



***Facultad
de
Ciencias***

CARACTERIZACIÓN Y DISCRIMINACIÓN DE FONDO DRELL-YAN EN EL ANÁLISIS $H \rightarrow WW$

(Characterization and discrimination of Drell-Yan background in the $H \rightarrow WW$ analysis)

Trabajo de Fin de Grado
para acceder al

GRADO EN FÍSICA

Autora: Clara Lasasa García

Directora: Alicia Calderón Tazón

Co-Director: Pedro J. Fernández Manteca

Septiembre - 2019

AGRADECIMIENTOS

Gracias a la Universidad de Cantabria por los medios, en concreto a mis profesores por haberme formado como física.

Gracias al grupo de altas energías del IFCA por haberme acogido con los brazos abiertos, por haberme hecho sentir una más.

Gracias Alicia por contar conmigo y confiar en mí desde el primer momento. Gracias por las oportunidades que me has brindado. Ha sido un placer trabajar contigo.

Gracias Pedro por todo el tiempo dedicado, por todo lo que me has enseñado y por la paciencia. Me siento muy afortunada por haberte tenido como codirector.

Gracias a la gente que me ha regalado Santander y en concreto la Física. Gracias Celia, Javi, Lucía, Jaime, Irene, Gallo, Carmen, Álvaro B., Fran, Álvaro M., Pablo, sin vosotros no habría sido lo mismo, gracias por hacerlo más llevadero.

Gracias a mi familia por el apoyo incondicional. En especial a ti, mamá, por estar siempre para absolutamente todo a más de 500 km de distancia.

RESUMEN

El objetivo de este trabajo ha sido caracterizar y discriminar dos procesos físicos producidos en el LHC: la producción de un bosón de Higgs cuando se desintegra a dos bosones W, y el Drell-Yan. Para ello, se ha realizado un análisis secuencial y un análisis multivariable basados en muestras de simulación Monte Carlo que reproducen las condiciones del experimento CMS durante la toma de datos de 2017 a una energía de centro de masas de $\sqrt{s} = 13$ TeV y a una luminosidad integrada de 41.5 fb^{-1} . Se ha obtenido que el análisis multivariable, el cual combina el poder discriminatorio de variables de entrada en un único discriminante, logra una separación mejor que el análisis secuencial, en el que se aplican selecciones simples en las distribuciones de dichas variables de entrada.

Palabras clave: CMS, LHC, discriminación, caracterización, análisis multivariable, análisis secuencial, Drell-Yan, bosón de Higgs.

ABSTRACT

The objective of this project has been to characterize and discriminate between two physical processes produced in the LHC: the production of a Higgs boson when it decays into two W bosons, and Drell-Yan. For this purpose, a sequential analysis and a multivariate analysis have been performed based on Monte Carlo simulation samples that reproduce the status of the CMS experiment during the 2017 data collection at a center-of-mass energy of $\sqrt{s} = 13$ TeV and an integrated luminosity of 41.5 fb^{-1} . It has been determined that the multivariate analysis, which combines the discriminatory power of the input variables into a single discriminant, achieves a better separation than the sequential analysis, in which some simple selections are applied to the distributions of such input variables.

Key words: CMS, LHC, discrimination, characterization, multivariate analysis, sequential analysis, Drell-Yan, Higgs boson.

ÍNDICE

Agradecimientos	I
Resumen	II
1. Motivación	1
2. Introducción teórica	2
2.1. Modelo Estándar	2
2.2. Necesidad y descubrimiento del bosón de Higgs	2
2.3. Producción del bosón de Higgs en el LHC	3
2.4. Desintegración del bosón de Higgs	4
3. Dispositivo experimental	5
3.1. El acelerador LHC	5
3.2. El detector CMS	6
4. Reconstrucción e identificación de objetos físicos	9
4.1. Objetos	9
5. Procesos de señal y fondo	12
5.1. Procesos de señal	12
5.2. Procesos de fondo	12
5.3. Muestras de Monte Carlo	14
6. Métodos de análisis multivariable	16
6.1. Análisis multivariable	16
6.1.1. Fases y flujos de trabajo	16
6.2. Árboles de decisión	17
6.3. Algoritmos de construcción de BDTs	19
6.3.1. Boosting - AdaBoost	19
6.3.2. Bagging	20

6.3.3. Random Forest	20
6.4. Opciones de configuración en BDTs	20
7. Análisis secuencial	21
7.1. Selección	21
7.2. Significancia de señal	28
8. Análisis basado en un BDT	30
8.1. Variables de entrada	30
8.2. Elección del número de jets para el entrenamiento	33
8.3. Rendimiento de los diferentes algoritmos de entrenamiento	36
8.4. Optimización de hiperparámetros del modelo	38
8.5. Evaluación de los resultados	41
8.5.1. Árboles de decisión	43
8.5.2. Matrices de Correlación	44
8.5.3. Ranking de variables	45
8.6. Significancia de señal seleccionando en la respuesta del BDT	46
9. Conclusiones	51
Referencias	52

1. MOTIVACIÓN

El objetivo principal de este trabajo es identificar y diferenciar dos procesos físicos que pueden producirse en el gran colisionador de hadrones LHC (del inglés *Large Hadron Collider*) a partir de colisiones entre dos haces de protones: el denominado Drell-Yan en el que se genera un bosón Z o un fotón, y la producción del bosón de Higgs cuando éste se desintegra en dos bosones W ($H \rightarrow WW$). Se han utilizado para este propósito las simulaciones de Monte Carlo que reproducen las condiciones del detector CMS (del inglés *Compact Muon Solenoid*) en toda la toma de datos del año 2017 a una energía de centro de masas de 13 TeV y una luminosidad de 41.5 fb^{-1} .

Este trabajo se centra en la selección de eventos con dos leptones muy energéticos de carga opuesta y mismo sabor (dos electrones o dos muones), así como energía transversa perdida o MET (del inglés *Missing Transverse Energy*) en el estado final. En el caso de la producción del bosón de Higgs en el canal a dos bosones W, cada uno de ellos se desintegra en un leptón cargado y un neutrino que no interacciona con el detector pero cuya presencia queda registrada en la MET. Por su parte, en el proceso de Drell-Yan, el bosón Z o el fotón virtual se desintegra en dos leptones.

Para alcanzar el objetivo, se han empleado BDTs (del inglés *Boosted Decision Trees*) que es una técnica de análisis multivariable que permite una mejor discriminación de la señal con respecto a firmas similares, llamadas fondos. El entrenamiento de los BDTs se ha realizado en base a un conjunto de variables de entrada generando como resultado un discriminante. Su rendimiento se ha cuantificado por medio de las curvas ROC (del inglés *Receiver Operating Characteristic*), que reflejan el poder discriminatorio de las variables. Posteriormente, se ha aplicado una selección sobre la misma para lograr un aumento de significancia de señal.

Existen análisis multivariable previos con BDTs para ganar significancia de señal $H \rightarrow WW$ frente a fondo Drell-Yan [1], utilizando datos de CMS del año 2016 a una energía de centro de masas de 13 TeV y una luminosidad de 35.9 fb^{-1} . Se pretende implementar un procedimiento similar para datos tomados en 2017, de manera que este trabajo pueda ser de utilidad en publicaciones futuras.

La presente memoria se distribuye en secciones. La Sección 2 introduce brevemente el Modelo Estándar de partículas así como los mecanismos de producción y desintegración del bosón de Higgs. En la Sección 3 se presenta el dispositivo experimental empleado: el acelerador LHC y el detector CMS. La Sección 4 describe la reconstrucción e identificación de objetos en CMS y las muestras de datos utilizadas en el análisis. En la Sección 5 se describen los procesos de señal y fondo, así como las muestras de simulación empleadas. La Sección 6 explica todos los aspectos importantes relativos al análisis multivariable. Finalmente, en las Secciones 7 y 8, se explica respectivamente la estrategia seguida en los análisis secuencial y multivariable; a su vez, se presentan los resultados obtenidos. Todo el código desarrollado en el contexto de este trabajo puede ser consultado en Github [36].

2. INTRODUCCIÓN TEÓRICA

2.1. Modelo Estándar

El Modelo Estándar de partículas o SM (del inglés *Standard Model*) [2] es la interpretación física, probada experimentalmente, que describe la estructura de la materia, así como su dinámica a partir de tres fuerzas fundamentales.

Este modelo, como se muestra en la Figura 1, describe las partículas fundamentales, denominadas fermiones, así como los bosones responsables de las fuerzas fundamentales. Entre los fermiones, con espín semientero, se distinguen leptones y quarks, que se agrupan en dobletes correspondientes a tres generaciones. Por su parte, los cuatro bosones poseen espín entero.

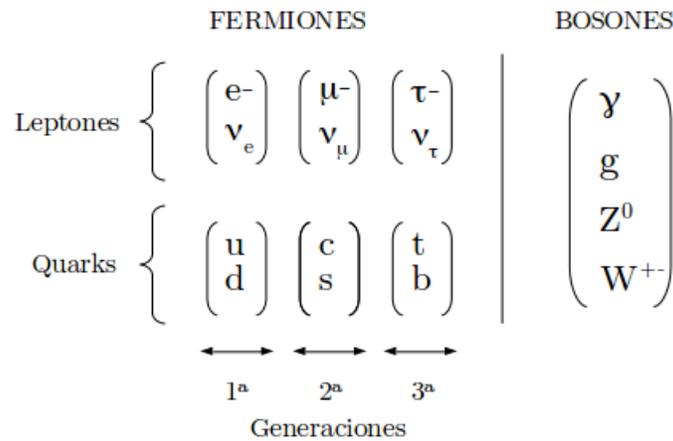


Figura 1: Modelo estándar de la física de partículas.

Las partículas del SM obedecen las leyes de la Mecánica Cuántica y de la Relatividad General. Sus interacciones se basan en tres fuerzas fundamentales: la Fuerza Fuerte, la Fuerza Electromagnética y la Fuerza Débil, unificándose las dos últimas en la Interacción Electro débil. Todo ello es descrito en el marco de la Teoría Cuántica de Campos o QFT (del inglés *Quantum Field Theory*).

2.2. Necesidad y descubrimiento del bosón de Higgs

El Modelo Estándar ha sido exitosamente probado para explicar las interacciones débil, fuerte y electromagnética entre partículas fundamentales, y también ha predicho la existencia de partículas que más tarde han sido experimentalmente descubiertas. No obstante, la formulación de dicho modelo conduce a bosones sin masa, asociados a la interacción débil: W^+ , W^- y Z^0 , lo cual no concuerda con los resultados experimentales, que indican que sí son masivos. Ante esto, es necesario un método para proporcionar masa a dichos bosones.

Han sido propuestos muchos mecanismos para resolver este problema. Sin embargo, únicamente el mecanismo de Higgs, que propone la inclusión de un campo escalar doble complejo (denominado campo de Higgs) con su correspondiente bosón escalar, el bosón de Higgs, ha sido experimentalmente confirmado. De acuerdo a esta teoría, las partículas adquieren su masa interactuando con el campo de Higgs que ocupa todo el espacio, de tal manera que cuanto mayor es la interacción de una partícula con el campo de Higgs, mayor es su masa [3] [4].

El único modo de verificar esta teoría es probar la existencia del campo de Higgs, es decir, encontrar el bosón de Higgs asociado. El mecanismo de Higgs fue teorizado en 1964, aunque las limitaciones tecnológicas provocaron que esta búsqueda se extendiese hasta el 4 de julio de 2012. En dicha fecha, los experimentos ATLAS [5] y CMS [6] del LHC anunciaron el descubrimiento de una nueva partícula en el rango de masas en torno a $125 \text{ GeV}/c^2$, y el 14 de marzo de 2013 el CERN confirmó que sus propiedades y comportamiento eran consistentes con el bosón de Higgs postulado por el SM.

2.3. Producción del bosón de Higgs en el LHC

Hay cuatro modos de producción del bosón de Higgs en el LHC. El mecanismo con mayor sección eficaz en el LHC (Figura 2a) y con una probabilidad del 87 % es el proceso de fusión de gluones, que es esencialmente un proceso de interacción fuerte donde el acoplamiento entre los gluones y el bosón de Higgs tiene lugar a través del intercambio de un quark top. El segundo modo de producción del bosón de Higgs más dominante (Figura 2b) con una probabilidad del 7 % es la fusión de dos bosones vectoriales o VBF (del inglés *Vector Boson Fusion*); procede por el scattering inelástico de dos quarks o antiquarks mediados por el intercambio de un bosón W o Z . El siguiente mecanismo de producción del Higgs más relevante (Figura 2c) con una probabilidad del 5 % es la producción asociada con un bosón de gauge W o Z . El último modo (Figura 2d) con una probabilidad inferior al 1 % es la producción asociada a un par de quarks top-antitop ($t\bar{t}$), que puede proporcionar información importante sobre el acoplamiento de Yukawa top-Higgs [7].

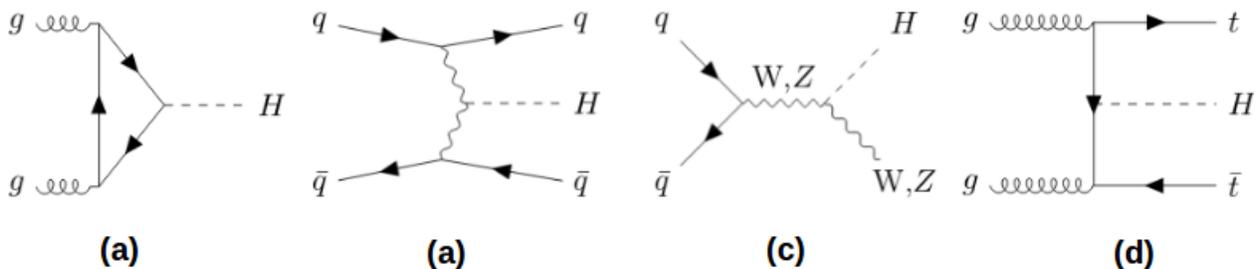


Figura 2: Principales diagramas de Feynman para la producción del bosón de Higgs.

En ese trabajo se han utilizado simulaciones de Monte Carlo que incluyen los cuatro modos de producción.

2.4. Desintegración del bosón de Higgs

El bosón de Higgs es muy inestable y se desintegra rápidamente. Su tiempo de vida media es aproximadamente $\tau = 10^{-22}$ s. En principio, puede desintegrarse a todas las partículas del SM. Sin embargo, debido a la proporcionalidad del acoplamiento a la masa de las partículas involucradas, las mayores tasas de desintegración corresponden a las partículas más masivas. Entre los diversos modos de desintegración que se han medido, la desintegración a dos bosones W (en la que se basa este trabajo) con probabilidad del 21.4 %, a dos bosones Z con probabilidad del 2.6 %, a dos fotones γ con probabilidad del 0.2 %, a un par de leptones tau-antitau ($\tau^+\tau^-$) con una probabilidad del 6.3 %, y a un par de quarks bottom-antibottom ($b\bar{b}$) con una probabilidad del 58.4 %, tienen fuerte consistencia con la predicción del SM. También se predice el canal de desintegración a dos gluones g con una probabilidad del 8.2 %, entre otros. Incluso hay búsquedas de desintegraciones poco comunes del bosón de Higgs como a un bosón Z y un fotón γ con una probabilidad de 0.1 %, o a un par muón-antimuón ($\mu^+\mu^-$) con una probabilidad de 0.02 % [7].

3. DISPOSITIVO EXPERIMENTAL

Para el desarrollo de este trabajo se han utilizado simulaciones que reproducen las condiciones del detector CMS del acelerador LHC. Mediante el LHC se producen colisiones de haces de protones, y a través de CMS se detectan partículas estables tras la colisión. Por tanto, ambos conforman el dispositivo experimental empleado.

3.1. El acelerador LHC

El LHC [8] [9] es el colisionador de hadrones circular más grande y potente del mundo. Está localizado en el CERN (del francés *Conseil Européen pour la Recherche Nucléaire*) en Ginebra, Suiza. Consiste en un anillo de 27 kilómetros de circunferencia construido bajo tierra a una profundidad de entre 50 y 100 metros. Dicho anillo está provisto de imanes superconductores que permiten focalizar los haces y doblar sus trayectorias a través del acelerador, así como de una serie de cavidades de radiofrecuencia que impulsan las partículas aumentando su energía.

El LHC hace colisionar dos haces de protones que viajan a una velocidad cercana a la de la luz, a una energía máxima de 7 TeV proporcionando una energía de centro de masas total $\sqrt{s} = 14$ TeV con una luminosidad instantánea de $10^{34} \text{ cm}^{-2}\text{s}^{-1}$. La luminosidad instantánea es un indicador clave del rendimiento de un acelerador y está definida como la relación entre el número de eventos detectados por unidad de tiempo y la sección eficaz de la interacción, mientras que la luminosidad integrada se corresponde con la integral de la luminosidad con respecto al tiempo.

Los haces de protones se componen de paquetes discretos de 10^{11} protones separados en intervalos de 25 ns, por lo que es frecuente que se produzca más de una colisión *pp*. Como resultado, se reconstruyen trazas que proceden de diversos puntos denominados vértices donde han tenido lugar las interacciones. No obstante, solo una de las millones de colisiones producidas es la principal, es decir, la de interés, y su vértice es conocido como vértice primario. El resto contribuye al incremento de partículas de fondo, conocido como efecto *pileup*. Los haces siguen trayectorias separadas a través de tuberías sometidas a ultra alto vacío, con configuraciones de campo magnético opuestas.

El tipo de partículas, la energía de centro de masas en las colisiones y la luminosidad son las características más importantes de un acelerador de partículas. La energía de centro de masas determina los tipos de partículas que pueden ser estudiadas o descubiertas; cuanto mayor es la energía, más son las posibilidades de producir partículas más masivas y por tanto nueva física no descrita por el SM. Las partículas masivas no pueden observarse directamente, casi inmediatamente se desintegran en partículas más ligeras que a su vez también pueden desintegrarse hasta ser estables para los tiempos de detección; las partículas resultantes de la cadena de desintegraciones son identificadas en los detectores.

Aunque el LHC fue diseñado para alcanzar una energía total de centro de masas $\sqrt{s} = 14$ TeV, la energía por protón alcanzada durante el período de toma de datos usado en este trabajo, conocido como Run II, es de 6.5 TeV. Así pues, la energía de centro de masas relevante para este estudio es $\sqrt{s} = 13$ TeV.

El LHC cuenta con cuatro puntos de colisión en los que se localizan los experimentos ALICE, ATLAS, CMS y LHC-b. En la Figura 3 se muestra un esquema del complejo sistema de aceleradores del CERN y los cuatro experimentos del LHC mencionados anteriormente.

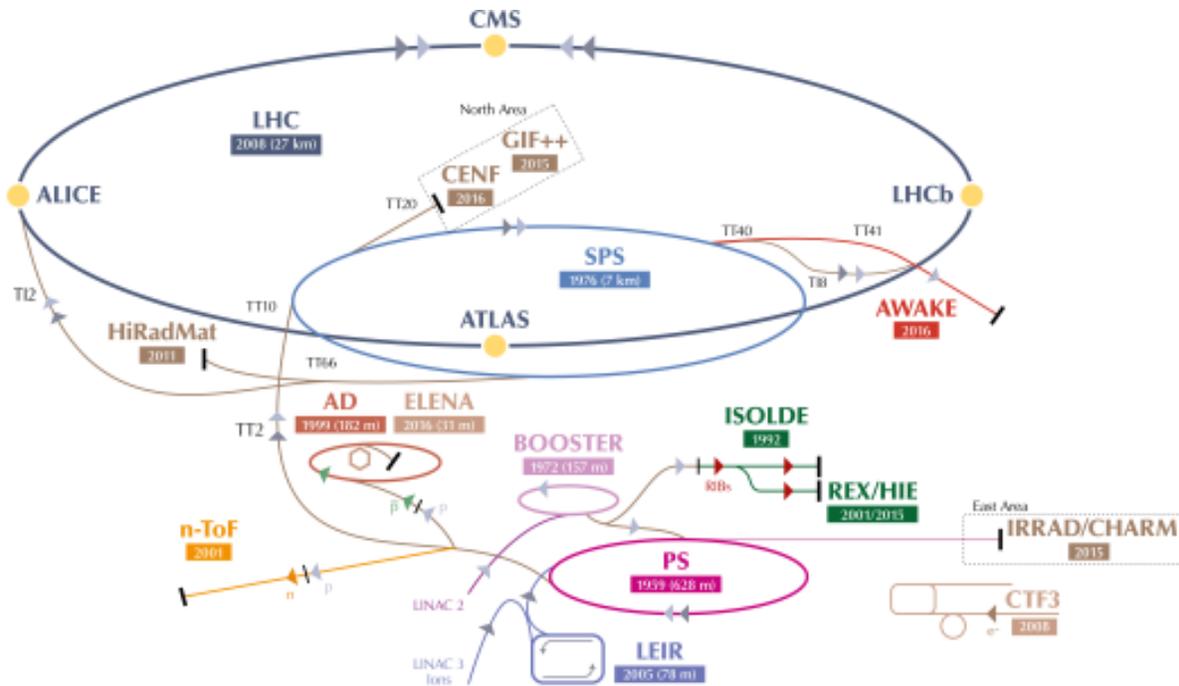


Figura 3: Sistema de aceleradores del CERN. [10]

3.2. El detector CMS

CMS [11] [12] [13] es un detector cilíndrico de propósito general, es decir, diseñado para verificar la base del SM y observar cualquier fenómeno de nueva física que el LHC pueda revelar. Su peso alcanza 14000 toneladas y mide aproximadamente 15 m de diámetro y 21.6 m de largo.

En la Figura 4 se observa la estructura del detector. Está compuesto, de la zona más interna a la más externa, por un *tracker* de píxeles y tiras de silicio para la detección de partículas cargadas, un calorímetro electromagnético de cristal de tungstato (ECAL) para la medida de electrones y fotones principalmente, un calorímetro hadrónico constituido de material denso y absorbente (HCAL) para la medida de hadrones (jets). En la parte más externa se encuentran las cámaras de muones, que son de tres tipos: *Drift Tubes Chambers* (DTC), *Cathode Strip Chambers* (CSC) y *Resistive Plate Chambers* (RPC).

Entre el HCAL y las cámaras de muones hay un imán superconductor que alcanza un campo magnético de 3.8 T que curva las partículas cargadas de muy alto momento y permite una buena resolución en la medida del mismo.

En cuanto a su geometría, el sistema de coordenadas aceptado tiene como origen el punto de colisión, el eje y apunta verticalmente hacia arriba, el eje x radialmente desde el origen, y el eje z sigue la dirección del haz. El ángulo azimutal ϕ se mide a partir del eje x en el plano XY transversal al haz, mientras que el ángulo polar θ se mide desde el eje z en el plano XZ .

