



**FÍSICA DEL QUARK TOP: PRODUCCIÓN DE
PAR TOP ANTI-TOP CON JETS
ADICIONALES**
(Top quark Physics: Top anti-top pair
production with additional jets)

Trabajo de Fin de Máster
para acceder al

**MÁSTER EN FÍSICA DE PARTÍCULAS
Y DEL COSMOS**

Autor: Conrado Muñoz Díaz

Director: Javier Andrés Brochero Cifuentes

Julio 2022

Agradecimientos

Con este trabajo concluye una etapa corta, pero intensa, de mucho aprendizaje y en la que he conocido a compañeros de muchos lugares diferentes.

Me gustaría comenzar dándole las gracias al director de este trabajo, Javier Brochero, por todas las horas dedicadas, su disposición a cualquier hora, cualquier día, la paciencia sobre todo al inicio cuando no tenía las nociones necesarias para el manejo de las herramientas de HEP, y en definitiva la ayuda que me ha prestado tanto a la hora de elaborar este Trabajo de Fin de Máster como a la hora de buscar opciones para mi continuación en estudios superiores. Agradecer también la ayuda de Sergio que nos ha permitido la migración de pesos de las redes entrenadas en Python a C++.

Agradecer a los coordinadores del máster, Patricio y Gervasio, por su implicación, su profesionalidad, y la ayuda que nos prestan a los alumnos en cualquier momento. Este es un máster relativamente reciente, con no más de una decena de alumnos por curso, pero que mediante los más de 20 profesores hacen que el aprendizaje sea de gran nivel. Gracias también a todos los profesores e investigadores invitados al máster.

También me gustaría dar las gracias al CSIC y al Instituto de Física de Cantabria por las becas de introducción a la investigación en el marco del Programa JAE Intro SOMdM, que nos permite a jóvenes estudiantes trabajar codo con codo con investigadores de primer nivel y en temas de total actualidad. Creo que la idea de la mayoría de jóvenes cuando empiezan la carrera de Física es acabar realizando investigación, y esta es sin duda una muy buena forma de adentrarse en este mundo.

Por último, agradecer también a mi familia por el apoyo y sustento no solo cuando decidí cambiar de ciudad para continuar con mis estudios en Física de Partículas, sino durante toda mi etapa como estudiante.

Resumen

En el siguiente trabajo se han estudiado diferentes métodos de identificación de jets adicionales que acompañan a decaimientos $t\bar{t}$ en modo semi-leptónico, producidos en colisiones protón-protón a energía en el centro de masas $\sqrt{s} = 13$ TeV en el LHC, cuyos datos fueron tomados por el detector CMS durante el año 2017. La intención de este estudio es caracterizar mejor uno de los principales fondos en eventos $t\bar{t}H$, los eventos $t\bar{t}b\bar{b}$. Para ello se han estudiado tres métodos: el primero de ellos más analítico, la reconstrucción cinemática, basada en la recomposición del evento $t\bar{t}$ a partir de sus productos de decaimiento; los otros dos, mediante diferentes técnicas de aprendizaje automático como BDT y DNN, consistieron en entrenar redes neuronales para identificar jets adicionales de los eventos $t\bar{t}$ +jets. Para la reconstrucción cinemática se han logrado eficiencias de hasta el 80 % identificando jets adicionales, mientras los métodos de aprendizaje automático llegaron al 60 %.

Palabras clave: quark top, $t\bar{t}$ +jets, categorías $t\bar{t}$, CMS, decaimiento semi-leptónico, reconstrucción cinemática, BDT, DNN.

Abstract

In this document, different identification methods for additional jets has been studied, that are produced with the semi-leptonic decay mode products of $t\bar{t}$. These events are produced in proton-proton collisions at energy in the centre of mass $\sqrt{s} = 13$ TeV in the LHC, whose data were collected by CMS detector during 2017. The interest of this study is a better characterization for $t\bar{t}b\bar{b}$ events, one of the main backgrounds in $t\bar{t}H$ events. To do this, three methods have been studied: the first of them more analytical, the kinematic reconstruction, based on the recomposition of the $t\bar{t}$ event from its decay products; the other two, using different machine learning techniques such as BDT and DNN, consisted of training neural networks to identify additional jets from the $t\bar{t}$ +jets events. For the kinematic reconstruction, efficiencies greater than 80 % have been achieved identifying additional jets, while for machine learning methods reached 60 %.

Keywords: top quark, $t\bar{t}$ +jets, $t\bar{t}$ categories, CMS, semi-leptonic decay mode, kinematic reconstruction, BDT, DNN.

Índice general

1. Introducción	1
2. El Modelo Estándar	3
3. La física del quark top	7
3.1. Producción de quark top	8
3.1.1. Producción de un único quark top	8
3.1.2. Producción de par top anti-top ($t\bar{t}$)	9
3.2. Decaimientos $t\bar{t}$	10
3.3. Decaimiento semi-leptónico de $t\bar{t}$	12
3.3.1. Principales fondos	13
3.4. Categorías $t\bar{t}$ +jets	14
4. El acelerador LHC y el detector CMS	17
4.1. El Solenoide Compacto de Muones (CMS)	19
4.1.1. Tracker de Silicio	20
4.1.2. Calorímetros	22
4.1.3. Sistema de muones	22
5. Análisis	25
5.1. Selección de objetos y eventos	25
5.1.1. DAQ y trigger	25
5.1.2. b-tagging	26

5.1.3. Selección de eventos	26
5.2. Gráficos de control	27
5.2.1. Leptón	27
5.2.2. Jets	28
5.2.3. Energía perdida transversa, MET	31
5.3. Emparejamiento GenParticles-RecoJets	31
5.4. Reconstrucción cinemática	33
5.5. Boosted Decision Tree (BDT)	36
5.6. Deep Neural Network (DNN)	43
6. Conclusiones	45
Bibliografía	49
A. Entrenamiento de las BDTs con datos de categoría $t\bar{t}b\bar{b}$	53
B. Redes neuronales profundas	57

Índice de figuras

2.1. Partículas Modelo Estándar	4
3.1. Diagramas de Feynman modos de producción único quark top	9
3.2. Diagramas de Feynman modos de producción par $t\bar{t}$	10
3.3. Decaimiento semi-leptónico $t\bar{t}$	12
4.1. Esquema experimentos CERN	18
4.2. Esquema detector CMS	20
4.3. Sección subdetectores CMS	23
5.1. Momento transversal leptones (MC y datos)	27
5.2. Pseudorapidez leptones (MC y datos)	28
5.3. Momento transversal jet más energético por leptones (MC y datos)	28
5.4. b-tagging del jet más energético por leptones (MC y datos)	29
5.5. Multiplicidad de b -jets (criterio laxo) por leptones (MC y datos)	29
5.6. Multiplicidad de b -jets (criterio medio) por leptones (MC y datos)	30
5.7. Multiplicidad de b -jets (criterio estricto) por leptones (MC y datos)	30
5.8. Energía perdida transversal por leptones (MC y datos)	31
5.9. Esquema GenParticles-GenJets-RecoJets	32
5.10. Variables entrada entrenamiento (jets separados) de BDT	37
5.11. Correlación variables entrada entrenamiento (jets separados) de BDT (señal y fondo)	37
5.12. Respuesta BDT (jets separados) para datos de verificación	38

5.13. Variables entrada entrenamientos (pares de jets) de BDT (fondo de W y de top)	39
5.14. Correlación variables entrada entrenamiento (pares de jets) de BDT (señal, fondo de W y de top)	40
5.15. Respuestas BDTs (pares de jets) para datos de verificación	41
5.16. Respuestas DNNs (pares de jets) para datos de verificación	44
A.1. Variables entrada entrenamientos $t\bar{t}b\bar{b}$ (pares de jets) de BDT (fondo de W y de top)	54
A.2. Correlación variables entrada entrenamiento $t\bar{t}b\bar{b}$ (pares de jets) de BDT (señal, fondo de W y de top)	55
A.3. Respuestas BDTs entrenadas con $t\bar{t}b\bar{b}$ (pares de jets) para datos de verificación	56

Índice de cuadros

3.1. Modos de decaimientos $t\bar{t}$	11
3.2. “ <i>Branching ratios</i> ” para el bosón W	11
3.3. “ <i>Branching ratios</i> ” para decaimientos $t\bar{t}$	11
3.4. Partículas detectables por categorías $t\bar{t}$ +jets (semi-leptónico)	15
5.1. Cortes a partir del trigger para e, μ y jets	26
5.2. Eficiencias emparejamiento GenParticles-GenJets y GenJets-RecoJets	33
5.3. Eficiencias del KinFitter identificando jets adicionales	34
5.4. Eficiencias del KinFitter identificando jets adicionales (condición laxa de b -jets)	35
5.5. Porcentaje de eventos con 2 o 1 de los mayores CSV en $t\bar{t}$	36
5.6. Pesos de las variables de entrada en la BDT de jets separados	37
5.7. Pesos de las variables de entrada en la BDT de pares de jets	41
5.8. Eficiencias de las BDTs identificando jets adicionales	42
5.9. Eficiencias de las BDTs (entrenadas con $t\bar{t}b\bar{b}$) identificando jets adicionales	43
5.10. Pesos de las variables de entrada en la DNN de pares de jets	44
5.11. Eficiencias de las DNNs identificando jets adicionales	44
6.1. Resumen eficiencias de los tres métodos identificando jets adicionales	46
A.1. Eficiencias de las BDTs entrenadas con $t\bar{t}b\bar{b}$ identificando jets adicionales	55

Capítulo 1

Introducción

El 5 de abril de 2015 se inició en el LHC (del inglés “*Large Hadron Collider*”, Gran Colisionador de Hadrones) la segunda puesta en funcionamiento del acelerador tras la primera parada entre los años 2013 y 2015, y duraría hasta finales de 2018. Este trabajo se ha realizado con simulaciones y datos recogidos durante el 2017 por el experimento CMS, cuando se producían colisiones protón-protón con energía en el centro de masas de 13 TeV, cuyos datos corresponden a una luminosidad de $L = 41,5 \text{ pb}^{-1}$. Los eventos estudiados en el documento contienen quarks top, partícula en la que se centran muchos estudios de física más allá del Modelo Estándar (SM por sus siglas en inglés, “*Standard Model*”).

Se han analizado diferentes métodos para la identificación de jets adicionales que aparecen en la producción $t\bar{t}$, particularizando para los decaimientos semi-leptónicos (uno de los bosones W a los que decae el quark top se desintegra leptónicamente, mientras que el otro bosón W lo hace hadrónicamente). En eventos $t\bar{t}$ se generan jets procedentes de los decaimientos de los quarks top y sus productos, así que es crucial identificar el origen de todos los jets del evento.

Las simulaciones Monte Carlo (MC) que se analizaron han sido a dos niveles: simulación de los procesos físicos que intervienen a nivel generación (Gen), teniendo en cuenta únicamente las partículas y sus condiciones físicas; y simulaciones a nivel reconstrucción (Reco), introduciendo la respuesta del detector CMS en la simulación.

En primer lugar, el trabajo ha consistido en identificar los jets generados en simulaciones a nivel Reco con sus correspondientes quarks generados a nivel Gen. Esto es lo que se llama emparejamiento (o “*matching*” en inglés) entre los GenParticles y los RecoJets.

Una vez determinados los orígenes reales de cada uno de los jets del evento, se prosiguió a evaluar los tres métodos estudiados para identificar jets adicionales.

El primero de ellos es la reconstrucción cinemática, en la que estudiando todas las permutaciones posibles de los diferentes jets de un evento, verifica cuál de ellas recompone mejor la estructura de decaimiento del sistema $t\bar{t}$. Para ello utiliza la información de las masas de las partículas involucradas y minimiza una función χ^2 probando todas las permutaciones posibles. Los jets del evento no utilizados por el algoritmo de reconstrucción cinemática son considerados como los jets adicionales del evento $t\bar{t}$. Comparando con los resultados del “*matching*”, obtendremos las eficiencias de este método de reconstrucción

cinemática, para el que podremos además imponer condiciones internas (por ejemplo, el sabor de los jets adicionales) en la selección de la mejor permutación en favor de mejorar las eficiencias.

El segundo de los métodos es la utilización de aprendizaje automático (“*machine learning*”) para obtener BDTs (“*Boosted Decision Trees*”) que permitan diferenciar en cada evento la señal (jets adicionales) del fondo (jets del par $t\bar{t}$). Una vez obtenidos los pesos de la red, evaluamos su desempeño con eventos estadísticamente independientes a los utilizados en el entrenamiento. La optimización del BDT se realiza en función de variables consideradas discriminantes entre pares de jets (par de jets adicionales, par de jets del bosón W , par de b -jets del $t\bar{t}$).

El tercer método es la utilización de redes neuronales profundas (“*deep neural networks*”, DNNs) con las librerías de KERAS para Python. De nuevo aquí se entrena una red con las mismas variables de entrada que en el caso de BDT para posteriormente evaluar las respuestas de la DNN con sus pesos.

Finalmente, se compararán los tres métodos mediante sus eficiencias respecto al “*matching*” (véanse capítulos 5 y 6).

El documento se encuentra dividido en varios capítulos introductorios (capítulos 2 al 4) y los correspondientes al estudio (capítulos 5 y 6). En el capítulo 2 se presenta una breve introducción acerca del SM, y en el capítulo 3 un desarrollo más extendido sobre la física del quark top y sus decaimientos semi-leptónicos. Luego, en el capítulo 4 se exponen algunos detalles del acelerador de partículas LHC y se explica la estructura del detector CMS.

Finalmente, y a modo de apéndice, se añaden algunos anexos (A y B) que si bien no son esenciales para entender la idea del texto, ayudan a la comprensión total del estudio.

Capítulo 2

El Modelo Estándar

La materia del Universo está compuesta por partículas elementales que junto con sus interacciones forma el conocido Modelo Estándar (SM) en Física de Partículas, formulado en la década de 1970 [1]. El SM contiene tres tipos de partículas: leptones, quarks (ambos fermiones de espín 1/2) y los mediadores (bosones, de espín entero, responsables de las interacciones entre los fermiones).

Los leptones, con carga eléctrica entera, son seis: cargados (electrón, muón y tau) y neutros (los correspondientes neutrinos: neutrino electrónico, muónico y del tau). La carga eléctrica de los leptones más masivos (e , μ , τ) es $-e$, mientras que los neutrinos, partículas de mucha menos masa, no tienen carga. Es por esto que los neutrinos solo interactúan débilmente, mientras que los leptones cargados pueden hacerlo además mediante interacción electromagnética.

Por su parte, también hay seis “sabores” de quarks: up (u), down (d), charm (c), strange (s), top (t) y bottom (b). Todos ellos tienen carga eléctrica racional: cargados positivamente, $\frac{2}{3}e$ (u , c y t) y con carga negativa, $-\frac{1}{3}e$ (d , s y b). Estos fermiones de carga semi-entera no solo pueden interactuar débilmente y electromagnéticamente, sino que a su vez lo hacen fuertemente. Los quarks son los constituyentes de los hadrones, agrupaciones de quarks. Existen dos clases de hadrones: por un lado, los mesones son la unión de un quark con un anti-quark, mientras que los bariones son la unión de tres quarks (o anti-quarks). Una peculiaridad del quark top (el más masivo de todos) es que no forma parte de hadrones, debido precisamente a su masa¹. Ejemplos de mesones son los conocidos piones (π^+ , π^- , π^0), mientras que los bariones más abundantes en la materia son los protones (uud) y neutrones (udd).

Los fermiones se suelen agrupar en tres generaciones o familias de manera natural, de forma que cada familia es una “réplica” de la anterior pero con una mayor masa. La presencia de tres generaciones de partículas en el SM sigue siendo a día de hoy una cuestión abierta en física teórica [3]. Esta clasificación la podemos observar, junto a las principales magnitudes características de cada partícula, en la figura 2.1.

¹El quark top, al ser más pesado que los bosones W, es el único quark que decae semi-débilmente (ver capítulo 3). Precisamente esta manera de decaer hace que su vida media sea muy corta, no dando tiempo a la hadronización antes del decaimiento [2].

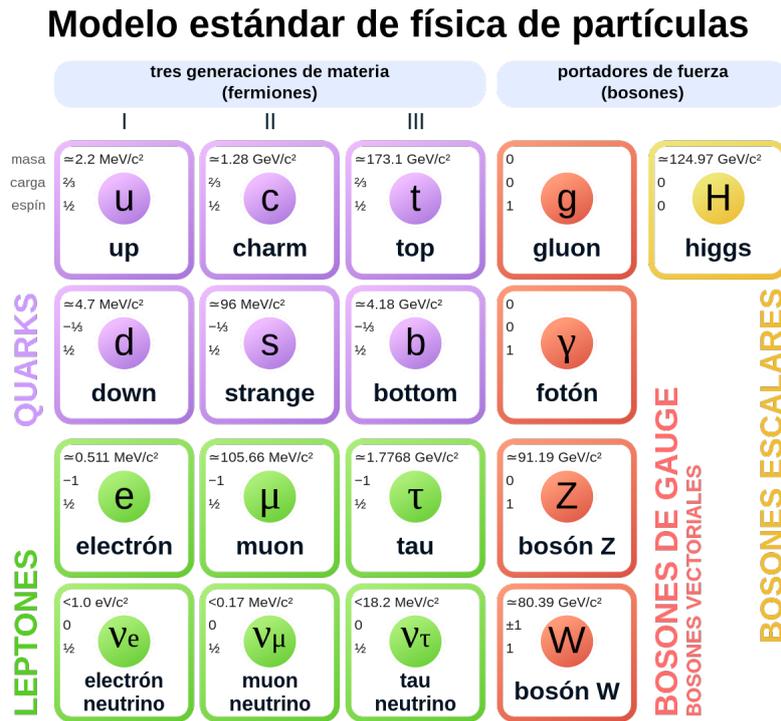


Figura 2.1: Partículas contenidas en el Modelo Estándar por generaciones y sus principales magnitudes (masa, carga eléctrica y espín). [Fuente: Wikipedia]

El último grupo de partículas que forman el SM son los bosones gauge, que hacen de mediadores en las interacciones del resto de partículas. Para la interacción débil, el SM cuenta con tres bosones gauge: el bosón Z (neutro) y los bosones W^\pm (con carga eléctrica ± 1). Estos tres bosones (junto con el Higgs) son los únicos con masa. En la interacción electromagnética, el fotón es el mediador incluido en el SM. Por último, la interacción fuerte queda mediada por los gluones. Estos últimos dos bosones no tienen ni masa, ni carga eléctrica. El bosón de Higgs es el único bosón con espín 0, mientras que el resto de mediadores expuestos tienen espín 1.

Para todas estas partículas que contiene el SM existen sus correspondientes anti-partículas, de igual masa pero de números cuánticos con signo opuesto. Por ejemplo, el positrón (e^+) es un electrón con carga positiva y número leptónico del electrón (L_e) -1 . A su vez, la asociación de anti-quarks formando bariones, por ejemplo, producen los anti-bariones, como el anti-protón (\overline{uud}), protón de carga negativa, y número bariónico -1 . También existen partículas, como el fotón, que son sus propias anti-partículas ($\gamma = \bar{\gamma}$).

Además de contener a las partículas de la figura 2.1, y las anti-partículas correspondientes, el SM es una teoría que pretende explicar también las interacciones de las mismas, mediante teoría cuántica de campos (QFT de sus siglas en inglés, “*Quantum Field Theory*”), una síntesis de la mecánica cuántica y la relatividad especial. La solución de aunar estas dos teorías modernas no es tan sencillo como encontrar una solución relativista de la ecuación de onda de Schrödinger (aparecen soluciones de energía negativa). Mientras que en las ecuaciones de ondas se cuantiza una partícula en un potencial externo, la QFT propone un cambio del punto de vista asumiendo las partículas como modos de un campo,

para cuantizar este. Esto permite resolver problemas físicos con muchos grados de libertad. Los resultados de la QFT son mejores cuanto menores son las interacciones, de forma que puedan ser tratadas como perturbaciones [4].

La electrodinámica cuántica (QED de sus siglas en inglés, “*Quantum Electrodynamics*”) es la teoría que explica las interacciones electromagnéticas. Esta interacción, debido a que es mediada por fotones, partículas sin masa, tiene alcance infinito². A escalas nucleares ($r \sim 10^{-13}$ cm ó energías $E \sim 200$ MeV) entran en juego las dos interacciones restantes del SM: por un lado, la interacción nuclear fuerte, responsable de mantener los neutrones y protones en el núcleo atómico, y por otro, la interacción nuclear débil, mediante la que se producen procesos de desintegración beta por ejemplo [4]. Esta interacción débil, dado que sus mediadores (bosones Z y W) son masivos, tienen un alcance limitado, mientras que la interacción fuerte, mediada por gluones (sin masa) en principio debería ser de alcance infinito, aunque el fenómeno de confinamiento [5] hace que esta interacción tenga una distancia típica del núcleo atómico.

La fuerza débil la entendemos como una aproximación a bajas energías de una teoría más completa, la teoría electrodébil (EW por sus siglas en inglés, “*Electro Weak*”), que junta la electrodinámica cuántica con la fuerza débil. La teoría EW junto con la de las interacciones fuertes, la cromodinámica cuántica (QCD, del inglés “*Quantum Chromodynamics*”) forman el SM [4].

²La fuerza gravitatoria también tiene alcance infinito, aunque de ella no hablaremos ya que no está incluida en el SM, dada la imposibilidad, por el momento, de cuantizar esta fuerza.

Capítulo 3

La física del quark top

En 1975 una publicación [6] con el título “*Evidence for anomalous lepton production in $e^+ - e^-$ annihilation*” (Evidencia de producción de leptón anómalo en aniquilación $e^+ - e^-$), se proponía la existencia de un nuevo leptón más masivo que los otros dos (e , μ) que se denominó leptón tau, τ . La confirmación de este nuevo fermión del SM, junto con el de su correspondiente neutrino ν_τ , hizo sospechar acerca de la formación de una tercera generación en el SM, que siguiendo de la misma estructura que las dos ya presentes ($u - d$; $e - \nu_e$ y $c - s$; $\mu - \nu_\mu$) predecía la existencia de un doblete de isospín formado por dos nuevos quarks [7].

