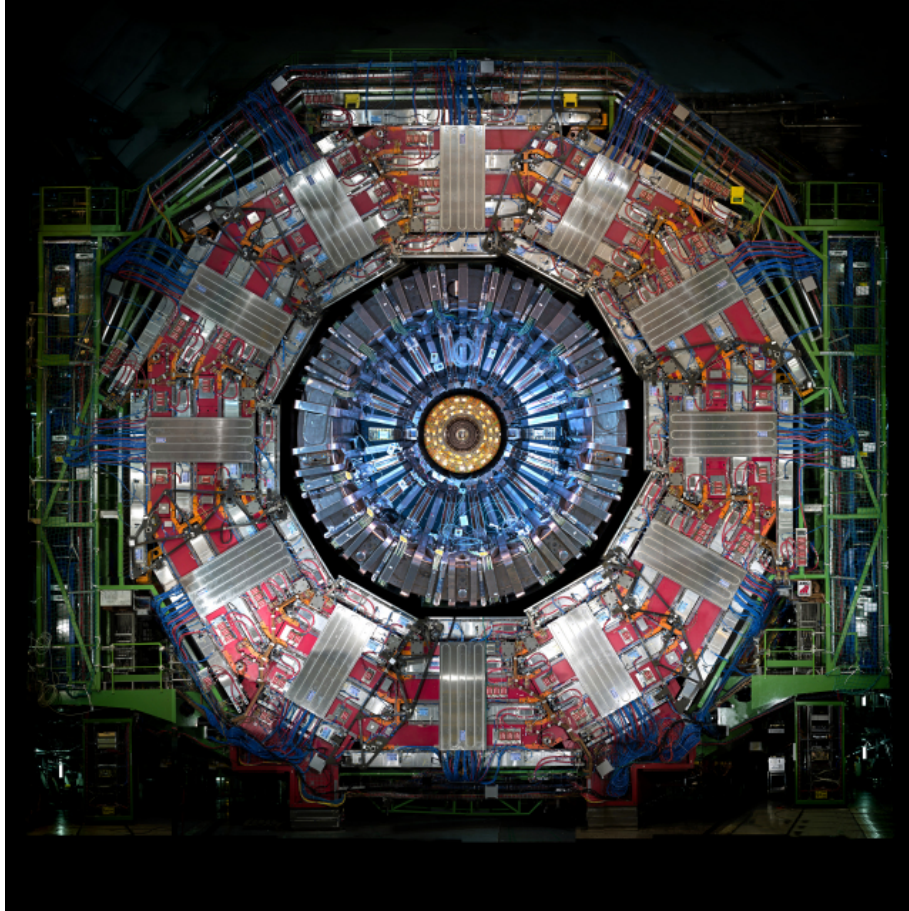


Figura 2.5: [25] Fotografía de una sección transversal del barril del detector CMS, donde se aprecian las sucesivas capas de componentes.

([26], [27]) El «Solenoides de Muones Compacto» (*Compact Muon Solenoid*) o CMS, es un experimento ubicado en uno de los cuatro puntos de colisión que hay en el LHC (fig. 2.5). Dimensionalmente es grande y tiene forma cilíndrica (atravesándolo, justo por el centro de su eje, los dos tubos del LHC). Mide 21.6 m de largo, tiene un diámetro de unos 15 m y está conformado de numerosos subdetectores y componentes de los cuales que se organizan en capas cilíndricas alrededor de los tubos del LHC, salvo en la «base» y la «tapa» del cilindro, donde la disposición es algo diferente. Debido a esto, se suele dividir estructuralmente el detector en *barril* y *tapas* (véase la figura 2.6).

Como ya comentamos, este detector es uno de los dos experimentos de propósito general que el LHC posee, junto a ATLAS. A la hora de diseñarlo, se buscó un aparato que detectase y midiese muones con gran eficacia, electrones y fotones con gran resolución, un buen detector de las trayectorias y un calorímetro hadrónico hermético, que evitase que los hadrones no detectados «escapasen» de él. Con estos objetivos en mente, el resultado fue un experimento subdividido en un detector de trazas, un calorímetro electromagnético, un calorímetro hadrónico, un solenoide superconductor que ofreciese un

potente campo magnético y un sistema de muones muy preciso.

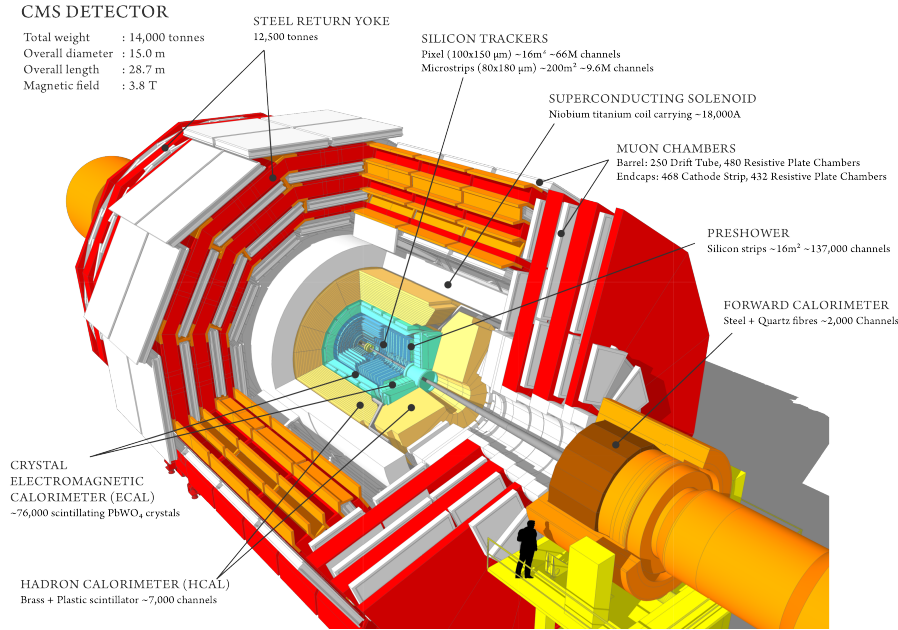


Figura 2.6: [28] Diagrama del detector CMS y sus subdetectores y componentes.

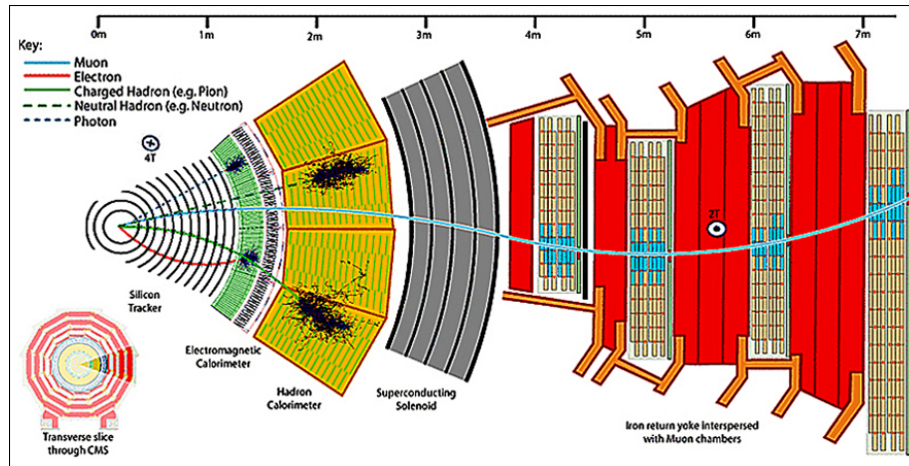


Figura 2.7: [29] Diagrama de un sector circular del detector, en la que se ve cómo distintas partículas atraviesan el mismo.

Sistema de coordenadas

Estos componentes los explicaremos con más detalle más adelante, pero antes debemos hacer mención al sistema de coordenadas que se emplea para analizar los datos. Debido a la estructura de los detectores, se redefinen las coordenadas del espacio euclídeo común ($\mathbf{R}^3$) (x, y, z) partiendo de las coordenadas cilíndricas. El eje z coincide con la recta tangente a la circunferencia del CMS en el punto donde se producen las colisiones (donde se ubica el origen) y, por consiguiente, es en el plano xy donde las secciones del barril del CMS viven. Muchas variables se definen en función a este plano transversal, como el momento transversal ($\vec{p}_T$). Se define el semieje positivo del eje z como aquel que apunta hacia

las montañas de Jura (en Suiza).

Pasando a coordenadas polares, podemos expresar cualquier punto como (r, θ, z) , donde $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, y $r \geq 0$. Se define la pseudorapidez como $\eta := -\ln \tan(\theta)$ y es una de las variables que se emplean. Considerando también el ángulo azimutal $\varphi \in [0, 2\pi]$, es fácil ver que todas las trayectorias rectilíneas que parten del origen se pueden caracterizar por un par (η, ϕ) . Para concretar un punto de cada una de esas trayectorias, se debe usar otra variable, la distancia r al centro.

Ahora describiremos brevemente los subdetectores que componen CMS. Comenzaremos por las capas más internas, cercanas al punto de colisión, y nos iremos alejando hacia el exterior radialmente.

2.2.2. Detector de trazas (*tracker*)

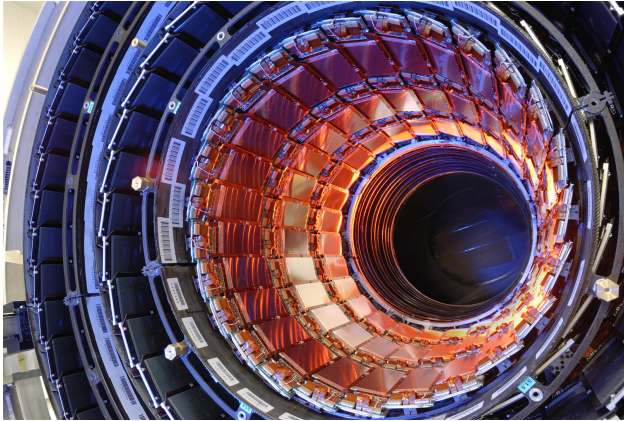


Figura 2.8: [30] Primera mitad de la capa exterior del barril del *tracker* de CMS.

A la hora de estudiar un proceso físico, uno de las magnitudes físicas más importantes a considerar es el momento de las partículas que, producto del citado proceso, se alejan del punto de colisión o interacción. El detector CMS calcula el momento de las partículas cargadas en función a cómo es su trayectoria bajo un campo magnético. Trivialmente, esta será curva, pero en función al momento que tenga la partícula en cuestión será más o menos curva.

El detector de trazas, más conocido como *tracker* se encarga de registrar las trayectorias de las partículas cargadas. Es el primer componente (alejándonos radialmente desde los tubos del LHC) del CMS y por lo tanto, el que más sufre la radiación producida por las colisiones. Además, debe ser capaz de perturbar lo más mínimo a las partículas (para que puedan acceder a otras partes del detector). Teniendo todo esto en mente, se escogió el silicio como material para hacer el *tracker*.

El funcionamiento en general es simple: los dispositivos de silicio que lo conforman son capaces de registrar con una precisión de $\pm 10 \mu\text{m}$ el paso de una partícula cargada. Con esos datos se puede reconstruir con mucha fidelidad la trayectoria. Para conseguirlo, dentro del *tracker* hay del orden de 6000 conexiones electrónicas por centímetro cuadrado. Esta cifra tan alta provocaría (por efecto Joule) un calentamiento del detector. Con ánimo de prevenir este, así como de minimizar los posibles daños por la intensa radiación, todo el detector de trazas se mantiene a $\sim -20^\circ\text{C}$, empleando un sistema de enfriamiento con

tubos que transportan gases (aire o nitrógeno) secos a los que se les reduce la temperatura previamente.

Este subdetector está subdividido en dos partes. La más cercana al punto de colisión, a escasos centímetros (la región más cercana, 4 cm) está formada por los llamados **píxeles de silicio**. Están dispuestos en capas (a 4 cm, 7 cm y 11 cm sobre los tubos del LHC) y en dos discos que actúan como «tapas» (se puede apreciar como la parte azul, no celeste, más interna de la figura 2.6). Las «teselas» que conforman esta parte del *tracker* suman en total unos 65 millones de «clics» que se activan cuando una partícula cargada los atraviesa, dando energía suficiente a electrones para que salgan de la órbita de los átomos de silicio. Es entonces cuando esos electrones se recogen en forma de una corriente eléctrica. Estas teselas o píxeles están distribuidas en cada una de las capas y miden $100\ \mu\text{m} \times 150\ \mu\text{m}$: su pequeño tamaño es el responsable de poder conocer con tanta precisión las trayectorias de las partículas cargadas.

La segunda parte del *tracker* está conformada por diez capas de **láminas de silicio** que se distribuyen hasta los 130 cm de radio desde los tubos del LHC, así como dos «tapas» del barril que cierran el detector de trazas. Su funcionamiento es muy similar al de los píxeles de silicio

2.2.3. Calorímetro electromagnético (ECAL)

El siguiente subcomponente del CMS es el calorímetro electromagnético, más conocido por ECAL. Su responsabilidad es la de medir la energía de electrones y fotones. Las condiciones de trabajo del LHC condicionaron trivialmente el desarrollo del ECAL. El material escogido, un centellador, es el tungstenato o wolframato de plomo ($PbWO_4$), en forma de cristales: estos son muy densos y centellean de forma rápida, en un área pequeña y en cascadas de fotones bien conocidas. Esto era ideal para los meros 25 ns que se dan entre cada choque de *bunches*, porque permite medir la energía de electrones y fotones rápidamente y de forma compacta espacialmente. Las cascadas son amplificadas después en un fotomultiplicador que se encarga después de transmitir la señal eléctrica.



Figura 2.9: [31] Operario tratando los cristales de wolframato de plomo que luego se instalarían en el ECAL del CMS.

En total, hay cerca de 80000 cristales así en el ECAL. Son alargados: los del barril miden $\sim 2 \times 2 \times 25\ \text{cm}^3$ y los de las tapas $\sim 3 \times 3 \times 20\ \text{cm}^3$. Así mismo, cada uno pesa 1.5 kg, lo que eleva la masa del calorímetro electromagnético a unas 100 t. Estos cristales requieren un tratamiento especial: por una parte, deben conservarse a una temperatura baja ($\sim 0.1^\circ\text{C}$) para que mantengan buenas propiedades centelladoras,

y por otra, la radiación en ellos provoca una leve coloración de los mismos. Durante el periodo de funcionamiento del LHC, existe un mecanismo para comprobar la transmisión óptica de cada cristal, y cuando el LHC se detiene y el detector vuelve a temperatura ambiente, los cristales son capaces de «autorreparar» las coloraciones, volviendo a su estado original.

Estos cristales se obtuvieron creciéndolos artificialmente durante dos días cada uno, tardando al final una década en obtener los 80000. Después se agruparon en subgrupos «pequeños», de 2.5 t cada uno para finalmente ser ensamblados dentro del CMS.

Una breve mención merece la «precascada» (*preshower*) del ECAL que está situada en cada una de las tapas. Cuando un pion neutro decae a dos fotones, estos pueden tener trayectorias muy similares y podrían registrarse como un solo fotón en el ECAL, lo cual sería un error. Para solventar esto y, además, poder diferenciar esta incidencia de los casos en los que hay una desintegración a dos fotones, se añade otra capa más en las tapas de ECAL. Esta, que está hecha en parte de silicio y en parte de plomo, mide con mayor resolución espacial debido a sus componentes las energías, y en conjunto con el resto de capas del ECAL y del *tracker*, también mejora la medición de las trayectorias.

2.2.4. Calorímetro hadrónico (HCAL)

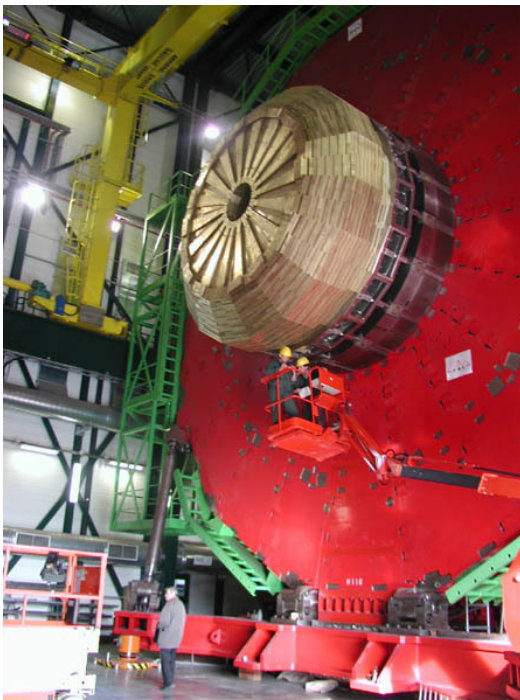


Figura 2.10: [32] Ensamblaje de la última «tapa» del calorímetro hadrónico de CMS.

Tras el calorímetro electromagnético se encuentra el hadrónico. Este se encarga de medir la energía de las partículas compuestas de *quarks* (es decir, hadrones). Como, por ejemplo, protones, kaones o piones. Además, se diseñó con la idea de que fuera lo más hermético posible. Esto permite medir el momento en eventos en los que no haya un balance en el mismo. P. ej., sucesos en los que se detecten partículas emergiendo por un lado, pero por el opuesto no. Esto nos permite inducir y calcular el momento de una supuesta partícula que nos haya sido imposible detectar, pero para ello debemos asegurarnos (con la hermeticidad del HCAL) de que somos capaces de medir el resto.

Este subdetector se organiza en capas y, como el resto, tiene forma de barril: tiene capas organizadas cilíndricamente y luego dos «tapas» en los laterales. Así mismo,

y análogamente al ECAL, tiene dos calorímetros «adelantados» (*forward calorimeters*), ubicados al final de CMS y justo sobre el tubo, para recoger todas las partículas que emergen del punto de colisión con ángulos muy bajos sobre los tubos.

El HCAL es verdaderamente masivo: pesa del orden de 1000 t y está hecho en buena parte de latón obtenido de cascotes de proyectiles de la Segunda Guerra Mundial cedidos gentilmente por la armada rusa. Su funcionamiento funciona a través de múltiples capas de absorbentes y de centelladores que las partículas atraviesan, dejando un rastro de energía. Cuando todos estos rastros de energía, que atraviesan varias capas, (llamados «torres») son sumados, se obtiene la energía de la partícula en cuestión.

2.2.5. Solenoide superconductor

Uno de los componentes esenciales del detector, como hemos visto, es el campo magnético al que deben estar sometidas las partículas. Este permite calcular el momento de las partículas cargadas, así como el signo de su propia carga (en función a cómo sea la curva que describen). El campo magnético está producido por un enorme solenoide superconductor que opera a unos ~ 4.7 T. Es de lejos la estructura más masiva del CMS, con 12000 t de peso (CMS pesa en total unas 12500 t), y es de hecho el mayor imán superconductor jamás construido.

Su característica superconductor le permite ofrecer un campo magnético muy potente, de 4 T, que trivialmente es de direcciones opuestas en el *tracker*, ECAL y HCAL, ubicados «dentro» del solenoide, y en los sistemas de muones, que se encuentran «fuera». Con el fin de que las líneas de campo magnético fuera del solenoide sigan un patrón comprensible que nos permita estudiar cómo se comportan las partículas bajo el mismo, los componentes del CMS externos al imán están incrustados en una estructura de hierro. Esta actúa como «guía» del campo magnético. Este armazón ofrece, como el propio solenoide, soporte estructural para el resto de componentes del CMS.



Figura 2.11: [33] Fotografía del solenoide del CMS durante su construcción.

2.2.6. Sistema de muones

Pasado el solenoide se encuentra el último componente del CMS: el sistema de muones (la parte más alejada del tubo roja, blanca y naranja de la fig. 2.6). Estas partículas, que atraviesan el «hermético» HCAL sin problemas, se miden en este subdetector que, al igual que el resto de partes del CMS, está dividido en un «barril» y dos tapas. Está formado de tres tipos de elementos que permiten tanto

actuar como un *trigger* rápido y detectar cuándo pasa un muon, como para medir la traza de los propios muones con precisión. Como esta parte del detector está «fuera» del solenoide, el campo magnético que se sufre tiene sentido contrario, por lo que el sentido de giro de la curva que realizarán las partículas cargadas será trivialmente el contrario. El alineamiento es casi perfecto con el *tracker* del centro del CMS, lo que permite que la reconstrucción de la trayectoria del muon sea muy fiable, pudiendo obtener su momento transversal.

Son tres los componentes del sistema de muones, que ahora describiremos brevemente. Los **tubos de deriva** (*Drift Tubes*, DT) son pequeños compartimentos de 4 cm de ancho que están rellenos de un gas y por el medio de los cuales pasa un cable. Cuando un muon atraviesa uno de esos tubos, proporcionará, al estar cargado, la energía suficiente para que algún electrón se desprenda de los átomos del gas. Estos se recogen en el cable (debido a la diferencia de potencial existente) y se registra esta señal. Cada cámara de tubos de deriva ($\sim 2 \times 2.5 \text{ m}^2$) tiene doce capas de tubos distribuidos en tres bloques de cuatro tubos. La orientación del primero y del último es distinta a la del grupo del medio, de modo que los cables que atraviesan el medio de los tubos del grupo del medio son paralelos a los tubos del LHC. Así, se pueden medir las tres coordenadas de los puntos de la trayectoria, puesto que cada bloque de tubos solo permite precisar dos.

En las «tapas» del sistema de muones no se emplean tubos de deriva, puesto que, si bien son precisos, dependen de la existencia de un campo magnético conocido y uniforme. En los extremos axiales del detector el campo magnético varía más, y se usan en su defecto **cámaras de tiras catódicas** (*Cathode Strip Chambers*, CSC). Estas consisten en un recipiente relleno de un gas que contiene una suerte de mallado hecho de cables que actúan como ánodos y tiras de cobre ortogonales a los primeros, que actúan como cátodos. El funcionamiento es análogo a lo que ocurre en los tubos de deriva, con la salvedad de que ahora los iones positivos se ven arrastrados también hacia el cobre. Esto permite conseguir dos coordenadas de la posición del muon cuando atravesó la cámara de tiras catódicas.

Finalmente, las **cámaras de placas resistivas** (*Resistive Plate Chambers*, RPC) están tanto en las tapas del detector como en el barril y se encargan de ofrecer un criterio de *trigger* rápido para tomar la decisión de conservar los datos del muon detectado. Su funcionamiento es similar al de los otros dos componentes. Las cámaras consisten en un compartimento hecho de dos placas cargadas (una positivamente, la otra negativamente) separadas por el volumen de un gas. Cuando un muon atraviesa el gas, se encontrará con átomos de este a los que dará suficiente energía como para desprenderse de un electrón. Estos electrones no pueden ser recogidos por las placas, que son transparentes a ellos, sino que se registran en unas tiras externas metálicas. La ventaja está en la rapidez del proceso (1 ns) y en su resolución espacial (1 cm, [34]), de ahí su uso como fuente de información para el *trigger*.

3 Descripción del trabajo experimental

¡Gordon! ¡Tío, me alegro de verte! ¿Qué demonios son estas cosas? ¿Por qué llevan puestos los uniformes del equipo de ciencia?

(Guardia de seguridad del complejo científico de Black Mesa, en *Half-Life*, VALVE)

LA investigación en física de altas energías experimental se realiza con el apoyo absolutamente necesario de herramientas informáticas. No solo por su utilidad *per se*, sino porque la inmensa cantidad de datos nos obliga indefectiblemente a ello. En este capítulo hablaremos sobre las bases del trabajo en HEP en aceleradores (particularmente, en el LHC).

