Studying high-redshift galaxies with ALMA: biases due to complex source structure and companion sources

THESIS

submitted in partial fulfillment of the
requirements for the degree of

BACHELOR OF SCIENCE
in
ASTRONOMY AND PHYSICS

Author :	Aniek van Ogtrop
Student ID :	s1854305
Supervisor :	Matus Rybak
2^{nd} supervisor :	Jacqueline Hodge
2^{nd} corrector :	Dorothea Samtleben

Leiden, The Netherlands, June 28, 2019

Studying high-redshift galaxies with ALMA: biases due to complex source structure and companion sources

Aniek van Ogtrop

Sterrewacht, Leiden University
P.O Box 9513, 2300 RA Leiden, the Netherlands

June 28, 2019

Abstract

Since the arrival of millimeter wave interferometers and their subsequent advancement, the resolution of observations has improved significantly. However, the effects of these improvements have not yet been investigated. This research examines the effect of improving resolutions to SubMillimeter Galaxies (SMGs) like simulated Gaussian sources. To achieve this, simulated 350 GHz ALMA observations with 20 different antenna configurations with resolutions ranging from 0.05 to 1.1 arcsec of simulated Gaussian sources performed with CASA are used.
For resolutions better than 0.2 arcsec, depending on the SNR, the flux and size from observations obtained with an image plane analysis, where the size is found using CASA, can be significantly lower than the true flux and size, due to the largest angular scale of the observations. Fitting a circular Gaussian to the *uv*-plane data demonstrates that the *uv*-plane analysis more accurately recovers the true FWHM than the image plane analysis using CASA. However, the major and minor axes of an elliptical source are not recovered and the size of an elliptical source is overestimated with the *uv*-plane analysis.
Different SMG companion fields from the ALESS and UDS survey originally observed by Hodge et al. (2013), Simpson et al. (2015) and Wardlow et al. (2018) and quasar fields with a companion detected by Decarli et al. (2017) were simulated with circular Gaussian sources to examine the influence of companion sources to the FWHM of the main source acquired with the *uv*-plane analysis. Generally, the companions in the companion fields examined do not show an influence on the FWHM of the main source.

Contents

1 Introduction

Almost two decades ago SubMillimeter-luminous Galaxies (SMGs) were discovered (Blain et al. 2002; Hodge et al. 2013). SMGs lend their name from the wavelength range they were first discovered in (Gullberg et al. 2018). These are dusty, star forming galaxies between redshift 2 and 5. Their high redshift indicates SMGs are from the early Universe between 3.2 and 1.2 Gyr after the Big Bang. As a reference, it is currently 13.5 Gyr after the Big Bang (Wright 2006).

SMGs are bright in the submillimeter range due to the thermal continuum emission of dust grains which is responsible for about 99% of the emitted energy in the submillimeter wavebands. These dust grains absorb the optical and ultraviolet emission of the young stars in SMGs and re-radiates it at infrared and submillimeter wavelengths. The remainder of the emitted energy in the submillimeter wavebands is caused by line emission from transitions of atoms and molecules in the interstellar gas. These emission lines can be used to determine the spectroscopic redshift of these galaxies (Blain et al. 2002).

SMGs are known to have massive star forming bursts (Hodge et al. 2013). The star formation rate (SFR) of SMGs can amount up to thousands of solar masses per year (Casey et al. 2014), compared to the Milky Way which has a SFR of around 1 solar mass per year (Robitaille and Whitney 2010). The star formation is thought to be triggered by galaxy interactions. However, images with better resolutions are yet to find evidence for this in the morphology of these SMGs (Hodge et al. 2018).

One of the major goals of modern physical cosmology is discovering how galaxies and stars were formed from the almost uniform gas in the beginning of the Universe. By researching galaxy and star formation in the early Universe more light can be shed on this. It has been shown that SMGs contain about 40-50% of the total stellar mass at redshift 2 (Swinbank et al. 2014). This makes stars from an SMG representative of stars at redshift 2. As SMGs are from the early Universe and have a high star formation rate they are ideal targets to observe to learn more about early star formation (Blain et al. 2002).

Prior to the discovery of SMGs local Luminous InfraRed Galaxies (LIRGs) were found. It is uncertain if LIRGs closely relate to SMGs, but they do provide insight on galaxy evolution of similar galaxies. These LIRGs have an infrared luminosity of $L_{IR} > 10^{11}$ $L_\odot$. A small subset has a luminosity of $10^{12} < L_{IR} < 10^{13}$ $L_\odot$ and is called UltraLuminous InfraRed Galaxies (ULIRGs). These have star formation rates of the order of 50 $M_\odot yr^{-1}$ and are caused by major mergers between equal mass galaxies.

These mergers trigger star formation and the formation of dust particles that absorb emission from stars and re-radiates it in the infrared and submillimeter wavelengths. Due to the limited amount of gas and high star formation rate, this phase is thought to be short-lived (Casey et al. 2014).

As submillitmeter wavelengths are long compared to the optical range, large diameters of radio telescopes are required for a small resolution. Before the arrival of radio interferometry, the resolutions of single dish telescopes did not give a small enough resolution to compete with the resolutions reached with optical and near infrared telescopes. This has led to the recent spike in interest for submillimetter research (Blain et al. 2002).

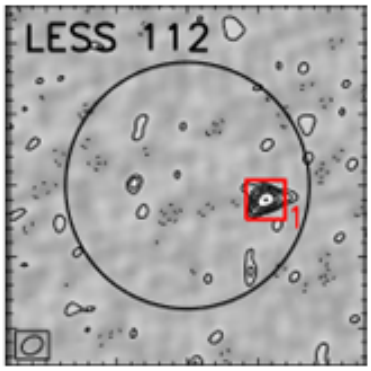


Figure 1.1: ALESS 112 from a 870 μm continuum observation with ALMA made by Hodge et al. (2013) with a resolution of 1.6 arcsec.

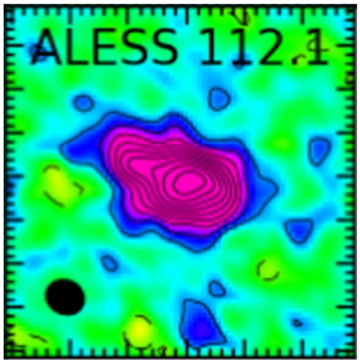


Figure 1.2: ALESS 112.1 from a 870 μm continuum observation with ALMA made by Hodge et al. (2016) with a resolution of 0.16 arcsec.

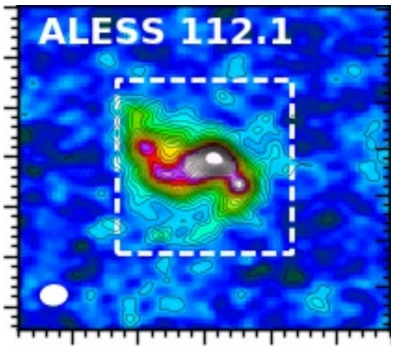


Figure 1.3: ALESS 112.1 from a 870 μm continuum observation with ALMA made by Hodge et al. (2018) with a resolution of 0.07 arcsec.

The limitations to the resolution of radio telescopes raises issues to observations of SMGs and any other object. When SMGs were observed with the LABOCA ECDFS (Extended Chandra Deep Field South) Submillimeter Survey (LESS) on the APEX telescope, a 12 m single dish radio telescope, with a spatial resolution of 27" (approximately $2.2 \cdot 10^2$ kpc assuming a redshift of 2.5 (Wright 2006)) (Weiss et al. 2009) only bright sources could be detected and none of which were resolved. As the SMGs are not resolved some assumptions needed to be made to analyse the data. The observed SMGs were assumed to be perfect circular Gaussian sources and each observed SMG was assumed to be alone. Then Hodge et al. (2013) used an ALMA (Atacama Large Millimeter/submillimeter Array) survey (ALESS) with a resolution of 1.6 arcsec to find that at least 35% (possibly up to 50%) of those bright sources from LESS have bright companion galaxies even though they were assumed to be solitary. An example of a field observed by Hodge et al. (2013) is shown in figure 1.1 for the source LESS 112 which is called ALESS 112.1 in the ALESS survey. This field is also observed by Hodge et al. (2016) with a resolution of 0.16 arcsec (see figure 1.2). Three years later Wardlow et al. (2018) made Band 3 ALMA observations with a resolution between 0.8 and 1.1 arcsec of six of these ALESS companion fields and proved that the sources were in fact

companions. Follow up ALMA observations in Band 7 on six luminous SMGs from the ALESS survey by Hodge et al. (2018) with a 0.07" resolution showed that these sources have a complex morphology and are not smooth sources as previously assumed (see figure 1.3). This also implies that the data analysis from these earlier surveys have been executed using overly simplistic assumption.

This research focuses on the impact these overly simplistic assumptions have on the measured properties of the galaxies. This is achieved using simulations of ALMA observations with different antenna placements of simulated sources with similar properties as ALESS sources generated using CASA[1]. This research consists of two parts. The first part investigates the influence of an improving resolution on single Gaussian sources. Especially, the influence of an improving resolution on the observed flux and size of the source. The size of the source is determined using two methods, namely using an image plane analysis and a *uv*-plane analysis. The second part of this research examines the influence of companion sources to the observed diameter of the main source. Companion fields originally observed by Hodge et al. (2013), Simpson et al. (2015), Wardlow et al. (2018) and Decarli et al. (2017) are simulated using Gaussian sources. Simulated ALMA observations are made utilising CASA and with a *uv*-plane analysis the size of the main source in the companion field is found.

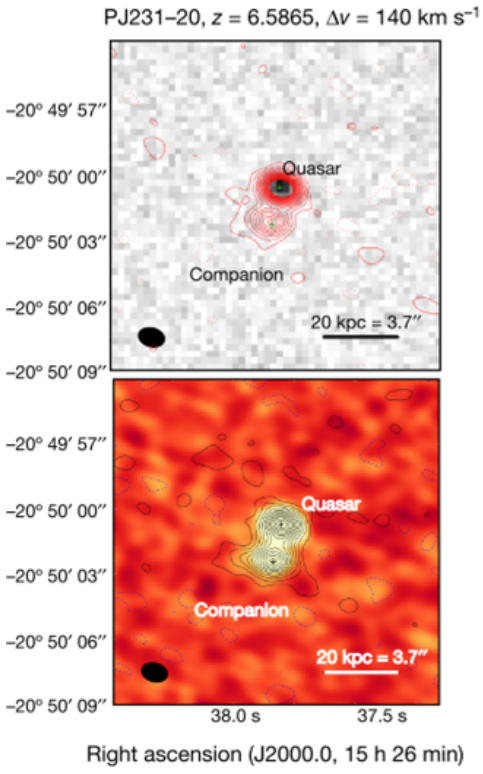


Figure 1.4: PSO J231-20 observed by Decarli et al. (2017) with ALMA.

The fields provided by Decarli et al. (2017) consist of a quasar and a highly star-forming companion at a redshift higher than 6. The star-formation rate of these companions could explain the abundance of massive galaxies at a redshift of 4. An example of one of the fields provided by Decarli et al. (2017) is shown in figure 1.4.

The cosmology used in Decarli et al. (2017) is adopted in this research, namely a Lambda cold dark matter cosmology (Planck Collaboration et al. 2015) with $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_m = 0.3$ and $\Omega_\Lambda = 0.7$.

[1]CASA can be downloaded here: `https://casa.nrao.edu/`.

2 Radio astronomy

Radio astronomy observes in a frequency range between 15MHz and 15THz (Wilson et al. 2009) and is used to observe objects such as neutron stars, black holes, quasars and active galactic nuclei, among others.

Radio astronomy allows astronomers to observe in new wavelength ranges and see new objects. A positive aspect of radio astronomy is that it can be conducted from the ground. Figure 2.1 illustrates the atmospheric opacity of the Earth's atmosphere. In the radio range, the transmission through the Earth's atmosphere is roughly 100%. This means that the atmosphere does not absorb the signal before it reaches the telescope. On the other hand, ultraviolet light can only be observed from space as the atmosphere almost fully absorbs its signal before it reaches the ground. The transmission for the submillimeter range is less than 100% but still observable from the ground. Radio interferometer ALMA (Atacama Large Millimeter/submillimeter Array) covers a wavelength range from 0.3 to 3.6 mm.

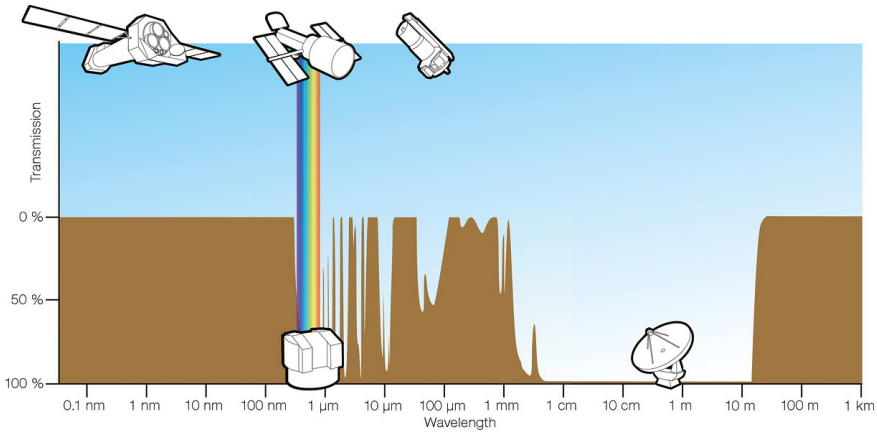


Figure 2.1: The brown curve shows the transmissivity of the Earth's atmosphere at a given wavelength. ALMA operates between 0.3 and 3.6 mm where the opacity depends on the altitude and water vapour in the atmosphere. Credit: ESA/Hubble (F. Granato)[1]

Dust particles and atmospheric gas molecules can cause Rayleigh scattering of sunlight in the ultraviolet and visible range which makes daytime observations impossible for faint objects. Another advantage of radio astronomy is that this Rayleigh scattering of sunlight does not effect the radio sky which means the radio sky is always dark.

[1]The image can be found here: `https://www.eso.org/public/images/atm_opacity/` .

This results in the possibility to make observations in the radio spectrum during the day or night (Condon and Ransom 2016).

The opacity is generally measured at zenith. Zenith is the point directly overhead a certain location. The total zenith opacity depends on a number of factors (Condon and Ransom 2016):

1. The opacity of dry air. This is nearly independent of wavelength.
2. Molecular oxygen (O_2). The permanent magnetic dipole creates rotational transitions that can absorb radio waves.
3. Water vapour. Precipitable water vapour absorbs at radio and submillimeter wavelengths. This is the main cause of the opacity of the Earth's atmosphere for submillimeter waves.

The data analysed in this research are obtained from simulations with ALMA, a radio interferometer. ALMA consists of multiple telescopes, often called antennas, that link up to work together. When using a single dish radio telescope, the resolution limiting factor is generally the diameter of the dish. As the diameter of the dish cannot be increased indefinitely, there is a limit to the angular resolution achievable with a single dish telescope. Yet with radio interferometry, the resolution limiting factor is generally the separation between each individual antenna. As the separation between antennas can amount to far greater lengths than the diameter of a single dish telescope physically can, radio interferometry is used for accurate, high-resolution observations.

This chapter will firstly discuss the types of radiation (section 2.1). Secondly, radio astronomy is explained. To fully understand radio interferometry, first an understanding of single dish radio astronomy is needed. Therefore, section 2.2 covers single dish radio astronomy and section 2.3 discusses radio interferometry. Then the ideal environment of a radio telescope/interferometer is discussed in section 2.4 and lastly radio interferometer ALMA is covered in section 2.5.

2.1 From source to image

Each object that has a temperature higher than 0 K (-273°C) emits electromagnetic radiation. Depending on the temperature of the object, it is visible in different bands of the electromagnetic spectrum. Radiation caused by the temperature of the object is called thermal radiation. There are also types of radiation not caused by temperature, which are called non-thermal radiation. The three main types of non-thermal radiation are synchrotron emission, caused by charged particles and the magnetic field of the object, Compton scattering, where a photon collides with an electron and loses energy to the electron, and stimulated emission, where a photon causes an electron to drop to a lower energy level emitting a photon with the energy difference of the two states. As this research is conducted using simulated observations of ALMA and ALMA observes mostly thermal radiation from dust and emission lines, the focus is on this.

Objects at a great distance from Earth have a higher redshift than objects closer to Earth as they are moving faster away from Earth due to the expansion of the Universe.

This redshift influences the emitted wavelengths. Therefore, high redshift objects that emit in the optical range are observed on Earth in the infrared range. Especially when looking at specific emission lines, this has to be taken into account when observing such a high redshift object. This research focuses on galaxies with a redshift between 2 and 6, which implies that wavelengths or frequencies need to be rescaled. This can be done using equation 2.1 or 2.2.

$$\lambda_{emit} = \frac{\lambda_{obs}}{1+z} \tag{2.1}$$

$$\nu_{emit} = \nu_{obs}(1+z) \tag{2.2}$$

Here λ_{emit} is the emitted wavelength, λ_{obs} the observed wavelength, z the redshift, ν_{emit} the emitted frequency and ν_{obs} the observed frequency.

2.2 Single dish radio telescope

A single dish radio telescope has a parabolic dish that focuses all incoming waves into one point above the dish, the focus. A subreflector is located at the focus, reflecting the waves to the center of the dish where the feed horn is located, often in the shape of a funnel. The narrow end of the funnel is the size of the critical wavelength of the desired channel. At the end of the funnel a receiver collects the wave and converts the electric field of the radio wave to a voltage with the same frequency. Radio sources are generally weak and therefore the voltage is amplified with a pre-amp. The resulting voltage is now mixed with another wave which is called the local oscillator. The resulting voltage now has a lower frequency, $|\nu_{signal} - \nu_{LO}|$. After another amplification, this time with an IF-amp, the signal is digitised and can then be processed (Burke and Graham-Smith 2009).

The dish needs to be very smooth when observing very small radio waves. Each little imperfection in the surface will scatter the wave away from the focus, which results in loss of efficiency due to information loss. When longer wavelengths are observed, the surface does not have to be as perfect since the waves are not scattered away by small imperfections. ALMA observes millimeter and submillimeter waves and is therefore very dependent on the surface quality of its dishes. According to the Handbook of ALMA Cycle 7[2] (Remijan et al. 2019), the surface of ALMA's antennas have a deviation of 25 or less microns away from a perfect parabola.

To examine the radiation pattern of an antenna, it is tested with a planar wave. The outcome is called the voltage reception pattern. The voltage reception pattern of the antenna is not a single lobe, but in fact consists of a main lobe and several side lobes. This implies that if an object is observed in the main lobe, unwanted signal can be detected with the side lobes. This main lobe is also referred to as the primary beam. For a circular aperture of diameter d, the main beam has a width of $1.22\lambda/d$. The feed horns used in ALMA's antennas are designed to have a nearly Gaussian primary beam and low side lobes to achieve the best resolution and sensitivity (Remijan et al. 2019).

[2]The handbook can be found here: `https://almascience.eso.org/documents-and-tools/cycle7/alma-technical-handbook`.

The angle from the on-axis pointing direction of the telescope influences the antenna response. The most accurate response is found with an on-axis incidence observation. When the incidence is off-axis interference disturbs the signal and it is not added up constructively and can even lead to destructive interference at an angle of λ/D (Remijan et al. 2019).

2.3 Interferometry[3]

Single dish radio astronomy generally has a worse angular resolution compared to optical telescopes due to its longer wavelengths. The largest steerable single dish radio telescope is the Green Bank Telescope with a diameter of 100 m[4]. Observing at a frequency of 350 GHz would result in a resolution of roughly 1.8 arcsec. On the other hand, the most compact configuration of ALMA antennas used in this research has a maximum baseline of around 160 m and consequently a resolution of approximately 1.1 arcsec for an observation at 350 GHz. The most extended configuration of ALMA antennas used in this research has a maximum baseline of approximately 3700 m and a resolution of around 0.05 arcsec for a 350 GHz observation. This is a significant improvement from the 1.8 arcsec resolution with the single dish telescope.

Hence, radio interferometry allows for a much better angular resolution than single dish radio astronomy. A radio interferometer consists of at least two radio antennas (telescopes). These antennas combine their signal, which reduces the resolution since the resolution limiting factor is now inversely proportional to the separation between the antennas rather than the diameter of the single dish. This separation is called the baseline, where the longest baseline determines the resolution. The angular resolution θ_{res} is inversely proportional to the baseline, as is shown in equation 2.3:

$$\theta_{res} = \frac{\lambda}{B}. \tag{2.3}$$

Here, λ is the wavelength and B is the distance between two antennas. The resolution is in radians. It can be seen that the shorter the wavelength the better the resolution.

This derivation follows the ALMA Cycle 7 Technical Handbook Remijan et al. (2019). Since radio sources are far from Earth, the radio waves can be approximated as planar waves which is used throughout the derivation.

Figure 2.2 shows a schematic representation of a two antenna interferometer with a baseline $\mathbf{b}$. Both antennas observe a source at position $\mathbf{s_0}$ at an angle θ from the zenith. The separation between the two antennas as perceived from the direction of $\mathbf{s_0}$ is equal to $u = b\cos\theta$. In this case, an on-axis wavefront reaches antenna 2 before it reaches antenna 1 (the white dashed lines). To reach antenna 1 the wavefront needs to travel an extra path of length $\mathbf{b} \cdot \mathbf{s_0} = b\sin\theta$. Hence there is a geometrical delay between antenna 1 and antenna 2 of $\tau_g = \frac{\mathbf{b}\cdot\mathbf{s_0}}{c}$. An artificial delay can be inserted in the signal path of antenna 2 to compensate for this geometrical delay, which allows the signals

[3]This section follows the ALMA Cycle 7 Technical Handbook Remijan et al. (2019).

[4]More information on the Green Bank Telescope can be found on the observatory's website: `https://greenbankobservatory.org/science/telescopes/gbt/`.

from both antennas to arrive with the same phase at the correlator. The power received by the correlator is proportional to the voltages squared and the power response of the antenna.

