Strong lensing with ALMA: resolving the nature of high-redshift galaxies

Matus Rybak

München

Strong lensing with ALMA: resolving the nature of high-redshift galaxies

Matus Rybak

Dissertation
an der Fakultät für Physik
der Ludwig–Maximilians–Universität
München

vorgelegt von
Matus Rybak
aus Poprad, Slowakei

München, den 1. März 2017

Erstgutachter: Prof. Dr. Simon D. M. White
Zweitgutachter: Dr. Torsten Enßlin
Tag der mündlichen Prüfung: 5. April 2017

Zusammenfassung

Die vorliegende Arbeit stellt eine neuartige Methode zur Modellierung radiointerferometrischer Beobachtungen von ausgedehnten Quellen vor, welche starken Gravitationslinseneffekten unterworfen sind. Konkret erweitern wir die Bayesische Linsenmodellierungsmethode aus Vegetti und Koopmans (2009) auf das Gebiet der Interferometrie. Durch direkte Anpassung des beobachteten Interferenzkontrastes *(visibility function)* und Rekonstruktion der Oberflächenhelligkeit auf der Quellenebene eines adaptives Gitters, stellt diese Technik eine signifikante Verbesserung gegenüber Linsenmodellierungsmethoden dar, die das synthetische Bild anpassen (Dye et al., 2014), sowie gegenüber visibility-fitting-Methoden, die annehmen die Oberflächenhelligkeit der Quelle einem folge einem analytischen Profil (Bussman et al., 2013; Hezaveh et al., 2013b; Bussman et al., 2015; Spilker et al., 2016). Die Fähigkeit komplexe Strukturen in ausgedehnten Quellen zu erfassen, macht diese Methode für Linsenmodellierungen in hochaufgelösten Beobachtungen mittels großer Radiointerferometer wie dem Atacama Large Millimeter/submillimeter Array (ALMA) oder dem Global VLBI besonders geeignet.

Unsere initialen Tests, durchgeführt an simulierten ALMA-Beobachtungen einer durch eine starke Gravitationslinse abgebildeten Submillimeter-Starburstgalaxie, bestätigen die Fähigkeit, das korrekte Linsenmodell basierend auf Daten zu erhalten, die eine breite Spanne in Winkelauflösung und Sensitivität abdecken. Wir identifizieren eine systematische Abweichung (*bias*) in der intrinsischen Quellengröße und -helligkeit, die aus den Beobachtungen niedrigster Auflösung (<0.5 Bogensekunden) abgeleitet werden: dies ist das Resultat kompakter Strukturen der Quelle, welche nicht aufgelöst werden. Bei höherer Auflösung wird die komplexe Quellenstruktur korrekt rekonstruiert.

Wir demonstrieren die Leistungsfähigkeit unserer Technik an ALMA-Beobachtungen einer sog. *dusty star-forming galaxy* (DSFG), SDP.81 ($z{\sim}3$). Indem die Oberflächenhelligkeitsverteilungen des Staub-Kontinuums und der CO-Emissionslinen auf 50- bzw. 100-parsec-Skalen aufgelöst werden, stellt diese Analyse eine Verbesserung um eine Größenordnung gegenüber früheren Untersuchungen von DSFGs bei hoher Rotverschiebung dar. Diese hochaufgelösten Rekonstruktionen offenbaren sowohl kompakte als auch ausgedehnte Sternenentstehungsgebiete in einer Gasscheibe, wie auch Beweise für eine kürzliche oder noch andauernde Verschmelzung.

Weiterhin untersuchen wir die systematischen Unterschiede zwischen Linsenmodellen, die auf pixel-basierten oder parametrischen Quellenmodellen beruhen. Namentlich modellieren wir sechs Linsen aus dem *South Pole Telescope* lens sample (Vieira et al., 2013), die mit ALMA bei einer Auflösung von 0.5-1.0 Bogensekunden beobachtet wurden. Obwohl die zusätzliche Freiheit pixel-basierter Methoden weitere Entartungen verursacht, werden für mehrere Quellen komplexe Strukturen ermittelt. Dies hat zur Folge, dass Vergrößerungen, die durch pixel-basierte Methoden erhalten wurden, zum Teil um ein Vielfaches von veröffentlichten Ergebnissen, die auf parametrischen Methoden beruhen (Hezaveh et al., 2013b; Spilker et al., 2016), abweichen können. Diese Erkenntnisse heben die Einschränkungen der Ableitung von Quelleneigenschaften aus niedrig-aufgelösten Daten hervor.

Schlussendlich nutzen wir unsere Technik um die CO (2-1) und Millimeterwellenkontinuumsemission eines Quasars bei geringer Rotverschiebung, RXJ1131-1231, der einem starken Gravitationslinseneffekt unterliegt, zu rekonstruieren. Diese Analyse zeigt die Spiralstruktur des Gases in der umgebenden Galaxie und setzt Grenzen für die Größe der sub-mm-Emissionsregion im aktiven galaktischen Kern, welche sich als kompakter als die broad-line-Region herausstellt.

Summary

This thesis introduces a novel technique for modelling radio-interferometric observations of strongly gravitationally lensed extended sources in the visibility-space.

By fitting directly the observed visibility function and reconstructing the source-plane surface brightness distribution on an adaptive grid, this technique presents a significant improvement over lens-modelling techniques that fit the synthesised image (e.g. Dye et al., 2014), or visibility-fitting methods that assume the source surface brightness distribution to follow an analytic profile (Bussmann et al., 2013; Hezaveh et al., 2013b; Bussmann et al., 2015; Spilker et al., 2016). The ability to capture the complex structure of extended sources makes it especially suited for analysis of high angular resolution observations with large radio-interferometers such as the Atacama Large Millimeter/submillimeter Array (ALMA) or the Global Very Long Baseline Interferometry (VLBI).

Our initial tests on simulated ALMA observations of a realistic strongly-lensed starburst galaxy, confirm the ability of our technique to recover the true lens model from a data spanning a wide range in angular resolution and sensitivity. We find bias in intrinsic source size and brightness inferred from the lowest resolution observations (<0.5 arcseconds), a result of compact source structures not being resolved. At higher resolutions, the complex structure of the source is reconstructed correctly.

We demonstrate the power of our technique on ALMA observations of the dusty star-forming galaxy SDP.81 (z∼3). By resolving the surface brightness distribution of the dust continuum and CO emission lines on 50- and 100-parsec scales, respectively, this analysis presents an order of magnitude improvement over earlier studies of high-redshift dusty star-forming galaxies. These high-resolution reconstructions reveal both compact and extended star-formation embedded in a gaseous disk, as well as evidence for a recent or on-going merger.

Second, we investigate the differences between lens- and source models obtained with pixellated and parametric source models, respectively, by modelling a set of six lenses from the *South Pole Telescope* lens sample (Vieira et al., 2013) observed with ALMA at a 0.5-1.0 arcsecond resolution. Although our lens models are in good agreement with the previously published ones (Hezaveh et al., 2013b; Spilker et al., 2016), several sources are found to have complex structure. Consequently, magnifications obtained with pixellated source models vary by up to a factor of a few compared to results based on parametric models. These findings highlight the limitations of parametric source models in inferring source properties even from low-resolution data.

Finally, we analyse the CO (2-1) and mm-wave continuum emission in a low-redshift strongly lensed quasar RXJ1131-1231. Based on the flux-ratio anomaly observed in the ALMA continuum, we find the continuum emission to be dominated by a synchrotron radiation associated with a very compact source (∼0.01 pc). The CO (2-1) emission reveals a large disk of gas with a complex, spiral-like structure. This initial study illustrates the power of strong lensing to leverage the high-resolution of ALMA for detailed studies of distant galaxies on sub-kiloparsec scales.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The star-formation activity of the Universe reached its peak between redshifts 2 to 5, approximately 1 - 3.5 billion years after the Big Bang. This epoch is characterised by a population of dusty star-forming galaxies (DSFGs), forming stars at rates hundreds of times higher than the Milky Way. Due to their large dust reservoirs, these galaxies are often completely obscured at optical wavelengths (Shapley, 2011), and the bulk of their luminosity is shifted to the far-infrared (far-IR). In spite of their important role in the formation and evolution of massive galaxies, the mechanism that drives such intense star-formation rates is still poorly understood. While several distinct models have been proposed (Tacconi et al., 2008; Dekel et al., 2009; Davé et al., 2010; Engel et al., 2010; Narayanan et al., 2010a), observational evidence remains inconclusive (Casey et al., 2014). One of the main obstacles in constraining the nature of DSFGs is the limited resolution of sub-mm/mm-wave instruments: until very recently, the best observations available were only capable of resolving structures on scales of few kpc, often obtained at a large expense of observing time (Carilli et al., 2011; Ivison et al., 2011; Riechers et al., 2011; Tacconi et al., 2013; Hodge et al., 2015).

Although many of these limitations have been overcome with the advent of large sub-mm/mm-wave interferometers such as ALMA, studying structures at sub-kpc scales still remains prohibitively expensive.

Fortunately, a large population of strongly gravitationally lensed DSFGs has been detected in large-scale surveys at far-IR and mm-wavelengths (Carlstrom et al., 2011; Negrello et al., 2010; Vieira et al., 2013). With a typical magnification of ∼10, strong gravitational lensing provides a unique tool to study DSFGs at higher resolution and with much shorter on-source times. This, however, comes at a price: the effect of strong gravitational lensing has to be deconvolved from the observed surface brightness distribution in order to recover the true unlensed source-plane surface brightness distribution.

To this end, several visibility-fitting gravitational lens modelling techniques have been developed (Bussmann et al., 2013, 2015; Hezaveh et al., 2013b; Spilker et al., 2016), assuming that the source-plane surface brightness distribution can be well described by simple analytic functions such as Gaussian or Sérsic profiles. But given the significant amount of both compact and extended structures exhibited by sub-mm galaxies, and biases in the synthesis imaging process used for imaging of radio-interferometric data, these codes as well as techniques developed for the modelling of optical data (Warren & Dye, 2003; Suyu et al., 2006; Vegetti & Koopmans, 2009; Dye et al., 2015) are no longer sufficient for analysing ALMA observations.

For this reason, this thesis focuses on the development of a new gravitational lens-modelling technique for the analysis of interferometric data with pixellated sources. In particular, this work extends the Bayesian lens-modelling pixellated-source technique of Vegetti & Koopmans (2009) to directly model the visibility function observed by radio-interferometers, and presents its application to several ALMA datasets spanning a wide range in resolution and sensitivity.

In this chapter, we introduce the three main areas that form the backbone of this thesis - strong gravitational lensing, radio-interferometry and dusty star-forming galaxies. Specifically, we first introduce the formalism of strong gravitational lensing, focusing on galaxy-galaxy lensing. We then proceed to describe the basics of radio-interferometry: the visibility function and the procedure used to recover the true sky image from its sparsely sampled Fourier transform. Finally, we summarize the current status of knowledge of DSFGs at redshifts 2 to 5, and describe the two competing scenarios for the origin of their extreme star-formation rates.

1.1 Strong gravitational lensing

1.1.1 Lensing formalism

Gravitational lensing is a deflection of light from a background source due to the space-time curvature caused by massive objects along the line-of-sight (lenses). Gravitational lensing both magnifies (or de-magnifies) the flux and distorts the shape of the lensed source. Depending on the mass of the foreground lens, the magnification factor and the strength of the distortion, gravitational lensing can be broadly divided into three regimes:

- **Strong lensing**: provided the background source and the lens are closely aligned, multiple highly-magnified images of the source are formed. A massive lens in form of a galaxy or a galaxy cluster is required for the image separation to be properly resolved. Depending on the relative position of the lens and the source, lensed images can be significantly stretched in a direction tangential to the lensing mass profile, producing Einstein rings or arcs. Strong lensing has been used to constrain the potential of lens galaxies (e.g. Kochanek, 1991), search for massive dark-matter substructure (e.g. Kochanek & Dalal, 2004; Vegetti et al., 2010), or, in case the background source is variable, to constrain the Hubble constant (e.g. Refsdal, 1964; Blandford & Narayan, 1992; Suyu et al., 2017)

- **Weak lensing**: if the source is not closely aligned with the lens, multiple images are not produced and the magnification and distortion are very weak. While the background galaxies are preferentially stretched in a tangential direction, this effect is much smaller (∼1%, Hoekstra & Jain, 2008) than the intrinsic source ellipticity and is detectable only as a statistical effect inferred from a large sample of background sources. By providing a direct measurement of the surface mass density in a given region, weak lensing is used to constrain mass profiles of e.g. galaxy clusters or large-scale structure (e.g. Bartelmann & Schneider, 1999; Hoekstra & Jain, 2008).

- **Micro lensing**: if the mass of the lens in the strong lensing case is relatively low (such as a single star or a planet), the separation of the multiple images of the source is too small to be resolved (typically on milli-arcseconds scales). However, the overall magnification of the background source is still detectable, mostly as a transient event: a typical scenario involves a planet passing in front of a background star, with a timescale of a few days (e.g. Mao & Paczynski, 1991; Gould & Loeb, 1992; Gaudi, 2010).

In this thesis, we focus on the strong lensing scenario in which multiple images of the background source are formed. The individual images are often highly deformed by stretching in a direction perpendicular to the centre of the lensing mass. Depending on the relative position of the source with respect to the lens, Einstein rings or arcs can be formed. The strong lensing scenario requires a massive deflector such as a galaxy cluster or a single massive galaxy. While lensing by galaxy clusters offers very high magnification (20 and upwards), we focus on the *galaxy-galaxy* lensing scenario. In this case, the single lensing galaxy (with at most a few companions) can be easily approximated by a parametric profile with an external shear, as will be discussed below.

For a general gravitational lensing case, assuming that the impact parameter of the ray from the background source $\boldsymbol{x} = (x_1, x_2)$ is much larger than the Schwarzschild radius of the enclosed mass (i.e. the gravitational field is weak) and the deflection angle is small, the two-dimensional deflection angle $\hat{\boldsymbol{\alpha}} = (\alpha_1, \alpha_2)$ is given as:

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{x}) = \frac{4G}{c^2} \iint \left(\int dz \rho(\boldsymbol{x}', z) \frac{\boldsymbol{x} - \boldsymbol{x}'}{|\boldsymbol{x} - \boldsymbol{x}'|^2} \right) d^2\boldsymbol{x}'. \tag{1.1}$$

where $\rho(\boldsymbol{x}', z)$ is the three-dimensional mass density at position $(\boldsymbol{x}', z)$, z is the line-of-sight coordinate, G is the gravitational constant and c the speed of light (Meylan et al., 2006).

In the *thin lens* approximation (Figure 1.1), we assume the deflector has a line-of-sight extent much smaller than the angular diameter lens-source distance D_{LS} and the angular diameter lens-observer distance D_L (no contribution of other line-of-sight structures is considered in this example). For a galaxy-galaxy lensing, the lens-source and lens-observer angular diameter distances D_{LS} and D_L are typically of the order of $10^3 - 10^4$ Mpc, while the impact parameter $|\boldsymbol{x}|$ is few kpc.

Neglecting the line-of-sight extent of the deflector, the three-dimensional integral in Equation 1.1 is reduced to:

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{x}) = \frac{4G}{c^2} \iint \Sigma(\boldsymbol{x}') \frac{\boldsymbol{x} - \boldsymbol{x}'}{|\boldsymbol{x} - \boldsymbol{x}'|^2} d^2\boldsymbol{x}', \tag{1.2}$$

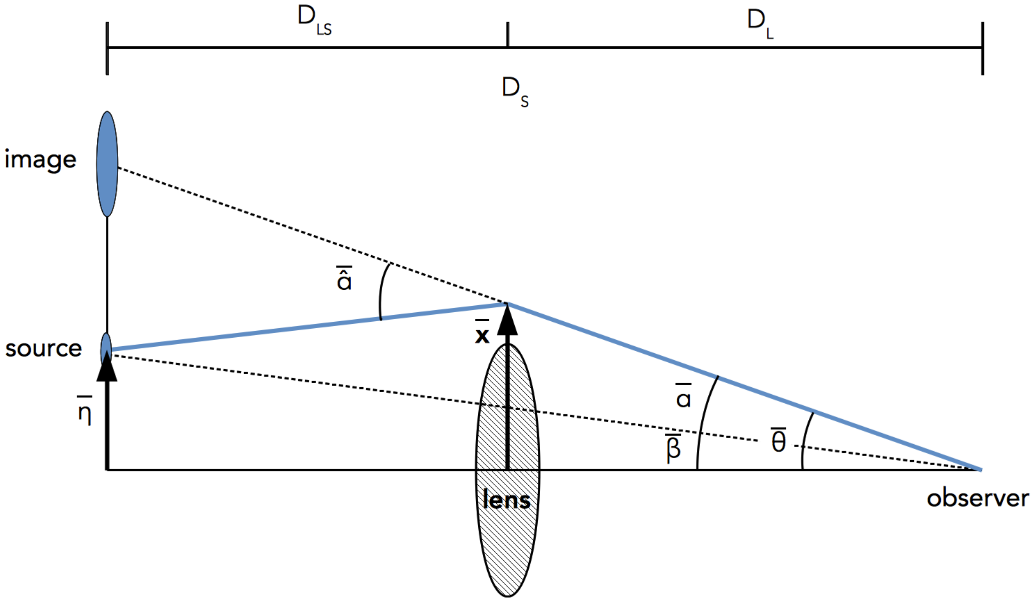

Figure 1.1: Illustration of a strong-lensing scenario in a thin-lens approximation. The ray trajectory from the source to the observer is denoted by the blue line; the lens mass-distribution is indicated by the shaded ellipse.

where we have defined the surface mass density of the deflector $\Sigma(\boldsymbol{x})$ as

$$\Sigma(\boldsymbol{x}) \equiv \int \rho(\boldsymbol{x}, z)\, dz. \tag{1.3}$$

Introducing angular lens- and source-plane coordinates $\boldsymbol{\theta}$, $\boldsymbol{\eta}$ as $\boldsymbol{x} = D_L\boldsymbol{\theta}$, $\boldsymbol{y} = D_s\boldsymbol{\eta}$, the mapping between the source-plane position $\boldsymbol{\beta}$ and the image position $\boldsymbol{\theta}$ is given as:

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \frac{D_{LS}}{D_S}\hat{\boldsymbol{\alpha}}(D_L\boldsymbol{\theta}) \equiv \boldsymbol{\theta}-\boldsymbol{\alpha}(\boldsymbol{\theta}), \tag{1.4}$$

where we have defined the scaled deflection angle as $\boldsymbol{\alpha}(\boldsymbol{\theta}) = \frac{D_{LS}}{D_S}\hat{\boldsymbol{\alpha}}(D_L\boldsymbol{\theta})$ (Figure 1.1).

The Equation 1.2 can be re-written in terms of the *convergence* κ as:

$$\boldsymbol{\alpha}(\boldsymbol{x}) = \frac{1}{\pi}\iint \kappa(\boldsymbol{\theta}')\,\frac{\boldsymbol{\theta}-\boldsymbol{\theta}'}{|\boldsymbol{\theta}-\boldsymbol{\theta}'|^2}\, d^2\boldsymbol{\theta}', \tag{1.5}$$

where

$$\kappa(\boldsymbol{\theta}) \equiv \frac{\Sigma(D_L\boldsymbol{\theta})}{\Sigma_{crit}}, \tag{1.6}$$

and the critical surface mass density Σ_{crit} is defined as:

$$\Sigma_{crit} \equiv \frac{c^2}{4\pi G}\frac{D_S}{D_{LS}D_L}. \tag{1.7}$$

Equation 1.5 can be further re-written as the gradient of the *lensing potential* $\psi(\boldsymbol{\theta})$,

$$\psi(\boldsymbol{\theta}) = \frac{1}{\pi} \iint \kappa(\boldsymbol{\theta}') \ln |\boldsymbol{\theta} - \boldsymbol{\theta}'| \, d^2\boldsymbol{\theta}', \tag{1.8}$$

as

$$\boldsymbol{\alpha}(\boldsymbol{\theta}) = \nabla\psi(\boldsymbol{\theta}). \tag{1.9}$$

The lensing equation (Equation1.4) then becomes:

$$\boldsymbol{\beta} = \boldsymbol{\theta} - \nabla\psi(\boldsymbol{\theta}). \tag{1.10}$$

The image positions therefore depend directly on the *first derivative of the lensing potential* ψ. In addition, the surface mass density can be expressed as:

$$\kappa(\boldsymbol{\theta}) = \frac{1}{2}\nabla^2\psi(\boldsymbol{\theta}). \tag{1.11}$$

Generally, for a source to have multiple images, the condition $\kappa(\boldsymbol{\theta}) > 1$ has to be satisfied for some position $\boldsymbol{\theta}$. This criterion can be used to quantitatively differentiate between strong- and weak-lensing regimes.

Assuming that the angular extent of the source is much smaller than the scale on which the lens properties change, the lensing equation can be linearized. The distortion of the images of the source is then given by the Jacobian matrix

$$\mathbf{A}(\boldsymbol{\theta}) = \frac{\partial\boldsymbol{\beta}}{\partial\boldsymbol{\theta}} = \left(\delta_{ij} - \frac{\partial^2\psi(\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}\right) = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix}, \tag{1.12}$$

where γ_1, γ_2 denote the components of the shear $\gamma = \gamma_1 + \mathrm{i}\gamma_2$,

$$\gamma_1 = \frac{1}{2}\left(\frac{\partial^2\psi}{\partial\theta_1^2} - \frac{\partial^2\psi}{\partial\theta_2^2}\right), \ \gamma_2 = \frac{\partial^2\psi}{\partial\theta_1\partial\theta_2}. \tag{1.13}$$

The magnification μ - the ratio of the the flux observed in the image-plane S_{sky} and that of the unlensed source S_{source} - is then given as

$$\mu = (\det \mathbf{A})^{-1}. \tag{1.14}$$

In other words, given a magnification factor of μ, the flux and the surface area of the lensed image scale as μ (as the surface brightness is conserved), while the angular size of the lensed image scales as $\sqrt{\mu}$. The magnification μ is not necessarily positive - images with a positive parity have the same orientation as the source, while those with negative parity are its mirror-images. The lens-plane positions for which μ diverges (i.e. $\det \mathbf{A} \to 0$) are known as *critical curves.* The corresponding curves in the source-plane are known as *caustics.*

For a more extended source - i.e. one whose angular extent is large compared with changes in the lensing potential ψ - the overall magnification is simply a surface-brightness weighted integral of $\mu(\boldsymbol{\beta})$ over the entire extent of the source:

$$\mu_{tot} = \frac{1}{\iint I_s(\boldsymbol{\beta})\,\mathrm{d}^2\boldsymbol{\beta}} \iint I_s(\boldsymbol{\beta})\mu(\boldsymbol{\beta})\,\mathrm{d}^2\boldsymbol{\beta}. \tag{1.15}$$

While the local magnification $\mu(\boldsymbol{\beta})$ does not depend on the wavelength of the light, the total magnification of the source μ_{tot} does. Namely, due to the weighting by the source-plane surface brightness $I_s(\boldsymbol{\beta})$, μ_{tot} will be generally identical for two different wavelengths λ_1, λ_2 only if the respective surface brightness distributions $I_s(\lambda_1)$, $I_s(\lambda_2)$ are identical. This *differential magnification* effect can have a dramatic impact on inferred source properties. For example, in case of SDP.81 (Chapter 3), differential magnification transforms the intrinsic double-horn profile of the CO (5-4) emission line into an almost Gaussian one as the CO emission crosses the caustic line. In the same system, the source-plane offset between the young stellar population (traced by UV continuum) and dust (traced by mm-wave continuum) causes a 50 percent increase in magnification from μ_{tot}^{UV} ~11 to μ_{tot}^{dust} ~17. Another example is B1938+666, where the radio jet in the source galaxy crosses the caustic, resulting in a very high magnification μ_{tot}^{radio}=173 (Barvainis & Ivison, 2002). On the other hand, the significantly more extended optical emission is magnified by μ_{tot}^{opt} ~13 (Lagattuta et al., 2012). In context of the sub-mm sources investigated in this thesis, differential magnification plays a significant role, as different tracers (dust continuum, various CO transitions, H_2O, HCN, HCO+ molecular lines) trace very different parts of the interstellar medium, and hence are likely to have different magnifications (Serjeant, 2012).

1.2 Interferometry and synthesis imaging

1.2.1 Visibility function

A radio-interferometer is an interconnected network of antennas (dishes or dipoles); the relative distances between individual pairs of antennas are referred to as *baselines*. An interferometer measures the so-called *visibility function* $V(u,v,w)$, which is effectively a discretely-sampled two-dimensional Fourier transform of the sky surface brightness distribution $I(l,\ m)$, such that,

$$V(u,v,w) = \iint A(l,m) I(l,m) e^{-2\pi i[ul+vm+w(\sqrt{1-m^2-l^2}-1)]} \frac{dl\ dm}{\sqrt{1-m^2-l^2}}, \tag{1.16}$$

where l and m are the coordinates in the plane of the sky (using the standard radio-astronomy notation, l pointing towards the west and m towards the north), $A(l,m)$ is the beam profile and u, v and w are the baseline coordinates in the Fourier plane[1] (Taylor et al., 1999).

The w-term - the depth of the array - only becomes important a larger angular distances from the phase-tracking centre. As the array pointing is accurate within

---

[1] A minus sign in the exponent is adopted, following the convention in the Common Astronomy Software Applications (CASA, McMullin et al., 2007).

a few arcseconds for a typical galaxy-galaxy lensing observation, the w-term can be safely discarded. The Equation 1.16 can be then reduced to

$$V(u,v) = \iint A(l,m) I(l,m) e^{-2\pi i(ul+vm)} dl\ dm. \tag{1.17}$$

Unfortunately, a straightforward inversion of Equation 1.17 is prevented by the incomplete sampling of the uv-plane - we can only measure $V(u,v)$ at positions actually sampled by baselines. As a result, one does not obtain the true sky image $I(l,m)$, but the so-called *dirty* image $I_D(l,m)$,

$$I_D(l,m) = A(l,m)^{-1} \iint S(u,v) V(u,v) e^{+2\pi i[ul+vm]} du\, dv, \tag{1.18}$$

where $S(u,v)$ is the sampling function in the uv-plane, which is equal to one for sampled u, v positions and zero otherwise (Figure 1.2). The dirty image can be re-written as the convolution of the true sky brightness distribution $I(l,m)$ with the dirty beam B, which is effectively an inverse Fourier transform of the sampling function,

$$B(l',m') = \iint S(u,v) e^{+2\pi i(ul'+vm')} du\, dv. \tag{1.19}$$

In the image-plane, the effect of the incomplete sampling of the uv-plane is to produce a side-lobe structure that is highly dependent on the baseline configuration of the interferometer and how the visibility data is weighted. As a consequence, the noise in the image-plane is highly correlated and the signal is imprinted with a side-lobe structure and has to be recovered through deconvolution. This is evident in Figure 1.2, which illustrates the different stages of the synthesis imaging process.

1.2.2 Synthesis imaging

The most direct approach to the reconstruction the sky surface brightness distribution is to assume that the dirty image is a collection of point sources convolved with the dirty beam. This is the basic assumption behind CLEAN (Högbom, 1974; Clark, 1980; Schwab, 1984), the most widely used deconvolution algorithm in radio-astronomy (including its numerous extensions). During the CLEANing process, the sky is iteratively decomposed into point-source components and the corresponding visibility-space response is subtracted from the observed *V(u,v)* until a certain threshold is reached. While the original CLEAN algorithm was proven to provide the optimal solution for point-like sources, this approach tends to provide unsatisfactory solutions for highly resolved, diffuse and extended structures, as any extended source will be decomposed into a collection of point sources. This is particularly problematic in the context of strong gravitational lensing, as gravitational arcs are not only extended, but also preferentially stretched in one direction.

A number of improved modifications of CLEAN have been developed such as W-CLEAN (Wakker & Schwarz, 1988) and MS-CLEAN (Cornwell, 2008), mostly with the aim to provide better imaging for extended structures. For example, MS-CLEAN

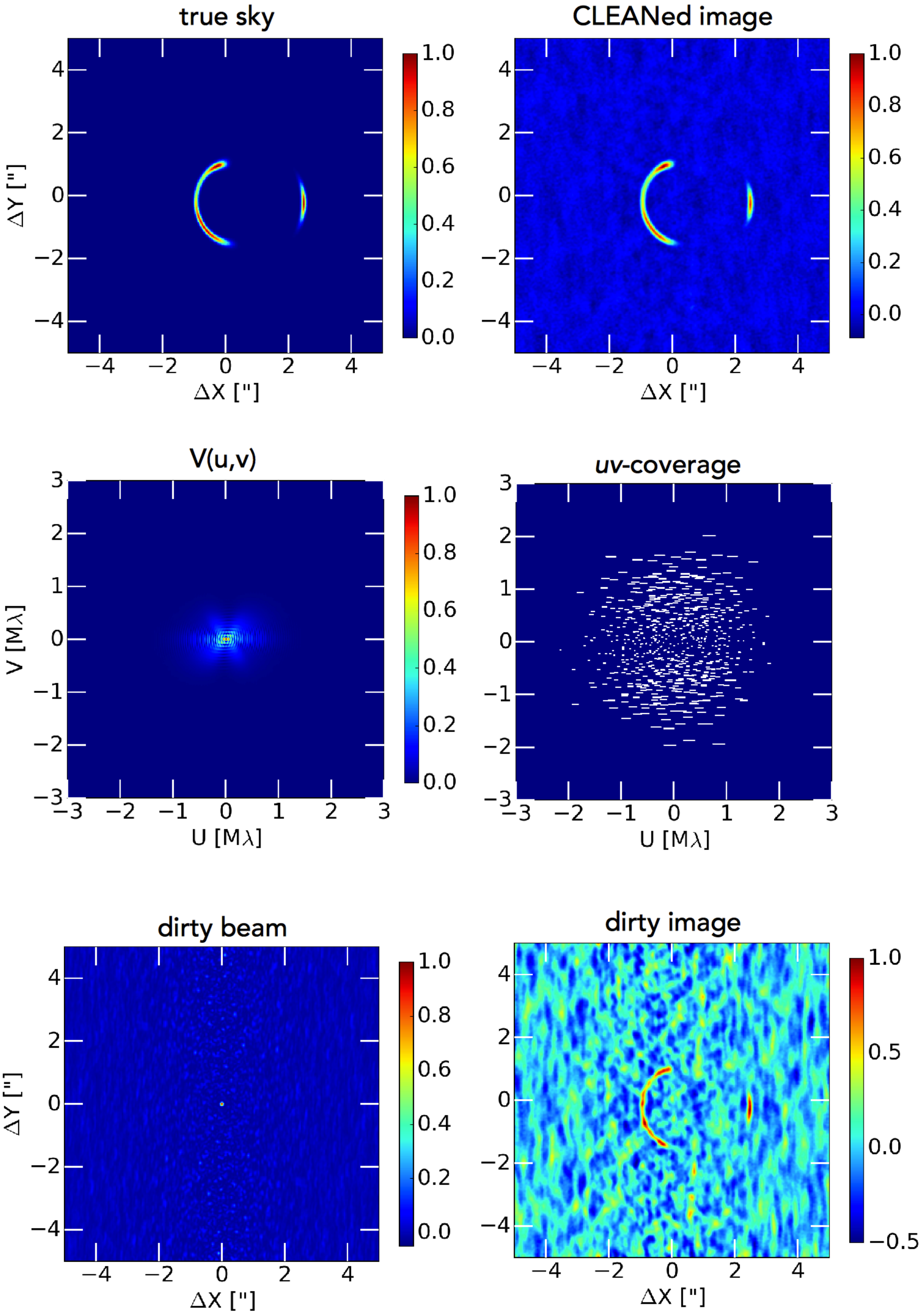

Figure 1.2: Illustration of a the synthesis imaging process. *Upper left*: true sky surface brightness distribution $I^\star(l,m)$. *Upper right*: CLEANed image, natural weighting, obtained using the MS-CLEAN; note that the noise in the de-convolved image is highly correlated. *Middle left*: amplitude of the complex visibility function $\boldsymbol{V}(u,v) = \mathbf{F}\boldsymbol{I}^\star(l,m)$, calculated for every point in the *uv*-plane. *Middle right*: sampling function $\boldsymbol{S}(u,v)$ (*uv*-coverage) corresponding to a 1-hour ALMA observation of a source close to the equator. *Bottom left*: Dirty beam $\mathbf{F}^{-1}\boldsymbol{S}(u,v)$. *Bottom right*: dirty image of the data $\boldsymbol{I}_D(l,m) = \mathbf{F}^{-1}\boldsymbol{S}(u,v)\boldsymbol{V}(u,v)$.

replaces point sources with tapered Gaussians with a different scaling. However, this approach still suffers from sub-optimal recovery of extended structures and unclear propagation of measurement errors, while introducing an additional arbitrariness in choosing the appropriate scaling: "*it is unclear which function the algorithm minimizes, or even if it minimizes anything at al*" (Garsden et al., 2015). More promisingly, several image reconstruction methods based on a visibility-fitting Bayesian approach have been developed, such as RESOLVE by Junklewitz et al. (2016) or the Gibbs sampling-based technique of Sutter et al. (2014).

We will discuss lensing-specific applications of synthesis imaging and techniques developed for lens-modelling of interferometric data in Chapter 1.

1.3 Dusty star-forming galaxies

Dusty-star forming galaxies at high redshift are, as their name implies, galaxies with an intense star formation enshrouded in a large dust reservoir that often completely obscures the rest-frame UV emission (for a comprehensive review, see Shapley, 2011; Casey et al., 2014). In particular, the UV radiation from young stars is absorbed by the surrounding dust, and then re-radiated at longer wavelengths. As a result, DSFGs are primarily detected by their continuum emission in the far-IR to the mm-wavelengths. Typical dust temperatures range between 30 - 50 K (Casey et al., 2014; Bussmann et al., 2013); the dust continuum spectral energy distribution peaks around 100 μm rest-frame wavelength.

The DSFG population is typically divided into several subcategories based on the method with which they have been detected. For example, the sub-mm bright galaxies detected by sub-mm/mm-wave surveys with instruments such as the Submillimeter Common User Bolometric Array (SCUBA, Holland et al., 1999) or the Max Planck Millimeter Bolometer (MAMBO; Kreysa et al., 1999) are often referred to as sub-millimetre galaxies (SMGs). As this thesis is focused on objects selected both in the FIR (SDP.81, Chapter 3) and sub-mm (SPT lenses in Chapter 4), we refer to them collectively as DSFGs.

DSFGs play an important role in the star-formation history of the Universe. Looking back in time, the star formation rate (SFR) density peaks at approximately $\sim 0.2\,\mathrm{M}_{\odot}\mathrm{yr}^{-1}\mathrm{Mpc}^{-3}$ between redshifts $z = 2$ - 4 (1.5 - 3.5 Gyr after the Big Bang), an order of magnitude higher than the present-day value of $\sim 0.01\,\mathrm{M}_{\odot}\,\mathrm{yr}^{-1}\,\mathrm{Mpc}^{-3}$ (e.g. Hopkins, 2004; Hopkins & Beacom, 2006; Bouwens et al., 2010, see Figure 1.3). Up to 50% of this star-forming activity takes place in DSFGs (Wardlow et al., 2011). A typical SFR of DSFGs inferred from far-IR/sub-mm continuum observations is between 100 - 1000 $\mathrm{M}_{\odot}\,\mathrm{yr}^{-1}$. For comparison, the Milky Way's SFR is only 2 $\mathrm{M}_{\odot}\,\mathrm{yr}^{-1}$ (e.g. Chomiuk & Povich, 2011).

1.3.1 Origin of the intense star-formation in DSFGs

The physical mechanism behind the extreme inferred star-formation rates (hundreds to thousands solar masses per year) of DSFGs is still not well understood. Two distinct scenarios have been put forward: short-lived major-merger induced starbursts

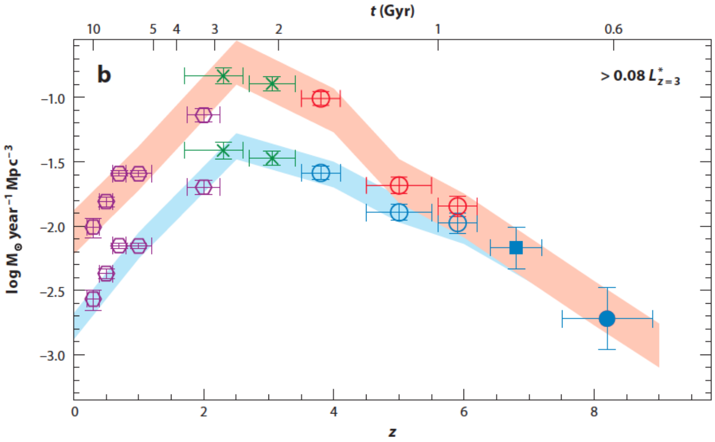

Figure 1.3: Star-formation density of the Universe as a function of redshift, inferred from an un-corrected (*blue*) and extinction-corrected (*red*) UV-luminosity functions. Image credit: Bouwens et al. (2010); Shapley (2011).

(e.g. Narayanan et al., 2010a) versus cold mode accretion (e.g. Dekel et al., 2009; Davé et al., 2010).

In the *major-merger scenario*, the sub-mm bright DSFGs are scaled-up versions of the ultra-luminous infrared galaxies seen in the present-day Universe, for which short-lived intense starbursts are triggered by major mergers of gas-rich galaxies (Sanders & Mirabel, 1996). In this canonical picture, the star formation is triggered by gas compression during the merging phase; as the gas continues to lose angular momentum and develops strong inflows to the center, the star formation rate increases until the remaining gas supply is either exhausted or expelled due to AGN or supernova feedback (e.g. Casey et al., 2014). Using non-cosmological hydrodynamic simulations, it has been demonstrated that starbursts triggered by major mergers reproduce the observed spectral energy distributions of DSFGs (Narayanan et al., 2010a,b). Additionally, the clumpy morphology (on scales of few kpc) observed in many DSFGs seems analogous to merger-induced gas compressions seen in many local ultra-luminous galaxies. Finally, given the inferred SFR and the mass of the available gas reservoir suggest that many SMGs are "maximum starburst systems" (Tacconi et al., 2006) and will exhaust their gas supply on 50-300 Myr time-scales (Tacconi et al., 2008).

However, it is unclear whether the major-merger rate at high redshift is high enough to fully account for the observed population of galaxies with such extreme SFRs, with cosmological simulations suggesting that only a small fraction of SMGs-host halos (M=10^{12} M$_\odot$) undergo a major merger between redshifts z = 2 - 4 (Narayanan et al., 2015). Moreover, while the rate of mergers increases with redshift, the fraction of extreme star-forming systems does not increase accordingly (e.g. Bournaud et al. 2014). Even the efficiency of high-z mergers in increasing the SFR significantly over that expected for isolated systems is still open for debate: if the ISM in isolated high-z galaxies is very clumpy, they might already be forming star at close-to-the-maximum efficiency (Bournaud et al., 2015).

