

# A kiloparsec-scale internal shock collision in the jet of a nearby radio galaxy

Eileen T. Meyer<sup>1,2</sup>, Markos Georganopoulos<sup>2,3</sup>, William B. Sparks<sup>1</sup>, Eric Perlman<sup>4</sup>, Roeland P. van der Marel<sup>1</sup>, Jay Anderson<sup>1</sup>, Sangmo Tony Sohn<sup>5</sup>, John Biretta<sup>1</sup>, Colin Norman<sup>1,5</sup> & Marco Chiaberge<sup>1,5,6</sup>

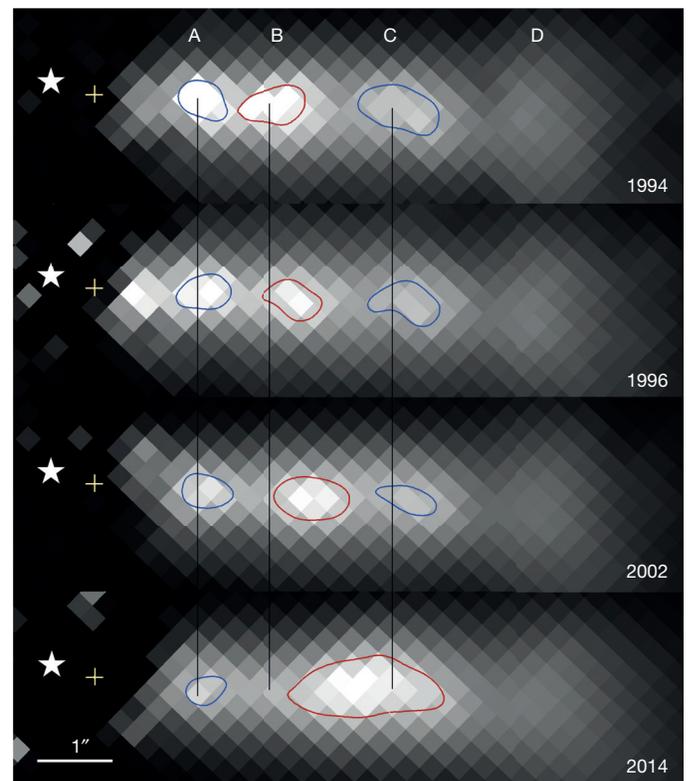
Jets of highly energized plasma with relativistic velocities are associated with black holes ranging in mass from a few times that of the Sun to the billion-solar-mass black holes at the centres of galaxies<sup>1</sup>. A popular but unconfirmed hypothesis to explain how the plasma is energized is the ‘internal shock model’, in which the relativistic flow is unsteady<sup>2</sup>. Faster components in the jet catch up to and collide with slower ones, leading to internal shocks that accelerate particles and generate magnetic fields<sup>3</sup>. This mechanism can explain the variable, high-energy emission from a diverse set of objects<sup>4–7</sup>, with the best indirect evidence being the unseen fast relativistic flow inferred to energize slower components in X-ray binary jets<sup>8,9</sup>. Mapping of the kinematic profiles in resolved jets has revealed precessing and helical patterns in X-ray binaries<sup>10,11</sup>, apparent superluminal motions<sup>12,13</sup>, and the ejection of knots (bright components) from standing shocks in the jets of active galaxies<sup>14,15</sup>. Observations revealing the structure and evolution of an internal shock in action have, however, remained elusive, hindering measurement of the physical parameters and ultimate efficiency of the mechanism. Here we report observations of a collision between two knots in the jet of nearby radio galaxy 3C 264. A bright knot with an apparent speed of  $(7.0 \pm 0.8)c$ , where  $c$  is the speed of light in a vacuum, is in the incipient stages of a collision with a slower-moving knot of speed  $(1.8 \pm 0.5)c$  just downstream, resulting in brightening of both knots—as seen in the most recent epoch of imaging.

We obtained deep V-band imaging of 3C 264, a radio galaxy 91 megaparsecs (Mpc) from Earth with the Advanced Camera for Surveys (ACS) on board the Hubble Space Telescope (HST) in May 2014. The comparatively deep ACS imaging provides a reference image to compare against previous HST imaging for evidence of proper motions of the four previously known optical knots within the 2''-long jet<sup>16–18</sup>. We localized over a hundred globular clusters in the host galaxy as a reference system on which to register previous V-band images taken with HST’s Wide Field and Planetary Camera 2 (WFPC2) in 1994, 1996, and 2002. The systematic error in the registration of the WFPC2 images is generally of the order of 5 milliarcseconds (mas) or less. After aligning all images to a common reference frame, the fast proper motion of knot B is clearly visible, as shown in Fig. 1 and Supplementary Videos 1 and 2. Previous radio observations have revealed that the initially narrowly collimated jet bends by about 10° at the location marked by the yellow cross<sup>19</sup>, which appears to align well with the central axis of the jet in our imaging, and which serves as our reference point for all measured positions.

To measure their apparent speeds, the position of each knot was measured with a centroiding technique. In the case of knots B and C, we also modelled the jet as a constant-density conical jet with superimposed resolved knots, in order to measure their fluxes and positions better, particularly in the final epoch when they appear to overlap. We plotted the position of each knot along the direction of the jet axis

versus time, and fitted the data with a least-squares linear model. A slope significantly larger than zero ( $P < 0.05$ ) indicates significant proper motions, and we used the conversion factor 1.442c years per milliarcsecond to convert angular apparent speeds  $\mu_{\text{app}}$  to dimensionless apparent speed relative to  $c$ ,  $\beta_{\text{app}}$  (see Table 1).

We found that knots A and D have an apparent speed  $\beta_{\text{app}}$  consistent with zero (Extended Data Figs 1 and 2), while the inner knots B and C have  $\beta_{\text{app}} = 7.0 \pm 0.8$  and  $1.8 \pm 0.5$ , respectively (Fig. 2). The value for knot B exceeds the fastest speeds measured in the jet in the nearby radio galaxy M87, the only other source for which speeds on kiloparsec (kpc) scales have been measured<sup>20,21</sup>. The current difference in speeds between knots B and C puts them on a collision course, an interaction



**Figure 1 | A comparison of HST images of the jet in 3C 264 from 1994 to 2014.** The galaxy and core emission from 3C 264 have been subtracted, with a white star representing the location of the black hole, and a yellow cross indicating the location of a bend in the jet seen with radio interferometry<sup>19</sup>. The contours show the 30% flux-over-background isophotes around each peak and are overlaid along with vertical guidelines to aid the eye. The first three images were taken with WFPC2, and the final epoch is a deep ACS/WFC image taken for the purposes of measuring proper motions in the jet.

