

# Computational Imaging for VLBI Image Reconstruction

Katherine L. Bouman<sup>1</sup>

Michael D. Johnson<sup>2</sup>

Daniel Zoran<sup>1</sup>

Vincent L. Fish<sup>3</sup>

Sheperd S. Doeleman<sup>2,3</sup>

William T. Freeman<sup>1,4</sup>

<sup>1</sup>Massachusetts Institute of Technology, CSAIL

<sup>2</sup>Harvard-Smithsonian Center for Astrophysics

<sup>3</sup>MIT Haystack Observatory

<sup>4</sup>Google

## Abstract

*Very long baseline interferometry (VLBI) is a technique for imaging celestial radio emissions by simultaneously observing a source from telescopes distributed across Earth. The challenges in reconstructing images from fine angular resolution VLBI data are immense. The data is extremely sparse and noisy, thus requiring statistical image models such as those designed in the computer vision community. In this paper we present a novel Bayesian approach for VLBI image reconstruction. While other methods often require careful tuning and parameter selection for different types of data, our method (CHIRP) produces good results under different settings such as low SNR or extended emission. The success of our method is demonstrated on realistic synthetic experiments as well as publicly available real data. We present this problem in a way that is accessible to members of the community, and provide a dataset website ([vlbiimaging.csail.mit.edu](http://vlbiimaging.csail.mit.edu)) that facilitates controlled comparisons across algorithms.*

## 1. Introduction

High resolution celestial imaging is essential for progress in astronomy and physics. For example, imaging the plasma surrounding a black hole’s event horizon at high resolution could help answer many important questions; most notably, it may substantiate the existence of black holes [10] as well as verify and test the effects of general relativity [22]. Recently, there has been an international effort to create an Event Horizon Telescope (EHT) capable of imaging a black hole’s event horizon for the first time [12, 13]. The angular resolution necessary for this observation is at least an order of magnitude smaller than has been previously used to image radio sources [24]. As measurements from the EHT become available, robust algorithms able to reconstruct images in this fine angular resolution regime will be necessary.

Although billions of dollars are spent on astronomical imaging systems to acquire the best images, current reconstruction techniques suffer from unsophisticated priors and

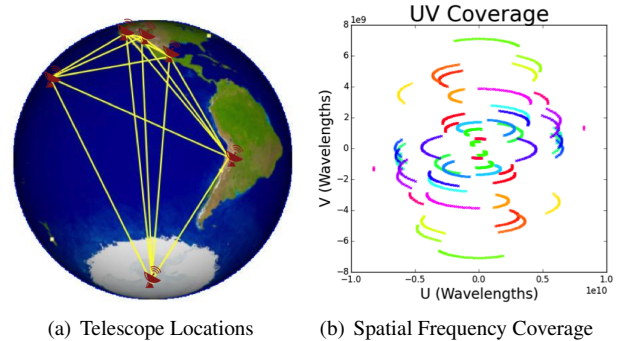


Figure 1. **Frequency Coverage:** (A) A sample of the telescope locations in the EHT. By observing a source over the course of a day, we obtain measurements corresponding to elliptical tracks in the source image’s spatial frequency plane (B). These frequencies,  $(u, v)$ , are the projected baseline lengths orthogonal to a telescope pair’s light of sight. Points of the same color correspond to measurements from the same telescope pair.

a lack of inverse modeling [36], resulting in sub-optimal images. Image processing, restoration, sophisticated inference algorithms, and the study of non-standard cameras are all active areas of computer vision. The computer vision community’s extensive work in these areas are invaluable to the success of these reconstruction methods and can help push the limits of celestial imaging. [16, 17, 27, 43].

Imaging distant celestial sources with high resolving power (i.e. fine angular resolution) requires single-dish telescopes with prohibitively large diameters due to the inverse relationship between angular resolution and telescope diameter [41]. For example, it is predicted that emission surrounding the black hole at the center of the Milky Way subtends  $\approx 2.5 \times 10^{-10}$  radians [15]. Imaging this emission with a  $10^{-10}$  radian resolution at a 1.3 mm wavelength would require a telescope with a 13000 km diameter. Although a single telescope this large is unrealizable, by simultaneously collecting data from an array of telescopes located around the Earth, it is possible to emulate samples from a single telescope with a diameter equal to the maximum distance between telescopes in the array. Using multiple telescopes in this manner is referred to as very long baseline interferometry (VLBI) [41]. Refer to Figure 1a.

VLBI measurements place a sparse set of constraints on the source image’s spatial frequencies. The task of reconstructing an image from these constraints is highly ill-posed and relies heavily on priors to guide optimization. Current VLBI image reconstruction techniques have been reasonably successful in imaging large celestial sources with coarse angular resolution. However, as the demand for higher resolving power increases, particularly for the EHT, traditional reconstruction algorithms are quickly approaching their limits [28, 38].

The difficulty of image reconstruction drastically increases as the angular resolution of a VLBI array improves. To improve angular resolution (i.e., increase resolving power), one must either increase the maximum distance between two telescopes or decrease the observing wavelength [41]. Due to the fixed size of Earth, increasing the maximum telescope baseline results in a smaller set of possible telescope sites to choose from. Therefore, algorithms must be designed to perform well with increasingly fewer measurements [28]. Extending VLBI to millimeter and sub-mm wavelengths to increase resolution requires overcoming many challenges, all of which make image reconstruction more difficult. For instance, at these short wavelengths, rapidly varying inhomogeneities in the atmosphere introduce additional measurement errors [31, 38].

In this paper, we leverage ideas from computer vision to confront these challenges. We present a new algorithm, CHIRP (Continuous High-resolution Image Reconstruction using Patch priors), which takes a Bayesian approach and novel formulation to solve the ill-posed inverse problem of image reconstruction from VLBI measurements. Specifically, the contributions of this algorithm are:

- An improved forward model approximation that more accurately models spatial-frequency measurements,
- A simpler problem formulation and optimization strategy to model the effects of atmospheric noise on VLBI data.

Furthermore, current interferometry testing datasets are small and have noise properties unsuitable for radio wavelengths [3, 6, 26].

- We introduce a large, realistic VLBI dataset website to the community ([vlbiimaging.csail.mit.edu](http://vlbiimaging.csail.mit.edu)).

This website allows researchers to easily access a large VLBI dataset, and compare their algorithms to other leading methods. Its automatic evaluation system facilitates unbiased comparisons between algorithms, which are otherwise difficult to make and are lacking in the literature. Furthermore, we hope to make this computational imaging problem accessible to computer vision researchers, cross-fertilizing the astronomy and computer vision communities.

