Effects of lens motion and uneven magnification on image spectra

Indranil Banik⋆ and Hongsheng Zhao⋆

Scottish Universities Physics Alliance, University of St Andrews, North Haugh, St Andrews, Fife KY16 9SS, UK

ABSTRACT

Counter to intuition, the images of an extended galaxy lensed by a moving galaxy cluster should have slightly different spectra in any metric gravity theory. This is mainly for two reasons. One relies on the gravitational potential of a moving lens being time dependent (the moving cluster effect, MCE). The other is due to uneven magnification across the extended, rotating source (the differential magnification effect, DME). The time delay between the images can also cause their redshifts to differ because of cosmological expansion. This differential expansion effect is likely to be small. Using a simple model, we derive these effects from first principles. One application would be to the Bullet Cluster, whose large tangential velocity may be inconsistent with the Λ cold dark matter paradigm. This velocity can be estimated with complicated hydrodynamic models. Uncertainties with such models can be avoided using the MCE. We argue that the MCE should be observable with Atacama Large Millimetre Array. However, such measurements can be corrupted by the DME if typical spiral galaxies are used as sources. Fortunately, we find that if detailed spectral line profiles were available, then the DME and MCE could be distinguished. It might also be feasible to calculate how much the DME should affect the mean redshift of each image. Resolved observations of the source would be required to do this accurately. The DME is of order the source angular size divided by the Einstein radius times the redshift variation across the source. Thus, it mostly affects nearly edge-on spiral galaxies in certain orientations. This suggests that observers should reduce the DME by careful choice of target, a possibility we discuss in some detail.

Key words: gravitational lensing: strong – galaxies: clusters: individual: 1E 0657−56 – galaxies: kinematics and dynamics – dark matter.

1 INTRODUCTION

The standard Λ cold dark matter (ΛCDM) paradigm (Ostriker & Steinhardt 1995) still faces many challenges in reproducing galaxy scale observations (for a recent review, see Famaey & McGaugh 2012). Particularly problematic is the anisotropic distribution of satellites around Local Group galaxies, a question recently revisited in detail (Pawlowski et al. 2014). A different analysis focusing on Andromeda came to similar conclusions (Ibata et al. 2014). The relevant observations for the Milky Way (Pawlowski & Kroupa 2013) and Andromeda (Ibata et al. 2013) are difficult to repeat outside the Local Group because of the need to obtain 3D positions and velocities.

On a larger scale, Cai et al. (2014) found that the collision speed distribution of interacting galaxy clusters can be quite sensitive to the underlying law of gravitation. Thus, the high collision speed of the components of the Bullet Cluster 1E 0657−56 (Tucker, Tananbaum & Remillard 1995) has been argued in favour of

modified gravity (Katz et al. 2013). However, this speed is not directly measured as the collision is mostly in the plane of the sky. Instead, the speed is estimated using simulations of the shock generated in the gas by the collision. The separation of the dark matter (DM) and gas (Clowe et al. 2006) also plays an important role – there is less gas drag at lower speeds, so the separation is generally reduced.

A collision speed close to $3000\,\mathrm{km\,s^{-1}}$ is thought to be required to explain the observed properties of the Bullet Cluster (Mastropietro & Burkert 2008). For the inferred masses of the components (Clowe, Gonzalez & Markevitch 2004), this appears difficult to reconcile with ΛCDM (Thompson & Nagamine 2012). This work suggested that a cosmological simulation requires a comoving volume of $(4.48\,h^{-1}\,\mathrm{Gpc})^3$ to see an analogue to the Bullet Cluster.

The recent work of Lage & Farrar (2014) finds a similar collision speed but suggests a higher mass for the Bullet Cluster's components. While higher mass objects are likely to collide faster, such heavy clusters are rare in cosmological simulations. For example, their own unpublished work (Lage & Farrar 2015) based on the Horizon Run N-body simulation (Kim et al. 2009) showed how there were only seven cluster pairs with masses comparable to

their higher estimate for the Bullet Cluster mass. Because a larger volume of $(6.59\,h^{-1}\,\text{Gpc})^3$ was used in the simulation and $\left(\frac{6.59}{4.48}\right)^3 \approx 3$, this result is not very surprising in light of previous works.

However, some recent unpublished studies have raised the probability estimate of observing a galaxy cluster merger with properties comparable to the Bullet Cluster. Bouillot et al. (2014) used a larger box size of $(21\,h^{-1}\,\text{Gpc})^3$, using the Dark Energy Universe Simulation-Full Universe Run (DEUS-FUR) simulation. Thompson, Davé & Nagamine (2014) took issue with the friends-of-friends algorithm (Davis et al. 1985) long used to search outputs of *N*-body simulations for analogues to the Bullet Cluster. After switching to the recently developed ROCKSTAR algorithm (Behroozi, Wechsler & Wu 2013), the rate of occurrence of analogues to the Bullet Cluster increased by a factor of ∼100. Despite this, Thompson et al. (2014) quoted the probability of a collision as fast as the observed one as only one in 2170, which is still fairly small.

Moreover, a few other massive colliding clusters with high infall velocities have been discovered in the last few years (Gómez et al. 2012; Menanteau et al. 2012; Molnar et al. 2013b). The El Gordo Cluster (ACT-CL J0102−4915) may be particularly problematic due to its combination of high redshift ($z = 0.87$; Menanteau et al. 2012), high mass (Jee et al. 2014) and high inferred collision speed (Molnar & Broadhurst 2015).

A detailed analysis of how likely it is that observers would have seen interacting clusters with the observed properties is still lacking. One would need to account for incomplete sky coverage and perhaps faster collisions being easier to discover due to greater shock heating of the gas. A key input to any such analysis must be the collision speeds of the components. This work focuses on measuring cluster motions more accurately.

Molnar et al. (2013b) argue that inferring collision speeds from observations of the shock can be non-trivial just due to projection effects, let alone other complexities of baryonic physics. To see if there is any tension with the ΛCDM model, the collision speeds should be determined in a more direct way. Ultimately, we would like to determine the proper motions of colliding clusters.

Although not feasible by traditional methods, such motions may be inferred using the moving cluster effect (MCE; Birkinshaw & Gull 1983). As derived later, this effect relies on the gravitational potential of an object being time dependent due to its motion. Consequently, if a source behind the object were multiply imaged, the images would have slightly different redshifts. Moreover, as the DM generally outweighs the gas on cluster scales (Blaksley & Bonamente 2010), the MCE is mostly sensitive to the motion of the DM. This is simpler to model than the gas, making the results easier to compare with simulations.

The MCE would likely be around 1 km s^{-1} for the Bullet Cluster (Molnar et al. 2013a). The effect may be searched for using cosmic microwave background (CMB) photons (e.g. Cai et al. 2010). However, as noted by those authors, temperature anisotropies in the CMB make it difficult to spot such a signal around an individual object. Thus, we focus instead on using a multiply imaged background galaxy as the source. Spectral features in this galaxy could be used to determine the redshifts of its multiple images.

We consider the feasibility of obtaining measurements of the required accuracy using Atacama Large Millimetre Array (ALMA) in Section 7. Measuring this effect seems to be within our reach. One might instead conduct the observations in the visible/near-infrared (IR) with large spectroscopic instruments such as the Thirty Meter Telescope (TMT) and European-Extremely Large Telescope (E-ELT).

Therefore, it is important to consider other effects that might also cause the redshifts of double images to differ. Perhaps the most important such mechanism is what we term the differential magnification effect (DME). This depends on details of the source. If this is a rotating disc galaxy not viewed face-on, then different parts of the source have different radial velocities and hence redshifts.

The lens magnifies the source non-uniformly. The exact way in which this occurs is different for each image. Consequently, the intensity-weighted mean redshift of the images is usually different.

If one could perform integral field spectroscopy of the source galaxy accurate to ∼1 km s^{-1}, then one would simply need to compare the redshift of the same part of the galaxy between the two images. By focusing on a small part of the galaxy, the DME would be greatly reduced. However, this will be a challenging observational goal. The high spectral accuracy demanded by MCE measurements means the source will likely be spatially unresolved in the near future.

Assuming this to be the case, we determine the order of magnitude of the DME for a typical disc galaxy. We find that it may well be significant. Thus, we explore exactly how it affects the profiles of individual spectral lines. The effect is quite different to the MCE, which simply shifts each line. This may provide a way to correct for the DME and also to verify that a redshift difference is indeed caused by the MCE. Without spectra detailed enough to see such small differences between line profiles, it might still be possible to calculate the DME, though the determination would be less secure.

The additional complications and uncertainty introduced by trying to correct for the DME necessitate a discussion on how it may be reduced. Aside from the obvious steps of using ellipticals/face-on spirals and smaller – likely slower rotating – galaxies, an important factor to consider is how much the magnification varies across the source. The larger the variation, the larger the DME.

For this reason, an edge-on fast-rotating spiral galaxy might still be a good target if it is oriented so the magnification is nearly constant across the image. At the other extreme, the magnification varies rapidly near a caustic. Therefore, caustic images are likely to be strongly affected by the DME (Molnar et al. 2013a).

We emphasize the need to model both the redshift structure of the source and the deflection map of the lens when trying to use precise lensed image redshifts to determine the tangential motion of the lens.

2 THE MOVING CLUSTER EFFECT IN STANDARD AND NON-STANDARD GRAVITY

2.1 The lensing geometry

Fig. 1 illustrates the basic geometry that will be considered here. Because we are mostly concerned with angles on the sky, the distances relevant to us are the angular diameter distances to the lens and source (D_l and D_s). Also important is the angular diameter distance to the source galaxy as perceived by an astronomer at the lens, measured at the epoch that the lens is currently observed at (z_l). This last distance we denote D_{ls}.