Los estudios que se hacen con los datos recogidos en el LHC pueden llegar a ser muy distintos entre sí, aunque sobresalen dos tipos generales: aquellos que se destinan a buscar partículas que no se sustentan en los fundamentos del modelo estándar, o aquellos que tratan de medir parámetros ya conocidos o que pertenecen a la física esperable del modelo estándar. Por ejemplo, la búsqueda de partículas o procesos relacionados con las teorías SUSY son del primer tiempo, mientras que el mismo proceso $t\bar{t}H$ lo es del segundo.

Los procedimientos, en general, son similares en cualquier caso. Se parte de la **recolección** de los datos «en bruto» en el LHC. Se les aplica después un proceso de **reconstrucción** en el cual a partir de la información de cada subdetector, se intenta obtener la información física del suceso en cuestión (i.e., qué leptones había en el estado final, con qué trayectorias, qué momento...). Este procedimiento se hace una vez recogidos y almacenados los datos (se dice que se hace *offline*). Sin embargo los análisis se suelen enfocar a uno (a veces más) proceso físico determinado, y la información que se consigue en los detectores es de cualquier tipo de suceso que haya sido identificado y registrado: no necesariamente del proceso que pueda interesar al estudio. Con este fin, se realiza una **selección** de los datos obtenidos atendiendo a las características del proceso de interés: los sucesos en los que se dé este proceso se consideran la *señal* y el resto, el *fondo*. La relación entre estos dos subconjuntos de la selección (puesto que resulta casi imposible librarse de todo el fondo) es importante, y siempre se intenta maximizar la señal sobre el fondo. Explicaremos con más detalle en otra sección de este capítulo los motivos.

Sin embargo, ¿cómo sabemos que esos sucesos que llamamos señal provienen exactamente de nuestro proceso? Para tratar de saber esto, se recurre a la simulación a través de procesos de Monte Carlo de los procesos físicos en las mismas condiciones que ocurrirían en el LHC. Este procedimiento (y también todos estos datos) se conoce sencillamente como **generación**. Después, se simula cómo sería el paso de estos sucesos por el detector (CMS, en nuestro caso), para conseguir la misma información que vería en primera instancia el mismo. Llegados a este punto, la información sigue el mismo proceso de reconstrucción que siguen los datos (para evitar diferencias) y con esa información física del proceso se puede aplicar la misma selección que a los datos y comparar la predicción con estos. Debido a su naturaleza simulada, la selección de sucesos que provienen de simulaciones es algo distinta, como veremos en el siguiente capítulo.

3.1. Simulación de procesos

Las simulaciones de Monte Carlo (se suele decir «MC» sencillamente) de procesos se hacen a partir de un número no muy grande de programas. Las muestras de datos simulados o muestras MC provienen de haber encadenado el uso de un simulador del proceso, y de simular el paso de los objetos físicos por el detector, así como de la reconstrucción consecuente. Ejemplos de programas que simulan los procesos físicos (y que se usan en los análisis y estudios de este trabajo) son **MadGraph**, **Powheg** o **aMC@NLO**, que en general se combinan con **Pythia** para reconstruir la cascada de interacciones que siguen al proceso «central». Después, se emplea el *software* **Geant4** para simular la interacción de esos procesos físicos con el detector CMS.

Debido a que la física del modelo estándar (QFT, QED, QCD) es perturbativa, estas simulaciones se suelen clasificar en virtud al nivel de perturbaciones tenidas en cuenta para la generación de los datos: orden fundamental (*leading order* / *tree level*, LO), primer orden (*next-to-leading order*, NLO), segundo orden (*next-to-next-to-leading order*, NNLO), etcétera. De este modo, **MadGraph** permite obtener información a LO, mientras que **Powheg** y **aMC@NLO** la dan a NLO. La diferencia entre estos últimos radica en cómo devuelven la información de la simulación y en el número de *jets* adicionales que son capaces de simular en el proceso. Depende de cada proceso, pero en general **aMC@NLO** admite un mayor número de *jets* que **Powheg**. Además, si bien para todos los procesos que hay muestras a NLO hay muestras de **aMC@NLO**, no es así para **Powheg**.

El número de procesos simulado no es del mismo orden que el número de datos que se tienen del LHC: se hacen muchísimas más simulaciones, con el fin de proveer a los físicos experimentales de información suficiente para prevenir problemas de falta de datos simulados. Debido a esto, cada uno de los sucesos

de un proceso simulado se pesa de acuerdo al número de sucesos total de la muestra de MC y a la sección eficaz del proceso en las energías que se dan en el LHC. La diferencia que hay entre **Powheg** y **aMC@NLO** en cuanto a la información de la simulación es que este último hace las simulaciones a NLO dando pesos negativos a algunos eventos. Esto no constituye ningún problema en general, siempre que el número de sucesos simulados en el análisis, pasada la selección sea suficiente: **aMC@NLO** genera más sucesos con pesos positivos que con pesos negativos, de modo que el balance total *de la muestra* es que siempre hay valores positivos (y por tanto reales, físicos). El motivo de esto es la forma de simular los procesos del generador, que da un conteo doble a determinados eventos.

Como ya comentamos anteriormente, una vez hecha la simulación y la generación sigue un proceso de reconstrucción de los mismos, para volver a obtener la información física de los objetos de los eventos. Este se lleva a cabo *offline*, es decir, una vez obtenidos y almacenados todos los datos. Describiremos ahora cómo se hace este procedimiento, al que se someten tanto datos como simulaciones.

3.2. Reconstrucción de eventos e identificación de objetos

El proceso de reconstrucción de los eventos en el detector CMS ([35]) se basa en el **algoritmo *particle-flow*** («flujo de partículas») o PF ([36], [37], [1]). Este parte de la información más inmediata que el detector puede recoger y la combina con otra más elaborada para rehacer los objetos físicos del evento, i.e., electrones, muones, fotones, hadrones cargados y hadrones neutros. El algoritmo construye unos **elementos** que emplea después para identificar los objetos.

Las señales o *hits* que deja una partícula en el *tracker* y en el sistema de muones se emplean para reconstruir las **trazas** o trayectorias de las partículas en ellos. Con este fin, se usa un proceso de cuatro o cinco iteraciones en las que se identifican las trayectorias. En la primera iteración, se utilizan criterios muy estrictos para encontrar las trazas, de modo que solo las trayectorias de partículas que claramente se observe que hayan pasado se admiten. En las siguientes iteraciones, se relaja este criterio, incrementando el número de trayectorias recogidas. Este pequeño algoritmo se ha probado eficaz a la hora de encontrar las trazas de partículas.

Otro elemento que se construye son los llamados **clusters** en los calorímetros. Estos se obtienen a partir de los depósitos de energía que las partículas dejan en cada calorímetro. Otro algoritmo las agrupa en torno a deposiciones elevadas, obteniendo al final conjuntos de estos restos de energía que se pueden identificar como la energía que ha dejado en uno de los dos calorímetros una partícula.

Una vez construidas las trazas en *tracker* y sistema de muones y los *clusters*, el algoritmo PF

intenta conectar elementos de ambos tipos (un número variable de ellos) para construir candidatos a objetos físicos (muones, hadrones...), que deberán pasar algunos requisitos más para que el algoritmo concluya que, en efecto, lo son (según se hayan «definido»). No se buscan a la vez todos los objetos físicos del proceso, sino que se sigue un orden en función a la facilidad o fiabilidad del reconocimiento del objeto en cuestión en el CMS. Primero se reconstruyen los muones, debido a que requieren *hits* en los sistemas de muones que el resto de partículas no (y ello es muy distintivo). Después, los electrones y los hadrones cargados, aprovechando las trazas y los depósitos en el ECAL y HCAL. Encontrados los electrones, los *clusters* que queden en el ECAL deberán corresponderse con fotones y finalmente, los que queden en el HCAL, con hadrones neutros. Conseguidos estos objetos, se puede proceder a calcular el cuatrimomento de cada uno.

En las siguientes subsecciones describiremos cómo se construyen estos candidatos a objetos y cómo se identifican al final si lo son, o no.

3.2.1. Vértice primario

Reconstrucción e identificación

Se llama vértice primario (*primary vertex*, PV) al punto espacial donde chocaron los dos protones que dieron lugar al evento registrado. Tiene especial importancia, sobre todo para detectar los *quarks* *b*, como veremos.

La forma de calcularlo es extrapolar las trayectorias detectadas hacia la zona central del detector: donde se cruzan los dos tubos del LHC y se producen, por consiguiente, las colisiones. Sin embargo, y como ya explicamos, debido al *pile-up*, podemos registrar más de una colisión a la vez. En tal caso (y de ahí el «primario» del PV), se define el PV como aquel cuyos objetos físicos reconstruidos tienen mayor momento transversal (i.e., momento en el plano transversal del detector).

Otras variables relacionadas con la calidad de la traza de las partículas se emplean para identificarlo concretamente. Una de ellas, por ejemplo, es el indicador de un test χ^2 de Pearson que se realiza con las señales recogidas en el *tracker* y la trayectoria reconstruida de alguna de las partículas del supuesto vértice principal.

Relación con la identificación de otros objetos

Más tarde, cuando se reconstruyan los leptones, a cada uno se le dará una serie de variables relacionadas con el vértice principal. Entre ellas, y que usamos en nuestro análisis, se encuentran los parámetros de impacto en centímetros respecto al mismo en el eje de los tubos ($d_z(l)$), y en los ejes transversales ($d_x(l)$ y $d_y(l)$), a veces se unen en $d_{xy}(l)$), así como la significancia de los parámetros de

impacto en el espacio del detector, $\text{SIP}_{3\text{D}}(l)$. Esta última se define como una significancia, esto es, nos dice cuán grande es el valor que hemos medido respecto a su incertidumbre: $\text{SIP}_{3\text{D}} := \frac{d(l)}{\Delta(d(l))}$. Ajustándolos, podemos exigir en la selección leptones que vengan con más o menos certeza del vértice primario.

3.2.2. Muones

Reconstrucción

El procedimiento para reconstruir muones parte de las trazas del evento del *tracker* y del sistema de muones. Con ellas, el primer paso es tratar de conectar una traza del *tracker* con otra de las cámaras de muones, consiguiendo así la trayectoria completa del muon en el detector. Esto se hace teniendo en cuenta las pequeñas pérdidas de energía que hace el muon mientras lo atraviesa, así como naturalmente, el cambio en el sentido del campo magnético al pasar el solenoide. Una vez se tienen dos trazas compatibles, se reconstruye una traza completa con toda la información (esta vez, conjunta) de las dos trazas.

Identificación

Una vez reconstruido, para su identificación se emplean distintas variables. Hablaremos ahora de las más importantes y de otras que usaremos en nuestro análisis.

Las exigencias más básicas están relacionadas con el **momento transversal** de los muones y el valor de η : estos se corresponden con los de la partícula en el momento inmediato tras darse el proceso físico (y que estimamos con ayuda de los subdetectores). Esto permite no considerar, por ejemplo, muones que debido a cómo haya sido la colisión, emerjan de ella en una dirección bastante paralela a la del tubo del LHC, donde la reconstrucción no es tan buena. Así mismo, un criterio de selección en el momento se puede hacer con muchos motivos. Exigir un $|\vec{p}_T|$ mínimo de 10 – 30 GeV permite librarse de muones que provienen de procesos físicos que probablemente no nos interesen, puesto que el que estudiamos solo se dé significativamente a una energía superior.

Los *quarks*, debido a la naturaleza de la interacción fuerte, tienden a sufrir procesos de hadronización poco después de la colisión. En ellos, es habitual que se vean involucrados otros muones o leptones, por ejemplo. Para «separar» los muones que salen del vértice primario de estos, se define, en el espacio (η, φ) del cual hablamos anteriormente, el **aislamiento**, I . Este se define en función a la distancia de tal espacio, dada por $\Delta(R) := \sqrt{(\eta_i - \eta_j)^2 + (\varphi_i - \varphi_j)^2}$. En el espacio tridimensional, los lugares geométricos de una distancia fija $R_{\text{máx.}}$ son conos. El aislamiento de un muon se define como la suma de los módulos de los momentos transversales de todas las partículas que se encuentren en un cono de radio $R_{\text{máx.}}$ centrado en el muon. Es decir,

$$I(\mu) := \sum_{\Delta(R_i) < \Delta(R_{\text{máx.}})} |\vec{p}_T(i)|. \quad (3.1)$$

Esta variable permite hacernos una idea de la cantidad y el momento de las partículas que rodean a un muon determinado. Trivialmente, valores bajos de la misma serán indicativos de muones aislados. Otra variable que se suele definir es el **aislamiento relativo**, $I_{\text{rel.}(l)}$. Esta se define como la intensidad dividida entre el momento transversal (i.e., $I_{\text{rel.}(l)} := \frac{I(l)}{|\vec{p}_T(l)|}$), lo cual de una medida relativa a cada muon del aislamiento. Por desgracia, en casos en los que la colisión tiene como resultado un movimiento del sistema de ref. de la partículas respecto al detector muy alto (estos casos se dicen *boosted*, «acelerados», en referencia a los *boosts* de relatividad especial), el aislamiento relativo no es suficiente. En tales situaciones, los productos de la colisión se tienden a agrupar, y las definiciones de aislamiento anteriores no resultan tan efectivas como en eventos sin colisiones *boosted*.

Para resolver este inconveniente, se define el **mini aislamiento**,

$$I_{\text{mini}}(l) := \frac{\sum_{\Delta(R_i) < \Delta(R^*(l))} |\vec{p}_T(h_i^\pm)| - \text{máx.} \left\{ 0, \sum_{\Delta(R_j) < \Delta(R^*(l))} |\vec{p}_T(h_j^0)| + \sum_{\Delta(R_k) < \Delta(R^*(l))} |\vec{p}_T(\gamma_k)| - \rho^{\frac{R^*(l)}{0.3}} \mathcal{A}(l) \right\}}{|\vec{p}_T(l)|}, \quad (3.2)$$

donde $h_i^\pm$, h_j^0 y γ_k hacen referencia a hadrones cargados, neutros y fotones, ρ es la **densidad de energía de pile-up**, R^* se define como

$$R^*(l) := \frac{10}{\text{mín.} \{ \text{máx.} \{ |\vec{p}_T(l)|, 50 \}, 200 \}}, \quad (3.3)$$

y $\mathcal{A}(l)$, un área efectiva que depende del valor de η del leptón en cuestión. La presencia del sumando $\rho^{\frac{R^*(l)}{0.3}} \mathcal{A}(l)$ se justifica por la contaminación de las otras colisiones que se dan, esto es, el *pile-up*. Si bien es posible no contabilizar los hadrones cargados de otros vértices (puntos de colisión), ya que la trayectoria se puede reconstruir con el *tracker*, eso no lo podemos hacer con los hadrones neutros y los fotones. Para ello se agrega ese término que depende de la energía asociada al *pile-up* y actúa como una corrección. Exigir un valor bajo de esta variable conlleva un mejor aislamiento de nuestros leptones, garantizando así que provengan del proceso principal y del vértice primario, reduciendo las posibilidades de que los confundamos con otros muones y/o electrones.

En el análisis que nos atañe, $t\bar{t}H$, el origen del mayor número de leptones mal identificados procede de la desintegración de los hadrones con *quarks b* (de procesos como $t\bar{t} + jets$ o $W + jets$, por ejemplo). En ellos, los leptones asociados a la hadronización se pueden identificar mal como leptones que provienen del proceso físico de la colisión. Para tratar de evitar estos leptones mal identificados, se emplean algunas

variables relacionadas con esta hadronización. Por ejemplo, se considera el **discriminante de *b*-tagging del algoritmo CSV** (véase más adelante, en el apartado sobre *quarks*, su descripción) del *b*-jet más cercano al leptón, el cociente $\mathbf{p}_T^{\text{ratio}}$ entre el momento transversal del leptón y el del *b*-jet en cuestión, y la variable $\mathbf{p}_T^{\text{rel}}$ definida como sigue.

$$p_T^{\text{rel}}(l) := \frac{(\vec{p}(jet) - \vec{p}(l)) \cdot \vec{p}(l)}{\|\vec{p}(jet) - \vec{p}(l)\|} \quad (3.4)$$

Dentro de la subcomunidad científica de investigadores que trabajan conjuntamente con los resultados del experimento CMS llamada «Colaboración CMS» (*CMS Collaboration*) existe un subgrupo llamado *Muon Physics Object Group* (Muon POG, o simplemente POG, [38])¹ que se encarga de estudiar las eficiencias de identificación de objetos en todo el detector. Para ello tienen en cuenta la forma física de los subdetectores, su distribución, materiales, etcétera. Para cada *run*, este grupo construye una serie de distintas identificaciones de muones, que van desde selecciones de objetos reconstruidos que podrían ser muones, pero no se tiene mucha certeza de ello (muones *loose*), hasta otras en las que se imponen exigencias muy fuertes, de modo que se está casi seguro de que los objetos conseguidos son muones (muones *tight*)² Hay otras identificaciones también, como algunas dedicadas exclusivamente a muones de alto momento transversal.

Todas estas selecciones se ponen a disposición de toda la Colaboración CMS, para que se empleen en los distintos análisis que se hacen. Nosotros haremos uso de las identificaciones ***loose muon***, que apenas impone condiciones sobre el objeto reconstruido por el algoritmo PF, y ***medium muon***. Esta última exige varias características al muon reconstruido para que asegurarnos de que sea un buen muon global.

Como ya comentamos en la descripción del proceso $t\bar{t}H$ (1.4), uno de los fondos más importantes que tenemos en el análisis se debe al proceso $t\bar{t}$. Uno de sus estados finales, en el cual hay dos leptones de distinto signo, nos afecta especialmente, debido a que si identificamos mal la carga de uno de los leptones, podríamos pensar que se da una situación en la que hay dos leptones de carga igual (y recordemos, buscar estos estados finales era una forma de limpiar una gran cantidad de fondo y una garantía de señal limpia). Con el fin de reducir estos casos (a tales leptones se les dice leptones *fake*), se tiene en cuenta en la identificación la variable conocida como ***tight-charge***, que usaremos en la identificación de muones y de electrones. Esta es una suerte de pre-categorización de los objetos reconstruidos que se basa en distintas variables (como exigir una reconstrucción muy buena de la traza del *tracker* o consistencia entre las medidas de la carga del ECAL y el *tracker*, en el caso de los electrones).

¹Existen análogos para electrones, fotones, etc..

²Es común usar las palabras *tight* o *loose* para definir a categorías, criterios de selección o tipos de esta. De aquí en adelante usaremos esa notación, donde un objeto *tight* será aquel que hemos seleccionado con exigencias altas, estrictas, y uno *loose*, aquel seleccionado con exigencias más relajadas o suaves.

([39]) En análisis en los que es complicado obtener sucesos de señal, se emplea una variable conocida por **Lepton MVA** que se define para cada lepton en función a otras variables del mismo. Los análisis o estudios multivariable (*multivariate analysis* MVA) son aquellos que se hacen en virtud a un espacio de variables de varias dimensiones (dos o más): el *Lepton MVA* se construye en función a las siguientes variables de entrada.

- $|\vec{p}_T(l)| = p_T(l)$.
- $\eta(l)$.
- $I_{\text{mini}}(l)$, considerando solo hadrones cargados en la suma de esta.
- $I_{\text{mini}}(l)$, considerando solo hadrones neutros y fotones en la suma de esta.
- $p_T^{\text{rel}}(l)$.
- $p_T^{\text{ratio}}(l)$.
- Discriminante de *b-tagging* del algoritmo CSV (v2).
- $\text{SIP}_{3\text{D}}(l)$.
- $d_{xy}(l)$.
- $d_z(l)$.
- Compatibilidad de segmentos (*segment compatibility* para muones): variable definida según la concordancia de la trayectoria reconstruida de los candidatos a muones con cada una de las señales de las cámaras de muones que pasan cerca de la misma, en función a un análisis estadístico de Pearson χ^2 .
- Valor del MVA-ID cuando no sobrepasa el valor de *trigger* (para electrones): es otro MVA basado también en un BDT que se construye con distintas variables destinadas a la selección específica de electrones. No entraremos en detalle, pues se aleja bastante del alcance de este trabajo.
- Varias variables relacionadas con la calidad de la traza.

Estas variables se someten a un procedimiento de análisis de datos conocido como árbol de decisión dopado (*boosted decision tree*, BDT). Esta herramienta de clasificación se emplea en el contexto del aprendizaje automático (*machine learning*) y permite al final obtener un árbol de decisión con el que clasificar los eventos en función a las variables antes escogidas. El algoritmo «entrena» el árbol para que *aprenda* a diferenciar sucesos que nos interesan (señal) de aquellos que no (fondo). En el caso que nos atañe, el árbol se ha entrenado con una muestra de MC del proceso $t\bar{t}H$ como señal y contra leptones *fake* de una muestra de MC de $t\bar{t} + \text{jets}$. Una vez entrenado, se aplica el BDT a los datos o muestras de

MC que queramos, y este asignará un valor en el intervalo $[-1, 1]$ (también $[0, 1]$) usualmente llamado discriminador del BDT. Aquellos leptones con un discriminante alto, cercano a la unidad, serán leptones similares a los que podríamos esperar del proceso que nos interesa ($t\bar{t}H$), y los cercanos al -1 , los menos parecidos.

El algoritmo de «entrenamiento» para obtener un BDT no es ni mucho menos sencillo, y no nos pararemos a detallarlo aquí. En cualquier caso, es importante señalar que si bien un árbol de decisión dopado que haya sido bien entrenado es una herramienta potente para mejorar la selección de los sucesos de señal, tiene una contrapartida importante. Desde un punto de vista matemático, el algoritmo y el funcionamiento de los BDT está correcta y debidamente asentado sobre una fundamentación estadística. Pero una vez se ha conseguido un buen BDT, y se han clasificado los datos de CMS con él, no se puede hacer una justificación física de por qué en una selección exigimos a nuestros leptones, por ejemplo, un discriminante mayor que 0.72: esa capacidad se ha perdido. Con las variables de entrada sí se puede: podemos tener una justificación física para hacer un criterio de selección en $|\vec{p}_T|$ por ejemplo, y podríamos intentar entender el resultado. Sin embargo, con el discriminante del BDT no: además de que la variación en el discriminante se traduce en un desplazamiento en un espacio multidimensional ya difícil de entender, el módulo y dirección de la variación proceden del proceso de diseño y entrenamiento. Por ello, es extremadamente complicado entender físicamente las clasificaciones basadas en el discriminante del BDT, que solo nos da información sobre el leptón por su cercanía a la unidad, o al -1 .

Además de todas estas variables que se usarán en la identificación de muones (y leptones) también tenemos en consideración la **compatibilidad de segmentos** o $\mathbf{p}_T^{\text{ratio}}$, de las que ya hablamos dentro de la descripción del *Lepton MVA*.

3.2.3. Electrones

Reconstrucción

La reconstrucción de los electrones se hace de forma similar a la de los muones, solo que sin las trazas del sistema de muones (obviamente). La trayectoria se consigue solo estudiando las trazas del *tracker* y los *clusters* del ECAL. Aunque se han de tener en cuenta dos cosas: debido a su ligera masa en comparación con los muones, la radiación que emiten por *Bremsstrahlung* es mucho más grande, y se ha de tener en cuenta a la hora de estimar las trayectorias. Y, además, los *clusters* del ECAL pueden deberse bien a electrones, o bien a fotones. Para diferenciar ambos casos hay que intentar correlacionar una traza del *tracker* con un *cluster*.

Identificación

Para identificar los objetos resultantes del alg. PF como electrones se emplean requisitos en algunas de las mismas variables que describimos para los muones, como en η , p_T , d_{xy} , d_z , SIP_{3D}, I_{mini} , p_T^{ratio} , el discriminante de *b-tagging* del *b-jet* más cercano al leptón, los requisitos de *tight-charge*, o el propio *Lepton MVA*. Sin embargo, para los electrones se emplean normalmente y además, otros requisitos.