## 2. A Brief Introduction to VLBI

We briefly describe VLBI to provide the necessary background for building an accurate likelihood model. Our

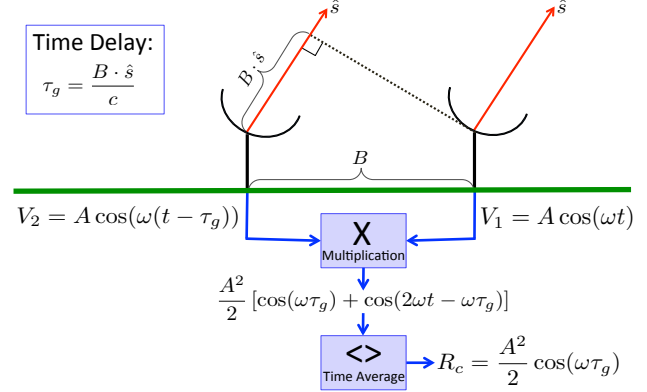


Figure 2. **Simplified Interferometry Diagram:** Light is emitted from a distant source and arrives at the telescopes as a plane wave in the direction  $\hat{s}$ . An additional distance of  $B \cdot \hat{s}$  is necessary for the light to travel to the farther telescope, introducing a time delay between the received signals that varies depending on the source’s location in the sky. The time-averaged correlation of these signals is a sinusoidal function related to the location of the source. This insight is generalized to extended emissions in the van Cittert-Zernike Thm. and used to relate the time-averaged correlation to a Fourier component of the emission image in the direction  $\hat{s}$ .

goal is to provide intuition; for additional details we recommend [41]. As with cameras, a single-dish telescope is diffraction limited. However, simultaneously collecting data from an array of telescopes, called an interferometer, allows us to overcome the single-dish diffraction limit.

Figure 2 provides a simplified explanation of a two telescope interferometer. Electromagnetic radiation travels from a point source to the telescopes. However, because the telescopes are separated by a distance  $B$ , they will not receive the signal concurrently. For spatially incoherent extended emissions, the time-averaged correlation of the received signals is equivalent to the projection of this sinusoidal variation on the emission’s intensity distribution.

This phenomenon is formally described by the *van Cittert-Zernike Theorem*. The theorem states that, for ideal sensors, the time-averaged correlation of the measured signals from two telescopes,  $i$  and  $j$ , for a single wavelength,  $\lambda$ , can be approximated as:

$$\Gamma_{i,j}(u, v) \approx \int_{\ell} \int_m e^{-i2\pi(u\ell + vm)} I_{\lambda}(\ell, m) d\ell dm \quad (1)$$

where  $I_{\lambda}(\ell, m)$  is the emission of wavelength  $\lambda$  traveling from the direction  $\hat{s} = (\ell, m, \sqrt{1 - \ell^2 - m^2})$ . The dimensionless coordinates  $(u, v)$  (measured in wavelengths) are the projected baseline,  $B$ , orthogonal to the line of sight.<sup>1</sup> Notice that Eq. 1 is just the Fourier transform of the source emission image,  $I_{\lambda}(\ell, m)$ . Thus,  $\Gamma_{i,j}(u, v)$  provides a single complex Fourier component of  $I_{\lambda}$  at position  $(u, v)$  on the 2D spatial frequency plane. We refer to these measurements,  $\Gamma_{i,j}$ , as *visibilities*. Since the spatial frequency,

<sup>1</sup>The change in elevation between telescopes can be neglected due to corrections made in pre-processing. Additionally, for small FOVs wide-field effects are negligible.

$(u, v)$ , is proportional to the baseline,  $B$ , increasing the distance between telescopes increases the resolving power of the interferometer, allowing it to distinguish finer details.

**Earth Rotation Synthesis** At a single time, for an  $N$  telescope array, we obtain  $\frac{N(N-1)}{2}$  visibility measurements corresponding to each pair of telescopes. As the Earth rotates, the direction that the telescopes point towards the source ( $\hat{s}$ ) changes. Assuming a static source, this yields measurements of spatial frequency components (visibilities) of the desired image along elliptical paths in the  $(u, v)$  frequency plane (see Fig. 1b).

**Phase Closure** All equations thus far assumed that light travels from the source to a telescope through a vacuum. However, inhomogeneities in the atmosphere cause the light to travel at different velocities towards each telescope. These delays have a significant effect on the phase of measurements, and renders the phase unusable for image reconstructions at wavelengths less than 3 mm [31].

Although absolute phase measurements cannot be used, a clever observation - termed phase closure - allows us to still recover some information from the phases. The atmosphere affects an ideal visibility (spatial frequency measurement) by introducing an additional phase term:  $\Gamma_{i,j}^{\text{meas}} = e^{i(\phi_i - \phi_j)} \Gamma_{i,j}^{\text{ideal}}$ , where  $\phi_i$  and  $\phi_j$  are the phase delays introduced in the path to telescopes  $i$  and  $j$  respectively. By multiplying the visibilities from three different telescopes, we obtain an expression that is invariant to the atmosphere, as the unknown phase offsets cancel, see Eq. 2 [14].

$$\begin{aligned} \Gamma_{i,j}^{\text{meas}} \Gamma_{j,k}^{\text{meas}} \Gamma_{k,i}^{\text{meas}} &= e^{i(\phi_i - \phi_j)} \Gamma_{i,j}^{\text{ideal}} e^{i(\phi_j - \phi_k)} \Gamma_{j,k}^{\text{ideal}} e^{i(\phi_k - \phi_i)} \Gamma_{k,i}^{\text{ideal}} \\ &= \Gamma_{i,j}^{\text{ideal}} \Gamma_{j,k}^{\text{ideal}} \Gamma_{k,i}^{\text{ideal}} \end{aligned} \quad (2)$$

We refer to this triple product of visibilities as the *bispectrum*. The bispectrum is invariant to atmospheric noise; however, in exchange, it reduces the number of constraints that can be used in image reconstruction. Although the number of triple pairs in an  $N$  telescope array is  $\binom{N}{3}$ , the number of independent values is only  $\frac{(N-1)(N-2)}{2}$ . For small telescope arrays, such as the EHT, this effect is large. For instance, in an eight telescope array, using the bispectrum rather than visibilities results in 25% fewer independent constraints [14].