<sup>1</sup>Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, Maryland 21218, USA. <sup>2</sup>University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, Maryland 21250, USA. <sup>3</sup>NASA Goddard Space Flight Center, 8800 Greenbelt Road, Greenbelt, Maryland 20771, USA. <sup>4</sup>Florida Institute of Technology, 150 West University Boulevard, Melbourne, Florida 32901, USA. <sup>5</sup>Johns Hopkins University, 3400 North Charles Street, Baltimore, Maryland 21218, USA. <sup>6</sup>Istituto Nazionale Astrofisica, Istituto di Radio Astronomia, Via Piero Gobetti 101, I-40129 Bologna, Italy.

**Table 1 | Summary of knot positions and speeds**

	Distance along the jet		Proper motion	
	Initial (1994) (mas)	Final (2014) (mas)	$\mu_{\text{app}}$ (mas yr <sup>-1</sup> )	$\beta_{\text{app}}$
Knot A	148 ± 5	149 ± 1	0.07 ± 0.20	0.1 ± 0.3
Knot B	257 ± 5	359 ± 2	4.85 ± 0.58	7.0 ± 0.8
Knot C	450 ± 6	478 ± 2	1.27 ± 0.32	1.8 ± 0.5
Knot D	582 ± 5	581 ± 1	-0.13 ± 0.31	-0.2 ± 0.5

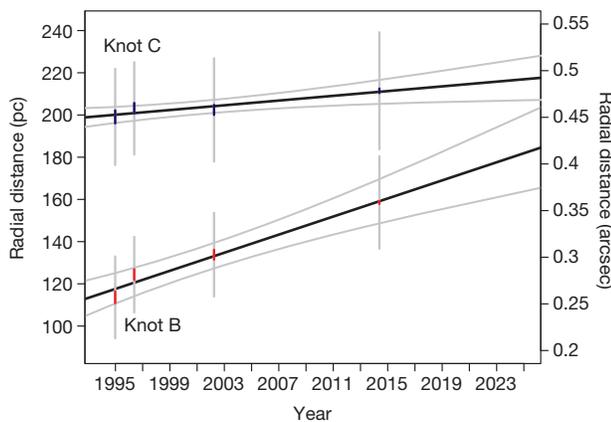
Knots A, D by contour method; Knots B, C from maximum-likelihood model. The distance along the jet is measured from the bend in the radio jet noted in Fig. 1.  $\beta_{\text{app}} = v/c$ , where  $v$  is velocity.

which has already begun in the final epoch from 2014, in which the knots begin to overlap (Fig. 2).

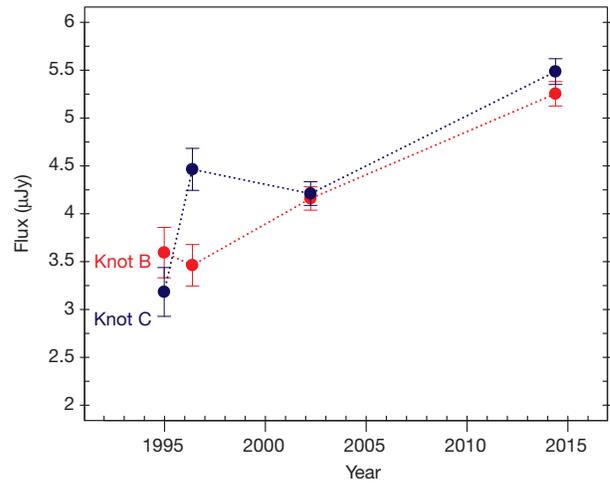
In the internal shock model, the collision of two components results in particle acceleration, which will manifest in a significant brightening of the components as they combine into a single moving component. Our modelling results show that in the final 2014 epoch, both knots B and C brighten at the same time by approximately 40% over the mean flux level of the previous three epochs (Fig. 3). The brightening can also be seen in Extended Data Fig. 3, where we show the flux contour along the length of the jet (averaged transversely over a distance of 0.1") for each epoch.

The flux increase is also corroborated by measuring the flux within a circular aperture centred on and enveloping knots B and C in all epochs, which shows a significant total flux change of approximately 20%–40% over prior epochs. Under equipartition, the cooling length for the optical-emitting electrons for knots B and C is longer than the distance they travelled in our observations, consistent with the lack of any decay in flux levels for these knots over the first three epochs. This is not the case for stationary knot A, which can be seen to decay with a timescale of around 70 years. Knot A appears to be analogous to knot HST-1 in the M87 jet; the latter is thought to be a stationary re-confinement shock where the jet pressure drops below that of the external environment<sup>22</sup>. The event that energized knot A may have been the passage of fast-moving knot B in about 1971<sup>+8</sup><sub>-17</sub>, yielding a timescale comparable to the knot A decay time.

Previous radio observations of the 3C 264 jet show that the locations of the knots in contemporaneous radio and optical imaging are very similar<sup>18,23</sup>. The earliest radio image of 3C 264 taken in October 1983 with the VLA clearly shows three distinct features (see contour comparison in Extended Data Fig. 4). While knot D appears to be completely stationary over this time-frame, the data suggest the possibility



**Figure 2 | Position versus time for knots B and C.** The position of the centre of each knot is noted by thick red and blue lines, respectively, with line extent corresponding to the  $3\sigma \chi^2$  modelling error plus the systematic error on the mean from the image registration. The grey extensions show the best-fit size of the resolved knot from the model. As shown, the best-fit linear slope yields a speed of  $7 \pm 0.8$  for knot B and  $1.8 \pm 0.5$  for knot C. In the final epoch the knots are directly adjacent and possibly overlapping, within the errors.



**Figure 3 | Change in optical flux at 6,000 Å in the colliding knots B and C over 20 years.** The knots show a simultaneous significant increase in flux in the 2014 image. Errors are the  $3\sigma \chi^2$  modelling error.

that knots B and C were moving faster in the past and may have decelerated (Extended Data Fig. 5).

In the internal shock model, the efficiency  $\eta$  of the conversion of the dissipated kinetic energy  $E_{\text{diss}}$  into radiation is generally unknown. For the collision in 3C 264 we can estimate it from our observations (further details are in the Methods): we take the minimum possible value of the Lorentz factor  $\Gamma_B = 7.1$  for knot B so that the jet is observed at an angle of  $8.1^\circ$ , and for slower knot C we take  $\Gamma_C = 2.8$ . We assume an equipartition magnetic field and one proton per electron in the plasma to obtain masses of  $2.5 \times 10^{30}$  g and  $6.7 \times 10^{30}$  g for knots B and C. After the collision, the single combined component will move with Lorentz factor  $\Gamma_m = 3.7$ . The total energy dissipated is taken as the difference in kinetic energy before and after the collision and is  $E_{\text{diss}} = 2.6 \times 10^{51}$  erg. If the knot B+C complex stays at the flux level observed in 2014 throughout the duration of the collision (which we take to be the knot superposition time of 30 years in the observer's frame, as one would expect if the cooling time of the optically emitting electrons is shorter than the collision time), the efficiency of conversion is only  $\eta = 10^{-3}$ , considerably lower than is usually assumed in models.

