

FACULTAD DE CIENCIAS

TRABAJO DE FIN DE GRADO DE FÍSICA

---

**Búsqueda de partículas supersimétricas en el marco de los modelos SUSY con el experimento CMS del LHC (CERN) a 13 TeV.**

---

*Autor:* Carlos Francisco ERICE CID

*Tutor:* Javier CUEVAS MAESTRO

12 de julio de 2016

# Índice

<b>1. Introducción</b>	<b>3</b>
<b>2. Extensiones al modelo estándar</b>	<b>5</b>
2.1. El Modelo Estándar	5
2.2. Limitaciones del SM	7
2.3. Extensiones supersimétricas	9
2.3.1. Búsquedas de SUSY	11
2.4. Teorías de campos efectivas	13
<b>3. El acelerador LHC y el experimento CMS</b>	<b>15</b>
3.1. El gran colisionador de hadrones	15
3.1.1. Detectores en el LHC	17
3.2. El detector CMS	18
3.2.1. Imán solenoidal	19
3.2.2. Tracker	20
3.2.3. Calorímetro electromagnético	21
3.2.4. Calorímetro hadrónico	22
3.2.5. Detectores de muones	22
3.3. <i>Trigger</i>	23
<b>4. Identificación de objetos y procesado de datos</b>	<b>25</b>
4.1. Algoritmo <i>particle flow</i>	25
4.1.1. Reconstrucción del vértice principal	27
4.1.2. Reconstrucción de muones	27
4.1.3. Reconstrucción de electrones	29
4.1.4. Reconstrucción de jets	29
4.1.5. Energía faltante	30
4.2. Simulaciones de Montecarlo	31
4.3. Herramientas informáticas	33
<b>5. Procesos de interés y muestras</b>	<b>34</b>
5.1. Procesos MSSM	34
5.2. Procesos de EFT	36
5.3. Procesos de fondo	37
5.4. Muestras de datos	39
5.5. Preselección de eventos	40
<b>6. Métodos de análisis</b>	<b>42</b>
6.1. Cortes secuenciales	42
6.2. <i>Boosted Decision Trees</i>	43
6.2.1. Árbol de decisión	43
6.2.2. Construcción del BDT	45
6.2.3. Características del BDT	48
6.3. Incertidumbres estadísticas y sistemáticas	49
6.3.1. <i>Full CLs</i>	49
6.3.2. Método asintótico	51
6.4. Proyección de resultados	51
6.5. Variables de interés	52

<b>7. Estudio de los procesos <math>T_{2tt}</math></b>	<b>56</b>
7.1. Parámetros del BDT . . . . .	56
7.1.1. Profundidad del BDT . . . . .	57
7.1.2. Árboles totales . . . . .	58
7.1.3. Parámetro de <i>boost</i> . . . . .	59
7.2. Punto 850 – 100 . . . . .	60
7.2.1. Muestras y variables . . . . .	60
7.2.2. Regiones de estudio . . . . .	63
7.2.3. Proyección de resultados . . . . .	69
7.2.4. <i>Poda</i> de variables . . . . .	70
7.3. Procesos de entrenamiento y test . . . . .	72
7.4. Punto 600 – 300 . . . . .	75
7.4.1. Proyección de resultados . . . . .	78
7.5. Resultados globales . . . . .	78
<b>8. EFT para DM</b>	<b>80</b>
8.1. Proyección de resultados . . . . .	83
<b>9. Conclusiones</b>	<b>84</b>
<b>10. Bibliografía y referencias</b>	<b>86</b>

## 1. Introducción

Durante la segunda mitad del s.XX los avances en el campo de la Física de Altas Energías se centraron, fundamentalmente, en el desarrollo y la comprobación de las propiedades del conocido como **Modelo Estándar**. Tras la observación experimental del bosón de Higgs en los detectores CMS y ATLAS (2012), se confirmaba la existencia de la última de las partículas necesarias para dicho modelo.

Pese a dar una descripción notablemente precisa de las interacciones fuertes y electrodébil, el propio Modelo Estándar presenta una serie de carencias a la hora de explicar diversos fenómenos observables en la naturaleza. De esta forma se motiva la necesidad de las búsquedas de física *más allá del Modelo Estándar* o de *nueva física*.

Enmarcado en el área de conocimiento de la Física de Partículas, el objetivo principal de este estudio es la ampliación de los conocimientos obtenidos en las asignaturas de Física Nuclear y de Partículas Elementales, Física de Altas Energías y Aceleradores y Métodos Numéricos y sus Aplicaciones a la Física de cara a un análisis de datos que nos permitiese determinar la presencia experimental de dichos procesos de *nueva física* a través de datos obtenidos en el experimento CMS.

En primer lugar expondremos brevemente las motivaciones de nuestro estudio: la situación actual del Modelo Estándar y su estructura, los hechos experimentales que conducen a la necesidad de su ampliación y los diferentes caminos abiertos en la investigación de la misma. En este último punto nos centraremos en dos paradigmas fundamentales que, en cierta forma, se encuentran relacionados: la búsqueda de **materia oscura** (**DM** por sus siglas en inglés) y de **procesos supersimétricos** (**SUSY**). Buscaremos realizar una caracterización adecuada del tipo de signatura experimental esperable para estos procesos en términos de la detección en aceleradores centrándonos especialmente en el **canal dileptónico**, objeto central de nuestro estudio posterior.

Seguidamente, expondremos las características básicas de los dispositivos experimentales para la toma de datos, su procesado y almacenamiento. Comenzando con una breve descripción del **gran colisionador de hadrones** (**LHC**) para centrarnos posteriormente en el detector **CMS** y en los diferentes subdetectores, o *capas*, del mismo. Conocida la estructura del detector, buscaremos caracterizar y explicar los procesos de **identificación y reconstrucción** de los diversos objetos físicos medidos en nuestro experimento. Adicionalmente, incorporaremos una breve descripción del funcionamiento de las **simulaciones** como complemento fundamental en el estudio experimental de estos datos y su importancia a la hora de la estimación de los procesos de fondo en nuestras medidas como modelado de las predicciones teóricas.

Posteriormente, introduciremos los distintos **métodos de análisis** utilizados en nuestro estudio a fin de separar procesos de *nueva física* de los procesos ya conocidos. El hecho central a estudiar será la transición de los conocidos métodos *univariante* a los estudios de tipo **multivariable**, dando especial importancia en este último caso a los árboles de decisión *empujados* (*Boosted Decision Trees*, **BDT**) que se utilizarán en el

análisis de datos posterior. A continuación, haremos una breve introducción a los diversos métodos utilizados para la incorporación de las **incertidumbres** (estadística y sistemática) a nuestros resultados y la forma de extender nuestros resultados a situaciones de mayor luminosidad (una mayor cantidad de datos). Finalmente, introduciremos las variables de interés físico en nuestro estudio, sobre las que aplicaremos los métodos anteriores.

Como punto central de nuestro trabajo, incluiremos los diferentes resultados obtenidos mediante la aplicación de las técnicas anteriormente descritas a datos obtenidos en el LHC complementados con simulaciones de los diferentes procesos de fondo. Nuestro objetivo fundamental en este caso es proporcionar los resultados obtenidos en términos de las posibilidades de **observación** o **exclusión** de procesos de *nueva física* en las regiones del espacio de parámetros consideradas. Para ello, consideraremos tanto procesos de tipo **SUSY** como **DM**.

Finalmente, realizaremos una recapitulación de los resultados obtenidos, indicando adicionalmente los posibles caminos de mejora y búsqueda alternativa a seguir en estudios posteriores.

## 2. Extensiones al modelo estándar

### 2.1. El Modelo Estándar

El Modelo Estándar de la Física de Partículas (**SM**) describe con gran exactitud los procesos correspondientes a las interacciones fuerte y electrodébil observados hasta el momento [1]. Comprende una descripción de las propiedades generales de las partículas que conforman la materia así como de las relaciones entre las mismas y las interacciones que sufren de tipo fuerte o electrodébil. Actualmente se encuentra confirmada experimentalmente la existencia de las partículas que forman parte del Modelo Estándar, con la reciente incorporación del bosón de Higgs en el año 2012. Un esquema de las partículas fundamentales del SM se puede observar en la Figura 1.

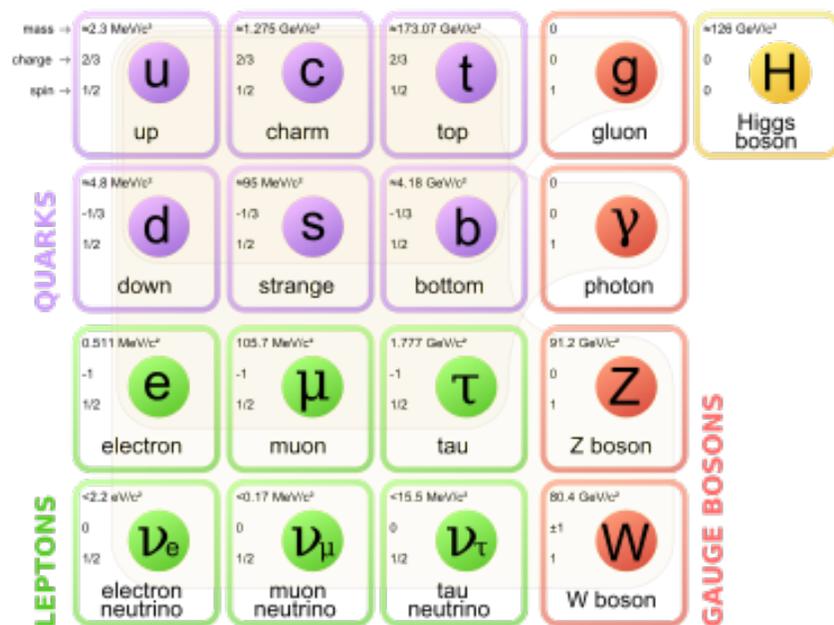


Figura 1: Esquema de las partículas del Modelo Estándar.

Las partículas elementales se encuentran divididas en dos grandes grupos o *sectores*. Por una parte tenemos los **fermiones** (partículas de spin semientero que se rigen por la estadística de Fermi-Dirac, en la Figura 1 en colores verde y morado), que tradicionalmente se han considerado como partículas *de materia*. Por otra parte se encuentran los **bosones** (partículas de spin entero que se rigen por la estadística de Bose-Einstein, en la Figura 1 en colores rojo y amarillo) que tradicionalmente se han considerado partículas *portadoras de interacciones*. En el sector fermiónico se observa una estructura básica repetida en tres subsectores o **generaciones**, siendo la diferencia fundamental entre una partícula de una generación y sus correspondientes en el resto la variación de su masa (que aumenta al aumentar la generación). En cada una de estas generaciones se encuentra repetido el mismo esquema: un leptón cargado, un leptón sin carga y dos quarks con cargas de signo opuesto. Adicionalmente, cada

una de las partículas del sector fermiónico lleva asociada su correspondiente antipartícula<sup>1</sup>, con propiedades físicas idénticas a la original salvo la carga eléctrica que es opuesta.

Los **quarks** (en morado en la Figura 1) son las únicas partículas fermiónicas del Modelo Estándar con carga de color, es decir, que sufren la interacción fuerte. En cada generación tenemos un quark con carga eléctrica positiva<sup>2</sup>  $+2/3$  y otro con carga eléctrica negativa  $-1/3$ . De esta forma tendremos los quarks *up* (primera generación, carga  $+2/3$ ), *down* (primera generación, carga  $-1/3$ ), *strange* (segunda generación, carga  $+2/3$ ), *charm* (segunda generación, carga  $-1/3$ ), *top* (tercera generación, carga  $+2/3$ ) y *bottom* (tercera generación, carga  $-1/3$ ). Es común abreviarlos por  $u, d, s, c, t, b$  (respectivamente), mientras que sus correspondientes antipartículas se denotan añadiendo una barra superior:  $\bar{u}, \bar{d}, \bar{s}, \bar{c}, \bar{t}, \bar{b}$ .

Por otra parte, para el caso de los **leptones** (en verde en la Figura 1) se encuentra en cada generación una partícula con carga eléctrica  $-1$  pero neutra respecto a la interacción fuerte. Cada una de estas partículas lleva asociado su correspondiente **neutrino** que se suele denominar a partir del leptón cargado al que acompaña y es neutro tanto desde el punto de vista electromagnético como fuerte. Los leptones cargados se conocen como electrón (primera generación), muón (segunda generación) y tau (en ocasiones tauón, tercera generación). Es común abreviarlos por  $e, \mu, \tau$ , mientras que los neutrinos se suelen denotar  $\nu_e, \nu_\mu, \nu_\tau$ .

Dentro del sector bosónico tenemos una división fundamental entre bosones de tipo vectorial (spin 1) y escalar (spin 0). Los bosones vectoriales (en la Figura 1 en rojo) actúan como portadores de las interacciones fundamentales mientras que el, hasta ahora, único bosón escalar conocido está asociado al mecanismo que proporciona masa a las partículas elementales a través de interacciones con las mismas.

El primero de los bosones vectoriales es el **fotón**,  $\gamma$ , responsable de la interacción electromagnética. Se trata de una partícula con carga eléctrica y sin carga de color, además de tener masa nula. El fotón es su propia antipartícula.

Las partículas  $W^+$ ,  $W^-$  y  $Z^0$  son los **bosones vectoriales masivos** responsables de la interacción débil. Ninguno de los tres tiene carga de color por lo que no interaccionan fuertemente. Las cargas eléctricas de cada uno son, respectivamente,  $+1$ ,  $-1$  y  $0$ . El hecho de que sean masivos hace que la interacción débil tenga un rango notablemente pequeño, pues a grandes distancias su intensidad disminuye con gran rapidez. Guardan una relación fundamental con el fotón en cuanto a que las cuatro partículas se obtienen directamente a partir de la ruptura de la simetría (relacionada con el bosón de Higgs) de la unificada interacción electrodébil. El bosón  $W^+$  es la antipartícula del bosón  $W^-$  y el bosón  $Z^0$  es su propia antipartícula. La interacción débil se acopla a todas las partículas del SM.

<sup>1</sup>En el caso de los neutrinos, se especula la posibilidad de que sean partículas de Majorana, es decir, que sean su propia antipartícula.

<sup>2</sup>En este caso, y en lo sucesivo, tomaremos como unidad de carga la carga del electrón:  $e$ .

Por último, los **gluones** (genéricamente denotados  $g$ ) son los responsables de la interacción fuerte. Si bien son partículas con masa nula y sin carga eléctrica tienen una particularidad de notable importancia: tienen carga de color. Es decir, que los gluones pueden sentir la propia interacción que portan. Este hecho da lugar al fenómeno de *confinamiento* que nos impide observar individualmente partículas con carga de color no nula. De forma efectiva, esto provoca que todo quark o gluón *libre* de lugar a un proceso de *hadronización* a fin de formar estructuras con carga de color total nula<sup>3</sup>. Este mismo fenómeno es responsable de que la interacción fuerte tenga también un rango limitado.

Finalmente, el bosón escalar presente en el Modelo Estándar es el **bosón de Higgs** (en la Figura 1 en amarillo). Dicho bosón no tiene carga eléctrica ni de color. Interacciona con las partículas fermiónicas y los bosones  $W^\pm$  y  $Z$  dándoles masa.

## 2.2. Limitaciones del SM

Si bien el Modelo Estándar ha demostrado ser un marco notablemente preciso para la explicación de los fenómenos propios de la Física de Altas Energías, varios hechos justifican la necesidad de una ampliación del mismo que pueda responder a estos a un nivel más fundamental. Esto motiva la búsqueda de modelos más allá del Modelo Estándar (**BYM**, *Beyond the Standard Model*) que se materializan experimentalmente en los procesos de *nueva física*. Presentamos a continuación una breve recopilación de algunos de los más relevantes motivos para la búsqueda de dichos procesos de *nueva física* [2], [3]:

- **Gravedad:** El principal reto a superar es el hecho de que la gravedad no queda explicada en modo alguno a través del Modelo Estándar. Desde el punto de vista del formalismo de las Teorías Cuánticas de Campos, o **QFT**, (marco teórico fundamental del Modelo Estándar), la existencia de una interacción gravitatoria debería llevar asociada una hipotética partícula bosónica, el gravitón, de la que hasta el momento no se ha encontrado indicio alguno. Adicionalmente, las posibles teorías QFT que incorporan la gravedad son no *renormalizables*, con lo que no encajan con las observaciones clásicas de los efectos de la interacción gravitatoria. Desde un punto de vista experimental, a las energías de trabajo del LHC (del orden de 10 TeV) la interacción gravitatoria tiene una intensidad despreciable respecto al resto (y por tanto puede ser considerada aparte) pero la evolución de las distintas constantes de acoplo con la energía podría hacer que fuese relevante a energías lo suficientemente altas (escala de Planck,  $10^{16}$  TeV). A medida que nos aproximásemos a este último caso el tratamiento conjunto sería necesario.

---

<sup>3</sup>Se generan pares quark-antiquark y se recombinan con el original hasta dar partículas neutras para la interacción fuerte.

- **Jerarquía:** Al existir un bosón escalar dentro de la teoría, se puede comprobar que la escala natural a la que se encontraría su masa tendría correcciones del orden del *cut-off*<sup>4</sup> en la misma. Por tanto, se esperaría encontrar un Higgs con una masa del orden de, al menos, la escala de Planck lo que obviamente discrepa con las medidas experimentales (125 GeV).
  
- **Materia oscura:** Observaciones de carácter astronómico y cosmológico apuntan a la existencia de la conocida como materia oscura (**DM**). Se trata de materia estable (a escalas de tiempo cosmológicas) que no interacciona de forma electromagnética pero sí lo hace gravitatoriamente. Sus efectos se pueden medir a través del estudio de gran cantidad de fenómenos en el universo: curvas de rotación de galaxias, perfiles de dispersión de velocidades estelares, efectos de lente gravitacional a gran escala, etc. De todas las partículas del Modelo Estándar, únicamente los neutrinos cumplirían las propiedades necesarias para ser candidatos a **DM** pero no son lo suficientemente abundantes como para componer la proporción del universo acotada por las observaciones cosmológicas.
  
- **Energía oscura:** De la misma forma que para el caso de la materia oscura, existen evidencias de la existencia de la energía oscura (**DE**). En este caso concreto, al estudiar el fondo cósmico de microondas, es posible observar que, aún incluyendo a la materia observable la cantidad adicional de materia oscura, es necesaria una densidad de energía mayor en el universo para justificar el ritmo expansivo del universo. A la energía *faltante* para justificar estas observaciones se la denomina energía oscura. Ninguno de los fenómenos o partículas del Modelo Estándar supone un candidato razonable para explicar dicha energía oscura.
  
- **Parámetros libres:** En su expresión actual, el Modelo Estándar presenta un total de 19 parámetros libres (contando las diferentes masas y constantes de acoplo presentes). Si bien están determinados experimentalmente, muchos de ellos con gran precisión, resulta notablemente extraña la imposibilidad de relacionarlos entre sí mediante un marco teórico adecuado. Modelos que incluyen la reducción de estos parámetros libres conforman lo que se conoce como teorías de unificación<sup>5</sup>.
  
- **Simetría CP:** Los modelos cosmológicos más aceptados para el origen del universo incorporan el proceso de bariogénesis que da lugar a la asimetría actual entre materia y antimateria. Para eliminar dicha simetría

---

<sup>4</sup>La escala de energías hasta la cual es aplicable la teoría. Tiene especial importancia al considerar los procesos de *renormalización* en la misma.

<sup>5</sup>Una Gran Teoría de Unificación (GUT), incluye la relación de todas las posibles constantes de acoplo de las interacciones en una sola, dando lugar a una *unificación* de las interacciones.

tría son necesarios procesos que violan la simetría CP<sup>6</sup>. Si bien existen procesos en el Modelo Estándar que permiten esta violación de CP (principalmente procesos regidos por la interacción débil) no son, en principio, lo suficientemente frecuentes como para justificar la alta asimetría observada actualmente.

Tal cantidad de evidencias nos conducen directamente a la necesidad de desarrollar marcos más allá del Modelo Estándar que permitan explicar, al menos en parte, estos hechos.

### 2.3. Extensiones supersimétricas

Una de los paradigmas más extendidos como posible ampliación al Modelo Estándar viene dado por el concepto de Supersimetría (**SUSY**). La idea fundamental en este caso consiste en contemplar la existencia de una partícula *compañera* para cada una de las partículas del Modelo Estándar que tenga un spin de comportamiento opuesto. Es decir, a cada fermión le correspondería un compañero supersimétrico bosónico y a cada bosón le correspondería un compañero supersimétrico fermiónico [4].

Centrémonos en nuestro caso particular en proporcionar un breve desarrollo de los componentes básicos del Modelo Estándar Supersimétrico Mínimo (**MSSM**), una de las posibles extensiones SUSY más sencillas del modelo estándar que proporcionaría una solución (teóricamente) a una buena parte de los problemas planteados en la sección anterior.

En el MSSM, cada uno de los fermiones del SM queda asociado a un bosón distinto de tipo escalar (con spin 0) a los que en conjunto denominaremos **sfermiones**<sup>7</sup>. Por otra parte, para cada uno de los bosones tendremos un compañero supersimétrico que será un fermión de spin 1/2; en conjunto estos últimos se denominan **gauginos**<sup>8</sup>.

Es bien sabido que no se ha detectado hasta el momento ninguna partícula supersimétrica. Asumiendo que sufren las interacciones del Modelo Estándar resultaría extremadamente improbable que no hubiésemos producido ninguna si sus masas se encontrasen a niveles de energía accesibles a los aceleradores de los que se disponen actualmente. Es por tanto lógico observar que tiene que aparecer una ruptura en la simetría entre las partículas del SM y sus compañeras supersimétricas en cuanto al comportamiento de su masa. Es decir, que el sector supersimétrico de MSSM tendría que estar compuesto por partículas de una masa muy superior a las del SM. Para la ruptura de esta simetría se introducen términos adicionales en el lagrangiano (no es una ruptura *espontánea*) cuyas características determinarán el comportamiento general del modelo.

---

<sup>6</sup>Transformación simultánea de conjugación de carga y de paridad espacial.

<sup>7</sup>Usualmente se utiliza el prefijo *s-* para indicar que estamos tratando de la compañera supersimétrica. Así, el squark stop sería el compañero supersimétrico del quark top:  $\tilde{t}$

<sup>8</sup>En estos casos se utiliza el sufijo *-ino* para su nomenclatura. Hablaremos, por ejemplo, del gluino  $\tilde{g}$ .

Recordemos que la ruptura espontánea de la simetría produce la aparición de  $\gamma$ ,  $W^+$ ,  $W^-$ ,  $Z^0$  a través de una combinación de las partículas asociadas a la interacción electrodébil  $B$ ,  $W^1$ ,  $W^2$ ,  $W^3$ . Un proceso similar se observa cuando los compañeros supersimétricos de estas últimas partículas se mezclan con el Higgsino (compañero supersimétrico del Higgs). De esta mezcla aparecen las partículas que denominamos **neutralinos** ( $\chi_1^0, \chi_2^0, \chi_3^0, \chi_4^0$ ) y *charginos* ( $\chi_1^\pm, \chi_2^\pm$ ).

Adicionalmente, en el caso de los sfermiones, aparece una ruptura en la simetría de masas en función de la **quiralidad** de las partículas. El resultado efectivo es que los estados de masa *observados* (las partículas) se corresponden con dos partículas de masa diferente mezcla de los posibles estados quirales de los sfermiones. Así del par  $\tilde{t}_R, \tilde{t}_S$  se desprenden dos partículas *reales* con diferente masa  $\tilde{t}_1$  y  $\tilde{t}_2$ . En el modelo MSSM estas diferencias son relativamente pequeñas salvo en el caso del squark stop[4]. Debido a la alta masa del quark top, la diferencia de masa en el par  $\tilde{t}_1$  y  $\tilde{t}_2$  es más elevada que en el resto. Una consecuencia importante en gran cantidad de modelos es que, como consecuencia, el squark  $\tilde{t}_1$  es el sfermión más ligero en gran cantidad de modelos MSSM.

De la misma forma que podemos definir un operador  $C$  (conjugación de carga) que transforme cada partícula en su correspondiente antipartícula, podemos intentar definir un operador  $Q$  que a cada partícula le haga corresponder su compañera supersimétrica. Por ejemplo:  $Q|t\rangle = |\tilde{t}\rangle$  y  $Q|\tilde{t}\rangle = |t\rangle$ . Asociado a este operador podemos definir cierto número cuántico  $R$  con valores  $\pm 1$  de forma que  $R = 1$  para el SM y  $R = -1$  para las partículas supersimétricas. De esta forma,  $Q|R = 1\rangle = |R = -1\rangle$ . Formalmente, es posible determinar el valor del número cuántico  $R$ :

$$R = (-1)^{3(B-L)+2S}$$

Donde  $B$  es el número bariónico,  $L$  el número leptónico y  $S$  el spin de la partícula. Observemos que, bajo estas definiciones, toda partícula del SM tiene  $R = 1$  (las partículas con número bariónico o leptónico no nulo son los fermiones en el SM). De la misma forma, toda compañera supersimétrica tendrá  $R = -1$ . En MSSM se incorpora el concepto de **simetría-R**, es decir, toda desintegración conservará el valor total de  $R$ . De forma efectiva, esta simetría implicará que en procesos que partan de estados del SM, las partículas SUSY se producirán siempre *a pares*.

Mención especial dentro del modelo, por las implicaciones que tiene tanto desde el punto de vista de la detección como de la explicación de los procesos de materia oscura, merece la partícula supersimétrica más ligera (**LSP**, *Lightest Supersymmetric Particle*). La importancia de esta partícula aparece en cuanto observamos que, si se tiene que conservar la simetría-R, ha de ser necesariamente estable. De otra forma, tendrían que existir partículas más ligeras con  $R = -1$  (para que al desintegrarse produjese una de ellas). Por tanto, la LSP apare-

cería al final de las cadenas de desintegración de las partículas SUSY, por lo que se suele exigir que tenga carga eléctrica y de color nulas. Con estas condiciones, el candidato habitual a LSP es el neutralino más ligero, el  $\chi_1^0$ .

Desde el punto de vista de la sección anterior, la teoría MSSM da respuesta a varios de los problemas mencionados:

- Jerarquía: MSSM implica la aparición de diagramas de Feynman adicionales al considerar las correcciones a la masa del Higgs (notablemente elevadas en el SM). Dichos diagramas dan lugar a correcciones que se cancelan con las del SM dando lugar a la posibilidad de una masa similar a la observada (*ajuste fino*).
- Materia oscura: MSSM proporciona candidatos notablemente buenos a componentes de la Materia Oscura. La LSP (usualmente el  $\chi_1^0$ ) es considerada un firme candidato a una partícula **WIMP**<sup>9</sup> como componente principal de la materia oscura, pues es eléctricamente neutra, sin carga de color y masiva.
- Parámetros libres: Aunque en general las teorías supersimétricas pueden complicar enormemente el total de parámetros libres, modelos razonables de MSSM pueden construirse consiguiendo una unificación de interacciones. En función de los términos que se incorporen al lagrangiano para la ruptura de la simetría de masas se puede llegar a reducir el número de parámetros libres a 5.
- Simetría CP: La introducción de términos que rompan la simetría de masas en SUSY puede conducir a la aparición de procesos de violación de CP que den cuenta de esta asimetría fundamental en edades tempranas del universo.

Nótese que, en principio, MSSM no proporciona soluciones naturales a dos de los problemas fundamentales planteados anteriormente. No proporciona una explicación ni para la energía oscura ni para la gravedad. Se trata, por tanto, como mínimo de un modelo efectivo de la misma forma que lo es el SM.

### 2.3.1. Búsquedas de SUSY

Desde el punto de vista de nuestro trabajo, nos interesa analizar el camino natural para la búsqueda de SUSY mediante la utilización de aceleradores-colisionadores de partículas, concretamente en el marco del LHC. La idea fundamental en este caso es que, gracias a las elevadas energías en el centro de masas que se producen en las colisiones,  $\sqrt{s} = 13TeV$ , debería ser posible generar las partículas más ligeras de MSSM, cuyas masas oscilan (según algunos modelos) en el rango de  $100GeV - 1TeV$ . Obviamente, la señal experimental a analizar dependerá completamente de las características del modelo MSSM que utilicemos (y de los valores otorgados

---

<sup>9</sup> *Weakly Interacting Massive Particle*. Los estudios de DM indican que partículas con masas del orden de  $100GeV$  interactuantes con las intensidades típicas de la interacción débil darían una respuesta notablemente precisa a las mediciones astrofísicas de DM, lo que se conoce como el *WIMP Miracle* [5].

a los parámetros libres en el mismo). Presentamos a continuación una breve caracterización de las propiedades relevantes de estos modelos desde el punto de vista experimental.

Asumiendo que la paridad-R se conservase en los procesos, esperaríamos llegar a observar, por lo menos, la presencia de la LSP en nuestros experimentos. Obviamente, en función de la naturaleza de la LSP esperaremos unos tipos u otros de señal experimental asociada con las propiedades de la misma. Usualmente, se considera como LSP el neutralino de menor masa  $\chi_1^0$  aunque algunos modelos consideran un hipotético gravitino (compañero del no menos hipotético gravitón) o incluso el chargino más ligero  $\chi_1^\pm$ .

Los rangos de masas de las partículas producibles son también especialmente relevantes. De poder alcanzar a producir partículas supersimétricas mucho más masivas que la LSP, las cadenas de desintegración podrían dar lugar a pares de partículas altamente energéticas cuya señal experimental será muy distinta al caso menos energético de masas más próximas.

Por último, destaquemos que no es estrictamente necesario que las partículas SUSY tengan vidas medias lo suficientemente pequeñas como para que se desintegren en el interior del detector. Es decir, que partículas SUSY que no fuesen la LSP podrían *escapar* a nuestra observación, complicando el estudio del propio proceso.

Usualmente se utilizan **Modelos Simplificados** que permitan reunir ideas básicas de MSSM (y supersimetría en general) eliminando gran parte de su complejidad. La idea básica de los modelos simplificados es estudiar directamente una cantidad muy reducida de partículas supersimétricas y cadenas de desintegración posibles. A cambio de la simplicidad del modelo, se pierde toda la generalidad de SUSY incluyendo todas las relaciones entre los parámetros de las diferentes partículas e interacciones.

La ventaja es que, desde el punto de vista fenomenológico, permiten definir con relativa sencillez búsquedas particulares de procesos supersimétricos que se pueden modelizar adecuadamente.

En nuestro caso particular, nos centraremos en el análisis de datos correspondientes a un modelo simplificado tipo  $T2tt$  [6]. En este modelo, consideramos únicamente las partículas supersimétricas dadas por  $\tilde{t}$  (en caso de ruptura de simetría de masas de las partículas quirales, el stop más ligero) y  $\chi_1^0$  (LSP). El proceso a considerar viene dado por:

$$pp \rightarrow \tilde{t}\tilde{t} \rightarrow t\bar{t}\chi_1^0\chi_1^0$$

Donde suponemos que en colisiones  $pp$  (las que se producen en el LHC) se pueden producir pares stop-antistop. Cada uno de los squarks decae a su vez en un quark top y un neutralino. La señal final vendrá dada, entonces, por la signatura típica de un proceso  $t\bar{t}$  más la aparición de dos neutralinos. Véase la Figura 2 (Derecha).

En nuestro modelo simplificado, podemos considerar que la *branching fraction*<sup>10</sup> de este proceso es total (es decir, que todo  $\tilde{t}$  decae en  $t\chi_1^0$ ) y que el tiempo de vida medio del  $\tilde{t}$  es lo suficientemente bajo como para que se produzca su desintegración en el interior del detector. De esta forma, los dos parámetros fundamentales que restan son las dos masas de las partículas supersimétricas  $m_{\tilde{t}}$  y  $m_{\chi_1^0}$ . Por tanto, se suele hablar del *plano o espacio* de parámetros  $m_{\tilde{t}} - m_{\chi_1^0}$  haciendo referencia a todos los posibles pares de valores de las masas en este modelo simplificado. Cada una de las posibles combinaciones de masas da lugar a un punto en ese plano.