### 3. Related Work

We summarize a few significant algorithms from the astronomical interferometry imaging literature.

**CLEAN** CLEAN is the de-facto standard method used for VLBI image reconstruction. It assumes that the image is made up of a number of bright point sources. From an initialization image, CLEAN iteratively looks for the brightest point in the image and “deconvolves” around that location by removing side lobes that occur due to sparse sampling in

the  $(u, v)$  frequency plane. After many iterations, the final image of point sources is blurred [18]. Since CLEAN assumes a distribution of point sources, it often struggles with reconstructing images of extended emissions [38].

For mm/sub-mm wavelength VLBI, reconstruction is complicated by corruption of the visibility phases. CLEAN is not inherently capable of handling this problem; however, self-calibration methods have been developed to greedily recover the phases during imaging. Self-calibration requires manual input from a knowledgeable user and often fails when the SNR is too low or the source is complex [38].

**Optical Interferometry** Interferometry at visible wavelengths faces the same phase-corruption challenges as mm/sub-mm VLBI. Although historically the optical and radio interferometry communities have been separate, fundamentally the resulting measurements and imaging process are very similar [31]. We have selected two optical interferometry reconstruction algorithms representative of the field to discuss and compare to in this work [28]. Both algorithms take a regularized maximum likelihood approach and can use the bispectrum, rather than visibilities, for reconstruction [4, 11]. Recent methods based on compressed sensing have been proposed, but have yet to demonstrate superior results [25, 28].

BSMEM (BiSpectrum Maximum Entropy Method) takes a Bayesian approach to image reconstruction [11]. Gradient descent optimization [37] using a maximum entropy prior is used to find an optimal reconstruction of the image. Under a flat image prior BSMEM is often able to achieve impressive super-resolution results on simple celestial images. However, in Section 6 we demonstrate how it often struggles on complex, extended emissions.

SQUEEZE takes a Markov chain Monte Carlo (MCMC) approach to sample images from a posterior distribution [4]. To obtain a sample image, SQUEEZE moves a set of point sources around the field of view (FOV). The final image is then calculated as the average of a number of sample images. Contrary to gradient descent methods, SQUEEZE is not limited in its choice of regularizers or constraints [28]. However, this freedom comes at the cost of a large number of parameter choices that may be hard for an unknowledgeable user to select and tune.

#### 3.1. Spectral Image Reconstruction

VLBI image reconstruction has similarities with other spectral image reconstruction problems, such as Synthetic Aperture Radar (SAR), Magnetic Resonance Imaging (MRI), and Computed Tomography (CT) [8, 29, 33, 39]. However, VLBI faces a number of challenges that are typically not relevant in these other fields. For instance, SAR, MRI, and CT are generally not plagued by large corruption of the signal’s phase, as is the case due to atmospheric differences in mm/sub-mm VLBI. In SAR the Fourier samples

are all coherently related and the absolute phase can generally be recovered, even under atmospheric changes [19, 32]. However, although fully understanding the connection remains an open problem, incorporating ideas of phase closure, as is done in this work, may open the potential to push SAR techniques [21] past their current limits.

## 4. Method

Reconstructing an image using bispectrum measurements is an ill-posed problem, and as such there are an infinite number of possible images that explain the data [28]. The challenge is to find an explanation that respects our prior assumptions about the “visual” universe while still satisfying the observed data.

### 4.1. Continuous Image Representation

The image that we wish to recover,  $I_\lambda(\ell, m)$ , is defined over the continuous space of angular coordinates  $\ell$  and  $m$ . Many algorithms assume a discretized image of point sources during reconstruction [38]. This discretization introduces errors during the reconstruction optimization, especially in fitting the higher frequency visibilities. Instead, we parameterize a continuous image using a discrete number of terms. *This parameterization not only allows us to model our emission with a continuous image, but it also reduces modeling errors during optimization.*

Since each measured complex visibility is approximated as the Fourier transform of  $I_\lambda(\ell, m)$ , a convenient parameterization of the image is to represent it as a discrete number of scaled and shifted continuous pulse functions, such as triangular pulses. For a scene defined in the range  $\ell \in [-\frac{F_\ell}{2}, \frac{F_\ell}{2}]$  and  $m \in [-\frac{F_m}{2}, \frac{F_m}{2}]$ , we parameterize our

space into  $N_\ell \times N_m$  scaled pulse functions,  $h(l, m)$ , centered around

$$l = i\Delta_\ell + \frac{\Delta_\ell}{2} - \frac{F_\ell}{2} \quad \text{for } i = 0, \dots, N_\ell - 1 \quad (7)$$

$$m = j\Delta_m + \frac{\Delta_m}{2} - \frac{F_m}{2} \quad \text{for } j = 0, \dots, N_m - 1 \quad (8)$$

for  $\Delta_\ell = \frac{F_\ell}{N_\ell}$  and  $\Delta_m = \frac{F_m}{N_m}$ . Using Eq. 7 and 8 we can represent a continuous image as a discrete sum of shifted pulse functions scaled by  $x[i, j]$ . We refer to this image as  $\hat{I}_\lambda(\mathbf{x})$  for vectorized coefficients  $\mathbf{x}$ .

Due to the shift theorem [34], plugging this image representation into the van Cittert-Zernike theorem (Eq. 1) results in a closed-form solution to the visibilities in terms of  $H(u, v)$ , the Fourier transform of  $h(l, m)$ , as seen in Eq. 6. Note that performing a continuous integration has been reduced to a linear matrix operation similar to a Discrete Time Fourier Transform (DTFT).

In Figure 3 we show that this representation allows us to approximate the true frequency components more accurately than a discretized set of point sources, especially for high frequencies. Any pulse with a continuous closed-form Fourier transform can be used in this representation (e.g. rectangle, triangle, sinc, Gaussian). The chosen pulse places an implicit prior on the reconstruction. For instance, a sinc pulse with frequency and spacing  $\Delta$  can reconstruct any signal with a bandwidth less than  $\frac{\Delta}{2}$  [34]. In this work we choose to use a triangle pulse with width  $(2\Delta_\ell, 2\Delta_m)$  since this is equivalent to linearly interpolating between pulse centers and also simplifies non-negativity constraints. Although we developed this image model for VLBI image reconstruction, it has potential applicability to a much more general class of imaging problems that rely on frequency constraints.

