Bayesian forward modeling of high-resolution radio interferometric gravitational lens observations

Devon Powell

June 3, 2021

SDSS 2021

Collaborators:
Simona Vegetti, Francesca Rizzo, Hannah Stacey (Max Planck Inst. for Astrophysics)
John McKean (Groningen/ASTRON), Cristiana Spingola (Bologna/INAF)

Publication:
Powell et al., "A novel approach to visibility-space modelling of interferometric gravitational lens observations at high angular resolution", *MNRAS* 2021.

Motivation

Science with strong lenses

- Detection of low-mass dark matter haloes, constraints on mass function

  e.g. Ritondale et al. (2019), Despali et al. (2018), Birrer et al. (2017), Hezaveh et al. (2016), Vegetti et al. (2014), Vegetti et al. (2012), Vegetti et al. (2010)

- Source science taking advantage of magnification

  e.g. Spingola et al. (2019), Rizzo et al. (2018), Johnson et al. (2017), Leethochawalit et al. (2016), Swinbank et al. (2015)

- Cosmology using time delays between images

  e.g. Rusu et al. (2019), Suyu et al. (2018), Wong et al. (2017), Treu and Marshall (2016), Chen et al. (2016), Courbin et al. (2011), Fassnacht et al. (2002)

- Polarimetry and rotation measure synthesis

  e.g. Mao et al. (2017)

Lots of science with lenses - review in Koopmans et al. (2009).

Science with strong lenses

- **Detection of low-mass dark matter haloes, constraints on mass function**

  e.g. Ritondale et al. (2019), Despali et al. (2018), Birrer et al. (2017), Hezaveh et al. (2016), Vegetti et al. (2014), Vegetti et al. (2012), Vegetti et al. (2010)

- Source science taking advantage of magnification

  e.g. Spingola et al. (2019), Rizzo et al. (2018), Johnson et al. (2017), Leethochawalit et al. (2016), Swinbank et al. (2015)

- Cosmology using time delays between images

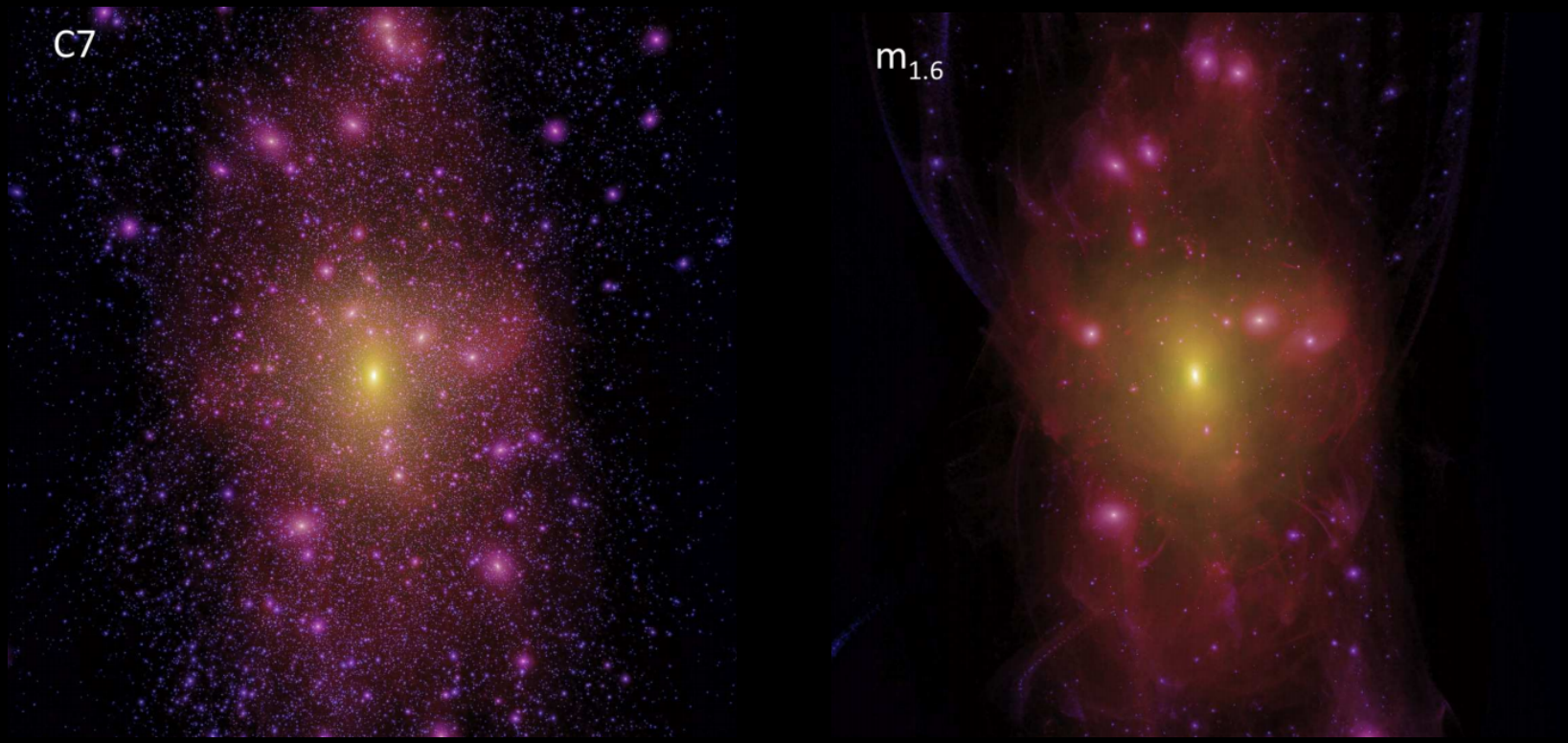
  e.g. Rusu et al. (2019), Suyu et al. (2018), Wong et al. (2017), Treu and Marshall (2016), Chen et al. (2016), Courbin et al. (2011), Fassnacht et al. (2002)

- Polarimetry and rotation measure synthesis

  e.g. Mao et al. (2017)

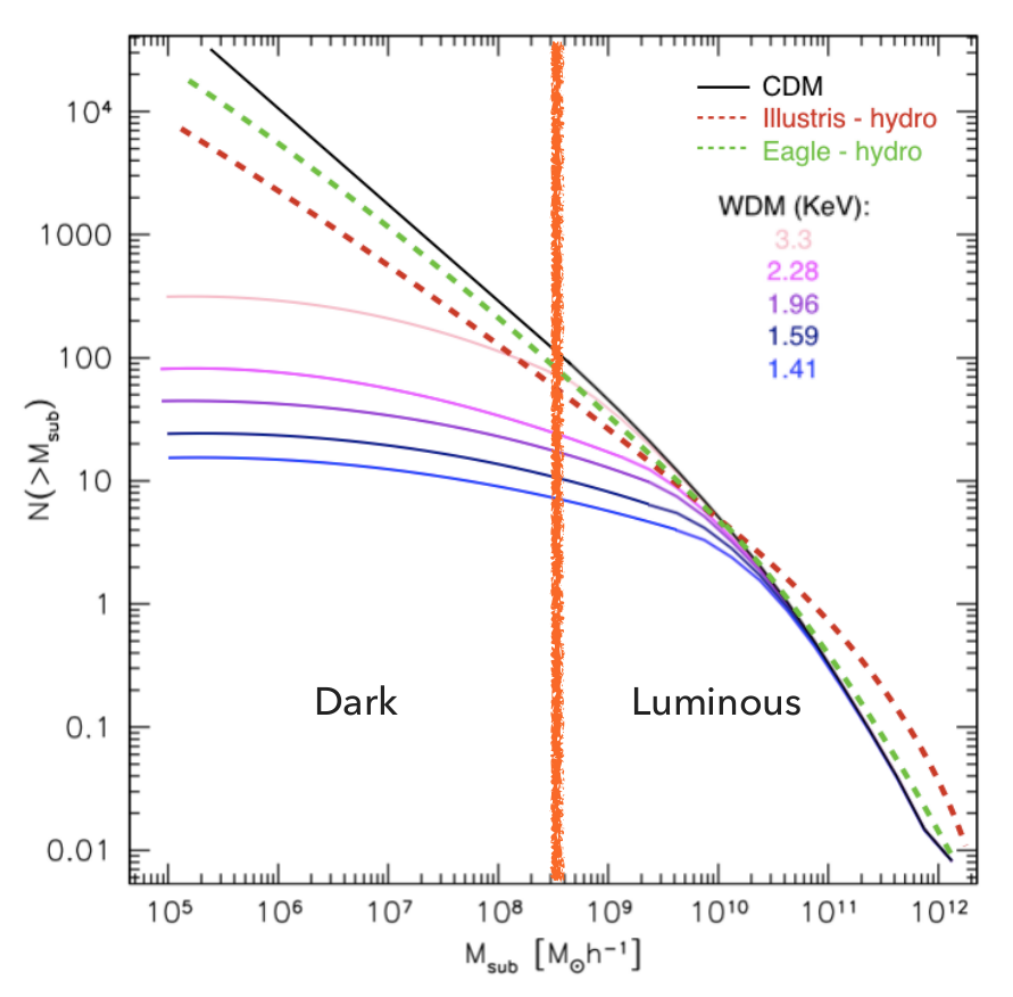
Lots of science with lenses - review in Koopmans et al. (2009).

