

Facultad de Ciencias

Propiedades en el infrarrojo lejano de cuásares de
líneas anchas

(Far-infrared properties of broad-line quasars)

Trabajo de Fin de Grado

para obtener el

Grado de Física

Autor: Luka Giannotti San Juan

Director: Francisco Jesús Carrera Troyano

Curso 2023-2024

Índice General

Resumen	1
1. Introducción	3
1.1 Importancia de los cuásares en el estudio del universo	3
1.2 Metodología y enfoque del trabajo	4
1.3 Objetivo del estudio	5
2. Fundamentos Teóricos	7
2.1. Cuásares: características generales	7
2.2. Características de las bandas utilizadas	8
2.2.1. Óptico	8
2.2.2. Rayos X.....	9
2.2.3. FIR	10
3. Metodología	11
3.1. Selección y descripción de las fuentes de datos	11
3.2. Técnicas de análisis de datos	14
3.3. Construcción del diagrama de Venn	15
4. Resultados	17
4.1. Obtención de las zonas del Diagrama de Venn	17
4.2. Redshift y luminosidad	19
4.2.1. Redshift	21
4.2.2. Luminosidad	24
4.3. Flujo en FIR	26
4.4. Propiedades espectrales X	29
4.4.1. Flujo	29
4.4.2. Densidad Columna (N_H)	30
4.4.3. Índice fotónico	32
4.4.4. Resultados de K-S	33
5. Discusión	35
5.1. Interpretación de los resultados	35
5.2. Comparación con estudios previos	35
5.3. Implicaciones para futuros estudios	36
6. Conclusiones	37
7. Bibliografía	39
8. Apéndice	42

Agradecimientos: Me gustaría agradecer a mi familia y a Adela por el apoyo durante todos estos años y también a Francisco por la paciencia y la buena disponibilidad que ha tenido siempre.

Resumen

Los cuásares son de extrema importancia para la comprensión del universo en sus etapas más jóvenes. En este estudio se tiene como objetivo comprender si existen diferencias significativas en la emisión FIR entre cuásares con distintas propiedades espectrales X, como habían encontrado estudios anteriores. Este objetivo se ha realizado mediante comparaciones de tres muestras principales de cuásares detectados en el óptico, en rayos X y en IR y analizando las propiedades espectrales de estos.

Por un lado, se ha concluido que para el redshift y para la luminosidad se tiene un sesgo proveniente de la limitación de las técnicas utilizadas para obtener el redshift que impide estudiar estas propiedades con precisión. Por otro lado, la comparación directa para las propiedades espectrales X entre cuásares detectados y no detectados en FIR indica que no hay evidencias de que las distribuciones sean diferentes entre sí.

Abstract

Quasars are of extreme importance for understanding the universe in its earliest stages. This study aims to determine whether there are significant differences in FIR emission between quasars with different X spectral properties, as previous studies have found. This objective has been achieved by comparing three main samples of quasars detected in optical, X-rays, and IR and analyzing their spectral properties.

On one hand, it has been concluded that for redshift and luminosity, there is a bias stemming from the limitations of the techniques used to obtain the redshift, which prevents studying these properties accurately. On the other hand, direct comparison of X spectral properties between quasars detected and not detected in FIR indicates that there is no evidence that the distributions are different from each other.

Palabras clave: Cuásar, Espectro, Redshift, Rayos X, Infrarrojo lejano

Key Words: Quasar, Spectrum, Redshift, X-ray, Far-Infrared (FIR)

Capítulo 1

Introducción

1.1 Importancia de los cuásares en el estudio del universo

Un Núcleo Galáctico Activo (AGN, por sus siglas en inglés) es una región compacta en el centro de una galaxia que emite cantidades significativas de energía a través de procesos no estelares. Los cuásares son una forma extrema de AGN, caracterizados por su intensa luminosidad y emisión de energía. Son objetos extremadamente luminosos y distantes que se encuentran en los centros de algunas galaxias y su energía proviene de agujeros negros supermasivos que se alimentan de materia circundante. A medida que la materia cae hacia el agujero negro, se calienta y emite una cantidad enorme de radiación que abarca desde ondas de radio hasta rayos X y gamma, teniendo en cuenta que las ondas de radio y los rayos gamma no proceden del material candente, sino de material relativista, es decir, material que se mueve a velocidades cercanas a las de la luz, que interacciona con campos magnéticos. [4]

Debido a su gran distancia, los cuásares nos permiten observar el universo tal como era en sus primeras etapas. La luz que recibimos de ellos ha viajado miles de millones de años, ofreciendo una visión del pasado cósmico y permitiéndonos estudiar cómo eran las galaxias y su entorno en el universo primitivo [4]. Según estudios previos, los cuásares están íntimamente relacionados con la evolución de las galaxias [1][2][3]. Estudiar los cuásares nos ayuda a entender cómo los agujeros negros supermasivos influyen en la formación y evolución de las galaxias.

Además, los cuásares nos permiten estudiar la distribución de gas en el universo. Su intensa radiación ilumina y revela las propiedades del gas intergaláctico, incluyendo la estructura, lo cual es vital para comprender los procesos de formación de estructuras a gran escala en el cosmos .

También permiten estudiar la evolución química del universo al proporcionar información sobre la abundancia de elementos en diferentes épocas. Esto es crucial para comprender la formación de estrellas y galaxias y cómo ha cambiado la composición química del universo con el tiempo. [5]

En resumen, los cuásares son fundamentales para la cosmología moderna no solo por permitirnos ver el universo en sus primeras etapas, sino también por su papel en entender la evolución de las galaxias, la distribución de materia en el cosmos y la evolución química del universo.

1.2 Metodología y enfoque del trabajo

El enfoque comparativo en el estudio de los cuásares implica analizar datos obtenidos en diferentes bandas del espectro electromagnético, tales como rayos X, infrarrojo lejano (FIR, del inglés Far-Infrared) y óptico. Este enfoque permite obtener una visión más completa de los cuásares, ya que cada banda proporciona información única sobre diferentes aspectos del objeto estudiado. Entre las ventajas del enfoque comparativo se encuentra la multidimensionalidad de los datos, permitiendo obtener información sobre los diversos procesos físicos que ocurren en los cuásares. Por ejemplo, los rayos X nos informan sobre los procesos de alta energía cerca del agujero negro supermasivo, mientras que las observaciones en el infrarrojo pueden revelar la presencia de polvo y gas frío cerca del agujero negro supermasivo o en la galaxia anfitriona [6]. Además, la recopilación de datos en varias bandas permite la verificación cruzada de los resultados, aumentando la fiabilidad de las conclusiones.

Se ha procedido con una muestra de cuásares seleccionados en el óptico, de los cuales se ha estudiado su emisión en rayos X y en IR.

Los rayos X son cruciales para estudiar los procesos de altas energías cerca del agujero negro supermasivo. Procesos como la acreción, que se debe a la acumulación de materia en un objeto masivo debido a la gravedad. En el caso de los cuásares, el material forma un disco de acreción alrededor del agujero negro, y la fricción dentro del disco calienta el material, emitiendo una gran cantidad de energía en forma de radiación. Además, los AGN son fuentes poderosas de rayos X, siendo posible revelar agujeros negros gracias a su estudio. [6]

Los AGN también pueden influir significativamente en sus galaxias anfitrionas al liberar enormes cantidades de energía en forma de radiación y vientos, afectando el medio interestelar. La energía liberada puede calentar el gas circundante, inhibiendo la formación de nuevas estrellas, o comprimir el gas y desencadenar episodios de formación estelar. El infrarrojo es sensible al gas frío, permitiendo estudiar la estructura y la composición de la galaxia anfitriona. De esta manera se puede entender como el AGN afecta a la galaxia en su conjunto. La emisión IR está asociada a regiones de intensa formación estelar pero lo que se observa no es la emisión del cuásar directa, sino que el emisor es el polvo calentado por las estrellas jóvenes.

Al determinar la relación entre estas emisiones y cómo varían con las propiedades espectrales de los cuásares en rayos X, se espera obtener una comprensión más profunda de los mecanismos de retroalimentación que regulan el crecimiento del agujero negro y la formación de estrellas en las galaxias. Esto contribuirá significativamente a la teoría de la coevolución de los agujeros negros supermasivos y sus galaxias anfitrionas, proporcionando nuevas perspectivas sobre la dinámica y evolución de las galaxias a lo largo del tiempo.

1.3 Objetivo del estudio

Anteriormente se ha observado que hay una relación entre las propiedades espectrales en X y la emisión en FIR. Analizar cómo las propiedades espectrales en rayos X de los AGN correlacionan con la emisión en FIR puede proporcionar información sobre cómo la actividad del AGN influye en la formación estelar y, en consecuencia, en la evolución de la galaxia. [1][2][3]

El objetivo de este estudio es investigar, con una muestra mayor y más reciente, si estas observaciones se siguen dando, estudiando si existen diferencias significativas en la emisión en el infrarrojo lejano (FIR) entre cuásares con distintas propiedades espectrales en rayos X. Al explorar estas diferencias, se pretende entender de una forma más clara y fiable la coevolución de los agujeros negros supermasivos situados en los centros de las galaxias y sus respectivas galaxias anfitrionas.

Capítulo 2

Fundamentos Teóricos

2.1 Cuásares: características generales

Los cuásares son objetos celestes extremadamente luminosos y distantes, situados en los centros de galaxias activas, alimentados por agujeros negros supermasivos. Estos agujeros negros, con masas que varían desde millones hasta miles de millones de veces la masa del Sol, atraen material de su entorno formando discos de acreción. El proceso de acreción es fundamental para la energía que emiten los cuásares: el material en el disco se calienta enormemente debido a la fricción y a la compresión gravitacional, emitiendo radiación en múltiples longitudes de onda, desde el radio hasta los rayos X y gamma (recordar que las emisiones en ondas de radio y los rayos gamma provienen del material relativista que interacciona con campos magnéticos). Esta radiación es lo que hace que los cuásares sean tan extremadamente luminosos, superando en brillo a todas las estrellas de su galaxia anfitriona combinadas [5].

Además de los discos de acreción, muchos cuásares presentan jets de partículas que son expulsados a velocidades cercanas a la de la luz. Estos jets pueden extenderse por miles de años luz y son una característica notable de los cuásares, contribuyendo significativamente a su emisión en radio y otras partes del espectro electromagnético. El entorno de los cuásares también incluye grandes cantidades de gas y polvo, creando un medio polvoriento que puede absorber y reemitir parte de la radiación emitida por el disco de acreción, afectando así su observación desde la Tierra. Este polvo puede calentar y emitir en FIR, proporcionando información adicional sobre la estructura y la composición de la galaxia anfitriona.

Espectralmente, los cuásares muestran una variedad de características distintivas. Las líneas de emisión en sus espectros son generadas por el gas caliente y ionizado cerca del agujero negro. Estas líneas son cruciales para determinar la composición química y las condiciones físicas del gas circundante.