La llegada del primer nuevo quark se produjo un año después por dos grupos del Síncrotrón Alemán de Electrones (DESY). Este fue el quark bottom, del que midiendo el área bajo la curva de resonancia en el experimento infirieron una carga eléctrica $Q = -1/3$ [7].

Por tanto, el quark que completara esta tercera familia del SM junto con los leptones τ , ν_τ y con el quark b debía ser de carga eléctrica $Q = 2/3$ y tercera componente de isospín $T_3 = 1/2$. Todo esto predecía lo que más tarde se acabó llamando quark top (t), compañero de isospín del quark b . Además de completar la tercera generación del SM, la existencia de este quark proporciona la explicación más sencilla para la violación de simetría CP (carga y paridad) en la interacción débil [7].

En colisiones protón anti-protón ($p\bar{p}$) se espera [7] que aparezcan pares top anti-top ($t\bar{t}$) mediante fusión gluón-gluón y aniquilación quark anti-quark. Y así fue en 1995 en el colisionador Tevatrón de Fermilab, donde los dos detectores CDF [8] y DØ [9] descubrieron el hasta ahora último quark cuya masa es de unos $174 \pm 10_{-12}^{+13}$ GeV/ c^2 según el grupo CDF [8].

Es precisamente el valor tan elevado de la masa del quark top (mayor que los ≈ 80 GeV/ c^2 de masa del bosón W) lo que hace que la física de esta partícula sea realmente importante en el desarrollo de nuevas teorías. En particular, su alta masa hace que el quark top decaiga semi-débilmente en un bosón W y un quark tipo down (d, s, b). Las probabilidades de que decaiga en cada uno de los distintos quark de tipo down las da la matriz de Cabbibo-Kobayashi-Maskawa (CKM), que contiene la probabilidad de transición de un quark q en otro q' simplemente tomando módulo cuadrado del elemento de la matriz

en cuestión, $|V_{qq'}|$:

$$V_{\text{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \quad (3.1)$$

Los valores de los módulos de cada uno de los elementos de la matriz CKM son [10]:

$$|V_{\text{CKM}}| = \begin{pmatrix} 0,97435 \pm 0,00016 & 0,22500 \pm 0,00067 & 0,00369 \pm 0,00011 \\ 0,22486 \pm 0,00067 & 0,97349 \pm 0,00016 & 0,04182^{+0,0005}_{-0,00074} \\ 0,00857^{+0,00020}_{-0,00018} & 0,04110^{+0,00083}_{-0,00072} & 0,999118^{+0,00031}_{-0,000036} \end{pmatrix}, \quad (3.2)$$

así que las probabilidades del quark top en decaer en los diferentes quarks tipo down son:

$$t \rightarrow W + q, \text{ siendo } q = d (0,007\%); s (0,17\%); b (99,82\%), \quad (3.3)$$

es decir, casi con total probabilidad el quark top decae en un bosón W y un quark b .

Con una masa como la del quark top, y dado su principal canal de desintegración, el SM predice a NLO (“*Next to Leading Order*”) una anchura de desintegración $\Gamma_t \approx 1,35\text{GeV}/c^2$, lo que supone un tiempo de vida media del quark top de $\approx 0,5 \times 10^{-24}$ s, por tanto se espera que decaiga antes de que forme hadrones con sabor top o estados ligados de quarks $t\bar{t}$. Este es el principal distintivo del quark top, su no presencia en hadrones, a diferencia del resto de quarks del SM [2].

3.1. Producción de quark top

Los principales mecanismos de producción de quark top según el SM en los colisionadores son: mediante interacción electro-débil, la producción de un único quark top (por evento); y mediante interacciones fuerte, la generación de pares top anti-top, $t\bar{t}$.

3.1.1. Producción de un único quark top

La interacción que media en la producción de un único quark top por evento es la electro-débil, mediante vértices Wtb . Aunque en este método de producción es complicado de separar la señal del fondo (aplican análisis multivariable en el artículo [11]), los resultados permiten obtener mejores valores para el elemento $|V_{tb}|$ de la matriz CKM, sin asumir únicamente la existencia de tres familias de quarks en el SM. La condición de unitariedad de la matriz CKM implica que ese elemento de matriz deba ser $|V_{tb} \simeq 1|$, de manera que una medida más precisa de este elemento de matriz podría indicar la presencia de una cuarta familia de quarks [11].

La producción de un único quark top se da mediante tres modos principalmente [12], que se pueden ver en la figura 3.1:

- Canal s: la colisión quark anti-quark genera un bosón W virtual, que decae a un par quark anti-quark $t\bar{b}$, como se ve en la figura 3.1 (a). Es el canal que menos contribuye a la formación de un único quark top, con un 4% de probabilidad¹.

¹Realmente estos porcentajes dependen de la energía en el centro de masas de la colisión pp . En este caso estamos tomando un valor $\sqrt{s} = 7$ TeV

- Producción de par tW : el estado final es un par fermión top con bosón W . Su contribución es del 20%. Su diagrama de Feynman lo podemos ver en la figura 3.1 (b).
- Canal t : es el modo que más contribuye, con un 75%. También recibe el nombre de fusión W gluón, dado que el bosón W colisiona con un quark b del mar de quarks del protón previamente generado por un gluón, produciendo finalmente un quark t . Este modo se observa en la figura 3.1 (c) y (d).

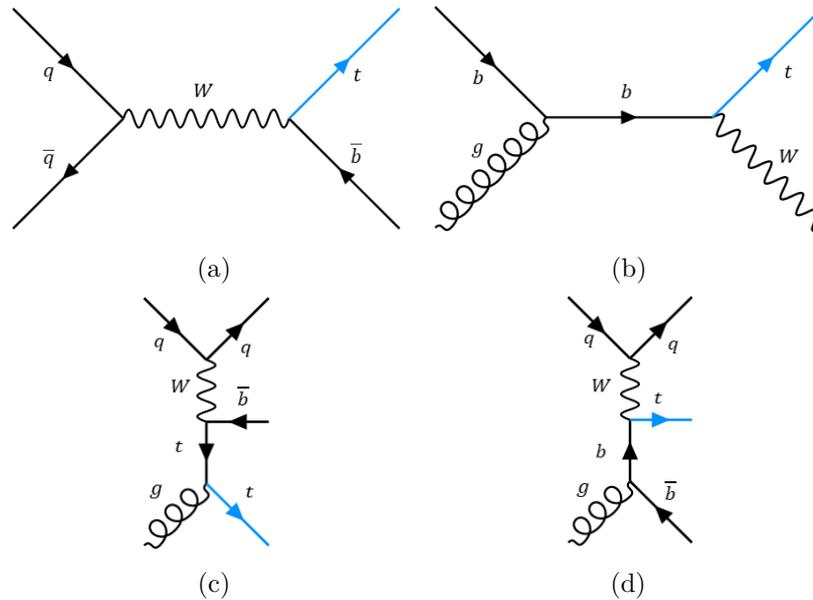


Figura 3.1: Diagramas de Feynman para los tres modos de producción de un único quark top: (a) canal s ; (b) modo producción tW ; (c) y (d) canal t . Se marca en azul la producción de un único quark top en cada modo.

Además, otros quarks tipo down pueden estar involucrados (quarks d, s) aunque su importancia viene dada por los elementos de la matriz CKM, cuyos valores son mucho menores que en el caso del quark b (véase ecuación 3.2) [2].

3.1.2. Producción de par top anti-top ($t\bar{t}$)

En los colisionadores de hadrones, los quarks top se producen principalmente en pares $t\bar{t}$ mediante interacciones fuertes (gobernada por la QCD). Los procesos más comunes de producción de pares $t\bar{t}$ son los que se muestran en la figura 3.2. Aproximadamente el 80% de la producción de pares en el LHC proviene de la fusión de gluones (figuras 3.2 (a) y (b)), mientras que el 85% de los pares del Tevatrón son debidos a la aniquilación de quarks (figura 3.2 (c)), diferencia debida principalmente a la asimetría de carga inicial² [2].

²La colisión que se produce en el Tevatrón es $p\bar{p}$, mientras que en el LHC se producen colisiones protón-protón, pp

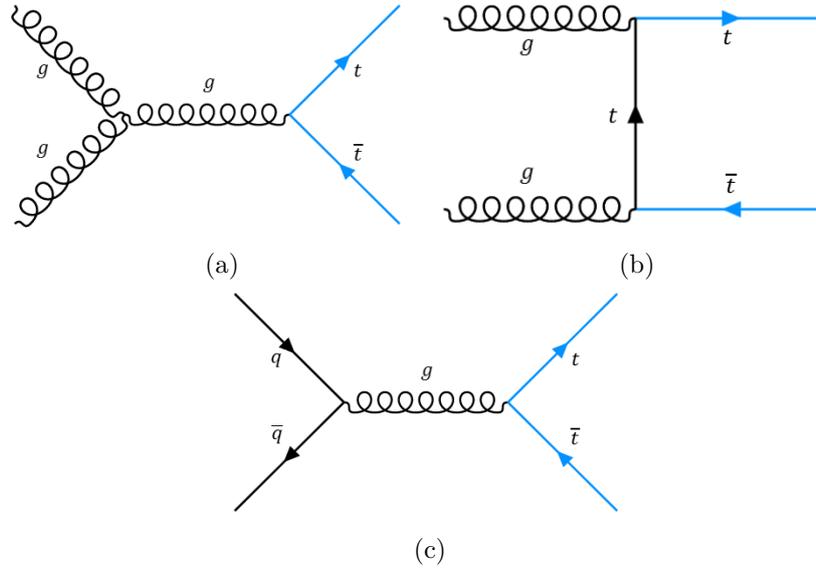


Figura 3.2: Diagramas de Feynman para los dos principales modos de producción de pares $t\bar{t}$. Arriba, (a) y (b) mediante fusión de gluones. Abajo, (c), mediante aniquilación de quarks. Se marca en azul la producción del par $t\bar{t}$ en cada modo.

En su mayoría, la producción de quarks top en los colisionadores se da por pares $t\bar{t}$, y en ocasiones acompañados de jets³ adicionales. En las siguientes secciones se resumen los decaimientos, fondos y categorías de estos eventos $t\bar{t}$.

3.2. Decaimientos $t\bar{t}$

La gran masa del quark top hace que tenga un tiempo de vida media muy corto, no dando tiempo a la hadronización. Esto hace que el quark más pesado de todos decaiga, en un 99.8% de las veces en un bosón W y un quark b , por lo que un par $t\bar{t}$ decaerá en dos bosones W y dos quarks b [12].

Experimentalmente, los quarks b se detectan como jets muy energéticos, dada la hadronización de este quark, mientras que los bosones W pueden decaer leptónicamente o hadrónicamente. Al haber dos bosones: el decaimiento leptónico se produce cuando ambos lo hacen a leptón (cargado) más neutrino; el decaimiento hadrónico, cuando ambos se desintegran en un par $q\bar{q}$, dando lugar a un jet por cada quark (debido a la hadronización de estos); y el decaimiento semi-leptónico, cuando uno de ellos decae leptónicamente pero el otro lo hace hadrónicamente. La misma nomenclatura se utiliza para hablar por tanto de decaimientos $t\bar{t}$, de forma que en eventos de desintegración $t\bar{t}$, al menos siempre habrá 2 jets (los procedentes de los dos quarks b del decaimiento de quarks top), y como máximo habrá 6 jets (dos b -jets, y un par del decaimiento de cada W , en el modo hadrónico en el que los dos W caen a pares $q\bar{q}$). En el cuadro 3.1 se resumen estos tres modos en que pueden decaer los pares $t\bar{t}$.

³Se suele denominar jets al conjunto de partículas resultantes de la hadronización de un quark, dada la imposibilidad de este a permanecer aislado (energéticamente es más rentable generar hadrones y demás partículas en un “choro” que tiene la dirección del propio quark, que mantener aislado al mismo).

Decaimientos	N_b	N_l	$N_{q\bar{q}}$	N_{Jets}
Leptónico	2	2	0	2
Hadrónico	2	0	2	6
Semi-leptónico	2	1	1	4

Cuadro 3.1: Resumen de las partículas resultantes según el modo de decaimiento del par $t\bar{t}$. La segunda columna indica el número de quarks b por evento, la tercera el número de pares leptón-neutrino, la cuarta el número de pares $q\bar{q}$, y la última el número de jets asociados a la desintegración del par $t\bar{t}$ según el modo de decaimiento (suma de las columnas segunda más el doble de la cuarta, al haber dos jets por cada par $q\bar{q}$).

En resumen, dado que los eventos $t\bar{t}$ decaen siempre a $W + b$, el decaimiento de los dos bosones W es lo que condiciona el modo de desintegración del evento $t\bar{t}$. La probabilidad de desintegración en diferentes estados finales es lo que se conoce como “*branching ratio*” (BR), y en el caso del bosón W se presentan en el cuadro 3.2 [13].

Decaimientos	BR (%)
$e + \nu_e$	$10,71 \pm 0,16$
$\mu + \nu_\mu$	$10,63 \pm 0,15$
$\tau + \nu_\tau$	$11,38 \pm 0,21$
Hadrones	$67,41 \pm 0,27$

Cuadro 3.2: “*Branching ratios*” para el bosón W .

Combinando las probabilidades para el caso de decaimientos de pares $t\bar{t}$ en los que hay dos W , las abundancias relativas de cada modo de desintegración se presentan en el cuadro 3.3.

Modo	Partículas	BR (%)
Leptónico	$ll\nu\nu$	$10,71 \pm 0,20$
	$ee\nu_e\nu_e$	$1,15 \pm 0,03$
	$\mu\mu\nu_\mu\nu_\mu$	$1,13 \pm 0,03$
	$\tau\tau\nu_\tau\nu_\tau$	$1,30 \pm 0,05$
	$e\mu\nu_e\nu_\mu$	$2,28 \pm 0,05$
	$e\tau\nu_e\nu_\tau$	$2,44 \pm 0,06$
	$\mu\tau\nu_\mu\nu_\tau$	$2,42 \pm 0,06$
Semi-leptónico	$l\nu q\bar{q}$	$44,10 \pm 0,40$
	$e\nu_e q\bar{q}$	$14,44 \pm 0,22$
	$\mu\nu_\mu q\bar{q}$	$14,33 \pm 0,21$
	$\tau\nu_\tau q\bar{q}$	$15,34 \pm 0,29$
Hadrónico	$q\bar{q}q'\bar{q}'$	$45,40 \pm 0,40$

Cuadro 3.3: “*Branching ratios*” para decaimientos $t\bar{t}$ calculados a partir de los mismos para el bosón W del cuadro 3.2.

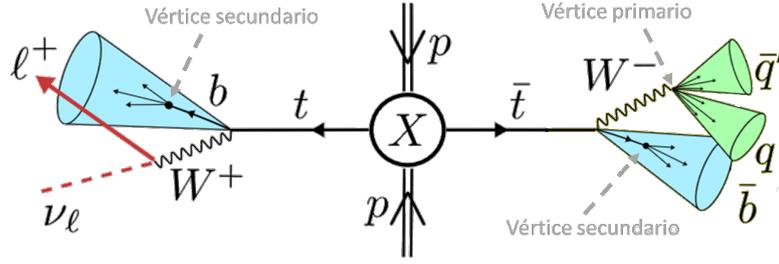


Figura 3.3: Esquema del decaimiento semi-leptónico de $t\bar{t}$ en el que uno de los bosones W se desintegra en un par leptón cargado más neutrino, mientras que el otro lo hace a un par $q\bar{q}$. Además se puede observar la diferencia entre la hadronización en vértice primario para jets ligeros (del W) y la hadronización en vértice secundario para b -jets. [Fuente: UAM]

Entre leptones y quarks, lo más eficiente a la hora de detectar partículas son los leptones, por lo que el modo leptónico sería el más indicado para nuestro estudio, además de tener menos procesos como fondo. Sin embargo, la baja estadística (aproximadamente un 10% del total de decaimientos son leptónicos) hace que el canal de desintegración en el que nos centremos sea el semi-leptónico, con prácticamente cuatro veces más posibilidades de obtener este tipo de desintegración que en el primero, y con una mayor eficiencia en detección que en el caso hadrónico. Por tanto, en la siguiente sección se explicará el modo de decaimiento semi-leptónico, con sus principales fondos que podrán interferir en las mediciones en las que estamos interesados.

3.3. Decaimiento semi-leptónico de $t\bar{t}$

El modo de decaimiento semi-leptónico consiste en la desintegración hadrónica (en un par $q\bar{q}$) de uno de los bosones W mientras el otro decae leptónicamente a un electrón (o μ , o τ) y su correspondiente neutrino, de tal forma que, en total en el proceso de desintegración $t\bar{t}$ se podrán detectar cuatro jets más el leptón cargado por cada evento⁴, tal y como se mostraba en el cuadro 3.1.

Los jets pueden ser ligeros (quarks u, d principalmente, aunque también hay quien incluye los quarks s, c) o pesados (quark b). Generalmente los jets resultantes del decaimiento de bosones W son ligeros, y una de las diferencias más claras con los jets pesados es la velocidad con la que hadronizan: mientras los jets ligeros hadronizan nada más formar el quark (lo que se llama el vértice primario), los jets pesados comienzan a hadronizar transcurrido un corto (pero medible) espacio, en el denominado vértice secundario. Esta diferencia se puede ver en el esquema del decaimiento semi-leptónico de $t\bar{t}$ de la figura 3.3.

Algunas de las características cinemáticas más relevantes de este tipo de decaimientos son: el alto momento transversal (p_T) del leptón (dada la baja masa del neutrino y la alta masa del bosón W , por conservación de la energía, el leptón es altamente energético); la energía faltante transversal (E_T) se asocian a la energía de neutrinos al no ser detectados; la medida del vértice secundario respecto del primario sirve para identificar b -jets mediante lo que se denomina “ b -tagging” [14].

⁴El neutrino interactúa mediante fuerza débil y se escapa del detector sin dejar señal. Una forma de estimar su energía es medir la faltante, y asignársela al neutrino.

De los tres modos de decaimiento de los eventos $t\bar{t}$, el que mejor balance tiene entre el número de señal y fondo es el decaimiento semi-leptónico, ya que el leptónico tiene poca estadística y el hadrónico, aunque es el que más estadística posee, tiene mucho fondo. Por ello, en la siguiente subsección mostraremos cuales son esos eventos que tienen un estado final similar al del decaimiento semi-leptónico.

3.3.1. Principales fondos

Los fondos son eventos que producen estados finales similares al que queremos estudiar, pero que su origen es muy distinto. Entre las causas de la semejanza entre estados finales podemos destacar no-detecciones de algunas partículas (partículas perdidas) en eventos para los que se acaben detectando el mismo estado final que en procesos $t\bar{t}$, o directamente asociar partículas a eventos de manera errónea, como por ejemplo los leptones, que pueden haberse producido en decaimientos de mesones de los jets, de rayos cósmicos, etc.

Los principales procesos del SM que contribuyen al fondo (o ruido) [15] de la señal que queremos estudiar son:

- **Decaimientos leptónicos y hadrónicos de $t\bar{t}$.** Los casos en los que el decaimiento de $t\bar{t}$ son a dos leptones pueden llevar a falseamientos de la señal al perderse uno de los leptones. En el caso del decaimiento hadrónico, es posible detectar leptones falseados o provenientes de otras fuentes.
- **Single top.** Un quark top es producido junto a un bosón W . El quark t decae a un b y otro W , y en el caso en que los dos bosones del proceso decaigan de manera diferente (uno leptónicamente y otro hadrónicamente), la señal es similar a la que buscamos.
- **W +jets.** Si mediante asociación de gluones o quarks se produce un bosón W que decae leptónicamente, el gluón puede generar un par de quarks pesados ($b\bar{b}$), dado lugar a un estado final semejante al de $t\bar{t}$.
- **Z +jets.** Similar a la situación anterior. También podría ocurrir una identificación errónea de los jets como leptones o viceversa.
- **Dibosones (VV).** Al producirse pares de bosones de interacción débil (WW , ZZ , y WZ) uno de ellos puede decaer leptónicamente y el otro hacerlo de manera hadrónica. La baja sección eficaz de estos procesos hace que su contribución sea pequeña.
- **QCD.** Dada la alta concentración de quarks en las colisiones, estos mismo o los gluones podrían producir pares $b\bar{b}$ que a su vez decaerían en quarks más livianos junto con bosones W , generando la misma estructura final que el decaimiento semi-leptónico de $t\bar{t}$ si uno de los bosones W decae hadrónicamente y el otro lo hace a leptón y neutrino.
- **Bosón vectorial o escalar ($t\bar{t}V$).** En decaimientos $t\bar{t}$ pueden producirse además bosones vectoriales (W, Z) o escalares (H) que podrían decaer hadrónicamente o leptónicamente. En particular, el decaimiento hadrónico es especialmente importante dado que en estos tipos de eventos, podrían detectarse seis jets y un leptón, igual que en los eventos $t\bar{t} + jets$ que vamos a estudiar.

El fondo con más presencia es el de decaimientos leptónicos y hadrónicos por posibles fallos de detección, seguido del proceso “*single top*”. El resto de procesos son mucho menos frecuentes, siendo el de menor relevancia el de un único bosón vectorial ($t\bar{t}V$) [15].

Por ello, en los gráficos no mostraremos la contribución de los fondos por separado, simplemente por un lado estará la de los otros modos de decaimiento de $t\bar{t}$ (leptónico y hadrónico) como $t\bar{t}Bkg$, y por otro la suma de todo el resto de fondos como $Nont\bar{t}Bkg$.

3.4. Categorías $t\bar{t}+jets$

Los eventos $t\bar{t}$ llevan usualmente asociados jets adicionales [16], que pueden ser livianos (“*light-flavor*”, LF) o pesados (“*heavy-flavor*”, HF), y las características de los eventos con diferentes jets adicionales suelen variar mucho. Así que en esta sección presentamos las principales categorías de $t\bar{t}+jets$ atendiendo al sabor de los jets adicionales.

La forma de asignar a diferentes jets el origen de un quark determinado se realiza mediante algoritmos de “emparejamiento fantasma”(o “*ghost-matching*” en inglés), que estudian los productos de la hadronización de los quarks con unas coordenadas similares, y asignan el sabor del quark inicial atendiendo al contenido de quarks de los hadrones en el jets. Por ejemplo, hadrones con quarks b probablemente vengan de un quark b , mientras que hadrones con quarks c lo harán de la hadronización de este último. Las principales características de este tipo de algoritmos son: clasifica los jets por su contenido en quarks de distintos sabores (b , c o ligeros); para cada hadrón encuentra un único quark y su correspondiente partícula madre (partícula que produjo el quark que cuando hadronizó dio lugar al hadrón detectado); los hadrones asociados a decaimientos leptónicos de jets también son identificados, y se asocian al leptón del decaimiento [17, 18].