Mass function: CDM vs. WDM



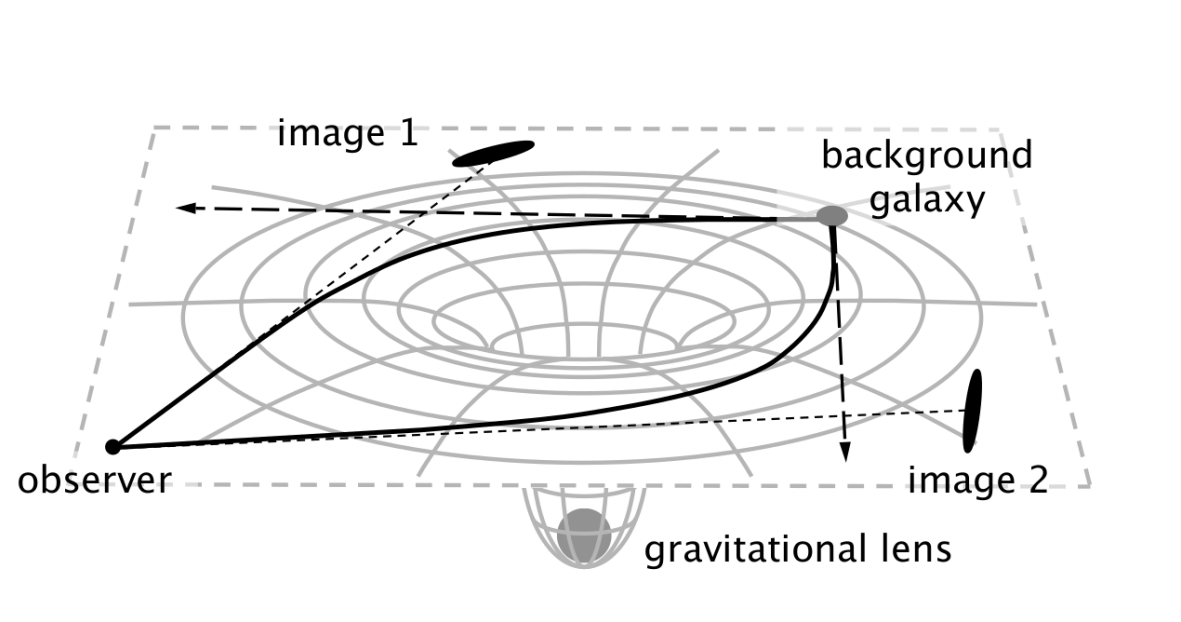
Lovell et al. (2014)

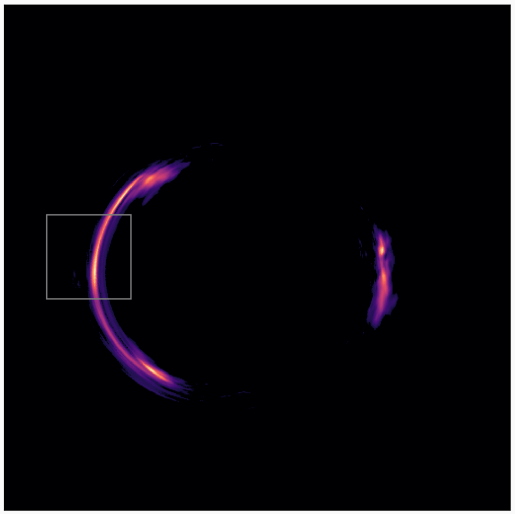
Mass function: CDM vs. WDM

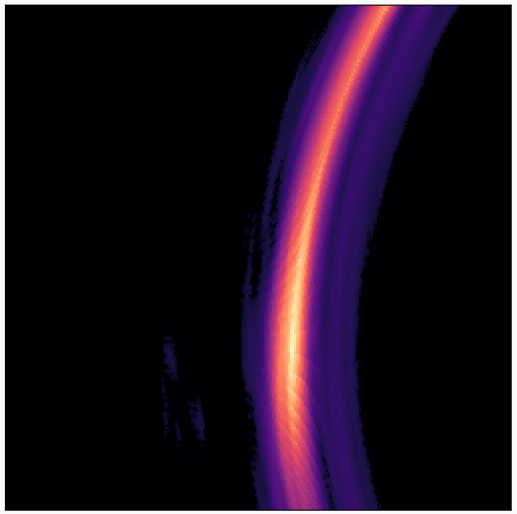


Despali and Vegetti (2017), Lovell et al. (2014)

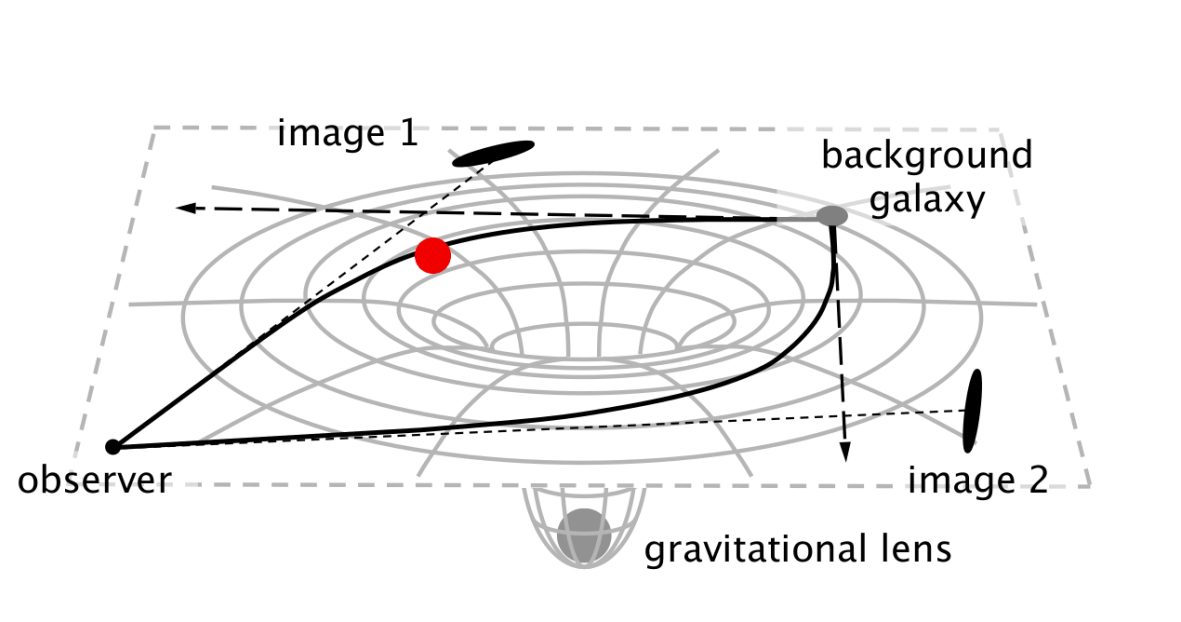
LoW-Mass Perturbers (LwMPs) with lensing

