

A dynamically cold disk galaxy in the early Universe

F. Rizzo¹, S. Vegetti¹, D. Powell¹, F. Fraternali², J. P. McKean^{2,3}, H. R. Stacey^{2,3,1} & S. D. M. White¹

¹*Max-Planck Institute for Astrophysics, Karl-Schwarzschild Str 1, D-85748 Garching, Germany*

²*Kapteyn Astronomical Institute, University of Groningen, Postbus 800, NL-9700 AV Groningen, the Netherlands*

³*ASTRON, Netherlands Institute for Radio Astronomy, Oude Hoogeveensedijk 4, NL-7991 PD Dwingeloo, the Netherlands*

The extreme astrophysical processes and conditions that characterize the early Universe are expected to result in young galaxies that are dynamically different from those observed today^{1–5}. This is because the strong effects associated with galaxy mergers and supernova explosions would lead to most young star-forming galaxies being dynamically hot, chaotic and strongly unstable^{1,2}. Here we report the presence of a dynamically cold, but highly star-forming, rotating disk in a galaxy at redshift (z) 4.2⁶, when the Universe was just 1.4 billion years old. Galaxy SPT–S J041839–4751.9 is strongly gravitationally lensed by a foreground galaxy at $z = 0.263$, and it is a typical dusty starburst, with global star-forming⁷ and dust properties⁸ that are in agreement with current numerical simulations⁹ and observations of its galaxy population¹⁰. Interferometric imaging at a spatial resolution of about 60 pc reveals a ratio of rotational-to-random motions of $V/\sigma = 9.7 \pm 0.4$, which is at least four times larger than expected from any galaxy evolution model at this epoch^{1–5}, but similar to the ratios of spiral galaxies in the local Universe¹¹. We derive a rotation curve with the typical shape of nearby massive spiral galaxies, which demonstrates that at least some young galaxies are dynamically akin to those observed in the local Universe, and only weakly affected by extreme physical processes.

SPT–S J041839–4751.9 (hereafter SPT0418–47) has been observed with the Atacama Large Millimeter/submillimeter Array (ALMA) to image the thermal continuum emission from the dust and the emission from the 158- μm fine-structure line of ionized carbon [CII] (Fig. 1a). The [CII] line is typically the brightest fine-structure line emitted in star-forming galaxies, and is a main gas coolant, coming from multiple phases of the inter-stellar medium (ISM), including the warm ionized, the warm and cold neutral atomic, and the dense molecular medium¹². We applied a three-dimensional (3D) lens-kinematic modelling technique¹³ to the interferometric data cube containing

the [CII] line to determine the gas surface brightness distribution in each spectral channel (Fig. 1, 2 and Extended Data Figs. 5, 6 and Table 1; see Methods for full discussion), from which we derive the gas kinematics, and perform a robust dynamical analysis of the galaxy. In addition, we reconstructed the far-infrared surface brightness distribution of the heated dust emission from the interferometric continuum dataset.

We find that the rotation curve of SPT0418–47 has the typical shape of a bulge-dominated spiral galaxy in the local Universe¹¹; it has a bump at 0.2 kpc from the galaxy centre and then declines before flattening at radii larger than 1.5 kpc (Figs. 2a, c). The [CII] velocity dispersion σ is well described by an exponential profile, with average values of $\sim 45 \text{ km s}^{-1}$ in the inner regions ($\lesssim 1 \text{ kpc}$) and $\sim 18 \text{ km s}^{-1}$ in the outer disk ($\gtrsim 1 \text{ kpc}$) (Figs. 2b, d, Extended Data Table 1). From a decomposition of the rotation curve, we derive the relative contribution of the different mass components to the total galactic gravitational potential, the stellar component, the gaseous disk and the dark matter halo (Extended Data Table 5, Fig. 2c; see Methods).

SPT0418–47 has global physical properties (total stellar mass $M_{\text{star}} = 1.2^{+0.2}_{-0.1} \times 10^{10} M_{\odot}$, dark matter mass $M_{\text{DM}} = 1.7^{+0.3}_{-0.3} \times 10^{12} M_{\odot}$, gas fraction $f_{\text{gas}} = 0.53^{+0.06}_{-0.08}$, and stellar effective radius $R_e = 0.22^{+0.04}_{-0.02} \text{ kpc}$; see Extended Data Table 5 for further parameters) that are in agreement with the predictions from the most recent theoretical models⁹, as well as observations of the population of starburst galaxies at this redshift¹⁰. From the de-lensed dust emission, we derive an intrinsic infrared luminosity of $(2.4 \pm 0.4) \times 10^{12} L_{\odot}$, a star formation rate (SFR) of $352 \pm 65 M_{\odot} \text{ yr}^{-1}$ and a gas depletion timescale of $38 \pm 9 \text{ Myr}$, which are all typical of mm-selected starburst galaxies at this redshift¹⁰ (see Methods).

Our high-resolution 3D kinematical analysis shows that SPT0418–47 has a ratio of rotational velocity (V) to velocity dispersion σ of $V/\sigma = 9.7 \pm 0.4$ (Extended Data Table 2), which is similar to that of spiral galaxies in the local Universe¹¹, but considerably larger than what is predicted by most numerical^{1,3} and analytical^{2,4,5} models. For example, for the majority of star-forming galaxies at $z \approx 4$ with stellar masses in the range 10^9 – $10^{10.5} M_{\odot}$, the most recent cosmological magnetohydrodynamical simulation TNG50¹ gives $V/\sigma \lesssim 3$ (light-blue band in Fig. 3); even though such simulated galaxies have rotationally supported gas-rich disks, they are dynamically hotter than their low-redshift counterparts. Complex astrophysical processes, such as stellar feedback by supernovae or radiation pressure, or active galactic nucleus (AGN) feedback from the central supermassive black-hole, galaxy mergers, and gas inflows and outflows, are expected to have a significant impact on the gas kinematics within galaxies at this early epoch, and are predicted to be responsible for a progressive increase of chaotic, random motions with redshift^{1,4,5}. Our finding

firmly rules out models in which a high star-formation feedback and a high gas fraction necessarily produce large turbulent motions^{1,4} (light-blue, green, and black bands in Fig. 3) and violent disk instabilities^{2,3}, resulting in dispersion dominated systems with $V/\sigma \lesssim 2$ at these redshifts (grey band in Fig. 3; see Methods for further discussions about these comparisons). Our result requires galaxy evolution models to produce dynamically cold galaxies that are not characterised by large turbulent motions^{1,4} and violent instabilities^{2,3}, already at early times.

Our kinematic analysis also allows the level of axisymmetric disk instabilities to be measured within SPT0418–47 via the Toomre parameter¹⁴ Q : a value of $Q > 1$ ensures that no instabilities will develop because large-scale collapse is prevented by differential rotation, whereas $Q \lesssim 1$ indicates that instabilities will be able to grow and lead to the formation of gas and star-formation clumps within the disk. For SPT0418–47, we find an average value of $Q = 0.97 \pm 0.06$ at $R > 1$ kpc (Fig. 2d), where the gas component is dominant (Fig. 2c), indicating a potentially unstable disk, prone to form clumpy star-forming regions. This result supports the hypothesis¹⁵ that the irregular morphologies found for dusty starburst galaxies in the optical/ultraviolet rest-frame wavelengths¹⁶ are poorly resolved clumpy star-forming regions, and not objects that are undergoing or have recently experienced a merging event.

Using our rotation curve decomposition (see Methods), we find that the stellar component of SPT0418–47 is well described by a Sérsic profile with a Sérsic index of $2.2^{+0.3}_{-0.2}$ and a stellar mass of $1.2^{+0.2}_{-0.1} \times 10^{10} M_{\odot}$. Several observational studies of the structural properties¹⁷ of galaxies have confirmed that the Hubble sequence is already in place at $z \approx 2.5$, with galaxies showing a large variety of morphologies. However, at $z \gtrsim 3$, the lack of spatially resolved data in the rest-frame optical emission for these galaxies has prevented the study of their structure and morphologies. The unprecedented spatial resolution of 60 pc of the dataset for SPT0418–47 allows us to study, for the first time, the morphological properties of a $z \approx 4$ galaxy. The bump in the inner region of the rotation curve clearly indicates that a bulge is already in place at $z \approx 4$, whereas the Sérsic index of ~ 2 is a signature of either a disk bulge¹⁸ or an embedded disk-like component.

Dusty starburst galaxies are believed to be the progenitors of early-type galaxies (ETGs), which are the most massive galaxies observed today, dominated by old stellar populations. The most popular evolutionary track for this transformation^{2,3,19} predicts that the dusty-starburst phase is followed by a quenching phase, during which AGN feedback leads to gas consumption and heating with the consequent formation of a population of compact quiescent galaxies²⁰ at $z \approx 2$. In the final phase, dry minor mergers are expected to be responsible for a growth in galaxy size and

the transformation into present-day ETGs. In Fig. 4a-d we compare the main physical quantities of SPT0418–47, as inferred from our dynamical model, with the corresponding quantities for the sample of local ETGs from the ATLAS^{3D} survey²¹. We consider only those local ETGs with estimated stellar ages $\gtrsim 12$ Gyr²², which is the lookback time corresponding to $z \approx 4$. The orange thin diamond in each panel shows the values derived for SPT0418–47 as observed today, and the red diamond shows the corresponding baryonic values (gas plus stars), under the assumption that all the gas that we observe today will be converted into stars, preserving the disk configuration. Given the SFR estimated for SPT0418–47, this conversion will happen in ~ 38 Myr (see Methods). The comparison between the ETGs and the stellar/baryonic quantities for SPT0418–47 in the size–mass plane (Fig. 4a) indicates that this starburst galaxy should increase its stellar mass by a factor of 6 (3 for the red diamond) and its effective radius by a factor of 11 (3 for the red diamond), in order to evolve into an average ETG (yellow cross). This is in agreement with a simple toy model²³ for mergers, in which a single dry major merger event would be responsible for an increase in both the size and stellar mass of SPT0418–47 by a factor of 3. In Fig. 4a, the grey stars show the populations of compact galaxies²⁰ at $2 < z < 3$. Notably, the median size for this population is comparable to the size of SPT0418–47 (red diamond), but its median stellar mass is a factor of 3 larger. Also, the position of SPT0418–47 on the mass–size plane is compatible with the low-mass end of both ETGs and $z \approx 2$ compact galaxies. This implies that either this galaxy will evolve smoothly into a low-mass ETG after the consumption and/or heating of the cold gas reservoir or, as predicted by the merger scenario, it will reach the bulk of the ETG population in the size–mass plane. Fig. 4c shows that a potential progenitor of an ETG has already at $z \approx 4$ a disky stellar component, a feature that is very common especially amongst local fast rotator ETGs¹⁸. Finally, we derived the fraction of dark-matter mass within the effective radius and found that, with a value of $f_{\text{DM}}(< R_e) = 0.095^{+0.004}_{-0.004}$ (red diamond), the central regions of SPT0418–47 are dominated by baryons. As shown in Fig. 4c, such a low fraction of dark matter is compatible with observations of local ETGs²⁴, implying that the physical mechanisms responsible for the mass and size growth of this galaxy with cosmic time should preserve the dark-matter contribution within the innermost ~ 1 kpc.