Estos también son conocidos por su variabilidad, con cambios en su luminosidad que pueden ocurrir en escalas de tiempo que van desde días hasta años. Esta variabilidad es una herramienta poderosa para estudiar la estructura interna del disco de acreción y la dinámica del material en el entorno inmediato del agujero negro. Además, los cuásares presentan desplazamientos al rojo significativos, lo que indica que se están alejando de nosotros debido a la expansión del universo. Este gran desplazamiento al rojo nos permite usar los cuásares como marcadores para estudiar el universo temprano y la evolución de las estructuras a gran escala del cosmos.[6][7]

2.2 Características de las bandas utilizadas

Se ha hablado previamente sobre la razón detrás del uso de cada una de las bandas empleados en este trabajo y su importancia. A continuación, se ofrece una explicación más detallada sobre el uso específico de cada banda en este trabajo y sus características distintivas.

2.2.1 Óptico

Del espectro Óptico se utiliza la detección, es decir, el catálogo de cuásares, y el redshift.

El redshift o el desplazamiento al rojo es una de las características más importantes estudiadas, ya que es una medida de desplazamiento de las líneas espectrales hacia longitudes de onda más largas debido a la expansión del universo. En el caso de los cuásares ayuda a determinar la distancia a la que se encuentra el cuásar y velocidad de recesión, siendo los cuásares con redshift más altos los que están más distantes y permitiéndonos así estudiar el universo primitivo.

2.2.2 Rayos X

Como características espectrales X de los cuásares se tienen el flujo, la luminosidad, la densidad columna (N_H) y el índice fotónico. Todas las características son la clave de este trabajo y son, principalmente, las que se van a estudiar con detalle, por lo que es necesario que se entiendan bien estos conceptos.

Se estudiarán tanto el flujo como la luminosidad, características que no dan la misma información. El flujo se refiere a la cantidad de energía emitida por el cuásar que llega a la tierra por unidad de área y de tiempo, y se mide en erg/s/cm^2 . La luminosidad,

por el contrario, es la cantidad total de energía que emite el cuásar en todas las direcciones por unidad de tiempo, se mide en erg/s. Esta es una propiedad intrínseca del cuásar, es decir, no depende del observador y proporciona una medida directa de su potencia. El flujo observado es crucial para estudiar la luminosidad intrínseca del cuásar para estudiar la variabilidad en la emisión de rayos X.

Hay que tener en cuenta que las dos propiedades mencionadas hasta ahora, el redshift y la luminosidad, están relacionadas. Generalmente, un redshift más alto implica una mayor luminosidad aparente en nuestras observaciones. Esto se debe a que, a distancias mayores, solo se van a poder observar cuásares más luminosos. [3]

Otras características importantes del espectro X que se han estudiado son la densidad columna (N_H) y el índice fotónico.

N_H se refiere a densidad columna y se mide en número de átomos por unidad de superficie. La absorción principal es debida a elementos pesados (O, Fe) que suelen ir acompañando al hidrógeno. El valor de esta columna está relacionado con la emisión FIR, lo cual se debe a que las regiones con mayor densidad columna contienen más gas y polvo. Al tener más material, tiende a tener mayores tasas de formación de estrellas. Estas estrellas calientan el polvo circundante, lo que resulta en una mayor emisión en FIR.

El índice fotónico es la medida de cómo varía el flujo de rayos X con la energía. Un índice fotónico más bajo indica que el flujo de fotones disminuye más lentamente con la energía, mientras que un índice más alto indica que el flujo de fotones disminuye más rápidamente con la energía (el umbral suele ser alrededor de 2). Este parámetro es útil para entender los mecanismos de emisión y la física del material alrededor del agujero negro supermasivo.

La combinación de N_H y el índice fotónico proporciona una visión detallada del entorno y las condiciones físicas de los cuásares.

2.2.3 FIR

El FIR es una banda crucial para el estudio de los cuásares y sus galaxias anfitrionas debido a su sensibilidad al polvo y gas frío, y su capacidad para penetrar a través de nubes densas que bloquean la luz visible. En las regiones de formación estelar, el polvo calentado por las estrellas jóvenes emite fuertemente en el FIR. Este tipo de emisión es esencial para estudiar las zonas donde nacen nuevas estrellas, ya que la

mayor parte de la radiación visible y ultravioleta de las estrellas jóvenes es absorbida y reemitida en el FIR. Así, el FIR permite mapear las estructuras polvorosas y las regiones de intensa formación estelar dentro de las galaxias anfitrionas de los cuásares.

La observación en FIR también es vital para entender la interacción entre el núcleo activo galáctico (AGN) y la galaxia anfitriona. Los cuásares emiten radiación intensa que puede calentar el polvo circundante. Al estudiar la radiación reemitida en el FIR, se puede inferir cómo el AGN influye en su entorno. Por ejemplo, puede revelar si el AGN suprime o fomenta la formación estelar en diferentes partes de la galaxia. Estudios realizados con observatorios en infrarrojo han mostrado que las galaxias con cuásares activos tienen diferentes perfiles de emisión en FIR comparadas con galaxias sin cuásares activos, sugiriendo una interacción significativa entre el AGN y el gas y polvo circundante.

Capítulo 3

Metodología

Para alcanzar el objetivo, se analizan muestras espectroscópicas detalladas de cuásares obtenidas a través de diversas observaciones astronómicas, como las proporcionadas por el Sloan Digital Sky Survey (SDSS). Se estudian específicamente las características de la emisión en rayos X, que indican los procesos de alta energía que ocurren cerca del agujero negro supermasivo. Estas observaciones se comparan con los datos del infrarrojo lejano, que reflejan la presencia y distribución del polvo y gas frío en la galaxia anfitriona, así como las regiones de formación estelar.

Este enfoque permite, no solo identificar posibles correlaciones entre la actividad del núcleo galáctico activo (AGN) y la estructura de la galaxia anfitriona, sino que también permite comprender cómo estos procesos están interrelacionados y afectan mutuamente su evolución. La emisión en rayos X, que es un indicador directo de la acreción de materia por parte del agujero negro, se utiliza para evaluar la actividad del AGN, mientras que la emisión en el infrarrojo lejano proporcionará una visión detallada de la tasa de formación estelar dentro de la galaxia anfitriona.

3.1 Selección y descripción de las fuentes de datos

Las fuentes principales utilizadas para el análisis comparativo de este trabajo son principalmente tres:

Fuentes Ópticas (DR12Q): [8]

El catálogo DR12Q del Sloan Digital Sky Survey (SDSS) es una compilación extensa de cuásares identificados a partir de observaciones espectroscópicas. El sistema fotométrico conocido como SDSS utiliza un conjunto de cinco filtros para cubrir un rango amplio del espectro de luz visible e infrarroja cercana. Estos filtros son u (ultravioleta), g (verde), r (rojo), i (infrarrojo cercano) y z (infrarrojo más lejano). Este catálogo incluye cuásares cuya magnitud aparente en el filtro i es menor que 22 y que presentan líneas de emisión amplias, típicas de estos objetos, confirmados mediante espectroscopía.

Para la identificación de los cuásares, el SDSS tiene un identificador único de objetos celestes, “ObjID”, el cual es el responsable de etiquetar de manera única cada fuente astronómica específica.

Las zonas del cielo observadas por el SDSS se clasifican en dos tipos: s (shallow) y d (deep). Las fuentes tipo s corresponden a vastas extensiones celestes estudiadas con menor profundidad, mientras que las fuentes tipo d representan reducidas secciones celestes analizadas con mayor detalle.

Además, DR12Q proporciona datos detallados sobre las propiedades espectrales de los cuásares, incluyendo propiedades fotométricas en múltiples bandas, y parámetros de las líneas de emisión y absorción, aunque la propiedad más importante del catálogo para este trabajo es el redshift.

Fuentes X (D6.4):[9][14][15][16]

El informe D6.4 del proyecto XMM2ATHENA proporciona un catálogo de fuentes en rayos X basado en datos del observatorio XMM-Newton. Este catálogo combina fuentes de los catálogos D6.1 y D6.2, los cuales han ajustado las fuentes del conjunto de datos proveniente del observatorio XMM-Newton: 4XMM-DR11 y 4XMM-DR11s. [10][11]

Las fuentes clasificadas como AGN en D6.4 incluyen parámetros espectrales como el redshift, la luminosidad en X, el flujo, la absorción de gas (NH) y el índice fotónico.

Hay que considerar que los z del catálogo D6.4 vienen de identificaciones de las fuentes X y de técnicas de redshift fotométrico aplicadas a los AGN identificados.

De la misma manera que “ObjID”, las fuentes X también cuentan con un identificador único “SRCID”, el cual tiene la misma función que “ObjID” pero para las fuentes X.

Fuentes IR lejano (Spire250): [17][19]

El instrumento SPIRE (Spectral and Photometric Imaging Receiver) a bordo del Observatorio Espacial Herschel, es una cámara y espectrómetro submilimétrico diseñado para explorar el universo en longitudes de onda del infrarrojo lejano. SPIRE contiene dos sub-instrumentos: un fotómetro de imagen y un espectrómetro de transformada de Fourier (FTS).

El fotómetro de imagen opera en tres bandas anchas centradas en 250,350 y 500 micrómetros, lo que permite realizar imágenes del cielo en tres “colores” submilimétricos simultáneamente. El FTS cubre un rango de 194-571 micrómetros y permite medir las características espectrales de átomos y moléculas en esta banda.

La banda de 250 micrómetros es especialmente importante para la astronomía infrarroja porque permite detectar la emisión térmica de granos de polvo calentados por la formación estelar. Al observar en esta longitud de onda, SPIRE puede identificar regiones de intensa actividad estelar y mapear la distribución de polvo y gas en las galaxias. [18]

Además de para las correlaciones, el archivo Spire250 se utiliza principalmente para el estudio del Flujo a 250 micras, el cual se mide en mJy (miliJansky), una unidad de medida para la intensidad de flujo que es proporcional a $10^{-26} W/m^2/Hz$.

Archivos XO: [20]

Además de las fuentes principales, se cuenta con una serie de archivos resultantes de una correlación previamente establecida entre las posiciones de las fuentes X y las de las fuentes del SDSS teniendo en cuenta sus posiciones, los cuales se denominarán “Archivos XO”. Los Archivos XO se clasifican en dos tipos según se han observado en el SDSS: s y d.