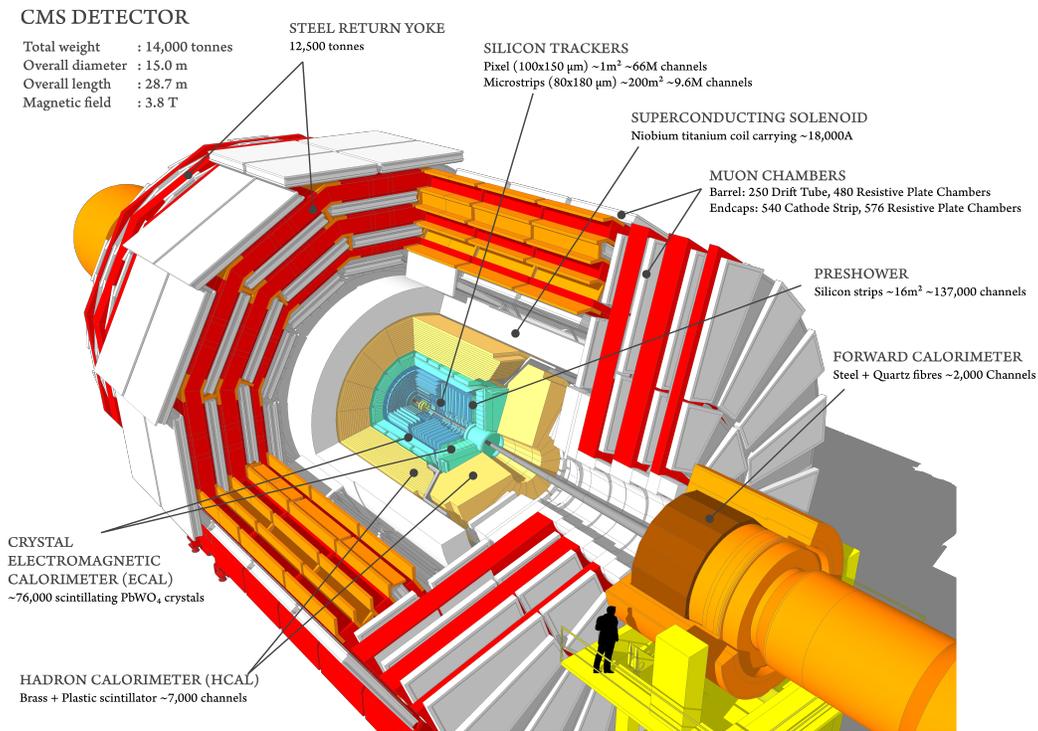


Figura 4: Diseño del detector CMS. [11]

Una variable angular importante, que se utiliza en el análisis por ser invariante bajo transformaciones de Lorentz en el eje z , es la pseudorapidez η , la cual se define en función del ángulo polar como $\eta = -\ln(\tan\frac{\theta}{2})$. Esta variable indica el volumen fiducial de detección del dispositivo experimental, y su distribución espacial se muestra en la Figura 5. Otras variables importantes se definen en el plano transversal a la dirección del haz de partículas, como el momento transverso p_T o la energía transversa E_T , así como la energía transversa perdida E_T^{miss} que será definida posteriormente. Estas componentes transversas son medidas con alta precisión en este tipo de detectores.

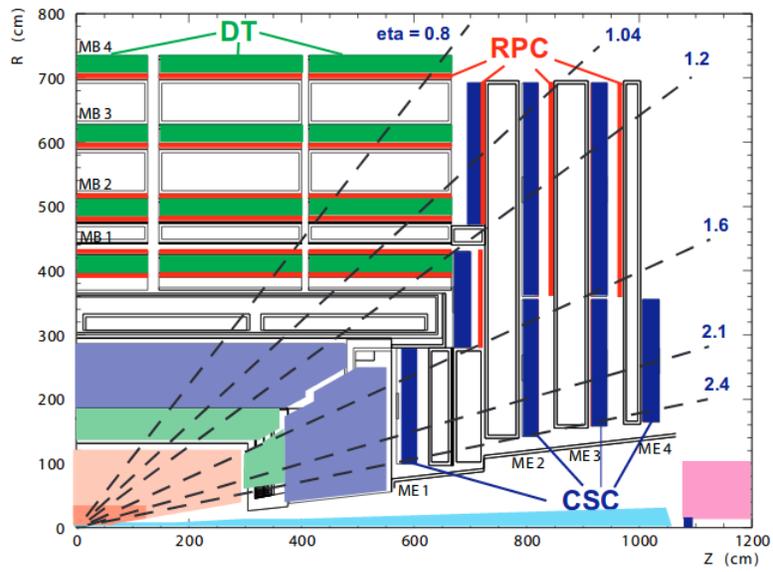


Figura 5: Cuadrante de CMS y distribución de valores de la pseudorapidez η . [12]

4. RECONSTRUCCIÓN E IDENTIFICACIÓN DE OBJETOS FÍSICOS

En esta sección se define el procedimiento de reconstrucción e identificación de los objetos físicos presentes en los eventos de las muestras empleadas en el trabajo.

Las partículas generan señales eléctricas en los subdetectores de CMS. El proceso de reconstrucción de los objetos transforma dichas señales en información física relevante (energía, momento o dirección). Mediante el algoritmo *Particle Flow* [14] se combina toda la información de las partículas producidas en la colisión, consiguiendo así identificar y reconstruir cada partícula individualmente. La reconstrucción se lleva a cabo durante el procesado de los datos y para cada tipo de análisis se seleccionan posteriormente los criterios más adecuados de identificación y aislamiento de objetos.

En la Figura 6 se ilustra el alcance de algunas partículas en las diferentes capas del detector, coincidiendo con las firmas características de cada uno de los objetos físicos que se utilizan en este trabajo.

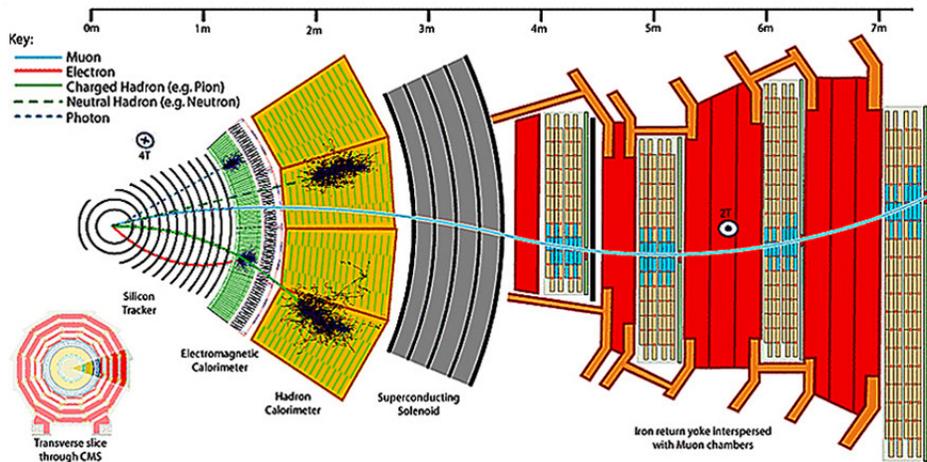


Figura 6: Señal de las partículas en los subdetectores de CMS. [15]

4.1. Objetos

Los objetos que se han utilizado en este trabajo son: leptones, jets, quarks b y MET.

■ Leptones

Los leptones que han sido estudiados en este proyecto son muones y electrones.

- Muones: se reconstruyen combinando la información de los impactos que dejan en las cámaras de muones y de las trazas que dejan en el *tracker* [16], tal como se indica en la Figura 6.

- Electrones: se reconstruyen a partir de la energía que depositan en el ECAL y de las trazas que generan en el *tracker* [17], tal como se indica en la Figura 6.

En todo análisis que contiene muones o electrones en el estado final, es fundamental discernir aquellos (*prompt*) que provienen de desintegraciones producidas directamente en el punto de interacción (de bosones W y Z principalmente), de aquellos otros (*non-prompt*) que proceden de desintegraciones secundarias (de mesones K y Π), o de leptones en jets. Por otra parte, asumiendo que los leptones en jets tienen mayor energía alrededor de su traza debido al resto de partículas que contiene el jet, se puede identificar si un leptón está contenido en un jet. Para ello, se estudian diferentes criterios de identificación (ID) y aislamiento (ISO) con diversos valores de eficiencia. El aislamiento puede definirse como la suma de la energía depositada de todas las partículas reconstruidas en un cono de radio $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ en torno al leptón, dividido entre el p_T del mismo.

En el análisis se han usado muones con criterios muy restrictivos de ID (*Tight ID* [16]), que incluye una serie de criterios de calidad a las trazas tanto en el *tracker* como en las cámaras de muones, alcanzando una eficiencia de ID de aproximadamente el 97 % para muones con $p_T > 20$ GeV.

Se ha impuesto un momento transverso mínimo de 20 GeV y ser reconstruidos en la región $|\eta| < 2.4$ para aprovechar el máximo volumen del detector. A su vez, para rechazar muones procedentes de vértices secundarios, se han impuesto criterios en cuanto a la distancia respecto del vértice principal: para la distancia radial entre el vértice primario y el muón se requiere $d_0 < 0.02$ cm, y para la distancia longitudinal se requiere $d_z < 0.1$ cm. Además, se aplica un aislamiento menor que 0.15 dentro de un cono $\Delta R = 0.4$, con una eficiencia de aproximadamente el 95 %.

Del mismo modo, se han aplicado requisitos estrictos de ID para los electrones (*Tight Electron ID* [17]) y se ha requerido un ISO menor que 0.06 en un cono $\Delta R = 0.3$. A su vez, se ha impuesto un momento transverso mínimo de 20 GeV, ser reconstruido en una región $|\eta| < 2.5$, así como los criterios de distancia $d_0 < 0.05$ cm y $d_z < 0.1$ cm.

■ Jets

Los jets son haces de partículas normalmente provenientes de la hadronización de quarks o gluones tras una colisión. La reconstrucción de estos objetos se lleva a cabo mediante el algoritmo recursivo *anti- k_T* [18], que se basa en las distancias entre los distintos pares de partículas y de cada partícula con la dirección del jet candidato.

En este trabajo se han definido los jets como objetos reconstruidos por el algoritmo *anti- k_T* , con pseudorapidez $|\eta| < 5.2$ y momento transverso $p_T > 30$ GeV, para

asegurarse de que son jets reales y bien medidos, reduciendo así el error respecto a usar jets de menor p_T .

- **Quarks b**

El algoritmo conocido como *b-tagging* [19] aprovecha el relativamente largo tiempo de vida de los hadrones para identificar jets procedentes de la hadronización de quarks b . La trayectoria media de los hadrones B es suficientemente larga para hacer factible la reconstrucción de vértices secundarios desplazados con respecto al vértice primario de la colisión. Esta información particular permite reconstruir vértices secundarios e identificar con una cierta eficiencia quarks b (*b-tag*).

Un aspecto fundamental en este estudio, por ser un fondo importante, es la eliminación de sucesos que involucran la producción de quarks t . Los quarks t se desintegran en un aproximadamente 100% de los casos en un bosón W y un quark b . En este trabajo, se ha usado el *b-tag* para identificar los jets producidos a partir de la hadronización de un quark b que pueda provenir de procesos de fondo que contengan quarks t y así poder vetarlos. El algoritmo empleado es el *Combined Secondary Vertices* (DeepCSV) [20] que utiliza aprendizaje profundo y se basa en la combinación de información general de los vértices secundarios, sus parámetros de impacto y la cinemática de los jets. En el punto de selección escogido se tiene la máxima eficiencia posible de veto, cercana a un 86%.

- **MET**

La energía transversa faltante, MET (del inglés *Missing Transverse Energy*), o E_T^{miss} [21] está definida en base a la conservación de energía antes y después de la colisión. Se reconstruye como la suma vectorial negativa del momento transverso p_T de todas las partículas detectadas en el evento y reconstruidas mediante el algoritmo *Particle Flow*, tal que:

$$\vec{E}_T^{miss} = - \sum \vec{p}_T \quad (1)$$

Se asocia principalmente a los neutrinos, que escapan del detector sin dejar señal debido a su baja probabilidad de detección. No obstante, todos los eventos pueden mostrar una cierta MET instrumental incluso en procesos físicos que no involucran la producción de neutrinos.

5. PROCESOS DE SEÑAL Y FONDO

En este capítulo se realiza una descripción de la señal física que quiere estudiarse y de los principales fondos que van a presentar la misma signatura final, con el objetivo de entender las características físicas relevantes en dichos procesos que son de utilidad para separar señal y fondo.

5.1. Procesos de señal

El proceso estudiado como señal en este trabajo es la producción del bosón de Higgs a través de los cuatro mecanismos que se han mostrado en la Figura 2 de la Subsección 2.3 y su posterior desintegración a dos bosones W . Atendiendo a los modos de desintegración de los dos bosones W , el estudio se centra en la desintegración leptónica de los mismos, donde ambos bosones decaen a un leptón cargado y un neutrino, siendo la probabilidad de desintegración del bosón W para cada familia de leptones (e, μ, τ) del $10.86 \pm 0.09\%$ [22]. Por tanto, el estado final de este trabajo incluye dos leptones de carga opuesta y E_T^{miss} debida a los neutrinos.

5.2. Procesos de fondo

Hay un gran número de procesos relevantes según el SM con signaturas finales similares a la señal previamente descrita. Todos estos procesos se conocen como fondos y los más importantes son descritos a continuación.

- Drell-Yan ($q\bar{q} \rightarrow Z/\gamma^* \rightarrow l^+l^-$): tiene lugar cuando un quark procedente de un protón y un antiquark procedente de otro protón se aniquilan, creando un fotón virtual o un bosón Z que se desintegra a un par de leptones de carga opuesta. En la Figura 7 se muestra el diagrama de Feynman del proceso. La principal diferencia entre este fondo y la señal es la ausencia de MET real. En este caso no hay neutrinos en el estado final, por tanto la posible MET medida es instrumental.

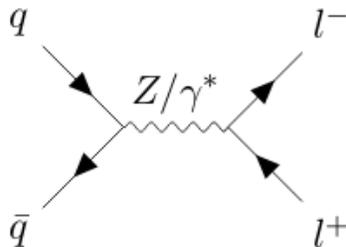


Figura 7: Diagramas de Feynman del proceso Drell-Yan.

- Dos bosones W ($q\bar{q}/gg \rightarrow W^+W^- \rightarrow l^+\nu l^-\bar{\nu}$): la producción de este proceso en colisiones pp en el LHC se produce en aproximadamente el 95% de los casos por

aniquilación quark-antiquark, y en torno a un 5 % por fusión de gluones. En la Figura 8 se muestran los posibles diagramas de Feynman del proceso. Este fondo tiene una signatura final similar a la señal cuando ambos bosones se desintegran leptónicamente; sin embargo, la cinemática es diferente debido a que los dos W no proceden directamente de la desintegración de un bosón de Higgs.

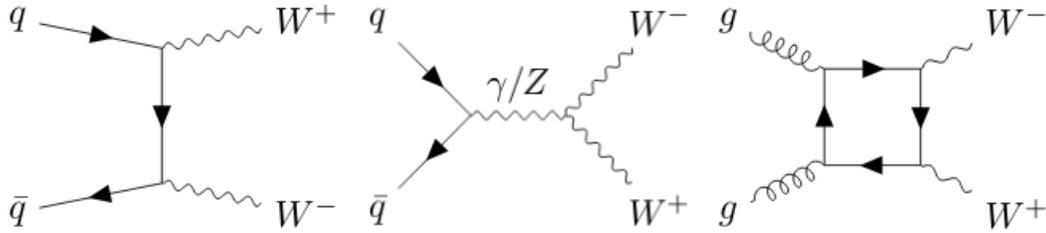


Figura 8: Diagramas de Feynman para la producción de dos bosones W .

- Top ($tW^\pm \rightarrow W^+bW^\pm \rightarrow l^+\nu l^\pm\nu + b$ y $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow l^+\nu l^-\bar{\nu} + 2b$): en este fondo se incluye la producción de un quark top asociado a un bosón W , así como la producción de un quark top y un quark antitop. En la Figura 9 se muestran ambos diagramas de Feynman. La desintegración del quark top es casi al 100 % en un bosón W y un quark b . De este modo, si se considera la desintegración leptónica del bosón W , se tienen dos leptones, MET, y un quark b debidos al proceso tW o dos quarks b debidos al proceso $t\bar{t}$, que pueden ser vetados aplicando el algoritmo de b -tag.

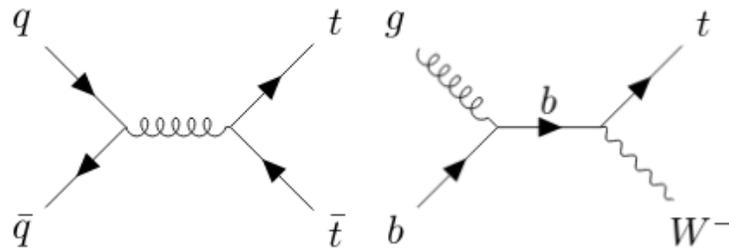


Figura 9: Diagramas de Feynman para la producción de fondo Top.

- W +jets ($W + jets \rightarrow l^\pm\nu l^\pm$): en este proceso, cuyo diagrama de Feynman se muestra en la Figura 10, se produce un bosón W junto con uno o más jets. Es un fondo instrumental dado que solo se tienen dos leptones y MET en el estado final si el W se desintegra leptónicamente y uno de los jets es identificado de manera errónea como un leptón aislado. La contribución de este fondo es más importante para leptones de bajo p_T , donde la probabilidad de ser mal identificados aumenta.

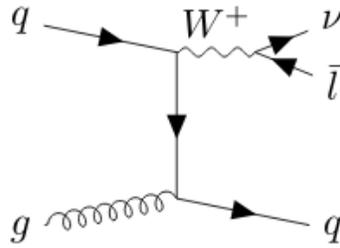


Figura 10: Diagramas de Feynman para el proceso W +jets.

- VZ : este proceso involucra la producción de un bosón W junto a un bosón Z o bien la producción de dos bosones Z . En el primer caso, $W^\pm Z \rightarrow l^\pm \nu l^+ l^-$, la contribución predominantemente ocurre cuando ambos bosones se desintegran leptónicamente y uno de los tres leptones del estado final no es detectado, o cuando el bosón W se desintegra hadrónicamente generando una signatura similar a la de Drell-Yan. En el segundo caso, $ZZ \rightarrow l^+ l^- \nu \bar{\nu}$, se considera fondo para la señal cuando uno de los bosones Z se desintegra leptónicamente y el otro se desintegra en neutrinos, resultando en un estado final con dos leptones y MET.
- $V\gamma$ y $V\gamma^*$: estos procesos se desintegran como $V\gamma^{(*)} \rightarrow l^\pm \nu l^+ l^-$, y son considerados como fondo para la señal cuando uno de los tres leptones que pueden ser producidos en el estado final no es identificado o cuando se producen exactamente dos leptones.
- VVV : este proceso involucra la producción de diversas combinaciones de tres bosones gauge (W^\pm, Z), y puede ser considerado como fondo si se detecta la misma signatura que para la señal en el estado final.

5.3. Muestras de Monte Carlo

Se han utilizado diferentes muestras basadas en métodos de Monte Carlo (MC) para simular tanto los procesos de señal como los de fondo. Estas simulaciones probabilísticas incluyen la parte teórica y la parte de reconstrucción del detector, y son producidas con diversos generadores de eventos.

En este trabajo, las muestras de MC empleadas simulan los datos que fueron tomados en 2017 por el experimento CMS del LHC durante el Run II a una energía de centro de masas de $\sqrt{s} = 13$ TeV y a una luminosidad integrada de 41.5 fb^{-1} . Cada una de ellas se ha escalado por dicha luminosidad y por su sección eficaz de producción.

Las principales muestras empleadas en el análisis se indican en la Tabla 1 junto con sus respectivas secciones eficaces de producción en pb [23].

Muestra	σ / pb
Drell-Yan (baja m_{ll}): $Z \rightarrow ll$	18610.00
Drell-Yan (alta m_{ll}): $Z \rightarrow ll$	6189.39
WW (prod. qq): $WW \rightarrow ll\nu\bar{\nu}$	12.18
Top: $t\bar{t}$	87.31
Top (single t, single \bar{t}): tW	35.60
VZ: $WZ \rightarrow llqq$	5.60
VZ: $WZ \rightarrow ll\nu$	4.43
VZ: $ZZ \rightarrow llqq$	3.22
VZ: $ZZ \rightarrow lll$	1.21
VZ: $ZZ \rightarrow ll\nu\nu$	0.56
$V\gamma$: $W\gamma \rightarrow l\nu\gamma$	405.27
HWW: ggH	0.99
HWW: $t\bar{t}H$	0.21
ZH, $H \rightarrow WW$: $ggZH$	1.00
ZH, $H \rightarrow WW$: $qqZH$	0.12

Tabla 1: Principales muestras de simulación empleadas en el análisis junto con su sección eficaz de producción σ .

Los generadores de los dos procesos principales que se quieren discernir en este trabajo por medio de un análisis multivariable son:

- POWHEG [24] para generar la señal HWW, esto es, la producción del bosón de Higgs por fusión de gluones;
- MADGRAPH [25] para generar las muestras que simulan el fondo Drell-Yan a baja masa del sistema dileptónico m_{ll} ;
- aMC@NLO [26] para generar la simulación de fondo Drell-Yan a alta masa del sistema dileptónico m_{ll} .

6. MÉTODOS DE ANÁLISIS MULTIVARIABLE

Una vez descritos los procesos de señal y fondo que son de interés en este trabajo, se introduce una serie de conceptos importantes en relación al análisis multivariable que va a ser desarrollado. Las fuentes consultadas para recabar la información aquí expuesta son [27], [28], [29], [30], [31], [32], [33].

6.1. Análisis multivariable

El análisis multivariable o MVA (del inglés *Multivariate Analysis*) permite combinar, mediante algún tipo de método o algoritmo predictivo, un conjunto de variables discriminatorias que describen propiedades de los eventos en un único discriminante que actúa como clasificador.

En este trabajo se ha empleado el paquete TMVA (del inglés *Toolkit for Multivariate Analysis*) que está integrado en ROOT [34] y proporciona un entorno para la aplicación de algoritmos de aprendizaje multivariable en problemas de clasificación y regresión.

En los problemas de clasificación, los distintos algoritmos tratan de etiquetar una serie de datos separándolos entre varias categorías o clases. En el marco de este trabajo, el problema de clasificación se centra en la separación de los sucesos de señal S , en este caso $H \rightarrow WW$, y los sucesos de fondo B , en este caso Drell-Yan. El objetivo es encontrar la mejor clasificación posible y para ello se ha utilizado el método multivariable BDT.

6.1.1. Fases y flujos de trabajo

Un algoritmo de clasificación consta de dos fases independientes: la fase de ejecución, donde los métodos multivariable son entrenados y probados, y la fase de aplicación, en la cual dichos métodos se aplican al problema de clasificación concreto, es decir, a datos reales.

La etapa de entrenamiento supervisado consiste en aprender las características de las diferentes clases de eventos en una muestra donde las clases o categorías (en este caso señal y fondo) son conocidas; el resultado final del modelo de predicción asigna una cierta probabilidad a cada suceso de pertenecer a una de las clases. Durante la prueba, se ejecuta el modelo sobre una muestra de datos estadísticamente independiente para validarlo.

En la Figura 11 se muestra un diagrama de las dos fases descritas y los diferentes pasos en cada una de ellas para el caso concreto de un BDT en TMVA. La integración en diversas rutinas de código, cuya comprensión constituye gran parte del trabajo que se ha desarrollado, incluye la importación de ficheros de entrada, la lectura de variables, la configuración de hiperparámetros del BDT, el entrenamiento y la prueba.