2.2 Including lens motion

Suppose that the lens moves transversely to $\boldsymbol{OL}$. Thus, one of the light paths gets 'stretched' while the other gets 'squeezed', leading to a redshift difference between the images. To calculate this effect, we make the thin lens or triangle approximation whereby each photon trajectory is treated as two straight lines. In this case, the
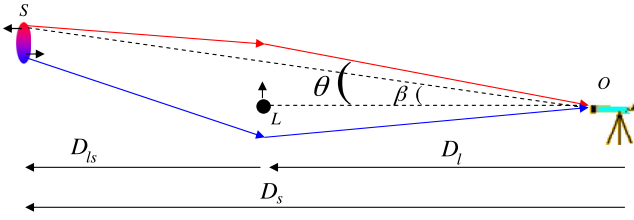
Figure 1. The lensing geometry is depicted here. Upper photon trajectory = primary image (same side as unlensed source), lower trajectory = secondary image. Relevant distances are indicated at bottom. The lens is treated as a point mass moving transversely to the viewing direction. The source is an extended disc galaxy. There is a redshift gradient across it due to rotation.

light arrival time surface is generally given by (Kovner 1990)

$$cT(\boldsymbol{\theta}) = \text{constant} + \frac{D_l D_s}{2 D_{ls}} (\boldsymbol{\theta} - \boldsymbol{\beta})^2 - \Psi(D_l \boldsymbol{\theta} - \boldsymbol{x}_1). \tag{1}$$

This consists of a geometric part and a relativistic part due to the lensing potential Ψ. We assume that Ψ depends only on position relative to the lens, which is located at $\boldsymbol{x}_1$ relative to some reference point in the lens plane.

The path length $cT(\boldsymbol{\theta})$ can be thought of as a function of the lens plane position $\boldsymbol{\theta}$ hit by a ray from the source. The actual rays for the images are at the extrema of this function (Fermat's principle).

We briefly describe how the geometric part of equation (1) is derived. Each section of a hypothetical undeviated photon trajectory can be mapped on to a part of the actual trajectory. The latter is slightly longer as there is an extra factor of the secant of the angle between them. This is expanded at second order, as the angle is small. The angle is $(\theta - \beta)$ for the stretch LO while for OS, it is $(\theta - \beta) \frac{D_l}{D_{ls}} \frac{1+z_l}{1+z_s}$. The last factor arises because photons emitted in the same direction gradually get further apart due to cosmic expansion. Thus, photons emitted in different directions end up more widely separated than in a static universe.

Equation (1) follows most naturally if using comoving distances, which can be added simply. This fact leads to the very useful relation between angular diameter distances:

$$(1 + z_l) D_l + (1 + z_s) D_{ls} = (1 + z_s) D_s. \tag{2}$$

If the lens moves in the transverse direction, then $\boldsymbol{x}_1$ changes and so the lensing potential Ψ changes at every point in the lens plane. Thus, transverse motion of the lens would cause the arrival time to change according to

$$c\dot{T} = \dot{\boldsymbol{x}}_1 \cdot \nabla \Psi \equiv -\boldsymbol{v}_t \cdot \boldsymbol{\alpha}. \tag{3}$$

$\boldsymbol{\alpha} \equiv -\nabla \Psi$ is the unreduced (true) deflection angle and $\boldsymbol{v}_t \equiv \dot{\boldsymbol{x}}_1$ is the transverse velocity of the lens (note time here refers to that measured by a clock at the lens).

The rate of increase of the path length $c\dot{T}$ is equivalent to a shift of the intrinsic spectrum of the source. While the source's intrinsic spectrum cannot be directly measured, the relative spectra of the images in a multiple image system can. Images 1 and 2 would have a relative redshift velocity:

$$\delta V_r \equiv V_1 - V_2 = -\boldsymbol{v}_t \cdot (\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_2), \tag{4}$$

where $\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_2$ is the relative deflection angle between the images. The observed angular separation between the images is $\frac{D_{ls}}{D_s}$ times as much. This allows the MCE to be calculated without knowing what the deflection angles are, as long as one is sure of the identification of the double image and the distances to the lens and source.

So far, we have used time measured by a clock at the lens. Using one on the Earth instead would introduce a factor of $(1 + z_l)$ to the time delay. Putting it in, we should think of $\boldsymbol{v}_t$ as the transverse peculiar velocity in comoving coordinates. This is the comoving lens distance times the relative proper motion of the lens with respect to that of the source.[1] This takes account of transverse motions of the observer and the source. In Section 3.2, we show that such motions affect image redshifts much less than motion of the lens.

For multiple lens planes, one would simply add the redshift differences due to each plane.

The above derivation is a property of metric theories of gravity. Hence, it is independent of details of the theory, something we now show. Consider a static Universe with no observer–source relative motion. Use a reference frame moving with the lens, so the source and observer both appear to move at $-\dot{\boldsymbol{x}}_1$. In general, there is a Doppler shift for a photon emitted by the source as perceived at the lens. A similar effect arises between lens and observer. The shifts cancel if the photon is not deflected by the lens (if the source emits a photon 'backward', then the observer 'ploughs into' the photon).

However, if the photon is deflected, then there is a net frequency shift between source and observer. This shift is different for another photon which gets deflected by a different amount. Therefore, the difference in deflection angles determines the redshift difference between the photons. If observers could be sure the photons initially had the same energy, then the redshift difference could be directly measured, thus constraining $\dot{\boldsymbol{x}}_1$.

For small deflections by a non-relativistic lens, the result in an expanding Universe is exactly the same as in a static one, once all the angles have been properly accounted for.

2.3 Point mass lenses

The unlensed source and images all lie along a line, so we only consider positions along this line. For a point mass lens, one can generically say that a ray of light with impact parameter $D_l\theta$ is deflected by

$$\alpha = -\frac{4G\widetilde{M}}{c^2 D_l \theta}, \tag{5}$$

where $\widetilde{M}$ is the 'equivalent lensing mass' in alternative gravity. Different gravity theories require different amount of real mass M to produce the observed equivalent mass $\widetilde{M}$, the Einstein radius θ_E or the deflection angles $\alpha_{1,2}$. In general, $\widetilde{M}$ depends on position in modified gravity theories, even with a point mass. For simplicity, we neglect this.

Combining equations (4) and (5), we get that

$$V_1 - V_2 = \frac{4G\widetilde{M}v_t}{c^2 D_l} \left(\frac{1}{\theta_1} - \frac{1}{\theta_2} \right). \tag{6}$$

Note that the signs of θ_1 and θ_2 are opposite because the images are on opposite sides of the lens (equation 10).

v_t is the component of the transverse velocity $\boldsymbol{v}_t$ projected along the line connecting images 1 and 2. If the lens proper motion is orthogonal to the image separation, then v_t would be zero.

---

[1] Physically, it would be the peculiar velocity of the lens in the direction orthogonal to our line of sight, in the absence of peculiar motions of either observer or source.

Noting that deflecting a photon at the lens only affects part of its trajectory, we get the classical lens equation

$$\beta = \theta - \frac{D_{ls}}{D_s}\frac{4G\widetilde{M}}{c^2 D_l \theta} \tag{7}$$

$$\equiv \theta - \frac{\theta_E{}^2}{\theta}, \quad \text{where } \theta_E \equiv \sqrt{\frac{4G\widetilde{M}}{c^2}\frac{D_{ls}}{D_l D_s}}. \tag{8}$$

The Einstein radius θ_E defines a typical angular scale for the problem. It will be convenient to use this to normalize all relevant angles. Thus, we let

$$u \equiv \frac{\beta}{\theta_E} \quad \text{and} \quad y \equiv \frac{\theta}{\theta_E}. \tag{9}$$

Images are formed where

$$y = \frac{1}{2}\left(u \pm \sqrt{u^2 + 4}\right). \tag{10}$$

For later use, we note that the magnification of a small part of the source located at an unlensed angular position of $\beta \equiv u\theta_E$ relative to the lens is given by

$$A = \left|\frac{\theta}{\beta}\frac{\partial\theta}{\partial\beta}\right| \tag{11}$$

$$= \frac{1}{2}\left(\frac{u^2 + 2}{u\sqrt{u^2+4}} \pm 1\right). \tag{12}$$

The result follows from the surface brightness of an unlensed source and a lensed one being equal (Liouville's theorem). Thus, A is the Jacobian of the mapping between where objects appear on the sky and where they would without a lens. The modulus signs are needed because otherwise $A < 0$ for the secondary image (indicating that it is inverted).

The secondary image becomes very faint if $u \gg 1$ (in this case, $A \sim \frac{2}{u^4}$). It is difficult to find a source with $u \ll 1$ as this corresponds to a very small part of the sky. Thus, any source used for MCE measurements will very likely have $u \sim 1$. We assume this is the case.

3 CAUSES OF DIFFERENCES BETWEEN MULTIPLE IMAGE REDSHIFTS

Before deriving the DME, we briefly consider a few factors that can affect redshift differences between double images of a source strongly lensed by the Bullet Cluster.

3.1 Lens motion

The MCE is maximal for a source displaced from the lens in the direction of its proper motion. This direction can be determined from images of the shock fairly easily. Moreover, observers should select targets to maximize the MCE. Thus, we assume the double images are indeed separated along the direction of motion of the lens. Otherwise, the MCE is reduced by the cosine of the angle between them (equation 4).

With these assumptions, we combine equations (6) and (10) to get that the difference in redshift velocity between the two images is

$$\Delta\overline{v_r}|_{MCE} = \frac{2v_t\sqrt{GM\left(u^2+4\right)}}{c}\sqrt{\frac{D_s}{D_{ls}D_l}}. \tag{13}$$

Table 1. Input parameters used for Fig. 4. The source galaxy is assumed positioned so as to maximize the MCE (i.e. it is separated from the lens on the sky along the direction of motion of the lens, which is clear from images). The lens mass should roughly correspond to the subcluster in the Bullet. A flat ΛCDM cosmology is adopted (Planck Collaboration XVI 2014).

Parameter	Meaning	Value
H_0	Present Hubble constant	67.3 km s^{-1} Mpc^{-1}
Ω_m	Present matter density	0.315
D_l	(Angular diameter) Distance to lens at $z_l = 0.296$	0.945 Gpc
D_s	Distance to source at $z_s = 1.7$	1.795 Gpc
D_{ls}	Distance to source from lens position in space–time	1.341 Gpc
M	Mass of lens	1.2×10^{14} M$_\odot$
r_d	Scale length of source galaxy	3.068 kpc
v_t	Tangential velocity of lens	3000 km s^{-1}
v_f	Flat line level of source galaxy rotation curve	100 km s^{-1}
$\sin i \cos \gamma$	See Fig. 2 and equation (21)	$\frac{1}{2}$

Using parameters appropriate to the Bullet Cluster (Table 1), the effect is around 1 km s^{-1}.