Archivos XO	Catálogo X
xmatch_xgoruc_4dr13_v1.0_galexsdss12d.fits	DR13
xmatch_xgoruc_4dr13_v1.0_galexsdss12s.fits	DR13
xmatch_xoruc_4dr13_v1.0_nogalexsdss12d.fits	DR13
xmatch_xoruc_4dr13_v1.0_nogalexsdss12s.fits	DR13
xmatch_xgoruc_4dr11s_v1.0_galexsdss12d.fits	DR11s
xmatch_xgoruc_4dr11s_v1.0_galexsdss12s.fits	DR11s
xmatch_xoruc_4dr11s_v1.0_nogalexsdss12d.fits	DR11s
xmatch_xoruc_4dr11s_v1.0_nogalexsdss12s.fits	DR11s

Tabla 1: Archivos XO utilizados y su catálogo X de proveniencia. [12][13]

Los archivos que acaban con “d” se refieren a “deep” como se ha explicado previamente, y los archivos que acaban con una “s” se refiere a “shallow”. Los que llevan “galex” en el nombre se refieren a que los datos del satélite GALEX están incluidos, y los que llevan “nogalex” a que no lo están.

Mediante los identificadores “ObjID” y “SRCID” y las columnas que describen la posición en el cielo de cada fuente como la ascensión recta y declinación, es posible realizar correlaciones precisas entre diferentes catálogos para posteriormente estudiar sus diferencias.

3.2 Técnicas de análisis de datos

El software utilizado para el análisis de datos es TopCat. TopCat es un software interactivo que está diseñado para la manipulación, análisis y comparación de grandes conjuntos de datos tabulares. Permite cargar, visualizar y manipular datos provenientes de diversas fuentes, incluyendo archivos locales o base de datos en línea. El software es compatible con varios tipos de archivo como FITS, ASCII o CSV, lo que lo convierte en un software extremadamente sensible y útil. Además, ofrece herramientas avanzadas de visualización como gráficos de dispersión, histogramas... facilitando la identificación de patrones y tendencias de los datos.

TopCat permite trabajar con grandes volúmenes de datos y por esa razón es utilizada repetidamente en astronomía. Sus capacidades de coincidencia cruzada y consultas espaciales permiten la integración y comparación de datos entre diferentes catálogos, mejorando la precisión del análisis. [23]

A partir de este software, para ayudar a determinar patrones y tendencias de las muestras, se ha hecho uso de medidas estadísticas como la media, la varianza y los cuartiles, que son fundamentales para resumir y describir las características de un conjunto de datos.

La media se define como la suma de todos los valores dividida por el número de valores. Es una tendencia central que proporciona una idea del valor típico del conjunto de datos.

La varianza mide la dispersión de los datos respecto a la media. Indica qué tan dispersos están los datos en torno a la media, es decir, una varianza baja indica que los datos están más concentrados cerca de la media y viceversa. [23]

Los cuartiles son valores que dividen un conjunto de datos en cuatro partes iguales. Son los más importantes para este estudio debido a que son esenciales para entender la distribución de los datos y para identificar la presencia de valores atípicos. El cuartil 1 es el que divide el primer 25% de los datos, es decir, el 25% de los datos es menor a ese valor. El cuartil 2, o la mediana, es el que divide el 50%, es decir, el 50% de los datos es menor que ese valor. El cuartil 3 es el que divide el 75%, es decir, el 75% de los datos es menor a ese valor. Por esta razón los cuartiles son especialmente útiles para distribuciones no simétricas. [23]

El método estadístico que se ha utilizado que proporciona una mayor fiabilidad es la prueba Kolmogorov-Smirnov (K-S). Es una prueba no paramétrica utilizada para comparar dos distribuciones de probabilidad entre una muestra y una distribución teórica o entre dos muestras entre sí. Esta prueba mide la distancia máxima entre las

funciones de distribución acumulada (FDA), que representa la proporción acumulada de observaciones, de las dos muestras. Es muy útil para determinar si las muestras provienen de la misma distribución o de distribuciones diferentes.[23]

La prueba K-S proporciona dos valores clave, D y P:

El valor D es la distancia máxima entre las dos funciones de distribución acumulativa y el valor P indica la probabilidad de observar un valor tan extremo como D bajo la hipótesis nula de que las distribuciones son iguales.

Típicamente un valor P utilizado para evaluar la similitud de dos muestras se encuentra alrededor de 0.0027 [23]. Si el valor p es mayor que 0.0027, no hay suficiente evidencia para rechazar la hipótesis nula de que las muestras provienen de la misma distribución, sugiriendo que son similares. Si el valor p es menor que 0.0027, se rechaza la hipótesis nula, indicando que las distribuciones son significativamente diferentes.

La prueba estadística K-S se ha podido realizar mediante un código de Python. Python es un lenguaje de programación utilizado en diversas áreas debido a su facilidad de uso y extensa biblioteca. [23][24][25][26]

3.3 Construcción del diagrama de Venn

Se ha de construir un diagrama como el de la Figura 1. Para ello, a partir de las fuentes ópticas se realiza una correlación con las fuentes X, tras la cual se hace una tercera correlación entre esas fuentes con la banda FIR, para obtener el número de fuentes de cada una de las subzonas a estudiar.

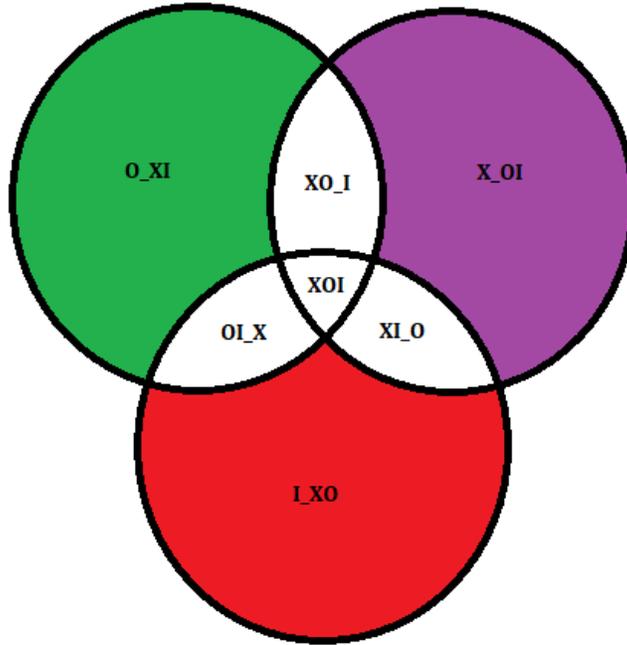


Figura 1: Diagrama de Venn a estudiar.

En la Figura 1 se observan las zonas a estudiar, que son subzonas de las ya mencionadas fuentes del óptico, fuentes X y fuentes FIR. Se define “_” como “no”, es decir, la zona verde es la zona donde se encuentran fuentes del óptico que no se hayan detectado ni en FIR ni en rayos X. De la misma manera, la zona morada son fuentes detectadas en rayos X que no se hayan detectado en ninguna otra banda a estudiar. El mismo proceso sigue a la zona roja.

Las zonas sin color se componen de tres zonas diferentes de las que forman parte fuentes que se han detectado en dos bandas y en una no, como por ejemplo la zona XO_I que son fuentes detectadas en el óptico y en rayos X, y no detectadas en FIR y la zona OI_X son fuentes detectadas en el óptico y en FIR y no en X.

Finalmente, la zona central XOI es formada por fuentes que son detectadas en las tres bandas del espectro a estudiar.

Además de esto, posteriormente se hace una comparación entre la zona X_I ($XO_I + X_{OI}$) y la zona XI ($XI_O + XOI$). Siguiendo la explicación anterior, estas zonas se refieren a fuentes detectadas en X y no en FIR (X_I) y fuentes detectadas en X y en FIR (XI).

Capítulo 4

Resultados

4.1 Obtención de las zonas del Diagrama de Venn

Lo primero que se hizo fue verificar si las fuentes de los Archivos XO presentaban solapamientos. Como se ha mencionado previamente, las fuentes tipo “s” corresponden a vastas extensiones del cielo estudiadas con menor profundidad y las de tipo “d” corresponden a reducidas secciones analizadas con mayor detalle, pero no se conoce si las regiones pudieran coincidir, produciendo así que las fuentes se solapen entre ambas.

La verificación se realizó utilizando el observatorio 3D de TopCat observando las fuentes y viendo si se solapan entre sí, y, para corroborarlo con una mayor certeza, se hizo una correlación utilizando el identificador ObjID del DR12Q que tienen incorporado estos archivos. Se concluye que no se solapan y que, efectivamente, estudian distintas secciones celestes.

A partir de las Fuentes X, se tiene que en la columna D6 de los archivos, que proviene del estudio D6.4 previamente explicado, puede obtener dos valores según su origen: $D6 = 2$ y $D6 = 1$.

Si $D6 = 2$, entonces significa que las fuentes vienen directamente del catálogo DR11s, por lo que simplemente utilizando el identificador X (SRCID) y correlacionándolo con los Archivos XO, se han obtenido las fuentes X que están asociadas a fuentes del óptico (SDSS). Luego, entre esas fuentes asociadas, mediante el identificador ObjID, se escogen las fuentes que están asociadas a cuásares del SDSS correlacionándolo con DR12Q.

Si $D6 = 1$, se debe tener en cuenta que el identificador X (SRCID) corresponde a la versión DR11 del catálogo XMM, mientras que el identificador ObjID viene del DR13. Para solucionar esto, primero se correlacionan los catálogos DR11 y DR13 mediante el identificador SRCID. Tras interrelacionarlos, se pueden asociar fuentes X con un identificador proveniente del DR11 a fuentes X con identificador proveniente del

DR13. De esta manera, se repite el procedimiento anterior utilizando los Archivos XO correspondientes.

A raíz de estas correlaciones, se tiene que la zona XO consta de 4.648 fuentes asociadas a cuásares del SDSS.

Para verificar si se iba por buen camino, se ha realizado una comprobación. Observando la separación máxima entre cuásares con un mismo SRCID, se ha realizado una correlación directa entre D6.4 y DR12Q utilizando esa distancia máxima para comprobar que iba en la misma línea que los resultados obtenidos, es decir, se han obtenido las fuentes X más cercanas a cada cuásar del DR12Q. Efectivamente, se ha verificado obteniendo un valor similar a los 4.648 previamente determinados.

La zona XO tendrá dos columnas para el redshift, una proveniente del archivo D6.4 (zX) y otra proveniente del archivo DR12Q (zO). Se ha realizado por tanto la comprobación de que, aunque su proveniencia sea distinta, tenga los mismos valores para el redshift ya que las fuentes son las mismas.

	zX	zO
Media	1.551	1.536
Varianza	0.803	0.751
Cuartil 1	0.786	0.786
Cuartil 3	2.281	2.270
Mediana	1.402	1.397

Tabla 2: Medidas estadísticas para los valores obtenidos para el redshift proveniente del archivo D6.4 y del archivo DR12Q.

Los valores, como cabía esperar, son muy similares. Esto concuerda ya que los cuásares que se han obtenido deberían ser los mismos, y entra dentro del rango de error de cada técnica utilizada para obtener estos valores del redshift.