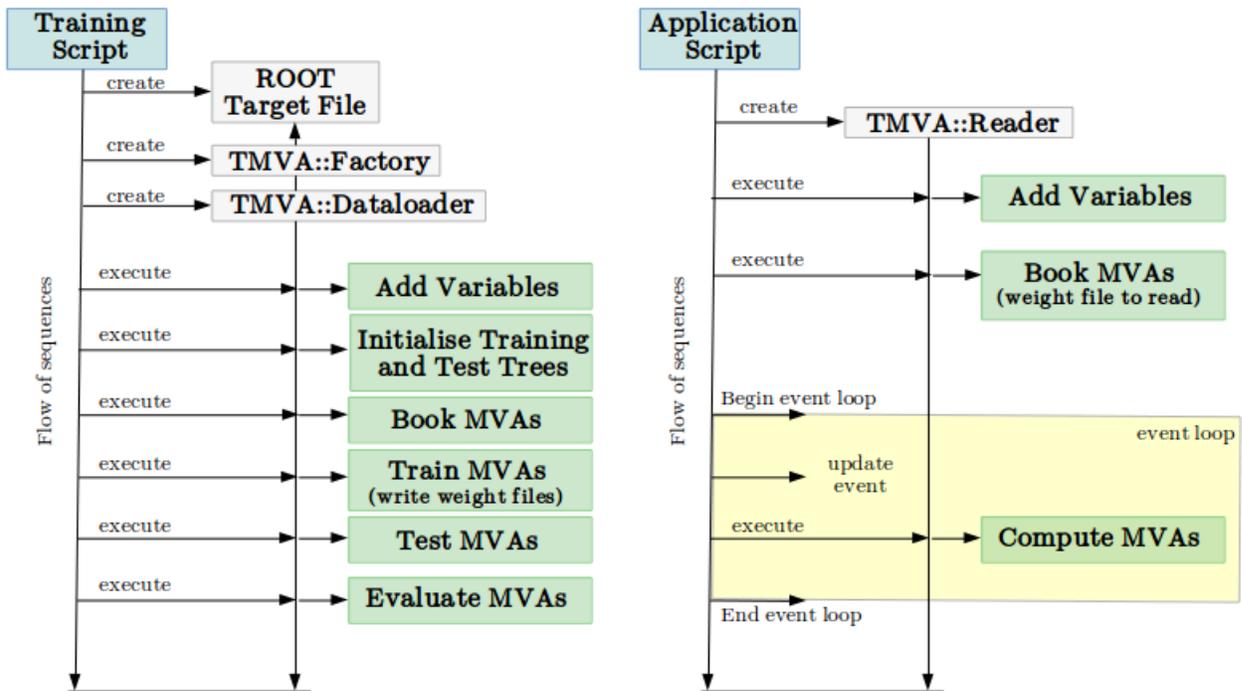


Figura 11: Flujo de secuencias en un script de ejecución y en un script de aplicación a través de TMVA.

6.2. Árboles de decisión

Un árbol de decisión es un tipo de algoritmo de aprendizaje supervisado que se usa principalmente en problemas de clasificación.

El principio básico de un árbol de decisión se centra en clasificar los datos según sus propiedades. Este tipo de algoritmo aplica criterios de clasificación de manera secuencial, de forma que en cada paso aquellos eventos que han sido clasificados de manera errónea tienen una nueva oportunidad para ver si otros criterios pueden ayudar a clasificarlos adecuadamente.

El proceso de crecimiento, como se muestra en la Figura 12, comienza con un nodo principal que se divide en dos nodos, de forma que se separan los datos en dos subconjuntos basándose en la variable que mejor separación proporciona. Para seleccionar dicha variable se estudia la ganancia de separación en diversos puntos de su rango, siendo la ganancia definida mediante el criterio *Gini Index* [27] como $p(1 - p)$, donde $p = S/(S+B)$ es la pureza de señal del nodo. La variable que proporciona la máxima ganancia es la escogida.

Las divisiones en nodos se repiten hasta que se alcanza un cierto criterio de parada, que puede ser el número mínimo de eventos por nodo, el máximo número de nodos, o

la máxima profundidad de árbol especificada, esto es, el número máximo de niveles de separación que puede tener lugar.

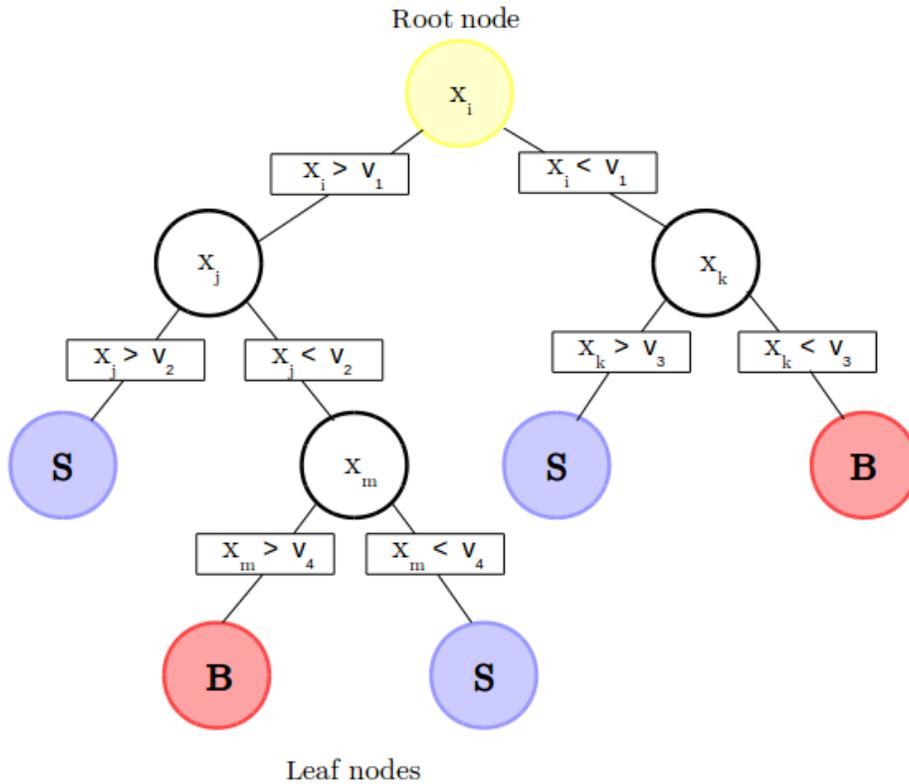


Figura 12: Esquema de un árbol de decisión. Se muestra el nodo principal (root node), los nodos intermedios, y los nodos finales (leaf nodes) que pueden ser de la clase señal *S* o fondo *B*. La variable discriminante seleccionada en cada nodo está representada por x_i , x_j , x_k y x_m ; el valor del corte óptimo de separación en cada variable está representado por v_1 , v_2 , v_3 y v_4 .

Un defecto de los árboles de decisión es su inestabilidad con respecto a fluctuaciones estadísticas en la muestra de entrenamiento. Por ejemplo, si dos variables de entrada exhiben un poder de separación similar, una fluctuación en la muestra de entrenamiento puede causar que el algoritmo de crecimiento del árbol decida separar seleccionando una cierta variable, mientras que otra variable podría haber sido seleccionada si no existiera la fluctuación; en tal caso, la estructura completa del árbol se ve alterada a partir de este nodo, resultando posiblemente en una respuesta clasificadora diferente. Este problema se supera construyendo un conjunto de árboles, es decir, un bosque, y clasificando un evento según la mayoría de las clasificaciones hechas por ellos.

La profundidad de los árboles en un bosque se reduce drásticamente durante el proceso de crecimiento, minimizando así la tendencia a especializarse en el conjunto de datos de entrenamiento presente en los árboles de decisión simples, que típicamente crecen hasta alcanzar una gran profundidad de forma que los nodos finales dejan de tener suficiente

estadística.

Estos bosques son frecuentemente denominados BDTs, y se construyen mediante algoritmos de *boosting* y *bagging*.

6.3. Algoritmos de construcción de BDTs

En este trabajo se ha estudiado el rendimiento de tres algoritmos: uno de *boosting* conocido como *AdaBoost*, *Bagging*, y una variante de este último denominada *Random Forest*. Estos algoritmos, que generan un conjunto de árboles, aumentan la estabilidad estadística del clasificador y son capaces de mejorar drásticamente el rendimiento de separación en comparación a un único árbol de decisión.

A continuación, se explican los diferentes algoritmos indicando las diferencias más relevantes entre los mismos: si operan secuencialmente o en paralelo, si repesan o no los eventos, y el modo de selección de las variables.

6.3.1. Boosting - AdaBoost

El procedimiento de *boosting* es secuencial, es decir, la entrada de cada árbol individual depende de la salida del árbol anterior.

De entre los diferentes algoritmos de *boosting*, en este trabajo se aborda únicamente *AdaBoost* (acrónimo de *Adaptive Boost*). En un problema de clasificación, este algoritmo asigna en el entrenamiento de un árbol un peso mayor a los eventos que fueron mal clasificados durante el entrenamiento del árbol anterior. Tras la primera iteración, los eventos mal clasificados se multiplican por un peso común α denominado *boost weight*, de forma que el árbol subsecuente es entrenado usando una muestra de eventos modificados. Este peso se deriva de la proporción de eventos mal clasificados, err , en el árbol previo, tal que $\alpha = \frac{1-err}{err}$. Los pesos de la muestra de eventos completa son entonces normalizados tal que la suma de pesos se mantiene constante.

El rendimiento aumenta forzando que el aprendizaje sea más lento. La velocidad de aprendizaje se controla por medio de un parámetro β , de manera que se repesan los eventos mal clasificados por α^β .

Las predicciones para un nuevo conjunto de datos, $y_{Boost}(x)$, se basan en las predicciones de los árboles individuales, $h_i(x)$, asignando un peso, $\ln(\alpha_i)$, a cada árbol en función de su rendimiento individual, tal que:

$$y_{Boost}(x) = \frac{1}{N} \cdot \sum_i^N \ln(\alpha_i) \cdot h_i(x) \quad (2)$$

6.3.2. Bagging

Este algoritmo se llama también *Bootstrap aggregating*, y a diferencia del procedimiento de *boosting*, opera en paralelo. Se crean m submuestras aleatoriamente con posibilidad de reemplazo (*bootstrapping*) a partir de los datos de entrenamiento, y se construye un árbol independiente para cada submuestra.

Al igual que en *boosting*, la predicción para un nuevo conjunto de datos se hace a partir del valor de predicción predominante en los m árboles individuales. Sin embargo, mientras que en *boosting* se asigna más peso a aquellos con mayor rendimiento, en *bagging* todos los árboles tienen el mismo peso.

6.3.3. Random Forest

El algoritmo *Random Forest* (RF) es una versión mejorada del algoritmo *Bagging*.

Al igual que en *Bagging*, opera en paralelo construyendo un árbol para cada submuestra de los datos de entrada. Sin embargo, a diferencia de los anteriores algoritmos, cada árbol se construye con una submuestra aleatoria del conjunto de variables de entrada. Así se evita que si una variable tiene un poder discriminatorio mucho mayor que los demás aparezca en la mayoría de árboles y por tanto el modelo dependa en gran medida de ésta.

6.4. Opciones de configuración en BDTs

Las principales opciones de configuración disponibles en el paquete TMVA para personalizar el clasificador BDT, que han sido definidas y modificadas en este trabajo, son:

- NTrees: número de árboles en el bosque;
- MaxDepth: profundidad máxima permitida de los árboles de decisión;
- MinNodeSize: mínimo porcentaje de eventos de entrenamiento requeridos en un nodo final u hoja;
- nCuts: número de puntos de división en el rango de las variables (granularidad), usado para encontrar el corte óptimo en la clasificación en un nodo;
- BoostType: algoritmo empleado para la generación de los árboles que constituyen un bosque. Los algoritmos óptimos suelen ser: *AdaBoost* y *Bagging*;
- AdaBoostBeta: velocidad de aprendizaje para el algoritmo *AdaBoost*;
- UseRandomisedTrees: escoge aleatoriamente de entre las variables de entrada un subconjunto de variables en cada nodo. Actúa como *Random Forest*.

7. ANÁLISIS SECUENCIAL

Antes de realizar el análisis multivariable y una vez introducidas las muestras de simulación empleadas en este trabajo, así como los procesos de señal y fondo de mayor interés, se ha realizado una secuencia de selecciones en distintas variables con el fin de obtener una región de señal lo más pura posible en Higgs.

En esta sección, se presentan los resultados del análisis secuencial llevado a cabo.

7.1. Selección

Como se ha descrito en la Sección 5, el estado final de estudio es de dos leptones de signo opuesto y mismo sabor ($ee/\mu\mu$), así como una cierta cantidad de energía transversa perdida debida a la presencia de neutrinos. Se requiere que ambos leptones estén bien identificados y aislados para rechazar leptones en jets, que tienen mayor energía en torno a la traza debido al resto de partículas.

Se han separado los eventos en dos categorías en función del número de jets (reconstruidos, según la definición de jet de la Sección 4), dividiéndolos en aquellos con 0-jets o 1-jet, donde la sección eficaz de producción del bosón de Higgs por fusión de gluones es máxima.

Para esta medida se han seleccionado dos leptones con $p_T^{l1} > 25$ GeV para el primer leptón, $p_T^{l2} > 13$ GeV para el segundo leptón en el canal ee y $p_T^{l2} > 10$ GeV para el segundo leptón en el canal $\mu\mu$. La diferencia entre electrones y muones viene dada por la mínima selección en p_T que se realiza en los triggers.

La contribución de fondos con tres o más leptones en el estado final, tal como los procesos VZ , $V\gamma(*)$ y VVV , se reduce rechazando aquellos eventos que tienen un tercer leptón reconstruido con $p_T > 10$ GeV.

Por otra parte, para rechazar resonancias de baja masa, como J/Ψ o Υ , se impone que la masa invariante del sistema dileptónico m_{ll} sea mayor que 20 GeV. Además, se reduce también fondo DY requiriendo que la diferencia en valor absoluto entre la masa del sistema dileptónico y la masa del bosón Z sea mayor que 15 GeV (Z -veto), ya que dicho bosón tiene una masa muy bien medida de (91.1876 ± 0.0021) GeV [22]. En la Figura 13 se observa la distribución en el nivel previo a realizar esta selección.

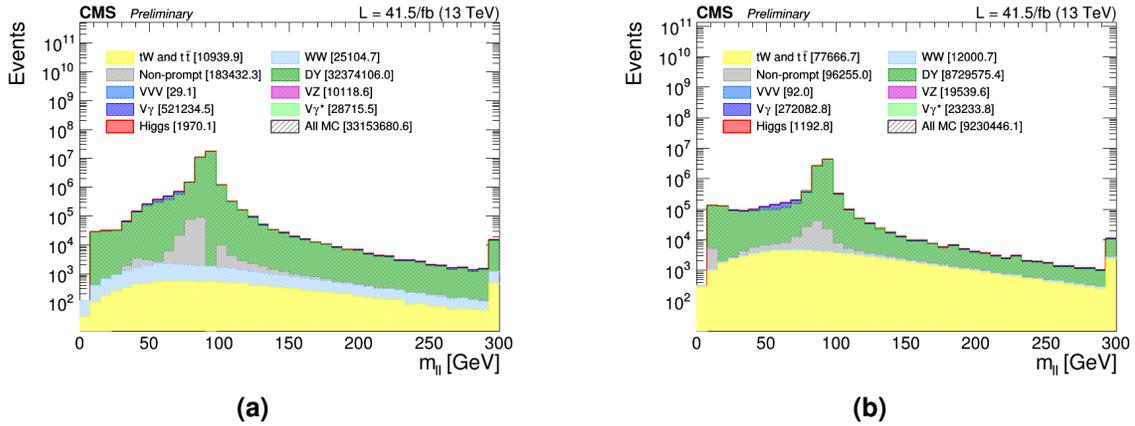


Figura 13: Distribución en escala logarítmica de m_{ll} en el canal de (a) 0-jets y (b) 1-jet antes de realizar la selección en esta variable.

Se requiere que el momento transverso del sistema dileptónico p_T^{ll} sea mayor que 30 GeV, ya que en promedio los pares de leptones provenientes de DY tienen un momento más bajo que aquellos procedentes de la desintegración del proceso $H \rightarrow WW$. Este nivel de preselección suprime a su vez fuertemente el fondo W +jets. En la Figura 14 se muestra la distribución de esta variable antes de realizar la selección, observándose así su poder discriminatorio contra dichos fondos, especialmente contra DY.

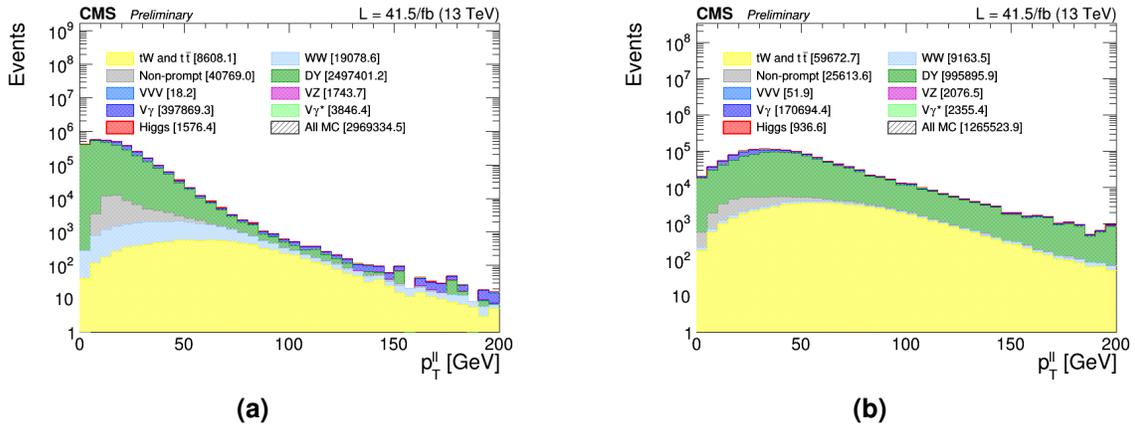


Figura 14: Distribución en escala logarítmica de p_T^{ll} en el canal de (a) 0-jets y (b) 1-jet antes de realizar la selección en esta variable.

Para obtener una mayor reducción de fondo DY se pide un mínimo de MET, puesto que dicho proceso no produce neutrinos en el estado final y por tanto la MET solo es instrumental. Se requiere una selección mayor que 50 GeV en la variable p_{up}^{met} , que está definida como la MET total que se mide empleando un algoritmo análogo a *Particle Flow* (ver Sección 4) pero aplicando otro procedimiento para reducir el efecto de *pile-up*. Así mismo, se hace una selección sobre la variable m_{pmet} , que se define como el mínimo entre la proyección de la MET computada únicamente a partir de partículas cargadas que

se miden en el *tracker* (TkE_T^{miss}), y la proyección de la MET total (E_T^{miss}) que considera tanto las partículas cargadas del evento como las neutras.

$$projected E_T^{miss} = \begin{cases} E_T^{miss} & \Delta\phi_{min}(leptons, E_T^{miss}) \geq \pi/2 \\ E_T^{miss} \sin\Delta\phi_{min} & \Delta\phi_{min}(leptons, E_T^{miss}) \leq \pi/2 \end{cases} \quad (3)$$

$$projected TkE_T^{miss} = \begin{cases} TkE_T^{miss} & \Delta\phi_{min}(leptons, TkE_T^{miss}) \geq \pi/2 \\ TkE_T^{miss} \sin\Delta\phi_{min} & \Delta\phi_{min}(leptons, TkE_T^{miss}) \leq \pi/2 \end{cases} \quad (4)$$

$$mpmet = \min(projected E_T^{miss}, projected TkE_T^{miss}) \quad (5)$$

La distribución de esta variable permite eliminar un gran porcentaje del fondo DY restante aplicando una selección de $mpmet > 20$ GeV. Esto se debe a que, por definición, en sucesos con MET real (debida a la presencia de neutrinos), como en el caso de $H \rightarrow WW$, ésta tiene dirección en sentido contrario al p_T de los objetos que se reconstruyen en el evento. Sin embargo, en el proceso de DY la MET procede de la resolución del detector, por lo que se espera una dirección aleatoria. En base a las ecuaciones 3 y 4, cuando la MET está angularmente próxima a algún leptón del DY, se multiplica por el seno del ángulo que forman las direcciones de los leptones y de la MET. De este modo, en sucesos de DY se consigue sesgar la MET hacia valores bajos, mientras que la correspondiente a procesos con MET real no se ve modificada.

Las distribuciones de $puppimet$ y $mpmet$ se observan en las Figuras 15 y 16, mostrando cómo a este nivel de selección el fondo DY es el predominante.

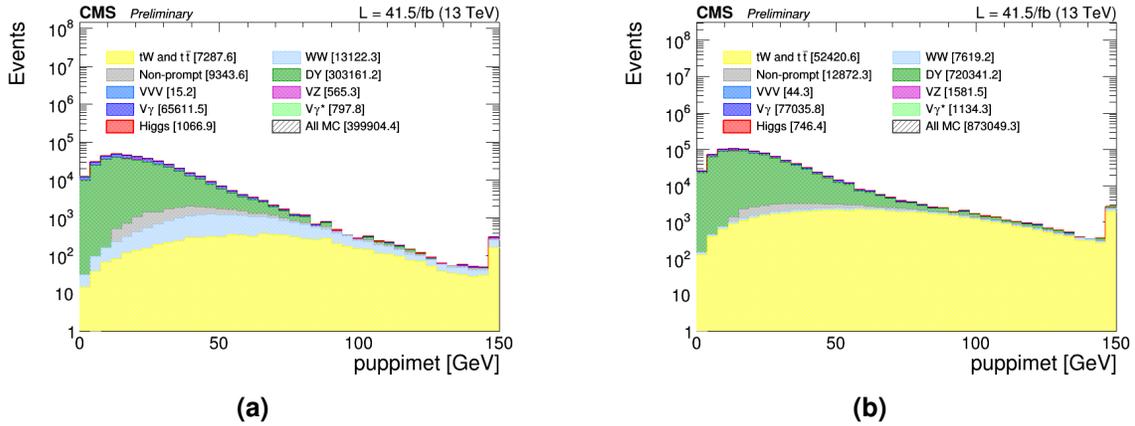


Figura 15: Distribución en escala logarítmica de p_{puppimet} en el canal de (a) 0-jets y (b) 1-jet antes de realizar la selección en esta variable.

