The Astrophysical Journal, 829:11 (15pp), 2016 September 20 doi:10.3847/0004-637X/829/1/11

© 2016. The American Astronomical Society. All rights reserved.

HIGH-RESOLUTION LINEAR POLARIMETRIC IMAGING FOR THE EVENT HORIZON TELESCOPE

Andrew A. Chael[1], Michael D. Johnson[1], Ramesh Narayan[1], Sheperd S. Doeleman[1,2],
John F. C. Wardle[3], and Katherine L. Bouman[4]
[1] Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA; achael@cfa.harvard.edu
[2] Massachusetts Institute of Technology, Haystack Observatory, Route 40, Westford, MA 01886, USA
[3] Brandeis University, Physics Department, Waltham, MA 02454, USA
[4] Massachusetts Institute of Technology, Computer Science and Artificial Intelligence Laboratory, 32 Vassar Street, Cambridge, MA 02139, USA

Received 2015 December 4; revised 2016 June 8; accepted 2016 July 6; published 2016 September 14

ABSTRACT

Images of the linear polarizations of synchrotron radiation around active galactic nuclei (AGNs) highlight their projected magnetic field lines and provide key data for understanding the physics of accretion and outflow from supermassive black holes. The highest-resolution polarimetric images of AGNs are produced with Very Long Baseline Interferometry (VLBI). Because VLBI incompletely samples the Fourier transform of the source image, any image reconstruction that fills in unmeasured spatial frequencies will not be unique and reconstruction algorithms are required. In this paper, we explore some extensions of the Maximum Entropy Method (MEM) to linear polarimetric VLBI imaging. In contrast to previous work, our polarimetric MEM algorithm combines a Stokes I imager that only uses bispectrum measurements that are immune to atmospheric phase corruption, with a joint Stokes Q and U imager that operates on robust polarimetric ratios. We demonstrate the effectiveness of our technique on 7 and 3 mm wavelength quasar observations from the VLBA and simulated 1.3 mm Event Horizon Telescope observations of Sgr A* and M87. Consistent with past studies, we find that polarimetric MEM can produce superior resolution compared to the standard CLEAN algorithm, when imaging smooth and compact source distributions. As an imaging framework, MEM is highly adaptable, allowing a range of constraints on polarization structure. Polarimetric MEM is thus an attractive choice for image reconstruction with the EHT.

Key words: black hole physics – Galaxy: center – techniques: high angular resolution – techniques: image processing – techniques: interferometric

1. INTRODUCTION

Magnetic fields in plasmas around compact objects such as pulsars and black holes are critical for powering these objects' energetic emissions. In active galactic nuclei (AGNs), magnetic fields in accretion disks around central supermassive black holes regulate outflow and the interaction of the AGN with its host galaxy (Fabian 2012; Kormendy & Ho 2013). Stochastic magnetic fields in the disk drive turbulence via the magnetorotational instability (Balbus & Hawley 1998), enabling efficient accretion and conversion of gravitational energy into near-Eddington luminosities. Near the horizon, magnetic fields can convert spin energy from the black hole into energetic jets of plasma (Blandford & Znajek 1977; Blandford & Payne 1982). Synchrotron emission from electrons in AGN cores and jets is characteristically linearly polarized, with a polarization direction that traces magnetic field lines in the plasma (Ginzburg & Syrovatskii 1965). Faraday rotation of the polarization vector as it propagates through the plasma is determined by the magnetic field strength and electron density along the line of sight (Rybicki & Lightman 1979). Thus, measurements of the linear polarization direction and its frequency dependence characterize the magnetic field around these objects and provide observational windows into these objects' fundamental physics. For a recent review of observations of magnetic fields in AGNs, see Wardle (2013).

Linear Polarimetric Very Long Baseline Interferometry (VLBI) at centimeter and millimeter wavelengths can measure polarization magnitude and orientation at high angular resolutions of fractions of a milliarcsecond (Roberts et al. 1994; Attridge 2001; Attridge et al. 2005). VLBI observations provide an incomplete sample of the Fourier transform of the sky flux density distribution, so image reconstruction (or deconvolution) algorithms are needed. The standard reconstruction algorithm is CLEAN (Högbom 1974), which models the image as a collection of point sources. CLEAN produces images of the three Stokes parameters I, Q, and U separately, so unphysical fractional polarizations $m = \sqrt{Q^2 + U^2}/I > 1$ are possible (especially in regions of low total intensity I). In contrast, Bayesian regularization methods can naturally incorporate prior information of the image's spatial distribution and physical constraints such as $m \leqslant 1$. One such regularization method is the Maximum Entropy Method, (MEM), which finds the image that is most consistent with the data that maximizes an entropy function, analogous to the log of a prior probability distribution. MEM imaging algorithms have been in use for decades (e.g., Gull & Daniell 1978 and Cornwell & Evans 1985), but because of early computational limitations, they are used infrequently compared to CLEAN. The theory behind polarimetric MEM was pioneered in several theoretical papers beginning in the 70s (Ponsonby 1973; Nityananda & Narayan 1983; Narayan & Nityananda 1986; Shevgaonkar 1987), but implementation on actual VLBI data has been limited to only a handful of studies since (Holdaway 1990; Holdaway & Wardle 1990; Sault et al. 1999; Coughlan & Gabuzda 2012, 2013).

Polarimetric MEM is particularly promising for imaging the accretion flow and jets near supermassive black holes observed by the Event Horizon Telescope (EHT). The EHT is a global 1.3 mm VLBI array that will obtain nominal resolutions of approximately 25 microarcseconds, allowing observations of nearby supermassive black holes at scales on the order of the projected Schwarzschild radius (Doeleman et al. 2009). Previous observations with three EHT baselines have constrained the size of the 1.3 mm emission region to scales on the order of the lensed Schwarzschild radius in Sgr A* (Doeleman et al. 2008; Fish et al. 2011) and M87 (Doeleman et al. 2012). The first polarimetric observations with the EHT in 2013 resolved the polarimetric emission in Sgr A*, providing strong evidence for ordered magnetic fields near the event horizon (Johnson et al. 2015). Future observations are expected to obtain enough data to construct an image of the Sgr A* black hole accretion flow (Fish et al. 2014) and jet base of M87 (Lu et al. 2014). As the EHT is now capable of observing with full polarization, polarimetric MEM provides an attractive solution to creating full-polarization images of these sources.

In this paper, we develop an application of MEM to polarimetric VLBI data. In Section 2, we review the fundamentals of polarimetric VLBI, as well as the standard CLEAN algorithm for image reconstruction. In Section 3, we review standard MEM, applied to total intensity images. We discuss applications of MEM to data without calibrated phase information, and quantify the resolution and fidelity of MEM images. We then move to polarimetric MEM in Section 4, where we introduce and compare two forms of the polarimetric entropy function.

In Section 5, we discuss our implementation of polarimetric MEM, including the details of our simulation of EHT data and our minimization algorithm. We present results from applying polarimetric MEM to 7 mm VLBA observations of the quasar 3C279 and 3 mm VLBA observations of 3C273, as well as simulated EHT data from several 1.3 mm model images of Sgr A* and M87. We compare the effects of different forms of the entropy function, and we test the ability of polarimetric MEM to resolve polarization field structure. We show that for these high-frequency VLBI observations, polarimetric MEM is capable of reproducing the general morphology of the CLEAN images, but with typically higher resolution and fidelity. Finally, in Section 6, we outline future directions for polarimetric MEM in VLBI and synthesis imaging in interferometry more broadly.

Our imaging code is written in Python and uses the Limited-Memory-Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) minimization routine in the Scipy package. Our imaging programs, as well as a variety of routines for simulating and manipulating VLBI data, are available at https://github.com/achael/eht-imaging.

2. FUNDAMENTALS OF INTERFEROMETRIC IMAGING

By the Van Cittert–Zernike theorem, measured interferometric visibilities $\tilde{I}_k$ are the Fourier components of the true source image $I(x, y)$ (in total intensity) plus thermal noise n_k (Thompson et al. 2007)(hereafter TMS):

$$\tilde{I}_k = \int\int I(x, y)e^{-2\pi i(u_k x + v_k y)}dx\, dy + n_k. \tag{1}$$

Here, x and y are real space angular coordinates and u_k, v_k are the interferometric baseline coordinates projected orthogonal to the source line of sight and measured in wavelengths.

An interferometer incompletely samples the u–v plane, so direct Fourier transform (i.e., the "dirty image") of the measured visibilities is a convolution of the true image and the Fourier transform of the u–v sample coverage (the "dirty beam"). VLBI imaging can thus be approached either as deconvolution of the dirty beam from the dirty image or as fitting a model to visibility data with regularizing constraints. Finite sampling also ensures that no image that reproduces the observed visibilities $\tilde{I}_k$ is unique; extra information is always required to constrain the image. In CLEAN, this extra information is the representation of the sky image in terms of a finite number of point sources. MEM allows for many potential regularizing constraints through the use of different entropy functions. Furthermore, MEM naturally incorporates uncertainties due to thermal noise and quantifies the goodness-of-fit in a standard χ^2 metric. This makes MEM a natural choice for sparse or heterogeneous VLBI arrays such as the EHT.

A further complication in high-frequency VLBI is that atmospheric fluctuations make stable phase information on individual baselines impossible. However, adding the phases observed on three baselines around a triangle cancels the atmospheric contribution at each station, so these *closure phases* only contain information about the source. Assuming that the visibility amplitudes can be calibrated to remove station-dependent gain terms (which vary more slowly than the unstable phase terms from the atmosphere), calibrated amplitudes can be combined with closure phases in the bispectrum, the product of three simultaneous visibilities around a triangle (Rogers et al. 1974), (TMS),

$$\tilde{I}_B = \tilde{I}_{12}\, \tilde{I}_{23}\, \tilde{I}_{31}, \tag{2}$$

where, for instance, $\tilde{I}_{12}$ is the measured visibility between stations 1 and 2 at a given time.

Because the image of linear polarization is a two-dimensional vector field defining both the magnitude and direction of the linear polarization at each location, it can be represented as a complex image, $P(x, y) = Q(x, y) + iU(x, y)$, where $Q(x, y)$ and $U(x, y)$ are the images of the linear Stokes parameters. The linearly polarized image can also be expressed in terms of the polarization fraction $m(x, y)$ and polarization position angle $\chi(x, y)$ as $P = m\, I\, e^{2i\chi}$, where $m = |P|/I = \sqrt{Q^2 + U^2}/I \leqslant 1$, and $\chi = \frac{1}{2}\arctan\frac{U}{Q}$. Polarimetric visibilities, $\tilde{Q}_k$ and $\tilde{U}_k$, are also related to the images $Q(x, y)$ and $U(x, y)$ via the van Cittert–Zernike theorem (Equation (1)). For synthesis imaging, the most significant difference between polarization and total flux is that the images of Q and U are not constrained to be positive and that the total polarization fraction in each pixel is constrained to be less than one.

Because the atmosphere is not significantly birefringent at the high frequencies we are considering (TMS), the atmospheric contribution to phase is identical for all of the visibilities $\tilde{I}_k$, $\tilde{Q}_k$, and $\tilde{U}_k$, so polarimetric ratios such as $\tilde{Q}_k/\tilde{I}_k$ provide the same immunity to station-based phase errors as closure phase (Roberts et al. 1994). We define the visibility domain polarimetric ratio and phase (following the notation of Johnson et al. 2014)

$$\breve{m}_k = \frac{\tilde{P}_k}{\tilde{I}_k}. \tag{3}$$

It is important to note that $\breve{m}$ is *not* the Fourier transform of the image plane polarization fraction m. In particular, $\breve{m}$ is not conjugate-symmetric under the reversal of baselines

$(u, v) \rightarrow (-u, -v)$. It also is possible for the magnitude $|\breve{m}|$ to exceed unity, if, for example, the total intensity visibility $\tilde{I}$ has a "null" on some baseline, due to the presence of some sharp feature in the image (Johnson et al. 2015).