### 4.2. Model Energy

We seek a maximum a posteriori (MAP) estimate of the image coefficients,  $\mathbf{x}$ , given  $M$  complex bispectrum measurements,  $\mathbf{y}$ . Following recent success in image restoration using patch priors [43, 44], we choose to use an expected patch log likelihood (EPLL) and minimize the energy:

$$f_r(\mathbf{x}|\mathbf{y}) = -D(\mathbf{y}|\mathbf{x}) - \text{EPLL}_r(\mathbf{x}) \quad (9)$$

Eq. 9 appears similar to the familiar form using a Bayesian posterior probability; however, since bispectrum measurements are not independent, the data term  $D$  is not a log-likelihood. Additionally,  $\text{EPLL}_r$  is the expected log likelihood of a randomly chosen patch in the image  $\hat{I}_\lambda(\mathbf{x})$ , not the log-likelihood of the image itself [43]. Details of the data and prior term are discussed below.

#### 4.2.1 Data Term $-D(\mathbf{y}|\mathbf{x})$

As discussed in Section 2, bispectrum measurements are invariant to atmospheric inhomogeneity. Therefore, we

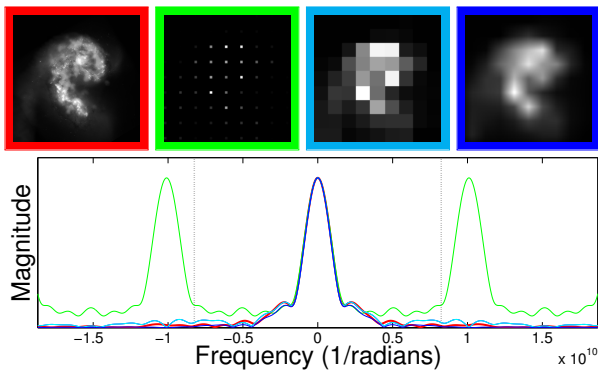


Figure 3. Accurately modeling the frequencies of an image is crucial for fitting VLBI measurements during image reconstruction. Here we show, that with the same number of parameters, we can much more accurately model the true frequency distribution. A slice of frequencies for the true image is shown in red. Overlayed we show the effect of using the traditional discretized imaging model (green), and our improved model for rectangle (cyan) and triangle (blue) pulses. The dotted lines denote the frequency range sampled in Fig 1b. Representing an image in this way reduces modeling errors for higher frequencies during image reconstruction.



$$\begin{aligned}\Gamma(u, v) &\approx \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-i2\pi(ul+vm)} \sum_{i=0}^{N_\ell-1} \sum_{j=0}^{N_m-1} x[i, j] h\left(\ell - \left(\Delta_\ell i + \frac{\Delta_\ell}{2} - \frac{FOV_\ell}{2}\right), m - \left(\Delta_m j + \frac{\Delta_m}{2} - \frac{FOV_m}{2}\right)\right) d\ell dm \\ &= \sum_{i=0}^{N_\ell-1} \sum_{j=0}^{N_m-1} x[i, j] e^{-i2\pi\left(u\left(\Delta_\ell i + \frac{\Delta_\ell}{2} + a_\ell\right) + v\left(\Delta_m j + \frac{\Delta_m}{2} + a_m\right)\right)} H(u, v) = A\mathbf{x} = \left(A^{\Re} + iA^{\Im}\right)\mathbf{x}\end{aligned}\quad (6)$$

choose to express an image's "likelihood" in terms of the bispectrum, rather than visibility, measurements. Let  $y_k$  be a noisy, complex bispectrum measurement corresponding to visibilities  $k_{1,2}$ ,  $k_{2,3}$ , and  $k_{3,1}$  for telescopes 1, 2 and 3. Ideal bispectrum values,  $\xi$ , can be extracted from  $\hat{I}_\lambda(\mathbf{x})$  using the following polynomial equation:

$$\xi_k(\mathbf{x}) = A_{k_{1,2}}\mathbf{x}A_{k_{2,3}}\mathbf{x}A_{k_{3,1}}\mathbf{x} = \xi_k^{\Re}(\mathbf{x}) + i\xi_k^{\Im}(\mathbf{x}) \quad (10)$$

where complex, row vector  $A_{k_{m,n}}$  extracts the 2D frequency,  $(u, v)$ , corresponding to the baseline between telescopes  $m$  and  $n$  from  $\hat{I}_\lambda(\mathbf{x})$ . By assuming Gaussian noise ( $\Sigma_k$ ) on the complex bispectrum measurements, we evaluate  $-D(\mathbf{y}|\mathbf{x})$  as:

$$\gamma \sum_{i=1}^M \left[ \frac{\alpha_k}{2} \begin{pmatrix} \xi_k^{\Re}(\mathbf{x}) - \mathbf{y}_k^{\Re} \\ \xi_k^{\Im}(\mathbf{x}) - \mathbf{y}_k^{\Im} \end{pmatrix}^T \Sigma_k^{-1} \begin{pmatrix} \xi_k^{\Re}(\mathbf{x}) - \mathbf{y}_k^{\Re} \\ \xi_k^{\Im}(\mathbf{x}) - \mathbf{y}_k^{\Im} \end{pmatrix} \right] \quad (11)$$

To account for the independence of each constraint, we set  $\alpha_k = \frac{(T_k-1)(T_k-2)}{2\binom{T_k}{3}} = \frac{3}{T_k}$  for  $T_k$  telescopes observing at the time corresponding to the  $k$ -th bispectrum value. If necessary, effects due to the primary beam and interstellar scattering [15] can be accounted for in Eq. 11 by weighting each bispectrum term appropriately. Refer to the supp. material [7] for further details. The polynomial structure of our "likelihood" formulation is simpler than previously proposed formulations using the bispectrum [5, 40].

**Noise** Although the error due to atmospheric inhomogeneity cancels when using the bispectrum, residual error exists due to thermal noise and gain fluctuations [41]. This introduces Gaussian noise on each complex visibility measurement. Since the bispectrum is the product of three visibilities, its noise distribution is not Gaussian; nonetheless, its noise is dominated by a Gaussian bounding its first-order noise terms. In the supp. material we show evidence that this is a reasonable approximation [7].

