



Universidad
Zaragoza



Facultad de Ciencias
Universidad Zaragoza

Trabajo de fin de grado de Física

Mejoras del análisis de datos del experimento ANAIS-112 con técnicas de aprendizaje automático

Universidad de Zaragoza-Facultad de Ciencias-Departamento de Física Teórica

Junio de 2025

Autora:

Marta Rojo Gómez

Directores:

Dra. María Lucía Martínez Pérez,
Dr. Iván Coarasa Casas

Índice

1. Introducción	1
1.1. La búsqueda de materia oscura	1
1.2. Objetivos	1
2. El experimento ANAIS-112 y el nuevo sistema de adquisición ANOD	2
2.1. Descripción de ANAIS-112	2
2.2. Parámetros para el estudio de la forma de los pulsos	3
2.3. Poblaciones de estudio	3
2.4. Necesidad de un sistema de adquisición de datos mejorado (ANOD)	4
2.5. Sucesos de baja energía en ANAIS y ANOD	6
3. Descripción de la técnica de aprendizaje automático	9
3.1. Árboles de decisión	9
3.2. <i>Boosted Decision Trees: AdaBoost</i>	10
3.3. Implementación del algoritmo	11
4. Aplicación de <i>AdaBoost</i> a ANOD	11
4.1. Selección y depuración de las poblaciones para el entrenamiento del algoritmo	11
4.2. Variables de entrenamiento y optimización del número de árboles	12
4.3. Optimización del corte en BDT	15
4.4. Evaluación del sobreentrenamiento e infraentrenamiento	18
4.5. Validación del corte en BDT	19
5. Comparativa de los resultados de ANOD respecto de ANAIS	19
5.1. Filtrado del fondo y eficiencia	19
5.2. Reducción del umbral del experimento	22
6. Conclusiones y líneas de mejora	24

1. Introducción

1.1. La búsqueda de materia oscura

Desde su primera predicción en 1933 por Fritz Zwicky [1], y hasta las estimaciones actuales que sitúan la materia oscura en torno al 84 % de la composición total de materia en el Universo [2], su naturaleza ha permanecido como uno de los grandes enigmas de la física de partículas desde su planteamiento inicial. A lo largo de los años, se han propuesto numerosas partículas más allá del Modelo Estándar como posibles candidatas. Entre ellas, los WIMPs (*Weakly Interacting Massive Particles*) han destacado como una de las opciones más estudiadas, debido a sus propiedades especialmente adecuadas para explicar la materia oscura: serían partículas estables, eléctricamente neutras, con masas en el rango de 1 GeV a 100 TeV y que interactuarían únicamente mediante la interacción débil y la gravedad [3]. Estas características justificarían la dificultad para detectarlas directamente mediante experimentos convencionales.

La primera propuesta para su detección directa parte de la idea de que, al moverse la Tierra a través del halo galáctico de materia oscura, los WIMPs deberían atravesar los detectores. Ocasionalmente, podrían colisionar con núcleos atómicos de estos con secciones eficaces del orden de 10^{-40} - 10^{-50} cm², provocando retrocesos medibles con energías inferiores a 100 keV. Una señal muy distintiva que permitiría confirmar su existencia sería la modulación anual en el ritmo de detección, causada por la variación en la velocidad relativa entre la Tierra y el halo a lo largo del año. En esta idea se basa el experimento DAMA/LIBRA, situado en el laboratorio del Gran Sasso, en Italia. Durante más de 20 años, este experimento ha registrado una modulación anual compatible con la esperada por WIMPs, con un nivel de confianza estadística significativo (hasta 11.8σ en la región de baja energía [1-6] keV) [4], utilizando como blanco cristales centelladores de NaI(Tl). Sin embargo, esta señal no ha sido confirmada por otros experimentos, que, a pesar de tener mayor sensibilidad, emplean blancos distintos. No obstante, esta comparación entre experimentos depende de los modelos de halo e interacción entre los WIMPs y los núcleos. Así, ante la necesidad de confirmar o refutar de forma independiente el resultado de DAMA/LIBRA con el mismo material, surge el experimento ANAIS-112 (*Annual modulation with -112 kg- NaI(Tl) Scintillators*), cuya toma de datos comienza en agosto de 2017 en el Laboratorio Subterráneo de Canfranc (España), llevando acumulados casi ocho años de medidas [5].

1.2. Objetivos

La extracción de una posible señal de modulación presenta numerosos desafíos. Las medidas del experimento están afectadas por un fondo que no solo se debe a isótopos radiactivos presentes en los propios centelleadores y en sus alrededores, sino también a eventos de falso centelleo, difíciles de distinguir de los centelleos auténticos a bajas energías. A pesar de los avances en el análisis de eventos de baja energía (<10 keV, donde se espera la señal de materia oscura dado el material blanco utilizado)—en particular, la incorporación de técnicas de aprendizaje automático, que han supuesto un hito al mejorar la eficiencia y reducir el fondo [6]—, a muy baja energía (<2 keV) persisten discrepancias entre las predicciones de los modelos de fondo y los datos experimentales, además de una eficiencia limitada en ANAIS. Esta situación evidenció la necesidad de implementar un sistema de adquisición mejorado, ANOD, para tratar de mejorar la discriminación de sucesos a baja energía. En comparación con el sistema de adquisición de ANAIS, ANOD presenta dos mejoras que permiten ampliar el estudio de eventos de baja energía: la salida de los fotomultiplicadores se digitaliza en una ventana temporal cuatro veces más larga, y no hay tiempo muerto entre sucesos.

Con su instalación ya completada, y motivado por un trabajo previo en el que, utilizando los primeros datos de ANOD para cuatro de los nueve detectores, ya se observó un defecto en el filtrado de sucesos [7], este trabajo propone mejorar la discriminación de sucesos a baja energía utilizando los eventos registrados por dicho sistema. Para ello, el uso de técnicas de aprendizaje automático resulta indispensable, dada la efectividad que han demostrado previamente en ANAIS.

2. El experimento ANAIS-112 y el nuevo sistema de adquisición ANOD

2.1. Descripción de ANAIS-112

El experimento ANAIS-112 está compuesto por nueve detectores centelleadores de NaI(Tl), dispuestos en una matriz 3×3. Cada detector tiene forma cilíndrica, con un diámetro de 12.07 cm, una longitud de 29.85 cm y una masa de 12.5 kg. Aunque el experimento se encuentra instalado bajo 800 metros de roca, lo que equivale a un apantallamiento de 2450 metros equivalentes de agua (m.w.e.) frente a la radiación cósmica, se ha diseñado un sistema de blindaje adicional para minimizar el fondo radiactivo. Encapsulados en cobre OFE (*Oxygen Free Electronic*), los detectores están rodeados por una estructura de apantallamiento compuesta, de dentro hacia fuera, por 10 cm de plomo romano, 20 cm de plomo de baja actividad y una caja anti-radón, mantenida mediante un flujo continuo de nitrógeno libre de radón. Además, y como mejora frente al experimento DAMA/LIBRA, ANAIS-112 está rodeado por un sistema de veto activo compuesto por 16 centelleadores plásticos para la detección de muones. Finalmente, el conjunto se completa con tanques de agua y bloques de polietileno de 40 cm de grosor, que actúan como moderadores de neutrones [8].

En lo que respecta al sistema de adquisición de datos (DAQ), cada detector está acoplado en sus extremos a tubos fotomultiplicadores (PMT) Hamamatsu R12669SEL2 de alta eficiencia cuántica, mediante ventanas de cuarzo. El disparo (*trigger*) de cada PMT se genera utilizando un discriminador de fracción constante CAEN N843, capaz de operar de manera efectiva al nivel de un solo fotoelectrón. Dado que el tiempo de centelleo del NaI(Tl) es de aproximadamente 230 ns [11], el *trigger* de cada módulo se establece con una condición de coincidencia lógica AND entre las señales de ambos PMT en una ventana temporal de 200 ns. La señal de cada PMT se divide en varias vías con distintas ganancias, lo que permite cubrir eficazmente tanto el rango de baja como el de alta energía. Aunque el estudio del rango de alta energía es de interés para el control del fondo radiactivo, la región de interés (ROI) de este experimento, donde se espera la señal de materia oscura, se da para baja energía (< 10 keV).

Por tanto, centrándose en el rango de baja energía, la señal de cada PMT se digitaliza mediante un módulo MATAcq-CAEN V1729A, con una frecuencia de muestreo de 2 GS/s y una alta resolución de 14 bits. Cada señal se almacena en búferes circulares de 2520 puntos para su posterior transferencia a un ordenador y, en total, cada suceso registrado por ANAIS-112 se compone de dos señales (una por cada PMT) de 1.26 μ s de duración, incluyendo un *pretrigger* de 0.25 μ s, y se adquieren con una resolución temporal de 0.5 ns por punto [8].

El disparo general del sistema de adquisición se genera mediante el OR lógico de los nueve módulos, en una ventana de 1 μ s. Un módulo electrónico (*pattern unit*) registra qué módulos han producido señal, de modo que solo se transfiere al ordenador la información de aquellos que hayan generado un disparo. Durante el tiempo que dura la transferencia, un módulo CAEN V977 I/O Register (IOREG0) inhibe nuevos disparos, generando un tiempo muerto de aproximadamente 4 ms por suceso [8].

A pesar de todos los esfuerzos por tratar de minimizar el fondo radiactivo del experimento utilizando cristales radiopuros y blindajes, todavía existe un ritmo elevado en la región de baja energía que dificulta la observación de posibles sucesos generados por la interacción de partículas de materia oscura. Por ello, es imprescindible estudiar exhaustivamente los sucesos de baja energía del experimento (que se verán en detalle en la sección 2.5), con el objetivo de desarrollar protocolos de filtrado que permitan seleccionar correctamente los sucesos de centelleo. Para ello, a partir del segundo año de toma de datos se introduce un décimo módulo llamado blank, compuesto por los dos tubos PMT, pero sin centelleador, de manera que permite analizar los sucesos de falso centelleo. Este módulo está situado al lado del montaje de ANAIS-112 y cuenta con un blindaje independiente de 20 cm de plomo.

2.2. Parámetros para el estudio de la forma de los pulsos

Para caracterizar los sucesos se definen los siguientes parámetros (definidos sobre el pulso suma, aunque la definición se extiende sobre cada PMT, añadiendo un sufijo 0 ó 1 en función del escogido):

- $P_1 = \frac{\sum_{100ns}^{600ns} A(t)}{\sum_{0ns}^{600ns} A(t)}$ Siendo $A(t)$ la amplitud del pulso a un tiempo t después del *trigger* t_0 , esta variable representa la fracción de pulso en la cola.
- $P_2 = \frac{\sum_{0ns}^{50ns} A(t)}{\sum_{0ns}^{600ns} A(t)}$ Análogamente, representa la fracción de pulso en la cabecera.
- $CAP_x = \frac{\sum_0^x A(t)}{\sum_0^{t_f} A(t)}$ Mide la proporción de la carga acumulada hasta un tiempo específico x en comparación con la carga total acumulada durante un tiempo t_f .
- $\mu_p = \frac{\sum_i A_i t_i}{\sum_i A_i}$ Siendo A_i el valor de amplitud dado un pico a tiempo t_i , esta variable representa el primer momento.

También se definen otros parámetros relevantes, como el número de picos n_0 y n_1 (identificados por un algoritmo de detección de picos) y de fotoelectrones $nphe_0$ y $nphe_1$ de cada PMT (definidos como el área del pulso dividido entre el área de un fotoelectrón); la energía total del suceso (calculada a partir del área del pulso suma hasta 1 μ s); la multiplicidad de un suceso *mult* (que cuenta el número de detectores en los que se ha producido señal para un mismo suceso); la asimetría en número de fotoelectrones entre PMTs $Asynphe = (nphe_0 - nphe_1)/(nphe_0 + nphe_1)$; el momento en el que se alcanza el máximo del pulso según el PMT $tmax_0$ y $tmax_1$; el momento en el que se alcanza el máximo del pulso total $Tmax$ y el segundo momento central SM , calculado a partir del pulso digitalizado.

2.3. Poblaciones de estudio

La población objeto de estudio, de la que se pretende extraer la posible señal de modulación, son las medidas de fondo. Sin embargo, son también necesarias medidas con fuentes conocidas con el objetivo de calibrar la región de baja energía y corregir posibles derivas en la ganancia. Por ello, cada dos semanas se interrumpe la medida de fondo para realizar una calibración durante 3–4 horas utilizando una fuente de ^{109}Cd , cuya desintegración emite líneas características a 11.9, 22.6 y 88.0 keV. Esta calibración se lleva a cabo mediante dos sirgas que contienen las fuentes y que se introducen en el montaje experimental, coincidiendo la posición de las fuentes en cada sirga con ventanas de Mylar en los detectores que permiten el paso de la radiación.

Además, durante un período prolongado, se emplearon también las poblaciones de coincidencias asociadas a las desintegraciones de ^{22}Na (0.87 keV) y ^{40}K (3.2 keV). Estos isótopos se desintegran por captura electrónica a un núcleo hijo excitado, lo que resulta en una detección simultánea entre dos detectores: uno detecta el suceso de baja energía y el otro el fotón de alta energía emitido. Actualmente, esta población no resulta tan significativa puesto que, aunque el isótopo ^{40}K es estable en este contexto, el ^{22}Na —más relevante por tener su pico cercano al umbral de energía del experimento (actualmente en 1 keV)— tiene una vida media de 2.6 años. Dado que el experimento acumula casi ocho años de medidas, la estadística asociada a este isótopo es ya muy limitada.