Tras realizar las comprobaciones necesarias, para correlacionar las zonas del óptico y de rayos X con la zona del infrarrojo lejano, se tuvo en cuenta la falta de un identificador en las fuentes FIR. Por ello, se decidió hacer una correlación con las

fuentes de rayos X (X) y del óptico (O) usando una distancia máxima de 6 arcosegundos.

A partir de esta correlación, se obtuvieron 729 fuentes en la zona XI y 1,045 fuentes en la zona OI, y la zona XOI se determinó utilizando nuevamente los identificadores de las fuentes de rayos X y del óptico.

Una vez completadas las correlaciones con ambas fuentes, se considera completado el diagrama de Venn.

Venn	Nº Fuentes
X_OI	31.003
O_XI	292.393
I_XO	418.132
XO_I	4.388
XI_O	469
OI_X	785
XIO	260

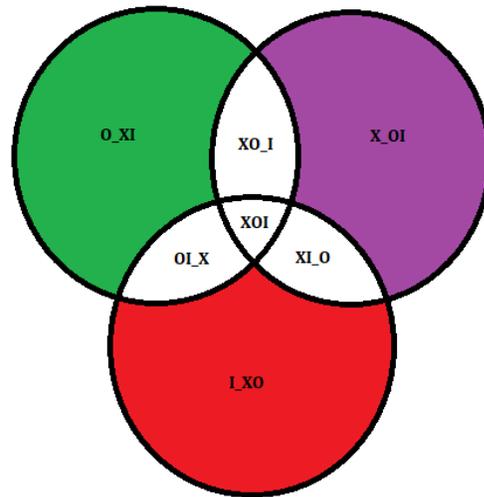


Tabla 3: Número de fuentes para cada subzona del Diagrama de Venn.

Figura 1: Diagrama de Venn a estudiar.

4.2 Redshift y luminosidad

El redshift y la luminosidad están interrelacionados, a mayor redshift, corresponde una mayor luminosidad. Se ha estudiado esta correspondencia con relación a la luminosidad en rayos X (detectada en el XMM, L_x).

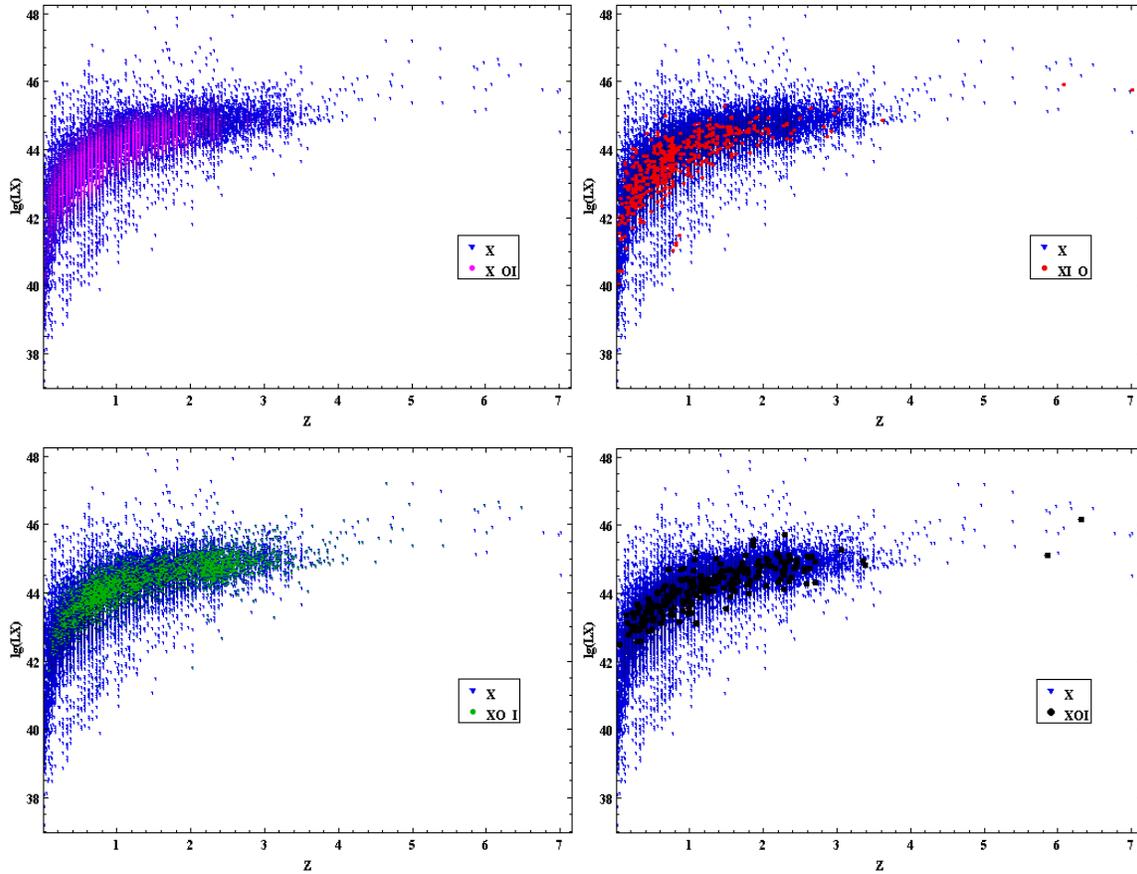


Figura 2: Representación de la dependencia de la luminosidad L_x entre 2-12KeV en $[\text{erg} \cdot \text{s}^{-1}]$ en función del redshift.

La diferencia del número de fuentes entre las zonas a estudiar en la Figura 2 es significativa, pero se pueden apreciar ciertos patrones de las distribuciones que sigue cada una de las subzonas.

Se aprecia que a menores redshifts, y en consecuencia a menores luminosidades, las fuentes X y FIR son las que tienen una mayor densidad de fuentes en esas zonas, llegando a los mínimos redshifts detectados en X. Mientras que en las zonas que se detectan en el óptico (XO_I y XOI) tienen la mayor densidad de fuentes en redshifts más altos, llegando a los máximos de la zona X.

Esto puede deberse a varios factores.

Para comprender el origen y la trascendencia de estos resultados, se han estudiado estas dos propiedades de manera individual. De esta manera se pueden obtener resultados más convincentes y realizar un análisis más preciso.

4.21 Redshift

Se han comparado los valores del redshift de las diferentes subzonas respecto a las zonas más grandes para intentar entender la relación entre la detección de estas fuentes en ciertos espectros con los valores de estos del redshift.

Primero se han realizado comparaciones del redshift de los AGN de la zona del óptico respecto a todas sus subzonas estudiadas en el diagrama de Venn.

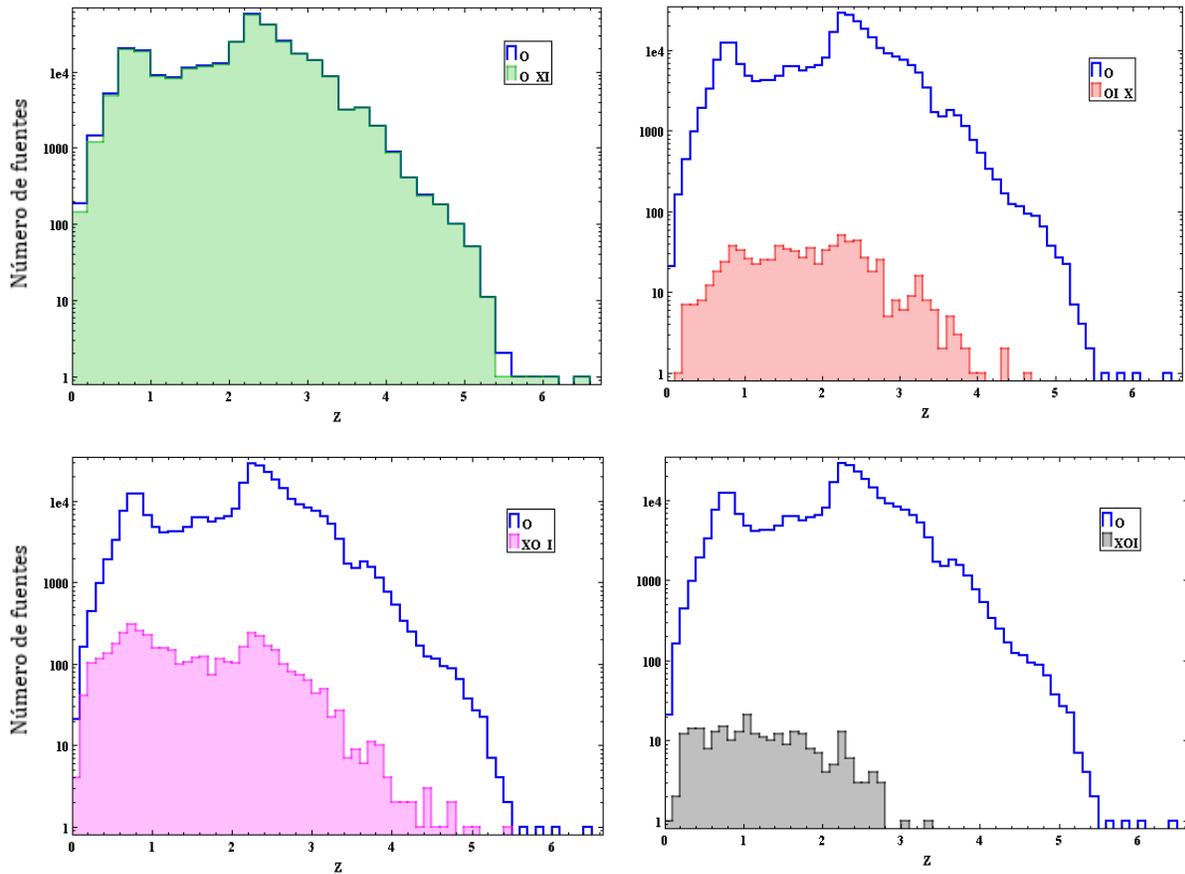


Figura 3: Representación de los resultados de la comparación del redshift de las fuentes del óptico respecto a las subzonas del diagrama de Venn en escala logarítmica respecto al número de fuentes.

Lo primero a observar es que el número de fuentes de O y de O_XI y que la zona O_XI tiene una mayor densidad de fuentes en redshifts más altos que cuando se tienen en cuenta las observaciones en rayos X y FIR. A medida que se estudian las fuentes detectadas en rayos X y FIR se detecta que los redshifts bajan de manera significativa. Esto concuerda con los resultados obtenidos en el apartado anterior (Figura 2).

No hay muchas diferencias visuales entre las distribuciones de las fuentes del óptico detectadas también en FIR o en rayos X, ni en las detectadas en ambos además del óptico.

Los valores numéricos de estos histogramas se pueden obtener de la Tabla 4.

Para entender si esto ocurre solo con las fuentes del óptico o es un patrón, se estudia el redshift detectado proveniente del D6.4 comparando las fuentes X con todas sus subzonas del diagrama de Venn.