3.2 Source and observer motion

The peculiar velocity (with respect to the CMB) of the Sun is well known (369 km s^{-1}; Planck Collaboration XXVII 2014) and could be included in a more careful analysis. We do not include it as we only seek a rough idea of the magnitude of the MCE. This will not be substantially affected by observer motion as this is much slower than that of the lens ($\sim$3000 km s^{-1}).

More problematic may be the unknown peculiar velocity of the source. Treating the Local Group peculiar velocity ($\sim$630 km s^{-1}) as typical for galaxies, the Bullet Cluster transverse motion likely greatly exceeds the source's peculiar velocity. In this case, only the component of this velocity parallel to the lens transverse motion has much effect on image redshifts, leading to a factor of $\frac{1}{2}$ on average.[2]

Another factor of $\frac{1+z_l}{1+z_s}\frac{D_l}{D_s} \approx \frac{1}{4}$ arises due to the geometry of the situation and cosmic expansion. Moreover, typical peculiar velocities were smaller long ago. Supposing they were $\frac{1}{2}$ as much as today at $z_s = 1.5$, we see that source motion cannot affect the inferred lens velocity much more than $\sim$50 km s^{-1}. This effect can be reduced by observing more than one double image pair. However, we consider the accuracy with even just one well-observed pair sufficient.

Thus, we ignore any motion other than that of the lens. We note that it might be good to avoid source galaxies which are interacting, as their peculiar velocities might be higher.

3.3 Cosmological expansion

A redshift difference between the images can also arise because the time of flight of photons emitted by the source is different depending on which path they took to get to the Earth. As the photons for both images arrive simultaneously, the photons for one image must have been emitted earlier than for the other. Thus, in an expanding Universe, one of the images will have a higher redshift. Because of both a longer path length and a stronger gravitational

---

[2] For an angle between a fixed vector and another statistically isotropic one, $\langle|\cos\theta|\rangle = \frac{1}{2}$.

field along the path, this image is the secondary (on the opposite side of the lens as the unlensed source would appear – see Fig. 1). We term this the differential expansion effect (DEE).

A quick way to estimate the DEE is to assume that cosmological distances like D_L are usually $\sim \frac{c}{H}$. The extra path length $\sim D_L \theta^2$. The effect of the difference in gravitational time delays can crudely be approximated as equal to that due to different geometric path lengths.

The DEE expressed as a redshift $= H\Delta t \approx \theta^2$. Meanwhile, the MCE $\sim \frac{v}{c}\theta$. Assuming a velocity of 3000 km s^{-1} and an image separation of 20 arcsec, we see that the MCE is ~ 50 times larger than the DEE. Thus, we calculate the DEE more precisely.

We first consider just the difference in time of flight ('delay') due to different strengths of gravity along the two possible photon paths (Shapiro 1964). The relative Shapiro delay between the images is

$$\Delta t_l = \frac{4GM}{c^3} \ln\left(\frac{b_2}{b_1}\right), D_s, D_{ls} \gg b_{1,2}. \tag{14}$$

The impact parameters of the rays are $b_{1,2}$. This result is valid if b is much larger than the Schwarzschild radius of the lens, so the rays are only weakly deflected. The ray with smaller b is delayed more as it goes deeper into the lens' gravitational potential well. It also has a longer geometric path length (it forms the secondary image in Fig. 1).

Note this is the time delay at the lens. In reality, both photons must reach the Earth now, so the time of emission at the source must have been different. This requires an extra factor of the relative rates of a clock at the lens and at the source, $\frac{1+z_l}{1+z_s}$.

We then combined this Shapiro delay with the geometric path difference between the trajectories. Thus, the difference in time of emission required for photons traversing the two trajectories to reach the Earth simultaneously is

$$\Delta t_s$$
$$= \frac{1+z_l}{1+z_s}\frac{2GM}{c^3}\left[u\sqrt{u^2+4} + 2\ln\left(\frac{u^2 + u\sqrt{u^2+4}+2}{2}\right)\right]. \tag{15}$$

We used equation (10) to relate the source position u to the positions of its images. A correction for cosmological time dilation was also applied.

The change in redshift is given by the fractional difference in the scale factor of the Universe at the time of emission of the photons:

$$\Delta z = H(z_s)\Delta t_s. \tag{16}$$

Using realistic parameters (Table 1 and $u \approx 1$), the DEE ~ 1 m s^{-1}. In Section 3.1, we showed that the MCE is ~ 1000 times larger, allowing us to neglect the DEE.

Even with more accurate instruments, a very large number of double image pairs would need to be observed before the random noise from source peculiar motions was reduced below such a small level. Thus, the DEE will not be important in the foreseeable future. An exception might possibly arise if the source galaxy peculiar motion could be estimated based on properties of galaxies near it.

4 DERIVATION OF THE DIFFERENTIAL MAGNIFICATION EFFECT FOR UNRESOLVED IMAGES

The effects mentioned in Section 3 cause the frequencies of identical photons emitted in different directions to end up different when
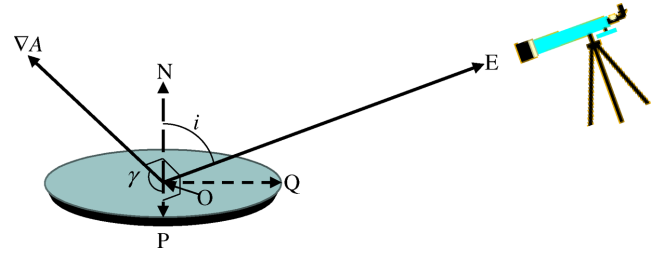

Figure 2. The observing geometry is shown here. The source galaxy has centre O and normal to its plane $\boldsymbol{ON}$. Earth is towards $\boldsymbol{OE}$, so the galaxy's inclination to the sky plane is i. $\boldsymbol{OQ}$ and $\boldsymbol{OP}$ are in the galaxy's plane and orthogonal to each other, with $\boldsymbol{OQ}$ as closely aligned with $\boldsymbol{OE}$ as possible. Thus, $\boldsymbol{OP}$ and $\boldsymbol{OE}$ are orthogonal. ∇A is directed within the source plane, so must also be orthogonal to $\boldsymbol{OE}$. ∇A is at an angle γ to $\boldsymbol{OP}$. The source is parametrized using cylindrical polar coordinates (r, ϕ), with centre O and initial direction ($\phi = 0$) along $\boldsymbol{OQ}$.

measured at the Earth. The DME does not do this. It is merely an observational artefact due to inability to simultaneously resolve the images and take highly accurate spectra of them. This causes parts of the source with different redshifts to get blended together in the spectra. The precise way in which this blending occurs is different between the images.

We assume the spectra are integrated over the entirety of each image. The source is modelled as a typical spiral galaxy with exponential surface density profile and a realistic rotation curve. The lens is treated as a point mass. The parameters considered (Table 1) are designed with the Bullet Cluster (Tucker et al. 1995) in mind.

The basic idea is that spatially unresolved spectra can determine the intensity-weighted mean redshift of each image. This may be affected by rotation of the source galaxy. The effect is not reliant on an expanding Universe, so it will be simplest to think of the Universe as static for the remainder of this section.

The mean redshift velocity of each image is given by

$$\overline{v_r} \equiv \frac{\int_{\text{Image}} A\Sigma v_r \, dS}{\int_{\text{Image}} A\Sigma \, dS}. \tag{17}$$

The integrals are over area elements of the source S. This is treated as a disc with surface density

$$\Sigma = \Sigma_0 \, e^{-\frac{r}{r_d}}. \tag{18}$$

The magnification A varies little over the source galaxy. This is because $\frac{r_d}{D_s} \ll \theta_E$ (see Table 1). Thus, a linear approximation to A is sufficient:

$$A \approx A_0 + \frac{\partial A}{\partial u}du \quad (A_0 \equiv A \text{ at centre of source}). \tag{19}$$

In our model, A varies linearly with position in the source plane, but only in the direction directly away from the (projection of the) lens. In the orthogonal direction, A is independent of position at first order (because u is, and A depends only on u).

The geometry of the source is shown in Fig. 2. The radial velocity of any part of it is

$$v_r = v_c(r) \sin\phi \sin i. \tag{20}$$

Only the component of ∇A along $\boldsymbol{OP}$ is important. To see why, suppose that ∇A was entirely along $\boldsymbol{OQ}$. Reflecting the galaxy about the line OP without altering ∇A should reverse the DME as this is equivalent to reversing ∇A. However, the radial velocity of every part of the galaxy remains unaltered after the reflection (as $\phi \to \pi - \phi$). Thus, the DME must also remain unaltered.

Noting that the component of ∇A along $\boldsymbol{ON}$ is irrelevant for the DME, we see that only the component along $\boldsymbol{OP}$ might be relevant. This component causes the approaching and receding halves of the galaxy to be magnified differently. It will be responsible for the DME. Thus, we assume the lens is located somewhere along the line $\boldsymbol{OP}$, making ∇A entirely along this direction. The result is then multiplied by $\cos \gamma$.

The magnitude of the DME is therefore $\propto \cos \gamma \sin i$. Assuming isotropy, all values of γ are equally likely. But values of i close to $\frac{\pi}{2}$ are more likely because there are more ways for two vectors to be orthogonal than to be aligned. This means the ratio between the average magnitude of the DME and the maximum it could be is given by the mean of $|\cos \gamma \sin i|$, with γ unweighted but a further $\sin i$ weighting over i.[3] Thus,

$$\langle | \sin i \cos \gamma | \rangle = \frac{\int_0^{\frac{\pi}{2}} \cos \gamma \, d\gamma}{\int_0^{\frac{\pi}{2}} d\gamma} \frac{\int_0^{\pi} \sin^2 i \, di}{\int_0^{\pi} \sin i \, di}$$

$$= \frac{1}{2}. \tag{21}$$

The angular separation between the lens and the unlensed source is given by

$$u = u_0 + \frac{\boldsymbol{r} . \widehat{\boldsymbol{OP}}}{D_s \theta_E} \quad (u_0 \equiv u \text{ at centre of source}). \tag{22}$$

The component of $\boldsymbol{r}$ (measured from the galaxy's centre) along $\boldsymbol{OP}$ is $r \sin \phi$. We have not kept careful track of signs because, for any conceivable orientation, the source galaxy could be rotating in the opposite sense, thereby reversing the DME. We explain which image has a lower redshift due to the DME later.