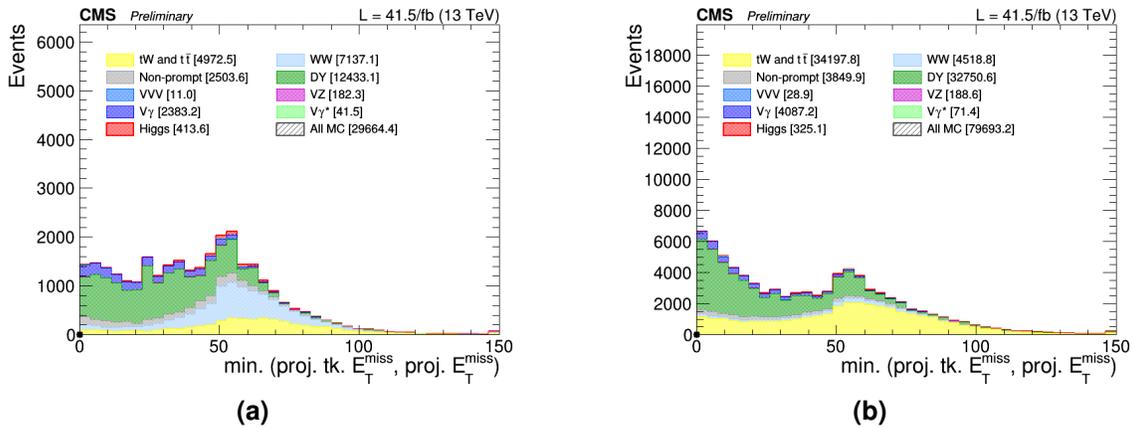


Figura 16: Distribución de m_{pmet} en el canal de (a) 0-jets y (b) 1-jet antes de realizar la selección en esta variable.

Finalmente, se puede reducir el fondo de Top vetando aquellos eventos que contienen al menos un jet identificado como proveniente de un quark b .

La selección final utilizada en esta análisis se resume en la Tabla 2.

Variable	Selección
Canal de desintegración	$ee / \mu\mu$
$q_{l_1} \times q_{l_2}$	< 0
$p_T(l_1)$ [GeV]	> 25
$p_T(l_2)$ [GeV]	> 13 (e) > 10 (μ)
nLepton	≥ 2
Veto tercer leptón ($p_T(l_3) > 10$ [GeV])	Aplicado
m_{ll} [GeV]	> 20
$ 91,1876 - m_{ll} $ [GeV]	> 15
p_T^{ll} [GeV]	> 30
p_{ppimet} [GeV]	> 50
m_{pmet} [GeV]	> 20
bVeto Deep CSV - L (número de b -jets = 0)	Aplicado
njets	0 / 1

Tabla 2: Selección aplicada en el análisis. La carga y el momento transversal de los leptones se representa como q y p_T , m_{ll} y p_T^{ll} son la masa invariante y el momento transversal del sistema dileptónico; p_{ppimet} y m_{pmet} son, respectivamente, la MET reconstruida con Particle Flow aplicando correcciones, y el mínimo entre las proyecciones de la MET medida en el tracker y de la MET total.

Una vez aplicada toda la selección descrita, se representan las principales distribuciones que queremos estudiar para procesos de Monte Carlo en las Figuras 17-22 para el canal de 0-jets, y en las Figuras 23-28 para el canal de 1-jet. En cada caso se muestran seis variables: el momento transversal de cada uno de los leptones $p_T(l_1)$ y $p_T(l_2)$, el momento transversal del sistema dileptónico p_T^{ll} , la masa invariante del sistema dileptónico m_{ll} , la diferencia angular entre los dos leptones $\Delta\phi_{ll}$, y la energía perdida transversal m_{pmet} .

0-jets

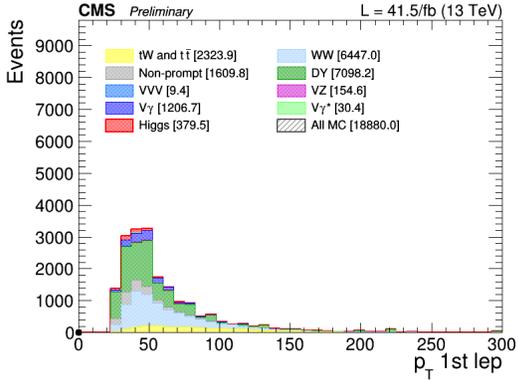


Figura 17: Distribución del momento transverso del leptón más energético $p_T(l_1)$ en el canal de 0-jets.

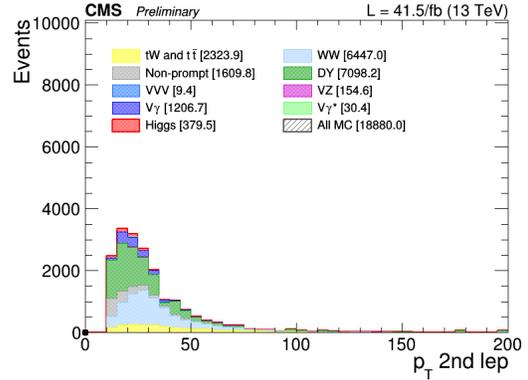


Figura 18: Distribución del momento transverso del segundo leptón $p_T(l_2)$ en el canal de 0-jets.

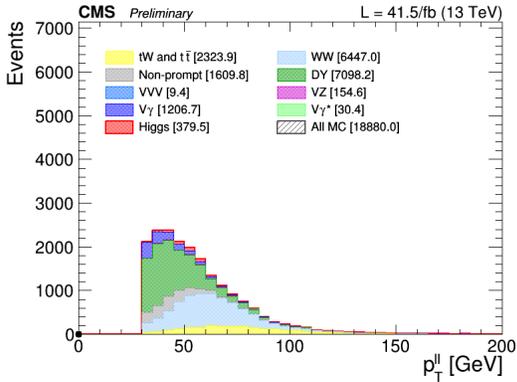


Figura 19: Distribución del momento transverso del sistema dileptónico p_T^{ll} en el canal de 0-jets.

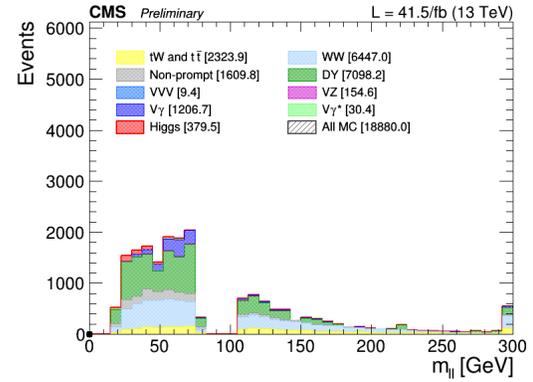


Figura 20: Distribución de la masa invariante del sistema dileptónico m_{ll} en el canal de 0-jets.

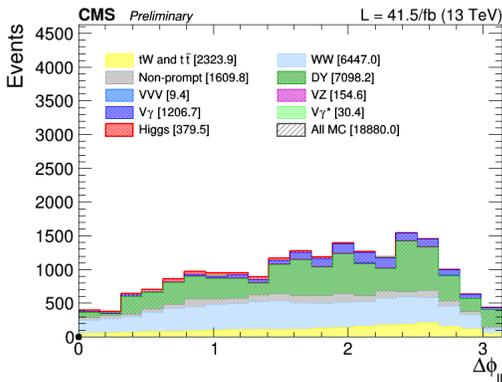


Figura 21: Distribución de la diferencia angular entre los dos leptones $\Delta\phi_{ll}$ en el canal de 0-jets.

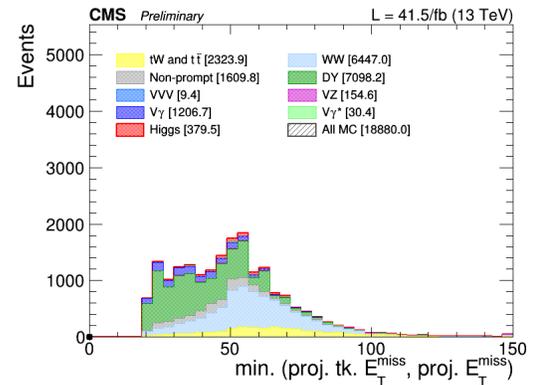


Figura 22: Distribución de la energía perdida transversa m_{pmet} en el canal de 0-jets.

1-jet

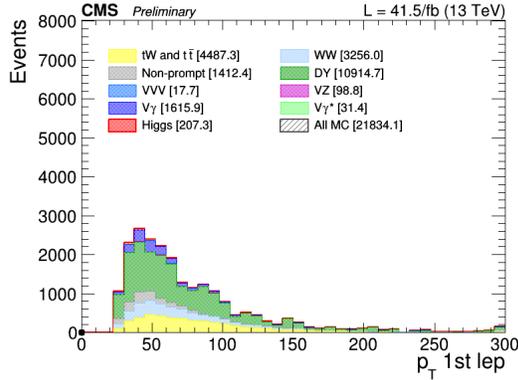


Figura 23: Distribución del momento transverso del leptón más energético $p_T(l_1)$ en el canal de 1-jet.

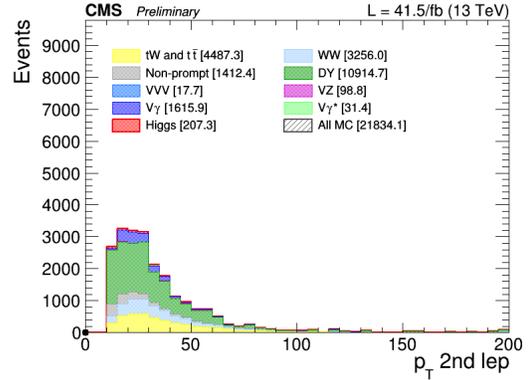


Figura 24: Distribución del momento transverso del segundo leptón $p_T(l_2)$ en el canal de 1-jet.

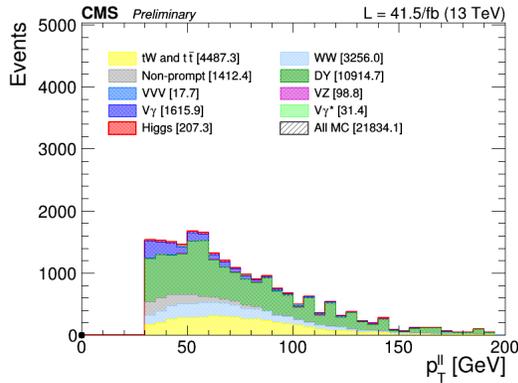


Figura 25: Distribución del momento transverso del sistema dileptónico p_T^{ll} en el canal de 1-jet.

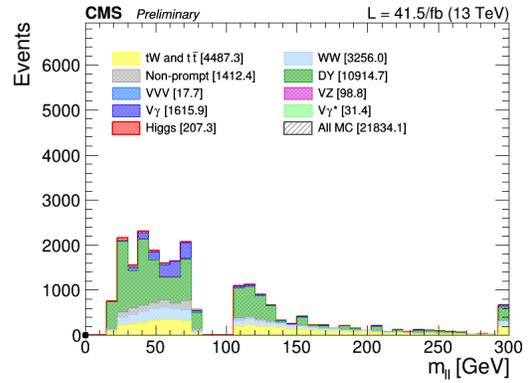


Figura 26: Distribución de la masa invariante del sistema dileptónico m_{ll} en el canal de 1-jet.

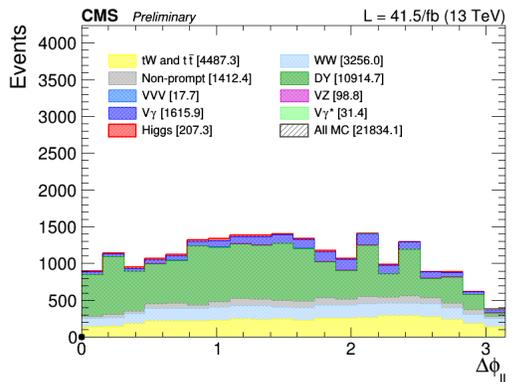


Figura 27: Distribución de la diferencia angular entre los dos leptones $\Delta\phi_{ll}$ en el canal de 1-jet.

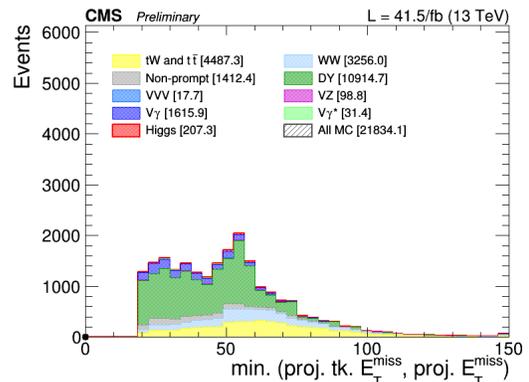


Figura 28: Distribución de la energía perdida transversa m_{pmet} en el canal de 1-jet.

7.2. Significancia de señal

En las Tablas 3 y 4, se muestra el número de eventos de los fondos y de la señal a cada nivel de selección para los análisis de eventos con 0-jets y 1-jet reconstruidos respectivamente, así como la significancia de señal, Sig , que está definida como:

$$Sig = \frac{S}{\sqrt{S+B}} \quad (6)$$

donde S y B representan respectivamente el total de eventos de señal ($H \rightarrow WW$) y de fondo (suma de contribuciones del resto de procesos físicos) que pasan una cierta selección.

Selección	S ($H \rightarrow WW$)	DY	WW	Top	Non-prompt	VV(V)	$V_{\gamma}^{(*)}$	B	Sig
$p_T(l_1) > 25$ GeV $p_T(l_2) > 13$ GeV (e) $p_T(l_2) > 10$ GeV (μ) nLepton ≥ 2	1987.7	32388104.2	25114.1	11360.7	183961.1	11150.1	561303.4	33180993.6	0.35
Veto tercer leptón	1970.1	32374106.0	25104.7	10939.9	183432.3	10147.7	549950.0	33153680.6	0.34
$m_{ll} > 20$ GeV	1701.3	32325436.5	24426.4	10695.8	195192.2	10070.6	549704.4	33115526.0	0.30
$ 91,1876 - m_{ll} > 15$ GeV	1576.4	2497401.2	19078.6	8608.1	40769.0	1761.9	401715.7	2969334.5	0.92
$p_T^H > 30$ GeV	1066.9	303161.2	13122.3	7287.6	9343.6	580.5	66409.3	399904.4	1.69
$p_{upmet} > 50$ GeV	413.6	12433.1	7137.1	4972.5	2503.6	193.3	2424.7	29664.4	2.39
$m_{pmet} > 20$ GeV	396.7	8130.7	6682.3	4505.6	1850.1	172.9	1370.9	22712.4	2.61
bVeto Deep CSV - L	379.5	7098.2	6447.0	2323.9	1609.8	164.0	1237.1	18880.0	2.74

Tabla 3: Número de eventos de los procesos de señal y fondo tras aplicar cada selección en el análisis de eventos con 0-jets reconstruidos. S representa los eventos de señal y B la suma de eventos de los distintos procesos de fondo; Sig es la significancia de señal.

Selección	S ($H \rightarrow WW$)	DY	WW	Top	Non-prompt	VV(V)	$V_{\gamma}^{(*)}$	B	Sig
$p_T(l_1) > 25$ GeV $p_T(l_2) > 13$ GeV (e) $p_T(l_2) > 10$ GeV (μ) nLepton ≥ 2	1209.1	8734492.8	12009.0	78940.7	96767.2	20154.8	298878.6	9241243.2	0.40
Veto tercer leptón	1192.8	8729575.4	12000.7	77666.7	96255.0	19631.6	295316.6	9230446.1	0.39
$m_{ll} > 20$ GeV	1023.5	8515755.1	11644.5	75386.7	106728.4	19184.6	294835.4	9023534.6	0.34
$ 91,1876 - m_{ll} > 15$ GeV	936.6	995895.9	9163.5	59672.7	25613.6	2128.4	173049.8	1265523.9	0.83
$p_T^H > 30$ GeV	746.4	720341.2	7619.2	52420.6	12872.3	1625.8	78170.1	873049.3	0.80
$p_{upmet} > 50$ GeV	325.1	32750.6	4518.8	34197.8	3849.9	217.5	4158.6	79693.2	1.15
$m_{pmet} > 20$ GeV	246.5	15252.2	3770.3	29002.6	2487.9	153.7	2042.3	52709.0	1.07
bVeto Deep CSV - L	207.3	10914.7	3256.0	4487.3	1412.4	116.5	1647.3	21834.1	1.40

Tabla 4: Número de eventos de los procesos de señal y fondo tras aplicar cada selección en el análisis de eventos con 1-jet reconstruido. S representa los eventos de señal y B la suma de eventos de los distintos procesos de fondo; Sig es la significancia de señal.

Tras la secuencia completa de selecciones que se muestra en la Tabla 2, se ha llegado a una región enriquecida en eventos de $H \rightarrow WW$. Las significancias obtenidas en el último

nivel de selección son $Sig_{secuencial}^{0j} = 2.74$ y $Sig_{secuencial}^{1j} = 1.40$ para el canal de 0-jets y 1-jet respectivamente.

Como se aprecia en las Tablas 3 y 4, el canal de 0-jets es más puro en señal que el canal de 1-jet. A último nivel de selección, el fondo predominante en ambos canales es DY; no obstante, el segundo fondo con mayor número de eventos es el WW en el caso de eventos con 0-jets reconstruidos y el Top en el caso de eventos con 1-jet reconstruido. Esto se debe a que los quarks b a partir de los cuales se reconstruyen los jets suelen provenir de desintegraciones de quarks t , con que es un fondo más importante en el canal de 1-jet.

8. ANÁLISIS BASADO EN UN BDT

Con el objetivo de intentar obtener una región más pura en Higgs que la que se tiene al nivel de selección final de la Sección 7, se ha empleado un BDT. Puesto que el principal modo de producción del bosón de Higgs es el mecanismo de fusión de gluones, en el análisis multivariable se ha considerado exclusivamente como señal este proceso. Por su parte, en cuanto a los procesos de fondo, se ha tomado únicamente DY por ser el predominante tras la selección final del análisis secuencial.

Dado que algunas de las variables más discriminatorias de fondo DY respecto de $H \rightarrow WW$ son las relativas a la MET, se han empleado muestras con una selección menos restrictiva que en el análisis secuencial. Concretamente, en el entrenamiento del BDT se han seleccionado eventos con $p_{\text{up}}^{\text{pupimet}} > 20$ GeV en lugar de $p_{\text{up}}^{\text{pupimet}} > 50$ GeV para tener en cuenta la información proporcionada por otras variables en los eventos con valores bajos de MET.

Tras dicha selección, se ha dispuesto de 30218 eventos de señal y 123142 eventos de fondo Drell-Yan, para una luminosidad total de 41.5 fb^{-1} . El conjunto de eventos ha sido dividido a la mitad generando así dos muestras independientes: una para la fase de entrenamiento y otra para la fase de testeo.

La evaluación del BDT se ha desarrollado con las clases en C++ del Apéndice I.

8.1. Variables de entrada

Como se ha explicado en la Sección 6, el BDT es un MVA que toma como entrada una serie de variables para su entrenamiento. En base al estudio desarrollado en la Sección 7 (ver Figuras 17-28 y Apéndice II), se han seleccionado las siguientes variables por mostrar un cierto poder de separación entre señal y fondo:

Variables cinemáticas:

- m^{ll} (m_{ll}): masa invariante del sistema dileptónico.
- p_T^{ll} (p_{Tll}): momento transversal del sistema dileptónico.
- p_T^{l1} ($\text{Lepton_pt}[0]$): momento transversal del leptón más energético en el evento.
- p_T^{l2} ($\text{Lepton_pt}[1]$): momento transversal del segundo leptón más energético en el evento.
- E_T^{miss} (MET_pt): energía transversal perdida reconstruida con el algoritmo *Particle Flow* conforme a la ecuación 1.
- $\text{Track}E_T^{\text{miss}}$ (TkMET_pt): E_T^{miss} computada únicamente a partir de las partículas cargadas medidas en el *tracker*.

- $\underline{ProjectedE_T^{miss}}$ (*projpfmet*): proyección de la E_T^{miss} total reconstruida con *Particle Flow*. Está definida por la ecuación 3
- $\underline{Min.projectedE_T^{miss}}$ (*mpmet*): valor mínimo entre $\underline{ProjectedE_T^{miss}}$ y la proyección de $\underline{TrackE_T^{miss}}$ que está definida por la ecuación 5.
- $\underline{PuppiE_T^{miss}}$ (*PuppiMET_pt*): E_T^{miss} aplicando otro procedimiento para reducir el efecto de *pile-up*.
- $\underline{PfMetDivSumMet}$: significancia de la MET. Está definida como:

$$PfMetDivSumMet = E_T^{miss} / \sqrt{E_T} \quad (7)$$

Variables angulares:

- $\underline{\Delta\phi_U}$ (*dphill*): diferencia en el ángulo azimutal entre los dos leptones.
- $\underline{\Delta\phi_{U,E_T^{miss}}}$ (*dphillmet*): diferencia en el ángulo azimutal entre el sistema dileptónico y la E_T^{miss} .
- $\underline{\Delta\phi_{l_1,E_T^{miss}}}$ (*dphilmet1*): diferencia en el ángulo azimutal entre el leptón más energético y la E_T^{miss} .
- $\underline{\Delta\phi_{l_2,E_T^{miss}}}$ (*dphilmet2*): diferencia en el ángulo azimutal entre la E_T^{miss} y el segundo leptón más energético.

Variables de alto nivel:

- $\underline{m_T^H}$ (*meth*): masa transversa del bosón de Higgs. Está definida como:

$$m_T^H = \sqrt{2p_T^{ll} E_T^{miss} (1 - \cos\Delta\phi_{U,E_T^{miss}})} \quad (8)$$

- $\underline{m_T^{W1}}$ (*mtw1*): masa transversa del bosón más energético. Está definida como:

$$m_T^{W1} = \sqrt{2p_T^{l1} E_T^{miss} (1 - \cos\Delta\phi_{l_1,E_T^{miss}})} \quad (9)$$

- $\underline{m_T^{W2}}$ (*mtw2*): masa transversa del segundo bosón más energético. Está definida como:

$$m_T^{W2} = \sqrt{2p_T^{l2} E_T^{miss} (1 - \cos\Delta\phi_{l_2,E_T^{miss}})} \quad (10)$$

En las Figuras 29, 30 y 31 se representan las distribuciones, normalizadas al número de eventos, de estas variables para los procesos de señal (azul) y fondo (rojo) seleccionados en el análisis multivariable, mostrándose el poder discriminatorio individual de cada una

de ellas. El grado de separación en todas las variables no es el mismo; sin embargo, para poder hacer uso de todas ellas, es precisamente por lo que se emplea un BDT.