Principalmente se distinguen cuatro categorías de decaimientos $t\bar{t}+jets$ según el sabor de los jets adicionales, que de menos a más frecuente son:

- $t\bar{t}b\bar{b}$: eventos con al menos 4 b -jets (2 de los adicionales y otros dos del decaimiento de cada quark top). Supone alrededor del 2% de los eventos $t\bar{t}+jets$ cuando se seleccionan únicamente eventos con 6 jets, de los cuales al menos hay 2 b -jets.
- $t\bar{t}b$: al menos tienen 3 b -jets (1 adicional y otros dos del decaimiento de cada quark top). Pueden deberse a no-detecciones de uno de los b -jets adicionales, o por no superar la aceptación del detector. Con las mismas condiciones de selección que antes, estos eventos suponen cerca del 6% del total.
- $t\bar{t}c\bar{c}$: eventos con al menos 2 c -jets (los adicionales) y 2 b -jets (del decaimiento de los quarks top). Suelen falsear la señal de eventos con b -jets adicionales (algunos de estos se detectan como c -jets) y por tanto, con las condiciones de selección de eventos antes expuestas, suponen más de un 14% de los eventos $t\bar{t}$.
- $t\bar{t}LF$: aunque pueden ser eventos con 4 jets livianos (2 de los adicionales y otros dos del decaimiento de uno de los bosones W), también se incluyen en esta categoría eventos $t\bar{t}$ sin jets adicionales, o incluso con un jet liviano junto a un c -jet. Son los eventos más comunes, con un 78% de las ocasiones (teniendo en cuenta eventos con 6 jets, de los cuales 2 son b -jets).

A modo de resumen, y particularizando para el modo de decaimiento semi-leptónico de $t\bar{t}$ +jets, en el siguiente cuadro se muestran los diferentes tipos de partículas y jets que podrían detectarse para cada categoría de jets adicionales, asumiendo que el par $q\bar{q}$ al que decae uno de los bosones W son jets livianos (es cierto para la gran mayoría de eventos).

Categoría	N_l	N_b	N_c	N_{LF}	N_{Jets}
$t\bar{t}b\bar{b}$	1	≥ 4	0	≥ 2	≥ 6
$t\bar{t}b$	1	≥ 3	0	≥ 2	≥ 5
$t\bar{t}c\bar{c}$	1	≥ 2	≥ 2	≥ 2	≥ 6
$t\bar{t}LF$	1	2	0 (1*)	≥ 2	≥ 4

Cuadro 3.4: Resumen de las partículas y jets que podrían detectarse según la categoría $t\bar{t}$ +jets. La segunda columna, número de leptones más neutrinos siempre es 1 ya que tratamos el modo semi-leptónico. La tercera columna, el número de b -jets en cualquier caso será como mínimo 2 (los dos del decaimiento de cada quark top). La cuarta, el número de c -jets es nulo salvo en la categoría $t\bar{t}c\bar{c}$ o cuando haya un único c -jet en $t\bar{t}LF^*$. La quinta columna son los jets livianos, y dado que asumimos que el bosón W que decae hadrónicamente lo hace a esta categoría de jets, como mínimo siempre habrá 2. La última columna es el total de jets del evento, que como mínimo será 4 (los 4 jets del decaimiento $t\bar{t}$: dos b -jets de los top y dos jets livianos del bosón W que decae hadrónicamente). El sentido de usar el símbolo \geq es que pueden detectarse eventos con más de dos jets adicionales.

Capítulo 4

El acelerador LHC y el detector CMS

El Gran Colisionador de Hadrones (*“Large Hadron Collider”*, LHC) [19] es el mayor acelerador de partículas del mundo, localizado en la frontera Franco-Suiza, que forma parte del Centro Europeo para la Investigación Nuclear (CERN), en el que participan 21 países de todo el mundo. Consta de más de 9300 dipolos magnéticos superconductores (que operan a temperaturas cercanas al cero absoluto) a lo largo de todo el anillo de aproximadamente 27 km de circunferencia, y cuya función es acelerar protones¹ hasta velocidades cercanas a la de la luz para hacerlas colisionar entre sí o contra determinados blancos en puntos bien definidos, donde se encuentran los detectores del LHC.

Los cuatro detectores principales del LHC son:

- **CMS:** *“Compact Muon Solenoid”* [20], uno de los principales detectores por los que se llevó a cabo el proyecto de LHC. En la siguiente sección se amplía la información, ya que fue en este experimento donde se recogieron los datos utilizados en este trabajo.
- **ALICE:** *“A Large Ion Collider Experiment”* [21] cuyo objetivo es el estudio de la física de la materia que interactúa fuertemente en densidades energéticas muy elevadas (plasma quark-gluón, que posiblemente fuera el primer estadio del Universo tras el Big Bang).
- **ATLAS:** *“A Toroidal LHC ApparatuS”* [22] es el otro de los dos detectores de objetivos generales del LHC, que investiga una amplia gama de física, desde la búsqueda del bosón de Higgs [23, 24] (lograda en el año 2012, junto al detector CMS, también del CERN) hasta dimensiones y partículas adicionales que podrían formar la materia oscura. Aunque tienen los mismos objetivos que el experimento CMS, utiliza diferentes técnicas y diseño del sistema magnético.
- **LHCb:** *Large Hadron Collider beauty experiment”* [25], especializado en la investigación mediante el estudio del quark *b* de las diferencias entre materia y antimateria.

El LHC fue instalado en el antiguo LEP (*“Large Electron-Positron Collider”*) para acelerar y hacer colisionar protones a energías de centro de masa de hasta 13 TeV, aunque

¹Estos protones se obtienen de átomos de hidrógeno a los que se le extraen el electrón.

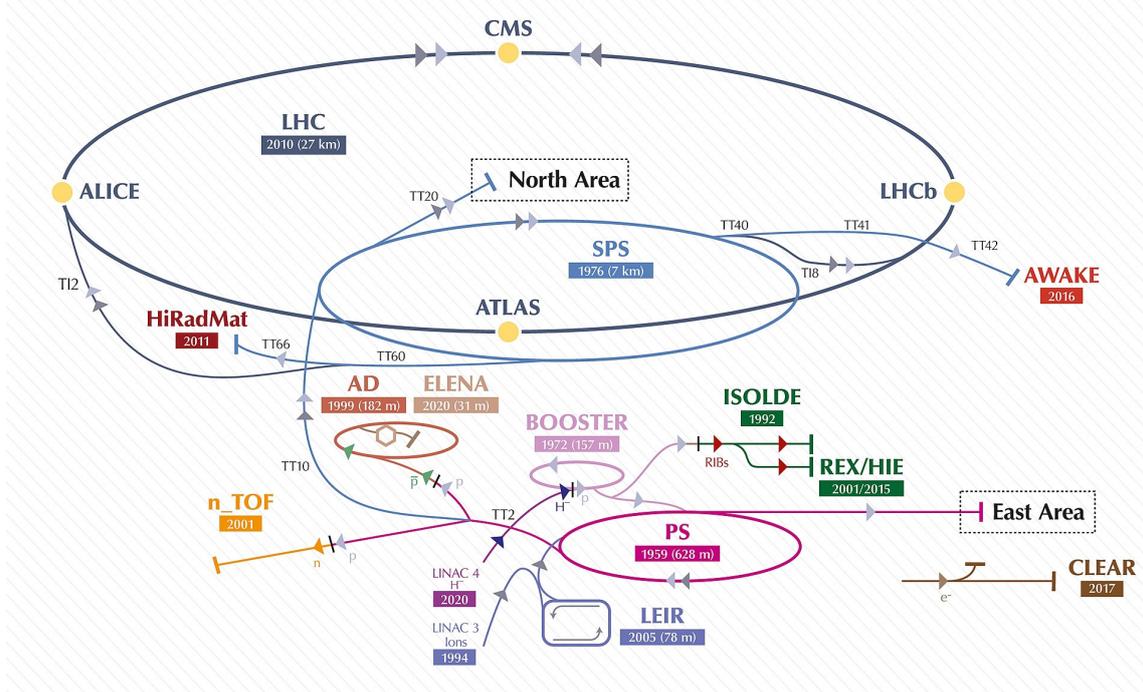


Figura 4.1: Esquema de los experimentos y aceleradores del CERN. [Fuente: CERN]

esta energía se ha ido aumentando de forma progresiva, desde los 900 GeV en el año 2009, hasta los 13 TeV de finales del 2018 antes de una parada para realizar nuevas actualizaciones que se ha prolongado hasta el 2022.

El número de eventos producidos por segundo en colisiones del LHC viene dado por:

$$N_{\text{event}} = L\sigma_{\text{event}}, \quad (4.1)$$

siendo σ_{event} la sección eficaz y L la luminosidad, uno de los parámetros más importantes de los colisionadores, que depende de parámetros del haz de la siguiente forma:

$$L = \frac{N_b^2 n_b f_{\text{rev}} \gamma_r}{4\pi \epsilon_n \beta^*} F, \quad (4.2)$$

donde N_b es el número de partículas por paquete², n_b es el número de paquetes por haz, f_{rev} es la frecuencia de revolución, γ_r es el factor gamma relativista, ϵ_n es la normalización de la emitancia de haz transverso, β^* es la función beta en el punto de colisión y F es el factor de reducción debido al ángulo de cruce.

El procedimiento de aceleración de protones en LHC es una sucesión de pasos en los que se va proporcionando energía al haz. Primero, del acelerador lineal (LINAC 2) pasa al Sincrotrón de Protones Booster con una energía de 0.12 GeV. Allí, se aceleran hasta alcanzar los 1.4 GeV, y entonces, en el Sincrotrón de Protones llegan hasta los 26 GeV. Luego, pasan por el Super Sincrotrón de Protones, donde son acelerados hasta los 450 GeV, y finalmente, en el LHC alcanzan una energía de centro de masas de 13 TeV. Una vez en el LHC, se hace girar a los dos haces en sentidos opuestos durante varias horas antes de

²Los protones no se aceleran en haces continuos, sino que se lanzan en paquetes de manera que entre colisión y colisión exista un tiempo sin choques, crucial para que la electrónica procese la lectura sin saturarse.

hacerlos colisionar en los diferentes puntos del LHC donde se encuentran los experimentos (lugares donde la trayectoria de los dos haces se cruzan).

4.1. El Solenoide Compacto de Muones (CMS)

El CMS [26, 27] es uno de los detectores de múltiples propósitos del acelerador LHC y cuenta con una participación de más de 4000 físicos de partículas, ingenieros, informáticos, técnicos y estudiantes de alrededor de 240 institutos y universidades de más de 50 países. Recibe el nombre de *compacto* dado su gran peso (unas 12500 toneladas) en un diámetro de no más de 15 metros, con una longitud total de 21.6 metros. Junto con el detector ATLAS, forman los dos grandes detectores de múltiples propósitos de la construcción de LHC. El hecho de que ambos tengan funciones parecidas, pero sean detectores diferentes permite la verificación de los estudios encontrados por uno en el otro, como pasó con el descubrimiento del bosón de Higgs en 2012 [23, 24].

El detector CMS cuenta con varios subdetectores especializados en una función y tipo de partículas concretas. A grandes rasgos, podemos distinguir: trackers, calorímetros y detector de muones (aunque también es un tracker, que detecta a los μ). Los primeros se utilizan para conocer el camino que siguen partículas cargadas al atravesarlos. Es por ello que deben ser lo suficientemente finos para no frenar a las partículas ni quitarles demasiada energía. Conocido el radio de la trayectoria curva (debido al campo magnético en el que se encuentran los detectores), podemos saber el momento de la partícula. Los segundos, calorímetros, tienen la función de frenar completamente a algunas partículas (no tienen por que ser cargadas), obteniendo de esta manera su energía. Una vez la partícula cede toda su energía en el calorímetro no podrá seguir viajando a través del detector, así que el orden de los diferentes subdetectores es importante.

En la figura 4.2 se muestra la estructura de los diferentes detectores que forman CMS. En la parte más cercana al haz (donde ocurren las colisiones) se encuentra el tracker de silicio (permite medir el momento de partículas cargadas como electrones, muones, hadrones cargados). Rodeando a este, se encuentra el primero de los calorímetros, el electromagnético (ECAL) que frena y obtiene la energía tanto de fotones (no detectados por el tracker) como de electrones. Tras este detector, el segundo de los calorímetros, el hadrónico (HCAL), encargado de medir la energía de todo tipo de hadrones y sus productos de decaimiento (tanto neutros, como los neutrones, como cargados como piones, kaones, etc.). Por último, tras el solenoide superconductor se localizan los cuatro detectores de muones, ya que estas partículas son capaces de atravesar todos los detectores más internos y llegar al final del detector. Estos detectores de muones también son trackers, pero al ser las únicas partículas (muones) que llegan a esta parte de CMS, sabiendo el momento (y su masa), magnitudes como la energía quedan bien definidas.

El sistema de coordenadas adoptado en CMS tiene como origen el centro de la colisión en el interior del experimento, de forma que el eje y apunta verticalmente hacia arriba, el eje x radialmente perpendicular al eje y en el plano de la colisión, y el eje z en la dirección del haz. Así, las coordenadas que se utilizan para describir las posiciones de los diferentes eventos y medidas son, además de la coordenada z :

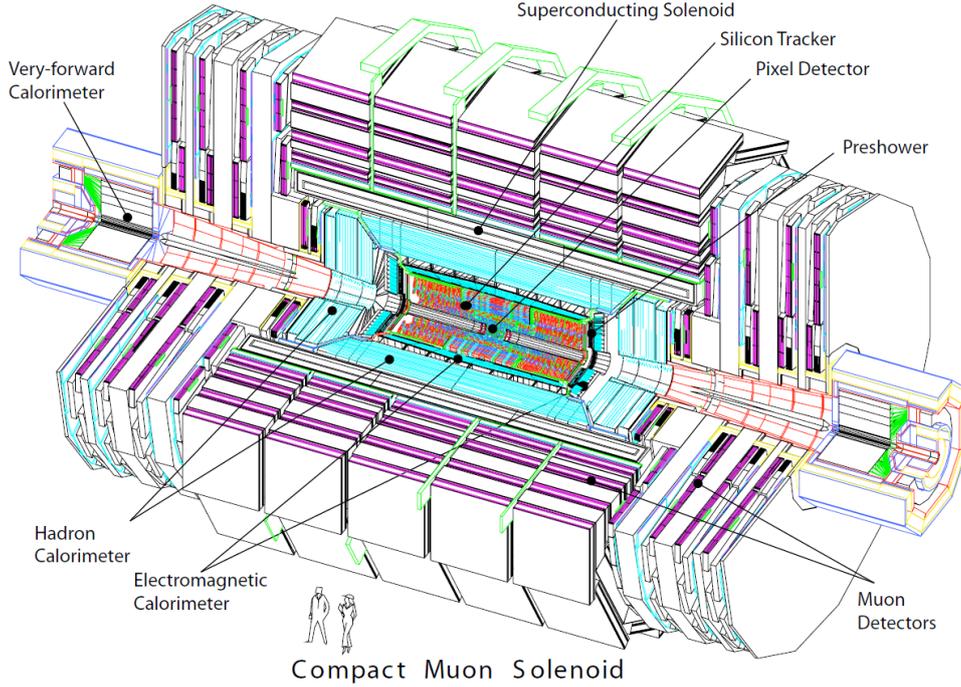


Figura 4.2: Esquema de los detectores del CMS en el LHC [26].

- Ángulo azimutal: medido desde el eje x en el plano $x - y$:

$$\phi \equiv \arctan(y/x). \quad (4.3)$$

- Pseudorapidez: definida a partir del ángulo polar $\theta = \arctan(\sqrt{x^2 + y^2}/z)$ como:

$$\eta \equiv -\ln \tan(\theta/2). \quad (4.4)$$

Además, tanto la energía como el momento que se miden en el plano transversal se denotan como E_T y p_T respectivamente, junto a la energía faltante en el plano transversal (que como dijimos podía asignársela a los neutrinos) E_T^{miss} .

La selección de coordenadas cilíndricas, cambiando el ángulo polar por la pseudorapidez es conveniente dado que a valores de η fijos, la multiplicidad de partículas de alta energía es aproximadamente constante.

En las siguientes subsecciones se explica un poco más detalladamente cada detector de CMS, su composición y función.

4.1.1. Tracker de Silicio

La parte más interna del detector, capaz de reproducir los caminos de las partículas cargadas al estar inmerso en un campo magnético de 3.8 T para poder estimar el momento transversal de estas. Está compuesto de tres tipos de detectores: píxeles, silicon strip y micro strip gas chambers (MSGCs), ocupando las regiones de más alta, media y más baja densidad de partículas respectivamente (es decir, partes más internas, medias y externas del tracker). Cuánto mayor sea la granularidad de estos detectores, mayor será la resolución

de la medida. Por ello, se colocan en la parte más interna los píxeles, mientras que en la parte más externa del tracker se sitúan los MSGCs, en favor también de reducir el material justo antes del calorímetro electromagnético. Gracias al silicio, en las partes más internas e intermedias se consiguen unas altas densidades de detectores con espesores de no más de $300 \mu\text{m}$ [28].

Sistema de píxeles

De los eventos más interesantes en el LHC son los que contienen b -jets generados a partir del decaimiento de partículas pesadas, así que la forma más eficiente de etiquetar a estos jets es obtener su camino lo más cerca posible del vértice donde se generaron. Dada la alta densidad de flujo de partículas, los píxeles confirmarán o rechazarán los caminos propuestos por las capas más externas del tracker. Una vez el camino de la partícula haya sido reconstruido a partir de varios píxeles, se podrá extrapolar este camino hasta el vértice con gran precisión. Este sistema de píxeles tiene dos capas cilíndricas alrededor del eje z , y dos discos finales a cada lado del cilindro [28].

Tracker de bandas de silicio

Los Silicon Strip Tracker (SST) están basados en micro-bandas de silicio en la región intermedia del tracker de CMS. En combinación con los sistemas de píxeles y los MSGCs permiten reconocer patrones, reconstruir caminos y medir momentos para partículas de momento transversal mayor a $2 \text{ GeV}/c$. La efectividad de este tipo de detectores fue previamente comprobada en experimentos como LEP y CDF. Cubre aproximadamente un área de 70 m^2 , con 5 capas cilíndricas y 10 discos tapando el cilindro en cada lado [28].

Micro strip gas chambers (MSGCs)

El principal reto para detectar los caminos de las partículas en LHC es el reconocimiento de patrones en ambientes altamente congestionados (con muchas partículas). Una mejora en este reconocimiento implica una alta granularidad en los detectores. La introducción de los MSGCs (contadores proporcionales de gas fabricados con tecnología microelectrónica) permiten superar esta limitación. Además, la distribución de carga en más de una banda debido a difusión permite el uso de algoritmos de interpolación para lograr una resolución de posición mucho mejor que teniendo en cuenta solo la lectura central. Esta es la principal ventaja de las bandas de gas respecto a los detectores de micro-bandas de estado sólido, permitiendo cubrir grandes áreas con un número razonable de canales electrónicos. La principal desventaja es la velocidad de recolección de carga primaria, al no poderse reducir por debajo de dos cruces de paquetes (“*bunch-crossings*”) [28].

4.1.2. Calorímetros

Calorímetro electromagnético (ECAL)

Rodeando al tracker nos encontramos con el calorímetro electromagnético (ECAL) [29] que juega un papel fundamental en el estudio de la física de la ruptura de simetría electrodébil. Formado por un cristal centelleador de tungsteno (PbWO_4) que ofrece la mejor resolución en la medida de energía para electrones y fotones que depositan su energía en el interior del volumen homogéneo de cristal del calorímetro. Las partes en las que se divide este calorímetro son: la parte del cilindro (“*barrel*”), los discos que tapan el cilindro (“*endcap*”) y el “*preshower*”.

Calorímetro hadrónico (HCAL)

En unas capas más externas está presente el calorímetro hadrónico (HCAL) [30], que con un funcionamiento parecido al electromagnético, mediante centelleadores es capaz de medir la energía de hadrones y sus productos de decaimiento. Tanto la parte cilíndrica como los discos a ambos lados del cilindro experimentan el intenso campo magnético de casi 4 T producido por el solenoide de CMS, así que estos detectores están hechos de material no magnético (aleación de cobre y acero inoxidable). El diseño del calorímetro hadrónico posee una buena granularidad, moderada resolución de energía y profundidad suficiente para frenar a los hadrones en su interior. Este calorímetro consta de 4 partes: igual que en el electromagnético, la parte cilíndrica (“*barrel*”) y los discos a ambos lados del cilindro (“*endcap*”), además de un cilindro más externo que rodea al resto (“*outer barrel*”) y un calorímetro delantero (“*hadronic forward*”).

4.1.3. Sistema de muones

Es el último de los detectores del CMS (el más externo), y es capaz de medir las trayectorias (y por tanto verificar con los resultados del tracker de silicio en la parte más interna del detector) de muones. Únicamente estas partículas y los neutrinos llegarán a la última parte del detector CMS sin depositar su energía en los calorímetros [31].

El detector de muones usa tres tecnologías diferentes: tubos de deriva (“*drift tubes*”, DT) en la zona cilíndrica, cámaras de línea catódica (“*cathode strip chambers*”, CSC) en la región a sendos lados del cilindro y cámaras de placas resistivas (“*resistive plate chambers*”, RPC) en ambas partes. Aunque la geometría y materiales de los tres tipos de subdetectores en el sistema de muones son diferentes, su funcionamiento es muy parecido: consisten en un cátodo y un ánodo en el interior de un volumen gaseoso, en el que al pasar un muón, golpea a los electrones de los átomos del gas, que siguiendo el cabo eléctrico producen una señal [31].

A modo de resumen y para poder visualizar todos estos subdetectores del CMS, en la siguiente figura se muestra una sección de CMS señalando cada parte.