In the competing *cold mode accretion* scenario, most sub-mm galaxies are long-

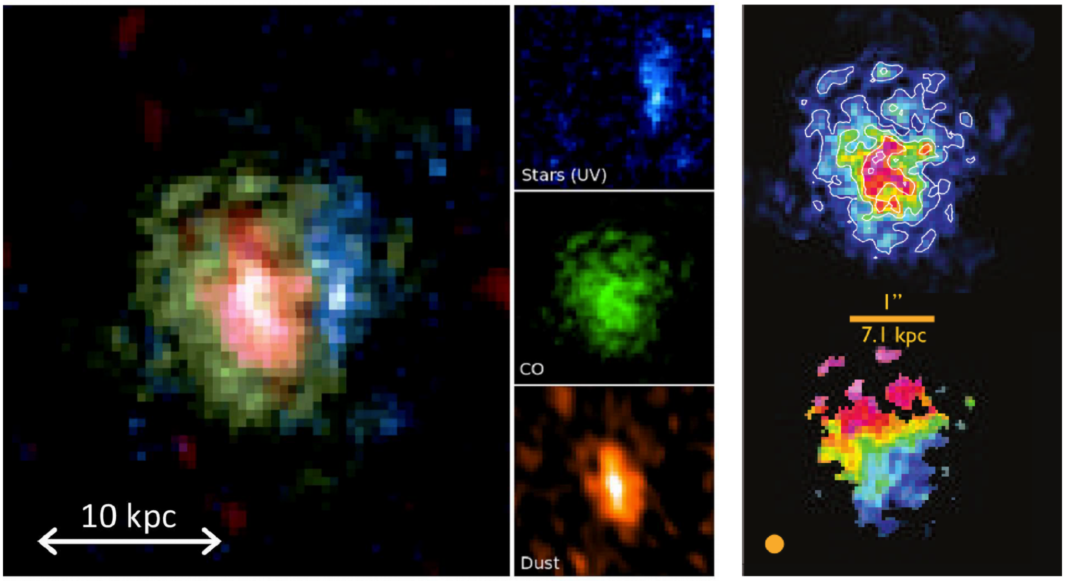

Figure 1.4: Resolving a z=4.05 starburst GN20 at ∼1.5 kpc resolution. *Left*: A composite image of GN20: the stellar component (HST, blue), molecular gas (VLA, green) and dust continuum (PdBI, red) are clearly spatially separated. *Right*: Velocity (*upper*) and velocity dispersion (*lower*) maps of GN20, based on VLA data. Synthesised beam size is indicated by the orange ellipse. Image courtesy: Hodge et al., 2015; Carilli & Walter, 2013

lived large disks with ordered rotation, where star formation is sustained by cold gas accretion and minor mergers (Kereš et al., 2009; Dekel et al., 2009). For example, using hydrodynamical simulations, Dekel et al. (2009) have found that most starburst galaxies are "*stream-fed galaxies*", with only those with extreme SFRs requiring a major merger. This scenario has succeeded in reproducing the observed stellar masses and clustering scales of sub-mm galaxies (Davé et al., 2010). More recently, hydrodynamical cosmological simulations by Narayanan et al. (2015) have shown that intense star formation in DSFGs is dominated by large disks with a duty cycle of ∼0.75 Gyr. While mergers play a more significant role at earlier redshifts, their contribution is negligible during the peak sub-mm brightness at z =2 - 3. Indeed, in their simulations, the sub-mm brightness of the galaxy peak typically more than one Gyr after the last major merger.

Finally, it is possible that DSFGs are not a monolithic population, but rather a mixture of several types of intensely star-forming systems with different physical mechanism driving their intense SFRs and following distinct evolutionary paths (Hayward et al., 2011, 2013). To differentiate between these different origin scenarios, statistical studies of DSFGs as a population are necessary.

1.3.2 Observational studies

Observational studies of redshift 2 to 4 galaxies at far-IR/sub-mm/mm-wavelengths have been for a long time limited by the transmissivity of the atmosphere at sub-mm/mm-wavelengths, and the limited angular resolution of sub-mm observations.

For example, the early surveys with SCUBA had a FWHM beam size of 20 arcsec, resulting in high values of confusion noise. Recently, this has been dramatically improved with the advent of sub-mm/mm-wave interferometric arrays at high altitudes such as ALMA, Plateau de Bure Interferometer (PdBI)/Northern Extended Millimeter Array (NOEMA) and the Submillimeter Array (SMA), offering angular resolutions down to tens of milliarcseconds (long-baseline ALMA configurations). The progressively improving angular resolution is a key element for studying the dynamical properties of DSFGs and resolving the physical conditions of their interstellar medium, chiefly by tracing molecular gas and dust.

So far, molecular gas studies have mostly been focusing on ^{12}CO emission lines due to their high intrinsic brightness and ease of observing over a wide redshift range. In particular, several groups have used $>$0.2 arcsec ($\sim$1.5 kpc at z=3) imaging of CO lines with PdBI/NOEMA, finding substantial evidence for the major-merger scenario. For example, Tacconi et al. (2006) and Tacconi et al. (2008) found "*strong dynamical evidence for major dissipative merging*" in four sub-mm systems. Similarly, Engel et al. (2010) have used sub-arcsecond imaging of CO lines of a larger sample of galaxies ($z = 1.2 - 3.4$) to reveal morphologies with either two distinct components or significant merger-like disturbances in all objects in their sample.

Furthermore, multi-wavelength studies of sub-mm galaxies that trace dust, gas and stellar emission, such as the analysis of SDP.81 presented in Chapter 3, or a larger analysis of 16 galaxies from the ALESS survey by Hodge et al. (2016) - reveal significant (few kpc) offsets between individual components, indicating recent/ongoing merger.

On the other hand, based on a large study of CO (3-2) line in combination with HST imaging of 52 galaxies, Tacconi et al. (2013) found that only less than 20% of z=1-2.5 galaxies exhibit clear merger signatures. A similar fraction of merger-induced starbursts was recently derived by Lofthouse et al. (2017) using visual classification of galaxies from the CANDELS survey. Ordered disk-like rotation was observed in clumpy disks even in very bright sources, such as GN20 - a redshift 4 starburst and the brightest sub-mm galaxy in the GOODS-N field (Carilli et al., 2011; Hodge et al., 2013; Figure 1.4). Even the clumpy morphologies observed in mm-wave dust continuum are not necessarily incompatible with disk-like morphologies. For example, Hodge et al. (2016) demonstrated that due to deconvolution effects, disk-like galaxies observed "*even at moderate SNR*" are prone to be resolved into several clumps.

However, already at low resolutions, resolving dust continuum and obtaining spatially resolved kinematic gas of molecular gas is extremely expensive, even with the full power of ALMA. For example, to resolve structures on 100-pc scales - comparable to the size of star-forming regions in the local Universe - at $\lambda_{obs} \sim 1.0$ mm, the most extended ALMA configuration would have to be used and the on-source time required to obtain a good sensitivity would be of the order of 10 hours.

Strong gravitational lensing, thanks to the increased sensitivity and angular resolution that is provided by its high magnifications, is a uniquely powerful tool for overcoming these limitations and study DSFGs in great detail. We will now turn our attention to DSFGs that undergo strong gravitational lensing.

1.3.3 The H-ATLAS, HerMES and SPT lens samples

Two particular facilities - the *Herschel Space Observatory* and the *South Pole Telescope* (SPT) - and three major surveys: *Herschel Astrophysical Terahertz Large Area Survey* (H-ATLAS, Eales et al., 2010) and *Herschel Multi-Tiered Extragalactic Survey* (HerMES, Oliver et al., 2012) on *Herschel* and the SPT Survey (Carlstrom et al., 2011) have been instrumental in detecting lens candidates in sub-mm wavelengths.

The Herschel Space Observatory was a European Space Agency space mission with a support from NASA: a 3.5-metre telescope observing in far-IR to sub-mm wavelengths between 55 and 672 μm, a regime inaccessible to ground-based observatories. Stationed at the second Lagrangian point, it conducted its science operations between July 2009 and April 2013. The H-ATLAS and HerMES covered about 500 deg^2 and 100 deg^2, respectively, in three bands (250 μm, 350 μm and 500 μm). The South Pole Telescope is a 10-metre single-dish telescope, observing at 1.4, 1.0 and 3.0 mm. The SPT survey covered 2,500 deg^2 at 1.4 and 2.0 mm at 0.7 - 1.6 arcsec resolution. The different wavelength regimes probed by Herschel and SPT (far-IR vs. mm-wave) result in different selection functions - namely, SPT-selected sources tend to be located at a higher redshift, and have a larger far-IR luminosity than Herschel-selected sources (Casey et al., 2014).

While selecting strongly lensed systems at e.g. optical wavelengths is often a very resource-intense task (Bolton et al., 2006; Lin et al., 2009; Marshall et al., 2016), at sub-mm wavelengths, several factors combine to allow an easy detection of strongly lensed systems. These are: the dramatic drop in the number density of the brightest unlensed sub-mm sources, and the negative K-correction which allows the detection of even very distant sub-mm bright DSFGs (Blain, 1996; Negrello et al., 2007).

The *negative K-correction* refers to the increasing brightness of a dust continuum at a fixed observed wavelength with increasing redshift. This effect makes DSFGs more-or-less uniformly detectable over a wide range of redshifts (Shapley, 2011; Casey et al., 2014). For a galaxy with an intrinsic spectral energy distribution f_ν and luminosity L, the measured flux-density S_ν at an observed frequency ν is given by:

$$S_\nu = \frac{L_\nu}{4\pi D_L^2}, \tag{1.20}$$

where D_L is the luminosity distance and L_ν is the luminosity of the source at a given frequency. In the Rayleigh-Jeans approximation, $L_\nu \sim \nu_0^{2+\beta}$, and $\nu = \nu_0/(1+z)$. On the other hand, the luminosity distance evolves with redshift as $D_L \sim (1+z)^2$. Thus Equation 1.20 can be rewritten as:

$$S_\nu(z) \sim \frac{(1+z)^{\beta+2}}{(1+z)^4} = (1+z)^{\beta-2}. \tag{1.21}$$

As the spectral index $\beta = 1.5 - 2.0$ for a typical dust spectral energy distribution (Casey et al., 2014), $S_\nu(z)$ remains roughly constant over a wide range of redshifts, especially at lower frequencies ν[2]. Because of the roughly constant $S_\nu(z)$, sub-

---

[2] Note that the $(1+z)^{\beta-2}$ dependency breaks down as the rest-frame frequency ν_0 approaches the

mm galaxies are almost uniformly detectable up to very high redshifts ($z{\sim}8$). In the context of strong lensing, this translates into a high surface number density of bright sources and a high number of potentially detectable lenses.

The combination of the steep drop in number counts for bright SMGs and the large number of potential background sources means that strong lenses can be identified in sub-mm / far-IR surveys by simply imposing a threshold on the observed flux, at which the number of unlensed sources is supposed to be negligible (Figure 1.5). The potential contaminants are then eliminated via a cross-correlation with catalogues of nearby galaxies and bright active galactic nuclei - AGNs (Vieira et al., 2013; Weiß et al., 2013). Similarly, using a flux-density threshold of 100 mJy at 500 μm and by removing the contaminants, H-ATLAS identified 80 lens candidates (Negrello et al., 2017, Figure 1.5).

After selecting the lens candidates, a higher-resolution imaging and spectroscopic follow-up is necessary to confirm their strongly lensed nature. The imaging is necessary to detect the presence of multiple images of the source or Einstein arcs, while the spectroscopy is used to measure the redshifts of the lens and the source (e.g. Bussmann et al., 2013).

Thanks to the resolution and flux-density boost provided by strong lensing, strongly lensed DSFGs have been used extensively for studying the properties and kinematics of the ISM at kpc-scales (Aravena et al., 2013; Bothwell et al., 2013; Bussmann et al., 2013; Messias et al., 2014; Bussmann et al., 2015; Spilker et al., 2015; Ma et al., 2016; Weiß et al., 2013). Additionally, strong lensing has been used to detect many fainter molecular emission lines, such as water lines (Omont et al., 2011, 2013; Yang et al., 2016) or HCN and HCO+ (Oteo et al., 2017).

This thesis is focused on the study of strongly gravitationally lensed galaxies that have been observed with the Atacama Large Millimeter/submillimeter Array (ALMA). Specifically, we are interested in constraining the physical and kinematic status of their inter-stellar medium at very high angular resolutions.

1.4 Outline

This thesis is structured as follows: In Chapter 2, we present our newly-developed visibility-fitting code for modelling of strongly gravitationally lensed systems and the results of initial tests on simulated data. Specifically, we investigate how well do we recover the input lens models and the source-plane surface brightness reconstruction depending on the angular resolution of the array and sensitivity. We then present multiple lensed objects observed with varying resolution and quality, covering both the *Herschel*- and SPT-selected systems. First, in Chapter 3, we focus on the ALMA 2014 Long Baseline Campaign observations of SDP.81, a redshift-3 starburst. Chapter 4 compares pixellated reconstructions of several SPT-selected lenses observed with the low-resolution ALMA Cycle-0 array with previously published models based on parametric sources. In Chapter 5, we present an analysis of ALMA

---

peak of the SED (200-100 μm, 1.5-3.0 THz), a regime in which the Rayleigh-Jeans approximation is no longer valid.

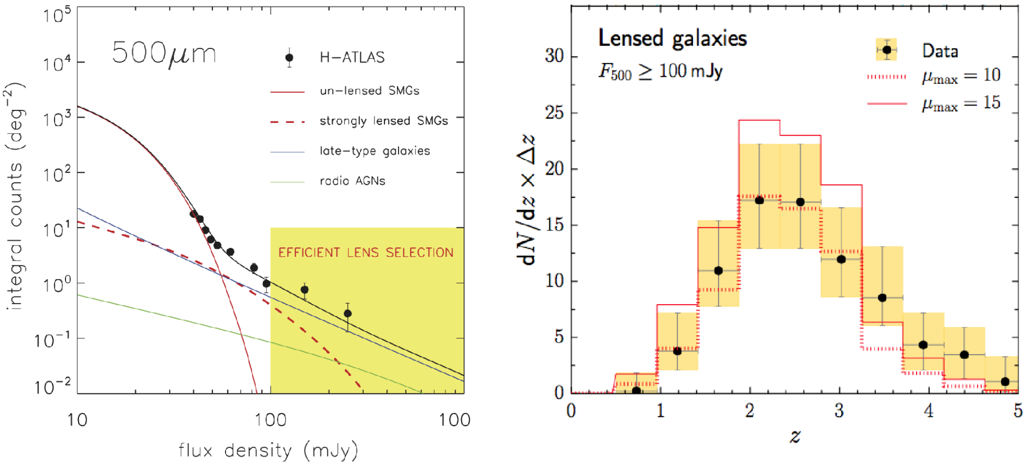

Figure 1.5: *Left*: Selecting sub-mm lenses from 500 μm H-ATLAS source counts. Three different populations contribute to the observed source-count function (black points): sub-mm galaxies (red), nearby late-type galaxies (blue) and AGN-powered radio sources (green). The contribution from the lensed sub-mm galaxies (red dashed line) completely dominates the source count at flux densities >100 mJy. In this regime, the lens selection is extremely efficient. *Right:* Redshift distribution of lensed sources from the full H-ATLAS survey. Image credit: Negrello et al., 2010, 2017

observations of RXJ1131-1231, a low-redshift quasar surrounded by a large disk of gas, revealing the spiral structure of the interstellar medium. Finally, Chapter 6 summarizes the results of this work and provides an overview of future applications and possible improvements of our lens-modelling technique.

Chapter 2

Lens modelling in the visibility-space

Abstract

In this chapter, we introduce a novel method for strong gravitational lens modelling of interferometric data in the visibility space. This method extends the technique of Vegetti & Koopmans (2009) to the interferometric domain and represents a significant improvement over both the lens modelling of CLEANed data and visibility-fitting methods that use parametric source models. We then apply our lens-modelling method to mock ALMA observations of a realistic surface brightness model of a high-redshift starburst galaxy with angular resolution ranging from 1.0 to 0.1 arcsec. We find a significant *bias* in the inferred source properties at coarser resolutions (>0.4 arcsec) as the clumpy structure of the source is not properly resolved; the intrinsic source flux-density is overestimated by more than 50 %. However, at higher angular resolutions, the inferred source flux-density and magnification are in a good agreement with the input model. These results highlight the limitations in inferring source properties from low-resolution data and the importance of pixellated source models for studying gravitationally lensed galaxies.

Note: This chapter is based on Rybak, Vegetti and McKean: "Strong lens modelling of interferometric data with pixellated sources", submitted to MNRAS, August 2016.

2.1 Introduction

The lens-modelling of optical data with source-plane surface brightness distribution reconstructed on a pixellated grid has a long tradition (e.g. Kochanek et al., 1989; Ellithorpe et al., 1996a; Warren & Dye, 2003). Further developments include the introduction of adaptive source-plane grids (Dye & Warren, 2005), model comparison in terms of Bayesian evidence (Suyu et al., 2006) and the inclusion of pixellated perturbations to the lensing potential (Koopmans, 2005; Vegetti & Koopmans, 2009). Nevertheless, in order to take advantage of the combined power of strong gravitational lensing and high-resolution interferometric observations, gravitational lens modelling codes currently tailored for imaging data need to be extended to model the interferometric data in the visibility space.

This idea was first put into practice as a lensing-specific modification of Clean – LensClean, introduced by Kochanek & Narayan (1992) and Ellithorpe et al. (1996b) and improved and applied to VLA and MERLIN data by Wucknitz (2004). In LensClean, the dirty image is used to identify individual source-plane components and, given a particular lens mass model, find their image-plane counterparts; these sources are then subtracted from the visibility function. This method still partially relies on the inverse image reconstruction, which introduces some arbitrariness into the gravitational lens modelling.

More recently, visibility-fitting lens modelling techniques have been developed by Bussmann et al. (2012, 2015) and Hezaveh et al. (2013a). The former makes use of a Markov Chain Monte Carlo approach to generate a large number of parametric lens and source configurations, which are individually mapped into the visibility space. The best solution is then identified based on the χ^2 calculated in the visibility space. Hezaveh et al. (2013b) introduced a visibility-fitting technique with parametric lens and source models, including internal self calibration, and applied it to several ALMA-observed lenses from the SPT sample. Although these techniques represent a significant improvement over lens modelling using synthesis imaging data, they remain limited by the choice of a parametric source model in the form of a Sérsic or a Gaussian profile. Finally, in their latest work, Hezaveh et al. (2016) have extended the technique of Hezaveh et al. (2013b) to pixellated sources.

2.2 Direct Fourier Transform formalism

The method presented in this thesis is an extension of the gravitational lens modelling technique developed by Vegetti & Koopmans (2009, henceforth VK09) to the interferometric domain. In this paper, we focus on the modelling of smooth mass distributions and source reconstruction. We refer to Koopmans (2005); Vegetti & Koopmans (2009) and Vegetti et al. (2012, 2014) for more details on the gravitational imaging of substructures.

2.2.1 The response operator

Under the thin-lens and small-deflection-angle approximations, the complex visibility function $\boldsymbol{d}$ and the unknown background source surface brightness distribution $\boldsymbol{s}$ can be related to each other using the following set of linear equations,

$$\mathbf{FL}(\psi(\boldsymbol{\eta}, \boldsymbol{x}))\boldsymbol{s} + \boldsymbol{n} = \boldsymbol{d}, \tag{2.1}$$

where $\mathbf{L}$ is the lensing operator that transforms the background source surface brightness distribution $\boldsymbol{s}$ into the image plane (see VK09 for more details), $\mathbf{F}$ is the response function that transforms the image-plane surface brightness distribution $\mathbf{L}\boldsymbol{s}$ into the visibility function, $\psi(\boldsymbol{\eta}, \boldsymbol{x})$ is the lensing potential specified by the parameters $\boldsymbol{\eta}$ and the image-plane position $\boldsymbol{x}$, and $\boldsymbol{n}$ is the noise vector. The individual elements of $\mathbf{F}$, corresponding to the k-th baseline and j-th sky pixel, are defined as,

$$\mathbf{F}_{kj} = A(l_j, m_j) e^{-2\pi i (u_k, v_k, w_k)\cdot(l_j, m_j, \sqrt{1 - l_j^2 - m_j^2} - 1)} \tag{2.2}$$

where u_k, v_k and w_k are the uv-plane coordinates of the k-th baseline specified in the units of wavelength, l_j and m_j are the sky-plane coordinates of the j-th pixel specified in radians, and finally $A(l, m)$ denotes the primary beam response as a function of sky-plane position. For all practical purposes, we assume the array to be coplanar and $w_k \sim 0$ due to the small field-of-view imaged, thus considerably simplifying the Equation 2.2 as:

$$\mathbf{F}_{kj} = A(l_j, m_j) e^{-2\pi i (u_k, v_k)\cdot(l_j, m_j)}. \tag{2.3}$$

As $\mathbf{F}$ computes the visibility function only for the uv-plane locations where it has been measured, it automatically includes information relative to the limited uv-plane coverage. Thus $\mathbf{F}$ effectively encodes the Fourier transform operation, sampling function and the beam profile response. Splitting each $\mathbf{F}_{kj}$ entry into its real and imaginary part yields:

$$\mathrm{Real}(\mathbf{F}_{kj}) = A(l_j, m_j)\cos(2\pi(u_k l_j + v_k m_k)), \tag{2.4}$$

and

$$\mathrm{Im}(\mathbf{F}_{kj}) = A(l_j, m_j)\sin(-2\pi(u_k l_j + v_k m_k)). \tag{2.5}$$

For practical purposes, the first N_{vis} entries of $\mathbf{F}$ correspond to the real-part $\mathrm{Real}(\mathbf{F}_{kj})$, while the remaining N_{vis} rows list their imaginary counterparts $\mathrm{Im}(\mathbf{F}_{kj})$. By construction, $\mathbf{F}$ is a dense, real-valued matrix with dimensions $2 \times N_{vis} N_{sky}$, where N_{vis} is the total number of visibilities (determined by the data) and N_{sky} is the number of sky pixels (defined by the user). Similarly, the noise vector $\boldsymbol{n}$ takes the form of a single column with $2 \times N_{vis}$ rows, where the first half contains the noise information for the real visibilities and the second half for the imaginary ones. In practice, an *image-plane mask* is often used to force regions of the sky from which no emission is expected to 0. This can be straightforwardly incorporated into the $\mathbf{F}$ operator, as the columns corresponding to the masked image-plane pixels can be simply set to 0. The size of the problem is then reduced proportionally to the fraction of pixels masked.

As in VK09, the source surface brightness distribution is reconstructed on an irregular adaptive grid defined by a Delaunay tessellation. Namely, the source grid is constructed by projecting the centre of every n-th pixel from the regular sky grid onto the source plane. These points then form the vertices of a triangular source-plane grid. An adaptive grid constructed in this way provides a high pixel density in the high-magnification regions close to the caustics, whereas areas further away from the caustics are sampled more sparsely. The surface brightness of each source-plane pixel is then determined as an interpolation between the values at the vertices. While in the case of optical data, the lens-plane grid is determined by the pixel size of the data, fitting the visibility function leaves both the sky-plane and the source-plane grid to be constructed, at least in principle, arbitrarily. Therefore, special care has to be taken when constructing these two grids to avoid any bias from improper sampling. Moreover, if not all the sky-plane pixels are cast back into the source plane, the size of the *effective pixels* (sky-plane regions corresponding to source-plane pixels) is increased by the casting factor n.

Consequently, in order to cover all resolved structures and avoid discretisation problems during the Fourier transform, the casting factor should be chosen in such a way that the largest effective pixels provide at least Nyquist sampling of the plane-of-the sky for the Fourier transformation. The number of sky-plane pixels cast back into the source plane is therefore chosen in such a way that the effective subsampling is about three or greater.

2.2.2 Solving the linear equation system

As both $\boldsymbol{\eta}$ and $\boldsymbol{s}$ are unknown, Equation 2.1 cannot be simply inverted. This would lead to an indefinite number of possible solutions, all equally minimising the χ^2, as well as to fitting of the noise. Instead, a penalty function, in the form of a regularised set of linear equations has to be defined. In terms of Bayesian statistics, the regularization corresponds to a prior on the level of smoothness of the source surface brightness distribution (see Suyu et al., 2006, VK09 for a more detailed discussion). Assuming that the noise on the real and the imaginary visibilities is Gaussian, it can be shown that the most probable source given the data and a lens mass model is found by minimizing the following penalty function (Warren & Dye, 2003; Suyu et al., 2006, VK09):

$$P(\boldsymbol{s}\,|\,\psi(\boldsymbol{\eta},\boldsymbol{x}),\lambda_s,\boldsymbol{d}) = \chi^2 + \lambda_s(\mathbf{R}\boldsymbol{s})^\top(\mathbf{R}\boldsymbol{s}) \tag{2.6}$$

with

$$\chi^2 = (\mathbf{FL}\boldsymbol{s}-\boldsymbol{d})^\top\mathbf{C}_d^{-1}(\mathbf{FL}\boldsymbol{s}-\boldsymbol{d}) \tag{2.7}$$

where λ_s and $\mathbf{R}$ are the regularization constant and regularization matrix for the source surface brightness distribution and $\mathbf{C}_d^{-1}$ is the covariance matrix of the visibility data. Under the assumptions that the noise is uncorrelated between the individual baselines, $\mathbf{C}_d^{-1}$ is reduced to a diagonal matrix (Wucknitz, 2002). In terms of Bayesian statistics, minimizing the penalty function is equivalent to maximizing the posterior probability density of the source given a form and a level λ_s of regularization and given a set of lens parameters $\boldsymbol{\eta}$. To minimize the Equation 2.6, we need to find a source solution $\boldsymbol{s}$ for which $P(\boldsymbol{s} \,|\, \psi(\boldsymbol{\eta}, \boldsymbol{x}), \lambda_s, \boldsymbol{d})$ is minimized, i.e.:

$$\frac{\partial}{\partial \boldsymbol{s}} P(\boldsymbol{s} \,|\, \psi(\boldsymbol{\eta}, \boldsymbol{x}), \lambda_s, \boldsymbol{d}) = \frac{\partial}{\partial \boldsymbol{s}} \left(\chi^2 + \lambda_s (\mathbf{R}\boldsymbol{s})^\top (\mathbf{R}\boldsymbol{s}) \right) = 0. \tag{2.8}$$

This leads to the following set of equations for $\boldsymbol{s}$:

$$((\mathbf{FL})^\top \mathbf{C}_d^{-1} \mathbf{FL} + \lambda_s \mathbf{R}^\top \mathbf{R}) \boldsymbol{s} = (\mathbf{FL})^\top \mathbf{C}_d^{-1} \boldsymbol{d}. \tag{2.9}$$

The solution to the Equation 2.9 is found using a Cholesky decomposition. Note that a Cholesky decomposition does not automatically force the solution for the source to be strictly non-negative. This can introduce a spurious negative signal in the image plane. However, in practice, the signal is dominated by a bright lensed source while the blank sky is covered with a Gaussian-like noise centred around zero, provided that the signal-to-noise ratio is sufficiently good.

2.2.3 Finding the most probable mass model

Given a choice of the lens parameters $\boldsymbol{\eta}$ and level of regularization λ_s, the most probable source surface brightness distribution can be derived via Equation 2.9. However, as the lens parameters $\boldsymbol{\eta}$ and source regularization λ_s are free parameters of the model, they also have to be inferred from the data; this is done using a three-step optimization scheme as follows.

First, starting with an initial guess for the lens parameters $\boldsymbol{\eta}$ and keeping the source regularization λ_s fixed to a relatively large value, we minimize the penalty function $P(\boldsymbol{s}, \boldsymbol{\eta} | \mathbf{R}, \boldsymbol{d})$ for $\boldsymbol{\eta}$ using a Downhill-Simplex with Simulated Annealing scheme (Press et al., 1992) and solving Equation 2.9 for $\boldsymbol{s}$. In this first step, we typically force the source to be very smooth by choosing a large value of λ_s. In the second step, we optimize for λ_s and $\boldsymbol{s}$ while keeping the lens mass model parameters fixed to the values $\boldsymbol{\eta}$ derived at the previous step. Finally, adopting the best-fit value of λ_s, we re-optimize for $\boldsymbol{\eta}$ and $\boldsymbol{s}$.

Considering the evidence assigned to each model, following the derivations presented in Suyu et al. (2006) and VK09, the task of maximizing the posterior probability of the parameters

$$P(\boldsymbol{\eta}, \lambda_s | \boldsymbol{d}, \mathbf{R}) = \frac{P(\boldsymbol{d} | \lambda_s, \boldsymbol{\eta}, \mathbf{R}) P(\lambda_s, \boldsymbol{\eta})}{P(\boldsymbol{d} | \mathbf{R})} \tag{2.10}$$

can be reduced, assuming a prior $P(\lambda_s, \boldsymbol{\eta})$ that is flat in $\boldsymbol{\eta}$ and $\ln \lambda_s$, to simply maximizing the *likelihood* term $\mathcal{L}$:

$$\mathcal{L} = P(\boldsymbol{d} | \lambda, \boldsymbol{\eta}, \mathbf{R}) = \frac{Z_{model}(\lambda_s, \boldsymbol{\eta})}{Z_{data} Z_{reg}(\lambda_s)}, \tag{2.11}$$

where Z_{data} and Z_{reg} are the normalization factors for the data and regularization probability distribution functions, respectively:

$$Z_{data} = (2\pi)^{N_{vis}} (\det \mathbf{C}_d)^{1/2}$$

and,

$$Z_{reg}(\lambda_s) = (\frac{2\pi}{\lambda_s})^{N_s/2}(\det \mathbf{A}_{reg})^{-1/2}. \tag{2.12}$$

Here, N_{vis} is the number of visibilities, N_s is the number of source pixels, $\boldsymbol{s}_{MP}$ is the source solution maximizing the posterior and E_{reg}, $\mathbf{A}_{reg}$ are defined as:

$$E_{reg}(\boldsymbol{s}) = \frac{1}{2}(\mathbf{R}\boldsymbol{s})^\top(\mathbf{R}\boldsymbol{s}), \tag{2.13}$$

$$\mathbf{A}_{reg} = \frac{\partial^2}{\partial \boldsymbol{s}^2}E_{reg}(\boldsymbol{s}) = \mathbf{R}^\top\mathbf{R}. \tag{2.14}$$

The normalization for the model probability distribution Z_{model} takes the following form:

$$Z_{model} = (2\pi)^{N_s/2}(\det \mathbf{A})^{-1/2}\exp(-E(\boldsymbol{s}_{MP})), \tag{2.15}$$

with

$$E(\boldsymbol{s}) = \frac{1}{2}\chi^2(\boldsymbol{s}) + \lambda_s E_{reg}(\boldsymbol{s}), \tag{2.16}$$

and

$$\mathbf{A} = \frac{\partial^2}{\partial \boldsymbol{s}^2}E(\boldsymbol{s}_{MP}) = (\mathbf{FL})^\top\mathbf{C}_d^{-1}(\mathbf{FL}) + \lambda_s\mathbf{R}^\top\mathbf{R}. \tag{2.17}$$

Taking the logarithm of the likelihood term (Equation 2.11), we obtain:

$$\log\mathcal{L} = -E(\boldsymbol{s}_{MP}) - \frac{1}{2}\log(\det \mathbf{A}) + \frac{1}{2}\log(\det \mathbf{A}_{reg}) + \frac{1}{2}\log(\det \mathbf{C}_d^{-1}) - N_{vis}\log(2\pi) + \frac{N_s}{2}\log\lambda_s, \tag{2.18}$$

which gives the evidence for the ranking of individual $\boldsymbol{\eta}$, λ_s models.

2.2.4 Lens mass model

The projected mass density of the gravitational lens galaxy is described by an elliptical power-law distribution, with a power-law of γ and flattening q, plus an external shear component with strength Γ and position angle Γ_θ. The dimensionless surface density κ and Einstein radius R_{ein} are defined, respectively, as:

$$\kappa(x,y) = \frac{\kappa_0(2-\frac{\gamma}{2})\,q^{\gamma-3/2}}{2(q^2(x^2+r_c^2)+y^2)^{(\gamma-1)/2}} \tag{2.19}$$

and

$$R_{ein} = \left(\frac{\kappa_0(2-\gamma/2)\,q^{(\gamma-2)/2}}{3-\gamma}\right)^{1/(\gamma-1)}, \tag{2.20}$$

where κ_0 is the mass density normalization and r_c the core radius (assumed to be negligibly small). For $\gamma = 2$ these expressions reduce to the SIE relations introduced by Kormann et al. (1994). Altogether, the lens mass model has eight free parameters. Corresponding deflection angles are calculated following Barkana (1998).

2.2.5 Calculating the error on the source

While parametric sources described by an analytic profile have a well-defined error, related to those on the relative parameters, the error on pixellated sources varies for each source-plane pixel and is correlated among pixels. In particular, the uncertainty on the most probable source is characterised by the covariance matrix $\mathbf{C}_s$, which can be expressed as follows

$$\mathbf{C}_s = \left\langle (\boldsymbol{s}_{true} - \boldsymbol{s}_{MP})(\boldsymbol{s}_{true} - \boldsymbol{s}_{MP})^{\top} \right\rangle, \tag{2.21}$$

where the right-hand side is averaged over a large number of the $\boldsymbol{s}_{MP}$ realizations. The data vector $\boldsymbol{d}$ is directly related to the true source $\boldsymbol{s}_{true}$ via

$$\boldsymbol{d} = \mathbf{FL}\boldsymbol{s}_{true} + \boldsymbol{n}, \tag{2.22}$$

while the most probable source given the data is a solution to Equation (2.21), and is an average over several noise realizations while keeping $\boldsymbol{s}_{true}$ fixed,

$$\boldsymbol{s}_{MP} = \left[(\mathbf{FL})^{\top}\mathbf{C}_d^{-1}\mathbf{FL} + \lambda_s \mathbf{R}^{\top}\mathbf{R}\right]^{-1}(\mathbf{FL})^{\top}\mathbf{C}_d^{-1}\mathbf{FL}\boldsymbol{s}_{true}. \tag{2.23}$$

Substituting these into Equation 2.21, and following the derivation of Suyu et al. (2006), we obtain:

$$\mathbf{C}_s = \left[(\mathbf{FL})^{\top}\mathbf{C}_d^{-1}\mathbf{FL} + \lambda_s \mathbf{R}^{\top}\mathbf{R}\right]^{-1}. \tag{2.24}$$

While in reality we do not know the matrix $\mathbf{C}_s$ directly, but we know its inverse $\mathbf{C}_s^{-1}$. As matrix inversion can be computationally very expensive, we use the following setup: given a Cholesky decomposition of the source covariance matrix $\mathbf{C}_s = \mathbf{X}\mathbf{X}^{\top}$, the lower triangular matrix $\mathbf{X}$ can be used to generate a correlated vector $\boldsymbol{s}$ with covariance $\mathbf{C}_s$ and mean $\boldsymbol{s}_{MP}$ from a set of uncorrelated normally distributed variables $\boldsymbol{x}$ with mean of zero and standard deviation of one. A random source realization from a distribution given by the most probable source $\boldsymbol{s}_{MP}$ and $\mathbf{C}_s$ can be obtained as

$$\boldsymbol{s} = \boldsymbol{s}_{MP} + \mathbf{X}\boldsymbol{x}; \tag{2.25}$$

which can be further re-arranged as

$$(\mathbf{X}^{-1})^{\top}(\boldsymbol{s} - \boldsymbol{s}_{MP}) = \boldsymbol{x}. \tag{2.26}$$

where $\mathbf{X}^{-1}$ is obtained by Cholesky decomposition of $\mathbf{C}_s^{-1}$. By drawing a large number of $\boldsymbol{x}$ realizations and solving the Equation 2.26 for every realization of $\boldsymbol{x}$, we estimate the error on the source by taking the r.m.s. variance of the resulting distribution of $(\boldsymbol{s} - \boldsymbol{s}_{MP})$, thus avoiding the expensive inversion of $\mathbf{C}_s^{-1}$.

2.3 Modelling mock ALMA data

2.3.1 Data preparation

All of the simulated datasets that are analysed here were created using the `simobserve` task within the CASA software package as this allows for a set of realistic array configurations, as well as frequency set-ups and observing conditions to be generated.

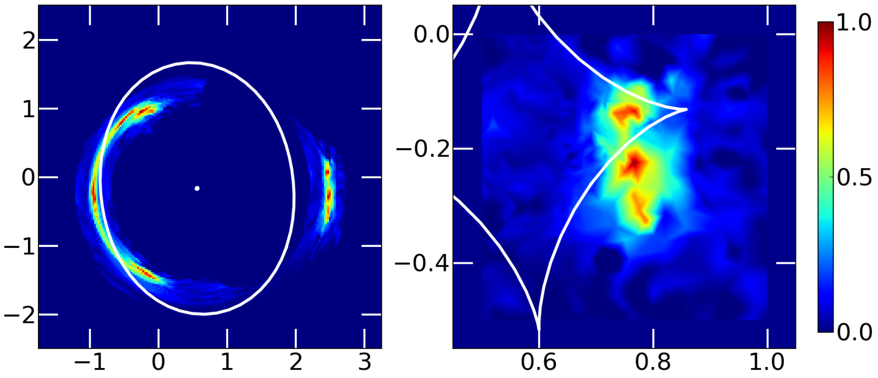

Figure 2.1: Sky and source-plane surface brightness distribution used to generate the mock ALMA datasets. Both images are normalized with respect to the surface brightness maximum, and the spatial axes are given in arcseconds with respect to the phase tracking centre. Critical lines in the sky-plane and caustics in the source plane are indicated in white.

First, we require a representative surface brightness distribution for a high redshift dusty, star-forming galaxy (DSFG). For this, we use the reconstructed source that we obtained for SDP.81 (see Chapter 3), which was observed as part of the ALMA Science Verification programme of the long baseline array (ALMA Partnership et al., 2015b). To summarize, SDP.81 is a DSFG at redshift $z_S = 3.042$ that is gravitationally lensed by a foreground elliptical galaxy at redshift $z_L = 0.299$ (Negrello et al., 2010). The reconstructed source was found to have a combination of extended and compact structure, likely associated with complex star-forming systems (see Figure 2.1). It is this observationally motivated source surface brightness distribution, which provides a realistic model for the structure of a high redshift starburst galaxy on sub-50 pc-scales, that we use for our simulations. Note that the complex structure of this source cannot be reproduced by a few simple Gaussian components.

A simulated image of the gravitationally lensed surface brightness distribution (i.e. in the sky-plane) is generated using the lens mass model parameters corresponding to our best-fit model for SDP.81; the lens model parameters are listed in Table 2.1. The lensed surface brightness distribution is sampled onto a regular grid with a pixel-size of 25 mas, providing sufficient subsampling for the simulated observations; the total sky flux density is set to 110 mJy, as this is representative of the lensed DSFGs found, for example, by the South Pole Telescope gravitational lens survey (Vieira et al., 2013).

