Distant, dusty star-forming galaxies

Jianhang Chen

München 2023

Distant, dusty star-forming galaxies

Jianhang Chen

Dissertation
der Fakultät für Physik
der Ludwig-Maximilians-Universität
München

vorgelegt von
Jianhang Chen
aus Chongqing, China

München, den 5. September 2023

Erstgutachter: Prof. Dr. Volker Springel
Zweitgutachter: Prof. Dr. Klaus Dolag
Tag der mündlichen Prüfung: 19. Oktober 2023

Zusammenfassung

Die Entstehung und Entwicklung von Galaxien über die kosmische Zeit hinweg zu verstehen, ist ein wichtiges Ziel der modernen Astronomie. Unsere derzeitigen Beobachtungen des frühen Universums stammen hauptsächlich aus dem optischen und nahen infraroten Wellenlängenbereich, der meist auf staubfreie Systeme ausgerichtet ist. Das Aufkommen von Millimeter- und Submillimeter-Teleskopen (Submm/mm-Teleskopen) in den letzten zwei bis drei Jahrzehnten hat die Erforschung staubiger sternbildender Galaxien (DSFGs) erheblich vorangebracht und uns eine vollständigere Bestandsaufnahme ferner Galaxien ermöglicht. Gleichzeitig hat dies neue Herausforderungen mit sich gebracht: Was sind die inneren und äußeren Eigenschaften von DSFGs? Wie hängen sie mit den Galaxien im lokalen Universum zusammen? Wie hoch ist ihre Anzahl und ihr Beitrag zur Dichte der kosmischen Sternentstehungsrate? Was verraten sie uns über das frühe Universum? In dieser Arbeit werde ich über zwei Durchmusterungen berichten: ALMACAL und ALMARED, die die Populationen von DSFGs am schwachen bzw. hellen Ende der Submm-Zahlen untersuchen. In der Zwischenzeit werde ich auch über ein neues Fenster berichten, durch das galaktische Magnetfelder mit Staubpolarisation im frühen Universum untersucht werden können.

Im Rahmen der ALMACAL-Durchmusterung habe ich die “freien” Kalibrierungsdaten des Atacama Large Millimeter/Submillimeter Array (ALMA) genutzt, um eine Kontinuumsdurchmusterung von DSFGs im Sub-Bogensekundenbereich, bei mehreren Wellenlängen und blind durchzufhren. Ich habe ALMA-Kalibrierungsscans verwendet, um die Sichtlinien zu den ALMA-Kalibratoren und darüber hinaus zu kartieren. ALMACAL hat inzwischen 1.001 Kalibratoren mit einer Gesamtabdeckung von etwa 0,3 deg^2 über den von der chilenischen Atacama-Wüste aus zugänglichen Himmel abgedeckt und dabei mehr als 1.000 Stunden an Integration gesammelt. Die Tiefe, die durch die Kombination mehrerer Besuche in jedem Feld erreicht wird, ermöglicht es ALMACAL, nach schwachen DSFGs in verschiedenen Wellenlängenbändern zu suchen. Basierend auf der aktuellsten ALMACAL-Datenbank berichte ich über die Entdeckung von 186 DSFGs mit flux-Dichten bis hinunter zu $S_{870\,\mu\mathrm{m}} \sim 0{,}2\,\mathrm{mJy}$, vergleichbar mit bestehenden großen ALMA-Durchmusterungen, aber weniger anfällig für kosmische Varianz. Ich berichte auch über die Anzahl der beobachteten Galaxien bei fünf Wellenlängen zwischen $870\,\mu$m und 3 mm, in den ALMA-Bändern 3, 4, 5, 6 und 7, die einen Maßstab für Modelle der Galaxienentstehung und -entwicklung darstellen. Durch die Integration der beobachteten Anzahlen und der Best-fit-Funktionen stelle ich den aufgelösten Anteil des kosmischen Infrarot-Hintergrunds (CIB) und die spektrale Form des CIB dar. Durch die Kombination bestehender Durchmusterungen zeige ich, dass ALMA derzeit etwa die Hälfte des CIB im Submm/mm-

Bereich aufgelöst hat.

In der ALMARED-Durchmusterung habe ich eine ALMA-Snapshot-Durchmusterung der rötesten ~ 3000 *Herschel*-ATLAS-Quellen mit spektroskopischer Nachverfolgung kombiniert, um nach Galaxien-Proto-Haufen zu suchen. Alle Ziele sind so ausgewählt, dass sie $500\,\mu$m Riser mit $S_{500\mu\mathrm{m}} > S_{350\mu\mathrm{m}} > S_{250\mu\mathrm{m}}$. Ich fand heraus, dass etwa 1% der $500\,\mu$m Riser mit ALMA in $\geq$ 3-Quellen aufgelöst wurden. Anschließend wählte ich 12 überdichte Felder mit photometrischen Rotverschiebungen von $z > 3$ aus. Ich führte spektroskopische ALMA-Band-3-Scans durch, um in diesen 12 Feldern nach Emissionslinien zu suchen, und bestätigte vier neue Proto-Cluster-Kerne. Ich werde zeigen, dass die Entdeckungsrate in ALMARED, zusammen mit einem zufälligen überdichten Feld in der ALMACAL-Durchmusterung, uns davor gewarnt hat, dass die Projektionseffekte von DSFGs die Auswahl von Proto-Haufen, die auf der photometrischen Überdichte basieren, stark verunreinigen können. Ich leite ihre Sternentstehungsrate, das gesamte verfügbare molekulare Gas und die Halomasse aus der Multi-Wellenlngen Photometrie und den Spektrallinien ab. Ich stelle fest, dass diese Proto-Cluster-Kerne massiv sind und eine extreme Sternentstehungsaktivität aufweisen. Ich zeige, dass ihre Sternentstehungseffizienz, das verfgbare molekulare Gas und die Halomasse die Idee untersttzen, dass sie die Vorläufer der heutigen hellsten Haufengalaxien (BCGs) sein könnten. Diese Entdeckung deutet darauf hin, dass wir im Massenzentrum dieser Proto-Haufen Zeugen der Bildung von BCGs im Frhstadium sind.

Zusätzlich zu diesen beiden Durchmusterungen bin ich in der glcklichen Lage, zu der Gruppe zu gehören, die die Untersuchung der Staubpolarisation in fernen DSFGs initiiert hat. Zunächst untersuchte ich die Möglichkeit, die Magnetfelder der fernen DSFGs mit Hilfe der Staubpolarisation zu kartieren, basierend auf der Staubpolarisation von nahen Starburst-Galaxien. Anschließend leitete ich eine der ersten vollpolarisierenden Staubbeobachtungen an zwei SMGs: 9io9 und SPT0326−52. Auf der Grundlage dieser bahnbrechenden Beobachtungen habe ich das Vorhandensein von polarisierter Staubemission bei $z = 2,55$ und $z = 5,56$ bestätigt, was darauf hindeutet, dass die Entwicklung von galaktischen Magnetfeldern in frühen Galaxien effizient sein kann. In 9io9 habe ich auch die komplexe Struktur der Magnetfelder im interstellaren Medium aufgeklärt. Ich diskutiere mehrere mögliche Mechanismen zur Erklärung der Entstehung der beobachteten frühen galaktischen Magnetfelder. Diese bahnbrechende Durchmusterung hat die Leistungsfähigkeit von ALMA bei der Beobachtung der polarisierten Staubemission im frhen Universum unter Beweis gestellt und ein neues Fenster zur Erforschung des magnetisierten Universums geöffnet.

Abschließend möchte ich einen Blick auf Bereiche werfen, die mit den DSFGs zusammenhängen, und einige interessante Projekte vorstellen, die in den kommenden Jahren verfolgt werden sollen.

Abstract

Understanding the formation and evolution of galaxies across cosmic time is a major goal of modern astronomy. Our current observations of the early Universe come mainly from the optical and near-infrared rest-frame wavelengths, which are mostly biased towards dust-unobscured systems. The advent of millimetre and submillimetre (submm/mm) telescopes in the last two–three decades has greatly advanced the study of dusty star-forming galaxies (DSFGs), providing us with a more complete census of distant galaxies. Meanwhile, it has posed new challenges: What are the internal and external properties of DSFGs? How do they connect to the galaxies in the local Universe? What is their number density and contribution to the cosmic star-formation rate density? What do they tell us about the early Universe? In this thesis, I will report two surveys: ALMACAL and ALMARED, which survey the populations of DSFGs at the faint and bright end of submm number counts, respectively. Meanwhile, I will also report a new window through which to study galactic magnetic fields with dust polarisation in the early Universe.

In the ALMACAL survey, I have exploited the ‘free’ calibration data from the Atacama Large Millimetre/submillimetre Array (ALMA) to conduct a sub-arcsec, multi-wavelength, and blind continuum survey of DSFGs. I used ALMA calibration scans to map the lines of sight towards and beyond the ALMA calibrators. ALMACAL has now covered 1,001 calibrators, with a total sky coverage of around $0.3\ \mathrm{deg}^2$, distributed across the sky accessible from the Atacama desert in Chile, and has accumulated more than 1,000 h of integration. The depth reached by combining multiple visits to each field makes ALMACAL capable of searching for faint DSFGs in different wavelength bands. Based on the most up-to-date ALMACAL database, I report the detection of 186 DSFGs with flux densities down to $\mathrm{S}_{870\,\mu\mathrm{m}} \sim 0.2\,\mathrm{mJy}$, comparable with existing ALMA large surveys but less susceptible to cosmic variance. I also report the number counts at five wavelengths between 870 μm and 3 mm, in ALMA bands 3, 4, 5, 6 and 7, providing a benchmark for models of galaxy formation and evolution. By integrating the observed number counts and the best-fit functions, I present the resolved fraction of the cosmic infrared background (CIB) and the CIB spectral shape. Combining existing surveys, I show that ALMA has currently resolved about half of the CIB in the submm/mm regime.

In the ALMARED survey, I have combined an ALMA snapshot survey of the reddest ~ 3000 *Herschel*-ATLAS sources with spectroscopic follow-up to search for galaxy proto-clusters. All the targets are selected to be 500 μm risers with $S_{500\mu\mathrm{m}} > S_{350\mu\mathrm{m}} > S_{250\mu\mathrm{m}}$. I found that about 1% of the 500 μm risers have been resolved into ≥ 3 sources with ALMA. I then selected 12 over-dense fields with photometric redshifts, $z > 3$. I conducted ALMA band 3 spectroscopic scans to search for emission lines in these 12 fields and confirmed four new proto-cluster cores.

I will show that the detection fraction within ALMARED, together with a serendipitous over-dense field in the ALMACAL survey, has cautioned us that the projection effects of DSFGs can greatly contaminate the selection of proto-clusters based on photometric over-density. I derive their star-formation rate, total available molecular gas and halo mass based on the multi-wavelength photometry and spectral lines. I find that these proto-cluster cores are massive and sustain extreme star-forming activity. I show that their star-formation efficiency, the available molecular gas, and halo mass all support the idea that they could be the progenitors of today's 'brightest cluster galaxies' (BCGs). This discovery indicates that we are witnessing the formation of early-stage BCG in the mass centre of these proto-clusters.

In addition to these two surveys, I am fortunate to be part of the group that has initiated the study of dust polarisation in distant DSFGs. I first investigated the feasibility of mapping the magnetic fields of the distant DSFGs with dust polarisation, based on the dust polarisation of nearby starburst galaxies. I then led one of the first full-polarisation dust observations towards two SMGs: 9io9 and SPT0326−52. Based on these pioneering observations, I have confirmed the presence of polarised dust emission at $z = 2.55$ and $z = 5.56$, indicating that the development of galactic magnetic fields can be efficient in early galaxies. In 9io9, I also resolved the complex structure of the magnetic fields in the interstellar medium. I discuss several possible mechanisms to explain the formation of the observed early galactic magnetic fields. This pioneering survey has fully demonstrated the power of ALMA to observe the polarised dust emission in the early Universe and opened a new window to explore the magnetised Universe.

Finally, I look ahead to areas related to DSFGs and outline several intriguing projects to be pursued in the coming years.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Dusty star-forming galaxies

As their name implies, a primary feature of dusty, star-forming galaxies (DSFGs) is an abundance of dust. The dust absorbs starlight and re-emits the energy in the far-infrared (far-IR), which makes DSFGs exceptionally bright in the far-IR waveband. The dust grains, with diameters typically in the range 0.3 nm - 0.3 μm, have their peak absorption at ultraviolet (UV) wavelengths, which makes dust emission sensitive to the birth of hot, massive stars, which produce abundant UV photons (Galliano et al., 2018). The far-IR window is not accessible from the ground, which limits studies of DSFGs in the nearby Universe to telescopes in space. Luckily, the expansion of the Universe redshifts the far-IR into millimetre and submillimetre (submm/mm) bands for DSFGs at $z > 1$. Because of this, the study of distant DSFGs has been revolutionised since the arrival of ground-based submm/mm telescopes, such as the James Clerk Maxwell Telescope (Smail et al., 1997; Barger et al., 1998; Hughes et al., 1998), as had been predicted by Blain & Longair (1996).

The realm of DSFGs includes galaxies at different evolutionary stages. They can be starburst galaxies, triggered by galaxy mergers and strong gas inflows (e.g. Tacconi et al., 2008; Hodge et al., 2012). They can be hyperluminous galaxies, where an active galactic nuclei (AGN) is just about to blow out from their dusty surroundings (e.g. Ivison et al., 2013; Oteo et al., 2017b). They can also be normal star-forming galaxies, strongly lensed by a foreground galaxy or galaxy cluster (e.g. Rizzo et al., 2020; Liu et al., 2023). Due to the different selection methods and source properties, DSFGs have been named in a variety of ways, such as hyperluminous galaxies, extremely red galaxies, submillimetre galaxies, optically-faint galaxies, infrared-bright galaxies, dust-obscured galaxies, amongst many others. Throughout this thesis, I will use ‘SMG’ to emphasise the brightest sub-group of DSFGs with $S_{1.2\,\mathrm{mm}} > 1$ mJy (or $S_{870\mu\mathrm{m}} > 3$ mJy), detectable with a single-dish telescope, and I will use ‘DSFG’ to designate general star-forming galaxies that have fainter dust emission, $S_{1.2\,\mathrm{mm}} > 10$ nJy, detectable with the latest submm/mm interferometers.

The submm/mm window not only led to a new era of SMGs and DSFGs, but it also opened an new and effective window to explore the early Universe. When we observe DSFGs at different redshifts in the submm/mm bands, the cosmic dimming is compensated by the increasing brightness of the dust emission as the rest-frame wavelength moves closer to its peak of the dust emission. This feature is named the 'negative K-correction'. Negative K-correction makes it possible to observe DFSGs with similar star-formation rates (SFR) across a wide range of redshifts. Recently, the detection of DSFGs have been pushed out to cosmic dawn (Cooray et al., 2014; Watson et al., 2015; Laporte et al., 2017b; Tamura et al., 2019; Inami et al., 2022). This is critical to our understanding of how the star-formation rate density (SFRD) evolves across cosmic time, as optical surveys are highly biased to UV-bright, unobscured systems. Indeed, multiple works have found that the DFGSs are the main contributor to cosmic SFRD at $z > 4$ (Zavala et al., 2021; Fujimoto et al., 2023).

In the following sections, I will start with the submm/mm surveys in the past few years, which are designed to give a more complete census of DSFGs. I will then summarise the general properties of DSFGs, including their multi-wavelength counterparts, their interstellar medium (ISM), kinematics, and how these properties connect to their subsequent evolution. In the final section, I will shift the focus to the environment of DSFGs. I will provide evidence that the formation of clustered DSFGs is associated with the core region of proto-clusters. I will also summarise all the currently confirmed proto-cluster cores and discuss their connection with more evolved clusters of galaxies.

1.2 Submm/mm sky surveys

The field of DSFGs has been extensively reviewed by multiple works. The various surveys before ALMA are summarized by Casey et al. (2014). The most recent review by Hodge & da Cunha (2020) focuses on the contributions to the research field of the Atacama Large (Sub)Millimetre Array (ALMA). In this section, I will mostly focus on the various surveys on DSFGs achieved in the past three years — a relatively short period, but with immense importance for the next few decades.

These recent surveys have advanced our understanding of DSFGs in two ways. The first one is the systematic follow-up of bright SMGs with high-resolution interferometers, such as ALMA and the Northern Extended Millimeter Array (NOEMA). These follow-up observations have greatly increased the number of SMGs with spectroscopic redshifts and our knowledge of their internal structure. Along with the increased coverage of the multi-wavelength photometry, it has improved their SED modelling and better constrained their gas properties (e.g. Neri et al., 2020; Chen et al., 2022a; Cox et al., 2023). Meanwhile, high-spatial-resolution and high-frequency-resolution observations have made it routine to study their morphologies, kinematics and spatially resolved properties (e.g. Rizzo et al., 2021; Dye et al., 2022). These efforts have been supplemented by the several ALMA Large Programmes, such as INvestigate [C II] at Early Times (ALPINE — Béthermin et al., 2020), an ALMA Large Program to Discover the Most Luminous [CII]+[OIII] Galaxies in the Epoch of Reionisation (REBELS — Bouwens et al., 2022), and a survey of gas, dust and stars on kiloparsec scales in star-forming galaxies at $z \sim 4$–5 (CRISTAL, Herrera-Camus et al. in prep).

The second direction has been afforded by deep surveys towards conventional deep fields.

Again, the achievements have mainly been accomplished by several ALMA Large Programmes. The first was in the UltraDeep Field (UDF): the ALMA SPECtral line Survey in the UDF (ASPECS — Walter et al., 2016; González-López et al., 2020), and the next was towards the lensing fields: ALMA Lensing Cluster Survey (ALCS — Fujimoto et al., 2023). These surveys have constrained the number counts (number density as a function of flux density) down to nJy. Meanwhile, they have also detected spectral lines and dust continuum from faint high-redshift DSFGs, confronting the formation and evolution of dust in the early Universe (e.g. Laporte et al., 2017b; Tamura et al., 2019). These surveys have also demonstrated the power of submm/mm interferometers for observations of early galaxies, which provide a great complement to the advent of the *James Webb Space Telescope (JWST)*.

The number counts of DSFGs have been constrained over the last two decades, starting at mJy levels and gradually extending down to nJy levels. These basic number counts have constituted strong tests of galaxy formation and evolutional models. In the following sections I will dive into DSFGs in three different groups, based on ther restframe far-IR flux density range. In each group, I will summarise the most recent observational constraints and their backend galaxy populations.

1.2.1 Traditional SMGs

The bright number counts were first probed by the ground-based single-dish telescope. The pioneer was the SCUBA camera on the 15-m James Clerk Maxwell Telescope, located high on Mauna Kea in Hawaii, which surveyed about 500 arcmin^2 of sky (Blain, 2002). Its successor, SCUBA2, with much faster scan efficiency has covered $> 5\,\mathrm{deg}^2$ of sky (e.g. Geach et al., 2017). Because of the success of SCUBA, many similar detector systems have been mounted on several other submm/mm telescopes, such as the AzTEC camera on the ASTE telescope (Aretxaga et al., 2011; Scott et al., 2012), the GISMO camera on the IRAM 30-m telescope (Staguhn et al., 2014; Magnelli et al., 2019), the LABOCA camera on the APEX telescope (Wei et al., 2009) and most recently, the NIKA2 camera on the IRAM 30-m telescope (Bing et al., 2023). SPT-3G is another special survey. It used the South Pole Telescope (SPT) to survey SMGs at longer wavelengths (Vieira et al., 2013; Spilker et al., 2016; Reuter et al., 2020). Because of the selection effect, it detected a large fraction of DSFGs with a higher median redshift, $z \sim 4$.

Single-dish telescopes are more efficient at covering a large sky area, but they also suffer from large uncertainties related to source confusion (source blending) and the effects of gravitational lensing. Because of this, significant effort had to be devoted to following up single-dish sources with interferometers. The most notable recent effort has been the ALMA follow-up of SCUBA2 sources (e.g. Stach et al., 2019; Simpson et al., 2020), ALMA follow-up of SPT-3G sources (Spilker et al., 2016; Reuter et al., 2020), and ALMA+NOEMA follow-up of bright *Herschel* sources (Urquhart et al., 2022; Cox et al., 2023). With these follow-up efforts, many SMGs were resolved into double or multiple sources, or have been revealed to be strongly lensed by foreground sources, indicating that single-dish surveys have over-estimated the bright-end submm/mm number counts. Currently, the useful synergy between interferometers and single-dish telescopes has tightly constrained the bright end of the number counts — see Fig. 1.1. For those single-dish surveys without high-resolution follow-up, a theoretical model is usually used to apply these corrections (e.g. Bing et al., 2023).

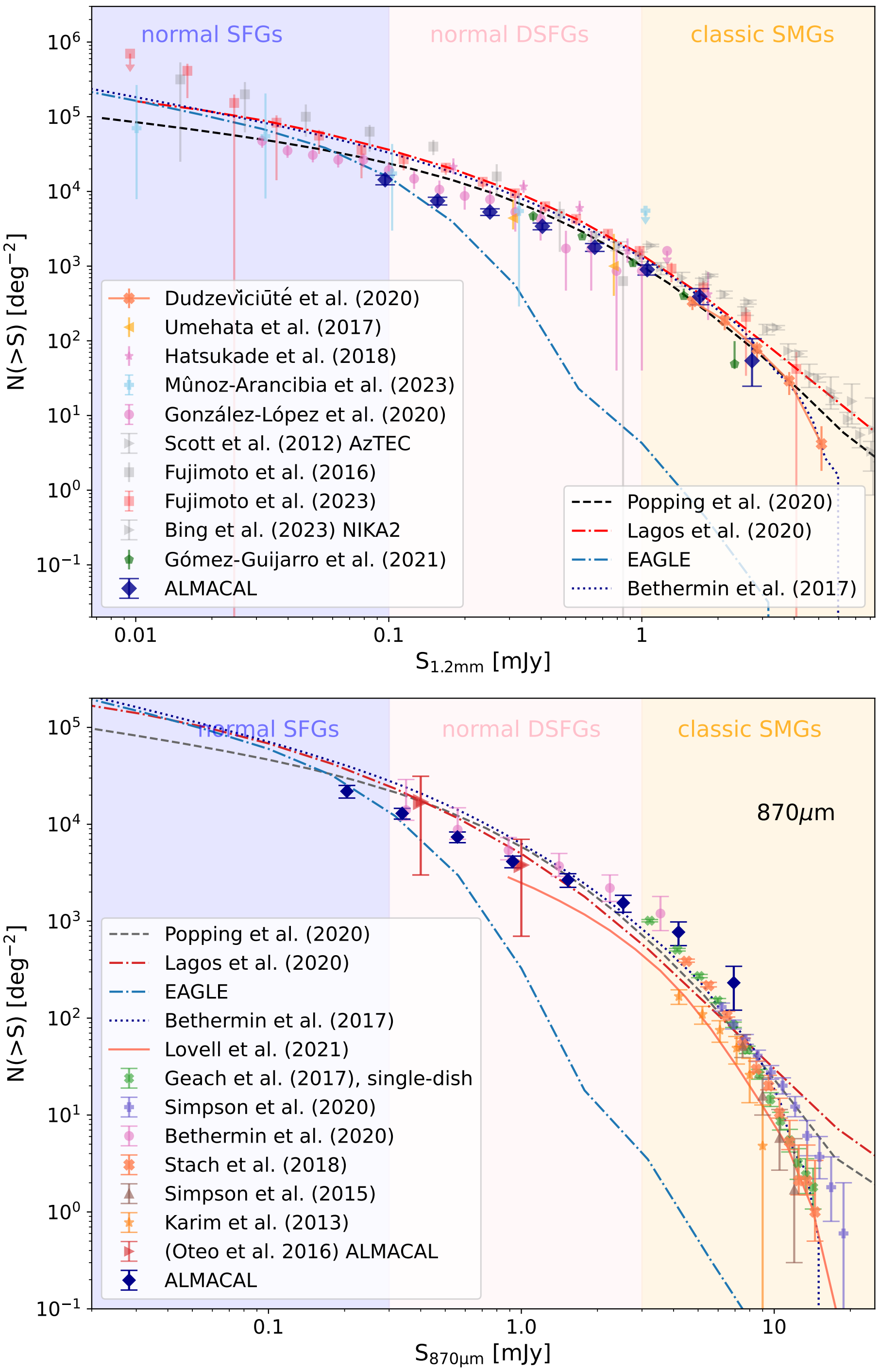

Fig. 1.1: The number counts at 870 μm and 1.2 mm. These two wavelengths have received most of the observations due to the best atmospheric transmission. Currently, ALMA follow-up is mainly focused on the traditional single-dish SMGs initially found at 850 or 870 μm; while most of the ALMA deep surveys towards conventional fields and lensed fields are mainly conducted at 1.2 mm. In general, different surveys report consistent number counts, but the scatter is generally larger than the reported error in each of these surveys, indicating that some or all of the surveys may have under-estimated their uncertainties. The largest discrepancy is found at the faintest end of S$_{1.2\,\mathrm{mm}}$, calling for another independent survey to resolve the current divergence.

Numerical simulations have taken considerable steps forward in reproducing the number counts of the SMGs in recent years. Flexible approaches such as semi-empirical and semi-analytical models have successfully explained the number counts of SMGs from various surveys (e.g. Popping et al., 2020; Lagos et al., 2020). Due to computational limitations, hydrodynamics with full radiative transfer has currently only been applied to massive starburst galaxies in the simulations (McAlpine et al., 2019; Lovell et al., 2021) and they are generally consistent with the observations at the bright end of the counts. Different models provide different number counts at the fainter end, which makes deep submm/mm surveys more sensitive to different models.

Fully explaining the intrinsically bright SMGs is an ongoing effort. Compared with the local analogues, possible explanations are major mergers of galaxies or inflows of cold gas, or some combination of the two. High spatial resolution follow-up usually resolves complex morphologies, favouring that they are mostly major mergers, as is the case for dusty starbursts in the local Universe (e.g. Oteo et al., 2016b; Perry et al., 2023). Meanwhile, cold-mode accretion (CMA) has also been proposed to explain their formation in the early Universe (Dekel et al., 2009). This scenario has been supported to some extent by deep observations of GN20 , but it cannot be well distinguished with merger of galaxies (Carilli et al., 2010; Hodge et al., 2013b). AGN could also be playing a non-negligible role in these systems, but their signatures are often elusive (Ivison et al., 2013, 2019, , see also 1.3.5).

In the rare, bright SMGs ($S_{870\mu\mathrm{m}} > 10\,\mathrm{mJy}$), three types of system could contribute. They can be intrinsically bright, and thus qualify as hyper-luminous IR galaxies (HyLIRG, e.g. Ivison et al., 2013; Oteo et al., 2017a); they can be strongly lensed systems, whose flux density is amplified by foreground galaxies or clusters (e.g. Vieira et al., 2013; Urquhart et al., 2022); they can also be barely resolved clusters of DSFGs (e.g. Walter et al., 2012; Oteo et al., 2018; Miller et al., 2018). All three galaxy populations need to be tackled in galaxy evolution models to fully reproduce the number counts of SMGs. Meanwhile, due to their rareness and exceptional brightness, they are also valuable scientific targets. For instance, hyperluminous DSFGs are thought to be the precursor of quasars, making them ideal targets to understand how the emergence of AGN activity could quench the host galaxy (Ivison et al., 1998, 2019); Strongly lensed DSFGs allow us to probe intrinsically fainter DSFGs and resolve their ISM (Dye et al., 2022; Rizzo et al., 2020). More examples can be found in section 1.4. The clustering of DSFGs will be the topic of section 1.5.

1.2.2 Normal DSFGs

Due to the confusion limit, single-dish telescopes cannot reach DSFGs with sub-mJy flux densities without resorting to the use of gravitational lensing (Smail et al., 1997) or high-frequency submm bands (Wang et al., 2017). The emergence of large interferometers such as ALMA and NOEMA has significantly improved the sensitivity, which has extended the number counts of DSFGs down to the level of 0.1 mJy. For instance, in 10 s, ALMA can reach deeper, and with better spatial resolution, than SCUBA could reach in 50 h of perfect weather, albeit over a smaller field of view. Thanks to this, several large surveys have covered several arcmin^2, such as the ALMA survey towards the GOODS-S (Hatsukade et al., 2018; Franco et al., 2018; Gómez-Guijarro et al., 2019) and the ALMA survey in the SSA22 field, which contains a protoclusterfirst revealed by an over-density of Lyman-break galaxies (Umehata et al., 2017). Meanwhile, ALMA has accumulated a significant volume of archival data towards various fields, including some of the deep fields, which has initialised several archive projects to search for DSFGs (e.g Fujimoto et al., 2016; Liu et al., 2019). ALMACAL is one such archive project, which I will introduce in Chapter 2.

The galaxies in the mid-range of flux density are thought to be the main source of the cosmic background in the submm window. They include the bulk of DSFGs with $100 < \mathrm{SFR} < 1000\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$ and $1 < z < 6$. They generally show an enhanced star-formation rate compared to the main-sequence galaxies but are not as extreme as the brighter and rarer SMGs. Current deep sky surveys are starting to find more of these galaxies, but deriving their properties and securing their redshift will be a long journey.

1.2.3 Normal star-forming galaxies

Pushing down the detection of DSFGs below 0.1 mJy over a sky area of tens of arcmin^2 remains difficult for modern interferometers, but it is a regime where we start detect numerous normal star-forming galaxies (SFG) at $z \sim 1$. ASPECS is currently the deepest ALMA survey available. It has pushed down the number counts to 0.03 mJy, and down to 0.01 mJy based on a $P(D)$ analysis. Blind surveys towards strongly lensed cluster fields have reached similar depths (Muñoz Arancibia et al., 2023; Fujimoto et al., 2023). Popping et al. (2020) used a semi-empirical model to explain the number counts of the ASPECS survey at 1.2 mm. Their best model predicts that the observed number counts of DSFGs ($S_{1.2\,\mathrm{mm}} < 0.1$ mJy) is dominated by star-forming galaxies with $M_{\mathrm{dust}} < 10^9\,\mathrm{M}_\odot$ and $\mathrm{SFR} < 100\,\mathrm{M}_\odot\mathrm{yr}^{-1}$ at $1 < z < 2$. In the semi-anaytical SHARK model, Lagos et al. (2020) also predicted that the galaxy population with $S_{1.2\,\mathrm{mm}} < 0.1$ mJy are started to be dominated by the main-sequency galaxies with stellar masses $10^9 - 10^{10}\mathrm{M}_\odot$. However, current results still suffer from uncertainties. Significantly different number counts at the faint end (< 0.1 mJy) have been found using different approaches. The measurements from lensed fields predict two orders of magnitude higher number counts than the results from conventional deep fields (González-López et al., 2020; Muñoz Arancibia et al., 2023; Fujimoto et al., 2023). The difference could be due to the different analysing methodologies (Fujimoto et al., 2023), but it can also be rooted in cosmic variance, due to the limited sky coverage in these deep fields.

1.2.4 DSFGs at different submm/mm wavelengths

Historically, the number counts were mostly conducted around 870 μm and 1.2 mm. At these two wavelengths the generally brighter dust emission and the good atmospheric transmission give the best compromise to search for DSFGs. However, the negative K-correction is different at different wavelengths, which makes different wavelengths sensitive to galaxies at different redshifts (Lagache et al., 2005) and/or with different dust temperatures (see later). For instance, models with different intrinsic luminosity functions at $z > 4$ predict considerably different number counts at longer wavelengths (Casey et al., 2018b). This makes multi-wavelength surveys for DSFGs another testbed for galaxy evolutional models.

Multi-band follow-up is essential to fully understand the cosmic infrared background (CIB). The CIB – covering mid-infrared to submm/mm wavelengths – comprises an important component of the energy emitted by galaxies over the history of the Universe (Hauser et al., 1998; Fixsen et al., 1998). Along with the optical background, it represents the energy produced by the formation and evolution of galaxies, and all related processes, across all cosmic time. Since the first measurements of the CIB, a primary goal in astrophysics has been to identify the sources responsible for it, and thus to understand the lifetime energy budget of the Universe (Cooray & Sheth, 2002; Lagache et al., 2005; Dole et al., 2006; Planck Collaboration et al., 2014). However, due to the lack of surveys that simultaneously cover all the submm/mm windows, we still lack the resolved SED of CIB at submm/mm wavelengths.

Multi-band submm/mm observations of DSFGs are also crucial to constrain their dust properties. A typical simplified modified black body model contains four key parameters, the absolute opacity $\kappa(\lambda_0)$, the dust mass $M_{\rm dust}$, dust temperature $T_{\rm dust}$ and the dust emissivity β (Galliano et al., 2018). They are critical to infer the available cold gas, the SFR, and the compositions of the dust. The most critical parameter, $T_{\rm dust}$, has been debated to vary systematically with redshift (Magnelli et al., 2014; Drew & Casey, 2022; Sommovigo et al., 2022) and sensitive to the CMB temperature at high redshifts (da Cunha et al., 2013; Zhang et al., 2016; Jin et al., 2019). Current studies still suffer from biased sample selection and over-simplified models due to the limited availability of data (Liang et al., 2019; Bakx et al., 2021).

Unfortunately, even conducting a cosmological survey at a single wavelength is already costly with current interferometers, which makes it increasingly expensive to complete the existing surveys with similar sensitivity at different wavelengths. In Chapter 2, I will introduce a new approach to conducting multi-band surveys for DSFGs based on free ALMA calibration observations.

1.2.5 Luminosity function

With increasing efforts to get the redshifts of a larger sample of DSFGs, we are also approaching the era to convert the number counts into a luminosity function. Number counts represent a powerful tool with which to test galaxy evolution models in the projected sky, but a more rigorous test requiresredshifts, which allows us to convert the number counts into luminosity functions at different redshifts. Great examples are available among optical surveys, where complete redshift surveys make luminosity functions a powerful tool for observational cosmology. Therefore, defining the luminosity function of DSFGs will be the next focus of submm/mm sky surveys.

Early results from the deepest submm/mm surveys already shed light on the importance of the luminosity function of DSFGs. Walter et al. (2020) compiled the redshift evolution of molecular gas based on the CO line luminosity function (Decarli et al., 2019, 2020) and dust continuum luminosity function (Magnelli et al., 2020) from the ASPEC survey, which demonstrates the importance of molecular gas modulating the cosmic SFRD. Fujimoto et al. (2023) revisited the cosmic SFRD with the luminosity function of DSFGs from the ALMA Lensing Cluster Survey. They provided compelling evidence that DSFGs are important contributors to the CIB at $z > 2$.

1.3 Multiwavelength counterparts

Multi-wavelength radiation from galaxies encodes the energy budget from various physical processes. The spectral energy distribution (SED) modelling is an essential way to decompose the energy output from stars, dust and AGN, which has been used widely to understand origin and evolution of SMGs. In the following, I will summarise the observational results and the main features of DSFGs learned at different wavelengths.

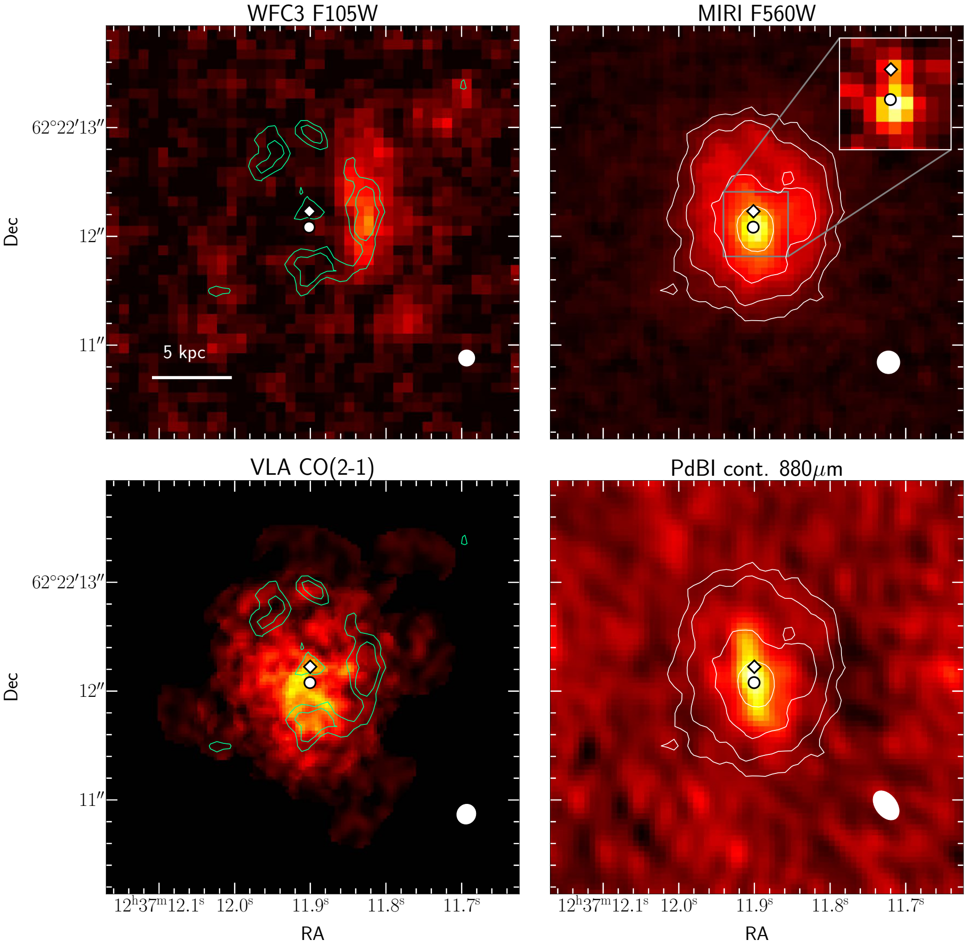

Fig. 1.2: A multi-wavelength view of GN20 (Daddi et al., 2009; Hodge et al., 2012, 2013b; Colina et al., 2023) from *HST*'s WFC3 F105W (top left), *JWST*'s MIRI F560W (top right), CO(2–1) (bottom left) and dust continuum from the IRAM PdBI (bottom right). Green contours show the near-IR light distribution (start at 3σ) and the white contours show the isophotes of MIRI F560W (start at 5σ). GN20 is one of the most extensively studied un-lensed SMGs, the rich multi-wavelength images demonstrate the different structures revealed by different wavelengths, highlighting the importance of multi-wavelength follow-up of DSFGs.

1.3.1 Radio

The radio emission from DSFGs, at frequencies below 30 GHz, consists of free-free and synchrotron radiation (see the review by Condon, 1992). The free-free emission comes mainly from H II regions, where thermalised electrons are decelerated by the electric field from ions. The synchrotron comes from relativistic electrons travelling through magnetic fields. Depending on the kinematic temperature, electron density and magnetic field, their contributions to the total radio emission vary at different wavelength in different systems.

A tight FIR/radio correlation (FRC) has been confirmed across a wide range of IR luminosity, including in DSFGs and the extreme sub-group SMGs (e.g. Yun et al., 2001; Ivison et al., 2010a). In actively star-forming galaxies, the relativistic electrons (and positrons) originating from supernova explosions dominate the synchrotron radiation. The supernova rate is directly linked to the same population of massive stars powering the FIR dust emission (Voelk, 1989). This theory takes both FIR and radio emission as star-formation calorimeters, which provides a natural explanation of the FRC. However, a deviation has been found in low-IR luminosity systems and there has been a long dispute about the redshift evolution of q_{FIR} (Ivison et al., 2010b; Magnelli et al., 2015; Algera et al., 2020; Molnár et al., 2021), which has motivated several more complicated models (Thompson et al., 2006; Lacki et al., 2010; Delvecchio et al., 2021). The parameter q_{FIR} has been found to correlate with stellar mass, where a lower value of q_{FIR} has been found in more massive systems (Delvecchio et al., 2021). Understanding radio emission over a larger sample of star-forming galaxies is one of the main scientific goals of the Square Kilometre Array (SKA). A more detailed study on FRC will require the upcoming more sensitive radio and far-IR sky surveys.

The first radio observations towards SMGs, especially with radio interferometers such as the Very Large Array (VLA), greatly improved the positional accuracy and enabled the first morphological analysis (Ivison et al., 1998, 2000; Smail et al., 2000; Ivison et al., 2002, 2007). The high-accuracy positions offered by interferometers makes it possible to conduct stacking analyses at other wavelengths over a large sample of DSFGs (e.g Ivison et al., 2010b; Lindner et al., 2012). These high-resolution radio observations typically indicate a relatively extended star-formation region with a fraction showing multiple components or disturbed morphologies, which points strongly to mergers as the origin of their boosted SFR (Ivison et al., 2007; Biggs & Ivison, 2008).

Besides supernovae, AGN are also strong synchrotron emitters. The synchrotron from a powerful AGN jet can substantially enhance the total radio luminosity, which makes q_{FIR} larger than for purely star-forming galaxies. This radio-loud signature, with $q_{\mathrm{FIR}} > 1.8$, has become a way to identify AGNs in DSFGs. However, only ~10 per cent of AGNs show radio-loud signatures (Best & Heckman, 2012; Heckman & Best, 2014), therefore radio follow-up cannot give a complete census of the AGN populations in SMG host galaxies. Nevertheless, radio-loud signatures have been more frequently reported in the most luminous DSFGs (Venemans et al., 2007; Ivison et al., 2012; Chapman et al., 2023) indicating that the brightest sub-sample of DSFGs could harbour AGNs.

Very Long Baseline Interferometry (VLBI) offers another promising way to identify AGNs. Thanks to the extraordinary angular resolution of VLBI, and the spatial filtering, only compact

AGN emission pops out in VLBI images. Such experiments were first carried out by Biggs & Ivison (2008); Biggs et al. (2010) and Momjian et al. (2010) at 1.4 GHz to search for compact synchrotron radiation from AGN. However, the success rate is relatively low amongst radio-detected SMGs (Biggs & Ivison, 2008; Biggs et al., 2010; Chi et al., 2013; Chen et al., 2020).

1.3.2 Near- and mid-infrared

Near- and mid-infrared observations towards DSFGs have taken monumental steps forward with the launch of the *Spitzer Space Telescope* in late 2003 and then, more recently, the *James Webb Space Telescope*. Following on from the pioneering *Spitzer* spectroscopic survey of SMGs conducted by Pope et al. (2008), Riechers et al. (2014) studied the mid-IR emission from the SMG, GN20, reporting the first detection of polycyclic aromatic hydrocarbon (PAH) at 6.3 μm. Meanwhile, the embedded continuum emission in the mid-IR support the presence of dusty torus around an obscured AGN (Riechers et al., 2014).

JWST offers both higher angular resolution and sensitivity comparing with *Spitzer* and *HST*. The availability of the mid-IR instrument, MIRI, makes it a more robust probe of the stellar morphologies of DSFGs. By cross-matching DSFGs with early public *JWST* surveys, several early attempts have revealed the complex morphologies of DSFGs. Cheng et al. (2022) reported the first statistical sample of DSFGs ($n = 16$) observed by *JWST* and ALMA. Most of their targets were normal DSFGs (SFR $\sim$ 10–300 $\mathrm{M_{\odot}yr^{-1}}$), fainter than traditional SMGs. Their morphologies are mostly disk-like, favouring secular evolution instead of galaxy mergers in developing their stellar disks. Chen et al. (2022b) reported another seven targets, mainly brighter SMGs, selected from the SCUBA-2 Cosmological Survey (S2CLS, Geach et al., 2017). Those targets also showed the presence of stellar disks; some of them have even developed stellar bars. Compared to the DSFGs from Cheng et al. (2022), SMGs are more likely to be morphologically disturbed and have a nearby companion. With the growing number of DSFGs observed by *JWST* (e.g. Gillman et al., 2023), DSFGs and SMGs seem to be capable of developing distinguishable stellar structures during the first 3 Gyr of the Universe, and may have diverse formation pathways.

Near- and mid-IR observations of DSFGs above $z > 4$ remain rare. (Álvarez-Márquez et al., 2023) conducted deep *JWST* observations of the most distant known SMG, SPT0311−58, a merging system. The NIR image revealed clumpy stellar structures in both galaxies. The line ratio between Paα and Hα indicates the presence of heavy dust extinction, $A_{\mathrm{V}} \sim 4$. Colina et al. (2023) and Peng et al. (2023) showed observations of two more $z \sim 4$ SMGs, GN20 and SPT0418−47, both also found to be merging systems. In SPT0418−47, the companion galaxy is only detected in the *JWST* observations due to its lower SFR, highlighting the important synergy between *JWST* and ALMA.

1.3.3 Optical

Rest-frame UV radiation, 1250–2500Å, is widely used to study the stellar morphologies of galaxies. UV radiation is dominated by young stars, which makes it one of the most widely used star-formation tracers, despite the known and pernicious effects of dust. UV radiation is not accessible by ground optical telescopes for the nearby Universe ($z < 0.5$), but it is easy to get for distant galaxies (Steidel et al., 1996). Unfortunately, since the wavelength is close to the size of the dust grains, it suffers greatly from dust attenuation. Available rest-frame UV observations towards DSFGs have only been achieved for a handful of targets in the conventional deep fields. Hodge et al. (2015) showed the morphologies of brightest SMGs – GN20 – in the CANDELS fields. Its rest-frame UV emission systematically offset from the dust and molecular line emissions (see also Fig. 1.2).

The longer rest-frame wavelength at $\lambda > 4000\,\text{Å}$ offers better means to probe the stellar morphologies of the DSFGs. The longer wavelength suffers less from dust extinction, but it is only accessible by the near-IR telescopes for galaxies at $z > 1$. Several large surveys with HST in the near-IR have yielded generally complex morphologies in the brightest sub-sample of DSFGs, indicating the essential role of galaxy interaction and mergers in triggering the starburst activities (Chen et al., 2015). It should be noted that the orientation effects can further complicate the interpretation of the morphologies, due to the higher extinction with a lower inclination angle (Lovell et al., 2021).

Optical spectral lines are also valuable in diagnosing the nature of DSFGs. The most intriguing line is the Lyman-α. Due to its resonant nature, it has been proven to be an effective way to confirm the redshifts of the SMGs (Chapman et al., 2003). Soon later, it enabled the first statistical sample of SMGs with accurate redshifts (Chapman et al., 2005), which initiated massive follow-up observations of these targets (Alexander et al., 2005; Greve et al., 2005; Tacconi et al., 2006). Meanwhile, with the detection of multiple optical lines, their line ratios offer valuable information about their ISM. For instance, one can derive the dust extinction of the ISM based on the line ratio of Hα and Hβ (Shapley et al., 2023; Sandles et al., 2023); gas excitation mechanism from the ratio of [N II]/Hα and [O III]/Hβ (e.g. Trump et al., 2023); the gas phase metallicity if [N II] or [O III] could be detected along with the Balmer lines (e.g. Taylor et al., 2022; Nakajima et al., 2023). The decomposition of the line shapes of the Balmer lines provides another way to identify and measure the mass of the AGN. An intriguing example has recently been demonstrated by Ivison et al. (2019), in which a hyperluminous SMG showed an exposed broad line region traced by the broad Hα emission. Moreover, the spatially resolved spectral line can be used to study the morphologies and the kinematics of the ionised gas (e.g. Swinbank et al., 2006).

In general, the optical window is not the preferred window to study the DSFGs, but the existing results caution us to interpret the optical observations of distant galaxies. The galaxies in the early Universe are supposed to be more turbulent and more vulnerable to frequent galaxy interactions. Therefore, the galactic properties in these systems could be quite heterogeneous. Currently, we are mainly observing the early Universe at $z > 6$ with the rest-frame optical wavelengths, which is biased to the unobscured regions in these galaxies. Recent observations from *JWST* do support that most $z > 6$ galaxies are young and metal-poor (e.g. Schaerer et al., 2022; Curti et al., 2023; Robertson et al., 2023), but there are also cases of evolved galaxies (Tamura et al., 2019; Bunker et al., 2023), indicating that part of these early galaxies could also suffer from dust extinction.

1.3.4 X-rays

There are two dominant sources of X-ray emission in DSFGs. In the star-forming regions, the mass loss through the stellar winds of massive stars feeds the nearby compact companion, triggering X-ray emission during the accretion. Another source is the AGN, where the innermost region of the accretion disk and the corona above the SMBH could both emit X-rays. X-rays coming from corona are stronger and harder than those from star-forming regions, making X-ray follow-up of SMGs the cleanest way to search for AGN signatures. The seminal X-ray studies of DSFGs were carried out soon after the first detection of SMGs in the lensed fields (Fabian et al., 2000). Then, the sample sizes were gradually extended to > 100 (Barger et al., 2001; Alexander et al., 2002, 2005; Laird et al., 2010; Wang et al., 2013; Ueda et al., 2018; Uematsu et al., 2023).

Despite of the numerous efforts, it remains difficult to quantify the importance of AGN in DSFGs. The reported AGN occupation rate – the number fraction of the DSFGs with a confirmed AGN – ranges from 15-75 % in different studies (e.g. Alexander et al., 2005; Wang et al., 2013; Ueda et al., 2018; Uematsu et al., 2023). Current constraints still suffer from different pre-sample selection and observation depths both in submm/mm and X-ray bands. Stacking analysis can potentially reveal the AGN signature from a much large sample of DSFGs, but the current analyses still give ambiguous results. Wang et al. (2013) collects the largest number of SMGs with X-ray observation and ~ 1 arcsec positional accuracy. Their stacked $L_{0.5-8\,\mathrm{keV}}$ is larger than the expected contribution from the typical SFR in DSFGs, supporting the presence of possibly more obscured AGNs in part of the fainter DSFGs. On the other hand, earlier works only find detection in the stacked soft band (0.5-2 keV, Georgantopoulos et al., 2011; Lindner et al., 2012), which disfavour the presence of more heavily obscured AGNs in the X-ray non-detection targets. It should be noted that even without hard X-ray detections, the presence of heavily obscured AGNs with $\log N_{\mathrm{H}} > 23.5\mathrm{cm}^{-2}$ cannot be ruled out (e.g. Ricci et al., 2023). However, with higher intrisic luminosity, such as the most extreme hyperluminous starbursts, the incidence of AGNs is thought to be higher when the AGN is about to blow out its dusty surroundings (e.g. Ivison et al., 2019). Moreover, a slightly higher AGN fraction has been found in the overdense environments (e.g. Umehata et al., 2019; Vito et al., 2020), but those results need to be confirmed by larger future samples.

1.3.5 Caveats

The increasing availability of multiwavelength data has greatly improved our understanding of the origin and the possible descendants of DSFGs over the past two decades. However, our current SED modellings are still limited by the photometric coverage with matched spatial resolution and the inadequate mid-IR photometries.

One uncertainty comes from the signature of AGNs. HyLIRG were widely thought to harbour the SMBH from the general merger-starburst-AGN paradigm (Hopkins et al., 2006). Ma et al. (2016) reported the X-ray observation and SED modelling of the most distant HyLIRG found so far, SPT0346−58, where no strong AGN evidence has been found. (Ivison et al., 2013) reported a similar result in the HyLIRG HATLAS J084933, where a galaxy with a comprehensive multi-frequency dataset initially showed no hint of any AGN contribution to its bolometric luminosity.

However, a massive and immensely powerful AGN was eventually revealed by X-ray data and the presence of a >10,000-km s^{-1} Balmer Hα line (Ivison et al., 2019). These examples caution the conclusions without panoramic wavelength coverage.

The next uncertainty is the basic assumption about the initial mass function (IMF). Variable IMFs have already been widely reported in different environments (e.g. Cappellari et al., 2012; Li et al., 2023; Yung et al., 2023), where a top-heavy IMF is generally preferred in environments with lower metallicity and higher gas densities (e.g. Dabringhausen et al., 2009; Jeřábková et al., 2018; Zhang et al., 2018b). Dye et al. (2021) showed that different assumptions of the IMF in the SMG can lead to about 3× differences in the total SFR, stellar mass and dust mass. Unfortunately, the current SED modelling of SMGs cannot effectively constrain the IMF. Caution is needed when interpreting the results.

1.4 Interstellar medium

With the presence of wide spectral bandwidth, submm/mm surveys have detected many FIR spectral lines. The spectral lines are dominated by the rotational transition of the molecules and the ionization ions of the Carbon, Nitrogen, and Oxygen. Bright lines, such as CO, [C I], [C II]158μm, and [O III]88μm, have been widely used to probe the multiphase gas and the kinematics of early galaxies, including DSFGs.

1.4.1 Carbon monoxide

CO lines are the most commonly observed molecular line in the far-IR. It is the second most abundant molecule apart from H_2 and usually the strongest lines in the observed millimetre wavelength. Spectral line surveys thus preferentially search for CO lines to get the redshift of DSFGs (e.g. Urquhart et al., 2022; Chen et al., 2022a; Cox et al., 2023). Until now, the low-J CO transitions are still widely demanded to study the general molecular gas reservoirs in DSFGs (e.g. Frias Castillo et al., 2023). With the advent of low-frequency receivers mounted on ALMA, low-J CO will become another favourite to survey the distant DSFGs. However, the conversion from CO luminosity to the molecular gas is still uncertain. The cross-validation from multiple tracers and the dynamical signature of the CO lines all support a generally lower conversion factor in the DSFGs, especially in the most active SMGs (Ivison et al., 2011; Dunne et al., 2022).

CO transitions have been widely used to study the morphologies, kinematics and ionization state of the cold ISM. The resolved studies generally reported more extended line emission than the dust continuum and the kinematical signature consistent with a rotating disk (e.g. Hodge et al., 2015; Calistro Rivera et al., 2018). With the increasing collection of multiple CO transitions in DSFGs, it is now possible to construct the CO spectral line energy distribution (SLED) (Yang et al., 2017; Stanley et al., 2023). The CO SLED generally show a large scatter, with most of them peaked around J=6-8. However, some CO SLED of SMGs similar to that of the AGN and quasars (Greve et al., 2014; Ma et al., 2016; Kirkpatrick et al., 2019), which can be interpreted as evidence that AGN may be present in SMGs, but they are not quenching the star-formation very effectively.