The difference in u between the centre of the source galaxy and any other point in it is given by

$$du = \frac{r \, \sin \phi}{D_s \theta_E}. \tag{23}$$

A_0 represents a constant magnification across the source. This does not contribute to the numerator in equation (17) because the radial velocity $v_r \propto \sin \phi$, while Σ is independent of ϕ due to axisymmetry. Thus, integrating over ϕ gives 0.[4] The DME arises when including the first-order correction to A.

The denominator in equation (17) is a normalization for each image (its total intensity).[5] Because the magnification is nearly constant across the source galaxy, we can approximate that $A = A_0$. The first-order correction to A would have a $\sin \phi$ dependence, which is irrelevant when integrated over all ϕ. This further justifies our approximation. Therefore, the denominator in equation (17) becomes

$$\int_{\text{Image}} A \Sigma \, dS = \Sigma_0 \pi r_d{}^2 \left(\frac{u^2 + 2}{u \sqrt{u^2 + 4}} \pm 1 \right). \tag{24}$$

To understand the sign of the DME, first note that regions closer to the lens are magnified more. In our approximation, the numerator in equation (17) is determined by $\frac{\partial A}{\partial u}$, which is the same for both

images (equation 12). Therefore, the image with the lower magnification (the secondary image) has a larger $|\overline{v_r}|$. Thus, if it was known which side of the rotating source was the approaching side, one could determine which image should have a higher mean redshift due solely to the DME.

Including the second-order dependence of A on sky position slightly alters the calculations done so far. Because a second-order term does not affect the approaching and receding halves of the source galaxy differently, the numerator in equation (17) is unaltered. But the denominator is affected, because the total intensity of each image may be altered by a second-order term. This means that our derivation of the DME has a fractional error which is second order in $\frac{r_d}{D_s \theta_E}$. We consider this acceptable and proceed to develop a model for the redshift structure of the source. This requires a rotation curve.

4.1 Model rotation curves

It will likely be difficult to directly observe the source galaxy rotation curve $v_c(r)$ as it is very far away. It is also difficult to precisely determine its surface density and thus predict the form of $v_c(r)$. Fortunately, we are considering a disc-integrated spectrum and so the exact shape of $v_c(r)$ will turn out not to be very important once the maximum level $v_{\max}$ is fixed.

To get a rough idea of $v_c(r)$, we take advantage of the tight empirical relation between the forces in rotating disc galaxies as required to sustain their rotation curves and those predicted by Newtonian gravity based on the visible (baryonic) mass (Famaey & McGaugh 2012, and references therein). This empirical formalism goes by the name of modified Newtonian dynamics (MOND; Milgrom 1983). Regardless of whether it is correct at a fundamental level, it does seem to provide a good empirical way of predicting rotation curves. Here, this is important because measuring the actual rotation curve of the source galaxy would be very challenging.

The particular empirical relation we adopt follows the work of Famaey & Binney (2005):[6]

$$\left(\frac{|\boldsymbol{g}|}{|\boldsymbol{g}| + a_0} \right) \boldsymbol{g} = \boldsymbol{g}_N, \tag{25}$$

where $\boldsymbol{g}$ is the true gravitational field strength while $\boldsymbol{g}_N$ is the prediction of Newtonian gravity based on the visible mass. a_0 is an acceleration scale ($\approx 1.2 \times 10^{-10}$ m s^{-2}) below which either gravity becomes non-Newtonian or DM must be considered in addition to the baryons. Thus, the magnitude of the gravitational field is given by

$$g = \frac{g_N}{2} + \sqrt{\left(\frac{g_N}{2} \right)^2 + g_N a_0}. \tag{26}$$

It is not worthwhile to accurately determine g_N for any particular mass distribution because the actual mass distribution in the source is uncertain. Thus, we approximated g_N using an analytic method. We assumed that, to determine g_N at a particular in-plane location, only material at smaller radii need be considered (we verified that the force from material at larger radii was very small). The Newtonian force at a distance r from the centre of a narrow ring of mass dM and radius x is

$$g_N \approx \frac{G \, dM}{r^2} + \frac{3G \, dM \, x^2}{4r^4} \quad (x < r, \text{ interior ring}). \tag{27}$$

---

[3] Edge-on galaxies are less likely to be detected due to dust obscuration. This makes low values of i – and thus a smaller DME – more likely, for a randomly selected multiple image.

[4] This is expected, as the DME does not arise if the image is uniformly magnified.

[5] What we perceive as the total intensity given D_s and z_s, but without correcting for magnification by the lens.

[6] This is the so-called 'simple μ-function' in MOND.

This is correct at second order in $\frac{x}{r}$. The total force at any point within the disc was found by decomposing the galaxy into a large number of rings with $dM = 2\pi x\,dx\,\Sigma(x)$. We then summed only the forces resulting from interior rings. Therefore,

$$g_N = \int_0^r \left(\frac{G\,dM}{r^2} + \frac{3G\,dM\,x^2}{4r^4} \right) \tag{28}$$

$$= 2\pi G \Sigma_0 f(\widetilde{r}), \quad \text{where} \tag{29}$$

$$\widetilde{r} \equiv \frac{r}{r_d} \quad \text{and} \tag{30}$$

$$f(\widetilde{r}) = \frac{1 - \frac{13}{4} e^{-\widetilde{r}}}{\widetilde{r}^2} - \frac{7 e^{-\widetilde{r}}}{4\widetilde{r}} + \frac{9 \left(1 - e^{-\widetilde{r}} - \widetilde{r} e^{-\widetilde{r}}\right)}{2\widetilde{r}^4}. \tag{31}$$

When obtaining the true value of g from g_N, the ratio $\frac{g_N}{a_0}$ is important. Therefore, we introduce a new variable, the dimensionless density k,

$$k \equiv \frac{G\Sigma_0}{a_0}. \tag{32}$$

Typical values for k are order 1. Using the empirical equation (26) to get g from g_N:

$$\frac{g}{a_0} = \pi k f(\widetilde{r}) + \sqrt{\left(\pi k f(\widetilde{r}) + 1\right)^2 - 1}. \tag{33}$$

To get the rotation curve, we equate g with the centripetal acceleration. Thus,

$$\frac{v_c^2}{\widetilde{r}\,r_d} = \left(\frac{g}{a_0}\right) a_0, \tag{34}$$

$$v_f = \sqrt[4]{2\pi k} \sqrt{r_d a_0}. \tag{35}$$

The rotation curve flat lines at the level $v_f = \sqrt[4]{GMa_0}$, where the total disc mass $M = 2\pi\Sigma_0 r_d^2$. The shape of the rotation curve is given by

$$\widetilde{v}_c(\widetilde{r}) \equiv \frac{v_c(\widetilde{r})}{v_f} = \frac{\sqrt{\pi k \widetilde{r} f(\widetilde{r}) + \widetilde{r}\sqrt{\left(\pi k f(\widetilde{r}) + 1\right)^2 - 1}}}{\sqrt[4]{2\pi k}}. \tag{36}$$

We show two such rotation curves in Fig. 3. Here, we also show how K affects the ratio between v_{max} and v_f.

4.2 The final result

Combining our results, we get that

$$|\overline{v_r}| = \frac{v_f r_d \sin i \cos\gamma}{D_s \theta_E}$$

$$\times \frac{\int_0^\infty \int_0^{2\pi} \overbrace{e^{-\widetilde{r}}}^{\propto\Sigma} \widetilde{v}_c(\widetilde{r})\widetilde{r}^2 \overbrace{\frac{4}{u^2(u^2+4)^{\frac{3}{2}}}}^{-\frac{\partial A}{\partial u}} \sin^2\phi\,d\phi\,d\widetilde{r}}{\pi \underbrace{\left(\frac{u^2+2}{u\sqrt{u^2+4}} \pm 1\right)}_{\propto A}}. \tag{37}$$

The magnification A changes by order 1 over an angular distance of θ_E, while the angular radius of the source galaxy $\sim \frac{r_d}{D_s}$. Thus, the DME as a fraction of the typical radial velocity of the source

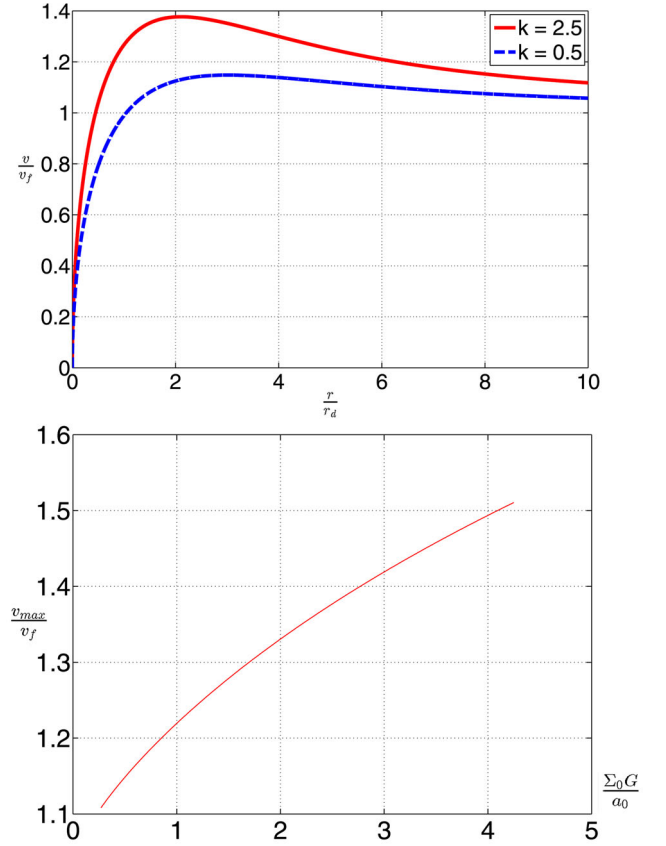**Figure 3.** Top: rotation curves resulting from equations (31) and (36), used in this work. $v_c(r)$ flat lines at v_f. The surface density $\Sigma = \Sigma_0 e^{-\frac{r}{r_d}}$. The parameter k controls the shape of the rotation curve ($k \equiv \frac{\Sigma_0 G}{a_0}$). Bottom: the ratio of maximum to flat line rotation speed as a function of central surface density.

is $\sim \frac{r_d}{D_s \theta_E}$. The galaxy's typical radial velocity is $v_f \sin i$. Another factor of $\cos(\gamma)$ is needed to account for the lensing geometry. As can be seen from equation (37), this provides a rough guide to the DME (as $u \sim 1$ for a realistic target).