Las poblaciones mencionadas se basan en retrocesos electrónicos. Sin embargo, se espera que los WIMPs interactúen produciendo retrocesos nucleares. La eficiencia de centelleo no es la misma en ambos casos, por lo que sería necesario aplicar una conversión a esta energía. Desafortunadamente, esta

conversión, conocida como factor de eficiencia relativa (o de *quenching*), aún no está bien definida [9]. Estudios recientes han mostrado una variación significativa en las mediciones del factor de quenching en cristales de NaI(Tl), especialmente por debajo de 10 keV, coincidiendo con la región de interés para la búsqueda de materia oscura [10]. Una mejor comprensión de este factor es crucial para resolver discrepancias entre diferentes experimentos, como los resultados nulos de la mayoría de las búsquedas de WIMPs y la afirmación de detección de materia oscura por parte de la colaboración DAMA/LIBRA.

Por este motivo, se realizan calibraciones con una fuente de neutrones de ^{252}Cf . Como ventaja adicional, la actividad de esta fuente es mayor, lo que permite recopilar una mayor estadística en un tiempo más corto. Además, a diferencia de los sucesos de ^{109}Cd , los de neutrones se distribuyen homogéneamente por todo el volumen de los detectores.

La última población relevante a tener en cuenta es la de sucesos adquiridos por el módulo sin centelleador (blank). Como se menciona en la sección 2.1, esta es de utilidad para el estudio de sucesos de falso centelleo. Es destacable que la ausencia de centelleador impide utilizar fuentes radiactivas para su calibración en energía, por lo que esta se realiza a partir de la cantidad de luz observada, concretamente a través de la recolección de luz de los módulos de ANAIS-112 (aproximadamente, unos 14.5 fotoelectrones por keV de energía). La energía en este módulo se denota como *energía equivalente* y según los parámetros definidos sobre el pulso en la sección 2.2, es $(nphe0 + nphe1)/14.5$.

2.4. Necesidad de un sistema de adquisición de datos mejorado (ANOD)

La primera estrategia empleada para seleccionar sucesos de centelleo consistió en aplicar cortes sobre las variables de los pulsos descritas en la sección 2.2 [8]. Sin embargo, esta técnica conduce a una eficiencia limitada —definida como la fracción de sucesos teóricamente buenos que superan el corte—, lo que la convierte en un método demasiado agresivo. Además, el desarrollo de simulaciones individualizadas para cada detector, que incorporan tanto su contaminación específica como la de los materiales que los rodean, pone de manifiesto una diferencia significativa con respecto a los datos experimentales. En este contexto, la aplicación de técnicas de aprendizaje automático ha supuesto una mejora considerable, incrementando la eficiencia y reduciendo el fondo [6].

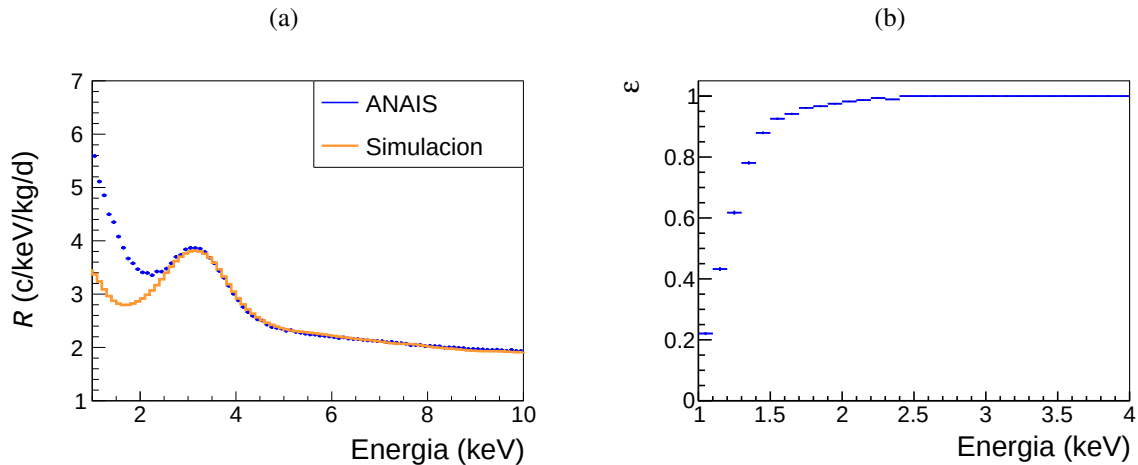


Figura 1: (a) Espectro de fondo en anticoincidencia, correspondiente al promedio de los detectores, adquirido por ANAIS. Se muestra para los seis primeros años de medidas y filtrado con técnicas de aprendizaje automático (en azul), comparado con la simulación (en naranja). Se observa una fuerte discrepancia en la región [1–2] keV. (b) Eficiencia de aceptación de sucesos. Se observa una fuerte disminución por debajo de 1.5 keV.

No obstante, tal y como se observa en la Figura 1a, aun con técnicas de aprendizaje automático persiste una discrepancia entre el espectro de fondo simulado (en naranja) y el medido (en azul), ambos en anticoincidencia ($mult = 1$), ya que se espera que la materia oscura interactúe de forma aislada con un único detector, dada su sección eficaz de interacción débil. Esta puede ser debida a una deficiencia en el filtrado de sucesos o a una contaminación no contemplada en el modelo de fondo. Especialmente, se observa un desacuerdo notable respecto de la simulación en la región [1-2] keV.

Por otro lado, en la Figura 1b se tiene una fuerte disminución de la eficiencia por debajo de 1.5 keV, llegando hasta un valor de 20 % a 1 keV. Esta es la razón por la que se establece el umbral del experimento en 1 keV, lo que continúa siendo una de las principales limitaciones del experimento ANAIS. Cabe destacar que esta Figura 1 se corresponde con el promedio de todos los detectores, para los seis primeros años de medidas.

Así, con el doble objetivo de mejorar la discriminación de sucesos de ruido o espurios y la eficiencia de selección de eventos, se propuso la instalación de un sistema paralelo de adquisición de datos: ANOD (*ANAIS No Dead-time*). ANOD fue instalado inicialmente en los detectores 0, 2, 5, 8 el 23/08/2023. El 02/04/2024 se dejó de digitalizar el detector 8 para medir el módulo blank —puesto que la tarjeta de adquisición disponible entonces solo contaba con ocho canales y cada detector necesita un canal por PMT—. Su implementación se completó el 17/12/2024 gracias a la adquisición de una nueva tarjeta de 32 canales, que permitió la incorporación de todos los módulos (0-8 y blank). De este modo, el sistema se encuentra plenamente operativo durante la realización del presente trabajo. En la Tabla 1 se muestra una comparativa entre los sistemas de adquisición de datos ANAIS y ANOD y, en la Figura 2, un esquema de los dos sistemas de adquisición completos.

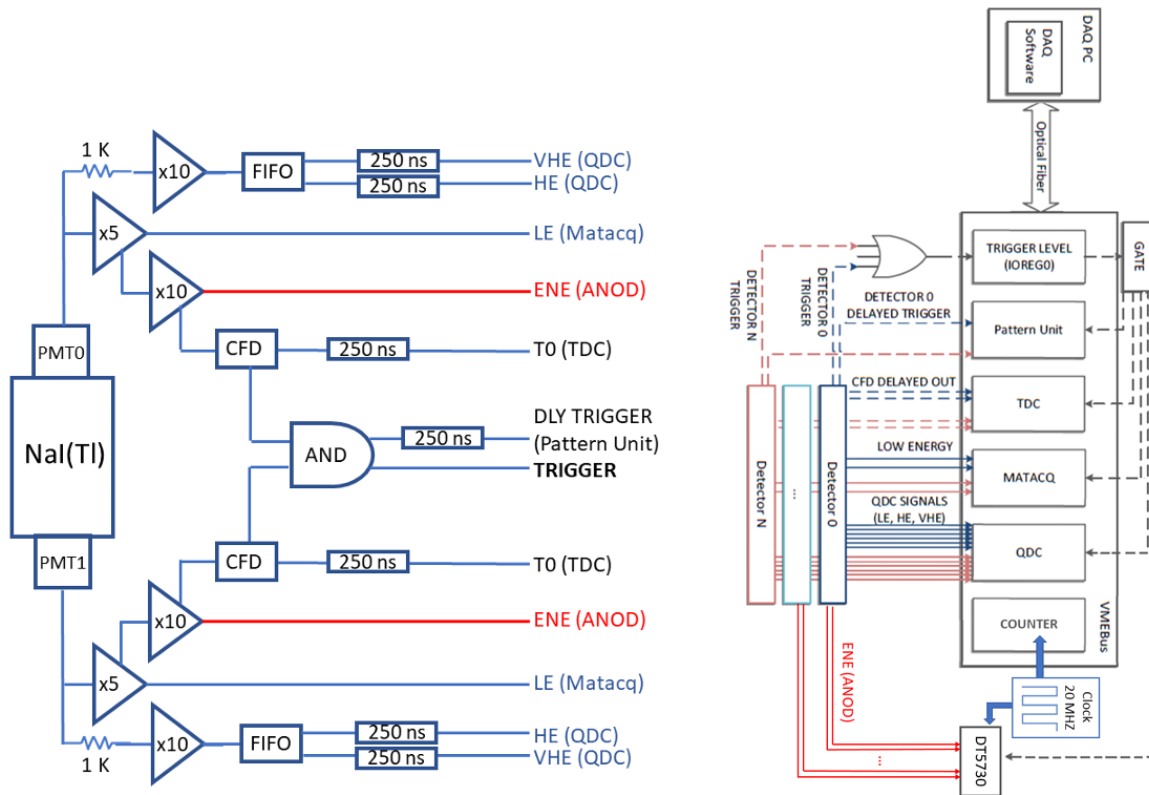


Figura 2: Esquema de la cadena de adquisición ANAIS (vías azules) y ANOD (vías rojas).

Entre las características más relevantes, ANOD amplía la ventana de digitalización a 8 μ s, pero disminuye la frecuencia de muestreo con respecto a ANAIS (2 GS/s vs 500 MS/s), y por tanto la resolución temporal del pulso (0.5 vs 2.0 ns/punto). Por otro lado, como se observa en la Figura 2, el sistema ANOD es independiente al de ANAIS, estando sincronizados por un reloj común.

Parámetro	ANAIS	ANOD
Ventana de digitalización (ns)	1260	8000
Pretrigger (ns)	250	715
Resolución temporal (ns/punto)	0.5	2
Tiempo muerto (ms)	~ 4	–

Tabla 1: Comparación de parámetros de adquisición entre los sistemas ANAIS y ANOD.

Finalmente, cabe mencionar que la instalación de ANOD permite la definición de nuevas variables. Entre las que serán relevantes para este trabajo, destacan la variable *MAD*, que representa la desviación absoluta media (*Mean Absolute Deviation*); *MinWidthAnod*, que indica el intervalo temporal entre el primer y el último pico observado en la traza del PMT; el número de picos detectados tras la ventana de ANAIS-112 según el PMT *n0tail* y *n1tail*; y la asimetría en número de picos detectados tras la ventana de ANAIS-112 $Asyntail = (n0tail - n1tail)/(n0tail + n1tail)$. Además, también es posible definir las mismas variables que para ANAIS (que se denotan con el superíndice *ANOD*) y variables adquiridas por ANOD restringidas a la ventana de ANAIS (por ejemplo: *Asynphe^{ANODANA}*).

2.5. Sucesos de baja energía en ANAIS y ANOD

Para poder distinguir los sucesos de centelleo reales de los espurios, es necesario comprender la fenomenología presente en la región de interés. Por ello, esta sección se dedica a su análisis, mostrando la ventaja del sistema de adquisición de datos mejorado ANOD.

Los sucesos canónicos de centelleo son aquellos cuya señal decae con la constante característica del NaI(Tl), de aproximadamente 230 ns [11]. Un ejemplo representativo de este tipo de evento, registrado simultáneamente por ambos sistemas DAQ, se muestra en la Figura 3.

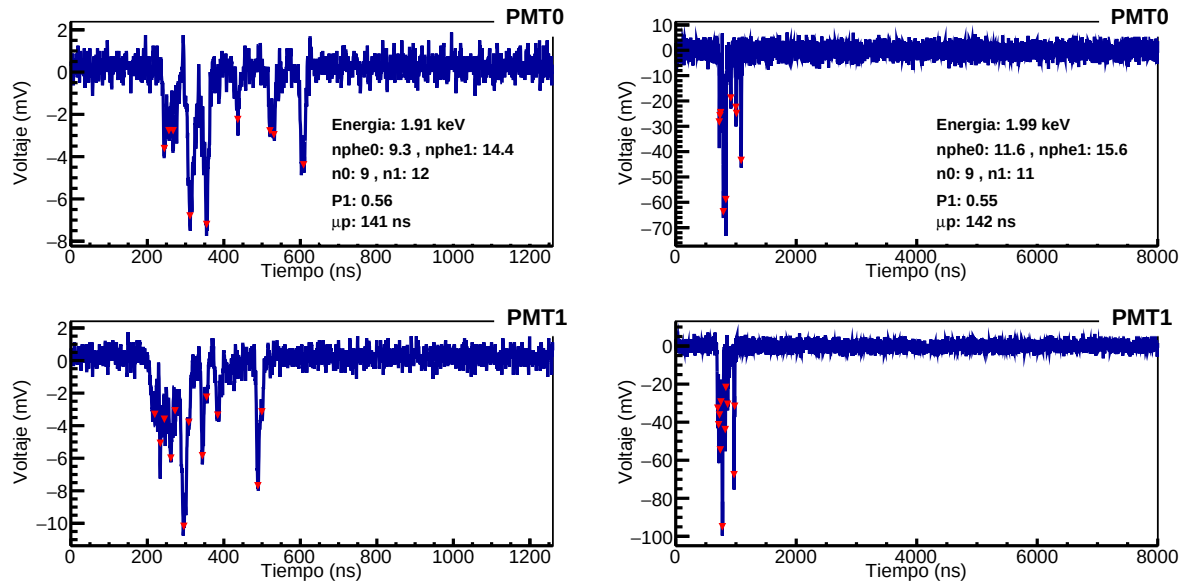


Figura 3: Ejemplo de un mismo suceso de centelleo adquirido por ambos DAQs: a la izquierda, el sistema ANAIS; a la derecha, el sistema ANOD. Se aprecia una diferencia de ganancia entre ambos sistemas, y los triángulos rojos indican los picos detectados por un algoritmo de detección de picos [8].