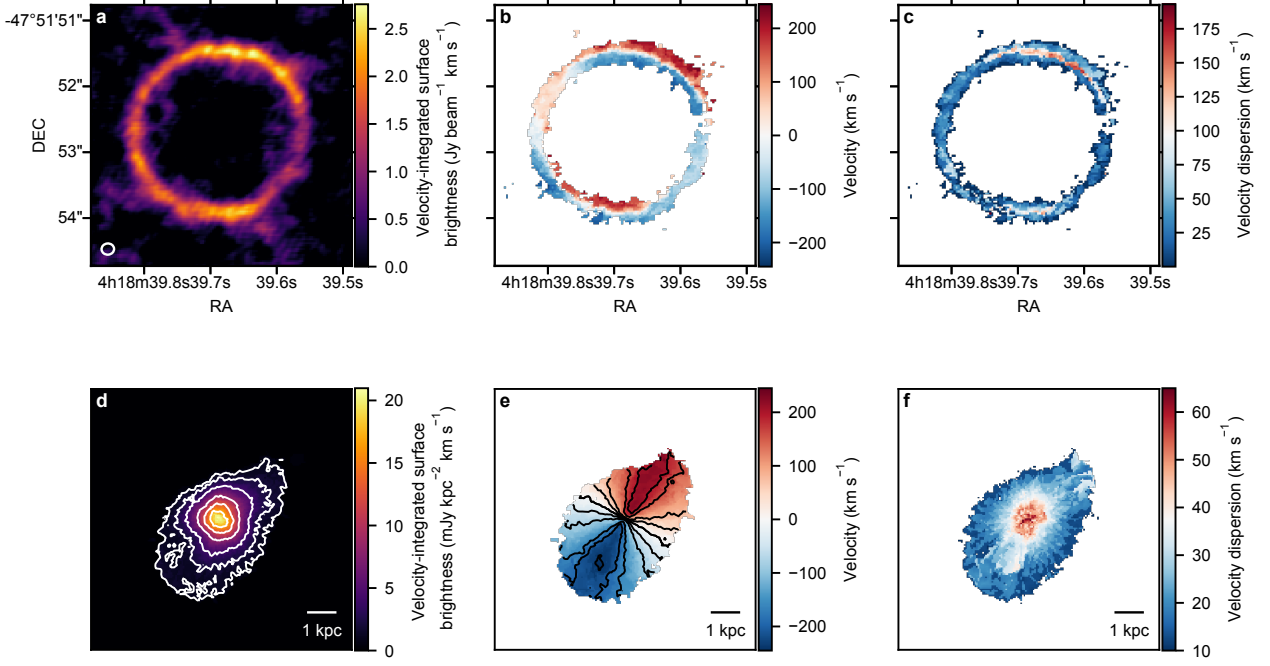


Figure 1: [CII] emission from the lensed galaxy SPT0418-47 and source plane reconstruction. Panel a: emission of the 158- μm fine-structure line of ionized carbon [C II] integrated across a velocity range of 721 km s^{-1} (zeroth-moment map). The beam size, shown as a white ellipse on the lower left corner, is $0.19'' \times 0.17''$ at a position angle of -85.22° . Panels b and c: same as in panel a, but the emission is colour-coded by the flux-weighted velocity and velocity dispersion (first- and second-moment maps) respectively. Panels d-f: zeroth-, first- and second-moment maps of the reconstructed source. In panel d the white contours are set at $n = 0.05, 0.1, 0.2, 0.4, 0.6, 0.8$ times the value of the maximum flux of the zeroth-moment map. In Panel e the black solid contours are at $V_{\text{sys}} \pm \Delta V$, where V_{sys} is the systemic velocity set to 0 km s^{-1} and $\Delta V = 40 \text{ km s}^{-1}$. We note that the scales of the second-moment maps (panels c and f) are different because panel c shows the observed values, while panel f shows the intrinsic ones (beam-smearing corrected). These six maps are intended only for visualisation, as the source reconstruction and its kinematic modelling are performed using the full 3D information of the data cube containing the [CII] emission line (see Methods and Extended Data Fig. 5 for further details).

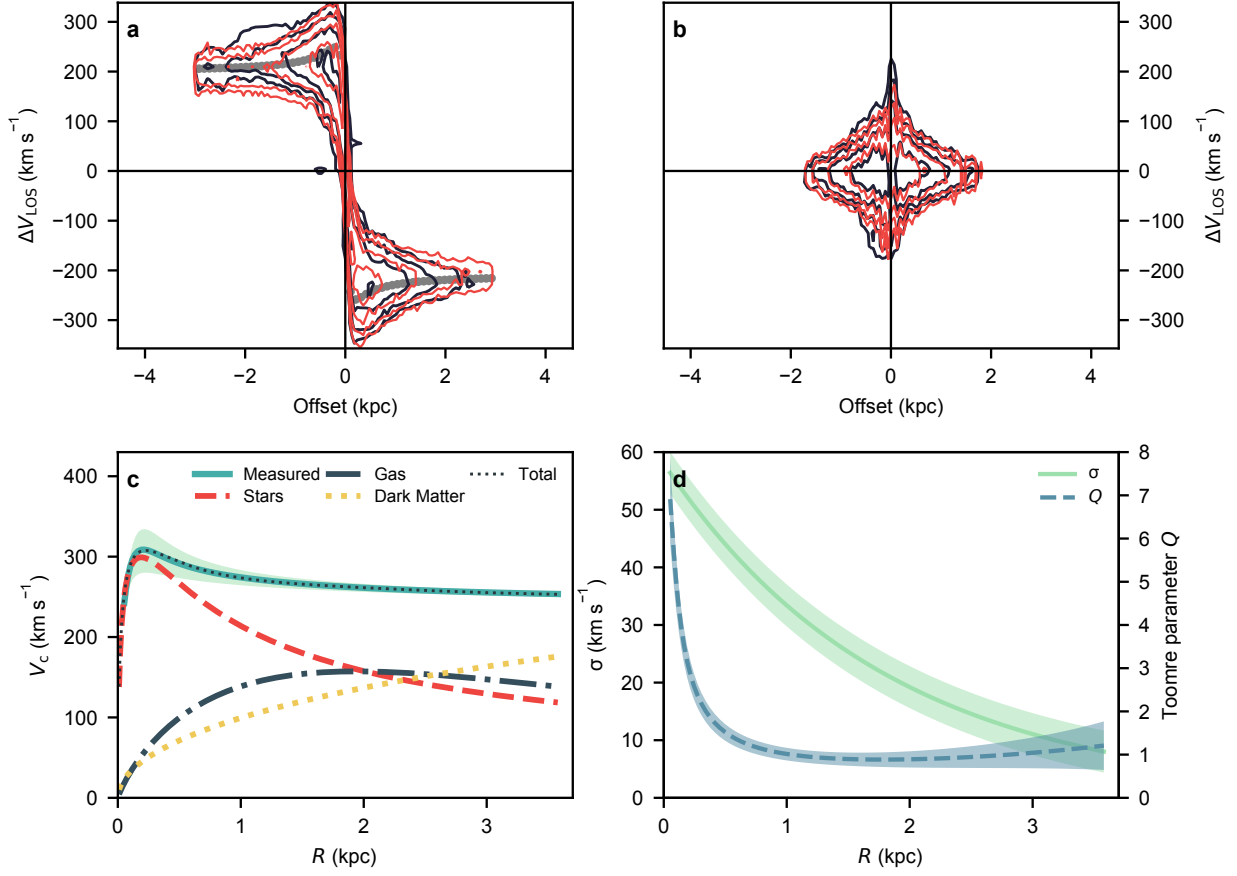


Figure 2: **Kinematic and dynamical properties of SPT0418-47.** Panel a and b: [CII] emission in the position-velocity diagrams along the major and minor axis respectively. These diagrams show slices, the equivalent of putting long slits along the two axes. The x axis shows the offset along the major and minor axis from the galaxy center, and the y axis represents the line-of-sight velocity centred at the systemic velocity of the galaxy. The dark blue contours show the reconstructed source, and the red contours show the best kinematic model (see Methods). The contour levels are set at $n = 0.1, 0.2, 0.4, 0.8$ times the value of the maximum flux in the major-axis position-velocity diagram. The grey circles show the rotation velocities derived using our 3D lens-kinematic methodology. Panel c: rotation curve decomposition. The green solid line shows the circular velocity profile. The black dotted line is the best dynamical model, obtained by fitting the different mass components contributing to the total gravitational potential as shown in the legend. Panel d: velocity dispersion profile (solid green line) and Toomre parameter profile (dotted blue line). The coloured bands in panel c and d represent uncertainties obtained by error propagation from the 1 s.d. uncertainties of the parameters that define the respective profiles (see Methods and Extended Data Table 1).