1.4.2 Ionized carbon

Far-infrared fine structures line [C II]158μm has been widely used to study global properties and the ISM of distant galaxies. [C II] is the main cooling channel of metal-enriched ISM. It is one of the brightest far-IR lines, observable at $z > 2$ with ALMA. However, at $z < 4$, the observation is mainly hindered by the transmission of the atmosphere at frequency $\nu > 380$ GHz, which is only available through the ALMA band 8 to ALMA band 10. Luckily, for galaxies $z > 4$, [C II] oan be observed efficiently with ALMA in many DSFGs, which makes it the essential tool to study the high-z DSFGs. Because of this, over the past six years, all the six ALMA large programs in the section of high redshift universe have used [C II] to probe the properties of very distant galaxies.

[C II] has been regarded as an effective SFR tracer of distant galaxies. The ionizing energy of [C II] is only 11.3 eV, which makes it easily to be excited in star-formation region (Narayanan & Krumholz, 2017). Over the past years, a main concentration has been focused on the testing of [C II] as a reliable tracer of the recent SFR (Ferrara et al., 2019; Zanella et al., 2018; Schaerer et al., 2020). Several promising works have indicated that the luminosity of [C II] is tightly correlated with SFR up to galaxies at $z \sim 7$ (Capak et al., 2015; Pentericci et al., 2016; Schaerer et al., 2020). However, some systems also show a relative [C II] deficit (Ota et al., 2014; Schaerer et al., 2015; Knudsen et al., 2016; Carniani et al., 2018), which inspired various speculations (Smith et al., 2017; Carniani et al., 2018; Lagache et al., 2018; Ferrara et al., 2019; Rybak et al., 2019; Romano et al., 2022; Liang et al., 2023). It is still not clear what is the main physical mechanism behind this divergence, as the tight correlation and deficit have been reported in a wide range of $L_{\rm IR}$ and dust properties. Therefore, $L_{\rm [C\,II]} -$ SFR scaling relation can not be applied blindly.

[C II] has also been widely used to constrain the properties of the ISM. It has been proposed to trace the total molecular gas (Zanella et al., 2018; Dessauges-Zavadsky et al., 2020) and the dust temperature (Sommovigo et al., 2022). All these relations either directly use the $L_{\rm [C\,II]}$-SFR relation or share similar underline assumptions, which make them also suffer from large uncertainties among different targets. Theoretical models show [C II] trace mainly the photodissociation regions (PDRs) and mainly associated with the atomic gas around the molecular clouds (Zhang et al., 2018a; Ferrara et al., 2019). This model is also generally supported by the observation of the nearby molecular clouds (e.g. Schneider et al., 2023) and the extended [C II] halos (e.g. Fujimoto et al., 2019). Heintz et al. (2021, 2022) explored the usage of [C II] to measure HI gas across the cosmic time. Such measurement should also suffer from uncertainties, but it offers statistical implication that the majority of the HI gas is not associated with galaxies in the early Universe ($z > 4$ Heintz et al., 2022). In short, [C II] is a powerful tracer to probe the ISM and even CGM back to the cosmic re-ionization, but we should be careful to not over-interprete the results. Therefore, future surveys should cover more spectral lines and multi-wavelength observations to consolidate these scaling relations.

1.4.3 Neutral carbon

[C I] has been regarded as an alternative molecular gas tracer as low-J CO transitions. Several unique advantages make [C I] attractive in many circumstances. [C I] has two popular transitions, the [C I]($3^{\rm P}_2 - 3^{\rm P}_1$) and [C I]($3^{\rm P}_1 - 3^{\rm P}_0$), hereafter [C I](2−1) and [C I](1−0). Their critical densities are $n_{10} \sim 500\,{\rm cm}^{-3}$ and $n_{21} \sim 10^3\,{\rm cm}^{-3}$, comparable to the most widely used CO(1−0) and CO(2−1), thus likely probing the same clouds of molecular gas (Papadopoulos et al., 2004). Both transitions are accessible with the submm/mm telescope. [C I](2−1) is also right close to CO(7−6), which makes it easy to observe along with the survey of CO lines.

The excitation energies of these two transitions are $E_1/k = 23.6$ K and $E_2/k = 62.4$ K, respectively, which make their ratio an independent measurement of the excitation temperature. Moreover, their excitation temperatures are higher than that of the ground transition of COs, which gives them higher contrast in the early Universe where the CMB temperature is much higher (da Cunha et al., 2013; Zhang et al., 2016). Moreover, [C I] transitions have lower optical depth compared to CO transitions, similar to 13CO (Ikeda et al., 2002), which makes them better total molecular gas tracer in the DSFGs.

[C I] has already given exciting results on the evolution of the DSFGs. Firstly, the sample size of DSFGs with available [C I] observations has reached ~ 100, which has been proved to be a robust tracer of the molecular gas across validated with other tracers (Alaghband-Zadeh et al., 2013; Valentino et al., 2018; Dunne et al., 2022; Gururajan et al., 2023). Secondly, extended [C I] emissions have been found in several systems, highlighting its potential to search for extended cold gas and trace the cold gas inflow from cosmic web (Emonts et al., 2018, 2023). Moreover, due to the unique advantages of [C I], it offers another effective way to probe the kinematics of the high-z DSFGs (e.g. Lelli et al., 2023).

1.4.4 Ionized gas

Oxygen and nitrogen are the next most common elements in DSFGs. [O III]88μm is another strong far-IR line after [C II]. Photon ionisation from O^+ to O^{++} [O III] is 35.1 eV, which can only exist in strong radiational fields. [O III] has been long thought to be an effective redshift and ISM prober as [C II] in distant galaxies (Inoue et al., 2014) and it has already become the second most frequently detected far-IR line in the galaxies at $z > 6$ (Inoue et al., 2016; Tamura et al., 2019; Witstok et al., 2022; Tamura et al., 2023). From the observations collected in the local Universe, the [O III] is found to be stronger than [C II] in the metal-poor dwarf galaxies and more closely correlated with SFR (De Looze et al., 2014; Cormier et al., 2015). In the early Universe, the detection of [O III] is mostly detected in the Lyα emitters and LBGs (Tamura et al., 2019; Hashimoto et al., 2019; Harikane et al., 2020; Witstok et al., 2022), supporting their general metal-poor. Currently, it is still unclear how strong the [O III] could be in the metal-enriched dusty galaxies, which is hampered by the limitation of available targets (Akins et al., 2022; Tamura et al., 2023). The resolved study from Tamura et al. (2023) showed that the spatial distribution of [O III] is separated from the dust continuum, which indicates that different regions contribute [O III] and dust emission, respectively. Another [O III] transition at 52 μm is also supposed to be bright, but it is harder to observe due to its shorter wavelength.

Far-IR [N II] is normally weaker than [O III] and currently only received a few reports in the literature. Current observations most focused on the [N II]122 μm and 205 μm (Doherty et al., 2020; Sugahara et al., 2021; Killi et al., 2023). Doherty et al. (2020) reported the detection of both lines in a strongly lensed HyLIRG 9io9, at $z = 2.556$. The ratio between these two lines yields an electron density $n \sim 300\,\mathrm{cm}^{-3}$, which is an order of magnitude higher than that of the Milky Way.

1.4.5 Dense gas

The dense gas, normally traced by HCN, HNC, and HCO^+, is directly associated with the star-forming clouds, which delivers the tightest and almost one-to-one correlation with the SFR (Gao & Solomon, 2004; Wu et al., 2005). The presence of the dense gas is also expected in the DSFGs, especially in the classical SMGs to support their boosted star formation rates. Oteo et al. (2017a) showed the presence of massive dense molecular gas in two SMGs at z=1.5-1.8. However, a recent larger survey of dense gas (Rybak et al., 2022) generally received non-detections, indicating a lower fraction of dense gas in DSFGs compared with local star-forming galaxies. In the prototypical SMG at z=3.04: SDP.81, the mid-J transitions HCN(4-3) and HCN(4-3) remain undetected with nearly 4 hr of ALMA on-source time (Rybak et al., 2023). These non-detections have been attributed to inefficient mechanical heating and sub-solar metallicity. With the advent of ALMA's low-frequency receivers: Band 1 and Band 2, such tests can be extended to a larger sample of SMGs with higher redshifts.

1.4.6 Kinematics

The formation of galactic disks is a key stage in galaxy morphological evolution. Their formation timescale and overall stability are sensitive to galaxy evolution models, which have motivated continuous efforts to search for disk galaxies in the early Universe (Hodge et al., 2012; Smit et al., 2018; Tsukui & Iguchi, 2021). In the traditional view of galaxy formation, primordial galaxies are thought to be turbulent systems, subject to gravitational instability, violent gas infall, frequent galaxy interactions, supernova explosions, and strong galactic winds (Pillepich et al., 2019). However, a surprisingly large fraction of star-forming galaxies at cosmic noon has turned out to be disk-dominated (e.g. Wisnioski et al., 2019), albeit clumpy and unstable (e.g. Swinbank et al., 2012; Förster Schreiber et al., 2018). In DSFGs, the observations of gas kinematics benefited from far-IR cooling lines such as [C II] and [C I]. Recent [C II] surveys of DSFGs have initiated the discoveries of cold disks at $z \sim 4$–5, detected directly via high-resolution observations (e.g. Neeleman et al., 2020; Fraternali et al., 2021; Tsukui & Iguchi, 2021; Lelli et al., 2021, 2023) or aided by strong lensing (e.g. Rizzo et al., 2020, 2021). Their velocity and velocity dispersion ratios (V/σ) are similar to those of local disk galaxies, indicating cold disks can develop on a short timescale.

Existing observations are still limited by the accuracy of the kinematic modelling. Due to the relative compact size of distant DSFGs, most observations can only marginally resolve the cold disk (Neeleman et al., 2020; Fraternali et al., 2021). In such situation, it is not easy to effectively distinguish the close merger and galactic disk (e.g. Rizzo et al., 2022). Moreover, to measure the velocity and velocity dispersion accurately, the inclination angle and a reliable rotation curve, with SNR$>$ 10 and include at least 3 independent resolution elements, are needed. Therefore, current analyses strongly depend on the underline assumptions on the kinematic status of the observed galaxies. Strong lensing can improve the spatial resolution and the SNR along the Einstein ring, but the consolidation of lensing and kinematic modelling could also significantly increase the free parameters (e.g. Rizzo et al., 2020). Deep observation remains expensive and inefficient to get the kinematics of galaxies over a statistical sample. One of the ALMA large programs, CRITAL, increased the sample to ∼ 20 with high angular resolution and deep [C II] observations. These deeper integration has revealed more complex substructures (Herrera-Camus 2023, in prep.).

Theoretical explanations of these early cold disks are also under debate. The current most versatile approach is the high-resolution zoom-in hydrodynamic simulations. (Kretschmer et al., 2022) conducted such a simulation recently, proposing that the formation of these cold disks is due to a period of co-planar, co-rotating gas accretion from the cosmic web. In such a scenario, the presence of the cold disk could be an ephemeral stage and its V/σ is sensitive to the gas flows around the galaxies. However, it is still unclear how frequently these alignments could happen in distant galaxies. Future surveys with a larger sample of targets from a complete sample are needed to understand the fraction of cold disks in different galaxies.

The cold gas disk should be carefully interpreted regarding the formation of galaxy disks in the early Universe. There is a long debate about the formation of the thin disk in the galaxies across cosmic time (Bird et al., 2013; Buck et al., 2020; Meng & Gnedin, 2021; Tamfal et al., 2021). The widely accepted ‘Upside-down’ scenario predicts the first formation of a thick stellar disk and gradually settles down to a cold and thin disk (Bird et al., 2013). A few works demonstrate that the stars also form the cold disk naturally and the stars are scattered gradually to a higher galactic plane through the gravitational instability and galaxy integration (Meng & Gnedin, 2021; Tamfal et al., 2021). Recent observation of the cold gas disk at $z > 4$ seems to support the second scenario, but more detailed comparisons are needed between the simulation and observation. For instance, from the zoom-in hydrodynamic simulation, the kinematics of the cold gas and stars are highly decoupled, where the gas shows stronger rotational motions than the collisionless stars (Pillepich et al., 2019). It should be noted that the selection of gas tracer also affects the measurement of the kinematic, where the cold molecular gas usually shows a larger value of the V/σ compared with the atomic gas and ionized gas (Kretschmer et al., 2022).

1.5 Clustering properties

Cosmological models predict the hierarchical evolution of structure over cosmic time. Proto-clusters of galaxies are supposedly the earliest over-densities to take shape, form stars and – one way or another – light up. They are thus excellent tracers of early structure formation, and are widely believed to evolve into the most massive clusters in today’s Universe (see the reviews by Overzier, 2016; Alberts & Noble, 2022).

For the few examples known, their extreme over-densities and active on-going star formation are consistent with model predictions for dense nodes at intersections of the so-called ‘cosmic web’ (Bond et al., 1996). Their activity is indicative of feeding by cold streams of gas from this web, which somehow sustains vigorous star formation and nurses the growth of super-massive black holes (e.g. Dekel et al., 2009). Proto-clusters are ideal laboratories, then, to study the interplay between baryons and various feedback processes.

Over the past decade, wide surveys at various wavelengths have led to the discovery of proto-clusters with diverse properties (see the review by Overzier, 2016). Early attempts to find proto-clusters relied mainly on searches for over-densities of broadband drop-outs, like the Lyman-break galaxies (LBGs — e.g. Steidel et al. 1998), which led to the discovery of giant structures such as that in SSA 22 at $z = 3.1$. Narrow-band and spectroscopic follow-up of these over-densities then led to the discovery of e.g. Lyα and Hα emitters in these structures (e.g. Rhoads et al., 2000; Ouchi et al., 2005). Such searches mainly and preferentially targeted the fields of distant QSOs and radio galaxies (e.g. Venemans et al., 2007; Wylezalek et al., 2013a). The haul from these early surveys stimulated further systematic searches for proto-clusters, as well as multi-wavelength follow-up (e.g. Toshikawa et al., 2018).

Contemporaneously, the discovery of SMGs have also been connected to the formation of proto-clusters. Due to their extreme star formation rates (SFRs) and massive gas reservoirs, SMGs are thought to trace peaks in the underlying density field and are viewed as the most likely precursors of present-day ellipticals (e.g. Lilly et al., 1999; Swinbank et al., 2006), which makes them another intriguing potential signpost for proto-clusters. Because of this, searching for SMGs in known proto-cluster environments – as well as for galaxy over-densities around SMGs – have been popular ways to find and confirm galaxy proto-clusters (e.g. Dannerbauer et al., 2014; Umehata et al., 2015).

Indeed, SMGs are found to trace galaxy overdensity and usually trace the most active region of the proto-clusters. Calvi et al. (2023) systematically analysed the environment properties of the 12 SMGs in the GOODS-N fields. Benefiting from the availability of the photometric redshift catalogues, they found 11/12 SMGs were found in the overdense region. Meanwhile, in several confirmed proto-clusters, the SMGs are spatially overlapped with the extended Lyα blobs (Ivison et al., 2000; Umehata et al., 2019; Daddi et al., 2022b). The spectral line and dust continuum observations all support that SMGs trace the region with the abundant cold gas and the boosted star-formation activities.

Moreover, interferometric follow-up of extraordinarily red objects discovered by single-dish submm surveys led serendipitously to the discovery of a new class of proto-clusters which are very-much dominated by DSFGs (Oteo et al., 2018; Miller et al., 2018). These systems are invisible to optical surveys due to the heavy dust obscuration, but are likely the most active core regions of a particular type of proto-cluster, or a phase of proto-cluster evolution. In these systems, the SFR of the central galaxies were simultaneously underoging starburst phase, which makes the total SFR of the proto-cluster cores are orders of magnitude higher than typical field galaxies at similar cosmic time. Meanwhile, their cold gas reservoirs and galaxy interactions are found to be much larger and more common than for field galaxies, which makes them natural place to trigger the earliest developement of galaxy cluster (e.g. Monson et al., 2021).

Because of the relatively evolved status, the proto-clusters cores stand out in the proto-clusters. And because of their extreme properties, they are the most valuable systems to study the transition of galaxy proto-clusters to clusters. In the following sections, I will focus on the observational results of the clustered DSFGs. I will first try to come out with a definition of proto-cluster core. Then, I will collect all the proto-cluster cores confirmed in the literature. At last, I will discuss the connection between the galaxy cluster and proto-cluster with all these proto-cluster cores.

1.5.1 The definition of proto-cluster core

Currently, there is no clear definition of galaxy proto-cluster cores in the literature. To aid our analysis, we define the proto-cluster core as the extreme overdensity $\delta_{\text{gal}} > 10$ over ~1 Mpc scale. This definition will only include the most central massive halo of the proto-cluster. In the regime of submm/mm, it means we expect to find multiple ($n \geq 3$) traditional SMGs within the field of view about one arcmin. Besides, the whole proto-cluster is not necessary to be virialised, but their central region is supposed to be massive enough to develop near-virialised states much quicker than the whole system. Based on the NFW dark matter profile with $1 \times 10^{13}\,\mathrm{M}_\odot$ and concentration of $c = 5$, the predicted escape velocity is $< 1500\,\mathrm{km}^{-1}$ (Miller et al., 2018). Taking into account the projected effects, I defined the velocity tolerance of possible member galaxies to be $\pm 2000\,\mathrm{km\,s^{-1}}$.

Constraining the halo mass of the proto-cluster is the key to tracing their evolution. However, unlike the virialised galaxy clusters, there is no standard way to measure their total halo masses. Inspired by several seminal works (Overzier, 2016; Wang et al., 2016; Oteo et al., 2018; Miller et al., 2018; Long et al., 2020), we have adopted three methods to estimate the halo mass of proto-cluster cores:

- If the core has more than ten spectroscopically confirmed members and the system is nearly virialised, we derive the halo mass from the velocity dispersion of the member galaxies (e.g. Wang et al., 2016).

  $$\sigma_v(M,z) = \sigma_{\text{DM},15}\left(\frac{h(z)M_{\text{total}}}{10^{15}M_\odot}\right),$$

  where the $h(z)$ is the dimonsionless Hubble parameter, $\sigma_{\text{DM},15} = 1083\,\mathrm{km}s^{-1}$ and $\alpha = 0.336$ is the canonical value derived from the simulations (Evrard et al., 2008). The method have been applied to several proto-cluster cores and generally give reasonable results as the following methods (Wang et al., 2016; Oteo et al., 2018).

- If the member galaxies have well-resolved velocity fields, the total halo mass could be calculated by summing the total dynamic mass of all the members (Hodge et al., 2013b). Assuming a spherically symmetric system:

  $$M_{\text{dyn}} = 2.33 \times 10^5 \gamma V_{\text{rot}}^2 R / \sin^2(i),$$

  where the $\gamma(< 1)$ is a dimensionless scaling factor to compensate for the structure of the disk and the dark matter halo. Based on the assumption of spherical halo and thin disk, I used $\gamma = 1$ gives an upper limit on the halo mass of each galaxy. Without the resolved kinematic fields, several works have also used the FWHM of the spectral line as a proxy for the total dynamic mass (Neri et al., 2003):

  $$M_{\text{dyn}} \sin^2 i(M_\odot) = 4 \times 10^4 \Delta v_{\text{FWHM}}^2 R$$

  This formula needs the independent measurement of the inclination angle. For those that cannot be reliably determined, we have also used the average correction $\left\langle \sin^2 i \right\rangle = 2/3$ (Tacconi et al., 2008).

- If the member galaxies have a reliable measurement of the total mass of cold gas or stellar mass, their halo mass can be measured from the corresponding scaling relations. The most popular method to derive the scaling relations is the abundance matching, which assumes the massive halos share the same distribution function as the massive galaxies defined by the total stellar mass, gas mass, or baryonic mass. However, proto-clusters are still in the early stage of accumulating their baryonic content, so the existing stellar-halo mass relation (SHMR) and baryon-halo mass relation (BHMR) may underestimate the total halo mass. Faucher-Giguère et al. (2011) has given the mass fraction of cold gas, stellar mass as a function of redshift and halo mass at different cosmic times. For my estimation, I used the maximum cold gas fraction of 5% to estimate the total mass if the mass of cold gas is reliable.

In most systems, various methods can be used to estimate the total halo mass. I will show in the following section that different methods normally give consistent results within 0.5 dex, which is accurate enough to predict the redshift evolution of the proto-cluster core with the models.

1.5.2 The family of galaxy proto-cluster cores

Tab. 1.1: A collection of all the proto-cluster cores.

Name	z	M_{gas}	M_{star}	M_{halo}	SFR	Reference
		$\times 10^{10}\mathrm{M}_\odot$	$\times 10^{10}\mathrm{M}_\odot$	$\times 10^{12}\mathrm{M}_\odot$	$\mathrm{M}_\odot \mathrm{yr}^{-1}$	
DRC	4.002	194 ± 24	77 ± 19	50 ± 20	6500	1,2
SPT2349-56	4.304	60	–	12 ± 7	6000	3
GN20ab[a]	4.053	9.3	28	5.1 ± 1.5	1600	4,5
AzTEC3	5.298	5.3	1	> 0.4	1500	6
RO-1001	2.910	–	76	40	1200	7
HATLAS J084933	2.410	29	40	5	6990	8,9
CL J1001	2.506	51	210	80^{+47}_{-29}	3400	10
QO-1000[b]	2.7603	–	79	12 ± 1	230	11

Note: The references are: 1–Oteo et al. (2018), 2–Long et al. (2020), 3–Miller et al. (2018), 4–Daddi et al. (2009), 5–Hodge et al. (2013b), 6–Capak et al. (2011), 7–Daddi et al. (2021), 8–Ivison et al. (2013), 9–Ivison et al. (2019), 10–Wang et al. (2016), 11–Ito et al. (2023). (a) the SFR and the stellar mass were adopted from Daddi et al. (2009) , which only includes the measurement in the two bightest galaxies. (b) I only included the three massive galaxies with $\delta z < 0.03$, which satisifies my definition of proto-cluster core.

The confirmed proto-cluster cores in the literature are still limited. The most two extreme proto-cluster cores are DRC and SPT2349-56. Oteo et al. (2018) and Long et al. (2020) have compared multiple methods to estimate the total halo mass of the DRC, an agreement has been found to be around $5 \pm 2 \times 10^{13}\,\mathrm{M}_\odot$ using both virialised assumption and 5% cold gas fraction.

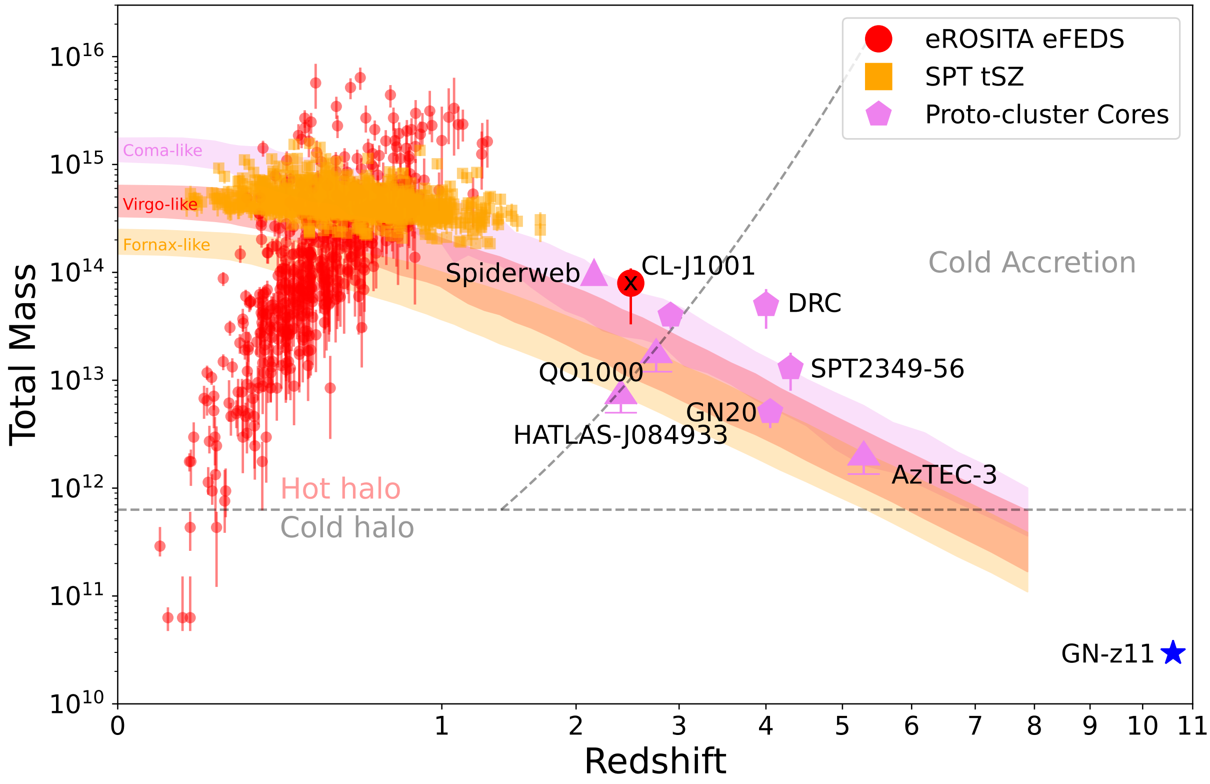

Fig. 1.3: The halo mass evolution of galaxy clusters as a function of cosmic time. I have included all the current confirmed proto-cluster cores (Daddi et al., 2009; Hodge et al., 2013b; Capak et al., 2011; Ivison et al., 2013; Wang et al., 2016; Oteo et al., 2018; Miller et al., 2018) and also the galaxy clusters found by the X-ray and kSZ surveys (Bocquet et al., 2019; Bleem et al., 2020; Huang et al., 2020). The shadows with different colours show the evolution of the halo mass of Coma-like, Virgo-like and Fonax-like clusters from the hydrodynamic simulation (Chiang et al., 2013). There is a clear division between galaxy cluster and galaxy proto-cluster around $z \sim 2$. It is still unclear whether there is a global transition from galaxy proto-cluster to cluster or it is just a selection effect opposed by the instrumental limitations.

Miller et al. (2018) also measured the dynamic mass of the core region based on the virialised assumption, it is around $1.2 \pm 0.7 \times 10^{13}\,\mathrm{M_\odot}$. Since all the member galaxies are dominated by the cold molecular gas, the total halo mass is also consistent with the result scaled from the assumption of 5 per cent cold gas mass, which yields a total mass of $1.3 \pm 0.5 \times 10^{13}\mathrm{M_\odot}$. Based on these two most famous examples, it supports the most simple assumption that 5 per cent cold gas mass is a reasonable assumption to estimate the total halo mass of gas-rich proto-cluster cores.

In addition to DRC and SPT2349-56, there are also several other proto-cluster cores that satisfied our definition. The first similar system is the GN20 at $z = 4.053$, found in the GOODS-N field. It contains three massive galaxies named: GN20, GN20.2a and GN20.2b. Their projected physical distance is only about 170 kpc and velocity difference is $\sim 2000\,\mathrm{km\,s^{-1}}$(Daddi et al., 2009). Daddi et al. (2009) and Hodge et al. (2013b) provided the dynamical mass of each of the three members, which added up to $5.1 \pm 1.5 \times 10^{12}\,\mathrm{M_\odot}$. Unlike the other systems, these three galaxy benefit from the deep multi-band photometries in the GOODS-N field, which makes it possible to estimate their stellar mass. The two brightest systems have already developed a significant amount of stellar mass. Summing up all the gas mass, the 5 per cent assumption yields a total halo mass of about $7.6 \times 10^{12}\,\mathrm{M_\odot}$. It is larger than the dynamic mass, but still consistent within the 2σ level. We adopted the total dynamic mass in this work.

Another high-z proto-cluster core is the AzTEC3, found in the COSMOS field (Capak et al., 2011). This proto-cluster core also contains two groups of galaxies at z=5.300 and z=5.298, respectively. These two groups were separated 137 kpc on the sky (870 kpc in comoving coordinates). Capak et al. (2011) also estimated the halo mass $> 4 \times 10^{11}\mathrm{M_\odot}$. Their halo mass is derived from the global dark matter to baryon ratio of 5.9, which can only be taken as the lower limit of the halo mass. As all the starburst galaxies in this core are gas-rich systems, the 5% cold gas fraction in the dark matter halo predicts a halo mass of $1 \times 10^{13}\,\mathrm{M_\odot}$. We will use the halo mass of $1 \times 10^{13}\,\mathrm{M_\odot}$ in this work to compare with other systems.

(Ivison et al., 2013) reported another proto-cluster core, HATLAS J084933, featured with a binary HyLIRGs. HATLAS J084933 is a unique system. Within a physical region of 120 kpc, five starburst galaxies have been identified based on the extensive observations from VLA, ALMA, and VLT/KMOS and *XMM-Newton* (Ivison et al., 2013, 2019). This system has already developed significant stellar $M_{\mathrm{star}} \sim 4 \times 10^{11}\,\mathrm{M_\odot}$, whose mass is slightly larger than the available total molecular gas $M_{\mathrm{gas}} \sim 2.9 \times 10^{11}\,\mathrm{M_\odot}$. HATLAS J084933 has no estimation on the total halo mass, the 5% cold gas fraction predicts a halo mass of $> 5 \times 10^{12}\,\mathrm{M_\odot}$, which is also compatible with stellar mass scaled halo mass $\sim 6 \times 10^{12}\,\mathrm{M_\odot}$. One of the two HyLIRGs, HATLAS J084933-W, has been found to host a relatively unobscured SMBH, which outshines the total X-ray luminosity and exposes its broad line region (Ivison et al., 2019). How could the dusty HyLIRG host such an unobscured AGN is unclear at the moment, but it offers an inspiring case to understand AGN's role during the transition from proto-cluster to clusters.

Another proto-cluster core featured in extended Lya emission is RO-1001, at z=2.91, reported by Daddi et al. (2021). This proto-cluster core was initially found in a proto-cluster candidate traced by an overdensity of radio sources. Then, ALMA's follow-up revealed three SMGs, two of which are ongoing merging galaxies. This is the most massive structure from a large survey of Lya emission in the overdense environments (Daddi et al., 2022a). Its reported halo mass is roughly around 4×10^{13} $M_\odot$. Daddi et al. (2021) find no more massive galaxies in the surrounding regions from their SED analysis. This target thus could be a galaxy halo instead of the core of a proto-cluster. Therefore, a spectral line survey in the surrounding region is needed to confirm its wider environment.

Wang et al. (2016) also found another dense core at $z \sim 2.5$, CL J1001. There are 11 massive galaxies in a very concentrated 80 kpc projected region. Most importantly, this could be a young cluster core instead of proto-cluster core as it has already developed extended X-ray emission detected by *XMM-Newton* and *Chandra*. The total halo mass is reported to be $8\times10^{13} M_\odot$ within the region that is 200 times above the critical density. This is a particularly intriguing target to understand how the proto-cluster core will finally evolve into a cluster. We will discuss it in more detail in the next section.

Another intriguing proto-cluster core is QO-1000, which is discovered in the COSMOS field (Ito et al., 2023). The most different feature of this proto-cluster core is that it consists of multiple quenched galaxies. Over a larger $7 \times 4\,\mathrm{Mpc}^2$, it has 14 photometrically quenched galaxies. In its central $\sim$ 1 Mpc, it contains three of its most massive galaxies with $\delta z = 0.0024$. It is the first proto-cluster core that mainly consists of red-sequence galaxies, with much ongoing star formation. Ito et al. (2023) provide the estimation of the total halo mass in this core is about$> 1.2\times10^{13}$ $M_\odot$.

I also included the cluster samples from the surveys of X-ray and the thermal SunyaevZeldovich effect in the $z < 2$ Universe. I included the X-ray galaxy clusters found through the *eROSITA* Final Equatorial-Depth Survey (Liu et al., 2022). These clusters were selected due to the presence of the extended X-ray emission from the intracluster medium (ICM). The halo mass of these X-ray selected clusters is derived from the mass-luminosity relation, where the luminosity is the total soft X-ray (0.1-2.4 keV) luminosity within the central 500 kpc (Liu et al., 2022). I also included the cluster selected by the SunyaevZeldovich effect from the South Pole Telescope (Bocquet et al., 2019; Bleem et al., 2020; Huang et al., 2020). The halo mass is defined as the integrated mass within the region that is 500× above the critical density. We have currently collected several hundred galaxy clusters, but we have much fewer confirmed proto-clusters.

1.5.3 The connection between proto-cluster and cluster

A tentative sharp transition between galaxy clusters and proto-clusters has become a puzzle in recent years. The X-ray sky surveys for extended thermalised X-ray emission and the large submm surveys for the Sunyaev-Zeldelvich effect have confirmed a large number of galaxy clusters at $z < 2$ over the past decade (Bocquet et al., 2019; Bleem et al., 2020; Huang et al., 2020). On the other hand, the galaxy proto-clusters are mostly found above $z = 2$ (Daddi et al., 2009; Hodge et al., 2013b; Capak et al., 2011; Ivison et al., 2013; Wang et al., 2016; Overzier, 2016; Oteo et al., 2018; Miller et al., 2018). It is still not clear whether such a sharp transition at $z \sim 2$ is physical or the result of the selection effects.

Within the proto-cluster cores, more evolved galaxies are more likely to be found at lower redshifts. In these proto-cluster cores at $z > 4$, most of the member galaxies are gas-rich starburst galaxies (Capak et al., 2011; Oteo et al., 2018; Miller et al., 2018); while at $z < 4$, galaxies become more massive, especially with their stellar mass, and it is more likely to host AGNs and quenched galaxies (Wang et al., 2016; Ito et al., 2023). However, such evidence is still weak due to the limited number of proto-cluster cores at each different stage. But all the current evidence is still generally consistent with the formation pattern that the massive galaxies are formed from the starburst galaxy triggered by major mergers and quenched subsequently by the depletion of cold gas and AGN feedback (Hopkins et al., 2006).

A bursty and intensive star-forming epoch is also preferred by stellar population analyses of the massive galaxies in modern-day clusters. The bulk of the stellar mass in those systems formed significantly earlier than $z \sim 2$ (e.g. Stanford et al., 2006; Rosati et al., 2009; Willis et al., 2020). Coupled with the finding that significant star formation usually lasts no more than ~1 Gyr, proto-clusters should host a cosmologically brief period of very intense star formation before becoming galaxy clusters. Given the total stellar mass of 'red and dead' galaxies in the central regions of a typical massive cluster, it is reasonable to assume that the high-redshift progenitors of local cluster have average SFRs of $> 1000\,\mathrm{M}_{\odot}\,\mathrm{yr}^{-1}$ over a period of a ≈1 Gyr, and significantly higher SFRs on shorter timescales given the stochastic nature of star-formation events. Plausibly such episodes are coincident with the discovery of proto-cluster cores.

Extreme over-densities of intensely star-forming galaxies in proto-cluster cores are also consistent with predictions for the densest filaments at intersections of the cosmic web. Recent observations do directly reveal such node in the Lyα mapping of the proto-clusters. Daddi et al. (2021) found an extended Lyα blub around the proto-cluster core RO-1001, where the morphology and kinematics of Lyα both prefer the origin of the gas inflow from the cosmic web. Moreover, ultra-deep Lyα observation towards the field SSA22 directly revealed the cosmic web traced by the extended Lyα emission, wherein the SMGs preferentially reside in these extended Lyα nebula (Umehata et al., 2019). Consequently, identifying and fully characterising these systems is not only crucial to our understanding of the evolution of galaxies and clusters of galaxies, but they also excellent regions in which to search for signatures of the very highest levels of cold gas feeding (e.g. Daddi et al., 2022a), and to explore the fate of gas and dust expelled from galaxies during their strongest periods of stellar feedback (e.g. Spilker et al., 2020), and to witness the emergence of AGN (Ivison et al., 2019; Wylezalek et al., 2022). I will come back to this topic and provide more observational results in Chapter 3.

1.6 This thesis

The field of DSFGs has received tremendous progress over the past 2-3 decades, but as I have pointed out, there are still many uncertainties and puzzles in the explanation of different subsamples of DSFGs. My PhD project focuses on two ALMA surveys: ALMACAL and ALMARED, which surveyed the populations of DSFGs in the faint and bright end of the number counts, respectively. These two surveys offer unique opportunities to conduct multi-waveband surveys of DSFGs and explore their internal and environmental properties. Meanwhile, I also initiated the observation of dust polarisation in DSFGs, which offers a new window to probe their ISM and galactic magnetic fields. This thesis is structured as follows:

- **Chapter 2:** In this chapter, I will further extended the disscussion of the multi-band submm/mm sky surveys and try to combine all the available surveys with ALMACAL to constrain the resolved fraction of the cosmic infared background.

- **Chapter 3:** In this chapter, I will first caution the future submm/mm surveys to search for galaxy proto-cluster cores, where the project effects will heavily contaminate the selection methods motivated by galaxy overdensity. Then, I will report a new sample of galaxy proto-cluster core from our specially designed ALMARED survey. I will give further evidence that these proto-cluster cores are connected to the formation of the brightest cluster galaxies in the today's cluster.
- **Chapter 4:** In this chapter, I will first introduce the effective window to study the cosmic magnetism with the full-polarisation observation of dust. Then, I will summarise the first results come from our pioneer projects and discuss the possible mechanisms to explain current observation.
- **Chapter 5:** In this chapter, I will conclude my PhD work and outlook several fields which can be further explored associated with DSFGs and submm/mm observations.

Chapter 2

Multi-band ALMA survey for DSFGs

2.1 Motivation

Several decades after the discovery of the CIB, various galaxy populations are known to make significant contributions. Among them, strongly star-forming galaxies with total infrared luminosities larger than $10^{11}\,L_\odot$ (Sanders & Mirabel, 1996) are thought to be the main contributors, with galaxies at different redshifts contribute to different parts of the CIB. In the mid- and far-infrared (-IR), the CIB largely comprises luminous and ultraluminous IR galaxies (LIRGs and ULIRGs) at $z < 1.5$ (Elbaz et al., 2002). In the submm/mm wavelength regime, the CIB is dominated by dusty, star-forming galaxies (DSFGs) at higher redshifts ($z \gtrsim 1.5$). In the submm/mm bands, dimming due to increasing redshift is compensated by stronger dust emission as the observing frequency traces rest frequencies progressively closer to the peak of a typical dust spectral energy distribution (SED). This makes the submm/mm wavelength regime particularly well suited to the detection of dust-rich galaxies at high redshift (e.g. Blain & Longair, 1993; Smail et al., 2002)

The first submm-selected galaxies (SMGs, the brightest DSFGs) were found by continuum surveys at $850\,\mu$m with the 15-m James Clerk Maxwell Telescope (e.g. Smail, Ivison & Blain 1997; Barger et al. 1998; Hughes et al. 1998; Eales et al. 1999), which ushered in the era of submm cosmology. Since then, SMGs have been studied extensively at different wavelengths, using radio, X-ray and optical telescopes to trace their properties (e.g. Ivison et al., 2002; Alexander et al., 2005; Chapman et al., 2005, respectively). New instruments and associated surveys have increased the family of SMGs selected at the original wavelength (e.g. Chen et al., 2016; Geach et al., 2017), at longer wavelengths ($\lambda > 1\,$mm, e.g. Scott et al., 2012; Magnelli et al., 2019), as well as at shorter wavelengths, e.g. with the *H*erschel Space Observatory ($\lambda < 500\,\mu$m, e.g. Eales et al., 2010; Oliver et al., 2010).

A significant advance came with the advent of the Atacama Large Mm/submm Array (ALMA), which was immediately able to pinpoint and even spatially resolve many of the earlier bright, single-dish-selected SMGs (e.g. Hodge et al., 2013a; Ikarashi et al., 2015; Simpson et al., 2015a; Stach et al., 2019). Some especially bright lensed sources proved to be ideal targets for detailed studies of their internal structures and physical states (e.g. Geach et al., 2018; Rizzo et al., 2020;

Dye et al., 2022), while other bright SMGs were resolved into multiple sources, leading to the discovery of proto-clusters of extreme DSFGs (e.g. Ivison et al., 2013; Oteo et al., 2018; Miller et al., 2018).

ALMA also offers the sensitivity and spatial resolution necessary to push below the confusion limit imposed by single-dish imaging, allowing the discovery of much fainter DSFGs. As happened with the single-dish telescopes that came before, two methods have been used to undertake surveys for DSFGs: long integrations in conventional ‘deep fields’ (e.g. Walter et al., 2016; Dunlop et al., 2017; Umehata et al., 2018) and the targeting of lensing clusters to probe smaller areas where the gravitational magnification is high (e.g. Gonzlez-Lpez et al., 2017; Laporte et al., 2021), enabling the discovery of faint DSFGs in the epoch of re-ionisation (e.g. Laporte et al., 2017b; Tamura et al., 2019; Fudamoto et al., 2021).

To create a full inventory of dusty galaxies, we need deep, wide and blind surveys. Before ALMA, single-dish telescopes had already covered 10s to 100s of square degrees of the sky and found many thousands of bright FIR/submm/mm sources. However, the diffraction limit prevented us from delving below the intensely star-forming ULIRG regime, aside from a small number of strongly lensed systems. It was also necessary to make model-based corrections to the counts to deal with issues associated with source blending, and a significant portion of bright SMGs found by single-dish telescopes were later resolved into multiple sources by interferometers (e.g. Ivison et al., 2007; Wang et al., 2011; Karim et al., 2013). Using ALMA, it has been possible to extend the submm/mm detection limit down to the sub-mJy level. Even with the small number of antennas available in Cycle 0, ALMA pushed down to $S_{870\mu\mathrm{m}} \sim 0.4\,\mathrm{mJy}$ in the Extended *Chandra* Deep Field South (Karim et al., 2013; Hodge et al., 2013a). Since then, extensive follow-up campaigns have been carried out for the bright SMGs found in earlier single-dish surveys (e.g. Weiß et al., 2013; Miettinen et al., 2017; Brisbin et al., 2017; Cowie et al., 2018; Stach et al., 2019; Simpson et al., 2020). Meanwhile, blind ALMA surveys have been continuously enlarging the survey area and/or improving the detection limit (e.g. Hatsukade et al., 2016, 2018; Walter et al., 2016; Dunlop et al., 2017; Umehata et al., 2018). Very recently, the ALMA Spectroscopic Survey in the Hubble Ultra Deep Field (ASPECS) Large Programme (Walter et al., 2016) has achieved a sensitivity of $10\,\mu\mathrm{Jy\,beam}^{-1}$ (González-López et al., 2019, 2020) and although ALMA is not optimised for large sky coverage, the GOODS-ALMA project has observed the 72 arcmin^2 Great Observatories Origins Deep Survey South field (GOODS-South) in two different array configurations (Franco et al., 2018; Gmez-Guijarro et al., 2022). Considerable time has also been invested in the well-known legacy fields, which have also inspired dedicated data mining projects (e.g. Ono et al., 2014; Zavala et al., 2018; Liu et al., 2019).

All these ALMA surveys have helped to constrain the number counts and the properties of DSFGs. Number counts – the projected galaxy surface density as a function of flux density – represent the most basic measurement we can glean from such observations. They therefore provide a simple test of the validaty of models of galaxy formation and evolution. Indeed, both semi-analytic models (e.g. Lacey et al., 2016; Somerville et al., 2012; Lagos et al., 2020) and post-processing models in hydrodynamic simulations (e.g. Shimizu et al., 2012; McAlpine et al., 2019; Cowley et al., 2019; Lovell et al., 2021) have struggled to reproduce the number counts of DSFGs. In addition, accurate and unbiased number counts are essential to resolve the CIB and determine the contributions of different galaxy populations to the total CIB. However, due the

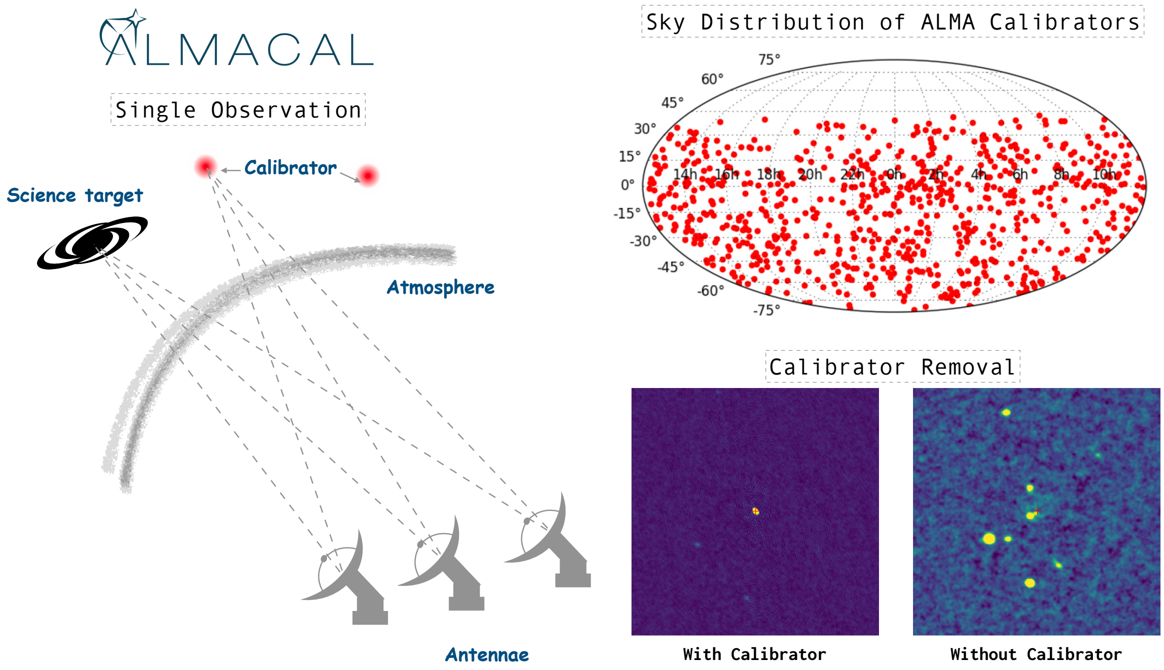

Fig. 2.1: Illustration of ALMACAL sky survey. All the ALMA calibrators are randomly distributed in the sky below R.A. < 45°. These calibrators are observed along with the science targets in every ALMA project. By removing the bright central calibrator in the field, ALMA can survey the sky in the vicinity of each calibrator. With the ability to combine multiple visits to each calibrator, ALMACAL is one of the widest and deepest submm/mm sky surveys at multiple wavelengths.

very high demand for ALMA time it has become difficult to go much deeper or wider, and it has proved even more difficult to justify covering the same areas at multiple submm/mm wavelengths.

ALMACAL is a novel submm/mm survey that exploits the ‘free’ data that ALMA must collect to ensure its observations can be processed to make spectra, images and cubes with accurate positions, polarisations and flux densities. With ALMACAL, we survey the immediate vicinity of each calibrator, which are typically blazars at $z < 1$ (Bonato et al., 2018; Klitsch et al., 2019a). Since ≈ 20 per cent of all ALMA observing time is spent on calibration, ALMACAL is already competitive with the widest and deepest submm/mm surveys. It does come with an obvious disadvantage: the lack of ancillary data at depths comparable with classic deep fields, such as the Cosmic Evolution Survey (COSMOS) or GOODS. However, it also comes with several major advantages. Number counts of ALMACAL are largely immune to cosmic variance, since it covers a great many pointings scattered across the observable sky, and the presence of an in-beam calibrator allows for perfectly calibrated submm/mm/radio imaging with a very high dynamic range. The blind detections share the same spatial and spectral set-ups as the science targets, whatever they may have been, which makes it possible to search for spectral lines (in absorption or emission) and to study morphologies at the very highest spatial resolution, up to 20 mas (see

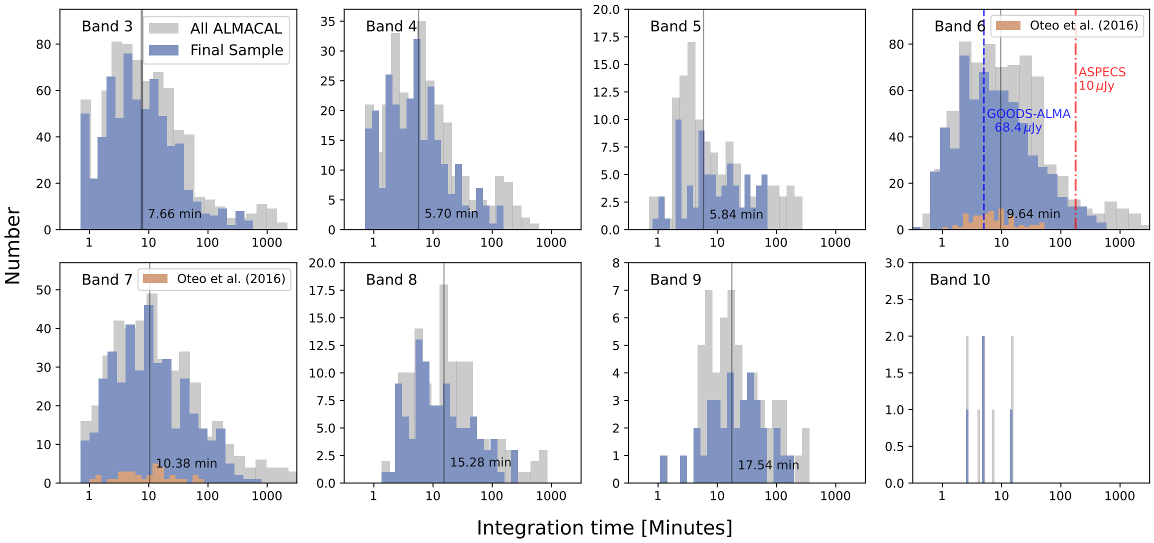

Fig. 2.2: Statistics of on-source time of ALMACAL observations from ALMA bands 3 to 10. For every panel, the abscissa is the accumulated total on-source time in units of minutes. The ordinate is the number of fields in every time step. We show the total data available in each band and the data used in this work. Some of them have been dropped because of calibration errors or strong residuals after calibrator removal (see §2.3.1). In bands 6 and 7 we also show the statistics for the data used by Oteo et al. (2016a), illustrating the immense increase since that time. Both the number of fields and the sensitivity in other bands have increased, making it possible to expand the number counts to a wider flux density range and to multiple ALMA bands. In band 6, the two vertical lines show the sensitivity reached by the ASPECS (González-López et al., 2020) and GOODS-ALMA surveys (Franco et al., 2018; Gmez-Guijarro et al., 2022). ALMACAL provides a good compromise between their sensitivity and sky coverage.

Oteo et al., 2017b). Meanwhile, these fields will be visited repeatedly by future submm/mm and radio interferometers, which will keep improving the sensitivities in these fields at different wavelengths.

In our earliest attempt at mining the ALMA calibration data, Oteo et al. (2016a) described the first search for DSFGs in ALMACAL. With multi-band data available for most of the fields, they derived dual-band number counts, in bands 6 (1.2 mm) and 7 (870 μm), using the data collected before 2015. Since then, both the number of calibrators and their on-source integration times have grown by more than an order of magnitude. Using the new data, Klitsch et al. (2020) reported the first number counts in ALMA band 8, at 650 μm. Besides number counts, searches for molecular absorption or emission lines have put constrains on the evolution of molecular gas density over cosmic time (Klitsch et al. 2019a; Hamanowicz et al. 2022, submitted). In this work, we extend the number counts of Oteo et al. (2016a) and Klitsch et al. (2020) to the most up-to-date ALMACAL dataset, covering ALMA bands 3, 4, 5, 6 and 7, at wavelengths from 3 mm down to 870 μm, respectively.

The paper is structured as follows: in §2.2 we introduce the ALMACAL project and its results to date; §2.3 details our data analysis, including source detection, corrections for flux boosting, sample completeness, effective area, source classification and selection bias. §2.4 presents the final source catalogue and the number counts in five different ALMA bands. §2.5 contains relevant discussion and comparison between our number counts with literature and model predictions. The various fractions of the CIB resolved by ALMA are also presented in §2.5. We summarise our key results in §2.6. Throughout this paper, we follow the terminology used in the review by Hodge & da Cunha (2020). In this paper, we assume a ΛCDM cosmology with H_0=67.7 and $\Omega_m = 0.31$ (Planck Collaboration et al., 2020).

2.2 ALMACAL

ALMACAL[1] aims to exploit all ALMA calibrator scans for science (Zwaan et al., 2022). Running now for more than ten years, ALMACAL has already accumulated more than 1,000 h of data. As of March 2020, ALMACAL includes 1,001 calibrators. About 97 per cent of them are classified as blazars (Bonato et al., 2018), which are quasars – active galactic nuclei – whose jets are oriented very close to our line of sight, such that relativistic beaming makes them extremely bright. The calibrators are spread all over the sky and used to calibrate science targets local to them. Depending on their brightness, compactness and flux stability, they can be used to calibrate bandpass, gain (complex amplitude and phase), flux density and polarisation. In a typical ALMA scheduling block (SB), two or more calibrators will be observed, along with the science targets, so calibrators share the same instrumental configuration as the science targets. The rich data buried in the calibration observations make them far more useful than their original intention.