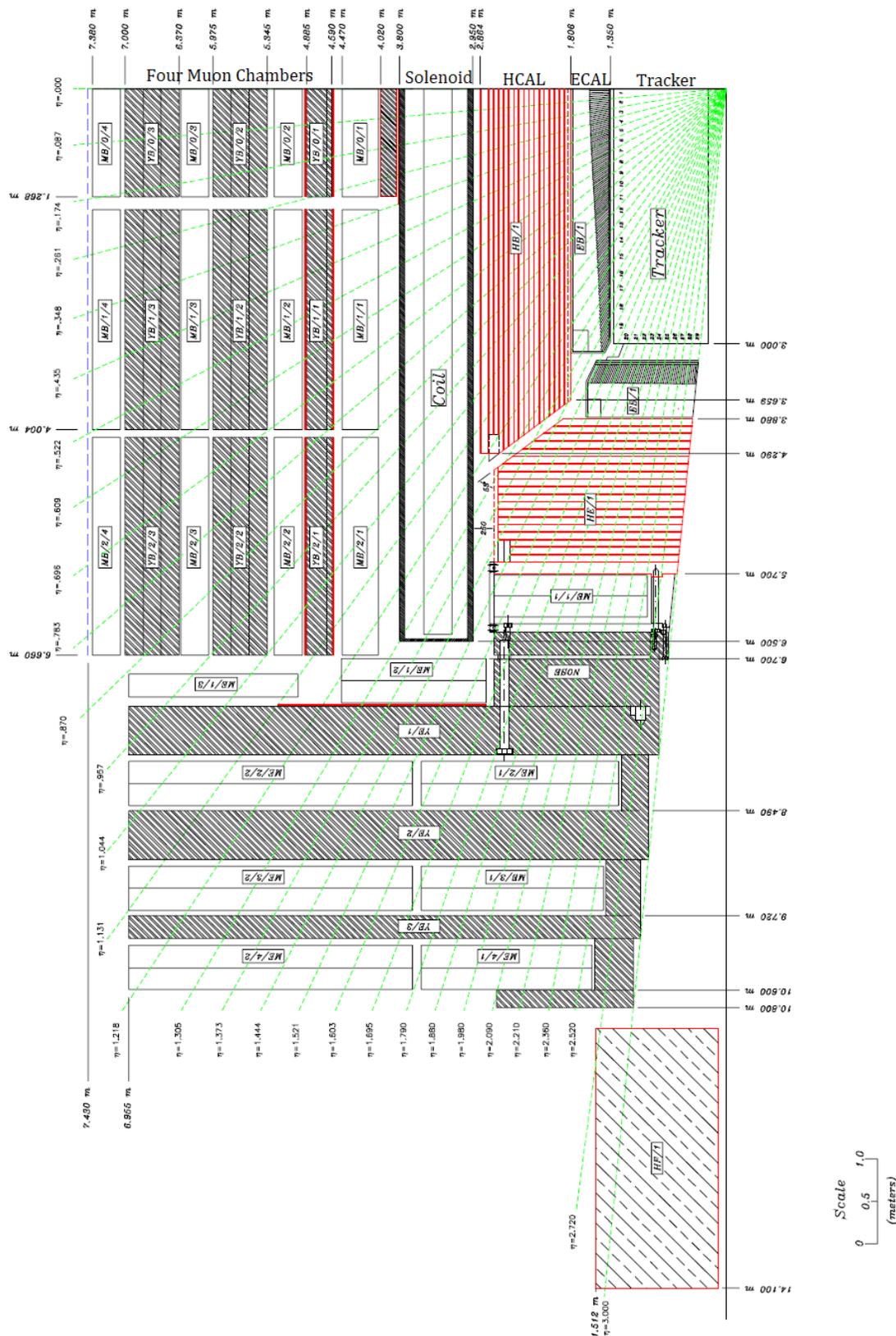


Figura 4.3: Sección del detector CMS en el que se muestran todos los subdetectores que contiene. La parte derecha de la figura sería la zona más interna del detector (al lado del haz de partículas) donde se encuentra el tracker. La parte de la izquierda, es la más externa. Además se señalan los diferentes valores de la coordenada pseudorapidez, η [30].

Capítulo 5

Análisis

5.1. Selección de objetos y eventos

Desde que se produce el fenómeno físico en la colisión en el detector CMS, hasta que analizamos los datos utilizados por ejemplo en este trabajo, se realizan una serie de pasos que permiten una primera verificación de las medidas guardadas. Además, dado que cada 25 ns se cruzan dos paquetes de partículas a colisionar, y un evento típico tiene un tamaño de 1 MB, tendríamos que almacenar unos 40 TB/s, algo completamente fuera del alcance de cualquier dispositivo. Por ello, son necesarios diferentes criterios de selección que permitan almacenar únicamente eventos que vayan a estudiarse.

5.1.1. DAQ y trigger

Antes que nada, el sistema de adquisición de datos (DAQ) obtiene información de una propiedad física, normalmente multi-dimensional, de la señal de los detectores. Es entonces cuando los componentes electrónicos del detector convierten los parámetros específicos de las medidas en señales eléctricas. Este primer paso permite al trigger la posibilidad de analizar los datos a guardar. Este trigger no es más que un conjunto de condiciones que se imponen a las medidas recogidas para saber si guardar o no el evento casi de manera inmediata. Entonces, el principal objetivo del sistema que forman el DAQ y trigger es almacenar la mayor cantidad de datos de los procesos a estudiar con el menor coste posible (en el menor tiempo). Para ello se requiere una eficiencia alta.

La forma de discernir si un evento es o no relevante se produce aplicando cortes (véase cuadro 5.1) a magnitudes como el momento transverso de las partículas medidas, la pseudorapidez, o verificando estados como el aislamiento de leptones (definiendo el radio de un cono, un leptón será aislado si en el interior de dicho cono no se encuentran más partículas) ya que provienen del mismo fenómeno los leptones a los que decae uno de los bosones W (leptón más neutrino) que posibles leptones que se puedan generar en la hadronización de los quarks.

5.1.2. b-tagging

Otra de las características importantes para el caso de la identificación de jets es conocer si se tratan de o b -jets [14], utilizando el algoritmo llamado “*Combined Secondary Vertex*” (DeepCSV) [32], que se basa en el alto tiempo que tardan los quarks b en hadronizar (lo hacen, no en el vértice primario, sino en un vértice secundario tras desplazarse), utilizando además “*deep neural networks*”. Como respuesta, el algoritmo devuelve para cada jet un valor real entre 0 y 1, dependiendo de si el jet muy probablemente sea liviano, o sea un b -jet respectivamente.

Además, se definen los denominados puntos de trabajo (“*working points*”) laxo, medio y estricto (“*loose*”, “*medium*” y “*tight*”), como los cortes en la variable discriminante cuando la probabilidad de identificar de manera errónea un jet liviano como b -jet es del 10 %, 1 % y 0.1 % respectivamente. Para las medidas realizadas durante el año 2017 por CMS, el valor de los puntos de trabajo son: 0.1522 para laxo (10 %), 0.4941 para medio (1 %) y 0.8001 para estricto (0.1 %) [32].

5.1.3. Selección de eventos

En este estudio trabajamos con decaimientos semi-leptónicos de eventos $t\bar{t}$ +jets, en favor de reducir el fondo de los decaimientos hadrónicos, pero teniendo algo más de estadística que en las desintegraciones puramente leptónicas. Para aumentar la eficiencia de los eventos seleccionados por el trigger, imponemos una serie de cortes más estrictos que los del trigger, para los leptones aislados y los dos jets más energéticos de cada evento (véase cuadro 5.1).

Magnitud	e^{Iso}	μ^{Iso}	1º jet	2º jet
$p_{\text{T}}^{\text{cut}}$ [GeV]	35	30	40	30
η^{cut}	2.1	2.4	2.4	2.4

Cuadro 5.1: Cortes realizados a partir del trigger de CMS para almacenar eventos con magnitudes físicas relevantes. A los propios cortes del trigger realizamos nuevos cortes (más restrictivos) para aumentar la eficiencia de los eventos con los que se trabajan. Se aplican cortes al momento transversal de electrones y muones aislados y a los dos primeros jets más energéticos del evento [33, 34, 35].

Además, queremos estudiar eventos con jets adicionales, así que teniendo en cuenta las partículas y jets presentes en las diferentes categorías de $t\bar{t}$ +jets atendiendo al sabor de los adicionales (véase cuadro 3.4), los eventos con los que trabajamos tendrán en principio las siguientes características:

- Estrictamente 6 jets: del propio decaimiento $t\bar{t}$ serán un b -jet de cada top, y el par $q\bar{q}$ al que decae uno de los bosones W hadronizarán dando lugar a dos jets livianos (casi siempre). Esto hace un total de 4 jets. Los otros dos lo imponemos para que haya dos jets adicionales, sumando un total de 6 jets en el evento.
- Al menos 2 b -jets: de los 6 jets del evento, dos tendrían que provenir de quark b (uno

por cada top que decae en $b + W$). El criterio de “*working point*” que se sigue para comprobar estos al menos 2 b -jets es el estricto, es decir, un error al identificar jets livianos como b -jets del 0.1 %.

Estos no serán los criterios de selección definitivos, sino que para mejorar las eficiencias de cada método de reconocimiento de jets adicionales se aplicarán algunos nuevos criterios más restrictivos (siempre partiendo de estos generales).

5.2. Gráficos de control

Para validar la concordancia entre los datos de simulaciones MC de cada proceso, y los datos experimentales, se representan en los siguientes histogramas las principales magnitudes tanto del leptón del evento como de los jets y de la energía perdida (energía que falta en el evento, que se suele asignar al neutrino). Al tener naturaleza de comprobación únicamente estos gráficos, para la representación de los mismos no se han aplicado los criterios de selección de eventos de la subsección 5.1.3, sino que sólo hemos tomado eventos con al menos 4 jets, siendo al menos 2 de los cuales b -jets con criterio laxo.

Para los gráficos de control, se representa el número de eventos frente a la magnitud en cuestión, separando los casos en los que el leptón del evento es un electrón (gráficos de la izquierda) y en los que es un muón (gráficos de la derecha). Agrupando a todos los fondos expuestos en la subsección 3.3.1 se representan de azul como $Nont\bar{t}Bkg$. En color violeta se encuentran los eventos procedentes de decaimientos leptónicos y hadrónicos del par $t\bar{t}$, $t\bar{t}Bkg$. Finalmente, las cuatro categorías de $t\bar{t}$ +jets de la sección 3.4 se encuentran con colores rojo ($t\bar{t}LF$), naranja ($t\bar{t}c\bar{c}$), amarillo ($t\bar{t}b$) y verde ($t\bar{t}b\bar{b}$).

5.2.1. Leptón

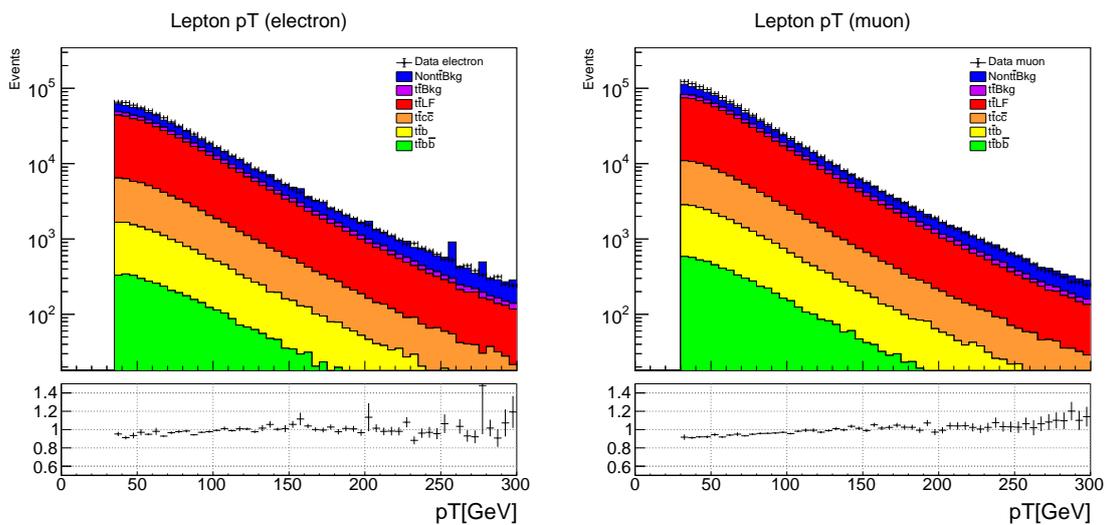


Figura 5.1: Representación del momento transverso del leptón del evento (izquierda, eventos con electrón, derecha, con muón) para simulaciones MC y datos.

Podemos ver en la figura 5.1 el corte en p_T que realizamos tras el trigger. Para el caso del electrón, un momento transversal de 35 GeV, mientras que para el caso de muón de 30 GeV.

De nuevo, podemos observar en la figura 5.2 el corte en η que realizamos tras el trigger. Para el caso del electrón, una pseudorapidez de 2.1, mientras que para el muón de 2.4 (se aprecia 2.5 por el tamaño de los bins que se ha escogido).

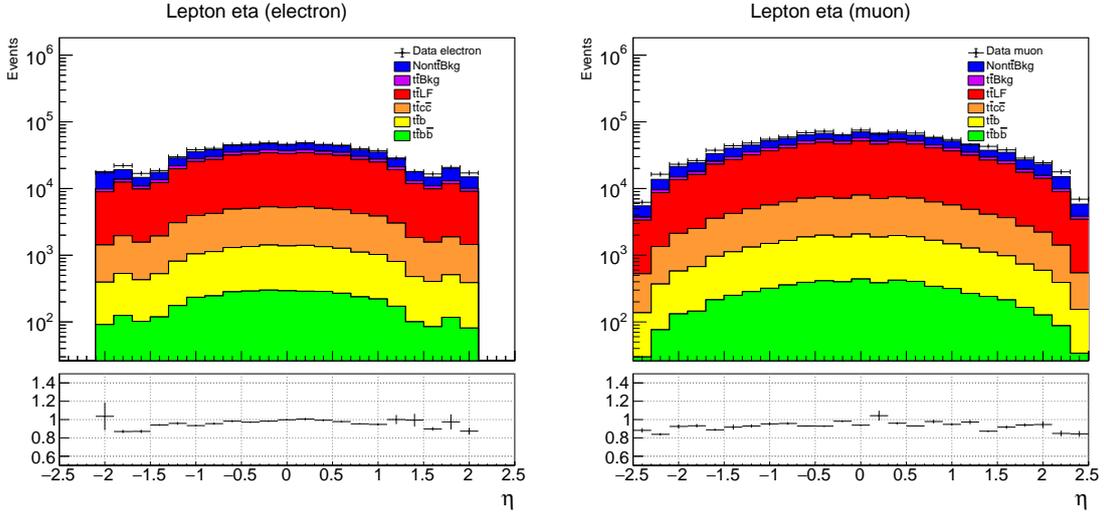


Figura 5.2: Representación de la pseudorapidez (η) del leptón del evento (izquierda, eventos con electrón, derecha, con muón) para simulaciones MC y datos.

5.2.2. Jets

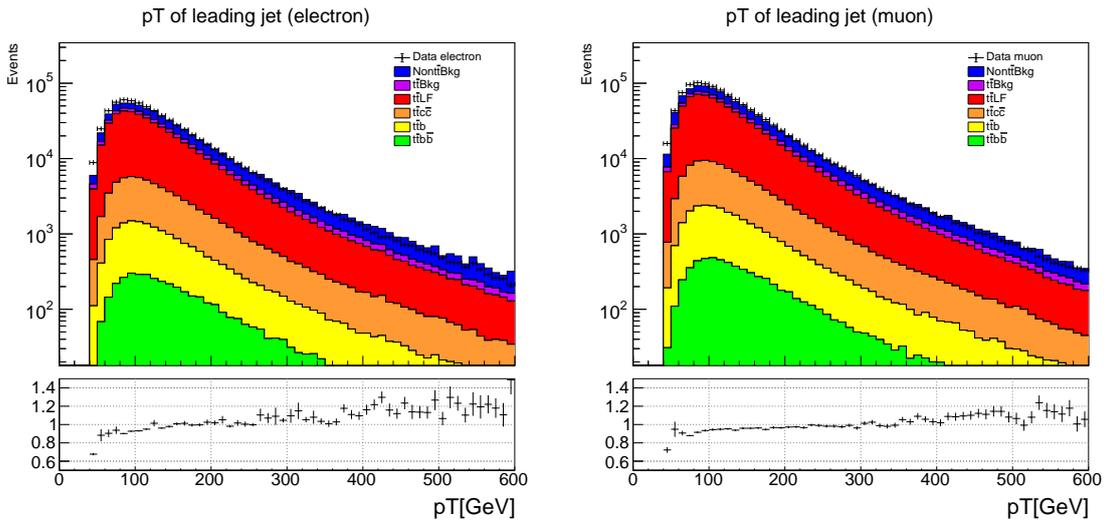


Figura 5.3: Representación del momento transversal del jet más energético del evento (“*leading jet*”) en función del leptón del evento (izquierda, eventos con electrón, derecha, con muón) para simulaciones MC y datos.

En esta ocasión (figura 5.3) se observa el corte en el p_T del jet principal que realizamos tras el trigger, a valor de 40 GeV. Además, procuramos que los eventos de comprobación tuvieran al menos 2 b -jets con el criterio laxo. De ahí los saltos que se observan en ambos gráficos de la figura 5.4 a valor de 0.15 aproximadamente.

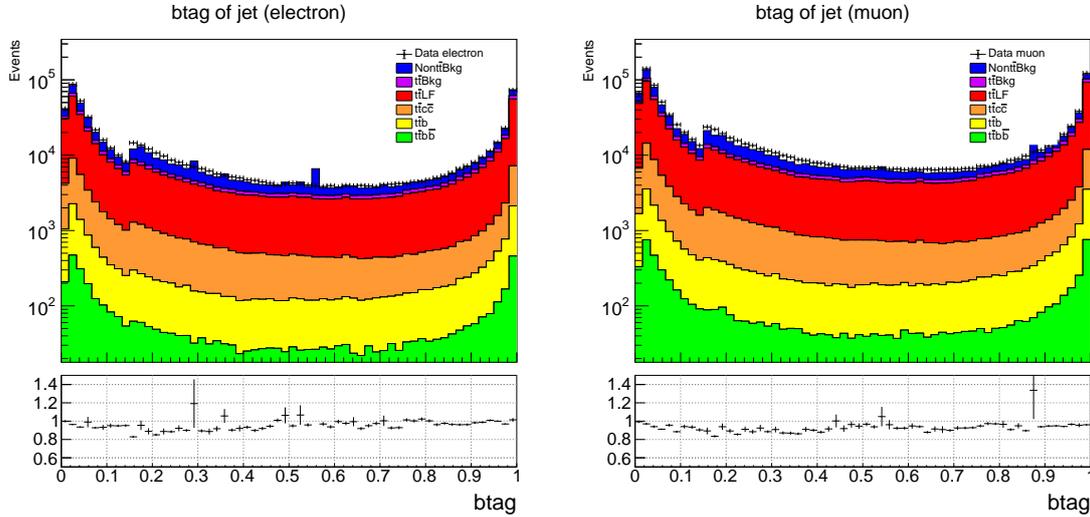


Figura 5.4: Representación del discriminante de b -tagging del jet más energético del evento en función del leptón del evento (izquierda, eventos con electrón, derecha, con muón) para simulaciones MC y datos.

Multiplicidad de b -jets por evento según los tres criterios

- Laxo:

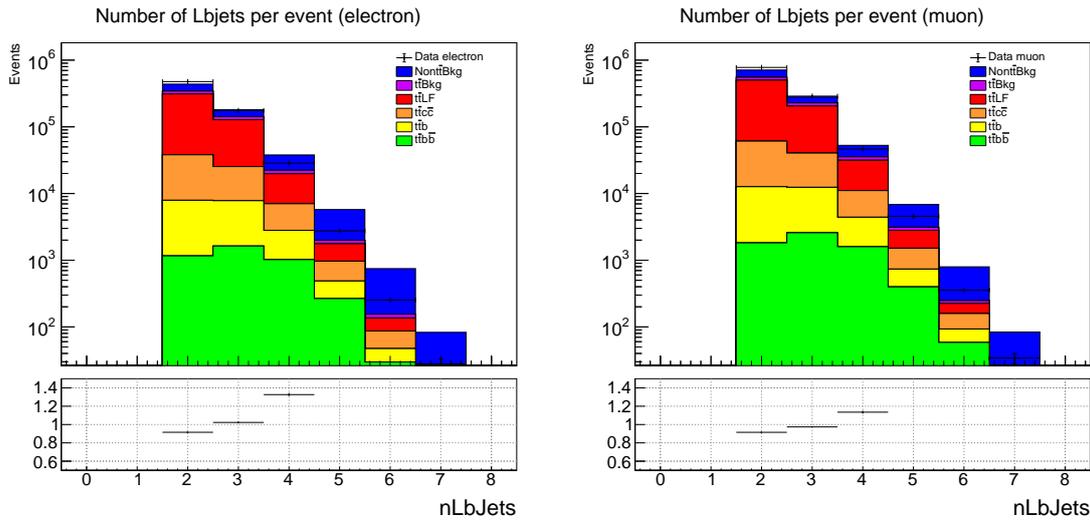


Figura 5.5: Representación de la multiplicidad de b -jets con el criterio laxo (“loose”) en función del leptón del evento (izquierda, eventos con electrón, derecha, con muón) para simulaciones MC y datos.