3. TOTAL INTENSITY MEM

The standard CLEAN algorithm operates on the dirty image obtained from Fourier transforming the sparse interferometer data and treats the imaging process as a deconvolution of the dirty beam from the image (Högbom 1974). CLEAN models the sky brightness distribution as a collection of point sources. It determines the locations and magnitudes of these point sources iteratively by finding the maximum intensity pixel of the dirty image, then subtracting the shifted and scaled "dirty beam." After a certain number of iterations, CLEAN convolves the point source model with a "clean" beam obtained by fitting a Gaussian to the central component of the dirty beam. The algorithm halts after the maximum brightness point in an image drops below some multiple of the residual rms level, or after negative components start to be removed. Finally, the dirty image residuals are added to the restored image to include low-intensity diffuse brightness distributions that are poorly captured by the point source decomposition. Because CLEAN relies on absolute visibility phase information to perform the inverse Fourier Transform to the dirty image at each step of the algorithm, visibility phases corrupted by atmospheric phase fluctuations must either be calibrated or self-calibrated in a loop with multiple iterations of CLEAN (TMS).

In contrast, MEM as described in this paper operates directly on the measured visibilities or robust quantities like closure phases or the bispectrum. In MEM and other Bayesian regularization imaging methods, an image is fitted to the data by minimizing a weighted sum of χ^2 and a regularizing function that incorporates prior information. With this approach, only the forward Fourier transform from the trial image to the visibility domain is used, and trial image visibilities can be directly compared with measured, calibrated visibilities or other data products derived from the measured visibilities. Furthermore, for sparse VLBI arrays, we can avoid sampling errors introduced by transforming with a Fast Fourier Transform (FFT) and compute trial image visibilities at the sampled baseline points with a discrete-time Fourier Transform (DTFT).[5]

In MEM, we maximize a regularizing function, or "entropy" of the image, with respect to data constraints. In what follows, we denote all arrays of image pixels or visibilities in bold. For a n^2 pixel test image $\boldsymbol{I}'$, a prior/bias image $\boldsymbol{B}$, and an array of N measured visibilities $\tilde{\boldsymbol{I}}$, we maximize the objective function (Narayan & Nityananda 1986):

$$J = S(\boldsymbol{I}', \boldsymbol{B}) - \alpha(\chi^2(\boldsymbol{I}') - 1), \tag{4}$$

where $S(\boldsymbol{I}', \boldsymbol{B})$ is the chosen regularizer or entropy function and χ^2 is the goodness-of-fit test statistic that compares the visibilities of the test image $\boldsymbol{I}'$ to the data. The mixing coefficient α controls the weighting between the regularizer (entropy) term and the data (χ^2) term. Considering MEM as a form of constrained optimization, α plays the role of a Lagrange multiplier. In practice, it can be fixed, iterated manually, or be allowed to vary in the maximization process.

The goodness-of-fit χ^2 term is defined in the visibility domain:

$$\chi^2(\boldsymbol{I}') = \frac{1}{2N}\sum_{k=1}^{N}\frac{1}{\sigma_k^2}|\tilde{I}_k - \tilde{I}'_k|^2, \tag{5}$$

where σ_k is the noise estimate on the kth u, v point and the model visibilities I'_k are the DTFT of the test image evaluated at the kth u, v point. The factor of 2 in the denominator of Equation (5) is included because the variance σ_k^2 is taken to be the variance along either the real or imaginary axis; in fitting the data, we must fit the real and imaginary parts of the visibilities separately. Assuming that the visibilities are normally distributed, χ^2 will possess a χ^2 distribution, and a good fit where the trial visibilities agree with the measurements within error has $\chi^2 \approx 1$.

The standard entropy function motivated from information theory is (Frieden 1972; Gull & Daniell 1978)

$$S(\boldsymbol{I}', \boldsymbol{B}) = -\sum_{i=1}^{n^2} I'_i \log\left(\frac{I'_i}{B_i}\right), \tag{6}$$

but many other entropy functions can be chosen, including $S(\boldsymbol{I}') = \sum \log(I'_i)$, $S(\boldsymbol{I}') = \sum\sqrt{I'_i}$ (Narayan & Nityananda 1986), or the ℓ_1 norm $S(\boldsymbol{I}') = \sum|I'_i|$ (Honma et al. 2014). In fact, it can be shown that for any convex function $S(\boldsymbol{I}', \boldsymbol{B})$ of the I_i, the reconstruction is guaranteed to converge (Narayan & Nityananda 1986).

To deal with phase uncertainty in MEM without needing to calibrate or self-calibrate the visibility phases, we can extend this technique to the image bispectrum, replacing the data term χ^2 with its bispectral extension. In this method, unlike in a self-calibration loop as used with CLEAN, the visibility phases are not calibrated prior to imaging. The objective function becomes

$$J_B = S(\boldsymbol{I}', \boldsymbol{B}) - \alpha(\chi_B^2(\boldsymbol{I}') - 1), \tag{7}$$

where the bispectrum data term is

$$\chi_B^2(\boldsymbol{I}') = \frac{1}{2N_B}\sum_{j=1}^{N_B}\frac{1}{\sigma_{Bj}^2}|\tilde{I}_{Bj} - \tilde{I}'_{Bj}|^2. \tag{8}$$

In the above equation, we have N_B *independent* bispectrum measurements $\tilde{\boldsymbol{I}}_{Bj}$, each with standard deviation σ_{Bj}. At any instant in time with detections on all baselines to T sites, there are $\binom{T}{3} = T!/3!(T-3)!$ triangles but only $(T-1)(T-2)/2$ independent bispectrum measurements (TMS). In our reconstructions, we constructed a set of independent bispectra at each time, using the criterion that each triangle contain the station with the highest signal-to-noise ratio.

Using the bispectrum for MEM image reconstruction was pioneered in optical interferometry with the BSMEM gradient descent algorithm (Buscher 1994; Baron & Young 2008), which has been successfully used on simulated EHT observations (Fish et al. 2014; Lu et al. 2014). Recent developments using the bispectrum directly in image reconstruction include the CHIRP algorithm (Bouman et al. 2016), which uses a data-driven regularizing function based on features found in a library of sample images instead of a standard entropy term based on a single prior image.

---

[5] While the term "discrete-time Fourier transform" refers to time as the discretely sampled interval, in our case the transform is spatial and the discretely sampled interval is the image angular coordinate.

Because the bispectrum may not include any significantly small triangles to tightly constrain the unresolved flux, it is particularly useful to add a total flux constraint to the objective function (Equation (4)):

$$J_B \rightarrow J_B + \gamma\left[\sum_{i=1}^{n^2} I'_i - F_{\mathrm{obs}}\right]^2, \tag{9}$$

where F_{obs} is the observed total flux density of the source and γ is a hyperparameter that controls the relative weighting of the flux constraint compared to the χ^2_B and entropy term in Equation (7). In addition, since the bispectrum lacks overall phase information, the bispectrum data carry no information about the absolute image position and the final image centroid is arbitrary. Choosing image frame coordinates where $x = 0$, $y = 0$ corresponds to the center of the frame, we can add a center-of-mass constraint to the objective function:

$$J_B \rightarrow J_B + \delta\left[\sum_{i=1}^{n^2} I'_i x_i + \sum I'_i y_i\right]^2, \tag{10}$$

where δ is another hyperparameter to control the weight of this constraint with respect to the data term, entropy, and any other constraints. Finally, imaging with the bispectrum carries an additional complication in that the bispectrum values are not necessarily described by a normal distribution, and thus the bispectrum χ^2_B statistic is not as straightforward to interpret as it is in the case where calibrated visibility phases are available. In the high signal-to-noise limit, however, the distribution of the bispectrum points approaches a Gaussian (Rogers et al. 1995).

We can implement a total intensity MEM algorithm using a quasi-Newton gradient descent method. The derivatives of S and χ^2 or S_B and χ^2_B with respect to the pixel values I_i can be computed analytically and evaluated at each step. The inverse Hessian can be approximated by neglecting the off-diagonal terms, as in the Cornwell and Evans algorithm (Cornwell & Evans 1985), or through numerical approximation as in the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm (Byrd et al. 1995).

Because nonlinear methods like MEM and CLEAN input prior information into the imaging process, we might expect some degree of image "superresolution," or the production of image features on scales less than the array nominal resolution $R_{\min} = \lambda/b_{\max}$, where $b_{\max}$ is the length of the longest baseline in the VLBI array. The frequently quoted result that MEM has a superresolution factor of 1/4 the nominal resolution is in fact *not* based on the image prior, but only on the analyticity of the data (Narayan & Nityananda 1986). The derivation of this factor requires the unrealistic assumption of infinite signal-to-noise (Holdaway 1990). In practice, the superresolution factor may be informed by both the analyticity of the data (degraded by noise) and the entropy function. Simple tests comparing blurred MEM to the model source distribution suggest that in practice, total intensity MEM can achieve a superresolution factor of 1/3 to 1/2 the nominal resolution (see Section 5.3).

4. LINEAR POLARIMETRIC MEM

To extend MEM to complex linear polarized images while avoiding atmospheric phase corruption, we maximize the objective function (Holdaway 1990):

$$J_m = S_m(\boldsymbol{P}') - \beta(\chi^2_m(\boldsymbol{I}', \boldsymbol{P}') - 1). \tag{11}$$

Here $\boldsymbol{P}'$ is our trial image of the polarized flux and χ^2_m is the test statistic that compares the polarimetric ratios of the test image to the data:

$$\chi^2_m(\boldsymbol{I}', \boldsymbol{P}') = \frac{1}{2N}\sum_{k=1}^{N}\frac{1}{\sigma^2_{m\,k}}|\breve{m}_k - \breve{m}'_k|^2, \tag{12}$$

where $\breve{m}_k = \tilde{P}_k/\tilde{I}_k$ is the polarimetric ratio on the kth u, v point, which is insensitive to phase errors. As in the bispectral imaging of total intensity, this MEM technique does not reconstruct the phases on $\tilde{Q}$ and $\tilde{U}$ directly before imaging, but instead relies on the robust, measurable polarimetric ratios $\breve{m}$ to guide the imaging directly. Thus, computing χ^2_m requires a $\boldsymbol{I}'$ reconstruction and its visibility phases—when using the polarimetric ratios in this manner, we cannot image $\boldsymbol{P}'$ independently from $\boldsymbol{I}'$.

The conventional polarimetric entropy, first developed from the eigenvalues of the Stokes parameter correlation matrix, is (Ponsonby 1973; Nityananda & Narayan 1983; Narayan & Nityananda 1986; Holdaway & Wardle 1990)

$$\begin{aligned} S_m(\boldsymbol{P}') = -\sum_{i=1}^{n^2} I'_i \Bigg[&\frac{m_{\max} + m'_i}{2}\log\left(\frac{m_{\max} + m'_i}{2}\right) \\ &+ \frac{m_{\max} - m'_i}{2}\log\left(\frac{m_{\max} - m'_i}{2}\right)\Bigg]. \end{aligned} \tag{13}$$

The quantity $m_{\max}$ is the maximum fractional polarization; this can be generically set to 1. For synchrotron sources we can instead set $m_{\max} \approx 0.75$, as was done by Holdaway & Wardle (1990) to ensure that the degree of polarization remains limited to the expected maximum for power-law synchrotron emission (Rybicki & Lightman 1979). This entropy naturally favors images with $|m'_i| < m_{\max}$. However, it contains no information about the polarization direction and it tends to favor low polarization magnitudes, as it is maximized in the absence of data constraints when $m'_i = 0$ in all pixels.

Another possibility is to use a simple log entropy, as one might use for total intensity images (e.g., Ponsonby 1973; Nityananda & Narayan 1983; Narayan & Nityananda 1986):

$$S_m(\boldsymbol{P}') = -\sum_{i=1}^{n^2}|P'_i|\log|P'_i|. \tag{14}$$

Because the polarization field traces the magnetic field structure, which we expect to not be completely disordered in resolved images, we may want to move beyond pixel-by-pixel entropy terms to gradient-based regularizing functions. One option is to use a regularizer that is proportional to the total variation of the trial image $\boldsymbol{P}'$. The isotropic total variation TV of a complex image matrix $\boldsymbol{X}$ typically used for image reconstruction and denoising (Rudin et al. 1992) is

$$\mathrm{TV}(\boldsymbol{X}) = \sum_{i=1}^{n}\sum_{j=1}^{n}\sqrt{|X_{i+1,j} - X_{i,j}|^2 + |X_{i,j+1} - X_{i,j}|^2}. \tag{15}$$

Adopting the total variation of the complex polarimetric image as our "entropy" term, we take $S(\boldsymbol{P}') = -\mathrm{TV}(\boldsymbol{P}')$.

The total variation entropy constrains MEM to prefer smooth polarization fields in both direction and magnitude. However, the gradient of Equation (15) becomes infinite for uniform images, so care must be taken in the minimization algorithm, especially when determining the initial test image. The actual imaging algorithm can again use quasi-Newton or conjugate gradient methods, and can operate on either the Q and U arrays or the m and χ images.