El procedimiento general de trabajo para estudiar el proceso  $T2tt$  se basa en analizar el espacio de parámetros  $m_{\tilde{t}} - m_{\chi_1^0}$  en busca de una señal experimental en cada uno de los posibles puntos contando los eventos observados y sustrayendo a estos la cantidad total de eventos de fondos conocidos. Cuando la luminosidad recogida es lo suficientemente alta como para que se tuviesen que producir estos procesos y no se observan se dice que hemos logrado **excluir** el punto correspondiente del espacio de parámetros.

Notemos, por otra parte, que la sección eficaz del proceso viene determinada únicamente por el acoplo existente entre  $\tilde{t}\tilde{t}$  y el sector SM. Observemos que este proceso no debería depender, en principio, de la masa del neutralino por lo que la sección total de nuestro proceso depende únicamente de la masa del squark  $\tilde{t}$  en nuestro modelo simplificado. Este hecho es fundamental, pues a medida que aumenta la posible masa del  $\tilde{t}$  la sección eficaz del proceso se haría tan pequeña que su producción en colisiones es cada vez más escasa.

## 2.4. Teorías de campos efectivas

Asumimos que un modelo BSM que pretenda explicar la materia oscura tiene que cumplir dos requisitos mínimos. El primero es la existencia de un *sector oscuro*, es decir, un conjunto de partículas que no pertenezcan al SM, que interactúen entre sí y que contengan, al menos, una partícula con las cualidades básicas como para componer la materia oscura (masiva, con cargas eléctrica y de color nulas). Por otra parte, para propiciar la búsqueda a través de aceleradores (que no podemos asegurar que sea posible) es necesario que exista una forma de interacción entre dicho sector oscuro y las partículas del SM (en otro caso no podríamos producirla).

Si estudiamos el breve desarrollo de la sección anterior sobre MSSM, resulta obvio que estos modelos proporcionan un primer marco posible para la explicación de la materia oscura. Sin embargo, a la hora de realizar búsquedas de materia oscura puede resultar de más utilidad alejarnos de la generalidad del problema y restringirnos a casos *simplicios* en los que, de nuevo, los parámetros a determinar sean menores y los procesos aparezcan simplificados.

Un caso paradigmático en estos modelos son las teorías de campos efectivas (**EFT**) [7]. En esencia, una EFT considera la existencia de una partícula (que compone la materia oscura) y una interacción de la misma a determinadas partículas del modelo estándar que se produce de forma *puntual* (es decir, sin la aparición de un bosón

<sup>10</sup>La proporción de eventos en los que una partícula decae según un modo dado.

mediador). Sugestivamente, esto es equivalente a considerar que el mediador de la interacción es excesivamente masivo para nuestras energías de trabajo y, por tanto, se produce únicamente de manera virtual en un corto intervalo de tiempo (con lo que la interacción es *casi puntual*). Desde el punto de vista formal, las diferentes EFT se formalizan considerando los posibles acoplos que respetan las simetrías del problema y son compatibles con los términos conocidos del SM, lo que da lugar a los distintos *operadores* que describen distintos modelos EFT. Materialmente, cada uno de estos acoplos se traducirá en las posibles formas de interacción entre las partículas lo que dará lugar a los distintos procesos experimentales.

En nuestro caso particular, nos centraremos en un modelo EFT en el cual la interacción entre SM y DM se produce *por pares* entre quarks top y partículas de materia oscura. Es decir, tenemos el proceso  $t\bar{t} \rightarrow \chi\bar{\chi}$  (donde  $\chi$  denota la hipotética partícula de DM), como se puede ver en la Figura 2 (Izquierda).

Notemos que en este caso tendremos dos parámetros libres fundamentales que constituyen el espacio paramétrico de estudio: la intensidad del acoplo (en términos de cierta constante de acoplo) y la masa de la partícula de materia oscura  $m_\chi$ . Denominaremos a este proceso  $tt + DM$  o, simplemente,  $DM$ .

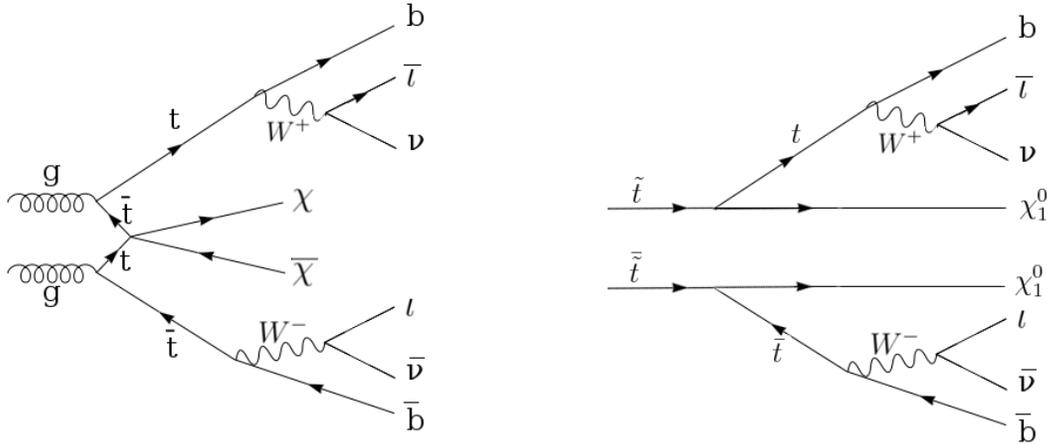


Figura 2: Diagramas de Feynman de los procesos de señal.  $tt + DM$  (Izquierda).  $T2tt$  (Derecha)

### 3. El acelerador LHC y el experimento CMS

La forma usual de generar procesos de *nueva física* pasa por la aceleración de partículas bien conocidas (y técnicamente sencillas de manipular) hasta alcanzar energías lo suficientemente altas como para que, en una colisión entre ellas, se puedan producir gran cantidad de nuevos procesos en cantidades relevantes<sup>11</sup>. Debido a que los procesos de *nueva física* no se han observado, por el momento, a las energías de trabajo previas de los aceleradores de partículas, resulta relativamente intuitiva la necesidad de alcanzar escalas energéticas más altas para tener esperanzas de encontrar este tipo de fenómenos. En los experimentos de aceleradores-colisionadores, esta condición se traduce en obtener energías de centro de masas  $\sqrt{s}$  lo más elevadas posibles con lo que la sección eficaz de los procesos de producción, en principio, aumentará (el espacio de fases disponible para las partículas en su estado final es mayor). Asimismo, es necesario obtener luminosidades integradas (en esencia datos totales) lo suficientemente elevadas como para poder confirmar nuestros resultados estadísticamente a cada energía de trabajo.

En la actualidad, el gran colisionador de hadrones (**LHC**) es el acelerador de partículas en funcionamiento que permite alcanzar energías de centro de masas más elevadas y es, por tanto, el candidato fundamental a proporcionar los primeros indicios de eventos de *nueva física*. Nos centraremos en trabajar con los datos producidos en el mismo que han sido recogidos por el experimento **CMS** (*Compact Muon Solenoid*) uno de los dos detectores de carácter general situados en el mismo.

Al tratarse de las fuentes primeras de los datos que utilizaremos en nuestro análisis, es fundamental entender tanto el funcionamiento del LHC como del CMS a la hora de la aceleración de partículas, su colisión y el estudio de los productos resultantes.

#### 3.1. El gran colisionador de hadrones

El gran colisionador de hadrones (**LHC**) es el acelerador-colisionador de partículas capaz de alcanzar mayores energías hasta la fecha. Se trata de un gigantesco acelerador circular con 27km de circunferencia situado subterráneamente en la frontera franco-suiza cercana a Ginebra.

Originalmente diseñado para la aceleración simultánea de dos haces de partículas<sup>12</sup> (en sentidos opuestos) hasta alcanzar energías  $\sqrt{s} = 14TeV$  en colisión (7 *TeV* por haz), hasta el momento se ha realizado un aumento progresivo de las energías de trabajo desde su primera puesta en marcha en 2009. Entre los años 2010 y 2013 se realizó la primera fase de toma de datos en el LHC con energías que pasaron de los  $\sqrt{s} = 7TeV$  iniciales a  $\sqrt{s} = 8TeV$  finales teniendo como resultado destacado el descubrimiento (y la confirmación de este) del bosón de Higgs.

---

<sup>11</sup>Las secciones eficaces típicas en estos procesos tienen una fuerte dependencia con la energía. A mayores energías aumenta con rapidez la sección eficaz total.

<sup>12</sup>En nuestro caso nos interesan las colisiones de protones, pero también es utilizado para la colisión de iones pesados.

Tras dos años dedicados al mantenimiento y mejora del LHC, en 2015 comenzó la segunda serie de tomas de datos a una energía  $\sqrt{s} = 13TeV$  con el objetivo de aumentarla hasta  $\sqrt{s} = 14TeV$  en el futuro. En el desarrollo de este trabajo se disponía de datos correspondientes a una luminosidad integrada  $L = 1,3fb^{-1}$  tomados durante el breve periodo de 2015. Se estaba comenzando (Junio de 2016) el periodo de toma de datos de 2016 a la luminosidad instantánea nominal de  $\mathcal{L} = 10^{-2}fb^{-1}s^{-1}$ .

El LHC cumple la función fundamental de proporcionar paquetes de partículas altamente energéticos y garantizar la colisión de los mismos en los puntos adecuadamente designados para la situación de los detectores. Para ello, tiene que ser capaz de acelerar dichos paquetes de partículas mientras mantiene su cohesión y dirección (mantener el haz dentro del propio acelerador) y modificar su desplazamiento para poder cruzar los haces en los puntos de colisión.

De cara al proceso de acelerado, el LHC no parte de partículas en reposo sino que un sistema de aceleradores previos proporciona ya los paquetes a una energía lo suficientemente elevada (en torno a 450 GeV por haz) para la operación del LHC [8]. Al principio de dicho sistema se encuentran dos aceleradores lineales (uno destinado a la aceleración de protones, LINAC2, y el otro a iones pesados, LINAC3). A continuación se encuentran dos aceleradores circulares (PS y SPS) que forman los paquetes de partículas en los haces y proporcionan aumentos de la energía de cada uno de forma sucesiva hasta alcanzar la energía de inyección del LHC.

El principio de funcionamiento de estos aceleradores es en todos casos muy similar. Para producir la aceleración de las partículas en el LHC se utilizan las **cavidades de radiofrecuencia** (véase la Figura 3). Se trata de secciones en las que se producen campos eléctricos alternantes a frecuencias de, aproximadamente,  $400MHz$  en secciones contiguas. El efecto que se consigue al coordinar esta sección con el paso de las partículas cargadas es que dichas partículas experimenten en todo momento un campo eléctrico *hacia la dirección de desplazamiento*. De forma efectiva, esto produce un aumento en la energía de las partículas del haz que, además, compensa sobradamente la pérdida fundamental en el acelerador, la producida por la radiación sincrotrón<sup>13</sup>.

Para conseguir dirigir los haces a lo largo del *tubo* del acelerador, es necesario poder modificar su dirección en cada giro de la misma. Para ello, se utilizan imanes dipolares superconductores que rodean al haz de partículas. Estos imanes producen un potente campo magnético (hasta  $8,3T$ ) en dirección vertical cuyo resultante es una fuerza de Lorentz central sobre las partículas del haz que les hace tomar su trayectoria curva.

Por último, es deseable (de cara a maximizar la luminosidad en el punto de colisión, aumentar las colisiones) compactificar lo más posible los paquetes de protones de forma que se obtengan paquetes más *densos*. Con este objetivo, se dispone de cuadrupolos magnéticos. Aunque la acción de un cuadrupolo magnético comprime

---

<sup>13</sup>Radiación que produce toda partícula cargada que sigue una trayectoria curva.

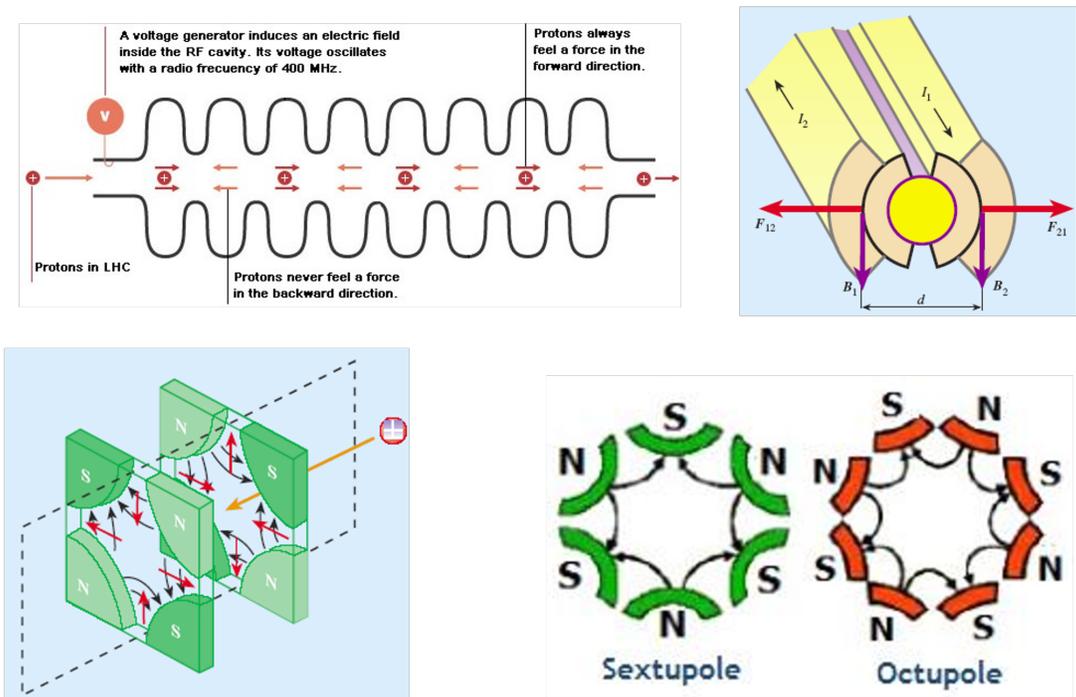


Figura 3: Sistemas electromagnéticos en el LHC. (Sup. Izquierda) Cavidades de radiofrecuencia. (Sup.Derecha) Dipolos magnéticos. (Inf. Izquierda) Cuadrupolos magnéticos. (Inf. Derecha) Multipolos magnéticos. [9]

los paquetes en uno de los dos ejes transversales a su dirección de desplazamiento, realmente tiene el efecto contrario en el otro eje. Por ello, los cuadrupolos sucesivos en el LHC se encuentran *cruzados* de forma que sus efectos dispersivos se cancelen sucesivamente y se obtenga un haz *fino*. Multipolos magnéticos de orden superior (sextupolos) se utilizan para realizar correcciones en las partículas con energías ligeramente desviadas de la energía media del haz, de forma que se proporciona a todo el haz una energía uniforme.

Imanes multipolares adicionales se encuentran situados en las cercanías de los puntos de colisión de forma que curven ligeramente la trayectoria de los dos haces para producir el cruce de los mismos (y las consecuentes colisiones de partículas) en el punto deseado.

### 3.1.1. Detectores en el LHC

En el LHC se encuentran, actualmente, cuatro detectores principales situados en cuatro puntos distintos en los que es posible producir el cruce de los haces para producir las colisiones de alta energía buscadas.

**ALICE** estudia colisiones de iones pesados (plomo-plomo) con el objetivo de observar las propiedades de la interacción fuerte. Concretamente, busca generar y estudiar el comportamiento del plasma de quark-gluones.

**LHCb** está centrado en el estudio de mesones B (que incluyen el quark b entre sus quarks de valencia), busca observar la posibilidad de violaciones de la simetría CP en procesos que impliquen a estos mesones.

**CMS** y **ATLAS** son experimentos de carácter general orientados a búsquedas mucho menos específicas que ALICE y LHCb. Se encuentran diseñados para la detección del rango más amplio posible de las diferentes partículas del SM sin centrarse específicamente en un sector específico del mismo. Dado tal diseño, pueden utilizarse para gran variedad de estudios distintos en la física de Partículas. Aunque sus objetivos son relativamente similares, ambos detectores utilizan sistemas ligeramente distintos lo que da la posibilidad de confirmación independiente de los resultados de uno por parte del otro.

### 3.2. El detector CMS

En nuestro caso particular trabajaremos con datos obtenidos originalmente en el detector CMS por lo que es de especial interés comprender adecuadamente su funcionamiento y la forma que tiene de detectar los diferentes tipos de partículas posibles.

El detector en sí es una enorme máquina de forma aproximadamente cilíndrica con una longitud de  $28,7m$  y un diámetro de  $14,6m$ . Su estructura interior está organizada en forma de subdetectores o *capas*, con cada uno de ellos especializado en un aspecto u otro del proceso de detección [10]. Una vista esquemática se puede observar en la Figura 4.

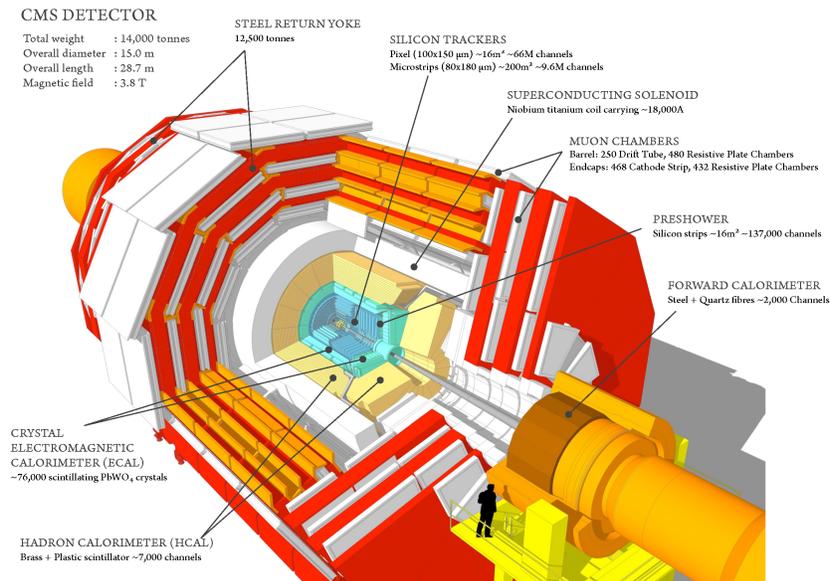


Figura 4: Interior del detector CMS con los diferentes subdetectores identificados.[11]

Desde el punto de vista funcional, es necesario destacar en primer lugar el convenio utilizado en la designación de coordenadas en el detector CMS, pues no coincide con las coordenadas *usuales* utilizadas en otros campos de la física.

En coordenadas ortogonales usuales, consideramos el eje  $OZ$  el correspondiente eje tangente al haz de partículas que discurre por el centro del detector. Los ejes  $OX$  y  $OY$  forman parte del plano transversal a los propios

haces con el  $OX$  siendo el normal a la superficie terrestre (y valores positivos *hacia fuera*).

A partir de este sistema  $(x, y, z)$  podemos considerar las coordenadas esféricas asociadas  $(R, \theta, \phi)$ . Tomamos en este caso el ángulo  $\phi$  definido positivamente sobre el plano transversal e inicio en el eje  $OX$  ( $\phi \in [0, 2\pi)$ ), mientras que el ángulo  $\theta$  queda medido respecto al eje  $OZ$  y varía en  $[-\pi/2, \pi/2]$ . Por último,  $R$  queda definida como la distancia del punto considerado al origen.

Dadas estas coordenadas realizamos una última transformación. En primer lugar, en vez de considerar  $R$ , tomaremos la proyección de esta cantidad sobre el plano  $XY$ ,  $\rho$ , correspondiente a la distancia de la proyección de nuestro punto sobre el plano transversal y el origen. La motivación fundamental para este cambio es el hecho de que en el plano transversal la suma de momentos será, en principio, nula (los haces iniciales no tienen momento transversal). Además, el momento transversal de las partículas cargadas resulta medible con relativa facilidad mientras que el longitudinal, en general, no lo es. Por último, definimos la pseudorapidez,  $\eta$ , como  $\eta = -\log(\tan(\frac{\theta}{2}))$ . La principal ventaja de la pseudorapidez es que, en los casos que nos ocupan (con altos momentos longitudinales), aproxima con gran exactitud la rapidez: una magnitud invariante Lorentz. Observemos que bajo esta definición  $\eta = 0$  en el plano transversal  $XY$  y  $\eta = \pm\infty$  a lo largo del eje del haz.

En lo sucesivo, utilizaremos el sistema de coordenadas  $(\rho, \eta, \phi)$  y cuando hablemos de componentes transversales nos estaremos refiriendo a la coordenada en  $\rho$  de la correspondiente cantidad.

### 3.2.1. Imán solenoidal

Comenzamos detallando las distintas capas del detector con el imán solenoidal que le da nombre. Este se observa de color blanco en la Figura 4. Si bien no es la capa más interna, su relevancia desde el punto de vista de toma de datos y la reconstrucción de partículas está íntimamente relacionada con los subdetectores más internos.

Situado entre el calorímetro hadrónico (**HCAL**) y las cámaras de muones se trata de un imán superconductor de forma solenoidal cuyo eje está alineado con el eje del haz de partículas. Con una longitud de  $13m$  y  $6m$  de diámetro el campo magnético que produce en la zona central de su eje es aproximadamente constante. Dicho campo es de una magnitud de  $3,8T$ .

La utilidad fundamental de este campo magnético es el hecho de que *curva* las trayectorias de las partículas cargadas lo que proporciona un método geométrico simple de determinar su momento transversal. Considerando la fuerza de Lorentz responsable de la trayectoria curva de la partícula, es fácil deducir que:

$$\frac{p_T^2}{mR} = m \frac{v_T^2}{R} = |\vec{F}_L| = |q\vec{p} \times \vec{B}| = q|\vec{p}| \sin(\theta)|\vec{B}| = qp_T B$$

Donde  $B$  es el campo magnético,  $p_T$  el momento transversal de la partícula,  $m$  su masa y  $R$  el radio de

curvatura de su trayectoria. Sin más que despejar obtenemos que  $p_T = qRB$ . Como es usual en Física de Partículas, nos interesa explicitar esta fórmula en múltiplos de unidades de  $eV$  (lo que es equivalente a multiplicar por  $c$  los momentos). Se obtiene directamente la fórmula clásica:

$$p_T [GeV] = 0,3q [e]R [m]B [T]$$

Donde el factor 0,3 aparece al considerar el valor de  $c$  y la carga de la partícula está expresada en unidades de la carga del electrón. De esta forma podemos determinar el momento transverso de toda partícula cargada que atraviese el detector si somos capaces de reconstruir su trayectoria.

### 3.2.2. Tracker

La primera de las capas del detector CMS permite precisamente la medición de los radios de curvatura necesarios. Dicha capa recibe el nombre de detector de trazas o **tracker**. Se puede observar de color azul en la Figura 4.

La idea básica detrás del *tracker* reside en registrar el paso de las partículas a través de un mallado notablemente denso de pequeños sensores de forma que se puedan reconstruir sus trayectorias en los puntos más cercanos a la colisión sin alterar demasiado el desplazamiento de la propia partícula (con deposiciones pequeñas de energía). Nótese que los detectores del *tracker* no nos permitirán discernir qué partícula ha realizado cada deposición, por lo que la resolución espacial es vital para distinguir unas partículas de otras. Este problema se agudiza enormemente al observar que en las colisiones típicas tenemos un fenómeno de *pile-up*<sup>14</sup> que multiplica el total de partículas detectadas simultáneamente. Adicionalmente, la posibilidad de la aparición de vértices secundarios en procesos de interés (generación de quarks  $b$  por ejemplo) da lugar a la necesidad de medidas de precisión de las trayectorias que permitan reconstruirlos adecuadamente.

El principio físico por el que se rigen los detectores del *tracker* se basa en la interacción entre dichas partículas cargadas y los electrones del silicio que lo compone. Al paso de partículas cargadas, la interacción con la materia produce pares electrón hueco que posteriormente se pueden detectar. Si dividimos el espacio en secciones independientes de materia podemos estudiar la deposición de energía que se produce en cada una y así reconstruir la trayectoria de las partículas a través de las mismas.

Basado en este principio en el *tracker* se disponen de dos tipos de detectores de silicio : **píxeles** y **strips**.

Los píxeles cubren la región más cercana al haz, con una capa interna situada a apenas  $4cm$  del punto de colisión. El sistema de píxeles está compuesto por tres capas cilíndricas y dos capas adicionales planas que

---

<sup>14</sup>Producción de varias colisiones en un pequeño intervalo de tiempo. Debido a la alta densidad de protones en los haces que colisionan es común que el detector registre simultáneamente datos correspondientes a decenas de colisiones.

recubren las bases de dichos cilindros (con pequeñas aberturas centrales en torno al eje del haz). Cada una de estas capas está compuesta por pequeños fragmentos de silicio (píxeles) de  $100\mu m \times 150\mu m$ . Con esta configuración, una partícula de bajo  $|\eta|$  (alto momento transverso),  $|\eta| < 2,5$ , atravesará al menos tres capas de píxeles y, por tanto, tendremos al menos tres puntos de alta precisión que forman parte de su trayectoria.

Rodeando a las secciones de píxeles tenemos las tiras o *strips* de silicio. Con un funcionamiento similar al de los píxeles, estas secciones presentan una resolución mucho menor pues son de un tamaño bastante mayor. Esta parte del subdetector se divide a su vez en cuatro subsecciones distintas en las que las tiras están diseñadas de acuerdo a la forma y función en su posición: el *Inner Barrel*, el *Outer Barrel*, el *Inner Endcap* y el *Outer Endcap*. Las dos primeras cubren el espacio  $|\eta| < 1,5$  y están compuestas, respectivamente, por 4 y 6 capas cilíndricas concéntricas de tiras de silicio la última de ellas situada a  $1,3m$  del haz. Las dos últimas forman las *tapas* correspondientes a estos cilindros,  $1,5 < |\eta| < 2,5$  y constan respectivamente de 3 y 1 capas de tiras.

Debido a ligeras diferencias entre el comportamiento de los sensores tipo *Barrel* y los tipo *Endcap*, las trayectorias de partículas con bajo  $|\eta|$  se miden con mejor resolución. En principio esto es altamente necesario en la búsqueda de resultados posteriores. Por eso, se suele decir que las partículas de bajo  $\eta$  detectadas tienen más *calidad* que las de alto  $\eta$ .

### 3.2.3. Calorímetro electromagnético

Tras el *tracker* se encuentra el calorímetro electromagnético o **ECAL** que se puede observar de color azul verdoso en la Figura 4. Como su nombre indica, su objetivo es la detección y medición de las propiedades de las partículas que interactúan electromagnéticamente, es decir, fotones y partículas cargadas eléctricamente.

El funcionamiento básico del ECAL se basa en el proceso de **centelleado**. En un material centelleador se producen excitaciones de los electrones atómicos al paso de partículas ionizantes (partículas cargadas o fotones). Las desexcitaciones de estos electrones producen la emisión de fotones lo que da lugar al proceso de *centelleo*. Obviamente, en este proceso las partículas ionizantes pierden sucesivamente energía con cada interacción por lo que pueden llegar a ser absorbidas en el material dando lugar a señales claras en el detector en forma de *depósitos de energía*.

El ECAL del CMS está compuesto por gran cantidad de cristales de wolframato de plomo ( $PbWO_4$ ), un material centelleador de alta densidad que detiene una gran parte de las partículas ionizantes incidentes. Los cristales se encuentran dispuestos radialmente respecto al punto central del detector al que presentan cada uno una superficie relativamente pequeña  $2cm \times 2cm$  lo que proporciona una resolución notablemente buena de la posición de interacción con las partículas ionizantes.

Sin embargo, aunque los cristales centelleadores proporcionan una señal luminosa, esta es demasiado leve. Para solucionar este problema cada uno de los cristales centelleadores lleva incorporado un fotomultiplicador

que amplifica la señal de centelleo. El proceso se da de tal forma que estudiando la intensidad de la señal final es posible determinar la energía depositada en cada uno de los cristales por las partículas ionizantes.

Estructuralmente, el ECAL de CMS, como el *tracker* puede considerarse dividido en una zona de *barrel*,  $|\eta| < 1,48$ , correspondiente a un cilindro central (los cristales son más estrechos y los fotomultiplicadores son fotodiodos de avalancha) y la zona de *endcap*,  $1,48 < |\eta| < 3,00$  (cristales más anchos y fotodiodos de vacío). Como en el caso del *tracker*, la zona del *barrel* proporciona una mayor precisión espacial en las medidas.

### 3.2.4. Calorímetro hadrónico

Envolviendo al ECAL se encuentra el calorímetro hadrónico o **HCAL**, de color beige y amarillo en la Figura 4.

La idea de funcionamiento básico es muy similar a la del ECAL: utilizar la radiación de partículas ionizantes producida en materiales centelleadores cuya detección nos permita definir la energía y posición de las partículas. Sin embargo, el HCAL tiene como objetivo principal la detección y absorción de hadrones neutros, que en principio no interactúan electromagnéticamente. Para solucionar este problema se instaure un sistema de capas alternas de materiales centelleadores y materiales de alta densidad (bronce, acero). Las partículas hadrónicas interactuarán (por interacción fuerte) con los materiales de alta densidad dando lugar a *cascadas de partículas* que se pueden detectar en el material centelleador posterior por los mismos procesos que en el ECAL.

De cara a detectar todas las partículas posibles de las que se producen en el proceso el HCAL incorpora una tercera sección al *barrel* ( $|\eta| < 1,4$ ) y *endcap* ( $1,3 < |\eta| < 3,0$ ): *forward* ( $2,9 < |\eta| < 5,0$ ) dedicada a la detección de las partículas con alto momento longitudinal.

### 3.2.5. Detectores de muones

El más externo de los subdetectores, los detectores de muones (de color rojo, amarillo y blanco en la Figura 4), se encuentra situado en el exterior del imán solenoidal (y por lo tanto el campo magnético en su zona tiene signo opuesto al que tenemos en el interior). El nombre de este dispositivo se debe a que, por la baja interacción que presentan con los calorímetros, se espera que los muones sean los únicos que puedan atravesar el resto de capas del detector<sup>15</sup>. Por tanto, esta última sección del detector está específicamente diseñada para la medida de las propiedades de este tipo específico de leptones.

El principio físico en el que se basan los detectores de muones es la producción de *corrientes de avalancha*. Al atravesar un material gaseoso, los muones (partículas cargadas) producen la ionización (separación del átomo neutro en electrón e ión) del mismo en un sistema situado entre dos electrodos de distinto signo. La diferencia de potencial entre dichos electrodos acelera el electrón y el ión resultantes haciendo que estos a su vez ionicen más

---

<sup>15</sup>En la práctica siempre es posible que otras partículas atraviesen el detector, pero es relativamente menos frecuente.

átomos del gas hasta dar lugar a la generación de una corriente entre ambos electrodos que se puede detectar con dispositivos electrónicos.

Funcionando bajo este principio, se dividen los detectores de muones en tres secciones fundamentales. En primer lugar, en la zona del *barrel* ( $|\eta| < 1,2$ ) se encuentran los *drift tubes* (DT). En la zona *endcap* ( $1,2 < |\eta| < 2,4$ ) están presentes las *cathode strip chambers* (CSC) con un tiempo de reacción menor. Finalmente, rodeando al sistema se encuentran las *resistive plate chambers* (RPC) con menor eficiencia de detección que los otros dos sistemas pero una reacción más rápida (lo que resulta especialmente útil de cara a la definición del *trigger*).

### 3.3. *Trigger*

El último componente del proceso de detección es el dispositivo de **trigger**. Si bien el resto de componentes del detector se encuentran encaminados a estudiar la *física* de las colisiones que se producen, el *trigger* responde a un problema puramente funcional: la cantidad de información sobre las colisiones que se produce es demasiado elevada y no puede ser almacenada por completo. El proceso por el que se seleccionan los eventos *de interés* que serán almacenados para su posterior análisis conforma el *trigger*. Obviamente nos interesa conservar la mayor proporción posible de estos dentro de las limitaciones físicas para el grabado de la información, es decir, maximizar la eficiencia de *trigger*.

En el caso del detector CMS el proceso de *trigger* se lleva a caso en dos pasos que distinguiremos como *Level 1* (L1) y *high level trigger* (HLT)[12].