- Medio:

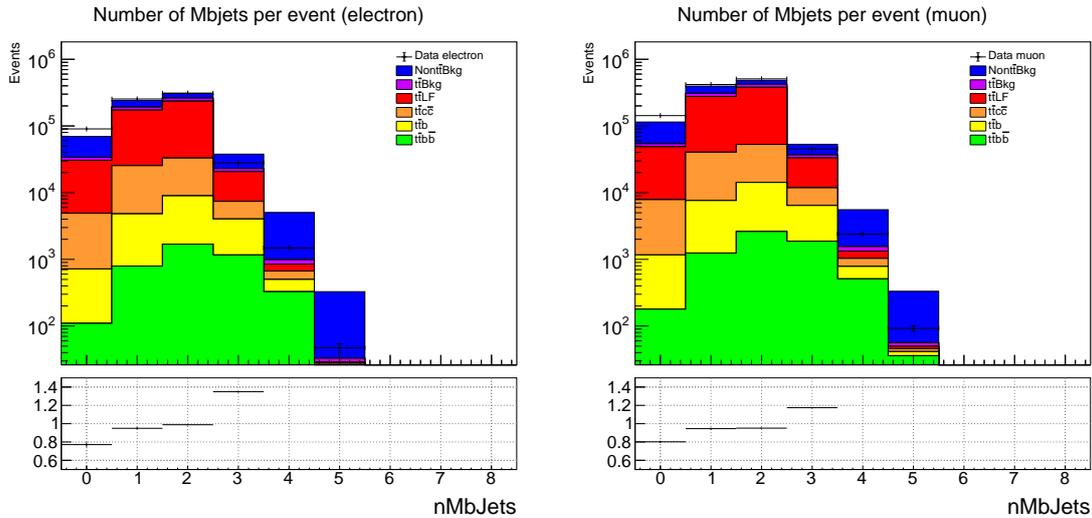


Figura 5.6: Representación de la multiplicidad de b -jets con el criterio medio (“*medium*”) en función del leptón del evento (izquierda, eventos con electrón, derecha, con muón) para simulaciones MC y datos.

- Estricto:

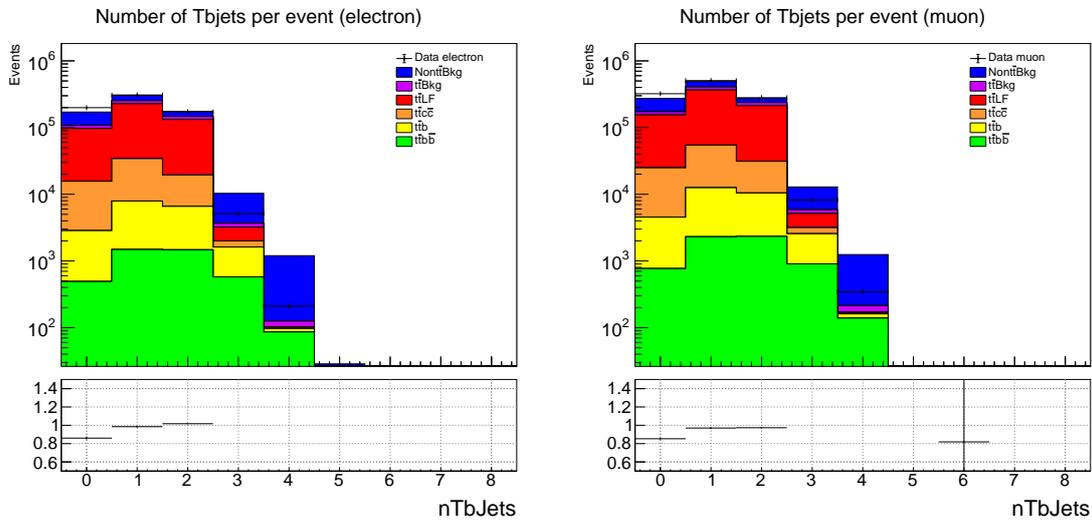


Figura 5.7: Representación de la multiplicidad de b -jets con el criterio estricto (“*tight*”) en función del leptón del evento (izquierda, eventos con electrón, derecha, con muón) para simulaciones MC y datos.

Lo más destacado de estos gráficos es observar como, al haber aplicado un criterio de selección sobre el número de b -jets que han de tener cada evento (2 según el criterio laxo), para este criterio no se observa ningún bin de datos por debajo de 2. Además, destacar también como la media de las distribuciones va descendiendo conforme aumenta la exigencia del “*working point*”. También señalar que el grado de concordancia entre datos

experimentales y simulación MC desciende drásticamente para valores altos de multiplicidad, principalmente debido a los eventos de fondo de QCD, para los que su simulación es algo más compleja que para el resto. De hecho, los datos experimentales en cualquier caso se encuentran dentro del margen de la contribución de $Nont\bar{t}Bkg$.

5.2.3. Energía perdida transversa, MET

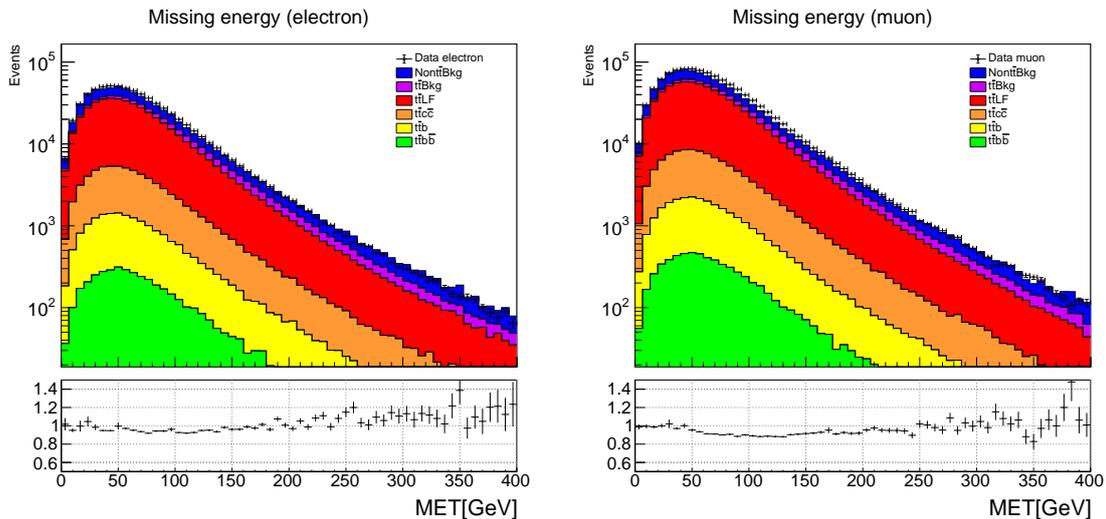


Figura 5.8: Representación de la energía perdida transversa en función del leptón del evento (izquierda, eventos con electrón, derecha, con muón) para simulaciones MC y datos.

Añadir que, en todos los gráficos, el balance entre eventos simulados por MC y datos experimentales está en un rango menor que entre 0.5 y 1.5, lo que indica una buena concordancia de ambas informaciones. Además, una característica general es que la estadística en eventos cuyo leptón principal es un muón es mayor que para aquellos con un electrón (el criterio de corte del trigger es más estricto con electrones, que solo se detectan en el tracker de silicio y en el ECAL, que con muones, que pasan por todo el detector hasta los “*muon chambers*”).

Los datos experimentales únicamente fueron utilizados para verificar las simulaciones MC, y poder trabajar con estas como si de datos se tratasen, permitiéndonos utilizar únicamente en el estudio de los diferentes métodos de identificación de jets adicionales con los eventos $t\bar{t}$. Así que, una vez hemos comprobado que las simulaciones MC se comportan como los datos experimentales, en las siguientes secciones utilizaremos solo los eventos de $t\bar{t}$ de modo de decaimiento semi-leptónico.

5.3. Emparejamiento GenParticles-RecoJets

Con el fin de identificar el origen real de los jets en los eventos de simulaciones MC, lo primero que tenemos que hacer es obtener los eventos con los que vamos a comparar los resultados de cada uno de los métodos. Para ello, trabajamos con simulaciones MC de los

diferentes procesos que generan señal $t\bar{t}$ +jets (las cuatro categorías explicadas en la sección 3.4). Estas simulaciones pueden ser a nivel GenParticles, GenJets o RecoJets. Estas son sus principales características:

- GenParticles: simulación teniendo en cuenta la física del problema, a nivel partónico (quarks, leptones, etc). Cada evento $t\bar{t}$ consta de dos quarks b , un leptón cargado (muón o electrón) y su correspondiente neutrino procedentes del decaimiento leptónico de uno de los bosones W y dos quarks livianos procedentes del decaimiento hadrónico del otro de los bosones W . Para cada partícula a nivel GenParticles, la información que tenemos es su ID (indica que tipo de partícula es, por ejemplo, del 1 al 6 son los quarks d, u, s, c, b, t respectivamente [36]), y sus principales magnitudes como p_T, η, ϕ y E .
- GenJets: simulación teniendo en cuenta la física del problema, pero desarrollando la hadronización de los quarks, considerándolos como jets (jets, leptones, etc). En este caso, cada partícula sigue teniendo su ID (los jets conservan el ID del quark que los produjo, aunque no necesariamente tiene que haber solo 4 jets, ya que pueden entrar también los adicionales). Las magnitudes de las partículas son las mismas que en el caso anterior.
- RecoJets: simulación añadiendo el experimento CMS, es decir, teniendo en cuenta las interacciones de las partículas y los jets con el propio detector. Se simula la respuesta del detector. Ya los jets no conservan el ID, en favor de una nueva magnitud para ellos, el discriminante b-tagging (explicado en la subsección 5.1.2). Las magnitudes físicas son las mismas que en los casos anteriores, pero la estructura de los datos es la misma que para datos experimentales reales.

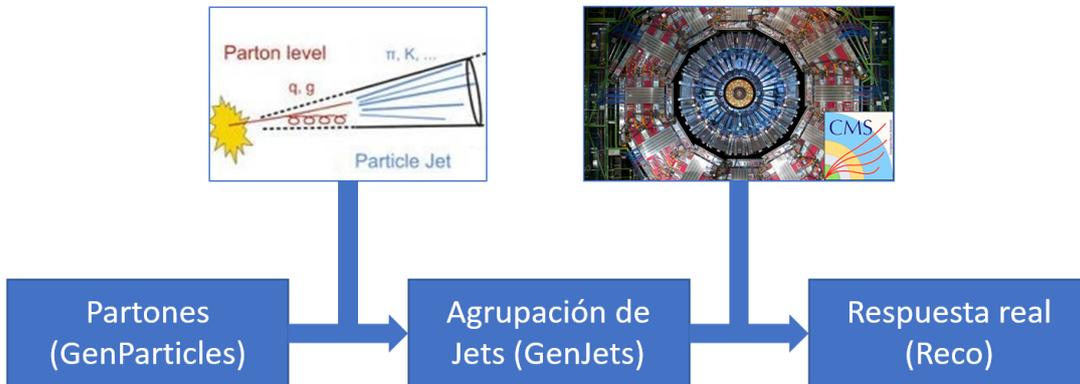


Figura 5.9: Esquema de los diferentes niveles de una simulación partiendo de nivel GenParticles, para agrupar los hadrones de los jets a nivel GenJets, y finalizar introduciendo en la simulación el detector CMS a nivel RecoJets [20].

De esta forma, comprobamos para eventos con estrictamente 6 jets, y al menos dos de ellos b -jets con criterio estricto (véase criterio de selección de eventos en la subsección 5.1.3) todas los posibles emparejamientos entre jets del set de GenJets, y quarks del set de GenParticles. El criterio para discernir si un jet se corresponde con un quark dado es comparar la proximidad de ambos mediante un ΔR , definido como:

$$\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}, \quad (5.1)$$

donde $\Delta\eta$ y $\Delta\phi$ son las diferencias entre las coordenadas η y ϕ respectivamente del quark en GenParticles y del jet en GenJets. Siempre y cuando se cumpla que $\Delta R < 0,4$, emparejaremos al jet con el quark. En caso de que varios jets tengan un $\Delta R < 0,4$ con un mismo quark, asignaremos el emparejamiento con el de menor ΔR , y al otro jet le asignaremos el siguiente quark con segundo menor ΔR (si existe uno con esas condiciones). Puede haber eventos que no consigan emparejar a los 4 quarks del evento a nivel GenParticles, sino que emparejen 3, 2, 1 o incluso ninguno. Esto es, la eficiencia del emparejamiento no es del 100%.

Una vez hemos emparejado quarks a nivel GenParticles con jets a nivel GenJets, hacemos lo propio entre jets de GenJets y RecoJets, teniendo en cuenta únicamente los GenJets emparejados previamente¹ (los no emparejados no se comprueban para este “*matching*”). El criterio para asignar jets de uno a otro set es el mismo que antes, y de nuevo, en el caso de que varios RecoJets se emparejen con un mismo GenJet, tomaremos el de menor ΔR , y al otro buscaremos su segundo mejor $\Delta R < 0,4$ (si existe). También pueden haber eventos que no emparejen a los 4 jets del decaimiento de $t\bar{t}$, pues la eficiencia no es del 100% aquí tampoco. De hecho, en el cuadro 5.2 se muestran las eficiencias de los emparejamientos entre GenJets con los quarks del set GenParticles, y de los GenJets con los RecoJets respecto a la totalidad de las simulaciones MC.

Categoría	$t\bar{t}LF$	$t\bar{t}c\bar{c}$	$t\bar{t}b$	$t\bar{t}b\bar{b}$
GenPart-GenJet (%)	92.76	91.61	89.79	87.63
GenPart-GenJet-RecoJet (%)	90.28	88.32	85.67	81.93

Cuadro 5.2: Eficiencias por categoría de sabor de jets adicionales en porcentaje de los dos emparejamientos realizados sobre eventos con exactamente 6 jets, siendo al menos dos de ellos b -jets con criterio estricto.

En resumen, con estos dos procesos de emparejamientos sucesivos, hemos ligado a cada quark a nivel GenParticles con su RecoJet (mediante la utilización intermedia de GenJets). Así, al haber tomado eventos con únicamente 6 jets, y guardar tras el proceso de emparejamientos los eventos que consiguieron un “*matching*” de los 4 quarks (para no influir en las eficiencias de los métodos de detección de jets adicionales), para cada evento que pase todos estos filtros tendremos 4 jets que pertenecen (según el *matching*), al decaimiento de $t\bar{t}$ y otros 2 que serán por descarte los adicionales. Será con estos eventos con los que trabajaremos en los métodos de detección de jets adicionales, una vez obtenido el resultado que deberían dar estos métodos.

5.4. Reconstrucción cinemática

El primer de los métodos para identificar jets adicionales es el de reconstrucción cinemática (“*Kinematic Fitter*”, abreviado KinFitter [37]) del par $t\bar{t}$. Este algoritmo, utiliza las magnitudes físicas de cada partícula, para encontrar la combinación de jets más adecuada para reproducir el decaimiento $t\bar{t}$, es decir, devuelve los 4 jets del evento $t\bar{t}$.

¹Recordemos que en los GenJets ya están presentes los jets adicionales, y lo que se pretende es identificar los RecoJets que vienen de quarks del evento $t\bar{t}$. Esa es la razón del primero de los emparejamientos, para no trabajar con los adicionales y aumentar así la eficiencia del “*matching*”.

En total, para eventos con 6 jets, el algoritmo comprueba 360 combinaciones ($6 \times 5 \times 4 \times 3$: 6 posibles jets para la primera asignación, 5 para la segunda, 4 para la tercera y 3 para la cuarta) y obtiene, si las condiciones sobre las masas invariantes de los diferentes subsistemas lo permiten, un valor de una función χ^2 , definida en la Ref. [37]. Aquella permutación con menor valor del parámetro χ^2 será la que devuelva el algoritmo.

El algoritmo tiene en cuenta algunos criterios para considerar la permutación o no, como pueda ser que la masa invariante del par de jets livianos que procederían del decaimiento hadrónico del otro bosón W , o la propia masa invariante del sistema leptón cargado más neutrino sea comparable con la masa del bosón W ($m_W = 80,4$ GeV) [37].

Entonces, usando la energía perdida transversa², y las magnitudes del leptón y los jets del evento, el algoritmo selecciona los 4 jets del evento que forman parte (según el KinFitter) de la mejor reconstrucción cinemática del decaimiento de $t\bar{t}$. Los 2 jets que no son considerados como solución del KinFitter, serán los adicionales según este algoritmo.

Una vez hecho esto para todos los eventos con exactamente 6 jets, siendo al menos dos de ellos b -jets con criterio estricto (subsección 5.1.3), y únicamente para aquellos que lograron un emparejamiento de los 4 jets del evento $t\bar{t}$, calculamos la eficiencia de este método:

- Los eventos cuyos 2 jets adicionales según el “*matching*” coincidan con los 2 jets adicionales según el KinFitter tienen una eficiencia de 1.
- Aquellos en los que únicamente 1 de los jets adicionales según el KinFitter coincide con 1 de los adicionales según el emparejamiento tienen una eficiencia de 1/2.
- Los eventos en los que ninguno de los jets adicionales según el KinFitter coincide con los adicionales según el emparejamiento tienen una eficiencia 0.

De esta forma, haciendo el promedio de eficiencias de identificación de jets adicionales sobre todos los eventos considerados, teniendo en cuenta el peso relativo³ de cada uno de ellos, se obtienen las eficiencias (según la categoría) para el método de reconstrucción cinemática mostradas en el cuadro 5.3.

Categoría	$t\bar{t}LF$	$t\bar{t}c\bar{c}$	$t\bar{t}b$	$t\bar{t}b\bar{b}$
$\xi_{\text{KinFitter}}$ (%)	72.07	71.10	70.91	68.51

Cuadro 5.3: Eficiencias por categoría en porcentaje del algoritmo de reconstrucción cinemática identificando los jets adicionales de cada evento con exactamente 6 jets, siendo al menos dos de ellos b -jets con criterio estricto.

Para tratar de mejorar estos resultados, se aplicó condición adicional: considerar que el CSV (b -tagging) de los jets tomados como b -jets a los que decaen los quarks top pasen el

²Con ella, el KinFitter es capaz de reconstruir el momento del neutrino procedente del decaimiento leptónico de uno de los bosones W (junto al leptón cargado de entrada).

³Cada evento tiene una variable real, peso del evento, que viene a representar la probabilidad de que este se produzca, de forma que al trabajar con los eventos, si los pesamos, los más probables tendrán más relevancia en nuestros cálculos.

“*working point*” laxo, es decir, que tome solo los jets que puedan considerarse b -jets según el criterio “*loose*” del b -tagging. Recordemos que el criterio de selección de eventos es de exactamente 6 jets por evento con al menos 2 b -jets de criterio estricto (subsección 5.1.3), es decir, si hay al menos 2 b -jets con criterio “*tight*”, siempre habrá al menos 2 b -jets con criterio laxo (este segundo criterio es menos restrictivo que el primero, y los b -jets estrictos son b -jets laxos, pero al contrario no tiene por qué).

Con esta mejora, en el cuadro 5.4 se muestran los nuevos resultados para las eficiencias del KinFitter.

Categoría	$t\bar{t}LF$	$t\bar{t}c\bar{c}$	$t\bar{t}b$	$t\bar{t}b\bar{b}$
$\xi_{\text{KinFitter}}$ (%)	82.68	79.84	75.60	70.35

Cuadro 5.4: Eficiencias por categoría de sabor de jets adicionales en porcentaje del algoritmo de reconstrucción cinemática mejorado (asumiendo b -jets del decaimiento $t\bar{t}$ con criterio laxo) identificando los jets adicionales de cada evento con exactamente 6 jets, siendo al menos dos de ellos b -jets con criterio estricto.

Comparando los cuadros 5.3 y 5.4, vemos la mejora en eficiencias de más de diez puntos en el caso de $t\bar{t}LF$ y de casi dos puntos en el caso de $t\bar{t}b\bar{b}$.

Se trataron otras posibles mejoras como cortes en la variable interna del KinFitter χ^2 , tomando solo eventos con bajo valor de la misma (cuanto menor es χ^2 , mayor es la probabilidad de que la permutación sea la correcta), pero las eficiencias no mejoraban de manera significativa (algunas fluctuaciones en la tercera cifra decimal de las eficiencias).

Otra posible forma de comprobar la bondad del emparejamiento es obtener la fracción de eventos (pesados correctamente) que asignan:

- a los dos jets con mayor CSV del evento con origen de los quarks top (serían los b -jets a los que decaen cada quark top, junto con los bosones W).
- al menos a uno de los dos jets con mayor CSV del evento con origen de los quark top (b -jet acompañando al bosón W en el decaimiento de uno de los quark top).

Esto tiene especial relevancia sobre todo en eventos $t\bar{t}LF$ y $t\bar{t}c\bar{c}$, ya que los b -jets que estén en estos eventos debería usarlos el KinFitter como aquellos jets a los que decae cada uno de los quarks top junto al bosón W (los adicionales no son b -jets, y los jets a los que decae hadrónicamente uno de los bosones W son en su mayoría jets livianos). Por el contrario, en eventos $t\bar{t}b$ y $t\bar{t}b\bar{b}$, al ser los adicionales b -jets, podría pasar que los de mayor CSV fueran estos, y no los del decaimiento $t\bar{t}$, al haber 3 y 4 b -jets respectivamente por evento (siguiendo los criterios de selección de eventos de la subsección 5.1.3).

En el cuadro 5.5 se muestran esos porcentajes de eventos por categoría. Como vemos, que al menos 1 de los dos jets con mayor CSV del evento forme parte del decaimiento $t\bar{t}$ está prácticamente asegurado en todas las categorías, como sería de esperar dado que estos jets a los que decae el top son b -jets, que por definición tienen un CSV alto. Sin embargo, para que ambos b -jets del evento $t\bar{t}$ sean los de mayor CSV, los adicionales han de no ser

b -jets, lo que permitiría que esta condición se diera. En el caso de que los adicionales sean b -jets, hay más jets con los que confundir al algoritmo de KinFitter.

Condición	$t\bar{t}LF$	$t\bar{t}c\bar{c}$	$t\bar{t}b$	$t\bar{t}b\bar{b}$
2 jets con $>$ CSV (%)	94.48	90.41	43.44	26.00
1 de los 2 jets con $>$ CSV (%)	99.95	99.86	98.99	88.84

Cuadro 5.5: Porcentaje de eventos (pesados) con: primera fila, los 2 jets con mayor CSV del eventos provenientes de $t\bar{t}$; segunda fila, al menos 1 de los dos jets con mayor CSV del evento proveniente de $t\bar{t}$; según el algoritmo de reconstrucción cinemática.

5.5. Boosted Decision Tree (BDT)

Las siguientes dos secciones se basan en métodos de aprendizaje automático (“*machine learning*”). Este tipo de algoritmos toman un set de datos inicial con el que entrenan una red para asignar diferentes pesos a los nodos de la red. Una vez entrenada la red correctamente, el algoritmo debe ser capaz de devolver una respuesta en función de los parámetros introducidos. Las redes con las que se trabajan en este proyecto son redes binarias, dado que la respuesta que necesitamos es que distinga entre señal (jets adicionales) y fondo (jets del evento $t\bar{t}$, ya sean los b -jets que vienen directamente de top o los jets livianos del bosón W que decae hadrónicamente). El entorno utilizado en ambas aplicaciones de aprendizaje automático es TMVA de ROOT [38].