This, however, is a lower limit for two reasons. First, theoretical modelling<sup>24</sup> of internal shocks suggests that the flux should steadily rise to a peak occurring halfway through the collision, which would increase  $\eta$ . Second, it is possible that the cooling time of the optically emitting electrons is longer than the collision time, which would increase the duration of the elevated optical–ultraviolet emission, and therefore  $\eta$ . The rate of cooling depends on the magnetic field in the shocked plasma. Long-term monitoring of the collision in 3C 264 over the coming decades could probe the evolution of the flux and hence constrain two free parameters of the internal shock model: the fraction of  $E_{\text{diss}}$  that goes to radiating electrons and the fraction that goes to generating the magnetic field in the shocked plasma.

**Online Content** Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 25 November 2014; accepted 31 March 2015.

1. Livio, M. Astrophysical jets: a phenomenological examination of acceleration and collimation. *Phys. Rep.* **311**, 225–245 (1999).
2. Rees, M.J. The M87 jet—internal shocks in a plasma beam. *Mon. Not. R. Astron. Soc.* **184**, 61–65 (1978).
3. Medvedev, M. V. & Loeb, A. Generation of magnetic fields in the relativistic shock of gamma-ray burst sources. *Astrophys. J.* **526**, 697–706 (1999).
4. Paczynski, B. & Xu, G. Neutrino bursts from gamma-ray bursts. *Astrophys. J.* **427**, 708–713 (1994).
5. Rees, M. J. & Meszaros, P. Unsteady outflow models for cosmological gamma-ray bursts. *Astrophys. J.* **430**, L93–L96 (1994).

6. Fender, R. P., Belloni, T. M. & Gallo, E. Towards a unified model for black hole X-ray binary jets. *Mon. Not. R. Astron. Soc.* **355**, 1105–1118 (2004).
7. Spada, M., Ghisellini, G., Lazzati, D. & Celotti, A. Internal shocks in the jets of radio-loud quasars. *Mon. Not. R. Astron. Soc.* **325**, 1559–1570 (2001).
8. Fomalont, E. B., Geldzahler, B. J. & Bradshaw, C. F. Scorpius X-1: the evolution and nature of the twin compact radio lobes. *Astrophys. J.* **558**, 283–301 (2001).
9. Migliari, S., Fender, R. & Méndez, M. Iron emission lines from extended X-ray jets in SS 433: reheating of atomic nuclei. *Science* **297**, 1673–1676 (2002).
10. Abell, G. O. & Margon, B. A kinematic model for SS433. *Nature* **279**, 701–703 (1979).
11. Hjellming, R. M. & Rupen, M. P. Episodic ejection of relativistic jets by the X-ray transient GRO J1655 – 40. *Nature* **375**, 464–468 (1995).
12. Mirabel, I. F. & Rodríguez, L. F. A superluminal source in the Galaxy. *Nature* **371**, 46–48 (1994).
13. Dhawan, V., Mirabel, I. F. & Rodríguez, L. F. AU-scale synchrotron jets and superluminal ejecta in GRS 1915+105. *Astrophys. J.* **543**, 373–385 (2000).
14. Cheung, C. C., Harris, D. E. & Stawarz, L. Superluminal radio features in the M87 jet and the site of flaring TeV gamma-ray emission. *Astrophys. J.* **663**, L65–L68 (2007).
15. Marscher, A. P. *et al.* Probing the inner jet of the quasar PKS 1510–089 with multi-wavelength monitoring during strong gamma-ray activity. *Astrophys. J.* **710**, L126–L131 (2010).
16. Crane, P. *et al.* Discovery of an optical synchrotron jet in 3C 264. *Astrophys. J.* **402**, L37–L40 (1993).
17. Baum, S. A. *et al.* HST and Merlin observations of 3C 264—a laboratory for jet physics and unified schemes. *Astrophys. J.* **483**, 178–193 (1997).
18. Perlman, E. S. *et al.* A multi-wavelength spectral and polarimetric study of the jet of 3C 264. *Astrophys. J.* **708**, 171–187 (2010).
19. Lara, L., Giovannini, G., Cotton, W. D., Feretti, L. & Venturi, T. The inner kiloparsec of the jet in 3C 264. *Astron. Astrophys.* **415**, 905–913 (2004).
20. Biretta, J. A., Sparks, W. B. & Macchetto, F. Hubble Space Telescope observations of superluminal motion in the M87 jet. *Astrophys. J.* **520**, 621–626 (1999).
21. Meyer, E. T. *et al.* Optical proper motion measurements of the M87 jet: new results from the Hubble Space Telescope. *Astrophys. J.* **774**, L21 (2013).
22. Stawarz, L. *et al.* Dynamics and high-energy emission of the flaring HST-1 knot in the M 87 jet. *Mon. Not. R. Astron. Soc.* **370**, 981–992 (2006).
23. Lara, L. *et al.* The radio-optical jet in NGC 3862 from parsec to subkiloparsec scales. *Astrophys. J.* **513**, 197–206 (1999).
24. Kobayashi, S., Piran, T. & Sari, R. Can internal shocks produce the variability in gamma-ray bursts? *Astrophys. J.* **490**, 92–98 (1997).

**Supplementary Information** is available in the online version of the paper.

**Acknowledgements** E.T.M. acknowledges HST grant GO-13327.

**Author Contributions** E.T.M. performed the HST data analysis and wrote the paper. B.S., J.A., R.P.v.d.M. and S.T.S. were consulted on and contributed to the HST data analysis. M.G. contributed to the interpretation and performed the theoretical calculations in consultation with E.T.M. E.P. contributed radio data. J.B., C.N. and M.C. provided insight into the design of the observations and interpretation. All authors discussed the results and commented on the manuscript.

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to E.T.M. (eileen.meyer@gmail.com).