#### 4.2.2 Regularizer –EPLL<sub>r</sub>( $\mathbf{x}$ )

We use a Gaussian mixture model (GMM) patch prior to regularize our solution of  $\hat{I}_\lambda(\mathbf{x})$ . Although this is not an explicit prior on  $\hat{I}_\lambda(\mathbf{x})$ , patch priors have been shown to work well in the past [43, 44]. Building on ideas from [43], we maximize the EPLL by maximizing the probability of each of the  $N$  overlapping pulse patches in  $\hat{I}_\lambda(\mathbf{x})$ :

$$\text{EPLL}_r(\mathbf{x}) = \sum_{n=1}^N \log p(P_n\mathbf{x}). \quad (12)$$

$P_n$  is a matrix that extracts the  $n$ -th patch from  $\mathbf{x}$  and  $p(P_n\mathbf{x})$  is the probability of that patch learned through the GMM's optimization. We use a patch size of 8x8.

### 4.3. Optimization

To optimize Eq. 9 we use "Half Quadratic Splitting" [43]. This method introduces a set of auxiliary patches  $\{z^i\}_1^N$ , one for each overlapping patch  $P_i\mathbf{x}$  in the image. We can then solve this problem using an iterative framework:

**(1) Solve for  $\{z^n\}$  given  $\mathbf{x}$ :** In order to complete this step we set  $\{z^n\}$  to the most likely patch under the prior, given the corrupted measurements  $P_n\mathbf{x}$  and weighting parameter  $\beta$  (described further in [43]).

**(2) Solve for  $\mathbf{x}$  given  $\{z^n\}$ :** If we were able to work with visibilities our problem would be quadratic in  $\mathbf{x}$ , and we could solve then for  $\mathbf{x}$  in closed-form. However, since we use bispectrum measurements rather than visibilities, our energy is a 6th order polynomial that cannot be solved in closed-form. One possibility is to solve for the gradient and then use gradient descent to solve for a local minimum. However, this method is slow. Alternatively, we perform a 2nd order Taylor expansion around our current solution,  $\mathbf{x}_0$ , to obtain an approximation of the local minimum in closed-form. A detailed derivation of the gradient and local closed-form solution can be found in the supp. material [7].

As suggested in [43] we iterate between these two steps for increasing  $\beta$  values of 1, 4, 8, 16, 32, 64, 128, 256, 512.

**Multi-scale Framework** Since our convexification of the energy is only approximate, we slowly build up  $\hat{I}_\lambda(\mathbf{x})$  using a multi scale framework. This framework also helps to avoid local minima in the final solution. We initialize the image  $\mathbf{x}_0$  with small random noise centered around the mean flux density (average image intensity), and iterate between solving for  $\{z^n\}$  and  $\mathbf{x}$ . Then, using our discretized formulation of the image, we increase the number of pulses used to describe the image. This framework allows us to find the best low-resolution solution before optimizing the higher frequency detail in the image. In this paper, we initialize our optimization using a set of  $20 \times 20$  pulses and slowly increase to  $64 \times 64$  pulses over 10 scales.

### 5. Dataset

We introduce a dataset and website ([vlbiimaging.csail.mit.edu](http://vlbiimaging.csail.mit.edu)) for evaluating the performance of VLBI image reconstruction algorithms. By supplying a large set of easy-to-understand training and testing data, we hope to make the problem more accessible to those less familiar with the VLBI field. The website contains a:

- Standardized data set of real and synthetic data for training and blind testing of VLBI imaging algorithms

- Automatic quantitative evaluation of algorithm performance on realistic synthetic test data
- Qualitative comparison of algorithm performance
- Online form to easily simulate realistic data using user-specified image and telescope parameters

Current interferometry dataset challenges are both small and assume noise characteristics unsuitable for radio wavelengths [3, 26]. The introduction of this new radio VLBI dataset will help to reveal shortcomings of current methods as well as encourage the development of new algorithms.

**Synthetic Measurements** We provide a standardized format [35] dataset of over 5000 synthetic VLBI measurements corresponding to a variety of array configuration, source images, and noise levels. Measurements are simulated using the MIT Array Performance Simulator (MAPS) software package [28]. This software has been developed to accurately model the visibilities and noise expected from a user-specified array and source. *Visibilities from the MAPS are not generated in the same manner as our forward model.*

We generate data using a collection of black hole [9], celestial [1, 2], and natural images. We have deliberately included a diversity of images in the imaging database, since imaging algorithms for black holes must be sufficiently non-committal that they can identify departures from canonical expectations. Natural images test robustness to complex scenes with varied image statistics.

**Real Measurements** We provide 33 sets of measurements from the VLBA-BU-BLAZAR Program [23] in the same standardized format [35]. This program has been collecting data on a number of gamma-ray blazars every month since 2007. Measurements are taken using the Very Long Baseline Array (VLBA) at 43 GHz. Although both the angular resolution and wavelength of these measurements are very different from those taken by the EHT (which collects at  $\approx 230$  GHz) [15], they provide a means to test algorithms on measured, experimental data.

**Test Set and Error Metrics** We introduce a blind test set of challenging synthetic data. Measurements with realistic errors are generated using a variety of target sources and telescope parameters and provided in the OIFITS format [35]. This test set introduces a means for fair quantitative comparisons between competing algorithms. Researchers are encouraged to run their algorithms on this data and submit results to the website for evaluation.

Traditional point-by-point error metrics, such as MSE and PSNR, are sometimes uninformative in the context of highly degraded VLBI reconstructions. Therefore, we supplement the MSE metric with the perceptually motivated structural similarity (SSIM) index [42]. Since the absolute position of the emission is lost when using the bispectrum, we first align the reconstruction to the ground truth image using cross-correlation. We then evaluate the MSE

and SSIM on the normalized, aligned images. Although we consider MSE and SSIM a good first step towards quantitative analysis, we believe a better metric of evaluation is subject for future research.

## 6. Results and Discussion

Measurements from the EHT have yet to become available. Therefore, we demonstrate the success of our algorithm, CHIRP, on a sample of synthetic examples obtained from our online dataset and real VLBI measurements collected by the VLBA-BU-BLAZAR Program.

**Synthetic Measurements** For image results presented in the paper we generated synthetic data using realistic parameters for the EHT array pointed towards the black hole in M87. Corresponding  $(u, v)$  frequency coverage is shown in Figure 1b. The geometry of an array imposes an intrinsic maximum resolution on the image you can reconstruct from its measurements. Figure 4 shows the effect of filtering out spatial frequencies higher than the minimum fringe spacing. These images set expectations on what is possible to reliably reconstruct from the measurements. Additional results and details about the telescopes, noise properties, and parameter settings can be found in the supp. material [7].

**Method Comparison** We compare results from our algorithm, CHIRP, with the three state-of-the-art algorithms described in Section 3: CLEAN, SQUEEZE, and BSMEM. Images were obtained by asking authors of the competing algorithms or knowledgeable users for a suggested set of reconstruction parameter (provided in the supp. material [7]). The website submission system allows the results from other parameter settings and algorithms to be compared, both qualitatively and quantitatively.

