



Universidad Zaragoza

Trabajo Fin de Máster

*Máster Universitario en Física del Universo: Cosmología, Astrofísica,
Partículas y Astropartículas.*

***J-PLUS: las regiones de formación estelar extrema
en el Universo a bajo desplazamiento al rojo***

Autor

Jorge Porrón Lafuente

Director

Alejandro Lumbreras-Calle

FACULTAD DE CIENCIAS

Departamento de Física Teórica

Junio 2023

Índice

1. Introducción.....	1
2. SED fitting y selección de objetivos.....	3
a. Construcción de la SED.....	4
b. Selección de objetivos.....	6
3. Inspección visual.....	8
a. Obtención de imágenes y filtros.....	8
b. Comparación visual con Legacy Survey.....	9
4. Detección y medición automatizada de clumps.....	11
a. Identificación de fuentes (NoiseChisel).....	12
b. Identificación de clumps (Segment).....	12
c. Catálogos (MakeCatalog).....	15
d. J-PLUS vs Legacy Survey.....	15
e. Muestra final.....	17
5. CIGALE.....	19
6. Análisis de los resultados.....	20
a. Masa del Clump – Masa de la galaxia que lo alberga.....	21
b. Masa del Clump – Distancia al centro de la galaxia.....	22
c. Star Formation Rate – Distancia al centro de la galaxia.....	23
d. SFR específica (sSFR) – Distancia al centro de la galaxia.....	24
e. Tamaño del clump – Distancia al centro de la galaxia.....	24
f. Masa de la Población Joven del Clump – Masa de la Población Joven de la galaxia.....	25
g. Edad del Clump – Distancia al centro de la galaxia.....	26
h. Luminosidad de las líneas - Distancia al centro de la galaxia.....	26
i. Resultados generales.....	28
7. Conclusiones.....	28
8. Referencias.....	29

1. Introducción

La observación y estudio de galaxias es muy importante en nuestra comprensión de la formación y evolución del cosmos. La gran revolución cosmológica se desarrolló durante la primera mitad del S. XX cuando astrónomos como Hubble o Eddington cambiaron nuestro entendimiento sobre el universo. Comenzando con los trabajos de Messier¹ o Herschel² y sus catálogos de objetos difusos, y más tarde con Shapley, Curtis y su *Gran Debate*³ sobre la naturaleza galáctica o extragaláctica de estos objetos, fue Edwin Hubble quien cambió nuestra comprensión del Universo. Con su observación de la expansión del universo⁴ y la distinción de estrellas en la (entonces conocida como) *Nebulosa de Andrómeda*, hizo uso de la relación periodo-luminosidad de Leavitt para cefeidas y estimó una distancia de 1.5 millones de años luz, mucho mayor que la de las estrellas más lejanas entonces conocidas en la galaxia, poniendo fin al *Gran Debate*.

Con el desarrollo de nuevas tecnologías y técnicas de observación se han ido abriendo nuevas ventanas en el análisis de estos objetos astronómicos. Históricamente, las galaxias se han clasificado de acuerdo con su morfología, siendo la clasificación más usada hoy en día la *Secuencia de Hubble*. Tiene este nombre porque originalmente se pensaba que las galaxias evolucionaban por ella, desde una morfología elíptica hasta morfologías irregulares pasando por formas lenticulares y espirales. Hoy sabemos que esta evolución no ocurre así y cada tipo de galaxia tiene unas propiedades, historia y origen más complejo.

La observación de galaxias más lejanas supone verlas como eran hace miles de millones de años, en las primeras etapas de su desarrollo, y se ha observado una característica interesante: mientras que las galaxias cercanas (a bajo *redshift*) son principalmente elípticas y espirales, observaciones del universo lejano (a alto *redshift*) presentan una morfología más caótica⁵. Galaxias de formación estelar típica (SFGs por su acrónimo en inglés *Star-Forming Galaxies*) tienen una apariencia simétrica y ordenada, habitualmente espiral, en el Universo cercano. Sin embargo, sus contrapartes a alto *redshift* son más irregulares, destacando la presencia de granos (a los que a partir de ahora nos referiremos como *clumps*) en su observación, que corresponden a regiones de formación estelar intensa.

La toma de imágenes de galaxias suele hacerse a través de distintos filtros que permiten estudiar distintas propiedades de ellas, muestreando distintas regiones del espectro electromagnético. Una herramienta crucial para el estudio de su evolución son los cartografiados del Universo, donde se toman imágenes en distintos filtros y se hacen catálogos de objetos con propiedades similares que nos permiten, por ejemplo, observar esta relación entre morfología caótica y *redshift*.

Para este trabajo se hace uso de una selección de galaxias en el universo cercano ($z < 0.06$) que presentan una característica interesante: sufren de formación estelar extrema (en adelante *EELGs* por *Extreme Emission Line Galaxies*). Para galaxias suficientemente grandes esto se traduce en la presencia de *clumps* en su observación (galaxias más pequeñas se observan como un único *clump* de formación estelar extrema). El hecho de que galaxias tan cercanas posean características típicas de galaxias lejanas puede convertirlas en objetivos ideales para el estudio del Universo primitivo.

Recientemente, se ha llevado a cabo una selección⁶ de 466 *EELGs* haciendo uso de la *Data Release 2* del catálogo *Javalambre Photometric Local Universe Survey (J-PLUS⁷)*, un proyecto de cartografiado fotométrico conducido en el Observatorio Astrofísico de Javalambre (OAJ⁸) usando el telescopio *JAST80* y

la cámara *T80Cam*⁹. El sistema de filtros de *J-PLUS* está compuesto por 12 filtros (*Figura 1*), entre los que destacamos el filtro *J0515*, con longitud de onda central de 5150Å y *Full Width at Half Maximum (FWHM)* de 200Å, pues una de las principales características de esos *clumps* de formación estelar extrema es la gran emisión de la línea de emisión nebular de *[OIII]4959+5007*, que para galaxias a un *redshift* de entre $0.007 < z < 0.06$ queda en esa banda.

Filter	Central wavelength [Å]	FWHM [Å]
<i>u</i>	3485	508
<i>J0378</i>	3785	168
<i>J0395</i>	3950	100
<i>J0410</i>	4100	200
<i>J0430</i>	4300	200
<i>g</i>	4803	1409
<i>J0515</i>	5150	200
<i>r</i>	6254	1388
<i>J0660</i>	6600	138
<i>i</i>	7668	1535
<i>J0861</i>	8610	400
<i>z</i>	9114	1409

Figura 1: Filtros de J-PLUS.

La elección de esta línea de *[OIII]5007* como criterio para considerar formación estelar se debe a los distintos procesos físicos que suceden en las regiones de galaxias con formación estelar activa. Por un lado, las estrellas evolucionan quemando su combustible en distintas fases y el tiempo durante el que lo queman depende de su masa, las estrellas más masivas agotan su combustible muy rápidamente (rápido en términos cosmológicos, las estrellas tipo O evolucionan quemando hidrógeno a lo largo de su *Secuencia Principal* durante unos 10^6 años mientras que las de tipo G como el Sol lo hacen durante 10^9 años). Por lo tanto, si se observan estrellas masivas de este tipo en una región particular de la galaxia podemos asumir que se están formando recientemente. La identificación de estas estrellas jóvenes e intensas se hace analizando el espectro, sabemos que en las regiones de formación estelar hay mucho gas y polvo interestelar que puede ser ionizado por la energía emitida por estas estrellas masivas. A lo largo de la evolución de la galaxia antes de la aparición de este brote, muchas de las estrellas masivas de corta vida habrán explotado en forma de supernova expulsando elementos más pesados como el oxígeno que enriquecen el gas y el polvo y pueden ser ionizados. Por las características de presión, densidad, y las propiedades atómicas, en este tipo de galaxias (pequeñas, con metalicidad no muy alta), la línea de *[OIII]5007* es muy intensa, habitualmente la más intensa⁶.

Así, se trata de una elección conveniente para nuestro estudio particular, pero no la única, otra posibilidad sería, por ejemplo, usar la línea de H_{α} para buscarlas. El problema es que el filtro de *J-PLUS* que ofrece información sobre este pico es más estrecho que *J0515* y deberíamos restringirnos a galaxias más cercanas, ya que el corrimiento debido al *redshift* sacaría el pico de nuestro filtro para galaxias a *redshifts* más pequeños que los que se pueden estudiar en *J0515*, cubriendo así un volumen de universo más pequeño y contaminando nuestra selección de galaxias con regiones pequeñas de galaxias muy cercanas a nosotros.

El objetivo de este *TFM* es el de realizar un análisis preliminar de estas *EELGs* y de sus *clumps*, con el fin de obtener información sobre la estructura, propiedades físicas (posición, tamaño, brillo, ...) y evolución de estos *clumps* en relación con su *EELG* anfitriona. Para ello, se escogen 125 de estas *EELGs* clasificadas como extensas (ya que las compactas o semicompactas son un único *clump* de formación estelar extrema) y se

hace uso de distintas herramientas de análisis de datos y edición de imagen para obtener información sobre ellas.

La estructura del trabajo es la siguiente: en el apartado 2 se introducen los conceptos teóricos que usaremos para el análisis. En el apartado 3 se realiza una primera inspección visual de las imágenes para dos cartografiados (*J-PLUS* y *Legacy Survey*). En el apartado 4 se introducen y aplican a nuestra muestra las herramientas de detección automatizada de *clumps*. En el apartado 5 se presentan los parámetros que introducimos en el software de análisis de los *clumps* identificados. En el apartado 6 se analizan los resultados obtenidos. Finalmente, en el apartado 7 se presentan las conclusiones finales.

2. Sed fitting y selección de objetivos

La definición de *EELG* varía entre diferentes trabajos, aunque suele aceptarse como un límite en la *Equivalent Width (EW)* de la línea $[OIII]5007$, siendo la *EW* una medida de la intensidad de las líneas espectrales. Estas líneas espectrales tienen un tamaño particular en función de la naturaleza de la fuente, la *EW* corresponde a la anchura sobre el continuo del espectro a la que corresponde el área de la línea espectral. Una *EW* grande implica una gran certeza sobre la presencia del pico al que corresponde.

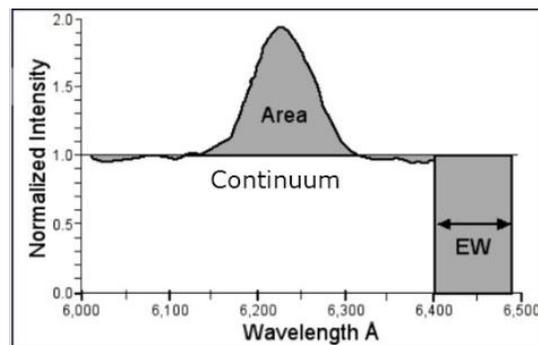


Figura 2: *EW* de una línea de emisión.

En la literatura se toman entre 100\AA y 300\AA de *EW* para esta línea de $[OIII]5007$, en nuestro caso calificaremos como fuertes emisores en *J0515* aquellos para los que aproximadamente $EW(J0515) > 200\text{\AA}$ ⁶ y tras un análisis espectral y la eliminación de posibles contaminantes nos quedaremos con aquellos para los que $EW([OIII]5007) > 300\text{\AA}$.

Las galaxias emiten radiación electromagnética en todo el espectro. La distribución de energía sobre la longitud de onda, λ , (o frecuencia) se llama *Spectral Energy Distribution (SED)* y es su análisis el que nos ofrece la información física que buscamos. Los diferentes procesos que ocurren en la galaxia dejan una impronta en la *SED*, cada uno dominando en unas longitudes de onda particulares. Para producir la *SED* se analiza la cantidad de fotones que se recibe de una galaxia con una energía (longitud de onda) particular. Se puede hacer directamente espectroscopía para una fuente particular, pero es muy costoso observacionalmente y es común ajustar el espectro a observaciones fotométricas como las nuestras con

este proceso de *SED fitting*, con el que se modela el espectro en base a las observaciones fotométricas obtenidas con nuestros telescopios y filtros.

Al observar una galaxia en distintos filtros vemos cuántos fotones se emiten desde una región particular con una energía particular, cuantos más filtros se utilicen y más estrechos sean, con más precisión podremos medir la *SED* de la galaxia. Lo mejor que podemos hacer es reducir al máximo el ruido de la señal recibida al observar una galaxia y usar la mayor cantidad posible de filtros lo más estrechos posibles. Esta es una de las principales razones para el uso de *J-PLUS*, sus 12 filtros (7 estrechos y 5 anchos, *Figura 1*) nos permiten obtener una *SED* razonablemente precisa de cada galaxia, a diferencia de *Legacy Survey*, por ejemplo, que sólo usa 4 filtros anchos con los que se puede reproducir una *SED*, pero ofrecerá muy poca información. Un ejemplo de proyecto con intención de mejorar la precisión en este sentido es *J-PAS*¹⁰ que se desarrollará en *OAJ* y hará uso de 56 filtros estrechos.

La radiación emitida por las galaxias es producto de los distintos objetos que las componen (y que emiten radiación electromagnética), en particular las estrellas, el gas y el polvo. Así, cabe esperar que una hipotética galaxia compuesta por estrellas muy masivas y calientes presentase una *SED* distinta a una galaxia con estrellas más frías, en particular se observaría un exceso en la cantidad de fotones de mayor energía (menor longitud de onda por $E = h \cdot c / \lambda$) emitidos por estas estrellas calientes. Este es un ejemplo sencillo de cómo un estudio de la *SED* nos puede dar información sobre la naturaleza (o composición en este caso) de la galaxia. En realidad, la situación no es tan sencilla y la *SED* depende de gran cantidad de parámetros, incluso en este ejemplo sencillo habría que tener en cuenta muchos más efectos como la cantidad de estrellas de cada tipo, la dependencia de la extinción con la longitud de onda, ...