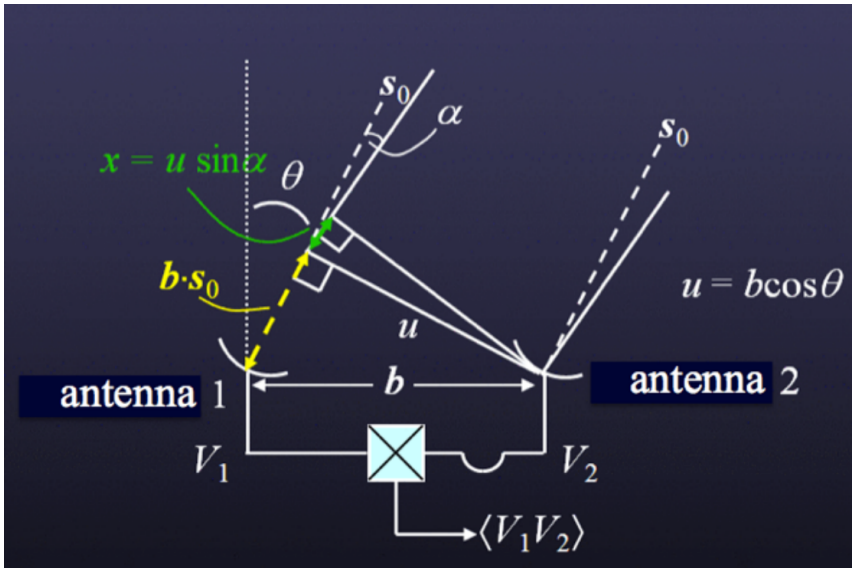


Figure 2.2: A schematic 1D representation of a two antenna interferometer separated by a baseline **b** both pointing towards a source at $\mathbf{s_0}$. The projected separation as seen from the source is u. The dashed line represents an on-axis wavefront and the solid line represents an off-axis wavefront at an angle α. The two antennas are connected to a correlator (Remijan et al. 2019).

Let us now assume the wavefront is moved off-axis by a small angle α (see the solid white lines in figure 2.2). The artificial delay implemented in the signal path of antenna 2 will not suffice to compensate for the new delay caused by the angle α. The extra path, on top of the on-axis extra path, is equal to $x = u\sin\alpha = ul$ with $l = \sin\alpha$. This extra path length causes the signals of the two antennas to have a phase difference when they arrive at the correlator. Thus the voltage response of the second antenna, V_2, can be expressed in terms of the voltage response of the first antenna, V_1, and a phase delay factor:

$$V_2 = V_1 e^{2\pi i(ul)}. \tag{2.4}$$

So far only a one dimensional case has been investigated. Expanding to a two dimensional case, introduces β, a direction on the sky orthogonal to α. The baseline now has two components b_1 and b_2, namely in the x and y direction. This gives $u = b_1\cos\theta$ and $v = b_1\cos\phi$ with ϕ the angle of the position of $\mathbf{s_0}$ orthogonal to θ. The extra path created by this second dimension is equal to $y = v\sin\beta = vm$ with $m = \sin\beta$. This second dimension can be incorporated in equation 2.4 like such:

$$V_2 = V_1 e^{2\pi i(ul+vm)}. \tag{2.5}$$

To filter out uncorrelated noise, the correlator takes the time average of the product of the signals from the two antennas:

$$\langle V_1 V_2 \rangle = \langle \int\int V_1(l,m)\,dl\,dm \int\int V_2(l,m)\,dl\,dm \rangle. \tag{2.6}$$

Recall that $l = \sin\alpha$ and $m = \sin\beta$. This equation can be split into cross terms of the coordinates l and m and terms of the same coordinates. Assuming that signals from different parts of the sky (cross terms of l and m) are incoherent, i.e. their phase has no similarities, the average of the cross terms will be zero and the input of the correlator becomes:

$$\langle V_1 V_2 \rangle = \int\int \langle V_1(l,m)^2 \rangle \mathcal{A}(l,m)\, e^{2\pi i(ul+vm)} dl\, dm\,. \tag{2.7}$$

Here $\mathcal{A}(l,m)$ is the power response of the antenna. It is known that $V^2 \propto P$ with P the power received by the correlator and $P \propto I_\nu$ with I_ν the intensity distribution on the sky. With this equation 2.7 can be rewritten as:

$$\langle V_1 V_2 \rangle \propto \int\int \mathcal{A}(l,m)\, I(l,m)\, e^{2\pi i(ul+vm)} dl\, dm\,. \tag{2.8}$$

Therefore, the correlator measures the Fourier transform of the intensity distribution on the sky, the so-called complex visibility $\mathcal{V}$:

$$\mathcal{V}(u,v) = \int\int \mathcal{A}(l,m)\, I(l,m)\, e^{2\pi i(ul+vm)} \frac{dl\, dm}{\sqrt{1-l^2-m^2}} = A\, e^{i\phi}\,, \tag{2.9}$$

with $\mathcal{V}$ a complex number that can be described by an amplitude A and a phase ϕ. The coordinates u and v correspond to the vectorial separation between each antenna pair in wavelengths. The factor $1/\sqrt{(1-l^2-m^2)}$ can be approximated to unity and are therefore neglected. The amplitude gives information about the source brightness and the phase about the location relative to the phase center at coordinates u and v.

From this, it follows that the sky brightness distribution is the inverse Fourier transform of the visibility distribution:

$$\mathcal{A}(l,m)\, I(l,m) = \int\int \mathcal{V}(u,v)\, e^{-2\pi i(ul+vm)}\, du\, dv\,. \tag{2.10}$$

The sky brightness distribution and the visibility distribution contain the same amount of information if the uv-plane is perfectly sampled.

The shorter baselines measure larger scales and the longer baselines measure smaller scales due to the inverse scaling relation between x, y and u, v of the Fourier transform. In general, the longest baseline is used to calculate the resolution. However, the actual resolution depends on the actual uv-plane coverage rather than the longest baseline. To avoid adopting an unrealistic resolution, ALMA uses the 80^{th} percentile of the baselines as a proxy for the calculation of the resolution. The resolution can be found using the relation obtained empirically with representative ALMA configuration (Remijan et al. 2019):

$$\theta_{res} \approx 0.574 \frac{\lambda}{B_{80}}\,, \tag{2.11}$$

with B_{80} the 80^{th} percentile of the baselines.

The shortest baseline determines the maximum recoverable scale, also called the largest angular scale (Remijan et al. 2019):

$$\theta_{LAS} \approx 0.6 \frac{\lambda}{B_{min}}\,, \tag{2.12}$$

with B_{min} the shortest baseline. Again, a configuration can have an unrepresentatively short baseline, which results in an overestimated largest angular scale. Thus ALMA uses the 5^{th} percentile of the baselines to determine the largest angular scale. Empirically using simulation, the following relation was found (Remijan et al. 2019):

$$\theta_{LAS} \approx 0.983 \frac{\lambda}{B_5}, \tag{2.13}$$

with B_5 the 5^{th} percentile of the baselines.

A pair of antennas gives two uv-datapoints, namely at (u, v) and at $(-u, -v)$, because the visibilities originate from a Hermitian complex-valued function. To recover the sky brightness distribution, the distribution of the visibilities across the uv-plane is needed. The better uv-coverage that can be achieved, the better the reconstruction of the sky brightness distribution. The uv-coverage can be improved by increasing the number of baselines, i.e. the number of antennas. The multiple antennas need to have different lengths and directions to prevent redundancy. The number of different baselines N is calculated using equation 2.14:

$$N = \frac{1}{2} n(n-1). \tag{2.14}$$

Here n is the number of antennas used in the configuration (Burke and Graham-Smith 2009).

The Earth's rotation aids in the increase of uv-coverage. The Earth's rotation causes the projected separation of the antenna pairs to change. This means that the antennas will have new positions and therefore new uv-points. This effect is called 'Earth rotation aperture synthesis'. Hence, repeated observations increase the uv-coverage.

Equation 2.10 allows the sky brightness distribution to be recovered from the visibilities. However, in practice it is impossible to cover the entire uv-plane. Even though an interferometer can behave as one large radio telescope with a diameter of the baseline B, there is a crucial difference. The surface area of an interferometer is not the same as one large radio telescope with diameter B. The antenna coverage is not complete. Each antenna pair contributes two points in the uv-plane. It is clear that only a fraction of the uv-plane is covered.

This introduces a limit to the details visible in the sky brightness distribution, i.e. the sky brightness distribution is limited by a minimum scale defined as the resolution. The images only contain information on the angular scales observed by the interferometer and not those that are unobserved. This is called 'spatial filtering'. Especially a lack in shortest baselines results in a low sensitivity to large-scale emission. A radio interferometer is unable to have a baseline smaller than the diameter of the antenna. Therefore, visibilities are not sampled at or near the origin of the uv-plane. These problems are called, respectively, the zero-spacing problem and short-spacing problem and they lead to a biased resulting image led by the small-scale emission of the sky brightness distribution. Thus an interferometer has a largest angular scale as mentioned before

(equation 2.12). A solution for this zero-spacing problem is adding a single antenna to the configuration and using it to sample the total power. The short-spacing problem can be solved by adding a compact component to the configuration.

Figure 2.3 shows the *uv*-plane data from three different Gaussian sources with diameters of 5, 10 and 20 arcsec. Observing with an extended configuration can lead to extended sources not being detected. When observing a 20 arcsec source with a configuration with a shortest baseline of 200 kλ, it will not be detected since the real part of the visibility is zero. Only the 5 arcsec source is properly detected. This illustrates the short-spacing problem. When observing with a single dish radio telescope, the total power is detected (at a *uv*-distance of 0 kλ) and the 20 arcsec source is still detected.

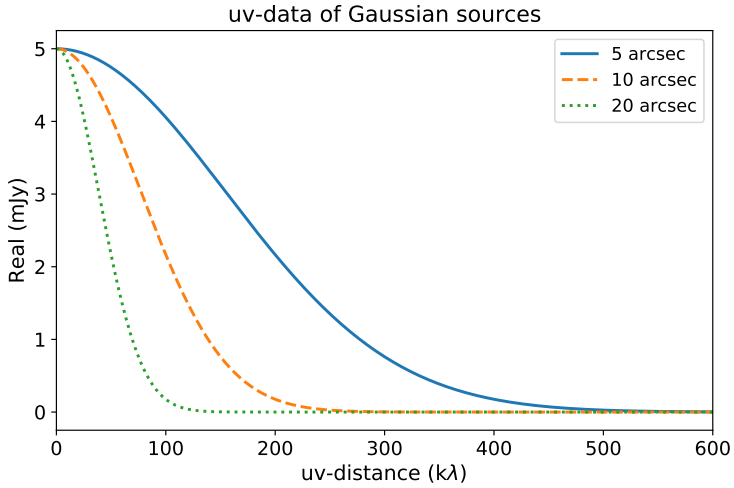


Figure 2.3: *uv*-plane data of the real part of the visibility for three different mock Gaussian sources with FWHM of 5 arcsec (blue solid line), 10 arcsec (orange dashed line) and 20 arcsec (green dotted line).

This limit to the details visible in the sky brightness distribution needs to be taken into account when the sky brightness is recovered from the visibilities. Imagine that the configuration being used has M baselines, which corresponds to $2M$ data points in the *uv*-plane. This can be written in a sampling distribution with Dirac delta functions like so:

$$B(u,v) = \sum_{k=1}^{2M} \delta(u-u_k, v-v_k)\,. \tag{2.15}$$

With this sampling distribution the dirty image can be found:

$$I_V^D(l,m) = \int\int \mathcal{V}(u,v)B(u,v)e^{2\pi i(ul+vm)}\,du\,dv\,. \tag{2.16}$$

Taking the inverse Fourier transform of the dirty image and following the convolution theorem ($\mathcal{F}\{f*g\} = \mathcal{F}\{f\}\cdot\mathcal{F}\{g\}$)[5], equation 2.16 can be rewritten as:

$$I^D = b(l,m) * I(l,m)\mathcal{A}(l,m)\,, \tag{2.17}$$

[5]Throughout this paper $\cdot$ will denote a product and $*$ will denote a convolution.

with $b(l,m) = \mathcal{F}^{-1}\{B(u,v)\}$ the point spread function also known as the synthesised beam or dirty beam.

Figure 2.4 shows the relation between the dirty beam, visibilities, true sky image and dirty image (Remijan et al. 2019).

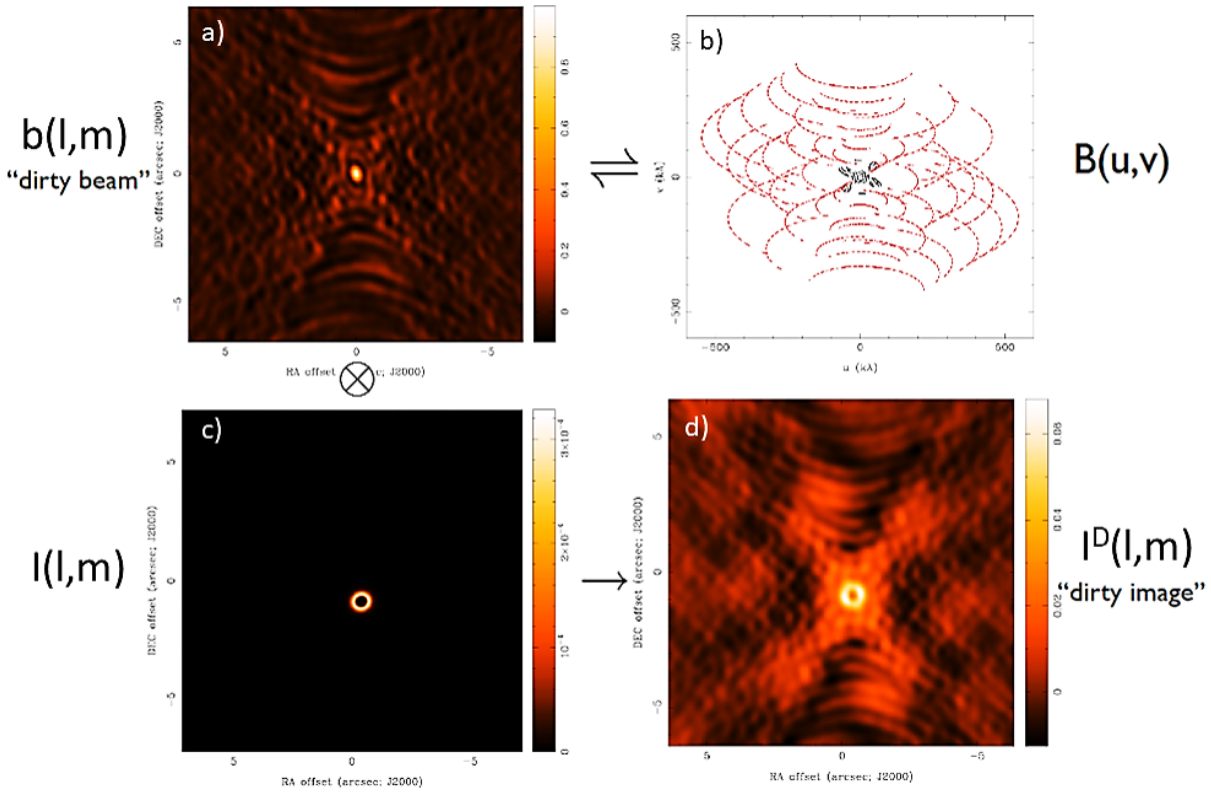


Figure 2.4: Figure a shows an example of a dirty beam. Figure b shows the corresponding uv-plane coverage. The red dots are from an extended configuration and the black dots are from a compact configuration. Figure c shows an example of a true sky image. Figure d shows the dirty image obtained from the convolution of $I(l,m)$ (figure c) with $b(l,m)$ (figure b). The antenna power response, $\mathcal{A}(l,m)$ has been neglected in this figure (Remijan et al. 2019).

An advantage of working with the uv-plane data rather than the image plane data, is that the image plane data needs to be cleaned before analysis can be done. The weighting chosen for the cleaning gives a different outcome image[6]. To do analysis in the uv-plane, no cleaning is required and therefore there is no bias.

2.4 Ideal environment

Even though the majority of the incoming radio and submillimeter waves are transmitted through the Earth's atmosphere, still a fraction gets absorbed. To minimise this absorption, an observation site at high altitude is preferable. Consequently, the radio waves have to travel a small distance through the Earth's atmosphere and more radio

[6]The different weightings are explained in section 3.3.

waves will reach the telescope. Hence, the observations made at high altitude are less noisy than observations made at low altitude.

Another aspect of the Earth's atmosphere that influences the transmissivity is water vapour. The small drops of water absorb the radio and submillimeter waves. As a result the ideal observing site is extremely dry throughout the year.

The observation site needs to be remote to avoid background radio waves from civilisation. Radio frequencies used for telecommunication interfere with observations in the radio range.

When a new observation site is needed, extensive research is conducted to find the ideal spot. ALMA is located in the middle of the Atacama desert on the Chajnantor Plateau in northern Chile. This is above 5000m altitude where the atmospheric conditions are arguable the best achievable on Earth. This area is one of the driest areas on Earth with exceptionally clear sky conditions. The area is easily accessible year-round and close to large cities which can provide energy and services. The latitude (23°south) allows observations of the southern sky as well as a large portion of the northern sky. The Chilean government declared this area a radio quiet zone (radio emission above 31.3 GHz is prohibited) and has light pollution protection laws in place (Bustos et al. 2014).

2.5 ALMA

The radio telescope this research is based on is the Atacama Large Millimeter/submillimeter Array or ALMA for short (figure 2.5). ALMA is a collective effort of the European Southern Observatory (ESO), the National Science Foundation (NSF) of the USA and the National Institute of Natural Sciences (NINS) of Japan and the Republic of Chile and has been active since 2013. ALMA is located in the middle of the Atacama desert on the Chajnantor Plateau. All ALMA operations are done in concession by the government of Chile[7].

ALMA is an interferometer with 66 antennas that can be moved with extreme precision to make new configurations with properties useful for each desired research. Fifty antennas have a diameter of 12 m and form the 12-m Array used for high resolution observations. The 12-m Array is accompanied by the Atacama Compact Array (ACA), also known as the Morita Array. This has twelve antennas with a diameter of 7 m closely together (the 7-m Array) and four antennas with a diameter of 12 m for single dish observations (or Total Power observations) called the TP Array. The TP Array solves the zero-spacing problem and the 7-m Array samples with baselines between 9 and 30 m to bridge the gap to the smallest baseline of the 12-m Array and solve the short-spacing problem (Remijan et al. 2019).

This research looks at 20 different ALMA antenna configurations from the 12-m Ar-

[7]More information on ALMA and the image can be found on the observatory's website: `https://www.almaobservatory.org/en/home/`.

ray listed in CASA[8].

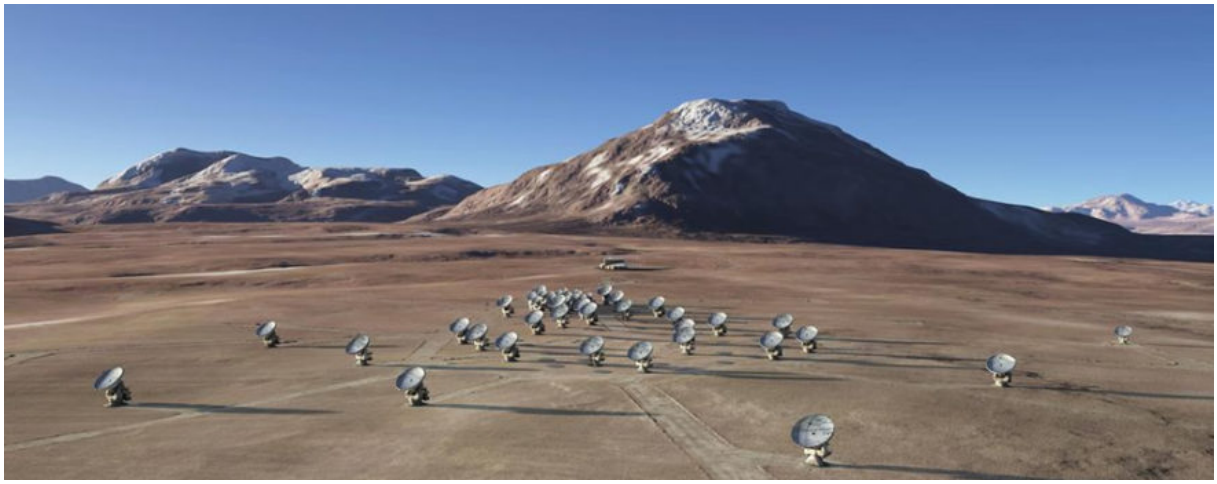


Figure 2.5: The Atacama Large Millimeter/submillimeter Array (ALMA) located in Chile[7].

[8]More information on the used antenna configurations can be found here: `https://casaguides.nrao.edu/index.php/Antenna_Configurations_Models_in_CASA`.

3 Methods

Using simulated Gaussian sources and observations of known SMG companion fields and quasar companion fields, different properties of galaxies are researched. The next sections will cover the methodology used in this research.

In this research, the Common Astronomy Software Applications package (CASA) has been utilised[1]. CASA's main goal is to help with the post-processing of radio astronomical data (McMullin et al. 2007).

3.1 Single sources

This research used different sources. The first part of the analysis is performed using the source NGC 4038-4039, also known as Antennae (see figure 3.1). This is a set of two gas-rich galaxies interacting and it is the nearest galaxy merger at only 22 Mpc from Earth (Schweizer et al. 2008).

Figure 3.1: An image of NGC 4038-4039 (Antennae). Credit: NASA, ESA, and the Hubble Heritage Team (STScI/AURA)-ESA/Hubble Collaboration[2]

The second part of the research analyses simulated Gaussian sources. These sources are made using a Python code. This code allowed the user to choose the position of the source in the sky, the flux and the Full Width at the Half Maximum (FWHM) of the major and minor axis of the source. The source is positioned in the center of the field.

[1]This research used CASA version 5.3.0-143. More information on CASA can be found on its website: `https://casa.nrao.edu/` .

[2]The image can be found here: `https://www.nasa.gov/multimedia/imagegallery/image_feature_1086.html`

With this, different sources with different shapes were created. The properties of the different Gaussian sources made in this research are displayed in table 3.1.

Table 3.1: Properties of the simulated Gaussian sources made in this research using a Python code.

Shape	Flux (mJy)	Major axis FWHM (″)	Minor axis FWHM (″)	Position (J2000)
Circular	5.0	0.2	0.2	03h31m24.721s -27d50m47.08s
Elliptical	5.0	0.3	0.15	03h31m24.721s -27d50m47.08s
Very elliptical	5.0	0.4	0.1	03h31m24.721s -27d50m47.08s

3.2 Setting up the simulations

Utilising CASA functions `simobserve` and `simanalyze` allowed for the creation of ALMA simulations with different antenna configurations. The function `simobserve` creates simulated observations of the skymodel given. The skymodel is a .fits file containing an image of the desired source. This can either be an image of an observation or a simulated field. Many different parameters can be set, including the *skymodel* (a model image to observe), the *inbright* (the surface brightness of the brightest pixel), the *incenter* (the frequency of the center of the channel[3]), the *inwidth* (the channel width), the *integration* (the integration time), the *total time* (the total time of observation) and the *antennalist* (the positions of the interferometer antennas). The antenna configurations used in this research are 20 configurations provided by CASA. These have 50 ALMA antennas with a diameter of 12 m (alma.out01.cfg - alma.out20.cfg)[4].

