Chapter 1
Scientific Introduction

1.1 Challenges to the $\Lambda$CDM model

In the current cosmological structure formation scenario, galaxy building is driven by the hierarchical process of merging, which in turn, is dominated by the dark matter component (e.g. Sanders et al., 1988; Hopkins et al., 2009b). This $\Lambda$CDM model of galaxy formation has been very successful in explaining the structure of the Universe at scales larger than galaxies, but it presents several challenges at galactic and sub-galactic scales. According to this model, dark matter haloes grow by the merging of smaller mass systems (White & Rees, 1978; Davis et al., 1985; Navarro et al., 1996, 1997; Vogelsberger et al., 2014). During the merger process, systems with masses $\leq 10^{10}$ M$_\odot$ are predicted to survive. For example, the Milky Way should host $\sim 10^3$ sub-haloes, but only a few tens of satellites have been observed (Klypin et al., 1999; Moore et al., 1999). A key question is, are these satellites too faint to be detected, or is the distribution of matter in galaxies smoother than what is predicted?

To address this tension, it is necessary to detect and quantify the possible dark matter-dominated satellites within and beyond the Local Group galaxies, and evaluate if their properties are consistent with the $\Lambda$CDM predictions or not.

Another key test to determine if the evolution of galaxies follows the expectations from $\Lambda$CDM is through tracing the cold molecular gas distribution in and around galaxies. The gas fraction in galaxies is expected to be larger at early stages of galaxy formation, when the galaxy growth is thought to be dominated by gas-rich mergers (e.g. Hopkins et al., 2009a). Furthermore, as a consequence of the hierarchical galaxy formation process, merger remnants should have a central starburst component at the inner central part ($< 1$ kpc) that is embedded
in a massive molecular gas reservoir \cite{Kormendy1992, Hopkins2009b, Hopkins2013}. However, the properties of the cold molecular gas are almost completely observationally unconstrained at high redshift. This is primarily because the cold interstellar medium has an intrinsically low surface brightness. Therefore, another key question is,

\textit{what are the scales at which the cold molecular gas is distributed at high-redshift, and how does it effect galaxy growth?}

The scales at which it is possible to directly test the theoretical predictions are small (sub-kpc), requiring sensitive high angular-resolution observations of the cold molecular gas to be made. Therefore, this kind of study is currently limited to a few extremely bright and rare galaxies, which may not be representative of the most common galaxy-building process.

In the hierarchical structure formation process, mergers are responsible also for the formation of super massive black holes (SMBHs) that are thought to reside at the centre of all massive galaxies \cite[e.g.][]{Sanders1988, Volonteri2010}. An intriguing consequence of this merger sequence is the inevitable formation of binary systems of SMBHs. In fact, if each halo contains a SMBH at its centre, when two galaxies collide, the two SMBHs could migrate towards the centre of the gravitational potential of the combined system, as a consequence of dynamical friction \cite{Begelman1980}. An alternative avenue for the formation of primordial SMBHs relies on the direct collapse of a massive gas cloud, leading to the formation of a black hole seed, which grows at the expense of the surrounding material. Even though the paucity of binary SMBH systems at low redshift seems to be in agreement with the theoretical expectations from the $\Lambda\text{CDM}$ model, it is challenging to obtain a sample of bound pairs of SMBHs at high redshift that can be statistically used for discerning between these two formation processes \cite{Volonteri2016, Rosas-Guevara2019}. It is even more challenging to study the coalescence of two such SMBHs, as to date, there is only one tentative detection of an orbital motion in a binary SMBH system \cite{Bansal2017}. Therefore, it is not clear what is the process that leads to the formation of SMBHs. Moreover, if their formation is due to merging, the question then becomes,

\textit{for how long can two SMBHs coalesce before finally merging, and can this process be directly observed?}

In both scenarios, the accretion of a SMBH is inevitable during the so-called
Active Galactic Nucleus (AGN) phase. In the current galaxy formation paradigm, a natural consequence of mergers is the presence of large-scale inflows of gas towards the centre of galaxies, which trigger both a central starburst and the AGN [Di Matteo et al. 2005]. The extremely large amount of energy released by the SMBH during this active phase may regulate the rate at which stars are formed through winds and collimated outflows [Silk & Rees 1998], as suggested by the tight relation between the mass of the SMBH and the velocity dispersion of the stars within the central bulge [Ferrarese & Merritt 2000]. However, the feeding and feedback of AGN cover a wide range of scales, and it is not clear how the AGN feeding on sub-kpc scales is connected to the feedback on Mpc scales. Therefore, high angular-resolution observations of cold molecular gas accretion in AGN host galaxies can provide a direct probe to understanding the connection between the fuelling of SMBHs and their feedback on the surrounding medium and, consequently, on the star formation process.