El primero de los dos subsistemas, *L1*, proporciona una decisión rápida que reduce en tres órdenes de magnitud la cantidad de eventos registrados. Para ello, se basa en las mediciones directas, sin reconstruir ni identificar las partículas observadas en el proceso, para estimar las cantidades físicas relevantes. La información se procesa por separado en el ECAL y HCAL (*trigger de calorímetro*) y los detectores de muones (*trigger de muones*)<sup>16</sup>. La información procesada por ambos se combina en el *trigger global* que decide almacenar o no un evento en función de los resultados. Los eventos almacenados se envían, junto con la información del L1, al HLT.

Tradicionalmente, se subdividen los procesos del HLT en *Level 2* (L2) y *Level 3* (L3) incorporando este último una mayor información sobre el evento<sup>17</sup> pero proporcionando una respuesta más lenta. No obstante, en el caso del CMS ambos niveles se encuentran incorporados secuencialmente en el HLT en vez de corresponderse con procesos separados. El HLT se subdivide a su vez en varios subprocesos distintos encargados del análisis de diferentes tipos de Cerns físicos: electrones y fotones (en conjunto), muones, jets<sup>18</sup>, b-jets (producidos por un

---

<sup>16</sup>En ambos casos existen subdivisiones de cada *trigger* que recogen la información *capa a capa* del detector. La información del *tracker* no se utiliza en el proceso de *trigger L1*.

<sup>17</sup>Por ejemplo, puede incorporar datos del *tracker*.

<sup>18</sup>El proceso de hadronización de los quarks que se producen en el LHC da lugar a que lo que se logre detectar sean *cascadas de partículas hadrónicas* orientadas en la dirección del momento del quark original (aproximadamente) conocidas comúnmente como

quark  $b$ ) y jets producidos por leptones  $\tau$ . La información de cada uno de estos subprocesos del HLT se combina para dar lugar al *trigger* o selección global que indica si se registra la información del evento o no de forma permanente.

Obviamente, no todos los procesos de interés físico estudiados tienen la misma forma de señal en el estado final. Por ello, se diseñan diversas configuraciones, o rutas, de cara a proporcionar *triggers* de diferentes características para las diferentes firmas esperadas. Así, no será lo mismo hablar de un *trigger dileptónico* que de un *trigger single lepton*, etc. En nuestro caso, dado que nos interesa estudiar los procesos  $T2tt$  y  $tt+DM$  específicamente en el canal dileptónico, nos centraremos únicamente en eventos que hayan pasado alguno de los *trigger* dileptónicos configurados en el CMS.

---

jets de hadrones.

## 4. Identificación de objetos y procesado de datos

Hasta el momento nos hemos centrado en dar una descripción de los principios básicos en el funcionamiento del LHC y del detector CMS. Aunque relevante de cara a entender todo el proceso de estudio y las posibles consecuencias que puedan tener en nuestros propios resultados, nos interesa especialmente introducir todo el proceso de elaboración que los datos que se obtienen a nivel de detector sufren de cara a recopilar la información física que contienen. Es decir, nos interesa entender el procedimiento de **reconstrucción** que, a partir de las mediciones de las deposiciones de energía en el detector, nos permite formar las partículas físicas que forman parte del evento.

Por otra parte, aunque los datos tomados en el detector resultan fundamentales en las búsquedas de *nueva física*, también es necesario destacar la importancia de las simulaciones en las mismas. Dado que, obviamente, las medidas reales no nos permiten discrepar completamente unos eventos de otros (ya que únicamente vemos las consecuencias de los mismos y pueden coincidir), las simulaciones suponen en muchos casos la mejor forma de la que disponemos para estudiar los comportamientos *en vacío* de algunos de los procesos más relevantes en las medidas. De la misma forma, la relevancia de las simulaciones adquiere un mayor relieve en cuanto a que, aunque desconocemos el comportamiento exacto de los procesos de *nueva física*, podemos simular los mismos mediante distintos modelos. Se obtienen así resultados de simulación que contrastar con los datos experimentales.

En el siguiente capítulo introducimos las herramientas y conceptos básicos para el tratamiento de datos, la reconstrucción de las propiedades de las partículas a partir de los datos medidos, la simulación de los procesos de colisión y las herramientas utilizadas para el análisis conjunto de los datos proporcionados por ambos métodos.

### 4.1. Algoritmo *particle flow*

El algoritmo de reconstrucción del evento utilizado en el CMS es el algoritmo de flujo de partículas o **particle flow** [12]. La idea subyacente fundamental radica en intentar identificar todos los objetos resultantes del proceso y proporcionar una medida lo más fiable posible de su energía-momento y de características adicionales del proceso (por ejemplo, que un jet provenga de un quark b) a través de una combinación de la información proporcionada por todos los subdetectores del sistema.

De cara a reconstruir las trayectorias de las partículas tenemos tres señales experimentales fundamentales: los *hits* de las partículas en los detectores de muones, los *hits* en el *tracker* y las deposiciones de energía en el ECAL y el HCAL.

En primer lugar se intentan reconstruir trayectorias a nivel local (en cada subdetector) que puedan corresponderse a una única partícula.

En el *tracker* este concepto se traduce en un algoritmo de tipo iterativo. En primer lugar se seleccionan los

conjuntos de señales (semillas) en el *tracker* candidatos a formar una trayectoria partiendo de criterios estrictos. Posteriormente, y con los datos de estas primeras trayectorias, se intenta incorporar a cada una de ellas el resto de señales circundantes en el *tracker* comprobando si reconstruyen adecuadamente (con criterios más laxos) una trayectoria. Se repite el procedimiento eliminando en cada paso las señales ya asignadas a una trayectoria. De forma similar se pueden reconstruir trazas a través de los datos de los detectores de muones (si bien la precisión en estos casos es bastante menor).

Una aproximación relativamente similar se lleva a cabo con los datos de los calorímetros. En este caso se parten de las zonas del detector en las cuales la energía depositada alcanza un máximo que sobrepase cierto umbral mínimo. Dichas zonas actuarán de semillas para el proceso de creación de un *clúster*: se analizan las celdas adyacentes a la semilla y se incorporan al mismo aquellas con deposiciones elevadas de energía. Se terminan construyendo dichos *clúster* de forma iterativa. Notemos que en este caso los *clústers* se definen por separado, es decir, el algoritmo actúa independientemente en cada una de las secciones de ECAL y HCAL.

En general, la mayor parte de las partículas dejarán señales en más de uno de los detectores (véase Figura 5 (Izquierda)). Por tanto, una vez se han reconstruido las trayectorias individuales en cada uno de los detectores, el siguiente paso lógico consiste en relacionarlas entre sí a fin de intentar determinar las deposiciones de energías globales de las partículas a lo largo de todos ellos. Un *algoritmo de enlace* tiene como objetivo encontrar las relaciones entre las señales de cada detector de cara a reconstruir estas.

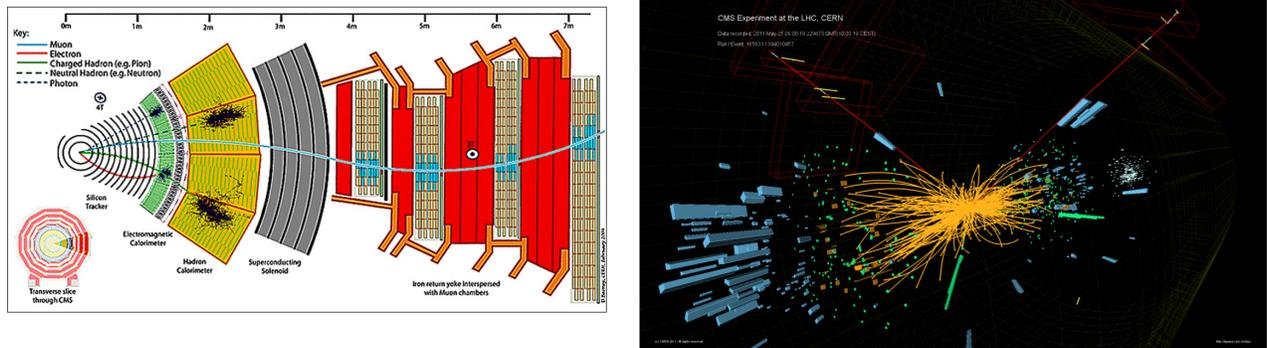


Figura 5: (Izquierda) Corte transversal del detector CMS incluyendo varias partículas detectadas y las deposiciones de energía que producen en los subdetectores. (Derecha) Representación simplificada de un evento real en el detector CMS, se observan con claridad las trayectorias de dos muones (en rojo).[11]

El funcionamiento de estos algoritmos es conceptualmente sencillo. Se consideran posibles pares de señales (sean trazas reconstruidas o *clústers*) que puedan ser compatibles entre sí desde un punto de vista cinemático y dinámico. Para cada uno de estos pares se estudian parámetros físicos que nos indiquen cuánto de probable es que pertenezcan a una misma trayectoria. Se designan como *emparejados* los pares de señales más próximos entre sí y, en los casos físicamente coherentes, se intentan acoplar a estos las señales más plausibles posibles.

Los criterios para establecer la verosimilitud en este emparejado son de tipo tanto físico como topológico. Por ejemplo, de un *clúster* en el ECAL podremos, suponiendo que se trata de una partícula cargada, intentar relacionarlo con las distintas trazas del *tracker*, siendo la mejor de todas ellas aquella cuya prolongación se encuentre más próxima al propio *clúster*. De la misma forma, deposiciones en el ECAL y en el HCAL se pueden relacionar en cuanto a lo próxima que esté espacialmente la señal en el HCAL respecto a la proyección desde el punto de choque de la superficie del ECAL.

El último paso en el proceso consiste en la identificación y reconstrucción completa de las propiedades de cada uno de los objetos presentes en el evento. Por sus propias características físicas, la dificultad varía enormemente de unos casos a otros. En las siguientes secciones presentamos una discusión básica de los algoritmos de reconstrucción, los objetos reconstruidos, y sus propiedades en orden de complejidad creciente.

#### 4.1.1. Reconstrucción del vértice principal

Como ya mencionamos, en el LHC los haces de protones que se hacen chocar contienen paquetes notablemente compactos de los mismos. Por este motivo, en la mayor parte de los choques se producen múltiples interacciones entre pares de protones que quedan registradas simultáneamente a nivel de detector. Este fenómeno, denominado **pile-up**, se debe solucionar para separar los datos correspondientes a nuestro suceso de interés o **evento** (generalmente una colisión inelástica) de las colisiones secundarias registradas en el detector.

Dado que, a escala del detector, las colisiones se producen de manera puntual, se puede solucionar este problema intentando encontrar los **vértices primarios** de la colisión (el punto donde se produce la interacción  $pp$ ) para cada uno de los choques participantes aislando así los datos de interés correspondientes al evento proveniente del vértice correspondiente.

Dada la alta resolución de las mediciones de las trayectorias en el *tracker*, se pueden prolongar los datos de las mismas hasta el eje del detector. Las regiones en las que coincidan múltiples trayectorias pasarán a ser consideradas vértices de interacción. A cada uno de ellos se asociarán por criterios de cercanía espacial los objetos reconstruidos asociados al mismo<sup>19</sup> para reconstruir el **evento** físico a estudiar.

#### 4.1.2. Reconstrucción de muones

Las partículas físicas más fácilmente identificables en el CMS son los muones,  $\mu$ . A las energías de trabajo actuales del LHC los muones interaccionan débilmente con los calorímetros, por lo que sus trayectorias quedan, relativamente, poco alteradas a su paso por los detectores, pierden una cantidad muy pequeña de energía en

---

<sup>19</sup>Es decir, se asociarán los objetos en términos de la proximidad de la extensión de las trazas al vértice de colisión.

estos subdetectores<sup>20</sup>.

Desde el punto de vista de la detección, esperamos observar señales de cada muón tanto en el *tracker* como en los propios detectores de muones. Notemos que, como los muones son partículas cargadas, su trayectoria se curva en sentidos opuestos en ambos sistemas al encontrarse el primero en el interior del solenoide y el segundo en el exterior. Teniendo en cuenta este fenómeno, es posible reconstruir las trayectorias de los muones combinando las señales en el *tracker* y en los detectores de muones.

Como, en principio, las trazas en el sistema de muones son más escasas, se parte de una de estas y se prolonga *hacia atrás en el tiempo* hasta alcanzar el *tracker*. En este punto se buscan las trayectorias más cercanas tomando como medida de proximidad en términos del espacio  $(\eta, \phi)$  donde la distancia considerada es  $d = \sqrt{\Delta(\phi)^2 + \Delta(\eta)^2}$ .

La trayectoria final del muón, que nos permitirá medir su  $p_T$  en términos de su curvatura, se reconstruye finalmente por interpolación con los datos de los detectores de muones y de la traza más próxima. El momento longitudinal  $p_Z$  se puede determinar (y con él la energía) con facilidad conocida la trayectoria, el momento transverso y la intensidad del campo magnético.

Ahora bien, nos interesaría poder distinguir entre los muones que se producen acompañando jets de partículas (en sucesos de desintegración de quarks, por ejemplo) y los que se dan *aislados* (por ejemplo, en procesos de desintegración leptónica de bosones W o Z). Para establecer esta diferencia, se utilizan de nuevo conos en el espacio  $(\eta, \phi)$  para definir la proximidad. Estableciendo un límite máximo  $\Delta(R_0)$ , diremos que un objeto (un jet o una partícula) está en el cono del muón si las coordenadas relativas entre ambos cumplen que  $\sqrt{\Delta(\phi)^2 + \Delta(\eta)^2} \leq \Delta(R_0)$ . Se define el aislamiento del muón  $I$  como la suma de los momentos (o energías) transversas de todos los objetos en dicho cono. Se define asimismo una cantidad relativa  $I_{rel}$  a partir del cociente por el momento transverso del propio muón:

$$I = \sum_{\Delta(R_i) < \Delta(R_0)} |p_T(i)| \quad , \quad I_{rel} = \frac{I}{|p_T(\mu)|}$$

Esperaremos que muones *aislados* tengan valores bajos de ambas variables. Los muones aislados son una señal de gran importancia, pues nos indican la presencia de procesos que involucran, usualmente, bosones  $W$  o  $Z$ , relativamente típicos en procesos de *nueva física*.

Una vez reconstruída adecuadamente la trayectoria del muón, se estima la contribución del mismo en ECAL y HCAL para eliminar las posibles señales residuales que haya podido dejar a su paso por los mismos.

---

<sup>20</sup>Un problema importante de futuro es el hecho de que, a medida que aumente la energía de trabajo del LHC, muones más energéticos se producirán cuya pérdida de energía mediante procesos radiativos será notablemente superior lo que puede afectar a estas medidas.

### 4.1.3. Reconstrucción de electrones

De la misma forma que los muones, los electrones pueden utilizarse como señales básicas de procesos de interés, ya que se producen en una cantidad relativamente escasa de procesos en las colisiones. A diferencia de los muones, los electrones sí que son absorbidos casi por completo en el ECAL, por lo que la señal que dejan en los subdetectores está compuesta habitualmente por la traza en el *tracker* y la energía del ECAL.

En su interacción con la materia, a las energías de trabajo del LHC, la pérdida de energía por procesos radiativos de los electrones en la materia (*Bremsstrahlung*) es lo suficientemente alta como para que se puedan alterar significativamente las trayectorias que detectamos a través de los procesos de ionización. El resultado directo es que la medición de la energía-momento de los electrones resulta más complicada y está abierta a más factores de error que en el caso muónico. Adicionalmente, los propios fotones emitidos por *Bremsstrahlung* pueden ser detectados por separado en el ECAL, lo que hay que tener en cuenta a la hora de la reconstrucción.

La reconstrucción de la trayectoria es, por lo demás, muy similar al trabajo realizado para los muones sustituyendo las trazas de los detectores de muones por los *clústers* formados en el ECAL. Notemos que, en este caso, no todos los *clúster* se corresponderán con electrones pues también los fotones se pueden absorber en este subdetector.

Por último destaquemos que, como en el caso de los muones, se define una variable de aislamiento  $I$  (de forma similar) que nos da una idea numérica de lo *alejado de otras partículas* que se ha producido el propio electrón.

### 4.1.4. Reconstrucción de jets

Una vez identificadas las partículas leptónicas, nos queda por interpretar la forma en la cual se detecta la presencia de quarks en el proceso. Como vimos en la Sección 2.1, el hecho de que los gluones tengan carga de color impide que se observen *quarks o gluones desnudos*.

Desde el punto de vista práctico, cuando se produce un quark o gluón solitario en los procesos de colisión, este se **hadroniza**, es decir, se produce una emisión de partículas (por un proceso de autointeracción) tipo quark y una recombinación de estas con el quark original hasta dar lugar a múltiples hadrones. Experimentalmente esto se traduce en la observación de agrupaciones en forma de *chorros de partículas* a las que denominamos genéricamente **jets**.

Si bien existen diversas formas de reconstruir los jets, indicamos a continuación un breve proceso de reconstrucción posible. Partimos de trazas observadas en el *tracker* (que en este hipotético caso se corresponderían con hadrones cargados). Extrapolando estas trazas hasta el ECAL<sup>21</sup> y el HCAL se pueden determinar las posiciones

---

<sup>21</sup>Una parte importante de la energía de los jets se debe a mesones  $\pi^0$  cuyo principal modo de desintegración produce  $2\gamma$  que se absorben en el ECAL.

en las cuales esperamos deposiciones de energía por parte del hadrón cargado correspondiente. De los clústers cercanos en el ECAL y el HCAL se sustrae la energía correspondiente a las deposiciones esperadas del hadrón (en términos del  $p_T$  medido para el mismo), en caso de que el exceso de energía sea significativo asumiremos la aparición de fotones (si el exceso de energía se da solo en el ECAL) o hadrones neutros (si se observa exceso de energía en el HCAL también).

Las señales que resten en el HCAL (no identificadas hasta el momento) se asociarán con hadrones neutros mientras que las que queden en el ECAL se asociarán a fotones adicionales.

Una vez reconstruidas las partículas, se definen los jets como los conjuntos de partículas en una sección cónica del espacio determinada (según la distancia angular que definimos anteriormente).

Si bien asumimos los jets en un caso bastante general, nos resultaría de especial interés poder determinar el sabor de la partícula que los origina (así como determinamos la diferencia entre un muón y un electrón). Esto no es posible en todos los casos pero sí que se puede lograr en el caso de jets originados por quarks b o **b-jets**. Adicionalmente, dado que un quark t decae en  $Wb$  en una fracción mayoritaria de los casos, la determinación de estos b-jets resulta especialmente relevante en los procesos de *nueva física* que involucran al quark top.

Un b-jet incorporará inicialmente al menos un hadrón que tenga un quark b como uno de sus quarks de valencia. Dichos hadrones tienen una vida media relativamente alta (sin alcanzar la de los piones cargados  $\pi^\pm$ ) de forma que pueden desplazarse en el *tracker* una cantidad de espacio relativamente elevada antes de desintegrarse. El efecto de dicha desintegración será la aparición de un nuevo jet (centrado en el punto de desintegración), por lo que la signatura experimental será la aparición de trayectorias dentro del jet que apuntan a un vértice secundario en su interior. Además, dado que la masa del quark b es bastante superior a la del resto de quarks (excepto el t), los jets resultantes de su desintegración tendrán un  $p_T$  más elevado al añadir al mismo la energía proveniente de la masa del quark b.

Estos dos hechos se aprovechan en un algoritmo conocido como *Complex Secondary Vertex (CSV)* que permite calcular una variable discriminante de lo *probable que es que un jet provenga de un quark b*. Eligiendo diversos valores mínimos de dicha variable obtendremos condiciones más o menos estrictas para la aceptación de un b-jet.

#### 4.1.5. Energía faltante

Hasta el momento hemos descrito la forma de detectar y reconstruir muones, leptones, jets y fotones. Sin embargo, existen más partículas estables (a las escalas de tiempo y distancia del proceso de detección) en el SM (y en muchos de los modelos de *nueva física*) que deberíamos poder obtener en nuestras colisiones. El ejemplo más inmediato es el caso de los neutrinos. Estos no interactúan con los diferentes subdetectores: constituyen

el caso más usual de **partículas invisibles**.

El procedimiento principal por el que podemos determinar la presencia de partículas invisibles se basa en el principio físico básico de la conservación del momento. Como los haces que colisionan entre sí tienen<sup>22</sup> un momento completamente longitudinal, el momento transverso inicial del sistema de colisión  $p - p$  será nulo. Dicho momento se conservará durante todo el proceso, por lo que en los objetos finales del mismo la suma total de momentos transversos será nula. Tendremos que:

$$0 = \sigma_\mu p_T^\mu + \sigma_e p_T^e + \sigma_{jets} p_T^{jets} + \sigma_\gamma p_T^\gamma + \sigma_{inv} p_T^{inv}$$

Donde cada una de las sumas se extiende a todos los objetos del tipo dado reconstruidos para el evento. Se define la energía faltante,  $\cancel{E}_T$  o  $E_T^{miss}$ , como la suma de los momentos transversos de las partículas invisibles de forma que podemos calcularlo como:

$$\cancel{E}_T = - \sum_{vis} p_T^{vis}$$

Donde  $i$  varía sobre todas las partículas reconstruidas en el evento.

Una reconstrucción completa (simplificada) se puede observar en la Figura 5 (Izquierda). El vértice principal (en la zona central) y dos muones (en rojo) se observan con claridad.

## 4.2. Simulaciones de Montecarlo

Como ya mencionamos, las simulaciones del proceso de colisión son fundamentales para nuestra comprensión del mismo. Su utilidad se encuentra presente a la hora de medir eficiencias (por comparación de datos con simulación de procesos bien conocidos) y, aspecto central en la búsqueda de *nueva física*, evaluar la cantidad de procesos de fondo conocidos que aparecerán en nuestra señal experimental.

La idea de una simulación de Montecarlo (MC) es intentar reproducir de la forma más exacta posible el comportamiento de un proceso de interés físico así como la interacción que este presenta con los distintos elementos del detector. Una buena simulación reproducirá adecuadamente el comportamiento experimental a través de una modelización adecuada de los aspectos estocásticos del mismo (canal de desintegración, tiempo de vida, etc.).

La simulación del evento global se realiza en varias etapas separadas, con una clara correspondencia física, para cada una de las cuales existe *software* específicamente diseñado. Globalmente queremos emular tanto el proceso de interés físico estudiado como su interacción con el detector, a fin de reproducir de la manera más fidedigna posible los datos obtenidos experimentalmente.

---

<sup>22</sup>Realmente se inclinan ligeramente los haces para producir la colisión y esto no es más que una aproximación.

Recordemos que un protón contiene una serie de *componentes* (partones): quarks de valencia (los quarks que usualmente se dicen que *forman* el protón), gluones y quarks *del mar* (pares quark-antiquark).

La colisión *principal* se da entre un partón de uno de los protones y otro partón del otro, cada uno de los con una fracción del momento del protón original, este proceso es el **scattering fuerte**. Al mismo tiempo, el resto de partones de cada protón puede dar lugar a procesos secundarios adicionales, dando lugar al **underlying event**. Adicionalmente, toda partícula cargada fuertemente puede radiar un gluón que puede, a su vez, emitir gluones adicionales dando lugar a **cascadas de partones**.

Una vez los procesos anteriores han alcanzado regímenes de energías más bajas, se produce la **hadronización** de las partículas libres con carga fuerte. Simultáneamente, las partículas generadas con vidas medias cortas se pueden **desintegrar** dando lugar a partículas adicionales en el detector.

Entre los procesos que involucran la interacción fuerte podemos establecer una clara diferencia: los sucesos que se producen a altas energías (régimen perturbativo) y los que ocurren a bajas energías (régimen no perturbativo). Para el primer caso (principalmente se produce en el *scattering fuerte*) existen diversos paquetes de simulación con precisiones etiquetadas como *Leading Order* (primer orden en los diagramas de Feynman considerados), *Next to Leading Order* (incluyendo el segundo orden), etc. Ejemplos de estos simuladores son MADGRAPH, POWHEG, etc. En el segundo régimen, que incorpora procesos no renormalizables, se encuentran procesos del *underlying event*, de *hadronización* y de *cascadas de partones*. El simulador más utilizado en estos casos es PYTHIA.

Una vez se obtienen los 4-momentos de las partículas participantes en el evento, se incorpora a las simulaciones la acción del propio detector. Es decir, introducimos las partículas creadas en un entorno *ideal* en un modelo que reproduce la acción del detector CMS de cara a obtener un equivalente a las medidas que obtendríamos en el mismo. Dos aproximaciones distintas se pueden utilizar en este caso.

En la *simulación completa* (*full-sim*) se reproduce completamente la estructura del detector y todas las interacciones posibles con las componentes del mismo. Requiere de un gasto computacional importante y se lleva a cabo con GEANT4.

La *simulación rápida* (*fast-sim*) se utiliza como una simplificación veloz del caso anterior. En este caso simplemente se introducen en las variables generadas por la simulación variaciones estadísticas globales de acuerdo con el comportamiento esperado por su interacción con la materia.

Tras estos procedimientos de generación y reconstrucción de los datos, obtenemos muestras simuladas (o

muestras de Montecarlo) cuyo comportamiento es (si estamos usando un modelo adecuado) similar al de los datos reales.

### 4.3. Herramientas informáticas

El análisis de datos que hacemos en este trabajo resulta relativamente pesado desde el punto de vista computacional. Adicionalmente, la necesidad de repetir en diversas condiciones análisis bastante complejos nos impone la necesidad de utilizar de forma eficiente nuestros recursos de cálculo. Por estos motivos, elegimos C++ como lenguaje de programación básico sobre el que elaborar nuestros programas de análisis.

Adicionalmente, utilizaremos las distintas librerías de ROOT que añaden a la potencia de C++ objetos predefinidos de utilidad para la física de altas energías, así como librerías específicas para el graficado de nuestros resultados. Dentro de ROOT, destacar el uso principal de TMVA [13] como librería dedicada al análisis multivariante. Esta contiene una implementación de alta eficiencia para la utilización de los métodos multivariante usados en el trabajo.

Por último, nos valdremos de Python como lenguaje auxiliar a la hora de automatizar múltiples procesos.

## 5. Procesos de interés y muestras

Partiendo de la introducción en las secciones 2.3.1 y 2.4 de los procesos a estudiar, incorporemos ahora los conceptos expuestos en los capítulos anteriores para traducir la física propia del proceso en las correspondientes señales que esperamos obtener experimentalmente. De la misma forma, una vez comprendida la señal buscada, estudiemos cuales son los principales procesos *de fondo* que comparten propiedades con nuestros eventos de *nueva física*.

Identificados tanto los procesos de señal como de fondo, introduciremos las muestras de trabajo principales de nuestro estudio que modelizan dichos procesos.

Estudiando la física de estos procesos, definiremos finalmente los criterios físicos de preselección que impondremos a los eventos de las mismas de cara a considerarlos en nuestro estudio final.

### 5.1. Procesos MSSM

En la Sección 2.3.1 introdujimos previamente el concepto de supersimetría (SUSY) y modelos como el MSSM. Adicionalmente, estructuramos el proceso  $T2tt$ , que se corresponde con el modelo simplificado que estudiaremos en nuestro trabajo.

Recordemos que la señal correspondiente se produce mediante la forma:

$$pp \rightarrow t\bar{t} \rightarrow t\bar{t}\chi_1^0\chi_1^0$$

Usualmente, el quark  $t$  se desintegra en un par  $bW$  (es posible que de lugar a pares  $dW$  o  $cW$  pero los términos correspondientes de la matriz CKM<sup>23</sup> indican que son procesos muy poco usuales). Por tanto, el proceso mayoritario será de la forma:

$$pp \rightarrow bW^+\bar{b}W^-\chi_1^0\chi_1^0$$

Al considerar las posibles desintegraciones de un bosón  $W$  tenemos dos alternativas: desintegración hadrónica,  $W \rightarrow qq'$ , (aproximadamente en 2/3 de los casos) o desintegración leptónica  $W \rightarrow l\nu$  (aproximadamente en 1/3 de los casos). A nivel de detector, los leptones que realmente podemos identificar con claridad son los muones y los electrones, por lo que el canal leptónico quedará reducido para nosotros a una proporción algo superior a 2/9 del total<sup>24</sup>. Combinando las posibilidades para los dos quarks  $W^+$  y  $W^-$  obtenemos tres posibles grupos de estados finales totales.

En el caso hadrónico, ambos decaen en pares de quarks:  $W^+W^- \rightarrow qq'q''q'''$ . Este canal es notablemente

<sup>23</sup>Cabibbo-Kobayashi-Maskawa. Indica la *Branching fraction* de las desintegraciones de cada quark a cada otro quark posible.

<sup>24</sup>Ya que el leptón  $\tau$  se puede desintegrar tanto a leptones como a hadrones.

frecuente<sup>25</sup>, pero su signatura experimental (4 jets) es también bastante común en los procesos típicos en las colisiones  $p-p$  por lo que su separación de los procesos de fondo puede resultar relativamente complicada.

El canal semileptónico se da en el caso de que los dos bosones se desintegren por caminos distintos:  $W^+W^- \rightarrow qq'\nu$ . De nuevo, este canal se da en una cantidad elevada de los casos<sup>26</sup>. Adicionalmente, la signatura experimental incorpora un leptón en el estado final, lo que en principio es relativamente raro en los procesos que se producen en el LHC.

Por último, el canal dileptónico se da cuando ambos bosones sufren desintegración leptónica:  $W^+W^- \rightarrow \nu\nu'\nu'$ . Aunque este es el canal con menos estadística de los tres<sup>27</sup>, presenta en su signatura final dos leptones que, además, tienen signos opuestos. Si bien se pueden producir otros procesos con esta signatura, son lo suficientemente escasos como para que este canal presente un estudio más simple que los otros dos. Por este motivo elegimos este canal para nuestro estudio posterior.

Por tanto, los procesos que esperamos observar serán de la forma:

$$pp \rightarrow t\bar{t} \rightarrow t\bar{t}\chi_1^0\chi_1^0 \rightarrow bW^+\bar{b}W^-\chi_1^0\chi_1^0 \rightarrow b\bar{b}l\nu\nu'\chi_1^0\chi_1^0$$

Se puede observar en la Figura 2 (Derecha) el esquema del proceso correspondiente. De las partículas en el estado final, tanto los leptones (en nuestro caso muones y electrones) como los neutrinos son estables a nivel de detector. Adicionalmente, asumimos que los neutralinos  $\chi_1^0$  lo son también al ser la LSP de nuestro modelo que respeta la paridad R (Sección 2.3.1). Los quarks b por su parte se desintegrarán dando lugar a dos b-jets.

Ya que ni neutrinos ni neutralinos interaccionan electromagnéticamente, formarán parte de la energía faltante  $\cancel{E}_T$  en la señal final. La signatura experimental final constará, entonces, de dos leptones con carga de distinto signo, dos b-jets y una  $\cancel{E}_T$  muy elevada. Desde este punto de vista,  $\cancel{E}_T$  es la variable fundamental en cuanto a que el proceso es virtualmente idéntico al canal dileptónico de  $t\bar{t}$  exceptuando una menor cantidad de esta magnitud.

Mencionemos que, aunque usualmente se consideren por separado, tomaremos simultáneamente todos los canales dileptónicos ( $ee$ ,  $\mu e$  y  $\mu\mu$ ) de cara a obtener la mayor estadística posible. El principal motivo para esta decisión es el hecho de que los procesos de señal a estudiar tienen muy baja sección eficaz (son muy escasos).

En cuanto a las muestras de señal utilizadas, centraremos nuestro estudio en el punto del espacio de parámetros correspondiente a  $m_{\tilde{t}} = 850\text{GeV}$ ,  $m_{\chi_1^0} = 100\text{GeV}$ , puntos situados fuera de la zona de exclusión actual (véase la Figura 6). De forma complementaria, estudiaremos un punto situado en una región con una dinámica bastante distinta,  $m_{\tilde{t}} = 600\text{GeV}$ ,  $m_{\chi_1^0} = 300\text{GeV}$ , de cara a comparar los resultados de nuestros análisis entre

<sup>25</sup>Contando los  $\tau$ ,  $4/9 = 2/3 \times 2/3$  del total de eventos con  $W^+W^-$  se desintegran según este camino.

<sup>26</sup>La combinatoria de casos (el hecho de que uno u otro de los dos bosones sea el que sigue cada canal), hace que se produzca de nuevo en una proporción de  $4/9 = 2 \times 2/3 \times 1/3$ .

<sup>27</sup>Únicamente el restante  $1/9$  de los eventos ocurren en este canal.

