the Spitzer/IRAC fluxes to be reliable. By contrast, SPT0311−58 E is one full IRAC resolution element, 1.7″, from the lens centroid, and we consider the residual emission at this position to be usable in our subsequent analyses. Images of the model and residuals are provided in Extended Data Fig. 4 and the resulting photometry is provided in Extended Data Table 2.

Gravitational lens modelling. Gravitational lens modelling of SPT0311−58 was performed using two different codes which model the source-plane emission in different ways. Both codes fit to the visibilities measured by ALMA or other interferometers directly to avoid the correlated noise between pixels in inverted images. In each, the lens galaxy is modelled as a singular isothermal ellipsoid, and posterior parameter distributions are sampled using a Markov chain Monte Carlo technique, marginalizing over several sources of residual calibration uncertainty (such as antenna-based phase errors).

Initial lens models were created using the visilens code, which is described in detail elsewhere[24]. The source plane is modelled as one or more elliptical Sérsic profiles. Because of the simplicity of this source-plane representation, the code is able to sample large and complex parameter spaces quickly. The continuum emission at 160 μm, 110 μm and 90 μm was modelled with four Sérsic components, one for SPT0311−58 E and three for SPT0311−58 W. These models leave approximately 8σ peak residuals in the 160-μm and 90-μm data, which both reaching peak signal-to-noise ratios of more than 150.

After determining the lens parameters using visilens, we used the best-fitting values as initial input to a pixelated reconstruction code[12]. This code represents the source plane as an array of pixels, rather than an analytic model, and determines the most probable pixel intensity values for each trial lens model while imposing a gradient-type regularization[42] to avoid over-fitting the data. For each dataset, we fit for the strength of this regularization. At 160 μm and 90 μm we re-fit for the lens model parameters and compare to the visilens models as a test of the robustness of the lens modelling. Within each code, the best-fitting lens parameters at the two independent wavelengths are consistent to within 10%. Further, the lens parameters and the source structure are consistent between the two independent codes, with intrinsic source flux densities, sizes and magnifications that agree to within 15%. The increased freedom in the source plane afforded by the pixelated reconstruction means that the lens parameters are not independently well constrained by the 110-μm data, which have lower signal-to-noise ratio and spatial resolution. For these data, we apply the lensing deflections determined from the other two datasets to reconstruct the source-plane emission. The pixelated reconstructions of the three continuum wavelengths are shown in Extended Data Fig. 5.

The channelized [C II] line is modelled using the same pixelated reconstruction technique, using 39 consecutive channels of 40 km s^{-1} width, each with a peak signal-to-noise ratio ranging from 9 to 34. For each channel, we apply the lensing deflections from the best-fitting model of the 160-μm data, which were observed simultaneously. We fit for the strength of the source-plane regularization[12,42] at each channel, which varies across the line profile as some velocities (such those multiply imaged from −280 km s^{-1} to +80 km s^{-1}) experience higher magnification than others (such as the entire eastern source at >+560 km s^{-1}). The models of each [C II] channel are represented in Extended Data Fig. 6.

We determine the source magnifications using the 90-μm pixelated model, in which the E source is detected at the highest signal-to-noise ratio and so the effects of varying the aperture used to measure the intrinsic flux density are minimized. Because the source-plane morphology is very similar between the three continuum wavelengths, the magnification is also essentially identical between them. We find flux-weighted, source-averaged magnifications for the E source, the W source and the system as a whole of $\mu_E = 1.3$, $\mu_W = 2.2$ and $\mu_{tot} = 2.0$, respectively. These magnifications are substantially lower than the median magnification of 5.5 within the sample of 47 SPT-discovered dusty galaxies[24] for which we have data adequate to construct lens models or to conclude that sources are unlensed. In this case the low magnification is a consequence of the low mass of the lensing halo, which is typically expressed as an 'Einstein' radius θ_E. The lens model for this source indicates $\theta_E = 0.29''$, which is around the 10th percentile for SPT lensed sources[24], and the background source is both much larger than and offset from the regions of highest magnification. A large portion of the source is therefore only weakly magnified and the source-averaged values are low.

Finally, we also construct a lens model of the 95-GHz ALMA data (rest-frame 380 μm; Extended Data Table 1). Because the spatial resolution of these data are low (3.5″), we model them using only the visilens code, which is more suited to low-resolution data. We allow only the lens parameters and source structural parameters (such as position and radius) to vary within the ranges determined from the higher-resolution 160-μm, 110-μm and 90-μm continuum data, leaving only the flux densities of the E and W sources as free parameters. This modelling indicates that essentially all of the observed 380-μm emission can be ascribed to the W source, with the E source 'detected' at about 1σ.