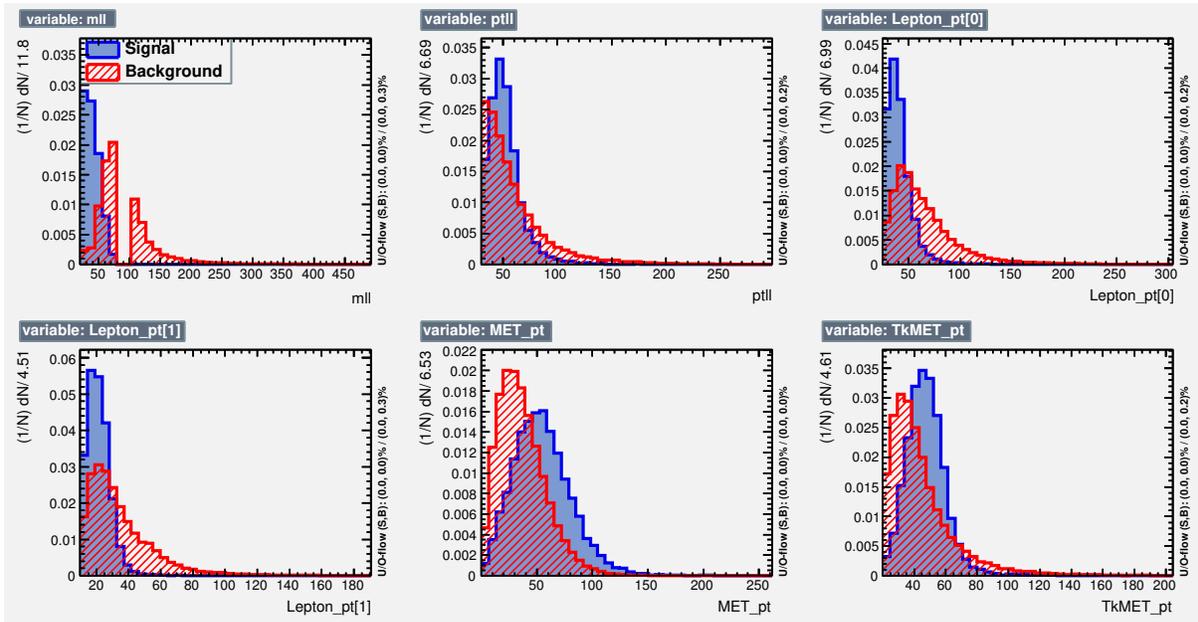


Figura 29: Distribución de las variables discriminatorias consideradas en el análisis multi-variable, tanto para señal en azul como para fondo en rojo.

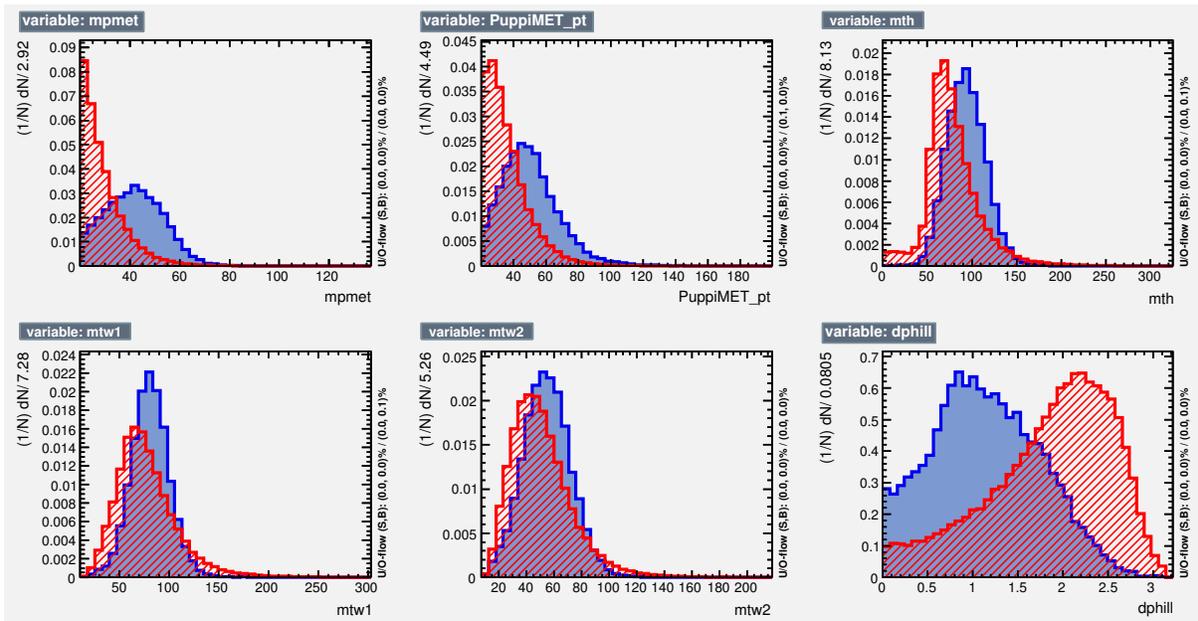


Figura 30: Distribución de las variables discriminatorias consideradas en el análisis multi-variable, tanto para señal en azul como para fondo en rojo.

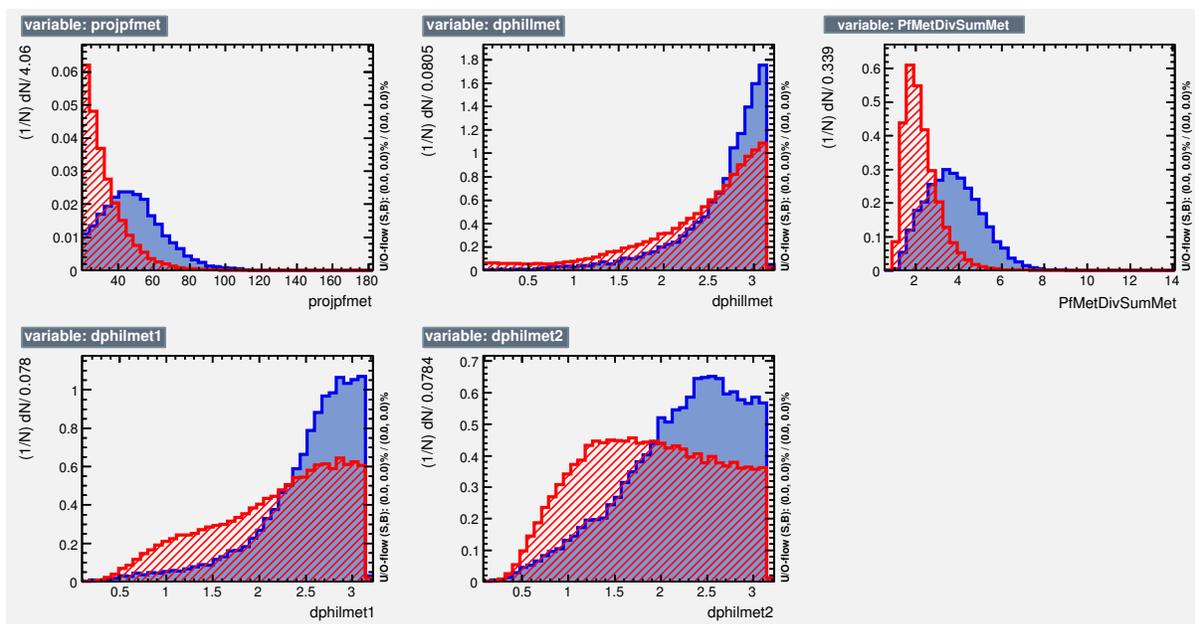


Figura 31: Distribución de las variables discriminatorias consideradas en el análisis multivariable, tanto para señal (azul) como para fondo (rojo).

8.2. Elección del número de jets para el entrenamiento

Como se ha explicado en la Sección 7, el análisis se ha realizado separadamente para eventos con 0-jets o con 1-jet reconstruidos.

El propósito de esta subsección es probar si un único BDT entrenado conjuntamente con eventos que tienen 0-jets y 1-jet reconstruidos, puede dar un resultado similar para eliminar sucesos de DY que dos BDTs entrenados por separado, uno con eventos que poseen 0-jets y otro con eventos con 1-jet. Si esto es así, sería posible usar un solo BDT por simplicidad. Por tanto, el objetivo es comparar la significancia y el rendimiento proporcionados por dichos BDTs al evaluar, según el caso, sobre eventos con 0-jets reconstruidos o sobre eventos con 1-jet reconstruido.

Para el entrenamiento de los tres BDTs se ha seleccionado el algoritmo *AdaBoost* y se ha mantenido el valor por defecto del resto de opciones de configuración que se indican en la Subsección 6.4.

En las Figuras 32 y 33 se muestra, para el conjunto de eventos con 0-jets reconstruidos, la distribución de la variable de salida que ha sido obtenida entrenando, respectivamente, con eventos que poseen únicamente 0-jets y eventos que poseen 0-jets o 1-jet. De manera análoga, en las Figuras 34 y 35 se muestra, para el conjunto de eventos con 1-jet reconstruido, la distribución de la variable de salida correspondiente a entrenar solo con eventos que poseen 1-jet y aquella obtenida entrenando con eventos con 0-jets y 1-jet. Los eventos representados han sido sometidos a la secuencia de selecciones de la Sección 7,

a excepción de la selección en $p_{\text{up}}^{\text{pup}} \text{met}$ que se ha relajado a $p_{\text{up}}^{\text{pup}} \text{met} > 20 \text{ GeV}$.

Eventos con 0-jets reconstruidos:

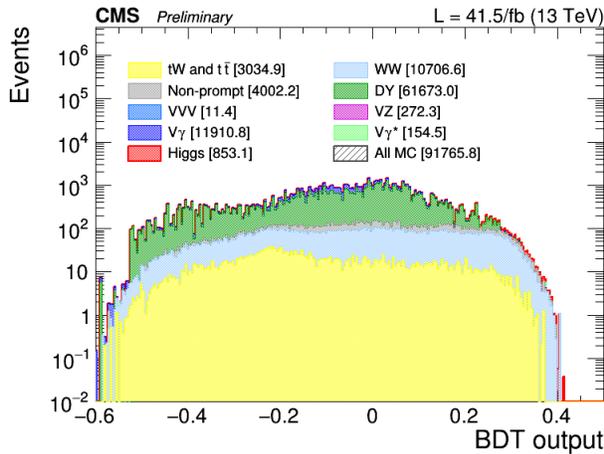


Figura 32: Distribución en escala logarítmica, para eventos con 0-jets reconstruidos, de la variable de salida al entrenar con eventos con 0-jets o 1-jet.

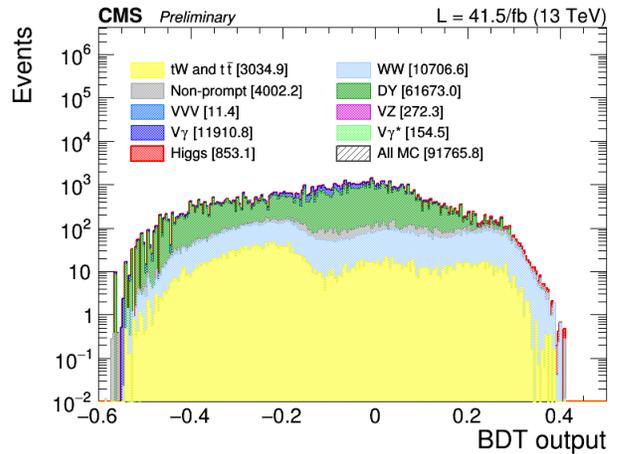


Figura 33: Distribución en escala logarítmica, para eventos con 0-jets reconstruidos, de la variable de salida al entrenar con eventos con 0-jets.

Eventos con 1-jet reconstruido:

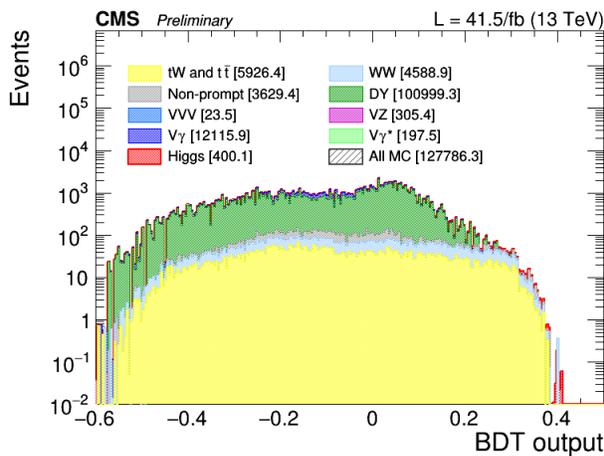


Figura 34: Distribución en escala logarítmica, para eventos con 1-jet reconstruido, de la variable de salida obtenida al entrenar con eventos con 0-jets o 1-jet.

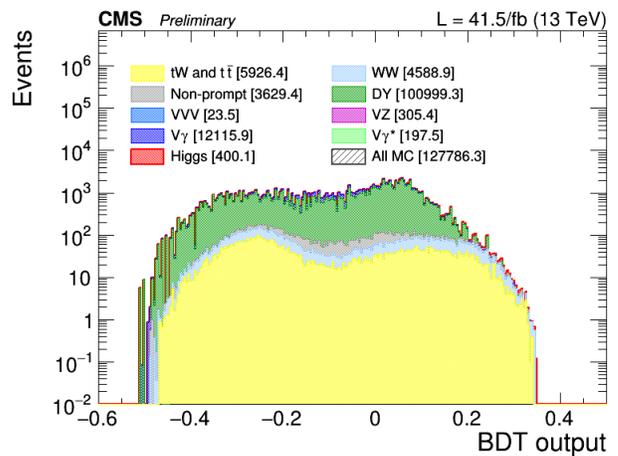


Figura 35: Distribución en escala logarítmica, para eventos con 1-jet reconstruido, de la variable de salida obtenida al entrenar con eventos con 1-jet.

La mejor representación para evaluar el rendimiento es la curva ROC, que se construye seleccionando los sucesos que pasan o no una determinada selección, y estos se dividen

entre la integral total (de menos infinito a infinito) para así extraer la eficiencia de señal y de rechazo de fondo que se tiene para cada punto posible de selección. Son muy útiles en casos como el de este análisis, en los que la forma de las distribuciones de señal y de fondo es muy similar, de modo que el punto de selección no es evidente.

En las Figuras 36 y 37 se representan, para el conjunto de eventos con 0-jets y el conjunto de eventos con 1-jet respectivamente, las curvas ROC construidas en base a las distribuciones mostradas anteriormente (Figuras 32-35). Para su obtención se ha tomado como fondo no solo el Drell-Yan sino todos los procesos físicos descritos en la Subsección 5.2.

Además, en las Figuras 38 y 39 se muestra una comparación análoga mediante el uso de las curvas de significancia de señal, donde para cada punto de corte en la variable en cuestión se calcula la significancia como indica la ecuación 6.

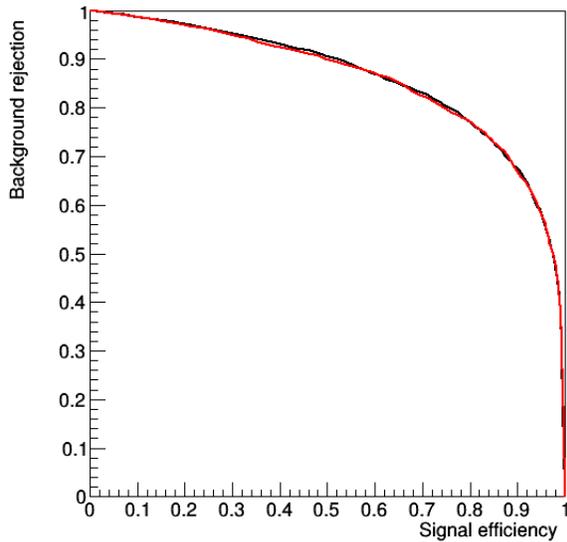


Figura 36: *Curvas ROC asociadas a la variable de salida del BDT entrenado con eventos con 0-jets y 1-jet reconstruidos (negro) y del entrenado exclusivamente con eventos con 0-jets (rojo). Aplicación sobre eventos con 0-jet reconstruidos.*

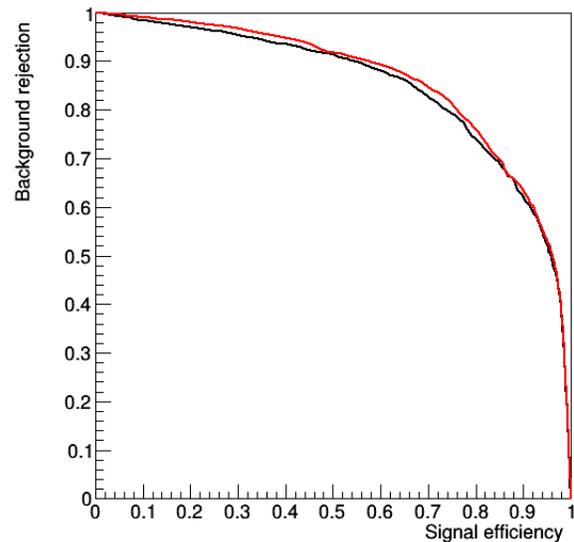


Figura 37: *Curvas ROC asociadas a la variable de salida del BDT entrenado con eventos con 0-jets o 1-jet reconstruidos (negro) y del entrenado exclusivamente con eventos con 1-jet (rojo). Aplicación sobre eventos con 1-jet reconstruido.*

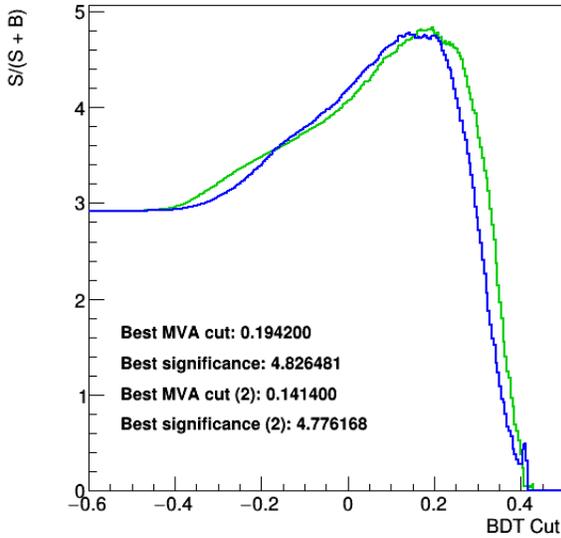


Figura 38: Curvas de significancia de señal para la variable de salida del BDT entrenado con eventos con 0-jets o 1-jet reconstruidos (verde) y del entrenamiento exclusivamente con eventos con 0-jets (azul). Aplicación sobre eventos con 0-jet reconstruidos.

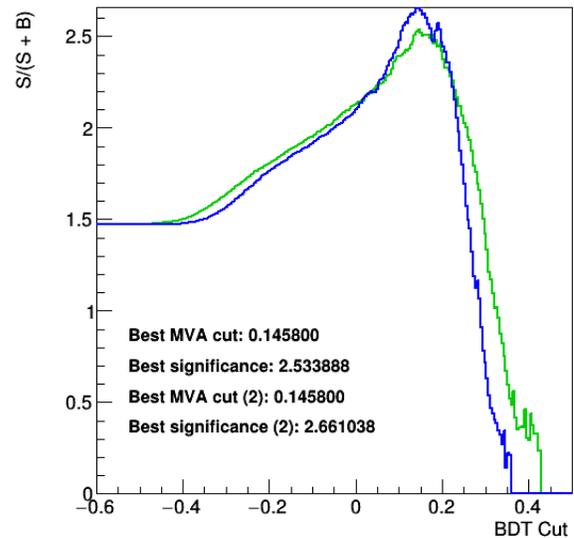


Figura 39: Curvas de significancia de señal para la variable de salida del BDT entrenado con eventos con 0-jets o 1-jet reconstruidos (verde) y del entrenamiento exclusivamente con eventos con 1-jet (azul). Aplicación sobre eventos con 1-jet reconstruido.

En cuanto al estudio sobre eventos con 0-jets reconstruidos, se ha obtenido una significancia de señal máxima $Sig_{0j+1j}^{0j} = 4.83$ entrenando conjuntamente con eventos con 0-jets y 1-jet, así como una significancia máxima $Sig_{0j}^{0j} = 4.78$ entrenando únicamente con eventos con 0-jets. Por su parte, en el estudio sobre eventos con 1-jet reconstruido, se ha obtenido una significancia de señal máxima $Sig_{0j+1j}^{1j} = 2.53$ entrenando con eventos con 0-jets y 1-jet en conjunto, así como una significancia máxima $Sig_{1j}^1 = 2.66$ entrenando exclusivamente con eventos con 1-jet.

Los resultados indican que el entrenamiento con eventos inclusivos en jets proporciona aproximadamente el mismo rendimiento que entrenando individualmente con eventos con 0-jets y 1-jet. La mayor discrepancia se observa en el estudio sobre eventos con 1-jet, donde el entrenamiento individual presenta un aumento de significancia del 5% respecto al entrenamiento inclusivo. No obstante, como el incremento no es muy grande, se ha decidido por simplicidad considerar en la continuación del trabajo entrenamientos con todos los eventos que posean 0-jets y 1-jet reconstruidos.

8.3. Rendimiento de los diferentes algoritmos de entrenamiento

Se han estudiado los diferentes algoritmos explicados en la Sección 6 con el fin de seleccionar el que presenta mayor rendimiento, es decir, el que mejor relación presenta entre rechazo de fondo y eficiencia de señal.

NTrees	MaxDepth	MinNodeSize	nCuts	AdaBoostBeta
800	3	0.5%	20	0.5

Tabla 5: Configuración inicial de parámetros.

En las Figuras 40 y 41 se muestran, respectivamente, las variables de salida junto con la evaluación de sobreentrenamiento y las curvas ROC generadas al entrenar el BDT con los algoritmos *AdaBoost* (negro), *Bagging* (rojo) y *Random Forest* (verde), empleando las mismas variables de entrada y como valores iniciales de los parámetros definidos en la Sección 6 los que se muestran en la Tabla 5.

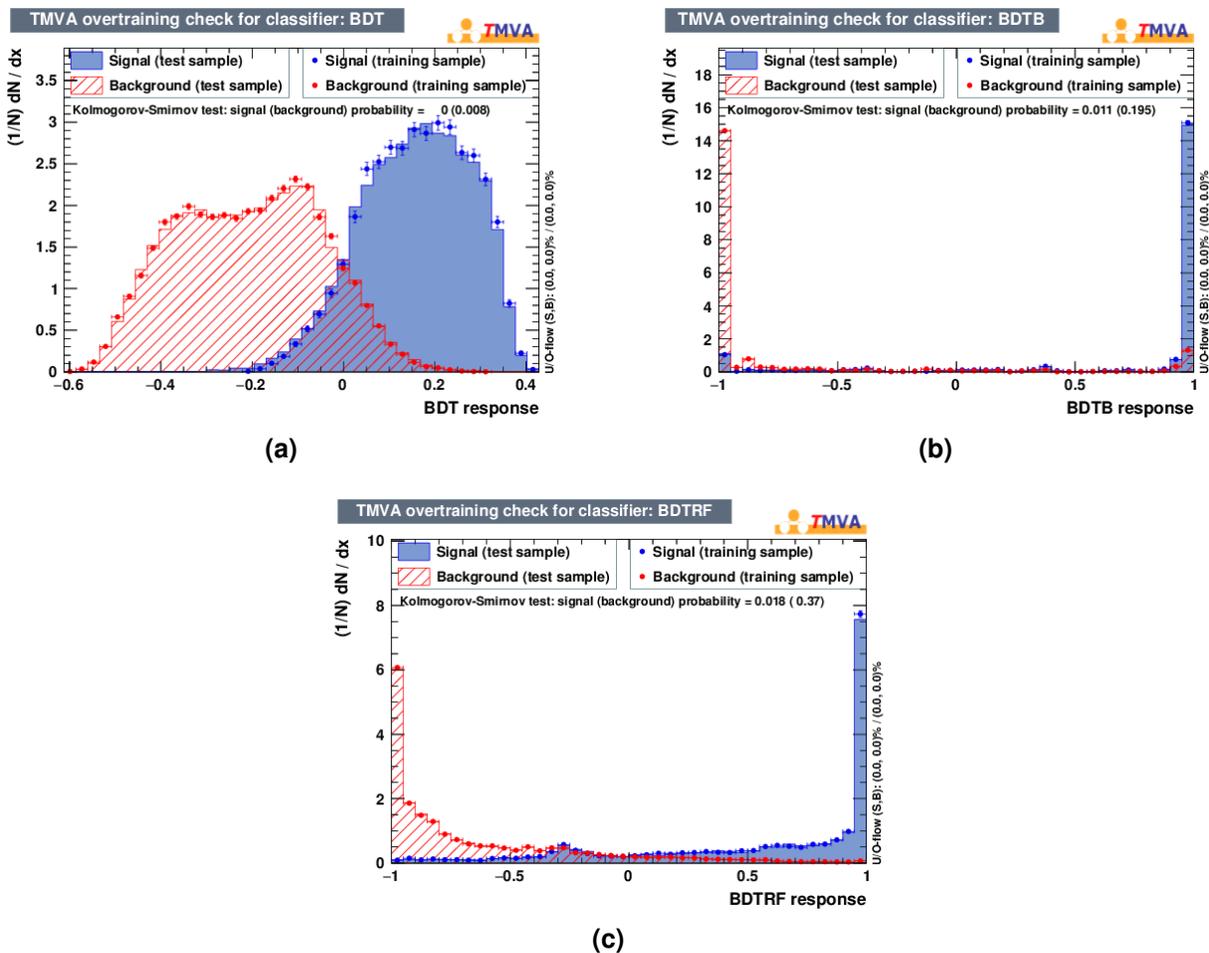


Figura 40: Distribución de la variable de salida y evaluación del sobreentrenamiento para el BDT entrenado con el algoritmo (a) *AdaBoost*, (b) *Bagging* y (c) *Random Forest*. En rojo se muestra el fondo y en azul la señal. Los rellenos rayado y liso representan la muestra de prueba mientras que los círculos representan la muestra de entrenamiento.