Una vez disponemos de la información sobre la cantidad de fotones emitidos en cada uno de nuestros filtros, podemos crear nuestra *SED* observada para cada galaxia o *clump* particular. Al disponer sólo de 12 filtros, esta *SED* será una aproximación de la *SED* real, pero es suficiente para obtener resultados razonables. Para producir la *SED* a partir de nuestra información fotométrica se hace uso del software *CIGALE*, que varía los distintos parámetros de los que depende la *SED* para generar *SEDs* sintéticas que compara con nuestros resultados observacionales, permitiéndonos obtener la *SED* generada por *CIGALE* más similar a nuestra *SED* obtenida con la información del flujo de fotones en nuestros filtros.

a. Construcción de la *SED*

Para modelar el espectro lo más importante es entender qué procesos y de qué manera le afectan. *J-PLUS* es un proyecto de cartografiado fotométrico en el rango óptico, que ocupan los filtros de la *Figura 1*, y en este rango los elementos de la galaxia que emiten radiación electromagnética y forman su espectro son:

1. **Estrellas:** Para modelar la luz emitida por las estrellas se utilizan predicciones para el espectro de un grupo de estrellas particular: *Simple Stellar Population (SSP)*, que dan cuenta de la luz emitida por un grupo de estrellas que nacen a la vez o siguiendo alguna regla simple (decaimiento exponencial, por ejemplo). En nuestra galaxia de estudio no esperamos que todas las estrellas tengan la misma edad y metalicidad, sino que se hacen uso de varios *SSP* para los distintos subconjuntos de estrellas que la compongan. Por ejemplo, podemos suponer que una galaxia está compuesta sólo por estrellas de $2 \cdot 10^6$ años con metalicidad solar y por estrellas de $5 \cdot 10^9$ años

y el doble de la metalicidad solar. Cada uno de estos subconjuntos será una *SSP* característica y al combinarlas obtendremos la *CSP* (*Composite Stellar Population*) de la galaxia. El peso relativo en masa y luminosidad de las distintas *SSP* puede variar de una galaxia a otra en función de las propiedades de las estrellas que la componen (como la cantidad de estrellas de cada tipo).

Para nuestro grupo idealizado de estrellas con una masa y metalicidad concreta hay un parámetro clave que los distingue: como las estrellas más masivas viven poco tiempo, las *SSPs* de gran edad carecerán de ellas y darán lugar a un espectro diferente del de las *SSPs* de menos edad que todavía contengan gran cantidad de estrellas masivas. El espectro de una *SSP* de edad t y metalicidad Z viene dado por la suma de los espectros de todas las estrellas individuales y dependerá de su masa. El flujo en una λ particular viene dado por:

$$f_{SSP}(t, Z) = \int_{m_{low}}^{m_{up}(t)} f_{star}[T_{eff}(M), \log g(M) | t, Z] \phi(M) dM$$

Se integra desde las estrellas de menor masa hasta las de mayor masa, donde como vemos sí hay una dependencia con t debido a lo comentado anteriormente, para una *SSP* de gran t las estrellas más masivas serán menos masivas que para una *SSP* de menor t donde estas estrellas masivas aún no han explotado. El flujo total es la suma del flujo de cada estrella individual (f_{star}), que depende de su temperatura y gravedad de superficie, multiplicado por el número de estrellas que esperamos de esa masa ($\phi(M)$).

Las estrellas más masivas son más luminosas y dominarán la *SSP*, pero también son menos numerosas. La cantidad de estrellas de una masa particular viene caracterizada por la *Initial Mass Function* (*IMF*), una distribución de probabilidad de la masa con la que una estrella entra en la *Secuencia Principal*. Se usan distintos perfiles, entre los que destacan el de Salpeter¹¹ ($\phi(M) \propto \left(\frac{M}{M_{solar}}\right)^{-2.35}$) o más recientemente algunos más complicados como el de Chabrier¹².

Finalmente, conocida la cantidad de estrellas de cada masa en nuestra *SSP* de una edad y metalicidad particulares, se puede hacer uso de librerías espectrales para conocer el espectro completo de cada clase estelar y combinarlos para obtener nuestra *SSP*:

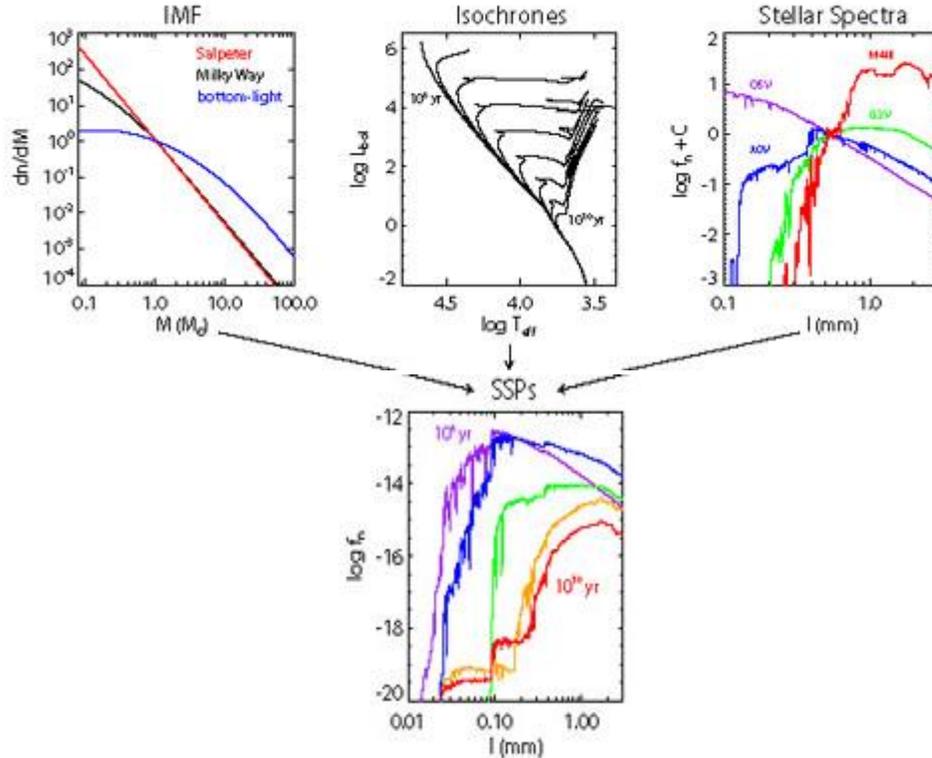


Figura 3: SSPs para distintos t , las de menor t (morado, azul) son dominadas por las altas energías de las estrellas masivas. Imagen tomada de Ref. ¹³.

2. **Polvo:** Otro elemento importante en la SED de nuestras galaxias será el polvo que emite y absorbe luz. Para modelar la extinción se usan curvas de extinción que dan cuenta de la cantidad de luz que esperamos sea bloqueada por el polvo en función de la λ . Depende del tamaño y tipo de los granos de polvo y de su cantidad en la galaxia. Estas curvas de extinción se obtienen observando estrellas dentro de la *Vía Láctea* y la *Gran Nube de Magallanes*, donde se pueden resolver individualmente.

Hay que tener en cuenta que, además de absorberla, el polvo puede dispersar la luz, así, una nube de polvo puede extinguir luz en una dirección, pero dispersarla en otra. Cuando se tienen en cuenta tanto la extinción como la dispersión debida a una nube de polvo se habla de atenuación, y es considerada en nuestros modelos para las SED. La atenuación debida al polvo es mayor para λ grandes, de modo que un modelo con mucho polvo afectará a la SED disminuyendo esa parte del espectro galáctico.

3. **Gas:** La mejor forma de estudiar el gas es con las líneas de emisión y absorción. El problema es que, al hacer fotometría, los filtros anchos no suelen ser suficientemente precisos para resolver líneas individuales, por esa razón nuestros filtros estrechos resultan especialmente útiles para analizar las propiedades del gas.

b. Selección de objetivos

Para la selección⁶ de *EELGs* se usa el lenguaje *Astronomical Data Query Language (ADQL)* para seleccionar objetos con ciertas propiedades en la base de datos de *J-PLUS*. Partiendo de 30336 objetos que presentan un exceso en el filtro *J0515* se hace un *Cross-Match* con la *Data Release 16* de *Sloan Digital Sky Survey (SDSS¹⁴)* y el catálogo *unWISE¹⁵* del satélite *Wide-Field Infrared Survey Explorer (WISE¹⁶)* para eliminar contaminantes como cuásares o estrellas que pueden presentar un exceso en *J0515* debido a otras causas.

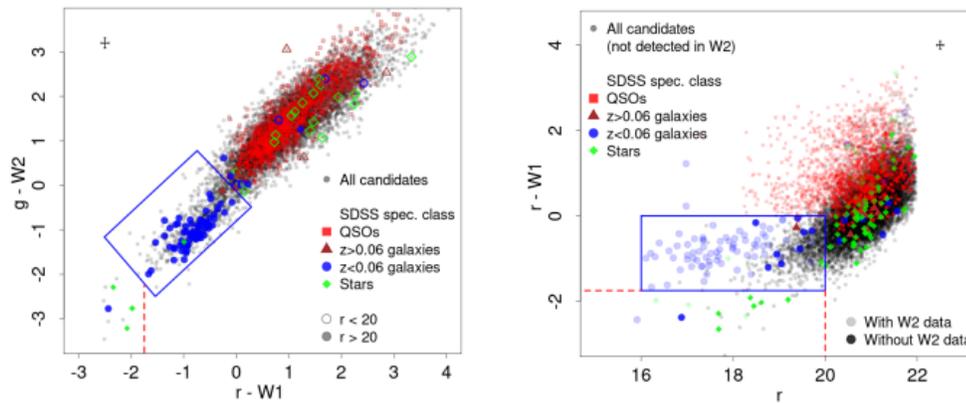


Figura 4: Diagrama color-color para separar los distintos tipos de objetos. Los datos de *r* y *g* se obtienen de *J-PLUS DR2* y los de *W1* y *W2* del catálogo *unWISE*. Los puntos grises son nuestros candidatos, de los que se colorean aquellos cuya naturaleza es conocida en la *Data Release 16* de *Sloan Digital Sky Survey (SDSS)*: en rojo los cuásares, en azul las galaxias a $z < 0.06$ (nuestro objetivo), en marrón las galaxias con $z > 0.06$ y en verde las estrellas. Los círculos azul claro en la imagen de la derecha son los objetos de los que se disponen medidas de *W2*. Figura tomada de *Lumbreras-Calle et al. 2021⁶*.

Se observa claramente cómo la mayoría de nuestros objetivos de interés identificados quedan en una región (rectángulo azul), de modo que nos quedamos con los candidatos contenidos en ella (la explicación de esta distinción queda fuera de los límites de este trabajo, ver ¹⁷).

Finalmente, para confirmar y caracterizar la naturaleza de estos candidatos a *EELG* se hace un ajuste de la *SED* de cada uno con el software *CIGALE^{18,19}* ajustando su *SED* a distintos parámetros razonables⁶. Un ejemplo de ajuste como éste para una galaxia particular se puede ver en la *Figura 5*.

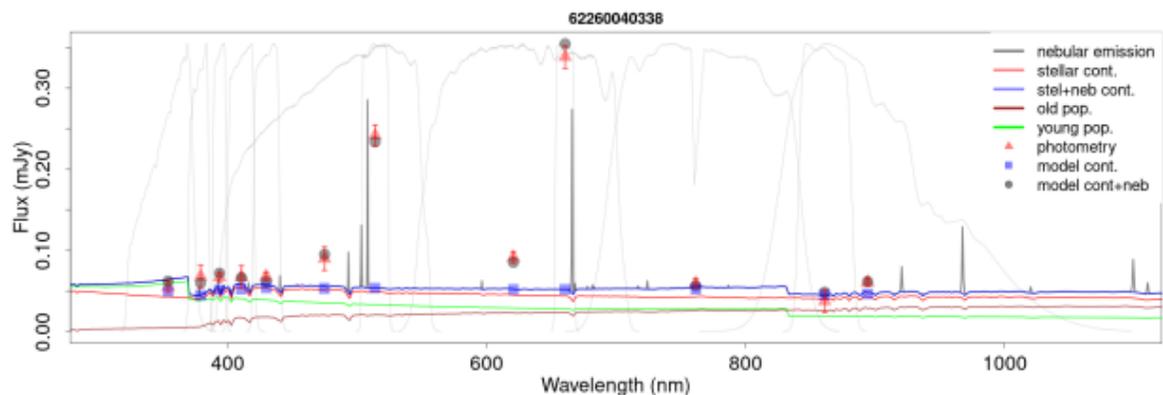


Figura 5: Ejemplo de la SED de una galaxia de J-PLUS con el ajuste de CIGALE. Los triángulos rojos corresponden a la fotometría observada en cada uno de los 12 filtros de J-PLUS, los puntos grises y cuadrados azules son la fotometría sintética de CIGALE con y sin las líneas de emisión. Las líneas gruesas son las diferentes componentes del espectro calculado por CIGALE más aproximado al observado para el subconjunto mencionado en la leyenda y las líneas grises finas son los perfiles de transmisión de los 12 filtros. Figura tomada de Lumbreras-Calle et al. 2021⁶.

Con el ajuste SED podemos calcular con precisión el ancho equivalente de $[OIII]5007$. Seleccionando aquellas para las que $EW([OIII]5007) > 300\text{\AA}$, queda una muestra pura de 466 EELGs, de las que tomamos 90 extensas, además de 35 extra que no cumplen el criterio de $EW([OIII]5007) > 300\text{\AA}$ en toda la galaxia, pero sí en alguno de sus *clumps*. Obtenemos una muestra final de 125 EELGs extensas que presentan un claro exceso en $[OIII]5007$ que identificamos como formación estelar extrema.

