Probing dark matter substructure in the gravitational lens HE 0435−1223 with the WFC3 grism

A. M. Nierenberg,[1]⋆† T. Treu,[2]‡ G. Brammer,[3] A. H. G. Peter,[1,4,5] C. D. Fassnacht,[6] C. R. Keeton,[7] C. S. Kochanek,[1,4] K. B. Schmidt,[8] D. Sluse[9] and S. A. Wright[10]

[1]*Center for Cosmology and AstroParticle Physics, The Ohio State University, 191 West Woodruff Avenue, Columbus, OH 43204, USA*
[2]*UCLA Physics and Astronomy, 475 Portola Plaza, Los Angeles, CA 90095-1547, USA*
[3]*Space Telescope Science Institute, 3700 San Martin Dr, Baltimore, MD 21211, USA*
[4]*Department of Astronomy, The Ohio State University, 4055 McPherson Laboratory, 140 West 18th Avenue, Columbus, OH 43210, USA*
[5]*Department of Physics, The Ohio State University, 191 West Woodruff Avenue, Columbus, OH 43204, USA*
[6]*UC Davis Department of Physics, 1 Shields Ave., Davis, CA 95616, USA*
[7]*Department of Astrophysics and Astronomy, Rutgers, 36 Frelinghuysen Rd, Piscataway, NJ 08854, USA*
[8]*Leibniz-Institut für Astrophysik Potsdam, An der Sternwarte 16, D-14482 Potsdam, Germany*
[9]*Department of Astrophysics, Geophysics, Oceanography, University of Liège, Allée du 6 Aôut 17, B5C, B-4000 Sart Tilman, Belgium*
[10]*UCSD Center for Astrophysics and Space Science, 9500 Gilman Dr, La Jolla, CA 92093-0424, USA*

ABSTRACT

Strong gravitational lensing provides a powerful test of cold dark matter (CDM) as it enables the detection and mass measurement of low-mass haloes even if they do not contain baryons. Compact lensed sources such as active galactic nuclei (AGNs) are particularly sensitive to perturbing subhaloes, but their use as a test of CDM has been limited by the small number of systems which have significant radio emission that is extended enough to avoid significant lensing by stars in the plane of the lens galaxy, and red enough to be minimally affected by differential dust extinction. Narrow-line emission is a promising alternative as it is also extended and, unlike radio, detectable in virtually all optically selected AGN lenses. We present first results from a Wide Field Camera 3 (WFC3) grism narrow-line survey of lensed quasars, for the quadruply lensed AGN HE 0435−1223. Using a forward modelling pipeline that enables us to robustly account for spatial blending, we measure the [O III] 5007 Å flux ratios of the four images. We find that the [O III] fluxes and positions are well fit by a simple smooth mass model for the main lens. Our data rule out a $M_{600} > 10^8(10^{7.2})\,\mathrm{M_{\odot}}$ Navarro–Frenk–White perturber projected within ∼1.0 (0.1) arcsec of each of the lensed images, where M_{600} is the perturber mass within its central 600 pc. The non-detection is broadly consistent with the expectations of ΛCDM for a single system. The sensitivity achieved demonstrates that powerful limits on the nature of dark matter can be obtained with the analysis of ∼20 narrow-line lenses.

Key words: galaxies: dwarf – galaxies: haloes – quasars: individual: HE 0435−1223 – dark matter.

1 INTRODUCTION

A key prediction of cold dark matter (CDM) is that the halo mass function should follow an unbroken power law with $\mathrm{d}N/\mathrm{d}M \propto M^{-\alpha}$, where $\alpha = 1.9 \pm 0.1$ from cluster down to planet mass haloes (Diemand et al. 2008; Springel et al. 2008). CDM models of the

distribution and growth of structure match observations with remarkable success over an enormous range of distance and size scales (e.g. Planck Collaboration XVI 2014). On smaller scales, tests of CDM become more difficult owing to the uncertain physics of star formation in sub-Milky Way mass haloes ($M_{\mathrm{vir}} < 10^{12}\,\mathrm{M_{\odot}}$). A famous example of this is the 'missing satellite problem', so-called due to the fact that CDM simulations predict that thousands of subhaloes should be gravitationally bound to the Milky Way, while only ∼tens of *luminous* satellite galaxies have been observed (Klypin et al. 1999; Moore et al. 1999; Strigari et al. 2007; Weinberg et al. 2008; Drlica-Wagner et al. 2015). Extrapolations based on the

depth and area of the Sloan Digital Sky Survey (SDSS) and the Dark Energy Survey (DES) forecast that the Milky Way may host as many as $\sim$100 ultrafaint ($L < 10^3 \, L_\odot$) satellite galaxies (see e.g. Hargis, Willman & Peter 2014; Drlica-Wagner et al. 2015). However, measurements of the halo mass function at these low luminosities become extremely difficult. In particular, the stars of low-luminosity galaxies occupy only the inner $\sim$100 pc of their dark matter halo. This is a small fraction of the total halo virial radius, which, in the absence of the effects of tidal stripping, extends out to several kpc even for a very low-mass $10^7 \, M_\odot$ halo. One consequence of this is that satellite galaxies inhabit haloes with kinematically consistent masses over five orders of magnitude in luminosity (Strigari et al. 2008).

Simulations with varying implementations of semi-analytic and/or numerical baryonic feedback have demonstrated that it is possible to suppress star formation in dark matter subhaloes sufficiently to match the observed Milky Way satellite luminosity function with an underlying CDM halo mass function at low redshift (e.g. Thoul & Weinberg 1996; Gnedin 2000; Kaufmann et al. 2008; Weinberg et al. 2008; Macciò et al. 2010; Springel 2010; Guo et al. 2011; Zolotov et al. 2012; Brooks et al. 2013; Starkenburg et al. 2013; Lu et al. 2014; Menci et al. 2014; Wetzel et al. 2016). However, models that can match the luminosity function of satellites at redshift zero around the Milky Way are not always successful at reproducing the luminosity function around higher mass hosts, or at higher redshifts (e.g. Nierenberg et al. 2013, 2016).

Strong gravitational lensing provides a powerful test of CDM as it enables a measurement of the subhalo mass function without requiring stars or gas to detect subhaloes. In a strong gravitational lens, a background source is multiple imaged with the image positions and magnifications depending on the first and second derivatives of the gravitational potential, respectively. The image magnifications are particularly sensitive to low-mass perturbations, with a lower sensitivity limit determined by the source size. If the source is of the scale of microarcsec (μas) (such as a quasar accretion disc), then the image magnifications will be significantly affected by stars in the plane of the lens galaxy (also known as microlensing). In contrast, milliarcsec (mas) scale sources are not significantly lensed by stars, but are sensitive to the presence of perturbing subhaloes that have characteristic Einstein radii of $\sim$mas, corresponding to typical masses $10^4 < M/M_\odot < 10^9$ given typical lens configurations.

Traditionally, radio-loud quasar sources have been used to detect subhaloes, as they are extended enough (tens to hundreds of pc; e.g. Jackson et al. 2015) to avoid microlensing, and also are not affected by differential dust extinction in the plane of the lens galaxy. Dalal & Kochanek (2002) used the lensed magnification of six radio-loud quasar lenses and PG 1115+08 (Weymann et al. 1980) to estimate the average fraction of mass in substructure relative to the mass in the smooth halo component ($f_{\rm sub}$) around these lenses, finding this fraction to be broadly consistent with predictions from CDM although with large uncertainties. This was an important proof of method. Progress requires a larger sample of gravitational lenses which enables a measurement of not only the fraction of mass in substructure but also the slope of the subhalo mass function. Furthermore, Xu et al. (2015), Hsueh et al. (2016) and Gilman et al. (2017) have demonstrated that systematic uncertainties may occur in flux ratio measurements if the deflector is not accurately modelled due to insufficiently deep optical imaging, as may be the case for studies that rely entirely on radio imaging.

There are several paths forward for increasing the sample of lenses which can be used to measure the subhalo mass function. Gravitational imaging, for instance, can be used to detect subhaloes

perturbing the lensed positions of background galaxies. This method currently has a limiting mass sensitivity of $\sim$$10^8 \, M_\odot$ (Vegetti et al. 2012, 2014; Hezaveh et al. 2016; Birrer, Amara & Refregier 2017). The sensitivity limit is determined in part by imaging spatial resolution, which enables the measurement of the small astrometric perturbations to the strongly lensed background sources caused by substructure. The next generation of telescopes and adaptive optics will lower this limit for background galaxy sources, and deep very long baseline interferometry (VLBI) imaging of extended radio jets will potentially enable the detection of masses as low as $\sim$$10^6 \, M_\odot$ with gravitational imaging.

Observations of quasar lenses at longer wavelengths can provide a microlensing-free probe of substructure. Redwards of $\sim$4 μm rest frame, the quasar continuum emission is expected to be dominated by the dusty torus, which is sufficiently large to be unlensed by stars (e.g. Sluse et al. 2012). For a typical source redshift, this implies that mid-IR imaging at wavelengths greater than 10 μm can probe dark matter substructure. This method has been applied successfully to several systems with larger image separations (e.g. Chiba et al. 2005; MacLeod, Kochanek & Agol 2009; Minezaki et al. 2009). *James Webb Space Telescope* (*JWST*) can provide the spatial resolution required to extend mid-IR flux ratio measurements to systems with smaller image separations. Jackson et al. (2015) demonstrated that deep radio observations of radio-quiet lensed quasars can also successfully be used as a microlensing-free probe of substructure, albeit with larger flux uncertainties than radio-loud systems ($\sim$8–10 per cent compared with $\sim$3–5 per cent for radio-loud systems). They estimate that this method can be applied to approximately half of optically selected quasar lenses.

Strongly lensed quasar *narrow-line* emission provides an alternate probe of substructure, with comparable precision to radio-loud lensing studies. This method, originally proposed by Moustakas & Metcalf (2003), is extremely promising as it enables the measurement of substructure with current observational facilities in virtually all of the tens of optically selected quadruple quasar lenses predicted to be found in DES (Agnello et al. 2015; Ostrovski et al. 2017, http://strides.astro.ucla.edu/) and other wide-field imaging surveys, and the hundreds forecast to be found using Large Synoptic Survey Telescope (LSST; Oguri & Marshall 2010). Sugai et al. (2007) demonstrated this method for the gravitational lens RXS 1131$-$1231 using seeing-limited observations with the Subaru integral field spectrograph. This lens has an unusually large separation of $>$1 arcsec between each of the images.

