

Búsqueda de materia oscura en producción asociada con quarks top en el experimento CMS del LHC (CERN)

*Dark matter search in production associated with
top quarks in the CMS experiment at LHC (CERN)*

Manuel Souto Juan



Universidad de Oviedo

Trabajo Fin de Grado
Curso 2018-2019

supervisado por
Francisco Javier Cuevas Maestro y Sergio Sánchez Cruz
Universidad de Oviedo

Resumen

La búsqueda de materia oscura es uno de los principales temas de estudio en el ámbito de la física actual. Evidencias cosmológicas apuntan a su existencia, lo que conlleva estudios en ese ámbito para intentar describir su naturaleza y comportamiento. Sin embargo, los estudios con colisionadores de partículas han permitido desde su puesta en funcionamiento un marco de búsqueda de nuevas partículas; tómese por ejemplo el caso del descubrimiento del bosón de *Higgs* en el LHC en el año 2012. Este hecho permite preguntarnos si es posible producir materia oscura en colisiones de partículas ya que, de esta forma, el LHC sería un ámbito de estudio adicional a los estudios cosmológicos en la producción de este tipo de materia. Bajo esta premisa, en este trabajo se estudiará uno de los principales modelos para describir la materia oscura que se ha estudiado en colisionadores; su originalidad se basa en un método de estudio no realizado en CMS hasta el momento, usando datos de CMS tomados en el año 2016 ($\mathcal{L} = 35.9 \text{ fb}^{-1}$). Este estudio está basado en una búsqueda previa realizada por la colaboración del detector ATLAS ([1]).

El objetivo perseguido es estudiar el canal de producción de materia oscura asociada a dos *quarks* top. Se tratará de definir una región de estudio de los datos que permita, bajo ciertos criterios, asegurar un descubrimiento de nuevas partículas, o que por el contrario permita invalidar el modelo empleado en el estudio. Los resultados y conclusiones obtenidos se recogerán en el último apartado de este trabajo. En ellos se muestra que los resultados obtenidos son compatibles con el modelo estándar, permitiendo poner límites a este modelo de búsqueda de nueva física.

Índice general

Resumen	I
Contenidos	III
Introducción	1
1. Introducción al modelo estándar (SM)	3
1.1. Descripción del modelo estándar actual	3
2. Descripción y estudio de DM en aceleradores	7
2.1. Descripción y detección de DM	7
2.2. Principales modelos de DM en aceleradores	9
2.2.1. Higgs y Z como portales a DM	10
2.2.2. Teorías de campo efectivas (EFT)	10
2.2.3. Modelos simplificados de mediación	11
2.2.4. Modelos supersimétricos (SUSY)	12
2.2.5. Modelo simplificado estudiado	14
3. El detector CMS	15
3.1. LHC: el gran colisionador de hadrones	15
3.2. Descripción general del detector	16
3.3. Luminosidad, <i>pile-up</i> y <i>trigger</i>	19
3.4. Identificación y simulación de objetos	20
3.4.1. Identificación de objetos en el detector CMS	20
3.4.2. Simulación de procesos en el detector CMS	24
4. Descripción del análisis de datos	27
4.1. Procesos involucrados en el análisis	27
4.1.1. El proceso $t\bar{t}$ +DM	27
4.1.2. El proceso $t\bar{t}$	28

4.1.3.	El proceso Drell-Yan (DY)	29
4.1.4.	Los procesos $t\bar{t}W$ y $t\bar{t}Z$	29
4.1.5.	Los procesos dibosónicos: WW, WZ, ZZ	30
4.1.6.	El proceso $W + Jets$	31
4.1.7.	El proceso tW	32
4.2.	Descripción de datos y muestras de Monte Carlo (MC) empleadas . .	32
4.3.	Variables cinemáticas de discriminación	34
4.3.1.	<i>Blinding</i>	41
4.4.	Selección de eventos	41
5.	Resultados obtenidos	53
5.1.	Extracción de señal. Obtención del límite superior (UL) del modelo .	53
5.2.	Incertidumbres sistemáticas	55
5.3.	Resultados y conclusiones	59
	Conclusiones finales	65
	Referencias y otra bibliografía consultada	69

Introducción

Una gran parte de los procesos físicos en la naturaleza están incluidos en el marco del modelo estándar (SM, *Standard Model*). Sin embargo, la existencia de fenómenos no incluidos en su marco es conocida desde hace varios años, siendo las observaciones cosmológicas la principal fuente de evidencias de física más allá del SM (BSM, *Beyond Standard Model*).

Estos hechos apuntan a la existencia de *materia no visible* llamada **materia oscura** (DM), es decir, materia que interacciona débilmente con la materia ordinaria visible; sin embargo, es conocido que su detección es debida a su interacción gravitatoria con la materia ordinaria. Medidas sobre la *radiación de fondo cósmico de microondas* (CMB, *Cosmic Microwave Background*) muestran que la materia bariónica representa un 4% de la materia del universo observable, mientras que la proporción de materia oscura es 5 veces mayor. Esto motiva la necesidad de poner en conjunto el conocimiento astrofísico con la física de partículas.

La búsqueda de DM se separa en tres tipos, en función del tipo de búsqueda realizada:

- Detección directa: basada en el estudio de colisiones de DM galáctica con blancos de materia ordinaria.
- Detección indirecta: basada en la detección de los productos de aniquilación de DM.
- Detección en colisionadores de partículas.

El uso de colisionadores ha revelado muchos aspectos de la materia ordinaria, por lo que son una herramienta clave en el estudio de física BSM relacionada con el comportamiento de nuevas partículas, independientemente del ámbito en que dichas partículas sean descubiertas.

Los conocimientos empleados en la realización de esta memoria están relacionados con las asignaturas «*Física Nuclear y de Partículas Elementales*», «*Física de*

Altas Energías y Aceleradores», «*Introducción a la Física Computacional*» e «*Inferencia Estadística*», entre otras. También se ha tenido que obtener competencias en el marco computacional empleado: el *framework* empleado en el análisis posterior y conocimientos en lenguajes de programación (C++, Python, bash).

A continuación se expone la estructura de este trabajo:

- El primer capítulo se centrará en la exposición del SM actual, introduciendo las partículas elementales y una descripción de las interacciones entre dichas partículas mediante sus correspondientes mediadores.
- El segundo capítulo introducirá los distintos modelos de DM estudiados en aceleradores, y se dedicará con más profundidad al estudio del modelo usado en este trabajo.
- El tercer capítulo incluirá una descripción breve del detector CMS y el proceso de identificación de los distintos objetos detectados.
- En el cuarto capítulo se procederá a introducir los distintos procesos que aparecerán en el posterior análisis, con el fin de comprender la señal y los distintos fondos estudiados. Se continuará con una descripción de los datos y las simulaciones empleadas y se expondrá el análisis realizado, incluyendo los distintos observables y criterios de selección empleados.
- En el quinto capítulo se recogerán los resultados experimentales obtenidos. Este capítulo estará centrado en la descripción del análisis anterior que permita obtener una cota superior en la sección eficaz del proceso, tras lo que se comprobarán las principales fuentes de incertidumbre sistemáticas que afectan dicho análisis.

Finalmente, se incluye un apartado de conclusiones donde se expondrán los resultados obtenidos en este trabajo y se comentarán posibles mejoras que permitan acotar de una forma más precisa la sección eficaz del proceso estudiado.

Capítulo 1

Introducción al modelo estándar (SM)

En este capítulo se comenzará haciendo un breve resumen del modelo estándar actual como una base para las búsquedas de nuevas partículas que se abordan en este estudio; una visión más extensa puede ser consultada en [13]. También se incluirá una descripción de los principales procesos que influirán en el análisis realizado posteriormente, así como una descripción del proceso señal considerado.

1.1. Descripción del modelo estándar actual

El **modelo estándar de física de partículas** es una teoría científica que describe la estructura fundamental de la materia y las interacciones entre partículas. Se trata de una teoría cuántica de campos que incorpora tres de las cuatro interacciones fundamentales entre partículas, exceptuando la *interacción gravitatoria*. La palabra "modelo" proviene del hecho histórico de su establecimiento en los años 70, donde no había suficiente evidencia experimental para considerarlo una teoría.

El modelo estándar establece la existencia de dos principales familias de partículas: los **fermiones**, que son los componentes de la materia; y los **bosones**, que son las partículas mediadoras de las interacciones fundamentales. En el marco del modelo estándar, las interacciones entre partículas se producen por el *intercambio* de un bosón. En la figura 1.1 se muestran las partículas fundamentales del SM agrupadas por clases aunque, como es conocido, el gravitón no está actualmente incluido en el modelo estándar debido a la dificultad de incorporarlo a una teoría de campos renormalizable. Esto implica uno de los mayores problemas del modelo estándar: la unificación de la gravedad.

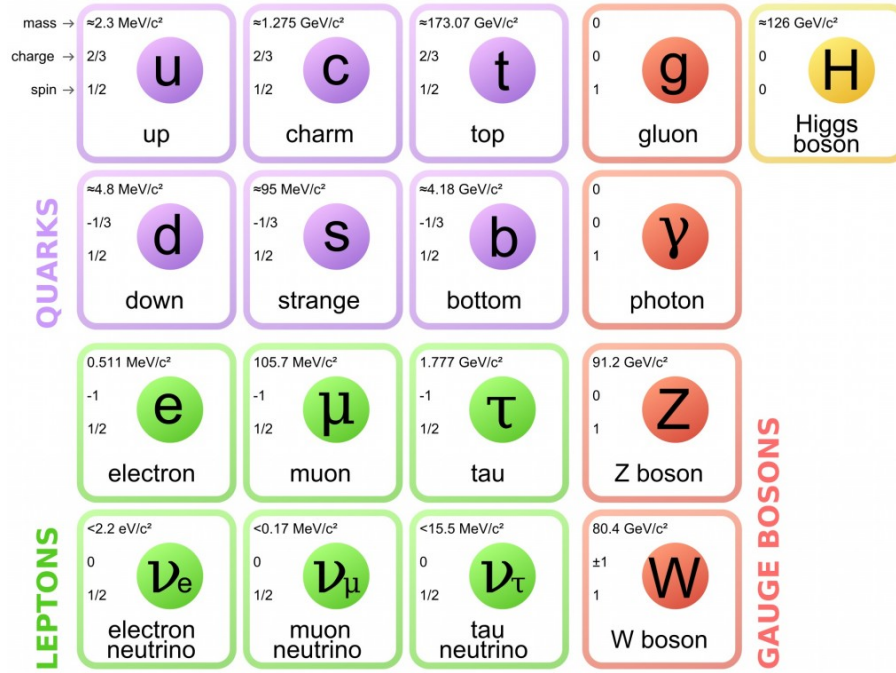


Figura 1.1: [6] Partículas elementales del modelo estándar. El conjunto de *quarks* (morado) y leptones (verde) forma los fermiones, mientras que los bosones mediadores de las interacciones fundamentales aparecen en rojo. Se incluye el bosón de Higgs, responsable de la masa de las partículas (amarillo).

Los fermiones son partículas con *spin* semientero que se rigen por la estadística de Fermi-Dirac; estas partículas componen la materia visible. El resto de partículas del modelo estándar, denominados bosones, son partículas con *spin* entero que se rigen por la estadística de Bose-Einstein, y su intercambio entre fermiones lleva a la aparición de las interacciones fundamentales. Por otro lado, aunque en la figura 1.1 no se muestre, para cada fermión existe una antipartícula asociada, cuya única diferencia con éste es que su carga eléctrica es de igual magnitud, pero de signo opuesto.

Para que una interacción pueda llevarse a cabo entre dos fermiones, es necesario que sea posible el intercambio de un cierto bosón. Esto es debido a que cada bosón está asociado a una interacción fundamental, y cada una de ellas tiene una carga asociada: electromagnética (EM)-carga eléctrica, débil-carga débil, fuerte-color. Por tanto, para cada proceso de interacción es necesario que los fermiones posean carga del tipo indicado.

A continuación se exponen las distintas interacciones fundamentales y los bosones que las rigen:

- **Interacción electromagnética:** está mediada por el fotón (γ), una partícula eléctricamente neutra de masa nula, lo que explica el alcance infinito de esta interacción. Todos los *quarks*, todos los leptones exceptuando los neutrinos y los bosones $W^\pm$ tienen carga eléctrica y pueden sufrir esta interacción.
- **Interacción débil:** está mediada por el bosón Z, eléctricamente neutro y los bosones $W^\pm$. Estos tres bosones son especialmente masivos comparados al resto de bosones mediadores, lo que explica el corto rango de esta interacción. Todos los fermiones poseen carga débil.
- **Interacción fuerte:** está mediada por el gluón (g), partícula eléctricamente nula, sin carga débil. Sin embargo, a diferencia del resto de bosones, posee carga de color. Este hecho hace que esta interacción sea la más compleja introducida en el modelo estándar, y la teoría que explica su comportamiento se conoce como la cromodinámica cuántica (*Quantum Chromodynamics*, QCD).

Sufren fuerza fuerte las partículas con carga de color, es decir, los seis *quarks* y sus *antiquarks* asociados. El hecho de que el gluón tenga color implica que interacciona consigo mismo, lo cual permite explicar el reducido alcance de esta interacción (el radio atómico) y el fenómeno del *confinamiento*.

Por último, el bosón de Higgs se acopla a todas las partículas con masa mediante el **mecanismo de Higgs** (incluido a sí mismo) y es el responsable de darles su masa. En la formulación actual del SM, se considera que los neutrinos no tienen masa; esta cuestión es objeto de estudio en la física de partículas.

Una última clasificación de las partículas fundamentales es su ordenación por masa. Esta clasificación se hace por columnas, como se muestra en la figura 1.1, y cada una recibe el nombre de **generación**. Las generaciones se ordenan en orden creciente de masa. Cada leptón cargado eléctricamente (electrón, e ; muón, μ ; tau, τ) tiene en su generación a su neutrino asociado.

Por último, existen las partículas compuestas por *quarks*, que se denominan hadrones. Una subclasificación de estos atiende al número de constituyentes: mesones (dos *quarks*) o bariones (tres *quarks*). Por ejemplo, un protón es un barión, constituido por quarks de primera generación en configuración *uud*. Debido al *confinamiento* cuántico, una característica de QCD, las partículas con carga de color no pueden existir en libertad y deben formar hadrones (*hadronización*). Esto implica que cuando se produce un gluón o un *quark*, estos interaccionan fuertemente con su entorno y todos tienden a formar un haz de partículas colimado en una dirección. Este haz recibe el nombre de **jet de partículas**.

Capítulo 2

Descripción y estudio de DM en aceleradores

Este capítulo estará dedicado a introducir la materia oscura, comenzando por una descripción fenomenológica basada en su interacción con la materia visible y su importancia en el entendimiento de la composición del universo. Por otro lado, se motivará la búsqueda de este tipo de materia en condiciones controladas, como los colisionadores de partículas, con la intención de comprender su naturaleza y comportamiento; esta descripción se ha obtenido y se puede consultar en [14]. Por último, se introducirán los principales modelos teóricos empleados en el LHC para la descripción de estas nuevas partículas, mostrando y justificando su uso en este ámbito; estos modelos se recogen en [4], [5].

2.1. Descripción y detección de DM

En el ámbito de la física actual, todos los procesos físicos se encuentran en el marco del modelo estándar. Sin embargo, existen ciertos fenómenos que no han podido ser englobados en su marco, como pueden ser la unificación de la interacción gravitatoria o las oscilaciones de neutrinos. Uno de los principales fenómenos de estudio en este marco es la búsqueda de *materia oscura*, materia no bariónica que compone en gran medida el universo observable.

Las primeras observaciones de estos fenómenos datan del año 1930, cuando cálculos realizados por astrónomos mostraban discrepancias sobre la masa de las galaxias, hecho corroborado por observaciones de las velocidades de rotación de las galaxias en los años 70. Estos cálculos tenían en cuenta la masa correspondiente a la materia visible y, por tanto, que interaccionaba electromagnéticamente. Por ello y, dado que

este hecho implicaba una interacción con partículas del modelo estándar, era necesario que esta *materia oscura* interaccionase con la materia visible gravitatoriamente. Este exceso de materia fue comprobado adicionalmente mediante las lentes gravitacionales, una consecuencia de la teoría de la relatividad general que implicaba la distorsión de una fuente de luz en presencia de materia.

Esta materia no es bariónica o visible, debido a la naturaleza de su comportamiento: no se trata de antimateria, dado que no muestra los rayos- γ característicos de las aniquilaciones materia-antimateria; tampoco se trata de agujeros negros, ya que no interacciona de forma directa con los fotones. Además, las observaciones realizadas por el telescopio Planck sobre el CMB muestran que la proporción de materia oscura en el universo visible estaría en torno al 68 %, mientras que la materia bariónica sería tan sólo el 5 %. Por tanto, el único conocimiento exacto de su naturaleza es que está compuesto por partículas masivas, dado que interacciona gravitatoriamente; por tanto, experimentos en colisionadores, que han permitido descubrir y conocer el comportamiento de la materia visible, constituyen un ámbito de estudio adecuado para el descubrimiento de nuevas partículas. Al tratarse de un sistema controlado, es posible conocer el tipo y la energía de las partículas que colisionan (en el caso del LHC se conoce la energía de cada protón, pero debido a la presencia de *partones*, no es posible conocer la energía exacta de la partícula que colisiona) y generar un alto número de colisiones que permitan el estudio de este tipo de procesos.

El estudio de procesos en el LHC se basa en la colisión de **protones**. Como ha sido comentado en el primer capítulo, los protones son hadrones, y por tanto, partículas compuestas por *quarks*. Sin embargo, y como se comentó anteriormente, la suma de las masas de sus constituyentes (dos *quarks* *u* y un *quark* *d*) no reconstruye la masa total del protón, sino que un gran porcentaje de su masa es **energía de enlace**. Este hecho conlleva la presencia de *quarks de mar* y gluones, que no son responsables de caracterizar la estructura del protón, pero sin embargo pueden ser los responsables de la interacción en el momento de la colisión. La probabilidad de interacción de un cierto componente del protón se cuantifica mediante las *parton density functions* (PDFs, [15]). De esta forma, la partícula primaria que interactúa es el *gluón*, produciéndose interacciones con *quarks* más pesados (*top*, *bottom*) a mayores valores de energía de centro de masas del sistema; esto justifica el hecho de mejorar los aceleradores para aumentar la energía del sistema.

Los procesos estudiados en física de partículas se representan mediante **diagramas de Feynmann** ([16]). Se trata de diagramas simplificados donde se muestra la interacción entre las partículas en los estados inicial y final, y se representan las

partículas mediadoras de la interacción. Aunque no se representan ejes coordenados, se interpreta el eje horizontal como el eje *temporal*, aumentando éste hacia la derecha. Se distinguen distintos trazos en función de la naturaleza de la partícula: las líneas rectas representan partículas fermiónicas, como leptones o *quarks*, las líneas sinusoidales representan mediadores electrodébiles (fotones, bosones $W^\pm, Z$), las líneas circulares representan mediadores fuertes (g), y las líneas rectas discontinuas representan mediadores bosónicos escalares (por ejemplo, el bosón de Higgs). De esta forma, cada vértice en un diagrama de Feynmann representa una interacción entre dos o más partículas. Estos diagramas permiten una visión clara de este tipo de procesos, de alta complejidad, con una interpretación más sencilla.

2.2. Principales modelos de DM en aceleradores

La producción de partículas candidatas a materia oscura en colisionadores de partículas, como el caso del LHC, se realizará en consonancia con procesos que involucren partículas del modelo estándar (ya que en otro caso no podrían ser detectadas). Por tanto, es necesario introducir un modelo teórico que permita describir la interacción entre estas partículas y las partículas del modelo estándar, siendo así posible analizar los datos tomados y contrastar hipótesis de validez sobre dichos modelos. En este apartado se introducirán los distintos modelos empleados en el LHC para la búsqueda de candidatos a DM. Éstos pueden clasificarse en dos tipos:

- Modelos totalmente especificados y consistentes, con una gran cantidad de parámetros que permiten el estudio en ciertas regiones del espacio de fases considerado; este es el caso de supersimetría (SUSY).
- Modelos simplificados con pocas componentes, que permiten caracterizar la fenomenología de un gran número de procesos evitando la necesidad de tomar asunciones excesivamente fuertes; este es el caso de las *effective field theories* (EFTs).

Estos modelos incluyen partículas candidatas a DM, y describen su interacción con la materia ordinaria mediante uno o varios mediadores de distintas características. El modelo estudiado en este trabajo será un modelo simplificado basado en un mediador escalar con la materia oscura de *spin* cero; será descrito de forma más precisa al final del capítulo.

2.2.1. Higgs y Z como portales a DM

Se trata de un caso particular de los modelos simplificados introducidos anteriormente. Si la extensión del modelo estándar se realiza añadiendo una sola partícula candidata a DM, el modelo estudiado puede ser un modelo en que la partícula mediadora entre las partículas del modelo estándar y dicho candidato sea un bosón de Higgs o un bosón Z . Por ello, estos modelos reciben el nombre de *portales*. Un modelo de este tipo puede expresarse mediante un diagrama de Feynmann como el mostrado en la figura 2.1.