En cuanto a los sucesos de no centelleo, contribuyen:

- Sucesos de luz de Cherenkov, generados principalmente en el vidrio de los PMTs (ver Figura 4a). Estos son rápidos, de manera que se pueden eliminar a través del corte $P_1 > 0.35$, $P_{10} > 0.15$ y $P_{11} > 0.15$, prácticamente sin pérdida de eficiencia, en ambos DAQs. Además, este corte también es capaz de eliminar de manera muy eficiente las coincidencias aleatorias de fotoelectrones individuales producidos por la corriente oscura, aunque normalmente tienen una energía equivalente de dos fotoelectrones y por tanto están por debajo del umbral de análisis (1 keV).
- Sucesos altamente energéticos, como los producidos por la interacción de un muon. En ANAIS, sus colas prolongadas son capaces de activar múltiples veces el DAQ, generando señales de falso centelleo a baja energía. Gracias al sistema de veto activo, cuando un muon es detectado por los centelleadores plásticos, se descartan los sucesos producidos durante el segundo posterior, eliminando estas colas eficientemente.
- Eventos con constantes de decaimiento diferentes, habitualmente más largas y generalmente con un carácter asimétrico en el reparto de luz entre ambos PMTs (ver Figura 4b). Se atribuyen a emisiones de luz en el PMT. En el experimento ANAIS, estos decaimientos lentos representan un problema, ya que son difíciles de distinguir de los sucesos de centelleo auténtico debido al tamaño limitado de la ventana (1.26 μ s).

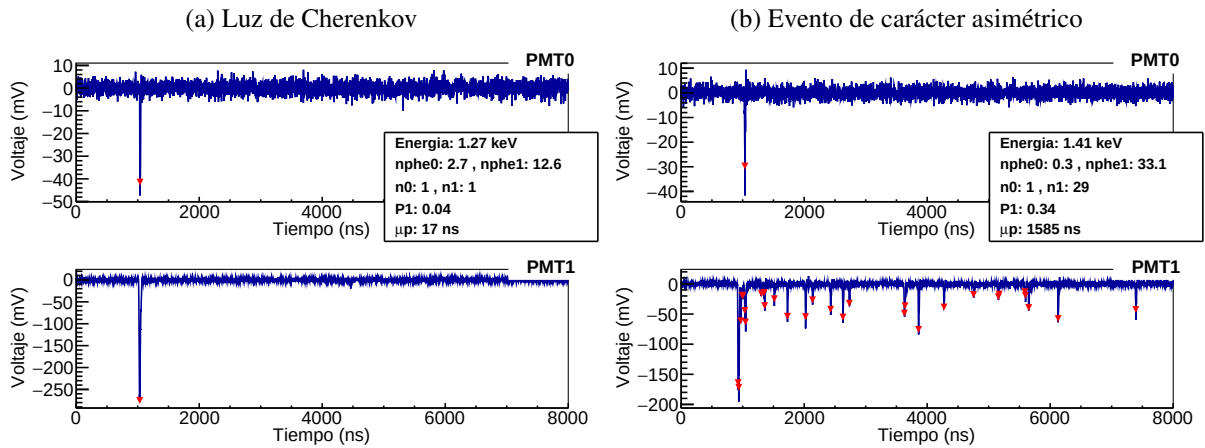


Figura 4: Ejemplos de sucesos de no centelleo: (a) luz de Cherenkov y (b) pulso asimétrico.

Es en este último caso en el que destaca la relevancia de ANOD. Con una ventana temporal más larga, por ejemplo, se desplaza el primer momento de los sucesos con decaimientos largos hacia valores mayores. Esto es lo que se muestra en la Figura 5, en la que se representa el ritmo de sucesos en [1-2] keV en función del primer momento para diversas poblaciones: neutrones (*Neu*), con condición de multiplicidad igual (en verde) o mayor que uno (en negro); cadmio (*Cd*, en azul); fondo (*Fondo*, en naranja); y blank (en rojo). A esta última se le ha aplicado un factor multiplicativo de nueve, dado que el módulo blank es uno solo frente a los nueve módulos con detectores centelleadores que registran el resto de eventos. Además, cabe destacar que tanto a la población de fondo como a la de blank se les ha impuesto el corte $P_1^{ANOD} > 0.35$, $P_{10}^{ANOD} > 0.15$ y $P_{11}^{ANOD} > 0.15$, de forma que los sucesos mostrados son aquellos que podrían confundirse con eventos compatibles con centelleo. En (a) se muestran los sucesos de ANOD restringidos a la ventana de ANAIS, y en (b), los de la ventana completa de ANOD.

La figura evidencia la pérdida de información en ANAIS. Mientras que en este DAQ no es posible distinguir los sucesos de centelleo de la población de sucesos de alta asimetría, ANOD consigue diferenciarlos y concentrar estos últimos en la región $\mu_p^{ANOD} = 1500$ ns. Es relevante mencionar, además, que tanto la población de calibración de ^{109}Cd como los sucesos en anticoincidencia de la calibración de neutrones contienen una fracción no despreciable de sucesos de falso centelleo. En este aspecto, la selección de los sucesos de neutrones múltiples supone una gran mejora, lo que tendrá relevancia en la sección 4.1.

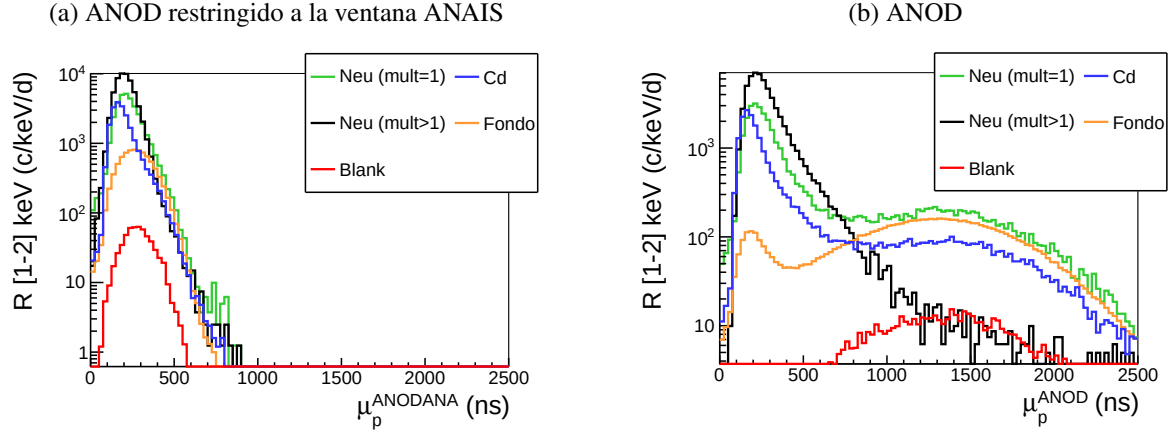


Figura 5: Distribución del primer momento en [1-2] keV en la ventana de ANAIS (izquierda) y de ANOD (derecha) para distintas poblaciones: neutrones (*Neu*), con condición de multiplicidad igual (en verde) o mayor que uno (en negro); cadmio (*Cd*, en azul); fondo (*Fondo*, en naranja); y blank (en rojo).

Es importante tener en cuenta que, aunque se han nombrado dos fuentes de falso centelleo, existen otras todavía sin identificar o sin caracterizar correctamente. Un ejemplo de ello son los *afterpulses*, generados por ionización de impurezas dentro de los PMTs [11]. Su caracterización e identificación precisa continúa siendo una línea activa de investigación en el marco del experimento.

También, se debe destacar que en ambos sistemas DAQ se registran en contadas ocasiones tanto sucesos de centelleo reales como espurios que presentan una caracterización deficiente. Un ejemplo es el de los eventos que disparan durante el *pretrigger*, lo que impide un cálculo correcto de la línea de base y conlleva una distorsión en los parámetros del pulso. Este efecto puede observarse en la Figura 6a, donde se registran valores de $P_1 > 1$ y número de *phe* y energía negativos. Asimismo, pueden detectarse sucesos que ocurren al final de la ventana de adquisición, quedando registrados de forma incompleta. A lo largo de este trabajo, estos eventos se eliminan de todas las poblaciones, garantizando que las variables P_1 y P_2 queden correctamente definidas en el intervalo [0,1], y que las áreas y el primer momento sean positivos.

Adicionalmente en ANOD se registran sucesos con oscilaciones en la línea de base (ver Figura 6b), que solo se manifiestan durante el tiempo muerto de ANAIS y cuya causa aun se desconoce. Estos eventos pueden ser eliminados a través de un corte en la variable *MAD* que se aplicara a todas las poblaciones a lo largo del trabajo. Eventualmente, este corte también capaz de eliminar de manera colateral algunos sucesos asimétricos.

Por último, cabe destacar que la clasificación de los eventos resulta no trivial, ya que con frecuencia se observan pulsos que no presentan de forma clara las características propias ni de sucesos de centelleo genuinos ni de falsos. Esta complejidad se ve acentuada en el sistema ANOD, donde la mayor duración de la ventana de adquisición permite la identificación de combinaciones complejas dentro de una misma ventana temporal. Entre otros, estas combinaciones pueden incluir sucesos de Cherenkov, *afterpulses*, coincidencias fortuitas con un único fotoelectrón, fosforescencias en el NaI(Tl) y eventos genuinos de centelleo.

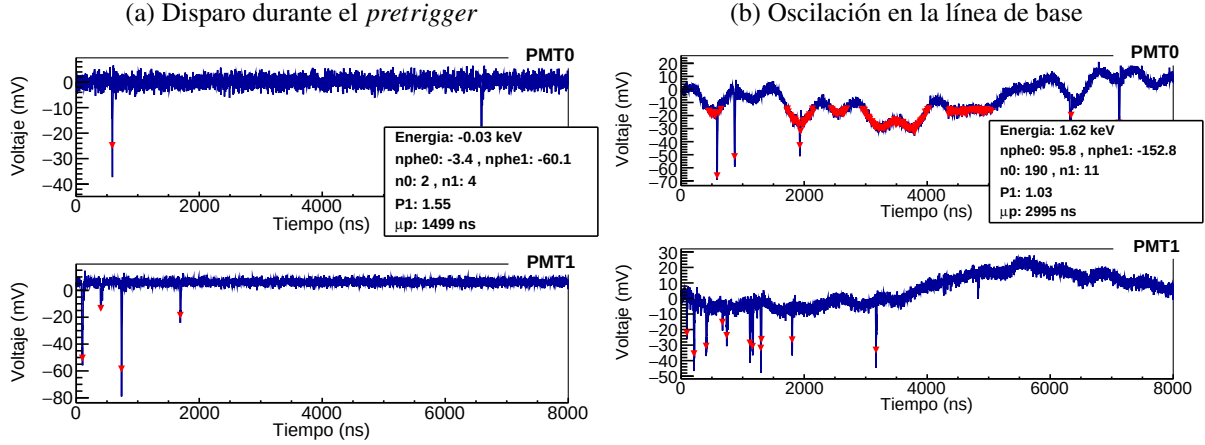


Figura 6: Ejemplos de sucesos con la línea de base mal caracterizada: (a) disparo durante el *pretrigger* y (b) oscilación en la línea de base (solo presente en ANOD).

3. Descripción de la técnica de aprendizaje automático

La alta variabilidad observada en los datos registrados por ANOD, junto con la comprobación previa de un defecto en el filtrado de sucesos en los primeros cuatro detectores digitalizados con ANOD al aplicar cortes simples con este sistema [7], hace que el empleo de técnicas de aprendizaje automático no solo sea recomendable, sino necesario. En particular, se emplean *Boosted Decision Trees* (BDT), una técnica multivariante que extiende los análisis basados en cortes simples y permite tratar eventos que no presentan de forma clara todas las características de señal o ruido [12], debido a la citada variabilidad. Esta metodología ha sido ya ampliamente validada en análisis similares [6], lo que refuerza su idoneidad para el presente estudio.

3.1. Árboles de decisión

Comenzando por el concepto de árbol de decisión, en nuestro caso de interés —clasificar un suceso como centelleo verdadero (señal) o falso centelleo (ruido)— se utilizan árboles de decisión binarios. Es decir, se parte de un conjunto de sucesos (nodo raíz), cada uno caracterizado por un conjunto de variables discriminantes, $\vec{x}_i$. En cada nodo, se selecciona una de estas variables (ya sea aleatoriamente o mediante un criterio de importancia) y se prueban distintos valores de corte sobre ella. Se elige el corte c_j que mejor separa para los eventos según una métrica de pureza o ganancia de información, dividiendo el nodo en dos ramas: una para los eventos que superan el corte y otra para los que no lo hacen. Este proceso se repite recursivamente en cada rama hasta que se cumple un criterio de detención, como alcanzar una profundidad máxima del árbol, un número mínimo de eventos por nodo, o que no se logre una ganancia significativa con nuevas divisiones. Cuando esto ocurre, se dice que se ha alcanzado una hoja y el suceso queda clasificado como señal o ruido (ver Figura 7).