One intriguing example is the calibrator named J1058+0133, one of the brightest blazars close to Cosmic Evolution Survey (COSMOS) field. Oteo et al. (2017b) found two $z = 3.4$

---

[1] https://almacal.wordpress.com

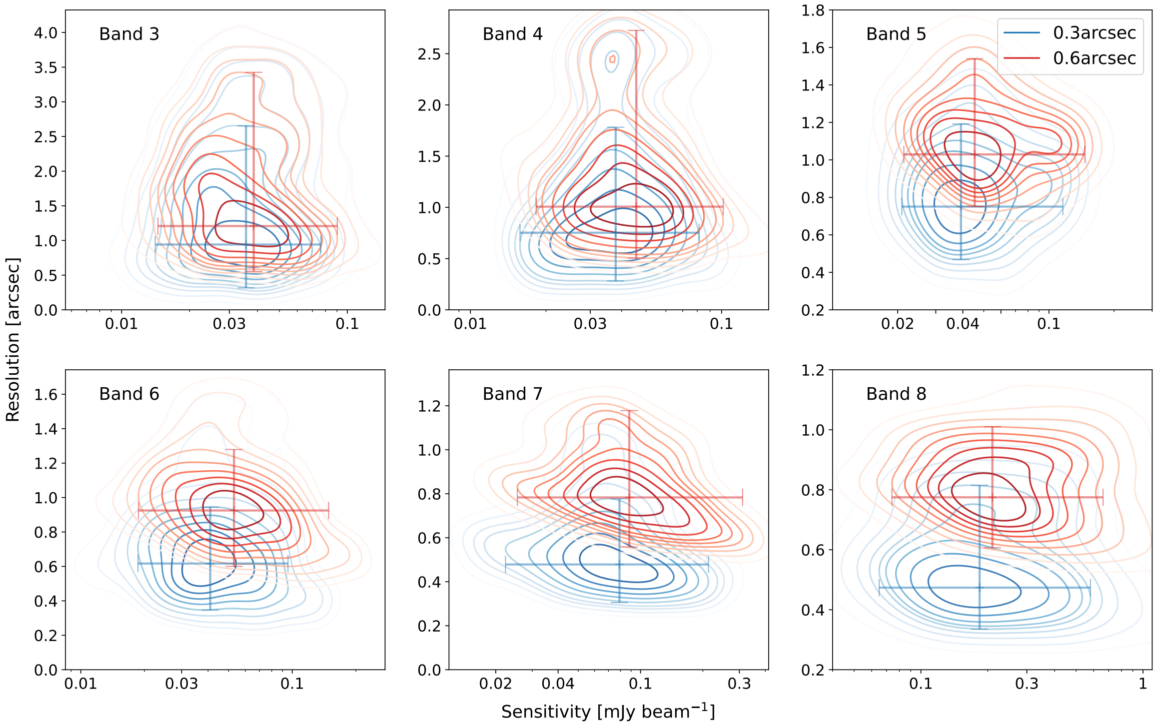

Fig. 2.3: Density distribution of sensitivity versus resolution for all the combined images in each ALMA band. We show the distributions for each set of tapered images in different colours. The lowest contours include 90 per cent of the data points and decrease in steps of 10 per cent. Our two *uv* tapers ensure the images have at least 0.3″ and 0.6″ spatial resolution, respectively. In the final images, most of the sources should remain unresolved.

SMGs behind this calibrator. ALMA has invested hundreds of hours in the COSMOS field, which makes J1058+0133 one of the most frequently visited calibrators, observed in total for around 150 h. With the rich range of configurations that ALMA has employed in the COSMOS field, Oteo et al. (2017b) were able to create a high signal-to-noise multi-band image of the two SMGs, self-calibrated on the timescale of 1 sec using the in-beam calibrator, J1058+0133, with 20 mas spatial resolution. Additionally, multiple CO lines have been identified in this system, yielding a remarkable spectral line energy distribution, making this amongst the most extraordinary datasets gathered for any SMG.

A detailed description of the ALMACAL data retrieval and calibration pipeline was given in the first ALMACAL paper, Oteo et al. (2016a), which we summarise briefly here. All the data reduction is carried out using the Common Astronomy Software Application (CASA – McMullin et al., 2007). Firstly, the calibration data are requested from the ALMA science archive. Next, we apply the standard calibration pipeline, following the *ScriptForPI.py* Python script for each project and splitting out the calibrator data. For calibrators that lack flux density calibration – mostly the bandpass calibrators – their flux densities are recovered from the internal flux tables in the data delivery package. After that, two cycles of self-calibration are performed in the multi-frequency synthesis (mfs) image. The first cycle is focused on phase-only solutions; the second cycle corrects the amplitude and phase together. Next, a point-source model of the calibrator is removed from the *uv* data by the CASA internal tool UVMODELFIT to make science-ready data. Finally, the fully calibrated measurements are re-binned to a channel width of 15.6 MHz. After these pre-processing steps, the scans contain the fully calibrated data, with the central calibrator subtracted out.

This specially designed pipeline takes full advantage of the bright, central source to do self-calibration, avoids possible effects due to variability, and it is straightforward to combine calibrated data from many different projects. The dynamic range of images made for calibrators exceeds 10^4, where the background sources or calibrator jets are typically 100 times weaker than the calibrators themselves. After removing the central, bright calibrator from each dataset, we can also easily identify the observations that are suitable for combination. At the moment, we simply reject any problematic data (see §2.3.1 for details). Due to the immense size of the dataset, it is not realistic to re-calibrate all the problematic observations manually.

ALMA has invested significant observing time in various deep cosmology fields (see a summary in Hodge & da Cunha, 2020). ASPECS is currently the deepest blind survey, covering 2.9 arcmin^2 with an r.m.s. sensitivity that reaches $\sigma_{1.2\text{mm}} \approx 9.3\,\mu$Jy beam^{-1} (González-López et al., 2020). Meanwhile, GOODS-ALMA represents the largest blind survey, with sky coverage of 72.4 arcmin^2, with a much shallower sensitivity, $\sigma_{1.2\text{mm}} \approx 70\,\mu$Jy beam^{-1} (Gmez-Guijarro et al., 2022). ALMA follow-up of bright sources from the single-dish surveys is the most efficient way to probe the very brightest SMGs and such an approach has been employed extensively in the existing legacy fields (e.g. Weiß et al., 2013; Stach et al., 2019; Simpson et al., 2020). ALMACAL manages to combine the best depth and sky coverage – see §2.3.6. In Fig. 2.2, we show the on-source time for all our calibrators in the different ALMA receiver bands. Compared with Oteo et al. (2016a) in bands 6 and 7, the total on-source time and the number of available calibrators have both increased dramatically. Ultimately, ALMACAL will go deeper and wider than any ALMA blind survey, unless a significant strategic investment is made.

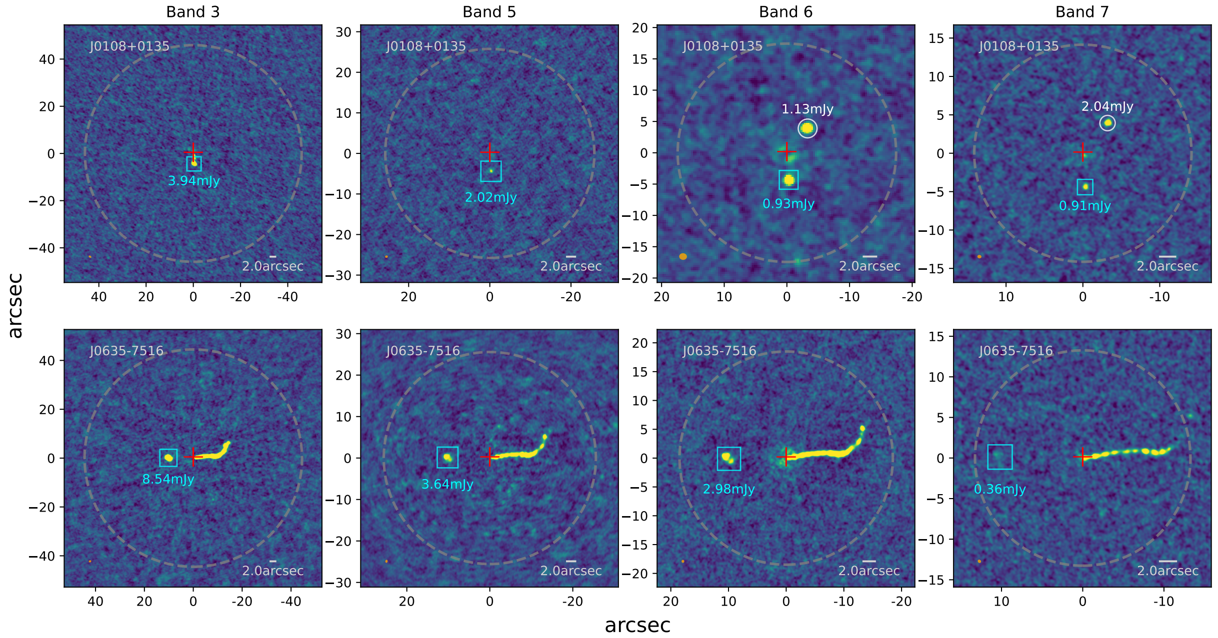

Fig. 2.4: Two examples of detections in ALMACAL in different bands. The wavelength is decreasing from left to right. In every image, the central bright calibrator has been removed – its position is marked with a red cross – and all the available observations have been combined. The dashed grey circle is the FoV adopted in our work, which is 1.8 × FWHM of the respective primary beam. In the top row, the images in ALMA bands 4, 5, 6 and 7 towards the field J0108+0135 are shown, where one SMG (marked with a white circle) and one jet (marked with a cyan square) have been found. The SMG has higher flux densities at higher frequencies, while, in contrast, the flux density of the radio jet decreases. In the bottom row, the images from the field J0635−7516 are shown, where two-sided extended radio jets have been discovered. We have classified each detection based on its spectral index and morphology.

2.3 Observations and Analysis

We detail the steps we have taken regarding data selection and imaging in §2.3.1, then present the methods and steps taken to construct the source catalogue in §2.3.2 and 2.3.3. After that, we derive the correction functions for our survey, including those for flux deboosting (§2.3.4), completeness (§2.3.5), effective area (§2.3.6) and effective wavelength (§2.3.7). Finally, we classify our sources in §2.3.8 and discuss selection biases in §2.3.9.

2.3.1 Data selection and imaging

A rose has its thorns. The bright calibrator at the centre of the field can be helpful for self-calibration but at times it brings problems. Firstly, the calibrator is not always a point source; blazars have jets, and they are not always (or have not always been) oriented along the line of sight; they can be slightly resolved, or amplitude and phase errors can remain after the calibration steps. All these issues can introduce residuals after point-source subtraction of the central calibrator. To avoid residuals corrupting the final, combined image, we applied two imaging cycles.

In the first imaging cycle, we image the visibilities of every single dataset after point-source removal, then we visually inspect all the images, discarding the ones with incorrect calibration or strong residuals. This step alone led to the loss of around half of our data, but it ensured the best quality of the final combined images. Having removed the poor quality data, we re-scale the weighting of different observations with the CASA task, STATWT, then combine the observations of each field using the task, CONCAT. All the images were made using TCLEAN in CASA version 5.7.0 in ‘mfs’ mode, which combines all the frequency channels into a single continuum image.

The second imaging cycle use the combined, re-scaled visibilities to create science-ready images. One of the major goals of ALMACAL is to search for DSFGs in the calibrator fields. From previous ALMA follow-up of known, bright DSFGs, we expect that they have compact dust morphologies, with a median size ~ 0.5 arcsec (FWHM, Ikarashi et al., 2015; Simpson et al., 2015b; Gullberg et al., 2019). To optimise the sensitivity to compact sources, we adopted a natural weighting scheme and the Hogbom (Högbom, 1974) deconvolution algorithm to clean our images. In addition, to avoid resolving DSFGs we further tapered the visibilities with Gaussian kernels of 0.3 and 0.6 arcsec. We show the final resolution versus sensity in Fig. 2.3, our uvtapers ensure the images have at at least $0.3''$ and $0.6''$ spatial resolution, respectively. These two tapered images were used for source detection, and to quantify the fraction of missing flux density. During the deconvolution, we first estimated the sensitivity of the final combined visibilities using the function SENSITIVITY in CASA *Analysis Utilities*[2] and used this as the threshold in TCLEAN. In every major cleaning cycle, the built-in AUTO-MULTITHRESH algorithm in TCLEAN was used to search for secure emission and determine the clean regions in the residual maps. We cleaned the image within an area defined as 1.8× the FWHM of the primary beam. After cleaning, IMPBCOR was used to correct the primary beam response. In the following steps, as a convenience, the image *without* the primary beam correction was used for source detection, while the primary-beam-corrected image was used for flux density measurements.

2.3.2 Source detection

Following our two earlier number counts papers, Oteo et al. (2016a) and Klitsch et al. (2020), we performed the source detection using SEXTRACTOR as it can deal easily with complex source structures (Bertin & Arnouts, 1996). The search for detections is made in the images before primary beam correction, corresponding to a signal-to-noise image. We detected residual signals from the calibrator in some fields, so we masked the central part of each map to a radius of 2 arcsec. This radius is 2× larger than the masking used in band 8 (Klitsch et al., 2020) but the loss is small for the larger fields of view (FoV) at the longer wavelengths explored here. To improve the reliability of the source detection in ALMACAL, we applied the source-finding algorithm in the two tapered images simultaneously. We search sources with SNRs higher than 3σ and then only accept detections with peak SNR higher than 5σ in at least one of the two

---

[2] https://casaguides.nrao.edu/index.php/Analysis_Utilities

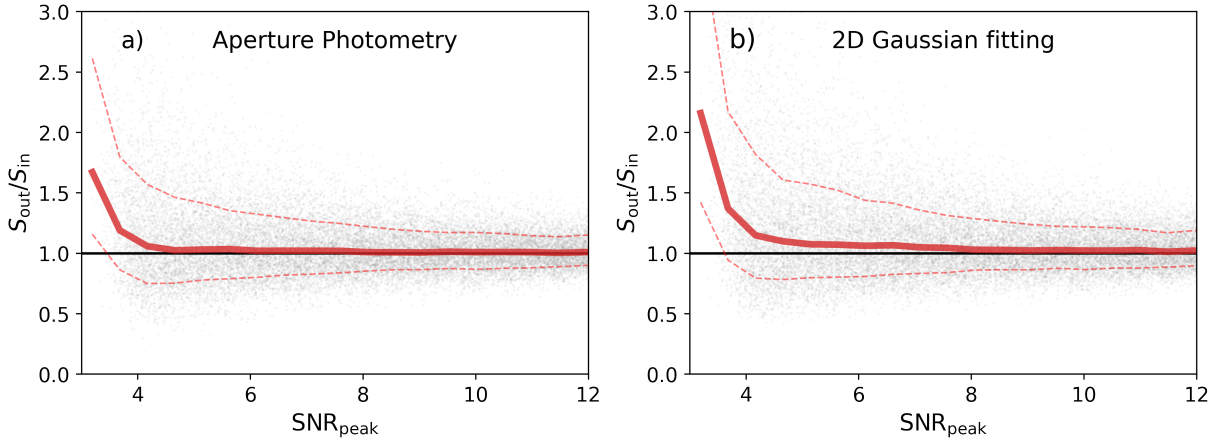

Fig. 2.5: Flux boosting as a function of peak SNR. The boosting effects have been estimated from the recovered flux density of artificial sources (0.2 arcsec FWHM) randomly injected in the observed image. The abscissa is the peak SNR of the injected source, and the ordinate is the ratio between the measured flux density, $S_{\rm out}$, and the injected flux density, $S_{\rm in}$. The two methods used for flux measurements are shown in the two panels. Panel a) is used aperture photometry; panel b) is used 2-dimensional Gaussian fitting. The red solid line is the median value in different SNR bins, the red dashed lines enclose 68 per cent of the points in each bin. Our simulations indicate that different flux measuring methods will lead to slightly different deboosting functions, where the Gaussian fitting tends to capture more positive noise during the fitting. In our final catalogue, we report the two measurements corrected by their own de-boosting functions.

tapered images. [3] Our adoption of a 5σ final detection cut prevents us from finding very faint sources, but ensure our detected sources are robust. In Fig. 2.4, we show the source detections at different wavelengths in two example fields.

We searched for sources within 1.8× the FWHM of the primary beam, which corresponds to 15 per cent of the peak sensitivity. This represents a reasonable trade-off, yielding a large total effective area whilst maintaining reasonable sensitivity. The final positions of all the detections are defined by their signal-to-noise centroid (Bertin & Arnouts, 1996).

2.3.3 Flux density

We measured the flux densities of our detections using aperture photometry. Aperture photometry was undertaken using the PYTHON package, PHOTUTILS, with elliptical apertures. The size of the aperture is important: selecting too large or too small an aperture introduces uncertainties into the total flux density. Since most of our sources are unresolved, we use the synthesised beam as the shape of our aperture. Based on our testing, an aperture size that is twice the synthesised beam with an appropriate aperture correction delivers the most reliable flux density. The aperture correction assumes an unresolved point source. To better capture all the flux for our detections, we use the $0.6''$ tapered images as the primary images to measure their flux densities. The flux difference between the two *uv* tapered images are small, less than 10 per cent, supporting our assumption that most of emissions from our detections remain unresolved.

We also report the flux densities from 2-dimensional Gaussian fitting. The Gaussian fitting was achieved by ASTROPY.MODELING with the GAUSSIAN2D model and the *LevMarLSQFitter* fitting algorithm. Before applying the fitting algorithm, a cut-out is taken from the original image, measuring 5× the major axis of the synthesised beam. During fitting, the amplitude, axial ratio, positional angle could change freely. The centre of the Gaussian function is allowed to shift by ±0.1 arcsec relative to the photometric centre. The flux density was calculated by integrating the best-fitting Gaussian.

The sensitivity of every ALMACAL field is not uniform, decreasing outwards from the phase centre. This means sources need to be brighter at larger radii to be detectable. Because the primary-beam correction scales up the flux density and noise by the same factor, the peak SNRs of detections remain the same. In the following two sections, we discuss flux boosting and sampling completeness as a function of peak SNR, which is convenient to apply to all our fields.

2.3.4 Flux deboosting

The well-known effect that we must consider when measuring accurate flux densities is the so-called 'flux boosting' of faint sources (Hogg & Turner, 1998). It has two origins (e.g. Coppin et al., 2006; Casey et al., 2014). First, we are more likely to detect faint sources that have been scattered towards higher flux densities by random noise (Eddington flux boosting), an effect that

---

[3] Most of our detections (>90 per cent) have been found in both of the two tapered images. Those detections found in just one of the tapered images, with a counterpart in another ALMA band, are also classified as robust detections; those without counterparts in any other ALMA band are marked as uncertain. The variation of the final number counts that include these uncertain sources are discussed in 2.20

can be mitigated to some extent by choosing a high SNR detection threshold. Second, where there is a rapid increase in the number of sources as we delve fainter, i.e. when the source counts are steep, then the resolution element of a telescope may include additional sources that are individually fainter than the detection threshold. The latter issue has been common for single-dish submm/mm telescopes, because of their typically large diffraction limit, $\gtrsim 10$ arcsec, and because the source counts of DSFGs at $S_{850\mu\mathrm{m}} \gtrsim 6$ mJy are steep (Simpson et al., 2020).

However, the blending effects should be very much less significant for deep, high-resolution interferometric observations, such as those obtained using ALMA. The spatial resolution of ALMA is considerably higher than that of single-dish telescopes, which makes the confusion noise significantly lower. At the same time, the number counts of DSFGs have been found to flatten at fainter flux densities (e.g. Stach et al., 2018; González-López et al., 2020). The combination of these two differences with historic work ensures that blending effects are smaller for interferometer-based work. Nevertheless, we have tested the degree of flux boosting affecting our detections via bespoke Monte Carlo simulations in all our fields.

To preserve the noise characteristics of our fields, as well as any undetected faint sources, we injected artificial sources directly into our observed images. Before we begin, we clean the image and run our source-finding algorithm to search for detections, masking any before passing the image to the simulation. For each cycle of the simulation, we generated 20 artificial sources for each field. The artificial sources were randomly assigned flux densities such that their SNRs lay between 2–20 σ. Their positions were also randomly assigned, within the 1.8× FWHM area of the primary beam. Next, we removed all close pairs (mutual distance < 3arcsec) of artificial sources to avoid unnatural source blending. After this pruning, the artificial sources were convolved with the synthesised beam and added to the image. This process was repeated $1,000\times$ for each field, to gather a statistically significant sample of simulated sources. We then repeat all the steps for different source sizes. We model the shape of the injected sources with 2-D Gaussians, with the intrinsic size varying across a grid of five different sizes (point sources with FWHM = 0.1, 0.2, 0.3, 0.6 arcsec). Finally, we searched these images for detections in the same way as we do for our sources (see §2.3.3 and 2.3.2), using the same methods to measure their flux densities.

We also tested the injection of artificial sources into the *uv* visibilities. Similar to the image-plane simulation, we also randomly generate 20 sources per cycle per field. Then, ALMA task FT was used to convert the point sources into visibilities and ALMA task UVSUB is used to add them into the visibilities. The two simulations give consistent results, but conducting simulations in the visibilities is time-consuming, especially considering that we need simulations for different fields, so we chose to conduct all the simulations in the image plane for all the ALMACAL fields.

Fig. 2.5 shows the effects of flux boosting as a function of SNR. We only show the results for our fiducial source size of 0.2 arcsec. The flux boosting also varies for different source sizes, but the flux boosting converges to a single boosting function above $\mathrm{SNR}_{\mathrm{peak}} \geq 5$ (see more dicussion below). Moreover, size measurements are problematic at low SNRs, which complicates any size-based correction. As a compromise, for simplicity, we focus on the flux boosting for sources with a fixed size (FWHM=0.2 arcsec) in our analysis.

The flux boosting effect is slightly different for different flux-measuring methods. We found a smaller boosting factor for aperture photometry than with the 2-D Gaussian fitting. For aperture photometry, we did not find noticeable boosting for bright sources (SNR > 5σ); while for 2-D Gaussian fitting, the boosting effects are still not negligible for sources with SNR up to 10. This boosting factor is consistent with Oteo et al. (2016a), whose results were based on two-dimensional Gaussian fitting with CASA task, IMFIT. The key difference between aperture photometry and Gaussian fitting, as deployed here, is that the latter allows the centroid to move slightly (±0.1 arcsec) and has the freedom to vary the source shape, such that it will suffer more from Eddington flux boosting. However, Gaussian fitting is more robust than aperture photometry in crowded regions. In our final catalogue, both of the flux densities are reported and corrected by their own deboosting function. In the following analysis we use the flux density from aperture photometry, but changing to Gaussian fitted flux density does not change our results significantly.

We also tested flux boosting for sources with different sizes. We use the same simulation as described in §2.3.4 to quantify the flux boosting of different source sizes. Each time we inject single-size sources and measure their boosting effects. In Fig. 2.6, it shows the boosting effects from different methods. Firstly, sources with different sizes have slightly different boosting factors, especially at low SNRs. Secondly, despite the very extended sources (FWHM~0.6 arcsec), all the boosting factors converged at SNR > 5. Since we only account for sources with SNR > 5, it is reasonable to correct the boosting effects based on the simulation of our fiducial source size (FWMH=0.2 arcsec).

2.3.5 Completeness & reliability

We use the same simulation as discussed in §2.3.4 to derive the completeness correction for our catalogue. We apply the source-finding procedures to these simulated images, using the methods described in §2.3.2. An injected source was marked as recovered if it was matched by a detection within the synthesised beam, and was otherwise marked as missed. If a source was detected without a corresponding injected source, it was marked as a false detection.

As shown in Fig. 2.7, the resulting completeness varies for different source sizes. Size-based corrections have been discussed in the literature (e.g. Béthermin et al., 2020), but the intrinsic source size cannot be well constrained for sources with SNR < 10 (Simpson et al., 2015b). We therefore adopt the completeness correction derived from our fiducial source size (FWHM=0.2 arcsec). Besides a robust size measurement, the intrinsic size distribution of DSFGs is needed to fully correct the completeness, which requires future observational effort (e.g. Gullberg et al., 2019; Smail et al., 2021). We did not find many extended sources in the existing surveys with sensitivities and resolutions similar to ALMACAL. Therefore, we stick to the fiducial completeness correction but caution that this as a limitation of our analysis.

With a 5σ detection threshold, we found no spurious detections in our simulations. The spurious fraction goes down to zero around SNR ~ 4.7 in most of our fields, hence the adopted 5σ threshold will give us a clean and robust sample.

2.3.6 Effective area

Unlike surveys made via mosaics of uniform ALMA pointings, the effective area of ALMACAL changes with sensitivity. Firstly, different fields have different total integration times, which leads

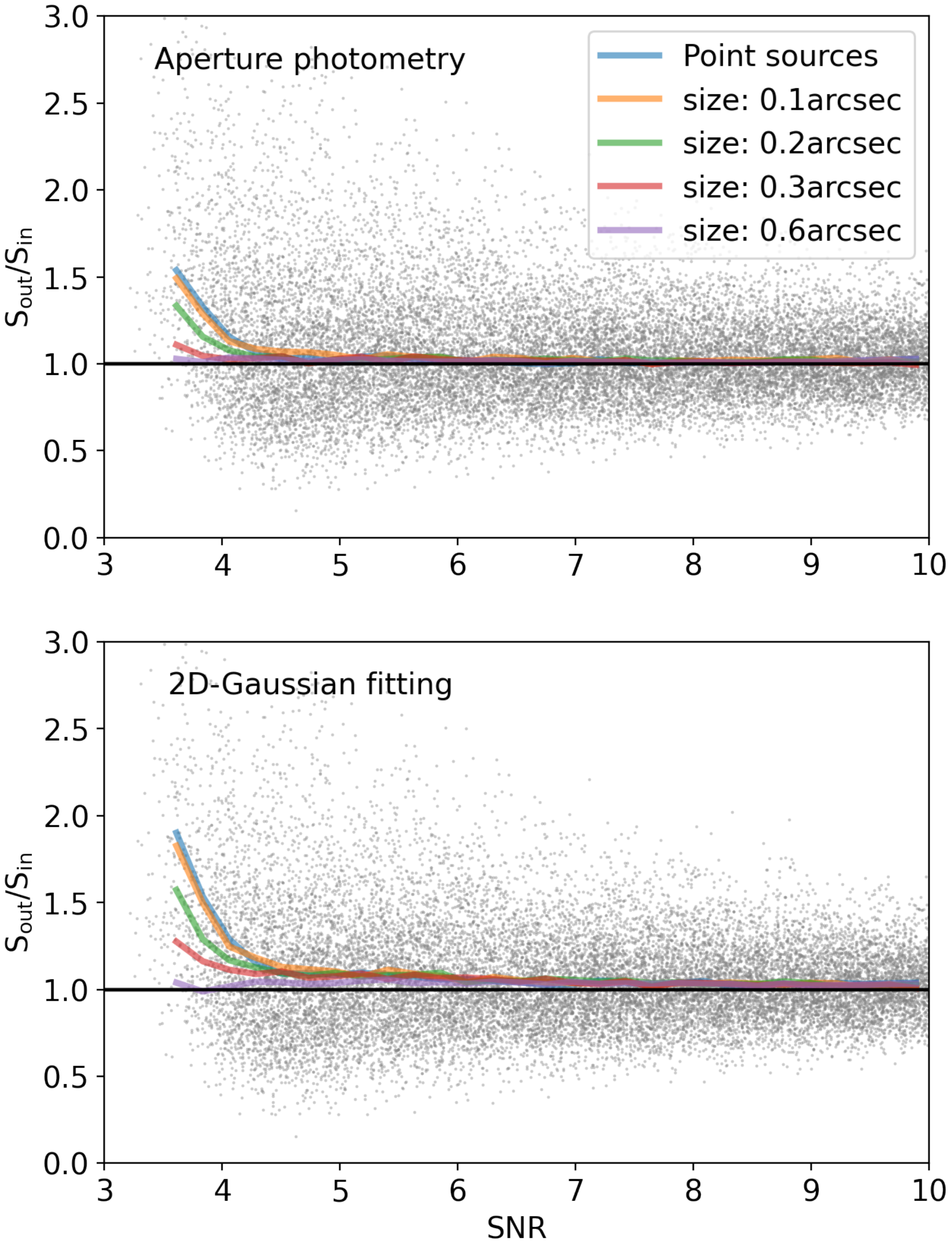

Fig. 2.6: The comparison of flux boosting for different extended sources. The labels are same as Fig. 2.7. Sources with different sizes have different boosting factors but converge at SNR > 5.

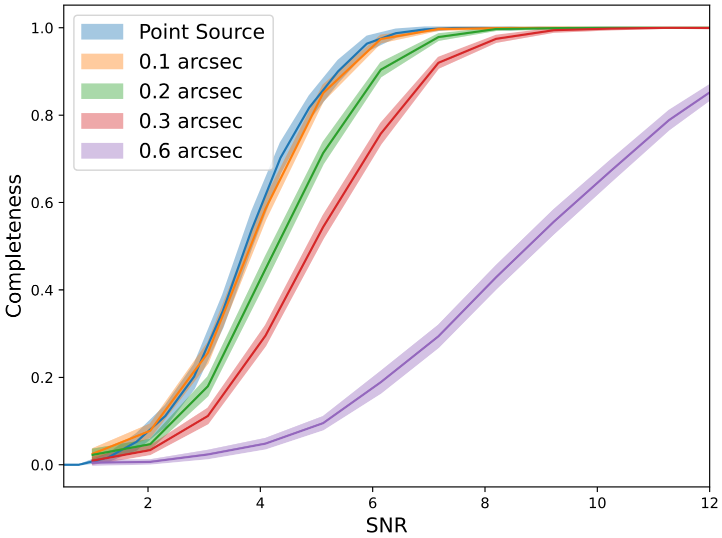

Fig. 2.7: Sample completeness as a function of peak SNR for various source sizes (in FWHM). The larger the source size, the lower the completeness at a given SNR. However, source size with SNR < 10 cannot be measured robustly. Based on size distribution of DSFGs from current surveys at similar sensitivity, we adopted 0.2″ as the fiducial source size to correct the sample completeness. We caution that our survey is less sensitive to very extended sources (FWHM > 0.6 arcsec).

to different sensitivities. On top of this, the sensitivity decreases with increasing distance from the phase centre, as per the primary beam response. To quantify this effect, we measured the effective area as a function of peak flux density. For every field, the effective area for a particular flux density is dictated by the radius at which its SNR drops to 5σ. The maximum effective radius is limited by the 1.8× FWHM for source detection. Both the SNR threshold and the maximum radius are the same as those used for source detection. We performed the same calculation for all of the available fields to determine the total effective area at different peak flux densities.

Fig. 2.8 shows the total effective area as a function of peak flux density for the different ALMA bands. Due to the larger FoV at longer wavelengths, and its popularity with ALMA users, ALMA band 3 has the largest effective area, close to 700 arcmin^2 at a flux density of 0.1 mJy. It is followed by bands 4 and 6. The ALMA FoV in band 6 is much smaller than in band 4, but band 6 has more observations – 39 per cent of all the ALMACAL observing time. Bands 5 and 7 have similar effective areas at flux densities larger than 0.3 mJy, but band 7 goes much deeper. Besides ALMACAL, we also plot the sky coverage of other pointed and blank-field surveys (González-López et al., 2020; Franco et al., 2018; Stach et al., 2019; Gmez-Guijarro et al., 2022). By comparison, ALMACAL offers a good balance between sensitivity and effective area, and will continue to be an essential complement to blind surveys.

Klitsch et al. (2020) estimated the cosmic variance of band 8 footprints from ALMACAL following the methodology described in Driver & Robotham (2010). They found the cosmic variance is less than 5 per cent level in ALMACAL band 8 observations. All the bands with

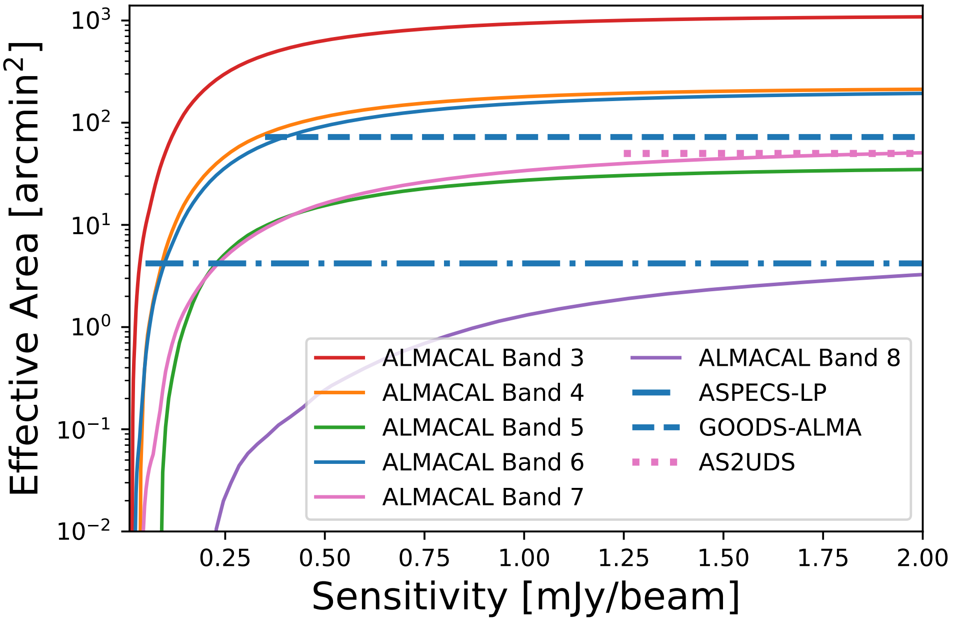

Fig. 2.8: Effective area as a function of sensitivity (5σ) for different ALMA bands. The sensitivity reached in each field is different, and decreases with increasing distance to the phase centre within each pointing. These effects lead to the final effective area changing with sensitivity. In addition to ALMACAL, we also show the sensitivities and effective areas of ASPECS (González-López et al., 2020) and GOODS-ALMA (Gmez-Guijarro et al., 2022) conducted in ALMA band 6, and the AS2UDS survey (Stach et al., 2019) in ALMA band 7. Compared with existing dedicated ALMA surveys, ALMACAL offers a good balance between sensitivity and effective area and will continue to be an essential complement to existing or ongoing blind surveys.

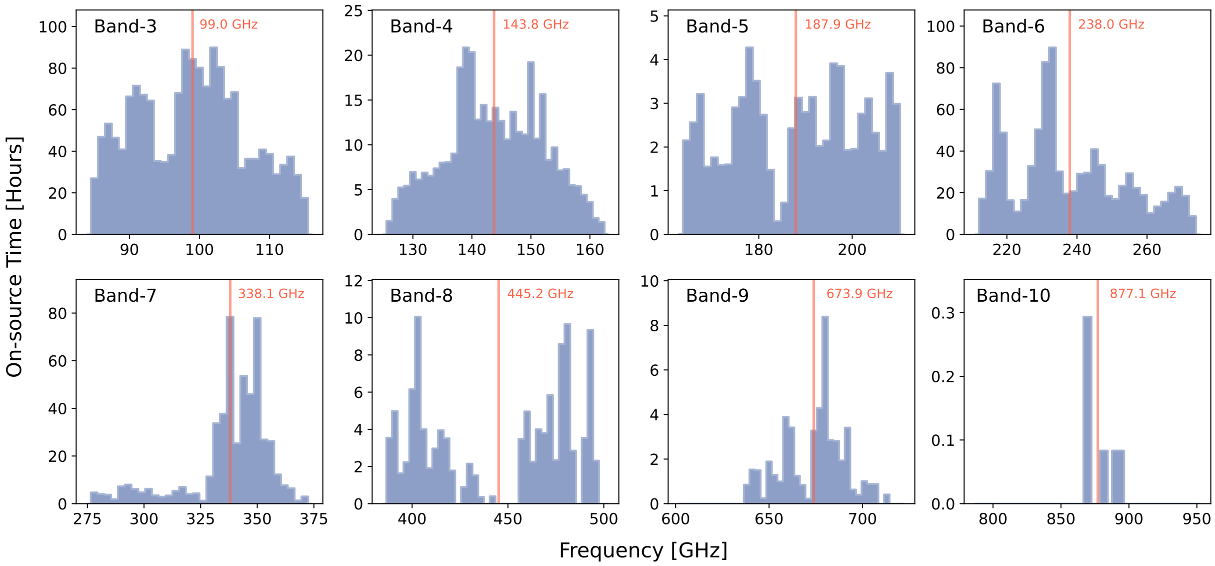

Fig. 2.9: ALMACAL frequency coverage in the different ALMA bands. The abscissa is the frequency coverage in each ALMA band. The ordinate is the total integration time for the selected observations at each frequency. For each band, the vertical red line is the time-weighted mean frequency. In contrast to existing ALMA blind surveys, the frequency coverage of ALMACAL is much wider. To make meaningful comparisons with literature results, number counts have been re-scaled to the characteristic frequencies based on the SED of a modified blackbody (see §2.4.2 for more detail).

Tab. 2.1: Statistical summary of ALMACAL detections.

ALMA Band	Ref. λ (mm)	Sky coverage (arcmin^2)	On-source time (h)	$\langle\nu\rangle$ (GHz)
3	3.00	817	250.6	99.0
4	2.00	157	53.1	143.75
5	1.50	27	20.8	200
6	1.20	149	344.7	250
7	0.87	45	275.2	345
8	0.65	5	52.7	460
9	0.45	1	20.9	666
10	0.35	0.1	0.5	857
total	–	1201.1	1018.5	–

longer wavelengths have at least 10× larger sky coverage, which should suffer from less than 1 per cent cosmic variance.

2.3.7 Effective wavelength

The spectral coverage of ALMACAL is much more complex than that of previous ALMA surveys where the observations were typically carried out at a fixed frequency. ALMA calibrators share the same configurations as their science targets, such that the same calibrator observed by different projects will potentially have different spectral configurations. Fig. 2.9 shows the frequency coverage of ALMACAL from ALMA band 3 to 10. The vertical lines are the exposure-time-averaged mean frequency for each band. For example, the observations in band 6 are spread across the whole band. The exposure-weighted mean frequency in band 6 is 237.95 GHz (1.26 mm), which is slightly lower than the most commonly used 250 GHz (1.2 mm). By comparison, the coverage in band 7 is simpler, with most of the observations undertaken at around 340 GHz. Its mean frequency is 338.08 GHz (887 μm). The effects of these differences need to be corrected before a comparison is made with literature results, as discussed in §2.4.2.

2.3.8 Source classification

Sources detected by ALMACAL are typically either thermal in nature, e.g. continuum emission from dust in star-forming galaxies, or non-thermal in nature, e.g. synchrotron emission from radio jets associated with the calibrator or other radio sources in the field. Thanks to the multi-band coverage of ALMACAL, these emission mechanisms can be separated by their spectral index, α, where $S_\nu \propto \nu^\alpha$. For a typical DSFG at $z \sim 2$, we are probing the Rayleigh-Jeans tail of emission from warm dust with $\alpha \approx +3.5$ (Ivison et al., 2010c; Swinbank et al., 2010), so the emission becomes brighter at higher frequencies. For radio AGN and jets, on the other hand, the flux density typically declines at higher frequencies or stays flat, such that $\alpha \lesssim 0$. We

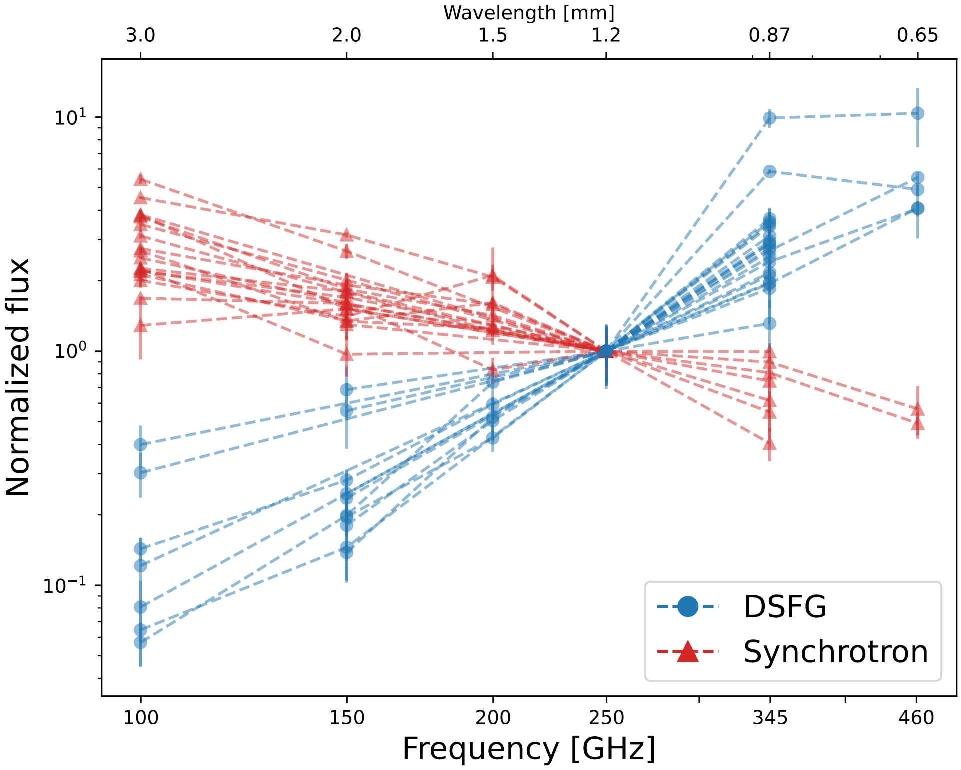

Fig. 2.10: The submm/mm SED of a few confirmed DSFGs and Synchrotron sources. Only the sources detected in band 6 plus at least two more bands are shown here. All the SEDs have been normalized by their band 6 flux densities. We see that DSFGs have positive spectral indices with their fluxes rising at higher frequencies, while the Synchrotron sources typically have negative spectral indices with their fluxes decreasing at higher frequencies.

Tab. 2.2: Classification ALMACAL detections.

ALMA Band	Number of detections	DSFGs	Synchrotron	Unclassified sources
3	63	8	44	12
4	54	20	23	11
5	21	10	9	2
6	228	132	54	42
7	130	93	22	15
8	17	13	3	1
9	2	2	0	0
10	0	0	0	0
total	371	186	102	83

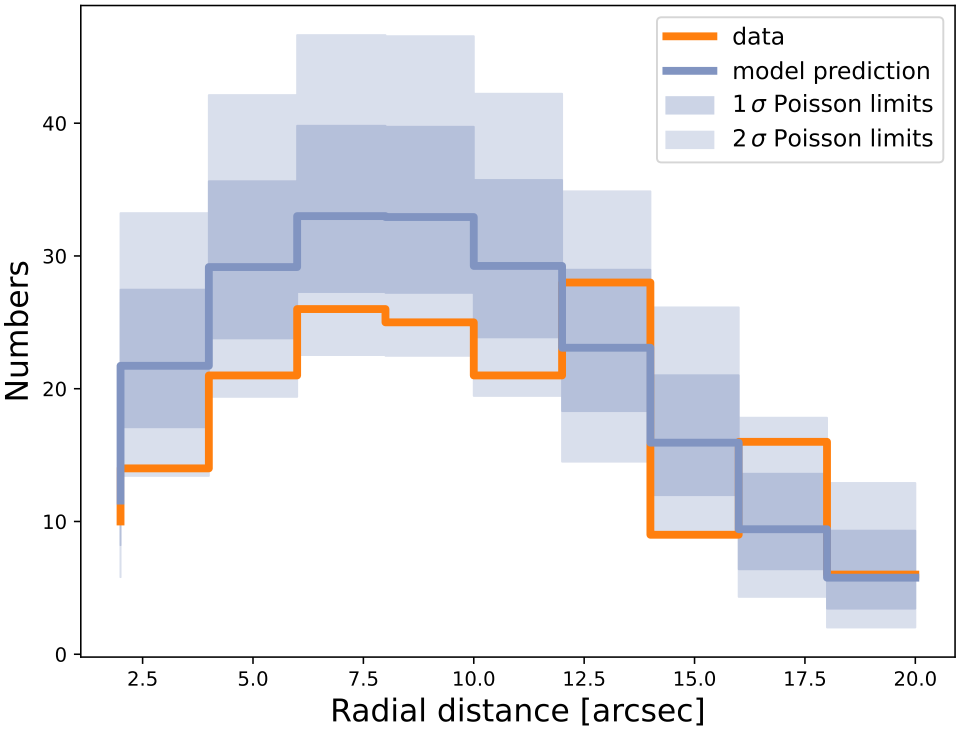

Fig. 2.11: Radial distribution of DSFGs detected in band 6 in the ALMACAL survey. The abscissa is the radial distance of the detections from their field centre. The orange histogram is the radial distribution of the confirmed DSFGs. The blue histogram is the predicted DSFGs distribution based on the expected number counts from ASPECS González-López et al. (2020). The blue and light blue shadows are the 1σ and 2σ Poisson variations of the model predictions. Small differences between ALMACAL and ASPECS are expected. However, the radial distribution of our DSFGs is generally consistent with the expected of randomly distributed DSFGs, which indicates that the thermal sources are not clustered around the blazars.

categorised a source as a DSFG if its multi-band flux densities are consistent with a dust SED, and as synchrotron source if its spectral index instead betrays synchrotron emission (see also Klitsch et al., 2020).

We show the SED from examples of confirmed sources in Fig. 2.10. To obtain the dust SEDs of sources covered by different ALMA bands, we extend the flux measurement to all the available bands whenever a detection is confirmed. The aperture size and aperture correction are calculated using the same way we do for true detections. We show all the measurements that have reliable flux densities (SNR ≥ 3) in at least three ALMA bands in Fig. 2.10. In general, DSFGs have distinguishable SEDs from radio sources. The median α in our sample is $\alpha = 3.2 \pm 0.3$ and $\alpha = -0.9 \pm 0.1$ for DSFGs and Synchrotron, respectively.

We commonly encounter the radio jets of our blazar calibrators, which are not always perfectly aligned with the line of sight. Most are relatively compact, beginning close to the calibrator; they tend to be much more prominent in bands 3 and 4 and are easily identified by their positions and elongated morphologies. However, extended radio jets have been found in about 10 per cent of the fields. These jets could be resolved into several blobs and morphologically they resemble point sources at higher frequencies. With detections in multiple bands, these jets can be easily excluded via their spectral indices and their morphological alignment (see the second example in Fig. 2.4).

The classification of an ALMACAL source is uncertain if the source is only covered by one ALMA band. In such cases, we first we cross-matched with radio images from the Monitoring Of Jets in Active galactic nuclei with VLBA Experiments (MOJAVE)[4], Faint Images of the Radio Sky at Twenty-cm (FIRST – Becker et al., 1995), the NRAO VLA Sky Survey (NVSS – Condon et al., 1998) and the NRAO VLA Archive Survey (NVAS)[5]. If the source has been detected by the radio surveys given their typical depth, it is classified as synchrotron. The additional radio images are crucial for sources found in band 3 due to the generally wider FoV and higher sensitivity. Examples of our synergetic classification, combining data from ALMA and the VLA, are illustrated in Fig. 2.12. In band 3, we are able to confirm 14 per cent more of the total sources with the help of radio surveys.

For those unclassified sources lacking radio coverage, because of their small numbers (< 20%), we did not find a large difference in number counts by including these uncertain sources, so in the main context we focus on reliable SMGs that are confirmed by at least two ALMA bands or one ALMA band with the radio images available. Details of the flux density distribution, radial distance distribution and the contribution to the number counts of the unclassified sources can be found in Fig. 2.14 and Fig. 2.20. Source classifications are available from our on-line version of the catalogue.

The radial distance and flux distribution of all the detections can offer additional information about the accuracy of the source classification. In Fig. 2.14, we show the histograms of band 6 sources, where the number of detections is largest, in different bins of flux and radial distance. All three categories share a similar trend in different flux bins, but they are different in the distribution of radial distance. Most of the radio sources are quite close to the calibrator (<5 arcsec), which

---

[4] https://www.physics.purdue.edu/astro/MOJAVE/

[5] https://www.vla.nrao.edu/astro/nvas/

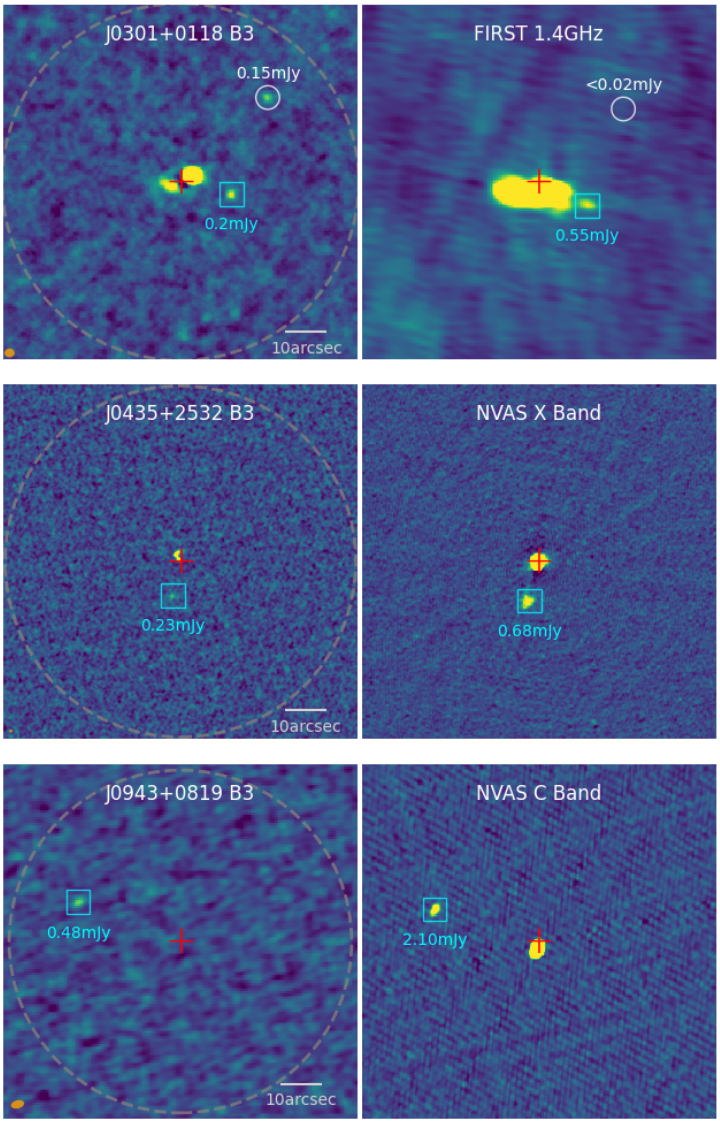

Fig. 2.12: Examples of confirmed radio sources from VLA archive images. The first column shows the band 3 images from ALMACAL. The second column shows the corresponding radio images from VLA archive and they have been cropped to the same scale as the ALMACAL images. The dashed grey circle is the FoV adopted in this work, which is 1.8×FWHM of the respective primary beam. SMG is marked with white circle and synchrotron emission is marked with cyan square. Radio images are useful to constrain the emission mechanisms of ALMACAL detections without multi-bands observations.

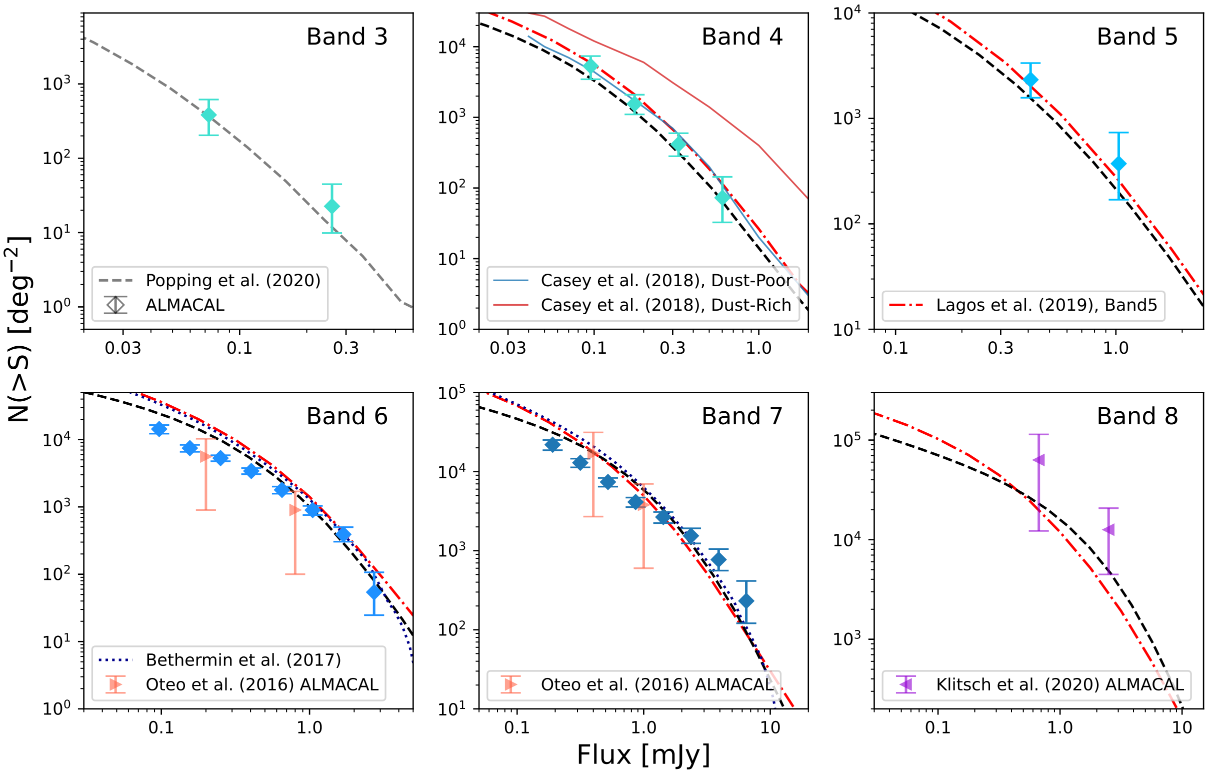

Fig. 2.13: Multi-band number counts from ALMACAL. The number counts reported in this work are displayed in each panel. In bands 6 and 7, we also show the first ALMACAL number counts by Oteo et al. (2016a); in band 8, we adopt the results from Klitsch et al. (2020). We also show model predictions of the number counts at various wavelengths, including the semi-analytical model from Lagos et al. (2020) and the semi-imperical models from Popping et al. (2020) and Bthermin et al. (2017). A successful model should be able to explain the number counts at the different wavelengths simultaneously. The most recent semi-analytical and semi-empirical models are broadly consistent with the number counts of ALMACAL at bright fluxes, but show apparent differences at fainter limits in band 6 and 7 (see also Fig. 2.15 and Fig. 2.16), which need to be confirmed by the future surveys. Therefore, multi-wavelength number counts are a powerful tool to validate those models.

should be the natural outcome if the majority of them are blobs of radio jets. Similar to our prediction in 2.3.9, we have more DSFGs at the median radial distance. The unclassified sources have a much flatter distribution, which should include both RSs and DSFGs. However, their relatively small number prohibits a quantitive analysis. In the next section, we will randomly sample these unclassified sources to test their influence on the final number counts.