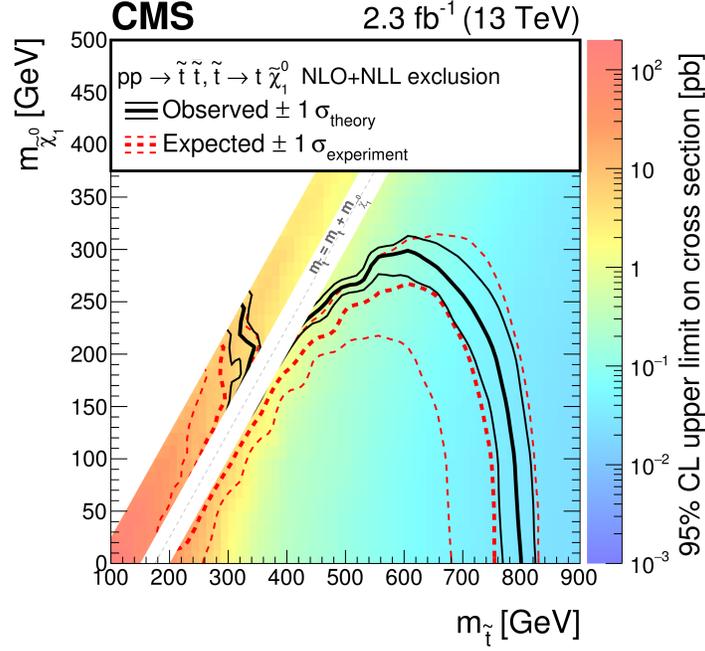


Figura 6: Límites de exclusión establecidos en el espacio de parámetros  $m_{\tilde{t}} - m_{\chi_1^0}$ . Los puntos por debajo de la línea negra se encuentran ya excluidos en la posibilidad de descubrimiento de procesos SUSY en el canal  $T2tt$ . [14]

un punto y otro. Sabiendo que la sección eficaz es en el primer caso  $\sigma_{850-100} = 0,01896pb$  y en el segundo  $\sigma_{600-300} = 0,01896pb$ , esperamos observar cómo afecta este hecho fundamental al poder predictivo de nuestros métodos de análisis.

Adicionalmente, de cara al entrenamiento de los métodos multivariante, enriqueceremos nuestras muestras de señal con puntos adicionales situados en la *diagonal* del plano de parámetros  $m_{\tilde{t}} - m_{\chi_1^0} = 750GeV$  y  $m_{\tilde{t}} - m_{\chi_1^0} = 300GeV$ . Esto se debe a que asumimos que la cinemática del proceso depende fuertemente de la diferencia de energías entre ambas y que, por tanto, el comportamiento de las diferentes variables estudiadas será relativamente similar en estos puntos. En todos los casos las muestras son simuladas para colisiones  $pp$  suponiendo una energía en el centro de masas  $\sqrt{s} = 13TeV$ .

## 5.2. Procesos de EFT

Siguiendo el desarrollo de la Sección 2.4, podemos contemplar una EFT en la que se presente un acoplo de la forma  $t\bar{t} \rightarrow \chi\chi$  donde  $\chi$  representa la partícula de materia oscura.

En concreto, estudiaremos procesos de la forma  $pp \rightarrow t\bar{t}\chi\chi$ , donde el estado final quede caracterizado por la presencia de dos partículas de DM junto con un par  $t\bar{t}$ .

Considerando que, o bien las partículas de DM son estables, o bien se desintegran en otras partículas de DM invisibles al detector, estas contribuirán a la señal final únicamente en la forma de  $\cancel{E}_T$ . Nuestro proceso queda,

entonces, reducido a un caso muy similar al que obteníamos en el modelo  $T2tt$ . Como entonces, nos centraremos en el caso dileptónico, que responderá a la forma:

$$pp \rightarrow t\bar{t}\chi\chi \rightarrow bW^+\bar{b}W^-\chi\chi \rightarrow b\bar{b}l\nu l'\nu'\chi\chi$$

Se puede observar en la Figura 2 (Izquierda) el diagrama correspondiente. La signatura experimental de este proceso es idéntica a la del caso anterior: dos leptones de carga opuesta, dos b-jets y  $\cancel{E}_T$ . De nuevo  $\cancel{E}_T$  será una variable fundamental en cuanto a que la cantidad adicional de energía faltante es lo que representa la presencia de partículas de DM en nuestro proceso. Continuando con nuestro estudio, realizaremos un análisis inclusivo al tratar simultáneamente todos los canales dileptónicos posibles.

De nuevo, nos restringiremos a un caso concreto del espacio de parámetros, estudiando una muestra MC con  $m_\chi = 1000 GeV$ . En todos los casos las muestras son simuladas para colisiones  $pp$  suponiendo una energía en el centro de masas  $\sqrt{s} = 13 TeV$ . Teniendo en cuenta las condiciones típicas impuestas sobre la WIMP, tomamos una sección eficaz tipo de  $\sigma_{DM} = 1 pb$ .

### 5.3. Procesos de fondo

Dada la signatura experimental esperada tanto para  $T2tt$  como para  $DM$ , podemos considerar los posibles procesos del SM con una sección eficaz apreciable y que contribuyan al fondo total en nuestras medidas. A continuación presentamos un breve resumen de dichos procesos que organizamos en varios grupos con propiedades similares.

- **Proceso  $t\bar{t}$ :** En el canal dileptónico, la desintegración de pares  $t\bar{t}$  tiene una signatura experimental idéntica a las señales consideradas sin más que eliminar las partículas correspondientes a *nueva física* (que se corresponden con  $\cancel{E}_T$ ).
- **Procesos  $ZZ$ :** Procesos con dos bosones  $Z$  pueden dar lugar a estados finales muy similares al observado. Se pueden recuperar los dos leptones sin más que considerar la desintegración leptónica de uno de los dos y los dos b-jets considerando la desintegración  $Z \rightarrow b\bar{b}$  del otro. Los incorporamos dentro del conjunto  $VV$  (dos bosones vectoriales masivos).
- **Procesos  $ZW$ :** De nuevo, los leptones pueden aparecer por causa de la desintegración leptónica del  $Z$  o incluso una desintegración leptónica del  $W$  y una mala identificación que de lugar a un *leptón falso*. Lo mismo puede ocurrir con los 2 b-jets. Los incorporamos dentro del conjunto  $VV$ .
- **Procesos  $WW$ :** De forma muy similar al  $ZW$ , es posible que den lugar a las diferentes partículas en la

signatura. Sin embargo, para observar todas simultáneamente tendríamos que identificar incorrectamente varias de ellas por lo que esperamos que su contribución total sea menor que la de  $WZ$ . Los incorporamos dentro del conjunto  $VV$ .

- **Procesos  $W + Jets$ :** En estos procesos, la desintegración leptónica del  $W$  puede dar lugar a la aparición de  $\cancel{E}_T$  y un leptón. Los jets proporcionan causas de la aparición de b-jets y un falsamente identificado leptón aislado completaría la signatura. Los incorporamos dentro del conjunto  $W + Jets$ . Son procesos con una sección eficaz notablemente alta por lo que, aunque nuestra preselección (Sección 5.5) elimine la gran mayoría de su contribución, es muy probable que una proporción elevada de ellos *pase* los cortes de preselección.
- **Procesos  $tW$ :** Alcanzar la signatura experimental es posible si  $W$  y  $t$  decaen leptónicamente. Únicamente sería necesario la aparición de un b-jet adicional. Los incorporamos al conjunto  $TW$ .
- **Procesos  $ttW$ :** Procesos en los que dos de las tres partículas decaen leptónicamente contribuyen directamente con la signatura experimental completa al proceso. Adicionalmente, la combinatoria del problema hace que una mayor proporción de los canales posibles se corresponda con productos finales compatibles con nuestra señal. Los incluimos en  $ttV$  (par  $t\bar{t}$  más bosón vectorial masivo).
- **Procesos  $ttZ$ :** Al igual que el caso anterior, nos basta con que, o bien el  $Z$ , o bien los dos quarks  $t$  decaigan leptónicamente para que estén presentes en nuestro detector los elementos de la señal. De nuevo, gran cantidad de los eventos en los que se generan estas partículas presentan las partículas de la señal en su desintegración. Los incluimos en  $ttV$ .
- **Procesos Drell-Yan:** Denominamos genéricamente Drell-Yan a procesos en los cuales aparecen bosones  $Z/\gamma^*$  producto de la colisión  $qq'$ . En estos casos,  $Z/\gamma^*$  puede desintegrarse en un par leptónico, lo que unido a la identificación de jets del *underlying event* como b-jets puede generar las partículas presentes en nuestra señal. Los incluimos dentro de  $DY$ . Son procesos con una sección eficaz notablemente elevada por lo que, aunque nuestra preselección elimine la gran mayoría de su contribución, es muy probable que una proporción elevada de ellos *pase* los cortes de preselección.

En todos los casos las muestras utilizadas fueron simuladas para colisiones  $pp$  suponiendo una energía en el centro de masas  $\sqrt{s} = 13TeV$ . Para hacernos una idea de la importancia que puede tener como fondo cada uno de estos procesos en nuestro estudio es lógico estudiar, además de la similitud de la signatura experimental, la sección eficaz de cada uno de ellos. Como se observa en la Figura 7, los procesos con más sección eficaz de los considerados serían  $W + jets$  ( $W$  en la figura) o Drell-Yan ( $Z$ ). Sin embargo, estos dos procesos presentan una signatura experimental muy distinta a la de la señal y, por tanto, es esperable que se vean reducidos fuertemente

por nuestra preselección. El fondo dominante será  $t\bar{t}$ , un proceso con una enorme sección eficaz y una signatura casi idéntica a la señal. Los procesos  $ttV$  ( $ttZ$  y  $ttW$ ) aunque irreducibles como fondo (al igual que  $t\bar{t}$ ), tienen una sección eficaz bastante menor por lo que su papel será relativamente menos relevante.

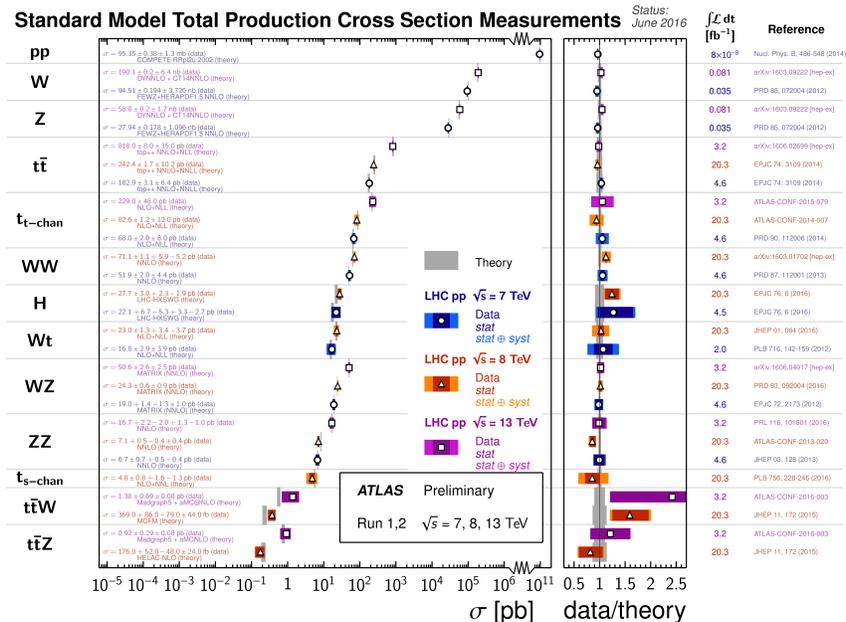


Figura 7: Secciones eficaces de procesos del SM para las distintas energías en el CM obtenidas en el LHC. Destaquemos que los procesos de señal considerados, que no aparecen en la figura, se encontrarían situados incluso por debajo de los últimos mostrados. [15]

Por otra parte, si recordamos las cifras presentadas en las dos secciones anteriores, todos los procesos de fondo tienen una sección eficaz similar o superior a cualquiera de nuestros procesos de señal. No sólo eso, sino que el proceso  $t\bar{t}$  tiene una sección eficaz entre 3 y 5 órdenes superior a la señal y una signatura experimental notablemente similar. Es decir, por cada evento de señal estaremos produciendo entre  $10^3$  (DM) y  $10^5$  ( $T2tt$  en el punto con  $m_{\tilde{t}} = 850 GeV$ ) eventos de  $t\bar{t}$ . Queda claro, entonces, que nuestra señal es prácticamente minúscula respecto al fondo por lo que es necesario un ajuste increíblemente fino de nuestros métodos de separación.

#### 5.4. Muestras de datos

Por último, disponemos de muestras de datos correspondientes a mediciones experimentales en el experimento CMS durante el año 2015 (primera parte de la *second run*). En total se corresponden con una luminosidad integrada de  $1,28 fb^{-1}$  a una energía en el centro de masas  $\sqrt{s} = 13 TeV$ . Si bien disponemos de muestras separadas por canales, como en casos anteriores realizaremos un estudio inclusivo que incluya datos  $ee$ ,  $\mu\mu$  y  $e\mu$  simultáneamente.

## 5.5. Preselección de eventos

De cara a simplificar nuestro estudio posterior, exigiremos una serie de requisitos a nuestros eventos para incluirlos en nuestras muestras de análisis. Es decir, reduciremos las muestras originales de fondo, datos y señal imponiendo condiciones que deban de cumplirse. Elegimos dichas condiciones con el criterio de buscar que restrinjan procesos distintos al considerado en la señal. Utilizaremos valores de corte previamente estudiados de cara a maximizar la eficiencia en la selección perdiendo la menor cantidad de *eventos reales* posibles.

En primer lugar, esperamos encontrar dos leptones en nuestro evento que además provienen de bosones  $W$ , con lo que prácticamente se producen en el vértice principal y serán bastante energéticos (ya que  $m_W \gg m_e$ ). Adicionalmente deberían encontrarse aislados en el detector. Por tanto exigimos exactamente dos leptones *buenos* en el evento y de signo opuesto. Los criterios para considerar un leptón *bueno* varían en función de su sabor.

En el caso de los muones, exigiremos las siguientes condiciones. En primer lugar, al tratarse de partículas energéticas pediremos  $p_T > 15\text{GeV}$ . La condición de aislamiento se traduce en exigir  $I_{rel} < 0,12$  (aislado, véase 4.1.2). Por último, el hecho de que se generen prácticamente en el vértice de interacción se puede exigir en término de sus parámetros de impacto<sup>28</sup> tomando  $|dxy| < 0,2\text{cm}$  y  $|dz| < 0,5\text{cm}$ . Finalmente, de cara a tener muones identificados con una resolución decente, exigiremos  $|\eta| < 2,4$  (detectado en *barrel* o *endcap*).

Por motivos similares, exigimos para electrones  $p_T > 15\text{GeV}$ ,  $I_{rel} < 0,15$ ,  $|dxy| < 0,02\text{cm}$ ,  $|dz| < 0,10\text{cm}$  y  $|\eta| < 2,5$ .

Adicionalmente, en caso de que los dos leptones sean del mismo sabor, aplicaremos un corte adicional con la intención de reducir la contribución de todos los procesos que incorporan un bosón  $Z$ . Como sabemos, entre los canales principales de desintegración de este se encuentra  $Z \rightarrow \bar{l}l$ , lo que da lugar a un pico notablemente pronunciado en la distribución de masa invariante del par  $\bar{l}l$ . Intentaremos eliminar dicho pico (y por tanto gran parte de la contribución de procesos con un  $Z$ ) imponiendo que la masa invariante del par se encuentre fuera de dicho pico:  $|m_{ll} - m_Z| > 15\text{GeV}$ .

Además, impondremos requisitos físicos derivados de la aparición de 2 b-jets. Exigiremos la presencia de al menos 2 jets de los cuales, al menos uno de ellos tendrá que ser un b-jet (exigimos únicamente un b-jet de cara a no perder una cantidad demasiado grande de señal en este corte).

En cuanto a cada uno de los jets, exigiremos que sea energético (los quarks b proceden de un quark t con una masa muy elevada, por lo que es de esperar que tengan un alto  $p_T$ ):  $p_T > 30\text{GeV}$ . Al igual que en el resto de partículas, exigiremos  $|\eta| < 2,4$  para tener una buena resolución en su medida. El valor del discriminante CSV (véase la Sección 4.1.4) exigido para calificar un jet de b-jet será 0,89.

<sup>28</sup>La distancia mínima entre la trayectoria de la partícula y el vértice central en los ejes longitudinal,  $dz$ , y transversal,  $dxy$ .



## 6. Métodos de análisis

Si bien hemos introducido ya los procesos físicos a estudiar, las muestras de datos utilizadas para ello y con el procedimiento seguido para obtenerlas, nos queda por presentar cómo vamos a analizar e interpretar dichos datos. En el siguiente capítulo buscamos introducir los procedimientos de análisis utilizados durante el trabajo que aplicaremos posteriormente.

La base de todo análisis en busca de procesos de *nueva física* pasa por cuantificar la frecuencia de aparición de los procesos de fondo y señal en cierta región (lo que coloquialmente se denomina *contar eventos*) y comparar dichos valores (esperados) con los que se obtienen experimentalmente (medidos). Cuando los datos medidos excedan a los datos de fondo esperados en una cantidad significativa (respecto a la precisión del experimento) y coincidan adecuadamente con el valor esperado de señal sumado al fondo, diremos que hemos encontrado un evento de *nueva física*.

### 6.1. Cortes secuenciales

La forma más sencilla de definir las regiones en las que realizamos nuestro estudio pasa por definir zonas del espacio de parámetros dinámicos o cinemáticos del proceso acotando los propios valores de estos mediante lo que se denominan **cortes**.

A partir del comportamiento físico de los procesos de señal y de fondo, se pueden postular valores justificados para el corte que disminuyan lo más posible la presencia de eventos de fondo,  $B$ , respecto a los sucesos de señal,  $S$ .

Así, por ejemplo, para discriminar entre un proceso de fondo  $Z \rightarrow \bar{l}l$  respecto a una señal  $Z \rightarrow \nu\nu$  podríamos exigir que  $\cancel{E}_T > 0$ , ya que, idealmente, el primero no contiene partículas *invisibles* mientras que en el segundo tendremos un par de neutrinos.

En vista de que nuestro objetivo es encontrar una región donde  $S$  sea bastante mayor que  $B$ , es usual definir cantidades **discriminantes** como  $S/B$ ,  $S/(S+B)$ ,  $S/\sqrt{S+B}$ , etc. En general, una región de corte será *mejor* que otra (esperaremos que permita con mayor facilidad el descubrimiento o exclusión de procesos de señal) cuanto mayor sea el valor de estos discriminantes.

Alternativamente, podemos definir la **eficiencia** del corte en la señal y el fondo:  $\epsilon_s$  y  $\epsilon_b$ . La eficiencia respecto a un proceso se definirá como la proporción de eventos del mismo que cumplen la condición de corte. Un corte será *mejor* cuanto mayor sea su eficiencia de señal y menor sea su eficiencia de fondo.

Como usualmente un único corte en una variable no da valores elevados de  $S/B$  (sobre todo en el caso de procesos de señal de muy baja sección eficaz), se suelen realizar varios cortes sucesivos o *secuenciales* para definir la región de trabajo. Diremos que este tipo de análisis, que involucran en cada paso una única variable

característica del proceso, son de tipo monovariante o univariante.

## 6.2. *Boosted Decision Trees*

El procedimiento de cortes secuenciales resulta especialmente potente en cuanto a que cada uno de los cortes presenta una alta eficiencia en la discriminación  $S/B$ . Sin embargo se trata de un procedimiento ciertamente limitado en cuanto a que estamos desechando directamente mucha información sobre el evento en términos de otras variables de menor importancia y las propias correlaciones entre todos los parámetros físicos. Hablamos de un **análisis multivariante** cuando estemos considerando en cada paso del análisis simultáneamente varias de las variables específicas de cada evento.

En nuestro caso particular, nos centraremos en los *Boosted Decision Trees* (**BDT**) como método multivariante para nuestro análisis de datos. La idea central de un BDT es sustituir la pequeña cantidad de elecciones de alta eficiencia dada por los cortes secuenciales con un número mucho mayor de selecciones multivariable de baja eficiencia.

Dado que será nuestro método de análisis fundamental, es necesario estudiar el funcionamiento básico de un BDT, así como los parámetros que lo caracterizan, para tener una idea clara de la estructura del proceso.

### 6.2.1. **Árbol de decisión**

La unidad fundamental en el trabajo del BDT es el **árbol de decisión** (DT). Un árbol de decisión no es más que una manera *formal* de separar un conjunto de objetos en varios subconjuntos distintos a través de decisiones binarias sobre las propiedades de sus elementos.

Supongamos que tenemos inicialmente un conjunto  $X$  tal que los elementos del mismo tienen un total de  $n$  propiedades distintas a las que denominaremos  $\{P_1, \dots, P_n\}$ . Así, un elemento  $x$  de  $X$  llevará asociada cierta tupla de elementos  $x = (x_1, \dots, x_n)$  que identifican dichas propiedades.

Llamamos al conjunto  $X$  **nivel 0** del árbol. Consideremos ahora una partición de  $X$  **binaria**<sup>29</sup> en términos de una de sus propiedades  $P_i$ . Es decir, tomamos dos condiciones  $A_1$  y  $A_2$  que puede cumplir  $P_i$ , excluyentes entre sí y que incluyan todos los posibles valores de  $P_i$ . Esto nos permite *ramificar* el primer nivel del árbol, dividir  $X$  en dos subconjuntos disjuntos  $X_1 = \{x \in X \mid x_i \in A_1\}$  y  $X_2 = \{x \in X \mid x_i \in A_2\}$ . Los subconjuntos  $X_1$  y  $X_2$  constituyen el nivel 1 del árbol y se denominan **nodos**.

Por iteración de este proceso, podemos aplicar tantas decisiones binarias como quedamos en cada uno de los subconjuntos resultantes a fin de ir generando los siguientes niveles del árbol. Denominamos **profundidad** al nivel máximo que alcanza el árbol.

---

<sup>29</sup>Técnicamente, podemos considerar decisiones múltiples en cada nivel, pero en este caso nos restringiremos únicamente a decisiones de tipo binario, ya que únicamente queremos realizar una separación en dos conjuntos: señal y fondo.

**Ejemplo:** Consideremos varias partículas del SM (véase Sección 2.1) caracterizadas por las siguientes propiedades:

- $P_1 = \text{Spin}$ : O bien se trata de partículas bosónicas o bien partículas fermiónicas (variable lógica).
- $P_2 = \text{Masa}$ : Valor de masa de la partícula (variable numérica continua).
- $P_3 = \text{Carga}$ : Valor de carga de la partícula en unidades de carga del electrón (variable numérica discreta).

Podemos clasificar el conjunto de partículas  $X = \{e^-, \mu^+, \nu, \gamma, Z, W^+, t\}$  en términos de un árbol de decisión de profundidad 3 como se observa en la Figura 8.

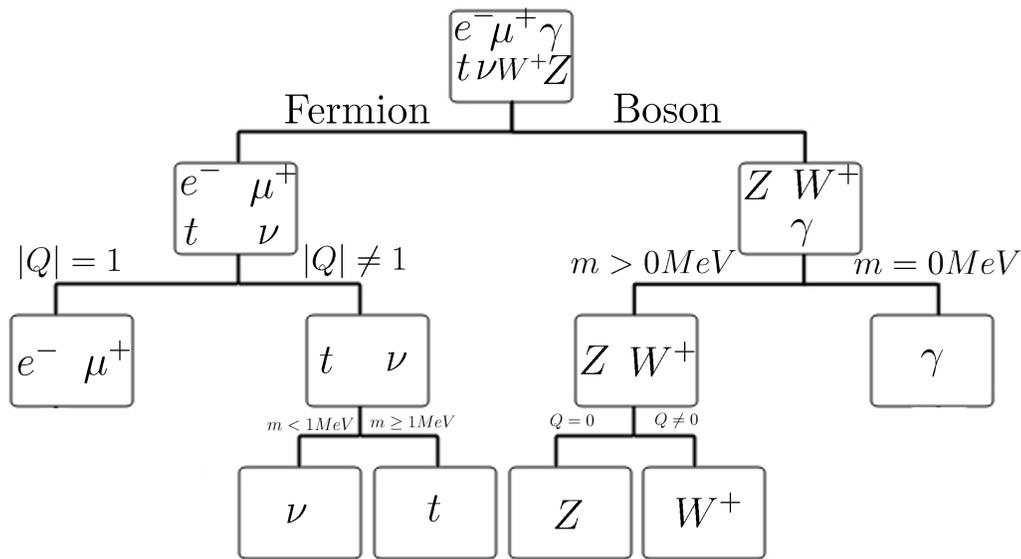


Figura 8: Ejemplo de un árbol de regresión.

Observamos que, obviamente, no todos los nodos tienen que ramificarse hasta el mismo nivel. Varios de los elementos originales del conjunto pueden terminar en la mismo **nodo** final u **hoja**.

Por otra parte, es sencillo ver cómo los árboles de regresión generalizan directamente a los cortes secuenciales, pues cada una de las elecciones de ramas que lleve a una de las hojas pueden representar cortes secuenciales en las distintas variables consideradas. Así, podemos considerar como conjunto inicial el **total de eventos esperados de señal y fondo**. Las propiedades de cada elemento serán las características físicas medidas para cada evento. Diremos que un nodo es **tipo señal** (S) si el número de eventos de señal en el mismo es mayoritario y que es **tipo fondo** (B) si la mayoría de los eventos del mismo son de fondo. De esta forma, mediante la selección de todas las hojas del árbol tipo S, podremos definir regiones de estudio en las que predomine la señal sobre el fondo.

Con esta definición, obtenemos una forma más general de definir regiones en nuestro espacio de parámetros total. Sin embargo, la ganancia en eficiencia respecto al caso de los cortes secuenciales aparece realmente cuando consideramos múltiples árboles de decisión, lo que se conoce como un **bosque**. La idea fundamental en este caso es definir una nueva variable para cada evento que nos indique en qué proporción de los árboles de decisión considerados pertenece a una región tipo S. Aunque en muchas ocasiones eventos de señal aparecerán en la región de fondo y viceversa, si elegimos adecuadamente la forma de los árboles esperamos que esta variable discrimine especialmente bien entre nuestros dos casos.

### 6.2.2. Construcción del BDT

El algoritmo de elección de los árboles es lo que da lugar al concepto de BDT. Presentamos a continuación las ideas básicas del algoritmo de construcción aplicadas al caso concreto de nuestro interés. Un desarrollo en mayor profundidad se puede observar en [16].

Consideremos que tenemos un conjunto de  $N_S$  sucesos de señal y  $N_B$  eventos de fondo que queremos separar mediante el uso de un BDT. Cada uno de los eventos está caracterizado por  $n$  variables físicas de forma que podemos identificar el evento  $x$  con la tupla  $x = (x_1, \dots, x_n)$ . Inicialmente, cada uno de los sucesos tendrá cierto peso  $w_x$  que, si no viene especificado por los procesos de simulación previos, consideraremos  $w_x = 1$ . Normalizamos dichos pesos de forma que  $\sum_{x \in S} w_x = \sum_{x \in B} w_x$  (equivalentemente, consideraremos muestras de señal y fondo del mismo tamaño efectivo). Cuando hablemos de una cantidad de sucesos en un nodo, consideraremos en todos los casos esta cantidad *pesada*, es decir, realmente consideraremos la suma de los pesos de todos los eventos en dicho nodo.

El primer paso del algoritmo consiste en la definición del árbol inicial  $T_1$ . Para cada nodo en el árbol se lleva a cabo el siguiente procedimiento.

En primer lugar, se comprueba que el nodo no ha alcanzado la **profundidad máxima** permitida en el árbol<sup>30</sup>, contiene una **cantidad de eventos mínima**<sup>31</sup> y no contiene ya únicamente eventos de un tipo. Si se cumplen estas dos condiciones, se procede a considerar cortes simples en cada una de las  $n$  variables del evento considerando varios valores límite para cada una (la **resolución en estos valores de corte** es otro parámetro a considerar en el BDT). De entre todos los cortes considerados, elegimos el que proporciona una mayor *calidad* en la separación y ramificamos el árbol según éste. Repetimos el proceso hasta que ningún nodo más pueda ser

<sup>30</sup>La potencia de los BDT depende principalmente en la capacidad de generar una gran cantidad de árboles simples en vez de centrarse en pequeñas cantidades de árboles complejos. Imponer una profundidad máxima reduce el tiempo de computación necesario.

<sup>31</sup>De cara a que las regiones definidas tengan una significación estadística razonable, se impone una cantidad de eventos mínima por nodo. Si el nodo original tiene menos del doble de esa cantidad, no permitiremos que se divida.

ramificado.

Para la definición de la *calidad* de la separación definimos en primer lugar la **pureza** del nodo  $N_i$  con un conjunto de eventos de señal  $S_i$  y de fondo  $B_i$  como:

$$p_i = \frac{\sum_{x \in S_i} w_x}{\sum_{x \in S_i} w_x + \sum_{x \in B_i} w_x}$$

Que no es más que una forma de expresar el valor de  $S/(S+B)$  para cada nodo. Observamos que  $p_i$  tomará valores cercanos al 0 en caso de que el fondo sea mayoritario y valores próximos a 1 si lo es la señal. Por este motivo, utilizaremos la pureza como criterio de designación de un nodo como tipo  $S$  ( $p_i > 0,5$ ) o  $B$  ( $p_i < 0,5$ ).

Si bien existen formas alternativas de proceder, asumiremos el **criterio de Gini** (o del índice de Gini) de cara a determinar la mejor separación de un nodo posible. Para cada nodo  $N_i$  definimos el índice de Gini  $G_i$ :

$$G_i = p_i(1 - p_i) \sum_{x \in N_i} w_x$$

Como  $p_i$  toma valores más próximos a 0 o 1 cuanto más *puro* en fondo o señal es el nodo, el índice de Gini tomará valores más elevados cuanto más *mezclado* esté el resultante. De cara a maximizar la calidad de la señal en los nodos resultantes maximizaremos el valor de la suma de índices de Gini de los nodos resultantes. Es decir, si los nodos ramificados a partir de  $N_i$  son  $N_{i1}$  y  $N_{i2}$  escogeremos el corte que maximice  $\delta(G_i) = G_{i1} + G_{i2}$ .

Una vez construido el primero de los árboles, surge el problema de definir el resto de los árboles de nuestro bosque.

Quizás la idea más lógica sería pasar a construir árboles adicionales eligiendo en algunos nodos cortes que no sean *óptimos*. Sin embargo, este procedimiento no nos garantiza que los procesos de señal que inicialmente clasificamos en hojas de fondo lleguen a formar parte en algún momento de hojas de señal. Es decir, podríamos estar definiendo árboles que, de forma redundante, clasificasen correctamente los mismos eventos una y otra vez y fallasen repetidamente en otros.

La importancia fundamental del BDT es que permite solucionar este problema mediante el procedimiento de **boosting**. En esencia, durante el *boosting* modificamos el peso de los eventos que inicialmente clasificamos erróneamente de forma que la construcción de un nuevo árbol (según el algoritmo anterior) los tenga en cuenta en mayor medida. Por iteración de construcción de árboles y de *boosting* se construyen una cantidad determinada de **árboles totales** (otro parámetro del método). Adicionalmente, en función de la *calidad de separación de cada árbol* se asigna a cada uno de estos un peso mayor o menor en el resultado final.

Existen diversos métodos para realizar el *boosting* y la asignación de pesos a árboles. Nos centraremos en el

conocido como AdaBoost [17].