Otro análisis MVA conocido simplemente como ***electron MVA ID***, que es el que mencionamos cuando describimos la variable **Lepton MVA**.

También se usa la variable $\frac{1}{E} - \frac{1}{p}$. En caso de que la reconstrucción sea fiable, será cercana a cero (en general, si es negativa o muy alta, se considera que se ha reconstruido mal el electrón). Otra es $\frac{H}{E}$: el cociente entre la energía depositada en el HCAL y en el ECAL que se podría asociar a la trayectoria de la partícula. Si esta es un electrón, se esperaría que tal cociente fuera muy pequeño.

Para ahondar en la calidad del vínculo entre la traza reconstruida del posible electrón y la deposición asociada en el ECAL de energía, se usan variables como $\Delta\phi_{i\eta}$ o $\Delta\eta_{i\eta}$, que son la variación en las variables ϕ y η entre el cuadrimomento del posible electrón (en el instante después del proceso físico del vértice principal) obtenido solo usando la trayectoria reconstruida en el *tracker*, y entre la del cuadrimomento extrapolado a partir de la trayectoria obtenida con la información del *tracker* y del ECAL. Valores pequeños de esta nos dirán que añadir la información del ECAL no supone una variación grande en el cuadrimomento (lo que tiene lógica si la deposición en él está asociado a una trayectoria reconstruida en el *tracker*).

Con el fin de garantizar también la calidad del rastro de energía del ECAL, se hacen criterios de selección sobre la variable $\sigma_{i\eta i\eta}$, que mide la esfericidad de los rastros medidos en el subdetector.

Muchas veces, algún fotón que procede de la colisión puede *convertirse* (decaer) en un par e^+e^- , y estos electrones pueden ser reconocidos como si fueran electrones del proceso físico, aunque no lo sean. Para evitarlo, se ha construido una variable, **rechazo de conversiones** (*conversion rejection* o *conversion veto*) que estima a partir de información de las capas más cercanas al *tracker* del ECAL si el electrón que consideramos proviene de un fotón «convertido» o no.

Es común también rechazar candidatos a electrones de PF que hayan dado señales en el *tracker* lejos de su trayectoria reconstruida. Se suele exigir a las identificaciones más estrictas que no haya ninguna **señal perdida** o *missing hit*.

3.2.4. Hadrones: *jets*

Reconstrucción

Por la naturaleza de la interacción fuerte, todos los *quarks* que se produzcan en el proceso del evento sufrirán procesos de hadronización, a excepción de los *tops* que, debido a su vida media comparativamente corta con el resto, decaen (en otros *quarks* o en objetos distintos como leptones) antes de ser capaces de hadronizarse. En el proceso de hadronización surgen otras partículas derivadas del mismo, como leptones por ejemplo, así como otros posibles hadrones. Al conjunto del hadrón y todas estas partículas se le conoce como *jet* (fig. 3.1).

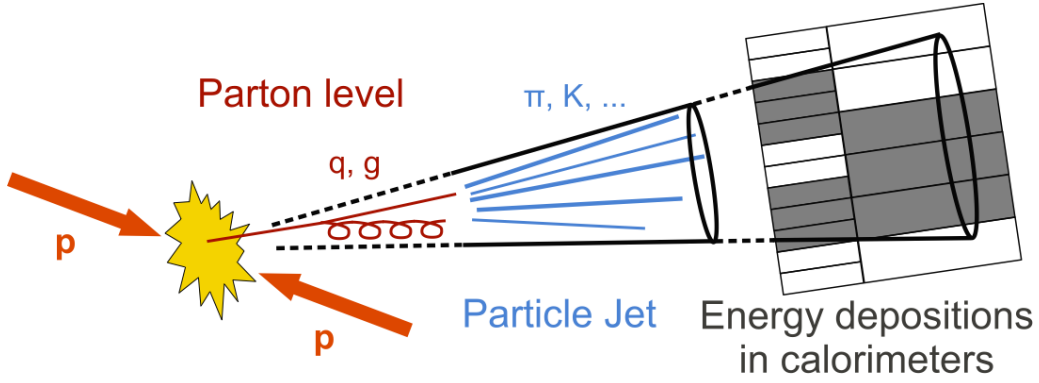


Figura 3.1: [40] Diagrama de la producción de un *jet*.

El proceso de reconstrucción de estos *jets* se hace partiendo de los candidatos dados por el algoritmo PF, que emplea a su vez el algoritmo *anti- k_t* ([41]). Este se encarga de conformar los *jets*, agrupando las partículas que a ellos pertenecerían, a partir de los momentos transversales de cada una, así como de algunos parámetros.

No somos capaces de distinguir en general si un *jet* que detectamos proviene de uno u otro *quark* derivado del proceso físico, con dos excepciones. La primera ya la mencionamos: los *quarks top* tienen una vida media muy baja y decaen antes de poder hadronizarse. La segunda son sus compañeros de generación: los *quarks bottom*. Cuando uno de estos *quarks* es producido por la colisión, sufre un proceso de hadronización. Los hadrones de estos *quarks* en general tienen una vida media comparativamente mayor, y son capaces de desplazarse una pequeña distancia (algunos milímetros) dentro de los tubos del LHC antes de decaer (véase fig. 3.2). El punto en el cual estos hadrones decaen se puede identificar como si fuera

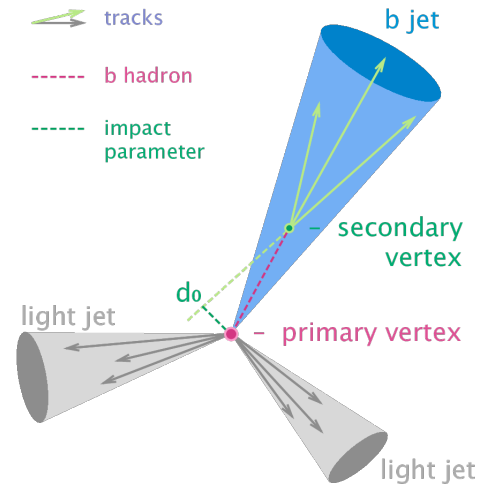


Figura 3.2: [42] Diagrama de la producción de un *b-jet*.

un vértice secundario, en donde no se da una colisión, sino una desintegración. La resolución del *tracker* (aprox. $20\,\mu\text{m}$ en el plano transversal y $30\,\mu\text{m}$ en el eje z para el vértice primario) permite detectar estos vértices secundarios, de modo que esta es una vía para poder identificar qué *jets* provienen de *quarks b*.

Con este fin se ha elaborado el **algoritmo CSV**, y concretamente, su última versión, el **algoritmo CSVv2** ([43]). Este tiene en cuenta la identificación de este vértice secundario, así como otros muchos relativos a la traza, parámetros de impacto, etcétera. Al final, el algoritmo devuelve un valor en el intervalo $[0, 1]$, que nos informa de cuán seguro está de que es un *b-jet*. Valores cercanos a la unidad significan que el *jet* en cuestión es muy probable que lo sea, y alejados de esta, que no (este número se obtiene a través de una clasificación hecha con las variables de entrada). Como de este modo podemos decir que un *jet* es probable que sea (o no) un *b-jet*, se conoce a este proceso como «etiquetado de *quarks b*», o simplemente *b-tagging*. También es común referirse a este procedimiento en general como *b-tagging* o *b-tag*.

Identificación

A la hora de identificar nuestros *jets*, se recurre en general a exigencias relacionadas con el p_T , los valores de η , así como claramente el discriminante de *b-tagging*. Un proceso análogo al que se hace con los leptones, pero a la inversa, es tratar de separar todos los *jets* de muones y electrones reconstruidos. Para ello se recurre a exigir una distancia mínima (en el espacio (η, φ) , esto es, conos en $\mathbf{R}^3$) entre algunos o todos los leptones reconstruidos e identificados, y los candidatos a *jets*. Esto se conoce como «limpieza» (más comúnmente *cleaning*) y no solo se aplica a los *jets*: otros objetos son susceptibles de ser igualmente «limpiados».

3.2.5. Taus

Reconstrucción

([44]) Los leptones tau no pueden ser identificados de un modo análogo a electrones y muones, usando el algoritmo PF. El motivo es que es el único leptón con la masa suficiente ($m_\tau = 1.777\,\text{GeV}$) para decaer en hadrones, a diferencia de los muones y electrones. De hecho, la fracción de desintegración (en porcentaje) de todos los modos en los que hay hadrones como producto de la desintegración es 64.8%, mientras que en un 17.8% los leptones τ decaen a electrones ($\tau \rightarrow e^- \bar{\nu}_e \nu_\tau$) y en un 17.4% a muones ($\tau \rightarrow \mu^- \bar{\nu}_\mu \nu_\tau$).

Debido a los modos de desintegración a muones y electrones (que sí se identifican con el algoritmo PF) en los que es complicado detectar que vengan de un tau, estos no se suelen considerar. Sin embargo, los taus que decaen a algún hadrón (y otras partículas), denotado este (y el resto de subproductos de la

hadronización) por τ_h (taus hadrónicos) son identificables. Para ello, se tienen en consideración muchos factores, como la separación con electrones y muones, variables cinemáticas y otras. Esto resulta en el conocido como **algoritmo «hadron-plus-strips»**, que permite la reconstrucción de estos leptones tau.

Identificación

Se emplean criterios de selección tanto en p_T , como en η , así como un análisis MVA entrenado con señales de $t\bar{t}H$ contra fondo de $t\bar{t}$, donde todos los eventos tienen un aislamiento del muon mayor.

También se pueden exigir limpieza (*cleaning*) de estos τ_h respecto a otros objetos del proceso físico de interés, como por ejemplo, los electrones y muones. Por «limpieza» nos referimos a lo análogo que explicamos para los *jets*: a garantizar que ningún τ_h se encuentra a una distancia (en el espacio (η, φ)) menor que un máximo determinado de los objetos respecto a los que se quiera limpiar.

3.2.6. Energía transversal faltante y otras cantidades características del proceso

Los neutrinos que se produzcan en cualquier colisión, ν_l , casi con absoluta seguridad, no interactuarán mientras se alejan del centro del detector, y atravesarán todos y cada uno de los subcomponentes del CMS sin dejar ningún rastro. Esto también se aplica a algunas de las partículas BSM de las cuales se espera que quizás en el LHC sea donde tengan ocasión de darse a conocer.

¿Cómo obtener información, pues, de aquello que no podemos medir? La solución que masivamente se emplea es construir una variable llamada **energía transversal faltante** que se define como el módulo del vector opuesto que resulta de la suma de todos los momentos transversales de todos los objetos reconstruidos por el algoritmo PF. Es decir, el módulo del vector

$$\vec{E}_T^{\text{miss}} = - \sum_i \vec{p}_T(i), \quad (3.5)$$

que es

$$E_T^{\text{miss}} := \left| - \sum_i \vec{p}_T(i) \right|. \quad (3.6)$$

Es casi imperativo hacer, una vez definida esta variable, la siguiente aclaración. Formalmente, E_T^{miss} **no** es una energía. Lo sería en el hipotético caso de que todos los objetos reconstruidos tuvieran masa nula, pero pese al abuso de notación que esto supone, el hecho es que el nombre se ha extendido en toda la física de partículas de aceleradores.

La importancia de esta variable es que nos da una medida, *siempre* imprecisa, de cómo sería el p_T de una hipotética partícula que hubiera sido producida en la colisión también, pero que no haya sido

detectada. Nótese que si se produjeron en el proceso físico de la colisión, por ejemplo, dos neutrinos, no tenemos forma de separarlos: E_T^{miss} sería el módulo del vector suma de los dos momentos transversales.

En virtud al proceso que nos interese, nos puede ser de utilidad estudiar esta cantidad. Es fácil percatarse de que, por desgracia, el *pile-up* influye negativamente en la fiabilidad de E_T^{miss} . Con la intención de solventar este problema, se define la variable $\mathbf{H}_T^{\text{miss}}$, definida de modo idéntico a E_T^{miss} , pero en vez de tener en cuenta todos los objetos reconstruidos del algoritmo PF, solo se usan los determinados por la selección. Esta cantidad es más robusta que la energía transversal faltante, puesto que los objetos seleccionados se presuponen fiables (en el sentido de que se les han puesto requisitos para asegurarse de tener objetos físicos certeros), aunque tiene peor resolución que E_T^{miss} (debido a que su definición depende de muchos menos objetos). Para resolver el problema del *pile-up*, se combinan ambas variables en un **discriminante lineal**,

$$\mathbf{E}_T^{\text{miss}}\text{LD} := E_T^{\text{miss}} \cdot 0.00397 + H_T^{\text{miss}} \cdot 0.00265, \quad (3.7)$$

combinando ambas variables, siendo en este discriminante donde se imponen los criterios de selección de selección de eventos.

Otra cantidad que se suele emplear para analizar o seleccionar eventos es la variable H_T ,

$$H_T := \sum_i p_T(jet_i), \quad (3.8)$$

que es una medida de la energía (momento en verdad) hadrónica (transversal) del evento.

3.3. Cálculo de secciones eficaces y límites superiores

La sección eficaz, en física de partículas en aceleradores, como hemos visto, es una estimación de cuán probable es que se dé un determinado proceso físico en alguna colisión. Para procesos contenidos dentro del modelo estándar, estas son predecibles y calculables, aunque no entraremos en las ecuaciones que permiten calcularlas.

Experimentalmente, en física de partículas podemos calcular la sección eficaz de la siguiente forma (se puede justificar),

$$\sigma = \frac{N_{\text{datos}} - N_{\text{fondo}}}{\varepsilon \mathcal{L}}, \quad (3.9)$$

donde N_{datos} y N_{fondo} son el recuento de los datos experimentales que tenemos y el fondo obtenido de simulaciones de MC, ε son un conjunto de eficacias y $\mathcal{L}$ es la luminosidad integrada. Si bien esta

magnitud ya es importante *per se*, normalmente se emplea la **fuerza de señal**, o *signal strength*, que se define como el cociente entre el valor experimental de la sección eficaz, y el valor teórico de esta. Es decir,

$$\mu := \frac{\sigma_{exp.}}{\sigma_{teo.}} \quad (3.10)$$

La sección eficaz es una de los parámetros característicos de los procesos físicos. El mayor requisito para poder declarar a un proceso como «observado» es haber medido la sección eficaz con una significancia mayor que 5σ (por convenio), usándose como comentamos de forma común μ . Esta clase de comparaciones con los resultados teóricos se hacen a través de tests de hipótesis, en los que tener un resultado con una significancia «de x sigma» quiere decir que, suponiendo la hipótesis nula, la probabilidad de tener un resultado como el nuestro, o peor, es la misma que si en una distribución normal $Z := N(0, 1)$ calculásemos la probabilidad de tener un resultado de $Z > x$ (esto es, lo que se conoce usualmente como p-valor).

O, dicho de otra forma, quiere decir que bajo la hipótesis de que mi resultado es erróneo (para secciones eficaces esto se correspondería con que hemos medido una señal que no es la del proceso que buscamos, la medimos por «error estadístico»), la probabilidad de que tal cosa sucediera sería la misma que la probabilidad de $Z > x$.

En análisis como el nuestro no es posible calcular la sección eficaz de modo que salga un valor con sentido; como ya hemos dicho, nuestro análisis necesita de más luminosidad para tener los suficientes datos como para calcularla. Por consiguiente, lo que se intenta no es calcular la sección eficaz (o μ), sino poner límites superiores a sus valores. Concretamente, esto se hace sobre la variable μ .

Nosotros hemos calculado el límite utilizando los programas preparados por la Colaboración CMS ([45]) utilizando el siguiente algorítmico. Este se basa en realizar un test con hipótesis nula $H_0 := [\text{fondo}]$ e hipótesis alternativa $H_1 := [\text{fondo} + \text{señal}]^3$, y se usa un intervalo de confianza del 95 %.

Para ello se emplea un estadístico basado en un cociente de ecuaciones de verosimilitud que se construyen partiendo de μ , la fuerza de la señal y generalizando el test de hipótesis a una hipótesis alternativa, en vez de $S + F$, $\mu S + F$. Después, se aplica un algoritmo conocido comúnmente por CLs, *Confidence Levels*, que es el que de verdad calcula el intervalo de confianza (del 95 %), y permite obtener el límite superior.

Este algoritmo sin embargo, recurre a muchas simulaciones y puede ser costoso computacional-

³Aunque a la hora de ejecutar el procedimiento CL_s , el test de hipótesis usado sea el inverso.

mente. Con el fin de agilizar el proceso, se puede emplear una aproximación asintótica (de Wald) para obtener el límite superior. Este método se conoce como método asintótico, y es el que usaremos. El motivo de buscar límites superiores (y no inferiores también) es que normalmente se espera encontrar algo más que el fondo, por lo que la sección eficaz experimental será superior a la teórica.

Los límites se diferencian entre límite observado y esperado. El primero se calcula haciendo un ajuste a los datos observados, mientras que el segundo se sustituyen estos con simulaciones de datos. Para optimizar análisis, se tienen en cuenta los límites esperados, ya que mejorar un estudio atendiendo a las observaciones implicaría introducir un determinado sesgo en el mismo. Una explicación más detallada se puede encontrar en [46] y [47].

4 Implementación computacional

Perdona el desorden. No he tenido tiempo de poner orden desde que me mataste. [Suena voz pre-grabada] Prueba de Sarcasmo Completada.

(La inteligencia artificial GLaDOS de Aperture Science, en *Portal 2*, VALVE)

TODOS los análisis y estudios que se han hecho en este trabajo se han programado en **C++**, junto con **ROOT**: un conjunto de librerías y programas contruidos sobre **C++** también que se usa esencialmente en física de partículas, debido a las funciones y facilidades que ofrece para análisis de este tipo. Así mismo, el marco de trabajo (comúnmente, *framework*) que se ha usado es el creado por el Grupo Experimental de Altas Energías de la Universidad de Oviedo llamado **AnalysisPAF** ([48]). Con *framework* nos referimos al código que estructura análisis en física de partículas y que, en general, puede dar cabida al estudio de un proceso, o de más (como es el caso del *framework* que se ha empleado). Para hacer todos estos análisis, se ha construido desde cero el código específico para este estudio, aprovechando las ventajas (como las librerías de dibujado de gráficos) del *framework*, así como el desarrollo llevado a cabo anteriormente en un marco anterior de trabajo. El marco de trabajo también emplea **PAF** (**P***ROOF* **A***nalisis* **F***ramework*, [49], [50]): un conjunto de librerías y programas que permiten construir un *framework* modular y ordenado con el que realizar análisis en física de partículas desarrollado también en el Grupo Experimental de Altas Energías de la Universidad de Oviedo.

El código programado está pensado para ser ejecutado en los equipos del mismo grupo, así como en el *cluster* de servidores de computación que este tiene. Con este fin, se ha empleado también **Bash** para controlar todas las ejecuciones de los múltiples análisis y estudios realizados. Así mismo, **Python** también se ha usado para la obtención de los resultados (*yields*) de los análisis, y para manejar todo el código, así como las partes de este que eran compartidas, se ha usado el gestor de repositorios **git**. La plataforma empleada en todo momento estaba constituida de máquinas operando sistemas operativos **GNU/Linux**.

En las siguientes secciones de este capítulo describiremos cómo se ha construido el marco computacional, así como la selección de eventos general que se ha usado. Se ha tomado como referencia el análisis de

CMS del proceso $t\bar{t}H$ ([1]), que en general se ha intentado replicar, aunque en algunas partes se ha modificado. Debido a la naturaleza ya comentada de este análisis (que es un análisis enfocado a alta luminosidad), por la cual nos encontramos ante un proceso con bajos recuentos de eventos en estos momentos, se emplean en la actualidad técnicas de análisis de datos avanzadas en los esfuerzos de la Colaboración CMS. Nosotros no hemos empleado esas técnicas.

4.1. Datos y muestras de Monte Carlo

Los datos empleados han sido todos aquellos recogidos en el detector CMS durante el año 2016 a 13 TeV de energía (en centro de masas). El camino que toman esos datos, una vez han sido reconstruidos, así como las simulaciones también reconstruidas no es el de ser directamente analizados. Para la mayor parte de estudios, el tamaño de la información sigue siendo enorme, y no se necesita tanta información. Por consiguiente, estos datos y simulaciones sufren varios «recortes» o *skims*. Primero, se realiza en el *Tier-0* una preselección general, de los datos y MC reconstruidos, resultando en muestras conocidas como **AOD**. El tamaño sigue siendo muy grande, así que se hacen más selecciones en los *Tier-1* y *2* que dan como resultado una serie de muestras llamadas **miniAOD** ([51]). Y finalmente, para optimizar a un nivel mucho más específico según análisis, se hace una última selección obteniendo unos archivos de ROOT que se conocen como **n-tuplas** o **HeppyTrees** ([45] [52]). Estas muestras de datos y MC ya tienen un tamaño aceptable para ejecutar análisis en los *Tier-2* y *3*. En la figura 4.1 se puede ver un esquema de parte del proceso.

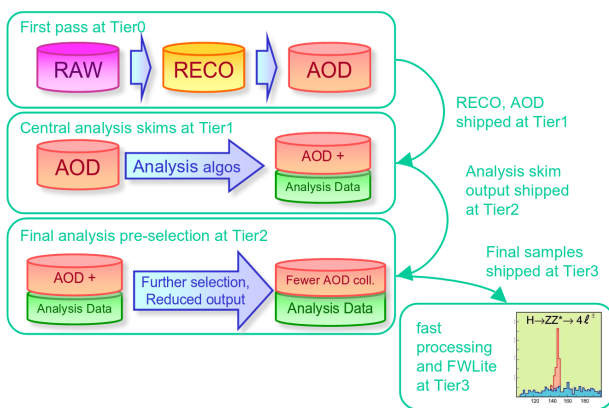


Figura 4.1: [53] Esquema del proceso de obtención de los datos *raw* (pre-reconstrucción) y parte de las preselecciones y recortes a los que son sometidos.

Las muestras de datos vienen agrupados en grupos, en función a la señal que se obtiene del *level-1 trigger* (de hecho, se consideran varios), de los cuales nosotros usaremos cinco: **SingleMuon**, **SingleElec**, **DoubleEG**, **DoubleMuon** y **MuonEG**.

Respecto a las simulaciones, se han tomado muestras de MC análogas a las que se hicieron en el análisis de CMS (no exactamente las mismas, puesto que las que hemos empleado han sido obtenidas aquí), intentando usar siempre las de mayor

número de eventos simulados para reducir la estadística. Además, el análisis de CMS estima algunos procesos de fondo a partir de datos reales (empleando subconjuntos de datos que se consideran que representan fielmente los mismos): nosotros no hemos hecho eso (por falta de tiempo). En su lugar, estimamos todos los fondos con simulaciones de MC.