An important quantity for the DME is the difference in $\frac{1}{A}\frac{\partial A}{\partial u}$ between the images:

$$\Delta\left(\frac{1}{A}\frac{\partial A}{\partial u}\right) = \frac{\partial A}{\partial u}\Delta\left(\frac{1}{A}\right) = \frac{4}{\sqrt{u^2+4}}. \tag{38}$$

In equation (37), the integration over ϕ yields π. The integral over r is not analytic. Thus, we define

$$I \equiv \int_0^\infty e^{-\widetilde{r}} \widetilde{v}_c(\widetilde{r})\,\widetilde{r}^2\,d\widetilde{r}. \tag{39}$$

Substituting for θ_E using equation (8), we get that

$$\Delta\overline{v_r}|_{DME} = \frac{v_f\, r_d\, \sin i\, \cos\gamma\, I\, c\, \sqrt{D_l}}{\sqrt{u^2+4}\, \sqrt{GMD_{ls}D_s}}. \tag{40}$$

The integral I depends somewhat on the central surface density in the sense that, for the same v_f, the DME is greater at higher k. However, the *maximum* rotation speed is very well correlated with the DME. In fact, the ratio $\frac{I}{v_{max}} = 1.89 \pm 0.02$ for $k = 0.1 \to 5$. As maximum rotation speeds are generally easier to determine than the flat line level, this makes correcting for the DME easier.

If the surface density declines sufficiently slowly with r, then the integral I might diverge. This is due to limited validity of a linear

approximation to A – a more careful treatment would be required. This might apply to rotating elliptical galaxies with $\rho \propto r^{-4}$. But even if $\Sigma \propto r^{-3}$, the divergence of I is fairly slow. Thus, although the integral would need a cut-off radius, its precise value would not affect the result much.

A linear approximation to A must break down if u changes by order 1. Thus, a logical cut-off might be the Einstein radius (at the source plane) or the distance between the source and the projected lens.

If a fibre-fed spectrograph was used or the field of view was otherwise restricted, then this may impose an obvious cut-off. For an inclined disc galaxy, a circular aperture would cover a non-circular region in the disc plane. One might need to take this into account.

We now decide on realistic parameters to gain a feel for the scale of the DME and MCE. Ideally, one would like to measure the motion of both components of the Bullet Cluster. However, we choose a mass corresponding roughly with the subcluster in the Bullet (Mastropietro & Burkert 2008). This is because the centre of mass likely has little peculiar velocity as there is little structure on such large scales. Thus, the lower mass component probably moves faster (with respect to the Hubble flow).

The MCE $\propto v\sqrt{M}$ (equation 13). Assuming also that $v \propto \frac{1}{M}$ for the components of the Bullet Cluster and that the value of u would be broadly similar whichever component is targeted, the MCE overall $\propto \frac{1}{\sqrt{M}}$. This makes it larger around the subcluster. Furthermore, using its motion to extrapolate the total collision velocity is much more reliable than using the motion of the main cluster, because the subcluster contributes most of the relative velocity.

A typical source galaxy orientation is chosen using equation (21). Although we used $v_f = 100$ km s^{-1}, it is around double this for our own Galaxy (e.g. McMillan 2011).

Using the parameter values in Table 1, we obtained the results shown in Fig. 4. The DME and MCE are comparable in magnitude.

Observing a similar source multiply imaged by the higher mass component instead does not reduce the relative importance of the DME. This is because the DME $\propto \frac{1}{\sqrt{M}}$ (equation 40), just like the MCE. In fact, this scaling highlights an additional problem: substructure in the lens (e.g. individual galaxies) with much less

mass than the entire cluster can enhance the DME. For example, an elliptical galaxy in the lens plane with $M = 10^{13}$ M$_\odot$ would cause a DME $\sim$3 times larger than the smooth cluster potential.

This problem could be mitigated to some extent by not selecting images which show indications of being lensed by small-scale structure (e.g. avoiding images appearing near a galaxy in the lens plane). We have implicitly assumed that such a selection has been done, such that a point mass model for the lens is appropriate. Even in this case, it might well be necessary to correct for the DME. This correction would be less relevant if targets could be selected for which the effect is small. We now consider how these things might be achieved.

5 CORRECTING FOR THE EFFECT

For spatially unresolved spectra, it is possible to calculate the DME by determining the parameters in equation (40). If radial velocities accurate to a few km s^{-1} are obtained for a galaxy, then determining $v_{max} \sin i$ to $\sim$10 km s^{-1} should be feasible using widths of spectral lines (see later).

r_d might be obtained from an image of the target, once distortion and magnification by the lens were corrected for. If the image were taken at more than one wavelength, it would suggest a value for k (which we do not need very accurately) as the colour can be used to estimate the baryonic M/L.

There is no need to determine the inclination as we are only interested in redshift gradients across the source. However, the orientation of the major axis of the image is important in determining the axis of rotation and thus the angle γ in Fig. 2.

To know the sense of rotation (i.e. which side of the source galaxy is the approaching side), we would need spectra of different parts of the source galaxy. Naturally, a disc-integrated spectrum would be insufficient for this purpose. However, one could make do with poorer spectral resolution.

The secondary image is inverted relative to the primary, providing an important consistency check if both images were used for such a determination. We strongly recommend doing this, because an error would lead to a 200 per cent error in the calculation of the DME. The chance of this is minimized with two determinations of the sense of rotation.

Finally, we also need ∇A, which must come from a lensing reconstruction.

5.1 Additional information from detailed spectral line profiles

It is often possible to extract more information from a spectral line than just the location of its centroid. The width of the line profile can be used to estimate e.g. $v_{max} \sin i$.

The MCE simply shifts the entire spectrum. The DME leads to a 'tilt' being introduced because one side of the galaxy is magnified more than the other. These effects are different. Therefore, detailed line profiles can tell us if the shift in the centroid of spectral lines is due to the MCE or the DME. This would avoid the need to determine parameters like the disc scale length and orientation. A detailed lensing reconstruction to determine ∇A would also be avoided.

We investigate how the DME and MCE affect line profiles of disc galaxies with rotation curves parametrized by equation (36). We assume the galaxy is viewed edge-on, so the radial velocity of any part of it is

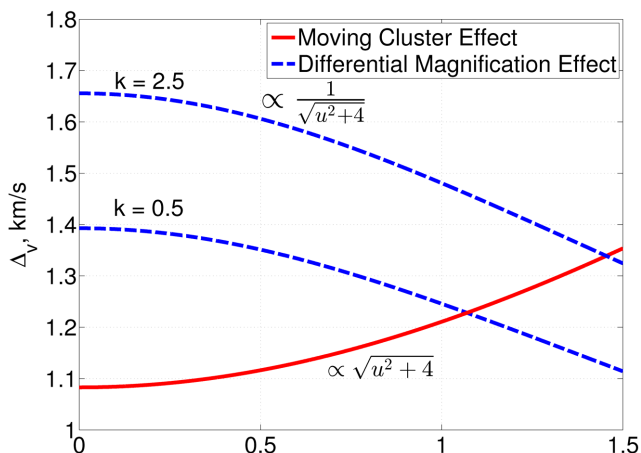$$v_r(r, \phi) = v_c(r) \sin(\phi). \tag{41}$$

Figure 4. The difference in redshift between double images of a typical background galaxy as a function of its position, due to the effects described in the text (equations 13 and 40). Parameter values used are listed in Table 1. The shape of the rotation curve (k) has a modest impact on the DME once its flat line level v_f is fixed. If instead v_{max} is held fixed, then the impact of k on the DME is very small.
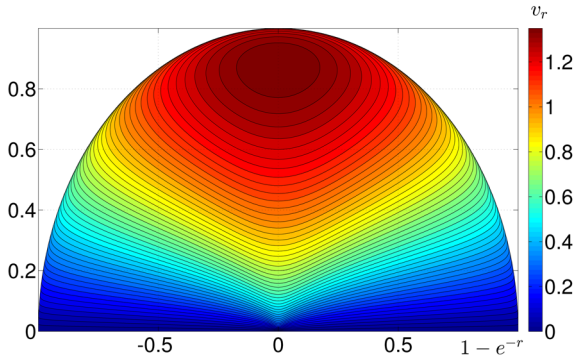
Figure 5. Radial velocity map of a disc galaxy viewed by an observer within its plane at large x, for the case $k = 2.5$. Radial velocities are antisymmetric about the x-axis. The radial coordinate is rescaled so all parts of the figure would be equally bright. The units are such that $r_d = 1$ and $v_f = 1$. Note the large region with v_r close to the maximum value. The result for $k = 0.5$ is very similar, although v_{max} is much closer to v_f.

The resulting radial velocity map is shown in Fig. 5. Only half of the galaxy is shown because v_r is antisymmetric about the viewing direction (the x-axis). v_r is symmetric about the y-axis, because $\sin\phi = \sin(\pi - \phi)$.

To determine the profile of a narrow spectral line, we divide the galaxy up into a large number of elements. We use cylindrical polars so v_r becomes separable. Thus, the rotation speed only needs to be calculated once at each r (for all ϕ). The radial velocity at the centre of each element is used to classify it among 200 bins in radial velocity.