Supongamos que hemos construido el árbol  $k$ -ésimo,  $T_k$ , y tenemos definidos los pesos correspondientes a los  $N_s + N_B$  sucesos totales de estudio. Se define el **error** del árbol  $k$ :

$$E_k = 1 - \frac{\sum_{x \in T_k} w_x Id(x)}{\sum_{x \in T_k} w_x}$$

Donde  $Id(x)$  es 1 si la hoja en la que se encuentra el suceso es del mismo tipo que este y 0 en otro caso. De forma efectiva,  $E_k$  mide la proporción de eventos falsamente clasificados. Cuanto mayor sea  $E_k$ , peor será la calidad del árbol obtenido. De cara a dar una importancia mayor a los árboles que proporcionan una mejor clasificación, se define el **peso del árbol** como<sup>32</sup>:

$$\alpha_k = \beta \log \left( \frac{1 - E_k}{E_k} \right)$$

Donde  $\beta$  representa la intensidad del *boost* y es un parámetro adicional del BDT. Por último, cada uno de los pesos del evento se redimensiona y normaliza a su vez bajo la forma:

$$w_x \rightarrow \frac{w_x e^{\alpha_k} (1 - Id(x))}{\sum_{y \in T_k} w_y e^{\alpha_k} (1 - Id(x))}$$

De forma que los eventos mal identificados ven su peso aumentado y los bien identificados lo ven disminuido. Adicionalmente, cuanto mejor sea el árbol (mayor valor de  $\alpha$ ), más se modifican los pesos, con lo que se consigue variar más sensiblemente los siguientes resultados para diversificar el poder de separación del BDT.

Con todas estas definiciones estamos en condiciones de definir la variable discriminante que utiliza los datos de todos los árboles generados. Para cada uno de los eventos estudiados se define la variable *BDT* o respuesta del BDT:

$$BDT(x) = \frac{\sum_{i=1}^{N_T} \alpha_i T_i(x)}{\sum_{i=1}^{N_T} \alpha_i}$$

Donde  $N_T$  es el total de árboles considerados y  $T_i(x)$  vale 1 si el evento se encuentra en una hoja tipo señal en el árbol  $i$ -ésimo o  $-1$  si se encuentra en una hoja tipo fondo.

Si el proceso de creación del BDT es correcto, los sucesos de fondo tendrán valores relativamente próximos a  $-1$  mientras que los sucesos de señal tendrán valores cercanos a 1. Debido a la naturaleza estadística de las medidas, esperaremos observar dos picos en la distribución de esta variable correspondientes a señal y fondo. Cuanto más separados estén dichos picos mayor poder de separación tendrá nuestro BDT.

---

<sup>32</sup>Observamos que, si  $E_k < 1/2$ ,  $\alpha_k$  no está bien definido como peso (es negativo). Técnicamente es necesario imponer que  $E_k > 1/2$ .

### 6.2.3. Características del BDT

Como vemos, la creación de un BDT implica necesariamente la utilización de muestras de datos que nos indiquen el comportamiento tanto de señal como de fondo. A este proceso por el que se construyen los distintos árboles constituyentes del mismo se le denomina **entrenamiento** para lo que se utilizarán muestras de señal y fondo específicamente dedicadas.

Un fenómeno relativamente frecuente tras el entrenamiento de un método multivariante cualquiera es lo que se conoce como *overtraining*. De forma intuitiva, el *overtraining* se da cuando el propio método se adapta específicamente a las peculiaridades de la muestra de entrenamiento usada que no son típicas del proceso de señal (o fondo) estudiado. El problema de un BDT con *overtraining* es que durante un proceso de aplicación del mismo no estará comprobando lo similar que son las muestras de test a la señal o el fondo, sino cuanto se asemejan sus peculiaridades particulares a las de las muestras utilizadas para el entrenamiento. Para evitar este fenómeno es necesario disponer de muestras que reproduzcan de forma fidedigna el comportamiento de los procesos (lo que a su vez requiere de una estadística adecuada, una cantidad de sucesos relevante) así como elegir adecuadamente los parámetros de trabajo del BDT, pues utilizar una profundidad exagerada o una cantidad de árboles demasiado grande puede dar lugar a la aparición de *overtraining*.

La forma más común de estudiar la validez del propio BDT es comprobar su funcionamiento con muestras adicionales (obviamente no podemos utilizar las mismas que para el caso del entrenamiento, pues el BDT está expresamente optimizado con ellas) en lo que se denomina procedimiento de **test**. Para realizar un test utilizamos muestras conocidas de señal o fondo, estadísticamente independientes de las muestras de entrenamiento, y comprobamos que, tras la aplicación del BDT, el resultado es cualitativamente similar al que se obtiene al aplicarlo sobre las de entrenamiento<sup>33</sup>. Evitar la presencia de *overtraining* es fundamental para obtener BDTs discriminantes.

Un concepto adicional a tener en cuenta es la **importancia** de las variables utilizadas dentro del conjunto de variables del BDT. Se define la importancia como el número de ramificaciones en los árboles del BDT que se dan a través de cortes en la variable (cada uno de estos cortes pesado a través del coeficiente correspondiente al árbol del que es parte). Usualmente se proporciona el valor normalizado de forma que la suma de la importancia de las variables sea la unidad de forma que la interpretación sea más rápida en términos del poder de decisión que proporcionan al BDT. Un interés directo de estudiar estos valores radica en el hecho de que variables de poca importancia pueden eliminarse del BDT para acelerar el tiempo de computación dedicado a su entrenamiento.

---

<sup>33</sup>Matemáticamente esta similaridad se puede expresar en términos de test de hipótesis bien conocidos sobre la distribución de la variable BDT para test y entrenamiento:  $\chi^2$ , Kolmogorov-Smírnov de doble distribución, etc.

Por último, es necesario mencionar la posibilidad de la aparición de **correlación** de las variables. Intuitivamente dos variables están correladas cuando sus distribuciones no son estrictamente independientes, es decir, se observa que para distintos valores fijos de una de ellas son más o menos probables valores determinados de la otra<sup>34</sup>. Desde el punto de vista del BDT la importancia de la correlación es directa: si dos variables se encuentran fuertemente correladas el valor de una determina, prácticamente, el valor de la otra y por tanto no puede proporcionar un poder de separación mayor al propio BDT. Como en el caso de variables de poca importancia, eliminar una de las variables correladas en un par de ellas no debería producir efectos mayores en el resultado proporcionado por el BDT y acelera el proceso de entrenamiento.

### 6.3. Incertidumbres estadísticas y sistemáticas

Una vez obtenida la forma del BDT para señal y fondo podemos dar una idea, mediante un corte en la variable de salida del BDT, de la proporción señal/fondo en una región típicamente muy rica en señal. Sin embargo, esto supone desechar gran parte de la información que proporciona el BDT: la propia distribución de dicha variable. Sería deseable un método estadístico que permitiese el análisis simultáneo de la distribución de la variable BDT y, al mismo tiempo, modelizase los efectos estadísticos (debidos a una cantidad pequeña de datos) que esto tuviese en la imprecisión de la medida.

Adicionalmente, nos interesa incorporar las diferentes fuentes sistemáticas (debidas al propio proceso de medida) de error en nuestro cálculo de forma que podamos estimar estadísticamente la posibilidad de garantizar exclusión (o detección) de procesos de *nueva física* dentro del error experimental en los puntos del espacio de parámetros considerados. Presentamos a continuación una explicación básica de los métodos estadísticos empleados. Dado que, en principio, requieren de una menor luminosidad total nos centraremos en la búsqueda de límites de exclusión.

#### 6.3.1. Full CLs

El **método Full CLs** (*Confidence Levels*, [17])o *CLs* es una herramienta estadística general (válida para muestras de señal y fondo de cualquier tamaño).

Consideremos un experimento para el cual queremos contrastar si tenemos únicamente cierto fondo  $B$  o realmente existe una señal adicional al fondo  $S + B$ . Definimos la **intensidad de la señal**,  $r$ , como el número de eventos observados por cada evento esperado. Nuestra idea es dar un *intervalo de confianza* para  $r$  acotado superiormente de forma que  $r \in [0, r_0]$  para cierto nivel de confianza  $\alpha$ . Consideraremos la hipótesis  $rS + B$  en

<sup>34</sup>Formalmente, se puede expresar la correlación en términos de las funciones de distribución conjunta y marginal de las variables. Existen gran cantidad de test de hipótesis destinados al propio estudio de la correlación entre variables, como por ejemplo tests  $\chi^2$ , tests de Pearson, etc.

vez de  $S + B$  (que recuperamos haciendo  $r = 1$ ).

Supongamos que tenemos un total de  $n$  canales de análisis (en nuestro caso las  $n$  columnas del histograma correspondiente a la variable BDT). Para el canal  $i$  esperamos  $rs_i + b_i$  cuentas bajo la hipótesis  $rS + B$  y  $b_i$  cuentas bajo la hipótesis  $B$ . Asumiendo que cada canal sigue una distribución de Poisson<sup>35</sup>, estos valores esperados serán el parámetro de cada distribución bajo cada hipótesis.

Supongamos que realizamos una medida experimental y obtenemos  $d_i$  cuentas en cada canal. La probabilidad de obtener estas medidas supuestas bajo cada hipótesis vendrá dada por las correspondientes funciones de probabilidad evaluadas en el punto  $d_i$ . Podemos comparar ambas probabilidades mediante un cociente para definir el estadístico de la razón de verosimilitudes  $X_i$  como una función de  $d_i$ :

$$X_i(d_i) = \frac{P_{rS+B}(d_i)}{P_B(d_i)} = \frac{e^{-(rs_i+b_i)}(rs_i + b_i)^{d_i}}{d_i!} \bigg/ \frac{e^{-(b_i)}(b_i)^{d_i}}{d_i!}$$

Que tomará valores altos si se favorece  $rS + B$  y bajos si se favorece  $B$ . La extensión a todos los canales no es más que el producto de verosimilitudes:

$$X(d_1, \dots, d_n) = \frac{P_{rS+B}(d_1, \dots, d_n)}{P_B(d_1, \dots, d_n)} = \prod_{i=1}^n \left( \frac{e^{-(rs_i+b_i)}(rs_i + b_i)^{d_i}}{d_i!} \bigg/ \frac{e^{-(b_i)}(b_i)^{d_i}}{d_i!} \right)$$

Tomemos  $X_{obs}$  el valor de  $X$  para nuestra observación experimental concreta  $\{d'_i\}_{i=1}^n$ . Podemos calcular varios niveles de confianza en términos probabilísticos. Definimos  $CL_{s+b}$  como la probabilidad de, supuesta cierta  $rS+B$  obtener unos datos tan *tipo B* o más (con valores iguales o menores de  $X$ ) como  $\{d'_i\}_{i=1}^n$ . Alternativamente tomamos  $CL_b$  la probabilidad, supuesta  $B$ , de obtener unos datos tan *tipo B* o más como los observados:

$$CL_{s+b} = P_{s+b}(X \leq X_{obs}) \quad CL_b = P_b(X \leq X_{obs})$$

Se define el coeficiente de confianza  $CL_s = \frac{CL_{s+b}}{CL_b}$ . Como  $s + b \geq b$ ,  $CL_{s+b} \leq CL_b$  y, por tanto,  $CL_s \leq 1$ . No sólo eso, sino que en el caso de un único canal se reduce al nivel de confianza correspondiente a obtener una intensidad de señal  $r$  igual o superior a la observada.

Es decir, que  $CL_s$  nos indica la *probabilidad* de que el valor real de  $r$  sea más elevado  $r_0$ . Sin más que encontrar el valor de  $r_0$  tal que  $CL_s = 0,05$  tendríamos un intervalo de confianza al 95% para el valor de  $r$ .

Usualmente diremos que hemos alcanzado *nivel de exclusión* cuando dicho intervalo no admita la hipótesis  $S + B$ , es decir, cuando  $r_0 < 1$ .

Ahora bien, hasta el momento hemos introducido únicamente el error estadístico en nuestro estudio. Adicio-

---

<sup>35</sup>La elección de una distribución de Poisson para cada celda (columna del histograma) está modelizando aquí el error estadístico debido al número de datos totales de los que disponemos.

nalmente hay fuentes de errores sistemáticos que afectan a toda la medida que tenemos que tener en cuenta en nuestro tratamiento.

La forma usual de incorporar estas incertidumbres es en términos estadísticos a través de variaciones estocásticas de las propias probabilidades definidas por las distribuciones de Poisson [18]. La idea es que cada una de estas probabilidades tendrá fluctuaciones debido a la presencia de cada fuente sistemática de imprecisión.

De forma práctica el cálculo analítico del efecto de estas variaciones suele resultar imposible o muy caro computacionalmente. La alternativa es realizar simulaciones de Montecarlo que reproduzcan su acción. Para ello, a partir de la observación experimental construimos gran cantidad de experimentos *de juguete* o *toys* con distintas observaciones modelizadas aleatoriamente mediante variaciones de la muestra. Para cada uno de estos podemos calcular los correspondientes valores al 95 % de confianza para  $r_0$  y obtener así una distribución empírica para  $r_0$ . A partir de los cuantiles de esta distribución podremos encontrar el valor *esperado* de  $r_0$  (la mediana). De la misma forma, podremos dar valores a  $\pm 2\sigma$  de cara a definir la imprecisión en la medida de  $r_0$  sin más que estudiar los cuantiles 0,025 y 0,975.

### 6.3.2. Método asintótico

El principal problema a la hora de aplicar el método Full CL es que todo el proceso de simulación de los *toys* consume una cantidad increíblemente elevada de recursos computacionales. El **método asintótico** [19] proporciona una alternativa notablemente más rápida pero, a cambio, exige una cantidad de datos relativamente amplia para ser aplicable.

Partiendo de un estudio Full CLs el método asintótico sustituye la generación de los modelos de juguete por un único set de datos: el conjunto *Asimov*. Se sacrifica la amplia generalidad del Full CLs a cambio de una notablemente mayor rapidez.

## 6.4. Proyección de resultados

Un hecho de relevancia en la Física de Partículas (y en general en casi toda ciencia experimental) es la necesidad de obtener muestras de datos lo suficientemente abundantes como para que las conclusiones que obtengamos de ellas sean relevantes desde el punto de vista estadístico. En el caso particular de los estudios de *nueva física*, dados los ingentes recursos que consume una serie de tomas de datos, es necesario en muchas ocasiones trabajar con muestras bastantes reducidas que no nos permiten garantizar resultados firmes.

Este problema es tanto o más acuciante cuando el tipo de procesos que estudiamos tienen una baja sección eficaz, como en los dos casos de interés en nuestro estudio. A causa de la baja luminosidad de la que disponemos es altamente probable que no podamos garantizar la detección (o exclusión) de procesos de *nueva física*. Querriamos poder dar una idea de cuánta luminosidad es necesaria para lograr la exclusión, es decir, garantizar

la ausencia de *nueva física* en la región de trabajo. Para ello, utilizaremos los métodos estadísticos usados para el estudio de los errores sistemáticos sobre muestras de luminosidad creciente (aumentada de forma artificial redimensionando los pesos de los eventos) hasta encontrar los puntos en los cuales encontramos exclusión (en el sentido del apartado anterior). En este caso diremos que hemos alcanzado la luminosidad de exclusión esperada.

## 6.5. Variables de interés

De cara a construir un BDT para la discriminación señal/fondo en nuestros procesos tenemos que definir las variables físicas características de estos que, en principio, tienen comportamientos distintos en ambos casos. Presentamos a continuación las variables utilizadas durante el estudio, seleccionadas ampliando el trabajo presente en [20] junto con una breve explicación de la importancia como discriminantes.

1.  $\cancel{E}_T$ . La energía faltante tal y como está definida en la Sección 4.1.5. Nuestros procesos de señal contienen, en ambos casos, un total de 4 partículas invisibles al detector. Además, al menos dos de ellas (los neutralinos o las partículas de DM) son además relativamente pesadas. Esperamos, entonces, que la energía faltante en los procesos de señal sea bastante más elevada que en los diferentes procesos de fondo que contienen únicamente partículas invisibles ligeras y, usualmente, en menor cantidad. Es por ello que  $\cancel{E}_T$  debería ser una de las variables con mayor poder discriminante de entre las que podamos estudiar.
2.  $m_{ll}$ . La masa invariante de los dos leptones exigidos a nuestro evento. Considerando los 4-momentos de estos  $P(l_1)$  y  $P(l_2)$ , se define la masa invariante de los mismos como  $m_{ll} = \sqrt{(P(l_1) + P(l_2))^2}$ .

En caso de que los dos leptones provengan de la misma partícula  $X$ , por ejemplo en el caso  $Z \rightarrow ll'$ , la suma de sus 4-momentos es el 4-momento de  $X$  y, por tanto,  $m_{ll} = m_X$ .

En principio, dado que en la señal ambos leptones proceden de distintos bosones  $W$ , no tenemos ninguna condición física sobre el valor de  $m_{ll}$ . Sin embargo, en procesos de fondo en los que ambos procedan de una misma partícula  $X$  (usualmente procesos que contienen un bosón  $Z$ ), debería aparecer un pico en la distribución de  $m_{ll}$  en torno a  $m_X$ . De esta forma, deberíamos poder discriminar la señal respecto a los fondos *Drell – Yan*,  $ZZ$ ,  $ZW$ ,  $WZ$  y  $TTZ$ .

3.  $M_{T2, ll}^2$ . En los casos en los que los resultados de la desintegración de una partícula son *invisibles* no podemos reconstruir completamente su 4-momento ya que únicamente nos es posible determinar las componentes transversas del momento de todos los objetos involucrados y, por tanto, no podemos reconstruir la masa original a partir de la masa invariante del sistema. Para ello se introduce el concepto de **masa transversa**. Consideremos, por ejemplo, el caso de la desintegración  $W \rightarrow l\nu$ . De conocer  $P^l$  y  $P^\nu$  tendríamos  $m_W^2 = (P^l + P^\nu)^2 = m_l^2 + m_\nu^2 + 2(E^l E^\nu - \vec{p}^l \vec{p}^\nu)$ . Ahora bien, únicamente podemos medir  $p_T^\nu$  (como  $\cancel{E}_T$ ). Tomando

$\mathcal{E}_T = |\vec{p}_T^\nu|$  se define la **masa transversa**: (véase [21])  $m_T^2(p_T^l, p_T^\nu) = 2(E_T^e \mathcal{E}_T - \vec{p}_T^l \vec{p}_T^\nu)$ . Con esta definición es sencillo probar<sup>36</sup> que  $m_T^2 \leq m_W^2$  para toda combinación posible de los momentos de leptón y neutrino.

Ahora bien, en los procesos que nos interesan se producen simultáneamente varias partículas *invisibles* que, además, no podemos medir individualmente. Considerando un proceso  $t\bar{t}$  dileptónico tendremos dos neutrinos contribuyendo a  $\mathcal{E}_T$ , por lo que aparece la condición  $p_T^{miss} = p_T^{\nu_1} + p_T^{\nu_2}$ . Como ambos neutrinos proceden del proceso  $W \rightarrow l\nu$ , tendremos que, de poder medir sus momentos transversos:  $\max(m_T^2(p_T^{l_1}, p_T^{\nu_1}), m_T^2(p_T^{l_2}, p_T^{\nu_2})) \leq m_W^2$ . Utilizando la ligadura  $p_T^{miss} = p_T^{\nu_1} + p_T^{\nu_2}$  podemos considerar todas las posibles descomposiciones de  $p_T^{miss}$  compatibles para definir la **masa transversa**:  $M_{T2,\ell}^2 = \min_q \left[ \max(m_T^2(p_T^{l_1}, p_T^{miss} - q), m_T^2(p_T^{l_2}, q)) \right]$ . Como entre todas las posibles descomposiciones se encuentran, en el caso ideal, los valores reales de  $p_T$  para ambos neutrinos tendremos que una cota superior para  $M_{T2,\ell}$  viene dada por  $M_{T2,\ell}^2 \leq m_W^2$ .

Consideremos ahora el caso de  $T2tt$  o  $tt + DM$ , nuestros procesos de señal. En ambos casos tenemos un total de 4 partículas invisibles por lo que la condición  $p_T^{miss} = p_T^{\nu_1} + p_T^{\nu_2}$  se ve sustituida por  $p_T^{miss} = p_T^{\nu_1} + p_T^{\nu_2} + p_T^{\chi_1} + p_T^{\chi_2}$ . Como consecuencia,  $M_{T2,\ell}^2 \not\leq m_W^2$ . Es decir, se elimina el límite superior en los valores de  $M_{T2,\ell}$ . Este motivo es la causa principal de que  $M_{T2,\ell}$  sea una variable con un increíble poder discriminante entre nuestros procesos de señal y los fondos en los que ambos leptones procedan de bosones  $W$  y no aparezcan partículas invisibles adicionales. Por tanto, es de esperar que  $M_{T2,\ell}$  permita discriminar los procesos de señal respecto a  $t\bar{t}$  (proceso de fondo ampliamente dominante),  $tW$ ,  $WW$ , etc.

4.  $M_{T2,bb}$ . Siguiendo un razonamiento análogo al anterior, en los casos en los que tenemos 2 b-jets podemos definir una variable similar a  $M_{T2,\ell}$ . En este caso, la idea es intentar reconstruir la masa del quark  $t$  de una forma similar a como antes reconstruíamos la del quark  $W$ . En este caso el papel de  $p_T^{miss}$  tendría que ser desempeñado por un objeto que reconstruyese el momento transversal de los dos bosones  $W$  que acompañan a los dos quarks  $b$  en la desintegración de  $t\bar{t}$ . Es sencillo reconstruir su momento transversal en un proceso en el que las únicas partículas invisibles son los neutrinos ya que el momento transversal de cada leptón está bien determinado. Basta con tomar  $p_T^{WW} = p_T^{miss} + p_T^{l_1} + p_T^{l_2}$ . Definimos:  $M_{T2,bb}^2 = \min_q \left[ \max(m_T^2(p_T^{b_1}, p_T^{WW} - q), m_T^2(p_T^{b_2}, q)) \right]$ .

Como en el caso de  $M_{T2,\ell}$ , en caso de que únicamente estén presentes dos partículas invisibles tenemos la restricción  $p_T^{WW} = p_T^{W_1} + p_T^{W_2}$  y por lo tanto aparece la cota superior  $M_{T2,bb}^2 \leq m_t^2$ . De nuevo, esta restricción no queda impuesta en caso de tener 4 partículas invisibles<sup>37</sup>, con lo que nuestros procesos de señal pueden tener  $M_{T2,bb}^2 > m_t^2$ . De nuevo, esto nos puede ayudar a discriminar respecto a los fondos  $t\bar{t}$

<sup>36</sup>En coordenadas pseudoeuclídeas (la métrica usual), no es más que eliminar el término al cuadrado correspondiente al momento longitudinal del neutrino del total de la energía.

<sup>37</sup>En el caso de  $T2tt$ , realmente estamos reconstruyendo  $p_T^{\tilde{t}}$  por lo que el límite quedaría sustituido por  $M_{T2,bb}^2 \leq m_{\tilde{t}}^2$ . Lo mismo ocurrirá con  $M_{T2,bb}$ .

(proceso de fondo ampliamente dominante),  $tW$ ,  $WW$ .

5.  $M_{T2,lb\bar{b}}$ . Con  $M_{T2,bb}$  nos aprovechamos de la posibilidad de reconstruir la masa del quark  $t$  pero no utilizamos el hecho de que conocemos por separado los momentos transversos de cada uno de los dos leptones involucrados. A la hora de intentar utilizarlo nos encontramos con el problema principal de que no podemos asignar cada uno de los leptones de forma unívoca al quark  $b$  que aparece en su cadena de desintegración. La solución en este caso pasa por considerar todas las posibles combinaciones quark  $b$  - leptón y tomar el mínimo de *masas transversas* obtenidas. Es decir, definimos:

$$M_{T2,lb\bar{b}}^2 = \min \left( \min_q \left[ \max(m_T^2(p_T^{l_1} + p_T^{b_1}, p_T^{miss} - q), m_T^2(p_T^{l_2} + p_T^{b_2}, q)) \right], (\dots) \right. \\ \left. (\dots) \min_q \left[ \max(m_T^2(p_T^{l_1} + p_T^{b_2}, p_T^{miss} - q), m_T^2(p_T^{l_2} + p_T^{b_1}, q)) \right] \right)$$

Como esta construcción contempla el caso de una asignación correcta de leptón-bosón con la partición adecuada de  $p_T^{miss}$  tendremos el límite  $M_{T2,bb}^2 \leq m_t^2$ .

Como en casos anteriores, la aparición de 4 partículas invisibles en la señal implica que no existe tal límite para la misma y, por tanto la variable  $M_{T2,lb\bar{b}}$  podrá resultar un discriminante adecuado entre señal y fondo.

6.  $H_T$ . Se define la **energía hadrónica transversa total** como  $H_T = |\sum_{jets} p_T^{jets}|$ . Es decir, contiene información sobre la cantidad de actividad hadrónica en cada evento. Desde el punto de vista de nuestra señal, tenemos en todos los casos como mínimo 2 b-jets que, al proceder de quarks  $t$  que son mucho más pesados que el quark  $b$ , tienen un momento bastante elevado. Por tanto, en ambos procesos de señal esperamos valores de  $H_T$  bastante elevados. En cuanto a los fondos, en principio todos los procesos de fondo pueden incluir la producción jets de una forma u otra. Sin embargo, no en todos los casos proceden de desintegraciones de quarks  $t$  (todo el fondo dibosónico:  $WW$ ,  $ZZ$ ,  $WZ$ ) o se deben a eventos del *underlying event* (Drell-Yan) por lo que esta variable puede resultar un buen discriminante en dichos casos.

7.  $M_{eff}$ . Definimos  $M_{eff}$  como la energía transversa total de las partículas principales en los eventos señal. Es decir, tomamos  $M_{eff} = p_T^{l_1} + p_T^{l_2} + p_T^{jet_1} + p_T^{jet_2} + \cancel{E}_T$ . En nuestra señal  $T2tt$  tenemos en principio partículas con un *boost* bastante alto ya que en última instancia provienen de el squark  $\tilde{t}$  que en los dos puntos estudiados es increíblemente masivo. Además, la contribución en ambas señales de  $\cancel{E}_T$  debido a la presencia de partículas invisibles masivas hace suponer que debemos esperar valores muy elevados de  $M_{eff}$  en ambos casos. En los procesos de fondo no esperamos valores tan elevados de  $\cancel{E}_T$  (véase el punto 1) y tampoco que las partículas aparezcan con un *boost* tan elevado (salvo en los casos en los que tengamos  $t\bar{t}$ ).

Por tanto, es razonable asumir que  $M_{eff}$  será una buena variable discriminante en nuestro estudio.

8.  $P_T^{llmiss}$ . Definimos  $P_T^{llmiss} = \cancel{E}_T + p_T^{l_1} + p_T^{l_2}$ . De la misma forma que  $M_{eff}$ , la variedad de partículas invisibles en el estado final debería dar lugar a una gran cantidad de  $\cancel{E}_T$  en la señal por lo que esperamos que sea una variable discriminante.
9. \* Consideradas las variables derivadas de la energía transversa en el evento, podemos intentar aprovechar las relaciones angulares entre los diferentes objetos como variables discriminantes. Si bien su interpretación no es tan directa como en los casos anteriores, las características topológicas de los procesos pueden justificar pequeñas diferencias entre unos procesos y otros que nos pueden llegar a proporcionar cierto poder de separación.
10.  $\Delta(\phi)_{j,\cancel{E}_T}$ . La distancia angular entre  $p_T^{miss}$  y  $p_T^{jet}$  para el jet más cercano.
11.  $m(\Delta(\phi)_{j,\cancel{E}_T})$ . La distancia angular entre  $p_T^{miss}$  y  $p_T^{jet}$  para el jet más cercano de entre los dos jets con más  $p_T$ .
12.  $\Delta(\phi)_{l,\cancel{E}_T}$ . La distancia angular entre  $p_T^{miss}$  y  $p_T^l$  para el leptón más cercano.
13.  $\Delta(\phi)_{l,l}$ . La distancia angular entre  $p_T^{l_1}$  y  $p_T^{l_2}$ .
14.  $\Delta(\phi)_{j,l}$ . La distancia angular entre  $p_T^{l_1}$  y  $p_T^{jet_1}$  (los más energéticos).
15.  $\Delta(\phi)_{llmiss,\cancel{E}_T}$ . La distancia angular entre  $P_T^{llmiss}$  y  $\cancel{E}_T$ .

En principio utilizaremos para nuestro estudio las 15 variables anteriores a fin de intentar aprovechar lo más posible las diferencias cinemáticas y dinámicas que puedan aparecer entre los procesos de señal y los de fondo. Si bien la elección de un número tan, relativamente, elevado de variables puede ralentizar el proceso de entrenamiento del BDT, esperaremos poder reducir el número de estas variables a través de estudios entre sus posibles correlaciones y eliminación de las no discriminantes.

## 7. Estudio de los procesos $T2tt$

Comenzamos el estudio en sí intentando aplicar los métodos anteriores sobre las muestras  $T2tt$ , procesos de señal correspondientes a un modelo simplificado de tipo SUSY (Sección 2.3.1).

En todos los casos se utilizan como muestras de fondo para entrenamiento y test todas las muestras correspondientes a los procesos estudiados en la Sección 5.3 (divididas en una mitad para cada caso). Para el test utilizaremos la muestra completa correspondiente al punto estudiado.

De cara a maximizar la estadística total en el entrenamiento, consideraremos para el mismo todas las muestras disponibles tales que  $m_{\tilde{t}} - m_{\chi_1^0} = \Delta(m)$  donde  $\Delta(m)$  vendrá dado de la aplicación de esta fórmula sobre los parámetros del punto estudiado. Esperamos que la cinemática de los procesos quede bastante fijada por esta diferencia, la energía *libre* que se produce y puede transmitir al quark  $t$  que se produce.

### 7.1. Parámetros del BDT

Antes de tratar nuestros datos es necesario comprender en la mayor profundidad posible el funcionamiento práctico de nuestro método multivariante: el BDT. Concretamente, nos interesa comprobar en qué medida los diferentes parámetros de funcionamiento del mismo pueden afectar a un proceso de discriminación similar al de nuestro estudio.

Con el objetivo de determinar los parámetros de trabajo óptimos, realizamos un estudio detallado del proceso de entrenamiento del BDT utilizando muestras correspondientes a un punto *excluido* intermedio entre los dos de estudio:  $m_{\tilde{t}} = 750\text{GeV}$ ,  $m_{\chi_1^0} = 200\text{GeV}$  (aplicamos el razonable, véase Sección 6.5, corte de preselección  $M_{T2,l} > 80\text{GeV}$  para acelerar el proceso de entrenamiento), variando tres parámetros principales: profundidad máxima ( $D$ ), número de árboles totales ( $N_T$ ) y el parámetro  $\beta$  o parámetro de *boost*.

Estudiaremos tres propiedades fundamentales del árbol como criterios de calidad: la forma de la distribución final en la muestra de test (es decir, su poder discriminante), la ausencia de *overtraining* (es decir, la generalidad del entrenamiento al test) y la rapidez del proceso de entrenamiento.

1. Tengamos en cuenta que, pese a que normalizaremos los datos de partida a la hora de clasificarlos para proporcionar una mejor visualización del resultado, la realidad es que los procesos de fondo tienen una sección eficaz muchísimo mayor que los de señal y, por tanto, la más pequeña cantidad porcentual de fondo en una región es suficiente para equipararse o superar a la señal. Por tanto, caracterizaremos el poder discriminante del BDT por dos parámetros principales altamente relacionados: la cantidad de *columnas* en el histograma de la distribución de la variable *respuesta del BDT* que son libres de fondo (pesadas por su contenido) y el hecho de que bajo el pico de señal la cola de fondo sea lo más pequeña posible (que los picos de señal y fondo se encuentren separados).

2. Por otra parte, la presencia de *overtraining* se puede cuantizar en términos estadísticos (test de Kolmogorov-Smírnov, K-S, de doble distribución). Alternativamente, se puede observar cualitativamente en la diferencia de formas que presentan los resultados de entrenamiento y test para los casos de señal y fondo por separado. La aparición de una gran cantidad de puntos *alejados* entre sí entre los resultados de ambas muestras es un indicio claro de *overtraining*. En este último caso nos centraremos especialmente en los puntos situados en la zona de señal más extrema, pues son los que queremos caracterizar especialmente bien en nuestras mediciones.
3. Por último, aunque la rapidez del proceso es mucho menos relevante que obtener un BDT con un poder de separación lo suficientemente elevado, como variable nos permite discriminar en casos dudosos respecto al resto de propiedades.

En principio el comportamiento cinemático y dinámico del punto de señal escogido no es necesariamente similar a ninguno de los dos que estudiaremos en profundidad. Sin embargo, observemos que sí reproduce las propiedades estadísticas de estos, en cuanto a que presenta una cantidad de eventos totales similares y, considerando la diagonal de masas en la que se encuentra, podemos realizar un entrenamiento con un número de puntos parecidos a los utilizados luego en los dos puntos de interés. Observemos, además, que los parámetros modificados afectan únicamente a la construcción del BDT. Aunque esta dependa de la forma de la señal, podemos suponer que los resultados obtenidos son extrapolables (al menos en términos de orden de magnitud) al resto de puntos estudiados.