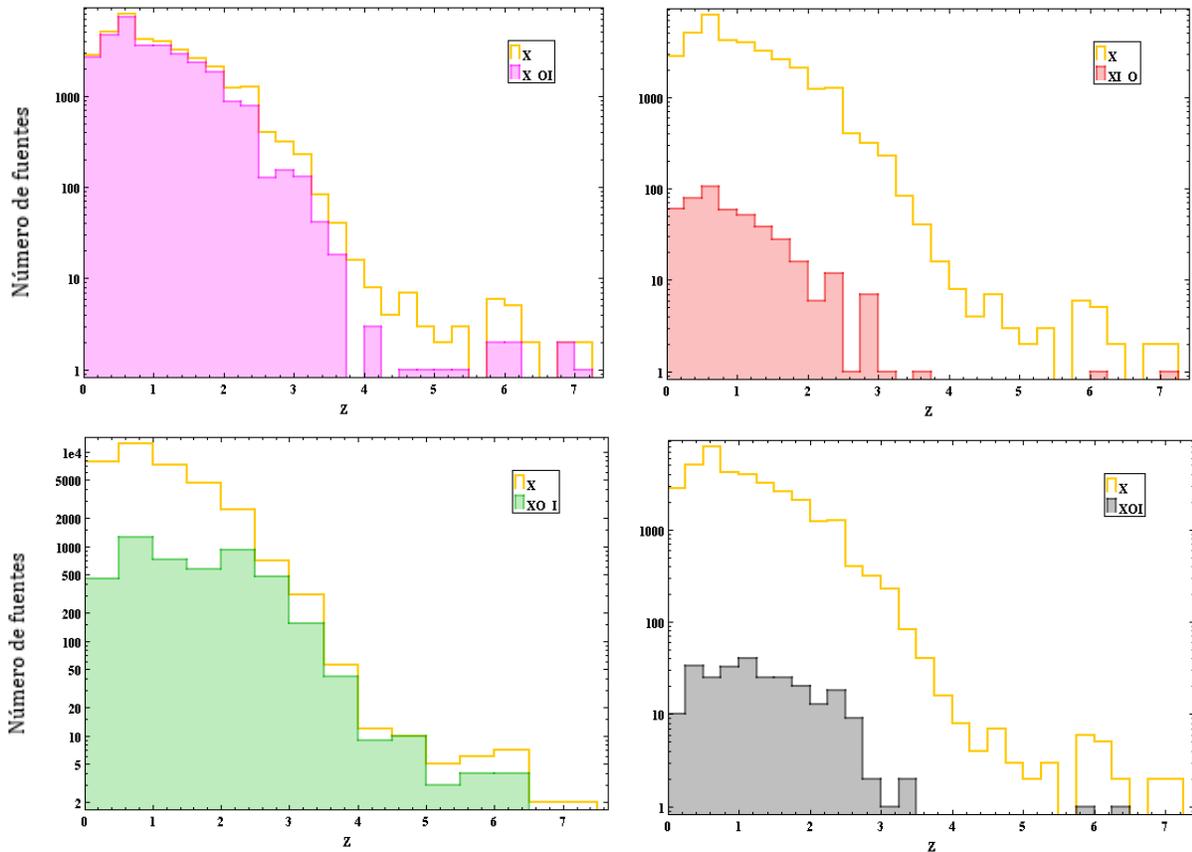


Figura 4: Representación de los resultados de la comparación del redshift proveniente del D6.4 respecto a las subzonas del diagrama de Venn en escala logarítmica respecto al número de fuentes.

En la figura 4 se corrobora lo propuesto anteriormente. Se observa que de todas las fuentes X, las que tienen el redshift más alto forman parte de las detectadas en el óptico como se ha visto en la Figura 3. Y de la misma manera, las fuentes detectadas con valores del redshift más bajo son las detectadas solamente en rayos X o tanto en rayos X como en FIR.

Se hace un estudio más exhaustivo mediante la estadística en la Tabla 4.

	O_XI	X_OI	I_XO	OI_X	XO_I	XI_O	XOI
Media	2,12	1,03	-	1,84	1,53	0,90	1,30
Varianza	0,64	0,48	-	0,64	0,75	0,52	0,68
Cuartil 1	1,60	0,53	-	1,21	0,79	0,38	0,71
Mediana	2,29	0,83	-	1,86	1,40	0,73	1,15
Cuartil 3	2,61	1,43	-	2,38	2,27	1,23	1,76

Tabla 4: Valores del redshift para todas las zonas del diagrama de Venn.

Una inspección numérica revela claramente que los resultados exhiben características de distribuciones disímiles y no muestran una uniformidad en su comportamiento.

Se observa que en cuanto se detectan en rayos X, las fuentes tienen un valor de redshift menor. La zona en la cual el valor medio del redshift es menor es precisamente en la que no se detecta en el óptico.

Los valores de los redshifts de las fuentes que engloba la zona O han sido calculadas mediante el Sloan SSDS, mientras que la zona X se han obtenido con técnicas de redshift fotométrico. Al tener dos proveniencias diferentes, puede haber un sesgo que evite que se puedan estudiar las zonas desde una misma base ya que unas tendrán de base un redshift más alto.

Otra razón de que esto suceda es que el óptico “ve” más profundo que las otras bandas estudiadas, de manera que podría estar observando redshifts de mayor valor a los que los espectros de rayos X y FIR no serían capaces de llegar.

Para determinar cuál de las dos teorías es más probable que sea correcta, se han comparado las fuentes que tienen como proveniencia el observatorio SDSS y las fuentes calculadas en D6.4 mediante técnicas de redshift fotométrico.

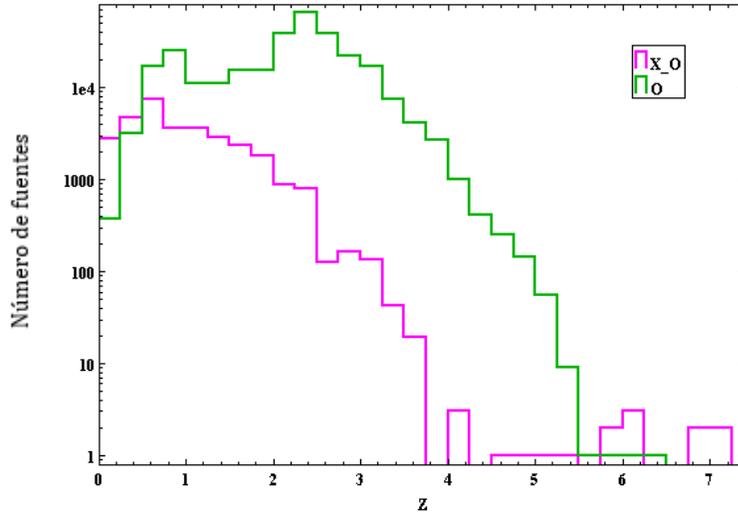


Figura 5: Histograma del redshift de las zonas O y X_O en escala logarítmica respecto al número de fuentes.

La Figura 5 ilustra de manera clara y precisa la tendencia mencionada previamente, proporcionando un resultado claro de que la técnica utilizada para obtener el redshift para D6.4 tiene una tendencia a estudiar redshifts más bajos. Este sesgo dificulta el análisis del estudio, ya que se parte de esa desviación.

4.22 Luminosidad

Se estudia la luminosidad en X de todas las subzonas de la zona X para analizar si las conclusiones obtenidas previamente se repiten también en el estudio de la luminosidad.

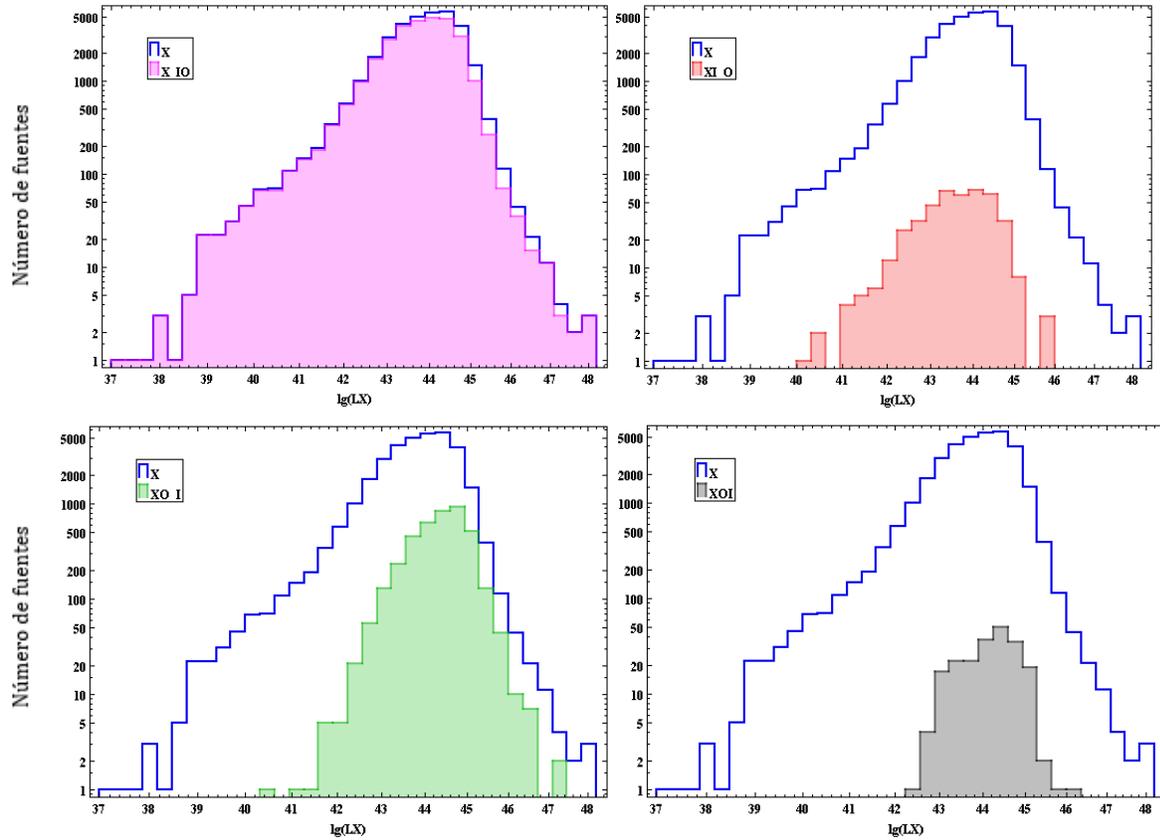


Figura 6: Histograma en escala logarítmica del logaritmo de la luminosidad en X entre 0.2-12keV en $[erg.s^{-1}]$ de AGN detectados en rayos X de la zona X y de sus respectivas subzonas respecto al número de fuentes.