3. Inspección visual

Nuestro objetivo principal es el de analizar en detalle las imágenes de las galaxias en distintos filtros con el fin de detectar y estudiar las propiedades de los *clumps*. Para ello, comenzamos por comparar las imágenes obtenidas en J-PLUS con las de *Legacy Survey*, que son las de mejor resolución espacial disponible.

a. Obtención de imágenes y filtros

Para obtener información interesante muchas veces no basta con la simple observación de los objetos, sino que es necesario almacenar esta información y trabajar con los datos obtenidos. En este campo de la *Astronomía Digital*, el mayor avance se realizó con el desarrollo en 1969 de los primeros dispositivos CCD (*Charge-Coupled Device*) por Boyle y Smith en los *Laboratorios Bell* (invento por el que fueron galardonados en 2009 con el premio Nobel²⁰) con los que se obtiene información de la cantidad de fotones que reciben.

Cuando se observa una región del cielo, los píxeles que apuntan a una zona que presenta un emisor de fotones (como puede ser una galaxia) detectarán más fotones que los píxeles que apuntan a regiones vacías. Si la resolución espacial es suficiente, se podrá descomponer el objeto en subregiones más o menos intensas. Además, se pueden poner distintos filtros que sólo permitan el paso a los CCDs de fotones con unas energías concretas, así, un mismo objeto puede tener dos morfologías totalmente distintas si lo miramos a través de uno u otro filtro:

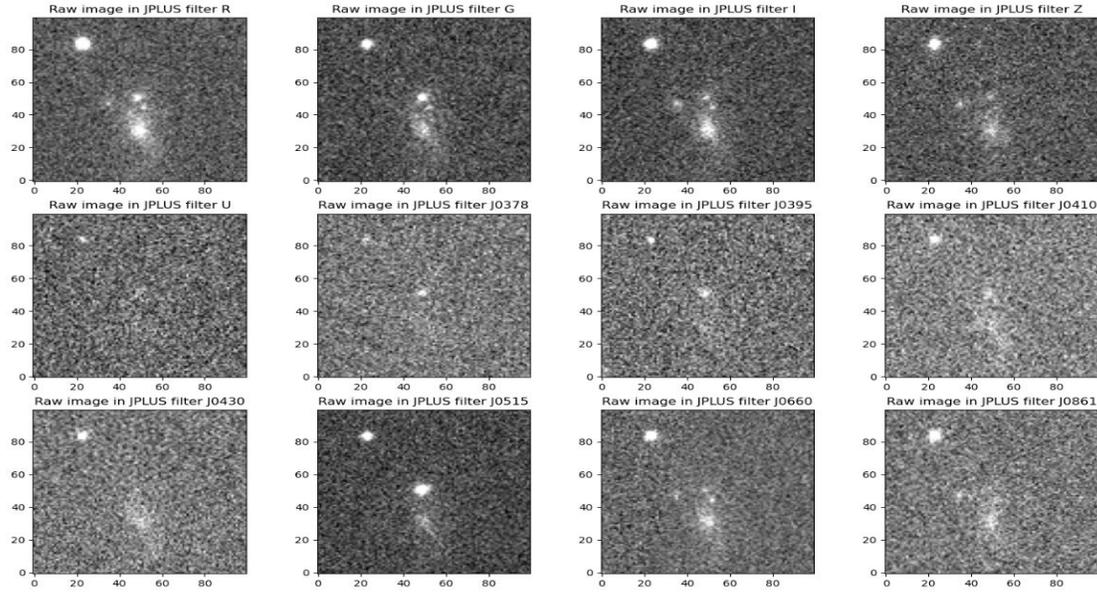


Figura 6: Observación de una misma EELG (73268016484) en cada uno de los 12 filtros de J-PLUS. La EELG es el objeto central, la región de la esquina superior izquierda que se observa en todos los filtros es un objeto externo como puede ser una estrella cercana que se observa en esa posición en la bóveda celeste. Como vemos dependiendo del filtro la galaxia tiene una u otra morfología, apenas emitiendo en energías contenidas en el filtro U o haciéndolo solo desde una compacta región en J0378.

b. Comparación visual con Legacy Survey

Nuestro primer paso es asegurarnos de que las galaxias seleccionadas presentan estas regiones de formación estelar extrema. Como hemos mencionado, la selección de EELGs se ha hecho a partir del cartografiado J-PLUS del OAJ, que hace uso de los filtros de la Figura 1, pero éste no es el único cartografiado disponible. En particular, al tratarse de galaxias cercanas todas han sido observadas por otros cartografiados con diferentes objetivos y características, aunque casi ninguno con filtros estrechos como los de J-PLUS.

El cartografiado disponible que presenta la mejor resolución espacial es Legacy Survey, que en su Data Release 10 identifica 2826169461 objetos²¹ combinando diferentes proyectos de mapeado del hemisferio norte realizados con distintos telescopios:

- The Beijing-Arizona Sky Survey (BASS)²²
- The DECam Legacy Survey (DECaLS)²³
- The Mayall z-band Legacy Survey (MzLS)²⁴

Pese a la mayor profundidad de las imágenes de Legacy Survey, solo se usan los filtros *g*, *r*, *i* y *z*. Las imágenes obtenidas en estos filtros para Legacy Survey tienen mejor resolución espacial que J-PLUS (0.25" por píxel frente a 0.55" por píxel) y *seeing atmosférico*. Gracias a ello, en ellas se aprecian claramente regiones más brillantes dentro de las galaxias. Sin embargo, como no se dispone de información de nuestro filtro de interés J0515, no podemos estar del todo seguros de que esas regiones correspondan a los *clumps*

de formación estelar extrema. En estos filtros se pueden observar regiones más brillantes dentro de las galaxias, pero el mecanismo de su producción puede ser distinto. Un ejemplo revelador es el hecho de que en ellas se observa una región más brillante en el centro que no siempre se debe a una formación estelar extrema, sino a la mayor densidad estelar en el centro de las galaxias.

En cualquier caso, podemos usar *Legacy Survey* para comparar las imágenes obtenidas en sus mismos filtros para *J-PLUS* y decidir en base a lo observado si *J-PLUS* identifica suficientemente bien zonas más brillantes dentro de la galaxia. Estas zonas más brillantes pueden ser unas u otras y tener una naturaleza distinta en función del filtro utilizado, pero si *J-PLUS* detecta razonablemente bien esas regiones en los filtros *g* y *r* para los que sabemos que *Legacy Survey* las detecta, podemos asumir que detectará lo suficientemente bien regiones con un exceso de brillo en el filtro *J0515*, que ahora sí, corresponderán a nuestros *clumps* de formación estelar extrema de interés.

Un ejemplo bastante ilustrativo es una de las galaxias estudiadas (*Figura 7*):

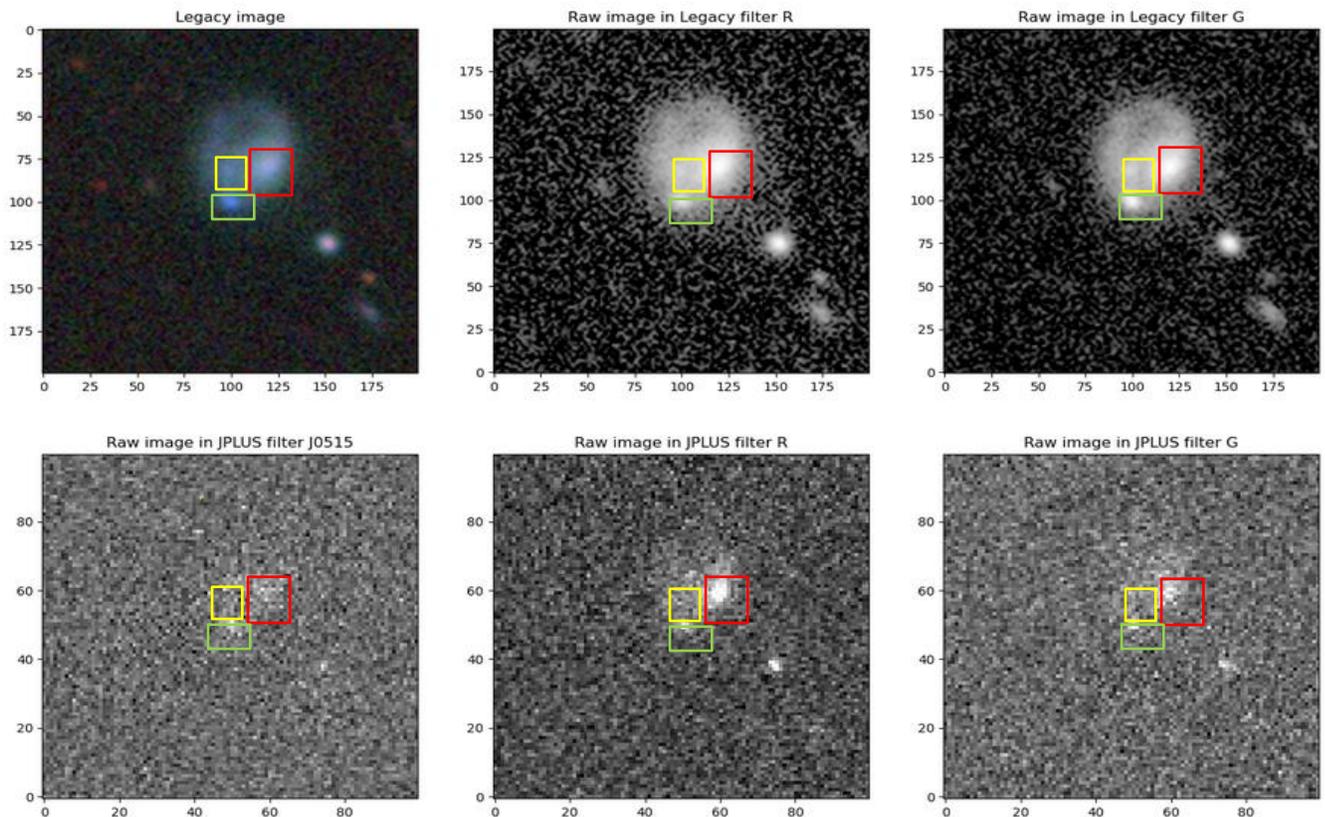


Figura 7: Comparación de la observación de una misma galaxia (68417017612) en los filtros r y g de Legacy Survey y J0515, r y g de J-PLUS. Las imágenes de los filtros r y g de Legacy Survey vienen dadas en escala logarítmica por claridad, pues la gran precisión de Legacy Survey comparado con J-PLUS hace que se detecten muchos más fotones en todos los filtros y las imágenes sean muy brillantes en escalas

polinómicas como las que se muestran para *J-PLUS*. La imagen de arriba a la izquierda es una composición de las imágenes en los distintos filtros de *Legacy Survey* (r, g, z, i).

En la imagen de arriba a la izquierda en la *Figura 7* podemos destacar 3 regiones:

- Una zona encuadrada en rojo en la que se observa claramente un exceso de brillo en los filtros g y r tanto para *Legacy Survey* como para *J-PLUS* que, sin embargo, no se aprecia en la imagen de *J0515* de *J-PLUS*, dada en la imagen inferior izquierda.
- Una región encuadrada en verde que se observa claramente en todos los filtros. El hecho de que sí se vea en el filtro *J0515* indica que es uno de los *clumps* de formación estelar extrema que nos interesa.
- Una región encuadrada en amarillo distinguible en las imágenes de *Legacy Survey* pero no en las de *J-PLUS*. Son este tipo de detecciones en *Legacy Survey* pero no en *J-PLUS* las que nos podrían dar problemas como el de rechazar algunos *clumps* que sí están disponibles usando las imágenes de *Legacy Survey* pero no las de *J-PLUS*.

Hay que tener en cuenta que la imagen de arriba a la izquierda en la *Figura 7* es una combinación de varios filtros y no podemos sencillamente decir que *J-PLUS* nos hace perder muchas detecciones al ver que en esa imagen hay una región de exceso de brillo bastante clara que no se ve en ninguna de las imágenes de *J-PLUS*, al ser una combinación de 4 filtros un pequeño exceso en cada filtro en esa región se amplifica al combinarlos, hay que comparar filtro a filtro.

Aun asumiendo que también hay una detección en esa zona amarilla en los filtros g y r de *Legacy Survey*, hay que recordar el objetivo del estudio: queremos realizar un análisis preliminar de las regiones de formación estelar extrema. Estas detecciones dudosas que podrían servir como argumento para exigir el uso de las imágenes de *Legacy Survey*, en realidad no son importantes. Queremos tener la seguridad de que la región que seleccionamos es realmente un *clump* de formación estelar extrema y es preferible perder un posible *clump* débil antes que considerar por error como *clump* una región cuyo exceso en el brillo fuera debido a otra causa, como un exceso en la densidad estelar en esa región.

Un caso muy claro de una región como ésta es la zona roja, en la que vemos claramente un exceso en todos los filtros salvo en *J0515*, lo que indica que no es una región de formación estelar extrema, a diferencia de la zona verde que se observa en todos los filtros y es para este ejemplo el único *clump* de interés.

Este análisis visual es interesante, pero demasiado subjetivo como para poder concluir que *J-PLUS* detecta lo suficientemente bien los *clumps* en nuestras *EELGs*, por lo tanto, utilizaremos herramientas automáticas de detección que se describen en la siguiente sección.

4. Detección y medición automatizada de clumps

Como acabamos de ver, la identificación visual de *clumps* puede ser una herramienta útil, pero para ser rigurosos en su selección y para poder trabajar con los datos es necesario hacer uso de softwares de análisis, en particular vamos a usar *GNU Astronomy Utilities*²⁵ (*GNUAstr*), un paquete del sistema

operativo de tipo Unix *GNU's Not Unix*²⁶ (GNU). En particular, *GNUAstro* es un paquete que contiene programas y librerías para la manipulación y análisis de datos astronómicos.

