Università
degli Studi
di Padova

Università degli Studi di Padova
Dipartimento di Fisica e Astronomia "Galileo Galilei"
Ph.D. course in Astronomy, XXXIV Cycle

Lighting up the dark side of the Universe: new insights on the galaxy build-up through the ALPINE survey

Coordinator: Prof. Giovanni Carraro
Supervisor: Prof. Paolo Cassata

External evaluators: Dr. Roberto Decarli
Dr. Eva Schinnerer

Ph.D. student: Michael Romano

To...

ACKNOWLEDGEMENTS

SOMMARIO

Negli ultimi anni, diversi studi hanno fornito importanti indizi sulla formazione e assemblaggio delle sorgenti primordiali tramite osservazioni della loro emissione ottica e nel vicino infrarosso (NIR), che tracciano la luce ultravioletta (UV) prodotta da stelle appena nate che riesce a fuggire dalle coltri di polvere che le circondano. Questi lavori hanno fornito un quadro accurato ma incompleto dell'Universo, che vede un veloce aumento della densità del tasso di formazione stellare (SFRD) dall'epoca della reionizzazione al cosiddétto mezzogiorno cosmico. Sebbene siano stati proposti diversi meccanismi di formazione delle galassie a $z \gtrsim 3$, essi non tengono conto della quantità di formazione stellare (SF) oscurata dalla polvere, che non viene osservata da telescopi ottici. In letteratura, sono presenti numerosi tentativi di vincolare la componente oscurata dello SFRD a $z > 4$. Tuttavia, la maggior parte di essi si basa su rilevamenti individuali di galassie submillimetriche altamente oscurate, che non rappresentano la maggioranza delle galassie ad alto redshift.

L'avvento dell'interferometro ALMA ha permesso di ottenere una descrizione più approfondita dei processi che regolano l'evoluzione dello SFRD, grazie all'identificazione di galassie primordiale molto polverose. La svolta è arrivata con le sempre maggiori rilevazioni di ALMA della riga del [CII] a 158 μm, il mezzo di raffreddamento principale delle galassie che formano stelle (SFGs). Precedenti osservazioni di [CII] in galassie comuni, fino all'epoca della reionizzazione dell'HI, hanno fornito le basi per il progetto ALPINE. Questo programma rivoluzionario è stato ideato con l'obiettivo di misurare l'emissione nel lontano infrarosso (FIR) di un campione rappresentativo di SFGs a $z > 4$ attraverso la rilevazione del [CII] e del continuo circostante, così da studiarne le proprietà della polvere e del gas, e capire come tali galassie evolvono nel tempo.

Lo scopo di questa tesi è quello di indagare i meccanismi responsabili della for-

mazione delle galassie e le sorgenti polverose che si nascondono nell'Universo primordiale, così da aggiungere un'informazione significativa al complicato puzzle della formazione ed evoluzione delle galassie. A tal fine, abbiamo utilizzato i dati ALPINE per studiare le proprietà della polvere e del gas freddo in galassie primordiali a $4 < z < 6$. Da un lato, abbiamo collegato le osservazioni a più lunghezze d'onda da telescopi ottici/NIR all'emissione nel FIR rilevata da ALMA per fare un censimento delle sorgenti completamente oscurate dalla polvere nell'Universo lontano. In questo modo, possiamo fornire una stima completa della SF totale (UV+FIR) a $z \sim 5$. Dall'altro lato, abbiamo sfruttato l'emissione del [CII] per ricavare la morfologia e la cinematica delle galassie ALPINE. Abbiamo usato queste informazioni per identificare major merger e per valutarne il contributo al processo di assemblaggio delle galassie, in relazione a quello della SF. La nostra analisi suggerisce che la SF sia il processo dominante nella formazione delle galassie a tutte le epoche, sebbene i merger possano aver fornito un contributo significativo sotto determinate assunzioni. Infine, abbiamo testato il [CII] come uno strumento indipendente per stimare lo SFR globale ad alto redshift, trovando che tale transizione può ancora essere considerata un buon tracciante della SF a $z \sim 5$.

I risultati di questa tesi, insieme a quelli della collaborazione ALPINE, ci hanno permesso di aumentare la nostra conoscenza delle prime fasi di formazione delle galassie. Tuttavia, sono necessarie ulteriori indagini sui meccanismi che governano l'evoluzione delle strutture primordiali attraverso il tempo cosmico. Osservazioni future ci permetteranno di caratterizzare la popolazione di galassie molto polverose all'alba dell'Universo, ponendo ulteriori vincoli al bilancio complessivo di SF e ai processi fisici che hanno condizionato l'Universo lontano, fornendo una visione completa e oggettiva del cosmo.

Abstract

The last few years have seen incredible improvements in our knowledge of the distant Universe. Plenty of studies provided important clues on the formation and assembly of primordial sources by looking at their observed-frame optical and near-infrared (NIR) emission, tracing the ultra-violet (UV) light arising from new-born stars that is not hidden behind the curtains of dust within the galaxies. These works led to an accurate yet incomplete picture of the early Universe which manifests itself in a sharp rise of the cosmic star-formation rate density (SFRD) from within the epoch of Reionization to the cosmic noon. Although several mechanisms were proposed to explain the rapid build-up of galaxies at $z \gtrsim 3$, they were biased against the amount of star formation (SF) obscured by the dust, missed by rest-frame UV/optical surveys. Many attempts at constraining the dust-hidden component of the SFRD at $z > 4$ are presented in the literature. However, most of them are based on individual detections of highly dust-obscured submillimeter galaxies, not being representative of the bulk of the galaxy population at high redshift.

The advent of the ALMA interferometer has paved the way for a more in-depth description of the processes that regulate the SFRD evolution, thanks to the identification of very dusty primordial galaxies. The game changer has come with the increasing number of detections of the [CII] 158 μm line with ALMA, the dominant coolant in the spectra of star-forming galaxies (SFGs). Previous observations of [CII] in normal galaxies, up to the epoch of HI reionization, provided the foundation for the ALPINE project. This revolutionary survey was designed with the aim of measuring the rest-frame FIR emission of a representative sample of $z > 4$ SFGs through the detection of [CII] and the surrounding continuum, to study their dust and gas properties and to understand how such galaxies evolve over cosmic time.

The aim of this thesis is to investigate the mechanisms responsible for the galaxy

build-up, and to shed light on the dusty sources hiding in the early Universe, in order to add a significant piece of information to the tangled picture of the formation and evolution of galaxies. To achieve this goal, we made use of ALPINE data to study the properties of dust and cold gas in primordial galaxies at $4 < z < 6$. On one side, we linked multi-wavelength observations from optical/NIR telescopes to the rest-frame FIR emission detected by ALMA to make a census of completely dust-obscured sources in the distant Universe. In this way, we can provide an unbiased view of the total (UV+FIR) amount of SF at $z \sim 5$. On the other side, we took advantage of the [CII] emission to retrieve the morphological and kinematic status of the ALPINE galaxies. We used that information to identify major mergers in the sample, and to assess their relative contribution, along with SF, in the process of galaxy mass-assembly at different ages. Our analysis suggests that *in-situ* SF is the dominant process in the build-up of galaxies at all epochs, although mergers may have provided a significant contribution under well-defined assumptions. Finally, we tested [CII] as another independent tool to estimate the global SFR at high redshift, finding that this atomic line could still be a good tracer of SF at $z \sim 5$.

The results of this thesis, along with those by the ALPINE collaboration, allowed us to increase our current knowledge of the early phases of galaxy formation. Yet, thorough investigations of the mechanisms ruling the evolution of primordial structures across cosmic time are needed in order to confirm our findings. Future observations will allow us to unearth the population of very dusty galaxies at the dawn of the Universe, setting further constraints on the overall budget of SF at early times and on the physics conditioning the star-formation history, providing a complete and unbiased view of the cosmos.

PAPERS

Here below, it is provided a complete list of the papers I wrote as a first author or co-author during the PhD. This thesis is mostly based on some of my original works.

First-author papers

- **Romano M.**, Grazian A., Giallongo E., et al. 2019, *Lyman continuum escape fraction and mean free path of hydrogen ionizing photons for bright z∼4 QSOs from SDSS DR14*, A&A, 645A, 45R

- **Romano M.**, Cassata P., Morselli L., et al. 2020, *The ALPINE-ALMA [CII] Survey: on the nature of an extremely obscured serendipitous galaxy*, MNRAS, 496, 875R

- **Romano M.**, Cassata P., Morselli L., et al. 2021, *The ALPINE-ALMA [CII] Survey: The contribution of major mergers to the galaxy mass assembly at z∼5*, A&A, 653, A111

- **Romano M.**, Morselli L., Cassata P., et al. submitt., *The ALPINE-ALMA [CII] Survey: the population of [CII]-undetected galaxies and their role in the* $\mathrm{L}_{[\mathrm{CII}]}$*-SFR relation*

Papers as part of the ALPINE collaboration

- Cassata P., Morselli L., Faisst A., **et al.** 2020, *The ALPINE-ALMA [CII] survey. Small Lyα-[CII] velocity offsets in main-sequence galaxies at* $4.4 < z < 6$, A&A, 643A, 6C

- Dessauges-Zavadsky M., Ginolfi M., Pozzi F., **et al.** 2020, *The ALPINE-ALMA [CII] survey. Molecular gas budget in the early Universe as traced by [CII]*, A&A, 643A, 5D

- Faisst A. L., Schaerer D., Lemaux B. C., **et al.** 2020, *The ALPINE-ALMA [CII] Survey: Multiwavelength Ancillary Data and Basic Physical Measurements*, ApJS, 247, 61F

- Fujimoto S., Silverman J. D., Béthermin M., **et al.** 2020, *The ALPINE-ALMA [CII] Survey: Size of Individual Star-forming Galaxies at* $z = 4 - 6$ *and Their Extended Halo Structure*, ApJ, 900, 1F

- Gruppioni C., Béthermin M., Loiacono F., **et al.** 2020, *The ALPINE-ALMA [CII] survey. The nature, luminosity function, and star formation history of dusty galaxies up to* $z \simeq 6$, A&A, 643A, 8G

- Jones G. C., Vergani D., **Romano M.**, et al. 2021, *The ALPINE-ALMA [CII] Survey: Kinematic Diversity & Rotation in Massive Star Forming Galaxies at* $z \sim 4.4 - 5.9$, MNRAS, 507, 3540J

- Khusanova Y., Béthermin M., Le Fèvre O., **et al.** 2021, *The ALPINE-ALMA [CII] Survey: Obscured Star Formation Rate Density and Main Sequence of star-forming galaxies at* $z > 4$, A&A, 649A, 152K

- Le Fèvre O., Béthermin M., Faisst A., **et al.** 2020, *The ALPINE-ALMA [CII] survey. Survey strategy, observations, and sample properties of 118 star-forming galaxies at* $4 < z < 6$, A&A, 643A, 1L

- Loiacono F., Decarli R., Gruppioni C., **et al.** 2021, *The ALPINE-ALMA [CII] survey: the luminosity function of serendipitous [CII] line emitters at* $z \sim 5$, A&A, 646A, 76L

- Pozzi F., Calura F., Fudamoto Y., **et al.** 2021, *The ALPINE-ALMA [CII] survey: dust mass budget in the early Universe*, A&A, 653A, 84P

Other papers

- Boutsia K., Grazian A., Calderone G., **et al.** 2020, *The Spectroscopic Follow-up of the QUBRICS Bright Quasar Survey*, ApJS, 250, 26B
- Cassata P., Liu D., Groves B., **et al.** 2020, *ALMA reveals the molecular gas properties of 5 star-forming galaxies across the main sequence at* $3 < z < 3.5$, ApJ, 891, 83C
- Grazian A., Giallongo E., Fiore F., **et al.** 2020, *On the AGN Nature of Two UV-bright Sources at* $z_{spec} \sim 5.5$ *in the CANDELS Fields: An Update on the AGN Space Density at* $M_{1450} \sim -22.5$, ApJ, 897, 94G
- Morselli L., Rodighiero G., Enia A., **et al.** 2020, *A panchromatic spatially resolved analysis of nearby galaxies - II. The main sequence - gas relation at sub-kpc scale in grand-design spirals*, MNRAS, 496, 4606M

CONTENTS

List of Figures

LIST OF TABLES

List of Abbreviations

AGN:	active galactic nucleus
BIC:	Bayesian information criterion
CGM:	circumgalactic medium
EW:	equivalent width
FIR:	far-infrared
FWHM:	full width at half maximum
GSMF:	galaxy stellar mass function
IMF:	initial mass function
ISM:	interstellar medium
NIR:	near-infrared
PA:	position angle
PAH:	polycyclic aromatic hydrocarbon
PDR:	photodissociation region
PSB:	post-starburst
PSF:	point spread function
PVD:	position-velocity diagram
rms:	root mean square
SED:	spectral energy distribution
SFG:	star-forming galaxy
SFH:	star-formation history
SFR:	star-formation rate
SFRD:	star-formation rate density
SMG:	submillimeter galaxy
SMHM:	stellar mass - halo mass relation

SNR:	signal-to-noise ratio
SPW:	spectral window
sSFR:	specific star-formation rate
UV:	ultra-violet

1 | INTRODUCTION

Galaxies are the building blocks of the Universe. Since many years, their origin and evolution have been the subject of a deep investigation aimed at collecting the missing pieces of cosmic structure formation (e.g., Silk & Mamon 2012; Madau & Dickinson 2014; Naab & Ostriker 2017; Dayal & Ferrara 2018). Nowadays, the ΛCDM paradigm is by far the most diffused and accepted cosmological scenario portraying the galaxy assembly, for which small fluctuations in the initial density field of the Universe grow through gravitational instability driven by cold dark matter particles, giving birth to the variety of structures we observe today (Peebles 1982). Within this framework, galaxies grow in a hierarchical, "bottom-up" manner where small dark matter halos form earlier and merge together making increasingly larger systems over time (e.g., Blumenthal et al. 1984). The effects of gravity in molding the cosmic web are now detailed with very high accuracy through N-body numerical simulations (e.g., Springel et al. 2005, 2008; Stadel et al. 2009; Klypin et al. 2011), and confirmed by observations of the large-scale structures through galaxy distributions and clustering (e.g., Peacock et al. 2001; Reid et al. 2010). On the other hand, at smaller scales, the collapse of baryons toward the center of dark matter halos is dictated by a series of complex physical mechanisms that are anything but easy to interpret and model, preventing a complete and thorough description of the galaxy formation process.

Over the last decades, considerable efforts to investigate the evolution of galaxies through cosmic time have been undertaken. To piece the history of the Universe together from primordial epochs up to the present day, multi-wavelength observations of the galaxy populations at different times have been collected, especially in the rest-frame ultra-violet (UV) and far-infrared (FIR) regimes, tracing respectively the instantaneous star formation produced by short-lived massive stars, and the UV light emitted by young stellar populations and absorbed by interstellar dust (then re-emitted at longer wavelengths). Particularly, the star-formation history (SFH) of the Universe is enclosed in the cosmic star-formation rate density (SFRD), that is the amount of star formation per unit time and per unit volume at a specific redshift. This quantity has been characterized in great detail up to 13 billion years ago, when the first galaxies formed (e.g., Bouwens et al. 2011b; Coe et al. 2013). The emerging picture outlines three main epochs of cosmic activity (see Figure 1.1, left panel), with the SFRD showing a steep rise during the cosmic reionization from $z \sim 10$ to 6 (e.g., Bouwens et al. 2011a, 2012a), peaking at $1 < z < 3$ within the so-called *cosmic noon* during which most of the stars observed today have formed (e.g., Reddy et al. 2008; Marchesini et al. 2009), and finally decreasing by a factor of ~ 10 to the local Universe (Lilly et al. 1996; Madau et al. 1996; Madau & Dickinson 2014). Furthermore, it is now ascertained that the bulk of the star formation takes place at all epochs along a well-defined "main-sequence" of galaxies, with a relatively tight dispersion (< 0.3 dex) in the star-formation rate (SFR) versus stellar mass plane (e.g., Noeske et al. 2007; Rodighiero et al. 2011; Speagle et al. 2014).

A remarkable fact is evidenced in Figure 1.1 (right panels): at the time of that study, most of the constraints on the SFRD at $z > 3$ came from rest-frame UV emission, missing the amount of star formation obscured by dust (e.g., Bouwens et al. 2012a,b, 2015; Schenker et al. 2013; Oesch et al. 2018). Indeed, the UV ionizing photons produced by young stars are easily absorbed by dust grains in the interstellar medium (ISM), and then re-emitted at longer wavelengths in the FIR. At those high redshifts, UV-selected surveys could be thus biased against dusty star-forming galaxies (SFGs) whose rest-frame FIR emission is redshifted to the sub-mm regime, and observatories at these frequencies were not sensitive enough to catch the bulk of this elusive population of sources. Nevertheless, in the last decades, few attempts of constraining the dust-obscured SFRD at high redshift have been made. Dedicated space- and ground-based telescopes have contributed to this aim, like the Cosmic Background Explorer (COBE), the *Herschel* and *Spitzer* satellites (e.g., Puget et al. 1996; Hauser et al. 1998; Werner et al. 2004; Pilbratt et al. 2010) or the Infrared Space Observatory (ISO; Kessler et al. 1996), radio interferometers like the Karl G. Jansky Very Large Array (VLA; Thompson et al. 1980; Perley et al. 2011), single-dish submillimeter telescopes such as the James Clark Maxwell Telescope

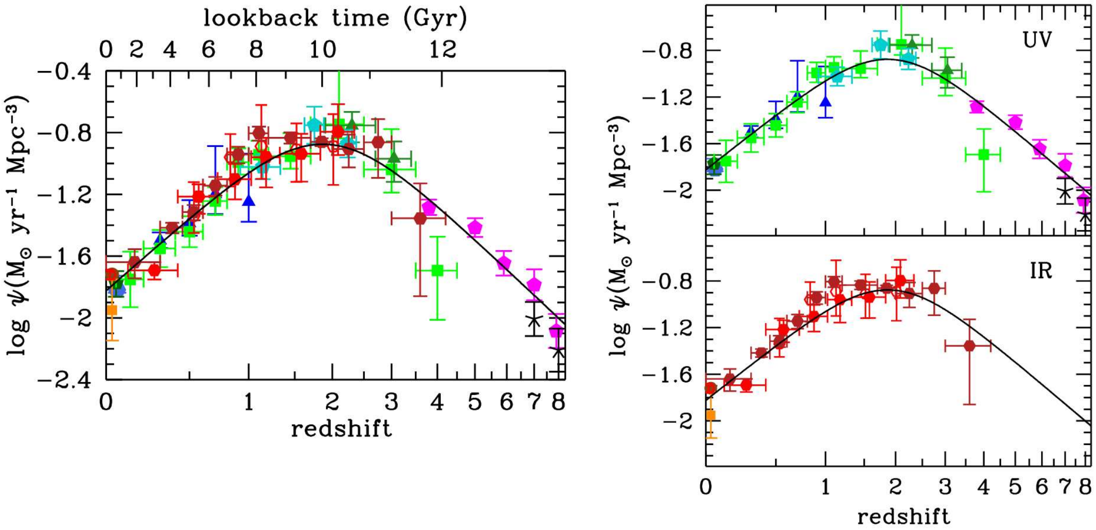


Figure 1.1: Cosmic SFH as a function of redshift and lookback time (left panel). Top and bottom panels on the right show the difference from rest-frame UV (blue, green and magenta points) and infrared (red points) measurements, tracing the dust- unobscured and obscured star-formation activity. The solid black line is the best fit to the data parameterized as in Equation 15 by Madau & Dickinson (2014) (from which the Figure is taken).

(JCMT; Holland et al. 1999; Robson et al. 2017), the IRAM 30-meter telescope (Baars et al. 1987), and submillimeter interferometers like the Submillimeter Array (SMA; Ho et al. 2004) or the NOrthern Extended Millimeter Array (NOEMA; Guilloteau et al. 1992). All these facilities have allowed us to discover a large number of dusty SFGs, bright in the sub-mm (thus named *submillimeter galaxies*, SMGs; Smail et al. 1997; Barger et al. 1998; Hughes et al. 1998; Blain et al. 2002; Simpson et al. 2014, 2017), characterized by large FIR luminosities ($\gtrsim 10^{12}$ L$_\odot$), high SFRs ($\gtrsim$ 100 M$_\odot$ yr^{-1}), and peaked at $z \sim 2$ (e.g., Wardlow et al. 2011; Casey et al. 2014; Swinbank et al. 2014; Brisbin et al. 2017). At earlier epochs ($z > 4$), most of the observed SMGs are very bright and with large SFRs (e.g., Capak et al. 2011; Riechers et al. 2013, 2017; Marrone et al. 2018), lacking information from the less extreme population of dusty SFGs. Direct derivations of the dust-hidden component of the SFRD were undertaken, as well. However, their robustness were hampered by inaccurate redshift determinations, complex completeness corrections, or poor constraints on the infrared luminosity functions (e.g., Gruppioni et al. 2013; Rowan-Robinson et al. 2016; Koprowski et al. 2017; Wang et al. 2019).

The turning point came with the Atacama Large Millimeter/submillimeter Array (ALMA; Wootten & Thompson 2009; Andreani 2010). Thanks to its unprecedented sensitivity and resolution, ALMA has opened a new window on the cold and dusty Universe up to the appearance of the first galaxies. Aside from most extreme star-forming sources with large SFRs and bright in the sub-mm (e.g., Riechers et al. 2014; Simpson et al. 2014; Dudzevičiūtė et al. 2020), a large population of faint SMGs has been also discovered at $z > 4$ thanks to recent ALMA surveys (e.g., Aravena et al. 2016; Walter et al. 2016; Dunlop et al. 2017; Franco et al. 2018; Yamaguchi et al. 2019; Fudamoto et al. 2021). Many of them have large amounts of dust, resulting invisible in the optical and near-infrared (NIR) bands even to the watchful eye of the Hubble Space Telescope (HST), and are thus renamed *HST-dark* galaxies. These sources usually present a visible counterparts in the mid-infrared (e.g., Wang et al. 2019; Yamaguchi et al. 2019), although in some cases they do not show emission even in the deeper *Spitzer*-IRAC bands (e.g., Williams et al. 2019; Romano et al. 2020; see Chapter 3). As a matter of fact, Dudzevičiūtė et al. (2020) completed an ALMA study of a large sample of SMGs selected from the ALMA-SCUBA-2 survey of the Ultra Deep Survey (UDS) field (AS2UDS; Stach et al. 2019), finding that $\sim 17\%$ of their sources are undetected in the optical/NIR down to $K_s = 25.7$ mag. Fudamoto et al. (2021) have recently discovered two dust-obscured SFGs at $z > 6$ as part of the ALMA large program REBELS (Reionization Era Bright Emission Line Survey; Bouwens et al. 2021). They found that this population of dusty galaxies could provide up to 25% of the $z > 6$ cosmic SFRD. Therefore, the contribution of such obscured galaxies to the star formation could be substantial, especially at high redshift, where the observations are biased against very dusty systems and there is no complete census of star formation.

To deduce the precise portion of hidden star formation not traced by the plenty of available deep UV observations, a large sample of sources detected based on their FIR emission in blind surveys would be needed. Several attempts have been made during the last few years through blind continuum imaging (e.g., Dunlop et al. 2017) or spectral line scans (e.g., Decarli et al. 2016; Walter et al. 2016; Pavesi et al. 2018). However, even with the amazing capabilities of the ALMA telescope, the relatively small fields of view would require an incredible observing time to collect a statistically significant sample of *normal* galaxies in the distant Universe ($z \gtrsim 4$), representative of the bulk population of SFGs at all epochs. Therefore, the focus has moved to the detection of targeted sources

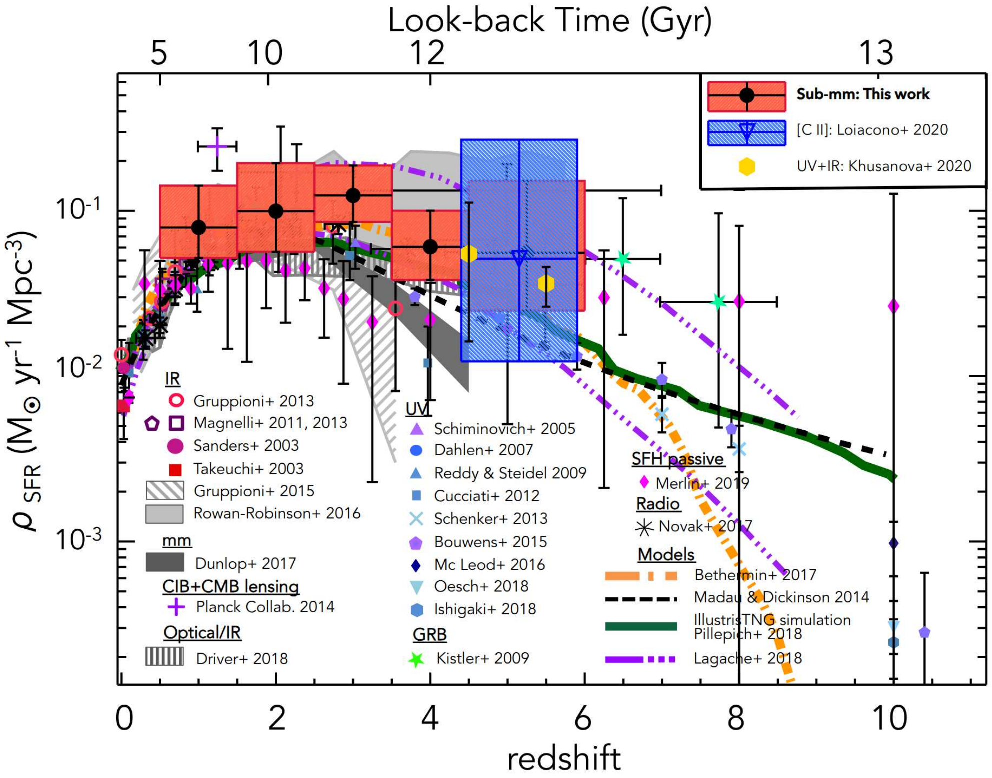


Figure 1.2: SFRD as a function of redshift and look-back time. Black circles and red boxes represent the SFRDs obtained by integrating the infrared luminosity function of continuum non-target detections in ALPINE and their 1σ uncertainties, respectively (Gruppioni et al. 2020). Open triangle with the blue box is the result obtained from the [CII] luminosity function of serendipitous line emitters by Loiacono et al. (2021), while the yellow filled hexagons are the values obtained by Khusanova et al. (2021) from the UV+FIR emission of the ALPINE targets. All the other points come from other surveys in the literature probing UV, optical, IR, mm, and radio emission, or from gamma-ray bursts at $z > 4$ (see references in Gruppioni et al. 2020). Models are also reported as colored lines. In particular, the dashed-black curve reports the model by Madau & Dickinson (2014), as shown in Figure 1.1. Figure from Gruppioni et al. (2020).

selected on their known properties, such as stellar masses, SFRs or luminosities, and in the serendipitous observations of fainter objects in the fields of the main galaxies (e.g., Hatsukade et al. 2013; Capak et al. 2015; Maiolino et al. 2015; Fujimoto et al. 2016).

Most recently, the *ALMA Large Program to INvestigate [CII] at Early times* (ALPINE, PI: O. Le Fèvre; Le Fèvre et al. 2020; Faisst et al. 2020b; Béthermin et al. 2020) has detected the dust continuum and the bright FIR singly ionized carbon ([CII]) line at 158 μm in a large sample of SFGs at $4 < z < 6$, providing new constraints on the shape of the global SFRD in the distant Universe (see Chapter 2). In particular, Gruppioni et al. (2020) identified 56 sources serendipitously detected in continuum at ~ 860 and $\sim 1000\ \mu$m in ALPINE and used them to derive the cosmic SFRD that is found to remain almost constant from $z \sim 1$ up to $z \sim 6$, and significantly higher than the UV/optical measurements at the same redshifts. Indeed, they found that $\sim 17\%$ of the total SFRD at $z > 3$ is provided by HST-dark galaxies, confirming that a non-negligible amount of star-formation activity at high redshift is still missed by survey probing the rest-frame UV emission from galaxies. These findings are also in agreement with independent measurements of the total SFRD obtained by summing the UV and FIR SFR components of the ALPINE targets (Khusanova et al. 2021), and by exploiting the [CII] luminosity function of serendipitous detections in the survey (Loiacono et al. 2021). A similar result is also found by Talia et al. (2021) who built up and analyzed a sample of radio-selected, UV-dark dusty galaxies at $z \gtrsim 3$. They found that the contribution of these systems to the SFRD could be as high as 40% of the corresponding UV-based quantity, posing further questions on the presence of such dusty galaxies in the first two billion years after the Big Bang, currently not predicted by galaxy formation models and simulations. All these outcomes suggest that the evolution of the cosmic SFRD could be less steep than previously measured at $z \gtrsim 3$, as illustrated in Figure 1.2. However, further observations are needed to increase the statistics, reducing the current uncertainties and possibly confirming these results.

Setting aside the contribution of the dust-obscured star formation at early epochs, another relevant point of discussion resides in the physical processes that shape the history of the Universe, defining the mass-assembly of galaxies through cosmic time. Indeed, it is still unclear which are the main mechanisms responsible for the rapid increase of the cosmic SFRD at $z < 6$. Two principal characters are thought to drive this evolution, that are the accretion of cold gas (both from the internal reservoir or from the outer environment of galaxies) and the merging of galaxies (e.g., Hopkins et al. 2006; Dekel et al. 2009; Naab & Ostriker 2017).

The first process requires the supply of new cold gas to the galaxy from the filaments of the cosmic web, refreshing and/or enhancing the star formation (e.g., Kereš et al. 2005; Dekel et al. 2009). Although this could be one of the main modes of galaxy assembly (e.g., Kereš et al. 2009; Bouché et al. 2010), direct observations of gas flowing through the filaments are tricky while indirect evidence has increased over the last years (Bouché et al. 2013, 2016; Zabl et al. 2019). However, *in-situ* star formation is observable through a variety of different proxies and up to the earliest epochs (e.g., Madau & Dickinson 2014 and references therein).

On the other hand, galaxy mergers are also thought to play an important role in the stellar mass growth over cosmic time. They are relatively common through the history of the Universe and are observable with different methods and at different stages (e.g., Conselice & Arnold 2009; Xu et al. 2012; Conselice 2014; Tasca et al. 2014; Ventou et al. 2017; Duncan et al. 2019; Romano et al. 2021). The merging of galaxies is a natural prediction of the hierarchical structure formation model (White & Frenk 1991; Springel et al. 2005; Klypin et al. 2011) for which dark matter halos grow their mass and affect the build-up of galaxies in the Universe (Kauffmann et al. 1993; Khochfar & Burkert 2005). Not only can mergers increase the stellar mass of galaxies (up to a factor of 2 for major mergers, i.e., for galaxies of nearly equal stellar mass; López-Sanjuan et al. 2012; Oser et al. 2012; Kaviraj et al. 2014), but they can also trigger starbursts and active galactic nuclei (AGNs; Silk & Rees 1998; Kartaltepe et al. 2012; Chiaberge et al. 2015; see also Shah et al. 2020 who found that galaxy interactions do not significantly enhance AGN activity up to $z \sim 3$), in some cases blowing the gas out from the galaxy which arrests, at least temporarily, the formation of new stars (Silk & Rees 1998; Fabian et al. 2006; Cattaneo et al. 2009). Therefore, observing mergers at different epochs is key to shed light on the relative contribution of the distinct processes that rule the assembly of galaxies in the Universe.

Many fundamental points have been raised so far regarding the formation and composition of primordial galaxies, and their evolution through the cosmic epochs. In summary:

- *What is the amount of star formation provided by dust-obscured galaxies at $z > 3$?*
- *What are the properties of such dusty objects?*
- *Which are the main processes molding the shape of the cosmic SFRD?*

In order to deal with these questions, we need the knowledge of one of the main ingredients in the galaxy evolution recipe, that is the cold gas, the fuel for star formation in galaxies.

1.1 Cold gas in the early Universe

The cold gas content in galaxies represents the reservoir of material from which stars form (e.g., Solomon & Vanden Bout 2005; Omont 2007; Carilli & Walter 2013). Its amount in the ISM of distant sources could have a large impact on the structure of the SFRD in the early Universe. For instance, a large gas content or a high star-formation efficiency can lead to an increase in the star-formation activity, with the consequence of a more rapid growth of galaxies (e.g., Davé et al. 2012; Riechers et al. 2019). Thanks to the current powerful observatories such as the ALMA and VLA telescopes, the cold gas can be easily traced up to very early times by looking at the emission of molecular transitions or atomic fine structure lines.

The molecular gas budget in the Universe is dominated by molecular hydrogen (H_2). However, given the high temperature excitation ($T \gtrsim 100$ K) of its high-energy levels and forbidden lowest energy transitions, H_2 is practically not observable in emission (e.g., Kennicutt & Evans 2012; Bolatto et al. 2013). On the other hand, carbon monoxide (CO) is the most abundant molecule after H_2 and has lower excitation requirements ($T \sim 5$ K for the first excited state) than molecular hydrogen, making it easy exploitable for constraining the molecular gas mass in nearby and intermediate galaxies (e.g., Daddi et al. 2010; Saintonge et al. 2011; Genzel et al. 2015; Tacconi et al. 2018). Yet, at $z > 4$, it becomes mostly detectable at the highest levels of transition and in high excitation states (e.g., Carilli & Walter 2013), although cases of low-J emissions are also present in the literature (e.g., Riechers et al. 2021).

Other species can be used to trace the molecular gas mass in the ISM. Atomic fine structure lines have been ascribed as the dominant coolants of the gas in SFGs, reradiating the far-UV energy gained by star formation at higher wavelengths, in the rest-frame FIR (e.g., Spitzer 1978). Depending on their ionization potential higher or lower than that of neutral hydrogen (HI, 13.6 eV), these lines can trace both the ionized medium (e.g., [NII], [OIII]; Ferkinhoff et al. 2010; Nagao et al. 2012), and the neutral ISM (e.g., [CI], [OI], [CII]; Malhotra 2001; Walter et al. 2011; Coppin et al. 2012).

In the following, I will describe in detail the [CII] 158 μm line as it represents by now, especially at $z > 4$, a workhorse for the study of the ISM, and because it has been used in this thesis as the main tool for characterizing distant galaxies through the ALPINE survey.

1.1.1 The [CII] 158 *µm* line

The [CII] 158 µm line has become a cornerstone for the study of galaxy evolution. It is the brightest line in the FIR band (e.g. Stacey et al. 1991), near the peak of dust continuum emission, representing the dominant coolant of the neutral diffuse ISM (e.g., Wolfire et al. 2003). Since it has a lower ionization potential than neutral hydrogen, i.e. 11.3 eV compared to 13.6 eV, this line can also be excited in other components of the ISM, such as in dense photodissociation regions (PDRs; Hollenbach & Tielens 1999), and in the diffuse ionized gas (e.g. Cormier et al. 2012). In principle, in order to remove the ambiguity on the interpretation of the [CII] emission, the relative contribution of the various gas phases should be assessed. However, different studies suggest that the bulk of the [CII] emission arises from the external layers of molecular clouds heated by UV photons in the PDRs (Stacey et al. 1991; Madden et al. 1997; Kaufman et al. 1999; Cormier et al. 2015; Pavesi et al. 2016).

Given the strength of the line (with fluxes $> 10^3$ times stronger than CO(1-0) in SFGs; Carilli & Walter 2013), the [CII] emission is easily accessible both in local and primordial galaxies (e.g., Boselli et al. 2002; Maiolino et al. 2005; Cox et al. 2011; Carniani et al. 2013; De Breuck et al. 2014; De Looze et al. 2014; Pineda et al. 2014). At first, most of the detections at $z > 4$ came from sources with extreme SFRs ($\gtrsim 1000$ M$_\odot$ yr^{-1}), but the situation rapidly changed with the advent of ALMA that enabled to target hundreds of normal SFGs (SFR < 100 M$_\odot$ yr^{-1}) up to the epoch of Reionization and beyond (e.g., Capak et al. 2015; Pentericci et al. 2016; Bradač et al. 2017; Matthee et al. 2017; Smit et al. 2018; Le Fèvre et al. 2020), offering a new way for studying high-z galaxies.

There are numerous ways to take advantage of [CII] as a probe of the ISM properties in distant galaxies. First, being a good tracer of the systemic velocity of a source, it can be used to spectroscopically confirm the redshift of galaxies through cosmic time (e.g., Pentericci et al. 2016; Matthee et al. 2019; Cassata et al. 2020; Matthee et al. 2020). Furthermore, the [CII] emission can provide important information on the presence of outflows (e.g., Gallerani et al. 2018; Ginolfi et al. 2020b), and on the kinematics of the ISM (e.g., Jones et al. 2021; Romano et al. 2021). Therefore, the comparison of [CII] observations with simulations is mandatory to disentangle the diverse processes that take place in galaxies, in order to understand how they shape the observed morphology and kinematics of the line (e.g., Olsen et al. 2017; Vallini et al. 2017; Ferrara et al. 2019; Kohandel et al. 2019).

Finally, although recent studies suggest that the [CII] line could be a good tracer of the molecular gas (e.g., Fahrion et al. 2017; Zanella et al. 2018), a tight relation between its luminosity and the global SFR is observed from local to intermediate galaxies (e.g. De Looze et al. 2014; Herrera-Camus et al. 2015). If also confirmed at higher redshifts, this would be a complementary and independent way to estimate the total SFR in galaxies, with respect to the combination between UV and FIR measurements. Many studies are focusing on this question, with some finding a good agreement between high-z and the local [CII]-SFR relation (e.g., Matthee et al. 2019; Schaerer et al. 2020; Romano et al. submitt.; see Chapter 6), and other showing the presence of the so-called "[CII]-deficit" for which galaxies deviate from the local relation (e.g., Díaz-Santos et al. 2013; Carniani et al. 2018; Harikane et al. 2018). In particular, this deficit is represented by a drop in the [CII] to infrared luminosity ratio ($L_{[CII]}/L_{IR}$) as a function of increasing L_{IR} (e.g., Malhotra et al. 2001a; Graciá-Carpio et al. 2011), that could be ascribed to low metallicity in the ISM of galaxies and/or to intense radiation fields, as suggested by theoretical studies (e.g., Vallini et al. 2015; Olsen et al. 2017; Stark et al. 2017; Vallini et al. 2017; Lagache et al. 2018; Ferrara et al. 2019). On the other hand, other works do not find a (significant) [CII]-deficit at high redshifts, concluding that this behavior is not universal (e.g., Zanella et al. 2018; Schaerer et al. 2020).

ALMA is by far the most powerful facility exploitable to finally solve this kind of issues thanks to the detection of the [CII] emission in a statistical way through large programs like ALPINE (Le Fèvre et al. 2020; Béthermin et al. 2020; Faisst et al. 2020b) during the rising of the SFRD, or with the forthcoming results by the REBELS survey (Bouwens et al. 2021) in the $z > 6.5$ Universe. In the following, I will explore the ALMA capabilities in the study of the high-redshift Universe, providing some basics concepts of its functioning.

1.2 ALMA: A NEW ERA FOR THE STUDY OF HIGH-REDSHIFT GALAXIES

The last ten years have seen a prodigious progress in our knowledge of the dust-obscured star formation in distant sources. Since the early science operations in 2011, the ALMA telescope has directed its antennas to the sky to catch signals from the dawn of the Universe to the present day, particularly in the (sub-)mm regime, where the emission from the dark and cold dust and gas spread over SFGs shines.

What makes ALMA so powerful and efficient in the field of galaxy formation and evolution?

ALMA is a radio telescope composed of 66 antennas operating at 0.3 to 3.6 mm, displaced on the Chajnantor plateau over 5000 m above sea level, in the Chilean Atacama Desert. The high elevation and climate conditions make it one of the drier sites in the world, providing the right requirements to detect (sub)millimeter light from Earth, that would be otherwise absorbed by water vapour in the atmosphere. The ALMA antennas are arranged in two principal arrays: the *12-m Array*, which comprises fifty antennas each measuring 12 m in diameter and that can be used for high-resolution imaging, and the *Atacama Compact Array (ACA)*, composed of twelve closely spaced 7-m antennas, plus four 12-m antennas for single-dish observations.

The need for observing with more dishes directly comes from the limits imposed by the resolution. Indeed, for single telescopes, the resolution goes as $\sim \lambda/D$, where λ is the wavelength of observation and D is the diameter of the telescope. As evident, by fixing the size of the dish, the resolution is worse for radio telescopes (observing at longer wavelengths) than for optical or infrared ones. In order to reach similar resolutions (or even higher) to those achievable at higher frequencies, the ALMA antennas act together as a single telescope (i.e., an *interferometer*) with a size given by the maximum distance between them which can go from 15 m up to 16 km. In particular, the antennas can be moved across the plateau to form different configurations, depending on the level of details to reach during an observation. The larger the distance between antennas, the higher their resolving power. Therefore, depending on the chosen configuration, it is possible to probe both the broader structures of an astronomical object and its finest features.[1]

In this thesis, we have used the ALMA data collected during cycles 5 and 6 of 118 SFGs, as part of the ALPINE large program. The observations spent 15-25 min on source for a total elapsed time of $\sim$ 70 hours, and were taken in Band 7 ($\nu_{\rm obs}$ = 275-373 GHz) with compact array configurations C43-1 and C43-2, reaching $>$ 0.7 arcsec resolution (Béthermin et al. 2020). I will introduce in the following some basic concepts of interferometry that are useful to understand how ALMA works and the results presented in the next chapters.

1.2.1 Basics of interferometry

The astonishing ALMA capabilities rely on the principles of interferometry. This technique is especially used in the millimeter/radio regime, where the long wavelengths observed would result in a poor angular resolution. Interferometry brings together the signals collected by two or more separated antennas and processes them with a supercomputer (named *correlator*) to make them act as a single telescope as large as the whole array. This allows to achieve a resolution $\sim \lambda/B$, where B is called *baseline* and, in this case, represents the maximum distance between a couple of antennas.

1.2.1.1 Visibilities

Contrary to single-dish telescopes, an interferometer does not directly measure the surface brightness distribution of an object in the sky, rather interference patterns produced by the correlation between the signals at certain angular scales, called *visibilities*, that are determined by the separations between the pairs of antennas composing the array.

To understand how the interference pattern is related to the observed source brightness, let us consider an array composed of only two antennas looking at the same object in a narrow frequency range centered on $\nu = \omega/2\pi$, as represented in Figure 1.3. The voltage V_1 and V_2 measured by antennas 1 and 2, respectively, can be written as

$$V_1 = Ve^{i\omega(t-\tau)} \qquad\qquad V_2 = Ve^{i\omega t} \tag{1.1}$$

---

[1] See `https://public.nrao.edu/telescopes/alma/` for more details.

where $\tau = \vec{b} \cdot \hat{s}/c$ is the geometric delay experienced by antenna 1 with respect to antenna 2, $\vec{b}$ is the baseline vector pointing from one antenna to the other, $\hat{s}$ is the unit vector indicating the direction of the observed source, and c is the speed of light. The delay has to be taken into account in order for the two signals to arrive with the same phase at the correlator. Moreover, considering θ as the angle between the source and the vertical direction, we can define the projected separation of the two antennas toward the source as $u = b\ cos(\theta)$. At the correlator, the two signals are multiplied and averaged over time, providing the correlator response

$$R(\tau) = \left\langle V_1\ V_2 \right\rangle = \int V^2\ e^{-i\omega\tau}\ d\tau. \tag{1.2}$$

Actually, the baseline vector is a 2D vector with components b_1 and b_2 in the plane of the antennas. We can thus define $u = b_1\ cos(\theta)$, and $v = b_2\ cos(\alpha)$, where α is the angle between the observed source and the direction orthogonal to θ. To these components will correspond the two directions in the sky plane, l and m, so that Equation 1.2 becomes

$$R(u, v) = \iint V^2\ e^{-2\pi i(ul+vm)}\ dldm. \tag{1.3}$$

By looking at Figure 1.4, it is evident that u and v represent the projection of each baseline onto the plane of the sky in the E-W and N-S directions, respectively, and are called *spatial frequencies*, measured in units of wavelength. On the other hand, the components l and m are defined along the same directions of the u, v components but in the tangent sky plane, measured in units of radians.

At this point, it is worth noting that $V^2 \propto P$, where P is the power received by the antennas, and that $P \propto I_\nu$, where I_ν is the intensity distribution on the sky. Therefore, the quantity measured by the correlator becomes

$$\mathcal{V}(u, v) = \iint I(l, m)\ e^{-2\pi i(ul+vm)}\ dldm = Ae^{i\phi}, \tag{1.4}$$

that is called *complex visibility*. As suggested by the name, being a complex number, the visibility can be expressed by an amplitude (A) and a phase (ϕ), which provide information on the source brightness and its location with respect to the phase center.

It is evident from Equation 1.4 that the visibility function $\mathcal{V}(u, v)$ is the 2D Fourier

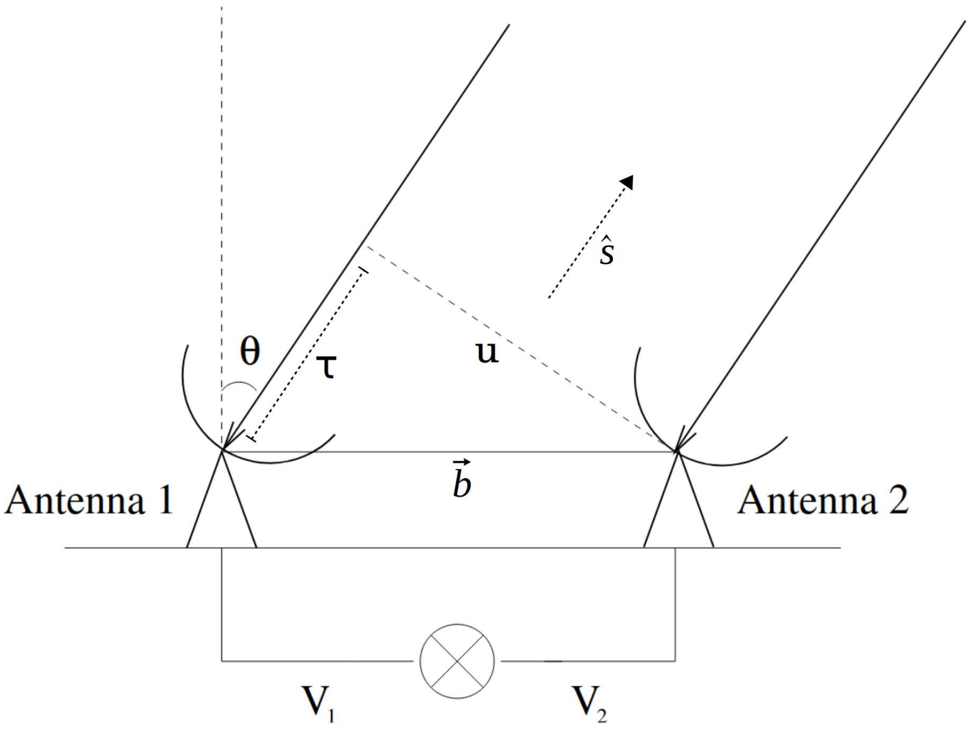


Figure 1.3: Sketch of a two-antenna interferometer. The two antennas are spaced by a physical distance b (i.e., the baseline) and look at the same source at an angle θ from the vertical direction. The projected distance of the two antennas along the direction of the source is represented by the segment u. The radiation from the source must travel an extra time τ to reach Antenna 1. The two antennas are connected to a correlator that combines the corresponding signals. Figure adapted from Birkinshaw & Lancaster (2005).

transform of the sky brightness distribution $I(l, m)$. It is also true that $I(l, m)$, that represents the physical quantity we are interested in, can be obtained as the inverse Fourier transform of $\mathcal{V}(u, v)$ as

$$I(l.m) = \iint \mathcal{V}(u, v)\ e^{2\pi i(ul+vm)}\ dudv. \tag{1.5}$$

The link between $I(l, m)$ and $\mathcal{V}(u, v)$ is dictated by the so-called van Cittert-Zernike theorem: by measuring the distribution of complex visibilities as a function of the coordinates u and v (i.e., in the *uv-plane*), the brightness distribution on the sky plane (also called *image plane*) can be recovered.

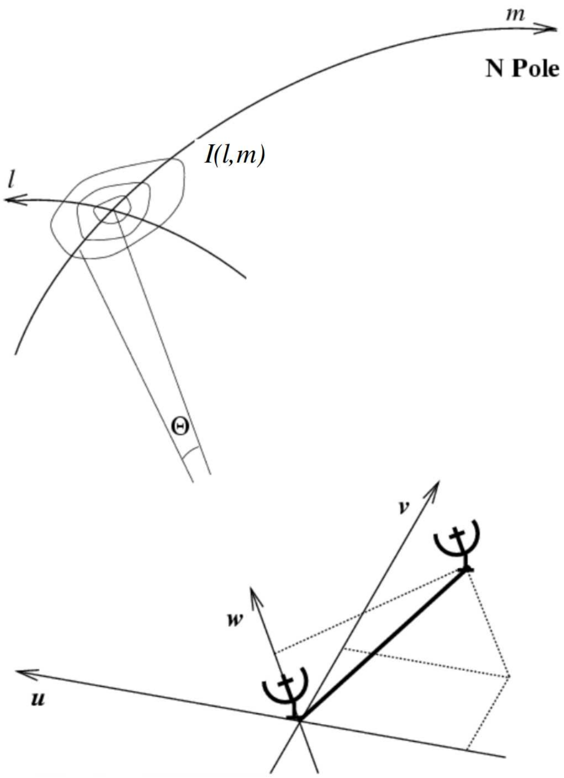


Figure 1.4: Geometrical relation between a source in the image plane with sky brightness $I(l, m)$ and an interferometer consisting of two antennas in the uv-plane. Figure from D. J. Wilner, *14th Synthesis Imaging Workshop*.

To each baseline will correspond a single visibility in the uv-plane. However, given the complex nature of $\mathcal{V}(u, v)$, two visibilities will result from a measurement, one at position (u, v) and the other arising from its complex conjugate at $(-u, -v)$, consisting in a total of $N(N-1)$ points in the uv-plane, where N is the number of antennas. In order to recover the most accurate representation of the true sky brightness, the uv-plane should be uniformly sampled. The coverage of this plane can be improved in different ways. On one hand, multiple antennas can be added to an array at different positions and with different configurations, to sample different scales. On the other hand, although the physical distance between antennas will not mutate, their projected baselines on the sky will change through Earth's rotation. In this way, while Earth rotates, the points in the uv-plane move along an ellipse sampling more and more visibilities in the plane,

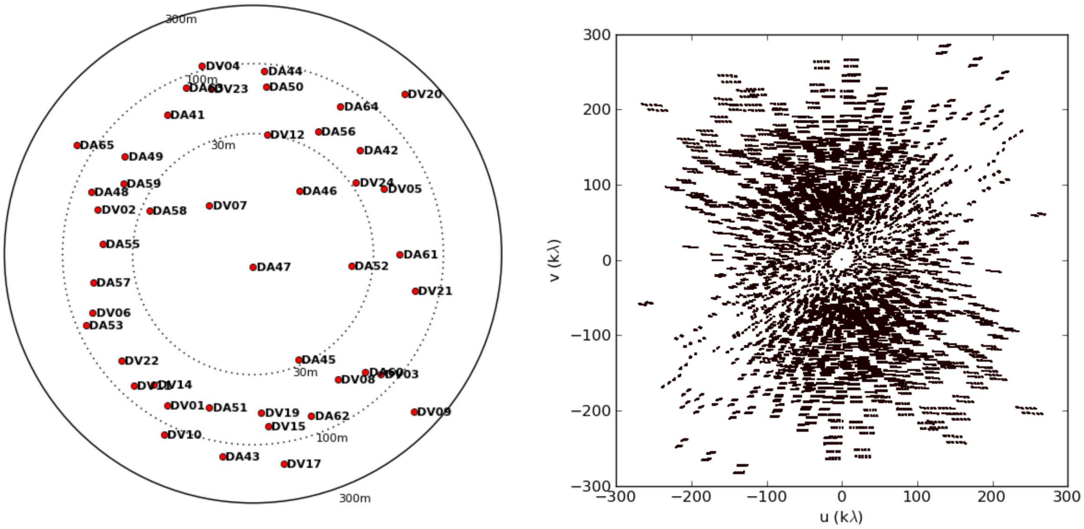


Figure 1.5: *Left:* Antenna positions for the ALPINE target DEIMOS_COSMOS_665626 in logarithmic scale. In particular, this is the C43-2 ALMA configuration, corresponding to > 0.7 arcsec resolution. Each antenna is labeled with its name. *Right:* uv-plane coverage for DEIMOS_COSMOS_665626 in kilowavelengths (meters). The on-source integration time was ~ 16 min.

depending on the observing time. This technique is shown in Figure 1.5, where the antenna configuration of one of the ALPINE targets (DEIMOS_COSMOS_665626) and its corresponding uv-plane are reported. As a result, the combination of the number of antennas, their configuration, and of the observing time, will increase the sampling, leading to a final image that can resemble the true sky brightness distribution. At the same time, it is not possible to completely sample the uv-plane and obtain all the visibilities. This is evident in Figure 1.5 (right panel), where the inner whole (including baselines shorter than the single antenna diameter) prevents us to obtain information on the large-scale structure of the source (also known as *zero-spacing problem*). Possible solutions to this issue are to extrapolate the emission samples by these short baselines with a long integration time, or to take advantage of single-dish telescopes.

1.2.1.2 Imaging and deconvolution

As introduced in the previous section, the observed visibilities are sampled only in certain places of the uv-plane. This issue can be quantified by the *sampling function* $S(u, v)$, that is equal to unity where data are taken, and zero otherwise. Accounting for this, Equation 1.5 can be written as

$$I^D(l,m) = \iint \mathcal{V}(u,v)S(u,v)\; e^{2\pi i(ul+vm)}\; du dv, \tag{1.6}$$

where $I^D(l,m)$ is the measured sky brightness, or *dirty image*. It is related to the true intensity distribution through the convolution theorem for Fourier transforms, such that

$$I^D = I \otimes B, \tag{1.7}$$

where I is the sky brightness of Equation 1.5, convolved with the *synthesized beam* or *point spread function* (PSF), defined as the inverse Fourier transform of the sampling function like

$$B(l,m) = \iint S(u,v)\; e^{2\pi i(ul+vm)}\; du dv. \tag{1.8}$$

Therefore, it is possible to solve Equation 1.7 for I from the knowledge of the dirty image and the PSF. This process is known as *deconvolution problem* and it is usually addressed with the CLEAN algorithm (Högbom 1974), that forms the basis for all the other deconvolution algorithms developed in radio astronomy. This assumes that the true sky brightness is represented by a collection of point sources, and uses an iterative process to find their positions and strengths in the observed field. The typical steps followed by such a method are:

1. Initializing the residual map to the dirty map, with an empty clean component list;

2. Identifying the pixels with the brightest intensity in the residual map, adding them to the clean component list, multiplied by the *loop gain* $\gamma \leq 1$[2];

3. Multiplying the clean components for the dirty beam, and subtracting them from the residuals;

---

[2] This factor ranges from 0 to 1, where 0 means that no flux subtraction and deconvolution will occur (increasing indefinitely the number of iterations required to converge), and 1 means that all the flux of a pixel is subtracted from the residual map (possibly leading to "holes" in the image). Typically, it assumes values between 0.1 and 0.3, meaning that, at each iteration, 10-30% of the peak flux value is subtracted from the residuals.

4. Iterating points 2 and 3 until some stopping criteria are met (e.g., when the peak flux in the residuals is lower than some user-provided noise level);

5. Convolving the clean components with the so-called *clean beam*, a suitably version of the sampling function (tipically a Gaussian);

6. Producing the clean map by adding the clean components to the final residual map.

Furthermore, it is often worth introducing a weighting function $W(u,v)$ to modify the weight of the visibilities during the imaging process, with the aim of improving the sensitivity or accounting for noise variations between samples. By default, a *natural weighting* is considered, with $W(u,v) = 1/\sigma^2(u,v)$, where σ is the noise variance of the visibilities. This choice provides the highest signal-to-noise ratio (SNR), but produces images with poor angular resolution as it attributes more weight to shortest baselines. *Uniform weighting* is inversely proportional to the density of visibilities in a uniform region of the uv-plane. In this way, it provides more weights to long baselines enhancing the angular resolution, but at the same time it increases the root mean square (rms) image noise, reducing the sensitivity. Finally, the *robust Briggs weighting* represents a hybrid form of the previous two weightings. It is ruled by the *robust* parameters, which takes values between -2 (close to uniform weighting) to +2 (close to natural weighting), allowing to find a compromise between angular resolution and point source sensitivity.