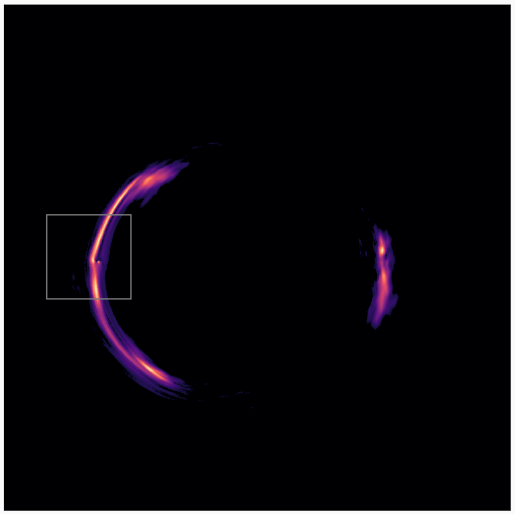


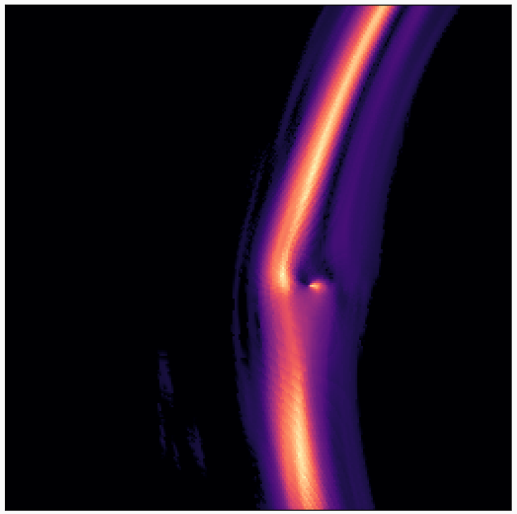




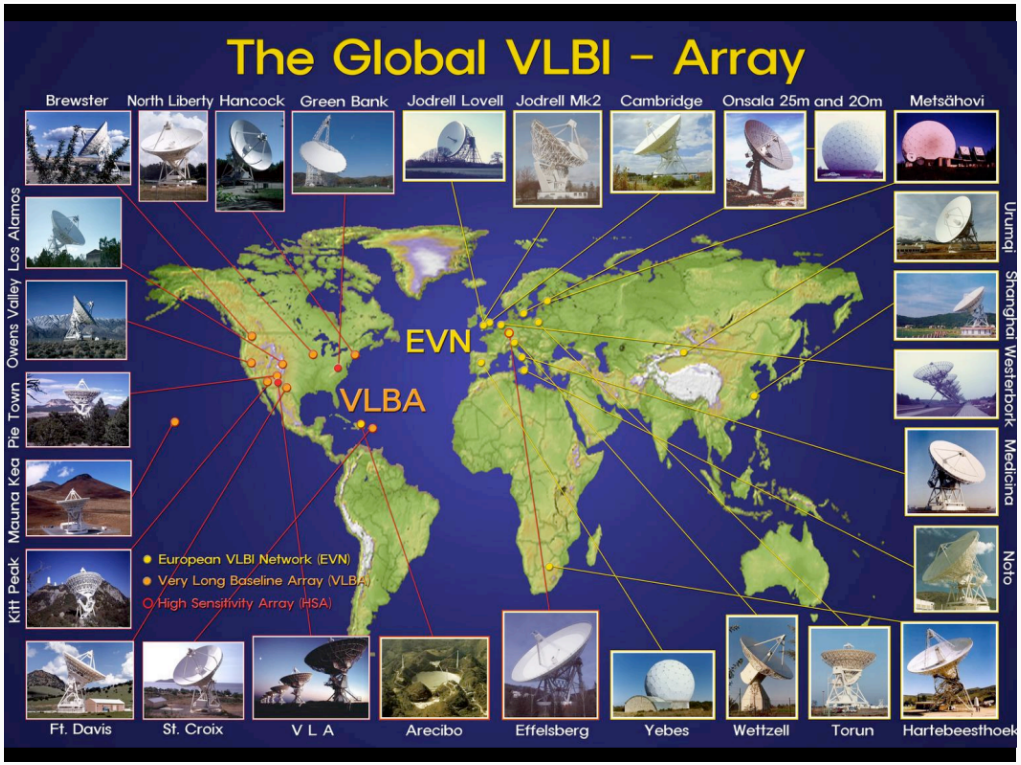
LoW-Mass Perturbers (LwMPs) with lensing







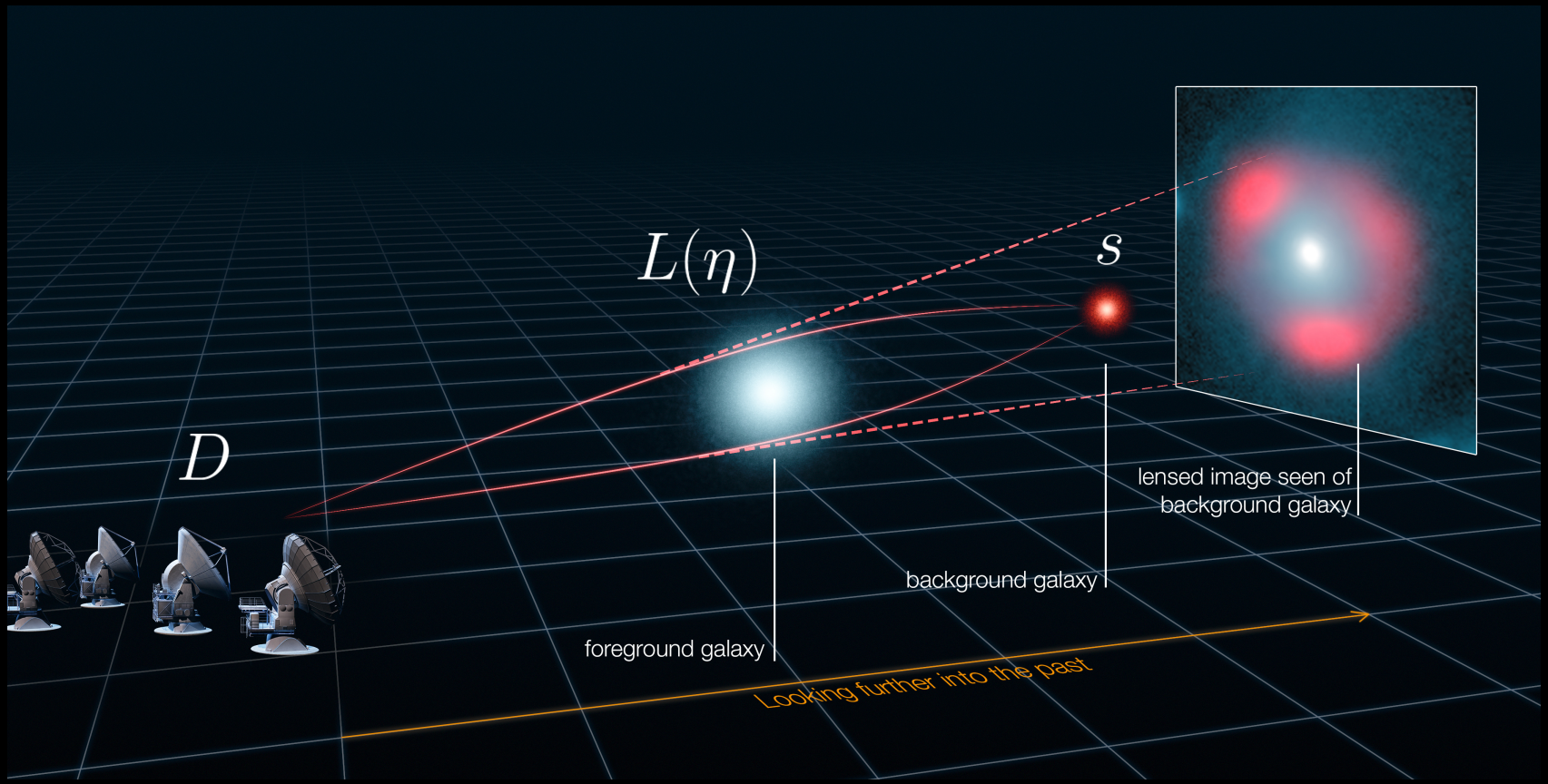
Global VLBI



- GVLBI array (Global Very Long Baseline Interferometry)
- Sub-milliarcsec resolution at shorter wavelengths
- Huge numbers of visibilities (up to $\sim 10^{10}$). A challenge for modeling.

Modeling

Forward modeling



ALMA (ESO/NRAO/NAOJ), L. Calada (ESO), Y. Hezaveh et al.