The different antenna configurations have different sets of baselines. The direction and length of the baselines determine the resolution and largest angular scale of the observation. The lengths of the baselines of the different configurations are visualised in figure 3.2. Here the minimum (0^{th} percentile), 5^{th} percentile, 80^{th} percentile and maximum (100^{th} percentile) of the baseline is shown. As can be seen between configuration 11 and 12 and between 16 and 17, there is a jump in the shortest baseline. Figure 3.3 shows the largest angular scale (left) as determined with the minimal baseline and the 5^{th} percentile with the corresponding equations, 2.12 and 2.13 respectively. On the right the resolution is shown as determined with the 80^{th} percentile and the maximum baseline with the corresponding equations, 2.11 and 2.3 respectively. The largest angular scale of the minimum baseline shows the same jumps visible in figure 3.2. It can be seen that the resolutions range from 1.1 to 0.05 arcsec.

[3]The channel is the frequency range in which observations are performed.

[4]More information on CASA functions can be found in the CASA Task Reference Manual: `https://casa.nrao.edu/docs/taskref/TaskRef.html#TaskRefli1.html`.

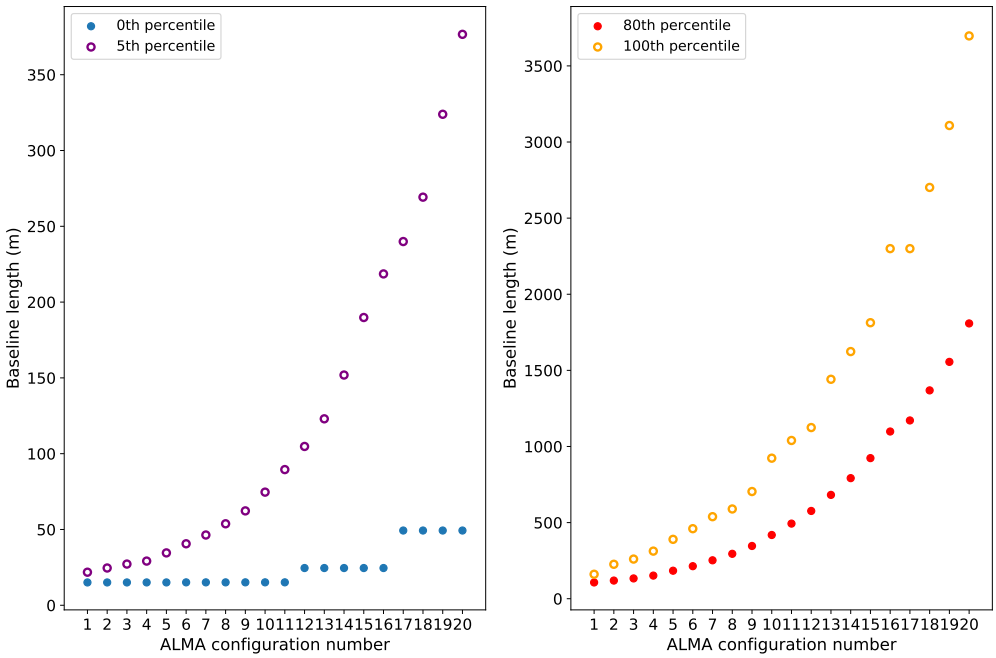


Figure 3.2: The baselines of the 20 ALMA antenna configuration used in this research showing the minimum baseline (0^{th} percentile) (blue closed circles), the 5^{th} percentile (purple open circles), the 80^{th} percentile (red closed circles) and the maximum baseline (100^{th}) (orange open circles).

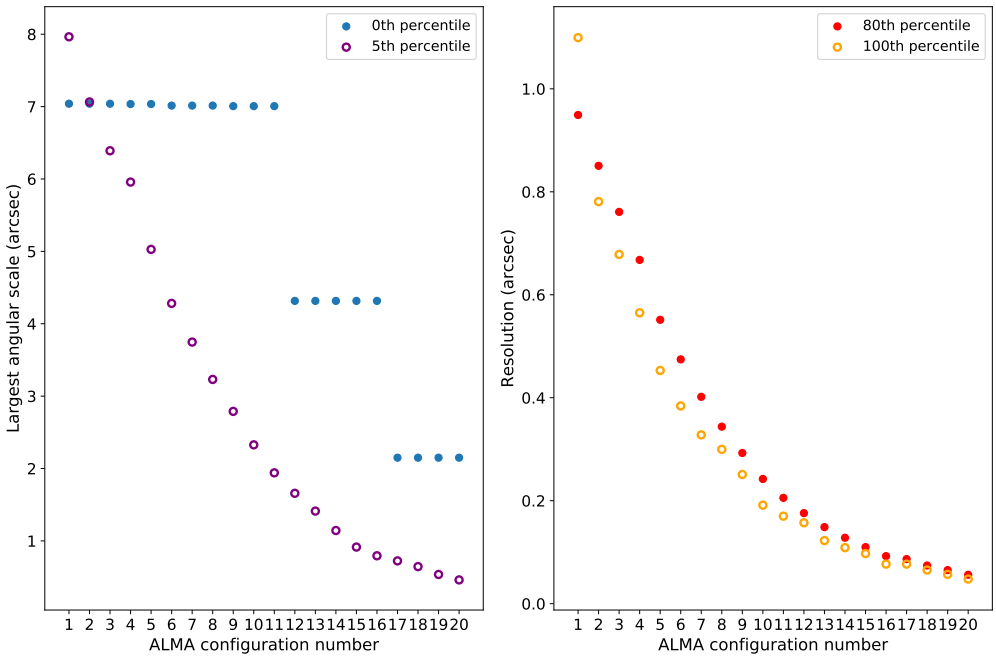
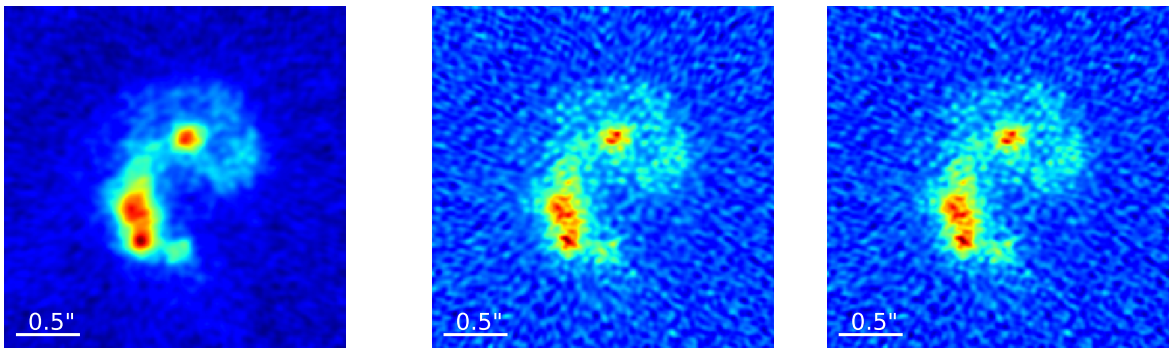


Figure 3.3: The largest angular scale (left) of the 20 ALMA antenna configurations used in this research calculated with equation 2.12 for the minimum baseline (0^{th} percentile) (blue closed circles) and with equation 2.13 for the 5^{th} percentile (purple open circles). The resolutions (right) of the 20 ALMA antenna configurations used in this research calculated with equation 2.11 for the 80^{th} percentile (red closed circles) and with equation 2.3 for the maximum baseline (100^{th} percentile) (orange open circles).

To analyse the simulated observation created with `simobserve`, the function `simanalyze` is used. Again, different parameters can be set, including *niter* (the maximum number of cleaning iterations), the *threshold* (the flux level to stop cleaning) and the *weighting* (the weighting to apply to the visibilities[5]). These analysed simulated observations were then saved to .fits files for further analysis. For each source 20 different observations were made with the 20 different ALMA antenna configurations.

3.3 Different cleaning methods of CASA

To analyse observation data in the image plane, first the data need to be cleaned. CASA has a built in function with six different weightings for this cleaning method. This research cleans the data using the function `simanalyze`. Three of these methods are explored in this section, namely natural, uniform and Briggs. Figure 3.4 shows the Antennae galaxies cleaned with each of the three different weightings. Note that each of these figures is made with the same data set. Only the weighting is changed.



(a) ***(b)*** ***(c)***

Figure 3.4: Antennae galaxies cleaned using three different weightings of CASA, namely natural (left, 3.4a), uniform (middle, 3.4b) and Briggs with R=0 (right, 3.4c).

3.3.1 Natural

If the weighting is chosen to be natural, only the data weights are taken into account. These weights are equal to the inverse of the noise variance of the visibility. This can be shown using equation 3.1.

$$w_{n,i} = \omega_i = \frac{1}{\sigma_i^2} \tag{3.1}$$

Here $w_{n,i}$ is the imaging weight of sample i, ω_i is the data weight and σ_i is the RMS (Root Mean Square) noise on visibility i.

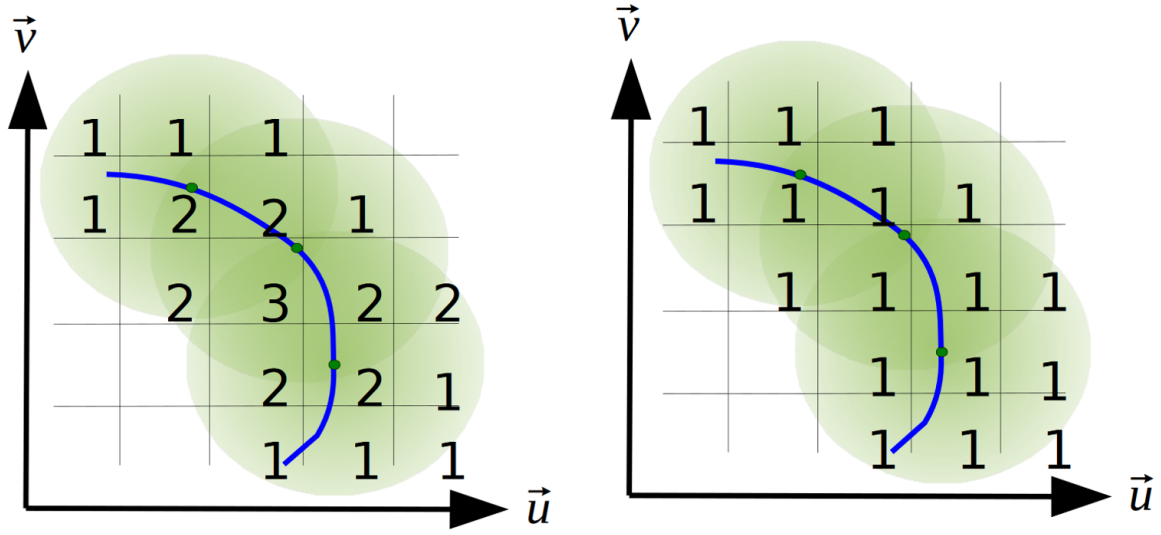
A natural weighting should result in an image with the the highest Signal to Noise Ratio (SNR), i.e. the lowest noise. To achieve this, more weight is generally put on the shorter baselines corresponding to the larger spatial scales, which results in images

[5]The different weightings of CASA are discussed in section 3.3.

with a poorer angular resolution. A natural weighting maximises the point source sensitivity, yet for detailed or more complicated objects details will get lost. This is visible when figure 3.4a and 3.4b are compared.

3.3.2 Uniform

When the weighting is set to uniform, the data weights are calculated the same way as with a natural weighting, i.e. the inverse of the noise variance of the visibility. To get 'uniform' imaging weight, the data is gridded to a number of cells in the uv-plane and then re-weighted. This way, the low weighted data gets more influence. Usually this is the data originating from longer baselines. Therefore, this sharpens the resolution but increases the RMS image noise.



(a) ***(b)***

Figure 3.5: A visual representation of a natural weighting (left, 3.5a) and a uniform weighting (right, 3.5b)[6].

Figure 3.5 shows the visual representation of a natural (left, 3.5a) and uniform (right, 3.5b) weighting. In figure 3.5a it can be seen that for a natural weighting different cells have different weightings, whereas for a uniform weighting all cells have the same weight independent of the data weights.

A uniform weighting is shown in equation 3.2;

$$w_{u,i} = \frac{\omega_i}{W_k}. \tag{3.2}$$

Here $w_{u,i}$ is the imaging weight of sample i, ω_i is the inverse of the variance. The data is then gridded onto a grid with a uv cell size of 2/(Field Of View). This gives the gridded weights W_k.

[6]The images can be found here: `https://science.nrao.edu/science/meetings/2017/vla-data-reduction/DRW2017_Imaging_RVU.pdf`

3.3.3 Briggs

The Briggs weighting is a flexible weighting developed by Dan Briggs (Briggs 1995). When using the Briggs weighting, the robustness needs to be specified using the robust parameter R. R can take on any value between -2, closest to a uniform weighting, and 2, closest to a natural weighting. When R takes the value 0 it gives a good trade-off between sensitivity and resolution. The weighting scheme is:

$$w_{B,i} = \frac{\omega_i}{1 + W_k f^2}, \tag{3.3}$$

with $w_{B,i}$ the imaging weight of sample i, ω_i is the inverse of the variance and W_k the gridded weights as defined in section 3.3.2. f is defined as follows:

$$f^2 = \frac{(5 \cdot 10^{-R})^2}{\frac{\sum_k W_k^2}{\sum_i \omega_i}}, \tag{3.4}$$

with R the robust parameter. It is evident that when R is maximum, f^2 is smallest and therefore, W_k doesn't have a large influence on $w_{B,i}$ and $w_{B,i}$ can approximately be reduced to a natural weighting (equation 3.1). The same can be done for minimum R and it can be found that then $w_{B,i}$ can approximately be reduced to a uniform weighting (equation 3.2).

Figure 3.4c shows a simulated ALMA observation of the Antennae galaxies cleaned with a Briggs weighting with $R = 0$. Comparing this with figure 3.4a and 3.4b, it can be seen that this weighting creates an image between natural and uniform. Although the cleaned images for uniform and Briggs do look alike, their difference is non-zero.

3.4 Calculating the flux

One of the properties of SMGs that is examined is the flux of the source. The flux is calculated using Python. As the configurations have different baseline collections and therefore different resolutions and largest angular scales, each of the simulated observations has a different image.

A box found using DS9[7] is used to select the source in the image. The box is selected such that the largest source is fully enclosed and the included background is minimised. To be able to compare the flux of the sources with different antenna configurations, the same box is used for all different antenna configurations per source. Figure 3.6 shows how the box would be choosen. The white dashed line represents the box.

[7]The program DS9 shows .fits images and allows regions to be made. More information on DS9 can be found on their website: `http://ds9.si.edu/site/Home.html`.

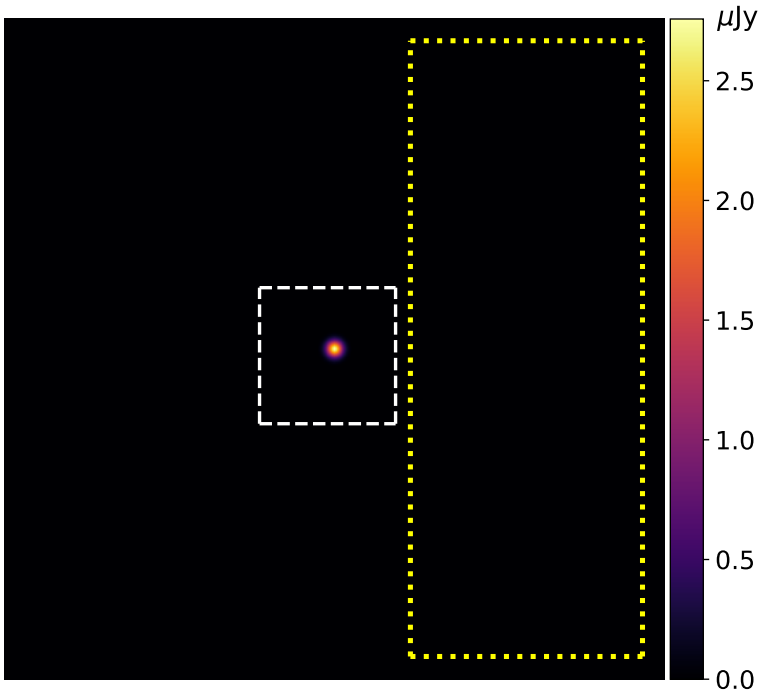


Figure 3.6: The field of the circular Gaussian is shown with the box used to calculate the flux indicated with the white dashed lines and the box used to calculate the RMS noise indicated with the yellow dotted lines.

The flux is then calculated by summing the values of each of the pixels inside the box. The image is given in Jy per beam, so a conversion for this needed to be implemented to compare the fluxes as the beam is different for each configuration. Flux density is generally given in jansky (Jy): 1 Jy = 10^{-26} W m^{-2}Hz^{-1}. The flux is found using equation 3.5:

$$F = \frac{\sum_{box} f_{pixel}}{A_{beam}}. \tag{3.5}$$

Here F is the flux in Jy, f_{pixel} is the flux of a pixel in Jy·beam^{-1} and A_{beam} is the beam area dependant on the configuration and can be calculated using equation 3.6. The sources that are observed are larger than the area of the telescope beam. This requires integration over the telescope beam to measure the flux. This results in the following relation: $F \propto A_{beam}^{-1}$. As mentioned before, the flux is calculated for every source with the same box for each configuration to allow comparison. The beam size for each configuration is calculated like such:

$$A_{beam} = \frac{\pi \cdot \mathrm{B_{MAJ}} \cdot \mathrm{B_{MIN}}}{4\ln(2) \cdot l^2}. \tag{3.6}$$

Here $\mathrm{B_{MAJ}}$ is the FWHM of the major axis of the beam as given by CASA, $\mathrm{B_{MIN}}$ is the FWHM of the minor axis of the beam as given by CASA and l is the length of the pixel edge found with DS9.

The uncertainty of the flux is calculated using a box only containing background (the yellow dotted box in figure 3.6), no pixels of the source are included. The final Root Mean Square (RMS) noise is calculated using equation 3.7.

$$\text{RMS} = \sqrt{\frac{\sum_{BGbox}(f_{pixel} - \mu_f)^2}{N_{BG}}} \cdot \frac{\sqrt{N_{source}}}{A_{beam}} \tag{3.7}$$

Here f_{pixel} is the flux of a pixel inside the background box and μ_f is the mean value of the flux of the pixels inside the background box both in Jy·beam^{-1}, N_{BG} is the number of pixels in the box of the background, N_{source} is the number of pixels in the box of the source and A_{beam} is the beam area as calculated using equation 3.6.

The resolution is set to be the FWHM of the major axis of the beam.

3.5 Calculating the source size

The source major and minor axis are found with the CASA function `imfit`. The function `imfit` fits elliptical Gaussian components on a selected region of the image and returns amongst others the Full Width at the Half Maximum (FWHM) of the major and minor axis of these fits deconvolved from the beam and the corresponding uncertainties.

The size is found using the equation for the area of an ellipse:.

$$A = \frac{1}{4}\pi ab\,. \tag{3.8}$$

Here a is the length of the major axis of the source and b is the length of the minor axis of the source. Since CASA gives the FWHM of the major and minor axis, the the size is calculated with these FWHMs. Throughtout this paper the size is calculated with the FWHM of the major and minor axis.

Note that `imfit` did not converge for all observations and those are excluded from the data.

3.6 *uv*-fitting

Since the simulated observations are performed with ALMA, a radio interferometer, the visibilities are obtained.[8] These are Fourier transformed to create the image in the image plane. To evaluate the effect of the different ALMA configurations on the *uv*-plane data and in particular the size of the source, a *uv*-fitting is applied to the visibilities.

Since the sources that are examined in this research are all Gaussians (as they are set to be Gaussians), the Fourier transform is also a Gaussian[9]. In the image plane the source can be described as follows:

[8] Visibilities are explained in further detail in section 2.3.

[9] The full Fourier transform is covered in section A.1.

$$f(x,y) = A_{xy}\, e^{\left(-\frac{x^2+y^2}{2\sigma_{xy}^2}\right)}. \tag{3.9}$$

Note that here it is assumed that the source is circular. In equation 3.9 A_{xy} is the amplitude of the Gaussian in the image plane, x and y are the coordinates used in the image plane and σ_{xy} determines the width of the Gaussian.

The uv-plane data is obtained from the measurement set created by `simobserve` when the simulated observation was made. To analyse the uv-data from the source, it is radially binned. The data is divided into bins with a binsize of n kλ (kilo wavelength). Here n is the number of kλ the binsize is and λ is the wavelength of the observation. In each bin, the mean of the real part of the visibility is taken and its corresponding mean of the uv-distance. The error on each bin is the standard deviation of all the data points in the bin. For the uv-plane analysis to work, the source needs to be in the center of the image as only then the Fourier transform of the Gaussian is a Gaussian. These averaged data will be called the binned data from now on.

A Gaussian is fitted to the binned data (equation 3.10). This is the Fourier transformed Gaussian from equation 3.9 since the visibilities are the Fourier transform of the image plane data[10].

$$F(u,v) = A_{uv}\, e^{-2\pi^2\sigma_{xy}^2(u^2+v^2)} \tag{3.10}$$

Here is A_{uv} the amplitude of the Gaussian in the uv-plane, σ_{xy} is the same as in equation 3.9 of the image plane and u and v are the coordinates in the uv-plane corresponding to the vectorial separation between each antenna pair.

Using the Python function `curve_fit` the fit parameters A_{uv} and σ_{xy} and their uncertainties are found. An example of such a fit is shown in figure 3.7. `curve_fit` allows for initial guesses for the fit parameters to be put in. The initial guesses are set to be the predetermined parameter values from the simulated observations. `curve_fit` gives a weighting to each bin. When only one data point is in the bin the standard deviation is zero. As the weighting is determined by the inverse of the standard deviation, this causes a division error. To avoid this, the standard deviation is set to 10^{-10} mJy when its initial value was zero. This is a typical value for the error. When `curve_fit` does not converge or produces a fit that is clearly not in line with the data (see figure 3.8 for an example of a bad fit), this is disregarded from the further analysis.

As the sources are perfect Gaussians, the amplitude should consist only of a real part of the visibility. Therefore, only the real part of the visibility is fitted.

As the relation between σ_{xy} and FWHM_{xy} is known to be: $\mathrm{FWHM}_{xy} = 2\sqrt{2\ln 2}\,\sigma_{xy}$, the FWHM of the source in the image plane can be found[10]. This can be compared to the FWHM found in the image plane analysis using `imfit`.

[10]The full Fourier transform is covered in section A.1.

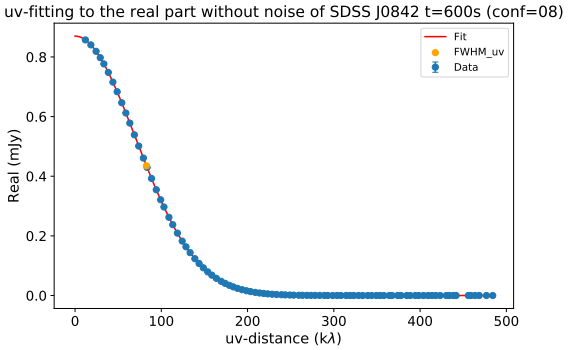


Figure 3.7: An example of a fit from a simulated ALMA observation with configuration 8 of 600 s of source SDSS J0842 for data without noise.