Higher resolution imaging is necessary for the majority of quad lens systems, which often have at least one pair of images separated only by a few tenths of an arcsecond. Adaptive optics can provide the necessary spatial resolution to isolate the lensed images. Nierenberg et al. (2014) used the integral field spectrograph, the Ohio State InfraRed Imager/Spectrometer (OSIRIS; Larkin et al. 2006), at Keck with adaptive optics to measure spatially resolved narrow-line flux ratios in the gravitational lens B1422+231 (Patnaik et al. 1992). Adaptive optics is an effective tool, however it can only be applied to those systems in which the narrow-line emission of interest falls in a suitable wavelength range for adaptive optics correction, for instance either *H* or *K* band in the case of Keck OSIRIS. Furthermore, adaptive optics requires the presence of a nearby, bright, tip-tilt star, although often the lensed quasar itself is bright enough for this purpose. Space-based spatially resolved spectroscopy provides an alternative for those systems that fall outside these wavelength bands, or are at a declination unsuitable to Keck OSIRIS spectroscopy.
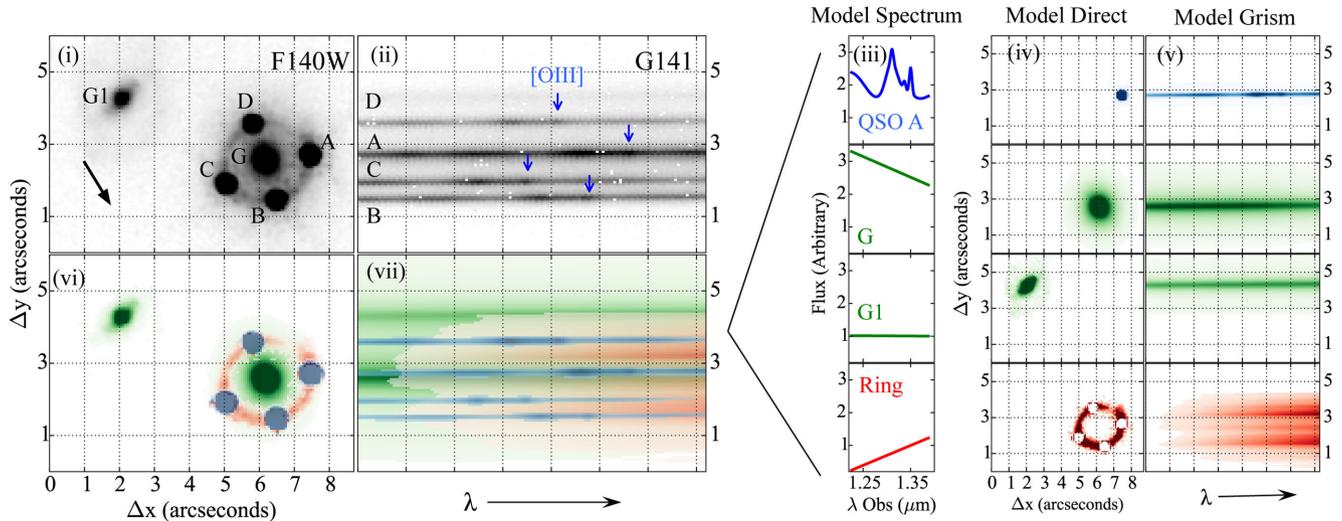
Figure 1. Demonstration of the forward modelling method used to infer spectral parameters. Note that the image contrasts have been altered between images to highlight different features. Panel (i): drizzled *F140W* image, arrow indicates north. Panel (ii): interlaced G141W grism image, with light dispersed along the *x*-axis of the *F140W* image. QSO spectra (A–D) are labelled. They overlap with spectra from the ring, the main deflector (G) and the spiral galaxy (G1). Blue arrows indicate the location of narrow [O III] 4959 and 5007 Å emission that are partially blended at this resolution. Column (iii): MCMC proposed 1D spectra for four of the seven components labelled in panel (i). Each of the QSO images A–D has a separate model spectrum (shown in Fig. 3), only spectrum A is shown here. Column (iv): model direct images for each separate spectral component, described in Section 3.1. The central QSO pixels are masked in the ring model to account for noisy PSF subtraction in this region. Column (v): model 2D grism images for each spectral component generated from convolving the model spectra in column (iii) with the model direct image in column (iv). Panels (vi) and (vii): final, combined model direct image and model grism image, generated from the sum of columns (iv) and (v), respectively (and the other three QSO images not shown). Colours are the same as in columns (iii–v). The goodness of fit is calculated by the χ^2 difference between true and model 2D G141 images.

In this work we demonstrate that the *Hubble Space Telescope* infrared grism on the Wide Field Camera 3 (WFC3) can be used to measure spatially resolved narrow-line image fluxes, with comparable sensitivity to substructure for ground-based results from Keck. We present an analysis of WFC3 grism observations of HE 0435−1223 (hereafter HE 0435; Wisotzki et al. 2002) to demonstrate our reduction mechanism. The deflector is an early-type galaxy at redshift 0.4546 (Morgan et al. 2005) and the source is a quasar at redshift 1.693 (Sluse et al. 2012). This system has been extensively studied since its discovery; it has been monitored over a decade, and its relatively long time delay and significant intrinsic variability have made it a powerful probe of the group density profile and the Hubble constant (Kochanek et al. 2006; Courbin et al. 2011; Bonvin et al. 2017; Wong et al. 2017). Thanks to this attention, there are numerous multiband and spectroscopic measurements available for comparison and significant effort has gone into measuring the properties of the environment of the deflector (e.g. Momcheva et al. 2006, 2015; Wong et al. 2011; Sluse et al. 2016; Wilson et al. 2016), which is important for comparing detected subhalo properties with predictions from CDM.

In Section 2 we describe the observations and initial data reduction for this system. In Section 3 we describe the spectral extraction pipeline developed to measure narrow-line fluxes. In Section 4 we report measured quasar spectral features and integrated emission fluxes, and compare the measured fluxes to results from broad-band and radio studies. In Section 5 we perform a simple gravitational lens inference to test for the presence of substructure. In Section 6 we test for the effects of resolved narrow-line emission on our results. In Section 7 we discuss the constraint from this system. In Section 8 we provide a brief summary of the main conclusions.

We assume a flat ΛCDM cosmology with $h = 0.7$ and $\Omega_m = 0.3$. All magnitudes are given in the AB system (Oke 1974).

2 OBSERVATIONS AND INITIAL REDUCTION

We observed the gravitationally lensed quasar HE 0435 as a part of HST-GO-13732 (P.I. Nierenberg), a grism survey of narrow-line emission in six quasar lenses. The target was observed on 2015 August 30 for 2062 s with the G141 grism, and for 400 s with *F140W* direct imaging. The observation was taken at a dispersion angle of 147°east of north so that the dispersed quasar images would be maximally separated from each other along the direction perpendicular to the dispersion axis. In order to recover subpixel information, we split the observations into a four-point dither pattern with half-integer subpixel offsets following the procedure of Brammer et al. (2012, see also Schmidt et al. 2014; Momcheva et al. 2016). For each dither position, we took a 100 s direct exposure with *F140W*, immediately followed by a 515 s G141 exposure. The *F140W* direct images at each dither position were used to obtain accurate wavelength solutions for each G141 exposure (Brammer et al. 2012).

Raw *F140W* and G141 exposures were individually processed with ASTRODRIZZLE (Gonzaga et al. 2012) in order to reject cosmic rays, remove geometric distortion and perform flat-field subtraction (Koekemoer et al. 2011; Brammer et al. 2012). The *F140W* exposures were drizzled on to a 0.06 arcsec pixel scale (approximately half the native pixel size). The upper left-hand panel of Fig. 1 shows the drizzled *F140W* image of the gravitational lens and nearby spiral galaxy G1.

The G141 exposures were interlaced and combined on to a 0.06 arcsec pixel scale, corresponding to an observed wavelength resolution of ~22 Å pixel^{-1} following Brammer et al. (2012), Schmidt et al. (2014) and Momcheva et al. (2016). Unlike drizzling, interlacing does not introduce correlated pixel flux errors. Fig. 1 shows the final interlaced grism data with arrows indicating the 5007 and 4959 Å [O III] doublet that is partially blended given the grism resolution.

3 SPECTRAL EXTRACTION

Our goal is to measure the narrow-line emission flux in each lensed quasar image, taking into account blending between distinct spectral components after they are dispersed by the grism. The lower left-hand panel of Fig. 1 highlights the blending by showing how the light from the ring, quasar images and galaxies combine in a model grism image. In order to rigorously account for the overlapping spectra in the grism image, we employ a forward modelling approach. We discuss this method in detail in the following subsections; in brief, we generate a model direct image and use the 3D-HST grism simulation code (Brammer et al. 2012) to iteratively map proposed component spectra into a simulated 2D grism image, which is then compared with the original interlaced grism image to compute a χ^2 goodness of fit. In Section 3.1 we discuss how the model direct image is generated, in Section 3.2 we discuss the 1D models we use for the spectral components and in Section 3.3 we describe the statistical inference.

3.1 Direct image model

The grism image is effectively a convolution of an object spectrum with its direct image. Thus given a model for the direct image, it is possible to generate a predicted grism spectrum. We model the direct image as having seven distinct spatial plus spectral components: four separate quasars, the main deflector, the quasar host galaxy that is lensed into a ring and the nearby galaxy G1. Here we discuss how the direct image model is generated for each of the components. These direct model images are then combined with model 1D spectra, as discussed in Section 3.2, in order to generate model grism images.

We model the four quasar images as point sources, using a nearby star to model the point spread function (PSF). We optimize the point source positions and fluxes using GALFIT (Peng et al. 2002, 2010). A possible concern for using a drizzled star image as the model for the point source is that the true PSF is not accurately captured at the exact location of the lensed images. Furthermore, the full width at half-maximum (FWHM) of the grism PSF varies slightly with wavelength, which the 3D-HST pipeline does not account for in the forward modelling process. We have checked that the exact PSF model does not impact our inference on the [O III] flux ratios by running the entire modelling process described in the next two subsections with a total of five different PSF models: (1) a median combination of stars in the *F140W* field of view; (2) and (3) the median star blurred by 10 and 15 per cent; (4) and (5) two different nearby stars. While the choice of PSF model affected the overall fit to the 2D grism image, we found that it had no impact on our inference of the relative [O III] fluxes. The quasar image positions are listed in Table A1, with uncertainties given by the variation in best-fitting GALFIT positions for the different PSF models.