Bayesian inference

$$\boxed{P(\mathbf{s} \mid \mathbf{d}, \eta, \lambda) = \frac{P(\mathbf{d} \mid \mathbf{s}, \eta)\, P(\mathbf{s} \mid \lambda)}{P(\mathbf{d}|\eta, \lambda)}}$$

- $\mathbf{d}$ is the data (complex radio visibilities)
- $\mathbf{s}$ is the source plane surface brightness
- η is a vector containing lens parameters (mass, slope, ellipticity, LwMPs...)
- λ is a *hyper-parameter* setting the strength of the regularization (fixed for MAP)

Multi-level (hierarchical) Bayes

We use a three-step inference process:

1. Optimize for linear parameters (source inversion)

$$\text{Max} \boxed{P(\mathbf{s} \mid \mathbf{d}, \eta, \lambda) = \frac{P(\mathbf{d} \mid \mathbf{s}, \eta)\, P(\mathbf{s} \mid \lambda)}{P(\mathbf{d}|\eta, \lambda)}}$$

2. Optimize for non-linear parameters (lens model and regularization)

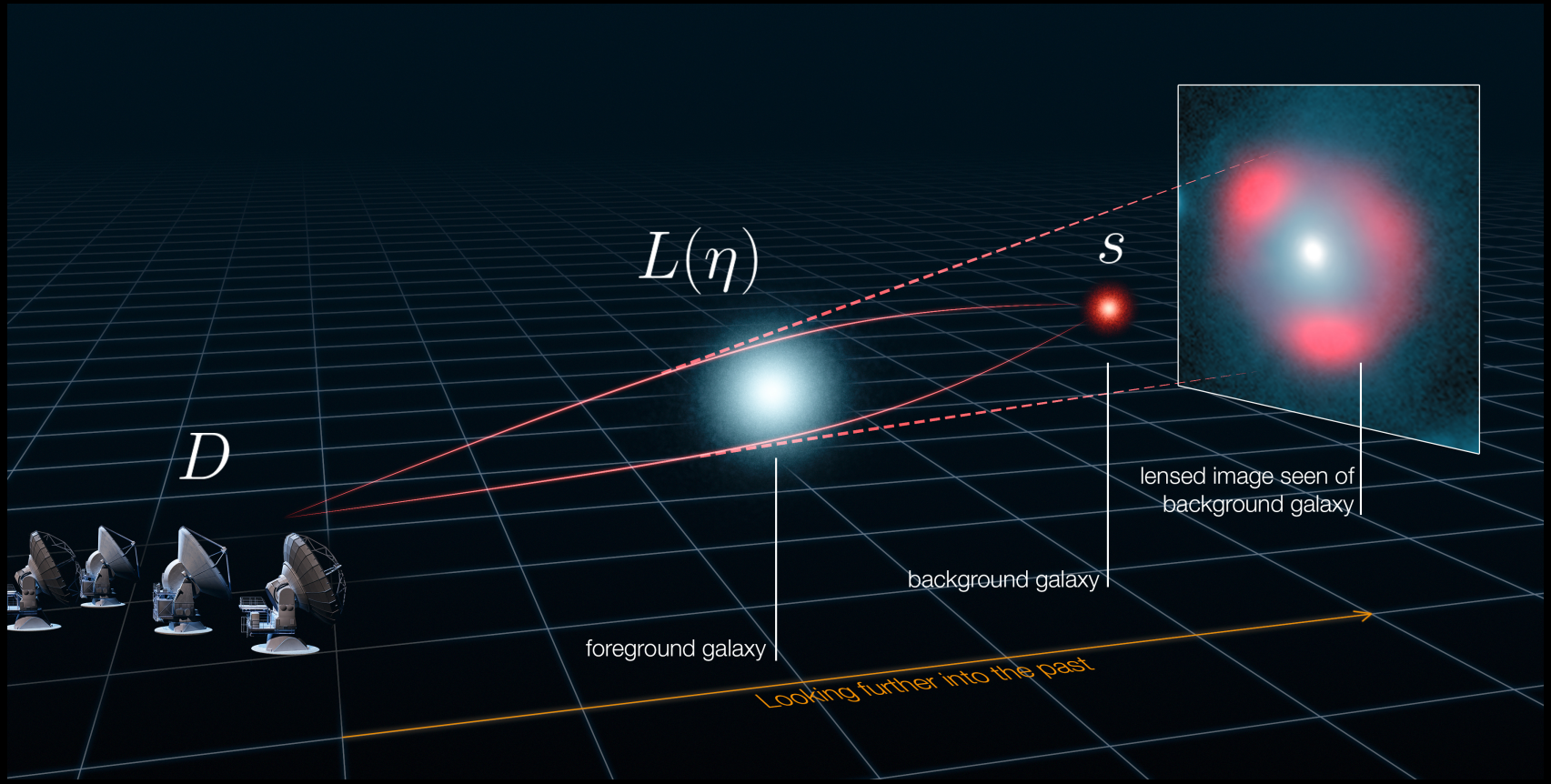
$$\text{Max} \boxed{P(\eta, \lambda \mid \mathbf{d}) = \frac{P(\mathbf{d} \mid \eta, \lambda)\, P(\eta, \lambda)}{P(\mathbf{d})}}$$

3. Evidence computation and model comparison

$$\text{Max} \boxed{P(\mathbf{d}) = \int P(\mathbf{d}|\eta, \lambda)\, P(\eta, \lambda)\, \mathrm{d}\lambda\, \mathrm{d}\eta}$$

See Vegetti and Koopmans (2009) for the specific implementation, MacKay (1991) for a theoretical overview.

Forward modeling



ALMA (ESO/NRAO/NAOJ), L. Calada (ESO), Y. Hezaveh et al.

First level of inference: Source inversion

Maximizing the log-posterior with respect to s implies

$$\frac{\partial}{\partial s}\left[(D\,L\,s - d)^T\,C_d^{-1}\,(D\,L\,s - d) + s^T\,R\,s\right] = 0$$

giving

$$\boxed{(L^T D^T C_d^{-1}\,D\,L + R)\,s = L^T D^T C_d^{-1}\,d}$$

Source inversion (when the operators are matrices)

$$(L^T D^T C_d^{-1} D L + R)\, s = L^T D^T C_d^{-1}\, d$$

Vectors:

- s - Reconstructed source solution
- d - Data (radio visibilities)

Matrices:

- L - Lensing operator
- C_d^{-1} - Visibility-space noise covariance
- R - Regularization matrix
- D - Nonuniform discrete Fourier transform

Source inversion (when the operators are functions)

Replace discrete FT with non-uniform FFT:

$$D \rightarrow G\,F\,W$$

- G - Gridding operator
- F - FFT
- W - Apodization operator (inverse FT of the gridding kernel)

Source inversion (when the operators are functions)

$$(L^T W^T F^T G^T C_d^{-1} G F W L + R)\, s = L^T W^T F^T G^T C_d^{-1} d$$

Vectors:

- s - Reconstructed source solution
- d - Data (radio visibilities)

Explicit matrices:

- L - Lensing operator
- C_d^{-1} - Visibility-space noise covariance
- R - Regularization matrix
- W - Apodization operator (inverse FT of the gridding kernel)

Source inversion (when the operators are functions)

$$(L^T W^T F^T G^T C_d^{-1} G F W L + R) s = L^T W^T F^T G^T C_d^{-1} d$$

Implicitly defined linear operators ("Matrix-Free"):

- G - Visibility gridding operator (implemented on GPU for speed)
- F - FFT

The Fourier Transform is a dense matrix. Large numbers of visibilites ($> 10^8$) demand that we use a gridding operation + FFT instead of direct DFT.

Computational challenges

For space, denote:

$$A \equiv (L^T W^T F^T G^T C_d^{-1} G F W L + R)$$

and

$$b \equiv L^T W^T F^T G^T C_d^{-1} d$$

- A is not explicitly defined. It exists only as a function for matrix-vector multiplication (linearly mapping vectors to vectors)
- We must solve $A s = b$
- We must also compute $\log \det A$

Preconditioned conjugate gradient solver

Since A is not explictly defined, we must use an iterative solver. Conjugate gradient seems to work best for this problem. However, we *must* precondition to get reasonably fast convergence, due to extremely high condition number (can be 10^{10}).

Instead of $A\,s = b$, we solve

$$P^{-1}\,As = P^{-1}b$$

We must design the preconditioner P such that it is a sparse matrix which approximates A reasonably well.

Preconditioned conjugate gradient solver

Our preconditioner is based on the image-plane noise covariance:

$$C_x^{-1} = W^T F^T G^T C_d^{-1} G F W$$

One can show that each row of this matrix simply contains the naturally-weighted dirty beam. We approximate

$$\tilde{C}_x^{-1} = \mathrm{diag}(C_x^{-1})$$

which is a diagonal matrix containing the brightest pixel in the dirty beam. Using only the diagonal works well and guarantees positive-definiteness of P.

Our full preconditioner is then

$$P = L^T \tilde{C}_x^{-1} L + R$$

which is sparse and can be decomposed for fast application of P^{-1}.