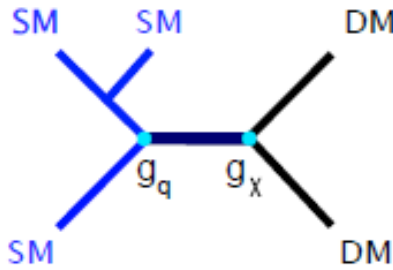


Figura 2.1: [5] Diagrama de Feynmann representando un modelo *portal* con un mediador entre las partículas del modelo estándar (SM) y la materia oscura (DM) con constantes de acoplamiento mediador-SM (g_q) y mediador-DM (g_χ).

Los modelos *portal* del Higgs pueden construirse añadiendo simplemente DM al modelo estándar; en el marco de este tipo de modelos, las búsquedas directas en colisionadores sobre los bosones de Higgs decayendo en partículas invisibles implicarían una gran sensibilidad a otros acoplamientos del Higgs a nuevas partículas, lo que motiva el uso de este tipo de modelos.

2.2.2. Teorías de campo efectivas (EFT)

Este tipo de modelos son muy utilizados en este ámbito debido a que no asumen demasiados detalles sobre la partícula mediadora, lo que permite una gran libertad en su definición. En el caso de que la masa de dicho mediador sea grande comparada con la energía de la colisión, el modelo se asemeja al producido por una fuerza de contacto.

Suelen utilizarse dado que su libertad a la hora de fijar el mediador implica que puede desarrollarse como una gran cantidad de distintos modelos, adaptados al caso que concierne en ese momento. También permiten simplificar los detalles técnicos de la interacción a alta energía y, dado que no fijan un mecanismo de mediación

entre el modelo estándar y la materia oscura, se pueden obtener conclusiones en primera aproximación; sin embargo, si la cinemática del proceso es accesible, debería reemplazarse por un modelo más completo que describa la mediación entre ambos sectores.

2.2.3. Modelos simplificados de mediación

En caso de que la energía de la colisión sea del orden de la masa del mediador o mayor, el estudio de este tipo de modelos se realiza de una forma específica. Los modelos más comunes incluyen mediadores neutros producidos únicamente que decaen en un par de partículas invisibles, como se muestra en la figura 2.1. Dichos modelos son muy utilizados debido a su similitud con la fenomenología de SUSY, y porque mediadores con color permiten incluir vértices con una única partícula de DM. Estos modelos pueden clasificarse en función del *spin* de su mediador: mediadores vectoriales de *spin* 1 (Z'), mediadores escalares (φ) y mediadores de *spin* 2.

Los acoplamientos de estos mediadores (g_q a los *quarks*, g_l a los leptones y g_χ a partículas invisibles), la masa de la partícula invisible (m_χ) y la masa del mediador (M_{med}) son parámetros libres en estos modelos. Los valores de estos acoplamientos se toman de forma que el modelo sea perturbativo en el espacio en que se realizan las búsquedas.

Los modelos que incluyen mediadores escalares sin carga de color son análogos al caso anterior de modelos *portal* de Higgs, pero con un mediador BSM. Ejemplos de estos procesos se muestran en la figura 2.2.

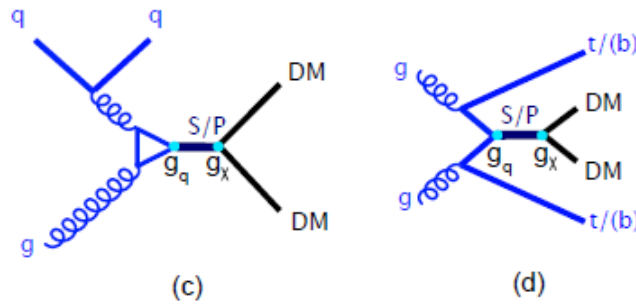


Figura 2.2: [5] Ejemplos de diagramas de Feynmann para un modelo con un mediador escalar (S) o pseudoescalar (P): (c) la interacción SM-mediador se realiza mediante un *loop* de gluones; (d) el mediador se produce en asociación a un par de *quarks* pesados.

Este tipo de modelos escalares está completamente determinado por la masa de la partícula invisible (m_χ), la masa del mediador (M_{med}), el acoplamiento φ -SM (g_q) y el acoplamiento φ -DM (g_χ). Estos modelos fijan el valor de g_q constante para todos los *quarks* y se considera un factor de los acoplamientos de Yukawa a los fermiones. Por otro lado, debe considerarse que este nuevo bosón escalar puede estar relacionado con el bosón de Higgs; en el caso de alta mezcla con él recuperaríamos el modelo *portal* del Higgs. Esto también permite restaurar la invariancia gauge a estos modelos.

2.2.4. Modelos supersimétricos (SUSY)

El desarrollo de teorías *supersimétricas* en física ha sido considerable en los últimos 30 años. La introducción de diversas teorías en este campo, a pesar de no poseer evidencias experimentales que corroboren su veracidad, está sin embargo justificada. Esto es debido a diversos efectos en teorías no supersimétricas, como el modelo estándar, que ponen de manifiesto diversos problemas no incluidos en el modelo. Sin embargo, la introducción de supersimetría permitiría explicar estos hechos y, por tanto, motiva la búsqueda de supersimetría en este ámbito.

La extensión de SUSY al modelo estándar postula la existencia de partículas *compañeras supersimétricas* para todas las partículas del modelo estándar con una diferencia de *spin* de $1/2$. En consonancia con lo comentado anteriormente, este tipo de modelos permiten solventar diversos problemas inherentes al modelo estándar actual, como la unificación de los acoplamientos.

Se denotan como *squarks* a los compañeros supersimétricos de los *quarks*. En el marco de SUSY, existen *squarks* a izquierdas o levógiros ($\tilde{q}_L$) y a derechas o dextrógiros ($\tilde{q}_R$); ambos se mezclan en dos autoestados de masa $\tilde{q}_1, \tilde{q}_2$ donde se denota por $\tilde{q}_1$ al más ligero. En el caso de *quarks* pesados, como el caso del top, se producen efectos de mezcla grandes; esto implica que su compañero supersimétrico, denotado *stop*, es el *squark* más ligero.

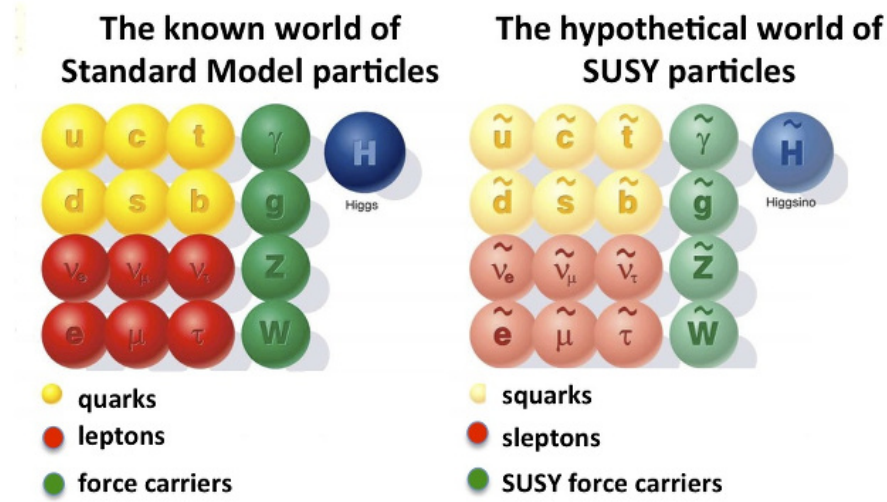


Figura 2.3: [12] Partículas elementales del modelo estándar (izquierda) y sus compañeros supersimétricos asociados (derecha).

En el marco del modelo estándar, el número cuántico de paridad $R = (-1)^{3(\mathcal{B}-L)+2S}$ ($\mathcal{B}$ representa el número bariónico, L el número leptónico y S el *spin*) toma siempre el valor 1 para todas las partículas, pero la introducción de SUSY al SM puede provocar un cambio en dicho valor. El modelo más simple que asegura la conservación de dicho número cuántico es el MSSM (*Minimal Supersymmetric Standard Model*); este modelo incluye al SM el compañero supersimétrico de cada partícula del modelo estándar, y por tanto dobla el número de partículas. Dentro de este modelo, las partículas supersimétricas se producen por pares.

Los compañeros supersimétricos de los bosones del SM se mezclan entre sí para dar distintos autoestados conocidos como *neutralinos* o *charginos*, en función de si son neutros eléctricamente o poseen carga, respectivamente. Las partículas supersimétricas que sufren esta mezcla son: *winos* (supercompañeros del campo Gauge SU(2) de los bosones W), *binos* (supercompañeros del campo Gauge U(1) de la carga electrodébil) y *higgsinos* (supercompañeros del bosón de Higgs). Las correspondientes mezclas dan lugar a dos *charginos*: $\tilde{\chi}_i^\pm$ ($i \in \{1, 2\}$) y cuatro *neutralinos*: $\tilde{\chi}_j^0$ ($j \in \{1, 2, 3, 4\}$); todos ellos ordenados en orden creciente de masa. En el marco del MSSM, el LSP es el más ligero de todos ellos, $\tilde{\chi}_1^0$.

La detección de estas partículas en el detector se basa en una traza de $\cancel{E}_T$ y alta multiplicidad de otros objetos (leptones o jets) debido a los decaimientos de supercompañeros pesados. Un diagrama de la extensión al modelo estándar introducida por el MSSM se puede observar en la figura 2.3.

2.2.5. Modelo simplificado estudiado

El modelo estudiado en este trabajo está basado en búsquedas de DM y partículas supersimétricas realizadas en ATLAS con datos de colisiones pp a una energía de centro de masas $\sqrt{s} = 13 \text{ TeV}$ tomados en el año 2016 en estados finales con *quarks* pesados (t, b) (véase [1], [2]).

Se trata de un modelo simplificado con un mediador de *spin* 0 y escalar (φ) entre el candidato a DM y las partículas del modelo estándar. Siguiendo la notación empleada anteriormente, este modelo dependerá de cuatro parámetros: la masa del mediador (m_φ), la masa del candidato (m_χ), la constante de acoplamiento φ -SM (g_q) y la constante de acoplamiento φ -DM (g_χ). Para el estudio realizado, se considera que $\mathbf{g}_\chi = \mathbf{g}_q = \mathbf{g} = \mathbf{1}$ y que $\mathbf{m}_\chi = \mathbf{1} \text{ GeV}$. El diagrama de Feynmann correspondiente a este proceso puede consultarse en la figura 4.1.

El canal estudiado es un canal de producción de materia oscura en asociación con un par de *quarks* pesados ($t\bar{t}$) en el canal de desintegración dileptónico; es decir, el proceso de desintegración $\mathbf{t} \rightarrow \mathbf{b}\mathbf{W}$ ($BR \approx 0.95$) y $\mathbf{W} \rightarrow \mathbf{l}\nu_l$ ($BR \approx 0.11$). El análisis realizado se restringirá al caso dileptónico $e^\pm\mu^\mp$ para eliminar la resonancia provocada por el bosón Z (en torno a 90 GeV). La motivación del uso de un canal dileptónico se basa en una limitación experimental: la detección de electrones y muones es más precisa que la detección de otros objetos, como *jets* hadrónicos. Por ello, muchos de los *triggers* involucrados en la detección en el LHC están basados en leptones. Este hecho se estudiará más a fondo en el siguiente capítulo.

Capítulo 3

El detector CMS

En este capítulo se incluirá una descripción del detector CMS del LHC empleado para la toma de datos utilizados en este trabajo. Se comenzará por un epígrafe sobre la instalación del LHC, seguido de una breve descripción del detector y las partes que lo componen; le seguirá la necesidad de incluir un *trigger* en la toma de datos. Por último, se describirá el proceso de detección de partículas en el detector, con el objetivo de comprender las distintas trazas que los distintos objetos involucrados en los sucesos dejan en las distintas componentes del detector; además, a lo largo del capítulo se introducirán diversos conceptos de carácter general (*jets, b-tagging,...*) empleados en este ámbito, con la finalidad de que el lector se sitúe en el marco actual de trabajo en el LHC.

3.1. LHC: el gran colisionador de hadrones

([18], [19]) El gran colisionador de hadrones o *large hadron collider* (LHC) es el colisionador de partículas más grande y potente del mundo. Está situado en las instalaciones del CERN, cerca de la ciudad de Ginebra, y consta de un anillo de unos 27 km de longitud situado bajo tierra, lo que le proporciona aislamiento frente a la radiación externa.

Está compuesto por dos tubos paralelos por donde viajan los haces de partículas que colisionan. Estos haces son acelerados mediante imanes superconductores, y se les proporciona una aceleración inicial mediante otros colisionadores del CERN empleados anteriormente al LHC que se hallan en esa localización (Linac 2, PSB, PS, SPS). Estos tubos se cruzan en cuatro lugares donde se encuentran los cuatro detectores del LHC: CMS (*Compact Muon Solenoid*), ATLAS (*A Thoroidal LHC Apparatus*), ALICE (*A Large Ion Collider Experiment*) y LHCb (*LHC- beauty*).

CMS y ATLAS se dedican a búsquedas de ámbitos más generales, estudiando toda clase de procesos, mientras que ALICE y LHCb se dedican a búsquedas en aspectos concretos, como su propio nombre indica.

En la actualidad, el acelerador opera a una energía de centro de masas $\sqrt{s} = 13$ TeV, aunque está diseñado para alcanzar en un futuro los 14 TeV.

3.2. Descripción general del detector

El detector CMS (*Compact Muon Solenoid*) es uno de los detectores situados en un punto de colisión del LHC (*Large Hadron Collider*), junto con ATLAS, ALICE y LHCb. Tiene forma cilíndrica y está colocado rodeando a los tubos del LHC, que lo atraviesan por su punto medio. Está compuesto por distintos subdetectores colocados en capas cilíndricas alrededor de los tubos del LHC. Este detector permite medir muones de forma efectiva, contiene un solenoide superconductor que produce un campo magnético de alta intensidad, y posee un calorímetro hadrónico hermético y un detector de trazas de alta precisión. Un corte transversal del detector se muestra en la figura 3.1.

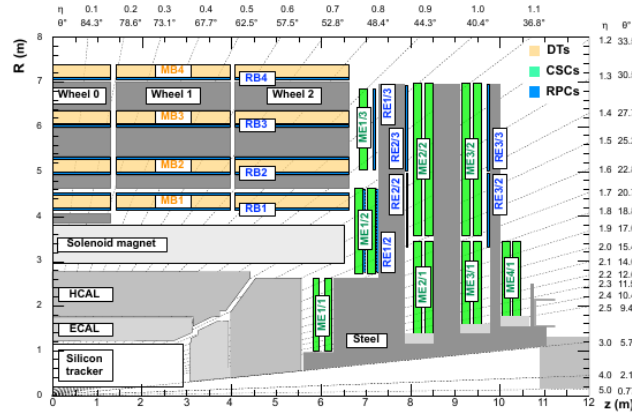


Figura 3.1: [20] Corte de un cuadrante en el plano R-z del detector CMS, con el tubo del LHC situado en el eje horizontal y el punto de interacción en el origen. Se muestran en distintos colores las componentes del detector de muones de CMS:

DTs (*Drift Tubes*), CSCs (*Cathode Strip Chamber*) y RPC (*Resistive Plate Chamber*).

El estudio de los procesos que tienen lugar en el LHC, y por tanto en CMS, se basa en un hecho importante: los protones viajan por el tubo del LHC, y por tanto no poseen una componente de momento en la dirección perpendicular a estos tubos.

Por tanto, debido a la ley de conservación de la energía y al tratarse de sistemas aislados, el momento de las partículas producidas en todas las direcciones perpendiculares al haz debe ser nulo. Este hecho motiva la definición de un plano compuesto por estas direcciones, perpendicular al tubo del LHC; este recibe el nombre de **plano transverso**. Una de las principales ventajas de emplear este plano es que, debido a que las componentes de *momento transverso* deben dar una suma nula, una desviación de esta magnitud de ese valor implicará la presencia de partículas invisibles no detectadas, como por ejemplo neutrinos o, en nuestro caso, candidatos a materia oscura.

El sistema de coordenadas empleado en la detección es un sistema cilíndrico basado en la geometría del detector. Por ello, se define cada punto del espacio mediante el siguiente sistema coordenado: $\mathbf{r}$ es la distancia al tubo del LHC, $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ el ángulo mostrado en la figura 3.1 y $\mathbf{z}$ su la distancia en el eje horizontal. Una transformación del ángulo θ muy empleada es la **pseudorapidez**: $\eta := -\ln \tan \theta$. A continuación se describirán las diferentes partes que componen el detector CMS, y se describirá brevemente cada una de ellas.

Detector de trazas o *tracker*

El detector de trazas, más comúnmente llamado *tracker*, es el detector encargado de recoger la traza de las partículas cargadas que lo atraviesan. Se encuentra situado en la parte más interna del detector, y permite el cálculo del momento de las partículas cargadas; esto es debido a que éstas siguen una trayectoria helicoidal en su interior debido a la presencia del campo magnético.

Está compuesto por una gran cantidad de píxeles de silicio, operando cada uno de ellos a una resolución espacial de $\pm 10\mu\text{m}$, lo que dota a este detector de una gran precisión en la determinación de las trazas de las partículas cargadas. Por contraparte, el gran número de conexiones eléctricas provocaría un enorme calentamiento del detector, por ello esta zona se encuentra refrigerada con gases a bajas temperaturas.

Calorímetro electromagnético (ECAL)

Este detector se emplea con la finalidad de detectar deposiciones de energía de partículas cargadas eléctricamente, y también es el encargado de detectar la energía de los fotones producidos. Está formado por cristales de centellador colocados de forma hermética, de forma que se minimice la posibilidad de que una partícula cargada no sea detectada. Estos centelladores tienen una respuesta muy rápida,

lo que permite la determinación de la energía de estas partículas en los instantes anteriores a una nueva colisión. Este componente absorbe toda la energía de los hadrones cargados que lo atraviesan.

Calorímetro hadrónico (HCAL)

Este detector tiene una finalidad similar al calorímetro electromagnético, esta vez en la detección de partículas hadrónicas. Está aislado herméticamente, de forma que es muy preciso en la detección del momento de las partículas y permite el estudio de procesos en que se produzca un desbalance de energía, por ejemplo por producción de una partícula invisible. Está compuesto por varias capas de materiales centelladores, de forma que la energía de una partícula puede ser determinada sumando las energías detectadas en cada capa. También permite absorber toda la energía de los hadrones neutros que lo atraviesan.

Solenoides superconductor

El solenoide es el componente del detector encargado de producir el campo magnético. Su funcionamiento permite la detección del momento de las partículas cargadas en el detector y su construcción permite que opere a un campo magnético de $3.8T$. Se trata además del mayor imán superconductor jamás construido. Estos campos magnéticos se alcanzan aplicando propiedades de materiales superconductores; notar además que el campo magnético tiene direcciones opuestas en la zona interior del solenoide (*tracker*, ECAL, HCAL) y en su exterior (cámaras de muones).

Cámaras de muones

Estas cámaras están dedicadas a detectar los muones, que por su comportamiento atraviesan el resto de componentes del detector. Están divididas en tres componentes distintas que, en combinación, permiten tanto determinar con precisión la trayectoria de los muones como hacerlo de forma rápida. Estas componentes se muestran en la figura 3.1 y son: *Drift Tubes* (DTs), *Cathode Strip Chambers* (CSCs) y *Resistive Plate Chambers* (RPCs). Mientras que los DTs se emplean en la zona de 'barril' del detector, correspondiente a $|\eta| < 1.2$, y los CSCs en la zona correspondiente a $0.9 < |\eta| < 2.4$, los RPCs los complementan y permiten una respuesta rápida en la detección de los muones.

3.3. Luminosidad, *pile-up* (PU) y *trigger*

Una de las magnitudes más importantes a tener en cuenta en los colisionadores de partículas es la **luminosidad**. Esta magnitud puede interpretarse como una medida del número de colisiones detectadas por unidad de tiempo y superficie. En su versión más simple, puede expresarse como:

$$\mathcal{L} := \frac{f N_1 N_2}{4\pi\sigma_x\sigma_y}, \quad [\mathcal{L}] = b^{-1}s^{-1} \quad (3.1)$$

donde N_1, N_2 representan el número de *bunches* de protones que colisionan, f es la frecuencia de rotación de dichos *bunches* alrededor del LHC y $\sigma_{x,y}$ representan las dimensiones de cada *bunch* en el punto de colisión. La luminosidad actual en el LHC es tal que se produce una interacción entre dos *bunches* de protones cada unos 25 ns. Al tratarse de paquetes de protones, en cada uno de estos sucesos ocurre un alto número de diversas interacciones entre ellos. El valor medio de estos sucesos recibe el nombre de *pile-up* (PU), y es un efecto no deseado debido a la dificultad que impone en la toma de medidas de los procesos. Este efecto empeora con la luminosidad; a mayor luminosidad, mayor número de sucesos tendremos y, por tanto, mejor serán las estimaciones realizadas, y sin embargo eso implicará un aumento del PU.

Por otro lado, el gran número de colisiones por segundo implica que no es posible registrar todas las colisiones detectadas, debido a que la tecnología actual no permite almacenar tal cantidad de datos. Por ello, existen numerosos dispositivos (tanto *hardware* como *software*) que permiten identificar con gran velocidad los eventos que en ese momento interesan, y por tanto descartar el resto y ser capaces de almacenarlos. Este tipo de herramientas, en su conjunto, se conoce como *trigger*. Una corta descripción del *trigger* de CMS se expone a continuación, aunque una visión más detallada de su funcionamiento se puede consultar en ([21]):

Hay una primera fase de selección de sucesos con *hardware* instalado en el propio detector, que en el margen de 4 μs toma la decisión de registrar o no un evento mediante la información obtenida en los calorímetros y las cámaras de muones; este *trigger* se conoce como *level-1 trigger* (L1). Este sistema está acoplado a un conjunto de procesadores comerciales, denominado *high level trigger* (HLT), que opera con máquinas en paralelo y analiza de forma *offline* los datos tomados por el L1 *trigger*. Por último, se toman copias de seguridad de los datos y se envían a los centros de distintos organismos para su estudio.