### 7.1.1. Profundidad del BDT

En primer lugar estudiamos la profundidad máxima del BDT,  $D$ . Para ello, consideramos valores que varían desde  $D = 1$  hasta  $D = 20$  y observamos el comportamiento de la aplicación del BDT con cada uno de estos valores del parámetro.

Los resultados obtenidos para  $D = 2$ ,  $D = 5$  y  $D = 15$  se pueden observar en la Figura 9. La evolución parece relativamente clara a la vista del resultado gráfico.

En primer lugar, el *overtraining* crece a medida que aumenta el valor de  $D$  (cuantos más cortes, más se adapta el BDT al entrenamiento). Se puede observar fácilmente que las distribuciones de test y entrenamiento coinciden con bastante exactitud (dentro de su incertidumbre) en el caso  $D = 2$ . Incluso se llega a obtener un  $p$ -valor de  $P = 0,072$  del test K-S<sup>38</sup>. En el paso a  $D = 5$  la forma empeora notablemente (obsérvese ese pico

<sup>38</sup>El  $p$ -valor nos indica la probabilidad de obtener dos muestras tan distintas como estas o más suponiendo cierta la hipótesis de que su distribución es la misma. Por tanto, nos interesa que sea lo más elevado posible. Sin embargo, la importante presencia de las

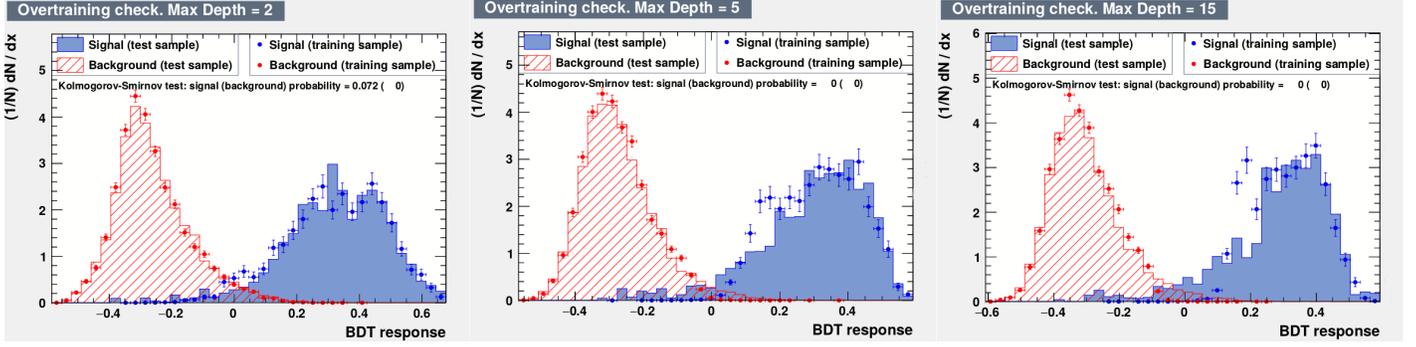


Figura 9: Estudio del *overtraining* del BDT en función de la máxima profundidad definida para el mismo. Se encuentran representadas tanto las muestras de fondo (rojo) como de señal (azul) de entrenamiento (puntos) y test (barras). Todas las muestras se encuentran normalizadas al mismo peso total. Los tiempos de entrenamiento empleados son, respectivamente, de 58s, 124s y 163s. Los parámetros utilizados son  $N_T = 2000$  y  $\beta = 0,1$ .

adicional en torno a  $BDT = 0,2$ ), lo que termina aún peor en  $D = 15$ .

En cuanto a la forma de  $BDT$ , notamos cierta mejoría de  $D = 2$  a  $D = 15$  (una mayor cantidad de celdas libres de fondo). Sin embargo, no es lo suficientemente pronunciada como para pasar por alto los fallos debidos al *overtraining*.

Del estudio global de los resultados concluimos que el valor máximo de  $D$  que arroja unos resultados razonables se encuentra en  $D = 2$  o, siendo especialmente estrictos, en  $D = 1$  (si bien la pérdida de poder de separación se hace palpable en este último).

### 7.1.2. Árboles totales

El segundo parámetro del BDT a considerar en nuestro estudio vendrá dado por el total de árboles,  $N_T$ . Como en el caso anterior, consideramos valores entre 10 y 100000 árboles totales a fin de observar la evolución del comportamiento. En la Figura 10 se pueden observar casos representativos de los resultados obtenidos:  $N_T = 200$ ,  $N_T = 2000$  y  $N_T = 20000$ .

El caso  $N_T = 200$  resulta especialmente representativo del hecho de considerar una cantidad demasiado pequeña de árboles totales. Si bien el *overtraining* presente es relativamente escaso (únicamente una pequeña columna), la forma de los dos picos resulta ya un indicio claro de carencia estadística: el pico de señal ni siquiera se advierte como tal. Aunque esto sería en principio favorable (una mejor separación), nos está señalando la falta de una *tendencia hacia la normalidad* típica<sup>39</sup> típica de la mayor parte de los procesos estadísticos. Se confirma esta falta de estadística en cuanto a que las colas de señal y fondo son enormemente largas: una cantidad relativamente elevada de procesos de fondo se encuentran bajo el pico de señal. Esto nos hace rechazar estos

incertidumbres experimentales, que K-S no tiene en cuenta, dificulta enormemente obtener altos valores en este test de hipótesis.

<sup>39</sup>A través del bien conocido Teorema del Límite Central.

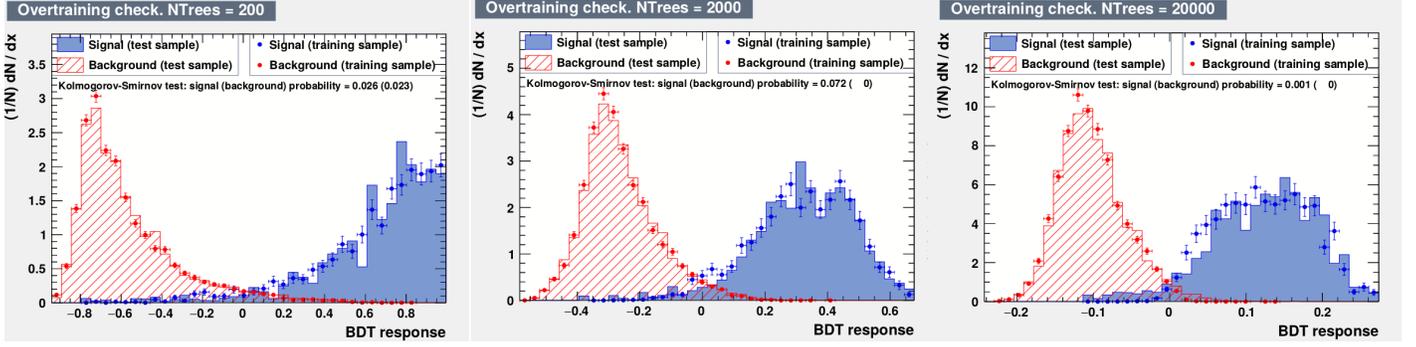


Figura 10: Estudio del *overtraining* del BDT en función de los árboles totales presentes en el mismo. Se encuentran representadas tanto las muestras de fondo (rojo) como de señal (azul) de entrenamiento (puntos) y test (barras). Todas las muestras se encuentran normalizadas al mismo peso total. Los tiempos de entrenamiento empleados son, respectivamente, de 5,9s, 58,1s y 588,0s. Los parámetros adicionales utilizados son  $D = 2$  y  $\beta = 0,1$ .

valores del parámetro  $N_T$ .

Los casos  $N_T = 20000$  y  $N_T = 2000$  son relativamente similares. La forma es relativamente similar (con 8 y 6 columnas libres de fondo respectivamente). Aparece una ligera cantidad de *overtraining* en  $N_T = 20000$  sobre  $N_T = 2000$ , tendencia que se confirma a valores más altos de  $N_T$ . La principal diferencia entre ambos es el tiempo de entrenamiento (y de evaluación) empleado que, como podemos observar, aumenta directamente con  $N_T$ . En vista de que aumentar el número de árboles no proporciona un poder de separación excesivamente superior (la mejora a partir de  $N_T = 10000$  es casi nula), aumenta ligeramente el *overtraining* y bastante el tiempo de trabajo, nos restringiremos a un rango entre  $N_T = 500$  y  $N_T = 10000$  para nuestro estudio.

### 7.1.3. Parámetro de *boost*

El último parámetro del BDT a tratar es el valor del parámetro de *boost*, AdaBoost:  $\beta$ . Como en casos anteriores, consideraremos varios valores del parámetro entre  $\beta = 0,01$  y  $\beta = 1$ . En la Figura 11 se encuentra una muestra representativa de los resultados.

Como es lógico, valores altos de  $\beta$  tienden a centrar los resultados del BDT en torno al 0 (ya que los eventos con mejores resultados en cada iteración tienen menos importancia en las siguientes y tienden a clasificarse peor después). El resultado a bajos valores de  $\beta$ , aunque tiene un pico de señal muy extremo, presenta colas de fondo muy alargadas que pueden cubrir completamente a dicha señal salvo en los eventos más extremos. Por otra parte, a valores altos de  $\beta$  aumenta también el *overtraining* (observemos sobre todo las colas de la señal). Adicionalmente el tiempo de entrenamiento no varía excesivamente de unos casos a otros.

A la vista de estos resultados, concluimos como aceptable un rango relativamente elevado de valores, de  $\beta = 0,05$  (se empieza a destacar el pico de señal) a  $\beta = 0,3$ .

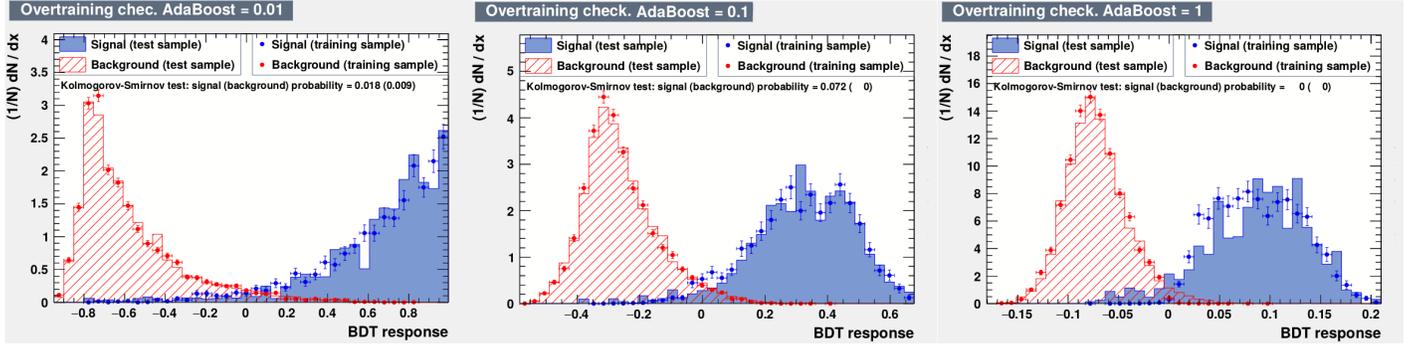


Figura 11: Estudio del *overtraining* del BDT en función del valor del parámetro de *boosting*  $\beta$ . Se encuentran representadas tanto las muestras de fondo (rojo) como de señal (azul) de entrenamiento (puntos) y test (barras). Todas las muestras se encuentran normalizadas al mismo peso total. Los tiempos de entrenamiento empleados son, respectivamente, de 57,7s, 58,1s y 58,4. Los parámetros utilizados son  $N_T = 2000$  y  $D = 2$ .

## 7.2. Punto 850 – 100

Comenzamos nuestro estudio físico considerando el punto  $m_{\tilde{t}} = 850$ ,  $m_{\chi_1^0} = 100$  o, abreviadamente, punto 850 – 100.

Como indicamos anteriormente, para el entrenamiento se consideran varias simulaciones de Montecarlo correspondientes a otros puntos en la misma diagonal del plano  $m_{\tilde{t}} - m_{\chi_1^0}$  que nuestro punto de interés. Es decir, tomaremos puntos con  $m_{\tilde{t}} - m_{\chi_1^0} = 750 \text{ GeV}$  para  $m_{\tilde{t}} \in [750, 800, 900, 950]$ . Esperamos que el comportamiento cinemático del proceso dependa fuertemente de  $m_{\tilde{t}} - m_{\chi_1^0}$  y que, por tanto, estas muestras sean similares en cierta medida a la del punto de interés (de esta forma enriquecemos la muestra de entrenamiento con más datos).

Por último, el procedimiento de test se lleva a cabo con los propios datos del punto  $m_{\tilde{t}} = 850$ ,  $m_{\chi_1^0} = 100$ .

### 7.2.1. Muestras y variables

Como paso previo al entrenamiento del BDT consideremos en primer lugar las muestras de datos utilizadas. En la Tabla 1 se puede encontrar un breve resumen de las mismas, así como de los eventos totales que las componen a una luminosidad integrada de  $L = 1,28 \text{ fb}^{-1}$  (luminosidad integrada de los datos). Notemos que no se trata de eventos reales, si no de eventos pesados, es decir, las muestras consideradas tienen una estadística bastante mayor (de cara a modelizar mejor la verdadera distribución de las variables físicas del proceso) pero cada evento tiene un peso bastante inferior a 1. La excepción es la muestra de datos, en la que todos los eventos tienen peso unidad.

El primer hecho a observar es que el fondo está completamente dominado (más del 90 %) por el proceso  $t\bar{t}$ . Si bien otros procesos como  $W + Jets$  y  $Drell - Yan$  tienen una mayor sección eficaz (Figura 7), la preselección prácticamente elimina el resto de los fondos. Es por ello que centraremos nuestro estudio fundamentalmente en la reducción del fondo correspondiente a este proceso.

Muestra	Eventos (pesados)/Ad
$t\bar{t}$ (MC)	10004
$tW$ (MC)	437
Drell-Yan (MC)	373
$t\bar{t}V$ (MC)	54
$VV$ (MC)	10
$W + jets$ (MC)	8
Fondo (Total)	10886
$T2tt$ (MC, $m_{\tilde{t}} = 850$ , $m_{\chi_1^0} = 100$ )	0.38
$T2tt$ (MC, $m_{\tilde{t}} = 600$ , $m_{\chi_1^0} = 300$ )	3.10

Tabla 1: Eventos totales correspondientes a una luminosidad de  $L = 1,28 fb^{-1}$  para los diferentes procesos de señal y fondo considerados en nuestro estudio tras la preselección de datos. Nótese lo pequeñas que son las relaciones  $S/B = 0,000035$  ( $m_{\tilde{t}} = 850$ ,  $m_{\chi_1^0} = 100$ ) y  $S/B = 0,00028$  ( $m_{\tilde{t}} = 600$ ,  $m_{\chi_1^0} = 300$ ) incluso después de la preselección. (el hecho de que las muestras MC contengan eventos de peso no unitario da lugar a un número de eventos esperados no entero).

Observemos que los eventos esperados a la luminosidad que tenemos se encuentran en el orden de 1 o incluso 0. Es por tanto lógico que, por muy potente que sea nuestro método de separación, nos será prácticamente imposible garantizar detección o exclusión de los procesos SUSY en este canal y a esta luminosidad y nos tendremos que centrar en extrapolar los resultados en términos de las luminosidades esperadas en las que alcanzaremos esto.

Adicionalmente, un hecho a notar es que ambos procesos de señal presentan una cantidad de eventos muy inferior (de 4 a 5 órdenes de magnitud) a los eventos totales de fondo esperados. Es decir, incluso tras la preselección específicamente dirigida a enriquecer  $S/B$ , la proporción de fondo en el total es abrumadoramente dominante.

Por otra parte, nos interesa conocer aproximadamente el comportamiento de las variables que participan en nuestro entrenamiento de cara a entender mejor las posibilidades discriminantes de cada una de ellas. En las Figuras 12 y 13 se pueden observar las distribuciones de 8 de las variables de interés físico que utilizaremos para nuestro entrenamiento.

De la forma de las variables representadas se puede inducir rápidamente las que, en principio, darán lugar a un poder de separación mayor. Es decir, las que tengan una forma más distinta para señal y fondo. En estos casos destacan  $M_{T2,u}$  y  $E_T^{miss}$  ( $\cancel{E}_T$ ) en la parte superior de la Figura 12 y  $M_{T2,blb}$  entre las variables de la Figura 13. En los tres casos podemos observar claras diferencias de forma entre el fondo (con valores más bajos) y la señal (cuya cola derecha se extiende hacia valores más elevadas), como esperábamos debido a la física del problema (Sección 6.5). Sin embargo, aún en estos casos una parte importante de la señal se encuentra superpuesta a los procesos de fondo. Exceptuando zonas muy limitadas en las colas de las distribuciones (que no alcanzan en ningún momento a sumar un evento esperado total), en casi todo punto aparece un fondo mínimo que tapanía

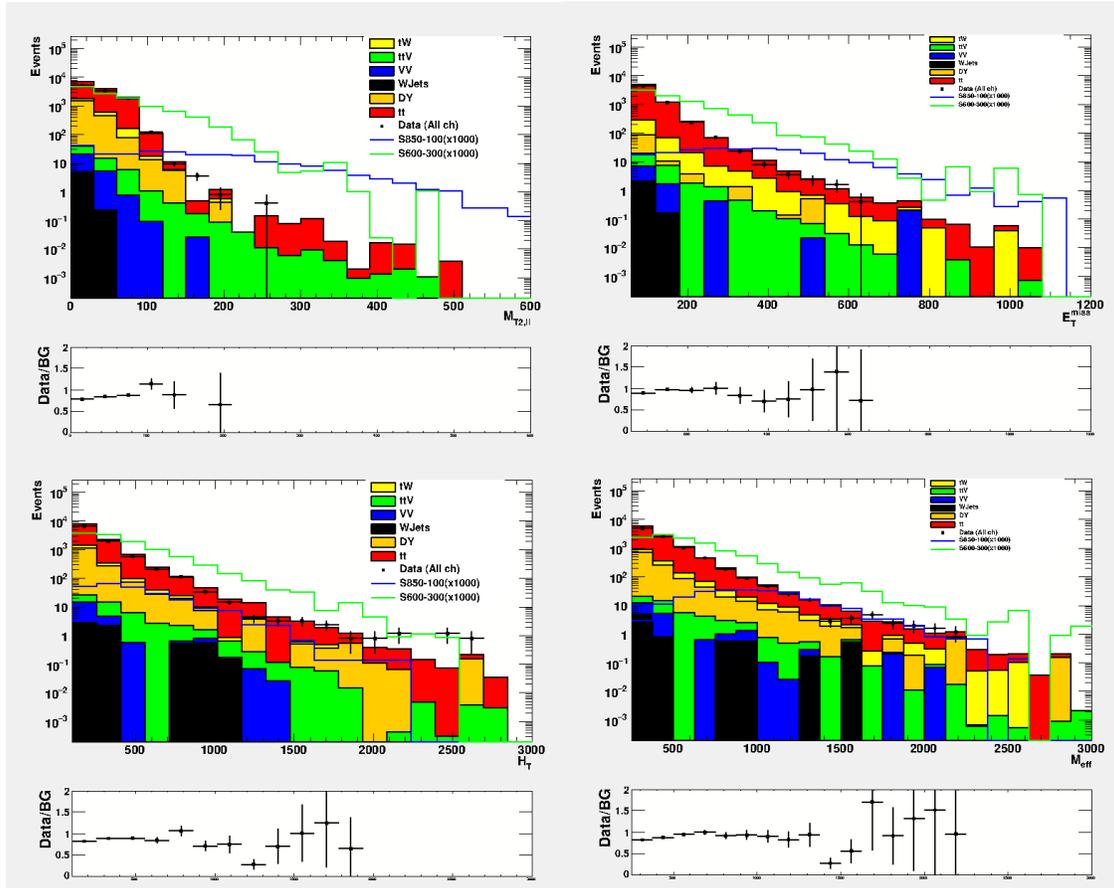


Figura 12: Distribución de algunas de las variables físicas consideradas para el entrenamiento del BDT en los datos, procesos de señal y de fondo (apilados). Los dos procesos de señal se encuentran multiplicados por un factor 1000 para poder apreciar su distribución. Incorporamos los datos a fin de observar la buena reproducción de la forma de los mismos.

completamente la señal observada al tener en cuenta la gran diferencia en sección eficaz.

Podemos observar adicionalmente variables, como  $M_{eff}$  y  $H_T$ , en cuyas distribuciones no se observan columnas libres de señal pero si se observa un desplazamiento de los picos de señal respecto a los del fondo. Con un método lo suficientemente sensible, esperaríamos que dichos desplazamientos permitiesen cierto grado de discriminación.

Por último, las dos variables angulares, así como  $M_{T2,bb}$  parecen seguir una distribución prácticamente idéntica a la del fondo, por lo que en principio resultaría difícil lograr separación alguna con cada una de ellas individualmente. No obstante, debido a posibles correlaciones con otras variables podrían dar lugar a cierta separación por lo que no las excluimos de nuestro entrenamiento por el momento.

Concluimos, entonces, que el fondo más relevante en nuestro estudio estará formado por los procesos  $t\bar{t}$ . Esperaremos, adicionalmente, que las variables  $M_{T2,u}$ ,  $\cancel{E}_{TY}$   $M_{T2,lb\bar{b}}$  sean unos discriminantes aceptables entre fondo y señal.

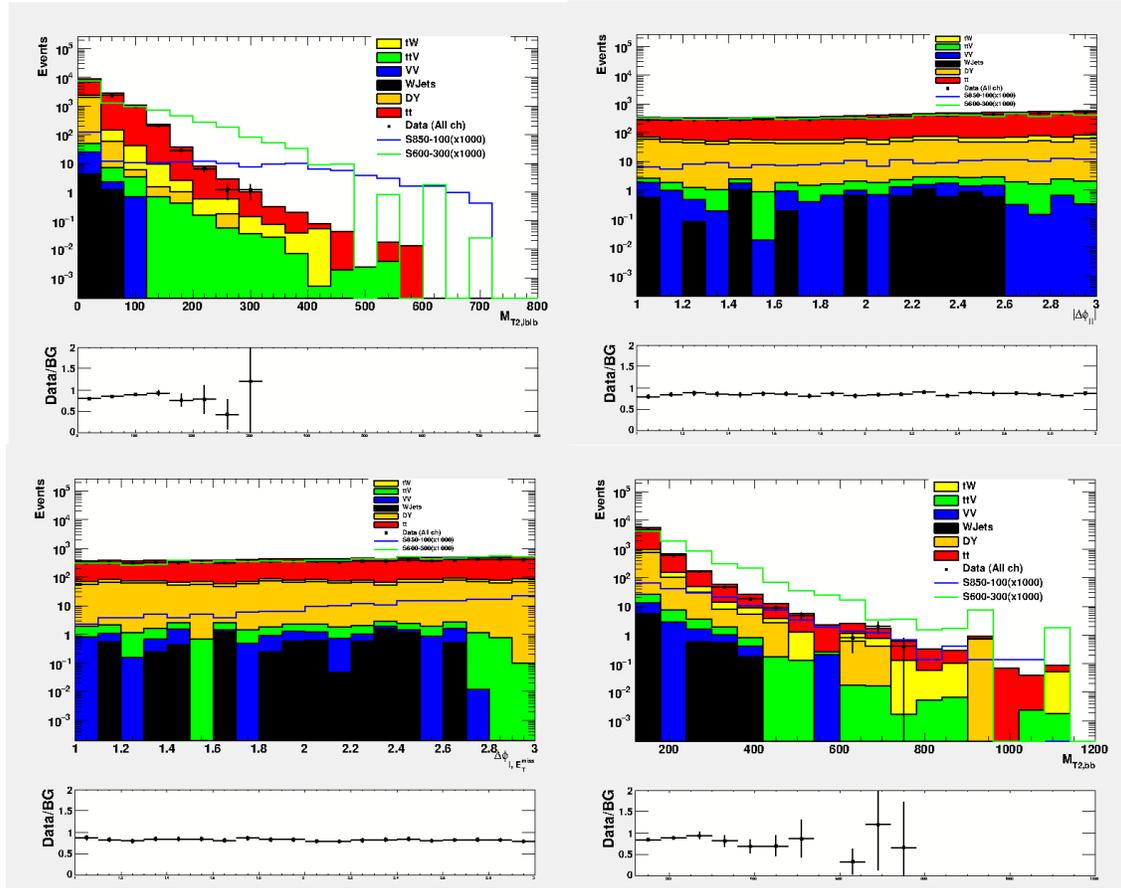


Figura 13: Distribución de algunas de las variables físicas consideradas para el entrenamiento del BDT en los datos, procesos de señal y de fondo (apilados). Los dos procesos de señal se encuentran multiplicados por un factor 1000 para poder apreciar su distribución. Incorporamos los datos a fin de observar la buena reproducción de la forma de los mismos.

### 7.2.2. Regiones de estudio

Dado lo predominantes que son los eventos de fondo sobre la señal, intentaremos definir regiones de estudio más acotadas que las definidas por nuestra preselección. Para definir estas nuevas regiones de estudio impondremos condiciones adicionales a nuestros eventos en términos de todas las variables consideradas para el evento.

La idea fundamental llegado a este punto es estudiar las combinaciones posibles de cortes secuenciales y entrenamientos de BDT que permitan la mayor separación posible. En principio, el propio BDT debería ser capaz de discriminar estos cortes por si mismo pero, como todo método multivariante, la acotación de la región de búsqueda debería mejorar su potencia de discriminación al permitirle estudiar particularidades propias de la zona de estudio. A este fin realizamos una serie de procedimientos sucesivos.

1. En primer lugar partimos de un entrenamiento básico utilizando las 15 variables definidas en la Sección

6.5 para dar lugar a un primer BDT. Los parámetros idóneos encontrados para el mismo son  $N_T = 5000$ ,  $\beta = 0,1$  y  $D = 2$ .

2. Estudiado este entrenamiento, procedemos a estudiar la posibilidad de introducir cortes en cada una de las 15 variables a través de criterios físicos (basados en condiciones cinemáticas o en la signatura esperada de la señal) así como en la distribución de las variables de interés que se observan en las Figuras 12 y 13. Cada uno de estos cortes define una nueva región de interés (una nueva selección de eventos) en la que entrenamos un nuevo BDT de dos formas distintas: incorporando la variable de corte y eliminándola.

La idea en este caso es intentar comprobar si se puede resumir adecuadamente la capacidad de separación de algunas de las variables en un único corte de forma que no sea necesario utilizarlas dentro del BDT (con lo que se simplifica el análisis de este). En caso de que esto no sea posible, buscamos al menos ampliar el ratio  $S/B$  en la región de estudio con cortes más generales de forma que el BDT pueda *centrarse* en particularidades más propias de las relaciones entre las variables difíciles de modelizar con cortes secuenciales.

3. Posteriormente, partimos de los mejores cortes para cada variable definidos en el paso anterior y definimos regiones mediante todas las combinaciones 2 a 2 posibles de dichos cortes. De nuevo, entrenamos en cada una de estas regiones el BDT incluyendo las variables correspondientes a los cortes y sin incluirlas, anotando los resultados correspondientes. En los casos más prometedores (cuando la mejora respecto al corte individual sea sustancial), realizamos un estudio de pequeñas variaciones de los cortes definidos a fin de intentar obtener el mejor resultado posible. Los resultados correspondientes a estos análisis se observan en la Tabla 2.
4. Finalmente, repetimos el procedimiento incorporando una tercera variable a los posibles cortes considerados. En este caso, intentamos seleccionamos los cortes que definieron las 10 mejores regiones en el caso anterior y los combinamos entre sí en los casos en los que alguna de las variables esté repetida (la combinatoria, en otro caso, transforma el considerar todas las posibilidades en una tarea difícil de asumir). Las características del corte que define la mejor región obtenida se encuentran resumidas en las Tablas 4 y 3.

Ahora bien, es necesario definir los *criterios de calidad* por los que nos guiamos para decidir si un entrenamiento es mejor o peor que otro. En general nos guiamos por los siguientes puntos en orden de relevancia:

1. Cada corte individual debe tener una eficiencia de señal mínima  $\epsilon_S \geq 0,5$  (cortes múltiples pueden reducir

	$M_{T2, ll}$	$M_{T2, bb}$	$\cancel{E}_T$	$M_{T2, lbb}$	$\Delta(\phi)_{l, \cancel{E}_T}$	$P_T^{llmiss}$	$\Delta(\phi)_{l, j}$	$m(\Delta(\phi)_{j, \cancel{E}_T})$
$M_{T2, ll}$	<b>30,39</b>	30,17	30,34	30,76	30,51	29,86	30,61	30,47
$M_{T2, bb}$	30,17	<b>29,71</b>	29,86	30,46	29,69	29,81	30,07	29,79
$\cancel{E}_T$	30,34	29,86	<b>29,21</b>	30,38	29,27	29,49	29,66	30,08
$M_{T2, lbb}$	30,76	30,46	30,38	<b>30,23</b>	30,32	30,35	30,43	30,62
$\Delta(\phi)_{l, \cancel{E}_T}$	30,51	29,69	29,27	30,32	<b>29,54</b>	29,46	29,55	30,39
$P_T^{llmiss}$	29,86	29,81	29,49	30,35	29,46	<b>29,53</b>	30,04	–
$\Delta(\phi)_{l, j}$	30,61	30,07	29,66	30,43	29,55	30,04	<b>29,6</b>	30,57
$m(\Delta(\phi)_{j, \cancel{E}_T})$	30,47	29,79	30,08	30,62	30,39	–	30,57	<b>30,23</b>
Corte	> 80 GeV	> 200 GeV	> 80 GeV	> 80 GeV	> 1,6 rad	> 150 GeV	< 2 rad	> 1,7 rad
Entrenada	Si	Si	Si	Si	No	Si	No	Si

Tabla 2: Valores de  $S/\sqrt{S+B}$  obtenidos para el corte óptimo en la variable BDT en las regiones definidas por los cortes secuenciales indicados en las distintas variables (simulado para un total de  $N_S = 1000$  y  $N_B = 1000$  en la muestra de test). Todas las regiones indicadas cumplen los criterios de calidad 1 y 2. En la diagonal se encuentran los correspondientes cortes en una variable, el resto de celdas se corresponden a regiones definidas por cortes en dos variables. Resaltados en verde se destacan los tres mejores cortes de una variable mientras que en azul se observan las tres mejores combinaciones posibles de cortes de dos variables. Por último, se indica el corte utilizado en cada variable y si la variable correspondiente se añade al entrenamiento del BDT. Todos los entrenamientos se realizaron con  $N_T = 1000$ ,  $\beta = 0,1$  y  $D = 2$ . No se encontró ningún corte en las variables  $m_{ll}$ ,  $\Delta(\phi)_{j, \cancel{E}_T}$ ,  $\Delta(\phi)_{llmiss, \cancel{E}_T}$ ,  $\Delta(\phi)_{ll}$ , y  $H_T$  que mejorase el entrenamiento.

ligeramente esta cantidad siempre que cada uno no la sobrepase). De esta forma evitamos una pérdida excesiva de señal en las regiones de estudio. Teniendo en cuenta que nuestra señal tiene una sección eficaz muy baja, perder una cantidad muy elevada de eventos de señal en nuestra región de trabajo requeriría de luminosidades mucho más elevadas tanto para el descubrimiento como para la exclusión de los procesos de *nueva física*.

2. El entrenamiento del BDT en la región delimitada por el corte da lugar a una respuesta *acceptable* en la prueba de test. Para ello, seguimos criterios de selección similares a los de la Sección 7.1.
3. Entre todos los cortes que cumplan los criterios 1 y 2, elegimos el que maximice la eficiencia en la separación. Para ello, consideramos el discriminante  $S/\sqrt{S+B}$ . A partir de los resultados de la prueba de test se estudia el corte óptimo en el BDT que define una región en la cual se maximiza  $S/\sqrt{S+B}$ . Además, exigimos que se mejore el poder de separación del BDT entrenado sin corte alguno.
4. De entre los cortes con resultados similares para el criterio 3, definimos como *mejor* aquel que mayor eficiencia de señal presente.