2.3.9 Selection bias

ALMACAL is not a truly blind survey. The pre-selection of the calibrator fields can bias our survey if the central calibrator has influenced the detection of nearby sources. Since roughly 97 per cent of the calibrators are blazars, the blazar activity may have been triggered by an interaction, perhaps also giving rise to a nearby starburst. In addition, the blazar host galaxy may act as a gravitational lens. However, a blazar is bright mainly due to the beaming effect, which does not always traces the most massive galaxies and dark matter halos. The redshift distribution of the ALMACAL blazars peaks at $0.5 < z < 1.0$ (Bonato et al., 2018), well short of the typical redshifts to DSFGs (Brisbin et al., 2017; Aravena et al., 2020; Dudzevičiūtė et al., 2020). Moreover, we note that a systematic search for galaxy over-densities around three distant quasars with known companions failed to find associated DSFGs (Meyer et al., 2022).

To further quantify the possible biases that might be introduced by the central bright sources, in Fig. 2.11 we compare the radial distribution of all our DSFGs with the prediction of the ALMA blind survey, ASPECS. We used the best-fitting number counts from ASPECS as the count model and then created a mock survey in the manner of ALMACAL. We show the radial distribution from band 6 where the number of detections is the highest. If the same density of sources detected by ASPECS is recovered here, scattered randomly throughout our fields, we would have unbiased DSFG positions, regardless of the central blazar. We made mock observations by using the same observational set-ups as ALMACAL, covering the same effective area at different flux densities, and the uncertainty in the radial distribution is given by the Poisson error. Fig. 2.11 shows the $1\,\sigma$ and $2\,\sigma$ Poisson confidence levels indicating that the radial distribution of our detections is broadly consistent with ASPECS. In particular, we do not see any excess of detections near the central calibrators, indicating there is little evidence for clustering of the DSFGs around the blazars (see also §**??** and Fig. 2.19 for the difference in number counts between inner and outer parts of all the fields.).

2.4 Results

In this section, we present the source catalogue and the number counts at different wavelengths.

2.4.1 Source catalogue

Our classification of sources is carried out by the procedures presented in §2.3.8 and the numbers of different detections are summarised in Table 2.1. In total, we have detected 371 sources, including 186 secure DSFGs with multi-band confirmations of their purity and spectral indices.

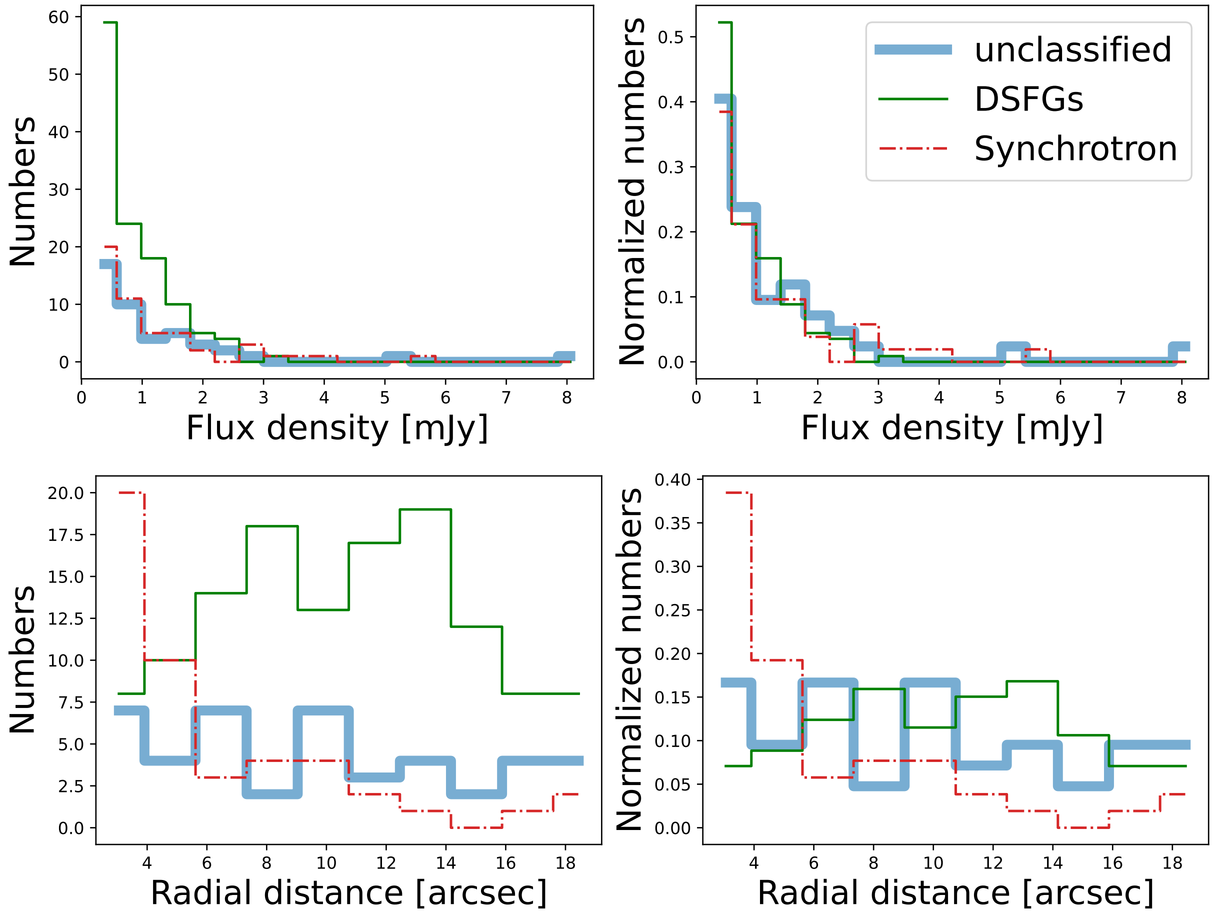

Fig. 2.14: Histograms of all the detections in flux bins and radial distance bins from band 6. The first row shows the number and the normalized number of detections that change with the flux density. We do not find an evident difference in the flux density distribution of the three classifications. The second row shows the number and the normalized number of detections change with their radial distance to the central calibrator. Synchrotron sources concentrate near the calibrator, which is consistent with the nature of radio jets associated with the blazar. Compared with Synchrotron, DSFGs have a flatter distribution as a function of radius. Due to the small number of unclassified sources, we can not classify them by their radial distance, but including them and excluding them will not change our main result.

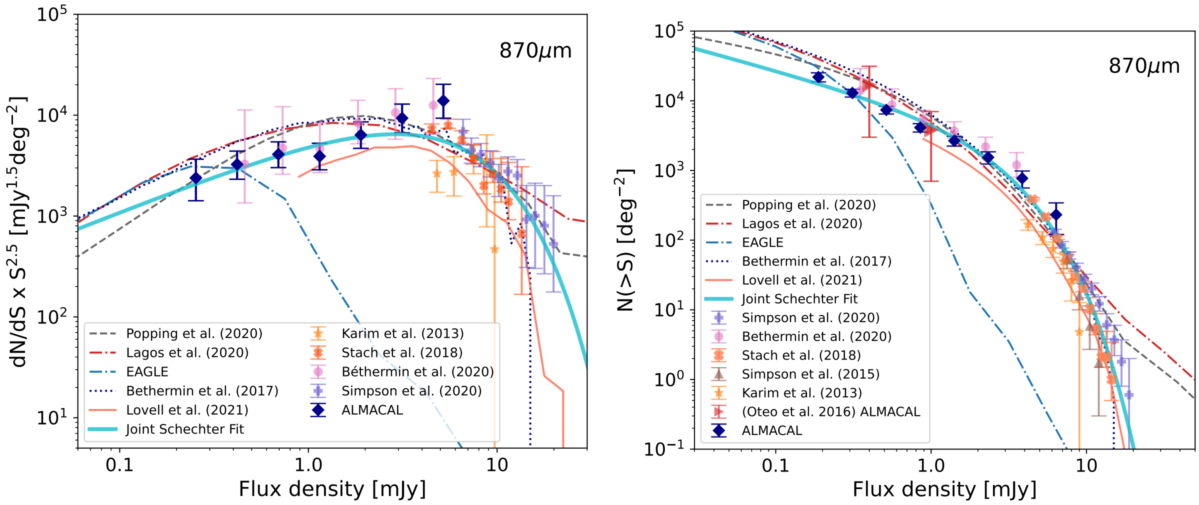

Fig. 2.15: Differential number counts (left) and cumulative number counts (right) at 870 μm in band 7. The differential counts have been normalised by $S^{2.5}$ to reduce the dynamic range. We show the new ALMACAL results and the previous ALMACAL band 7 number counts from Oteo et al. (2016a). We also plot the number counts from surveys that used ALMA (Karim et al., 2013; Simpson et al., 2015a; Stach et al., 2018; Béthermin et al., 2020; Simpson et al., 2020). Model predictions from Fig. 2.13 are included here, with the same line styles. The predictions based on EAGLE (Camps et al., 2018; McAlpine et al., 2019) and SIMBA (Lovell et al., 2021) cosmological hydrodynamic simulations are also included as a comparison. The radiative transfer post-processing is only applied to SIMBA galaxies with SFR$>$ 20 M$_\odot$ yr^{-1}, which can only give complete number counts above $S_{870\mu m}$ >1mJy. The number counts from ALMACAL are consistent with previous results in a wide flux density range and represent the deepest survey available at the moment. The best joint Schechter fits are shown in each plot.

The largest number of detections is in ALMA band 6, where we have detected 228 sources, including 132 DSFGs. This is followed by band 7, where 93 DSFGs from 130 detections have been confirmed. The detection rate in band 4 is also promising, with 20 secure DSFGs detected. Band 5 has ten DSFGs, which enables us to derive the number counts at 1.5 mm for the first time. In band 3, sky coverage is significantly larger than higher frequency bands and they are typically deeper, which makes the radio images much more important to classify the detections. Luckily, around half of the ALMACAL band 3 footprints have been covered by VLA archival surveys. Combining ALMACAL with the VLA archive, we have confirmed eight DSFGs in band 3. Band 9 has only two detections found in 1 arcmin^2 – too few to meaningfully constrain the number counts. For ALMA band 8, we adopt the results of Klitsch et al. (2020), which already benefited from the updated version of ALMACAL.

As we can see here, the number of detections provided by ALMACAL is competitive with dedicated ALMA cosmological surveys, offering a complementary way to refine the counts. We also report the detected numbers of synchrotron sources in Table 2.1. In this paper, we mainly focus on the number counts of DSFGs.

2.4.2 DSFG number counts

One of the major goals of ALMACAL is to constrain the number counts of DSFGs. The multi-band coverage of the calibrators makes it possible to constrain the number counts at multiple wavelengthes.

We derive the number counts following the formula used in Oteo et al. (2016a). For every individual detection, its contribution to the total cumulative number counts is:

$$N_i(S_i) = \frac{1 - f_{\mathrm{sp}(S_i)}}{C(S_i) \cdot A(S_i)}, \tag{2.1}$$

where S_i is the flux density of the detection; $f_{sp}(S_i)$ is the fraction of spurious sources at S_i, which is equal to unity with our 5σ detection threshold; $A(S_i)$ and $C(S_i)$ are the effective area and completeness at S_i, respectively.

The final number counts are calculated in flux density bins, S_j, evenly spaced in log scale. For each bin, the cumulative number counts is defined as:

$$N_j(> S_j) = \sum_i^{n_i} N_i(S_i), \tag{2.2}$$

where n_i is the number of N_i that satisfies $S_i \leq S_j$. The differential number counts for each flux density bin is given by:

$$\frac{dN_j}{dS} = \frac{\sum_i^{n_{ij}} N_i}{S_{j+1} - S_j}, \tag{2.3}$$

where n_{ij} is the number of N_i that meets $S_j < S_i \leq S_{j+1}$.

The uncertainties in the number counts are calculated via Monte-Carlo simulations. Firstly, the flux density of every detection was randomly sampled according to its 1σ uncertainty. Then, the whole DSFG sample is grouped into the different flux density bins, S_j. The flux density bins are constructed to include at least three detections. Their cumulative and differential number counts are calculated using Equations 2.2 and 2.3, respectively. We ran this simulation $1,000$ times to determine the scatter of the number counts. Finally, Poisson errors were added to the uncertainties of number count according to the number of detections in each bin. We discuss the number counts that include the unclassified sources in §2.20.

To compare with literature results and theoretical models, we re-scaled the flux densities of all our detections to reference wavelengths in each ALMA band. The reference wavelengths used in this work are summarised in Table 2.2. The re-scaling factor is determined by the flux density ratio at different wavelengths assuming the composite SED from the AS2UDS survey at a reference redshift of $z = 2.0$ (Dudzevičiūtė et al., 2020), which also has a power-law index $\alpha \sim 3.2$. During the scaling, we add 20 per cent of error to the number counts to reflect the variation of the composite SED (Dudzevičiūtė et al., 2020). The wavelength re-scaling does not provide a perfect correction for every detection, because each will have a different, unknown redshift, but since the correction is small (~ 10 per cent), it should be reasonable in a statistical sense. The number counts before and after re-scaling are displayed in Fig. 2.20.

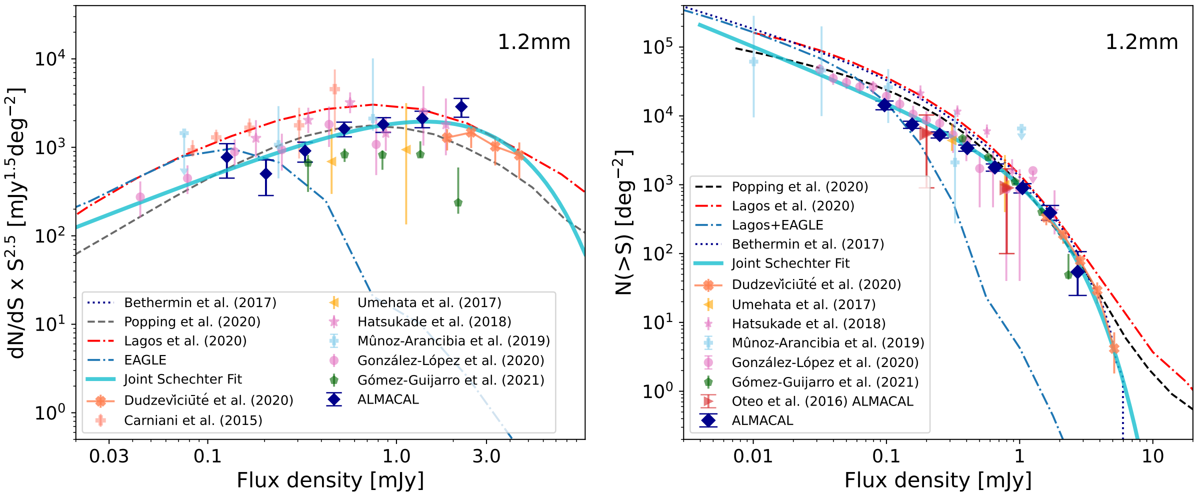

Fig. 2.16: Differential number counts (left) and cumulative number counts (right) at 1.2 mm. We show the new ALMACAL results and the previous ALMACAL band 6 number counts from Oteo et al. (2016a). Interferometric results reported by Umehata et al. (2018); Hatsukade et al. (2018); Muñoz Arancibia et al. (2018) are also shown. From Umehata et al. (2018), only the results from the field are shown. From Muñoz Arancibia et al. (2018), only the combined results are shown (see the updated results from the corrigendum Muoz Arancibia et al., 2019). We also show the predicted SED-scaled number counts from AS2UDS (Dudzevičiūtė et al., 2020) to provide constraint in the higher flux density range. Model predictions from Fig. 2.13 are also included here. In band 6, ALMACAL overlaps with GOODS-ALMA and empirical AS2UDS prediction at the brighter end and is consistent with ASPECS at the fainter end. The best joint Schechter fits are shown in each plot.

Table 2 reports our cumulative and differential number counts and they are plotted in Fig. 2.13, which shows the number counts in the different ALMA bands. This is the first time that we have number counts calculated consistently across several bands. In the following section, we discuss comparisons with the literature, and with several different model predictions, and the implications of these number counts in the context of the underlying galaxy populations and cosmic infrared background.

2.4.3 Robustness of number counts

We also tested the number counts in the inner and outer parts of the fields. The dark matter halo associated with the central calibrator can also act as a gravitational lens. It can stretch the background DSFGs and produce multiple images, which can potentially boost the number counts in the inner regions. We test this effect in band 7, where the effect is the strongest. We first divide our fields into the inner and outer regions defined by the separation radius r_{sep}. We then calculate the number count in the inner and outer regions separately. If the total area of the inner regions is significantly smaller than the outer regions, we bootstrap the outer regions that sample the

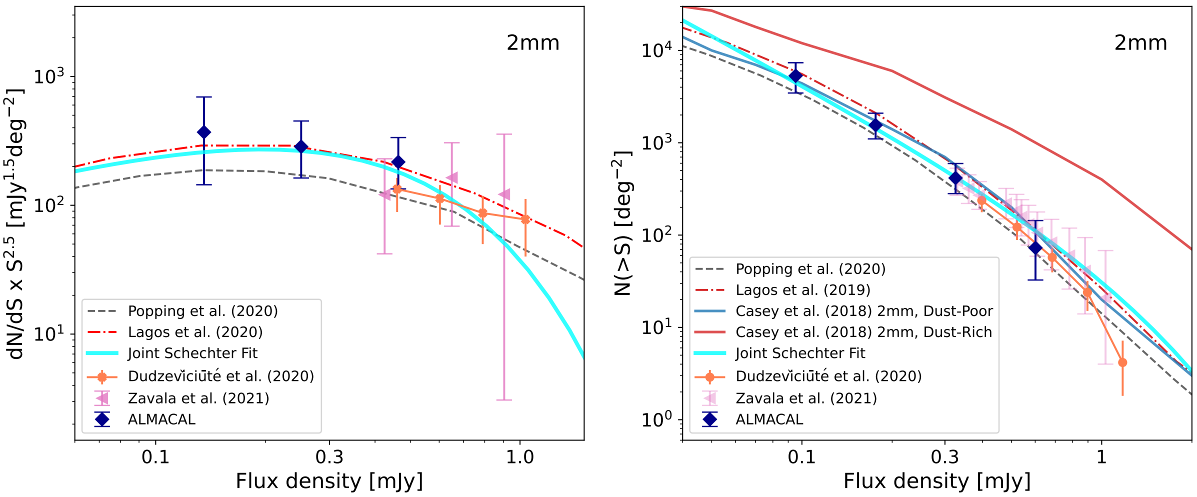

Fig. 2.17: Differential number counts (left) and cumulative number counts (right) at 2 mm. We show the results from ALMACAL and the Mapping Obscuration to Reionisation with ALMA (MORA) survey (Zavala et al., 2021). We also include the SED predicted 2 mm number counts from AS2UDS survey (Dudzevičiūtė et al., 2020). The two different models proposed by (Casey et al., 2018a) are both included. Existing data support the dust-poor model and are consistent with the model predictions from Popping et al. (2020) and Lagos et al. (2020). The best joint Schechter fits are shown in each plot. ALMACAL is consistent with Zavala et al. (2021) and empirical SED-scaled AS2UDS counts at the bright end and present the deepest survey at 2 mm.

Tab. 2.3: Cumulative number counts in various ALMA bands.

Band 3 (3 mm)

S [mJy]	N	N($>S$) [deg^{-2}]	δN($>S$)$_{\rm lower}$	δN($>S$)$_{\rm upper}$
0.07	7	380	170	230
0.26	3	20	10	20

Band 4 (2 mm)

S [mJy]	N	N($>S$) [deg^{-2}]	δN($>S$)$_{\rm lower}$	δN($>S$)$_{\rm upper}$
0.10	21	5310	1850	2040
0.18	17	1550	450	530
0.33	10	410	130	180
0.60	3	70	40	70

Band 5 (1.5 mm)

S [mJy]	N	N($>S$) [deg^{-2}]	δN($>S$)$_{\rm lower}$	δN($>S$)$_{\rm upper}$
0.41	10	2320	750	1020
1.03	3	370	200	360

Band 6 (1.2 mm)

S [mJy]	N	N($>S$) [deg^{-2}]	δN($>S$)$_{\rm lower}$	δN($>S$)$_{\rm upper}$
0.10	132	14350	2090	2090
0.16	125	7470	880	880
0.25	118	5300	530	530
0.40	100	3410	350	350
0.65	70	1780	210	210
1.05	42	890	130	130
1.69	20	390	80	100
2.73	3	50	20	50

Band 7 (870 μm)

S [mJy]	N	N($>S$) [deg^{-2}]	δN($>S$)$_{\rm lower}$	δN($>S$)$_{\rm upper}$
0.19	92	21900	3250	3250
0.32	84	12950	1640	1640
0.52	69	7390	930	930
0.86	53	4120	560	560
1.43	39	2660	420	420
2.36	25	1540	300	370
3.91	13	770	210	270
6.47	4	230	110	180

Tab. 2.4: Differential number counts in various ALMA bands.

Band 3 (3 mm)

S [mJy]	N	dN/dS [mJy^{-1}deg^{-2}]	δ(dN/dS)$_{\text{lower}}$	δ(dN/dS)$_{\text{upper}}$
0.17	4	1990	1130	1690
−−	−−	−−	−−	−−

Band 4 (2 mm)

S [mJy]	N	dN/dS [mJy^{-1}deg^{-2}]	δ(dN/dS)$_{\text{lower}}$	δ(dN/dS)$_{\text{upper}}$
0.14	4	54440	33170	47700
0.25	7	9050	3880	5260
0.46	7	1480	560	810
−−	−−	−−	−−	−−

Band 5 (1.5 mm)

S [mJy]	N	dN/dS [mJy^{-1}deg^{-2}]	δ(dN/dS)$_{\text{lower}}$	δ(dN/dS)$_{\text{upper}}$
0.72	7	4080	1570	2250
−−	−−	−−	−−	−−

Band 6 (1.2 mm)

S [mJy]	N	dN/dS [mJy^{-1}deg^{-2}]	δ(dN/dS)$_{\text{lower}}$	δ(dN/dS)$_{\text{upper}}$
0.13	7	136600	57430	78580
0.20	7	26900	11620	15700
0.33	18	14820	3760	4620
0.53	30	7990	1500	1780
0.85	28	2720	510	620
1.37	22	960	200	250
2.21	17	390	90	120
−−	−−	−−	−−	−−

Band 7 (870 μm)

S [mJy]	N	dN/dS [mJy^{-1}deg^{-2}]	δ(dN/dS)$_{\text{lower}}$	δ(dN/dS)$_{\text{upper}}$
0.25	8	73860	29940	39660
0.42	15	28600	8210	10180
0.69	16	10230	2670	3350
1.15	14	2780	740	960
1.90	14	1280	340	440
3.14	12	530	150	200
5.19	9	220	70	100
−−	−−	−−	−−	−−

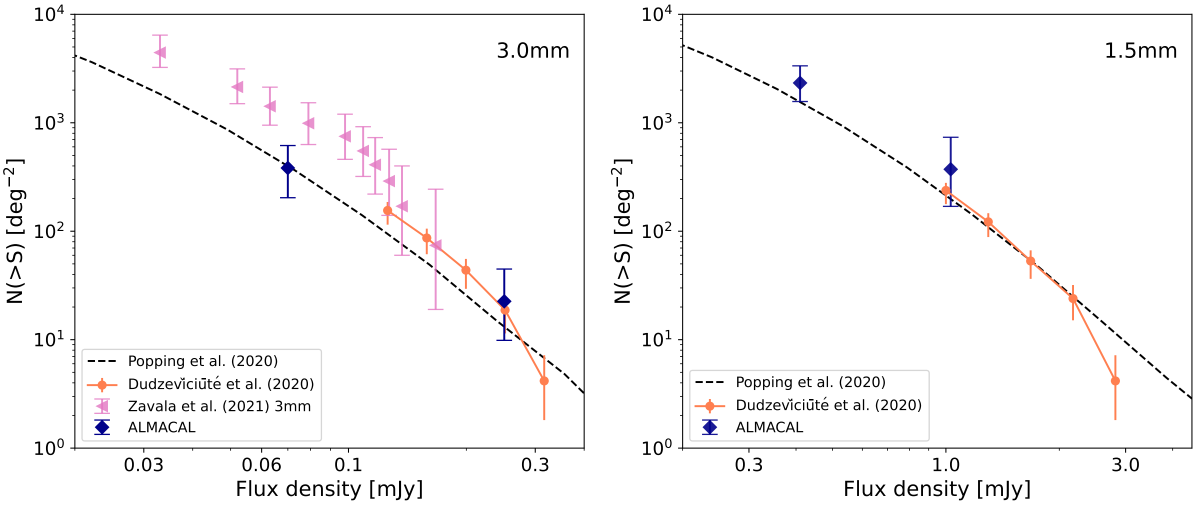

Fig. 2.18: Cumulative number counts at 3 mm and 1.5 mm. We also include semi-empirical model predictions from Popping et al. (2020) and the SED-based prediction from AS2UDS survey (Dudzevičiūtė et al., 2020). At 1.5 mm, ALMACAL is the first survey to constrain the number counts of DSFGs. At 3.0 mm, we also include the measurements from Zavala et al. (2021), which is based on ALMA archival data. Number counts from ALMACAL overlap with (Zavala et al., 2021) and empirical SED-scaled AS2UDS predictions at brighter end, but are systematically lower than Zavala et al. (2021) at the faint end.

same total area as the inner regions. We show the difference between inner and outer regions in Fig. 2.19. We see no statistically significant evidence for excess source counts in the inner 5–8″ regions, as might be caused by either lensing from a halo associated with the calibrator, or from clustering or interactions of the SMGs around the blazar.

To quantify the contribution of these unclassified detections to our final number counts, we calculate their probability of being radio sources or DSFGs based on the overall number of confirmed sources in each band. As we see in §2.14, DSFGs have different radial distribution as radio sources, we thus calculate the probability of being a DSFG at different radial bins. For instance, the probability that a single-band detection in band 6 is a DSFG is equal to the fraction of DSFGs amongst all the classified sources in band 6 at the same radial distance. We choose the radial bins to have at least 10 sources in each bin. We repeat this process 1, 000 times to capture the variation of the final number counts. The comparison of the bootstrapped number counts and the original number counts is in Fig. 2.20. The bootstrapped number counts are slightly larger than the number counts of the confirmed DSFGs sample, but they are still consistent with the original number counts within the uncertainties.

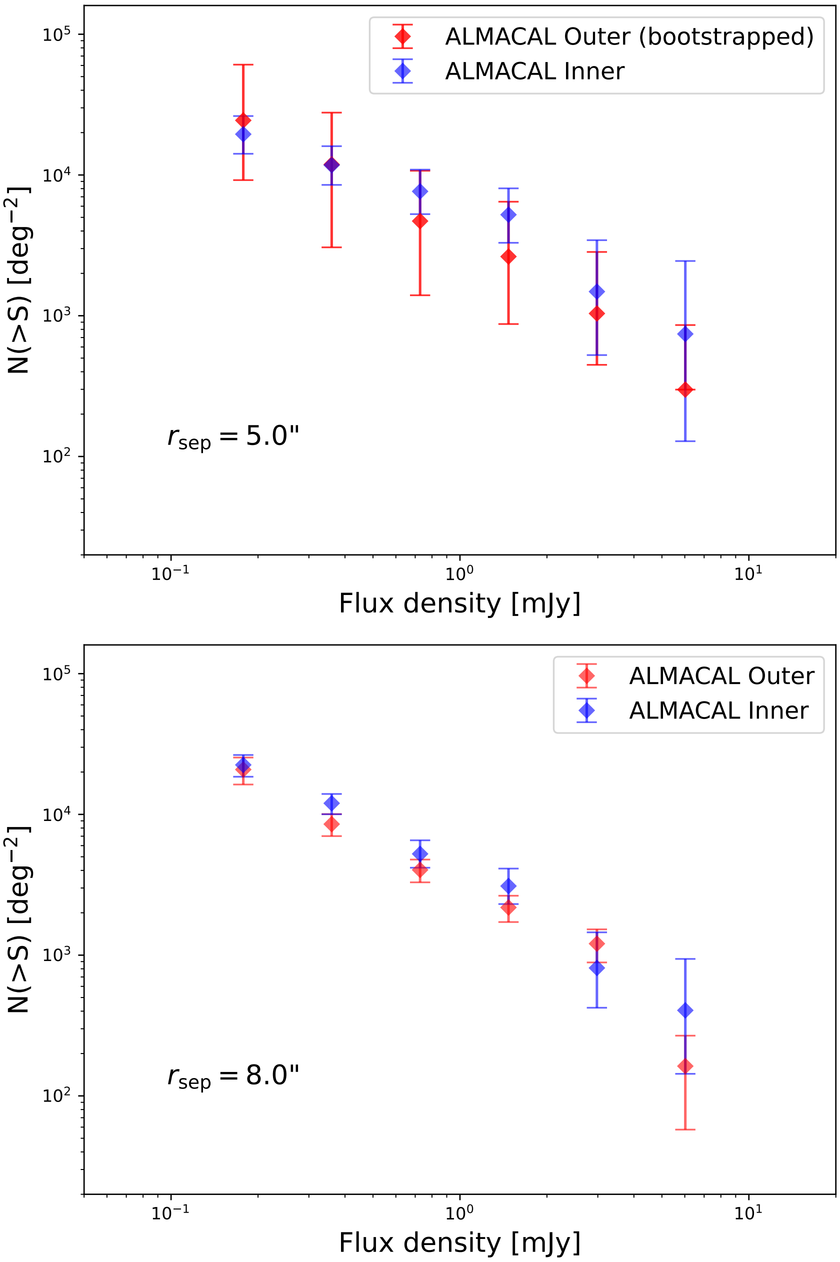

Fig. 2.19: The comparison between the number counts in the inner parts and the outer parts of all the fields. We use all the inner regions but randomly sample the available outer regions to reach the same total sky coverage as the inner ones. The difference in number counts between the inner and outer regions is well within their intrinsic uncertainties.

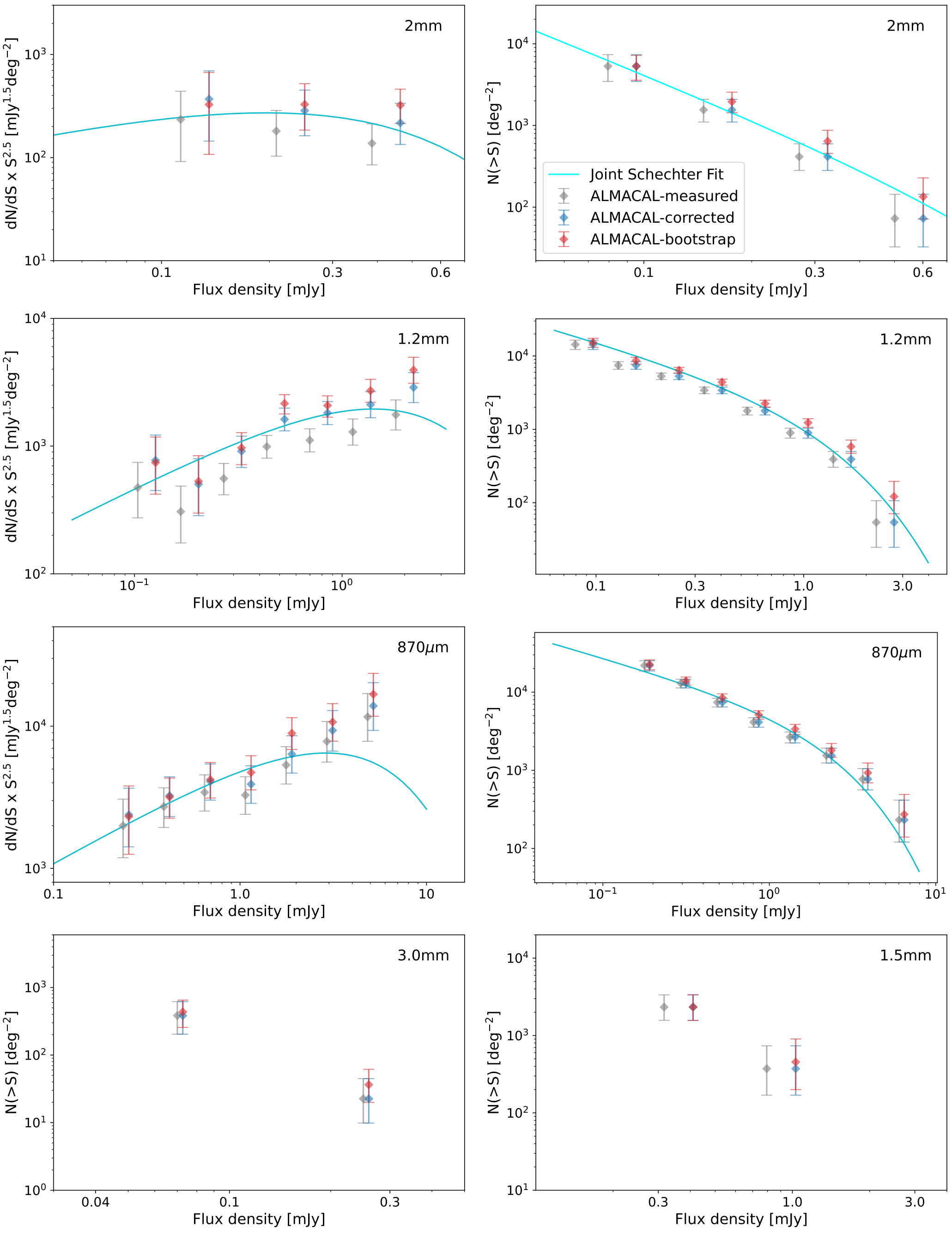

Fig. 2.20: The stability of number counts. We show the final results reported in §2.4.2 along with the results before the effective wavelength correction (see more in §2.3.7) and the bootstrapped number counts in each band (see more in §2.4.3). Bootstrapping is achieved by randomly sampling the unclassified sources into the final DSFG sample based on their radial distribution. The bootstrapped number counts are consistent with the original results within the uncertainties, which indicates our number counts are robust against unclassified sources. We also show the best-fitt Schechter functions reported from Fig. 2.15 to Fig. 2.17.

2.5 Discussion

2.5.1 Comparison with the literature

We compare our number counts reported in Table. 2.3 and Table. 2.4 only with those from other interferometric studies – all of them from ALMA – which are relatively free from source blending. In the following subsections, we will discuss number counts in each band separately, comparing ALMACAL with the literature.

870 μm

Number counts around 870 μm – ALMA band 7 – were the first to be established in the field of submm cosmology, where useful data have been collected since the first single-dish survey (Smail et al., 1997; Hughes et al., 1998; Barger et al., 1998; Blain et al., 1999). In Fig. 2.15 we compare the number counts from ALMACAL with the results from previous ALMA surveys at 870 μm. Karim et al. (2013) published the first ALMA interferometric number counts at 870 μm, following up the bright single-dish sources which had been detected in the Extended *Chandra* Deep Field South using LABOCA on the 12-m APEX telescope (LESS: Wei et al., 2009). After that, Simpson et al. (2015a), Stach et al. (2018) and Simpson et al. (2020) targeted a larger sample of similarly bright SCUBA-2 sources with ALMA. More recently, Béthermin et al. (2020) probed fainter flux densities using serendipitous detections from the ALMA Large Programme to INvestigate [C II] at Early times (ALPINE) survey. Note that this latter survey is expected to be more biased than ALMACAL, since they targeted known high-redshift sources. With ALMACAL we are going even deeper, detecting sources down to 0.2 mJy at the 5σ level, whilst remaining able to detect relatively bright sources due to the large number of fields covered. As Fig. 2.15 show at the faint end, our number counts follow the general trend seen in earlier work and are consistent with the ALPINE survey Béthermin et al. (2020). At the bright end, our results match well with the follow-up ALMA surveys of known, bright submm sources (Karim et al., 2013; Stach et al., 2018; Simpson et al., 2020).

1.2 mm

Number counts around 1.2 mm – ALMA band 6 – have also been well studied for many years. It started with single-dish surveys, such as the early single-dish surveys at the IRAM 30-m telescope using the Max-Planck Millimeter Bolometer array (MAMBO) (e.g. Greve et al., 2004), the Bolocam at the 10-m Caltech Submillimeter Observatory (e.g. Laurent et al., 2005), and AzTEC at the 15-m James Clerk Maxwell Telescope and the 10-m Atacama Submillimetre Telescope Experiment (ASTE) (e.g. Scott et al., 2012). The first deep interferometric observations were made by Hatsukade et al. (2013) towards the Subaru/*XMM-Newton* Deep Survey field. Since then, band 6 number counts have been extended by mining existing deep fields in the ALMA archive (Ono et al., 2014; Carniani et al., 2015; Fujimoto et al., 2016). More recently, dedicated ALMA blind surveys, ASPECS (González-López et al., 2019, 2020) and the ALMA Frontier Fields survey (Gonzlez-Lpez et al., 2017; Muñoz Arancibia et al., 2018) have greatly improved the depth, while GOODS-ALMA (Franco et al., 2018; Gmez-Guijarro et al., 2022) has increased the sky coverage substantially, up to 72-arcmin^2. All of these surveys are plotted in Fig. 2.16. ALMACAL offers a good balance between survey depth and sky coverage. In the flux density range between 0.1 mJy and 1.0 mJy, most of the number counts are consistent within 2σ. At the brighter end, ALMACAL is consistent with GOODS-ALMA and represents the widest survey available.

At 1.2 mm, discrepancies have been reported in number counts both at the fainter end ($<$ 0.1 mJy) and at the brighter end ($>$ 1.0 mJy). At the fainter end, Fujimoto et al. (2016) reported number counts that appear systematically high relative to other surveys, possibly due to the use of a relatively low SNR detection threshold (SNR=3σ) and the bias introduced using the targeted fields. At these flux limites, we find that ALMACAL is consistent with the results from ASPECS (González-López et al., 2020) and the ALMA Frontier Fields Survey (Gonzlez-Lpez et al., 2017; Muoz Arancibia et al., 2019). At the bright end, there are no direct number counts measurements from ALMA for sources brighter than $S_{1.2\,\mathrm{mm}} > 2$ mJy. To better constrain the 1.2 mm bright-end number counts, we adopted the SED-scaled number counts from AS2UDS (Stach et al., 2018, 2019). Dudzevičiūtė et al. (2020) derived SED fits for all SMGs within the $\approx$1 square degree AS2UDS, providing the opportunity to derive the expected counts in other wavelengths. We thus predict 1.2 mm flux densities for each SMG based on its best-fitting SED and then derive the number counts after correcting for the original survey completeness down to $S_{1.2\,\mathrm{mm}} \approx 3.6$ mJy (see also Stach et al., 2018). The AS2UDS prediction at 1.2 mm is over-plotted in Fig. 2.16. ALMACAL is also consistent with the AS2UDS predictions in the overlapping regime. In the future, a larger blind ALMA survey, or follow-up of the bright sources from wide field 1.2-mm single-dish surveys, is needed to verify our results and confirm the number counts at the brightest flux densities.

2 mm

There are relatively few published studies of number counts at longer wavelengths $\lambda \gg 1$ mm. The first single-dish 2 mm survey was conducted by the Goddard IRAM Superconducting Millimeter Observer (GISMO) at the IRAM 30-m telescope (Staguhn et al., 2014; Magnelli et al., 2019). The number counts were later confirmed by the Mapping Obscuration to Reionisation with ALMA (MORA) survey (Zavala et al., 2021). We revisit the 2 mm number counts with the detections from ALMACAL. In Fig. 2.17, we plot the new measurements from ALMACAL, as well as the data from Magnelli et al. (2019) and Zavala et al. (2021). Our new measurements are consistent with the two previous surveys at high flux densities and go 4× deeper, down to $S_{2\,\mathrm{mm}} \sim 0.1$ mJy. We also predict the 2 mm number counts based on the AS2UDS sample, following the same method we used for 1.2 mm in §2.5.1. The AS2UDS predictions are shown in Fig. 2.17, which matches well with MORA, suggesting that there is no substantial new population of sources appearing at $\lambda \gg 1$ mm.

Tab. 2.5: Best-fitting models for cumulative and differential number counts.

Cumulative number counts

	Schechter			Double power-law			
	α	S_0 mJy	N_0 deg^{-2}	α_1	α_2	S_0 mJy	N_0 deg^{-2}
B4 (2 mm)	$-1.7^{+0.3}_{-0.3}$	$1.0^{+0.6}_{-0.3}$	$0.09^{+0.01}_{-0.01}$	0.0	$2.1^{+0.2}_{-0.2}$	1.0	$0.04^{+0.01}_{-0.01}$
B6 (1.2 mm)	$-0.8^{+0.1}_{-0.1}$	$1.0^{+0.1}_{-0.1}$	$2.7^{+0.3}_{-0.3}$	$1.0^{+0.1}_{-0.1}$	$3.8^{+0.4}_{-0.3}$	$1.2^{+0.2}_{-0.2}$	$1.6^{+0.2}_{-0.2}$
B7 (870 μm)	$-0.6^{+0.1}_{-0.1}$	$2.2^{+0.1}_{-0.1}$	$9.6^{+0.9}_{-0.9}$	$1.0^{+0.1}_{-0.1}$	$4.7^{+0.2}_{-0.2}$	$4.5^{+0.4}_{-0.4}$	$3.6^{+0.7}_{-0.6}$

Differential number counts

	Schechter			Double power-law			
	α	S_0 mJy	N_0 $\mathrm{mJy}^{-1}\mathrm{deg}^{-2}$	α_1	α_2	S_0 mJy	N_0 $\mathrm{mJy}^{-1}\mathrm{deg}^{-2}$
B4 (2 mm)	-1.7	$0.2^{+0.4}_{-0.3}$	$6.0^{+0.9}_{-0.7}$	0.0	$3.1^{+0.8}_{-1.1}$	1.0	$0.1^{+0.2}_{-0.1}$
B6 (1.2 mm)	$-1.7^{+0.1}_{-0.1}$	$1.7^{+0.5}_{-0.3}$	$2.4^{+0.8}_{-0.9}$	$1.9^{+0.1}_{-0.2}$	$4.6^{+1.8}_{-1.0}$	$2.3^{+0.8}_{-0.9}$	$1.1^{+1.6}_{-0.5}$
B7 (870 μm)	$-1.7^{+0.2}_{-0.2}$	$3.9^{+0.7}_{-0.7}$	$2.2^{+1.4}_{-0.9}$	$2.0^{+0.2}_{-0.3}$	$4.7^{+0.9}_{-0.5}$	$5.6^{+2.2}_{-1.5}$	$0.8^{+1.6}_{-0.8}$

Note: Values without errors are fixed values during the fitting.

1.5 mm and 3 mm

ALMACAL is the first survey to constrain the number counts of DSFGs at 1.5 mm, with detections down to $S_{1.5\,\mathrm{mm}} \sim 0.4$ mJy. At 3 mm, Zavala et al. (2018) provided constraints on the DSFGs number counts using the ALMA Science Archive towards regions around targeted sources in three extragalactic legacy fields: COSMOS, CDF-S and the UDS. These results were later updated in Zavala et al. (2021) after removed three spurious detections. We compare our cumulative number counts with those of Zavala et al. (2021) in Fig. 2.18, along with the SED-based predictions from AS2UDS and various theoretical model predictions (Lagos et al., 2020; Popping et al., 2020). Our counts agree well with the those predicted from AS2UDS and Zavala et al. (2021) at the bright end. However, they are systematically lower than Zavala et al. at the faint end, which we attribute to their counts being biased high due to associations and potentially source blending in the original target selection.

Joint fitting

Combining different surveys offers a way to mitigate some of the biases that are particular to each individual survey. We have therefore perform a joint fit of the count data in ALMA bands where data is sufficient: 870 μm, 1.2 mm abd 2 mm. In band 7 (870 μm), we include the data from Karim et al. (2013); Simpson et al. (2015a); Stach et al. (2018); Béthermin et al. (2020);

Tab. 2.6: Best-fitting models for cumulative and differential number counts.
Joint fit

	α/α_1	$/\alpha_2$	$S_{2.0\,\mathrm{mm}}$ mJy	$N_{2.0\,\mathrm{mm}}$ mJy^{-1}deg^{-2}
Schechter	$-1.7^{+0.1}_{-0.1}$		$0.3^{+0.1}_{-0.1}$	$5.0^{+6.9}_{-4.0}$
Double power-law	$1.9^{+0.1}_{-0.1}$	$4.2^{+0.3}_{-0.3}$	$0.3^{+0.1}_{-0.1}$	$3.7^{+6.9}_{-3.9}$
	$S_{1.2\,\mathrm{mm}}$ mJy	$N_{1.2\,\mathrm{mm}}$ mJy^{-1}deg^{-2}	$S_{870\,\mu\mathrm{m}}$ mJy	$N_{870\,\mu\mathrm{m}}$ mJy^{-1}deg^{-2}
Schechter	$1.9^{+0.5}_{-0.3}$	$1.8^{+0.8}_{-0.7}$	$3.9^{+0.5}_{-0.5}$	$2.1^{+0.7}_{-0.5}$
Double power-law	$2.2^{+0.5}_{-0.4}$	$1.1^{+0.7}_{-0.5}$	$4.3^{+0.7}_{-0.6}$	$1.3^{+0.6}_{-0.5}$

Simpson et al. (2020); in band 6 (1.2 mm), we include the data from Umehata et al. (2017); Muñoz Arancibia et al. (2018); González-López et al. (2020) and the SED-scaled number counts from AS2UDS (Stach et al., 2018); in band 4, we include the data from Zavala et al. (2021). We fit the differential and the cumulative number counts with the Schechter function (Schechter, 1976) and double power law:

$$N(>S)\,dS = N_0\left(\frac{S}{S_0}\right)^{\alpha}\exp\left(-\frac{S}{S_0}\right)d\left(\frac{S}{S_0}\right)\ , \tag{2.4}$$

$$N(>S)\,dS = N_0\left[\left(\frac{S}{S_0}\right)^{\alpha_1}+\left(\frac{S}{S_0}\right)^{\alpha_2}\right]^{-1} d\left(\frac{S}{S_0}\right)\ , \tag{2.5}$$

where N_0 is the normalisation factor, in units of deg^{-2} for the cumulative number counts and in units of mJy^{-1}deg^{-2} for the differential number counts. S_0 is the flux density at the turnover. The index, α, is the power-law index of the Schechter function in Equation 2.4; while α_1 and α_2 are the slopes of the two independent power laws in Equation 2.5. During the fitting, N_0, S_0 and α, or α_1 and α_2 are free parameters. The best-fit models for the different bands are summarised in Table 2.5 and the best-fit Schechter functions are plotted in Figs 2.15, 2.16 and 2.17. We adopt the best-fit Schechter function as our fiducial model, because it provides the best fit to the flattening trend in the very deep 1.2-mm number counts (see discussion in González-López et al. 2020).

We constrain the free parameters using the maximum-likelihood minimisation algorithm, MINIMIZE, from SCIPY.OPTIMIZE, and adopt the Markov chain Monte Carlo (MCMC) sampler, EMCEE, to derive the confidence levels. We treat different surveys equally, using their own reported errors, but increase the uncertainties if re-scaling is applied (see more in §2.4.2). In band 4, the available data are not sufficient to constrain the models. We thus fix $\alpha = 1.7$ for the Schechter function and $\alpha_1 = 0$ and $S_0 = 1$ mJy for the double power law (see more discussion in §2.5.3).

From our best fits, we confirm the shallower trend seen in the faint flux density range at 870 μm (Béthermin et al., 2020) and 1.2 mm (González-López et al., 2020). Based on the joint Schechter function fitting for the differential number counts, the turnover flux density at 870 μm is $S_{870\mu\mathrm{m}} = 3.9^{+0.7}_{-0.7}$ mJy. At 1.2 mm, González-López et al. (2020) found that the number counts flattened below 0.1 mJy, and argued for a triple power law to fit the differential number counts, based on their $P(D)$ analysis. In our fitting, a Schechter function with a turnover at $S_{1.2\mathrm{mm}} = 1.7^{+0.5}_{-0.3}$ mJy with a power-law index of $\alpha = 1.7 \pm 0.1$ gives a reasonable fit to the existing data. However, as we mentioned earlier, considerable uncertainties exist both at the fainter and brighter end of the 1.2 mm number counts. At 2 mm, the constrained fitting gives a turnover flux $S_{2\mathrm{mm}} = 0.2^{+0.4}_{-0.3}$ mJy, which still suffers from large uncertainties.

2.5.2 Comparing with model predictions

Simulations and analytical models of galaxy formation and evolution must also explain the DSFG population. Depending on basic assumptions and underlining techniques, these modelling methods can be roughly divided into semi-analytic models (SAMs), semi-empirical models and hydrodynamic (N-body) simulations.

SAMs are the most widely used tools (Cole et al., 2000). They start with the merging trees of dark matter halo from either cosmological N-body simulations or self-consistent Monte-Carlo simulations and use analytical equations for baryonic processes. Their flexibility and relatively cheap computational cost make them powerful to explore large regions of parameter space. However, it was difficult to reproduce submm/mm number counts by incorporating standard feedback processes along with radiative transfer to balance the FUV-to-NIR and the FIR-to-submm/mm emission of galaxies simultaneously across cosmic time (Baugh et al., 2005; Somerville et al., 2012; Lacey et al., 2016; Cowley et al., 2019). Lagos et al. (2019, 2020) adopted attenuation curves obtained from the radiative transfer analysis of hydrodynamical simulations (Trayford et al., 2020) to reproduce the panchromatic emission of galaxies. Their model is better able to match the multi-wavelength DSFG number counts. We show predictions from Lagos et al. (2020) in Fig. 2.13. They are broadly consistent with ALMACAL in the various ALMA bands, but exhibit over-prediction at the fainter end of bands 6 and 7.

Due to the increasing complexity of SAMs and the computational cost of fully three-dimensional dust radiative transfer, observational empirical scaling relations have been coupled into SAMs – creating semi-empirical models – to simplify the calculations (Bthermin et al., 2017; Popping et al., 2020). Popping et al. (2020) provide a new framework to explain the number counts at 1.2 mm (González-López et al., 2020), modelling the submm/mm emission of galaxies with simple functions from full radiative transfer simulations, where the submm/mm flux density of a galaxy is a function of SFR and total dust mass (Hayward et al., 2011). The Popping et al. model is extended here to provide the predictions on the Rayleigh-Jeans tail of the submm SED by fitting the original 850-μm and 1.1-mm predictions with a grey body. Their predictions for different ALMA bands are shown in Fig. 2.13. They are slightly lower than the predictions from Lagos et al. (2020) at fainter flux densities and are more consistent with the observations but at the cost that there are no true ab initio predictions. Casey et al. (2018a) also explored the number counts with different fractions of DSFGs at $z > 4$. They proposed two simple models: their dust-poor model has a steep slope in the luminosity function in the early Universe, while their dust-rich model has a much shallower slope and predicts many more DSFGs. The two models thus predict different number counts in the mm bands, with the dust-rich model having a higher

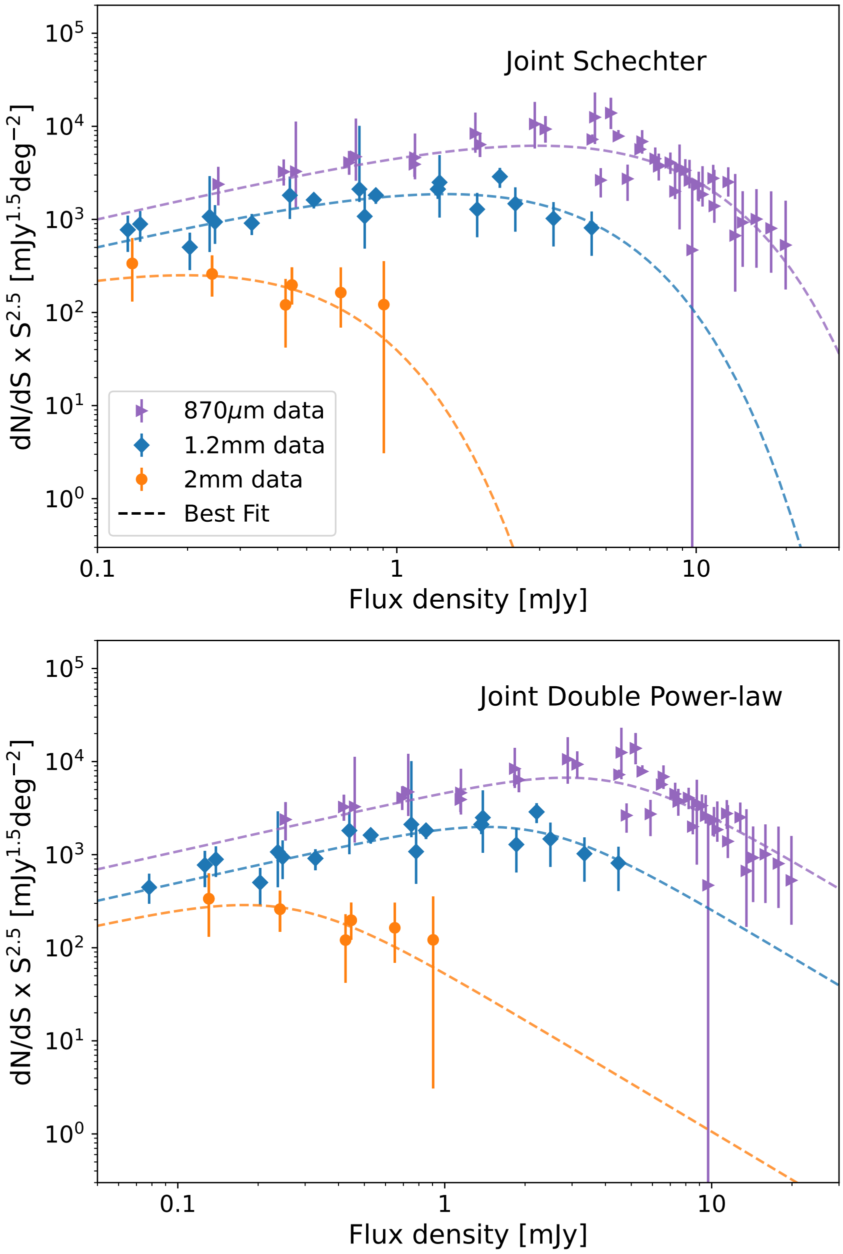

Fig. 2.21: Joint fits combining differential number counts at 2 mm, 1.2 mm and 870 μm. *Left:* Joint Schechter fits for all the data mentioned in §2.5.1. During the fitting, the α index of the Schechter function is bound for all the three wavelengths. The best final fits give $\alpha = -1.7 \pm 0.1$. *Right:* Joint double power-law fits. In the flux density regime of ALMACAL, all the measurements can be fitted with the same power-law index $\alpha = -1.9 \pm 0.1$, indicating the same or similar galaxy population behind the number counts.

number density of DSFGs. Our 2 mm number counts favour the dust-poor model, which is also closer to the prediction from Popping et al. (2020) and Lagos et al. (2020).