1.3 OUTLINE OF THE THESIS

The investigation of the processes that led to the formation and evolution of galaxies through the cosmic epochs is still an active field of modern astronomy. Indeed, the major result emerging from the previous sections is that we are lacking a huge piece of information regarding the interplay between dust, gas, and star formation in the distant Universe, which are fundamental ingredients for the assembly of galaxies.

With this thesis, we are taking part in the completion of the intricate puzzle of the history of the Universe. We took advantage of the ALMA prowess to characterize a large sample of galaxies at the end of the Reionization epoch, representative of the overall SFG population at $z \sim 5$ and part of the ALPINE large program, acquiring information on their dust and gas content, and on the mechanisms driving their growth.

Specifically, this thesis is structured in the following way:

In Chapter 2, we present the ALPINE survey in a nutshell, providing an introduction to the data we used for our analysis. We also make a brief review of the results achieved so far from the entire collaboration, highlighting the relevance of the [CII] line in the study of the high-redshift Universe.

In Chapter 3, we discuss the nature of a galaxy serendipitously discovered in the field of one of the ALPINE main targets. The peculiarity of this source is the presence of a large amount of dust which makes it invisible in the optical bands, posing questions on the origin of the observed emission line. Through our analysis, we found evidence for the line being [CII], emphasizing the importance of looking at this kind of objects as they can contribute significantly to the total budget of star formation at early times.

In Chapter 4, we use the previous analysis as a benchmark for the characterization of the other serendipitous objects with no optical counterpart observed in ALPINE. Being the emission lines of these sources fainter than that observed from the galaxy of Chapter 3, their interpretations result to be less secure, but still of relevant importance for possible future spectroscopic follow-up studies.

Chapter 5 is dedicated to the contribution of the major mergers observed in ALPINE to the galaxy mass-assembly at $z \sim 5$ and through cosmic time. We used for the first time the [CII] emission to recover the fraction of major mergers in the early Universe, and studied their relevance at different cosmic epochs assuming diverse merging timescales. We found that the process of *in-situ* star formation seems to be the dominant mode of galaxy mass-assembly at all times, although mergers may have contributed significantly under some specific assumptions.

In Chapter 6, we deepen the relation between [CII] and star formation obtained with the new ALPINE data by studying in detail the properties of the [CII] non-detections.

We found that, by including the contribution of these sources to the sample of [CII]-detected galaxies, the [CII]-SFR relation is in agreement with that derived in the local Universe, suggesting that this atomic transition is still a good tracer of star formation at early epochs.

Finally, in Chapter 7 we provide our concluding remarks of the thesis and future perspectives.

2 | ALPINE: THE ALMA LARGE PROGRAM TO INVESTIGATE [CII] AT EARLY TIMES

ALPINE is an ALMA Large Program designed to detect the [CII] line at 158 μm rest-frame and the surrounding FIR continuum emission from a sample of 118 SFGs at $4.4 < z < 5.9$, avoiding the redshift range $4.6 < z < 5.1$ due to a low-transmission atmospheric window (Le Fèvre et al. 2020; Béthermin et al. 2020; Faisst et al. 2020b). The campaign took $\sim$ 70 hours of observation with ALMA in Band 7 (275-373 GHz) during cycles 5 and 6. The ALPINE targets are drawn from the Cosmic Evolution Survey (COSMOS; Scoville et al. 2007a,c) and Extended Chandra Deep Field South (E-CDFS; Giavalisco et al. 2004; Cardamone et al. 2010) fields and have secure spectroscopic redshifts from the VIMOS UltraDeep Survey (VUDS; Le Fèvre et al. 2015; Tasca et al. 2017) and DEIMOS 10K Spectroscopic Survey (Hasinger et al. 2018). The rest-frame UV selection ($\mathrm{L_{UV}} > 0.6\ \mathrm{L}^*$, i.e. $M_{1500} \leq -20.2$ mag) ensures that these sources lie on the so-called main-sequence of SFGs, a tight correlation between the SFR and the stellar mass (M_*; e.g., Noeske et al. 2007; Rodighiero et al. 2011; Tasca et al. 2015), being thus representative of the underlying star-forming, UV-detected galaxy population at $z \sim 5$ (e.g., Speagle et al. 2014). Due to the location of the ALPINE targets in these well-studied fields, archival multi-wavelength data are available from the UV to the NIR (e.g., Koekemoer et al. 2007; Sanders et al. 2007; McCracken et al. 2012; Guo et al. 2013; Nayyeri et al. 2017), reaching the X-ray and radio bands (e.g., Hasinger et al. 2007; Smolčić et al. 2017), allowing us to recover physical quantities such as stellar masses ($9 \lesssim \log(\mathrm{M_*/M_\odot}) \lesssim 11$) and star-formation rates ($1 \lesssim \log(\mathrm{SFR}/\mathrm{M}_\odot\ \mathrm{yr}^{-1}) \lesssim 3$) through spectral energy distribution (SED)-fitting (Faisst et al. 2020b; see Section 2.1.2). Moreover, the wide area covered by ALPINE, allowed us to detect serendipitous sources for free in the fields of observations (i.e., $\sim 25\ \mathrm{arcmin}^2$) of the main targets (Gruppioni et al. 2020; Loiacono et al. 2021). 14 serendipitous line detections were robustly observed over the entire ALPINE pointings. Among these, 8 sources were classified as [CII] emitters based on the available photometric or spectroscopic redshifts of the optical/NIR counterparts, 2 lines as CO transitions, and the remaining 4 detections could not be unambiguously identified because of the lack of an optical/NIR component (Romano et al. 2020; Loiacono et al. 2021; see also Chapter 3 and Chapter 4). On the other hand, Gruppioni et al. (2020) analyzed a sample of 56 serendipitous sources detected in continuum in ALPINE and used them to derive the total infrared luminosity function and to quantify the cosmic SFRD up to $z \sim 6$.

2.1 DATA AND OBSERVATIONS

2.1.1 ALMA DATA

The collected ALPINE-ALMA data were reduced and calibrated using the Common Astronomy Software Applications (CASA; McMullin et al. 2007) software. Each data cube was continuum-subtracted in the uv-plane by masking the channels containing line emission, then fitting a flat continuum model to the visibilities. These channels are within 3σ from the central frequency of the best Gaussian fit of the identified line, obtained through an iterative line finder algorithm aimed to search for [CII] emission in a cylinder of 1 arcsec aperture around the phase center of the cube, covering the full bandwidth to allow for spatial and spectral offsets. In case the [CII] spectrum presents a non-Gaussian excess in the wings, the line channels are extended of an additional $\sim 0.1-0.2$ GHz interval to obtain conservative continuum-free cubes (see Section 6 of Béthermin et al. 2020 for further details). This process returned line-only cubes with channel width varying from 25 km s^{-1} to 35 km s^{-1} and average beam size $0.85'' \times 1.13''$ (with a pixel scale of $\sim 0.15''$; Béthermin et al. 2020). These cubes were inspected searching for SNR > 3.5 emission (corresponding to 95% purity of the sample) in the 1 arcsec aperture around the phase center and at the expected frequency from the UV

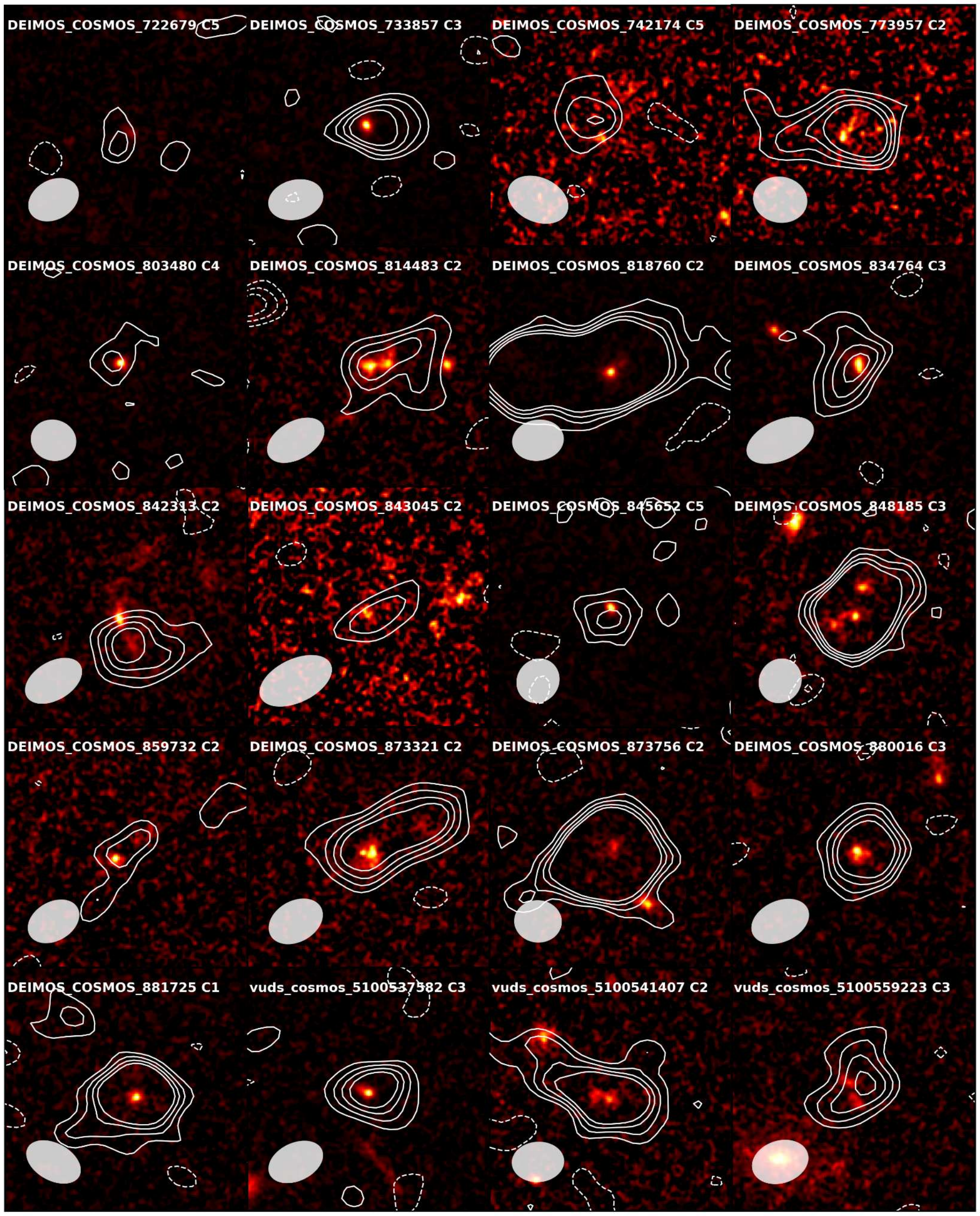


Figure 2.1: Example of [CII] integrated intensity maps for some of the ALPINE targets. White contours report the ALMA emission. In the background, HST F814W images (Koekemoer et al. 2007, 2011) show the optical emission of the sources. Each image is $5'' \times 5''$ wide and it is centered on the UV rest-frame position of the target. The filled ellipse in the bottom left corner represents the ALMA synthesized beam. The name of each target is reported at the top of each image, along with its morpho-kinematic classification (see Section 2.2). Figure from Le Fèvre et al. (2020).

spectroscopy (Faisst et al. 2020b), resulting in 75 [CII] detections out of 118 ALPINE targets (of which 23 sources were also detected in continuum). For each data cube, the map of the integrated [CII] intensity (i.e., *moment-0* map) was produced by summing all the spectral channels containing the emission line as

$$M(x, y) = \sum_{k=1}^{N_{channel}} S_v(x, y, k)\ \Delta v_{channel}(k), \tag{2.1}$$

where $S_v(x, y, k)$ is the flux density in the k-th channel at the position (x, y), while $\Delta v_{channel}(k)$ is the velocity width of the same channel. The mean reached sensitivity is 0.14 Jy km s^{-1} for the line, and 39 μJy/beam for the continuum.

Figure 2.1 shows an example of [CII] moment-0 maps for some of the ALPINE targets compared to the underlying optical emission as observed with HST (Le Fèvre et al. 2020). More information on the data cubes reduction can be found in Béthermin et al. (2020).

2.1.2 ARCHIVAL DATA

As introduced above, the ALPINE galaxies are distributed in the COSMOS and E-CDFS fields. Consequently, they can rely on a wealth of ancillary data spanning the whole electromagnetic spectrum. Indeed, the two fields are matched by the COSMOS2015 (Laigle et al. 2016) and 3D-HST (Brammer et al. 2012; Skelton et al. 2014) catalogs, respectively, providing a wide photometric information. In particular, HST F814W i-band imaging (Scoville et al. 2007b; Koekemoer et al. 2007), Subaru and Very Large Telescope (VLT) optical imaging (Taniguchi et al. 2007; Nonino et al. 2009; Taniguchi et al. 2015; Laigle et al. 2016), and NIR imaging from the UltraVISTA (UVista) Survey (McCracken et al. 2012) and the Canada-France-Hawaii Telescope (CFHT; Hsieh et al. 2012) are available, as well as Spitzer imaging (Sanders et al. 2007; Capak et al. 2012; Steinhardt et al. 2014). X-ray coverage comes from the XMM-Newton (Hasinger et al. 2007) and Chandra (Elvis et al. 2009; Civano et al. 2016) telescopes, while at the longest wavelengths radio emission is observed with the VLA at 3 GHz (Smolčić et al. 2017). More details can be found in Faisst et al. (2020b).

By combining the already available multi-wavelength data with the new constraints from ALMA, Faisst et al. (2020b) provided several physical measurements of the ALPINE

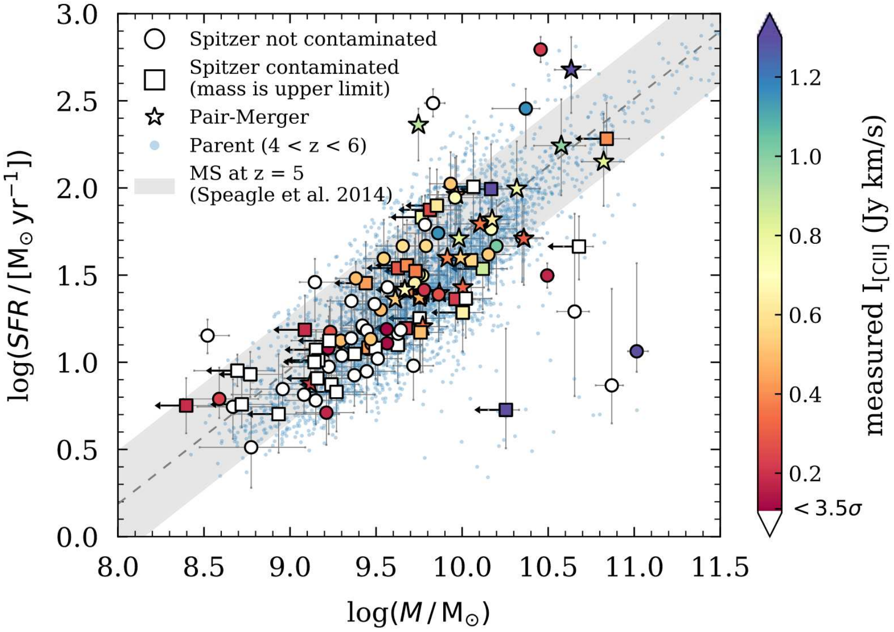


Figure 2.2: Relation between SFR and stellar mass for the ALPINE galaxies (large markers) and the COSMOS galaxies (small blue points) at $4 < z < 6$. ALPINE detections are color-coded for the measured [CII] flux in Jy km s^{-1}, while non-detections are shown in white. Galaxies with blended photometry in the Spitzer images (possibly leading to an overestimation of the stellar mass) are marked as squares (with upper limit on their stellar mass), while stars indicate mergers as classified by Le Fèvre et al. (2020). The gray dashed-line and band (±0.3 dex wide) display the main-sequence parameterizations by Speagle et al. (2014) at $z \sim 5$. Figure from Faisst et al. (2020b).

targets, including SFRs and stellar masses, through SED-fitting with the LePHARE fitting code (Arnouts et al. 1999; Ilbert et al. 2006). Figure 2.2 reports the relation between SFR and stellar mass obtained by Faisst et al. (2020b) for the ALPINE sample, showing how these galaxies lay along the main-sequence at $z \sim 5$ (Speagle et al. 2014).

2.2 Morpho-kinematic diversity in ALPINE

Figure 2.1 displays the large diversity in [CII] emission observed in the ALPINE sample. Some sources show extended morphology, others are more compact, and many report signs of merging activity as also visible from the underlying HST images.

Le Fèvre et al. (2020) performed a preliminary qualitative morphological and kinematic classification of the 75 ALPINE targets showing $> 3.5\sigma$ [CII] emission. In particular, the channel maps around the emission line, the intensity (moment-0) and velocity (moment-1) maps, the integrated [CII] spectra, the position-velocity diagrams (PVDs) along the major and minor axes of the velocity maps, and the multi-wavelength data at the position of the source were visually inspected by an internal team of the ALPINE collaboration that placed these galaxies in the following classes:

1. *Rotators* (Class 1). The galaxies in this class are characterized by a single source in the ancillary data and show a clear velocity gradient in the moment-1 maps which then implies a tilted (straight) PVD along the major (minor) axis. They can also exhibit a double-horned profile in the integrated spectra.

2. *Mergers* (Class 2). In case of interacting systems the presence of two or more components in the optical images within the ALMA field of view and/or in the moment-0 maps and PVDs is expected. A complex behavior in the channel maps and in the [CII] spectra could be visible.

3. *Extended Dispersion Dominated* (Class 3). These sources extend in [CII] emission over multiple ALMA beams and are typically characterized by straight PVDs along the major and minor axis and by a Gaussian line profile.

4. *Compact Dispersion Dominated* (Class 4). As opposed to the class 3 galaxies, these objects are unresolved in the moment-0 maps.

5. *Weak* (Class 5). These targets are too weak to be visually classified into one of the above classes.

In particular, following e.g., Epinat et al. (2012), all the people involved in the classification analyzed at first the above information individually, providing an initial statistics from which the mode of each object was estimated. Then, the final classification was refined through another iteration with the objects presenting the most disagreement between the classifiers. In the example reported in Figure 2.1, the class number is shown at the top right of each panel, while the final morpho-kinematic distribution obtained is presented in Figure 2.3. Le Fèvre et al. (2020) found that $\sim 40\%$ of the ALPINE galaxies

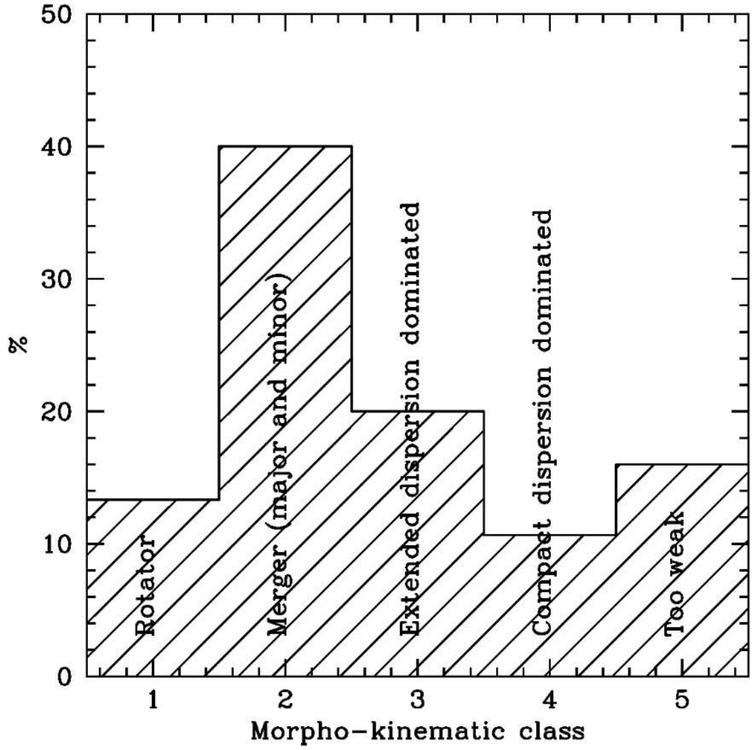


Figure 2.3: Morpho-kinematic distribution of the ALPINE [CII]-detected galaxies. Figure from Le Fèvre et al. (2020).

are likely undergoing a merging phase and that $\sim 13\%$ of the targets are possibly forming an ordered disk at $z \sim 5$. Extended and compact sources make up 20% and 11% of the sample, respectively, with the remaining galaxies (16%) being too complex or weak to be classified. The different morphology and kinematics of these normal galaxies in the early Universe suggests that many physical processes altering the evolution of these sources were already in place between 0.9 and 1.4 Gyr after the Big Bang. In particular, the high fraction of mergers (both minor and major) could indicate that galaxy mass-assembly through merging is frequent at these epochs for main-sequence SFGs (see Chapter 5).

On the other hand, in Jones et al. (2021) we expanded this initial classification of the ALPINE targets by adopting a more quantitative analysis based on the tilted ring fitting code ^{3D}Barolo (3D-Based Analysis of Rotating Objects from Line Observations; Di Teodoro & Fraternali 2015), a morphological classification, and a set of disk identification criteria. In particular, among the 75 [CII]-detected galaxies, 29 (representative of the high stellar mass region of ALPINE) were successfully fit with ^{3D}Barolo retrieving morpho-kinematic parameters like inclination, position angle (PA), velocity dispersion profile, rotation curve, and systemic velocity. Figure 2.4 shows an example of fit results for different classes of galaxies, comparing data, models and residuals.

Besides, two other quantitative classification methods were applied to these sources, i.e., the Gini-M_{20} method by Lotz et al. (2004) and the five disk-like criteria by Wisnioski et al. (2015). Although these criteria were not designed for the characterization of high-z SFGs, we used them to test their applicability to the ALPINE data. Briefly, the Gini-M_{20} criterion was adopted in the local Universe to distinguish between different morphological types of galaxies. The Gini coefficient is a measure of the brightness distribution in a galaxy and it is low (high) for diffuse (concentrated) emission. M_{20} is instead the second-order moment of the pixels making up the brightest 20% of the galaxy, resulting in large values when non-axisymmetric features are present. At low redshifts, Lotz et al. (2008) found that different galaxies are displaced in different regions of the diagram, successfully classifying them depending on their morphology. However, in the case of the ALPINE galaxies, such analysis did not lead to a clear and net separation of the morphological type of these galaxies, mostly because of the low resolution of our observations. Nonetheless, this was one of the first applications of this method to a statistical sample of FIR observations of $z > 4$ SFGs and will be improved with future high-resolution data. On the other hand, the five disk identification criteria of Wisnioski et al. (2015) were also exploited to obtain additional information on the morpho-kinematics of each galaxy. These criteria can be summarized as:

- The velocity field shows a velocity gradient along a single axis with a significant ($> 3\sigma$) slope;
- $v_{rot}/\sigma_0 > 1$ across all rings, where v_{rot} is the rotational velocity, and σ_0 is the velocity dispersion;
- The midpoint between the extreme values of the velocity along the kinematic axis is coincident, within one beam width, with the peak of the velocity dispersion map;
- There is a good agreement ($< 30^\circ$) between the morphological and kinematic PAs;

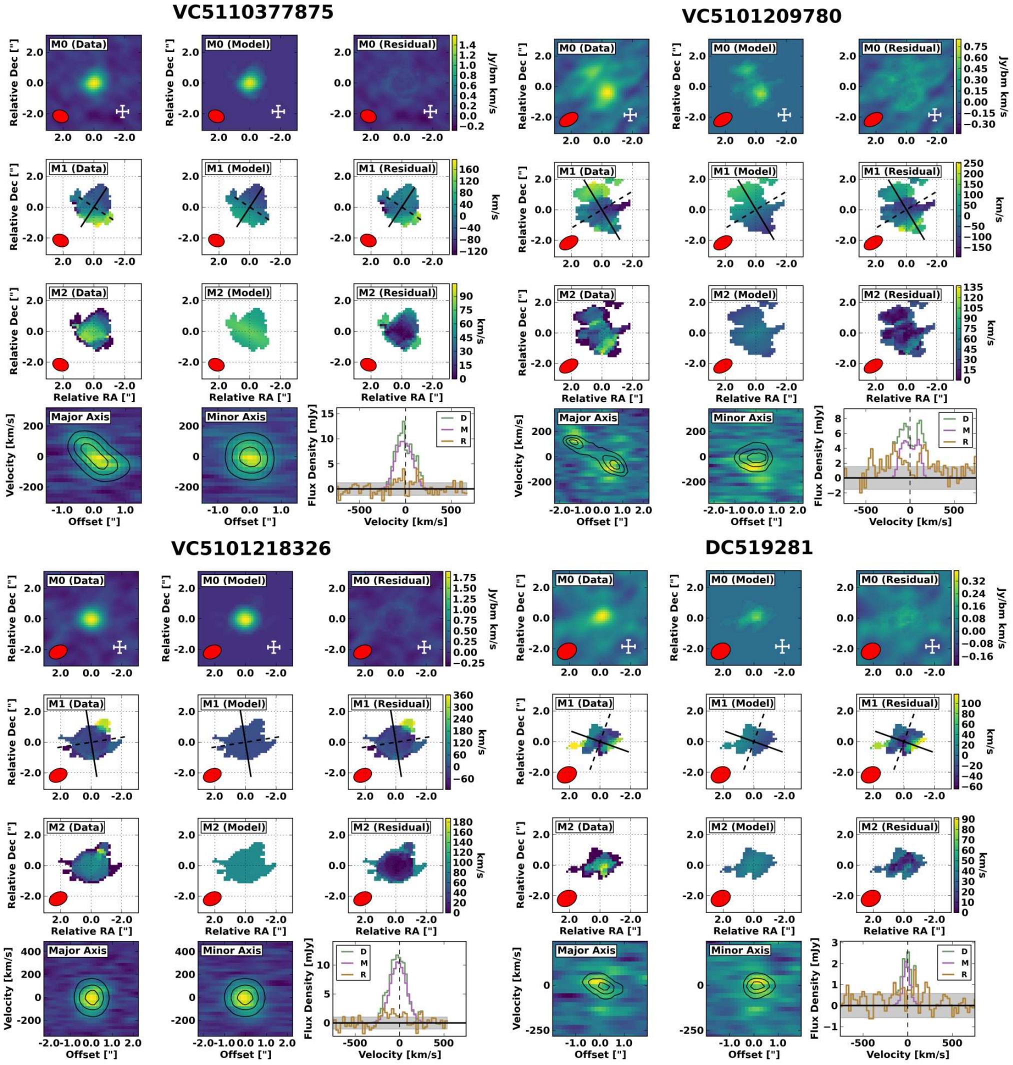


Figure 2.4: Example of ^{3D}Barolo fit for four different types of galaxies: rotator (upper left), merger (upper right), dispersion dominated (lower left), and uncertain (lower right). Along each column, moment maps, PVDs and spectra are shown for data, model and residual. First three rows of each figure report the integrated intensity (moment-0), velocity field (moment-1), and velocity-dispersion field (moment-2) maps. The white cross at the bottom right of the first rows show a 5 kpc × 5 kpc physical scale, while the red ellipse on the left represents the synthesized beam. In the second row, the kinematic major (solid) and minor (dashed) axis are also shown. For the PVDs, the background colors represent the observed data, while the solid contours come from the model at 20%, 50%, and 80% of its maximum value. Data, model, and residual spectra are represented by green, purple, and orange lines, respectively. The shaded gray area is the 1σ uncertainty. Figure from Jones et al. (2021).

- The midpoint between the extreme values of the velocity along the kinematic axis is coincident, within one beam width, with the peak of the intensity map.

In principle, if all the criteria are met, the source is most likely a rotating galaxy. On the contrary, if none of the criteria is satisfied, the source is likely a merger or it has a disturbed morphology. Unfortunately, the resolution and sensitivity of the ALPINE data were insufficient to draw conclusions on the morphological type of these galaxies from this analysis alone.

However, by combining the results from those criteria and the ^{3D}Barolo fit, we were able to robustly classify 14 galaxies, with the remaining marked as "uncertain" (i.e., objects unable to be classified due to low SNR, low spectral resolution, and/or conflicting evidence from our combination of analysis). These 14 sources include 6 rotators, 5 mergers, and 3 dispersion-dominated galaxies (making the 43%, 36% and 21% of the sample). This statistics results in a bit different morpho-kinematic distribution with respect to the qualitative one obtained by Le Fèvre et al. (2020), but highlights the large diversity of SFGs at $z \sim 5$ even in this more massive ALPINE sub-sample.

In the following, I will present some of the outstanding results obtained through the ALPINE survey so far.

2.3 Scientific results overview

ALPINE represents the first, statistically significant sample of [CII]-detected normal galaxies with panchromatic observations between the end of the Reionization epoch and the cosmic noon. This makes the ALPINE one a golden sample to study the properties of the Universe during this epoch, which directly links primordial galaxy formation with the more mature galaxy growth at the peak of the cosmic SFRD. Thanks to these properties, outstanding results have been obtained so far through the ALPINE data.

Regarding the ISM properties of these galaxies, Schaerer et al. (2020) studied the [CII]-SFR relation over the redshift range $4 < z < 8$, finding no evolution with respect to the local relation by De Looze et al. (2014), with a bit steeper slope at high-z when [CII] non-detections are taken into account. Romano et al. (submitt.) confirmed the previous little evolution found by Schaerer et al. (2020) by analyzing the [CII] non-detections through stacking, and exploiting the results by Cassata et al. (2020) on the observed Lyα-[CII] velocity offset in these SFGs (see Chapter 6). Ginolfi et al. (2020b) investigated the connection between star formation-driven outflows and the chemical enrichment of the circumgalactic medium (CGM), suggesting that the baryon cycle and gas exchanges with the CGM were already at work in $z \sim 5$ SFGs. These outflows could also be responsible for the presence of ubiquitous extended [CII] halos around these sources, as suggested by a study of the [CII] size of the ALPINE targets (Fujimoto et al. 2020). Fudamoto et al. (2020) used both individual FIR continuum observations of the ALPINE galaxies and stacking analysis to deepen the dust attenuation properties at $z \sim 5$. As major results, they found indication for a rapid build-up of dust during the Reionization epoch in massive UV-selected SFGs, which are also characterized by steeper attenuation curves than at $z < 4$. The content of dust in these galaxies was later constrained by Pozzi et al. (2021). By comparing the derived dust scaling relations with predictions from chemical evolution models, they claimed that the ALPINE sources are most likely the progenitors of present-day elliptical rather than disc galaxies, also finding that UV selection misses the most dust-obscured objects at $z \sim 5$. Dessauges-Zavadsky et al. (2020) used instead the [CII] line as a molecular gas tracer, in order to estimate the molecular gas budget in the early Universe. Furthermore, in-depth studies of peculiar mergers and very dust-obscured sources were carried out by Ginolfi et al. (2020b), Jones et al. (2020), and Romano et al. (2020) (see Chapter 3). Morphological and kinematic analysis of the ALPINE targets was at first qualitatively addressed by Le Fèvre et al. (2020), and the quantitatively by Jones et al. (2021) and Romano et al. (2021) (see Section 2.2 and Chapter 5 for more details). Statistical studies were made with the UV-selected ALPINE galaxies to estimate the cosmic SFRD (Khusanova et al. 2021) and their luminosity function (Yan et al. 2020). On the other hand, Loiacono et al. (2021) built the luminosity function of [CII] line emitters discovered serendipitously in the ALPINE pointings, while Gruppioni et al. (2020) used the continuum serendipitous detections to obtain the infrared luminosity function at $z \sim 5$.[1]

---

[1] See the arXiv for a complete list of the ALPINE papers and more details.

3 | THE NATURE OF AN EXTREMELY OBSCURED HIGH-REDSHIFT GALAXY

Based on:
Romano M., Cassata P., Morselli L., et al. 2020, MNRAS, Volume 496, pg. 875-887

In this chapter, we report the serendipitous discovery of a dust-obscured galaxy observed as part of the ALPINE program. While this galaxy is detected both in line and continuum emissions in ALMA Band 7, it is completely dark in the observed optical/NIR bands and only shows a significant detection in the UVista K_s band. We discuss the nature of the observed ALMA line, i.e., [CII] at $z \sim 4.6$ or CO at $z \sim 2.2$. At both redshifts, we derive the SFR from the ALMA continuum and the physical parameters of the galaxy, such as the stellar mass, by fitting its SED. Exploiting the results of this work, we believe that our source is a main-sequence, dusty SFG at $z = 4.6$ (i.e. [CII] emitter) with $\log(\mathrm{SFR}/\mathrm{M}_\odot\ \mathrm{yr}^{-1}) \sim 1.4$ and $\log(\mathrm{M}_*/\mathrm{M}_\odot) \sim 9.9$. As a support to this scenario our galaxy, if at this redshift, lies in a massive proto-cluster recently discovered at $z \sim 4.57$, at only ~ 1 proper Mpc from its center. This work underlines the crucial role of the ALPINE survey in making a census of this class of objects, in order to unveil their contribution to the global SFRD at the end of the Reionization epoch.

3.1 General context

At $z > 3$, the cosmic SFRD is almost exclusively constrained by UV-selected samples (Bouwens et al. 2012a,b; Schenker et al. 2013; Oesch et al. 2015), lacking information about the star formation obscured by the dust (see Section 1). Rest-frame UV-selected galaxies must be corrected for the dust absorption: wrong dust corrections can lead to large uncertainties on the SFR estimates and, consequently, to an incorrect picture of the SFH of the Universe (e.g. Gallerani et al. 2010; Castellano et al. 2014; Scoville et al. 2015; Álvarez-Márquez et al. 2016). At the same time, heavily dust-obscured SFGs may be completely missed by surveys probing the rest-frame UV/optical emissions.

With the advent of new facilities, a population of faint and dusty SFGs has been confirmed at high redshift, e.g. SMGs (Dunlop et al. 2004; Daddi et al. 2009; Riechers et al. 2010; Weiß et al. 2013; Huang et al. 2014; Simpson et al. 2014; Santini et al. 2016; Cooke et al. 2018), ALMA-only sources (e.g. Williams et al. 2019), the extremely red objects selected with H and IRAC colors (HIEROs galaxies) from Wang et al. (2016). While the bulk of these objects peaks at $2 < z < 3$, a significant tail of dusty galaxies without optical/NIR detections appears to be in place at $z > 4$ (Capak et al. 2008; Daddi et al. 2009; Riechers et al. 2010; Walter et al. 2012; Riechers et al. 2013; Simpson et al. 2014; Cooke et al. 2018; Pavesi et al. 2018; Dudzevičiūtė et al. 2020).

In particular, a large population of high-redshift SMGs has been discovered during the last decade. For instance, Walter et al. (2012) combined measurements from the IRAM Plateau de Bure Interferometer (PdBI) and the VLA to spectroscopically identify the dust-obscured starburst HDF850.1, one of the brightest sources with no optical/NIR counterpart observed at 850 μm in the Hubble Deep Field (Hughes et al. 1998). By using millimeter wave molecular line scan, this source was placed at $z_{\rm spec} = 5.18$ among an overdensity of galaxies at the same redshift, with a [CII]/FIR luminosity ratio comparable to that observed in local SFGs. Moreover, many of these objects are extreme starbursts, such as HFLS3. This source is confirmed to be at $z_{\rm spec} = 6.34$ exploiting information from different molecular and atomic fine structure cooling lines and shows a large FIR luminosity (i.e. $L_{\rm FIR} \sim 2 \times 10^{13}\ L_\odot$) and SFR $> 10^3\ M_\odot\ {\rm yr}^{-1}$ (Riechers et al. 2013). Simpson et al. (2014) analyzed a sample of 96 SMGs from the ALMA-LABOCA *Chandra* Deep Field-South Survey (ALESS; Hodge et al. 2013), finding a median photometric redshift of $z_{\rm phot} \sim 2.5$, with 35% of the SMGs lying at $z > 3$ (and $\sim$ 20 sources most likely at $z > 4$). Cooke et al. (2018) examined the ALMA data cubes of $\sim$ 700 high-redshift SMGs from the AS2UDS survey, searching for serendipitous emission lines. They found 10 candidate line emitters, 8 of which are likely [CII] emitters at $z \sim 4.5$ and with the remaining two sources associated with high-J CO emission. Dudzevičiūtė et al. (2020) found that $\sim$ 17% of the SMGs in the AS2UDS survey are undetected in the observed optical/NIR bands and fainter than $K \sim 25.7$ mag. These sources preferentially lie at higher redshifts ($z > 3$) with respect to the median photometric redshift of the sample, i.e. $z_{\rm phot} \sim 2.6$. An in-depth study of this elusive population of galaxies is necessary in order to complete the census of the SFGs at high redshift contributing to the cosmic SFH as well as to better understand the early phases of the galaxy formation (e.g. Blain et al. 2002; Casey et al. 2014).

The ALPINE survey combined FIR continuum and UV measurements, together with the [CII] observations, to provide an estimate of the total (obscured and unobscured) star formation in a large sample of high-z SFGs. The rest-frame UV selection allows ALPINE to account for 80-95% of the total flux density at $z > 4$ emitted by galaxies directly in the UV, or in the FIR after dust reprocessing (Bouwens et al. 2009; Casey et al. 2012; Bouwens et al. 2016; Capak et al. 2015; Aravena et al. 2016; Novak et al. 2017; see Figure 1 of Faisst et al. 2020a). The remaining 5-20% of the missing star formation at these redshifts is yielded by a free blind survey covering an additional area of 25 arcmin^2 beyond the rest-frame UV-selected ALPINE targets, where many galaxies have been serendipitously detected so far (Loiacono et al. 2021). Among these, different sources are completely obscured in the observed optical/NIR bands. The study of these objects is crucial for obtaining a robust estimate of the total SFRD at $z > 4$ and for characterizing the overall population of high-redshift SFGs.

In the following, we discuss the nature of a galaxy (hereafter, Gal−A) randomly discovered in the field of the ALPINE target DEIMOS_COSMOS_665626 (DC_665626). The galaxy has a spatial offset of $\sim$ 6 arcsec (1 arcsec is $\sim$ 7 kpc at $z = 4.583$, the redshift of the target) from DC_665626. It does not show any optical counterpart at the position of the emission detected with ALMA and, for this reason, its nature results to be ambiguous. Besides, since Gal−A is the brightest galaxy detected in line emission among all those having no optical counterpart and serendipitously observed in ALPINE so far

(Loiacono et al. 2021), this work can be exploited as a benchmark for future analysis on this type of source. A similar analysis on the remaining ALPINE dust-obscured galaxies with no optical counterpart is presented in Chapter 4.

Throughout this work and for the rest of the thesis, we assume a Λ-CDM cosmology with $\mathrm{H}_0 = 70$ km/s/Mpc, $\Omega_\mathrm{m} = 0.3$, and $\Omega_\Lambda = 0.7$. We furthermore use a Chabrier (2003) initial mass function (IMF) and AB magnitudes.

3.2 OBSERVATIONS AND DATA REDUCTION

3.2.1 ALMA DATA

DC_665626 was observed with ALMA in Band 7 ($\nu_\mathrm{obs} = [275 - 373]$ GHz) on 25 May 2018 (Cycle 5; Project 2017.1.00428.L, PI O. Le Fèvre) using 45 antennas with the C43-2 array configuration (with a maximum baseline of ~ 250 m). The on-source integration time was 16 minutes, with a total elapsed time of 37 minutes (see Figure 1.5 in Section 1.2.1.1). The spectral setup consisted of two sidebands with a frequency range of $\Delta_\nu^\mathrm{l} \simeq [339 - 343]$ GHz and $\Delta_\nu^\mathrm{u} \simeq [351 - 355]$ GHz for the lower and upper sidebands, respectively. Both sidebands were made up of two spectral windows (SPWs) of width 1.875 GHz, each of which containing 128 channels 15.625 MHz wide (the sidebands overlapped for 7 channels), with a typical rms of 0.6 mJy beam^{-1} per channel. The flux and phase were calibrated using the standard calibrators J1058+0133 and J0948+0022, respectively.

The data are analyzed using standard pipelines for ALMA reduction included in the software CASA (McMullin et al. 2007), version 5.4.0. The imaging is obtained by running the `TCLEAN` task on the visibilities, setting a threshold of $3\sigma_\mathrm{rms}$ on the noise level when cleaning the data (where σ_rms is obtained from the dirty image), and with a natural weighting scheme to increase the sensitivity (see Section 1.2.1.2).

3.2.2 IDENTIFICATION OF THE SERENDIPITOUS SOURCE

As part of the COSMOS field (Scoville et al. 2007d,b), which is one of the most thoroughly studied regions of the sky so far, multi-wavelength data are available for the

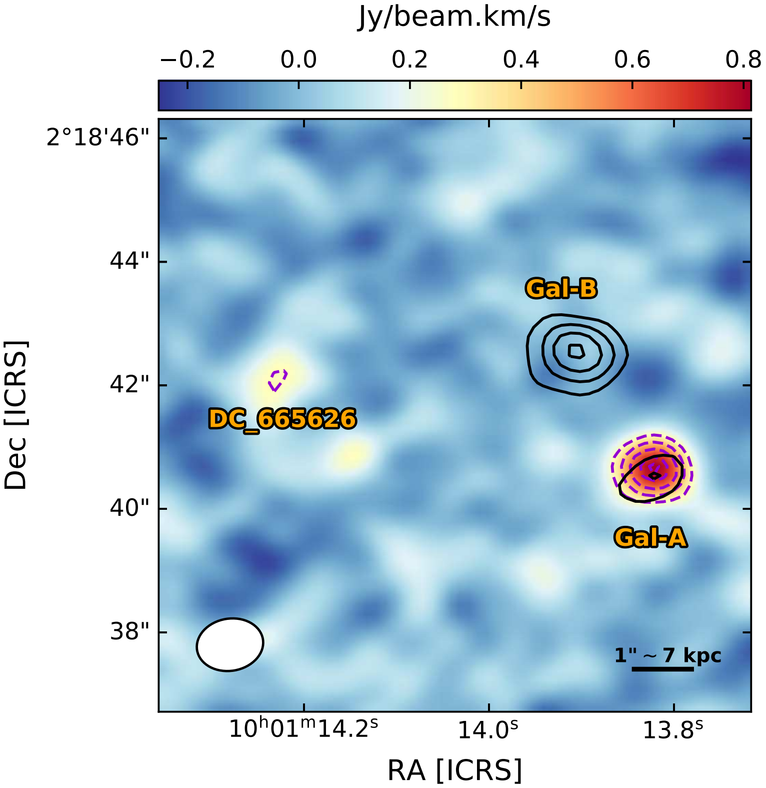


Figure 3.1: Continuum-subtracted moment-0 map of Gal−A. The ALPINE target DC_665626, Gal−A and Gal−B are labeled. Line and continuum emissions are shown with dashed violet and solid black contours starting from 4σ and 3σ (at step of 2σ), respectively. The white ellipse in the bottom left corner is the synthesized beam.

whole ALPINE sample, including high-resolution HST imaging (Koekemoer et al. 2007, 2011) and photometry from the CFHT, the Spitzer telescope and other facilities (Capak et al. 2007; Laigle et al. 2016). Spectroscopic redshifts are available from large optical spectroscopic campaigns at the VLT (VUDS; Le Fèvre et al. 2015) and Keck (DEIMOS 10k survey; Hasinger et al. 2018). Multi-band photometry and spectroscopic data allow us to build SEDs and to derive robust physical parameters including SFRs and stellar masses through SED-fitting (Faisst et al. 2020b). As a result of this analysis, we find that DC_665626 has $\log(\mathrm{M}_*/\mathrm{M}_\odot) = 9.21^{+0.16}_{-0.18}$, $\log(\mathrm{SFR}/[\mathrm{M}_\odot\ \mathrm{yr}^{-1}]) = 0.71^{+0.29}_{-0.18}$, and a spectroscopic redshift of $z_{\mathrm{spec}} = 4.583$, obtained from the Ly$\alpha$ emission and the ISM absorption lines in the observed-frame optical spectrum.

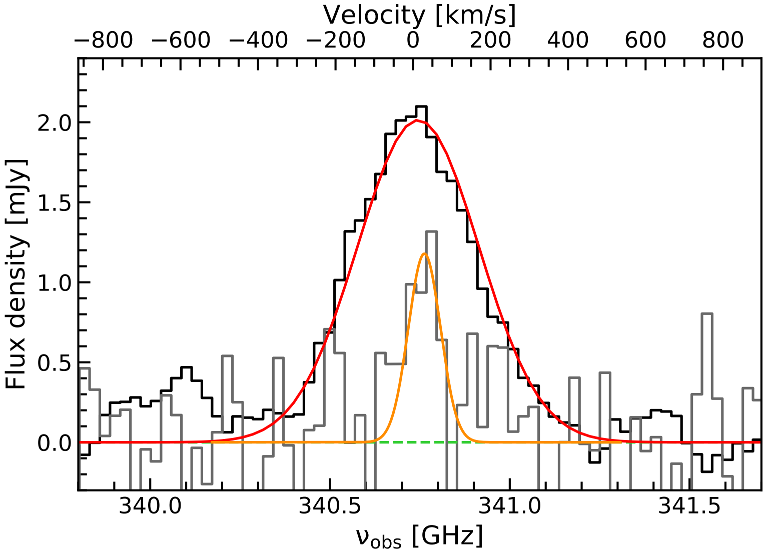


Figure 3.2: Emission line flux at the position of Gal−A (black histogram) as a function of the observed frequency. The solid red curve represents the Gaussian fit on the line. The dashed green line marks the zero-flux level. The velocity offset is reported on the top axis. For comparison, we also show the spectrum of the [CII] line arising from DC_665626 (gray histogram) with its associated Gaussian fit (solid orange curve).

Since the ALPINE target DC_665626 is at $z_{\rm spec}$ = 4.583, the [CII] emission from this source ($\nu_{\rm rest}$ = 1900.54 GHz) is expected to be redshifted at around $\nu_{\rm obs}$ = 340.42 GHz, falling inside the lower sideband of the observed ALMA spectrum. When we inspect the data cube, together with the [CII] emission coming from DC_665626 (at 4.4σ; Béthermin et al. 2020), we identify a more significant line emission feature with a spatial offset of $\sim$ 6 arcsec ($\sim$ 40 proper kpc at $z \sim 4.6$) with respect to the ALPINE target. We refer to the source of this emission as Gal−A (RA: 10:01:13.82, Dec: +02:18:40.66), that is detected both in continuum and in line emissions at 5σ and 12σ, respectively. Figure 3.1 shows the continuum-subtracted moment-0 map of Gal−A. This is computed by summing the integrated intensity in all the spectral channels in the data cube containing the line emission, following Equation 2.1 (see Section 3.2.4 for more details). The synthesized beam with a size of 1.08” × 0.85” at PA = - 80° and another galaxy (hereafter, Gal−B) detected at 9σ in continuum only northwards of the offset emission ($\sim$ 2 arcsec away from Gal−A when considering the peak positions of the two emissions) are also displayed.

We show in Figure 3.2 the spectrum of the emission line observed at the position of Gal–A (black histogram). It is extracted from a circular region 2 arcsec wide, including the 2σ contours from the moment-0 map of the source. Using the *spectral profile tool* within the CASA viewer, we fit the line profile with a Gaussian function finding a full width at half maximum ($\mathrm{FWHM_{line}}$) of 308 ± 34 km $\mathrm{s^{-1}}$ and a peak frequency at ν_{peak} = 340.76 GHz. For comparison, we also show the [CII] line spectrum (gray histogram) of the main ALPINE target, obtained in the same way as for Gal–A.

Although DC_665626 is detected in [CII] in spatial coincidence with its rest-frame UV emission, we consider the possibility that the emission centered at the position of Gal–A is connected with that of the ALPINE target. The displacement between [CII] and UV/Lyα emission has already been observed in high-redshift galaxies (Gallerani et al. 2012; Willott et al. 2015; Cassata et al. 2020; but see also Bradač et al. 2017). It is also reproduced by radiative transfer simulations as a consequence of the strong stellar feedback which could quench the [CII] emission in the central region of the galaxies, allowing it to arise mostly from infalling or satellite clumps of neutral gas around them (Vallini et al. 2013; Maiolino et al. 2015). However, these models predict spatial offsets up to $\sim 1-2$ arcsec ($\sim 7-14$ kpc at the redshift of the target), well below the offset that we measure in this case ($\gtrsim$ 6 arcsec). Therefore, we exclude that the observed ALMA emission at the position of Gal–A is directly linked to DC_665626.

3.2.3 Multi-wavelength photometry of Gal-A

We exploit all the multi-wavelength photometry available for Gal–A as part of the COSMOS field (Laigle et al. 2016) to identify its associated counterpart. In Figure 3.3, we present some cutouts centered on this galaxy in different photometric filters, from the UV to the FIR. Gal–B is visible in most of the photometric bands and has a photometric redshift $z_{\mathrm{phot}} = 2.25^{+0.22}_{-0.15}$, as reported in the COSMOS2015 catalog (Laigle et al. 2016). Another foreground galaxy, labeled Gal–C in Figure 3.3, is well detected in the images from the optical to the NIR wavelengths and has $z_{\mathrm{phot}} = 2.02^{+0.12}_{-0.12}$, from COSMOS2015. Conversely, Gal–A is not clearly identified in any optical filter. It is fairly visible in the UVista K_s band, even if it is not listed as a detection in the UVista DR4 catalog (McCracken et al. 2012).

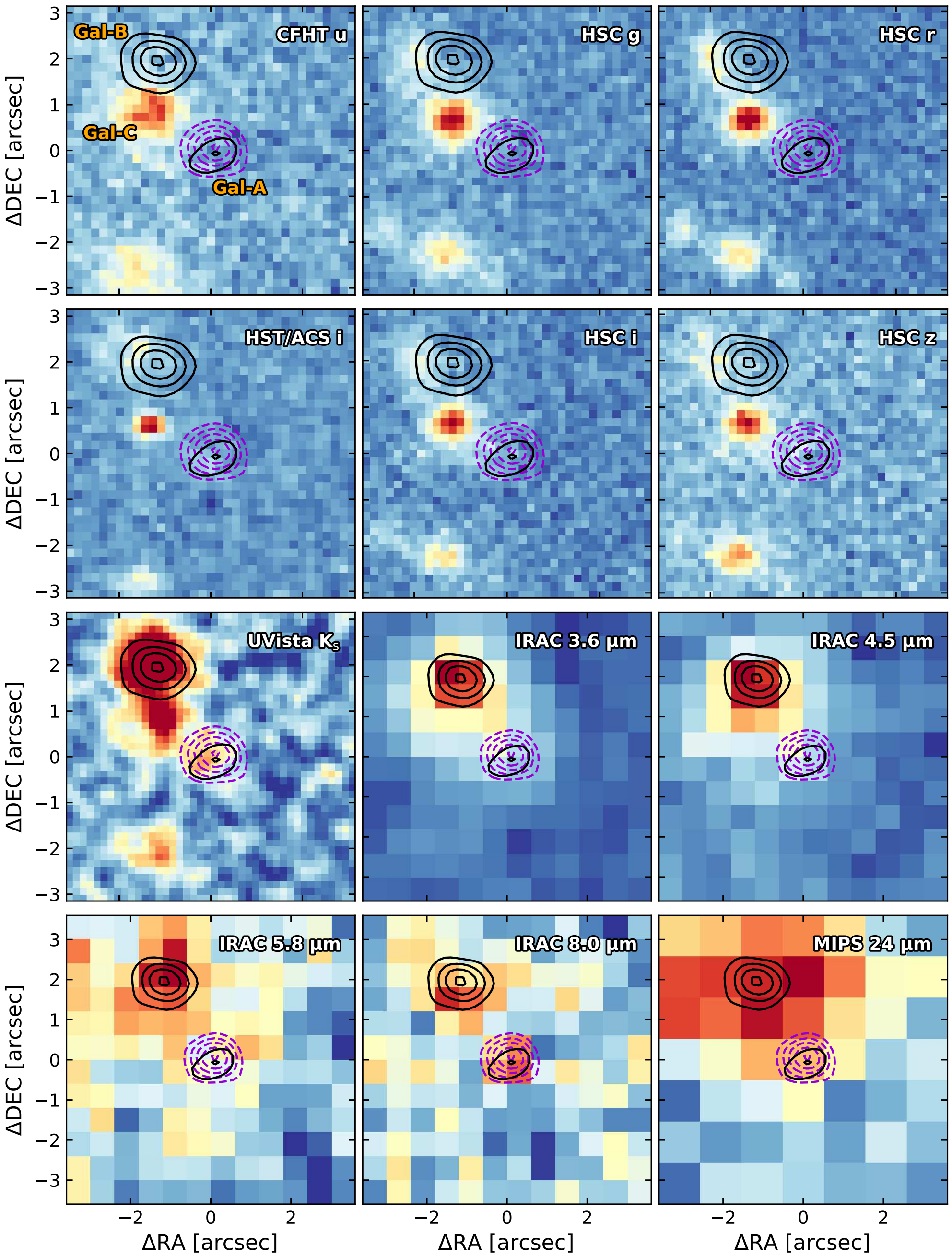


Figure 3.3: Cutouts centered on Gal–A in different photometric filters, from HST/ACS (Koekemoer et al. 2007) and Subaru, UVista and Spitzer (Capak et al. 2007; Laigle et al. 2016). The violet dashed and black solid contours are $> 3\sigma$ line and continuum emissions (at steps of 2σ), respectively. Gal–A, Gal–B, and Gal–C are labeled in the upper-left plot of the figure. Wavelengths increase from the upper-left to the bottom-right corner.

More in detail, to reproduce the SED of Gal–A, we use observations in the u^* band from MegaCam on CFHT, as well as the g, r, i and z filters from the Hyper Suprime-Cam (HSC) on the Subaru telescope, in order to set an upper limit to the optical emission of the source. NIR constraints come from the Y, J, H and K_s bands from VIRCAM on the VISTA telescope. Finally, we obtain information on the SED up to $\sim 8\ \mu$m in the observed-frame from the IRAC channels on Spitzer. For each band, we center a fixed aperture of 1.4 arcsec of diameter on Gal–A (enclosing the 3σ contours of the emission line detected by ALMA) and estimate its flux. We then compute the limiting magnitude as the standard deviation of the fluxes measured in 10000 apertures (of 1.4 arcsec of diameter) randomly distributed in a wide region of the sky (masked for the emission of bright sources) and assume it as the error on the flux. As expected, we do not find any significant detection ($> 5\sigma$) of our source in the optical filters. For this reason, we consider the measured 5σ limiting magnitudes as upper limits in these bands. The same argument applies to the VISTA filters except for the K_s band in which, as mentioned above, a faint emission arises at the position of Gal–A. Making use of `SExtractor` (Bertin & Arnouts 1996), we manage to deblend the analyzed galaxy from the other two nearby sources obtaining an estimate of its apparent magnitude in this band. Through this analysis, Gal–A is detected at $\sim 8.3\sigma$ with an AB magnitude $K_s = (24.4 \pm 0.1)$, which is very close to the 5σ limiting magnitude of ~ 24.5 (computed in 2.0 arcsec diameter apertures) from the UVista DR4 catalog. It is worth noting that in this case we do not assume the uncertainty on the flux resulting from `SExtractor`. We consider instead the corresponding 1σ limiting magnitude as the error on the flux as done for the other bands although, in practice, the two uncertainty estimates are comparable.