Computational challenges

For space, denote:

$$A \equiv (L^T W^T F^T G^T C_d^{-1} G F W L + R)$$

and

$$b \equiv L^T W^T F^T G^T C_d^{-1} d$$

- A is not explicitly defined. It exists only as a function for matrix-vector multiplication (linearly mapping vectors to vectors)
- We must solve $A s = b$
- We must also compute $\log \det A$

Second level of inference: Computing the evidence

In optimizing the lensing operator L, we compute the Bayesian evidence. Correct normalization of this evidence requires us to compute $\mathrm{logdet}(A)$.

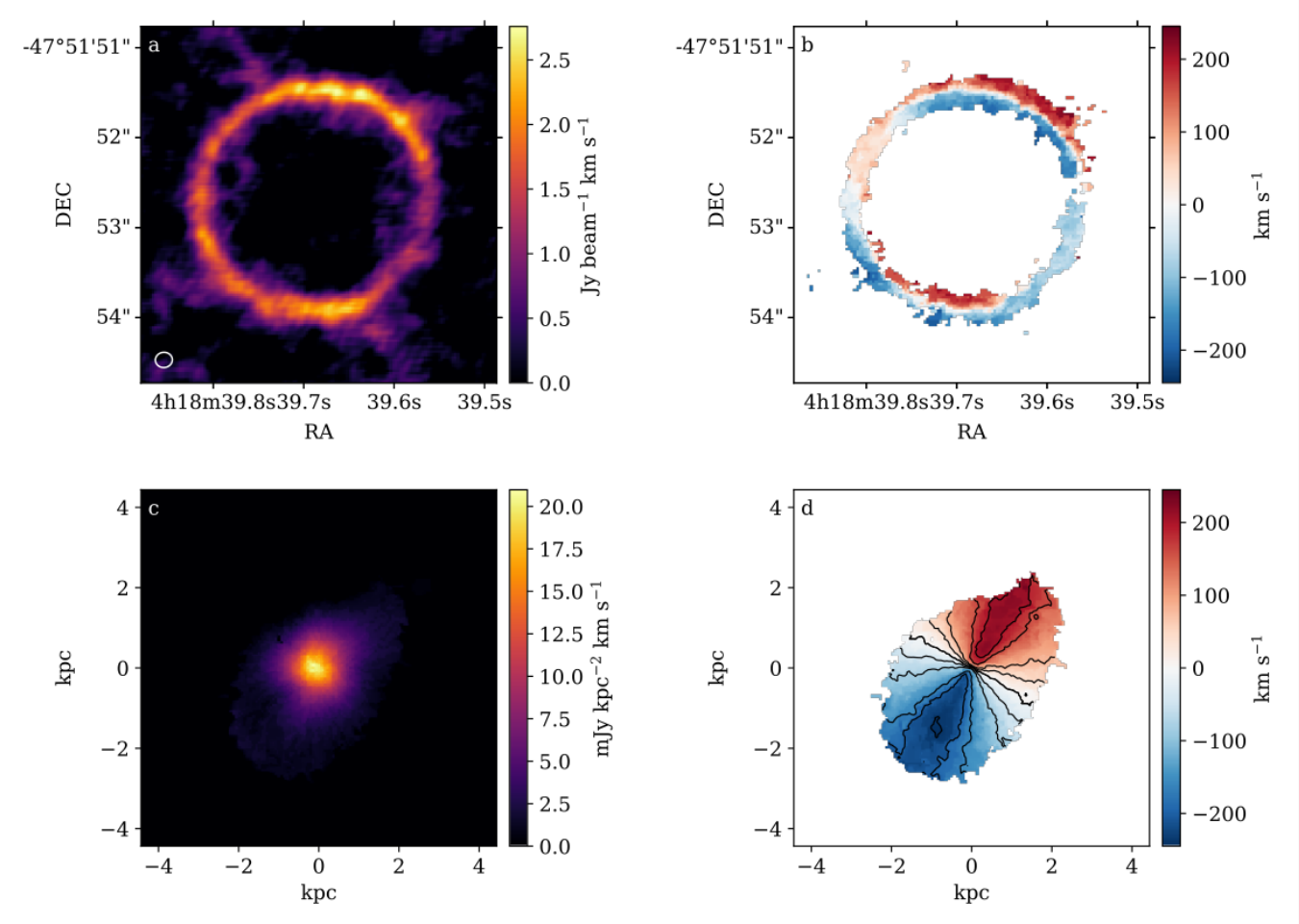
We find that the determinant of our sparse preconditioner also provides an adequate approximation for this normalization factor.

$$\mathrm{logdet}(A) \approx \mathrm{logdet} P = \mathrm{logdet}(L^T \tilde{C}_x^{-1} L + R)$$

The determinant is computed for free when performing the Cholesky decomposition.

Results

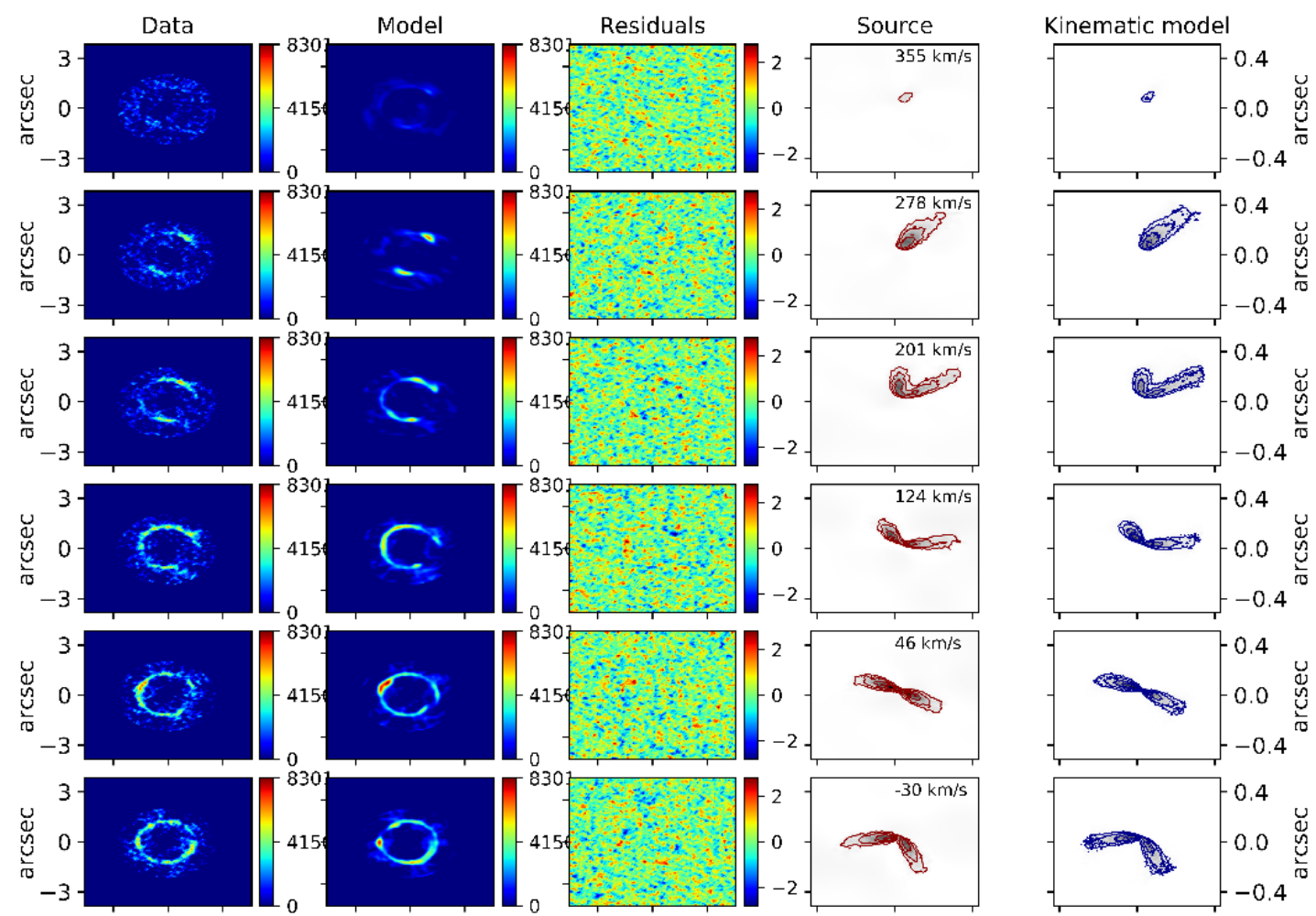
Kinematics of SPT0418-47 (Francesca Rizzo)



Dusty starburst galaxy at $z = 4.2$, effective ~ 60 pc resolution.