Finally, we can combine the polarimetric and total flux terms into a joint objective function for simultaneous imaging of $\boldsymbol{I}'$ and $\boldsymbol{P}'$ with the bispectrum and polarimetric ratios:

$$J_{\mathrm{tot}} = S_{\mathrm{tot}}(\boldsymbol{I}', \boldsymbol{P}', \boldsymbol{B}) - \alpha(\chi^2_B(\boldsymbol{I}') - 1) - \beta(\chi^2_m(\boldsymbol{I}', \boldsymbol{P}') - 1) + \text{constraints}. \tag{16}$$

Here $S_{\mathrm{tot}}(\boldsymbol{I}', \boldsymbol{P}', \boldsymbol{B})$ is a joint entropy function, such as a combination of Equations (6) and (13), and the constraints can include terms constraining the total flux density (Equation (9)) or image centroid (Equation (10)). Joint imaging can use data in the polarimetric ratios to constrain the total intensity image, and is thus the most theoretically sound method of MEM imaging polarized fields. In practice, however, convergence to the true image is poor if we allow both $\boldsymbol{I}'$ and $\boldsymbol{P}'$ to vary starting from a flat or Gaussian initial image (Holdaway & Wardle 1990). The method favored by Holdaway (1990) alternates iterations where $\boldsymbol{I}'$ and $\boldsymbol{P}'$ are changed independently. In our experience, this method does not offer any practical benefit over imaging $\boldsymbol{I}'$ and $\boldsymbol{P}'$ separately, as the $\boldsymbol{P}'$ reconstruction is not allowed to constrain the $\boldsymbol{I}'$ image. To aid convergence, we tested a strategy of performing initial imaging steps with total intensity and polarization separately and then using the result as the initial image of a joint imaging process. Even in this case, we found the performance of the joint imaging routine to be highly sensitive to the weighting between the bispectrum and polarimetric ratio data terms (α and β in Equation (16)), and the advantage over careful independent imaging of $\boldsymbol{I}'$ and $\boldsymbol{P}'$ seems minimal.

5. IMPLEMENTATION AND RESULTS

In this section, we discuss the implementation of our MEM algorithm and results from applying it to real and simulated data sets. In Section 5.1, we describe the algorithm's implementation in our python software (available at https://github.com/achael/eht-imaging). In Section 5.2, we check our method for consistency with CLEAN on established reconstructions of VLBA observations of the quasar 3C279 at 7 mm and of 3C273 at 3 mm. In Section 5.3, we characterize our algorithm's ability to "superresolve" image structure and compare its performance with CLEAN's in the regime of model images of Sgr A* at 1.3 mm, as might be observed by the EHT. Finally, in Section 5.4, we apply our method to several different simulated EHT data sets from the array that observed Sgr A* and M87 in 2016 March and the expected expanded array in 2017, and interpret the results.

5.1. *Implementation*

We imaged a variety of simulated and real polarimetric VLBI data sets by numerically maximizing the polarimetric ratio objective function, Equation (11), given Stokes I images generated by MEM on the bispectrum obtained by minimizing Equation (4). Because of the relatively small size of our VLBI data sets and the limited fields of view of our reconstructed images, we used DTFTs instead of FFTs in calculating the χ^2 terms and the gradients of J_B and J_m. This eliminates the error introduced by interpolating data from the measured u, v points to the FFT grid. Our MEM routines use the Limited-Memory BFGS algorithm (L-BFGS) (Byrd et al. 1995) implemented in the Scipy scientific python package (Zhu et al. 1997; Jones et al. 2001). L-BFGS is a quasi-Newton gradient descent algorithm that relies on a gradient function provided by the user and estimates the Hessian matrix as it iterates. L-BFGS does not store a full Hessian matrix, but approximates it with a series of vectors from the preceding m steps, making it a preferred choice for our reconstructions due to the large size of the Hessian ($n^2 \times n^2$ for an $n \times n$ image). On our limited data sets, we found that L-BFGS ran sufficiently quickly on images up to 500×500 pixels, and it was consistently more accurate in its reconstructions of model images than either a polarimetric modification of the Cornwell–Evans algorithm or a conjugate gradient method.

A major difference between our MEM implementation and CLEAN is that MEM uses only a forward transform from the sample image to visibility space, while CLEAN uses inverse transforms from the visibilities to the image domain. In CLEAN, the inverse transforms require the visibility data to be gridded, but in MEM the sample visibilities can naturally be computed with a DTFT using the exact sampled u, v points. Our method is also different than some past MEM algorithms (e.g., Cornwell & Evans 1985), which have generally used the forward transform, approximating the data χ^2 as a difference between the dirty image and the test image convolved with the dirty beam. Our method instead uses DTFTs and is ideal for sparse VLBI arrays.

While entropy terms like Equations (6) and (13) can be designed to prefer images that obey physical constraints like $I > 0$ and $0 < m < 1$, the initial steps of an unbounded minimization algorithm often take the image into an unphysical configuration and complex values of the entropy functions. This problem can be addressed by using a bounded minimization algorithm or placing a manual clip on the values of I or m (as in Cornwell & Evans 1985). Instead, we chose to perform a change of variables in both the total intensity and polarimetric images to naturally enforce the image constraints.

In Stokes I, to satisfy the total intensity constraint $I \geqslant 0$, we transformed to $I = e^{\xi}$, where $-\infty < \xi < \infty$. For the polarimetric data, we choose to reconstruct images in m and χ, instead of Q and U, as both the physical constraint $m < 1$ and the entropy functions Equations (13) and (15) are most naturally defined in terms of these variables. To naturally satisfy the constraint $0 < m < 1$, we transformed to $m = \frac{1}{2} + \frac{1}{\pi}\arctan\kappa$, where $-\infty < \kappa < \infty$. In both cases, we modified the gradient given to the algorithm by multiplying by the derivatives of these expressions (see Appendix D).

To compare polarimetric regularizers, we used both the standard entropy term, Equation (13), which we refer to as the Ponsonby–Nityananda–Narayan (PNN) entropy, and a total variation entropy term, Equation (15). In all of our reconstructions, we first imaged $\boldsymbol{I}'$ directly, using the bispectrum. Because bispectral imaging of $\boldsymbol{I}'$ can converge poorly given a poor choice of image prior, we used a sequence of five runs of the

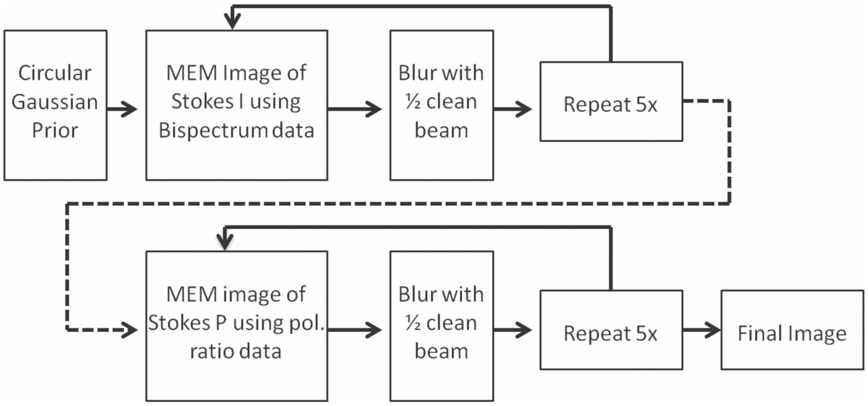

Figure 1. Flowchart summarizing our imaging procedure.

algorithm, substituting the prior image in Equation (6) with the final image from a previous run blurred with a $1/2$ scaled clean beam. To initialize the P imaging process, we use an initial image that has constant fractional polarization magnitude and direction, set equal to the zero-baseline value, which is multiplied by the final $\boldsymbol{I}'$ image. Our tests have shown, however, that the final polarimetric image is generally insensitive to the initial polarimetric image. If zero-baseline polarization data are not available, an image with a constant 5% polarization fraction and zero polarization position angle may be used instead. We again used a sequence of five runs of imaging $\boldsymbol{P}'$, using the final image blurred to $1/2$ the array resolution as the initial image of each subsequent run. Finally, we again convolved the final $\boldsymbol{I}'$ and $\boldsymbol{P}'$ images with a Gaussian beam of $1/2$ the size of the fitted clean beam, limiting MEM's tendency to superresolve spurious features. Note that since the bispectrum does not contain absolute phase information constraining the location of the total flux image centroid, MEM images produced with the bispectrum are frequently offset from the model image, despite attempts to constrain this tendency with Equation (10). As a result, we have manually centered our images to provide clear comparisons with the model images. The essential steps of our procedure are summarized in Figure 1.

Our python code with routines for simulating and manipulating data and producing MEM images with the Scipy L-BFGS algorithm is available to download at https://github.com/achael/eht-imaging.

5.2. 7 and 3 mm VLBA Quasar Observations

To check consistency with CLEAN on real data, we produced polarimetric images from 7 mm quasar observations from the Very Long Baseline Array obtained by the Boston University Blazar Research Group in 2013 (data reduction is described in Jorstad et al. 2005).[6] We compared our polarimetric images with established CLEAN reconstructions that were convolved with the fitted clean beam. While the VLBA data were phase-calibrated, we still used MEM algorithms that only used bispectrum and polarimetric ratio data. Our results for the quasar 3C279 are displayed in Figure 2. We found that when our MEM reconstructions were convolved with the same beam used by CLEAN, the MEM images were an excellent match for the overall polarization magnitude and direction structure of the established images. Both the PNN (Equation (13)) and total variation regularizers (Equation (15)) performed well in reconstructing the direction of the polarization field, and were consistent with each other. In general, the total variation regularizer, which does not prefer low polarization magnitudes, produced higher fractional polarization than the PNN regularizer in areas with weak Stokes I flux.

To further test our method at higher frequencies, we produced a polarimetric image from the 3 mm observation of 3C273 taken with the VLBA in conjunction with the Green Bank Telescope (GBT) reported in Hada et al. (2016). The results are displayed in Figure 3. When convolved with the clean beam, the results from our MEM method are broadly consistent, but some discrepancies are apparent in regions of low polarized intensity.

5.3. "Superresolution" and Comparisons with CLEAN on Simulated 1.3 mm Data

Unlike CLEAN images, MEM images in theory do not require restoration with the fitted interferometer beam. However, when testing our method on synthetic data, we found that at a certain point in the imaging process both total intensity and polarimetric MEM algorithms began producing spurious high-frequency features that were not present in the true source distribution. Restoring the final MEM images by convolving with a Gaussian beam will offset this tendency, but it is important not to make the beam too large and wash out real high-spatial-frequency features that may be "superresolved" on scales smaller than the interferometer beam.

To test MEM's capacity for "superresolution" and determine the appropriate restoring beam size, we produced total intensity MEM (using the entropy term in Equation (6)) and CLEAN images of a model of Sgr A* using simulated data with thermal noise from the EHT array projected to be available in 2017 (See Section 5.4). For this simple test we neglected the effects of inaccurate amplitude calibration, atmospheric phase corruption, and interstellar scattering. We used a MEM algorithm with full visibility phase information, directly minimizing Equation (4) with the χ^2 term in Equation (5). This choice, while infeasible in practice due to phase errors, allowed us to directly compare to CLEAN without introducing the need for self-calibration.