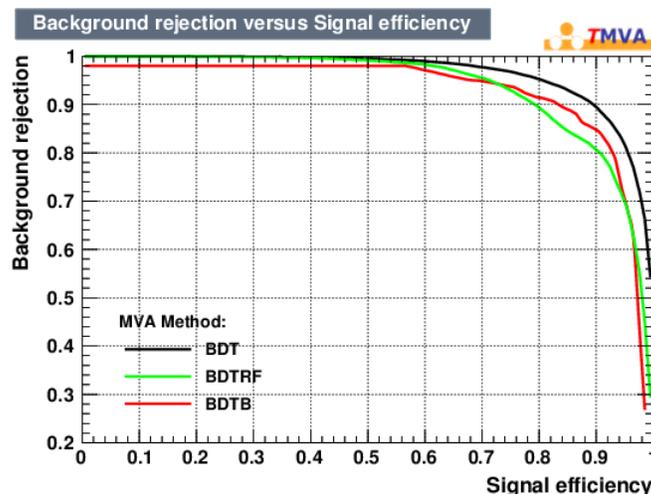


Figura 41: Curvas ROC - Rechazo de fondo respecto de eficiencia de señal. En negro se muestra la correspondiente al algoritmo *AdaBoost*, en rojo la correspondiente a *Bagging*, y en verde *Random Forest*.

Como se observa en la Figura 41, el mejor rendimiento lo proporciona el algoritmo *AdaBoost*, ya que es el que mejor acuerdo entre rechazo de fondo y eficiencia de señal presenta. Con el fin de mejorar el rendimiento de este algoritmo, se ha procedido a realizar una optimización de los principales hiperparámetros del modelo, como se muestra en la siguiente subsección.

8.4. Optimización de hiperparámetros del modelo

Partiendo del algoritmo *AdaBoost*, se ha realizado una optimización de los hiperparámetros y se ha estudiado su rendimiento en términos de eficiencia y significancia. Para ello, se han configurado varios BDTs utilizando distintos valores de los siguientes hiperparámetros que han sido definidos en la Subsección 6.2: *NTrees*, *MaxDepth*, *MinNodeSize*, *nCuts* y *AdaBoostBeta*. Mientras se ha variado el valor de un parámetro concreto, se han mantenido, para el resto de parámetros, los valores iniciales del BDT que se recogen en la Tabla 5.

En las Tablas 6, 7, 8, 9 y 10 se recogen los diversos valores seleccionados respectivamente para *NTrees*, *MaxDepth*, *MinNodeSize*, *nCuts* y *AdaBoostBeta*. Así mismo, las Figuras 42, 43, 44, 45 y 46 muestran las curvas ROC asociadas a estas variaciones.

NTrees
250
500
800
1000
2000

Tabla 6: Valores estudiados para el parámetro NTrees.

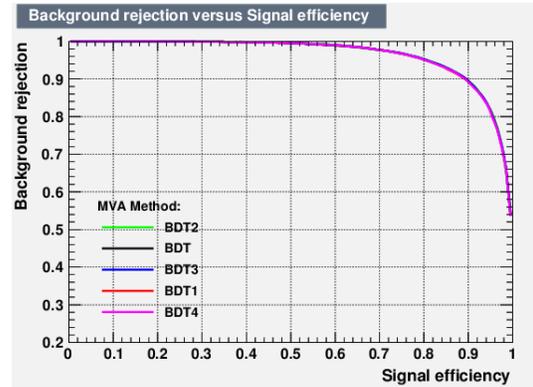


Figura 42: Curvas ROC. El rojo corresponde a 250 árboles, el verde a 500, el negro a 800, el azul a 1000, y el rosa a 2000.

MaxDepth
3
4
5

Tabla 7: Valores estudiados para el parámetro MaxDepth.

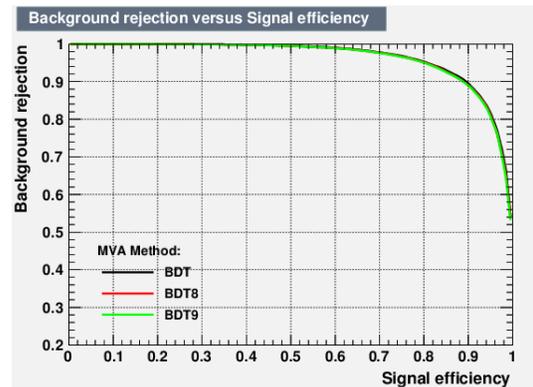


Figura 43: Curvas ROC. El negro corresponde a una profundidad del árbol de 3, el rojo a una de 4, y el verde a una de 5.

MinNodeSize
0.25 %
0.50 %
2.50 %
5.00 %

Tabla 8: Valores estudiados para el parámetro MinNodeSize.

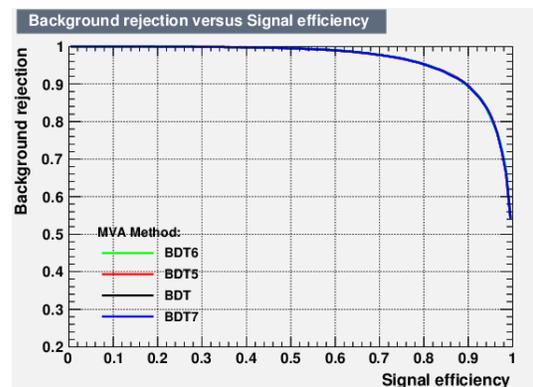


Figura 44: Curvas ROC. El rojo corresponde a un tamaño mínimo de muestra en un nodo del 0.25%, el verde a un 2.50%, el negro a un 0.50%, y el azul a un 5.00%.

nCuts
20
100
1000
10000

Tabla 9: Valores estudiados para el parámetro nCuts.

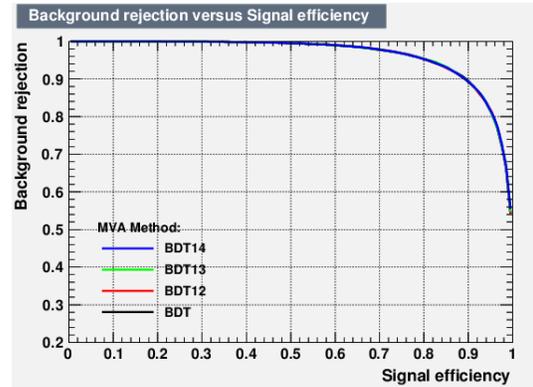


Figura 45: Curvas ROC. El negro corresponde a un número de cortes de 20, el rojo a 100, el verde a 1000, y el azul a 10000.

AdaBoostBeta
0.2
0.5
0.8

Tabla 10: Valores estudiados para el parámetro AdaBoostBeta.

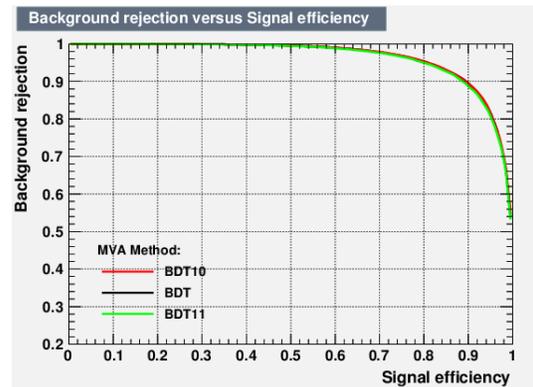


Figura 46: Curvas ROC. El rojo corresponde a un ritmo de aprendizaje de 0.2, el negro a un ritmo de 0.5, y el verde a un ritmo de 0.8.

A partir de las curvas ROC de las Figuras 42, 43, 44, 45 y 46, se puede seleccionar una configuración optimizada tomando para cada parámetro el valor que mejor rendimiento presenta. Con este estudio, se ha observado que el algoritmo proporciona resultados muy sólidos y el cambio de rendimiento al variar los valores de los hiperparámetros es muy sutil. En la Tabla 11 se comparan los valores que corresponden a la configuración inicial y los que corresponderían a una configuración optimizada.

	Configuración inicial	Configuración optimizada
NTrees	800	1000
MaxDepth	3	3
MinNodeSize	0.5 %	0.5 %
nCuts	20	20
AdaBoostBeta	0.5	0.2

Tabla 11: Configuración inicial y configuración optimizada de parámetros.

8.5. Evaluación de los resultados

Finalmente, entrenando un nuevo BDT con las configuraciones mostradas en la Tabla 11, se ha obtenido que la optimizada presenta mejor rendimiento que la inicial. En la Figura 47 se muestran las curvas ROC asociadas.

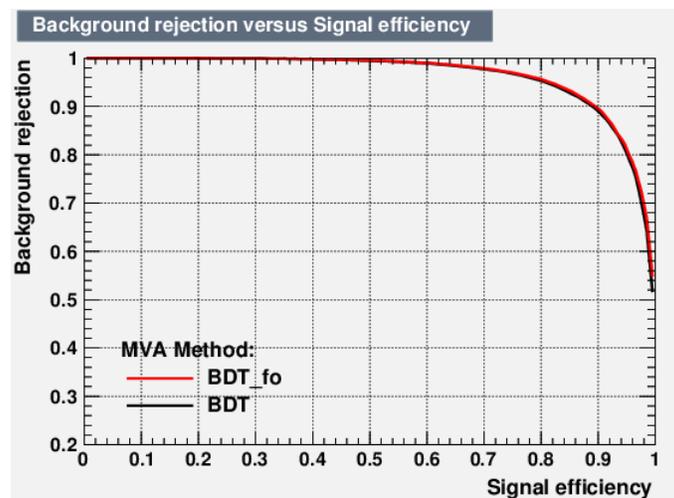


Figura 47: Curvas ROC. En negro se muestra la configuración inicial y en rojo la optimizada.

En las Figuras 48 y 49 se observa para cada una de las configuraciones la distribución de la variable de salida generada con el BDT para la señal y el fondo considerados en el conjunto de datos de entrenamiento. Así mismo, se presenta una evaluación del sobreentrenamiento, que ocurre cuando el modelo se especializa en el conjunto de datos de entrenamiento, fallando al ser aplicado sobre un conjunto de datos independiente. Puesto que las distribuciones de la variable de salida correspondiente a la muestra de entrenamiento y la correspondiente a la muestra de prueba tienen un comportamiento muy similar, se concluye que el modelo no sobreentrena. En caso contrario, el rendimiento de la muestra de entrenamiento sería superior al de la muestra de prueba.

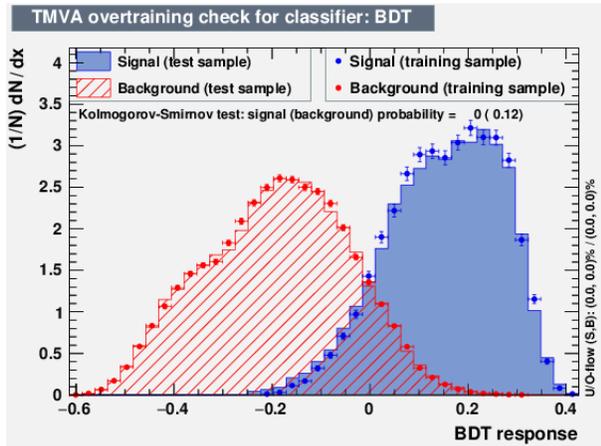


Figura 48: Distribución de la variable de salida correspondiente a la configuración inicial del BDT y evaluación del sobreentrenamiento. En rojo se muestra el fondo y en azul la señal. Los rellenos rayado y liso representan la muestra de prueba mientras que los círculos representan la muestra de entrenamiento.

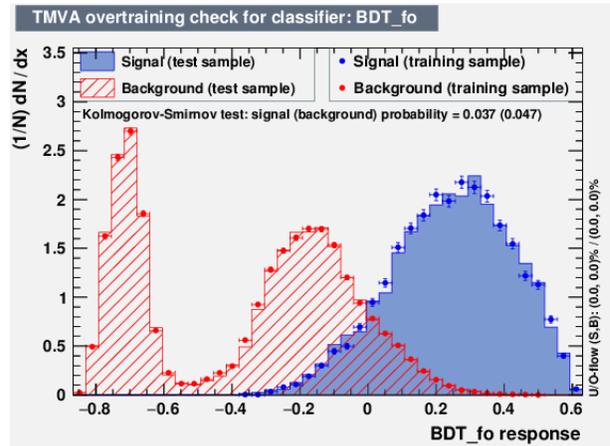


Figura 49: Distribución de la variable de salida correspondiente a la configuración optimizada del BDT y evaluación del sobreentrenamiento. En rojo se muestra el fondo y en azul la señal. Los rellenos rayado y liso representan la muestra de prueba mientras que los círculos representan la muestra de entrenamiento.

Estrechamente relacionadas con las distribuciones de las Figuras 48 y 49, se representan en las Figuras 50 y 51 las curvas de significancia de señal, de eficiencia de fondo y de eficiencia de señal para cada una de las configuraciones.

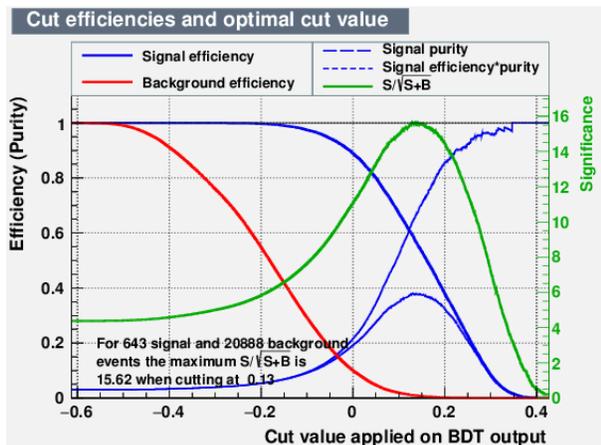


Figura 50: Curvas de significancia de señal (verde), eficiencia de señal (azul) y eficiencia de fondo (rojo) con la configuración inicial.

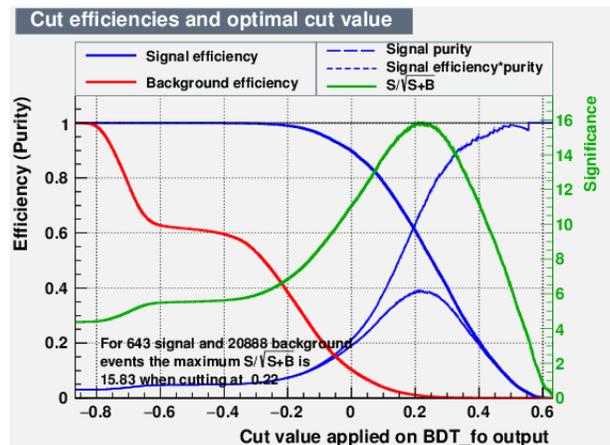


Figura 51: Curvas de significancia de señal (verde), eficiencia de señal (azul) y eficiencia de fondo (rojo) con la configuración optimizada.

La configuración optimizada muestra mayor rendimiento que la configuración inicial,

así como una mayor significancia de señal. Para la configuración inicial, se alcanza una significancia máxima $S_{inicial} = 15.62$ para el valor 0.13 de la variable de salida, mientras que para la configuración optimizada, se logra una significancia máxima $S_{optimizada} = 15.83$ seleccionando en el valor 0.22 de la variable de salida.

A continuación se realiza una evaluación en mayor detalle de los resultados obtenidos tras el entrenamiento y testeo del BDT con la configuración optimizada del algoritmo *AdaBoost*.

8.5.1. Árboles de decisión

En la Figura 52, se muestra como ejemplo el crecimiento de uno de los mil árboles del entrenamiento del BDT. La gama de colores empleada para cada uno de los nodos se debe a la pureza de señal presente en cada uno de ellos; tiende al rojo cuando la pureza se acerca a cero, es decir, cuando el nodo está dominado por eventos de fondo, y tiene al azul cuando la pureza se acerca a uno, esto es, cuando predominan los eventos de señal.

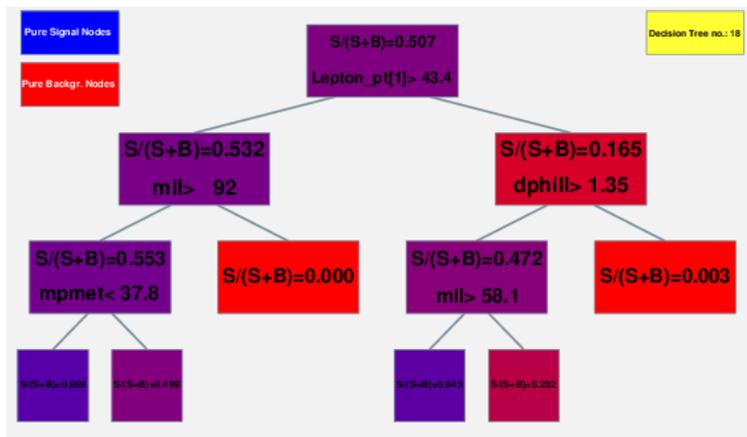


Figura 52: Crecimiento de un árbol de decisión durante en el entrenamiento del BDT.

En este árbol concreto, las variables seleccionadas por dar la mejor separación entre señal y fondo en los diferentes nodos han sido p_T^{l2} , m^l , $\Delta\phi_{ll}$, repitiéndose la masa del sistema dileptónico en tres de ellos.

8.5.2. Matrices de Correlación

En las Figuras 53 y 54 se muestran las correlaciones entre las variables de entrada para el proceso de señal y de fondo respectivamente. Las correlaciones lineales entre variables se determinan por medio del coeficiente de Pearson [35].

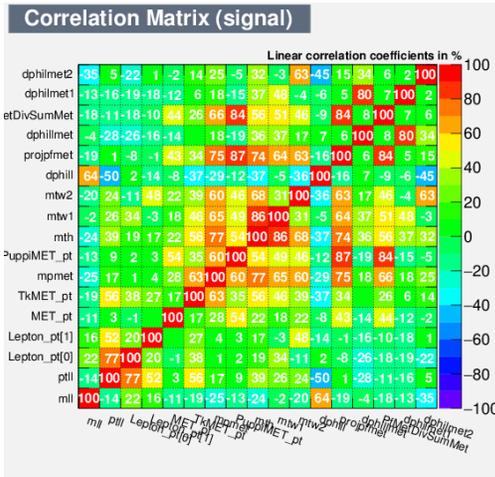


Figura 53: Matriz de correlación entre variables de entrada en el proceso de señal.

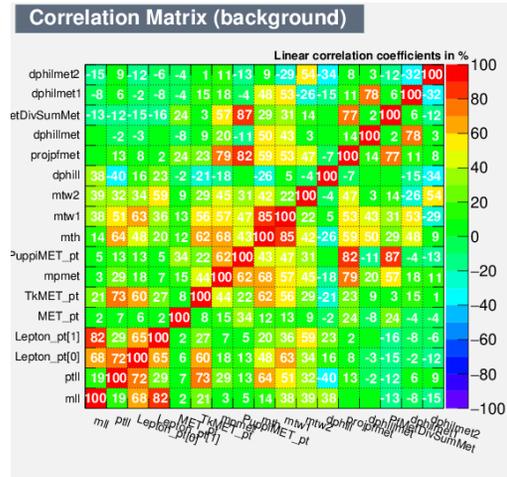


Figura 54: Matriz de correlación entre variables de entrada en el proceso de fondo.

Como se observa, hay algunas variables de entrada fuertemente correlacionadas, tanto para señal como para fondo. En este análisis, ambas matrices presentan correlaciones similares entre variables. No obstante, se muestra una correlación mayor entre las variables relacionadas con el momento transverso de los leptones y la masa invariante del sistema dileptónico para el proceso de fondo, mientras que la correlación entre variables relacionadas con la MET y variables de alto nivel es mayor para el proceso de señal.

8.5.3. Ranking de variables

En función del número de veces que las variables de entrada son elegidas como variables discriminatorias en los nodos, se genera un ranking de importancia de las variables. El ranking correspondiente al entrenamiento del BDT con la configuración optimizada se indica en la Tabla 12.

Ranking	
Variable	Importancia
m_{ll}	0.1852
$PfMetDivSumMet$	0.1064
d_{phill}	0.0812
$Lepton_pt[1]$	0.0720
$mpmet$	0.0679
m_{th}	0.0517
MET_pt	0.0516
$projpfmet$	0.0515
$PuppiMET_pt$	0.0483
$mtw1$	0.0483
$Lepton_pt[0]$	0.0476
$d_{phillmet}$	0.0436
$mtw2$	0.0358
$d_{phillmet1}$	0.0347
$TkMET_pt$	0.0332
$d_{phillmet2}$	0.0231
pt_{ll}	0.0180

Tabla 12: Ranking de variables en el entrenamiento del BDT.

Las variables principales que ocupan las tres primeras posiciones en el ranking de importancia son: m_{ll} , $PfMetDivSumMet$ y $\Delta\phi_{ll}$. Son las más importantes porque presentan un gran poder de separación entre señal $H \rightarrow WW$ y fondo DY. La masa invariante m_{ll} del proceso de señal pica en 125 GeV mientras que la correspondiente al proceso de fondo pica aproximadamente en 91.2 GeV. Por su parte, $PfMetDivSumMet$ tiende a tomar valores más altos en el caso de señal que en el caso de fondo, debido a que el proceso DY no cuenta con la presencia de neutrinos que conforman la MET real. La diferencia en el ángulo azimutal entre los dos leptones, $\Delta\phi_{ll}$, tiende a ser mayor en el proceso de fondo puesto que los dos leptones suelen tener sentidos contrarios, mientras que en el proceso de señal tienden a ser paralelos debido a los neutrinos que adquieren sentido opuesto al momento transversal de los leptones.

