Imaging the Schwarzschild-radius-scale Structure of M87 with the Event Horizon Telescope Using Sparse Modeling

Kazunori Akiyama[1,2,3,4,15], Kazuki Kuramochi[2,3], Shiro Ikeda[5,6], Vincent L. Fish[1], Fumie Tazaki[2], Mareki Honma[2,7], Sheperd S. Doeleman[1,4,8], Avery E. Broderick[9,10], Jason Dexter[11], Monika Mościbrodzka[12], Katherine L. Bouman[13], Andrew A. Chael[4,7], and Masamichi Zaizen[14]

[1] Massachusetts Institute of Technology, Haystack Observatory, Route 40, Westford, MA 01886, USA; kazu@haystack.mit.edu
[2] Mizusawa VLBI Observatory, National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan
[3] Department of Astronomy, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
[4] Black Hole Initiative, Harvard University, 20 Garden Street,Cambridge, MA 02138,USA
[5] Department of Statistical Science, School of Multidisciplinary Sciences, Graduate University for Advanced Studies,
10-3 Midori-cho, Tachikawa, Tokyo 190-8562, Japan
[6] Graduate University for Advanced Studies, 10-3 Midori-cho, Tachikawa, Tokyo 190-8562, Japan
[7] Department of Astronomical Science, School of Physical Sciences, Graduate University for Advanced Studies, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan
[8] Harvard Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA
[9] Perimeter Institute for Theoretical Physics, 31 Caroline Street, North Waterloo, Ontario N2L 2Y5, Canada
[10] Department of Physics and Astronomy, University of Waterloo, 200 University Avenue West, Waterloo, Ontario N2l 3G1, Canada
[11] Max Planck Institute for Extraterrestrial Physics, Giessenbachstr. 1, D-85748 Garching, Germany
[12] Department of Astrophysics/IMAPP, Radboud University Nijmegen, P.O. Box 9010, 6500 GL Nijmegen, The Netherlands
[13] Massachusetts Institute of Technology, Computer Science and Artificial Intelligence Laboratory, 32 Vassar Street, Cambridge, MA 02139, USA
[14] Department of Astronomy, School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

Abstract

We propose a new imaging technique for radio and optical/infrared interferometry. The proposed technique reconstructs the image from the visibility amplitude and closure phase, which are standard data products of short-millimeter very long baseline interferometers such as the Event Horizon Telescope (EHT) and optical/infrared interferometers, by utilizing two regularization functions: the ℓ_1-norm and total variation (TV) of the brightness distribution. In the proposed method, optimal regularization parameters, which represent the sparseness and effective spatial resolution of the image, are derived from data themselves using cross-validation (CV). As an application of this technique, we present simulated observations of M87 with the EHT based on four physically motivated models. We confirm that ℓ_1 + TV regularization can achieve an optimal resolution of $\sim$20%–30% of the diffraction limit λ/D_{max}, which is the nominal spatial resolution of a radio interferometer. With the proposed technique, the EHT can robustly and reasonably achieve super-resolution sufficient to clearly resolve the black hole shadow. These results make it promising for the EHT to provide an unprecedented view of the event-horizon-scale structure in the vicinity of the supermassive black hole in M87 and also the Galactic center Sgr A*.

1. Introduction

Supermassive black holes (SMBHs) reside in the majority of the galactic nuclei in the universe. In a subset of such galaxies, accretion drives a highly energetic active galactic nucleus often associated with powerful jets. Understanding the nature of these systems has been a central goal in astronomy and astrophysics. The SMBHs at the centers of our galaxy (Sgr A*) and the giant elliptical galaxy M87 provide unprecedented opportunities to directly image the innermost regions close to the central black hole, since the angular size of the event horizon is the largest among known black holes. The angular size of the Schwarzschild radius (R_s) is $\sim$10 μas for Sgr A* for a distance of 8.3 kpc and a mass of $4.3 \times 10^6 M_\odot$ (e.g., Chatzopoulos et al. 2015), and $\sim$3–7 μas for M87 with a distance of 16.7 Mpc (Blakeslee et al. 2009) and a mass of 3–$6 \times 10^9 M_\odot$ (e.g., Gebhardt et al. 2011; Walsh et al. 2013). The apparent diameter of the dark silhouette of the black hole is $\sqrt{27} R_s$ for the non-rotating black hole. It corresponds with $\sim$52 μas for Sgr A* and $\sim$16–36 μas for M87, which changes

by only 4% with the black hole spin and viewing orientation (Bardeen 1973).

Very long baseline interferometric (VLBI) observations at short/submillimeter wavelengths ($\lambda \lesssim 1.3$ mm, $\nu \gtrsim 230$ GHz) can achieve a spatial resolution of a few tens of microarcseconds and therefore are expected to resolve event-horizon-scale structures, including the shadow of SMBHs. Indeed, recent significant progress on 1.3 mm VLBI observations with the Event Horizon Telescope (EHT; Doeleman et al. 2009) has succeeded in resolving compact structures of a few R_s in the vicinity of the SMBHs in both Sgr A* and M87 (e.g., Doeleman et al. 2008, 2012; Fish et al. 2011, 2016; Akiyama et al. 2015; Johnson et al. 2015). Direct imaging of these scales will be accessible in the next few years with technical developments and the addition of new (sub)millimeter telescopes such as the Atacama Large Submillimeter/millimeter Array (ALMA) to the EHT (e.g., Fish et al. 2013).

Regardless of the observing wavelength, angular resolution (often referred to as "beam size" in radio astronomy and "diffraction limit" in optical astronomy) is simply given by $\theta \approx \lambda/D_{\mathrm{max}}$, where λ and D_{max} are the observed wavelength and the diameter of the telescope (or the longest baseline length

---

for the radio interferometer), respectively. A practical limit for a ground-based, 1.3 mm VLBI array like the EHT is $\sim 25\,\mu$as ($= 1.3$ mm/10,000 km), which is comparable to the radius of the black hole shadow in M87 and Sgr A*. Hence, imaging techniques with good imaging fidelity at a spatial resolution smaller than λ/D_{max} would be desirable, particularly for future EHT observations of M87 and Sgr A*.

The imaging problem of interferometry is formulated as an under-determined linear problem when reconstructing the image from full-complex visibilities that are Fourier components of the source image. In the context of compressed sensing (also known as "compressive sensing"), it has been revealed that an ill-posed linear problem may be solved accurately if the underling solution vector is sparse (Candes & Tao 2006; Donoho 2006). Since then, many imaging methods[16] have been applied to radio interferometry after pioneering work by Wiaux et al. (2009a) and Wiaux et al. (2009b); (see Garsden et al. 2015, and references therein). We call these approaches "sparse modeling" since they utilize the sparsity of the ground truth.

In Honma et al. (2014), we applied LASSO (Least Absolute Shrinkage and Selection Operator; Tibshirani 1996), a technique of sparse modeling, to interferometric imaging. LASSO solves under-determined ill-posed problems by utilizing the ℓ_1-norm (see Section 2.2 for details). Minimizing the ℓ_1-norm of the solution reduces the number of non-zero parameters in the solution, equivalent to choosing a sparse solution. The philosophy of LASSO is therefore similar to that of the traditional CLEAN technique (Högbom 1974), which favors sparsity in the reconstructed image and has been independently developed as Matching Pursuit (Mallat & Zhang 1993) in statistical mathematics for sparse reconstruction. In Honma et al. (2014), we found that LASSO can potentially reconstruct structure ~ 4 times finer than λ/D_{max}. Indeed, it works well for imaging the black hole shadow for M87 with the EHT in simulations.

Our previous work (Honma et al. 2014) has three relevant issues. The first issue is reconstructing the image only from the visibility amplitudes and closure phases (see Section 2.1), which have been the standard data products of EHT observations (Lu et al. 2012, 2013; Akiyama et al. 2015; Wagner et al. 2015; Fish et al. 2016) and optical/infrared interferometry. The algorithm of Honma et al. (2014) is applicable only for full-complex visibilities, which are the usual data products from longer-wavelength radio interferometers. We have recently developed a fast and computationally cheap method to retrieve the visibility phases from closure phases (PRECL; Ikeda et al. 2016), which can reconstruct the black hole shadow of M87 combined with LASSO in simulated EHT observations. However, since the phase reconstruction in PRECL adopts a different prior on visibilities than LASSO, the resultant image may not be optimized well in terms of ℓ_1-norm minimization and sparse image reconstruction. Another potential approach is to solve for the image and visibility phases with ℓ_1-norm regularization simultaneously, enabling us to reconstruct the image with the full advantage of the regularization function.

The second issue is that the ℓ_1-norm regularization might not provide a unique solution and/or could reconstruct an image that is too sparse of an image if the number of pixels

with non-zero brightness is not small enough compared to the number of pixels. This violates a critical assumption in techniques with ℓ_1-norm regularization that the solution (i.e., the true image) should be sparse. Such a situation may occur for an extended source or also even for a compact source if the imaging pixel size is set to be much smaller than the size of the emission structure. Pioneering work has made use of transforms to wavelet or curvelet bases in which the image can be represented sparsely (e.g., Li et al. 2011; Carrillo et al. 2012, 2014; Dabbech et al. 2015; Garsden et al. 2015). As another strategy to resolve this potential issue, in Honma et al. (2014), we proposed the addition of another regularization, Total Variation (TV; e.g., Rudin et al. 1992), which is another popular regularization in sparse modeling. TV is a good indicator for sparsity of the image in its gradient domain instead of the image domain (see Section 2.2 for details) and it has been applied to astronomical imaging (e.g., Wiaux et al. 2010; McEwen & Wiaux 2011; Carrillo et al. 2012, 2014; Uemura et al. 2015; Chael et al. 2016) that includes optical interferometric imaging without the visibility phases (e.g., MiRA; Thiébaut 2008; and also see Thiébaut 2013 for a review). TV regularization generally favors a smooth image (i.e., with larger effective resolution) but with a sharp edge, in contrast with maximum entropy methods (MEM; e.g., Narayan & Nityananda 1986), which favor a smooth edge (see, e.g., Uemura et al. 2015, for comparison between TV and MEM). Inclusion of TV regularization enables reconstruction of an extended image while preserving sharp emission features preferred by ℓ_1-norm regularization, thereby extending the class of objects where sparse modeling is applicable. Indeed, regularization with both the ℓ_1-norm and TV has been shown to be effective for imaging polarization with full-complex visibilities in our recent work (Akiyama et al. 2017).