As a consequence of the open questions raised above, the most stringent tests to the ΛCDM model rely on the properties of galaxies on kpc and sub-kpc scales. Therefore, there is the need to obtain detailed information on the dark matter distributions of galaxies, and to resolve their stellar and molecular gas content. In order to directly probe the expected AGN-regulated galaxy evolution, it is necessary to directly image the outflows at cosmological epochs when both SMBHs and star formation were extremely active. In this thesis, we address these small-scale challenges to the ΛCDM scenario by combining very high angular resolution imaging at optical/near-infrared (NIR) and radio wavelengths with the resolving power of gravitational lensing.

According to the theory of general relativity, a gravitational field bends spacetime and, therefore, the geodesic of a light ray is also bent because of the presence of a massive object. As a consequence, multiple distorted magnified images of a background source may be observed if a massive object (acting as a lens) lies between the distant source and the observer (see Fig. 1.1). The gravitational lensing effect allows us to probe the total (baryonic and dark) matter density distribution in distant galaxies, but also gives us a magnified view of gas, stars and AGN in high redshift objects. When gravitational lensing is combined with high angular-resolution observations, it is then possible to access the crucial scales needed to test the validity of the ΛCDM structure formation model. In this introductory chapter, we review the key concepts of gravitational lensing in Section 1.2 and studies of gravitational lensing at radio wavelengths in Section 1.3. Finally, in Section 1.4 we give a brief outline of this thesis.
1.2 The Universe under a cosmic telescope

The aim of this section is to give an introduction to the basic concepts of gravitational lensing and to explain why gravitationally lensed sources allow us to observe the small-scale structure within galaxies. For detailed reviews on the subject, we redirect the interested reader to Schneider et al. (1992), Blandford & Narayan (1992) and Treu (2010).

1.2.1 Basic concepts of gravitational lensing

A gravitational lensing system consists of a background source, whose emitted radiation is deflected by a lensing galaxy, before arriving to an observer (see Fig. 1.1). Since the physical size of a lensing galaxy is much smaller than the relative distances between the source, lens and observer, the mass distribution of the lens is approximated to a bi-dimensional distribution, called the thin screen approximation. The source is also assumed to lie on a plane, which throughout this thesis, we refer to as the source-plane. The light from a background source is deflected by an angle $\alpha$ following the lens equation,

$$\beta = \theta - \alpha(\theta),$$

meaning that a source at true angular position $\beta$ is seen by the observer at angular positions $\theta$. The angles involved in the lens equation are shown in Fig. 1.1. As Eq. 1.1 has multiple solutions for fixed $\beta$, multiple images of a background source can be observed in what shall be referred to as the lens-plane.

An axially-symmetric lensing galaxy can produce multiple images when the central value of the surface mass-density is larger than $\Sigma_{\text{crit}}$, that is,

$$\Sigma \geq \Sigma_{\text{crit}}, \quad \Sigma_{\text{crit}} = \frac{c^2}{4\pi G D_{LS}D_L} D_S,$$

where $\Sigma_{\text{crit}}$ is the critical surface mass-density, which depends on the angular diameter distance to the lens $D_L$, to the source $D_S$, and their relative angular diameter distance $D_{LS}$ (as also shown in Fig. 1.1). $G$ is the gravitational constant and $c$ is the speed of light.