Es importante entender las principales utilidades del programa que usaremos si se quiere mejorar la experiencia del usuario, adaptando su uso a nuestras necesidades particulares.

a. Identificación de fuentes (NoiseChisel²⁷)

Hasta aquí, nuestra imagen no es más que un array de píxeles con un valor para cada píxel que caracteriza el número de fotones detectado por el CCD. Solo con esta información no podemos obtener resultados interesantes como la cantidad de galaxias que hay en la imagen o su magnitud. La clasificación de los píxeles entre ruido y señal es la detección, y es el paso más importante en el análisis de la imagen. Como las galaxias observadas no tienen una separación clara con el fondo hay que aplicar un algoritmo de detección que nos permita identificar y limitar objetos difusos como un conjunto de píxeles en nuestra imagen. *NoiseChisel* se ocupa de esto y su *output* es una imagen binaria, donde los píxeles valen = 1 si corresponden a un objeto detectado o = 0 si son ruido (*Figura 8*).

Para la identificación de señal, *NoiseChisel* utiliza un algoritmo en pasos, comenzando por la identificación de los píxeles que superan un valor de señal concreto. Estas *detecciones iniciales* se caracterizan luego como detección real o no en base a su *Signal-to-Noise Ratio (S/N)*, donde el ruido base de cada detección es la media del valor de los píxeles alrededor de él que no se calificaron como detección en el primer paso. Dependiendo de la intensidad de la fuente, muchas de estas *detecciones iniciales* estarán aisladas en píxeles particulares rodeados de ruido en las regiones menos intensas de la fuente. Para producir objetos compactos en la detección, *NoiseChisel* aumenta artificialmente la señal tanto de las *detecciones iniciales* como de los píxeles contiguos a ellas que son ruido. De esta forma, una fuente poco intensa quedará en su mayor parte camuflada en el ruido, pero algunos píxeles separados se detectarán como *detección inicial*, y al aumentar artificialmente la señal de sus píxeles contiguos (sean detección o no) se consiguen identificar regiones compactas de señal que corresponden a la fuente.

Pese a que ahora sabemos qué píxeles corresponden a una señal y cuáles no, aún no sabemos cuántos objetos se han detectado o qué píxeles corresponden a qué objeto particular.

b. Identificación de clumps (Segment²⁸)

Si los objetos están claramente separados y no están en contacto entre ellos, la distinción entre ellos es muy fácil, pero hay que tener en cuenta que distintas fuentes pueden estar tan próximas que se detecten como un solo objeto. Las regiones más externas de dos galaxias, por ejemplo, pueden tener una señal lo suficientemente grande como para no ser consideradas como ruido, sin embargo, estas regiones son muy difusas y pueden confundirse las de dos galaxias distintas que estén muy próximas en la imagen (esto no significa que las galaxias estén físicamente próximas, recordemos que nuestras imágenes son una proyección en dos dimensiones de la bóveda celeste y aunque dos galaxias parezcan próximas en la imagen una puede estar miles de millones de años luz más lejos o cerca que la otra, lo mismo puede ocurrir con estrellas que estén delante del objeto como se menciona en la *Figura 6*).

La herramienta que distingue regiones separadas de otras dentro de una región señalada por *NoiseChisel* como fuente y no *background* o cielo es *Segment* y, análogamente al caso mencionado de distinguir entre dos objetos distintos, se puede usar esta herramienta para encontrar subregiones dentro del objeto, que en nuestro caso puede corresponder a los *clumps* de interés.

El *threshold* de detección para *NoiseChisel* es extremadamente pequeño, lo que da lugar a este problema de detección conjunta para diferentes fuentes. Para identificar subregiones, *Segment* clasifica como *máximo local* el píxel cuyo valor es mayor que el de sus 8 píxeles contiguos, y éstos se clasifican como *clump*. Dentro de cada objeto se analizan todos los *máximos locales* que contiene y se ordenan del más intenso (que se etiqueta con el número mayor) al menos intenso (con etiquetas cada vez más pequeñas). Si sus 8 píxeles contiguos no tienen etiqueta (son *clump* y no *máximo local*) toman la etiqueta de éste, creciendo así las subregiones a partir de los *máximos locales*. Cuando alrededor de un píxel se tienen píxeles contiguos con diferente etiqueta (han entrado en contacto dos regiones distintas), este píxel se identifica como un “río” que separa dos “montañas” que son las regiones que contienen píxeles cuyos 8 píxeles contiguos tienen su misma etiqueta. Una vez el objeto se ha dividido en estas “montañas” separadas por “píxeles río”, se analiza el *S/N ratio* de la región, y si la región corresponde a una zona especialmente intensa de la fuente, se superará el *threshold* para el *S/N ratio* para el que se identifique esta región como suficientemente intensa como para ser un *clump* del objeto en cuestión.

Así, el *output* de *Segment* consta de dos partes, la primera (“*OBJECTS*”) son los objetos detectados por *NoiseChisel* dentro de los cuales busca subregiones, además, en vez de asignar un valor =1 a todos los píxeles que corresponden a la señal, ahora asigna a todos los píxeles de cada objeto separado un valor particular (distinguiendo entre posibles objetos próximos diferentes), y una segunda parte (“*CLUMPS*”) que selecciona las subregiones dentro de cada objeto (*Figura 8*).

Nuestras *EELGs* de estudio son bastante claras y no presentan objetos externos que puedan confundirse con ellas, sin embargo, hay que tener cuidado y asegurarnos de que las subregiones que detecta *Segment* dentro de la galaxia detectada por *NoiseChisel*, corresponden realmente a *clumps* de la *EELG*.

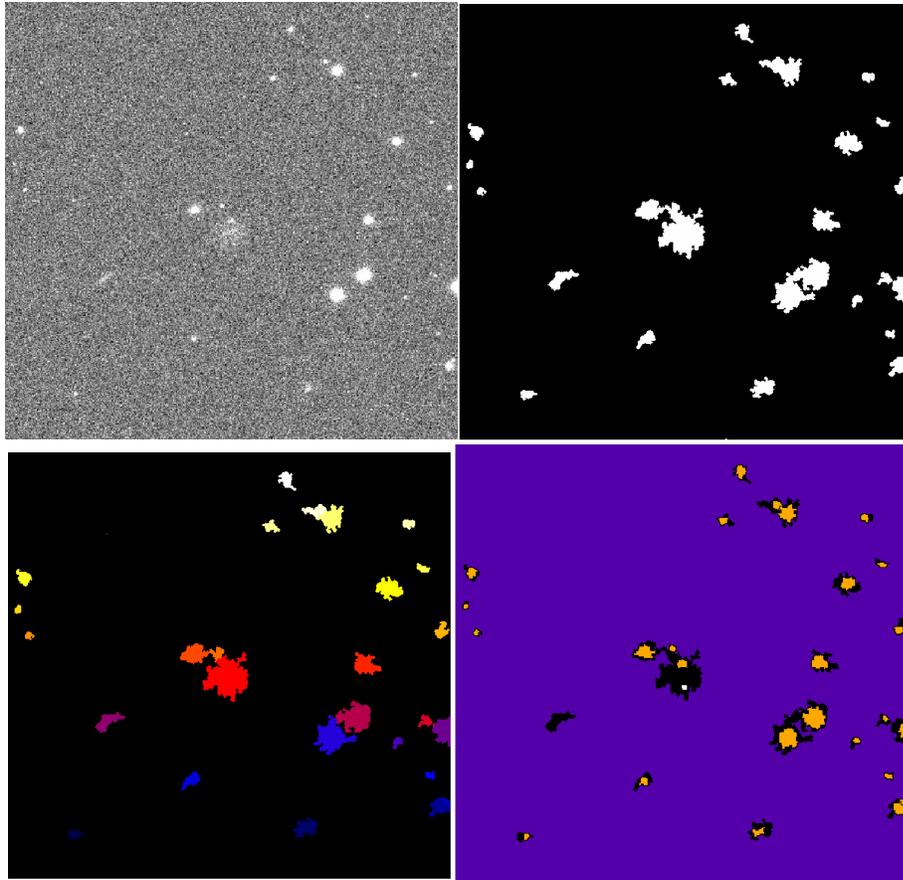


Figura 8: Una de nuestras imágenes que contiene una EELG (la EELG 64391017639, en el centro en todas las imágenes) observada con el programa DS9 en el filtro J0515 de J-PLUS (arriba a la izquierda), la detección de objetos observados tras aplicar NoiseChisel (arriba a la derecha), la identificación de los distintos objetos pese a que puedan estar en contacto dada por Segment (abajo a la izquierda, se cambia la escala de colores por claridad) y la detección subregiones detectadas por Segment que identifica nuestros clumps (abajo a la derecha).

En la *Figura 8* vemos en la imagen superior derecha (*output* de *NoiseChisel*) que *NoiseChisel* identifica las diferentes fuentes, sin embargo, el objeto central que debería contener solo nuestra galaxia está contaminado por dos objetos claramente externos. Al aplicar *Segment* (cuyo *output* son las dos imágenes inferiores) se caracteriza cada objeto con un número particular, como vemos en la imagen inferior izquierda realmente tenemos tres objetos distintos (distintos tonos de naranja y rojo). Además, *Segment* identifica las distintas subregiones dentro de los objetos en la imagen inferior derecha. El número de *clumps* dentro de cada objeto se caracteriza con un número en función del número de *clumps* que tenga dicho objeto, así vemos que en el objeto rojo central (nuestra *EELG*) se identifican dos *clumps*. En la escala de color seleccionada la primera subregión en un objeto se identifica con un = 1 que se ve naranja, en nuestra *EELG* se identifican dos *clumps* siendo el segundo caracterizado con un = 2 que en esta escala de color es blanco. Vemos también que no tienen porqué detectarse subregiones en todos los objetos, como ocurre con el

objeto morado en la zona central izquierda, y que los dos objetos que confundían nuestra *EELG* tienen un *clump* cada uno (más ejemplos de *EELG* y sus *clumps* y objetos detectados se adjuntan en el *Anexo 1*).

c. Catálogos (*MakeCatalog*²⁹)

Una vez identificados los distintos objetos y *clumps* en nuestra imagen debemos trabajar con esta información. Para ello, se usa la herramienta *MakeCatalog*, que produce catálogos en formato tabla con información sobre los *clumps* y objetos detectados. La información que ofrece el catálogo se puede elegir de entre los distintos *inputs* posibles, solicitando, por ejemplo, la suma de la señal de todos los píxeles que corresponden a un objeto o *clump* particular (está directamente relacionado con la magnitud del objeto).

Es importante recordar que hasta aquí tenemos un conjunto de píxeles con un valor particular que caracteriza las cuentas en el CCD para ese píxel, pero para obtener información física como la magnitud de las distintas fuentes, hay que transformar estas unidades en unidades útiles. En astrofísica es común trabajar en *Janskys*, una unidad de medida de la densidad de flujo espectral, para lo que es necesario introducir el *zeropoint* apropiado para cada filtro.

Seleccionamos en *MakeCatalog* los siguientes *inputs*:

--ids --x --y --clumps-x --clumps-y --area --magnitudo --magnitudo-error --half-sum-radius --semi-major

- **ids**: Devuelve el ID identificado por *Segment* para cada objeto y *clump* de la imagen.
- **x, y, clumps-x, clumps-y**: Devuelve el píxel de la posición central pesada por luminosidad de cada objeto (*x, y*) y *clump* (*clumps-x, clumps-y*).
- **area**: Devuelve el número de píxeles que conforman el área de todos los objetos y *clumps*.
- **magnitudo**: Devuelve la magnitud (en escala logarítmica por razones históricas) de cada objeto y *clump*.
- **magnitudo-error**: Devuelve el error en la magnitud.
- **half-sum-radius**: Devuelve el radio (en unidad de píxeles) del área que contiene la mitad de los píxeles que conforman el área del objeto. Es una medida del tamaño normalizado de cada objeto.
- **semi-major**: Devuelve el valor del semieje mayor que mejor se ajustaría a la morfología del objeto en caso de que éste fuera una elipse. Es una medida del tamaño normalizado de cada objeto.

d. *J-PLUS* vs *Legacy Survey*

Una vez introducidas las herramientas de análisis estamos en condiciones de hacer una comparación entre *J-PLUS* y *Legacy Survey* más rigurosa que la que hicimos en la inspección visual, aplicando *NoiseChisel* y *Segment* a las imágenes de ambos cartografiados en los filtros *g* y *r* (no se puede comparar directamente *J0515* porque *Legacy Survey* no tiene este filtro, pero como se menciona en la inspección visual asumimos que los resultados de la comparación son análogos en todos los filtros).

Durante este análisis se encontró un problema de alineamiento en los datos descargados de *J-PLUS* del que se informó a los encargados de los datos y fue solucionado por ellos.