As with our algorithm, SQUEEZE [4] and BSMEM [11] use the bispectrum as input. CLEAN cannot automatically handle large phase errors, so CLEAN results were obtained using *calibrated* (eg. no atmospheric phase error) visibilities in CASA [20]. In reality, these ideal calibrated visibilities would not be available, and the phase would need to be recovered through highly user-dependent self-calibration

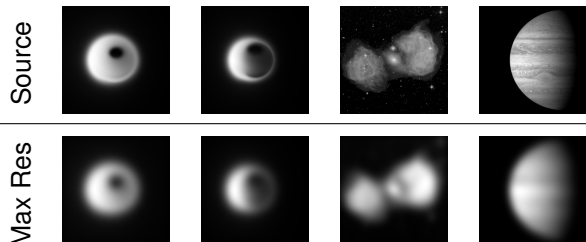


Figure 4. **Intrinsic Maximum Resolution:** The geometry of a telescope array imposes an intrinsic maximum resolution on the image you can reconstruct from the measurements. Recovering spatial frequencies higher than this resolution is equivalent to superresolution. For results presented, the minimum recoverable fringe spacing (corresponding to the maximum frequency) is  $24.72 \mu\text{-arcseconds}$ . The original ‘Source’ images ( $183.82 \mu\text{-arcsecond}$  FOV) are used to synthetically generate realistic VLBI measurements. We show the effect of filtering out spatial frequencies higher than the minimum fringe spacing for these source images in ‘Max Res’.

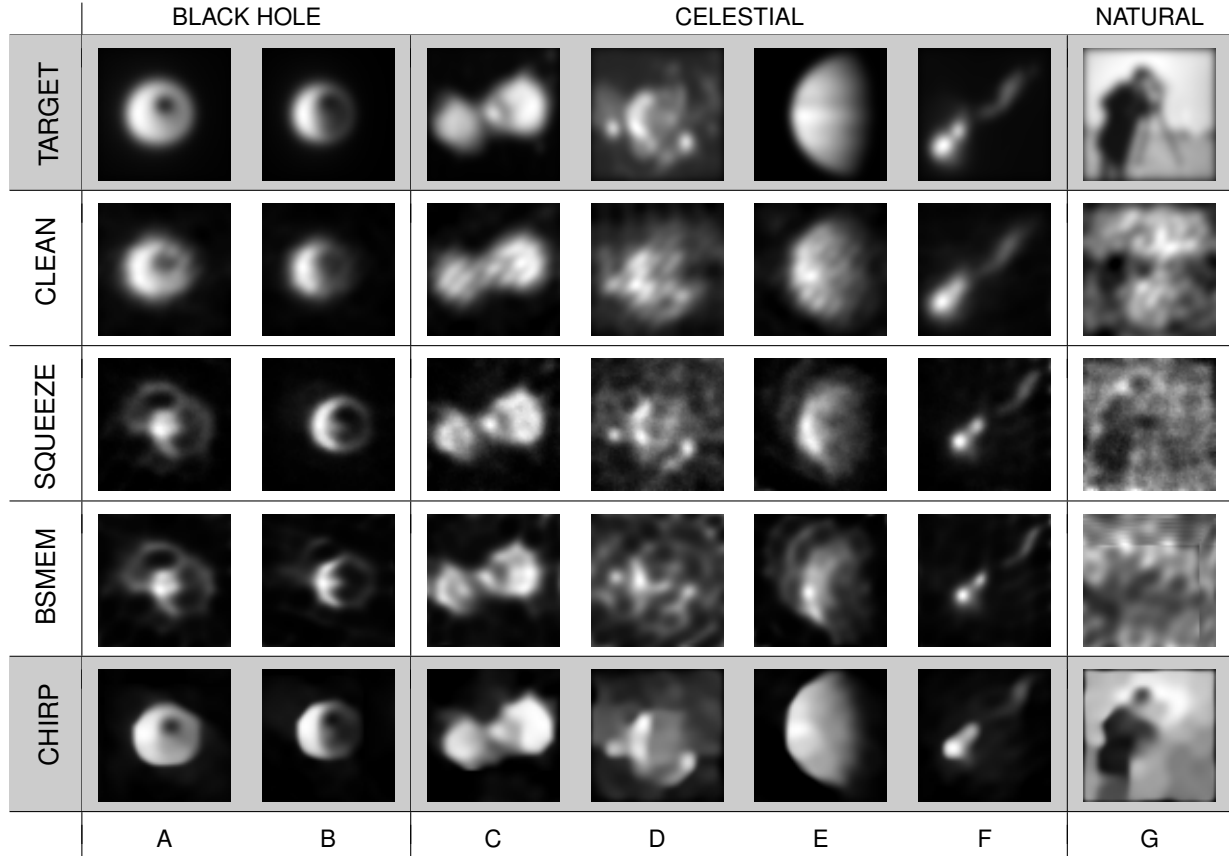


Figure 5. **Method Comparison:** Comparison of our algorithm, ‘CHIRP’ to three state-of-the-art methods: ‘CLEAN’, ‘SQUEEZE’, and ‘BSMEM’. We show the normalized reconstruction of a variety of black hole (a-b), celestial (c-f), and natural (g) source images with a total flux density (sum of pixel intensities) of 1 jansky and a  $183.82 \mu\text{-arcsecond}$  FOV. Since absolute position is lost when using the bispectrum, shifts in the reconstructed source location are expected. The ‘TARGET’ image shows the ground truth emission filtered to the maximum resolution intrinsic to this telescope array.