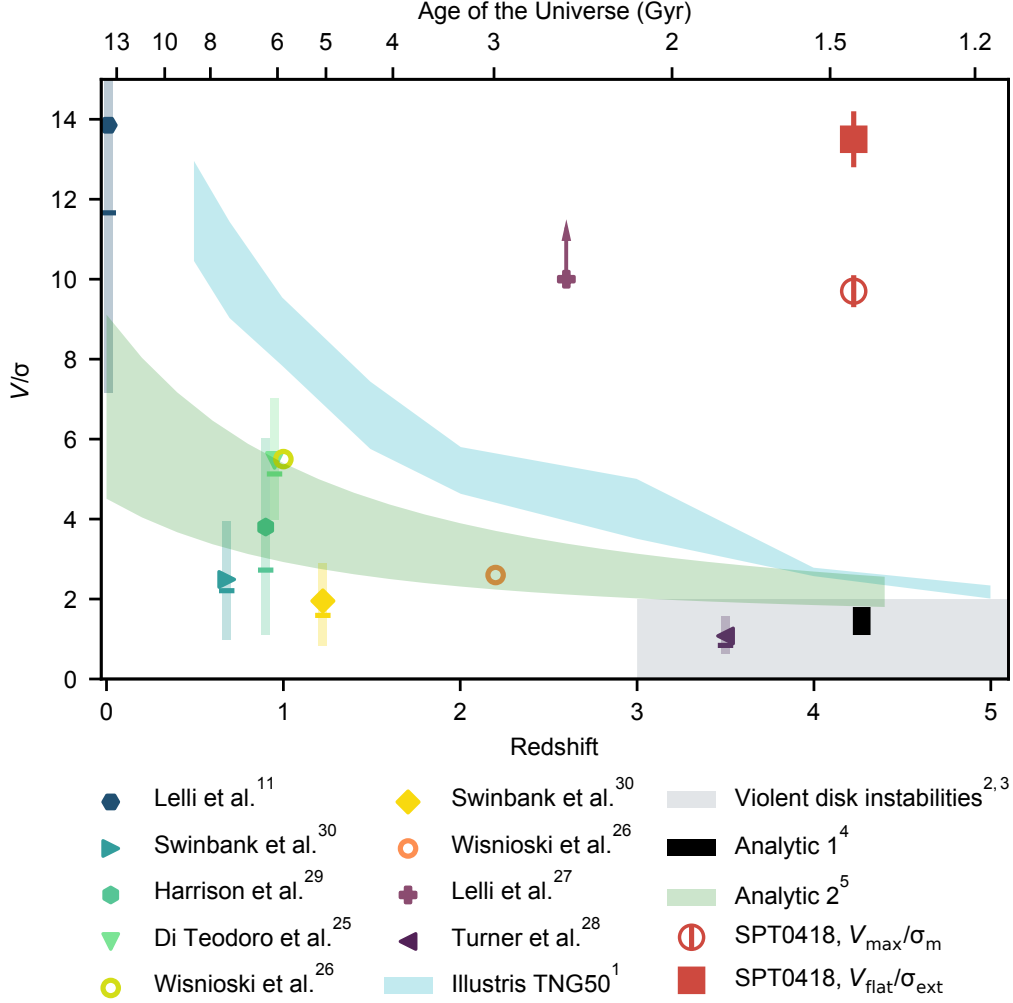


Figure 3: **Comparison between SPT0418-47 and samples of observed and simulated galaxies.** Mean ratios of the rotational to random motions (V/σ) versus redshifts for the comparison samples^{11,25–30} of observed star-forming galaxies indicated in the legend and for SPT0418-47 (red square and red empty circle). The gas tracers are: HI ¹¹, $\text{H}\alpha$ ^{25,26,29,30}, $[\text{OII}]$ ³⁰, $[\text{OIII}]$ ²⁸, $[\text{CI}]$ ²⁷. For SPT0418-47, the V/σ was obtained from the $[\text{CII}]$ emission line. For the comparison samples, the shaded regions show the area between the 16th and 84th percentiles of the distributions, while the horizontal bars show the median values (when available). In Extended Data Table 3 we show the different extraction methods used to compute V/σ . For the empty markers the V/σ values were calculated using for each galaxy a proxy for the maximum rotation velocity, V_{max} . The V/σ values shown by the full markers were calculated using the flat or the outer part of the rotation curve V_{flat} . The violet cross is a lower limit for a single galaxy²⁷. The light-blue band shows the V/σ for simulated galaxies from TNG50 simulation¹ in the mass range $10^9 - 10^{11} M_{\odot}$. For these simulated galaxies, V/σ is calculated as the ratio between the maximum rotation velocity, V_{max} and the mean velocity dispersion σ_{m} . The grey area show the expected V/σ for galaxies dominated by violent disk instabilities^{2,3}. The black and green areas show a prediction and an assumption, respectively, from two different analytical models (Analytic 1⁴ and Analytic 2⁵). The red square is the $V_{\text{max}}/\sigma_{\text{m}}$ for SPT0418-47, and the empty red circle shows the position of the V/σ value for SPT0418-47, calculated as the ratio between V_{flat} and the velocity dispersion at outer radii σ_{ext} (see Extended Data Table 2). The error bars for the red square and empty circles show the 1 s.d. uncertainties.

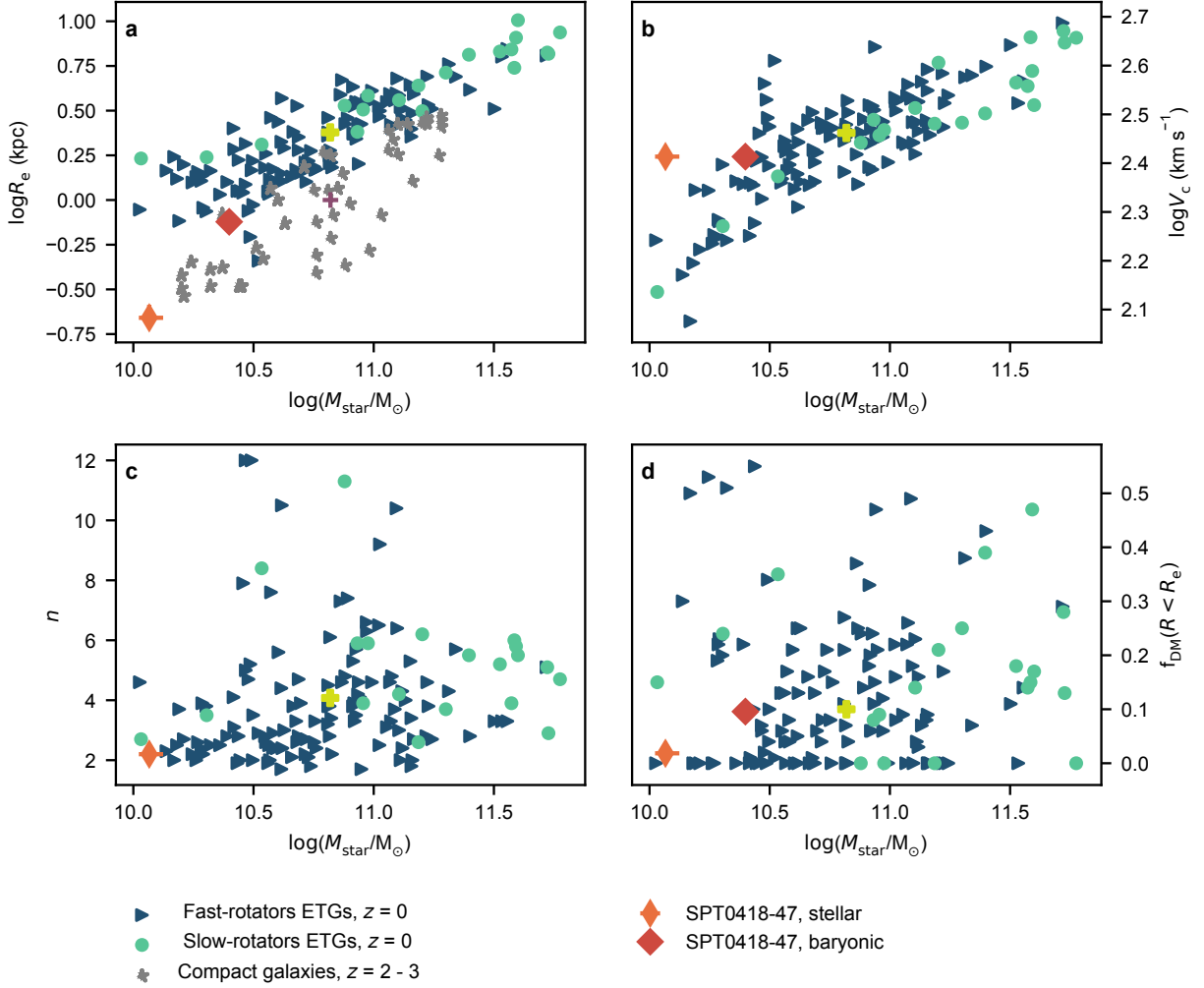


Figure 4: **Comparison between SPT0418-47 and samples of its plausible descendants.** In all panels, the orange thin diamond shows the position of SPT0418-47, and the red diamond shows the baryonic quantities (gas plus stars), under the assumption that all the gas that we observe today will be converted into stars, preserving the disk configuration. The error bars show the 1 s.d. uncertainties. The $z \approx 0$ ETGs (from the ATLAS^{3D} survey^{18,21,22,24}), the plausible descendants of SPT0418-47, are shown according to their kinematic classification: blue triangles show fast rotators, green circles show slow rotators and yellow crosses show the median values. Panel a: size-stellar mass plane. The grey stars show the compact galaxies²⁰ at $z \approx 2-3$ and the violet cross show the median values for this sample. Panel b: circular velocity versus stellar mass. Panel c: Sérsic index versus stellar mass. Panel d: fraction of dark matter within the effective radius versus stellar mass.

Methods

Observations. SPT0418–47 was identified during the South Pole Telescope Survey³¹ as a far-infrared luminous source with dust-like spectral indices at 1.4 and 2.0 mm³². Follow-up observations³² with ALMA confirmed SPT0418–47 to be a dusty starburst galaxy at redshift (z) 4.2248⁶ that is strongly gravitationally lensed by a foreground galaxy at $z = 0.263$.

The data used for the analysis described here were taken from the ALMA Archive. SPT0418–47 was observed with ALMA on 5 July 2017 under project code 2016.1.01499.S (PI: K. Litke). Observations were taken at a central frequency of 358 GHz with an antenna configuration with a maximum baseline of 1400 m. The data were correlated with both linear polarisations (XX and YY), with a visibility integration time of 6 s, in four spectral windows, each with 240 channels and 1.875 GHz bandwidth. The spectral windows were centred on 350.8, 352.7, 362.8 and 364.6 GHz, where the last spectral window covers the redshifted rest frequency of the [CII] line ($\nu_{\text{rest}} = 1900.5369$ GHz). J0519–4546 was observed to calibrate the flux density scale and J0522–3627 was used to correct the spectral bandpass. J0428–5005 was observed as a secondary check source. Phase switching to J0439–4522 was carried out at 7-min intervals to calibrate the complex gains resulting from atmosphere-induced phase and amplitude fluctuations. The total on-source integration time was 21 min.

We calibrated the raw visibility data using the ALMA pipeline in the Common Astronomy Software Applications CASA package³³. The data were then inspected to confirm the quality of the pipeline calibration and that no further flagging was required. Phase-only self-calibration was performed on the continuum with a solution interval of 100 s to correct for residual phase errors. The complex gain corrections from the continuum were applied to the line spectral window. The line data were prepared by fitting a model to the line-free spectral windows and subtracting it from the visibilities, to produce a data set with only the spectral line emission. To improve the signal-to-noise ratio we averaged the data into groups of four velocity channels resulting in twenty-eight independent channels each with a width of 25.7 km s^{-1} .

SPT0418-47 was imaged on a pixel scale of 0.03 arcsec per pixel with natural weighting of the visibilities and deconvolved using CLEAN³⁴. The zeroth-, first- and second- moment maps of the [CII] emission are shown in Figs. 1a-c; we note that these images are intended only for visualisation, as all the modelling and analysis is performed on the visibility data directly. Throughout this study, we assume a Λ CDM cosmology, with Hubble constant $H_0 = 67.8 \text{ km s}^{-1} \text{ Mpc}^{-1}$, matter density $\Omega_m = 0.308$, and vacuum energy density $\Omega_\Lambda = 0.691$ ³⁵.

Lens and kinematic model. We model the data cube containing the [CII] emission line using a 3D Bayesian lens-kinematic modelling technique that fits the data directly in visibility space

(Extended Data Fig. 5). See Rizzo et al.¹³ and Powell et al.³⁶ for a more detailed description of the modelling approach, which is briefly summarized below.

The method used for source reconstruction is grid-based: the background source surface brightness distribution is reconstructed on a triangular adaptive grid³⁷ defined by a Delaunay tessellation. The source grid automatically adapts with the lensing magnification, so that there is a high pixel density in the high-magnification regions, resulting in a spatial resolution of ~ 40 pc in the inner regions and ~ 90 pc in the outer regions for this system. The kinematics of the background galaxy are obtained by fitting the lensed data directly in a hierarchical Bayesian fashion, where a 3D kinematic model, \mathbf{s}_{kin} , that describes a rotating disk is used as a regularizing prior for a pixellated source reconstruction.

Taking advantage of the fact that gravitational lensing conserves surface brightness, and considering the observational noise \mathbf{n} , the data in the visibility space \mathbf{d} and the source surface brightness distribution \mathbf{s} can be related to each other via the following set of linear equations:

$$\text{DL}(\boldsymbol{\eta}_{\text{lens}})\mathbf{s} + \mathbf{n} = \mathbf{d} \quad (1)$$

where L is the lensing operator, which is a function of the lensing parameters $\boldsymbol{\eta}_{\text{lens}}$, while D is the non-uniform discrete Fourier transform operator. Because both $\boldsymbol{\eta}_{\text{lens}}$ and \mathbf{s} are unknown, equations (1) can not be simply inverted. We herefore fit the data within the framework of Bayesian statistics by using two levels of inference.

At the first level of inference, it can be shown that given the data, a lens-mass model and a source kinematic model, the most probable a posteriori source \mathbf{s}_{MP} is obtained by solving the following set of linear equations:

$$[(\text{DL})^T \text{C}^{-1} \text{DL} + \lambda \text{R}^T \text{R}] \mathbf{s} = (\text{DL})^T \text{C}^{-1} \mathbf{d} + \lambda \text{R}^T \text{R} \mathbf{s}_{\text{kin}} \quad (2)$$

where C^{-1} is the noise covariance and R is the source regularization form with regularization strength λ .