An important detail is the determination of regularization parameters (e.g., weights on regularization functions), which is common in the vast majority of existing techniques. Since one cannot know the true image of the source a priori, one should evaluate goodness-of-fitting and select appropriate regularization parameters from the data themselves. In well-posed problems, one can use statistical quantities considering residuals between data and models as well as model complexity to avoid over-fitting, such as reduced χ^2, the Akaike information criterion and the Bayesian information criterion using the degrees of freedom to constrain model complexity. However, for ill-posed problems like interferometric imaging, degrees of freedom cannot be rigorously defined, preventing the use of such statistical quantities.

In this paper, we propose a new technique to reconstruct images from interferometric data using sparse modeling. The proposed technique directly solves the image from visibility amplitudes and closure phases. In addition to the ℓ_1-norm (LASSO), we also utilize another new regularization term, TV, so that a high-fidelity image will be obtained even with a small pixel size and/or for extended sources. Furthermore, we propose a method to determine optimal regularization parameters with cross-validation (CV; see Section 2.3), which can be applied to many existing imaging techniques. As an example, in this paper, we applied our new technique to data obtained from simulated observations of M87 with the array of the EHT expected in Spring 2017.

---

[16] A list of papers are available in https://ui.adsabs.harvard.edu/#/public-libraries/wmxthNHHQrGDS2aKt3gXow.

2. The Proposed Method

2.1. A Brief Introduction of the Closure Phase

A goal of radio and optical/infrared interferometry is to obtain the brightness distribution $I(x, y)$ of a target source at a wavelength λ or a frequency ν, where (x, y) is a sky coordinate relative to a reference position, so-called the phase-tracking center. The observed quantity is a complex function called visibility $V(u, v)$, which is related to $I(x, y)$ by two-dimensional Fourier transform given by

$$V(u, v) = \int dx dy\, I(x, y) e^{-i2\pi(ux+vy)}. \qquad (1)$$

Here, the spatial frequency (u, v) corresponds to the baseline vector (in units of the observing wavelength λ) between two antennas projected to the tangent plane of the celestial sphere at the phase-tracking center.

Observed visibilities are discrete quantities, and the sky image can be approximated by a pixellated version where the pixel size is much smaller than the nominal resolution of the interferometer. The image can therefore be represented as a discrete vector I, related to the Stokes visibilities V by a discrete Fourier transform F:

$$V = FI. \qquad (2)$$

The sampling of visibilities is almost always incomplete. Since the number of visibility samples V is smaller than the number of pixels in the image, solving the above equation for the image I is an ill-posed problem.

Here, we consider that the complex visibility V_j is obtained from observation(s) with multiple antennas. Let us define its phase and amplitude as ϕ_j and $\bar{V}_j$, respectively, denoted as follows

$$V_j = \bar{V}_j e^{i\phi_j}, \qquad (3)$$

where j is the index of the measurement. Each measurement corresponds to a point (u_j, v_j) in the (u, v)-plane and recorded at time t_j. In actual observations, some instrumental effects and the atmospheric turbulence primary from the troposphere induce the antenna-based errors in the visibility phase, leading to the observed phase $\tilde{\phi}_j$ being offset from the true phase ϕ_j of the true image. In particular, this is a serious problem in VLBI observations performed at different sites (see Thompson et al. 2001).

However, the robust interferometric phase information can be obtained through the measurements of the closure phase, defined as a combination of triple phases on a closed triangle of baselines recorded at the same time. It is known that the closure phase is free from antenna-based phase errors (Jennison 1958), which can be seen from the following definition of the closure phase,

$$\psi_m^{123} = \tilde{\phi}_j^{12} + \tilde{\phi}_k^{23} + \tilde{\phi}_l^{31} = \phi_j^{12} + \phi_k^{23} + \phi_l^{31}, \qquad (4)$$

where m is the index of the closure phase, and upper numbers (1, 2, 3) mean the index of stations involved in the closure phase or the visibility phase. The closure phase is also known as a phase term of the triple product of visibilities on closed baselines recorded at the same time, $V_j^{12} V_k^{23} V_l^{31}$, known as the bi-spectrum.[17] Closure phases have been used to calibrate visibility phases in VLBI observations (e.g., Rogers et al. 1974).

In short/submillimeter VLBI or optical/infrared interferometry, the stochastic atmospheric turbulence in the troposphere over each station induces a rapid phase rotation in the visibility, making it difficult to calibrate or even measure the visibility phase reliably (e.g., see Rogers et al. 1995; Thiébaut 2013). Thus, image reconstruction using more robust closure phases, free from station-based phase errors, is useful for interferometric imaging with such interferometers.

2.2. Image Reconstruction from Visibility Amplitudes and Closure Phases

In this paper, we propose a method to solve the two-dimensional image $I = \{I_{i,j}\}$ by the following equations:

$$\min_I C(I) \text{ subject to } I_{i,j} \geqslant 0. \qquad (5)$$

The cost function $C(I)$ is defined as

$$C(I) = \chi^2(I) + \eta_l \|I\|_1 + \eta_t \|I\|_{\mathrm{tv}}, \qquad (6)$$

where $\|I\|_p$ is l_p-norm of the vector I given by

$$\|I\|_p = \left(\sum_i \sum_j |I_{i,j}|^p \right)^{\frac{1}{p}} \text{ (for } p > 0), \qquad (7)$$

and $\|I\|_{\mathrm{tv}}$ indicates an operator of TV.

The first term of Equation (6) is the traditional χ^2 term representing the deviation between the reconstructed image and observational data (i.e., the visibility amplitude $\bar{V} = \{|V_j|\}$ and closure phase $\Psi = \{\psi_m\}$), defined by

$$\chi^2(I) = \|\bar{V} - A(FI)\|_2^2 + \|\Psi - B(FI)\|_2^2, \qquad (8)$$

where A and B indicate operators to calculate the visibility amplitude and closure phase, respectively. Deviations between the model and observational data are normalized with the errors. This form of the residual sum of squares (RSS) is originally proposed in the Bi-spectrum Maximum Entropy Method (Buscher 1994) and also for modeling EHT data (e.g., Lu et al. 2012, 2013). Note that it could be replaced with an RSS term for bi-spectra (e.g., Bouman et al. 2015; Chael et al. 2016).

The second term represents LASSO-like regularization using the ℓ_1-norm. Under the non-negative condition, ℓ_1-norm is equivalent to the total flux. η_l is the regularization parameter for LASSO, adjusting the degree of sparsity by changing the weight of the ℓ_1-norm penalty. In general, a large η_l prefers a solution with very few non-zero components, while $\eta_l = 0$ introduces no sparsity. In this paper, we use the normalized regularization parameter $\tilde{\eta}_l$ defined by

$$\eta_l \equiv \tilde{\eta}_l (N_{\mathrm{amp}} + N_{\mathrm{cphase}}) / \max(\bar{V}), \qquad (9)$$

which is less affected by the number of visibility amplitude and closure phase data points, N_{amp} and N_{cphase}, respectively, and also by the total flux density of the target source.

---

[17] Data products of visibility amplitudes and closure phases are also sometimes named as "bi-spectrum" in the literature (e.g., Buscher 1994). In this paper, we strictly distinguish them.

The third term is the TV regularization, defined by the sum of all differences of the brightness between adjacent image pixels. In this paper, we adopt a typical form for two-dimensional images (Rudin et al. 1992) that has been used in astronomical imaging, (e.g., Wiaux et al. 2010; McEwen & Wiaux 2011; Uemura et al. 2015; Chael et al. 2016), defined as

$$||\boldsymbol{I}||_{\rm tv} = \sum_i \sum_j \sqrt{|I_{i+1,j} - I_{i,j}|^2 + |I_{i,j+1} - I_{i,j}|^2}. \qquad (10)$$

TV is a good indicator of image sparsity in its gradient domain instead of the image domain. TV is highly affected by the effective spatial resolution of the image. For instance, an image with a small TV value has blocks of image pixels whose brightness are similar, since the image is sparse in the gradient domain. The size of such blocks is equivalent to the effective spatial resolution, getting smaller for images with higher TV values. Thus, the regularization parameter $\eta_{\rm t}$ adjusts the effective spatial resolution of the reconstructed image. In general, a larger (smaller) $\eta_{\rm t}$ prefers a smoother (finer) distribution of power with less (higher) discreteness, leading to larger (smaller) angular resolution. In the present work, we use the normalized regularization parameter $\tilde{\eta}_{\rm t}$ defined by

$$\eta_{\rm t} \equiv \tilde{\eta}_{\rm t}(N_{\rm amp} + N_{\rm cphase})/4\max(\bar{\boldsymbol{V}}), \qquad (11)$$

similar to the LASSO term. A factor of four is based on a property of TV that displays a difference in the brightness in all four directions at each pixel. Note that a major difference to maximum entropy methods, which also favor smooth images, is that TV regularization has a strong advantage in edge-preserving; strong TV regularization favors a piecewise smooth structure, but with clear and often sharp boundaries between non-emitting and emitting regions.