Finally, a weak emission seems to arise at the position of Gal–A in the IRAC bands. However, as shown in Figure 3.3, this could be partially contaminated by the emission of the two nearby galaxies at $z_{\mathrm{phot}} \sim 2$ in the 3.6 and 4.5 μm bands, while it seems to emerge from the background at 8.0 μm, where Gal–B and Gal–C become fainter. We find that, in the *Rainbow* catalog (Pérez-González et al. 2008; Barro et al. 2011a,b), Gal–B and Gal–C have been deblended in all the four IRAC channels using the Subaru r band as a prior for the two sources, while no counterpart of Gal–A is present. We briefly summarize here the main steps of the deblending procedure adopted in Barro et al. (2011a,b). At first, several catalogs are cross-correlated to the IRAC 3.6+4.5 μm source position using

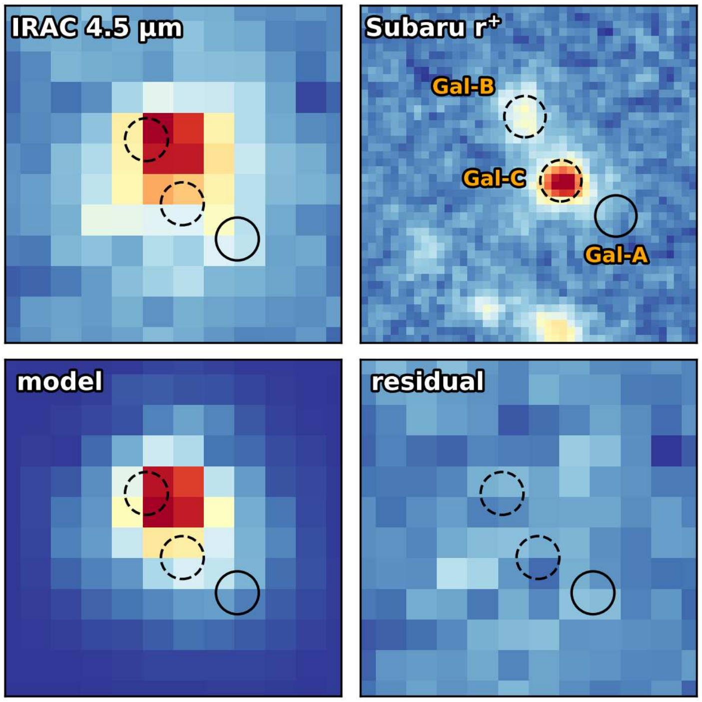


Figure 3.4: Example of deblending procedure for Gal–A in the IRAC 4.5 μm image. *Top right:* Subaru r^+ band image showing the optical emission of Gal–B and Gal–C (identified by the dashed black circles). Gal–A (solid black circle) is not visible. *Top left:* IRAC 4.5 μm image showing the blended sources. *Bottom left:* model of the blended sources obtained by convolving the fluxes of the two visible optical sources with the IRAC PSF. *Bottom right:* residual map obtained from the subtraction of the model from the true image.

a 2 arcsec search radius, in order to identify the associated counterpart in the available bands. When multiple counterparts lying at least 1 arcsec away (half of the typical IRAC PSF, i.e., $\sim$ 2 arcsec) from each other are found in the optical/NIR images, the deblending procedure is applied. The PSF of the higher resolution image (in our case, the Subaru *r* band) is convolved with the IRAC PSF to create a model image. Then, the intensity of each source is scaled to match the flux measured on the IRAC image in 0.9 arcsec radii centered on the positions of the optical counterparts. At the end, total magnitudes are computed by applying aperture corrections in each IRAC band. For more details on the deblending procedure and on the counterpart identification see Barro et al. (2011a,b) and references therein.

In a similar way, we attempt a deblending procedure on Gal−A using the `2D GALFIT` fitting algorithm (Peng et al. 2002) in order to extract the photometric information on this object from the IRAC bands. We model Gal−B and Gal−C as point-like sources, using their optical positions and deblended fluxes from *Rainbow* as a first guess, and considering for each IRAC channel its typical PSF. To obtain the flux in each channel, we perform aperture photometry at the position of Gal−A in the residual maps[1]. As an example, Figure 3.4 shows the results of this procedure in the IRAC 4.5 μm filter. We are aware that with this method we may underestimate the flux of Gal−A in the IRAC channels as we are spreading the global flux of the three components on only two sources. To account for this, when performing SED-fitting (see Section 3.3.1.1) we decide to consider the IRAC fluxes ranging between the deblended (lower) and blended (higher) values. We find, however, that our conclusions do not depend on this assumption; in fact, we obtain similar results when using the deblended fluxes in the SED-fitting. As an alternative approach, we try to fit a three-component model leaving as a free parameter the flux corresponding to Gal−A and using the ALMA continuum peak position as a prior. However, probably due to the small distance between the galaxies, the code is not able to perform the fit. Table 3.1 summarizes the photometric information we obtain for Gal−A; this is exploited in Section 3.3.1.1 to estimate the physical properties of this galaxy from the SED-fitting.

3.2.4 Analysis of the emission line

Since Gal−A shows no optical counterpart, we do not know a priori the nature of the emission line; it could be [CII] emission at a similar redshift of DC_665626 (i.e. $z_{\rm spec} = 4.583$), but also high-J CO transitions are expected ($J_{\rm up} > 3$) at the observed frequencies in ALMA Band 7, although at lower redshift (Carilli & Walter 2013).

In this work, we consider only the two high-J CO transitions with $J_{\rm up} = 9, 10$ which fall into the SPW of observation at $z \gtrsim 2$. Indeed, Ilbert et al. (2013) claim that galaxies at $z < 2$ (corresponding in our case to lower CO transitions) should be more easily detected

---

[1] As we use a fixed aperture of 1.4 arcsec of diameter (which is smaller than the typical PSF of the IRAC channels), we compute aperture corrections from point-source objects lying in the field of our galaxy, in order to estimate the total fluxes in the IRAC filters. In particular, we divide the flux measured in these bands by 0.28, 0.30, 0.29 and 0.19, going from 3.6 to 8.0 μm.

Table 3.1: Summary of available data for Gal–A in each photometric band used for the SED-fitting (see Section 3.3.1.1). The first two columns are the instruments (with relative telescopes) and filters used. Central wavelength is the mean wavelength weighted by the transmission of the filter. In the last column we report the 5σ upper limits for all the bands from the UV to the NIR, except for the K_s detection which is obtained with `SExtractor`. For the IRAC channels, we report the upper limits obtained by measuring the flux at the position of Gal–A before the deblending procedure (see text). All the flux measurements are computed in apertures of 1.4 arcsec of diameter.

Instrument /Telescope	Filter	Central λ [μm]	Observed flux [μJy]
MegaCam/CFHT	u^*	0.3811	$< 4.72 \times 10^{-2}$
HSC/Subaru	g	0.4816	$< 3.50 \times 10^{-2}$
	r	0.6234	$< 4.79 \times 10^{-2}$
	i	0.7741	$< 7.24 \times 10^{-2}$
	z	0.8912	$< 1.04 \times 10^{-1}$
VIRCAM /VISTA	Y	1.0224	$< 3.28 \times 10^{-1}$
	J	1.2556	$< 3.98 \times 10^{-1}$
	H	1.6499	$< 5.40 \times 10^{-1}$
	K_s	2.1578	$(6.31 \pm 0.76) \times 10^{-1}$
IRAC/Spitzer	ch1	3.5573	< 2.40
	ch2	4.5049	< 2.67
	ch3	5.7386	< 3.59
	ch4	7.9274	< 10.11

in the UV/optical filters, with more than the 95% of the sources detected in at least four photometric bands, from the UV to the NIR (Ilbert et al. 2006). Therefore, if our source was at $z < 2$, we would expect it to be visible in the optical bands shown in Figure 3.3.

For these reasons, in the following we discuss the nature of Gal–A considering three transitions as possible interpretations for the observed emission: [CII] at $\nu_{\rm rest}$ = 1900.5 GHz, CO(9-8) at $\nu_{\rm rest}$ = 1036.9 GHz, and CO(10-9) at $\nu_{\rm rest}$ = 1152.0 GHz. As the observed emission line has a peak frequency of 340.76 GHz, Gal–A would be at redshift $z_{\rm gal}$ = 4.577, $z_{\rm gal}$ = 2.043 and $z_{\rm gal}$ = 2.381 for [CII], CO(9-8) and CO(10-9), respectively. Table 3.2 lists the considered transitions and their rest frequencies, as well as the corresponding redshift for Gal–A in the three cases.

Table 3.2: Summary of the physical parameters estimated for the three possible emission lines attributed to Gal−A. The first three columns report the considered emission line, its rest-frequency emission, and the redshift z_{gal} derived using the observed peak frequency, respectively. The fourth and fifth columns list the line luminosity (L_{line}) and the total infrared luminosity (L_{FIR}) for each emission lines, respectively. Finally, the last column report the SFRs, directly computed from the FIR luminosities following Kennicutt (1998).

	ν_{rest} [GHz]	z_{gal}	$\log(L_{\mathrm{line}})$ [$L_\odot$]	$\log(L_{\mathrm{FIR}})$ [$L_\odot$]	$\log(\mathrm{SFR})$ [$M_\odot$ yr^{-1}]
CO(9-8)	1036.9	2.043	8.04 ± 0.04	11.44 ± 0.50	1.45 ± 0.50
CO(10-9)	1152.0	2.381	8.20 ± 0.04	11.42 ± 0.50	1.43 ± 0.50
[CII]	1900.5	4.577	8.88 ± 0.04	11.38 ± 0.50	1.38 ± 0.50

To estimate the intensity of the line and continuum emissions from Gal−A, we separate these components using the CASA `IMCONTSUB` task. Giving in input all the channels in the SPWs free of the emission line, this task creates a continuum map of the source and a continuum-subtracted data cube. We then select all the consecutive channels having emission above $1\sigma_{\mathrm{spec}}$ (i.e. the rms estimated from the line spectrum) encompassing the emission line in order to compute the moment-0 map with the CASA `IMMOMENTS` task.

The line and continuum fluxes are computed using the CASA `IMFIT` task. We define a region surrounding the emissions and then select only the pixels with a flux density larger than 2σ. As the size of the emission region is comparable with the clean beam size, we assume that the source is unresolved and we take the peak flux as the total flux. We obtain $S_{\mathrm{cont}} = 245 \pm 24$ μJy and $S_{\mathrm{line}}\Delta v = 1.19 \pm 0.11$ Jy km s^{-1} for the continuum and line, respectively.

We derive the total infrared (between 8 and 1000 μm) luminosity of the source, in the three cases, assuming a shape of its SED from Magdis et al. (2012) and normalizing its flux to S_{cont}, which is the observed flux at $\sim 845 - 880$ μm. According to Kennicutt (1998), this luminosity also provides a good estimate of the obscured SFR. We obtain $\log(L_{\mathrm{FIR}}/L_\odot) = 11.38 \pm 0.5$ in case of [CII] emission, $\log(L_{\mathrm{FIR}}/L_\odot) = 11.44 \pm 0.50$ for CO(9-8) and $\log(L_{\mathrm{FIR}}/L_\odot) = 11.42 \pm 0.50$ for CO(10-9) emissions. The uncertainties on the FIR luminosities are calculated by adding in quadrature the error on the continuum flux (~ 0.04 dex, which directly affects the L_{FIR} estimates) and a conservative systematic error of 0.5 dex which takes into account possible variations in the luminosity caused by different SED templates (e.g. Capak et al. 2015; Casey et al. 2018). As can be seen, this latter term dominates over the uncertainty on the continuum flux. Following Equation 4 in Kennicutt (1998), these FIR luminosities translate into SFRs[2] ranging from 24 to 28 $M_{\odot}$ yr^{-1}. Finally, we estimate the line luminosities as in Solomon, Downes, & Radford (1992) using the following relation:

$$L_{line} = 1.04 \times 10^{-3}\, S_{line}\, \Delta v\, D_L^2\, \nu_{obs}\ [L_{\odot}], \tag{3.1}$$

where D_L is the luminosity distance of the source in Mpc and ν_{obs} the observed peak frequency in GHz. We thus obtain $\log(L_{[CII]}/L_{\odot}) = 8.88 \pm 0.04$, $\log(L_{CO}/L_{\odot}) = 8.04 \pm 0.04$ for CO(9-8) and $\log(L_{CO}/L_{\odot}) = 8.20 \pm 0.04$ for CO(10-9), where the uncertainties are computed by propagating the line flux error using Equation 3.1. All the above-mentioned physical quantities computed for Gal−A are reported in Table 3.2.

3.3 RESULTS

3.3.1 ON THE NATURE OF THE SERENDIPITOUS SOURCE

With no detections in the optical bands and with the only information of the ALMA Band 7 line and continuum, unveiling the nature of Gal−A is a challenging task. We use here the physical quantities estimated in Section 3.2.4 to deduce plausible conclusions from our source.

Figure 3.5 (left panel) shows the correlation between L_{CO} (for the (9-8) and (10-9) transitions) and L_{FIR} for a compilation of SFGs in the literature, together with the expected position of Gal−A. The respective best-fitting lines on the individual data are also shown (solid lines, Liu et al. 2015). It is worth noting that the reported values are for local galaxies, spanning a FIR luminosity range between $\sim 10^8 - 10^{12}$ $L_{\odot}$. However, the empirical correlations continue to apply even including high-redshift galaxies (filled circles in the figure). In this case indeed, as shown in Liu et al. (2015), the results of the

---

[2] We scale the SFR from Salpeter to Chabrier IMF by dividing by 1.7 (e.g., Zahid et al. 2012).

fit do not significantly change. We then note that the computed L_{FIR} by Liu et al. (2015) are integrated between 40 − 400 μm, which is a smaller wavelength range with respect to the one adopted in this paper to compute L_{FIR}. In order to take this difference into account, we rescale the FIR luminosities of Gal−A in Figure 3.5 to the same integration interval as in Liu et al. (2015), for consistency ($L_{FIR}^{8-1000}/L_{FIR}^{40-400} \sim 1.4$, on average). It can be seen that, for both possible CO transitions, our galaxy would be an outlier of the empirical relations found by Liu et al. (2015), if it was at $z \sim 2$. However, considering the large uncertainties on L_{FIR} (i.e. 0.5 dex), Gal−A could still be part of the lower envelope of local SFGs in the figure, tracing high-density regions ($n_{H2,crit} \sim 10^5 - 10^6$ cm^{-3}; Carilli & Walter 2013) where the star formation may occur.

In the right panel of Figure 3.5 we plot the [CII] luminosity as a function of L_{FIR} in the case that Gal−A was a [CII] emitter at $z \sim 4.6$, along with other results in the literature for different types of objects (e.g. Malhotra et al. 2001a, Stacey et al. 2010). Our source perfectly sits on the local SFGs relation, with $\log(L_{[CII]}/L_{FIR}) \sim -2.5$; possibly, this galaxy may belong to the high-redshift SFGs population which extends to $\log(L_{FIR}/L_\odot) \sim 11$. As previously said, the [CII] line is mostly produced by the UV radiation field in star-forming regions (e.g. Cormier et al. 2015), and so it could be a good tracer of the SFR (e.g., Schaerer et al. 2020). As the FIR emission marks out the SFR of a source, the relation between the [CII] luminosity and L_{FIR} translates into a correlation between $L_{[CII]}$ and the SFR of a galaxy. Gal−A follows this relation, not showing the typical [CII]-deficit which arises at $L_{FIR} > 10^{11}$ $L_\odot$ (e.g. Luhman et al. 1998; Malhotra 2001; Luhman et al. 2003; Lagache et al. 2018).

These results suggest that our source, randomly detected in the DC_665626 field, may more likely be a strongly obscured [CII] emitter at high redshift. However, to validate this hypothesis, more data are needed.

For the sake of simplicity, in the next sections we adopt an intermediate CO redshift (i.e. $z = 2.2$) between those of the two CO transitions reported in Table 3.2, as both cases lead to similar results.

3.3.1.1 ESTIMATE OF THE PHYSICAL PROPERTIES

We derive the physical parameters of Gal−A, such as its stellar mass, through SED-fitting, comparing the results from the `LePHARE` (Arnouts et al. 1999; Ilbert et al. 2006) and

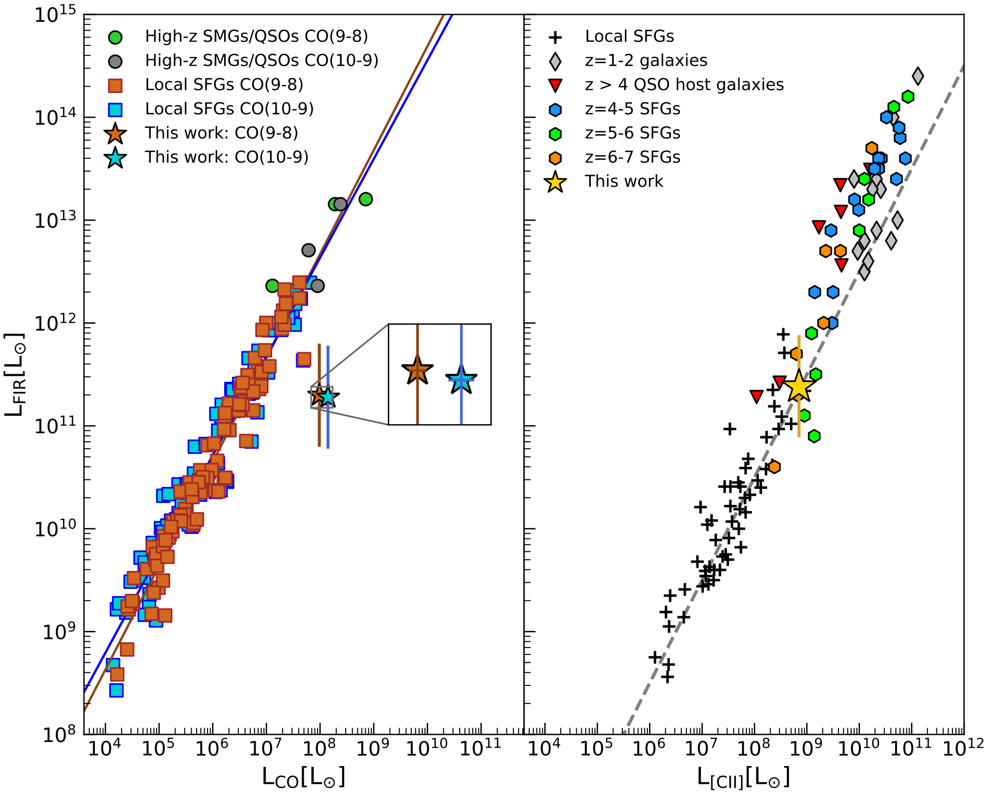


Figure 3.5: *Left panel:* empirical relations between CO(9-8) (solid brown line), CO(10-9) (solid blue line) and FIR luminosity (Liu et al. 2015) with overlaid the values for individual local galaxies as brown and blue squares, respectively (Liu et al. 2015, private communication). The two stars are the values found for Gal–A in this work (same color legend). Error bars are estimated by propagating the error of the line flux on L_{CO}, and assuming a variation of 0.5 dex for L_{FIR}. Also shown are the values obtained for high-redshift SMGs/QSOs as the green and gray filled circles in case of CO(9-8) and CO(10-9) transitions, respectively (Carilli & Walter 2013; ALMA Partnership et al. 2015; Carniani et al. 2019). *Right panel:* [CII] as a function of FIR luminosity for several kinds of objects at different redshifts. Black crosses are local SFGs (Malhotra et al. 2001a); gray diamonds are $z = 1 - 2$ galaxies, including starburst- and AGN-dominated sources (Stacey et al. 2010); red triangles are $z = 4.1 - 7.1$ QSO host galaxies (Pety et al. 2004; Maiolino et al. 2005; Iono et al. 2006; Maiolino et al. 2009; Wagg et al. 2010; Willott, Omont, & Bergeron 2013); $z = 4 - 7$ SFGs are the cyan, green, and orange hexagons (Lagache, Cousin, & Chatzikos 2018). The dashed gray line represents the average [CII]/FIR ratio for local galaxies (Ota et al. 2014). The yellow star shows the position of our source. Error bars are estimated by propagating the error of the line flux on $L_{[CII]}$, and assuming a variation of 0.5 dex for L_{FIR}.

Table 3.3: Comparison of the physical parameters of Gal–A estimated from the SED-fitting at $z = 2.2$ and $z = 4.6$ with `LePHARE` and `MAGPHYS`. Each value represents the mean of the probability distribution obtained by perturbing the photometry of Gal–A 1000 times and fitting that photometry with the models. The uncertainties are given by the 16th and 84th percentiles of the distributions.

Physical parameters	$z = 2.2$		$z = 4.6$	
	`LePHARE`	`MAGPHYS`	`LePHARE`	`MAGPHYS`
$E_s(B-V)$	$0.5^{+0.1}_{-0.1}$	$0.6^{+0.2}_{-0.1}$	$0.4^{+0.1}_{-0.1}$	$0.3^{+0.1}_{-0.1}$
$\log(\mathrm{Age/yr})$	$8.0^{+0.4}_{-0.3}$	$8.3^{+0.4}_{-0.5}$	$7.9^{+0.3}_{-0.2}$	$8.0^{+0.2}_{-0.3}$
$\log(\mathrm{SFR/M_{\odot}\ yr^{-1}})$	$1.5^{+0.3}_{-0.5}$	$1.4^{+0.4}_{-0.4}$	$2.2^{+0.3}_{-0.5}$	$2.0^{+0.1}_{-0.1}$
$\log(\mathrm{M_{*}/M_{\odot}})$	$9.2^{+0.3}_{-0.2}$	$9.7^{+0.3}_{-0.3}$	$9.9^{+0.2}_{-0.2}$	$10.1^{+0.2}_{-0.2}$
$\log(\mathrm{L_{FIR}/L_{\odot}})$	$11.1^{+0.1}_{-0.1}$	$11.2^{+0.7}_{-0.5}$	$11.9^{+0.1}_{-0.1}$	$11.9^{+0.2}_{-0.2}$

the `MAGPHYS` (da Cunha, Charlot, & Elbaz 2008) codes. In the first case, we use a synthetic set of templates of SFGs based on the stellar population synthesis models from Bruzual & Charlot (2003). We explore constant, exponentially declining (with $\tau = 0.1, 0.3, 1, 3$ Gyr) and delayed (with $\tau = 0.1, 0.5, 1, 3$ Gyr) SFHs. To account for the metallicity dependence, we use models with solar ($Z_{\odot}$) and sub-solar ($0.2\ Z_{\odot}$) metallicity. We then account for the dust attenuation using the Calzetti et al. (2000) attenuation law with a stellar $E_s(B-V)$ ranging from 0 to 0.7 in steps of 0.05. Following Ilbert et al. (2009), we also add the contribution of the rest-frame UV and optical emission lines in the different filters. However, we note that our final results do not change significantly if we choose not to include the contribution of the emission lines to the SED-fitting process. Finally, we perform the fit in the flux density space and add systematic errors (depending on the filter) in order to avoid the χ^2 computation to be dominated by small errors (Faisst et al. 2020b).

Figure 3.6 shows the SEDs obtained with `LePHARE` (blue curves) from the best-fit between the models and the photometry of Gal–A (Table 3.1) at $z = 4.6$ and $z = 2.2$. As Gal–A is very faint from the observed optical to the NIR wavelengths, we decide to perturb the flux in each filter by its relative rms to test the dependence of the fitting on the observed photometry of the galaxy. We thus run a Monte Carlo simulation, building 1000 perturbed SEDs that we then refit, in order to obtain a better estimate of the above-mentioned physical parameters from their probability distributions. More in detail, we extract the perturbed flux in each band from a Gaussian distribution centered on the measured flux and with a standard deviation equal to the measured rms. We list our results in Table 3.3. At $z = 4.6$ these results point toward the solution for which Gal−A is a young, dusty SFG. Moreover, the SFR and the FIR luminosity are quite in agreement with the corresponding quantities in Table 3.2. Adopting the same procedure for the SED-fitting at $z = 2.2$, we find that Gal−A should be a less massive galaxy but, as expected, with a high level of dust obscuration.

We then compare the results from the SED-fitting by `LePHARE` with those obtained with the `MAGPHYS` code, in which we also include the observed ALMA continuum in Band 7. This code is based on the energy balance between the emission by the stellar populations and the emission and attenuation by the dust in the galaxies. In particular, we use the updated version of the `MAGPHYS` code which is optimized for the SED-fitting of high-redshift ($z > 1$) SFGs (da Cunha et al. 2015). The best models that fit the observations are shown in Figure 3.6 as the green curves. At first glance, it is evident that the SEDs reproduced by `LePHARE` and `MAGPHYS` are quite similar to each other at both redshifts. This is confirmed by the physical parameters of our galaxy obtained with `MAGPHYS` after perturbing the photometry 1000 times. These parameters are listed in Table 3.3, along with those computed with `LePHARE`. The results from the two codes are in good agreement within the uncertainties, both at $z = 2.2$ and $z = 4.6$. This reassures us about the robustness of our estimates.

Hereafter, as our conclusions do not change if considering the outputs from one or the other code, we decide to use the physical parameters estimated with `LePHARE` in order to make consistent comparisons to other works.

With the stellar mass obtained from the SED-fitting (i.e. $\log(\mathrm{M}_*/\mathrm{M}_\odot) \sim 9.9$) and the SFRs measured from the FIR luminosity of the source (i.e. $\log(\mathrm{SFR}/\mathrm{M}_\odot\ \mathrm{yr}^{-1}) \sim 1.4$), we determine the position of Gal−A along the main-sequence of SFGs (given the large uncertainties involved, our conclusions do not change if using the SFRs from the SED-fitting). In Figure 3.7 we show the main-sequence relations, assuming a Chabrier IMF,

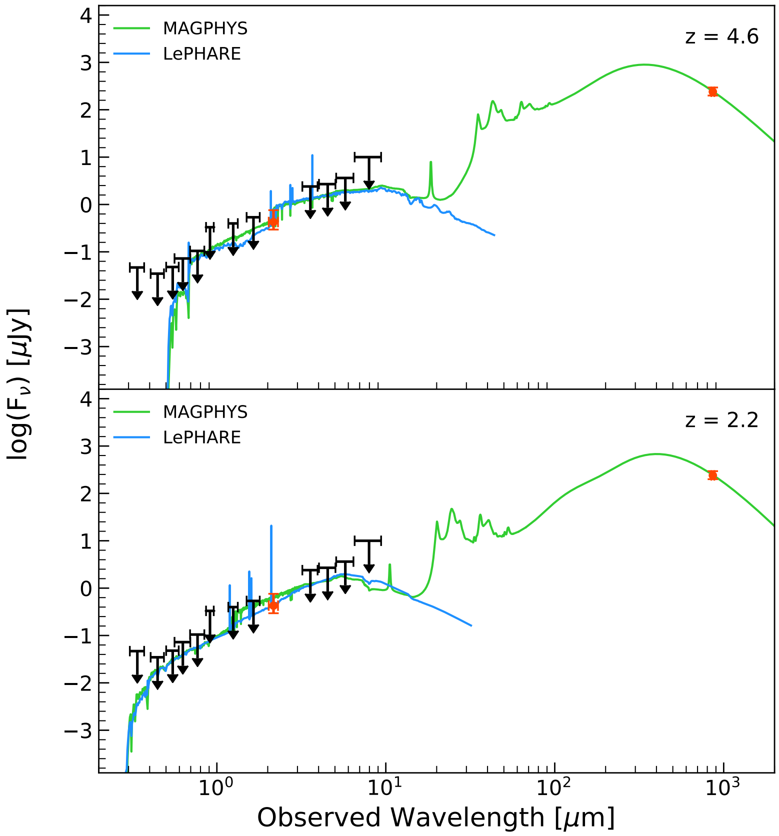


Figure 3.6: SEDs of Gal−A at $z = 4.6$ (top panel) and $z = 2.2$ (bottom panel). The green and blue curves are the best-fit models computed with the MAGPHYS and LePHARE codes, respectively. Upper limits on the flux, as reported in Table 3.1, are shown in black. The orange points with the error bars are the detection in the UVista K_s band and the observed ALMA continuum in Band 7.

at $z = 2.2$ (left panel) and $z = 4.6$ (right panel), as obtained by Speagle et al. (2014) by combining measurements from previous works in the literature. Should the source be at $z = 2.2$, it would lie $\sim 2\sigma$ above the main-sequence, toward the region populated by starburst galaxies. Whether the source is at $z = 4.6$, instead, it would sit on its corresponding main-sequence. In this case, we also show the location of the ALPINE sample (in the redshift range $4.4 \leq z \leq 4.6$) in the figure. The ALPINE galaxies have ages in the range $7.8 \lesssim \log(\mathrm{Age/yr}) \lesssim 9.0$ and $E_s(B-V)$ between 0 and 0.5 (Faisst et al. 2020b). Gal−A has a similar age to those estimated for the ALPINE targets. Moreover,

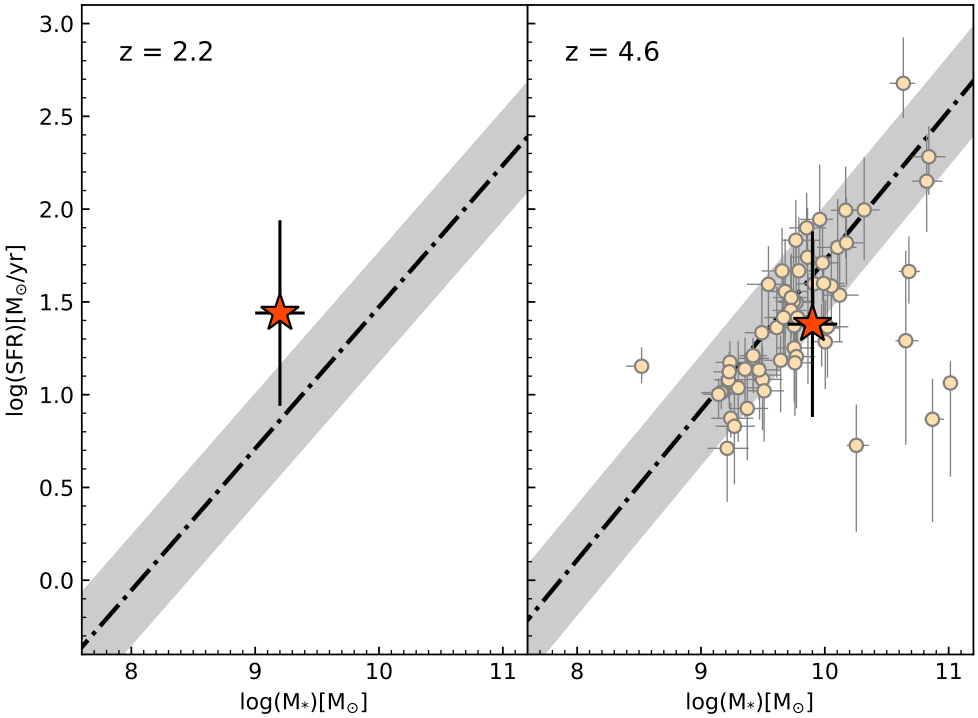


Figure 3.7: Star-forming main-sequence relations (dot-dashed lines; Speagle et al. 2014) at redshift 2.2 (left panel) and 4.6 (right panel). The gray bands indicate the scatter from the main-sequence (±0.3 dex width). The orange stars represent the positions of our source in the diagram, given by the estimated stellar mass from the SED-fitting with LePHARE and the SFR from the FIR luminosity of Gal–A. For the case at $z = 4.6$, we also show the positions of the ALPINE galaxies at $4.4 \leq z \leq 4.6$ (small circles).

its SFR and M_* are comparable with those of the ALPINE sources and place it along the main-sequence at $z = 4.6$. However, the mean $E_s(B-V)$ of the ALPINE galaxies is ~ 0.1, while Gal–A has $E_s(B-V) \sim 0.4$, lying on the tail of the distribution of the color excess, and making it undetected in the optical bands. In this scenario, we should expect an entire population of optically-invisible SFGs still to be observed, which might contribute to the cosmic SFRD at early times.

3.3.1.2 ESTIMATE OF THE DYNAMICAL MASS

We attempt an estimate of the galaxy dynamical mass ($M_{\rm dyn}$) obtained from the FWHM of the observed emission line. Following Wang et al. (2013), we assume a rotating disk geometry for the gas as a first approximation. In this way, $M_{\rm dyn} = 1.16 \times 10^5\ v_{\rm cir}^2\ D$, where D is the disk diameter in kpc, and $v_{\rm cir} = 0.75\ {\rm FWHM_{line}}\ \sin^{-1}(i)$ is the circular velocity of the gas disk in km s^{-1} (with i the inclination angle between the gas disk and the line of sight). Since Gal−A is not resolved, we take the FWHM of the major axis of the 2D Gaussian fitted to the emission line as the size of our galaxy (1.06 ± 0.04 arcsec, which corresponds to 7.09 ± 0.27 kpc at $z \sim 4.6$, and to 8.99 ± 0.34 kpc at $z \sim 2.2$). We derive dynamical masses (uncorrected for the galaxy inclination) of $M_{\rm dyn}\ \sin^2(i) = 4.4 \times 10^{10}\ M_\odot$ and $5.6 \times 10^{10}\ M_\odot$ for $z = 4.6$ and $z = 2.2$ respectively, with a 25% of uncertainty obtained from the individual errors on the $\rm FHWM_{line}$ and on the size of the source. Following Capak et al. (2015), we assume as values for the inclination angle $\sin(i)$=0.45 and $\sin(i)$=1, ranging from a nearly face-on to an edge-on disk. When $\sin(i)$=1, the previous dynamical masses remain unchanged. However, in the case with $\sin(i)$=0.45, $M_{\rm dyn}$ increases of a factor 5. This reflects the large uncertainties on the size and geometry of the source, which cannot be well constrained with the current data and our poor resolution.

Furthermore, this approximation could cease to be valid in the case that the stellar mass of the source is smaller than the mass threshold above which the galaxies are thought to form ordered disks. For instance, Simons et al. (2015) found a so-called "mass of disk formation" of $\log(M_*/M_\odot) = 9.5$ above which the majority of the galaxies of their sample are rotation-dominated. Below this threshold there is instead a large scatter and the galaxies could be either rotation-dominated disks and asymmetric or compact galaxies without any sign of rotation. At $z = 2.2$ Gal−A should have $\log(M_*/M_\odot) = 9.2$, therefore it is prone to this kind of issue.

For comparison, we also run the $^{\rm 3D}$Barolo algorithm (Di Teodoro & Fraternali 2015) on the continuum-subtracted data cube to obtain a more accurate estimate of the dynamical mass. This code creates synthetic 3D observations of the galaxy and compares them with the input data cube, finding the kinematical and geometrical parameters which best describe the data. It is particularly useful to retrieve information on low-resolution data where the kinematics is biased by the size of the beam, as in this case. We find $\log(M_{\rm dyn}/M_\odot) = 10.4 \pm 1.0$ for $z = 4.6$ and $\log(M_{\rm dyn}/M_\odot) = 10.5 \pm 1.0$ for $z = 2.2$. These results are quite in agreement with the former, given the large error on $M_{\rm dyn}$. As a result, at both redshifts $M_{\rm dyn}/M_* >> 1$, likely indicating that the galaxy has recently begun forming stars, resulting in small stellar masses and large gas fractions. In particular, in the case that our galaxy is a [CII] emitter at $z = 4.6$, we estimate the fraction of molecular gas as $f_{\rm mol} = M_{\rm mol}/(M_{\rm mol} + M_*)$, where $M_{\rm mol}$ is the molecular gas mass. Following Zanella et al. (2018), which found a tight correlation between the [CII] luminosity and the molecular gas mass (Dessauges-Zavadsky et al. 2020), we derive $\log(M_{\rm mol}/M_\odot) \sim 10.4$. Assuming a small contribution of the dark matter in the galaxy (e.g. Barnabè et al. 2012), this value is comparable with the gas mass resulting from the subtraction of the stellar mass from the dynamical mass. We thus obtain $f_{\rm mol} \sim 0.75$ that is in perfect agreement with the average molecular gas fraction estimated for the ALPINE targets at $z \sim 4.5$ (Dessauges-Zavadsky et al. 2020). However, given the large uncertainties involved in the computation of the dynamical mass, this result is not conclusive.

3.4 Discussion

For the [CII]/FIR diagnostic, our source presents similar properties to a large population of SFGs in the literature. The SED-fitting reveals a large dust attenuation as expected for such an obscured galaxy and places Gal−A along the main-sequence at $z \sim 4.6$. In addition, we compute the rest-frame equivalent width (EW) of the [CII] line as the ratio between the flux of the [CII] emission and the flux density of the underlying continuum, i.e. $EW_{\rm [CII]} \sim 0.46\ \mu$m. This result is in agreement with the EW estimated for the high-redshift SMGs in Swinbank et al. (2012) identified as [CII] emitters at $z \sim 4.4$ and with the median EW of the candidates [CII] emitters at $z \sim 4.5$ in Cooke et al. (2018). We also compare the observed photometry of our source with that of a sample of ∼ 700 SMGs from the AS2UDS survey in Dudzevičiūtė et al. (2020), where they show how the observed K band magnitude decreases with the redshift (see their Figure 5). We find that Gal−A follows very well the observed trend. In particular, the estimated K_s magnitude sets our galaxy on the composite SMG SED they obtain by measuring the median value of each individual SED at different wavelengths. The set of the above results strongly suggests that Gal−A is a dust-obscured galaxy at $z \sim 4.6$.

Nevertheless, we cannot exclude the possibility that the observed emission line is associated to a dusty, less massive source at $z \sim 2.2$, with a $\sim 2\sigma$ scatter from the main-sequence and with implied CO luminosities that are a factor about 10 or more higher than the typical CO luminosities of local SFGs or high-z SMGs and QSOs. In this respect, Boogaard et al. (2020) investigated the CO excitation in a sample of 22 SFGs at $0.46 < z < 3.60$ as part of the ALMA Spectroscopic Survey in the Hubble Ultra Deep Field (ASPECS; Walter et al. 2016). They found evidence for intrinsic higher excitations in $z > 2$ galaxies compared to lower redshift sources, driven by an evolution in the ISM conditions of SFGs through cosmic time. Along with the detections of high-J CO transitions in their sample (with upper limits up to J = 10), their findings could justify the possibility of a CO(9-8) or CO(10-9) emission observed in ALMA Band 7 at $z \sim 2.2$. In this latter case, the (spectroscopic) redshift of Gal−A would also be comparable with the (photometric) redshifts of Gal−B and Gal−C, maybe suggesting the presence of an on-going merging at that epoch. However, to test this hypothesis, more kinematic information is needed.

In the most likely scenario in which Gal−A is at $z \sim 4.6$, it may be part of the same dark matter halo of DC_665626. In this case, we can assume a stellar mass − halo mass (SMHM) relationship to estimate some physical properties of the halo. There are several ways to derive this relation. For instance, Behroozi, Conroy, & Wechsler (2010); Behroozi, Wechsler, & Conroy (2013) used the abundance matching technique to explore the SMHM relation out to $z \sim 8$, assuming that the most massive galaxies are monotonically assigned to the most massive halos. Another frequent approach is the Halo Occupation Distribution modeling which assumes that the number of galaxies in a given dark matter halo depends only on the halo mass. Harikane et al. (2016) used this method to reproduce the SMHM relation out to $z = 7$, obtaining results in agreement with Behroozi et al. (2013). In particular, since Gal−A has a larger stellar mass than DC_665626, we can suppose that the ALPINE target is a satellite galaxy of our serendipitous source embedded in its dark matter halo. In this case, from the stellar mass of Gal−A (i.e. $\log(M_*/M_\odot) = 9.9 \pm 0.2$), the previously discussed models predict a halo mass between $\log(M_H/M_\odot) \sim 11.5$ and $\log(M_H/M_\odot) \sim 11.7$. Using the empirical model by Mashian, Oesch, & Loeb (2016), which links the SFR of the central galaxy to its host halo mass via abundance matching techniques, M_H also translates into a SFR between $\sim$ 20 and 40 $M_{\odot}$ yr^{-1}, in agreement within the uncertainties with the value estimated from the FIR continuum for Gal−A, i.e. SFR $\sim$ 24 $M_{\odot}$ yr^{-1}. Exploiting this information and following Lapi et al. (2018), we compute the virial radius of the halo as $R_H \equiv [3M_H/4\pi\rho_c\Delta_H E_z]^{1/3}$, where $\rho_c \approx 2.8 \times 10^{11}$ h^2 $M_{\odot}$ Mpc^{-3} is the critical density, $\Delta_H \simeq 18\pi^2 + 82[\Omega_m(1+z)^3/E_z - 1] - 39[\Omega_m(1+z)^3/E_z - 1]^2$ is the non-linear density contrast at the collapse, and $E_z = \Omega_\Lambda + \Omega_m(1+z)^3$ is a redshift dependent factor; we obtain $R_H \sim 39 - 45$ kpc. Comparing this result to the observed spatial offset between our source and DC_665626 ($\sim$ 40 kpc), we may conclude, according to this scenario, that the main ALPINE target could be a low mass satellite in the dark matter halo of Gal−A.

It is worth noting that we obtain similar results even in the opposite case in which Gal−A is a satellite galaxy of DC_665626. Following the same procedure explained above and since DC_665626 has $\log(M_*/M_{\odot}) \sim 9.2$, we obtain $\log(M_H/M_{\odot}) \sim 11.4$ and $\log(SFR/M_{\odot}\ yr^{-1}) \sim 1.0$ (which is consistent with the SFR of the ALPINE target obtained through the SED-fitting). In turn, this provides $R_H \sim 36$ kpc, which is again comparable to the observed offset between the two galaxies.

Finally, Gal−A may also be part of the massive proto-cluster of galaxies PCIJ1001 + 0220 located at $z = 4.57$ in the COSMOS field (Lemaux et al. 2018). In fact, our source lies well inside the 2 Mpc boundary used for spectroscopic membership in that work, with a systemic velocity offset < 350 km s^{-1}. This strengthens the hypothesis that this source is at $z \sim 4.6$.

3.5 Summary

In this chapter, we presented the characterization of a heavily dust-obscured galaxy, named Gal−A, serendipitously discovered in one of the ALPINE pointings. This source is detected both in line and continuum emissions and does not show any associated counterpart, from the UV to FIR wavelengths, except for the K_s band from UVista (DR4). This leads to high uncertainties on the real nature of the observed emission line, i.e., [CII] at $z_{gal} = 4.577$, CO(9-8) or CO(10-9) at $z_{gal} = 2.043$ and 2.381, respectively.

Although we cannot definitively exclude that Gal−A is a dust-obscured galaxy at $z \sim 2.2$, the analysis undertaken in this work suggests that this source is more likely a $z \sim 4.6$ main-sequence SFG missed by the UV/optical surveys because of its high level of dust-obscuration. There is much evidence in favor of this latter scenario:

- The observed emission line has a relatively high rest-frame equivalent width, i.e., $\mathrm{EW}_{\mathrm{[CII]}} \sim 0.46\,\mu$m, compatible with other [CII] emitters at similar redshifts (e.g. Swinbank et al. 2012; Cooke et al. 2018);

- The [CII]/FIR ratio places our source right on the local SFGs relation (Ota et al. 2014), also among several high-redshift SFGs. At the same time, the estimated CO luminosities for Gal−A are more than 10 times larger (at fixed $\mathrm{L}_{\mathrm{FIR}}$) than the typical luminosities of high-J CO transitions for local SFGs and high-z SMGs and QSOs (e.g. Liu et al. 2015; Carilli & Walter 2013), disfavoring the $z \sim 2.2$ case;

- This galaxy is detected in the K_s band of the UVista survey with an AB magnitude of $\sim$ 24.4. This value exactly follows the expected trend with the redshift of the composite SED of the SMG sample analyzed by Dudzevičiūtė et al. (2020). Exploiting the available photometry of Gal−A, we thus obtain $\log(\mathrm{Age/yr}) \sim 7.9$ and $\log(\mathrm{M}_*/\mathrm{M}_\odot) \sim 9.9$. These results, together with $\log(\mathrm{SFR}/\mathrm{M}_\odot\ \mathrm{yr}^{-1}) \sim 1.4$ (estimated from the observed FIR luminosity), place Gal−A on the $z \sim 4.6$ main-sequence, along with the rest-frame UV-selected ALPINE targets;

- Our galaxy may be part of a massive proto-cluster, i.e. PCI J1001+0220, recently discovered at $z \sim 4.57$ in the COSMOS field (Lemaux et al. 2018). In fact, Gal−A lies at only ~ 1 proper Mpc from the center of the overdensity, with a spectroscopic redshift close to that of the systemic redshift of the proto-cluster;

- At this epoch, several dust-obscured galaxies without optical/NIR detections have been already confirmed, mostly as starbursts/SMGs (e.g. Swinbank et al. 2012; Riechers et al. 2013; Simpson et al. 2014; Riechers et al. 2017; Cooke et al. 2018; Alcalde Pampliega et al. 2019). Gal−A could be part of this elusive population of sources likely being a main-sequence galaxy with less extreme properties, such as a smaller mass and/or luminosity.

Whether the emission comes from CO or [CII], both the cases presented above are undoubtedly interesting. If it was at $z \sim 2.2$, our galaxy would increase the sample of high-J CO emitters at high redshift, leading to a more in-depth study of the excitation conditions of the molecular gas in these sources. Only a handful of these kinds of objects have been detected so far and most of them seem to be associated with AGNs activities (Weiß et al. 2007; Riechers et al. 2011, 2013). Should the serendipitous emission be [CII] instead, we would identify a SFG invisible to optical/NIR observations.

As found in Loiacono et al. 2021, a large fraction ($\sim 57\%$) of the serendipitous sources found in ALPINE is confirmed to be [CII] emission at $z > 4$, with an additional $\sim 29\%$ of objects (including our galaxy) showing no counterparts and considered as promising [CII] candidates. Only $\sim 14\%$ of the sources are confirmed to be CO emitters at lower redshift. In particular, using the global sample of [CII] emitters (both confirmed and candidates), they obtain a cosmic SFRD which is ~ 2 times higher than the previous estimates from the UV surveys (if considering their results for the average galaxy population at $z \sim 5$). This result stands out the large contribution of the obscured star formation in the early Universe and the importance of searching for elusive sources, as the one analyzed in this work.

Eventually, we plan to spectroscopic follow-up this source to firmly establish the nature of its emission line. For instance, ALMA observations in Band 6 could reveal [NII] emission at 205 μm rest-frame if the galaxy is at $z \sim 4.6$; in this case the ratio [CII]/[NII] would also provide the fraction of [CII] emission arising from the ionized gas, i.e., from star-forming regions (Oberst et al. 2006, 2011; Zhao et al. 2016). X-shooter at the VLT could also be useful to unveil the redshift of this source by observing the [OII] emission at $z = 4.6$, or even the Hα emission redshifted in the NIR region of the spectrum at $z \sim 2.2$. However, these observations could be hampered by the large $E_s(B-V)$ found for this source, which makes it invisible in the optical filters. Finally, the Near-Infrared Spectrograph (NIRSpec) on the James Webb Space Telescope (JWST) will be a powerful facility for the follow-up of this kind of sources as well.

4 | THE POPULATION OF DUST-OBSCURED GALAXIES FROM ALPINE

The analysis of the fields of view of the 118 ALPINE galaxies revealed a large amount of emission lines, besides those from the main targets, serendipitously discovered and covering a wide redshift range (see Chapter 2). Loiacono et al. (2021) found 14 serendipitous lines with fidelity (i.e., the probability that a detection is an authentic source) greater than 85%, corresponding to a SNR threshold of 6.3. Among these, they identified 8 lines as [CII] emitters (at $4.3 < z < 5.4$) and other 2 as CO transitions with $\mathrm{J_{up}} = 5, 7$ (at $z < 1.3$) based on the photometric or spectroscopic redshifts of the corresponding optical/NIR counterparts observed from the ancillary data. The remaining 4 sources result to be ambiguous because of the lack of counterparts in the available multi-wavelength photometry, suggesting large amounts of dust, or the presence of gas-rich, low-stellar mass galaxies.

In Chapter 3, we investigated the properties of one of these dark galaxies, i.e., Gal−A (named S665626 in Loiacono et al. 2021, where *S* stands for serendipitous), which does not present any optical counterpart in all but the K_s UVista band. Through our analysis, we found more evidence in favor of the [CII] interpretation rather than the CO one. As this is the brightest source in line luminosity among those with no counterpart in ALPINE, we considered its characterization as a benchmark study for the analysis of the other dusty objects. Therefore, in this chapter, I report the individual characterization of the remaining 3 dust-obscured galaxies detected in ALPINE (see Table 4.1), adopting

Table 4.1: Available information on the 3 sources with no counterpart from Loiacono et al. (2021). From left to right are reported the name of the serendipitous source, the coordinates, the observed frequency and FWHM, the flux of the line and of the continuum (if present).

target name	RA [deg]	Dec [deg]	ν_{obs} [GHz]	FWHM [km s^{-1}]	F$_{line}$ [Jy km s^{-1}]	F$_{cont}$ [mJy]
S818760	150.47784	2.54207	341.450	202±12	0.78±0.06	0.425±0.104
S859732	149.99890	2.60632	343.096	99±15	1.21±0.24	-
S5100822662	149.74141	2.08131	344.256	56±7	0.17±0.03	-

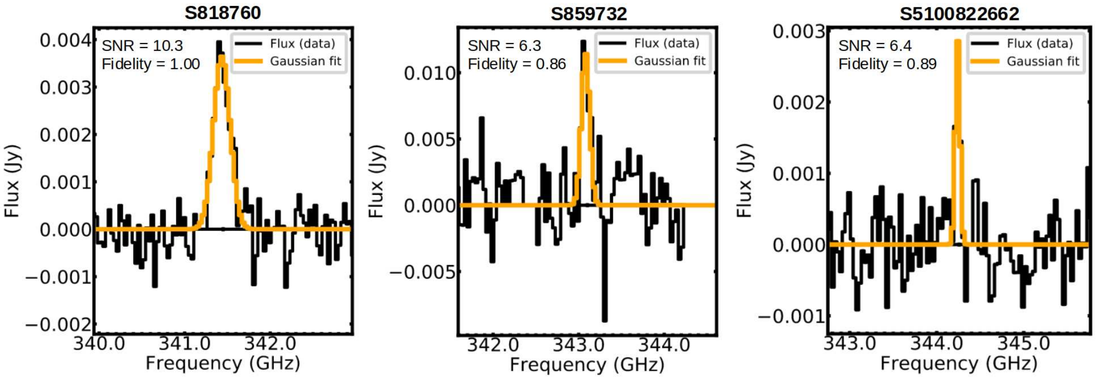


Figure 4.1: Spectra of the serendipitous emission lines with no optical counterpart analyzed in this chapter. In each panel, black histogram is the observed emission, fitted by using a single Gaussian component (orange line). We also show the SNR and fidelity obtained for each source. Figure adapted from Loiacono et al. (2021).

the same methodology used in Romano et al. (2020) and reported in Chapter 3. Figure 4.1 shows the spectra of the serendipitous emission lines, along with the corresponding fidelity and SNR as obtained by Loiacono et al. (2021). Although the following analysis is qualitative and not conclusive to unambiguously define the nature of these sources, it can be used as a reference for future investigations of dusty galaxies for which spectroscopic redshifts are not available, or cannot be easily retrieved.

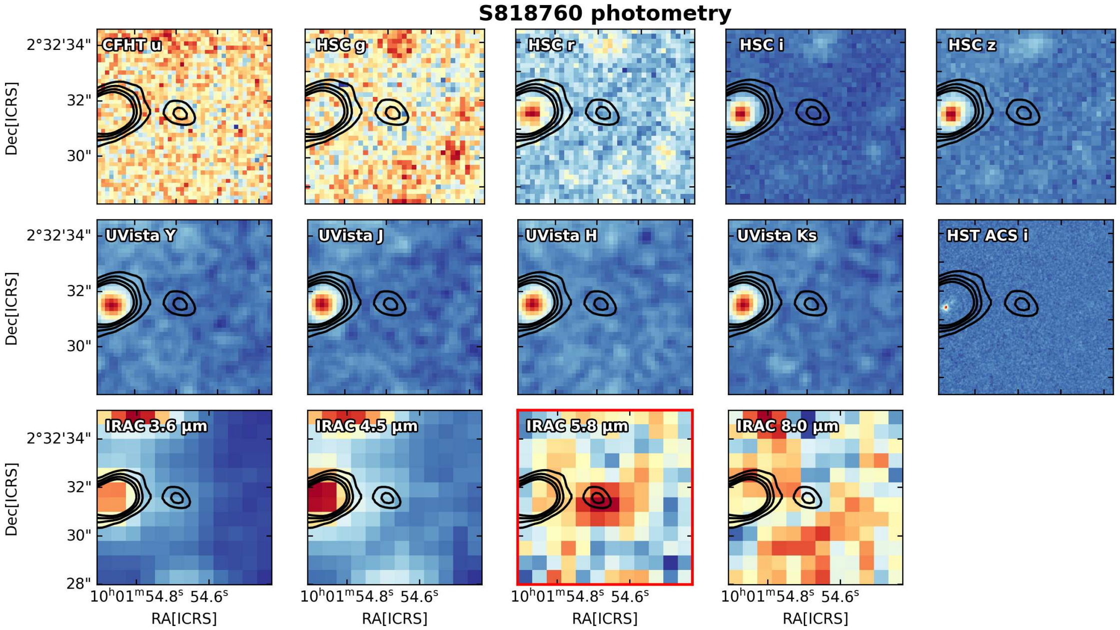


Figure 4.2: Cutouts centered on S818760 in several photometric bands, from UV to NIR. Each image is $6'' \times 6''$ wide. Black contours are $> 3\sigma$ line emission (at steps of 2σ) from the moment-0 map. The panel highlighted in red shows a detection at the position of the source. Wavelengths increase from the upper-left to the bottom-right corner.

4.1 S818760

This source was the object of a first in-depth morphological and kinematic analysis by Jones et al. (2020), and further characterized later as part of statistical studies by Jones et al. (2021) and Romano et al. (2021). It is bright in the sub-mm (detected both in line and continuum), but optically faint, showing any counterpart neither in the available catalogs nor in the images, from the UV to the NIR, except for a clear emission at 5.8 μm in the IRAC Channel 3 band. Furthermore, it is the weakest of a three-component system, lying at $\sim$ 3 arcsec (i.e., $\sim$ 18 kpc at $z = 4.56$, the redshift of the main ALPINE target) from the major galaxy which is experiencing an on-going major merger with the secondary source (Romano et al. 2021; see also Chapter 5). Spatial and kinematic considerations suggest that, if at the same redshift of the other two nearby galaxies, S818760 will also coalesce with them, resulting in a future minor merger. In the following, we test the possibility that the observed emission line comes from both high-J CO transitions (i.e., CO(9-8) and CO(10-9)) or from [CII], in order for the source to be at $z > 2$.

Figure 4.2 shows the cutouts of S818760 (placed at the center of the images) from the UV to the NIR. We examine photometry from CFHT, HSC, HST, UVista (DR4) and SPLASH. In the images, it is also visible the major ALPINE target on the left. A complete image of the possible three-merging system and its morphological and kinematic analysis are also visible in Figure A.7 (right panel) and discussed in Appendix A. No significant emission is found in the different bands except for the IRAC 5.8 μm, where a signal arises from the noise at the position of the source. As done for S665626 (i.e., Gal−A) in Chapter 3, we use the measured line and continuum fluxes to recover the [CII], CO, and FIR luminosities. With this information, we reproduce the diagnostic plots of Figure 3.5 for the source S818760, comparing L_{FIR} and L_{line}. In this case, we do not show the corresponding figure as its outputs are similar to those of Figure 3.5, with the CO emitters not following the empirical relation found by Liu et al. (2015), and the [CII] emitter along the local relation by Malhotra et al. (2001a), leading (accounting only for this diagnostics) to conclusions comparable to that of S665626 (i.e., that this source may more likely be a strongly obscured [CII] emitter at high redshift, strengthening the analysis undertaken by Jones et al. (2020), which consider this source to be part of an on-going merging at $z \sim 4.6$).

About the photometry, we center a fixed aperture of 2 arcsec of diameter on the source (enclosing the 3σ contours of the emission line detected by ALMA) and estimate its flux in each band. We then compute the limiting magnitude as the standard deviation of the fluxes measured in 10000 apertures (of 2 arcsec of diameter) randomly distributed in a wide region of the sky (masked for the emission of bright sources) and assume it as the error on the flux. Table 4.2 summarizes this information. We use these data to retrieve the physical parameters of the galaxy (e.g. stellar mass, SFR) through SED-fitting using the `LePHARE` and `MAGPHYS` codes. However, given that the two algorithms provide similar results (as already shown in Chapter 3) and since we are including the ALMA continuum in Band 7, in the following we will discuss only the results from `MAGPHYS`. For each band, we set the 5σ magnitude as an upper limit to the flux. The galaxy is detected at 5.8 μm, thus in this case we use the flux measured in the 2 arcsec aperture for the SED-fitting, assuming the 1σ limiting magnitude as the error on the IRAC Channel 3

Table 4.2: Summary of available photometry for the serendipitous sources in each photometric band used for the SED-fitting. The first two columns are the instruments (with relative telescopes) and filters used. Central wavelength is the mean wavelength weighted by the transmission of the filter. In the last columns we report the 5σ upper limits (or detected fluxes) for all the bands from the UV to the NIR, and for each of the dust-obscured serendipitous sources. All flux measurements are computed in apertures of 2 arcsec of diameter, except for the detections in S5100822662 that are computed with `SExtractor`.