3.4. Identificación y simulación de objetos

Como ha sido comentado en apartados anteriores, la gran cantidad de información recogida en el detector hace indispensable el uso de herramientas informáticas para procesarla y analizarla. El trabajo de recogida y análisis de estos datos se realiza de la siguiente forma: existe una primera recogida de datos a nivel de *trigger*, tras lo que esta información es recogida en centros distribuidos por el mundo, y a continuación se incluye el proceso de reconstrucción de dichos datos, donde se trata de emplear los datos tomados por los distintos subdetectores de CMS para reconstruir el momento de las partículas del proceso en el estado final; por ejemplo, el número de leptones, el momento de las partículas, etc.

A continuación se describirán el proceso de identificación de objetos en el detector CMS y el proceso de simulación de las muestras de procesos de MC empleadas en el análisis posterior de los datos tomados.

3.4.1. Identificación de objetos en el detector CMS

En este apartado se introducirá el proceso de identificación de los distintos objetos de un suceso a partir de los datos tomados por el detector. El algoritmo empleado en CMS para esta tarea se conoce como **particle flow** (PF) ([22]). Dicho algoritmo se basa en la toma de datos directos del detector y permite la reconstrucción de los objetos principales del evento: electrones, muones,... Por un lado, el algoritmo reconstruye las trazas o trayectorias de las partículas que atraviesan el detector mediante las señales que se detectan tanto en el *tracker* como en las cámaras de muones. Por otro lado, se reconstruyen objetos llamados *clusters*, correspondientes con las deposiciones de energía en los calorímetros.

Tras esta clasificación, el algoritmo combina la información recogida para llevar a cabo la reconstrucción de los objetos físicos del suceso. La reconstrucción sigue además un orden predeterminado: en primer lugar, se reconstruyen los muones debido a que las trazas en las cámaras de muones son muy características, y eso permite discriminar las pequeñas deposiciones de energía en el ECAL correspondientes a ellos; después, se reconstruyen los electrones y hadrones cargados, debido a sus trazas en el *tracker* y sus deposiciones de energía en el ECAL y HCAL respectivamente. Por último, los *clusters* restantes se corresponderán con fotones en el caso del ECAL, y hadrones neutros en el caso del HCAL.

Los criterios de reconstrucción de estos objetos se enuncian brevemente a continuación:

Vértice primario

Se conoce como vértice primario al punto de interacción entre los dos protones del suceso. Como se comentó anteriormente, el *pile-up* es un problema en el posterior análisis de datos; uno de los problemas asociados al *pile-up* es la identificación del vértice primario. Por ello, aunque los vértices de las interacciones se reconstruyen mediante el algoritmo PF extrapolando las trayectorias de las partículas detectadas hasta el tubo del LHC, podremos reconstruir diversos vértices en estas condiciones. Por ello, se toma el vértice primario como aquel cuyos objetos poseen mayor momento transversal. La reconstrucción de vértices es especialmente importante en el caso de los *jets* provenientes de *quarks* b, como se observará posteriormente.

Muones

En la reconstrucción de muones se toman las trazas dejadas por la partícula en las cámaras de muones, y se extrapolan a las trazas dejadas en el *tracker*; todo ello teniendo en cuenta efectos de pérdida de energía y su trayectoria curva debido al campo magnético producido por el solenoide. Los criterios de selección de este tipo de objetos se basan en gran medida en los valores de su momento transversos, su pseudorapidez y variables que estudian la compatibilidad de la traza del *tracker* y la traza del sistema de muones, entre otros. Estos criterios pretenden asegurar que el muon reconstruido y aceptado como tal proviene efectivamente de la interacción en el vértice primario, y no se debe a procesos posteriores en los que no se está interesado.

Otra variable de discriminación empleada se conoce como **aislamiento**, y se define como:

$$\Delta R := \sqrt{(\eta_1 - \eta_2)^2 + (\varphi_1 - \varphi_2)^2} \quad (3.2)$$

Esta variable se interpreta como la distancia del muon a la distribución en forma de cono del *jet* que se produce en el proceso hadrónico, permitiendo discriminar muones provenientes de la desintegración de un *quark*. La mayor parte de los procesos que en el caso aquí expuesto influirán en este fondo se corresponden con *jets* provenientes de *quarks* b, en procesos como $W + Jets$; este proceso se describirá con mayor profundidad posteriormente.

El momento transversal de los muones es un criterio de selección importante en el caso de eliminar procesos no interesantes. Por ello, el criterio usual es exigir muones cuyo momento transversal sea al menos mayor a 15 GeV. Otro criterio tomado

usualmente es sobre la pseudorapidez, tomando $|\eta| < 2.4$ debido a que es la zona del espacio que cubre el *tracker*.

Electrones

Para el caso de la reconstrucción de electrones se procede de forma similar al caso de reconstrucción de muones, esta vez con trazas en el *tracker* y deposiciones de energía en el ECAL. Sin embargo, dado que no se dispone de trazas en las cámaras de muones, deben tenerse varios factores en cuenta: por un lado, dado que su masa es menor que la masa de los muones, la radiación que emiten por efecto *Bremsstrahlung* es más considerable y puede afectar a su trayectoria; además, la reconstrucción con los *clusters* del ECAL debe tener en cuenta la presencia de deposiciones debida a fotones.

Hadrones y *jets*

Como se comentó en el apartado anterior, los *quarks* sufren hadronizaciones en tiempos mucho más cortos que los tiempos necesarios para sufrir desintegraciones debido al *confinamiento*. Esto es cierto para todos ellos exceptuando el *quark top*, cuyo tiempo de vida es más corto que el tiempo medio de hadronización; este efecto es debido a su alta masa. Este tipo de procesos puramente fuertes conllevan una cascada de desintegraciones y hadronizaciones, produciendo diversas partículas; la interpretación de estos procesos es muy complicada debido a que la fuerza fuerte es no perturbativa. Estas partículas tienden a colimarse en un haz con forma de cono cuya dirección coincide con la dirección de producción del *quark* principal del proceso; a este haz se le conoce como *jet*. La reconstrucción de estos objetos se realiza agrupando las partículas provenientes de estos *jets*.

En general no es posible distinguir el *quark* que produjo en primer lugar el *jet*, excepto en dos casos particulares: los *quarks top*, con su bajo tiempo de vida, decaen inmediatamente después de ser producidos; el otro caso es los *quarks bottom*. Este tipo de *quarks* producen una señal característica que permite su discriminación del resto de *jets*; está basada en la detección de un vértice secundario. Los *quarks b* tienden a hadronizarse tras su producción en hadrones con *quarks b*, cuyo tiempo de vida es superior al del resto de hadrones. Esto permite, con la precisión del detector, la reconstrucción de un segundo vértice correspondiente al *jet* producido por dicho hadrón en el interior del *jet* primario. Este hecho permite discriminar entre los conocidos como *b-jets* y el resto de *jets* posibles. Puede observarse este proceso en la figura 3.2.

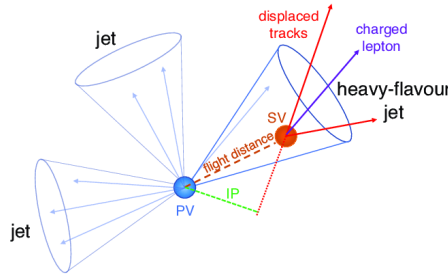


Figura 3.2: [23] Diagrama de representación de la producción de un *b-jet*. PV representa el vértice primario e IP el parámetro de impacto.

El algoritmo encargado de la identificación de este tipo de *jets* es el algoritmo CSV. Combina información relativa al vértice principal, parámetro de impacto y vértice secundario, entre otras para la reconstrucción de estos objetos; su salida es un valor en el intervalo $[0, 1]$ que indica cuán seguro se está de que se trate de un *b-jet*, siendo más probable cuanto mayor sea. Este proceso se conoce como etiquetado, y por tanto se refiere a él como *b-tagging*. Exigencias en esta variable, junto al momento transversal y la pseudorapidez se emplean en la identificación de este tipo de objetos.

Leptones τ

Este tipo de objetos no pueden ser detectados como los anteriores debido a la masa del tau, $m_\tau = 1.777$ GeV, este valor implica que es el único leptón capaz de desintegrarse hadrónicamente y en el interior del *tracker*. El caso en que dicho leptón se desintegre en muones o electrones es detectado por el algoritmo PF pero es difícil de discriminar del caso de electrones o muones por separado; en el caso de su desintegración hadrónica, es posible reconstruir la procedencia de sus productos de desintegración. Este tipo de objetos se conocen como *taus hadrónicos* (τ_h).

Neutrinos y otras partículas invisibles

En los procesos que ocurren en el detector, una gran parte de ellos incluirán la producción de neutrinos (o, posiblemente, otras partículas invisibles BSM). Por ello, la identificación de este tipo de objetos es crucial en el análisis de los datos. Sin embargo, dado que este tipo de partículas no interactuarán en el interior del detector y lo atravesarán sin dejar rastro, es necesario incluir una magnitud que permita reconstruir este tipo de objetos. La variable que se define para este propósito se llama *momento transversal faltante* o *energía transversal faltante*, denotado comúnmente

en la literatura como MET o $\cancel{E}_T$. Se define como el módulo del vector inverso de la suma de los momentos transversos de todos los objetos detectados en el suceso:

$$\cancel{E}_T := \left| - \sum_i \vec{p}_T(i) \right| \quad (3.3)$$

Una consideración a tener en cuenta es que esta partícula nos daría el hipotético momento transverso de una partícula invisible producida en el suceso, pero si en dicho proceso se producen varias partículas invisibles, esta variable representaría el momento transverso de todos ellos a la vez, no permitiendo discriminar entre ellos; por ejemplo, en el caso de dos neutrinos. Esta variable es una de las principales utilizadas en el estudio de procesos con este tipo de partículas.

3.4.2. Simulación de procesos en el detector CMS

La simulación de datos mediante muestras de Monte Carlo es necesaria para el posterior análisis de los datos tomados. Estas simulaciones incluyen los cálculos de las secciones eficaces teóricas a distintos órdenes, y permiten emular el paso de las partículas consideradas por el detector, considerando todos los posibles efectos. Este proceso se conoce como de **generación**. Este tipo de muestras permiten predecir de antemano los sucesos de fondo que contendrá nuestra búsqueda, y para estimar si la búsqueda realizada es sensible a la señal considerada.

Por tanto, el método de trabajo es el siguiente: se generan los procesos físicos en que estamos interesados, y con el conocimiento de la disposición y el funcionamiento de todas las componentes del detector, se simula su paso por el detector y se registran los datos que el detector tomaría (trayectorias en el *tracker*, número de leptones, deposiciones de energía en el ECAL y HCAL,...). Después, se trabajaría con estos datos de igual forma que se tratan los datos reales obtenidos, para evitar diferencias en su tratamiento y obtener una conclusión imparcial. Esto permite incluir unas simulaciones de MC a los datos obtenidos que permitirán discriminar de forma efectiva las señales de los fondos pertinentes. Es importante remarcar que este tipo de simulaciones dependen en gran medida del conocimiento del detector para realizar estas simulaciones con la precisión adecuada a nuestras medidas: la disposición de cada componente del detector (píxeles, cristales,...), su eficiencia, etc. Todos estos factores pueden influir en las simulaciones y, por tanto, en las conclusiones obtenidas de los datos.

La simulación de MC de los objetos en el detector se realiza mediante distintos programas, entre los que se encuentran **Powheg** y **aMC@NLO**, empleados en

conjunto a un *software* de simulación, como **Pythia**, en la reconstrucción de las 'cascadas' producidas en los procesos y que serán los empleados en este trabajo. Se entienden por cascadas a los diversos procesos posteriores a la interacción principal que dan lugar al estado final observado. Tras ello, el *software* **Geant4** se emplea para simular la interacción de estos sucesos con el detector CMS. Este tipo de simulaciones están basadas en teorías perturbativas de los modelos que rigen las partículas del modelo estándar (QCD, QED,...), y por tanto se clasifican en relación al orden en teoría de perturbaciones alcanzado: a primer orden (LO, *leading order*), a segundo orden (NLO, *next-to-leading order*), tercer orden (NNLO, *next-to-next-to-leading order*), etc. Como es lógico, el uso de muestras simuladas a mayores órdenes de teoría de perturbaciones implicará un mayor acuerdo con el valor real por el modelo estándar actual, y por tanto una mayor discriminación en la búsqueda de señales con nuevas partículas o en la determinación de propiedades de partículas del modelo estándar.

El uso de muestras de MC es esencial en el análisis realizado en los detectores, y es inmediato notar que un mayor número de simulaciones permitirá reducir la incertidumbre asociada a este tipo de muestras. Por ello, el número de simulaciones realizadas es mucho mayor que el número de datos tomados en el detector, y se sigue actualmente incrementando para permitir, por ejemplo, la discriminación respecto a señales cuya sección eficaz es muy baja y puede confundirse con incertidumbre estadística debida a esta simulación. Esto también implica que los sucesos medidos deben ser *pesados* al número de sucesos simulados empleados y a la sección eficaz teórica del proceso estudiado.

Capítulo 4

Descripción del análisis de datos

En este capítulo se procederá a describir el trabajo realizado en el contraste de veracidad del modelo considerado. Se comenzará con un primer apartado donde se expondrá el modelo señal estudiado, seguido de los principales fondos involucrados; comprenderá una breve descripción de dichos procesos y se argumentará la importancia de su inclusión en nuestro estudio. A continuación, se describirán los datos y las muestras de Monte Carlo empleadas, así como el *framework* empleado en la realización del trabajo. Posteriormente, se procederá a la descripción de las variables que se emplearán en la discriminación de nuestra señal respecto a los distintos fondos, los criterios de selección empleados y las estimaciones de los fondos y las incertidumbres sistemáticas consideradas.

4.1. Procesos involucrados en el análisis

En este apartado se describirán brevemente los procesos considerados en el desarrollo del trabajo, incluyendo uno de los principales diagramas de Feynmann representativo de dicho proceso y argumentando la razón de su inclusión en el análisis realizado.

4.1.1. El proceso $t\bar{t} + \text{DM}$

El proceso indicado como $t\bar{t} + \text{DM}$ es el proceso de señal considerado en este trabajo. Se trata de un proceso de producción de un par de *quarks top-antitop* ($t\bar{t}$) acompañado de la producción de una partícula invisible candidato a DM. La producción de este par puede ser debida al proceso $q\bar{q} \rightarrow t\bar{t}$ o el proceso $gg \rightarrow t\bar{t}$, conocido este último como fusión de gluones (*gluon fusion*) y siendo la fuente

principal de producción de este par, siendo en el caso del LHC a una energía $\sqrt{s} = 13$ TeV del orden del 90 % ([24]). En la figura 4.1 se muestra uno de los posibles diagramas de Feynmann a primer orden de este proceso.

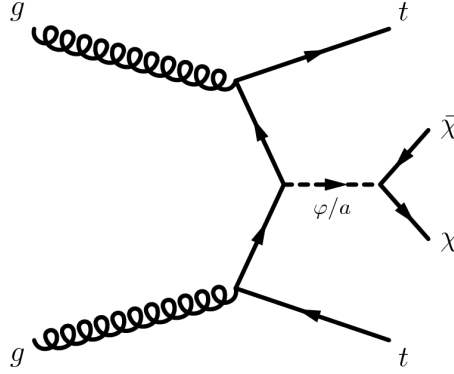


Figura 4.1: [2] Diagrama de Feynmann a primer orden del proceso de producción de un par $t\bar{t}$ y una partícula invisible (χ) mediante un mediador: escalar (φ) o pseudoescalar ligero (a).

Suponiendo la estabilidad del candidato a DM, este proceso estará caracterizado por un estado final similar al correspondiente a un proceso $t\bar{t}$, incluyendo un par de partículas invisibles adicional. Como se comentará en el apartado siguiente, este tipo de procesos conllevan un estado final con neutrinos, y como es conocido, este tipo de partículas no son detectadas. Debido al principio de conservación de la energía, la traza de este tipo de partículas es *momento transverso faltante* en nuestro suceso, representado como $\cancel{E}_T$. Por tanto, este tipo de procesos conllevan la producción de un par de *quarks* top-antitop acompañado de dos partículas invisibles adicionales; por ello, es de esperar que aunque el estado final de ambos procesos sea indistinguible, la magnitud de $\cancel{E}_T$ sería mayor para el proceso considerado en este apartado.

4.1.2. El proceso $t\bar{t}$

El proceso $t\bar{t}$ es uno de los procesos más estudiados en el ámbito del LHC. Este proceso tiene varias formas de producción representadas por sus correspondientes diagramas de Feynmann; en particular, una de ellas posee un diagrama de Feynmann similar al indicado en la figura 4.1 en el caso de *gluon fusion*. Por tanto, como se comentó anteriormente, este proceso será uno de los principales fondos a considerar en este trabajo.

Este proceso viene caracterizado por las fracciones de desintegración (BR, *branching ratio*) de los *quarks* t , que en un 96 % de las veces ([3]) se corresponde con

una desintegración en un *quark* b y un bosón $W^\pm$. Por tanto, el estado final de este proceso está caracterizado por dos *quarks* b y los productos de desintegración de los bosones $W^\pm$: leptónicos o hadrónicos. En el caso aquí considerado, estudiando el proceso con ambos bosones W desintegrándose leptónicamente (conocido como el canal *dileptónico*), el estado final estará acompañado de dos neutrinos.

4.1.3. El proceso Drell-Yan (DY)

Este proceso recibe este nombre por el estudio realizado por *Sidney Drell* y *Tung-Mow Yan* (1970) para describir la producción de pares leptón-antileptón en aceleradores hadrónicos a altas energías. Se trata de un proceso de aniquilación de un *quark* proveniente de uno de los hadrones con un *antiquark* proveniente del otro hadrón; de esta forma se produce un bosón Z o un fotón virtual que produce el par leptón-antileptón. Este proceso está recogido en la figura 4.2.

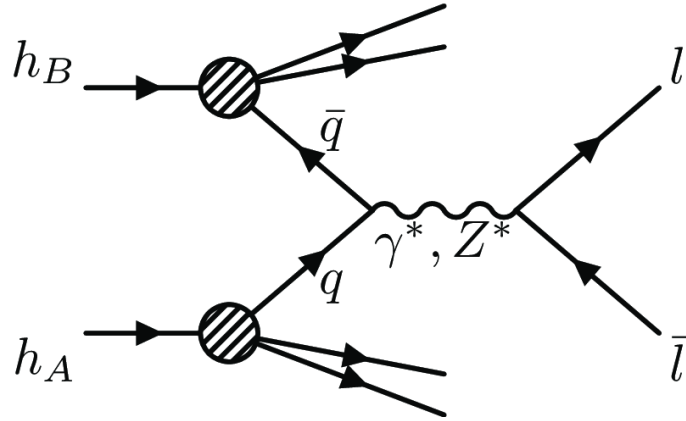


Figura 4.2: [7] Diagrama de Feynmann del proceso Drell-Yan (DY) entre dos partículas hadrónicas (h_A, h_B) para producir un par $l\bar{l}$ mediante un fotón virtual (γ^*).

Este proceso es un fondo importante en el estudio de procesos con estados finales dileptónicos, por lo que es importante de considerar en nuestro análisis. Por otro lado, se mostrará que el uso de ciertas variables de selección permite discriminar de forma apreciable este tipo de procesos de los procesos de señal considerados.

4.1.4. Los procesos $t\bar{t}W$ y $t\bar{t}Z$

Ambos procesos se basan en la producción de un bosón $W^\pm, Z$ en conjunción a un par de *quarks* $t\bar{t}$. Sus diagramas de Feynmann dominantes a primer orden (LO, *leading order*) se recogen en la figura 4.3.

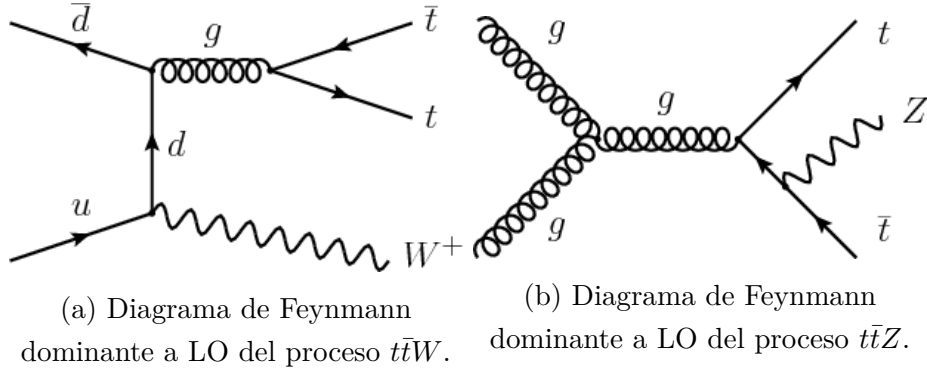


Figura 4.3: [8] Diagramas de Feynmann dominantes en la producción de los procesos $t\bar{t}W$, $t\bar{t}Z$.