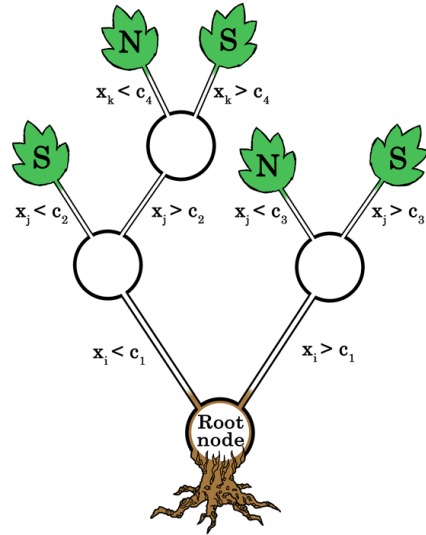


Figura 7: Esquema de un árbol de decisión. Los sucesos se van clasificando en nodos hijos según cortes c_j de las variables discriminantes $\vec{x}_i$, hasta llegar a una hoja, que se identifica como señal S o ruido N en función de la proporción de eventos de señal que contiene tras el entrenamiento. Figura extraída de [6].

El algoritmo necesita ser entrenado con un conjunto de eventos previamente etiquetados como señal (centelleos reales) o ruido (falsos centelleos). Durante el entrenamiento, a cada hoja del árbol (es decir, a cada conjunto de eventos que ya no se divide más) se le asigna un valor de pureza definido como $p = \frac{s}{s+r}$, donde s es la suma de los pesos de los eventos de señal y r la suma de los pesos de los eventos de ruido en esa hoja. Se interpreta que los eventos con $p \geq 0.5$ son clasificados como señal, y aquellos con $p < 0.5$, como ruido. Una vez entrenado, el árbol se puede aplicar a nuevos eventos no etiquetados para clasificarlos.

3.2. Boosted Decision Trees: AdaBoost

Cuando en vez de un solo árbol se tiene un conjunto de árboles (bosque) y un algoritmo de boosting (técnica que combina iterativamente clasificadores débiles para formar uno más robusto), el algoritmo de árboles de decisión se ve potenciado, recibiendo el nombre de Boosted Decision Trees. Siguiendo [12], se describe un algoritmo concreto de *Boosted Decision Trees*, el *Adaptive Boosting* (AdaBoost).

Esquemáticamente se tiene lo siguiente:

- **Fase de entrenamiento:** a cada evento de la muestra se le asigna un subpeso inicial $\omega_i^{(0)}$. Como es habitual, para asegurar que los eventos de señal y de ruido tengan la misma influencia en el entrenamiento, se utiliza una muestra balanceada con N eventos de cada tipo. En consecuencia, todos los eventos parten con el mismo subpeso inicial: $\omega_i^{(0)} = \frac{1}{2N} \forall i$. Al terminar la clasificación del árbol, se calcula un error asociado al árbol basado en si los sucesos han sido bien clasificados, $\varepsilon^{(0)}$, que a su vez asocia un peso al árbol $\alpha^{(0)}$. El peso del árbol modifica los subpesos en función de la correcta clasificación del evento. El siguiente se centra en los sucesos que han sido mal clasificados en el árbol anterior, por lo que parte de los subpesos modificados. El proceso se repite hasta llegar al árbol final.

$$\varepsilon^{(k)} = \frac{\sum_{i=1}^N \omega_i^{(k)} \cdot \text{isMissClassified}_k(i)}{\sum_{i=1}^N \omega_i^{(k)}} \quad \alpha^{(k)} = \beta \cdot \ln \left(\frac{1 - \varepsilon^{(k)}}{\varepsilon^{(k)}} \right) \quad \omega_i^{(k+1)} = \omega_i^{(k)} \cdot e^{\alpha^{(k)} \cdot \text{isMissClassified}_k(i)}$$

Donde *isMissClassified* asigna un valor +1 si el evento ha sido correctamente clasificado (señal como señal, y ruido como ruido) y -1 si no; y β es la tasa de aprendizaje.

- **Aplicación del entrenamiento:** dado un evento i , cada árbol k le asocia un valor $T_k = \pm 1$ en función de si ha sido clasificado como señal o como ruido. Ponderando con el peso de cada árbol, se asocia al evento un valor $BDT(i)$.

$$BDT(i) = \frac{1}{N} \sum_k^{nTrees} \ln(\alpha^{(k)} \cdot T_k(i))$$

Que estará comprendido en el intervalo $[-1, 1]$, siendo los eventos con BDT altos más susceptibles de ser señal, y los de BDT bajos de ser ruido. Al igual que en los árboles de decisión, donde los eventos que se clasifican en hojas con una probabilidad $p > 0.5$ se consideran buenos, en un algoritmo *AdaBoost* se establece un corte en el valor BDT. Los eventos que superen este umbral son clasificados como buenos. Por lo tanto, el enfoque se centra en seleccionar el valor adecuado del corte, en adelante *BDTCut*.

3.3. Implementación del algoritmo

La programación del algoritmo *AdaBoost* se realiza mediante el paquete TMVA (Toolkit for Multivariate Data Analysis) [13] de ROOT, el software de análisis de datos utilizado durante la realización de este trabajo. Los códigos base han sido desarrollados por el Dr. Iván Coarasa Casas, director de este trabajo. A continuación, se detallan y comentan los parámetros preestablecidos utilizados en el análisis.

- NTrees: Número de árboles en el bosque. Se prueba con valores: {1, 5, 10, 50, 100, 200, 500, 850, 1000, 2000}.
- MinNodeSize = 2.5%: Tamaño mínimo del nodo (porcentaje del total de eventos) requerido para dividir.
- MaxDepth = 3: Profundidad máxima del árbol. Limita la complejidad y evita sobreajuste.
- AdaBoostBeta = 0.5: Tasa de aprendizaje (β). Valores bajos priorizan eventos mal clasificados.
- nCuts = 20: Número de divisiones evaluadas por variable para optimizar el corte.
- SeparationType = GiniIndex: Métrica para evaluar divisiones. A partir de la definición de pureza extendida en árboles de decisión, donde se pondera cada evento con su subpeso correspondiente $\omega_i^{(k+1)}$, se evalúa la calidad de una división mediante el índice de Gini. Esta métrica se define como:

$$Gini = p \cdot (1 - p)$$

Donde p representa la pureza tras un corte. Entonces, el corte óptimo es aquel que minimiza el valor del índice de Gini, es decir, que maximiza la pureza del nodo resultante.

4. Aplicación de *AdaBoost* a ANOD

4.1. Selección y depuración de las poblaciones para el entrenamiento del algoritmo

Para la fase de entrenamiento del algoritmo de identificación de eventos de centelleo, es fundamental contar con poblaciones representativas tanto de señal como de ruido. Como se menciona en la sección 2.3, se dispone de las siguientes poblaciones: fondo, calibración con ^{109}Cd , coincidencias de ^{22}Na y ^{40}K , neutrones y blank.

Como población de ruido se ha seleccionado el blank, ya que proporciona una muestra libre de eventos físicos reales. Cabe destacar que en estudios anteriores se ha utilizado también la señal de fondo con el mismo propósito, dado que en su mayoría está compuesta por eventos de falso centelleo en la región [1-2] keV. Sin embargo, el fondo contiene una fracción no despreciable de eventos reales de centelleo (como decaimientos naturales), lo que podría introducir sesgos en el entrenamiento del algoritmo. Por esta razón, se prefiere utilizar el blank como representación más pura del ruido.

En lo que respecta a la señal, la población de coincidencias ha sido tradicionalmente descartada debido a su escasa estadística, una limitación que se ha acentuado en la actualidad (ver sección 2.3). Aunque en estudios anteriores se ha recurrido a la población de calibración con ^{109}Cd , la disponibilidad actual de medidas con neutrones hace que su utilización resulte más adecuada. Esto se debe a que los neutrones interactúan principalmente por dispersión elástica produciendo retrocesos nucleares, al igual que se espera de las partículas WIMP, lo que los convierte en una mejor aproximación a la señal buscada. Las diferencias intrínsecas entre las poblaciones de cadmio y neutrones se aprecian en la Figura 5, donde no solo se observa una mayor actividad asociada a la fuente de neutrones, sino también un desplazamiento

entre ambas distribuciones. Este desplazamiento refleja las distintas formas de interacción subyacentes a cada tipo de evento.

Sin embargo, incluso la población de neutrones presenta cierta contaminación por sucesos de ruido. Dado que la eficacia del algoritmo mejora cuanto más limpia sea la muestra de señal, es necesario realizar un estudio previo que permita depurar esta población. La mayor parte del ruido se concentra por debajo de los 2 keV, aumentando a medida que disminuye la energía (ver Figura 1a). En particular, por debajo de 1.5 keV la eficiencia de detección cae de forma significativa (ver Figura 1b), llegando a alcanzar un 20 % en 1 keV. Esto dificulta la identificación de eventos relevantes por debajo de este valor [6, 8], que se establece como umbral del experimento. Por este motivo, este trabajo se centra en el intervalo [1-2] keV.

Descartados los sucesos con variables mal definidas u otros sucesos anómalos, a partir del análisis de esta región:

- Vista la sección 2.5, es natural trabajar con una población de neutrones filtrada con los cortes $P_1^{ANOD} > 0.35$, $P_{10}^{ANOD} > 0.15$ y $P_{11}^{ANOD} > 0.15$ para eliminar eventos rápidos asociados a luz de Cherenkov. Estos cortes se denotan como corte de centelleo.
- La Figura 5 muestra que los eventos de neutrones con multiplicidad mayor que uno presentan características más consistentes con eventos de centelleo. Estas coincidencias múltiples, aunque aún pueden estar contaminadas por sucesos de falso centelleo, permiten una mejor discriminación frente a los sucesos de falso centelleo, que típicamente se manifiestan como eventos aislados.

Es relevante tener presente que estos cortes ayudan a tener una población más limpia, pero en ningún caso libre de eventos espurios.

En cuanto a la población de blank, se establece un corte en el valor de la RMS o media cuadrática del pulso, una medida de la dispersión temporal de la línea de base, que permite descartar pulsos de ruido eléctrico. Este corte contribuye a depurar la muestra de ruido y a mejorar la comparación con la población de señal.

En lo que respecta a la estadística, es importante señalar que, con el sistema ANOD disponible para todos los detectores, solo se tiene una única medida de neutrones, realizada durante 42 h. Esto supone, tras aplicar los precortes, un total de $N=46939$ eventos con $mult > 1$. Dado que el número de sucesos registrados en el blank con el sistema ANOD (desde su digitalización el 02/04/2024 y hasta el 20/05/2025) excede este valor, se selecciona aleatoriamente un subconjunto de igual tamaño para garantizar la comparación equitativa, como se describe en la fase de entrenamiento de la sección 3.2.

Sin embargo, ya que el blank está compuesto principalmente por sucesos rápidos de tipo Cherenkov fáciles de eliminar mediante el corte en la variable P_1^{ANOD} , se impone, en la medida de lo posible, que la muestra de ruido contenga el máximo número de sucesos lentos, definidos como aquellos con un valor $P_1^{ANOD} > 0.3$. Finalmente, la muestra de ruido está compuesta por un 70 % de sucesos rápidos y un 30 % de sucesos lentos, limitados por la estadística actual.

4.2. Variables de entrenamiento y optimización del número de árboles

Una vez seleccionados los sucesos de señal y de ruido, junto con sus respectivos precortes en la sección 4.1, se aplica el algoritmo configurado con los parámetros descritos en la sección 3.3, centrándose en la región energética [1–2] keV, especialmente conflictiva. El algoritmo no emplea la totalidad de los eventos disponibles para el entrenamiento, sino que reserva aleatoriamente una fracción de ellos como población de test (30 %), destinada a la detección de posibles casos de sobreentrenamiento o infraentrenamiento.

Ranking	Parámetro	Separación independiente (%)	Importancia de variable (%)
1	P_2^{ANOD}	67.09	7.421
2	$Asynphe^{ANODANA}$	54.38	6.937
3	P_1^{ANOD}	60.22	6.173
4	$n1^{ANODANA}$	59.75	5.699
5	$n0^{ANODANA}$	55.43	5.452
6	SM^{ANOD}	24.70	5.079
7	μ_p^{ANOD}	56.91	4.294
8	CAP_{800}^{ANOD}	36.43	3.878
9	$SM^{ANODANA}$	46.46	3.782
10	$n1tail$	7.106	3.584
11	$n0tail$	7.945	3.584
12	$\mu_p^{ANODANA}$	56.19	3.534
13	$tmax1^{ANOD}$	25.02	3.415
14	$Asynphe^{ANOD}$	12.51	3.278
15	CAP_{200}^{ANOD}	19.40	3.078
16	μ_{p0}^{ANOD}	54.40	3.077
17	CAP_{700}^{ANOD}	33.86	3.062
18	$Asyntail$	6.550	3.013
19	CAP_{50}^{ANOD}	49.40	2.871
20	$Tmax^{ANOD}$	29.33	2.821
21	CAP_{300}^{ANOD}	13.32	2.821
22	CAP_{500}^{ANOD}	24.71	2.433
23	CAP_{100}^{ANOD}	40.46	2.423
24	CAP_{600}^{ANOD}	29.97	2.257
25	μ_{p1}^{ANOD}	54.74	2.221
26	$tmax0^{ANOD}$	23.44	2.113
27	CAP_{400}^{ANOD}	18.09	1.702

Tabla 2: Porcentaje de separación de variables de manera independiente (previa al entrenamiento) y importancia de variables (posterior al entrenamiento), ordenadas según importancia.

Antes de llevar a cabo el entrenamiento completo, el algoritmo evalúa individualmente la capacidad de separación de cada variable, proporcionando un porcentaje que cuantifica su poder discriminante. A partir de estos resultados, se seleccionan únicamente aquellas variables que muestran una separación significativa, priorizando las que presentan un mayor potencial para distinguir entre señal y ruido.

Aunque el uso de un mayor número de variables no penaliza directamente el rendimiento del algoritmo más allá del aumento en el tiempo de cómputo, se opta por mantener un conjunto razonable de variables, favoreciendo la eficiencia y la interpretabilidad del modelo final. El conjunto final de variables elegidas junto con su porcentaje de separación inicial se recoge en la Tabla 2 (segunda y tercera colum-

na). El resto de variables relevantes para caracterizar un pulso se incluyen como variables espectadoras, que no participan en el entrenamiento, pero son interesantes para la posterior validación de este.