Proceso físico	Modo de desintegración	Software de generación	σ (pb)	Clasificación
$t\bar{t}W$	$W \rightarrow l + \nu_l$	aMC@NLO	0.2043	$t\bar{t}W$
	$W \rightarrow q + q'$	aMC@NLO	0.4062	$t\bar{t}W$
$t\bar{t}Z$	$W^+W^- \rightarrow l + l' + \nu_l + \nu_{l'}$	aMC@NLO	0.2529	$t\bar{t}Z$
	$Z \rightarrow q + \bar{q}$	aMC@NLO	0.5297	$t\bar{t}Z$
WZ	$WZ \rightarrow l + l' + l'' + \nu_{l''}$	Powheg	4.4297	WZ
$W + \gamma$	$W \rightarrow l + \nu_l$	MadGraph	585.8	<i>conversions</i>
$Z + \gamma$	$Z \rightarrow l + \bar{l}$	aMC@NLO	131.3	<i>conversions</i>
$t + \gamma$		aMC@NLO	2.967	<i>conversions</i>
$t\bar{t} + \gamma$		aMC@NLO	3.697	<i>conversions</i>
ZZZ		aMC@NLO	0.0140	<i>rares</i>
WZZ		aMC@NLO	0.0557	<i>rares</i>
WWZ		aMC@NLO	0.1651	<i>rares</i>
WWW		aMC@NLO	0.2086	<i>rares</i>
$tZ + jets$	$Z \rightarrow l + \bar{l}$	aMC@NLO	0.0758	<i>rares</i>
$tttt$		aMC@NLO	0.0091	<i>rares</i>
ZZ		MadGraph	16.523	<i>rares</i>
WW		MadGraph	115	<i>fakes</i>
$W^+W^+ + jets$		MadGraph	0.0371	<i>fakes</i>
$t\bar{t} + jets$		aMC@NLO	831.76	<i>fakes</i>
$t\bar{t} + jets$		Powheg	831.76	<i>fakes</i>
$DY + jets$	$\gamma \rightarrow l + \bar{l} \ (m_{ll} \in [5, 50) \text{ GeV})$	aMC@NLO	22635.09	<i>fakes</i>
$DY + jets$	$\gamma \rightarrow l + \bar{l} \ (m_{ll} = 50 \text{ GeV})$	aMC@NLO	6025.2	<i>fakes</i>
$W + jets$	$W \rightarrow l + \nu_l$	MadGraph	61526.7	<i>fakes</i>
tW		Powheg	35.6	<i>fakes</i>
$\bar{t}W$		Powheg	35.6	<i>fakes</i>
t	Canal t	Powheg	136.02	<i>fakes</i>
$\bar{t}$	Canal t	Powheg	136.02	<i>fakes</i>
t	Canal s, $W \rightarrow l + \nu_l$	aMC@NLO	3.6806	<i>fakes</i>
WW	$WW \rightarrow l + l' + \nu_l + \nu_{l'}$	Powheg	12.178	<i>fakes</i>
$t\bar{t}H$	$H \rightarrow b\bar{b}$	Powheg	0.2151	$t\bar{t}H$

Cuadro 4.1: Muestras de simulaciones de Monte Carlo empleadas en este trabajo.

Las muestras de MC que hemos empleado están recogidas en el cuadro 4.1. Se ha intentado siempre emplear datos simulados que provengan de generadores que traten los procesos al mayor orden posible (NLO), pero en algunos de los estudios que hemos realizado el número de eventos seleccionados era tan pequeño, que nos encontrábamos con problemas cuando usábamos el generador aMC@NLO¹. Concretamente, se llegó a tener un número de eventos simulados con pesos negativos, debido a ello. Para solventarlo hicimos algunos cambios en las muestras. Dentro de lo posible, intentamos que fueran del generador Powheg, que trata órdenes NLO, pero si no decidimos bajar a LO. El *software* que simula las cascadas hadrónicas que siguen al proceso principal es Pythia8.

Todos los procesos se han clasificado en varios grupos. $t\bar{t}H$, en tanto señal, se considera aparte. Los procesos $t\bar{t}W$ y $t\bar{t}Z$, debido a la importancia de estos fondos, también, así como WZ se considera

¹Los mismos inconvenientes que mencionamos en el capítulo anterior

aparte por motivos de comparación con el análisis de CMS. El resto se agrupan en tres bloques: *fakes*, *conversions* y *rare*s. Los procesos que forman parte del primer subgrupo están en él debido a que si hemos encontrado un suceso de ese proceso con nuestra selección, hemos tenido que malidentificar, por ejemplo, un leptón de un *jet* como si viniera del proceso (y no esté relacionado con el *jet*) y que ha de tener al menos la misma carga que otro leptón. Otra posibilidad es que hayamos malinterpretado, en general, la carga de cualquier leptón de alguno de los estados finales del proceso. Por ejemplo, para el proceso WW , donde ambos bosones W decaen semileptónicamente, podríamos malinterpretar la carga de uno de ellos (que en principio ha de ser contraria por la naturaleza del proceso) y entonces tendríamos un suceso con un leptón *fake* que pasaría nuestro criterio de selección de dos leptones con misma carga.

Los procesos de *conversions* son aquellos en los que en general se dan fotones que podrían decaer en un par e^+e^- (una «conversión» o *conversion*) y que, en conjunción con las desintegraciones de, por ejemplo, un bosón W , nos pueda dar a un evento candidato a nuestra selección.

Finalmente, los procesos marcados con *rare*s se corresponden con aquellos que tienen estados finales que podrían pasar nuestra selección (sin necesidad de tener leptones *fakes* o *conversions*), pero cuya sección eficaz es relativamente baja.

4.2. Triggers

Una de las selecciones se hace escogiendo sucesos basándonos en qué requisito pasaron en el HLT para ser guardados los eventos. El HLT se encarga de comprobar, con toda la información del evento, si se cumple alguna de las muchas condiciones para guardar la información del mismo. Escoger los eventos en función a qué condiciones fueron las que permitieron que se almacenase el suceso en cuestión es una forma de preseleccionar eventos que permite descartar muchos que no nos interesan.

En este análisis *grosso modo* consideramos eventos que han sido recogidos debido a señales de uno, dos o tres o más leptones, así como señales de dos leptones con determinadas condiciones en ambos. Hay muchos «conjuntos de condiciones» que hacen referencia a esos supuestos, y todos ellos al final, cuando se consiguen las muestras de datos, están registrados como variables lógicas. Si un evento ha pasado determinado conjunto de condiciones, esa variable lógica será 1 y si no, 0. Nosotros agrupamos todos estos conjuntos de condiciones que usamos en cuatro grupos: $\mu\mu$, ee , $e\mu$ y $3l4l$, en función a los estados finales que aspiramos a reconocer en aquellos sucesos que pasen **alguno** de los conjuntos de condiciones de cada grupo. La lista de HLT² que hemos usado se muestra en la figura ??.

²Cada uno de los conjuntos de requisitos que constituyen un HLT se resume en una de esas variables lógicas que se conocen comúnmente como HLT *paths* o, sencillamente, HLT.

Cualquier evento que cumpla **alguno** de los conjuntos de condiciones de cada grupo diremos que pasa el *trigger* de nuestro análisis (pese a que *todos* han pasado el *trigger* de CMS, pues si no no estarían guardados). Eso sí, la comprobación de si cumplen o no estas exigencias se hace en función a los leptones seleccionados (véase más adelante la identificación en detalle). P. ej.: un evento con dos muones deberá de cumplir alguno de los conjuntos de condiciones del grupo de dos muones, y así análogamente. Las simulaciones de MC pasan también los *trigger*, de modo que se les pueda aplicar el mismo criterio de selección en función a qué HLT pasaron o cuál no.

Una cosa importante a tener en consideración cuando se tratan los *triggers* de datos es el conocido como **solapamiento** (*overlap*) de las muestras de datos que se obtienen una vez llegan a los físicos experimentales que hacen análisis con ellas. Los cinco archivos que usamos nosotros, **SingleMuon**, **SingleElec**, **DoubleMuon**, **DoubleEG** y **MuonEG** contienen eventos que al menos tienen un muon, un electrón, dos muones, dos electrones (o fotones, puesto que son indistinguibles para el L-1 *trigger*) y un muon y un electrón o fotón. Esto origina un problema a la hora de hacer una selección de eventos en función a qué *trigger* pasaron, puesto que, por ejemplo, un evento que tenga dos muones y un electrón (o fotón) podría estar tanto en la muestra **DoubleMuon** como en **DoubleEG**, con lo que si el evento acaba siendo seleccionado en ambos casos, aun exigiéndole que pase conjuntos de condiciones HLT distintos, o los mismos, entonces estaremos contando dos veces el mismo suceso: de ahí el «solapamiento» que existe entre todas estas muestras.

Para resolverlo, es decir, quitar este solapamiento u *overlap*, se crea una biyección entre las muestra de datos y los estados finales del evento que consideramos. En nuestro caso, estos dependerán de cuántos leptones identifiquemos de su sabor (porque estamos considerando *triggers* en función al número de leptones). De este modo, conseguimos evitar contar un mismo evento más de una vez.

	HLT <i>paths</i>
$\mu\mu$	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_v HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ_v HLT_IsoMu22_v HLT_TkIsoMu22_v HLT_IsoMu24_v HLT_IsoTkMu24_v
ee	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v HLT_Ele27_WPTight_Gsf_v HLT_Ele25_eta2p1_WPTight_Gsf_v HLT_Ele27_eta2p1_WPLoose_Gsf_v
$e\mu$	HLT_Mu23_TrkIsoVVL_Ele8_CaloIdL_TrackIdL_IsoVL_v HLT_Mu23_TrkIsoVVL_Ele8_CaloIdL_TrackIdL_IsoVL_DZ_v HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_v HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v HLT_IsoMu22_v HLT_TkIsoMu22_v HLT_IsoMu24_v HLT_IsoTkMu24_v HLT_Ele27_WPTight_Gsf_v HLT_Ele25_eta2p1_WPTight_Gsf_v HLT_Ele27_eta2p1_WPLoose_Gsf_v
$3l/4l$	HLT_DiMu9_Ele9_CaloIdL_TrackIdL_v HLT_Mu8_DiEle12_CaloIdL_TrackIdL_v HLT_Mu8_TripleMu_12_10_5_v HLT_Ele16_Ele12_Ele8_CaloIdL_TrackIdL_v HLT_Mu23_TrkIsoVVL_Ele8_CaloIdL_TrackIdL_IsoVL_v HLT_Mu23_TrkIsoVVL_Ele8_CaloIdL_TrackIdL_IsoVL_DZ_v HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_v HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_v HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ_v HLT_IsoMu22_v HLT_TkIsoMu22_v HLT_IsoMu24_v HLT_IsoTkMu24_v HLT_Ele27_WPTight_Gsf_v HLT_Ele25_eta2p1_WPTight_Gsf_v HLT_Ele27_eta2p1_WPLoose_Gsf_v

Cuadro 4.2: [1] Listado de los HLT *paths* empleados en el trabajo.

4.3. Identificación de objetos

En las siguientes subsecciones hablaremos de cómo caracterizamos los objetos de los eventos físicos. Todo análisis debe «definir» lo que considera un muon, un electrón o un *jet*, puesto que aunque el algoritmo PF nos devuelva al final reconstrucciones, estas están sujetas a posibles errores y quizás no nos interesen determinados eventos (p. ej., con leptones de p_T muy bajo) porque están sujetos a más fondo, u otros motivos.

4.3.1. Identificación de muones y electrones

La identificación de los muones y electrones se hace, al igual que el resto del análisis, tratando de imitar el análisis de CMS ([1]³). La caracterización de muones se puede observar en el cuadro 4.3 y la de los electrones en 4.4.

A diferencia de en otros análisis, en este seleccionamos tres tipos diferentes de leptones: leptones *loose*, leptones *fakeable*⁴ y leptones *tight*, en orden creciente de exigencias. Serán los elementos de la categoría *tight* los que consideremos como posibles leptones que vienen del proceso. Los leptones *fakeable* y *loose* se emplean para otros criterios de selección, o como categorías respecto a las cuales hacer «limpieza» (*cleaning*) de otros objetos reconstruidos. Estas categorías, al tener una selección menos exigente, tendrán más elementos en ellas. Por lo tanto, si hacemos criterios de selección con objetos de ellas en vez de con categorías de selección más estricta, estaremos cortando respecto a más elementos y, por tanto, estaremos haciendo un criterio de selección más exigente. Esto nos permite mejorar la eficacia de nuestros criterios de selección según nuestras necesidades.

La categoría *fakeable* está hecha de objetos muy similares a los *tight*, pero que no lo son por poco. Una de sus diferencias más significativas, es que no tienen el requisito en el criterio de selección del *Lepton MVA*. Por el contrario, el grupo de objetos *loose* es mucho más amplio que los otros dos debido a que las exigencias para pertenecer a él son bastante menores. Nótese (como se puede comprobar en los cuadros) que los grupos más estrictos de objetos son subconjuntos de los menos exigentes. Así, el conjunto de los muones *tight* es un subconjunto de los *fakeable*, y este lo es de los *loose* (lo mismo ocurre para los electrones).

Respecto a los criterios de selección en sí, las variables donde se hacen ya han sido brevemente explicadas en la sección anterior. Los cortes que están antes de la línea continua con el criterio de selección en $p_T > 5 \text{ GeV}$ son imperativos porque son las exigencias que se pusieron a las muestras que se usaron para entrenar el BDT del *Lepton MVA*. En este análisis, además, se hace una exigencia

³Existen diferencias entre el análisis real de CMS y el descrito en la nota interna (o *internal note*).

⁴Sería posible traducirlo por «falseables». Su significado se expresa en el siguiente párrafo.

Criterio de selección	Muon <i>loose</i>	Muon <i>fakeable</i>	Muon <i>tight</i>
p_T	$> 15 \text{ GeV}$	$> 15 \text{ GeV}$	$> 15 \text{ GeV}$
$ \eta < 2.4$	✓	✓	✓
$ d_{xy} < 0.05 \text{ cm}$	✓	✓	✓
$ d_z < 0.1 \text{ cm}$	✓	✓	✓
$\text{SIP}_{3D} < 8$	✓	✓	✓
$I_{\text{mini}} > 0.4$	✓	✓	✓
Muon <i>loose</i> de POG	✓	✓	✓
Muon <i>medium</i> de POG	—	—	✓
Disc. <i>Lepton MVA</i> > 0.90	—	—	✓
Disc. CSVv2 del <i>jet</i> más cercano	—	$< 0.8484 / < 0.3^*$	< 0.8484
$p_T^{\text{ratio}} > 0.5$	—	✓*	—
Compatib. de segmentos	—	✓*	—
<i>Tight-charge</i> (solo 2ISS)	—	—	≥ 1

Cuadro 4.3: Requisitos para la identificación de los muones *tight*, *fakeable* y *loose*. Los requisitos con asterisco (*) solo se aplican a aquellos objetos *fakeable* que no pasan el requisito del discriminante del *Lepton MVA*.

más para los leptones *tight* que pertenezcan a eventos de la categoría 2ISS, como se verá más adelante. Concretamente, es el requisito de *tight-charge*: esto hace que los leptones de los eventos de esta categoría comparada con las otras sean levemente distintos, además de otras posibles complicaciones en función a cómo se realice la separación en categorías, como veremos cuando describamos la selección de eventos.

4.3.2. Identificación de leptones tau

Cuando nos referimos aquí a los leptones tau, buscamos las reconstrucciones solo de τ_h , esto es, desintegraciones de estos leptones que den como resultado algún hadrón, como explicamos en el capítulo anterior. Los requisitos para la selección de nuestros τ_h son los siguientes: $p_T > 20 \text{ GeV}$, $|\eta| < 2.3$, WP⁵ medio del MVA contra $t\bar{t}$ (descrito en el capítulo anterior), identificación del modo de desintegración válida y WP *tight* de la identificación contra muones y *medium* contra electrones. Una vez pasan esos requisitos, se les exige una «limpieza» con todos los muones y electrones *loose* a una distancia de $\Delta R > 0.4$.

La caracterización de estos leptones que decaen hadrónicamente es única, a diferencia de la que hacemos para muones y electrones, donde tenemos tres subcategorías (*loose*, *fakeable* y *tight*). El motivo es que el único uso que le daremos será vetar la presencia de cualquier tau hadrónico que encontremos, puesto que tales eventos se tienen en cuenta en un análisis diferente ([54]) del multileptónico que nosotros tratamos.

4.3.3. Identificación de jets

La identificación de los *jets*, en general, se hace exigiéndoles $p_T > 25 \text{ GeV}$, $|\eta| < 2.4$ y un *cleaning* respecto a todos los muones y electrones *fakeable* bajo una distancia mínima de $\Delta R > 0.4$.

⁵ *Working point*, «referencia de trabajo». Un WP es un determinado criterio de selección que se estandariza. Se suelen clasificar en función a la exigencia que se haga al objeto (*loose*, medio, *tight*...).

Criterio de selección	Electrón <i>loose</i>	Electrón <i>fakeable</i>	Electrón <i>tight</i>
p_T	$> 15 \text{ GeV}$	$> 15 \text{ GeV}$	$> 15 \text{ GeV}$
$ \eta < 2.4$	✓	✓	✓
$ d_{xy} < 0.05 \text{ cm}$	✓	✓	✓
$ d_z < 0.1 \text{ cm}$	✓	✓	✓
$\text{SIP}_{3D} < 8$	✓	✓	✓
$I_{\text{mini}} > 0.4$	✓	✓	✓
Disc. de MVA ID $> (0.0, 0.0, 0.7)$	✓	✓	✓
$\sigma_{in\eta} < (0.011, 0.011, 0.030)$	–	✓	✓
$\frac{H}{E} < (0.10, 0.10, 0.07)$	–	✓	✓
$\Delta\eta_{in} < (0.01, 0.01, 0.008)$	–	✓	✓
$\Delta\varphi_{in} < (0.04, 0.04, 0.07)$	–	✓	✓
$-0.05 < \frac{1}{E} - \frac{1}{p} < (0.010, 0.010, 0.005)$	–	✓	✓
Disc. <i>Lepton MVA</i> > 0.90	–	–	✓
Rechazo de conversiones	–	–	✓
Número de <i>missing hits</i>	< 2	$= 0$	$= 0$
Disc. CSVv2 del <i>jet</i>	–	< 0.8484 / $< 0.3^\dagger$	< 0.8484
$p_T^{\text{ratio}} > 0.5$	–	✓ [†]	–
<i>Tight-charge</i> (solo 2ISS)	–	–	≥ 2

Cuadro 4.4: Requisitos para la identificación de los electrones *tight*, *fakeable* y *loose*. Los requisitos con daga (†) solo se aplican a aquellos objetos que no pasan la selección *tight*. Además, algunos de los criterios de selección se hacen en función al valor de η . Tales criterios se muestran dependiendo de una terna, en función al valor de η del leptón: $0 < |\eta| < 0.8$, $0.8 \leq |\eta| < 1.479$, $1.479 \leq |\eta| < 2.5$.

Respecto a los *b-jets*, en particular, se tienen en consideración dos *working point*: el *loose* y el *medium* del algoritmo CSVv2, que conllevan exigencias respectivas de > 0.5426 y > 0.8484 para el discriminante del citado algoritmo.

4.4. Los factores de escala (*scale factors*, SF)

Cuando se generan procesos físicos con un programa cualquiera, se simula el paso de los objetos por el detector, y se hace la reconstrucción de estos, obviamente, no se consiguen resultados perfectos. Aun en el supuesto de que nuestro *software* fuera en un momento dado casi perfecto, tendría otro fallo: la propia degradación de los materiales y componentes del detector por la radiación a la que están sometidos.

Por consiguiente, resulta obvio que la *eficacia* de las muestras de MC en lo que se refiere a obtener iguales resultados en un análisis con ellas y con datos, no es del 100% y, de hecho, no es en general despreciable. Con el fin de paliar esta diferencia, se emplean unos factores que escalan el peso de cada evento en función a las distintas fuentes de discrepancias que puede haber con los datos. Estos números se llaman, en un alarde de originalidad, **factores de escala** o *scale factors*, SF. Haremos una breve descripción de los que hemos usado en nuestros estudios a continuación. No son los únicos, de hecho otros, como los SF aplicados al proceso de *b-tagging* son muy comunes, pero estos no los hemos podido implementar por falta de tiempo.

Categoría	Factor de escala
ee	1.01 ± 0.02
$e\mu$	1.01 ± 0.01
$\mu\mu$	1.00 ± 0.01
$3l, 4l$	1.00 ± 0.03

Cuadro 4.5: [1] Listado de los factores de escala del *trigger* usados según la categoría donde se haya seleccionado el evento para el cual se aplica.

SF de leptones

Estos factores de escala dependen de la definición de los muones y electrones (de ahí que sean distintos para 2lSS y el resto de categorías, como justificaremos más tarde) y se encargan de paliar la diferencias en la identificación de los mismos. Dependen del p_T y del valor de η de cada leptón. Se adjunta la distribución de uno de los factores que se aplican a los leptones como muestra en la figura 4.2.

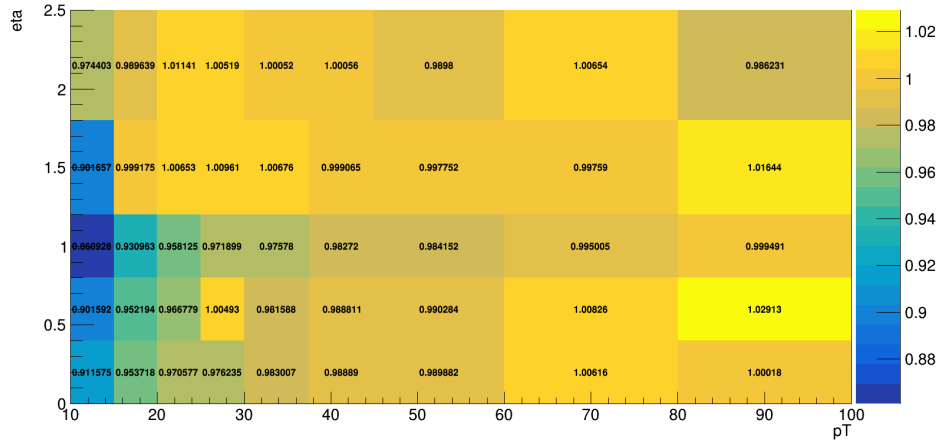


Figura 4.2: Gráfico bidimensional que muestra los valores de uno de los factores de escala (SF) aplicados a los leptones (concretamente, a muones seleccionados en la categoría 2lSS). Depende del valor de p_T (en GeV) y η . Para valores mayores de p_T que el máximo, se usa el factor de escala de $p_T = 100$ GeV.

SF de *trigger*

También existen pequeñas diferencias en la eficacia del *trigger*, y por ello se recurre a SF para reducirlas. En nuestro caso, se aplican SF simples según la categoría (2lSS por canales ee , $e\mu$ y $\mu\mu$ y $3l$ y $4l$ juntas), a diferencia de en otros análisis, donde se tienen variables del evento o de los objetos de este consideración para aplicarlos. Los hemos resumido en el cuadro 4.5.

SF de *pile-up*

Otro efecto a tener en consideración en las diferencias entre las simulaciones y los datos reales (que dependen del funcionamiento en el momento de la medición del LHC) es el *pile-up*. Precisamente debido a ello estos SF son independientes del análisis.

4.5. Selección de eventos

El *framework* identifica para cada evento todos los objetos que hemos enunciado hasta ahora y, cuando termina, decide si este evento nos interesa (porque se asemeja a lo que un evento donde se dé un proceso $t\bar{t}H$ multileptónico esperamos que sea). Para ello realizamos una serie de criterios de selección sucesivos en distintas variables y cantidades referidas tanto a los objetos del suceso, como al evento en sí. Este tipo de análisis se dice **secuencial** y también, en inglés, *cut & count* («corta y cuenta»).

Algunas de estas exigencias se hacen al principio, de modo general a todo suceso, y los eventos que las pasan después se categorizan. El análisis de CMS ([1]) detalla una subdivisión mucho más exhaustiva (hasta llegar a unas 20 categorías) que la que hacemos aquí, pero ello se debe a que en tal análisis se emplean más tarde herramientas de MVA (dos BDT bidimensionales) para refinar la relación señal-fondo. Nosotros no haremos tal cosa en nuestros estudios, de modo que nuestras categorías se reducen a tres: $2lSS$ (*two lepton same sign*), $3l$ y $4l$.