At the second level of inference, the lens parameters $\boldsymbol{\eta}_{\text{lens}}$, the regularization strength λ and the kinematic parameters $\boldsymbol{\eta}_{\text{kin}}$ defining \mathbf{s}_{kin} are obtained by maximizing the posterior probability distribution $P(\mathbf{d}|\boldsymbol{\eta}_{\text{lens}}, \boldsymbol{\eta}_{\text{kin}}, \lambda)$ defined as:

$$2 \log P(\mathbf{d}|\boldsymbol{\eta}_{\text{lens}}, \boldsymbol{\eta}_{\text{kin}}, \lambda) = -\chi^2 - \lambda (\mathbf{s}_{\text{MP}} - \mathbf{s}_{\text{kin}})^T \text{R}^T \text{R} (\mathbf{s}_{\text{MP}} - \mathbf{s}_{\text{kin}}) - \log \det \text{A} \\ + \log \det(\lambda \text{R}^T \text{R}) + \log \det(2\pi \text{C}^{-1}). \quad (3)$$

In equation (3), $\chi^2 = (\text{DLs} - \mathbf{d})^T \text{C}^{-1} (\text{DLs} - \mathbf{d})$ and the matrix A is defined as $[(\text{DL})^T \text{C}^{-1} \text{DL} + \lambda \text{R}^T \text{R}]$. Because of the large number of visibilities per channel, $n_{\text{vis},i} = 163190$, we replace the non-uniform discrete Fourier transform operator D with a non-uniform Fast Fourier Transform (FFT) operator^{38,39}

(NUFFT), which first uses a gridding kernel to interpolate the visibilities onto a regular gridded uv-plane, then applies an FFT, and finally uses an apodization correction to remap the model and the data onto the original uv-sampling. Because the non-uniform FFT does not have an explicit matrix representation, the source inversion, equation (2), is solved using a preconditioned conjugate gradient solver. We note that for the conjugate gradient method we use a tolerance of 10^{-12} , which results in typical accuracies of 10^{-6} for the source inversions.

As usually done when modelling galaxy-galaxy strong lensing observations, the lens is described by a projected mass density profile with a cored elliptical power-law distribution plus the contribution of an external shear component of strength Γ_{sh} and position angle θ_{sh} (see Extended Data Table 1). This assumption has been shown to provide a good fit to large samples of known lenses, as discussed, for example, by Koopmans et al.⁴⁰ and Barnabé et al.⁴¹. The dimensionless projected mass density profile is defined as

$$\kappa(x, y) = \frac{\kappa_0 \left(2 - \frac{\gamma}{2}\right) q^{\gamma - \frac{3}{2}}}{2 \left[q^2 (x^2 + r_c^2) + y^2 \right]^{\frac{\gamma-1}{2}}}, \quad (4)$$

where κ_0 is the normalization, q is the projected flattening, γ is the density slope (a value of $\gamma = 2$ corresponds to an isothermal profile), x_0 and y_0 define the center of the mass distribution, r_c is the core radius and θ is the position angle of the major axis. The lens mass model parameters are in agreement within 2 s.d. with those derived from modelling the dust continuum⁴².

The kinematic model \mathbf{s}_{kin} is defined by the parameters describing the rotation velocity, the velocity dispersion profile and the geometry of the galaxy. We choose the multi-parameter function

$$V_{\text{rot}}(R) = V_t \frac{\left(1 + \frac{R_t}{R}\right)^\beta}{\left[1 + \left(\frac{R_t}{R}\right)^\xi\right]^{1/\xi}}, \quad (5)$$

as the functional form for the rotation velocity (see Extended Data Table 1). In equation (5) V_t is the velocity scale, R_t is the turnover radius between the inner rising and outer part of the curve, β specifies the power-law behaviour of the curve at large radii and ξ defines the sharpness of the turnover. We choose the above multi-parameter function because it is flexible enough to reproduce the large variety of observed rotation curves, and therefore allows much more freedom than other typically used functions, for example the arctangent.

The velocity dispersion profile, $\sigma(R)$ is described by an exponential function, which is more flexible than the more commonplace choice of a constant value^{26,28}. We also tested a linear function

but found it to be significantly less favoured by the data, by a Bayes factor of 1.8 relative to the exponential model.

For the geometry of the kinematic model, defined by the inclination (i) and position angle (PA), we assume that there is no radial variation (Extended Data Table 1). The surface density of the gas is not a free parameter; instead, we impose a pixel-by-pixel normalization, which is given by the surface brightness distribution obtained from the lens modelling of the zeroth-moment map. The advantage of using this normalization is that it allows us to account for possible asymmetries in the gas distribution. The derivation of the lens mass model and the source kinematics is done using a four-step optimisation scheme, and the uncertainties on the parameters are obtained from the posterior distributions calculated with MULTINEST⁴³ (Extended Data Fig. 6), by adopting the user-defined tolerance, sampling efficiency and live points of 0.5, 0.8 and 200. We then verified that the evidences estimated by MULTINEST flatten as a function of the prior volume. As described in our method paper¹³, for each parameter, we adopt priors that are flat in the intervals $[\bar{\eta}_{\text{lens/kin}} - 0.2\bar{\eta}_{\text{lens/kin}}, \bar{\eta}_{\text{lens/kin}} + 0.2\bar{\eta}_{\text{lens/kin}}]$, where $\bar{\eta}_{\text{lens/kin}}$ are the best-fitting parameters, inferred from the nonlinear optimization. To be conservative, we report as errors in the parameters the sum in quadrature of the following two contributions: the 1 s.d. uncertainty on the posterior distributions derived by MULTINEST and the difference between the maximum a posteriori parameter values obtained by MULTINEST and by the non-linear optimizer. These values of the uncertainties are in line with what expected from the tests presented in Rizzo et al.¹³, where we applied our methodology to data characterized by spatial and spectral resolutions of at least a factor of three worse than those analysed in this manuscript.

Direct Forward modelling For a consistency check, we also derived the lens (η_{lens}) and kinematic parameters (η_{kin}) by using a direct forward modelling methodology. The main difference to the methodology described in Rizzo et al.¹³ and summarized above is that a parametric, rather than a pixellated source is used to fit the data. The parametric source is defined by the same 3D kinematic model, \mathbf{s}_{kin} , used as a regularizing prior for the source in our fiducial method. In this implementation, to marginalize over η_{lens} and η_{kin} , we maximize the posterior probability distribution (see equation (3) for comparison), defined as $2 \log P(\mathbf{d}|\eta_{\text{lens}}, \eta_{\text{kin}}) = -(\text{DLs}_{\text{kin}} - \mathbf{d})^T \mathbf{C}^{-1} (\text{DLs}_{\text{kin}} - \mathbf{d})$. The parameters obtained with the forward modelling method are within the 1 s.d. uncertainties with respect to those obtained with our fiducial model (see Extended Data Fig. 6).

In the direct forward modelling method, the source is forced to have a regular configuration, chosen a priori using the assumptions on the kinematic parametrizations. As explained in detail in Rizzo et al.¹³, by assuming our fiducial methodology, we instead use the kinematic model as a prior for the source reconstruction. The main advantage of the fiducial approach is that it does not directly

impose a parametric model which may not be a good fit to the data.

Beam smearing correction. In recent years, observational studies have used the V/σ ratio to study the kinematic evolution of galaxy populations and distinguish between rotation- and dispersion-dominated systems. However, the measurement of the V/σ ratio is a challenge for high-redshift, marginally resolved studies because of the so-called beam-smearing effect⁴⁴. This effect can cause a strong overestimation of the velocity dispersion and underestimation of the rotation velocity. For galaxies in the redshift range 1 to 3, there is no consensus whether or not there is an increase of the V/σ ratio with time^{25–30}. For galaxies at $z \gtrsim 4$ there have been very few attempts to go beyond the fitting of the integrated line profile^{45,46}. However, marginally resolved observations and a low signal-to-noise ratio have prevented a measurement of a beam-smearing-corrected V/σ ratio. The combination of the high signal-to-noise ratio and the spatial resolution that is achievable thanks to gravitational lensing and to an intrinsically beam-smearing corrected kinematic analysis has allowed us to probe the intrinsic V/σ ratio for the first time at this high redshift. Different definitions of the V/σ ratio can be found in literature. Here we use of the two most popular ones (see Extended Data Table 2). We obtain $V/\sigma = 9.7 \pm 0.4$ when we calculate V/σ as the ratio between the maximum rotation velocity V_{\max} and the mean velocity dispersion σ_{m} , whereas we get $V/\sigma = 13.7 \pm 0.7$, when we calculate it as the ratio between the flat part of the rotation velocity V_{flat} and the velocity dispersion at outer radii ($R > 1$ kpc) σ_{ext} . The values of V/σ for SPT0418–47 seem to indicate that there is no decrease of V/σ with time. However, larger samples will be needed to confirm this trend.

V/σ for the comparison samples. To determine the evolution of the dynamical properties of galaxies across cosmic time, different studies have used different gas tracers (see column 2 of Extended Data Table 3) and different extraction methods to obtain their physical parameters. The comparison of V/σ for SPT0418–47 with values derived from different gas tracers is based on the assumption that the ionized gas kinematics reflects the cold gas motions of galaxies, as observed for a few galaxies^{47,48}. However, by comparing the results obtained from different tracers and with different techniques, a recent study⁴⁹ found that velocity dispersions measured from ionized tracers tend to be $\sim 10 - 15 \text{ km s}^{-1}$ higher than those measured from molecular tracers. If we add this 15 km s^{-1} to our inferred [CII] velocity dispersions, we find $V_{\max}/\sigma_{\text{m}} = 6.6$ and $V_{\text{flat}}/\sigma_{\text{ext}} = 7.8$, which are still much higher than those observed and those predicted by galaxy evolution models. We also stress that whereas the above analysis mostly compared the molecular and the ionized gas tracers, in this study we focus on the [CII] emission, which traces both cold and warm gas. In Extended Data Table 3, we show the main assumptions made to extract the velocities and velocity dispersions that enter the calculation of the V/σ ratio for the samples shown in Fig. 3.

V/σ in galaxy evolution models. There is a general consensus^{50–53} that galaxy disks at high redshift are much more turbulent than their local counterparts in all components (stars, warm and cold gas).

In Fig. 3 we show the V/σ predicted (or assumed) by current galaxy evolution models:

- The light-blue area shows the predicted evolution of V/σ in the TNG50 simulation¹, where the gas kinematics is derived from H α emitting gas. However, because, by construction, the ionized and molecular gas of TNG50 galaxies have the same dynamics, we can compare these estimations with our measurement coming from the [CII] emission line. TNG50 galaxies at $z \sim 4$ have median $V_{\text{max}}/\sigma_{\text{m}} \sim 3$, with a standard deviation of 1.5, which means that SPT0418–47 is 4.5 standard deviations away from the median value.
- By using an analytic approach and a cosmological mesh-refinement simulation, two recent studies^{2,3} found that galaxies at $z \gtrsim 3$ are dominated by violent disk instabilities, which lead to V/σ values $\lesssim 2$ in all components (grey area in Fig. 3).
- By using an analytic model that combines stellar feedback and gravitational instabilities, Krumholz et al.⁴ derived a prediction for the correlation between cold gas velocity dispersion and SFR, such that $\text{SFR} = (0.42 f_{\text{eff,gas}} V^2 \sigma) / (\pi G Q_{\text{min}})$ (their equation 60), where we have used the constants appropriate for high-redshift galaxies^{4,49}. By using $f_{\text{eff,gas}} = f_{\text{gas}} \times 1.5 = 0.53 \times 1.5$, $\text{SFR} = 352 M_{\odot}/\text{yr}$, $V = 259 \text{ km s}^{-1}$, $Q_{\text{min}} = 0.88$, we derived a value of $\sigma = 183_{-42}^{+53} \text{ km s}^{-1}$, which is a factor of ~ 6 higher than our measured value of 32 km s^{-1} , and implies $V/\sigma = 1.4_{-0.3}^{+0.4}$ (black band in Fig. 3).
- Hayward and Hopkins⁵ used an analytic model to study the effects of stellar feedback in regulating star formation and driving outflow. In this case, the gas velocity dispersion is not a predicted quantity, but it is derived by assuming the following relation between σ and the circular velocity: $\sigma \sim f_{\text{gas}} V_c / \sqrt{2}$. For the values of the gas fraction and V_c measured for SPT0418–47 by our analysis, this analytical model implies a value of $\sigma = 120 \text{ km s}^{-1}$, which is a factor of ~ 4 higher than the measured value, resulting in a $V/\sigma = 2.6$. In Fig. 3 (green band), we show the predictions of V/σ for this model, obtained by using the relations of ref.⁵ for velocity versus stellar mass and the redshift evolution of velocity dispersion with respect to the gas fraction.