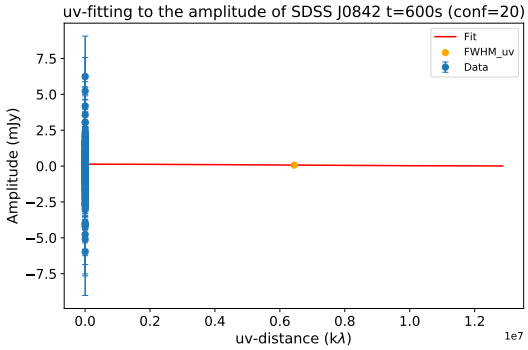


Figure 3.8: An example of a fit that is discarded from a simulated ALMA observation with configuration 20 of 600 s of source SDSS J0842 for data with noise.

3.7 Companion sources

In addition to investigating single Gaussian sources, SMG companion fields and quasar companion fields are also examined to find the influence of the companion sources on the observed FWHM of the main source. To achieve this, ten companion fields are simulated, five of which are from ALESS and are originally observed by Hodge et al. (2013) and Wardlow et al. (2018) and one is from the Ultra Deep Survey (UDS) from the SCUBA-2 Cosmology Legacy Survey (S2CLS) originally observed by Simpson et al. (2015) and Wardlow et al. (2018).

The S2CLS is a survey done with the SCUBA-2 submillimeter camera on the James Clark Maxwell Telescope (JCMT). The aim of the survey is to observe formation of massive galaxies and black holes in wavebands 450 and 850 μm. The survey consists of two parts. The 850 μm survey is around 35 square degrees. The 450 μm survey is deeper and of 1.3 square degrees in the GOODS fields, UKIDSS UDS and COSMOS regions (Geach et al. 2017).

The remaining four companion fields consist of a quasar and a companion originally observed by Decarli et al. (2017). Companions of quasars at a redshift higher than 6 (which these fields are) are highly star-forming. These companion galaxies show similarities with host galaxies of quasars. However, these host galaxies also host accreting supermassive black holes and our four targets do not show any evidence for such a black hole. The rapid star formation of these types of galaxies in the early Universe could account for the abundance of massive galaxies at $z \approx 4$ (Decarli et al. 2017).

3.7.1 ALESS and UDS companion fields

To evaluate the influence companion sources have on the *uv*-fitting of a Gaussian profile to the main object, simulated observations of fields with companion sources are needed. The same code as in section 3.2 to make a single Gaussian source is used. However, for these simulations two or more Gaussian sources are put into the same field. These sources are created from ALESS (Hodge et al. 2013; Wardlow et al. 2018)

and from S2CLS (Simpson et al. 2015; Wardlow et al. 2018) and the properties of the objects are adopted in the simulations. The image is made such that the main source is in the center, generally the brightest source. The properties of the sources used in this research are shown in table 3.2.

These simulated fields were then observed with ALMA simulations with a frequency of 350 GHz using `simobserve` and `simanalyze` as in section 3.2. The same *uv*-plane analysis as in section 3.6 is used to fit a single Gaussian to the data. As the aim is to investigate the influence of companion sources on the size of the main source, a single Gaussian is fitted and not multiple. Since the behaviour of a single Gaussian is known from the previous part, fitting the single Gaussian will shows the difference created by the companion sources. The initial guesses of `curve_fit` are set to be equal to the predetermined values of the main source. The main source is simulated as a perfect Gaussian, so the amplitude should consist only of a real part of the visibility. However, the observation of the source and the noise accompanying it as well as the companions could add an imaginary component. To investigate this, the *uv*-fitting is not only performed w.r.t. the real part of the visibility but also w.r.t. the amplitude. This allows the real part of the visibility and the amplitude to be compared and therefore the contribution of the imaginary part can be investigated. The fits that did not converge (e.g., figure 3.8) are again disregarded from the further analysis.

The FWHM of the sources of the ALESS survey provided by Wardlow et al. (2018) and Hodge et al. (2013) is estimated according to Rivera et al. (2018). Analysing four ALESS sources with a redshift between 2.1 and 2.9 with a similar *uv*-plane analysis as used in this research a FWHM of their ALESS SMG sources was found between 0.8-1.6 arcsec for their CO(3-2) data. The sources are assumed to be circularly symmetrical, i.e. the major and minor axis are identical, and a FWHM of 0.8 arcsec is adopted as FWHM for a frequency of 350 GHz.

Not all sources provided by Wardlow et al. (2018) are observed in 870 μm but at 3.3 mm. To be able to combine the sources in one field, the flux of all the sources needs to be observed in the same wavelength. Therefore, the flux at 3.3 mm was rescaled to a flux at 870μm. This is done with a modified blackbody spectrum (equation 3.11) obtained from Casey et al. (2014):

$$S(\nu, T) \propto \frac{(1 - e^{-\tau(\nu)})\nu^3}{e^{h\nu/kT} - 1}. \tag{3.11}$$

Here $S(\nu, T)$ is the flux density in Jy, ν is the frequency of the observation, $\tau(\nu)$ is the optical depth (transparancy of a medium), h is Planck's constant; $6.626 \cdot 10^{-34}$ m^2 kg s^{-1}, k is the Boltzmann constant; $1.381 \cdot 10^{-23}$ m^2 kg s^{-2} K^{-1} and T is the temperature of the dust in the galaxy. Da Cunha et al. (2015) found that an average typical ALESS SMG (of a sample of 99 SMGs) has a luminosity-averaged dust temperature of 43 ± 2 K. The optical depth can be calculated using $\tau(\nu) = (\nu/\nu_0)^\beta$ with ν_0 the frequency where the optical depth is unity and β the spectral emissivity index and is assumed to be 1.5 (empirically found between 1 and 2) (Casey et al. 2014). With the assumption that SMGs are optically thin, the $(1 - e^{-\tau(v)})$ term can be reduced to ν^β. This assumption is valid for rest-frame wavelengths longer than 450 μm.

Table 3.2: Properties of the simulated SMG companion fields from ALESS (Hodge et al. 2013; Wardlow et al. 2018) and S2CLS (Simpson et al. 2015; Wardlow et al. 2018).

Source	Position (J2000)	S_{870} (mJy)	FWHM (arcsec)
ALESS41.1[c]	03h31m10.07s -27d52m36.7s	4.9	0.8
ALESS41.3[c]	03h31m10.30s -27d52m40.8s	2.7	0.8
ALESS41.C[d]	03h31m09.81s -27d52m25.4s	3.2	0.8
ALESS49.1[c]	03h31m24.72s -27d50m47.1s	6.0	0.8
ALESS49.2[c]	03h31m24.47s -27d50m38.1s	1.8	0.8
ALESS49.C[d]	03h31m24.58s -27d50m43.4s	1.3[a]	0.8
ALESS49.L[d]	03h31m24.72s -27d50m43.7s	1.3[a]	0.8
ALESS71.1[c]	03h33m05.65s -27d33m28.2s	2.9	0.8
ALESS71.3[c]	03h33m06.14s -27d33m23.1s	1.4	0.8
ALESS75.1[c]	03h31m27.19s -27d55m51.3s	3.2	0.8
ALESS75.2[c,f]	03h31m27.67s -27d55m59.2s	5.0	0.8
ALESS75.4[c]	03h31m26.57s -27d55m55.7s	1.3	0.8
ALESS75.C[d]	03h31m26.65s -27d56m01.1s	1.0[a]	0.8
ALESS87.1[c]	03h32m50.88s -27d31m41.5s	1.3	0.8
ALESS87.3[c]	03h32m51.27s -27d31m50.7s	2.4	0.8
ALESS87.C[d]	03h32m50.65s -27d31m34.9s	1.8[a]	0.8
ALESS87.L[d]	03h32m52.42s -27d31m49.1s	1.8[b]	0.8
UDS306.0[e]	02h17m17.07s -05d33m26.6s	8.3	0.8
UDS306.1[e]	02h17m17.16s -05d33m32.5s	2.4	0.8
UDS306.2[e]	02h17m16.81s -05d33m31.8s	2.3	0.8
UDS306.L[d]	02h17m17.10s -05d33m31.5s	0.7[a]	0.8

[a] The upper limit was given and is adopted as the flux for the simulations.
[b] This flux is rescaled using equation 3.12.
[c] The information on these sources were provided by Hodge et al. (2013).
[d] The information on these sources were provided by Wardlow et al. (2018).
[e] The information on these sources were provided by Simpson et al. (2015).
[f] ALESS 75.2 is a less-reliable Supplementary source from the Hodge et al. (2013) catalogue.

As only the rescaling factor $S(\nu, T)$ and the observation flux is known, the rescaling needs to be done with equation 3.12:

$$S_{new} = \frac{S_{old}\tilde{S}(\nu_{new}, T)}{\tilde{S}(\nu_{old}, T)}. \tag{3.12}$$

Here S_{new} and S_{old} are the fluxes of the rescaled flux and the observation, respectively, and $\tilde{S}(\nu, T)$ is the proportionality function from equation 3.11.

Substituting equation 3.11 into equation 3.12 and using the assumption that SMGs are optically thin gives the following:

$$S_{new} = S_{old}\frac{\nu_{new}^{3+\beta}}{\nu_{old}^{3+\beta}}\frac{e^{h\nu_{old}/kT} - 1}{e^{h\nu_{new}/kT} - 1}. \tag{3.13}$$

The temperature, T, used is 43 K from Da Cunha et al. (2015).

Some sources were given an upper limit for the flux observed at 870μm. If the rescaled flux was greater than the upper limit provided by Wardlow et al. (2018), the upper limit was adopted as the 870 μm flux (indicated with a superscript a in table 3.2). The flux that is rescaled using equation 3.13 is indicated with a superscript b.

The properties shown in table 3.2 are used to make images of the six fields. A single Gaussian is fitted to the main source (X.1 for all fields except for UDS where it is X.0). The fitting was executed to both the real part of the visibility and the amplitude of the *uv*-plane data. Since observations are simulated, the noise an observer would get when observing with ALMA is added by CASA during the `simobserve` procedure. The output of `simobserve` provides a measurement set for both the clean data (without added observation noise) and noisy data (with added observation noise). For each field the data without noise and the noisy data is fitted to examine if the noise influences the fitting of the main sources and if the influence of the companion sources might be overpowered by the noise.

3.7.2 Quasar companion fields

Another set of companion fields is simulated in this research. This set is acquired from Decarli et al. (2017) and is visualised in table 3.3. These fields contain a quasar and a (highly star-forming) companion galaxy. The FWHMs of these sources are given in kpc. Since it is assumed that the sources are circularly symmetric, the major and minor axis are set to be equal. Using the cosmology calculator of Wright (2006), this is converted to arcsec applying the cosmology used in Decarli et al. (2017)[11]. The redshift used in the calculator are shown in table 3.3. The FWHM of the diameter of the sources in arcsec are also displayed in table 3.3.

Again the same code is used to make the simulated fields and simulated observations. For each field the quasar was set in the center of the image. The simulated observations are made with `simobserve` and `simanalyze`. Each field is observed with its frequency denoted in table 3.3. Again the data with and without noise are both fitted with a single Gaussian using `curve_fit`. The initial guesses of `curve_fit` are set to be the predetermined values. The fitting is done for the real part of the visibility as well as the amplitude of the *uv*-data.

[11] The calculator can be found here: `http://www.astro.ucla.edu/~wright/CosmoCalc.html`.

Table 3.3: Properties of the simulated fields with a quasar and a companion (Decarli et al. 2017).

Source		redshift	Frequency (GHz)	Position (J2000)	Flux (mJy)	Diameter (kpc)	FHWM (arcsec)
SDSS J0842+1218	Quasar	6.1 ± 0.0	269	08h42m29.43s +12d18m28.95s	0.87	6.0	1.1
	Companion			08h42m28.95s +12d18m55.1s	0.36	7.0	1.2
CFHQ J2100-1715	Quasar	6.1 ± 0.0	268	21h0m54.70s -17d15m21.9s	1.20	4.0	0.71
	Companion			21h0m55.45s -17d15m21.7s	2.05	4.6	0.81
PSO J231-20	Quasar	6.6 ± 0.0	251	15h26m37.84s -20d50m0.8s	4.41	5.0	0.92
	Companion			15h26m37.87s -20d50m2.3s	1.73	7.7	1.4
PSO J308-21	Quasar	6.2 ± 0.0	262	20h32m10.00s -21d14m2.4s	1.34	4.8	0.86
	Companion			20h32m10.17s -21d14m2.7s	0.19	6.4	1.1

4 Results

4.1 Antennae

Using an image of Antennae taken with the PACS instrument of the Herschel Space Observatory originally observed by Schirm et al. (2014), different simulated observations were made with different configurations of ALMA's antennas and various observing times. This image is used as an exercise only and does not represent the physical properties (size and flux) of Antennae. The image plane analysis is executed with the uniform weighting to preserve the outer details of Antennae[1]. The flux, found using the method explained in section 3.4, is plotted against the resolution in figure 4.1 for observations of 600 s and 3600 s. As seen, improving resolution (i.e. long-baseline configurations) decreased the flux. At resolutions better[2] than 0.2 arcsec the flux has decreased to about 50% of the original. The simulation has a high flux as the inbright was set to be 5 mJy. (Recall that inbright determines the surface brightness of the brightest pixel.) Consequently, the errors on the flux are very small and not visible in the figure.

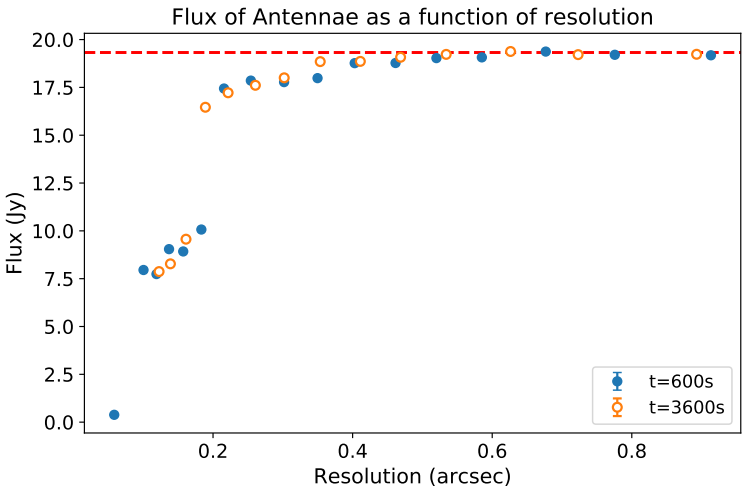


Figure 4.1: The flux of simulated ALMA observations with different antenna configurations of Antennae as a function of the resolution of an observation of 600 s (closed blue circles) and 3600 s (open orange circles). The flux of the original image is shown with the red dashed line.

[1]See section 3.3 for further explanation on the weightings.

[2]To avoid ambiguity throughout the paper a worse resolution means a resolution with higher values and a better resolution means a resolution with lower values.

Applying the method described in section 3.5, the size of Antennae is determined for different ALMA antenna configurations with `imfit`. This is shown in figure 4.2.

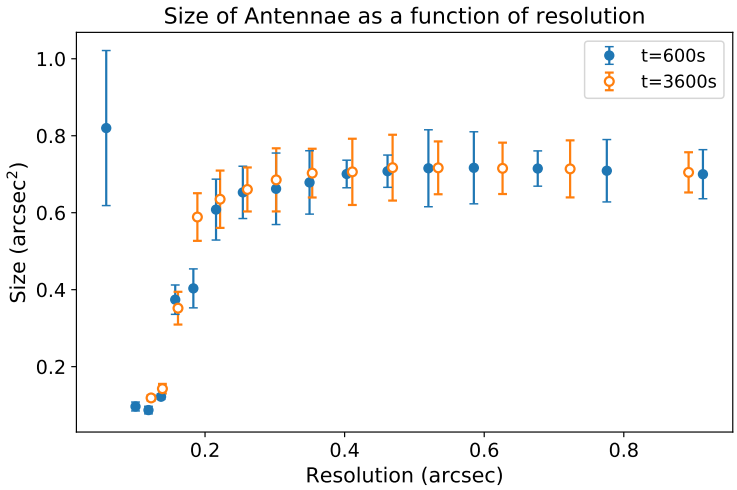


Figure 4.2: The size of simulated ALMA observations with different antenna configurations of Antennae as a function of the resolution for observations of 600 s (closed blue circles) and 3600 s (open orange circles).

Figure 4.2 demonstrates that for improving resolution the size decreases. The best resolution observation of 600 s is 10% larger than the trend seen for a resolution between 0.4 and 0.8 arcsec, as `imfit` did not recognise a source for this observation and only selected a section of noise. Note that Antennae consists of two parts. Once those two parts are resolved, `imfit` selects only one of these (the bottom left one). This explains the sizes found at a resolution better than 0.2 arcsec. The size of the original image is not included since the two parts are resolved and `imfit` does not give a representable size of Antennae. Recall that the physical properties of Antennae are changed for the purpose of this analysis.

4.2 Single Gaussian sources

This next section will cover the image plane analysis as well as the *uv*-plane analysis of the three different simulated Gaussian sources. The characteristics of these sources are shown in table 3.1. Images of the sources are shown in the corresponding sections. Again, flux analysis, as explained in section 3.4, is performed for these sources for the 20 different ALMA antenna configurations and different observing times. First the analysis of the circular source is explained. Secondly, the analysis of the elliptical sources is described.

4.2.1 Circular Gaussian

The circular source is shown in figure 4.3 and has a FWHM of the diameter of 0.2 arcsec and a flux of 5 mJy.

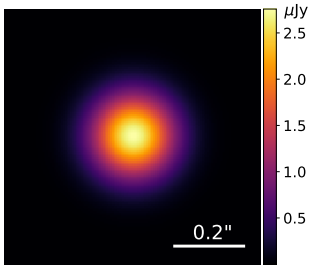


Figure 4.3: Image of the circular Gaussian source with a FWHM of 0.2 arcsec and a flux of 5 mJy.

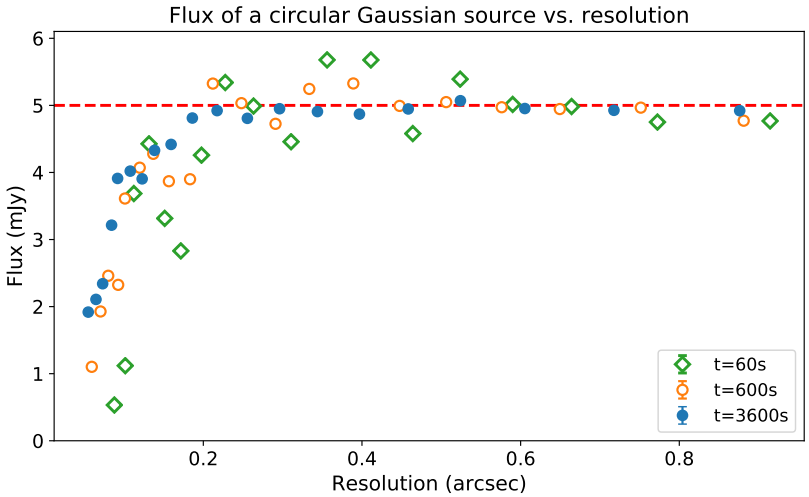


Figure 4.4: The flux as a function of resolution of the circular Gaussian source shown in figure 4.3 from simulated ALMA observations with different antenna configuration of 60 s (green open diamonds), 600 s (orange open circles) and 3600 s (blue closed circles). The flux of the original source is denoted by the red dashed line.

Figure 4.4 shows observations with different ALMA antenna configurations with observation times of 60 s, 600 s and 3600 s. The errors are of the order of 10^{-5} mJy due to the absence of background sources and therefore not visible in the figure. Similar to Antennae, with improving resolution the flux decreases. For observations with a resolution worse than 0.2 arcsec, the flux is scattered around the original flux set to the source when the field was created. When the resolution is better than 0.2 arcsec (i.e. long-baseline configurations), the flux decreases below the flux of the original source. The observation of 3600 s decreases to about 2 mJy, whereas the observation of 60 s decreases to about 0.5 mJy.

The size of the simulated circular Gaussian is found with `imfit` and plotted in figure 4.5 for observations of 60 s, 600 s and 3600 s. The size of the original is also determined with `imfit` to be able to compare it to the sizes found with the simulated observations. From the figure it is evident that for resolutions better than 0.2 arcsec, the size decreases to below the size of the original source. The size of the observations of 3600 s does not drop off as much as the size of the observations of 600 s and 60 s. In general, the observations of 3600 s recover the size more accurately than the shorter observations.

The size of the circular source is also examined with the *uv*-plane analysis described in section 3.6. Figure 4.6, 4.7 and 4.8 show the major axes found with `imfit` and the FWHM found with the *uv*-plane analysis method explained in section 3.6.

When analysing these figures, it is evident that the FWHM obtained with the *uv*-plane analysis are closer to the original FWHM and have a smaller error than the major axes acquired with `imfit`. What is most notable is the major axis found with `imfit` decreases significantly below the original major axis at a resolution better than 0.2 arcsec.

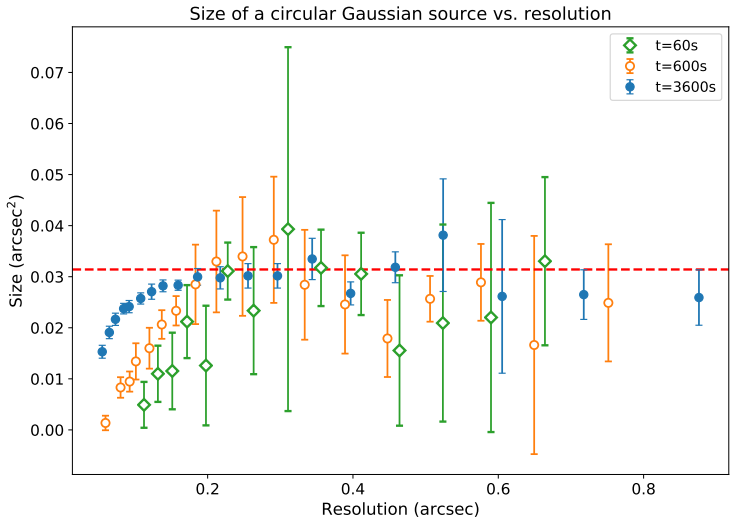


Figure 4.5: The size of simulated ALMA observations with different antenna configurations of 60 s (green open diamonds), 600 s (orange open circles) and 3600 s (blue closed circles) of the circular Gaussian source shown in figure 4.3 as a function of resolution. The size of the original source is denoted by the red dashed line.