In order to disentangle the lens galaxy light from the prominent ring and bright quasar images, we start with the empirical model of the deflector by Wong et al. (2017) derived from very deep (9337 s), F160W imaging. This model is generated from a superposition of Chameleon profiles (Dutton et al. 2011), and is based on the simultaneous fitting of the quasar point images, a model for the lensed background quasar host galaxy, and the lens. To generate a model for the *F140W* light profile, we start by fitting two Sérsic profiles to the empirical F160W model. Next, we hold all of the parameters for the Sérsic profiles obtained in the previous step fixed except for the total flux, and fit the *F140W* direct image as a combination of the lens and quasar images. We subtract the

best-fitting galaxy and quasi-stellar object (QSO) models from the direct *F140W* image and are left with a residual image composed primarily of the ring. In the third step, we subtract the residual ring image from the original *F140W* image, and re-fit the ring-subtracted image in the galaxy and QSO models, this time allowing all of the Sérsic parameters for the galaxy model to vary freely. The final model for the galaxy is taken from the third step. We have tested several different iterations of this process, including allowing for less flexibility in the galaxy model to verify that the inferred narrow-line fluxes and image positions are not sensitive to the exact galaxy model, although the overall fit to the 2D grism image varies.

The ring model is generated by subtracting the best-fitting lens galaxy and quasar models from the direct image. We mask a small region near the centre of each QSO image where the PSF subtraction is noisy. We have confirmed that the inferred [O III] fluxes are not sensitive to the exact size of the masked region. Finally for G1, we simply use a small cut-out of the direct image, which is possible because it is isolated in the direct image.

3.2 1D spectral models

In this section we describe the analytic models we use for the 1D spectra. Although our grism data extend over a significantly larger wavelength region, we confine our model and comparison with the data to a small wavelength region around [O III]. Fig. 2 shows the fitted regions for each QSO image. The extent of the region is approximately rest frame 4500–5200 Å, but it varies between the lensed QSO images and is selected to achieve two goals. First, to provide sufficient spectral coverage to obtain a good constraint on the broad Fe, Hβ and continuum features that overlap with the [O III] emission (Fig. 2). Secondly, the modelled region is extended where there is a possibility of an emission feature blending with the [O III] emission of a neighbouring 2D spectrum. An example of the latter case is found in the redward extension of the fitted region for the QSO spectrum C, in order to include the possible broad Fe flux contribution from image C to the image B narrow-line emission (Fig. 1).

Although gravitational lensing is not wavelength dependent, the spectra of the four QSO images may not necessarily be related by a simple multiplicative magnification factor owing to the intrinsic variability of the QSO (which can lead to different image spectra owing to the varying image arrival times) and to the differential effect of stellar microlensing as a function of intrinsic source size. Thus we construct a model for the four QSO spectra which enables variations due to these effects. The QSO spectrum in the wavelength range of interest is composed of broad Fe and Hβ emission, continuum emission and narrow [O III] and Hβ emission.

We model the broad Hβ emission as a Gaussian with an intrinsic redshift offset from the [O III] emission left as a free parameter. Although other studies have found that the Hβ line profile can have a complicated structure, the low resolution of the grism data does not warrant a higher order model; Fig. 2 demonstrates that a Gaussian is sufficient to fit the line profile in this case. The broad Fe emission is modelled using IZw1 templates which we interpolate in velocity space following Bennert et al. (2011) and Woo et al. (2006). The broad Fe emission velocity is independent of the broad emission width in our model. The model allows for a redshift offset between the broad Fe emission and the [O III] emission.

We model the continuum emission as a straight line rather than a power law, given the small fractional size of the wavelength region of interest. The slope and amplitude of the line are left as free parameters. The broad-line and continuum emission both arise from
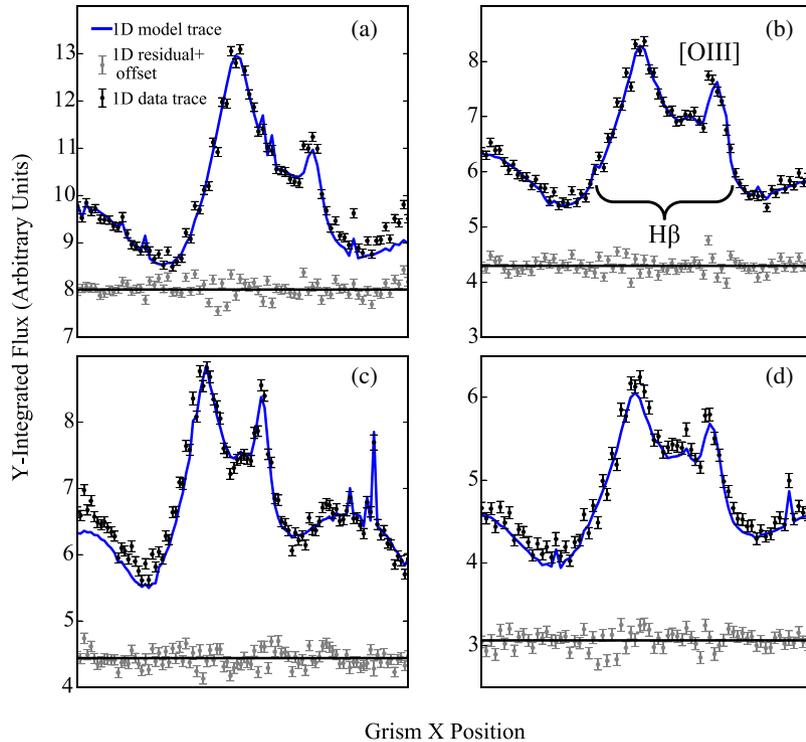
Figure 2. Lensed quasar spectra extracted along the *x*-axis of the 2D grism image (Fig. 1) via PSF weighted averaging along the *y*-axis. Absolute fluxes are arbitrary. 1D spectra include contamination from neighbouring dispersed QSO images, lens galaxy light, the lens ring and the nearby spiral galaxy G1 as illustrated in Fig. 1. The residual is derived by subtracting the 2D grism model from the 2D grism data, and then performing the PSF weighted *y*-axis averaging; thus it is not a simple subtraction of the blue line from the black points. The residual has been offset from zero by the amount indicated by the horizontal lines for ease of visualization. The modelled region varies slightly for each image depending on its position in the 2D image, as discussed in Section 3.2.

regions that can be affected by microlensing, which we discuss in more detail in Section 4. To account for this possibility we allow the width and amplitude of the Hβ emission and the slope and amplitude of the continuum emission to vary independently between the model spectra. We did not find significant evidence for variations in the broad Fe velocities between images, and so kept them fixed for our final analysis, however we allowed the Fe amplitudes to vary independently from the Hβ amplitudes.

Unlike the broad and continuum emission, narrow [O III] and Hβ emission come from a sufficiently extended source (greater than tens of pc) not to be affected by either stellar microlensing or intrinsic variability (Moustakas & Metcalf 2003; Bennert et al. 2006a,b; Müller-Sánchez et al. 2011). Owing to this, we assume that both the line widths and the relative amplitudes of [O III] and narrow Hβ should be constant between the lensed images. We model the [O III] doublet and Hβ narrow lines as Gaussians, and assume that they have the same redshift, which is valid given the spectral resolution of the grism. The ratio of the [O III] doublet 4959 and 5007 amplitudes is fixed to the quantum-mechanically predicted value of 1/3.

The 1D models for the deflector, ring and G1 spectra are modelled as straight lines over the short wavelength region of interest, with amplitudes and slopes as free parameters. We do not find evidence requiring the inclusion of emission or absorption features in any of these spectra relative to the measurement uncertainties and given the brightness of the QSO spectra (see e.g. Fig. 2).

We assume that the image fluxes are not affected by differential dust extinction. In the rest frame of the lens, the [O III] emission lines lie at roughly ∼9300 Å. At this wavelength, total dust extinction in lens galaxies, and early-type galaxies in general, is typically of order

only a few hundredths of a magnitude (e.g. Falco et al. 1999; Ferrari et al. 1999), which is well within our overall flux measurement uncertainty. This assumption is further supported by the similarity of the broad-band optical colours of the images (Wisotzki et al. 2003). The images also have mutually consistent C IV (lens rest frame ∼2790 Å) and Hβ (lens rest frame ∼9300 Å) broad-line flux ratios.

3.3 Inference of QSO spectral parameters

We infer the probability distribution of the parameters of the 1D spectral models using a Bayesian forward modelling approach with the EMCEE Markov Chain Monte Carlo software package (Foreman-Mackey et al. 2013). For each step, the Markov Chain Monte Carlo (MCMC) algorithm proposes parameters for the 1D spectra of all seven distinct spectral components (four QSO images, the main galaxy, the lens ring and G1). We then simulate dispersed images of each separate component and add them to generate a full model 2D grism image. Finally, the χ^2 of the fit is computed relative to the original 2D interlaced image. Fig. 1 illustrates how the model 2D direct image components are dispersed into the model 2D grism image for each MCMC step.