Hydrodynamic models, on the other hand, still struggle to reproduce the number counts of DSFGs and usually underpredict their numbers (Shimizu et al., 2012). With recent advances in computational resources, fully 3-D radiative transfer (RT) has been integrated with modern cosmological simulations with different recipes for stellar and AGN feedback (Camps et al., 2018; McAlpine et al., 2019; Lovell et al., 2021). Here, we compare our results with EAGLE (Evolution and Assembly of GaLaxies and their Environment, Schaye et al. 2015) and SIMBA (Dav et al., 2019) cosmological simulations that have 3-D radiative transfer implementations (Camps et al., 2018; Lovell et al., 2021). To compute the number counts of EAGLE, we first retrieve the publicly available observer-frame fluxes in ALMA bands 6 and 7 of all galaxies from $z = 10$ to $z = 0$ (see details about the flux calculation in Camps et al., 2018). Then, we use the fluxes and all the snapshots of the simulation to construct a lightcone. For each redshift, we compute the projected sky area of 100 Mpc^2 and the redshift bin implied by a 100 Mpc comoving distance. We then calculate the number of galaxies per unit area, per unit redshift, at each flux bin contributed by each snapshot $N(z, S)$. Finally, We integrate $N(z, S)$ along redshift to obtain the cumulative number counts $N(S)$. For SIMBA, we adopt the results from Lovell et al. (2021), which reproduced the population of bright DSFGs but – due to the computational cost – their predicted number counts are only complete down to $S_{850\mu\mathrm{m}} \sim 1$ mJy. We show all the predictions in Fig. 2.15 and Fig. 2.16.

Modelling the basic observations can help us to understand the relevant physical processes behind galaxy formation and evolution. Number counts are one of the most basic measurements from observations and play an essential role in validating models. By construction, semi-empirical models closely match the observed scaling relations, producing reasonable number counts at most of the observable flux densities, which can help design future surveys. SAMs start with basic physical assumptions, representing our understanding of the physics behind the observables. They can already give reasonable predictions for the flux densities covered by existing surveys. Future surveys with the flexibility offered by SAMs will continue to be a powerful tool to understand the fundamental physics behind the number counts. Hydrodynamic cosmological simulations remain the most computational expensive method. However, they also offer the most detailed description of the internal/external galactic structures and their environments. The discrepancy between the model predictions and the observations will continue to motivate more detailed sub-grid physics in cosmological simulations. Multi-wavelength number counts – especially in the submm/mm range – offer additional constraints on the redshift distribution of DSFGs, thanks to the negative-K correction (Baugh et al., 2005; Casey et al., 2014). Future, deep multi-band DSFG survey will therefore continue to be a critical benchmark for modern galaxy evolution models.

2.5.3 The galaxy populations behind the number counts

It has long been argued that submm/mm surveys at different depths and different wavelengths probe different galaxy populations and redshifts (Chapman et al., 2005; Casey et al., 2014). This claim has been explored by semi-empirical models and semi-analytical models (Béthermin et al.,

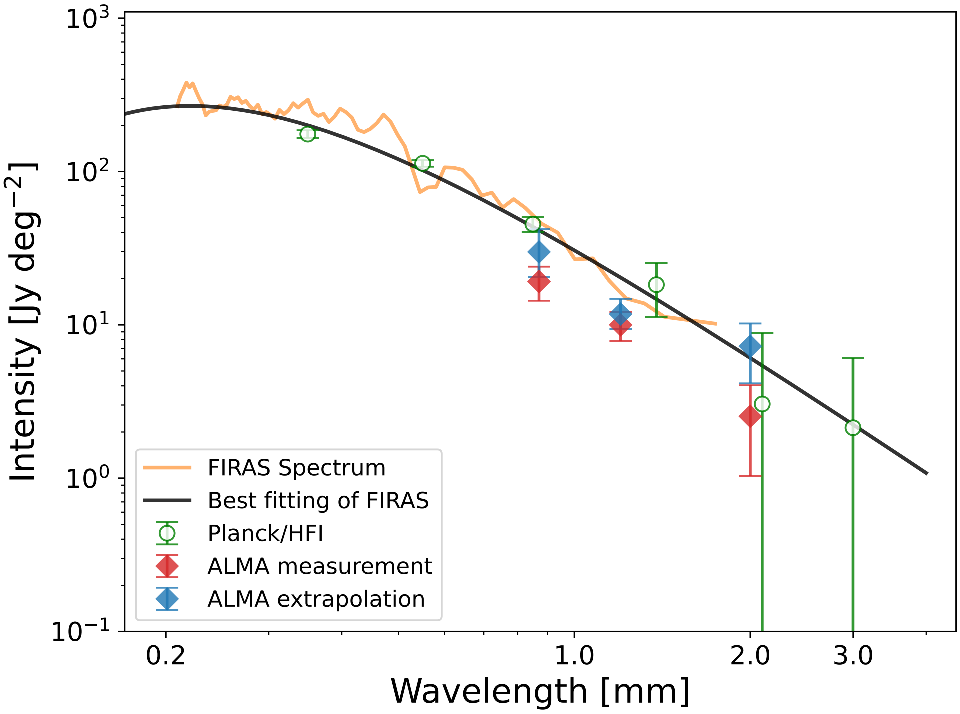

Fig. 2.22: ALMA resolved cosmic infrared background. The resolved CIB at different ALMA bands is the integration of joint Schechter fitted differential number counts within the observed flux density ranges. The predicted CIB is the whole integration of the best-fitting differential number counts. The direct measurement from FIRAS on *COBE* is shown in orange. The black curve is the best-fitting FIRAS spectrum made by Fixsen et al. (1998). We also show the recalibrated *Planck*/HFI data (Odegard et al., 2019). Comparing with the direct measurements from FIRAS, ALMA currently resolved nearly one half of the CIB from 870μm to 2 mm, but the fraction is highly uncertain due to the large uncertainties of the FIRAS spectrum at the submm/mm wavelengths.

Tab. 2.7: ALMA resolved cosmic infrared background.

λ mm	Resolved Intensity Jy deg^{-2}	Extrapolated Intensity Jy deg^{-2}	Resolved fraction ^{a} Observed	Resolved fraction ^{b} Extrapolation
2000	2.4 ± 1.7	7.0 ± 3.6	$0.40^{+0.23}_{-0.11}$	$0.36^{+0.27}_{-0.25}$
1200	11.2 ± 2.2	14.1 ± 2.9	$0.54^{+0.32}_{-0.15}$	$0.79^{+0.18}_{-0.15}$
870	18.3 ± 4.7	27.0 ± 7.8	$0.44^{+0.26}_{-0.12}$	$0.67^{+0.17}_{-0.17}$

Note: The resolved intensity is the integration within the flux density range that is covered by ALMA; a): the resolved fraction compared with the FIRAS measurement; b): the resolved fraction compared with the prediction of this work.

2015; Popping et al., 2020; Lagos et al., 2020). Béthermin et al. (2015) argued that shallower surveys at longer wavelengths yield DSFGs with a higher median redshift. Popping et al. (2020) emphasised that the contribution from galaxies at $z > 4$ to the number counts is small and the flattening of the number counts at $S_{1.2\text{mm}} < 0.1$ is likely caused by the shallow faint-end luminosity functions of $z = 1$–2 DSFGs. Lagos et al. (2020) also suggested that the redshift increases with flux density for bright SMGs, but this trend flattens out for fainter DSFGs. In general, simulations reproduced the observational trend that reshift increases with flux density for bright SMGs (Simpson et al., 2020, see also Ivison et al. 2007; Simpson et al. 2017; Brisbin et al. 2017; Stach et al. 2019), but we still have limited observational constraints about the redshift distribution of fainter DSFGs.

Thanks to the unique wavelength coverage of ALMACAL, we have been able to constrain the multi-band number counts for faint DSFGs simultaneously. Even though all the parameters are fitted independently, band 6 and 7 share remarkable similar power-law indices (see Table 2.5). If we were probing a similar galaxy population at different wavelengths and if their redshift distributions do not vary significantly, the conversion of number counts between different wavelengths could then be approximated by a simple colour difference. Inspired by this idea, we jointly fit the differential number counts in band 6 and band 7 by linking their power-law indices. In Fig. 2.21, we show the joint Schechter and double power-law fitting of ALMACAL measurements. In the flux density regime of ALMACAL, the differential number counts in band 6 and band 7 can be well fit with the same power-law index with $\alpha_1 = -1.9 \pm 0.1$, implying the number counts in band 6 and band 7 are dominated by the same galaxy population. A tentative fit with the same power-law index can also be applied to the fainter end number counts at 2 mm, but the statistical significance is limited. The joint Schechter fitting gives $\alpha = 1.7 \pm 0.1$, which is the same for independent fitting for band 6 and 7. To give more robust total CIB prediction in band 4, we fixed $\alpha = 1.7$ for the Schechter fitting of differential number counts.

2.5.4 Resolved cosmic infrared background

Two methods have been used to constrain the CIB: the direct measurement and the integrated intensity based on galaxy number counts. For direct measurements, the pioneering experiment were achieved by the Far Infrared Absolute Spectrophotometer (FIRAS) aboard *Cosmic Background Explorer* (*COBE*) (Puget et al., 1996; Fixsen et al., 1998) and were later continued by instruments including *ISOPHOT* (Juvela et al., 2009), *AKARI* (Matsuura et al., 2011) and *Planck*/HFI (Planck Collaboration et al., 2014) at different wavebands. A direct measurement requires accurate foreground modeling for the scattered light from solar interplanetary dust and the emission from the Milky Way, which usually needs additional assumptions or ancillary data for calibration (Fixsen et al., 1998; Odegard et al., 2019). On the other hand, if we have good knowledge of the galaxy number counts, we can also reconstruct the CIB by integrating the intensities from all the galaxies at different wavelengths. Due to the limiting sensitivity of the instruments used to derive number counts, extrapolations are normally needed to recover the undetected faint sources. Comparing the integrated energy density of galaxy number counts with direct measurements, we can quantify the energy budgets of different galaxy populations. The accurate comparison between the two can also provide us clues about the presence of any unidentified populations.

After nearly two decades of efforts, a major part of the CIB has been resolved into single sources (Casey et al., 2014). However, large uncertainties still exist in the submm/mm bands. Firstly, the accuracy of direct CIB measurement in the original FIRAS spectrum is decreasing toward longer wavelengths, as one can see in Fig. 2.22. Secondly, the integration based on galaxy number counts also suffers from large uncertainties due to the poorly constrained slope at faint flux densities andalso cosmic variance uncertainties. Because of this, different works report quite different fractions of resolved CIB, both at 870 μm (e.g. Smail et al., 2002; Chen et al., 2013; Oteo et al., 2016a; Béthermin et al., 2020) and 1.2 mm (e.g. Fujimoto et al., 2016; González-López et al., 2020).

To make optimal use of the existing data and to minimize their biases, we choose to derive the resolved CIB using our joint best-fitting model. Two values are calculated: the resolved intensity and the extrapolated total intensity. The resolved intensity is the integration of the best-fitting differential number counts within the flux density range that has been covered by existing surveys. The extrapolated total intensity is based on the extrapolation of the best-fitting differential number counts to include both faint and bright sources that are beyond the detection limits of current surveys. It should be noted that the extrapolation can introduce large uncertainties if we adopt incorrect models. The deep ALMA blind surveys ASPECS and the ALMA Frontier Fields Survey at 1.2 mm (González-López et al., 2020; Muñoz Arancibia et al., 2018) suggested that the number counts flatten out at the very faint flux density end ($S_{1.2\,\mathrm{mm}} < 0.1$mJy), implying that the counts can be best described by a Schechter function. We thus made the predictions based on our best-fitting Schechter functions, using the theoretical integration formula:

$$\int_0^{\infty} N(>S)\,S\,dS = N_0 \times S_0 \times \Gamma(\alpha+2) \quad . \tag{2.6}$$

The resolved CIB and the predictions at 2 mm, 1.2 mm, and 870 μm are summarized in Table 2.7.

In Fig. 2.22, we compare our results with the direct measurements from *COBE*/FIRAS and the most recent re-calibrated *Planck*/HFI data (Odegard et al., 2019). Both the resolved intensities and the predictions from ALMA are shown.

Our results are consistent with the literature and our predictions are close to the direct measurements from FIRAS. At 870 μm, the resolved CIB for $S_{870\mu\mathrm{m}} > 0.2$ mJy is 18.3 ± 4.7 Jy deg^{-2}, which matches the resolved intensity of 16.4 ± 2.7 Jy deg^{-2} from the secure sample of Béthermin et al. (2020). Compared with direct measurements from FIRAS, the resolved intensity at 870 μm accounts for ~44 per cent of the total CIB. At 1.2 mm, our resolved intensity for $S_{1.2\,\mathrm{mm}} > 45\,\mu$Jy is 11.2 ± 2.2 Jy deg^{-2} and the extrapolated DSFGs' contribution to the CIB is 14.1 ± 2.9 Jy deg^{-2}. Both of these values are slightly larger than the measurement of ASPECS (González-López et al., 2020), but closer to the direct measurements. At 2 mm, the resolved CIB for $S_{\mathbf{2mm}} > 0.1$mJy is 2.4 ± 1.7 Jy deg^{-2}, which is close to half of the direct measurement. However, since large uncertainties also exist in the FIRAS measurement at this wavelength, the comparison the calculation of the resolved fraction is not very accurate. More recently, (Odegard et al., 2019) have improved the accuracy of the CIB direct measurement with Planck/HFI data, but the uncertainties at the longer wavelengths have not significantly improved. Therefore, to have a more precise comparison, we still need more accurate direct measurements of the total CIB at submm/mm wavelengths. Comparing our predictions based on extrapolations, we have resolved 36^{+27}_{-25}, 79^{+18}_{-15}, and 67^{+17}_{-17} per cent of the CIB at 2 mm, 1.2 mm, and 870 μm, respectively.

2.6 Conclusions

ALMACAL is a survey that is exploiting calibration data that are accumulating for free with every scheduled ALMA observing project. Before the observatory shutdown due to Covid-19 in March 2020, ALMACAL had accumulated more than 1,000 h of observing time and covered 1001 calibrator fields. The sensitivity, total sky coverage, and wide frequency sampling make ALMACAL a promising dataset to undertake blind surveys for DSFGs. Within the ALMACAL footprints, we detect 371 sources, including 186 DSFGs confirmed by their spectral indices. We report the number counts based on these DSFGs for ALMA band 3 to band 7 (wavelengths spanning from 3 mm to 870 μm), which are mostly in agreement with existing surveys at overlapping flux densities. In band 4 (2 mm) and 7 (870 μm), ALMACAL represents the deepest survey available to this date. In band 5 (1.5 mm), ALMACAL is the first survey to be able to constrain the number counts of DSFGs. Together with the previously reported band 8 results from Klitsch et al. (2020), we are now able to present number counts covering almost the entire wavelength range of covered by ALMA, from 0.65 to 3 mm.

We compare our number counts with various model predictions. The semi-analytic models from Lagos et al. (2020) and semi-empirical models from Bthermin et al. (2017) and Popping et al. (2020) match the number counts in ALMA bands reasonably well. Hydrodynamic simulations, although finding it generally harder to reproduce the number counts of DSFGs, have shown promising results from the recent efforts (e.g. McAlpine et al., 2019; Lovell et al., 2021). As demonstrated in this work, multi-wavelength number counts can be a powerful benchmark to validate galaxy formation and evolution models. Future surveys, including the ongoing ALMACAL, will continue to improve the DSFG number counts by covering a larger sample and a wider flux density range.

We also turn ALMACAL into a cosmological survey to constrain the energy budget from DSFGs to the cosmic infrared background. We provide joint fits for number counts of DSFGs by combining ALMACAL with literature ALMA surveys. Compared with the direct measurements from FIRAS/*COBE* (Fixsen et al., 1998) and HFI/*Planck*, we report ALMA has directly resolved 40^{+23}_{-11}, 54^{+32}_{-15}, and 44^{+26}_{-12} per cent of CIB at 2 mm, 1.2 mm, and 870 μm, respectively. Due to the large uncertainties of direct measurements at submm/mm wavebands, we still suffer from large errors, thus demanding more accurate direct CIB measurements.

The large number of detections in ALMACAL suggests the promising opportunity to conduct a large submm/mm survey with ALMA calibration data. We only explore the continuum image of these calibrator fields, more treasures are still buried in this database. In addition, ALMACAL is just a small part of the ALMA archive database. How to make use of the enormous ALMA archive effectively is challenging, but also a promising way to find more hidden gold.

Chapter 3

The enviroments of DSFGs: ALMA survey of galaxy proto-cluster cores

3.1 Motivation

In Chapter 1, I have summarised the different methods in searching for galaxy proto-clusters and their indications on the transition between the cluster and proto-cluster. In this Chapter, I will introduce my efforts to search for the proto-cluster cores with ALMA, and also the lessons I have learned during this process. The first part about the projection effects of the DSFGs is based on my published paper Chen et al. (2023) and the second part is based on the early results of the ALMARED survey.

Submm surveys are essential in the search for proto-cluster cores. In the proto-cluster cores, the enhanced level of star formation and the rapid metal enrichment can make these systems bright in dust emission. Coincidentally, detailed studies of the SMGs have shown that they are predominantly starburst galaxies. Moreover, since the very first discoveries of SMGs by single-dish telescopes, high-resolution interferometric follow-ups have resolved some of them into multiple components. Combined with their high SFR, it is natural to connect the birthplace of SMGs with the dense filaments and nodes of the cosmic web in the early Universe. As indicated by their morphologies in the resolved studies, a large number of SMGs may have undergone a recent merger, which can also be a natural result if they reside in a denser environment. Therefore, studying the number density and evolutionary state of proto-clusters provides stern tests of the earliest stages of galaxy cluster formation models, which is crucial for understanding the formation of the most massive structures in our Universe.

Oteo et al. (2018) presented the first known proto-cluster core – the Distant Red Core (DRC; Fig. 1) – identified as an over-density of ultraluminous sources in SCUBA-2 and LABOCA surveys of the ultra-red sources from *Herschel* (Ivison et al., 2016a). In the central 300 kpc of DRC there are at least 12 DSFGs, each confirmed spectroscopically at $z = 4.002$, with IR luminosities ranging from 3×10^{11} to $1.6\times10^{13}\,\mathrm{L}_\odot$ and individual SFRs $50 < \mathrm{SFR} < 3000\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$ (Oteo et al., 2018; Ivison et al., 2020). Miller et al. (2018) found another example, SPT 2349−56, a system of similar richness and scale, at $z = 4.3$, identified initially by the South Pole Telescope

(Vieira et al., 2013).

In both cases, multiple sub-mm bright sources are located at the same redshift and within a few hundred co-moving kpc. The dynamics of the populations indicate that the systems are the cores of already massive proto-clusters that are likely to evolve into clusters with masses $> 10^{14}\,\mathrm{M}_{\odot}$, quite possibly $\sim 10^{15}\,\mathrm{M}_{\odot}$. Proto-clusters identified through optical techniques have a far lower density of DSFGs at this high luminosity and SFR (e.g. Toshikawa et al., 2020; Hu et al., 2021), which raises the possibility that optical selection methods miss a key stage in the evolution of clusters and their galaxy populations.

Despite their importance, we are still a long way from understanding their nature. Such systems have been detected at $z \sim 4$, and less extreme systems have even been detected up to $z \sim 5$ (Capak et al., 2011; Walter et al., 2012). The collapse timescale of a cluster is close to the Hubble time, yet these proto-cluster cores were already dense and massive when the Universe was less than 10% of its current age. Do we see these proto-clusters just as they started to form, or do they also contain older, more evolved galaxies? How did they accrete such enormous masses of cold gas? If starbursts are ephemeral, with typical timescales of ≈ 100 Myr, as implied by their M_{gas}/SFR, how and why can we be seeing so many simultaneous starbursts across a ≈Mpc structure?

In this section, we first report a serendipitous discovery in the ALMACAL footprints, which highlights the contamination of projections from the negative-K correction; Then, I report another recent efforts to search for more proto-cluster cores from the ALMARED survey.

3.2 Projection effects of DSFGs: a cautionary tale

3.2.1 Galaxy overdensity

Traditionally, and for obvious reasons, statistical over-densities of galaxies have been used to identify proto-clusters. Pioneering studies utilised broad-band photometry to search for over-densities of Lyman-break galaxies (LBGs, e.g. Steidel et al., 1998), which led to the discovery of giant structures such as that at $z = 3.1$ in the SSA 22 field. Many more were found subsequently via deep narrow-band surveys, searching for various strong line emitters, like Ly α emitters (e.g. Steidel et al., 2000; Matsuda et al., 2004; Ouchi et al., 2005) and H α emitters (e.g. Kurk et al., 2004; Tanaka et al., 2011; Hayashi et al., 2012). Large multi-band photometric surveys also inspired selection based on colour (e.g. Zirm et al., 2008; Capak et al., 2011; Wylezalek et al., 2013a; Laporte et al., 2022). Those surveys preferentially targeted existing massive galaxies or structures, such as high-redshift radio galaxies (HzRGs), quasars and giant Ly α blobs (LABs). Among them, HzRGs are known to be hosted by massive galaxies (Seymour et al., 2007) and are therefore thought to trace the most massive dark matter halos. Indeed, they have often been found to signpost over-dense regions and are frequently embedded in proto-clusters (e.g. Stevens et al., 2003; Miley & De Breuck, 2008). LABs are also giant structures, sometimes extending several hundred kpc from the central engine. They are widely thought to be powered by starbursts and/or strong active galactic nuclei (AGN) feedback, both preferentially triggered in denser environments (e.g. Overzier et al., 2013; Umehata et al., 2019). Specifically, in the proto-typical proto-cluster, the Spiderweb, at $z = 2.2$, all these methods have been used to reveal various galaxy populations (Kurk et al., 2000; Pentericci et al., 2002; Kurk et al., 2004; Zirm et al., 2008; Hatch et al., 2011).

On the smaller scales, recent advances in submm interferometry have accelerated the confirmation and resolved studies of the most extreme and rare proto-cluster cores. Proto-cluster cores are supposed to be mass centre of the galaxy proto-clusters, which is likely tracing the cores of the most massive dark matter halos. These rare galaxy over-densities were first discovered as single bright sources by single-dish telescopes with limited spatial resolution, then later resolved into typically a dozen starburst galaxies by deeper interferometric observations. In the Distant Red Core (DRC) at $z = 4.0$, for example, first identified via imaging with *Herschel*, the James Clerk Maxwell Telescope (JCMT) and APEX by Ivison et al. (2016b), 12 SMGs were later confirmed in the central 300 kpc region and four more gas-rich galaxies in the outskirts (e.g. Oteo et al., 2018; Ivison et al., 2020); in SPT 2349−56, at $z = 4.3$, 14 dusty star-forming galaxies were found gathered within a 130-kpc region (Miller et al., 2018). These recent observational breakthroughs suggest that we may be witnessing the most extreme evolutionary stage of a proto-cluster, where the central region is being transformed into something typical of today's clusters: a massive central cD-type galaxy, such as NGC 1275 in the Perseus cluster. With the advent of the *James Webb Space Telescope (JWST)*, we will likely find many more such systems at higher redshifts (e.g. Laporte et al., 2022; Jin et al., 2023; Morishita et al., 2023; Helton et al., 2023).

However, the number of confirmed proto-clusters remains limited. Due to the expense of the spectroscopic follow-up required to confirm the redshifts of member galaxies, searching for projected over-densities remains the most commonly used technique for identifying proto-clusters (e.g. Lammers et al., 2022). Because they can cover larger areas than interferometers, single-dish submm telescopes continue to target promising candidates: e.g. HzRGs, LABs, as well as over-densities found at other wavelengths, searching for additional SMGs (e.g. Robson et al., 2014; MacKenzie et al., 2017; Zeballos et al., 2018; Li et al., 2020; Wang et al., 2021; Nowotka et al., 2022; Zhang et al., 2022; Li et al., 2023).Over-densities have been reported in many of these systems, although the densities are only marginally higher than those of the control fields and further follow-up often gives mixed results. For instance, Wylezalek et al. (2013b) used far-IR/submm spectral energy distributions (SEDs) to constrain the redshifts of galaxies towards the $z = 3.8$ radio galaxy, 4C 41.17, finding just one of the many neighbouring far-IR galaxies to be at similar redshift as 4C 41.17, with most of the sources in the foreground. Similar false confirmations have also been reported in several proto-cluster candidates (Chapman et al., 2015; Meyer et al., 2022).

In this subsection, we report the discovery of an extreme over-density within the ALMACAL survey: seven point-like sources around an ALMA calibration source – the blazar, J0217−0820, at $z = 0.6$. Six of these are shown to be dusty star-forming galaxies (DSFGs) based on multi-band submm photometry. The over-density is comparable to the most extreme known, confirmed proto-cluster cores.

I first summarise the ALMACAL survey and the data used in this work, in §3.2.2. We describe the over-density of submm galaxies around J0217−0820 in §3.2.3. We unveil the redshifts of those galaxies in §3.2.3, then discuss the conclusions and implications in §3.2.4, using light-cones made from recent cosmological simulations to better understand the statistics.

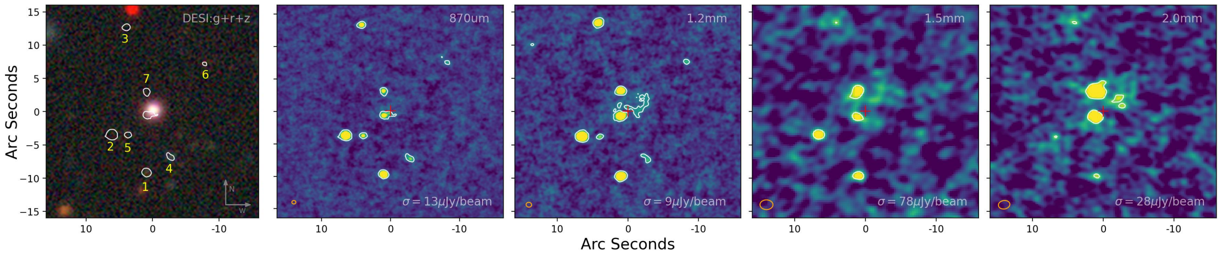

Fig. 3.1: Multi-wavelength images of J0217−0820. Left to right: RGB images from the DESI Legacy Imaging Surveys; ALMA images at 870 μm, 1.2 mm, 1.5 mm, and 2.0 mm. In the ALMA images, the central blazar has been removed; The orange ellipse shows the FWHM of the synthesized beam and the white contours show the emission 3 σ above the RMS noise. Except for the blazar, all the other sources are invisible in the optical image.

Tab. 3.1: Observed properties of the seven objects and the blazar.

ID	R.A. h:m:s	Dec. d:m:s	z	$S_{2\,\mathrm{mm}}$ (mJy)	$S_{1.5\,\mathrm{mm}}$ (mJy)	$S_{1.2\,\mathrm{mm}}$ (mJy)	$S_{870\mu m}$ (mJy)
HLW-1	02:17:02.73	−08:21:02.1	2.240	0.20±0.05	0.42±0.20	1.44±0.05	4.00±0.19
HLW-2	02:17:03.10	−08:20:56.2	2.460	0.17±0.06	0.58±0.16	1.43±0.05	3.93±0.17
HLW-3	02:17:02.94	−08:20:39.2	1.238	0.30±0.08	–	1.23±0.07	3.05±0.26
HLW-4	02:17:02.47	−08:20:56.9	?	–	–	0.13±0.02	0.47±0.07
HLW-5	02:17:02.93	−08:20:56.2	2.460?	–	–	0.11±0.03	0.42±0.21
HLW-6	02:17:02.12	−08:20:44.7	1.477	–	–	0.13±0.10	0.36±0.16
HLW-7	02:17:02.73	−08:20:49.1	?	0.98±0.05	0.79±0.10	0.53±0.08	0.43±0.11
Blazar	02:17:02.66	−08:20:52.35	0.6065	281±14	226±10	157±8	115±6

3.2.2 An serendipitous discovery from ALMACAL survey

ALMACAL exploits ‘free’ calibration data to monitor the ALMA calibration sources – most of which are blazars – and survey their immediate vicinities (Zwaan et al., 2022). The ALMA calibrators are distributed fairly randomly across the sky, serving to calibrate the bandpass response, complex gains, the flux density scale, and the degree and angle of polarisation. Overall, ≈ 20% of the available ALMA telescope time is and has been spent on calibrators. For some calibrators, the depth reached by combining all the available visits is comparable with the deepest cosmological fields. Since calibration scans typically cover many frequencies, the resulting spectral coverage can be considerable, offering the opportunity to search for spectral lines. Towards the calibrator, J1058+0133, for example, Oteo et al. (2017b) found two SMGs and detected a host of spectral lines, yielding a remarkable spectral line energy distribution; Towards calibrator J0238+1636, Klitsch et al. (2019b) observed multiple CO emission lines around one Ly α absorber, which shows excited interstellar medium.

Chen et al. (2023) conducted a multi-band survey for dusty starbursts as part of the ALMACAL project. J0217−0820 was the densest region found among the available ALMACAL footprints. Until August 2022, around 8.6 h of data had accumulated. Combining these data, ALMA has sampled the spectrum from band 3 to band 7. In band 7, the 1-σ depth reaches 13 μJy beam^{-1}. Within $d = 30''$ around the centre, seven continuum sources have been found ($> 5\sigma$ significance), including three at $S_{870\mu\mathrm{m}} > 1$ mJy, which we will call traditional SMGs, following Hodge & da Cunha (2020), and three fainter sources, which we will call DSFGs, see also Table 3.3.

The data retrieval and calibration in J0217−0820 is similar to the scheme described in Oteo et al. (2016a). Summarising, we first collected all the observations that used J0217−0820 as a calibrator, then run the standard calibration offered by `ScriptForPI.py` in each project, splitting out the calibrated data for J0217−0820. Within each project, the flux density was calibrated by the dedicated calibrator, or scaled to the most recent online catalogue. Next, we self-calibrated the data in phase-only and in phase-and-amplitude mode to improve the fidelity of the image.

Blazars tend to be variable at submm wavelengths (e.g. Robson et al., 1983). To minimise the resulting effects, we used Common Astronomy Software Applications (CASA) tool, UVMODELFIT, to subtract the central calibrator from each independent observation before imaging, adopting a point source model.

After these steps, we combined the observations to create the final continuum image and datacube in each band, as follows. We used two imaging cycles to create the continuum images. We first made dirty images of all the observations using TCLEAN in ‘mfs’ mode, with zero clean iterations. Next, we rejected images with calibration issues and unsuccessful point-source subtraction, based on our visual inspection. Less than 2% of the total data were lost. After that, we used the CASA task, STATWT, to re-calculate the weights, based on the noise in each observation. Finally, TCLEAN constructed the combined image, wherein the AUTO-MULTITHRESH algorithm searched for robust emission during each major clean cycle. We also tapered the image to 0.4 and 0.8 arcsec spatial resolution (FWHM) to aid recovery of any extended sources. Our images of J0217−0820 are shown in Fig. 3.1.

We classified the continuum sources based on their multi-band colours (see Appendix 3.2.3).

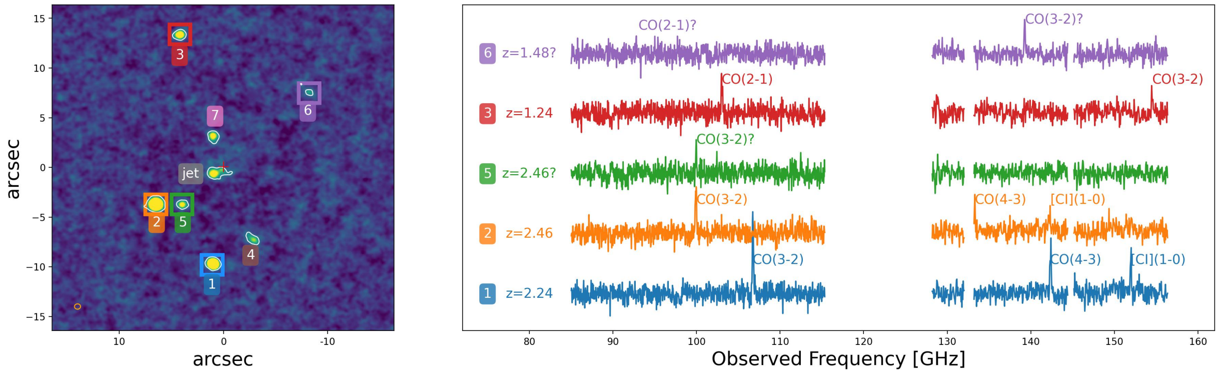

Fig. 3.2: Redshifts of the continuum sources in the J0217−0820 field. Left: continuum image at 870 μm. Right: confirmed spectral lines associated with those continuum sources in ALMA band 3 and band 4. A zoom version with the additional lines is available in Appendix Fig. 3.4. We have confirmed five redshift groups. The central blazar lies at $z = 0.6$, while the surrounding dusty starbursts are located at four different redshifts.

Among the seven point-like sources, six have thermal spectra consistent with dusty starbursts, and one is a synchrotron source. Continuum images and measurements at different wavelengths are presented in Fig. 3.1 and Table 3.3. We name the seven sources as the Calabash brothers, or Huluwa, abbreviated to HLW[1]. There is also a radio jet associated with the central blazar, heading south-east.

To create the datacubes, we also followed two cycles. First, we used all the available data to create a dirty datacube in each ALMA band. After that, we extracted spectra at the positions of the seven continuum sources. Since the calibrator always shares the same instrumental configuration as the intended science target, the sensitivity can be quite different from observation to observation, resulting in quite different sensitivities at different frequencies. To make the best use of the spectral data, we normalised the extracted spectra by the sensitivities in each channel to create a ‘signal-to-noise (SNR) spectrum’ for each continuum source. These SNR spectra preserve strong spectral lines whilst suppressing large spikes caused by the low sensitivity at some frequencies. The resulting spectra are shown in Fig. 3.2.

Starting with the 5 σ peaks in our SNR spectra, we searched for additional lines, assuming the original peak to be ^{12}CO or [C I]. After searching for all the possible emission lines, we returned to the original calibrated data and created a 1-GHz-wide datacube for each line. Similar to the cleaning process for our continuum images, AUTO-MULTISHRESH was applied to identify channels with robust emission. We stopped cleaning when the peak emission fell below the 2σ noise level of the residual image.

---

[1] There are seven Hulu Wa, each with different characteristics and abilities, see: Calabash Brothers

3.2.3 Analysis

Over-densities

We define the over-density of galaxies as follows:

$$\delta_{\rm gal} = \frac{n_{\rm gal} - \langle n_{\rm gal} \rangle}{\langle n_{\rm gal} \rangle}, \tag{3.1}$$

where $\delta_{\rm gal}$ is the galaxy over-density, $n_{\rm gal}$ is the number of galaxies found in the field of interest and $\langle n_{\rm gal} \rangle$ is the average number of galaxies from the adopted reference survey.

Towards J0217−0820, we have detected six DSFGs with $S_{870\mu\rm m} > 0.3$ mJy, three of them being classical SMGs with $S_{870\mu\rm m} > 3$ mJy. Based on the number counts from large surveys at 870 μm (Chen et al. 2023, see also Stach et al. 2018; Béthermin et al. 2020; Simpson et al. 2020), the expected average number of DSFGs in the FoV of ALMA band 7 ($d = 30''$) is $\langle n_{\rm gal} \rangle \approx 0.050 \pm 0.002$ at $S_{870\mu\rm m} > 3$ mJy and $\langle n_{\rm gal} \rangle \approx 0.67 \pm 0.08$ at $S_{870\mu\rm m} > 0.3$ mJy. Then, the inferred over-density is $\delta_{\rm gal} \approx 59$ and $\delta_{\rm gal} \approx 8.0$, respectively. If we were to count the synchrotron source, the over-density would be even higher. Empirically, finding $\delta_{\rm gal} \gtrsim 8$ indicates strongly that a field is not consistent with the reference field, and likely contains a proto-cluster (e.g. Chiang et al., 2013; Lovell et al., 2018). Solely based on its over-density, J0217−0820 is thus an extremely promising proto-cluster candidate, similar to DRC (Lewis et al., 2018).

Redshift constraints

We identified 15 spectral lines in this field thanks to the wide spectral coverage of ALMACAL. We list all the confirmed spectral lines and their velocity-integrated flux densities in Appendix Table 3.2.

The brightest three SMGs (HLW-1, HLW-2, HLW-3) have at least three spectral lines each, sufficient to tie down their redshifts unambiguously. HLW-4 displays no robust spectral lines. HLW-5 has only one robust spectral line, but it is at the same frequency as a line seen from its neighbour, HLW-2, so they are likely an interacting pair, as is common or possibly even ubiquitous amongst SMGs (Engel et al., 2010). HLW-6 has one strong spectral line ($> 5\sigma$) as well as a weak line at 92.8 GHz, consistent with CO(2–1) and (3–2) at z=1.477. Even without the second, weaker line, it is clear that HLW-6 is not associated with any other galaxies in this field.

Considering its synchrotron spectrum, HLW-7 is likely at $z = 0.6065$ (as determined for the central blazar by the Sloan Digital Sky Survey – Albareti et al., 2017), i.e. it may well be another jet hotspot, or it can be a companion to the blazar as they are close enough together to have perhaps triggered each other's AGN activity.

Source Classification

We classified the sources based on multi-band photometry from ALMA. For thermal emission from dust, the submm/mm flux densities of DSFGs decrease with increasing wavelength; for synchrotron emission, on the other hand, their flux densities will increase. These emission mechanisms can therefore be well separated by their submm/mm colours. We have constructed a colour-colour plot based on ALMA bands 4, 6 and 7 for all the sources in Fig. 3.3. The three bright SMGs share a similar spectral index in the ALMA bands, which is different from the radio sources, including the blazar and jet. The other three DSFGs are only detected in bands 6 and 7, so we used their upper limits in band 4 to calculate the colours. Although they suffer from large uncertainties, their submm colours are close to those of the SMGs. In band 3, we did not detect the continuum emission of the DSFGs; also, the central band-3 region is contaminated by the blazar. For these reasons, we did not include the band-3 continuum measurements in Table. 3.3.

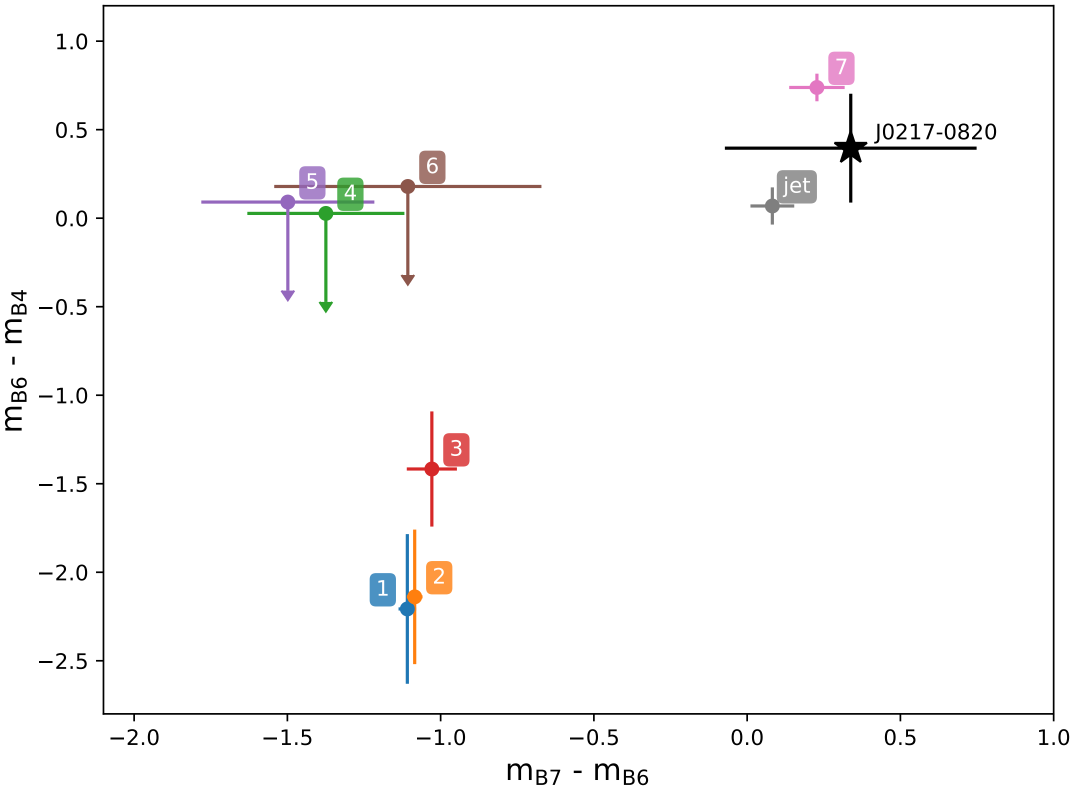

Fig. 3.3: Submm/mm colours of the detected sources. The colours of the synchrotron radio sources are quite distinct from those of the DSFGs.

Simulations

Recently, numerical simulations have successfully reproduced the number counts of DSFGs over a wide range of flux densities (Chen et al., 2023). These simulations can thus be used to test the projection effect with various setups. Depending on the adopted methodologies, different works have used different assumptions and approximations. It is therefore worth doing the tests in different types of simulations.

We have adopted the semi-analytical simulation from the SHARK project (Lagos et al., 2019,

Tab. 3.2: Measured flux of the confirmed spectral lines.

ID	CO(2–1) (Jy km s^{-1})	CO(3–2) (Jy km s^{-1})	CO(4–3) (Jy km s^{-1})	CO(5–4) (Jy km s^{-1})	CO(6–5) (Jy km s^{-1})	[CI](1–0) (Jy km s^{-1})	CO(7–6) (Jy km s^{-1})
HLW-1	–	1.47±0.08	1.70±0.15	–	2.53±0.22	0.71±0.12	1.33±0.11
HLW-2	–	1.44±0.10	1.25±0.18	–	–	1.52±0.24	–
HLW-3	1.06±0.20	1.07±0.20	1.20±0.20	–	–	0.61±0.04	–
HLW-5	–	0.72±0.20[a]	–	–	–	–	–
HLW-6	0.34±0.09	0.36±0.08	–	–	–	–	–

Notes: (a) assumes the same redshift as HLW-2.

2020), which has a cubic volume of (210 cMpc h^{-1})^3, and calculated the dust emission from the SED template ProSpect^2. We have also adopted the semi-empirical simulation from Popping et al. (2020), which plugged the observational scaling relations into UnivereMachine (Behroozi et al., 2019) and produced a simulated cubic Universe with a volume of (250 cMpc h^{-1})^3. The submm/mm flux densities of the galaxies are scaled from their SFR and dust mass (Popping et al., 2020). For the hydrodynamical simulations, we used the light cones from Lovell et al. (2021), which performed full radiative transfer simulations for all the galaxies with SFR> 20 $M\odot$ yr^{-1} in the Simba (100 cMpc h^{-1})^3 box (Davé et al., 2019). This simulation is only complete for DSFGs with $S_{850} > 0.25$ mJy, similar to the faintest DSFGs found around J0217−0820.

With all these simulated light cones, we searched for DSFGs following the same observational setups. We used the same FoV as we did in band 7 ($d = 30''$). We randomly placed the pointing within the simulated light cone and then counted the number of DSFGs as a function of flux density in each pointing. We repeated this process 1000 times to obtain the average number of DSFGs at each flux level and to sample their variations. Note that the simulations did not all model the radio emission of the galaxies, so we can only compare the results for DSFGs. We did not apply redshift constraints when searching for over-densities in the simulated light cones as the dominant population of DSFGs lies at $z > 1$ (Lagos et al., 2019; Popping et al., 2020).

Spectral lines

We measured the line fluxes of all the detected emission lines in Fig. 3.4. We extracted the spectra using a fixed aperture, twice the size of the synthesised beam. Since most of the sources remain unresolved or only marginally resolved in our images, this aperture gives the most robust results based on our tests. All the emission lines detected in bands 3 and 4 have negligible continuum emission, so we fitted the extracted spectra with a single Gaussian line profile. For the lines detected in bands 5 and 6, we fitted the line with a Gaussian profile plus a flat continuum. The resulting line fluxes are reported in Table 3.2.

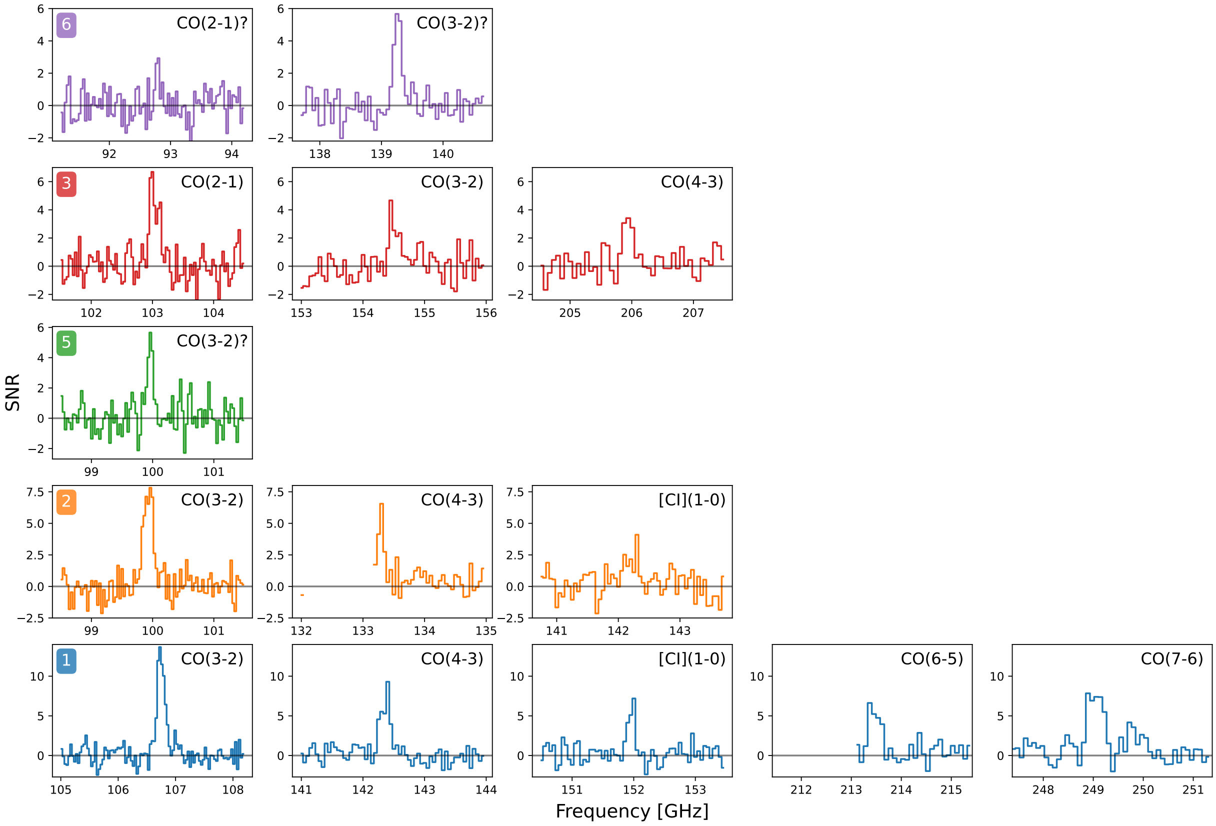

Fig. 3.4: Zoomed-in view of the spectral lines in Fig. 3.2, with additional lines found in ALMA band 5 and band 6.

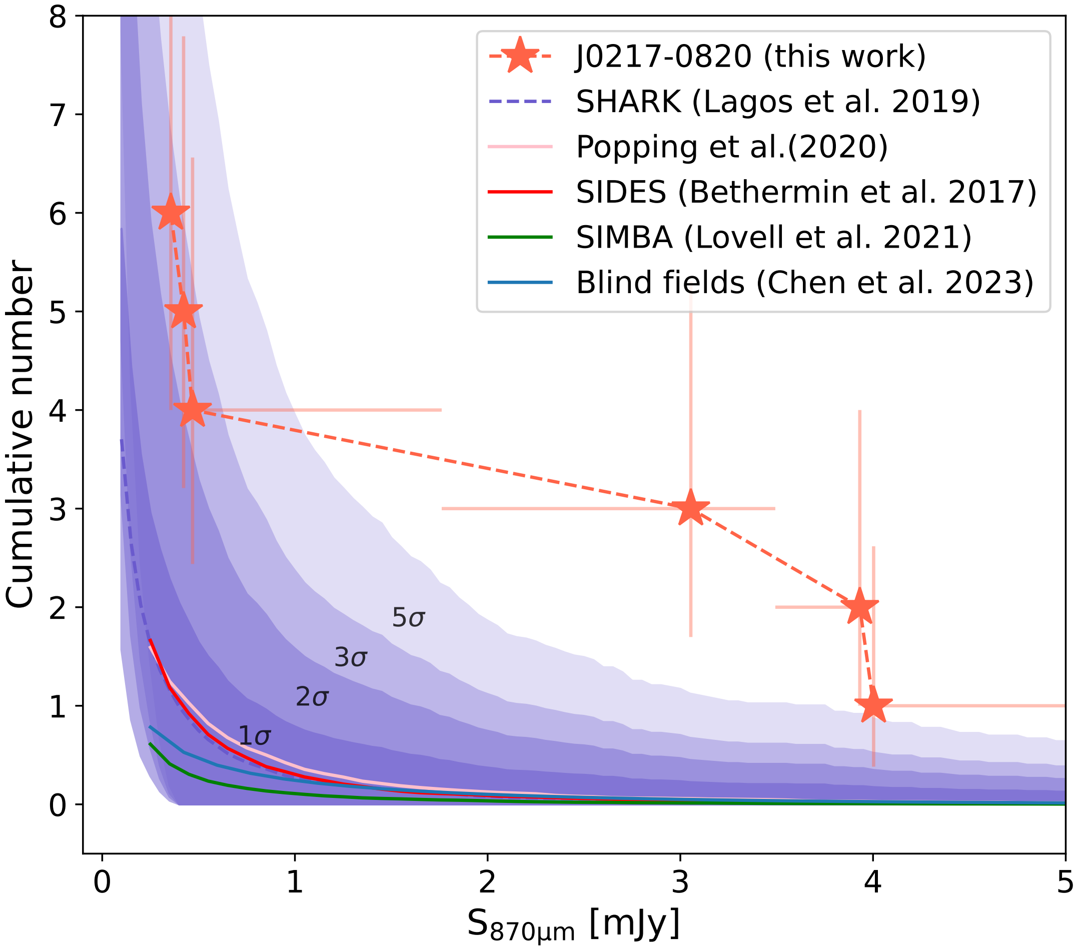

Fig. 3.5: Projected number of galaxies in various simulations. We search for galaxies following the same configuration of ALMA band 7 ($d = 30''$). We include the results from the lightcones of semi-empirical models (Bthermin et al., 2017; Popping et al., 2020), semi-analytic models (Lagos et al., 2019, 2020), and hydrodynamic simulations (Lovell et al., 2021). The shadows show the $1\,\sigma$, $2\,\sigma$, $3\,\sigma$, and $5\,\sigma$ scatter in the lightcones of the SHARK simulation (Lagos et al., 2019). Error bars indicate the Poisson errors. Within the simulations it is very rare ($> 5\sigma$) to find the number of SMGs seen around J0217−0820, but the deeper the observation goes, the stronger the projection effect.

3.2.4 Implications of the projection effects

J0217−0820 was discovered within the 45-arcmin^2 ALMACAL band 7 footprint, which is about 100 Mpc^2 (comoving coordinates) at $z \sim 2$. It is orders of magnitude smaller than the parent surveys that led to the discoveries of the two extremes, confirmed proto-cluster cores (Oteo et al., 2018; Miller et al., 2018), because we start from data with an order of magnitude better spatial resolution. The negative K correction has greatly benefited the search for DSFGs across a wide redshift range, but because of this it also amplifies the projection effects (Hayward et al., 2013; Lovell et al., 2021). Clearly, given what we have found here, over-densities such as that around J0217−0820 could contaminate the search for genuine proto-cluster cores in future deep submm/mm sky surveys.