Instrument /Telescope	Filter	Central λ [μm]	Observed flux S818760 [μJy]	Observed flux S859732 [μJy]	Observed flux S5100822662 [μJy]
MegaCam/CFHT	u^*	0.3811	$< 8.00 \times 10^{-2}$	$< 8.00 \times 10^{-2}$	$< 8.00 \times 10^{-2}$
HSC/Subaru	g	0.4816	$< 6.08 \times 10^{-2}$	$< 6.08 \times 10^{-2}$	$< 6.08 \times 10^{-2}$
	r	0.6234	$< 8.32 \times 10^{-2}$	$< 8.32 \times 10^{-2}$	$(5.40 \pm 1.66) \times 10^{-2}$
	i	0.7741	$< 1.36 \times 10^{-1}$	$< 1.36 \times 10^{-1}$	$(2.05 \pm 0.27) \times 10^{-1}$
	z	0.8912	$< 1.89 \times 10^{-1}$	$< 1.89 \times 10^{-1}$	$(2.21 \pm 0.38) \times 10^{-1}$
VIRCAM /VISTA	Y	1.0224	$< 1.74 \times 10^{-1}$ [a]	$< 4.79 \times 10^{-1}$ [b]	$(1.49 \pm 0.35) \times 10^{-1}$
	J	1.2556	$< 2.09 \times 10^{-1}$ [a]	$< 5.75 \times 10^{-1}$ [b]	$(2.96 \pm 0.42) \times 10^{-1}$
	H	1.6499	$< 3.02 \times 10^{-1}$ [a]	$< 8.32 \times 10^{-1}$ [b]	$(3.25 \pm 0.60) \times 10^{-1}$
	K_s	2.1578	$< 3.98 \times 10^{-1}$ [a]	$< 5.75 \times 10^{-1}$ [b]	$(6.31 \pm 0.80) \times 10^{-1}$
IRAC/Spitzer (SPLASH)	ch1[c]	3.5573	$< 1.83 \times 10^{-1}$	$< 1.83 \times 10^{-1}$	$< 9.50 \times 10^{-1}$ [d,e]
	ch2[c]	4.5049	$< 2.00 \times 10^{-1}$	$< 2.00 \times 10^{-1}$	$< 8.40 \times 10^{-1}$ [d,e]
	ch3[c]	5.7386	8.46 ± 0.44 [d]	< 2.20	< 2.20
	ch4[c]	7.9274	< 2.41	< 2.41	< 1.98 [d,e]

[a] 5σ limiting magnitude from UVista (DR4) Ultra-deep field;

[b] 5σ limiting magnitude from UVista (DR4) deep field;

[c] 5σ limiting magnitude from Laigle et al. (2016), correcting for aperture (from 3 arcsec to 2 arcsec);

[d] this flux is corrected for the aperture by dividing by 0.44, 0.43, 0.37 and 0.31, going from 3.3 to 8.0 μm;

[e] flux uncorrected for blending.

flux. Figure 4.3 (left panel) shows the results of the SED-fitting at $z \sim 4.6$ and $z \sim 2.2$ (assuming the latter as an intermediate redshift between the two considered high-J CO transitions, as done for S665626). As evident, the measured photometry cannot be well reproduced at both redshifts. Indeed, it is not easy to explain the observed emission excess at 5.8 μm for a high-z galaxy.

In the hypothesis that the observed emission line is instead due to a different transition, we try to estimate the photometric redshift of this galaxy using the `MAGPHYS`+photo-z code (Battisti et al. 2019). The right panel of Figure 4.3 shows the computed SEDs (in green) that are now able to reproduce the observed photometry, resulting in an average redshift of $z = 0.85$. In this case, the 5.8 μm excess could be explained by a polycyclic aromatic hydrocarbon (PAH) feature at 3.3 μm rest-frame observed in normal SFGs, that is redshifted in the IRAC Channel 3 band. The observed ALMA line at $\nu_{\rm obs} = 341.45$ GHz could be then ascribed to the CO(5-4) transition. This emission has a rest-frequency $\nu_{\rm rest} = 576.27$ GHz, that would imply a redshift of $z = 0.69$ for our galaxy (for which the 3.3 μm PAH emission would still be inside the IRAC Channel 3 bandwidth). This latter case is also shown in Figure 4.3 (orange curves), along with the previous SEDs at $z = 0.85$.

This low-z solution is even in agreement with the results from Liu et al. (2015). In fact, if this galaxy was at $z = 0.69$ (i.e. CO(5-4) emitter), it would lie on the $L_{\rm CO}/L_{\rm FIR}$ relation observed by Liu et al. (2015), as displayed in Figure 4.4. However, this case is not favored by the evidence of merging with the two nearby galaxies at $z \sim 4.6$ found by Jones et al. (2020). On the other hand, they found that this source has a wide separation both spatially (~ 3 arcsec) and kinematically (~ 300 km s^{-1}) from its nearby major companion, suggesting that, if at $z \sim 4.6$, this galaxy is not experiencing an on-going merging but rather it will merge with the other two objects at later time. Finally, although they consider this source as a [CII] emitter, the lack of an optical counterpart does not allow us to definitively establish the high-z nature (i.e., $z \sim 4.6$) of S818760. The information about S818760 (e.g., FIR and line luminosities) for the different cases is reported in Table 4.3.

Assuming that this galaxy is a CO(5-4) emitter at $z = 0.69$, we then perturb the photometry 100 times and re-do the SED-fitting with `MAGPHYS` in order to explore the dependence of the physical parameters of the source on the adopted photometry. In

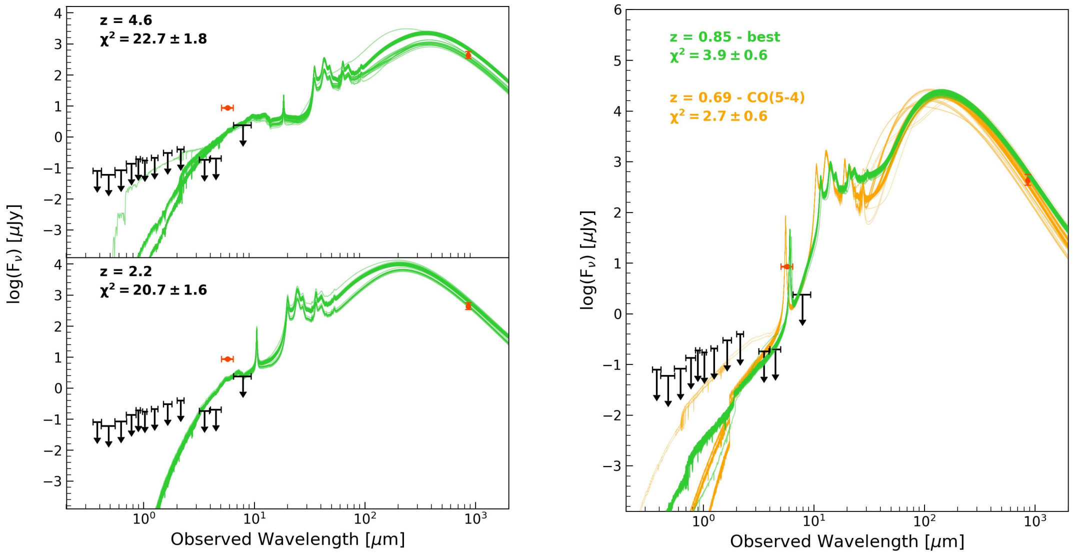


Figure 4.3: *Left figure:* SEDs of S818760 at $z = 4.6$ (top panel) and $z = 2.2$ (bottom panel). The green curves are the best-fit models computed with the `MAGPHYS` code, after perturbing the photometry 100 times. Upper limits on the flux, as reported in Table 4.2, are shown in black. The orange points with the error bars are the detection in the IRAC 5.8 μm band and the observed ALMA continuum in Band 7. *Right figure:* SEDs of S818760 with an average redshift of 0.85 obtained from `MAGPHYS`+photo-z (in green) and for a CO(5-4) emitter with $z = 0.69$ (in orange). Average χ^2 values are also reported for each redshift.

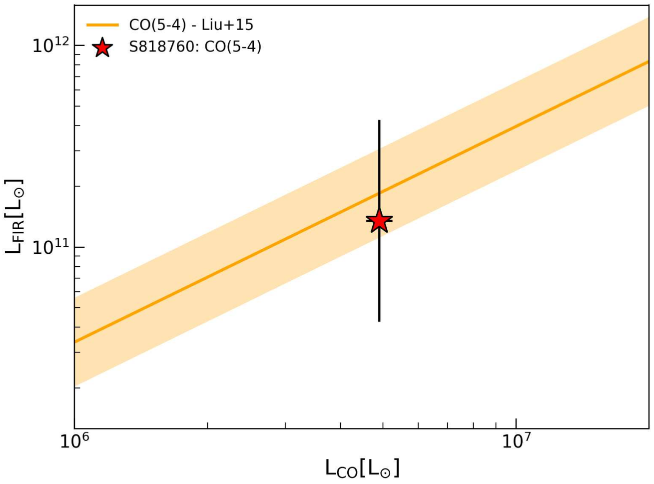


Figure 4.4: Empirical relation between CO(5-4) and FIR luminosity for local galaxies as obtained by Liu et al. (2015). The solid line and shaded area are the best-fit on the data and its uncertainty, respectively. The red star represents the position in the diagram of S818760.

Table 4.3: Physical parameters of the sources estimated from the computed line and continuum fluxes. As the continuum emission is not detected in S859732 and S5100822662, we cannot directly obtain the FIR luminosity and the SFR as done for S818760.

target	transition	$\nu_{\rm rest}$ [GHz]	$z_{\rm gal}$	D_L [Mpc]	$\log(L_{\rm line})$ [$L_\odot$]	$\log(L_{\rm FIR})$ [$L_\odot$]	$\log({\rm SFR})$ [$M_\odot\ {\rm yr}^{-1}$]
S818760	CO(5-4)	594.1	0.6877	4185.1	6.69 ± 0.03	11.27 ± 0.50	1.28 ± 0.50
	CO(9-8)	1036.9	2.0368	15887.6	7.84 ± 0.03	11.69 ± 0.50	1.70 ± 0.50
	CO(10-9)	1152.0	2.3738	19149.2	8.01 ± 0.03	11.67 ± 0.50	1.68 ± 0.50
	[CII]	1900.5	4.5660	41914.5	8.69 ± 0.03	11.63 ± 0.50	1.64 ± 0.50
S859732	CO(9-8)	1036.9	2.0222	15748.4	8.03 ± 0.09	-	-
	CO(10-9)	1152.0	2.3577	18991.4	8.19 ± 0.09	-	-
	[CII]	1900.5	4.5393	41625.7	8.87 ± 0.09	-	-
S5100822662	CO(9-8)	1036.9	2.0120	15651.2	7.17 ± 0.07	-	-
	CO(10-9)	1152.0	2.3463	18879.8	7.34 ± 0.07	-	-
	[CII]	1900.5	4.5206	41423.6	8.02 ± 0.07	-	-

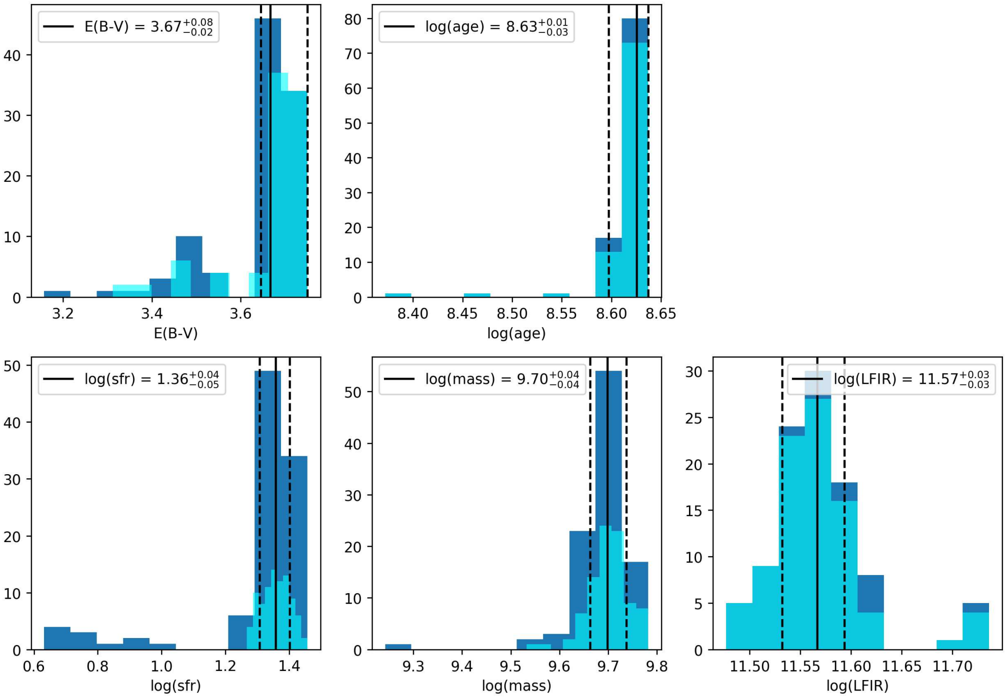


Figure 4.5: Statistical distributions of the physical parameters computed with MAGPHYS from the 100 perturbed SEDs of S818760. Blue histograms represent the whole sample. Cyan histograms are only the SEDs with $\log(\mathrm{SFR}/\mathrm{M}_{\odot}\ \mathrm{yr}^{-1}) > 1.2$. Solid black lines are the mean of these latter data. Dashed lines represent the 16th and 84th percentiles of the respective distributions.

particular, for each SED we take the 50th percentile of the probability distribution of each physical parameter and used that to build our own statistical distribution. Our results are shown in Figure 4.5 (blue histograms). In order to reproduce the observed SFR from the FIR luminosity, we select only the SEDs that give a $\log(\mathrm{SFR}/\mathrm{M}_{\odot}\ \mathrm{yr}^{-1}) > 1.2$ (cyan histograms) and compute the average, 16th and 84th percentiles of those data (black lines). As evident, in order to be at $z = 0.69$, our source should be a massive (i.e., $\log(\mathrm{M}/\mathrm{M}_{\odot}) \sim 9.7$) SFG with a large amount of dust.

Although the observed photometry seems to credit the low-z scenario, the extreme color excess value (i.e., $E(B-V) \sim 3.6$) obtained from the SED-fitting and the close proximity to the main ALPINE target, revive the hypothesis that S818760 is a $z \sim 4.6$ dust-obscured galaxy about to join the nearby merging system at the same redshift. However, with the available data, we cannot provide a firm conclusion on this issue.

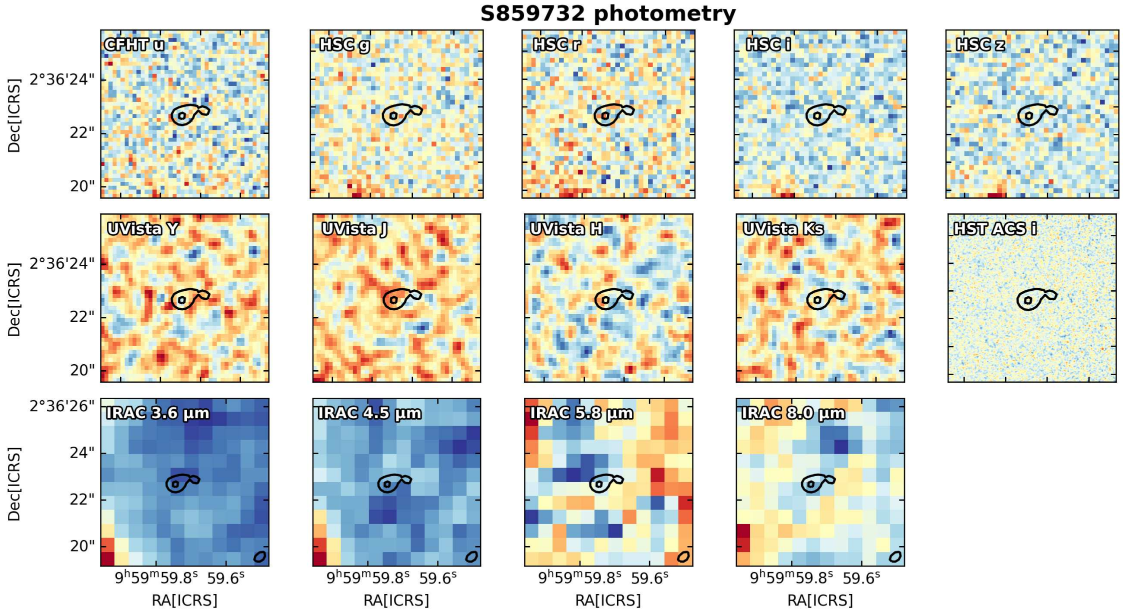


Figure 4.6: Cutouts centered on S859732 in several photometric bands, from UV to NIR. Each image is 6″ × 6″ wide. Black contours are $> 3\sigma$ line emission (at steps of 2σ) from the moment-0 map. Wavelengths increase from the upper-left to the bottom-right corner.

4.2 S859732

S859732 does not show any significant emission in all inspected bands, from UV to NIR (see Figure 4.6). We recover the line luminosity assuming that the emission arises from both [CII] and high-J CO lines (see Table 4.3). However, having completely no information in any band and without the ALMA continuum, we are not able to make the SED-fitting for this source and provide possible constraints on its nature or physical parameters. The only hint we have is that this galaxy lies within 2 arcsec from the main ALPINE target, suggesting the presence of an interactive system. The morphological and kinematic analysis of the principal galaxy suggests indeed the presence of an ongoing merging with another closely spaced component (Romano et al. 2021; see also Chapter 5 and Figure A.9, left panel). In this scenario, S859732 could be a highly dust-obscured galaxy that will eventually merge with nearby sources in the future. However, to confirm or not this possibility, more data and observations are needed.

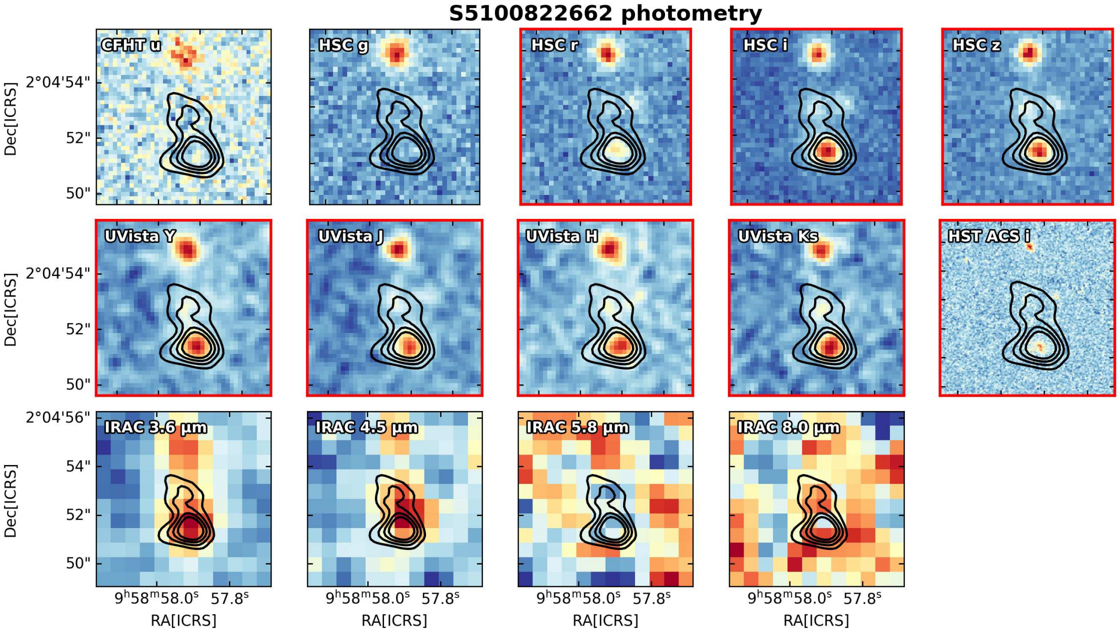


Figure 4.7: Cutouts centered on S5100822662 in several photometric bands, from UV to NIR. Each image is $6'' \times 6''$ wide. Black contours are $> 3\sigma$ line emission (at steps of 2σ) from the moment-0 map. The panels highlighted in red show a faint emission at the position of the source. Wavelengths increase from the upper-left to the bottom-right corner.

4.3 S5100822662

Figure 4.7 shows the images at different wavelengths of S5100822662. It is evident the high proximity (i.e., ~ 1.3 arcsec) of the serendipitous source (at the center) to its associated main target (at the bottom of the figure). In this case, a clear optical counterpart is visible in the maps, emerging from the background from the HSC r band to the UVista K_s band. This galaxy is listed as a detection in the COSMOS2015 catalog (Laigle et al. 2016), with a photometric redshift $z = 0.69$. At this redshift, the emission line observed by ALMA could likely be explained by CO(5-4) emission at $\nu_{rest} = 594.1$ GHz. Therefore, we test the hypothesis that this source is indeed at $z = 0.69$, along with the cases at $z \sim 2.2$ and $z \sim 4.6$ (i.e. high-J CO or [CII] emitter, respectively), by fitting the observed photometry with `MAGPHYS`. In particular, we exploit the photometry reported in Table 4.2, where for the u and g bands we report the 5σ upper limits to the flux, while from

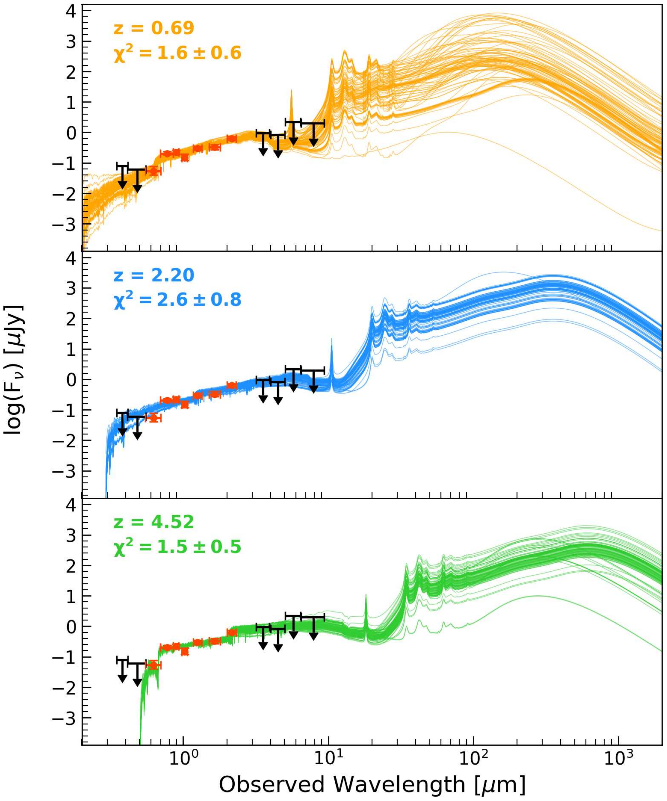


Figure 4.8: SEDs of S5100822662 at $z = 0.69$ (top panel), $z = 2.20$ (middle panel), and $z = 4.52$ (bottom panel). Upper limits on the flux, as reported in Table 4.2, are shown in black. The orange points with the error bars are the detections from the r to the K_s band. Average χ^2 values are also reported for each redshift.

the r to the K_s band we use the fluxes obtained with `SExtractor`. Finally, for the IRAC bands we use the fluxes measured in a 2 arcsec aperture without making a deblending procedure (however, we let these fluxes to vary from zero to the blended values when doing the SED-fitting) and correcting for the aperture. Unfortunately, as this galaxy is detected only in line emission, no information on the ALMA continuum is available to constrain the rest-frame FIR region of the SED.

Figure 4.8 shows the SEDs that fit the observed photometry of S5100822662 at the three different redshifts discussed before (for each case, perturbing the photometry 100 times). At $z = 0.69$ and $z = 4.52$, the models seem to better reproduce the observed

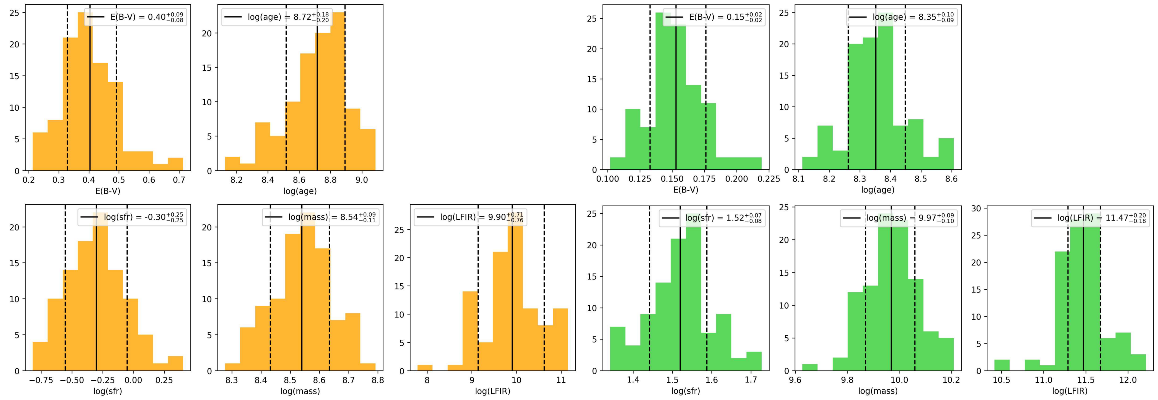


Figure 4.9: Statistical distributions of the physical parameters computed with MAGPHYS from the 100 perturbed SEDs of S5100822662 at $z = 0.69$ (left) and $z = 4.52$ (right). Solid black lines are the means of the distributions. Dashed lines represent the 16th and 84th percentiles of the respective distributions.

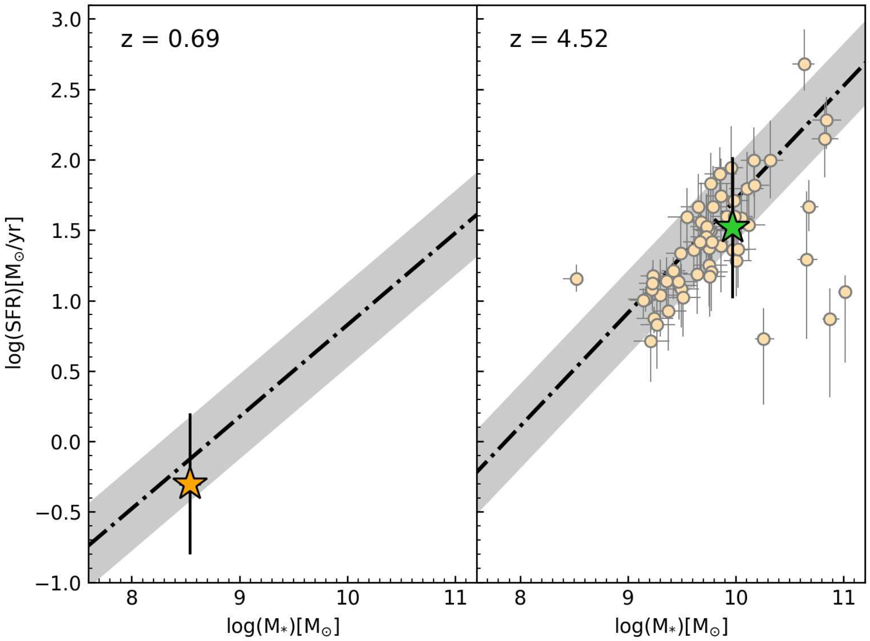


Figure 4.10: Star-forming main-sequence relations (dot-dashed lines; Speagle et al. 2014) at redshift 0.69 (left panel) and 4.52 (right panel). The gray bands indicate the scatter from the main-sequence (±0.3 dex width). The orange and green stars represent the positions of S5100822662 in the diagram, given by the estimated stellar mass and SFR from the SED-fitting. For the case at $z = 4.52$, we also show the positions of the ALPINE galaxies at $4.4 \lesssim z \lesssim 4.6$ (small circles).

photometry, while at $z = 2.20$ the SEDs do not fit well the g and r bands. This is also witnessed by the average χ^2 values at the three redshifts, which are smaller for the $z = 0.69$ and $z = 4.52$ cases. Therefore, we check the statistics of the lower and upper redshift to find the possible physical parameters that describe the galaxy. Figure 4.9 shows the distributions of such parameters at both redshifts. At $z = 0.69$, the galaxy is more dusty, less massive and less star-forming compared to the case at $z = 4.52$. With the results from the SED-fitting, we also check the position of this source along the main-sequence of SFGs following (Speagle et al. 2014), as done for S665626. As shown in Figure 4.10, at both redshifts the source lies along its relative main-sequence, within the ±0.3 dex uncertainty.

Finally, as done for S818760, we make a diagnostic diagram using the FIR luminosity (from the SED-fitting) and the line luminosity (CO or [CII], depending on the redshift) of the source. This is shown in Figure 4.11. As you can see, the $z = 4.52$ case (right panel) is quite in agreement with the local relation, although the source lies at the edge of the observed scatter. On the other hand, if this galaxy was a CO(5-4) emitter at $z = 0.69$ (left panel), it would be an outlier of the empirical relation between FIR and CO luminosities found by Liu et al. (2015), with a much higher CO emission (at fixed L_{FIR}) than observed in local galaxies. This seems to suggest that S5100822662 is a $z = 4.52$ galaxy (i.e., a [CII] emitter), spatially close to the ALPINE main target. This scenario is corroborated by the work by Romano et al. (2021), who analyzed in detail the morphology and kinematics of the observed [CII] emission from this galaxy, suggesting the presence of an on-going major merger with the central ALPINE target, and thus attesting the high redshift (i.e., $z > 4$) nature of the source (see Chapter 5 and Figure 5.2). This conclusion is also confirmed by the qualitative morpho-kinematic classification by Le Fèvre et al. (2020) and by the more quantitative analysis by Jones et al. (2021).

4.4 SUMMARY

In this chapter we exploited some of the methods previously introduced for the characterization of very dusty galaxies observed by chance in the ALPINE program. In particular, we showed that their application to the ALMA-detected dust-obscured sources with no optical counterpart, possibly combined with other information from the [CII]

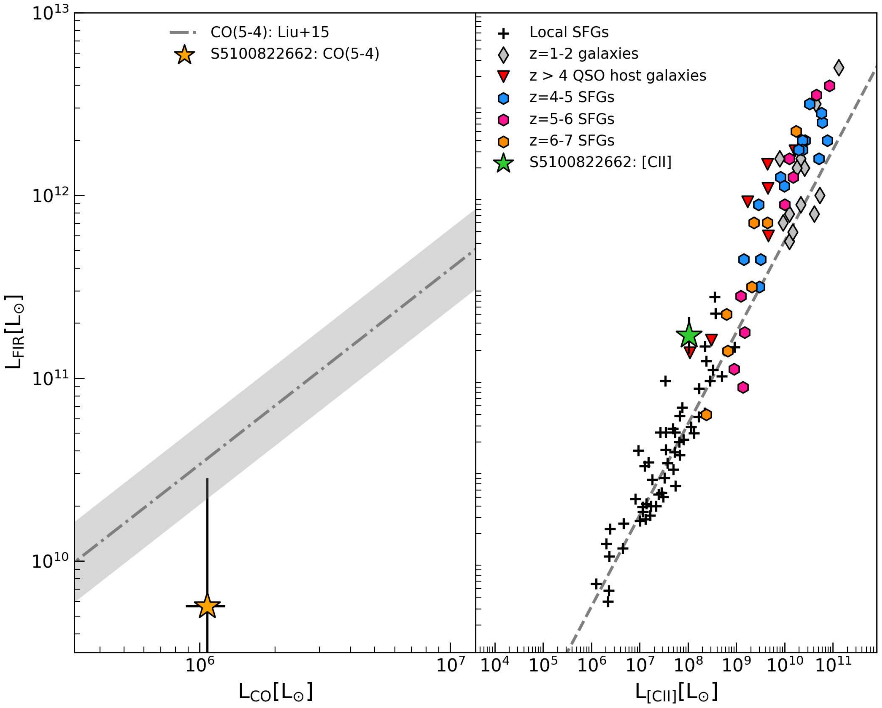


Figure 4.11: *Left panel:* empirical relations between CO(5-4) (dot-dashed line) and FIR luminosity (Liu et al. 2015). The star represents the value found for S5100822662. Error bars are estimated by propagating the error of the line flux on L_{CO}, and assuming the 16th and 84th percentiles of the FIR distribution as its errors. *Right panel:* [CII] as a function of FIR luminosity for several kinds of objects at different redshifts. Same legend as in Figure 3.5. The green star shows the position of S5100822662. Error bars are estimated by propagating the error of the line flux on $L_{[CII]}$, and assuming the 16th and 84th percentiles of the FIR distribution as its errors.

morphology and kinematics, could provide useful insights on their real nature. 4 out of the 14 serendipitous emission lines detected in ALPINE have no spectroscopic or photometric redshifts, with faint or completely absent optical counterparts. One of these galaxies was studied in details in Chapter 3. Here, we analyzed the remaining three sources in a more qualitative way than done for S665626, in order to show the potential of the method. Briefly:

- S818760 presents a clear detection at 5.6 μm only, resulting invisible in all the other bands but ALMA, which detected its continuum at $\sim 800\ \mu$m. Morpho-kinematic studies of the emission line, along with those carried out on the two nearby galaxies at

$z \sim 4.6$, suggest that this could be a dusty source at the same redshift (i.e., [CII]-emitter), forming a three-merging system and missed by UV/optical surveys for its high level of dust. On the other hand, the observed photometry raises a different scenario, for which the galaxy is a CO-emitter at $z \sim 0.69$ with an even greater dust content than the high-z case.

- S859732 has completely no photometric information. We only observed the emission line, from which is not possible to draw conclusions on its real nature. However, this galaxy could be a satellite of the main ALPINE target that is likely a major merger, as indicated by Le Fèvre et al. (2020) and Romano et al. (2021). In this case, this source could be at $z \sim 4.5$, on the way to merge with the nearby major companion.
- S5100822662 is the only dust-obscured serendipitous galaxy which presents some weak emission in multiple optical bands. It is indeed listed as a detection in the COSMOS2015 catalog with photometric redshift $z = 0.69$, for which the emission line could originate from CO(5-4) emission. However, we found that the photometry is better fit by a $z = 4.5$ solution, considering the source a [CII]-emitter instead. This scenario is also supported by the results of the $L_{FIR}/L_{[CII]}$ diagnostic diagram and by the presence of the nearby main target with which the galaxy is likely undergoing a major merger (Romano et al. 2021).

Although not conclusive, this kind of analysis could pave the way to future follow-up and/or spectroscopic campaigns aimed to finally solve the ambiguity on the observed serendipitous emission lines and redshifts of the corresponding galaxies. Moreover, if all of the four serendipitous galaxies with no optical counterpart in ALPINE were [CII] emitters, the fraction of non-target galaxies at $z > 4$ would rise to 86%, providing further evidence for the large population of dust-obscured sources missed by rest-frame UV/optical surveys in the early Universe.

5 | THE CONTRIBUTION OF MAJOR MERGERS TO THE GALAXY MASS-ASSEMBLY AT $z \sim 5$

Based on:
Romano M., Cassata P., Morselli L., et al. 2021, A&A, Volume 653, pg. 111-142

Galaxy mergers are thought to be one of the main mechanisms of the mass-assembly of galaxies in the Universe, but there is still little direct observational evidence of how frequent they are at $z \gtrsim 4$. Recently, many works have suggested a possible increase in the fraction of major mergers in the early Universe, reviving the debate on which processes (e.g., cold accretion, star formation, mergers) most contribute to the mass build-up of galaxies through cosmic time.

In this chapter, we discuss about the importance of major mergers in this context. In the following analysis we used, for the first time, the morpho-kinematic information provided by the [CII] emission, along with archival multi-wavelength data, to identify major mergers at $z \sim 5$. In particular, we find a merger fraction of $f_{\mathrm{MM}} \sim 0.44$ (0.34) at $z \sim 4.5$ (5.5) from ALPINE. By combining our results with those at lower redshifts, we computed the cosmic evolution of the merger fraction which is described by a rapid increase with cosmic time, a peak at $z \sim 3$, and a slow decrease toward earlier epochs. Depending on the timescale prescription used (T_{MM}), this fraction translates into a merger rate per galaxy (R_{MM}) ranging between ~ 0.1 and $\sim 4.0\,\mathrm{Gyr}^{-1}$ at $z \sim 5$, which in turn corresponds to an average number of major mergers per galaxy between 1 and 8 in $\sim$ 12.5 Gyr (from $z = 6$ to the local Universe). When convolved with the galaxy number density at different epochs, $R_{\rm MM}$ provides the merger rate density ($\Gamma_{\rm MM}$) which becomes approximately constant over time at $1 < z < 4$, including values from 10^{-4} to 10^{-3} Gyr^{-1} Mpc^{-3}, depending on the assumed $T_{\rm MM}$. We finally compare the specific star formation and SFRD with the analogous quantities from major mergers, finding a good agreement at $z > 4$ if we assume a merger timescale that quickly decreases with increasing redshift. Our new constraints on the merger fraction from the ALPINE survey at $z \sim 5$ reveal the presence of a significant merging activity in the early Universe. Whether this population of mergers can provide a relevant contribution to the galaxy mass-assembly at these redshifts and through the cosmic epochs is strongly dependent on the assumption of the merger timescale. However, our results show that an evolving $T_{\rm MM} \propto (1+z)^{-2}$ agrees well with state-of-the-art cosmological simulations, suggesting a considerable role of mergers in the build-up of galaxies at early times.

5.1 General context

How galaxies grow their stellar mass through cosmic time is still one of the most puzzling questions of modern cosmology. During the last decades, two major physical mechanisms have been proposed to contribute or dominate the build-up of galaxies at different epochs of the Universe: the accretion of cold gas and the merging of galaxies (see Chapter 1).

Although there are many ways of quantifying the star formation at different ages of the Universe, measuring the rate at which galaxies merge through cosmic time is by far more intricate. Depending on which phase the merger is passing through, there are two main methods to estimate the incidence of such events at a given epoch. On-going and post-mergers could lead to the formation of post-starburst (PSB) galaxies (at least in groups or clusters; e.g., Wu et al. 2014; Lemaux et al. 2017) and leave morphological or kinematic imprint on the merging or coalesced system such as tails of stripped material, irregular shapes, disturbed velocity fields, and/or deep absorption lines in the spectra of PSBs (e.g., Le Fèvre et al. 2000; Conselice et al. 2003, 2008; Jogee et al. 2009; Wild et al. 2009; Casteels et al. 2014). However, observations of these features could be hampered by the spatial resolution needed to resolve the morphological details highlighting the presence of a merger (especially at high redshifts), or confused by diverse galaxy types that can mimic the kinematics or morphology of a merging system.

On the other hand, galaxy pairs that are going to merge in a certain timescale can also be identified based on some selection criteria regarding their spatial and velocity separation (e.g., Le Fèvre et al. 2000; Patton et al. 2000; Lin et al. 2008; de Ravel et al. 2009; López-Sanjuan et al. 2012; Ventou et al. 2017; Mantha et al. 2018; Duncan et al. 2019). Typically, close pairs must lie within a defined radius with a projected separation in the sky that could be as wide as 50 kpc and, in case of spectroscopic surveys, have a relative velocity not larger than Δv = 500 km s^{-1} (i.e., $\Delta z \sim 0.0017$) which excludes those pairs that are not gravitationally bound (e.g., Patton et al. 2000).

Both these methods can provide an estimate of the merger fraction through cosmic time but, as they probe different phases and timescales of the merging, they can lead to discordant results if not interpreted properly. However, even by adopting the same observational method, a large scatter is present in the literature for the merger fraction and the corresponding merger rate at all redshifts. This could be attributed to several reasons: i) the different criteria used for selecting galaxy pairs; ii) the distinction between major and minor mergers for which a dual uncertainty has to be taken into account, both in the pair ratio which can be obtained from the stellar mass (López-Sanjuan et al. 2013; Tasca et al. 2014; Man et al. 2016; Ventou et al. 2017; Duncan et al. 2019) or the flux (Bluck et al. 2012; Man et al. 2012, 2016; López-Sanjuan et al. 2015), and in the major (or minor) merger definition having a threshold in mass or flux ratio ranging between 2-6 (Xu et al. 2012; Ventou et al. 2017, 2019); iii) the merger timescale used to convert the pair fraction into a merger rate, which in turn depends on the method used for characterizing mergers, on their selection criteria, and on the properties of the galaxies undergoing a merger (Kitzbichler & White 2008; Jiang et al. 2014; Snyder et al. 2017).

Several studies tried to constrain the fraction of major mergers and the merger rate out to $z \lesssim 2$ (e.g., de Ravel et al. 2009; López-Sanjuan et al. 2011, 2013; Lotz et al. 2011; Xu et al. 2012). During this time, the merger fraction is found to increase with redshift as $\propto (1+z)^m$, where $1 \lesssim m \lesssim 5$, depending on the luminosity, the stellar mass and the spectral type of galaxies, leading to a large scatter in the literature. Only a handful of works have extended these measurements to higher redshifts. For instance, Conselice & Arnold

(2009) used both pair counts and morphological studies to investigate the presence of mergers at $4 < z < 6$ and their contribution to the galaxy assembly, finding a possible peak in the merger fraction at $z \sim 3$ followed by a decline which suggested a deviation from the power-law trend found at lower redshifts. However, the results by Conselice & Arnold (2009) do not distinguish between major and minor mergers and could have been subject to large uncertainties due to photometric redshifts. A peak in the cosmic evolution of the merger fraction is also found at $1 \lesssim z \lesssim 3$ by more recent observational studies (e.g., Mantha et al. 2018; Duncan et al. 2019) which, however, do not provide conclusive statements given the large above-mentioned differences on the selection and definition of the mergers (see for example the results by Mantha et al. 2018 for both stellar mass- or flux-selected pairs).

In this chapter we take advantage of the recent data collected by the ALPINE project (Béthermin et al. 2020; Faisst et al. 2020b; Le Fèvre et al. 2020) to further constrain the importance of major mergers to the galaxy mass-assembly at the end of the Reionization epoch and through cosmic time. This survey allows us to add a key additional piece of information to the wealth of multi-wavelength data already available for each targeted galaxy thanks to the ALMA observations of the ionized carbon [CII] at 158 μm rest-frame and the surrounding continuum. The 3D (i.e., RA, Dec, velocity) information enclosed in each ALPINE data cube is of great importance for identifying morphological and kinematic disturbances highlighting the presence of possible merging components. In particular, Le Fèvre et al. (2020) made a first fundamental step in the morpho-kinematic classification of the ALPINE galaxies, finding a high fraction ($\sim$ 40%) of mergers (both minor and major) at $z \sim 5$ among the ALPINE sample (see Section 2.2). An in-depth analysis of two merging systems found in ALPINE at $z \sim 4.6$ was made later by Jones et al. (2020) and Ginolfi et al. (2020a), showing the strength of these observations in the identification and characterization of this kind of sources in the high-redshift Universe.

Previous works estimated the merger fraction by exploiting morphological information or direct pair counts obtained through rest-frame optical observations and photometric or spectroscopic redshifts, as derived from optical ionized gas tracers (e.g., Conselice & Arnold 2009; Xu et al. 2012; Ventou et al. 2017). Here, for the first time, we make use of the rest-frame FIR [CII] line to compute the major merger fraction in the early Universe. As this line is less affected by dust extinction than optical tracers, morphological and kinematic analysis of this emission can reveal the presence of mergers with dusty components that are partially or completely obscured in the optical bands. Following these previous studies, we thus refine the fraction of mergers at the redshifts explored by the ALPINE survey and analyze in detail their [CII] emission combined with optical data for estimating their incidence at $z > 4$ and, along with other values in the literature, through cosmic time.

5.2 METHODS

To examine the merger fraction at $z \sim 5$, we take advantage of the continuum and [CII] emission maps of the targets from the ALPINE data release 1 (DR1; Béthermin et al. 2020)[1]. Briefly, for each source the CASA task `UVCONTSUB` was used to obtain line-only cubes. The continuum was first estimated in the uv-plane by masking the spectral channels containing line emission[2] and fitting a model to the visibilities, and it was then subtracted to the data cubes. With the CASA `IMMOMENTS` task the spectral channels that included the emission line were summed up to produce the velocity-integrated intensity (moment-0) maps, i.e. $M_0(x, y) = \Delta v \sum_i I_i(x, y)$, where I_i is the intensity of the i-th pixel at the position (x, y), Δv is the velocity width between two consecutive channels, and the sum is over all the channels containing the spectral line (Béthermin et al. 2020). By masking all the pixels below 3σ (where σ is the rms estimated in the moment-0 map), we also compute the intensity-weighted velocity (moment-1) and velocity dispersion (moment-2) maps, defined as $M_1 = \sum_i I_i v_i / M_0$ and $M_2 = \sqrt{\sum_i I_i (v_i - M_1)^2 / M_0}$, respectively. It is worth noting that in the analysis by Le Fèvre et al. (2020) the morphology and kinematic of the [CII] line were inspected within the 2σ emission contours of the moment maps. In this work, we decide to select only the pixels above the 3σ level in order to be less affected by possible spurious emissions that could alter the effective morphology and velocity field of the targets (however, we find that this choice does not significantly change our final results).

With the CASA `IMFIT` task we fit one 2D Gaussian component to each source in the

---

[1] DR1 data are available at `https://cesam.lam.fr/a2c2s/data_release.php`.

[2] For a detailed description of the procedure used to estimate the range of spectral channels containing line emission for all the [CII]-detected galaxies, see Section 6.1 in Béthermin et al. (2020).

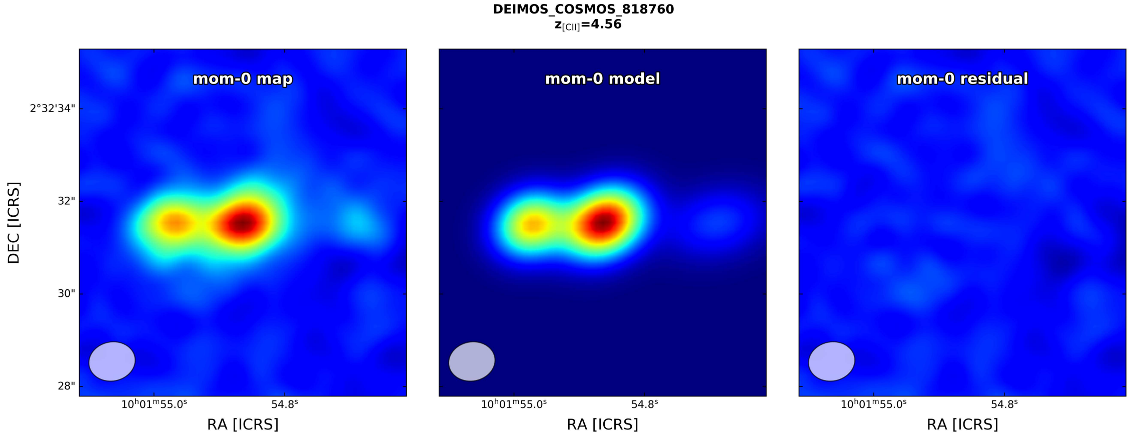


Figure 5.1: Example of 2D Gaussian fit as obtained through the CASA `IMFIT` tool for the ALPINE merger DC_818760. *Left:* moment-0 map centered on the rest-frame UV position of the main ALPINE target and showing the [CII] emission arising from three different components undergoing a merging. *Center:* [CII] modeling obtained by fitting three 2D Gaussian components to the observed moment-0 map. *Right:* residuals obtained by subtracting the model from the observed map. We also show the ALMA synthesized beam in the bottom left corner of each panel.

moment-0 map, retrieving best-fit parameters such as the peak intensity and integrated flux density, and morphological information like the coordinates of the [CII] emission peak, the FWHM of the major and minor axis of the Gaussian and its PA. If the [CII] morphology shows the presence of different spatially-resolved components in the vicinity of the main ALPINE target (as in case of mergers), we fit an individual 2D Gaussian to each of them in order to retrieve the above morphological and kinematic information for all the observed sub-structures. An example of this is shown in Figure 5.1, where we report the modeling of the ALPINE target DEIMOS_COSMOS_818760 (DC_818760). As evidenced by Jones et al. (2020), this is likely a triple merging system (see also Appendix A). Therefore, in this case, we use a three-component 2D Gaussian model to fit the observed [CII] emission visible from the moment-0 map. We then use the morphological central position provided by the 2D Gaussian best-fit on each ALPINE target[3] to produce the PVDs along the major and minor axes with the CASA `IMPV` task, with an averaging

---

[3] In case of multiple spatially-resolved [CII] emissions, we use the position provided by the Gaussian fit on the global system rather than the positions obtained by fitting individual 2D Gaussians to each of the observed components.

width of five pixels and a pseudo-slit length of 3″ (if not otherwise stated). Conversely to the analysis made by Le Fèvre et al. (2020) in which the major axis was aligned with the morphological PA, we decide to produce the two PVDs with the major axis oriented along the direction of major velocity gradient in the moment-1 map, and with the minor axis perpendicular to the major one. In this way, we increase the chances of identifying possible mergers or rotating galaxies in the ALPINE sample[4].

Lastly, we extract the spatially-integrated 1D [CII] spectrum of each target within the 3σ contours at the position of the source from the moment-0 map using the CASA task `SPECFLUX`.

5.2.1 MERGERS CHARACTERIZATION

Figure 5.2 shows two examples of the data analyzed for each [CII]-detected source. Taking as a reference the left set of panels, we display at the top the [CII] morphology superimposed on the HST/ACS F814W (left) and UVista K_s (right) cutouts centered on the UV rest-frame coordinates of the ALPINE target. We also show the $> 3\sigma$ emission contours from the optical images. In the second row, the moment-1 and moment-2 maps are reported color-coded for the velocity and velocity dispersion, respectively. In the velocity field map, the major and minor axes used to produce the underlying PVDs (color-coded for the flux intensity in Jy/beam) are shown. The axes are centered on the coordinates retrieved from the 2D Gaussian fit to the intensity map and are represented with the direction along which the PVDs are computed. The PA of the major axis is measured from North through East. Both the moment-1 and moment-2 maps are smoothed by using a bilinear interpolation. Finally, the lower panel reports the [CII] spectrum (in black) extracted from the continuum-subtracted data cube within the 3σ contours of the corresponding moment-0 map, and modeled with a single and a multiple-component Gaussian function (in this latter case, we show also the possible single components shaping the global spectral profile).

The described panels report the morpho-kinematic analysis of one of the sources classified as merger both by Le Fèvre et al. (2020) and by the analysis undertaken in this

---

[4] Two merging systems with different velocity fields or a rotating galaxy with a clear velocity gradient would be ideally represented with separated components in the PVDs and with an ‘S-shape’ in the PVD along the major axis, respectively.

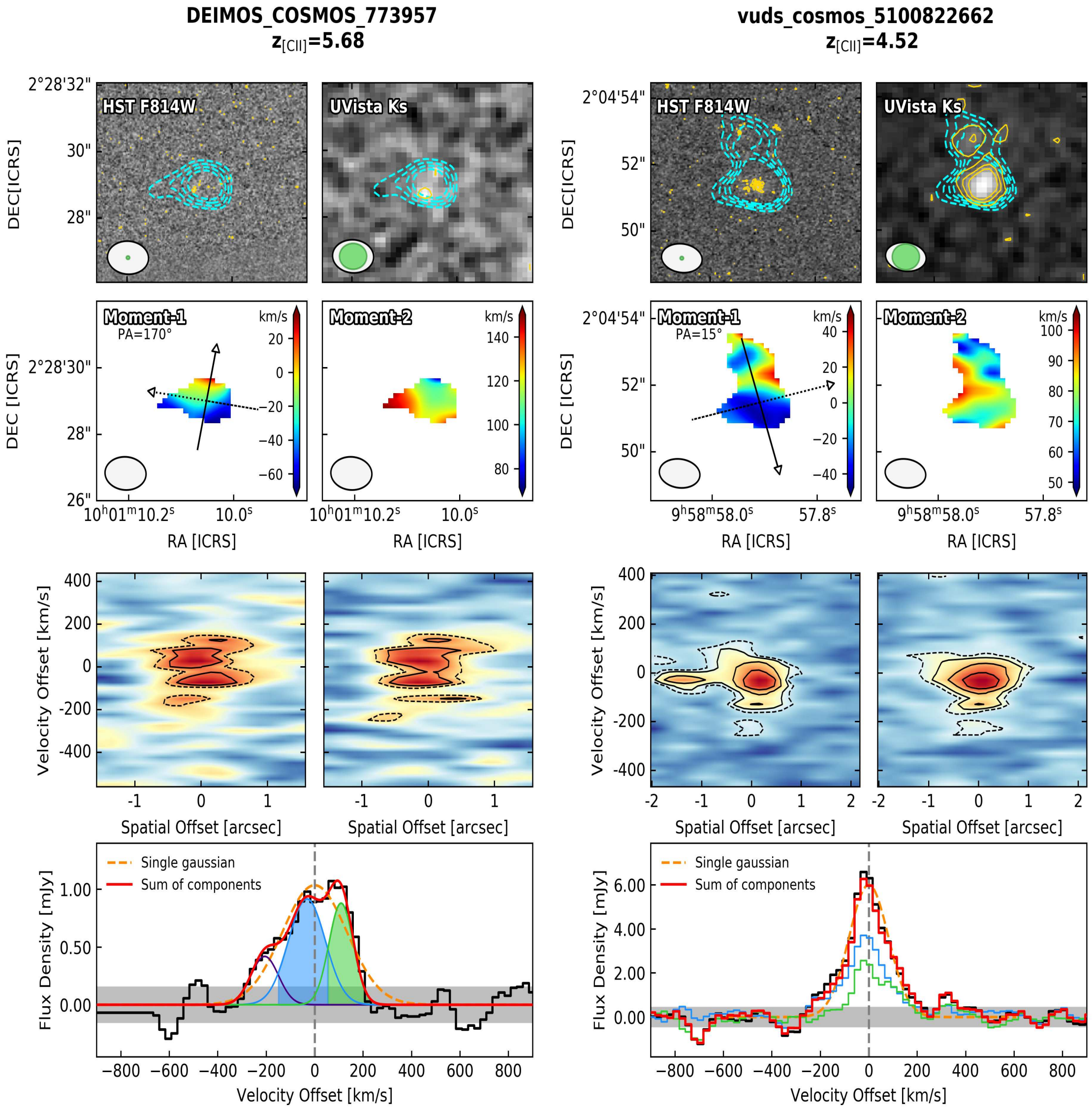


Figure 5.2: *Left:* Morpho-kinematic analysis of the ALPINE target DC_773957 at $z_{[CII]} = 5.68$. *First row*: HST/ACS F814W (left) and UVista DR4 K_s-band (right) images centered on the UV rest-frame position of the target. Each cutout is $6'' \times 6''$ wide. The cyan contours show the [CII] ALMA emission starting from 3σ above the noise level. Yellow contours in the optical maps represent 3, 5 and 7σ emission. In the lower left corner, the ALMA beam (white) and HST or UVista resolutions (green) are displayed. *Second row:* moment-1 (left) and moment-2 (right) maps color-coded for the velocity and velocity dispersion in km s^{-1}. The velocity map reports the direction of the major (solid) and minor (dashed) axis (centered on the coordinates returned by the best-fit 2D Gaussian model on the moment-0 map) along which the PVDs are computed.

Figure 5.2 ***(cont.)*** **:** *Third row:* PVDs along the major (left) and minor (right) axis color-coded for the flux intensity in each pixel. Dashed contours include the 2σ emission in the maps while 3, 5 and 7σ emission is represented by solid lines. *Fourth row:* [CII] spectrum (black histogram) extracted within the 3σ contours of the intensity map. The gray shaded band marks the 1σ level of the spectrum while the dashed vertical line shows the zero velocity offset computed with respect to the redshift of the [CII] line. Purple, blue and green lines are three individual possible components of the merging system, resulting in the global profile in red. The shaded areas under the curves represent the channels used to compute the [CII] intensity maps of the corresponding individual components. A single Gaussian fit is also visible with a dashed-orange line. *Right:* Morpho-kinematic analysis of the ALPINE target vc_5100822662 at $z_{[\mathrm{CII}]} = 4.52$. Same panels as in the left figure. In the bottom panel we report the [CII] spectra of the major (blue) and minor (green) merger components extracted at the positions of the two resolved emissions, as further described in the text.

work[5], i.e., DEIMOS_COSMOS_773957 (DC_773957) at $z_{[\mathrm{CII}]} = 5.68$. There are many lines of evidence for considering this galaxy in a merging phase. First, the [CII] emission is elongated toward the East at $> 3\sigma$, resulting in a clearly disturbed morphology. Furthermore, two peaks of emission at 3σ are resolved in the UVista K_s band, which could indicate a close pair of objects or the presence of a dust screen. The velocity field is also quite disturbed, with the PVD along the major axis presenting two 3σ components (solid contours) separated in velocity by ~ 140 km s^{-1} and with a slight spatial offset. These two components are also visible in the [CII] spectrum with green and blue lines. A third minor component at ~ -200 km s^{-1} could be instead due to a faint satellite nearby the ALPINE target and it is fairly visible in the PVDs within the 2σ contours. It is worth noting that, even though with a single Gaussian function we find an integrated flux of the line that is comparable with that obtained from the multi-component fit, the above-mentioned three components are needed in order to reproduce the observed [CII] spectrum.