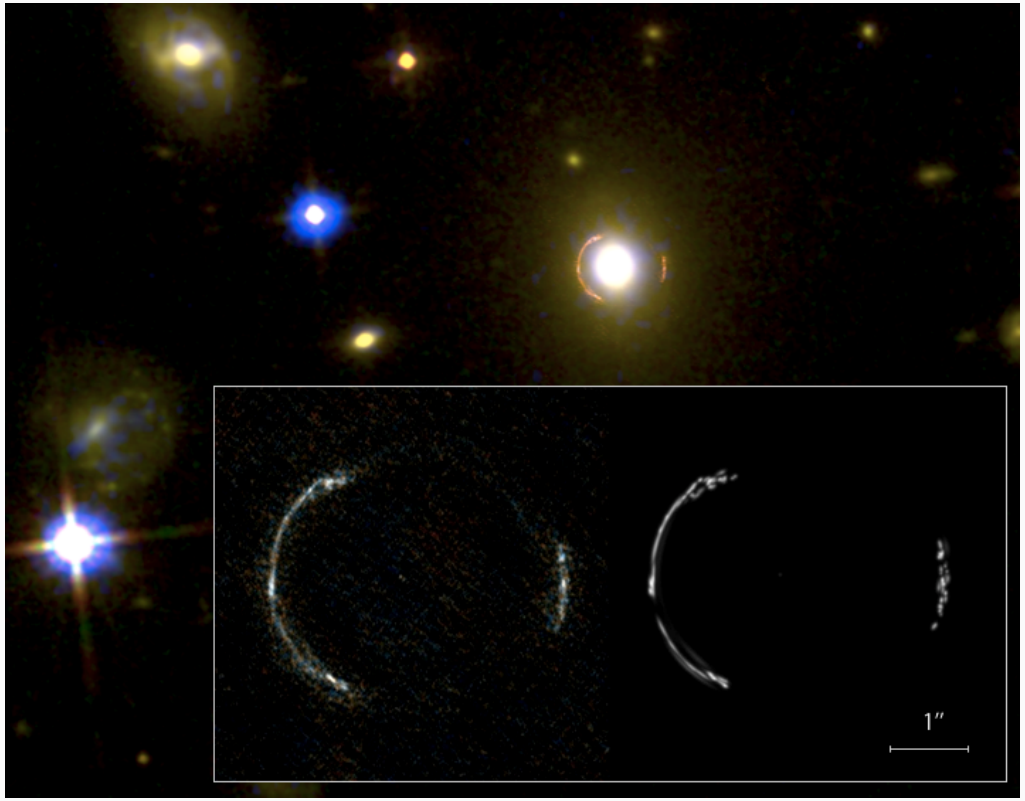
First study of $v_{\rm rot}/\sigma$ at this redshift.

Kinematics of SPT0418-47 (Francesca Rizzo)



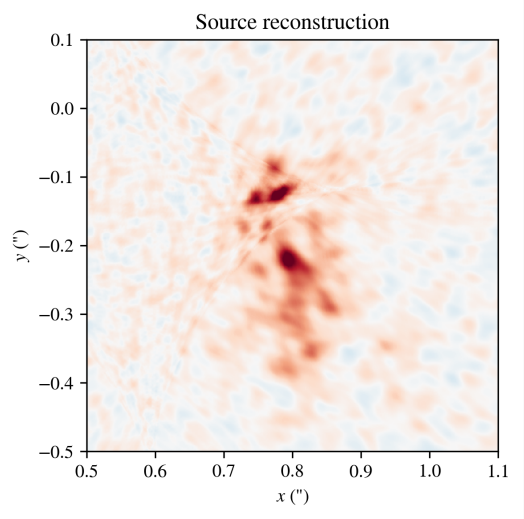
Rizzo et al., "A dynamically cold disk galaxy in the early Universe", *Nature* 2020.