La primera, $2lSS$, recoge aquellos sucesos que nos resulten interesantes, que tengan solo dos leptones *tight*, y que además estos tengan la misma carga (lo cual garantiza una distinción clara de estos eventos). Frecuentemente nos referiremos a los canales $\mu\mu$, $e\mu$ y ee , que pertenecen a $2lSS$. $3l$ y $4l$ recogen los sucesos que nos interesen que tengan exactamente tres y cuatro o más leptones *tight*. Nótese que, precisamente, este criterio de selección en el número de leptones *tight* garantiza que las tres categorías sean ortogonales entre sí (no habrá ningún evento que pueda estar en las dos). A continuación describiremos las exigencias que hacemos en cada una de estas subdivisiones.

4.5.1. Preselección

- Todo evento debe pasar el *trigger* correspondiente en función al número de leptones *tight* y su sabor.
- El número de leptones *tight* debe ser mayor o igual que 2, con el fin de hacer un análisis *multileptónico*.
- Todo evento con un par de leptones *loose* cuya masa invariante sea menor que 12 GeV es rechazado, puesto que así eliminamos procesos que no nos interesan (p. ej., J/ψ), además de que los *software* de generación tienen problemas para hacer una buena simulación a energías bajas, por la escala de energía establecida de QED.
- Se exigen al menos dos *jets* identificados, puesto que esperamos en el proceso que analizamos que haya *quarks top* que decaigan a *quarks bottom* (lo cual ocurre en la inmensa mayoría de las veces, como comentamos en el primer capítulo). Además, para asegurarnos de que haya de estas

desintegraciones a *quarks* b , o bien un *jet* identificado pasa el WP medio del algoritmo CSVv2 de *b-tagging*, o dos pasan el WP *loose*.

4.5.2. 2ISS

- Debe haber exactamente 2 leptones *tight*.
- Uno de los leptones debe tener $p_T > 25$ GeV.
- Ambos deben poseer igual signo (en carga).
- Deben pasar el criterio de selección de *tight-charge* mencionado en las figuras 4.3 y 4.4 (en función al sabor de cada uno), para reducir las posibilidades de confusión con la carga. Esto es importante en esta categoría, porque esta se define en función a que tengan (o no) igual carga.
- Debe haber, al menos, cuatro *jets* identificados, puesto que si solo hemos encontrado dos leptones buenos (*tight*), entre el resto de objetos del estado final debe haber necesariamente más *quarks*: al menos, dos más.
- En el caso de que los dos leptones *tight* sean electrones, se aplican un par de criterios de selección más. Concretamente, se requiere que la masa invariante de ambos electrones no esté separada por menos de 10 GeV de la del bosón Z ($m_Z \simeq 91.2$ GeV [10]), ya que a diferencia de con muones, que somos capaces de medirlos mejor, hemos de garantizar que la carga de alguno de los dos electrones no haya sido mal medida y que en verdad procedan de la desintegración de un bosón Z . También se exige en este canal $METLD > 0.2^6$, puesto que un estado final de $t\bar{t}H$ que pertenezca a *2ISS* deberá tener (a consecuencia del desintegración desde un bosón W en leptón y neutrino, p. ej.) un valor mayor de E_T^{miss} que otros estados finales.

4.5.3. 3l

- Debe haber exactamente 3 leptones *tight*.
- Uno de los leptones debe tener $p_T > 25$ GeV.
- No debe haber ningún par de leptones *loose* con una masa invariante cercana en menos de 10 GeV a la masa del bosón Z , lo cual reduce el fondo de procesos que tienen un Z que decae a dos leptones.
- La suma de las cargas de los tres leptones debe ser -1 o 1 . De este modo, garantizamos tener dos leptones *tight* con el mismo signo de la carga, lo cual es algo típico de pocos estados finales (y así reducimos bastante fondo).

⁶Véanse distribuciones de esta variable en figuras del capítulo siguiente, como por ejemplo en la figura 5.3.

- Tampoco se admiten eventos que tengan dos pares de leptones *loose* con distinta carga pero del mismo sabor y de masa invariante menor de 140 GeV. De modo análogo al anterior criterio de selección, esto lo hacemos para librarnos del fondo de otros procesos con un Higgs decayendo en cuatro leptones.
- En aquellos eventos en los que haya menos (estrictamente) de cuatro *jets*, aplicamos el criterio de selección $\text{METLD} > 0.2$, salvo que exista algún par de leptones *loose* de carga opuesta y mismo sabor, en cuyo caso el criterio de selección es más escrito: $\text{METLD} > 0.3$. Este criterio de selección se hace igual que en la categoría anterior para eliminar fondo de otros procesos físicos con Z en ellos.

4.5.4. 4l

- Debe haber 4 o más leptones *tight*.
- Uno de los leptones debe cumplir $p_T > 25$ GeV.
- Al igual que antes, exigimos que ningún par de leptones *loose* tenga una masa invariante cercana a la del bosón Z en menos de 10 GeV.
- También exigiremos que no haya dos pares de leptones *loose* de carga distinta pero igual sabor que tengan una masa invariante menor que 140 GeV (para eliminar fondo relacionado con otros procesos con algún bosón de Higgs).

5 Resultados experimentales

*Proof is boring. Proof is tiresome. Proof is an irrelevance.
People would far rather be handed an easy lie than search
for a difficult truth, especially if it suits their own purposes.*

(El implacable Inquisidor Sand dan Glokta, en
Last Argument of Kings, saga *The First Law*,
JOE ABERCROMBIE)

Los frutos del análisis se recogen en este capítulo. En cada una de las secciones se exponen todos los resultados a los que se ha llegado a la hora de hacer el análisis del proceso $t\bar{t}H$, en el estudio de la identificación de leptones y en análisis tangenciales.

Las gráficas mostradas son histogramas, donde se ven las distribuciones de distintas variables de los eventos u objetos de estos. La clasificación de fondo, señal y datos se muestra en la leyenda y sigue los grupos definidos en el capítulo anterior. En los cuadros se puede apreciar el recuento de eventos¹, así como indicadores que permiten comparar categorías o ejecuciones del análisis con parámetros modificados entre sí.

Se muestra el cociente de la predicción (esto es, el fondo y la señal, que son simulados) respecto a los datos ($\frac{\text{Predicción}}{\text{Datos}}$) y el indicador $\frac{S}{\sqrt{F+(\Delta F)^2}}$, donde el número de eventos asociados a la señal (simulada) es S y al fondo, F . Este indicador mide la relación entre la señal y el fondo, que es importante para poder distinguir ambos y ser capaces de medir bien la sección eficaz. Si la relación señal-fondo fuera lo suficientemente baja, en función al análisis y a los fondos que este tenga, el recuento de eventos de señal podría ser inferior a la incertidumbre del fondo (cosa que sucede en este análisis), con lo que no necesitaríamos la señal para ajustar los datos a lo que el fondo nos dice. El hecho de usar $\frac{S}{\sqrt{F+(\Delta F)^2}}$ para tratar de optimizar los resultados y no simplemente $\frac{S}{F}$ es que el primer indicador se puede usar para optimizar las selecciones en función a la relación señal-fondo, pero teniendo también en cuenta la incertidumbre de este ([55]).

¹El criterio seguido para mostrarlos es el de indicar números naturales cuando el número de eventos lo permita, y si no, números reales con tres cifras decimales. Se recuerda que si bien el número de eventos que medimos ha de ser entero, no lo es así en el caso de la predicción simulada, pues todos los eventos se pesan por valores en general mucho menores que la unidad.

Muestra/Grupo de muestras	Incertidumbre
$t\bar{t}H$	+5.8 %
$t\bar{t}W$	-9.8 %
$t\bar{t}Z$	± 12 %
Resto	± 10 %
	± 30 %

Cuadro 5.1: [1] Incertidumbres asociadas a la normalización de las simulaciones empleadas. Las incertidumbres que no son de las muestras $t\bar{t}H$, $t\bar{t}W$ o $t\bar{t}Z$ se han establecido en un 30 %, debido a que no se estiman a partir de datos (como sí hacen en la referencia).

Tanto en histogramas, como en el cálculo de límites a secciones eficaces y en los resultados, se consideran incertidumbres estadísticas. Respecto a las sistemáticas, se han tenido en consideración las incertidumbres asociadas a los factores de escala de *trigger*, de los leptones y del *pile-up*, así como la incertidumbre existente al medir la energía de los *jets* debido a la escala de esta que miden los calorímetros hadrónicos (también conocido como la incertidumbre de JES, *jet energy scale*), que es distinta. Las incertidumbres en la normalización de las muestras, que se deben a una diferencia general entre los sucesos simulados y los reales de esos procesos, también se han considerado, aunque solo para cálculos de resultados y de límites a secciones eficaces, no para histogramas. Estas incertidumbres se recogen en el cuadro 5.1.

Las incertidumbres se muestran numéricamente en los recuentos de resultados. En los histogramas, una barra sombreada representa las asociadas a la predicción, y líneas negras las de los datos.

5.1. Análisis del proceso $t\bar{t}H$

En esta sección comentaremos los resultados fruto de la construcción que describimos en el capítulo anterior de nuestro análisis de $t\bar{t}H$. Para realizarlo, como ya indicamos, nos hemos basado en parte del análisis multileptónico realizado por la colaboración CMS ([1]).

Los resultados en número de eventos para todas las categorías se pueden consultar en el cuadro 5.2. Así mismo, mostramos en las figuras 5.1, 5.2, 5.3, 5.4 y 5.5 los histogramas de distintas variables de los eventos (y objetos de estos) que nos permiten entender cómo es la distribución de procesos. La clasificación de los procesos es la misma explicada en el capítulo anterior.

Atendiendo a los resultados en número de eventos y también a los histogramas, lo primero que es destacable es la distribución de eventos entre las categorías: la inmensa mayoría (en datos), el 78.9 %, pertenece a $2lSS$, un no despreciable 21.0 % a $3l$, y los pocos sucesos que quedan (un 0.03 %) son de $4l$. Esto entra en lo esperable (pues las categorías $3l$ y $4l$ exigen un mayor número de leptones que

Categoría	Fondo												Señal (tH)		Datos		Predicción Datos	S √ F+(ΔF)²
	ttW		ttZ		WZ		Conversions		Fakes		Rares							
	Valor	Δ	Valor	Δ	Valor	Δ	Valor	Δ	Valor	Δ	Valor	Δ	Valor	Δ	Valor	Δ		
Todas	166	22	94	11	23	8	53	19	295	103	28	9	658	109	79	9	752	27
2lSS	137	18	60	7	14	5	40	14	254	89	20	6	526	93	59	6	591	24
2lSS (μμ)	48	7	17	2	5	2	5	2	35	13	6	2	115	15	20	2	170	13
2lSS (ee)	22	3	12	2	3	1	14	6	64	22	4	1	118	23	10	1	115	11
2lSS (eμ)	68	9	31	4	7	2	20	8	155	61	11	3	293	62	29	3	306	17
3l	29	4	32	4	8	3	13	5	41	14	8	2	131	16	20	2	158	13
4l	0.008	0.007	1.454	0.229	–	–	–	–	–	–	0.093	0.030	1.555	0.231	0.634	0.090	3	2

Cuadro 5.2: Resultados en número de eventos del análisis del proceso $t\bar{t}H$ para las categorías $2lSS$, $3l$ y $4l$, los subcanales de $2lSS$ (ee , $\mu\mu$ y $e\mu$) así como el conjunto.

se da en menos estados finales). De todos modos, el número de sucesos en cualquier categoría es muy reducido (comparado con otros procesos, que pueden obtener recuentos de varios órdenes de magnitud por encima), siendo en el mejor de los casos del orden de 10^2 eventos.

El acuerdo entre datos y predicción es razonablemente bueno, como se ve en las gráficas y en el cociente $\frac{S}{F} = 0.981$. El fondo es mucho mayor que la señal, como vemos en la relación $\frac{S}{F} = 0.121$ (i.e., nuestra señal es solo el 12% de nuestro fondo) y con el valor del estimador $\frac{S}{\sqrt{F+(\Delta F)^2}} = 0.713$. Es esencial percatarse de que el error del fondo es mayor que nuestro recuento de eventos de señal, por lo que simplemente con el primero podemos explicar unos valores de datos idénticos a la predicción. Esto es otra muestra de nuevo de la peculiaridad de hacer este análisis con la cantidad de datos recogidos hasta la fecha. Se aprecia claramente también la predominancia de los fondos de *fakes*, $t\bar{t}W$ y $t\bar{t}Z$ sobre el resto, particularmente los *fakes* y $t\bar{t}W$.

Se ha intentado calcular la sección eficaz, pero debido a las condiciones del análisis y del experimento en la actualidad (pocos datos recogidos, proceso complejo, muchos fondos...) que ya hemos comentado, se obtiene un valor ($\sigma \sim 0.25$ pb) que tiene un error tal que, si bien «integra» el valor teórico esperado ($\sigma \simeq 0.22$ pb)², también se incluye dentro de él la sección eficaz nula y muchísimos otros valores. En definitiva, es un resultado muy impreciso.

Lo que sí hemos podido obtener es un límite superior al valor de la fuerza de señal (*signal strength*), que es el siguiente.

$$\mu_{\text{obs.}} = 3.21 \quad \mu_{\text{esp.}} = 2.04^{+0.85}_{-0.59} \quad (5.1)$$

El límite observado se encuentra dentro de los márgenes de 2σ del esperado, lo cual es aceptable.

²No se acerca a los 0.5 pb debido a que la muestra de señal no considera los estados finales en los que el Higgs se desintegra en $b\bar{b}$.

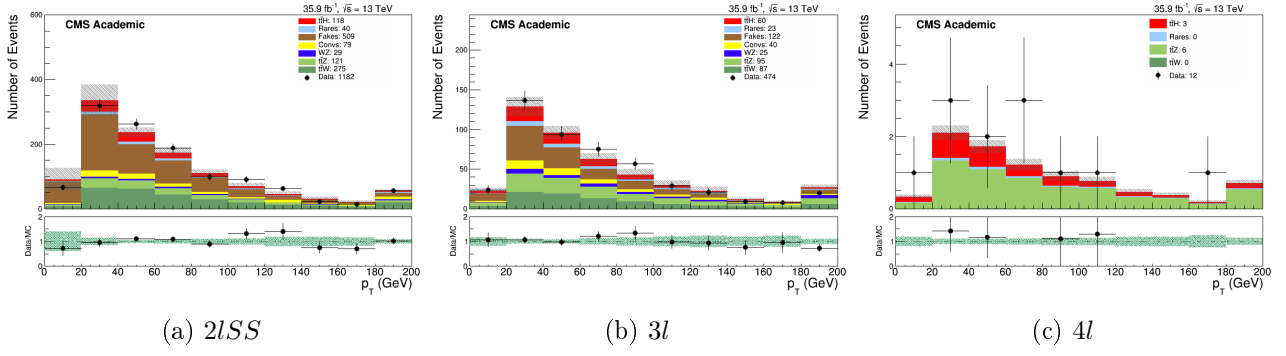


Figura 5.1: Histogramas del p_T de los primeros 2, 3 o 4 leptones (según los haya) para las categorías $2lSS$, $3l$ y $4l$ en el análisis de $t\bar{t}H$.

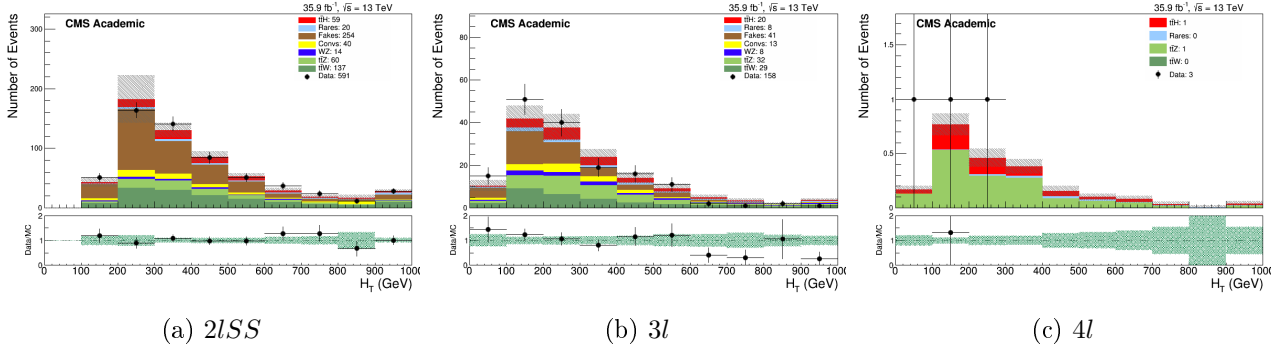


Figura 5.2: Histogramas de la variable H_T para las categorías $2lSS$, $3l$ y $4l$ en el análisis de $t\bar{t}H$.

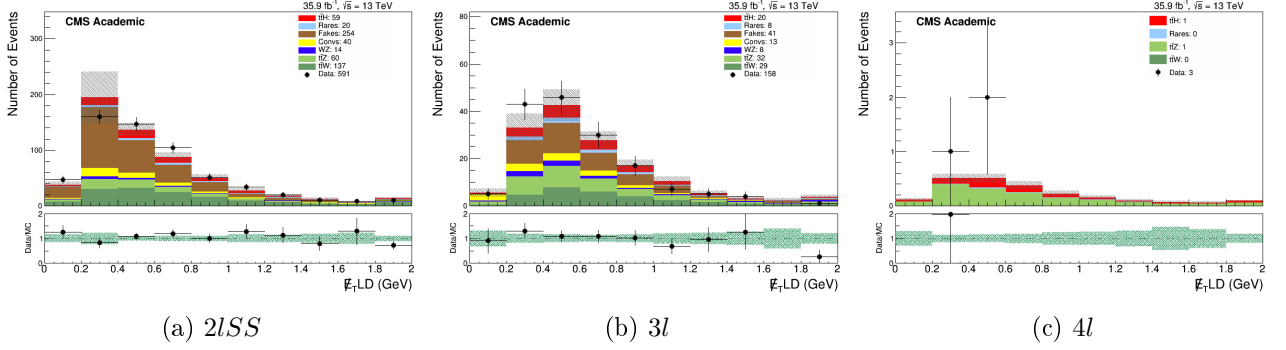


Figura 5.3: Histogramas de la variable METLD para las categorías $2lSS$, $3l$ y $4l$ en el análisis de $t\bar{t}H$.

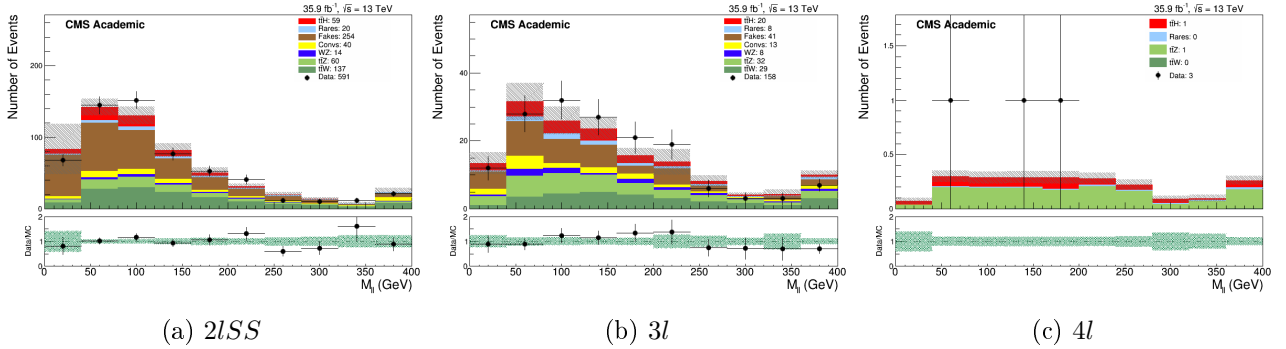
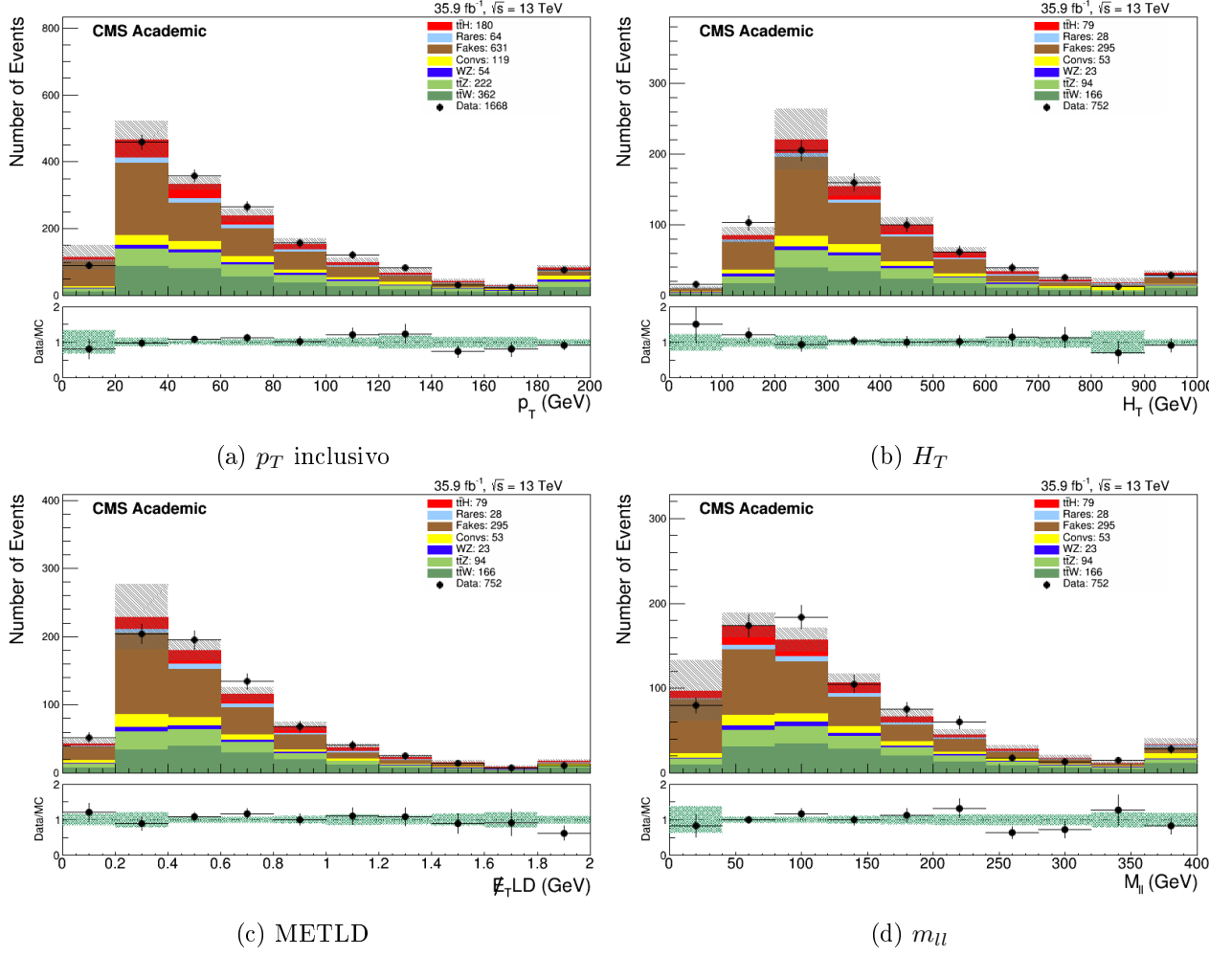


Figura 5.4: Histogramas de la masa invariante de los dos leptones seleccionados más energéticos para las categorías $2lSS$, $3l$ y $4l$ en el análisis de $t\bar{t}H$.