If the background source lies exactly behind a lensing galaxy with a circularly symmetric gravitational potential, it will be distorted into a ring-shaped image, which is known as an Einstein ring (e.g., see Fig. 1.2). The Einstein radius, $\theta_E$, depends on the relative distances between the source and the lens, but also on the mass of the deflector,

$$\theta_E = \left( \frac{4GM}{c^2} \frac{D_{LS}}{D_L D_S} \right)^{1/2},$$

(1.3)
Section 1.2. The Universe under a cosmic telescope

Figure 1.1: (Upper) Schematic representation of the various components involved in a strong gravitational lensing configuration, such as a background source, a lensing galaxy, the multiple images (which in the upper panel are extended creating gravitational arcs) and the observer (at Earth). Courtesy of R. Schulz. (Lower) Angles involved in a gravitational lensing system under the assumption of a thin screen approximation. Side view of the trajectory of a light ray emitted by a source (indicated by the solid line) is bent by a lens, producing a lensed image seen by an observer. We show only one component of the angles, which are related to each other by the lens equation (Eq. 1.1).
Figure 1.2: Example of gravitationally lensed objects observed with the Keck-II Telescope at $K'$-band with the adaptive optics mode. (Upper left) JVAS B1938+666 is an example of an Einstein ring. The emission at the centre is due to the elliptical lensing galaxy [Lagattuta et al. 2012]. (Upper right) CLASS B0128+437 is in a fold-configuration, showing four point-like lensed images, with a hint of an Einstein ring and faint emission from the lensing galaxy [Lagattuta et al. 2010]. This system shows both flux ratio and astrometric anomalies. (Lower) CLASS B2045+265 consists of four images (A, B, C, D) in a cusp-configuration. With this deep NIR imaging, [McKean et al. 2005] demonstrated that the violation of the cusp-relation (see Section 1.2.3) in this system can be completely attributed to the dwarf galaxy (G2), which is a satellite of the main lensing galaxy (G1).
where $M$ is the mass of the lensing galaxy within the Einstein radius. Even though only a few lensing systems show a complete Einstein ring, the Einstein radius can be estimated for any multiply-imaged system, as the typical angular separation between the lensed images is of the order of $2\theta_E$, but more commonly, this is solved for during lens modelling. The Einstein radius is one of the most important properties of a gravitational lensing system, as it provides an estimate of the mass enclosed within $\theta_E$, which includes both the luminous and dark matter components. A more detailed description on how to obtain information on the mass density profile of a lensing galaxy is given in Section 1.2.2.

Gravitational lensing preserves surface brightness, what it changes is the solid angle under which the source is observed, that is, the solid angle element $\delta\beta^2$ is mapped onto the solid angle $\delta\theta^2$. The change of the solid angle under which the source is seen implies that a source can be magnified or de-magnified, such that,

$$
\mu = \frac{\delta\theta^2}{\delta\beta^2}.
$$

Therefore, the magnification $\mu$ of a lensed object is simply given by the ratio between the area in the image-plane and the area in the source-plane. An important consequence of the lensing magnification is that extremely distant galaxies, that are usually too faint to be detected even by the largest facilities, can be enlarged and magnified. Therefore, massive galaxies act as natural telescopes, giving the unique opportunity to study intrinsically faint distant sources and to reconstruct their undistorted morphology, even at high redshift.

The magnification of a gravitational lens is dependent of the structure of the lens, and can be expressed as,

$$
\mu = \frac{1}{[(1 - \kappa)^2 - \gamma^2]},
$$

where $\gamma$ is called the shear (from external mass structure) and $\kappa$ is called the convergence, and is dependent on the surface-mass density of the lens ($\Sigma$), such that,

$$
\kappa = \frac{\Sigma}{\Sigma_{\text{crit}}},
$$

(1.6)

In addition, the convergence can be expressed in terms of the lensing potential, $\Phi$,

$$
\kappa = \nabla^2 \Phi.
$$

(1.7)

Therefore, to determine the magnification of an object, some knowledge of the lens structure is needed.

The magnification is formally infinite at the so-called critical curves (e.g. for $\kappa = 1$ when $\gamma = 0$), which consist of two closed lines that do not intersect.
Figure 1.3: Example of lensing configurations of an elliptical lens mass density distribution for
an extended background source. The image configurations of a source crossing the fold caustics
are shown on the left, while on the right there are the theoretical image configurations when
a source crosses the cusp caustic. The critical and caustic curves are indicated on each image.
Figure adapted from Narayan & Bartelmann (1996).