Este proceso es importante de considerar en el estudio realizado por varios motivos. En primer lugar, a pesar de tener baja sección eficaz, constituyen un fondo importante en el estudio de procesos que no se corresponden con $t\bar{t}$; debe tenerse en cuenta que la sección eficaz de producción en el LHC de dos *jets* cualesquiera, por ejemplo $b\bar{b}$, es mucho mayor que la sección eficaz de producción de sucesos $t\bar{t}$. Por otro lado, esto implica la posibilidad de producción de bosones $W^\pm$ junto al par de *quarks* $t\bar{t}$, y por tanto confundirse con la señal considerada.

4.1.5. Los procesos dibosónicos: WW, WZ, ZZ

Este tipo de procesos se caracterizan por tener un estado final característico de eventos dibosónicos, como se comenta a continuación. Se muestra en la figura 4.4 unos diagramas de Feynmann dominantes en este tipo de procesos en el canal de desintegración leptónico.

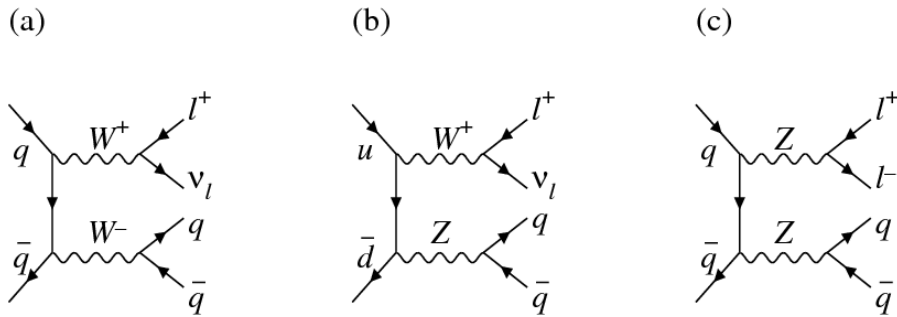


Figura 4.4: [9] Diagramas de Feynmann dominantes de los procesos dibosónicos en el canal leptónico.

Estos procesos deben ser considerados debido a la presencia de leptones en el estado final provenientes de bosones W, Z . Notar que otros estados finales posibles pueden ser: dileptónicos (en el caso de una desintegración leptónica de ambos bosones), completamente hadrónicos (lo que conlleva un mayor número de jets en el estado final) o desintegraciones de un bosón Z en partículas invisibles (debido a su canal de desintegración en dos neutrinos; el *branching ratio* de este suceso es aproximadamente un 20 %), además de estados finales trileptónicos y con hasta cuatro leptones. Por ello, de nuevo trazas en $\cancel{E}_T$ indican que este es un fondo considerable.

4.1.6. El proceso $W + Jets$

Este proceso está caracterizado por la producción de un bosón $W^\pm$ y una emisión a segundo orden de un gluón, que como se comentó anteriormente da lugar a un *jet hadrónico*. Se puede observar un diagrama de Feynmann de este proceso en la figura 4.5.

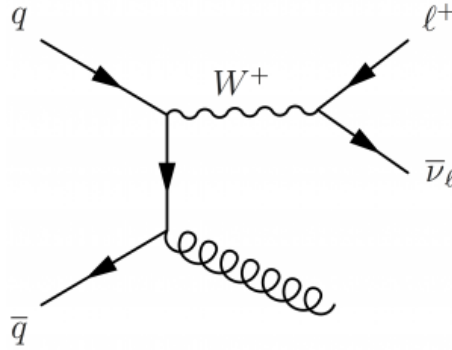


Figura 4.5: [10] Diagrama de Feynmann dominante del proceso $W + Jets$ en el canal leptónico.

Este tipo de procesos deben ser considerados como un fondo de nuestras señales en el canal dileptónico debido a que, además de la presencia de un leptón proveniente del bosón W , existe la posibilidad de producción de un leptón adicional asociado a un *jet*, como se comentó en la definición de la variable de aislamiento. Otro motivo atendiendo al estado final del proceso de señal considerado es la presencia de una traza de $\cancel{E}_T$ proveniente del neutrino asociado al bosón W .

4.1.7. El proceso tW

Este proceso es otro de los principales fondos para el proceso $t\bar{t}$, debido a que los diagramas de Feynmann a NLO (figura 4.6) interfieren con los diagramas de producción del proceso $t\bar{t}$.

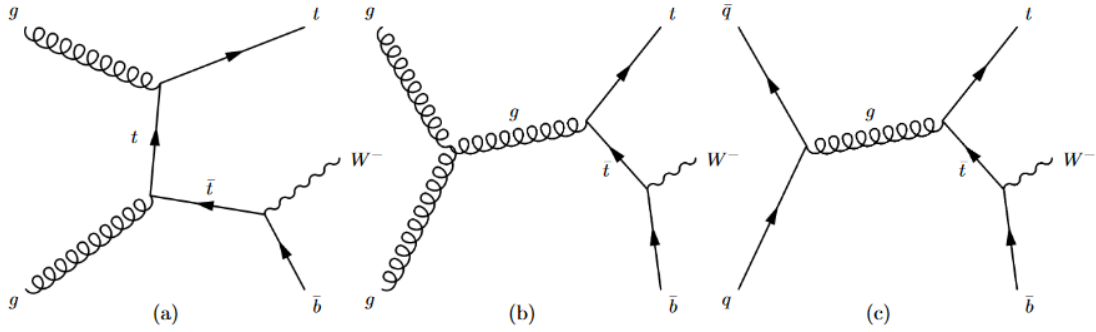


Figura 4.6: [11] Ejemplos de diagramas de Feynmann del proceso tW a NLO que interfieren con $t\bar{t}$.

Notar que en el estado final de este tipo de procesos se producen un *quark* top y un *quark* bottom asociado a un bosón W (ya que como se comentó anteriormente, se cumple que el *branching ratio* del proceso $t \rightarrow Wb$ es aproximadamente el 96 %); esto implica que en el estado final se observarán dos *jets* correspondientes a dichos *quarks* y los productos de desintegración del bosón W , ya sea leptónico o hadrónico. Por ello, este fondo también será considerado.

4.2. Descripción de datos y muestras de Monte Carlo (MC) empleadas

En este trabajo se ha llevado a cabo un estudio empleando programas enmarcados en **ROOT**, un paquete de programas en C++ empleado ampliamente en física de partículas para realizar este tipo de análisis. Se ha empleado como base del análisis un *framework* denominado *AnalysisPAF* ([25]), desarrollado por el Grupo Experimental de Altas Energías de la Universidad de Oviedo; un conjunto de programas que incluyen herramientas que permiten el estudio de diversos procesos de interés. Empleando este código como base, se han desarrollado programas en **Python** y **C++** con el objetivo de diseñar un análisis propio de los datos obtenidos. Este estudio se enmarca en un conjunto de librerías que permiten el estudio modular de estos procesos denominado *PAF* (*Proof Analysis Framework*, [26]). El código

diseñado se ejecuta en los equipos del grupo, operando con máquinas de sistemas operativos GNU/Linux. Para la ejecución y control de los programas realizados se ha empleado **Bash**, y para la obtención del código común se ha empleado la librería de repositorios **git**.

Los datos empleados en el transcurso del trabajo fueron tomados en CMS durante el año 2016 en colisiones pp a una energía de centro de masas $\sqrt{s} = 13$ TeV. Se emplearán tres muestras de datos, agrupadas en función del *trigger* que cumplen: *SingleMuon*, *SingleElec* y *MuonEG*.

Respecto a las muestras de MC tomadas, los procesos físicos que representan, su agrupación para su posterior análisis y su sección eficaz de producción se encuentran recogidos en la tabla 4.1; en ella se incluyen los procesos comentados en la sección inicial del capítulo.

Proceso	Canal de desintegración	Grupo	Sección eficaz (pb)
$t\bar{t}W$	$W \rightarrow l\nu_l$	$VV + t\bar{t}V$	0.2043
$t\bar{t}W$	$W \rightarrow qq'$	$VV + t\bar{t}V$	0.4062
$t\bar{t}Z$	$Z \rightarrow qq'$	$VV + t\bar{t}V$	0.5297
$t\bar{t}Z$	$Z \rightarrow W^+W^- \rightarrow l\nu_l l'\nu_{l'}$	$VV + t\bar{t}V$	0.2529
WW	Canal dileptónico	$VV + t\bar{t}V$	115
WZ	Canal dileptónico	$VV + t\bar{t}V$	47.13
ZZ	Canal dileptónico	$VV + t\bar{t}V$	16.523
$DY + Jets$	$\gamma \rightarrow l\bar{l}, m_{l\bar{l}} \in [10, 50]$ GeV	$Z/\gamma^* \rightarrow e^\pm \mu^\mp$	22635.09
$DY + Jets$	$\gamma \rightarrow l\bar{l}, m_{l\bar{l}} = 50$ GeV	$Z/\gamma^* \rightarrow e^\pm \mu^\mp$	6025.2
$t\bar{t}$		$t\bar{t}$	831.76
$t\bar{t}$	<i>Non-prompt leptons</i>	Non W/Z	-
$W + Jets$	$W \rightarrow l\nu_l$	Non W/Z	61526.7
tW		tW	35.85
$\bar{t}W$		tW	35.85
$t\bar{t} + DM$	$t\bar{t} \rightarrow l\nu_l l'\nu_{l'}, m_\varphi = 5$ GeV	$t\bar{t}\varphi (5, 1)$ GeV	3.478
$t\bar{t} + DM$	$t\bar{t} \rightarrow l\nu_l l'\nu_{l'}, m_\varphi = 10$ GeV	$t\bar{t}\varphi (10, 1)$ GeV	2.05603
$t\bar{t} + DM$	$t\bar{t} \rightarrow l\nu_l l'\nu_{l'}, m_\varphi = 20$ GeV	$t\bar{t}\varphi (20, 1)$ GeV	1.09991
$t\bar{t} + DM$	$t\bar{t} \rightarrow l\nu_l l'\nu_{l'}, m_\varphi = 50$ GeV	$t\bar{t}\varphi (50, 1)$ GeV	0.308667

Tabla 4.1: [27] Tabla de muestras de simulaciones MC empleadas en el trabajo.

Los procesos se clasifican en varios grupos en función del fondo que representan: en primer lugar se encuentran los procesos $t\bar{t}$ acompañados de un bosón y los procesos dibosónicos, en segundo lugar los procesos Drell-Yan para distintos valores

de masa invariante de ambos leptones, en tercer lugar el grupo Non W/Z , donde se incluyen el proceso $W + Jets$ y los leptones no correspondientes al proceso principal, llamados *non-prompt leptons*; estos últimos se incluyen para estudiar su influencia en el proceso de señal, por ejemplo, en el caso de muones provenientes de la desintegración de un *quark* b . A continuación se incluye el grupo tW , y por último las cuatro muestras de señal del proceso $t\bar{t} + DM$ para cuatro posibles valores de masa del mediador escalar φ : 5, 10, 20 y 50 GeV, a una masa fija del candidato a DM χ de 1 GeV. Todas estas muestras se toman en el canal de desintegración dileptónico por los motivos expuestos anteriormente.

El *trigger* tomado en la selección de estos sucesos está basado en leptones, como es en la mayoría de casos en este tipo de detectores por las ventajas expuestas anteriormente. En este caso, el *trigger* empleado está basado en eventos que cumplen uno de los siguientes criterios de selección:

- $p_{\mu}^T \geq 24 \text{ GeV}$
- $p_e^T \geq 32 \text{ GeV}$
- $p_{\mu}^T \geq 8 \text{ GeV}, p_e^T \geq 23 \text{ GeV}$
- $p_{\mu}^T \geq 23 \text{ GeV}, p_e^T \geq 12 \text{ GeV}$

es decir, se trata de un *trigger* de muones, electrones o una combinación de ambos. Los requerimientos en el caso electrónico son mayores debido a la facilidad de identificación de los muones respecto a los electrones, como fue expuesto anteriormente.

4.3. Variables cinemáticas de discriminación

En esta sección se expondrán los distintos observables empleados en la discriminación de las señales propuestas y los fondos considerados en este trabajo. El empleo de estos observables tiene un fundamento experimental motivado por los modelos empleados, todos ellos están fundamentados en estudios realizados en este ámbito ([1],[2]). También se incluirá una descripción teórica de cada variable, y se incluirán dos gráficos en cada caso: el primero representa la variable indicada en cada caso (incluyendo las muestras de datos empleadas), que se incluye para comprobar la correcta modelización del observable; el segundo es una comparación del fondo principal del proceso ($t\bar{t}$) frente a la señal más característica ($t\bar{t}\varphi$ (50, 1) GeV). Se

representan los histogramas normalizados al número total de sucesos de cada proceso, y se emplea la señal con mediador de masa 50 GeV debido a que, como muestra la tabla 4.1, se trata de la señal con menor sección eficaz y por tanto permitirá una determinación de la región de señal de una forma más efectiva. El objetivo de esta segunda gráfica es mostrar la región del espacio de fases correspondiente donde se maximizan las señales respecto a los fondos considerados; en el siguiente apartado, esto permitirá definir una región de señal donde realizar el análisis de datos considerado óptimo en este trabajo.

Todas las gráficas mostradas de ahora en adelante contemplarán eventos que se requerirá cumplan las siguientes condiciones:

- (i) **Leptones de distinto signo:** como se comentó anteriormente, el análisis realizado se basa en el canal dileptónico con leptones de distintas generaciones ($e\mu$); esta producción implica por la ley de conservación de la carga que ambos deben tener signos opuestos.
- (ii) Cumple los requisitos de **filtro de $\cancel{E}_T$** : no todos los sucesos con alto valor de $\cancel{E}_T$ indican procesos interesantes, como la producción de partículas invisibles aquí estudiada, sino que pueden provenir de efectos de *ruido de fondo*, rayos cósmicos, etc. Por ello, se desarrolla un criterio mediante algoritmos para discriminar este tipo de señales de las que realmente interesa estudiar. Por tanto, se requiere que cada suceso cumpla los requisitos impuestos por el algoritmo.
- (iii) Cumple los requisitos de **trigger**.

Las variables empleadas en la discriminación son:

- **$\cancel{E}_T$:** esta variable representa la actividad invisible de cada suceso, y como se ha comentado anteriormente, es la principal variable de estudio en procesos con partículas invisibles; si bien es conocido que se necesita emplear otras variables para fijar de mejor forma una región de señal óptima. La distribución de esta variable se muestra en la gráfica 4.7.
- **Número de *jets*:** esta variable se emplea en la discriminación para las señales aquí estudiadas debido a la presencia de *quarks top*; este hecho implica la presencia de al menos dos *jets* provenientes de dichos *quarks*, lo que permitirá discriminación respecto a fondos de procesos con menor número de *jets*. La distribución de esta variable se muestra en la gráfica 4.8.

- **Número de *jets* provenientes de *quarks* b (*b-jets*):** esta variable se emplea en la discriminación para las señales aquí estudiadas debido a la presencia de *quarks top*; como ya fue comentado, la presencia de este tipo de *quarks* lleva a su desintegración en *quarks bottom*, cuyos *jets* son característicos y posibles de determinar. La distribución de esta variable se muestra en la gráfica 4.9.
- **$\Delta\phi$ ([28]):** esta variable representa la dirección azimutal entre el vector de *momento transverso faltante* ($\vec{E}_T$) y el vector opuesto a la suma de la actividad hadrónica transversa del evento. Este observable se emplea debido a que, en la producción de dos *jets* con alto momento transverso, estos se producen en direcciones opuestas en el plano transverso y, por tanto, la variable toma el valor π ; por tanto, una desviación de este valor implica la presencia de partículas de baja energía no reconstruidas. La distribución de esta variable se muestra en la gráfica 4.10.
- **$\cos(\theta_b)$ ([29]):** representa el coseno entre el ángulo que forma uno cualquiera de los dos leptones con el *boost* en el sistema de centro de masas de ambos leptones. Esta variable es sensible al *spin* del par producido, y por tanto ofrece discriminación frente a fondos dibosónicos. Su distribución se muestra en la figura 4.11.
- **m_{T2}^l :** variable cinemática que acota la masa de partículas invisibles suponiendo que se desintegran de forma semi-invisible:

$$m_{T2}(p_T^1, p_T^2, q_T) = \min_{q_T^1 + q_T^2 = q_T} \left\{ \max[m_T(p_T^1, q_T^1), m_T(p_T^2, q_T^2)] \right\}$$

donde m_T representa la masa transversa y p_T^i el momento transverso de dos partículas (suponiendo que no tienen masa). En el caso estudiado en este trabajo se toma como p_T^1, p_T^2 el momento transverso de los dos leptones de cada suceso y como q_T al *momento transverso faltante*; se denota entonces m_{T2}^l . En los procesos $t\bar{t}$, esta variable está acotada superiormente por la masa del bosón W, lo que permitirá discriminar de forma importante frente al principal fondo estudiado. La distribución de esta variable se muestra en la figura 4.12.

- **$E_{T,\min}$:** variable que representa el mínimo valor de $E^2 - p_z^2$ de las cuatro posibles combinaciones de *b-jets* y leptones correspondientes a eventos con dos leptones y dos *b-jets*. Esta variable puede interpretarse en el ámbito de la reconstrucción de la masa del *quark top*, donde:

$$m_t^2 = E^2 - p_T^2 - p_z^2,$$

donde m_t representa la masa del *quark top*, E la energía total del sistema tras la desintegración, p_T el módulo del vector momento transverso total de todas las partículas y p_z la componente en el eje z del momento de todas las partículas. Desarrollando esta ecuación en su parte de **partículas visibles** y **partículas invisibles**, se obtiene la relación:

$$E_{T,min} + (E_z^\nu)^2 \leq m_t^2,$$

donde E_z^ν representa la componente z de la energía total del sistema debida a las partículas invisibles. Esto muestra que esta variable está relacionada con la masa del *quark top*.

En el caso de sucesos con dos leptones y menos de dos *b-jets*, se toman los *jets* con mayor momento transverso. Se muestra la distribución de esta variable en la figura 4.13.

- **m_{eff}** : variable que representa la suma escalar de $\cancel{E}_T$, el momento transverso de los dos leptones y el momento transverso de los dos *jets* con más momento del suceso. Esta variable representa una medida de la energía total del sistema, y por tanto es sensible a la producción de partículas invisibles masivas en valores altos. Su distribución se representa en la figura 4.14.

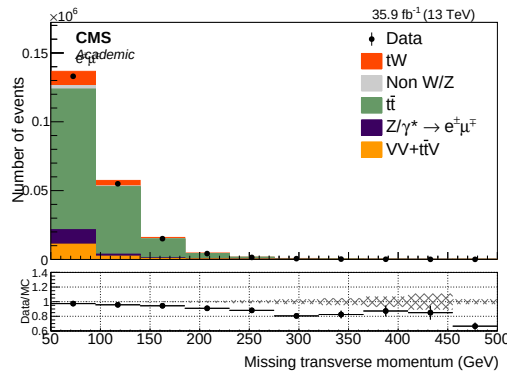
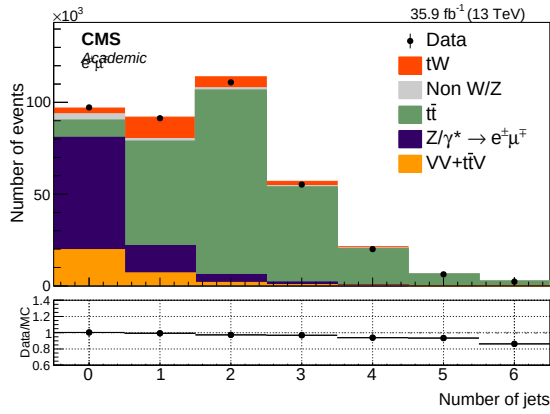
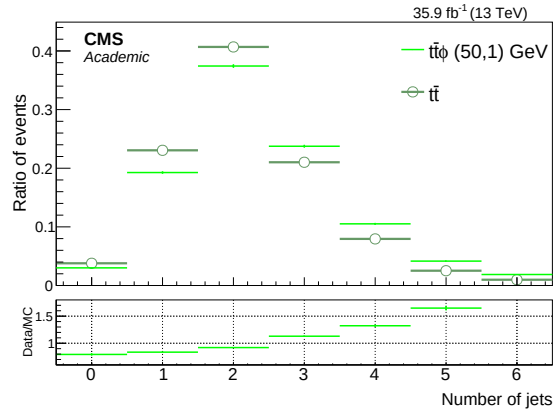


Figura 4.7: Representación gráfica de la variable $\cancel{E}_T$ con los datos y muestras de MC empleadas.

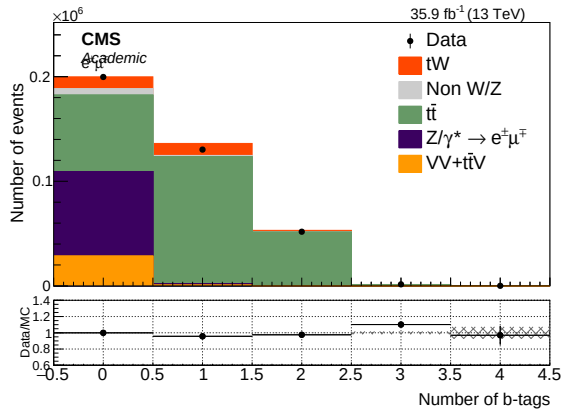


(a) Distribución del número de *jets* en datos, junto con la predicción del SM, dada por las simulaciones de Monte Carlo.