Six different ALMA Cycle 3 array configurations, corresponding to CASA antenna configuration files `alma.3.1` through `alma.3.6`, with 36 antennas each, are used for the simulated observations. The synthesised Full Beam-width at Half-Maximum (FWHM) beam-sizes were 0.96×0.86, 0.54×0.49, 0.42×0.36, 0.26×0.24, 0.19×0.16 and 0.092×0.088 arcseconds (Briggs weighting and robust parameter = 0), respectively. For each antenna configuration with the exception of the most extended one (`alma.3.6`), we simulate two observations with a total on-source time of 60 s (10 snapshots of 6 seconds each) and 600 s (20×30 s); these choices correspond

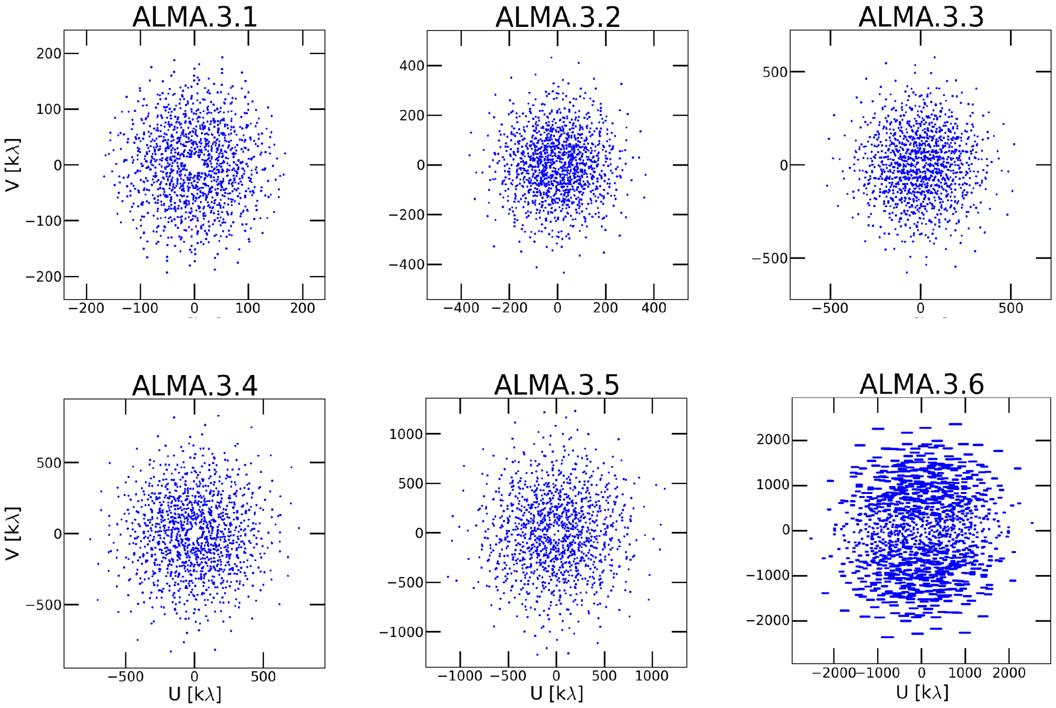

Figure 2.2: The *uv*-coverage for individual array configurations used to create the ALMA mock datasets. The *uv*-coverage for array configurations `alma.3.1` to `alma.3.5` is essentially identical for the 60 and 600 seconds on-source time realizations.

to a typical follow-up observation of gravitational lens candidates. For the `alma.3.6` array configuration, we reproduce a high sensitivity high-resolution observation such as those of the ALMA Long Baseline Campaign by setting the on-source time to 1 hour (120×30 s). The frequency band is centred at 350 GHz, with a bandwidth of 4×2 GHz, mimicking a typical Band 7 continuum observation. The noise model for these observations is generated by the `simobserve` task using the pre-defined noise tables for typical ALMA observations under good observing conditions (i.e. precipitable water vapour of 0.5 mm and an ambient temperature of 269 K). The total number of visibilities to be modelled is 6,300 and 12,600 for the 60 and 600 seconds datasets, respectively, and 75,600 visibilities for the `alma.3.6` dataset. The *uv*-coverage plots of these simulated datasets are shown in Figure 2.2.

2.3.2 Lens modelling

The gravitational lens parameters of the simulated visibility datasets are determined in the following way. First, we extract the real and imaginary part of the visibility function for each measurement, and then we estimate the noise for each baseline by taking the rms variance of the real and imaginary visibilities for each individual baseline. While for these mock data, both the signal and noise could be safely assumed to be constant with time for a given baseline, in case the signal or the noise on a given baseline changes significantly as a function of time, a better noise estimate will be yielded by estimating the rms noise from visibilities from which a model (based on e.g. CLEANed image of the sky) has been subtracted.

The initial values of the lens mass model parameters are estimated as follows: κ_0 is estimated as the image distance measured from CLEANed images; ellipticity of the lens q, lens orientation angle θ and lens centre position x, y estimates are based on optical imaging of the lens galaxy (taken from the HST imaging as presented in Dye et al., 2014). The mass density slope γ is assumed to follow a singular isothermal ellipsoid profile, while the initial shear strength Γ is set to zero. The uncertainty on the lens mass model parameters is estimated from multiple reconstructions with different initial parameter values and forms of regularization.

As discussed in Section 2.1, the sky grid for the lens modelling has to be chosen manually. To ensure a proper sub-sampling in the sky-plane, we set the sky pixel size to 200 mas for the `alma.3.1` dataset, 100 mas for the `alma.3.2` and `alma.3.3` datasets, 50 mas for the `alma.3.4` and `alma.3.5` datasets and 25 mas for the `alma.3.6` dataset. The reconstructed lens and source models do not change when the sub-sampling is increased even further. All the mock datasets are modelled using both gradient and curvature-type regularization; the best form of regularization is determined by comparing the Bayesian evidence. Based on the Bayesian evidence, for lower resolution datasets (`alma.3.1` to `alma.3.3`), the curvature regularization is preferred. On the other hand, starting with the 600-second `alma.3.3` dataset, the gradient form is preferred, as the source structure is progressively more resolved. The maximum *a posteriori* lens model parameters for each simulated dataset are presented in Table 2.1, with corresponding best models shown in Figure 2.3.

Note that the reconstructions were carried out with all baselines being assigned the same weight, $w = 1$. As shown by Junklewitz et al. (2016), the signal reconstruction of a Gaussian signal field effectively introduces an intrinsic *robust* weighting, thus assigning higher weights to baselines with lower noise and denser *uv*-plane coverage. All of the dirty images presented in Figure 2.3 and in the following sections, were created using $w = 1$, with all measurements having the same weight, independent of the noise on a given baseline and the *uv*-coverage. These images are for illustration purposes only, as our algorithm minimizes the penalty function given by Equation 2.6, rather than the dirty image residuals.

2.3.3 Results

Thanks to the extended bright arc providing a lot of information to constrain the lens mass model, the true lens mass model is recovered for all datasets within the 1- to 2-σ confidence intervals, even in the case of the lowest-resolution array (`alma.3.1`). As shown by the dirty-image residuals in Figure 2.3, we almost completely recover the structure of the emission. Specifically, depending on the array configuration and the integration time, between 93 and 100% of the sky-plane flux is recovered. Due to the fixed number of baselines between different array configurations, the *uv*-coverage becomes progressively sparse for more extended arrays. Figure 2.5 shows the total flux density measured from the CLEANed images for each array. For the realistic

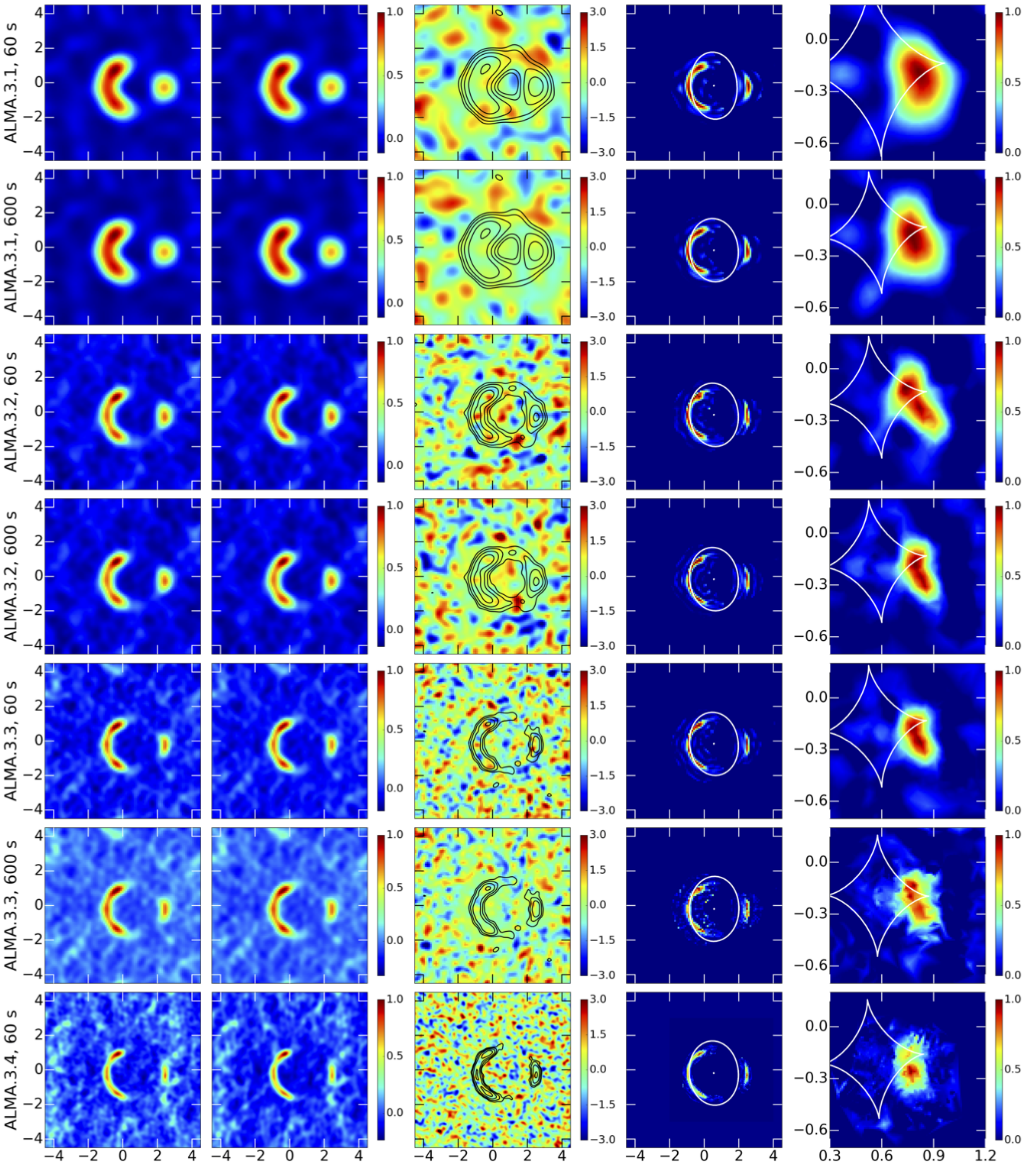

Figure 2.3: Mock data - the best models for each dataset. *Left to right*: The data dirty image, model dirty image, dirty image residuals (data-model) with CLEANed image contours, best sky model and best source model. Critical lines in the sky-plane and caustics in the source plane are indicated in white, as in Figure 2.1. The CLEANed images were produced using Briggs weighting with robust parameter = 0; contours start at $3-\sigma$ level and increase by a factor of 2. The two left-most columns are normalized to the data surface brightness maximum; the residuals are given in units of σ and the two columns on the right are normalized to the peak brightness. Note that the dirty image residuals are shown for illustration only as we minimize the penalty function as given by Equation 2.6, rather than dirty image residuals. The spatial scale is in arcsec with respect to the phase-tracking centre.

input source, none of the arrays considered fully recovers the input sky-plane flux density: the most compact `alma.3.1` and `alma.3.2` arrays detect at most 90% of the total emission, which decreases to around 60% for the `alma.3.5` and `alma.3.6` arrays.

Unlike the lens mass model, the reconstructed source surface brightness distribution is strongly dependent on the observational setup (Figure 2.3). Qualitatively speaking, at lower resolutions, the source is overly smooth and extended; increasing the resolution and sensitivity results in a progressively smaller source, with compact regions becoming more resolved. To provide a first-order quantitative comparison between different source models, we fit each source with an elliptical Gaussian profile; the best-fit model parameters are presented in Table 2.2. At the lowest resolution, both the major axis FWHM, the axis ratio of the source and the intrinsic source flux are systematically overestimated by more than 50% (Figure 2.6); for example, for `alma.3.1` datasets, not a single source reconstruction agrees with the true extent of the source. Looking at Figure 2.6, the 600 seconds on-source time models recover the source size and flux systematically better than 60 s observations, due to a higher SNR; this effect is only marginal compared to that of the array resolution. At resolution better than 0.4 arcsec, the clumpy structures are resolved, and the source size and intrinsic flux are recovered within 1- to 2-σ uncertainty. Finally, at the highest resolution, the limited *uv*-plane coverage at short spacings ($<$150 m) causes a more than 25% of the total flux to be resolved out (Figure 2.5).

2.3.4 Discussion

To determine whether the overestimation of the source size and flux-density is due to the structure of the source or a modelling artifact, we replace the clumpy source from Figure 2.1 by an elliptical Gaussian profile corresponding to the fit to the input source (with parameters listed in the first row of Table 2.1) and create a new set of mock ALMA observations with an identical array configuration and on-source time as those outlined earlier in this section. We then perform the full lens modelling procedure. The discrepancy between the reconstructed and true source size is significantly reduced; for example, at the lowest resolution (`alma.3.1` array), the discrepancy is down to 20% (consistent within 2-σ), compared to 80% for the realistic source. For more extended arrays, we recover the source size within 1-σ uncertainty (Figure 2.7). Therefore, we conclude that the discrepancy in source sizes is due to the clumpy structure not being resolved at resolutions with a FWHM beam size 0.4 arcsec, rather than due to systematics of the modelling technique.

It is important to note that lens mass models can be well constrained by the lens modelling of the deconvolved data in the image plane, as positions and flux ratios of individual images are largely insensitive to any deconvolution issues. A comparison of lens mass models for the same system obtained using image- and *uv*-plane modelling, shows a good agreement between different methods that fit the visibility function (Rybak et al., 2015a,b), CLEANed images (Dye et al., 2015) or use image-tagging (Wong et al., 2015). However, any artifacts from the CLEANing process will be transferred to the source plane, resulting in unphysical features in

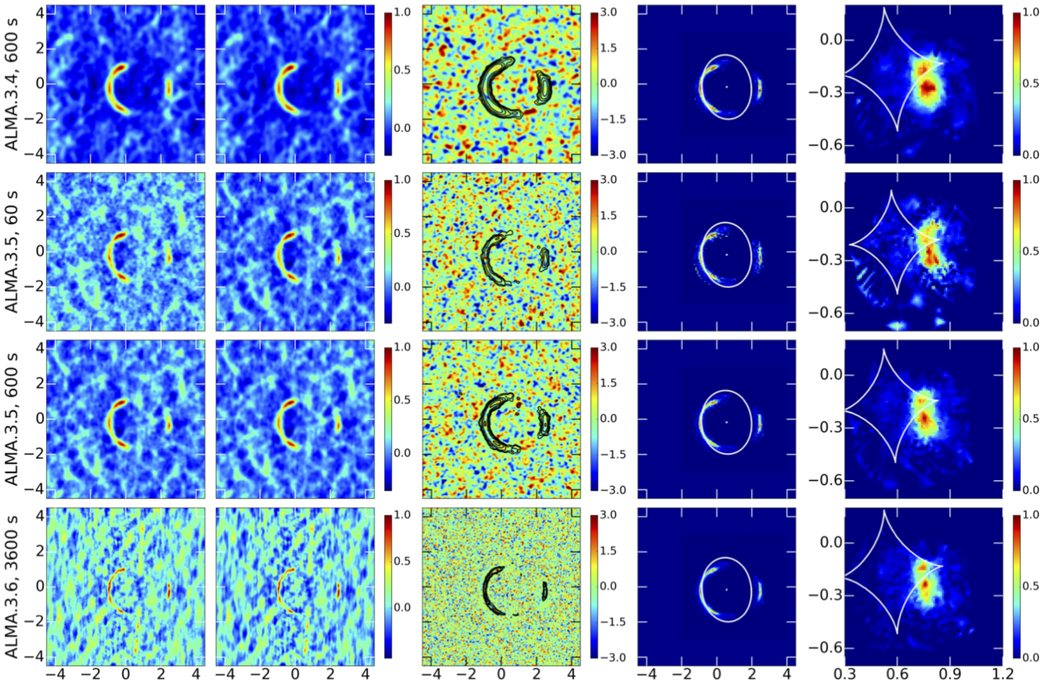

Figure 2.4: continued

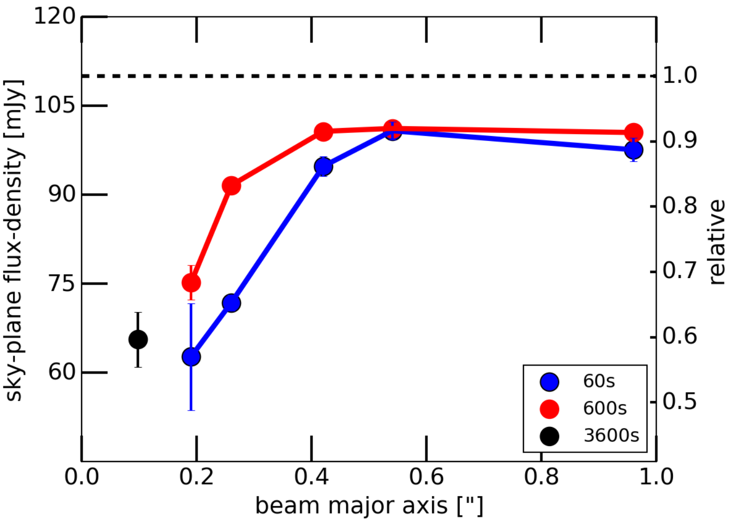

Figure 2.5: Observed sky flux-density as a function of array resolution and integration time, estimated from Cleaned images (natural weighting) and compared to the true sky flux-density (110 mJy). The individual datapoints correspond to the array configurations shown in Figure 2.2; the blue, red and black colours correspond to 60, 600 and 3600 seconds on-source time, respectively. Due to the sparse *uv*-coverage at higher resolutions, up to 40% of the flux is resolved out. Note that even for the most compact configuration, at least ∼10% of the flux is resolved out.

Table 2.1: Mock data - best lens mass model parameters, with the true lens mass model for comparison. For definitions of individual parameters, see Section 2.2.

Array	Time [s]	κ_0 [arcsec]	q	θ [deg]	ΔX [arcsec]	ΔY [arcsec]	Γ	Γ_θ [deg]
True	–	1.606	0.820	8.3	+0.562	-0.165	0.036	3.0
`alma 3.1`	60	1.617±0.030	0.69±0.09	2±3	0.515±0.027	-0.163±0.011	0.031±0.02	10±11
-	600	1.616±0.024	0.76±0.04	0±4	0.541±0.026	-0.171±0.016	0.041±0.017	6±5
`alma 3.2`	60	1.615±0.014	0.79±0.05	-1±5	0.537±0.023	-0.164±0.022	0.038±0.014	12±8
-	600	1.619±0.013	0.78±0.04	7±5	0.544±0.019	-0.181±0.010	0.031±0.010	2±7
`alma 3.3`	60	1.620±0.013	0.82±0.02	7±5	0.555±0.017	-0.195±0.022	0.039±0.010	1±10
-	600	1.629±0.012	0.81±0.03	1±5	0.535±0.019	-0.167±0.016	0.038±0.007	12±7
`alma 3.4`	60	1.618±0.013	0.79±0.03	7±3	0.538±0.021	-0.186±0.023	0.034±0.012	1±7
-	600	1.620±0.013	0.80±0.02	6±3	0.541±0.020	-0.177±0.021	0.029±0.012	4±6
`alma 3.5`	60	1.607±0.013	0.82±0.02	5±2	0.545±0.008	-0.197±0.021	0.044±0.014	-1±5
-	600	1.608±0.011	0.81±0.03	5±3	0.545±0.016	-0.181±0.015	0.042±0.013	5±6
`alma 3.6`	3600	1.605±0.003	0.82±0.02	7±1	0.551±0.009	-0.171±0.005	0.0367±0.002	6±1.7

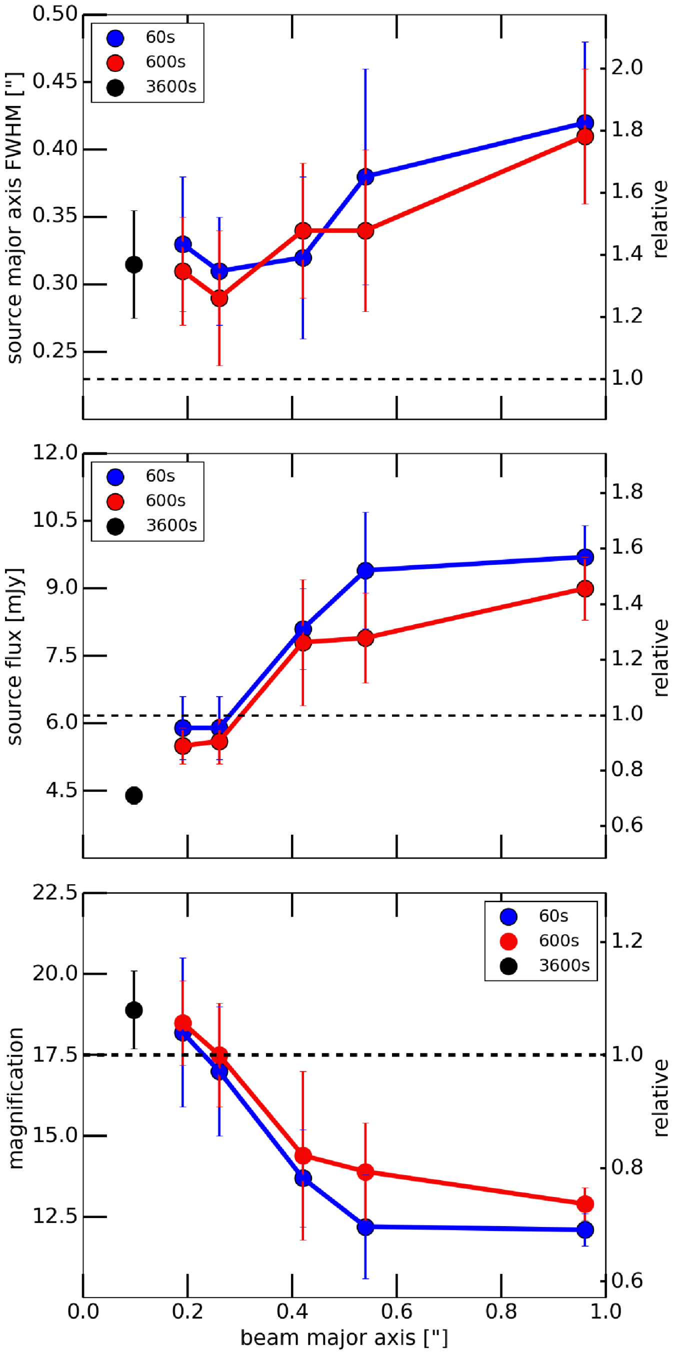

Figure 2.6: *Upper*: The reconstructed source size as a function of array resolution and on-source time compared to the true value (dashed line), based on an elliptical Gaussian fit to the pixellated reconstructed source. The blue, red and black colours correspond to 60, 600 and 3600 seconds on-source time, respectively. *Middle*: The estimated intrinsic source flux density, compared to the true value of 6.18 mJy (dashed line). *Lower*: The estimated magnification compared to the true value $\mu = 17.5$ (dashed line).

Table 2.2: Mock data - best source models, with the true source for comparison. The surface brightness distribution of each source model was fitted with a single Gaussian elliptical profile. The individual columns list the estimated integrated magnification μ, total source-plane flux-density S_ν and the major axis FWHM, the axis ratio f and the orientation angle θ_s (east of north) of the fitted Gaussian.

Array	Time [s]	μ	S_ν [mJy]	FWHM [arcsec]	f	θ_s [deg]
True	–	17.5	6.28	0.23	0.45	7
`alma 3.1`	60	12.1±0.5	9.7±0.7	0.42±0.06	0.70±0.11	10±10
-	600	12.9±0.5	9.0±0.7	0.41±0.05	0.64±0.12	0±10
`alma 3.2`	60	12.2±1.6	9.4±1.3	0.38±0.08	0.55±0.17	0±20
-	600	13.9±1.4	7.9±1.0	0.34±0.06	0.56±0.17	0±20
`alma 3.3`	60	13.7±1.5	8.1±0.9	0.32±0.06	0.57±0.02	0±22
-	600	14.4±2.6	7.8±1.4	0.34±0.05	0.67±0.16	12±13
`alma 3.4`	60	17.0±2.0	5.9±0.7	0.31±0.04	0.56±0.11	0±12
-	600	17.5±1.6	5.6±0.5	0.29±0.05	0.55±0.13	21±20
`alma 3.5`	60	18.2±2.3	5.9±0.7	0.35±0.02	0.50±0.13	0±20
-	600	18.5±1.3	5.5±0.4	0.36±0.02	0.46±0.12	0±19
`alma 3.6`	3600	18.9±1.2	4.4±0.2	0.32±0.04	0.53±0.14	0±18

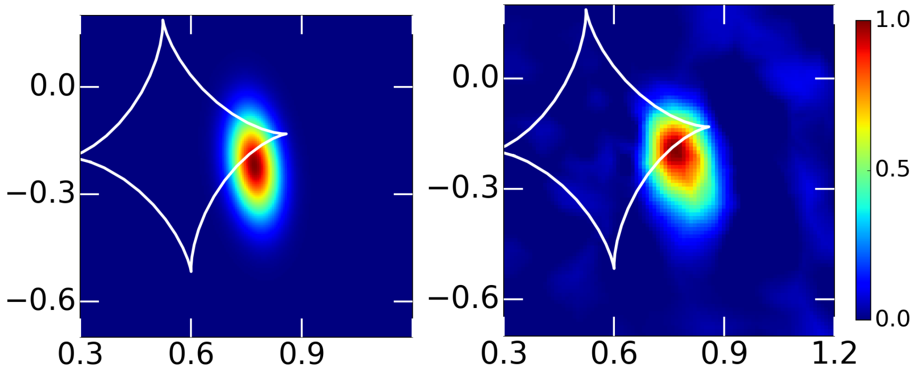

Figure 2.7: *Left:* input Gaussian source-plane surface brightness distribution. *Right:* the resulting reconstructed source for mock data made using the `alma.3.3` array and an on-source time of 60 seconds. The reconstructed source does not exhibit any spurious structure that could be introduced by the modelling procedure, and is well-described by a Gaussian profile: the input and reconstructed Gaussian profiles agree within the 1-σ uncertainties.

the reconstructed source. Therefore, while the modelling of CLEANed images can still deliver reliable lens global mass distributions, this can lead to an incorrect reconstruction of the source surface brightness distribution.

2.4 Conclusions

We have developed a novel Bayesian technique for the modelling of gravitational lensed galaxies with a pixellated source using interferometry data by directly fitting the observed visibility function. By fitting the visibility function rather than the CLEANed data and by using a pixellated source model, this approach represents a considerable improvement over previous visibility-techniques techniques utilizing parametric sources (Bussmann et al., 2013; Hezaveh et al., 2013b) or techniques based on modelling of CLEANed images (Dye et al., 2015). As parametric source models can not capture the range of source structures resolved by observations with state-of-the-art interferometric arrays such as ALMA, our pixellated-source technique offers a unique tool for studying this complex high-redshift sources.

The analysis of mock ALMA Cycle-3 data with 0.1-1.0 arcsec resolution, based on a physically realistic model for a sub-mm galaxy, demonstrated the robustness of the lens modelling procedure, with reconstructed lens mass models being consistent with the input model within 2-σ limits. While source models for more extended configurations reproduced the true source structure with both the central compact region and the extended emission, at low resolutions, we found the source to be overly smooth and extended, leading to a bias in inferred integrated intrinsic flux-density by more than 50%.

The substantial bias in inferred source properties from low-resolution data reveals the limitations in inferring source properties from low-resolution data. Therefore, although dedicated campaigns with the ALMA Cycle-0 and SMA provided a large sample of more than a hundred gravitationally lensed sub-mm galaxies observed at low resolution (Negrello et al., 2010; Vieira et al., 2013), the inferred intrinsic source sizes and flux densities might suffer from a large systematic error and have to be taken with caution.

Chapter 3

SDP.81: a highly resolved redshift 3 starburst

Abstract

In this chapter, we present the an analysis of the gravitational lens system SDP.81 at z=3.042 using ALMA 2014 Long Baseline Campaign data. In particular, we present the reconstruction of the source-plane dust emission at 1.0-mm and 1.3-mm continuum at sub-50-pc resolution, as well as the reconstruction of the CO (5-4), (8-7) and (10-9) emission at sub-100-pc resolution.

We find the dust emission to be concentrated into a ∼2.0×0.7 kpc disk-like structure that contains three compact regions with densities that peak between 120 - 190 $\mathrm{M_\odot\,yr^{-1}\,kpc^{-2}}$, surrounded by extended star formation, with a density of 20 - 30 $\mathrm{M_\odot\,yr^{-1}\,kpc^{-2}}$. There is also a tentative variation in the spectral slope of the different star-forming regions, which is likely due to a change in the dust temperature and/or opacity across the source.

For the CO lines, we find clear evidence for an excitation-dependent structure in the molecular gas distribution; the CO (5-4) is significantly more diffuse and structured than in CO (8-7) or CO (10-9). A kinematic analysis of the velocity fields shows evidence for a star-forming disk with multiple velocity components that is consistent with a on-going/post-coalescence merger. Source reconstructions from ALMA and the Hubble Space Telescope reveal that the stellar component is offset from the molecular gas and dust components. Together with VLA CO (1-0) data, they provide corroborative evidence for a complex ∼2 kpc-scale starburst that is embedded within a larger ∼15 kpc structure.

Note: This chapter is based on Rybak et al., MNRAS 451, 2015 and Rybak et al., MNRAS 453, 2015.

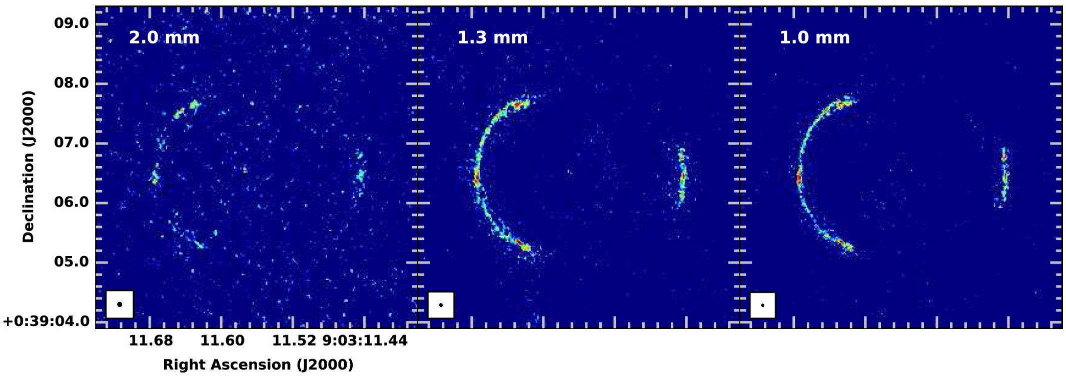

Figure 3.1: ALMA imaging of the SDP.81 continuum in Band 4 (*left*), Band 6 (*centre*) and Band 7 (*right*), using Briggs weighting, robust parameter = 1. The beam size is 56×50, 39×30 and 31×23 mas for Bands 4, 6 and 7, respectively, and is indicated by the black ellipses in the lower left corners of each image. Note the faint emission close to the centre of the lens due to the AGN in the lensing galaxy. Image credit: ALMA Partnership et al., 2015b

3.1 Introduction

SDP.81 (also known as J090311.6+0039006) is a redshift $z_S = 3.042$ sub-mm bright starburst galaxy, lensed by a massive elliptical galaxy located at redshift $z_L = 0.299$. It was discovered by Negrello et al. (2010) as a part of the *Herschel* Astrophysical Terahertz Large Area Survey (H-ATLAS; Eales et al. 2010). The source redshift was determined by Lupu et al. (2012) using multiple CO transitions observed with the 10-metre Caltech Submillimeter Observatory Telescope. Furthermore, Omont et al. (2013) used the IRAM Plateau de Bure Interferometer (PdBI) to detect several CO transitions and a water line $H_2O(2_{02} \rightarrow 1_{11})$.

Prior to the ALMA observations described in this work, SDP.81 was modelled by Bussmann et al. (2013) using 880 μm SMA observations, and Dye et al. (2014), using HST imaging. While the lens model of Bussmann et al. (2013) was limited by the relatively low angular resolution of the SMA array (0.6 arcsec FWHM beam size), they use a visibility-fitting procedure with the lens described by a singular isothermal ellipsoid profile (Chapter 2), with the source approximated by a single Sérsic profile. The inferred magnification is $\mu_{880\mu\mathrm{m}} = 11.1 \pm 1.1$; as the authors note, their model still shows some significant (>3-σ) residuals. On the other hand, Dye et al. (2014) applied a pixellated lens-modelling technique to higher-resolution *Hubble Space Telescope* (HST) imaging to derive an optical magnification of $\mu_{\mathrm{opt}} = 10.6^{+0.6}_{-0.7}$; for a detailed description of this data, see Section 3.3.

3.2 2014 ALMA Long Baseline Campaign

3.2.1 Observations

In autumn 2014, ALMA conducted its 2014 Long Baseline Campaign (LBC). For the first time in ALMA operations, very long baselines of up to 15 km were used, giving a nominal angular resolution of 10 mas. As a part of the ALMA Science Verification programme, the data was made publicly available in February 2015, after standard checks and calibration by the ALMA team. Given the unprecedented quality of the data, and the immediate availability to a wide scientific community, the LBC dataset presented an ideal target for a demonstration of our visibility-fitting lens modelling technique (Chapter 2).

A detailed description of the observational setup and data processing was given in ALMA Partnership et al. (2015a) and ALMA Partnership et al. (2015b), which respectively cover the 2014 Long Baseline Campaign in general and the SDP.81 observations in particular. Here, we provide a brief summary of these observations.

As one of the LBC targets, observations of SDP.81 were taken over several observing blocks during 2014 October and November. These observations used the most extended ALMA configuration up to date: with 31 to 36 of the 12-meter antennas in the array and baseline lengths ranging from 15 m to 15 km. About 10% of the baselines were shorter than 200 m. This gave an array that was sensitive to structures on angular scales between 19 arcsec and 16 mas (at 236 GHz).

The total time on-source for individual observations was as follows: 5.9 hours for Band 4, 4.4 h for Band 6 and 5.6 h for Band 7. The integration time varied between 2 and 6 seconds. Both the XX and YY polarizations were observed.

The data were taken in Bands 4, 6 and 7 and comprised both continuum imaging datasets centred at 2.0, 1.3 and 1.0 mm (156, 236 and 290 GHz), and spectral line datasets centred on the CO (5-4)[1] (rest-frame frequency $f_0 = 576.267$ GHz), CO (8-7) ($f_0 = 921.799$ GHz), CO (10-9) ($f_0 = 1151.985$ GHz) and $H_2O\,(2_{02} \rightarrow 1_{11})$ ($f_0 = 987.927$ GHz) emission lines. Of these, CO (5-4) fell into the Band 4; CO (8-7) and $H_2O(2_{02} \rightarrow 1_{11})$ into Band 6 and CO (10-9) into Band 7. An attempt to observe the continuum emission in Band 3 was abandoned after preliminary tests revealed the source to be too faint in this particular Band (ALMA Partnership et al., 2015b). Figure 3.1 shows the CLEANed images of the full-resolution continuum data as presented in ALMA Partnership et al. (2015b).

In each Band, four spectral windows (SPWs) were used; three of which were configured in the continuum mode with a channel width of 15.6 MHz. The spectral setups for the line observations was as follows: for the CO (5-4) line, a SPW centered on the line with a channel width of 0.976 MHz was used for all executions. In Band 6, a SPW centered on the $H_2O(2_{02} \rightarrow 1_{11})$ line with 0.976 or 1.95 MHz resolutions was used (the spectral resolution varied between individual executions); the second line in this Band, the CO (8-7) line was covered by one of the continuum SPWs. Finally, the CO (10-9) line was observed at a 1.95 MHz spectral resolution in Band 7.

The observing conditions were better than average for all three Bands - the precipitable water vapor values at zenith were between 0.6 - 3.2, 0.5 - 3.1 and 0.3 - 0.7 mm for Bands 4, 6, and 7, respectively.

---

[1] In this work, we refer exclusively to the rotational CO transitions, following the CO ($J_{upper} - J_{lower}$) notation.

3.2.2 Data inspection and selection

From our imaging tests, we found that the recovered structure of the source is highly dependent on the baseline data that are used. Namely, after continuum imaging using only baselines longer than 5 Mλ, only a handful of very compact regions were detected. On the other hand, continuum imaging based on baselines shorter than 0.5 Mλ revealed a faint Einstein ring. Figure 3.2 presents the imaging data for Bands 6 and 7 using a cut in the *uv*-plane at baseline lengths of 0.5 and 2 Mλ. The 0.5 Mλ images show that at mm-wavelengths the observed system comprises four images in a cusp configuration with evidence for a low-surface brightness Einstein ring. The 2 Mλ image shows that the large arc is composed of several structures with varying surface brightness, and although the Einstein ring is effectively resolved out at these scales, there is still evidence of emission from this part of the source. For the visibilities at >1 Mλ (i.e. removing the shortest baselines), the extended arc was mostly resolved out, which demonstrates that most of the emission at mm-wavelengths is extended on scales larger than 0.2 arcsec. Only the most compact components within the individual images were detected with the >2.5 km baselines. We carried out additional tests to determine the scales on which most of the compact structures are resolved out. We found that they are detected up to around 5 Mλ, that is, around 30 mas-scales. Based on these results, we restrict our lens modelling to the visibility data at *uv*-distances $\leq$2 Mλ for the continuum, as there is little or no extended structure detected on baselines longer than this. The CO and H_2O lines suffer from an even lower SNR at long baselines than the continuum, as we will describe later, to obtain a robust reconstruction, only baselines shorter than 1 M$\grave{\lambda}$m were considered for the line datasets.

The visibility data for the final observing block of the Band 6 dataset had a larger rms noise compared to the rest of the observations; consequently, this observing block was removed from the dataset. Also, the absolute flux-density calibration of the first spectral window of the Band 6 dataset was higher by $\sim$ 28% relative to the other SPWs in the same Band, an offset that we corrected for during our analysis. The post-reduction quality of the Band 7 dataset was excellent; no additional corrections were required for this dataset.

A low-level continuum emission, co-spatial with the centre of the lensing galaxy, was detected prominently in the Band 7 continuum data. While reminiscent of a potential central lensed image, comparison of the continuum spectrum of this region with that of the four images showed that this is in fact an AGN hosted by the lensing galaxy (Wong et al., 2015). Due to its low flux-density and a very compact size, we could safely disregard this emission from our analysis.

The insufficient signal-to-noise ratio on the longer baselines of the data did not allow this data to be self-calibrated (ALMA Partnership et al., 2015b). Although we tried to apply self-calibration at least to the baselines within the innermost 2 Mλ, we did not manage to obtain a robust solution for the phase errors. As a result, no additional phase-corrections were applied. While this does not present a major issue for the source reconstruction, as shown by Hezaveh et al. (2016), the uncorrected phase errors can effectively mimic perturbations to the Einstein arc caused by a substructure in the lens potential.