## METHODS

**HST data analysis.** These methods are essentially identical to that used in the previous study of M87, and we refer to ref. 21 for more extensive details beyond the outline given here. The imaging used for this study consisted of two orbits of F606W ACS/WFC imaging obtained in 2014 (eight exposures of approximately 600 s each, proposal ID 13327), and the following WFPC2 planetary camera (PC) imaging: nine 140-s exposures in the F702W filter in 2002 (proposal ID 9069), one 350-s and one 400-s exposure taken in the F791 filter in 1996 (proposal ID 5927), and two 140-s exposures in the F702W filter taken in 1994 (proposal ID 5476).

To build the reference frame, in each ACS exposure approximately 100 common globular clusters were detected and localized using a point spread function (PSF) peak-fitting routine. Taking into account the filter-dependent geometric distortion, a linear six-parameter transformation solution was found for each individual ACS exposure based on these positions into a common master frame. The final systematic error in the registration of the ACS exposures is about 0.007 pixels or 0.18 mas. The WFPC2/PC imaging was similarly matched to the master reference frame, using approximately 10–15 bright globular clusters per raw image. The final systematic error in the registration of the WFPC2/PC images is of the order of about 0.14 pixels or 3.5 mas. We include a systematic error of 5 mas for all WFPC2 images and 0.2 mas for the ACS image, added in quadrature to the measured positional errors in each epoch. For each of the four epochs, the transformation solutions were used to stack the exposures into a mean image (25 mas per pixel resolution), correcting for exposure time differences (scaled to the longest exposure), and iteratively sigma-clipping to reject cosmic rays (when more than two images were used). The final image stacks are referred to below by the year of the observations (1994, 1996, 2002, and 2014 images).

The host galaxy of 3C 264 is elliptical, with a light distribution increasing smoothly towards the centre up to the inner 0.9", where a dust disk is clearly present<sup>25</sup> and absorbing some of the galaxy light. The host galaxy light outside the dust disk was modelled individually for the 1996, 2002, and 2014 images using the IRAF (image reduction and analysis facility) software tasks *ellipse* and *bmodel*, and applying a selective mask to avoid the regions of the jet and globular clusters (Extended Data Fig. 6). The 2002 model was also used for the 1994 epoch since the former is deeper. The dust disk was modelled individually for each epoch, by measuring the flux along progressively smaller contour line circles around the centre of the galaxy. The model flux was taken to be a moving average around any individual pixel, interpolating over masked-out areas. The dust disk was modelled up to two pixels (50 mas) from the centre of the galaxy. Inside this radius it is very difficult to reliably separate the core of the jet (a point source) from the galaxy owing to the irregular dust disk, which prevents fitting of the galaxy with a Sérsic or exponential model. Since the core flux level is not of primary interest for this study, we simply take the image values inside this radius as the model, and then subtract the entire model of galaxy and dust disk from our images for each epoch. Finally, we scaled all of the images to flux units (in  $\mu\text{Jy}$ ) at a reference frequency of 6,000 Å using the IRAF task *calcphot*, assuming a power-law form ( $f_\nu \propto \nu^{-1}$ ),  $E(B - V) = 0.02$ , and making the correction for sensitivity changes between decontamination events.

**Proper motions.** For all knots (A–D) and all epochs, we applied the same contour technique to measure the locations of each knot as applied in the M87 study<sup>21</sup>, except that the contour line levels were drawn at 30% flux over background. In the final 2014 epoch, knots B and C were inside a single contour. The knots are also resolved and cannot be easily modelled as point sources to calculate their flux level. To better measure the fluxes of knots and obtain independent locations for the knots, we modelled the jet as a two-dimensional image, focusing only on the region of knots A, B and C. The model, described below, has 14 parameters, and our maximum-likelihood optimization yields an estimate of the knot flux and location relative to the core.

**Maximum-likelihood modelling of the 3C 264 jet.** In all epochs, the model was optimized to match the two-dimensional image over a subset of the image plane centred on knots B and C (Extended Data Fig. 7), covering a total of 226 reference frame pixels, which we call the ‘image box’. We fixed the central jet axis at a position angle (North through East) of 40.43°, running from one end of the image box to the other. The ‘background’ jet is modelled as a cone parameterized by radius  $l_1$  at the distance from the core coinciding with the start of the image box and increasing linearly to  $l_2$  at the other end of the box. The intensity of the jet (in  $\mu\text{Jy}$  per pixel) is characterized by a single value  $I_{\text{jet}}$ . The effective intensity at any given point changes both linearly along the jet (in such a way as to mimic an expanding jet with a constant rate of flow) and laterally from the jet axis, to mimic a decreasing line-of-sight depth through the jet. Given any point inside the defined jet area, we have  $w_{\text{jet}}$  as the radius of the jet at that distance from the core, and  $x_{\text{jet}}$  as the

perpendicular distance of the particular point from the axis. The specific intensity is thus

$$I = I_{\text{jet}} \frac{1}{w_{\text{jet}}} \sqrt{\frac{w_{\text{jet}}^2 - x_{\text{jet}}^2}{w_{\text{jet}}^2}} \quad (1)$$

On top of this jet, we model knots A, B, and C as circular disks parameterized by radii ( $r_A, r_B, r_C$ ), distances from the core ( $d_A, d_B, d_C$ ), and flux densities in  $\mu\text{Jy}$  per pixel ( $I_A, I_B, I_C$ ) which is constant over the disk and zero outside. Knots B and C were also allowed to move in the transverse direction perpendicular to the jet axis. Knot D is not included because the image box does not include this area. The full jet model thus has 14 free parameters. The simulated image is created by adding the contribution of each of the four components to each individual pixel, before applying the relevant PSF.

To accurately fit the jet model to the image data, the simulated image must be convolved with an appropriate PSF, which is different for each image epoch. We used the TinyTim<sup>28</sup> software package to create model PSFs for both WFPC2/PC and ACS/WFC using the appropriate filter to match the observations. We chose a roughly central pixel location—(400, 400) for WFPC2/PC or (1,000, 3,000) for ACS/WFC—standard focus, and a power-law object spectrum with  $F_\nu \approx \nu^{-1}$ , the spectral index of  $-1$  being that previously observed for the knots in the jet<sup>18</sup>. For the WFPC2/PC PSFs, the initial PSF was subsampled by a factor of ten before being rebinned into the WFPC2/PC native PC scale, smoothed with the recommended  $3 \times 3$  pixel Gaussian, and finally resampled onto the 0.025" pixel scale of our stacked images. Optimal values of the 14 free parameters which lead to a simulated image that best fits the data were found by the method of maximum likelihood. We treated each pixel as an independent Poisson trial with a predicted value  $\lambda_i$  and measured value  $x_i$ , where

$$\lambda_i = g \frac{p_i}{a_{\text{scale}}} + b_i \quad (2)$$

and

$$x_i = g \frac{f_i}{a_{\text{scale}}} + b_i \quad (3)$$

Here  $g$  is the gain ( $g = 7$  electrons per DN, where DN is data number, for the WFPC2/PC images,  $g = 1$  for the ACS/WFC images) and  $p_i$  is the predicted flux in the simulated image pixel. The constant  $a_{\text{scale}}$  is the scaling factor, specific to each epoch, which was originally used to multiplicatively scale the images from units of DN or electrons to flux in  $\mu\text{Jy}$ , while  $b_i$  is the (unscaled) background estimation that was subtracted from the images before scaling. The same equation is used to calculate  $x_i$  except that we use the measured flux in each pixel  $f_i$  rather than the simulated prediction.