Strong lensing could amplify the flux density of the background sources and boost the number counts in the vicinity. Around the foreground blazar, we did not find other galaxies at the same redshift or share similar colours in the Dark Energy Spectroscopic Instrument (DESI) and VISTA/VIRCam images. It is therefore unlikely that this over-density is caused by the strong lensing of a foreground galaxy cluster. Meanwhile, the morphologies of the submm sources do not support the idea that they have been strongly lensed by the blazar.

Could we find such alignments in cosmological simulations? To address this question, we have made mock lightcones from recent simulations that predict submm/mm emission characteristics, then searched those lightcones for projected over-densities. We adopted simulations from semi-analytical models (SAMs – Lagos et al., 2019, 2020), semi-empirical models (Bthermin et al., 2017; Popping et al., 2020) and hydro-dynamical simulations (Lovell et al., 2021) (see more details in Appendix 3.2.3). The results are presented in Fig. 3.5.

We first note that almost all of the simulations predict much lower numbers of traditional, bright SMGs, with a significance higher than 5σ. This inconsistency could indicate that what we have found here is a real, rare alignment. However, the failure to reproduce the number density of SMGs has long been reported in the literature. Possible explanations are rooted in the stellar initial mass function (e.g. Baugh et al., 2005; Lacey et al., 2016), stellar/AGN feedback (Lovell et al., 2021; Hayward et al., 2021) and various dust-related models (including production, destruction and temperature; see more discussion in Safarzadeh et al., 2017; Lagos et al., 2019; Lovell et al., 2021). Therefore, if such chance alignment is much higher than what has been predicted, this discrepancy may also suggest that we could have under-predicted the cosmic variance of submm surveys.

At the fainter end, the probability of finding a similar overdensity is much higher than at the bright end. At $S_{870\mu m} \sim 0.1$mJy, the chance of finding more than six DSFGs is close to 5 per cent. Considering only the Poisson process, the probability is only around 0.1 per cent, which indicates that clustered structures in the Universe contribute to the projection effects. This effect should be strongest at submm/mm wavelengths, due to the large redshift space probed in the submm/mm windows. This cautions against the use of a simplistic statistical approach to select proto-clusters and quantify cosmic variance in deep submm surveys.

In summary, although J0217−0820 is an extremely over-dense region – where its over-density, $\delta_{\rm gal} \simeq 8$, is comparable to known, extreme proto-cluster cores – we have determined

---

[2] https://github.com/asgr/ProSpect

that this is almost entirely due to chance alignments. In future deep galaxy surveys – those with ALMA in the submm/mm bands – similar projection effects may contaminate the search for proto-clusters and contribute to the cosmic variance.

3.3 ALMARED survey for proto-cluster cores

3.3.1 Motivation of ALMARED

Proto-cluster core is a new sighpost of proto-cluster. The search for the proto-cluster is severely hampered by the volume size of the proto-cluster. (Ata et al., 2022) demonstrated that the spatial size of the proto-cluster at $z \sim 2$ is almost 10 times larger than the final size of the cluster at $z \sim 0$ at each dimension. However, within the gravitational centre of the proto-cluster, it could have developed a massive dark matter halo and concentrated a substantial amount of the cold gas. They are supposed to evolve much quicker and earlier than the field galaxies. Meanwhile, the starburst activity could make them stand out in the submm/mm sky surveys. Therefore, a submm sky survey covering hundreds of $\deg^2$ is the key to finding these systems.

Of the existing confirmed proto-cluster cores (Oteo et al., 2018; Miller et al., 2018), the number density of the proto-cluster core is excessively low. DRC was found through a total 600 $\deg^2$ footprint of the H-ATLAS survey (Ivison et al., 2016a); while SPT2349-56 was discovered by a 2500 $\deg^2$ footprint of the SPT-3G survey (Miller et al., 2018). However, these estimates can only be used as a lower limit, as the current selection criteria may have missed many potential candidates. In addition, the current estimate suffers from the small number of statistics. Therefore, the next crucial step to go is to find more such systems. Fortunately, due to the overdensity of the DSFGs in such systems, their total submm flux is among the brightest SGMs in the single-dish sky surveys. However, the strongly lensed sources or a single HyLIRG could also share a similar flux density as the proto-cluster core.

Inspired by the discoveries of Oteo et al. (2018), we have launched a massive ALMA 1.3-mm continuum snapshot survey of 3083 ultrared *Herschel* sources, with $S_{500\mu\mathrm{m}} > S_{350\mu\mathrm{m}} > S_{250\mu\mathrm{m}}$ and $S_{500\mu\mathrm{m}} > 45\,\mathrm{mJy}$ – the so-called '500-μm risers' (Ivison et al. 2016a; Ivison et al. in prep.). The requirement on the three *Herschel* bands ensures that we mainly selected the peak of the cold dust emission at $\lambda \geq 400\mu$m, which corresponds to a redshift of about $z \geq 3$, assuming a dust temperature of around 30 K. Compared to Ivison et al. (2016a), we have mainly released the constraint of spectral index, allowing the selection of more candidates with redshifts down to $z \sim 3$. A key feature of these ALMA observations is the spatial resolution of ∼ 1". It can be used to separate the brightest member galaxies from the strongly lensed galaxies. Furthermore, based on the two confirmed systems: DRC and SPT2349-56 (Oteo et al., 2018; Miller et al., 2018), the 3-5 brightest member galaxies can be well detected with 20s ALMA on-source time in band-6. A more detailed description of the source properties, including their SED, multiplicities, and morphologies, can be found in Ivison et al. (in prep.). In this section, I will mainly focus on the 12 fields with multiplicity larger than 3 and photometric redshift $z_{\mathrm{photo}} > 3$, based on the existing multi-band photometry from ALMA and *Herschel*. All these fields have been observed with ALMA band-3 spectral scan and band-7 continuum imaging.

3.3.2 Data calibration and analysis

The observations were made between December 2018 and March 2019, under the cycle-8 project 2021.1.00018 (P.I. R. J. Ivison). Four tunings were used to cover the whole 8 GHz bandwidth of ALMA band-3 on each target. All the data were calibrated and analysed by the standard ALMA pipeline v6.2.1-7. I used the official task TCLEAN to make the continuum image and spectral datacubes. For the continuum images, I used natural weighting to achieve a higher sensitivity for the continuum sources. For band-3 data, in addition to the continuum image, I also created the spectral datacube with two different spectral and spatial resolutions. The low-resolution datacubes were created with natural weighting and spectral resolution of 100 km s^{-1}; while the high-resolution datacube were created with Briggs weighting (robust=0.5) and spectral resolution of 50 km s^{-1}. These low-resolution data products were later used for flux measurements of the spectral lines; while the high-resolution datacubes were mainly used for kinematics and morphology analysis. During the cleaning, I also used the 'usemask=auto-multithresh' with the default setting to aid the masking of possible sources in the image.

For the continuum images, I used SEXTRACTOR to search for possible sources and measure the shape of the sources. I searched for the sources in the images before the primary beam response correction, where the whole map is more uniform. Then, having obtained the position of the sources, I used aperture photometry to measure their flux density. The aperture was three times larger than the best-fitted ellipticals by SEXTRACTOR. I also randomly placed the same aperture 100 times within the images to derive the error of the aperture photometry. Then, I quantified the aperture loss and flux boost by injecting artificial point sources into the observed image. This process was repeated 100 times for each map to derive the median correction factor between the injected and measured values. After correcting for the aperture loss and flux boost, I also corrected the primary beam response based on the ”*.pb” maps from TCLEAN. In several fields, I also find close merging galaxies. For these sources, I manually fit two 2D-Gaussian to separate the flux densities of the two nearby sources to disentangle the flux contamination.

For the spectral lines, I first extracted the spectral lines at the position of the continuum sources. I extracted only the spectra from the pixel with the highest continuum emission. I searched for 5σ peaks in the extracted spectra. After obtaining all the 5σ spectral peaks, I first assumed that the strongest peak is one of the CO lines or [C I] and then searched for the other lines at the expected frequencies. If only one line is detected, or if there is a degeneracy between the CO lines, I determine the single-line redshift by combining it with our photometric redshift. I also searched the 5σ line peaks throughout the whole datacube for line emitters without continuum emission.

The line flux was measured by collapsing all the channels associated with the spectral line. Then, I used aperture photometry in the same way as for the continuum sources. I also created the moment 1 (M1, velocity) and moment 2 (M2, velocity dispersion) maps of each target. The two moments were created following the formula:

$$\mathrm{M1} = \frac{\int v I_v dv}{\int I_v dv}$$

$$\mathrm{M2} = \frac{\int I_v (v - \mathrm{M1})^2 dv}{\int I_v dv}$$

In the equations, v is the velocity of the channel relative to reference channel. I_v is the observed flux density in each pixel.

Before applying these two formulas, I first created the 3D mask of the datacube (see §3.5). The 3D mask was created to select the pixels containing true emissions. I first selected all the pixels that were 2σ above the RMS of each channel. Then, I applied the binary opening to the 2σ structure to remove the random 2σ peaks associated with the random noise. After this, a cycle of binary dilation was applied to expand the existing binary structure to include the regions with weak signals. I used the tools packaged in SCIPY.NDIMAGE to conduct all the mathematical morphology operations, including opening and dilation. After the binary opening and dilation, I labelled all the existing structures in the 3D datacube with SCIPY.NDIMAGE.LABEL. I then iterated all the structures to remove all the structures with less than 25 pixels, which is less than the FWHM of the synthesised beam. The final remaining structures in the 3D cube were re-labelled and used to extract pixels and channels to create the moment maps.

3.3.3 A new sample of proto-clusters cores

We selected proto-cluster cores based on the overdensity of the SMGs. We defined the proto-cluster core to have $n_{\mathrm{SMG}} \geq 3$ in the FoV of ALMA band 6 (d=30”). From our simulation in 3.2.3, it is very unlikely (¡0.1%) to have three SMGs within the FoV produced solely by the projection effect. $n_{\mathrm{SMG}} = 2$ could also indicate the presence of massive galaxies, but it expected be more common as most SMGs were triggered by galaxy mergers. Therefore, in this study, I will only focus on the massive halos consists of at least three SMGs. Based on the analysis in §3.3.4, I will show that this definition does lead to the discovery of extreme proto-clusters.

Based on my empirical definition of proto-cluster, we have confirmed five proto-cluster cores from our ALMA band-3 spectral scan. I include all the band 3 continuum images and the spectra of the possible member galaxies in Fig.3.7 and Fig.3.6. In each field, at least three SMGs have the same redshift. One of them is the DRC, which has been discovered before (Oteo et al., 2018).

Our ALMA spectral scan confirmed again the strong projection effects in the submm/mm surveys. Our candidates were selected from band 6 overdensity, but the final spectral scan only confirmed five of them are true proto-cluster cores. The other seven fields are all misclassified as the proto-cluster due to the projection effects. Therefore, I emphasise again the importance of spectral follow-up of the overdensity in confirming the proto-cluster cores.

3.3.4 Properties of the proto-cluster cores

Total molecular gas

The estimation of the total molecular gas in DSFGs is highly uncertain. The cold phase of H_2 is not directly observable due to the lack of permanent dipole moment. Its quadrupole transitions can only be excited with $T_{\mathrm{ext}} > 500\,\mathrm{K}$ and they are weak. As a consequence, observations can

Tab. 3.3: Observed properties of the proto-cluster cores.

ID	J2000	Band 3 mJy	Band 6 mJy	Band 7 mJy	CO(4-3) Jy km s^{-1}	[C I](2-1) Jy km s^{-1}
AR1635						
1	01h27m23.4s -30d58m09.4s	0.16±0.02	2.76±0.26	8.36±0.36	0.90±0.30	< 0.60
2	01h27m24.0s -30d58m07.1s	0.03±0.01	1.21±0.25	2.42±0.26	0.29±0.30	< 0.66
3	01h27m23.5s -30d58m05.3s	0.13±0.02	2.78±0.32	7.14±0.33	0.69±0.31	0.64±0.21
4	01h27m23.0s -30d57m50.4s	0.06±0.01	–	–	0.73±0.41	< 0.78
AR2440						
1	23h59m00.0s -35d00m06.6s	0.10±0.02	2.76±0.42	7.76±0.40	1.16±0.19	0.82±0.16
2	23h58m59.5s -35d00m04.2s	0.07±0.02	1.62±0.34	4.10±0.28	0.89±0.21	0.57±0.21
3	23h58m59.3s -35d00m03.8s	0.07±0.02	0.75±0.37	4.98±0.32	1.12±0.20	0.27±0.19
AR2623						
1	12h24m59.3s -00d56m52.3s	0.17±0.02	3.40±0.31	10.80±0.39	1.97±0.22	0.89±0.27
2	12h24m58.7s -00d56m48.3s	0.26±0.03	4.31±0.47	12.87±0.34	2.74±0.22	1.19±0.25
3	12h24m59.0s -00d56m47.5s	0.14±0.02	2.68±0.32	9.26±0.31	1.35±0.21	0.42±0.27
4	12h24m59.2s -00d56m40.9s	0.04±0.02	1.44±0.29	3.00±0.40	< 0.69	1.11±0.25
AR2747						
1	13h38m08.3s +25d51m48.6s	0.20±0.03	0.79±0.26	2.77±0.32	1.24±0.47	–
2	13h38m09.0s +25d51m54.6s	0.12±0.03	3.12±0.39	10.05±0.69	2.00±0.44	–
3	13h38m09.2s +25d51m56.1s	0.16±0.03	1.83±0.25	5.87±0.39	2.21±0.53	–
4	13h38m08.2s +25d52m09.8s	0.21±0.04	2.82±0.28	7.42±0.47	3.13±0.61	–

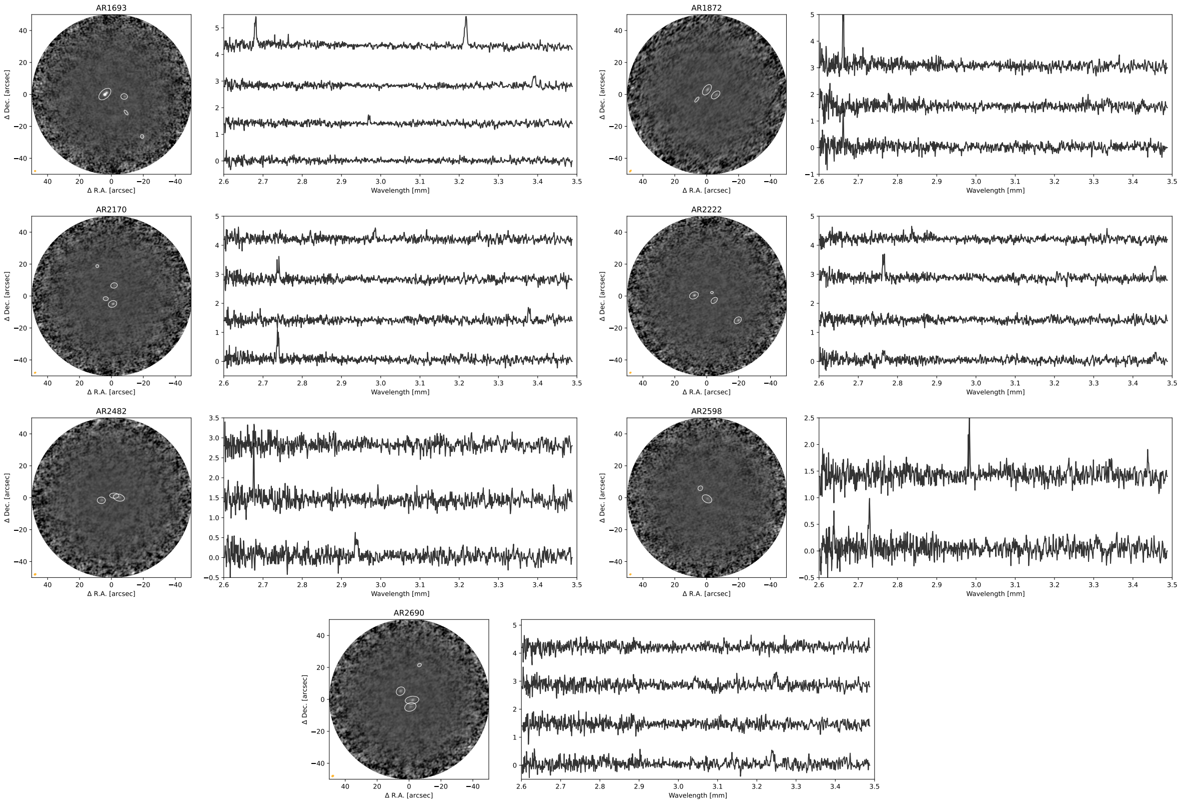

Fig. 3.6: The fields with overdensity are produced by the projection effect. In each subpanel, the left image shows the band-3 continuum image without the spectral lines; the right spectra were extracted from the pixel with the peak continuum emission. The spectra from bottom to the top are corresponding to the continuum sources from the bottom to the top. Among them, only AR1872 and AR2170 both have two galaxies at the same redshifts.

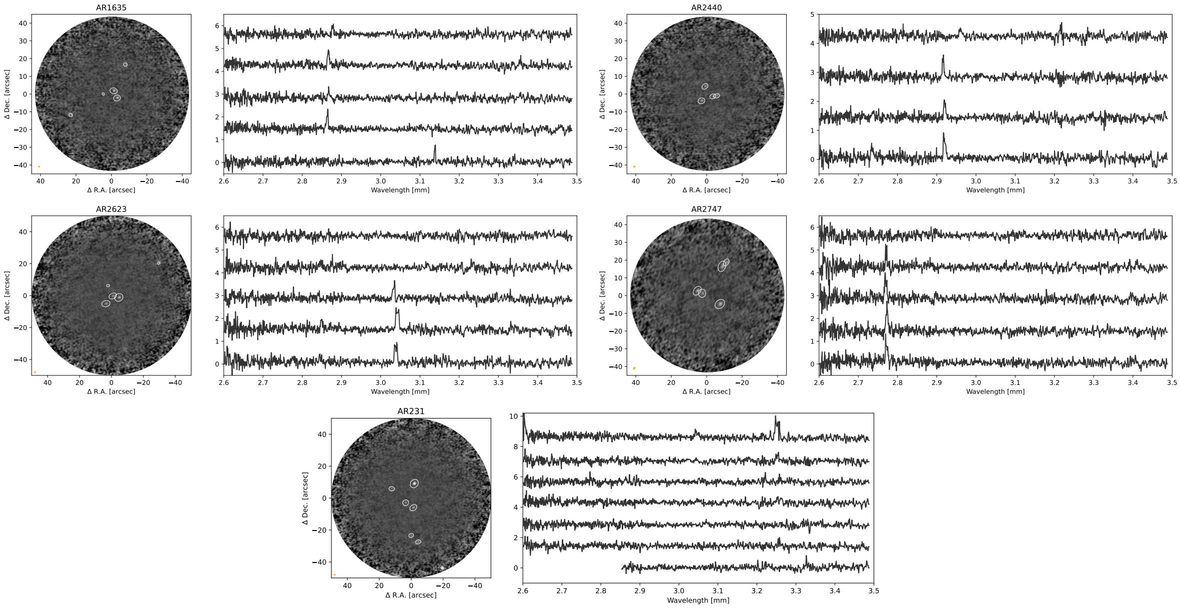

Fig. 3.7: The fields include three and more than three DSFGs at the same redshift. In each subpanel, the left image shows the band-3 continuum image without the spectral lines; the right spectra were extracted from the pixel with the peak continuum emission. The last target (AR231) is the previously confirmed DRC (Oteo et al., 2018; Ivison et al., 2020). We have detected more than half of its member galaxies. It confirms the effectiveness of our combined efforts with snapshot surveys and spectral scans.

Tab. 3.4: The physical properties of the member galaxies in the proto-cluster cores.

ID	z	$M_{\mathrm{molgas,CO}}$ $\times 10^{11}\mathrm{M}_{\odot}$	$M_{\mathrm{molgas,850}\,\mu\mathrm{m}}$ $\times 10^{11}\mathrm{M}_{\odot}$	$M_{\mathrm{molgas,[C\,I](1-0)}}$ $\times 10^{11}\mathrm{M}_{\odot}$	SFR $\mathrm{M}_{\odot}\,\mathrm{yr}^{-1}$	t_{dep} Myr^{-1}
1	-	1.6±0.5	1.8±0.2	< 1.8	1400	110
2	-	0.5±0.5	0.3±0.1	< 1.8	400	120
3	-	1.2±0.6	1.4±0.2	1.8±0.6	1260	100
AR1635 (total)	4.509	3.3±0.9	3.5±0.3	–	3120	110
1	-	1.7±0.3	1.6±0.3	1.6±0.3	1200	150
2	-	1.3±0.3	1.0±0.3	1.1±0.4	640	206
3	-	1.7±0.3	1.1±0.3	0.5±0.4	730	230
AR2440 (total)	3.490	4.7±0.5	3.8±0.5	3.2±0.6	2560	190
1	-	3.2±0.3	2.5±0.3	1.8±0.5	1670	190
2	-	4.4±0.4	3.8±0.4	2.5±0.5	2090	210
3	-	2.2±0.3	2.0±0.3	0.9±0.6	1450	150
4	-	0.2±0.4	0.5±0.2	2.3±0.5	470	40
AR2623 (total)	3.663	9.9±0.7	8.8±0.6	7.4±1.1	5680	150
1	-	1.6±0.6	3.2±0.5	–	1500	110
2	-	2.7±0.6	2.0±0.5	–	890	300
3	-	2.9±0.7	2.6±0.5	–	1170	250
4	-	4.2±0.8	3.4±0.6	–	1570	270
5	-	1.1±0.9	2.2±0.6	–	990	110
AR2747 (total)	3.260	12.5±1.7	13.4±1.2	–	6100	210

only be performed with other cold gas tracers such as CO and [C I], and a scale relation is needed to apply the conversion. However, most SMGs do not have available CO(1-0) observations, so a further conversion from the high-J transitions to the ground-level CO(1-0) is needed. In our four systems, we measured the molecular gas based on the available high-J CO emissions. In AR1635, it is CO(5-4). In the rest systems, we used CO(4-3). We first scaled the observed CO lines to the ground-level CO, following the conversion factor from (Tacconi et al., 2020). Then, we calculated the total molecular gas using the CO(1-0) to H_2 conversion factor of $\alpha_{\mathrm{CO\,1}} = 0.8$ and the formular from (Tacconi et al., 2020).

$$\frac{M_{\mathrm{molgas}}}{M_\odot} = 1.58\times10^{9}\left(\frac{\alpha_{\mathrm{CO\,1}}\times R_{1\mathrm{J}}}{\alpha_0}\right)\times\left(\frac{F_{\mathrm{CO\,J}}}{\mathrm{Jy\,km\,s^{-1}}}\right)\times(1+z)^{-3}\times\left(\frac{\lambda_{\mathrm{obs\,J}}}{\mathrm{mm}}\right)^{2}\times\left(\frac{D_{\mathrm{L}}}{\mathrm{Gpc}}\right)^{2}$$

In the formular, $\alpha_0 = \mathrm{M_\odot\,K\,km\,s^{-1}\,pc^{2^{-1}}}$ is the unit of the conversion factor. $R_{1\mathrm{J}}$ is a correction from CO SLED. We used $R_{1\,4} = 5.6$ for CO(4-3) and $R_{1\,5} = 7.0$ for CO(5-4). The results of the calculations are summarised in Tab.3.4.

It is not easy to estimate the error of these calculations. Statistically, the scatter of $H_2 \sim L_{\mathrm{CO(1-0)}}$ is only 0.1 dex within Milky Way's disk, but significantly larger in starburst galaxies, $\sim$ 0.5 dex (Bolatto et al., 2013). The α_{CO} is most reported to be around 0.8 in the DSFGs (e.g. Ivison et al., 2011), which we also adopted here. The correction of the CO rotational ladder is also uncertain, but it should most within 0.2 dex (Tacconi et al., 2020). Therefore, the statistical error of our estimation should be within 0.8 dex.

I also cross-validated our measurement from other methods. Firstly, almost all our targets have the dust continuum measurement around 3 mm, which is also statistically correlated with the total mass of molecular gas (Scoville et al., 2016; Tacconi et al., 2020; Dunne et al., 2021). I converted the dust continuum to molecular following the formulas in (Dunne et al., 2021). I first convert the 3 mm flux density to the rest-frame 870 μm, assuming a typical dust temperature of 38 K and dust emissivity of $\beta = 1.8$ recommended for SMGs (Dunne et al., 2021). The derived total molecular gas is shown as $M_{\mathrm{molgas,850\mu m}}$ in Tab.3.4, which are in the order of $10^{10} - 10^{11}\mathrm{M_\odot}$ and generally consistent with our CO-based calculation within 0.2 dex.

A third of our sources also have secure [C I](1-0) measurements. I used the [C I](1-0) to cross-check the accuracy of the total molecular gas. I followed the scaling relation in Dunne et al. (2021). I used Q_{10}=0.48 to account for the non-LTE excitation of [C I](1-0). I also adopted $X_{\mathrm{CI}} = 3\times10^{-5}$ recommended for SMGs (Dunne et al., 2021). The [C I](1-0)-based measurements are also shown in Tab.3.4. They are also broadly consistent with the measurement from CO and dust continuum.

In addition, for the member galaxies with strong CO emissions, we can also measure their rotational velocities. Based on the assumption that they are rotating disks, we can derive the dynamic mass of the member galaxies. Their maximum velocities are ranging from 300 − 400 km s^{-1}, see more detail from Fig.3.13 to Fig.3.16. Based on the simple Kaplerian disk, the dynamic mass can be derived following $M_{\mathrm{dyn}} = v_{\mathrm{max}}\,R/G$. In our observations, the typical radius of the disk is about $1 - 1.5''$ ($\sim 7 - 10$ kpc in physical size). Then, the estimated molecular gas is about $1.5 - 4\times10^{11}\mathrm{M_\odot}$, which is consistent with CO-based calculation. We did not correct the inclination angle, which could further increase the total dynamic mass. In summary, there are still large uncertainties in the measurement of the total molecular gas, but different independent methods give consistent results.

Star-formation rate

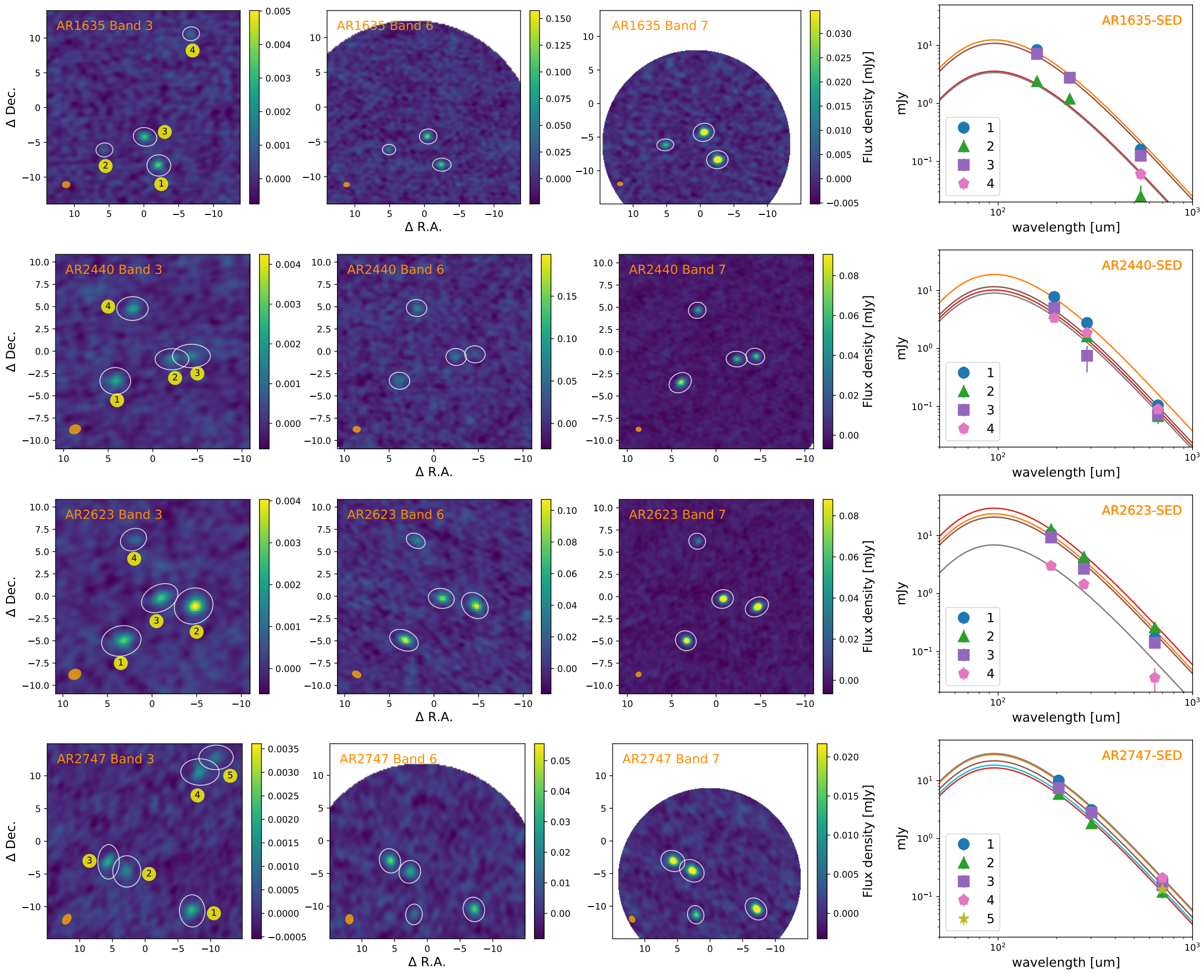

Fig. 3.8: The continuum images of the fource new proto-cluster cores. From left to right, it shows the continuum images at ALMA band 3, band 6 and band 7. ALMA band 3 has the largest FoV. I have labelled the member galaxies in each proto-cluster core in the band 3 continuum image. In the last column, I show the SED extracted from the three ALMA continuum observations and the best fits based on the median SED of the AS2ADS (Dudzevičiūtė et al., 2020).

We estimated the SFR based on their FIR SED. Due to the limited data available, we fitted their photometries from three ALMA bands with the median SED of the SMGs in AS2UDS (Dudzevičiūtė et al., 2020). Then, I derived the SFR from the SFR of the median SED (see Fig.3.8). The derived values are summarised in Tab.3.4.

All the member galaxies are extreme starburst galaxies. They are quantitatively similar to the top brightest members in DRC and SPT2349-56 (Oteo et al., 2018; Miller et al., 2018). Due to our selection bias, we are still missing the galaxies with much lower SFR or dust extinction. Therefore, their total SFR can be considered as a lower limit. In the future, follow-up observations with with larger FoV will reveal the larger structures around the proto-cluster cores.

Gas depletion timescales

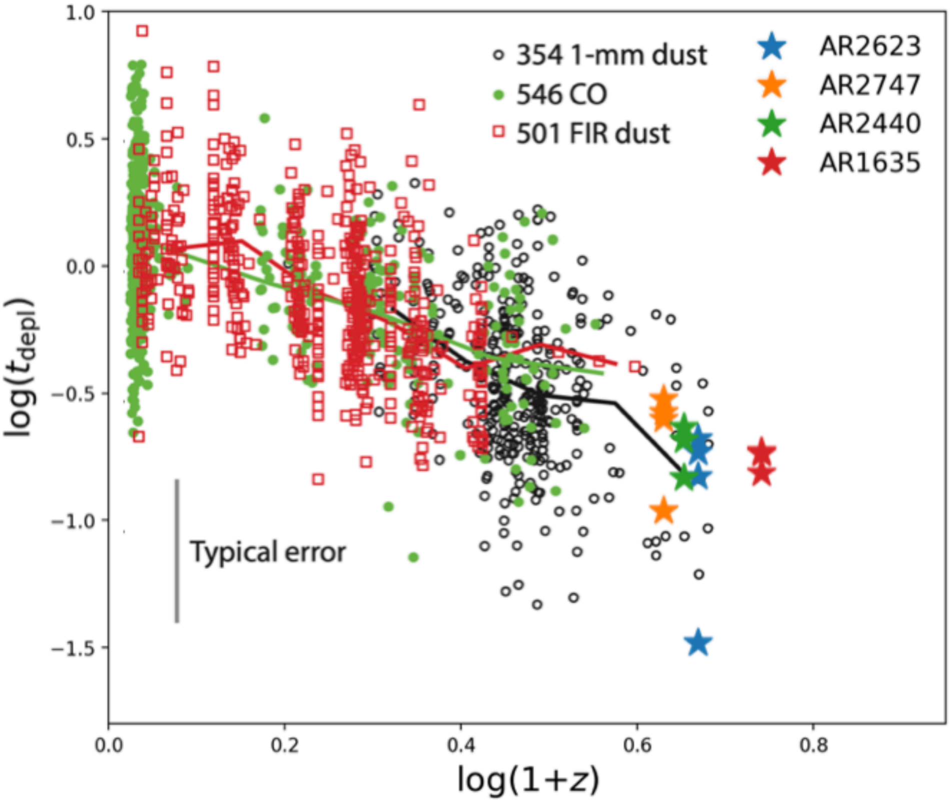

Fig. 3.9: The gas depletion timescale of the member galaxies in newly confirmed proto-cluster cores. The galaxies from each proto-cluster core are marked with a different colour. I also included the gas depletion timescale of the star-forming galaxies summarised in (Tacconi et al., 2020). The gas depletion time scales of the member galaxies follow the general evolutional trend of the star-forming galaxies as a function of redshift.

Based on our estimates of the total SFR and molecular gas in the member galaxies of the proto-cluster cores, we can also constrain their gas depletion timescale. I show our results in Fig.3.9. In general, the gas depletion timescale in these overdense regions is shorter than those of the typical main-sequence galaxies, indicating an efficient accumulation of stellar mass. Meanwhile, they still follow the evolutional trend as a function of cosmic time. (Tacconi et al., 2020) argued that the main driver of the gas depletion timescale connects with the global transition of cosmic star formation over cosmic time. Our results possibly indicate most of the early star formation activity before cosmic noon could happen within this overdense region.

Kinematics

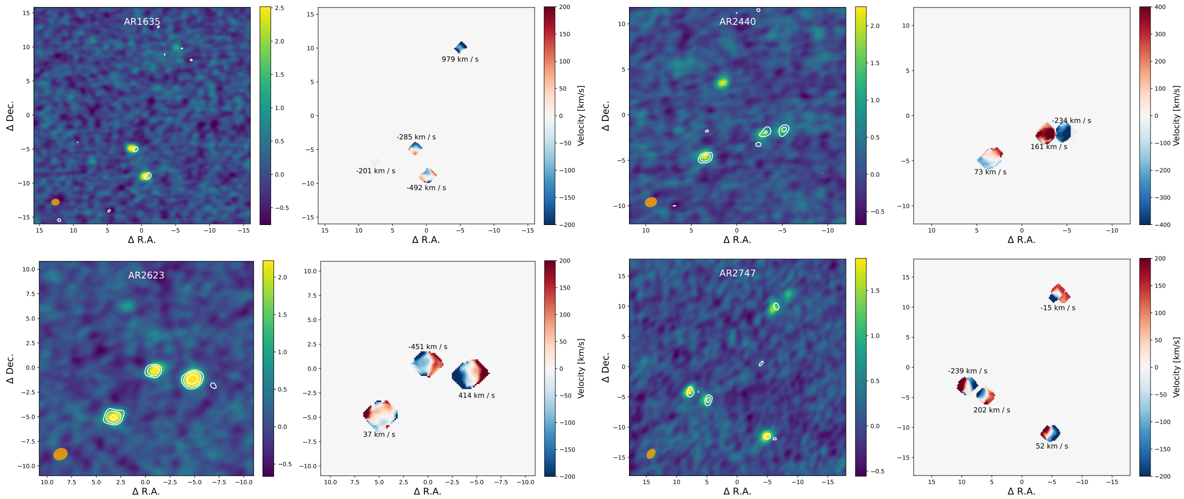

Fig. 3.10: The velocity maps of the four new proto-cluster cores. In each subpanel, I show the band-3 continuum image on the left and the velocity maps of each target on the right. The white contours show the measurement of the corresponding CO emissions, starting from 2 σ above the RMS level. For the member galaxies with strong line emission, most of them show the pattern of rotation disk. Due to our limited resolution, we cannot distinguish the close mergers from the ordered rotating disks.

The kinematics of all the four fields are shown in Fig.3.10. For each target, the velocity maps are derived relative to its redshift. The velocity of the whole galaxy relative to the mean redshift of the proto-cluster core is denoted by the velocity right close to the individual galaxies. Our initial observing setups were not optimised to resolve the kinematics of the galaxies, as we have asked for compact configurations to maximise the detecting sensitivity. However, in the fields of AR2623 and AR2747, the spectral lines in the member galaxies are strong enough to infer their kinematic structures. The velocity maps generally support the presence of the galaxy disk, similar to the results reported by Xiao et al. (2022) in the proto-cluster core CL-J1001. Due to the limitations of our spatial resolution, we cannot fully exclude the possibility of close mergers. Future high spatial resolution observations would be helpful to fully unveil their kinematic structures and the possible presence of close mergers.

Do we witness the formation of the brightest cluster galaxies?

As its name indicates, the proto-cluster is thought to be the progenitor of the local massive clusters. In the centre of the clusters, many of them are populated by an excessively bright and massive elliptical galaxy, which is called the brightest cluster galaxy (BCG). Based on the assumption that the proto-cluster core traces the early mass centre of the ancient proto-clusters, it is natural to connect the proto-cluster core with the ongoing formation of BCGs. In this section, I will provide three pieces of evidence that the proto-cluster cores are very likely to be the progenitors of BCGs.

The total stellar mass

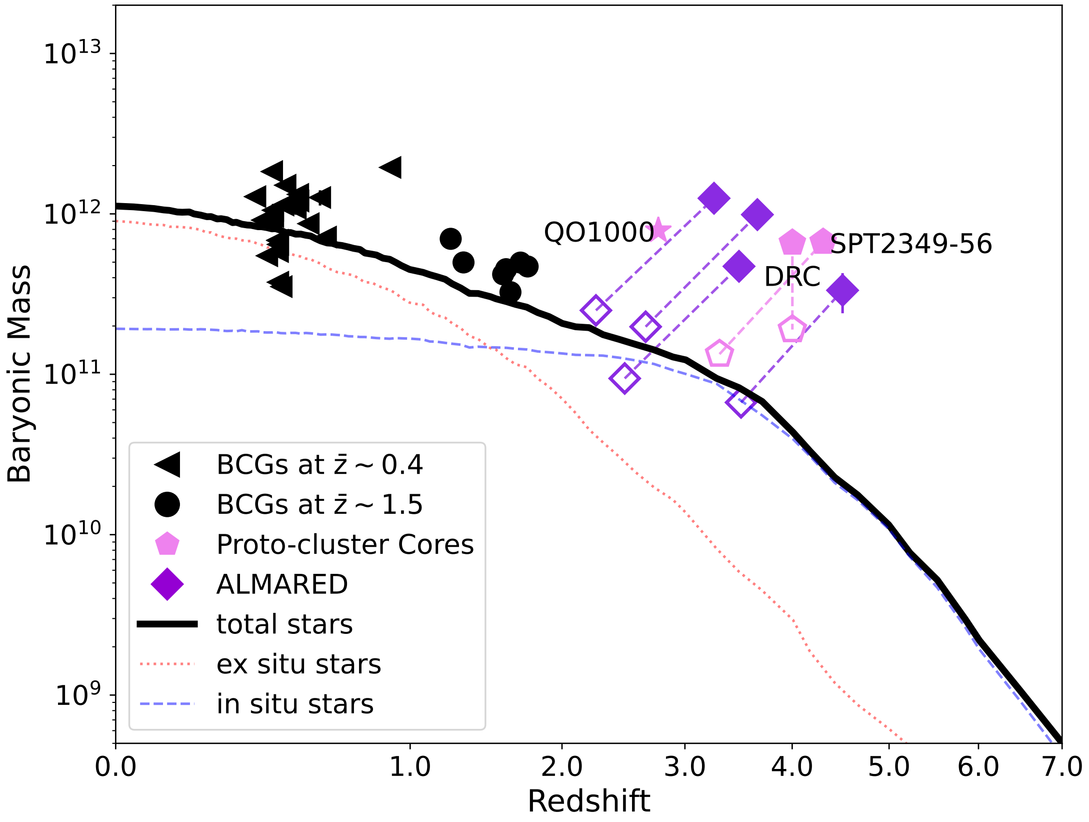

Fig. 3.11: The baryonic mass evolution of the BCGs. The black solid line shows the stellar mass growth as a function of the redshift from the IllustrisTNG-100 (Montenegro-Taborda et al., 2023). The red dotted line and the blue dashed line show the components of the ex-situ and in-situ stars. The coloured shadows show the 1σ deviation. I also show the stellar mass measurement of the low-z BCGs from (DeMaio et al., 2018, 2020). Besides the clusters, I also show our measurement of the proto-cluster cores. The total mass of the molecular gas is shown as the solid violet diamond and the hexagon. The hollow version of the diamond and hexagon shows the predicted stellar mass in these cores if 20% of their current total molecular gas can be converted into the stellar mass. This comparison supports that all the proto-cluster cores will finally become BCGs.

How the BCGs develop their stellar mass is not fully settled. In the low redshift range, $z < 2$, the BCGs were found to accumulate their stellar mass mainly via the merger with the infalling galaxies (DeMaio et al., 2018, 2020). However, at $z > 2$, we currently lack observational evidence about their formation.

Numerical simulations provide a useful tool to track the evolution of stellar mass across the formation history of galaxy clusters (Montenegro-Taborda et al., 2023). In the simulation, the dominant channel to support the stellar mass growth at $z > 2$ is the in-situ star formation. To accumulate $\sim 10^{11}\mathrm{M}_\odot$ within the first 2 Gyr, the averaged SFR should be close to $100\,\mathrm{M}_\odot\mathrm{yr}^{-1}$. Considering the episodic nature of the starburst and the time needed for the gas cooling and infalling, the simulation predicts the presence of one or several periods with SFR $> 1000\mathrm{M}_\odot\mathrm{yr}^{-1}$. Meanwhile, these starburst cores need to site inside the overdense regions to support the subsequent mergers. In our proto-cluster cores, the observational results are consistent with all these predictions. Their total SFR is $> 3000\,\mathrm{M}_\odot\mathrm{yr}^{-1}$ and harbour $\sim 10^{12}\,\mathrm{M}_\odot$, which makes them the best place to nourish the BCGs. Moreover, the timescale and frequency of the starbursts are sensitive to the AGNs models (De Lucia & Blaizot, 2007). Our current statistics show that most of the proto-cluster cores reach their maximum SFR between $z = 3 - 4$, which seems to support the AGNs feedback is not strong in proto-cluster cores at $z > 3$.

The typical stellar mass of the BCG at $z \sim 1.5$ within the central 100 kpc is about $4.8 \pm 1.0 \times 10^{11}\mathrm{M}_\odot$. This is close to half of the total molecular gas in our sample. In Fig.3.11, I overplot our observation above the evolution of the stellar mass of BCG in the IllustrisTNG-300 (Montenegro-Taborda et al., 2023). We currently lack the direct measurement of the stellar mass in our targets, but if these systems are gas-dominated and if they can convert 20% of their molecular gas into stars in the next ~1 Gyr, their stellar mass will be consistent with the mass growth of BCGs in the simulation. The star formation efficiency $\epsilon \sim 0.2$ is still higher than most of the normal star-forming galaxies, but it is thought to be higher in the starburst systems.

The evolution of the halo mass

If the proto-cluster cores do trace the formation galaxy clusters, then their halo mass should also follow the redshift evolution of the galaxy clusters. I first converted the total molecular gas to the halo mass based on the general 5% cold gas fraction in the dark matter haslo. I show the predicted halo mass of our proto-clusters cores in Fig.3.12. These new proto-cluster cores generally follow the halo mass evolution as the Coma- or Virgo-like cluster (Chiang et al., 2013), which will turn into the most massive clusters at $z \sim 0$. Since these cores have already accumulated the majority of the baryonic mass, it is thus reasonable to conclude that these cores will finally become the BCG.

The number density of proto-cluster cores

The typical number density of the Virgo-like cluster is about $10^{-8}(\mathrm{h/Mpc})^3$ (Bahcall & Cen, 1993; Casey, 2016). All the ALMARED targets were selected from the *H*-ATLAS Data Release 1 (Valiante et al. 2016, Ivison et al. in prep.), covering roughly 600 deg^2. Our selection criteria (see more detail in §3.3.2) are sensitive to the targets with $z > 3$. Based on the redshift range of our detections $3 < z < 4.5$, the total cosmic volume is about $\sim 1 \times 10^{10}\,\mathrm{Mpc}^3$. Including DRC, we have detected 5 proto-clusters cores, which results in a volume density of the $n_{\mathrm{core}} > 5 \times 10^{-10}\,\mathrm{Mpc}^{-3}$. We can only derive the lower limits as there may be more proto-cluster cores in the *H*-ATLAS footprint but have been missed by our sample selection. The estimated number density of the proto-cluster cores is lower than the Coma-like cluster, which indicates we are

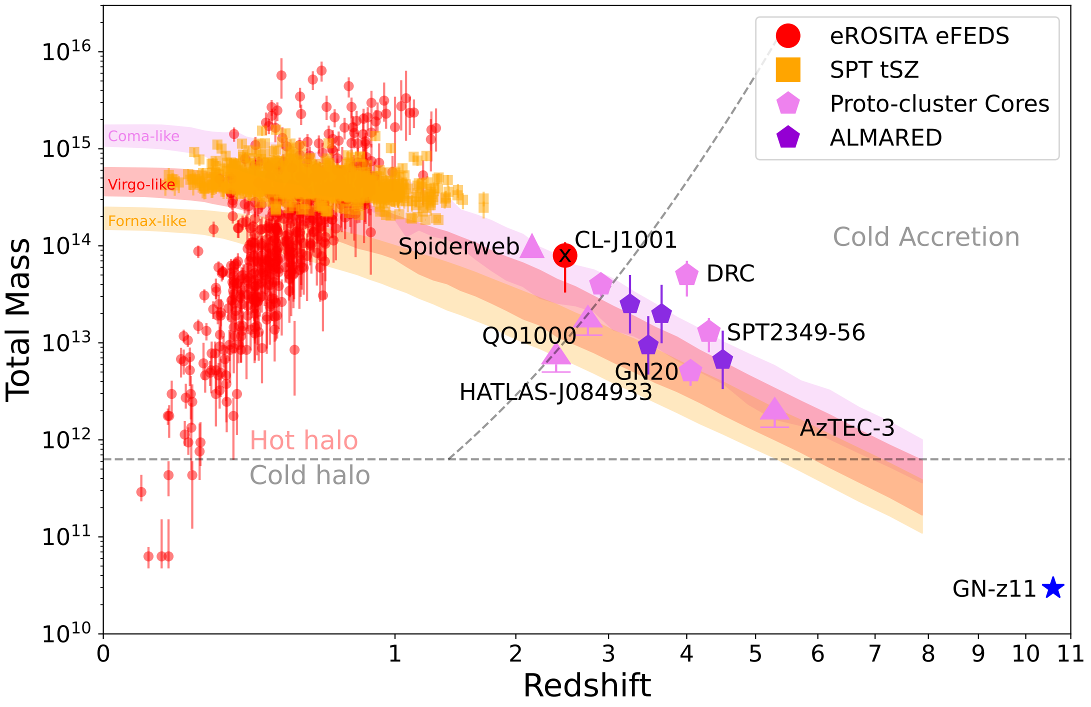

Fig. 3.12: The positions of proto-cluster cores found by ALMARED on the halo mass function. By assuming a general 5% baryonic mass within the halo, the predicted halo mass of the proto-cluster cores follows the evolution of the most massive halos from the hydrodynamical simulations (Chiang et al., 2013). This gives another evidence that we are witnessing the formation of the massive core of the galaxy clusters.

witnessing the emergence of even rare density peaks.

Indication of the proto-cluster cluster connections

The proto-cluster cores reported here are all very concentrated (< 1 cMpc) and are almost certainly will merge into a single bright cluster galaxy. They are different from the proto-clusters signposted by the overdensity of LAE and LBGs, where the structures are spread across $>$ 10000 cMpc (e.g. Casey, 2016) and their overdensities are only slightly higher than the fields. Based on our arguments above, it is reasonable to conclude that what we are seeing here is the birthplace of the massive BCG. However, our estimations still suffer from large uncertainties, as these early-developed structures may not follow the general Baryon/dark matter fraction from the abundance matching (Behroozi & Silk, 2018).

Besides, the starburst stage of the proto-cluster core gives a natural explanation of the dramatic difference between proto-cluster and cluster. In galaxy clusters, the central massive galaxies are generally found to be dead and red, sustaining a low level of star formation. However, in the proto-cluster, they are mostly dominated by the young and gas-rich star-forming galaxies. Between these two stages, the proto-cluster core can be a quick transitional stage. Combining their high SFR and the availability of a large cold gas reservoir, the proto-cluster core can form a substantial amount of stellar mass and deplete most of their cold gas efficiently. In the meantime, the growth of the halo mass and the wake of the AGN will also stop the formation of cold gas, reducing the star formation over the subsequent cosmic time.

3.4 Conclusion

The over-density of DSFGs likely traces the core region of massive halos, which is a unique signpost of the most active proto-cluster in the early Universe. Due to its massive cold gas reservoirs and boosted SFR, it could betray the emergence of the brightest cluster galaxies in the most massive clusters. Consequently, the survey of the proto-cluster core is highly complementary to our existing surveys towards the mega-structures based on the overdensity of line emitters and drop-outs or the massive halos traced by QSO or radio galaxies.

However, as we have found in the field of J0217−0820 and in the dedicated ALMARED survey, the line-of-sight projection effects can significantly contaminate the survey of proto-cluster cores in the submm/mm bands. The physical reason for this is the combination of the negative K-correction and the cosmic variance. The weakness of the current selection criteria is the lack of sub-arcsec surveys in the near- and mid-IR, which prevents the accurate determination of the photometric redshift of the member galaxies. The advent of JWST and its efficiency in broad-band photometry can potentially initiate a new sample of proto-clusters.

Although we still have a limited number of proto-cluster cores, their existence suggests that a considerable fraction of stellar mass in the cluster core could have been formed on a short timescale ($\sim$ 100 Myr). Comparing their mass with that of the BCGs in massive clusters indicates that we are witnessing their very early formation. Following the general inside-out growth and quenching mechanism of the clusters, a multi-wavelength follow-up of these systems could provide valuable insights into the development of AGNs and environmental effects in the transitional stage of cluster formation.

3.5 Appendix

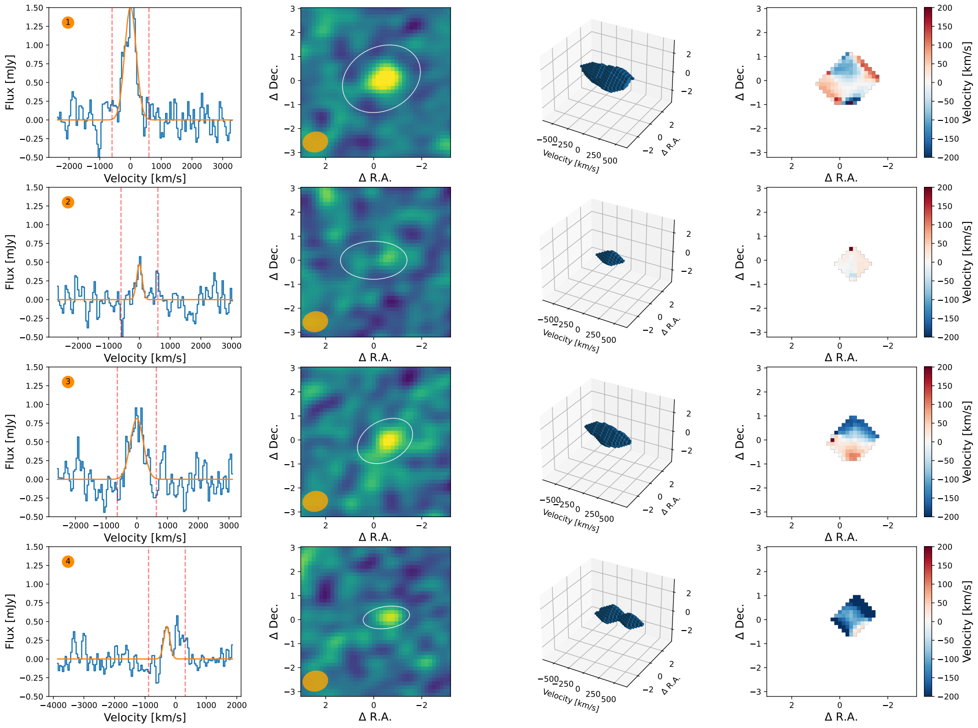

Fig. 3.13: The spectral line analysis of the member galaxies in AR1635. The first column shows the integrated spectrum extracted from the aperture, which is illustrated in the second column. The second column shows the collapsed image of the CO(5-4) within the velocity range that is shown as the vertical dash line in the first column. The third column shows the 3D structure of the line emission in the 3-dimensional space (ra + dec + velocity). The fourth column shows the moment-1 image of the emission line. Only the regions shown in the third column were used to derive the moment-1.