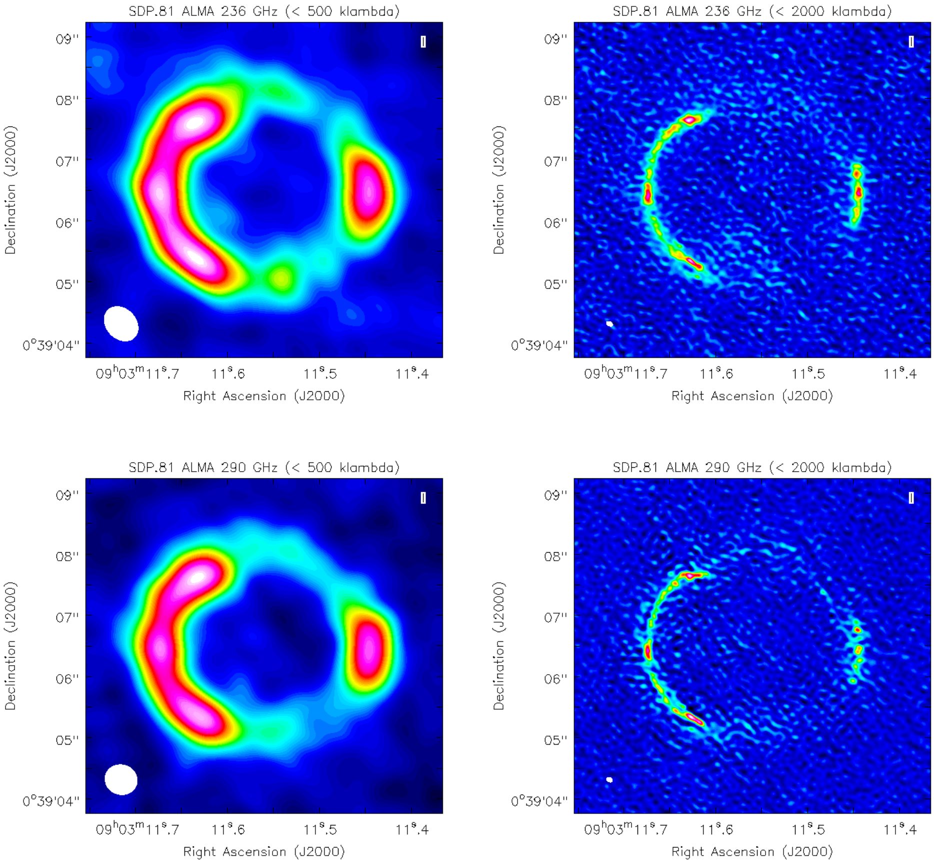

Figure 3.2: ALMA continuum imaging of SDP.81 of Band 6 (*upper*) and Band 7 continuum (*lower*) (1.3 and 1.0 mm, respectively). The images have been produced at two angular resolutions using a cut in the *uv*-data of 0.5 Mλ (*left*) and 2 Mλ (*right*), and weighting was applied (natural and Briggs, robust parameter = 0) to emphasise the structure seen on different scales. The synthesised restoring beam is shown as the white ellipse in the bottom left corner of each map.

3.3 Ancillary observations

3.3.1 Hubble Space Telescope

SDP.81 was observed with the Wide Field Camera 3 aboard the HST on 2011 April 4 (GO: 12194; PI: M. Negrello). The data were taken through the F110W and F160W filters, with exposure times of 712 and 4418 seconds, respectively (for further details, see Negrello et al., 2014). The HST imaging detected a single dominant lensing galaxy, two gravitational arcs from the background source and extended low-surface brightness structures around the system. The data were reprocessed using the standard ASTRODRIZZLE[2] pipeline (using a pixel size of 65 mas) and the subtraction of the lensing galaxy emission was done using GALFIT (Peng et al., 2010). Figure 3.3 shows the overlay of the HST, VLA (as described in the following section) and ALMA image-plane surface brightness distributions. Although the Einstein arcs at HST/ALMA wavelengths are offset from each another, due to their different positions in the source-plane (see below), the separation between the main and counter-arc is the same in both the infrared and mm components, which demonstrates that they are at the same redshift. We modelled the infrared emission using the pixelated gravitational lensing code of Vegetti & Koopmans (2009). Consistent with other studies (Dye et al., 2014, 2015; Wong et al., 2015; Tamura et al., 2015), we found that the rest-frame UV emission has a magnification of $\mu_{tot}^{UV} = 11.3 \pm 0.1$. We discuss the source-plane properties of the rest-frame UV/optical emission, with respect to the dust and gas components in Section 3.7.

3.3.2 Very Large Array

Spectral line imaging of the CO (1-0) emission from SDP.81 was carried out using the Extended Very Large Array (VLA) in the most compact D-configuration on 2010 July 18 (PI: R. Ivison; 3.3×2.3 arcsec beam-size at position angle -10.1 deg). The integration time was $\sim$85 min and 16 VLA antennas that had been upgraded with the new receiver systems were used. Further details about these observations are described by Valtchanov et al. (2011). We obtained these data from the VLA archive and reprocessed them in the standard manner using the CASA package. In Figure 3.3, we show the image-plane emission of the CO (1-0) transition, with respect to the infrared and mm-continuum emission as observed by ALMA, and the UV emission as observed by HST-WFC3. It is clear that the CO (1-0) emission is co-located with the ALMA arcs and the low-surface brightness infrared emission to the south, demonstrating that these extended infrared components are at the same redshift as SDP.81. Unfortunately, the SNR of the VLA dataset is too low to provide a reliable source reconstruction using our visibility lens modelling technique or to determine if the rest-frame UV component to the north of SDP.81 is also associated with this lens system.

---

[2] link: `drizzlepac.stsci.edu`

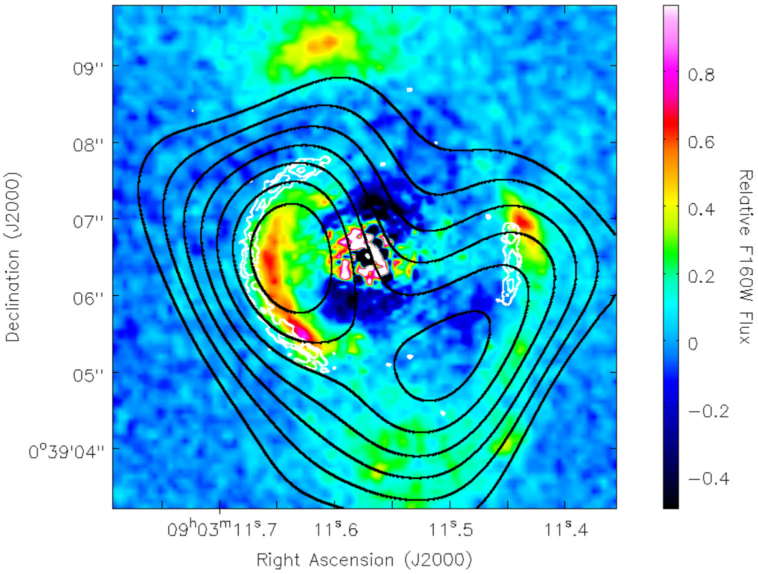

Figure 3.3: The image-plane HST (WFC3 F160W) rest-frame UV/optical emission from SDP.81 (lensing galaxy subtracted) with the ALMA dust continuum (white contours) and VLA CO (1-0) emission (black contours; at 30 to 90% relative to the peak emission, increasing by 10 percentage points).

3.4 Lens modelling

3.4.1 Obtaining the lens model

A preliminary lens model was obtained using the CLEANed data and the lens-modelling code of Vegetti & Koopmans (2009). As in Chapter 2, our lens model is described by an elliptical power-law with a slope γ and flattening q, plus an external shear component with a strength Γ and shear angle Γ_θ. Altogether, this model has eight free parameters (Table 3.1). After obtaining the preliminary model, the final optimization for lens parameters was carried out using the visibility-fitting method on the Band 7 continuum data, as this dataset had the best SNR. The best *a posteriori* lens model parameters are listed in Table 3.1. These are consistent with the gravitational lens modeling of the SMA data by Bussmann et al. (2013) and of the HST data by Dye et al. (2014) at the 2-σ level, with the small differences stemming from different choices of mass parametrization. For example, Bussmann et al. did not include the external shear in their model. In agreement with Dye et al. (2014), we found the lens mass density slope to be consistent with an isothermal profile: $\gamma = 2.00 \pm 0.03$. After finding the best lens model, we kept the lens parameters fixed for all the remaining datasets while we re-optimized for the source surface brightness distribution and its regularization.

Table 3.1: Maximum *a posteriori* lens model parameters

Lens model parameter		Best-model value
κ_0	[arcsec]	1.606±0.005
q		0.82±0.01
θ	[deg]	8.3±0.4
Δx	[arcsec]	≡0.562
Δy	[arcsec]	≡-0.165
γ		2.00±0.03
Γ		0.036±0.004
Γ_θ	[deg]	3.0±0.2

3.4.2 Bands 6 and 7 continuum modelling

In order to minimize the line contamination, channels containing the CO and water lines were removed from the data. To reduce the number of visibilities N_{vis} to be modeled, we averaged the data in time and frequency: namely, we used a time-bin of 20 seconds for averaging in the time domain; in the frequency domain, each SPW was averaged into a single channel. To further limit the number of visibilities used for the gravitational lens modelling, we imposed a *uv*-distance *cut* of 2 Mλ (Figure 3.2); the corresponding synthesized beam-size is 95×71 mas at a position angle of 64.5 deg east of north (at 236 GHz, Briggs weighting, robust parameter = 0). This *uv*-distance cut reduced the date volume to a manageable size of 100,000 to 150,000 visibilities per spectral window. Due to the adaptive grid used for the source reconstruction, the resolution in the source-plane varied depending on the local magnification. For a 2 Mλ *uv*-cut (Figure 3.2), we achieved a typical source-plane resolution of 50 parsecs, a factor of 10 to 20 improvement over previous studies (Hodge et al., 2015).

With the lens model in hand, we reconstructed the source surface brightness map for each of the two spectral window datasets in Bands 6 and 7 separately by optimizing for the source and its regularization parameter λ_s only. This was done partly to provide a consistency check between the source reconstructions for each spectral window, but also to limit the size of the visibility datasets that are fitted. We then improved the SNR of the source-plane reconstruction for each Band by combining reconstructed source surface brightness distributions for both SPWs. We estimated the noise level for each baseline at a given time and frequency by calculating the rms of the visibilities measured within the same observing block (typical length of approximately 30 minutes). In principle, this could lead to a biased noise estimate if the system temperature varies sharply on short time-scales; this was not found to be the case.

The sky-plane pixels size was set to 25 mas, providing more than sufficient sub-sampling of the sky signal. Additionally, a mask was used to impose zero surface brightness on the regions of sky where no emission was expected.

In Figure 3.5, we show the average reconstructed source surface brightness distribution and the resulting sky-model for the Band 6 and 7 datasets. We found that the structure of the source surface brightness distribution is in general consistent between the two Bands, which provided an independent check of our source reconstruction. However, there are variations in the relative surface brightness of the structure within each Band (see Section 3.7 below for discussion). The high surface brightness inner region of the source is reconstructed at higher resolution and at a higher signal-to-noise ratio compared to the rest of the source plane, as it is close to the caustic and hence highly magnified.

We determined the residuals between the data and the model in the visibility plane, and found that they are within the uncertainties of the individual visibilities in the real and imaginary parts. The residuals were also within the expected rms over the whole dataset. For example, 4.7 and 0.3% of the model visibilities were within the 2- and 3-σ off the measured ones, respectively, which is in good agreement with what would be expected from a Gaussian distribution of the residuals (white noise). However, to illustrate the goodness of fit of our model, we also show in Figure 3.5 the dirty-image residuals.

3.4.3 CO emission lines modelling

In order to model the three-dimensional spectral maps of the CO lines, we first subtracted the continuum emission from the *uv*-plane data using the line-free channels. Given the lower signal-to-noise ratio of the lines emission, and to limit the number of visibilities used in our analysis and to ensure that we compare the various transitions on the same angular-scale, a 1 Mλ cut in the *uv*-distance was imposed (Figure 3.4). This translated into a typical source-plane resolution of about 100 parsecs. A similar *uv*-plane taper had to be employed by ALMA Partnership (2015b) to achieve a good image quality. For a Briggs weighting scheme with robust parameter = 0, this corresponds to a synthesised beam with a full width at half-maximum (FWHM) of 180 × 131 mas at a position angle of 77.7 deg. Only the CO (5-4) and (8-7) transitions had SNR good enough to obtain a robust reconstruction of the velocity map, while the CO (10-9) could be only modelled as an integrated line.

For the CO (5-4) and (8-7) lines, as the emission lines were quite broad, we averaged the data in frequency across several channels, resulting in a velocity resolution of $41\,\mathrm{km\,s^{-1}}$ for both lines. This was found to be a reasonable compromise between velocity space sampling (to study the structure of the lines) and signal-to-noise ratio (SNR) required for a robust source reconstruction. Keeping the lens model fixed, we re-optimised for the source regularization parameter λ_s and the source surface brightness distribution. The gradient form of the regularization was found to provide a better fit to the data due to the complex structure of the source. For the CO (5-4) frequency bins #17 and #18 and CO (8-7) bin #4, the signal-to-noise ratio was found to be too low to obtain a source reconstruction; these channels were then excluded from the following kinematic analysis.

For the CO (10-9) line, we averaged the channels 150-549 (400 channels in total) into a single frequency bin, centered at $f_{10-9} = 285.126$ GHz. Because of the same 1 $M\lambda$ *uv*-distance cut, the CO (10-9) dataset has the same resolution as the CO (5-4) and (8-7) ones. Consequently, the sky-plane mask and grid were set up in the same way as for the (5-4) and (8-7) transitions; we only re-optimized for the source

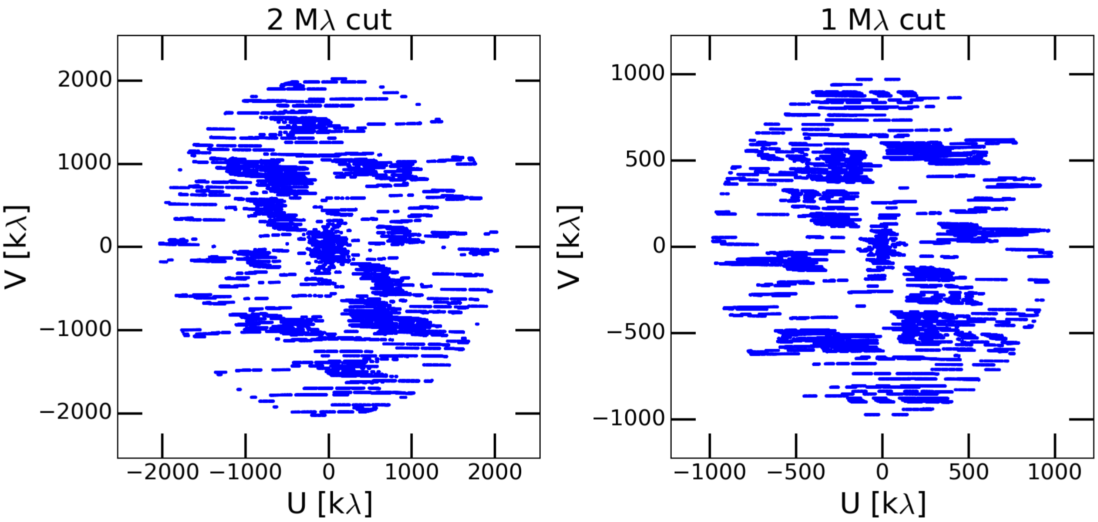

Figure 3.4: Band 7 *uv*-plane coverage used for the lens modelling. *Left*: 2 Mλ cut used for the continuum analysis. *Right*: 1 Mλ cut used for CO lines modelling. Band 6 includes additional baselines and hence has a marginally better *uv*-plane coverage.

regularization λ_s and the source-plane surface brightness distribution.

3.5 Results

3.5.1 Bands 6 and 7 continuum

The pixelated reconstruction of the continuum source-plane surface brightness distribution in Bands 6 and 7 reveals a complex morphology. There are three distinct regions of the source that are resolved in our pixelated reconstructions, embedded in a wider region of low-surface brightness emission (see Figure 3.5). This outer region is consistent with the extended emission that forms the Einstein ring seen in our low resolution images (see Figure 3.2). It is clear that the sub-kpc structure of the galaxy is not uniform, but contains several regions of intense dust emission that are presumably associated with recent star-formation. In particular, the central part of the source is elongated and has an extent of $\sim$1.9×0.7 kpc. This morphology is reminiscent of a star-forming disk (e.g. as in GN20; Hodge et al., 2015), but it could also be due to the multiple components of an ongoing merger (e.g. Swinbank et al., 2006; Tacconi et al., 2006; Engel et al., 2010). We note that photo-dissociation region models for SDP.81 from an analysis of the [O III] lines in the far-infrared spectrum, as measured with *Herschel*, constrain the size of the emitting region to be 500–700 pc (effective radius; Valtchanov et al., 2011), which is consistent with the overall size of the central star forming region that we detect. However, the individual regions with the most intense dust emission are smaller than this size, that is, $\sim$500 pc.

Based on our reconstruction of the Bands 6 and 7 continuum, we estimate the total magnification to be $\mu_{tot} = 17.6 \pm 0.4$. This magnification is dominated by the

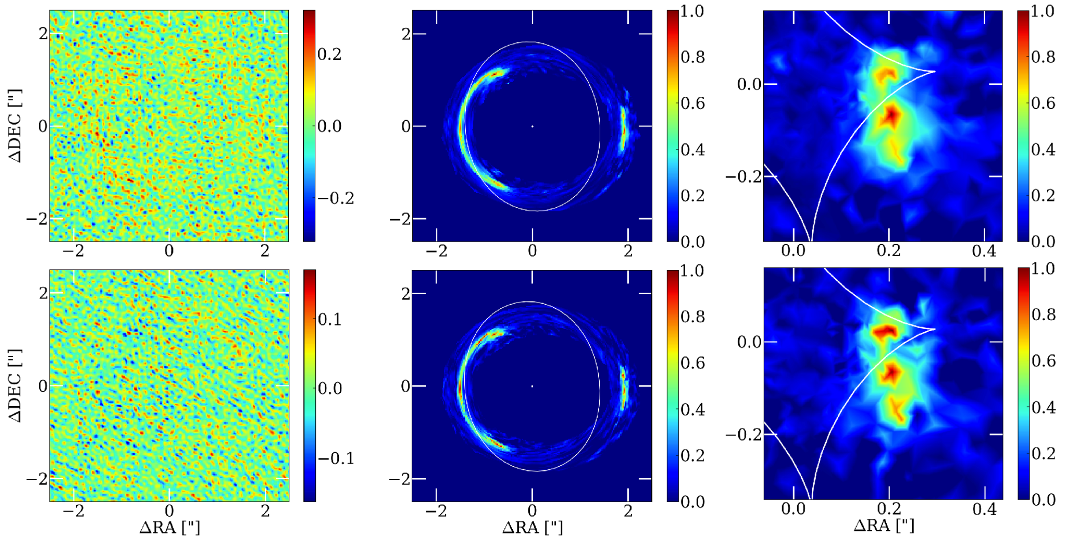

Figure 3.5: Reconstruction of the Bands 6 (*upper*) and Band 7 (*lower*) continuum. *Left*: the image-plane residuals from the Fourier transform of the residual visibilities (data-model) and gridded assuming uniform weighting. *Centre*: the modelled image-plane brightness distribution with critical curves. *Right*: the reconstructed source-plane brightness distribution with caustics. The image-plane residuals are in units of mJy beam^{-1}. The surface brightness of the reconstructed source is normalised with respect to the brightest pixel. The lens centre is at position (0,0).

central star-forming disk, which has a magnification of $\mu_{disk} = 25.2 \pm 2.6$. The total flux-density of the source at Bands 6 and 7 at <0.5 Mλ from our fit to the visibility data are 21.7 ± 2.2 and 38.9 ± 3.9 mJy, respectively, assuming a 10% uncertainty on the flux-density calibration.

3.5.2 CO emission lines

The source-plane velocity integrated emission of the CO (5-4), (8-7) and (10-9) transitions are presented in Figure 3.6. These maps show clear excitation dependent structure in the molecular gas distribution of SDP.81. The CO (5-4) emission has both diffuse and multiple compact structures that extend over 3 kpc in the north-south direction, with an additional low surface brightness component that is 1–2 kpc in size towards the north-east of the source. The emission from the CO (8-7) and (10-9) transitions is much more compact, with an extent of about 1.5 kpc, and there is only a single brightness clump. There is also evidence of a faint tail of emission that extends by 2 kpc to the south-west; this tail is not seen in the CO (5-4) transition, which implies there is higher-excitation gas in that region, possibly due to shocks. As a check, we compared our reconstructions with the image-plane moment-zero maps of the counter-arc (ALMA Partnership et al., 2015b); the counter-arc does not cross a critical curve (see Figure 3.5), hence the magnification is approximately constant across this image. We find that our source-plane reconstructions are in good agreement with the observed image-plane structure of the counter-arc. The total magnification of the CO (5-4) and CO (8-7) gas components are 17.0±0.4 and

16.9±1.1, respectively.

The total velocity integrated emission from the CO (5-4) and CO (8-7) transitions has been determined by directly integrating the zeroth-moment maps over the total extent of the reconstructed source. We find that the velocity integrated line intensities are 0.54 ± 0.06, 0.45 ± 0.03 and 0.075 ± 0.008 Jy km s^{-1} for the CO (5-4), CO (8-7) and CO (10-9), respectively (Table 3.2). These correspond to line luminosities of $L'_{\rm CO(5-4)}$= (0.91±0.10)×10^{10} K km s^{-1} pc^2, $L'_{\rm CO(8-7)}$= (0.27±0.02)×10^{10} K km s^{-1} pc^2 and $L'_{\rm CO(10-9)}$= (0.31±0.03)×10^{10} K km s^{-1} pc^2 . This gives an intrinsic line luminosity ratio of $r_{8-7/5-4} = 0.30 \pm 0.04$, which is consistent with other starburst galaxies at z =2–3 that typically show low excitation (e.g. Scott et al., 2011).

Ideally, we would like to compare line luminosities of this higher-excitation transitions with the line luminosity of the CO (1-0) emission to study the CO ladder; in this way, a comparison with CO-ladder models for high-redshift galaxies (e.g. Narayanan & Krumholz, 2014) could be made. However, since the CO (1-0) line if offset from the dust continuum, it would be incorrect to apply the magnification derived from the latter to the observed CO (1-0) spectrum. It would also be incorrect to apply the magnification derived from the higher transitions as the CO (1-0) is expected to be much more extended, up to 15 kpc FWHM (Ivison et al., 2011).

From the intrinsic source-plane line profiles of the CO (5-4) and (8-7) transitions (Figure 3.7), we find that both transitions have symmetric double-horn profiles, typical of rotating disks, with velocity peaks that are separated by 290 km s^{-1}. Again, this result is consistent with the image-plane line profiles of the counter arc. The spatially dependent structure of the CO emission lines extends over the caustic curves, resulting in a strong differential magnification effect: as the source crosses the caustic, the red channels have much higher magnification than the blue ones. As a result, the observed image-plane line profiles are severely distorted (Figure 3.7): in the case of the CO (5-4) line, the differential magnifications completely erases the double-horned profile. The CO (5-4) and CO (8-7) emission lines are also offset by about ∼80 km s^{-1} from the systemic velocity, as defined by the CO (1-0) measurement (Frayer et al., 2011). This could be due to the presence of a turbulent inflows or outflows of higher-excitation gas, or to the fact that the CO (1-0) line also undergoes differential magnification, biasing the inferred systemic velocity.

The moment-one and moment-two maps (velocity and velocity dispersion) of the CO (5-4) and CO (8-7) transitions are presented in Figure 3.8 and Figure 3.9, respectively. The velocity maps for both emission lines show a common central solid-body-like component in the north-south axis. However, the diffuse and extended component of the CO (5-4) and the extended tail of the CO (8-7) show a different velocity from what would be expected for a solid body rotation. Although the CO (5-4) extended feature has a slightly higher velocity dispersion than the rest of the emission, the high-velocity dispersion tail of the CO (8-7) tail is different in nature: the large 200 km s^{-1} velocity dispersion of the CO (8-7) tail is not due to a single broad velocity component, but is due to two velocity components along the same line-of-sight. The velocity dispersion of the CO (8-7) is overall higher than the CO (5-4) by about 40 km s^{-1}; however, the relatively coarse resolution of the data

Table 3.2: Total magnifications and intrinsic source properties for the continuum and CO lines.

Emission	μ	S_ν [mJy]	$S_\nu\Delta$v [Jy km s^{-1}]	L' [10^{10} K km s^{-1} pc^2]
Band 6 + 7 continuum	17.6±0.4	3.6±0.2	-	-
CO (5-4)	17.0±0.4	-	0.54±0.06	0.91±0.10
CO (8-7)	16.9±1.1	-	0.45±0.03	0.27±0.02
CO (10-9)	23.5±2.5	-	0.075±0.008	0.32±0.03

cube used for the line reconstruction might contribute to this discrepancy.

3.6 Kinematic analysis of CO 5-4 and CO 8-7

3.6.1 Kinemetry

In order to quantify the kinematic properties of SDP.81, we used the kinemetry method of Krajnović et al. (2006) to model the CO (5-4) and (8-7) moment maps. Kinemetry is based on the idea that the velocity/velocity dispersion fields of the reconstructed source can be described by a series of concentric ellipses of increasing major-axis length; the first (velocity) and second (velocity dispersion) moment profiles along each ellipse is then decomposed into a Fourier series with a small number of harmonic terms.

The first and second moment maps $M_1(x,y)$, $M_2(x,y)$ are obtained straightforwardly as:

$$M_1(x,y) = \left(\sum_\nu I_\nu(x,y)\right)^{-1} \sum_\nu v_\nu \cdot I_\nu(x,y), \tag{3.1}$$

and

$$M_2(x,y) = \left[\left(\sum_\nu I_\nu(x,y)\right)^{-1} \sum_\nu I_\nu(x,y)\,(v_\nu - M_1(x,y))^2\right]^{\frac{1}{2}} \tag{3.2}$$

where x and y denote the source-plane position, ν and v_ν the frequency and rest-frame velocity of a given channel and $I_\nu(x,y)$ the surface brightness at a position (x,y) for a given frequency.

In the kinemetry approach, the i-th moment map is decomposed as:

$$M_i(a,\psi) = A_{0,i}(a) + \sum_{n=1}^{N} \left(A_{n,i}(a)\sin(n\psi) + B_{n,i}(a)\cos(n\psi)\right) \tag{3.3}$$

where a and ψ are the semi-major axis and the azimuthal angle of the ring (measured from its major axis). This can be re-written as:

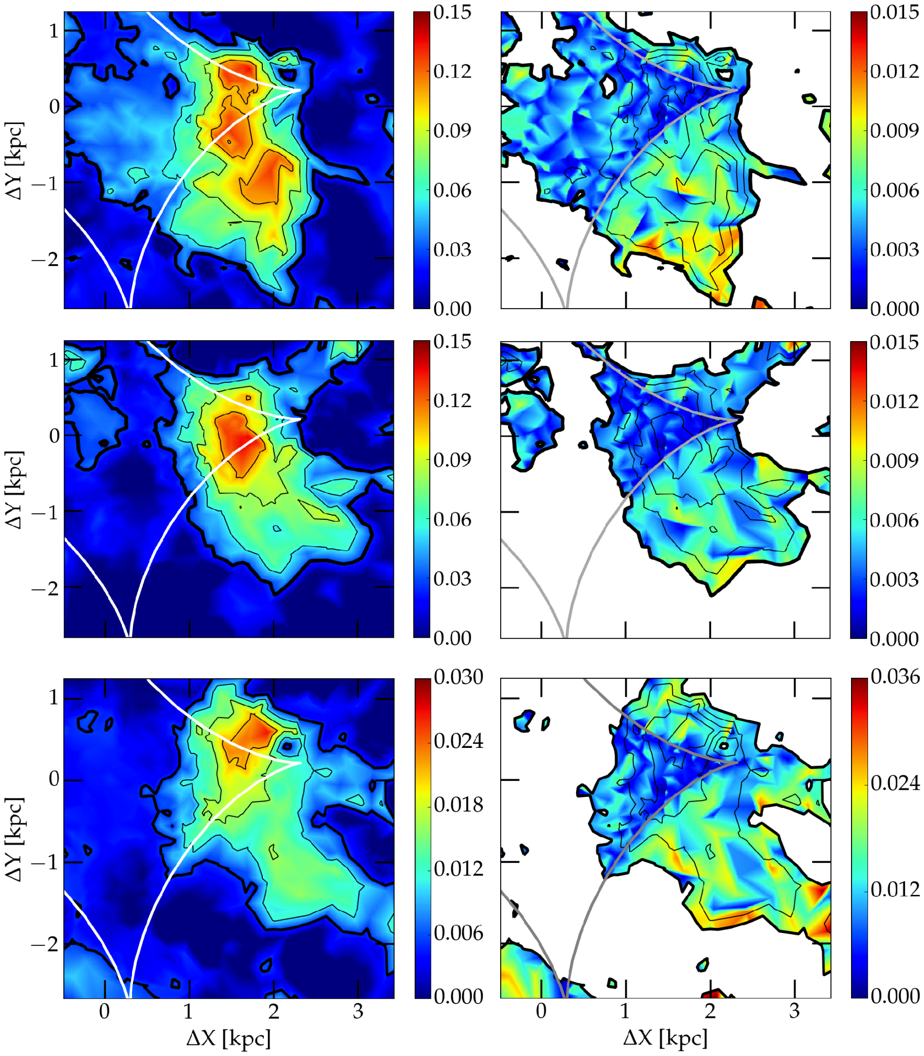

Figure 3.6: *Left*: The moment-zero maps for the CO (5-4) and (*centre*) CO (8-7) and (*lower*) CO (10-9) emission lines. *Right*: corresponding moment-zero uncertainties. The colourmap is in units of Jy km s^{-1} kpc^{-2}. The lens centre is at co-ordinate (0, 0) and the source-plane caustic is shown by the (*left*) white and (*right*) grey lines. The contours are drawn at 0.2, 0.4, 0.6 and 0.8 relative to the peak emission. Note that the in all three cases, the CO emission crosses the caustics, resulting in a strong differential magnification effect, as seen in Figure 3.7.

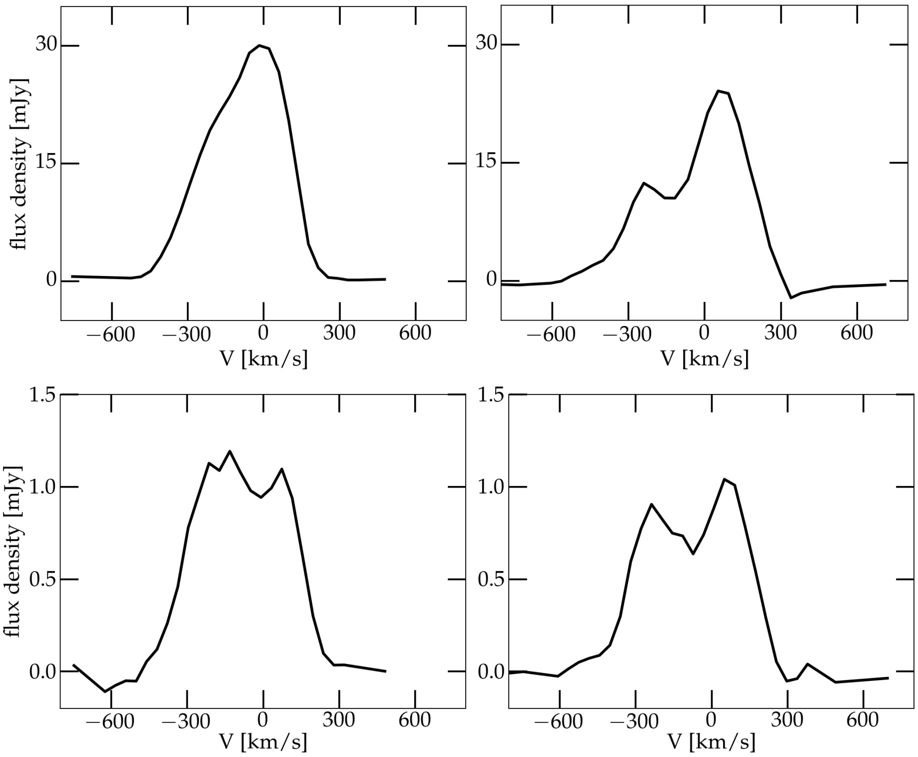

Figure 3.7: Line profile for the CO (5-4) (*left*) and (8-7) (*right*) lines. *Upper*: the observed plane-of-the-sky line profiles. *Lower*: the intrinsic source-plane profiles. The line profiles were smoothed with a boxcar filter 120 km s^{-1} wide (equal to three frequency bins), assuming a systemic velocity corresponding to a redshift of 3.042 and using the radio definition of the velocity.

$$M_i(a,\psi) = A_{0,i}(a) + \sum_{n=1}^{N} k_{n,i}(a)\cos\left[n(\psi - \phi(a))\right] \tag{3.4}$$

with

$$k_{n,i} = \sqrt{A^2_{n,i}(a) + B^2_{n,i}(a)} \tag{3.5}$$

and

$$\phi_{n,i} = \arctan\left(\frac{A_{n,i}}{B_{n,i}}\right) \tag{3.6}$$

For an ideal rotating disk, the only non-zero terms in the expansion would be the $B_{1,v}$ and $A_{0,\sigma}$ terms, corresponding to a circular velocity and a constant velocity dispersion, respectively.

3.6.2 Distinguishing disks from mergers

To classify the kinematic state of galaxies with kinemetry, Shapiro et al. (2008, hereafter S08) have used simulated SINFONI observations of $z \sim 2$ galaxies to devise a criterion that would allow to determine whether a galaxy is dominated by disk-like or merger-like kinematics. This method is based on the total kinematic asymmetry in the three-dimensional moment maps of the molecular gas, which is quantified as

$$K_{asym}(S08) = \sqrt{v_{asym}^2 + \sigma_{asym}^2} \tag{3.7}$$

with

$$v_{asym} = \left\langle \frac{k_{\mathrm{avg},v}}{B_{1,v}} \right\rangle_a \text{ and } \sigma_{asym} = \left\langle \frac{k_{\mathrm{avg},\sigma}}{B_{1,v}} \right\rangle_a \tag{3.8}$$

and $k_{\mathrm{avg},v} = (k_{2,v}+k_{3,v}+k_{4,v}+k_{5,v})/4$ and $k_{\mathrm{avg},\sigma} = (k_{1,\sigma}+k_{2,\sigma}+k_{3,\sigma}+k_{4,\sigma}+k_{5,\sigma})/5$. Here the $B_{1,v}$ term corresponds directly to the ordered rotation of an ideal rotating disk, while $A_{1,v}$ corresponds to a velocity gradient along the minor axis. While $A_{1,v}$ is in principle a measure of asymmetry in the system, this term can be significantly influenced by radial outflows/inflows due to AGN or bar instabilities, which would be expected even for a very ordered disk (S08).

This criterion has been subsequently revised by Bellocchi et al. (2012, hereafter B12), in order to better distinguish between isolated disks and post-coalescence mergers on the basis of large asymmetries at large radii. In this new approach the asymmetry of each ring is weighted by its circumference a, leading to the following equations:

$$v_{asym} = \sum_a \left(\frac{k_{\mathrm{avg},v}(a)}{B_{1,v}(a)} \right) \cdot \frac{1}{\sum_a C(a)}, \tag{3.9}$$

and

$$\sigma_{asym} = \sum_a \left(\frac{k_{\mathrm{avg},\sigma}(a)}{B_{1,\sigma}(a)} \right) \cdot \frac{1}{\sum_a C(a)}\,. \tag{3.10}$$

The total asymmetry term $K_{asym}(B12)$ is then defined similarly to Equation 3.7.

The physical motivation for this weighting scheme is related to the expectation that after a merging event, the inner regions of the merger remnant will rapidly relax into a rotating disk, while the outer parts are expected to be still out of equilibrium.

3.6.3 Kinematic status of SDP.81

Using the kinemetry, we have performed a kinematic analysis of both the CO (5-4) and CO (8-7) lines, independently. As we mentioned earlier, the CO (10-9) line is too faint for its velocity profile to be reliably reconstructed. We used the CASA `immoments` task to obtain the moment maps from our source-plane reconstructions. From this kinematic analysis we found that both transitions are characterized by a significant level of rotation as well as a significant level of asymmetry at large radii. In the central regions of both transitions, the first harmonic term, k_1, which describes the rotational motion, is substantial and the ratio k_5/k_1, which represents higher-order deviations from simple rotation, is consistent with zero, as expected for a disk-like rotation. However, in the outer regions, the k_1 term abruptly drops and at the same time the k_5/k_1 term significantly increases. This behaviour is expected for multiple kinematical components or for post-coalescence perturbed disks (Krajnović et al., 2006; Bellocchi et al., 2012).

We quantify the total kinematic asymmetry to be $K_{asym}(S08) = 0.65 \pm 0.10$ and $K_{asym}(B12) = 0.63 \pm 0.14$ for the CO (5-4) transition and $K_{asym}(S08) = 0.65 \pm 0.14$ and $K_{asym}(B12) = 0.72 \pm 0.20$ for the CO (8-7) transition. The uncertainties are a combination of the errors on the Fourier coefficients and the uncertainty due to the choice of centre of the velocity map. These values are consistent with a merger/post-coalescence merger interpretation according to both criteria: $K_{asym}(S08) > 0.50$ and $K_{asym}(B12) > 0.15$.

It is worth mentioning that the classification criteria of both S08 and B12 were devised for data with angular resolutions much coarser than the data presented here ($>$ 0.5 arcsec in S08, B12, compared to $\sim$0.15 arcsec in this work). However, we found that downgrading our resolution does not affect the derived K_{asym} values.

In addition to the kinemetry analysis, we assessed the kinematic state of SDP.81 using the visual criteria for kinematic classification proposed by Bellocchi et al. (2013). Based on the visual analysis, SDP.81 is classified as a perturbed disk; while the velocity maps show a clear disk-like velocity gradient, the velocity dispersion map is relatively smooth without an obvious central peak, indicating that the disk is offset from the actual centre of mass.

Finally, we estimated the dynamical mass of the disk structure by considering the maximum radius (1.56 ± 0.07 kpc), maximum velocity (150 ± 20 km s^{-1}) and the inclination angle (45 ± 8 deg) of the outermost kinemetry ellipse that is consistent with an ordered rotation. We found an enclosed dynamical mass of $M(< 1.56\,\mathrm{kpc}) = (1.6 \pm 0.6) \times 10^{10} M_{\odot}$, which, although a factor of ~ 2 smaller than that determined by Swinbank et al. (2015), is consistent with this result at the $2-\sigma$ level. However, as the system is far from being in equilibrium, such dynamical mass estimates should be interpreted with care.