Moreover, to obtain a more robust sample of mergers we compute, for each of the analyzed merger candidate, the rms of the observed spectra and use it to compute the SNR of the possible merging components. In particular, we obtain the integrated [CII] fluxes of the two principal Gaussian components of each spectrum and divide them by the product between the rms and the corresponding FWHM. We thus consider as mergers

---

[5] Note that Jones et al. (2021) classified this target as 'uncertain', following their classification criteria.

only those systems for which the minor (in terms of the [CII] integrated flux) component has SNR ≥ 3. The SNRs for the major and minor components of each system are reported in Table 5.1.

Another example of merging galaxies is provided on the right of Figure 5.2 by the morpho-kinematic analysis of vuds_cosmos_5100822662 (vc_5100822662) at $z_{\rm [CII]} = 4.52$. This source was classified as a merger both by Le Fèvre et al. (2020) and Jones et al. (2021). Indeed, two close and spatially resolved components are clearly visible in the moment-0 and optical maps. From the PVD along the major axis[6], the two merging systems are at the same velocity but spatially separated by $\sim 1.5''$ (i.e., $\sim$ 10 kpc at $z \sim 4.5$). In such a case, it is not possible to compute the contribution supplied by each individual source to the global shape of the line directly from the [CII] spectrum. However, we can provide an estimate of the integrated flux from the spectrum of each single component by modeling the total [CII] emission in the moment-0 map with two 2D Gaussian functions. In this way, we can extract the spectrum of the resolved component from the continuum-subtracted data cube at the position and with the shape provided by the best-fit parameters of the model (i.e., centroid, major and minor FWHM). For this specific ALPINE target, the spectra of the two merging components are shown in the bottom row of the figure as the blue and green histograms. As can be seen, their sum is quite consistent with the shape of the original [CII] spectrum.

A complete description of each individual merging system found from this work in the ALPINE survey is reported in Appendix A.

5.2.2 COMPARISON WITH PREVIOUS CLASSIFICATIONS

With the methodology introduced in the previous section, we proceed for a new morpho-kinematic classification of the ALPINE targets. As the main aim of this work is to find the contribution of mergers to the galaxy mass-assembly at $z \sim 5$ from the ALPINE survey, we particularly focus on this class of sources, in order to obtain the final sample of mergers needed to estimate the major merger fraction in the early Universe.

As done by Le Fèvre et al. (2020), we visually inspect the ancillary data, the intensity maps, and the velocity and velocity dispersion fields presented in Section 5.2.1 to search

---

[6] In this case, we use a pseudo-slight of 4" of length in order to cover all the 3σ emission from the velocity map.

Table 5.1: Physical parameters of the mergers. The source name is listed in column 1. Columns 2 and 3 report the redshifts of the two merger components as computed from the observed spectral features. In columns 4 and 5 we list their differences in velocity (in km s^{-1}) and their projected distances (in kpc). The ratios between the [CII] fluxes and the K_s-band fluxes (when available) of the two merger components are provided in columns 6 and 7, respectively. The right side of the table shows the measurements of FWHM (in km s^{-1}), [CII] emission size (in kpc), and SNR for both the minor and major components, intended as such in terms of their integrated [CII] fluxes. The asterisk near the size of some of the sources indicates a bad r_e estimate of the best-fit parameters of the 2D Gaussian fit on the moment-0 maps of the individual components.

							Minor			Major		
Source (1)	z_1 (2)	z_2 (3)	Δv [km s^{-1}] (4)	r_p [kpc] (5)	$\mu_{[\mathrm{CII}]}$ (6)	μ_K (7)	FWHM [km s^{-1}] (8)	r_e [kpc] (9)	SNR (10)	FWHM [km s^{-1}] (11)	r_e [kpc] (12)	SNR (13)
CG_38	5.5731	5.5698	152.3	2.7	1.8	-	96.9	1.4	3.8	165.9	2.0	3.9
DC_308643	4.5238	4.5221	92.3	3.0	1.2	-	85.4	1.0	4.5	85.6	1.4	5.3
DC_372292	5.1345	5.1374	144.0	1.8	2.7	2.9	82.8	2.2	4.1	157.9	3.7	5.7
DC_378903	5.4311	5.4293	94.5	0.9	1.9	-	70.0	4.4$^{(*)}$	3.6	111.2	1.3	4.2
DC_417567	5.6676	5.6700	106.7	5.3	2.1	-	84.8	1.0	3.1	163.9	4.0	3.4
DC_422677	4.4361	4.4378	92.8	0.0	1.4	-	95.7	1.0	4.7	93.1	0.5	6.7
DC_434239	4.4914	4.4876	206.1	5.5	6.9	-	103.5	3.9	4.5	441.8	3.4	7.2
DC_493583	4.5122	4.5141	103.4	1.0	4.7	-	64.8	3.3	3.6	139.5	1.0	8.0
DC_519281	5.5731	5.5765	158.0	0.9	4.5	-	95.7	0.9	3.1	163.8	2.3	8.3
DC_536534	5.6834	5.6886	234.6	4.4	2.7	-	113.9	1.9	3.6	228.5	2.6	4.9
DC_665509	4.5244	4.5261	95.9	1.0	2.0	-	70.5	4.5$^{(*)}$	3.0	107.0	2.4	4.0
DC_680104	4.5288	4.5308	106.4	0.0	1.0	-	89.5	0.5	3.7	111.6	8.7$^{(*)}$	4.7
DC_773957	5.6802	5.6770	141.2	3.0	1.6	5.6	118.0	1.3	5.9	179.5	1.7	6.2
DC_814483	4.5823	4.5796	145.5	7.9	2.4	-	128.1	1.1	3.3	192.7	2.8	4.2
DC_818760	4.5626	4.5609	92.3	9.9	1.3	2.6	127.1	2.4	10.4	271.0	2.9	18.3
DC_834764	4.5076	4.5055	119.0	1.0	1.7	-	96.2	1.4$^{(*)}$	5.1	143.8	3.1	5.8
DC_842313	4.5547	4.5406	751.5	11.6	32.9	1.5	229.9	4.3$^{(*)}$	5.1	1089.7	1.7	5.9
DC_859732	4.5353	4.5315	205.2	4.9	3.9	-	55.4	1.7$^{(*)}$	3.0	176.4	3.3$^{(*)}$	3.7
DC_873321	5.1545	5.1544	4.5	6.5	1.2	3.1	171.2	1.3	4.9	219.3	2.8	7.6
vc_5100541407	4.5628	4.5628	1.9	13.8	1.6	1.4	203.8	4.4$^{(*)}$	3.5	166.5	1.2	6.9
vc_5100822662	4.5210	4.5205	22.3	10.9	1.6	1.7	189.6	4.0	5.1	192.7	2.0	8.0
vc_5101209780	4.5724	4.5684	217.3	10.8	4.1	2.5	126.1	2.3	4.4	303.9	6.6	7.4
vc_5180966608	4.5294	4.5293	8.9	7.2	3.0	3.7	243.6	1.8	3.4	231.6	4.3	10.9

for the presence of multiple components or disturbed morphology near the position of the targets. The channel maps, the spectra and the PVDs are checked together searching for consistent emission features. By taking into account the results of the initial qualitative classification by Le Fèvre et al. (2020) and of the more recent quantitative analysis of a sub-sample of the ALPINE targets by Jones et al. (2021), we proceed with a more in-depth characterization of the [CII]-detected galaxies aimed at obtaining a robust merger fraction at $z \sim 5$. Adopting the same criteria described in Section 2.2 to differentiate the targets and considering the SNR of the minor merger component as described in Section 5.2.1, we find a slightly lower fraction ($\sim$ 31%, 23 out of 75 [CII]-detected sources) of mergers[7] if compared to the 40% found by Le Fèvre et al. (2020), with 12, 20 and 7% of the sample made by rotating, extended and compact dispersion dominated sources, respectively. To be more conservative in the classification of the galaxies (especially for obtaining a more robust merger statistics), we define the remaining 30% of the sample as 'uncertain', a new category that includes the weak galaxies (as described in Le Fèvre et al. 2020) and also objects that, by visual inspection, present features that are intermediate to those of various classes. This category is similar to the 'uncertain' (UNC) class introduced in Jones et al. (2021) that, based on the results of the ^{3D}Barolo fits, contains sources they are unable to classify because of the low SNR and/or spectral resolution, or contrasting evidence in their classification criteria. Although the uncertain category is populated by a significantly larger fraction of sources with respect to the weak class ($\sim$ 16%) in Le Fèvre et al. (2020), we recover the same qualitative morpho-kinematic distribution of the previous analysis, confirming the high fraction of rotators and mergers at these early epochs.

5.2.3 Physical parameters estimate

For each source classified as a merger, we measure several physical quantities that help us to estimate the merger fraction at $z \sim 5$. From the spectra, we compute the [CII] flux

---

[7] Some of the mergers found by Le Fèvre et al. (2020) were also analyzed in terms of the tilted ring model fitting by Jones et al. (2021) who found that the kinematic of those sources could be even compatible with rotating disks or dispersion dominated galaxies, or that the sensitivity and resolution of the data are too low for a conclusive classification (see Appendix A).

ratio between the major and minor components of the merger[8], i.e. $\mu_{[\mathrm{CII}]} = F_{1,[\mathrm{CII}]}/F_{2,[\mathrm{CII}]}$, finding a good agreement with the corresponding ratio computed from the PVDs. To obtain the integrated fluxes of the individual galaxies, we use a Gaussian decomposition of the global [CII] spectrum, as described in Section 5.2.1. When allowed by the resolution, we also take the UVista K_s-band flux ratio of the two merging sources, i.e. $\mu_K \equiv F_{1,K}/F_{2,K}$, where $F_{1,K}$ and $F_{2,K}$ are the integrated fluxes obtained by fitting a 2D Gaussian function to the major and minor components of the merger in the considered band, respectively. Then, from the spectral features, we also obtain the separation in velocity (Δv) between the intensity peaks of the merger components, and the FWHM both for the individual objects and single (overall spectrum) source.

We measure the projected distance (r_p) between the centers of the merger components as $r_p = \theta \times d_A(z_m)$, where θ is the angular separation in the sky (in arcsec) between the two galaxies and $d_A(z_m)$ is the angular diameter distance (in kpc arcsec^{-1}) computed at the mean redshift z_m of the two sources. In case the two components are not spatially resolved in [CII], we consider the distance between the two best-fit centroids returned by the 2D Gaussian fit to their moment-0 maps. Similarly to what has been done in Ginolfi et al. (2020a), the latter maps are obtained by collapsing the channels including the emission of the individual components. Such a case is shown, for instance, in Figure 5.2 for the target DC_773957. The shaded blue and green areas under the curves in the spectrum mark the collapsed channels used for building the [CII] intensity maps of the two components. As it can be seen, the velocity channels assigned to each component are not overlapping with each other. If instead the two objects are spatially resolved but at the same velocity (as for vc_5100822662), we are not able to create the [CII] moment-0 maps of the individual components (i.e., the two sources emit at the same frequency and we cannot disentangle their contribution in [CII]) and we measure the distance directly from the [CII] intensity map or from the PVDs.

Finally, we compute the size of the single components by fitting a simple 2D Gaussian model to the [CII] intensity maps obtained as described above. Following Fujimoto et al. (2020), we refer to the merging component size as the circularized effective radius r_e defined as $r_\mathrm{e} \equiv \sqrt{a\,b}$, where a and b are the best-fit semi-major and semi-minor axis,

---

[8] In the case the merging system is composed by more than two sources, we only consider the two major components in terms of their [CII] fluxes.

respectively. When the source is resolved, we use the best-fit beam-deconvolved a and b parameters, otherwise we take the beam-convolved sizes provided by the fit and correct them for the ALMA beam (the average beam minor axis of ALPINE is $\sim 0.8''$). It is worth noting that Fujimoto et al. (2020) measured the [CII] effective radii by fitting the line visibilities with the CASA task `UVMULTIFIT` and assuming an exponential-disk profile. However, by applying our more simplistic size measurements to the ALPINE galaxies analyzed by Fujimoto et al. (2020) and considering the mergers as single components, we found a good agreement with their results. A larger scatter is found only for some of the objects classified as mergers whose size measurements are flagged as unreliable by Fujimoto et al. (2020), likely because of the complicated [CII] morphology.

From now on, we refer to $r_{\rm e}$ when talking about the size of a source. All information presented in this section is reported in Table 5.1.

5.2.4 The merger sample

We show in Figure 5.3 (upper panel) the difference in velocity between the merger components as a function of their projected distance r_p for all the possible merging systems in ALPINE. In the upper-left corner, we report the uncertainties assumed for these data. The error on Δv is obtained by the sum in quadrature of the average uncertainty on each spectral element (i.e., 25 km s^{-1}). To estimate the typical error on the projected distance, we consider two random centroids corresponding to the initial positions of the minor and major components of the merger. Then, we obtain the distribution of the errors on the centroids as provided by the 2D Gaussian best-fits on the intensity maps of the merging components, and extract from that N = 1000 values used to perturb the initial positions. At each iteration, we re-compute the projected distance between the perturbed positions of the two components. We thus estimate the average uncertainty on r_p as the standard deviation of the distribution of the N projected separations (corresponding to $\sim$ 0.3 kpc).

To derive a robust major merger fraction ($f_{\rm MM}$), we decide to consider as mergers only those systems in which the components are separated by no more than $\Delta {\rm v} \leq 500$ km s^{-1}, assuming that this is the limit for a system to be gravitationally bound (e.g., Patton et al. 2000; Lin et al. 2008; Ventou et al. 2017). As shown, all the mergers in our sample satisfy

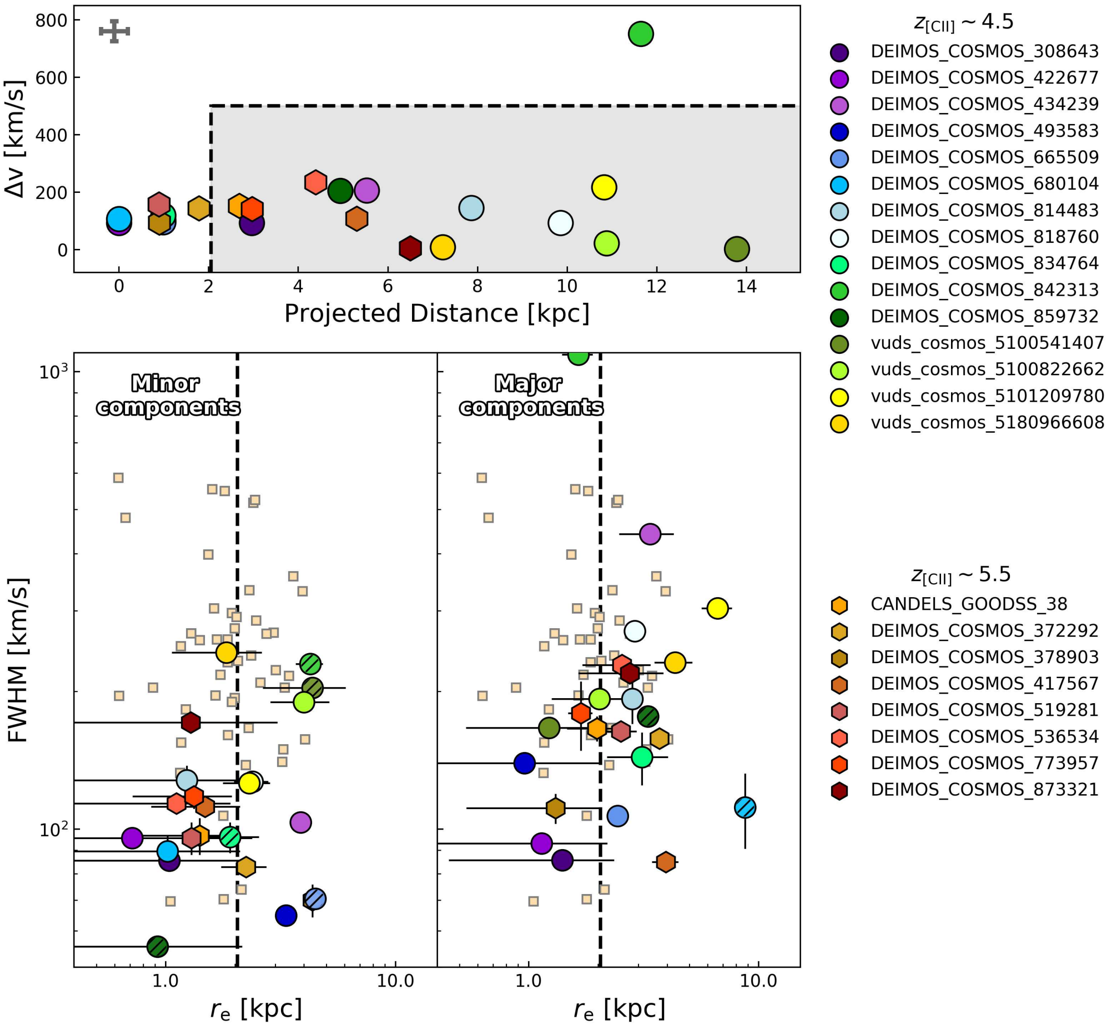


Figure 5.3: *Top panel:* Velocity separation between the components of each ALPINE merger as a function of their projected distance. Each target is represented by its own color while the symbols are different for $z \sim 4.5$ (circles) and $z \sim 5.5$ (hexagons) galaxies. The gray shaded region shows the area of the plot populated by galaxies which are more likely to be real mergers, with a projected distance larger than > 2.1 kpc and $\Delta v < 500$ km s^{-1}. The error bar on the top-left corner represents the typical error on each value. *Bottom panel:* FWHM as a function of r_e for the minor (left) and major (right) components of each merger. The colors and symbols are the same as in the top panel. Hatched markers identify the sources for which the best-fit size estimate is not reliable, after a visual inspection of the residuals. The dashed black lines mark the average [CII] size of the ALPINE targets (as estimated in this work), represented with small light squares in the background.

this constraint except one source. As analyzed in detail in Appendix A, this system is composed by the ALPINE target DC_842313 and a peculiar neighbour source with a large FWHM and [CII] flux, and could even be a merger of three galaxies with a Δv between the target and the further third component smaller than 500 km s^{-1}. We thus account for this object in the computation of the merger fraction.

Also, we consider as reliable mergers only those systems with a projected distance larger than the typical [CII] sizes of individual galaxies at these redshifts. Indeed, closer components could just represent clumps of star formation within the same galaxy, faking the presence of a merger by affecting the morphology and kinematics of the [CII] line. From the size measurements of the ALPINE galaxies we find an average [CII] size of $\sim$ 2.1 kpc, that is in good agreement with the median value found by Fujimoto et al. (2020) and with the typical distance between galaxies and satellites or mergers (>2 kpc) at $z > 4$ (e.g., Carniani et al. 2018; Whitney et al. 2019; Zanella et al. 2021). Therefore, we use this value as a threshold to classify the secure merging systems.

The two selection criteria on the distance and velocity separations define the gray region in Figure 5.3 which thus includes the sources that are classified as robust mergers. However, we cannot exclude *a priori* the systems with $r_p < 2.1$ kpc as they could be structures in an advanced stage of merging that we are not able to resolve with ALMA because of their close proximity and observational limitations (i.e., large synthesized beam, limited sensitivity). It is worth noting that, in some cases, spatially close clumps are observed with HST within the typical ALPINE beam size. However, the lack of redshift information for these sources prevent us from considering them as associate galaxies experiencing a merging. This could have an impact on our results, possibly reducing the estimated major merger fraction (and, consequently, the merger rate) at $z \sim 5$. For these reasons, we also show in Figure 5.3 (bottom panels) the individual sizes of the minor and major components of each of the possible mergers as a function of their FWHM. The colors and marker symbols are the same as in the top panel, while the hatched markers highlight those objects with a bad r_e estimate from the fit, after a visual inspection of the residuals. We also show for comparison the data of the individual ALPINE galaxies (excluding mergers) as light squares. As it can be seen, the sizes of the individual components lie, within the uncertainties, among those of the ALPINE targets. This suggests that all the individual components we are analyzing could be merging galaxies and that, with the current data, we cannot exclude them from the computation of the merger fraction at $z > 4$.

Finally, we provide an estimate of the stellar mass range covered by our merger components. We only have information on the stellar mass of the merging system as a whole through SED-fitting measurements (Faisst et al. 2020b) because the majority of the merger candidates are not resolved in most of the optical and NIR bands. We make the assumption that $M_{*,1} + M_{*,2} = M_*$ and define the mass ratio as $\mu = M_{*,1}/M_{*,2}$, where $M_{*,1}$ and $M_{*,2}$ are the stellar masses of the primary and secondary galaxy (i.e., $M_{*,1} > M_{*,2}$) and M_* is the total mass of the system from the SED-fitting. Then, we set $\mu = 4$ as a threshold to separate major ($1 \leq \mu \leq 4$) from minor ($\mu > 4$) mergers (e.g., Lotz et al. 2011; Mantha et al. 2018; Duncan et al. 2019) and use this value to solve the above equations for $M_{*,1}$ and $M_{*,2}$. With the assumptions above, we obtain a stellar mass distribution of the major components ranging between $\log(M_{*,1}/M_{\odot}) \sim 9.0$ and 10.7, with a mean value $\log(M_{*,1}/M_{\odot}) \sim 10$ (reducing by ~ 0.2 dex if assuming $\mu = 1$). We want to stress that this is just an exercise that is intended to provide an estimate of the stellar mass range of the merger components we are analyzing and that helps us in the comparison of our estimates with those of previous works. Indeed, we are not considering that some of the merger components are resolved in the K_s band and have their own mass ratio (see Section 5.3.1).

Therefore, following the above considerations, we can conclude that the close pairs of our sample are characterized by $r_p < 20$ kpc, $\Delta v < 500$ km s^{-1}, and average stellar masses $\log(M_*/M_{\odot}) \sim 10$.

5.2.5 Accounting for completeness

When deriving the merger fraction, we have to consider the possibility of missing a certain number of mergers because of the faintness of the minor component. Indeed, major mergers with a principal component near the threshold of the observable [CII] flux may imply the presence of a secondary component that could be lost by instrumental limitations. To account for this, we must correct for the incompleteness of our sample.

Previous works typically assumed corrections based on the limiting flux or stellar mass of the corresponding survey (e.g., Patton et al. 2000; López-Sanjuan et al. 2011;

Mundy et al. 2017; Duncan et al. 2019). ALPINE is not a flux-limited survey as each individual ALMA pointing reaches different depths (e.g., Béthermin et al. 2020). For this reason, we adopt the classical completeness corrections found in the literature but considering for each merger its limiting flux. In particular, we apply a weight to each minor component of our sample defined as

$$w_{\mathrm{comp}}(L,z) = \frac{\int_{L_2}^{L_1} \Phi(L,z)\, dL}{\int_{L_{\mathrm{lim}}}^{L_1} \Phi(L,z)\, dL}, \tag{5.1}$$

where $\Phi(L,z)$ is the [CII] luminosity function derived from the UV-selected central ALPINE targets (which is in turn corrected for completeness; Yan et al. 2020), L_{lim} is the luminosity corresponding to the limiting flux of each ALPINE pointing, L_1 is associated to the flux of the primary component $F_{1,[\mathrm{CII}]}$, and L_2 corresponds to $F_{2,[\mathrm{CII}]} = F_{1,[\mathrm{CII}]}/4$. We compute the [CII] luminosities as in Equation 3.1 by following Solomon et al. (1992). We also assume that $w_{\mathrm{comp}} = 1$ when $L_2 > L_{\mathrm{lim}}$. Such a correction corresponds to the ratio between the number density of galaxies with [CII] flux higher than $F_{2,[\mathrm{CII}]}$, and the number density of galaxies with a flux higher than the flux limit of each ALMA pointing.

5.3 RESULTS

5.3.1 THE FRACTION OF MAJOR MERGERS IN ALPINE

The aim of this work is to derive the fraction of galaxies undergoing a major merger at $z \sim 5$ and to estimate their contribution to the galaxy mass-assembly through cosmic time. Although the cosmic evolution of minor mergers is still poorly constrained at these redshifts, several studies suggest that they are more frequent in the nearby Universe, showing a decreasing fraction for $z \gtrsim 3$ (López-Sanjuan et al. 2011; Lotz et al. 2011; Ventou et al. 2019). We thus have to take into account a possible contamination from them in our sample. First, as we do not have the stellar mass information for the majority of our merger components, we define as major mergers those candidates for which the K_s-band flux ratio is smaller than 4, i.e., $1 < \mu_K < 4$. The K_s band is indeed a good tracer of the stellar mass of galaxies up to $z \sim 4$ (e.g., Laigle et al. 2016). At higher redshifts, the

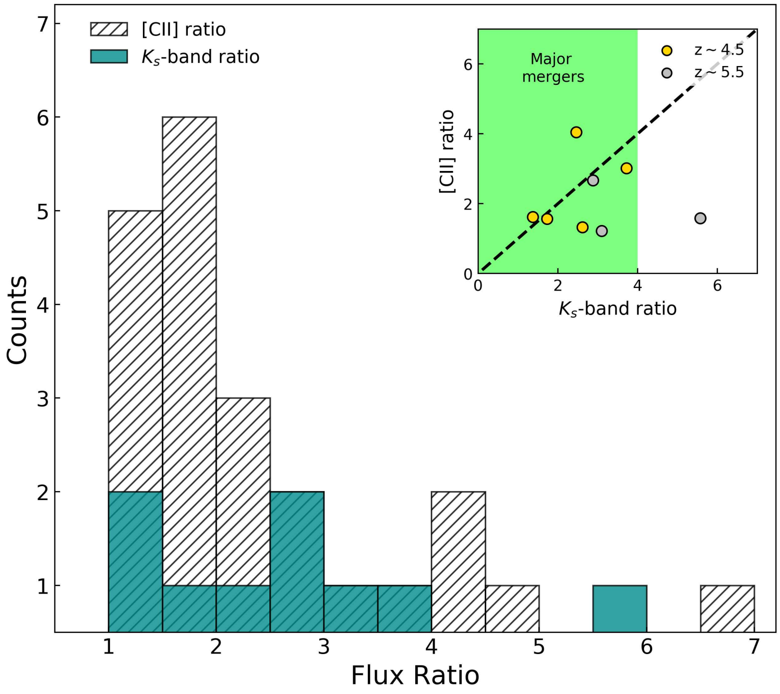


Figure 5.4: Distributions of the K_s-band (turquoise histogram) and [CII] (hatched histogram) flux ratios between the components of each merging system. *Inset:* comparison between the [CII] and K_s-band ratios for the sources having this information in common, both at $z \sim 4.5$ and $z \sim 5.5$ (represented as yellow and gray circles, respectively). The green area marks the region of the figure with $\mu_K < 4$. The dashed black line reports the 1:1 relation between the two quantities.

rest-frame K_s band samples the emission below the Balmer break (K_s-band corresponds to rest-frame ~ 3600 at $z \sim 5$, taking as a reference the mean wavelength of the filter) which is no more directly linked to the stellar mass of the galaxy. In these cases, mid-infrared bands could be better tracers of the galaxy stellar masses than the K_s filter, but their worst resolution do not allow us to spatially resolve the close galaxies for which we have the K_s fluxes. For comparison, López-Sanjuan et al. (2013) used the i-band ratios to identify major mergers. At the redshifts explored in their work ($0.9 < z < 1.8$), the observed i band corresponds to $\sim 3000 - 3500$ rest-frame (which is quite similar to the wavelength range covered by the K_s band at the highest redshift of our sample). They compared these ratios with those obtained from the K_s band (sampling the emission between 8000 and 10000 rest-frame, thus tracing better the stellar mass content of the galaxies), finding no significant changes in their major merger classification, and then supporting the i-band results. Therefore, following these arguments, we consider μ_K reasonably comparable to the stellar mass ratio of our sources, treating cautiously these estimates for those galaxies at the highest redshifts of our sample.

The K_s-band ratio is available for 9 out of 23 mergers, while for the remaining sources we only have the [CII] flux ratio, from which we cannot draw conclusions about the nature of the merger. These two ratios are shown in Figure 5.4 where it is evident that the majority of the sources have both μ_K and $\mu_{[\mathrm{CII}]}$ smaller than 4. Furthermore, in the inset we compare the [CII] and K_s-band flux ratios for the objects having these two measures in common[9]. There is a good agreement between the two ratios for some of the objects in the figure, but others show a large scatter from the 1:1 relation, suggesting that complex physical processes could take place in these systems. However, we find that 7 out of 8 sources (i.e., $\sim$ 88% of this sub-sample) have $\mu_K < 4$ while only one system shows a larger K_s-band flux ratio. Although a larger statistical sample would be needed, we can interpret this as an indication of the minor merger contamination of our sample. Indeed, if we assume that only the sources outside the green region contribute to the minor merger contamination (thus lying at $\mu_K > 4$), we can assume that the 12% of all the objects having only the $\mu_{[\mathrm{CII}]}$ information is affected by minor mergers (or, equivalently, that only the 88% of that sample contains major mergers). We note that, when excluding from this statistics the sources at $z \sim 5.5$ (whose K_s-band ratio is less reliable as a tracer of their stellar mass), all the objects in the inset of Figure 5.4 have $\mu_K < 4$, raising the fraction of major mergers in this sub-sample to 100%. This is taken into account later, when estimating the uncertainty on the merger fraction.

We define the major merger fractions in the two ALPINE redshift bins (at mean redshift $z \sim 4.5$ and 5.5, respectively) as

$$f_{\mathrm{MM}} = 0.88\,\frac{\sum_{j=1}^{\mathrm{N_p}} w^{j}_{\mathrm{comp}}}{\mathrm{N_g}}, \tag{5.2}$$

---

[9] Both in the main figure and in the inset we do not report the measure of the [CII] flux ratio for the merger DC_842313 (i.e., $\mu_{[\mathrm{CII}]} \sim 33$). We also do not consider this source in the statistics described in the text because, as elaborated in Appendix A, this is a peculiar source.

Table 5.2: Number of sources at $z \sim 4.5$ and $z \sim 5.5$ from the ALPINE survey used for the computation of the major merger fraction (whose values are reported in the last row) as in Equation 5.2. In particular, $\mathrm{N_g}$ is the number of ALPINE targets in the redshift bin, $\mathrm{N_p}$ is the number of observed mergers, and $\mathrm{N_{p,corr}}$ represents the number of mergers after correcting for incompleteness.

	$z \sim 4.5$	$z \sim 5.5$
$\mathrm{N_g}$	46	29
$\mathrm{N_p}$	15	8
$\mathrm{N_{p,corr}}$	23	11
f_{MM}	$0.44^{+0.11}_{-0.16}$	$0.34^{+0.10}_{-0.13}$

where $\mathrm{N_p}$ represents the number of all the mergers in our sample, $\mathrm{N_g}$ is the number of ALPINE galaxies in the considered redshift bin, and w^{j}_{comp} is the weight associated to each merger to correct for incompleteness (see Section 5.2.5). The factor 0.88 accounts for the statistical uncertainty on the real merger nature (the fact that the 12% of the $\mu_{\mathrm{[CII]}}$-only sub-sample could be affected by the presence of minor mergers). Note also that, as detailed in Section 5.2.4, we are not excluding any source with $r_p < 2$ kpc because the individual sizes of these merger components are comparable with those of the ALPINE galaxies.

To attribute an uncertainty to f_{MM} we consider two extreme cases. At first, as an upper limit to the major merger fraction, we count all the mergers in each redshift bin and divide them by $\mathrm{N_g}$, independently of their components projected distance and flux ratio. This uncertainty includes also the case in which our sample is not contaminated by minor mergers (i.e., all the analyzed mergers are major). As a lower limit instead, we only take the mergers whose components have $r_p > 2$ kpc, assuming that all the individual components with a smaller projected separation are not major mergers (see discussion above). Then, given the small statistics on both the total number of ALPINE galaxies and the sub-sample of mergers, we compute the corresponding Poissonian errors and add them in quadrature to the above-mentioned uncertainties.

With these assumptions and from Equation 5.2 we obtain $f^{z_4}_{\mathrm{MM}} = 0.44^{+0.11}_{-0.16}$ and $f^{z_5}_{\mathrm{MM}} = 0.34^{+0.10}_{-0.13}$, at $z \sim 4.5$ and $z \sim 5.5$, respectively. These results are shown in Figure 5.5 and reported in Table 5.2 along with the numbers used for their computation at both redshifts. We note that, by using Equation 5.2, we recover a merger fraction $\sim 30\%$ higher than obtained if not accounting for completeness corrections. It is also interesting that, assuming that 12% of the mergers are minor would imply a minor merger fraction of at least $\sim 5\%$, in agreement with the estimates by Ventou et al. (2019) at $z > 3$ (i.e., $0.08^{+0.07}_{-0.05}$). It should be noted that this fraction could be even higher. Indeed, the relatively low resolution of the ALPINE survey does not allow us to constrain the minor merger fraction at $z \sim 5$ (Le Fèvre et al. 2020), as we are not able to detect faint satellites that are instead expected to be in the neighborhood of $z > 4$ galaxies by simulations (e.g., Pallottini et al. 2017; Kohandel et al. 2019).

Finally, it is worth mentioning that $\sim 10\%$ of the ALPINE targets are part of a massive proto-cluster of galaxies (PCI J1001+0220) at $z \sim 4.57$ located in the COSMOS field (Lemaux et al. 2018). In principle, high-density environments like those in high-z protoclusters, or groups and low-mass clusters at lower redshift, could enhance the merging activity, resulting in a merger rate that could be several times larger than what expected in lower-density regimes (e.g., Lin et al. 2010; Lotz et al. 2013; Tomczak et al. 2017; Pelliccia et al. 2019). To account for this possible caveat, we compare the ALPINE members of the proto-cluster with the mergers in Table 5.1, finding that only two of them are part of the observed over-density. However, by removing these two sources from our analysis, we do not find any significant change in the merger fraction at $z \sim 4.5$ and, consequently, in the estimate of the merger rate.

5.3.2 Cosmic evolution of the major merger fraction

During the last years, a large number of studies focused on the estimate of the merger fraction at different redshifts and on its evolution through cosmic time (e.g., Le Fèvre et al. 2000; Conselice et al. 2003; Conselice & Arnold 2009; López-Sanjuan et al. 2013; Mundy et al. 2017; Ventou et al. 2017; Mantha et al. 2018; Duncan et al. 2019). However, the choice of various selection criteria on the spatial and velocity separation for the pair counts, the different mass ranges probed by each author, and the diverse threshold in μ to distinguish between major and minor mergers make a direct comparison between the many results found in the literature quite problematic. Therefore, we now compare the major merger fraction found from this work at $4 < z < 6$ with those computed from

Table 5.3: Description of the selection criteria and sample properties in previous pair counts studies and in this work. Column 1 lists the works with which we compare our f_{MM} estimates; column 2 reports the redshift range studied in those works or the redshift interval for which we take the major merger fraction estimates, accounting for the stellar mass of the primary galaxies in the merger (reported in column 6); column 3 provides the limit mass or flux ratio used to separate major from minor mergers; column 4 and 5 show the velocity separation and projected distance criteria adopted for the pair counts, respectively. The asterisks in column 4 indicate the works in the literature that made use of photometric redshifts to define the selection criterion on the velocity separation between pairs.

Literature (1)	Redshift range (2)	Major/Minor separation (3)	Δv [km s^{-1}] (4)	r_p [kpc] (5)	Primary Galaxy Mass (6)
de Ravel et al. (2009)	$0.5 < z < 0.9$	$F_1/F_2 \leq 4.0$	≤ 500	< 20	$\log(M_*/M_\odot) \geq 9.5$
Xu et al. (2012)	$0 \leq z \leq 0.6$	$M_1/M_2 \leq 2.5$	$\leq 500^{(*)}$	$5 - 20$	$9.8 < \log(M_*/M_\odot) \leq 10.2$
López-Sanjuan et al. (2013)	$0.9 < z < 1.8$	$F_1/F_2 \leq 4.0$	≤ 500	< 20	$9 < \log(M_*/M_\odot) < 11$
Tasca et al. (2014)	$1.8 \leq z \leq 4.0$	$M_1/M_2 \leq 4.0$	≤ 500	< 25	$9 < \log(M_*/M_\odot) < 11$
Ventou et al. (2017)	$0.2 \leq z \leq 4.0$	$M_1/M_2 \leq 6.0$	≤ 500	< 25	$\log(M_*/M_\odot) \geq 9.5$
Duncan et al. (2019)	$0.5 \leq z < 3.5$	$M_1/M_2 \leq 4.0$	$\leq 500^{(*)}$	$5 - 30$	$9.7 < \log(M_*/M_\odot) < 10.3$
This work	$4.4 \leq z < 5.9$	$M_1/M_2 \leq 4.0$	≤ 500	< 20	$9 \lesssim \log(M_*/M_\odot) \lesssim 11$

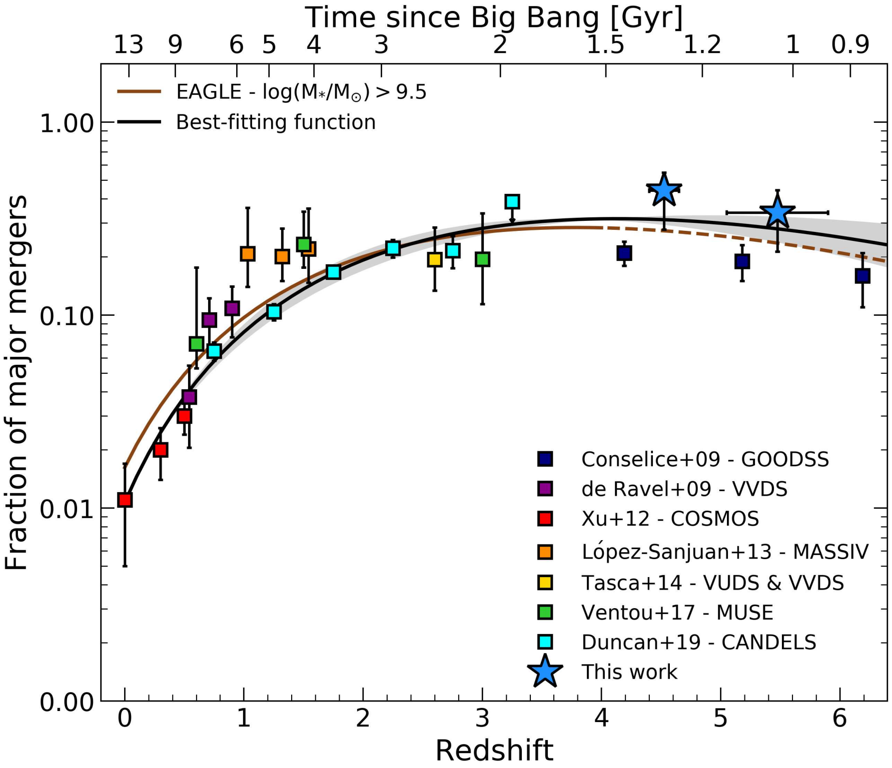


Figure 5.5: Cosmic evolution of the major merger fraction $f_{\rm MM}$ from the local to the early Universe. Colored squares show the data collected from the literature at different cosmic times through pair counts, both with spectroscopic (de Ravel et al. 2009; López-Sanjuan et al. 2013; Tasca et al. 2014; Ventou et al. 2017) and photometric (Xu et al. 2012; Duncan et al. 2019) redshifts, and/or morphological studies (Conselice & Arnold 2009). Blue stars are the $f_{\rm MM}$ estimates found in this work. The solid black line and the shaded region are the best-fit to the data with a combined power-law/exponential function and the associated 1σ error, respectively. Finally, the solid brown line illustrates the parameterized redshift evolution (up to $z = 4$) of the major merger fraction from the EAGLE simulation for galaxies with $\log(M_*/M_\odot) > 9.5$. The dashed line is just an extension of that curve to higher redshifts.

samples of galaxies with similar stellar masses and merger selection criteria. It is worth noting that the results of this work are derived from the ALPINE survey, and they mostly apply to relatively massive, rest-frame UV-selected galaxies. Therefore, the comparison with other data and models from the literature we show in the following, must be taken with caution as they could be biased by different selection techniques.

At $z < 1$, Xu et al. (2012) evaluated the major merger fraction from a sample of close pairs drawn from the COSMOS survey (Ilbert et al. 2009). From that work we select only the results in the stellar mass bin $9.8 < \log(M_*/M_\odot) \leq 10.2$, to be consistent with the average stellar mass of our sample. At this epoch we also account for the merger fraction computed by de Ravel et al. (2009) who analyzed a sample of spectroscopically confirmed galaxy pairs from the VIMOS VLT Deep Survey (VVDS; Le Fèvre et al. 2005) with $\log(M_*/M_\odot) > 9.5$. At $1 < z < 4$, López-Sanjuan et al. (2013) and Tasca et al. (2014) exploited the VUDS, VVDS and the Mass Assembly Survey with SINFONI in VVDS (MASSIVE; Contini et al. 2012) surveys to estimate f_{MM} from merging systems with spectroscopic measurements and average stellar masses $\log(M_*/M_\odot) \sim 10$. Ventou et al. (2017) observed the *Hubble* Ultra Deep Field (Beckwith et al. 2006) and the *Hubble* Deep Field South (Williams et al. 2000) with the Multi Unit Spectroscopic Explorer (MUSE) to constrain the galaxy major merger fraction up to $z \sim 6$. However, we take only their data points with $\log(M_*/M_\odot) \geq 9.5$ which extend up to $z \sim 3$. Finally, we get the pair fraction computed by Duncan et al. (2019) in a mass-selected ($9.7 < \log(M_*/M_\odot) < 10.3$) sample of galaxies drawn from the CANDELS survey (Grogin et al. 2011; Koekemoer et al. 2011) using a probabilistic approach based on their photometric redshifts. All these works assume a projected separation $r_p < 30$ kpc between the merger components and a difference in velocity $\Delta v \leq 500$ km s^{-1}. Most of them adopt $\mu < 4$ to identify the major merger population, with μ computed either as the ratio between the stellar mass of the primary and secondary galaxy in the system or as the difference in magnitude between the two components. The only exceptions are from Xu et al. (2012) and Ventou et al. (2017) that make use of a mass ratio of 2.5 and 6, respectively (however, as stated by Ventou et al. 2017, adopting a mass ratio limit of 4 implies the loss of only a few sources in their sample, not significantly altering the final statistics). Table 5.3 summarizes the main features of the samples from which the observational data at $z < 4$ introduced above are taken.

As shown in Figure 5.5, we combine these data with our measurements at $z \sim 5$ to provide the cosmic evolution of the major merger fraction. Traditionally, this is parameterized with a power-law fitting formula of the form

$$f_{\mathrm{MM}}(z) = f_0\ (1+z)^m, \tag{5.3}$$

Table 5.4: Best-fit parameters and 1σ uncertainties obtained by fitting the merger fractions and merger rates (computed with different merger timescales). For both quantities, we adopt a power-law plus exponential functional form of the type $\alpha(1+z)^m \exp(\beta(1+z))$.

$T_{\rm MM}$	α	m	β
	Merger fraction $f_{\rm MM}$		
-	0.024 ± 0.003	4.083 ± 0.501	-0.797 ± 0.189
	Merger rate $R_{\rm MM}$		
Kitzbichler & White (2008)	0.020 ± 0.003	4.282 ± 0.488	-1.036 ± 0.182
Jiang et al. (2014)	0.015 ± 0.002	4.350 ± 0.492	-0.740 ± 0.184
Snyder et al. (2017)	0.010 ± 0.001	6.083 ± 0.501	-0.797 ± 0.189

where f_0 is the merger fraction at $z = 0$ and m is the index that rules the redshift evolution. Our data suggest a slight decline of the merger fraction at $z > 4$, with a possible peak at lower redshift, as also previously found by other studies (e.g., Conselice & Arnold 2009; Ventou et al. 2017; Mantha et al. 2018; Duncan et al. 2019). Therefore, we use a combined power-law/exponential function to fit our observational data, such as

$$f_{\rm MM}(z) = \alpha\ (1+z)^m \exp(\beta(1+z)), \tag{5.4}$$

where β controls the exponential side of the curve and $f_0 = \alpha \exp(\beta)$. By using a nonlinear least square algorithm, we fit our data both with Equation 5.3 and Equation 5.4. Taking advantage of the Bayesian information criterion (BIC), we find a significant statistical evidence (ΔBIC > 10) in favor of the power-law/exponential form for describing the evolution of the merger fraction through cosmic time. These results are in contrast with those by Duncan et al. (2019) who, also relying on the BIC, found that there was no strong statistical evidence in favor of one of the two above-described parameterizations. This difference could be attributed to the fact that, for stellar masses $\log(M_*/M_\odot) \sim 10$, they measured the major merger fraction up to $z \sim 3$, while in this work our new ALPINE data allow us to constrain this quantity up to earlier epochs (i.e., $z \sim 5$).

For these reasons, we show in Figure 5.5 the best-fitting function to our data obtained through Equation 5.4, and use the results from this parameterization for the rest of the work. In particular, the parameters of this fit are $\alpha = 0.024 \pm 0.003$, $m = 4.083 \pm 0.501$ and $\beta = -0.797 \pm 0.189$ which are in good agreement with those found by Duncan et al. (2019) when fitting the CANDELS data points for galaxies with stellar mass $9.7 < \log(M_*/M_\odot) < 10.3$. For comparison, we also show the results obtained by Conselice & Arnold (2009) exploiting morphological analysis and pair counts for a sample of Lyman-break drop-out galaxies at $4 < z < 6$ with masses $\log(M_*/M_\odot) > 9{-}10$. Their data points are lower than our results but still comparable to them within the uncertainties. Anyway, as they do not differentiate between major and minor mergers, we do not include them in the computation of the cosmic evolution of the merger fraction.

We show in Figure 5.5 the best-fit parameterization of the major merger fraction found by Qu et al. (2017) for galaxies with $\log(M_*/M_\odot) \geq 9.5$ from the Evolution and Assembly of Galaxies and their Environments (EAGLE) hydrodynamical simulation (Schaye et al. 2015). They adopted a stellar mass ratio $\lesssim 4$ to identify major mergers and the same parameterization as in Equation 5.4, obtaining a cosmic evolution of $f_{\rm MM}$ that is similar to ours up to $z \sim 4$ (the higher redshift reached by the simulation). Extending their prediction to earlier epochs, a slightly milder redshift evolution with respect to our results is found. However, the two trends are still in good agreement with each other.

Finally, it is worth noting that we are recovering the cosmic evolution of the merger fraction by comparing galaxies with a constant stellar mass range at all redshifts. We are aware that, in this way, we are not necessarily tracing the same galaxy population through cosmic time, as instead achieved with constant cumulative comoving number density selections (Papovich et al. 2011; Conselice et al. 2013; Ownsworth et al. 2014; Mundy et al. 2015; Torrey et al. 2015)[10]. Nevertheless, as the majority of the observational and theoretical works in the literature make use of the constant stellar mass selection to estimate the fraction and rate of major mergers through cosmic time, such a choice allows us to make a fair comparison to these works and provides a quite easily quantifiable observational benchmark for future studies.

---

[10] Although this is no more assured in case of major mergers or strong changes in star formation (e.g., Leja et al. 2013; Ownsworth et al. 2014).

5.3.3 GALAXY MAJOR MERGER RATE

The major merger fraction can be translated into the merger rate $R_{\rm MM}$ (i.e., the number of mergers per galaxy and Gyr) as

$$R_{\rm MM}(z) = \frac{C_{\rm merg}\, f_{\rm MM}(z)}{T_{\rm MM}(z)}, \tag{5.5}$$

where $f_{\rm MM}(z)$ is the pair fraction at redshift z estimated in Section 5.3.2, and $C_{\rm merg}$ is the fraction of close pairs that will eventually merge into a single system in a typical timescale $T_{\rm MM}(z)$. This last term represents the major source of uncertainty in the merger rate computation and is usually obtained through simulations based on the dynamical friction timescales affecting the merging components (e.g., Kitzbichler & White 2008; Lotz et al. 2010; Jiang et al. 2014; Snyder et al. 2017).

Kitzbichler & White (2008) applied a semi-analytic model to the Millennium simulation (Springel et al. 2005) outputs finding an average merging time that depends linearly on the redshift and projected distance of the pair and weakly on the stellar mass of the main galaxy, i.e., $T_{\rm MM} \propto r_p\, M_*^{-0.3}\, (1+z/8)$. This relation is valid at $z \leq 1$, while at higher redshift it becomes

$$T_{\rm MM}^{-1/2} = T_0^{-1/2} + f_1 z + f_2(\log M_* - 10), \tag{5.6}$$

where T_0 is the merging time at $z = 0$, M_* is the stellar mass of the primary galaxy, while f_1 and f_2 are two coefficients of the parameterization. A mass dependence similar to the $z \leq 1$ Kitzbichler & White (2008) timescale ($T_{\rm MM} \sim M_*^{-0.2}$) was also found by Lotz et al. (2010) for mergers in the stellar mass range $9.7 < \log(M_*/M_\odot) < 10.7$ but in a smaller sample. More recent works explored a different redshift evolution of the merger timescale. For instance, Jiang et al. (2014) took into account the mass loss due to dynamical friction of the virial masses of galaxies experiencing a merger and, by using a high-resolution $N-$body cosmological simulation, found $T_{\rm MM} \propto H(z)^{-1/3}$, where $H(z)$ is the Hubble parameter at redshift z. An even stronger redshift evolution was found by Snyder et al. (2017) by comparing mass-selected close pairs with the intrinsic galaxy merger rate from the Illustris simulations (Genel et al. 2014; Vogelsberger et al. 2014). They found $T_{\rm MM} = 2.4\ (1+z)^{-2}$ but for primary galaxies with stellar mass range $10.5 < \log(M_*/M_\odot) \leq 11$. It is better to specify that, in the case of Snyder et al.

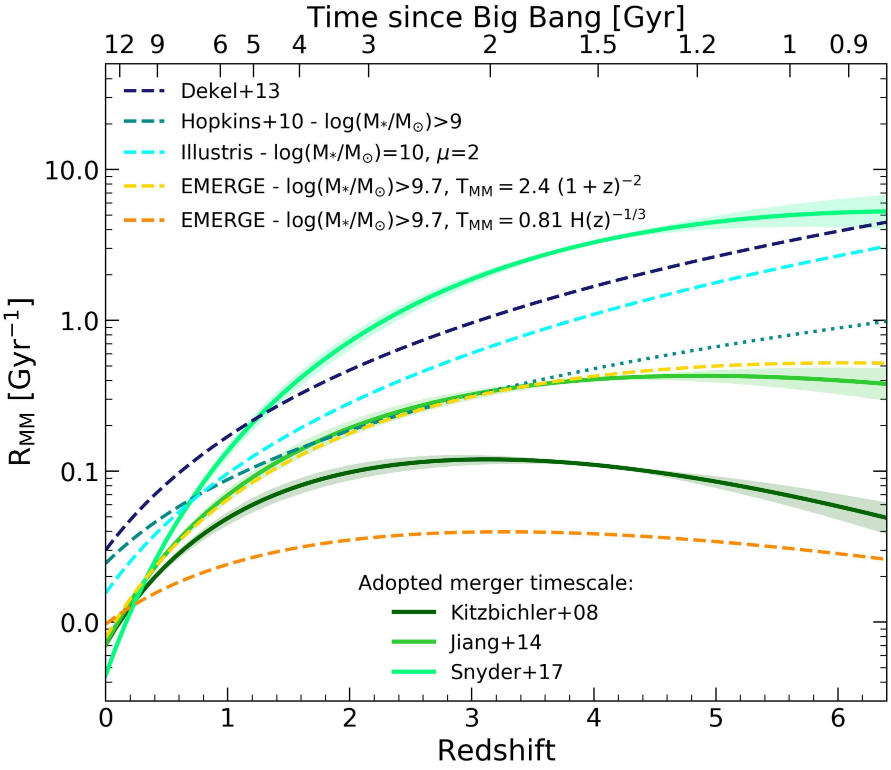


Figure 5.6: Redshift evolution of the major merger rate. The solid lines with the shaded regions represent the best-fitting functions to the data assuming three different merger timescales and their associated 1σ uncertainties, respectively. The cosmic evolution of $R_{\rm MM}$ computed from the halo-halo merger rate by Dekel et al. (2013), from the empirical model by Hopkins et al. (2010), and from the Illustris (Rodriguez-Gomez et al. 2015) and EMERGE (O'Leary et al. 2021) simulations are also shown with dashed dark blue, turquoise, cyan, yellow and orange lines. The dotted turquoise line shows the extrapolation of the Hopkins et al. (2010) merger rate to higher redshifts.

(2017), $T_{\rm MM}$ does not correspond to the merger timescale of Equation 5.5, rather it is the merger observability timescale which already includes the effects of the $C_{\rm merg}$ term. Nevertheless, we refer to that as $T_{\rm MM}$ for simplicity.

For the rest of this work, we compare the results obtained with the merger timescales provided by Kitzbichler & White (2008), Jiang et al. (2014) and Snyder et al. (2017), in order to highlight the large differences in the final outcome due to the choice of this parameter. About the probability for a pair to merge over a certain timescale (i.e., $C_{\rm merg}$), it is usually fixed to 0.6 (e.g., Lotz et al. 2011; Mantha et al. 2018). However, as in most cases this quantity is already included in the merger timescale prescriptions provided in different works (e.g. Duncan et al. 2019), we set $C_{\mathrm{merg}} = 1$ throughout.

With these assumptions, we present in Figure 5.6 the cosmic evolution of the merger rate, as obtained by combining our new $z > 4$ ALPINE data with those at lower redshifts from the literature. For each data point we compute R_{MM} from Equation 5.5, adopting the three different merger timescales defined above[11] and fit the resulting data with the same functional form used for the merger fraction. We show the best-fitting functions to these data (and their 1σ uncertainties) in Figure 5.6 and report the corresponding parameters in Table 5.4.

As evident, all the three functions show a steep increase of the merger rate from the local Universe to $z \sim 1$. However, at higher redshifts they lead to large differences about the incidence of these sources at early epochs. When considering the Kitzbichler & White (2008) timescale, we find a redshift evolution which is similar to what obtained in Section 5.3.2 for the merger fraction, with a peak at $z \sim 2.5$ and a quite fast decrease at higher redshifts. Both the merger rate evolutions obtained with $T_{\mathrm{MM}}(z)$ by Jiang et al. (2014) and Snyder et al. (2017) show instead a milder redshift attenuation at $z \gtrsim 4$. In particular, the increasingly smaller merger timescale found by Snyder et al. (2017) involves a very large merger rate in the early Universe compared to the other functions, possibly implying a significant role of major mergers in the galaxy mass-assembly at $z \sim 5$.

We then compare our results with the cosmic evolution of several expected merger rates from simulations. At $z < 3$, Hopkins et al. (2010) found that $R_{\mathrm{MM}} \propto (1+z)^{1.8}$ for galaxies with $\log(M_*/M_\odot) > 9$ and $\mu < 3$. This redshift evolution is comparable to our best-fitting function with $T_{\mathrm{MM}}(z) \propto H(z)^{-1/3}$ at $1 \lesssim z \lesssim 3$. However, if we extrapolate their results up to $z \sim 6$, we find that their predicted major merger rates could be larger (by ~ 0.5 dex) with respect to our observations. A stronger and increasing redshift dependence was found by Rodriguez-Gomez et al. (2015) for the merger rate from the Illustris simulations. By exploiting their parametric fitting formula with a stellar mass[12]

---

[11] When adopting the merger timescale by Jiang et al. (2014), we just consider the T_{MM} redshift evolution found in their work (i.e., $T_{\mathrm{MM}} \sim H(z)^{-1/3}$) and normalize it to $T_{\mathrm{MM}}(z = 0)$ from Kitzbichler & White (2008). This is also similar to setting a constant merger timescale of ~ 1 Gyr.