En particular, las BDTs construyen mediante análisis multivariable, una variable discriminante combinando las de entrada, de tal forma que maximice la separación entre señal y fondo: la señal tendrá en general altos valores de esta variable discriminante, mientras que el fondo será desplazado a valores bajos. Su respuesta será determinada por las variables de entrada con las que se entrena la red.

Como paso inicial, las variables de entrada que se utilizaron fueron las propias magnitudes de los jets: su momento y energía transversa, sus dos coordenadas angulares (η y ϕ) y el CSV (b-tagging). Trabajando con la cuarta parte de los eventos con 6 jets, siendo al menos 2 de ellos b -jets según el criterio estricto (véase subsección 5.1.3), para entrenar la red le proporcionamos los 2 jets adicionales y los 4 jets de $t\bar{t}$ del evento, según el criterio del emparejamiento. Por tanto, por cada 2 entradas de señal hay 4 de fondo en cada evento.

En la figura 5.10, vemos las distribuciones de las variables de entrada en el conjunto de datos de entrenamiento separando la señal (azul) del fondo (rojo). Principalmente la red tomará el CSV y en menor medida la coordenada η para construir la variable discriminante, que son las que mayor diferencia tienen entre señal y fondo. Además, en las matrices de la figura 5.11 se representa la correlación entre las variables de entrada, haciendo diferencia entre señal y fondo. Como era de esperar, la relación es notable entre energía y momento de los jets. En el cuadro 5.6 se muestran los pesos (relevancia relativa) de las diferentes variables de entrada en la construcción de la variable discriminante.

Variable	CSV	η	ϕ	p_T	E
Peso ($\times 10^{-1}$)	3.188	1.852	1.827	1.764	1.369

Cuadro 5.6: Relevancia de mayor (izquierda) a menor (derecha) de las variables de entrada (jets separados) tras el entrenamiento (con todas las categorías) de la BDT.

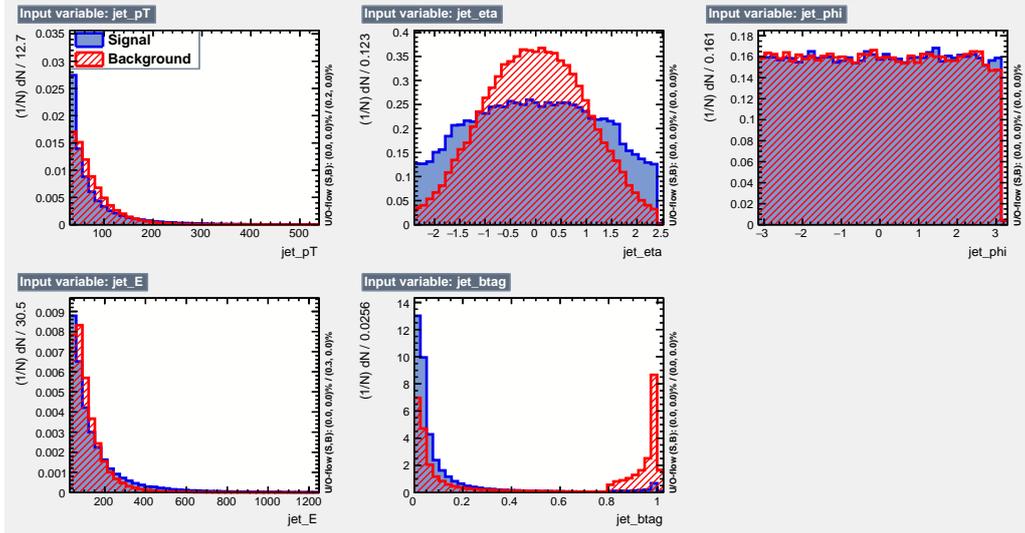


Figura 5.10: Distribución de las variables de entrada (jets separados) para el entrenamiento (con todas las categorías) de la red BDT separando señal (jets adicionales) y fondo (jets del decaimiento de $t\bar{t}$). De izquierda a derecha y de arriba a abajo: p_T , η , ϕ , E , CSV.

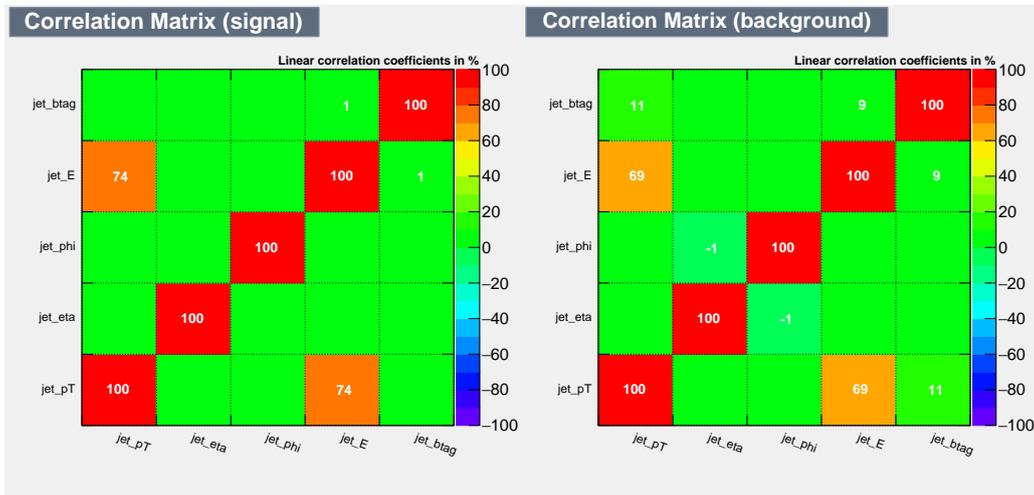


Figura 5.11: Matrices de correlación entre las variables de entrada (jets separados) tanto para señal (izquierda) como para fondo (derecha) para el entrenamiento (con todas las categorías) de la red BDT.

Tras entrenar la red con las variables de entradas descritas, se procede a comprobar la respuesta de la red con nuevos datos (estadísticamente independiente de los primeros), para verificar que no se ha producido sobre-entrenamiento⁴. Esto es lo que podemos comprobar

⁴Este fenómeno ocurre en aprendizaje automático cuando una red se especializa tanto con los datos de

en la figura 5.12, donde vemos como la distribución de la respuesta del set de datos de verificación (histogramas) se ajusta bastante a los datos con los que se entrenó la red (distribución puntual), es decir, no se produjo sobre-entrenamiento.

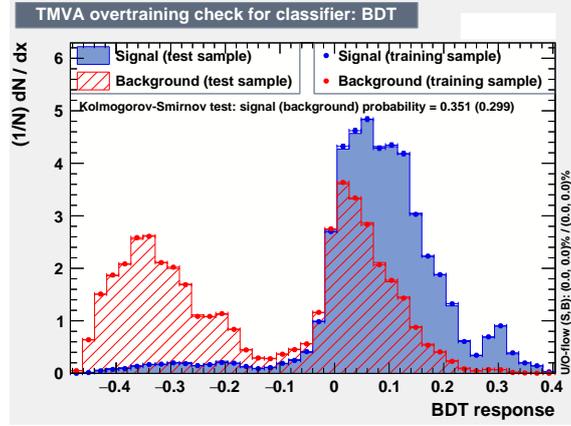


Figura 5.12: Distribución de la respuesta de la BDT (jets separados) entrenada (con todas las categorías) para los sets de datos de entrenamiento (valores puntuales) y datos de verificación (histogramas) para señal y fondo.

Sin embargo, la finalidad principal de la BDT entrenada no está cubierta, ya que la variable discriminante no es capaz de separar completamente la señal del fondo. Podríamos tomar los eventos con una respuesta de la BDT mayor a aproximadamente 0.02 (sería el corte más óptimo para esta red, donde el balance entre señal y fondo se maximiza), aunque la presencia de fondo no sería nula o prácticamente inexistente. Este primer entrenamiento de la red con las variables de los jets sueltos puede resultar no muy conveniente por muchos motivos como la diferente proporción en eventos de las distintas categorías $t\bar{t}$ +jets (explicado en la sección 3.4) dadas sus diferencias: eventos $t\bar{t}LF$ deberían tener un CSV de los jets adicionales cercano a 0 (son jets livianos), mientras que eventos $t\bar{t}b\bar{b}$ deben tener un CSV de jets adicionales cercano a 1.

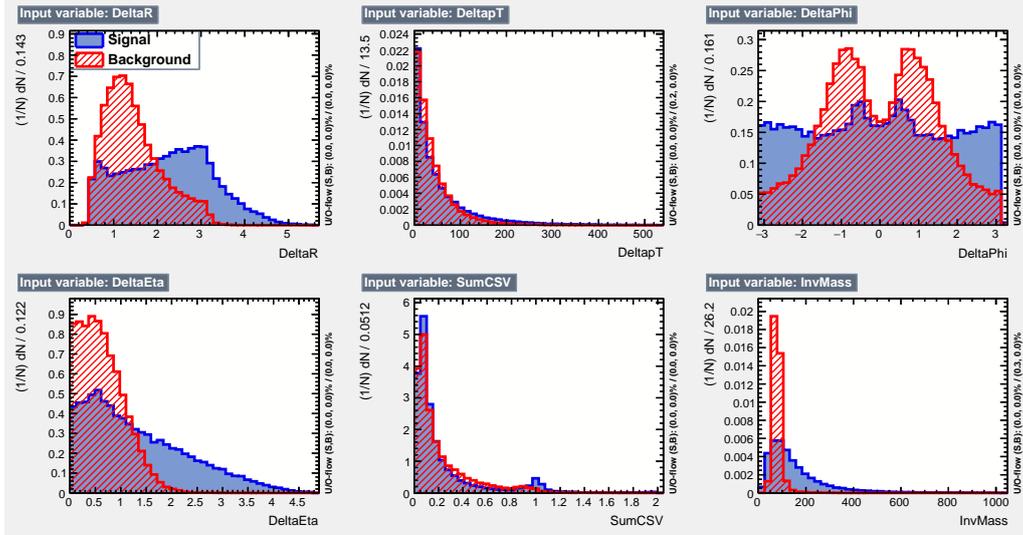
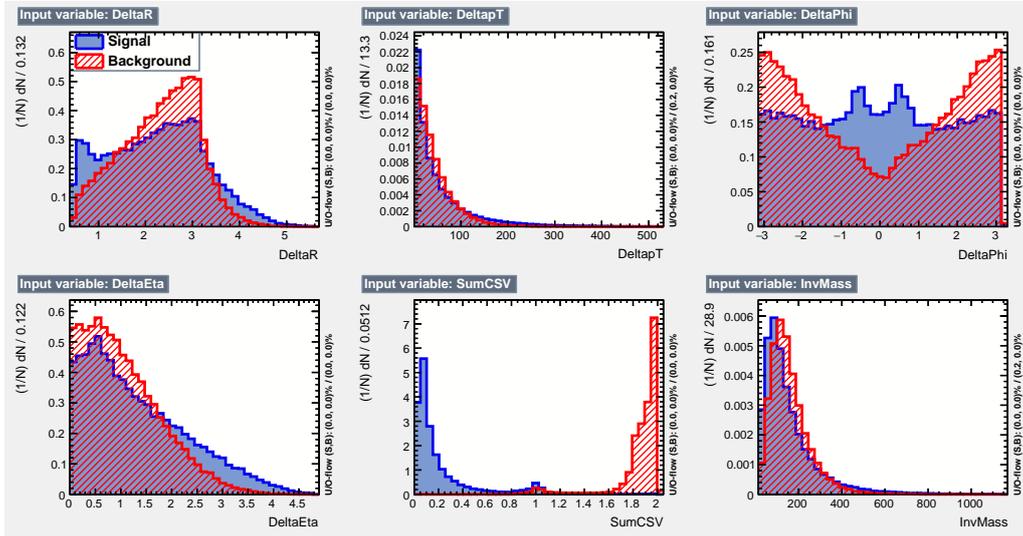
Por ello, se replantea la forma de introducir la información en la red para su entrenamiento. Además, aprovechamos la presencia de dos fuentes de jets para el fondo (b -jets del top y jets livianos del bosón W) para usar dos redes distintas. En la primera se entrenarían las nuevas variables de entrada para la señal (adicionales) y el fondo procedente de W , mientras que en la segunda se haría lo propio pero con el fondo procedente de top. Estas variables de entrada se tomaron relacionado los pares de jets (un par de jets de señal por un par de jets de fondo), y fueron las siguientes: distancia en el plano transversal (ΔR , definida en la ecuación 5.1), diferencias entre momento transversal, η y ϕ (Δp_T , $\Delta\eta$ y $\Delta\phi$ respectivamente), la suma de los CSV del par de jets y la masa invariante del sistema de par de jets:

$$M \equiv |\vec{p}_1 + \vec{p}_2|, \quad (5.2)$$

que no es más que el módulo de la suma de los tri-momentos.

En la figura 5.13 se muestran las 6 variables que se introducen en cada red. En el caso de la red para fondo de W , la separación entre jets del W es mucho menor que en los

entrenamiento que al final solo devuelve buenas respuestas para los datos de entrenamiento, dejando de lado los aspectos que en realidad se quieren aprender a diferenciar.

(a) Señal de adicionales y fondo de W .

(b) Señal de adicionales y fondo de top.

Figura 5.13: Distribución de las variables de entrada (pares de jets) para los entrenamientos (con todas las categorías) de las redes BDT separando señal y fondo. Para ambos, de izquierda a derecha y de arriba a abajo: ΔR , Δp_T , $\Delta\phi$, $\Delta\eta$, suma CSV y masa invariante.

pares de jets adicionales, por eso la distribución de fondo en ΔR pica antes que la de señal. Otra de las variables interesante en este caso es la masa invariante del sistema de jets. Obviamente, esta variable en el caso del fondo es una gaussiana alrededor de la masa del bosón W (aprox. $80.4 \text{ GeV}/c^2$ [13]). Así, las variables que influirán más en el discriminante de la BDT con este fondo de W serán la masa invariante, el ΔR y en menor medida los $\Delta\phi$ y $\Delta\eta$.

Por otro lado, en el caso de la red para fondo de top, aunque $\Delta\phi$ y ΔR pueden influir en la construcción de la variable discriminante de la BDT, la principal es la suma de CSV. Y es que, dada la desproporción entre eventos de las diferentes categorías según el sabor de jets adicionales (véase sección 3.4), la mayoría de eventos son $t\bar{t}LF$, cuyos jets adicionales (señal) son livianos (suma de CSV prácticamente nula), mientras que los jets del

top (fondo) son, por definición b -jets (suma de CSV prácticamente 2). Esta desproporción entre las categorías las veremos cuando halleemos las eficiencias de este (y el posterior) método.

Las matrices de correlación entre variables de entrada tanto para señal como para los dos fondos se muestran en la figura 5.14. Se observan algunas correlaciones importantes entre ΔR , $\Delta\eta$ y la masa invariante en la señal y el fondo de top, mientras que para el fondo de W la única correlación que existe (aunque pequeña) es entre ΔR y Δp_T .

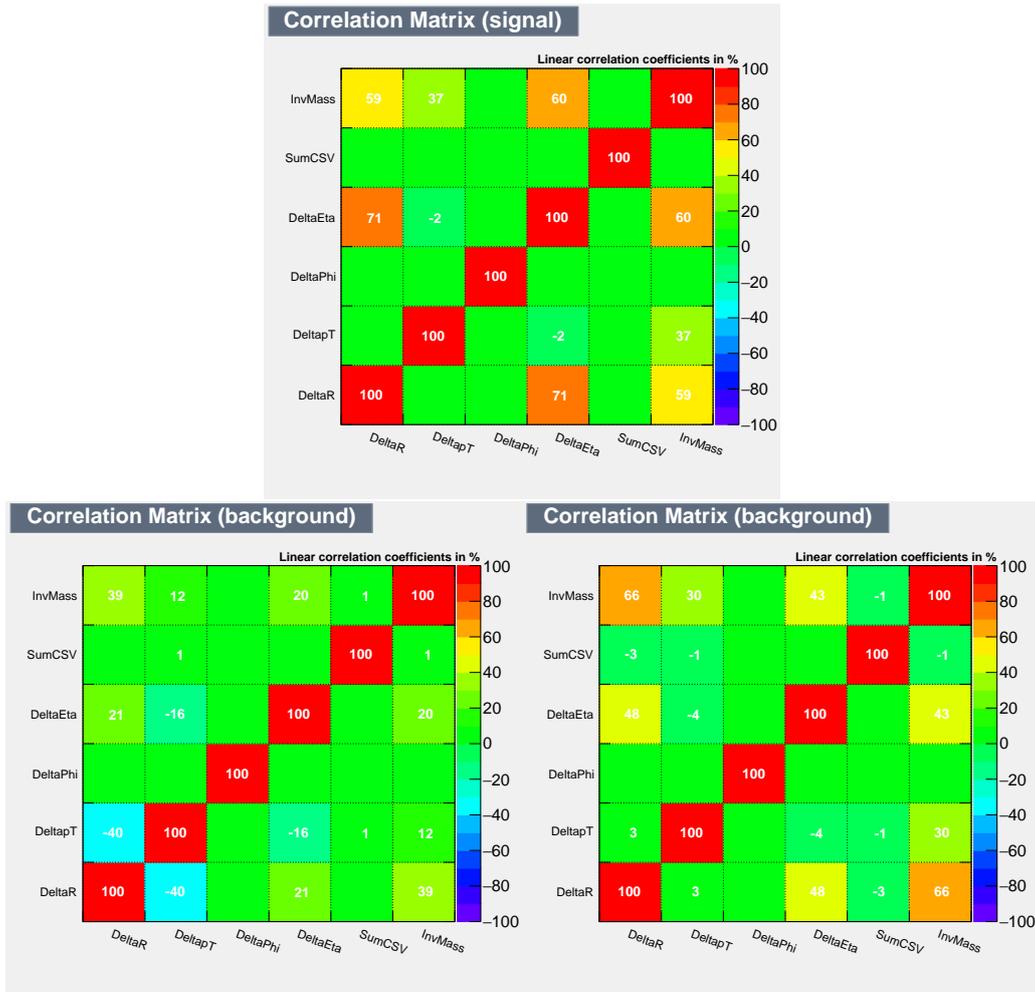


Figura 5.14: Matrices de correlación entre las variables de entrada (pares de jets) tanto para señal (arriba) como para fondo (abajo izquierda: fondo de W ; abajo derecha: fondo de top) para el entrenamiento (con todas las categorías) de la red BDT.

Igual que antes, una vez entrenada la red (con una cuarta parte de los datos disponibles), nos disponemos a comprobar el sobre-entrenamiento. En la figura 5.15 se muestran las respuestas de ambas redes al introducir los sets de test en las BDTs ya entrenadas, y en el cuadro 5.7, los pesos que toma cada una de las redes para las diferentes variables de entrada según su relevancia para construir la variable discriminante de la red.

Las redes no están sobre-entrenadas, y además separan bastante bien la señal del fondo. En el caso de la red con fondo de W , la separación no es tan evidente como en la red con fondo de top, pero aún así, por encima del corte óptimo de la variable discriminante en la

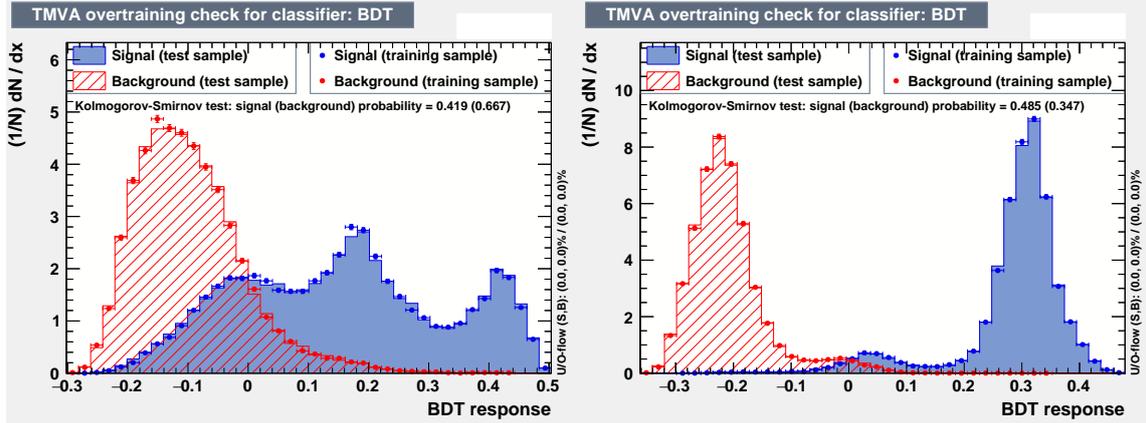


Figura 5.15: Distribución de las respuestas de las BDTs (pares de jets) entrenadas (con todas las categorías) para los sets de datos de entrenamiento (valores puntuales) y datos de verificación (histogramas) para señal y fondo (fondo de W a la izquierda, y fondo de top a la derecha).

Variable (fondo W)	$\Delta\eta$	InvMass	ΔR	$\Delta\phi$	ΣCSV	Δp_T
Peso ($\times 10^{-1}$)	2.281	1.898	1.663	1.544	1.413	1.201
Variable (fondo top)	ΣCSV	$\Delta\phi$	ΔR	$\Delta\eta$	InvMass	Δp_T
Peso ($\times 10^{-1}$)	2.496	1.799	1.677	1.647	1.196	1.185

Cuadro 5.7: Relevancia de mayor (izquierda) a menor (derecha) de las variables de entrada (pares de jets) tras el entrenamiento (con todas las categorías) de la redes BDTs.

primera BDT (corte en 0.1) la presencia del fondo se reduce hasta un 5% del total. En la BDT de fondo de top, la separación es casi perfecta, aunque como ya dijimos, esto se debe principalmente a que esta red toma como principal variable para construir el discriminante, la suma de CSV del par de jets, el ser los jets del top b -jets y haber mayoría de eventos $t\bar{t}LF$, con jets adicionales livianos, esa diferencia es máxima (los adicionales picaban en 0 mientras que los del fondo lo hacían en 2, véase la quinta figura 5.13b, SumCSV).