Dynamical model and Toomre parameter. Under the assumption that the total galactic gravitational potential Φ is axisymmetric, the rotation velocity $V_{\text{rot}}(R)$ of the gas, in cylindric coordinates

(R, ϕ, z) , is related to Φ by the equation $R(\partial\Phi/\partial R)_{z=0} = V_c^2 = V_{\text{rot}}^2 + V_A^2$, where V_c is the circular velocity, and V_A is the asymmetric-drift correction that accounts for the pressure support due to the random motions. Under the assumptions that the gas of the rotating disk has a thickness that is independent of the radius⁵⁴, it is thin and it has a spatial distribution described by an exponential profile, $\Sigma_{\text{gas}} = \Sigma_0 \exp(-R/R_{\text{gas}})$, the expression for V_A is given by

$$V_A^2 = -R\sigma^2 \partial \ln(\sigma^2 \Sigma_{\text{gas}}) / \partial R = -R\sigma^2 \partial \ln(\sigma^2 \exp(-R/R_{\text{gas}})) / \partial R. \quad (6)$$

To measure the scale radius R_{gas} we divide the zeroth-moment map of the reconstructed source (Fig. 1d) into rings (with centres, PA and i defined by the values of the kinematic model; see Extended Data Table 1) and we calculate the surface densities at a certain radius as azimuthal averages inside that ring. The surface density profile obtained in this way is then fitted using the exponential profile, resulting in a value of $R_{\text{gas}} = 0.9$ kpc. The resulting asymmetric-drift correction, equation (6), gives a small contribution ($\lesssim 1\%$) with respect to $V_{\text{rot}}(R)$.

To derive the contribution of the gas, stellar and dark-matter components to the total gravitational potential, we model the circular velocity as $V_c = \sqrt{V_{\text{star}}^2 + V_{\text{gas}}^2 + V_{\text{DM}}^2}$. Here V_{star} is the stellar contribution to the circular velocity, under the assumption that this component is described by a Sérsic profile^{55,56}, defined by the total stellar mass M_{star} , the effective radius R_{eff} and the Sérsic index n . V_{gas} is the gas contribution, under the assumption that the gas in this galaxy has a distribution described by an exponential profile, as traced by the [CII] emission line. The scale length that enters in V_{gas} is fixed at the value R_{gas} found in the previous paragraph, and the only free parameter of the fit for V_{gas} is the conversion factor ($\alpha_{[\text{CII}]}$) between the total [CII] luminosity and the total gas mass. A number of recent studies have found that [CII] is a good tracer of the total gas mass^{57,58}. For SPT0418–47, the [CII] luminosity is $1.8 \times 10^9 L_{\odot}$, obtained by computing the zeroth-moment map of the [CII] emission as the signal integrated along the spectral axis at each pixel of the reconstructed source. The dark-matter contribution V_{DM} is modelled as a Navarro-Frenk-White⁵⁹ (NFW) spherical halo, with a concentration parameter of $c = 3.06$. The latter is obtained by averaging the values of c at $z = 4.22$ for dark matter haloes with masses between $10^{10} M_{\odot}$ and $10^{13} M_{\odot}$, assuming the mass-concentration relation estimated in N-body cosmological simulations⁶⁰. We notice that at this redshift, c is almost independent of the dark-matter halo mass, varying by just 6 per cent for a variation of 3 orders of magnitude in the halo mass. To test the effect of our assumption of a constant concentration, we repeated the analysis with a c that is free to vary according to either a uniform prior or a Gaussian prior centred on the predicted mass-concentration relation, and found that all inferred dynamical parameters do not change significantly with c .

We summarize all our assumptions for the dynamical model in the second column of Extended

Data Table 4, and the corresponding free parameters are shown in the third column of the same table. We compute the Bayesian posterior distribution of these parameters using DYNESTY, a python implementation of the Dynamic Nested Sampling algorithm⁶¹ (see Extended Data Fig. 7 and Table 5). We use log-uniform priors for the mass parameters and uniform priors for the scale radii (fourth column of Extended Data Table 4). The inferred $M_{\text{star}} = 1.2^{+0.2}_{-0.1} \times 10^{10} M_{\odot}$ is in excellent agreement with the value of $(9.5 \pm 3.0) \times 10^9 M_{\odot}$ found in a recent independent study⁷ by fitting the spectral energy distribution of this galaxy. For $\alpha_{\text{[CII]}}$ we employ a uniform prior in the range corresponding to ± 3 standard deviations around the mean value of $30 M_{\odot}/L_{\odot}$, derived from a large sample of low- and high-redshift galaxies⁵⁷. We infer a value of $\alpha_{\text{[CII]}} = 7.3^{+1.0}_{-1.2} M_{\odot}/L_{\odot}$, in agreement with other studies of $z \sim 4$ dusty starburst galaxies⁵⁸.

In Extended Data Table 5 we show some relevant physical quantities (see also Fig. 4, panel d) derived from our dynamical analysis. In particular, for each of these we quote the 16th, 50th and 84th percentile, which were obtained from the full posterior sample points returned by DYNESTY.

We also calculate the Toomre parameter¹⁴ (Fig. 2d), using the general definition $Q(R) = (\sigma\kappa)/(\pi G\Sigma_{\text{gas}})$. Here, κ is the epicycle frequency defined as $\kappa = \sqrt{Rd\Omega^2/dR + 4\Omega^2}$, where $\Omega = V_{\text{rot}}/R$ is the angular frequency.

Comparison with local ETGs. In Fig. 4 we show some physical quantities for local ETGs taken from multiple studies of the ATLAS^{3D} survey^{18,21,22,24,62,63}. For comparison with SPT0418–47, we show for each quantity an upper limit, calculated under the assumption that all gas is converted into stars, and the observed quantities (orange diamonds), which can also be interpreted as a lower limit if all the gas that we observe will be ejected by feedback processes. However, we caution that this upper limit does not take into account the accretion of gas if no baryons are expelled by outflows. In this case, the red diamond could move further towards the upper right corner in the size-mass plane (Fig. 4a).

Dust continuum and SFR. Using the parameters of the lens mass model shown in Extended Data Table 1, we perform a pixellated source reconstruction of the dust continuum at $160 \mu\text{m}$ (rest frame), which allows us to calculate a total magnification factor μ of 32.3 ± 2.5 . SPT0418–47 has an observed (unlensed) infrared luminosity of $L_{\text{IR,obs}} = (7.7 \pm 1.3) \times 10^{13} L_{\odot}$ ⁶⁴. This value was obtained from a spectral energy distribution fitting of seven photometric data points in the wavelength rest frame range between 48 and $574 \mu\text{m}$. Under the assumption that the morphology of the source is the same for all photometric points in the infrared band, we can use the

magnification factor estimated from the emission at $160 \mu\text{m}$ to estimate an intrinsic luminosity of $L_{\text{IR}} = L_{\text{IR,obs}}/\mu = (2.4 \pm 0.4) \times 10^{12} L_{\odot}$. By assuming that the infrared emission comes from the thermal emission of dust, heated by the radiation field coming from young stars, we compute $\text{SFR} = (352 \pm 65) M_{\odot} \text{yr}^{-1}$. This value is derived from L_{IR} , by applying a conversion factor of $1.48 \times 10^{-10} M_{\odot} (\text{yr } L_{\odot})^{-1}$, which is valid for a Kroupa Initial Mass Function⁶⁵. The gas-depletion timescale is $M_{\text{gas}}/\text{SFR} = 38 \pm 9 \text{ Myr}$.

Velocity dispersion in starburst galaxies. By comparing the gas kinetic energy and the energy injected by stellar feedback, we determine whether the turbulence driven by supernova explosions can explain the velocity dispersion measured for SPT0418–47. The gas kinetic energy is given by^{66,67}: $E_{\text{kin}} = 3/2 M_{\text{gas}} \sigma_{\text{m}}^2 \sim 5 \times 10^{56}$ (see Extended Data Table 5). We estimate the energy injected by supernova explosions as^{66,67} $E_{\text{SNe}} = \eta_{\text{SN}} \text{SFR } \epsilon_{\text{SN}} E_{\text{SN}} t_{\text{D}}$. Here, η_{SN} is the supernova rate, ϵ_{SN} is the SFR efficiency (that is, the fraction of the supernova energy, $E_{\text{SN}} = 10^{51}$ erg, converted into turbulence) and t_{D} is the dissipation rate of the turbulence. We assume $\eta_{\text{SN}} = 0.01 M_{\odot}^{-1}$ ⁶⁶. For the highly uncertain parameter ϵ_{SN} we assume an upper limit value of 0.1, based on high-redshift observations⁶⁸. t_{D} is calculated as $t_{\text{D}} = h/\sigma_{\text{m}}$ where h is the scale height of the gas, assumed to be ~ 500 pc. Under this assumption, the estimate E_{SNe} is $\sim 5 \times 10^{56}$ erg, in agreement with E_{kin} , suggesting that the velocity dispersion measured for SPT0418–47 can be explained as being produced by turbulence driven by stellar feedback. Furthermore, we note that starburst galaxies both at low^{69,70} and high redshifts²⁷ have values of velocity dispersions, measured from molecular or neutral gas, that are comparable with those found for SPT0418–47.

Table 1: Lens and source kinematic parameters. The lens parameters describe a projected mass density profile with a cored elliptical power-law distribution, equation (4), plus the contribution of an external shear component ($\Gamma_{\text{sh}}, \theta_{\text{sh}}$). The kinematic parameters describe a rotating disk with a rotation curve defined by a multi-parameter function, equation (5), a velocity dispersion profile defined by an exponential function (σ_0, R_0), and a geometry defined by the inclination (i) and the position angles (PA).

Lens parameters		Kinematic parameters	
κ_0 (arcsec)	1.22 ± 0.04	i ($^\circ$)	54 ± 2
θ ($^\circ$)	22.6 ± 1.6	PA ($^\circ$)	-27 ± 3
q	0.91 ± 0.02	V_t (km s^{-1})	245.1 ± 0.8
γ	2.05 ± 0.09	R_t (kpc)	0.14 ± 0.03
Γ_{sh}	0.0082 ± 0.0003	β	0.80 ± 0.02
θ_{sh} ($^\circ$)	29.9 ± 3.0	ξ	2.0 ± 0.4
		σ_0 (km s^{-1})	58.1 ± 1.8
		R_0 (kpc)	1.7 ± 0.1

Table 2: Kinematic properties for SPT0418-47 derived under different assumptions. Parameters representing the rotation velocities and velocity dispersion profile, as well as the rotation support for this galaxy, calculated under different definitions. ^aMaximum rotation velocity. ^bMean velocity dispersion. ^cRotation-to-random-motions ratio, calculated from V_{max} and σ_{m} . ^dFlat rotation velocity, calculated using the flat part of the rotation curve ($R > 2$ kpc, see Fig. 2c). ^eVelocity dispersion at outer radii ($R > 1$ kpc). ^fRotation-to-random-motion ratio, calculated from V_{flat} and σ_{ext} .