En la Figura 8 se muestra la distribución de las dos variables inicialmente más discriminantes en función de la energía: P_2^{ANOD} en la primera fila y P_1^{ANOD} en la segunda fila, comparando la distribución para la población de neutrones (primera columna), fondo (segunda columna) y blank (tercera columna). En ellas se observa la dependencia de las variables con la energía y la dificultad para separar señal de ruido, especialmente en el rango problemático [1–2] keV. Esta observación refuerza la decisión de no utilizar la energía como variable de entrenamiento, ya que no se quiere inducir sesgos debidos a su fuerte correlación con la probabilidad de ser señal o ruido, prefiriéndose evaluar el rendimiento del algoritmo en ventanas energéticas específicas una vez entrenado.

Por otro lado, la Figura 8 también remarca, como se ha discutido en la sección 4.1, la necesidad de aplicar el corte de centelleo en la variable P_1^{ANOD} , ya que la población de neutrones (teóricamente buena) contiene una fracción no despreciable de sucesos espurios. Esto se deduce por la comparativa con la distribución de sucesos del módulo blank para esta variable.

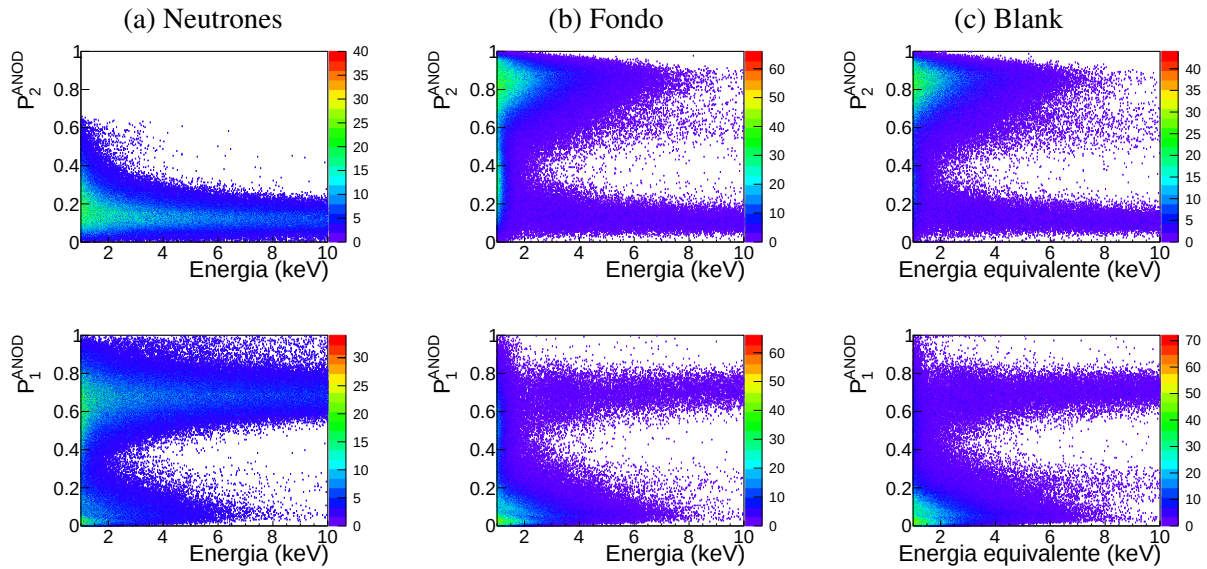


Figura 8: Distribución de las dos variables inicialmente más discriminantes en función de la energía: P_2^{ANOD} (fila superior) y P_1^{ANOD} (fila inferior), para las poblaciones de (a) neutrones, (b) fondo y (c) blank.

Finalizado el entrenamiento, se obtiene un ranking de variables que mejor separan los sucesos en función del número de árboles usados en el algoritmo. Para seleccionar el número óptimo de árboles ($nTrees$), se desarrolla una función que evalúa el rendimiento en términos del ritmo de fondo *corregido por la eficiencia* (denotado con R , diferente al ritmo R mostrado en gráficas anteriores, en las que no se corrige por la eficiencia) en la ventana [1–2] keV, bajo la condición de que la eficiencia promedio en [1–10] keV supere la alcanzada actualmente por ANAIS-112 en esa región (97.4%). Cabe destacar que la eficiencia se calcula como la eficiencia del corte $BDTCut$ (número de eventos de población de neutrones depurada que supera el corte dividido por su total) por la eficiencia de *trigger* (que tiene en cuenta la probabilidad de que un suceso de una determinada energía active la electrónica de adquisición y sea registrado).

Entonces, se aplica el entrenamiento sobre las medidas de fondo (desde la instalación completa de ANOD el 17/12/2024 y hasta el 20/05/2025) y la población de neutrones para cada número de árboles. En ambos casos, se aplica en la región [0–20] keV. Tras este, para la población de fondo se seleccionan

los sucesos con multiplicidad uno, pues se espera que la partícula de materia oscura interactúe una sola vez y de forma aislada dada su sección eficaz débil. Adicionalmente, se incorpora el corte de centelleo ($P_1^{ANOD} > 0.35$, $P_{10}^{ANOD} > 0.15$ y $P_{11}^{ANOD} > 0.15$) para el cálculo del ritmo, aunque no sería estrictamente necesario, puesto que la población de fondo ya hereda de manera efectiva los criterios de limpieza aplicados a la muestra de neutrones durante el entrenamiento.

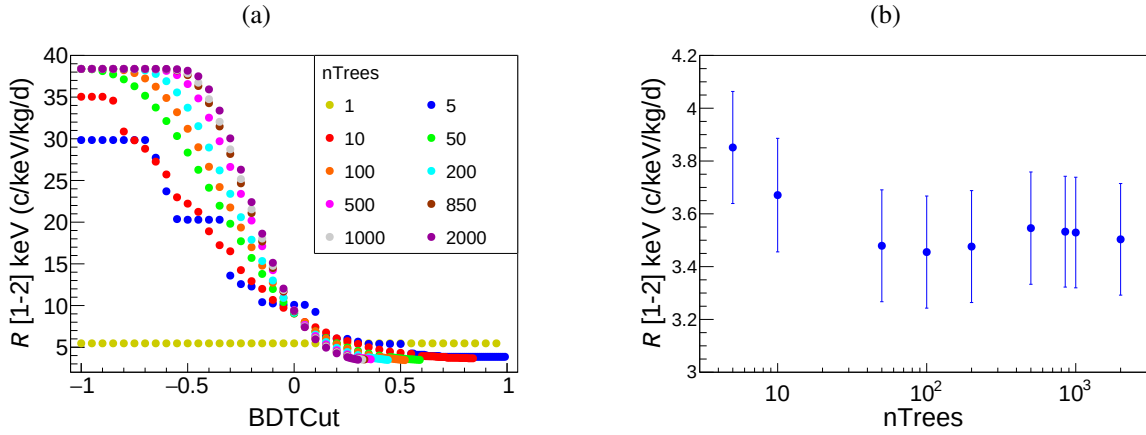


Figura 9: (a) Ritmo de fondo corregido por la eficiencia en función del corte en BDT, para distinto número de árboles. (b) Ritmo de fondo corregido por la eficiencia entre [1-2] keV para el cual la eficiencia alcanza el valor promedio de la eficiencia de ANAIS-112 en [1-10] keV. Se observa que a partir de $nTrees = 10$ todos los valores de ritmo corregido son compatibles.

Como es de esperar, en la Figura 9a se observa (a excepción de $nTrees=1$) que el ritmo de fondo corregido por la eficiencia decae según aumenta el valor en el corte hasta que se alcanza la eficiencia promedio de ANAIS en [1-10] keV (criterio de detención). Si se dibuja ese valor de ritmo en función del número de árboles (Figura 9b), se puede ver que se alcanza un *plateau*: a partir de $nTrees = 5$, todos los valores de ritmo son compatibles.

Se podría seleccionar por tanto este primer valor de número de árboles. Sin embargo, es conveniente elegir un número mayor, pues permite un abanico más amplio de valores de BDT, y por tanto una exploración más completa del valor del corte. Con este criterio, se escoge un valor $nTrees = 200$, que además es el que se utiliza en ANAIS. Si se incrementa aún más el número de árboles, los resultados obtenidos son equivalentes, pero el tiempo de computación aumenta y se incrementa también la probabilidad de subajuste debido a la estadística limitada.

Finalmente, se muestra para el valor de número de árboles elegido las variables de entrenamiento junto con el porcentaje de separación mencionado al principio, en la Tabla 2. La diferencia entre los porcentajes de separación de las variables individuales (antes del entrenamiento) y del conjunto total tras el entrenamiento es notable. Entre las variables con mayor capacidad de discriminación en el algoritmo *AdaBoost*, destacan la fracción del pulso en la cabeza y la asimetría en el número de fotoelectrones restringida a la ventana de ANAIS. Cabe señalar que la capacidad de separación de P_1 podría estar infraevaluada debido a que ya ha sido utilizada como precorte para neutrones y a que se ha forzado que el blank contenga un 30% de sucesos lentos.

4.3. Optimización del corte en BDT

Con el modelo entrenado y el número de árboles seleccionado, $nTrees = 200$, resulta útil analizar la distribución del discriminante BDT para las distintas poblaciones, como se muestra en la Figura 10: en (a) para [1-2] keV y en (b) para [2-3] keV. En ambas, los sucesos de tipo blank (ruido, en granate)

se concentran en valores bajos de BDT y los de neutrones (señal, en verde) en valores altos, mientras que el fondo (en naranja) aparece como una mezcla de ambos. Estas distribuciones orientan la elección del punto de corte en BDT, que debe eliminar el fondo similar al ruido sin reducir en exceso la señal (eficiencia). En (a), la superposición entre ruido y señal es considerable, lo que sugiere que buena parte del fondo en esta región energética es espurio. En cambio, en (b) la separación mejora, y el corte es menos crítico. Esto confirma de nuevo que la región más conflictiva para la clasificación se sitúa por debajo de 2 keV.

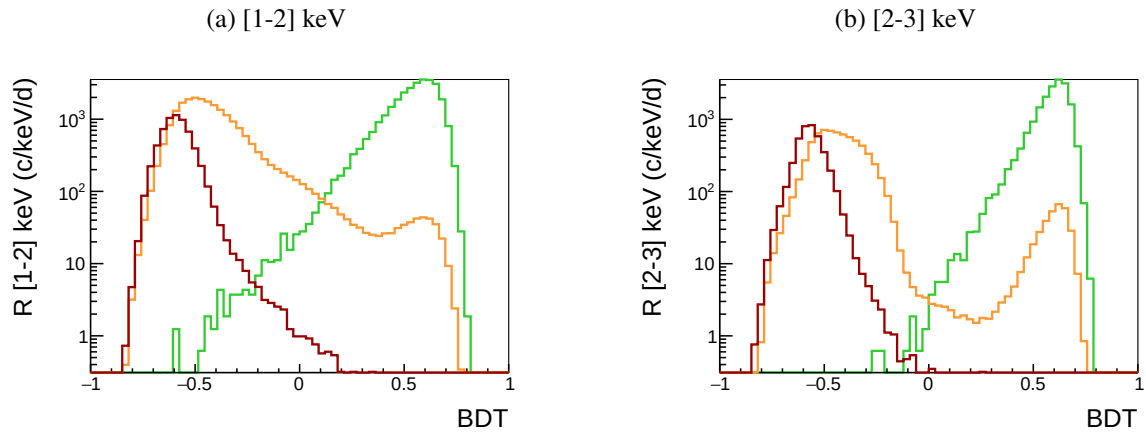


Figura 10: Distribución del discriminante BDT para distintas poblaciones: blank (ruido, en granate), neutrones (señal, en verde) y fondo (en naranja). En (a) [1-2] keV, la señal y el ruido se superponen, dificultando la discriminación; en (b) [2-3] keV, la separación mejora.

Por otro lado, como se observa en la Figura 8, la distribución de sucesos depende significativamente de la energía por debajo de 2 keV. Por lo tanto, es necesario establecer el corte en BDT de manera progresiva, ajustándolo por intervalos definidos con un bineado de 0.1 keV. Para energías superiores a 2 keV, el valor del corte se heredará del último intervalo de energía, es decir, del rango [1.9-2.0] keV.

Además, aunque en las figuras anteriores se han sumado las medidas de los nueve detectores, la contaminación de fondo varía individualmente en cada uno de ellos, lo que implica que el cálculo del corte debe realizarse de forma independiente para cada detector.

La manera de proceder, entonces, es la siguiente: dado un detector y un intervalo de energía, se calcula la eficiencia y el ritmo de fondo corregido por la eficiencia en función del valor del corte en BDT. En la Figura 11a se muestra el valor del ritmo de fondo corregido por la eficiencia obtenido para cada detector en el intervalo [1.0-1.1] keV, lo que refleja la alta variabilidad entre detectores mencionada anteriormente. En la Figura 11b se muestra la eficiencia y el ritmo de manera conjunta para el detector 1 en el intervalo [1.9-2.0] keV. La comparativa entre estas dos figuras también es prueba de la distinta distribución de sucesos en función de la energía.

Dado que en ningún caso se quiere que la eficiencia sea inferior a la de ANAIS, para cada detector e intervalo de energía, se calcula la eficiencia alcanzada en ANAIS (que se escoge constante durante los seis años de toma de datos) y se busca el valor de corte en BDT que proporcione esa eficiencia en ANOD. Una vez encontrado este valor de corte en BDT, se encuentra el ritmo asociado a él. Esto está representado de modo esquemático en la Figura 11b, mediante una línea roja vertical. El valor de ritmo asociado a la eficiencia queda remarcado en negro.