The maximum-likelihood method requires that we maximize the likelihood function, which is simply the product of the probability density values for each pixel. Because the number of total counts is large (typically  $\gg 1,000$ ) and in order to accurately calculate the likelihood contribution for each pixel, we use the normal approximation to the Poisson probability distribution in calculating the likelihood. We also chose to maximize the log of the likelihood function, which is equivalent to maximizing the likelihood but more practical to implement. Given the above definitions of  $\lambda_i$  and  $x_i$  for the  $i$ th pixel in the image box, the contribution to the likelihood is simply

$$pr_i = \frac{1}{\sqrt{2\pi\lambda_i/n}} e^{-\frac{(x_i - \lambda_i)^2}{2\lambda_i/n}} \quad (4)$$

where  $n$  is the total number of images in the stack. The final log-likelihood function to be maximized is thus

$$\mathcal{L} = \sum_i \log(pr_i)^{1/f} \quad (5)$$

where  $f$  is the correction for the different pixel scales between WFPC2/PC and ACS/WFC and our super-sampled reference frame described in the methods, with  $f = 3.3856$  for WFPC2/PC and  $f = 3.9204$  for ACS/WFC.

Although a number of approaches are possible for finding the optimal values of the 14 free parameters that will maximize  $\mathcal{L}$ , we chose to use the freely available Simulated Annealing code<sup>26</sup>, which is well-suited to multi-parameter optimizations of non-smooth functions that are likely to have many false maxima ‘traps’. All 14 parameters were allowed to roam inside a wide range of reasonable values, and we confirmed that the final optimal value did not show signs of running into upper or lower bounds.

The final results of the jet modelling are given in Extended Data Table 1. We have used the data for knots B and C only in the main part of the paper, as the subtraction of the background on the inner side of knot A is somewhat unreliable due to the nearness of the core, leading us to place one edge of the image box over the inner side of knot A. Knot A was thus included in the model more for

improving the fit of the other components than for measurement. In the main text, location values for knots A and D are taken from the contour method, while for knots B and C are taken from the best-fit model (radial positions only). Note that the transverse motions are only marginally significant once the systematic registration error is included, except in the final epoch where both knots are slightly deviated to the North.

Errors for each parameter were calculated by freezing all other parameters to their best-fit values, and progressively increasing (or decreasing) the value of the parameter of interest. This allowed us to plot the change in the likelihood function  $\Delta\mathcal{L}$  versus the parameter value. The shape of this function is roughly parabolic with a peak at the optimal parameter value. The distribution of  $2\Delta\mathcal{L}$  is a  $\chi^2$  variable with one degree of freedom. Thus an upper and lower error value can be calculated by finding the parameter value corresponding to a  $3\sigma$  value of  $2\Delta\mathcal{L}$ . The errors reported in Extended Data Table 1 are the  $3\sigma$  errors. Since the errors are two-sided, the criterion used was that  $2\Delta\mathcal{L} = 10.27288$ . Fluxes were calculated by multiplying the flux density best-fit value by  $\pi r^2$  where  $r$  is the best-fit radius, and appropriately propagating the errors.

The flux increase implied by the modelling results can be checked using aperture photometry. We used a 6.5-pixel circular aperture, centred on knots B and C so that both are almost entirely contained within the aperture during all four epochs without moving it. Using the 2014 modelling results, we estimate that the smooth-jet contribution to the flux in this aperture is 17.4  $\mu\text{Jy}$ . The total fluxes attributed to the jet for the four epochs are 7.9  $\mu\text{Jy}$ , 7.8  $\mu\text{Jy}$ , 6.7  $\mu\text{Jy}$ , and 9.4  $\mu\text{Jy}$ , indicating a 20%–40% flux change in the final epoch. The estimated error on these fluxes is approximately 0.3  $\mu\text{Jy}$ , where errors from the Poisson noise (0.2  $\mu\text{Jy}$  and 0.05  $\mu\text{Jy}$  for WFPC2 and ACS respectively), irregularities in the dust disk (0.15  $\mu\text{Jy}$ ), and possible contributions of knot A to the aperture (0.2  $\mu\text{Jy}$ ) have been added in quadrature. Note that these estimates do not correct for the flux which falls outside the aperture, and are thus slightly less than the totals for knots B and C taken from Extended Data Table 1.

**The extended baseline with VLA observations.** 3C 264 was observed in October 1983 at U band (14.9 GHz) in A configuration as part of program AB239. The quasar 3C 286 was used as the primary flux density calibrator source while QSO 1144+402 was used for a primary phase calibrator and secondary flux density calibrator. The observations used two intermediate frequencies in each band, each with a bandwidth of 50 MHz, and the average of the two was used to produce the final maps. The reduction of the 1983 U-band data followed closely the recommended procedure in chapter 4 of the AIPS Cookbook (<http://www.aips.nrao.edu>). Very little flagging was required. Once calibrated, a clean map was made with several rounds of self-calibration, and deconvolution using the AIPS tasks *clean* and *calib*. We used the ‘robust’ or Briggs’ weighting scheme to weight the visibilities.

The earlier radio image clearly shows three distinct features (see contour comparison in Extended Data Fig. 4). In the case of knot B, a quadratic fit to the HST and VLA data suggests that in late 1983 the knot had a speed of 10.2c, slowing by 0.16c per year to reach 5.6c at the beginning of 2014. While a quadratic fit is very poor for the combined data on knot C, a linear fit between the 1983 and 1994 epochs similarly suggests a speed of 9.8c. Strong conclusions are prevented by the disparity in imaging frequencies and the large errors on the radio positions.