Tras aplicar *NoiseChisel* y *Segment* a la misma galaxia y en el mismo filtro para ambos cartografiados se detectan unos u otros *clumps* en función de la imagen, en particular, debido a la mejor resolución espacial

de *Legacy Survey* esperamos detectar más *clumps*. En la *Figura 9* podemos ver una comparación para una de nuestras *EELG* de estudio, la *EELG* 65084029592:

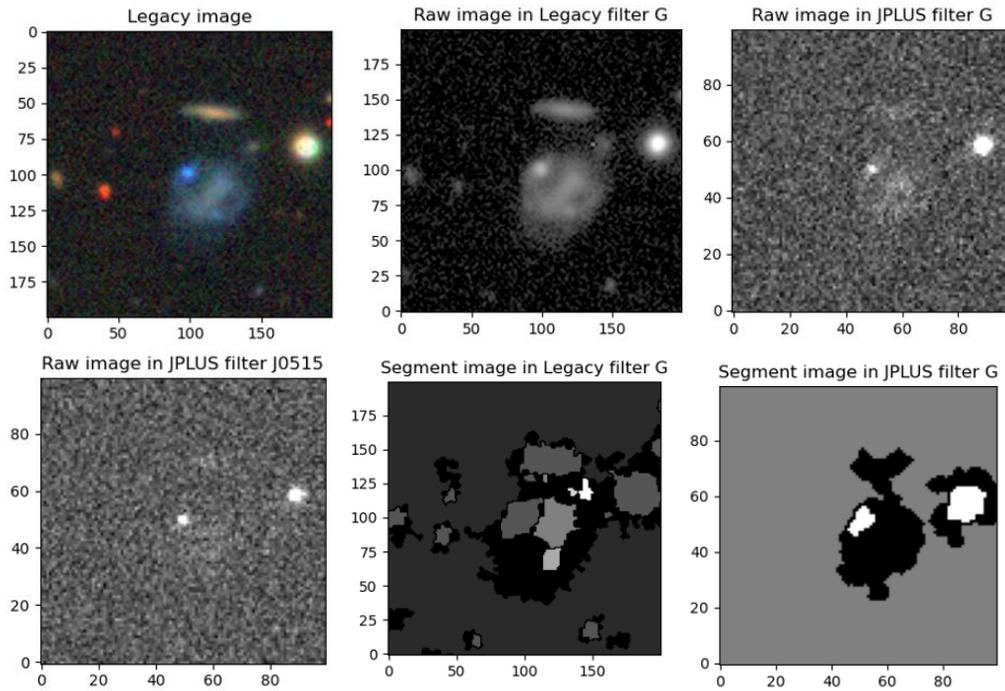


Figura 9: Comparación entre Legacy Survey y J-PLUS. Arriba a la izquierda la imagen de la EELG (centro de la imagen) en Legacy Survey formada por la composición de varios filtros, arriba en el centro la misma EELG en el filtro g de Legacy Survey (en escala logarítmica como en la Figura 7), arriba a la derecha la misma en J-PLUS, abajo a la izquierda la misma en el filtro de interés J0515, abajo en el centro los clumps detectados por Segment para la imagen de Legacy Survey y abajo a la derecha los clumps para J-PLUS.

Al estudiar la *Figura 9* vemos en la imagen superior izquierda un *clump* muy claro en nuestra *EELG*. Al comparar los *outputs* de *Segment* para ambos cartografiados en el filtro *g*, vemos que *J-PLUS* es muy preciso en la identificación del *clump* claramente visible (que en este caso particular es uno de nuestros *clumps* de interés, al estar presente en *J0515*), por otro lado, *Legacy Survey* identifica más *clumps* que en la inspección visual no son tan evidentes. Estas identificaciones de *Legacy Survey* no son tan precisas como en *J-PLUS* y uno de ellos (el grande central) es claramente el centro de la galaxia, y no un *clump*.

Vemos que *Legacy Survey* tiene una gran resolución espacial, pero a la hora de identificar *clumps* es más problemático, pues los detecta muy grandes y pervierte nuestros datos al considerar como *clumps* cosas que no lo son.

El otro *clump* pequeño que detecta *Legacy Survey* y que no se identifica para *J-PLUS* parece un ejemplo que puede servir como argumento en contra de nuestra conclusión de usar *J-PLUS* para identificar los *clumps*. Sin embargo, a pesar de que trabajar únicamente con las imágenes de *J-PLUS* puede suponer la pérdida de alguna región de interés en nuestras *EELGs* de estudio, *J-PLUS* es más que suficientemente preciso para detectar los *clumps* verdaderamente extremos que nos interesan, y la posible pérdida de algún *clump* no tan extremo como éste es aceptable teniendo en cuenta la naturaleza preliminar de este estudio.

El hecho de que de *Legacy Survey* pueda identificar como *clumps* regiones que no lo son y las dificultades que supondría manipular las imágenes para asegurarnos de que la imagen de *J-PLUS* y la imagen de *Legacy Survey* para esa *EELG* son realmente equivalentes (han sido tomadas en condiciones distintas: *seeing atmosférico*, telescopio, ...), nos permite concluir que el uso únicamente de datos de *J-PLUS* está justificado.

e. Muestra final

Al analizar las 125 *EELGs* de estudio se encuentran algunos problemas que deben ser solucionados: La idea es partir de una selección de 125 *EELGs*, sin embargo, es posible que en algún caso se haya identificado como *EELG* algo que en realidad no lo es, sino que se trata de una región de una galaxia espiral cercana de gran diseño que es identificada como una galaxia diferente más pequeña. En ese caso puede ocurrir que esta región sí que cumpla con nuestros criterios para ser considerada de formación estelar intensa, pero al ser significativamente más pequeña que la galaxia, no se trata del tipo de *clump* masivo que estamos buscando. Un ejemplo claro se aprecia en la *Figura 10*:

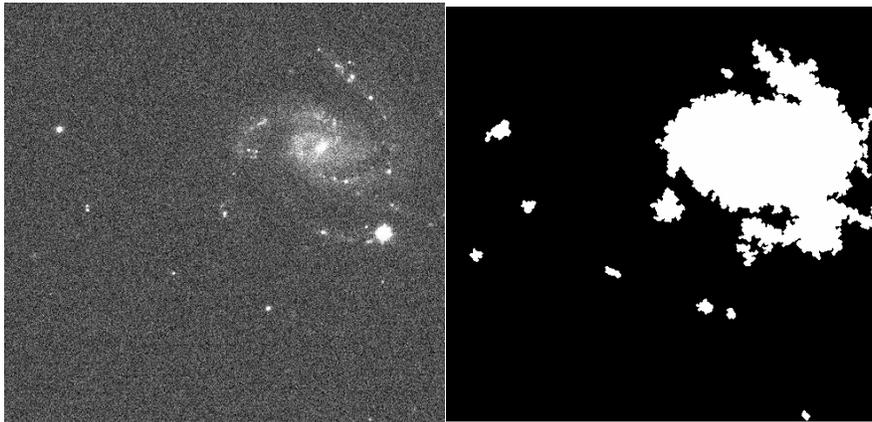


Figura 10: Imagen (a la izquierda) y output de NoiseChisel (derecha) para la EELG 73147013358.

Como vemos *NoiseChisel* está detectando como un objeto separado esa región que claramente pertenece al brazo de la galaxia espiral y 73147013358 no es realmente una *EELG*. Se identifican 5 casos como éste, dejando nuestra muestra en 120 *EELGs*.

En algunos casos la galaxia puede ser tan tenue en nuestras imágenes que *NoiseChisel* no detecte el objeto que corresponde a ella. En este caso no tendremos un objeto central y, por supuesto, tampoco ningún *clump* de interés. Se eliminan por este motivo 24 *EELGs*.

Además, hay 4 que presentan problemas en la propia imagen, una de ellas que no cumple con las condiciones para poder aplicarle *NoiseChisel* y 3 más que están muy cerca del borde de la imagen (*Figura 11*) y se eliminan para evitar posibles problemas.

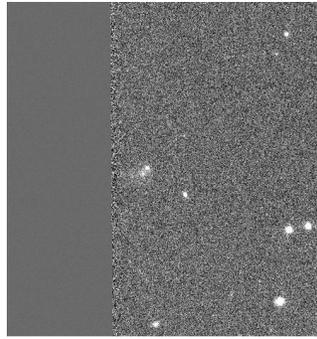


Figura 11: Imagen de la EELG 70076008955.

Finalmente, la muestra final es de 92 EELG, en las que se identifican 144 clumps.

Al analizar estas 92 EELGs con al menos un clump se encuentra una última complicación: en algunos casos *NoiseChisel* identifica como objetos diferentes dos regiones que son claramente el mismo objeto. Esto es un problema ya que en nuestro análisis queremos que toda la EELG sea identificada como un único objeto, para calcular la distancia del clump a su centro conocidos los outputs $-x$ y $-y$ de *MakeCatalog* del clump y del objeto, por ejemplo. Si la galaxia se separa en dos objetos distintos el resultado ofrecido por *MakeCatalog* no será el centro de la EELG, sino de la región que está considerando como objeto.

A la hora de identificar una región de la imagen como señal o ruido, *NoiseChisel* (y *Segment* para identificar las subregiones) hace uso de la *Point Spread Function (PSF)*, que da cuenta del tamaño que tendrá una fuente puntual en nuestra imagen. Al atravesar una apertura circular como el telescopio, la luz emitida por la fuente da lugar a un patrón de difracción circular, de modo que la señal de la fuente puntual se transmite a varios píxeles contiguos. Por defecto, la *PSF* de *NoiseChisel* y *Segment* tiene forma Gaussiana de 2x2 píxeles de tamaño, sin embargo, esto se puede personalizar con la función *kernel*. Eligiendo un *kernel* más grande (con una *PSF* de 5x5 píxeles, por ejemplo) se consigue que la distinción entre objetos se relaje, consiguiendo que no se separe un mismo objeto en dos o más distintos.

Un buen ejemplo puede verse en la *Figura 12*, donde se analiza la misma EELG y el output de *NoiseChisel* para dos *kernel* distintos.

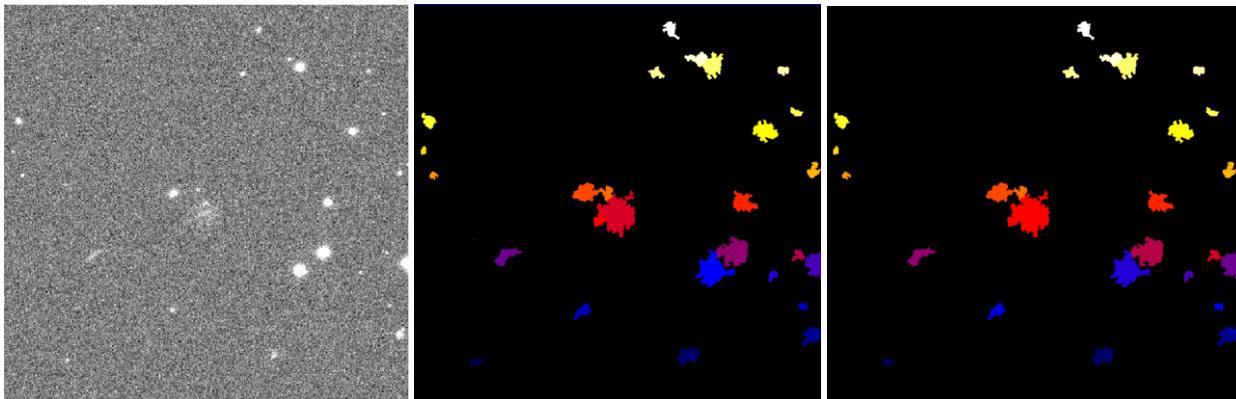


Figura 12: Imagen de la galaxia 64391017639 (a la izquierda) y output para un kernel por defecto (centro) y uno más grande (derecha).

Como vemos en la imagen de la izquierda, el objeto central es nuestra *EELG* de estudio. Al elegir un *kernel* menos estricto se consiguen unir los dos objetos distintos detectados en la imagen del centro (objeto rojo central y su esquina inferior izquierda naranja) en un solo objeto en la de la derecha (el rojo central). Esto es muy importante pues *MakeCatalog* ofrece información sobre la posición central del objeto. Si los dos objetos estuvieran separados la posición central del objeto en cuestión no correspondería a la posición central de la *EELG*, como sí ocurre si toda la *EELG* es identificada con un solo objeto. Debemos ser cuidadosos a la hora de usar el *kernel* más apropiado para cada caso.

Se ejecuta *CIGALE* una primera vez para estos 144 *clumps* de 92 *EELGs*. En este paso (como se verá más adelante) se identifican dos galaxias de gran diseño más, quedando la muestra final definitiva en 90 *EELGs* con 140 *clumps*.

5. CIGALE

Como vimos, el proceso para obtener modelos teóricos de las *SEDs* de galaxias es complejo y depende de muchos parámetros: distintos modelos dan lugar a *SEDs* completamente diferentes. De esto se encarga *CIGALE*, que genera distintos modelos dentro de unos parámetros razonables que le introducimos como *input* y compara las *SEDs* que obtiene con nuestros resultados fotométricos, calculando cuál es el mejor modelo para nuestros datos. De este modo obtenemos estimaciones de diversos parámetros físicos de los *clumps* que estudiamos (masas, edades de las poblaciones estelares, tasas de formación estelar, ...).

Para generar nuestro archivo de *input* y de configuración de *CIGALE* es necesario introducir la información sobre nuestros filtros en su base de datos. Una vez almacenada esa información introducimos nuestros resultados fotométricos obtenidos con *MakeCatalog* para los *clumps* detectados por *Segment*:

```
"id" "redshift" "line.H-alpha" "line.H-alpha_err" "line.OIII-500.7" "line.OIII-500.7_err" "line.OII-372.7" "line.OII-372.7_err" "line.H-beta" "line.H-beta_err" "line.OIII-495.9" "line.OIII-495.9_err" "line.NII-658.4" "line.NII-658.4_err" "line.Ly-alpha" "line.Ly-alpha_err" "rSDSS_jplus" "gSDSS_jplus" "iSDSS_jplus" "zSDSS_jplus" "uJAVA_jplus" "J0378_jplus" "J0395_jplus" "J0410_jplus" "J0430_jplus" "J0515_jplus" "J0660_jplus" "J0861_jplus" "rSDSS_jplus_err" "gSDSS_jplus_err" "iSDSS_jplus_err" "zSDSS_jplus_err" "uJAVA_jplus_err" "J0378_jplus_err" "J0395_jplus_err" "J0410_jplus_err" "J0430_jplus_err" "J0515_jplus_err" "J0660_jplus_err" "J0861_jplus_err"
```

Debemos introducir la información fotométrica en términos del flujo (y su error) en unidades de *mili Janskys (mJy)*. En *MakeCatalog* obtenemos medidas de la magnitud de los *clumps*, de modo que debemos convertirlas a unidades de flujo. La relación entre flujo y magnitud (*m*) en el sistema *AB* viene dada por:

$$m = -2.5 * \log_{10} F + 8.9 \rightarrow F (mJy) = 10^{\frac{16.4-m}{2.5}}$$

Así, con nuestras medidas de magnitud podemos hallar el flujo y su error sin mayores problemas.