Se observa que un conjunto de valores de corte BDT conducen al mismo valor de ritmo dentro de su incertidumbre. Por tanto, llamando $BDT_{inicial}$ al que correspondiente a la eficiencia de ANAIS,

se retrocede en el valor de BDT con el objetivo de maximizar la eficiencia, seleccionando como corte BDT_{final} aquel que cumple $R(BDT_{final}) = R(BDT_{inicial}) + \delta R(BDT_{inicial})$.

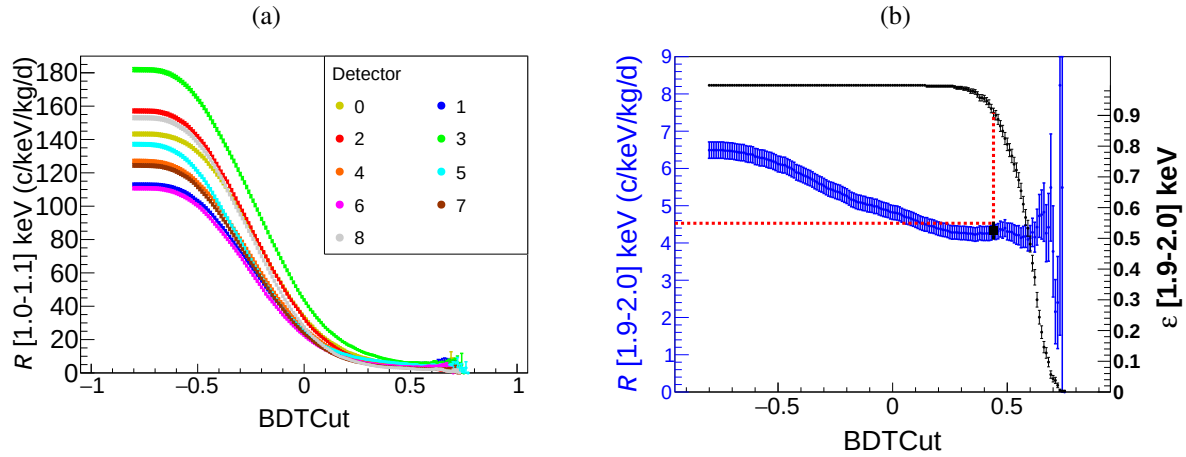


Figura 11: (a) Ritmo de fondo corregido por eficiencia entre [1.0-1.1] keV, para todos los detectores. Se observa que el fondo de cada detector es distinto. (b) Ritmo de fondo corregido por eficiencia (en azul) y eficiencia (en negro) para el detector 1, en la región [1.9-2.0] keV. Esquemáticamente muestra el proceso del $BDTCut$ óptimo. La línea roja vertical indica el valor del corte para el que la eficiencia de ANOD es igual a la de ANAIS en el intervalo de energía dado. Se toma el valor de ritmo con el mismo $BDTCut$ (punto negro), y se retroce en $BDTCut$ hasta el último valor compatible conforme al criterio $R_{final} > R_{inicial} + \delta R_{inicial}$ para aumentar la eficiencia (línea roja horizontal).

El algoritmo está diseñado para que, sin disminuir la eficiencia mínima promedio alcanzada en ANAIS, se obtenga el menor $BDTCut$ posible, que garantice que el ritmo resultante maximice la eficiencia. No obstante, este procedimiento puede resultar algo agresivo, dado que los valores finales de ritmo y eficiencia dependen en gran medida de la eficiencia seleccionada, además de la limitada estadística disponible ya mencionada. Por ello, se permite seleccionar un valor de corte en BDT que se encuentre dentro de un $\pm 10\%$ con respecto a los valores obtenidos mediante el procedimiento descrito.

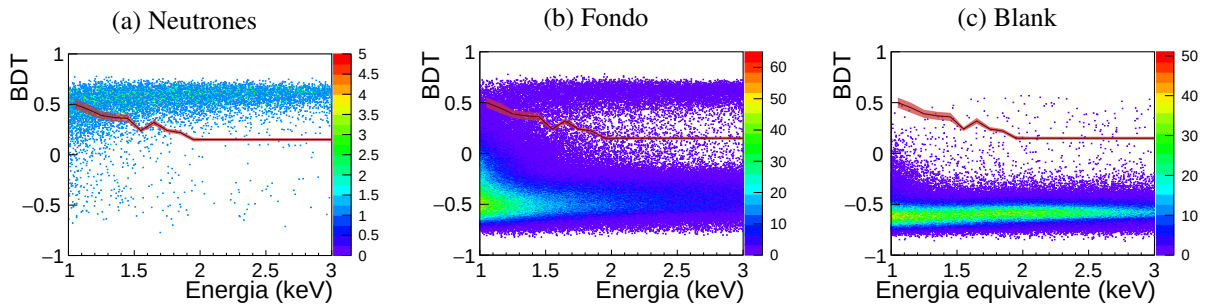


Figura 12: Distribución de la variable BDT en función de la energía para la población de (a) neutrones, (b) fondo y (c) módulo blank. La línea roja muestra el corte en BDT en función de la energía. En el intervalo [1-2] keV se observa la variación asociada a flexibilizar el corte en un $\pm 10\%$ (grosor en la línea $BDTCut$).

Como resultado, en la Figura 12 se muestra, tomando como ejemplo el detector 0, la distribución de la variable BDT en función de la energía, junto con el corte en BDT (representado por una línea roja). En esta figura se puede observar una separación adecuada entre los sucesos espurios (por debajo de la línea roja) y los de centelleo (por encima de la línea roja), por comparativa en la distribución de sucesos de calibración de neutrones (Figura 12a), de fondo (Figura 12b) y de blank (Figura 12c). Cabe destacar

que es apreciable, pero sutil, la flexibilización en el corte BDT (grosor en la línea $BDTCut$).

Además, en la Figura 12 puede observarse que la estadística asociada a la población de neutrones es considerablemente más reducida que la correspondiente a las poblaciones de blank y de fondo. No obstante, conviene recordar (como se explica en la sección 4.1) que la verdadera limitación estadística del análisis proviene de la muestra del blank, debido a la escasez de eventos con carácter lento en dicha población.

Esta conclusión no se deduce directamente de la figura mencionada, ya que en ella no se ha aplicado el corte de centelleo $P_1^{ANOD} > 0.35$, $P_{10}^{ANOD} > 0.15$ y $P_{11}^{ANOD} > 0.15$ ni a la población de fondo ni a la de blank. Esto se ha hecho con el fin de mostrar de forma más completa el potencial discriminante del corte en BDT. Cabe señalar que el corte final aplicado al espectro de fondo incluye tanto el corte en BDT como el de centelleo.

4.4. Evaluación del sobreentrenamiento e infraentrenamiento

Para comprobar que no se ha producido sobreentrenamiento ni infraentrenamiento, se muestra en la Figura 13, para cada detector y cada intervalo de energía con su respectivo corte en BDT, la eficiencia sobre la población de entrenamiento (en rojo) y sobre la población de test (en azul). Se ha añadido una sombra en rojo para el entrenamiento y en azul para el test que muestra la flexibilización del corte en BDT. Dado que los valores son compatibles, se confirma que no ha habido sobreentrenamiento y, puesto que los valores de eficiencia alcanzan la unidad, se confirma que no se ha producido infraentrenamiento.

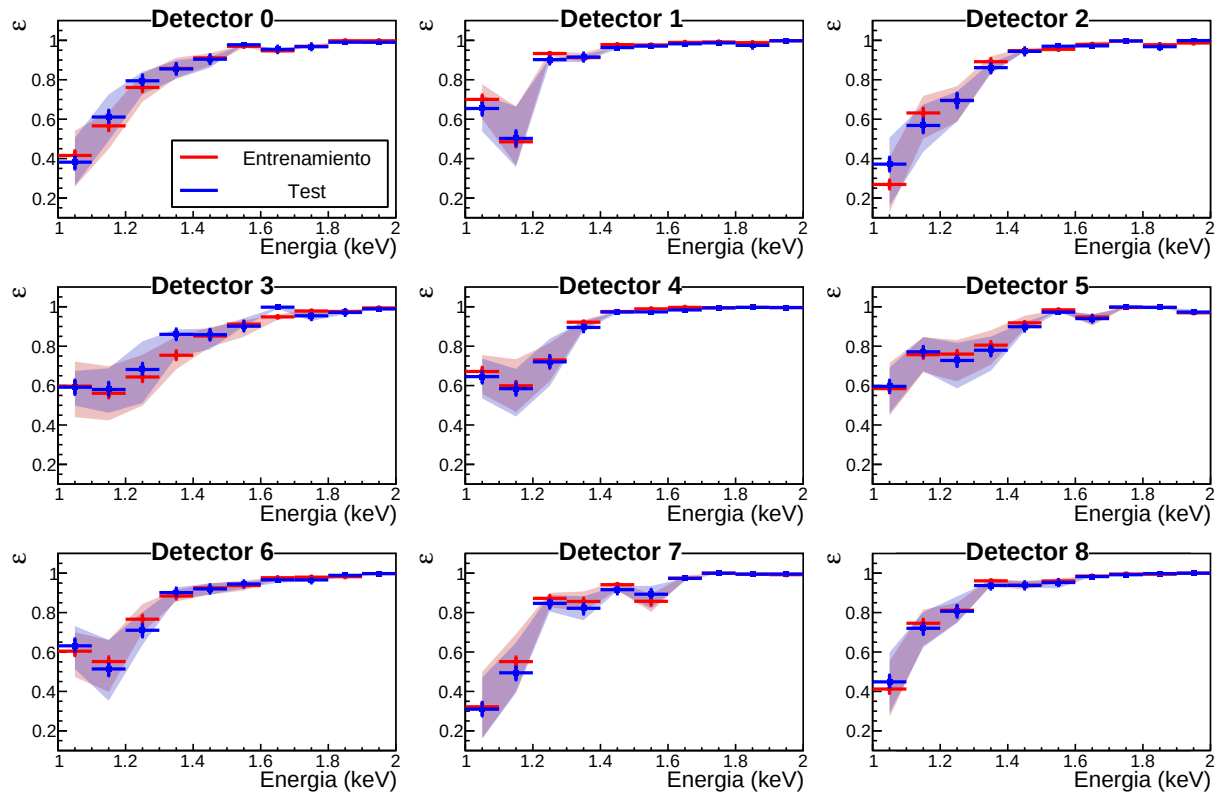


Figura 13: Eficiencia sobre la población de entrenamiento (en rojo) y sobre la población de test (en azul) para cada detector y cada intervalo de energía con su respectivo corte en BDT. La sombra en rojo para el entrenamiento y en azul para el test muestran la flexibilización del corte en BDT.

4.5. Validación del corte en BDT

La Figura 12 muestra una correcta separación entre eventos de centelleo y de ruido. No obstante, una validación adicional de este resultado consiste en representar algunas de las variables antes y después de aplicar el corte en BDT junto con el corte de centelleo. Esto se ilustra en la Figura 14, donde, para cada una de las cuatro variables identificadas como más discriminantes (ver Tabla 2, primera, segunda y cuarta columna), se representa la distribución de la población de neutrones (en verde) y de fondo (en naranja) antes (línea discontinua) y después del corte (línea continua). En dicha figura se observa: tras el corte, (a) se eliminan los pulsos de subida rápida; (b) el reparto de fotoelectrones entre los PMTs se vuelve simétrico; y (c) desaparece la población correspondiente a sucesos rápidos.

Estos cambios indican que se cumplen los requisitos esperados para eventos de centelleo reales, como lo demuestra la coincidencia de las poblaciones de fondo tras el corte con la de neutrones. Cabe destacar, además, que la población de neutrones apenas se ve afectada, lo que sugiere que el corte es eficiente.

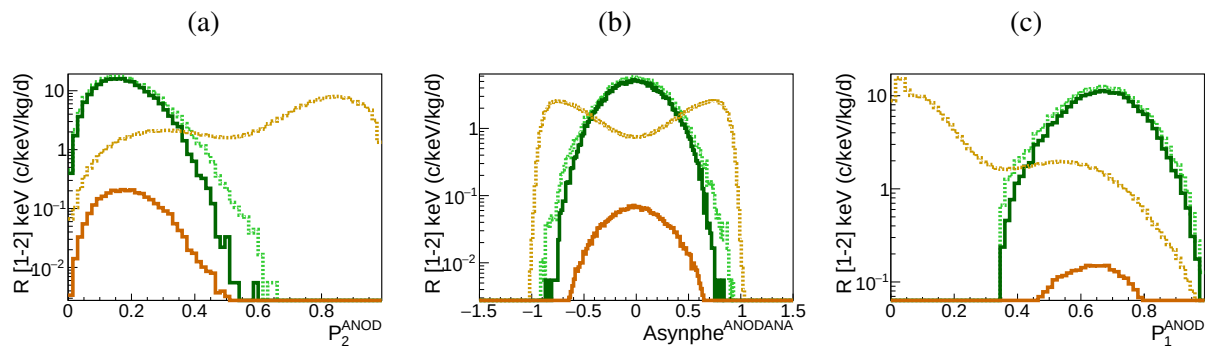


Figura 14: Representación de las distribuciones de la población de neutrones (en verde) y de fondo (en naranja) antes (línea discontinua) y después del corte en BDT (línea continua), para las tres variables más discriminantes del entrenamiento: (a) P_2^{ANOD} , (b) $Asynphe^{ANOD}$ y (c) P_1^{ANOD} .