After obtaining MEM and CLEAN reconstructions from the same data, we convolved the reconstructed images with a

---

[6] http://www.bu.edu/blazars/VLBAproject.html

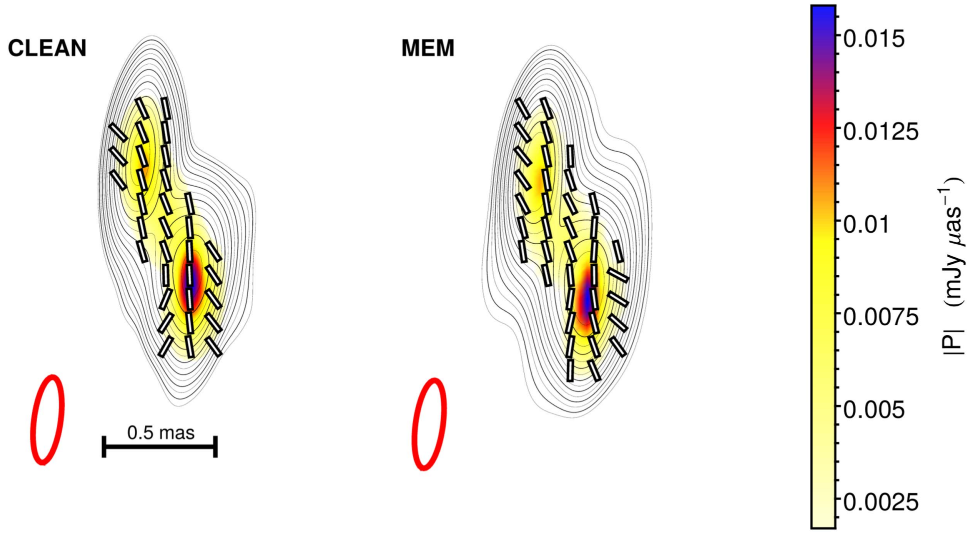

Figure 2. Reconstructions of 7 mm observations of the quasar 3C279 taken with the VLBA in 2013 April (Jorstad et al. 2005). Contours are of total flux in steps of $\sqrt{2}$ from 3× the background rms level. Ticks representing the direction of the polarization position angle and color corresponding to the polarized intensity $|P|$ are plotted in regions where $|P|$ is greater than 4× its background rms level. The left panel shows the reconstruction convolved with the fitted elliptical clean beam (384 × 119 μas FWHM, produced with Briggs weighting), and the right panel displays a MEM reconstruction of the same data set, smoothed with the fitted clean beam. While the CLEAN reconstruction used a self-calibration loop to determine visibility phases on $\tilde{I}$, $\tilde{Q}$, and $\tilde{U}$, the MEM reconstruction directly used bispectrum and polarimetric ratio data. The MEM reconstruction used the Ponsonby–Nityananda–Narayan (PNN) entropy term. The results are consistent with the CLEAN reconstruction when convolved with the same beam.

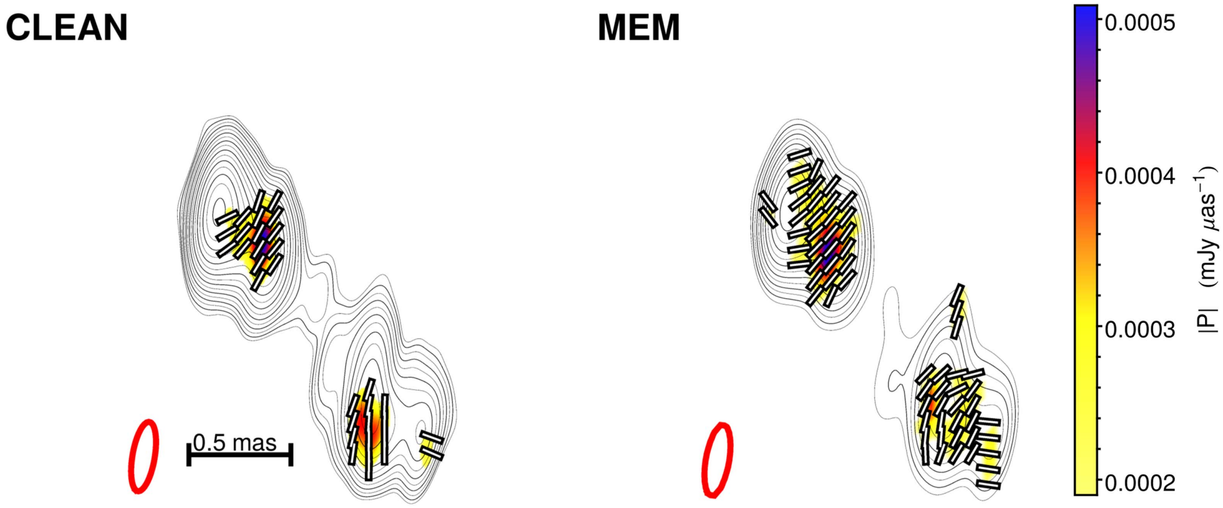

Figure 3. Reconstructions of 3 mm observations of the quasar 3C273 taken with the VLBA+GBT (Hada et al. 2016). Contours are of total flux in steps of $\sqrt{2}$ from 3× the background rms level. Ticks representing the direction of the polarization position angle and color corresponding to the polarized intensity $|P|$ are plotted in regions where $|P|$ is greater than 4× its background rms level. The left panel shows the reconstruction convolved with the fitted elliptical clean beam (340 × 108 μas FWHM, produced with natural weighting), and the right panel displays a MEM reconstruction of the same data set, smoothed with the fitted clean beam. While the CLEAN reconstruction used a self-calibration loop to determine visibility phases on $\tilde{I}$, $\tilde{Q}$, and $\tilde{U}$, the MEM reconstruction directly used bispectrum and polarimetric ratio data. The MEM reconstruction used the Ponsonby–Nityananda–Narayan (PNN) entropy term.

sequence of Gaussian beams scaled from the elliptical Gaussian fitted to the Fourier transform of the u, v coverage (the "clean beam"). We then computed the normalized root-mean-square error (NRMSE) of each restored image:

$$\mathrm{NRMSE} = \sqrt{\frac{\sum_{i=1}^{n^2} |I_i' - I_i|^2}{\sum_{i=1}^{n^2} |I_i|^2}}, \tag{17}$$

where $\boldsymbol{I}'$ is the final restored image and $\boldsymbol{I}$ is the true image. For the CLEAN reconstructions, we chose *not* to add the dirty image residuals back to the convolved model, as the residuals are a sensible quantity only for the full restoring beam. To minimize the effect of this choice on the CLEAN reconstruction, we chose a compact model image with no diffuse structure. After tuning our CLEAN reconstruction for this image, the total flux left in the residuals was less than 2% of the total image flux. In performing the CLEAN reconstruction, we used Briggs weighting and a loop gain of 0.025, with the rest of the parameters set to the default in the algorithm's CASA implementation.[7]

---

[7] http://casa.nrao.edu/docs/TaskRef/clean-task.html

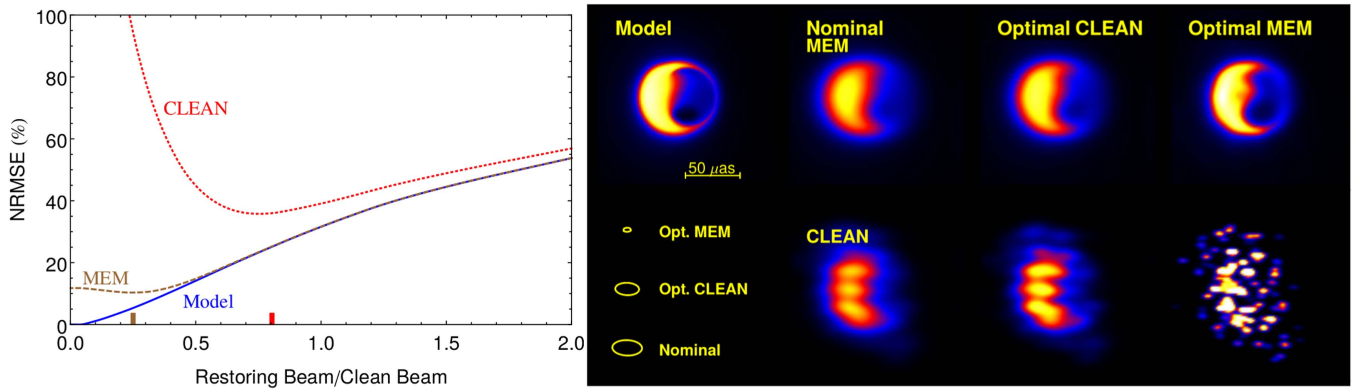

Figure 4. (Left) Normalized root-mean-square error (NRMSE, Equation (17)) of MEM and CLEAN reconstructed Stokes I images as a function of the fractional restoring beam size. For comparison, the NRMSE of the model image is also plotted. The reconstructed images were produced using simulated data from the EHT array; for straightforward comparison with CLEAN, realistic thermal noise was added to the simulated visibilities but gain calibration errors, random atmospheric phases, and blurring due to interstellar scattering were all neglected. The images were convolved with scaled versions of the fitted clean beam. The minimum for each NRMSE curve indicates the optimal restoring beam, which is significantly smaller for MEM (25% of nominal) than for CLEAN (0.78% of nominal). (Right) Example reconstructions restored with scaled beams from curves in the left panel. The center left panels are the MEM and CLEAN reconstructions restored at the nominal resolution, with the fitted clean beam. The center right panels show the reconstructions restored with the optimal beam for the CLEAN reconstruction and the far right panels show both reconstructions restored with the optimal MEM beam. The CLEAN reconstructions consist of only the CLEAN components convolved with the restoring beam and do not include the dirty image residuals, as discussed at the end of Section 3.

The results are displayed in Figure 4. In the left panel, we see that the MEM curve has a minimum in NRMSE at a significantly smaller beam size than the CLEAN reconstruction, demonstrating MEM's superior ability to superresolve source structure over CLEAN. Furthermore, the value of NRMSE from the MEM reconstruction is consistently lower than that from CLEAN for all values of restoring beam size. Most importantly, while the CLEAN curve NRMSE increases rapidly for restoring beams smaller than the optimal resolution, the MEM image fidelity is relatively unaffected by choosing a restoring beam that is too small. Choosing a restoring beam that is too large produces an image with the same fidelity as the model blurred to that resolution. The right panel of Figure 4 shows the model image, the interferometer "clean" beam, and the reconstructions blurred with the clean beam (nominal) and the measured optimal fractional beams. In addition to lower resolution and fidelity, the CLEAN reconstructions show prominent striping features from isolated components being restored with the restoring beam.

While Figure 4 demonstrates that in this case the MEM reconstruction has superior resolution and fidelity to the CLEAN reconstruction, the optimal restoring beam fractional size for the CLEAN reconstruction is still less than unity. This result was observed in several similar reconstructions, suggesting that shrinking the restoring beam used in CLEAN reconstructions to 75% of the nominal fitted beam can enhance resolution without introducing imaging artifacts, at least on images of compact objects similar to those used in these tests.

Repeating the exercise of Figure 4 with observations taken with increased or decreased signal-to-noise ratio resulted in NRMSE curves that are only slightly higher and lower than the curves in Figure 4, but shared the same form—in particular, the location of the minimum NRMSE values was barely shifted. This insensitivity to additional noise is likely due to the overall high S/N of our original observations, which had an average S/N of 178 and a minimum S/N of 13. Our results show that with a high average S/N, increasing or decreasing the noise by up to an order of magnitude does not significantly affect the image reconstruction. Observations with an average S/N ~ 1, on the other hand, may show a drastic change in quality with small adjustments to the noise level.

Extending the exercise from Figure 4, we calculated the NRMSE for polarimetric MEM from several test images as a function of restoring beam size, replacing the Stokes I flux with $P = Q + iU$ in Equation (17). Once again, we neglected the effects of inaccurate amplitude calibration, atmospheric phase corruption, and interstellar scattering in our simulated data; however, our MEM algorithm used only polarimetric ratios $\breve{m}$ while CLEAN reconstructed Q and U separately with full $\tilde{Q}$ and $\tilde{U}$ amplitude and phase information. The results are displayed in Figure 5. Though they are not displayed, reconstructions using different regularizer terms (i.e., Equations (13)–(15)) performed similarly. The degree of superresolution in the polarimetric MEM reconstructions is less than that in the total intensity case, typically with a minimum in NRMSE around a restoring beam size of 1/2 the nominal resolution. This reduced degree of superresolution is likely due to a combination of lower S/N on the polarized data points, the loss of absolute phase information in the MEM imaging process, and the low dynamic range of the $|m|$ images (Holdaway 1990).

The simple tests presented in Figures 4 and 5 suggest that MEM can "superresolve" source structure in I and P on scales greater than about 1/2 the nominal interferometer resolution. Furthermore, they suggest that at least on this class of images, featuring compact flux distributions, MEM achieves superior resolution and fidelity to CLEAN. For the remainder of this work, we adopted a strategy of restoring both total intensity and polarimetric images with a scaled beam 1/2 the size of the fitted beam.