8.6. Significancia de señal seleccionando en la respuesta del BDT

Se ha evaluado el modelo entrenado introduciendo la variable de salida en todas las muestras de simulación.

En las Figuras 55 y 56 se muestra la distribución de la variable de salida, para el conjunto de eventos con 0-jets y el conjunto de eventos con 1-jet respectivamente. La selección que ha sido aplicada en las muestras de simulación de todos los procesos de señal y fondo se corresponde con la indicada en la Tabla 2 a excepción de la selección en *puppi_{met}* que se ha relajado a *puppi_{met}* > 20 GeV.

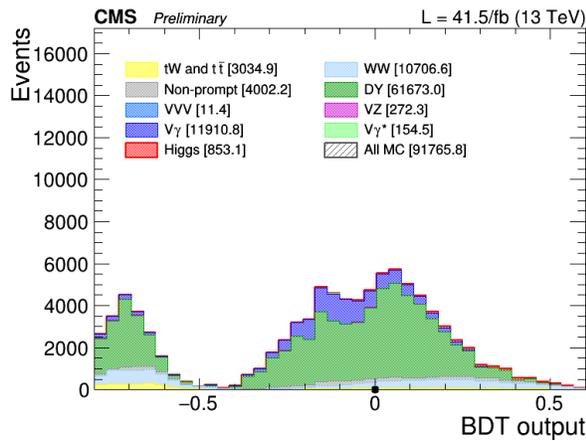


Figura 55: Distribución de la variable de salida para el conjunto de eventos con 0-jets considerando todos los fondos al último nivel de selección.

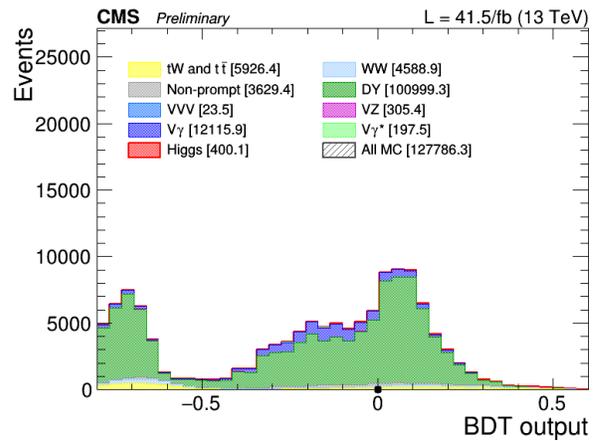


Figura 56: Distribución de la variable de salida para el conjunto de eventos con 1-jet considerando todos los fondos al último nivel de selección.

A partir del último nivel de selección aplicado en el análisis secuencial con la excepción del cambio en la selección de *puppi_{met}*, se ha procedido a seleccionar sobre la variable de salida para obtener una mayor discriminación de fondo. Se ha tomado el punto de selección óptimo proporcionado por las curvas de significancia de señal asociadas a la variable de salida del BDT en las que se consideran todos los fondos. Dichas curvas se muestran en las Figuras 57 y 58, e indican que el valor de selección que da máxima significancia de señal es aproximadamente 0.28 y 0.27 para el análisis de eventos con 0-jets y 1-jet respectivamente.

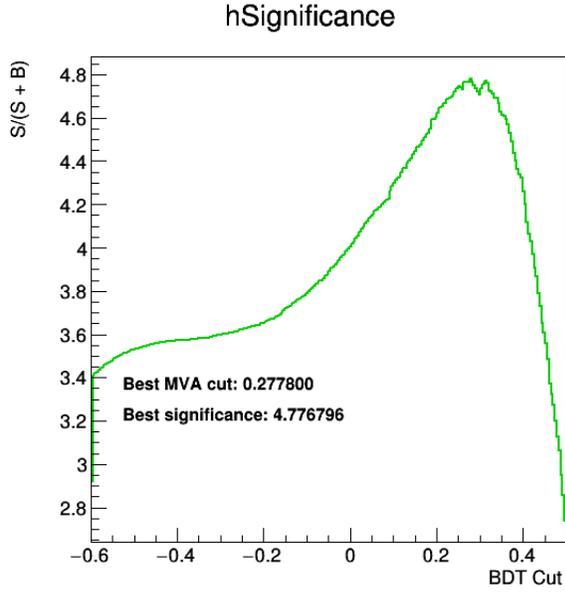


Figura 57: Curva de significancia de señal asociada a la variable de salida del BDT para el conjunto de eventos con 0-jets.

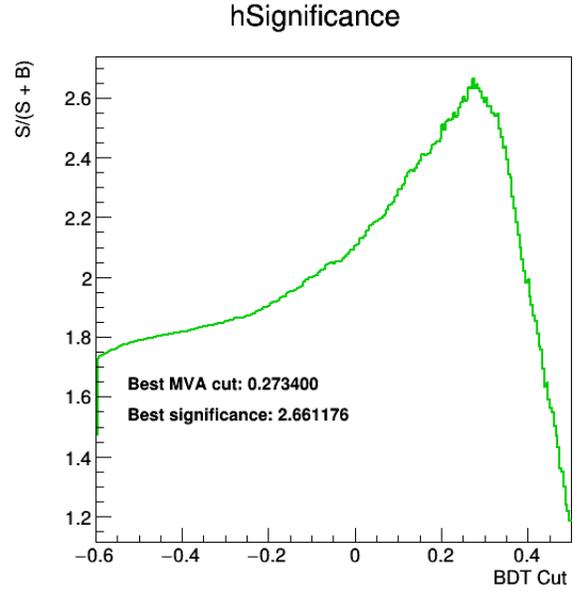


Figura 58: Curva de significancia de señal asociada a la variable de salida del BDT para el conjunto de eventos con 1-jet.

En las Figuras 59 - 64 se representan las distribuciones de seis variables en el canal de 0-jets, tras añadir a los ya aplicados en la Sección 7 con el cambio en *puppimet* un nuevo nivel de selección sobre la variable de salida del BDT. Análogamente, en las Figuras 65 - 70 se representan las mismas distribuciones en el canal de 1-jet.

0-jets

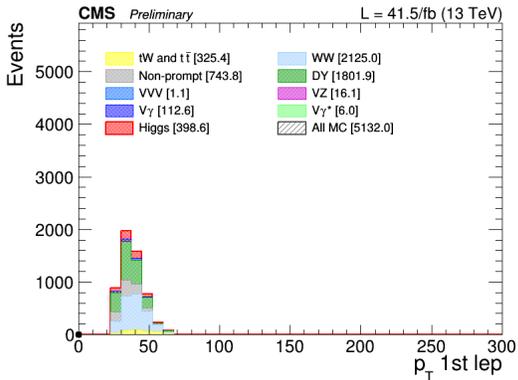


Figura 59: Distribución del momento transverso del leptón más energético $p_T(l_1)$ en el análisis de eventos con 0-jets.

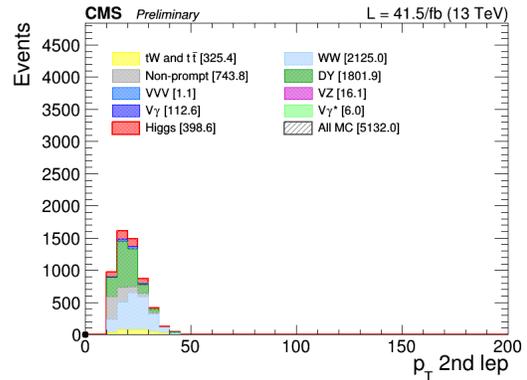


Figura 60: Distribución del momento transverso del segundo leptón $p_T(l_2)$ en el análisis de eventos con 0-jets.

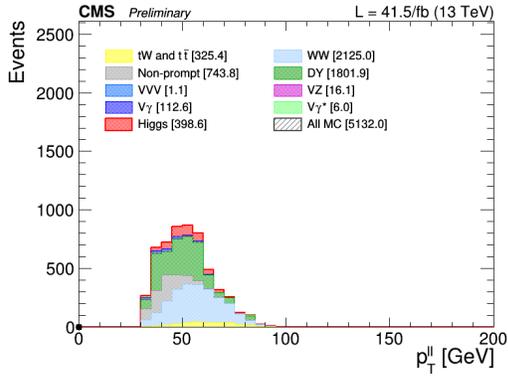


Figura 61: Distribución del momento transverso del sistema dileptónico p_T^{ll} en el análisis de eventos con 0-jets.

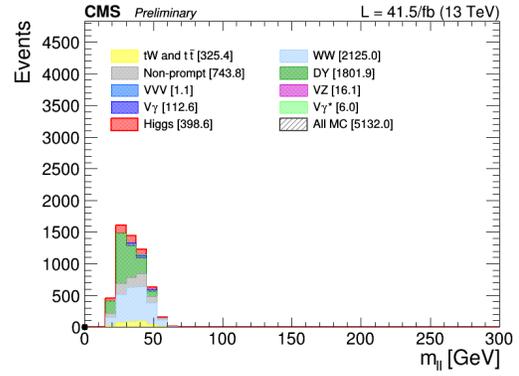


Figura 62: Distribución de la masa invariante del sistema dileptónico m_{ll} en el análisis de eventos con 0-jets.

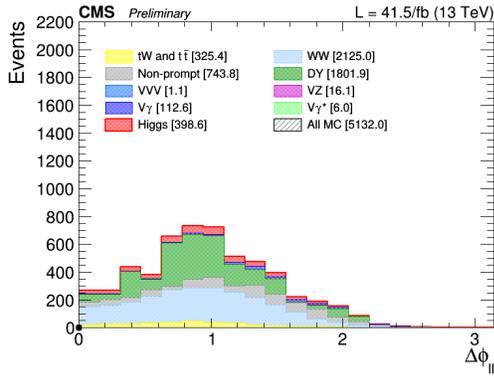


Figura 63: Distribución de la diferencia angular entre los dos leptones $\Delta\phi_{ll}$ en el análisis de eventos con 0-jets.

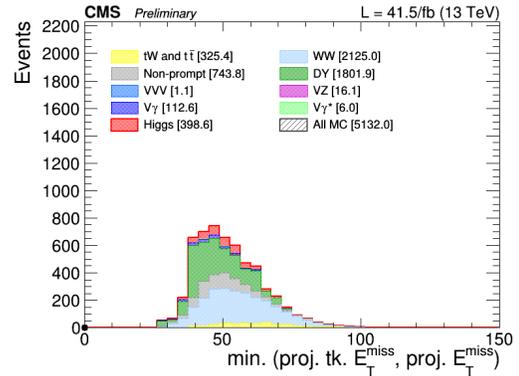


Figura 64: Distribución de la energía perdida transversa m_{pmet} en el análisis de eventos con 0-jets.

1-jet

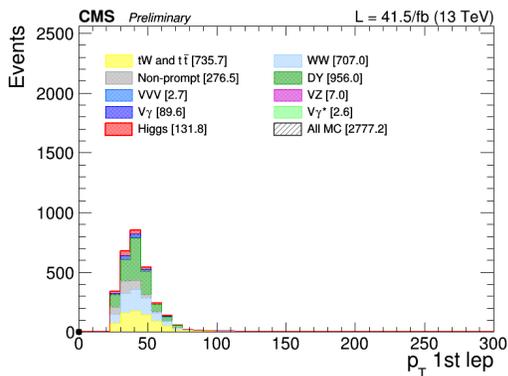


Figura 65: Distribución del momento transverso del leptón más energético $p_T(l_1)$ en el análisis de eventos con 1-jet.

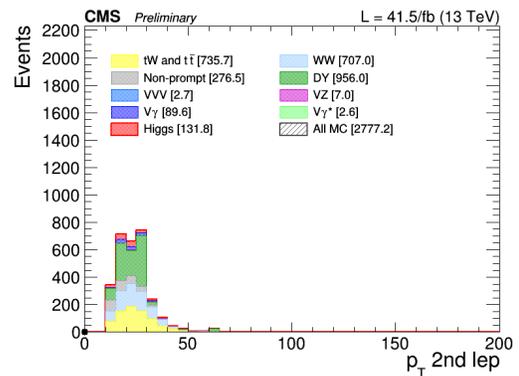


Figura 66: Distribución del momento transverso del segundo leptón $p_T(l_2)$ en el análisis de eventos con 1-jet.

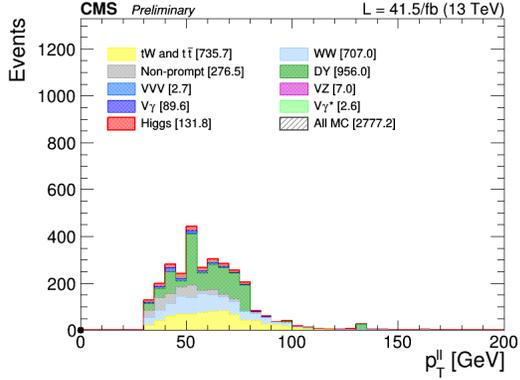


Figura 67: Distribución del momento transverso del sistema dileptónico p_T^{ll} en el análisis de eventos con 1-jet.

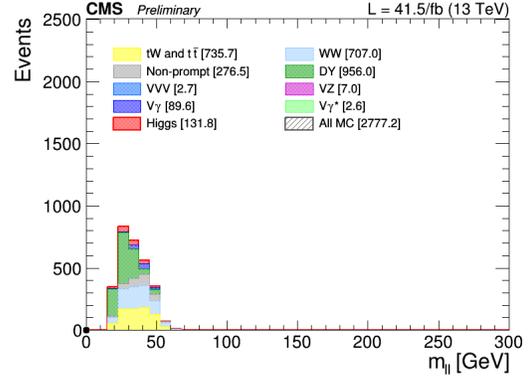


Figura 68: Distribución de la masa invariante del sistema dileptónico m_{ll} en el análisis de eventos con 1-jet.

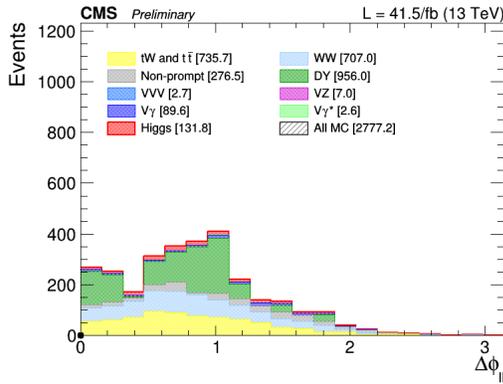


Figura 69: Distribución de la diferencia angular entre los dos leptones $\Delta\phi_{ll}$ en el análisis de eventos con 1-jet.

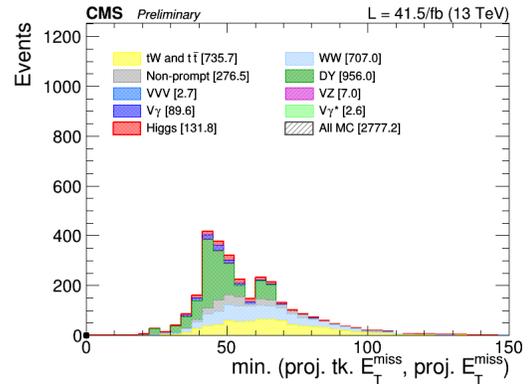


Figura 70: Distribución de la energía perdida transversa m_{pmet} en el análisis de eventos con 1-jet.

En las Tablas 13 y 14, correspondientes al conjunto de eventos con 0-jets y 1-jet respectivamente, se indica el número de eventos de cada fondo y de señal junto con la significancia de señal, en el último nivel de la selección llevada a cabo en la Sección 7 y tras seleccionar en la variable de salida del BDT.

Selección	S	DY	WW	Top	Non-prompt	VV(V)	$V\gamma^{(*)}$	B	Sig
bVeto Deep CSV - L	379.5	7098.2	6447.0	2323.9	1609.8	164.0	1237.1	18880.0	2.74
BDT output > 0.28	398.6	1801.9	2125.0	325.4	743.8	17.2	118.6	5132.0	5.36

Tabla 13: Número de eventos de los procesos de señal y fondo tras aplicar el último nivel de selección de las Sección 7 (con el cambio en p_{pmet}) y la selección en la variable de salida proporcionada por el BDT para el conjunto de eventos con 0-jets. S representa los eventos de señal y B la suma de eventos de los distintos procesos de fondo; Sig es la significancia de señal.

Selección	S	DY	WW	Top	Non-prompt	VV(V)	$V_{\gamma}^{(*)}$	B	Sig
bVeto Deep CSV - L	207.3	10914.7	3256.0	4487.3	1412.4	116.5	1647.3	21834.1	1.40
BDT output > 0.27	131.8	956.0	707.0	735.7	276.5	9.7	92.2	2777.1	2.44

Tabla 14: Número de eventos de los procesos de señal y fondo tras aplicar el último nivel de selección de las Sección 7 (con el cambio en $p_{\text{up}}^{\text{pimet}}$) y la selección en la variable de salida proporcionada por el BDT para el conjunto de eventos con 1-jet. S representa los eventos de señal y B la suma de eventos de los distintos procesos de fondo; Sig es la significancia de señal.

Tras aplicar la selección sobre la variable de salida del BDT entrenado con el algoritmo *AdaBoost* y la configuración optimizada de hiperparámetros, se ha obtenido una región enriquecida en señal $H \rightarrow WW$.

Las significancias obtenidas tras seleccionar en la variable de salida del BDT son $Sig_{\text{multivariable}}^{0j} = 5.36$ y $Sig_{\text{multivariable}}^{1j} = 2.44$ para el canal de 0-jets y 1-jet respectivamente. Por tanto, la significancia de señal alcanzada mediante el análisis multivariable es considerablemente mayor a la obtenida por medio del análisis secuencial de la Sección 7.

Como se aprecia en las Tablas 13 y 14, el canal de 0-jets sigue siendo más puro en señal que el canal de 1-jet. Seleccionando sobre la variable de salida del BDT, el fondo predominante para eventos con 0-jets reconstruidos es el proceso WW seguido por el proceso DY, mientras que en el caso de eventos con 1-jet reconstruido, el fondo predominante es el proceso DY seguido por el proceso Top.

9. CONCLUSIONES

Se han realizado dos análisis en paralelo para muestras de Monte Carlo que simulan la toma de datos de 2017 por el experimento CMS del LHC durante el Run II a una energía de centro de masas de $\sqrt{s} = 13$ TeV y a una luminosidad integrada de 41.5 fb^{-1} . El objetivo principal ha sido encontrar la mejor separación entre el proceso de señal $H \rightarrow WW$ y los procesos de fondo, principalmente Drell-Yan.

El primer análisis ha consistido en aplicar en serie una secuencia de selecciones sobre diversas variables discriminatorias que describen propiedades de los eventos. Por su parte, el segundo análisis ha permitido combinar mediante un BDT el conjunto de variables de entrada en un único discriminante que actúa como clasificador.

Se ha comprobado que el BDT es un método multivariable eficaz para la clasificación de eventos en física de partículas. Se ha probado que un único BDT entrenado conjuntamente con eventos que poseen 0-jets y 1-jet proporciona una significancia de señal similar a dos BDTs entrenados individualmente usando eventos con 0-jets o 1-jet, lo cual ha permitido simplificar el análisis llevado a cabo.

A partir de una configuración de parámetros inicial del BDT se ha estudiado el rendimiento de diferentes algoritmos, obteniéndose un mejor resultado con el algoritmo *AdaBoost*. Para este algoritmo se ha optimizado la configuración de hiperparámetros con el objetivo de obtener el mayor rendimiento posible, aunque los resultados que proporcionan estos algoritmos clasificadores son bastante sólidos y la mejora tras la optimización no resulta demasiado significativa.

Finalmente, los resultados presentados en las Secciones 7 y 8 indican que con el análisis multivariable se alcanza una región más pura en señal. El aumento de significancia en este análisis con respecto al análisis secuencial es considerable y se muestra en las Tablas 13 y 14 separadamente para eventos con 0-jets y 1-jet reconstruidos. Se concluye por tanto que el análisis más adecuado de cara a tener una región más pura en señal para realizar a posteriori estudios más avanzados es el multivariable.

El siguiente paso tras la fase de ejecución desarrollada en este trabajo, en la que se ha entrenado y probado el BDT con la configuración optimizada sobre simulación de Monte Carlo, sería proceder con la aplicación del modelo entrenado y probado sobre datos reales.