5. Comparativa de los resultados de ANOD respecto de ANAIS

5.1. Filtrado del fondo y eficiencia

Una vez comprobado que el algoritmo no presenta ni sobreentrenamiento ni infraentrenamiento (sección 4.4), y validado el corte en BDT (sección 4.5) —es decir, verificado el correcto funcionamiento del algoritmo de aprendizaje automático—, en esta sección se analizan los resultados obtenidos por ANOD tras aplicar dicho algoritmo. Estos se comparan con los datos de ANAIS correspondientes al mismo periodo de tiempo respecto de la toma de datos de ANOD (año 7.5), utilizando el filtrado en BDT definido en [5], así como con la simulación.

En la Figura 15 se presentan los resultados promediados para los nueve detectores: en (a) se muestra el espectro de fondo en anticoincidencia en el intervalo de energía [1–4] keV, y en (b) la eficiencia de aceptación de sucesos. Los datos de ANOD se representan en negro, incluyendo la flexibilización del corte en BDT ($\pm 10\%$) como una banda sombreada en rojo, los de ANAIS en azul y la simulación en naranja.

Los resultados individuales de cada detector se detallan en las Figuras 16 y 17. En la primera, se compara el espectro de fondo en anticoincidencia en [1–4] keV, manteniendo la misma codificación de colores que en la figura promediada. La Figura 17 muestra la eficiencia de aceptación de sucesos para ambos métodos sistemas de adquisición en el mismo intervalo energético, también separada por detector y con la misma codificación de colores.

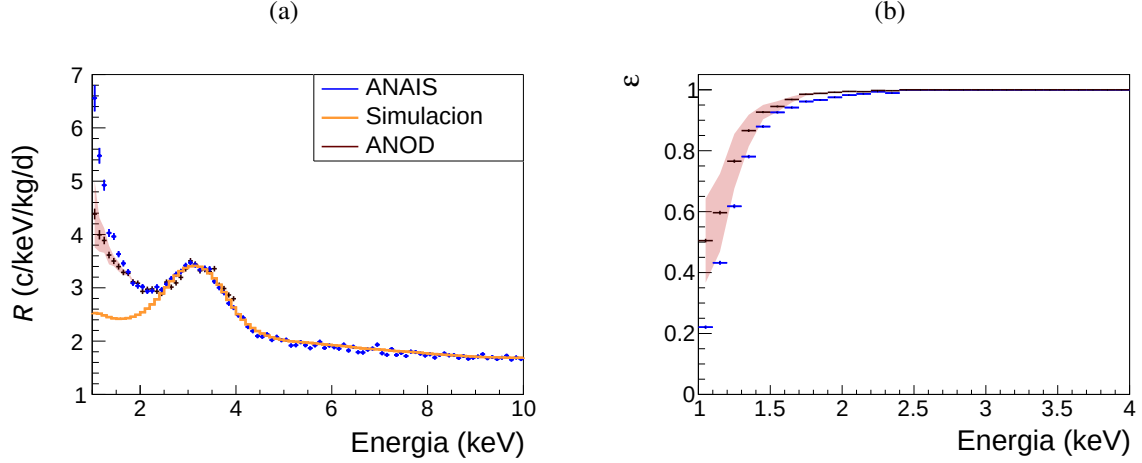


Figura 15: (a) Espectro de fondo en anticoincidencia promedio de todos los detectores en la región [1-4] keV. Se compara el fondo medido por ANOD (en negro, con una banda sombreada en rojo que representa una variación del $\pm 10\%$ en el corte en BDT) y por ANAIS (en azul) para el mismo periodo de tiempo de medidas de ANOD (año 7.5), y se muestra la simulación para ese mismo intervalo temporal (en naranja). (b) Eficiencia de aceptación de sucesos en la región [1-4] keV para ANOD (en negro) y ANAIS (en azul).

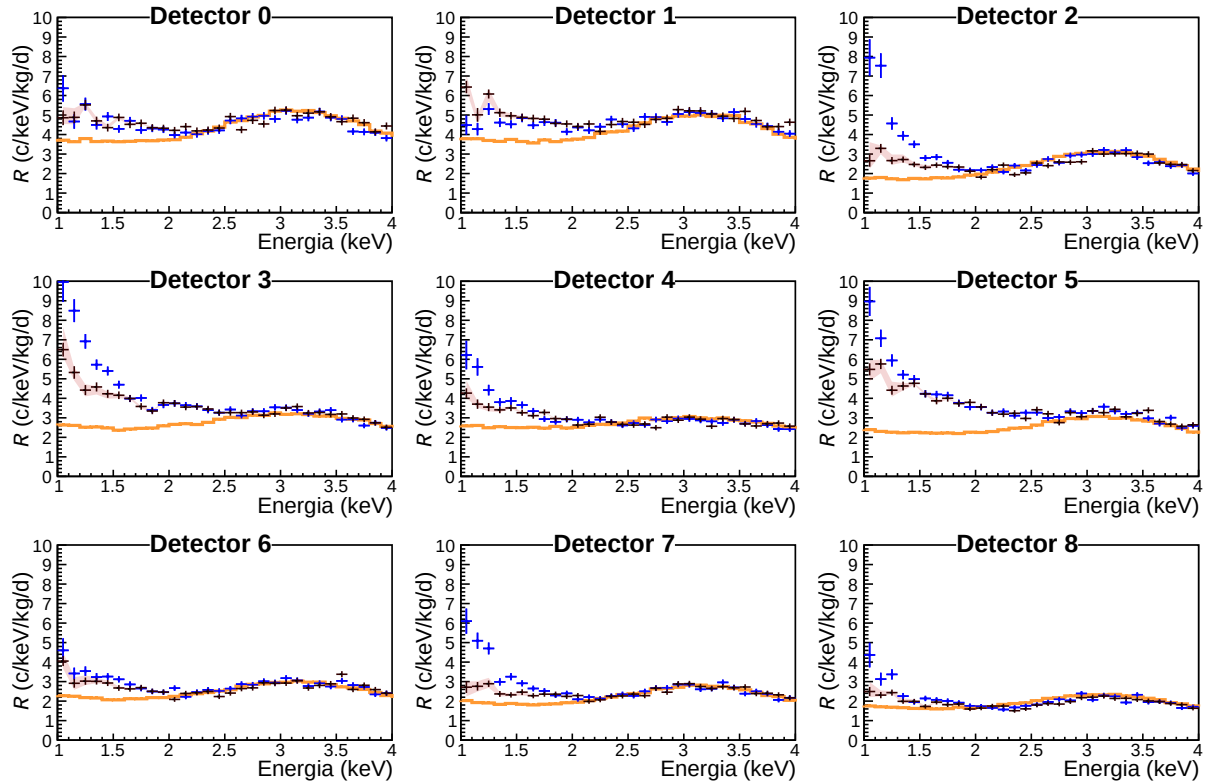


Figura 16: Espectro de fondo en anticoincidencia en el intervalo [1-4] keV para cada detector por separado. Se compara el fondo medido por ANOD (en negro, con una banda sombreada en rojo que representa una variación del $\pm 10\%$ en el corte en BDT) y por ANAIS (en azul) para el mismo periodo de tiempo de medidas de ANOD (año 7.5), y se muestra la simulación para ese mismo intervalo temporal (en naranja).

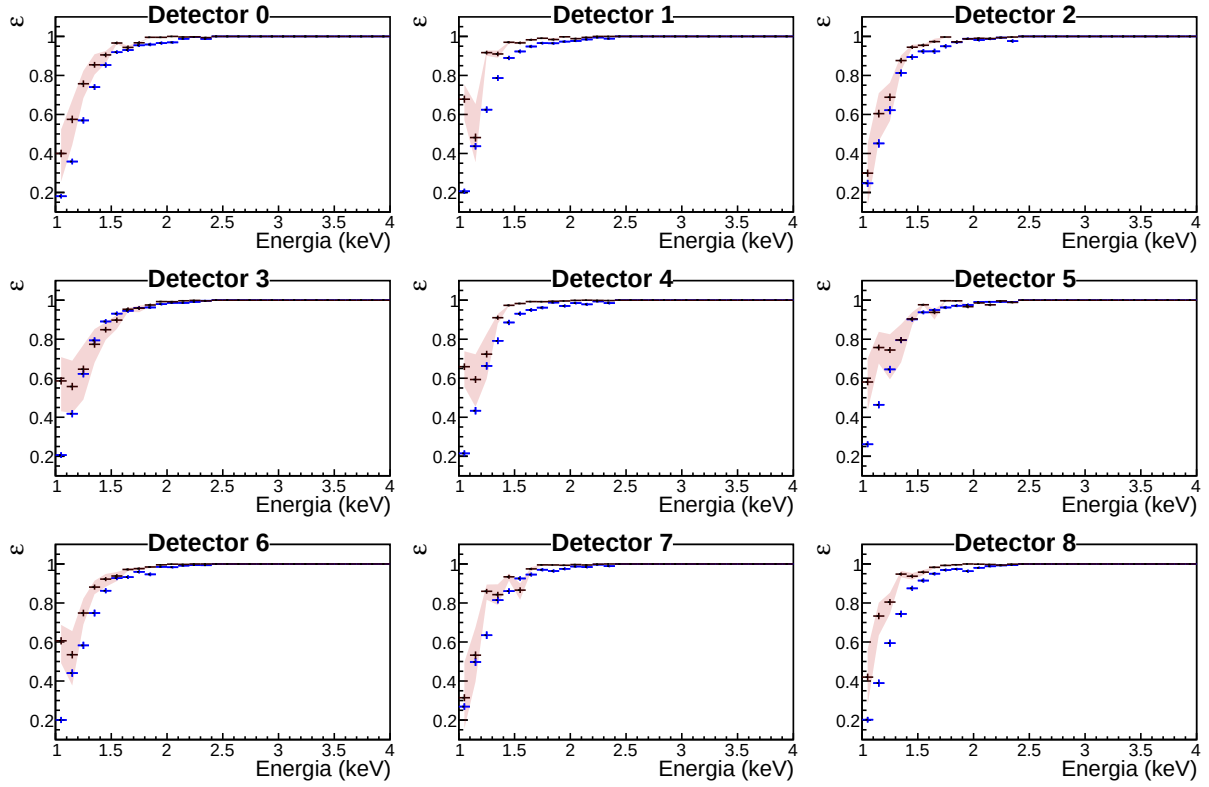


Figura 17: Eficiencia de aceptación de sucesos en el intervalo [1-4] keV para cada detector por separado. Se compara la eficiencia de ANOD (en negro, con una banda sombreada en rojo que representa una variación del $\pm 10\%$ en el corte en BDT) y por ANAIS (en azul) para el mismo periodo de tiempo de medidas de ANOD (año 7.5).

Detector	Fondo (c/keV/kg/d)			Eficiencia (%)		
	ANAIS	ANOD	v (%)	ANAIS	ANOD	v (%)
D0	4.78 ± 0.10	4.70 ± 0.08	-2	74.3 ± 0.4	83.6 ± 0.4	13
D1	4.59 ± 0.08	5.12 ± 0.07	12	77.2 ± 0.4	88.8 ± 0.3	15
D2	4.00 ± 0.13	2.54 ± 0.06	-37	77.8 ± 0.5	83.0 ± 0.4	7
D3	5.63 ± 0.14	4.39 ± 0.07	-22	77.1 ± 0.4	81.9 ± 0.4	6
D4	3.96 ± 0.10	3.40 ± 0.06	-14	77.9 ± 0.4	88.2 ± 0.3	13
D5	5.21 ± 0.11	4.47 ± 0.07	-14	78.6 ± 0.4	86.6 ± 0.4	10
D6	3.16 ± 0.08	2.89 ± 0.06	-9	75.9 ± 0.4	85.6 ± 0.4	13
D7	3.49 ± 0.09	2.47 ± 0.06	-29	78.6 ± 0.4	83.1 ± 0.4	6
D8	2.50 ± 0.08	2.01 ± 0.05	-19	75.7 ± 0.4	87.7 ± 0.3	16
Promedio	4.15 ± 0.03	3.55 ± 0.02	-14	77.01 ± 0.14	85.38 ± 0.13	11

Tabla 3: Fondo y eficiencia promedio en la región conflictiva [1-2] keV para cada detector, así como el valor promedio de los nueve detectores. Se incluye el porcentaje de variación de los resultados de ANOD respecto de los de ANAIS. Los datos son resultado de la aplicación del valor de corte en BDT central.

Para facilitar la comprensión de los resultados se recoge en la Tabla 3 el fondo y eficiencia promedio en la región conflictiva [1-2] keV para cada detector para ANAIS y ANOD, así como los valores promedio de los nueve detectores. La tabla también muestra el porcentaje de variación de los resultados de ANOD respecto de los de ANAIS. Los datos son resultado de la aplicación del valor de corte en BDT central.

A la vista de estos, se aprecia una elevada variabilidad entre los distintos detectores. Esta dispersión se debe, en parte, a las diferencias intrínsecas entre ellos, aunque la causa principal es la limitada estadística disponible. Esta limitación es consecuencia directa de la reciente puesta en marcha del nuevo sistema de adquisición, unida, como es habitual en las fases iniciales de instalación, con diversos problemas que han obligado a descartar parte de los datos.

Por un lado, esto refuerza la elección de explorar los resultados asociados a cortes en el valor BDT ligeramente más y menos restrictivos, que se muestran sombreados en las gráficas. Por otro lado, hace que resulte más adecuado centrarse en los resultados promediados, en lugar de interpretar los individuales de forma aislada. En este sentido, se observa que el balance global es positivo entre 1 y 2 keV: se consigue una reducción media del fondo del 14 % y un incremento medio de la eficiencia del 11 % respecto de ANAIS. Cabe destacar que, cerca del umbral del experimento, a 1 keV, la mejora es mucho más notable, habiéndose reducido el fondo de ANOD hasta un 40 % con respecto al de ANAIS, y aumentado la eficiencia desde un 20 % en ANAIS hasta un 50 % en ANOD. Esto puede visualizarse en la Figura 15, quedando reflejado el potencial del estudio.