El modelo de *CIGALE* se hace asumiendo dos tipos de poblaciones estelares, una vieja (de entre 100 y 5000 *Myrs*) y una joven que caracteriza la formación estelar. Como vimos al introducir la *SED*, cada una corresponde a una *SSP* particular y al variar la relación entre ellas y variando los parámetros de cada una se obtienen *SEDs* diferentes.

Una vez introducidos en el programa los resultados fotométricos obtenidos para cada filtro para los 144 *clumps* detectados por *J-PLUS* en el filtro *J0515* para nuestra muestra de 92 *EELGs*, probamos modelos entre distintos parámetros (lista completa en *Anexo 2*). Se eligen estos parámetros para que sean lo más parecidos a los elegidos en *Lumbreras-Calle et al. 2021⁶*, con el objetivo de que la comparación entre nuestros resultados para los *clumps* y los suyos para las *EELGs* sea lo más correcta posible. Utilizamos como *redshift* de cada *clump* el que fue derivado para cada galaxia en *Lumbreras-Calle et al. 2021⁶*.

Se prueban 2352000 modelos para cada uno de los *clumps*, obteniendo el modelo más similar a nuestros resultados fotométricos. Los resultados obtenidos para los parámetros más relevantes de los ajustes para cada *clump* se adjuntan en el *Anexo 3*.

6. Análisis de los resultados

En nuestro estudio se han analizado los *clumps*, pero para comparar con la galaxia que lo contiene hacemos uso de los resultados de *Lumbreras-Calle et al. 2021⁶*, donde se identificaron las *EELGs* de nuestra muestra y que fueron analizadas con *CIGALE*, obteniendo resultados análogos a los nuestros para las *EELGs*.

Debido a limitaciones técnicas (de incompatibilidad de versiones de *CIGALE*) entre ambos estudios no se utilizan exactamente los mismos parámetros para generar los modelos de *CIGALE*. Por ejemplo, en nuestro caso solo usamos $E_BVs_old_factor = 0.44$, mientras que en *Lumbreras-Calle et al. 2021⁶* se usa $E_BVs_old_factor = 0.44, 1$ (estos cambios afectan a parámetros que no tenemos en cuenta en el siguiente análisis, de modo que su efecto es menor). Estas diferencias entre los posibles modelos pueden dar lugar a diferencias en los resultados obtenidos para alguno de nuestros *clumps* particulares, sin embargo, la muestra general presenta en muchos casos tendencias claramente diferenciables a partir de las cuales podemos obtener información sobre ellos (todos los datos obtenidos en este *TFM* sobre de los *clumps* utilizados para la realización de las gráficas se adjuntan en el *Anexo 4*, además, se hace uso para generar las gráficas de dato obtenidos al ejecutar *CIGALE* que se adjuntan en el *Anexo 3*).

Pese a que es posible obtener información interesante con la observación directa de las gráficas, se añadirá un análisis estadístico para buscar una correlación entre los parámetros. En particular, usaremos el *Spearman's Rank Correlation Coefficient*³⁰, donde una correlación perfecta se caracteriza con un $+1$ o -1 e indica si las variables están relacionadas por una función monótona (que no tiene porque ser lineal). Por ejemplo, si todos los puntos con mayor valor de x que un cierto valor tienen valores de y también mayores, habrá una correlación de $+1$. Nuestro *output* para cada gráfica serán dos valores, el de esta *correlación de Spearman* y un *p-value* que indica la probabilidad de que no haya una correlación, cuanto mayor el valor del *coeficiente de Spearman* y menor el *p-value* mayor la confianza en que exista una correlación entre los valores de los dos parámetros comparados. Para valores puramente aleatorios ocurrirá lo contrario y el coeficiente de Spearman tenderá a $= 0$ y el *p-value* será mayor. El valor máximo de *p-value* que hemos considerado como válido para aceptar la existencia de correlación es de 0.05.

Una parte muy importante será la de estudiar algunas de las propiedades de los *clumps* en función de la distancia al centro de su galaxia. Una aplicación muy interesante de un estudio como éste es, por ejemplo, estudiar si hay una relación entre la edad del *clump* y su distancia al centro. Si se encuentra que los *clumps*

más viejos están más cerca del centro de su *EELG* madre, esto puede indicar una evolución desde las zonas más externas a las internas.

Para ello es importante calcular las distancias con precisión. En este punto conocemos la distancia entre el *clump* y el centro de la *EELG* que lo contiene en unidades de píxeles, pero es necesario trabajar en unidades físicas. Para calcular la distancia entre el *clump* y el centro de la *EELG* se hace uso de la distancia diametral angular (D_A), que relaciona el tamaño físico del objeto con su tamaño angular visto desde la Tierra con:

$$D_A = \frac{x}{\vartheta}$$

En nuestro caso ϑ es la distancia angular entre el *clump* y el centro de la *EELG*, y x será la distancia física que buscamos. En *J-PLUS* un píxel corresponde a 0.55 segundos de arco, de modo que podemos hallar ϑ en unidades angulares con $0.55 * \text{distancia en píxeles}$. En cuanto a D_A , se puede hallar haciendo uso del *redshift* asumiendo una cosmología con:

$$D_A = \frac{S_k(r)}{1+z}$$

Aquí, $S_k(r)$ depende de la cosmología asumida, en nuestro caso *Flat Λ CDM* con $\Omega_{m0} = 0.27$.

Por otro lado, la morfología y tamaño de las distintas *EELG* no es regular, de modo que es necesario normalizar su morfología a alguno de nuestros parámetros de normalización (***half-sum-radius*** o ***semi-major***) con el fin de corregir el hecho de que para *EELGs* más grandes las distancias al centro serán mayores que para las más pequeñas. Una vez normalizada, tendremos una distancia entre el centro del *clump* y el centro del objeto (que caracteriza la *EELG*) comparable entre galaxias con distintos tamaños.

Además, calcularemos la luminosidad de alguna línea de emisión para nuestros *clumps* (en particular de la de *[OIII]5007* que fue la utilizada para detectar las galaxias, y correlaciona con la intensidad de su formación estelar). Uno de los *outputs* de *CIGALE* es el flujo de las líneas introducidas en el análisis para el modelo que mejor se ajusta a las observaciones. Sin embargo, para medir la intensidad intrínseca debemos conocer la luminosidad de la línea. Para ello, hacemos uso de la distancia de luminosidad (D_L), que relaciona el flujo con la luminosidad, teniendo en cuenta que para dos fuentes igual de luminosas mediremos más flujo para la que sea más cercana, y que es también un *output* de *CIGALE*. La relación viene dada por:

$$F = \frac{L}{4 * \pi * D_L^2}$$

Conocidos el flujo de la línea (F) y la distancia de luminosidad de la fuente (que es también un *output* de *CIGALE*, conocido el *redshift*) se puede hallar la luminosidad.

a. Masa del Clump – Masa de la galaxia que lo alberga

Un ejemplo para el que se evidencia esta diferencia entre los datos de nuestros *clumps* y los de las galaxias obtenidos por *Lumbreras-Calle et al. 2021*⁶ se aprecia claramente en la siguiente gráfica.

Resulta evidente que la masa del *clump* debe ser menor que la de la galaxia que lo contiene. Al ver la *Figura 13* vemos que en algunos casos los resultados no tienen sentido físico, pues algunos de nuestros *clumps*

son más masivos que su galaxia. Un pequeño exceso de la masa del *clump* con respecto a la de la galaxia puede ser razonable para casos en que el *clump* es muy grande y ocupa prácticamente toda la *EELG*, las diferencias entre los modelos pueden ser la causa de que sobreestimemos la masa del *clump* (o que *Lumbreras-Calle et al. 2021*⁶ infraestime la de la galaxia). Vemos, sin embargo, algunos casos para los que el exceso es muy grande, en particular para la *EELG* 75211009334 se tiene que $M_{clump} = 756.07 M_{galaxia}$ que es claramente erróneo. Analizando una a una se identifican esta 75211009334 y 75875020483 como galaxias espirales de gran diseño y se eliminan del análisis. Así, la muestra final es de 90 *EELGs* y 140 *clumps*:

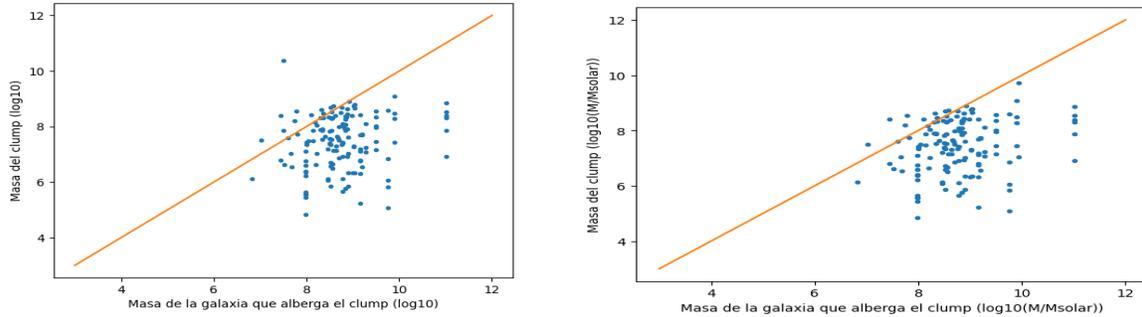


Figura 13: Masa del clump en función de la de la galaxia que lo alberga en escala logarítmica antes (izquierda) y después (derecha) de la eliminación de objetos problemáticos. La recta indica la relación 1:1.

Pese a que no tienen sentido físico mantenemos en la muestra los *clumps* ligeramente por encima de la recta 1:1, pues pueden ayudarnos a encontrar tendencias entre otros parámetros. El análisis estadístico con la correlación de Spearman corrobora la existencia de esa tendencia:

SpearmanrResult(correlation=0.199, pvalue=0.017)

b. Masa del Clump – Distancia al centro de la galaxia

La morfología de la galaxia se normaliza con los parámetros *half-sum-max* y *semi-major* de *MakeCatalog* con los que estimamos el ratio de esta distancia. Presentamos la gráfica para ambas normalizaciones:

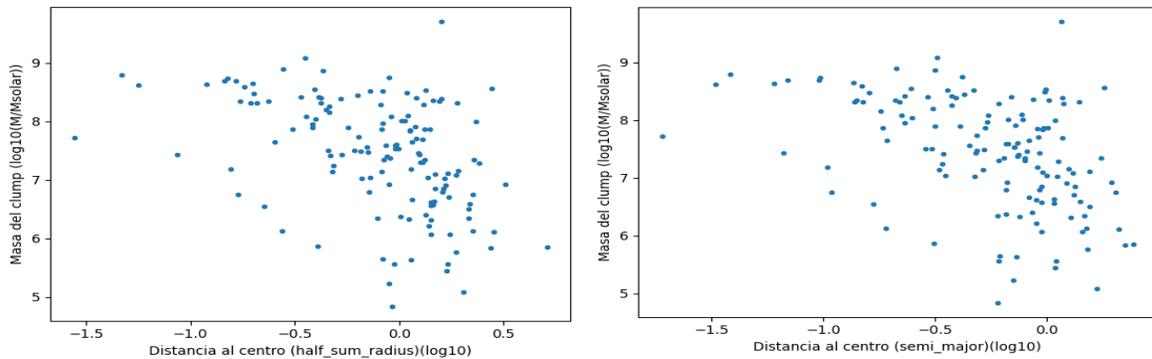


Figura 14: Masa del clump en escala logarítmica en función de la distancia normalizada con *half-sum-radius* y *semi-mayor* al centro de la galaxia que lo alberga en escala logarítmica.

Al observar la *Figura 14* vemos de nuevo el problema de la disparidad entre nuestros modelos y los de *Lumbreras-Calle et al. 2021*⁶. La normalización con respecto al radio de la galaxia se hace con `--half-sum-radius` y `--semi-major`, en que se asume una morfología regular para la *EELG*. Por su propia definición (`--half-sum-radius` se define como el radio de *la mitad* de los píxeles que conforman la galaxia), esperamos que algunos de los *clumps* se sitúen fuera del radio de la distancia que definen estos parámetros de normalización (quedan a la derecha de 0.0). Además, el hecho de que las galaxias sean irregulares puede suponer que el *clump* se localice muy lejos del radio de la galaxia calculado, si el *clump* es muy externo y la *EELG* muy irregular, por ejemplo. En algunos casos, las diferencias observadas son muy grandes, obteniendo *clumps* que están hasta a una distancia al centro de 5.09 veces el tamaño de la galaxia para la normalización en `--half-sum-radius` (este $5.09 \approx 0.706$ en escala logarítmica).