**Theoretical considerations.** We assume that the flow consists of an electron–proton plasma, with one proton per electron, in energy equipartition with the comoving magnetic field. We derive the equipartition magnetic field values at knots B and C using standard expressions including beaming<sup>27</sup>. We use the 2002 flux values at 22.46 GHz of 4.4 mJy and 4.1 mJy, and diameters of 39.6 pc and 48.9 pc, respectively, for knots B and C, as well as the observed radio spectral index of  $\alpha = 0.5$ . We assume that the knots are spherical, noting that the measured diameter in the comoving frame for a spherical blob equals that in the observer’s frame. We assume a low-frequency cutoff of 10 MHz in the observer’s frame. We

derive values of  $1.2/\delta$  mG and  $1.0/\delta$  mG (where  $\delta$  is the Doppler factor) for knots B and C respectively. These values are nearly identical to the values given in ref. 23 once corrected for beaming (they assumed a jet angle to the line of sight  $\theta = 50^\circ$ , which has now been ruled out by the superluminal motions we observe).

We calculated masses  $m_B = 2.5 \times 10^{30}$  g and  $m_C = 6.7 \times 10^{30}$  g for knots B and C respectively, through  $m = 2U_B V/c^2$ , where  $U_B$  is the energy density of the magnetic field in equipartition and  $V$  is the volume of the knot in the comoving frame. Using these masses and the Lorentz factors  $\Gamma_B = 7.1$ ,  $\Gamma_C = 2.8$  (as derived by the relation  $\Gamma_{B,\text{min}} = (\beta_{B,\text{app}}^2 + 1)^{1/2}$ , and assuming that  $\Gamma_B = \Gamma_{B,\text{min}}$ , which sets the jet angle  $\theta$  to the line of sight to  $\arcsin \theta = 1/\Gamma_B = 8.1^\circ$ ), the Lorentz factor  $\Gamma_m$  of the single component resulting from the collision is<sup>27</sup>

$$\Gamma_m = \left( \frac{m_B \Gamma_B + m_C \Gamma_C}{m_B/\Gamma_B + m_C/\Gamma_C} \right)^{1/2} = 3.7 \quad (6)$$

as found by conservation of momentum. The energy that will be dissipated cumulatively throughout the collision is

$$E_{\text{diss}} = m_B(\Gamma_B - \Gamma_m) + m_C(\Gamma_C - \Gamma_m) = 2.6 \times 10^{51} \text{ erg} \quad (7)$$

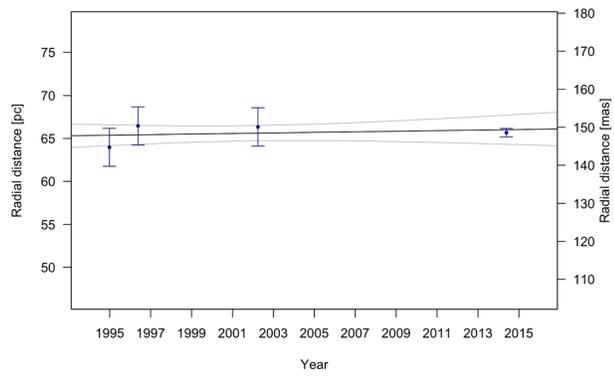
Assuming that the collision lasts from the moment the collision commences up to the point where the back end of the fast component has reached the back end of the slow component, the duration of the collision in terms of coordinate time in the galaxy frame is  $\Delta t = 2R'/(L_B(\beta_B - \beta_C)c) \approx 340$  years, where  $R'$  is the comoving radius of the fast component, while the duration for the observer is  $\Delta t_{\text{obs}} = 2R'/(L_B \Gamma_C \delta_C (\beta_B - \beta_C)c) \approx 25$  years, in approximate agreement with the 30-year time required for the knots to superimpose based on the observed HST proper motions. The power dissipated in the galaxy frame is then  $P_{\text{diss}} = E_{\text{diss}}/\Delta t = 2.4 \times 10^{41} \text{ erg s}^{-1}$ . If the dissipated energy was completely converted to radiation, the solid-angle integrated luminosity in the galaxy frame  $L_{4\pi}$  would be equal to  $P_{\text{diss}}$ . We can estimate the current  $L_{4\pi}$  due to the collision by using the flux increase  $f_{\text{coll}}$  in knots B and C from 2002 to 2014, to obtain the luminosity increase assuming isotropy  $L = 1.3 \times 10^{40} \text{ erg s}^{-1}$ , which corresponds to a solid-angle integrated luminosity  $L_{4\pi} = L f_{\text{coll}}^2 \delta_m^4 = 1.6 \times 10^{38} \text{ erg s}^{-1}$ . The efficiency of converting the available energy to radiation is then  $\eta = L_{4\pi}/P_{\text{diss}} \approx 10^{-3}$ ; as explained in the main text, this is a lower limit.

**The cooling length of a knot.** Using standard formulae, the comoving frame cooling time at an observed frequency  $\nu$  of a knot in equipartition is

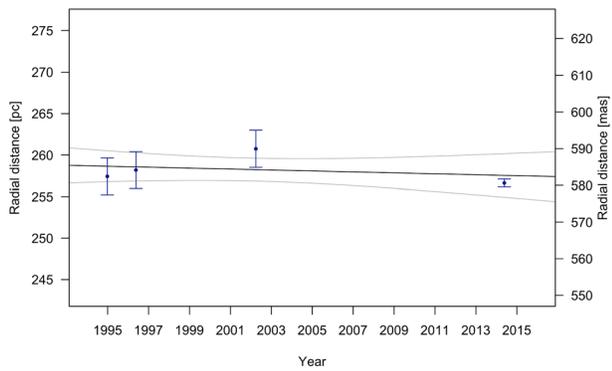
$$t' = \frac{6\pi m_e c}{\sigma_T B_{\text{eq},0}^2} \left( \frac{m_e c^2 B_{\text{eq},0}}{h B_{\text{cr}} \nu} \right)^{1/2} \delta^2 = t_0 \delta^2 \quad (8)$$

where  $B_{\text{eq},0}$  is the equipartition magnetic field assuming  $\delta = 1$ ,  $m_e$  is the electron mass,  $\sigma_T$  is the Thomson cross-section,  $h$  is Planck’s constant,  $B_{\text{cr}} = 2\pi m_e^2 c^3/(eh) = 4.4 \times 10^{13}$  G is the critical magnetic field and  $e$  is the electron charge. This corresponds to a propagation length of  $l = c\beta\Gamma t' = c\beta\Gamma \delta^2 t_0$ , which in turn corresponds to a projected cooling length  $l_{\text{obs}} = l \sin\theta = t_0 c \delta \beta_{\text{app}}$ , where we have used the standard expression  $\beta_{\text{app}} = \beta \sin\theta/(1 - \beta \cos\theta)$ . Using values appropriate for knot B we obtain  $l_{\text{obs}} = 874$  pc and for knot C  $l_{\text{obs}} = 104$  pc, in both cases longer than the propagation length recorded by our observations (about 43 pc for knot B and about 11 pc for knot C).