[12] It is worth noting that, as also stated by Snyder et al. (2017) and O'Leary et al. (2021), the mass ratio

of $\log(M_*/M_\odot) = 10$ and an average mass ratio $\mu = 2$ over cosmic time (see Section 5.3.5), the simulation tends to over-predict our observed merger rates at $z \sim 0$, while it underestimates them at intermediate redshifts. We also show the merger rate predictions found by O'Leary et al. (2021) using the results from the Empirical ModEl for the foRmation of GalaxiEs (EMERGE; Moster et al. 2018). In particular, they modeled the merger fraction evolution as in Equation 5.4 and then convolved it with a typical merger timescale to obtain $R_{\rm MM}$ as a function of redshift. We consider their best-fit parameters describing $f_{\rm MM}$ for galaxies with $\log(M_*/M_\odot) \geq 9.7$ and projected distance $r_p < 30$ kpc. When convolving this $f_{\rm MM}(z)$ with a merger timescale $T_{\rm MM} \propto H(z)^{-1/3}$ and with a power-law scaling as $T_{\rm MM} \propto (1+z)^{-2}$ we obtain a lower and milder $R_{\rm MM}$ cosmic evolution with respect to our major merger rate results obtained with a Kitzbichler & White (2008) timescale, and a comparable trend with the Jiang et al. (2014)-based evolution, respectively. Lastly, we show the predicted halo-halo merger rate by Dekel et al. (2013) which over-predicts the observed merger rates at low redshifts but is in good agreement with the Snyder et al. (2017)-based $R_{\rm MM}$ at $z \sim 6$.

We compute the average number of mergers a galaxy undergoes between $0 < z < 6$ by integrating the merger rate over time as

$$N_{\rm MM} = \int_{z_1}^{z_2} \frac{R_{\rm MM}(z)}{(1+z)H_0E(z)} dz, \tag{5.7}$$

where $E(z) = \sqrt{\Omega_{\rm m}(1+z)^3 + \Omega_\Lambda}$, assuming a flat Universe (i.e., $\Omega_{\rm k} = 0$). We show in Figure 5.6 the cumulative number of major mergers as a function of redshift adopting three different timescales, with the 1σ uncertainties derived from the errors on the major merger rate cosmic evolution. The average number of mergers computed with the Kitzbichler & White (2008) and Jiang et al. (2014) prescriptions is low, ranging between 0.5 and 1.4 down to $z = 0$. This is consistent, for instance, with the results by Mundy et al. (2017) who found that galaxies with $\log(M_*/M_\odot) > 9.5$ at $z \sim 3.5$ (selected at constant number densities, thus representing the progenitors of $\log(M_*/M_\odot) > 11$ galaxies at $z \sim 0$) undergo ~ 0.5 major mergers between $0 < z < 3.5$. However, if we assume an

---

defined by Rodriguez-Gomez et al. (2015) is evaluated at the time when the secondary galaxy has achieved its maximum stellar mass. This prevent us to make a direct comparison among the results of this and the other simulations.

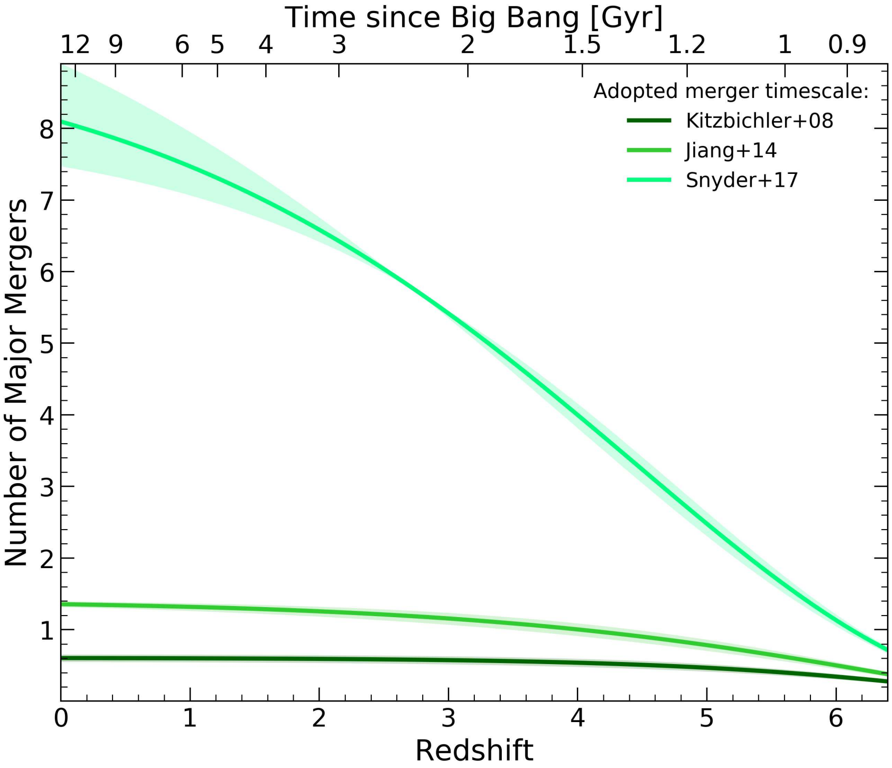


Figure 5.7: Cumulative number of major mergers per galaxy over cosmic time. Solid lines represent the cumulative distributions obtained by integrating Equation 5.7 adopting the merger timescale prescriptions of Kitzbichler & White (2008), Jiang et al. (2014) and Snyder et al. (2017). The shaded regions are the associated uncertainties computed from the best-fit errors on the merger rate cosmic evolution.

evolving merger timescale as $2.4\ (1+z)^{-2}$ (Snyder et al. 2017), the average number of mergers since $z = 6$ increases significantly, reaching $N_{\mathrm{MM}} \sim 8$ at the present day. Interestingly, this is similar to the result obtained by Conselice & Arnold (2009) that found $N_{\mathrm{MM}} \sim 7$ by integrating Equation 5.7 from $z = 6$ to 0, and assuming a constant merger timescale $T_{\mathrm{MM}} = 0.35$ (which is equal to the average timescale by Snyder et al. 2017 at $z < 6$; see their Figure 15). When examining our results at higher redshifts, we find that the number of mergers diminishes quickly. In particular, at $z > 4$, $N_{\mathrm{MM}} \lesssim 1$ if we consider the Kitzbichler & White (2008) and Jiang et al. (2014) merger timescale evolution, suggesting that, in these scenarios, not every galaxy will undergo a merger during this

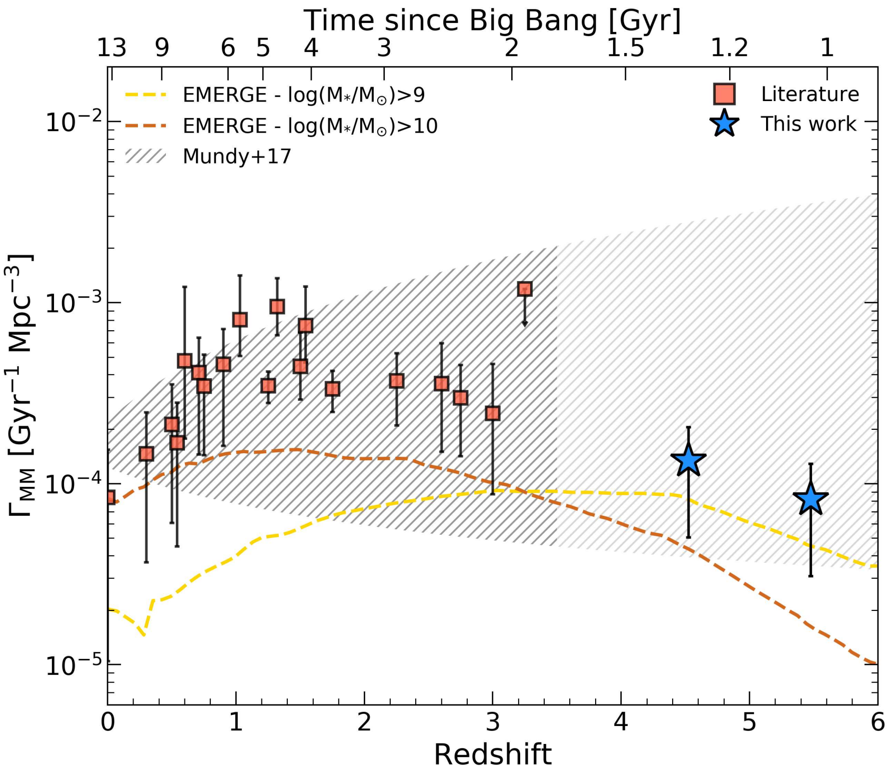


Figure 5.8: Merger rate density as a function of redshift. The squares represent the volume-averaged merger rates computed from the literature data at $z < 4$ adopting the Kitzbichler & White (2008) prescription for the merger timescale. The blue stars are the ALPINE data at $z \sim 5$ from this work. The error bars are obtained by combining in quadrature the uncertainties on the merger fractions and on the number densities. The darker hatched area reports the results including errors by Mundy et al. (2017) at $z < 3.5$. The lighter area only shows their extrapolation to higher redshifts. The dashed yellow and brown lines display the evolution of $\Gamma_{\rm MM}$ found with the EMERGE simulation (O'Leary et al. 2021) for galaxies with $\log(M_*/M_\odot) > 9$ and $\log(M_*/M_\odot) > 10$, respectively.

epoch.

5.3.4 Merger rate density

As discussed in Section 5.3.3, the merger rate traces the number of merger events per galaxy at a given mass and redshift. A more informative quantity is the volume-averaged merger rate, which is defined as

$$\Gamma_{\rm MM}(z) = \frac{f_{\rm MM}(z)\; n(z)}{T_{\rm MM}(z)}, \tag{5.8}$$

where $f_{\rm MM}$ and $T_{\rm MM}$ are the previously defined merger fraction and merger timescale, and $n(z)$ is the number density of galaxies in Mpc^{-3}. The latter is computed by integrating the galaxy stellar mass function (GSMF) in a certain redshift bin and between a minimum and maximum stellar mass limit in the range $M_*^{\rm min} < M_* < M_*^{\rm max}$. In particular, we exploit the GSMFs best-fit parameters from Mortlock et al. (2015), Santini et al. (2012) and Davidzon et al. (2017) at $z < 3$, $3 < z < 4$ and $z > 4$, respectively, after converting into a Chabrier (2003) IMF. However, it is worth noting that, as stated in Mundy et al. (2017), making use of different GSMF parameterizations does not significantly change the final results. We thus integrate these functions from $\log(M_*/M_\odot) = 9$ to $\log(M_*/M_\odot) = 11$, in order to include the full possible range of galaxy masses for the considered data at different redshifts. We estimate the errors on $n(z)$ by perturbing the corresponding GSMF with the associated uncertainties on its best-fit parameters at each redshift. We repeat this procedure 10^4 times, recomputing the integrated number densities at each step. From the perturbed distributions, we then obtain the 16th and 84th percentiles as the 1σ errors on $n(z)$.

Figure 5.8 shows the merger rate density computed with Equation 5.8 for each of the data introduced in Section 5.3.2 and for the two $z > 4$ ALPINE bins. For the sake of simplicity, in this case we report only the data points obtained by adopting the Kitzbichler & White (2008) prescription for the merger timescale. When using the $T_{\rm MM}(z)$ redshift evolution from Jiang et al. (2014) and Snyder et al. (2017), we obtain similar trends but shifted to higher merger rate densities. The error bar on each point is computed by propagating the merger fraction and number density uncertainties on Equation 5.8. At $1 \lesssim z \lesssim 5$, Duncan et al. (2019) found that the volume-averaged merger rate is quite constant. On the other hand, from our data we find a slight decrease both at $z < 1$ and at $z > 4$. At these early epochs, this difference could be caused by the poor constraints on the GSMFs adopted by Duncan et al. (2019) which result in large uncertainties on their data, making it impossible to draw robust conclusions at $z \sim 5$. Nevertheless, our derived merger rate densities are in agreement, within the uncertainties, with other results derived in the literature. For example, Mundy et al. (2017) studied the evolution of the merger rate density up to $z \sim 3.5$ for a large sample of $\log(M_*/M_\odot) > 10$ galaxies. They found that, for close pairs with projected separations $5 < r_p < 20$ kpc, the $\Gamma_{\rm MM}(z)$ evolution is better described by a power-law of the form $\Gamma_0(1+z)^\gamma$, with $\Gamma_0 = 1.64^{+0.58}_{-0.41} \times 10^{-4}$ and $\gamma = 0.48^{+1.00}_{-1.15}$. We report their results, along with the uncertainties, in Figure 5.8 and extrapolate them to the redshifts explored by ALPINE. As evident, our data points are comparable with the results by Mundy et al. (2017). If we also fit our merger rate densities (derived assuming the Kitzbichler & White (2008) merger timescale) with a power-law function we obtain $\Gamma_0 = (2.40 \pm 0.63) \times 10^{-4}$ and $\gamma = -0.16 \pm 0.23$, that are consistent with the above outcomes. The power-law fit on the data computed with the timescales by Jiang et al. (2014) is in agreement with the Mundy et al. (2017) findings, as well. If we instead consider the results obtained with the Snyder et al. (2017) timescale, we find a rapid increase of the merger rate density with redshift, which departs from the upper envelope of Mundy et al. (2017) already at $z \gtrsim 1$.

Finally, we also show the results from the EMERGE simulation (O'Leary et al. 2021) for galaxies in two different stellar mass bins. The two curves in the figure represent the intrinsic merger rate from the simulation for major mergers (i.e., $\mu < 4$) selected based on their main progenitor mass, thus describing a population of galaxies that will undergo a merger after a certain time. Galaxies with $\log(M_*/M_\odot) > 9$ are characterized by an increase of the merger rate density up to $z \sim 1$, an almost constant $\Gamma_{\rm MM}$ until $z \sim 4$ and a slow decrease to the highest redshifts. At larger stellar masses, a slighter increase of $\Gamma_{\rm MM}$ is present at low redshifts, while a steeper decrease in the early Universe is found. Both these trends are comparable to the one obtained in this work after exploiting data at different redshifts and including the new estimates at $z > 4$ from the ALPINE survey. The fact that they are systematically smaller than our results could be ascribed to several factors, such as the lower merger fraction evolution found in their simulation or the measure of the intrinsic merger rate which is computed when the merging process is already happened (thus possibly enhancing the merger timescales with respect to our assumptions).

5.3.5 The major merger specific mass accretion rate

By taking advantage of the merger rates and merger fraction deduced in the previous sections, we now provide an estimate of the average stellar mass gained by a galaxy during a merger event over cosmic time. To this aim, we compute the major merger specific mass accretion rate as $R_{\mathrm{MM}}\,\overline{\mu}^{-1}$, where $\overline{\mu}$ is the average mass ratio. We find that, for each data sample used to compute the merger rate at different redshifts, $\overline{\mu}^{-1} \sim 0.5$, including the average mass ratio from our sample of major mergers for which we know μ_K. Therefore, we just take the best-fitting functions to the cosmic evolution of the merger rate and multiply them by a factor ~ 0.5 to obtain the specific mass accretion rate in Gyr^{-1}. Our results are reported in Figure 5.9 along with the specific star formation rate (sSFR) evolution obtained by different authors (Dekel et al. 2009; Speagle et al. 2014; Tasca et al. 2015; Khusanova et al. 2021). As shown, star formation seems to be the dominant mode of mass growth at all epochs. However, if we assume a merger timescale $T_{\mathrm{MM}} \propto (1+z)^{-2}$ (Snyder et al. 2017), the specific mass accretion rate reaches the sSFR at $z > 3$, being comparable to it within the uncertainties. Therefore, we cannot exclude that major mergers may have significantly contributed to the galaxy mass-assembly in the early Universe. In particular, if the galaxy growth is dominated by cold gas accretion, the sSFR should evolve with redshift as $(1+z)^{2.25}$ (e.g., Dekel et al. 2009). Some authors find instead a flattening of the trend or even a possible decrease with redshift as from the ALPINE data (Tasca et al. 2015; Faisst et al. 2016; Khusanova et al. 2021), leaving space to other mechanisms that could regulate the assembly of galaxies at high redshifts, such as major mergers (e.g., Tasca et al. 2015; Faisst et al. 2016).

5.3.6 The major merger mass accretion rate density

Similarly to what did in Section 5.3.5 for the specific mass accretion rate, we exploit here the major merger rate density to obtain an estimate of the mass accreted through major mergers per unit time and volume for galaxies in a given stellar mass range, $\rho_{\mathrm{MM}}(z)$. This quantity can be compared to the other mechanisms of mass accretion, such as the mass gained through the process of star formation, and is then of fundamental importance to understand the role of mergers in the Universe.

Following Duncan et al. (2019), we assume that the increase in stellar mass for each

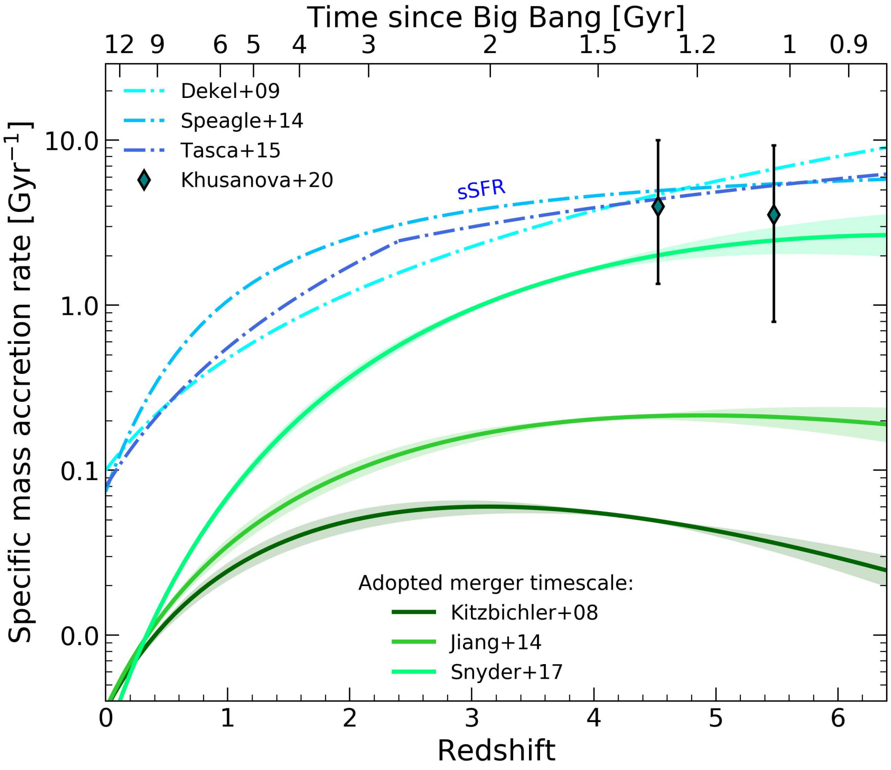


Figure 5.9: Comparison between the redshift evolution of the specific mass accretion rate (as derived from R_{MM}) and the specific star formation rate. The solid lines with the shaded regions are the best-fitting functions (and corresponding uncertainties) to the same R_{MM} data of the left panel of Figure 5.6 at different merger timescales divided by the average mass ratio $\overline{\mu}$ through cosmic time. The dot-dashed lines are several specific star formation rates from the literature (Dekel et al. 2009; Speagle et al. 2014; Tasca et al. 2015). The sSFR by Dekel et al. (2009) is normalized to sSFR(z = 0) = 0.1, as in Tasca et al. (2015). The dark cyan diamonds are the values by Khusanova et al. (2021) from the ALPINE survey.

merger and at each redshift is given by $\overline{M}_* \, \overline{\mu}^{-1}$, where $\overline{\mu}$ is the average mass ratio defined in Section 5.3.5 and $\overline{M}_*$ is the average stellar mass calculated through the GSMF. The latter is computed as

$$\overline{M}_*(z) = \frac{\int_{M_*^{\mathrm{min}}}^{M_*^{\mathrm{max}}} \Phi(z, M_*) \; M_* \; dM_*}{\int_{M_*^{\mathrm{min}}}^{M_*^{\mathrm{max}}} \Phi(z, M_*) \; dM_*}, \tag{5.9}$$

where $\Phi(z, M_*)$ represents the shape of the GSMF in a certain redshift bin. We estimate

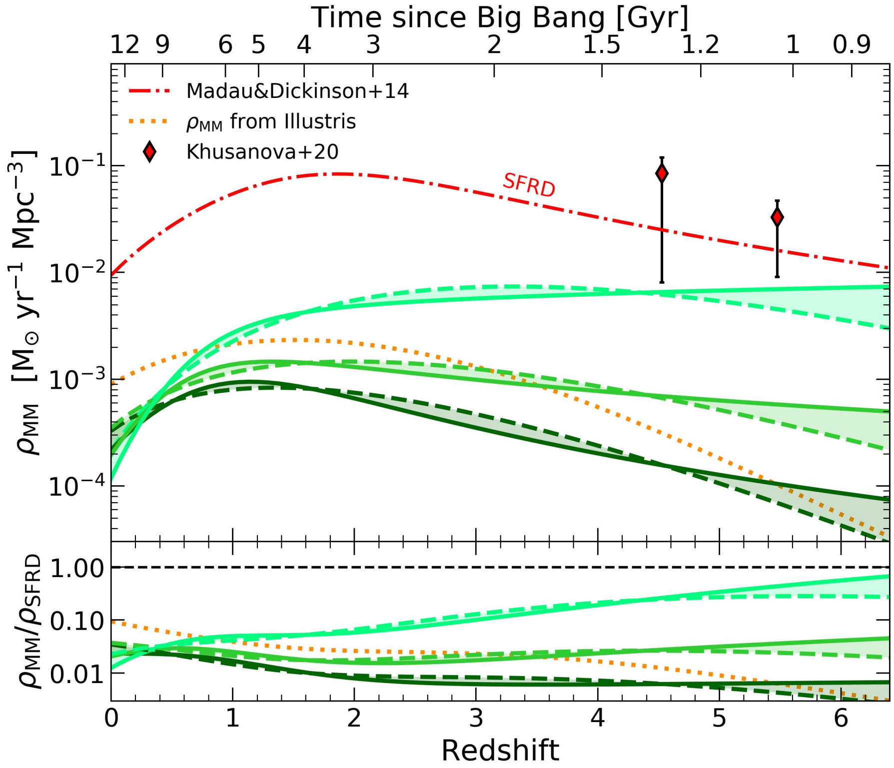


Figure 5.10: Stellar mass accretion rate density ($\rho_{\rm MM}$) as a function of redshift (top panel). Solid and dashed lines represent the best-fits to the data assuming double power-law and power-law/exponential functions, respectively. The shaded regions highlight the uncertainties resulting from the choice of the fitting form. The colors correspond to alternative merger timescales and are the same as in the left panel. The dotted curve reports the mass accretion rate density, $\rho_{\rm ill}$, obtained from the Illustris simulation, as described in the text. The dot-dashed red line shows the SFRD ($\rho_{\rm SFRD}$) by Madau & Dickinson (2014). The red diamonds are the total SFRD values obtained from the ALPINE survey (Khusanova et al. 2021). The bottom panel displays the ratio between $\rho_{\rm MM}$ and $\rho_{\rm SFRD}$ as a function of cosmic time. The dashed horizontal line marks a ratio equal to 1.

the uncertainties on $\overline{M}_*(z)$ as previously done for the number density, i.e., by recomputing it N times after perturbing the corresponding GSMF with its associated errors. At each redshift, we obtain a typical uncertainty of ~ 0.5 dex. In this way, the stellar mass accretion rate density can be estimated as

$$\rho_{\rm MM}(z) = \overline{\mu}^{-1}\, \overline{M}_*\, \Gamma_{\rm MM}(z). \tag{5.10}$$

We report the cosmic evolution of ρ_{MM} in Figure 5.10 (top panel). The dashed lines show the best-fitting curves to the data assuming a power-law/exponential function, as done for the merger fraction and merger rate, for different merger timescales. As shown, at $z > 1$ the curves start to decrease quite rapidly to lower mass accretion rate densities.

Following Mundy et al. (2017), we then provide the major merger accretion rate density computed with the results from the Illustris simulation, $\rho_{\mathrm{ill}}(z)$. We convolve the specific mass accretion rate by Rodriguez-Gomez et al. (2016) with the number density of galaxies by Torrey et al. (2015) as

$$\rho_{\mathrm{ill}}(z) = \int_{M_*^{\mathrm{min}}}^{M_*^{\mathrm{max}}} \int_{\mu_{\mathrm{min}}}^{\mu_{\mathrm{max}}} \Phi_{\mathrm{ill}}(z, M_*)\ \dot{m}_{\mathrm{accr}}(z, M_*, \mu)\ d\mu\ dM_*, \tag{5.11}$$

where Φ_{ill} is the shape of the GSMF at redshift z and stellar mass M_*, and $\dot{m}_{\mathrm{accr}}$ represents the average amount of mass accreted by a single galaxy when the merger occurs. As done for ρ_{MM}, we integrate the GSMF in the range $9 < \log(M_*/M_\odot) < 11$ and the mass accretion rate for mass ratios between $1 < \mu < 4$, accounting for major mergers. The result of the integration is reported in Figure 5.10 as a function of redshift. As can be seen, the computed ρ_{ill} is comparable with our findings, lying between the best-fit curves obtained by adopting the Kitzbichler & White (2008) and Snyder et al. (2017) prescriptions for the merger timescale, respectively. Moreover, the shape of the curve is quite similar to those deduced by our analysis, with a fast decrease at high redshifts.

To compare the relative contribution of mergers and star formation to the mass-assembly of galaxies through the cosmic time, we also show the SFRD from Madau & Dickinson (2014), after converting to a Chabrier IMF[13]. Moreover, we find that our data are also well fit by a double power-law of the form assumed by Madau & Dickinson (2014) for the SFRD, although we find no significant differences between the two functional forms when comparing them with a BIC statistics. The results of these fits are reported in the figure (solid lines), showing a flatter trend to the early Universe with respect to the power-law/exponential curves. In the bottom region of the figure we show the ratio between the major merger and star formation contributions to the stellar mass accretion in galaxies at different epochs. Looking at both curves, if we assume the merger

---

[13] To convert SFRs from Salpeter (1955) to Chabrier IMF we multiply by 0.63 (e.g., Madau & Dickinson 2014).

timescales by Kitzbichler & White (2008) and Jiang et al. (2014), major mergers seem to contribute less than 10% to the galaxy mass-assembly compared to the star formation mechanism at all redshifts. For comparison we also report, in the top panel, the total (UV+IR) SFRD obtained with the ALPINE data (Khusanova et al. 2021). These measurements indicate a possible $z > 4$ evolution of the SFRD that is shallower than previously thought, further decreasing the importance of major mergers to the mass-assembly of galaxies at these epochs. On the other hand, adopting the Snyder et al. (2017) timescale prescription, a larger contribution of mergers in the early Universe is in place, which becomes comparable to that provided by the star formation at $z \gtrsim 6$.

5.4 The importance of the merger timescale

As evidenced by the results presented in Section 5.3, one of the major sources of uncertainty in the investigation of the contribution of major mergers to the mass-assembly of galaxies through cosmic time is the typical time needed for a pair of objects to coalesce with each other. Indeed, this parameter is critical for converting the pair fraction into a merger rate, affecting all the derived quantities as well. Because of this, it could represent the main reason of disagreement between models and observations noticed in many works.

Previous studies often assumed $T_{\rm MM}$ as a constant over time. However, the emerging picture of a possibly decreasing (as found in this work) or nearly flat merger fraction through high redshifts (Conselice & Arnold 2009; Ventou et al. 2017; Mantha et al. 2018), compared to the increase of the merger rate found in most simulations over the same epoch (e.g., Hopkins et al. 2010; Rodriguez-Gomez et al. 2015), suggests that a redshift-dependent merger timescale is more suitable to reconcile models and observations. Here we show the results derived from three different merger timescales (Kitzbichler & White 2008; Jiang et al. 2014; Snyder et al. 2017), as introduced in Section 5.3.3, and compare them with state-of-the-art simulations. Kitzbichler & White (2008) found a merger timescale which slightly depends on the primary galaxy mass and that increases with redshift, leading to lower values of the merger rate when moving to early epochs. On the other hand, Jiang et al. (2014) and Snyder et al. (2017) found a slow and fast decrease of $T_{\rm MM}$ over time, respectively, implying a larger contribution from mergers to the galaxy mass-assembly at $z > 4$.

Looking at Figure 5.6, the merger rates obtained from the Kitzbichler & White (2008) timescale prescription are not easily reproducible by models. In fact, although we find values of $R_{\rm MM}$ comparable to those of the EMERGE simulation (O'Leary et al. 2021) at $z \gtrsim 5$, the latter are computed with a Jiang et al. (2014)-like evolving timescale which can provide $T_{\rm MM}$ values up to 5 times lower than the Kitzbichler & White (2008) ones at these epochs. A good agreement with the $T_{\rm MM} \propto H(z)^{-1/3}$ merger rates is found by convolving the EMERGE pair fraction evolution with the decreasing Snyder et al. (2017) timescale (again using smaller merging times in the simulation with respect to the observations) and with the outputs of the Hopkins et al. (2010) simulations, at least at $z < 3$ (indeed, they predicted larger merger rates with respect to our findings at earlier epochs). However, in this case, we note that the merger rates from Hopkins et al. (2010) are properly computed up to $z \sim 3$ and that the high-redshift extrapolation could be different from the real outputs of their simulation. Moreover, the observed shift between our findings and the EMERGE results due to the different choice of the merger timescale could be ascribed to the merger fraction evolution which, from their simulation, is smaller than ours. Finally, the fast-decreasing merger timescale by Snyder et al. (2017) lead to merger rates which are more than an order of magnitude larger than those from the Kitzbichler & White (2008) and Jiang et al. (2014) prescriptions. Adopting this formalism, we find that our merger rates are larger than those predicted by the Illustris simulations (Rodriguez-Gomez et al. 2015) (with an average mass ratio ~ 2) and comparable with the merger rates of $M_{\rm halo} \sim 10^{12}\ M_\odot$ by Neistein & Dekel (2008); Dekel et al. (2013) at $z \sim 6$.

Duncan et al. (2019) found a good agreement between their observations and simulated merger rates from Rodriguez-Gomez et al. (2015) for galaxies with $\log(M_*/M_\odot) > 10.3$, while they also obtained higher values than those predicted for lower mass galaxies. However, the results of that simulation depend significantly on the adopted mass ratio. Indeed, assuming $\mu > 2$, the Illustris simulation is able to reproduce our observations, at least at $z \gtrsim 4$. Moreover, Mantha et al. (2018) also found merger rates comparable to those predicted by Rodriguez-Gomez et al. (2015) up to $z \sim 3$ by dividing their observed pair fraction by the Snyder et al. (2017) merger timescale.

These results suggest that, to reconcile observations and simulations, a redshift-evolving timescale should be employed when converting the pair fraction into a merger rate, with a $T_{\rm MM}$ decreasing toward early epochs likely being the most suitable choice. However, many other factors can affect the $R_{\rm MM}$ estimates, like the probability of merging $C_{\rm merg}$ or the dependency of the merger timescale on several properties like the stellar mass of the galaxies, the on-going phase of the merger or the selection criteria of the close pairs. For these reasons, further investigation is needed in order to obtain more robust conclusions on the incidence of merger through cosmic time.

5.5 The contribution of major mergers to the galaxy mass-assembly

The relative contribution of the different mechanisms driving the galaxy mass-assembly through cosmic time has yet to be ascertained. The most favored scenario predicts cold accretion and *in-situ* star formation as the principal sources of stellar mass increase in galaxies, with major mergers only playing a minor role in the galaxy build-up (e.g. Dekel et al. 2009; Conselice et al. 2013; Kaviraj et al. 2013; Sánchez Almeida et al. 2014). However, during the last years several works started to suggest a prominent contribution of mergers within this context, opening again the debate on the importance of these events in the framework of galaxy evolution (e.g., Tasca et al. 2014; Mantha et al. 2018; Duncan et al. 2019).

Depending on the merger timescale adopted, with our new ALPINE data, we can find large merger rates up to high redshifts which could imply a substantial contribution of these events to the galaxy mass accretion at early epochs. We find that the stellar mass accretion rate density due to major mergers has a similar redshift evolution as the cosmic SFRD (Madau & Dickinson 2014) and that, assuming the Snyder et al. (2017) merger timescale, they become comparable with each other at redshifts approaching the Reionization epoch. However, the lack of further constraints on the most suitable choice for the merger timescale prevent us from providing firm conclusions. Indeed, looking at Figure 5.10, the importance of major mergers in the assembly of galaxies could be very different, especially at high redshifts, depending on all the above-mentioned caveats and on the adopted fitting function. It is worth noting that the comparison between the

SFRD and the mass accretion rate density in Figure 5.10 is not intended to provide the contribution of major mergers to the global star formation through cosmic time, rather it represents the relative importance of the two processes to the galaxy mass-assembly at a given epoch.

Despite this, we can certainly affirm that major mergers are frequent processes in the early Universe and must be accounted for when considering the mechanisms of galaxy assembly across cosmic time.

5.6 SUMMARY

In this chapter, we investigated the role of major mergers in the stellar mass-assembly of galaxies at $z \sim 5$. To this aim, we took advantage of the recent [CII] observations of a significant statistical sample of normal SFGs in the early Universe, as carried out by the ALMA large program ALPINE. The large amount of data (both spectroscopic and photometric) available for these sources combined with their new morphological and kinematic characterization made possible by the 3D [CII] information, allowed us to put one of the first constraints on the fraction and rate of mergers shortly after the end of the epoch of Reionization. Indeed, only a handful of works have been undertaken so far about the study of the merger contribution to the build-up of galaxies at $z > 4$, among which some are affected by large uncertainties due, for instance, to the lack of spectroscopic redshifts which properly identify close pairs. In the following, we summarize our major results:

- We identify 23 mergers, corresponding to $\sim 31\%$ of the 75 ALPINE [CII]-detected galaxies, with an average stellar mass of the most massive galaxies of $\log(M_*/M_\odot) \sim 10$. We find major merger fractions $f_{\rm MM} = 0.44^{+0.11}_{-0.16}$ and $f_{\rm MM} = 0.34^{+0.10}_{-0.13}$ at $z \sim 4.5$ and $z \sim 5.5$, respectively, after taking into account the presence of minor mergers and correcting for completeness. These results are in good agreement with morphological studies by Conselice & Arnold (2009) at the same redshifts and, when combined with previous works down to the local Universe, suggest a cosmic merger fraction evolution with a rapid increase from $z = 0$ to $z \sim 2$, a peak at $z \sim 2-3$, and a possible slow decline for $z > 3$. This trend is well described by a power-law/exponential function and is quite comparable with the outputs of the EAGLE hydrodynamical simulations from which Qu et al. (2017) claim that major mergers contribute less than 10% to the stellar mass gain in galaxies less massive than $\log(M_*/M_\odot) = 10.5$, attributing a leading role to the *in-situ* star formation.

- We convert the merger fraction into the merger rate per galaxy by adopting different redshift scaling for the typical time of merging ($T_{\rm MM}$). The results we obtain are strongly dependent on the choice of $T_{\rm MM}$. For instance, Kitzbichler & White (2008) found a merger timescale evolution which diverges more and more from that of Snyder et al. (2017) moving to high redshift. When using the first timescale prescription, we obtain $R_{\rm MM} \sim 0.09$ and $R_{\rm MM} \sim 0.07$ Gyr^{-1} at $z \sim 4.5$ and $z \sim 5.5$, respectively. However, these numbers are more than an order of magnitude larger if $T_{\rm MM} \propto (1+z)^{-2}$ is used. Accordingly, this uncertainty propagates to all our subsequent outcomes that involve the choice of a merger timescale.

- By integrating the merger rate over time, we obtain the average number of mergers a galaxy undertakes during its cosmic history from the early to the local Universe, which goes from less than 1 to ~ 8 mergers at $z = 0$, depending on the adopted merger timescale prescription. Our results are quite in agreement with the outputs of several cosmological simulations, especially when adopting the redshift-dependent Snyder et al. (2017) merger timescale with which we can match the large merger rates predicted at high redshifts by the Illustris simulations (Rodriguez-Gomez et al. 2015) and the halo-halo merger rate by Dekel et al. (2013).

- Another important quantity we estimate is the volume-averaged merger rate $\Gamma_{\rm MM}$, that is the merger rate $R_{\rm MM}$ multiplied by the number density of galaxies at a given redshift and for a specific stellar mass range. We find a nearly constant $\Gamma_{\rm MM}$ at intermediate redshifts in agreement with previous works in the literature (Mundy et al. 2017; Duncan et al. 2019), and a possible decrease at both $z < 1$ and $z > 4$.

- We take advantage of the two observationally-estimated merger rates to provide a constraint on the contribution of major mergers to the galaxy mass-assembly through the cosmic epochs. From the merger rate per galaxy, we obtain the average stellar mass accreted per major merger per unit of mass and compare it with

> the sSFR computed by different authors. The latter seems to dominate over the specific mass accretion rate at all times. However, when considering a rapidly evolving merger timescale like $T_{\rm MM} \propto (1+z)^{-2}$, the two quantities become comparable within the uncertainties at $z > 4$. We then estimate the mass accretion rate density from $\Gamma_{\rm MM}$ and compare this quantity to the well-known SFRD cosmic evolution (Madau & Dickinson 2014). We note that the contribution of major mergers to the global star formation rate ranges between being approximately equal to the SFRD to less than 1% of it, depending on the exact choice of the merging timescale, the parametric form used to fit the data, and the redshift.

To conclude, major mergers could have played a significant role in the galaxy mass-assembly through cosmic time, especially when merger timescales from recent literature, in which $T_{\rm MM}$ decreases rapidly with increasing redshift, are taken into account. However, future investigation is needed to finally establish the importance of such events in the complex picture of galaxy evolution. In particular, a larger statistical sample of spectroscopically confirmed galaxies, observed with deeper resolution, will allow us to confirm the large fraction of mergers at early times and their incidence at different redshifts, unveiling the relative contribution of each process to the build-up of galaxies through the Universe.

6 | THE POPULATION OF [CII]-UNDETECTED GALAXIES

Based on:
Romano M., Morselli L., Cassata P., et al., submitt.

The [CII] 158 μm emission line represents so far one of the most profitable tools for the investigation of high-redshift galaxies in the early Universe (see Section 1.1.1). Being one of the brightest cooling lines in the rest-frame FIR regime of SFGs, it has been successfully exploited as a tracer of SFR in local sources. The picture is more complex at higher redshifts, where its usability in this context is still under investigation. Recent results from the ALPINE survey suggest that there is no (or weak) evolution of the $L_{[CII]}$-SFR relation up to $z \sim 6$ but their reliability is hampered by the presence of a large population of [CII] non-detected galaxies.

In this chapter, we characterize the population of [CII] non-detections in ALPINE. By stacking their ALMA spectra, we obtain a signal detected at $\sim 5.8\sigma$, resulting in a [CII] luminosity of $\log(L_{[CII]}/L_{\odot}) \sim 7.7$. When combining this value with those from the [CII] detections, we find a $L_{[CII]}$-SFR relation with a slope of ~ 1.3, in agreement within the uncertainties both with the linear relation found in the local Universe, and with the previous findings from ALPINE at $z \sim 5$. This suggests that the [CII] line can be considered a good tracer of star formation up to the distant Universe. Finally, we show that the galaxies of our sample that most deviate from the observed $L_{[CII]}$-SFR relation could suffer from a poor redshift estimation, perhaps artificially reducing their

[CII] luminosity. In this respect, we claim that there is no evidence in favor of a deficit of [CII] content in high-z galaxies, in contrast with earlier studies.

6.1 General context

Over the last years, observations of the [CII] line emission at 158 μm rest-frame in galaxies have been progressively improved to the point of being able to characterize the first sources of light during or near the epoch of cosmic Reionization (e.g., Wagg et al. 2012; Carilli & Walter 2013; Wang et al. 2013; Capak et al. 2015; Smit et al. 2018; Hashimoto et al. 2019; Bakx et al. 2020; Le Fèvre et al. 2020).

In this context, the ALPINE survey (Béthermin et al. 2020; Faisst et al. 2020b; Le Fèvre et al. 2020) has recently provided the first statistically significant sample of high-redshift main-sequence galaxies detected in [CII] at the end of the reionization epoch ($4.4 < z < 5.9$). Considering as detections those galaxies with a [CII] emission $\geq 3.5\sigma$ (corresponding to a 95% sample purity), ALPINE reached a successful rate of 64%, resulting in 75 detections and 43 non-detections (Béthermin et al. 2020; Le Fèvre et al. 2020).

Béthermin et al. (2020) first investigated the $L_{[CII]}$-SFR relation by taking advantage of the ALPINE continuum-detected galaxies. They computed the average SFRs in different [CII] luminosity bins as the sum of the UV rest-frame data (Faisst et al. 2020b) and of the mean obscured SFRs derived through the stacking of the continuum data. Their results are in good agreement with the local and predicted relations. Then, Schaerer et al. (2020) took advantage of the full ALPINE sample (including both [CII]-detected galaxies and upper limits on non-detections; see Section 5.3) to study the evolution of the $L_{[CII]}$-SFR relation over cosmic time, and to understand if the [CII] line is a good tracer of SFR at high redshift as it is in the local Universe (e.g., De Looze et al. 2014; Pineda et al. 2014). They found that the [CII] luminosity of the ALPINE galaxies scales linearly with their total SFRs (as traced by the sum of UV and infrared contributions; see Schaerer et al. 2020), with a slight steepening of the slope depending on the [CII] non-detections upper limits used (Béthermin et al. 2020). However, to fully establish the connection between $L_{[CII]}$ and SFR in distant galaxies, a more in-depth investigation of the ALPINE non-detections is needed.

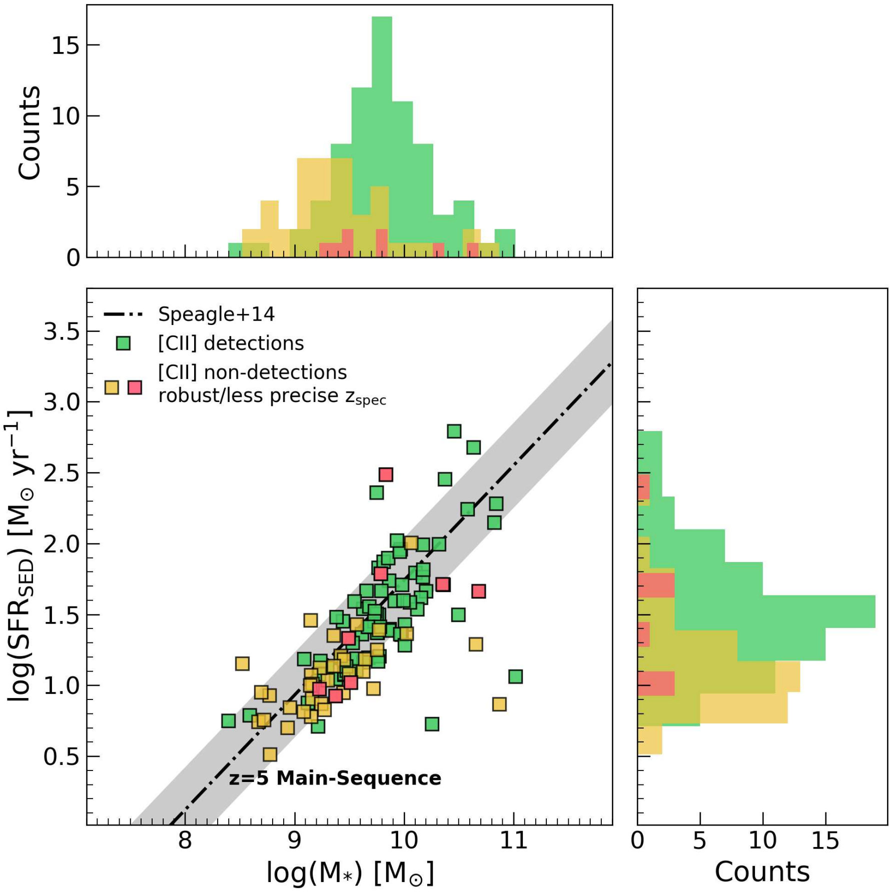


Figure 6.1: SFRs vs stellar masses of the ALPINE [CII]-detected (green squares) and undetected (yellow squares) galaxies. Red symbols represent [CII] non-detections with less precise spectroscopic redshift (see Section 6.2.2). The main-sequence of SFGs by Speagle et al. (2014) is shown as the black dot-dashed line, with its ±0.3 dex width represented by the shaded gray region. The top and right panels report the distributions in stellar mass and SFR, respectively, for both the detections and non-detections.

In this chapter, we derive the average properties of the population of [CII] undetected galaxies in ALPINE through line stacking and use them to investigate the [CII] as a tracer of SFR at high redshift, and to put constraints on the already thoroughly studied $L_{[CII]}$-SFR relation (Schaerer et al. 2020).

6.2 DATA AND OBSERVATIONS

6.2.1 MULTI-WAVELENGTH AND ALMA DATA

The ALPINE survey was designed to observe the [CII] line at 158 μm rest-frame and the surrounding FIR continuum emission from a sample of 118 SFGs at $4.4 < z < 5.9$, avoiding the redshift range $4.6 < z < 5.1$ due to a low-transmission atmospheric window. The targets are drawn from the well-studied COSMOS (Scoville et al. 2007a,c) and E-CDFS (Giavalisco et al. 2004; Cardamone et al. 2010) fields and have been observed in large optical/NIR spectroscopic campaigns such as VUDS (Le Fèvre et al. 2015; Tasca et al. 2017) and DEIMOS 10K Spectroscopic Survey (Hasinger et al. 2018). They are selected in the rest-frame UV ($L_{UV} > 0.6\ L^*$) and lie on the so-called main-sequence of SFGs (e.g., Noeske et al. 2007; Rodighiero et al. 2011; Tasca et al. 2015), thus being representative of the average population of SFGs at $z \sim 5$ (e.g., Speagle et al. 2014). A wealth of multi-wavelength data is also available for these sources (see Section 2.1.2).

In Figure 6.1 we show the distributions of [CII]-detected and undetected galaxies along the $z \sim 5$ main-sequence of SFGs. The stellar masses and SFRs are those from the SED-fitting (Faisst et al. 2020b). As evident, the ALPINE non-detections lie on the bottom-left side of the main-sequence, at lower stellar masses and SFRs with respect to those detected in [CII] (except for a few massive sources below the sequence).

6.2.2 REST-FRAME UV SPECTROSCOPIC DATA

The 118 ALPINE galaxies have confirmed rest-frame UV spectroscopic redshifts from the VUDS (Le Fèvre et al. 2015; Tasca et al. 2015) and DEIMOS 10K (Hasinger et al. 2018) surveys. These are obtained both from the Lyα line and from UV rest-frame ISM absorption lines. However, these features are not always the best tracers of the systemic

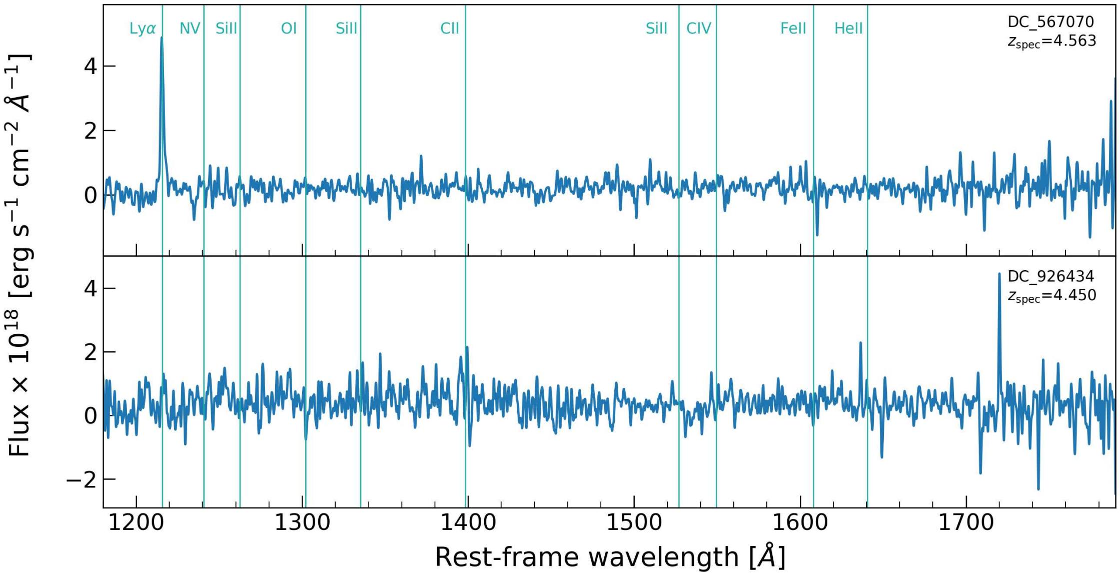


Figure 6.2: Examples of optical spectra of two [CII] non-detections with robust (top) and less precise (bottom) spectroscopic redshift. Both panels report UV emission and absorption features, such as the Lyα line or the ISM Si II, C IV and He II absorption lines. The spectra are smoothed with a Gaussian filter with size of 2 for a better visualization of the emission/absorption features.

redshift of a source. The Lyα emission line is typically redshifted (with respect to the systemic velocity) because of the resonant scattering of the Lyα photons. On the opposite, ISM lines are usually blueshifted, suggesting the presence of outflowing gas. The [CII] line is not affected by this kind of issues and, in principle, it can be used to stack together the ALMA spectra of the ALPINE non-detections to search for a significant signal. Moreover, it is not absorbed by dust and can be observed across the entire galaxy, resulting to be a better tracer of the systemic redshift than optical nebular lines, as well (e.g., Cassata et al. 2020; Faisst et al. 2020b).

We do not have the systemic redshift information for the 43 ALPINE non-detections, therefore we can just rely on their UV redshifts. In particular, for our analysis, we start with the redshift obtained from the peak of the Lyα line, and then correct it for the observed velocity offsets between Lyα and the systemic velocity traced by [CII] ($0 < \Delta v_{\rm Ly\alpha} < 400$ km s^{-1}) for the ALPINE detections (Cassata et al. 2020). To account for this, we need to know the UV spectroscopic redshifts of our sources with good accuracy, in order to exclude objects with less precise redshift estimates for which the [CII] line could lie outside of the ALMA observational window or that could alter the $\Delta v_{\mathrm{Ly}\alpha}$ statistics in our stacking analysis (see Section 6.3). For this reason, we visually inspected the optical spectra of the non-detections. We found that 35 out of 43 sources present multiple high SNR spectral features, allowing for a precise and accurate estimate of their spectroscopic redshifts[1]. The remaining 8 galaxies have very weak or no Lyα in emission and less prominent and sharp ISM absorption lines: although it is likely that the redshift is generally accurate for these sources as well, the low SNR of the spectral features does not allow us to obtain a redshift as precise as for the other 35 galaxies. Figure 6.2 reports an example of UV/optical spectra of two [CII] non-detections with a strong and absent Lyα line, respectively. In the first case, the galaxy shows a clear Lyα line in emission and some other possible absorption features at longer wavelengths that provide a precise estimate of the spectroscopic redshift. On the other hand, the spectrum of the second source is quite noisy, with a few recognizable spectral features. We decide to exclude this kind of sources for the rest of the work, in order not to include additional uncertainties (due to possible strong [CII] offsets with respect to their expected positions) to our analysis. We thus obtain a final sample of 35 non-detections. The 8 galaxies with less precise z_{spec} are also reported in Figure 6.1 for completeness as red squares.

6.3 STACKING OF NON-DETECTIONS

We proceed with a mean stacking of the ALMA spectra of the 35 ALPINE non-detections to search for a signal emerging from the noise of individual galaxies. In particular, we extract each spectrum from the original data cubes of the ALPINE data release 1 (DR1; Béthermin et al. 2020), within a fixed aperture of 1" of radius (defining the central regions of the ALPINE targets; see Béthermin et al. 2020) centered at the rest-frame UV position as traced by HST photometry.

At first, we use the Lyα-based spectroscopic redshifts of the non-detections to align them to the same reference frame, and then we stack them together. Because the Lyα line is typically shifted to the red relative to the systemic redshift as defined by the [CII] line (Cassata et al. 2020), it is likely that our stacked emission is offset from the systemic

---

[1] It is worth noting that all of these spectra show a prominent Lyα line in emission.

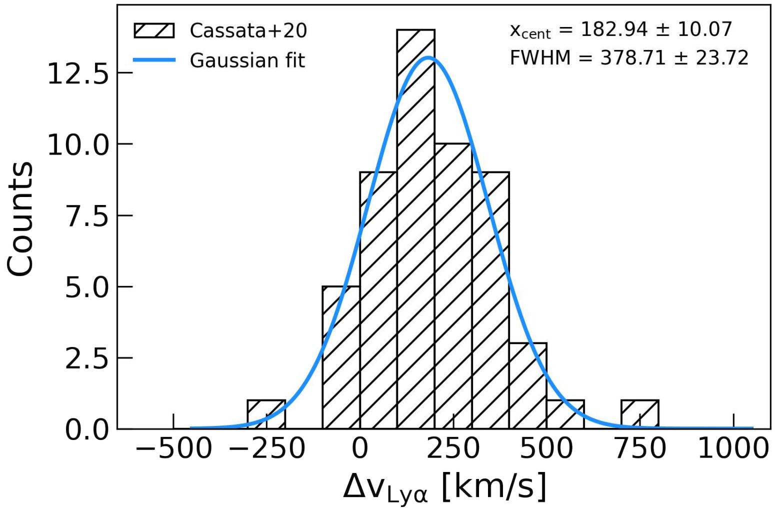


Figure 6.3: Offset velocities between Lyα and [CII] for a sub-sample of the ALPINE galaxies as computed by Cassata et al. (2020). The blue line shows a Gaussian fit on the distribution which peaks at $x_{cent} \sim 183$ km s^{-1}.

by some amount. Further, the velocity offset between Lyα and systemic is not constant, but is rather a complicated function of various physical conditions within a given galaxy (e.g., Erb et al. 2004; Pentericci et al. 2016; Marchi et al. 2019), meaning that the signal recovered in a stack where these offsets were not accounted for would be broadened and dampened. For these reasons, we apply to each source a spectral offset randomly extracted from the $\Delta v_{Ly\alpha}$ distribution obtained by Cassata et al. (2020) from a sub-sample of the ALPINE detections, that is shown in Figure 6.3. Then we compute the mean stack on the shifted spectra and repeat the whole process 1000 times. At each realization, we compute a Gaussian fit on the stacked line, estimating its full width at half maximum as FWHM = 2.355σ, where σ is the standard deviation of the Gaussian. For each stacked spectrum, we require that FWHM $\leq$ 400 km s^{-1}, defining the 84th percentile of the observed FWHM distribution of the ALPINE targets (see Béthermin et al. 2020). This check allows us to exclude those realizations for which the difference between the rest-frame UV spectroscopic redshift (as traced by the Lyα line) and the systemic one (as traced by the [CII] line) is too large, resulting in an artificially broader stacked line.