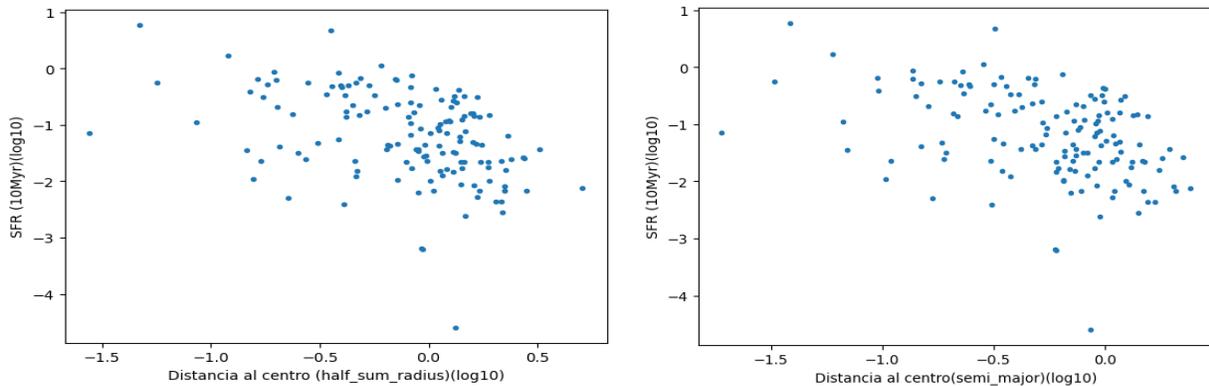
Pese a contar con algunos valores extremos, vemos que existe una tendencia clara: las galaxias más cercanas al centro son más masivas. El análisis estadístico con la correlación de Spearman corrobora la existencia de esa tendencia para ambas normalizaciones, cuanto más lejano el *clump* menos masivo. Este resultado está de acuerdo con los resultados de *Hinojosa-Goñi et al. 2018*³¹, de modo que esta tendencia de mayor masa cuanto más central el *clump* se cumple también para *EELGs* a bajo *redshift*.

SpearmanResult(correlation=-0.495, pvalue=3.852e-10) para --half-sum-radius.

SpearmanResult(correlation=-0.484, pvalue=1.034e-9) para --semi-major.

c. Star Formation Rate – Distancia al centro de la galaxia

Análogamente al caso anterior, se encuentra una correlación entre la tasa de formación estelar y la distancia del *clump* al centro de la galaxia. La tasa de formación de estrellas (*SFR*, por su nombre en inglés) es una medida de formación estelar, en particular cuántas masas solares de estrellas nuevas se forman en un año. Es uno de los valores de salida de *CIGALE* considerando los modelos de población estelar que mejor se ajustan y tiene 3 opciones: una instantánea, una que calcula la media en los últimos 100Myrs y otra sobre los últimos 10Myrs . Se elige esta última por ser consistente con el tiempo típico que tardan en desaparecer las líneas de emisión tras un brote de formación estelar instantáneo, siendo el valor más comparable con uno de los indicadores de *SFR* más utilizados: la luminosidad en la línea H_{α} .



*Figura 15: SFR del clump en función de la distancia normalizada con **half-sum-radius** y **semi-major** al centro de la galaxia, ambas en escala logarítmica.*

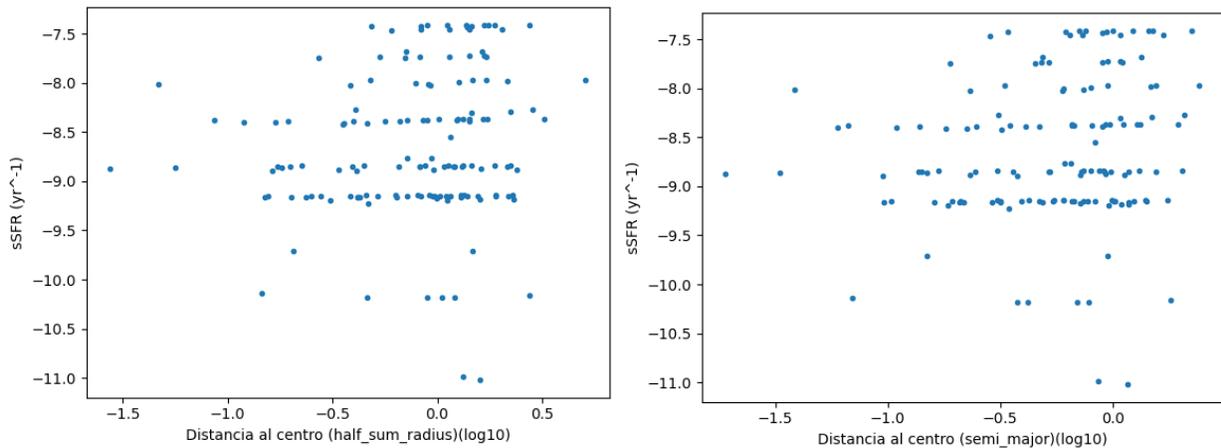
SpearmanrResult(correlation=-0.454, pvalue=1.409e-8) para -half-sum-radius.

SpearmanrResult(correlation=-0.429, pvalue=1.038e-7)) para -semi-major.

De nuevo, el análisis estadístico confirma la tendencia apreciable al observar la *Figura 15*, cuanto más lejano el *clump* menor el *SFR*. No obstante, esta correlación puede ser debida a la correlación conocida entre masa estelar y tasa de formación estelar, conocida como “*Main Sequence of Star Formation*”³². Este resultado está de acuerdo con los resultados de *Hinojosa-Goñi et al. 2018*³¹ de modo que esta tendencia de mayor *SFR* cuanto más central el *clump* se cumple también para *EELGs* a bajo *redshift*.

d. SFR específica (sSFR) – Distancia al centro de la galaxia

La *sSFR* se define como SFR/M_{clump} y da una idea de cómo de intensa es la formación estelar en el *clump*, pero teniendo en cuenta la masa de éste (es un valor intrínseco, que no depende de que el *clump* sea grande o pequeño, sino de si tiene formación estelar intensa o no).



Se observa que a mayor distancia del centro mayor es la *sSFR*. El análisis estadístico lo corrobora:

SpearmanrResult(correlation=0.207, pvalue=0.0137)

SpearmanrResult(correlation=0.220, pvalue=0.0085)

e. Tamaño del clump – Distancia al centro de la galaxia

Otro resultado interesante se obtiene al estudiar estos parámetros. Conocido el tamaño físico que corresponde a un píxel podemos hallar el tamaño del *clump* fácilmente multiplicando por el número de píxeles que lo componen (que nos da *-area* en *MakeCatalog*):

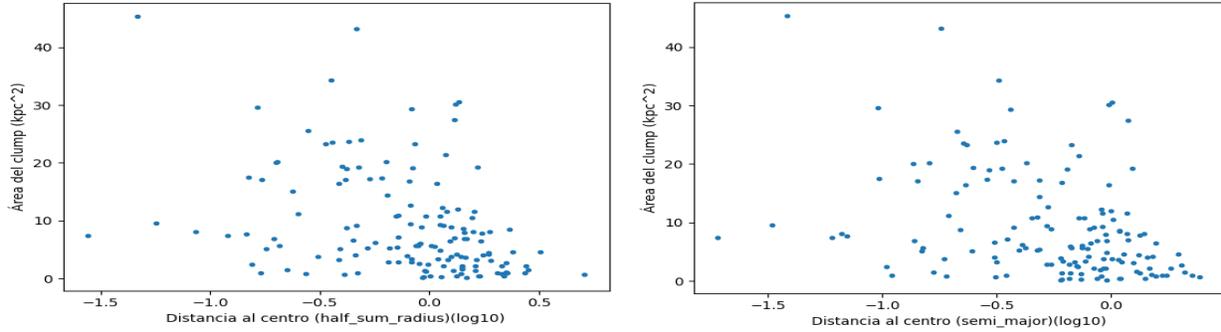


Figura 17: Área del clump en función de la distancia normalizada con **half-sum-radius** y **semi-major** al centro de la galaxia que lo alberga en escala logarítmica.

Vemos de nuevo una clara correlación entre el tamaño del *clump* y su distancia al centro. Cuanto más lejano es el *clump* menor su tamaño. Esta correlación es consistente con la “*Mass-Size Relation*”, donde galaxias con mayor masa estelar tienen tamaños más grandes. De nuevo, el análisis estadístico con la correlación de Spearman corrobora la existencia de esa tendencia para ambas normalizaciones, cuanto más lejano el *clump* más pequeño.

SpearmanResult(correlation=-0.403, pvalue=6.798e-7) para **-half-sum-radius**.

SpearmanResult(correlation=-0.368, pvalue=6.453e-6) para **-semi-major**.

f. Masa de la Población Joven del Clump – Masa de la Población Joven de la galaxia

Al comparar la masa de la población joven del *clump* con respecto a la de la *EELG* esperamos que la del *clump* sea menor, al ser la del *clump* una fracción de la total de la galaxia.

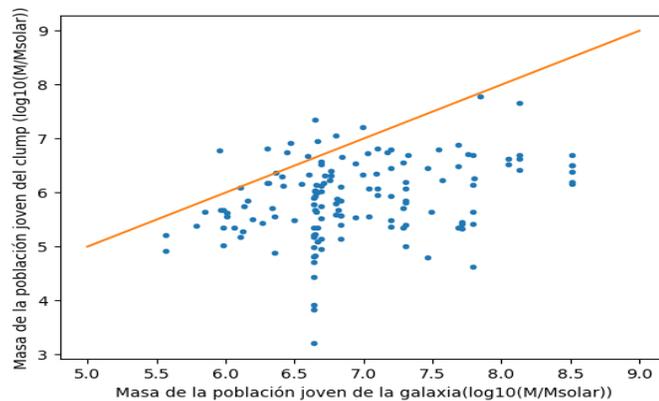


Figura 18: Masa de la población joven del clump en función de la de la galaxia, ambas en escala logarítmica. La recta naranja es la relación 1:1.

Vemos que en general se cumple el resultado esperado. Por otro lado, es interesante notar que para masas de la galaxia más bajas (de hasta 10^7 masas solares) la tendencia es a lo largo de la recta,

indicando que casi toda la población joven tiende a estar en el *clump*. Sin embargo, para masas de la galaxia más grandes la masa de la población abandona la recta y se aplana. Es decir, para *EELGs* más masivas hay mayor separación entre la masa de la población joven del *clump* y la total de la galaxia: puede ser debido a que haya múltiples *clumps* o que haya formación estelar fuera del *clump*.

SpearmanResult(correlation=0.343, pvalue=2.895e-05)

El análisis estadístico corrobora la correlación entre ambos parámetros.

g. Edad del Clump – Distancia al centro de la galaxia

Uno de los resultados interesantes que mencionamos como ejemplo de la utilidad de este tipo de estudio era el de buscar una correlación entre la edad de los *clumps* y la distancia al centro de la *EELG*, con el fin de estudiar si hubiera evidencia de dinámica de los *clumps* de fuera hacia dentro de la *EELG* (o viceversa).

Las posibles edades que probamos en nuestros modelos eran de 1, 2, 3, 4, 5, 7 o 12 *Myrs*. Tras estudiar el mejor modelo que ofrece *CIGALE* obtenemos:

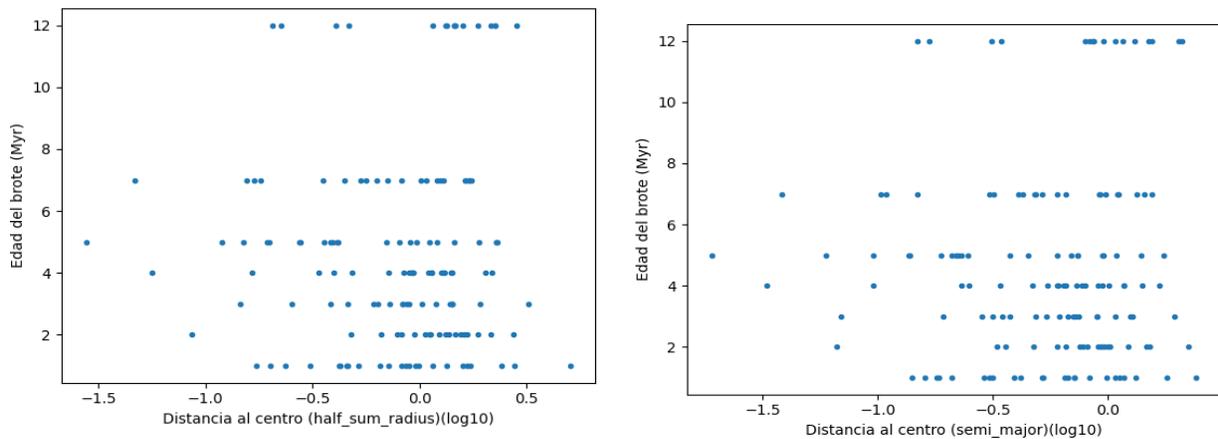


Figura 19: Edad del clump en función de la distancia normalizada con **half-sum-radius** y **semi-mayor** al centro de la galaxia que lo alberga en escala logarítmica.

Visualmente no se aprecia ninguna correlación clara entre estos parámetros. Además:

SpearmanResult(correlation=-0.003, pvalue=0.968) para **-half-sum-radius**.

SpearmanResult(correlation=0.001, pvalue=0.986) para **-semi-mayor**.

El estudio estadístico confirma que no se encuentra ninguna correlación de este tipo. Medir la edad del brote solo con mediciones fotométricas es complicado y puede ser la razón por la que no se aprecia ninguna correlación (o simplemente puede que no exista dicha correlación).

h. Luminosidad de las líneas - Distancia al centro de la galaxia

Por último, veamos si hay una relación entre la luminosidad de la línea de $[OIII]5007$ y la de H_α y la distancia al centro. Recordemos que estas líneas correlacionan la intensidad de la formación estelar en el *clump*.