However, the FWHM of the major axis found with the *uv*-plane analysis does not drop off as much for the 60 s and 600 s observations and does not drop off below one sigma for observations of 3600 s.

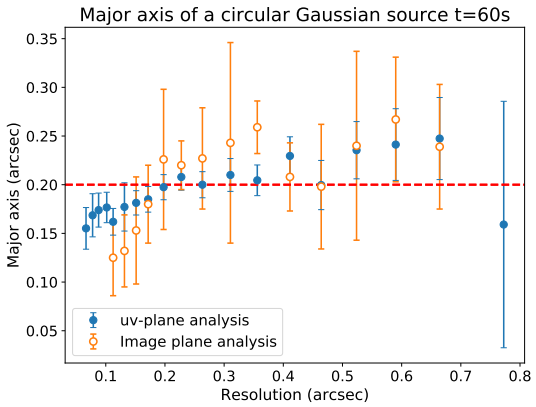


Figure 4.6: The FWHM of the major axis found with the image plane analysis using `imfit` (open orange circles) and with the *uv*-plane analysis (closed blue circles) of simulated ALMA observations of 60 s. The original FWHM is shown with the red dashed line.

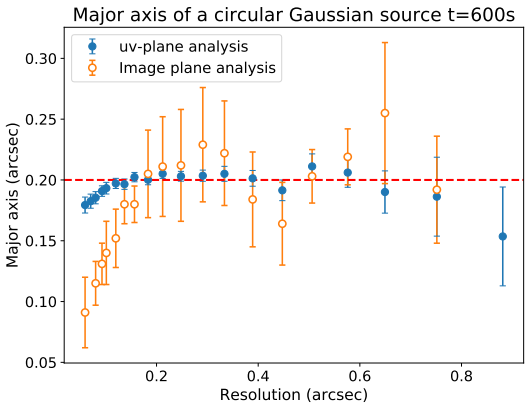


Figure 4.7: The FWHM of the major axis found with the image plane analysis using `imfit` (open orange circles) and with the *uv*-plane analysis (closed blue circles) of simulated ALMA observations of 600 s. The original FWHM is shown with the red dashed line.

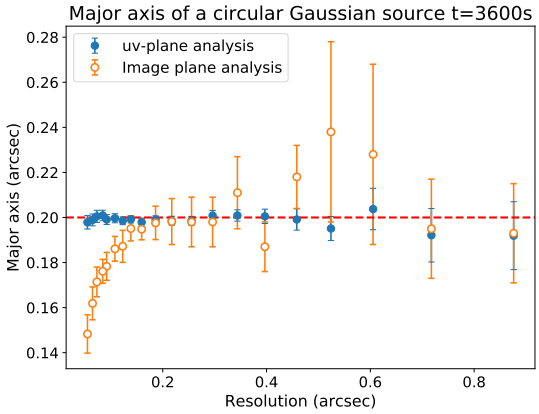


Figure 4.8: The FWHM of the major axis found with the image plane analysis using `imfit` (open orange circles) and with the *uv*-plane analysis (closed blue circles) of simulated ALMA observations of 3600 s. The original FWHM is shown with the red dashed line.

With this FWHM obtained using the *uv*-plane analysis, the size can be found. From figure 4.9 it is evident that longer observations recover the size of the source more accurately. In the 60 s and 600 s observations, a decrease is visible for resolutions better than 0.2 arcsec, where the 60 s observations reduce to 65% of the true size of the source and the 600 s observations drop off to 80% of the true size. The observations of 3600 s show no deviation larger than one sigma from the true size.

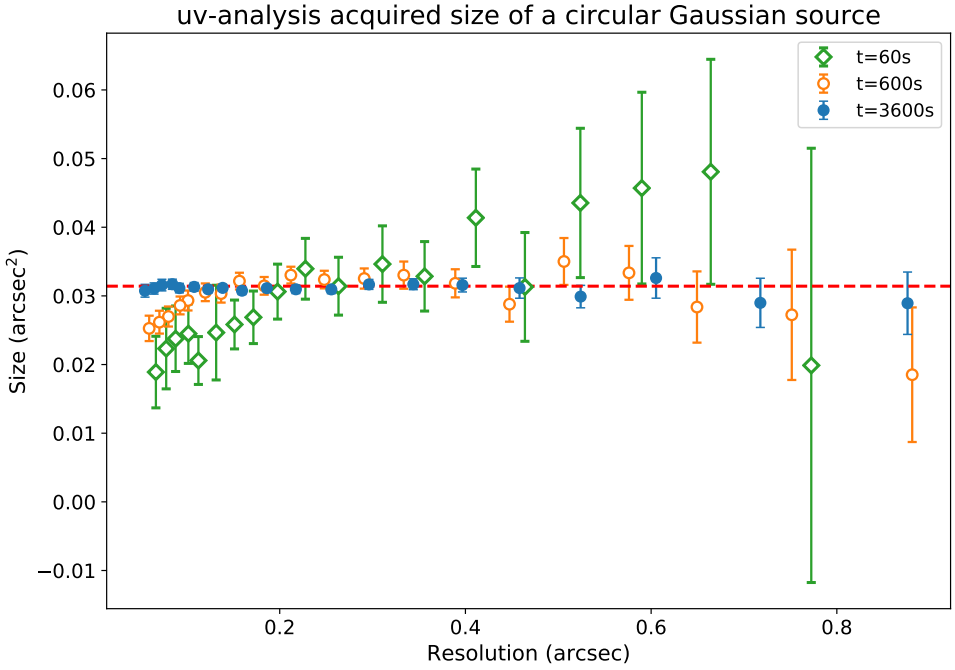


Figure 4.9: The size of the circular Gaussian source shown in figure 4.3 found using the *uv*-plane analysis with simulated ALMA observations with different antenna configurations of 60 s (green open diamonds), 600 s (orange open circles) and 3600 s (blue closed circles). The red dashed line represents the true size of the source.

4.2.2 Elliptical Gaussians

Looking at the elliptical Gaussian sources, the same analysis is applied to find the flux for simulated observations of 60 s, 600 s and 3600 s as mentioned in section 3.4. This is visualised in figure 4.11 for the first elliptical source (shown in figure 4.10). The errors are of the order of 10^{-5} mJy and are therefore not visible in the plot. Just like with the circular Gaussian, the flux decreases below the set flux of the original source. This again arises for resolutions better than 0.2 arcsec (i.e. long-baseline configurations). Moreover, the observations of 3600 s drop off to approximately 2 mJy. The observations of 60 s, on the other hand, drop off to approximately 0.1 mJy.

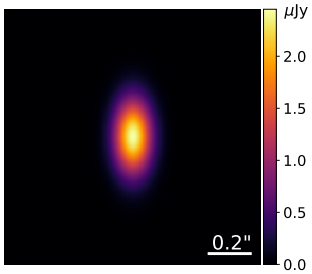


Figure 4.10: Image of the elliptical Gaussian source with a FWHM of the major axis of 0.3 arcsec, a FWHM of the minor axis of 0.15 arcsec and a flux of 5 mJy.

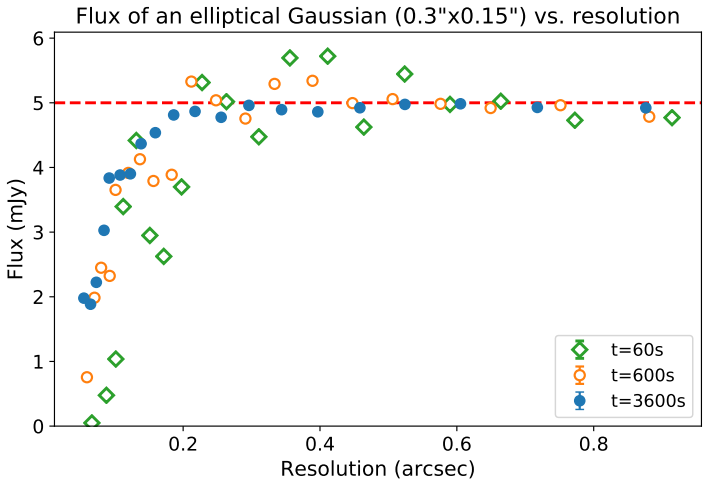


Figure 4.11: The flux as a function of resolution of the elliptical Gaussian source shown in figure 4.10 from simulated ALMA observations of 60 s (green open diamonds), 600 s (orange open circles) and 3600 s (blue closed circles). The flux of the original source is denoted by the red dashed line.

The size is acquired with `imfit` and plotted for observations of 60 s, 600 s and 3600 s in figure 4.12. Here again a decrease in size for resolutions better than 0.2 arcsec is visible. Especially at better resolutions, it is clear that the size of the observations of 3600 s have a smaller error than size of the observations of 600 s and 60 s. Moreover, the size of the observations of 3600 s does not decrease as much as the sizes obtained with observations of 600 s and 60 s do. The size of the 3600 s observations reduces to about 0.015 arcsec^2, whereas the size of the 600s observations drops off to around 0.007 arcsec^2 and `imfit` did not converge for observations of 60 s for the best resolutions.

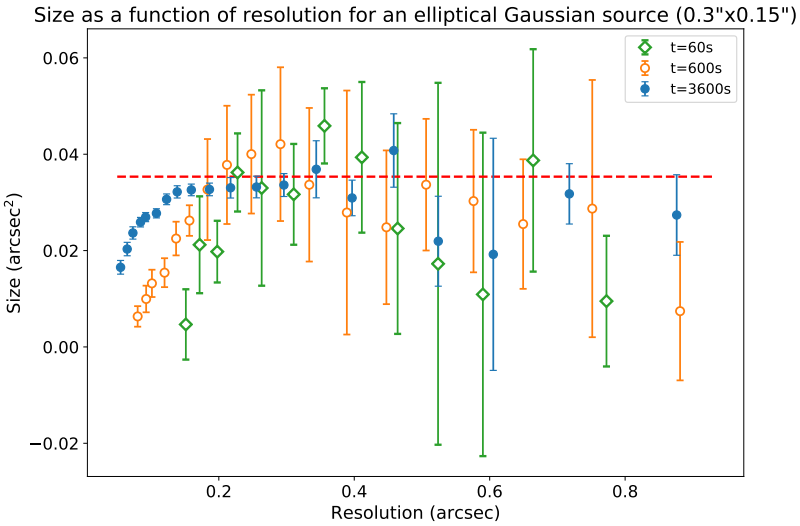


Figure 4.12: The size of the elliptical source shown in figure 4.10 found with `imfit` of simulated ALMA observations with different antenna configurations of 60 s (open green diamonds), 600 s (open orange circles) and 3600 s (closed blue circles). The original size is denoted with the red dashed line.

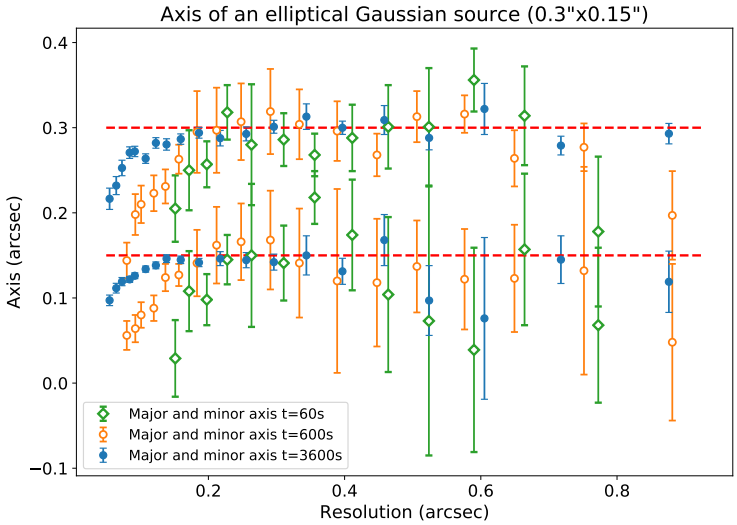


Figure 4.13: The major and minor axis obtained with `imfit` from simulated ALMA observations with different antenna configurations of observations of 60 s (open green diamonds), 600 s (open orange circles) and 3600 s (closed blue circles). The red dashed lines denote the true major and minor axis of the source.

It is interesting to look at the major and minor axes, as they differ for elliptical sources, and see if they can be recovered. This is shown in figure 4.13, where the major and minor axes are recovered mostly within one sigma of the true axis of the source for resolutions worse than 0.2 arcsec. For longer baseline configurations (i.e. better resolutions), the major and minor axes decrease (as expected from figure 4.12)

below their original values. The major axis found with observations of 600 s drop off to below 0.15 arcsec (the specified minor axis). For the observations of 60 s with the most extended configurations, `imfit` did not converge. Increasing the observation time (i.e. the SNR) decreases the errors and reduces the drop off occurring at resolutions better than 0.2 arcsec.

The FWHM is also found with the *uv*-plane analysis described in section 3.6. The results are shown in figures 4.14, 4.15 and 4.16. As a reference, the major and minor axes obtained with `imfit` in the image plane analysis are also shown. It can be seen that the diameter found with the *uv*-plane analysis lies in between the true major and minor axes. It is evident that the image plane analysis is more influenced by the improvement of the resolution as the major and minor axes drop off significantly, whereas the *uv*-plane analysis acquired FWHM of the diameter only changes 10% for observations of 3600 s (figure 4.16) and 30% for observations of 60 s (figure 4.14).

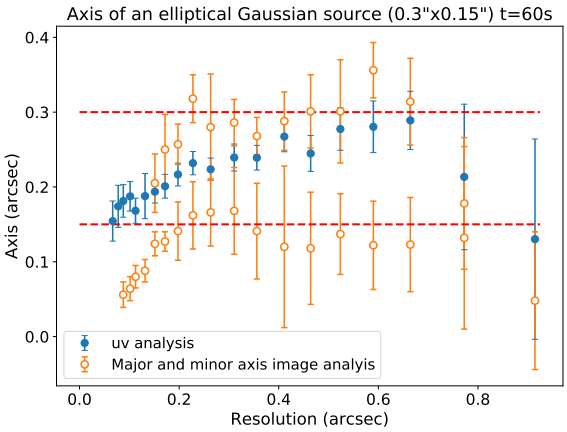


Figure 4.14: The FWHM of the diameter of the elliptical source found with the *uv*-plane analysis (closed blue circles) and the major and minor axis found with `imfit` (open orange circles) from simulated ALMA observations of 60 s. The red dashed lines denote the true major and minor axis of the source.

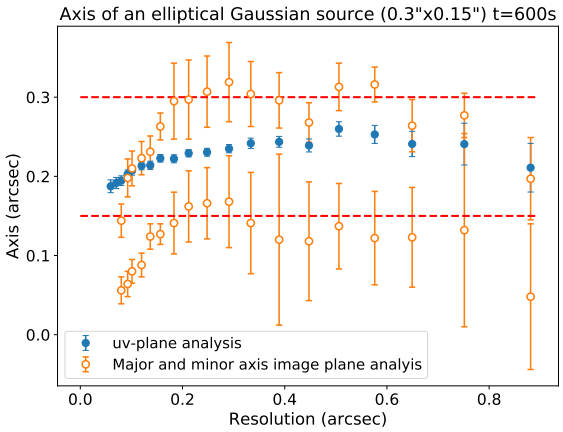


Figure 4.15: The FWHM of the diameter of the elliptical source found with the *uv*-plane analysis (closed blue circles) and the major and minor axis found with `imfit` (open orange circles) from simulated ALMA observations of 600 s. The red dashed lines denote the true major and minor axis of the source.

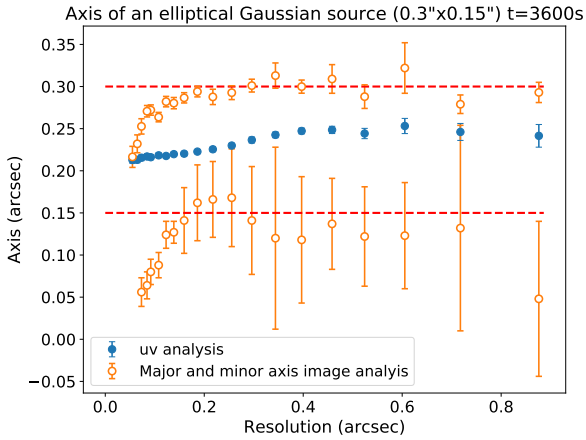


Figure 4.16: The FWHM of the diameter of the elliptical source found with the *uv*-plane analysis (closed blue circles) and the major and minor axis found with `imfit` (open orange circles) from simulated ALMA observations of 3600 s. The red dashed lines denote the true major and minor axis of the source.

With the diameter obtained using the *uv*-plane analysis, the size of the source can be calculated. This is again done for simulated observations of 60 s, 600 s and 3600 s (figure 4.21). It is evident that the size of an elliptical source is generally overestimated for a resolution worse than 0.2 arcsec. The observations of 3600 s recover the size most accurately. For resolutions better than 0.2 arcsec, the size acquired with the *uv*-plane analysis drops down to the true size for observations of 3600 s and drops below the true size for observations of 60 s and 600 s.

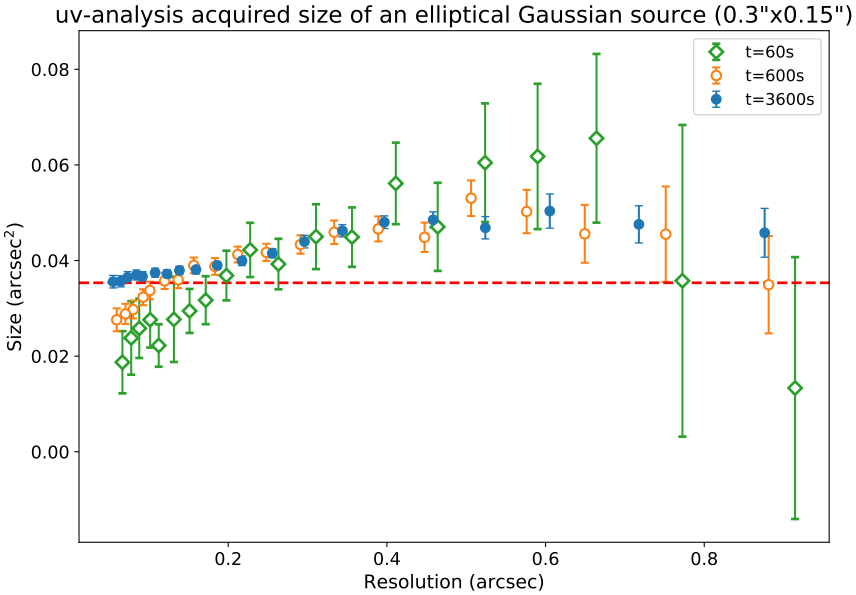


Figure 4.17: The size of the source as calculated with the diameter obtained using the *uv*-plane analysis for simulated ALMA observations of 60 s (open green diamonds), 600 s (open orange circles) and 3600 s (closed blue circles). The true size of the source is denoted with the dashed red line.

To investigate the influence of the ellipticity of a simulated Gaussian source on the flux and size obtained from simulated observations with different ALMA antenna configurations, 600 s observations are made of a more elliptical simulated Gaussian source (figure 4.18), which will be referred to as 'very elliptical source' from now on. Figure 4.19 shows the flux obtained from 600 s observations of the elliptical source (figure 4.10) and 600 s observations of the very elliptical source (figure 4.18). The errors are not visible in the plot as they are of the order of 10^{-5} mJy.

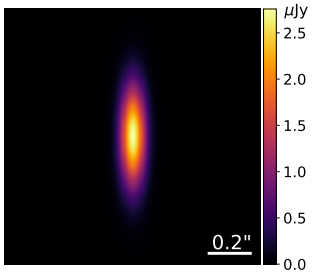


Figure 4.18: Image of the very elliptical source with a FWHM of the major axis of 0.4 arcsec, a FWHM of the minor axis of 0.1 arcsec and a total flux of 5 mJy.

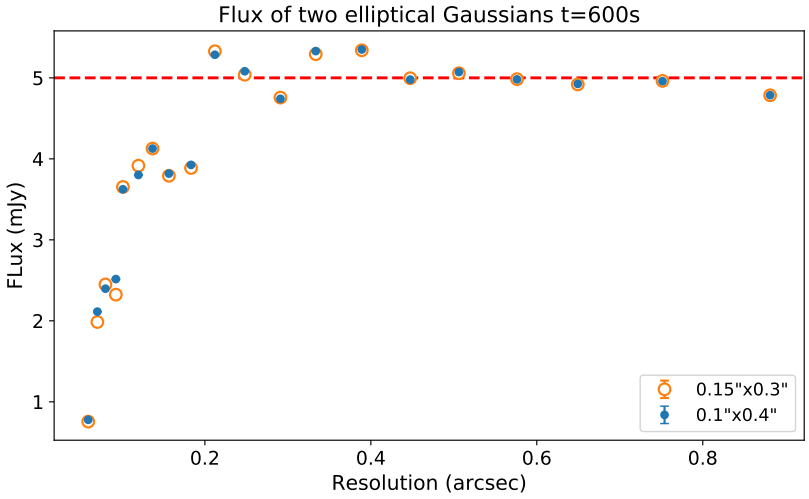


Figure 4.19: The flux of the very elliptical source (0.1"x0.4") (closed blue circles) and the elliptical source (0.15"x0.3") (open orange circles) obtained from simulated ALMA observations of 600 s. The red dashed line denotes the true flux of the source.

From figure 4.19, it is evident that the ellipticity of the source does not influence the obtained flux of the source, as the acquired flux for the elliptical and very elliptical source are very close together. Only at resolutions better than 0.2 arcsec is a small deviation visible.

The diameter of the very elliptical source is obtained with the *uv*-plane analysis and the major and minor axes are found using `imfit` (figure 4.20). Similar to the elliptical source, the FWHM obtained with the *uv*-plane analysis is found between the major and minor axes of the source and the FWHM found with the *uv*-plane analysis does not drop off as much as the major and minor axis found with the image plane analysis do.

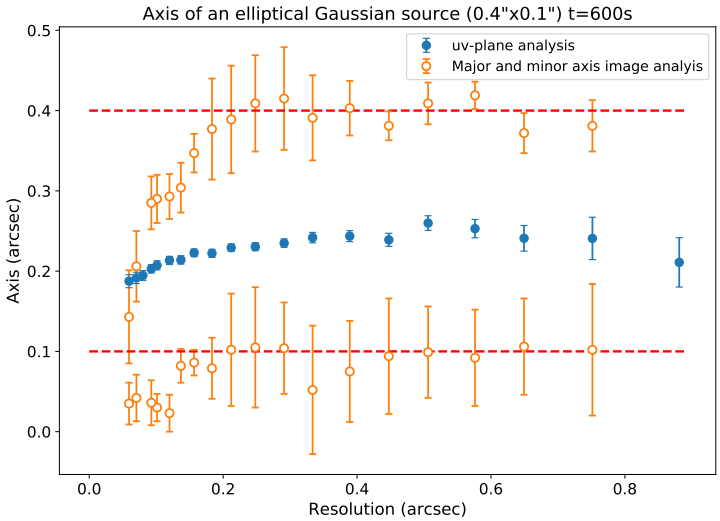


Figure 4.20: The diameter obtained with the *uv*-plane analysis (closed blue circles) and the major and minor axes found with the image plane analysis using `imfit` (open orange circles) for simulated ALMA observations of 600 s. The true major and minor axis are denoted by the red dashed line.