5.4. Imaging Different Models of Sgr A* and M87 at 1.3 mm

We applied our techniques to several 1.3 mm simulated images from supermassive black hole accretion disk and jet models with simulated data from the planned 2017 EHT array. We chose several images featuring different types of structure in total intensity and polarization, including semi-analytic

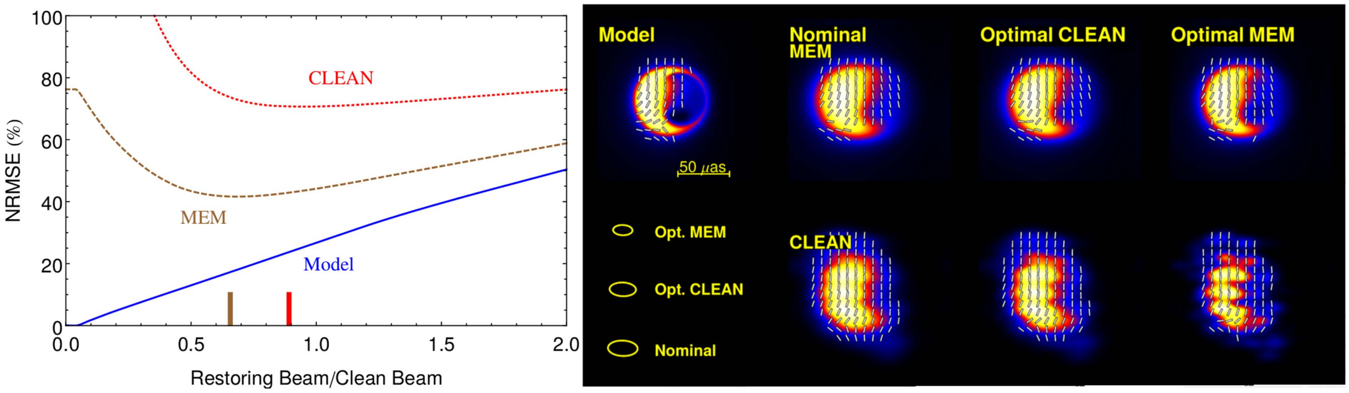

Figure 5. (Left) Normalized root-mean-square error (Equation (17), with $I \rightarrow P$) of MEM and CLEAN reconstructed polarimetric images vs. the size of the anisotropic restoring beam, as a fraction of the nominal fitted beam size. As in Figure 4, the CLEAN curve was computed by restoring the CLEAN point source model with scaled restoring beams without adding the dirty image residuals. The reconstructed images were produced using data simulated from the EHT array with realistic thermal noise; for simplicity of comparison with CLEAN, the data were not corrupted with gain uncertainties, random atmospheric phases, or blurring from interstellar scattering. Comparing to Figure 4, we see that, using this metric, the reconstruction of the linear polarization distribution is less accurate than the reconstructions of Stokes I, but the MEM reconstruction still provides superior resolution and fidelity to CLEAN, with optimal beam sizes at 70% and 95% of the nominal clean beam size, respectively. (Right) Example reconstructions restored with scaled beams from curves in the left panel. The center left panels are the MEM and CLEAN reconstructions restored at the nominal resolution, with the fitted clean beam. The center right panels show the reconstructions restored with the optimal beam for the CLEAN reconstruction and the far right panels show both reconstructions restored with the optimal MEM beam. Polarization position angle ticks are plotted in regions with I greater than 4× its rms value and $|P|$ greater than 2× its rms value.

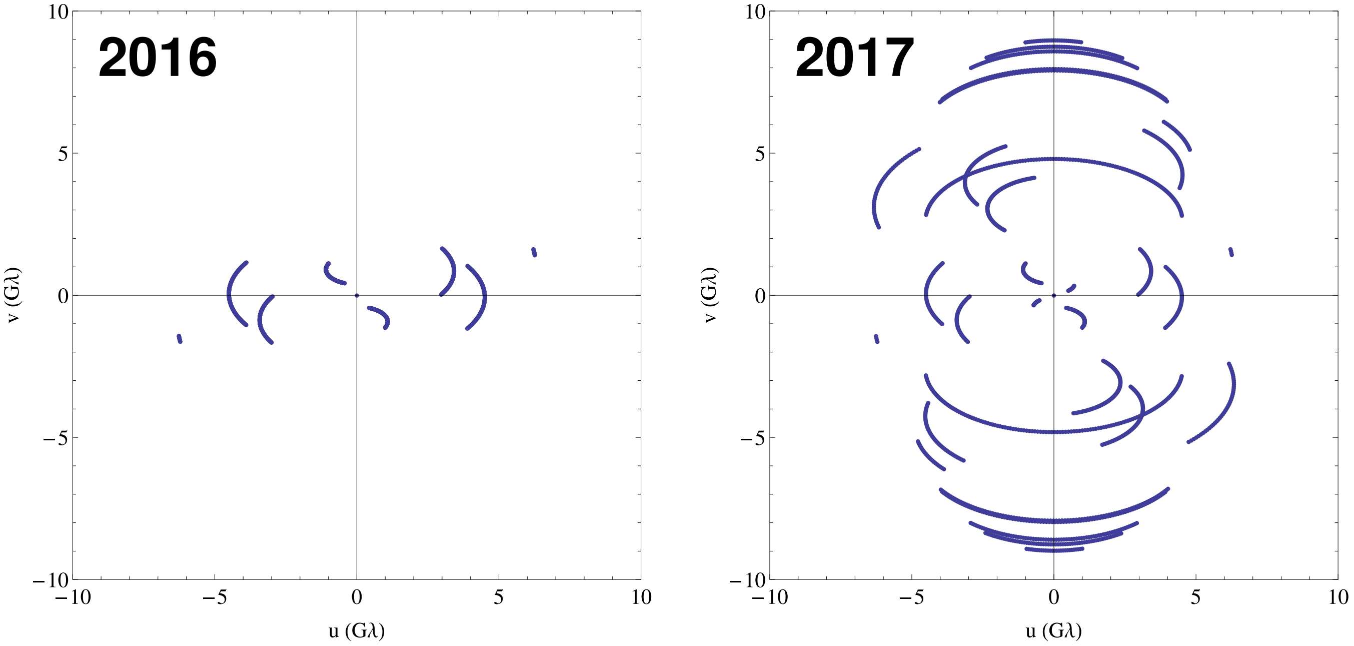

Figure 6. Event Horizon Telescope 24 hr u, v coverage for observations of Sgr A* in 2016 (left) and 2017 (right). The 2016 array includes the Submillimeter Array in Hawaii, the Submillimeter Telescope in Arizona, the Large Millimeter Telescope in Mexico, and the Pico Veleta millimeter dish in Spain. In 2017 the array is projected to expand to include the Plateau de Bure interferometer in France, the ALMA interferometer in Chile, and the South Pole Telescope.

radiatively inefficient accretion flow (RIAF) and jet models courtesy of Avery Broderick (Broderick & Loeb 2009; Broderick et al. 2011; Lu et al. 2014), ray-traced images from a magnetically arrested disk (see, e.g., Tchekhovskoy et al. (2011)) GRMHD simulation courtesy of Jason Dexter (Dexter 2014), and a GRMHD simulation from Roman Gold (Shcherbakov & McKinney 2013; Gold et al. 2016).

We sampled the Fourier transforms of these model images on projected baselines corresponding to the expected EHT arrays in 2016 and 2017. Our 2016 array included stations in Hawaii, Arizona, and Mexico, all operating with 2 GHz of bandwidth. The 2017 array is expected to include these stations with the addition of stations in France, the South Pole, and the ALMA interferometer in Chile, all operating with 4 GHz of bandwidth (see u, v coverage in Figure 6). We added realistic baseline-dependent Gaussian thermal noise on the complex visibilities. The standard deviation σ of the thermal noise is determined according to the standard equation (TMS):

$$\sigma = \frac{1}{0.88}\sqrt{\frac{\mathrm{SEFD}_1 \times \mathrm{SEFD}_2}{2\Delta\,\nu\, t_{\mathrm{int}}}}, \tag{18}$$

where SEFD_1 and SEFD_2 are the telescope system equivalent flux densities, $\Delta\nu$ is the observing bandwidth, and t_{int} is the integration time. The factor of $1/0.88$ in Equation (18) comes from losses due to two-bit quantization in the correlation process. For our simulations, we used an integration time of 60 s, unless otherwise stated.

We simulated the effects of gain calibration errors by assigning each site both a time-dependent gain G_i drawn from a Gaussian distribution with a mean of 1 and 10% standard deviation, and a time-dependent atmospheric opacity τ_i drawn from a Gaussian with a mean 0.1 and a standard deviation of 0.01. The "true" time-dependent SEFDs were then computed from the measured SEFD′s (denoted with primes) from the equation

$$\mathrm{SEFD}_i = \mathrm{SEFD}_i' \frac{e^{\tau_i/\sin\theta_i}}{G_i}, \tag{19}$$

where θ_i is the source elevation at the observation time. Thermal noise was added to the observations from a zero-mean circular complex Gaussian distribution of standard deviation given by Equation (18), with the measured SEFD′ replaced with the true SEFD at each time. The noisy visibilities were then multiplied by the ratio of the estimated SEFDs to true SEFDs

$$\tilde{I}_{ij} \to \tilde{I}_{ij} \times \sqrt{\frac{\mathrm{SEFD}_i' e^{0.1/\sin\theta_i}\ \mathrm{SEFD}_j' e^{0.1/\sin\theta_i}}{\mathrm{SEFD}_i\ \mathrm{SEFD}_j}}, \tag{20}$$

where we have used our assumption that the mean opacity at each site is 0.1 to adjust each of the estimated for elevation dependence. The measured EHT station SEFD′s we used were reported in Lu et al. (2014). In computing the expected thermal noise with these SEFD′s by Equation (18) for computing χ^2 terms, we again modified each SEFD′ by the $e^{0.1/\sin\theta}$ factor from elevation dependence, assuming an opacity $\tau = 0.1$. We did not include any correction for possible gain calibration error in our estimated noise terms.

In simulating phase corruption from atmospheric turbulence, we multiplied the visibilities by random phases drawn uniformly at each site and at each time step. For simulated observations of Sgr A* we also included the blurring effects of interstellar scattering, which we mitigated by dividing out the scattering kernel according to the method of Fish et al. (2014). This process has the net effect of increasing the noise level on long baselines. In practice, for Sgr A* refractive interstellar scattering contributes additional epoch-dependent image distortions (Johnson & Gwinn 2015), which we will analyze separately in a future work.

To compare the polarimetric reconstructions with different regularizers and with different arrays to the model image, we computed the NRMSE in Stokes I and P for each image via Equation (17). To compare the fidelity of the polarization position angle reconstruction, we also computed the mean square error of the polarization position angle of the reconstruction weighted by the magnitude of the total flux:

$$\text{Weighted Angular Error} = \sqrt{\frac{\sum(\mathrm{mod}\,(\chi_i' - \chi_i, \pi)^2\ |I_i|}{\sum |I_i|}}. \tag{21}$$

This error metric gives an rms estimate for the angular error in the polarization position angle reconstruction, and hence the magnetic field morphology of the source. It is weighted by the Stokes I flux because the polarization position angle in the reconstructions can swing wildly in regions with negligible polarized flux. This reasoning also led us to display polarization position angle ticks only in pixels with greater than 10% of the maximum Stokes I flux in all of our reconstructed images.

The EHT simulated data has a lower signal-to-noise than the VLBA data considered above, but we found that our MEM reconstructions were still nearly independent of the choice of regularizer and relative weighting. The polarimetric reconstructions were able to conclusively distinguish between the well-ordered field structure in a RIAF model (Figure 7) and the stochastic field configuration in a GRMHD simulation (Figure 8). Both the PNN and TV entropy terms reproduce the polarization magnitude and direction well, and the NRMSE in both I and P (Equation (17)) and the intensity-weighted polarization position angle error (Equation (21)) were similar for reconstructions with both entropy terms. For the RIAF model, which featured low polarization magnitudes and smoothly varying polarization position angle, the NRMSE values were 28.7% for Stokes I and 47.9% for Stokes P for the PNN reconstruction and 28.53% in Stokes I and 48.2% for P for the TV reconstruction. The corresponding weighted angular errors were 14°.9 and 14°.7. When we compare the reconstructions to the model image smoothed to the same resolution as the reconstruction's resolution (center left in Figure 7), the I and P NRMSE values drop to 25.1% and 46.7% for the PNN reconstruction and 24.8% and 47.0% for the TV reconstruction. The polarization position angle weighted error drops to 13°.7 and 13°.5 for the PNN and TV regularizers, respectively.