3.7 Discussion

3.7.1 Source morphology

In Figure 3.10, we present the composite source-plane image of SDP.81. It is apparent that the reconstructed rest-frame UV/optical stellar component (HST F160W) the combined CO (5-4) and CO (8-7) molecular gas component and the heated dust from the rest-frame FIR continuum emission form a complex structure. Taking the UV/optical stellar emission first, we see that this component is dominated by an elongated structure about 4 kpc long, that is oriented by 45 deg with respect to the north-south axis. There is another brightness peak about 4 kpc south of the main UV/optical peak; the two components are connected by a low-surface brightness bridge. The diffuse UV/optical structure extends about 10 kpc south; this structure is coincident with the CO (1-0) emission, as seen in Figure 3.3. The CO (5-4) and (8-7) molecular gas distribution looks disturbed, with the region to the east being much more diffuse, whereas there is a sharp cut-off in the gas towards the west,

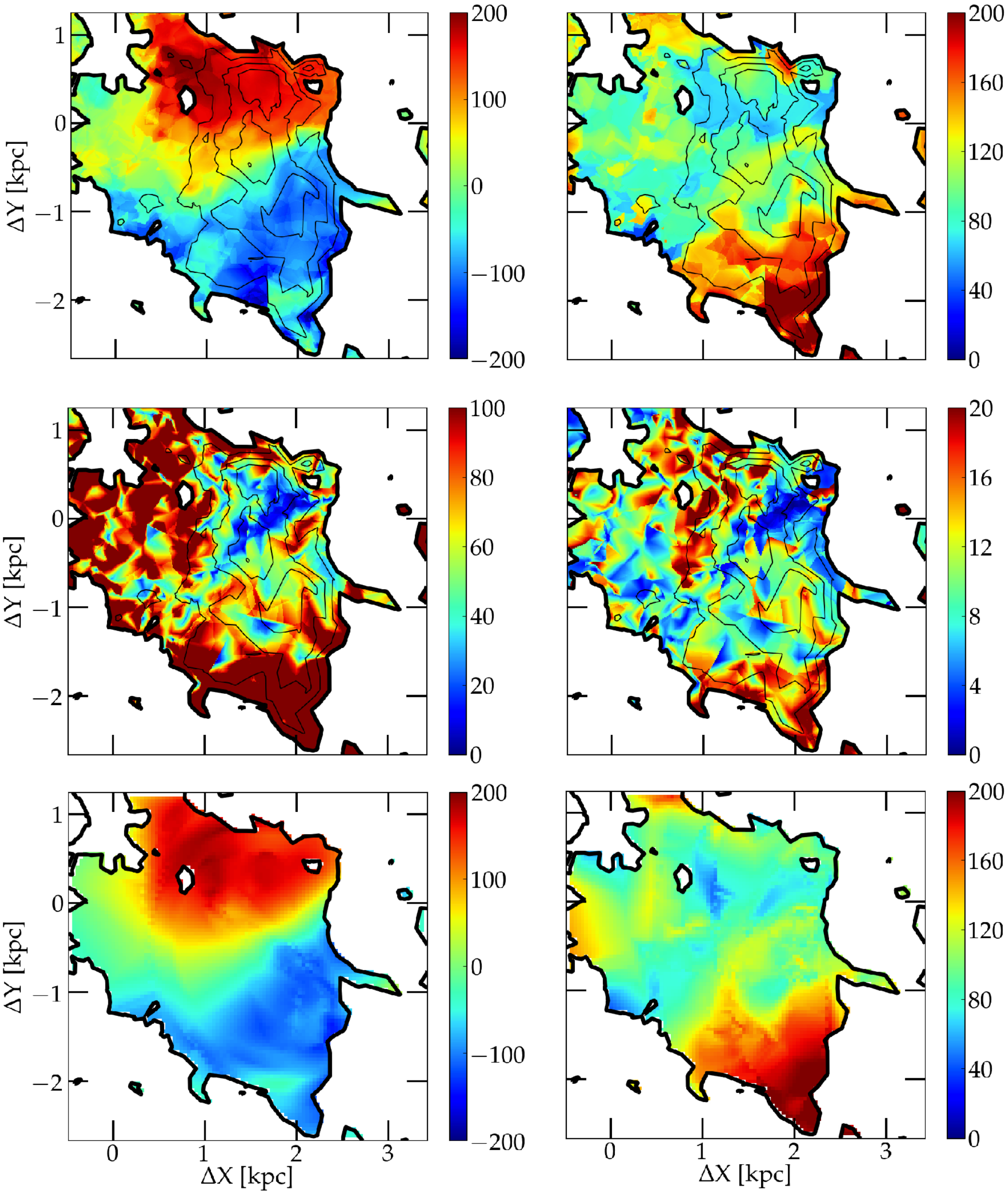

Figure 3.8: The un-smoothed first-moment (*upper-left*), second-moment (*upper-right*), first-moment uncertainty (*middle-left*) and second-moment uncertainty (*middle-right*) maps for the CO (5-4) transition. The lower panel shows the modelled (*left*) first- and (*right*) second-moment maps from our kinemetry analysis. The colour-bar is in units of $\mathrm{km\,s^{-1}}$, and for the first moment maps is shifted by $-80\,\mathrm{km\,s^{-1}}$, from a systemic velocity corresponding to a redshift of 3.042 (see Figure 3.7) using the radio definition of the velocity.

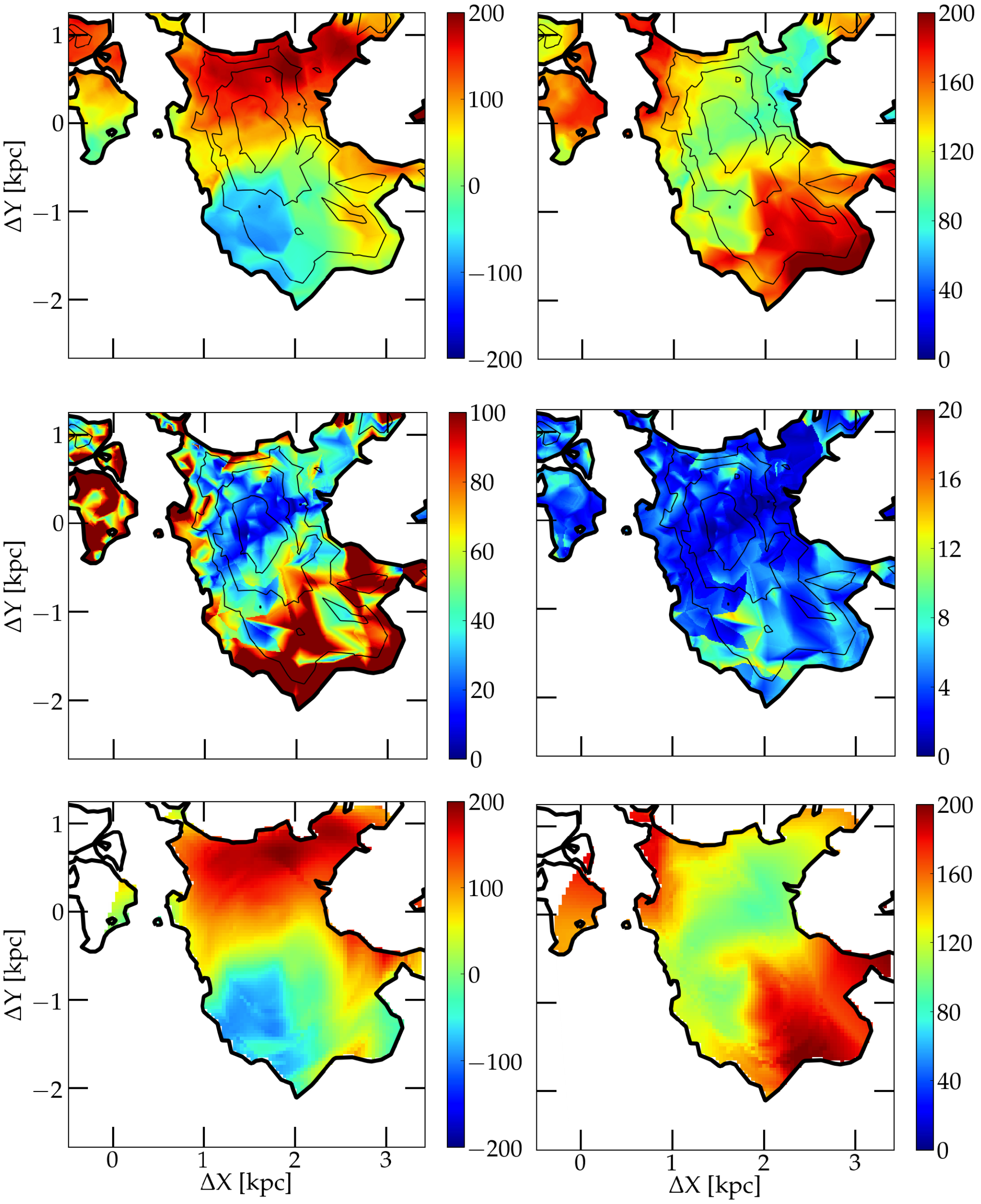

Figure 3.9: The same as Figure 3.8, but for the CO (8-7) transition.

which forms a *ridge* at the contact point with the stellar component (shown by the small distance between the intensity contours in Figure 3.10). Interestingly, the extended tail that is seen in the CO (8-7) and CO (10-9) molecular gas distribution neatly fits into the area between the main UV/optical component and the bridge feature. Compared to the other two CO transitions and the continuum emission, the CO (10-9) emission is offset further to the west of the dust/CO(5-4) emission, similarly to the CO (8-7) emission.

The ongoing star-formation, as traced by the heated dust emission, is concentrated in a region of about 2 kpc in size, and is co-incident with the disk-like rotating structure within the inner part of the molecular gas distribution. This component also shows extension to the east and a sharp cut-off to the western edge. We note the detection of a compact molecular gas and star-formation component (20-60 $\mathrm{M_{\odot} yr^{-1}\,kpc^{-2}}$) near the brightness peak of the UV/optical component, suggesting that star-formation is occurring in both structures. However, we find no further evidence of either heated dust or excited molecular gas emission across the full 15 kpc of the optical component, even though there is an established stellar population and a large reservoir of CO (1-0) available for star formation. This suggests that the current burst of star-formation is localised to the northern part of the structure.

Some further insights can be gained by comparing the morphology of SDP.81 with a $z = 4.05$ dusty star-forming galaxy GN20 (see Chapter 1), for which a similar multi-wavelength dataset is available (Hodge et al., 2013, 2015). In particular, the central structure of SDP.81 is remarkably similar to that of GN20: the rest-frame UV/optical emission is offset from the molecular gas component (as traced by lower excitation CO line) and the heated dust emission. Indeed, these offsets are not uncommon in high-redshift galaxies and may be attributed to dust extinction of the UV/optical stellar emission or to strong negative stellar feedback together with gas accretion from gas clumps or satellites (Maiolino et al., 2015). Like GN20, there is also evidence for a central star-forming disk at the centre of SDP.81. However, there are also marked differences, which suggest that the starburst in SDP.81 may be within an interacting/merging system. First, our kinemetry analysis of the extended molecular gas distribution suggests a post-coalescence perturbed disk (no such analysis is available for GN20). Also, the UV/optical emission shows multiple compact and diffuse components over an elongated ∼15 kpc region. The CO (1-0) data in hand are not sensitive enough or have high enough angular resolution to determine the relative velocities of these individual components, but it is clear from the overall CO (1-0) emission that they form part of the same extended structure.

As outlined in Chapter 1, a merger-based scenario has been put forward as the triggering mechanism behind the strong star-formation observed within sub-mm galaxies. This is based on the observed multiple-component morphology and disturbed kinematics of unlensed starburst galaxies and from simulations (e.g. Engel et al., 2010; Narayanan et al., 2010a). Indeed, this conclusion was also recently drawn by Messias et al. (2014) who studied the source-plane properties of a lensed merging system H-ATLAS J1429−0028. For SDP.81, the situation is still not clear, but further observations of the CO (1-0) emission with the VLA at higher angular

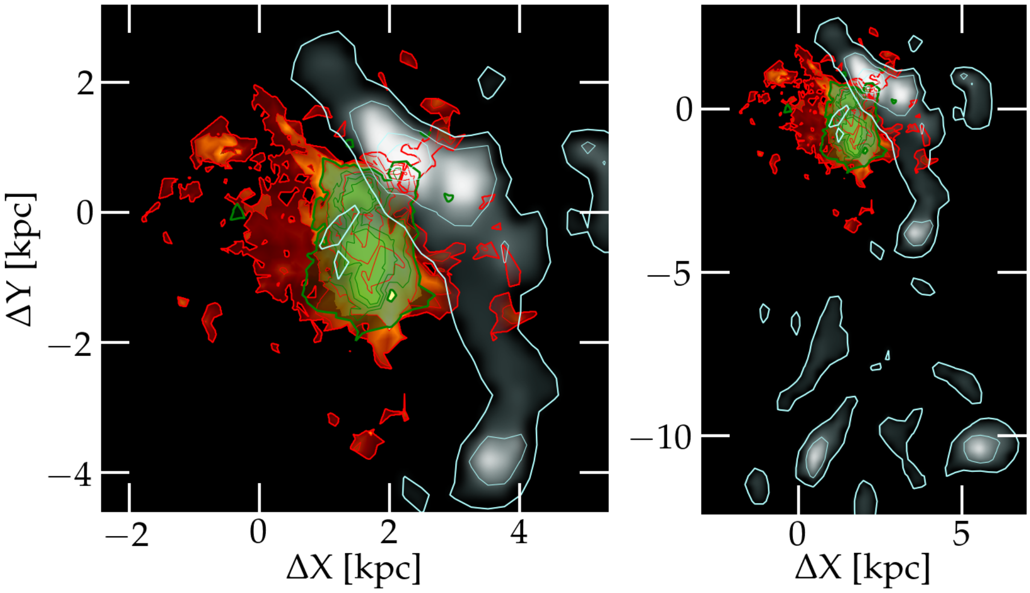

Figure 3.10: A composite source-plane reconstruction of the rest-frame FIR continuum dust (green, ALMA Bands 6+7), the combined CO (5-4) and CO (8-7) molecular gas (red) and the rest-frame UV/optical emission (grey).

resolution could measure the relative dynamics of the various UV/optical components. This would confirm that they form a merging system, with large velocity offsets, or if they are part of a clumpy structure with a more chaotic velocity distribution.

3.7.2 Star-formation rate

Using our measurement of the total magnification, we have determined the intrinsic properties of SDP.81 from a fit to the UV/optical-to-sub-mm broad-band spectrum that was carried out by Negrello et al. (2014). We find a star-formation rate of $315 \pm 60\,\mathrm{M_\odot\,yr^{-1}}$, a far-infrared luminosity of $(3.1 \pm 0.4) \times 10^{12}\ \mathrm{L_\odot}$ and a dust mass of $(6.4 \pm 0.5) \times 10^{8}\ \mathrm{M_\odot}$[3]. These results are consistent with what is typically found for sub-mm galaxies at redshift $z \sim 2.5$ (e.g. 400 $\mathrm{M_\odot yr^{-1}}$, Coppin et al., 2008).

In Figure 3.11, we show the star-formation rate density of the reconstructed source, under the assumption that all the heated dust emission is solely due to star formation. Also shown is the uncertainty in the star-formation rate density, which was calculated from the rms over all spectral channels. We find that the extended component has a moderate SFR surface density of $20 - 30\,\mathrm{M_\odot\,yr^{-1}\,kpc^{-2}}$, but there are multiple regions of intense star formation ($>100\mathrm{M_\odot\,yr^{-1}\,kpc^{-2}}$) as opposed to a single starburst site. The star-forming disk has a mean SFR surface density of $100\,\mathrm{M_\odot\,yr^{-1}\,kpc^{-2}}$. This is several orders of magnitude higher than is seen within disk galaxies in the local Universe (Leroy et al., 2013), but is comparable to the

---

[3] Here, we only consider the properties derived from the far-infrared part of the spectrum, since this is the parameter space where our magnification estimate is valid.

extended disk seen in GN20 at redshift $z = 4$ (Hodge et al., 2015). The maximum star-formation rate density in our reconstructed map is $190 \pm 20\,\mathrm{M_\odot\,yr^{-1}\,kpc^{-2}}$, which is below the theoretical expectation for Eddington-limited star-formation by a radiation-pressure supported starburst ($1000\ \mathrm{M_\odot\,yr^{-1}kpc^{-2}}$). There are several other examples of high redshift sub-mm luminous galaxies that have sub-Eddington starbursts (Bussmann et al., 2012; Hodge et al., 2015), but this is the first time that the star-formation density of a galaxy has been directly mapped on sub-50 parsec-scales. Consequently, high-resolution observations of a larger sample of high-redshift sub-mm galaxies are required to determine if the general population of dusty star-forming galaxies are considerably below the Eddington limit for radiation-pressure supported starbursts.

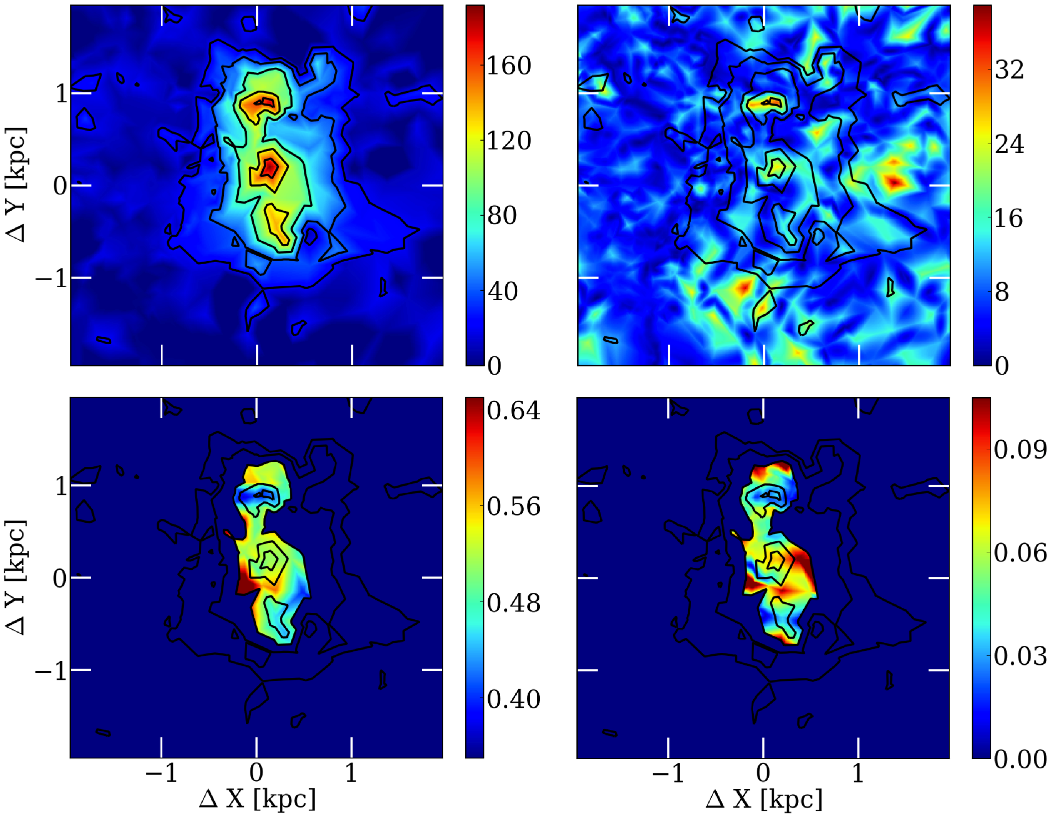

Figure 3.11: *Top:* The star-formation rate density (*left*) and the uncertainty (*right*) of the SDP.81 reconstructed source at redshift 3 shows multiple regions of intense star-formation. The colour-bar is in units of $\mathrm{M_\odot\,yr^{-1}\,kpc^{-2}}$; the peak star-formation rate density is $190\pm20\,\mathrm{M_\odot yr^{-1}kpc^{-2}}$. *Bottom*: The flux-ratio between 236 and 290 GHz (*left*) and the uncertainty (*right*) for a region with a star-formation rate density $> 80\,\mathrm{M_\odot\,yr^{-1}\,kpc^{-2}}$. The colour-bar is in units of flux-ratio. For all plots the isocontours of the star-formation rate density are set to 20, 40, 80, 120 and 160 $\mathrm{M_\odot\,yr^{-1}\,kpc^{-2}}$.

3.7.3 Indications of varying dust properties

We also see tentative evidence for a change in the dust continuum slope over the resolved extent of the intense star-forming regions within SDP.81 (see Figure 3.5 for the relative brightness of the regions). As shown in Figure 3.11, between 1.0 and 1.3 mm, the flux-ratio of the entire source is 0.56 ± 0.08 (or a spectral index of $\alpha = 2.8$, where $S_\nu = \nu^\alpha$), which is as expected for the best fit modified black-body model for the dust emission (temperature $T_{dust} = 34 \pm 1$ K and emissivity β=1.5; Bussmann et al., 2012). We find that the central dominant clump of dust emission in our maps has a flux-ratio of 0.51 ± 0.07 between 1.3 and 1.0 mm, but the two additional clumps of intense dust emission to the north and south have flux-ratios of 0.41 ± 0.03 and 0.45 ± 0.05, respectively (equivalent to $\alpha = 4.3 - 4.9$).

The change in the spectral slope of the continuum emission suggests that the different sites of intense star-formation have different physical conditions, but further observations at high angular resolution with ALMA Bands 9 and 10 will be needed to determine the extent of the variation in the dust temperature and/or optical depth at these locations. Still, our results show that global temperature models are likely too simple to explain the emission from galaxies with multiple sites of star-formation.

3.7.4 Comparison with SMA 880-μm continuum source

Although our lens model agrees within errors with that of Bussmann et al. (2013) based on SMA 880 μm data, there are significant differences between the reconstructed dust continuum surface-brightness distribution presented in this work, and that obtained by Bussmann et al., 2013 (Figure 3.12) .

As mentioned in Section 3.5, based on our Bands 6 and 7 continuum reconstruction, we derive a total magnification that is ∼60% higher than reported by Bussmann et al. (2013). This difference is most likely due to the differing source structure that we find compared to their study, as compact sources that are located close to the lens caustic can have very high magnifications. Indeed, the extent and surface brightness distribution of the reconstructed source are considerably different: Bussmann et al. (2013) model the source with a single Sérsic profile with an effective radius of 4.1 kpc and at a position angle that is almost perpendicular to the extension seen in the ALMA reconstruction. It is unlikely that these two very different source reconstructions can be reconciled by surface brightness distribution physically changing between 1.0/1.3 mm and 880 μm: even accounting for a non-uniform dust temperature across the source, the 880 μm continuum emission should follow a distribution similar to that of 1.0/1.3-mm continuum. The difference in the overall shape of the reconstructed source is therefore likely caused by the low resolution of the SMA data, combined with the limitations due to the choice of a parametric source model. Indeed, Bussmann et al. (2013) point out that their single parametric source model is too simple to fit the observed surface brightness distribution at 0.6 arcsec angular resolution; this highlights the advantage of using pixelated source reconstructions and 100 mas angular resolution observations to study the intrinsic properties of gravitationally lensed starburst galaxies.

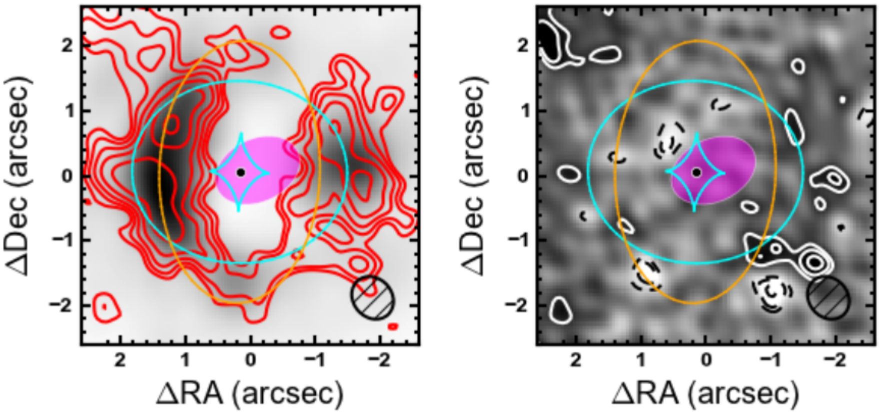

Figure 3.12: Source reconstruction of 880 μm continuum SMA observations of SDP.81, taken from Bussmann et al. (2013). *Left*: CLEANed SMA data (red contours) superimposed on the best sky model. *Right*: Sky-plane residuals. Magenta ellipse marks the half-light extent of the reconstructed source (approximated by a Sérsic profile), cyan and orange lines denote the source-plane caustics and image-plane critical curves, respectively.

3.7.5 Comparison with other reconstructions of the ALMA Long Baseline Campaign data

Thanks to the immediate availability of the ALMA LBC data, a number of different teams applied a wide range of techniques to reconstruct the source and the lens model of SDP.81. Here we compare our reconstructed continuum source with that presented in the literature and discuss the extent of agreement on the lens mass model obtained with different lens-modelling techniques.

We found significant differences in our visibility-plane reconstructions presented in this thesis and the image-plane reconstructions of Dye et al. (2015) and Swinbank et al. (2015). Our reconstructions detect more extended structure within the molecular gas distribution that is not seen in the image-plane analysis. Also, for the image-plane analysis the compact structure varies significantly even between 1.0 and 1.3 mm continuum data. Assuming the dust temperature does not change rapidly across the source, and given the excellent post-calibration sensitivity of 1.0 and 1.3 mm continuum datasets, a strong correlation between the reconstructed emission in these two Bands is expected.

These differences between the *uv*- and image-plane results are likely due to the choice of weighting and deconvolution errors in the low SNR images of the image-plane channel data; the *uv*-taper and weighting that were applied to the image-plane data produces a point spread function with significant side-lobe structure that can introduce spurious artefacts in the channel images. Although Dye et al. (2015) assumed non-correlated errors on the image-plan surface brightness distribution, the sidelobe structure of the ALMA dirty beam indicates that this is not the case. For these reasons, we believe that visibility-plane reconstructions are more robust and less biased, and the interpretation of intensity and velocity maps from image-plane reconstructions requires caution.

A more recent source reconstruction was presented in Hezaveh et al. (2016), who also used a visibility-fitting technique with pixellated sources. The reconstructed Bands 6 and 7 continuum in Hezaveh et al. (2016) is largely similar to that presented in Figure 3.5, providing an independent confirmation of the source-plane dust continuum surface brightness distribution in SDP.81.

Considering the lens model, more than five different methods were used to determine the mass density profile of the main lens. Apart from the lens models obtained by Dye et al. (2015) and Hezaveh et al. (2016) that were mentioned above, Wong et al. (2015) constrained the lens model using image positions of distinct parts of the source. Finally, Tamura et al. (2015) used the parametric GLAFIC package (Oguri, 2010). All the derived lens models agree within their respective errors. While the high resolution and sensitivity of the ALMA LBC data helps in deriving a robust lens model, a smooth lens mass density profile can be largely constrained by image position only. Hence, given a decent data quality, different lens modelling procedures are likely to recover a consistent lens model.

3.8 Conclusions

In this chapter we have presented the first application of a visibility-fitting technique with a pixellated source model to high-resolution ALMA observations. A reconstruction of the dust continuum and the CO emission on 50 - 100 parsec scales present an improvement by an order of magnitude over previous studies of high-redshift dusty star-forming galaxies, both lensed (e.g. Hezaveh et al., 2013b; Bothwell et al., 2013; Messias et al., 2014) and non-lensed (e.g. Hodge et al., 2015).

The combination of a very clumpy interstellar medium and ordered rotation in the inner parts of gas disk is in agreement with studies of other sources (Tacconi et al., 2013; Hodge et al., 2015, 2016). On the other hand, in the outer parts of the system, the visual inspection and the kinemetry analysis provide evidence for velocity-field perturbations consistent with an on-going or post-coalescence merger. Optical imaging suggests a companion located to the north of SDP.81; however, the quality of the data at hand does not allow us to obtain conclusive evidence.

Consequently, to fully understand the nature of SDP.81, as well as to study the physical processes of the dust and gas, observations of additional transitions and of short-wavelength sub-mm continuum are necessary. In particular, the CO (1-0) line is a key to tracing the overall distribution and amount of the cold dust in this system, as CO (1-0) is expected to be spatially extended on scales of more than 10 kpc. Therefore, CO (1-0) data will be crucial to determining the merger/disk-like nature of the SDP.81 and to confirm or refute the presence of a companion system. Unfortunately, the quality of archival VLA observations (Valtchanov et al., 2011) was too low to study the CO (1-0) distribution in the source plane.

Due to its strongly lensed nature which provides a zoomed-in view into its star-forming regions, and the unprecedented quality of the data available or expected from on-going campaigns, SDP.81 is probably the currently best-studied high-redshift star-forming system, and will potentially serve as a valuable testbed to put constraints on various proposed models of high-redshift star-forming processes.

Chapter 4

South Pole Telescope lenses: the low-resolution limit

Abstract

In this chapter, we apply our lens-modelling technique to six strongly gravitationally lensed sub-mm galaxies from the *South Pole Telescope* sample that were observed with the ALMA Cycle-0 array at 0.4 - 2.0 arcsec resolution. These systems were previously modelled by Hezaveh et al. (2013b) and Spilker et al. (2016) using parametric source models. Specifically, we investigate whether our technique reproduces the previously published lens models, and whether pixellated source models can recover extra structure from very low-resolution data. We find our lens models to be largely in agreement with the previously published ones. However, the intrinsic source flux-density predictions differ by up to a factor of a few compared to previous work, due to additional source–plane structure resolved by parametric models. Finally, we find that low-resolution data alone do not necessarily provide enough information to fully constrain the lens model and additional information from e.g. optical imaging of the lensing galaxy is required to obtain a robust lens model.

Note: This chapter is based on Rybak, Vegetti and McKean: "Strong lens modelling of interferometric data with pixellated sources", submitted to MNRAS, August 2016.

4.1 Introduction

The six strongly lensed sources analysed in this chapter were discovered as a part of the *South Pole Telescope* (SPT) survey (Carlstrom et al., 2011); the details of this survey were presented in Chapter 1. Due to the limited resolution of the SPT, a higher-resolution (0.4 - 2.0 arcsec) follow-up with ALMA Cycle-0 array was required to confirm their lensed nature (Vieira et al., 2013; Weiß et al., 2013).

These ALMA Cycle-0 data are a representative example of the follow-up observations of lensed candidates identified in the H-ATLAS, HerMES and SPT surveys; more than a hundred strong lenses were confirmed by imaging with the Submillimeter Array (SMA; Negrello et al., 2010; Bussmann et al., 2013) and ALMA (Hezaveh et al., 2013b; Bussmann et al., 2015; Spilker et al., 2016). The H-ATLAS and HerMES lens models were published by Bussmann et al. (2013, 2015) using the SMA and ALMA Cycle-1 observations, respectively. For the SPT sample, models for four lenses were presented by Hezaveh et al. (2013b, HZ13), while a much larger sample of 47 lenses was analysed by Spilker et al. (2016, SP16).

However, these studies have been limited by parametric models of the source-plane surface brightness distribution. Namely, Bussmann et al. (2013) and SP16 approximate the sources by a Sérsic profile; Bussmann et al. (2015) used an elliptical Gaussian profile, while Hezaveh et al. (2013b) used a circular Gaussian profile approximation. While the low resolution of the data largely justifies the use of simplistic source models, in several cases, even 0.5 - 1.0 arcsec resolution imaging reveals structure that can not be approximated by a simple analytic profile. For example, as shown in the previous chapter, Bussmann et al. (2013), were not able to fully account for the structure in the SMA observations of SDP.81.

Although limited by their low resolution and the parametric approximation, these initial studies greatly contributed to our understanding of properties of dusty-star forming galaxies as a population. The inferred lens- and source-models, often in combination with ancillary multi-wavelength observations, were used to investigate e.g. the far-infrared surface brightness and dust temperature of individual galaxies (Bussmann et al., 2013), kinematics of the CO gas (Messias et al., 2014), the CO-H_2 conversion factor (Spilker et al., 2015) or the distribution of sizes of sub-mm sources (Spilker et al., 2016).

4.2 ALMA Cycle-0 observations

The data analysed in this chapter were taken with the ALMA Cycle-0 array (project numbers 2011.0.00957.S, 2011.0.00958.S; PI: D. Marrone) in ALMA Band 7 (275 - 370 GHz) using four spectral windows (SPWs). Each SPW had a total bandwidth of 2 GHz and 128 frequency channels. Both linear polarizations (XX and YY) were used. The central frequencies of the SPWs were 336.8, 338.7, 348.9 and 350.8 GHz, respectively. Dates of individual observations, array configurations and on-source time for each object are listed in Table 4.1. Figure 4.1 shows the overlay of the ALMA 870-μm continuum imaging with archival HST data (Spilker et al., 2016). For the purposes of brevity, shortened object names are adopted throughout this

Table 4.1: SPT lenses - details of individual observations.

Source	Array	Date	Antennas	Integration [s]
SPT 0346-50	comp.	16/11/2011	14	15×6
–	ext.	04/07/2012	20	10×6
SPT 0418-47	comp	16/11/2011	14	15×6
–	ext.	04/07/2012	20	10×6
SPT 0532 -50	ext.	15/07/2011	20	10×6
SPT 0538-50	comp.	28/11/2011	16	10×6
–	ext.	15/07/2012	20	10×6
SPT 2031-51	hyb.	06/05/2012	17	10×6
SPT 2134-50	hyb.	06/05/2012	17	10×6

chapter: for example, SPT 0346-52 stands for SPT-S J034640-5205.1.

Three distinct array configurations were used for these observations. The *compact* array consisted of 14 to 16 antennas; the maximum baseline length for the compact array was ∼125 m which resulted in a synthesised beam size with 1.5×1.3 arcsec and 2.0×0.9 arcsec FWHM, depending on the number of operational antennas. The *extended* array consisted of 20 antennas with a maximum baseline length of ∼320 m, giving a FWHM synthesised beam size of 0.7×0.5 arcsec. Finally, in the case of SPT 2031-51 and SPT 2134-50, a *hybrid* array with 17 antennas was used. This array had a maximum baseline length of ∼400 m and a FWHM synthesised beam size of 0.4 × 0.4 arcsec. In terms of angular resolution, these array configurations correspond roughly to arrays `alma.3.1` to `alma.3.3` from Chapter 2, but with a much sparser *uv*-plane coverage due to the lower number of antennas. The extended-array observations of the fourth lens from the original HZ13 sample, SPT 0529-54, suffers from a very low signal-to-noise ratio and hence is not analysed here.

4.3 Lens modelling

For each lens, we modelled the data taken in both the compact- and extended-array configurations, combining both linear polarizations to form the total intensity Stokes I.

To remove any residual phase errors, we imaged and self-calibrated the compact and extended array datasets independently, using a solution interval that was equal to the on-source time and by combining the separate spectral windows to maximize the signal-to-noise ratio (SNR) for the solutions. Since the structure of the continuum emission source region is not expected to change significantly over the 2 GHz bandwidth of individual SPWs, we averaged the 128 frequency channels of each SPW together in order to increase the SNR.

Due to considerable offsets of several arcsec between the true target positions and the phase-tracking centre, the primary beam response can vary by up to 20% across the lensed images. Therefore, the primary beam profile had to be taken into

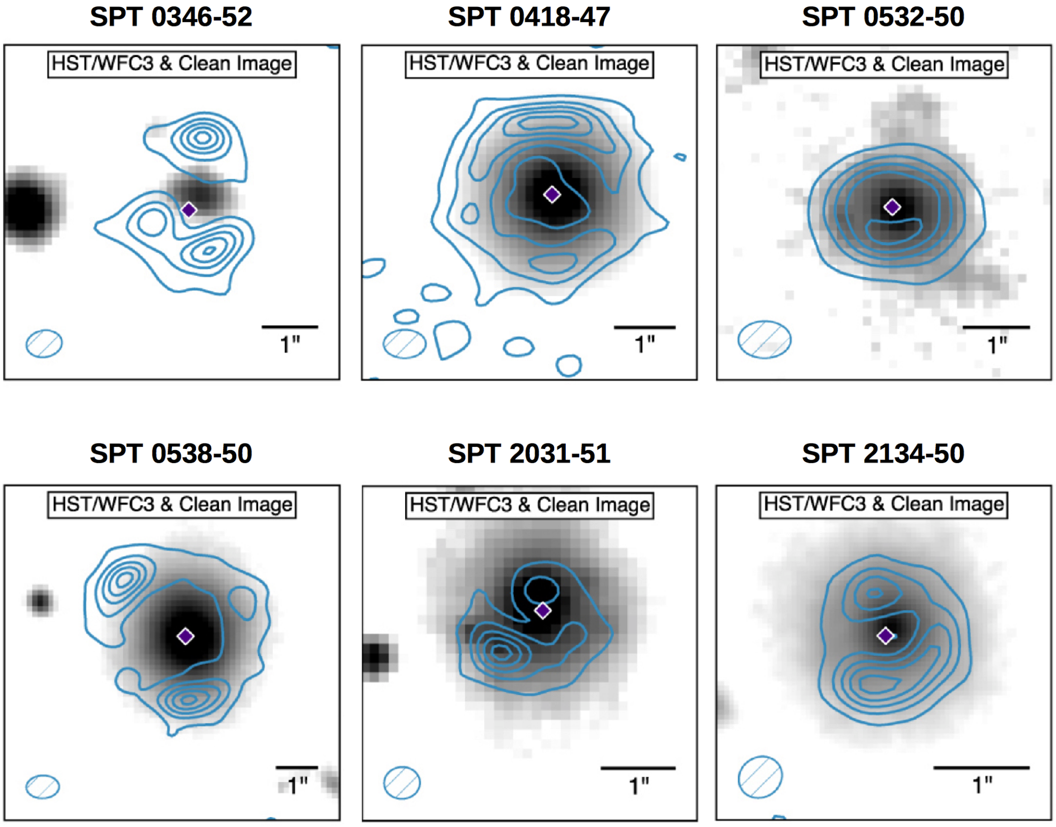

Figure 4.1: SPT lenses: overlay of the HST-WCF3 optical imaging (greyscale, F110W+F160W filters) and ALMA 870-μm continuum imaging (blue contours) based on the concatenated compact- and extended-array data. ALMA data were CLEANed using natural weighting. The contours are drawn at 10, 30, 50, 70 and 90% of the peak value. Synthesised beam size is shown in the lower left corner of each image. Image courtesy: Spilker et al., 2016.

account. The primary beam was found to be well-described by a circular Gaussian profile with a typical FWHM of ∼17 arcsec.

The root-mean-square (rms) noise for a given baseline was estimated directly from using the visibilities that were not flagged during the calibration process. As the signal on a given baseline is not expected to change significantly due to the short on-source time (60 or 90 seconds), and the noise on each baseline is much stronger than the signal, the rms was estimated from all samples for a given baseline, without subtracting the signal.

We now present the lens modelling results for each of the six lens systems individually. Results of the gravitational lens modelling for every lens, array configuration and corresponding source properties are provided in Tables 4.2 and 4.3, respectively. The best models for individual systems are presented in Figure 4.2 to 4.8.