With the diameter and the major and minor axes, the size of the very elliptical source is acquired with the *uv*-plane analysis and `imfit` respectively, for observations of 600 s. Figure 4.21 shows the size of the very elliptical source for observations of 600 s. Similar to the size found for the elliptical source, the size for the very elliptical source found with the *uv*-plane analysis overestimates the true size and only reduces down to the true size at a resolution of 0.1 arcsec. The image plane acquired size recovers the size more accurately for resolutions better than 0.2 arcsec, but drops down to 10% of the true size at the most extended configurations. The errors on the image plane acquired size are of the order of magnitude of the size.

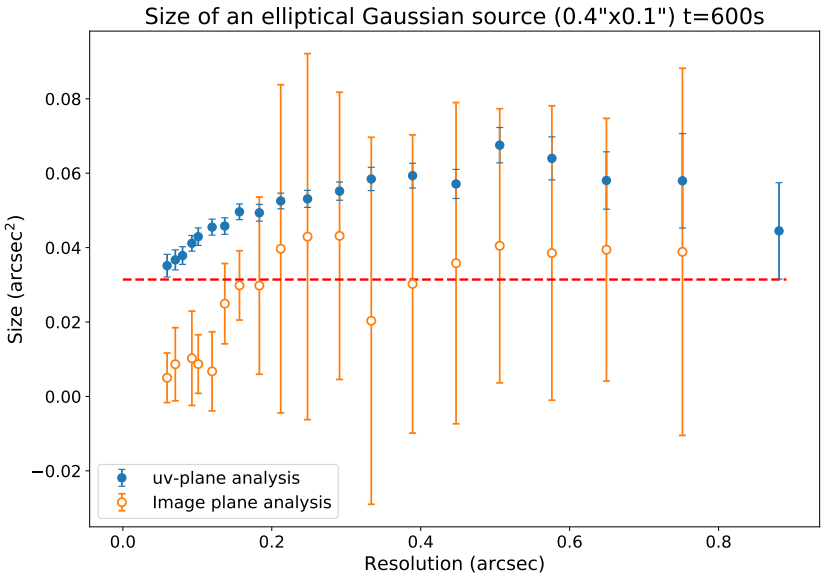


Figure 4.21: The size of the very elliptical source shown in figure 4.18 calculated with the diameter obtained from the *uv*-plane analysis (closed blue circles) and with the major and minor axes obtained from the image plane analysis using `imfit` (open orange circles) for simulated ALMA observations of 600 s. The red dashed line denotes the true size of the source.

4.3 Companion Gaussian sources

This section covers the *uv*-plane analysis of ten different companion fields. The different fields with companion sources are obtained from Hodge et al. (2013), Wardlow et al. (2018) and Decarli et al. (2017) and their properties are displayed in table 3.2 and 3.3. The sources acquired from Hodge et al. (2013), Simpson et al. (2015) and Wardlow et al. (2018) are from ALESS and S2CLS. Rivera et al. (2018) found a FWHM of 0.8-1.6 arcsec for four ALESS SMG sources from CO(3-2) data. Therefore, the FWHM of the diameter of these sources was set to 0.8 arcsec since they were unknown. Decarli et al. (2017) did provide FWHMs for their sources which were adopted in this research. These fields all contain a quasar with a companion source.

4.3.1 ALESS and UDS companion fields

This section will cover the *uv*-plane analysis of the ALESS and UDS companion fields simulated after observations of Hodge et al. (2013), Simpson et al. (2015) and Wardlow et al. (2018). As the aim is to find the influence that companion sources have on the main source[3], the *uv*-data of the field is fitted with a singular Gaussian to recover the size of the main source. This procedure is performed for data without noise and data with noise and the Gaussian is fitted to the real part of the visibility and the amplitude. This analysis is explained in detail in section 3.7.

[3]The main source here is considered the source put in the center of the field (X.1 for all ALESS fields and X.0 for the UDS field).

Figure 4.23 shows the *uv*-fitting of the ALESS 41 field (shown in figure 4.22) with sources ALESS 41.1, 41.3 both originally observed by Hodge et al. (2013) and 41.C originally observed by Wardlow et al. (2018). Figure 4.23a shows the *uv*-fitting to the real part of the visibility and figure 4.23b shows the *uv*-fitting to the amplitude for observations of 600 s with 20 different ALMA antenna configurations. As can be observed in figure 4.23a, the fits to the data without noise and to the data with noise are scattered around the original FWHM of the source as was predetermined. The fits to the data with noise have larger errors and are scattered more widely around the original in comparison to the FWHM inferred from the data without noise. The FWHMs inferred from the data with noise found for resolutions better than 0.1 arcsec have errors of around 25%.

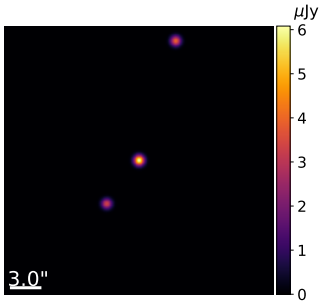
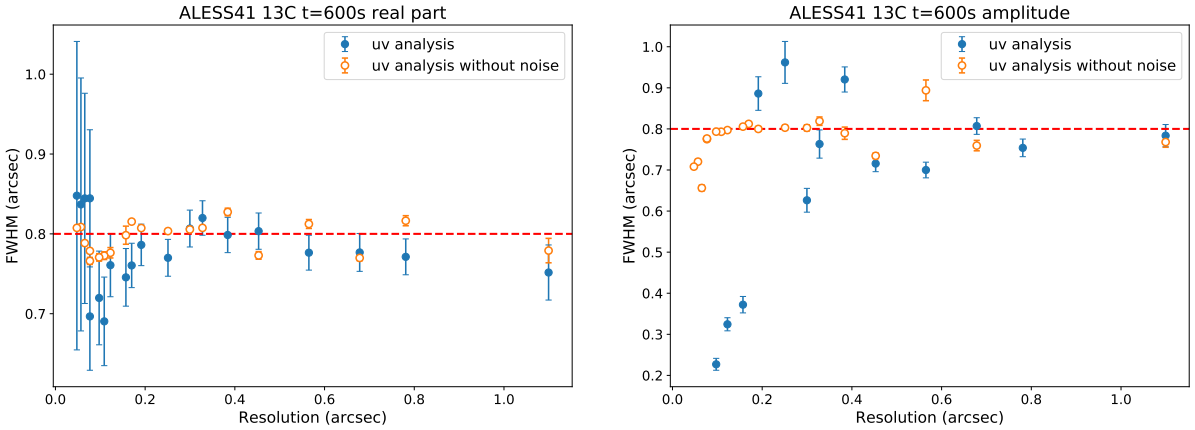


Figure 4.22: Image of the Gaussian representation of ALESS 41 with sources 41.1, 41.3 and 41.C.

Comparing the *uv*-fitting to the real part (figure 4.23a) with the *uv*-fitting to the amplitude (figure 4.23b), it is evident that FWHM obtained from the amplitude fitted data deviate more from the original FWHM of the source. The FWHM found from the data with and without noise decrease below the original FWHM for resolutions worse than 0.1 arcsec unlike the *uv*-fitting to the real part of the visibility.



(a) ***(b)***

Figure 4.23: The FWHM of the main source (ALESS 41.1) shown in figure 4.22 obtained with the *uv*-fitting to the real part of the visibility (left, 4.23a) and to the amplitude (right, 4.23b) for data without noise (open orange circles) and data with noise (closed blue circles) for simulated ALMA observations of 600 s. The red dashed line denotes the true FWHM of the main source.

ALESS 49 (shown in figure 4.24) is more closely together than ALESS field 41. Figure 4.25 shows the *uv*-plane analysis of the ALESS 49 field, with sources 49.1, 49.2 both detected by Hodge et al. (2013) and 49.C and 49.L originally observed by Wardlow et al. (2018) of observations of 600 s with 20 different ALMA antenna configurations. On the left (figure 4.25a) the real part of the visibility is fitted. As demonstrated in figure 4.25a, for resolutions better than 0.1 arcsec the obtained FWHM with the data without noise increases above the original FWHM up to about 180% and the FWHM obtained with the data with noise increases up to 120%. For the amplitude fitting (shown in figure 4.25b) this effect is not as clear. The FWHM inferred from the data without noise does not show a significant increase in FWHM and the FWHM inferred from the data with noise the fit for the most extended configurations did not converge. Nevertheless, an increasing trend in the FWHM seems to be starting at a resolution of 0.2 arcsec and better.

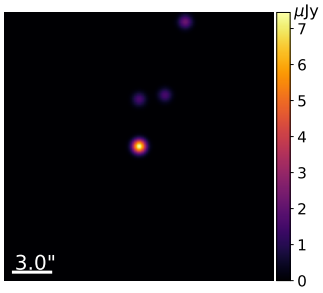
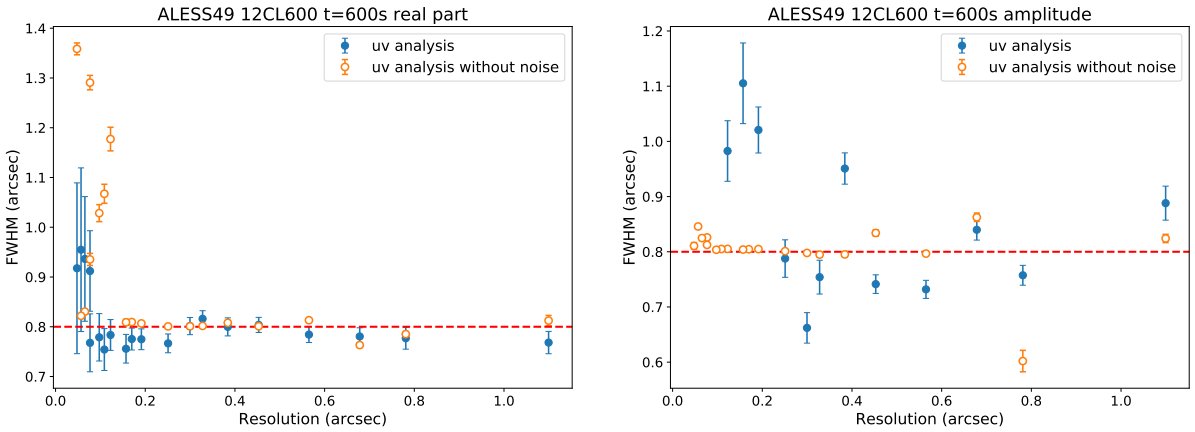


Figure 4.24: Image of the Gaussian representation of ALESS 49 with sources 49.1, 49.2, 49.C and 49.L.



(a) ***(b)***

Figure 4.25: The FWHM of the main source (ALESS 49.1) shown in figure 4.24 obtained with the *uv*-fitting to the real part of the visibility (left, 4.25a) and to the amplitude (right, 4.25b) for data without noise (open orange circles) and data with noise (closed blue circles) for simulated ALMA observations of 600 s. The red dashed line denotes the true FWHM of the main source.

ALESS field 71 only contains two sources, 71.1 and 71.3 both originally observed by Hodge et al. (2013). The field is shown in figure 4.26. Figure 4.27 shows the *uv*-plane analysis for observations of 600 s. A small rise in the FWHM is found in the fitting to the real part of the visibility (figure 4.27a) for the data with noise. This small increase, however, is accompanied by an increase in errors. Consequently, the original FWHM is still recovered within one sigma. The FWHM inferred from the data without noise does not show any notable deviation from the original FWHM in both the *uv*-plane analysis for the real part of the visibility and the amplitude (figure 4.27b). However, FWHM found with the amplitude analysis with the noisy data decreases significantly to 0.1 arcsec for resolutions below around 0.3 arcsec.

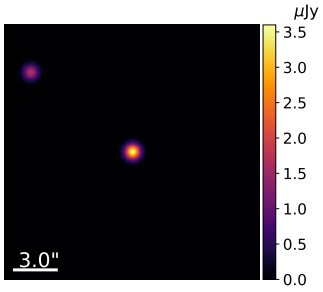
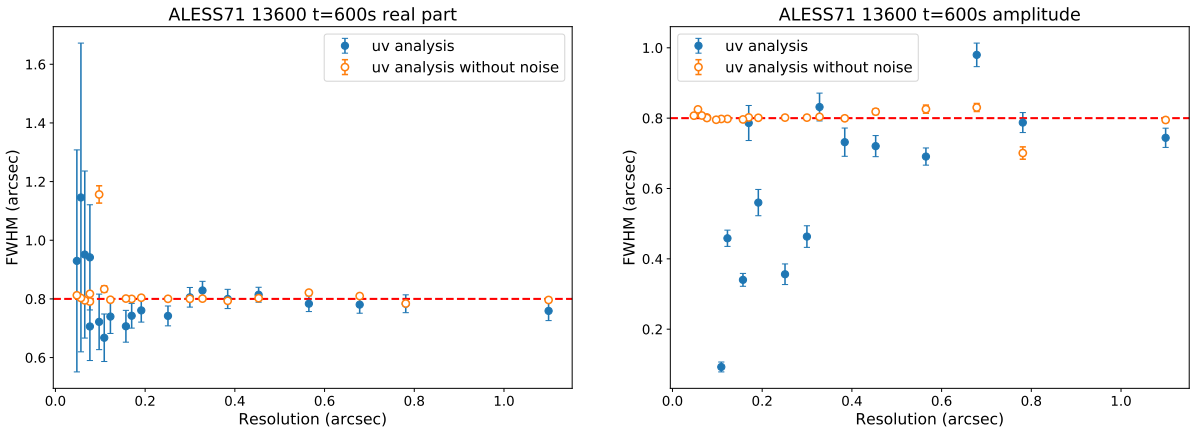


Figure 4.26: Image of the Gaussian representation of ALESS 71 with sources 71.1 and 71.3.



(a) ***(b)***

Figure 4.27: The FWHM of the main source (ALESS 71.1) shown in figure 4.26 obtained with the *uv*-fitting to the real part of the visibility (left, 4.27a) and to the amplitude (right, 4.27b) for data without noise (open orange circles) and data with noise (closed blue circles) for simulated ALMA observations of 600 s. The red dashed line denotes the true FWHM of the main source.

Figure 4.28 shows ALESS field 75 with 75.1, 75.2, 75.4 originally observed by Hodge et al. (2013) and 75.C detected by Wardlow et al. (2018). Source 75.2 (bottom left) is actually brighter than 75.1 (center), with fluxes of 5.0 mJy and 3.2 mJy respectively. Nonetheless, there is no clear effect of this visible in the *uv*-plane analysis. Figure 4.29 shows the *uv*-plane analysis of the real part of the visibility (figure 4.29a) and the amplitude (figure 4.29b). In the real part analysis (left) the FWHM inferred from the noisy data show a small increase above the original FWHM which is within the error margins at the best resolutions (i.e. the most extended baseline configurations) and the FWHM onbtained from the data without noise has three points at double the original FWHM. This, however, is not found with the amplitude analysis (figure 4.29b). On the contrary, here the FWHM obtained from the data with noise show a decrease below the original for better resolutions whilst the FWHM inferred from the data without noise is scattered closely around the original. The fit for the data with noise, however, did not converge for the most extended configurations and therefore no strong conclusions can be made.

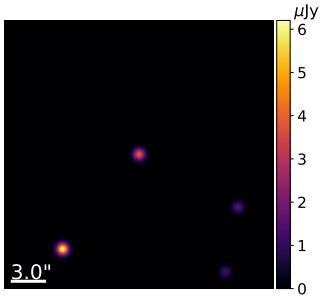
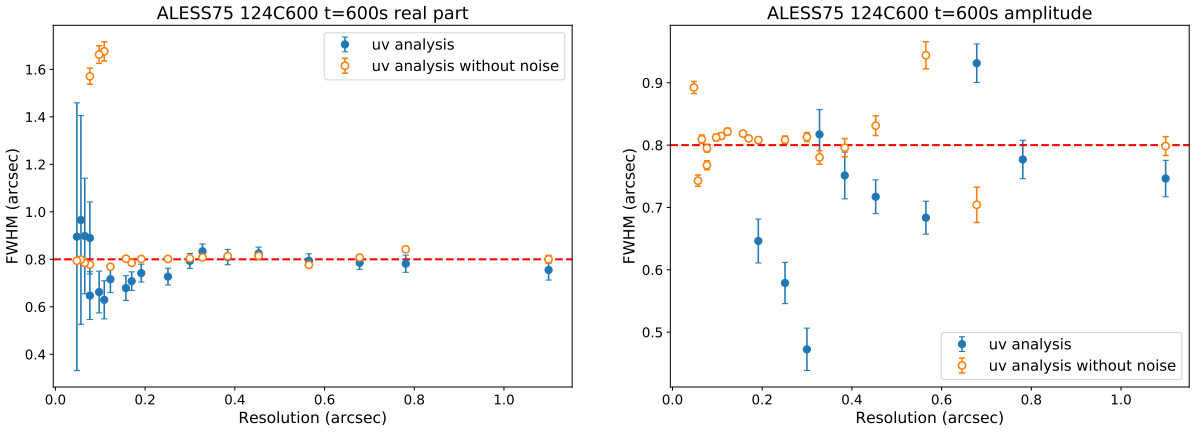


Figure 4.28: Image of the Gaussian representation of ALESS 75 with sources 75.1, 75.2, 75.4 and 75.C.



(a) ***(b)***

Figure 4.29: The FWHM of the main source (ALESS 75.1) shown in figure 4.28 obtained with the *uv*-fitting to the real part of the visibility (left, 4.29a) and to the amplitude (right, 4.29b) for data without noise (open orange circles) and data with noise (closed blue circles) for simulated ALMA observations of 600 s. The red dashed line denotes the true FWHM of the main source.

ALESS 87 (shown in figure 4.30) consists of four sources, namely 87.1 and 87.3 originally observed by Hodge et al. (2013) and 87.C and 87.L detected by Wardlow et al. (2018). In ALESS 87, the main source, 87.1, is actually the faintest of all sources in the field. Nevertheless, no significant effect of this is detected in the *uv*-plane analysis (figure 4.31). The FWHM found with the real part of the visibility analysis (shown in figure 4.31a) has an increase for the data without noise for resolutions around 0.1 arcsec. The FWHM obtained from the data with noise only shows an increase in the errors of the FWHMs of about 100% found with the best resolutions (better than 0.1 arcsec). The FWHM found with the *uv*-plane analysis conducted on the amplitude (figure 4.31b) shows no significant deviation from the original FWHM for the data without noise. The fits to the noisy data do show some unexpected behaviour. For a resolution between 0.1 and 0.6 arcsec the FWHM is consistently underestimated, except for one data point that has about 2.5 times the original FWHM at a resolution of around 0.2 arcsec. However, by examining the fitted Gaussian to the data, it is evident that these are not reliable[4].

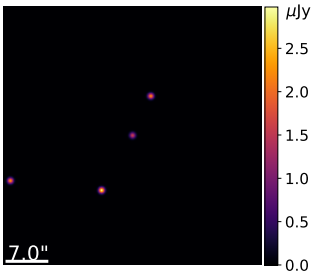
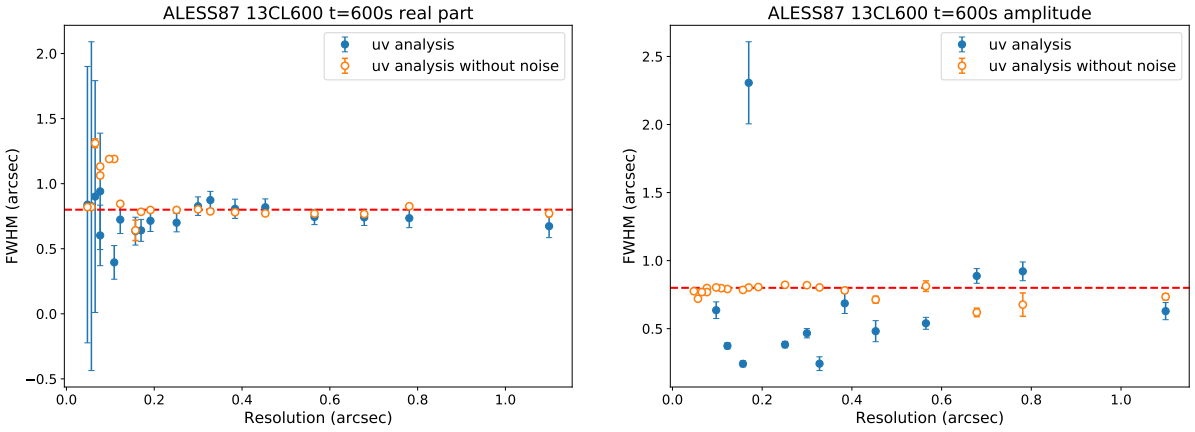


Figure 4.30: Image of the Gaussian representation of ALESS 87 with sources 87.1, 87.3, 87.C and 87.L.



(a) ***(b)***

Figure 4.31: The FWHM of the main source (ALESS 87.1) shown in figure 4.30 obtained with the *uv*-fitting to the real part of the visibility (left, 4.31a) and to the amplitude (right, 4.31b) for data without noise (open orange circles) and data with noise (closed blue circles) for simulated ALMA observations of 600 s. The red dashed line denotes the true FWHM of the main source.

[4]More information on this can be found in the discussion (section 5).

The UDS306 field (figure 4.32) contains sources 306.0, 306.1 and 306.2 originally observed by Simpson et al. (2015) and 306.L detected by Wardlow et al. (2018). In this field the main source, 306.0, has a flux four times as high as the companion sources. The FWHM found with the *uv*-plane analysis of the real part of the visibility (figure 4.33a) shows no significant deviation for the data without noise and, as seen with the ALESS fields, the FWHM inferred from the data with noise has an increase in the FWHM at the resolutions better than 0.1 arcsec. However, the errors increase too and the true FWHM stays within one sigma. The FWHM obtained from the *uv*-fitting to the amplitude (figure 4.33b) shows again no significant deviation from the original FWHM for the data without noise. The fits to the noisy data are scattered around the true FWHM of the main source.