Todo esto ha sido para entrenar las BDTs diferenciando los jets adicionales de aquellos del top o del bosón W . Para ello, hemos usado la mitad de los datos simulados de los que disponíamos (una cuarta parte para entrenar la red, y otra cuarta parte para verificar que no hubiera sobre-entrenamiento). Ahora toca comprobar la eficiencia de este método en la detección de jets adicionales en eventos con los criterios de selección de la sección 5.1.3. Los eventos deben ser estadísticamente independientes a los del entrenamiento, por lo que usamos la mitad que aún no hemos utilizado, y la cuarta parte usada para la verificación (que no ha participado en el entrenamiento de la red).

El proceso de obtención de las respuestas de las dos redes comienza por el cálculo de las variables de los pares de jets. Así, para eventos con 6 jets, existen 15 combinaciones posibles para formar la pareja de jets adicionales (6 posibilidades para elegir el primero, y 5 para el segundo, $6 \times 5 = 30$, dividido por dos al no importar el orden). Las 15 combinaciones posibles por evento se utilizan para obtener dos respuestas por evento (una para la red de fondo de W , y la otra para el fondo de top). Cuanto mayor sea la respuesta de cualquiera de ellas, más adicionales serán realmente, mientras que cuanto menor sea, mayor posibilidad

de que pertenezcan al fondo existe. Como criterio de selección de la mejor combinación, al tener dos respuestas, tomamos aquella cuya suma de respuestas sea mayor, es decir, el $\max \left\{ \text{BDT}_i^W + \text{BDT}_i^{\text{top}} \right\}$ con $i \in [1, 15]$, siendo BDT_i^W la respuesta de la combinación i -ésima por la red de fondo de W y análogamente $\text{BDT}_i^{\text{top}}$ para la red de fondo de top.

Notar que en el proceso de elección de la mejor combinación en cada evento no se han tenido en cuenta los jets que formarían el par de fondo de W ni los de fondo de top. Esto es principalmente por los objetivos (identificar jets adicionales), aunque también influye el aumento de combinaciones que provocaría que tuviéramos en cuenta los jets de fondo (de 15 combinaciones nos iríamos a 90 por cada evento).

Una vez devueltas todas las respuestas para todos los eventos, y habiendo seleccionado la mejor de ellas (según el criterio establecido), comparamos cuantos de los 2 jets adicionales según la BDT coinciden con los adicionales según el método de emparejamiento, asignando una eficiencia por evento igual que en el método anterior: 1 si los dos jets son los mismos, 1/2 si únicamente coincide uno de ellos, y 0 si ninguno está presente en los adicionales según emparejamiento.

Calculamos las eficiencias por categorías y la total, teniendo en cuenta el peso de cada evento, y las recogemos en el cuadro 5.8, donde además, a modo de justificación del criterio de mejor combinación por evento (mejor suma de respuestas), hemos añadido las eficiencias si únicamente tomáramos en cuenta la respuesta de una u otra red.

Criterio	$t\bar{t}$ +jets (Tot)	$t\bar{t}LF$	$t\bar{t}c\bar{c}$	$t\bar{t}b$	$t\bar{t}b\bar{b}$
Suma BDTs (%)	60.24	61.44	59.49	51.70	42.42
BDT^W (%)	45.63	42.41	42.62	45.17	43.90
BDT^{top} (%)	55.67	57.90	53.72	40.22	25.18

Cuadro 5.8: Eficiencias por categoría de sabor de jets adicionales en porcentaje de las BDTs (separando fondos de W y de top) identificando los jets adicionales de cada evento con exactamente 6 jets, siendo al menos dos de ellos b -jets con criterio estricto.

En cualquiera de las categorías vemos como el criterio de la mejor suma de respuestas de las BDTs tiene una eficiencia mayor que las eficiencias de las redes separadas. Véase como la mayor contribución a la eficiencia total es de la categoría $t\bar{t}LF$ (por la alta cantidad de estadística que posee). Además, las eficiencias van cayendo cuanto más pesados son los jets adicionales.

Estas bajas eficiencias, sobre todo en la categoría $t\bar{t}b\bar{b}$, nos llevó a tratar de mejorar este método, entrenando las redes solo con los eventos de $t\bar{t}b\bar{b}$ (recordemos del capítulo 3, esta categoría es importante al suponer uno de los principales fondos de la señal de $t\bar{t} + H$).

No se describe el proceso porque es análogo al anterior, y además los resultados de esta potencial mejora no fueron los esperados (cuadro 5.9). Aunque si se consiguió aumentar la eficiencia en la categoría $t\bar{t}b\bar{b}$ unos 6 puntos, en el resto de categorías descendió notablemente, no justificándose entonces la mejora (véase apéndice A para más información sobre los resultados del intento de mejora).

Criterio	Total	$t\bar{t}LF$	$t\bar{t}c\bar{c}$	$t\bar{t}b$	$t\bar{t}b\bar{b}$
Suma BDTs (%)	41.69	41.27	41.76	45.04	48.16

Cuadro 5.9: Eficiencias por categoría de sabor de jets adicionales en porcentaje de las BDTs (entrenadas con $t\bar{t}b\bar{b}$, y separando fondos de W y de top) identificando los jets adicionales de cada evento con exactamente 6 jets, siendo al menos dos de ellos b -jets con criterio estricto.

5.6. Deep Neural Network (DNN)

Es el último de los métodos utilizados para la identificación de jets adicionales. También usa aprendizaje automático, pero esta vez con la utilización de redes neuronales profundas. Estas redes suelen tener tantos nodos iniciales como variables de entrada se necesiten, y un nodo final (para redes binarias como nuestro caso), que toma valores entre 0 y 1, indicando una categoría u otra (para este estudio, devuelve 1 si es señal y 0 si es fondo).

Aclarar que para el entrenamiento de estas redes neuronales profundas se hizo uso de la biblioteca de Redes Neuronales de Código Abierto escrita en Python, Keras [39]. Una vez entrenadas las redes y obtenidos los pesos de las mismas, hubo que migrar los archivos de pesos a C++ para trabajar con el entorno que hemos utilizado en el resto del trabajo.

La estructura interna puede ser tan compleja como se quiera, y su morfología puede variar notablemente en el resultado final de eficiencias. Sin embargo, aquí trabajamos con una red básica, con 6 capas internas de $70 \times 50 \times 30 \times 30 \times 20 \times 10$ nodos respectivamente. Además se utilizan las mismas variables de entradas que en el método anterior, las seis variables de pares de jets (ΔR , Δp_T , $\Delta\eta$, $\Delta\phi$, la suma de CSV y la masa invariante del par de jets). Es decir, la red neuronal profunda cuenta con un total de 8 capas: la inicial, de 6 nodos (como el número de variables de entradas), las intermedias descritas anteriormente, y la final, de un único nodo al ser una red binaria, cuya respuesta se encuentra entre 0 y 1 (para más información sobre DNNs, véase el apéndice B).

Una vez definido los modelos de redes (una para cada fondo posible: fondo de W y de top), e inicializadas sus capas, se procedió a entrenarlas siguiendo los mismos pasos que para las BDTs. Las variables de entrada fueron las que se mostraron en la figura 5.13 y las correlaciones entre estas las de la figura 5.14, que ya comentamos previamente.

Con las redes entrenadas, se procedió a la comprobación del sobre-entrenamiento con el set de datos de verificación. En la figura 5.16 se muestran las distribuciones de señal y fondo para los sets de datos de entrenamiento (puntuales) y de verificación (histogramas), y en el cuadro 5.10, los pesos de las variables de entrada en las diferentes redes según la importancia de cada una en la variable discriminante de la red.

Como se observa, la concordancia entre ambos conjuntos de datos a la respuesta de la red, y la separación entre señal y fondo es máxima (mayor en el caso de la red con fondo de top, por las mismas razones que en el caso de la BDT de top).

Al igual que en el método anterior de las BDTs, tras entrenar y comprobar la red, se sigue con la obtención de las respuestas de ambas redes al resto de datos estadísticamente independientes. Al haber dos redes (dos respuestas), usamos el mismo criterio del máximo

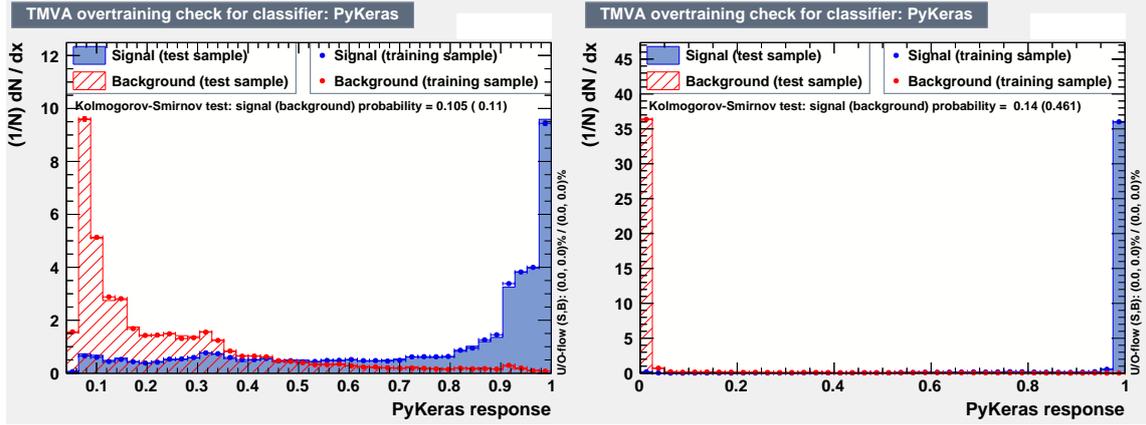


Figura 5.16: Distribución de las respuestas de las DNNs (pares de jets) entrenadas (con todas las categorías) para los sets de datos de entrenamiento (valores puntuales) y datos de verificación (histogramas) para señal y fondo (fondo de W a la izquierda, y fondo de top a la derecha).

Variable (fondo W)	InvMass	ΔR	$\Delta\eta$	$\Delta\phi$	ΣCSV	Δp_T
Peso ($\times 10^{-1}$)	4.061	2.098	1.975	0.689	0.241	0.227
Variable (fondo top)	ΣCSV	InvMass	$\Delta\phi$	ΔR	$\Delta\eta$	Δp_T
Peso ($\times 10^{-1}$)	9.210	0.437	0.404	0.402	0.373	0.306

Cuadro 5.10: Relevancia de mayor (izquierda) a menor (derecha) de las variables de entrada (pares de jets) tras el entrenamiento (con todas las categorías) de la redes DNNs.

de la suma de ambas respuestas para elegir la mejor permutación entre las 15 posibles (exactamente igual que antes). Una vez tomadas las mejores permutaciones para los jets adicionales en cada evento, se asigna una eficiencia de la DNN por cada evento con criterio idéntico al método de las BDTs, y se promedia (teniendo en cuenta los pesos de cada evento). Las eficiencias por categoría de sabor de jets adicionales y la total se muestran en el cuadro 5.11, donde además se ha añadido de nuevo las eficiencias si solo tuviéramos en cuenta una de las dos respuestas de las DNNs.

Criterio	$t\bar{t}$ +jets (Tot)	$t\bar{t}LF$	$t\bar{t}c\bar{c}$	$t\bar{t}b$	$t\bar{t}b\bar{b}$
Suma DNNs (%)	60.16	61.24	59.83	51.86	43.26
DNN W (%)	37.95	37.32	37.96	43.35	47.02
DNN top (%)	57.44	59.78	54.27	43.62	26.42

Cuadro 5.11: Eficiencias por categoría de sabor de jets adicionales en porcentaje de las DNNs (separando fondos de W y de top) identificando los jets adicionales de cada evento con exactamente 6 jets, siendo al menos dos de ellos b -jets con criterio estricto.

Como ocurría antes, las eficiencias de la categoría $t\bar{t}LF$ son mejores que en el caso de los eventos $t\bar{t}b\bar{b}$, al haber entrenado las redes con mucha más estadísticas de la primera respecto a la segunda. Otro aspecto importante, que también ocurría con el método anterior es que la red de fondo de top es más relevante en eventos con jets adicionales livianos o c -jets, mientras que la red de fondo de W toma más importancia en eventos con b -jets adicionales.

Capítulo 6

Conclusiones

La identificación de jets adicionales en eventos $t\bar{t}$ con una alta eficiencia permite realizar medidas precisas sobre la sección eficaz inclusiva y diferencial de estos eventos. Es por lo que el principal objetivo de este trabajo fue la utilización de tres métodos diferentes que realizasen esa tarea. El primero de ellos más analítico, utilizando toda la información física de los jets de cada evento para reconstruirlo a partir de la mejor combinación de jets que da las masas del sistema $t\bar{t}$. Los dos siguientes, utilizando técnicas de “*machine learning*”, tanto BDTs como a nivel profundo con DNNs.

El trabajo se realizó con simulaciones MC y datos experimentales (para verificar las características de la simulación) recogidos por CMS en el año 2017 a energías en el centro de masas de $\sqrt{s} = 13$ TeV.

En primer lugar, encontramos las eficiencias (última fila del cuadro 5.2) del emparejamiento entre jets del RecoJets (simulación teniendo en cuenta CMS) y los quarks de los que provendrían a nivel GenParticles (simulación a efectos físicos únicamente), teniendo en cuenta el emparejamiento intermedio con GenJets (hadronización de los quarks en jets). Dadas las altas eficiencias del emparejamiento, trabajamos únicamente con eventos con los 4 jets emparejados correctamente, para no afectar al cálculo de las eficiencias de cada uno de los métodos estudiados¹.

Tras obtener los orígenes reales de los jets en la simulación MC, con los que comparar cada respuesta de los métodos evento a evento, los métodos estudiados fueron:

KinFitter: se basaba en identificar y clasificar a los 4 jets del evento $t\bar{t}$ en: 2 jets livianos de uno de los bosones W , y 2 b -jets de cada top que decae en $b + W$. Así, para eventos con 6 jets, si el KinFitter identifica los cuatro jets con más probabilidad de ser del decaimiento de $t\bar{t}$, los dos restantes son considerados como jets adicionales. Sus eficiencias se presentan en el cuadro 5.4.

¹Si tuviéramos en cuenta por ejemplo un evento con 2 jets emparejados correctamente, los otros dos que se emparejaron en el evento $t\bar{t}$ serían los adicionales, para eventos con 6 jets. Así, al analizar ese evento con el método de reconstrucción cinemática, en caso de que este método hubiera detectado los 2 jets adicionales correctamente, al ser comparados con el emparejamiento (“*matching*”), daría una eficiencia de dicho evento del 0% para el KinFitter, afectando entonces haber tomado eventos con una eficiencia en el emparejamiento por debajo del 100%.

BDT: se entrenaron 2 redes con el fin de identificar los jets adicionales de aquellos proveniente del decaimiento del bosón W y del quark top, y una vez obtenidos los pesos se evaluaron los eventos estadísticamente independientes a los de entrenamiento. Al usar redes binarias, la respuesta de las redes nos indicaron cuan señal o fondo era cada pareja de jets de un evento dado. Sus eficiencias se encuentran en el cuadro 5.8.

DNN: se entrenaron 2 redes siguiendo el mismo procedimiento del BDT, aunque esta vez se hizo utilizando las librerías de Keras para Python. Tras obtener los pesos, hubo que migrar esos archivos de pesos de las redes para poder trabajar en nuestro entorno desarrollado en C++. Salvo este cambio de lenguaje, la metodología fue similar a la seguida para la BDT. Sus eficiencias se muestran en el cuadro 5.11.

Tanto para BDT como para DNN se identificaban directamente a los jets adicionales, mientras que el KinFitter resolvía los jets del evento $t\bar{t}$ y descartando estos, seleccionamos los adicionales. Así, el tiempo de computación de los métodos de aprendizaje automático fue mucho menor (al haber menos combinaciones que probar).

Se resumen en el cuadro 6.1 los principales resultados sobre las eficiencias (desglosadas por sabor de los jets adicionales) de cada método (KinFitter, cuadro 5.4; BDTs, cuadro 5.8; y DNNs, cuadro 5.11) identificando a los jets adicionales.

Criterio	$t\bar{t}$ +jets (Tot)	$t\bar{t}LF$	$t\bar{t}c\bar{c}$	$t\bar{t}b$	$t\bar{t}b\bar{b}$
KinFitter (%)	81.63	82.68	79.84	75.60	70.35
Suma BDTs (%)	60.24	61.44	59.49	51.70	42.42
Suma DNNs (%)	60.16	61.24	59.83	51.86	43.26

Cuadro 6.1: Eficiencias por categoría de sabor de jets adicionales en porcentaje de los tres métodos identificando los jets adicionales de cada evento con exactamente 6 jets, siendo al menos dos de ellos b -jets con criterio estricto.

Las diferencias entre eficiencias al considerar los dos últimos métodos (los de aprendizaje automático) son mínimas, pero si son bastante menores (entre un 20 % y un 30 %) respecto al método de reconstrucción cinemática (KinFitter). La principal razón es que el KinFitter utiliza toda la información física del sistema, los límites de las masas invariantes al agrupar dos partículas, y los criterios de b -tagging para los jets del decaimiento de los top en $W + b$, mientras que los métodos de aprendizaje automático se basan en el entrenamiento de una red a partir de una fracción de eventos de la simulación MC de la que disponemos para predecir los resultados en el set restante de eventos.

Tras la realización de este trabajo ha resultado como método más preciso el más analítico (KinFitter), aunque ello ha conllevado un mayor tiempo de ejecución (mayor número de combinaciones posibles por probar) que los métodos de aprendizaje automática, para los que únicamente hubo que emplear cierto tiempo en el entrenamiento de las redes. Una vez obtenido los pesos de las redes, las respuestas de estos métodos a los eventos era inmediata. Esto no significa que los métodos de redes neuronales sean menos eficientes que los analíticos, sino que habría que mejorar el modelo construido, atendiendo a los siguientes parámetros:

- Variables de entrada. Quizás las variables que hemos usado en las redes no han sido

las más eficientes, a pesar de haber definido estas a partir de pares de jets para mejorar los resultados que se obtenían con las magnitudes físicas medidas.

- Mayor estadística de las categorías con b -jets adicionales. Esto ha sido evidente desde que se han obtenido mejores eficiencias para los eventos de categorías con jets adicionales livianos, por tener un mayor conjunto de datos.
- Identificación de la categoría a partir de las variables de los jets. Podría estudiarse la caracterización de las diferentes categorías con las magnitudes de sus partículas para separar los eventos por categorías antes de introducirlos en la red neuronal. Se podría añadir entonces una variable entera que representase la categoría a la hora del entrenamiento, lo que implicaría prácticamente que la red usase diferentes pesos para cada categoría.

Bibliografía

- [1] Donald H. Perkins. *“Introduction to High Energy Physics”*. Cambridge University Press, 2000.
- [2] C. Patrignani *et al.* (Particle Data Group). *“The top quark”*. Chin. Phys. C, **40**, 100001, 2016.
- [3] David Griffiths. *“Introduction to Elementary Particles”*. Wiley-VCH, 2008.
- [4] Michele Maggiore. *“A Modern Introduction to Quantum Field Theory”*. Oxford University Press, 2005.
- [5] P.A. Zyla *et al.* (Particle Data Group). *“Quantum Chromodynamics”*. Prog. Theor. Exp. Phys. **2020**, 083C01, 2021.
- [6] M.L. Perl *et al.* *“Evidence for Anomalous Lepton Production in $e^+ - e^-$ Annihilation”*. Phys. Rev. Lett. **35**, 1489, 1975.
- [7] Quang Ho-Kim and Xuan-Yem Pham. *“Elementary Particles and Their Interactions”*. Springer, 1998.
- [8] F. Abe *et al.* (CDF Collaboration). *“Observation of Top Quark Production in $\bar{p}p$ Collisions with the Collider Detector at Fermilab”*. Phys. Rev. Lett. **74**, 2626, 1995.
- [9] S. Abachi *et al.* (D0 Collaboration). *“Observation of the Top Quark”*. Phys. Rev. Lett. **74**, 2632, 1995.
- [10] R.L. Workman *et al.* (Particle Data Group). *“CKM Quark-Mixing Matrix”*. 2022.
- [11] V.M. Abazov *et al.* (D0 Collaboration). *“Evidence for Production of Single Top Quarks and First Direct Measurement of $|V_{tb}|$ ”*. Phys. Rev. Lett. **98**, 181802, 2007.
- [12] P.A. Zyla *et al.* (Particle Data Group). *“List Top Quark”*. Prog. Theor. Exp. Phys. **2020**, 083C01, 2022.
- [13] R.L. Workman *et al.* (Particle Data Group). *“List W Boson”*. 2022.
- [14] A.M. Sirunyan *et al.* (CMS Collaboration). *“Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV”*. Journal of Instrumentation, Vol. 13, 2018.
- [15] A.M. Sirunyan *et al.* (CMS Collaboration). *“Measurement of the cross section for $t\bar{t}$ production with additional jets and b jets in pp collisions at $\sqrt{s} = 13$ TeV”*. J. High Energ. Phys. **2020**, 125, 2020.