Global kinematic properties			
^a V_{max} (km s^{-1})	308 ± 4	^d V_{flat} (km s^{-1})	259 ± 1
^b σ_{m} (km s^{-1})	32 ± 1	^e σ_{ext} (km s^{-1})	18 ± 1
^c $V_{\text{max}}/\sigma_{\text{m}}$	9.7 ± 0.4	^f $V_{\text{flat}}/\sigma_{\text{ext}}$	13.5 ± 0.7

Table 3: **Kinematic measurements for the comparison samples.** ^aThe value of σ are not provided and we assume a value of 10 km s^{-1} , typical of HI in local spiral galaxies¹¹. ^b $V_{\text{PV,max}}$ and $V_{\text{PV,min}}$ are the maximum and minimum velocity along the position-velocity diagram.

Study	Tracer	V	σ
Lelli et al. ¹¹	HI	Flat part of the rotation curve (V_{flat})	^a 10 km/s
Swinbank et al. ³⁰	H α , [OII]	Extracted at $3R_{\text{disk}}$	Average (σ_{m})
Harrison et al. ²⁹	H α	Extracted at $3.4R_{\text{disk}}$	Median value at $R > 3.4 R_{\text{disk}}$
Di Teodoro et al. ²⁵	H α	Flat part of the rotation curve (V_{flat})	Average (σ_{m})
Wisnioski et al. ²⁶	H α	$(V_{\text{PV,max}} - V_{\text{PV,min}})/2^b$	Average from outer regions (σ_{ext})
Lelli et al. ²⁷	[CI]	Flat part of the rotation curve (V_{flat})	Average (σ_{m})
Turner et al. ²⁸	[OIII]	Extracted at $3.4R_{\text{disk}}$	Median

Table 4: **Assumptions for the dynamical fit**

Component	Density profile	Free parameters	Priors
Stars	Sérsic	M_{star}	$[10^7, 10^{11}] M_{\odot}$
		R_{e}	$[0.04, 2.0] \text{ kpc}$
		n	$[0.5, 10]$
Gas	Exponential	$\alpha_{\text{[CII]}}$	$[3.8, 238.0] M_{\odot}/L_{\odot}$
DM	NFW	M_{DM}	$[10^{10}, 10^{13}] M_{\odot}$

Table 5: Physical quantities for SPT0418-47 derived from the kinematic and dynamical modelling.

Left: parameters inferred from a dynamical fit to the circular velocity. The stellar component is described by a Sérsic profile, the gas disk by an exponential profile and the dark matter is defined by a NFW profile. Right: the quantities in square brackets are calculated considering the gas component, under the assumption that all the gas that we observe today will be converted into stars, preserving the disk configuration. ^aTotal gas mass, computed as $M_{\text{gas}} = \alpha_{[\text{CII}]} L_{[\text{CII}]}$. ^bTotal baryonic mass, computed as $M_{\text{bar}} = M_{\text{star}} + M_{\text{gas}}$. ^cBaryonic half mass radius. ^dDark matter fraction within the half mass radius. ^eStellar-to-halo mass ratio: $M_{\text{star}}/M_{\text{DM}}$ [$(M_{\text{star}} + M_{\text{gas}})/M_{\text{DM}}$]. ^fGas fraction: $M_{\text{gas}}/(M_{\text{star}} + M_{\text{gas}})$. ^gVirial velocity of the dark-matter halo. ^hVirial radius of the dark-matter halo. ⁱMaximum velocity for an NFW halo, computed as $V_{200} \sqrt{0.216 c / (\ln(1 + c) - c / (1 + c))}$.

Parameters of the dynamical model		Derived parameters	
$M_{\text{star}} (10^{10} M_{\odot})$	$1.2^{+0.2}_{-0.1}$	^a $M_{\text{gas}} (10^{10} M_{\odot})$	$1.3^{+0.2}_{-0.2}$
$R_{\text{e}} (\text{kpc})$	$0.22^{+0.04}_{-0.02}$	^b $M_{\text{bar}} (10^{10} M_{\odot})$	$2.5^{+0.2}_{-0.1}$
n	$2.2^{+0.3}_{-0.2}$	^c $R_{\text{e,bar}} (\text{kpc})$	$[0.75^{+0.06}_{-0.06}]$
$M_{\text{DM}} (10^{12} M_{\odot})$	$1.7^{+0.3}_{-0.3}$	^d $f_{\text{DM}}(< R_{\text{e}})$	$0.018^{+0.005}_{-0.003} [0.095^{+0.008}_{-0.007}]$
$\alpha_{[\text{CII}]} (M_{\odot}/L_{\odot})$	$7.3^{+1.0}_{-1.2}$	^e $f_{\star} (10^{-3})$	$7.1^{+1.0}_{-0.8} [14.9^{+3.7}_{-2.6}]$
		^f f_{gas}	$0.53^{+0.06}_{-0.08}$
		^g $V_{200} (\text{km s}^{-1})$	320^{+17}_{-18}
		^h $R_{200} (\text{kpc})$	70^{+4}_{-4}
		ⁱ $V_{\text{max}} (\text{km s}^{-1})$	323^{+18}_{-19}

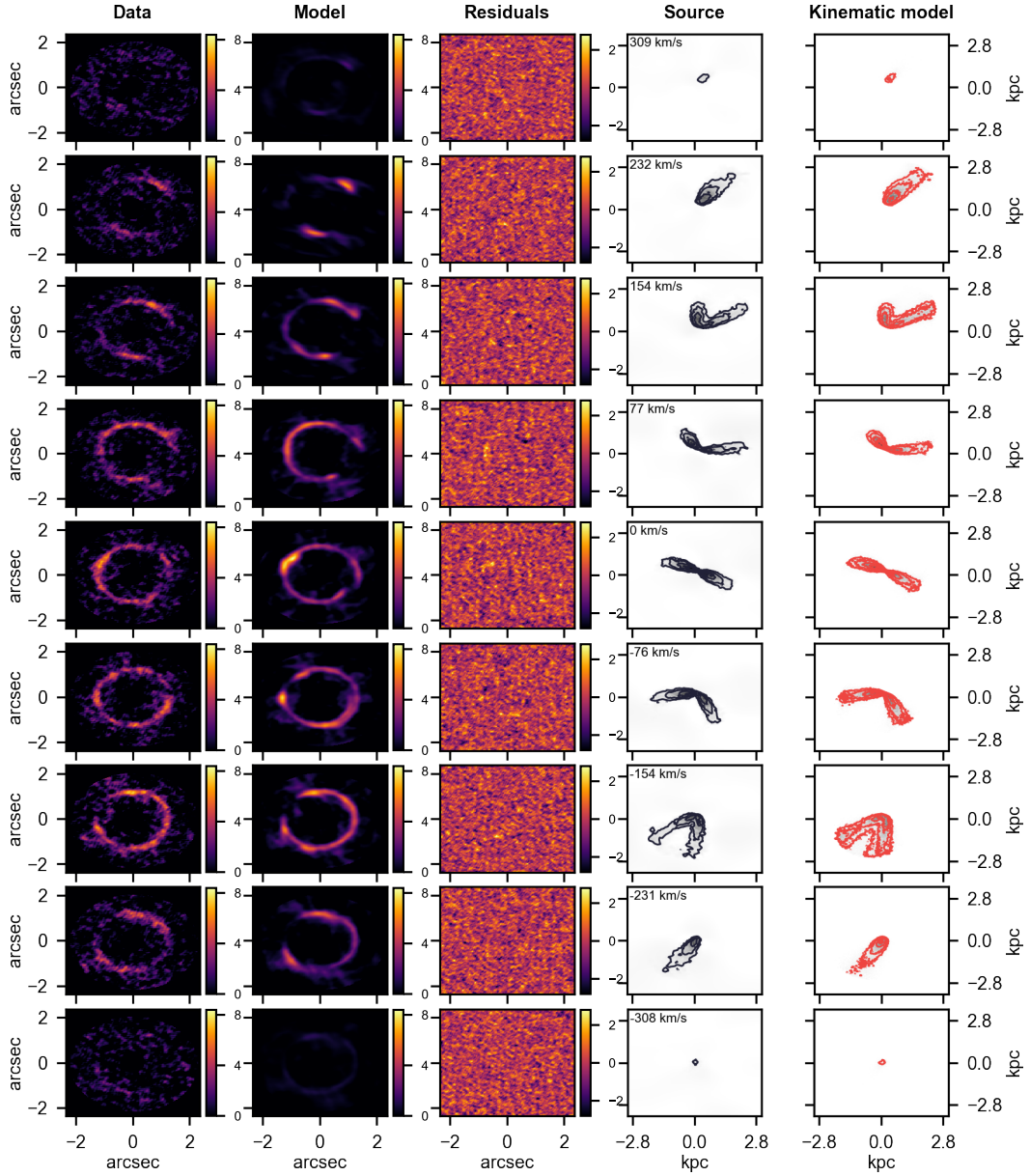


Figure 5: **Reconstruction of the [CII] emission and kinematic model.** The rows show some representative channel maps at the velocity shown on the upper left corner of column 4. Columns 1 and 2 show the dirty image of the data and the model, respectively, colour-coded by the flux in units of mJy/beam. Column 3 shows the residuals (data - model) normalized to the noise. Column 5 and 6 show the contours of the reconstructed source and of the kinematic model used to constrain the source reconstruction. The contour levels in the last columns are set at $n = 0.1, 0.2, 0.4, 0.6, 0.8$ times the value of the maximum flux of the kinematic model.