The problem described in Equations (5) and (6) is nonlinear minimum optimization. In this work, we adopt a nonlinear programming algorithm L-BFGS-B (Byrd et al. 1995; Zhu et al. 1997) that is an iterative method for solving bound-constrained nonlinear optimization problems. L-BFGS-B is one of the quasi-Newton methods that approximates the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm using a limited amount of computer memory. In L-BFGS-B, the cost function and its gradient are used to determine the next model parameters at each iterative process. We approximately set partial derivatives to 0 at non-differentiable points for both the ℓ_1-norm and TV. The partial derivatives of χ^2 are calculated numerically with central differences. We use the latest Fortran implementation of L-BFGS-B (L-BFGS-B v3.0; Morales & Nocedal 2011). We note that the problem is non-convex like other imaging techniques using closure quantities (e.g., Buscher 1994; Thiébaut 2008; Bouman et al. 2015; Chael et al. 2016), and a global solution is generally not guaranteed.

2.3. Determination of Imaging Parameters

In the proposed method, the most important tuning parameters are the regularization parameters for the ℓ_1-norm ($\tilde{\eta}_{\rm l}$) and TV ($\tilde{\eta}_{\rm t}$), which determine the sparseness and effective spatial resolution of the image, respectively. Smaller regularization parameters generally favor images with larger numbers of non-zero image pixels and more complex image structure, which could give better χ^2 values by over-fitting. On the other hand, large regularization parameters provide images that are

too simple and that do not fit the data well. To determine optimal parameters, we need to evaluate the goodness-of-fit using Occam's razor to prevent over-fitting.

In this work, we adopt CV to evaluate goodness-of-fit. CV is a measure of the relative quality of the models for a given set of data. CV checks how the model will generalize to an independent data set by using separate data sets for fitting the model and for testing the fitted model. CV consists of three steps: (1) randomly partitioning a sample of data into complementary subsets, (2) performing the model fitting on one subset (called the training set), and (3) validating the analysis on the other subset (called the validation set). To reduce variability, multiple rounds of CV are performed using different partitions, and the validation results are averaged over the rounds. If the regularization parameters are too small, the established model from the training set would be over-fitted and too complicated, resulting in a large deviation in the validation set. On the other hand, if the regularization parameters are too large, the established model would be too simple and not well-fitted to the training set, also resulting in a large deviation in the validation set. Thus, reasonable parameters can be estimated by finding a parameter set that minimizes deviations (e.g., χ^2) of the validation set.

In this work, we adopted 10-fold CV for evaluating the goodness-of-fit. The original data were randomly[18] partitioned into 10 equal-sized subsamples. Nine subsamples were used in the image reconstruction as the training set, and the remaining single subsample was used as the validation set for testing the model using χ^2. We repeated the procedure by changing the subsample for validation data 10 times, until all subsamples were used for both training and validation. The χ^2 values of the validation data were averaged and then used to determine optimal tuning parameters.

An important advantage of this method compared with previously proposed methods is that it is applicable to any type of regularization functions and also imaging with multiple regularization functions. For instance, Carrillo et al. (2012, 2014) and subsequent work solve images by utilizing the ℓ_1-norm on wavelet-transformed image or TV regularization alone. In this case, the parameter can be uniquely determined from the ℓ_2-norm of the estimated uncertainties on observational data (see Carrillo et al. 2012, for details). However, it is not straightforward to extend the idea for the problems with multiple regularization functions. For another example, Garsden et al. (2015) proposes another heuristic method to determine the regularization parameters on ℓ_1-regularization on the wavelet/curvelet-transformed image by estimating its noise level on each scale, which is successful. However, the method would not work for all types of regularization functions. On the other hand, CV is a general technique that can be applied to imaging with any other regularization functions or any combination of them, in principle, which include MEM (e.g., Buscher 1994; Chael et al. 2016) and patch priors (e.g., Bouman et al. 2015). This advantage is particularly important for Sgr A*, which needs to involve a regularization function to mitigate the interstellar

---

[18] In this work, subsamples are obtained with a uniform probability regardless of baselines following the most basic style of the CV. However, there could be a more effective way of choosing subsamples for interferometric imaging. The optimum partitioning for CV is in the scope of our next studies in the near future.

Table 1
Stations Used in Simulated Observations

Telescope	SEFD (Jy)
Phased ALMA	100
Phased SMA and JCMT	4000
LMT	1400
IRAM 30 m	1400
NOEMA single dish	5200
ARO/SMT	11000

scattering effects (scattering optics; Johnson 2016) in addition to the general regularization function(s) for imaging.

A relevant disadvantage of this method is its computational cost, since n-fold CV requires the reconstruction $(n+1)$ of images for each set of regularization parameters. Recently, an accurate approximation of CV, which can be derived from a single imaging on full data set for each parameter set, has been proposed for imaging from full-complex visibilities with $\ell_1 + $ TV regularizations (Obuchi & Kabashima 2016; Obuchi et al. 2016). Future development of such heuristic approximations for other types of data and regularization functions could overcome this issue.

3. Imaging Simulations

3.1. Physically Motivated Models

In this paper, we adopt four physical models previously proposed for 1.3 mm emission on event-horizon scales.

The first model is a simple, but qualitatively correct, force-free jet model (hereafter BL09) in the magnetically dominated regime presented in Broderick & Loeb (2009) and Lu et al. (2014). We adopted a model image presented in Akiyama et al. (2015), which is based on the model parameters fitted to the results of 1.3 mm observations with the EHT in Doeleman et al. (2012) and the SED of M87 (A. E. Broderick et al. 2017, in preparation). The approaching jet is predominant for this model (see Akiyama et al. 2015, for more details).

The second and third models are based on results of GRMHD simulations presented in Dexter et al. (2012). We used the representative models DJ1 and J2, which are based on the same GRMHD simulation but with different energy and spatial distributions for radio-emitting leptons. The dominant emission region is the accretion flow in DJ1 and the counter jet in J2 illuminating the last photon orbit in J2. We adopt model images in Akiyama et al. (2015), where the position angle of the large-scale jet for models is adjusted to $-70°$ inferred for M87 (e.g., Hada et al. 2011).

The last model is based on results of GRMHD simulations presented in Mościbrodzka et al. (2016), which models M87 core emission as radiation produced by the jet sheath. We use the image averaged for $\sim$3 months for our simulation (hereafter M16). The image has its dominant contribution from the counter jet illuminating the last photon orbit similar to J2 of Dexter et al. (2012), but the M16 model assumes energy distributions of leptons quite different from J2. We rotate the original model image of Mościbrodzka et al. (2016) to adjust the position angle of the large-scale jet to $-70°$.

3.2. Simulated Observations

We simulate observations with the EHT at 1.3 mm (230 GHz) using the MAPS (MIT Array Performance
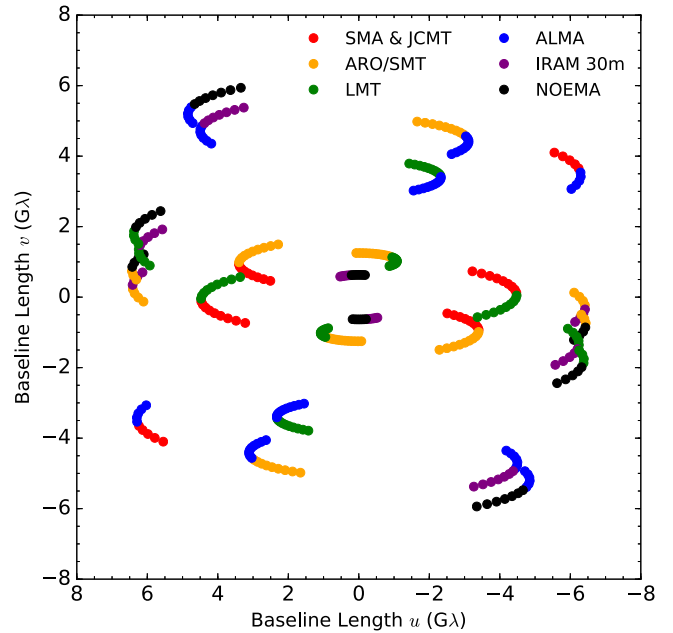

Figure 1. uv-coverage of the simulated observations with the EHT array expected in spring 2017. Each baseline is split into two colors to show involving stations.

Simulator) package[19] based on the above models. The simulated observations are performed for the array expected to comprise in Spring 2017.

We assume an array consisting of stations at six different sites: a phased array of the Submillimeter Array (SMA) antennas and the James Clerk Maxwell Telescope (JCMT) on Mauna Kea in Hawaii; the Arizona Radio Observatory's Submillimeter Telescope (ARO/SMT) on Mt. Graham in Arizona; the Large Millimeter Telescope (LMT) on Sierra Negra, Mexico; a phased array of the Atacama Large Millimeter/submillimeter Array (ALMA) in the Atacama desert, Chile; the Institut de Radioastronomie Millimétrique (IRAM) 30 m telescope on Pico Veleta, Spain; and a single dish telescope of the Northern Extended Millimeter Array (NOEMA) in France. We adopt the system equivalent flux density (SEFD) of each station shown in Table 1 based on the proposer's guide of 1 mm VLBI observations in ALMA Cycle 4.

The simulations assume a bandwidth of 3.5 GHz for Stokes I, which is half of the standard setting in ALMA Cycle 4.[20] We assume a correlation efficiency of 0.7, including a quantization efficiency of 0.88 for 2 bit sampling and other potential losses such as bandpass effects and pointing errors.