Figura 5.5: Histogramas de distintas variables para el conjunto de categorías del análisis $t\bar{t}H$.

5.2. Estudio de los criterios de identificación de leptones

5.2.1. Comparación con otros análisis

El primer estudio que hicimos para analizar la variable **Lepton MVA** fue comparar los resultados obtenidos del análisis de $t\bar{t}H$ con los conseguidos ejecutando el mismo análisis, pero con una identificación de leptones distinta, que no contuviera el **Lepton MVA**. Para ello, se hizo uso del código compartido en el repositorio común, y se implementaron las identificaciones de leptones para un análisis de procesos de supersimetría (relacionados con el *squark stop*) y la de un estudio del proceso $t\bar{t}$.

Por desgracia, en esos análisis no se emplean tres categorías de leptones como hacemos en el de $t\bar{t}H$, por lo que fue necesario, partiendo de las categorías ya existentes, construir las que faltaban en cada identificación. Estas nuevas definiciones se recogen en el cuadro 5.3.

Los resultados de la comparación se pueden observar en los cuadros 5.4 y 5.5 y en las figuras 5.6, 5.7, 5.8 y 5.9. Como se puede observar en comparación con los histogramas del capítulo anterior, el fondo de sucesos con leptones *fakes*, usando una identificación de estos sin **Lepton MVA**, es muy numeroso si se compara con los resultados de la sección anterior. Concretamente, se consigue un incremento del 357 % usando la identificación de leptones del análisis de *stop* y de hasta el 900 % para la del *top* (en todas las categorías) en el número de sucesos con leptones *fake*.

La relación señal-fondo, así como el estadístico $\frac{S}{\sqrt{F+(\Delta F)^2}} = 0.713$ empeoran mucho usando tanto las definiciones del análisis de *top* (0.025, 0.095) como las del de *stop* (0.050, 0.223), aunque claramente se ve algo más eficaz en discriminar señal del fondo la identificación de leptones del análisis de *stop*. La compatibilidad entre la predicción y los datos también se ve afectada, aunque algo menos en el caso de la identificación de *top*.

Si calculamos los límites esperados a la sección eficaz en este caso (calcularla directamente llevaría a resultados peores aún que los de la anterior sección), tenemos para la ejecución con la identificación de leptones del análisis de *top*

$$\mu_{\text{esp.}} = 5.27^{+2.79}_{-1.73}, \quad (5.2)$$

y, para la del *stop*:

$$\mu_{\text{esp.}} = 3.52^{+1.61}_{-1.08}, \quad (5.3)$$

Como vemos, los dos pares de límites son mucho mayores a los obtenidos en la sección anterior, así como sus incertidumbres.

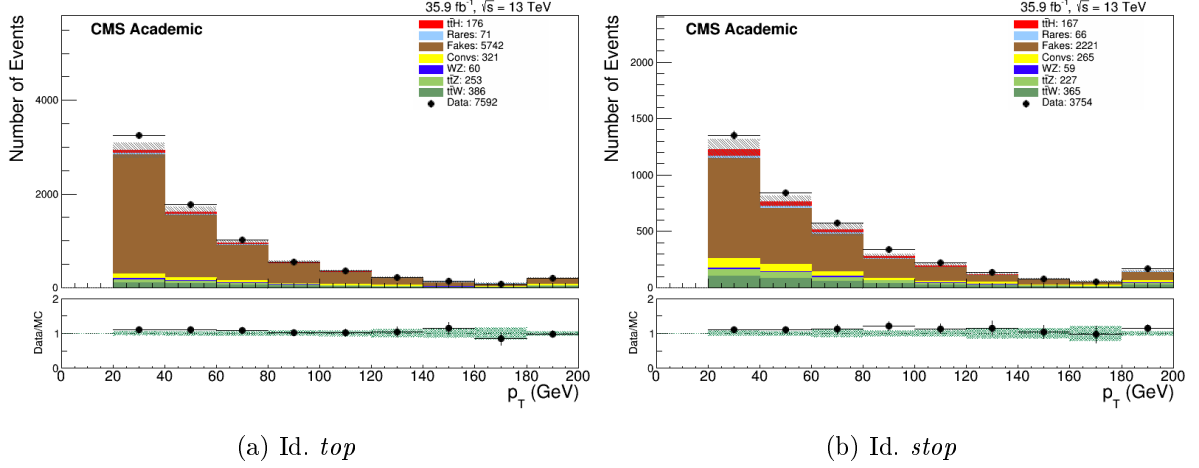


Figura 5.6: Histogramas del p_T de los primeros 2, 3 o 4 leptones (según los haya) para todas las categorías con la identificación de leptones de los análisis de *top* y *stop*.

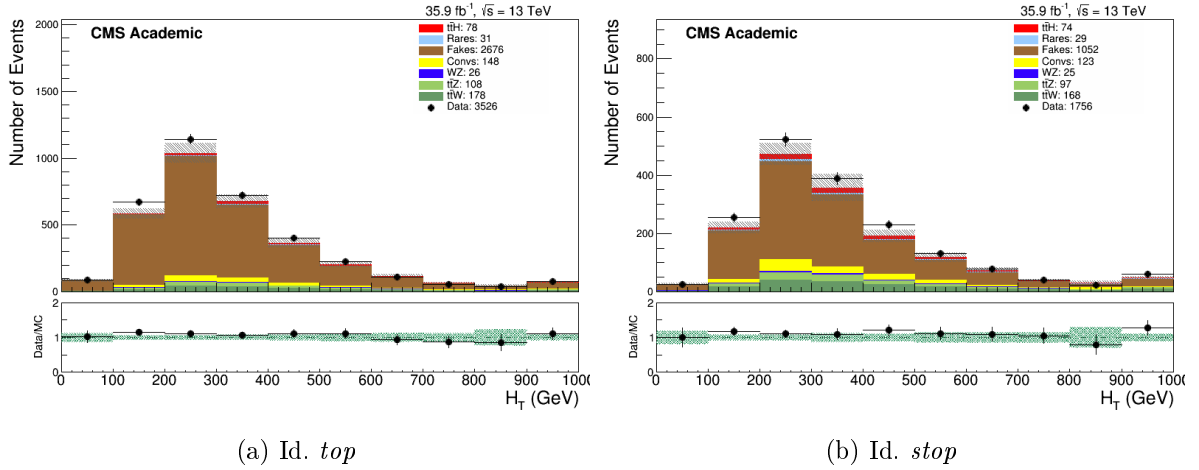


Figura 5.7: Histogramas de la variable H_T para todas las categorías con la identificación de leptones de los análisis de *top* y *stop*.

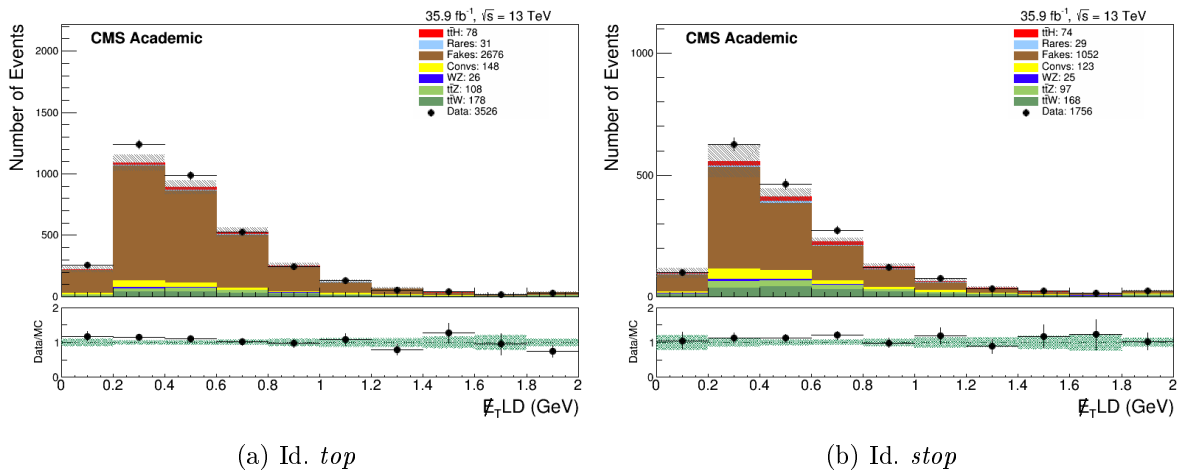


Figura 5.8: Histogramas de la variable METLD para todas las categorías con la identificación de leptones de los análisis de *top* y *stop*.

		Id. <i>stop</i>		Criterio de selección	Id. <i>top</i>	
		<i>Medium</i>	<i>Loose</i>		<i>Tight</i>	<i>Loose</i>
Leptones <i>tight</i>	μ	< 0.12		Id. muon (POG) <i>I</i>	< 0.25	
	e	personalizada < 0.12		Id. elec. (POG) <i>I</i>	$< 0.0588/0.0571^*$	
	Ambos	✓		$ \eta \notin [1.4442, 1.506]$	✓	
		–		<i>Missing hits</i> ≤ 1	✓	
		✓		$p_T > 20 \text{ GeV}$	✓	
		✓		$ \eta < 2.4$	✓	
Leptones <i>fakeable</i>	μ	<i>Medium</i> < 0.12		Id. muon (POG) <i>I</i>	<i>Medium</i> < 0.25	
	e	<i>Loose</i> $< 0.0695/0.0821^*$		Id. elec. (POG) <i>I</i>	<i>Medium</i> $< 0.0695/0.0821^*$	
	Ambos	✓		$ \eta \notin [1.4442, 1.506]$	✓	
		–		<i>Missing hits</i> ≤ 1	✓	
		✓		$p_T > 15 \text{ GeV}$	✓	
		✓		$ \eta < 2.4$	✓	
Leptones <i>loose</i>	μ	<i>Loose</i> < 0.4		Id. muon (POG) <i>I</i>	<i>Loose</i> < 0.4	
	e	<i>Loose</i> < 0.4		Id. elec. (POG) <i>I</i>	<i>Loose</i> < 0.4	
	Ambos	✓		$ \eta \notin [1.4442, 1.506]$	✓	
		–		<i>Missing hits</i> ≤ 1	✓	
		✓		$p_T > 5 \text{ GeV}$	✓	
		✓		$ \eta < 2.4$	✓	

Cuadro 5.3: Identificaciones de muones y electrones aplicadas para comparar con las de nuestro análisis. Las entradas con fondo gris son aquellas ya definidas en los estudios del *quark top* y del *squark stop*. Debido a que, en su conjunto, es más reducida esta identificación, se presenta solo en un cuadro. Los criterios de selección en el aislamiento (*I*) que tienen un asterisco se deben a los valores aplicados en función a $|\eta| \leq 1.479$ y $|\eta| > 1.479$. El criterio de «Buen vértice» se corresponde con exigencias en las variables d_{xy} y d_z . Para electrones, el criterio de selección del WP *tight* depende del valor de $|\eta|$: $|\eta| \leq 1.479 \Rightarrow d_{xy} < 0.05 \text{ cm} \wedge d_z < 0.1 \text{ cm}$, $|\eta| > 1.479 \Rightarrow d_{xy} < 0.10 \text{ cm} \wedge d_z < 0.2 \text{ cm}$. En cuanto a los muones, los cortes son, para el criterio de selección *medium* $d_{xy} < 0.2 \text{ cm} \wedge d_z < 0.5 \text{ cm}$, y para el *tight* $d_{xy} < 0.05 \text{ cm} \wedge d_z < 0.1 \text{ cm}$.

Categoría	Fondo												Señal (tH)		Datos		Predicción Datos	$\frac{S}{\sqrt{F+(\Delta F)^2}}$			
	ttW		ttZ		WZ		Conversions		Fakes		Rares								Total		
	Valor	Δ	Valor	Δ	Valor	Δ	Valor	Δ	Valor	Δ	$\frac{Fakes}{Fakes+tH} (\%)$	Valor	Δ	Valor	Δ	Valor			Δ	Valor	Δ
Todas	178	22	108	12	26	8	148	48	2676	820	907	31	10	3168	822	78	8	3526	59	0.921	0.095
Todas (tH)	166	22	94	11	23	8	53	19	295	103	100	28	9	658	109	79	9	752	27	0.981	0.713
2lSS	149	19	74	8	17	5	124	40	2287	701	899	23	7	2674	703	59	6	2988	55	0.915	0.083
3l	29	4	33	4	9	3	25	8	389	121	959	8	2	492	121	19	2	536	23	0.953	0.152
4l	0.034	0.015	1.638	0.245	–	–	0.058	0.045	–	–	–	0.361	0.304	2.091	0.393	0.675	0.093	2	1	1.383	0.450

Cuadro 5.4: Resultados en número de eventos del análisis del proceso $t\bar{t}H$ para las categorías 2lSS, 3l y 4l, los subcanales de 2lSS (ee , $\mu\mu$ y $e\mu$) así como el conjunto con las identificaciones de leptones (μ y e) del análisis del *quark top*. Se añade el total del análisis de $t\bar{t}H$ estándar para comparación (en las filas asociadas a las categorías, el porcentaje se refiere sobre el recuento de sucesos que el análisis estándar de $t\bar{t}H$ tiene en tales categorías).

Categoría	Fondo														Señal (tH̄)		Datos		Predicción Datos	S √(F+ΔF)²	
	ttW		ttZ		WZ		Conversions		Fakes		Rares		Total								
	Valor	Δ	Valor	Δ	Valor	Δ	Valor	Δ	Valor	Δ	Fakes FakesttH̄ (%)	Valor	Δ	Valor	Δ	Valor	Δ	Valor			Δ
Todas	168	21	97	10	25	8	123	41	1052	327	357	29	9	1495	331	74	8	1756	42	0.894	0.223
Todas (ttH)	166	22	94	11	23	8	53	19	295	103	100	79	9	28	9	658	109	752	27	0.981	0.713
2lSS	141	17	65	7	16	5	105	35	934	290	367	22	7	1283	293	56	6	1517	39	0.883	0.190
3l	28	4	31	4	9	3	18	6	118	38	290	7	2	210	39	18	2	236	15	0.964	0.425
4l	0.009	0.010	1.483	0.224	–	–	–	–	–	–	–	0.354	0.301	1.847	0.375	0.619	0.086	3	2	0.822	0.439

Cuadro 5.5: Resultados en número de eventos del análisis del proceso $t\bar{t}H$ para las categorías $2lSS$, $3l$ y $4l$, los subcanales de $2lSS$ (ee , $\mu\mu$ y $e\mu$) así como el conjunto con las identificaciones de leptones (μ y e) del análisis del *squark stop*. Se añade el total del análisis de $t\bar{t}H$ estándar para comparación (en las filas asociadas a las categorías, el porcentaje se refiere sobre el recuento de sucesos que el análisis estándar de $t\bar{t}H$ tiene en tales categorías).

5.2.2. Comparación con distintos criterios de selección de la variable Lepton MVA

El segundo estudio que hemos hecho con esta variable ha sido comparar los resultados obtenidos cuando cortábamos de forma distinta con ella. Para ello, nos hemos servido de guía de varios WP ([39]) que hay definidos para la misma. Los criterios de selección que hemos usado son los siguientes.

- (Añadido) > 0.97 .
- (Añadido) > 0.95 .
- $t\bar{t}H$: > 0.90 .
- *Extra tight*: > 0.90 (μ), > 0.85 (e).
- *Very tight*: > 0.45 (μ), > 0.75 (e).
- *Tight*: > 0.15 (μ), > 0.65 (e).
- *Medium*: > -0.20 (μ), > 0.50 (e).

Esperamos, por cómo está construida la variable, que en las selecciones menos estrictas, el número de eventos del grupo de procesos *fake* aumente. En el cuadro 5.7 se pueden apreciar los resultados, al igual que en el resto de categorías, pero con una columna más que muestra el porcentaje sobre el número de sucesos de los procesos del grupo de muestras *fakes* del criterio de selección en **Lepton MVA** del análisis de $t\bar{t}H$. Los histogramas de las figuras 5.10, 5.11, 5.12, y 5.13 permiten corroborar los resultados.

El avance del grupo de procesos *fakes* es muy significativo. Claramente se aprecia un aumento en esa columna al ir disminuyendo la exigencia del criterio de selección, hasta llegar al extremo del 508 % de lo que obtenemos con el criterio de selección de $t\bar{t}H$ normal, o hasta bajar al 4 % en la selección más estricta (*tight*) (considerando las tres categorías, i.e., 2lSS, 3l y 4l).

La concordancia entre la predicción y los datos obtenidos oscila un poco en los distintos criterios de selección, pero se mantiene más o menos aceptable. En cuanto a la relación señal-fondo, esta se mejora en los cortes más estrictos y se empeora en los más suaves, lo que confirma el comportamiento que

Criterio de selección	$\mu_{\text{esp.}}$
> 0.97	$1.60^{+0.67}_{-0.46}$
> 0.95	$1.93^{+0.79}_{-0.55}$
$t\bar{t}H$	$2.04^{+0.85}_{-0.59}$
<i>Extra tight</i>	$2.27^{+0.97}_{-0.67}$
<i>Very tight</i>	$2.51^{+1.13}_{-0.75}$
<i>Tight</i>	$2.84^{+1.30}_{-0.86}$
<i>Medium</i>	$3.01^{+1.43}_{-0.95}$

Cuadro 5.6: Límites superiores obtenidos en función al criterio de selección hecho en la variable **Lepton MVA** (se incorpora el hecho con el criterio de selección de $t\bar{t}H$ usual para comparación).

Criterio de selección	Fondo					Señal ($t\bar{t}H$)		Datos		$\frac{\text{Predicción}}{\text{Datos}}$	$\frac{S}{\sqrt{F+(\Delta F)^2}}$
	$Fakes$			Total							
	$Valor$	Δ	$\frac{Fakes}{Fakes(t\bar{t}H)} (\%)$	$Valor$	Δ	$Valor$	Δ	$Valor$	Δ		
> 0.97	14	7	5	246	22	46	5	363	19	0.804	1.694
> 0.95	177	71	60	471	75	62	7	524	23	1.017	0.788
$t\bar{t}H$	295	103	100	658	109	79	9	752	27	0.981	0.731
$Extra\ tight$	517	168	175	938	173	96	10	1056	32	0.979	0.545
$Very\ tight$	751	238	255	1200	242	102	11	1338	37	0.973	0.417
$Tight$	1032	323	350	1500	326	107	12	1624	40	0.990	0.325
$Medium$	1498	466	508	1988	468	111	12	2074	46	1.012	0.235

Cuadro 5.7: Resultados en número de eventos del análisis del proceso $t\bar{t}H$ para todas las categorías según el criterio de selección realizado en la variable **Lepton MVA**.

esperaríamos de ella. El estimador que usamos muestra un funcionamiento similar, resultando como destacado el valor del criterio de selección más estricto, que sobresale (para el conjunto de las categorías). Los valores de los límites esperados que podemos calcular se muestran en el cuadro 5.6 y vemos que los criterios de selección más estrictos tienen valores esperados mayores.

Se puede observar claramente que, atendiendo a estos límites, parecería lógico utilizar un WP más *tight* para estos casos. El hecho de que se haya escogido el establecido por la Colaboración CMS no atiende solo al valor de estos límites (que también). En un experimento como CMS y bajo las condiciones actuales del LHC, donde el proceso $t\bar{t}H$ tiene un fondo de *fakes* muy importante y depende mucho de la precisión de la medida de los fondos, se decidió dar prioridad a cuidar la cantidad de estos *fakes* sobre el número total de cuentas. Si bien no es nuestro caso, en la Colaboración CMS los fondos de *fakes* se estiman a partir de datos reales (además de desde MC), lo cual es más fiable que obtenerlos solo de simulaciones. Se vio que entonces, el WP de $t\bar{t}H$ reducía más la cantidad de eventos con leptones *fakes*, de modo que se escogió ese criterio de selección. Además, un criterio de selección más estricto conllevaría un número de sucesos mucho menor (como se aprecia en el cuadro 5.7) y esto podría dar lugar a fluctuaciones estadísticas en los resultados.

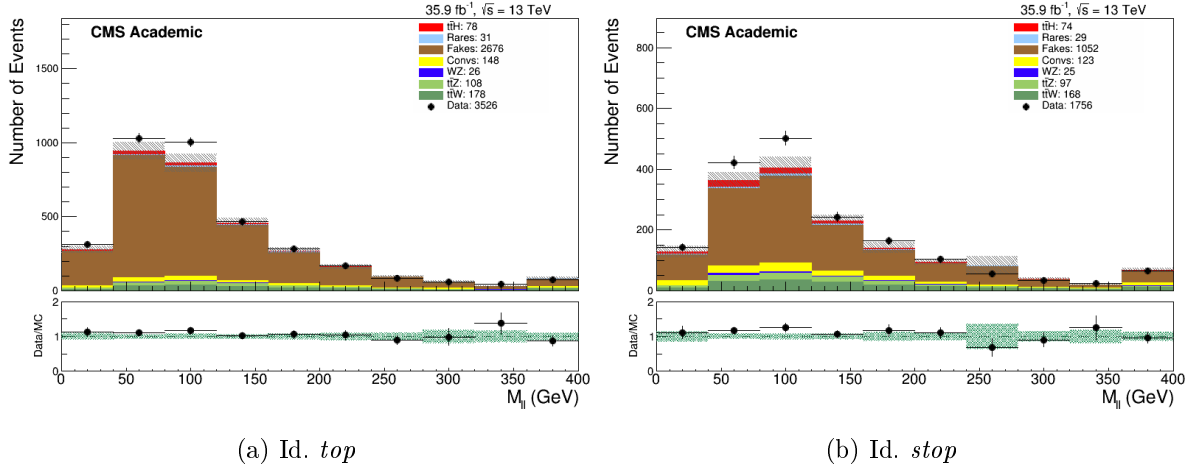


Figura 5.9: Histogramas de la masa invariante de los primeros dos leptones para todas las categorías con la identificación de leptones de los análisis de *top* y *stop*.

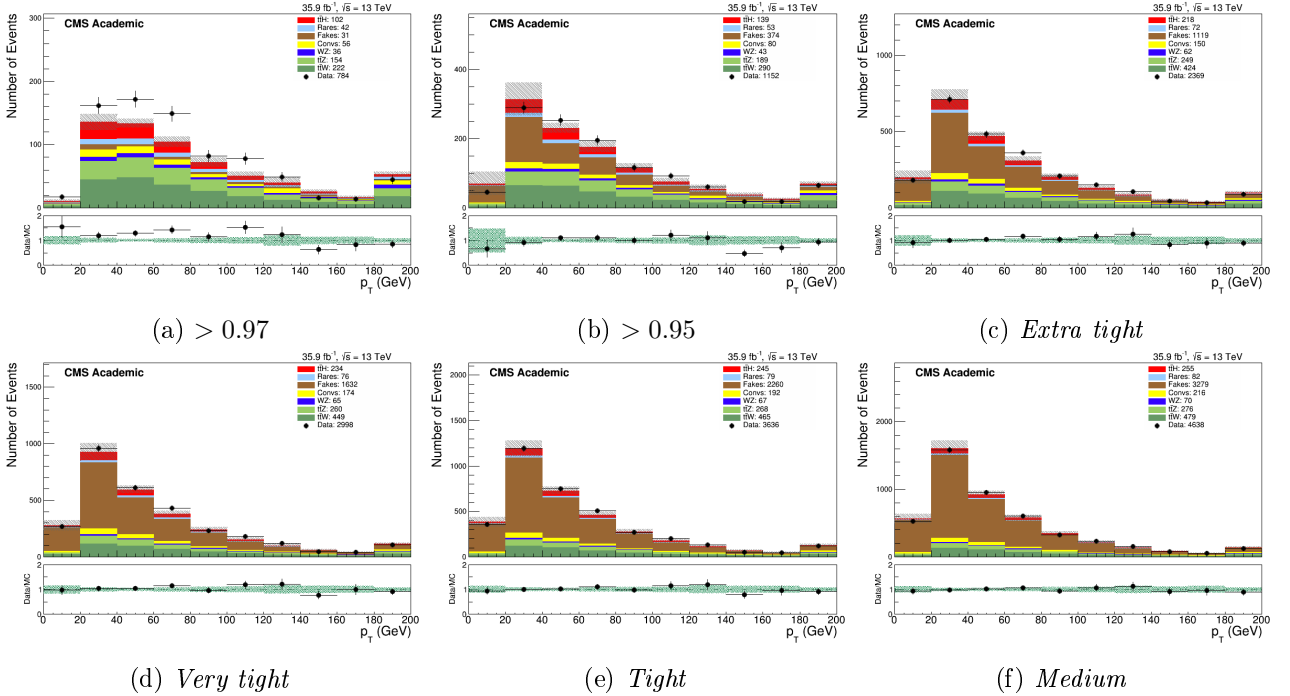


Figura 5.10: Histogramas del p_T de los primeros 2, 3 o 4 leptones (según los haya) para todas las categorías con criterios de selección en los WP considerados de la variable **Lepton MVA**.