- [16] G. Aad *et al.* (ATLAS Collaboration). “*Study of heavy-flavor quarks produced in association with top-quarks pairs at $\sqrt{s} = 7$ TeV using the ATLAS detector*”. Phys. Rev. D **89**, 072012, 2014.
- [17] Matteo Cacciari *et al.* “*The catchment area of jets*”. J. High Energ. Phys. P04(2008)005, 2008.
- [18] N. Bartosik and J. Hauk. “*Hadron based origin identification of heavy flavour jets at generator level*”. TWiki. CMSPublic, 2017.
- [19] O.S. Brüning *et al.* “*LHC Desing Report*”. Geneva: CERN, 2004.-548 p., 2004.
- [20] Página web CMS.
- [21] Página web ALICE.
- [22] Página web ATLAS.
- [23] G. Aad *et al.* (ATLAS Collaboration). “*Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*”. Physics Letters B Vol. 716, Issue 1, 2012.
- [24] S. Chatrchyan *et al.* (CMS Collaboration). “*Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*”. Physics Letters B Vol. 716, Issue 1, 2013.
- [25] Página web LHCb.
- [26] CMS Collaboration. “*CMS Physics: Technical Desing Report Volume I: Detector Performance and Software*”. Geneva: CERN, 2006. - 521 p., 2006.
- [27] CMS Collaboration. “*CMS Physics Technical Desing Report, Volume II: Physics Performance*”. Journal of Physics G: Nuclear and Particle Physics, Vol. 34, Num. 6, 2007.
- [28] CMS Collaboration. “*The CMS tracker: addendum to the Technical Desing Report*”. Geneva: CERN, 2000. - 82 p., 2000.
- [29] CMS Collaboration. “*The CMS electromagnetic calorimeter project: Technical Desing Report*”. Geneva: CERN, 1997. - 364 p., 1997.
- [30] CMS Collaboration. “*The CMS hadron calorimeter project: Technical Desing Report*”. Geneva: CERN, 1997. - 451 p., 1997.
- [31] CMS Collaboration. “*The CMS muon project: Technical Desing Report*”. Geneva: CERN, 1997. - 441 p., 1997.
- [32] CMS Collaboration. “*Performance of the DeepJet b tagging algorithm using 41.9/fb of data from proton-proton collisionns at 13 TeV with Phase 1 CMS detector*”. CMS performance Note, 2018.
- [33] CMS Collaboration. “*Cut Based Electron ID*”. TWiki, 2017.
- [34] CMS Collaboration. “*Cut Based Muon ID*”. TWiki, 2017.
- [35] CMS Collaboration. “*Jet Selection criteria CMS*”. TWiki, 2017.

- [36] P.A. Zyla *et al.* (Particle Data Group). “*Monte Carlo Particle Numbering Scheme*”. Prog. Theor. Exp. Phys. **2020**, 083C01, 2021.
- [37] Javier Brochero *et al.* “*Measurement of the cross section ratio $\sigma_{t\bar{t}b\bar{b}}/\sigma_{t\bar{t}jj}$ in $l+$ Jets final state at $\sqrt{s} = 13$ TeV*”. CMS Draft Analysis Note, CMS AN-16-339, 2019.
- [38] A. Hoecker *et al.* “*TMVA 4. Toolkit for Multivariate Data Analysis with ROOT. Users Guide*”. CERN-OPEN-2007-007, 2018.
- [39] Página web Keras.
- [40] Yann LeCun *et al.* “*Deep learning*”. Nature, vol. 521, p. 436-444, 2015.

Apéndice A

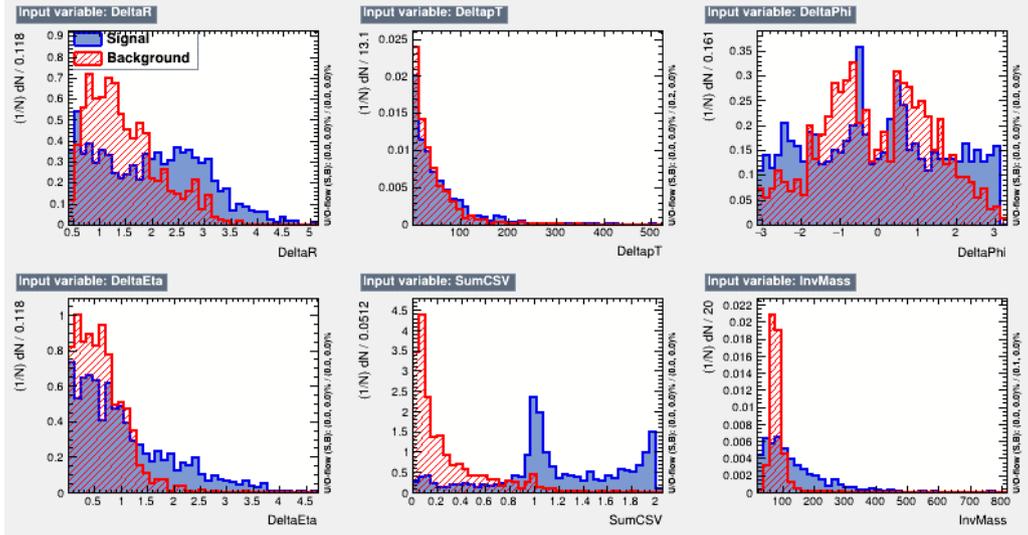
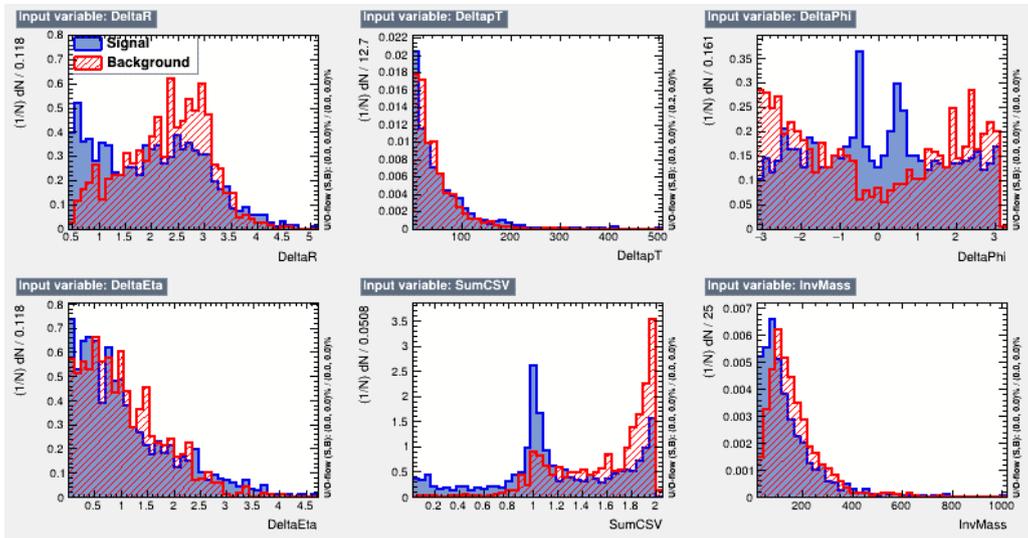
Entrenamiento de las BDTs con datos de categoría $t\bar{t}b\bar{b}$

En este apéndice se describe el intento de mejora de eficiencias identificando jets adicionales con las BDTs, tomando únicamente eventos $t\bar{t}b\bar{b}$ para entrenar las redes.

Las variables de entrada fueron las que definimos a partir de pares de jets en la sección 5.5 (ΔR , Δp_T , $\Delta\eta$, $\Delta\phi$, la suma de los CSV del par de jets y la masa invariante del sistema de jets), aunque la distribución de las mismas no tiene nada que ver con las mostradas en la figura 5.13 cuando consideramos todas las categorías para el entrenamiento. En la figura A.1 se muestran las distribuciones para las dos redes (fondo de W y de top) de las variables de entrada tomando solo los eventos $t\bar{t}b\bar{b}$. Para ambos tipos de fondos (de W y de top), la principal diferencia que se observa en las distribuciones de las variables de entrada es en la suma del CSV de los jets, especialmente en la señal (jets adicionales), para la que ahora se obtiene unos valores por encima de la unidad en la mayoría de los eventos (aunque debería ser alrededor de 2 para todos ellos, la eficiencia del b-tagging tampoco es del 100 %).

En la figura A.2 se representan las correlaciones entre las variables de entrada, tanto para la red de fondo de W como para la de fondo de top. En esta ocasión, para las variables de señal, la correlación más aguda está presente entre ΔR , $\Delta\eta$ y la masa invariante (también presente en el fondo para la red de fondo top), aunque también hay algunas menos intensas entre $\Delta\eta$ y la suma del CSV. En el caso de las variables de fondo para la red de fondo top, además existe una pequeña correlación entre Δp_T y la masa invariante. Por otro lado, para el fondo de la red de fondo W , la correlación más notable es entre ΔR , Δp_T y la masa invariante.

Por último, las distribuciones de los set de datos de entrenamiento y de verificación en señal y fondo para las dos redes entrenadas con eventos de $t\bar{t}b\bar{b}$ aparecen en la figura A.3. En esta ocasión, la mejor de las dos redes en separar señal de fondo es la de fondo W . Con un corte en la variable discriminante un poco anterior al valor 0, se consiguen eficiencias de señal que superan el 90 %, mientras que la del fondo cae hasta valores cercanos al 10 %. El corte más eficiente en la red de fondo top también estaría en torno a 0, pero en esta ocasión la eficiencia de la señal sería del 70 % mientras que la del fondo no sería menor del 30 %. Esto es debido principalmente a que en eventos $t\bar{t}b\bar{b}$, los jets adicionales y los b -jets a los que caen los quarks top en decaimientos $W + b$, tienen la misma naturaleza (CSV), y es

(a) Señal de adicionales y fondo de W .

(b) Señal de adicionales y fondo de top.

Figura A.1: Distribución de las variables de entrada (pares de jets) para los entrenamientos (con eventos $t\bar{t}b\bar{b}$) de las redes BDT separando señal y fondo. Para ambos, de izquierda a derecha y de arriba a abajo: ΔR , Δp_T , $\Delta\phi$, $\Delta\eta$, suma CSV y masa invariante.

más difícil diferenciarlos entre sí. Sin embargo, entre la señal y fondo con la red de fondo W ocurre lo contrario, los jets a los que decaen el bosón W que decae hadrónicamente suelen ser livianos, distinguiéndose así de los b -jets adicionales.

Esta mejora de entrenar la red únicamente con eventos $t\bar{t}b\bar{b}$ confiere obviamente un incremento en la eficiencia de detección de jets adicionales en esta categoría, pero empeora las de las demás categorías, como vemos en el cuadro A.1. Las eficiencias de la categoría con la que se entrenó logra un aumento de 5 puntos, aunque para el resto de categorías el descenso fue desde los casi 20 puntos en el caso de la $t\bar{t}LF$ y $t\bar{t}c\bar{c}$, hasta los 6 en el caso de $t\bar{t}b$. Por esta razón no tomamos estos resultados como el mejor que podría proporcionarnos las BDTs.

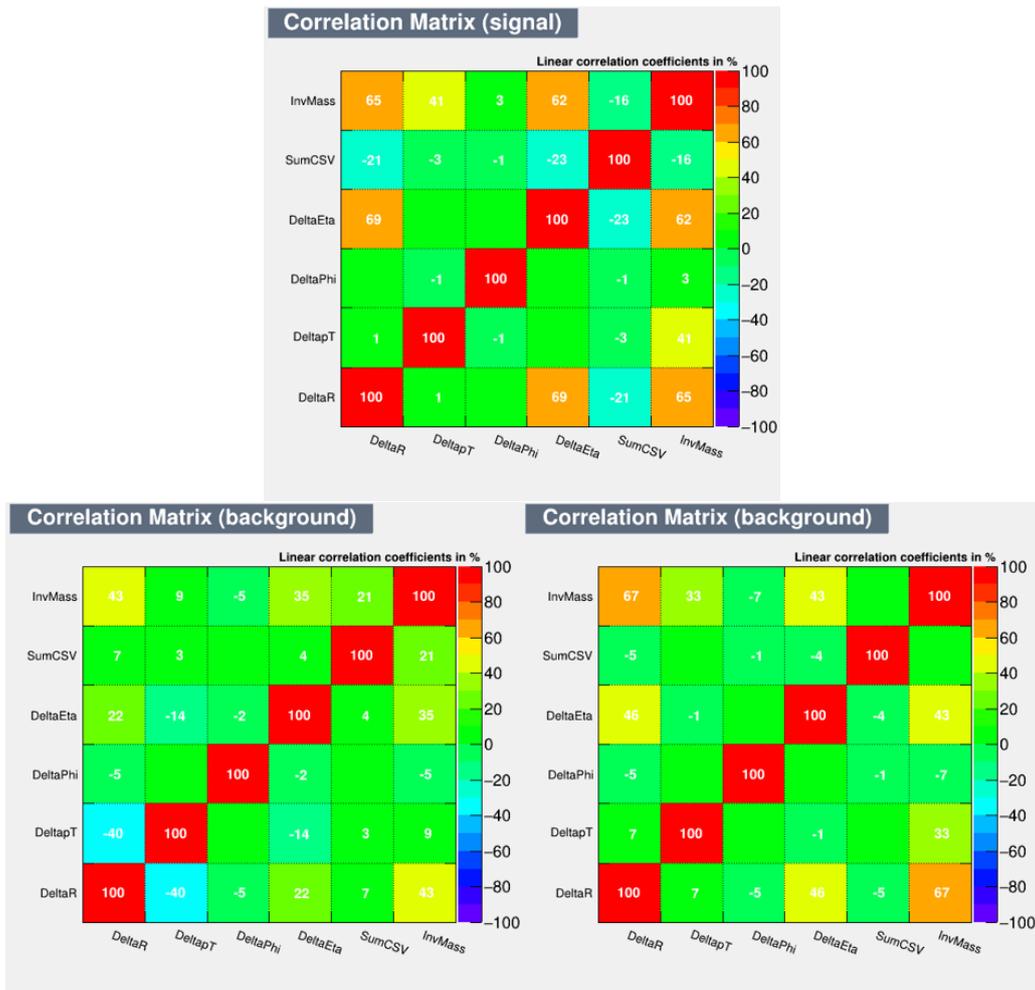


Figura A.2: Matrices de correlación entre las variables de entrada (pares de jets) tanto para señal (arriba) como para fondo (abajo izquierda: fondo de W ; abajo derecha: fondo de top) para el entrenamiento (con eventos $t\bar{t}b\bar{b}$) de la red BDT.

Criterio	Total	$t\bar{t}LF$	$t\bar{t}c\bar{c}$	$t\bar{t}b$	$t\bar{t}b\bar{b}$
BDT ^W con $t\bar{t}b\bar{b}$ (%)	18.68	16.64	20.79	33.23	42.99
BDT ^{top} con $t\bar{t}b\bar{b}$ (%)	50.02	51.66	49.75	36.56	25.52
Suma BDTs con $t\bar{t}b\bar{b}$ (%)	41.69	41.27	41.76	45.04	48.16

Cuadro A.1: Eficiencias por categoría de sabor de jets adicionales en porcentaje de las BDTs entrenadas con $t\bar{t}b\bar{b}$ (separando fondos de W y de top) identificando los jets adicionales de cada evento con exactamente 6 jets, siendo al menos dos de ellos b -jets con criterio estricto.

Un factor importante a tener en cuenta es que para los datos medidos experimentalmente por CMS, no tenemos esa variable entera que nos indicaba a que categoría de las cuatro expuestas en la sección 3.4 pertenece el evento. Esta variable está únicamente presente en las simulaciones MC. Por tanto, la información de la categoría de cada evento no es algo que se utilice para el entrenamiento de la red, simplemente es una forma de clasificar eventos y sus resultados a las respuestas de las redes atendiendo al sabor de sus jets adicionales.

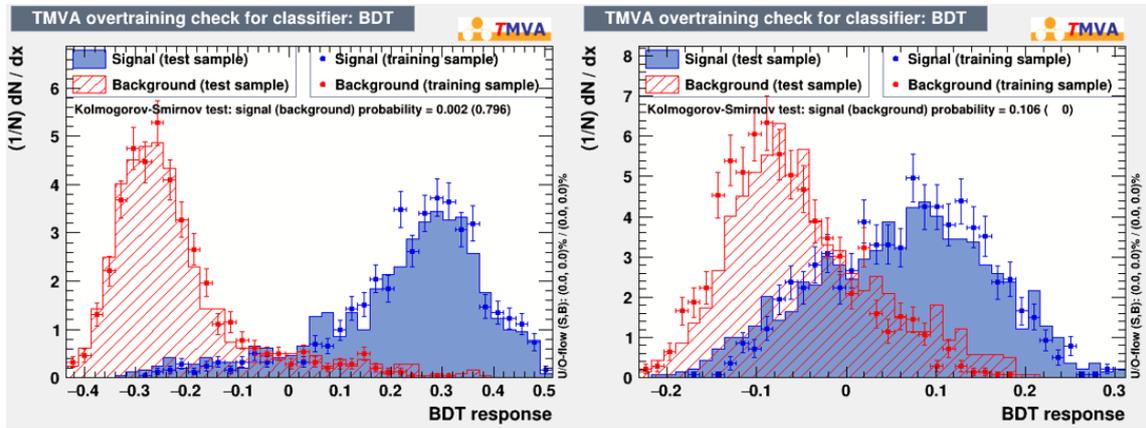


Figura A.3: Distribución de las respuestas de las BDTs (pares de jets) entrenadas (con eventos $t\bar{t}b\bar{b}$) para los sets de datos de entrenamiento (valores puntuales) y datos de verificación (histogramas) para señal y fondo (fondo de W a la izquierda, y fondo de top a la derecha).

Apéndice B

Redes neuronales profundas

El aprendizaje profundo (*“deep learning”*) permite a los modelos computacionales compuestos por múltiples capas aprender la representación de un conjunto de datos con un alto grado de abstracción [40]. El término “profundo” no se refiere a un aprendizaje con mayor grado de detalle de los datos proporcionados, sino que hace referencia al número de capas presentes en la red, lo que se denomina profundidad de la red.

Existen dos tipos de aprendizajes automáticos según la forma de entrenar a las redes:

- Aprendizaje supervisado: se parte de un conjunto completo de datos y sus respuestas con las que la red se puede entrenar. Se pueden resolver tanto problemas de clasificación (es el caso de este trabajo, discernir si el par de jets es señal o fondo) como problemas de regresión.
- Aprendizaje no supervisado: en algunos casos únicamente tenemos acceso a los datos, pero no a las respuestas. Es el propio algoritmo el que ha de encontrar las estructuras que caracterizan a los datos de entrada.

La forma más común de aprendizaje automático es el aprendizaje supervisado. La intención es que cada categoría en la clasificación sea detectada correctamente mediante la red. Para ello se utiliza una función que mide el error entre el resultado de la red y los resultados que deberían dar (proporcionados en el entrenamiento junto con los datos de entrada), para después modificar los parámetros internos (pesos de la red) en favor de reducir el error. Además se utiliza lo que se denomina vector gradiente que indica la cantidad que aumenta o disminuye el error si el peso aumenta una pequeña cantidad.

En la práctica se utiliza un procedimiento llamado gradiente de descenso estocástico (SGD), que consiste en calcular las salidas y errores para un subconjunto de los datos de entradas, para hallar el gradiente promedio y ajustar los pesos. Esto se repite para todos los subconjuntos de datos de entrenamiento hasta que deja de descender la función objetivo [40]. El número por el que se divide el set de datos de entrenamiento para generar los diferentes subconjuntos de datos con los que se entrena en cada iteración es el número de lotes (*“batch number”*), que para nuestra red fue 512 (es decir, los eventos se tomaban en lotes de 512 para entrenar). Normalmente este número de lotes suele ser una potencia de 2 (en nuestro caso $2^9 = 512$). En resumen, el procedimiento tras entrenar los primeros 512

datos y actualizar los pesos de la red, fue seguir con los siguientes 512 datos, y de forma sucesiva seguir actualizando los pesos hasta acabar con el set de datos de entrenamiento.

Otro parámetro de la red importante es el número de épocas, que viene a indicar el número de veces que se repite todo el entrenamiento de la red con los mismos datos, pero partiendo de la red con los pesos encontrados en la época anterior, es decir, una vez hemos acabado con los últimos 512 eventos del set de datos de entrenamiento (primera época), se vuelven a entrenar los 512 primeros datos con los pesos de la red encontrados en la última iteración (los más actualizados). Para nuestro caso, el número de épocas fue de 500, aunque utilizamos una función de parada temprana (*“early stopping”*) que permite salir al algoritmo de las iteraciones en el momento en el que la eficiencia del mismo comience a decaer.

Un último aspecto a tener en cuenta en las redes neuronales son las funciones de activación, que definen la salida de un nodo dado a partir de una entrada o conjunto de entradas. Estas funciones han de ser no lineales, lo que permiten a la red trabajar con problemas no triviales usando un número bajo de nodos. Las funciones de activación lineal serían equivalentes a trabajar con una única neurona (o nodo) en la red. Existen diferentes tipos de funciones de activación, y cada una tiene sus características que le permiten trabajar en diferentes situaciones, es decir, una misma red neuronal puede tener varias funciones de activación en diferentes capas, dependiendo si son capas externas (de entrada o salida), o capas internas. Estas son algunas de las principales funciones de activación, y donde se suelen usar más:

- ReLU: para el valor de entrada x , devuelve el máximo entre 0 y x , es decir, $f(x) = \max\{0, x\}$, por lo que siempre la respuesta de estas capas es un valor positivo. Una variante de esta es la activación *“leaky”* ReLU, que en vez de asignar 0 a valores negativos, toma $0,01x$ para esos valores. Con estas capas se evita el problema de gradiente de fuga, al ser la derivada de la función 1 en todo valor positivo de x , por lo que se suelen utilizar para capas interiores y en problemas de regresión para la última capa (aunque en estos tipos de problemas pueden no tener función de activación en la última capa).
- Tangente hiperbólica y sigmoide: la función tangente hiperbólica devuelve $g(x) = \tanh(x)$, mientras que la sigmoide es una función como $\sigma(x) = \frac{1}{1+\exp^{-x}}$. La forma de ambas es muy parecida: en valores negativos se acercan asintóticamente a -1 y 0 respectivamente, mientras que en valores positivo lo hacen ambas a 1. La principal diferencia es su derivada (velocidad de crecimiento), que para la primera es $g'(x) = 1 - (\tanh(x))^2$ y para la segunda es $\sigma'(x) = \sigma(x)(1 - \sigma(x))$. Sólo se utilizan en la última capa para redes de clasificación binarias (como es nuestro caso).
- SoftMax: es una generalización de la función logística definida como la función sigmoide, que permite comprimir un vector n -dimensional, \mathbf{z} , de valores reales arbitrarios en un vector n -dimensional, $\sigma(\mathbf{z})$, de valores reales en el rango $[0, 1]$. Es por ello que se usa en la última capa para redes de múltiple clasificación (más de dos categorías finales).

Para construir el modelo que hemos utilizado en las redes neuronales profundas usamos la estructura descrita en la sección 5.6: una capa inicial con 6 nodos (igual que el número de variables de entradas), 6 capas internas de número de nodos $70 \times 50 \times 30 \times 30 \times 20 \times 10$ y

la capa final con un único nodo. La activación de todas las capas fue una función ReLU, excepto la de salida, con una función de activación sigmoide, lo que nos proporciona valores entre 0 (si es fondo) y 1 (si es señal).