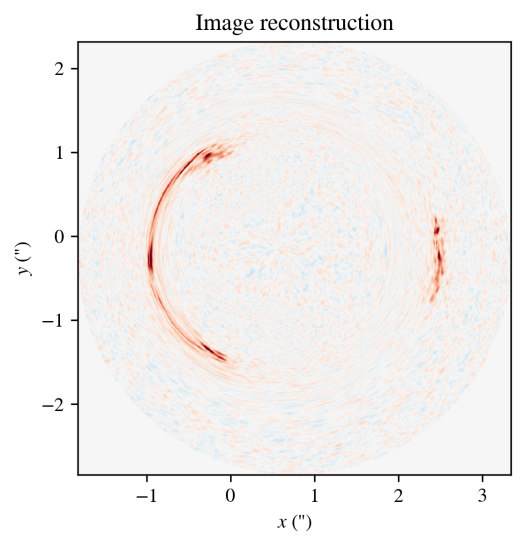
SDP.81

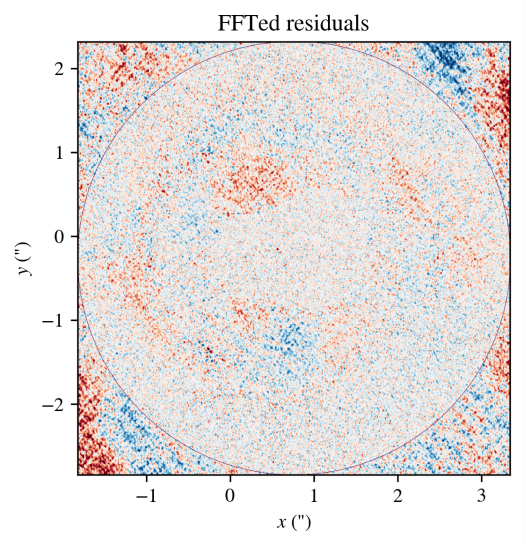


Y. Tamura (The University of Tokyo)/ALMA (ESO/NAOJ/NRAO)

SDP.81



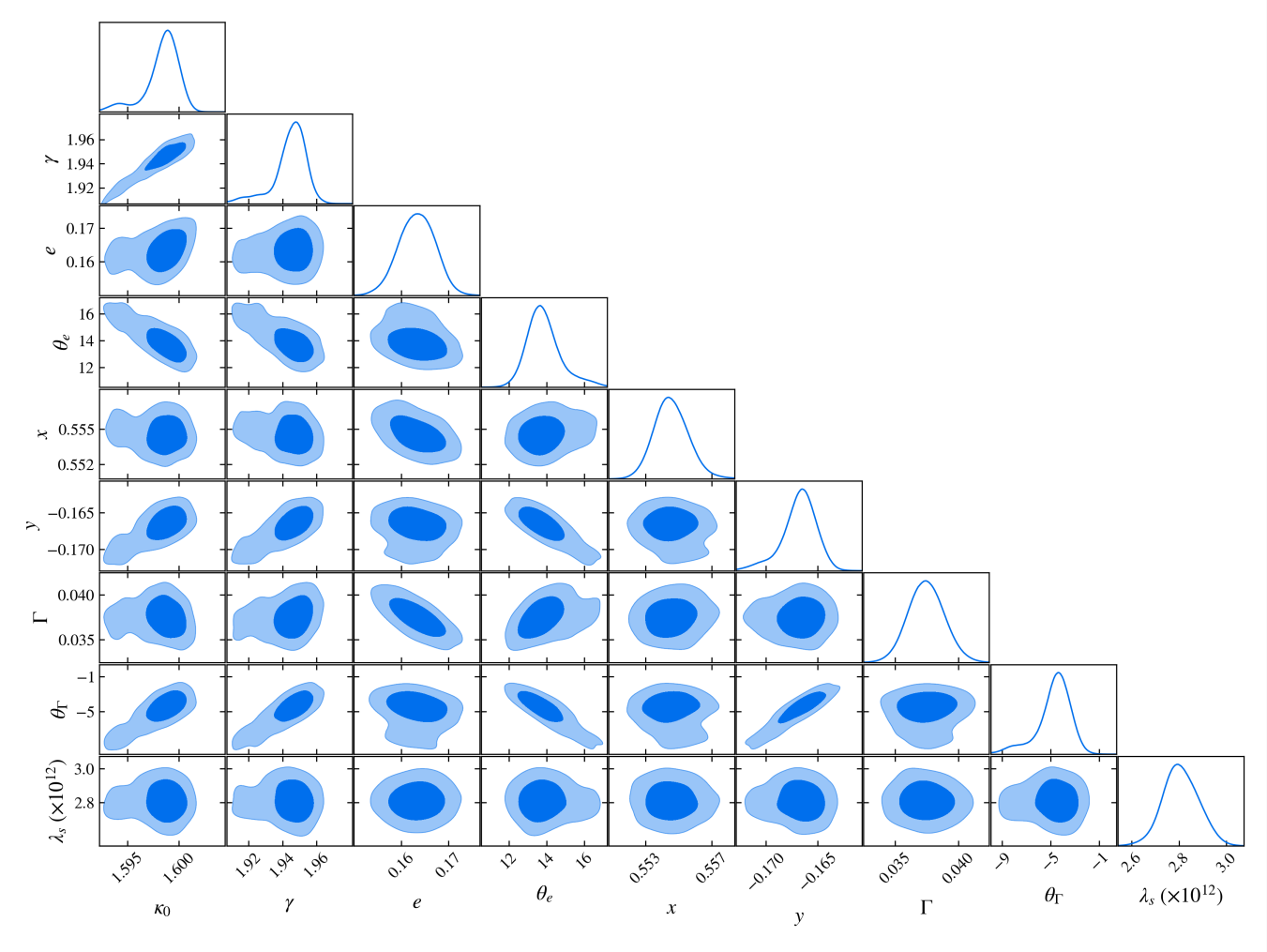




SDP.81 Band 6+7 observation. Attempting to reproduce substructure detection of Hezaveh et al. (2016).

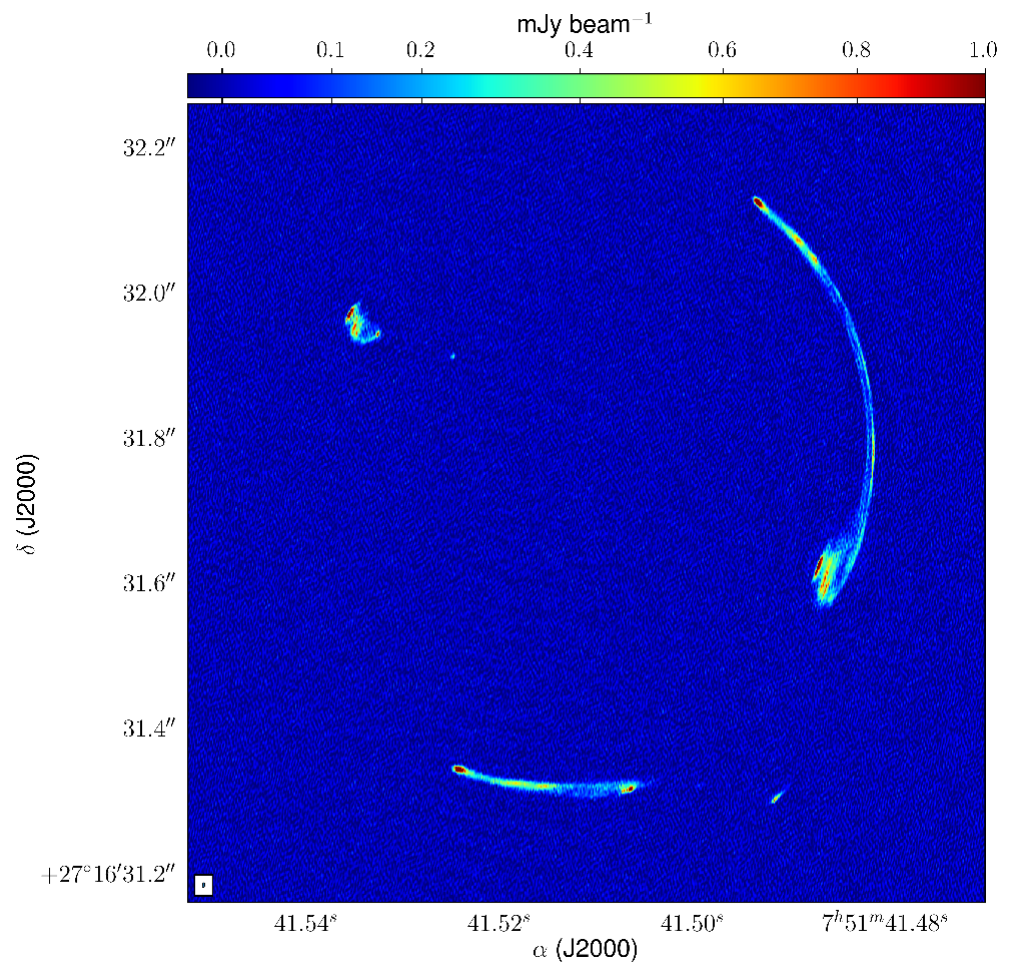
Computed from 8.8×10^8 visibilities. ~ 7 mas resolution.

SDP.81



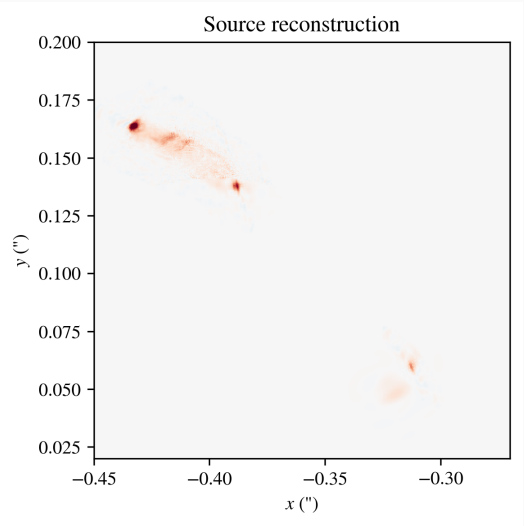
Posterior sampling in $\sim$ 12 hours on 100 cores + 16 GPUs.

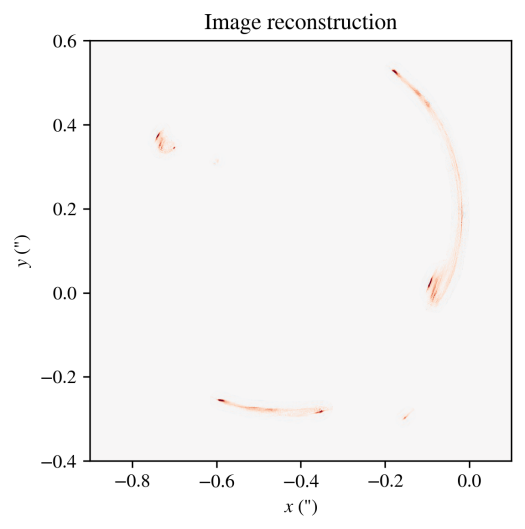
MG J0751+2716

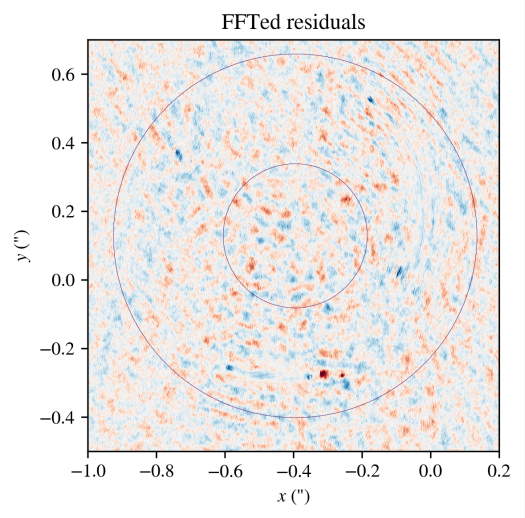


GVLBI data at $\sim$ 5 mas resolution, component-based modeling on CLEANed images (shown here) done by Spingola et al. (2018).

MG J0751+2716







4×10^8 visibilities, 5 mas resolution. Cored elliptical power law lens model does not quite explain the data. Modeling ongoing...

Outlook

Current status

- Working forward modeling code for radio interferometric data
- GPU acceleration for visibility gridding/de-gridding operation
- Data distributed over global MPI communicator for efficient memory use
- Currently applying code to several lens systems for DM searches and source science (Devon, Simona, Francesca)

Future work

- Self-calibration
- Polarimetry and rotation measure synthesis
- Plasma lens effects: Time delay, plasma lensing

References

Birrer, S., Amara, A., and Refregier, A. (2017). Lensing substructure quantification in RXJ1131-1231: a 2 keV lower bound on dark matter thermal relic mass. *Journal of Cosmology and Astro-Particle Physics*, 2017(5):037.