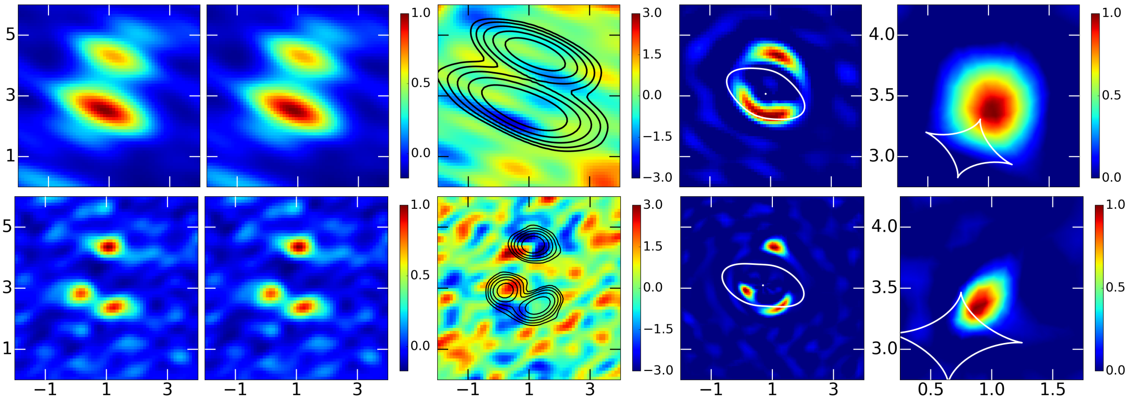

Figure 4.2: Results of modelling the SPT 0346-52: *Upper*: best compact-array model. *Lower*: best extended-array model. *From left to right*: The data dirty image, model dirty image, dirty image residuals (data-model) with CLEANed image contours, best sky model and best source model. Critical lines in the plane-of-the-sky and caustics in the source plane are indicated in white. The CLEANed images were produced using Briggs weighting with robust parameter = 0; contours start at 3-σ level and increase by a factor of 2. The two left-most columns are normalized to the data surface brightness maximum; the residuals are given in units of σ and the two columns on the right are normalized to the peak brightness.

SPT 0346-52

The extended-array lens model agrees with the one presented in HZ13 within errors. On the other hand, in the case of the compact-array observations, the highly elliptical synthesized beam reduces the lensed images to two elongated components. Consequently, our compact-array lens model is inconsistent with the extended-array one, as well as the one presented in HZ13. In particular, our best sky model does not recover the splitting of the bottom arc (Figure 4.2). By imposing the extended-array model on compact-array data and re-optimizing for the source regularization parameter λ_s, the splitting of the bottom arc is recovered. However, this model is heavily disfavoured by the evidence; the compact-array model presented in Figure 4.2 is preferred by $\Delta \log \mathcal{L} = 55$. Even restricting the slope of the mass density profile to an isothermal one (as in HZ13) does not recover the arc-splitting; the extended-array model is still disfavoured at $\Delta \log \mathcal{L} = 50$. We attribute the discrepancy between the compact-array lens models obtained with the pixellated and the parametric source models to two factors: (1) the very low resolution of the compact-array data, coarser than any of the mock datasets analysed in the previous section and (2) the structure of the source which is very close to a Gaussian profile and lends itself readily to parametric source modelling.

SPT 0418-47

Both compact- and extended-array lens model are found to be in agreement with each other as well as with the HZ13 and SP16 models. The compact-array source model consists of a bright core embedded in an extended region of fainter emission

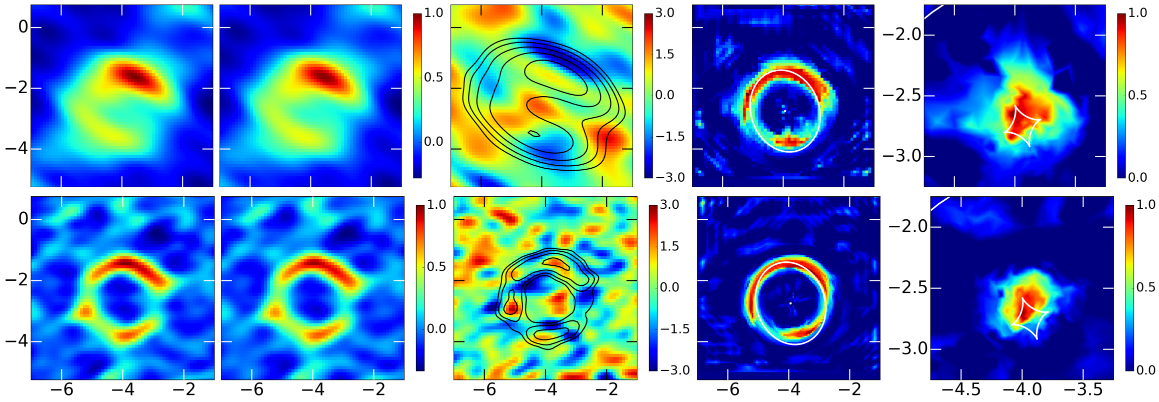

Figure 4.3: Results of modelling the SPT 0418-47 compact-array (*upper*) and extended-array (*lower*) datasets. Columns ordered as previously in Figure 4.2.

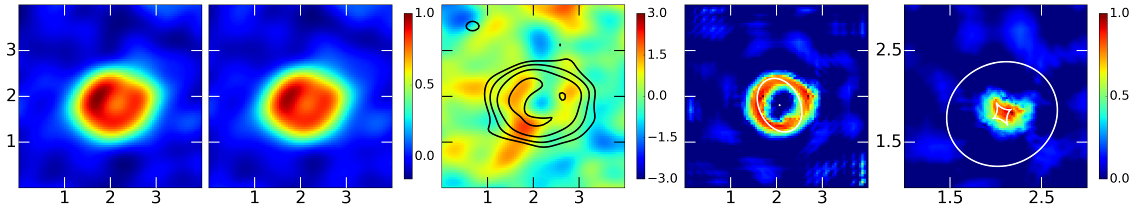

Figure 4.4: Results of modelling the SPT 0532-50, extended-array dataset. Columns ordered as previously in Figure 4.2.

(∼0.5 arcsec FWHM); its extent agrees with that of the parametric reconstructions in HZ13 and SP16. With the extended array, almost 60% of the compact-array flux density is resolved out. As a result, the faint emission in the source–plane is no longer recovered, leading to a 50% increase in the total magnification between the compact- and extended-array models.

SPT 0532-50

This system is unresolved in the compact-array observation due to its small size; with the extended-array, a very tight Einstein ring is resolved. Therefore, we model the extended-array data only. Compared to the SP16 model, we find the source to be asymmetric, causing the total magnification to be higher by about 50%.

The plane-of-the-sky residuals show an extended structure (at the $\sim$1.5-σ level) to the south-east that is not accounted for by the source model. This structure is not affected by our choice of a level of regularization, grid setup or noise estimation method. Given the quality of the data, it is unclear whether this is an extended region of faint emission or simply a noise feature.

SPT 0538-50

The compact-array observations of SPT 0538-50 shows two bright lensed images, as well as a fainter secondary image to the west (see Figure 4.6). Because of this double structure, HZ13 and SP16 found that a single source-plane component is not sufficient to fit the data. Instead, their models included a fainter extended component to the south-east. Using our pixellated technique, we found that without imposing a prior on the lens centre position, the maximum *a posteriori* model for both compact- and extended-array data (best-*uv*-fit model) is significantly different from the one obtained by HZ13 and SP16, with a lens centre offset from the SP16 position by 0.8 arcsec (Figure 4.5).

Namely, the lens configuration consists of two double imaged source-plane components. The brighter source is somewhat elongated, with the fainter companion offset by about 1 arcsec towards the north-west. This secondary component is detected at 2-σ level in the compact-array data. The flux ratio of the two source components is approximately 70:30 (compact-array model). The compact component is not detected with high enough significance in the extended-array data - as about 50% of the total flux density is resolved out, this is not particularly surprising. While a model analogous to that obtained by HZ13 and SP16 also fits the data, based on *uv*-data only, the best-*uv*-fit model is slightly preferred over the SP16 one by $\Delta \log \mathcal{L} = 3$ and 7 for the compact and extended arrays, respectively. After modelling the concatenated compact- and extended-array data, we find our model to be preferred with $\Delta \log \mathcal{L} = 18$.

However, the best-*uv*-fit model becomes heavily disfavoured once the imaging of the lensing galaxy is taken into account. HST-WFC3 imaging of the lensing galaxy (F110W and F160W filters; Vieira et al., 2013; Bothwell et al., 2013) reveals that the centre of the light distribution coincides with the lens mass centre as quoted in HZ13 and SP16. On the other hand, the centre of the light is offset by about 5 kpc from the best-*uv*-fit model lens centre. As it is unlikely that the centre of the light and the mass are more than 5 kpc distant from each other, our best-*uv*-fit model requires an unlikely error in the alignment of the HST and ALMA fields.

Specifically, a typical uncertainty in HST astrometry is 0.2-0.3 arcseconds; according to Dunlop et al. (2017) who studied the relative astrometry of the Hubble Ultra Deep Field and ALMA imaging, the HST astrometry shows an offset of <0.25 arcseconds from ALMA positions. We therefore conclude that the 0.8 arcsec shift in HST astrometry required to reconcile our best-*uv*-fit lens model with the optical imaging of the foreground galaxy would require dramatic systematic errors in the HST/ALMA alignment. By imposing a prior on the lens centre position based on the optical imaging, the *a posteriori* likelihood of the two models shifts dramatically: the best-*uv*-fit model is now disfavoured by $\Delta \log \mathcal{L} = -40$ and the maximum *a posteriori* model is in agreement with that presented in HZ13/SP16 (Figure 4.6).

Nevertheless, the pixellated source reconstruction shows significant deviations from the parametric source model of HZ13 and SP16. Namely, the faint component required by the parametric models is not present in our model, leading to a decrease in inferred compact-array magnification by 50%. Consequently, the flux-density of the pixellated source is a factor of two higher than that of parametric models presented in HZ13 and SP16.

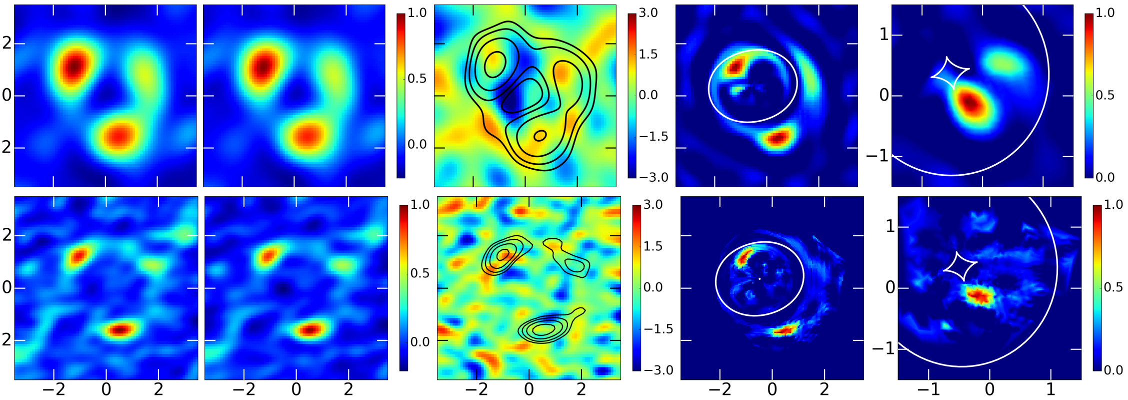

Figure 4.5: Results of modelling the SPT 0538-50, based on fitting the ALMA data only (best-*uv*-fit model). *Top*: compact-array dataset. *Bottom*: extended-array dataset. Columns ordered as previously in Figure 4.2. The source-plane reconstructions shows two distinct components. However, this model is heavily disfavoured by the likelihood once the HST imaging of the lensing galaxy is taken into account.

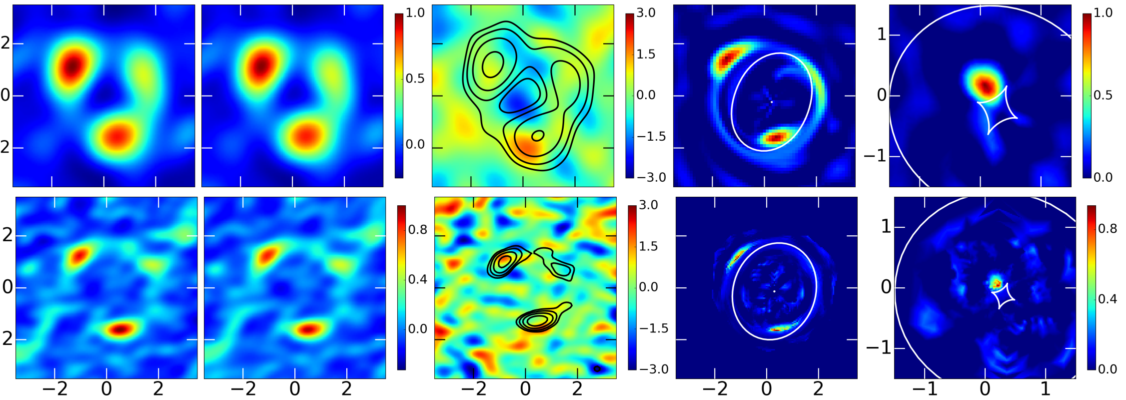

Figure 4.6: Results of modelling the SPT 0538-50, after imposing a prior on the lens centre position based on the HST imaging of the lensing galaxy. *Top*: compact-array dataset. *Bottom*: extended-array dataset. Columns ordered as previously in Figure 4.2.

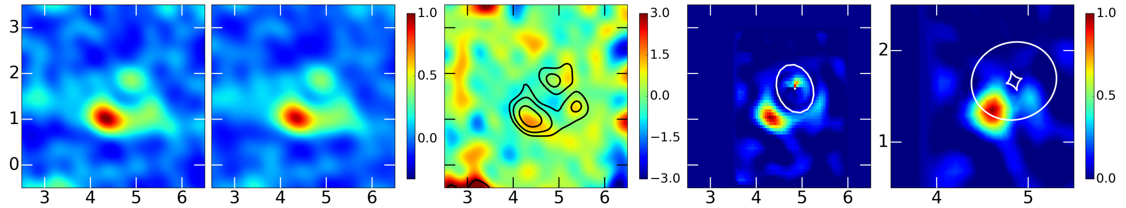

Figure 4.7: Results of modelling the SPT 2031-51, extended-array dataset. Columns ordered as previously in Figure 4.2.

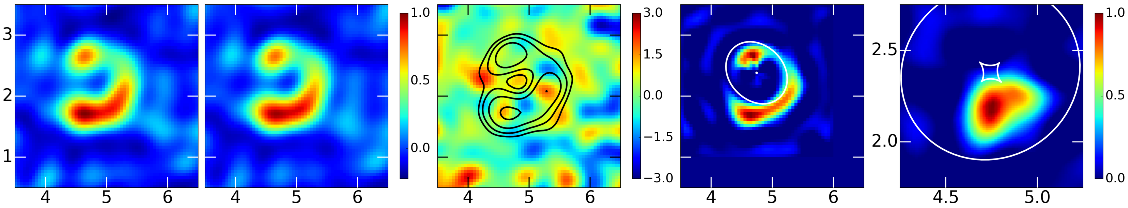

Figure 4.8: Results of modelling the SPT 2134-50, extended-array dataset. Columns ordered as previously in Figure 4.2.

SPT 2031-51

The source consists of two components with a flux ratio of 80:20. The fainter component, located to the west of the bright one, is detected at the 3-σ level. This is in contrast to the single Sérsic profile source model of SP16. Indeed, the CLEANed image of the shows non-symmetric structure in the counterimage (Figure 4.7). Consequently, our magnification estimate is reduced by a factor of 50% compared to SP16.

SPT 2134-50

Although the best-fit lens mass model agrees with that derived by SP16, our reconstructed source is more extended and shifted with respect to the caustics, reducing the global magnification estimate by a factor of ∼4. In our model, the source is double-imaged, rather than quadruply-imaged as presented in SP16. The source is asymmetric, with a FWHM axis ratio of almost 2:1 and the surface brightness peak is offset from the centre of the source; which corresponds to the structure of the main arc.

4.4 Results

4.4.1 Lens mass models

Considering the lens mass models, our best models agree with those presented in HZ13 and SP16 within the errors. For lenses for which both compact- and extended-array observations were modelled, we find lens mass models to be consistent between the two array configurations; the only exception is the compact-array observations of SPT 0346-52, for which the poor quality of the compact-array data introduces a large degeneracy in the lens mass model. Although the projected mass density profile slope γ was a free parameter of the model, all lens mass models are consistent, within the errors, with an isothermal profile ($\gamma = 2$).

4.4.2 Source properties

Compared to HZ13, all the source models presented in this paper exhibit significant deviations from circular symmetry, with a typical FWHM axis ratio of 1.5 to 2.0.

In the case of SPT 2031-51 and SPT 2134-50, our lens models agree with those derived by SP16; however, a different geometry and extent of the pixellated source compared to the Sérsic profile causes the magnification and flux density estimates to differ by a factor of 2 and 4, respectively. Figure 4.9 provides a visual comparison of the discrepancy in inferred source-plane flux density between this work and values reported in SP16. This discrepancy is largely driven by the structure of the pixellated sources, which show significant deviations from Gaussian or Sérsic profiles, revealing the complex nature of these objects even when observed at low resolution[1]. Although the synthesized beam of ALMA Cycle-0 observations is coarse, there are still baselines which sample the structure on smaller scales, and this is taken into account with our visibility fitting method. We stress that the complexity of the structure is due to the nature of the source itself, rather than due to the lens modelling technique; as shown in the discussion of the mock data models, our technique recovers Gaussian-like sources without introducing spurious structures.

For those systems for which we modelled both the compact- and extended-array data (SPT 0346-52, SPT 0418-47, SPT 0538-50), we find the total flux density measured with the extended array to be 40 to 70% lower than the flux density measured with the compact array, as some of the more diffuse emission is resolved out due to the lack of shorter spacings in the extended array. The compact-array fluxes are consistent with those measured with LABOCA at 870 μm (Greve et al., 2012; Weiß et al., 2013). As a result, the extended array sources are considerably more compact in all three cases. We therefore base our intrinsic flux density estimates S_{860} on the compact-array reconstructions (see Table 4.3).

---

[1] We note that the assumption of a simple power-law model for the mass density slope of the lensing galaxy can introduce artificial asymmetry into the source in case the actual mass density profile of the lensing galaxy deviates significantly from a simple power-law (M. Auger, private communication)

Finally, Figure 4.10 shows the comparison of the size distribution of the six SPT lenses presented in this chapter (as well as SDP.81 from Chapter 3) with that of non-lensed sub-mm galaxies from the Tacconi et al. (2008) sample; we find these two distributions to be in agreement. We make the comparison by fitting the reconstructed sources with an elliptical Gaussian profile; the average FWHM source size is 0.4±0.2 arcsec, which corresponds to a physical distance of about 3 kpc. Although the magnification of individual systems ranges between 2 to 12 for the compact-array reconstructions, there is no obvious trend towards smaller sizes of lensed sources compared to the Tacconi et al. (2006) sample (c.f. Hezaveh et al., 2013a); these findings are independently confirmed by SP16. However, the low resolution of the ALMA Cycle-0 arrays used here could lead to an overestimation of the source sizes, as shown by our analysis of the mock data (see Chapter 2).

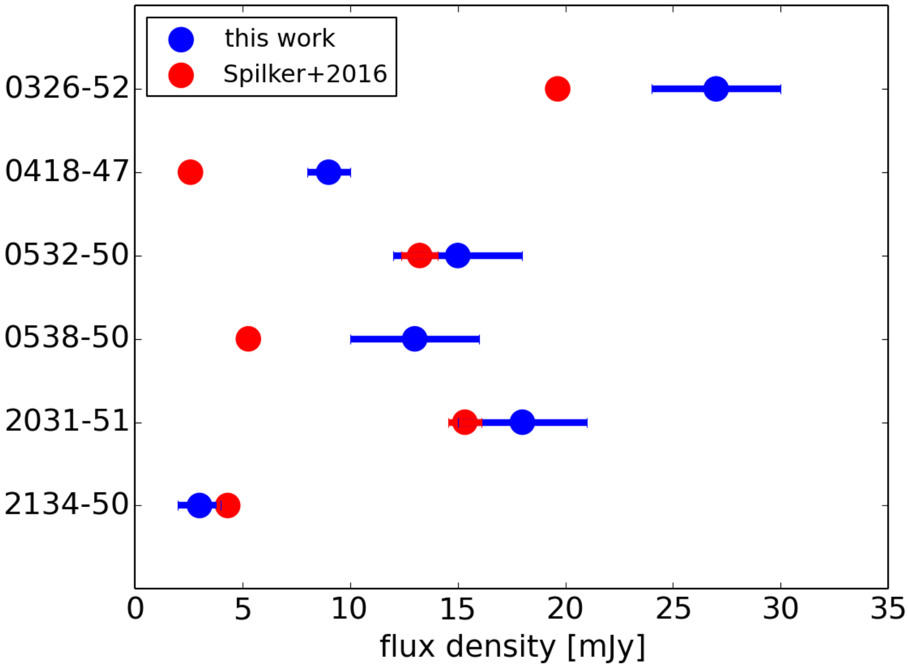

Figure 4.9: Comparison of inferred source-plane flux density for individual SPT lenses derived in this work (blue) and in Spilker et al. (2016, red). The error bars denote 1-σ uncertainties. The labels on vertical axis denote individual lenses.

Table 4.2: Maximum *a posteriori* lens mass model parameters with the lens redshift z_L. The lens centre coordinates ΔX, ΔY are specified with respect to the phase tracking centre.

Object	z_L	Array	κ_0 [arcsec]	q	θ [deg]	ΔX [arcsec]	ΔY [arcsec]	γ	Γ	Γ_θ [deg]
SPT 0346−52	–	Compact	1.04 ± 0.06	0.45 ± 0.04	63 ± 11	0.94 ± 0.08	3.06 ± 0.09	2.16 ± 0.20	0.12 ± 0.10	36 ± 17
		Ext.	0.96 ± 0.03	0.47 ± 0.05	70 ± 10	0.70 ± 0.03	3.09 ± 0.04	2.01 ± 0.14	0.11 ± 0.05	56 ± 22
SPT 0418−47	0.263	Compact	1.24 ± 0.05	0.77 ± 0.05	5 ± 10	-3.91 ± 0.17	-2.85 ± 0.19	2.01 ± 0.03	0.02 ± 0.01	32 ± 30
		Ext.	1.24 ± 0.02	0.83 ± 0.06	19 ± 2	-3.94 ± 0.01	-2.75 ± 0.04	2.00 ± 0.01	0.00 ± 0.01	25 ± 25
SPT 0532−50	1.15	Extended	0.51 ± 0.01	0.75 ± 0.04	26 ± 3	2.07 ± 0.01	1.81 ± 0.02	2.01 ± 0.02	0.02 ± 0.02	29 ± 2
SPT 0538−50	0.404	Compact	1.68 ± 0.01	0.80 ± 0.02	151 ± 3	0.28 ± 0.02	-0.24 ± 0.08	2.01 ± 0.01	0.05 ± 0.01	170 ± 4
		Ext.	1.74 ± 0.01	0.84 ± 0.03	159 ± 4	0.30 ± 0.03	-0.14 ± 0.04	2.01 ± 0.002	0.01 ± 0.01	11 ± 5
SPT 2031−51	0.624	Hybrid	0.46 ± 0.02	0.73 ± 0.05	19 ± 3	4.85 ± 0.02	1.67 ± 0.01	2.00 ± 0.01	0.01 ± 0.01	-13 ± 21
SPT 2134−50	0.776	Hybrid	0.50 ± 0.01	0.82 ± 0.03	36 ± 11	4.74 ± 0.02	2.38 ± 0.02	2.00 ± 0.01	0.03 ± 0.02	104 ± 21

Table 4.3: Magnification estimates and the intrinsic source properties. The individual columns list source redshift z_S, the compact- and extended-array magnifications μ_{comp} and μ_{ext} and the intrinsic flux-density at 860 μm, S_{860}. For objects with both compact- and extended-array models, S_{860} estimates are based on the compact-array model.

Object	z_S	μ_{comp}	μ_{ext}	S_{860} [mJy]
SPT 0346−52	5.656	7.2 ± 0.9	4.9 ± 0.3	27 ± 3
SPT 0418−47	4.225	11.5 ± 1.1	16.1 ± 2.4	9 ± 1
SPT 0532−50	3.399	–	8.6 ± 2.5	15 ± 3
SPT 0538−50	2.782	9.4 ± 2.0	16.7 ± 7.6	25 ± 5
SPT 2031−51	–	–	2.2 ± 0.4	18 ± 3
SPT 2134−50	2.780	–	4.3 ± 0.4	19 ± 1

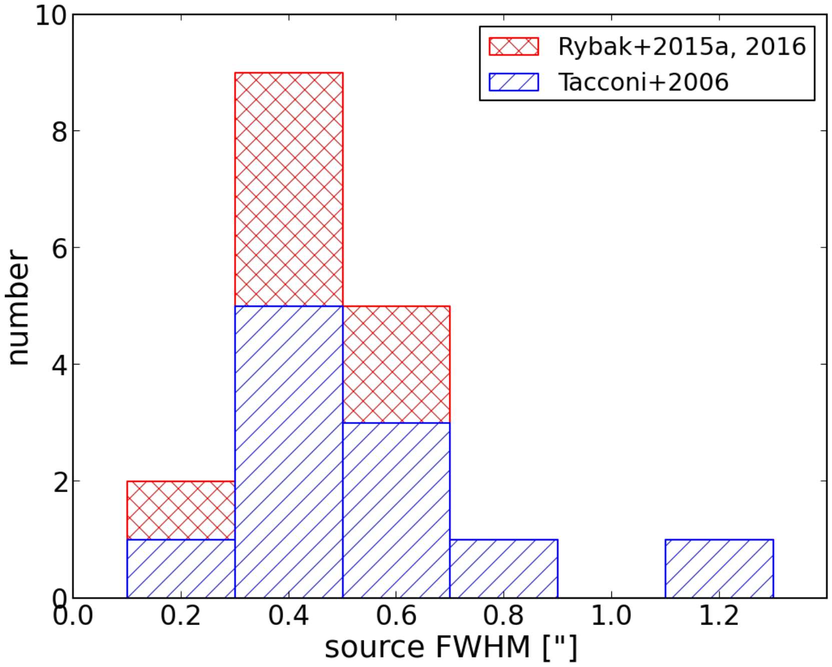

Figure 4.10: Comparison of the source size distribution of the six SPT lenses and SDP.81 (see Chapter 3), compared to the non-lensed sources from the Tacconi et al. (2008) sample. We do not find any obvious bias toward a smaller source size for the lensed objects.

4.5 Conclusions

In this chapter, we applied our method to six strong gravitational lenses observed with three ALMA Cycle-0 array configurations, which were previously modelled with analytic source models (Hezaveh et al., 2013b; Spilker et al., 2016). These six lenses are representative of low-resolution snapshot observations carried out in ALMA Cycles 0 and 1 and with the SMA as a follow-up of lens candidates identified in the H-ATLAS and SPT wide-field surveys. Compared to the high-resolution observations of SDP.81 (Chapter 3), these observations have a much lower angular resolution, a very sparse uv-plane coverage and a lower SNR.

We reproduced the lens mass models derived by Hezaveh et al. (2013b) and Spilker et al. (2016). Note that different source models (circular Gaussian / elliptical Sérsic / pixellated) are expected to result in a systematic scatter in inferred lens models (Spilker et al., 2016). The only exception is the compact-array observation of SPT 0346-52; the extremely poor resolution of this dataset, combined with the extra freedom of the pixellated model, results in large degeneracies that prevent us from recovering the correct lens model. For the higher-resolution extended-array observation, we recover the previously published lens model within errors. We note that the compact-array data, Hezaveh et al. (2013b) find the correct lens model to be only marginally preferred over alternative ones (Y. Hezaveh, private communication). Furthermore, in the case of SPT 0538-50, the additional freedom of the pixellated source model finds two distinct lens/source model; it is only by taking the optical imaging of the foreground galaxy into account that we are able to break this degeneracy.

In spite of the general agreement in lens model parameters, in three out of the six lenses studied in this paper (SPT 0538-50, SPT 2031-51 and SPT 2134-50), there are considerable differences between the pixellated models presented here, and those presented by Hezaveh et al. (2013b) and Spilker et al. (2016), with magnifications and inferred intrinsic source flux-densities deviating by a factor of a few. This discrepancy is caused by the complex structure of the background sources that - even at 0.5 arcsec resolution - deviates significantly from the symmetric Gaussian or elliptical Sérsic profiles used in previous studies. While this discrepancy can be attributed partially to the systematic scatter between parametric/pixellated source models, we see evidence for a more complex structure of the source that results in dramatically different magnifications.

The large scatter in the inferred magnification between the pixellated and parametric models, together with the bias towards brighter sources seen in the analysis of the mock ALMA data in Chapter 2, reveals the limitations in inferring source properties from low-resolution data. Therefore, although dedicated campaigns with the ALMA Cycles 0 and 1 arrays and the SMA provided a large sample of more than a hundred gravitationally lensed sub-mm galaxies observed at low resolution (Negrello et al., 2010; Vieira et al., 2013), the inferred intrinsic source sizes and flux densities might suffer from a large systematic error by up to a factor of few, and have to be taken with caution.

Chapter 5

RXJ1131-1231: dissecting a low-redshift quasar and its host

Abstract

This chapter presents the results of a sub-kpc analysis of a strongly gravitationally lensed quasar RXJ1131-1231 ($z_s = 0.658$) observed with ALMA Band 4 in CO (2-1) emission lines and 2.1-mm continuum. In particular, we investigate the flux-ratio anomaly observed in the 2.1-mm (ALMA Band 4) continuum, and the morphology and kinematics of the molecular gas as traced by the CO (2-1) line. The continuum emission is found to be compact and co-incident with the quasar detected at the optical and X-ray wavelengths, whereas the molecular gas emission forms a complete, thick Einstein ring. Applying our lens-modelling technique to the CO (2-1) data, we reveal a large disk of molecular gas with a diameter of ∼15 kpc, at an inclination of 54 deg. Fitting a dynamical model to the rotation curve, we find a dynamical mass enclosed within 5 kpc radius $M(R < 5\,\mathrm{kpc}) = (1.46 \pm 0.31) \times 10^{11}\,\mathrm{M}_{\odot}$. The molecular gas distribution is found to be highly structured, with clumpy structures that are co-incident with high velocity-dispersion regions, indicating a turbulent interstellar medium. The peak of the CO (2-1) emission is offset from the AGN by ∼2 kpc, indicative of a feedback from the central engine. We detect the multiple images of the AGN in the continuum, as well as extended emission from both the AGN host and the lensing galaxy. The flux ratio of the AGN images in the continuum are vastly different from the model predictions and from the optical flux ratios; we attribute this to micro-lensing by stars in the lensing galaxy. This flux-ratio anomaly, combined with the excess 2.1-mm continuum emission is consistent with an optically thick synchrotron emission from a compact region (∼0.01 pc) from the AGN close to the central black hole, although we find evidence for an extended continuum emission. However, high-resolution follow-up observations are required to confirm the presence of this extended emission.

Note: This chapter is based on D. Paraficz, M. Rybak, et al.: "Resolved ALMA view of the lensed starburst RXJ1131-1231", in prep.

5.1 Introduction

RXJ1131-1231 (RA = 11:31:51.50, DEC = -12.31.58.70) is an exceptional strongly lensed quasar. It consists of four images of a $z_s = 0.654$ quasar (Figure 5.1) with a total magnification of ∼50 (Sluse et al., 2007). The source quasar is hosted by a luminous late-type galaxy lensed by a foreground early-type galaxy at $z_l = 0.295$ (Sluse et al., 2003, 2007). It is lensed by a foreground late-type galaxy located at a redshift $z_L = 0.295$. As RXJ1131-1231 "*brings together rare properties (i.e. quad, bright optical Einstein ring, small redshift, high amplification), nearly unique among the known gravitational lens systems*" (Sluse et al., 2003), it is one of the most studied lensed quasars, with a wealth of observations on a broad range of wavelengths at many epochs. Previous observations include Chandra X-ray monitoring for a microlensing analysis (Sluse et al., 2006; Dai et al., 2010; Chartas et al., 2012; Reis et al., 2014); Integral Field Spectroscopy for the characterization of lens substructures (Sugai et al., 2007); and X-ray observations to estimate the black-hole spin (Reis et al., 2014). In addition, RXJ1131-1231 is being photometrically monitored at optical wavelengths by COSMOGRAIL (Tewes et al., 2013) in order to improve the time delay measurement. The COSMOGRAIL monitoring and time delays, together with detailed lens modeling based on HST imaging and the lens dynamics Suyu et al. (2014), make RXJ1131-1231 one of the most accurate cosmological probes available (Suyu et al., 2013; Birrer et al., 2016; Chen et al., 2016; Bonvin et al., 2016).

This works adds another dimension to the datasets available for RXJ1131-1231 in form of ALMA Band 4 (125-163 GHz) observations in an extended configuration. These observations aim at resolving the CO (2-1) emission line in the lensed host galaxy in order to map the line intensity and velocity across the Einstein ring. In addition, we use our ALMA data in the continuum to measure the flux ratios between the lensed images of RXJ1131-1231 and compare these with the flux ratios at different wavelengths.

We note that Leung et al. (2017) have recently presented an analysis of CO (3-2) and (2-1) lines based on few-arcsec resolution observations with the Combined Array for Research in Millimeter-wave Astronomy (CARMA) and the Plateau de Bure Interferometer (PdBI); we provide a comparison between our analysis and theirs in Section 5.5.

5.2 Observations

5.2.1 ALMA Band 4 observations

RXJ1131-1231 was observed during ALMA Cycle 2 campaign on 2015 July 19 (Proposal code: 2013.1.01207.S; PI: D. Paraficz). The purpose of the observations was to spatially resolve the redshifted CO (2-1) emission line (rest-frame frequency f_0=230.538 GHz) of the background source.

The array configuration consisted of 37 of the 12-metre antennas with projected baselines ranging from 27.5 to 1600 m, offering a sensitivity to structures on scales between 16.0 and 0.25 arcsec. The synthesised beam size is 0.28”×0.24 arcsec (Briggs weighting, robust parameter = 0). The total on-source time of 1.3 hours was split into two observing blocks: UCT 19:46 - 20:39 and 22:38 - 23:21, with an integration time of 6.05 seconds; total duration of the observations including calibrator observations was about 2.5 hours.

The data were taken in Band 4 (125 - 163 GHz) using four separate Spectral Windows (SPWs), with central frequencies of 150.993, 149.118, 139.044 and 137.169 GHz, respectively. SPW 0 was used in spectral line mode, centered on the redshifted CO (2-1) line at 139.044 GHz, with 480 spectral channels and a total bandwidth of 1.875 GHz. This provided a velocity resolution of 8.4 $\mathrm{km\,s^{-1}}$. The three remaining SPWs were used in the standard continuum mode, each with 128 spectral channels and a 2 GHz bandwidth tuned to central observing frequencies of 137.169, 149.118 and 150.993 GHz. The data were taken using both linear polarisations (XX and YY).

Titan and Ganymede were used to determine the absolute flux-density calibration, and the unresolved quasars J1058+0133 and J1118-1232 were used for calibrating the bandpass and correcting for the residual delays, respectively. The observations were phase-referenced using the J1130-1449 quasar, with a cycle time of 4.5 to 6 min. The data were reduced and calibrated using the standard ALMA pipeline using the Common Astronomy Software Application package Casa (McMullin et al., 2007). The corrected visibilities for the target and calibrators that were produced by the pipeline reduction were found to have no major calibration errors. Self-calibration was not used as there was an insufficient signal-to-noise ratio on the continuum emission to provide good calibration solutions.

5.2.2 Hubble Space Telescope

HST Advanced Camera for Surveys (ACS) images were obtained in two filters, F814W and F555W (GO 9744; PI: Kochanek) with a total exposure time of 1980 seconds in each filter. For our analysis, we used the post-processed images presented in Suyu et al. (2013).

5.2.3 Very Large Array

RXJ1131-1231 was also observed with the VLA at 4.86 GHz on 2008 December 29 (Proposal Code: AW741; PI: Wucknitz - of the LensClean fame, see Chapter 2). The VLA was used in the most extended A-configuration, which provided baselines between about 0.5 and 33.5 km. In total, 25 of the VLA antennas were available. The observations were phased referenced using the J1130-148 quasar, which was used to determine the complex gains (amplitude and phase) as a function of time, and 3C286 was used to determine the absolute flux-density scale. The visibility integration time was 10 seconds and two spectral windows with 50 MHz bandwidth were used. The data were calibrated in the standard way with the Casa package. Given the expected brightness of the target, and the likelihood that sidelobes from field sources would limit the dynamic range, wide-field imaging using a natural weighting scheme of the visibilities was used. This resulted in a beam size of 0.74 × 0.38 arcsec and an rms noise of 14 μJy beam^{-1}.

5.2.4 Herschel Space Observatory

Finally, RXJ1131-1231 was observed at the far-infrared wavelengths using the *Herschel Space Observatory* on 2012 December 08 (Proposal Code: OT1_abercian_1; PI: Berciana-Alba) as part of a study of over a hundred lensed quasars (Stacey et al., 2017). The target was observed at 250, 350 and 500 microns for 90 seconds each using the Spectral and Photometric Imaging Receiver (SPIRE) in the *smallmap* mode. The data were retrieved from the Herschel archive and processed within the Herschel Interactive Processing Environment (HIPE 14.2.1). The flux density of the target was measured at the three far-infrared wavelengths using the time-line extractor.

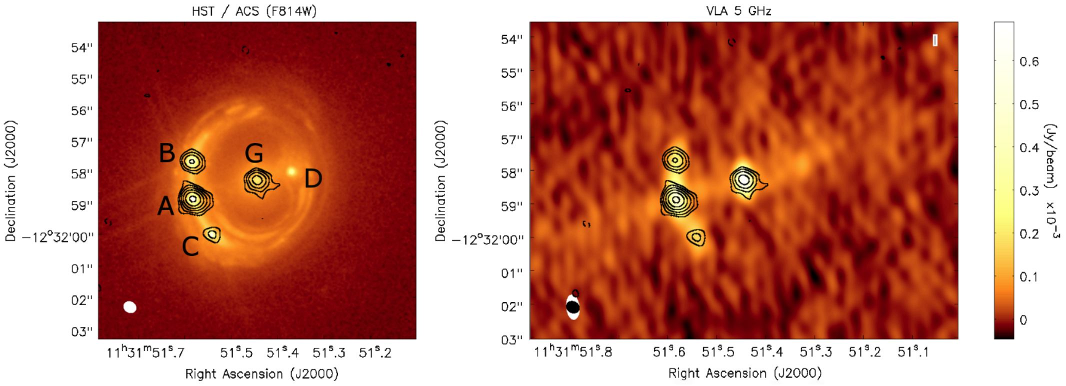

Figure 5.1: *Left:* HST/ACS image of RXJ1131-1231 in the F814W filter, with individual quasar images A, B, C, D and foreground galaxy G indicated. *Right:* VLA image at 5 GHz. In each panel, the ALMA 2.1 mm continuum contours at the (−3, 3, 6, 12, 24, 48, 96) × 10 μJy beam^{-1} levels are overlaid for reference.