De los resultados obtenidos para los cortes secuenciales en 1 y 2 variables (Tabla 2) podemos obtener varias conclusiones relevantes.

En primer lugar, es obvio que la variable  $M_{T2,u}$  tiene un gran poder de separación en todos los casos: mejora los resultados con todos los cortes de los que forma parte en gran medida y proporciona el mejor de los mismos al imponer la condición cinemática que ya expresábamos en la Sección 6.5. Adicionalmente, es necesario mantener dicha variable en el entrenamiento del BDT, pues la información adicional que contiene su distribución proporciona un poder de separación relevante.

Del resto de variables, destaca  $M_{T2,blb}$  que proporciona también buenos resultados si eliminamos la zona de valores más bajos. Curiosamente, aunque  $\cancel{E}_T$  debería caracterizar el proceso de señal, los cortes en dicha variable no resultan especialmente útiles en este procedimiento.

Únicamente dos variables (ambas angulares) pueden ser codificadas en términos de cortes sin provocar pérdida de poder de separación al entrenar el BDT en la región definida. En los casos en los que incorporemos dichos cortes para la definición de la región prescindiremos de incluir las variables en el entrenamiento.

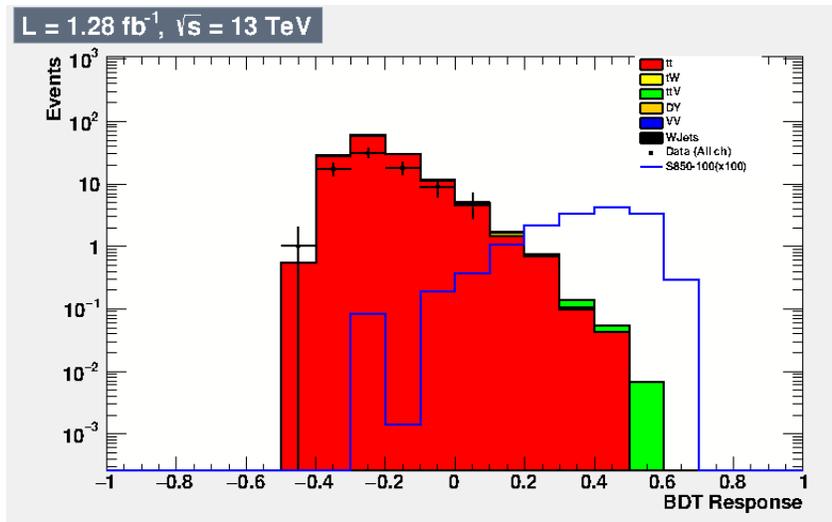


Figura 14: Punto 850–100. Distribución de la variable respuesta del BDT para los distintos procesos considerados en nuestro análisis en la región definida por  $M_{T2,u} > 80\text{GeV}$  y  $M_{T2,blb} > 80\text{GeV}$ . Los eventos de fondo se observan sumados para su mejor comparación con la señal. La señal se encuentra reescalada en un factor 100 para su observación.

Presentamos en la Figura 14, como ejemplo de los resultados obtenidos con estos cortes, la respuesta del BDT obtenida con el mejor corte doble considerado:  $M_{T2,u} > 80\text{GeV}$  y  $M_{T2,blb} > 80$ . Aún utilizando unos cortes muy restrictivos se observa que la señal (si tenemos en cuenta que está multiplicada por 100 para observar su comportamiento a bajos valores del BDT) es dominante sobre el fondo únicamente en las tres últimas columnas de la distribución (que, de hecho, ni siquiera llegan a sumar un evento total en peso). El fondo relevante en todo el proceso es fundamentalmente  $t\bar{t}$ , si bien un evento de  $ttZ$  alcanza casi el límite de la distribución. El resto de fondos no parecen proporcionar una contribución apreciable.

Por otra parte, los datos observados son, si tenemos en cuenta que únicamente representamos la incertidum-

Proceso	Eventos (antes) /Ad	Eventos (Después) / Ad	Eficiencia /Ad
$t\bar{t}$	10004	42,92	0,0043
$tW$	437	0,98	0,0022
Drell-Yan	373	0,00	0,0000
$t\bar{t}V$	54	0,34	0,0063
$VV$	10	0,00	0,0000
$W + jets$	8	0,00	0,0000
Fondo (Total)	10886	44,24	0,0041
$T2tt$ (MC, $m_{\tilde{t}} = 850$ , $m_{\chi_1^0} = 100$ )	0,38	0,08	0,1981

Tabla 3: Punto 850 – 100. Eficiencias del corte  $M_{T2,u} > 80GeV$ ,  $M_{T2,lblb} > 80GeV$ ,  $m(\Delta(\Phi)_{j,\cancel{E}_T}) > 1,7rad$  para los distintos procesos de señal y fondo (el hecho de que las muestras MC contengan eventos de peso no unitario da lugar a eventos esperados no enteros en la región de estudio).

bre estadística de los mismos y estamos obviando la sistemática que puede alcanzar valores relativos de hasta el 20 %, bastante compatibles con el fondo observado. En vista de que no obtenemos dato alguno para  $BDT > 0,1$ , no tenemos ningún indicio claro de la aparición de procesos SUSY (si bien los eventos esperados de SUSY en esta región y luminosidad no alcanzan la unidad). No obstante, los datos de los que disponemos son notablemente escasos.

De cara a lograr una mejor separación intentamos acotar aún más las regiones de estudio utilizando tres cortes secuenciales. El mejor de los resultados que conserva una forma aceptable del BDT viene dado por los cortes  $M_{T2,u} > 80GeV$ ,  $M_{T2,lblb} > 80GeV$ ,  $m(\Delta(\Phi)_{j,\cancel{E}_T}) > 1,7$ .

Desde el punto de vista físico cada uno de los cortes tiene una interpretación relativamente directa. El primero de ellos, como vimos en la Sección 6.5 se corresponde con el límite dinámico impuesto sobre  $M_{T2,u}$  en los procesos  $t\bar{t}$ . En el segundo, la misma razón nos induciría a considerar  $M_{T2,lblb} > 175GeV$ , sin embargo este corte (que es el que maximiza  $S/\sqrt{S+B}$ ) elimina demasiada señal por lo que rebajamos la condición hasta este valor. Por último,  $m(\Delta(\Phi)_{j,\cancel{E}_T})$  nos indica la producción de jets y  $\cancel{E}_T$  en direcciones opuestas (*back to back*) lo que nos indica que tienen un *boost* elevado respecto al sistema de referencia de la colisión lo que encaja con la idea de la producción a partir de partículas muy pesadas como  $\tilde{t}$ .

En las Tablas 3 y 4 así como en la Figura 15 se encuentran resumidas las propiedades principales de esta región y del BDT entrenado en la misma.

En la tabla 3 se puede observar la eficiencia de este corte para los diferentes procesos estudiados. Como se puede comprobar, el corte tiene la propiedad de eliminar completamente los fondos Drell-Yan,  $VV$  y  $W + jets$ . El resto de fondos se ven reducidos enormemente, en casi 3 órdenes de magnitud en todos los casos. Por el

Rango	Variable	Importancia
1	$\cancel{E}_T$	0,106
2	$M_{eff}$	0,091
3	$M_{T2,ll}$	0,088
4	$\Delta(\phi)_{ll}$	0,085
5	$H_T$	0,078
6	$\Delta(\phi)_{lj}$	0,075
7	$m(\Delta(\phi)_{j,\cancel{E}_T})$	0,074
8	$M_{ll}$	0,070

Tabla 4: Punto 850 – 100. Variables de mayor importancia en el BDT entrenado en la región determinada por  $M_{T2,ll} > 80 GeV$ ,  $M_{T2,lb\bar{b}} > 80 GeV$ ,  $m(\Delta(\Phi)_{j,\cancel{E}_T})$  junto con la importancia asociada a cada una de ellas.

contrario, la señal se ve reducida únicamente a la quinta parte de la original (lo que ya supone una cantidad muy pequeña de datos teniendo en cuenta la sección eficaz de la misma).

Cuantitativamente, observamos que la eficiencia de señal  $\epsilon_s = 0,198$  es unas 50 veces mayor que la eficiencia de fondo con lo que nuestros cortes secuenciales definen una región mucho más *tipo-señal* que la preselección original.

Por otra parte, en la Tabla 4 se encuentra la importancia de las variables más relevantes en el entrenamiento del BDT (las que aparecen en mayor número de ramificaciones del mismo). El primer hecho que salta a la vista es que, pese a que introdujimos un corte en  $M_{T2,ll}$ , esta variable sigue poseyendo un elevado poder de separación en la región de estudio, pues se encuentra en tercera posición. Por otra parte, las otras dos variables en las que cortamos no aparecen en esta lista, luego los cortes resumen bastante bien sus diferencias entre señal y fondo.

Observamos, adicionalmente, que la forma de variables como  $H_T$  o  $M_{eff}$  (para las cuales ningún corte específico proporciona una mejora en el resultado del entrenamiento) permite obtener bastante poder de separación, ya que se encuentran en los rangos quinto y segundo respectivamente. Este hecho nos señala la potencia adicional del BDT sobre el procedimiento de los cortes secuenciales individuales.

Por último, centrándonos en la forma final de la respuesta del BDT tenemos el resultado que se observa en la Figura 15. Si bien resulta muy similar a lo que se observa en la Figura 14 podemos notar varias mejoras.

En primer lugar, el pico de señal se encuentra más elevado (en términos relativos respecto al de fondo) que en el caso anterior. Como consecuencia, en las últimas cuatro columnas de la distribución la señal es predominante sobre el fondo. De la misma forma, la cola de señal debajo del pico de fondo prácticamente desaparece en esta región. El comportamiento de los fondos es, de todos modos, muy similar: la cola de la distribución de  $t\bar{t}$  y varios eventos aislados de  $t\bar{t}V$  son los que más dificultan el proceso de separación.

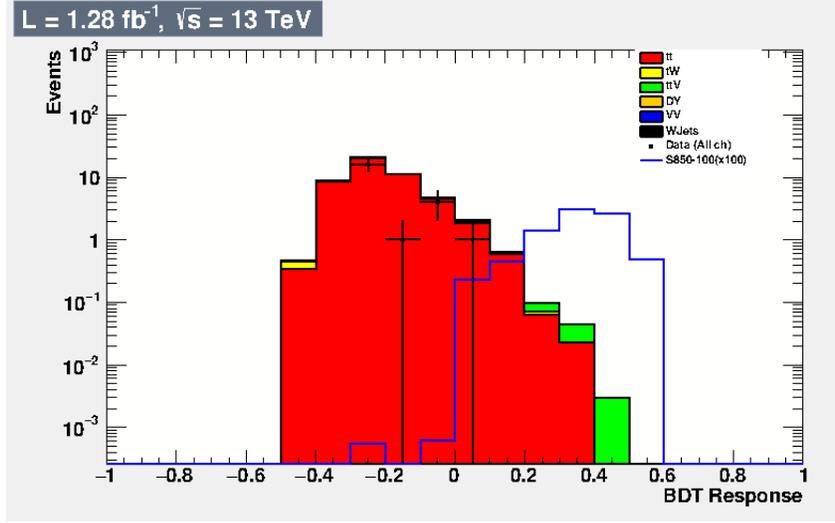


Figura 15: Punto 850–100. Distribución de la variable respuesta del BDT para los distintos procesos considerados en nuestro análisis en la región definida por  $M_{T2, ll} > 80\text{GeV}$ ,  $M_{T2, lbb} > 80\text{GeV}$  y  $m(\Delta(\phi))_{j, \cancel{E}_T} > 1,7\text{rad}$ . Los eventos de fondo se observan sumados para su mejor comparación con la señal. La señal se encuentra reescalada en un factor 100 para su observación.

La conclusión más importante a destacar en esta sección es la importancia capital de las variables tipo  $M_{T2}$ . Específicamente,  $M_{T2, ll}$  parece ser completamente relevante tanto desde el punto de vista del entrenamiento del BDT como para la definición de la región de estudio.

### 7.2.3. Proyección de resultados

Una vez optimizada la región de estudio y entrenado el correspondiente BDT en ella, podemos intentar utilizar la forma de la distribución del mismo para, aplicando los métodos estadísticos de la Sección 6.3 buscamos estudiar si hemos logrado excluir la presencia de señal en este punto del espacio paramétrico o si, por el contrario, necesitaremos una mayor luminosidad para lograrlo.

Dado que la cantidad de datos que tenemos es muy pequeña, nos será necesario utilizar el método *Full CLs* para el tratamiento de las imprecisiones estadísticas y sistemáticas. No obstante, y debido a la carga computacional que este acarrea, realizaremos también el estudio a través del método asintótico a fin de observar las posibles diferencias entre ambos y si, al menos, el método asintótico proporciona una aproximación razonable en nuestras condiciones para la posterior sustitución de uno por el otro en caso de necesidad.

Desde el punto de vista práctico, consideraremos las incertidumbres estadísticas de tipo Poissoniano (es decir, en una columna con  $n$  entradas, el error estadístico es  $\sqrt{n}$ ). En cuanto a los errores sistemáticos realizaremos una aproximación enormemente conservadora hacia los mismos asumiendo un error relativo del 20% en la luminosidad, del 15% en las muestras de señal y del 20% en las de fondo modelizados por una función de

distribución logarítmico-normal. Aplicaremos estos criterios a lo largo de todo el estudio.

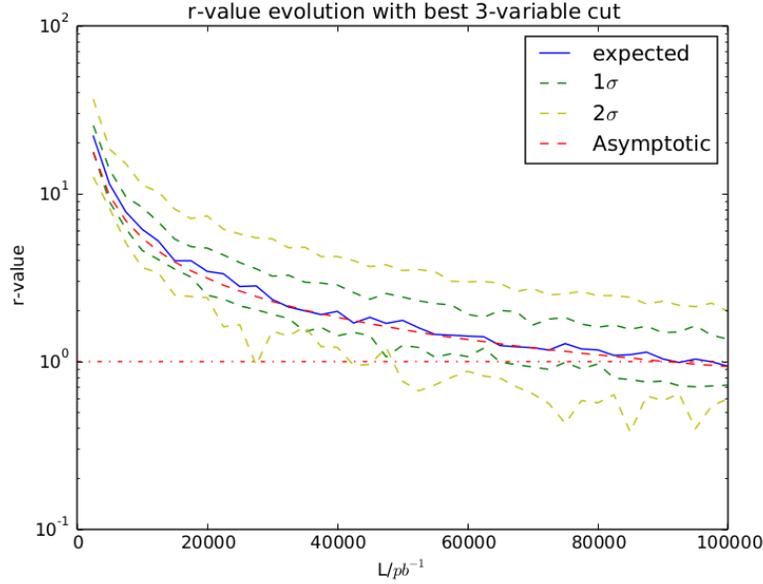


Figura 16: Punto 850 – 100. Evolución respecto a la luminosidad del límite de confianza al 95 % para el valor  $r$ , intensidad de señal. Valores observados representados tanto para el método asintótico como para el *Full CLs*. Para este último se encuentran dibujados además los límites de confianza  $\pm\sigma$  y  $\pm 2\sigma$ . En ambos métodos se alcanza la luminosidad de exclusión en torno a  $L = 87 fb^{-1}$ .

Los resultados que se observan en la Figura 16 nos permiten comprobar que ambos métodos tienen, al menos en el punto de estudio, comportamientos notablemente similares. Si bien el método *Full CLs* es más irregular, posiblemente por utilizar una cantidad de *toys* demasiado pequeña, la luminosidad de exclusión hallada es casi idéntica:  $87 fb^{-1}$ . Observemos que, a la luminosidad correspondiente a los datos que tenemos ( $1,28 fb^{-1}$ ), el límite en  $r$  al nivel de confianza del 95 % está lejos de proporcionarnos exclusión alguna pues es del orden de  $r_0 \approx 30$ .

#### 7.2.4. Poda de variables

Un procedimiento típico de simplificación para el BDT es la *poda* o *pruning*. La idea detrás del *pruning* es eliminar ciertas secciones del proceso que resultan redundantes o bien proporcionan escasa información de cara al poder de separación del BDT. En nuestro caso particular nos restringiremos a realizar una *poda de variables*, es decir, intentaremos eliminar las variables de menor relevancia en el BDT. De realizar el procedimiento adecuadamente es relativamente probable que disminuye la adaptación a la muestra (*overtraining*) lo que nos permitirá modificar el resto de los parámetros del BDT para compensar la pérdida de poder predictivo producida al desechar variables.

En este caso partimos del entrenamiento realizado utilizando únicamente los cortes de la preselección. A

partir de este, eliminamos la variable con menor importancia en el entrenamiento y volvemos a realizar un entrenamiento con las variables resultantes. Repetimos el procedimiento hasta que la importancia final de las variables sea relativamente similar. Consideramos el conjunto de variables resultantes final y realizamos un procedimiento de análisis por regiones como el descrito en la sección anterior utilizando las mejores regiones encontradas en nuestro estudio previo.

La principal diferencia en el entrenamiento del BDT es que, al disminuir el *overtraining* total, podemos aumentar el número de árboles totales sin producir excesivas cantidades de este. Tomamos entonces como parámetros del BDT  $N_T = 10000$ ,  $\beta = 0,1$  y  $D = 2$ .

El mejor resultado tras la poda se obtiene, como antes, al tomar el corte triple  $M_{T2, ll} > 80 GeV$ ,  $M_{T2, lbb} > 80 GeV$  y  $m(\Delta(\phi))_{j, \cancel{E}_T} > 1,7 rad$ . El entrenamiento del BDT se realizó considerando las variables<sup>40</sup>  $M_{T2, ll}$ ,  $M_{T2, lbb}$ ,  $H_T$ ,  $\cancel{E}_T$  y  $M_{eff}$ .

Los resultados del test del BDT se pueden observar en la Figura 17. En la Figura 18 se observan los resultados de la proyección en una búsqueda de la luminosidad necesaria para la exclusión.

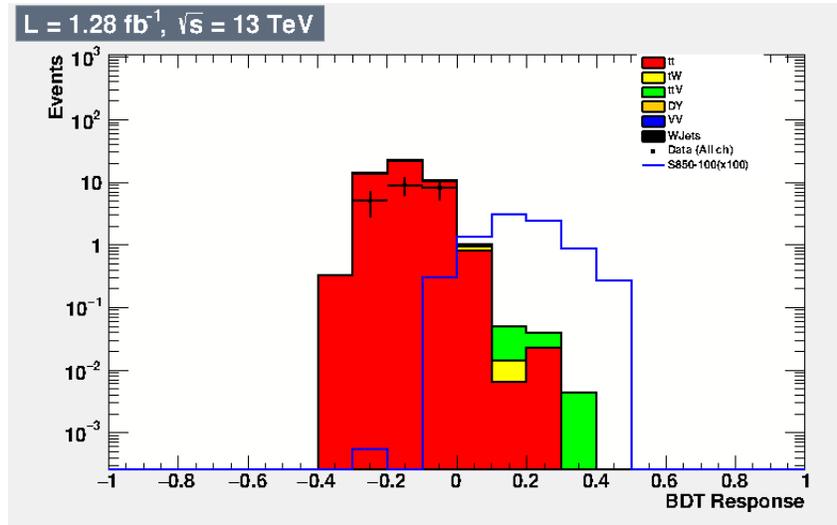


Figura 17: Punto 850–100. Distribución de la variable respuesta del BDT para los distintos procesos considerados en nuestro análisis con 5 variables en la región definida por  $M_{T2, ll} > 80 GeV$ ,  $M_{T2, lbb} > 80 GeV$  y  $m(\Delta(\phi))_{j, \cancel{E}_T} > 1,7 rad$ . Los eventos de fondo se observan sumados para su mejor comparación con la señal. La señal se encuentra reescalada en un factor 100 para su observación.

Comparando las Figuras 17 y 15 podemos observar un comportamiento muy similar de las dos. De nuevo, el fondo mayoritario es  $t\bar{t}$  durante toda la distribución salvo a altos valores de la respuesta del *BDT* para los cuales aparece un fondo residual  $ttV$ . No parece haber diferencias sustanciales entre un entrenamiento y otro salvo quizá una cola derecha de señal ligeramente más cercana al fondo en este segundo.

<sup>40</sup>Destaquemos que tras el proceso de poda, la importancia de todas las variables es notablemente similar, muy cercana a 0,2 en todos los casos.

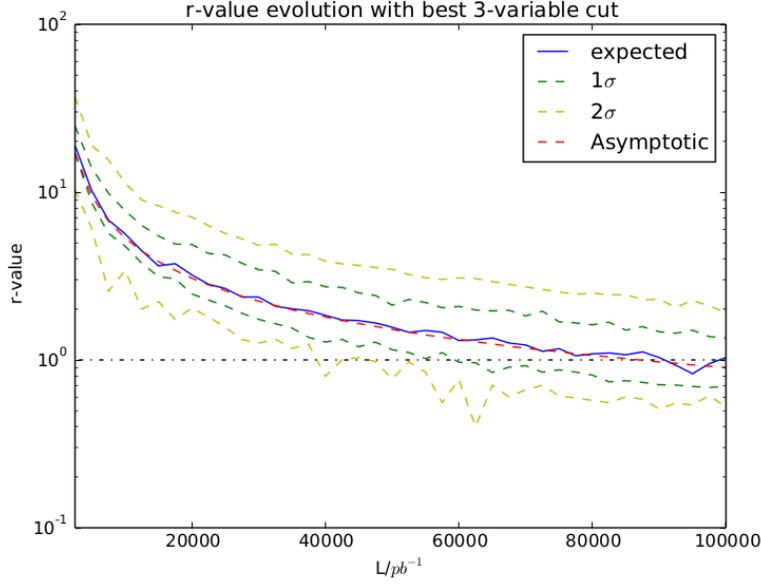


Figura 18: Punto 850 – 100. Evolución respecto a la luminosidad del límite de confianza al 95 % para el valor  $r$ , intensidad de señal. Valores observados representados tanto para el método asintótico como para el *Full CLs*. Para este último se encuentran dibujados además los límites a distancia  $\pm\sigma$  y  $\pm 2\sigma$ . En el método asintótico se observa exclusión a  $L = 87 fb^{-1}$ , mientras que en el *Full CLs* se da en torno a  $L = 90 fb^{-1}$ .

Por otra parte, estudiando el comportamiento de las Figuras 18 y 16 las diferencias son mínimas. La luminosidad necesaria para la exclusión es prácticamente idéntica con lo que el procedimiento de *poda* no parece mejorar nuestros resultados.

### 7.3. Procesos de entrenamiento y test

Hasta el momento hemos asumido implícitamente que dos fenómenos de nuestro entrenamiento tienen un comportamiento adecuado.

En primer lugar, hemos asumido que la estadística (el total de eventos generados) en las muestras de señal es lo suficientemente elevada como para la correcta modelización del comportamiento de las mismas.

Por otra parte, hemos asumido que la dinámica en las diagonales del espacio de parámetros del modelo  $T2tt$  es relativamente similar con el objetivo de mejorar la estadística de nuestra muestra de entrenamiento. Si bien parece lógico pensar que la dinámica de las partículas en cada evento esté determinada relativamente bien por la *energía libre* máxima accesible a los procesos SM del evento, esta hipótesis podría no ser cierta.

Comprobamos el correcto funcionamiento de estos dos procesos en dos puntos alternativos del espacio de parámetros que se prestan específicamente a estos análisis (por disponer de muestras con bastantes eventos simulados y diagonales con múltiples puntos de los que disponemos de varias muestras. Adicionalmente, estos

análisis nos permitirá observar lo bien que se adapta nuestro estudio a otros puntos posibles en el modelo  $T2tt$ .

En primer lugar, consideramos el punto  $m_{\tilde{t}} = 750\text{GeV}$ ,  $m_{\chi_1^0} = 200\text{GeV}$ , en el que disponemos de una estadística total bastante elevada: de cerca de  $N = 2000$  eventos simulados.

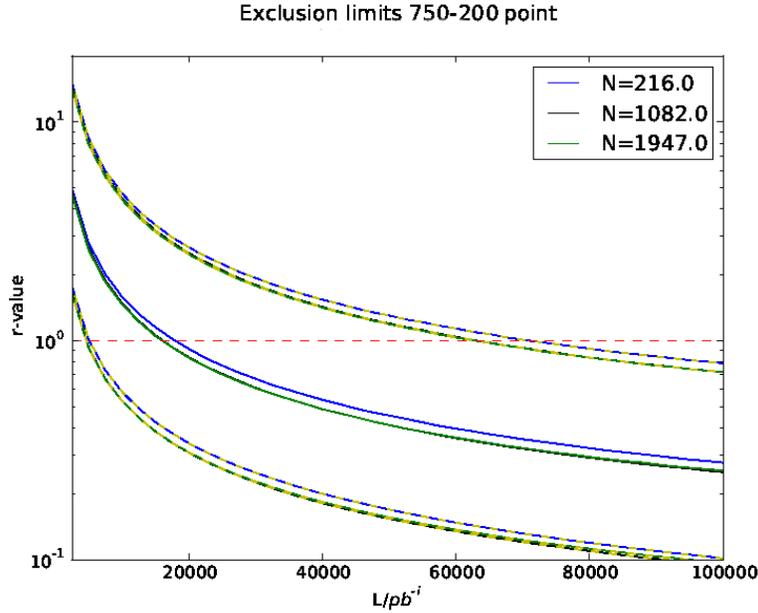


Figura 19: Punto 750 – 200. Evolución del límite superior para la intensidad de la señal ( $95\%CL$ , tratamiento asintótico) en función del número de eventos en la muestra de test de la señal. Se alcanza luminosidad de exclusión a  $15\text{fb}^{-1}$  para las muestras de  $N = 1082$  y  $N = 1947$  eventos, para  $N = 216$  son necesarios  $17\text{fb}^{-1}$ . Las curvas de colores alternos se corresponden con los valores de  $r$  a un incertidumbre  $2\sigma$ .

Realizamos un entrenamiento del BDT utilizando los parámetros óptimos encontrados para el punto  $m_{\tilde{t}} = 850\text{GeV}$ ,  $m_{\chi_1^0} = 100\text{GeV}$  (que obviamente no tienen por qué corresponderse con los de este punto, pero podemos considerar similares). Posteriormente, aplicamos el BDT obtenido finalmente sobre subconjuntos de la muestra de señal de diferente tamaño,  $N$ , obtenidos mediante un algoritmo tipo *bootstrap* (es necesario normalizar de nuevo los pesos de los eventos a posteriori). Sobre cada uno de los resultados realizamos el análisis estadístico (usamos en este caso el método asintótico como primera aproximación, en vista de su buen funcionamiento anterior y para ahorrar tiempo total de computación). Los resultados para algunas de las submuestras (consideramos un total de 10 casos posibles) se pueden observar en la Figura 19.

Como podemos observar, no se obtiene ganancia alguna al hacer la transición de  $N = 1082$  a  $N = 1947$ , si bien para  $N = 216$  los resultados son bastante peores (se pierde aproximadamente un  $10\%$  en la luminosidad de exclusión). Obviamente, no podemos trasladar este resultado en cifras absolutas al punto  $m_{\tilde{t}} = 850\text{GeV}$ ,  $m_{\chi_1^0} = 100\text{GeV}$ , ya que tiene una sección eficaz mucho mayor (lo que da cuenta de la reducida luminosidad necesaria para la exclusión). Sin embargo, sí que podemos inferir el comportamiento relativo debido al tamaño de la

muestra utilizada, pues el tratamiento estadístico es relativamente similar.

Concluiríamos que con  $N = 1000$  obtenemos una reproducción bastante buena del comportamiento del proceso  $T2tt$ . Esto reafirma el uso que hicimos en el punto inicial en la sección anterior (teniendo en este una muestra de  $N \approx 1500$ ).

Por otra parte, para estudiar la posible influencia de considerar los distintos puntos en la diagonal utilizaremos el punto  $m_{\tilde{t}} = 850\text{GeV}$ ,  $m_{\chi_1^0} = 350\text{GeV}$ . En este caso tenemos un total de  $TSS = 8$  muestras distintas correspondientes a puntos en dicha diagonal (incluyendo el propio punto de test) de forma que podremos observar el efecto de considerar más o menos puntos en el entrenamiento.

Esperamos que los puntos más próximos a  $m_{\tilde{t}} = 850\text{GeV}$ ,  $m_{\chi_1^0} = 350\text{GeV}$  tengan un comportamiento más parecido a este (ya que los parámetros del modelo simplificado son más similares). Adicionalmente, considerar una cantidad menor de puntos disminuye la estadística total de la muestra de entrenamiento, lo que empeora la calidad del mismo.

El procedimiento, en este caso, se basa en realizar varios entrenamientos de BDTs considerando distintas cantidades de puntos en la diagonal para conformar la muestra de entrenamiento. Para cada uno de los BDTs resultantes estudiamos tanto su forma (20) como un tratamiento asintótico de su proyección a altas luminosidades que nos proporcione una luminosidad de exclusión esperada (21). Como en el caso anterior, realizamos el entrenamiento utilizando los parámetros óptimos obtenidos  $m_{\tilde{t}} = 850\text{GeV}$ ,  $m_{\chi_1^0} = 100\text{GeV}$ .

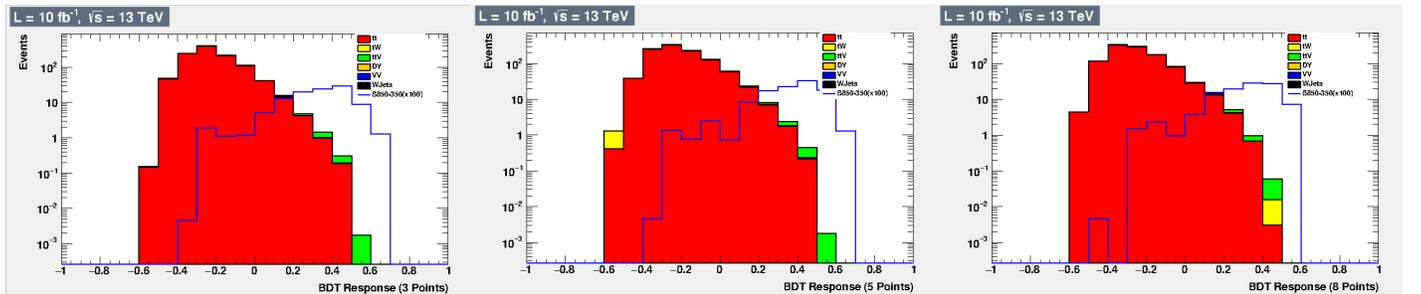


Figura 20: Punto 850–350. Respuesta del BDT entrenado utilizando muestras de entrenamiento correspondientes a una cantidad distinta de puntos en la diagonal  $m_{\tilde{t}} - m_{\chi_1^0} = 500\text{GeV}$ . Respectivamente 3 (Izquierda), 5 (Centro) y 7 (Derecha). Nótese en todos los casos la similitud de la forma global de la respuesta del BDT.

En la Figura 20 se puede observar cómo la forma final de la respuesta del BDT es bastante similar en los casos  $TSS = 3$  y  $TSS = 5$ , con dos columnas relativamente libres de fondo y una forma bastante parecida en la distribución de la señal. En la muestra  $TSS = 8$  la forma es ligeramente peor en cuanto a que aparece una cantidad de fondo ligeramente superior en una de las columnas que previamente era casi libre de este.