(see Fig. 1.3). These two critical curves are distinguished into tangential and radial components. An image located along the tangential critical curve is highly
distorted and magnified along this critical line, while an image situated in the
proximity of a radial critical line would be stretched in a radial direction. The
critical curves have corresponding closed lines in the source plane, which are called
cautics, and they can intersect. Lensed images at the location of the tangential
critical curve can lead to a precise measurement of the total mass enclosed within
$\theta_E$, while images close to the radial critical curve can accurately constrain the
inner slope of the mass density profile of the lensing galaxy. Nevertheless, in
order to physically quantify the mass within $\theta_E$, it is necessary to know precisely
the redshift of both the lensing galaxy and the background source.

1.2.2 Gravitational lens modelling

Lens modelling aims to determine an approximate mass density distribution for
the object acting as the lens, which can be a single galaxy or groups/clusters of
galaxies. The determination of the mass density distribution is important not
only for studying the lensing galaxy, but also for the analysis of the background
source, whose structure is reconstructed automatically while solving for the lens
equation (see Eq. 1.1). Therefore, changes to the lens mass model (the optics
of the system) will change our perception of the background source. However,
despite its apparent simplicity, the lens equation is often not analytically solvable
for any given gravitational potential. Therefore, it is necessary to make some approximations in order to find the most realistic mass model for a gravitational lens.

A classic approach that is often used when modelling the mass density distribution of a lensing galaxy is to parametrize the gravitational potential with an analytic formalism. In other words, this method consists of assuming that the lens potential (and the surface brightness distribution of the source) can be described by a set of parameters. The complexity of the mass model chosen depends on the available observational constraints, which consist of the positions of the lensed images and their flux densities. The number of degrees of freedom, that is, the difference between the number of observational constraints and the number of free parameters in the model, is, therefore, the important quantity of a lens mass model. The higher the number of observational constraints available, the more robustly the lens mass-density distribution can be tested.

The model that describes the total mass density distribution of the lensing galaxy is also referred to as the macro-model throughout this thesis. A simple macro-model that can describe the mass density distribution of lensing galaxies is the singular isothermal ellipsoid (SIE). This density profile depends on the radius $r$ as $\rho(r) \propto r^{-2}$. Based on observational results, the SIE is often used to describe the mass density distribution of lenses. For example, Auger et al. (2010) found that the lenses in the Sloan Lens ACS Survey (SLACS) sample are close to an SIE profile, as the mass density profile $\rho(r) \propto r^{-\gamma}$ has been found to have a slope $\gamma = 2.078 \pm 0.027$. The SIE seems, therefore, a valid approximation for the simple modelling of strong gravitational lenses. An SIE mass model has 7 variables, which are the mass scale, the position of the lensing galaxy, the ellipticity and its position angle, and an external shear and its position angle. The external shear term takes into account the effect of a gravitational potential that is external to the lensing galaxy, as for example a nearby galaxy companion or cluster. The position of each background source provides two constraints to the model (and each flux-ratio can provide another). Therefore, extended sources that are gravitationally magnified in both tangential and radial directions can provide a large number of constraints to the mass model.

An alternative method to parametric models consists of grid-based lens models, where the entire surface brightness distribution of the source is used to constrain the lens potential (e.g. Warren & Dye 2003; Koopmans 2005; Vegetti & Koopmans 2009a). Therefore, given the very large number of constraints available, it is possible to more confidently detect and quantify possible small-scale deviations to a smooth macro-model (Vegetti & Koopmans 2009b, but also see Section 1.2.3). An important advantage of grid-based models consists of the
accurate reconstruction of the source structure. By using pixellated source reconstructions it becomes possible to actually investigate the physical properties of high redshift galaxies, such as their physical size or intrinsic luminosity, which can be used to understand their evolutionary stage.

1.2.3 Deviations from a smooth model

There are several techniques that can be used to detect deviations from a globally smooth macro-model, which may indicate the presence of sub-haloes within the lensing galaxy or additional complex structure that has not been accounted for. One of these methods is based on the observed flux-ratios of the lensed images. The flux densities of the lensed images are related to the magnification factor $\mu$, which depends, via the convergence, on the second derivative of the gravitational potential of the lensing galaxy (see Eq. 1.7).