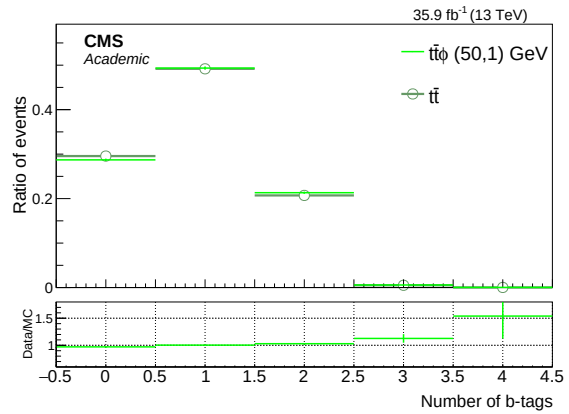


(b) Distribución del número de *jets* para $t\bar{t}$ y la señal con $m_\varphi = 50$ GeV.

Figura 4.8: Representaciones gráficas de la variable número de *jets* con los datos y muestras de MC empleadas.

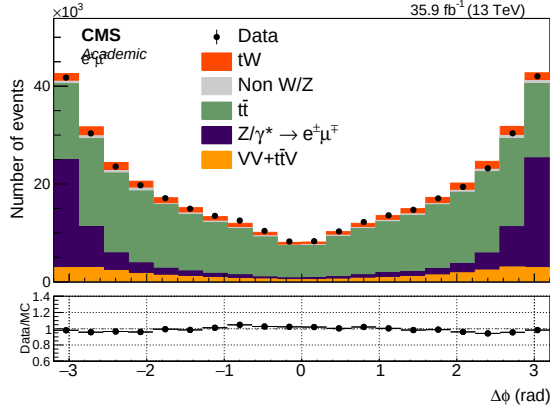


(a) Distribución del número de *b-jets* en datos, junto con la predicción del SM, dada por las simulaciones de Monte Carlo.

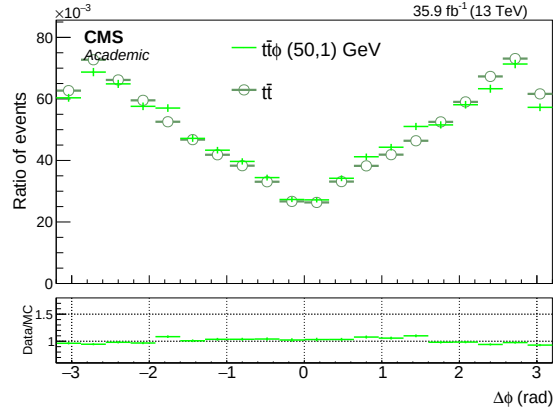


(b) Distribución del número de *b-jets* para $t\bar{t}$ y la señal con $m_\varphi = 50$ GeV.

Figura 4.9: Representaciones gráficas de la variable número de *b-jets* con los datos y muestras de MC empleadas.

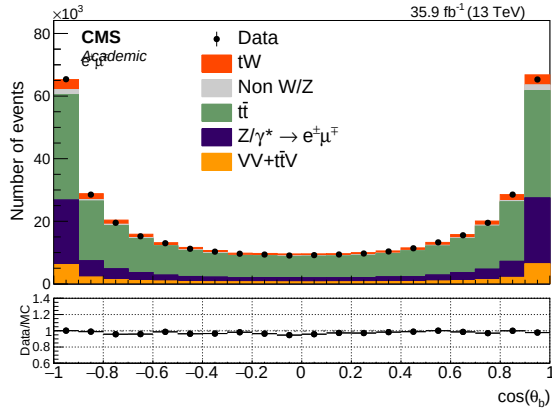


(a) Distribución de la variable $\Delta\phi$ en datos, junto con la predicción del SM, dada por las simulaciones de Monte Carlo.

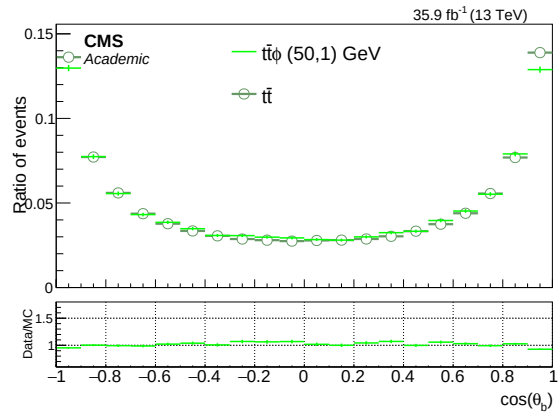


(b) Distribución de la variable $\Delta\phi$ para $t\bar{t}$ y la señal con $m_\varphi = 50$ GeV.

Figura 4.10: Representaciones gráficas de la variable $\Delta\phi$ con los datos y muestras de MC empleadas.

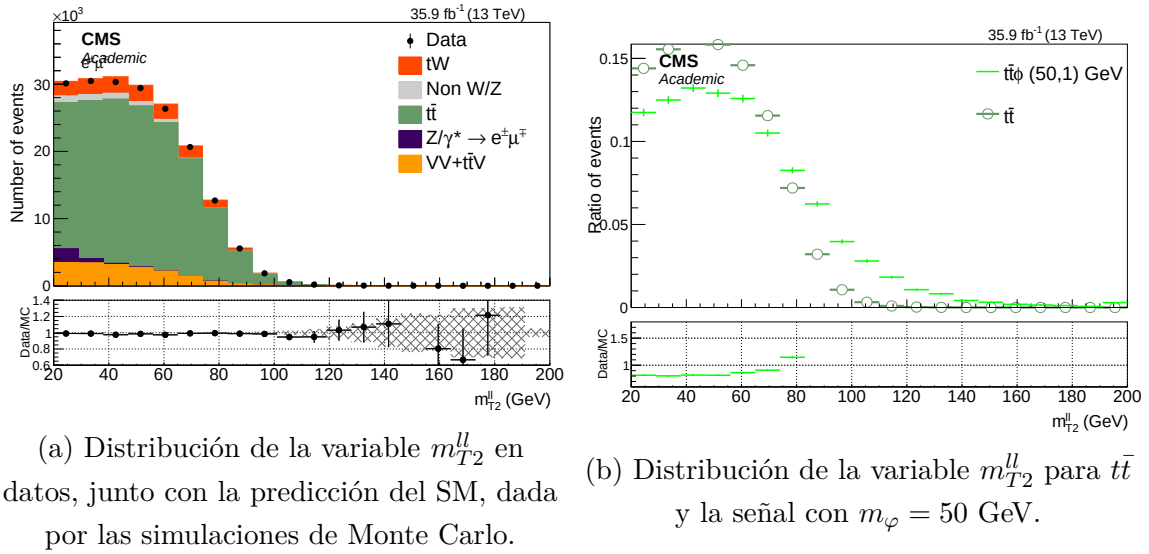


(a) Distribución de la variable $\cos(\theta_b)$ en datos, junto con la predicción del SM, dada por las simulaciones de Monte Carlo.



(b) Distribución de la variable $\cos(\theta_b)$ para $t\bar{t}$ y la señal con $m_\varphi = 50$ GeV.

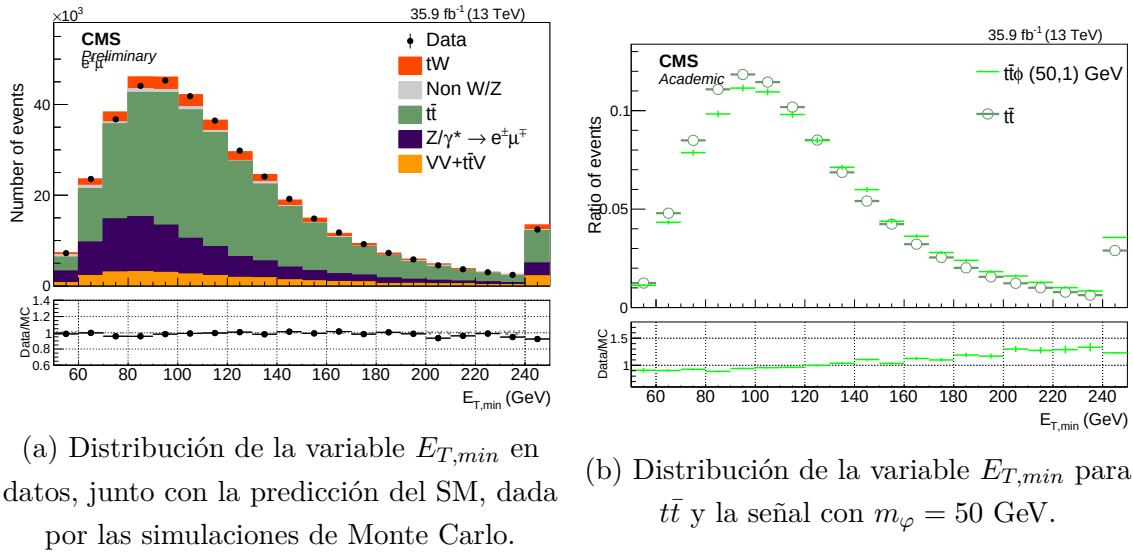
Figura 4.11: Representaciones gráficas de la variable $\cos(\theta_b)$ con los datos y muestras de MC empleadas.



(a) Distribución de la variable m_{T2}^l en datos, junto con la predicción del SM, dada por las simulaciones de Monte Carlo.

(b) Distribución de la variable m_{T2}^l para $t\bar{t}$ y la señal con $m_\phi = 50$ GeV.

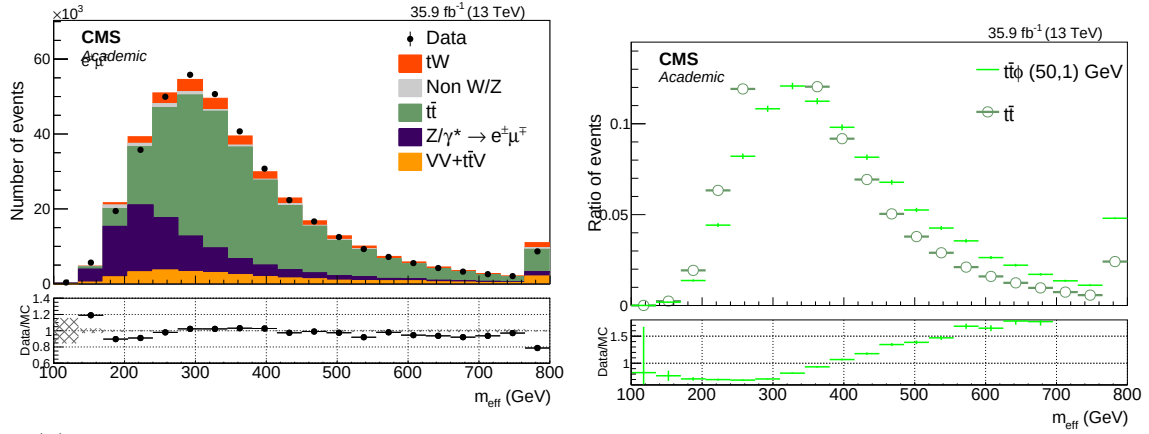
Figura 4.12: Representaciones gráficas de la variable m_{T2}^l con los datos y muestras de MC empleadas.



(a) Distribución de la variable $E_{T,min}$ en datos, junto con la predicción del SM, dada por las simulaciones de Monte Carlo.

(b) Distribución de la variable $E_{T,min}$ para $t\bar{t}$ y la señal con $m_\phi = 50$ GeV.

Figura 4.13: Representaciones gráficas de la variable $E_{T,min}$ con los datos y muestras de MC empleadas.



(a) Distribución de la variable m_{eff} en datos, junto con la predicción del SM, dada por las simulaciones de Monte Carlo.

(b) Distribución de la variable m_{eff} para $t\bar{t}$ y la señal con $m_\phi = 50$ GeV.

Figura 4.14: Distribución de la variable m_{eff} con los datos y muestras de MC empleadas.

4.3.1. *Blinding*

En la sección anterior se han mostrado gráficas en las que se incluyen datos y simulaciones de MC para las muestras contempladas en este trabajo, con la finalidad de mostrar la correcta modelización de dichas variables de discriminación. Sin embargo, en los estudios físicos realizados con colisionadores de partículas, la inclusión de los datos tomados se realiza tras haber desarrollado un análisis considerado óptimo, con el objetivo de extraer conclusiones. Esto se hace para tener imparcialidad en el desarrollo del análisis; se pretende evitar una posible influencia de los datos en el criterio empleado en el desarrollo del análisis, por ejemplo, mostrando un exceso de datos en una cierta región que lleve a definir un análisis sobre ella. El proceso de 'cegado' a los datos se conoce como *blinding*, y se empleará en el desarrollo de este trabajo. Por tanto, a pesar de haber mostrado datos en las gráficas anteriores, en lo que sigue sólo se trabajará con las simulaciones de MC hasta la obtención de un análisis considerado óptimo; a continuación, se incluirán los datos y se procederá a extraer conclusiones.

4.4. Selección de eventos

En esta sección del trabajo se procederá a la definición de la región de señal de estudio que permita maximizar las señales consideradas respecto a los fondos

mostrados. Esta selección estará basada en diversos criterios de selección sobre las variables introducidas en la sección 4.3, que permitirán definir una región del espacio de fases conveniente para el análisis realizado.

La aplicación de este tipo de criterios de selección se realiza de forma *secuencial*: tras observar la distribución de las variables consideradas, se define un criterio de selección sobre una de ellas en una zona que maximice señal frente a fondo, y se estudia la influencia de dicho criterio sobre el resto de variables. Este proceso se continúa hasta aplicar criterios sobre todas las variables consideradas, definiendo así la región de señal.

En primer lugar, a la vista de las gráficas mostradas en la sección anterior, se comenzará con un criterio de selección sobre el número de *jets* del suceso. Atendiendo a la gráfica 4.8b, se fijará $N^{\circ} \text{ *Jets*} \geq 2$. En las figuras 4.15, 4.16 (por cuestiones de espacio, se muestran en figuras separadas) se muestra el efecto de esta selección sobre el resto de variables. Como se puede observar, el comportamiento de dichas variables es similar al anterior, pero se reduce sustancialmente el número de sucesos dibosónicos y de procesos Drell-Yan.

En segundo lugar, tomando como referencia la figura 4.16, el siguiente criterio se aplicará sobre el número de *b-jets* en el suceso; se fijará $N^{\circ} \text{ *b-jets*} \geq 1$ como se puede observar en la gráfica 4.16a. La distribución del resto de variables tras esta selección se muestran en la figura 4.17. Se observa una mayor discriminación frente a eventos dibosónicos y procesos Drell-Yan.

A continuación se aplicarán criterios de selección sobre las variables angulares, ya que se observa en la figura 4.17 que la diferencia entre señal y fondo no es muy elevada, y por tanto no inducen un criterio demasiado estricto, pero aun así permiten reducir el fondo estudiado. El primer criterio se aplicará sobre la variable $\Delta\phi$, atendiendo a su distribución en la gráfica 4.17a. Por tanto, se tomará $\Delta\phi < 1 \text{ rad}$, ya que se observa que el centro de la distribución define un buen criterio de selección para maximizar la señal. Después, se realizará el mismo criterio para la variable $\cos(\theta_b)$; esta vez tomando $\cos(\theta_b) < 0.3$, debido a su distribución en la figura 4.18a.

Las distribuciones de las variables correspondientes a la selección en la variable $\Delta\phi$ se muestran en la figura 4.18, mientras que las correspondientes a la selección en $\cos(\theta_b)$ se muestran en la figura 4.19. Se observa que el comportamiento del resto de variables cinemáticas es similar al anterior, habiendo reducido el fondo de procesos Drell-Yan y dibosónicos considerablemente.

El siguiente criterio de selección se tomará sobre la variable $m_{T2}^{\ell\ell}$; como se puede observar en la figura 4.19a, esta variable ofrece un alto valor de discriminación en valores mayores a la masa del bosón W frente a los procesos $t\bar{t}$, que constituyen el fondo principal del proceso. Se han realizado distintos criterios de selección sobre esta variable, y se ha determinado el caso más favorable a la discriminación de los fondos considerados; por ello, se tomará $\mathbf{m}_{T2}^{\ell\ell} > 100 \text{ GeV}$. La distribución del resto de variables tras la aplicación de este criterio se muestra en la figura 4.20.

Tras la aplicación del criterio de selección sobre la variable $m_{T2}^{\ell\ell}$, se observa un claro decrecimiento en el número de sucesos esperados del proceso $t\bar{t}$, como era de esperar por la definición de esta variable. El hecho de tomar un criterio más estricto reduciría en mayor medida estos procesos, pero debe hacerse de forma que el número de sucesos final con que se trabaje en el análisis sea suficiente para tomar conclusiones; un número demasiado bajo puede hacer que la incertidumbre estadística asociada a las medidas impida extraer conclusiones sobre el modelo estudiado.

El penúltimo criterio de selección se aplicará sobre la variable $E_{T,min}$, tomando como referencia su distribución en la figura 4.20a. Por ello, el criterio empleado será $\mathbf{E}_{T,min} > 110 \text{ GeV}$. La distribución de las variables m_{eff} y $\cancel{E}_T$ tras su aplicación se muestran en la figura 4.21.

A la vista de la figura 4.21a, el último criterio de selección tomado será $\mathbf{m}_{eff} > 450 \text{ GeV}$, lo cual es compatible con dicha figura; la distribución de $\cancel{E}_T$ y la comparación de los histogramas normalizados de señal y fondo $t\bar{t}$ para dicha variable se muestran en la figura 4.22.

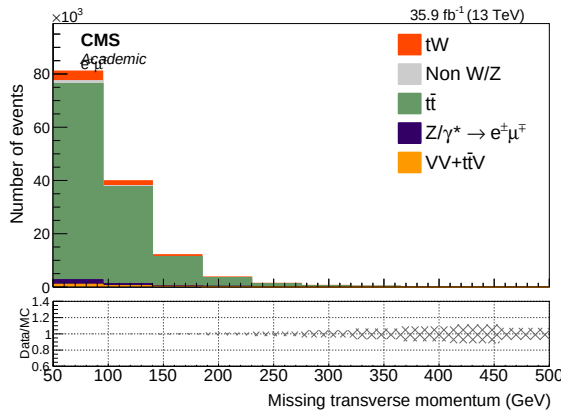
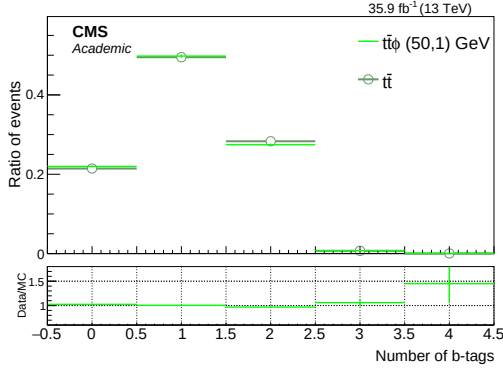
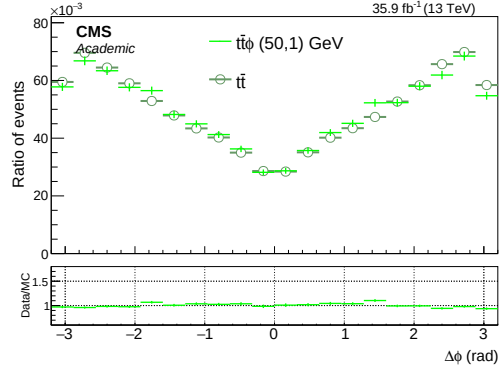


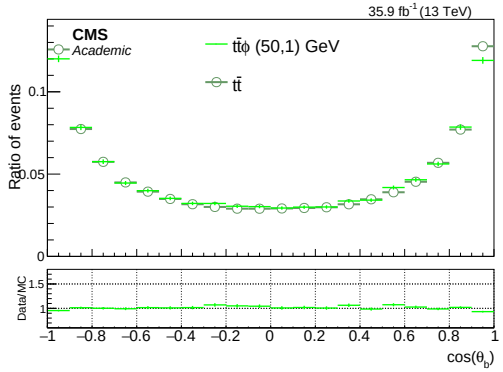
Figura 4.15: Distribución de la variable $\cancel{E}_T$, junto con la predicción del SM, dada por las simulaciones de Monte Carlo con la selección: $\mathbf{N}^0 \text{ Jets} \geq 2$



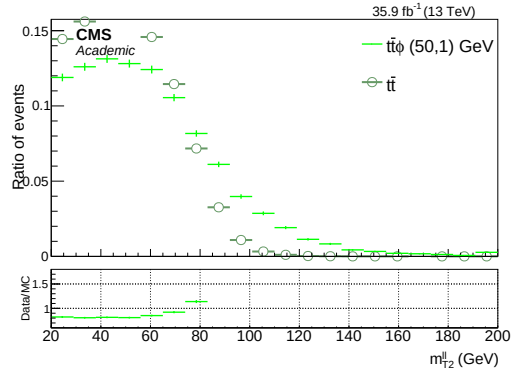
(a) Distribución del número de b -jets para $t\bar{t}$ y la señal con $m_\varphi = 50$ GeV.