Por otra parte, observando la Figura 21 podemos ver cómo el total de puntos considerados afecta directamente a la luminosidad de exclusión lograda. El caso  $TSS = 3$  parece proporcionar resultados relativamente peores

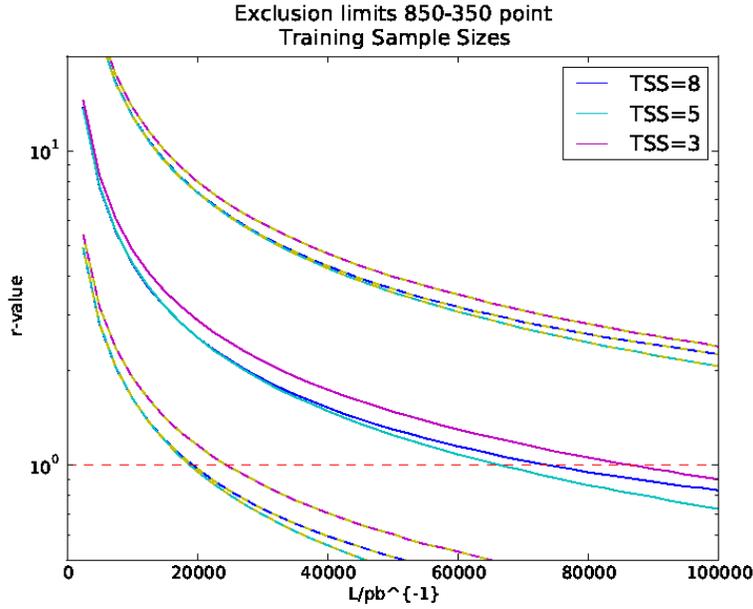


Figura 21: Punto 850 – 350. Evolución del límite superior para la intensidad de la señal (95 % $CL$ , tratamiento asintótico) en función del número de puntos considerados en el entrenamiento del BDT. Se alcanzan luminosidades de exclusión a  $87fb^{-1}$  ( $TSS = 3$ ),  $65fb^{-1}$  ( $TSS = 5$ ) y  $75fb^{-1}$  ( $TSS = 8$ ).

que los otros dos, posiblemente debido a la baja estadística en la muestra de entrenamiento. Por otra parte, al considerar  $TSS = 8$  estamos incluyendo muestras con valores de  $m_{\tilde{t}}$  y  $m_{\chi_1^0}$  ligeramente alejados de los de la señal lo que justifica que tengan una dinámica relativamente distinta al punto de señal y, en última instancia, proporcione una peor luminosidad de exclusión.

Concluimos que, en principio, para los puntos más cercanos en la diagonal la dinámica es lo suficientemente similar al punto de señal como para que la forma del BDT no se vea especialmente alterada. Sin embargo, los puntos más alejados pueden empeorar este comportamiento. Adicionalmente, considerar pocas muestras para el entrenamiento ( $TSS = 3$ ) también empeora la estadística arrojando peores resultados.

#### 7.4. Punto 600 – 300

De forma complementaria al estudio del punto  $m_{\tilde{t}} = 850GeV$ ,  $m_{\chi_1^0} = 100GeV$  buscamos observar la posibilidad de realizar un tratamiento similar en otra zona del espacio de parámetros que tampoco ha sido excluida. Para ello, repetimos todo el desarrollo anterior para el punto  $m_{\tilde{t}} = 600$ ,  $m_{\chi_1^0} = 300$ , con dos diferencias fundamentales respecto al anterior. Por una parte, la sección eficaz de este punto es unas 10 veces más elevada que en el anterior, luego esperamos que este hecho ayude al proceso de detección o exclusión.

Por otra parte, la diferencia de masas entre las dos partículas es bastante menor, por lo que la energía restante que podrán obtener los quarks  $t$  producidos será menor y, por tanto, el resultado debería tener una dinámica más similar a la del proceso  $t\bar{t}$  (exceptuando la alta presencia de  $\cancel{E}_T$ ).

Proceso	Eventos (antes) /Ad	Eventos (Después) / Ad	Eficiencia /Ad
$t\bar{t}$	10004.00	121.80	0.012
$tW$	437.00	2.78	0.0064
Drell-Yan	373.00	0.00	0.0000
$ttV$	54.00	0.88	0.0163
$VV$	10.00	0.00	0.0000
$W + jets$	8.00	0.00	0.0000
Fondo (Total)	10886.00	125.45	0.0115
$T2tt$ (MC, $m_{\tilde{t}} = 600$ , $m_{\chi_1^0} = 300$ )	3.10	0.89	0.2870

Tabla 5: Punto 600 – 300. Eficiencias del corte  $M_{T2,\ell} > 80GeV$ ,  $M_{T2,\ell b\ell b} > 80GeV$  para los distintos procesos de señal y fondo.

Nos centramos directamente en el mejor de los entrenamientos conseguidos que, en este caso, se da para  $N_T = 2000$ ,  $D = 2$ ,  $\beta = 0,1$  (parámetros del BDT) y la región definida por los cortes  $M_{T2,\ell} > 80GeV$  y  $M_{T2,\ell b\ell b} > 80GeV$  (ninguno de los conjuntos de tres cortes simultáneos mejoró de forma sensible los resultados de esta selección).

En la Tabla 5 se puede observar una breve recopilación de cómo afectan los cortes de selección a las muestras consideradas.

Como era obvio, los cortes resultan menos estrictos que en el caso del punto anterior. No obstante, se consigue reducir enormemente la presencia del fondo en nuestra región de estudio (se reduce el número de eventos en 2 órdenes de magnitud) mientras que perdemos únicamente las 2/3 partes de la señal. Desde el punto de vista cuantitativo, logramos aumentar la relación  $S/B$  en un factor 24. De todas formas, los eventos de fondo esperados siguen siendo ampliamente mayoritarios respecto a los de señal. Como en ocasiones anteriores, el fondo más relevante continúa siendo  $t\bar{t}$ , con una cantidad residual de  $ttV$  y  $tW$ .

Desde el punto de vista del entrenamiento, observamos en primer lugar la importancia de las variables utilizadas en la Tabla 6.

Como nos ocurría anteriormente, la variable  $M_{T2,\ell}$  sigue siendo fundamental para lograr la separación. Aún cuando forma parte del corte de selección de nuestra región, la forma de su distribución resulta fundamental para discriminar nuestros procesos de señal. Aunque en menor medida, el corte en  $M_{T2,\ell b\ell b}$  tampoco reúne toda la información que puede proporcionar la variable, pues sigue siendo relevante en el entrenamiento.

En comparación con la Tabla 4, se observa que  $\cancel{E}_T$  resulta mucho menos relevante en este caso, posiblemente debido a que los neutrinos que se generan están menos *boosteados* al provenir de un quark  $t$  menos energético. Notemos que, aunque en un orden distinto, las variables de relevancia se corresponden bastante bien con las que obteníamos entonces, por lo que podemos intuir que caracterizan al proceso  $T2tt$  de una forma relativamente

Rango	Variable	Importancia
1	$M_{T2, ll}$	0,117
2	$\Delta(\phi)_{ll}$	0,091
3	$\Delta(\phi)_{lj}$	0,090
4	$M_{eff}$	0,080
5	$\Delta(\phi)_{l\cancel{E}_T}$	0,077
6	$M_{T2, lbb}$	0,067
7	$m(\Delta(\phi)_{j\cancel{E}_T})$	0,066
8	$\cancel{E}_T$	0,064

Tabla 6: Punto 600 – 300. Variables de mayor importancia en el BDT entrenado en la región determinada por  $M_{T2, ll} > 80 GeV$ ,  $M_{T2, lbb} > 80 GeV$  junto con la importancia asociada a cada una de ellas.

global respecto a los fondos considerados.

En cuanto a la distribución de la respuesta del BDT podemos observar los resultados en la Figura 22.

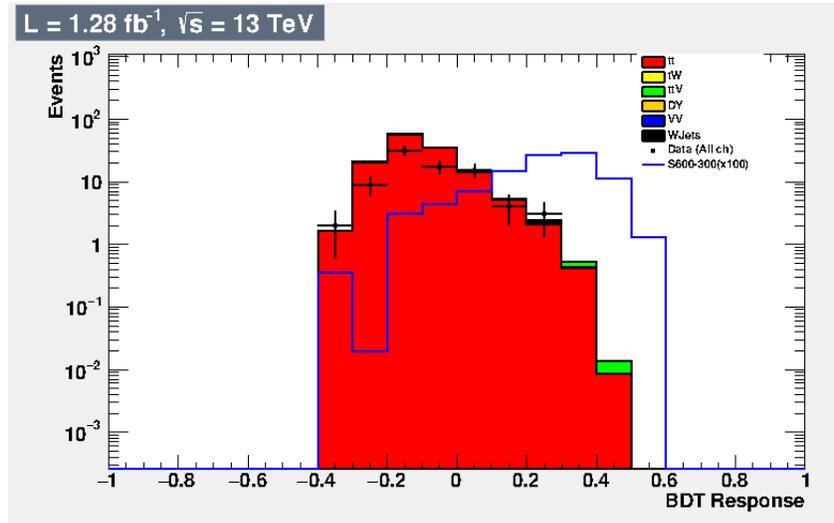


Figura 22: Punto 600–300. Distribución de la variable respuesta del BDT para los distintos procesos considerados en nuestro análisis con 5 variables en la región definida por  $M_{T2, ll} > 80 GeV$  y  $M_{T2, lbb} > 80 GeV$ . Los eventos de fondo se observan sumados para su mejor comparación con la señal. La señal se encuentra reescalada en un factor 100 para su observación.

Por pura comparación con el caso de la Figura 14 (definida en la misma región) observamos que en este nuevo punto el pico de señal se encuentra más elevado que en el anterior, posiblemente debido a la mayor sección eficaz de la que disponemos. Obviando este hecho, la forma obtenida para ambos es relativamente similar, apareciendo de nuevo esa pequeña cantidad de fondo residual  $ttV$  en la cola de la distribución. Como en casos anteriores, los datos medidos no parecen mostrar rastro alguna de señal (si bien el tamaño de la muestra es demasiado pequeño para dar una respuesta clara).

### 7.4.1. Proyección de resultados

Por último, incorporamos los errores estadístico y sistemático a nuestro estudio y realizamos la extensión de los resultados como en ocasiones anteriores. Nos centraremos en este caso en el tratamiento asintótico (que ya comprobamos que proporciona unos resultados aceptables en regiones con menor cantidad de datos).

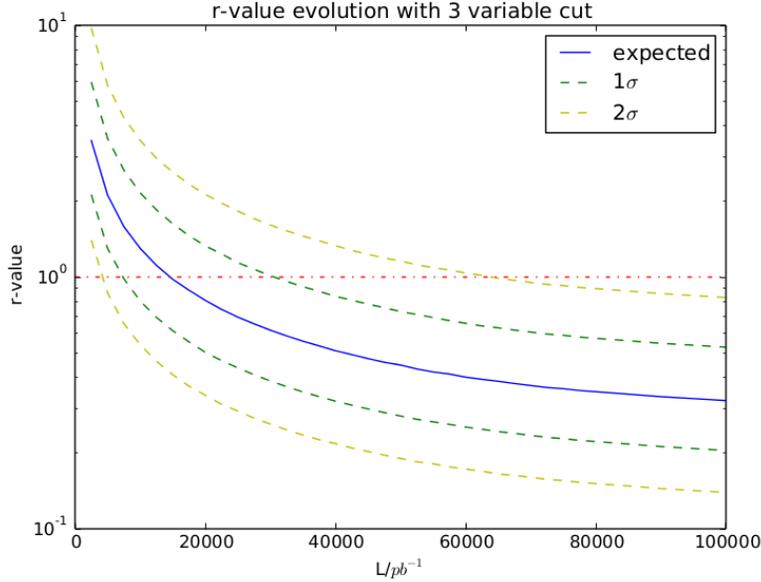


Figura 23: Punto 600 – 300. Evolución respecto a la luminosidad del límite de confianza al 95 % para el valor  $r$ , intensidad de señal, método asintótico. Se alcanza la luminosidad de exclusión a  $L = 18fb^{-1}$

Por comparación entre las Figuras 23 y 16 observamos con claridad que los límites de exclusión mejoran enormemente al considerar puntos con una sección eficaz bastante más elevada, pues logramos mejorar en casi un factor 4 el límite necesario para la exclusión.

No obstante, este sigue sin ser lo suficientemente bajo como para poder garantizar nada a la luminosidad a la que estamos trabajando, por lo que seguiríamos necesitando una cantidad de datos adicional para poder tomar una decisión concreta sobre la presencia de  $T2tt$  con las características de este punto.

### 7.5. Resultados globales

Del estudio de los parámetros del BDT hemos deducido una región óptima para cada uno de ellos que minimiza el proceso de *overtraining* sin ralentizar el funcionamiento del propio BDT.

Hemos podido comprobar que, en las condiciones de las muestras disponibles, la cantidad de eventos por muestra es estadísticamente significativa para reproducir su comportamiento (a través del límite encontrado

$N > 1000$ ).

Adicionalmente, hemos observado cómo se logra un equilibrio entre el total de datos óptimo para la muestra de entrenamiento (a partir del uso de puntos en las diagonales) y de reproducción adecuada de las propiedades de un punto a través de datos de la diagonal utilizando los dos puntos más próximos al de test. Dicho hecho es utilizado para designar las muestras de entrenamiento en los dos estudios (de hecho en  $850 - 100$  no tenemos otra opción, no tenemos más).

Finalmente, los entrenamientos propiamente dichos para los puntos  $850 - 100$  y  $600 - 300$  nos han permitido encontrar la luminosidad esperada a la cual podríamos observar la exclusión de procesos  $T2tt$  en los puntos correspondientes. Dichas luminosidades son, correspondientemente,  $87fb^{-1}$  y  $18fb^{-1}$ . Observamos en ambos casos que son necesarias observaciones de mayor luminosidad que la disponible tanto para garantizar exclusión como detección de procesos  $T2tt$ .

Durante el proceso de entrenamiento y selección de regiones en ambos puntos tenemos que destacar la importancia de dos variables de importancia:  $M_{T2,ll}$  y  $M_{T2,lb\bar{b}}$ . Ambas presentan un poder de discriminación notablemente elevado y representan una de las mejores condiciones físicas con un comportamiento distinto en señal y fondo.

Proceso	Eventos (antes) /Ad	Eventos (Después) / Ad	Eficiencia /Ad
$t\bar{t}$	10004.00	109.05	0.011
$tW$	437.00	4.94	0.011
Drell-Yan	373.00	0.00	0.000
$t\bar{t}V$	54.00	2.11	0.039
$VV$	10.00	0.03	0.003
$W + jets$	8.00	0.00	0.000
Fondo (Total)	10886.00	116.13	0.011
$DM$ ( $m_\chi = 1000GeV$ )	10.165	6.263	0.616

Tabla 7:  $DM$ . Eficiencias del corte  $M_{T2,l} > 80GeV$ ,  $M_{T2,blb} > 80GeV$  para los distintos procesos de señal y fondo.

## 8. EFT para DM

Por último, podemos aprovechar los métodos desarrollados en el análisis del canal  $T2tt$  de SUSY para búsquedas de nueva física alternativa. En nuestro caso estudiaremos una posible EFT centrada en la descripción de una partícula de materia oscura  $\chi$  con  $m_\chi = 1000GeV$ , asumiendo una sección eficaz del proceso  $t\bar{t} \rightarrow \chi\chi$  de  $\sigma = 1pb$  (que es aproximadamente la que esperamos en la búsqueda de la WIMP). Denominaremos a este proceso  $tt + DM$  o, simplemente,  $DM$ .

No tenemos un procedimiento equivalente a la idea de utilizar los puntos en la diagonal del espacio de parámetros para realizar el entrenamiento, por lo que tendremos que partir la muestra de señal en 2 partes de las cuales una irá destinada al proceso de entrenamiento y la otra al test y análisis. Por lo demás, el procedimiento seguido es completamente análogo al que llevamos a cabo en nuestro estudio de los procesos  $T2tt$ .

Presentamos únicamente el mejor de los resultados obtenidos. Este se consiguió utilizando los cortes de selección  $M_{T2,l} > 80GeV$ ,  $\cancel{E}_T > 100GeV$  y  $M_{eff} > 400GeV$  para definir la región de trabajo. Obsérvese que todos ellos responden directamente a las características físicas del proceso. Ya mencionamos anteriormente el límite cinemático que se da en  $M_{T2,l}$  en  $m_W$ . Por otra parte, las partículas de materia oscura  $\chi$  son invisibles al detector y notablemente masivas, por lo que esperamos una alta cantidad de energía faltante que, a su vez, contribuye tanto a  $\cancel{E}_T$  como a  $M_{eff}$ .

Como en ocasiones anteriores, comenzamos estudiando la eficiencia de estos cortes en señal y fondo (Tabla 7).

Como podemos observar, el proceso estudiado presenta una cantidad de eventos ligeramente mayor que los que analizamos de tipo  $T2tt$ . No obstante, recordemos que la sección eficaz del proceso es un parámetro del modelo y, por tanto, este dato no resulta especialmente significativo.

Como en casos anteriores, en la región de estudio el fondo está dominado completamente por los procesos

Rango	Variable	Importancia
1	$\cancel{E}_T$	0,135
2	$M_{T2,ll}$	0,095
3	$H_T$	0,092
4	$\Delta(\phi)_{l\cancel{E}_T}$	0,091
5	$\Delta(\phi)_{ll}$	0,086
6	$m(\Delta(\phi)_{j\cancel{E}_T})$	0,078
7	$m_{ll}$	0,069
8	$M_{T2,tbb}$	0,064

Tabla 8: *DM*. Variables de mayor importancia en el BDT entrenado en la región determinada por  $M_{T2,ll} > 80\text{GeV}$ ,  $\cancel{E}_T > 100\text{GeV}$  y  $M_{eff} > 400\text{GeV}$  junto con la importancia asociada a cada una de ellas.

$t\bar{t}$ , siendo casi inexistentes los eventos correspondientes a otras muestras en la región de estudio, exceptuando la presencia de  $tW$  y  $ttV$ .

Observamos que los cortes definidos son notablemente eficientes en la discriminación pues únicamente perdemos un 40 % de la señal mientras que reducimos el número de eventos de fondo en casi dos órdenes de magnitud. Cuantitativamente logramos amplificar  $S/B$  en un factor 56 con una pérdida bastante pequeña de señal.

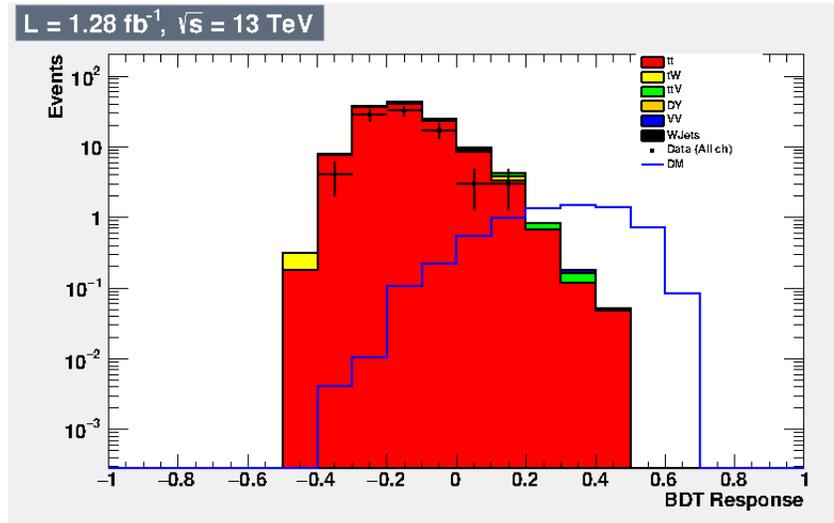


Figura 24: *DM*. Distribución de la variable respuesta del BDT para los distintos procesos considerados en nuestro análisis en la región definida por  $M_{T2,ll} > 80\text{GeV}$ ,  $\cancel{E}_T > 100\text{GeV}$  y  $M_{eff} > 400\text{GeV}$ . Los eventos de fondo se observan sumados para su mejor comparación con la señal. A diferencia del resto de casos, la señal **no** se encuentra reescalada.

Como en casos anteriores, nos es de especial interés estudiar la relevancia de las diferentes variables en el entrenamiento del BDT. Los resultados que se observan en la Tabla 8 nos indican, de nuevo, que pese a utilizar  $\cancel{E}_T$  y  $M_{T2,ll}$  para realizar los cortes que definen la región de estudio estas dos variables siguen siendo decisivas de

cara a discriminar entre los procesos de señal y fondo. Sin embargo, sí que conseguimos expresar adecuadamente el poder de separación de  $M_{eff}$  mediante nuestro corte estricto.

Si comparamos las variables de esta tabla con las de la Tabla 4 es inmediato comprobar una similitud que, en cierto modo, caracteriza los procesos de *nueva física* estudiados: en ambos casos son especialmente relevantes  $\cancel{E}_T$  y  $M_{T2,u}$ .

En la Figura 24 se puede observar la forma de la distribución obtenida en la prueba de test del BDT. Destaca inmediatamente el hecho de que, al contrario que en los casos anteriores, sí que esperamos observar una cantidad real de eventos de señal en nuestros datos. Sin embargo, como podemos observar, las zonas donde estos están predichos ( $BDT > 0,2$ ) no presentan registro de datos alguno. Este hecho es, de por sí, una buena razón para argumentar que es posible la exclusión de este proceso con los datos de los que disponemos.

## 8.1. Proyección de resultados

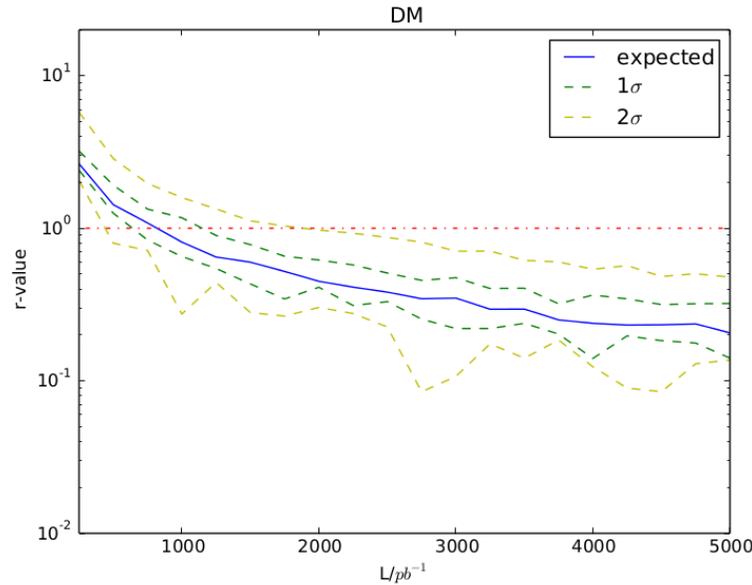


Figura 25:  $DM$ . Evolución respecto a la luminosidad del límite de confianza al 95 % para el valor  $r$ , intensidad de señal. Valores observados representados tanto para el método *Full CLs*. Para este último se encuentran dibujados además los límites a distancia  $\pm\sigma$  y  $\pm 2\sigma$ . Se observa exclusión en torno a  $L = 0,8fb^{-1}$ .

Para comprobar la posibilidad de exclusión utilizamos los métodos estadísticos descritos en la Sección 6.3. El resultado que se observa en la Figura 25 confirma nuestras sospechas: podemos garantizar la exclusión del punto considerado hasta con una luminosidad de  $0,8fb^{-1}$ . De hecho, el límite superior al nivel de confianza del 95 % para la intensidad de la señal  $r$  a la luminosidad de datos es inferior a 1 y, por tanto, encontramos exclusión de este punto.

Aprovechando este hecho, podemos intentar generalizar estos resultados a secciones eficaces menores con el objetivo de determinar la máxima sección eficaz posible que podemos excluir con nuestro análisis. Para ello redimensionamos la muestra de  $DM$  de forma que se corresponda a secciones eficaces cada vez menores y repetimos los test y el estudio con los  $CLs$ .

Los resultados que se observan en la Figura 26 nos indican que **es posible la exclusión** hasta secciones eficaces de  $\sigma \geq 0,7pb$  con la luminosidad de la que disponemos actualmente. De la prolongación incluida podemos observar las terriblemente elevadas luminosidades necesarias para la exclusión de los puntos con muy baja sección eficaz, siendo necesarios hasta  $1000fb^{-1}$  para la exclusión a secciones eficaces del orden de  $10^{-3}pb$ .

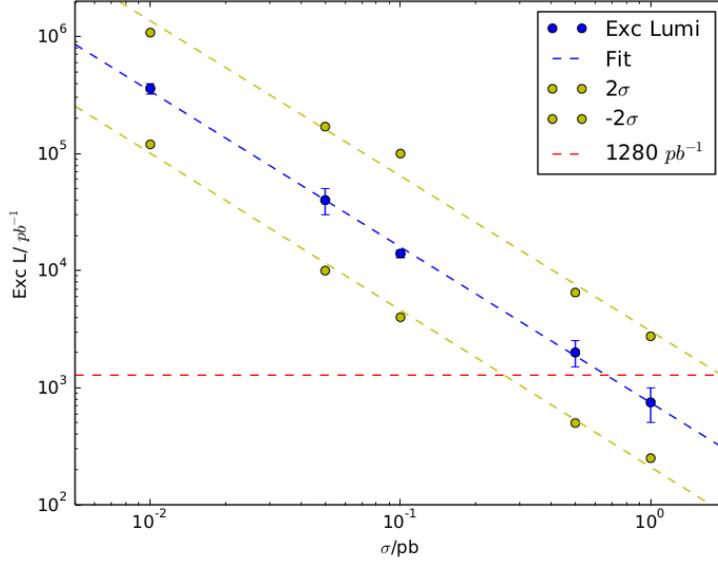


Figura 26:  $DM$ . Extensión del análisis de la luminosidad de exclusión a varias secciones eficaces posibles del proceso  $DM$  en la región definida por los cortes  $M_{T2,l} > 80\text{GeV}$ ,  $\cancel{E}_T > 100\text{GeV}$  y  $M_{eff} > 400\text{GeV}$ . Se observa exclusión para  $\sigma > 0,7\text{pb}$ .

## 9. Conclusiones

A lo largo de nuestro trabajo hemos planteado la posibilidad de las búsquedas de *nueva física*, concretadas en el caso de la búsqueda en aceleradores-colisionadores y hemos estudiado las posibilidades de la misma en términos de modelos simplificados SUSY y EFT para la búsqueda de DM.

En primer lugar, hemos proporcionado una descripción básica de los procesos *Beyond the Standard Model* y de los hechos que justifican la necesidad de su estudio. De esta forma, hemos motivado la introducción de los métodos de detección directa a través de aceleradores-colisionadores. Se ha introducido una descripción breve de los procesos de medición y reconstrucción llevados a cabo en este tipo de experimentos concretados en el caso del experimento *CMS*, cuyos datos son la base central de nuestro análisis.

De cara a entender el análisis utilizado hemos introducido los métodos multivariante, ejemplificados en el concepto de BDT. Adicionalmente, hemos introducido también los procedimientos utilizados para evaluar la presencia de incertidumbres en nuestras medidas y la extrapolación de los resultados a casos de más alta luminosidad.

Como punto central de este trabajo, los dos últimos capítulos del mismo se centran en el análisis de datos utilizando los conceptos y métodos introducidos en el resto de la memoria.

Utilizando datos reales obtenidos en el experimento CMS y correspondientes al *Run 2* del LHC (a una energía  $\sqrt{s} = 13\text{TeV}$ ) hemos obtenido varios resultados destacables.

En primer lugar hemos definido valores adecuados para los parámetros trabajo propios del BDT, así como para las muestras de entrenamiento y test necesarias para obtener los mejores resultados posibles desde el punto de vista de los límites de exclusión o detección.

Desde el punto de vista físico, nos hemos encontrado con limitaciones importantes en la luminosidad correspondiente a la muestra de datos utilizada que nos han impedido determinar la detección o exclusión de procesos de *nueva física* tipo SUSY en el canal  $T2tt$ . No obstante, hemos logrado determinar luminosidades esperadas de exclusión que podrán ser alcanzables, en principio, a medida que progrese el proceso de toma de datos.

Igualmente, hemos logrado imponer límites superiores en los parámetros del EFT considerado en la búsqueda de materia oscura, logrando establecer una exclusión de la misma para  $\sigma > 0,7pb$ .

Finalmente, destacar de nuestro estudio las apreciaciones cualitativas realizadas acerca de la relevancia de las variables tipo  $M_{T2}$  desde el punto de vista del poder discriminante que presentan. La consecuencia más clara de nuestro estudio es lo vitales que resultan, especialmente,  $M_{T2,u}$  para caracterizar y diferenciar los procesos de *nueva física* estudiados respecto a la conocida física del SM.

Por último, indicar la gran posibilidad de progreso en los estudios realizados en este trabajo. Al margen de los estudios de alta luminosidad, dentro de los propios modelos  $tt + DM$  y  $T2tt$  existen gran cantidad de regiones del espacio paramétrico no excluidas aún, que se podrían estudiar con las técnicas de este trabajo. No sólo eso, sino que la generalidad propia de nuestro método de trabajo puede extenderse a modelos de *nueva física* alternativos de forma bastante directa, lo que supone una extensión natural de todo el trabajo realizado hasta este punto.

## 10. Bibliografía y referencias

- [1] Y.NAGASHIMA, *Elementary Particle Physics Volume 2: Foundations of the Standard Model* Wiley (2013).
- [2] N. R. SHAH, *Lectures on Minimal Supersymmetric Standard Model (MSSM)* (2003).
- [3] C. CSAKI, *The Minimal Supersymmetric Standard Model (MSSM)* Mod.Phys.Lett.A11:599 (1996).
- [4] X.TATA, *Supersymmetry: Where it is and how to find it* hep-ph/9510287 (1995).
- [5] M. DREES, G. GERBIER, *Mini-Review of Dark Matter: 2012* [arXiv:1204.2373] (2012).
- [6] THE CMS COLLABORATION, *Search for top-squark pair production in the single-lepton final state in pp collisions at  $\sqrt{s} = 8$  TeV* Eur. Phys. J. C 73 (2013).
- [7] A. ASKEW, S. CHAUHAN, B. PENNING, W. SHEPHERD, M. TRIPATHI, *Searching for Dark Matter at Hadron Colliders* Int. Journal of Modern Physics A, Vol. 29 (2014).
- [8] K. SCHINDL, *The injector chain for the LHC* PS Division, CERN, (1999).
- [9] X. CID, R. CID , *Taking a closer look at LHC* ( [http://www.lhc-closer.es/taking\\_a\\_closer\\_look\\_at\\_lhc/](http://www.lhc-closer.es/taking_a_closer_look_at_lhc/) ). Accesible a 07/07/2016.
- [10] THE CMS COLLABORATION, *CMS Detector performance and software, Volume I* (2006).
- [11] THE CMS COLLABORATION, ( <http://cms.web.cern.ch> ). Accesible a 07/07/2016.
- [12] THE CMS TRIGGER AND DATA ACQUISITION GROUP, *The CMS High Level Trigger* Eur.Phys.J. C46 (2006).
- [12] THE CMS COLLABORATION, *Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and  $E_{miss}^T$*  CMS-PAS PFT-09-001 (2009).
- [13] A. HOECKER ET AL, *TMVA - Toolkit for Multivariate Data Analysis* PoS ACAT:040 (2007).
- [14] THE CMS COLLABORATION, *Search for new physics with the MT2 variable in all-jets final states produced in pp collisions at  $\sqrt{s} = 13$  TeV* [ arXiv:1603.04053] (2016).
- [15] THE ATLAS COLLABORATION, *Summary plots from the ATLAS Standard Model physics group.* ( <https://atlas.web.cern.ch/Atlas/GROUPSPHYSICS/CombinedSummaryPlots> ) Accesible a 07/07/2016.

- [16] B. P. ROE, H.J. YANG, J. ZHU, Y. LIU, I. STANCU, G. MCGREGOR, *Boosted Decision Trees as an Alternative to Artificial Neural Networks for Particle Identification* Nucl.Instrum.Meth. A543 577-584 (2005).
- [17] Y. FREUND, R.E. SCHAPIRE, *Experiments with a new boosting algorithm* Proc COLT, 209-217. ACM Press, New York (1996).
- [17] T. JUNK, *Confidence Level Computation for Combining Searches with Small Statistics* Nucl.Instrum.Meth.A434:435-443 (1999).
- [18] R. D. COUSINS, V. L. HIGHLAND, *Incorporating systematic uncertainties into an upper limit* Nucl.Instrum.Meth. A320:331-335 (1992).
- [19] G. COWAN, K. CRANMER, E. GROSS, O. VITELLS, *Asymptotic formulae for likelihood-based tests of new physics* Eur.Phys.J.C71:1554 (2011).
- [20] THE ATLAS COLLABORATION, *Search for direct top-squark pair production in final states with two leptons in pp collisions at  $\sqrt{s} = 8TeV$  with the ATLAS detector* JHEP06 (2014).
- [21] C.G.LESTER, D.J.SUMMERS, *Measuring masses of semi-invisibly decaying particles pair produced at hadron colliders* Phys.Lett.B463:99-103,1999.

En la carpeta adjunta, code, se puede observar un extracto del código empleado para los diferentes tratamientos realizados en este trabajo.