Se corrobora el análisis realizado para los redshift de las mismas zonas, lo cual concuerda con que los redshifts y las luminosidades están relacionadas. Se observa que se repite el patrón, las zonas en las que no se detecta en el óptico siguen una distribución parecida a la original de rayos X. En cambio, cuando se detecta en el óptico la distribución cambia y la luminosidad detectada en estas fuentes aumenta significativamente.

Estas diferencias se analizan de una manera más precisa haciendo uso de los valores estadísticos.

	X_OI	XO_I	XI_O	XOI
Media	43,81	44,38	43,59	44,16
Varianza	0,83	0,43	0,78	0,43
Cuartil 1	43,32	43,97	43,06	43,67
Mediana	43,93	44,46	43,67	44,28
Cuartil 3	44,44	44,82	44,24	44,66

Tabla 5: Valores numéricos del logaritmo de la luminosidad entre 2-12KeV [erg.s^{-1}] de las subzonas de X.

A simple vista, no se aprecia una gran diferencia. Sí se aprecia lo mencionado anteriormente pero no de manera clara, por lo que se somete a la prueba K-S estadística para analizar si, excluyendo las zonas detectadas en el óptico, hay alguna zona en la que se puede considerar que sigue una misma distribución.

Para todas las correlaciones se obtienen valores mucho menores que 0.0027. Se observa por tanto que, aunque pueda parecer que siguen la misma distribución en la Figura 6, la prueba estadística K-S dice que se puede rechazar la hipótesis nula, es decir, que no hay similitudes de ninguna muestra estudiada y que por tanto no siguen una misma distribución. El hecho de que las fuentes sean detectadas en uno o más bandas espectrales cambia la distribución del conjunto.

Se concluye lo que se había propuesto, que lo más probable es que se tenga un sesgo en las detecciones, lo cual quizás dificulte el análisis preciso y el estudio del trabajo realizado. Se van a estudiar más propiedades espectrales, pero se ha de tener en cuenta estos sesgos comprobados antes de realizar ninguna conclusión precipitada.

4.3 Flujo en FIR

El flujo en FIR es crucial para estudiar los objetos fríos como nubes de gas y polvo que no emiten luz visible o estudiar regiones de formación estelar ya que el FIR está asociado al crecimiento de la galaxia.

Se estudia el flujo a 250 micras detectado en FIR para las fuentes FIR y las subzonas de la zona del infrarrojo lejano.

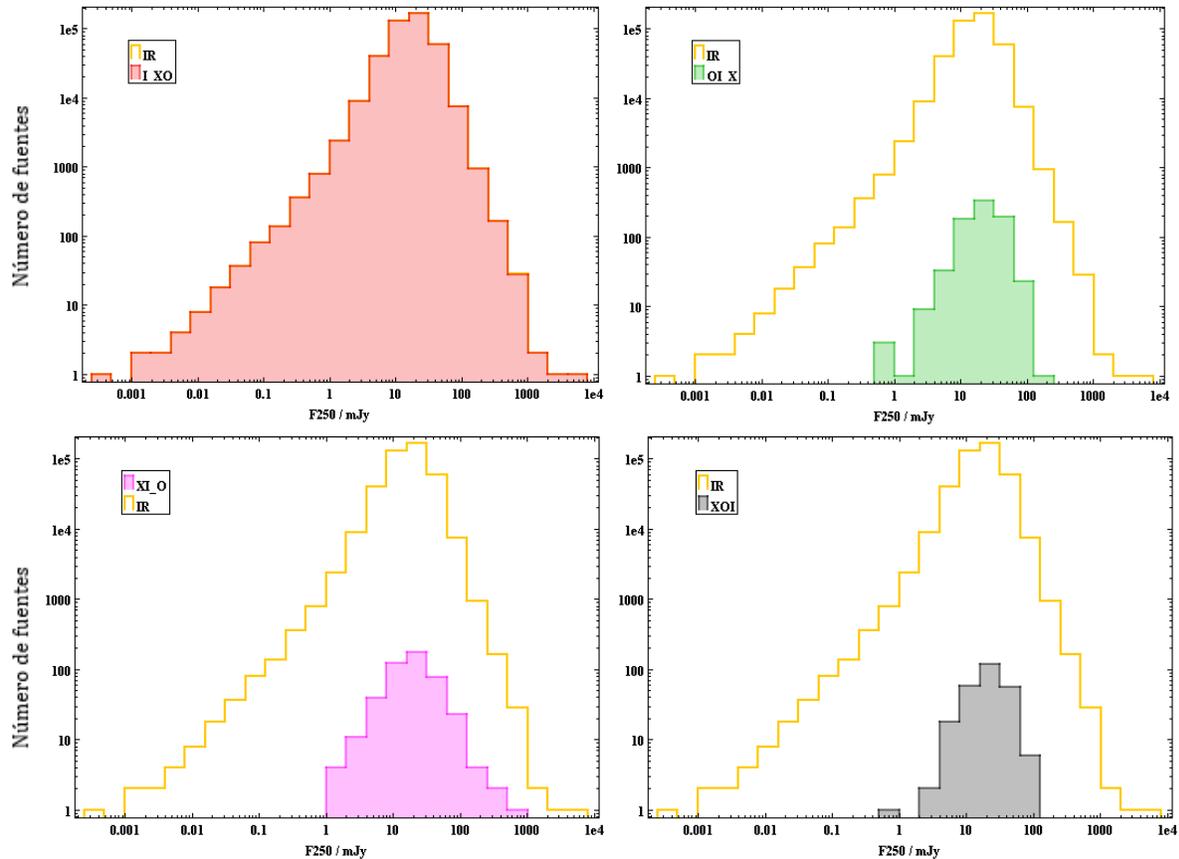


Figura 7: Flujo detectado a 250 micras en el infrarrojo de las zonas I_XO (rojo), XI_O (amarillo), OI_X (gris) y XIO (negro) respecto al número de fuentes.

Aparentemente no se observan grandes diferencias entre las subzonas de la zona I. Las distribuciones podrían seguir un mismo patrón. Lo que sí se aprecia es una gran variedad de flujos, ya que la densidad de flujos de las fuentes IR varía desde 0 mJy hasta 10.000 mJy. La de los cuásares prácticamente no baja de 1 mJy.

Para analizar de manera más precisa las distribuciones, se realiza un análisis numérico con propiedades estadísticas.

	I_XO	OI_X	XI_O	XOI
Media	21,37	26,23	24,98	24,61
Varianza	1501	283	542	245
Cuartil 1	11,23	14,75	11,55	13,87
Mediana	17,61	22,10	18,97	21,07
Cuartil 3	56,41	34,13	30,56	31,49

Tabla 6: Valores numéricos de la densidad del flujo en 250 micrómetros [mJy] para las subzonas de las fuentes Infrarrojas.

Respecto a los valores numéricos del flujo a 250 micrómetros se puede apreciar de manera más clara que la visual la concordancia con el análisis realizado anteriormente, observando que cuando la muestra tiene detecciones en el óptico, el flujo aumenta de manera significativa, y cuando están detectadas en FIR o en X, especialmente en FIR, disminuye. Esto sucede porque el flujo está relacionado con la luminosidad y anteriormente se ha estudiado que la luminosidad en muestras con detecciones en FIR disminuye significativamente (Figura 6).

Para estudiar si alguna de las muestras de fuentes sigue una misma distribución o no se parecen, se somete a la prueba estadística K-S.

Muestra 1	Muestra 2	Valor P
I_XO	OI_X	0
I_XO	XI_O	0.005
I_XO	XOI	0
OI_X	XI_O	0
OI_X	XOI	0.68
XI_O	XOI	0.19

Tabla 7: Representación del valor P de la prueba K-S para cada comparación del flujo a 250 micras.

Observación: Para $valores P < 0,0001$ se ha tomado 0.

Los resultados indican que las distribuciones siguen un patrón. Las muestras comparadas que coinciden en que ninguna de las dos se detecta en el óptico, o, por el contrario, que ambas son detectadas en el óptico, siguen una misma distribución. Esto visibiliza la importancia en la detección de cada una de las bandas. De hecho, la comparación que es más probable que siga una misma distribución es cuando se compara la subzona OI_X con la XOI (ambas detectadas en el óptico).

Como excepción, la comparación de las muestras XI_O y XOI es probable que sigan una misma distribución.

La razón por la cual OI_X y XI_O sigan la misma distribución que XOI y que OI_X y XI_O no sigan la misma distribución puede deberse a que ambas tienen una cola, pero OI_X tiene la cola hacia valores menores y XI_O hacia valores mayores.

4.4 Propiedades espectrales X

Tres propiedades X imprescindibles para el estudio de los AGN y del crecimiento de sus galaxias anfitrionas son el flujo, la densidad columna y el índice fotónico.

Esta comparación se realiza para los AGN detectados y no detectados en el infrarrojo, XI y X_I respectivamente, que es el principal objetivo de este estudio.

A raíz de la conclusión obtenida previamente en el estudio de la luminosidad y el redshift, se ha decidido estudiar también la comparación entre las propiedades X de las zonas detectadas y no detectadas en IR sin incluir las detectadas en el óptico, es decir, X_IO y XI_O.

Hay que destacar que se han eliminado las fuentes que no han tenido un error en la detección de las propiedades X, que eran aproximadamente un 75% de las originales en cada muestra.

4.41 Flujo

Se compara el flujo X detectado comparando las zonas X_I/XI y X_OI/XI_O.

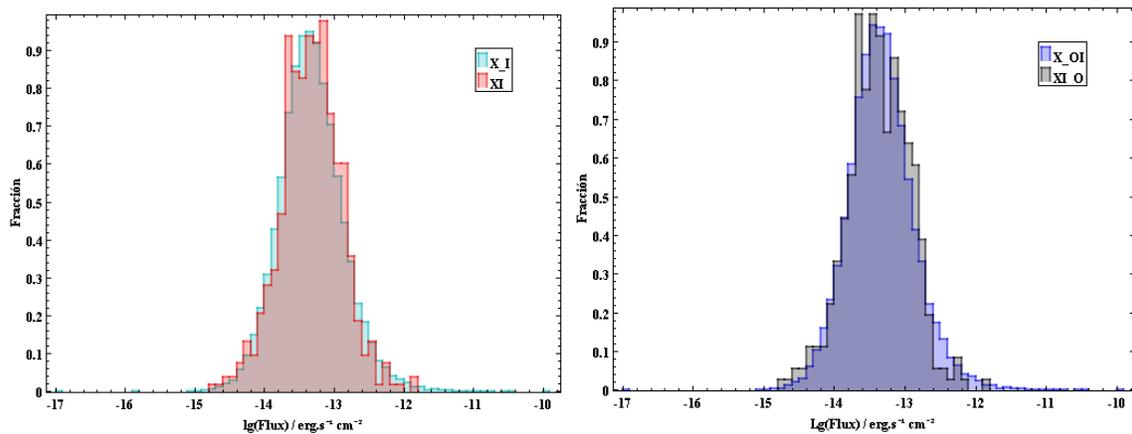


Figura 8: Histogramas del logaritmo del flujo en 0.2-12keV comparando las fuentes detectadas y no detectadas en FIR para cuando el óptico se tiene en cuenta y para cuando no, normalizado por área.