4 SPECTRAL FORWARD MODELLING RESULTS

Fig. 2 shows the 1D model, data and residual 'traces' for the four lensed QSO images. These traces are obtained by integrating the flux along the *y*-axis in the 2D image, weighted by the relative flux
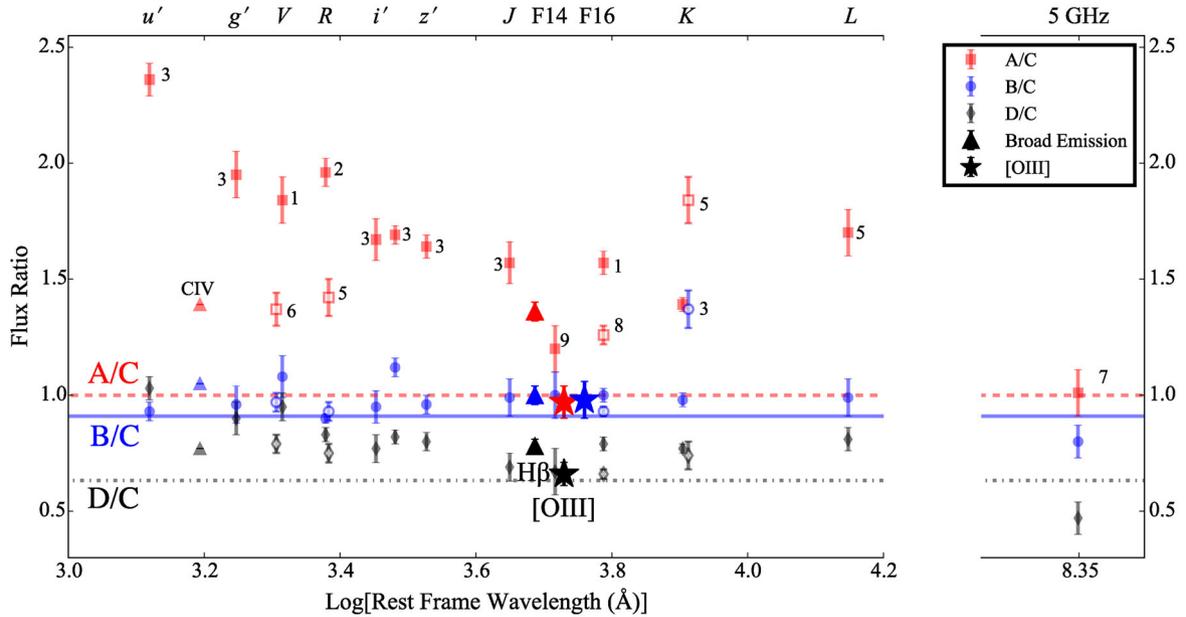
Figure 3. Flux ratio measurements for HE 0435 selected to represent variations with wavelength and time. References along with measurement dates are listed in Table A2. Numbers correspond to dates plotted in Fig. 4. Squares, circles and diamonds indicate broad-band continuum flux measurements, which are subject to time-delay-induced variability and microlensing bluewards of ∼4 μm rest frame. Stars and triangles represent [O III] and broad-line flux ratios. Measurements in the same filter, but from different years are slightly offset from each other for clarity, with the later measurement plotted with an open symbol. [O III] values have been shifted redwards to avoid overlap with broad Hβ results. The B/C [O III] flux ratio has been artificially shifted redwards so it does not lie on top of the A/C value. Dashed, solid and dash–dot lines represent the best smooth model prediction for the A/C, B/C and D/C flux ratios, respectively, given the image positions and narrow-line fluxes measured in this work. Top labels list observed bands, where F14 and F16 are abbreviations for the *HST* filters *F140W* and *F160W*, respectively.

of the direct F140 model PSF along that axis. Jumps in flux are due to small misalignments between the dispersion axis and the detector axis. This comparison shows that the input model provides an excellent fit to the observed spectra.

From the spectral modelling we obtain flux ratios between the broad Hβ fluxes and the [O III] fluxes from the image pairs A/C, B/C and D/C. Given that the intrinsic quasar luminosity is not known, gravitational lensing analyses rely on ratios of image fluxes rather than their absolute values. In Fig. 3, we compare these flux ratios with measurements from other studies across a range of filters and for fixed filters at multiple dates. These measurements are chosen to represent how the flux ratios vary with wavelength and time, and are only a small subset of the many measurements of this system obtained for time variability studies (e.g. Kochanek et al. 2006; Courbin et al. 2011; Bonvin et al. 2017). Table A2 contains references and observing dates for all flux ratios plotted in Fig. 3.

The narrow [O III] flux ratios are strikingly different from optical to near-IR flux ratios that are subject to contamination by microlensing and intrinsic QSO time variability.

HE 0435 has been monitored for 15 yr (Kochanek et al. 2006; Courbin et al. 2011; Bonvin et al. 2017), and during that time has shown highly variable broad-band flux ratios due to stellar microlensing and intrinsic variability. The intrinsic variability particularly affects images B and D that have time delays of over a week relative to images A and C. Fig. 3 highlights several repeat measurements of the system that show significant variability.

Based on simulations of QSO accretion discs and dusty tori, bluewards of rest frame ∼4 μm (observed ∼10 μm), the accretion disc makes a dominant contribution to the QSO emission (Sluse et al. 2013). From chromatic microlensing studies of this system, the quasar continuum emission has a half-light radius of ∼$10^{16.3 \pm 0.3}$ cm

(or 0.003 pc) at a rest-frame wavelength of 8000 Å. Sluse et al. (2012) estimate the Mg II broad-line region size to be ∼$10^{18 \pm 1}$ cm (or 0.03 pc). These sizes correspond to 0.5–5 μas at the lens redshift, and are thus affected by stellar microlensing. Bonvin et al. (2017) have inferred the approximate amplitude of the observed *R*-band (rest frame ∼2500 Å) stellar microlensing as a function of time for each of the images since 2003. Fig. 4 shows an estimate from Bonvin et al. (2017) for the microlensing effect on the *R*-band flux ratios as a function of time assuming a 'true' flux ratio value indicated by the straight lines. The amplitude of microlensing depends on the source size thus bluer filters are more strongly affected by microlensing while redder continuum measurements are less affected. Blackburne et al. (2014) have performed a detailed study of differential microlensing as a function of wavelength for this system and their data are included in Fig. 3, and in Table A2.

The broad-line emission flux ratios of both Hβ from this study and C III] and C IV from Wisotzki et al. (2003) are closer to the narrow-line emission flux ratios, which is consistent with microlensing being a function of emission region size.

We can test for the effects of microlensing on our data by comparing the relative amplitudes of emission features in our inferred spectra. In Fig. 5, we plot the marginalized models for the lensed image spectra from our analysis, normalized to the peak of the [O III] flux at 5007 Å in order to highlight how the emission features vary between the lensed images. Image A shows significant morphological differences, with continuum and broad Hβ fluxes that are much higher relative to the [O III] flux than the other three images. This indicates that there is significant source-size-dependent lensing. This finding is consistent with the inferred *R*-band microlensing of image A as observed by Bonvin et al. (2017) and shown in Fig. 4.
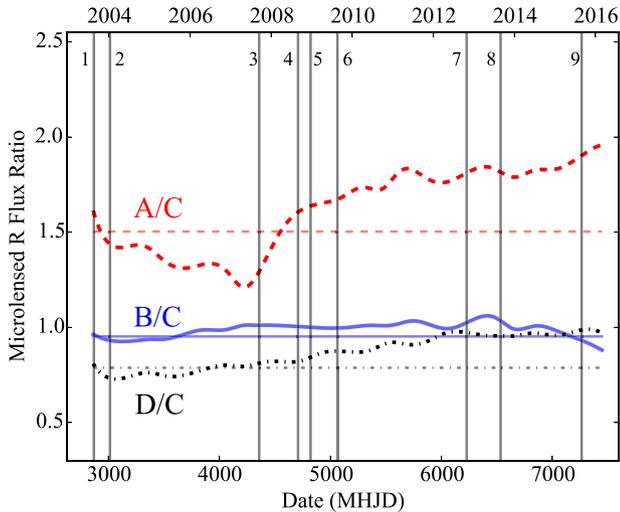
Figure 4. An estimate of fluctuations induced by microlensing in the *R*-band flux ratios as a function of time for HE 0435, by Bonvin et al. (2017). Vertical bars with numbers correspond to approximate dates for the measurements plotted in Fig. 3 and listed in Table A2. Measurements within 2 months of each other are combined to the same time point. Horizontal lines indicate the model 'true' flux ratios at the start of the monitoring campaign. MHJD = HJD − 245 0000.
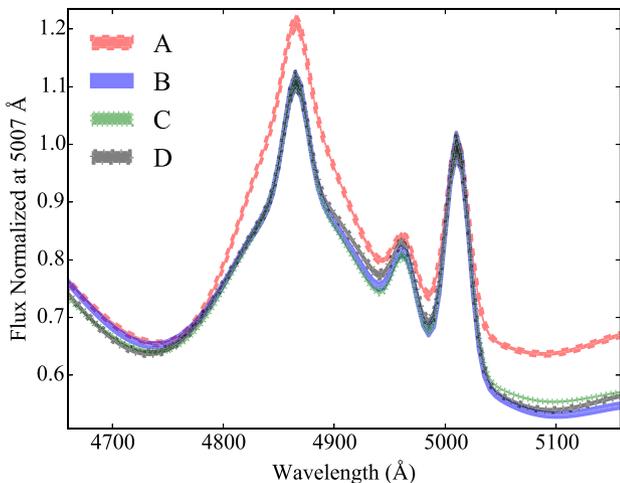
Figure 5. Posterior model spectra for each lensed quasar image, normalized to the peak of the narrow-line flux at 5007 Å to highlight variations in emission regions relative to the narrow line. Image A shows the most significant difference, with continuum and broad-line fluxes higher relative to the [O III] flux than in the other three images.

The narrow-line [O III] flux ratios are consistent with 5 GHz radio measurements from Jackson et al. (2015), with A/C, B/C and D/C differing at 0.25σ, 1.8σ and 2.2σ, respectively. This is expected given that both sources are expected to be extended enough to avoid all microlensing contamination. Although the results do not differ significantly, we note that Jackson et al. (2015) found that their radio emission was somewhat resolved, with an intrinsic source size of $\sigma \sim 288$ pc, assuming the source had a Gaussian flux distribution. This affects the flux ratios predicted from gravitational lensing relative to a point source for a fixed deflector mass model. We discuss this further in Section 6, where we also place limits on the size of

Table 1. Gravitational lens model parameters for the main deflector and external shear inferred from image positions and [O III] fluxes that are given in Tables A1 and A2, respectively.

Parameter	Value	Description
θ_E	$1\farcs200 \pm 0.003$	Spherical Einstein radius
q	0.91 ± 0.03	b/a
PA	-8 ± 5	Degrees east of north
γ_{ext}	0.063 ± 0.007	External shear amplitude
θ_γ	-18 ± 2	Direction of external shear (degrees east of north)

the narrow-emission region in our data and we examine the effects of a resolved narrow emission line region on our results.

5 GRAVITATIONAL LENS MODELLING

The lensed image positions and [O III] flux ratios are sensitive to the mass distribution of the deflector. As discussed in the Introduction, the image fluxes are particularly sensitive to small-scale perturbations caused by dark matter subhaloes. In this section we perform a gravitational lensing analysis of the system using the image positions and [O III] fluxes. We do not include the time delays as a constraint, given that they are minimally sensitive to perturbers near the lensed images, relative to the image fluxes and positions, and given the time delay measurement uncertainties for this system (Keeton 2009; Keeton & Moustakas 2009). Furthermore, unlike image positions and fluxes, which are most sensitive to the local mass distribution and can thus be well matched without including G1 explicitly (e.g. Sluse et al. 2012), the time delays are sensitive to the larger-scale environment and using them as a constraint would require including the complex multiplane lensing effect of G1 that is at a higher redshift than the main deflector (Jackson et al. 2015; Sluse et al. 2016; Bonvin et al. 2017; Wong et al. 2017). The macro-model parameters including the external shear term are left free to vary, and thus absorb the large-scale contributions from G1.