We simulate observations as a series of 5 minute scans with a cadence of 20 minutes over a GST range of 13–0 hr, corresponding to the time range when M87 can be observed by ALMA or LMT at an elevation greater than 20°. ALMA and LMT are sensitive stations near the middle of the east–west extent of the array, and they may be important anchor stations for fringe detection. This provides an observational efficiency of 25% in time, expected for VLBI observations with ALMA in 2017. Data are integrated for the duration of each scan (i.e., 5 minutes) following previous EHT observations (e.g., Doeleman et al. 2012;

Akiyama et al. 2015). Figure 1 shows the resultant uv-coverage of simulated observations. The maximum baseline length of observations is 7.2 Gλ, corresponding to $\lambda/D = 28.5\ \mu$as.

We note that the conditions of our simulation are much worse than previous simulations in Lu et al. (2014) in terms of the baseline sensitivity, uv-coverage, angular resolution (i.e., the maximum baseline length), and the exposure time of observations. Nevertheless, our simulated conditions are much closer to the observational conditions in Spring 2017.

3.3. Imaging

We reconstruct images from simulated data sets based on the method described in Section 2.1. We adopt a field of view (FOV) of 200 μas gridded by 100 pixels in both the R.A. and decl. directions for all models, giving a pixel size of ~1.6 μas corresponding to a physical scale of ~0.21 R_s. Images are reconstructed at four regularization parameters for both $\tilde{\eta}_l$ and $\tilde{\eta}_t$, ranging as 10^{-1}, 10^0, …, 10^{+2}. As a result, we obtain $4 \times 4 = 16$ images for each model.

Since the problem described in Section 2.1 is non-convex, our algorithm may be trapped in a local minimum and therefore may end up at an initial-condition dependent solution, similar to other algorithms using techniques minimizing non-convex functions with gradient descent methods. To avoid this, we start reconstructing images at $\tilde{\eta}_l = \tilde{\eta}_t = 10^2$, which is expected to derive the simplest image among parameters we adopt, assuming a point source as initial images. This is then used as the initial image at other values of the regularization parameters.

At each parameter, we do iterations until achieving convergence or 1000 iterations, and then filter the output image with a hard thresholding defined by

$$I_i^{\text{filtered}} = \begin{cases} 0 & (|I_i| < t) \\ I_i & (\text{otherwise}) \end{cases}, \tag{12}$$

where t is a threshold. We repeat this process until the normalized root mean square error (NRMSE) between the previous and latest filtered images becomes smaller than 1%. NRMSE is defined by (e.g., Chael et al. 2016)

$$\text{NRMSE}(\boldsymbol{I}, \boldsymbol{K}) = \sqrt{\frac{\sum_i |I_i - K_i|^2}{\sum_i |K_i|^2}}, \tag{13}$$

where $\boldsymbol{I}$ and $\boldsymbol{K}$ are the image to be evaluated and the reference image, respectively. We adopt the latest non-filtered image as the final product. Although this procedure makes the computational time longer, we found that this works very well for avoiding convergence at some local minima. In this paper, we set 10% of the peak flux as a threshold t. For the simulated observational data in this paper, it takes typically about several to 10 minutes on a standard desktop computer with six intel core-i7 CPU cores to reconstruct an image at each set of two regularization parameters.

We evaluate the goodness-of-fit for each image and then selected the best-fit images with 10-fold CV as described in Section 2.3. The quality of the reconstructed images is evaluated with the NRMSE. Since all model images have finer resolutions with narrower FOVs than the reconstructed images, we calculated these metrics as follows. First, we adjusted the pixel size of the reconstructed image to that of the model image with bi-cubic spline interpolation. Second, we

adjusted the position offsets between these two images so that the positions of their centers of mass coincide with each other,[21] because absolute positions cannot be defined from visibility amplitudes and closure phases alone. Finally, the metrics were evaluated. In addition to NRMSE, we also measure structural dissimilarity (Wang et al. 2004) between the model and reconstructed images using the DSSIM metric adopted work by Lu et al. (2014) and Fish et al. (2014). Since both metrics show similar trends, we show only the behavior of the NRMSE in the figures that follow.

3.3.1. Imaging with the Cotton–Schwab CLEAN

For evaluating the performance of our techniques, we also reconstructed images with the most widely used Cotton–Schwab CLEAN (henceforth CS-CLEAN; Schwab 1984) implemented in the Common Astronomy Software Applications (CASA) package[22] with uniform weighting. Since CLEAN requires complex visibilities, we adopted the simulated complex visibilities with thermal noises. We set a gain of 0.1 and a threshold of 0.08 mJy beam^{-1}, comparable to the image sensitivity of simulated observations. Since the fast Fourier transform is often used in CLEAN, a very small FOV can require a grid size in uv-plane that is too large, which could cause additional deconvolution errors. Hence, we set 1024 pixels with the same pixel size in each axis for the entire map, and put a CLEAN box in the central 100×100 pixels to put CLEAN components in the same region as other techniques. We use the model image instead of the CLEAN map for calculating metrics, since the residual map, which is generally added to the CLEAN map, cannot be calculated for the proposed method.

4. Results

4.1. The Best-fit Images

We show the best-fit images selected with CV in Figure 2. A clear shadow feature is well reproduced for the counter-jet- and accretion-flow-dominated models (J2, M16 and DJ1). This demonstrates that the EHT will achieve effective sufficient spatial resolution to image the black hole shadow of M87 if the mass-loading radius of the jet is not too large (Broderick & Loeb 2009; Lu et al. 2014; Akiyama et al. 2015).

Figure 3 shows the NRMSE metric for reconstructed images. The black curve labeled "Model" shows the NRMSE calculated when the model image is convolved with a circular Gaussian beam with a full width at half maximum (FWHM) as shown on the abscissa and compared against the original (unconvolved) model image. The Model curve effectively quantifies the best-case scenario in which the differences from the original input are due solely to a loss of resolution, not to errors in reconstructing the image. Figure 3 also shows the NRMSE of each of the reconstructed images convolved with circular Gaussian beams.

Figure 3 clearly shows that closure-phase imaging with ℓ_1 + TV regularization works well compared to CS-CLEAN, particularly at

---

[21] Previous works (e.g., Fish et al. 2014; Lu et al. 2014) derived position offsets between the model and reconstructed images by taking cross-correlations of the two images. However, we found that the position offsets derived from cross-correlations induce an additional error in the metrics due to errors in position offsets. We found that the center of mass for the image is a better indicator of the position offsets than the cross-correlation minimum.
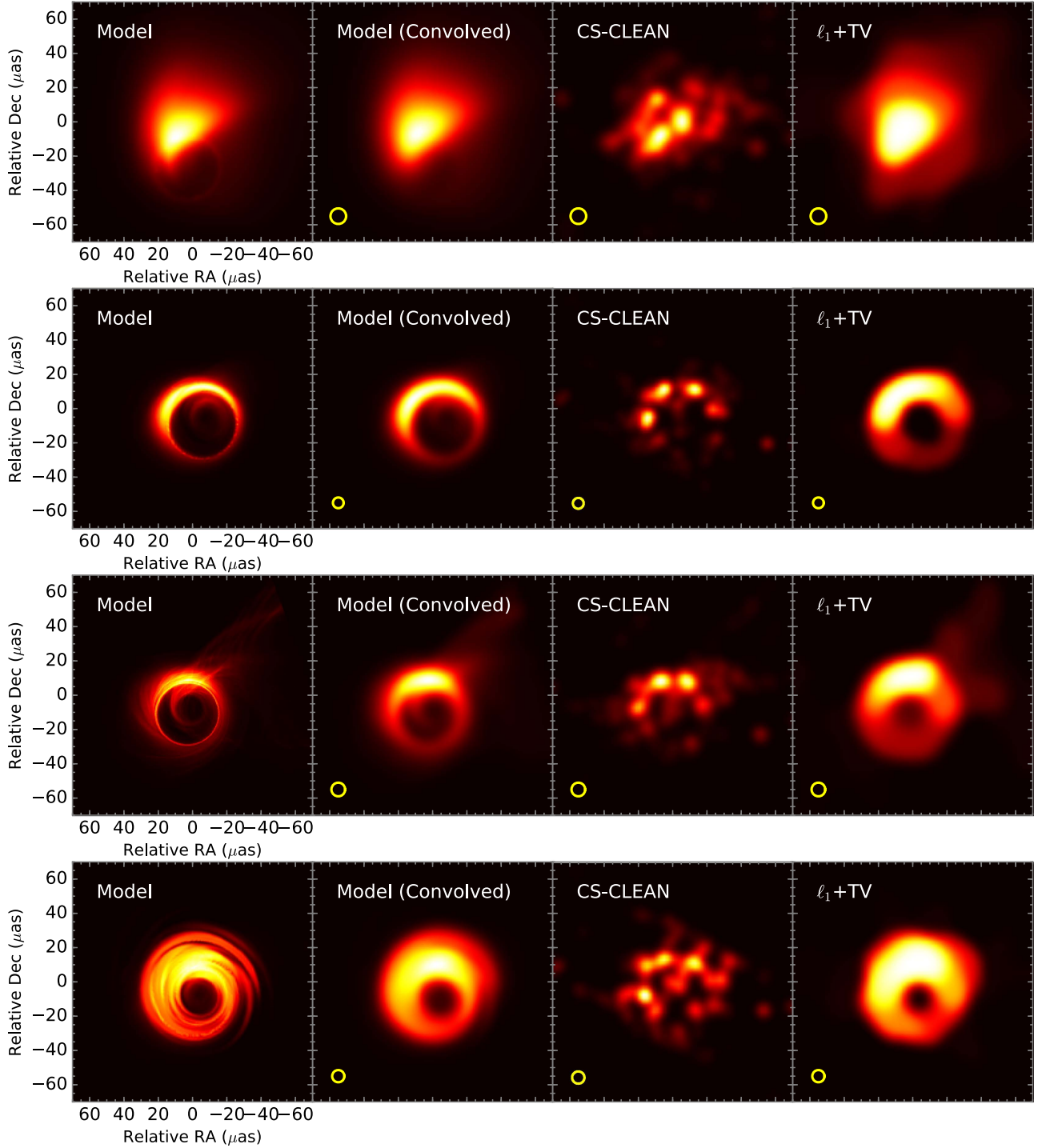[22] https://casa.nrao.edu/