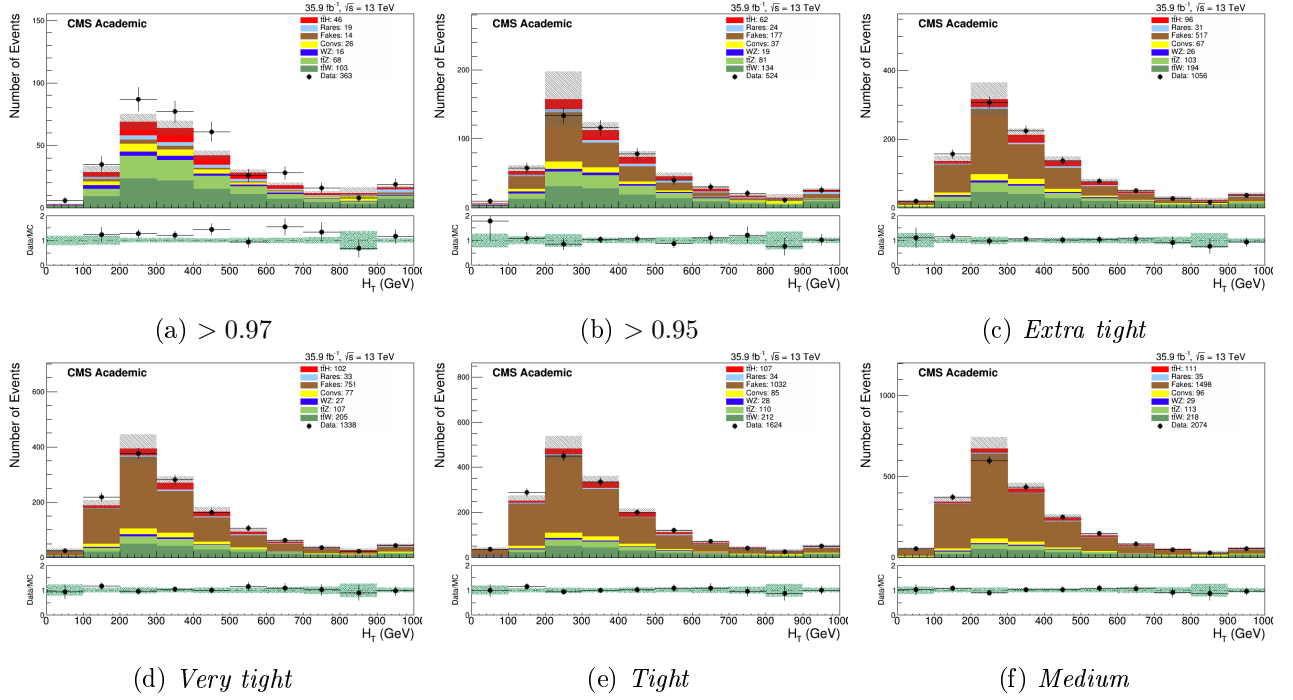


Figura 5.11: Histogramas de la variable H_T para todas las categorías con criterios de selección en los WP considerados de la variable Lepton MVA.

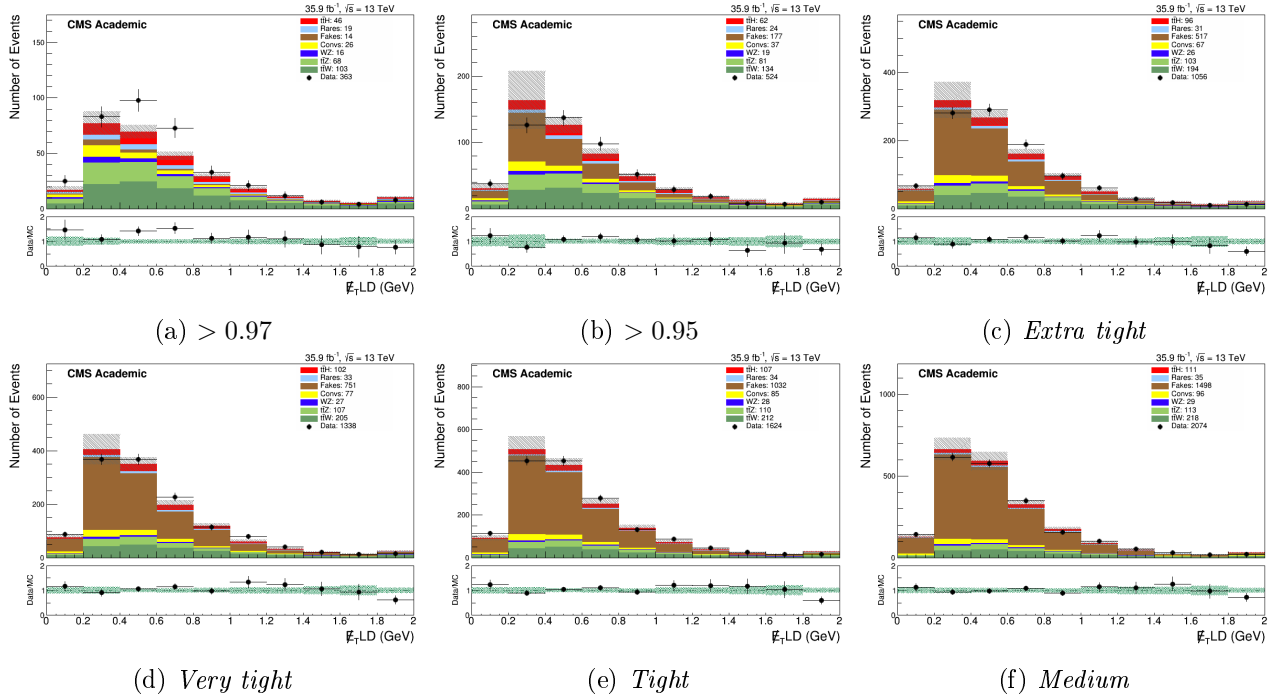


Figura 5.12: Histogramas de la variable METLD para todas las categorías con criterios de selección en los WP considerados de la variable Lepton MVA.

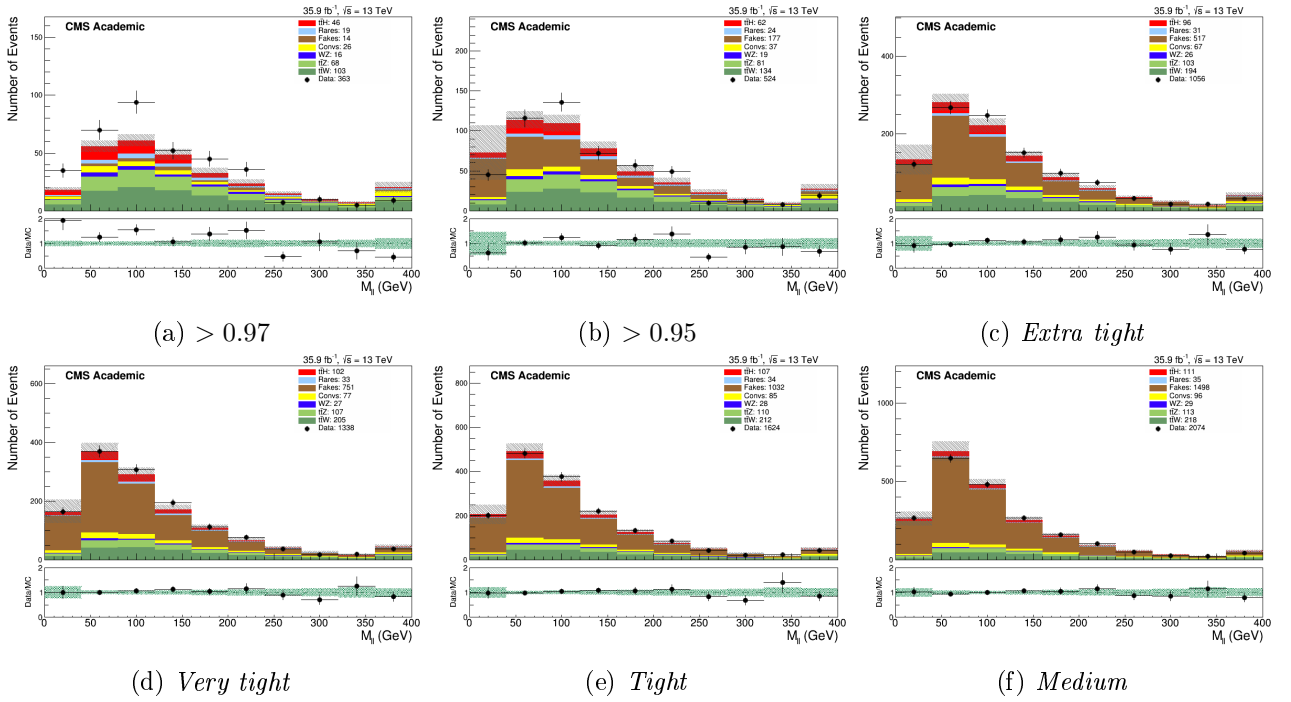


Figura 5.13: Histogramas de la masa invariante de los dos leptones seleccionados más energéticos para todas las categorías con criterios de selección en los WP considerados de la variable **Lepton MVA**.

5.3. Otras comparaciones y estudios

5.3.1. Cambios en la selección de leptones tau

A la hora de implementar nuestro análisis, realizamos una pequeña modificación respecto al análisis «base» de CMS consistente en hacer una identificación de los leptones tau que decaen hadrónicamente más estricta que la del análisis, con lo que el veto en nuestra selección a la presencia de esos leptones tau es más suave. Concretamente, en el análisis de CMS no se corta en las variables entrenadas contra la identificación de electrones ni muones.

De este modo, hemos obtenido una relación señal - fondo mejor, y también considerando el tamaño del fondo y el error de este, como podemos apreciar en los resultados de el cuadro 5.8, comparándolos con los de 5.2. No se aprecia diferencia importante en la concordancia entre predicción y datos, aunque sí una mejoría en el estadístico que consideramos (del 1.2%). El límite esperado con la identificación de leptones tau de la Colaboración CMS es

$$\mu_{\text{esp.}} = 2.05^{+0.87}_{-0.59}, \quad (5.4)$$

que si bien tiene iguales errores, es levemente mayor, lo cual está de acuerdo con la mejora en el estadístico.

5.3.2. Comparación de generadores aMC@NLO y Powheg para el proceso $t\bar{t} + \text{jets}$.

Durante la elaboración del análisis y las comparaciones, nos hemos visto obligados a cambiar de muestras de simulaciones de Monte Carlo. Los motivos en general se pueden reducir a que, al modificar nuestro análisis (o sin hacerlo), la selección era tan estricta que el número de sucesos de la muestra, sin pesarlos, era muy pequeño. Esto origina que la incertidumbre estadística se incremente, y además, en función al tipo de generador, puede tener mayores o menores efectos sobre el resultado (por ejemplo, si el generador es aMC@NLO y hay un recuento bajo de sucesos simulados no pesados). Por ello, consideramos apropiado hacer una comparación entre los dos generadores a NLO que hemos usado: Powheg y aMC@NLO con nuestro análisis, observando qué diferencias había entre las distribuciones de ambas muestras, y si

Categoría	Fondo												Señal (ttH)		Datos		Predicción Datos	$\frac{S}{\sqrt{F+(\Delta F)^2}}$
	ttW		ttZ		WZ		Conversions		Fakes		Rares							
	Valor	Δ	Valor	Δ	Valor	Δ	Valor	Δ	Valor	Δ	Valor	Δ	Valor	Δ	Valor	Δ		
Todas	165	22	91	11	22	7	52	18	294	103	27	8	651	108	78	9	743	27
Todas (nuestra id.)	166	22	94	11	23	8	53	19	295	103	28	9	658	109	79	9	752	27
2ISS	136	18	58	7	14	5	39	14	253	89	20	6	520	92	58	6	582	24
3l	29	4	31	4	8	3	13	5	41	14	8	2	130	16	20	2	158	13
4l	0.008	0.007	1.454	0.229	-	-	-	-	-	-	0.093	0.030	1.556	0.231	0.634	0.090	3	2
																	0.730	0.500

Cuadro 5.8: Resultados en número de eventos del análisis del proceso $t\bar{t}H$ para todas las categorías, con una identificación de leptones tau análoga a la realizada por la Colaboración CMS ([1]). Se agrega el total de categorías con nuestra identificación para comparación, el resto se pueden cotejar con los del cuadro 5.2.

Criterio de selección	Powheg		aMC@NLO	
	Valor	Δ	Valor	Δ
> 0.97	69	22	68	32
> 0.95	127	40	118	51
$t\bar{t}H$	237	74	275	100

Cuadro 5.9: Resultados en número de eventos del análisis de $t\bar{t}H$ aplicado exclusivamente a las muestras del proceso $t\bar{t} + jets$ con generadores **aMC@NLO** y **Powheg** para todas las categorías.

podrían afectar a nuestros resultados. Hemos escogido la muestra del proceso $t\bar{t} + jets$, que es una de las que más fondo aporta.

Los resultados se pueden apreciar en el cuadro 5.9, donde vemos que las diferencias claramente existen, sobre todo en el criterio de selección de $t\bar{t}H$, si bien se mantienen dentro de la incertidumbre de 1σ . En las figuras 5.14, 5.15 y 5.16 podemos ver las comparaciones en varias variables de las dos muestras de MC los WP de la variable **Lepton MVA** más estrictos (hemos dibujado estos WP porque son los que menos eventos recogen). En algunas columnas la diferencia es sustancial, y esta depende bastante del ancho que se asigne a cada columna o *bin*³.

Para comparar las dos distribuciones de forma cuantitativa, hemos aplicado un test de Pearson (χ^2) que compara la forma de ambas. Poniendo como ejemplo la variable *METLD*, hemos obtenido unos valores del estadístico χ^2 (dividido por el número de grados de libertad) para los WP de 2.68 ($t\bar{t}H$), 2.76 (> 0.95) y 1.86 (> 0.97). Cuando el valor se aproxima a la unidad, las distribuciones serán más parecidas y cuando se aleje, más dispares. Hemos comprobado que los indicadores dependen mucho de la variables escogida, y del WP, así como de la distribución de columnas o *bines* de los histogramas, pero no hay en general valores muy cercanos a la unidad (si se observan los valores del estimador, dividido por el número de grados de libertad, de todas las distribuciones). Todo esto es producto del reducido número de sucesos sin pesar. Otro ejemplo de cómo afecta a las simulaciones es que, en el canal $\mu\mu$, se obtienen para el WP 0.95 unos -0.98 sucesos simulados pesados de la muestra de **aMC@NLO**, lo cual no ocurriría si se tuvieran más eventos sin pesar al terminar la selección.

³De hecho, hemos tenido especial cuidado al distribuir los *bines* de los histogramas de la comparación para evitar en la medida de lo posible columnas sin eventos de alguna muestra.

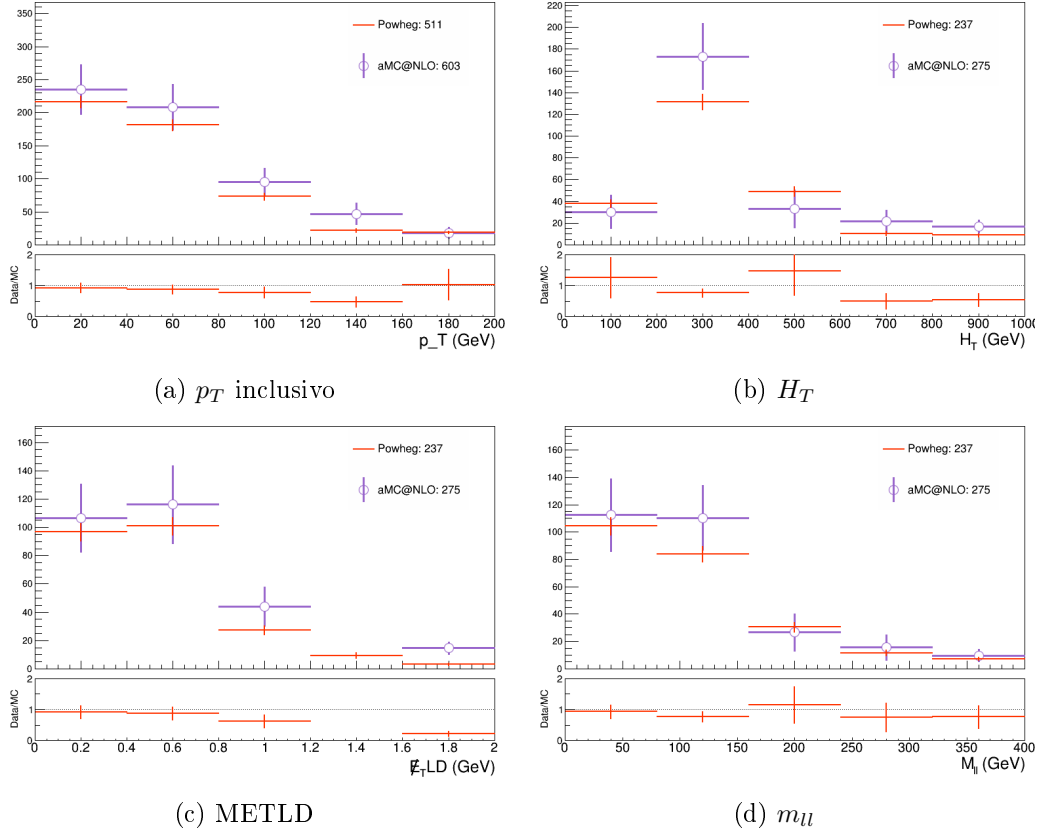


Figura 5.14: Histogramas de varias variables para el criterio de selección en Lepton MVA del análisis $t\bar{t}H$ que muestran la comparación entre las muestras de Powheg y aMC@NLO.

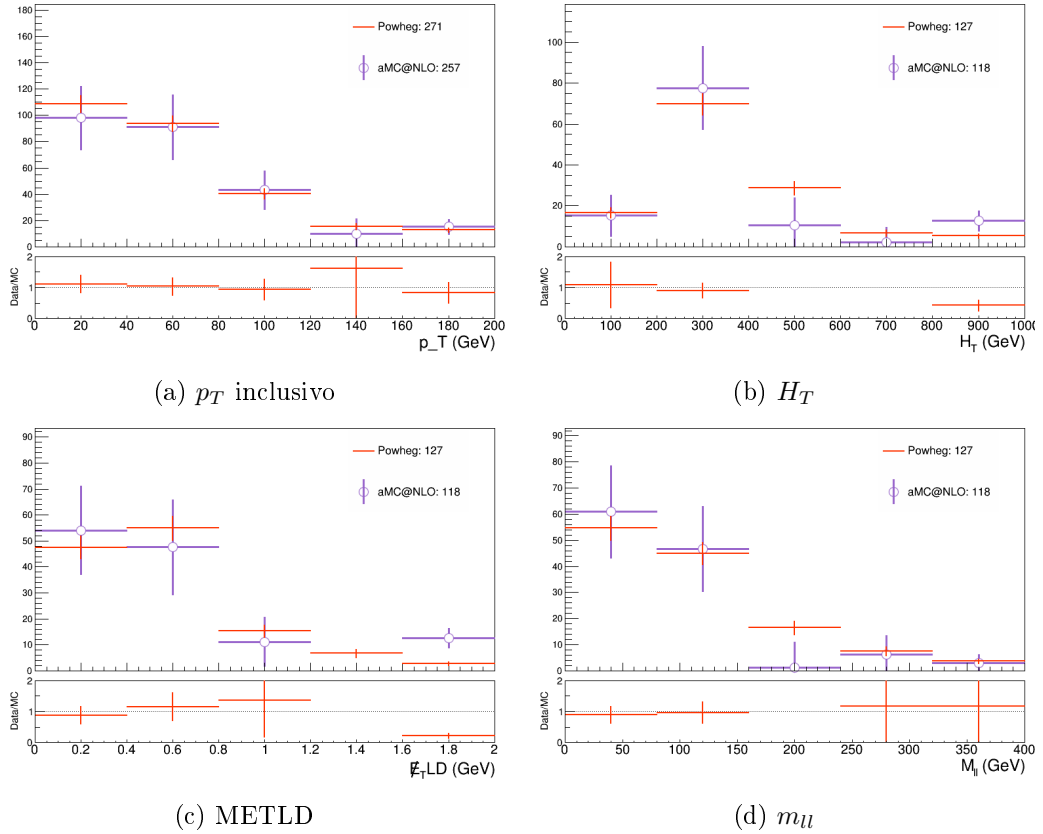


Figura 5.15: Histogramas de varias variables para el criterio de selección en Lepton MVA > 0.95 que muestran la comparación entre las muestras de Powheg y aMC@NLO.

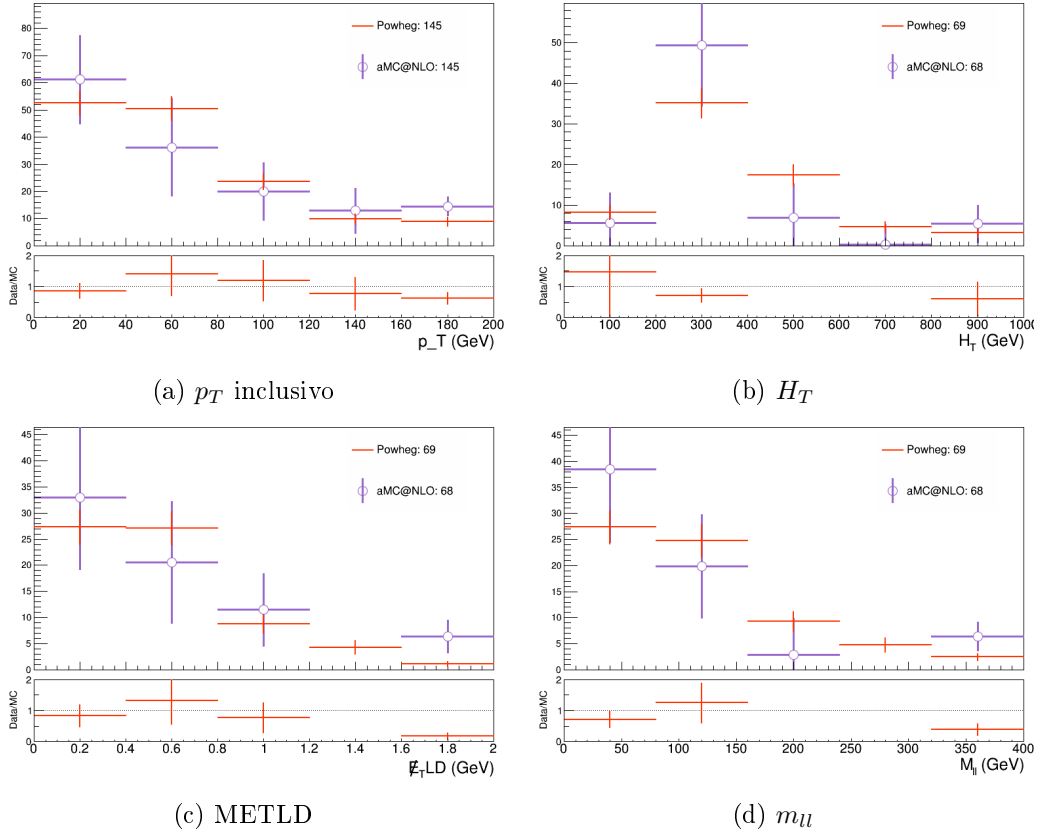


Figura 5.16: Histogramas de varias variables para el criterio de selecci3n en $\text{Lepton MVA} > 0.97$ que muestran la comparaci3n entre las muestras de Powheg y aMC@NLO.