In Section 5.1 we discuss the optimum smooth mass model fit to the data. In Section 5.2 we place limits on the presence of substructure near the lensed images. In Section 5.3 we discuss the effects of a finite narrow-line emission region.

5.1 Smooth model

We start with a simple, smooth mass distribution model of the system. We model the main deflector as a singular isothermal ($\rho(r) \propto r^{-2}$) ellipsoid (SIE) that has been shown to provide an excellent match to the combined stellar and dark matter mass distributions of elliptical galaxies well beyond the Einstein radius (e.g. Rusin, Kochanek & Keeton 2003; Rusin & Kochanek 2005; Gavazzi et al. 2007; Treu 2010; Gilman et al. 2017). The SIE has five free parameters: the lens centroid (two parameters), the ellipticity and position angle and the Einstein radius. The smooth model also includes parameters describing the magnitude and direction of external shear that can be generated by the group environment of the galaxy.

We optimize these model parameters relative to the observed lensed image positions and [O III] fluxes using GRAVLENS (Keeton 2001a,b), and find an overall best-fitting χ^2 of 1.7 for 1 degree of freedom (DOF). Table 1 lists the mean and 1σ uncertainties for the lens model parameters. Fig. 3 shows the best-fitting model fluxes as straight lines. The χ^2 of 1.7 for 1 DOF indicates that the data have a roughly ~20 per cent chance of being drawn from the best-fitting model, thus we do not find significant evidence indicating

that a more complex model with a significant subhalo contribution is necessary to fit the data. Our inferred lens model parameters are comparable to other values in the literature which are inferred from microlensing-free data. In particular, our inferred Einstein radius and external shear parameters are consistent with the single lens model parameters from Sluse et al. (2012). As expected, our inferred spherical equivalent Einstein radius of 1.200 ± 0.003 arcsec for the single lens model is somewhat higher than other values in the literature that explicitly include G1 and find values for the Einstein radius of the main deflector ranging from 1.07 ± 0.02 (Jackson et al. 2015) to 1.182 ± 0.002 arcsec (Wong et al. 2017). We also infer a somewhat higher value for the external shear than models that include G1 in addition to an external shear: $\gamma_{ext} = 0.063 \pm 0.007$ compared with 0.039 ± 0.004 and 0.030 ± 0.003 for Jackson et al. (2015) and Wong et al. (2017), respectively.

This result differs from the result of Fadely & Keeton (2012) who found that the *L*-band image fluxes and positions could not be fit with a smooth model although they did not report their best-fitting lens model parameters. They note that the observed *L* band corresponds to the rest frame 14000 Å at the redshift of the quasar, which should have a significant flux contribution from the dusty torus in addition to the accretion disc continuum emission (Wittkowski et al. 2004; Hönig et al. 2008; Sluse et al. 2013). In order to test for possible microlensing of the continuum component of the *L* emission, they analysed 2 yr worth of monitoring data from Kochanek et al. (2006). Unluckily over this time-scale there was not significant evidence for microlensing-induced variability in image A. Longer baseline data over 15 yr (Bonvin et al. 2017) reveal significant microlensing of image A, as shown in Fig. 4, which likely affected *L*-band flux ratios.

In the following subsection we place limits on the presence of perturbing subhaloes near the lensed images.

5.2 Limits on the presence of substructure

Each of the image positions and fluxes provides a local constraint on the presence of small scale structure. In this subsection we test the limits on the presence of a single perturbing subhalo given our [O III] flux and position measurements. We test the measurement sensitivity to two different perturber masses of $M_{600} \sim 10^8$ and $10^7 \, \mathrm{M_\odot}$, where M_{600} is the integrated mass within 600 pc of the centre of the perturber. These masses are chosen to be above and below the limit where the 'missing satellite problem' is observed in the Milky Way (e.g. Strigari et al. 2008).

As demonstrated in Nierenberg et al. (2014), the perturber mass distribution can significantly affect the predicted lensing signal with fixed M_{600}. This is due to the fact that shallower mass profiles must have a higher overall normalization than steeper mass profiles in order to achieve the same interior M_{600} integrated mass. This in turn causes the shallower mass profile to have a longer range impact on the observed image fluxes and magnifications. Here we demonstrate the lensing effect for two mass profiles, a singular isothermal sphere (SIS) that has an $m(r) \propto r^{-2}$, and a Navarro–Frenk–White (NFW; Navarro, Frenk & White 1996) halo that has a shallow interior profile of $m(r) \propto r^{-1}$, which transitions to a steeper value $m(r) \propto r^{-3}$ outside of a scale radius. We obtain the scale radius from the mass–concentration relation predicted by Macciò, Dutton & van den Bosch (2008) assuming a 5-yr *Wilkinson Microwave Anisotropy Probe* (*WMAP*5) cosmology (Dunkley et al. 2009).

For each perturber mass, and for each mass profile, we iteratively place the perturber at a fixed position, re-optimize the smooth model parameters and compare the new best-fitting χ^2 with the original χ^2

in the absence of a perturbation. We choose the grid spacing qualitatively to ensure that relevant angular dependences are captured in the χ^2 distribution as a function of position. We find that variations are well captured by a spacing of 0.1 arcsec for the $M_{600} \sim 10^8 \, \mathrm{M_\odot}$ mass perturber and 0.01 arcsec for the $\sim 10^7 \, \mathrm{M_\odot}$, respectively.

Fig. 6 shows the projected 2σ and 3σ exclusion regions ($p < 5$ and 0.3 per cent) for an SIS perturber with Einstein radius of 0.01 and 0.001 arcsec, respectively. Assuming the perturber is in the plane of the lens galaxy, these Einstein radii correspond to integrated masses of $\sim 10^{8.2}$ and $10^{7.2} \, \mathrm{M_\odot}$, respectively, within 600 pc of the centre of the perturber, making them comparable to the Milky Way satellites Fornax and Sagittarius (Strigari et al. 2008), albeit with steeper mass profiles. Based on the average minimum radius at which including a perturber results in a model probability that is lower than 0.3 per cent, the average exclusion radius is ~ 0.4 (0.1), 0.3 (0.08), 0.4 (0.09) and 0.3 (0.06) arcsec for images A, B, C and D for the 0.01 (0.001) arcsec Einstein radius perturber.

Fig. 7 shows the projected 2σ and 3σ exclusion regions for an NFW perturber with scale radius of 1.0 and 0.1 arcsec corresponding to integrated masses of $\sim 10^8$ and $10^{7.2} \, \mathrm{M_\odot}$ within the central 600 pc of the perturber. Again, assuming a best-fitting model probability lower than 0.3 per cent be excluded, the average radius of exclusion is A: 1.2 (0.1) arcsec; B: 0.3 (0.08) arcsec; C: 1.1 (0.09) arcsec; D: 0.8 (0.06) arcsec for images A, B, C and D for the 1.0 (0.1) arcsec perturber. As expected, the NFW perturber must be further from the lensed images than an SIS perturber with the same mass to avoid significantly perturbing their fluxes and positions relative to the best-fitting smooth mass model.

6 FINITE SOURCE EFFECTS

Until this point we have assumed that the narrow-line emission in HE 0435 is unresolved with *HST* resolution, even after being strongly lensed and thus magnified. In this section, we explore the effect of a resolved emission region on our results. Jackson et al. (2015) found that the 5 GHz radio emission fluxes for HE 0435 were best fit with a smooth deflector mass distribution and a resolved source with an intrinsic size of $\sigma \sim 34$ mas (288 pc) for a Gaussian profile. This model significantly improved the fit to their data relative to an unresolved emission model. Given this, it is important that we test whether the narrow-line emission is resolved in our data, and what the effects would be of incorrectly assuming the narrow-line emission was unresolved.

The narrow-line region is extended and in some cases has been observed to extend out several hundreds of pc (e.g. Bennert et al. 2006b). The narrow-line flux is not uniform, and can be dominated by the central tens of pc even for very luminous quasars (Müller-Sánchez et al. 2011). Sluse et al. (2007) demonstrated that the [O III] emission in the gravitational lens RXS J1131−1231 is resolved in their data, and thus inferred a minimum size of ~ 150 pc. Nierenberg et al. (2014) found that the narrow [O III] emission for B1422+231 (Patnaik et al. 1992) was marginally resolved with a dispersion of ~ 10 mas (50 pc).