For the disordered field in the GRMHD simulation (Figure 8), the NRMSE and weighted angular error fidelity metrics again give similar results for both reconstructions, but slightly favor the PNN image. For PNN, the NRMSE values were 30.4% for Stokes I and 74.23% for Stokes P, with a weighted angular error of 34°.1. For TV, the NRMSE values were 32.0% for Stokes I and 76.0% for P, with a weighted angular error of 34°.3. These high angular error values occur because of the mismatch of the smoothed-out polarization field of the reconstruction and the fine-scale structure in the model image. When compared to the smoothed model image at center left in Figure 8, the Stokes I and P NRMSE drop to 22.6% and 39.9% for the PNN reconstruction and 24.8% and 46.0% for the TV reconstruction; the polarization position angle weighted errors drop to 17°.2 for the PNN and 18°.5 for TV.

We also compared images produced with the PNN regularizer using simulated data from the full 2017 array and the smaller four-element array that observed in 2016 of March (Figure 6, left panel). As expected, the fidelity metrics show distinct improvement between the 2016 and 2017 reconstructions. Both in simulations of the near-horizon jet emission in M87 and accretion disk emission in Sgr A* (Figure 9), we found that the larger amount of information in polarimetric VLBI data over total intensity visibilities (due to the ability to accurately calibrate the phases of polarimetric ratios) was significant with the sparse baseline coverage in 2016. Namely, we were able to achieve more detail in the polarized emission reconstruction than its total intensity counterpart in both cases. In the absence of long baselines needed to resolve distinguishing features in the total intensity image, polarimetric imaging

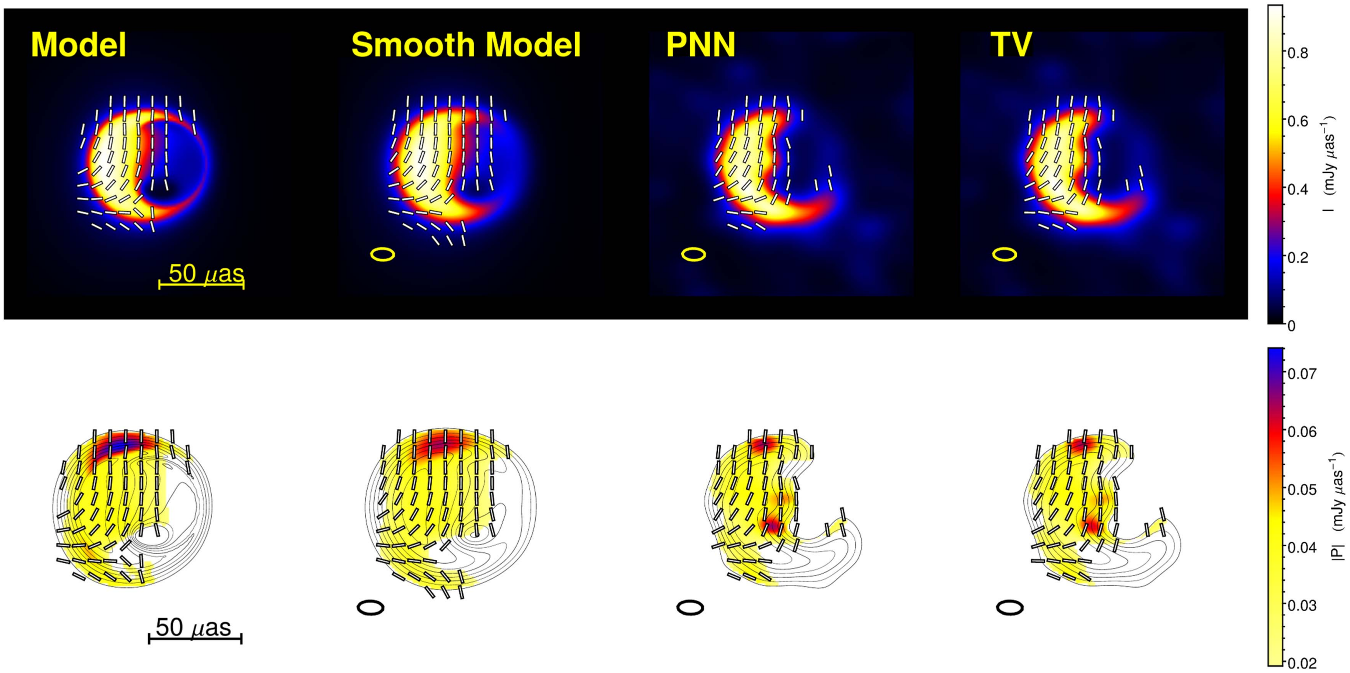

Figure 7. (Top) 1.3 mm MEM reconstructions of a Sgr A* image (left) from a simulation courtesy of Avery Broderick (Broderick et al. 2011). Color indicates Stokes I flux, and ticks marking the polarization position angle are plotted in regions with I greater than 4× its rms value and $|P|$ greater than 2× its rms value. Visibilities from the planned full EHT array were simulated, including the blurring effects of interstellar scattering, with realistic thermal noise, amplitude calibration errors, and random atmospheric phases included. Stokes I was imaged with the bispectrum and linear polarization was subsequently imaged using polarimetric ratios with the Ponsonby–Narayan–Nityananda (PNN) entropy function (right center) and a total variation (TV) regularizer (right). The final reconstructions were restored with a Gaussian beam of $1/2$ the size of the fitted clean beam (27×14 μas FWHM); for comparison, the model image smoothed to this resolution is displayed on the center left. (Bottom) The same reconstructions displayed in contours of total intensity, in steps of $\sqrt{2}$ up from 4× the background rms level. Color indicates the magnitude of the polarized flux $|P|$ and is displayed, along with ticks marking the polarization position angle, in regions where I is greater than 4× its background rms level and $|P|$ is greater than 2× its rms value. Both MEM priors successfully reproduce the smooth polarization morphology of the simulated image.

can help distinguish between different models of the emission region such as the disk and jet models in Figure 9. Polarimetric images, even with poor resolution, can begin to characterize the general magnetic field structure in Sgr A* and M87 with near-term EHT observations even before completely resolving the emission region or black hole shadow.

For our test Sgr A* model, a magnetically arrested disk GRMHD simulation (Figure 9, top panel), the NRMSE of the reconstructions shows distinct improvement between 2016 and 2017, but the weighted angular error (Equation (21)) metric is surprisingly similar across the reconstructions. For 2016, the NRMSE values were 52.30% for Stokes I and 77.3% for Stokes P, with a weighted angular error of 29.°3. In 2017, the NRMSE values were 36.06% for Stokes I and 66.9% for P, with an angular error of 28.°3. While the 2017 array long, high-sensitivity baselines to the ALMA array produce a qualitatively and quantitatively superior reconstruction, MEM techniques reproduce qualitative features of the polarization structure even with the sparse 2016 data.

When we instead compare the reconstructions to the model image smoothed to the same resolution as the respective restoring beam, the I and P NRMSE values drop to 24.0% and 59.0% for the 2016 reconstruction and 19.8% and 61.9% for the 2017 image. The polarization position angle weighted error drops to 20.°0 and 21.°6 for the 2016 and 2017 images, respectively. Even with minimal baseline coverage, MEM is able to reconstruct a reasonably accurate image when compared to the true image viewed at the same resolution.

The 2016 image of an M87 jet model (Figure 9, bottom panel) gave NRMSE values of 55.61% for Stokes I and 77.34% for Stokes P, with a weighted angular error of 23.°5. In 2017, the NRMSE values were 36.71% for Stokes I and 54.40% for P, with an angular error of 17.°9. When we instead compare the reconstructions to the model image smoothed to the same resolution as the restoring beam, the I and P NRMSE values drop to 21.3% and 34.5% for the 2016 image and 18.3% and 27.7% for the 2017 image, while the polarization position angle weighted error drops to 21.°6 and 14.°8 for the 2016 and 2017 images, respectively.

6. CONCLUSION

As the EHT opens up new, extreme environments to direct VLBI imaging, a renewed exploration of VLBI imaging strategies is necessary for extracting physical signatures from challenging data sets. In this paper, we have shown the effectiveness of imaging linear polarization from VLBI data using extensions of the MEM. We explored extensions of MEM using previously proposed polarimetric regularizers like PNN and adaptations of regularizers new to VLBI imaging like total variation. We furthermore adapted standard MEM to operate on robust bispectrum and polarimetric ratio measurements instead of calibrated visibilities. MEM imaging of polarization can provide increased resolution over CLEAN (Figure 5) and is more adapted to continuous distributions, as expected for the black hole accretion disks and jets targeted by the EHT. Furthermore, MEM imaging algorithms can naturally

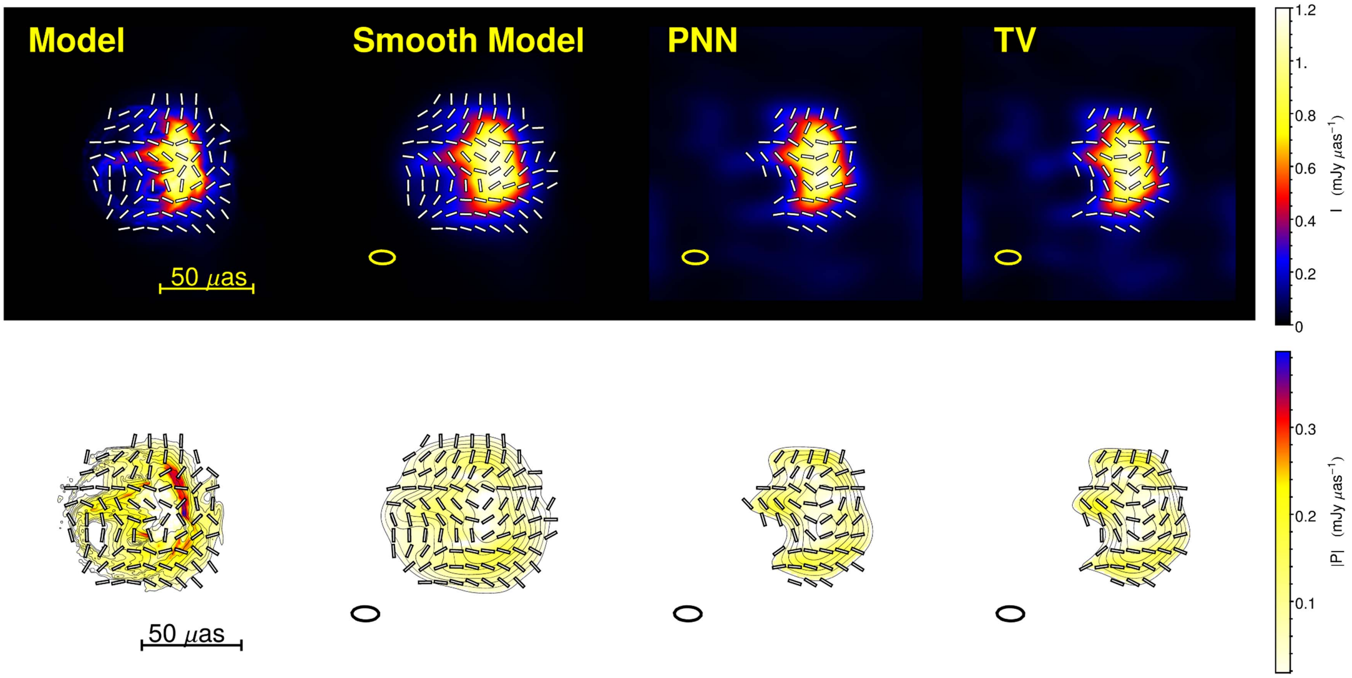

Figure 8. (Top) 1.3 mm MEM reconstructions of a ray-traced image computed from a GRMHD simulation of Sgr A* (left), provided courtesy of Roman Gold (Gold et al. 2016). Color indicates Stokes *I* flux. Ticks marking the direction of linear polarization are displayed in regions with *I* greater than 4× its rms value and $|P|$ greater than 2× its rms value. Visibilities from the planned full EHT array were simulated, including the blurring effects of interstellar scattering, with realistic thermal noise, amplitude calibration errors, and random atmospheric phases included. Stokes *I* was imaged with the bispectrum and linear polarization was subsequently imaged using polarimetric ratios with the Ponsonby–Narayan–Nityananda (PNN) entropy function (right center) and a total variation (TV) regularizer (right). The final reconstructions were restored with a Gaussian beam of 1/2 the size of the fitted clean beam (27 × 14 μas FWHM); for comparison, the model image smoothed to this resolution is displayed on the center left. (Bottom) The same reconstructions displayed in contours of total intensity, in steps of $\sqrt{2}$ up from the 4× background rms level. Color indicating the magnitude of the polarized flux, $|P|$, is displayed along with polarization position angle ticks in regions with *I* greater than 4× its rms value and $|P|$ greater than 2× its rms value. The reconstructions more accurately reproduce the direction of linear polarization than the fractional polarization, as fractional polarization in the reconstructions tends to become large in regions of low total flux. Nonetheless, both reconstructions recover an accurate picture of global structure of the model polarized flux distribution blurred to the EHT's resolution.

incorporate both physical constraints on flux and polarization fraction as well as constraints from prior information or expected source structure. Extending our code to run on data from connected element interferometers like ALMA is a logical next step, but it will require new methods to efficiently handle large amounts of data and image pixels across a wide field of view. Polarimetric MEM is also a promising tool for synthesis imaging of a diversity of other astrophysical systems typically observed with connected element interferometers. For example, the polarized dust emission from protostellar cores frequently exhibits a smooth morphology (Girart et al. 2006; Hull et al. 2013), so MEM may be better-suited to studying both the large-scale magnetic field morphologies and their small deviations, rather than studying typical reconstructions using CLEAN.