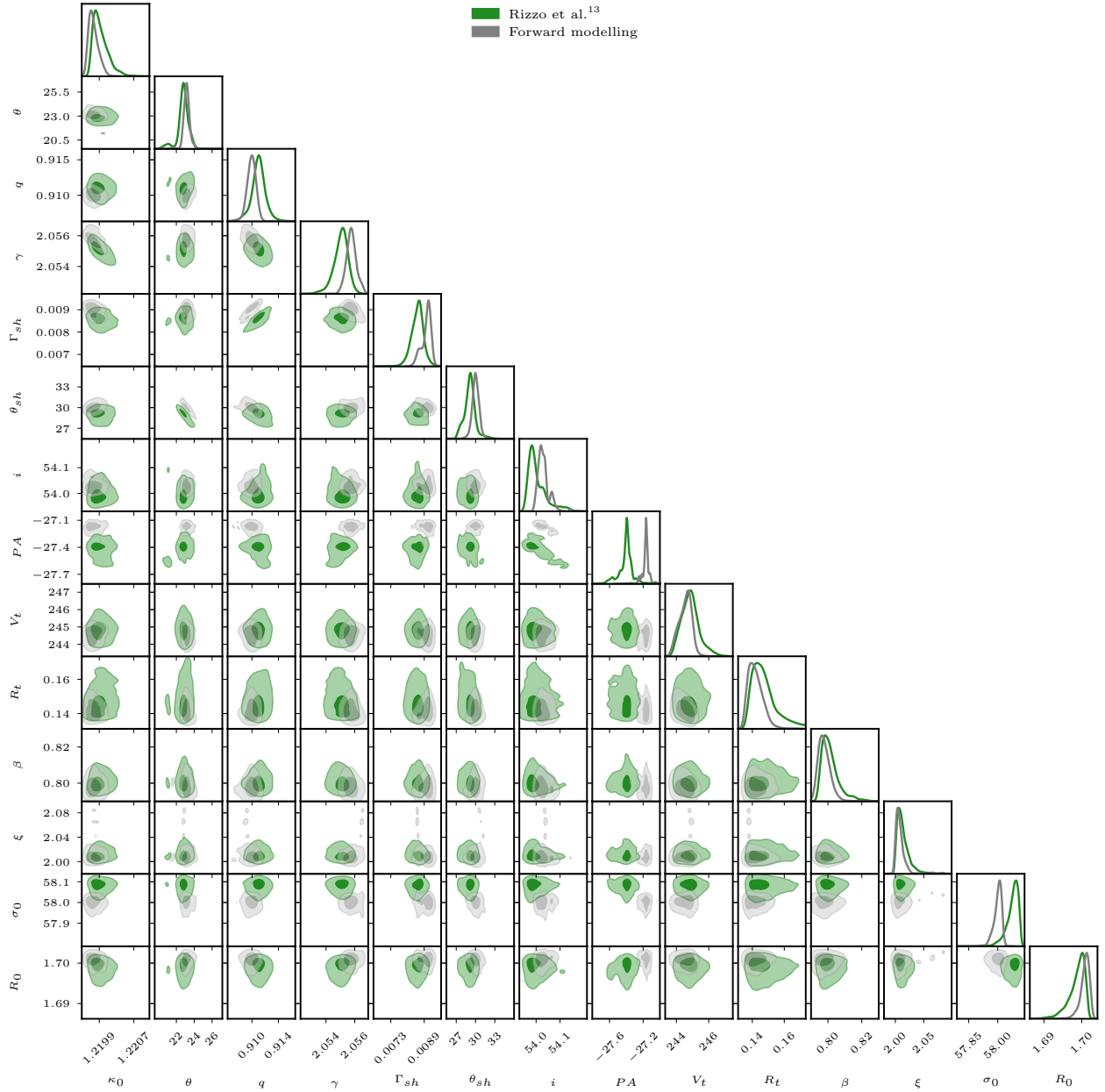


Figure 6: **Corner plot for the posterior distributions of the lens and kinematic parameters.** The dark and light areas the 2D distributions show the 39% and 86% confidence levels, corresponding to 1 s.d. and 2 s.d., respectively, obtained with the fiducial methodology described in Rizzo et al. 2018¹³ (green) and with the direct forward modelling methodology (gray). From left to right, the first six panels show the lens parameters, and the other panel show the source kinematic parameters (see also Extended Data Table 1).

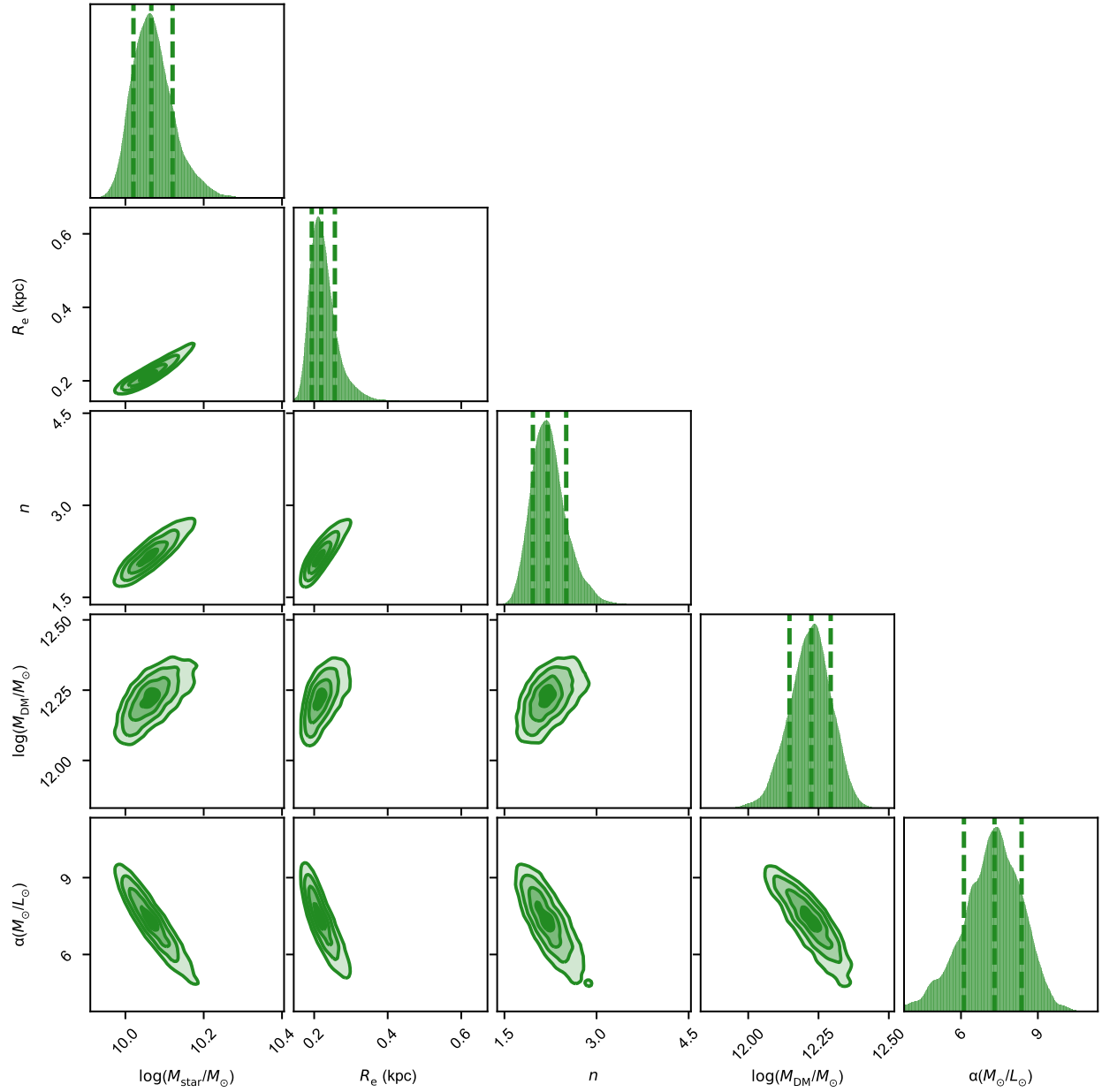


Figure 7: Corner plot for the posterior distributions of the dynamical parameters. The posterior distributions are the results of the decomposition of the circular velocity (Fig. 2c) into the physical components contributing to the total gravitational potential: the stars, the gas disk and the dark matter halo. The fitted parameters are the stellar mass M_{star} , the effective radius R_e , the Sérsic index n (Sérsic profile), the mass of an NFW dark-matter halo M_{DM} and the conversion factor between the [CII] luminosity and the gas mass $\alpha_{[\text{CII}]}$. The dashed lines in the 1D histograms show the 16th, 50th and 84th percentiles (see Extended Data Table 5).

1. Pillepich, A. *et al.* First results from the TNG50 simulation: the evolution of stellar and gaseous discs across cosmic time. *Mon. Not. R. Astron. Soc.* **490**, 3196–3233 (2019).
2. Dekel, A. & Burkert, A. Wet disc contraction to galactic blue nuggets and quenching to red nuggets. *Mon. Not. R. Astron. Soc.* **438**, 1870–1879 (2014).
3. Zolotov, A. *et al.* Compaction and quenching of high- z galaxies in cosmological simulations: blue and red nuggets. *Mon. Not. R. Astron. Soc.* **450**, 2327–2353 (2015).
4. Krumholz, M. R., Burkert, B., Forbes, J. C. & Crocker, R. M. A unified model for galactic discs: star formation, turbulence driving, and mass transport. *Mon. Not. R. Astron. Soc.* **477**, 2716–2740 (2018).
5. Hayward, C. C. & Hopkins, P. F. How stellar feedback simultaneously regulates star formation and drives outflows. *Mon. Not. R. Astron. Soc.* **465**, 1682–1698 (2017).
6. Weiß, A. *et al.* ALMA Redshifts of Millimeter-selected Galaxies from the SPT Survey: The Redshift Distribution of Dusty Star-forming Galaxies. *Astrophys. J.* **767**, 88 (2013).
7. De Breuck, C. *et al.* A dense, solar metallicity ISM in the $z = 4.2$ dusty star-forming galaxy SPT 0418-47. *Astron. Astrophys.* **631**, A167 (2019).
8. Gullberg, B. *et al.* The nature of the [C II] emission in dusty star-forming galaxies from the SPT survey. *Mon. Not. R. Astron. Soc.* **449**, 2883–2900 (2015).
9. McAlpine, S. *et al.* The nature of sub-millimeter and highly star-forming galaxies in the EAGLE simulation. *Mon. Not. R. Astron. Soc.* 1653 (2019).
10. Hodge, J. A. *et al.* The Kiloparsec-scale Star Formation Law at Redshift 4: Widespread, Highly Efficient Star Formation in the Dust-obscured Starburst Galaxy GN20. *Astrophys. J. Lett.* **798**, L18 (2015).
11. Lelli, F., McGaugh, S. S. & Schombert, J. M. SPARC: Mass Models for 175 Disk Galaxies with Spitzer Photometry and Accurate Rotation Curves. *Astron. J.* **152**, 157 (2016).
12. Stacey, G. J. *et al.* A 158 μm [C II] Line Survey of Galaxies at $z \sim 1-2$: An Indicator of Star Formation in the Early Universe. *Astrophys. J.* **724**, 957–974 (2010).
13. Rizzo, F., Vegetti, S., Fraternali, F. & Di Teodoro, E. A novel 3D technique to study the kinematics of lensed galaxies. *Mon. Not. R. Astron. Soc.* **481**, 5606–5629 (2018).
14. Toomre, A. On the gravitational stability of a disk of stars. *Astrophys. J.* **139**, 1217–1238 (1964).

15. Carilli, C. L. & Walter, F. Cool Gas in High-Redshift Galaxies. *Annu. Rev. Astron. Astr.* **51**, 105–161 (2013).
16. Chen, C.-C. *et al.* An ALMA Survey of Submillimeter Galaxies in the Extended Chandra Deep Field South: Near-infrared Morphologies and Stellar Sizes. *Astrophys. J.* **799**, 194 (2015).
17. Lang, P. *et al.* Bulge Growth and Quenching since $z = 2.5$ in CANDELS/3D-HST. *Astrophys. J.* **788**, 11 (2014).
18. Krajnović, D. *et al.* The ATLAS^{3D} project - XVII. Linking photometric and kinematic signatures of stellar discs in early-type galaxies. *Mon. Not. R. Astron. Soc.* **432**, 1768–1795 (2013).
19. Toft, S. *et al.* Submillimeter Galaxies as Progenitors of Compact Quiescent Galaxies. *Astrophys. J.* **782**, 68 (2014).
20. Barro, G. *et al.* CANDELS+3D-HST: Compact SFGs at $z \sim 2-3$, the Progenitors of the First Quiescent Galaxies. *Astrophys. J.* **791**, 52 (2014).
21. Cappellari, M. *et al.* The ATLAS^{3D} project - I. A volume-limited sample of 260 nearby early-type galaxies: science goals and selection criteria. *Mon. Not. R. Astron. Soc.* **413**, 813–836 (2011).
22. McDermid, R. M. *et al.* The ATLAS^{3D} Project - XXX. Star formation histories and stellar population scaling relations of early-type galaxies. *Mon. Not. R. Astron. Soc.* **448**, 3484–3513 (2015).
23. Naab, T., Johansson, P. H. & Ostriker, J. P. Minor Mergers and the Size Evolution of Elliptical Galaxies. *Astrophys. J. Lett.* **699**, L178–L182 (2009).
24. Cappellari, M. *et al.* The ATLAS^{3D} project - XX. Mass-size and mass- σ distributions of early-type galaxies: bulge fraction drives kinematics, mass-to-light ratio, molecular gas fraction and stellar initial mass function. *Mon. Not. R. Astron. Soc.* **432**, 1862–1893 (2013).
25. Di Teodoro, E. M., Fraternali, F. & Miller, S. H. Flat rotation curves and low velocity dispersions in KMOS star-forming galaxies at $z \sim 1$. *Astron. Astrophys.* **594**, A77 (2016).
26. Wisnioski, E. *et al.* The KMOS^{3D} Survey: Design, First Results, and the Evolution of Galaxy Kinematics from $0.7 \leq z \leq 2.7$. *Astrophys. J.* **799**, 209 (2015).