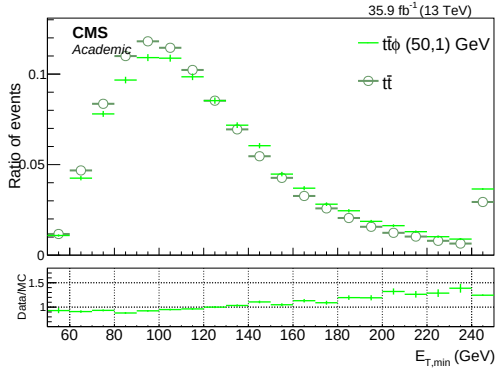
(b) Distribución de la variable $\Delta\phi$ para $t\bar{t}$ y la señal con $m_\varphi = 50$ GeV.



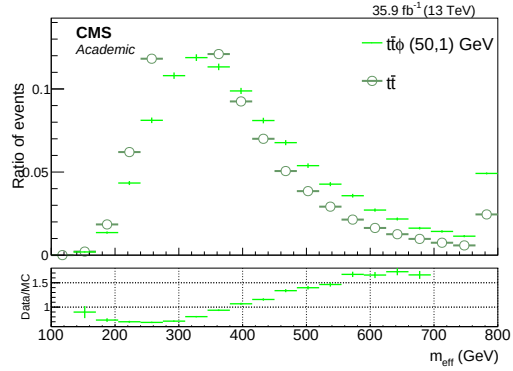
(c) Distribución de la variable $\cos(\theta_b)$ para $t\bar{t}$ y la señal con $m_\varphi = 50$ GeV.



(d) Distribución de la variable m_{T2}^{ll} para $t\bar{t}$ y la señal con $m_\varphi = 50$ GeV.

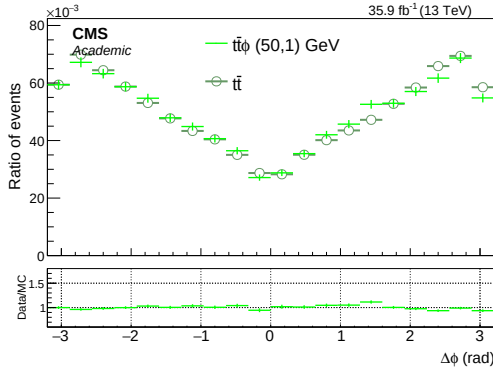


(e) Distribución de la variable $E_{T,min}$ para $t\bar{t}$ y la señal con $m_\varphi = 50$ GeV.

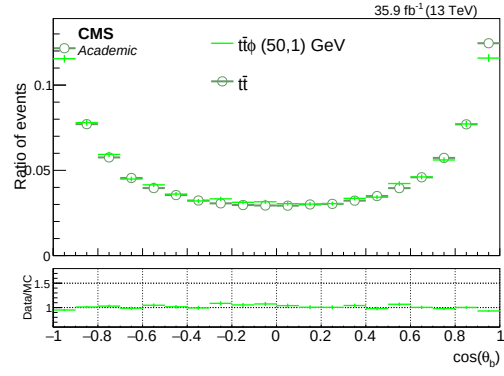


(f) Distribución de la variable m_{eff} para $t\bar{t}$ y la señal con $m_\varphi = 50$ GeV.

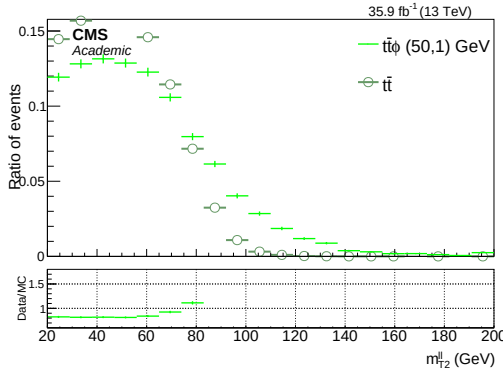
Figura 4.16: Representaciones gráficas de las variables de discriminación empleadas con la selección: $N^{\text{ Jets}} \geq 2$.



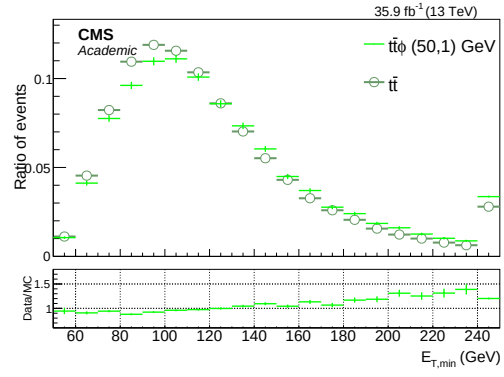
(a) Distribución de la variable $\Delta\phi$ para $t\bar{t}$ y la señal con $m_\varphi = 50$ GeV.



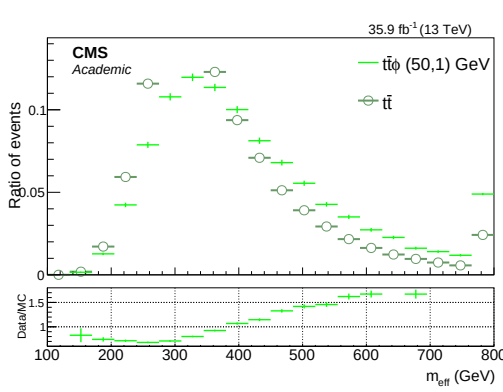
(b) Distribución de la variable $\cos(\theta_b)$ para $t\bar{t}$ y la señal con $m_\varphi = 50$ GeV.



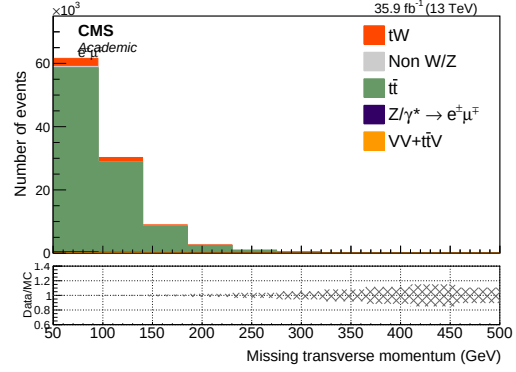
(c) Distribución de la variable m_{T2}^{ll} para $t\bar{t}$ y la señal con $m_\varphi = 50$ GeV.



(d) Distribución de la variable $E_{T,min}$ para $t\bar{t}$ y la señal con $m_\varphi = 50$ GeV.

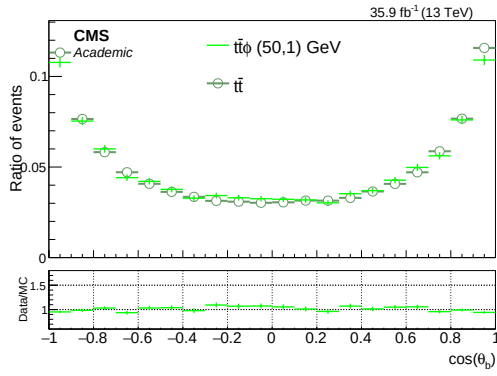


(e) Distribución de la variable m_{eff} para $t\bar{t}$ y la señal con $m_\varphi = 50$ GeV.

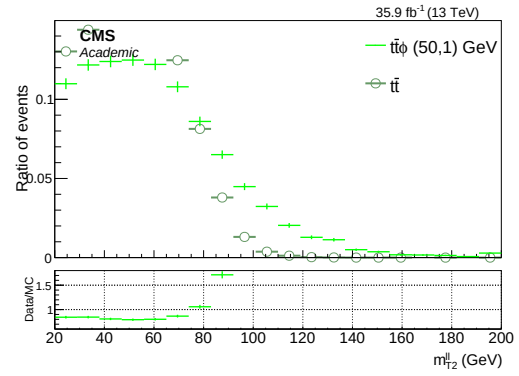


(f) Distribución de la variable E_T , junto con la predicción del SM, dada por las simulaciones de Monte Carlo.

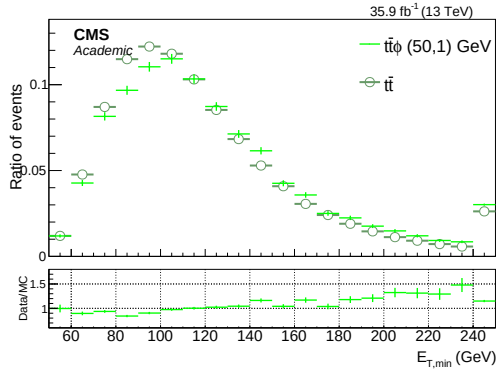
Figura 4.17: Representaciones gráficas de las variables de discriminación empleadas con la selección: $N^o \text{ Jets} \geq 2$, $N^o \text{ b-jets} \geq 1$.



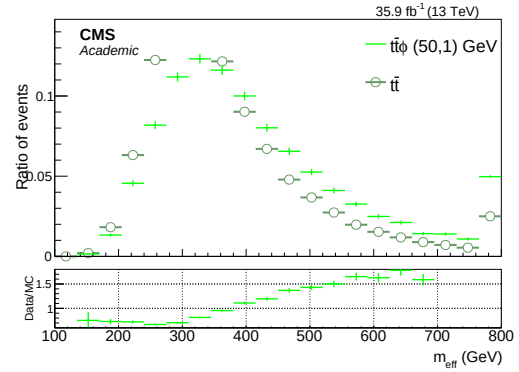
(a) Distribución de la variable $\cos(\theta_b)$ para $t\bar{t}$ y la señal con $m_\varphi = 50$ GeV.



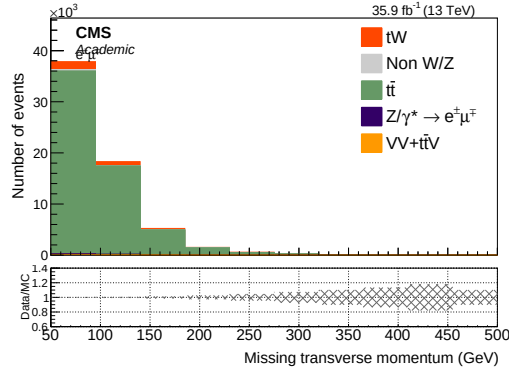
(b) Distribución de la variable m_{T2}^{ll} para $t\bar{t}$ y la señal con $m_\varphi = 50$ GeV.



(c) Distribución de la variable $E_{T,min}$ para $t\bar{t}$ y la señal con $m_\varphi = 50$ GeV.



(d) Distribución de la variable m_{eff} para $t\bar{t}$ y la señal con $m_\varphi = 50$ GeV.



(e) Distribución de la variable $\cancel{E}_T$, junto con la predicción del SM, dada por las simulaciones de Monte Carlo.

Figura 4.18: Representaciones gráficas de las variables de discriminación empleadas con la selección: $N^{\circ} \text{ Jets} \geq 2$, $N^{\circ} \text{ b-jets} \geq 1$, $\Delta\phi < 1$ rad.

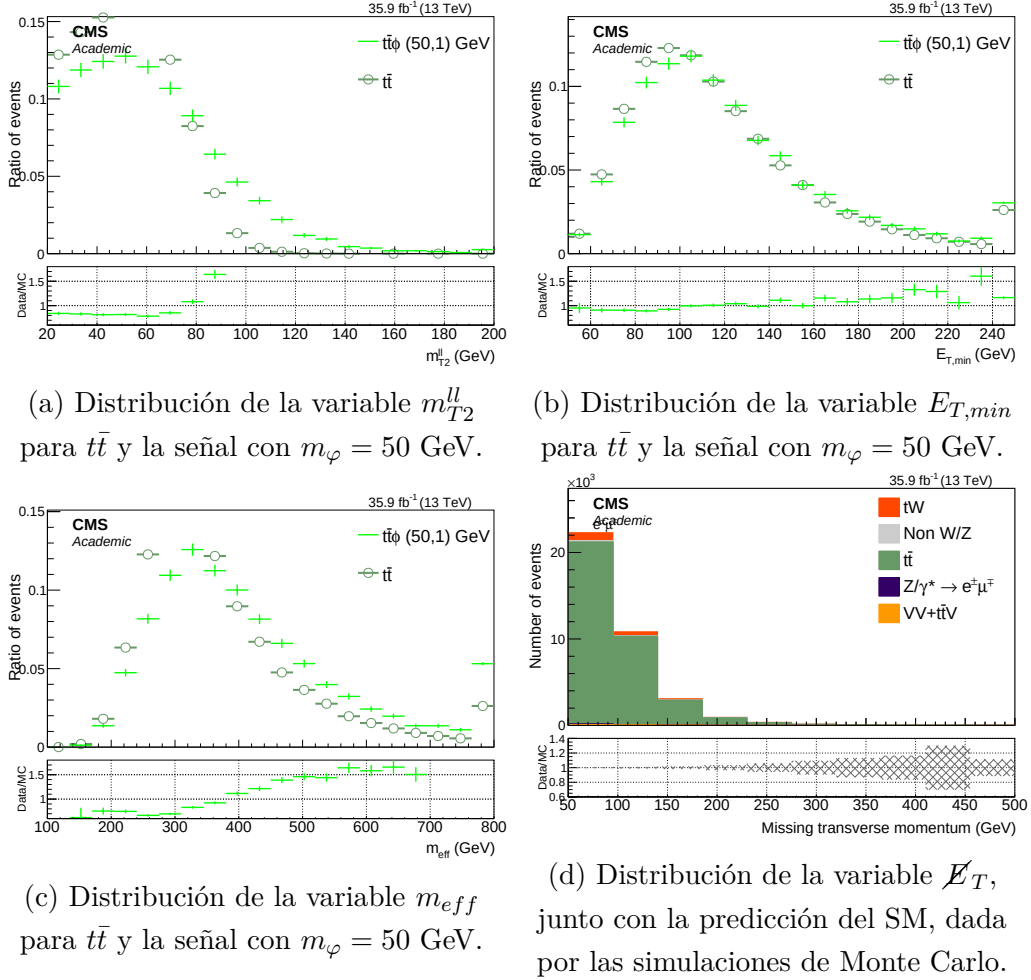
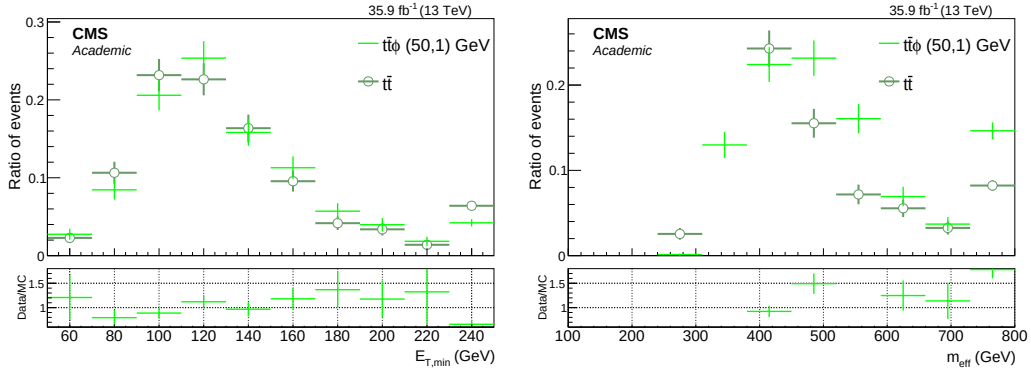
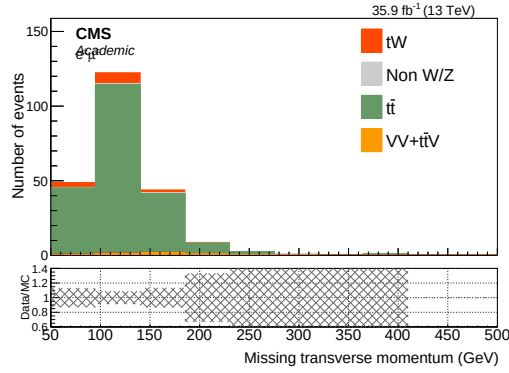


Figura 4.19: Representaciones gráficas de las variables de discriminación empleadas con la selección: $N^\circ \text{ Jets} \geq 2$, $N^\circ \text{ b-jets} \geq 1$, $\Delta\phi < 1 \text{ rad}$, $\cos(\theta_b) < 0.3$.



(a) Distribución de la variable $E_{T,min}$ para tt y la señal con $m_\varphi = 50$ GeV.

(b) Distribución de la variable m_{eff} para tt y la señal con $m_\varphi = 50$ GeV.



(c) Distribución de la variable E_T , junto con la predicción del SM, dada por las simulaciones de Monte Carlo.

Figura 4.20: Representaciones gráficas de las variables de discriminación empleadas con la selección: $N^\circ \text{ Jets} \geq 2$, $N^\circ b\text{-jets} \geq 1$, $\Delta\phi < 1$ rad, $\cos(\theta_b) < 0.3$, $m_{T2}^{\text{ll}} > 100$ GeV.

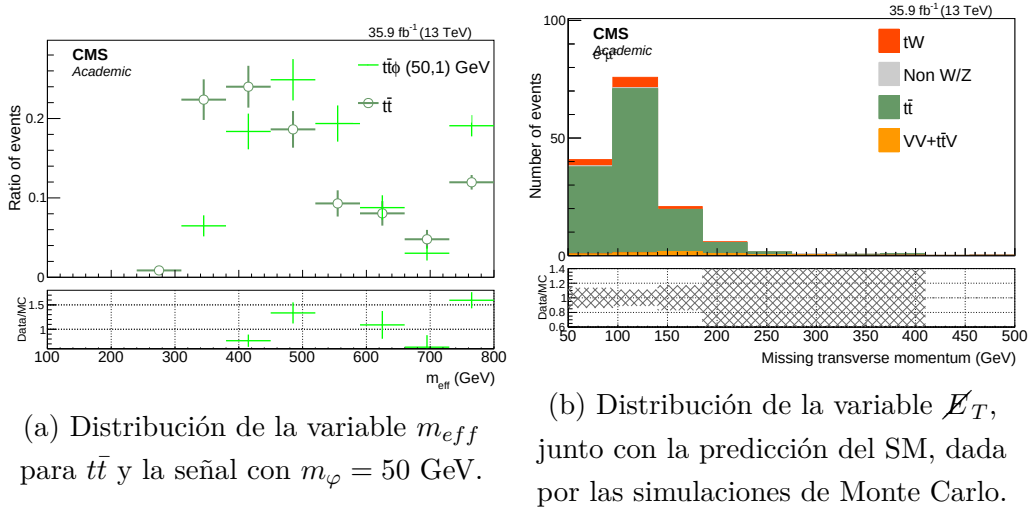


Figura 4.21: Representaciones gráficas de las variables de discriminación empleadas con la selección: $N^o \text{ Jets} \geq 2$, $N^o \text{ } b\text{-jets} \geq 1$, $\Delta\phi < 1 \text{ rad}$, $\cos(\theta_b) < 0.3$, $m_{T2}^{\text{ll}} > 100 \text{ GeV}$, $E_{T,\text{min}} > 110 \text{ GeV}$.

Una vez realizados estos criterios sobre las variables, debe considerarse si el fondo de *non-prompt leptons* es considerable a este nivel; en los pasos anteriores no es necesario estudiarlo debido a la baja cantidad de sucesos de este tipo. Para ello, se muestra en la figura 4.23 el número de leptones para las señales extremas ($m_\phi = 5, 50 \text{ GeV}$) y la distribución de la variable $\cancel{E}_T$ en el caso equivalente a la figura 4.22a y su distribución exigiendo **exactamente** dos leptones. Como se puede observar, el fondo debido a estos procesos no es un factor importante a este nivel.

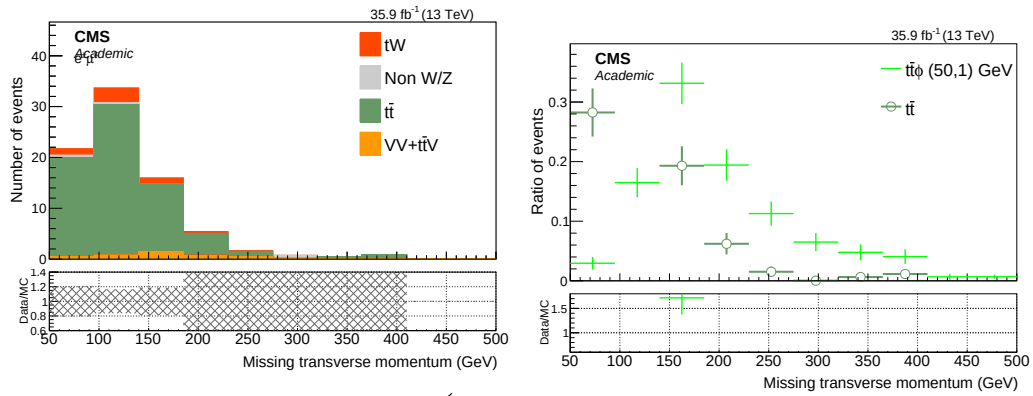
Por otro lado, la baja cantidad de sucesos es lógica debido a la baja sección eficaz del proceso, pero se espera que una de las principales fuentes de incertidumbre provenga de incertidumbres estadísticas debidas precisamente al bajo número de sucesos de MC; la forma de mejorar estas incertidumbres pasa por aumentar el orden de las simulaciones de MC. Como ya se comentó anteriormente, el aumento de simulaciones de MC es deseable para eliminar este tipo de incertidumbres para señales tan bajas, y por eso es necesario el desarrollo de nuevos algoritmos que permitan calcular los diagramas de procesos a órdenes más altos.