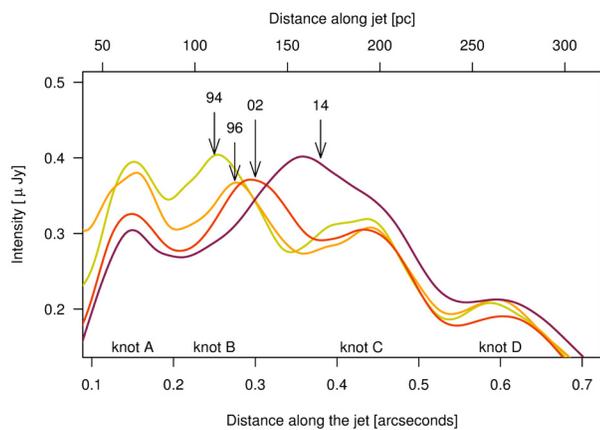
25. Sparks, W. B., Baum, S. A., Biretta, J., Macchetto, F. D. & Martel, A. R. Face-on dust disks in galaxies with optical jets. *Astrophys. J.* **542**, 667–672 (2000).
26. Goffe, W. L., Ferrier, G. D. & Rogers, J. Global optimization of statistical functions with simulated annealing. *J. Econ.* **60**, 65–99 (1994).
27. Harris, D. E. & Krawczynski, H. X-ray emission processes in radio jets. *Astrophys. J.* **565**, 244–255 (2002).
28. Krist, J. E., Hook, R. N. & Stoehr, F. 20 years of Hubble Space Telescope optical modeling using Tiny Tim. *Proc. SPIE* **8127** (2001).



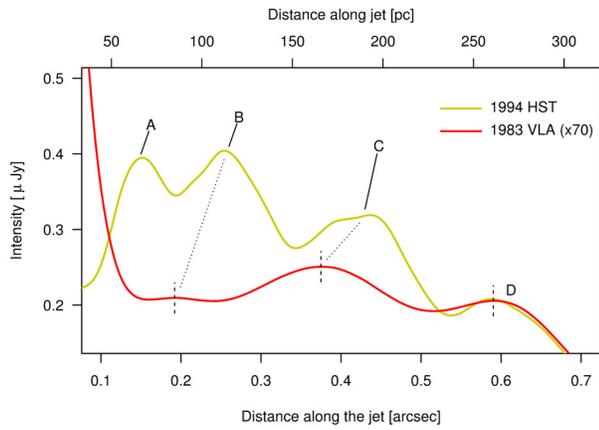
**Extended Data Figure 1 | Position versus year for knot A.** Positions are taken from the contour analysis method. Errors are  $1\sigma$  from contour-derived position measurement added to the systematic error on the mean position from the image registration in quadrature.



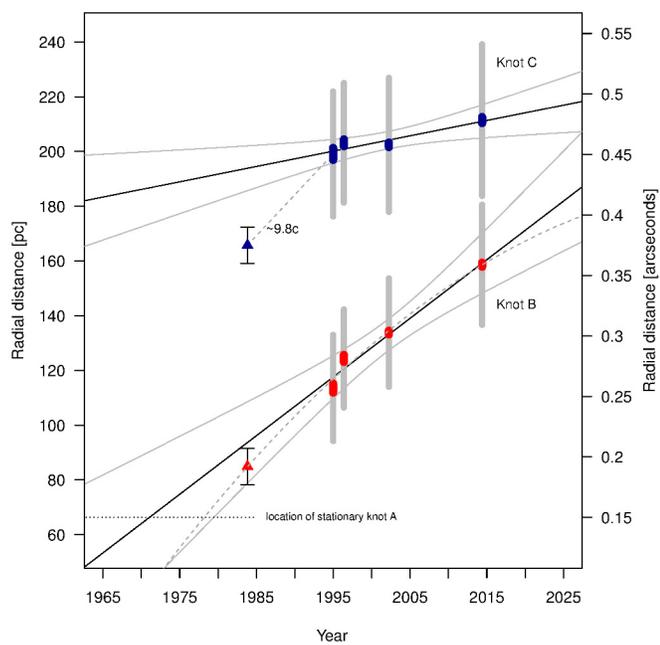
**Extended Data Figure 2 | Position versus year for knot D.** Positions are taken from the contour analysis method. Errors are  $1\sigma$  from contour-derived position measurement added to the systematic error on the mean position from the image registration in quadrature.



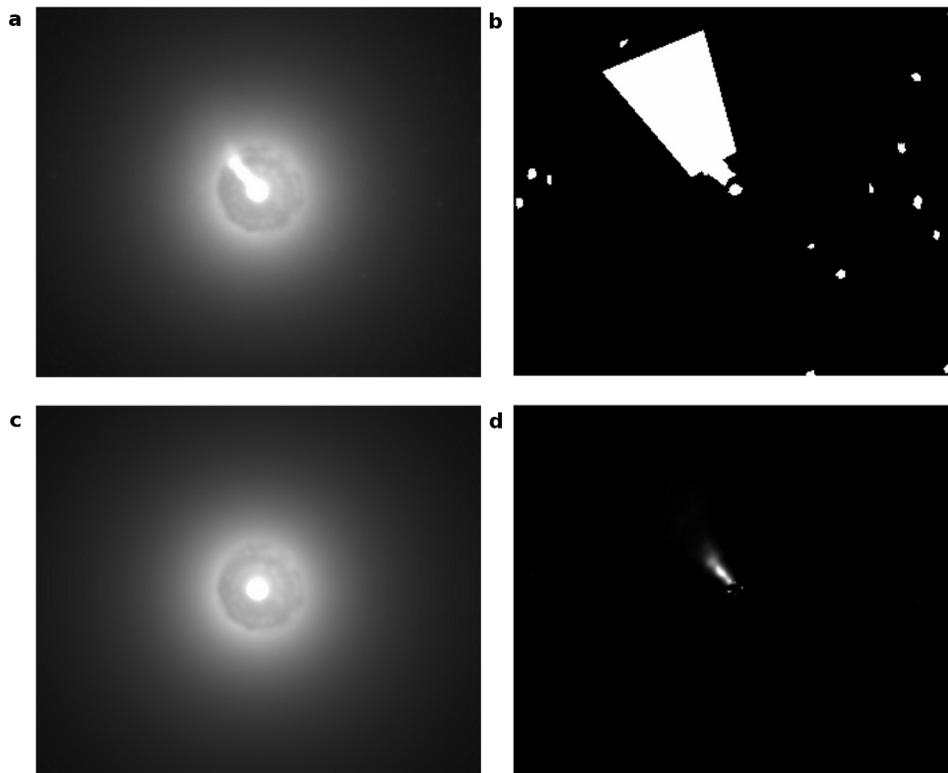
**Extended Data Figure 3 | Mean flux along the midline of the jet.** Flux is averaged transversely over  $0.1''$  at each step along the jet. The epoch is denoted by arrows. Note the decay of knot A with time, while the knot B+C complex appears substantially brighter in the last epoch.