REFERENCIAS

- [1] The CMS Collaboration, "Measurements of properties of the Higgs boson decaying to a W boson pair in pp collisions at $\sqrt{s} = 13$ TeV", CMS-HIG-16-042, CERN-EP-2018-141. Phys. Lett. B 791 (2019) 96, [doi: 10.1016/j.physletb.2018.12.073]
arXiv:1806.05246v2.
- [2] M. Thomson, "Modern Particle Physics", Cambridge University Press, United Kingdom 2016.
ISBN: 978-1-107-03426-6.
- [3] "CERN site", <https://home.cern/science/physics/higgs-boson>
- [4] "CMS Experiment - CERN site", <https://cms.cern/physics/higgs-boson>
- [5] The ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", CERN-PH-EP-2012-218. Phys. Lett. B 716 (2012) 1-29, [doi: 10.1016/j.physletb.2012.08.020].
arXiv:1207.7214v2.
- [6] The CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", CMS-HIG-12-028; CERN-PH-EP-2012-220. Phys. Lett. B 716 (2012) 30, [doi: 10.1016/j.physletb.2012.08.021].
arXiv:1207.7235v2.
- [7] M. Carena, C. Grojean, M. Kado, and V.Sharma, "Status of Higgs Boson Physics", Septiembre 2017. M. Tanabashi et al "Particle Data Group", Phys. Rev. D 98, 030001 (2018).
<http://pdg.lbl.gov/2019/reviews/rpp2018-rev-higgs-boson.pdf>
<https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWG>
- [8] "CERN site", <https://home.cern/science/accelerators>
- [9] "CERN site", <https://home.cern/science/accelerators/large-hadron-collider>
- [10] "CERN Document Server", <https://cds.cern.ch/record/2197559>
- [11] "CMS Experiment - CERN site", <https://cms.cern/detector>

- [12] The CMS Collaboration, "The CMS experiment at the CERN LHC". Ginebra, 2010.
JINST 3 (2008) S08004, [doi: 10.1088/1748-0221/3/08/S08004]
<https://cds.cern.ch/record/1129810?ln=es>
- [13] The CMS Collaboration, "CMS Physics: Technical Design Report Volume 1: Detector Performance and Software", CERN-LHCC-2006-001; CMS-TDR-8-1. Ginebra, 2006.
<http://cds.cern.ch/record/922757/>
- [14] The CMS Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus and MET", CMS-PAS-PFT-09-001. CERN, 2009.
<https://cds.cern.ch/record/1194487?ln=es>
- [15] "CERN Document Server", https://cds.cern.ch/record/2270046/files/Figure_001.png
- [16] The CMS Collaboration, "Performance of CMS muon reconstruction in pp collision events at 7 TeV", CMS-MUO-10-004; CERN-PH-EP-2012-173.
JINST 7 (2012) P10002 [doi: 10.1088/1748-0221/7/10/P10002]
arXiv:1206.4071v2
- [17] The CMS Collaboration, "Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at 8 TeV", CMS-EGM-13-001, CERN-PH-EP-2015-004.
JINST 7 (2012) P10002 [doi: 10.1088/1748-0221/10/06/P06005]
arXiv:1502.02701v2
- [18] M. Cacciari, G. P. Salam and G. Soyez, "The anti-kT jet clustering algorithm", LPTHE-07-03, Febrero 2008.
JHEP 0804:063,2008 [doi: 10.1088/1126-6708/2008/04/063]
arXiv:0802.1189
- [19] The CMS Collaboration, "Identification of b-quark jets with the CMS experiment", CMS-BTV-12-001; CERN-PH-EP-2012-262.
JINST 8 (2013) P04013 [doi: 10.1088/1748-0221/8/04/P04013]
arXiv:1211.4462v2
- [20] The CMS Collaboration, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV", CMS-BTV-16-002, CERN-EP-2017-326.
JINST 13 (2018) P05011 [doi: 10.1088/1748-0221/13/05/P05011]
arXiv:1712.07158v2

- [21] The CMS Collaboration, "Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13$ TeV using the CMS detector", CMS-JME-17-001, CERN-EP-2018-335.
JINST 14 (2019) P07004 [doi: 10.1088/1748-0221/14/07/P07004]
arXiv:1903.06078v2
- [22] M. Tanabashi et al "Particle Data Group", Phys. Rev. D 98, 030001 (2018).
<http://pdg.lbl.gov/2019/tables/rpp2019-sum-gauge-higgs-bosons.pdf>
- [23] <https://github.com/latinos/LatinoAnalysis/blob/master/NanoGardener/python/framework/samples/samplesCrossSections2017.py>. Fecha de consulta: Junio 2019.
- [24] S. Alioli, P. Nason, C. Oleari, E. Re, "NLO vector-boson production matched with shower in POWHEG", Mayo 2008.
JHEP0807:060,2008 [doi: 10.1088/1126-6708/2008/07/060]
arXiv:0805.4802
- [25] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, T. Stelzer, "MadGraph 5: Going Beyond", Junio 2011.
JHEP06(2011)128 [doi: 10.1007/JHEP06(2011)128]
arXiv:1106.0522v1
- [26] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, M. Zaro, "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", CERN-PH-TH/2014-064, CP3-14-18, LPN14-066, MCNET-14-09, ZU-TH 14/14, Julio 2014.
JHEP07(2014)079 [doi: 10.1007/JHEP07(2014)079]
arXiv:1405.0301v2
- [27] K. Albertsson, S. Gleyzer, A. Hoecker, L. Moneta, P. Speckmayer, J. Stelzer, J. Theraag, E. von Toerne, H. Voss, S. Wunsch, "TMVA-v4 Users Guide - Toolkit for Multivariate Data Analysis with ROOT", CERN-OPEN-2007-007.
<https://root.cern/tmva>
arXiv:physics/0703039 [Data Analysis, Statistics and Probability]
- [28] E. von Toerne, "Multi-Variate Analysis Methods", Terascale Statistics School, Universität Bonn, febrero 2016.
<https://indico.desy.de/indico/event/13610/contribution/4/material/slides/0.pdf>

- [29] K. Woodruff, "Introduction to boosted decision trees", NM State University, Machine Learning Group Meeting, septiembre 2017.
<https://indico.fnal.gov/event/15356/contribution/1/material/slides/0.pdf>
- [30] A. Bevan, "Multivariate Analysis Techniques Lecture 2", BABAR Analysis School, SLAC National Laboratory, Octubre 2019.
<http://pprc.qmul.ac.uk/bevan/BAS/lecture2.pdf>
- [31] Y. Coadou, "Boosted Decision Trees", CPPM Marseille, ESIPAP, Archamps, febrero 2016.
https://indico.cern.ch/event/472305/contributions/1982360/attachments/1224979/1792797/ESIPAP_MVA160208-BDT.pdf
- [32] C. Böser, S. Fink, S. Röcker, "Introduction to Boosted Decision Trees. A multivariate approach to classification problems", Institut für Experimentelle Kernphysik, KIT.
https://indico.scc.kit.edu/event/48/contributions/3410/attachments/1690/2312/BDT_KSET_A_Freudenstadt.pdf
- [33] The SWAN Service - CERN: Interactive Data Analysis, Machine Learning.
<https://swan.web.cern.ch/content/machine-learning>
- [34] ROOT Data Analysis Framework.
<https://root.cern.ch/>
- [35] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, "Numerical Recipes. The Art of Scientific Computing", Cambridge University Press, United Kingdom 2007.
ISBN: -13 978-0-511-33555-6
- [36] <https://github.com/claralasa/TFG>. Fecha de consulta: Septiembre 2019.

APÉNDICE I : ” EVALUACIÓN DE ENTRENAMIENTO EN TMVA ”

La evaluación del entrenamiento se obtiene mediante una serie de funciones integradas en el paquete TMVA. Algunas de ellas han sido modificadas para adaptarlas al caso concreto de este trabajo. Los cambios implementados pueden ser consultados en github [36].

Las principales funciones que se han empleado en este trabajo son:

- variables.C : representa la distribución de las diferentes variables de input usadas en el MVA, superponiendo la correspondiente a señal y la correspondiente a fondo;
- correlations.C : representa la matriz de correlación de las variables de input usadas en el MVA. Se distingue la relativa a señal de la relativa a fondo;
- mvas.C : representa las distribuciones de la variable de output de los diferentes métodos seleccionados en TMVA, es decir, las distribuciones de las respuestas clasificadores. Se superpone señal y fondo;
- mvaeffs.C : representa las eficiencias de señal y de fondo en función del valor de corte de la variable de output clasificadora. Se muestra también la pureza de señal y la significancia;
- eficiencias.C : representa las diferentes combinaciones relativas a las eficiencias de señal y de fondo. Las eficiencias se obtienen cortando en las variables de output clasificadoras. Es tradicionalmente la mejor representación para evaluar el rendimiento de discriminación a través de las curvas ROC (del inglés *Receiver Operations Characteristics*);
- BDTControlPlots.C : representa la fracción de error y el *weight boost* en función del número de árboles;
- BDT.C : representa los diagramas de crecimiento de los árboles de decisión.

APÉNDICE II : "DISTRIBUCIÓN DE OTRAS VARIABLES"

En las Figuras 71-81 y 82-92 se muestra, a último nivel de la selección mostrada en la Tabla 2, la distribución de las variables E_T^{miss} , $trackE_T^{miss}$, $projectedE_T^{miss}$, $puppiE_T^{miss}$, $\Delta\phi_{ll,E_T^{miss}}$, $\Delta\phi_{l1,E_T^{miss}}$, $\Delta\phi_{l2,E_T^{miss}}$, m_T^H , m_T^{W1} , m_T^{W2} y $PfMetDivSumMet$ para el canal de 0-jets y 1-jet respectivamente.

0-jets

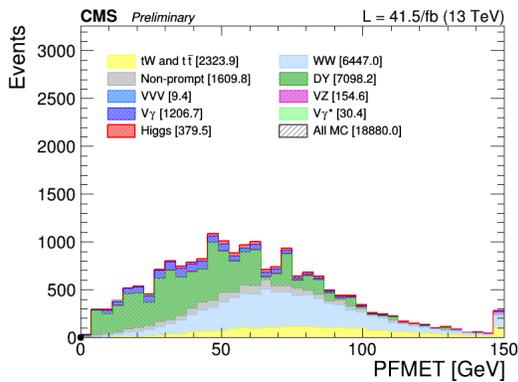


Figura 71: Distribución en el canal de 0-jets de E_T^{miss} .

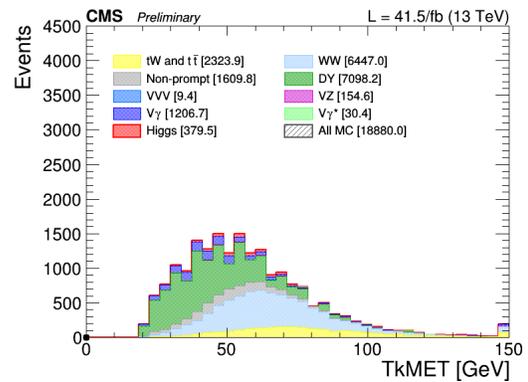


Figura 72: Distribución en el canal de 0-jets de $trackE_T^{miss}$.

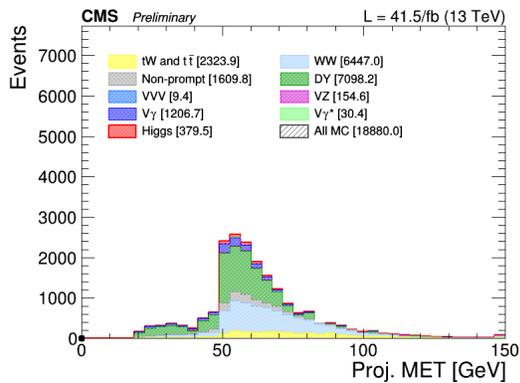


Figura 73: Distribución en el canal de 0-jets de $projectedE_T^{miss}$.

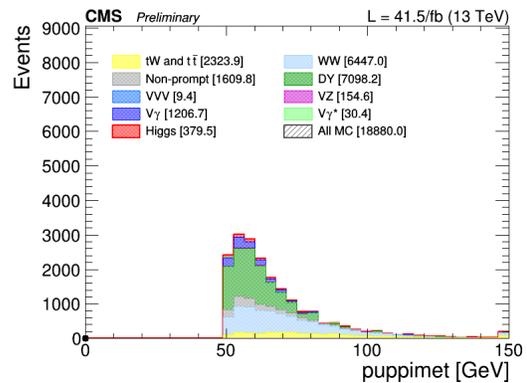


Figura 74: Distribución en el canal de 0-jets de $puppiE_T^{miss}$.

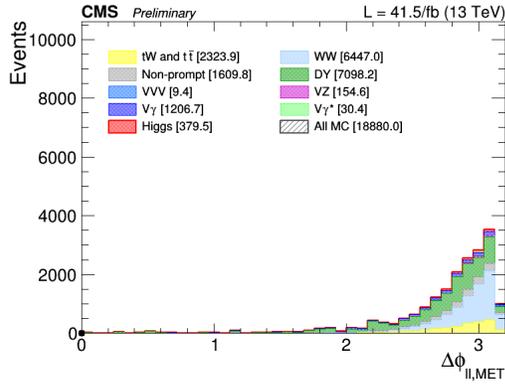


Figura 75: Distribución en el canal de 0-jets de $\Delta\phi_{ll, E_T^{miss}}$.

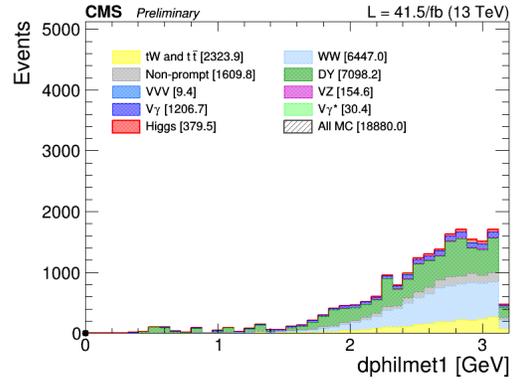


Figura 76: Distribución en el canal de 0-jets de $\Delta\phi_{l1, E_T^{miss}}$.

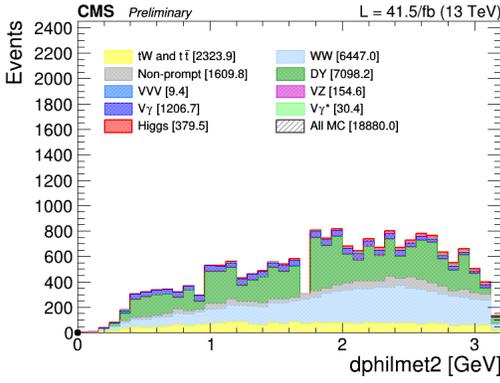


Figura 77: Distribución en el canal de 0-jets de $\Delta\phi_{l2, E_T^{miss}}$.

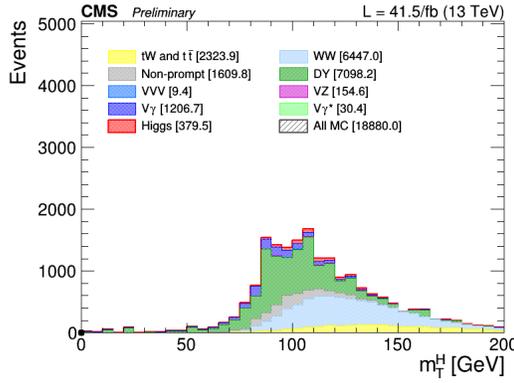


Figura 78: Distribución en el canal de 0-jets de m_T^H .

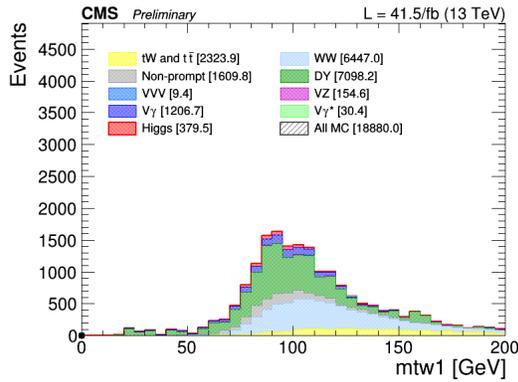


Figura 79: Distribución en el canal de 0-jets de m_T^{W1} .

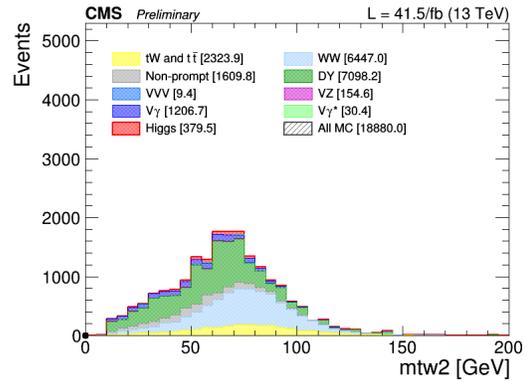


Figura 80: Distribución en el canal de 0-jets de m_T^{W2} .

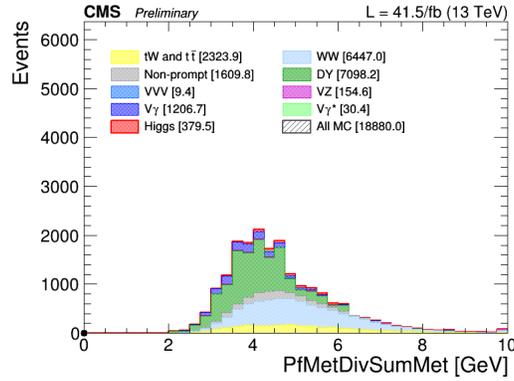


Figura 81: Distribución en el canal de 0-jets de $P fMetDivSumMet$.

1-jet

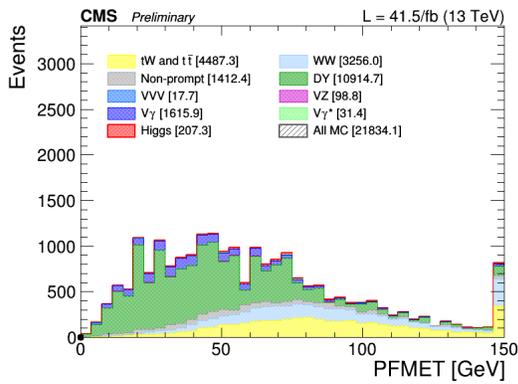


Figura 82: Distribución en el canal de 1-jet de E_T^{miss} .

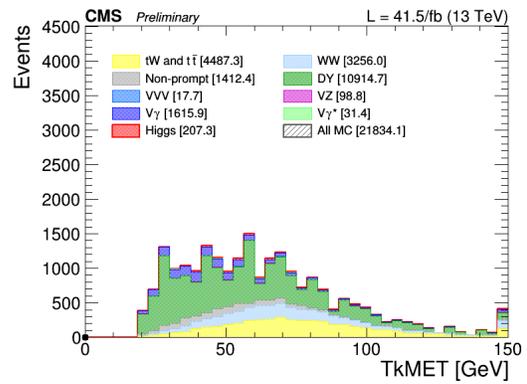


Figura 83: Distribución en el canal de 1-jet de $track E_T^{miss}$.

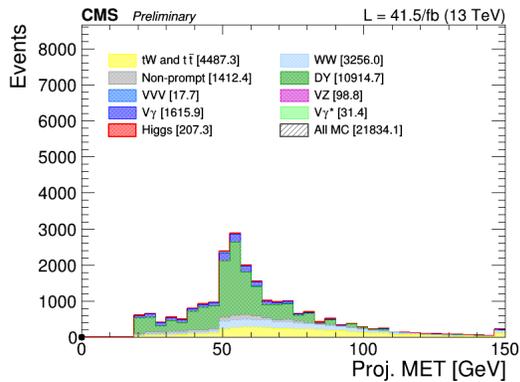


Figura 84: Distribución en el canal de 1-jet de $projected E_T^{miss}$.

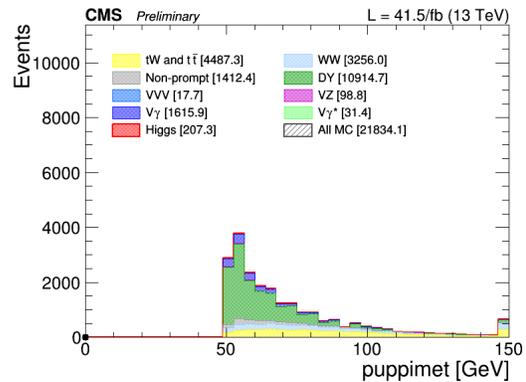


Figura 85: Distribución en el canal de 1-jet de $puppi E_T^{miss}$.

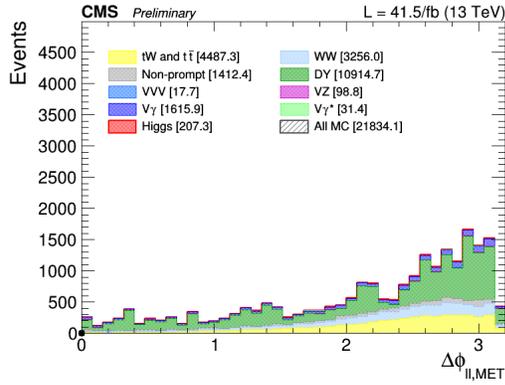


Figura 86: Distribución en el canal de 1-jet de $\Delta\phi_{ll, E_T^{miss}}$.

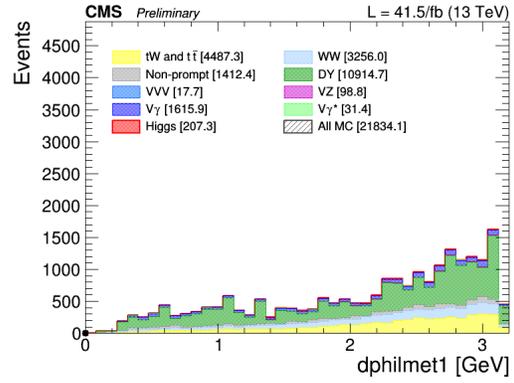


Figura 87: Distribución en el canal de 1-jet de $\Delta\phi_{l1, E_T^{miss}}$.

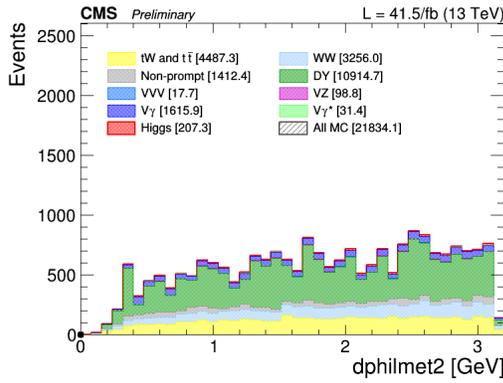


Figura 88: Distribución en el canal de 1-jet de $\Delta\phi_{l2, E_T^{miss}}$.

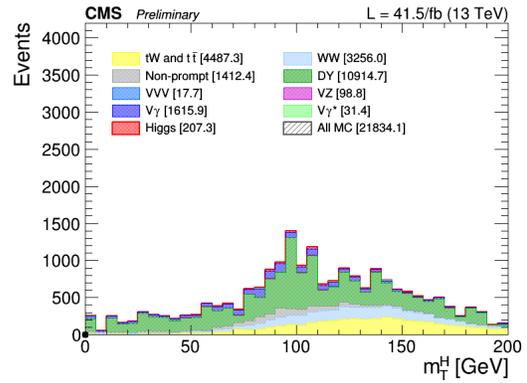


Figura 89: Distribución en el canal de 1-jet de m_T^H .

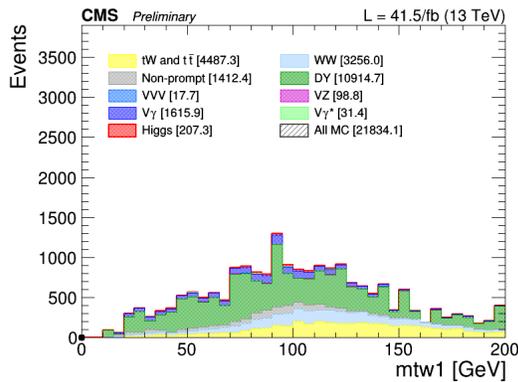


Figura 90: Distribución en el canal de 1-jet de m_T^{W1} .

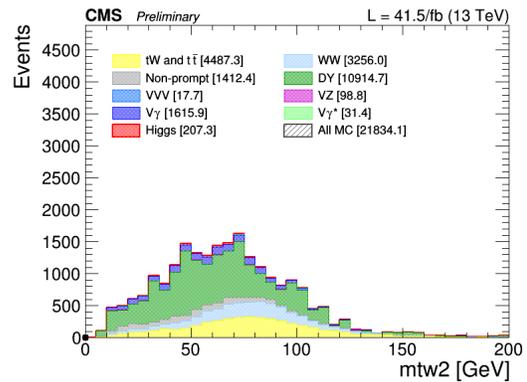


Figura 91: Distribución en el canal de 1-jet de m_T^{W2} .

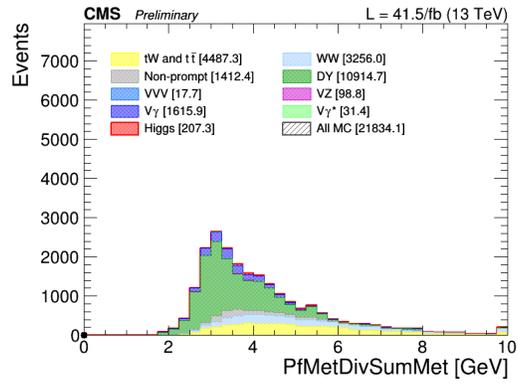


Figura 92: Distribución en el canal de 1-jet de $PfMetDivSumMet$.