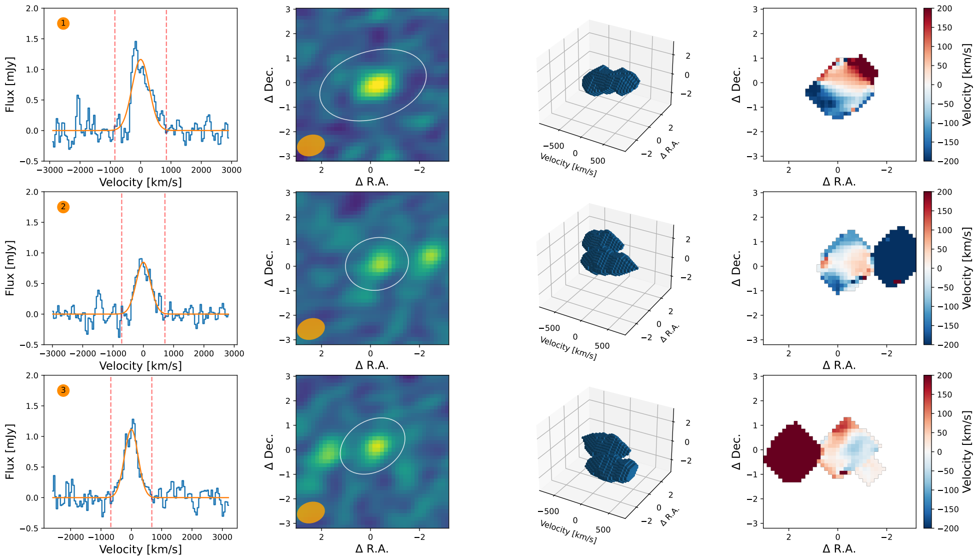

Fig. 3.14: It is the same as Fig3.13 but for the member galaxies in AR2440.

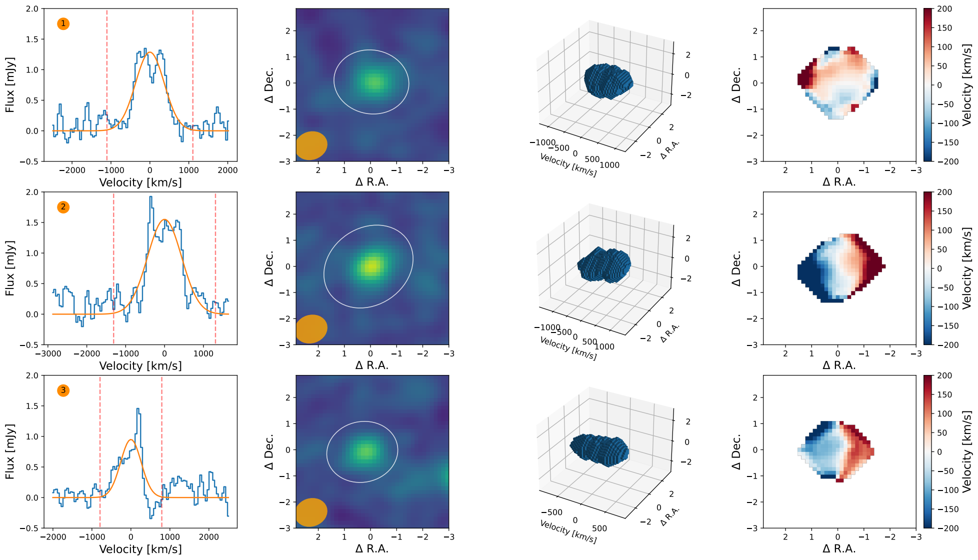

Fig. 3.15: It is the same as Fig3.13 but for the member galaxies in AR2623.

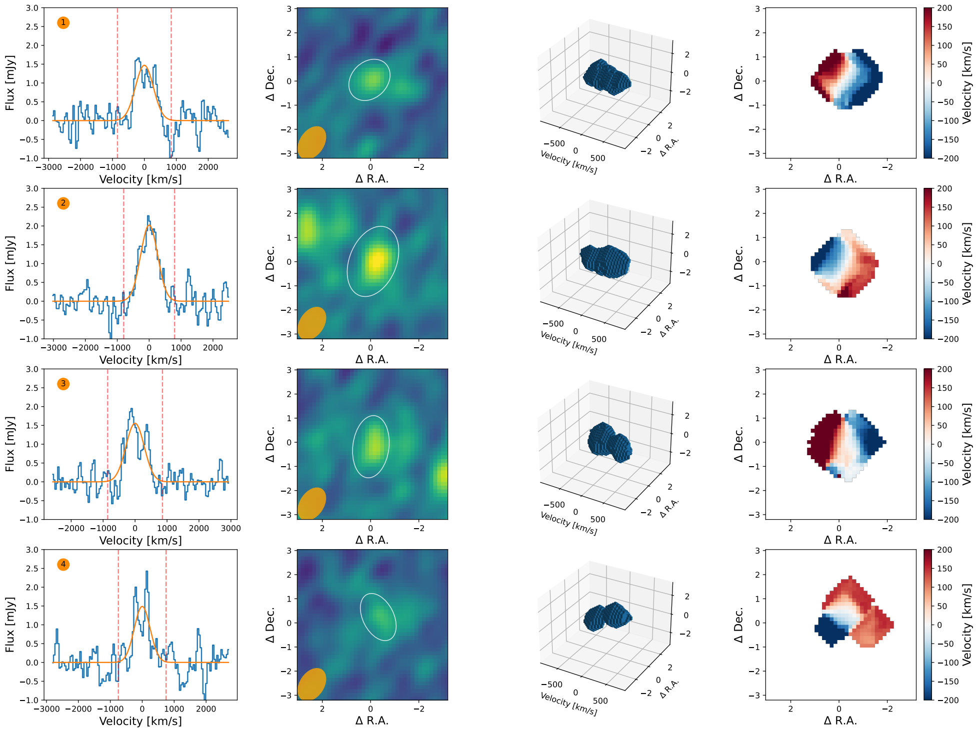

Fig. 3.16: It is the same as Fig3.13 but for the member galaxies in AR2747.

Chapter 4

A new window to explore the magnetised early Universe

4.1 Motivation

Magnetic fields have been found in almost all astronomical scales and play an essential role in many astrophysical phenomenons (Beck, 2015; Han, 2017). On the surface of the Sun, the magnetic fields dominate the solar activities, transporting energy from the converting zone to the heliosphere (Solanki et al., 2006); around the proto-starts, they dissipate the angular momentum and initiate the formation of planetesimals within the disk (Andrews, 2020). within molecular clouds, they modify the thermal instability, affecting their star formation efficiency (Crutcher, 2012); near the supermassive black holes, they control the accretion flows and drive the bipolar relativistic jets (Tchekhovskoy et al., 2011); in galaxies, they can thread the interstellar medium (ISM), mediating the baryonic cycles; on even larger scales, magnetic fields interfere with the accretion flows from the circumgalactic medium (CGM) and the intergalactic medium (IGM), complicating the environmental effects on galaxy evolution (Berlok & Pfrommer, 2019; van de Voort et al., 2021). However, observations of magnetic fields lag far behind theoretical studies, especially on galactic scales, demanding new observational strategies to constrain their formation at various spatial scales and on different timescales.

Our current understanding of galactic magnetic fields comes mainly from observations of the nearby Universe. Large-scale ordered magnetic fields have been routinely found in the ISM and CGM of the local galaxies. The strength of magnetic fields in the ISM is generally around 10μG, mostly aligned with galactic disks or galactic outflows (e.g. Beck, 2015; Lopez-Rodriguez, 2021). Ordered magnetic fields are also common in the CGM, and sometimes even in the IGM; their average strength is orders-of-magnitude weaker, but extend over much larger scales (e.g. Heesen et al., 2023a). The observation of magnetic fields in the distant Universe is much more difficult. Handful direct measurements indicate that they have already developed μG ordered magnetic fields when the Universe is about half of its current age (Bernet et al., 2008; Mao et al., 2017). Despite the difficulties in observation, we know next to nothing about their structure in and around galaxies in the early Universe, which is fundamental to trace their origins and amplification mechanisms.

Various theoretical models have been proposed to explain the formation of galactic magnetic fields. They are generally thought to start from seed fields and then be amplified through various galactic activities. The seed fields could be primordial, generated from the battery process (Biermann, 1950), where the magnetic fields are produced by the drifting electrons in the cosmic plasma. In addition, during the early formation of galaxies, the supernova explosion and AGN jet could also seed stronge magnetic fields quickly into the ISM (Beck et al., 2013). Then, the magnetic field can be amplified through the small-scale dynamo (or fluctuation dynamo), where the turbulent kinetic energy of the gas can be effectively converted to the energy of magnetic fields to reach energy equipartition. Meanwhile, the shear inside the galaxies, such as the formation of galactic disks, galaxy mergers, and galactic outflow can stretch the turbulent magnetic fields to form kpc-scale magnetic fields (Chyży & Beck, 2004; Drzazga et al., 2011; Jones et al., 2019b; Lopez-Rodriguez et al., 2020). To test various dynamo mechanisms in amplifying magnetic fields, we need to measure the strength and resolve the structure of magnetic fields across a wide range of cosmic time.

However, detecting and resolving magnetic fields is still difficult with existing methods. Only handful detections have been reported, so far, in detecting the magnetic fields of distant galaxies. The first, *Zeeman splitting*, is caused by spectral line splitting in the presence of ambient magnetic fields. The main difficulty with this method is finding sufficiently strong spectral lines with large line splitting, which limits us mainly to the H I 21-cm line and maser lines that can be observed in the nearby Universe. Initially, Wolfe et al. (2008) found a 21-cm absorption system in the direction of the luminous quasar, 3C 286, at $z = 0.692$. The H I Zeeman splitting shows this galaxy has developed a global magnetic field with an average field strength of about 84 μG. However, this detection was later proved to be a false detection (Wolfe et al., 2011) The other method is *Faraday rotation*, where the polarisation angle changes when light passes through a magnetised medium. High-redshift polarised quasars are the best targets, but their total rotation measures offer no information about the origin of the rotation (Oren & Wolfe, 1995; Kronberg et al., 2008). One way around this difficulty is to measure the Faraday rotation of Mg II absorbers. Bernet et al. (2008) showed statistically that quasars with strong Mg II absorption lines are associated with large rotation measures. Having assumed that Mg II absorption is associated with the halo of normal galaxies along the sightlines of quasars, their results favour a highly organised magnetic field when the Universe was one-third of its current age. More recently, Mao et al. (2017) targeted the differential Faraday rotation between multiple images of a lensed system, finding μG-level coherent magnetic fields in a galaxy at $z = 0.439$. Although inspiring, such detections are possible only because of rare fortuitous alignments; they do not offer us a route to reliable results from a well-understood sample. In addition, the Zeeman splitting and Faraday rotation can be hardly extended to even higher redshifts, even with the next-generation radio telescopes.

In this chapter, I propose to use the dust polarisation of the distant star-forming galaxies to trace the galactic magnetic fields and report two successful detections of large-scale ordered magnetic fields in two starburst galaxies at $z =\sim 2.5$ and $z \sim 5.6$, respectively.

4.2 Dust polarisation as a new B-field prober

Polarised dust emission is sensitive to the surrounding magnetic fields. In the presence of magnetic fields, the minor axis of dust aligns with external magnetic fields, leading to a preferential axis of dust extinction. This phenomenon is best explained by radiative torque alignment (Andersson et al., 2015), where radiative torques rotate the dust grains, inducing a net internal magnetisation and fast precession towards the alignment. This makes dust polarisation a robust tracer of surrounding magnetic fields. Since this discovery, dust polarisation has been used widely to study magnetic fields in molecular clouds and nearby galaxies (e.g. Planck Collaboration et al., 2015; Lopez-Rodriguez et al., 2022).Recently, Borlaff et al. (2023) systematically compared the magnetic fields traced by radio polarisation and dust polarisation in the Survey on extragALactic magnetiSm with SOFIA (SALSA) legacy programme, finding that far-IR dust polarisation is a better tracer of magnetic fields in the dense gas flows.

Dust is ubiquitous in the early Universe. Cold dust emission dominates the Far-IR SED of dusty galaxies, which redshifts into submm windows for galaxies at $z > 1$. From submm windows, dust emission has been detected in galaxies dating back to the epoch of re-ionisation (e.g. Laporte et al., 2017a), thanks to the advent of sensitive submm/mm telescopes like ALMA. Until now, thousands of DSFGs have been found (Hodge & da Cunha, 2020). Among them, hundreds of DSFGs are strongly lensed, where a foreground galaxy has amplified their flux densities and spatial scales. These ‘strongly lensed DSFGs’ have become the most popular targets for the resolved study of galactic substructures, ISM, and kinematics (Rizzo et al., 2020; Dye et al., 2022). Meanwhile, benefiting from the negative-K correction, the SMGs with similar total SFR share nearly constant flux density across from $z \sim 1 - 5$ (Guiderdoni et al., 1997; Blain, 2002).

Combing these two advantages, probing dust polarisation in DSFGs is a promising method to unveil the structure of magnetic fields in the early Universe. One possible concern of this method is the requirement for spatial resolution. The magnetic fields in starburst galaxies are thought to be turbulent, which may wash out the weak ordered fields if we averaging in a large physical region. Strong lensing provides a solution, but it is still unclear how well could the lensing preserve the weak polarisation signals. We thus tested the feasibility of this method using the local starburst galaxies. We used two test galaxies, M 82 and NGC 1068. They are both starburst galaxies and have different inclination angles, which makes them have different geometries of the global magnetic fields. In M 82, the magnetic fields are perpendicular to the stellar disk, aligned with the bipolar outflows (e.g. Jones et al., 2019b; Lopez-Rodriguez, 2021). In NGC 1068, the magnetic fields are generally aligned with the stellar disk (Lopez-Rodriguez et al., 2020). We first placed our test galaxies at the redshift of 2.553, which is the same as 9io9 (see more detailed discussion in §4.3). Then, We lensed them using the lens model of 9io9. We show our results in Fig.4.1. We found that strong lensing can preserve the global magnetic fields.

We also test the results with different spatial resolutions. Fig.4.2 shows the observed polarisation strength at different physical resolutions. In general, the better resolution, the more details of the magnetic fields we can recover. Our simulation indicates that a physical resolution of 1-1.5 kpc, which is about 0.2 arcsec at $z \sim 3$, is the optimal resolution to observe the magnetic field structure traced by dust polarisation. If the resolution is too low, the averaging effect can wash away the polarised signals. However, a much higher resolution is achievable with the current

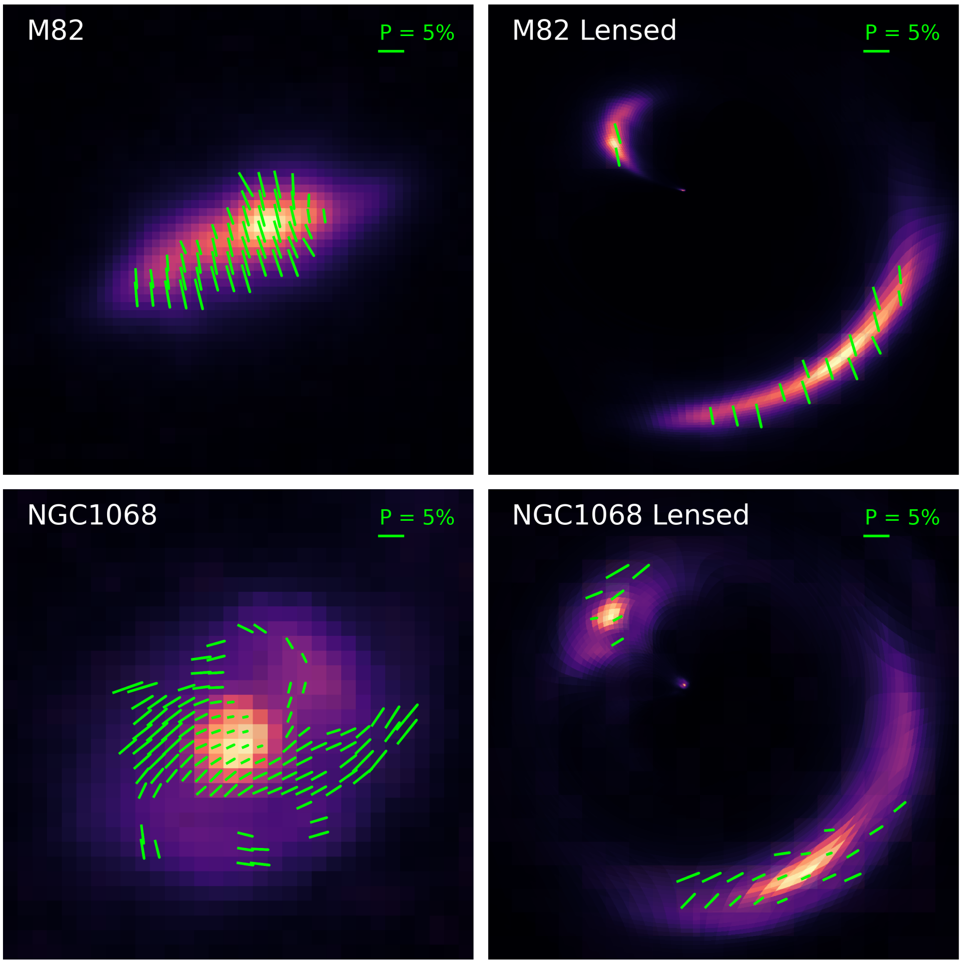

Fig. 4.1: The strongly lensed dust polarisation maps of local starburst galaxies. Left: Inferred magnetic field orientations in M 82 and NGC 1068, as revealed by HAWC+ onboard SOFIA (Jones et al., 2019b; Lopez-Rodriguez et al., 2020). The polarisation vectors have been rotated by 90° to show the magnetic field direction. Right: mock *lensed* images of these galaxies, if they were amplified gravitationally by a foreground galaxy at $z = 0.2$, with a similar configuration as that inferred for the 9io9 system. The simulations demonstrate that we will be able to map polarisation spatially across numerous independent resolution elements. The lengths of the green vectors show the polarisation fraction, P (only fractions larger than 5% are shown in M82, and fractions larger than 1% in NGC1068). This simulation indicates that the strong gravitational lensing preservers the large scale polarisation signals.

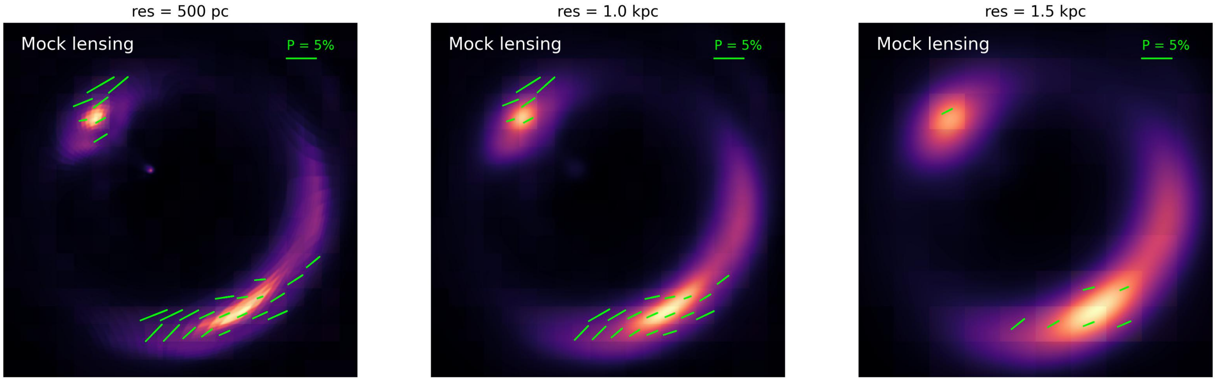

Fig. 4.2: The same lensed image as Fig.4.1 but observed with three different spatial resolutions. The recoverable polarisation signal decreases with a worse spatial resolution. This simulation demonstrates we need kpc to sub-kpc physical resolution to be able to probe the similar magnetic field structures in the ISM as we saw in NGC 1068.

interferometer, but the required sensitivity would be too expensive for a single source.

In the following sections, we applied this method to two exceptional bright and strongly lensed SMGs: 9io9 and SPT0346-52, at z=2.553 and z=5.6559, respectively, to test the validity of this method at different cosmic time.

4.3 Observation and data reduction

We conducted full-polarisation observations of 9io9 and SPT0346−52 in ALMA band 7 (project ID: 2021.1.00458.S, P.I. Jianhang Chen). The first target, 9ioi, is a hyper-luminous galaxy at z=2.553, found through the citizen science project *SpaceWarps* (Geach et al., 2015). It is one of the brightest SMGs in the sky, strongly lensed by a luminous red galaxy at $z \sim 0.2$. It has been thought to host an AGN in the centre (Geach et al., 2018). Because of its unusual brightness, it has been a preferential target to study the interstellar medium and AGN feedback in the early Universe (Geach et al., 2018; Doherty et al., 2020).

The other target is SPT0346−52. It is a strongly lensed galaxy at $z = 5.56$; it is the brightest and most intensive star-forming galaxy discovered in the 2,500 deg^2 of South Pole Telescope (Vieira et al., 2010; Spilker et al., 2016). Correcting for the lens amplification, SPT0346−52 is an instrinsic hyper-luminous galaxy ($L_{IR} \simeq 3.6 \times 10^{13} L_\odot$) and forming stars with highest rates (SFR$\simeq$ 4500 ± 1000$M_\odot yr^{-1}$). Because of this, it is one of the most popular targets to study the resolved properties of the multiphase ISM (Litke et al., 2019; Apostolovski et al., 2019; Litke et al., 2022).

The data were obtained in August and September 2022. In each observation, the coverage of the parallactic angle is wider than 120 deg, which delivers independent full-polarisation calibration. The data calibration were conducted with CASA (ver.6.2.1, The CASA Team et al., 2022). I followed the standard calibration pipeline of the two parallel polarisations, XX and YY. For the observation towards 9io9, calibrators J0238+1636 and J0217+0144 were used to calibrate

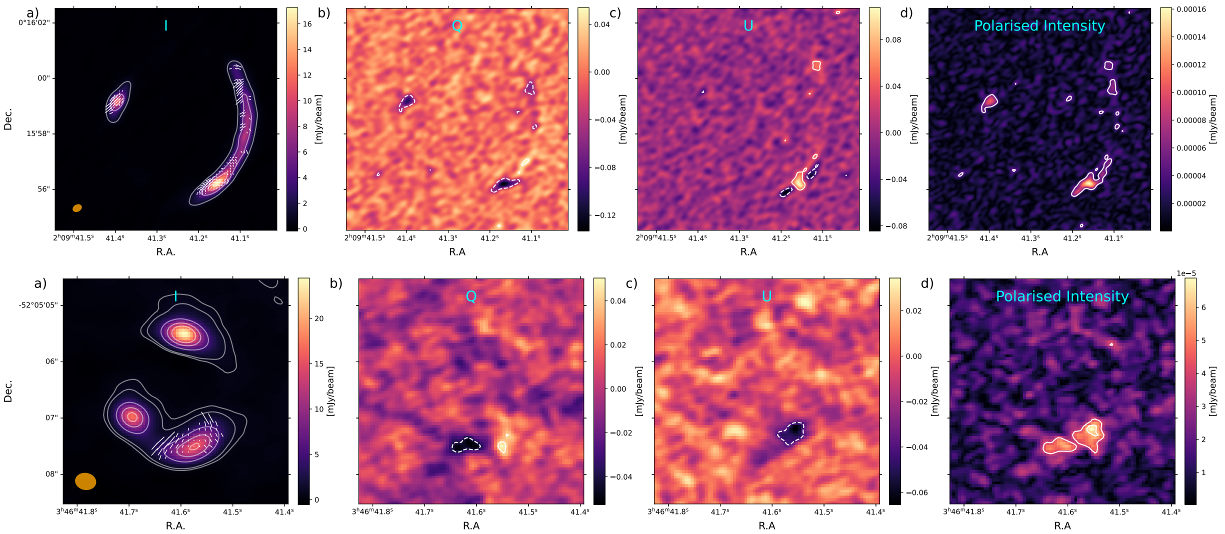

Fig. 4.3: The observed dust polarisation and the inferred magnetic fields. The first row shows the results of 9io9; the second row shows the results of SPT0346−52. **a-c)** are the maps of Stokes parameter: I, Q, and U. **d)** is the linear polarised intensities. In the maps, all the contours start at 2σ level of the RMS of the whole image and increase with a step of 2σ. The dashes contour indicates the negative value, while solid contours show the positive levels. The two detections of polarised dust emission indicates the starburst galaxies already developed kpc-scale right after the cosmic re-ionization.

the bandpass and phase variation, respectively. For SPT0346−52, calibrators J0519−4546 and J0253−5441 were used. The total flux of each target was scaled from their bandpass calibrator. After that, the calibrated data were split for the subsequent polarisation calibration, where the internal calibration template was used. J0423-0120 was used as the polarisation calibrator for both science targets. During the observation sessions, the best-fit polarisation fraction of J0423-0120 is around 4.0%, consistent with the long-term monitoring of its polarisation status. The final uncertainty of the linear polarised fraction is 0.1 per cent and about 1 deg in polarisation angle according to the ALMA technical handbook.

After the polarisation calibration, we used task TCLEAN within CASA to make the images and cubes. We first create cubes of the total intensity to search for possible detection of weak lines. The cubes were created with the 'hogbom' deconvolver and natural weighting. In SPT0346−52, we did not identify any strong lines ($> 5\sigma$) within our spectral window, which is consistent with our expectation as we designed our spectral window to avoid all the possible strong lines and to maximize the continuum sensitivity. In 9io9, we did detect a weak spectral line at 337.6 GHz. The line was masked when we created the continuum maps of the Stokes parameters. Finally, we created the full Stokes images with the 'clarkstokes' deconvolver and natural weighting.

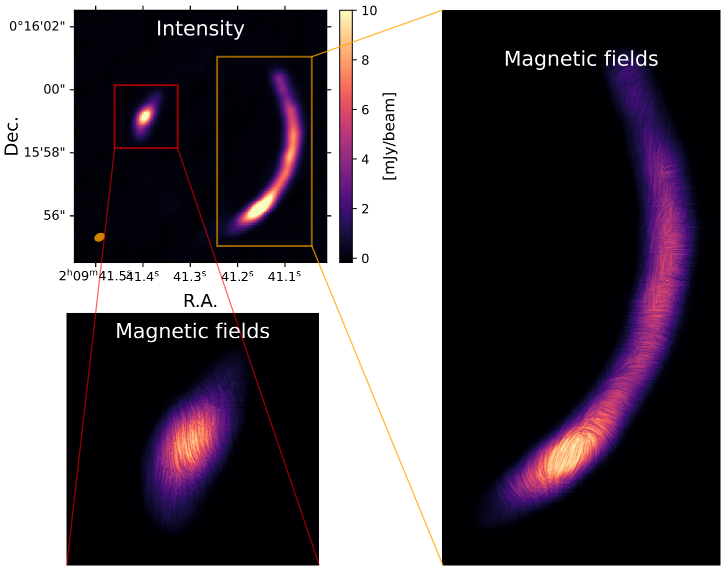

Fig. 4.4: The magnified view of the magnetic structure in 9io9. The field lines were constructed through the line convolution between the total intensity and the polarisation vectors. Our high-resolution observation supports that the magnetic field could change its direction quickly between different regions.

4.4 Results and Analysis

We have detected the polarised dust emission in both targets. Fig.4.3 shows the total intensity and the Stoke Q and V maps of the two targets. The linear polarised intensity is derived from $I_{\text{poli}} = \sqrt{Q^2 + U^2}$. The linear polarised fraction varies across the two galaxies, between 1-2%, which is similar to the median values in the nearby starburst galaxies (Lopez-Rodriguez, 2023). The polarised angle is defined as $\chi = \frac{1}{2}\arctan(U/Q)$. The inferred orientation of the magnetic field is perpendicular to the polarisation angle, which is shown as the white vectors in Fig.4.3. In 9io9, we also resolved the structure of the magnetic fields. The orientation of the field lines changes quickly in different regions. The zoom-in plot of the field structure is shown in Fig.4.4.

In 9io9, we modelled the magnetic fields in the source plane. We used the pixelated lensing inversion to recover the structure of magnetic fields in the source plane (Geach et al., 2018). The pixelated method needs no assumption about the source structure, which allows us to probe more complex source structures. We used the dust continuum observation to derive the lens modelling. We then constructed the mapper between the source plane and image plane with the best-fit model. The mapper was later used to fit the Stokes parameter and recover them in the source plane. We show the best-fit model and reconstructed source and Stokes parameters in Fig.4.5. The source plane model supports that there is another companion galaxy in the north of the main source. The reconstruction of the Stokes parameters in the source plane is highly sensitive to lens modelling, which is the dominant uncertainty in the recovered magnetic field structures. We have seen some artificial structures in the reconstructed source plane in Q and U maps in Fig.4.5.

In SPT0346−52, we only detected weak polarisation in two of the sub-images. Due to the low SNR of the polarisation, we cannot derive the polarisation maps in the source plane. Based on the analysis of the pixelized mapping, we found the detected polarisation signals mainly come from the edge of the continuum source (s).

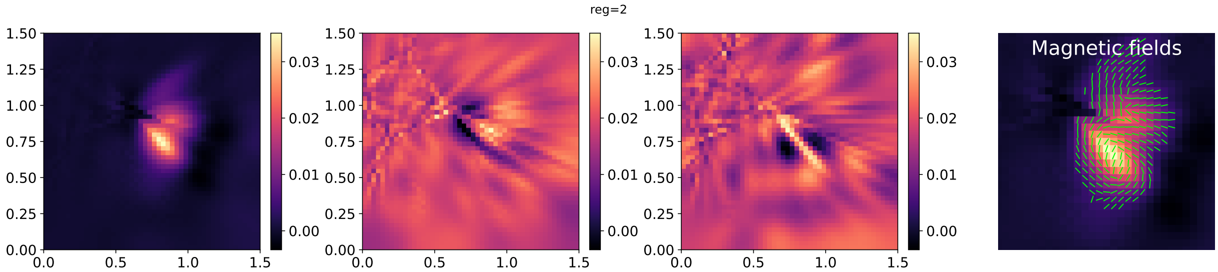

Fig. 4.5: The source plane model of the different Stokes parameters and the reconstructed structures of the magnetic fields. The total intensity model in the source plane shows two distinguished components, which indicates 9io9 system may be a merger system. The reconstructed magnetic fields also support that the magnetic fields in these two components are connected. Due to the weak polarised signals, the reconstructed structure of the polarisation maps still suffers from the lensing modelling.

4.5 Discussion

4.5.1 Star formation and magnetic fields

In the local disk galaxies, the total magnetic field strength correlates with the SFR surface density, $B_{\mathrm{tot}} \approx \sum_{\mathrm{SFR}}^{0.14}$(Beck, 2015). Together with the strong FIR-radio correlation (Condon, 1992), it supports the important role of the galactic magnetic fields in connecting synchrotron radiation to star formation activity. The FIR-radio correlation has been observed down to $z \sim 4$ by stacking analysis, and it shows mild variations with the redshift (Ivison et al., 2010c; Magnelli et al., 2015; Delvecchio et al., 2021), suggesting that similar galactic magnetic fields may already be present in the star-forming galaxies up to $z \sim 4$. Our two targets are all extreme starburst systems. If the turbulent dynamo is working effectively in the early Universe, they will be the most likely systems to develop strong magnetic fields.

However, only the turbulent magnetic fields correlate with the SFR surface density, the ordered magnetic fields show no obvious relationship with the SFR density (Beck, 2015). According to the mean-field theory (Beck et al., 2019), the development of large-scale ordered magnetic fields requires the presence of strong small-scale fields. Therefore, our two detections support the claim that strong magnetic fields may be widespread in star-forming galaxies across a wide range of cosmic time. Unfortunately, due to the limited sample size, we do need future observations to test our claim. If this claim stands, it also suggests that the FIR-radio correlation may exist at an even higher redshift than currently reported. The future upgraded radio facilities, such as the SKA and the ngVLA, will be able to test our claims.

It is also worth noting the contamination and contribution of the AGN in our systems. The evolution of AGN jets could also inject significant magnetic fields into the ISM, which is another way of accelerating the formation of μG magnetic fields. In 9io9, its radio excess betrayed the presence of AGN in the galactic centre (Geach et al., 2015). In SPT0346−52, however, the presence of AGN is more elusive. The joint SED analysis supports negligible AGN contribution to the total SED, supported by the non-detection of X-ray and low radio emission (Ma et al., 2016). Our current result cannot place reasonable constraints on the role of AGNs in the development of galactic magnetic fields. Future surveys covering a wider range of galactic properties will help in this regard.

4.5.2 Galactic disk

Theoretically, star-forming disk galaxies can develop and maintain large-scale ordered magnetic fields. Such fields originate from seeds, which can be amplified by the small-scale dynamo (Beck et al., 2019). Those initial turbulent fields can then be gradually ordered by the α-Ω dynamo (Beck, 2015): the presence of a galactic disk can provide stable differential rotation – the Ω-effect – to extend the twisted magnetic fields (the α effect) to the galactic scale. This paradigm has also been tested extensively in magneto-hydrodynamic simulations (Pakmor et al., 2017; Marinacci et al., 2018). However, observational constraints are limited to the nearby Universe and do not span the necessary cosmic timescale to discriminate between different implementations.

Geach et al. (2018) has re-constructed velocity fields of 9io9, which generally follows the pattern of a rotating disk. If the detected field structure of the main galaxy in the 9io9 system is connected to the formation of the disk, our observation supports a relatively efficient amplification process. Various numerical simulations have also pointed out that the synergy between turbulence and α-Ω dynamos can efficiently amplify the field strength, but the reported amplifications and timescales vary between different schemes and resolutions (e.g. Pakmor et al., 2017; Khoperskov & Khrapov, 2018; Whittingham et al., 2021). Our observation shows that such process could already develop μG magnetic fields within 2.5 Gyr, which provides a valuable reference for the numerical simulations.

4.5.3 Galactic outflow

Galactic outflows are common in starburst and AGN host galaxies, from the local Universe out to the epoch of re-ionisation (Veilleux et al., 2020). They play a crucial role in the baryon cycle of galaxies, mediating star formation and the chemical evolution of galaxies. Galactic outflows can also accelerate the formation and amplification of magnetic fields.

Outflows are an intriguing explanation for the observed magnetic fields in our systems, for several reasons. First, outflows, driven by either starbursts or AGNs, can seed strong magnetic fields in the ISM and CGM through the supernova explosion and relativistic jets (e.g. Beck et al., 2013). Furthermore, these systems are highly turbulent, so the turbulence dynamo can efficiently amplify these seed magnetic fields by equalising the energy between the turbulence and the magnetic fields (e.g. Schober et al., 2013). Subsequently, the shear produced by supernova explosions and galactic outflows can also stretch the field lines in space, converting small-scale magnetic fields into galaxy-scale fields (Brandenburg & Subramanian, 2005).

This scenario is supported by both magneto-hydrodynamic simulations and observations. Simulations by Bertone et al. (2006) showed that magnetic fields along gas outflows contribute to the magnetisation of the IGM. Meanwhile, Hanasz et al. (2013) demonstrated that starburst-driven relativistic cosmic rays can produce large-scale magnetic fields at intermediate and high redshifts. Observationally, strong magnetic fields have been found in the bipolar outflow of M82, producing kpc-scale fields that extend into the circumgalactic medium (CGM, Lopez-Rodriguez, 2021). These fields are in close equipartition with the turbulence, indicating their dynamical importance in the outflow. Furthermore, a statistical study of nearly 200 nearby galaxies has found that magnetic field strength decreases slowly as a function of distance from the galactic disk, but is enhanced in the bipolar direction, indicating that the CGM is magnetised by the galactic outflow (Heesen et al., 2023b).

Given the extreme SFR of our systems, it is natural to think that they could have developed galactic outflows driven by stellar feedback. Our current observations cannot directly support the ongoing function of the outflow in shaping the observed magnetic fields. Although, in SPT0346−52, Jones et al. (2019b) argued that the broad wing of the [C II] could support the presence of galactic outflows, it could also be a close merger (see more discussion in the next section). Future deep observations towards the unlensed galaxies may help to overcome this problem. Unfortunately, the lensing sources are still the most viable target to increase the sample size of similar measurements. We will return to this discussion in the next Chapter.

4.5.4 Galaxy mergers

Galaxy mergers are another efficient way to produce large-scale ordered magnetic fields. In the literature, mergers have been associated with starbursts (e.g. Teyssier et al., 2010), violent gas in/outflows (e.g. Herrera-Camus et al., 2019), AGN (e.g. Satyapal et al., 2014), and morphological transitions (e.g. Kocevski et al., 2012). During the merging, the gravitational torque can generate large-scale shear that stretches the small-scale field lines to large-scale ordered fields. Galaxy encounters are thought (and expected) to be more frequent in the early Universe (De Lucia & Blaizot, 2007), so this could be another efficient channel to produce the large-scale magnetic fields observed in 9io9 and SPT0346−52. In the local prototypical major merger – NGC4038/39 (the Antennae) – kpc-scale ordered magnetic fields have been found connecting the two merging galaxies (Chyży & Beck, 2004). Their ordered magnetic field strength goes up to 10 μG, which is several times larger than the Milky Way and other disk galaxies. Because of the long dissipation timescales of magnetic fields, they are also thought to be excellent tracers of tidal interaction and ram pressure (Drzazga et al., 2011).

In SPT0346−52, detailed analysis suggests that the starburst was triggered by a major merger (Spilker et al., 2015; Litke et al., 2019). The two merging galaxies are close to each other, separated by only $\approx$ 1 kpc from the lens modelling (Spilker et al., 2015). However, as we mentioned before, this system has also been discussed as a single galaxy but with a rotating disk (Jones et al., 2019a). Due to the limitations in the angular and velocity resolution, our full-polarisation data still cannot distinguish these two possibilities. Fortunately, in 9io9, our deeper observations do support the possible role of the merger in shaping the magnetic fields. Our source plane reconstruction supports that there may be a companion galaxy close to the main galaxy in the 9io9 system (see Fig.4.5). Moreover, the recovered magnetic fields connect the two merging galaxies, supporting the ordered magnetic fields could result from the merger process.

4.5.5 Implication about galaxy secular evolution

Although magnetic fields are widely accepted as an essential part of galaxy evolution, their importance is still under debate, especially on the galactic scales. To affect the gas collapse or infall on a free-fall timescale, the field strength needs to be $\approx 100\,\mu$G (e.g. Ceverino et al., 2012), which is 10–100× larger than observed. Krumholz & Federrath (2019) summarised the effects of magnetic fields in various star formation processes, concluding that they are not the dominant player in controlling star formation inside galaxies, but they may play a critical indirect role. Meanwhile, the timescale for reaching the equilibrium between the magnetic fields and the kinetic energy of the ISM could be $\sim$10 Gyr, making it unlikely that the magnetic fields could imprint any noticeable effects on short timescales.

The detection of the dust polarisation in SPT0346−52 indicates the presence of kpc-scale magnetic fields could have already been developed during the first Gyr of the galaxy's evolutionary history. The timescale for developing such large-scale magnetic fields is several times faster than the model predictions (e.g. Rodrigues et al., 2019). At similar redshifts, limits on much larger spatial scales are available from statistical analysis. Based on the γ-ray observations, where the strength of the intergalactic magnetic field can be inferred from the signature of the γ-ray cascade, the current lower limit showing the strength of magnetic fields in IGM is ($7 \times 10^{-16} > 10^{-9}$ G, S. H. E. S. et al., 2023). On the other hand, based on the Faraday rotation of the high-z radio source, Pshirkov et al. (2016) reported an upper limit of nG over 1 Mpc. Our result implies that magnetic fields can be developed much more efficiently within the galaxies. With the energetic AGN or starburst feedback, such ordered magnetic fields may be also transported to the CGM and IGM, which can be tested by future radio telescopes.

Our observations are only sensitive to kpc-scale ordered magnetic fields. Strong small-scale magnetic fields are required to develop the observed galactic-scale magnetic fields. Therefore, it is reasonable to conclude that the strength of the small-scale magnetic fields should be even stronger and more ubiquitous in the early DSFGs. Taking into account the lifetime ahead of these galaxies, it indicates that magnetic fields can play a noticeable role in the galaxy's secular evolution. Moreover, the approach I used here can be easily applied to many other strongly lensed dusty star-forming galaxies and even to the brightest unlensed targets. I will discuss the future projects in Chapter 5.

4.6 Conclusion

In this Chapter, I have discussed the use of full-polarisation of dust to probe the magnetic fields in the early DSFGs. I first tested our methods based on the simulations of two local starburst galaxies: M 82 and NGC 1068. I placed these two galaxies at the same redshift at 9io9 and strongly lensed by the same foreground galaxy. My simulation suggests that the strong-lensing system could preserve the signature of the large-scale magnetic fields. I have also shown that the full-polarisation and sub-arcsecond resolution observations towards the strongly lensed system are currently the optimal choices for observing the magnetic fields in the distant DSFGs.

I have also reported the results from our pioneer programme towards two lensed SMGs: 9io9 and SPT0346−52. Polarised dust emissions have been detected in both galaxies, suggesting that strong magnetic fields have already been developed in these two systems. I have discussed the possible physical mechanisms to explain the observed kpc-scale ordered magnetic fields. In 9io9, our source plane reconstructions support most of the galaxy merger as the cause, but we cannot exclude the contribution from other processes, such as galactic outflow and disk formation.

This programme has fully demonstrated the promising future of using dust polarisation to study cosmic magnetism. This method could be applied to many more systems to cover a range of galaxy properties and cosmic time.

Chapter 5

Outlook

Over the past two decades, the submm/mm window has provided exquisite advantages to the study of dusty galaxies in the early Universe. This journey is like a baby first opened its bleary eyes by SCUBA on JCMT to now observe with crystal clear eyes through ALMA. The two-decade development of submm/mm telescopes has greatly advanced our understanding of the DSFGs. Even for ALMA, it has already served us for more than 10 years, but we continuously gain new insights on the properties of more general galaxy populations through the submm and mm windows. Like what is already going on for the ALMA 2030 roadmap, we are now in a golden time to explore new possibilities with the submm and mm windows. Meanwhile, along with the new arrival of *JWST*, Euclid in space and the VLT/ERIS and SKA on the ground, we are already in a new era which enables us to observe the early Universe deeper and wider, as well as to get glimpse of the new hidden mysteries of the early Universe.

5.1 Exploring the early magnetised Universe

In this thesis, I have reported the detections of dust polarisation in two SMGs at z=2.6 and z=5.6, respectively. These inspiring discoveries demonstrate the promising future of detecting cosmic magnetism with dust polarisation. In my opinion, there are three directions we can push forward in this regard.

Large surveys of kpc-scale magnetic fields in DSFGs are the first promising direction to explore. The combination of single-dish telescope and the interferometric follow-ups have found hundreds of bright DSFGs, with $S_{870} > 10\,\mathrm{mJy}$. One per cent of their dust polarisation can be detected with $\sim 3\,\mathrm{hrs}$ of ALMA telescope time, which is also the minimum time required for polarisation calibration. With the help of the strong lensing, the time spent on individual sources can be even shorter if they can be observed with the same polarisation calibrator. Moreover, these bright-end DSFGs have been found in a wide range of redshifts. It enables a systematic study of dust polarisation as a function of cosmic time, which can be used to constrain the amplification timescale of the galactic magnetic fields. It should be noticed that these bright DFSFs are highly biased targets, there are either intrinsically bright HyLIRGs or high-amplified systems, thus it will complicate the analysis of the global development of the magnetic fields in galaxies. On the other hand, a large variability of these galaxies offers the parameter space to test the relation of the galactic magnetic fields with many other galactic properties, such as the SFR, stellar mass, AGN fraction, multiphase gas mass and their averaged ionization properties. Most importantly, such large surveys of dust polarisation can be used to select the targets with strong galactic ordered magnetic fields, which can be followed up by much higher angular resolution and deeper sensitivity.

These high-resolution observations of dust polarisation in DSFGs are the second direction worth the effort. As I have mentioned in Chapter 4, various galactic mechanisms are responsible for the development of kpc-scale magnetic fields. The most three practical mechanisms are the formation of galactic disks, galaxy mergers, and galactic outflow. They can be tested with the high-resolution observation of the magnetic field structures. For instance, future observation could target galaxies with confirmed rotating disks. By investigating the polarised fraction and coherent length of the ordered field as a function of the local velocity gradient, these observations can be used to test the efficiency of the mean-field theory. Meanwhile, high-resolution observing campaigns with ALMA have already revealed detailed morphological and kinematic structures of the distant DSFGs. Aided by the *JWST*, it is possible to distinguish the galaxy merger and strong galactic outflow on the shaping of galactic magnetic fields. Moreover, the dust emission and the kinematics of DSFGs can be obtained simultaneously, which can double the efficiency of such observations.

The dust polarisation is not only sensitive to the geometry of the magnetic fields, it also depends on the dust composition of the ISM. The third direction is the observation of polarisation spectrum. The first key parameter is the size distribution of the dust. The wavelength dependent polarisation of starlights indicates that the large dust grains (>100 nm) are preferentially aligned (Kim & Martin 1995). In the meantime, different dust grains tend to have different polarisation properties. For instance, the silicate component is more likely aligned with the external magnetic fields (Guillet et al., 2018), and it is sensitive to the iron inclusions (Hoang & Lazarian, 2008, 2016). Recently, motivated by the recent results that the dust polarisation is frequency independent roughly from 250 μm to 3 mm (Planck Collaboration et al., 2020), Hensley & Draine (2023) proposed the unified dust model Astrodust, in which the large dust grains (larger than 20 nm) has been modelled by a single composite material. All these modelling efforts also make the full spectral dust polarisation observation an effective way to disentangle different dust models and constrain the dust compositions in the ISM.

Along with these potential science goals, it also demands innovation and upgrades of the instruments. For instance, the current highest ALMA band available for full polarisation observation is band 7. It can only reach the $> 100\mu$m region when we target the DSFGs at $z > 1$, which cannot distinguish the existing dust models very efficiently. The future full polarisation capability from ALMA band 8 and band 9 can potentially distinguish different models by targeting the galaxies in the very early Universe. Meanwhile, the higher frequency also benefit the observation of the magnetic fields. For a typical DSFG, its flux density rises following ν^2 in the submm regime, which can effectively compensate for the loose of the atmospheric transmission at a higher frequency. Moreover, high frequency also improve the spatial resolution. Along with the ALMA long-base-line capability, it is, therefore, possible to directly resolve the turbulence scales in the unlensed DSFGs, which makes it even possible to estimate the strength of magnetic fields without the nuisance of lens deprojection.

5.2 The infant AGN

Active galactic nuclei (AGN) are a crucial player in modulating galaxy formation in the massive end. Massive quenched systems have been frequently reported recently at relatively high redshift ($z > 2$, Forrest et al., 2020; Carnall et al., 2022; Nanayakkara et al., 2022; Valentino et al., 2023). Part of these galaxies has been identified in the overdense environments (Kubo et al., 2021; Ito et al., 2023). It has been a puzzle how could these galaxies develop their stellar mass and quench so quickly. SED analysis of these systems implies the presence of a short period of extensive star formation burst right before their quenching (Forrest et al., 2020; Carnall et al., 2022). The bursty SFH and the number density of these massive quenched systems are coincident with the massive SMGs (Toft et al., 2014). Under such an evolutionary connection, we should expect a wide presence of "infant AGNs" in SMGs that are ready to quench the host galaxy effectively. Such an infant stage of AGN could be heavily obscured by its accretion disk and the host galaxy. Current X-ray observation (sensitive to 0.1-12 kev) could only detect obscured AGN with the nuclear column density $N_{\rm H} < 10^{23.5}{\rm cm}^{-3}$ (Ricci et al., 2023), which indicates that we may have missed a significant fraction of these heavily obscured AGNs. Even harder X-ray is more insensitive to dust obscuration, which could be an approach to observe these heavily obscured AGNs. However, we still miss such a high-energy X-ray telescope with enough sensitivity and angular resolution for SMGs.

Submm/mm offers another opportunity. Recently, ALMA follow-up a sample of "high-fidelity" AGN, selected from the *SWIFT/BAT*, has discovered a tight correlation between the hard X-ray (10-150 kev) and mm emission near 100 GHz (Ricci et al., 2023). Hard X-ray is thought to originate from the corona above the accretion disk, directly connected with the AGN activity. This tight correlation implies the 100 GHz emission could also come from the AGN. Besides, their ratio $L_{100\,{\rm GHz}}/L_{14-150\,{\rm kev}}$ is independent of the BH mass, accretion rate, and the host galaxy, which further support these two emission could have the same physical origin. The current most competitive explanation for the 100 GHz continuum emission is the synchrotron emission from the corona of the AGN accretion disk. Around the accretion disk, the presence of strong magnetic can first heat the corona with the energy released from the magnetic reconnection. In the meantime, the relativistic electrons scattered from the corona travel through the strong magnetic fields producing substantial synchrotron emission. Due to the Compton self-absorption, this excess is only detectable around the restframe 100 GHz. The excess around 100 GHz has also been confirmed in the nearby type-I AGNs, which also supports their synchrotron origin (Inoue & Doi, 2018; Inoue et al., 2020). One striking advantage of the mm excess is that it is insensitive to the extreme column density in the central region, which makes it an intriguing alternative indicator of AGNs in the dust-rich SMGs.

ALMA has just offered its low-frequency receiver in band-1, along with the veteran VLA, such an experiment can already be applied to the high-z SMGs. A current difficulty for the existing interferometers is the limitation of spatial resolution, even with the longest baseline offered by ALMA and VLA, we still cannot resolve the accretion disk of the AGN. As a result, the mm-excess will be contaminated by a substantial free-free emission from the star-forming regions. The strongly lensed SMGs could help in some sense, where the free-free emission can be subtracted by carefully comparing the central region with the off-centre star-forming regions. Other more promising facilities are the Global mm VLBI Array (GMVA) and Event Horizon Telescope (EHT), which can better filter out the more extended emission from the central star-forming regions.

5.3 Chemical evolution in the early Universe

5.3.1 Far-IR spectral line diagnosis

Optical line ratios have been widely used to estimate the gas-phase metallicity in the galaxies up to cosmic noon (Maiolino & Mannucci, 2019; Kewley et al., 2019). However, in a dusty environment, the optical diagnosis suffers from dust extinction, which may bias the measurement of the gas phase metallicity (Chartab et al., 2022).

Far-IR spectral lines offer another window to constrain the gas phase metallicity (Nagao et al., 2011). The proposed lines such as [O III]52μm, [O III]88μm and [N III]57μm are most available in the Far-IR. Their individual ratio [O III]52μm/[N III]57μm and [O III]88μm/[N III]57μm are both sensitive to the metallicity, but they also depend on the gas density and ionization parameter (Rigopoulou et al., 2018). Luckily, they have opposite dependence on gas density, which makes the joint ration [2.2×[O III]88μm+[O III]52μm]/[N III]57μm an ideal tracer of metallicity (Pereira-Santaella et al., 2017). Besides, this line ratio is robust against the presence of AGN and insensitive to ionising conditions. Based on this ratio, Chartab et al. (2022) had demonstrated that the local ULIRG still generally follow the global mass-metallicity relation as normal star-forming galaxies.

However, even for ALMA, detecting far-IR spectral lines associated with N^{++} and O^{++} remains a challenge. Only for galaxies with $z > 5$, the [N III]57μm starts to reshift into ALMA band 10. Currently, due to the limitation of the availability, most of the works focused on the [O III]88μm and [N II]122μm (e.g. Killi et al., 2023). This ratio is also sensitive to metallicity, but it needs an independent way to constrain the ionization parameter. In the meantime, this ratio can only be used to in the galaxies without strong AGN contribution.

Luckily, recent *JWST* has substantially changed our view towards the galaxies in the early Universe. *JWST* has extended the metalicity measurement to the epoch of re-ionisation (Curti et al., 2023; Nakajima et al., 2023). It has been found that the mass-metallicity relation in the early Universe shows a significant scatter compared with model predictions (Trump et al., 2023; Nakajima et al., 2023). The large uncertainties have been partly attributed to the limited sample size and calibration errors, which demands an independent measurement with other methods. Meanwhile, some of the $z > 5$ systems are chemically evolved systems (e.g. Bunker et al., 2023), such that their far-IR could be bright enough to be detected by ALMA.

Far-IR line ratios offer a complementary to the optical metallicity diagnosis. The ”direct” T_e method based on the auroral line [O III]λ4363Å mainly traces the low-metallicity regions of the galaxies (Kewley et al., 2019). Therefore, applying these far-IR diagnoses is necessary to give an unbiased view of the metallicity evolution in the very early Universe. It should be noted that the far-IR line diagnosis directly measures the O/N ratio, which needs to be converted to O/H to be compared with the optical line diagnosis. The current best fit of the O/N to O/H model is derived from HII regions in the nearby galaxies (Pilyugin & Grebel, 2016), thus special attention should be warranted when comparing the results from different methods.

5.3.2 Isotope abundance

Along with the formation and death of stars, metals from the associated nucleogenesis are ejected into the ISM. The major isotopes, such as the ^{12}C, ^{14}N and ^{16}O, could be produced through multiple channels and represent the most common tracer of the chemical evolution. The yields of the minor isotope, such as ^{13}C, ^{15}N, ^{17}O, and ^{18}O can only be produced from some specific conditions and channels. For instance, ^{13}C is released mainly by low- and medium-mass stars ($M_* < 8\,\mathrm{M}_\odot$), while ^{18}O is mainly produced by massive stars ($M_* > 8\,\mathrm{M}_\odot$). Their relative ratios are sensitive to the initial mass of the stars, therefore recorded the valuable information of the initial mass function and star formation history (Henkel & Mauersberger, 1993; Romano et al., 2017).

Recently, with the advances in galactic chemical evolution modelling and the sensitivity of submm/mm telescopes, these ratios have been used to study the IMF quantitatively. Zhang et al. (2018a) analysed the ^{13}C/^{18}O ratio, as observed by ALMA, in four bright strongly lensed SMGs. They found that only a top-heavy IMF could explain why their line ratios are much ($\approx 10\times$) lower than the measurements in the Milky Way and nearby star-forming galaxies. It should be noted that uncertainties remain for the yields of rotating massive stars (e.g. Romano et al., 2019).

The full potential of ALMA in the search for isotope lines in the DSFGs is still largely unexplored. Taking advantage of the strong lensing, a large number of strong-lensed systems with different intrinsic properties could be the potential targets. Statistical analysis of their isotope ratios could reveal the intrinsic scatter of the isotope ratios and connect the isotope ratio with many other global galactic properties. More intriguingly, such a method can be also applied to the proto-cluster cores, where the stacking analysis of all the members can make these observations less expensive. Compared to their field galaxies, it could be used to explore the environmental imprint in their chemical evolution.