Figure 2. Model and reconstructed images. All images are convolved with circular Gaussian beams with the FWMH sizes corresponding to diameters of the yellow circles, which coincide with the optimal resolutions for $\ell_1 + $ TV regularization shown in Figure 3. Top panels: the approaching-jet-dominated model BL09 taken from Akiyama et al. (2015; originally from Broderick & Loeb 2009 and Lu et al. 2014). The second and third panels from the top: the counter-jet-dominated models J2 taken from Akiyama et al. (2015; originally from Dexter et al. 2012) and M16 (Mościbrodzka et al. 2016), respectively. Bottom panels: the accretion-flow-dominated model DJ1 taken from Akiyama et al. (2015; originally from Dexter et al. 2012).

finer resolutions, despite the fact that CS-CLEAN uses full-complex visibilities with more information and higher SNRs than closure phases. Both CS-CLEAN and $\ell_1 + $ TV images achieve similar NRMSEs on scales comparable to or greater than the diffraction limit. On the other hand, the NRMSEs of the reconstructed images start to deviate from the model images in the super-resolution regime—namely on scales smaller than the

diffraction limit. In this regime, the NRMSEs differ by technique. CS-CLEAN has a common trend for all four models, which is broadly consistent with results of Chael et al. (2016). They achieve minimum errors at a resolution of ∼30%–60% of the diffraction limit and then show a rapid increase in errors at smaller scales. In contrast, closure-phase imaging with $\ell_1 + $ TV regularizations show much more modest variations in the super-resolution regime.
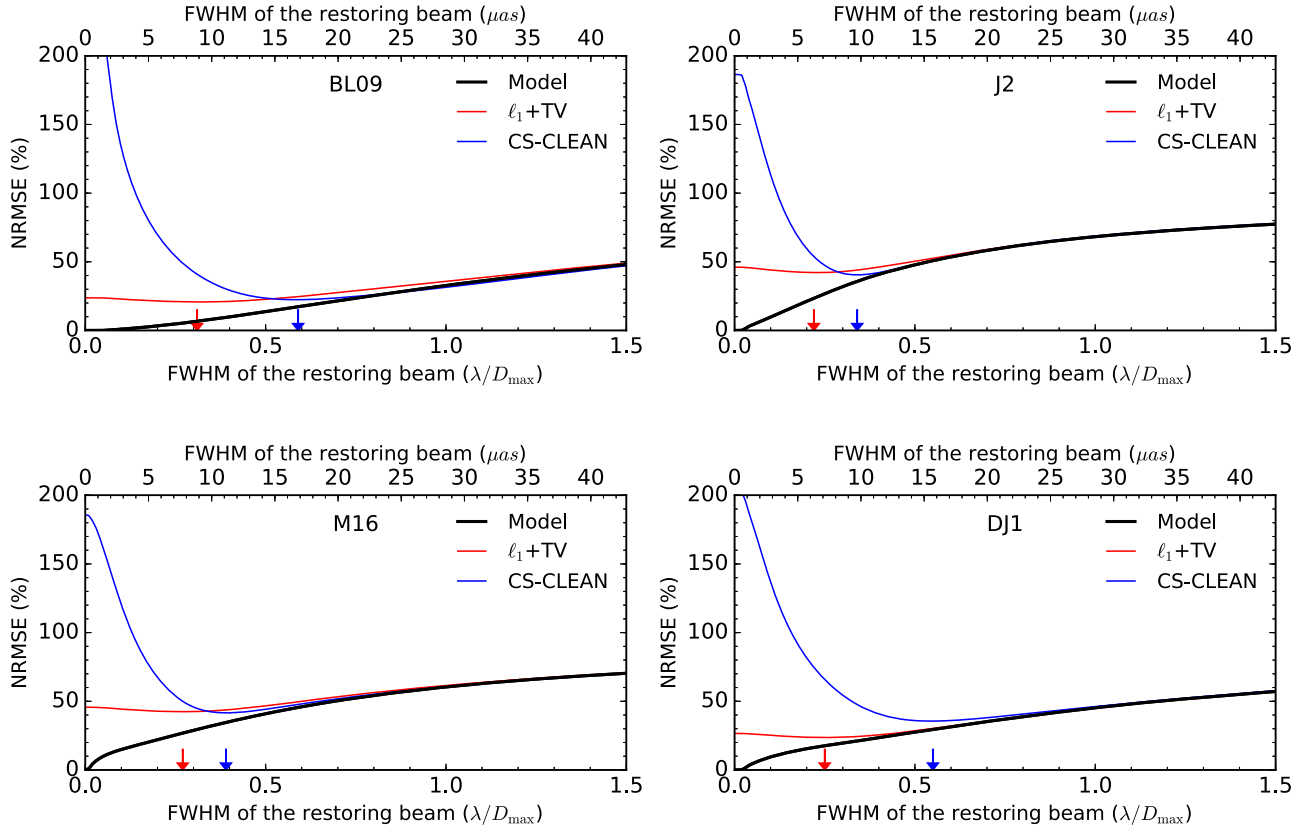
Figure 3. NRMSE between the non-beam-convolved original model image and beam-convolved model/reconstructed images of all four models, as a function of the FWHM size of the convolving circular beam. The red and blue arrows indicate the optimal resolution of $\ell_1 + \mathrm{TV}$ regularization and CS-CLEAN, respectively, which minimize the NRMSE.

They achieve minimum errors at a resolution of ~20%–30% of the diffraction limit, smaller than CS-CLEAN, and show only a slight increase at smaller scales. $\ell_1 + \mathrm{TV}$ reconstructions produce images that have a smooth distribution similar to the model images, resulting in smaller errors than CS-CLEAN, even if the $\ell_1 + \mathrm{TV}$ reconstructions are not convolved with a restoring beam.

In super-resolution regimes, the errors in $\ell_1 + \mathrm{TV}$ images mostly arise from the presence of tiny substructures in the image. For instance, we show model, reconstructed and residual images filtered with different baseline lengths for DJ1 in Figure 4. As shown in Figure 4, residuals are small when filtering baseline lengths are shorter than the maximum baseline length. On the other hand, for longer filtering baseline lengths, systematic residuals due to tiny substructures much smaller than the diffraction limit start to appear, which can be traced only with baselines longer than the simulated observations.

Although NRMSEs are better than CS-CLEAN at finer resolutions, $\ell_1 + \mathrm{TV}$ images have broader emission region sizes regardless of models. This is due to a typical feature of images reconstructed with the isotropic TV, which prefers flat images with sharp edges. Since the simulated data do not have visibilities at baseline lengths long enough to resolve the width of ring- or crescent-like features, TV enlarges their widths until images start to deviate from observed visibilities. This property of the isotropic TV regularization would be useful to constrain the upper-limit size of the emission regions and black hole shadow (see Section 5.1 for further discussions). On the other hand, in terms of the image fidelity, these results suggest that regularizations preferring much smoother edges are preferable;

smoother gradients in the image lead to images with higher contrast (i.e., brighter/fainter pixels become even brighter/ fainter, respectively) to conserve the total flux, which often makes the effective emission region size smaller. In Section 5.2, we discuss alternative regularizations of sparse imaging reconstruction for improving the image fidelity in the super-resolution regime.

4.2. Regularization Parameters and CV

Reconstructed images for all 16 sets of the regularization parameters are shown in Figure 5 for the accretion-flow-dominated model. As shown in Figure 5, the reconstructed images with $\tilde{\eta}_l$, $\tilde{\eta}_t \lesssim 1$ have noisy artifacts. This is true for all four models. Such artifacts may appear because the images are poorly constrained by both regularization functions at lower regularization parameters.