Haciendo uso de la relación entre flujo, distancia de luminosidad (ambos nos los da *CIGALE* para el mejor modelo) y luminosidad obtenemos:

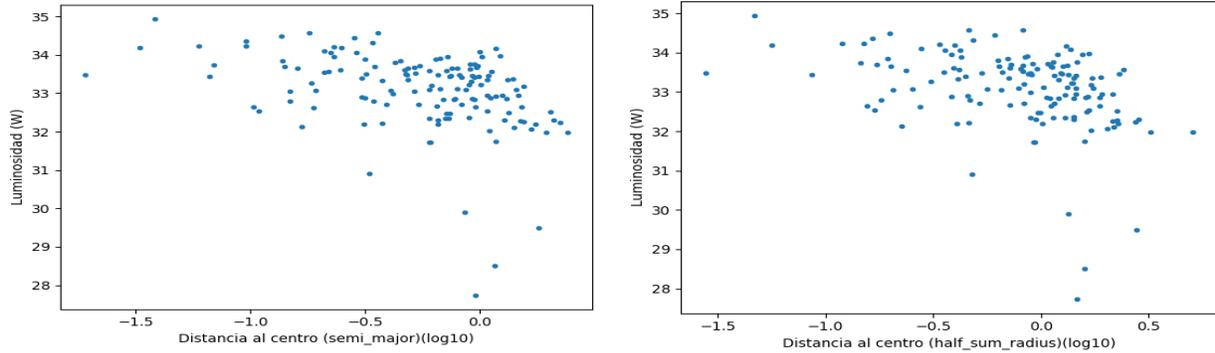


Figura 20: Luminosidad de la línea de $[OIII]5007$ en función de la distancia normalizada con **half-sum-radius** y **semi-mayor** al centro de la galaxia que lo alberga, ambas en escala logarítmica.

De nuevo se aprecia una correlación: cuando más cercano al centro el *clump* más intensa su línea de emisión de $[OIII]5007$, lo que evidencia mayor intensidad en la formación estelar. Análogamente para H_α :

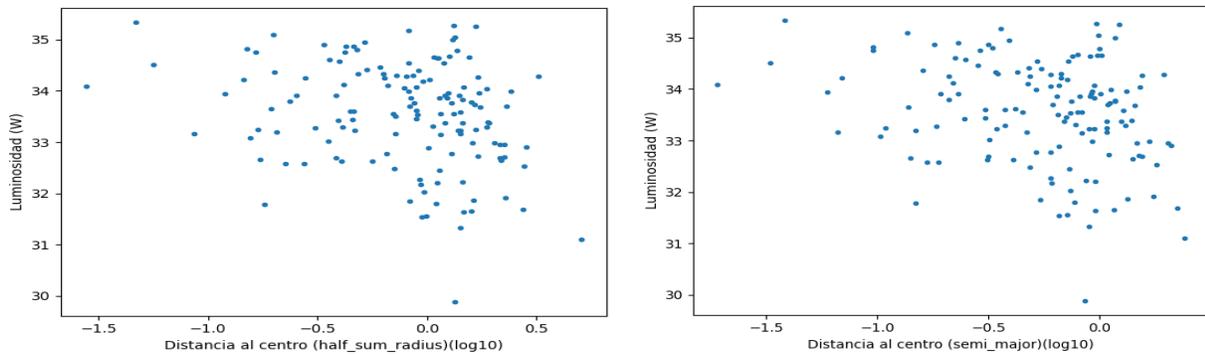


Figura 21: Luminosidad de la línea de H_α en función de la distancia normalizada con **half-sum-radius** y **semi-mayor** al centro de la galaxia que lo alberga, ambas en escala logarítmica.

El análisis estadístico para $[OIII]5007$ y H_α :

SpearmanrResult(correlation=-0.435, pvalue=6.265e-8) para **-half-sum-radius**.

SpearmanrResult(correlation=-0.412, pvalue=3.561e-7) para **-semi-mayor**.

SpearmanrResult(correlation=-0.257, pvalue=0.002) para **-half-sum-radius**.

SpearmanrResult(correlation=-0.239, pvalue=0.004) para **-semi-mayor**.

La correlación con H_α es peor que la de $[OIII]5007$, lo que puede indicar cambios en la metalicidad o potencial de ionización de las regiones (pero que haría falta análisis espectroscópico para saberlo con seguridad). Además, se ve que la intensidad media de H_α que es $= 1.95 * 10^{34} W$ es mayor que la de $[OIII]5007$ que es $= 4.84 * 10^{33} W$.

i. Resultados generales

Algunos resultados generales:

- De las 90 *EELGs* finales 14 no presentan *clumps*.
- La media de *clumps* por *EELG* que contiene al menos 1 *clump* es de 1.84.
- La *EELG* con más *clumps* es 72851011490 con 11 *clumps*.
- La distancia media de los *clumps* al centro normalizada con **–half-sum-radius** es de 1.054 * *Tamaño normalizado de la galaxia*.
- La distancia media de los *clumps* al centro normalizada con **–semi-major** es de 0.734 * *Tamaño normalizado de la galaxia*.
- La edad media del brote para los *clumps* es de 4.59 *Myrs*.
- La masa media de los *clumps* es de $1.54 \cdot 10^8 M/M_{solar}$.
- La masa media de las *EELGs* es de $1.75 \cdot 10^9 M/M_{solar}$.
- El *Star Formation Rate* medio de los *clumps* es de 2.47 *Myrs*.
- La masa media de la población joven de los *clumps* es de $2.71 \cdot 10^6 M/M_{solar}$.
- La masa media de la población joven de las *EELGs* es de $1.62 \cdot 10^7 M/M_{solar}$.
- El área medio de los *clumps* es de 9.075 *kpc*².
- El valor medio de χ^2 para cada ajuste de *CIGALE* es de 6.334 y para χ^2 reducido de 0.425.

7. Conclusiones

Las características similares entre las *EELGs* a bajo *redshift* y las galaxias más lejanas observadas en el universo convierten a estas *EELGs* en objetivos ideales para el estudio del Universo primitivo. En este trabajo hemos analizado 125 *EELGs* a bajo *redshift* identificadas en *Lumbreras-Calle et al. 2021*⁶. Haciendo uso de las herramientas *NoiseChisel* y *Segment* de *GNUAstro* para la detección automatizada de *clumps*, concluimos que las imágenes del cartografiado *J-PLUS* son suficientemente precisas como para identificar los *clumps* de formación estelar extrema más intensos de estas *EELGs*.

Tras eliminar 35 *EELGs* problemáticas de nuestra muestra, identificamos 140 *clumps* en 90 *EELGs*. Estos 140 *clumps* se analizan con el programa de ajuste de *SED*: *CIGALE*, y los resultados obtenidos se comparan con los resultados análogos obtenidos por *Lumbreras-Calle et al. 2021*⁶ para las galaxias que los contienen. El objetivo del trabajo es el de realizar un estudio preliminar de los resultados obtenidos, en particular se estudian posibles correlaciones entre parámetros que permitan obtener información sobre los *clumps*.

En nuestro análisis se identifican con gran confianza las siguientes correlaciones (o ausencia de ellas):

- Cuanto más lejano el *clump* respecto al centro de la galaxia que lo contiene, menos masivo, menor su *SFR*, su *sSFR* y su área.

- Cuanto más masiva la galaxia menor la contribución del *clump* a la masa de su población joven.
- No se encuentra una correlación entre edad del *clump* y distancia respecto al centro de la galaxia que lo contiene.
- Cuanto más lejano el *clump* respecto al centro de la galaxia que lo contiene, menor la intensidad de su línea de $[OIII]5007$ y de H_{α} .

El hecho de que se identifiquen algunas de ellas que eran esperadas y previamente conocidas permite validar el método de estudio de los *clumps*, en particular el uso de las herramientas *NoiseChisel* y *Segment* para la identificación automatizada de *clumps*.

Se han identificado bastantes correlaciones entre la posición de los *clumps* en la *EELG* y sus propiedades, mostrando *clumps* más masivos y formando más estrellas cuando nos acercamos el centro de las galaxias. Sin embargo, debido a la no evidencia de una correlación entre la edad de los *clumps* y su distancia al centro, este estudio no corrobora la idea (propuesta por *Dekel et al. 2009*³⁴ y apoyada por ejemplo en *Hinojosa-Goñi et al. 2018*³¹) de que los *clumps* se producen en las regiones exteriores y evolucionan hacia el centro, atrayendo más gas y siendo más intensos los brotes más centrales. Esto no descarta radicalmente la hipótesis pues las edades de los brotes que obtenemos se refieren a poblaciones estelares muy jóvenes, y es posible que la evolución a largo plazo, previa a dichos brotes, sea como se describe en esos otros estudios.

También se pueden comparar estos resultados con los obtenidos en *Guo et al. 2011*³³, aunque hay que ser cuidadoso al hacerlo, pues pueden aparecer diferencias debidas a que nosotros seleccionamos *clumps* con formación estelar reciente, pero en su caso se seleccionan otros que pueden haber tenido formación estelar hace decenas o centenas de millones de años, pues emiten en el ultravioleta.

En conclusión, en este trabajo se demuestra que el método utilizado para el estudio de *clumps* de formación estelar intensa en *EELGs* a bajo *redshift* funciona correctamente, obteniendo resultados esperados de acuerdo con otros trabajos similares (extendiendo el estudio a más bajo *redshift* y masa). La conclusión de que es correcto utilizar únicamente las imágenes obtenidas por *J-PLUS* para el estudio de este tipo de *clumps* es correcta y se valida el funcionamiento de las utilidades *NoiseChisel* y *Segment* recientemente añadidas al paquete *GNUAstro* para la manipulación y análisis de datos astronómicos.

Un análisis más detallado futuro podría analizar más variables de los *clumps* y galaxias (usando los mismos parámetros de *CEGALÉ* para ambos) en otras longitudes de onda (ultravioleta para estudiar con más detalle las poblaciones jóvenes, infrarrojo para las poblaciones antiguas). Además, es posible extender el trabajo a más fuentes (con la *Data Release 3* de *JPLUS* o el futuro *J-PAS*, por ejemplo) para detectar *clumps* y galaxias más débiles. Este puede ser el primer estudio que abra la puerta a otros proyectos sucesivos, que contribuyan a explicar las propiedades de los *clumps* de formación estelar intensa, y si los que detectamos en el universo local nos pueden ayudar a entender la formación de las primeras generaciones de galaxias.

8. Referencias

1. Charles Messier, *Catalogue des Nébuleuses & des amas d'Étoiles. Connaissance des Temps, 1781.*

2. *A Treatise on Astronomy (3rd ed.)*. Philadelphia: Carey, Leah and Blanchard, 1835.
3. Trimble, V. *The 1920 Shapley-Curtis Discussion: Background, Issues, and Aftermath*, 1995.
4. E Hubble, *A relation between distance and radial velocity among extra-galactic nebulae*, 1929.
5. Takatoshi Shibuya et al., *Morphologies of ~190,000 Galaxies at $z = 0-10$ Revealed with HST Legacy Data. III. Continuum Profile and Size Evolution of Ly α Emitters*, 2019.
6. Lumberras-Calle et al., *J-PLUS: Uncovering a large population of extreme [OIII] emitters in the local Universe*, 2021.
7. Cenarro, A. J. et al., *JPLUS: The Javalambre Photometric Local Universe Survey*, 2019.
8. Cenarro, A. J. et al., *First light instrumentation for the Observatorio Astrofísico de Javalambre (OAJ)*, 2014.
9. Marín-Franch, A., Taylor, K., Cenarro, J., Cristobal-Hornillos, D., & Moles, M. 2015, *IAU Gene. Assembly*, 29, 2257381
10. Benitez, N., et al. *J-PAS: The Javalambre-Physics of the Accelerated Universe Astrophysical Survey*, 2014
11. Edwin E. Salpeter, *The Luminosity Function and Stellar Evolution*, 1955.
12. Gilles Chabrier, *Galactic Stellar and Substellar Initial Mass Function*, 2003.
13. <https://ned.ipac.caltech.edu/level5/Sept14/Conroy/Conroy2.html>
14. Ahumada, R. et al., *The 16th Data Release of the Sloan Digital Sky Surveys*, 2020.
15. Meisner, A. M. et al., *unWISE Coadds: The Five-year Data Set*, 2019.
16. https://www.nasa.gov/mission_pages/WISE/main/index.html
17. Hernán-Caballero, A. et al., *Near-to-mid infrared spectrum of quasars*, 2016.
18. Boquien, M. et al., *CIGALE: a python Code Investigating GALaxy Emission*, 2019.
19. Noll, S. et al., *Analysis of galaxy spectral energy distributions from far-UV to far-IR with CIGALE: studying a SINGS test sample*, 2019.
20. George E. Smith, *Nobel Lecture: The invention and early history of the CCD*, 2010.
21. <https://www.legacysurvey.org/dr10/description/>
22. http://english.nao.cas.cn/ic2015/icp2015/202007/t20200706_240233.html
23. Blum et al., *The DECam Legacy Survey*, 2016.
24. Silva et al., *The Mayall z-band Legacy Survey (MzLS)*, 2016.
25. <https://www.gnu.org/software/quaastro/>
26. <https://www.gnu.org/home.es.html>
27. Mohammad Akhlaghi, Takashi Ichikawa, *Noise Based Detection and Segmentation of Nebulous Objects*, 2015.
28. Mohammad Akhlaghi, *Carving out the low surface brightness universe with NoiseChisel*, 2019.
29. Mohammad Akhlaghi, *Separating detection and catalog production*, 2016.
30. https://en.wikipedia.org/wiki/Spearman%27s_rank_correlation_coefficient
31. Hinojosa-Goñi et al., *Starburst galaxies in the COSMOS field: clumpy star-formation at redshift $0 < z < 0.5$* , 2018
32. Speagle, J. S. et al. *A Highly Consistent Framework for the Evolution of the Star-Forming "Main Sequence" from $z \sim 0 - 6$* , 2014.
33. Yicheng Guo et al., *Color and Stellar Population Gradients in Passively Evolving Galaxies at $z \sim 2$ from HST/WFC3 Deep Imaging in the Hubble Ultra Deep Field*, 2011.
34. Dekel et al., *Cold streams in early massive hot haloes as the main mode of galaxy formation*, 2009.