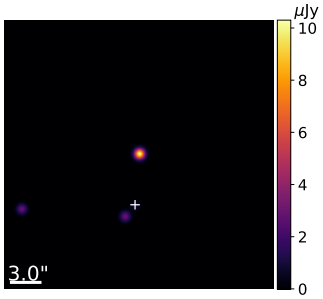
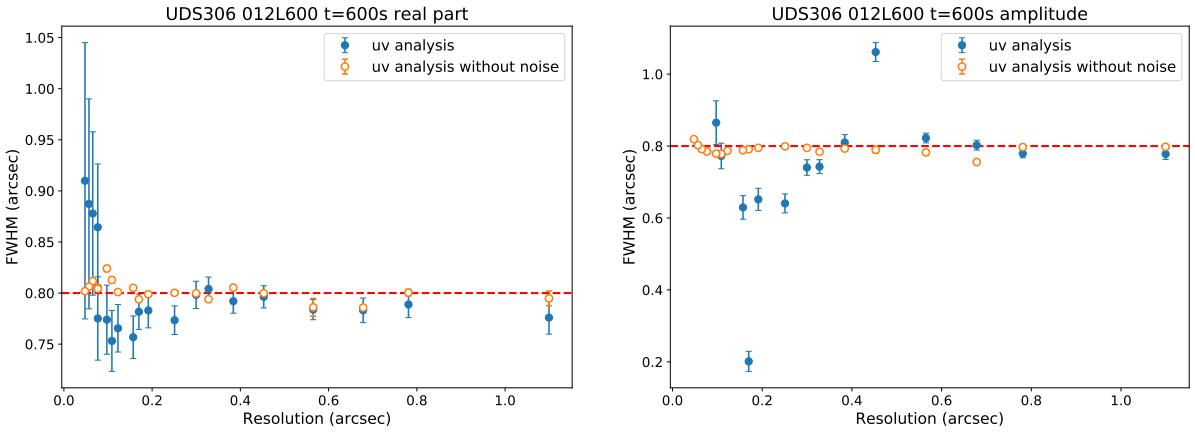


Figure 4.32: Image of the Gaussian representation of UDS306 with sources 306.0, 306.1, 306.2 and 306.L. UDS306.L is for clarity denoted with a white +.



(a) ***(b)***

Figure 4.33: The FWHM of the main source (UDS306.0) shown in figure 4.32 obtained with the *uv*-fitting to the real part of the visibility (left, 4.33a) and to the amplitude (right, 4.33b) for data without noise (open orange circles) and data with noise (closed blue circles) for simulated ALMA observations of 600 s. The red dashed line denotes the true FWHM of the main source.

4.3.2 Quasar companion fields

The fields obtained from Decarli et al. (2017) are simulated with the properties shown in table 3.3. The quasar is located in the center of the image. The same *uv*-plane analysis is conducted as for the ALESS and UDS sources.

The simulated observations of SDSS J0842+1218 (J0842) were executed at a frequency of 269 GHz. Figure 4.34 shows J0842. The *uv*-plane analysis is shown in figure 4.35. The FWHM found with the *uv*-fitting to the real part of the visibility (figure 4.35a) shows no significant deviation from the original FWHM for the data without noise. The FWHM inferred from the data with noise shows the now familiar pattern. At a resolution between 0.1 and 0.2 arcsec, the acquired FWHM decreases below the original and below 0.1 arcsec it increases above the original, but still within one sigma. The FWHM found with the *uv*-plane analysis for the amplitude (figure 4.35b) shows a big drop off for the noisy data for resolutions between 0.1 and 0.2 arcsec. The FWHM inferred from the data without noise is distributed closely to the original FWHM.

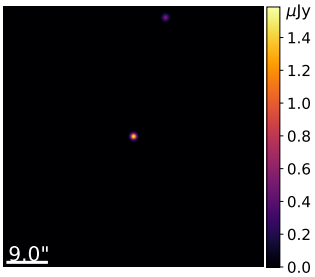
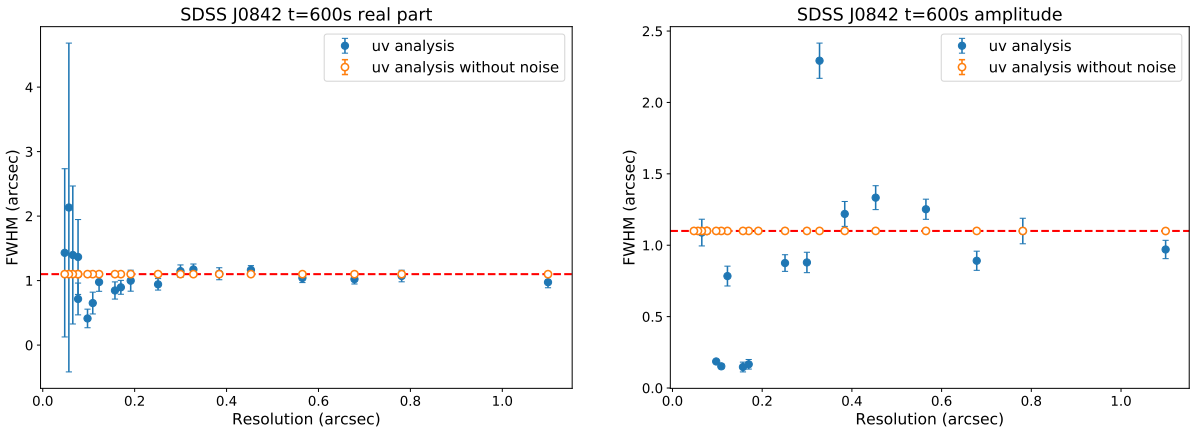


Figure 4.34: Image of the Gaussian representation of SDSS J0842+1218.



(a) ***(b)***

Figure 4.35: The FWHM of the quasar (SDSS J0842+1218) shown in figure 4.34 obtained with the *uv*-fitting to the real part of the visibility (left, 4.35a) and to the amplitude (right, 4.35b) for data without noise (open orange circles) and data with noise (closed blue circles) for simulated ALMA observations of 600 s. The red dashed line denotes the true FWHM of the main source.

In the CFHQ J2100-1715 (J2100) field (figure 4.36) the companion galaxy has a higher flux (2.05 mJy) than the quasar (1.20 mJy). The simulated ALMA observations were made at a frequency of 268 GHz. The FWHM found with the *uv*-plane analysis of the real part of the visibility (figure 4.37a) looks similar to the FWHM found with the *uv*-plane analysis obtained for J0842. Again, the data without noise is distributed around the original size of the quasar. The FWHM inferred from the noisy data has an increase above the original FWHM for a resolution below 0.1 arcsec whilst between a resolution of 0.1 and 0.2 arcsec the FWHMs found are lower than the original. For resolutions above 0.2 arcsec the FWHM of the quasar is closely scattered around the true FWHM. The FWHM found with the *uv*-fitting to the amplitude (figure 4.37b) seems to show the same behaviour but with a weaker trend. The FWHM inferred from the data without noise is again closely distributed around the original FWHM. The FWHM obtained from the data with noise is scattered around the original for a resolution worse than 0.2 arcsec. For a resolution better than 0.2 arcsec the found FWHMs are smaller than the original except for one that is about twice the FWHM of the original.

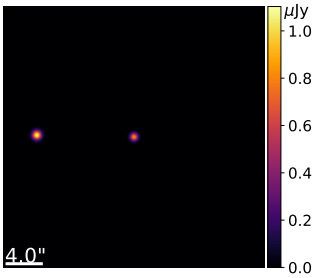
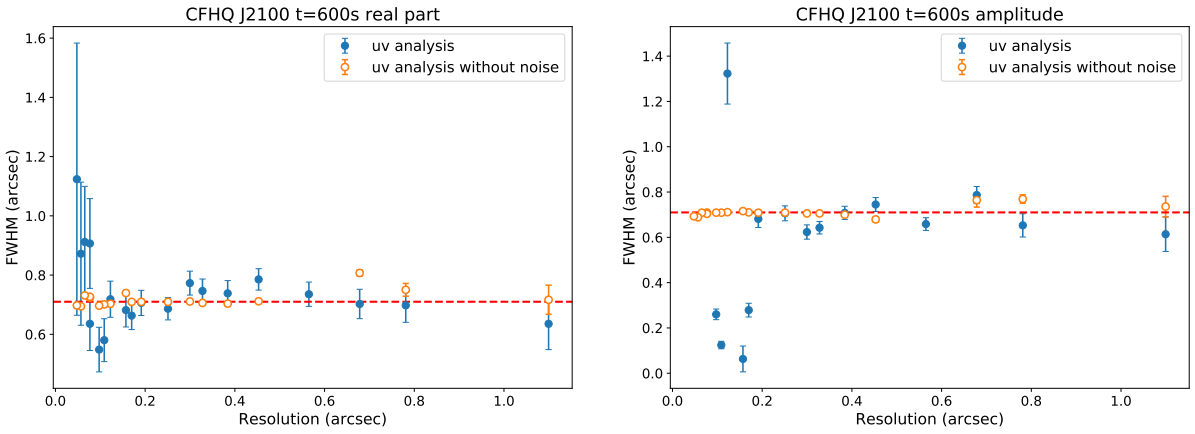


Figure 4.36: Image of the Gaussian representation of CFHQ J2100-1715.



(a) ***(b)***

Figure 4.37: The FWHM of the quasar (CFHQ J2100-1715) shown in figure 4.36 obtained with the *uv*-fitting to the real part of the visibility (left, 4.37a) and to the amplitude (right, 4.37b) for data without noise (open orange circles) and data with noise (closed blue circles) for simulated ALMA observations of 600 s. The red dashed line denotes the true FWHM of the main source.

The PSO J231-20 (J231) field (figure 4.38) has the smallest separation between quasar and companion (1.6 arcsec) and is observed at a frequency of 251 GHz. The companion is less than half the flux of the quasar. The FWHM found with the *uv*-plane analysis of the real part of the visibility (figure 4.39a) shows that for resolutions worse than 0.2 arcsec for the noisy data increases above the original size, up to 125% of the original. The FWHM obtained from the data without noise shows a smaller increase up to 110% starting at a resolution of 0.6 arcsec and worse. At a resolution better than 0.2 arcsec the FWHM acquired from the noisy data dips under the original FWHM, and for resolutions better than 0.1 arcsec the FWHM increases above the original again. The FWHM inferred with *uv*-plane analysis of the amplitude (figure 4.39b) does not show this same effect. The FWHM found with the data without noise is closely distributed around the original FWHM. The FWHM acquired from the data with noise is higher than the original for a resolution worse than 0.4 arcsec and better than 0.1 arcsec. In between 0.1 and 0.4 arcsec the FWHM is below than the original.

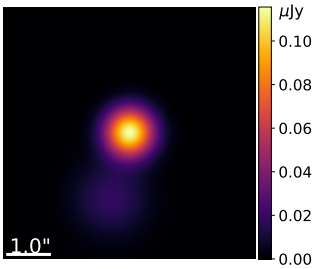
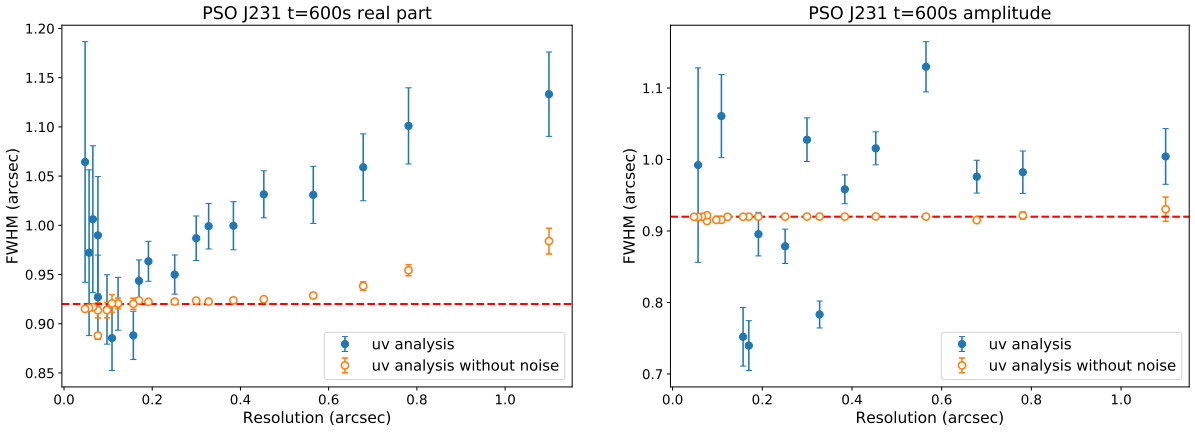


Figure 4.38: Image of the Gaussian representation of PSO J231-20.



(a) ***(b)***

Figure 4.39: The FWHM of the quasar (PSO J231-20) shown in figure 4.38 obtained with the *uv*-fitting to the real part of the visibility (left, 4.39a) and to the amplitude (right, 4.39b) for data without noise (open orange circles) and data with noise (closed blue circles) for simulated ALMA observations of 600 s. The red dashed line denotes the true FWHM of the main source.

The companion source in the PSO J308-21 (J308) field is very faint (0.19 mJy), so a white + is plotted at the center of the companion to make it visible in the field (figure 4.40). The field is observed at a frequency of 262 GHz. The FWHM found with the *uv*-plane analysis of both the real part of the visibility (figure 4.41a) and the amplitude (figure 4.41b) shows for the data without noise no significant deviation from the true FWHM. The FWHM inferred from the data with noise of the real part of the visibility shows a small decrease of 20% below the true FWHM between resolutions of 0.1 and 0.2 arcsec and an increase of 20% for resolutions better than 0.1 arcsec. However, the error in these FWHMs compensate for this. The FWHM obtained from the data with noise of the amplitude shows a wide scatter around the true value of 30%.

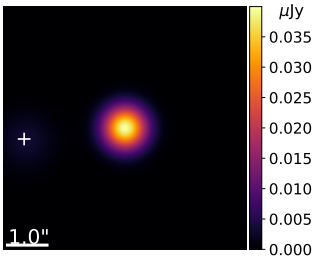
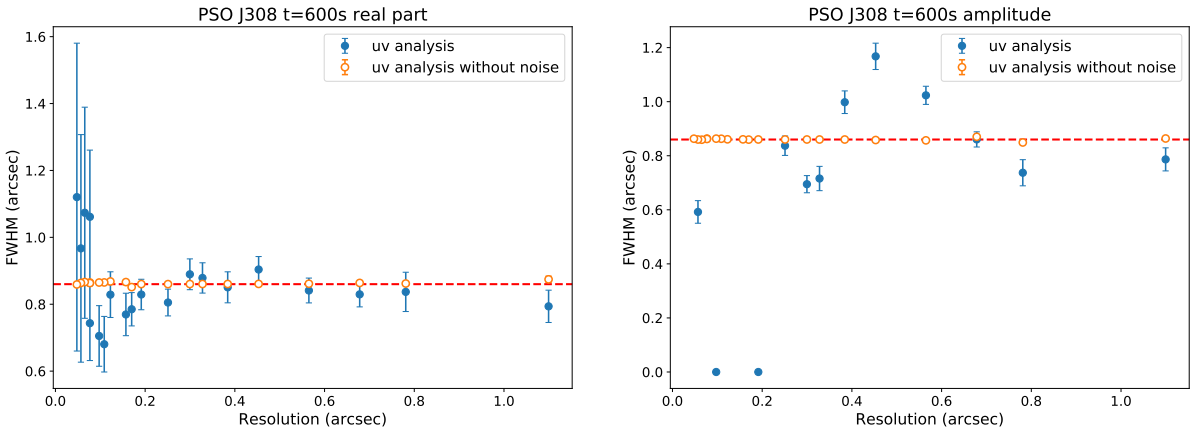


Figure 4.40: Image of the Gaussian representation of PSO J308-21. For clarity the companion is denoted with a white +.



(a) ***(b)***

Figure 4.41: The FWHM of the quasar (PSO J308-21) shown in figure 4.40 obtained with the *uv*-fitting to the real part of the visibility (left, 4.41a) and to the amplitude (right, 4.41b) for data without noise (open orange circles) and data with noise (closed blue circles) for simulated ALMA observations of 600 s. The red dashed line denotes the true FWHM of the main source.

5 Discussion

5.1 Single Gaussian sources

5.1.1 Image plane analysis

Comparing the results of the flux obtained from the simulated circular Gaussian source and the two simulated elliptical sources as a function of resolution (so per configuration) (figure 4.4, 4.11 and 4.19), there is a clear decreasing effect for resolutions better than 0.2 arcsec. This decrease can get to 90% below the true flux of the source for observations of 600 s.

A gradual decrease was expected due to the decrease in surface-brightness sensitivity. Improving the resolution leads to less signal per resolution element. The surface-brightness decreases for a decreasing angular area and therefore areas with low surface-brightness emission are resolved out.

For all three single Gaussian sources that were investigated in this research, the decrease began at a resolution of around 0.2 arcsec (e.g., figure 4.4), which implies that the start of the decrease is independent of the size of the source.

The observed flux can be influenced by the largest angular scale of the configuration and the surface-brightness sensitivity to varying extents. To investigate which is responsible for this sudden decrease, the peak flux SNR (Signal to Noise Ratio) is found for the circular Gaussian source. The peak flux is taken to be the maximum value in the box selected for the source when obtaining the flux (see figure 3.6). When the peak flux SNR is too low, the source will not be detected through the noise. The lowest value of the peak flux SNR is found for the most extended configuration and is 8.8 for an observation of 60 s, 17 for an observation of 600 s and 43 for an observation of 3600 s. There is no sudden decrease in the peak flux SNR at a resolution of 0.2 arcsec to account for the decrease in the flux. This does not indicate that the surface brightness is the cause of the sudden decrease in the flux.

For a resolution worse than 0.2 arcsec, the flux is well recovered within 10% of the true flux. When the observation time is increased (i.e. the SNR increases), this decreasing effect in the obtained flux is reduced.

The size obtained using the CASA function `imfit` in the image plane (described in section 3.5) for the circular and both the elliptical simulated Gaussian sources (figure 4.5, 4.12, 4.21) shows the same effect as present for the flux. Observing with a resolution better than 0.2 arcsec results in a sudden decrease in the size below that of the original.

The size decreases to 5% of the true size for an observation of 600 s.

Again, a gradual decrease was expected due to the lower surface-brightness sensitivity as explained with the flux above. The low surface-brightness areas at the edges of the source fade away into the noise and are not detected by `imfit` as part of the source.

Increasing the observation time, and consequently increasing the SNR, reduces the decrease in the size. An observation of 3600 s decreases to 33% of the true size whereas an observation of 600 s only recovers 5% of the true size.

5.1.2 *uv*-plane analysis

For the simulated circular Gaussian and the two simulated elliptical Gaussian sources, the diameter is found with the *uv*-plane analysis described in section 3.6. Comparing the results obtained with the image plane analysis using `imfit` and with the *uv*-plane analysis, it is evident that the *uv*-plane analysis is less influenced by the resolution. The size does not drop off as much as the size found with the image plane analysis and the errors are smaller. For the circular Gaussian, an observation of 600 s with the most extended configuration recovers 80% of the original size with the *uv*-plane analysis (figure 4.9) whereas only 5% of the true size is recovered with the image plane analysis using `imfit` (figure 4.5). However, for an observation of 3600 s the size obtained with the *uv*-plane analysis does not show a significant deviation from the true size larger than one sigma (figure 4.9).

It can be seen that the size of the source experiences a similar decrease as the flux for resolutions better than 0.2 arcsec and that both methods of obtaining the size result in this sudden decrease. This indicates that the cause of this decrease has influence on the image plane and the *uv*-plane.

Configurations 12 to 20 have a resolution better than 0.2 arcsec and a largest angular scale 40% lower than for configurations 1 to 11 (figure 3.3). Therefore, the sudden decrease in flux and size is caused by the sudden decrease in the largest angular scale. In examining the flux of any of the single Gaussian sources (e.g., figure 4.4), a jump can be found at a resolution of 0.1 arcsec. This can be explained by the second jump in the largest angular scale between configuration 16 and 17 (figure 3.3).

Since `imfit` fits a Gaussian to the source, the major and minor axes are recovered of an elliptical (Gaussian) source. This is an advantage of `imfit` over the *uv*-plane analysis employed in this research, which fits a circular Gaussian to the *uv*-data. Subsequently, the diameter of the source is found, yet the major and minor axes are not.

It should be noted that in this research the *uv*-fitting was conducted by fitting a circular Gaussian to the *uv*-data. As demonstrated by the results of the two simulated elliptical Gaussian sources, the size is mostly overestimated. However, the errors on the sizes found with the image plane analysis using `imfit` are on the same order of magnitude as the size and therefore the results obtained with `imfit` are not reliable.

Thus, when a source is assumed to be elliptical (or not circular), caution needs to be exercised when applying a circular fit. It is wise to perform both an image plane analysis and a *uv*-plane analysis and see if they are similar. If these differ, an elliptical fit to the *uv*-plane data could be made. This would require a different method of binning the data. The data is now radially binned, but then it needs to be radially and azimuthally binned. This would make the procedure longer and more complicated, but would account for any ellipticity.

5.2 Companion Gaussian sources

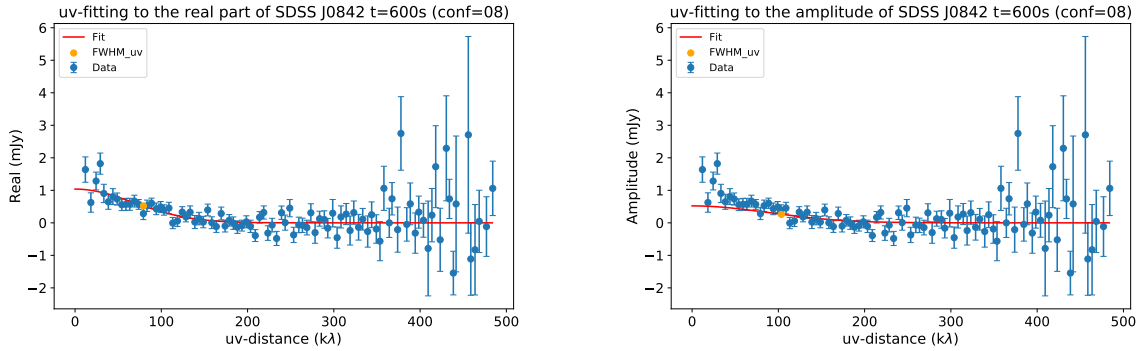
Ten different simulated companion fields detected by Hodge et al. (2013), Simpson et al. (2015), Wardlow et al. (2018) and Decarli et al. (2017) were examined in this research .

In most fields, there is no scatter around the true FWHM larger than 20% of the companion sources on the size of the main source (e.g., figure 4.23). The only field where the companion shows an influence on the size of the main source was PSO J231-20 (J231) originally observed by Decarli et al. (2017) (figure 4.39). This field had the smallest separation between the main source and its companion, and the flux of the companion source is 40% of the flux of the main source. For resolutions worse than 0.2 arcsec, the FWHM found for J231 for the noisy data is up to 20% higher than the original and decreases for improving resolutions. Since this is a clear trend in the FWHM rather than a scatter, it is concluded that the companion source influences the FWHM of the main source. When the resolution gets better than 0.2 arcsec, the companion source of J231 does not influence the size of the main source. This trend in the FWHM is visible for both the real part of the visibility and the amplitude. The FWHM inferrred from the data without noise only shows a 10% increase for the most compact configuration for the real part of the visibility, the FWHM obtained from the data without noise for the amplitude does not show an increase.