Interestingly, in quadruply-imaged systems, the magnification factors of the closest pair of images (A and B, in fold systems) and of the triplet images (A, B and C, in cusp systems) should sum asymptotically to zero as the source approaches the cusp/fold, for a smooth mass density distribution, such that,

$$|\mu_A| = |\mu_B| \text{ (fold relation)}; \quad |\mu_A| = |\mu_B| + |\mu_C| \text{ (cusp relation)}.$$  \hspace{1cm} (1.8)

Therefore, there is a clear prediction for the expected flux-ratios between the images in these lensing systems. Some examples of quadruply-imaged sources in the cusp and fold configuration are shown in Fig. 1.2.

Almost all quadruply-imaged sources do not obey the fold/cusp relation (Eq. 1.8), and these are typically referred to as flux-ratio anomalies (see Fig. 1.4). This anomaly can be attributed to the presence of a perturbation (like from the population of dark matter sub-haloes predicted by $\Lambda$CDM), which can modify the magnification of one or more of the lensed images, consequently changing the observed flux density (e.g. Metcalf & Madau, 2001a). For this reason, the study of flux-ratio anomalies in lenses has become an established method to detect and quantify dark matter sub-haloes associated with high-redshift lensing galaxies or along the line-of-sight. For example, an investigation of the flux-ratio anomalies observed in seven quadruply-imaged sources by Dalal & Kochanek (2002) estimated that $\sim 2$ per cent of the projected mass density of the lensing galaxy must be in the form of dark matter sub-haloes. However, this mass fraction seems to be high when compared to $\Lambda$CDM simulations (Mao et al., 2004). Therefore, alternative dark and baryonic matter perturbations may be playing a role in changing the properties of the lensed images. For example, in some cases the flux-ratio anomaly can be clearly related to luminous structures, like a luminous satellite in CLASS B2045+265 (see Fig. 1.2).
or an edge-on disk in the lensing galaxy of \textit{CLASS B1555+375} (Hsueh et al., 2016).

A mass perturbation can also change the position of the lensed images, that is, they are observed in a different position from what is predicted for a smooth mass distribution. Here, the lens system shows an astrometric anomaly, as for example in the case of \textit{CLASS B1152+199} (see Fig. 1.4). The analysis of this system seems to require multiple compact structures with masses $\leq 10^7 \, M_\odot$ in order to reproduce the observed anomalous bending of the radio jets (Metcalf, 2002). In another system, \textit{MG B2016+112}, which is the focus of Chapter 4 of this thesis, the astrometric anomalies have been attributed almost completely to a luminous satellite galaxy that is close in projection to the main lens (More et al., 2009).

The $\mu$Jy beam$^{-1}$ surface brightness sensitivity that can be provided by radio interferometers is crucial to recover the full extent of radio-loud gravitational arcs at mas-levels, which is necessary to understand the small angular-scale structure of the mass distribution of the lensing galaxy. Such sensitive VLBI observations at high angular resolution of gravitationally lensed arcs can theoretically enable the detection of possible perturbations due to the smallest sub-haloes, as shown in Fig. 1.5. The presence of a sub-halo close to the lensed arcs in projection can significantly distort the position and the surface brightness distribution of the lensed images. For example, a sub-halo with mass $10^6 \, M_\odot$ has an Einstein radius of $\sim 3$ mas, which is currently detectable only with VLBI observations. The detection of such low-mass haloes can place important constraints on the mass-function of dark matter (sub-)haloes and be used to determine the properties of dark matter (see Fig. 1.5).

1.3 Gravitational lensing at radio wavelengths

In order to accurately constrain the mass density distribution of lensing galaxies, systems with a radio-loud background source have some important advantages. First, the radio emission is not reddened nor obscured by the dust or the bright optical emission from the lensing galaxy. Therefore, the lensed images can be easily identified and their flux density can be confidently measured. Moreover, radio-loud sources often have extended structure, which can, for instance, be due to the presence of AGN jets, that extend from several pc- to kpc-scales. This means that gravitationally lensed radio sources can show extended multiple images, potentially connected into gravitational arcs, making them the most suitable systems for detailed studies of the mass density distribution of the foreground lensing galaxies.
Chapter 1