We show the result of this procedure in Figure 6.4, where the gray histograms rep-

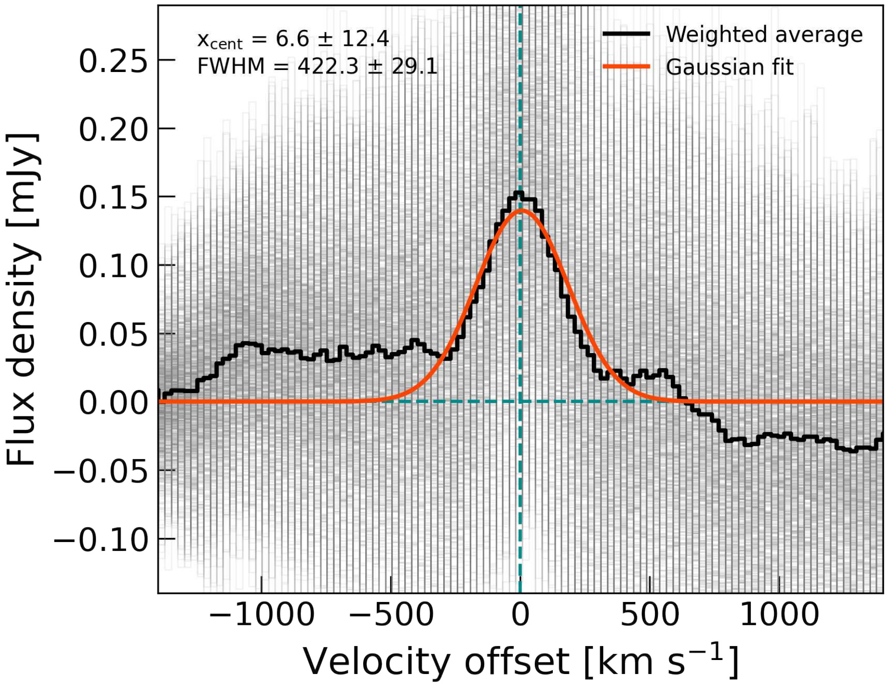


Figure 6.4: Average spectrum of the ALPINE [CII] non-detections weighted for the SNR of the individual stacked spectra with FWHM $\leq$ 400 km s^{-1} (solid black line). The thin lines represent the individual realizations after taking into account the observed shift between the [CII] and Lyα lines (Cassata et al. 2020). The solid red line represents the Gaussian fit on the average line profile. The dashed dark cyan lines mark the zero flux and velocity offset levels. The centroid and FWHM computed from the Gaussian fit on the average line profile are shown on the top left corner.

resent the spectra that satisfy the above requirements on the FWHM. To provide an average [CII] luminosity and SNR of the stacked line, we attribute different weights to each realization. We produce the averaged line profile assuming $w = 1$ in case of no weighting, $w = 1/\sigma_{rms}^2$ for an inverse-variance weighting, and w = SNR to weight each stacked spectrum by its SNR. In the latter case, we estimate the SNR as the ratio between the peak flux of the stacked line and the standard deviation of the spectral channels at velocities greater and smaller than ±600 km s^{-1} from the peak, in order to avoid contamination from the stacked emission line. The line profiles obtained by applying these different weightings are similar to each other, although resulting in a slightly higher

SNR for the case with w = SNR. For this reason, we decide to show in Figure 6.4 only the average stacked spectrum obtained by weighting each realization by its SNR. By fitting the stacked line profile with a Gaussian function, we find the signal peaking at $x_{\rm cent} = 6.6 \pm 12.4$ km s^{-1}, consistent with the systemic velocity traced by the [CII] line, and having FWHM = 422.3 ± 29.1 km s^{-1}. The computed SNR is $\sim$ 5.8, revealing the presence of an underlying population of [CII] emitters likely suffering from low SNR (in terms of [CII] emission) possibly caused by the low SFRs and stellar masses which characterize them (see Figure 6.1). Following Solomon et al. (1992), we also compute $L_{\rm [CII]}$ as in Equation 3.1. We obtain $\log(L_{\rm [CII]}/L_{\odot}) = 7.7 \pm 0.1$, where the uncertainty is computed by propagating the error on the integrated flux on Equation 3.1. We test that this result is not significantly different from the one obtained by removing the constraint on the FWHM of the stacked spectra, which thus affects only the width of the line.

Figure 6.4 shows that the average [CII] profile presents a negative continuum at the right of the line. By analyzing each spectrum individually, we attribute this behaviour to $\sim$ 10 sources that present a negative continuum at positive velocity offset with respect to the expected line. However, we believe that this issue is only due to a statistical fluctuation originated from the small size of the sample, rather than to calibration problems in the ALMA data reduction or physical processes in these sources. For these reasons, we include them in the stacking procedure. The exclusion of these sources from the stacking mitigates the asymmetry of the line but does not change our final results. Indeed, in that case, we retrieve a [CII] luminosity that is compatible within 1σ with that obtained by considering the whole sample of 35 non-detections.

6.4 RESULTS

By taking advantage of the [CII] luminosity computed from the stack of the ALPINE non-detections, we explore the relation between $L_{\rm [CII]}$ and SFR in these galaxies in order to compare them to the results obtained for the combined [CII] detections and non-detections upper limits (Schaerer et al. 2020).

Following Schaerer et al. (2020), we report in Figure 6.5 the [CII] luminosities and SFRs of the ALPINE detections and non-detections (as 3σ upper limits; see also Figure 4 by Schaerer et al. 2020). Contrarily to Figure 6.1 in which we show the SFRs ob-

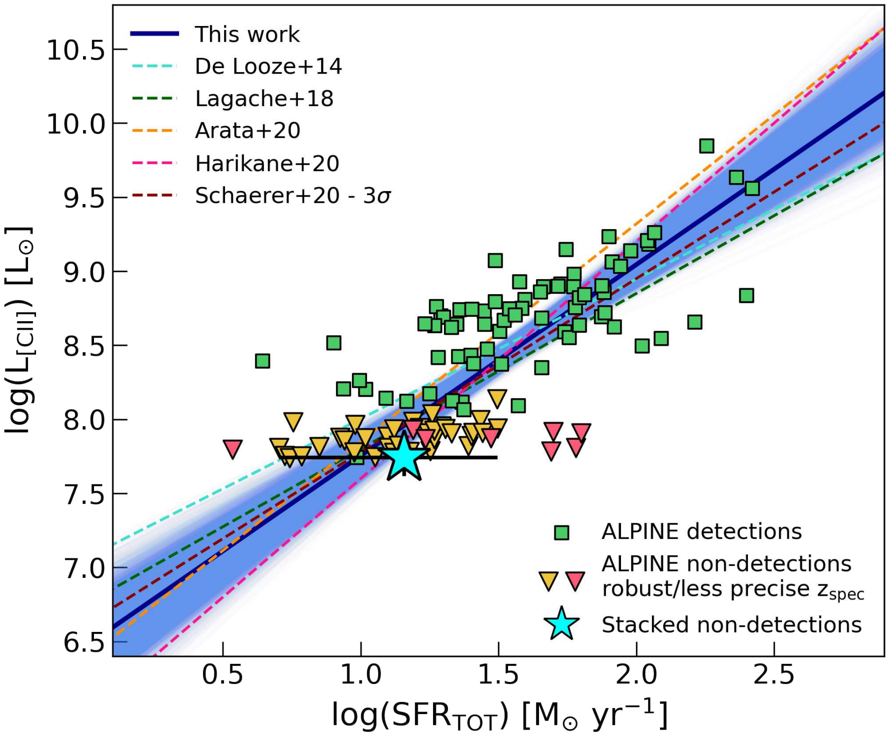


Figure 6.5: [CII] luminosity as a function of total SFR (UV+FIR; see text) for the ALPINE detections (green squares) and non-detections (as 3σ upper limits; yellow triangles). The red markers represent the [CII]-undetected galaxies with less precise z_{spec} (see Section 6.2). The cyan big star shows the result of this work for the stacked non-detections. The error on $L_{[CII]}$ is smaller than the star. We also report different $L_{[CII]}$-SFR relations from the literature: the local relation by De Looze et al. (2014) (dashed cyan line), the predicted relation by the models of Lagache et al. (2018) at $z \sim 5$ (dashed green line), the predicted relation by Arata et al. (2020) for galaxies at $z > 6$ (dashed yellow line), the fitted relation by Harikane et al. (2020) on $z > 6$ galaxies (dashed pink line), the latest results obtained by Schaerer et al. (2020) at $z \sim 5$ with ALPINE considering non-detections as 3σ upper limits (dashed red line). Finally, the solid blue line and shaded area report our best-fit and uncertainties to the [CII] detections + stacked non-detections.

Table 6.1: Summary of $L_{[CII]}$-SFR relations from the literature and from this work, as parameterized in Equation 6.1.

Literature (1)	Sample (2)	Redshift (3)	a (4)	b (5)
De Looze et al. (2014)	HII/starburst	< 0.5	7.06 ± 0.33	1.00 ± 0.04
Lagache et al. (2018)	`G.A.S. + CLOUDY`	4 - 6	6.75 ± 0.07	1.05 ± 0.07
Arata et al. (2020)	`GADGET-3 + ART`^{2}	>6	6.38	1.47
Harikane et al. (2020)	LBGs/SMGs	6 - 9	6.00	1.60
Schaerer et al. (2020)	ALPINE (3σ limits)	4 - 6	6.61 ± 0.20	1.17 ± 0.12
This work	ALPINE det + stacked non-det	4 - 6	6.46 ± 0.18	1.29 ± 0.12

tained through SED-fitting (for consistency with the stellar mass values obtained with the same procedure; Faisst et al. 2020b), we use in this case the total SFRs obtained as $\mathrm{SFR_{TOT}} = \mathrm{SFR(UV)} + \mathrm{SFR(IR)}$, where SFR(UV) comes from the observed UV absolute magnitude at 1500 , and SFR(IR) is obtained both through the ALMA continuum measurements and, for galaxies undetected in continuum, through predictions of the IRX-β relation by Fudamoto et al. (2020). In this way, we are able to compare our results with those by Schaerer et al. (2020) and with other $\mathrm{L_{[CII]}}$-SFR relations already present in the literature. Among these, we show: i) the local $\mathrm{L_{[CII]}}$-SFR relation found by De Looze et al. (2014) for a sample of low-z HII/starburst galaxies[2]; ii) the predicted relation found by Lagache et al. (2018) at $z = 5$; iii) the predictions from simulations for $z > 6$ galaxies by Arata et al. (2020); iv) the relation found by Harikane et al. (2020) for galaxies observed in [CII] at $6 < z < 9$; iv) the relation fitted to the ALPINE data only (including 3σ upper limits on non-detections) by Schaerer et al. (2020). Furthermore, we also show our results from the stacking of ALPINE non-detections, adopting the median SFR of the sample (i.e., $\log(\mathrm{SFR_{med}}/\mathrm{M_\odot\ yr^{-1}}) = 1.16$). As shown, our value is consistent with the 3σ upper limits found by Béthermin et al. (2020) and used by Schaerer et al. (2020) in their analysis.

We fit the combined [CII] luminosity and SFR from the stacking of non-detections with the ALPINE [CII]-detected galaxies with a linear relation of the form

$$\log(\mathrm{L_{[CII]}}/\mathrm{L_\odot}) = a + b \times \log(\mathrm{SFR}/\mathrm{M_\odot\ yr^{-1}}), \tag{6.1}$$

where a and b are the intercept and slope of the relation, respectively. For consistency with Schaerer et al. (2020), we use the `linmix` package[3] by Kelly (2007) which makes use of a Bayesian approach to account for measurement errors in both variables in linear regressions. The errors on the [CII] luminosities are taken from Béthermin et al. (2020) for the ALPINE detections, and from the stacking for the non-detected sources. Regarding the uncertainties on the SFRs, we use the combined errors from the corresponding UV and FIR quantities (that is, on average ~ 0.2 dex for both detections and non-detections).

---

[2] As this relation is based on a Kroupa & Weidner (2003) IMF, we scaled it to a Chabrier (2003) IMF by dividing the SFR by a factor 1.06 (e.g., Madau & Dickinson 2014), for consistency with other measurements from the literature.

[3] `https://github.com/jmeyers314/linmix`.

In Figure 6.5 we just represent the total SFR coverage of the non-detections used for the stacking as the horizontal bar.

With our stacked non-detections, we obtain a best-fit relation with $a = 6.46 \pm 0.18$ and $b = 1.29 \pm 0.12$. This is $\sim 2\sigma$ steeper than the slope obtained from local galaxies ($b = 1.00 \pm 0.04$; De Looze et al. 2014), and in agreement with that previously found by Schaerer et al. (2020) considering detections and 3σ upper limits on non-detections (i.e., $b = 1.17 \pm 0.12$). Table 6.1 summarizes the parameters describing the $L_{[CII]}$-SFR relations found in the literature and in this work.

6.5 SUMMARY

It is now well established that, in the local Universe, a linear relation between the [CII] luminosity and the SFR of galaxies is in place (e.g., De Looze et al. 2014). Whether this relation holds at earlier epochs is still debated. Lagache et al. (2018) used the semi-analytical model `G.A.S.` (Galaxy Assembly from dark-matter Simulations) in combination with the photoionization code `CLOUDY` (Ferland et al. 2013, 2017) to predict the [CII] luminosity of a large number of galaxies at $z \sim 5$. They found an average relation with a slope in agreement with the one by De Looze et al. (2014) for HII/starburst galaxies at low redshift, but that is dependent on several parameters, such as the metallicity of the galaxies and the intensity of their interstellar radiation field. More recently, Arata et al. (2020) combined cosmological hydrodynamic simulations performed with the `GADGET-3` code (Springel 2005), with the All-wavelength Radiative Transfer with Adaptive Refinement Tree ($\mathtt{ART}^2$) code (Li et al. 2008; Yajima et al. 2012) to predict the relation between [CII] and SFR for galaxies well within the Reionization epoch, at $z > 6$. They found a steep slope (i.e., $b = 1.47$) of the relation, suggesting that the deviation from the local Universe is caused by changes in the distribution of neutral gas in high-z galaxies. Observationally, similar results were obtained by Harikane et al. (2020) who found a very steep slope (i.e., $b = 1.6$) by analyzing a sample of $6 < z < 9$ LBGs and SMGs. Schaerer et al. (2020) investigated the $L_{[CII]}$-SFR relation for the first time with a statistical sample of SFGs at $4 < z < 6$ through the ALPINE survey. By exploiting the whole ALPINE sample (consisting of of 75 detections and 43 non-detections), they found that the $L_{[CII]}$-SFR relation continues to apply even at high redshift, with a possible weak deviation from the local Universe. Possible explanations for the differences between the Schaerer et al. (2020) results and those described above from the literature could reside in the more statistically significant dataset available from the ALPINE survey, and/or in the uncertainties introduced by the derivation of the SFRs (see Schaerer et al. 2020 for more details).

One of the major uncertainties in the $\mathrm{L}_{[\mathrm{CII}]}$-SFR relation obtained by Schaerer et al. (2020) derives partially from the group of ALPINE non-detected galaxies, which makes $\sim 36\%$ of the total sample. In principle, these sources could deviate from the $\mathrm{L}_{[\mathrm{CII}]}$-SFR relation followed by the other ALPINE detections, or they could just have [CII] emissions weaker than the other galaxies of the sample.

To remove this source of uncertainty, in this chapter we characterized the population of [CII] non-detected galaxies in ALPINE by stacking their ALMA spectra, taking into account the typical observed rest-frame UV-FIR spectral offset between the Lyα and [CII] lines (Cassata et al. 2020). The stack reveals a [CII] detection at $\sim 5.8\sigma$, resulting in a line luminosity of $\log(\mathrm{L}_{[\mathrm{CII}]}/\mathrm{L}_{\odot}) \sim 7.7$. This highlights also that the [CII] non-detections are not drawn from a different population of galaxies with respect to the ALPINE detections. Rather, these are galaxies lying on the bottom-left region of the $z \sim 5$ main-sequence with lower SFRs and stellar masses, and therefore fainter [CII] emission. By fitting the [CII] luminosity from the stacking of the non-detections and from the [CII]-detected galaxies as a function of their SFRs, we find a linear relation having a $\sim 2\sigma$ steeper slope than the local relation (1.29 ± 0.12 from this work against 1.00 ± 0.04 from De Looze et al. 2014) and in agreement with the previous results by Schaerer et al. (2020). It is interesting to note that our slope is also consistent with that found by De Looze et al. (2014) for metal-poor dwarf galaxies in the local Universe (i.e., $b \sim 1.25$). This is in line with the results from Faisst et al. (2020b) who compared the Hα (which is a good tracer of the star-formation properties of galaxies) and [CII] luminosities for a sub-sample of the ALPINE galaxies. They found that bright [CII] galaxies are in good agreement with the local relation between Hα and [CII] found by De Looze et al. (2014). However, for lower [CII] luminosities ($< 5 \times 10^8$ $\mathrm{L}_{\odot}$, as those probed by the [CII] non-detections), the galaxies seem to be more consistent with the relation found for local metal-poor dwarf galaxies, although with a large scatter. This further suggests that the metallicity of galaxies (along with the strength of their [CII] emission) could play an important role in the derivation of the $L_{[CII]}$-SFR relation (Vallini et al. 2015; Olsen et al. 2017; Lagache et al. 2018). Future investigations of this topic will be possible thanks to the forthcoming launch of JWST, that will be able to provide measurements of the metallicity content of distant galaxies.

It is worth noting that, both in the stack and in the fit, we do not include 10 non-detections having less precise spectroscopic redshifts than those from the other galaxies in the sample (see Section 6.2.2). As evidenced in Figure 6.5, some of these sources show the largest deviation from the derived $L_{[CII]}$-SFR relation, suggesting the possible presence of the so-called [CII]-deficit (e.g., Malhotra et al. 2001b; Vallini et al. 2015; Lagache et al. 2018; Harikane et al. 2020). However, these galaxies are likely affected by a poor estimate of their spectroscopic redshifts with respect to the other ALPINE non-detections. Indeed, for the analysis undertaken in this work, a not accurate derivation of $z_{\rm spec}$ could induce a not physical offset between the rest-frame UV spectroscopic redshift and the systemic one as traced by the [CII] line that we are not able to correct based on the observed $\Delta v_{\rm Ly\alpha}$ distribution. At worst, the expected emission line could be moved outside of the ALMA SPW of observation. In this scenario, we claim that there is no evidence of [CII]-deficit within the ALPINE sample, as also suggested by Schaerer et al. (2020).

Further and deeper observations are needed in order to firmly establish the nature of [CII] as a tracer of SFR in the early Universe, especially in the low SFR regime. Indeed, only a handful of strongly lensed galaxies have been detected so far at $\log(\mathrm{SFR}/\mathrm{M}_\odot\ \mathrm{yr}^{-1}) \lesssim 0.5$ (e.g., Knudsen et al. 2016; Fujimoto et al. 2021), where the $L_{[CII]}$-SFR relation is consistent from the spatially-resolved and the entire galaxy scales. In this context, the constraint provided in this work on the low $L_{[CII]}$ and SFR tail of normal high-z SFGs could serve as an input for cosmological simulations in order to provide fundamental insights on the physics of [CII] in the distant Universe.

7 | CONCLUSIONS AND FUTURE PERSPECTIVES

In this thesis, I have exploited the data collected by the ALPINE survey to investigate the properties of primordial galaxies in the early Universe in the context of galaxy formation and evolution. I have taken advantage of the morphological and kinematic information provided by the cold gas, as traced by the [CII] 158 μm line, to characterize the nature of dust-obscured galaxies and their incidence at high redshift, as well as the physical mechanisms that govern the rise of the SFRD during the period of rapid mass-assembly at $4 < z < 6$, with a closer look at the galaxy merging.

The recent discoveries of very dusty objects in the early Universe have revived the interest in the amount of obscured star formation at $z > 3$. A precise estimate of the SFRD across cosmic time is essential to depict the SFH of the Universe and give information on the origin and evolution of galaxies. Therefore, a complete census of dust-obscured sources at high redshift is needed to understand if the individual detections of such galaxies are part of the tail of the observed distribution peaked at $2 < z < 3$ (e.g., Swinbank et al. 2014), or if they are representative of a larger population of dusty objects in the distant Universe, which could provide a significant contribution to the cosmic SFRD at $z > 4$.

In this respect, we investigated the nature of dark galaxies observed serendipitously in the field of view of some ALPINE targets. These objects (4 out of 14 serendipitous detections) stand out for their large amount of dust, which makes them completely (or mostly) invisible in the optical bands, lacking information on their spectroscopic and/or photometric redshift. The emission lines arising from these sources and observed through the ALMA telescope could, in principle, originate from many atomic or molecular species other than [CII], which place them at different cosmic epochs, posing questions about their contribution to the global SFRD. We analyzed the brightest serendipitous emission line among the ALPINE galaxies with no counterpart (Romano et al. 2020), and used that as a pilot study for the characterization of the remaining sources. We exploited the available multi-wavelength and spectroscopic information to deduce the most likely nature of the line, both by using the FIR and line luminosities as diagnostics, and through SED-fitting procedures. All in all, we found more evidence in favor of a dusty SFG located ~ 40 kpc away from the main UV-selected target, and possibly part of a massive overdensity at $z \sim 4.6$. We then used similar methods to set the remaining three dusty objects to their correct cosmic epoch. We found that, including the pilot serendipitous galaxy, 3 out of 4 sources are most likely [CII] emitters at the same redshift of the ALPINE targets (with the remaining one also showing possible hints of a high-z nature). In case the 4 galaxies with no optical counterparts were all confirmed to be at $z > 4$, the fraction of serendipitous [CII] emitters would raise from $\sim 57\%$ to $\sim 86\%$. These galaxies could provide a SFRD at $z \sim 5$ that is up to 2 times larger than previous estimates from UV surveys, suggesting an almost constant distribution from $z \sim 1$ to the end of HI reionization, and implying a significant and increasing contribution of dust-obscured star formation through early epochs (Gruppioni et al. 2020; Fudamoto et al. 2021; Loiacono et al. 2021).

To obtain the global picture, we also need to know which are the physical processes that rule the evolution of galaxies from their origin to the present day. The debate is still ongoing on which are the main mechanisms responsible for the build-up of galaxies, with the star formation and major mergers as the most popular ways through which primordial sources increase their stellar mass across the cosmic ages (e.g., Bouché et al. 2010; Duncan et al. 2019).

To shed light on this topic we used, for the first time, both the morphological and kinematic information provided by [CII] to identify major mergers at $z \sim 5$, and to assess their relevance in the framework of galaxy mass-assembly (Romano et al. 2021). At first, we computed the major merger fraction at $z \sim 5$ from the ALPINE data. The

[CII] emission, in combination with the other multi-wavelength data exploitable for the ALPINE targets, allowed us to obtain the most robust sample of mergers at early epochs not biased by dust. We found that $\sim 40\%$ of normal galaxies at that redshift is composed of merging sources, providing the first constraint on the merger fraction from [CII] and in the stellar mass range probed by ALPINE (i.e., $\log(M_*/M_\odot) \sim 10$) at this epoch. When combined with data from the literature at lower redshift, our results produce a cosmic evolution of the merger fraction which is characterized by a rapid increase from the local Universe to $z \sim 2$, a peak at $z \sim 2-3$, and a slow decline toward earlier epochs, in good agreement with state-of-the-art hydrodynamical simulations. We then converted the merger fraction into the galaxy merger rate (i.e., the number of mergers per galaxy and Gyr), that is needed to estimate the contribution of the merging process to the mass-assembly. We also assumed three different timescales for the merger to happen and studied their impact on our final results. We used the merger rate to obtain the mass accreted through major mergers per unit time and volume, and compared that with the same quantity from star formation (i.e., the SFRD). We found that major mergers contribute less than 10% to the galaxy mass-assembly at all epochs, with the star formation as the main driver of the SFRD evolution. On the other hand, we also found that the contribution of major mergers to the galaxy build-up could be substantial (but always smaller than *in-situ* star formation) at $z \sim 5$ if considering the most recent results on the merger timescale, for which it decreases rapidly toward earlier epochs (Snyder et al. 2017).

As a final quest, we probed the properties of the cold gas in the distant Universe through [CII] observations. In particular, we investigated the use of [CII] as a SFR indicator at high redshift (Romano et al., submitt.). Indeed, it is well established that there is a good correlation between the [CII] luminosity and SFR in the local Universe, but the situation may be different for high-z galaxies. Simulations show that, for instance, the lower metallicity content of these sources or the strength of the interstellar radiation field could change the slope of the relation (e.g., De Looze et al. 2014; Vallini et al. 2015). Previous results from ALPINE showed a good agreement with the local $L_{[CII]}$-SFR relation, although when considering [CII] non-detections within the sample a marginally steeper slope was found (Schaerer et al. 2020).

To clarify this discrepancy, we characterized the population of [CII]-undetected galaxies in ALPINE through a stacking of their ALMA spectra. We found a resulting 5.8σ detection from which we obtained a [CII] luminosity of $\log(L_{[CII]}/L_{\odot}) \sim 7.7$. By combining this result with those from the ALPINE detections we found a $L_{[CII]}$-SFR slope marginally steeper than the local one, but in agreement with that within 2σ. We also did not find evidence in favor of the [CII]-deficit observed in other works at the same redshifts. We concluded that [CII] could still be considered a good tracer of star formation in the early Universe, although further constraints at low SFR and stellar masses are needed to give a solid answer to this issue.

To conclude, in this thesis we have used the radiation arising from dust grains and ionized carbon atoms in primordial galaxies to light up the dark and distant Universe. We have exploited the [CII] emission collected by the ALPINE project as a probe to explore the early stages of galaxy formation and to understand the mechanisms that rule the evolution of young sources through cosmic ages. We have found that a large population of dusty galaxies, in most cases missed by UV/optical surveys, is in place at $z > 4$. Such an abundance of heavily dust-obscured objects at high redshift is not predicted by state-of-the-art galaxy formation models (e.g., Henriques et al. 2015; Pillepich et al. 2018), highlighting the necessity of revising our current knowledge on galaxy evolution and dust production in the early Universe. We have ultimately provided new hints on the assembly of galaxies through cosmic time, setting a point of reference for innovative cosmological hydrodynamic simulations.

Certainly, in order to confirm these results, surveys sampling wider sky areas and looking for the elusive population of optically-invisible dusty galaxies are needed. A larger statistics will allow us to understand if such objects are common in the distant Universe, and to quantify with accuracy their effective contribution to the obscured star formation at different cosmic epochs, also posing constraints on the early phases of galaxy formation. Furthermore, high-resolution observations will allow us to better constrain the major merger fraction at $z \sim 5$ and to improve cosmological simulations to provide a better description of the merger timescale, representing by now the main obstacle to obtain a fair estimate of the merger rate.

In the near future, the combination between current and forthcoming facilities, like

ALMA and JWST, will allow us to further investigate the dark side of the Universe and to characterize the structure of dusty galaxies and their environments through large programs and follow-up studies.

A | INDIVIDUAL DESCRIPTION OF MERGERS

We present here the characterization of the mergers analyzed in this work, as illustrated in Section 5.2.1 and Figure 5.2. For each source, we show the moment maps, the PVDs, and the integrated [CII] spectrum. Due to the peculiarity of these objects, we briefly comment on them in the following, also highlighting possible differences and similarities with the previous morpho-kinematic classifications.

CG_38: Le Fèvre et al. (2020) classified this galaxy as a merger. Although the low SNR and small spatial extent, we include this source in our analysis as it shows multiple peaks in the PVDs that are associated to different components in the [CII] spectrum.

DC_308643: this source was classified as a merger by Le Fèvre et al. (2020). The moment-1 map shows a clear velocity gradient suggesting the presence of a rotating disk. However, a similar velocity map could also be reproduced by two merging components in an advanced phase of merging. This is suggested by the two close components visible in the PVD along the major axis and by the shape of the [CII] line which deviates from that of a single Gaussian.

DC_372292: we confirm the merger classification of this object by Le Fèvre et al. (2020). As for DC_308643, a possible velocity gradient is present in the moment-1 map. However, two components separated in velocity by $\sim$ 200 km s^{-1} are clearly visible in the PVDs and in the spectrum.

DC_378903: we classify this source as a merger as in Le Fèvre et al. (2020). Despite the low SNR and spatial extent, we manage to decompose the [CII] spectrum in two Gaussian components that are also visible at 3σ in the PVDs and in the channel maps.

DC_417567: this source was classified as a merger by Le Fèvre et al. (2020) because of the disturbed [CII] morphology and the presence of multiple peaks in the PVDs. These are reproduced by different Gaussian components in the overall [CII] spectrum. This target was also analyzed by Jones et al. (2021) which classified it as uncertain because of its low SNR. However, the poor ^{3D}Barolo fit supports the merger classification for this object.

DC_422677: as in Le Fèvre et al. (2020), we classify this source as a merger, mainly because of the presence of multiple components in the PVDs and in the spectrum.

DC_434239: this object was classified as a merger both by Le Fèvre et al. (2020) and Jones et al. (2021). The [CII] morphology is extended and disturbed, as well as that of the UVista K_s band. The PVDs show multiple peaks and we need at least three components to reproduce the global shape of the [CII] spectrum.

DC_493583: we classify this source as a merger as in Le Fèvre et al. (2020). Two separate components are visible both in the PVDs and in the spectrum, with a fainter minor object emerging from the noise at $v \sim -200$ km s^{-1}.

DC_519281: this source was classified as a merger by Le Fèvre et al. (2020). The complex velocity map and PVDs sustain this interpretation. Further evidence for ongoing merging activity is provided by the analysis of Jones et al. (2021) (even if they classify this object as uncertain because of the low SNR and spatial resolution).

DC_536534: the [CII] emission from this galaxy is spatially resolved in different clumps, allowing Le Fèvre et al. (2020) to classify it as a merger. The PVDs are clearly disturbed and multipeaked, and the integrated spectrum shows the presence of at least three different emitting components.

DC_665509: previously classified as a merging system (Le Fèvre et al. 2020), this source shows multiple components both in the spectrum and in the PVDs with the two major emissions at $v \sim 0$ and $v \sim 150$ km s^{-1}.

DC_680104: this object was classified as a merger by Le Fèvre et al. (2020). It shows a faint secondary component at 3σ in the moment-0 map which is also visible in the PVD along the major axis at a spatial offset of ~ -1 arcsec with respect to the main target. Although the SNR of this source is low, we consider it as a merger, mainly because of two peaks of emission in the PVDs at the same spatial position but separated by ~ 150 km s^{-1} in velocity. These are also visible in the [CII] spectrum.

DC_814483: the presence of multiple components in the optical images and the quite disturbed morphology of the [CII] emission led Le Fèvre et al. (2020) to classify this galaxy as a merger. We confirm the previous classification also noting the complex velocity field and the multiple components in the spectrum, possibly associated to peaks of emission in the PVDs.

DC_818760: this source was studied in detail by Jones et al. (2020) and then further analyzed by Jones et al. (2021). It is likely a triple merger, with the two main components visible in the optical maps and close in velocity (as shown by the PVD along the major axis), and another fainter source completely obscured in the optical and separated both in space and velocity from the other two objects. We thus confirm the previous merger classification by Le Fèvre et al. (2020) and Jones et al. (2021). In this analysis, we use a pseudo-slit 7.5 arcsec wide to recover the full [CII] emission, and we consider only the two close components with the larger [CII] fluxes and optical emission.

DC_834764: this source was classified as an extended dispersion-dominated galaxy by Le Fèvre et al. (2020). However, the peaks of emission in the PVDs, the shape of the emission line, and the quite disturbed velocity field suggest the presence of multiple components. We thus classify this object as a merger.

DC_842313: this galaxy is quite peculiar and was classified as a merger by Le Fèvre et al. (2020). The optical maps show the presence of two or three components. Indeed, the ALPINE target is at the center of the maps, clearly visible in the HST F814W image, while a brighter and larger component also detected in [CII] is present toward the north, particularly visible in the UVista K_s-band map. This is the well-studied SMG J1000+0234 (e.g., Capak et al. 2008; Schinnerer et al. 2008; Jones et al. 2017), and it was serendipitously observed in ALPINE (Loiacono et al. 2021). This galaxy has a large line width (i.e., ~ 1000 km s^{-1}) and is separated in velocity by ~ 750 km s^{-1} from the main ALPINE target. As we are considering only those systems with a velocity separation $\Delta v \leq 500$ km s^{-1} (assuming these components as gravitationally bound; Patton et al. 2000), we should exclude it from our analysis. However, considering only the emission arising from the spectral channels associated to the main target DC_842313, we find further evidence of an ongoing merging. In particular, we note a disturbed [CII] morphology elongated toward the extended K_s-band contours in the southwest region of the optical image, and peaks of emission in the PVDs coinciding with multiple components in the spectrum.

For these reasons, we think that this could be a triple merging system, and we keep it in the analysis of the merger fraction. In this case, the moment maps and PVDs show the emission from both the main target and J1000+0234, obtained considering all the spectral channels including the [CII] emission from the two sources. The bottom panel displays instead the individual normalized integrated spectra of the two components.

DC_859732: this galaxy was classified as a merger by Le Fèvre et al. (2020). Two distinct components are present in the [CII] spectrum, partially visible in the PVDs due to the low SNR.

DC_873321: this galaxy was classified as a merger both by Le Fèvre et al. (2020) and Jones et al. (2021). The [CII] emission is elongated toward the two bright sources visible in the UVista K_s band. These two components are at the same velocity, being indistinguishable in the PVDs and their spectra are shown in the bottom panel of Figure A.11 (right).

vc_5100541407: the extended [CII] morphology coincident with the two optical components visible in the maps, allowed Le Fèvre et al. (2020) to characterize this source as a merger. Jones et al. (2021) classified it as uncertain instead, mainly because of their adopted classification criteria. Given the morphological and kinematic information on this galaxy, we also consider it as a merger of two components emitting at the same velocity. A pseudo-slit 4.5 arcsec wide is used to compute the PVDs.

vc_5101209780: this galaxy was classified as a merger both by Le Fèvre et al. (2020) and Jones et al. (2021), and it was the subject of a further and in-depth analysis by Ginolfi et al. (2020a). The [CII] moment-0 map clearly shows the presence of two major components, as also confirmed by the optical images. Two fainter components are also found in the spectrum and PVDs, which are computed by adopting a pseudo-slit 4.5 arcsec wide.

vc_5180966608: Le Fèvre et al. (2020) classified this object as a merger, while Jones et al. (2021) considered it as uncertain, mainly because of the [CII] spectrum which resembles that of a single, extended source. We find more evidence for a merging system. Indeed, at least two optical components are visible both from HST and UVista within the 3σ [CII] emission contours. The velocity gradient is quite disturbed and several peaks of emission are present in the PVDs, more visible along the minor axis (which is aligned with the two optical sources).

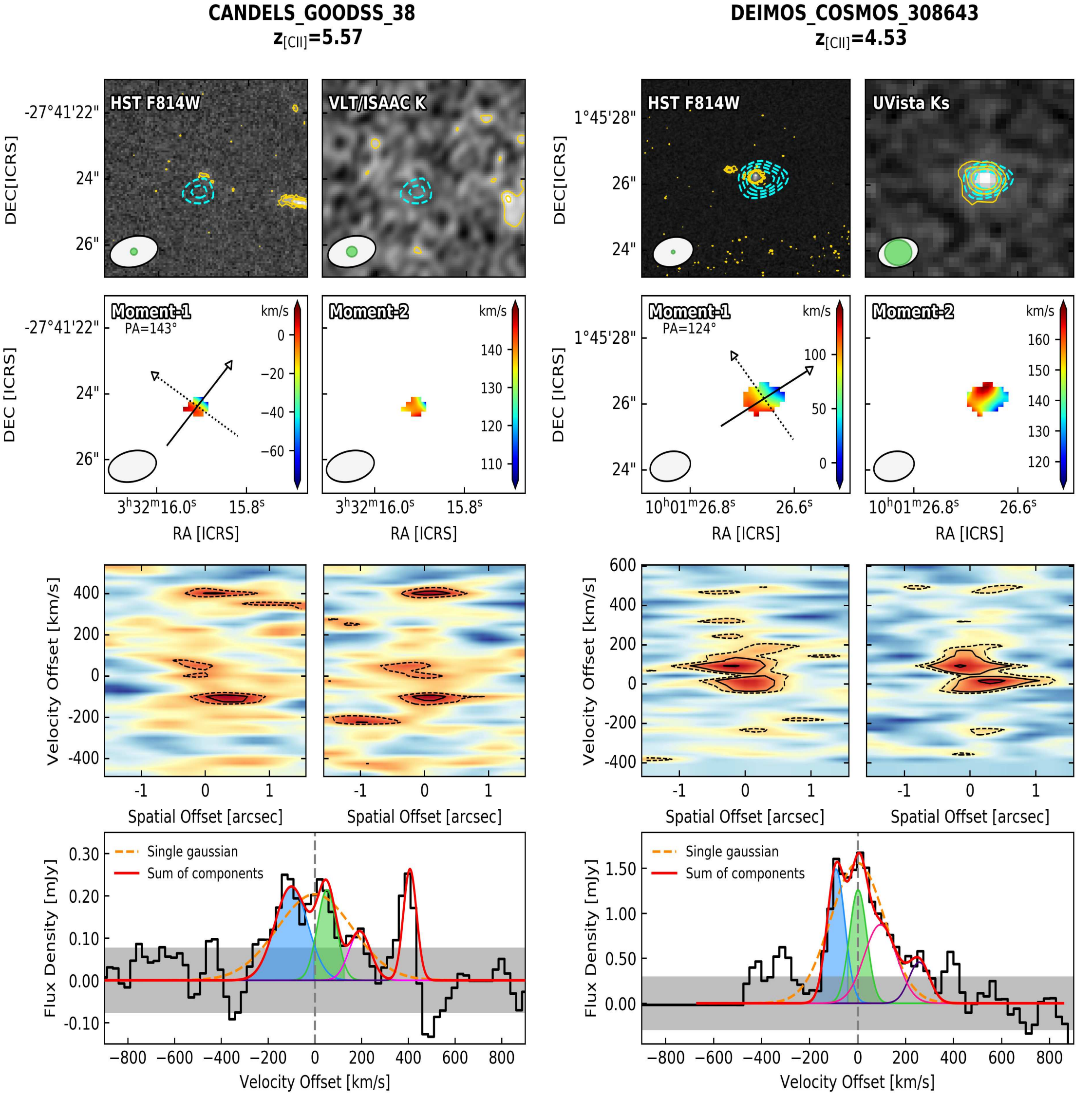


Figure A.1: Morpho-kinematic analysis of the ALPINE targets. *First row*: HST/ACS F814W (left) and UVista DR4 K_s-band (right) images centered on the UV rest-frame position of the target. Each cutout is $6'' \times 6''$ wide. The cyan contours show the [CII] ALMA emission starting from 3σ above the noise level. Yellow contours in the optical maps represent 3, 5 and 7σ emission. In the lower left corner, the ALMA beam (white) and HST or UVista resolutions (green) are displayed. *Second row:* moment-1 (left) and moment-2 (right) maps color-coded for the velocity and velocity dispersion in km s^{-1}. The velocity map reports the direction of the major (solid) and minor (dashed) axis (centered on the coordinates returned by the best-fit 2D Gaussian model on the moment-0 map) along which the PVDs are computed.

Figure A.1 ***(cont.)*** **:** *Third row:* PVDs along the major (left) and minor (right) axis color-coded for the flux intensity in each pixel. Dashed contours include the 2σ emission in the maps while 3, 5 and 7σ emission is represented by solid lines. *Fourth row:* [CII] spectrum (black histogram) extracted within the 3σ contours of the intensity map. The gray shaded band marks the 1σ level of the spectrum while the dashed vertical line shows the zero velocity offset computed with respect to the redshift of the [CII] line. Colored thin lines show individual possible components of the merging system, resulting in the global profile in red. The shaded areas under the curves represent the channels used to compute the [CII] intensity maps of the major and minor individual components. A single Gaussian fit is also visible with a dashed-orange line.

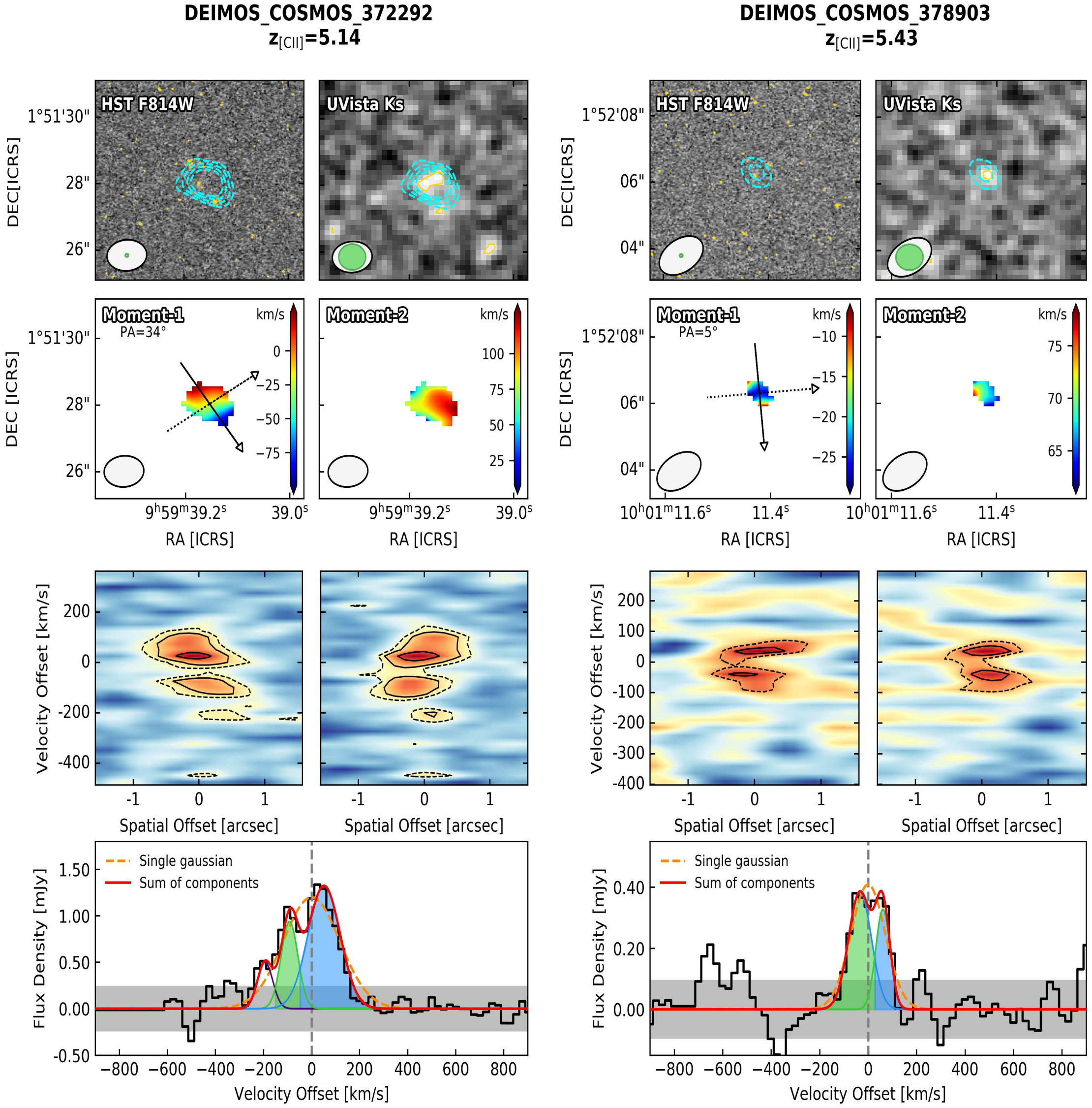


Figure A.2: Moment maps, PVDs and spectral decomposition for the ALPINE mergers, as described in Figure A.1.

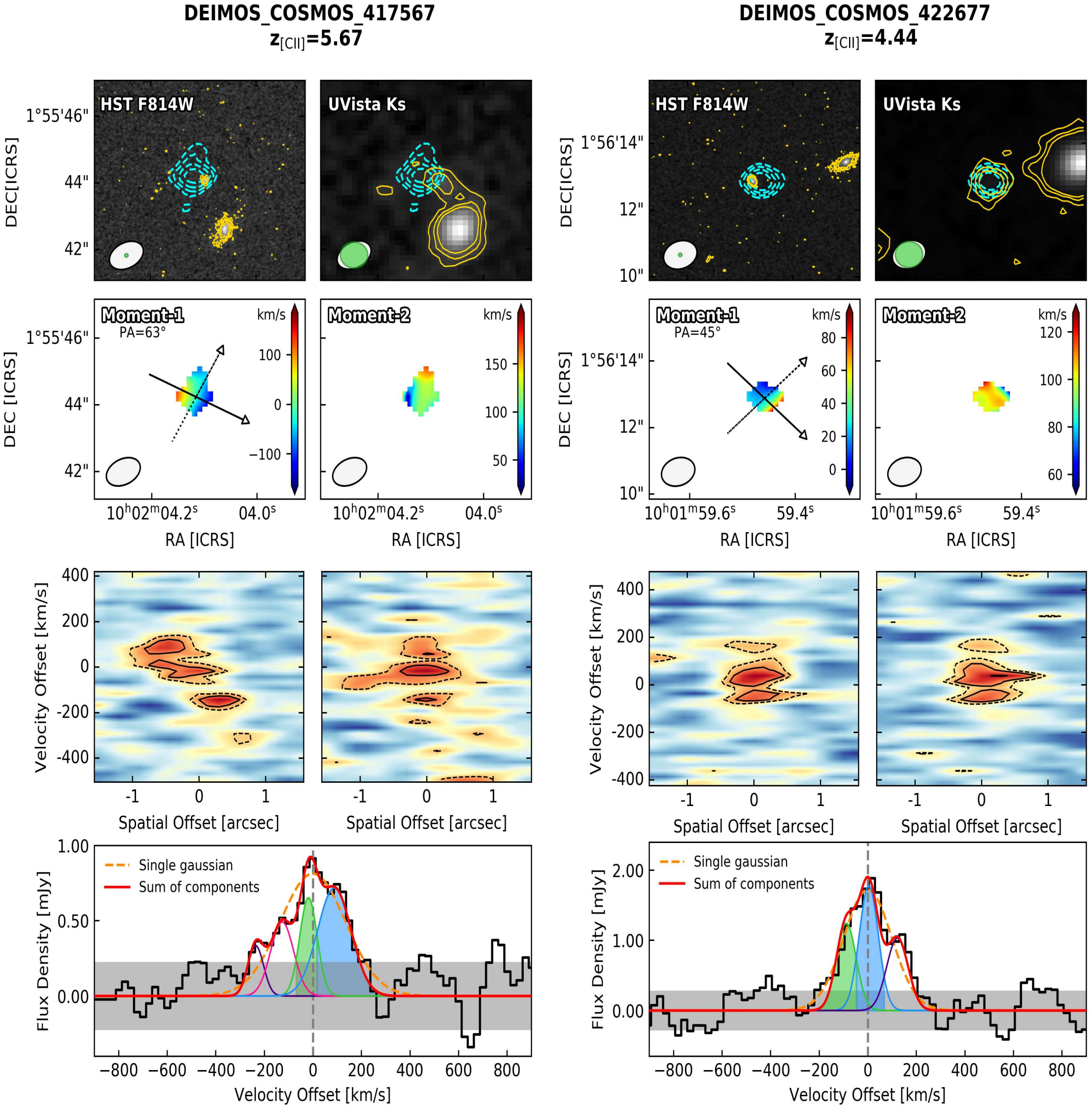


Figure A.3: Moment maps, PVDs and spectral decomposition for the ALPINE mergers, as described in Figure A.1.

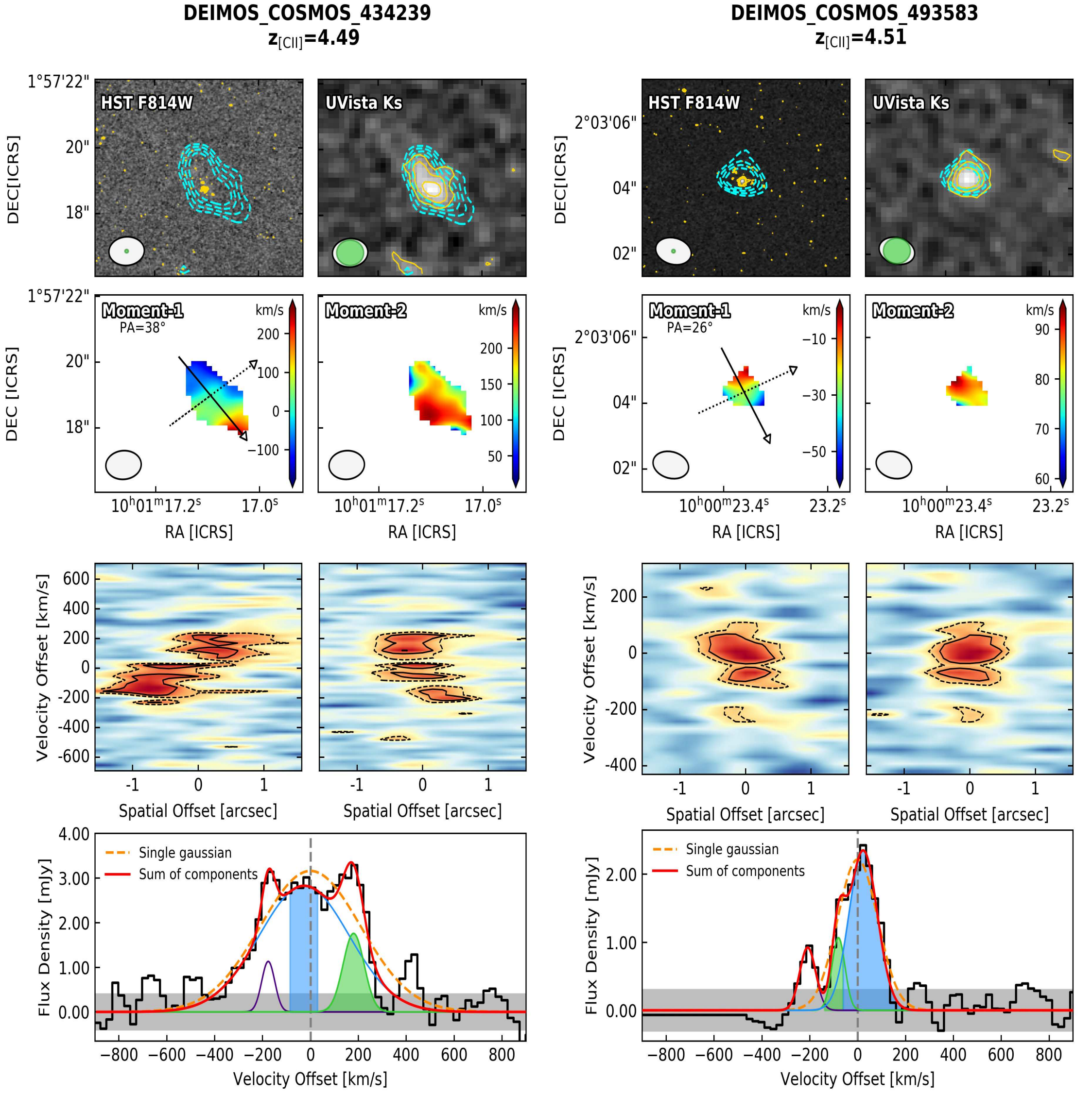


Figure A.4: Moment maps, PVDs and spectral decomposition for the ALPINE mergers, as described in Figure A.1.

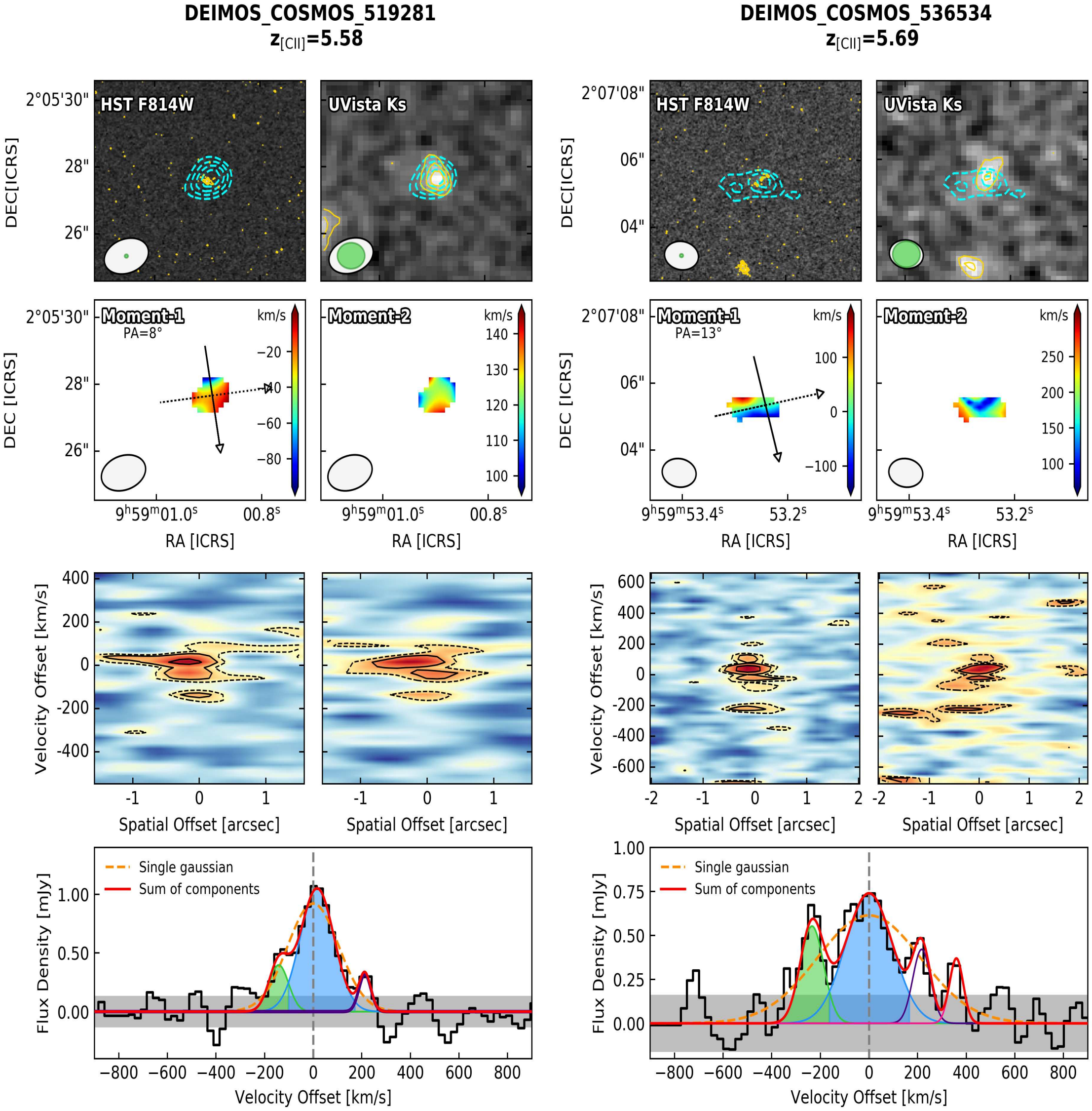


Figure A.5: Moment maps, PVDs and spectral decomposition for the ALPINE mergers, as described in Figure A.1.