The natural ability to incorporate various image constraints makes extensions of MEM useful for investigating new forms of image reconstruction that will be relevant for future EHT observations. Although our algorithm is relatively insensitive to calibration errors and we have shown that our reconstructions are reliable even after including realistic station gain uncertainties and fluctuations, we have not yet incorporated amplitude self-calibration that could further improve reconstructions of $\tilde{I}$. Future work will also investigate a Stokes *I* MEM imaging algorithm that uses only closure phase and closure amplitude data, which would be immune to phase and amplitude calibration errors, thereby eliminating the need for self-calibration. Another goal is the addition of dynamic deblurring that can disentangle effects of strong interstellar scattering with more complicated structure than the simple convolution that holds in the long-term average regime (Johnson & Gwinn 2015). With polarization, MEM could be used to image Faraday rotation across a frequency band. As an imaging framework, MEM is highly flexible and we anticipate that continued investigation will lead to new algorithms that can be tailored to the particular challenges of EHT image reconstruction.

We thank the National Science Foundation (AST-1310896, AST-1312034, AST-1211539, and AST-1440254) and the Gordon and Betty Moore Foundation (#GBMF-3561) for financial support of this work. R.N.'s research was supported in part by NSF grant AST1312651 and NASA grant TCAN NNX14AB47G. K.B. was supported by NSF CGV-1111415 and a NSF Graduate Fellowship. We thank Svetlana Jorstad, Alan Marscher, and Kazuhiro Hada for providing the data imaged in Section 5.2 and for their helpful comments. We thank Avery Broderick, Jason Dexter, and Roman Gold for generously providing model images. We also thank Lindy Blackburn for his help on simulating gain and phase errors and Kazunori Akiyama for his suggestion of applying total variation as a polarimetric regularizer. We thank the anonymous referee, whose thorough suggestions significantly improved this paper. This study makes use of 43 GHz VLBA

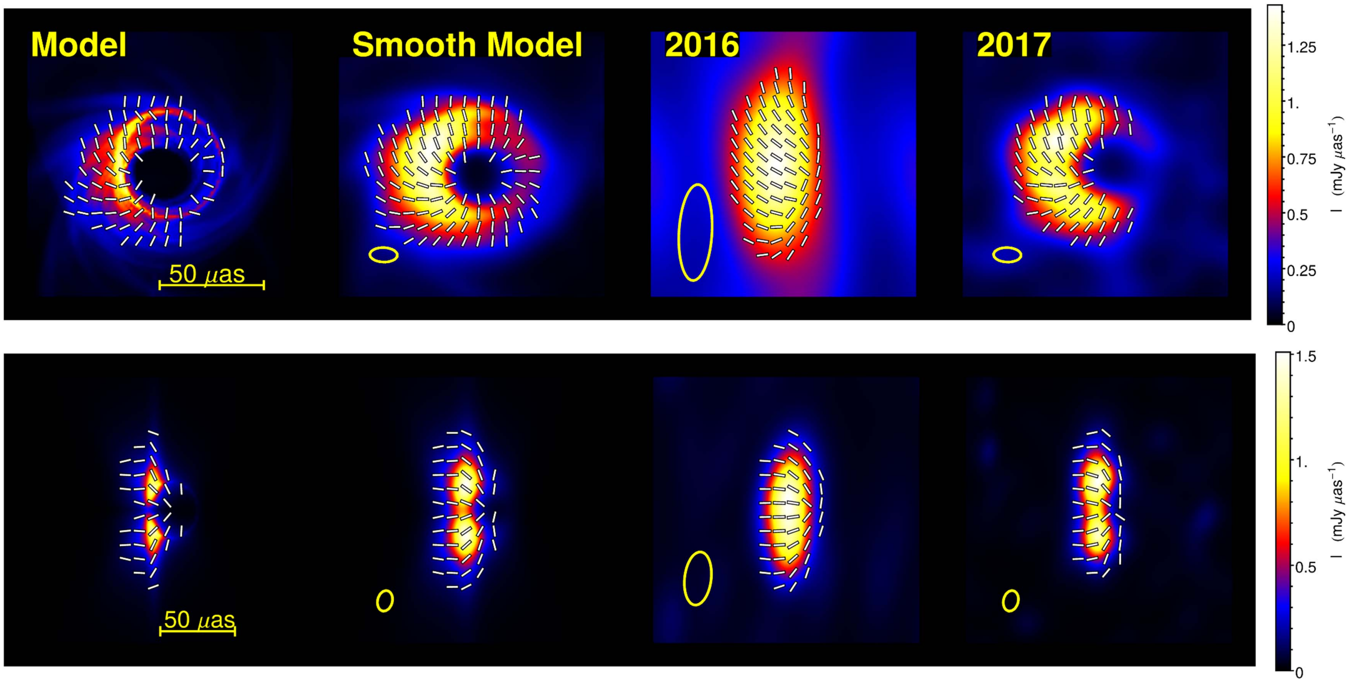

Figure 9. (Top) 1.3 mm MEM reconstructions of a magnetically arrested disk simulation of the Sgr A* accretion flow, courtesy of Jason Dexter (Dexter 2014). Color indicates Stokes I flux and ticks marking the direction of linear polarization are plotted in regions with I greater than 4× its rms value and $|P|$ greater than 2× its rms value. After blurring the image with the Sgr A* scattering kernel at 1.3 mm, data were simulated with realistic thermal noise, amplitude calibration errors, and random atmospheric phases. The center right panel shows a reconstruction with data simulated on EHT baselines expected in 2016 and the rightmost panel shows the reconstruction with the full array expected in 2017. Each reconstruction was restored with a Gaussian beam of 1/2 the size of the fitted clean beam (93 × 32 μas FWHM in 2016; 27 × 14 μas FWHM in 2017). For comparison, the center left panel shows the model smoothed to the same resolution as the 2017 image. (Bottom) 1.3 mm MEM reconstructions of a simulation of the jet in M87, courtesy of Avery Broderick (Broderick & Loeb 2009; Lu et al. 2014). Data were simulated on 2016 and 2017 EHT baselines as done in the top panel, but without the contributions from interstellar scattering that are significant for Sgr A*. Both reconstructions were restored with a Gaussian beam of 1/2 the size of the fitted clean beam (72 × 36 μas FWHM in 2016; 28 × 20 μas FWHM in 2017).

data from the VLBA-BU Blazar Monitoring Program (VLBA-BU-BLAZAR; http://www.bu.edu/blazars/VLBAproject.html), funded by NASA through the Fermi Guest Investigator Program. The VLBA is an instrument of the National Radio Astronomy Observatory. The National Radio Astronomy Observatory is a facility of the National Science Foundation that is operated by Associated Universities, Inc.

APPENDIX A
POLARIMETRIC VLBI OBSERVABLES

In practice, visibilities are estimated by correlating the measured electric fields at different sites. In VLBI, circular feeds are the most common, and the total intensity visibility $\tilde{I}$ is then given as the average of the parallel-hand correlations while $\tilde{Q}_k$ and $\tilde{U}_k$ are estimated using combinations of the cross-hand visibilities. In terms of the cross-hand correlations at sites 1 and 2, the four interferometric Stokes parameters measured on the 1–2 baseline are (Roberts et al. 1994)

$$\tilde{I}_{12} = \frac{1}{2}(\langle R_1 R_2^* \rangle + \langle L_1 L_2^* \rangle) \tag{22}$$

$$\tilde{Q}_{12} = \frac{1}{2}(\langle L_1 R_2^* \rangle + \langle R_1 L_2^* \rangle) \tag{23}$$

$$\tilde{U}_{12} = \frac{i}{2}(\langle L_1 R_2^* \rangle - \langle R_1 L_2^* \rangle) \tag{24}$$

$$\tilde{V}_{12} = \frac{1}{2}(\langle R_1 R_2^* \rangle - \langle L_1 L_2^* \rangle). \tag{25}$$

We ignore circular polarization in what follows. Via the van Cittert–Zernike theorem (Equation (1)), the complex visibilities $\tilde{I}(u, v)$, $\tilde{Q}(u, v)$, $\tilde{U}(u, v)$ are the Fourier transforms of the separate Stokes images $I(x, y)$, $Q(x, y)$, $U(x, y)$. The image linear polarization can also be represented with the fractional polarization m and polarization position angle χ (conventionally measured east of north), where

$$m(x, y) = \frac{\sqrt{Q(x, y)^2 + U(x, y)^2}}{I(x, y)},$$
$$\chi(x, y) = \frac{1}{2}\arctan\frac{U(x, y)}{Q(x, y)}. \tag{26}$$

The distinction between the polarization position angle χ and the data term χ^2 should be clear from the context. Similarly, we can decompose the Fourier conjugate $\tilde{P}$ (Equation (3)):

$$\tilde{P}(u, v) = \tilde{Q}(u, v) + i\tilde{U}(u, v) = \tilde{I}(u, v)\breve{m}(u, v). \tag{27}$$

Again, note that the complex quantity $\breve{m}(u, v)$ is not the Fourier conjugate of the real position-space fractional polarization $m(x, y)$.

Since I, Q, U are real, $\tilde{I}, \tilde{Q}, \tilde{U}$ are conjugate-symmetric under $(u, v) \rightarrow (-u, -v)$. This is not the case for

$\tilde{P} = \tilde{Q} + i\tilde{U}$. Instead, using Equation (22), for telescopes 1, 2 corresponding to a baseline vector (u, v), we see that

$$\tilde{P}(u, v) = \langle R_1 L_2^* \rangle \tag{28}$$

$$\tilde{P}(-u, -v) = \langle L_1^* R_2 \rangle. \tag{29}$$

In an imaging algorithm, we model the $n \times n$ total intensity and polarization images with length n^2 arrays $\boldsymbol{I}'$, $\boldsymbol{P}'$. The N measured total intensity and polarimetric visibilities form arrays $\tilde{\boldsymbol{I}}$, $\tilde{\boldsymbol{P}}$. When comparing our measurements to a test image, we use the arrays of the sample visibilities $\tilde{\boldsymbol{I}}' = \boldsymbol{A}\boldsymbol{I}'$ and $\tilde{\boldsymbol{P}}' = \boldsymbol{A}\boldsymbol{P}'$, where $\boldsymbol{A}$ is a Fourier matrix

$$A_{ij} = e^{-2\pi i (u_i x_j + v_i y_j)}. \tag{30}$$

APPENDIX B
THERMAL NOISE

Thermal noise on a VLBI baseline produces circular Gaussian error in the visibility plane with standard deviation σ given by Equation (18). In principle, the thermal noise σ is the same for $\tilde{I}$, $\tilde{Q}$, and $\tilde{U}$. The factor of $1/0.88$ comes from losses due to 2 bit quantization (TMS). The error in $\tilde{P}$ is also circular, with standard deviation

$$\sigma_P = \sqrt{2}\sigma. \tag{31}$$

Since the error is assumed to be circular in the high S/N limit, the error in the visibility amplitude $|\tilde{I}|$ is equal to the error in the real and imaginary parts, and the error in the visibility phase ϕ is

$$\delta|\tilde{I}| = \sigma,\ \delta\phi = \frac{\sigma}{|\tilde{I}|}. \tag{32}$$

The thermal noise on the bispectrum (Equation (2)) will in general not be described by a circular Gaussian distribution. However, in the limit of high S/N, we can approximate the distribution as a circular Gaussian with standard deviation (TMS)

$$\sigma_B = \delta|\tilde{I}_B| = |\tilde{I}_B| \sqrt{\frac{\sigma_1^2}{|\tilde{I}_1|^2} + \frac{\sigma_2^2}{|\tilde{I}_2|^2} + \frac{\sigma_3^2}{|\tilde{I}_3|^2}}, \tag{33}$$

where $|\tilde{I}_1|$, $|\tilde{I}_2|$, $|\tilde{I}_3|$, are the visibility amplitudes on the three baselines that make up the bispectrum, σ_1, σ_2, and σ_3 are their corresponding standard deviations, and $|\tilde{I}_B| = |\tilde{I}_1||\tilde{I}_2||\tilde{I}_3|$ is the bispectral amplitude. The error in the closure phase is just

$$\sigma_c = \frac{\sigma_B}{|\tilde{I}_B|}. \tag{34}$$

Similarly, the distribution of the visibility domain polarimetric ratio $\tilde{P}/\tilde{I} = \breve{m}$ is not generally a complex circular Gaussian. In the limit of high S/N, however, we again approximate it as such with standard deviation given by

$$\sigma_m = \delta|\breve{m}| = \sigma \sqrt{\frac{2}{|\tilde{I}|^2} + \frac{|\tilde{P}|^2}{|\tilde{I}|^4}}. \tag{35}$$

APPENDIX C
ENTROPY AND χ^2 GRADIENTS

Our implementation of the maximum entropy method uses a gradient descent algorithm to minimize the objective function J. For total intensity imaging, the necessary gradients of χ^2_B with respect to the image pixels I_k can be found in Bouman et al. (2016). Below we list the gradients of the polarimetric χ^2 and entropy terms used in the polarimetric imaging step.