CV works as an Occam's razor and prevents over-fitting. In Figure 6(a), we show the residual χ^2 between the validation set and the model image reconstructed from the training set, which is averaged for all 10 trials (henceforth, the CV value). As expected in Section 2.2, the CV value tends to be large for large regularization parameters, since the regularization functions too strongly constrain the image so that the model image is inconsistent with observational data. Once the CV value achieves its minimum value (at a parameter set marked with stars in Figure 6(a)), it starts to increase again for lower regularization parameters, since the model image is over-fitted to the training set and then shows larger deviations from the validation set.
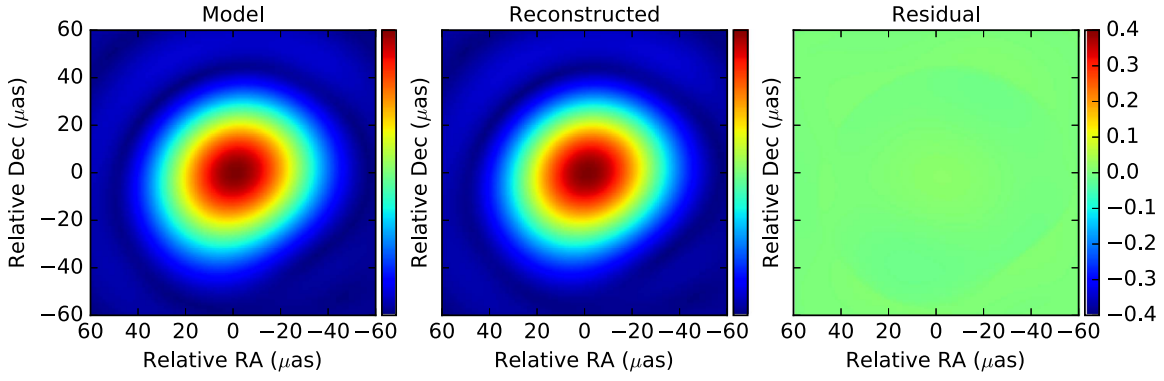
(a) Filtered with $|\mathbf{u}|_{\mathrm{cutoff}} = 0.5 D_{\mathrm{max}}/\lambda$
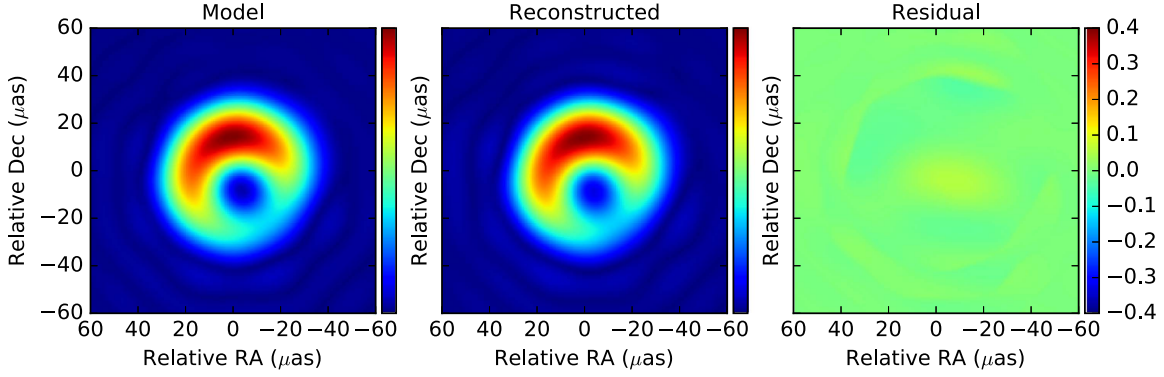
(b) Filtered with $|\mathbf{u}|_{\mathrm{cutoff}} = 1.0 D_{\mathrm{max}}/\lambda$
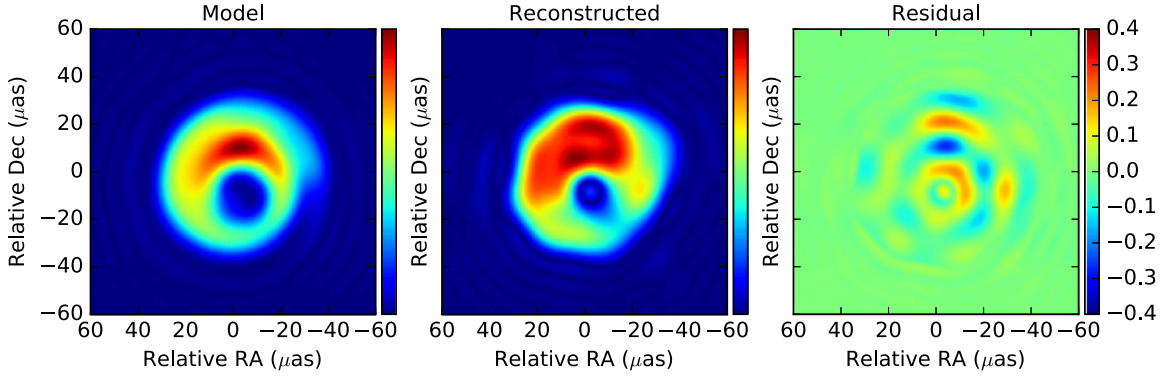
(c) Filtered with $|\mathbf{u}|_{\mathrm{cutoff}} = 2.0 D_{\mathrm{max}}/\lambda$
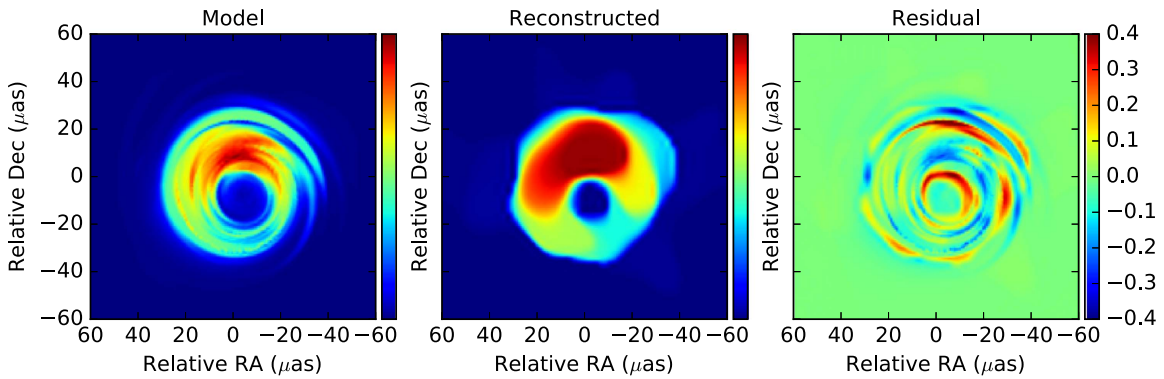
(d) Original Images (no filtering)

Figure 4. Model, reconstructed, and residual images of DJ1 for different filtering baseline lengths. The left and middle panels show the model and best-fit images, while the right panels show the difference between these two images normalized with the peak brightness of the model image. Images at panels (a), (b), and (c) are low-pass filtered with spatial frequencies (or baseline lengths) of $|\mathbf{u}|_{\mathrm{cutoff}} = (0.5, 1, 2) \times D_{\mathrm{max}}/\lambda$, which are equivalent to convolving images with modified Bessel functions with FWHM sizes of $\sim(1.2, 0.6, 0.3) \times \lambda/D_{\mathrm{max}}$, respectively. The panels (d) show original images without filtering.
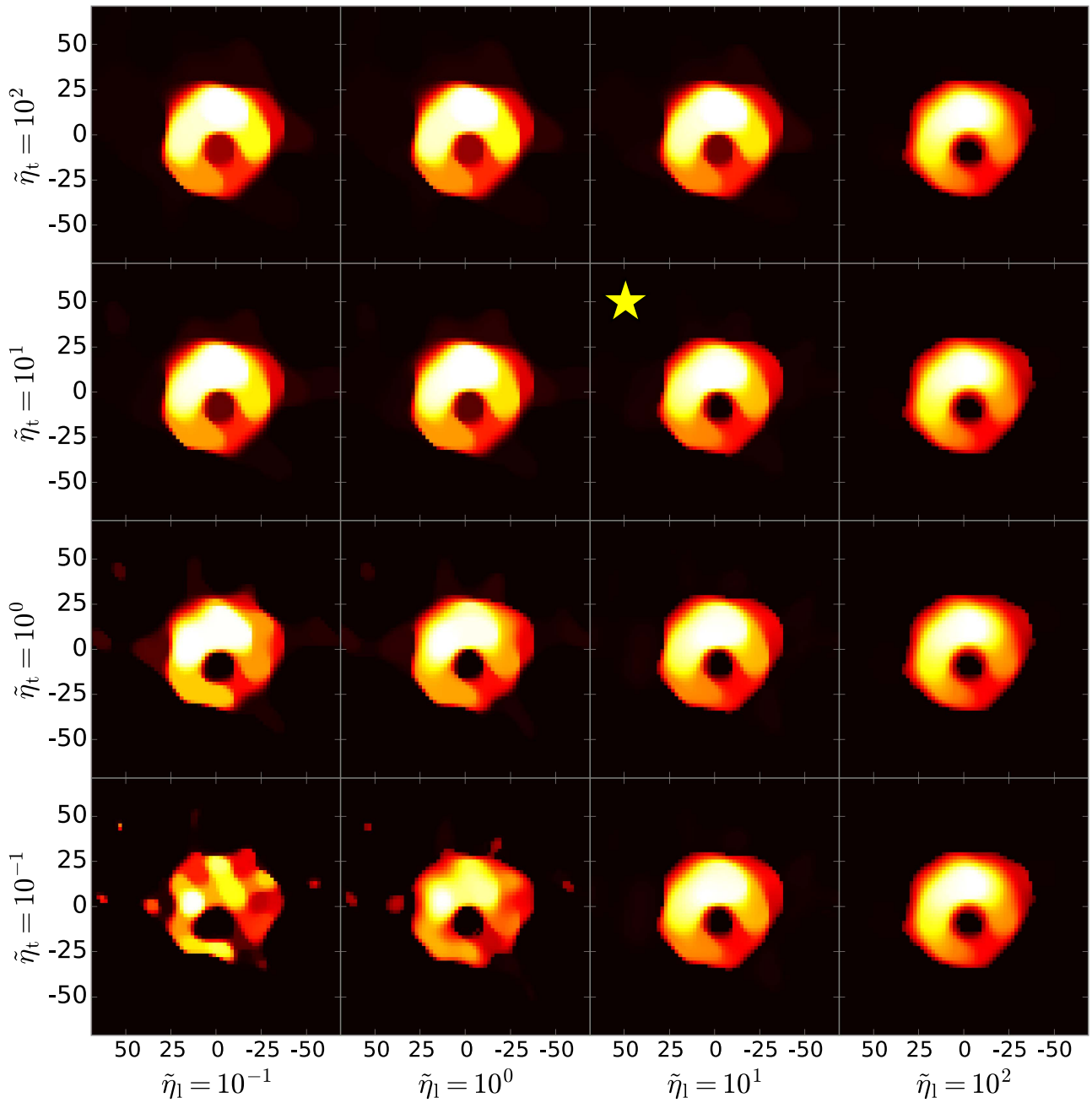
Figure 5. Parameter dependences of the reconstructed images in the regularization parameters for the accretion-flow-dominated model (DJ1). The yellow star indicates the regularization parameters for the best-fit images with the minimum CV. The units of the tick labels are μas.