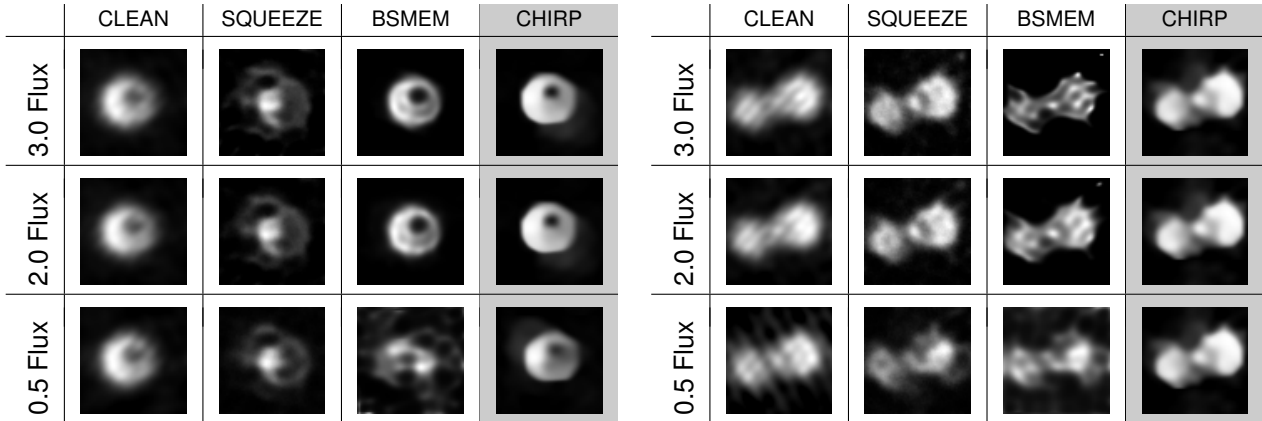


Figure 6. **Noise Sensitivity:** The effect of varying total flux density (in janskys), and thus noise, on each method’s recovered reconstructions. Decreasing flux results in higher noise. Notice how our method is fairly robust to the noise, while the results from other methods often vary substantially across the noise levels. The ground truth target images along with the results for a total flux density of 1 jansky can be seen in column A and C of Figure 5.

methods. However, in the interest of a fair comparison, we show the results of CLEAN in a “best-case” scenario.

Figure 5 shows a sample of results comparing our reconstructions to those of the current state-of-the-art methods. Our algorithm is able to handle a wide variety of sources, ranging from very simple celestial to complex natural images, without any additional parameter tuning. CLEAN produces consistently blurrier results. Both SQUEEZE and

BSMEM tend towards sparser images. This strategy works well for superresolution. However, it comes at the cost of often making extended sources overly sparse and introducing spurious detail. Although algorithms such as BSMEM and SQUEEZE may perform better on these images with specific hand-tuned parameters, these tests demonstrate that the performance of CHIRP requires less user expertise and provides images that may be less sensitive to user bias.

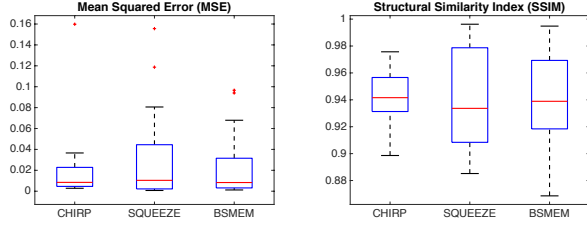


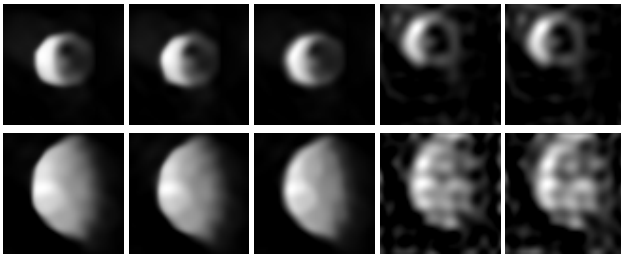
Figure 7. **Quantitative Analysis on Blind Test Set:** Box plots of MSE and SSIM for reconstruction methods on the blind dataset presented in Section 5. In SSIM a score of 1 implies perceptual indistinguishability between the ground truth and recovered image. Scores are calculated using the original ‘Source’ image (Refer to Fig. 4).

Figure 7 shows a quantitative comparison of our method to SQUEEZE and BSMEM for the challenging, blind test set presented in Section 5. Since CLEAN cannot automatically handle large phase errors, we were unable to fairly compare its results on this test set.

**Noise Sensitivity** The standard deviation of thermal noise introduced in each measured visibility is fixed based on measurement choices and the corresponding telescopes’ properties. Specifically,  $\sigma = \frac{1}{0.88} \sqrt{\rho_1 \rho_2 / \nu T}$  for bandwidth  $\nu$ , integration time  $T$ , and each telescope’s System Equivalent Flux Density ( $\rho$ ). Consequently, an emission with a lower total flux will result in a lower SNR signal.

Previous measurements predict that the total flux densities of the black holes M87 and SgA\* will be in the range 0.5 to 3.0 janskys [12, 13]. Figure 6 shows the effect of varying total flux density, and thus noise, on each method’s recovered reconstructions. Notice how our method is fairly robust to the noise, while the results from other methods often vary substantially across noise levels.

**Effect of Patch Prior** Flexibility of the patch prior framework allows us to easily incorporate a variety of different “visual” assumptions in our reconstructed image. For instance, in the case of the EHT, simulations of a black hole for different inclinations and spins can be used to train a patch model that can be subsequently used for reconstruction. In Figure 8 we show results using a patch prior trained on natural [30], celestial, and synthetic black hole images [9]. Only small variations can be observed among the resulting images. Given our selected parameters, this



Natural Celestial Black Hole  $\ell_2$  Norm  $\ell_{0.8}$  Norm

Figure 8. **Effect of Patch Prior:** Reconstructions using patch priors trained on natural, celestial, and synthetic black hole images as well as  $\ell_2$  and  $\ell_{0.8}$  norm priors on the residuals. The ground truth target image are shown in Figure 5 column B and E. The patch priors outperform results obtained using simpler  $\ell$ -norm priors. Since absolute position is lost during imaging, shifts in the reconstructed source location are expected.

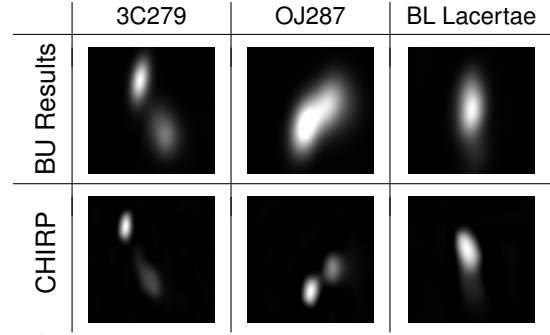


Figure 9. **Real Measurements:** A comparison of our reconstructed images to [23]’s results using CLEAN self-calibration. Note that we are able to reconstruct less blurry images, and are even able to resolve 2 separate, previously unresolved, bright emissions in blazar OJ287. Measurements were taken using the VLBA telescope array. The FOV for each image is 1.5, 1, and 1 milli-arcsecond respectively.

suggests that the prior guides optimization, but does not impose strong assumptions that may greatly bias results.