27. Lelli, F. *et al.* Neutral versus ionized gas kinematics at $z \approx 2.6$: the AGN-host starburst galaxy PKS 0529-549. *Mon. Not. R. Astron. Soc.* **479**, 5440–5447 (2018).
28. Turner, O. J. *et al.* The KMOS Deep Survey (KDS) - I. Dynamical measurements of typical star-forming galaxies at $z \approx 3.5$. *Mon. Not. R. Astron. Soc.* **471**, 1280–1320 (2017).
29. Harrison, C. M. *et al.* The KMOS Redshift One Spectroscopic Survey (KROSS): rotational velocities and angular momentum of $z \approx 0.9$ galaxies. *Mon. Not. R. Astron. Soc.* **467**, 1965–1983 (2017).
30. Swinbank, A. M. *et al.* Angular momentum evolution of galaxies over the past 10 Gyr: a MUSE and KMOS dynamical survey of 400 star-forming galaxies from $z = 0.3$ to 1.7. *Mon. Not. R. Astron. Soc.* **467**, 3140–3159 (2017).
31. Carlstrom, J. E. *et al.* The 10 Meter South Pole Telescope. *Publ. Astron. Soc. Pac.* **123**, 568 (2011).
32. Vieira, J. D. *et al.* Dusty starburst galaxies in the early Universe as revealed by gravitational lensing. *Nature* **495**, 344–347 (2013).
33. McMullin, J. P., Waters, B., Schiebel, D., Young, W. & Golap, K. CASA Architecture and Applications. In Shaw, R. A., Hill, F. & Bell, D. J. (eds.) *Astr. Soc. P.*, vol. 376 of *Astronomical Society of the Pacific Conference Series*, 127 (2007).
34. Högbom, J. A. Aperture Synthesis with a Non-Regular Distribution of Interferometer Baselines. *Astron. Astrophys. Sup.* **15**, 417 (1974).
35. Planck Collaboration *et al.* Planck 2015 results. XIII. Cosmological parameters. *Astron. Astrophys.* **594**, A13 (2016).
36. Powell, D., Vegetti, S., McKean, J. P. & Spingola, C. A novel approach to visibility-space modelling of interferometric gravitational lens observations at high angular resolution. *arXiv e-prints* arXiv:2005.03609 (2020).
37. Vegetti, S. & Koopmans, L. V. E. Bayesian strong gravitational-lens modelling on adaptive grids: objective detection of mass substructure in Galaxies. *Mon. Not. R. Astron. Soc.* **392**, 945–963 (2009).
38. Greengard, L. & Lee, J.-Y. Accelerating the nonuniform fast fourier transform. *SIAM Review* **46**, 443–454 (2004).

39. Beatty, P. J., Nishimura, D. G. & Pauly, J. M. Rapid gridding reconstruction with a minimal oversampling ratio. *IEEE Transactions on Medical Imaging* **24**, 799–808 (2005).
40. Koopmans, L. V. E., Treu, T., Bolton, A. S., Burles, S. & Moustakas, L. A. The Sloan Lens ACS Survey. III. The Structure and Formation of Early-Type Galaxies and Their Evolution since $z \sim 1$. *Astrophys. J.* **649**, 599–615 (2006).
41. Barnabè, M. *et al.* Two-dimensional kinematics of SLACS lenses - II. Combined lensing and dynamics analysis of early-type galaxies at $z = 0.08-0.33$. *Mon. Not. R. Astron. Soc.* **399**, 21–36 (2009).
42. Spilker, J. S. *et al.* ALMA Imaging and Gravitational Lens Models of South Pole Telescope—Selected Dusty, Star-Forming Galaxies at High Redshifts. *Astrophys. J.* **826**, 112 (2016).
43. Feroz, F., Hobson, M. P., Cameron, E. & Pettitt, A. N. Importance Nested Sampling and the MultiNest Algorithm. *arXiv e-prints* arXiv:1306.2144 (2013).
44. Di Teodoro, E. M. & Fraternali, F. ^{3D} BAROLO: a new 3D algorithm to derive rotation curves of galaxies. *Mon. Not. R. Astron. Soc.* **451**, 3021–3033 (2015).
45. Jones, G. C. *et al.* Dynamical Characterization of Galaxies at $z \sim 4-6$ via Tilted Ring Fitting to ALMA [C II] Observations. *Astrophys. J.* **850**, 180 (2017).
46. Smit, R. *et al.* Rotation in [C II]-emitting gas in two galaxies at a redshift of 6.8. *Nature* **553**, 178–181 (2018).
47. Übler, H. *et al.* Ionized and Molecular Gas Kinematics in a $z = 1.4$ Star-forming Galaxy. *Astrophys. J. Lett.* **854**, L24 (2018).
48. Girard, M. *et al.* Towards sub-kpc scale kinematics of molecular and ionized gas of star-forming galaxies at $z \sim 1$. *Astron. Astrophys.* **631**, A91 (2019).
49. Übler, H. *et al.* The Evolution and Origin of Ionized Gas Velocity Dispersion from $z \sim 2.6$ to $z \sim 0.6$ with KMOS^{3D}. *Astrophys. J.* **880**, 48 (2019).
50. Hung, C.-L. *et al.* What drives the evolution of gas kinematics in star-forming galaxies? *Mon. Not. R. Astron. Soc.* **482**, 5125–5137 (2019).
51. Teklu, A. F. *et al.* Declining Rotation Curves at $z = 2$ in Λ CDM Galaxy Formation Simulations. *Astrophys. J. Lett.* **854**, L28 (2018).

52. Bird, J. C. *et al.* Inside out and Upside down: Tracing the Assembly of a Simulated Disk Galaxy Using Mono-age Stellar Populations. *Astrophys. J.* **773**, 43 (2013).
53. Martizzi, D. Global simulations of galactic discs: violent feedback from clustered supernovae during bursts of star formation. *Mon. Not. R. Astron. Soc.* **492**, 79–95 (2020).
54. Iorio, G. *et al.* LITTLE THINGS in 3D: robust determination of the circular velocity of dwarf irregular galaxies. *Mon. Not. R. Astron. Soc.* **466**, 4159–4192 (2017).
55. Lima Neto, G. B., Gerbal, D. & Márquez, I. The specific entropy of elliptical galaxies: an explanation for profile-shape distance indicators? *Mon. Not. R. Astron. Soc.* **309**, 481–495 (1999).
56. Terzić, B. & Graham, A. W. Density-potential pairs for spherical stellar systems with Sérsic light profiles and (optional) power-law cores. *Mon. Not. R. Astron. Soc.* **362**, 197–212 (2005).
57. Zanella, A. *et al.* The [C II] emission as a molecular gas mass tracer in galaxies at low and high redshifts. *Mon. Not. R. Astron. Soc.* **481**, 1976–1999 (2018).
58. Gullberg, B. *et al.* The Dust and [C II] Morphologies of Redshift ~ 4.5 Sub-millimeter Galaxies at ~ 200 pc Resolution: The Absence of Large Clumps in the Interstellar Medium at High-redshift. *Astrophys. J.* **859**, 12 (2018).
59. Navarro, J. F., Frenk, C. S. & White, S. D. M. The Structure of Cold Dark Matter Halos. *Astrophys. J.* **462**, 563 (1996).
60. Dutton, A. A. & Macciò, A. V. Cold dark matter haloes in the Planck era: evolution of structural parameters for Einasto and NFW profiles. *Mon. Not. R. Astron. Soc.* **441**, 3359–3374 (2014).
61. Speagle, J. S. dynesty: A Dynamic Nested Sampling Package for Estimating Bayesian Posteriors and Evidences. *arXiv e-prints* arXiv:1904.02180 (2019).
62. Cappellari, M. *et al.* The ATLAS^{3D} project - XV. Benchmark for early-type galaxies scaling relations from 260 dynamical models: mass-to-light ratio, dark matter, Fundamental Plane and Mass Plane. *Mon. Not. R. Astron. Soc.* **432**, 1709–1741 (2013). 1208.3522.
63. Emsellem, E. *et al.* The ATLAS^{3D} project - III. A census of the stellar angular momentum within the effective radius of early-type galaxies: unveiling the distribution of fast and slow rotators. *Mon. Not. R. Astron. Soc.* **414**, 888–912 (2011).

64. Aravena, M. *et al.* A survey of the cold molecular gas in gravitationally lensed star-forming galaxies at $z > 2$. *Mon. Not. R. Astron. Soc.* **457**, 4406–4420 (2016).
65. Kennicutt, R. C. & Evans, N. J. Star Formation in the Milky Way and Nearby Galaxies. *Annu. Rev. Astron. Astr.* **50**, 531–608 (2012).
66. Tamburro, D. *et al.* What is Driving the H I Velocity Dispersion? *Astrophys. J.* **137**, 4424–4435 (2009).
67. Utomo, D., Blitz, L. & Falgarone, E. The Origin of Interstellar Turbulence in M33. *Astrophys. J.* **871**, 17 (2019).
68. Rafelski, M. *et al.* The Star Formation Rate Efficiency of Neutral Atomic-dominated Hydrogen Gas in the Outskirts of Star-forming Galaxies from $z \sim 1$ to $z \sim 3$. *Astrophys. J.* **825**, 87 (2016).
69. Leroy, A. K. *et al.* The Multi-phase Cold Fountain in M82 Revealed by a Wide, Sensitive Map of the Molecular Interstellar Medium. *Astrophys. J.* **814**, 83 (2015).
70. Lelli, F., Verheijen, M. & Fraternali, F. Dynamics of starbursting dwarf galaxies. III. A H I study of 18 nearby objects. *Astron. Astrophys.* **566**, A71 (2014).

Acknowledgements This letter makes use of the following ALMA data: 2016.1.01499.S. ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada), NSC and ASIAA (Taiwan), and KASI (Republic of Korea), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO and NAOJ. SV has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 758853). F.R. thanks T. Naab and T. Costa for useful comments and discussions.

Author contributions F.R., F.F. and S.V. analysed the data. D.P., F.R. and S.V. developed the software used for the lens-kinematic modelling. F.R. developed the software for the dynamical analysis. H.R.S. and F.R. reduced the data. F.R., J.P.M., S.V. and H.R.S. contributed to the writing of the manuscript. F.F. and S.D.M.W. helped with the interpretation of the scientific results. All authors discussed the results and provided comments on the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare that they have no competing financial interests. Correspondence and requests for materials should be addressed to F.R. (email: frizzo@mpa-garching.mpg.de).

Data Availability This paper makes use of the following ALMA data: 2016.1.01499.S, available at <http://almascience.eso.org/aq/>.

Code Availability The methodology about the source reconstruction and its kinematic modelling is fully explained in Reference 13. The code is not publicly available. However, the reader interested in using this code can contact the first author.