Since the ground truth images are known in this work, one can compare CV-selected images with images on other parameter sets, which cannot be done for real observational data with unknown ground truth images. Although CV selects the optimal parameter based on the noise level of the data, it does not guarantee that the selected parameter achieves the best imaging fidelity among all parameter sets examined. In Figure 6(b), we show the NRMSE of non-Gaussian-convolved images for all parameter sets. Although metrics for the image fidelity of best-fit images are slightly larger than the minimum values for three of four models (J2, M16, and DJ1), they are consistent to within a few percent. These slight differences between the best-fit and best-fidelity parameters would not produce substantial differences in the image, and the resulting images are good enough.

The above results clearly demonstrate that CV is a useful technique to determine the regularization parameters so that the reconstructed image does not overfit noises in the data. We emphasize that CV is a general technique and can be applied to imaging regardless of the specific data products used (e.g., full-complex visibilities, visibility amplitudes, and/or closure quantities) or chosen regularization (for instance, sparse modeling; MEM: e.g., Buscher 1994, Chael et al. 2016; patch priors: e.g., Bouman et al. 2015; scattering optics: Johnson 2016).

5. Discussions

5.1. Implications for Future EHT Observations of M87

The proposed method successfully reproduces a clear signature for the black hole shadow for the models J2, M16,
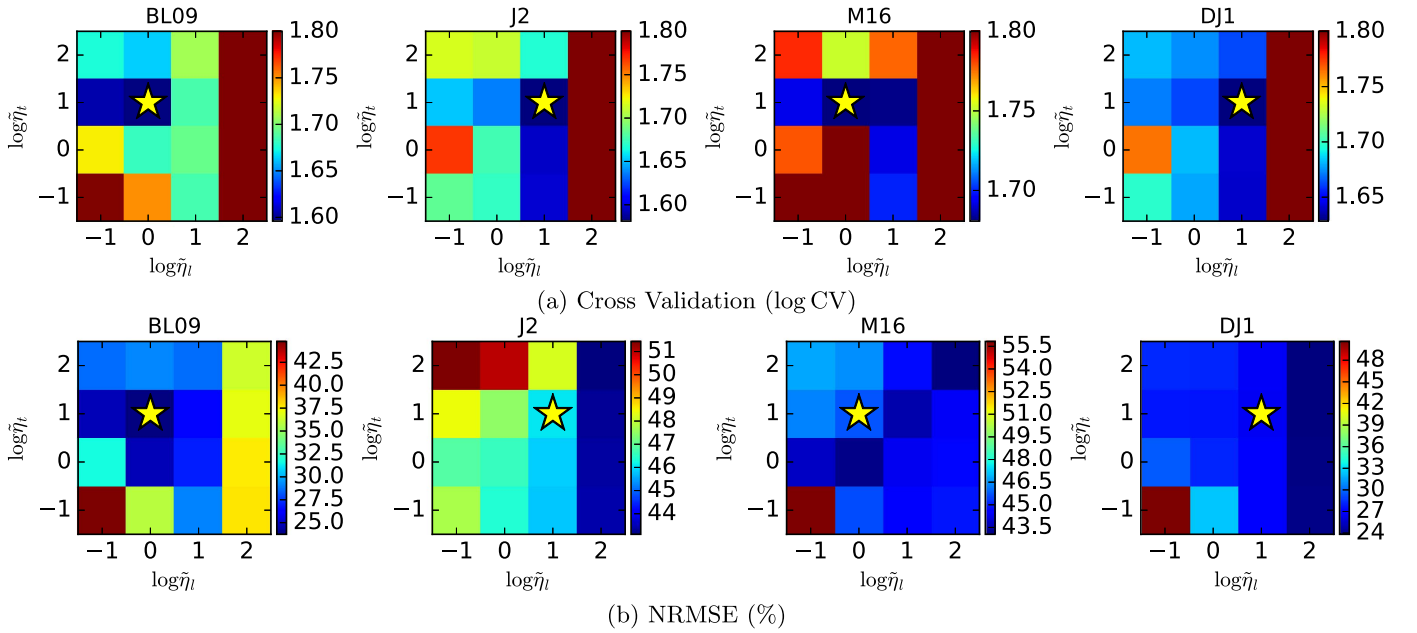
Figure 6. Logarithm of the CV value (a) and the NRMSE of the non-Gaussian convolved image (b) for models BL09, J2, M16, and DJ1. The yellow stars indicate the regularization parameters for the best-fit images with the minimum CV value. Note that, in panel (a), we set the upper limit of the color contours to values (1.8) lower than the maximum CV value (>2 for all models) in order to highlight differences in CV around the best-fit parameters.

and DJ1. This is as expected from the visibility distribution of these models, which have null amplitude regions, created by the shadow feature, at intermediate baseline lengths of ~3–4 Gλ (Akiyama et al. 2015). MEM also succeeds in reproducing the black hole shadow in these models (Lu et al. 2014). Presence of a clear shadow feature is tightly connected to where the dominant emission originates, since the silhouette of the black hole is created by the photons produced a few R_s from the black hole, illuminating the last photon orbit (see the discussion in Akiyama et al. 2015). As demonstrated in previous imaging simulations (Lu et al. 2014), future EHT observations can constrain the loading radius of the high-energy leptons producing synchrotron emission at 1.3 mm via the appearance of the black hole.

The most important implication of this work is that our regularization function TV and parameter selection with CV will enlarge the ring- or crescent-like emission illuminating the last-photon orbit of the black hole as much as possible, within the range that the model image neither over-fits nor deviates too much from the observed visibilities. Hence, the obtained width of the surrounding emission in the reconstructed image is close to an upper limit on the width of the emission region, and simultaneously, the obtained diameter of the black hole shadow should be interpreted as a reasonable lower limit for it. The clear shadow features in the reconstructed images for models J2, M16, and DJ1, therefore, strongly indicate that the EHT can sample a large enough range of visibilities with appropriately low noise levels to image the black hole shadow. In addition, the raw reconstructed image with the proposed method can be used to constrain the lower-limit size of the shadow. This would be useful to constrain the mass of M87, which has an uncertainty of an factor of about two between the stellar-dynamical (e.g., Gebhardt et al. 2011) and gas-dynamical (e.g., Walsh et al. 2013) modeling, and therefore be informative to clarify that which of modeling methods is desirable to measure the mass of supermassive black holes.

Another important implication is that, at least for the black hole images, post-processing Gaussian convolution would not be required with the $\ell_1 + TV$ regularization, although the CLEAN techniques do require it to reduce many compact artifacts in the image. As shown in Figure 3, the NRMSE curves for the $\ell_1 + TV$ regularization are shallow for smaller convolving sizes, and applying a circular Gaussian beam therefore makes only small improvements of a few percent in the NRMSE regardless of the input model images. Similar results are also shown in recent work with the MEM (Chael et al. 2016). Our results support that application of the beam is not required for the recent state-of-the-art imaging methods utilizing multi-resolution regularization functions for imaging the R_s-scale structure of M87 and Sgr A* with the EHT.

5.2. Relevant Future Issues

5.2.1. Other Sparse Regularization for Smoothed Images

In this paper, we adopt TV regularization, which favors a smooth image, so that images can be reconstructed with smaller pixel sizes and/or for more extended sources. As shown in Section 4, the reconstructed images have good image fidelity. In particular, it is noteworthy that CV-selected high values of $\tilde{\eta}_{lt} = 10^0$–10^1 for all models (see Figure 6(a)), suggesting that the solutions would be over-fitted without TV regularization. These results demonstrate that inclusion of the TV regularization can extend the range of objects where sparse modeling is applicable.

However, as described in Section 4.1, the reconstructed images have larger emission regions than the original model, since the isotropic TV (Rudin et al. 1992) adopted in this work prefers a flat brightness distribution in super-resolution regimes where the image cannot be constrained well. Hence, a regularization preferring more smoothed edges is required to improve the fidelity for the black hole imaging with the EHT.

In the context of sparse reconstruction, there are several candidates for improving the image fidelity as a natural extension of this work. First of all, there are other forms of regularization functions, which prefer sparse images in the gradient domain with smoother edges. For instance, an alternative form, given by,

$$||\boldsymbol{I}||_{\mathrm{tv}^2} = \sum_i \sum_j (|I_{i+1,j} - I_{i,j}|^2 + |I_{i,j+1} - I_{i,j}|^2). \qquad (14)$$

is also convex like the isotropic TV term adopted in this work, and prefers images with smoother edges. Furthermore, previous studies of sparse image reconstruction techniques have shown that regularization with ℓ_1+wavelet/curvelet transformation is also a promising approach (e.g., Li et al. 2011; Carrillo et al. 2012, 2014; Garsden et al. 2015; Dabbech et al. 2015). We will test these sparse regularizations in a forthcoming paper.

5.2.2. Enhancing Dynamic Range with Self-calibration

In VLBI, the visibility phase is initially calibrated with fringe fitting (also called fringe search), which is a self-calibration technique using phase closure (see Thompson et al. 2001). The fringe fitting can mitigate most station-based errors due to atmospheric and instrumental effects, though errors may remain if an incorrect source model is assumed. Traditionally, self-calibration with hybrid/differential mapping (e.g., Walker 1995) has been employed to solve for residual structural phase errors and images simultaneously, which has been successful for VLBI imaging.