C.1. Polarimetric Ratio χ^2

In imaging P, we work directly with interferometric polarimetric ratios. The reduced $\chi^2_m(\boldsymbol{I}', \boldsymbol{P}')$ that we use is (Equation (12))

$$\chi^2(\boldsymbol{I}', \boldsymbol{P}') = \frac{1}{2N} \sum_i^N \frac{|\tilde{P}_i/\tilde{I}_i - \tilde{P}_i'/\tilde{I}_i'|^2}{\sigma_i^2} = \frac{1}{2N} \sum_i^N \frac{|\breve{m}_i - \breve{m}_i'|^2}{\sigma_i^2}, \tag{36}$$

where errors on the polarimetric ratios are calculated according to Equation (35). Computing the gradient with respect to the image domain fractional polarizations m_k and polarization position angles χ_k gives

$$\frac{\partial}{\partial m_k} \chi^2 = -\frac{1}{N} \mathrm{Re} \left[I_k e^{-2i\chi_k} \sum_i A_{ki}^\dagger \left(\frac{\tilde{P}_i/\tilde{I}_i - \tilde{P}_i'/\tilde{I}_i'}{\tilde{I}_i'^* \, \sigma_i^2} \right) \right] \tag{37}$$

$$\frac{\partial}{\partial \chi_k} \chi^2 = -\frac{2}{N} \mathrm{Im} \left[I_k m_k e^{-2i\chi_k} \sum_i A_{ki}^\dagger \left(\frac{\tilde{P}_i/\tilde{I}_i - \tilde{P}_i'/\tilde{I}_i'}{\tilde{I}_i'^* \, \sigma_i^2} \right) \right]. \tag{38}$$

C.2. Ponsonby/Nityananda/Narayan Entropy

Setting $m_{\max} = 1$ in Equation (13), we have the traditional form of the PNN entropy:

$$S(\boldsymbol{I}', \boldsymbol{P}') = -\sum I_i' \left(\frac{1 + m_i'}{2} \log \frac{1 + m_i'}{2} + \frac{1 - m_i'}{2} \log \frac{1 - m_i'}{2} \right). \tag{39}$$

It has gradients with respect to polarimetric ratio m and polarization position angle χ given by

$$\frac{\partial S}{\partial m_k'} = -I_k' \operatorname{arctanh} m_k' \tag{40}$$

$$\frac{\partial S}{\partial \chi_k'} = 0. \tag{41}$$

C.3. Total Variation Entropy

Because the total variation (Equation (15)) involves differences between pixels in both the x and y image direction, we must adjust our notation to account for the 2D nature of the image. With both dimensions restored, the complex polarized image is $P_{i,j} = I_{i,j}\, m_{i,j}\, e^{2i\chi_{i,j}}$. The total variation entropy term is then

$$S(\boldsymbol{I}', \boldsymbol{P}') = -\sum_{ij} \sqrt{|P_{i+1,j}' - P_{i,j}'|^2 + |P_{i,j+1}' - P_{i,j}'|^2}. \tag{42}$$

The gradients with respect to m and χ are

$$\begin{aligned}\frac{\partial S}{\partial m'_{k,l}} = &-\frac{2|P'_{k,l}| - |P'_{k+1,l}|\cos[2(\chi'_{k+1,l} - \chi'_{k,l})] - |P'_{k,l+1}|\cos[2(\chi'_{k,l+1} - \chi'_{k,l})]}{\sqrt{|P'_{k+1,l} - P'_{k,l}|^2 + |P'_{k,l+1} - P'_{k,l}|^2}}\\ &-\frac{|P'_{k,l}| - |P'_{k-1,l}|\cos[2(\chi'_{k,l} - \chi'_{k-1,l})]}{\sqrt{|P'_{k,l} - P'_{k-1,l}|^2 + |P'_{k-1,l} - P'_{k-1,l+1}|^2}} - \frac{|{}'P_{k,l}| - |P'_{k,l-1}|\cos[2(\chi'_{k,l} - \chi'_{k,l-1})]}{\sqrt{|P'_{k+1,l-1} - P'_{k,l-1}|^2 + |P'_{k,l} - P'_{k,l-1}|^2}}\end{aligned} \tag{43}$$

$$\begin{aligned}\frac{\partial S}{\partial \chi'_{k,l}} = &+\frac{2|P'_{k,l}P'_{k+1,l}|\sin[2(\chi'_{k+1,l} - \chi'_{k,l})] + 2|P'_{k,l}P'_{k,l+1}|\sin[2(\chi'_{k,l+1} - \chi'_{k,l})]}{\sqrt{|P'_{k+1,l} - P'_{k,l}|^2 + |P'_{k,l+1} - P'_{k,l}|^2}}\\ &-\frac{2|P'_{k,l}P'_{k-1,l}|\sin[2(\chi'_{k,l} - \chi'_{k-1,l})]}{\sqrt{|P'_{k,l} - P'_{k-1,l}|^2 + |P'_{k-1,l} - P'_{k-1,l+1}|^2}} - \frac{2|P'_{k,l}P'_{k,l-1}|\sin[2(\chi'_{k,l} - \chi'_{k,l-1})]}{\sqrt{|P'_{k+1,l-1} - P'_{k,l-1}|^2 + |P'_{k,l} - P'_{k,l-1}|^2}}\end{aligned} \tag{44}$$

APPENDIX D
IMAGE CHANGE OF VARIABLES

To extend the range of the image variables for total intensity and polarization to the entire real line and avoid the use of bounded minimization, we use the change of variables

$$I_i = e^{\xi_i}, \quad m_i = \frac{1}{2} + \frac{1}{\pi}\tan^{-1}\kappa_i. \tag{45}$$

Consequently, we need to multiply the gradients given in Appendix C by the factors

$$\frac{dI_i}{d\xi_i} = e^{\xi_i}, \quad \frac{dm_i}{d\kappa_i} = \frac{1}{\pi(1 + \kappa_i^2)}. \tag{46}$$

REFERENCES

Attridge, J. M. 2001, ApJL, 553, L31
Attridge, J. M., Wardle, J. F. C., & Homan, D. C. 2005, ApJL, 633, L85
Balbus, S. A., & Hawley, J. F. 1998, RvMP, 70, 1
Baron, F., & Young, J. S. 2008, Proc. SPIE, 7013, 3
Blandford, R. D., & Payne, D. G. 1982, MNRAS, 199, 883
Blandford, R. D., & Znajek, R. L. 1977, MNRAS, 179, 433
Bouman, K. L., Johnson, M. D., Zoran, D., et al. 2016, in IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 913
Broderick, A. E., Fish, V. L., Doeleman, S. S., & Loeb, A. 2011, ApJ, 735, 110
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 (Dordrecht: Kluwer), 91
Byrd, R. H., Lu, P., & Nocedal, J. 1995, SIAM Journal on Scientific and Statistical Computing, 16, 1190
Cornwell, T. J., & Evans, K. F. 1985, A&A, 143, 77
Coughlan, C., & Gabuzda, D. 2012, JPhCS, 355, 012020
Coughlan, C. P., & Gabuzda, D. C. 2013, in European Physical Journal Web of Conf. 61, 7009
Dexter, J. 2014, in IAU Symp. 303, ed. L. O. Sjouwerman, C. C. Lang, & J. Ott (Cambridge: Cambridge Univ. Press), 298
Doeleman, S., Agol, E., Backer, D., et al. 2009, The Astronomy and Astrophysics Decadal Survey, Science White Papers, 68
Doeleman, S. S., Weintroub, J., Rogers, A. E. E., et al. 2008, Natur, 455, 78
Doeleman, S. S., Fish, V. L., Schenck, D. E., et al. 2012, Sci, 338, 355
Fabian, A. C. 2012, ARA&A, 50, 455
Fish, V. L., Doeleman, S. S., Beaudoin, C., et al. 2011, ApJL, 727, L36
Fish, V. L., Johnson, M. D., Lu, R.-S., et al. 2014, ApJ, 795, 134
Frieden, B. R. 1972, JOSA, 62, 511
Ginzburg, V. L., & Syrovatskii, S. I. 1965, ARA&A, 3, 297
Girart, J. M., Rao, R., & Marrone, D. P. 2006, Sci, 313, 812
Gold, R., McKinney, J. C., Johnson, M. D., & Doeleman, S. S. 2016, arXiv:1601.05550
Gull, S. F., & Daniell, G. J. 1978, Natur, 272, 686
Hada, K., Kino, M., Doi, A., et al. 2016, ApJ, 817, 131
Högbom, J. A. 1974, A&AS, 15, 417
Holdaway, M. A. 1990, PhD thesis, Brandeis Univ. Waltham
Holdaway, M. A., & Wardle, J. F. C. 1990, Proc. SPIE, 1351, 714
Honma, M., Akiyama, K., Uemura, M., & Ikeda, S. 2014, PASJ, 66, 95
Hull, C. L. H., Plambeck, R. L., Bolatto, A. D., et al. 2013, ApJ, 768, 159
Johnson, M. D., Fish, V. L., Doeleman, S. S., et al. 2014, ApJ, 794, 150
Johnson, M. D., & Gwinn, C. R. 2015, ApJ, 805, 180
Johnson, M. D., Fish, V. L., Doeleman, S. S., et al. 2015, Sci, 350, 1242
Jones, E., Oliphant, T., Peterson, P., et al. 2001, SciPy: Open source scientific tools for Python, [Online; accessed 2015-08-25]
Jorstad, S. G., Marscher, A. P., Lister, M. L., et al. 2005, AJ, 130, 1418
Kormendy, J., & Ho, L. C. 2013, ARA&A, 51, 511
Lu, R.-S., Broderick, A. E., Baron, F., et al. 2014, ApJ, 788, 120
Narayan, R., & Nityananda, R. 1986, ARA&A, 24, 127
Nityananda, R., & Narayan, R. 1983, A&A, 118, 194
Ponsonby, J. E. B. 1973, MNRAS, 163, 369
Roberts, D. H., Wardle, J. F. C., & Brown, L. F. 1994, ApJ, 427, 718
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, 256
Rybicki, G. B., & Lightman, A. P. 1979, Radiative Processes in Astrophysics (New York: Wiley-Interscience)
Sault, R. J., Bock, D. C.-J., & Duncan, A. R. 1999, A&AS, 139, 387
Shcherbakov, R. V., & McKinney, J. C. 2013, ApJL, 774, L22
Shevgaonkar, R. K. 1987, A&A, 176, 159
Tchekhovskoy, A., Narayan, R., & McKinney, J. C. 2011, MNRAS, 418, L79
Thompson, A. R., Moran, J. M., & Swenson, G. W. 2007, Interferometry and Synthesis in Radio Astronomy (New York: Wiley)
Wardle, J. F. C. 2013, in European Physical Journal Web of Conf. 61, 6001
Zhu, C., Byrd, R. H., & Nocedal, J. 1997, ACM Transactions on Mathematical Software, 23, 550