En la Figura 8 se puede apreciar que a simple vista parece que podrían seguir una misma distribución. Se ha realizado las dos comparaciones, una para cuando se tiene en cuenta el óptico y otra para cuando no se ha tenido en cuenta, como variación por los resultados obtenidos en los apartados anteriores.

Parece que las distribuciones pertenecen al mismo patrón, para corroborar esto, se estudian los valores estadísticos del flujo numéricamente.

	<i>XI</i>	<i>X_I</i>	<i>X_OI</i>	<i>XI_O</i>
Media	-13,33	-13,31	-13,32	-13,36
Varianza	0,19	0,21	0,21	0,19
Cuartil 1	-13,61	-13,61	-13,62	-13,64
Mediana	-13,33	-13,34	-13,35	-13,37
Cuartil 3	-13,06	-13,04	-13,06	-13,07

Tabla 8: Valores estadísticos del logaritmo del flujo [$\text{erg} \cdot \text{s}^{-1} \text{cm}^{-2}$] detectado en X para las zonas XI, X_I, X_OI y XI_O.

En principio se corrobora la sensación al ver la Figura 8, es decir, parece que pertenecen a la misma distribución, incluso si no se tiene en cuenta el óptico, lo cual indicaría que el sesgo que se tenía previamente no afecta en las propiedades espectrales X.

Para tener una certeza estadística de este hecho, se somete a la prueba estadística K-S. Los valores de los resultados de la prueba K-S de las propiedades espectrales se encuentran en la Tabla 11.

4.42 Densidad columna (N_H)

Se estudia esta propiedad X para cúasares detectados y no detectados en IR, y repitiendo el procedimiento realizado para el flujo, debido al sesgo analizado entre el óptico y los rayos X, también se estudia la densidad para zonas detectadas y no detectadas en IR sin detecciones en el óptico.

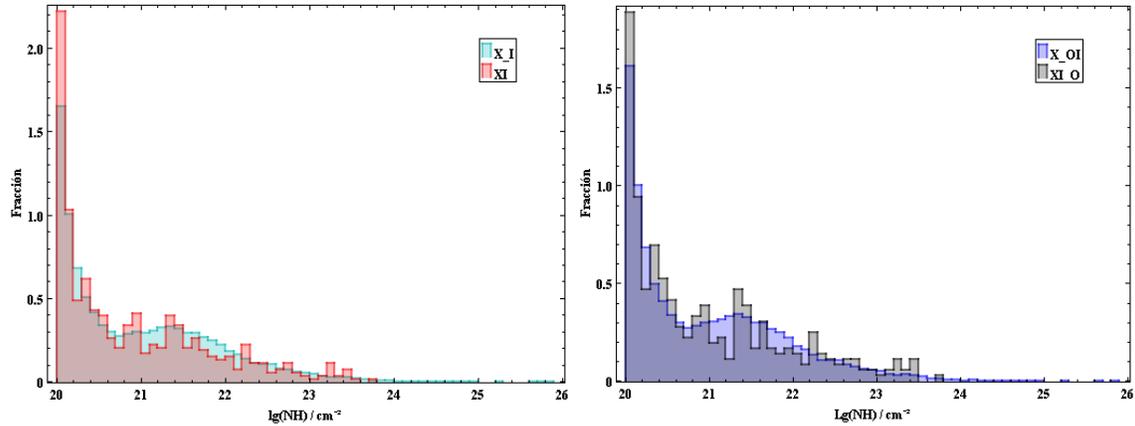


Figura 9: Histogramas del logaritmo de la densidad columna N_H [cm^{-2}] detectado de la comparación de las fuentes detectadas y no detectadas en FIR para cuando la banda del óptico se tiene en cuenta y para cuando no, normalizado por área.

De forma similar al anterior, la Figura 9 muestra que las fuentes podrían seguir una misma distribución. También se aprecia que la detección del óptico no afecta significativamente en la columna densidad.

Se procede de igual forma estudiando los valores numéricos para ver si respaldan la sensación que se tiene al ver la gráfica mediante un análisis estadístico.

	<i>XI</i>	<i>X_I</i>	<i>X_OI</i>	<i>XI_O</i>
Media	20,88	20,96	20,97	20,95
Varianza	0,79	0,77	0,77	0,84
Cuartil 1	20,18	20,12	20,19	20,16
Mediana	20,56	20,73	20,75	20,62
Cuartil 3	21,56	21,42	21,55	21,49

Tabla 9: Valores del logaritmo de la densidad columna [cm^{-2}] detectado para las zonas *XI*, *X_I*, *X_OI*, *XI_O*.

De la misma manera no se aprecia las diferencias numéricamente, por lo que se somete a la prueba K-S con el objetivo de tener una certeza basada en una prueba estadística de que siguen la misma distribución. Los valores de los resultados de la prueba K-S de las propiedades espectrales se encuentran en la Tabla 11.

No parece haber diferencias significativas en las distribuciones en las que se tiene en cuenta el óptico y en las que no. Se realiza una hipótesis de que el sesgo del óptico no afecta al estudio de las propiedades X entre zonas detectadas y no detectadas en IR.

4.43 Índice fotónico

Para comprender la variación del flujo X con la energía, se analiza los valores del índice fotónico (Photon Index) entre las fuentes detectadas y no detectadas en IR. Y para un análisis más exhaustivo, teniendo en cuenta el sesgo entre las detecciones en el óptico y en rayos X, se estudia también las zonas detectadas y no detectadas en IR sin detecciones en el óptico.

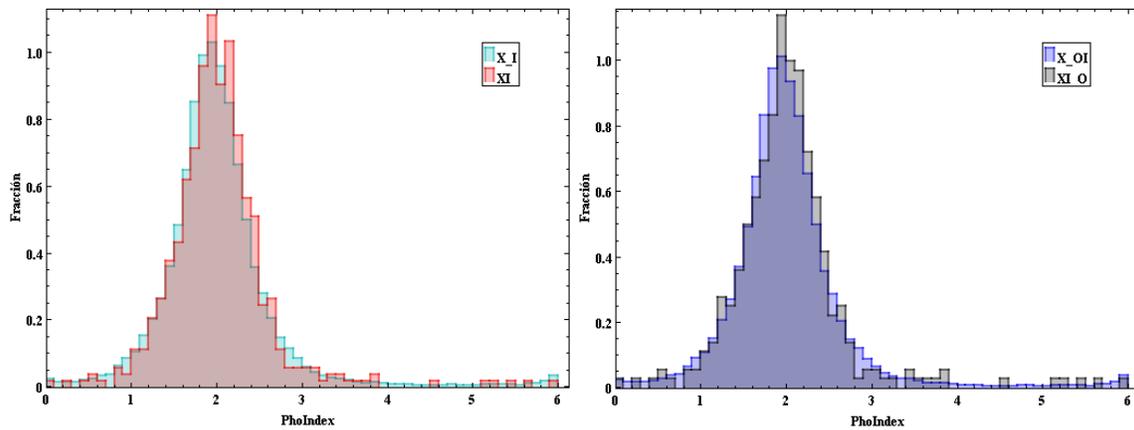


Figura 10: Histogramas del índice fotónico de la comparación de las zonas detectadas y no detectadas en FIR teniendo y no teniendo en cuenta la banda del óptico, normalizado por área.

La sensación que se ha tenido con las otras propiedades espectrales sigue vigente, parece que las muestras siguen una misma distribución.

Al no determinar visualmente una diferencia significativa respecto al índice fotónico, se procede a hacer el estudio numérico de este.

	<i>XI</i>	<i>X_I</i>	<i>X_OI</i>	<i>XI_O</i>
Media	2,03	2,01	2,01	2,02
Varianza	0,35	0,44	0,46	0,44
Cuartil 1	1,69	1,73	1,68	1,70
Mediana	1,98	1,96	1,96	1,99
Cuartil 3	2,24	2,26	2,24	2,25

Tabla 10: Valores del índice fotónico detectado para las zonas XI, X_I, X_OI, XI_O.

Se corrobora la hipótesis, de la misma manera que en el estudio del flujo en X y en la densidad columna, la figura y los datos numéricos que representan el índice fotónico detectado para fuentes detectadas y no detectadas en el infrarrojo considerando y no considerando las detecciones en el óptico, parece que siguen una misma distribución.

De la misma manera, de forma numérica no se aprecia la diferencia, por lo que se somete a la prueba K-S para tener certeza de este hecho. Los valores de los resultados de la prueba K-S de las propiedades espectrales se encuentran en la Tabla 11.

4.44 Resultados de la prueba K-S

Propiedades espectrales	XI / X_I	XI_O / X_OI
logFlux	0,66	0,74
logNH	0,014	0,028
PhoIndex	0,17	0,49

Tabla 11: Resultados del valor P de las propiedades espectrales a partir de la prueba estadística K-S para las comparaciones realizadas.

Teniendo en cuenta que típicamente el umbral para definir si dos muestras son similares o no es 0.0027 como se ha explicado previamente, se puede concluir que ambas muestras siguen la misma distribución, y el sesgo entre la detección en el

óptico y en rayos X no afecta, es más, cuando no se tiene en cuenta las detecciones en el óptico, las distribuciones se parecen aún más entre sí.

Capítulo 5

Discusión

5.1 Interpretación de los resultados

Los resultados del redshift y de la luminosidad indican un sesgo en el estudio debido a que las fuentes detectadas en rayos X y en FIR tienden a tener redshifts (y por tanto distancias) más pequeñas que las fuentes detectadas en el óptico.

Inicialmente, esto podría interpretarse como una indicación de que los cuásares detectados en el óptico se observan en una fase más temprana en el universo que los detectados en rayos X. Sin embargo, es importante considerar factores cruciales tal que la muestra de cuásares incluye objetos del SDSS, seleccionados principalmente por sus colores azules en el óptico, lo cual está asociado a menores absorciones en rayos X, introduciendo un sesgo. También hay que mencionar que los observatorios son más sensibles en el óptico que en rayos X o en FIR, de manera que en las últimas bandas sólo se pueden detectar a menores distancias.

A pesar de este sesgo, las propiedades espectrales X no han mostrado ninguna diferencia entre las distribuciones entre cuásares detectados y no detectados en FIR. Las muestras estudiadas respecto a las tres propiedades X, tal que el flujo en X, la densidad columna y el índice fotónico, pertenecen a la misma distribución, lo cual es un indicio de que el crecimiento de los AGN y de los agujeros negros supermasivos (rayos X) está relacionado con el crecimiento de estrellas y la formación estelar (FIR), y se podría interpretar que la actividad puede ser similar en diferentes fases de su detección espectral.