Assuming constant mass-to-light ratio (M/L) for the baryons, the total intensity of the element multiplied by the magnification A is then assigned to the corresponding radial velocity bin. Because radial velocities and wavelengths are directly related, in this way one obtains a synthetic line profile.

Spectral lines have an intrinsic width and can be further broadened by random motions within the galaxy. To account for these effects and also for instrumental errors, we convolved our synthetic line profiles with Gaussians of various widths σ. The results are shown in Fig. 6. The sharp peaks at $v_r \approx \pm v_{max}\sin i$ give rise to the name of a double-horned profile.

These horns are caused by the rotation curve having a peak, leading to a small range in v_r corresponding to a large range in r. The greatest attained values of $|v_r|$ also correspond to large ranges in ϕ, because $\sin\phi$ is nearly independent of ϕ when $\phi \approx \pm\frac{\pi}{2}$.

Thus, a small range in v_r corresponds to a large region in the galaxy. Moreover, the peak rotation speed occurs at a radius close to that which maximizes the light emitted per unit radius ($r = r_d$). Fig. 5 shows the 'bull's eye' corresponding to the fairly large region with near-maximal $|v_r|$. This is responsible for the very pronounced horns in the line profile. They are somewhat less pronounced at high σ.

Although one might expect a feature corresponding to v_f (at least at low σ), this is absent. A quick look at Fig. 5 shows why $v_c(r) \approx v_f$ only for sufficiently large r. There is very little light from such regions, so a disc-integrated spectrum is hardly sensitive to them. In fact, due to the steep decline in surface brightness with r, most of the spectral intensity at $v_r = v_f$ actually comes from the rising part of the rotation curve (when $v_c \sin\phi = v_f$) rather than from the flat part. Thus, in the line profile, there is nothing special about v_f. This is not true for v_{max}.
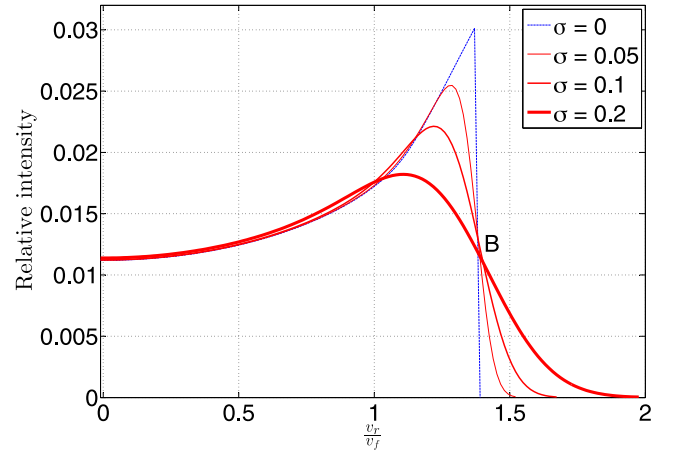

Figure 6. The synthetic line profile of an intrinsically narrow line in an unlensed galaxy with $k = 2.5$, viewed edge-on. The profile is symmetric about $v_r = 0$. Velocities are scaled to the flat line level v_f. The sharp drop in the line profile (blue) would probably get blurred (e.g. by random motions), so we convolved the profile with Gaussians of widths σ. The results are shown as red lines with thicknesses $\propto \sigma$. Notice how all four profiles pass close to the point marked B. The result for $k = 0.5$ is similar, if the profiles are scaled to have the same v_{max} rather than v_f.

Determining $v_{max}\sin i$ from a line profile is non-trivial as the horns move to lower $|v_r|$ as σ increases. Instead of using the horn positions, one could use the values of v_r where the intensity is a certain fraction of the intensity at the line centre ($v_r = 0$). If this fraction is chosen carefully, then one could simply scale the resulting v_r by a constant factor and accurately recover $v_{max}\sin i$ over a wide range in σ and k. To see why, note that spectra with different σ all pass close to the point marked B in Fig. 6.

Ultimately, it might be better to compare the observed line profile with a suite of synthetic profiles built for a range of k, σ and $v_{max}\sin i$. The initial guess for σ might come from considering the shape of the tail. If $v_{max}\sin i$ is accurately recovered, then the DME hardly depends on k.[7]

The horns are cause by a relatively small part of the galaxy but they greatly affect the mean radial velocity of its image. Thus, if the galaxy was not axisymmetric and e.g. had a dusty spiral arm obscuring light from this region, then the redshift measurement of each image may be biased. Partly for this reason, it may be a good idea to consider the rest of the line profile and not just the mean redshift (which is basically the same as considering just the horns).

We now consider how the DME and MCE affect the line profile. The mean redshift velocity of the line is raised by 1 per cent of v_{max} (1.08 per cent of v_f for $k = 0.5$ and 1.26 per cent for $k = 2.5$). We consider separately the cases where either the DME or the MCE is wholly responsible for this shift in line centroid. We also construct control line profiles like those in Fig. 6. These are obtained by setting

$$A = 1 \; \forall_{r,\phi}. \tag{42}$$

[7] Line profiles can also be used to find $v_f \sin i$ without detailed rotation curves. In this case, the value of k is important as a 'Σ correction' must be applied to get from v_{max} to v_f (e.g. bottom panel in Fig. 3). k does not affect the line profile much and so it would need to come from an image and photometry.
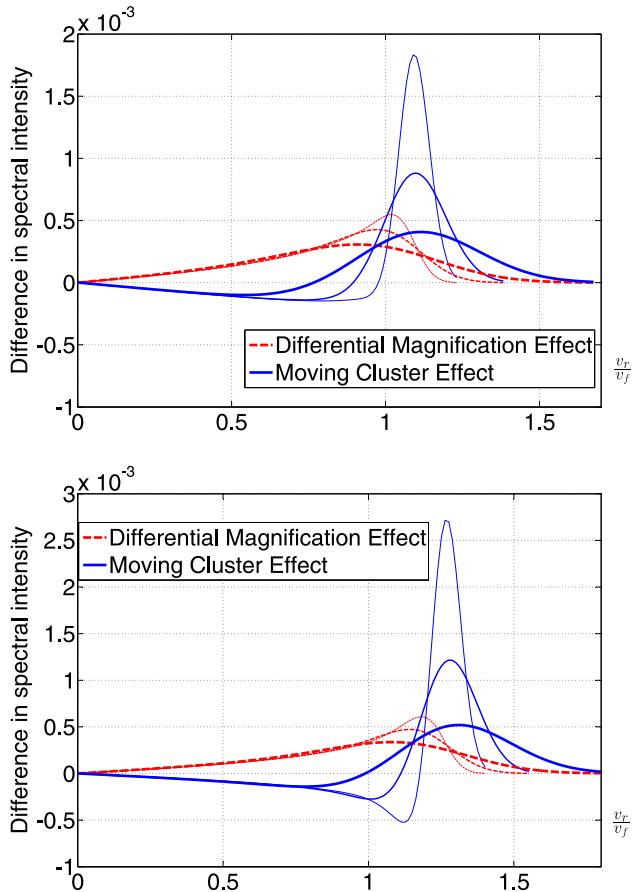
Figure 7. The residuals in the spectral profile due to the DME (equation 43) and the MCE (horizontal shift of profile) are shown here. These were obtained by subtracting a control line profile (equation 42). The patterns are antisymmetric about $v_{\mathrm{r}} = 0$. Results are for an edge-on galaxy with $k = 0.5$ (top) and $k = 2.5$ (bottom). Both effects change the mean redshift by 1 per cent of the maximum rotation speed, representing 1.08 per cent of v_{f} for $k = 0.5$ and 1.26 per cent for $k = 2.5$. The spectra were convolved with Gaussians of widths 0.05, 0.1 and 0.2 v_{f} (higher σ indicated by thicker line). The MCE cannot change the amplitudes of the horns. The DME makes one more pronounced and the other less.

In Fig. 7, we show the pattern of residuals (relative to the control) created by each effect. The total line intensity is kept the same for the comparison.

To obtain the corresponding observations, one would need to account for the images having different overall magnifications. Thus, the spectra would have to be rescaled. We assume this can be done perfectly (i.e. the photometry is very accurate).

The MCE corresponds to a horizontal shift in the spectrum relative to the control. This means the amplitudes of the horns are unaffected. *The pattern of residuals corresponds to the gradient in the spectrum.* Thus, the residuals are largest near the positions of the horns, but small precisely at them. The residuals are of opposite signs on either side of each horn because the gradient in the spectrum changes sign there.

For the DME, we set

$$A = 1 + n\widetilde{r}\sin\phi. \qquad (43)$$

Note that no assumptions are made about any of the factors controlling the amplitude of the DME, beyond it being a small effect relative to $v_{\mathrm{max}}\sin i$ (i.e. $n \ll 1$) and that we need not consider the

second-order dependence of A on position in the source plane. The purpose here is to illustrate how the DME affects the line profile, not how much (this is controlled by n). If the DME $\sim 0.01 v_{\mathrm{f}}\sin i$, then the residuals would be ~ 1 per cent of the line profile.

The image overall is not magnified for any (small) n. We adjust n until the line centroid shifts by the correct amount, to allow comparison with the MCE.

The DME causes one side of the galaxy to be magnified more than the other. Thus, the residuals due to it are of the same sign for each half of the galaxy (e.g. for $v_{\mathrm{r}} > 0$). There is no change in sign at the horns.

These correspond to material displaced from the centre of the galaxy along the direction $\boldsymbol{OP}$ in Fig. 2. As argued previously, we only need to consider the component of ∇A along this direction. Thus, the effects of differential magnification are substantial for the material corresponding to the horns (in so far as the DME affects the image at all). This is in contrast to the MCE, which hardly affects the line profile at these positions (because the gradient of the line profile there is 0).

For some v_{r}, the MCE leads to very large residuals if $\frac{\sigma}{v_{\mathrm{f}}\sin i}$ is low (Fig. 7). Thus, observing faster rotating galaxies might make it easier to distinguish between the DME and the MCE (as $\frac{\sigma}{v_{\mathrm{f}}\sin i}$ would likely be smaller). However, the DME would be larger and so it would have to be accounted for more accurately.