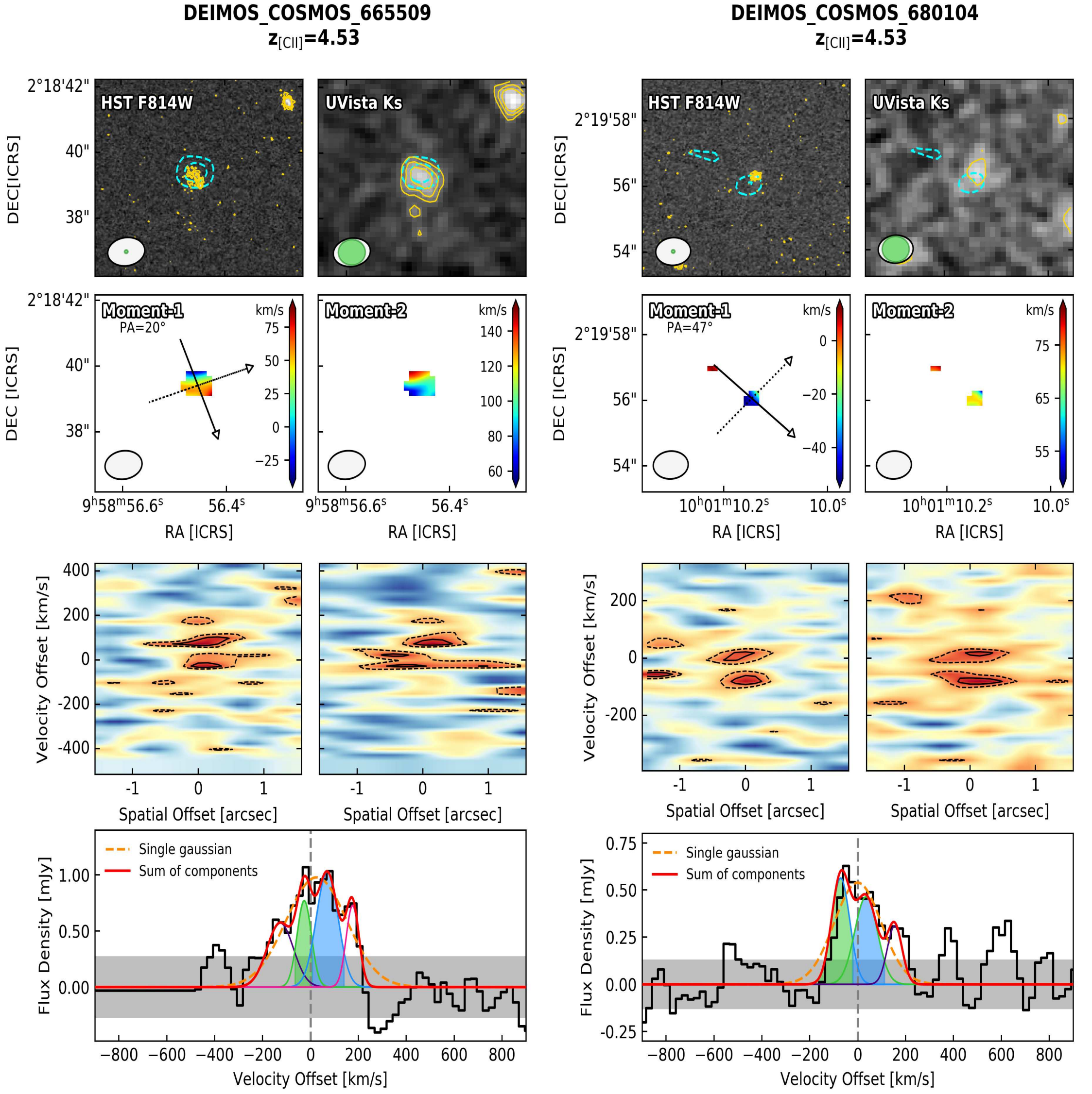
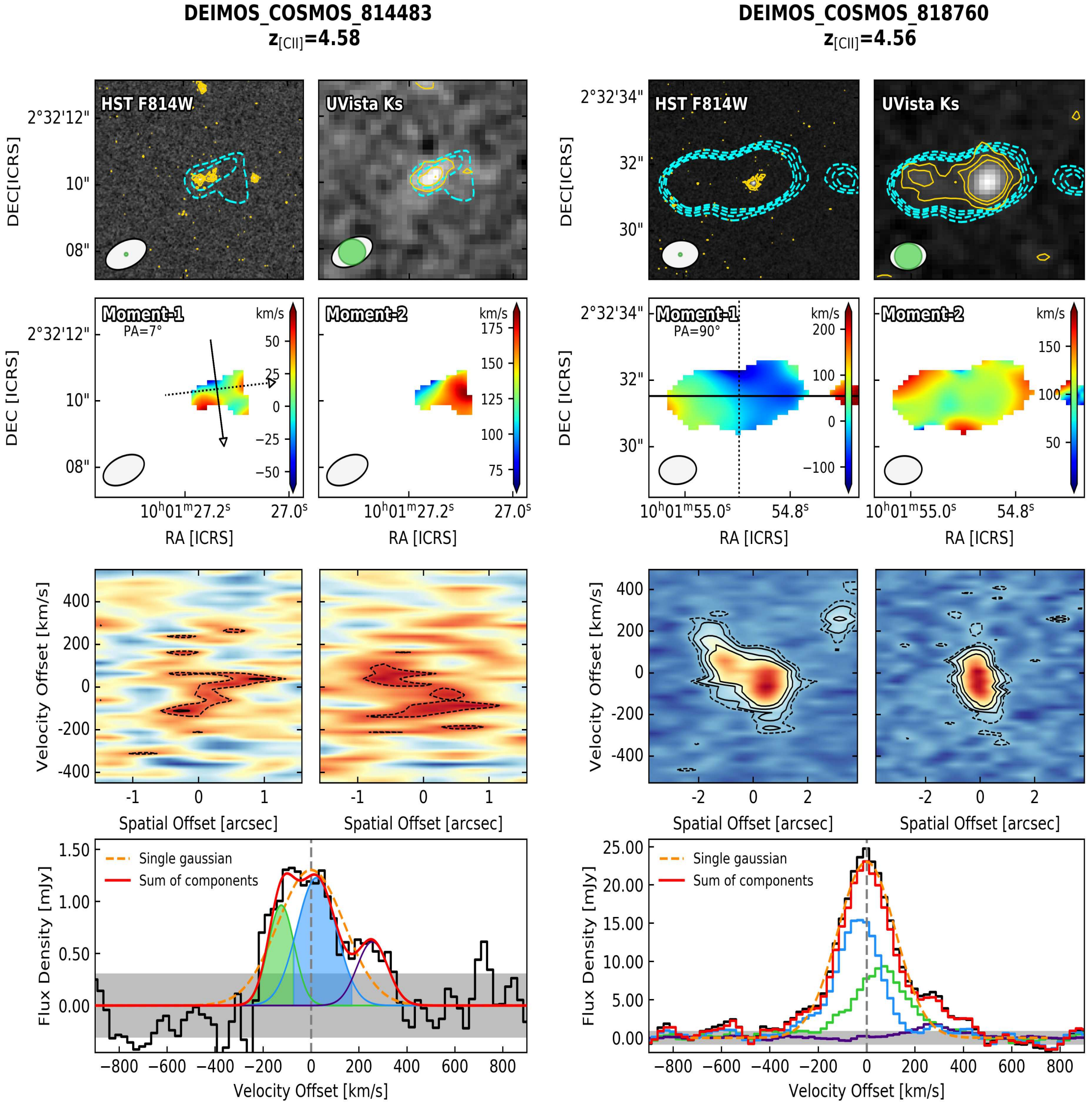


Figure A.6: Moment maps, PVDs and spectral decomposition for the ALPINE mergers, as described in Figure A.1.


DEIMOS_COSMOS_814483 $z_{[CII]}$=4.58

DEIMOS_COSMOS_818760 $z_{[CII]}$=4.56

Figure A.7: Moment maps, PVDs and spectral decomposition for the ALPINE mergers, as described in Figure A.1.

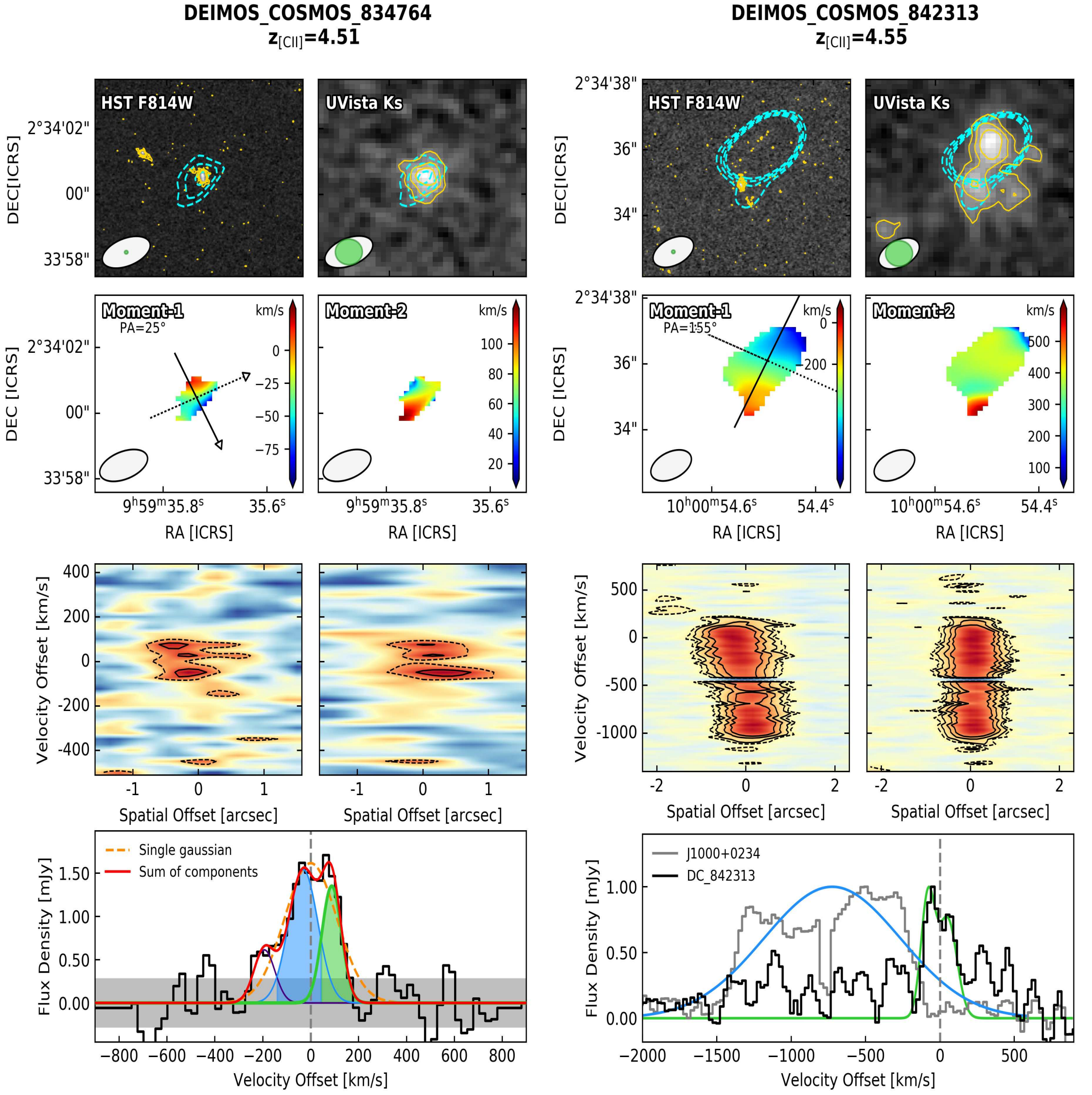


Figure A.8: Moment maps, PVDs and spectral decomposition for the ALPINE mergers, as described in Figure A.1.

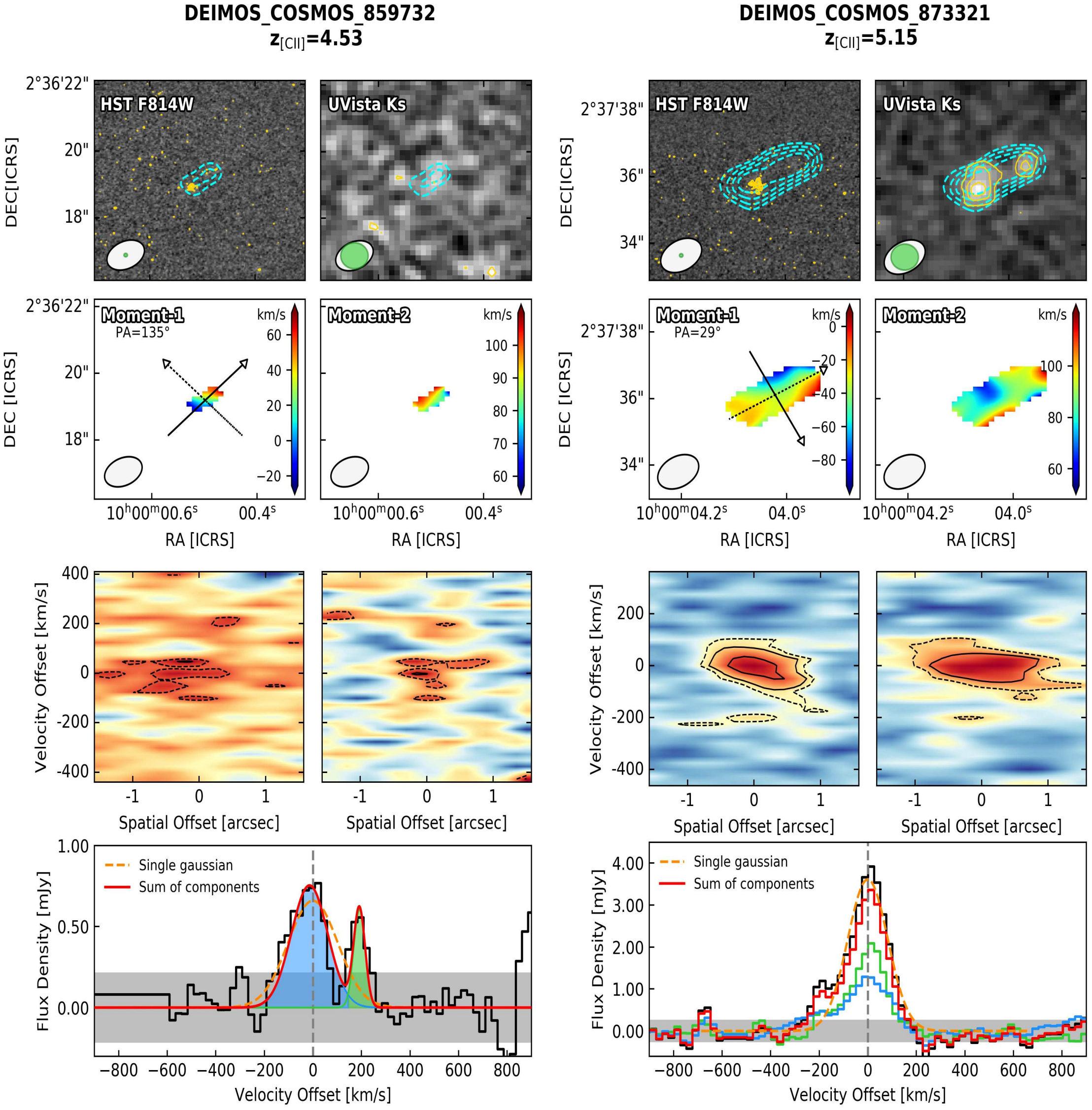


Figure A.9: Moment maps, PVDs and spectral decomposition for the ALPINE mergers, as described in Figure A.1.

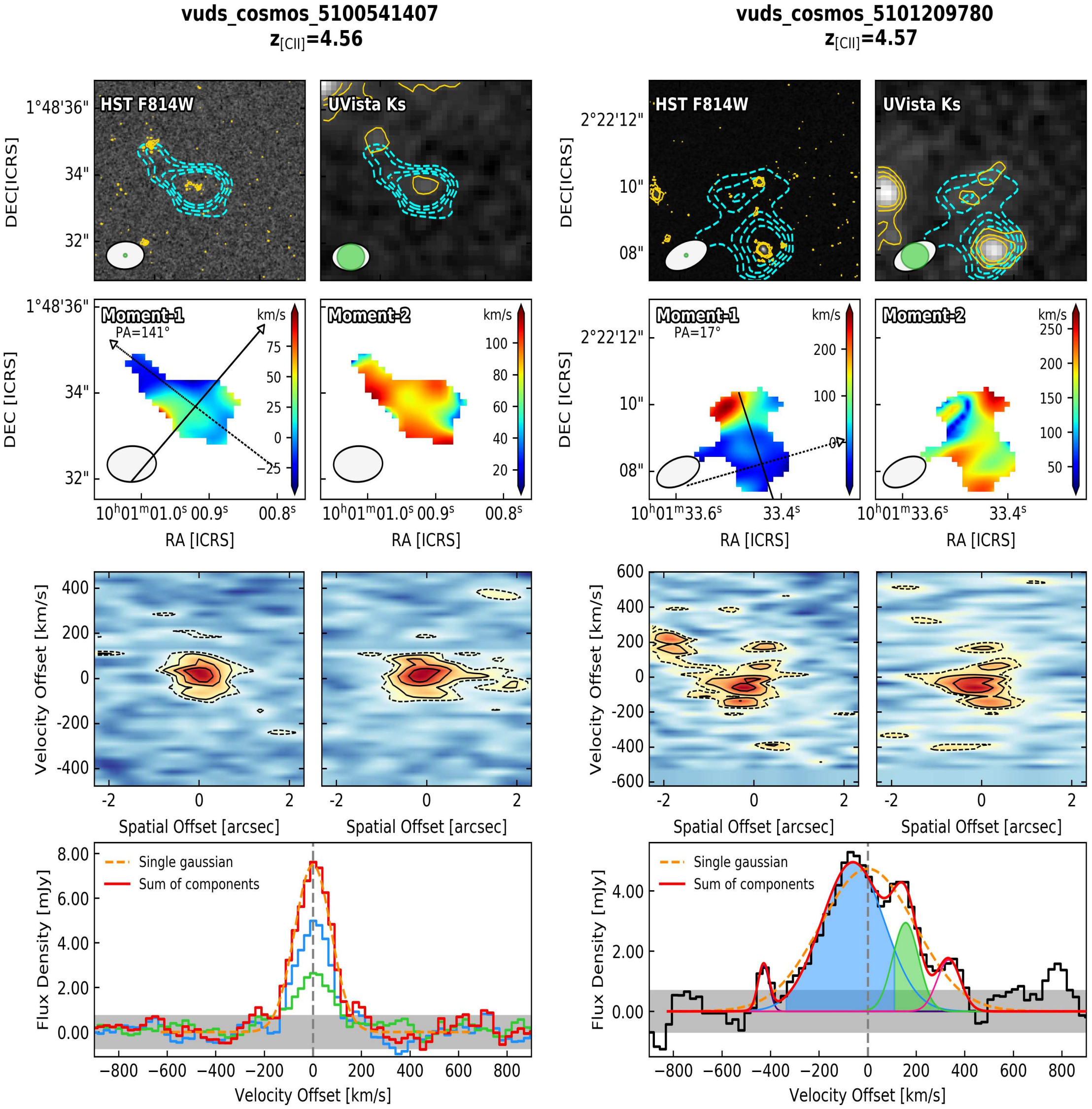


Figure A.10: Moment maps, PVDs and spectral decomposition for the ALPINE mergers, as described in Figure A.1.

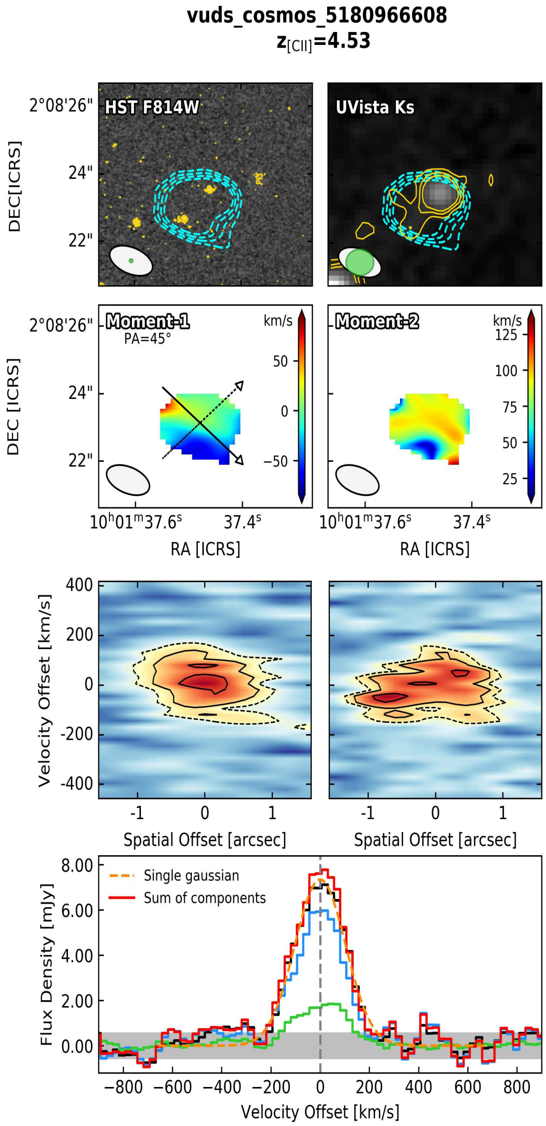


Figure A.11: Moment maps, PVDs and spectral decomposition for the ALPINE mergers, as described in Figure A.1.

BIBLIOGRAPHY

Alcalde Pampliega, B., Pérez-González, P. G., Barro, G., et al. 2019, ApJ, 876, 135

ALMA Partnership, Vlahakis, C., Hunter, T. R., et al. 2015, ApJ, 808, L4

Álvarez-Márquez, J., Burgarella, D., Heinis, S., et al. 2016, A&A, 587, A122

Andreani, P. 2010, in Astronomy with Megastructures. Joint Science with the E-ELT and SKA, 13

Arata, S., Yajima, H., Nagamine, K., Abe, M., & Khochfar, S. 2020, MNRAS, 498, 5541

Aravena, M., Decarli, R., Walter, F., et al. 2016, ApJ, 833, 71

Arnouts, S., Cristiani, S., Moscardini, L., et al. 1999, MNRAS, 310, 540

Baars, J. W. M., Hooghoudt, B. G., Mezger, P. G., & de Jonge, M. J. 1987, A&A, 175, 319

Bakx, T. J. L. C., Tamura, Y., Hashimoto, T., et al. 2020, MNRAS, 493, 4294

Barger, A. J., Cowie, L. L., Sanders, D. B., et al. 1998, Nature, 394, 248

Barnabè, M., Dutton, A. A., Marshall, P. J., et al. 2012, MNRAS, 423, 1073

Barro, G., Pérez-González, P. G., Gallego, J., et al. 2011a, ApJS, 193, 13

Barro, G., Pérez-González, P. G., Gallego, J., et al. 2011b, ApJS, 193, 30

Battisti, A. J., da Cunha, E., Grasha, K., et al. 2019, ApJ, 882, 61

Beckwith, S. V. W., Stiavelli, M., Koekemoer, A. M., et al. 2006, AJ, 132, 1729

Behroozi, P. S., Conroy, C., & Wechsler, R. H. 2010, ApJ, 717, 379

Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013, ApJ, 770, 57

Bertin, E. & Arnouts, S. 1996, A&AS, 117, 393

Béthermin, M., Fudamoto, Y., Ginolfi, M., et al. 2020, A&A, 643, A2

Birkinshaw, M. & Lancaster, K. 2005, in Background Microwave Radiation and Intracluster Cosmology, ed. F. Melchiorri & Y. Rephaeli, 127

Blain, A. W., Smail, I., Ivison, R. J., Kneib, J. P., & Frayer, D. T. 2002, Phys. Rep., 369, 111

Bluck, A. F. L., Conselice, C. J., Buitrago, F., et al. 2012, ApJ, 747, 34

Blumenthal, G. R., Faber, S. M., Primack, J. R., & Rees, M. J. 1984, Nature, 311, 517

Bolatto, A. D., Wolfire, M., & Leroy, A. K. 2013, ARA&A, 51, 207

Boogaard, L. A., van der Werf, P., Weiss, A., et al. 2020, ApJ, 902, 109

Boselli, A., Gavazzi, G., Lequeux, J., & Pierini, D. 2002, A&A, 385, 454

Bouché, N., Dekel, A., Genzel, R., et al. 2010, ApJ, 718, 1001

Bouché, N., Finley, H., Schroetter, I., et al. 2016, ApJ, 820, 121

Bouché, N., Murphy, M. T., Kacprzak, G. G., et al. 2013, Science, 341, 50

Bouwens, R. J., Aravena, M., Decarli, R., et al. 2016, ApJ, 833, 72

Bouwens, R. J., Illingworth, G. D., Franx, M., et al. 2009, ApJ, 705, 936

Bouwens, R. J., Illingworth, G. D., Labbe, I., et al. 2011a, Nature, 469, 504

Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015, ApJ, 811, 140

Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2012a, ApJ, 754, 83

Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2011b, ApJ, 737, 90

Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2012b, ApJ, 752, L5

Bouwens, R. J., Smit, R., Schouws, S., et al. 2021, arXiv e-prints, arXiv:2106.13719

Bradač, M., Garcia-Appadoo, D., Huang, K.-H., et al. 2017, ApJ, 836, L2

Brammer, G. B., van Dokkum, P. G., Franx, M., et al. 2012, ApJS, 200, 13

Brisbin, D., Miettinen, O., Aravena, M., et al. 2017, A&A, 608, A15

Bruzual, G. & Charlot, S. 2003, MNRAS, 344, 1000

Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682

Capak, P., Aussel, H., Ajiki, M., et al. 2007, ApJS, 172, 99

Capak, P., Aussel, H., Bundy, K., et al. 2012, SPLASH: Spitzer Large Area Survey with Hyper-Suprime-Cam, Spitzer Proposal

Capak, P., Carilli, C. L., Lee, N., et al. 2008, ApJ, 681, L53

Capak, P. L., Carilli, C., Jones, G., et al. 2015, Nature, 522, 455

Capak, P. L., Riechers, D., Scoville, N. Z., et al. 2011, Nature, 470, 233

Cardamone, C. N., van Dokkum, P. G., Urry, C. M., et al. 2010, ApJS, 189, 270

Carilli, C. L. & Walter, F. 2013, ARA&A, 51, 105

Carniani, S., Gallerani, S., Vallini, L., et al. 2019, MNRAS, 489, 3939

Carniani, S., Maiolino, R., Amorin, R., et al. 2018, MNRAS, 478, 1170

Carniani, S., Marconi, A., Biggs, A., et al. 2013, A&A, 559, A29

Casey, C. M., Berta, S., Béthermin, M., et al. 2012, ApJ, 761, 140

Casey, C. M., Hodge, J., Zavala, J. A., et al. 2018, ApJ, 862, 78

Casey, C. M., Narayanan, D., & Cooray, A. 2014, Phys. Rep., 541, 45

Cassata, P., Morselli, L., Faisst, A., et al. 2020, A&A, 643, A6

Casteels, K. R. V., Conselice, C. J., Bamford, S. P., et al. 2014, MNRAS, 445, 1157

Castellano, M., Sommariva, V., Fontana, A., et al. 2014, A&A, 566, A19

Cattaneo, A., Faber, S. M., Binney, J., et al. 2009, Nature, 460, 213

Chabrier, G. 2003, PASP, 115, 763

Chiaberge, M., Gilli, R., Lotz, J. M., & Norman, C. 2015, ApJ, 806, 147

Civano, F., Marchesi, S., Comastri, A., et al. 2016, ApJ, 819, 62

Coe, D., Zitrin, A., Carrasco, M., et al. 2013, ApJ, 762, 32

Conselice, C. J. 2014, ARA&A, 52, 291

Conselice, C. J. & Arnold, J. 2009, MNRAS, 397, 208

Conselice, C. J., Bershady, M. A., Dickinson, M., & Papovich, C. 2003, AJ, 126, 1183

Conselice, C. J., Mortlock, A., Bluck, A. F. L., Grützbauch, R., & Duncan, K. 2013, MNRAS, 430, 1051

Conselice, C. J., Rajgor, S., & Myers, R. 2008, MNRAS, 386, 909

Contini, T., Garilli, B., Le Fèvre, O., et al. 2012, A&A, 539, A91

Cooke, E. A., Smail, I., Swinbank, A. M., et al. 2018, ApJ, 861, 100

Coppin, K. E. K., Danielson, A. L. R., Geach, J. E., et al. 2012, MNRAS, 427, 520

Cormier, D., Lebouteiller, V., Madden, S. C., et al. 2012, A&A, 548, A20

Cormier, D., Madden, S. C., Lebouteiller, V., et al. 2015, A&A, 578, A53

Cox, P., Krips, M., Neri, R., et al. 2011, ApJ, 740, 63

da Cunha, E., Charlot, S., & Elbaz, D. 2008, MNRAS, 388, 1595 da Cunha, E., Walter, F., Smail, I. R., et al. 2015, ApJ, 806, 110

Daddi, E., Bournaud, F., Walter, F., et al. 2010, ApJ, 713, 686

Daddi, E., Dannerbauer, H., Stern, D., et al. 2009, ApJ, 694, 1517

Davé, R., Finlator, K., & Oppenheimer, B. D. 2012, MNRAS, 421, 98

Davidzon, I., Ilbert, O., Laigle, C., et al. 2017, A&A, 605, A70

Dayal, P. & Ferrara, A. 2018, Phys. Rep., 780, 1

De Breuck, C., Williams, R. J., Swinbank, M., et al. 2014, A&A, 565, A59

De Looze, I., Cormier, D., Lebouteiller, V., et al. 2014, A&A, 568, A62

de Ravel, L., Le Fèvre, O., Tresse, L., et al. 2009, A&A, 498, 379

Decarli, R., Walter, F., Aravena, M., et al. 2016, ApJ, 833, 70

Dekel, A., Birnboim, Y., Engel, G., et al. 2009, Nature, 457, 451

Dekel, A., Zolotov, A., Tweed, D., et al. 2013, MNRAS, 435, 999

Dessauges-Zavadsky, M., Ginolfi, M., Pozzi, F., et al. 2020, A&A, 643, A5

Di Teodoro, E. M. & Fraternali, F. 2015, MNRAS, 451, 3021

Díaz-Santos, T., Armus, L., Charmandaris, V., et al. 2013, ApJ, 774, 68

Dudzevičiūtė, U., Smail, I., Swinbank, A. M., et al. 2020, MNRAS, 494, 3828

Duncan, K., Conselice, C. J., Mundy, C., et al. 2019, ApJ, 876, 110

Dunlop, J. S., McLure, R. J., Biggs, A. D., et al. 2017, MNRAS, 466, 861

Dunlop, J. S., McLure, R. J., Yamada, T., et al. 2004, MNRAS, 350, 769

Elvis, M., Civano, F., Vignali, C., et al. 2009, ApJS, 184, 158

Epinat, B., Tasca, L., Amram, P., et al. 2012, A&A, 539, A92

Erb, D. K., Steidel, C. C., Shapley, A. E., Pettini, M., & Adelberger, K. L. 2004, ApJ, 612, 122

Fabian, A. C., Celotti, A., & Erlund, M. C. 2006, MNRAS, 373, L16

Fahrion, K., Cormier, D., Bigiel, F., et al. 2017, A&A, 599, A9

Faisst, A., Béthermin, M., Capak, P., et al. 2020a, in Panchromatic Modelling with Next Generation Facilities, ed. M. Boquien, E. Lusso, C. Gruppioni, & P. Tissera, Vol. 341, 12–16

Faisst, A. L., Capak, P., Hsieh, B. C., et al. 2016, ApJ, 821, 122

Faisst, A. L., Schaerer, D., Lemaux, B. C., et al. 2020b, ApJS, 247, 61

Ferkinhoff, C., Hailey-Dunsheath, S., Nikola, T., et al. 2010, ApJ, 714, L147

Ferland, G. J., Chatzikos, M., Guzmán, F., et al. 2017, Rev. Mexicana Astron. Astrofis., 53, 385

Ferland, G. J., Porter, R. L., van Hoof, P. A. M., et al. 2013, Rev. Mexicana Astron. Astrofis., 49, 137

Ferrara, A., Vallini, L., Pallottini, A., et al. 2019, MNRAS, 489, 1

Franco, M., Elbaz, D., Béthermin, M., et al. 2018, A&A, 620, A152

Fudamoto, Y., Oesch, P. A., Faisst, A., et al. 2020, A&A, 643, A4

Fudamoto, Y., Oesch, P. A., Schouws, S., et al. 2021, arXiv e-prints, arXiv:2109.10378

Fujimoto, S., Oguri, M., Brammer, G., et al. 2021, ApJ, 911, 99

Fujimoto, S., Ouchi, M., Ono, Y., et al. 2016, ApJS, 222, 1

Fujimoto, S., Silverman, J. D., Bethermin, M., et al. 2020, ApJ, 900, 1

Gallerani, S., Maiolino, R., Juarez, Y., et al. 2010, A&A, 523, A85

Gallerani, S., Neri, R., Maiolino, R., et al. 2012, A&A, 543, A114

Gallerani, S., Pallottini, A., Feruglio, C., et al. 2018, MNRAS, 473, 1909

Genel, S., Vogelsberger, M., Springel, V., et al. 2014, MNRAS, 445, 175

Genzel, R., Tacconi, L. J., Lutz, D., et al. 2015, ApJ, 800, 20

Giavalisco, M., Ferguson, H. C., Koekemoer, A. M., et al. 2004, ApJ, 600, L93

Ginolfi, M., Jones, G. C., Béthermin, M., et al. 2020a, A&A, 643, A7

Ginolfi, M., Jones, G. C., Béthermin, M., et al. 2020b, A&A, 633, A90

Graciá-Carpio, J., Sturm, E., Hailey-Dunsheath, S., et al. 2011, ApJ, 728, L7

Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, ApJS, 197, 35

Gruppioni, C., Béthermin, M., Loiacono, F., et al. 2020, A&A, 643, A8

Gruppioni, C., Pozzi, F., Rodighiero, G., et al. 2013, MNRAS, 432, 23

Guilloteau, S., Delannoy, J., Downes, D., et al. 1992, A&A, 262, 624

Guo, Y., Ferguson, H. C., Giavalisco, M., et al. 2013, ApJS, 207, 24

Harikane, Y., Ouchi, M., Inoue, A. K., et al. 2020, ApJ, 896, 93

Harikane, Y., Ouchi, M., Ono, Y., et al. 2016, ApJ, 821, 123

Harikane, Y., Ouchi, M., Shibuya, T., et al. 2018, ApJ, 859, 84

Hashimoto, T., Inoue, A. K., Mawatari, K., et al. 2019, PASJ, 71, 71

Hasinger, G., Capak, P., Salvato, M., et al. 2018, ApJ, 858, 77

Hasinger, G., Cappelluti, N., Brunner, H., et al. 2007, ApJS, 172, 29

Hatsukade, B., Ohta, K., Seko, A., Yabe, K., & Akiyama, M. 2013, ApJ, 769, L27

Hauser, M. G., Arendt, R. G., Kelsall, T., et al. 1998, ApJ, 508, 25

Henriques, B. M. B., White, S. D. M., Thomas, P. A., et al. 2015, MNRAS, 451, 2663

Herrera-Camus, R., Bolatto, A. D., Wolfire, M. G., et al. 2015, ApJ, 800, 1

Ho, P. T. P., Moran, J. M., & Lo, K. Y. 2004, ApJ, 616, L1

Hodge, J. A., Karim, A., Smail, I., et al. 2013, ApJ, 768, 91

Högbom, J. A. 1974, A&AS, 15, 417

Holland, W. S., Robson, E. I., Gear, W. K., et al. 1999, MNRAS, 303, 659

Hollenbach, D. J. & Tielens, A. G. G. M. 1999, Reviews of Modern Physics, 71, 173

Hopkins, P. F., Bundy, K., Croton, D., et al. 2010, ApJ, 715, 202

Hopkins, P. F., Hernquist, L., Cox, T. J., et al. 2006, ApJS, 163, 1

Hsieh, B.-C., Wang, W.-H., Hsieh, C.-C., et al. 2012, ApJS, 203, 23

Huang, J. S., Rigopoulou, D., Magdis, G., et al. 2014, ApJ, 784, 52

Hughes, D. H., Serjeant, S., Dunlop, J., et al. 1998, Nature, 394, 241

Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, A&A, 457, 841

Ilbert, O., Capak, P., Salvato, M., et al. 2009, ApJ, 690, 1236

Ilbert, O., McCracken, H. J., Le Fèvre, O., et al. 2013, A&A, 556, A55

Iono, D., Yun, M. S., Elvis, M., et al. 2006, ApJ, 645, L97

Jiang, C. Y., Jing, Y. P., & Han, J. 2014, ApJ, 790, 7

Jogee, S., Miller, S. H., Penner, K., et al. 2009, ApJ, 697, 1971

Jones, G. C., Béthermin, M., Fudamoto, Y., et al. 2020, MNRAS, 491, L18

Jones, G. C., Carilli, C. L., Shao, Y., et al. 2017, ApJ, 850, 180

Jones, G. C., Vergani, D., Romano, M., et al. 2021, MNRAS, 507, 3540

Kartaltepe, J. S., Dickinson, M., Alexander, D. M., et al. 2012, ApJ, 757, 23

Kauffmann, G., White, S. D. M., & Guiderdoni, B. 1993, MNRAS, 264, 201

Kaufman, M. J., Wolfire, M. G., Hollenbach, D. J., & Luhman, M. L. 1999, ApJ, 527, 795

Kaviraj, S., Cohen, S., Windhorst, R. A., et al. 2013, MNRAS, 429, L40

Kaviraj, S., Huertas-Company, M., Cohen, S., et al. 2014, MNRAS, 443, 1861

Kelly, B. C. 2007, ApJ, 665, 1489

Kennicutt, Robert C., J. 1998, ARA&A, 36, 189

Kennicutt, R. C. & Evans, N. J. 2012, ARA&A, 50, 531

Kereš, D., Katz, N., Fardal, M., Davé, R., & Weinberg, D. H. 2009, MNRAS, 395, 160

Kereš, D., Katz, N., Weinberg, D. H., & Davé, R. 2005, MNRAS, 363, 2

Kessler, M. F., Steinz, J. A., Anderegg, M. E., et al. 1996, A&A, 500, 493

Khochfar, S. & Burkert, A. 2005, MNRAS, 359, 1379

Khusanova, Y., Bethermin, M., Le Fèvre, O., et al. 2021, A&A, 649, A152

Kitzbichler, M. G. & White, S. D. M. 2008, MNRAS, 391, 1489

Klypin, A. A., Trujillo-Gomez, S., & Primack, J. 2011, ApJ, 740, 102

Knudsen, K. K., Richard, J., Kneib, J.-P., et al. 2016, MNRAS, 462, L6

Koekemoer, A. M., Aussel, H., Calzetti, D., et al. 2007, ApJS, 172, 196

Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, ApJS, 197, 36

Kohandel, M., Pallottini, A., Ferrara, A., et al. 2019, MNRAS, 487, 3007

Koprowski, M. P., Dunlop, J. S., Michałowski, M. J., et al. 2017, MNRAS, 471, 4155

Kroupa, P. & Weidner, C. 2003, ApJ, 598, 1076

Lagache, G., Cousin, M., & Chatzikos, M. 2018, A&A, 609, A130

Laigle, C., McCracken, H. J., Ilbert, O., et al. 2016, ApJS, 224, 24

Lapi, A., Pantoni, L., Zanisi, L., et al. 2018, ApJ, 857, 22

Le Fèvre, O., Abraham, R., Lilly, S. J., et al. 2000, MNRAS, 311, 565

Le Fèvre, O., Béthermin, M., Faisst, A., et al. 2020, A&A, 643, A1

Le Fèvre, O., Tasca, L. A. M., Cassata, P., et al. 2015, A&A, 576, A79

Le Fèvre, O., Vettolani, G., Garilli, B., et al. 2005, A&A, 439, 845

Leja, J., van Dokkum, P., & Franx, M. 2013, ApJ, 766, 33

Lemaux, B. C., Le Fèvre, O., Cucciati, O., et al. 2018, A&A, 615, A77

Lemaux, B. C., Tomczak, A. R., Lubin, L. M., et al. 2017, MNRAS, 472, 419

Li, Y., Hopkins, P. F., Hernquist, L., et al. 2008, ApJ, 678, 41

Lilly, S. J., Le Fevre, O., Hammer, F., & Crampton, D. 1996, ApJ, 460, L1

Lin, L., Cooper, M. C., Jian, H.-Y., et al. 2010, ApJ, 718, 1158

Lin, L., Patton, D. R., Koo, D. C., et al. 2008, ApJ, 681, 232

Liu, D., Gao, Y., Isaak, K., et al. 2015, ApJ, 810, L14

Loiacono, F., Decarli, R., Gruppioni, C., et al. 2021, A&A, 646, A76

López-Sanjuan, C., Cenarro, A. J., Varela, J., et al. 2015, A&A, 576, A53

López-Sanjuan, C., Le Fèvre, O., de Ravel, L., et al. 2011, A&A, 530, A20

López-Sanjuan, C., Le Fèvre, O., Ilbert, O., et al. 2012, A&A, 548, A7

López-Sanjuan, C., Le Fèvre, O., Tasca, L. A. M., et al. 2013, A&A, 553, A78

Lotz, J. M., Davis, M., Faber, S. M., et al. 2008, ApJ, 672, 177

Lotz, J. M., Jonsson, P., Cox, T. J., et al. 2011, ApJ, 742, 103

Lotz, J. M., Jonsson, P., Cox, T. J., & Primack, J. R. 2010, MNRAS, 404, 575

Lotz, J. M., Papovich, C., Faber, S. M., et al. 2013, ApJ, 773, 154

Lotz, J. M., Primack, J., & Madau, P. 2004, AJ, 128, 163

Luhman, M. L., Satyapal, S., Fischer, J., et al. 1998, ApJ, 504, L11

Luhman, M. L., Satyapal, S., Fischer, J., et al. 2003, ApJ, 594, 758

Madau, P. & Dickinson, M. 2014, ARA&A, 52, 415

Madau, P., Ferguson, H. C., Dickinson, M. E., et al. 1996, MNRAS, 283, 1388

Madden, S. C., Poglitsch, A., Geis, N., Stacey, G. J., & Townes, C. H. 1997, ApJ, 483, 200

Magdis, G. E., Daddi, E., Béthermin, M., et al. 2012, ApJ, 760, 6

Maiolino, R., Carniani, S., Fontana, A., et al. 2015, MNRAS, 452, 54

Maiolino, R., Caselli, P., Nagao, T., et al. 2009, A&A, 500, L1

Maiolino, R., Cox, P., Caselli, P., et al. 2005, A&A, 440, L51

Malhotra, S. 2001, in ESA Special Publication, Vol. 460, The Promise of the Herschel Space Observatory, ed. G. L. Pilbratt, J. Cernicharo, A. M. Heras, T. Prusti, & R. Harris, 155

Malhotra, S., Kaufman, M. J., Hollenbach, D., et al. 2001a, ApJ, 561, 766

Malhotra, S., Kaufman, M. J., Hollenbach, D., et al. 2001b, ApJ, 561, 766

Man, A. W. S., Toft, S., Zirm, A. W., Wuyts, S., & van der Wel, A. 2012, ApJ, 744, 85

Man, A. W. S., Zirm, A. W., & Toft, S. 2016, ApJ, 830, 89

Mantha, K. B., McIntosh, D. H., Brennan, R., et al. 2018, MNRAS, 475, 1549

Marchesini, D., van Dokkum, P. G., Förster Schreiber, N. M., et al. 2009, ApJ, 701, 1765

Marchi, F., Pentericci, L., Guaita, L., et al. 2019, A&A, 631, A19

Marrone, D. P., Spilker, J. S., Hayward, C. C., et al. 2018, Nature, 553, 51

Mashian, N., Oesch, P. A., & Loeb, A. 2016, MNRAS, 455, 2101

Matthee, J., Sobral, D., Boogaard, L. A., et al. 2019, ApJ, 881, 124

Matthee, J., Sobral, D., Boone, F., et al. 2017, ApJ, 851, 145

Matthee, J., Sobral, D., Gronke, M., et al. 2020, MNRAS, 492, 1778

McCracken, H. J., Milvang-Jensen, B., Dunlop, J., et al. 2012, A&A, 544, A156

McMullin, J. P., Waters, B., Schiebel, D., Young, W., & Golap, K. 2007, in Astronomical Society of the Pacific Conference Series, Vol. 376, Astronomical Data Analysis Software and Systems XVI, ed. R. A. Shaw, F. Hill, & D. J. Bell, 127

Mortlock, A., Conselice, C. J., Hartley, W. G., et al. 2015, MNRAS, 447, 2

Moster, B. P., Naab, T., & White, S. D. M. 2018, MNRAS, 477, 1822

Mundy, C. J., Conselice, C. J., Duncan, K. J., et al. 2017, MNRAS, 470, 3507

Mundy, C. J., Conselice, C. J., & Ownsworth, J. R. 2015, MNRAS, 450, 3696

Naab, T. & Ostriker, J. P. 2017, ARA&A, 55, 59

Nagao, T., Maiolino, R., De Breuck, C., et al. 2012, A&A, 542, L34

Nayyeri, H., Hemmati, S., Mobasher, B., et al. 2017, ApJS, 228, 7

Neistein, E. & Dekel, A. 2008, MNRAS, 388, 1792

Noeske, K. G., Weiner, B. J., Faber, S. M., et al. 2007, ApJ, 660, L43

Nonino, M., Dickinson, M., Rosati, P., et al. 2009, ApJS, 183, 244

Novak, M., Smolčić, V., Delhaize, J., et al. 2017, A&A, 602, A5

Oberst, T. E., Parshley, S. C., Nikola, T., et al. 2011, ApJ, 739, 100

Oberst, T. E., Parshley, S. C., Stacey, G. J., et al. 2006, ApJ, 652, L125

Oesch, P. A., Bouwens, R. J., Illingworth, G. D., et al. 2015, ApJ, 808, 104

Oesch, P. A., Bouwens, R. J., Illingworth, G. D., Labbé, I., & Stefanon, M. 2018, ApJ, 855, 105

O'Leary, J. A., Moster, B. P., & Krämer, E. 2021, MNRAS, 503, 5646

Olsen, K., Greve, T. R., Narayanan, D., et al. 2017, ApJ, 846, 105

Omont, A. 2007, Reports on Progress in Physics, 70, 1099

Oser, L., Naab, T., Ostriker, J. P., & Johansson, P. H. 2012, ApJ, 744, 63

Ota, K., Walter, F., Ohta, K., et al. 2014, ApJ, 792, 34

Ownsworth, J. R., Conselice, C. J., Mortlock, A., et al. 2014, MNRAS, 445, 2198

Pallottini, A., Ferrara, A., Gallerani, S., et al. 2017, MNRAS, 465, 2540

Papovich, C., Finkelstein, S. L., Ferguson, H. C., Lotz, J. M., & Giavalisco, M. 2011, MNRAS, 412, 1123

Patton, D. R., Carlberg, R. G., Marzke, R. O., et al. 2000, ApJ, 536, 153

Pavesi, R., Riechers, D. A., Capak, P. L., et al. 2016, ApJ, 832, 151

Pavesi, R., Riechers, D. A., Sharon, C. E., et al. 2018, ApJ, 861, 43

Peacock, J. A., Cole, S., Norberg, P., et al. 2001, Nature, 410, 169

Peebles, P. J. E. 1982, ApJ, 263, L1

Pelliccia, D., Lemaux, B. C., Tomczak, A. R., et al. 2019, MNRAS, 482, 3514

Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2002, AJ, 124, 266

Pentericci, L., Carniani, S., Castellano, M., et al. 2016, ApJ, 829, L11

Pérez-González, P. G., Rieke, G. H., Villar, V., et al. 2008, ApJ, 675, 234

Perley, R. A., Chandler, C. J., Butler, B. J., & Wrobel, J. M. 2011, ApJ, 739, L1

Pety, J., Beelen, A., Cox, P., et al. 2004, A&A, 428, L21

Pilbratt, G. L., Riedinger, J. R., Passvogel, T., et al. 2010, A&A, 518, L1

Pillepich, A., Springel, V., Nelson, D., et al. 2018, MNRAS, 473, 4077

Pineda, J. L., Langer, W. D., & Goldsmith, P. F. 2014, A&A, 570, A121

Pozzi, F., Calura, F., Fudamoto, Y., et al. 2021, arXiv e-prints, arXiv:2105.14789

Puget, J. L., Abergel, A., Bernard, J. P., et al. 1996, A&A, 308, L5

Qu, Y., Helly, J. C., Bower, R. G., et al. 2017, MNRAS, 464, 1659

Reddy, N. A., Steidel, C. C., Pettini, M., et al. 2008, ApJS, 175, 48

Reid, B. A., Percival, W. J., Eisenstein, D. J., et al. 2010, MNRAS, 404, 60

Riechers, D. A., Bradford, C. M., Clements, D. L., et al. 2013, Nature, 496, 329

Riechers, D. A., Capak, P. L., Carilli, C. L., et al. 2010, ApJ, 720, L131

Riechers, D. A., Carilli, C. L., Capak, P. L., et al. 2014, ApJ, 796, 84

Riechers, D. A., Leung, T. K. D., Ivison, R. J., et al. 2017, ApJ, 850, 1

Riechers, D. A., Nayyeri, H., Burgarella, D., et al. 2021, ApJ, 907, 62

Riechers, D. A., Pavesi, R., Sharon, C. E., et al. 2019, ApJ, 872, 7

Riechers, D. A., Walter, F., Carilli, C. L., et al. 2011, ApJ, 726, 50

Robson, I., Holland, W. S., & Friberg, P. 2017, Royal Society Open Science, 4, 170754

Rodighiero, G., Daddi, E., Baronchelli, I., et al. 2011, ApJ, 739, L40

Rodriguez-Gomez, V., Genel, S., Vogelsberger, M., et al. 2015, MNRAS, 449, 49

Rodriguez-Gomez, V., Pillepich, A., Sales, L. V., et al. 2016, MNRAS, 458, 2371

Romano, M., Cassata, P., Morselli, L., et al. 2021, A&A, 653, A111

Romano, M., Cassata, P., Morselli, L., et al. 2020, MNRAS, 496, 875

Rowan-Robinson, M., Oliver, S., Wang, L., et al. 2016, MNRAS, 461, 1100

Saintonge, A., Kauffmann, G., Wang, J., et al. 2011, MNRAS, 415, 61

Salpeter, E. E. 1955, ApJ, 121, 161

Sánchez Almeida, J., Elmegreen, B. G., Muñoz-Tuñón, C., & Elmegreen, D. M. 2014, A&A Rev., 22, 71

Sanders, D. B., Salvato, M., Aussel, H., et al. 2007, ApJS, 172, 86

Santini, P., Castellano, M., Fontana, A., et al. 2016, A&A, 596, A75

Santini, P., Fontana, A., Grazian, A., et al. 2012, A&A, 538, A33

Schaerer, D., Ginolfi, M., Béthermin, M., et al. 2020, A&A, 643, A3

Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, MNRAS, 446, 521

Schenker, M. A., Robertson, B. E., Ellis, R. S., et al. 2013, ApJ, 768, 196

Schinnerer, E., Carilli, C. L., Capak, P., et al. 2008, ApJ, 689, L5

Scoville, N., Abraham, R. G., Aussel, H., et al. 2007a, ApJS, 172, 38

Scoville, N., Abraham, R. G., Aussel, H., et al. 2007b, ApJS, 172, 38

Scoville, N., Aussel, H., Brusa, M., et al. 2007c, ApJS, 172, 1

Scoville, N., Aussel, H., Brusa, M., et al. 2007d, ApJS, 172, 1

Scoville, N., Faisst, A., Capak, P., et al. 2015, ApJ, 800, 108

Shah, E. A., Kartaltepe, J. S., Magagnoli, C. T., et al. 2020, ApJ, 904, 107

Silk, J. & Mamon, G. A. 2012, Research in Astronomy and Astrophysics, 12, 917

Silk, J. & Rees, M. J. 1998, A&A, 331, L1

Simons, R. C., Kassin, S. A., Weiner, B. J., et al. 2015, MNRAS, 452, 986

Simpson, J. M., Smail, I., Swinbank, A. M., et al. 2017, ApJ, 839, 58

Simpson, J. M., Swinbank, A. M., Smail, I., et al. 2014, ApJ, 788, 125

Skelton, R. E., Whitaker, K. E., Momcheva, I. G., et al. 2014, ApJS, 214, 24

Smail, I., Ivison, R. J., & Blain, A. W. 1997, ApJ, 490, L5

Smit, R., Bouwens, R. J., Carniani, S., et al. 2018, Nature, 553, 178

Smolčić, V., Novak, M., Bondi, M., et al. 2017, A&A, 602, A1

Snyder, G. F., Lotz, J. M., Rodriguez-Gomez, V., et al. 2017, MNRAS, 468, 207

Solomon, P. M., Downes, D., & Radford, S. J. E. 1992, ApJ, 398, L29

Solomon, P. M. & Vanden Bout, P. A. 2005, ARA&A, 43, 677

Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, ApJS, 214, 15

Spitzer, Lyman, J. 1978, JRASC, 72, 349

Springel, V. 2005, MNRAS, 364, 1105

Springel, V., Wang, J., Vogelsberger, M., et al. 2008, MNRAS, 391, 1685

Springel, V., White, S. D. M., Jenkins, A., et al. 2005, Nature, 435, 629

Stacey, G. J., Geis, N., Genzel, R., et al. 1991, ApJ, 373, 423

Stacey, G. J., Hailey-Dunsheath, S., Ferkinhoff, C., et al. 2010, ApJ, 724, 957

Stach, S. M., Dudzevičiūtė, U., Smail, I., et al. 2019, MNRAS, 487, 4648

Stadel, J., Potter, D., Moore, B., et al. 2009, MNRAS, 398, L21

Stark, D. P., Ellis, R. S., Charlot, S., et al. 2017, MNRAS, 464, 469

Steinhardt, C. L., Speagle, J. S., Capak, P., et al. 2014, ApJ, 791, L25

Swinbank, A. M., Karim, A., Smail, I., et al. 2012, MNRAS, 427, 1066

Swinbank, A. M., Simpson, J. M., Smail, I., et al. 2014, MNRAS, 438, 1267

Tacconi, L. J., Genzel, R., Saintonge, A., et al. 2018, ApJ, 853, 179

Talia, M., Cimatti, A., Giulietti, M., et al. 2021, ApJ, 909, 23

Taniguchi, Y., Kajisawa, M., Kobayashi, M. A. R., et al. 2015, PASJ, 67, 104

Taniguchi, Y., Scoville, N., Murayama, T., et al. 2007, ApJS, 172, 9

Tasca, L. A. M., Le Fèvre, O., Hathi, N. P., et al. 2015, A&A, 581, A54

Tasca, L. A. M., Le Fèvre, O., López-Sanjuan, C., et al. 2014, A&A, 565, A10

Tasca, L. A. M., Le Fèvre, O., Ribeiro, B., et al. 2017, A&A, 600, A110

Thompson, A. R., Clark, B. G., Wade, C. M., & Napier, P. J. 1980, ApJS, 44, 151

Tomczak, A. R., Lemaux, B. C., Lubin, L. M., et al. 2017, MNRAS, 472, 3512

Torrey, P., Wellons, S., Machado, F., et al. 2015, MNRAS, 454, 2770

Vallini, L., Ferrara, A., Pallottini, A., & Gallerani, S. 2017, MNRAS, 467, 1300

Vallini, L., Gallerani, S., Ferrara, A., & Baek, S. 2013, MNRAS, 433, 1567

Vallini, L., Gallerani, S., Ferrara, A., Pallottini, A., & Yue, B. 2015, ApJ, 813, 36

Ventou, E., Contini, T., Bouché, N., et al. 2017, A&A, 608, A9

Ventou, E., Contini, T., Bouché, N., et al. 2019, A&A, 631, A87

Vogelsberger, M., Genel, S., Springel, V., et al. 2014, MNRAS, 444, 1518

Wagg, J., Carilli, C. L., Wilner, D. J., et al. 2010, A&A, 519, L1

Wagg, J., Wiklind, T., Carilli, C. L., et al. 2012, ApJ, 752, L30

Walter, F., Decarli, R., Aravena, M., et al. 2016, ApJ, 833, 67

Walter, F., Decarli, R., Carilli, C., et al. 2012, Nature, 486, 233

Walter, F., Weiß, A., Downes, D., Decarli, R., & Henkel, C. 2011, ApJ, 730, 18

Wang, L., Pearson, W. J., Cowley, W., et al. 2019, A&A, 624, A98

Wang, R., Wagg, J., Carilli, C. L., et al. 2013, ApJ, 773, 44

Wang, T., Elbaz, D., Schreiber, C., et al. 2016, ApJ, 816, 84

Wardlow, J. L., Smail, I., Coppin, K. E. K., et al. 2011, MNRAS, 415, 1479

Weiß, A., De Breuck, C., Marrone, D. P., et al. 2013, ApJ, 767, 88

Weiß, A., Downes, D., Neri, R., et al. 2007, A&A, 467, 955

Werner, M. W., Roellig, T. L., Low, F. J., et al. 2004, ApJS, 154, 1

White, S. D. M. & Frenk, C. S. 1991, ApJ, 379, 52

Whitney, A., Conselice, C. J., Bhatawdekar, R., & Duncan, K. 2019, ApJ, 887, 113

Wild, V., Walcher, C. J., Johansson, P. H., et al. 2009, MNRAS, 395, 144

Williams, C. C., Labbe, I., Spilker, J., et al. 2019, ApJ, 884, 154

Williams, R. E., Baum, S., Bergeron, L. E., et al. 2000, AJ, 120, 2735

Willott, C. J., Carilli, C. L., Wagg, J., & Wang, R. 2015, ApJ, 807, 180

Willott, C. J., Omont, A., & Bergeron, J. 2013, ApJ, 770, 13

Wisnioski, E., Förster Schreiber, N. M., Wuyts, S., et al. 2015, ApJ, 799, 209

Wolfire, M. G., McKee, C. F., Hollenbach, D., & Tielens, A. G. G. M. 2003, ApJ, 587, 278

Wootten, A. & Thompson, A. R. 2009, IEEE Proceedings, 97, 1463

Wu, P.-F., Gal, R. R., Lemaux, B. C., et al. 2014, ApJ, 792, 16

Xu, C. K., Zhao, Y., Scoville, N., et al. 2012, ApJ, 747, 85

Yajima, H., Li, Y., Zhu, Q., & Abel, T. 2012, MNRAS, 424, 884

Yamaguchi, Y., Kohno, K., Hatsukade, B., et al. 2019, ApJ, 878, 73

Yan, L., Sajina, A., Loiacono, F., et al. 2020, ApJ, 905, 147

Zabl, J., Bouché, N. F., Schroetter, I., et al. 2019, MNRAS, 485, 1961

Zahid, H. J., Dima, G. I., Kewley, L. J., Erb, D. K., & Davé, R. 2012, ApJ, 757, 54

Zanella, A., Daddi, E., Magdis, G., et al. 2018, MNRAS, 481, 1976

Zanella, A., Pallottini, A., Ferrara, A., et al. 2021, MNRAS, 500, 118

Zhao, Y., Lu, N., Xu, C. K., et al. 2016, ApJ, 819, 69