Comparing the *uv*-plane analysis of the real part of the visibility and of the amplitude, it is clear that the real part analysis is more accurate and more of the fits converge. The amplitude analysis underestimates the FWHM more than the real part analysis does. To get a more accurate FWHM, the *uv*-plane analysis of the real part of the visibility is preferred over the *uv*-fitting to the amplitude. Figure 5.1 shows the *uv*-fitting of the real part of the visibility (left, 5.1a) and the *uv*-fitting to the amplitude (right, 5.1b) for a simulated ALMA observation of 600 s made with configuration 8 of source SDSS J0842. By eye, it is clearly visible that the fit to the real part is more accurate than the fit to the amplitude.

The Noise to Signal Ratio (NSR) of the amplitude, $\frac{\sigma_A}{A}$, is compared to the NSR of the real part of the visibility, $\frac{\sigma_R}{R}$. By dividing these two ratios figure 5.2 is found. Here it is evident that the NSR of the amplitude is larger than NSR of the real part of the visibility for a *uv*-distance shorter than 100 kλ. This is the part containing the most information about the source size (see figure 5.1), since the FWHM in the *uv*-plane is around 80 kλ. Therefore, this could explain the larger deviation of the amplitude from

the true value.



(a) ***(b)***

Figure 5.1: The *uv*-fitting of source SDSS J0842 for a simulated ALMA observation made with configuration 8 to the real part of the visibility (left, 5.1a) and to the amplitude (right, 5.1b). The FWHM extracted from the fitting to the real part of the visibility (left, 5.1a) is 1.1 arcsec and the FWHM extracted from the fitting to the amplitude (right, 5.1b) is 0.88 arcsec. The true FWHM of the quasar is 1.1 arcsec.

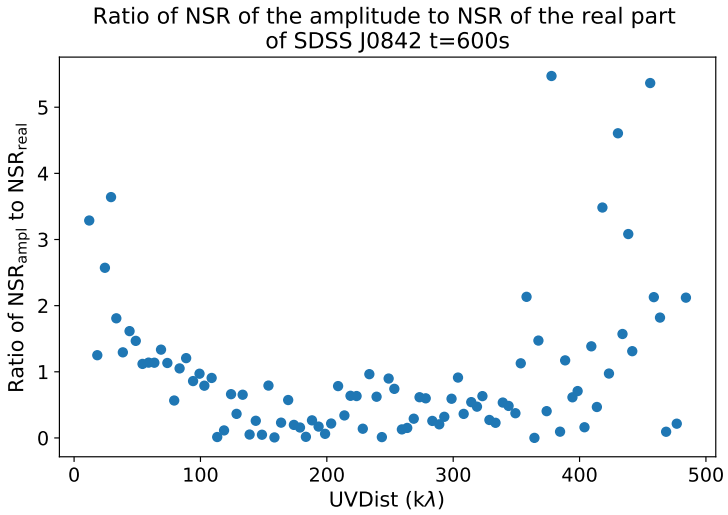


Figure 5.2: The ratio of the NSR (Noise to Signal Ratio) of the amplitude to the NSR of the real part of the visibility of a simulated ALMA observation of 600 s made with configuration 8 of source SDSS J0842 corresponding to the *uv*-plane data shown in figure 5.1.

5.3 Applications

This research was solely based on simulations. However, it can be linked to recent research conducted on SMGs. Two examples from Gullberg et al. (2018) and Rybak et al. (2019) are explained in this section.

Gullberg et al. (2018) used high resolution (0.03 arcsec) Band 7 continuum and [CII] ALMA observations to observe SMGs from ALESS and AS2UDS (ALMA/SCUBA-2

UDS). Due to the low surface-brightness sensitivity, the image plane analysis did not recover the size. Therefore, they utilised a similar *uv*-plane analysis to that used in this research obtaining the size of the sources by fitting to the amplitude. To solve the short-spacing and zero-spacing problem, they combined their high resolution data with low resolution observations of Band 7 dust continuum and [CII] emission and single dish observations. Our research shows evidence that the *uv*-plane analysis is in fact a more accurate method to acquire the FWHM than using an image plane analysis.

Rybak et al. (2019) claim from their observations of [CII] in ALESS 49.1 that ALESS 49.1 has a compact and extended component, where the extended component accounts for 80% of the total [CII] luminosity. In our research, the ALESS 49 field is examined to find the influence of the companion sources on the main source (ALESS 49.1) (figure 4.23). Figure 5.3 shows both the *uv*-data from Rybak et al. (2019) of the [CII] 157.74 μm line and the data found with our research. The [CII] data from Rybak et al. (2019) are fitted with two Gaussians. The data from this research is rescaled to the values of the compact Gaussian. The *uv*-data from configuration 13 is used as the resolution (0.16 arcsec) closest resembles the resolution of Rybak et al. (2019) (0.15 arcsec). Figure 5.3 shows that the extended [CII] component (corresponding to the sharp peak in the figure) is not caused by the companion sources in the field.

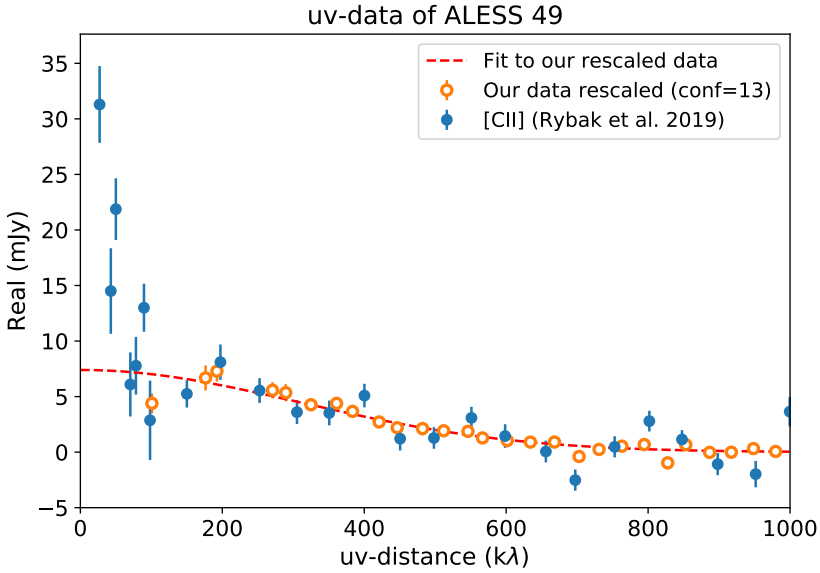


Figure 5.3: The real part of the visibility as a function of the *uv*-distance with [CII] data provided by Rybak et al. (2019) (closed blue circles) and our simulated ALESS 49 from this made with configuration 13 which resembles the resolution of Rybak et al. (2019) the most (open orange circles). Our simulated ALESS 49 data is rescaled to match the Gaussian corresponding to the compact component of the two Gaussians fitted to the [CII] data from Rybak et al. (2019). The red dashed line denotes the Gaussian fitted to the our simulated ALESS 49 data.

5.4 Future research

Future research needs to be dedicated to a more extensive set of different sources. This research only explored the effects of improving resolutions on Gaussian sources, yet in reality sources are rarely Gaussian. Hodge et al. (2016) found a Sérsic index of $n = 0.9 \pm 0.2$ corresponding to a disk-like dust emission from the median light profile of 16 luminous SMGs from ALESS with 0.16 arcsec resolution ALMA observations in Band 7. However, the Sérsic index for a Gaussian source is $n = 0.5$, which implies that SMGs are in fact not Gaussian. It is therefore valuable to investigate the influence of improving resolution for observation of non-Gaussian sources to the flux and size found with both an image plane analysis and a uv-plane analysis.

It was outside the scope of this research to do an elliptical fitting to the uv-data. Yet, realistically sources are seldom circular. It would therefore be encouraged to investigate the major and minor axes found for elliptical sources and companion sources.

To investigate the influence of companion sources to the observed size of the main source, the distance and flux of the companion source can be altered. The influence of many faint companion sources could also be valuable as these may stay undetected.

6 Conclusions

This research presents simulated ALMA observations with different antenna configurations of simulated Sub-Millimeter Galaxies (SMGs) as Gaussians. A circular Gaussian source and two elliptical Gaussian sources are investigated as well as simulated SMG companion fields with Gaussian sources originally observed by Hodge et al. (2013), Simpson et al. (2015) and Wardlow et al. (2018) and simulated quasar fields with a companion originally observed by Decarli et al. (2017).

The simulated observations of the single simulated Gaussian sources are used to find the flux and size. The size is obtained using both an image plane analysis with the CASA function `imfit` and a *uv*-plane analysis by fitting the *uv*-data with a single circular Gaussian. The results demonstrate that the flux decreases below the true flux of the source to 5% of the true flux for a resolution better than 0.2 arcsec for an observation of 600 s. This sudden decrease is most likely caused by the largest angular scale of the observations. The configurations with a resolution better than 0.2 arcsec have a sudden decrease in largest angular scale. The surface-brightness sensitivity does not show a sudden decrease around 0.2 arcsec and is therefore believed to not be the cause of the sudden decrease in the flux.

Comparing the sizes found with `imfit` and with the *uv*-plane analysis, it is evident that the *uv*-plane analysis recovers the size more accurately. However, the *uv*-fitting is performed with a circular Gaussian and thus does not recover ellipticity, which `imfit` does. Another advantage of the *uv*-plane analysis is that it is less influenced by the resolution and does not show as big a decrease as the image plane acquired size. Consistent for the flux and the two methods of obtaining the size is that an increase in observation time (and thus in SNR) reduces the errors and recovers the true value more accurately.

The simulated companion fields are examined to find the influence of the companion sources on the size of the main source found with the *uv*-plane analysis by fitting a single circular Gaussian to the *uv*-plane data of the main source. This procedure is performed on data without noise and data with noise to see if the noise cancels out the influence or enhances it. The *uv*-fitting is performed to the real part of the visibility and the amplitude. As the simulated sources are Gaussians, the Fourier transform should be negligible in the imaginary part. As the influence of companion sources on the main source is unknown, the amplitude is also used. Of the ten simulated companion fields that are investigated, nine recover the size of the main source accurately within 20% of

the true FWHM. The only simulated companion field that shows a difference is PSO J231-20, originally observed by Decarli et al. (2017). This was the field with the smallest separation between quasar and companion. The size of the quasar is mostly overestimated as the flux of the companion is combined with the flux of the quasar and leads to a contribution is the size.

The results demonstrate that the *uv*-fitting to the real part of the visibility is more reliable than to the amplitude.

All in all, this research shows that in general companion sources do not influence the FWHM found for the main source.

Acknowlegdements

I would like to thank Matus Rybak for supervising me and answering any and all questions I had. This research could not have been completed without his help. Furthermore, I would like to thank Jacqueline Hodge for sharing her wisdom during the weekly group meetings and for being my second supervisor. I would also like to thank Dorothea Samtleben for being my second corrector. Last but not least, a special thanks to Merel for her mental support and rarely letting me win with Set.

Appendix

A.1 2D Gaussian Fourier transform

When looking at the visibilities[1], the size of a source is acquired using a two dimensional Gaussian model in the uv-plane. This model needs to be Fourier transformed to get the size in the image plane. This section covers the derivation of this Fourier transform.

Assumed is that the source in the image plane is a two dimensional Gaussian of the form as visualised in equation A.1.

$$f(x,y) = A \exp\left(-\left(\frac{x^2}{2\sigma_x^2} + \frac{y^2}{2\sigma_y^2} \right) \right) \tag{A.1}$$

Here A is the amplitude of the Gaussian and is equal to the total flux of the source. x and y are the coordinates in the image plane and σ_x and σ_y are related to the Full Width at Half Maximum (FWHM) as is portrayed in equation A.2.

$$\sigma = \frac{\text{FWHM}}{2\sqrt{2\ln 2}} \tag{A.2}$$

When the Gaussian is circularly symmetric $\sigma_x = \sigma_y$. To find the values of σ_x, σ_y and A, the visibilities need to be fitted to the Fourier transform of $f(x,y)$ (equation A.1). The Fourier transform is done using equation A.3[2].

$$F(u,v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) \exp(-2\pi i u x) \exp(-2\pi i v y)\, dx\, dy \tag{A.3}$$

Here u and v are coordinates in the uv-plane, corresponding to the vectorial separation between each antenna pair in wavelengths.

Substituting in equation A.1 gives:

$$F(u,v) = \int\int_{-\infty}^{+\infty} A \exp\left(-\left(\frac{x^2}{2\sigma_x^2} + \frac{y^2}{2\sigma_y^2} \right) \right) \exp(-2\pi i(ux + vy))\, dx\, dy\,.$$

These two integrals can be computed separately like so:

$$F(u,v) = A \int_{\infty}^{\infty} \exp\left(-\frac{x^2}{2\sigma_x} \right) \exp(-2\pi i u x)\, dx \int_{\infty}^{\infty} \exp\left(-\frac{y^2}{2\sigma_y} \right) \exp(-2\pi i v y)\, dy\,. \tag{A.4}$$

[1] Visibilities are explained in detail in section 2.3.

[2] From now on the double integral will be displayed as $\int\int_{-\infty}^{+\infty}$.

Firstly, a general example will be explained and then it will be applied to the integral displayed in equation A.4.

Let us first look at a general Gaussian function:

$$f(x) = \exp(-ax^2 + bx). \tag{A.5}$$

Now taking the integral of this function gives equation A.6.

$$I = \int_{-\infty}^{\infty} \exp(-ax^2 + bx)\, dx \tag{A.6}$$

It is convenient to calculate the square of I, as will become clear in the next steps.

$$\begin{aligned} I^2 &= \left(\int_{-\infty}^{\infty} \exp(-ax^2 + bx)\, dx \right)^2 \\ &= \int_{-\infty}^{\infty} \exp(-ax^2 + bx)\, dx \int_{-\infty}^{\infty} \exp(-ay^2 + by)\, dy \\ &= \int\int_{-\infty}^{+\infty} \exp[-a(x^2 + y^2) + b(x + y)]\, dx\, dy \end{aligned}$$

Define new parameters $h = -\dfrac{b}{2a}$ and $k = ah^2 + bh$. With these new parameters equation A.5 can be rewritten as $f(x) = a(x-h)^2 + k$. Substituting this in in I^2 makes it easier to solve.

$$I^2 = \int\int_{-\infty}^{+\infty} \exp[-a(x-h)^2 - a(y-h)^2] \exp(2k)\, dx\, dy$$

Shifting the Gaussian to be centered at (0,0) by substituting $\widetilde{x} = x - h$ and $\widetilde{y} = y - h$ gives:

$$I^2 = \int\int_{-\infty}^{+\infty} \exp(-a\widetilde{x}^2 - a\widetilde{y}^2) \exp(2k)\, d\widetilde{x}\, d\widetilde{y}.$$

Changing to polar coordinates with $r = \sqrt{\widetilde{x}^2 + \widetilde{y}^2}$ and $d\widetilde{x}\, d\widetilde{y} = r\, dr\, d\theta$ gives:

$$I^2 = \int_0^{2\pi} \int_0^{\infty} r \exp(-ar^2) \exp(2k)\, dr\, d\theta.$$

Since $\exp(2k)$ is a constant and there is no θ dependence, this results in:

$$I^2 = 2\pi e^{2k} \int_0^{\infty} r e^{-ar^2}\, dr.$$

Now substituting $s = -r^2$ and $dr = -\dfrac{1}{2r} ds$ gives:

$$\begin{aligned} I^2 &= \pi e^{2k} \int_{-\infty}^{0} e^{as}\, ds \\ &= \frac{\pi}{a} e^{2k}. \end{aligned}$$

Taking the square root of I^2 gives the desired answer for the integral: $I = \sqrt{\frac{\pi}{a}} e^k$.

Looking back at the initial integral of equation A.4, it can be seen that for the first integral of dx $a = \frac{1}{2\sigma_x^2}$ and $b = -2\pi i u$ and for the second integral of dy $a = \frac{1}{2\sigma_y^2}$ and $b = -2\pi i v$. This gives the following:

$$F(u,v) = A\left(\sigma_x\sqrt{2\pi}e^{-2\pi^2\sigma_x^2u^2}\right)\left(\sigma_y\sqrt{2\pi}e^{-2\pi^2\sigma_y^2v^2}\right), \tag{A.7}$$

$$= A\,2\pi\sigma_x\sigma_y e^{-2\pi^2(\sigma_x^2u^2+\sigma_y^2v^2)}. \tag{A.8}$$

Since the relation between σ_x and σ_y with the FWHM is known to be FWHM $= 2\sqrt{2\ln 2}\,\sigma_{xy}$, when the visibilities are fitted with equation A.8, the FWHM can be extracted and the size of the source can be found.

When assuming circular symmetry the Gaussian function and its Fourier transform become:

$$f(x,y) = A\exp\left(-\frac{x^2+y^2}{2\sigma_{xy}^2}\right), \tag{A.9}$$

$$F(u,v) = A2\pi\sigma_{xy}^2 e^{-2\pi^2\sigma_{xy}^2(u^2+v^2)}. \tag{A.10}$$

Bibliography

Blain, A. W., Smail, I., Ivison, R. J., et al. (2002). Submillimeter galaxies. *Physics Report*, 369(2):111–176.

Briggs, D. S. (1995). *High Fidelity Deconvolution of Moderately Resolved Sources.* Ph.D. The New Mexico Institute of Mining and Technology.

Burke, B. F. and Graham-Smith, F. (2009). *An Introduction to Radio Astronomy*. Cambridge: Cambridge University Press.

Bustos, R., Rubio, M., Otárola, A., and Nagar, N. (2014). Parque Astronómico de Atacama: An Ideal Site for Millimeter, Submillimeter, and Mid-Infrared Astronomy. *Publications of the Astronomical Society of the Pacific*, 126(946):1126–1132.

Casey, C. M., Narayanan, D., and Cooray, A. (2014). Dusty Star-Forming Galaxies at High Redshift. pages 1–154.

Condon, J. J. and Ransom, S. M. (2016). *Essential Radio Astronomy*. Princeton University Press.

Da Cunha, E., Walter, F., Smail, I. R., et al. (2015). AN ALMA SURVEY of SUBMILLIMETER GALAXIES in the EXTENDED CHANDRA DEEP FIELD SOUTH: PHYSICAL PROPERTIES DERIVED from ULTRAVIOLET-TO-RADIO MODELING. *Astrophysical Journal*, 806(1).

Decarli, R., Walter, F., Venemans, B. P., et al. (2017). Rapidly star-forming galaxies adjacent to quasars at redshifts exceeding 6. *Nature*, 545(7655):457–461.

Geach, J., Dunlop, J., Halpern, M., et al. (2017). The scuba-2 cosmology legacy survey: 850 μm maps, catalogues and number counts. *Monthly Notices of the Royal Astronomical Society*.

Gullberg, B., Swinbank, A. M., Smail, I., et al. (2018). The dust and [CII] morphologies of redshift $\sim$4.5 sub-millimeter galaxies at $\sim$200pc resolution: The absence of large clumps in the interstellar medium of high-redshift galaxies. *The Astrophysical Journal*, 859(1):12.

Hodge, J. A., Karim, A., Smail, I., et al. (2013). AN ALMA SURVEY OF SUBMILLIMETER GALAXIES IN THE EXTENDED CHANDRA DEEP FIELD SOUTH: SOURCE CATALOG AND MULTIPLICITY. *The Astrophysical Journal*, 768(1):91.

Hodge, J. A., Smail, I., Walter, F., et al. (2018). ALMA reveals evidence for spiral arms, bars, and rings in high-redshift submillimeter galaxies. (1).

Hodge, J. A., Swinbank, A. M., Simpson, J. M., et al. (2016). Kiloparsec-scale dust disks in high-redshift luminous submillimeter galaxies. *The Astrophysical Journal*, 833(1):1–15.

McMullin, J., Waters, B., Schiebel, D., et al. (2007). CASA Architecture and Applications. *Astronomical Data Analysis Software and Systems XVI*, 376:127.

Planck Collaboration et al. (2015). Planck 2015 results. XIII. Cosmological parameters.

Remijan, A., Biggs, A., Cortes, P. A., et al. (2019). ALMA Technical Handbook, ALMA Doc. 7.3, ver. 1.0.

Rivera, G. C., Hodge, J. A., Smail, I., et al. (2018). Resolving the ism at the peak of cosmic star formation with alma: the distribution of co and dust continuum in z $\sim$2.5 submillimeter galaxies. *The Astrophysical Journal*, 863(1).

Robitaille, T. P. and Whitney, B. A. (2010). The present-day star formation rate of the milky way determined from spitzer-detected young stellar objects. *Astrophysical Journal Letters*, 710(1 PART 2):1–12.

Rybak, M., Rivera, G. C., Hodge, J. A., et al. (2019). Strong FUV fields drive the [CII]/FIR deficit in z$\sim$3 dusty, star-forming galaxies.

Schirm, M. R. P., Wilson, C. D., Parkin, T. J., et al. (2014). Herschel -SPIRE Fourier Transform Spectrometer Observations of Excited CO and [C I] in the Antennae (NGC 4038/39): Warm and Cold Molecular Gas. *The Astrophysical Journal*, 781(2).

Schweizer, F., Burns, C. R., Madore, B. F., et al. (2008). A new distance to the antennae galaxies (NGC 4038/39) based on the type ia supernova 2007. *Astronomical Journal*, 136(4):1482–1489.

Simpson, J. M., Smail, I., Swinbank, A. M., et al. (2015). THE SCUBA-2 COSMOLOGY LEGACY SURVEY: ALMA RESOLVES the BRIGHT-END of the SUB-MILLIMETER NUMBER COUNTS. *Astrophysical Journal*, 807(2).

Swinbank, A. M., Simpson, J. M., Smail, I., et al. (2014). An ALMA survey of sub-millimetre galaxies in the extended chandra deep field south: The far-infrared properties of SMGs. *Monthly Notices of the Royal Astronomical Society*, 438(2):1267–1287.

Wardlow, J. L., Simpson, J. M., Smail, I., et al. (2018). An ALMA survey of CO in submillimetre galaxies: Companions, triggering, and the environment in blended sources. *Monthly Notices of the Royal Astronomical Society*, 479(3):3879–3891.

Weiss, A., Kovacs, A., Coppin, K., et al. (2009). The LABOCA Survey of the Extended Chandra Deep Field South. *arXiv*, astro-ph.C.

Wilson, T. L., Rohlfs, K., and Hüttemeister, S. (2009). *Tools of Radio Astronomy*. Springer Berlin Heidelberg.

Wright, E. L. (2006). A Cosmology Calculator for the World Wide Web.