Como se puede observar en la figura 4.22b, los fondos se han reducido sustancialmente en comparación a la señal, y se muestra una predominación de la señal en la zona con $\cancel{E}_T > 150 \text{ GeV}$; por tanto, se muestra en la figura 4.24 la distribución de esta variable en dicha zona del espacio de fases incluyendo las cuatro señales estudiadas en este trabajo, donde claramente se observa un aumento de los sucesos de señal esperados frente a los fondos considerados.

Se recogen en la tabla 4.2 los criterios de selección que definen la región de señal empleada en este análisis.

Variable cinemática	Criterio de selección
$N^{\circ} \text{ jets}$	≥ 2
$N^{\circ} b\text{-jets}$	≥ 1
$\Delta\phi$	$< 1 \text{ rad}$
$\cos(\theta_b)$	< 0.3
$m_{T2}^{\ell\ell}$	$> 100 \text{ GeV}$
$E_{T,min}$	$> 110 \text{ GeV}$
m_{eff}	$> 450 \text{ GeV}$

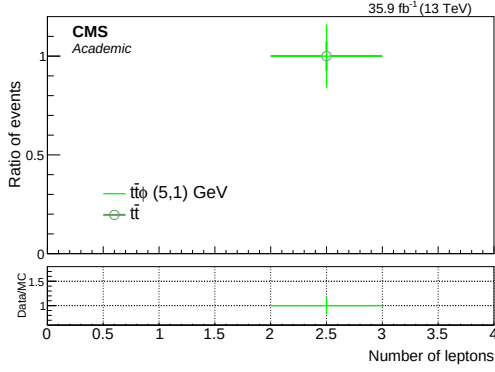
Tabla 4.2: Región de señal correspondiente al análisis realizado.



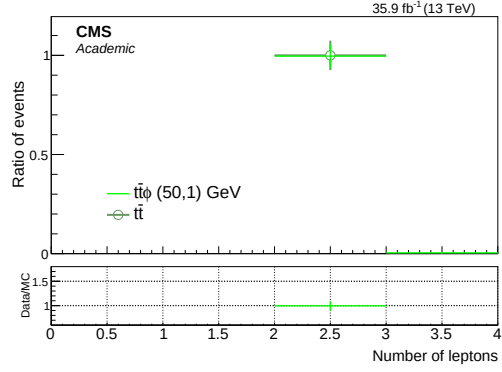
(a) Distribución de la variable $\cancel{E}_T$, junto con la predicción del SM, dada por las simulaciones de Monte Carlo.

(b) Distribución de la variable $\cancel{E}_T$ para $t\bar{t}$ y la señal con $m_{\varphi} = 50 \text{ GeV}$.

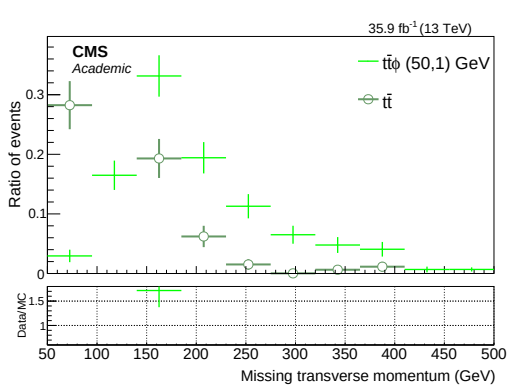
Figura 4.22: Representaciones gráficas de las variables de discriminación empleadas con la selección: $N^{\circ} \text{ Jets} \geq 2$, $N^{\circ} b\text{-jets} \geq 1$, $\Delta\phi < 1 \text{ rad}$, $\cos(\theta_b) < 0.3$, $m_{T2}^{\ell\ell} > 100 \text{ GeV}$, $E_{T,min} > 110 \text{ GeV}$, $m_{eff} > 450 \text{ GeV}$.



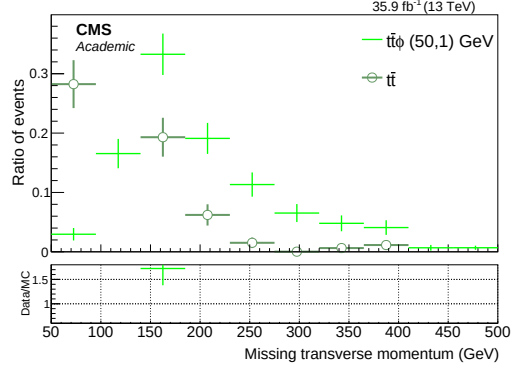
(a) Distribución del número de leptones detectados para $t\bar{t}$ y la señal con $m_\phi = 5$ GeV.



(b) Distribución del número de leptones detectados para $t\bar{t}$ y la señal con $m_\phi = 50$ GeV.



(c) Distribución de la variable $\cancel{E}_T$ para $t\bar{t}$ y la señal con $m_\phi = 50$ GeV.



(d) Distribución de la variable $\cancel{E}_T$ para $t\bar{t}$ y la señal con $m_\phi = 50$ GeV exigiendo exactamente dos leptones.

Figura 4.23: Influencia del fondo de *non-prompt leptons* sobre la distribución de $\cancel{E}_T$.

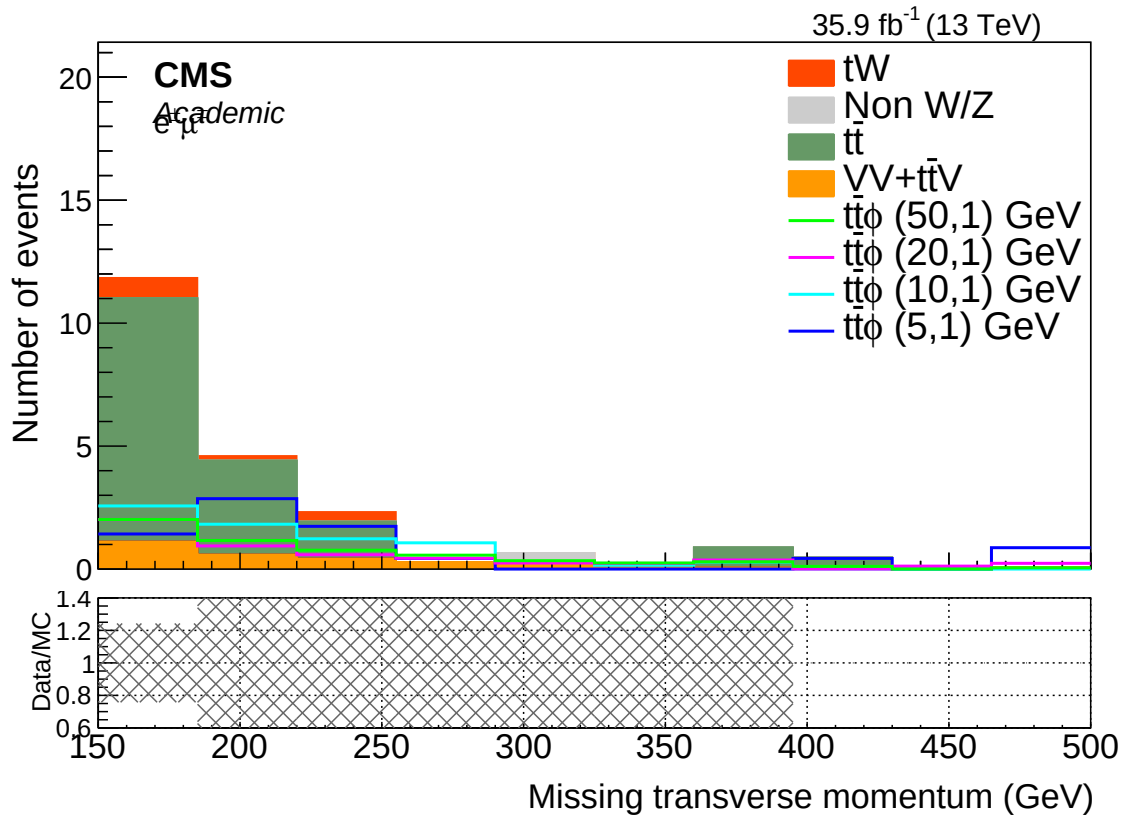


Figura 4.24: Distribución de la variable $\cancel{E}_T$, junto con la predicción del SM dada por las simulaciones de Monte Carlo y con las cuatro señales propuestas.

Capítulo 5

Resultados obtenidos

En este último capítulo se expondrán los resultados obtenidos mediante el análisis expuesto en el capítulo anterior. En primer lugar, se definirá el *upper-limit* o límite superior de un modelo experimental y el método de su obtención utilizando herramientas de CMS. A continuación, se estudiarán las principales fuentes de incertidumbre que afectan al análisis realizado, y por último se tratará de obtener un análisis más estricto que permita reducir las incertidumbres asociadas.

5.1. Extracción de señal. Obtención del límite superior (UL) del modelo

En este apartado del capítulo se procederá a la obtención del resultado final de este trabajo. Se realizará un estudio sobre el análisis expuesto anteriormente que permita obtener conclusiones sobre la veracidad del modelo. Para ello, se procederá a introducir conceptos como la *signal-strength* o fuerza de señal, el *upper-limit* de un modelo y los gráficos de exclusión.

En la mayoría de procesos estudiados en el LHC, las secciones eficaces de los procesos son uno de los principales observables que se pueden medir. Sin embargo, la obtención directa de esta magnitud en ciertos procesos puede ser complicada; por eso, normalmente se recurre a hacer medidas sobre una magnitud denominada *signal-strength* o **fuerza de señal**, definida como:

$$\mu := \frac{\sigma_{exp}}{\sigma_{th}},$$

donde σ_{exp} representa la sección eficaz experimental del proceso y σ_{th} su sección eficaz predicha por el SM.

Debido a la naturaleza cuántica de los procesos en física de partículas, el estudio de los datos requiere un tratamiento probabilístico. Más concretamente, el estudio de la adecuación de los datos tomados a las muestras simuladas por MC se basa en la realización de un *test de hipótesis* sobre la igualdad de sus distribuciones. Esto conlleva, en la práctica, una realización del siguiente contraste de hipótesis:

- H_0 : Física del modelo estándar (SM)
- H_1 : Física más allá del modelo estándar (BSM)

El desarrollo de este estudio es complejo y no se analizará con profundidad en este trabajo, pero se introducirán las ideas que lo fomentan. En primer lugar, se construye una *función de verosimilitudes* de la distribución de la variable por clases. La diferencia de esta función de verosimilitudes con la empleada en la inferencia estadística clásica es que existen **incertidumbres** debidas a la determinación de los parámetros (momento transversal, energía de los *jets*, etc.). Por tanto, esta función de verosimilitud incluye un término de información *a priori* (estadística *bayesiana*; por ejemplo, que la energía de los jets se distribuye en torno a cero como una gaussiana) y otro término en que se introduce un parámetro a fondo y señal denotado como $\vec{\theta}$. Este parámetro recibe el nombre de *nuisance* o parámetro de 'confusión', y es el responsable de incluir las incertidumbres en medidas en esta verosimilitud.

El objetivo del contraste es obtener un *intervalo de confianza* a un cierto nivel de significación para la sección eficaz (o μ) del proceso. Dicho intervalo de confianza será de la forma $[0, \tilde{\mu}]$, como es de esperar; el valor $\tilde{\mu}$ se conoce como el **límite superior para la fuerza de señal** del modelo. Para obtenerlo, y basándose en el estadístico clásico de *razón de verosimilitudes* (*Likelihood ratio*), se define un estadístico llamado **Profile Likelihood ratio** definido como:

$$q(\mu) := \frac{L(\mathbf{x}|\mu, \tilde{\theta}(\mu))}{L(\mathbf{x}|\tilde{\mu}, \tilde{\theta}(\tilde{\mu}))}, \quad (5.1)$$

donde L representa la función de verosimilitud, $\tilde{\mu}$ la *estimación máximo verosímil* del parámetro μ y $\tilde{\theta}(\mu)$ la *estimación máximo verosímil* del parámetro θ asociada al valor μ .

Este análisis, conocido como **extracción de señal**, debe llevarse a cabo para todos los sucesos y teniendo en cuenta todas las incertidumbres asociadas. Por ello, existen diversos algoritmos que permiten obtener el valor del límite superior a un cierto valor de significación. En el caso expuesto en este trabajo, se emplea un *software* desarrollado por CMS para el análisis del bosón de Higgs llamado **Higgs**

Combine Tool ([30]). Este código permite la obtención del UL correspondiente a cada señal considerada en función de la masa del mediador escalar φ , a un nivel de confianza del 95 %, incluyendo los cuantiles de su distribución correspondientes a 1σ y 2σ .

La forma de representar estos valores en la literatura es mediante un gráfico de exclusión: un gráfico del UL para la fuerza de señal del proceso para cada modelo considerado, incluyendo las incertidumbres a 1σ y 2σ . En el caso que se estudia en este trabajo, el gráfico muestra la estimación del límite superior para la fuerza de señal de cada modelo con un nivel de confianza del 95 % en función de la masa del mediador escalar con la materia oscura; se muestra el valor esperado o mediana de la distribución del estadístico anterior (ecuación 5.1) mediante una línea continua. Dicho gráfico se muestra en la figura 5.1, sin incluir los datos observados, por los motivos expuestos en el apartado sobre el *blinding*.

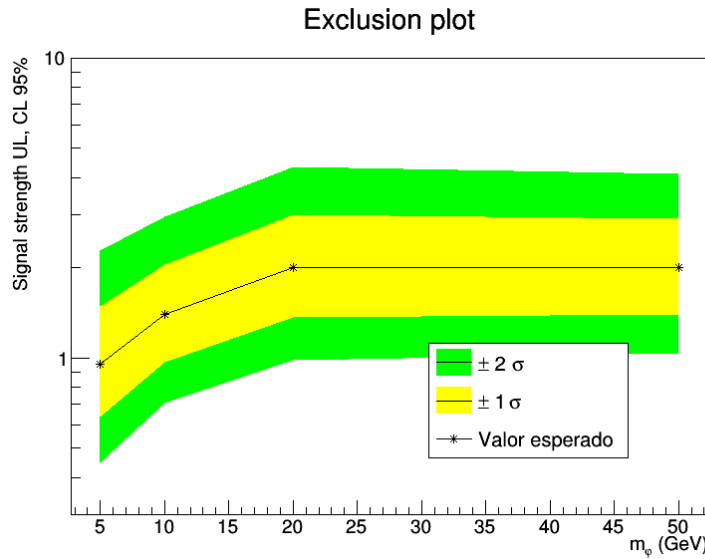


Figura 5.1: Gráfico de exclusión al nivel de confianza 95 % para el UL en la fuerza de señal de la masa del mediador escalar con la materia oscura, φ .

5.2. Incertidumbres sistemáticas

La incertidumbre asociada al estudio del límite superior del modelo depende de diversos factores, entre los que se hallan las incertidumbres sistemáticas asociadas. Por ello, la determinación de las principales fuentes de estas incertidumbres puede aportar un conocimiento más profundo que permita mejorar los análisis realizados.

Las fuentes de estas incertidumbres son variadas, pero en este trabajo se consideran variaciones en una cantidad fija hacia arriba y abajo en los parámetros de confusión; estos variarán el valor de la fuerza de señal y permitirán estudiar la influencia de cada sistemático en la incertidumbre obtenida. Para estudiar la influencia de estas incertidumbres, se empleará un *framework* basado en el uso del programa introducido anteriormente, *Combine*. Este *framework* recibe el nombre de **CombineHarvester** ([31]), y permitirá obtener una estimación en la fuerza de señal de los modelos planteados y estudiar el efecto de cada sistemático en su incertidumbre.

Las incertidumbres sistemáticas introducidas correspondientes a los cuatro modelos estudiados son:

- Escala de energía de los *jets* (JES)
- Btag: eficiencia en el etiquetado de *b-jets*
- Mistag: eficiencia en el valor de el falso etiquetado de *b-jets*
- PU: respecto al valor de *pile-up*
- ElecEff: eficiencia de reconstrucción de electrones
- MuonEff: eficiencia de reconstrucción de muones
- Trig: eficiencia de *trigger*

Las incertidumbres que se asignan a los distintos fondos considerados en el análisis son:

- $t\bar{t}$: 4 %
- tW : 10 %
- *Non W/Z* : 50 %
- $VV + t\bar{t}V$: 50 %

Se recogen en la tabla 5.1 las estimaciones de la fuerza de señal de los cuatro modelos. En las tablas 5.2, 5.3, 5.4 y 5.5 se muestran las fuentes de incertidumbre estudiadas, y la incertidumbre que inducen en la estimación de la fuerza de señal (tomando el valor $\tilde{\mu}_{exp} = 1.0$, dado que se trata del valor esperado). En dichas tablas, sólo se mostrarán las **contribuciones superiores al 1 %**. Se observa que las principales fuentes de incertidumbre son debidas en gran parte a las simulaciones

de Monte Carlo, a causa del bajo número de sucesos con que se trabaja, y otras fuentes como JES y $VV + t\bar{t}V$.

Masa del mediador (GeV)	$\tilde{\mu}_{\text{exp}}$	Incertidumbres sup/inf
5	1.0	+0.6/-0.5
10	1.0	+0.8/-0.7
20	1.0	+1.1/-0.9
50	1.0	+1.0/-0.9

Tabla 5.1: Estimación máximo verosímil de la fuerza de señal **esperada** obtenida para los cuatro modelos considerados.

Fuente de incertidumbre	Incertidumbre en $\tilde{\mu}_{\text{exp}}$	Contribución
MonteCarlo Bin 1	0.16	16 %
MonteCarlo Bin 3	0.16	16 %
JES	0.13	13 %
MonteCarlo Bin 4	0.12	12 %
MonteCarlo Bin 9	0.10	10 %
$VV + t\bar{t}V$	0.10	10 %
Luminosidad	0.09	9 %
$t\bar{t}$	0.08	8 %
MonteCarlo Bin 2	0.07	7 %
MonteCarlo Bin 7	0.06	6 %
MonteCarlo Bin 0	0.03	3 %
$tW, \bar{t}W$	0.02	2 %

Tabla 5.2: Influencia de las incertidumbres sistemáticas en la estimación máximo verosímil de la fuerza de señal **esperada** para el mediador escalar $m_\varphi = 5$ GeV.

Fuente de incertidumbre	Incertidumbre en $\tilde{\mu}_{exp}$	Contribución
MonteCarlo Bin 2	0.20	20 %
MonteCarlo Bin 3	0.18	18 %
$VV + t\bar{t}V$	0.17	17 %
MonteCarlo Bin 3	0.17	17 %
MonteCarlo Bin 5	0.11	11 %
Luminosidad	0.10	10 %
JES	0.10	10 %
$t\bar{t}$	0.09	9 %
MonteCarlo Bin 1	0.09	9 %
MonteCarlo Bin 6	0.07	7 %
MonteCarlo Bin 7	0.05	5 %
Non W/Z	0.03	3 %

Tabla 5.3: Influencia de las incertidumbres sistemáticas en la estimación máximo verosímil de la fuerza de señal **esperada** para el mediador escalar $m_\varphi = 10$ GeV.

Fuente de incertidumbre	Incertidumbre en $\tilde{\mu}_{exp}$	Contribución
MonteCarlo Bin 2	0.24	24 %
$VV + t\bar{t}V$	0.23	23 %
MonteCarlo Bin 3	0.21	21 %
MonteCarlo Bin 6	0.20	20 %
MonteCarlo Bin 1	0.15	15 %
MonteCarlo Bin 7	0.15	15 %
MonteCarlo Bin 4	0.14	14 %
$t\bar{t}$	0.14	14 %
Luminosidad	0.14	14 %
MonteCarlo Bin 5	0.12	12 %
MonteCarlo Bin 9	0.09	9 %
MonteCarlo Bin 0	0.05	5 %

Tabla 5.4: Influencia de las incertidumbres sistemáticas en la estimación máximo verosímil de la fuerza de señal **esperada** para el mediador escalar $m_\varphi = 20$ GeV.

Fuente de incertidumbre	Incertidumbre en $\tilde{\mu}_{exp}$	Contribución
$VV + t\bar{t}V$	0.26	26 %
MonteCarlo Bin 2	0.26	26 %
MonteCarlo Bin 3	0.23	23 %
MonteCarlo Bin 6	0.21	21 %
MonteCarlo Bin 4	0.19	19 %
MonteCarlo Bin 5	0.19	19 %
MonteCarlo Bin 7	0.13	13 %
Luminosidad	0.12	12 %
$t\bar{t}$	0.12	12 %
MonteCarlo Bin 1	0.07	7 %
JES	0.06	6 %
Non W/Z	0.05	5 %

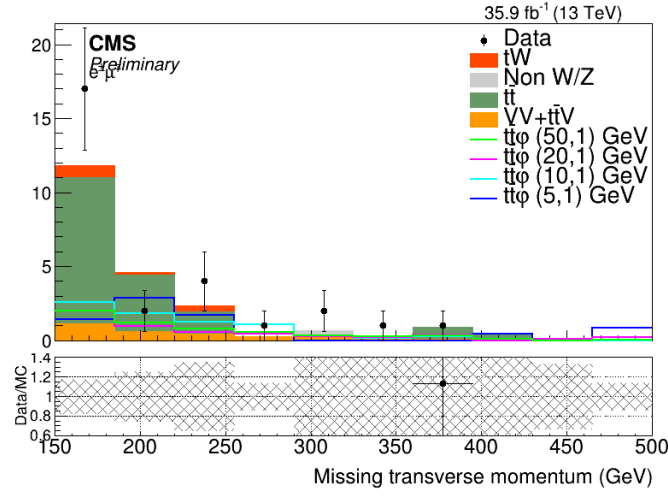
Tabla 5.5: Influencia de las incertidumbres sistemáticas en la estimación máximo verosímil de la fuerza de señal **esperada** para el mediador escalar $m_\varphi = 50$ GeV.