Detailed profiles of individual spectral lines may therefore help in determining the balance between the MCE and the DME in accounting for redshift differences between multiple images. In reality, a large number of spectral lines would probably need to be stacked. Even then, it seems likely that, in so far as redshift differences between the images are discernible, the cause of such differences can also be determined.

5.2 The second-order effect

For simplicity, we continue assuming the source is located along $\boldsymbol{OP}$. Thus, regions with high $|v_{\mathrm{r}}|$ are magnified more and regions with lower $|v_{\mathrm{r}}|$ are magnified less than the centre of the source due to the second-order dependence of A on position. To investigate what this means for spectral line profiles, we set A to depend quadratically on position along the direction $\boldsymbol{OP}$. This introduces a $\sin^2\phi$ dependence:

$$A = \frac{1 + n\widetilde{r}^2\sin^2\phi}{1 + 3n}. \qquad (44)$$

A quadratic dependence along the orthogonal direction would give a $\cos^2\phi$ dependence. Because $\cos^2\phi = 1 - \sin^2\phi$, a second derivative of A in either direction would affect the line profile in the same way (i.e. the residuals would have the same pattern, up to sign); once any overall magnification was corrected for.

When comparing the spectra, observers would first scale them to have equal intensities. Thus, we must avoid changing the total intensity. This leads to the factor of $1 + 3n$ in the denominator of equation (44).

The results are shown in Fig. 8. The effect is symmetric in v_{r}, so both horns are equally affected. These end up more pronounced in the secondary image than in the primary (in this example).

In reality, both the first- and second-order DME would be present for any given pair of images of the same object. Thus, the residuals would have both an antisymmetric and a symmetric part. However, the latter would likely be very small for cluster mass lenses (except for caustic images).
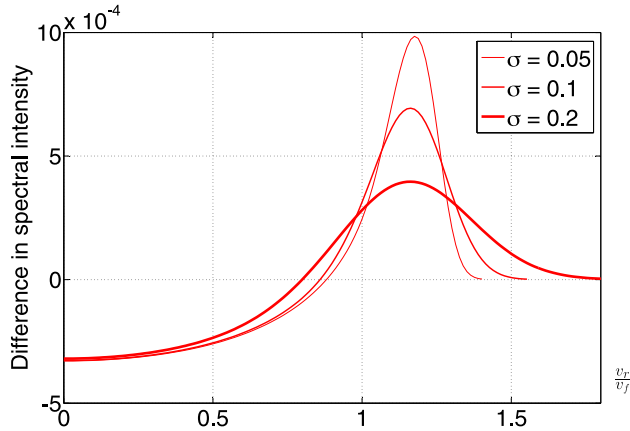
Figure 8. The pattern of residuals for the second-order DME (equation 44) and $k = 1$. A control profile obtained using equation (42) was subtracted and the result convolved with a Gaussian. The residuals are symmetric about $v_r = 0$, so both horns become more pronounced in this example.

5.2.1 Non-rotating sources

We briefly consider how the DME might affect a non-rotating pressure-supported galaxy, such as an elliptical. If a galaxy is symmetric such that $\rho(r) = \rho(-r)$ and $\sigma(r) = \sigma(-r)$, then at first order the DME does not affect an unresolved image at all. To see this, consider an inversion mapping $r \rightarrow -r$ while leaving ∇A unchanged. The situation is identical to reversing the direction of ∇A instead, so one expects the DME to act in exactly the opposite way on the spectrum. But the situation has not physically changed, so the DME must also remain unchanged.

This conclusion breaks down at second order. Suppose parts of the galaxy further from its centre are magnified more. Then, as the velocity dispersion generally decreases outwards, the derived velocity dispersion of the image will be reduced by the DME. The effect is larger for the fainter (secondary) image, which will thus appear to have a smaller velocity dispersion than the primary in this example.

This is likely to be more important for galaxy–galaxy lensing as θ_E is smaller, making du over the source larger. In this case, the DME might be useful to constrain the form of $\sigma(r)$ using a double image of a distant elliptical galaxy.

Alternatively, if the source galaxy was well understood, one might be able to constrain ∇A and thus have a better understanding of the lens. Doing both simultaneously would likely be very challenging and model dependent.

6 TARGETS WITH A SMALLER EFFECT

Fig. 4 shows that the DME may well need to be accounted for when using the redshifts of double images to determine the motion of the lens. However, doing this accurately may be difficult because of the cosmological distance to the source galaxy. Therefore, we suggest sources for which the DME should be smaller, allowing us to correct for it less accurately.

Some strategies outlined here involve selecting targets which are harder to observe, thereby making their spectra less accurate. It is up to observers to decide which targets best minimize the uncertainty introduced by the DME while still being feasible to obtain accurate spectra for. We also note that minimizing the uncertainty introduced by correcting for the DME is not necessarily equivalent

to minimizing the magnitude of the DME, because there may be sources for which the DME can be estimated more reliably.

6.1 QSOs

The DME $\propto rv$, where the source has typical size r and radial velocity spread v. For a given mass M, the Virial theorem gives $v \propto \frac{1}{\sqrt{r}}$. Thus, the DME $\propto \sqrt{r}$. For sources with a particular M, the DME would be reduced if the source were smaller, even though it would spin faster.

One obvious type of very small target visible over cosmological distances is a quasi-stellar object (QSO). If a doubly imaged QSO could be found lensed by the Bullet Cluster, it might make an excellent target.

QSO spectra can sometimes lack distinctive features which are required for precise redshift measurements. The Lyα forest might provide a solution, but only if the same feature appeared in both images. Because the rays of light corresponding to the images diverge from the source,[8] this is only feasible if the gas cloud causing the absorption feature was located fairly close to the QSO.

Another problem might be that the small size of QSOs makes their radiation time variable. Thus, the time delay between the images could make it difficult to compare their spectra. This might require observers to wait out the time delay, which would first have to be determined (though it could be estimated, perhaps using equation 15).

6.2 Smaller and fainter galaxies

The DME is proportional to both the rotation velocity and the size of the source galaxy. Brighter galaxies generally rotate faster (Tully & Fisher 1977), so targeting fainter galaxies might help. One advantage of this approach is that the number density of fainter galaxies is greater than for brighter ones (Schechter 1976). This makes it more likely that suitably oriented multiple images can be found.

However, it would be harder to obtain accurate spectra – and thus redshifts – for fainter targets. Given the high accuracy required in the redshift measurements and the cosmological distance to the source, this is perhaps not the best option at present.

6.3 Elliptical galaxies

Elliptical galaxies might make good targets as they usually rotate slower than spirals, if at all. They might be distinguished using colour or image shape (though one might need to correct for distortion by the lens). The surface brightness declines outwards much more gradually for ellipticals than for spirals, potentially providing another way of finding them.

Before conducting detailed observations, targets selected like this might be followed up to check if the spectral line profiles were double horned (characteristic of rotation along the line of sight). A good target should have a Gaussian-looking line profile, characteristic of a pressure-supported object.

However, even ellipticals can rotate, so the DME might not be eliminated by observing one. Also, most galaxies are not elliptical, so finding a bright doubly imaged one is somewhat dependent on luck. Nonetheless, we consider this the best option. This is partly because the work of Gonzalez et al. (2009) identified a multiply imaged galaxy which may be a good target for determining the MCE.

---

[8] By an angle $\frac{1+z_l}{1+z_s} \frac{D_l}{D_{ls}} (\theta_1 - \theta_2)$.

6.4 Galaxy orientation and viewing angle

Fig. 2 shows the geometry of the situation. The radial velocity of any part of the galaxy is scaled by $\sin i$, so a face-on spiral could not have a redshift gradient across it and thus would be unaffected by the DME.

Determining i requires an image of the source galaxy to determine its shape. One could imagine trying to select targets which look round. Even then, the target might be an elliptical galaxy with some rotation along the line of sight.

The direction of ∇A is also very important. In theory, we should seek situations where ∇A is orthogonal to the major axis of the image. In such situations, an edge-on disc galaxy would appear as a line on the sky aligned orthogonally to the direction towards the lens. With more complicated lenses, the galaxy would appear as a line on the sky orthogonal to ∇A, which hopefully could be estimated using a lensing reconstruction.

6.5 Galaxy position and caustics

As was already pointed out in Molnar et al. (2013a), the MCE is maximal for image separations aligned with the direction of the lens' proper motion. As the collision is nearly in the plane of the sky, the likely direction of this motion is plain to see and observers should target double images separated approximately along this direction.

In the simple lens model that we use, the DME $\propto \frac{1}{\sqrt{u^2 + 4}}$ while the MCE $\propto \sqrt{u^2 + 4}$. Thus, galaxies less closely aligned with the lens make better targets in terms of the systematic error of the DME. For such sources, the images are more widely separated.

Unfortunately, sources with larger u make worse targets under a number of other considerations. Both images – but especially the secondary – are fainter. This image also becomes very close to the lens, making it more likely to be obscured.

A lensing reconstruction could be used to suggest particular locations where the magnification is nearly constant. A galaxy with double images near such locations might make a good target for measuring the MCE. The difficulty with this is that such 'sweet spots' might be small and not have any observable galaxies in them. Also, a magnification map of sufficient accuracy might be difficult to obtain.

Regions where A varies rapidly with position enhance the DME. Caustics occur where the magnification of a small part of the source plane is infinity. This means that the magnification varies rapidly with position in the source plane, greatly increasing the DME. For this reason, it has been suggested to avoid caustic images (Molnar et al. 2013a).

However, it may be worthwhile to try correcting for the DME in caustic images because they are generally very bright, making for more accurate spectra. The correction would need to be done very accurately in this case, because the MCE might be much smaller than the DME. Because this is likely to lead to controversy surrounding the measurements, we also recommend avoiding caustic images unless the observational case is compelling.

6.6 Substructure within the source galaxy

If the source galaxy has e.g. a bright star-forming region which emits strongly in the ultraviolet (UV) while the rest of the galaxy does not, then another possibility arises. UV spectral lines would correspond to material in a small part of the galaxy. Consequently, different parts of it would have much the same radial velocity and the magnification across it would be more uniform than across the whole galaxy. This would reduce the DME for the UV lines.