We perform two tests here. First, we simulate narrow-line emission for three different source sizes and then redo the forward modelling inference using these simulated extended narrow-line images in place of the original point source model for the direct image to test whether this leads to an improved fit to the observed spectra. Secondly, we simulate gravitational lenses with three different narrow-line source sizes, and model them under the (in this case, incorrect) assumption that the narrow emission is unresolved to test for the effect on the inferred flux ratios.
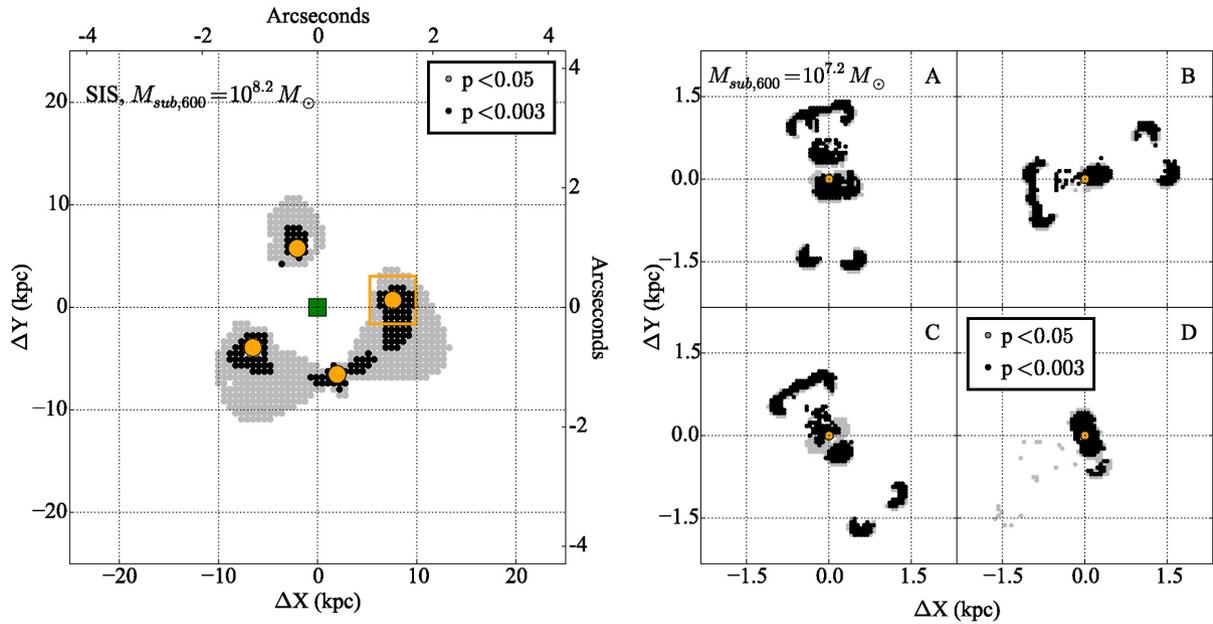
Figure 6. Projected exclusion regions for a singular isothermal spheroid perturber with a fixed mass. Light grey and dark grey positions are ruled out with greater than 95 and 99.7 per cent, respectively, based on the χ^2 probability of the best-fitting gravitational lens model to the image positions and [O III] fluxes after adding a perturber with $M_{600} = 10^{8.2}\,M_\odot$ (left-hand panel) and $10^{7.2}\,M_\odot$ (right-hand panel). The left-hand panel shows the entire lens system with the green square indicating the lens centroid, and the orange circles representing the quasar images. The orange box in the left-hand panel represents the size of the zoomed regions shown in the right-hand panel. The average projected radial limits are $\sim$0.4 (0.1), 0.3 (0.08), 0.4 (0.09) and 0.3 (0.06) arcsec for images A, B, C and D, respectively, for the $10^{8.2}$ ($10^{7.2}$) $M_\odot$ perturber. These exclusion regions correspond to cylinders with radii of $\sim$2(0.5) kpc around each lensed image, projected along the entire host halo.
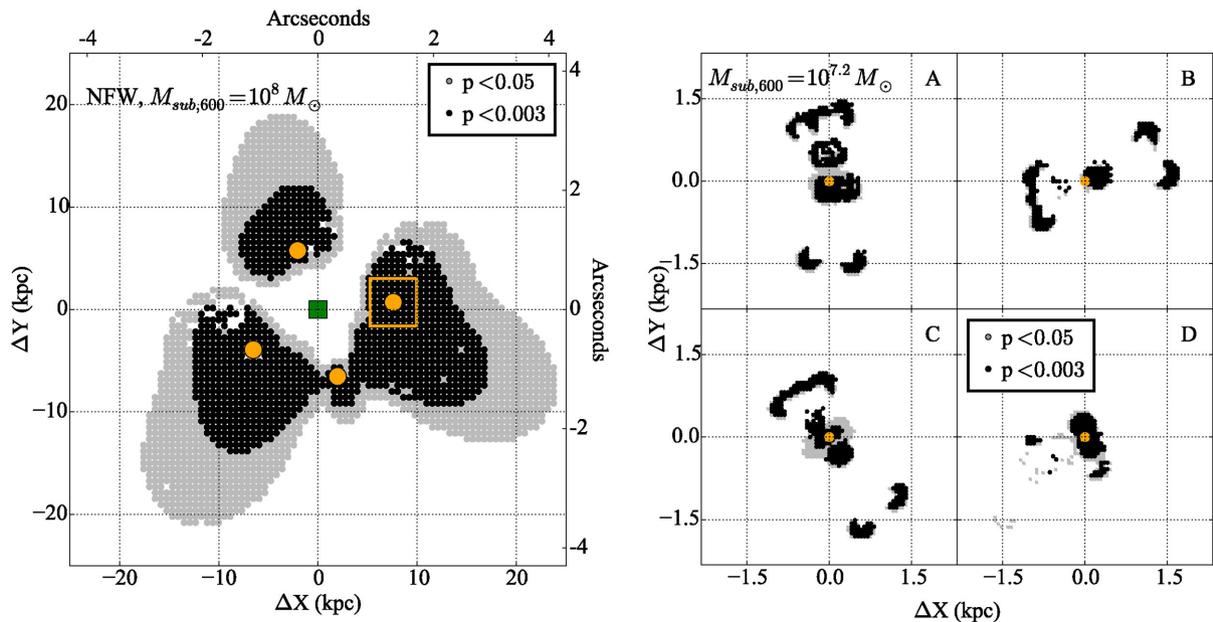
Figure 7. Exclusion regions for an NFW perturbing subhalo with $M_{600} = 10^8$ and $10^{7.2}\,M_\odot$, corresponding to NFW scale radii of 1.0 and 0.1 arcsec, respectively, determined the same way as in Fig. 6. The average radial limits are 1.2 (0.1), 0.3 (0.08), 1.1 (0.09) and 0.8 (0.06) arcsec for images A, B, C and D, respectively, for a 10^8 ($10^{7.2}$) $M_\odot$ perturber. These angular scales correspond to an average projected exclusion region of $\sim$6 (0.6) kpc at the redshift of the lens.

To test the model fit with a resolved narrow-line emission, we repeat the forward modelling inference described in Section 3 with an additional model component. As before, the QSO broad and continuum emission regions are modelled as being emitted by a point source. In order to account for a resolved narrow-line source, we generate a new model direct image for the narrow-line emission only. This extended emission model is generated assuming that the intrinsic source is a Gaussian with width σ_{NL}. We use the best-fitting gravitational lens model inferred from the continuum image positions and [O III] flux ratios in order to generate lensed extended
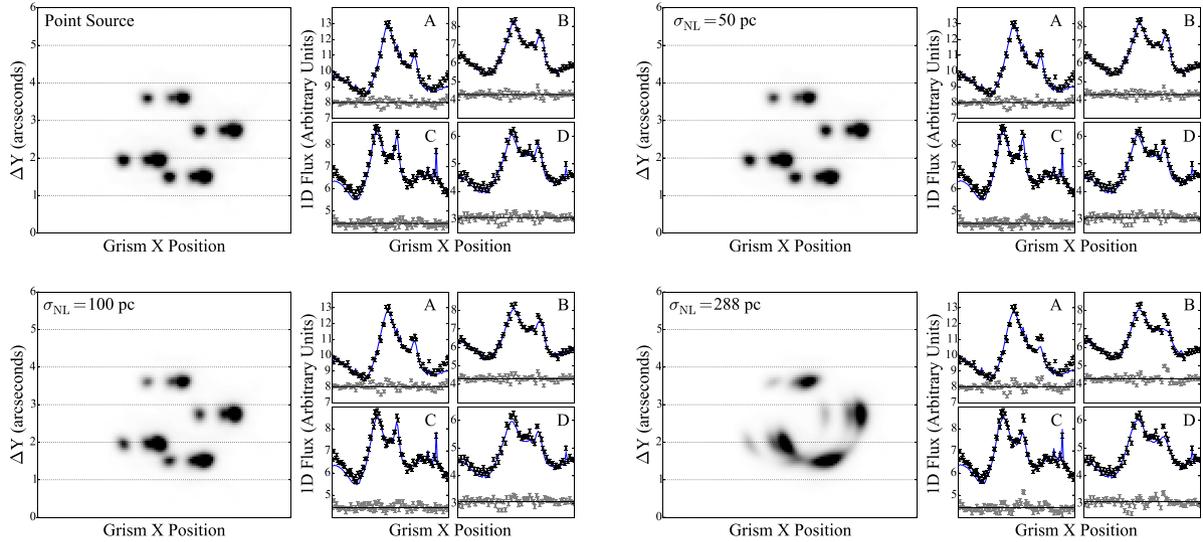
Figure 8. Effects of resolved narrow-line emission on model fit. We repeat the analysis of Section 4, now using a simulated extended narrow-emission region with a Gaussian flux distribution with dispersion $\sigma_{NL} = 50$, 100 and 288 pc. Here we show in the large left-hand panels, the best-fitting grism model for the extended narrow-line emission with a fixed finite size, after re-optimizing the 1D spectral parameters. While the full spectral model is used in the inference, we show only the narrow-line emission in the left-hand panels for clarity. In the adjacent right-hand panels we show a comparison of the model 1D trace spectra with the data. As in Fig. 2, the black points, blue line and light grey points correspond to the data, model and residual, respectively. The point source model and comparison between model and data are shown in the upper left-hand corner for comparison. The best-fitting χ^2 values calculated from the difference between the 2D model grism and true grism images are 13 228, 13 368 and 13 990 for 3909 DOF for the 50, 100 and 288 pc narrow-line regions, and 13 112 for 3906 DOF for the point source model. The point source model has three extra DOF as the narrow-line fluxes are allowed to vary independently.

emission models for the direct image. We generate models for σ_{NL} of 50, 100 and 288 pc. The latter source size is the best-fitting size for the 5 GHz radio emission for this system found by Jackson et al. (2015).

We then re-estimate the best-fitting 1D spectral parameters following the steps in Section 3 with two adjustments. First we use the new simulated extended source model as the direct image model for the narrow-line emission. The other QSO spectral components are modelled as being emitted by point sources as before. Secondly, because the simulated direct image for the [O III] fluxes by definition fixes the relative image fluxes, the spectral model has only a single parameter for the overall normalization of the narrow-line flux. Thus there are three fewer model parameters than the point source model in which the [O III] fluxes vary independently.

The left-hand panels of Fig. 8 show the best-fitting simulated grism images for the narrow-line components only of the three extended models and the point source model for comparison. Note that the narrow Hβ emission is also modelled as being extended as we assume it is emitted from the same region as the [O III] emission. While the 50 and 100 pc models differ only marginally from the point source model, the 288 pc source size is clearly extended with *HST* resolution. The [O III] doublet for images B and D in particular is nearly completely blended.

In the right-hand panels of Fig. 8 we compare the best-fitting model 1D trace to the data (analogous to Fig. 2) for each of the source sizes. The χ^2 comparison between the data and the simulated grism image grows progressively worse as the source size increases, with best-fitting χ^2 values of 13 228, 13 368 and 13 990 for 3909 DOF for the 50, 100 and 288 pc narrow-line regions, compared with 13 112 for 3906 DOF for the point source model.[1] Increasing

the narrow-emission source size results in a decreasing best-fitting peak narrow-line flux, despite the fact that the intrinsic narrow-line emission width is a free model parameter. This is due to the significantly extended emission that cannot be well fit in 2D.