**Extended Data Figure 4 | Comparison of 1983 VLA midline flux to 1994 HST midline flux.** Flux is averaged transversely over  $0.1''$  at each step along the jet. We fit parabolic forms to the peaks seen in the radio contour to derive rough estimates of knot location, with values depicted with short dashed lines. Dotted lines connecting matched knots are shown to guide the eye.

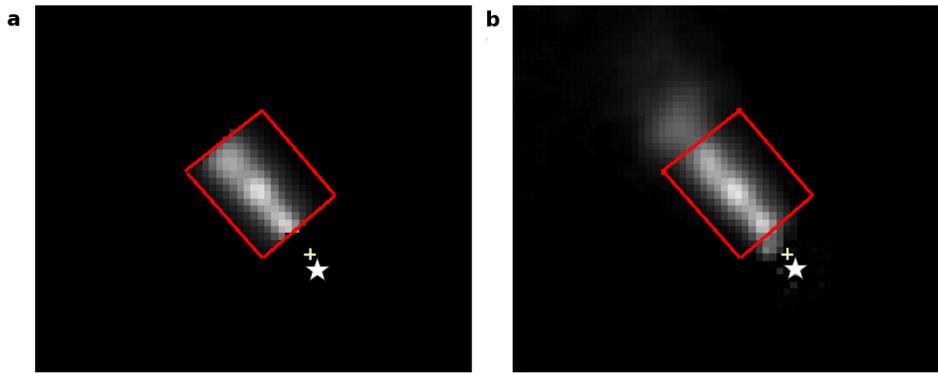


**Extended Data Figure 5 | Comparison of 1983 radio positions of knots with HST data.** The locations of knots B and C in 1983 are denoted by triangles, and we include an estimated 15 mas error on the position based on a typical centroiding error of 10% of the beam size (0.15"). Optical data are identical to that described in Fig. 2. For knot C, a dashed line gives the fit to epochs 1983 and 1994. For knot B, the dashed line is a parabolic fit to all data. The location of knot A is noted with a dotted line.



**Extended Data Figure 6 | Depiction of background subtraction method.**  
**a**, The final 2014 image stack before subtraction. **b**, The mask used in modelling the galaxy and dust disk. (White areas were not used in fits.) **c**, The final 2014

galaxy and dust model. **d**, The same image as **a** after **c** has been subtracted. Physical scaling of all images and flux scaling of **a**, **b** and **d** are identical.



**Extended Data Figure 7 | Depiction of modelling results.** **a**, The resulting background-subtracted model image for the 2002 epoch. **b**, The real background-subtracted 2002 image. In both cases the image box is shown by the red outline. Only pixels within this area were used in the fit.

Extended Data Table 1 | Maximum-likelihood modelling results

Name	Year	Radial Position (pixels)	Transverse Position (pixels)	Intensity ( $\mu\text{Jy}/\text{pixel}$ )	Radius (pixels)	Radial Position* (mas)	Total Flux ( $\mu\text{Jy}$ )	Size (pc)
Knot A <sup>†</sup>	1994	$5.45 \pm 0.13$	...	$0.65 \pm 0.04$	$1.28 \pm 0.04$	$136 \pm 5$	$3.3 \pm 0.3$	$28.3 \pm 1.0$
	1996	$4.95 \pm 0.12$	...	$0.43 \pm 0.02$	$2.01 \pm 0.05$	$124 \pm 5$	$5.5 \pm 0.3$	$44.4 \pm 1.0$
	2002	$4.89 \pm 0.09$	...	$0.24 \pm 0.01$	$2.05 \pm 0.04$	$122 \pm 4$	$3.1 \pm 0.1$	$45.3 \pm 0.9$
	2014	$4.97 \pm 0.12$	...	$0.47 \pm 0.02$	$1.42 \pm 0.03$	$124 \pm 3$	$3.0 \pm 0.1$	$31.4 \pm 0.6$
Knot B	1994	$10.28 \pm 0.15$	$0.45 \pm 0.13$	$0.37 \pm 0.02$	$1.76 \pm 0.05$	$257 \pm 5$	$3.6 \pm 0.3$	$38.8 \pm 1.2$
	1996	$11.26 \pm 0.11$	$-0.32 \pm 0.11$	$0.42 \pm 0.02$	$1.62 \pm 0.04$	$281 \pm 5$	$3.5 \pm 0.2$	$35.9 \pm 1.0$
	2002	$12.11 \pm 0.06$	$0.28 \pm 0.05$	$0.41 \pm 0.01$	$1.79 \pm 0.02$	$303 \pm 4$	$4.2 \pm 0.1$	$39.6 \pm 0.5$
	2014	$14.35 \pm 0.07$	$-0.70 \pm 0.05$	$0.42 \pm 0.01$	$1.98 \pm 0.02$	$359 \pm 2$	$5.3 \pm 0.1$	$43.9 \pm 0.5$
Knot C	1994	$18.02 \pm 0.19$	$-0.08 \pm 0.15$	$0.24 \pm 0.02$	$2.06 \pm 0.07$	$450 \pm 6$	$3.2 \pm 0.3$	$45.6 \pm 1.6$
	1996	$18.39 \pm 0.11$	$-0.41 \pm 0.09$	$0.36 \pm 0.01$	$1.98 \pm 0.04$	$460 \pm 4$	$4.5 \pm 0.2$	$43.7 \pm 0.9$
	2002	$18.32 \pm 0.07$	$-0.28 \pm 0.06$	$0.27 \pm 0.01$	$2.21 \pm 0.03$	$458 \pm 4$	$4.2 \pm 0.1$	$48.9 \pm 0.6$
	2014	$19.14 \pm 0.10$	$-0.56 \pm 0.07$	$0.28 \pm 0.01$	$2.51 \pm 0.03$	$478 \pm 2$	$5.5 \pm 0.1$	$55.4 \pm 0.6$

\* Systematic error of 3.5 (0.18) mas for WFPC2 (ACS) images has been added to the measurement error from column 3. † Knot A values are included for completeness; see Methods for note on reliability of knot A positions.