Thus, in this example, the redshifts for the images should be calculated using only the UV lines. In practice, one would exploit the fact that a small part of the galaxy should have only a narrow range of redshifts. Thus, one might use only spectral lines which have a similar redshift. If the intrinsic linewidth was small, then the line should be very narrow as there would not be much rotational broadening.

Another possibility is using spectral lines that are more prominent in the bulge of the galaxy (if it has one). The bulge is mostly pressure supported with little rotation and is also much smaller than the whole galaxy. In this case, the spectral lines to use might be quite broad, but have a Gaussian line profile even if the galaxy is rotating (so most spectral lines have a double-horned profile).

For this technique to be of much benefit, the galaxy needs to be quite inhomogeneous in some way. It might be difficult to tell whether it is from an image. Also, the technique reduces the number of spectral lines that are used to calculate the redshift, making for less precise measurements. This makes it difficult to target fainter galaxies, perhaps forcing observers to choose between observing all of a fainter galaxy or (effectively) part of a brighter one.

Any decision to restrict which spectral lines are used to determine the MCE should be justified based on more detailed observations of nearby galaxies. This increases confidence that the decision does indeed effectively restrict the observations to a small part of the source.

Even if all usable spectral lines are used to measure the MCE, it is still likely that some lines are less affected by the DME than others. It may be important to allow for this in the analysis e.g. by grouping spectral lines based on their linewidth and shape and obtaining an inference on the MCE for each group.

Because of the uncertainties introduced by such procedures, we recommend reducing the DME by careful choice of target so that the exact method used to correct for it does not much affect the inferred lens velocity.

7 OBSERVATIONAL PROSPECTS

We now consider the technical feasibility of detecting the MCE with high-resolution spectroscopic measurements. The target we consider is presented in Gonzalez et al. (2009). This has a flux of $\sim$100 mJy at mm wavelengths. Because of dust in the source galaxy, it is best to do the observations at such wavelengths. For this purpose, we consider using ALMA. The Bullet Cluster rises to within $\sim$35° of the zenith at this site (minimum airmass $\approx$1.2).

The parameters we supplied are given in Table 2. To clarify the tension with ΛCDM, it would be necessary to constrain the collision

Table 2. Input parameters used for the ALMA exposure time calculator, available at https://almascience.eso.org/proposing/sensitivity-calculator. The dual polarization mode should be used as polarization is unimportant here. The angular resolution does not affect the result, which was 6.17 h.

Parameter	Value
Declination	-56°
Frequency	150 GHz
Bandwidth per polarization	100 m s^{-1}
Water vapour column density	Default: 5th octile (1.796 mm)
Number of antennas	50 $\times$ 12 m
rms sensitivity	1.5 mJy

speed to within $\sim 250\,\mathrm{km\,s^{-1}}$ (representing an 8 per cent accuracy if the actual speed is $3000\,\mathrm{km\,s^{-1}}$). This corresponds to determining the redshift difference between the images to $0.1\,\mathrm{km\,s^{-1}}$.

A flux accuracy of 1.5 mJy corresponds to ~ 2 per cent accuracy near the peak of the spectral energy distribution. The online calculator suggests that this level of precision can be attained in just over 6 h under typical weather conditions. Thus, we believe that a night with all 50 of the 12 m dishes might allow us to constrain, in this case, the proper motion of the main cluster.

In principle, the subcluster's motion can also be determined using the MCE. However, we could not find known multiple images with separation close to the east–west direction, the likely direction of the collision. Thus, suitable multiple images would first need to be found around the subcluster. This might be accomplished using a fairly deep exposure with the Atacama Pathfinder Experiment (APEX) or other telescopes. If suitable targets were found, they could be targeted for detailed spectroscopic follow-up.

The actual direction of motion of each component of the Bullet Cluster is not known for certain. Thus, observers should target multiple images separated in roughly orthogonal directions. Observing more than one object can also minimize systematics associated with details of the source and the lens.

Once suitable targets are found, we believe that a few nights of observations should be sufficient. The field of view might be large enough that multiple images of different sources can be observed in the same pointing, reducing the required telescope time further. Observing the images simultaneously can also reduce systematics associated with changing atmospheric conditions.

The main difficulty would be in achieving a very accurate calibration of the spectra. However, it is the *relative* redshift between multiple images that is critical for determining the MCE. *Absolute* redshifts are not needed very precisely.

8 CONCLUSIONS

The MCE may provide an essentially direct method to determine the tangential motion of high-z lensing clusters such as the Bullet Cluster, thereby clarifying the tension that appears to exist with ΛCDM (Molnar et al. 2013a). This requires a precise determination of redshift differences between multiple images of the same object.

We expect the MCE to cause multiple images created by the Bullet Cluster to have a redshift velocity difference of $\approx 1\,\mathrm{km\,s^{-1}}$. We find that, for multiple images of a realistic target, this level of accuracy should be feasible with a night on ALMA, using all 50 of the 12 m dishes. To determine the motions of both the main and the subcluster, multiple pointings may be required.

We considered the effect of the time delay between multiple images. In an expanding Universe, this causes them to have different redshifts (the DEE). However, the effect is second order in the deflection angle, whereas the MCE is first order. Thus, the DEE can be neglected compared with the MCE.

The DME arises when observing an object with a redshift gradient across it, most likely due to rotation. The precise way in which the magnification varies across the source is different for different images. This leads to them having different mean redshifts. Under plausible circumstances, the effect is large enough that it must be considered when trying to infer the lens motion (Fig. 4).

We consider various methods for determining how the DME affects image redshifts. All techniques require the profiles of spectral lines, if only to estimate the redshift gradient across the image based on the linewidth. If the line profile could be observed in more detail,

then one could exploit the fact that the DME and MCE affect the line profile in different ways.

Otherwise, the DME could be estimated by determining the parameters which control it (disc scale length and orientation, maximum line of sight rotation speed and sense of rotation, how magnification varies with position in the source plane for each image and, to a smaller extent, the source surface density).

The DME is smaller for some sources than for others. We discuss which types of source might reduce the DME in Section 6. We believe the best option is to use multiple images of an elliptical galaxy as these are likely to rotate slower than spirals, if at all. In particular, the triple image identified in Gonzalez et al. (2009) might be a good source to observe.

The DME is larger for lower mass lenses, making it more important for galaxy–galaxy lensing. Measuring peculiar velocities of galaxies using the MCE might thus be very challenging, especially as these are likely smaller than for the Bullet Cluster.

However, the DME might be easier to observe. This might give more information about the gravitational potential of the lensing galaxy and perhaps the redshift structure of the source. We speculate that the DME might provide a way to estimate the radial gradient in the velocity dispersion of a distant lensed elliptical galaxy.

REFERENCES

Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013, ApJ, 762, 109
Birkinshaw M., Gull S. F., 1983, Nature, 302, 315
Blaksley C., Bonamente M., 2010, New Astron., 15, 159
Bouillot V. R., Alimi J.-M., Corasaniti P.-S., Rasera Y., 2014, MNRAS, preprint (arXiv:1405.6679)
Cai Y.-C., Cole S., Jenkins A., Frenk C. S., 2010, MNRAS, 407, 201
Cai Y.-C., Li B., Cole S., Frenk C. S., Neyrinck M., 2014, MNRAS, 439, 2978
Clowe D., Gonzalez A., Markevitch M., 2004, ApJ, 604, 596
Clowe D., Bradač M., Gonzalez A. H., Markevitch M., Randall S. W., Jones C., Zaritsky D., 2006, ApJ, 648, L109
Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371
Famaey B., Binney J., 2005, MNRAS, 363, 603
Famaey B., McGaugh S. S., 2012, Living Rev. Relativ., 15, 10
Gómez P. L. et al., 2012, AJ, 144, 79
Gonzalez A. H., Clowe D., Bradač M., Zaritsky D., Jones C., Markevitch M., 2009, ApJ, 691, 525
Ibata R. A. et al., 2013, Nature, 493, 62
Ibata R. A., Ibata N. G., Lewis G. F., Martin N. F., Conn A., Elahi P., Arias V., Fernando N., 2014, ApJ, 784, L6
Jee M. J., Hughes J. P., Menanteau F., Sifón C., Mandelbaum R., Barrientos L. F., Infante L., Ng K. Y., 2014, ApJ, 785, 20
Katz H., McGaugh S., Teuben P., Angus G. W., 2013, ApJ, 772, 10
Kim J., Park C., Gott J. R., III, Dubinski J., 2009, ApJ, 701, 1547
Kovner I., 1990, ApJ, 351, 114
Lage C., Farrar G., 2014, ApJ, 787, 144
Lage C., Farrar G. R., 2015, J. Cosmol. Astropart. Phys., 2, 38
McMillan P. J., 2011, MNRAS, 414, 2446
Mastropietro C., Burkert A., 2008, MNRAS, 389, 967
Menanteau F. et al., 2012, ApJ, 748, 7
Milgrom M., 1983, ApJ, 270, 365
Molnar S. M., Broadhurst T., 2015, ApJ, 800, 37

Molnar S. M., Broadhurst T., Umetsu K., Zitrin A., Rephaeli Y., Shimon M., 2013a, ApJ, 774, 70

Molnar S. M., Chiu I.-N. T., Broadhurst T., Stadel J. G., 2013b, ApJ, 779, 63

Ostriker J. P., Steinhardt P. J., 1995, Nature, 377, 600

Pawlowski M. S., Kroupa P., 2013, MNRAS, 435, 2116

Pawlowski M. S. et al., 2014, MNRAS, 442, 2362

Planck Collaboration XVI, 2014, A&A, 571, A16

Planck Collaboration XXVII, 2014, A&A, 571, A27

Schechter P., 1976, ApJ, 203, 297

Shapiro I. I., 1964, Phys. Rev. Lett., 13, 789

Thompson R., Nagamine K., 2012, MNRAS, 419, 3560

Thompson R., Davé R., Nagamine K., 2014, preprint (arXiv:1410.7438)

Tucker W. H., Tananbaum H., Remillard R. A., 1995, ApJ, 444, 532

Tully R. B., Fisher J. R., 1977, A&A, 54, 661

This paper has been typeset from a TEX/LATEX file prepared by the author.