5.4 DSFGs as cosmic probes

5.4.1 Probing the CMB temperatures

According to the standard ΛCDM cosmology, the CMB temperature varies with redshift following $T_{\mathrm{CMB}} = T^{z=0}_{\mathrm{CMB}}(1 + z)$. One proposed way to test the ΛCDM cosmology is to measure the temperature deviation from the $1 + z$ power law (Lima et al. 2000). For DSFGs, the CMB has been an obstacle to the observations of DSFGs at high redshifts. Submm/mm window has facilitated the observation of the DSFGs in the early Universe. But at higher redshifts, the higher CMB temperature starts to dominate the background, reducing the contrast of the mm continuum emission from SMGs. da Cunha et al. (2013) and Zhang et al. (2016) have demonstrated its effects on the interpretation of the total mass of dust and molecular gas, galactic morphologies, and kinematics. Observationally, Jin et al. (2019) observed four distant starburst galaxies ($3.6 < z < 5.9$) with cold dust. They argued that the observed unusual steep Rayleigh-Jeans slope is due to the effect of CMB.

However, the strong effects of CMB on the DSFGs also provide the opportunity to test the ΛCDM cosmology. Riechers et al. (2022) demonstrated the first such successful experiment. Their observation was initially designed to perform a spectral scan of a massive DSFG at $z = 6.34$, but they have detected a strong absorption line at the frequency of H_2O(110-101). The best explanation is thought to be the synergy between the CMB excitation and Far-IR pumping. Firstly, the CMB excites the ortho-H_2O 1_{01} to 1_{10}, then the 1_{10} are effectively pumped up to 2_{21} by the strong dust emission. This model has predicted another strong absorption at restframe $180\,\mu$m and a strong forbidden emission at $108\,\mu$m, which can be tested with future follow-ups. Besides, a sample of SMGs have been confirmed at $z > 5$, which makes this new tracer valuable to be applied to a large sample to reduce its systematic errors.

A similar approach can be applied to other temperature-sensitive transitions. Such an approach has already been applied to the line-of-sight absorbers of quasars (e.g. Molaro et al. 2002). By assuming an excitation temperature equal to the CMB temperature, an upper limit of CMB temperature can be derived. Possible tracers in the submm/mm bands could be the rotational transitions such as HNC, HCN, HCO^+. Meanwhile, with the availibility of the sensitive ALMA band-1 reciever, such experiments can also be applied to a larger redshift.

5.4.2 Number counts of galaxy proto-cluster

Galaxy clusters trace the extreme peaks of mass density fields, making their number counts a powerful tool for constraining the cosmological and astrophysical models (Voit, 2005). A similar power also exists in the number counts of proto-clusters and proto-cluster cores. A large sample of the galaxy proto-clusters identified in the early Universe can also be used to test cosmology and the earliest feedback processes. However, we still lack a systematic way to search for such massive halos in the submm bands. I discussed the advantage of submm observations in Chapter 3. Our ALMARED survey is one such attempt. But we can only derive the lower limit of the number density, which still suffers from the small number statistics. To make a complete survey of these massive haloes, we still need larger campaigns to continue this effort.

A key question about the feedback process in the (proto)cluster is the timescale for the evolution of the host ICM. The X-ray detection of the hot gas in the ICM has been pushed back to the $z \sim 2.5$ (Wang et al., 2016). More recently, the tSZ effect has also been confirmed in the Spiderweb proto-cluster, at $z = 2.16$ (Di Mascolo et al., 2023). These inspiring discoveries have revealed the early evolution of hot gas in the central massive halos of the proto-clusters. The tSZ is redshift independent, allowing such experiments to be adopted at much higher redshifts. Unfortunately, the ALMA time required to detect the tSZ signals in proto-cluster cores is still costly due to their intrinsic faintness. However, by combining the large sky surveys of the proto-cluster cores, it is possible to stack similar structures at different redshift bins to detect their tSZ effects.

References

Akins H. B., et al., 2022, ApJ, 934, 64

Alaghband-Zadeh S., et al., 2013, MNRAS, 435, 1493

Albareti F. D., et al., 2017, ApJ, 233, 25

Alberts S., Noble A., 2022, From Clusters to Proto-clusters: the Infrared Perspective on Environmental Galaxy Evolution (arxiv:2209.02781)

Alexander D. M., Aussel H., Bauer F. E., Brandt W. N., Hornschemeier A. E., Vignali C., Garmire G. P., Schneider D. P., 2002, ApJ, 568, L85

Alexander D. M., Bauer F. E., Chapman S. C., Smail I., Blain A. W., Brandt W. N., Ivison R. J., 2005, ApJ, 632, 736

Algera H. S. B., et al., 2020, ApJ, 903, 138

Álvarez-Márquez J., et al., 2023, MIRI/JWST observations reveal an extremely obscured starburst in the z=6.9 system SPT0311-58 (arxiv:2301.02313)

Andersson B.-G., Lazarian A., Vaillancourt J. E., 2015, ARA&A, 53, 501

Andrews S. M., 2020, Annual Review of Astronomy and Astrophysics, 58, 483

Apostolovski Y., et al., 2019, A&A, 628, A23

Aravena M., et al., 2020, ApJ, 901, 79

Aretxaga I., et al., 2011, MNRAS, 415, 3831

Ata M., Lee K.-G., Vecchia C. D., Kitaura F.-S., Cucciati O., Lemaux B. C., Kashino D., Müller T., 2022, Nat. Astron., 6, 857

Bahcall N. A., Cen R., 1993, ApJ, 407, L49

Bakx T. J. L. C., et al., 2021, arXiv:2108.13479 [astro-ph]

Barger A. J., Cowie L. L., Sanders D. B., Fulton E., Taniguchi Y., Sato Y., Kawara K., Okuda H., 1998, Nat, 394, 248

Barger A. J., Cowie L. L., Mushotzky R. F., Richards E. A., 2001, AJ, 121, 662

Baugh C. M., Lacey C. G., Frenk C. S., Granato G. L., Silva L., Bressan A., Benson A. J., Cole S., 2005, MNRAS, 356, 1191

Beck R., 2015, A&AR, 24, 4

Beck A. M., Dolag K., Lesch H., Kronberg P. P., 2013, MNRAS, 435, 3575

Beck R., Chamandy L., Elson E., Blackman E. G., 2019, Galaxies, 8, 4

Becker R. H., White R. L., Helfand D. J., 1995, ApJ, 450, 559

Behroozi P., Silk J., 2018, MNRAS, 477, 5382

Behroozi P., Wechsler R. H., Hearin A. P., Conroy C., 2019, MNRAS, 488, 3143

Berlok T., Pfrommer C., 2019, MNRAS, 489, 3368

Bernet M. L., Miniati F., Lilly S. J., Kronberg P. P., Dessauges-Zavadsky M., 2008, Nat, 454, 302

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

Bertone S., Vogt C., Enßlin T., 2006, MNRAS, 370, 319

Best P. N., Heckman T. M., 2012, MNRAS, 421, 1569

Béthermin M., De Breuck C., Sargent M., Daddi E., 2015, A&A, 576, L9

Béthermin M., et al., 2020, A&A, 643, A2

Biermann L., 1950, Zeitschrift Naturforschung Teil A, 5, 65

Biggs A. D., Ivison R. J., 2008, MNRAS, 385, 893

Biggs A. D., Younger J. D., Ivison R. J., 2010, MNRAS, 408, 342

Bing L., et al., 2023, NIKA2 Cosmological Legacy Survey: Survey Description and Galaxy Number Counts (arxiv:2305.07054), doi:10.48550/arXiv.2305.07054

Bird J. C., Kazantzidis S., Weinberg D. H., Guedes J., Callegari S., Mayer L., Madau P., 2013, ApJ, 773, 43

Blain A., 2002, PhR, 369, 111

Blain A. W., Longair M. S., 1993, MNRAS, 264, 509

Blain A. W., Longair M. S., 1996, Monthly Notices of the Royal Astronomical Society, 279, 847

Blain A. W., Kneib J. P., Ivison R. J., Smail I., 1999, ApJ, 512, L87

Bleem L. E., et al., 2020, ApJS, 247, 25

Bocquet S., et al., 2019, ApJ, 878, 55

Bolatto A. D., et al., 2013, Nat, 499, 450

Bonato M., et al., 2018, MNRAS, 478, 1512

Bond J. R., Kofman L., Pogosyan D., 1996, Nature, 380, 603

Borlaff A. S., et al., 2023, Extragalactic magnetism with SOFIA (SALSA Legacy Program) – V: First results on the magnetic field orientation of galaxies (arxiv:2303.13586), doi:10.48550/arXiv.2303.13586

Bouwens R. J., et al., 2022, ApJ, 931, 160

Brandenburg A., Subramanian K., 2005, PhR, 417, 1

Brisbin D., et al., 2017, A&A, 608, A15

Buck T., Obreja A., Macciò A. V., Minchev I., Dutton A. A., Ostriker J. P., 2020, MNRAS, 491, 3461

Bunker A. J., et al., 2023, JADES NIRSpec Spectroscopy of GN-z11: Lyman-α emission and possible enhanced nitrogen abundance in a $z=10.60$ luminous galaxy, doi:10.48550/arXiv.2302.07256

Bthermin M., et al., 2017, A&A, 607, A89

Calistro Rivera G., et al., 2018, ApJ, 863, 56

Calvi R., Castignani G., Dannerbauer H., 2023, Submillimeter Galaxies do trace Galaxy Protoclusters (arxiv:2302.10323)

Camps P., et al., 2018, ApJS, 234, 20

Capak P. L., et al., 2011, Nat, 470, 233

Capak P. L., et al., 2015, Nat, 522, 455

Cappellari M., et al., 2012, Nat, 484, 485

Carilli C. L., et al., 2010, ApJ, 714, 1407

Carnall A. C., et al., 2022, A first look at JWST CEERS: massive quiescent galaxies from 3 < z < 5 (arxiv:2208.00986), doi:10.48550/arXiv.2208.00986

Carniani S., et al., 2015, A&A, 584, A78

Carniani S., et al., 2018, MNRAS, 478, 1170

Casey C. M., 2016, ApJ, 824, 36

Casey C. M., Narayanan D., Cooray A., 2014, PhR, 541, 45

Casey C. M., et al., 2018a, ApJ, 862, 77

Casey C. M., Hodge J., Zavala J. A., Spilker J., da Cunha E., Staguhn J., Finkelstein S. L., Drew P., 2018b, ApJ, 862, 78

Ceverino D., Dekel A., Mandelker N., Bournaud F., Burkert A., Genzel R., Primack J., 2012, MNRAS, 420, 3490

Chapman S. C., Blain A. W., Ivison R. J., Smail I. R., 2003, Nat, 422, 695

Chapman S. C., Blain A. W., Smail I., Ivison R. J., 2005, ApJ, 622, 772

Chapman S. C., et al., 2015, MNRAS, 453, 951

Chapman S. C., et al., 2023, Brightest Cluster Galaxy Formation in the z=4.3 Protocluster SPT2349-56: Discovery of a Radio-Loud AGN, doi:10.48550/arXiv.2301.01375

Chartab N., et al., 2022, Gas Phase Metallicities of Local Ultra-Luminous Infrared Galaxies Follow Normal Star-Forming Galaxies (arxiv:2201.07478)

Chen C.-C., Cowie L. L., Barger A. J., Casey Caitlin. M., Lee N., Sanders D. B., Wang W.-H., Williams J. P., 2013, ApJ, 762, 81

Chen C.-C., et al., 2015, ApJ, 799, 194

Chen C.-C., et al., 2016, ApJ, 820, 82

Chen H., et al., 2020, A&A, 638, A113

Chen C.-C., et al., 2022a, ApJ, 929, 159

Chen C.-C., et al., 2022b, ApJ, 939, L7

Chen Y.-H., Heinz S., Hooper E., 2023, A Numerical Study of the Impact of Jet Magnetic Topology on Radio Galaxy Evolution (arxiv:2304.03863)

Cheng C., et al., 2022, JWST's PEARLS: A JWST/NIRCam view of ALMA sources (arxiv:2210.08163), doi:10.48550/arXiv.2210.08163

Chi S., Barthel P. D., Garrett M. A., 2013, A&A, 550, A68

Chiang Y.-K., Overzier R., Gebhardt K., 2013, ApJ, 779, 127

Chyży K. T., Beck R., 2004, A&A, 417, 541

Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168

Colina L., et al., 2023, A&A, 673, L6

Condon J. J., 1992, ARA&A, 30, 575

Condon J. J., Cotton W. D., Greisen E. W., Yin Q. F., Perley R. A., Taylor G. B., Broderick J. J., 1998, AJ, 115, 1693

Cooray A., Sheth R., 2002, Phys. Rep., 372, 1

Cooray A., et al., 2014, ApJ, 790, 40

Coppin K., et al., 2006, MNRAS, 372, 1621

Cormier D., et al., 2015, A&A, 578, A53

Cowie L. L., Gonzlez-Lpez J., Barger A. J., Bauer F. E., Hsu L. Y., Wang W. H., 2018, ApJ, 865, 106

Cowley W. I., Lacey C. G., Baugh C. M., Cole S., Frenk C. S., Lagos C. d. P., 2019, MNRAS, 487, 3082

Cox P., et al., 2023, z-GAL – A NOEMA spectroscopic redshift survey of bright Herschel galaxies: [I] Overview (arxiv:2307.15732)

Crutcher R. M., 2012, ARA&A, 50, 29

Curti M., et al., 2023, MNRAS, 518, 425

Dabringhausen J., Kroupa P., Baumgardt H., 2009, Monthly Notices of the Royal Astronomical Society, 394, 1529

Daddi E., et al., 2009, ApJ, 694, 1517

Daddi E., et al., 2021, A&A, 649, A78

Daddi E., et al., 2022a, The bending of the star-forming main sequence traces the cold- to hot-accretion transition mass over 0 (arxiv:2203.10880)

Daddi E., et al., 2022b, ApJ, 926, L21

Dannerbauer H., et al., 2014, A&A, 570, A55

Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, MNRAS, 486, 2827

Dav R., Angls-Alczar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, MNRAS, 486, 2827

De Looze I., et al., 2014, A&A, 568, A62

De Lucia G., Blaizot J., 2007, MNRAS, 375, 2

DeMaio T., Gonzalez A. H., Zabludoff A., Zaritsky D., Connor T., Donahue M., Mulchaey J. S., 2018, MNRAS, 474, 3009

DeMaio T., et al., 2020, MNRAS, 491, 3751

Decarli R., et al., 2019, ApJ, 882, 138

Decarli R., et al., 2020, ApJ, 902, 110

Dekel A., Sari R., Ceverino D., 2009, AJ, 703, 785

Delvecchio I., et al., 2021, A&A, 647, A123

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

Di Mascolo L., et al., 2023, Nat, 615, 809

Doherty M. J., Geach J. E., Ivison R. J., Dye S., 2020, ApJ, 905, 152

Dole H., et al., 2006, A&A, 451, 417

Drew P. M., Casey C. M., 2022, arXiv:2203.16655 [astro-ph]

Driver S. P., Robotham A. S. G., 2010, MNRAS, 407, 2131

Drzazga R. T., Chyży K. T., Jurusik W., Wiórkiewicz K., 2011, A&A, 533, A22

Dudzevičiūtė U., et al., 2020, MNRAS, 494, 3828

Dunlop J. S., et al., 2017, MNRAS, 466, 861

Dunne L., Maddox S. J., Vlahakis C., Gomez H. L., 2021, MNRAS, 501, 2573

Dunne L., Maddox S. J., Papadopoulos P. P., Ivison R. J., Gomez H. L., 2022, MNRAS, 517, 962

Dye S., et al., 2021, arXiv:2112.03936

Dye S., et al., 2022, MNRAS, 510, 3734

Eales S., Lilly S., Gear W., Dunne L., Bond J. R., Hammer F., Le Fvre O., Crampton D., 1999, ApJ, 515, 518

Eales S., et al., 2010, PASP, 122, 499

Elbaz D., Cesarsky C. J., Chanial P., Aussel H., Franceschini A., Fadda D., Chary R. R., 2002, A&A, 384, 848

Emonts B. H. C., et al., 2018, MNRAS, 477, L60

Emonts B. H. C., et al., 2023, Sci, 379, 1323

Engel H., et al., 2010, ApJ, 724, 233

Evrard A. E., et al., 2008, ApJ, 672, 122

Fabian A. C., et al., 2000, MNRAS, 315, L8

Faucher-Giguère C.-A., Kereš D., Ma C.-P., 2011, MNRAS, 417, 2982

Ferrara A., Vallini L., Pallottini A., Gallerani S., Carniani S., Kohandel M., Decataldo D., Behrens C., 2019, MNRAS, 489, 1

Fixsen D. J., Dwek E., Mather J. C., Bennett C. L., Shafer R. A., 1998, ApJ, 508, 123

Forrest B., et al., 2020, ApJ, 890, L1

Förster Schreiber N. M., et al., 2018, ApJS, 238, 21

Franco M., et al., 2018, A&A, 620, A152

Fraternali F., Karim A., Magnelli B., Gómez-Guijarro C., Jiménez-Andrade E. F., Posses A. C., 2021, arXiv:2011.05340 [astro-ph]

Frias Castillo M., et al., 2023, ApJ, 945, 128

Fudamoto Y., et al., 2021, Nature, 597, 489

Fujimoto S., Ouchi M., Ono Y., Shibuya T., Ishigaki M., Nagai H., Momose R., 2016, ApJS, 222, 1

Fujimoto S., et al., 2019, ApJ, 887, 107

Fujimoto S., et al., 2023, ALMA Lensing Cluster Survey: Deep 1.2 mm Number Counts and Infrared Luminosity Functions at $z\simeq1-8$, doi:10.48550/arXiv.2303.01658

Galliano F., Galametz M., Jones A. P., 2018, ARA&A, 56, 673

Gao Y., Solomon P. M., 2004, ApJ, 606, 271

Geach J. E., et al., 2015, MNRAS, 452, 502

Geach J. E., et al., 2017, MNRAS, 465, 1789

Geach J. E., Ivison R. J., Dye S., Oteo I., 2018, ApJ, 866, L12

Georgantopoulos I., Rovilos E., Comastri A., 2011, A&A, 526, A46

Gillman S., et al., 2023, Sub-Millimetre Galaxies with Webb: Near-Infrared Counterparts and Multi-wavelength Morphology, doi:10.48550/arXiv.2303.17246

Gómez-Guijarro C., et al., 2019, ApJ, 872, 117

González-López J., et al., 2019, ApJ, 882, 139

González-López J., et al., 2020, ApJ, 897, 91

Gonzlez-Lpez J., et al., 2017, A&A, 608, A138

Greve T. R., Ivison R. J., Bertoldi F., Stevens J. A., Dunlop J. S., Lutz D., Carilli C. L., 2004, ApJ, 354, 779

Greve T. R., et al., 2005, Monthly Notices of the Royal Astronomical Society, 359, 1165

Greve T. R., et al., 2014, ApJ, 794, 142

Guiderdoni B., Bouchet F. R., Puget J.-L., Lagache G., Hivon E., 1997, Nat, 390, 257

Guillet V., et al., 2018, A&A, 610, A16

Gullberg B., et al., 2019, MNRAS, 490, 4956

Gururajan G., et al., 2023, A&A, 676, A89

Gmez-Guijarro C., et al., 2022, A&A, 658, A43

Han J., 2017, ARA&A, 55, 111

Hanasz M., Lesch H., Naab T., Gawryszczak A., Kowalik K., Wóltański D., 2013, ApJ, 777, L38

Harikane Y., et al., 2020, ApJ, 896, 93

Hashimoto T., et al., 2019, PASJ, 71, 71

Hatch N. A., Kurk J. D., Pentericci L., Venemans B. P., Kuiper E., Miley G. K., Röttgering H. J. A., 2011, MNRAS, 415, 2993

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

Hatsukade B., et al., 2016, PASJ, 68, 36

Hatsukade B., et al., 2018, PASJ, 70, 105

Hauser M. G., et al., 1998, ApJ, 508, 25

Hayashi M., Kodama T., Tadaki K.-i., Koyama Y., Tanaka I., 2012, ApJ, 757, 15

Hayward C. C., Kere D., Jonsson P., Narayanan D., Cox T. J., Hernquist L., 2011, ApJ, 743, 159

Hayward C. C., Behroozi P. S., Somerville R. S., Primack J. R., Moreno J., Wechsler R. H., 2013, MNRAS, 434, 2572

Hayward C. C., et al., 2021, Monthly Notices of the Royal Astronomical Society, 502, 2922

Heckman T. M., Best P. N., 2014, ARA&A, 52, 589

Heesen V., et al., 2023a, A&A, 669, A8

Heesen V., et al., 2023b, A&A, 670, L23

Heintz K. E., Watson D., Oesch P., Narayanan D., Madden S. C., 2021, arXiv:2108.13442 [astro-ph]

Heintz K. E., et al., 2022, The gas and stellar content of a metal-poor galaxy at $z=8.496$ revealed by JWST and ALMA (arxiv:2212.06877), doi:10.48550/arXiv.2212.06877

Helton J. M., et al., 2023, The JWST Advanced Deep Extragalactic Survey: Discovery of an Extreme Galaxy Overdensity at $z = 5.4$ with JWST/NIRCam in GOODS-S, doi:10.48550/arXiv.2302.10217

Henkel C., Mauersberger R., 1993, Astronomy and Astrophysics, Vol. 274, p. 730-742 (1993), 274, 730

Hensley B. S., Draine B. T., 2023, ApJ, 948, 55

Herrera-Camus R., et al., 2019, ApJ, 871, 37

Hoang T., Lazarian A., 2008, Monthly Notices of the Royal Astronomical Society, 388, 117

Hoang T., Lazarian A., 2016, The Astrophysical Journal, 831, 159

Hodge J. A., da Cunha E., 2020, R. Soc. open sci., 7, 200556

Hodge J. A., Carilli C. L., Walter F., de Blok W. J. G., Riechers D., Daddi E., Lentati L., 2012, ApJ, 760, 11

Hodge J. A., et al., 2013a, ApJ, 768, 91

Hodge J. A., Carilli C. L., Walter F., Daddi E., Riechers D., 2013b, ApJ, 776, 22

Hodge J. A., Riechers D., Decarli R., Walter F., Carilli C. L., Daddi E., Dannerbauer H., 2015, ApJ, 798, L18

Högbom J. A., 1974, Astronomy and Astrophysics Supplement, Vol. 15, p.417, 15, 417

Hogg D. W., Turner E. L., 1998, PASP, 110, 727

Hopkins P. F., Hernquist L., Cox T. J., Di Matteo T., Robertson B., Springel V., 2006, ApJS, 163, 1

Hu W., et al., 2021, Nat. Astron., 5, 485

Huang N., et al., 2020, AJ, 159, 110

Hughes D. H., et al., 1998, 394, 7

Ikarashi S., et al., 2015, ApJ, 810, 133

Ikeda M., Oka T., Tatematsu K., Sekimoto Y., Yamamoto S., 2002, ApJS, 139, 467

Inami H., et al., 2022, arXiv:2203.15136 [astro-ph]

Inoue Y., Doi A., 2018, ApJ, 869, 114

Inoue A. K., Shimizu I., Tamura Y., Matsuo H., Okamoto T., Yoshida N., 2014, ApJ, 780, L18

Inoue A. K., et al., 2016, Sci, 352, 1559

Inoue Y., Khangulyan D., Doi A., 2020, ApJ, 891, L33

Ito K., et al., 2023, ApJ, 945, L9

Ivison R. J., Smail I., Le Borgne J.-F., Blain A. W., Kneib J.-P., Bezecourt J., Kerr T. H., Davies J. K., 1998, MNRAS, 298, 583

Ivison R. J., Dunlop J. S., Smail I., Dey A., Liu M. C., Graham J. R., 2000, The Astrophysical Journal, 542, 27

Ivison R. J., et al., 2002, MNRAS, 337, 1

Ivison R. J., et al., 2007, MNRAS, 380, 199

Ivison R. J., et al., 2010a, MNRAS, 402, 245

Ivison R. J., et al., 2010b, Astronomy and Astrophysics, 518, L31

Ivison R. J., et al., 2010c, A&A, 518, L35

Ivison R. J., Papadopoulos P. P., Smail I., Greve T. R., Thomson A. P., Xilouris E. M., Chapman S. C., 2011, MNRAS, 412, 1913

Ivison R. J., et al., 2012, MNRAS, 425, 1320

Ivison R. J., et al., 2013, ApJ, 772, 137

Ivison R. J., et al., 2016a, ApJ, 832, 78

Ivison R. J., et al., 2016b, ApJ, 832, 78

Ivison R. J., Page M. J., Cirasuolo M., Harrison C. M., Mainieri V., Arumugam V., Dudzevičiūtė U., 2019, MNRAS, 489, 427

Ivison R. J., Richard J., Biggs A. D., Zwaan M. A., Falgarone E., Arumugam V., van der Werf P. P., Rujopakarn W., 2020, MNRAS, 495, L1

Jeřábková T., Hasani Zonoozi A., Kroupa P., Beccari G., Yan Z., Vazdekis A., Zhang Z. Y., 2018, Astronomy and Astrophysics, 620, A39

Jin S., et al., 2019, ApJ, 887, 144

Jin S., et al., 2023, A&A, 670, L11

Jones G. C., Maiolino R., Caselli P., Carniani S., 2019a, A&A, 632, L7

Jones T. J., et al., 2019b, ApJ, 870, L9

Juvela M., Mattila K., Lemke D., Klaas U., Leinert C., Kiss C., 2009, A&A, 500, 763

Karim A., et al., 2013, MNRAS, 432, 2

Kewley L. J., Nicholls D. C., Sutherland R. S., 2019, ARA&A, 57, 511

Khoperskov S. A., Khrapov S. S., 2018, A&A, 609, A104

Killi M., et al., 2023, MNRAS, 521, 2526

Kirkpatrick A., Sharon C., Keller E., Pope A., 2019, ApJ, 879, 41

Klitsch A., et al., 2019a, MNRAS, 490, 1220

Klitsch A., et al., 2019b, Monthly Notices of the Royal Astronomical Society, 490, 1220

Klitsch A., et al., 2020, MNRAS, 495, 2332

Knudsen K. K., Richard J., Kneib J.-P., Jauzac M., Clément B., Drouart G., Egami E., Lindroos L., 2016, MNRAS, 462, L6

Kocevski D. D., et al., 2012, ApJ, 744, 148

Kretschmer M., Dekel A., Teyssier R., 2022, MNRAS, 510, 3266

Kronberg P. P., Bernet M. L., Miniati F., Lilly S. J., Short M. B., Higdon D. M., 2008, ApJ, 676, 70

Krumholz M. R., Federrath C., 2019, Frontiers in Astronomy and Space Sciences, 6

Kubo M., et al., 2021, ApJ, 919, 6

Kurk J. D., et al., 2000, A&A, 358, L1

Kurk J. D., Pentericci L., Overzier R. A., Röttgering H. J. A., Miley G. K., 2004, A&A, 428, 817

Lacey C. G., et al., 2016, MNRAS, 462, 3854

Lacki B. C., Thompson T. A., Quataert E., 2010, ApJ, 717, 1

Lagache G., Puget J.-L., Dole H., 2005, ARA&A, 43, 727

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

Lagos C. d. P., et al., 2019, MNRAS, 489, 4196

Lagos C. d. P., da Cunha E., Robotham A. S. G., Obreschkow D., Valentino F., Fujimoto S., Magdis G. E., Tobar R., 2020, MNRAS, 499, 1948

Laird E. S., Nandra K., Pope A., Scott D., 2010, MNRAS, 401, 2763

Lammers C., Hill R., Lim S., Scott D., Cañameras R., Dole H., 2022, MNRAS, 514, 5004

Laporte N., et al., 2017a, ApJ, 837, L21

Laporte N., et al., 2017b, ApJ, 837, L21

Laporte N., et al., 2021, MNRAS, 505, 4838

Laporte N., Zitrin A., Dole H., Roberts-Borsani G., Furtak L. J., Witten C., 2022, A lensed protocluster candidate at $z=7.66$ identified in JWST observations of the galaxy cluster SMACS0723-7327 (arxiv:2208.04930)

Laurent G. T., et al., 2005, ApJ, 623, 742

Lelli F., Di Teodoro E. M., Fraternali F., Man A. W. S., Zhang Z.-Y., De Breuck C., Davis T. A., Maiolino R., 2021, Sci, 371, 713

Lelli F., et al., 2023, Cold gas disks in main-sequence galaxies at cosmic noon: Low turbulence, flat rotation curves, and disk-halo degeneracy (arxiv:2302.00030)

Lewis A. J. R., et al., 2018, ApJ, 862, 96

Li Q., et al., 2020, ApJ, 900, 12

Li J., Liu C., Zhang Z.-Y., Tian H., Fu X., Li J., Yan Z.-Q., 2023, Nat, 613, 460

Liang L., et al., 2019, MNRAS, 489, 1397

Liang L., et al., 2023, $\rm [C_{II}]$ 158 $\rm \mu m$ emission as an indicator of galaxy star formation rate, doi:10.48550/arXiv.2301.04149

Lilly S. J., Eales S. A., Gear W. K. P., Hammer F., Le Fèvre O., Crampton D., Bond J. R., Dunne L., 1999, ApJ, 518, 641

Lindner R. R., Baker A. J., Beelen A., Owen F. N., Polletta M., 2012, ApJ, 757, 3

Litke K. C., et al., 2019, ApJ, 870, 80

Litke K. C., et al., 2022, Multi-Phase ISM in the z = 5.7 Hyperluminous Starburst SPT0346-52

Liu D., et al., 2019, ApJ, 887, 235

Liu A., et al., 2022, A&A, 661, A2

Liu D., et al., 2023, ApJ, 942, 98

Long A. S., et al., 2020, ApJ, 898, 133

Lopez-Rodriguez E., 2021, Nat. Astron., pp 1–11

Lopez-Rodriguez E., 2023, The magnetic fields of starburst galaxies. I. Identification and characterization of the thermal polarization in the galactic disk and outflow (arxiv:2306.10099), doi:10.48550/arXiv.2306.10099

Lopez-Rodriguez E., et al., 2020, ApJ, 888, 66

Lopez-Rodriguez E., et al., 2022, Extragalactic magnetism with SOFIA (SALSA Legacy Program). VI. The magnetic fields in the multi-phase interstellar medium of the Antennae galaxies (arxiv:2211.00012), doi:10.48550/arXiv.2211.00012

Lovell C. C., Thomas P. A., Wilkins S. M., 2018, MNRAS, 474, 4612

Lovell C. C., Geach J. E., Davé R., Narayanan D., Li Q., 2021, MNRAS, 502, 772

Ma J., et al., 2016, ApJ, 832, 114

MacKenzie T. P., et al., 2017, MNRAS, 468, 4006

Magnelli B., et al., 2014, A&A, 561, A86

Magnelli B., et al., 2015, A&A, 573, A45

Magnelli B., et al., 2019, ApJ, 877, 45

Magnelli B., et al., 2020, ApJ, 892, 66

Maiolino R., Mannucci F., 2019, A&AR, 27, 3

Mao S. A., et al., 2017, Nat. Astron., 1, 621

Marinacci F., et al., 2018, MNRAS, 480, 5113

Matsuda Y., et al., 2004, AJ, 128, 569

Matsuura S., et al., 2011, ApJ, 737, 2

McAlpine S., et al., 2019, MNRAS, 488, 2440

McMullin J. P., Waters B., Schiebel D., Young W., Golap K., 2007, Astronomical Data Analysis Software and Systems XVI ASP Conference Series, 376, 127

Meng X., Gnedin O. Y., 2021, MNRAS, 502, 1433

Meyer R. A., et al., 2022, ApJ, 927, 141

Miettinen O., et al., 2017, A&A, 606, A17

Miley G., De Breuck C., 2008, A&AR, 15, 67

Miller T. B., et al., 2018, Nat, 556, 469

Molnár D. C., et al., 2021, MNRAS, 504, 118

Momjian E., Wang W.-H., Knudsen K. K., Carilli C. L., Cowie L. L., Barger A. J., 2010, AJ, 139, 1622

Monson E. B., et al., 2021, ApJ, 919, 51

Montenegro-Taborda D., Rodriguez-Gomez V., Pillepich A., Avila-Reese V., Sales L. V., Rodríguez-Puebla A., Hernquist L., 2023, MNRAS, 521, 800

Morishita T., et al., 2023, Early results from GLASS-JWST. XVIII: A spectroscopically confirmed protocluster 650 million years after the Big Bang (arxiv:2211.09097)

Muñoz Arancibia A. M., et al., 2018, A&A, 620, A125

Muñoz Arancibia A. M., et al., 2023, A&A, 675, A85

Muoz Arancibia A. M., et al., 2019, A&A, 631, C2

Nagao T., Maiolino R., Marconi A., Matsuhara H., 2011, A&A, 526, A149

Nakajima K., Ouchi M., Isobe Y., Harikane Y., Zhang Y., Ono Y., Umeda H., Oguri M., 2023, JWST Census for the Mass-Metallicity Star-Formation Relations at z=4-10 with the Self-Consistent Flux Calibration and the Proper Metallicity Calibrators, doi:10.48550/arXiv.2301.12825

Nanayakkara T., et al., 2022, A population of faint, old, and massive quiescent galaxies at 3 < z < 4 revealed by JWST NIRSpec Spectroscopy (arxiv:2212.11638)

Narayanan D., Krumholz M. R., 2017, MNRAS, 467, 50

Neeleman M., Prochaska J. X., Kanekar N., Rafelski M., 2020, Nat, 581, 269

Neri R., et al., 2003, ApJ, 597, L113

Neri R., et al., 2020, A&A, 635, A7

Nowotka M., Chen C.-C., Battaia F. A., Fumagalli M., Cai Z., Lusso E., Prochaska J. X., Yang Y., 2022, A&A, 658, A77

Odegard N., Weiland J. L., Fixsen D. J., Chuss D. T., Dwek E., Kogut A., Switzer E. R., 2019, ApJ, 877, 40

Oliver S. J., et al., 2010, A&A, 518, L21

Ono Y., Ouchi M., Kurono Y., Momose R., 2014, ApJ, 795, 5

Oren A. L., Wolfe A. M., 1995, ApJ, 445, 624

Ota K., et al., 2014, ApJ, 792, 34

Oteo I., Zwaan M. A., Ivison R. J., Smail I., Biggs A. D., 2016a, ApJ, 822, 36

Oteo I., et al., 2016b, ApJ, 827, 34

Oteo I., et al., 2017a, ApJ, 850, 170

Oteo I., et al., 2017b, arXiv e-prints, 1709, arXiv:1709.04191

Oteo I., et al., 2018, ApJ, 856, 72

Ouchi M., et al., 2005, ApJ, 620, L1

Overzier R. A., 2016, A&AR, 24, 14

Overzier R. A., Nesvadba N. P. H., Dijkstra M., Hatch N. A., Lehnert M. D., Villar-Martín M., Wilman R. J., Zirm A. W., 2013, ApJ, 771, 89

Pakmor R., et al., 2017, MNRAS, 469, 3185

Papadopoulos P. P., Thi W. F., Viti S., 2004, MNRAS, 351, 147

Peng B., et al., 2023, ApJL, 944, L36

Pentericci L., Kurk J. D., Carilli C. L., Harris D. E., Miley G. K., Röttgering H. J. A., 2002, A&A, 396, 109

Pentericci L., et al., 2016, ApJ, 829, L11

Pereira-Santaella M., Rigopoulou D., Farrah D., Lebouteiller V., Li J., 2017, MNRAS, 470, 1218

Perry R. W., Chapman S. C., Smail I., Bertoldi F., 2023, Resolving a merger in a hyper-luminous submillimeter galaxy at z=2.82 (arxiv:2210.08191)

Pillepich A., et al., 2019, MNRAS, 490, 3196

Pilyugin L. S., Grebel E. K., 2016, MNRAS, 457, 3678

Planck Collaboration X., et al., 2014, A&A, 571, A30

Planck Collaboration et al., 2015, A&A, 576, A104

Planck Collaboration V., et al., 2020, A&A, 641, A6

Pope A., et al., 2008, ApJ, 675, 1171

Popping G., et al., 2020, ApJ, 891, 135

Pshirkov M. S., Tinyakov P. G., Urban F. R., 2016, Phys. Rev. Lett., 116, 191302

Puget J.-L., Abergel A., Bernard J.-P., Boulanger F., Burton W. B., Desert F.-X., Hartmann D., 1996, A&A, 308, L5

Reuter C., et al., 2020, ApJ, 902, 78

Rhoads J. E., Malhotra S., Dey A., Stern D., Spinrad H., Jannuzi B. T., 2000, The Astrophysical Journal, 545, L85

Ricci C., et al., 2023, A Tight Correlation Between Millimeter and X-ray Emission in Accreting Massive Black Holes from <100 Milliarcsecond-resolution ALMA Observations (arxiv:2306.04679)

Riechers D. A., et al., 2014, ApJ, 786, 31

Riechers D. A., Weiss A., Walter F., Carilli C. L., Cox P., Decarli R., Neri R., 2022, Nat, 602, 58

Rigopoulou D., Pereira-Santaella M., Magdis G. E., Cooray A., Farrah D., Marques-Chaves R., Perez-Fournon I., Riechers D., 2018, MNRAS, 473, 20

Rizzo F., Vegetti S., Powell D., Fraternali F., McKean J. P., Stacey H. R., White S. D. M., 2020, Nat, 584, 201

Rizzo F., Vegetti S., Fraternali F., Stacey H., Powell D., 2021, arXiv:2102.05671 [astro-ph]

Rizzo F., Kohandel M., Pallottini A., Zanella A., Ferrara A., Vallini L., Toft S., 2022, arXiv:2204.05325 [astro-ph]

Robertson B. E., et al., 2023, Nat. Astron., pp 1–11

Robson E. I., et al., 1983, Nature, 305, 194

Robson E. I., et al., 2014, ApJ, 793, 11

Rodrigues L. F. S., Chamandy L., Shukurov A., Baugh C. M., Taylor A. R., 2019, MNRAS, 483, 2424

Romano D., Matteucci F., Zhang Z. Y., Papadopoulos P. P., Ivison R. J., 2017, MNRAS, 470, 401

Romano D., Matteucci F., Zhang Z.-Y., Ivison R. J., Ventura P., 2019, MNRAS, 490, 2838

Romano M., et al., 2022, A&A, 660, A14

Rosati P., et al., 2009, Astronomy and Astrophysics, Volume 508, Issue 2, 2009, pp.583-591, 508, 583

Rybak M., et al., 2019, ApJ, 876, 112

Rybak M., et al., 2022, A&A, 667, A70

Rybak M., et al., 2023, PRUSSIC II – ALMA imaging of dense-gas tracers in SDP.81: Evidence for low mechanical heating and a sub-solar metallicity in a z=3.04 dusty galaxy, doi:10.48550/arXiv.2308.02886

S. H. E. S. SHES., et al., 2023, Constraints on the intergalactic magnetic field using Fermi-LAT and H.E.S.S. blazar observations (arxiv:2306.05132), doi:10.48550/arXiv.2306.05132

Safarzadeh M., Lu Y., Hayward C. C., 2017, MNRAS, 472, 2462

Sanders D. B., Mirabel I. F., 1996, ARA&A, 34, 749

Sandles L., et al., 2023, JADES: Balmer Decrement Measurements at redshifts 4 < z < 7, doi:10.48550/arXiv.2306.03931

Satyapal S., Ellison S. L., McAlpine W., Hickox R. C., Patton D. R., Mendel J. T., 2014, MNRAS, 441, 1297

Schaerer D., Boone F., Zamojski M., Staguhn J., Dessauges-Zavadsky M., Finkelstein S., Combes F., 2015, A&A, 574, A19

Schaerer D., et al., 2020, A&A, 643, A3

Schaerer D., Marques-Chaves R., Oesch P., Naidu R., Barrufet L., Izotov Y. I., Guseva N. G., 2022, First look with JWST spectroscopy: $z \sim 8$ galaxies resemble local analogues (arxiv:2207.10034), doi:10.48550/arXiv.2207.10034

Schaye J., et al., 2015, MNRAS, 446, 521

Schechter P., 1976, ApJ, 203, 297

Schneider N., et al., 2023, Nat. Astron., 7, 546

Schober J., Schleicher D. R. G., Klessen R. S., 2013, Astronomy & Astrophysics, Volume 560, id.A87, $¡$NUMPAGES$¿$13$¡$/NUMPAGES$¿$ pp., 560, A87

Scott K. S., et al., 2012, MNRAS, 423, 575

Scoville N., et al., 2016, ApJ, 820, 83

Seymour N., et al., 2007, ApJS, 171, 353

Shapley A. E., Sanders R. L., Reddy N. A., Topping M. W., Brammer G. B., 2023, JWST/NIRSpec Balmer-line Measurements of Star Formation and Dust Attenuation at z˜3-6, doi:10.48550/arXiv.2301.03241

Shimizu I., Yoshida N., Okamoto T., 2012, MNRAS, 427, 2866

Simpson J. M., et al., 2015a, ApJ, 807, 128

Simpson J. M., et al., 2015b, ApJ, 807, 128

Simpson J. M., et al., 2017, ApJ, 839, 58

Simpson J. M., et al., 2020, MNRAS, 495, 3409

Smail I., Ivison R. J., Blain A. W., 1997, ApJ, 490, L5

Smail I., Ivison R. J., Owen F. N., Blain A. W., Kneib J. P., 2000, ApJ, 528, 612

Smail I., Ivison R. J., Blain A. W., Kneib J. P., 2002, MNRAS, 331, 495

Smail I., et al., 2021, MNRAS, 502, 3426

Smit R., et al., 2018, Nat, 553, 178

Smith J. D. T., et al., 2017, ApJ, 834, 5

Solanki S. K., Inhester B., Schüssler M., 2006, Rep. Prog. Phys., 69, 563

Somerville R. S., Gilmore R. C., Primack J. R., Domnguez A., 2012, MNRAS, 423, 1992

Sommovigo L., et al., 2022, MNRAS, 513, 3122

Spilker J. S., et al., 2015, ApJ, 811, 124

Spilker J. S., et al., 2016, ApJ, 826, 112

Spilker J. S., et al., 2020, ApJ, 905, 85

Stach S. M., et al., 2018, ApJ, 860, 161

Stach S. M., et al., 2019, MNRAS, 487, 4648

Staguhn J. G., et al., 2014, ApJ, 790, 77

Stanford S. A., et al., 2006, ApJ, 646, L13

Stanley F., et al., 2023, ApJ, 945, 24

Steidel C. C., Giavalisco M., Pettini M., Dickinson M., Adelberger K. L., 1996, ApJ, 462, L17

Steidel C. C., Adelberger K. L., Dickinson M., Giavalisco M., Pettini M., Kellogg M., 1998, ApJ, 492, 428

Steidel C. C., Adelberger K. L., Shapley A. E., Pettini M., Dickinson M., Giavalisco M., 2000, ApJ, 532, 170

Stevens J. A., et al., 2003, Nature, 425, 264

Sugahara Y., et al., 2021, ApJ, 923, 5

Swinbank A. M., Chapman S. C., Smail I., Lindner C., Borys C., Blain A. W., Ivison R. J., Lewis G. F., 2006, MNRAS, 371, 465

Swinbank A. M., et al., 2010, Nat, 464, 733

Swinbank A. M., Sobral D., Smail I., Geach J. E., Best P. N., McCarthy I. G., Crain R. A., Theuns T., 2012, MNRAS, 426, 935

Tacconi L. J., et al., 2006, The Astrophysical Journal, 640, 228

Tacconi L. J., et al., 2008, ApJ, 680, 246

Tacconi L. J., Genzel R., Sternberg A., 2020, ARA&A, 58, 157

Tamfal T., Mayer L., Quinn T. R., Babul A., Madau P., Capelo P. R., Shen S., Staub M., 2021, arXiv:2106.11981 [astro-ph]

Tamura Y., et al., 2019, ApJ, 874, 27

Tamura Y., et al., 2023, ApJ, 952, 9

Tanaka I., et al., 2011, PASJ, 63, S415

Taylor A. J., Barger A. J., Cowie L. L., 2022, ApJ, 939, L3

Tchekhovskoy A., Narayan R., McKinney J. C., 2011, MNRAS, 418, L79

Teyssier R., Chapon D., Bournaud F., 2010, ApJ, 720, L149

The CASA Team T. C. T., et al., 2022, PASP, 134, 114501

Thompson T. A., Quataert E., Waxman E., Murray N., Martin C. L., 2006, ApJ, 645, 186

Toft S., et al., 2014, ApJ, 782, 68

Toshikawa J., et al., 2018, PASJ, 70, S12

Toshikawa J., Malkan M. A., Kashikawa N., Overzier R., Uchiyama H., Ota K., Ishikawa S., Ito K., 2020, ApJ, 888, 89

Trayford J. W., Lagos C. d. P., Robotham A. S. G., Obreschkow D., 2020, MNRAS, 491, 3937

Trump J. R., et al., 2023, ApJ, 945, 35

Tsukui T., Iguchi S., 2021, Sci, 372, 1201

Ueda Y., et al., 2018, ApJ, 853, 24

Uematsu R., et al., 2023, ALMA Lensing Cluster Survey: Properties of Millimeter Galaxies Hosting X-ray Detected Active Galactic Nuclei (arxiv:2301.09275)

Umehata H., et al., 2015, ApJ, 815, L8

Umehata H., et al., 2017, ApJ, 835, 98

Umehata H., et al., 2018, PASJ, 70, 65

Umehata H., et al., 2019, Sci, 366, 97

Urquhart S. A., et al., 2022, MNRAS, 511, 3017

Valentino F., et al., 2018, ApJ, 869, 27

Valentino F., et al., 2023, ApJ, 947, 20

Veilleux S., Maiolino R., Bolatto A. D., Aalto S., 2020, A&AR, 28, 2

Venemans B. P., et al., 2007, Astronomy and Astrophysics, Volume 461, Issue 3, January III 2007, pp.823-845, 461, 823

Vieira J. D., et al., 2010, ApJ, 719, 763

Vieira J. D., et al., 2013, Nat, 495, 344

Vito F., et al., 2020, A&A, 642, A149

Voelk H. J., 1989, A&A, 218, 67

Voit G. M., 2005, Reviews of Modern Physics, 77, 207

Walter F., et al., 2012, Nat, 486, 233

Walter F., et al., 2016, ApJ, 833, 67

Walter F., et al., 2020, ApJ, 902, 111

Wang W.-H., Cowie L. L., Barger A. J., Williams J. P., 2011, ApJ, 726, L18

Wang S. X., et al., 2013, ApJ, 778, 179

Wang T., et al., 2016, ApJ, 828, 56

Wang W.-H., et al., 2017, ApJ, 850, 37

Wang G. C. P., et al., 2021, MNRAS, 508, 3754

Watson D., Christensen L., Knudsen K. K., Richard J., Gallazzi A., Michałowski M. J., 2015, Nat, 519, 327

Weiß A., et al., 2013, ApJ, 767, 88

Wei A., et al., 2009, ApJ, 707, 1201

Whittingham J., Sparre M., Pfrommer C., Pakmor R., 2021, MNRAS, 506, 229

Willis J. P., et al., 2020, Nat, 577, 39

Wisnioski E., et al., 2019, ApJ, 886, 124

Witstok J., et al., 2022, MNRAS, 515, 1751

Wolfe A. M., Jorgenson R. A., Robishaw T., Heiles C., Prochaska J. X., 2008, Nat, 455, 638

Wolfe A. M., Jorgenson R. A., Robishaw T., Heiles C., Prochaska J. X., 2011, The Astrophysical Journal, 733, 24

Wu J., Evans II N. J., Gao Y., Solomon P. M., Shirley Y. L., Vanden Bout P. A., 2005, ApJ, 635, L173

Wylezalek D., et al., 2013a, ApJ, 769, 79

Wylezalek D., et al., 2013b, ApJ, 769, 79

Wylezalek D., et al., 2022, First results from the JWST Early Release Science Program Q3D: Turbulent times in the life of a $z \sim 3$ extremely red quasar revealed by NIRSpec IFU (arxiv:2210.10074)

Xiao M.-Y., et al., 2022, A&A, 664, A63

Yang C., et al., 2017, A&A, 608, A144

Yun M. S., Reddy N. A., Condon J. J., 2001, ApJ, 554, 803

Yung L. Y. A., Somerville R. S., Finkelstein S. L., Wilkins S. M., Gardner J. P., 2023, Are the ultra-high-redshift galaxies at z > 10 surprising in the context of standard galaxy formation models? (arxiv:2304.04348), doi:10.48550/arXiv.2304.04348

Zanella A., et al., 2018, MNRAS, 481, 1976

Zavala J. A., Casey C. M., da Cunha E., Spilker J., Staguhn J., Hodge J., Drew P. M., 2018, ApJ, 869, 71

Zavala J. A., et al., 2021, ApJ, 909, 165

Zeballos M., et al., 2018, MNRAS, 479, 4577

Zhang Z.-Y., Papadopoulos P. P., Ivison R. J., Galametz M., Smith M. W. L., Xilouris E. M., 2016, R. Soc. open sci., 3, 160025

Zhang Z.-Y., et al., 2018a, MNRAS, 481, 59

Zhang Z.-Y., Romano D., Ivison R. J., Papadopoulos P. P., Matteucci F., 2018b, Nat, 558, 260

Zhang Y., et al., 2022, MNRAS, 512, 4893

Zirm A. W., et al., 2008, ApJ, 680, 224

Zwaan M., et al., 2022, Messenger, pp. 10-13, 4 pages

da Cunha E., et al., 2013, ApJ, 766, 13

van de Voort F., Bieri R., Pakmor R., Gómez F. A., Grand R. J. J., Marinacci F., 2021, MNRAS, 501, 4888

Acknowledgements

This thesis was completed under the supervison of Prof. Dr. Rob Ivison with support from European Southern Observatory (ESO).

Pursuing a PhD is an adventure for me. It is filled with cheerfulness, peacefulness, sadness, and depression. All of them combined make this journey memorable and colourful. I've been very fortunate to have received countless help along this journey. I could not have completed my PhD project without the continuous support from my supervisors, collaborators, friends, ESO and my beloved parents.

I want to give special thanks to my supervisor Prof. Dr. Rob Ivison. He has introduced me to the world of dusty star-forming galaxies and provided me with detailed guidance throughout my PhD. His presence and engaging instructions are essential to this thesis. He also continuously encouraged me to develop my projects and design my future career pathway. I cannot forget the nights when we worked together on multiple proposals right before the deadline, and the moments when we scoff at random papers. My very first ALMA proposal requesting the full-polarisation observation of dust in three SMGs would not have been successful without his encouragement and help. Rob is not only my supervisor, he is also a great friend. I started my PhD during the hardest time of covid19. He is always there when I need guidance and help. His encouraging words and praise are keeping me moving forward.

I want to thank all my supervisors and mentors. Dr. Martin Zwaan has offered me vital help in the ALMACAL project. He initialized and led the data reduction of the ALMACAL project and he is always happy to help me whenever I have questions about the ALMA and ALMACAL. Dr. Celine Peroux is a fantastic collaborator and friend. She is always ready to help and offers me lots of instant feedback and guidance. I learned a lot about the circum-galactic medium because of her endless efforts in organising the CGM meetings. I also want to thank Andy Biggs, he is a wonderful person to collaborate with. He offered me crucial help in designing the full-polarisation ALMA projects when I knew nothing about polarisation observation. I also want to thank my former supervisor Prof. Dr. Zhang Zhi-Yu. He taught me all the basics of the interferometer and trained me to think critically, which are the stepping stones to my PhD. Dr. Gabriela Calistro Rivera is my charming ESO mentor. She always thinks of me during my PhD and is always willing to share her rich experience. She has given me an excellent example of how to write a competitive job application.

I want to thank Dr. Linda Tacconi and Prof. Dr. Volker Springel for being part of my PhD committee. I always admire their great achievement in their career. They both gave me very insightful comments and suggestions on my PhD project and offered me plenty of encouragement

and useful tips during my PhD. Linda also helped me a lot by supporting me to continue my research in the Max Plack IR group. I cannot wait longer to work more closely with her.

I want to thank all my wonderful collaborators during my PhD. Dr. Anne Klitsch has given me crucial help at the beginning of the ALMACAL project. She continuously shared her experience with me and helped me to debug and design my code. Prof. Dr. Ian Smail has also actively joined the ALMACAL project and always gives critical comments. It was a wonderful experience discussing science questions with him. I learned a lot from his extensive knowledge and insightful opinions. I am also grateful to Prof. Dr. Mark Swinbank, who has taught me the basics of kinematic/dynamic analysis and offered useful suggestions for my projects related to kinematic modelling. I am mainly an observational astronomer, but I have received loads of help from Dr. Gergö Popping, Dr. Claudia Lagos Urbina, and Dr. Christopher Lovell regarding the theoretical analysis with various modelling technics. I also want to thank my officemate Roland Szakacs, who continuously shared his valuable experience with me and offered me countless help. I am also lucky to work with James Geach, Enrique Lopez Rodriguez, George Bendo, Simon Dye, and Johan Richard on the fascinating magnetised early Universe.

I want to thank my close collaborator and – friend Natsuki Hayatsu. I enjoyed all the discussions with her. Her passion and carefulness towards astronomy will keep motivating me in my future research life. She also hosted me in Ishigaki Observatory during the last month of my PhD. It is a memorable place where I have written an important part of my PhD thesis. I also want to thank my amazing friends: Avinash, Dominika, Pierrick, and Arseniy for their companionship during the past three years. They have given me new meanings towards life beyond research.

I am grateful to be part of ESO. Its open and collaborative environment has fostered fruitful discussions. The office of science and student community have greatly supported me during the COVID lockdown and provided me with all the essential needs regarding my research and life. I want to thank our beloved secretary Ms. Nelma Silva. She has done an incredible job in helping me with my daily life in dealing with all the administrative things.

In the end, I want to thank my parents for their endless love and support.