6 Conclusiones

They say “doubt everything”, but I disagree. Doubt is useful in small amounts, but too much of it leads to apathy and confusion. No, don't doubt everything. QUESTION everything. That's the real trick. Doubt is just a lack of certainty. [...] But questions have answers, you see. [...] Questioning is a lifelong process. That's precisely what makes it so unlike doubt. Questioning engages with reality, interrogating all it sees. Questioning leads to a constant assault on intellectual status quo, where doubt is far more likely to lead to resigned acceptance. After all, when the possibility of truth is doubtful (excuse the pun), why not simply play along with the most convenient lie?

(Anónimo, en *The Talos Principle*, CROTEAM)

En este trabajo hemos estudiado el proceso $t\bar{t}H$, para el cual se ha elaborado un análisis secuencial (*cut & count*) considerado de alta luminosidad. Bajo las condiciones actuales del LHC, no tenemos los datos suficientes como para poder observar este proceso. Cuando la luminosidad integrada recogida en los experimentos del acelerador aumente (en futuros *run*, y en la actualización del acelerador al HL-LHC), será posible la observación y el estudio detallado del mismo. Esto redundará en las investigaciones de física más allá del modelo estándar, debido tanto a la mayor precisión de la medida de la sección eficaz del proceso (que puede participar como fondo en otros secundarios) y a la medida directa de la constante de acoplo de Yukawa entre el bosón de Higgs y el *quark top*.

En lo referente al primer objetivo del trabajo, el análisis elaborado es funcional, aunque las incertidumbres y el número de sucesos no permiten calcular una sección eficaz. Sin embargo, sí hemos podido calcular un límite superior esperado a la *signal strength* ($\mu = \frac{\sigma_{\text{exp.}}}{\sigma_{\text{teo.}}}$) con los resultados del cuadro 5.2, que es

$$\mu_{\text{obs.}} = 3.21 \quad \mu_{\text{esp.}} = 2.04^{+0.85}_{-0.59}.$$

El resultado no permite decir que hemos observado el proceso, aunque es un resultado esperado por el marco experimental actual y porque nuestro análisis no está optimizado para trabajar en estos niveles de luminosidad. Es lógico pues que un análisis como el de la Colaboración CMS, que emplea herramientas

estadísticas para optimizar mucho más el resultado, pueda alcanzar valores mejores, concretamente $0.8_{-0.2}^{+0.3}$. La comparación de los resultados (cuadro 5.2) con los de este análisis, que se pueden ver en la figura 6.1 arroja diferencias en todas las categorías salvo 4l. Esto, al igual que la diferencia en el límite, se puede explicar debido a distintas causas:

- El análisis realizado por CMS emplea, una vez se hace el estudio secuencial, técnicas multivariable (BDT) para mejorar la relación señal-fondo, lo que les permite obtener límites superiores de la sección eficaz más agresivos que si el análisis fuera solo *cut & count* como el nuestro.
- Nuestra identificación de leptones tau es más estricta que la suya.
- Nuestras muestras de simulaciones de MonteCarlo son distintas a las que se usaron para el análisis de la Colaboración. Concretamente, en la selección hecha para recortar (*skim*) las mismas y en algunos *paths* de HLT. Esto ha obligado a redefinir partes de la selección de leptones y *trigger*.
- Nosotros no hemos estimado los fondos de datos.
- No se han tenido en consideración los mismos factores de escala (SF) ni incertidumbres sistemáticas.

	$\mu\mu$	$e\mu$	ee
$t\bar{t}W$	$51.0 \pm 0.6 \text{ (stat.)} \pm 6.9 \text{ (syst.)}$	$72.8 \pm 0.7 \text{ (stat.)} \pm 10.2 \text{ (syst.)}$	$20.5 \pm 0.4 \text{ (stat.)} \pm 3.1 \text{ (syst.)}$
$t\bar{t}Z/\gamma^*$	$17.7 \pm 0.8 \text{ (stat.)} \pm 2.9 \text{ (syst.)}$	$47.3 \pm 1.6 \text{ (stat.)} \pm 9.0 \text{ (syst.)}$	$17.5 \pm 1.0 \text{ (stat.)} \pm 3.6 \text{ (syst.)}$
WZ	$4.2 \pm 0.6 \text{ (stat.)} \pm 4.1 \text{ (syst.)}$	$7.0 \pm 0.8 \text{ (stat.)} \pm 6.8 \text{ (syst.)}$	$1.8 \pm 0.4 \text{ (stat.)} \pm 1.7 \text{ (syst.)}$
Rare SM bkg.	$4.2 \pm 1.5 \text{ (stat.)} \pm 3.0 \text{ (syst.)}$	$13.3 \pm 1.9 \text{ (stat.)} \pm 9.3 \text{ (syst.)}$	$4.8 \pm 1.1 \text{ (stat.)} \pm 3.6 \text{ (syst.)}$
$WWss$	$3.5 \pm 0.6 \text{ (stat.)} \pm 2.5 \text{ (syst.)}$	$4.1 \pm 0.6 \text{ (stat.)} \pm 3.2 \text{ (syst.)}$	$1.4 \pm 0.3 \text{ (stat.)} \pm 1.2 \text{ (syst.)}$
Conversions		$7.8 \pm 2.5 \text{ (stat.)} \pm 2.3 \text{ (syst.)}$	$3.6 \pm 3.5 \text{ (stat.)} \pm 1.7 \text{ (syst.)}$
Charge mis-meas.		$16.4 \pm 0.2 \text{ (stat.)} \pm 9.1 \text{ (syst.)}$	$10.5 \pm 0.2 \text{ (stat.)} \pm 5.9 \text{ (syst.)}$
Non-prompt leptons	$38.7 \pm 1.6 \text{ (stat.)} \pm 20.5 \text{ (syst.)}$	$61.8 \pm 2.0 \text{ (stat.)} \pm 13.0 \text{ (syst.)}$	$17.7 \pm 1.1 \text{ (stat.)} \pm 5.4 \text{ (syst.)}$
All backgrounds	$120.3 \pm 2.5 \text{ (stat.)} \pm 11.7 \text{ (syst.)}$	$231.2 \pm 4.3 \text{ (stat.)} \pm 13.3 \text{ (syst.)}$	$77.9 \pm 4.0 \text{ (stat.)} \pm 9.0 \text{ (syst.)}$
$t\bar{t}H$ signal	$20.1 \pm 0.5 \text{ (stat.)} \pm 2.1 \text{ (syst.)}$	$27.9 \pm 0.5 \text{ (stat.)} \pm 3.0 \text{ (syst.)}$	$8.0 \pm 0.3 \text{ (stat.)} \pm 1.1 \text{ (syst.)}$
Data	150	268	89

	3L	4L
$t\bar{t}W$	$32.8 \pm 1.0 \text{ (stat.)} \pm 4.9 \text{ (syst.)}$	
$t\bar{t}Z/\gamma^*$	$49.8 \pm 3.9 \text{ (stat.)} \pm 11.1 \text{ (syst.)}$	$2.15 \pm 0.24 \text{ (stat.)} \pm 0.44 \text{ (syst.)}$
WZ	$9.1 \pm 0.9 \text{ (stat.)} \pm 4.0 \text{ (syst.)}$	
Rare SM bkg.	$8.8 \pm 4.3 \text{ (stat.)} \pm 5.9 \text{ (syst.)}$	$0.27 \pm 0.16 \text{ (stat.)} \pm 0.19 \text{ (syst.)}$
$WWss$		
Conversions	$5.3 \pm 1.2 \text{ (stat.)} \pm 4.0 \text{ (syst.)}$	
Charge mis-meas.		
Non-prompt leptons	$30.8 \pm 1.5 \text{ (stat.)} \pm 10.9 \text{ (syst.)}$	
All backgrounds	$137.3 \pm 6.2 \text{ (stat.)} \pm 12.4 \text{ (syst.)}$	$2.42 \pm 0.28 \text{ (stat.)} \pm 0.56 \text{ (syst.)}$
$t\bar{t}H$ signal	$19.5 \pm 1.0 \text{ (stat.)} \pm 3.0 \text{ (syst.)}$	$1.00 \pm 0.09 \text{ (stat.)} \pm 0.11 \text{ (syst.)}$
Data	148	3

Figura 6.1: [56] Resultados por categorías del análisis de la Colaboración CMS.

En el marco del análisis elaborado, hemos estudiado la identificación de leptones. Particularmente, la influencia de la variable **Lepton MVA**, que está diseñada para reducir el fondo de sucesos con leptones *fakes*. La comparación de los resultados de los cuadros 5.7, 5.4 y 5.5 han demostrado la eficacia de la variable para discriminar este fondo. Las comparaciones con identificaciones de leptones de otros análisis (que no usan **Lepton MVA**) suponen un cambio en porcentaje del número de sucesos con leptones *fake* de hasta el 907% de los que tiene el análisis de $t\bar{t}H$. Y la comparación con los distintos WP de esta variable nos permite ver cómo evoluciona la misma sin variar ningún otro factor. Los

requisitos más suaves considerados (WP *medium*) aumentan hasta en un 508 % el número de sucesos con leptones *fake*, y las exigencias más *tight* permiten reducirlos aún más, hasta un 4 % de los obtenidos.

Concluimos pues que la variable **Lepton MVA** es muy eficaz a la hora de reducir este fondo, particularmente en el análisis $t\bar{t}H$ que hemos elaborado.

Como estudio tangencial, se ha comprobado que una identificación más *tight* de los leptones tau que decaen hadrónicamente permiten una mejoría en el indicador $\frac{S}{\sqrt{F+(\Delta F)^2}}$ de un 1.2 %, así como en el límite superior de la sección eficaz. Este hecho podría dar pie a explorar un cambio en la identificación de taus que mejore la selección de eventos de todo el análisis.

También hemos visto que en nuestro análisis y en los WP más estrictos de la variable **Lepton MVA** que hemos considerado, la selección del generador de muestras de Monte Carlo juega un papel no despreciable, debido al contexto de baja luminosidad en el que nos encontramos. Concretamente, la comparación entre generadores de **aMC@NLO** y **Powheg** muestra desviaciones entre ambas apreciables en algunas regiones de determinadas distribuciones de variables, como se explicó en la tercera sección del capítulo anterior.

Todos los estudios realizados en este trabajo se aplicarán en los próximos análisis de la Colaboración CMS de este proceso. Para ello, se emplearán los datos recogidos a lo largo del segundo *run*.

7 Bibliografía

- [1] Bloch, D., Botta, C., and et. al., N. C. (2017) «Search for $t\bar{t}H$ in multilepton final states with the full 2016 dataset». *Internal CMS Draft Analysis Note*,.
- [2] Dalton, J. (1805) «On the Absorption of Gases by Water and other Liquids». *Memoirs of the Literary and Philosophical Society of Manchester*, **I**, 277–286.
- [3] CERN *Robert Millikan coins the term ‘cosmic rays’*. [En línea; consulta el 15-06-17] <https://timeline.web.cern.ch/robert-millikan-coins-the-term-%E2%80%98cosmic-rays%E2%80%99>.
- [4] Wikimedia *Diagram showing how a cyclotron works.* [En línea; consulta el 15-06-17] https://en.wikipedia.org/wiki/Cyclotron#/media/File:Cyclotron_diagram.png.
- [5] Wikimedia Commons *The Standard Model of elementary particles (more schematic depiction), with the three generations of matter, gauge bosons in the fourth column, and the Higgs boson in the fifth.* https://en.wikipedia.org/wiki/Standard_Model#/media/File:Standard_Model_of_Elementary_Particles.svg [En línea; consulta el día 11-02-17].
- [6] Wikimedia Commons *Summary of interactions between particles described by the Standard Model.* https://en.wikipedia.org/wiki/Standard_Model#/media/File:Standard_Model_of_Elementary_Particles.svg [En línea; consulta el día 11-02-17].
- [7] Gripaio, B. (2015) «Physics Beyond the Standard Model». *arXiv:1503.02636 [hep-ph]*, pp. 12–13.
- [8] DESY, Ethan Siegel *Summary of interactions between particles described by the Standard Model.* http://scienceblogs.com/startswithabang/files/2013/05/susyparticles_sm.png [En línea; consulta el día 13-02-17].
- [9] Desconocido *Diagrama de las constantes de acoplo.* [En línea; consulta el 15-06-17] <https://i.stack.imgur.com/UrOol.gif>.
- [10] et. al., C. P. (2016) «*Particle Data Group*». *Chinese Physics C*, **40**, 100001,.
- [11] de CMS), P. C. C. (2016) « $t\bar{t}H$ production at 13 TeV». *arXiv:1611.08528v1 [hep-ex]*,.
- [12] de Florian, D., Grojean, C., and et. al., F. M. (2017) «Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector». *arXiv:1610.07922 [hep-ph]*,.

- [13] CERN, Colaboración CMS Gráfico con diversas secciones eficaces a distintas energías para varios procesos que se dan en el acelerador LHC del CERN. [En línea; consulta el 25-06-17] <https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsCombined>.
- [14] CERN (comunicado de prensa) «*CERN Council Gives Go-ahead for Large Hadron Collider*». <https://press.cern/press-releases/1994/12/cern-council-gives-go-ahead-large-hadron-collider> [En línea; consulta el 29-05-17].
- [15] L. R. Evans, P. Bryant *LHC Machine*. <https://cds.cern.ch/record/1129806/?ln=es> CERN Document Server [En línea; consulta los días 28,30,31-05-17].
- [16] CERN «*Taking a closer look at LHC*». http://lhc-closer.es/taking_a_closer_look_at_lhc/1.home [En línea; consulta los días 28,30,31-05-17].
- [17] Mapa del LHC sobre la región cercana a la ciudad de Ginebra. <https://alexeinstein.files.wordpress.com/2014/10/lhc.jpg> [En línea; consulta el 30-05-17].
- [18] CERN (comunicado de prensa) «*LHC sets new world record*». <https://press.cern/press-releases/2009/11/lhc-sets-new-world-record> [En línea; consulta el 30-05-17].
- [19] Diagrama de todos los aceleradores que sirven de preinyector al LHC. <http://void.printf.net/~conor/sa/LHCb/cernschema.jpg> [En línea; consulta el 30-05-17].
- [20] Jean-Luc Caron, CERN «*Magnetic field induced by the LHC dipole's superconducting coils*». <https://cds.cern.ch/record/841511/?ln=es> CERN Document Server [En línea; consulta el 30-05-17].
- [21] CERN Esquema del funcionamiento de un cuadripolo magnético del LHC. http://www.lhc-closer.es/webapp/files/1435504123_b887b3b5c6aab9b0259320ea21935bbd.png [En línea; consulta el 30-05-17].
- [22] CERN, Colaboración CMS «*CMS Integrated Luminosity (p-p. 2016, 13 TeV)*». https://cms-service-lumi.web.cern.ch/cms-service-lumi/publicplots/int_lumi_per_day_cumulative_pp_20160onlineLumi.png [En línea; consulta el 31-05-17].
- [23] S. Cittolin, A. Rácz, S. Paris (CERN, Colaboración CMS) «*CMS The TriDAS Project : Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger*». <https://cdsweb.cern.ch/record/578006?> CERN Document Server [En línea; consulta el 28-05-17].
- [24] G. Apollinari, O. Brüning, L. Rossi (CERN) «*High Luminosity LHC Project Description*». <https://cds.cern.ch/record/1974419/?ln=es> CERN Document Server [En línea; consulta el 28-05-17].
- [25] CERN, Colaboración CMS «*CMS experiment*». <https://cds.cern.ch/record/1474902/?ln=es> CERN Document Server [En línea; consulta el 31-05-17].

- [26] Colaboración CMS «*The CMS experiment at the CERN LHC*». <https://cds.cern.ch/record/1129810/?ln=es> CERN Document Server [En línea; consulta el 31-05-17].
- [27] Colaboración CMS «*The CMS experiment at the CERN LHC*». <http://cms.web.cern.ch/news/what-cms> CERN Document Server [En línea; consulta el 31-05-17].
- [28] CERN, Colaboración CMS Diagrama que muestra las partes del detector CMS. https://cms-docdb.cern.ch/cgi-bin/PublicDocDB/RetrieveFile?docid=11514&version=1&filename=cms_120918_03.png [En línea; consulta el 31-05-17].
- [29] Desconocido Diagrama de una sección del CMS y del paso de partículas por él. [En línea; consulta el 15-06-17] http://www.particlecentral.com/images/cms_slice.jpg.
- [30] CERN, Colaboración CMS *First half of CMS inner tracker barrel*. <https://cds.cern.ch/record/995912/?ln=es> CERN Document Server [En línea; consulta el 01-06-17].
- [31] Colaboración CMS Fotografía de los cristales del ECAL del CMS. <http://cms.web.cern.ch/sites/cms.web.cern.ch/files/styles/large/public/field/image/ECALcrystals.jpg?itok=EH3HV6Z7> [En línea; consulta el 01-06-17].
- [32] CERN, Colaboración CMS *Assembled HCAL endcap (HE) in November 2002*. <https://cds.cern.ch/record/947004/?ln=es> CERN Document Server [En línea; consulta el 01-06-17].
- [33] Colaboración CMS Fotografía del solenoide superconductor del CMS durante su construcción. http://cms.web.cern.ch/sites/cms.web.cern.ch/files/styles/large/public/field/image/0509015_14-A4-at-144-dpi_thumb.jpg?itok=Z1Y7rBJV [En línea; consulta el 01-06-17].
- [34] CMS, C. (2014) «*The performance of the CMS muon detector in proton-proton collisions at $\sqrt{s} = 7$ TeV at the LHC*». *arXiv:1306.6905v2 [physics.ins-det]*,.
- [35] CMS, C. (2008) «The CMS Experiment at the CERN LHC». *INSPIRE*, DOI: 10.1088/1748-0221/3/08/S08004,.
- [36] CMS, C. (2009) «Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET». *CMS Physics Analysis Summaries*,.
- [37] CMS), F. B. C. (2013) «The CMS Particle Flow Algorithm». *arXiv:1401.8155 [hep-ex]*,.
- [38] CERN, Colaboración CMS Página de la wiki del CERN del Muon Physics Object Group de CMS. [En línea; consulta el 15-06-17] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuonPOG>.
- [39] CERN, Colaboración CMS Página de la wiki del CERN dentro de CMS sobre Lepton MVA. [En línea; consulta el 15-06-17] <https://twiki.cern.ch/twiki/bin/view/CMS/LeptonMVA>.

- [40] CERN, Colaboración CMS Diagrama de la producción de un *jet*. [En línea; consulta el 15-06-17] http://cms.web.cern.ch/sites/cms.web.cern.ch/files/styles/large/public/field/image/Sketch_PartonParticleCaloJet.png?itok=SbSUP7_9.
- [41] M. Cacciari, G. P. Salam, G. S. (2008) «The anti- k_t jet clustering algorithm». *arXiv:0802.1189 [hep-ph]*,.
- [42] Wikimedia Diagrama de la producción de un *b-jet*. [En línea; consulta el 15-06-17] https://upload.wikimedia.org/wikipedia/commons/b/ba/B-tagging_diagram.png.
- [43] CMS, C. (2016 (última versión)) «Identification of b quark jets at the CMS Experiment in the LHC Run 2». *CERN Document Server, CMS-PAS-BTV-15-001*,.
- [44] CMS, C. (2016) «Reconstruction and identification of tau lepton decays to hadrons and tau neutrino at CMS». *arXiv:1510.07488 [physics.ins-det]*,.
- [45] Colaboración CMS Repositorio de GitHub «CMGTools» secundario. [En línea; consulta el 25-06-17] <https://github.com/CERN-PH-CMG/cmgttools-lite>.
- [46] Cowan, G., Cranmer, K., Gross, E., and Vitells, O. (2010) «*Asymptotic formulae for likelihood-based tests of new physics*». *arXiv:1007.1727*,.
- [47] Cranmer, K. (2015) «*Practical Statistics for the LHC*». *arXiv:1503.07622*,.
- [48] Grupo Experimental de Altas Energías de la Universidad de Oviedo Repositorio «AnalysisPAF». <https://github.com/Oviedo-PAF/AnalysisPAF>.
- [49] Grupo Experimental de Altas Energías de la Universidad de Oviedo Página web de PAF. <http://www.hep.uniovi.es/PAF/>.
- [50] Delgado, J., del Castillo, E. F., González, I., and Rodríguez, A. (2015) «*PROOF Analysis Framework (PAF)*». *J. Phys.: Conf. Ser.*,.
- [51] Petrucciani, G., Rizzi, A., and Vuosalo, C. (2017) «*Mini-AOD: A New Analysis Data Format for CMS*». *arXiv:1702.04685*,.
- [52] Colaboración CMS Repositorio de GitHub «CMGTools» al completo. [En línea; consulta el 25-06-17] https://www.desy.de/dvsem/WS0708/hegner_talk.pdf.
- [53] Benedikt Hegner (DESY) Charla «*CMS Data Format and Analysis Environments*». [En línea; consulta el 25-06-17] https://www.desy.de/dvsem/WS0708/hegner_talk.pdf.
- [54] CMS, C. (2017) «Search for the associated production of a Higgs boson with a top quark pair in final states with a τ lepton at $\sqrt{s} = 13$ TeV». *CERN Document Server*,.

- [55] Punzi, G. (2003) «*Sensitivity of searches for new signals and its optimization*». *arXiv:physics/0308063 [physics.data-an]*,.
- [56] Bloch, D., Botta, C., and et. al., N. C. (2017) «Search for ttH in multilepton final states with the full 2016 dataset». *Internal CMS Physics Analysis Summary*,.
- [57] Lago, A. B. (2000) «Historia de la Filosofía 2», Editorial Everest S.A., León (España) 1a edition.
- [58] Cahn, R. and Goldhaber, G. (2009) «The Experimental Foundations of Particle Physics», Cambridge University Press, Cambridge (Reino Unido) 2^a edition.
- [59] Angelis, A. D. and Pimenta, M. J. M. (2015) «Introduction to Particle and Astroparticle Physics - Questions to the Universe», Springer-Verlag, Milán (Italia) 1^a edition.
- [60] Povh, B., Rith, K., Scholz, C., and Zetsche, F. (1995) «Particles and Nuclei - An Introduction to the Physical Concepts», Springer-Verlag, Berlín (Alemania) 1^a edition.
- [61] Griffiths, D. (2008) «Introduction to Elementary Particles», WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim (Alemania) 2^a edition.