We can also test what the effect would be on the inferred flux ratios, if we use a point source model for the narrow-line flux when the flux is actually resolved. For this we considered only the 50 and 100 pc models as the 288 pc model provides a markedly worse fit to the data. We simulated a mock spectrum using the resolved narrow-line models for the narrow-line component, and then inferred the image fluxes under the assumption used elsewhere in this work, that the narrow emission is unresolved in our data. The resulting flux ratios were A/C = 0.97 ± 0.08, B/C = 0.99 ± 0.07 and D/C = 0.65 ± 0.05 for the 50 pc source, and A/C = 1.0 ± 0.1, B/C = 1.0 ± 0.1 and D/C = 0.65 ± 0.08 for the 100 pc source. Both results are consistent with the input fluxes used to generate the resolved narrow-line source mock lenses, indicating that our results would not be biased by a marginally resolved narrow-line region.

7 DISCUSSION

We have demonstrated that the WFC3 IR grism provides sufficient spatial and spectral resolution to precisely measure lensed narrow-line emission with similar precision to continuum emission studies while avoiding the effects of microlensing, variability and differential dust extinction. We measure the [O III] flux ratios to be significantly offset from optical/near-IR continuum measurements of this system, particularly for image A. This is consistent with the results from long-term monitoring for this system, which have indicated that image A is significantly affected by microlensing, which systematically affects smaller emission regions. The [O III] flux ratios are consistent with the 5 GHz radio measurement from

---

[1] The point source model has three extra DOF as the narrow-line fluxes are allowed to vary independently.

Jackson et al. (2015), which are also extended and thus not affected by microlensing.

In order to fit the measured *F140W* image positions and [O III] fluxes, we rule out the presence of a perturbing NFW subhalo projected within roughly $\sim$1.0 (0.1) arcsec for an $M_{600} \sim 10^8$ (10^7) $M_\odot$ perturber near the lensed images. At the redshift of the lens these angular sizes correspond approximately to $\sim$6 and 0.6 kpc, respectively. It is informative to compare these limits with an approximate prediction of the number of subhaloes at these radii based on CDM models.

We perform a very basic estimate assuming that all possible perturbers are within the virial radius of the HE 0435 group and neglecting line-of-sight associations. From Sluse et al. (2016), the HE 0435 group has a virial mass of $\log[M_{200}/M_\odot] = 13.7 \pm 0.4$, where M_{200} is the mass within the region of the halo that has a mean density 200 times the critical density.

We estimated the number of subhaloes based on the results of Han et al. (2016), scaled to the virial mass of HE 0435, assuming that approximately half of subhaloes are destroyed through tidal interactions and merging following the recommendation of Han et al. (2016), and that an additional $\sim$30 per cent are destroyed by tidal interactions with a central baryonic potential (e.g. Garrison-Kimmel et al. 2017)

We assume that the subhalo mass profiles are NFW with the mass–concentration relation given by Macciò et al. (2008) scaled to the redshift of the group following the relation by Prada et al. (2012) and neglecting scatter. We fix M_{600} for subhaloes at infall, where M_{600} is the mass within the interior 600 pc of the subhalo. We make a simplifying assumption that M_{600} is not affected by tidal stripping after infall. This yields an estimated number of $\sim$250 surviving subhaloes in the group with masses greater than $M_{600} > 10^8$ $M_\odot$ and $\sim$19 000 subhaloes with $M_{600} > 10^7$ $M_\odot$ within $R_{200} = 540$ kpc.

We examined two different spatial distributions for the subhaloes. First, that the subhalo number density follows the mass distribution of the host halo everywhere, as predicted by pure CDM simulations in the absence of tidal stripping. Secondly, that the subhalo number density follows the mass distribution of the host halo except within the three-dimensional scale radius, where we assume all subhaloes are destroyed. This mimics an extreme version of the impact of the disc seen by Garrison-Kimmel et al. (2017). Both spatial distributions are normalized to have the same number of subhaloes. These two subhalo spatial distributions are chosen to bracket limits of the possible effect of the baryonic potential on our predicted detection rate.

In the first case, where the subhalo number density simply follows the NFW profile of the lens halo, we expect approximately $\sim$0.8 total subhaloes with $M_{600} > 10^8$ $M_\odot$ to be found within each $\sim$6 kpc projected excluded region. The mass function and size of the sensitivity region scales so that we expect to find $\sim$1 $M_{600} > 10^7$ $M_\odot$ subhalo within each of the four smaller $\sim$0.6 kpc projected regions around the images where we are sensitive to these lower masses.

In the second case in which all subhaloes are removed within the NFW scale radius of the host, we expect to detect approximately 0.08 (0.1) subhaloes with $M_{600} > 10^8$ $M_\odot$ (10^7 $M_\odot$) per exclusion region.

We also explored the effect of a warm dark matter (WDM) mass function, using the result from Schneider, Smith & Reed (2013) to estimate the shape of the subhalo mass function at infall given a 3 keV c^{-2} thermal relic dark matter particle, which is consistent with Lyα forest measurements (e.g. Viel, Bolton & Haehnelt 2009).

In this case, almost no subhaloes with masses $M_{600} < 10^8$ $M_\odot$ survive. Unlike the central baryonic potential, which is predicted to have a largely mass-independent effect on the subhalo population, WDM selectively destroys low-mass subhaloes. Promisingly, subhaloes with $M_{600} > 10^8$ $M_\odot$ are all expected to contain a significant number of stars (Strigari et al. 2008) in order to match comparisons between the Milky Way satellite luminosity function and CDM predictions (e.g. Strigari et al. 2008). This mean that CDM can be tested by comparing the rate of detections of substructure in narrow-line lenses with the rate predicted by luminous satellite studies, or measured by gravitational imaging studies that are sensitive to $M_{600} > 10^8$ $M_\odot$ subhaloes, after accounting for variations in host halo mass and sensitivity. This test can be performed in a sample of $\sim$20 lenses in which we would expect to detect $\sim$10 $M_{600} > 10^7$ $M_\odot$ subhaloes in the case of a CDM subhalo mass function, even in the case in which all subhaloes are destroyed within the NFW scale radius.

We emphasize, however, that a true comparison requires a CDM model that incorporates effects such as tidal stripping, as well as a marginalization over possible halo orientations, masses and formation histories (e.g. Jiang & van den Bosch 2016), and a range of baryonic physics implementations (e.g. Chua et al. 2016; Despali & Vegetti 2017). We leave such an analysis to a future paper, in which we jointly infer the properties of the subhalo mass function given our sample of narrow-line gravitational lenses measured with OSIRIS and the WFC3 grism.

8 SUMMARY

(i) We present a forward modelling method that uses the 3D-HST pipeline (Brammer et al. 2012) to measure spectra in the presence of significant spatial blending for G141 grism data. We apply this method to infer the lensed narrow, broad and continuum fluxes of the images in HE0435-1223.

(ii) The narrow [O III] flux ratios for HE 0435 are consistent with radio measurements from Jackson et al. (2015), and are significantly different from other emission measures that are subject to contamination by microlensing and intrinsic QSO time variability.

(iii) We find that the [O III] fluxes and image positions are well modelled with a simple gravitational lens model consisting of an SIE for the main galaxy in the presence of external shear.

(iv) Our data strongly disfavours a perturber with mass greater than $M_{600} = 10^{8.2}$ ($10^{7.2}$) $M_\odot$ within $\sim$1 (0.1) arcsec of the lensed images, where M_{600} is the projected perturber mass within its central 600 pc (best-fitting model probability <0.3 per cent).

(v) This demonstration that WFC3 grism measurement of narrow-line lensed quasars can be used to detect low-mass $M_{600} \sim 10^7$ $M_\odot$ subhaloes is extremely promising for future constraints of dark matter given the large number of quadruply imaged quasar lenses to be discovered in optical surveys such as DES and LSST, and with the follow-up that will be enabled by *JWST*.

REFERENCES

Agnello A. et al., 2015, MNRAS, 454, 1260

Bennert N., Jungwiert B., Komossa S., Haas M., Chini R., 2006a, A&A, 456, 953

Bennert N., Jungwiert B., Komossa S., Haas M., Chini R., 2006b, A&A, 459, 55

Bennert V. N., Auger M. W., Treu T., Woo J.-H., Malkan M. A., 2011, ApJ, 742, 107

Birrer S., Amara A., Refregier A., 2017, preprint (arXiv:e-prints)

Blackburne J. A., Pooley D., Rappaport S., Schechter P. L., 2011, ApJ, 729, 34

Blackburne J. A., Kochanek C. S., Chen B., Dai X., Chartas G., 2014, ApJ, 789, 125

Bonvin V. et al., 2017, MNRAS, 465, 4914

Brammer G. B. et al., 2012, ApJS, 200, 13

Brooks A. M., Kuhlen M., Zolotov A., Hooper D., 2013, ApJ, 765, 22

Chiba M., Minezaki T., Kashikawa N., Kataza H., Inoue K. T., 2005, ApJ, 627, 53

Chua K. T. E., Pillepich A., Rodriguez-Gomez V., Vogelsberger M., Bird S., Hernquist L., 2016, preprint (arXiv:1611.07991)

Courbin F. et al., 2011, A&A, 536, A53

Dalal N., Kochanek C. S., 2002, ApJ, 572, 25

Despali G., Vegetti S., 2017, MNRAS, 469, 1997

Diemand J., Kuhlen M., Madau P., Zemp M., Moore B., Potter D., Stadel J., 2008, Nature, 454, 735

Drlica-Wagner A. et al., 2015, ApJ, 813, 109

Dunkley J. et al., 2009, ApJ, 701, 1804

Dutton A. A. et al., 2011, MNRAS, 417, 1621

Fadely R., Keeton C. R., 2011, AJ, 141, 101

Fadely R., Keeton C. R., 2012, MNRAS, 419, 936

Falco E. E. et al., 1999, ApJ, 523, 617

Ferrari F., Pastoriza M. G., Macchetto F., Caon N., 1999, A&AS, 136, 269

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306

Garrison-Kimmel S. et al., 2017, preprint (arXiv:1701.03792)

Gavazzi R., Treu T., Rhodes J. D., Koopmans L. V. E., Bolton A. S., Burles S., Massey R. J., Moustakas L. A., 2007, ApJ, 667, 176