5.2. Reducción del umbral del experimento

El umbral de ANAIS-112 se estableció en 1 keV debido a su limitada eficiencia, del orden del 15-20 %, cuando todavía se utilizaban cortes simples para la selección de sucesos [8]. La posterior incorporación de técnicas de aprendizaje automático al sistema de adquisición permitió mejorar esta eficiencia, alcanzando valores mínimos inicialmente en torno al 25–30 %, aunque posteriormente hubo que implementar cambios en la calibración en energía y bajó ligeramente, al 22 % observado en la Figura 1b.

En cambio, en la sección 5.1 se ha demostrado que la aplicación de métodos de aprendizaje automático a los datos de ANOD produce una mejora sustancial, alcanzándose una eficiencia media cercana al 40 % a 1 keV. Este resultado respalda la decisión de explorar tanto la eficiencia como el ritmo de fondo por debajo del umbral convencional, con el objetivo de aumentar la sensibilidad frente a posibles señales de materia oscura.

En consecuencia, se propone reducir el umbral del experimento hasta los 0.7 keV. El procedimiento seguido para estimar el ritmo de fondo y la eficiencia en esta región se basa en el descrito en la sección 4.3, con una ligera modificación. Dado que no se dispone de datos de eficiencia de ANAIS por debajo de 1 keV, se ha tomado como punto de partida para el algoritmo de optimización la eficiencia mínima alcanzada en ANAIS en sus primeros pasos de selección de eventos (cortes simples). Esto es un 15 % en el caso más restrictivo.

Cabe señalar que, para este análisis, se han empleado los modelos previamente entrenados en la región [1-2] keV. Esto queda justificado a través de la Figura 18, en la que se muestra que los sucesos en [0.7-1.0] keV se distribuyen de forma similar a lo que se observa en [1-2] keV. En esta misma figura se muestra el valor de corte en BDT óptimo con una línea roja, cuyo grosor representa la variación en un $\pm 10\%$ respecto del valor de corte elegido.

Los resultados obtenidos tras el filtrado se presentan en la Figura 19. Como se observa en la Figura 19b, la eficiencia promedio de los datos de ANOD se mantiene en torno al 20 % en el intervalo de

energía a 0.7 keV, valor que coincide con la eficiencia mínima alcanzada por ANAIS en la región umbral original de 1 keV. Sin embargo, la Figura 19a muestra un aumento acusado del ritmo de fondo en ese mismo intervalo, generando una discrepancia significativa respecto al valor simulado. En concreto, para la región [1-2] keV se tiene una discrepancia del fondo medido en ANOD respecto al simulado del 44 %, mientras que para la región [0.7-1.0] keV, se observa una diferencia mucho mayor, lo que apunta a que el filtrado de sucesos es insuficiente en dicha región de energía.

Esto se interpreta como una consecuencia de no haber analizado en profundidad los sucesos en esta región energética, tradicionalmente excluida por estar bajo el umbral experimental. Es precisamente la mejora en la eficiencia lo que ahora permite, por primera vez, plantear un estudio detallado de los eventos por debajo de 1 keV. Así, aunque los resultados actuales aún no son concluyentes, abren la puerta a una futura caracterización precisa de esta región, con el objetivo de evaluar su viabilidad en la búsqueda de señales de materia oscura.

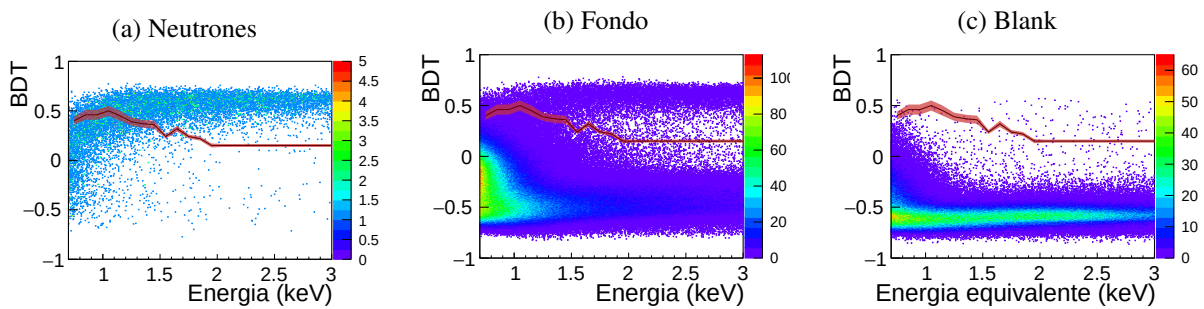


Figura 18: Distribución de la variable BDT en función de la energía para la población de (a) neutrones, (b) fondo y (c) módulo blank. La línea roja muestra el corte en BDT en función de la energía. Se observa que los sucesos en [0.7-1.0] keV se distribuyen de forma similar a lo observado por encima de 1 keV. Se muestra el valor de corte en BDT óptimo con una línea roja, cuyo grosor representa la variación en un $\pm 10\%$ respecto del valor de corte elegido.

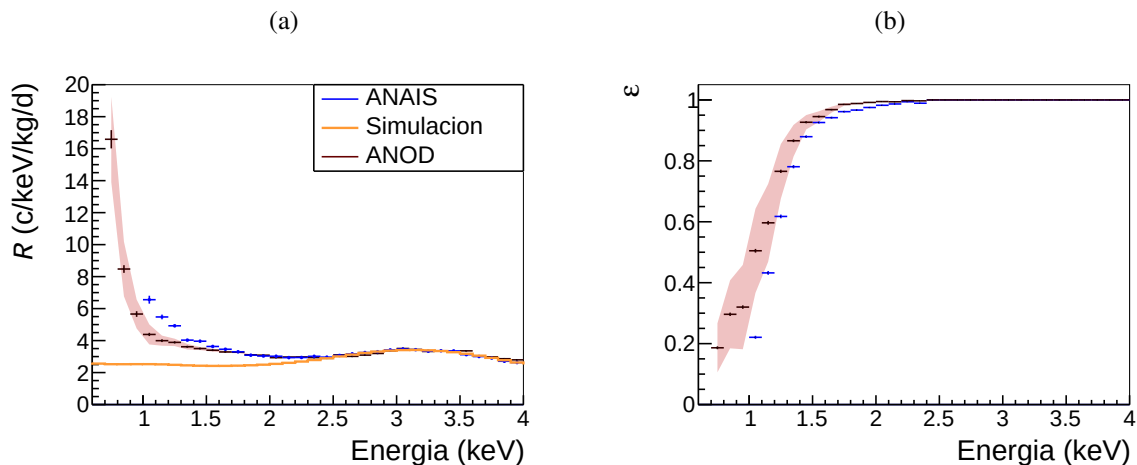


Figura 19: (a) Espectro de fondo en anticoincidencia promedio de todos los detectores en la región [0.7-4.0] keV. Se compara el fondo medido por ANOD (en negro, con una banda sombreada en rojo que representa una variación del $\pm 10\%$ en el corte en BDT) y por ANAIS (en azul) para el mismo periodo de tiempo de medidas de ANOD (año 7.5), y se muestra la simulación para ese mismo intervalo temporal (en naranja). (b) Eficiencia de aceptación de sucesos en la región [0.7-4.0] keV para ANOD y ANAIS.

6. Conclusiones y líneas de mejora

La discriminación entre sucesos de señal y eventos espurios constituye uno de los principales retos en los experimentos de detección directa de materia oscura. En el contexto del experimento ANAIS-112, este desafío adquiere una relevancia particular, ya que por debajo de 2 keV se observa una discrepancia entre el ritmo de sucesos observado y el esperado a partir de una simulación, que tiene en cuenta las distintas contaminaciones radiactivas presentes en los detectores y los materiales que los rodean. Aunque ANAIS-112 no ha registrado hasta la fecha una señal compatible con la modulación anual esperada, su comparación definitiva con DAMA/LIBRA sigue condicionada por posibles mejoras en el análisis de estos eventos de baja energía. Con esta motivación, se desarrolló e instaló ANOD, un sistema de adquisición de datos mejorado con respecto al de ANAIS. Su mayor ventana de adquisición y la ausencia de tiempo muerto permiten una caracterización más precisa de los pulsos, lo que mejora de forma significativa la capacidad de discriminación, especialmente frente a sucesos con reparto de luz asimétrico entre los fotomultiplicadores, que resultan difíciles de distinguir de los centelleos genuinos en ANAIS.

Este trabajo ha estudiado el potencial de ANOD mediante el uso de técnicas de aprendizaje automático, concretamente con el algoritmo *AdaBoost*, una implementación de árboles de decisión (*Boosted Decision Trees*) previamente validada en el marco de ANAIS. El análisis se ha centrado en la optimización de dos parámetros clave: el número de árboles del clasificador y el valor del corte sobre la variable BDT. Como resultado, se observan mejoras notables en cuanto a reducción de fondo (en un 14 %) y aumento de eficiencia (en un 11 %), al comparar ANOD con ANAIS en condiciones análogas en la región conflictiva ([1-2] keV). Además, en el umbral del experimento, actualmente en 1 keV, ANOD mantiene una eficiencia media en torno al 50 % frente al 20 % alcanzado previamente por ANAIS, además de una reducción del fondo del 40 % con respecto a ANAIS.

No obstante, se debe tener en cuenta que los datos filtrados de ANOD en esta región conflictiva todavía difieren respecto de los que predice la simulación. Por ello, resulta imprescindible seguir optimizando el análisis de sucesos en la región. Esto incluye el desarrollo de nuevas variables específicas del sistema ANOD, con mayor poder de discriminación, que permitan mitigar uno de los principales problemas asociados al entrenamiento de modelos supervisados: la posible contaminación de la población de señal con eventos espurios.

Por un lado, el aumento en la calidad de los datos abre la posibilidad, en un futuro, de aplicar clasificadores entrenados con ANOD a un reanálisis de los datos de ANAIS, gracias a la sincronización temporal entre DAQs. Por otro lado, esta mayor calidad permite plantear de forma realista la posibilidad de reducir el umbral del experimento. A modo de prueba de esto último, en este trabajo se ha explorado la región [0.7-1.0] keV, utilizando el algoritmo entrenado en la región [1-2] keV. Aunque esta aproximación aún no es óptima, los resultados sugieren que la mejora en eficiencia posibilita, por primera vez, el análisis de una región tradicionalmente excluida.

En conclusión, los primeros datos medidos y analizados del sistema ANOD con su instalación completada revelan, pese a la limitación estadística actual asociada a llevar únicamente medio año de toma de datos, una mejora sustancial respecto al sistema de adquisición de ANAIS. Como consecuencia, con el tiempo se espera que la acumulación de estadística y un análisis más exhaustivo de los eventos registrados permitan una mayor reducción del fondo y e incremento en la eficiencia. Esto será crucial para evaluar con mayor sensibilidad y robustez la posible señal observada por DAMA/LIBRA, e incluso explorarla a energías inferiores al umbral actual.

Referencias

- [1] F. Zwicky, “Die Rotverschiebung von extragalaktischen Nebeln,” *Helv. Phys. Acta* **6** (1933), 110-127 doi:10.1007/s10714-008-0707-4
- [2] N. Aghanim *et al.* [Planck], “Planck 2018 results. VI. Cosmological parameters,” *Astron. Astrophys.* **641** (2020), A6 [erratum: *Astron. Astrophys.* **652** (2021), C4] doi:10.1051/0004-6361/201833910 [arXiv:1807.06209 [astro-ph.CO]].
- [3] L. Baudis, “Dark matter detection,” *J. Phys. G* **43** (2016) no.4, 044001 doi:10.1088/0954-3899/43/4/044001
- [4] R. Bernabei *et al.*, “Further results from DAMA/Libra-phase2 and perspectives,” *Nucl. Phys. Atom. Energy* **22** (2021) no.4, 329-342 doi:10.15407/jnpae2021.04.329
- [5] J. Amaré *et al.*, “Towards a robust model-independent test of the DAMA/LIBRA dark matter signal: ANAIS-112 results with six years of data,” [arXiv:2502.01542 [astro-ph.IM]].
- [6] I. Coarasa *et al.*, “Improving ANAIS-112 sensitivity to DAMA/LIBRA signal with machine learning techniques,” *JCAP* **11** (2022), 048 [erratum: *JCAP* **06** (2023), E01] doi:10.1088/1475-7516/2022/11/048 [arXiv:2209.14113 [astro-ph.IM]].
- [7] D. M. Pascual, “Análisis de datos del experimento ANAIS-112 con un sistema de adquisición de datos mejorado,” Trabajo Fin de Grado, Grado en Física, Universidad de Zaragoza (2024). Dirigido por I. Coarasa y M. L. Martínez. [zaguan:152696]
- [8] J. Amaré *et al.*, “Performance of ANAIS-112 experiment after the first year of data taking,” *Eur. Phys. J. C* **79** (2019) no.3, 228 doi:10.1140/epjc/s10052-019-6697-4
- [9] J. I. Collar, “Quenching and channeling of nuclear recoils in NaI(Tl): Implications for dark-matter searches,” *Phys. Rev. C* **88**, no.3, 035806 (2013) [arXiv:1302.0796 [physics.ins-det]].
- [10] D. Cintas *et al.*, “Measurement of the sodium and iodine scintillation quenching factors across multiple NaI(Tl) detectors to identify systematics,” *Phys. Rev. C* **110**, no.1, 014613 (2024) [arXiv:2402.12480 [hep-ex]].
- [11] G. F. Knoll, “Radiation Detection and Measurement, 3rd ed.,” John Wiley and Sons, 2000, ISBN 978-0-471-07338-3
- [12] Y. Coadou, *Decision trees*, EPJ Web Conf. **4** (2010), 02003, doi:10.1051/epjconf/20100402003.
- [13] A. Hocker *et al.* [TMVA], “TMVA - Toolkit for Multivariate Data Analysis,” [arXiv:physics/0703039 [physics.data-an]].