5.2 Comparación con estudios previos

Se han realizado investigaciones previas a este estudio, tal que [1], [2], [3], que indicaban diferencias entre AGN detectados y no detectados en FIR. Esta discrepancia puede deberse a varios factores:

Por un lado, los datos utilizados para este estudio son de XMM-Newton, el cual tiene una cobertura espectral más ancha que la de ROSAT, que es la utilizada para esos estudios. Al tener una banda más ancha que incluye un mayor porcentaje de energías más altas, los datos de XMM-Newton proporcionan medidas más precisas.

Por otro lado, las muestras utilizadas en los estudios anteriores son objetos identificados espectroscópicamente, mientras que los utilizados para este son cuásares de SDSS, seleccionados sobre todo por sus colores azules en el óptico, que suelen ir asociados a absorciones X más pequeñas. Por esta razón, se podría haber dado este sesgo en contra de los objetos más absorbidos en X, el cual se ha evitado haciendo la comparación entre XI₀ y X_{IO}.

5.3 Implicaciones para futuros estudios

Las limitaciones mencionadas resaltan la necesidad de futuros estudios para enfocarse en observaciones multibanda (óptico, infrarrojo, rayos X, submilimétrico) para mapear la evolución de los cuásares y sus galaxias anfitrionas a diferentes redshifts tratando de tener menos sesgos, como por ejemplo estudiar todo tipo de cuásares y no seleccionarlos por sus colores, etc. De esta manera, se podrían comparar distintas muestras de cuásares detectados sin ningún sesgo previo y obtener resultados más precisos.

Es crucial investigar la relación entre la absorción y la formación estelar en distintas fases evolutivas, desarrollar modelos teóricos y simulaciones más precisas. Comparar diferentes tipos de AGN también ayudará a identificar patrones comunes y diferencias en sus mecanismos de crecimiento y evolución.

Capítulo 6

Conclusiones

A raíz de los artículos [1][2][3], se ha propuesto comprobar la relación entre el crecimiento del AGN con el de la galaxia anfitriona con un número de datos mayor y más preciso. Para ello, se ha cumplido el objetivo principal, investigar si existen diferencias significativas en la emisión en el infrarrojo lejano (FIR) entre cuásares que presentan distintas propiedades espectrales en rayos X.

Para cumplir este objetivo, se han descargado una serie de archivos de una correlación previamente realizada entre las bandas del óptico y de rayos X, tras la cual se han correlacionado con catálogos de cuásares del SDSS. Una vez se ha tenido este subconjunto de fuentes, se ha introducido una nueva banda espectral, FIR, con la que se ha vuelto a hacer una correlación para obtener los subconjuntos necesarios para el análisis que se ha llevado a cabo.

Teniendo todos los subconjuntos, se han estudiado las propiedades de estos, empezando con el redshift y la luminosidad y observando un sesgo en estas detecciones ya que las fuentes X y FIR tendían a observarse a redshifts más bajos y por ende con menores luminosidades. Respecto a las propiedades espectrales X, tal que el flujo, la densidad columna y el índice fotónico, se han comparado los conjuntos X_I/XI y XI_O/X_OI para evitar el sesgo detectado anteriormente. Se ha observado tanto de manera visual como mediante un análisis estadístico que no hay una diferencia significativa en los conjuntos, por lo que coincidiendo con lo que se esperaba en un principio, parece que el crecimiento del AGN y la fase de formación estelar de la galaxia anfitriona podrían estar relacionados.

Gracias a este trabajo se ha aprendido el uso de lenguajes de programación como Python para hacer análisis estadísticos y también a utilizar una herramienta muy útil para el estudio de procesos astronómicos como es TopCat. Además, se han asimilado conceptos fundamentales como los cuásares y las propiedades espectrales necesarias para su estudio, tal que el redshift, la luminosidad, el flujo... para entender la

evolución de la galaxia de manera precisa y obtener la mayor información posible para realizar una conclusión. Además, se ha comprendido la importancia del estudio en diferentes bandas como el óptico, los rayos X o el infrarrojo lejano, para el análisis exhaustivo de las propiedades de las galaxias y AGN lejanos.

Capítulo 7

Bibliografía

[1] Stevens, J. A., Page, M. J., Mittaz, J. P. D., Carrera, F. J., Smail, I., & McHardy, I. M. (2004). Submillimeter Evidence for the Coeval Growth of Massive Black Holes and Galaxy Bulges. *Astrophysical Journal*, 604, 93-106.

[2] Page, M. J., Stevens, J. A., Ivison, R. J., & Carrera, F. J. (2004). The evolutionary sequence of active galactic nuclei and galaxy formation revealed. *Astrophysical Journal Letters*, 611(2), L85-L88.

[3] Stevens, J. A., Page, M. J., Ivison, R. J., Carrera, F. J., Mittaz, J. P. D., Smail, I., & McHardy, I. M. (2005). Submillimetre photometry of X-ray absorbed QSOs: their formation and evolutionary status. *Monthly Notices of the Royal Astronomical Society*, 360(2), 610-620.

[4] NASA's Webb Will Use Quasars to Unlock the Secrets of the Early Universe. NASA. Disponible en: [NASA]([NASA's Webb Will Use Quasars to Unlock the Secrets of the Early Universe - NASA](#)).

[5] NASA's Webb to Study Quasars and Their Host Galaxies in Three Dimensions. NASA. Disponible en: (<https://www.nasa.gov/universe/nasas-webb-to-study-quasars-and-their-host-galaxies-in-three-dimensions/>)

[6] Carroll, B. W., & Ostlie, D. A. (2017). *An Introduction to Modern Astrophysics* (2nd Edition). Pearson.

[7] Peterson, B. M. (1997). *An Introduction to Active Galactic Nuclei*. Cambridge University Press

[8] Sloan Digital Sky Survey (SDSS) SkyServer DR12. *SkyServer: DR12 Data Release*. Disponible en: <https://skyserver.sdss.org/dr12/en/home.aspx>

[9] "XMM2ATHENA WP6 D6.4: Fitting products with classification", Mountrichas et al., 2023

- [10] Massida, Z., et al. (2022). "Fitting products for extracted spectra", *XMM2ATHENA Project, Deliverable D6.1*. Universidad de Cantabria.
- [11] Massida, Z., et al. (2022). "Spectra and Fitting products." *XMM2ATHENA Project, Deliverable D6.2*. Universidad de Cantabria.
- [12] Instituto de Física de Cantabria (IFCA), University of Leicester, et al. "4XMM-DR11 XMM-Newton Serendipitous Source Catalogue", Julio 2024.
http://xmmssc.irap.omp.eu/Catalogue/4XMM-DR11/4XMM_DR11.html
- [13] Instituto de Física de Cantabria (IFCA), University of Leicester, et al. "4XMM-DR13 XMM-Newton Serendipitous Source Catalogue", Julio 2024.
http://xmmssc.irap.omp.eu/Catalogue/4XMM-DR13/4XMM_DR13.html
- [14] Carrera, F. J., Traulsen, I., Schwobe, A. D., Lamer, G., et al. (2020). "The XMM-Newton serendipitous survey. X. The second source catalogue from overlapping XMM-Newton observations and its long-term variable content." *Astronomy & Astrophysics*, 641, A137.
- [15] Carrera, F. J., Webb, N. A., Traulsen, I., et al. (2020). "The XMM-Newton serendipitous survey. IX. The fourth XMM-Newton serendipitous source catalogue." *Astronomy & Astrophysics*, 641, A136.
- [16] Carrera, F. J., Webb, N. A., et al. (2023). "XMM2ATHENA, the H2020 project to improve XMM-Newton analysis software and prepare for Athena." *Astronomische Nachrichten*, 344(2), 20102.
- [17] SPIRE in depth, Herschel Space Observatory, disponible en: [Herschel Telescope](#)
- [18] The Herschel-SPIRE instrument and its in-flight performance, arXiv, disponible en: [arXiv SPIRE](#)
- [19] S.J. Oliver, et al. "HerMES, "Monthly Notices of the Royal Astronomical Society," Volume 424, Issue 3, August 2012, Pages 1614-1635.
<https://cdsarc.u-strasbg.fr/viz-bin/cat/VIII/103#/browse>
- [20] Pineau, F.-X., Derriere, S., et al. (2017). "Probabilistic multi-catalogue positional cross-match." *Astronomy & Astrophysics*, 597, A89.
- [21] Taylor, M. B. (2005). "TOPCAT & STIL": Starlink Table/VOTable Processing Software". *Astronomical Data Analysis Software and Systems XIV*, 347, 29.
- [22] Astropy Documentation. "Astropy FITS". Recuperado de: <https://docs.astropy.org/en/stable/io/fits/>

[23] Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (1992). *Numerical Recipes in C: The Art of Scientific Computing* (2nd ed.). Cambridge University Press.

[24] SciPy Documentation - ks_2samp. "SciPy KS Test". Recuperado de: https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.ks_2samp.html

[25] NumPy Documentation. "NumPy". Recuperado de: <https://numpy.org/doc/stable/>

[26] Matplotlib Documentation. "Matplotlib". Recuperado de: <https://matplotlib.org/stable/contents.html>

Capítulo 8

Apéndice

El Código utilizado para realizar la prueba K-S es el siguiente:

```
from astropy.io import fits
from scipy.stats import ks_2samp
import numpy as np
import matplotlib.pyplot as plt
ruta_archivo_1 = 'Ruta del archivo'
ruta_archivo_2 = 'Ruta del archivo'
with fits.open(ruta_archivo_1) as hdul1, fits.open(ruta_archivo_2) as hdul2:
    data1 = hdul1[1].data # Ajusta el índice si es necesario
    data2 = hdul2[1].data # Ajusta el índice si es necesario
    columna_a = data1['Nombre de la columna']
    columna_b = data2['Nombre de la columna']
    columna_a = columna_a[~np.isnan(columna_a)]
    columna_b = columna_b[~np.isnan(columna_b)]
    resultado_ks = ks_2samp(columna_a, columna_b)
    print("D:", resultado_ks.statistic)
    print("P:", resultado_ks.pvalue)
    sorted_data1 = np.sort(columna_a)
    sorted_data2 = np.sort(columna_b)
    cumulative1 = np.arange(1, len(sorted_data1) + 1) / len(sorted_data1)
    cumulative2 = np.arange(1, len(sorted_data2) + 1) / len(sorted_data2)
```

```
plt.figure(figsize=(10, 6))
plt.step(sorted_data1, cumulative1, label='Muestra 1', where='post')
plt.step(sorted_data2, cumulative2, label='Muestra 2', where='post')
plt.title('Titulo')
plt.xlabel('Eje x')
plt.ylabel('Función de Distribución Acumulativa (FDA)')
plt.legend()
plt.grid(True)
plt.show()
```