Figure 1.4: (Upper) Example of a flux-ratio anomaly between images A and B, observed from radio imaging at 1.7 GHz at mas-resolution in the lensed radio source CLASS B0712+472, which is in the fold configuration (Hsueh et al., 2017). (Lower) The doubly-imaged radio jets of the radio source CLASS B1152+199 show an astrometric anomaly (Metcalf, 2002).
As lens galaxies are typically massive ellipticals, with a mass within the Einstein radius of $10^{11}-12 \, M_\odot$, the maximum image separations in most galaxy-scale systems is about 1.2 arcsec. Therefore, in order to spatially resolve the multiple lensed images, it is necessary to observe the lens systems at sub-arcsec resolution. Moreover, for a detailed study of the mass density distribution of the deflector, a precise astrometry of the lensed images, at possibly sub-mas levels, is required. Therefore, only mas angular-resolution observations can investigate the small-scale properties of the mass density distribution of a lensing galaxy.

1.3.1 High angular resolution with interferometry

The angular resolution, $\theta$, is set by the Rayleigh criterion $\theta \sim \lambda/D$, where $\lambda$ is the observing wavelength and $D$ is the diameter of the collecting aperture. As radio wavelengths are much longer than optical wavelengths, in order to obtain a comparable angular resolution, it is necessary to build much larger sized radio dishes compared to the size of optical mirrors. For example, in order to achieve a resolution of 0.5 arcsec at an observing wavelength of 21 cm, the radio antenna diameter must be $\sim 100$ km, which is impractical to build. As a comparison, the largest radio antenna in the world is the Five-hundred-metre Aperture Spherical Telescope (FAST), which has a dish diameter of 500 m and is located in Kedu Guizhou, China. Nevertheless, it is possible to achieve mas angular resolution by using arrays of antennas that are separated by a baseline of distance $b$. All together, these antennas act as a single-dish radio telescope of diameter $b$. The process of combining these separate antennas coherently is called interferometry. A property of this observing technique is the possibility to spread the antennas over continents, hence the name Very Long Baseline Interferometry (VLBI), to achieve mas and sub-mas angular resolution imaging at cm-wavelengths. Such high angular resolutions are needed to investigate gravitational lensing by sub-haloes at the $10^6 \, M_\odot$ level and test the smoothness of galaxy-scale dark matter haloes on small-scales (see Fig. 1.5).

However, interferometry also comes with a cost. In general, radio interferometers are composed of a number $N > 2$ antennas, and the possible number of independent baselines are $N(N - 1)/2$. Each pair of radio antennas produce voltage outputs, which are multiplied and integrated in time by a correlator. The cross-correlated output to an extended source with brightness distribution $I_\nu$ of an interferometer is a complex quantity on a two-dimensional grid (the Fourier- or $uv$-plane) called the visibility function, such that,

$$V_\nu(u,v) = \int \int I_\nu(x,y)e^{-2\pi i(ux+vy)} \, dx \, dy,$$

which corresponds to the Fourier transform of the surface brightness distribution.
\( I_\nu \). The visibility function is not known over the entire \( uv \)-plane, but it is sparsely sampled according to the location of the antennas, the integration time and the observing frequency. The \( uv \)-coverage determines the angular scales sampled, where the maximum extent sets the angular resolution, while the detection of the largest angular structures depends on the shortest \( uv \) spacings. Therefore, for an extremely sparse array, such as is typically the case for the interferometers used here, the resulting gaps in the \( uv \)-plane result in information being lost on some angular-scales. In other words, each baseline of an interferometer samples a single point in the Fourier-plane, and it is only by having many baselines, using aperture synthesis from the rotation of the Earth or a large frequency bandwidth can the Fourier-plane be better sampled. As giving an in-depth discussion of radio interferometry is beyond the scope of this chapter, we refer the interested reader to, for example, [Wilson et al. (2009)] for an advanced review.

### 1.3.2 Radio interferometers used in this thesis

Over the last decade, significant developments have been made to increase the sensitivity of the currently available radio interferometers. Here, the arrays used for carrying out the observations of the gravitational lenses investigated in this thesis are introduced. The mas angular-resolution and tens of \( \mu \)Jy beam\(^{-1} \) sensitivity that are now possible are the most valuable properties needed to achieve the science goals of the proceeding science chapters.