5.3 Image-plane analysis

5.3.1 Imaging

The imaging of the continuum was done using all four SPWs, excluding those channels the contained the emission line; this resulted in a central frequency of 144.081 GHz. Images were made using Briggs weighting with a robust parameter of 0 and 2 (2 being equivalent to natural weighting), which resulted in a beam size of 0.28×0.24 and 0.41×0.34 arcsec, respectively, and an rms map noise of 15 and 10 μJy beam^{-1}, respectively.

The spectral-line cube was formed using natural weighting and all 480 spectral channels within the spectral window that contained the emission line. These data were CLEANed using a mask and a detection threshold of 3-σ to limit any contribution from noise peaks within the cube. The resulting beam size is 0.44×0.36 arcsec with an rms map noise of about 340 μJy beam^{-1} for a 8.4 km s^{-1} channel width.

From the CLEANed spectral-line cube, a velocity integrated map (moment-zero) of the line emission was made using those channels covering the line, and intensity-weighted velocity (moment-one) and velocity dispersion (moment-two) maps were made using those pixels with SNR > 3 (Figure 5.3).

5.3.2 Continuum emission

The main goal of this work is to map the CO (2-1) emission of the gas in the spiral quasar host galaxy. However, we also detected continuum emission both from the triplet images A/B/C of the background source and the lensing galaxy itself. Note that the fourth lensed image (D) remains undetected in the continuum.

The total flux density, spatially integrated over the three quasar images emitting in the continuum at 2.1 mm is 2.41 mJy beam^{-1} (Table 5.1). For comparison, Figure 5.1 shows the overlay of Band 4 continuum and F814W HST/ACS image.

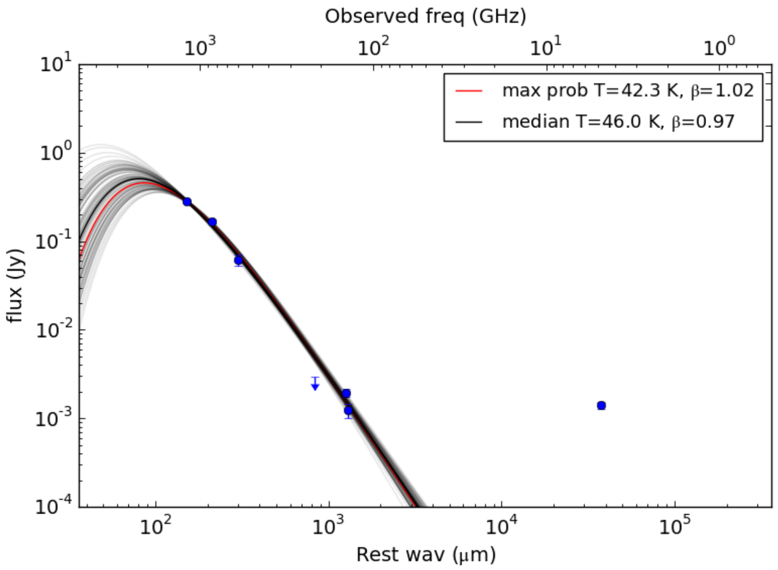

Figure 5.2: Observed SED of RXJ1131-1121 with the multi-wavelength photometry of the integrated 3 lensed images A, B and C not corrected for gravitational lensing (blue points with error bars). *Left to right*: *Herschel* 250, 350 and 500 μm, CARMA 1.4 mm, ALMA 2.1 mm, PdBI 2.1 mm and VLA 6 cm observations. The maximum a posteriori model (red line) yields an effective dust temperature of $T_{dust} = 42.3$ K and emissivity β=1.02; the median values are $T_{dust} = 46.0$ K and β=0.97. Figure and SED models courtesy of Hannah Stacey.

Figure 5.2 presents the spectral energy distribution (SED) fit to the multi-wavelength photometry of the lensed quasar. The ALMA and VLA measurement show a significant excess of emission at rest-frame wavelengths over 1 mm, compared to expectation from a pure dust emission SED. Namely, the dust accounts for between 10 and 50 % of the ALMA continuum flux; its contribution to the VLA continuum is negligible. This suggests that the Band 4 continuum emission arises from two distinct regions: an extended dust component, as well as a compact region associated with synchrotron emission, possibly offset with respect to each other.

We attribute the excess VLA 4.86 GHz continuum to synchrotron emission from star-forming regions in the lensed galaxy. An extended ($\sim$4 arcseconds) Einstein arc surrounding the triplet images confirms the extended nature of the 4.86 GHz continuum emission. Moreover, the e-VLBI observations at 18 cm (Wucknitz & Volino, 2008) did not detect any compact structure in the lensed images $\sim$100 μJy level; this indicated that the radio-emission in the lensed galaxy. This extended emission is associated with supernova shocks from massive (M>8 $M_{\odot}$) stars with a typical lifetime of 10^7 yr, indicating a recent or on-going star-forming activity (Condon, 1992).

Table 5.1: ALMA continuum peak flux and magnifications for the 3 detected quasar images (A, B and C) and for the lensing galaxy (G). The positions of individual images are given with respect to the brightest A image; the position of the A image is given in standard J2000 coordinates. For the D image , we give an upper limit for the non-detection at the 3-σ level.

ID	RA	DEC	$S_\nu pk$ [Jy beam^{-1}]	μ (Suyu et al., 2013)
A	$11^h31^m51.582$	-12:31:58.88	1.1×10^{-3}	-21.29
B	$+0.029\pm0.006$	$+1.196\pm0.006$	2.6×10^{-3}	14.43
C	-0.615 ± 0.018	-1.113 ± 0.012	7.5×10^{-3}	11.29
D	-	-	$< 3.0 \times 10^{-5}$	-1.19
G	-2.035 ± 0.007	$+0.604\pm0.007$	3.4×10^{-4}	-

Table 5.2: RXJ1131-1231: Flux ratio of quasar images as observed at optical, mm- and radio wavelengths. All wavelengths listed are given as observed.

Observation	Measurement date	A/B	A/C	B/C	A/D	R_{cusp}	Reference
Suyu+2013 model		1.475	1.885	0.782	17.933	0.095	Suyu et al., 2013
ALMA 2.1 mm	2015 Jul 19	4.16 ± 0.21	13.95 ± 1.74	0.29 ± 0.04	>0.46	0.519 ± 0.040	this work
Optical (average)	2004 - 2014	1.55 ± 0.95	4.02 ± 2.37	0.38 ± 0.17	10.37 ± 8.00	0.056 ± 0.235	Tewes et al., 2013
800 nm cont.	2005 Feb 8	1.33 ± 0.05	3.24 ± 0.20	0.41 ± 0.02	9.21 ± 0.75	0.029 ± 0.015	Tewes et al., 2013
[O III] 820 nm		$1.63^{+0.04}_{-0.02}$	$1.63^{+0.04}_{-0.02}$	$1.19^{+0.03}_{-0.12}$	-	0.146 ± 0.018	Sugai et al., 2007
Hβ 806 nm		$1.74^{+0.07}_{-0.12}$	$3.78^{+0.04}_{-0.06}$	$0.46^{+0.02}_{-0.03}$		0.087 ± 0.023	
800 nm cont.	2015 Jul 19	1.33 ± 0.1	3.06 ± 0.1	0.42 ± 0.01	39.8 ± 1.0	0.038 ± 0.026	Tewes et al., 2013
500 nm cont.	2003 Apr 26	1.82 ± 0.07	1.77 ± 0.09	0.31 ± 0.01	-	0.054 ± 0.015	Sluse et al., 2007
[O III] 820 nm		1.97 ± 0.02	1.48 ± 0.02	1.33 ± 0.02		0.083 ± 0.004	
Hβ 806 nm		1.61 ± 0.02	3.76 ± 0.09	0.43 ± 0.01		0.060 ± 0.005	

2.1 mm continuum triplet flux ratios

RXJ 1131-1231 shows three highly-magnified images of the lensed quasar, A, B and C. This configuration arises because the source quasar lies close to, and inside, a cusp point of the caustic. For a system in such a configuration, a simple but asymptotic *cusp relation* between the magnifications of the lensed images can be derived (Mao & Schneider, 1998):

$$R_{cusp} = \frac{\mu_A + \mu_B + \mu_C}{|\mu_A| + |\mu_B| + |\mu_C|} \rightarrow 0, \tag{5.1}$$

The power-law lens mass model by Suyu et al. (2013) predicts flux ratios of A/B=1.475, A/C=1.885, C/B=0.782 and A/D=17.933 and a cusp relation value of $R_{\rm cusp}^{model} = 0.094$. However, the observed flux ratios deviate significantly from these predictions especially at 2.1 mm, where $R_{\rm cusp}^{mm} \sim 0.52$ (see Table 5.2).

A violation of the cusp relation can be attributed to several different reasons. Gravitational effects include additional light deflection by moving stars in the lensing galaxy, producing microlensing events or massive substructures in the lens or along the line-of-sight (Mao & Schneider, 1998; Xu et al., 2015), producing the so-called *flux-ratio anomalies* (Sluse et al., 2013; Vegetti et al., 2014). Violation of the cusp relation may also stem from astrophysical effects such as extinction by dust or electron scattering by the ISM. Although lensing itself is achromatic, all the above effects can result in wavelength-dependent flux ratios as emission at different wavelengths traces different parts of the source. Differential extinction caused by dust in the lensing galaxy can be significant in the UV/blue/optical range (typically up to 0.1-0.3 magnitudes) but is negligible at mm- and radio wavelengths. However, scattering by electron in the ISM may blur the lensed images and modify their flux ratios at radio wavelengths from a few hundred MHz to a few GHz. Note that the time delay between the triplet images does not play a role in the flux-ratio anomalies seen in the three brightest images of RXJ 1131-1231, as the time delays in this triplet are close to 0 (Tewes et al., 2013). Image D would be affected by a time delay, which is ~91 days (Tewes et al., 2013), but is undetected in our ALMA continuum observations.

The wavelength range covered by the observations in Table 5.2, allows us to safely exclude a significant impact of dust extinction and electron scattering on the flux ratios. In particular, this excludes dust extinction as a source of the flux-ratio anomaly: the rest-frame 1.3 mm continuum will not be significantly affected by either dust extinction or electron scattering.

A more likely explanation for a drastic change in the observed flux ratio with wavelength is a perturbation to the macro lens model caused by microlensing and/or milli-lensing. The micro- and milli-lensing will cause significant perturbation only for source-plane regions with size comparable to that of the Einstein radius of the perturber. In particular, microlensing by stars magnifies the source on scales smaller than 10^{-2} pc, while the more massive substructures of $M \sim 10^6 - 10^8$ M$_\odot$ magnify source-plane regions 10-100 pc across.

What is a typical size of regions associated with different emission in a representative AGN? Starting from the central black hole, the optical emission arises from the accretion disc with a size $r_{acc} \sim 10^{-4}$ pc. Further out are the broad line region (BLR, $r_{\mathrm{BLR}} \sim 10^{-3} - 10^{-2}$ pc) associated with the Hβ emission line; followed by the narrow-line region ($r_{\mathrm{NLR}} \geqslant 100$ pc) associated with the [OIII] emission line. Finally, the mm-wave dust emission should arise from a torus located between the BLR and NLR with a typical scale $r_{dust} \sim 1 - 10$ pc (e.g. Jaffe et al., 2004; Burtscher et al., 2013).

The evidence that the [OIII] emission is partially resolved (Sugai et al., 2007) while the mm-wave emission is not, confirms that the latter emission originates from a region smaller than the narrow-line region (100 pc). In addition, while the [OIII] flux ratios can be considered as free of microlensing, they still deviate by several tens of percents from the macro-model. This is likely caused by the extended and asymmetric nature of the NLR emission (Sugai et al., 2007). Finally, aperture effects yield differences between flux ratios of Sugai et al. (2007) and Sluse et al. (2013). With these considerations in mind, we now try to interpret the individual flux ratios reported in Table 5.2.

The A/B flux ratio

In the optical, the flux ratio between the images A and B is consistent within the errors with the flux ratio expected from the smooth lens model of Suyu et al. (2013), fluctuating over the last 10 years around the predicted value of A/B=1.9 (Table 5.2). In the ALMA continuum, however, A/B∼4 is very difficult to reconcile with the microlensing hypothesis as the dust emission is thought to be much larger than the optical continuum region. However, there is a possibility for a larger source size to be more microlensed than a more compact source: an extended source will cross a micro-caustic before/after a more compact disc emission, such that for some time, the largest source is observed to be more magnified. For this effect to reach a factor of a few, as needed to explain the A/B ratio at optical and mm-wavelengths, the more extended emission (i.e. mm-wave emission) still needs to be not much larger than the microlens Einstein radius η_0, and it must have a large fraction of its flux offset with respect to the optical continuum source.

To quantify the probability of microlensing causing such a dramatic perturbation, we use a toy microlensing simulation. In this simulation, we convolve a microlensing pattern obtained for the image A with a model for the optical and continuum source, respectively. In doing so, we assume that 30 % of the surface mass density around the image A is contributed by stars (Dai et al., 2010). Namely, we use an optical continuum source following a Gaussian profile with FWHM=$0.1\times\eta_0$, and a continuum source in form of a ring with an inner radius R_{in} and an outer radius $R_{out} = 1.2 \times R_{in}$. This dust model was based on the mid-IR dust emission model of Sluse & Tewes (2014). Multiple source realizations with varying R_{in}/η_0 ratio were produced. Based on these convolved patterns, we extract the distribution of single-epoch microlensing magnification of image A.

Although this model is rather simplistic, the distribution of micro-magnification depends little on the exact morphology of the source provided its half-light radius is conserved (Mortonson et al., 2005). This means that our model should also be representative of microlensing of any light emission arising from the base of the jet, even though the details of the jet configuration are not captured by the model. For example, replacing the uniform dust ring with clumpy components with size of $0.01 \times \eta_0$ has a negligible impact on our results.

We find that $R_{in} \leq 0.01$ pc is needed for the emission to have a micro-lensing magnification of 2; this provides the upper limit on the size of the Band 4 continuum emission region.

Even for a source this compact, a predicted micro-lensing magnification of a factor up to 2 is not enough to explain the difference of A/B flux ratio between the (∼4) and the optical continuum (∼1.5). A simultaneous de-magnification of the optical region is also needed. Although such large magnification of the emission are not systematically associated with de-magnification of the optical region, there is a possibility that the component is magnified by up to a factor $\mu_{\rm dust} \sim 2.4$, while the continuum is de-magnified (i.e. $\mu_{\rm cont} < 1$); this could be observed if the source is located in a dense area of the caustic network. This situation happens more frequently if the size of the mm-emission region is small (i.e. $R_{\rm in} \sim 0.005\,{\rm pc}$), but still sufficiently large compared to the optical continuum (10^{-4} pc), putting a tentative lower boundary on the size of the mm-wave emission region (10^{-4} pc $< r_{\rm 2.1mm} <$ 0.01 pc).

To summarize, the A/B flux ratio measured in the ALMA 2.1 mm continuum is compatible with the microlensing scenario, but this scenario requires a very compact emission (i.e. $R_{\rm 2.1mm} \leq \eta_0$) and a very specific location of the source with respect to the micro-lens caustic network. To confirm or exclude the micro-lensing scenario, a follow-up observation over a fraction of time required for the source to shift by η_0 - about 10 years - is required.

The C/B flux ratio

The other ratio of interest is C/B, which is ∼ 0.4 at optical wavelengths and in Hβ and ∼0.29 in 2.1-mm continuum as observed with ALMA. All these values are significantly smaller that the predicted ratio of ∼0.78 and suggest a de–magnification of the image C consistent with milli-lensing by a substructure massive enough to de-magnify compact regions, while leaving extended regions i.e. [OIII] almost unchanged. Such a speculative possibility would require a substructure with an Einstein radius $\eta_0^{\rm sub} < 10$ pc, if perfectly aligned with image C, giving a mass $M_{sub} < 10^6 M_\odot$ (under the assumption of a point-like lens). While a more detailed investigation of the milli-lensing scenario is required to explain the different value of C/B between and other wavelengths, microlensing ($\mu \leq 1.3$) of the image B may also be a possibility.

Optical offset

In the above discussion, we have neglected the impact of the source position with respect to the macro-caustics on the flux ratios. As explained above, it is likely that the mm-wave continuum and optical emission do not share the same centroid. Any small centroid shift or any asymmetry in the emission regions may impact not only the flux of the lensed images, but also their positions. While the positions in the optical images are measured with milli-arcsec accuracy, the positions measured with our ALMA data (Table 5.2) have a precision that is typically four times poorer. Comparing the ALMA imaging with the centroid positions presented in Suyu et al. (2013), the maximum plausible shift between the optical and positions of the triplet images A, B, C are 10, 10 and 50 mas, respectively. Using the lens model of Suyu et al. (2013), we find the maximum effect of such a shift in position to be about 20 % for C/B. This is too low to explain the change in flux-ratios between the optical and mm-wave continuum.

Such a shift translates in the source plane into a maximum offset between the continuum and emission of ∼5 mas, (35 pc) using the power law lens model derived by Suyu et al. (2013). The impact of such a shift on the image flux ratios remains small, with a maximum change of 20 % in C/B ratio.

To summarize, it is hard to provide a single simple interpretation for the observed ALMA flux ratios with respect to other wavelengths. Based on this preliminary analysis, the most plausible explanation requires a combination of micro- and milli-lensing, and a very compact emission region of the order of the size of the BLR or smaller. This would mean that a large fraction of the emission is not produced by dust as the dust sublimation radius is larger than the BLR radius; this interpretation is supported by the excess flux-density in ALMA continuum (Figure 5.2). On the other hand, ALMA imaging suggests that although most of the continuum flux arises from an unresolved compact component, a small fraction of the flux is extended. We therefore conclude that the continuum emission arises from two regions: a compact one ($\sim 10^{-3}$ pc) being associated with synchrotron emission taking place in the vicinity of the accretion disc (but plausibly offset from it), and a more extended dust component.

5.3.3 CO (2-1) emission line

The prominent Einstein ring formed by the lensed quasar host galaxy is well detected in CO (2-1) in our ALMA observations. Figure 5.3 presents the image-plane velocity moments of the CO (2-1) line, calculating using CASA's `immoments` task and using a natural weighting of the visibilities. The brightness of the CO (2-1) line changes dramatically along the arc, with two prominent peaks in the south-east (co-incident with the C image) and in the north. As these regions correspond to lower-magnification values, this image-plane offset between the continuum and the peak of the line intensity indicates that the AGN is significantly offset from the brightest part of the gas in the source-plane as well.

The velocity map reveals a clear velocity gradient, from about -300 to +400 km s^{-1}, using the radio definition of the velocity. Note that the intensity peaks of the CO (2-1) line correspond to the red end of the velocity map (+300 to +400 km s^{-1}). The velocity dispersion map shows evidence regions with velocity dispersion between 15 and 50 km s^{-1}; these are co-incident with the CO (2-1) intensity peaks, indicating a turbulent motion in the molecular gas.

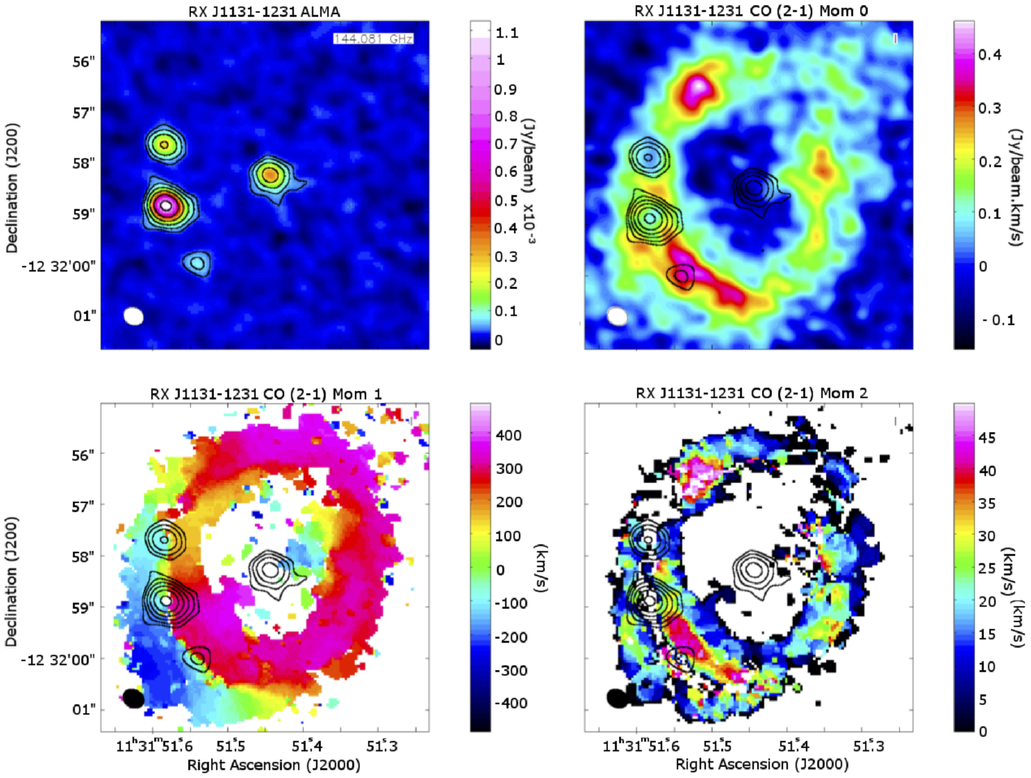

Figure 5.3: *Upper left*: The CLEANed image of the ALMA 2.1 mm continuum, using natural weighting. The beam size is 0.41×0.34 arcsec at a position angle 62 deg east of north (indicated by the white ellipse), the rms noise level is 10 μJy beam^{-1}. *Upper right*: CO (2-1) integrated line emission (moment-zero), natural weighting. The beam size is 0.44×0.36 arcsec at 60 deg; the rms noise level is 38 mJy beam^{-1} km s^{-1}. *Lower left*: CO (2-1) moment-one (velocity) map, showing the clear velocity-dependent structure of the molecular Einstein ring emission. *Lower right*: CO (2-1) moment-two (velocity dispersion) map, showing evidence for turbulent molecular regions with a velocity dispersion up to 50 km s^{-1}. In each panel, the ALMA 2.1 mm continuum contours at 30, 60, 120, 240, 480 and 960 μJy beam^{-1} km s^{-1} are overlaid for reference.

5.4 Lens modelling

5.4.1 Obtaining the lens model

For the mass density profile of the lensing galaxy, we adopt an elliptical power-law model with an external shear (see Chapter 2). The HST-ACS imaging (Figure 5.1, Suyu et al., 2013) reveals the presence of a luminous satellite to the north of the lensing galaxy, which we include into our analysis; we adopt a singular isothermal ellipsoid profile for the satellite. Consequently, our lens model has a total of thirteen free parameters, as listed in Table 5.3.

We first obtain a preliminary lens model using the CLEANed CO (2-1) data. The initial position of the lens centre is set to be co-incidental with the AGN in the lens galaxy, which is offset by 0.810 arcsec to the West and 0.450 arcsec to the North of the phase-tracking centre; the initial lens-satellite offset was based on HST imaging. The final lens model parameters optimization was in the visibility plane using the CO (2-1) data averaged over channels 121-200, with time averaging of 120 seconds, in order to obtain the maximal SNR. We then apply this lens model to obtain a reconstruction of the continuum source and individual CO (2-1) velocity bins, as described below.

The intrinsic properties of the reconstructed source in the continuum and in each velocity channel are discussed in details in the following section. We find that the inferred mass lens parameters (Table 5.3) are generally in agreement within 2-σ with those derived by Suyu et al. (2013) and Birrer et al. (2016) from a modelling of the HST data as well as those derived by Chen et al. (2016) from a modelling of Keck-II AO data (with the largest discrepancy on the mass density slope of the main galaxy). Both HST and Keck-II data have resolution almost a factor of 10 better than the ALMA Band 4 observations presented in this paper; the coarser resolution of ALMA data shows itself in a large degeneracy in the satellite parameters. By contrast, our model is not consistent with the parameters recently derived by Leung et al. (2017), which also shows tensions with Suyu et al. (2013), Birrer et al. (2016) and Chen et al. (2016). In particular, Leung et al. (2017) do not account for the presence of the satellite and do not allow for an external shear contribution. This leads to a mass model which is significantly flattened (q = 0.56±0.16) and which might bias their CO (2-1) line reconstruction.

Table 5.3: RXJ1131-1231: best lens model parameters. All parameters are listed as defined in Chapter *2*. The main lens centre position coordinates ΔX and ΔY derived from ALMA Band 4 data are defined with respect to the phase-tracking centre; the coordinates of the satellite mass density centre are given relative to the centre of the main lens.

		Main lens	Satellite
κ_0	[arcsec]	1.702±0.09	0.19±0.05
f		0.767 ± 0.014	0.890 ± 0.081
θ	[deg]	111±1	1±22
ΔX	[arcsec]	0.81 ± 0.01	-0.07 ± 0.10
ΔY	[arcsec]	0.28 ± 0.14	0.74 ± 0.14
γ		2.04 ± 0.03	$\equiv 2.00$
Γ		0.084 ± 0.016	
Γ_θ	[deg]	93±3	

5.4.2 Continuum modelling

As our lens modelling procedure can not include the presence of the AGN in the lensing galaxy, the contribution of the lensing galaxy to the observed visibility function had to be removed[1]. We use MS-Clean to obtain a model of the AGN, which

---

[1] An alternative - and more consistent - approach would include a single point-source model of the foreground AGN into the visibility-fitting procedure, in a manner similar to including the light from the lensing galaxy when modelling optical data.

was then subtracted from the observed visibility function using the `uvsub` task of the CASA package. To account for AGN-associated structures on a wider range of scales, we extracted the AGN model by first CLEANing the continuum data using uniform weighting, followed by further CLEAN runs with Briggs (robust = 0) and natural weightings.

To obtain a noise estimate, we first used MS-CLEAN to obtain a model of the sky surface brightness distribution; the corresponding visibility response was then subtracted from the data to obtain a pure noise-map. As the observation was carried out in two observing blocks at a different target elevation (77-80 vs. 45-54 deg above the horizon), the noise level changes appreciably between the two blocks, by about 30%. To obtain a robust estimate of the noise on the visibilities, the rms noise was estimated for each observing block separately.

To reduce the number of visibilities for the lens-modelling of the continuum, we used a time-averaging of 120 s. The image-smearing due to time-averaging is still negligible. In the frequency domain, each SPW was averaged into a single channel; all the four SPWs were then modelled together. Finally, the two polarizations were combined into the Stokes I parameter, $I = \frac{1}{2}(XX + YY)$. The final number of visibilities used for was about 10^5. The image-plane pixel size was set to 50 mas; the source grid is obtained by casting back all the sky-plane pixels, resulting in a sub-sampling of a factor of 5-6.

5.4.3 CO (2-1) modelling

The CO (2-1) was detected in SPW 2. After subtracting the continuum, we split the line-containing channels (#121-200) along with twenty channels on each side of the line (#101-120, 201-220), from the main dataset. These were then averaged in frequency into a total of 12 frequency bins of 10 channels each. The velocity resolution of the frequency-averaged data was $84\,\mathrm{km\,s^{-1}}$. While this allows us to reconstruct the CO (2-1) surface brightness distribution and the velocity map with a good signal-to-noise ratio, a finer velocity resolution would be required to obtain a robust reconstruction of the velocity dispersion (moment 2), necessary for a more detailed kinematic analysis such as the kinemetry analysis presented in Chapter 3. As in the case of the continuum, data volume was further reduced by time-averaging with a bin-size of 120 seconds; this reduced the number of visibilities to about 25,000 per frequency bin.

Keeping the lens model derived in Section 5.4, we re-optimize for the source and its regularization λ_s. As in case of the continuum model, we reconstruct each channel both with a gradient and curvature form of the regularization. The image-plane pixel size was set to 100 mas; the source grid is obtained by casting back all the sky-plane pixels, resulting in a sub-sampling of a factor of 2-3.

Table 5.4: CO (2-1): slices used for the line reconstruction. Left to right: slice number, observed frequency, rest-frame velocity offset with respect to the systemic velocity (z=0.654), integrated line luminosity and total magnification for each slice.

Line	ν_0 [GHz]	Δv [km s^{-1}]	$I_\nu \Delta v$ [mJy km s^{-1}]	μ
CO (2-1) slice 1	139.451	-298.7	115	4.5
CO (2-1) slice 2	139.412	-214.9	85	5.4
CO (2-1) slice 3	139.472	-130.6	99	4.8
CO (2-1) slice 4	139.333	-46.6	90	6.1
CO (2-1) slice 5	139.294	+37.4	104	7.4
CO (2-1) slice 6	139.255	+121.5	124	14.0
CO (2-1) slice 7	139.216	+234.9	110	7.1
CO (2-1) slice 8	139.177	+289.5	196	4.7

5.5 Results

5.5.1 Continuum

Figure 5.4 shows the best-model continuum reconstruction. We note that the image-plane residuals show a leftover extended structure associated with the AGN in the lensing galaxy, despite our best efforts to remove the foreground contribution.

The 2.1 mm continuum source is strongly elongated, with a major-axis FWHM of ~200 mas, corresponding to a physical distance of 1.5 kpc at z=0.658. The elongated source-plane structure is consistent with the imaging of the lensed ALMA continuum indicating that image A is extended in the radial direction. However, a similar extended emission is not detected in either B or C images, potentially due to their lower magnification. Therefore, while we do not find any evidence that images B and C are not point-like, we can not exclude that they contain extended emission below the noise level.

We performed several tests to determine whether the extended nature of the continuum source is due to: (1) uncorrected phase errors; (2) incorrect lens model and (3) artifact of the pixellated reconstruction of a point-like source.

First, we checked if the extended source structure is caused by uncorrected phase errors. Namely, we first split the continuum into two parts corresponding to the two observing blocks. We then proceeded to reconstruct the continuum source independently for each dataset - if the resolved structure in image A is a phase-error artifact, we would expect the source structure to differ between the two reconstructions. However, the size of the reconstructed source was found to be consistent between the two observing blocks.

Second, we explored the dependence of the source reconstruction on the lens model (especially the lens mass centroid). Using several lens models shifted by up to 100 mas from the best-model lens centre position, we found that the reconstructed source remains extended in all cases. To obtain a compact source, the lens centre had to be shifted enough to perturb the positions of the triplet images; this would prevent us from properly recovering the CO (2-1) emission. Consequently, it is unlikely that the elongated structure of the source is caused by an erroneous lens model.

Finally, we investigated the possibility that our lens-modelling technique might incorrectly recover a truly point-like source. We created a set of mock ALMA observations of a lensed point-source, directly corresponding to the ALMA Cycle-2 data analyzed in this work. We then applied our lens-modelling technique to these mock data in order to reconstruct the mock continuum source, under the assumption of a correct lens model. We found that the compact nature of the point-source is consistently recovered.

Therefore, we conclude that the elongation of the continuum source is not an artifact of the pixellated source reconstruction, but it reflects the true extended structure of the A image. Nevertheless, the extended structure of the source has to be interpreted with caution, for the following reasons: (1) due to the small magnification in the radial direction, the constraints on the radial extent of the source depend directly on the resolution; (2) the lack of extended emission in B and C images matching that observed in the A image. We therefore consider the 200 mas linear extent of the reconstructed source as an *upper limit* on the actual source size; a more precise continuum-source size estimate will be provided by a more detailed microlensing analysis (Sluse et al., in prep).

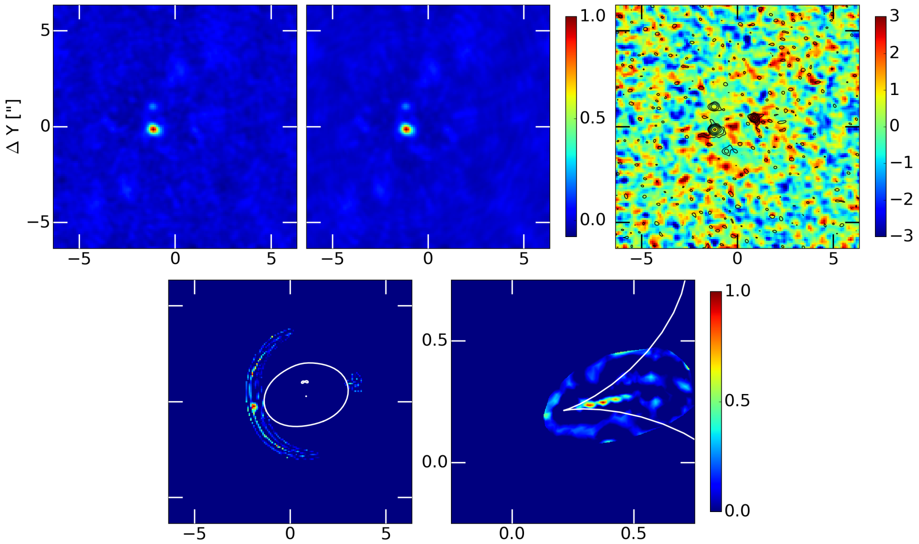

Figure 5.4: RXJ1131-1231: Band 4 continuum reconstruction. *Left to right*: Dirty image of the data, dirty image of the best model (normalized to the peak of the data), dirty image of the residuals (colourmap in units of dirty-image rms); the best sky surface brightness model, the best source-plane surface brightness distribution (normalized to their respective peaks).

5.5.2 CO (2-1) emission

The magnification provided by gravitational lensing allows us to clearly resolve a rotating disk and probe several physical properties that characterize the ionized gas in this galaxy. Figure 5.5 shows the reconstructed CO (2-1) emission for the eight frequency bins corresponding to channels 121-200 (Table 5.4). Starting from the bluest channels, the source is initially located just outside the caustic and hence only doubly imaged. The reddest channels are located inside the caustic and consequently form an almost perfect Einstein ring. Moreover, the reddest channels are also inherently brighter than the blue ones, up to a factor of two (Figure 5.5). A typical fish-tail structure associated with an ordered rotating disk is clearly seen. No emission was detected in the frequency bins corresponding to channels 101-120 and 201-220, which are therefore omitted from Figure 5.5.

The source-plane line profile (Figure 5.6) reveals a double-horn profile, typical of a rotating disk. As the CO (2-1) emission crosses the caustic, the observed sky-plane line profile is significantly affected by the differential magnification: μ ranges between 4 and 14 for different slices. This effect is however not as extreme as in the case of SDP.81 (Chapter 3), as the CO (2-1) emission in RXJ1131-1231 is significantly more extended than the caustics.

Figure 5.7 shows the CO (2-1) velocity moment maps, where the first panel of the figure displays only the moment-zero map of the stack, i.e. the intensity. We also derive the moment-one map to measure the velocity map. The source-plane CO emission is very regular and extends over almost 3 arcseconds (21 kpc in physical units), displaying a large rotating disk. A range of clumpy structure is also detected, illustrating the quality of the data and the increased resolution provided by gravitational lensing. A large (∼5 kpc) bar-like structure is clearly seen; in the image plane, it corresponds to the region with high velocity dispersion (σ=15-50 km s^{-1}), indicating a turbulent motion within this structure. Strikingly, this feature is offset from the continuum source (AGN) by about 0.3 arcseconds (2 kpc, Figure 5.7).

Figure 5.7 also provides a rotation curve constructed from the CO (2-1) emission. The curve is corrected for the inclination of the disk, i=54 deg, that we estimate from its ellipticity. While a full dynamical analysis of the disk (joint with the continuum) is out of the scope of this work, we nevertheless estimate that the total mass within 5 kpc is $M(R < 5\,\mathrm{kpc}) = (1.4 \pm 0.4) \times 10^{11}$ M$_\odot$. This is in agreement with the Leung et al. (2017) estimate.

Although this feature is reminiscent of a spiral-galaxy bar, the overlay of HST imaging and CO (2-1) moment-zero map (Figure 5.8) does not reveal any corresponding structure in the stellar component. A more detailed kinematic study is required for a more robust interpretation of this bar-like structure.

5.5.3 Comparison with Leung et al., 2017

As mentioned above, Leung et al. (2017, L17) recently presented an analysis of low-resolution CO (3-2) and (2-1) observations with CARMA and PdBI, based on a parametric visibility-fitting code (Bussmann et al., 2013). Note that the resolution and sensitivity of these observations is an order of magnitude lower compared to the ALMA data presented in this work: the FWHM beam size is 3.2×1.9 and 4.4×2.0 arcseconds for the CARMA and PdBI data, respectively.

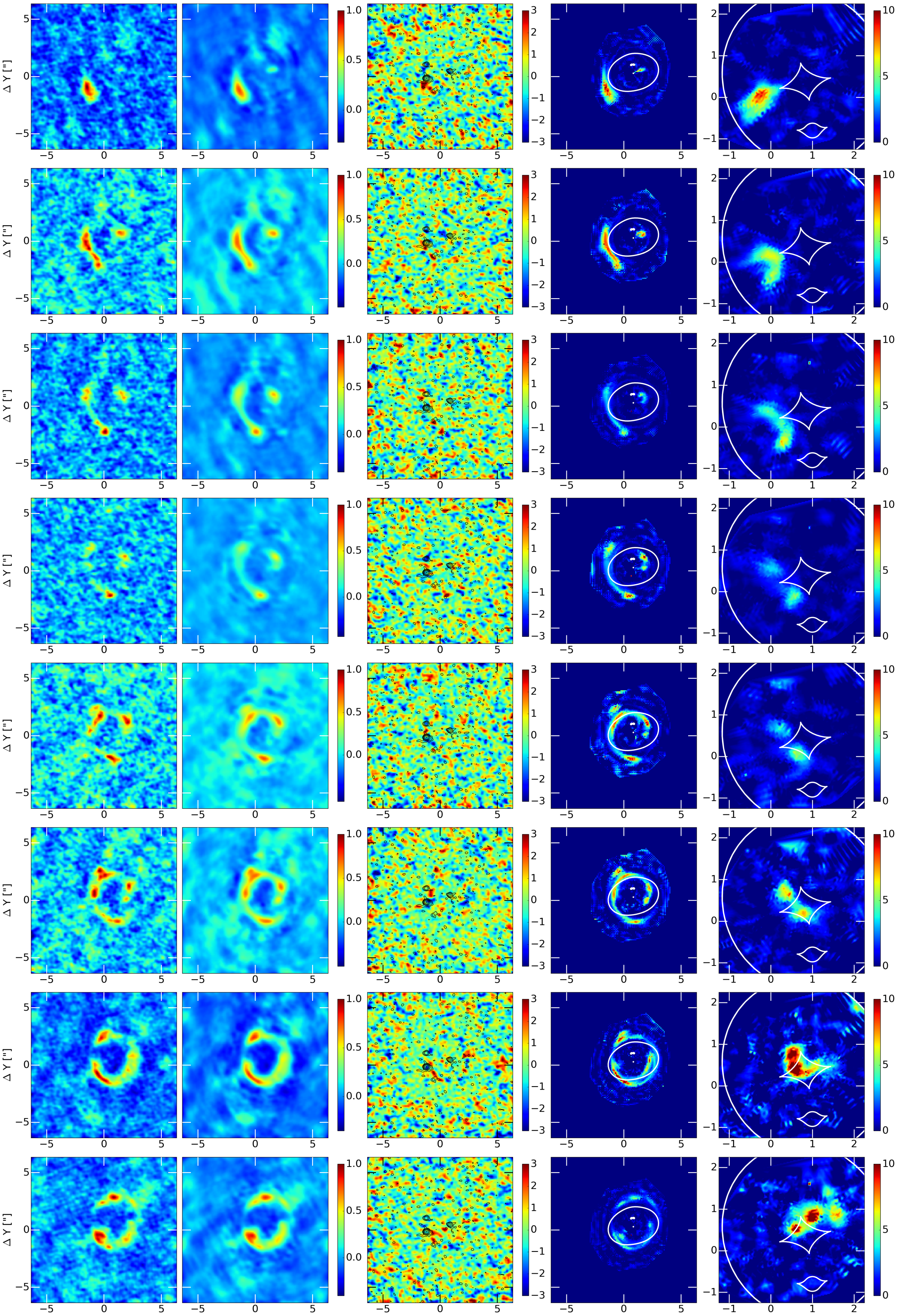

Figure 5.5: RXJ1131-1231: CO (2-1) reconstructions, velocity slices 1 to 8. *From left to right*: dirty image of the data; dirty image of the best model (normalized to the peak of the dirty image of the data). Dirty-image residuals (scale in σ). Best sky reconstruction (normalized to the peak). Best source reconstruction (mJy km s^{-1} kpc^{-2}).