This work and previous works on closure-phase imaging techniques using other regularizations such as MEMs (Fish et al. 2014; Lu et al. 2014, 2016; Chael et al. 2016) and patch priors (CHIRP; Bouman et al. 2016) demonstrate that an image can be reconstructed with high fidelity even from closure quantities. However, since closure phases and other closure quantities have less information about the source structure and also larger thermal noises than full-complex visibilities, imaging with closure quantities can limit the dynamic range, image sensitivity, and optimal spatial resolution.

A promising approach to improve the dynamic range is to use a reconstructed image from closure imaging techniques as an initial image for hybrid/differential mapping. It can also be used as a model for fringe fitting because self-calibrated images are often applied to detect more fringes on faint sources (e.g., Hada et al. 2016). In a forthcoming paper, we will evaluate the performance of such a hybrid-mapping technique including closure-phase/full-closure imaging priors. We emphasize that, for this purpose, one does not need to reconstruct the image with pixels much smaller than scales where the brightness distribution cannot be constrained by data and therefore will not affect results of self-calibration.

6. Conclusions

We have presented a new imaging technique reconstructing images from visibility amplitudes and closure phases by utilizing two regularizations of sparse modeling: the ℓ_1-norm and TV. Furthermore, we also propose a method to select optimal regularization parameters with CV, which can be applied to most existing imaging algorithms. As an example, we applied our technique to simulated observations of M87

with the EHT at 1.3 mm. Here, we summarize our conclusions.

1. We find that ℓ_1 + TV regularization can achieve an optimal resolution of $\sim$20%–30% of the diffraction limit λ/D_{max}, which is the nominal spatial resolution of a radio interferometer. This optimal resolution is better than that of the most widely used Cotton–Schwab CLEAN, which uses full-complex visibilities.
2. We confirm that CV works as an Occam's razor and prevents over-fitting when selecting the optimal regularization parameters. CV is a general method that can be applied to interferometric imaging more generally, such as imaging with full-complex visibilities and/or using other regularizations.
3. Using ℓ_1 + TV regularization, the reconstructed image maximizes the width of the emission region within the range that it neither over-fits nor deviates too strongly from the data. Hence, the clear reproduction of the black hole shadow in the reconstructed image suggests that future EHT observations will have the *uv*-coverage and sensitivity sufficient for imaging it. In addition, the reconstructed image will be able to constrain the sizes of the black hole shadow and surrounding emission region.

Finally, we remark that all of the above results demonstrate the clear promise of the EHT for providing an unprecedented view of the event-horizon-scale structure of the supermassive black hole in M87 and also the Galactic center Sgr A*.

References

Akiyama, K., Ikeda, S., Pleau, M., et al. 2017, AJ, 153, 159
Akiyama, K., Lu, R.-S., Fish, V. L., et al. 2015, ApJ, 807, 150
Bardeen, J. M. 1973, in Black Holes (Les Astres Occlus), ed. C. Dewitt & B. S. Dewitt (New York: Gordon and Breach), 215

Blakeslee, J. P., Jordán, A., Mei, S., et al. 2009, ApJ, 694, 556
Bouman, K. L., Johnson, M. D., Zoran, D., et al. 2016, IEEE Conf. on Computer Vision and Pattern Recognition, arXiv:1512.01413
Broderick, A. E., & Loeb, A. 2009, ApJ, 697, 1164
Buscher, D. F. 1994, in IAU Symp. 158, Very High Angular Resolution Imaging, ed. J. G. Robertson & W. J. Tango (Cambridge: Cambridge Univ. Press), 91
Byrd, R. H., Lu, P., Nocedal, J., & Zhu, C. 1995, SIAM Journal on Scientific Computing, 16, 1190
Candes, E. J., & Tao, T. 2006, ITIT, 52, 5406
Carrillo, R. E., McEwen, J. D., & Wiaux, Y. 2012, MNRAS, 426, 1223
Carrillo, R. E., McEwen, J. D., & Wiaux, Y. 2014, MNRAS, 439, 3591
Chael, A. A., Johnson, M. D., Narayan, R., et al. 2016, ApJ, 829, 11
Chatzopoulos, S., Fritz, T. K., Gerhard, O., et al. 2015, MNRAS, 447, 948
Dabbech, A., Ferrari, C., Mary, D., et al. 2015, A&A, 576, A7
Dexter, J., McKinney, J. C., & Agol, E. 2012, MNRAS, 421, 1517
Doeleman, S., Agol, E., Backer, D., et al. 2009, Astronomy, astro 2010: The Astronomy and Astrophysics Decadal Survey, 2010, 68
Doeleman, S. S., Fish, V. L., Schenck, D. E., et al. 2012, Sci, 338, 355
Doeleman, S. S., Weintroub, J., Rogers, A. E. E., et al. 2008, Natur, 455, 78
Donoho, D. L. 2006, ITIT, 52, 1289
Fish, V., Alef, W., Anderson, J., et al. 2013, arXiv:1309.3519
Fish, V. L., Doeleman, S. S., Beaudoin, C., et al. 2011, ApJL, 727, L36
Fish, V. L., Johnson, M. D., Doeleman, S. S., et al. 2016, ApJ, 820, 90
Fish, V. L., Johnson, M. D., Lu, R.-S., et al. 2014, ApJ, 795, 134
Garsden, H., Girard, J. N., Starck, J. L., et al. 2015, A&A, 575, A90
Gebhardt, K., Adams, J., Richstone, D., et al. 2011, ApJ, 729, 119
Hada, K., Doi, A., Kino, M., et al. 2011, Natur, 477, 185
Hada, K., Kino, M., Doi, A., et al. 2016, ApJ, 817, 131
Högbom, J. A. 1974, A&AS, 15, 417
Honma, M., Akiyama, K., Uemura, M., & Ikeda, S. 2014, PASJ, 66, 95
Ikeda, S., Tazaki, F., Akiyama, K., Hada, K., & Honma, M. 2016, PASJ, 68, 45
Jennison, R. C. 1958, MNRAS, 118, 276
Johnson, M. D. 2016, ApJ, 833, 74
Johnson, M. D., Fish, V. L., Doeleman, S. S., et al. 2015, Sci, 350, 1242
Li, F., Cornwell, T. J., & de Hoog, F. 2011, A&A, 528, A31
Lu, R.-S., Broderick, A. E., Baron, F., et al. 2014, ApJ, 788, 120
Lu, R.-S., Fish, V. L., Akiyama, K., et al. 2013, ApJ, 772, 13
Lu, R.-S., Fish, V. L., Weintroub, J., et al. 2012, ApJL, 757, L14
Lu, R.-S., Roelofs, F., Fish, V. L., et al. 2016, ApJ, 817, 173
Mallat, S. G., & Zhang, Z. 1993, ITSP, 41, 3397
McEwen, J. D., & Wiaux, Y. 2011, MNRAS, 413, 1318
Morales, J. L., & Nocedal, J. 2011, ACM Trans. Math. Softw., 38, 7
Mościbrodzka, M., Falcke, H., & Shiokawa, H. 2016, A&A, 586, A38
Narayan, R., & Nityananda, R. 1986, ARA&A, 24, 127
Obuchi, T., Ikeda, S., Akiyama, K., & Kabashima, Y. 2016, PLoSo, submitted (arXiv:1611.07197)
Obuchi, T., & Kabashima, Y. 2016, JSMTE, 2016, 053304
Rogers, A. E. E., Doeleman, S. S., & Moran, J. M. 1995, AJ, 109, 1391
Rogers, A. E. E., Hinteregger, H. F., Whitney, A. R., et al. 1974, ApJ, 193, 293
Rudin, L. I., Osher, S., & Fatemi, E. 1992, PhyD, 60, 259
Schwab, F. R. 1984, AJ, 89, 1076
Thiébaut, E. 2008, Proc. SPIE, 7013, 70131I
Thiébaut, É. 2013, in EAS Publications Ser. 59 New Concepts in Imaging: Optical and Statistical Models, ed. D. Mary, C. Theys, & C. Aime (Les Ulis: EDP Sciences), 157
Thompson, A. R., Moran, J. M., & Swenson, G. W., Jr. 2001, Interferometry and Synthesis in Radio Astronomy (2nd ed.; New York: Wiley-Interscience)
Tibshirani, R. 1996, Journal of the Royal Statistical Society. Series B (Methodological), 58, 267
Uemura, M., Kato, T., Nogami, D., & Mennickent, R. 2015, PASJ, 67, 22
Wagner, J., Roy, A. L., Krichbaum, T. P., et al. 2015, A&A, 581, A32
Walker, R. C. 1995, in ASP Conf. Ser. 82, Very Long Baseline Interferometry and the VLBA, ed. J. A. Zensus, P. J. Diamond, & P. J. Napier (San Francisco, CA: ASP), 247
Walsh, J. L., Barth, A. J., Ho, L. C., & Sarzi, M. 2013, ApJ, 770, 86
Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. 2004, ITIP, 13, 600
Wiaux, Y., Jacques, L., Puy, G. Scaife, A., M. M., & Vandergheynst, P. 2009a, MNRAS, 395, 1733
Wiaux, Y., Puy, G., Boursier, Y., & Vandergheynst, P. 2009b, MNRAS, 400, 1029
Wiaux, Y., Puy, G., & Vandergheynst, P. 2010, MNRAS, 402, 2626
Zhu, C., Byrd, R. H., Lu, P., & Nocedal, J. 1997, ACM Trans. Math. Softw., 23, 550