Gilman D., Agnello A., Treu T., Keeton C. R., Nierenberg A. M., 2017, MNRAS, 467, 3970

Gnedin N. Y., 2000, ApJ, 535, L75

Gonzaga S. et al., 2012, The DrizzlePac Handbook. STScI, Baltimore, MD

Guo Q. et al., 2011, MNRAS, 413, 101

Han J., Cole S., Frenk C. S., Jing Y., 2016, MNRAS, 457, 1208

Hargis J. R., Willman B., Peter A. H. G., 2014, ApJ, 795, L13

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

Hönig S. F., Smette A., Beckert T., Horst H., Duschl W., Gandhi P., Kishimoto M., Weigelt G., 2008, A&A, 485, L21

Hsueh J.-W., Fassnacht C. D., Vegetti S., McKean J. P., Spingola C., Auger M. W., Koopmans L. V. E., Lagattuta D. J., 2016, MNRAS, 463, L51

Jackson N., Tagore A. S., Roberts C., Sluse D., Stacey H., Vives-Arias H., Wucknitz O., Volino F., 2015, MNRAS, 454, 287

Jiang F., van den Bosch F. C., 2016, preprint (arXiv:1610.02399)

Kaufmann T., Bullock J. S., Maller A., Fang T., 2008, in Minchin R., Momjian E., eds, AIP Conf. Proc. Vol. 1035, The Evolution of Galaxies Through the Neutral Hydrogen Window. Am. Inst. Phys., New York, p. 147

Keeton C. R., 2001a, preprint (arXiv:astro-ph/0102341)

Keeton C. R., 2001b, preprint (arXiv:astro-ph/0102340)

Keeton C. R., 2009, preprint (arXiv:0908.3001)

Keeton C. R., Moustakas L. A., 2009, ApJ, 699, 1720

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

Kochanek C. S., Morgan N. D., Falco E. E., McLeod B. A., Winn J. N., Dembicky J., Ketzeback B., 2006, ApJ, 640, 47

Koekemoer A. M. et al., 2011, ApJS, 197, 36

Larkin J. et al., 2006, Proc. SPIE, 6269, 62691A

Lu Y. et al., 2014, ApJ, 795, 123

Macciò A. V., Dutton A. A., van den Bosch F. C., 2008, MNRAS, 391, 1940

Macciò A. V., Kang X., Fontanot F., Somerville R. S., Koposov S., Monaco P., 2010, MNRAS, 402, 1995

MacLeod C. L., Kochanek C. S., Agol E., 2009, ApJ, 699, 1578

Menci N., Gatti M., Fiore F., Lamastra A., 2014, A&A, 569, A37

Minezaki T., Chiba M., Kashikawa N., Inoue K. T., Kataza H., 2009, ApJ, 697, 610

Momcheva I., Williams K., Keeton C., Zabludoff A., 2006, ApJ, 641, 169

Momcheva I. G., Williams K. A., Cool R. J., Keeton C. R., Zabludoff A. I., 2015, ApJS, 219, 29

Momcheva I. G. et al., 2016, ApJS, 225, 27

Moore B., Ghigna S., Governato F., Lake G., Quinn T., Stadel J., Tozzi P., 1999, ApJ, 524, L19

Morgan N. D., Kochanek C. S., Pevunova O., Schechter P. L., 2005, AJ, 129, 2531

Moustakas L. A., Metcalf R. B., 2003, MNRAS, 339, 607

Müller-Sánchez F., Prieto M. A., Hicks E. K. S., Vives-Arias H., Davies R. I., Malkan M., Tacconi L. J., Genzel R., 2011, ApJ, 739, 69

Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563

Nierenberg A. M., Treu T., Menci N., Lu Y., Wang W., 2013, ApJ, 772, 146

Nierenberg A. M., Treu T., Wright S. A., Fassnacht C. D., Auger M. W., 2014, MNRAS, 442, 2434

Nierenberg A. M., Treu T., Menci N., Lu Y., Torrey P., Vogelsberger M., 2016, MNRAS, 462, 4473

Oguri M., Marshall P. J., 2010, MNRAS, 405, 2579

Oke J. B., 1974, ApJS, 27, 21

Ostrovski F. et al., 2017, MNRAS, 465, 4325

Patnaik A. R., Browne I. W. A., Walsh D., Chaffee F. H., Foltz C. B., 1992, MNRAS, 259, 1p

Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2002, AJ, 124, 266

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

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

Prada F., Klypin A. A., Cuesta A. J., Betancort-Rijo J. E., Primack J., 2012, MNRAS, 423, 3018

Ricci D. et al., 2011, A&A, 528, A42

Rusin D., Kochanek C. S., 2005, ApJ, 623, 666

Rusin D., Kochanek C. S., Keeton C. R., 2003, ApJ, 595, 29

Schmidt K. B. et al., 2014, ApJ, 782, L36

Schneider A., Smith R. E., Reed D., 2013, MNRAS, 433, 1573

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

Sluse D., Hutsemékers D., Courbin F., Meylan G., Wambsganss J., 2012, A&A, 544, A62

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

Sluse D. et al., 2016, preprint (arXiv:1607.00382)

Springel V., 2010, ARA&A, 48, 391

Springel V. et al., 2008, MNRAS, 391, 1685

Starkenburg E. et al., 2013, MNRAS, 429, 725

Strigari L. E., Bullock J. S., Kaplinghat M., Diemand J., Kuhlen M., Madau P., 2007, ApJ, 669, 676

Strigari L. E., Bullock J. S., Kaplinghat M., Simon J. D., Geha M., Willman B., Walker M. G., 2008, Nature, 454, 1096

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

Thoul A. A., Weinberg D. H., 1996, ApJ, 465, 608

Treu T., 2010, ARA&A, 48, 87

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

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

Viel M., Bolton J. S., Haehnelt M. G., 2009, MNRAS, 399, L39

Weinberg D. H., Colombi S., Davé R., Katz N., 2008, ApJ, 678, 6

Wetzel A. R., Hopkins P. F., Kim J.-h., Faucher-Giguére C.-A., Kereš D., Quataert E., 2016, ApJ, 827, L23

Weymann R. J. et al., 1980, Nature, 285, 641

Wilson M. L., Zabludoff A. I., Ammons S. M., Momcheva I. G., Williams K. A., Keeton C. R., 2016, ApJ, 833, 194

Wisotzki L., Schechter P. L., Bradt H. V., Heinmüller J., Reimers D., 2002, A&A, 395, 17

Wisotzki L., Becker T., Christensen L., Helms A., Jahnke K., Kelz A., Roth M. M., Sanchez S. F., 2003, A&A, 408, 455

Wittkowski M., Kervella P., Arsenault R., Paresce F., Beckert T., Weigelt G., 2004, A&A, 418, L39

Wong K. C., Keeton C. R., Williams K. A., Momcheva I. G., Zabludoff A. I., 2011, ApJ, 726, 84

Wong K. C. et al., 2017, MNRAS, 465, 4895

Woo J.-H., Treu T., Malkan M. A., Blandford R. D., 2006, ApJ, 645, 900

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

Zolotov A. et al., 2012, ApJ, 761, 71

Table A1. *F140W* image positions in units of arcseconds with north up and east left coordinates (rotated relative to Fig. 1).

Image	dRA	dDec.	Uncertainty
A	2.476	0.608	0.008
B	0.997	1.157	0.008
C	0	0	0.008
D	1.530	−1.005	0.008
G	1.314	0.067	0.06

APPENDIX: TABLE OF MEASUREMENTS OF FLUX RATIOS FOR HE 0435−1223

Here we provide (see Tables A1 and A2) the *F140W* lensed quasar image position and a selection of the many measurements of system chosen to illustrate how the flux ratios vary with wavelength and time, as shown in Figs 3 and 4.

Table A2. Subset of flux ratio measurements of HE 0435 from the literature, selected to represent variation across wavelength and with time, plotted in Fig. 3.

Filter	Date	f_A/f_C	f_B/f_C	f_D/f_C	Reference
u'	9/2007	2.36 ± 0.07	0.93 ± 0.04	1.03 ± 0.05	Blackburne et al. (2011)
C IV	9/2002	1.40[a]	1.05	0.770	Wisotzki et al. (2003)
g'	9/2007	1.95 ± 0.1	0.96 ± 0.08	0.9 ± 0.07	Blackburne et al. (2011)
V	8–9/2009	1.37 ± 0.07	0.97 ± 0.04	0.79 ± 0.04	Ricci et al. (2011)
F555W	8/2003	1.84 ± 0.1	1.08 ± 0.09	0.95 ± 0.06	Kochanek et al. (2006)
R	10/2007	1.89 ± 0.05	0.99 ± 0.04	0.88 ± 0.05	Blackburne et al. (2011)
R	10–12/2009	1.42 ± 0.08	0.93 ± 0.04	0.75 ± 0.04	Courbin et al. (2011)
i'	9/2007	1.67 ± 0.09	0.95 ± 0.07	0.77 ± 0.06	Blackburne et al. (2011)
F814W	8/2003	1.69 ± 0.04	1.12 ± 0.04	0.82 ± 0.03	Kochanek et al. (2006)
z'	9/2007	1.64 ± 0.05	0.96 ± 0.04	0.8 ± 0.04	Blackburne et al. (2011)
J	9/2007	1.57 ± 0.09	0.99 ± 0.08	0.69 ± 0.06	Blackburne et al. (2011)
Hβ	8/2015	1.36 ± 0.04	1.00 ± 0.04	0.78 ± 0.03	This work
[O III]	8/2015	0.97 ± 0.07	0.98 ± 0.07	0.66 ± 0.05	This work
F160W	8/2003	1.57 ± 0.05	1.0 ± 0.03	0.79 ± 0.03	Kochanek et al. (2006)
F160W	10/2012	1.30 ± 0.05	0.92 ± 0.05	0.66 ± 0.05	Wong et al. (2017)
K_s	9/2007	1.39 ± 0.03	0.98 ± 0.03	0.77 ± 0.02	Blackburne et al. (2011)
K	8/2008	1.84 ± 0.1	1.37 ± 0.08	0.74 ± 0.06	Fadely & Keeton (2011)
5 GHz	11/2012	1 ± 0.1	0.8 ± 0.07	0.47 ± 0.07	Jackson et al. (2015)

[a]Formal measurement uncertainties for Wisotzki et al. (2003) less than 0.1 per cent and dominated by unknown systematics.

This paper has been typeset from a TeX/LaTeX file prepared by the author.