In addition to the ALMA data, we use Herschel photometry[6] to constrain the SED of SPT0311−58 E and SPT0311−58 W to rest-frame 30 μm (250 μm observed). The resolution of Herschel SPIRE is not adequate to separate the two components, so we divide the total flux density observed in the three SPIRE bands between the E and W sources according to the ratios observed in the ALMA bands. These photometric points are then corrected for the continuum magnification derived from the ALMA data and used in the SED modelling described below. The total and intrinsic flux densities are reported in Extended Data Table 3.

Modelling the SED. In Extended Data Fig. 7 we present the SEDs of SPT0311−58 E, SPT0311−58 W and the foreground lens galaxy.

A photometric redshift for the lens is calculated with EAZY[43] using the data in Extended Data Table 2. The resulting redshift is 1.43, with a 1σ confidence interval of 1.08–1.85. The lens SED fitting is performed with the Code Investigating GALaxy Emission (CIGALE[44,45]) assuming $z = 1.43$.

The multiple rest-frame ultraviolet to rest-frame optical detections of SPT0311−58 E allow us to constrain the stellar mass using reasonable assumptions about the star-formation history at this early point in cosmic history. The SED is fitted by varying the e-folding time and age of a previously reported stellar population model[46] under single- and two-component formation histories, assuming solar metallicity and previously reported[47] initial mass function. The minimum radiation field, power-law slope and gamma, the fraction of dust mass exposed to radiation intensities above the minimum, from one dust model[48], and the colour excess and attenuation slope from other dust models[49,50] are kept free in the SED fitting. The AGN contribution is set to zero because there are no photometric points to constrain the spectral range that is most affected by AGN power (mid-infrared) and thus any fraction between 0% and 60% of the dust luminosity is attributable to AGNs with nearly equal probability. However, this ignores the spatial distributions of the dust and line emission, which are not strongly peaked as is usually observed in AGN-dominated galaxies, so we deem this wide range to be unphysical. The inferred stellar mass and star formation rates are $(3.5 \pm 1.5) \times 10^{10} M_\odot$ and $(540 \pm 175) M_\odot$ yr^{-1}, respectively, for the two-component star-formation history. These values agree within the uncertainties for a single-component star-formation history. The infrared luminosity (L_{IR}; integrated over 8–1,000 μm) is $(4.6 \pm 1.2) \times 10^{12} L_\odot$ and the extinction is $A_V = 2.7 \pm 0.2$ mag.

For the W source, we have only upper limits and the potentially contaminated IRAC detections to constrain the rest-frame optical and ultraviolet emission. Accordingly, we use the IRAC photometry as upper limits, along with the HST limits and far-infrared data in Extended Data Table 3, and model the SED with CIGALE. We find a luminosity of $L_{IR} = (33 \pm 7) \times 10^{12} L_\odot$, seven times larger than for the E source. A consistent luminosity is obtained by fitting the far-infrared SED with a modified blackbody[51]. The inferred star-formation rate, which is closely connected to L_{IR}, is $(2{,}900 \pm 1{,}800) M_\odot$ yr^{-1}. As for the E source, the SED allows the AGN fraction to fall between 0% and 60% with roughly equal probability, so we take the absence of a dominant infrared emission region (see Fig. 1c and Extended Data Fig. 5) as an indication that the AGN contribution is unlikely to be important and fix the AGN fraction to zero. The dust luminosity due to star formation could therefore in principle be up to a factor of two smaller if the spatial distribution of the emission is ignored. Given that the photometry reaches to only the rest-frame V band, it is possible to hide a very large stellar mass behind dust obscuration for plausible values of the visual extinction ($A_V \leq 6$, as seen in other massive dusty galaxies[14,16,52,53]). Considering the IRAC flux densities alone, we can calculate rest-frame mass-to-light ratios for the observed bands to see what masses could exist without relying on the poorly constrained CIGALE SED modelling. We use a stellar population synthesis code[54,55] to compute a stellar mass-to-light ratio under a range of assumptions: stellar ages of 0.1–0.8 Gyr (from a reasonably 'young' population to the approximate age of the Universe at the time) and metallicity of 0.1–1 times that of the Sun, with no dust attenuation. The mass then ranges from $(2–10) \times 10^{10} M_\odot$ per μJy of measured flux density. Taking the measured and de-magnified flux density (averaged between the two wavelengths) of 0.5 μJy, we find a stellar mass of $(1–5) \times 10^{10} M_\odot$ before correcting for extinction. If the extinction is as large as 5 mag, the true stellar mass could be unphysically large ($>10^{12} M_\odot$), demonstrating that we have no useful constraint without greater certainty about the reliability of the IRAC flux densities or more photometric data points.

Galaxy and halo masses. In Figs 2 and 3 we compile mass measurements for high-redshift galaxies discovered by various techniques. The galaxy sample comprises primarily galaxies identified through their luminous dust emission (DSFGs) and optically identified quasars (QSOs), which are typically the objects with the largest gas, dust or stellar masses at these redshifts. At the very highest redshifts, where very few galaxies have been found, objects selected on the basis of their ultraviolet emission are also included. The subsets of galaxies included in each