#### Very Large Array

The Karl G. Jansky Very Large Array (VLA) consists of 27 antennas with a diameter of 25 m located in New Mexico, USA (see Fig. 1.6). These antennas are placed on rail tracks with a Y-shape morphology, which allows the maximum baseline distances to be changed from 1 up to 36 km in four configurations from A (the largest) to D (the most compact). The recent upgrade of the high frequency Ka-band receiver (26.5 to 40 GHz) and the upgrade to 8 GHz of instantaneous bandwidth (3 bit sampling; 64 independent spectral windows) with full polarization capabilities has opened-up new observing possibilities with the VLA. These properties are particularly important when performing spectral line observations of the molecular gas distribution in high-redshift galaxies (e.g. Riechers et al., 2011). Thanks to the high angular resolution (that can be as high as \( \sim 60 \) mas at Ka-band in A-array configuration), and the excellent sensitivity (of the order of tens of \( \mu \)Jy beam\(^{-1} \)), it is possible to image both the extended and compact radio emission at low surface brightness. Moreover, the wide frequency range available enables such studies over a large range of redshifts. These capabilities
Figure 1.5: (Upper) Simulation of gravitational lensing arcs observed with VLBI, where there is a sub-halo of mass $10^6 \, M_\odot$ (left), $10^7 \, M_\odot$ (centre) and $10^8 \, M_\odot$ (right) in close projection. The strong distortion due to the presence of the sub-halo is clearly visible for the most massive sub-halo (McKean et al., 2015). (Lower) Number of dark matter sub-haloes as a function of the sub-halo mass (the dark matter sub-halo mass function). The black solid line represents the dark matter sub-halo mass function for cold dark matter. Two hydrodynamical simulations are shown in dashed lines. Warm dark matter models are shown in coloured solid lines, where each colour indicates a different dark matter particle mass (as indicated on the right panel). The vertical coloured bands define sub-halo mass regimes that strong gravitational lensing can probe with observations at different wavelengths. The representative telescopes for each band are listed on the right panel. Credits: G. Despali & S. Vegetti.
have been used extensively in Chapter 3 where sensitive spectral line observations of the CO (1–0) molecular gas are combined with complementary optical and NIR observations at a similar angular resolution to constrain the physical properties of two gravitationally lensed radio sources at redshifts 2 and 3.2.

**Very Long Baseline Array**

The Very Long Baseline Array (VLBA; see Fig. 1.7) is a homogeneous interferometer consisting of 10 antennas with a diameter of 25 m spread over the USA, which is operated by the Long Baseline Observatory. The observing frequency range is from 0.3 to 100 GHz in non-contiguous bands, and its longest baseline is between Mauna Kea (the most western antenna) and Saint Croix (the most eastern antenna) with a length of 8600 km. In the past, the data were processed on an hardware correlator that was built in the early 1990s, but they are now processed using a software correlator (DiFX) in Socorro using a current standard data rate of 2048 Mbits$^{-1}$. Also, the receivers at C-band were recently upgraded to cover a frequency range of 4 to 8 GHz, with two separate basebands that provide simultaneous multi-frequency imaging, and a typical baseline sensitivity of 1 mJy beam$^{-1}$ for 2 min integrations. In this thesis, these capabilities are used in Chapter 5 where the sensitive and wide-band system at C-band is used to characterize gravitational lens candidates found in the mJIVE–20 survey (Deller &
Middelberg (2014). When the on-source time is increased significantly, the VLBA can reach an imaging sensitivity of $\sim 10 \, \mu\text{Jy beam}^{-1}$, which allows faint radio emission to be robustly detected. In Chapter 4, sensitive VLBA observations ($\sim$10 hours on source) at 1.7 GHz are used to study the properties of the faint emission from lensed radio jets. The combination of $\mu\text{Jy beam}^{-1}$ sensitivity and mas angular resolution is crucial to recover the extended emission of the lensed radio jets on these scales, and provide a positional accuracy of the order of tens of $\mu\text{as}$, which is necessary to precisely constrain complex lens mass models.