Es importante notar que las incertidumbres no suman 100 %, debido a que el estudio realizado no considera las incertidumbres asociadas a los datos.

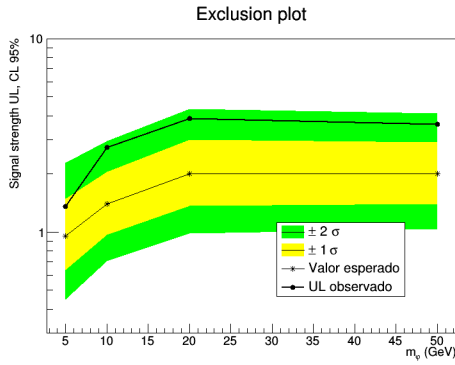
5.3. Resultados y conclusiones

En esta última sección del trabajo se empleará el análisis introducido anteriormente para llevar a cabo el análisis de los datos y obtener conclusiones sobre la veracidad del modelo estudiado.

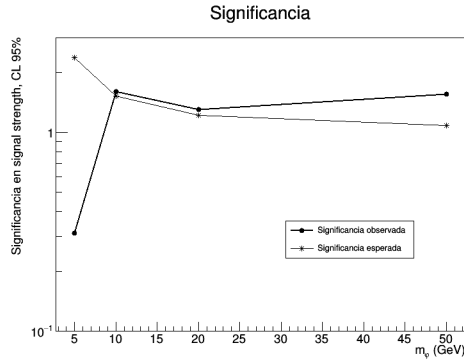
En primer lugar, en la figura 5.2 se muestran, en este orden: la distribución de la variable $\cancel{E}_T$, el gráfico de exclusión para el límite superior de los modelos y la significancia asociada a la estimación del límite superior de cada modelo, asumiendo existencia de la señal; todas ellas correspondientes al análisis óptimo introducido con los datos empleados. En el caso de la figura 5.2a se observan incertidumbres elevadas, con un acuerdo datos/MC superior en zonas de alto *momento transverso faltante*. En la figura 5.2b se observa un exceso de datos en los cuatro modelos estudiados, pero se encuentran dentro de la incertidumbre asociada a 2σ , y por tanto en acuerdo con el SM. Esto también implica que el análisis realizado permite discriminar modelos con secciones eficaces entre 1.5 y 3 veces la sección eficaz esperada de los modelos planteados.



(a) Representación gráfica de la variable E_T , con las cuatro señales propuestas e incluyendo las muestras de datos tomados.



(b) Gráfico de exclusión al nivel de confianza 95 % para el UL en la fuerza de señal de la masa del mediador escalar con la materia oscura, φ .



(c) Significancias al nivel de confianza 95 % para el UL en la fuerza de señal de la masa del mediador escalar con la materia oscura, φ . Se asume existencia de señal.

Figura 5.2: Acuerdo entre datos y simulaciones de MC en el análisis óptimo implementado.

Por otro lado, se muestran los valores de la fuerza de señal observada en la tabla 5.6 y las incertidumbres asociadas a los sistemáticos considerados para la señal observada en las tablas 5.7, 5.8, 5.9 y 5.10 (de nuevo, sólo se muestran contribuciones superiores al 1 %); la contribución se realiza en este caso a la estimación mostrada en la tabla 5.6. En este caso, se observan unos porcentajes de influencia superiores a los valores esperados observados en el apartado anterior, principalmente debidos a

Monte Carlo y a los procesos del canal $VV + t\bar{t}V$, aunque en este caso la influencia de los procesos $t\bar{t}$, JES y la luminosidad también son elevados.

Masa del mediador (GeV)	$\tilde{\mu}_{\text{obs}}$	Incertidumbres sup/inf
5	0.2	+0.6/-0.5
10	1.2	+0.9/-0.8
20	1.5	+1.3/-1.2
50	1.6	+1.2/-1.0

Tabla 5.6: Estimación máximo verosímil de la fuerza de señal **observada** obtenida para los cuatro modelos considerados.

Fuente de incertidumbre	Incertidumbre en $\tilde{\mu}_{\text{obs}}$	Contribución
JES	0.25	125 %
MonteCarlo Bin 3	0.18	90 %
$VV + t\bar{t}V$	0.17	85 %
MonteCarlo Bin 1	0.17	85 %
MonteCarlo Bin 4	0.15	75 %
MonteCarlo Bin 9	0.14	70 %
$t\bar{t}$	0.11	55 %
MonteCarlo Bin 7	0.10	50 %
MonteCarlo Bin 2	0.09	45 %
Luminosidad	0.09	45 %
MonteCarlo Bin 0	0.04	20 %
$tW, \bar{t}W$	0.02	10 %

Tabla 5.7: Influencia de las incertidumbres sistemáticas en la estimación máximo verosímil de la fuerza de señal **observada** para el mediador escalar $m_\varphi = 5$ GeV.

Fuente de incertidumbre	Incertidumbre en $\tilde{\mu}_{obs}$	Contribución
MonteCarlo Bin 2	0.26	22 %
MonteCarlo Bin 4	0.23	19 %
MonteCarlo Bin 3	0.22	18 %
$VV + t\bar{t}V$	0.20	17 %
MonteCarlo Bin 5	0.15	13 %
Luminosidad	0.13	11 %
$t\bar{t}$	0.13	11 %
MonteCarlo Bin 1	0.10	8 %
MonteCarlo Bin 7	0.10	8 %
JES	0.08	7 %
Non W/Z	0.07	6 %
MonteCarlo Bin 6	0.07	6 %
$tW, \bar{t}W$	0.02	2 %

Tabla 5.8: Influencia de las incertidumbres sistemáticas en la estimación máximo verosímil de la fuerza de señal **observada** para el mediador escalar $m_\varphi = 10$ GeV.

Fuente de incertidumbre	Incertidumbre en $\tilde{\mu}_{obs}$	Contribución
MonteCarlo Bin 2	0.36	24 %
$VV + t\bar{t}V$	0.35	23 %
MonteCarlo Bin 3	0.31	21 %
MonteCarlo Bin 7	0.23	15 %
MonteCarlo Bin 1	0.22	15 %
$t\bar{t}$	0.22	15 %
MonteCarlo Bin 6	0.21	14 %
Luminosidad	0.21	14 %
MonteCarlo Bin 4	0.21	14 %
MonteCarlo Bin 5	0.20	13 %
MonteCarlo Bin 9	0.18	12 %
Non W/Z	0.12	8 %
MonteCarlo Bin 0	0.06	4 %
$tW, \bar{t}W$	0.04	3 %

Tabla 5.9: Influencia de las incertidumbres sistemáticas en la estimación máximo verosímil de la fuerza de señal **observada** para el mediador escalar $m_\varphi = 20$ GeV.

Fuente de incertidumbre	Incertidumbre en $\tilde{\mu}_{obs}$	Contribución
MonteCarlo Bin 2	0.33	21 %
$VV + t\bar{t}V$	0.27	17 %
MonteCarlo Bin 3	0.22	14 %
MonteCarlo Bin 5	0.22	14 %
MonteCarlo Bin 6	0.19	12 %
MonteCarlo Bin 4	0.18	11 %
MonteCarlo Bin 7	0.16	10 %
Luminosidad	0.15	9 %
$t\bar{t}$	0.14	9 %
JES	0.13	8 %
MonteCarlo Bin 1	0.10	6 %
Non W/Z	0.10	6 %
MonteCarlo Bin 8	0.04	3 %
MonteCarlo Bin 9	0.03	2 %
$tW, \bar{t}W$	0.03	2 %

Tabla 5.10: Influencia de las incertidumbres sistemáticas en la estimación máximo verosímil de la fuerza de señal **observada** para el mediador escalar $m_\varphi = 50$ GeV.

Se incluyen a continuación distintos aspectos de referencia que permitirían el desarrollo de posteriores análisis en este ámbito:

- (i) **Introducción de nuevas variables de discriminación:** un aspecto fundamental en el análisis realizado se basa en la definición de una región de señal adecuada. Dado que la señal estudiada tiene una baja sección eficaz, una región de señal que reduzca de forma eficaz los distintos fondos es deseable. Por ello, no es descartable la definición de nuevas variables cinemáticas que permitan discriminar de forma eficaz esta señal de los diversos fondos.
- (ii) **Constante de acoplamiento del modelo:** un factor importante a considerar es la sección eficaz teórica predicha para las señales. Estas se obtienen suponiendo un valor para el acoplamiento del mediador escalar con las partículas del SM (g), tomando en este caso el valor uno para todas ellas. Por tanto, una variación en esta magnitud permitiría obtener modelos similares que permitirían estudiar este problema bajo distintos puntos de vista.
- (iii) **Aumento de las simulaciones de Monte Carlo:** como se ha observado en este trabajo, un factor limitante en la reducción de la incertidumbre asociada

proviene precisamente de las simulaciones de MC. El número de simulaciones empleadas es muy alto, permitiendo de esta forma el estudio de los procesos en el LHC con elevada precisión, por lo que un aumento del número de simulaciones generadas permitiría reducir de forma apreciable este factor. Este hecho entra en consonancia con las muestras de datos empleadas, tomadas en el año 2016, ya que se dispone de datos tomados en fechas posteriores que aún se hallan en la fase de filtrado y análisis *offline*; el empleo de estos datos en el futuro permitiría extender el análisis planteado en este trabajo.

- (iv) **Aumento de la luminosidad integrada:** la luminosidad es una medida del número de colisiones por unidad de tiempo y superficie, por tanto un aumento de esta magnitud implicaría un aumento en el número de sucesos detectados. Esto permitiría el estudio de procesos con bajas secciones eficaces, donde se podría aumentar el número de sucesos observados y así evitar trabajar en regiones de señal donde el MC influye considerablemente en la incertidumbre (como es el caso aquí estudiado).

Existe un proyecto para un acelerador con estas características (HL-LHC, *High Luminosity LHC*), aunque no estará operacional hasta más allá de 2025; y el aumento de la luminosidad conllevará además un aumento del *pile-up* asociado (es de aproximadamente unos 20 sucesos actualmente, y se esperan una media de 200 sucesos de PU en el HL-LHC). Por ello, se requeriría además mejoras sustanciales en el *hardware* y *software* del detector que permitan la toma y análisis de estos datos: *trigger*, nuevos algoritmos de reconstrucción, etc.

Conclusiones finales

En este último capítulo se recapitulan los resultados experimentales obtenidos en el estudio dedicado a este trabajo. Se ha llevado a cabo un estudio de modelos simplificados de producción de candidatos a DM de masa $m_\chi = 1$ GeV mediados por un bosón escalar de masas $m_\varphi = 5, 10, 20, 50$ GeV. Se emplean 35.9 fb^{-1} de datos tomados en CMS en el año 2016 a una energía de centro de masas de $\sqrt{s} = 13$ TeV, estudiados en el canal dileptónico. Se definen variables cinemáticas ([1],[2]) que permiten definir una región de señal (tabla 4.2) adecuada al estudio de los modelos planteados, reduciendo en un factor apreciable el número de sucesos de fondo hasta ser comparable a los sucesos de señal esperados. Los resultados obtenidos son compatibles con el modelo estándar, por lo que se procede a fijar límites en los modelos planteados. Se obtienen límites superiores para las secciones eficaces de los modelos consistentes con las predicciones del SM, y se estudia la influencia en las estimaciones de la fuerza de señal de los modelos debido a incertidumbres sistemáticas. Todos los resultados han sido obtenidos a un nivel de confianza del 95 %.

En caso de que se cumpliera el modelo, la fuerza de señal (equivalentemente, la sección eficaz) debería estar distribuida en torno al valor uno, indicando un acuerdo entre la sección eficaz experimental y la teórica predicha por el modelo de DM propuesto; valores menores a uno indicarían una falta de datos respecto a lo esperado por el modelo de DM, y por tanto, lo contradecirían. Por otro lado, un valor mayor a la unidad indicaría un exceso de datos respecto a lo esperado, y por tanto, conduciría a la posibilidad de una verificación del modelo.

A continuación se muestran los resultados obtenidos en la realización de este trabajo:

- **Estimaciones de la fuerza de señal**

Las incertidumbres asociadas a la fuerza de señal **esperada** (5.1) son compatibles con las predicciones del SM. La influencia de los sistemáticos introducidos en estas estimaciones (5.2 a 5.5) es apreciable, llegando a tomar valores en

torno al 25 % y siendo las principales contribuciones las simulaciones de MC y el fondo de procesos $VV + t\bar{t}V$; estos valores son esperados debido al bajo número de sucesos estudiados en la región de señal.

Por otro lado, los valores de fuerza de señal **observados** (5.6) están en acuerdo con el SM y permiten poner límites de exclusión sobre el resto de modelos propuestos. Sin embargo, son notables las influencias de los sistemáticos en estos valores (5.7 a 5.10), especialmente en el caso del mediador de masa $m_\varphi = 5$ GeV. De nuevo, estas incertidumbres se deben principalmente al Monte Carlo y los procesos $VV + t\bar{t}V$.

■ Determinación del límite superior observado

La significancia asociada a los modelos (5.2c) muestra que el límite superior de la fuerza de señal observado es consistente con lo predicho por el modelo estándar, y permite excluir el modelo con mediador escalar de masa $m_\varphi = 5$ GeV al nivel de confianza prefijado. Por otro lado, el límite superior observado correspondiente a cada modelo (5.2b) es consistente con el modelo estándar y no permite verificar la exclusión de dicho modelo en una incertidumbre de 2σ ; este hecho se atribuye a fluctuaciones estadísticas y sistemáticas en los datos tomados.

Se concluye, por tanto, que se ha desarrollado un método capaz de excluir modelos con mediadores escalares de masa $m_\varphi = 5$ GeV y fijar límites en las secciones eficaces correspondientes al resto de modelos, todos ellos compatibles con las predicciones del SM.

Una de las fuentes principales de incertidumbre en este trabajo ha sido las simulaciones de Monte Carlo, debido al estudio de señales con bajas secciones eficaces, y por tanto, bajo número de sucesos respecto a los fondos principales. Esto conlleva definir regiones de señal donde se trabaja con pocos sucesos, y por tanto aumenta la incertidumbre asociada a la estadística. Otra de ellas han sido los procesos agrupados como $VV + t\bar{t}V$, mostrados en la tabla 4.1. Este hecho es debido a la presencia del canal de desintegración del proceso $t\bar{t}Z$ en que el bosón Z se desintegra en un par de neutrinos, y por tanto constituye un **fondo irreducible** para el proceso de señal considerado.

Por último, cabe destacar que se ha realizado un análisis comparable a otro tipo de búsquedas similares realizadas en CMS y ATLAS y se han obtenido resultados compatibles en acuerdo con las predicciones del SM. El método desarrollado

permite fijar límites en la nueva física basada en los modelos de DM planteados, fundamentando las bases para posteriores análisis en este ámbito que empleen todos los datos correspondientes al *Run 2* del LHC y, por tanto, minimicen las contribuciones estadísticas que limitan el análisis.

Referencias y bibliografía consultada

- [1] The ATLAS Collaboration. *Search for direct top squark pair production and dark matter production in final states with two leptons in $\sqrt{s} = 13$ TeV pp collisions using 13.3 fb^{-1} of ATLAS data* (ATLAS-CONF-2016-076) /
- [2] The ATLAS Collaboration. *Search for dark matter produced in association with bottom or top quarks in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*. Eur. Phys. J. C 78 (2018) 18. doi: 10.1140/epjc/s10052-017-5486-1. arXiv: 1710.11412
- [3] *Particle Data Group*. (<http://pdg.lbl.gov/>)
- [4] *LHC Dark Matter Working Group: Next-generation spin-0 dark matter models*. CERN-LPCC-2018-02. arXiv: 1810.09420
- [5] A. Boveia, C. Doglioni. *Dark Matter Searches at Colliders*. arXiv: 1810.12238
- [6] *The Standard Model: a beautiful but flawed theory*. <https://www.quantumdiaries.org/2014/03/14/the-standard-model-a-beautiful-but-flawed-theory/>. [En línea; consulta el 09-07-2019].
- [7] F. Bechtel, P. Schleper. (2019). *The Underlying Event in Proton-Proton Collisions*.
- [8] Khachatryan, Vardan et al. *Observation of top quark pairs produced in association with a vector boson in pp collisions at $\sqrt{s} = 8$ TeV*. CMS-TOP-14-021 CERN-PH-EP-2015-248. arXiv:1510.01131
- [9] Aaltonen, T. et al. *Search for Standard Model Higgs Boson Production in Association with a W Boson Using a Matrix Element Technique at CDF in p-bar p Collisions at $\sqrt{s} = 1.96$ TeV*. FERMILAB-PUB-11-664. EarXiv:1112.4358
- [10] CMS Wiki Pages. *How to make Feynman diagrams in LaTeX*. <https://wiki.physik.uzh.ch/cms/latex:feynman> [En línea; 13-06-2019]
- [11] V. Bouza (2018). *Studies of the tW production process in p-p collisions at $\sqrt{s} = 13$ TeV at the CMS detector* (Universidad de Oviedo)
- [12] Ars Technica. *A sort of particle-free supersymmetry found in exotic materials*. <https://cdn.arstechnica.net/wp-content/uploads/2014/04/SM-SUSY-diagram.jpg> [En línea; 13-06-2019]
- [13] Hoddeson, Lillian H. et al. (1997). *The Rise of the Standard Model: Particle Physics in the 1960's and 1970's*. Cambridge University Press; 1ª edición.

- [14] Doglioni C., Tovey D. *Searching for Dark Matter with the ATLAS detector*. <https://atlas.cern/updates/atlas-feature/dark-matter> [En línea; consulta el 30-06-2019].
- [15] G. Salam. (2009). *QCD (for LHC)Lecture 2: Parton Distribution Functions*. <https://gsalam.web.cern.ch/gsalam/repository/talks/2009-Bautzen-lecture2.pdf>. European School of High-Energy Physics, Bautzen, Alemania [En línea; consulta el 30-06-2019].
- [16] T. Potter. *Feynman Diagrams*. https://www.hep.phy.cam.ac.uk/~chpotter/particleandnuclearphysics/Lecture_05_FeynmanDiagrams.pdf. High Energy Physics Group, University of Cambridge. [En línea; consulta el 30-06-2019].
- [17] Drees, M. (1996). *An Introduction to Supersymmetry*. APCTP-96-05, KEK-TH-501. arXiv:hep-ph/9611409v1
- [18] CERN Home Page. *The Large Hadron Collider*. <https://home.cern/science/accelerators/large-hadron-collider> [En línea; consulta el 09-07-2019].
- [19] CERN Home Page. *The accelerator complex*. <https://home.cern/science/accelerators/accelerator-complex> [En línea; consulta el 09-07-2019].
- [20] The CMS Collaboration. *The performance of the CMS muon detector in proton-proton collisions at $\sqrt{s} = 7$ TeV at the LHC*. arXiv:1306.6905v2. doi:10.1088/1748-0221/8/11/P11002.
- [21] The CMS Collaboration. *The CMS trigger system*. arXiv:1609.02366v2 [physics.ins-det].
- [22] The CMS Collaboration, Beaudette F. (2014). *The CMS Particle Flow Algorithm*. arXiv:1401.8155v1.
- [23] The CMS Collaboration. *Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV*. arXiv:1712.07158v1
- [24] T. M. Liss, F. Maltoni, A. Quadt (2017). *The Top Quark*. <http://pdg.lbl.gov/2017/reviews/rpp2017-rev-top-quark.pdf> [En línea; consulta el 09-07-2019].
- [25] Grupo Experimental de Altas Energías de la Universidad de Oviedo. *Repositorio 'Analysis-PAF'*. <https://github.com/Oviedo-PAF/AnalysisPAF>
- [26] Grupo Experimental de Altas Energías de la Universidad de Oviedo. *PAF*. <http://www.hep.uniovi.es/PAF/>
- [27] CMS Run2 Samples, Summer 2016. <https://docs.google.com/spreadsheets/d/1b4qnWfZrimEGYc1z4dH121-A9qyJgpqNUbh0lvCzjbE/edit?pref=2&pli=1#gid=1199164479> [En línea; consulta el 04-07-2019].
- [28] The CMS Collaboration, Armando B. (2017). *$\Delta\phi$ and multi-jet correlations with CMS*. PoS(DIS2017)167
- [29] T. Melia. *Spin before mass at the LHC*. JHEP 01 (2012) 143. arXiv: 1110.6185 [hep-ph]
- [30] The CMS Collaboration. Repositorio de GitHub *HiggsAnalysis-CombinedLimit*. <https://github.com/cms-analysis/HiggsAnalysis-CombinedLimit>
- [31] The CMS Collaboration. Repositorio de GitHub *CombineHarvester*. <https://github.com/cms-analysis/CombineHarvester>
- [32] StackExchange. *TikZ-Feynman: layers and alignment*. <https://tex.stackexchange.com/questions/343670/tikz-feynman-layers-and-alignment>. [En línea; consulta el 25-06-2019].