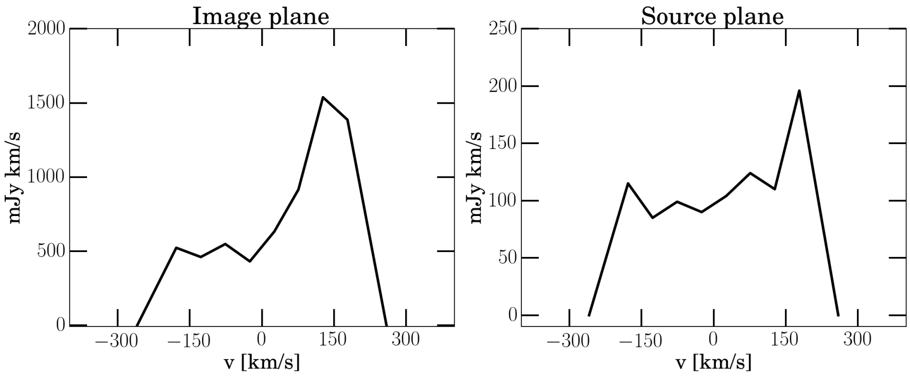

Figure 5.6: CO (2-1) line profiles. *Left*: line intensity measured in the image-plane. *Right*: line intensity measured in the source-place. Negative velocities correspond to the blue end of the line. Notice the considerable effect of the differential lensing on the amplitude of the red peak (1150 km s^{-1}); however, as seen in in the right panel and in Figure 5.5, the reddest channels are indeed inherently brighter than the bluest ones, approximately by a factor of two.

We discussed the major differences in the lens models between L17 and this work already in Section 5.4; here, we address the differences in the morphology of the reconstructed source. In particular, based on their lens-modelling of the CO (2-1) data using a visibility-fitting, parametric-source method of Bussmann et al. (2013)[2], L17 argue for a presence of a source-plane companion to the main source in RXJ1131-1231.

This companion would fall into the two reddest velocity slice listed in Table 5.4. Given the much better quality of the ALMA CO (2-1) dataset compared to the PdBI one, we would expect any kinematically distinct source-plane companion to be detected in our source-plane reconstruction. However, as shown in Figure 5.5, no such companion is recovered in our reconstruction. Instead, these slices correspond to the bar-like feature discussed in Section 5.5.3. This major difference is likely due to two factors: the major differences between the lens model used in L17 and those presented in this work and previous studies (Suyu et al., 2013, 2014; Chen et al., 2016), and to the inability of the parametric source-model to properly account for the complex structure of the CO (2-1) emission (even for low-resolution observations).

5.6 Conclusions

In this chapter, we presented new high-resolution ALMA Band 4 observations of a z=0.658 strongly-lensed quasar RXJ1131-1231. By achieving a typical source-plane resolution of 400-700 pc, we are able to study the structure and kinematics of the CO (2-1) and the source of the rest-frame 1.3 mm continuum emission at an unprecedented detail.

We found that the ALMA Band 4 continuum is dominated by synchrotron emission ($\sim$ 80% of the flux), rather than dust emission associated with either the AGN

---

[2] for more detail, see Chapters 2 and 3

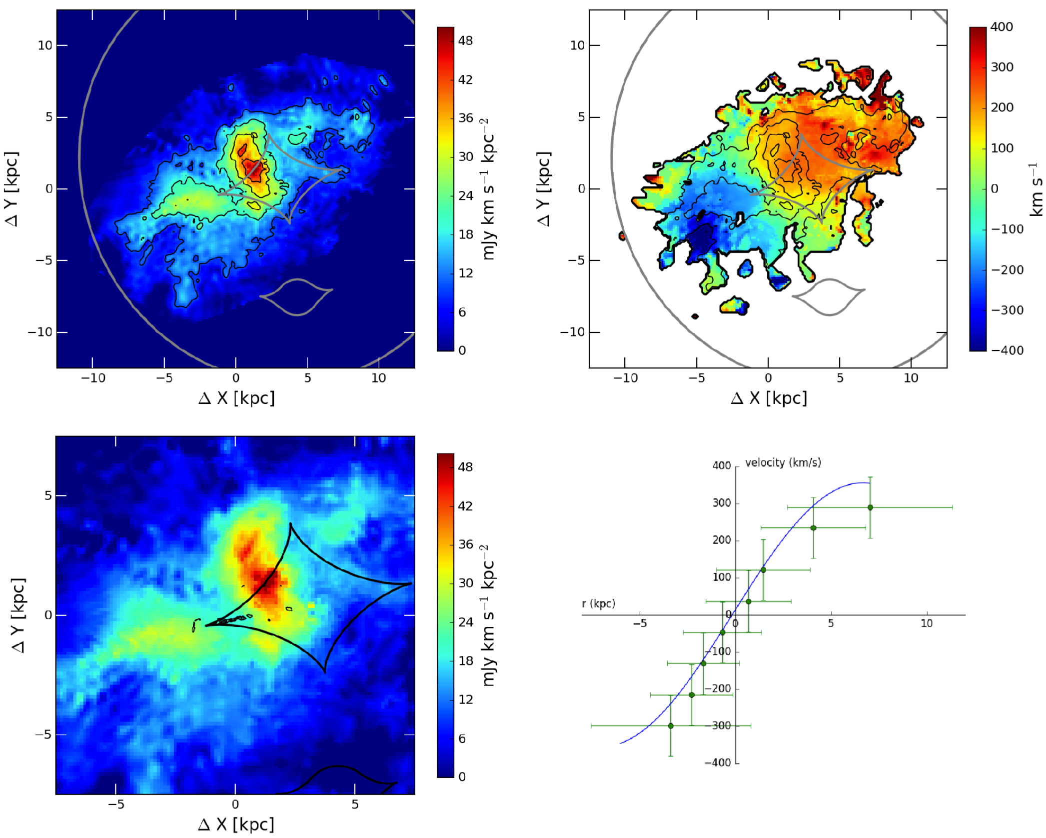

Figure 5.7: CO (2-1) moment maps. *Upper left*: integrated CO (2-1) line intensity (moment-0), in units of mJy km s^{-1} arcsec^{-2}. *Upper right*: moment-1 map, in units of km s^{-1}. Only regions for which the source-plane SNR>1 were taken into account. The black contours are drawn at 25, 50 and 75% of the peak surface brightness. *Lower left*: Overlay of the reconstructed CO (2-1) moment-0 map and 2.1 mm continuum. Caustics are traced by grey lines. *Lower right*: one-dimensional rotational curve obtained from CO (2-1) reconstruction, displaying a profile typical for disk-like dynamics. The solid blue line represents the rotational curve corrected for an inclination i=45 deg. The source-plane spatial scale is 1 arcsec = 7.030 kpc.

region or star-forming activity. Based on the flux-ratios of the triplet images measured in both the ALMA and optical continuum and using toy microlensing simulations, the size of the synchrotron-emission region is found to be more compact than the broad-line region (<100 pc). We attribute this emission to the optically thick base of the AGN jet. While the extended structure in the brightest lens image provides evidence for extended continuum emission, higher-resolution ALMA follow-up is necessary to disentangle the synchrotron and dust-related components.

Comparing the 2.1-mm flux-ratios of the triplet images with those measured at optical wavelengths, the A/B flux-ratio was found to be consistent with a microlensing scenario. A multi-epoch mm-wave follow-up on timescales of few years will be required to confirm this hypothesis. On the other hand, the C/B flux-ratio is unlikely to be explained by microlensing alone; suggesting a presence of a low-mass (10^6 M$_\odot$) substructure in the lensing galaxy.

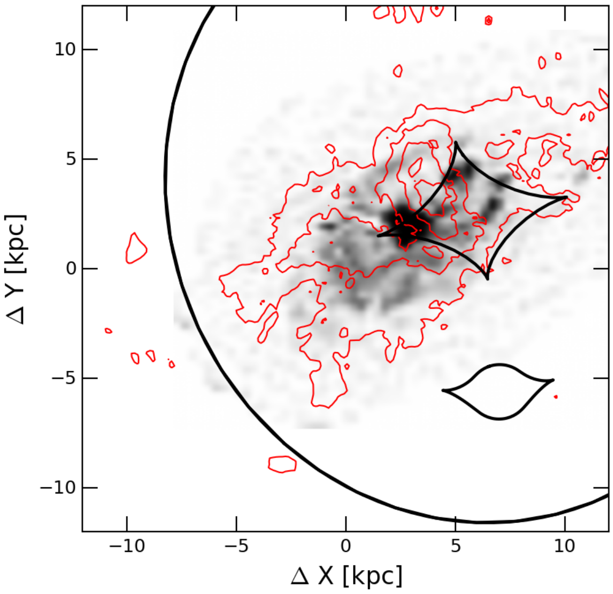

Figure 5.8: Overlay of source-plane ALMA CO (2-1) line emission (red contours, see Figure 5.7), and HST F814W continuum (greyscale, Suyu et al., 2013). The optical surface brightness peak corresponds to the quasar emission.

The reconstructed CO (2-1) reveals a large disk of molecular gas with ∼10 kpc radius, inclined at i ∼54 degrees; we estimate the dynamical mass within 5 kpc as $M(R < 5\,\mathrm{kpc}) = (1.46{\pm}0.31){\times}10^{11}\ \mathrm{M}_{\odot}$. The image-plane analysis provides evidence for turbulent motions associated with the bar-like feature ($\sigma = 15 - 50$ km s^{-1}). The magnification of individual velocity channels varies between 4 and 14. Although the velocity sampling used in this chapter (84 km s^{-1}) is insufficient for a detailed kinematic analysis, the high signal-to-noise ratio of the CO (2-1) dataset warrants further analysis.

The results presented in this chapter differ dramatically from the recently published analysis of Leung et al. (2017). In particular, the ALMA data rule out the presence of a source-plane companion proposed by Leung et al. (2017); highlighting the limitations of lens-modelling of very low resolution data with a parametric source in Leung et al. (2017) and the power of high-resolution observations with ALMA in revealing the detailed nature of the quasar and its host.

Finally, the ability to dissect different velocity components in lensed AGN hosts with ALMA on sub-kpc scales is of crucial importance for future detailed studies of AGN structure, especially in conjuction with the future infrared observations with the James Webb Space Telescope.

Chapter 6

Outlook and conclusions

In this final chapter, we summarize the main results of this thesis and outline the future developments. Specifically, we discuss modifications required to process extremely large datasets from the Atacama Large Millimeter/submillimeter Array (ALMA) and the Global Very Long Baseline Interferometry (VLBI) arrays, the application of gravitational imaging of substructures to interferometric data and lens-modelling of multi-band datasets.

6.1 Summary of main results

Lens-modelling of interferometric data in the visibility space

We have introduced a fully Bayesian lens-modelling technique tailored specifically of radio-interferometric observations, based on the technique of Vegetti & Koopmans (2009). Instead of using the synthesised (e.g. CLEANed) images which do not conserve the surface brightness and can contain deconvolution artifacts, we fit directly the observed visibility function. A source-plane regularization is used to prevent noise-fitting and unphysical solutions. By reconstructing the source on a pixellated grid rather than using an analytic profile, this technique is especially suited for analysing high-resolution observations of complex sources such as dusty, star-forming galaxies.

Modelling simulated ALMA observations

We first applied our technique to a set of simulated ALMA observations of a physically realistic high-redshift dust continuum emission for a wide range of angular resolution (0.1 - 1.0 arcseconds) and on-source time (60 - 3,600 seconds). While the input lens model is consistently recovered within 2-σ uncertainty for all datasets, the inferred source properties change dramatically with an increasing resolution. In particular, at low resolution, the compact, bright structures in the source surface brightness distribution are not properly resolved and the corresponding flux is distributed to larger spatial scales; resulting in an overestimation of the intrinsic source brightness and size by up to 50%. This bias highlights the large systematic uncertainties in inferring intrinsic source properties from low-resolution snapshot observations with ALMA and the Submillimeter Array (Negrello et al., 2010; Vieira et al., 2013; Bussmann et al., 2013, 2015; Spilker et al., 2016).

SDP.81: a sub-100-pc study of a redshift-3 starburst galaxy

In Chapter 3, we presented results of an analysis of the ALMA Long Baseline Campaign observations of SDP.81 (z=3.042). We provided the first maps of dust and CO emission in a high-redshift dusty star-forming galaxies at sub-100-pc resolution, an order of magnitude improvement over previous efforts. The reconstructed structure of the dust continuum surface brightness distribution suggests different conditions for star-formation in SDP.81, with several high-star formation rate (SFR) clumps surrounded by an extended star-formation. By combining ALMA data with archival Hubble Space Telescope and Very Large Array observations, a complex structure of SDP.81 is revealed - the gas, dust and stellar components are offset from each other, with an extended ($\sim$15 kpc) optical tail and a tentative counterpart systems suggestive of an on-going merger. Furthermore, a kinematic analysis of the CO (5-4) and (8-7) emission lines indicates significant perturbations to the system, supporting a major-merger scenario for the origin of the intense star-formation in SDP.81.

SPT lenses: comparison with parametric source models

In Chapter 4, we investigated the systematics of modelling low-resolution data using visibility-fitting codes with pixellated/parametric source models and the amount of structure that can be extracted from low-resolution ALMA observations. Namely, we applied our lens-modelling technique to a six lenses from the *South Pole Telescope* sample that were observed at 0.5-1.0 arcsecond resolution. We then compared our lens models to those obtained from the same data by Hezaveh et al. (2013b) and Spilker et al. (2016), who approximated the source by a Gaussian/Sérsic profile. We found that pixellated models reveal strong deviations from simple parametric models in three out of six systems, leading to corrections in the inferred source flux-density by up to a factor of a few. On the other hand, given the low quality of this particular dataset and the additional freedom of pixellated source models, several lens models are strongly degenerate. We demonstrated that some of these degeneracies can be broken if the imaging of the foreground galaxy is taken into account. The large scatter in inferred source properties limits the amount of information that can be extracted from low-resolution observations with ALMA Cycles 0 and 1 campaigns (Hezaveh et al., 2013b; Bussmann et al., 2015; Spilker et al., 2016) or the Sub-Millimeter Array (Bussmann et al., 2013).

RXJ1131-1231: AGN and gas properties

In Chapter 5, we applied our technique to ALMA Cycle 2 observations of a low-redshift strongly-lensed quasar RXJ1131-1231 (z=0.654), both the CO (2-1) emission line and 2.1 mm continuum. The CO (2-1) emission reveals a spectacular disk some 20 kpc across, with a spiral structure and a prominent bar-like feature. The continuum emission is composed of two separate components: a synchrotron emission from a compact source in the vicinity of the AGN accretion disk, and dust continuum emission associated with potential extended star-formation. While we found evidence for an extended continuum emission in the brightest lensed image, it is unclear whether this emission is associated with the source, or with the AGN activity in the lensing galaxy; additional observations are required to disentangle the lens/source contribution.

6.2 Future improvements and applications

Handling large datasets: gridding in the *uv*–plane and FFT

The main limitation of the Direct Fourier Transform (DFT) formalism used throughout this thesis, is that the size of the problem depends directly on the number of visibilities N_{vis}, as well as the inefficient computation of the Fourier transform. This becomes particularly problematic for datasets from full ALMA array ($N_{vis} = 10^6 - 10^7$) or the Global VLBI ($10^8 - 10^{10}$), which require prohibitive amounts of memory and computational time.

However, if the observed visibility function does not change significantly on short scales, the number of visibilities can be reduced by re-sampling them onto a (in principle arbitrary) *uv*-grid. The *uv*-plane gridding is a well-established procedure (e.g. Taylor et al., 1999). To briefly summarize, the observed visibility function $V(u,v)$ is first convolved with a kernel function $K(u',v')$ (e.g. a Kaiser-Bessel function). The resulting $V'(u,v) = K * V(u,v)$ can be then re-sampled onto a new *uv*-grid, after an appropriate re-normalization. Depending on the *uv*-coverage, N_{vis} is reduced by up to several orders of magnitude. The re-sampled visibilities can be straightforwardly processed with the DFT algorithm.

Provided the visibilities are re-sampled on a regular grid, a Fast Fourier Transform (FFT) algorithm can be used to further increase the speed of the calculation. However, the implementation of FFT techniques into our lens-modelling technique necessitates significant changes beyond the scope of this work.

Searching for massive substructures

As our lens-modelling technique is based on the method of Vegetti & Koopmans (2009), it can be readily applied to the *gravitational imaging* of substructures in the halo of the lensing galaxy as well as along the line-of-sight (Koopmans, 2005; Vegetti & Koopmans, 2009). A large number of substructures is predicted by the Cold Dark Matter (CDM) paradigm; however, there are persistent tensions between the amount of substructure predicted by cosmological simulations and inferred from observation (e.g. Klypin et al., 1999; Kravtsov et al., 2004). This discrepancy leaves open the possibility for alternative dark-matter models that modify the substructure mass-function, such as the warm dark matter (WDM) (Lovell et al., 2012, 2014).

To discriminate between different dark matter models, it is necessary to constrain the low-mass end of the substructure mass-function. Although low-mass substructure is unlikely to contain enough luminous material to be detectable by its light, the gravitational effect of its mass is still detectable by gravitational lensing. Indeed, the gravitational imaging technique of Vegetti & Koopmans (2009) has been used to detect massive substructure ($M_{sub} \geq 10^8$ M$_\odot$) in optical (HST: Vegetti et al., 2010, 2014) and near-IR imaging data (Keck-II adaptive optics telescope: Vegetti et al., 2012). Pushing the substructure detection threshold to even lower M_{sub} requires the superior resolution of large interferometers such as ALMA (~10 mas) and the Global VLBI (~1 mas). Indeed, Hezaveh et al. (2016) have recently reported a detection of a $M_{sub} \sim 10^9$ M$_\odot$ substructure in ALMA observations of SDP.81.

To demonstrate the feasibility of the gravitational imaging with our visibility-fitting code, we carry out preliminary tests on mock ALMA Band 7 observations simulating a 10^9 M$_\odot$ substructure in SDP.81. By including pixellated perturbations to the lensing potential and implementing a gridding procedure outlined in the previous section to reduce the data volume, we recover the correct position of the substructure (Figure 6.1).

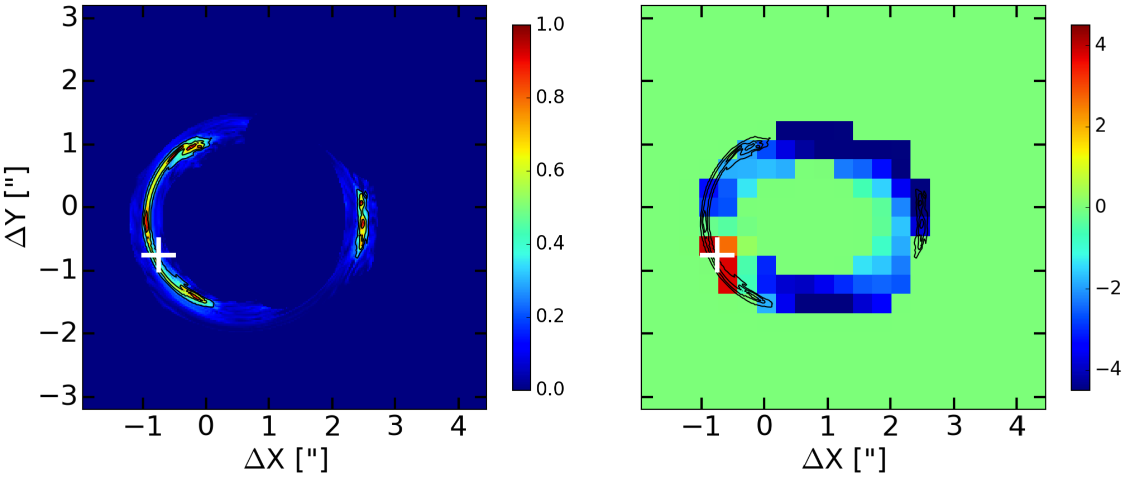

Figure 6.1: Detection of a massive substructure in simulated ALMA observations. *Left*: input sky-plane surface brightness distribution, with a massive substructure corresponding to the detection reported by Hezaveh et al. (2016) located at the position indicated by the white cross. *Right*: convergence correction $\Delta\kappa$ map shows a positive correction at the position of the substructure.

Multi-band lens-modelling

The lens mass models presented in this thesis were all derived from a single-band datasets, such as the ALMA Band 7 continuum in case of SDP.81 (Chapter 3) or the integrated CO (2-1) emission line in case of RXJ1131-1231 (Chapter 5). However, as the lens model does not depend on the observed wavelength, combining observations in multiple bands will increase the number of constraints on the lens model.

The ability to obtain the lens mass model by simultaneous modelling of several datasets is of particular importance for low signal-to-noise ratio observations, where combining different tracers can significantly increase the number of constraints on the lens model.

Acknowledgements

Co-authors and collaborators

The results presented in this thesis are based on a collaborative effort. Here, I would like to particularly acknowledge the work of Simona Vegetti on the development of our lens-modelling code, and on the lens modelling of HST observations and CLEANed data of SDP.81 and RXJ1131-1231 (Chapters 3 and 5); that of John McKean on the processing of ALMA data (Chapters 3, 4 and 5), Hannah Stacey on the SED-fitting of RXJ1131-1231 (Chapter 5) and Dominique Sluse and Danuta Paraficz on the micro-lensing simulations (Chapter 5). Many thanks to all the other collaborators on papers and proposals: Leon Koopmans, Chris Fassnacht,Cristiana Spingola, Matt Auger, Frederic Courbin, Luca Graziani, Sherry Suyu and many others.

Many thanks to Leon Koopmans, Chris Fassnacht, Henrik Junklewitz, André Offringa and Maksim Greiner for useful discussions on the development of our visibility-fitting method (Chapters 2 and 6); to Yashar Hezaveh and Justin Spilker for discussion on the lens model of SPT lenses (Chapter 4); to Sherry Suyu and Frederic Courbin for discussions on RXJ1131-1231 (Chapter 5).

Finally, I am greatly indebted to Paola Andreani (ESO Garching) for her initiative that got me involved with the ALMA data and for all her support.

Data products

This thesis makes use of the following ALMA data:

- Chapter 3: ADS/JAO.ALMA #2011.0.00016.SV,
- Chapter 4: ADS/JAO.ALMA #2011.0.00957.S and #2011.0.00958.S,
- Chapter 5: ADS/JAO.ALMA#2013.1.01207.S.

ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada) and NSC and ASIAA (Taiwan), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO and NAOJ. The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc.

Chapters 3 and 4 are based on observations made with the NASA/ESA Hubble Space Telescope, obtained from the data archive at the Space Telescope Science Institute. STScI is operated by the Association of Universities for Research in Astronomy, Inc. under NASA contract NAS 5-26555.

Figures

In this thesis, I reproduced several figures from other publications: ALMA Partnership et al. (2015b); Bouwens et al. (2010); Bussmann et al. (2013); Carilli & Walter (2013); Hodge et al. (2015); Negrello et al. (2010, 2017); Shapley (2011); Spilker

et al. (2016). Several figures in this work were produced by my collaborators: many thanks John McKean for Figures 3.2, 3.4, 5.1 and 5.3, Hannah Stacey for Figure 5.8, Sherry Suyu for the optical data in Figure 5.8 and Danuta Paraficz for Figure 5.7.

Translation

Finally, I would like to thank Korbinian Huber and Marlen Lauffer for the German translation of the Summary.

Danksagung

To Simona Vegetti and John McKean, my PhD-parents: thank you very much for your guidance, support, patience and all the opportunities they provided, and for the often-rushed München-Groningen trips.

To Simon White, for his trust, time and energy invested in me, for listening to my concerns and covering unexpected conference bills.

To my amazing co-workers and friends: Elisa Ritondale, Guilia Despali and Francesca Rizzo. To Cristiana Spingola, who is in Groningen, but definitely feels like one of us. "*Once I had all sorts of friends, now it's all Italian girls.*"

To the SHARP collaboration - Chris Fassnacht, Leon Koopmans, Matt Auger, Lindsay Oldham, Dave Lagattuta, Dorota Bayer, Jen-Wei Hsueh and all the others - for the spirit of community, an their professional and unprofessional advice, solicited and unsolicited.

To the MPA administration staff - Cornelia Rickl, Maria Depner, Stella Veith, Sonja Gruendl and all the others, for helping with a miriad tasks and making our lives much, much easier. And to our tech support - Bernt Christiandl and Andi Weiss - for their prompt response to all requests, big and small.

To my ex-supervisors and those who helped me to get here and taught me so much: Graham Bruce, Frieder König, Duncan Robertson, Ian Bonnell. Special thanks to Adrienne Erickcek and people from the Perimeter Institute: without them, I would have never joined the MPA.

To my fellow musketeers: Mikkel Nielsen and Alex Kolodzig, and to all the astrophysics friends I met here, especially Dijana Vrbanec, Haakon Andersen, Matteo Bugli, Amy Jones, Martha Lippich, Isabella Soldner-Remboldt, Jeffrey Chan, Vanessa Böhm, Durand d'Souza, Fabian Schmidt and Jan Grieb. To Tami and Jon, for being wonderful friends for all these years. To Diana Pinto, one of the bravest people I know.

To Pünktchen and Anton, who really enjoyed the making of this thesis. They have been eating a lot of paper lately.

To my family, for their constant support and inspiration over so many years.

To Marlen, for being an amazing and patient partner in crime and a constant voice of reason.

Bibliography

ALMA Partnership et al., 2015a, ApJl, 808, L1

ALMA Partnership et al., 2015b, Apjl, 808, L4

Aravena M., et al., 2013, MNRAS, 433, 498

Barkana R., 1998, ApJ, 502, 531

Bartelmann M., Schneider P., 1999, A&A, 345, 17

Barvainis R., Ivison R., 2002, ApJ, 571, 712

Bellocchi E., Arribas S., Colina L., 2012, A&A, 542, A54

Bellocchi E., Arribas S., Colina L., Miralles-Caballero D., 2013, AA, 557, A59

Birrer S., Amara A., Refregier A., 2016, J.~Cosmology Astropart. Phys., 8, 020

Blain A. W., 1996, MNRAS, 283, 1340

Blandford R. D., Narayan R., 1992, ARA&A, 30, 311

Bolton A. S., Burles S., Koopmans L. V. E., Treu T., Moustakas L. A., 2006, ApJ, 638, 703

Bonvin V., Tewes M., Courbin F., Kuntzer T., Sluse D., Meylan G., 2016, A&A, 585, A88

Bothwell M. S., et al., 2013, ApJ, 779, 67

Bournaud F., et al., 2014, ApJ, 780, 57

Bournaud F., Daddi E., Weiß A., Renaud F., Mastropietro C., Teyssier R., 2015, A&A, 575, A56

Bouwens R. J., et al., 2010, ApJl, 709, L133

Burtscher L., et al., 2013, A&A, 558, A149

Bussmann R. S., et al., 2012, ApJ, 744, 150

Bussmann R. S., et al., 2013, ApJ, 779, 25

Bussmann R. S., et al., 2015, ApJ, 812, 43

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

Carilli C. L., Hodge J., Walter F., Riechers D., Daddi E., Dannerbauer H., Morrison G. E., 2011, ApJl, 739, L33

Carlstrom J. E., et al., 2011, PASP, 123, 568

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

Chartas G., Kochanek C. S., Dai X., Moore D., Mosquera A. M., Blackburne J. A., 2012, ApJ, 757, 137

Chen G. C.-F., et al., 2016, MNRAS, 462, 3457

Chomiuk L., Povich M. S., 2011, AJ, 142, 197

Clark B. G., 1980, AA, 89, 377

Coppin K., et al., 2008, MNRAS, 384, 1597

Cornwell T. J., 2008, IEEE Journal of Selected Topics in Signal Processing, 2, 793

Dai X., Kochanek C. S., Chartas G., Kozłowski S., Morgan C. W., Garmire G., Agol E., 2010, ApJ, 709, 278

Davé R., Finlator K., Oppenheimer B. D., Fardal M., Katz N., Kereš D., Weinberg D. H., 2010, MNRAS, 404, 1355

Dekel A., et al., 2009, *Nature*, 457, 451

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

Dye S., Warren S. J., 2005, ApJ, 623, 31

Dye S., et al., 2014, MNRAS, 440, 2013

Dye S., et al., 2015, MNRAS, 452, 2258

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

Ellithorpe J. D., Kochanek C. S., Hewitt J. N., 1996a, ApJ, 464, 556

Ellithorpe J. D., Kochanek C. S., Hewitt J. N., 1996b, ApJ, 464, 556

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

Frayer D. T., et al., 2011, ApJl, 726, L22

Garsden H., et al., 2015, A&A, 575, A90

Gaudi B. S., 2010, Microlensing by Exoplanets. pp 79–110

Gould A., Loeb A., 1992, ApJ, 396, 104

Greve T. R., et al., 2012, ApJ, 756, 101

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

Hayward C. C., Narayanan D., Kereš D., Jonsson P., Hopkins P. F., Cox T. J., Hernquist L., 2013, MNRAS, 428, 2529

Hezaveh Y., Dalal N., Holder G., Kuhlen M., Marrone D., Murray N., Vieira J., 2013a, ApJ, 767, 9

Hezaveh Y. D., et al., 2013b, ApJ, 767, 132

Hezaveh Y. D., et al., 2016, ApJ, 823, 37

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

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

Hodge J. A., et al., 2016, ApJ, 833, 103

Hoekstra H., Jain B., 2008, Annual Review of Nuclear and Particle Science, 58, 99

Högbom J. A., 1974, AAPS, 15, 417

Holland W. S., et al., 1999, MNRAS, 303, 659

Hopkins A. M., 2004, ApJ, 615, 209

Hopkins A. M., Beacom J. F., 2006, ApJ, 651, 142

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

Jaffe W., et al., 2004, *Nature*, 429, 47

Junklewitz H., Bell M. R., Selig M., Enßlin T. A., 2016, AA, 586, A76

Kereš D., Katz N., Davé R., Fardal M., Weinberg D. H., 2009, MNRAS, 396, 2332

Klypin A., Kravtsov A. V., Valenzuela O., Prada F., 1999, ApJ, 522, 82

Kochanek C. S., 1991, ApJ, 373, 354

Kochanek C. S., Dalal N., 2004, ApJ, 610, 69

Kochanek C. S., Narayan R., 1992, ApJ, 401, 461

Kochanek C. S., Blandford R. D., Lawrence C. R., Narayan R., 1989, MNRAS, 238, 43

Koopmans L. V. E., 2005, MNRAS, 363, 1136

Kormann R., Schneider P., Bartelmann M., 1994, A&A, 284, 285

Krajnović D., Cappellari M., de Zeeuw P. T., Copin Y., 2006, MNRAS, 366, 787

Kravtsov A. V., Berlind A. A., Wechsler R. H., Klypin A. A., Gottlöber S., Allgood B., Primack J. R., 2004, ApJ, 609, 35

Kreysa E., et al., 1999, Infrared Physics and Technology, 40, 191

Lagattuta D. J., Vegetti S., Fassnacht C. D., Auger M. W., Koopmans L. V. E., McKean J. P., 2012, MNRAS, 424, 2800

Leroy A. K., et al., 2013, AJ, 146, 19

Leung T. K. D., Riechers D. A., Pavesi R., 2017, ApJ, 836, 180

Lin H., et al., 2009, ApJ, 699, 1242

Lofthouse E. K., Kaviraj S., Conselice C. J., Mortlock A., Hartley W., 2017, MNRAS, 465, 2895

Lovell M. R., et al., 2012, MNRAS, 420, 2318

Lovell M. R., Frenk C. S., Eke V. R., Jenkins A., Gao L., Theuns T., 2014, MNRAS, 439, 300

Lupu R. E., et al., 2012, ApJ, 757, 135

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

Maiolino R., et al., 2015, MNRAS, 452, 54

Mao S., Paczynski B., 1991, ApJl, 374, L37

Mao S., Schneider P., 1998, MNRAS, 295, 587

Marshall P. J., et al., 2016, MNRAS, 455, 1171

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

Messias H., et al., 2014, A&A, 568, A92

Meylan G., Jetzer P., North P., Schneider P., Kochanek C. S., Wambsganss J., eds, 2006, Gravitational Lensing: Strong, Weak and Micro (`arXiv:astro-ph/0407232`)

Mortonson M. J., Schechter P. L., Wambsganss J., 2005, ApJ, 628, 594

Narayanan D., Krumholz M. R., 2014, MNRAS, 442, 1411

Narayanan D., Hayward C. C., Cox T. J., Hernquist L., Jonsson P., Younger J. D., Groves B., 2010a, MNRAS, 401, 1613

Narayanan D., et al., 2010b, MNRAS, 407, 1701

Narayanan D., et al., 2015, *Nature*, 525, 496

Negrello M., Perrotta F., González-Nuevo J., Silva L., de Zotti G., Granato G. L., Baccigalupi C., Danese L., 2007, MNRAS, 377, 1557

Negrello M., et al., 2010, Science, 330, 800

Negrello M., et al., 2014, MNRAS, 440, 1999

Negrello M., et al., 2017, MNRAS, 465, 3558

Oguri M., 2010, glafic: Software Package for Analyzing Gravitational Lensing, Astrophysics Source Code Library (ascl:1010.012)

Oliver S. J., et al., 2012, MNRAS, 424, 1614

Omont A., et al., 2011, A&A, 530, L3

Omont A., et al., 2013, A&A, 551, A115

Oteo I., et al., 2017, preprint, (`arXiv:1701.05901`)

Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2010, AJ, 139, 2097

Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, Numerical recipes in FORTRAN. The art of scientific computing

Refsdal S., 1964, MNRAS, 128, 307

Reis R. C., Reynolds M. T., Miller J. M., Walton D. J., 2014, *Nature*, 507, 207

Riechers D. A., Hodge J., Walter F., Carilli C. L., Bertoldi F., 2011, ApJl, 739, L31

Rybak M., McKean J. P., Vegetti S., Andreani P., White S. D. M., 2015a, MNRAS, 451, L40

Rybak M., Vegetti S., McKean J. P., Andreani P., White S. D. M., 2015b, MNRAS, 453, L26

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

Schwab F. R., 1984, AJ, 89, 1076

Scott K. S., et al., 2011, ApJ, 733, 29

Serjeant S., 2012, MNRAS, 424, 2429

Shapiro K. L., et al., 2008, ApJ, 682, 231

Shapley A. E., 2011, ARAA, 49, 525

Sluse D., Tewes M., 2014, A&A, 571, A60

Sluse D., et al., 2003, A&A, 406, L43

Sluse D., et al., 2006, A&A, 449, 539

Sluse D., Claeskens J.-F., Hutsemékers D., Surdej J., 2007, A&A, 468, 885

Sluse D., Kishimoto M., Anguita T., Wucknitz O., Wambsganss J., 2013, A&A, 553, A53

Spilker J. S., et al., 2015, ApJ, 811, 124

Spilker J. S., et al., 2016, ApJ, 826, 112

Stacey H. R., et al., 2017, preprint, (`arXiv:1705.10530`)

Sugai H., Kawai A., Shimono A., Hattori T., Kosugi G., Kashikawa N., Inoue K. T., Chiba M., 2007, ApJ, 660, 1016

Sutter P. M., et al., 2014, MNRAS, 438, 768

Suyu S. H., Marshall P. J., Hobson M. P., Blandford R. D., 2006, MNRAS, 371, 983

Suyu S. H., et al., 2013, ApJ, 766, 70

Suyu S. H., et al., 2014, ApJl, 788, L35

Suyu S. H., et al., 2017, MNRAS, 468, 2590

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., 2015, ApJ, 806, L17

Tacconi L. J., et al., 2006, ApJ, 640, 228

Tacconi L. J., et al., 2008, ApJ, 680, 246

Tacconi L. J., et al., 2013, ApJ, 768, 74

Tamura Y., Oguri M., Iono D., Hatsukade B., Matsuda Y., Hayashi M., 2015, PASJ, 67, 72

Taylor G. B., Carilli C. L., Perley R. A., eds, 1999, Synthesis Imaging in Radio Astronomy II Astronomical Society of the Pacific Conference Series Vol. 180

Tewes M., Courbin F., Meylan G., 2013, A&A, 553, A120

Valtchanov I., et al., 2011, MNRAS, 415, 3473

Vegetti S., Koopmans L. V. E., 2009, MNRAS, 392, 945

Vegetti S., Koopmans L. V. E., Bolton A., Treu T., Gavazzi R., 2010, MNRAS, 408, 1969

Vegetti S., Lagattuta D. J., McKean J. P., Auger M. W., Fassnacht C. D., Koopmans L. V. E., 2012, *Nature*, 481, 341

Vegetti S., Koopmans L. V. E., Auger M. W., Treu T., Bolton A. S., 2014, MNRAS, 442, 2017

Vieira J. D., et al., 2013, *Nature*, 495, 344

Wakker B. P., Schwarz U. J., 1988, AA, 200, 312

Wardlow J. L., et al., 2011, MNRAS, 415, 1479

Warren S. J., Dye S., 2003, ApJ, 590, 673

Weiß A., et al., 2013, ApJ, 767, 88

Wong K. C., Suyu S. H., Matsushita S., 2015, preprint, (`arXiv:1503.05558`)

Wucknitz O., 2002, Lens Models for Compact and Extended Sources

Wucknitz O., 2004, MNRAS, 349, 1

Wucknitz O., Volino F., 2008, preprint, (`arXiv:0811.3421`)

Xu D., Sluse D., Gao L., Wang J., Frenk C., Mao S., Schneider P., Springel V., 2015, MNRAS, 447, 3189

Yang C., et al., 2016, A&A, 595, A80