Chen, G. C. F., Suyu, S. H., Wong, K. C., Fassnacht, C. D., Chiueh, T., Halkola, A., Hu, I. S., Auger, M. W., Koopmans, L. V. E., Lagattuta, D. J., McKean, J. P., and Vegetti, S. (2016). SHARP - III. First use of adaptive-optics imaging to constrain cosmology with gravitational lens time delays. *MNRAS*, 462(4):3457–3475.

References ii

Courbin, F., Chantry, V., Revaz, Y., Sluse, D., Faure, C., Tewes, M., Eulaers, E., Koleva, M., Asfand iyarov, I., Dye, S., Magain, P., van Winckel, H., Coles, J., Saha, P., Ibrahimov, M., and Meylan, G. (2011). COSMOGRAIL: the COSmological MOnitoring of GRAvItational Lenses. IX. Time delays, lens dynamics and baryonic fraction in HE 0435-1223. *AAP*, 536:A53.

Despali, G. and Vegetti, S. (2017). The impact of baryonic physics on the subhalo mass function and implications for gravitational lensing. *MNRAS*, 469(2):1997–2010.

Despali, G., Vegetti, S., White, S. D. M., Giocoli, C., and van den Bosch, F. C. (2018). Modelling the line-of-sight contribution in substructure lensing. *MNRAS*, 475(4):5424–5442.

References iii

Fassnacht, C. D., Xanthopoulos, E., Koopmans, L. V. E., and Rusin, D. (2002). A Determination of H_0 with the CLASS Gravitational Lens B1608+656. III. A Significant Improvement in the Precision of the Time Delay Measurements. *ApJ*, 581(2):823–835.

Feroz, F., Hobson, M. P., and Bridges, M. (2009). MULTINEST: an efficient and robust Bayesian inference tool for cosmology and particle physics. *MNRAS*, 398(4):1601–1614.

References iv

Hezaveh, Y. D., Dalal, N., Marrone, D. P., Mao, Y.-Y., Morningstar, W., Wen, D., Bland ford, R. D., Carlstrom, J. E., Fassnacht, C. D., Holder, G. P., Kemball, A., Marshall, P. J., Murray, N., Perreault Levasseur, L., Vieira, J. D., and Wechsler, R. H. (2016). Detection of Lensing Substructure Using ALMA Observations of the Dusty Galaxy SDP.81. *ApJ*, 823:37.

Johnson, T. L., Sharon, K., Gladders, M. D., Rigby, J. R., Bayliss, M. B., Wuyts, E., Whitaker, K. E., Florian, M., and Murray, K. T. (2017). Star Formation at z = 2.481 in the Lensed Galaxy SDSS J1110 = 6459. I. Lens Modeling and Source Reconstruction. *ApJ*, 843(2):78.

References v

Koopmans, L. V. E., Barnabe, M., Bolton, A., Bradac, M., Ciotti, L., Congdon, A., Czoske, O., Dye, S., Dutton, A., Elliasdottir, A., Evans, E., Fassnacht, C. D., Jackson, N., Keeton, C., Lasio, J., Moustakas, L., Meneghetti, M., Myers, S., Nipoti, C., Suyu, S., van de Ven, G., Vegetti, S., Wucknitz, O., and Zhao, H. S. (2009). Strong Gravitational Lensing as a Probe of Gravity, Dark-Matter and Super-Massive Black Holes. In *astro2010: The Astronomy and Astrophysics Decadal Survey*, volume 2010, page 159.

References vi

Leethochawalit, N., Jones, T. A., Ellis, R. S., Stark, D. P., Richard, J., Zitrin, A., and Auger, M. (2016). A Keck Adaptive Optics Survey of a Representative Sample of Gravitationally Lensed Star-forming Galaxies: High Spatial Resolution Studies of Kinematics and Metallicity Gradients. *ApJ*, 820(2):84.

Lovell, M. R., Frenk, C. S., Eke, V. R., Jenkins, A., Gao, L., and Theuns, T. (2014). The properties of warm dark matter haloes. *MNRAS*, 439(1):300–317.

MacKay, D. J. (1991). Bayesian interpolation. *NEURAL COMPUTATION*, 4:415–447.

References vii

Mao, S. A., Carilli, C., Gaensler, B. M., Wucknitz, O., Keeton, C., Basu, A., Beck, R., Kronberg, P. P., and Zweibel, E. (2017). Detection of microgauss coherent magnetic fields in a galaxy five billion years ago. *Nature Astronomy*, 1:621–626.

Ritondale, E., Vegetti, S., Despali, G., Auger, M. W., Koopmans, L. V. E., and McKean, J. P. (2019). Low-mass halo perturbations in strong gravitational lenses at redshift z 0.5 are consistent with CDM. *MNRAS*, 485(2):2179–2193.

Rizzo, F., Vegetti, S., Fraternali, F., and Di Teodoro, E. (2018). A novel 3D technique to study the kinematics of lensed galaxies. *MNRAS*, 481(4):5606–5629.

References viii

Rusu, C. E., Wong, K. C., Bonvin, V., Sluse, D., Suyu, S. H., Fassnacht, C. D., Chan, J. H. H., Hilbert, S., Auger, M. W., Sonnenfeld, A., Birrer, S., Courbin, F., Treu, T., Chen, G. C. F., Halkola, A., Koopmans, L. V. E., Marshall, P. J., and Shajib, A. J. (2019). H0LiCOW XII. Lens mass model of WFI2033-4723 and blind measurement of its time-delay distance and H_0. *arXiv e-prints*, page arXiv:1905.09338.

Spingola, C., McKean, J. P., Auger, M. W., Fassnacht, C. D., Koopmans, L. V. E., Lagattuta, D. J., and Vegetti, S. (2018). SHARP - V. Modelling gravitationally lensed radio arcs imaged with global VLBI observations. *MNRAS*, 478(4):4816–4829.

References ix

Spingola, C., McKean, J. P., Vegetti, S., Auger, M. W., Koopmans, L. V. E., Fassnacht, C. D., Lagattuta, D. J., Powell, D., Rizzo, F., Stacey, H. R., and Sweijen, F. (2019). SHARP $-$ VI. Evidence for CO (1$-$0) molecular gas extended on kpc-scales in AGN star forming galaxies at high redshift. *arXiv e-prints*, page arXiv:1905.06363.

Suyu, S. H., Chang, T.-C., Courbin, F., and Okumura, T. (2018). Cosmological Distance Indicators. *SSR*, 214(5):91.

References x

Swinbank, A. M., Dye, S., Nightingale, J. W., Furlanetto, C., Smail, I., Cooray, A., Dannerbauer, H., Dunne, L., Eales, S., Gavazzi, R., Hunter, T., Ivison, R. J., Negrello, M., Oteo-Gomez, I., Smit, R., van der Werf, P., and Vlahakis, C. (2015). ALMA Resolves the Properties of Star-forming Regions in a Dense Gas Disk at z 3. *ApJ*, 806:L17.

Treu, T. and Marshall, P. J. (2016). Time delay cosmography. *AAPR*, 24(1):11.

Vegetti, S. and Koopmans, L. V. E. (2009). Bayesian strong gravitational-lens modelling on adaptive grids: objective detection of mass substructure in Galaxies. *MNRAS*, 392(3):945–963.

References xi

Vegetti, S., Koopmans, L. V. E., Auger, M. W., Treu, T., and Bolton, A. S. (2014). Inference of the cold dark matter substructure mass function at z = 0.2 using strong gravitational lenses. *MNRAS*, 442(3):2017–2035.

Vegetti, S., Koopmans, L. V. E., Bolton, A., Treu, T., and Gavazzi, R. (2010). Detection of a dark substructure through gravitational imaging. *MNRAS*, 408(4):1969–1981.

Vegetti, S., Lagattuta, D. J., McKean, J. P., Auger, M. W., Fassnacht, C. D., and Koopmans, L. V. E. (2012). Gravitational detection of a low-mass dark satellite galaxy at cosmological distance. *Nature*, 481(7381):341–343.

References xii

Wong, K. C., Suyu, S. H., Auger, M. W., Bonvin, V., Courbin, F., Fassnacht, C. D., Halkola, A., Rusu, C. E., Sluse, D., Sonnenfeld, A. r., Treu, T., Collett, T. E., Hilbert, S., Koopmans, L. V. E., Marshall, P. J., and Rumbaugh, N. (2017). H0LiCOW - IV. Lens mass model of HE 0435-1223 and blind measurement of its time-delay distance for cosmology. *MNRAS*, 465(4):4895–4913.