**European VLBI Network and global VLBI**

The European VLBI Network (EVN) is a heterogeneous VLBI array that originally included only European radio telescopes, but now comprises antennas in Russia, China, South Africa and Puerto Rico (18 in total; see Fig. 1.8). Data can currently be recorded at 2048 Mbit s$^{-1}$ and each EVN session (three week period, every four months) covers four frequency bands (1.4, 5, 6 and 22 GHz). The EVN can observe simultaneously with the phased-VLA, the VLBA and the e-Multi-Element Remotely Linked Interferometer Network (e-MERLIN). In this case, the array is called *global VLBI*, and can achieve an angular resolution of between 2 and 10 mas, with a $\mu\text{Jy beam}^{-1}$ sensitivity at cm-wavelengths. In this
thesis, these properties of the global VLBI array at 1.7 GHz are used in Chapter 2 to search for astrometric anomalies in the lensed radio jets of MG J0751+2716, which shows a combination of compact and extended radio emission that produces extended gravitational arcs.

1.4 Thesis outline

In this thesis, the combination of gravitational lensing and high angular-resolution imaging is used to investigate both the mass distribution of foreground lensing galaxies and the properties of the highly magnified background radio-loud AGN. A brief summary of each science chapter is now given.

In Chapter 2 the mass density distribution of the massive lensing galaxy of MG J0751+2716 is determined by using new high quality VLBI imaging of extended gravitational arcs at mas-scale angular resolution. These data provide a wealth of observational constraints that are used to determine the inner (baryonic and dark matter) mass profile of a group galaxy and also investigate the smoothness of the dark matter distribution on mas-scales, which is sensitive to structures of $10^6$ $M_\odot$ within the lensing halo or along the line-of-sight. This chapter of the thesis demonstrates how very high angular resolution imaging can test models for dark matter.
Chapter 3 is devoted to a multi-wavelength study of two gravitationally lensed starburst/AGN composites at redshifts of around 2 and 3, a period when both the star-formation and AGN activity peaked. A sophisticated Bayesian grid-based lens modelling technique is applied to reconstruct the different emitting regions of these two targets, MG J0751+2716 and JVAS B1938+666, with the aim of better understanding AGN feedback and the role of merging in the star formation process. The reconstruction reveals two complex objects on 100 pc-scales, both with compact central stellar cores, extended radio jets and massive molecular gas reservoirs that extend out to 3 to 10 kpc-scales. Due to the magnifying power of gravitational lensing, all these source components are reconstructed at much higher angular resolution than that achievable for unlensed sources at similar redshifts.

The first detection of proper motion in a gravitationally lensed radio-source is presented in Chapter 4. Here, sensitive VLBI observations at two epochs separated by $\sim 15$ years of the lensed quasar MG B2016+112 are used to observe motion in the distant radio-components. We find evidence for a change in the position of the magnification $\sim 350$ lensed images of between 1 and 5 mas. Our parametric lens model reveals an intrinsic proper motion of 10 to 50 $\mu$as year$^{-1}$ in a complex source morphology. The properties of the proper motion, combined with archival multi-wavelength data, suggest that the motion can be due to two different scenarios. Either the motion is due to shocks along the jets of an FR I-type radio galaxy, or it is due to two radio-loud AGN orbiting around each other. If the latter interpretation is shown to be true, this would be the first detection of a dual AGN system made at high redshift. Further multi-wavelength observations are needed to build a more realistic mass model for the lens and to understand the geometry of this candidate dual AGN system.

In Chapter 5, the first survey for galaxy-scale gravitational lenses with wide-field VLBI is undertaken. Following a strategy similar to the JVAS and CLASS surveys, from an initial sample of 3,640 sources detected as part of the mJIVE-20 survey, 14 good lens candidates were selected. Among these candidates are two lenses previously found by CLASS, which demonstrates the efficiency and success of using wide-field VLBI to identify gravitational lenses. After following-up the remaining 12 lens candidates at higher frequencies with the VLBA, only one is not rejected on the basis of the surface brightness and spectral indices as a function of frequency, or from lens modelling. We find that the detection-rate of at least $2 \pm 1$ is in agreement with the expectations for the lensing statistics for compact radio sources, and we outline future surveys that can be done, based on the results from this pilot survey.

Finally, Chapter 6 provides a summary of this thesis and outlines future work.