**Real Measurements** We demonstrate the performance of our algorithm on the reconstruction of three different sources using real VLBI data from [23] in Figure 9. Although we do not have ground truth images corresponding to these measurements, we compare our reconstructions to those generated by the BU group, reconstructed using an interactive calibration procedure described in [23]. Alternatively, we are able to use bispectrum measurements to automatically produce image reconstructions with minimal user input. Notice that we are able to recover sharper images, and even resolve two potentially distinct sources that were previously unresolved in blazar OJ287.

## 7. Conclusion

Astronomical imaging will benefit from the cross-fertilization of ideas with the computer vision community. In this paper, we have presented an algorithm, CHIRP, for reconstructing an image using a very sparse number of VLBI frequency constraints. We have demonstrated improved performance compared to current state-of-the-art methods on both synthetic and real data. Furthermore, we have introduced a new dataset for the testing and development of new reconstruction algorithms. With this paper, the dataset, and algorithm comparison website, we hope to make this problem accessible to researchers in computer vision, and push the limits of celestial imaging.

**Acknowledgments** We would like to thank Andrew Chael, Katherine Rosenfeld, Lindy Blackburn, and Fabien Baron for all of their helpful discussions and feedback. This work was partially supported by NSF CGV-1111415. Katherine Bouman was partially supported by an NSF Graduate Fellowship. We also thank the National Science Foundation (AST-1310896, AST-1211539, and AST-1440254) and the Gordon and Betty Moore Foundation (GBMF-3561) for financial support of this work. This study makes use of 43 GHz VLBA 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 operated by Associated Universities, Inc.



## References

- [1] Jpl image gallery. <http://www.jpl.nasa.gov/spaceimages/>. Accessed: 2015-04-01.
- [2] Nrao image gallery. <https://public.nrao.edu/gallery>. Accessed: 2015-04-01.
- [3] F. Baron, W. D. Cotton, P. R. Lawson, S. T. Ridgway, A. Aarnio, J. D. Monnier, K.-H. Hofmann, D. Schertl, G. Weigelt, E. Thiébaut, et al. The 2012 interferometric imaging beauty contest. In *SPIE Astronomical Telescopes+ Instrumentation*, pages 84451E–84451E. International Society for Optics and Photonics, 2012.
- [4] F. Baron, J. D. Monnier, and B. Kloppenborg. A novel image reconstruction software for optical/infrared interferometry. In *SPIE Astronomical Telescopes+ Instrumentation*, pages 77342I–77342I. International Society for Optics and Photonics, 2010.
- [5] F. Baron and J. S. Young. Image reconstruction at cambridge university. In *SPIE Astronomical Telescopes+ Instrumentation*, pages 70133X–70133X. International Society for Optics and Photonics, 2008.
- [6] J.-P. Berger, F. Malbet, F. Baron, A. Chiavassa, G. Duvert, M. Elitzur, B. Freytag, F. Gueth, S. Hnig, J. Hron, H. Jang-Condell, J.-B. Le Bouquin, J.-L. Monin, J. Monnier, G. Perrin, B. Plez, T. Ratzka, S. Renard, S. Stefl, E. Thibaut, K. Tristram, T. Verhoelst, S. Wolf, and J. Young. Imaging the heart of astrophysical objects with optical long-baseline interferometry. *The Astronomy and Astrophysics Review*, 20(1), 2012.
- [7] K. L. Bouman, M. D. Johnson, D. Zoran, V. L. Fish, S. S. Doeleman, and W. T. Freeman. Computational imaging for vlbi image reconstruction supplemental material. <http://vlbiimaging.csail.mit.edu/static/papers/CHIRPsupp.pdf>.
- [8] R. Bracewell. *Fourier analysis and imaging*. Springer Science & Business Media, 2004.
- [9] A. E. Broderick, V. L. Fish, S. S. Doeleman, and A. Loeb. Evidence for Low Black Hole Spin and Physically Motivated Accretion Models from Millimeter-VLBI Observations of Sagittarius A\*. *The Astrophysical Journal*, July 2011.
- [10] A. E. Broderick, A. Loeb, and R. Narayan. The event horizon of sagittarius a\*. *The Astrophysical Journal*, 2009.
- [11] D. Buscher. Direct maximum-entropy image reconstruction from the bispectrum. In *Very High Angular Resolution Imaging*, pages 91–93. Springer, 1994.
- [12] S. S. Doeleman, V. L. Fish, D. E. Schenck, C. Beaudoin, R. Blundell, G. C. Bower, A. E. Broderick, R. Chamberlin, R. Freund, P. Friberg, et al. Jet-launching structure resolved near the supermassive black hole in m87. *Science*, 2012.
- [13] S. S. Doeleman, J. Weintraub, A. E. Rogers, R. Plambeck, R. Freund, R. P. Tilanus, P. Friberg, L. M. Ziurys, J. M. Moran, B. Corey, et al. Event-horizon-scale structure in the supermassive black hole candidate at the galactic centre. *Nature*, 2008.
- [14] M. Felli and R. E. Spencer. *Very Long Baseline Interferometry*. Springer, 1989.
- [15] V. L. Fish, M. D. Johnson, R.-S. Lu, S. S. Doeleman, K. L. Bouman, D. Zoran, W. T. Freeman, D. Psaltis, R. Narayan, V. Pankratius, et al. Imaging an event horizon: Mitigation of scattering toward sagittarius a\*. *The Astrophysical Journal*, 2014.
- [16] W. T. Freeman, T. R. Jones, and E. C. Pasztor. Example-based super-resolution. *Computer Graphics and Applications, IEEE*, 2002.
- [17] D. Glasner, S. Bagon, and M. Irani. Super-resolution from a single image. In *Computer Vision, 2009 IEEE 12th International Conference on*. IEEE, 2009.
- [18] J. Högbom. Aperture synthesis with a non-regular distribution of interferometer baselines. *Astron. Astrophys. Suppl.*, 1974.
- [19] T. Isernia, V. Pascazio, R. Pierri, and G. Schirrinzi. Synthetic aperture radar imaging from phase-corrupted data. *IEE Proceedings - Radar, Sonar and Navigation*, Aug 1996.
- [20] S. Jaeger. The common astronomy software application (casa). In *Astronomical Data Analysis Software and Systems XVII*, 2008.
- [21] M. Jehle, D. Perler, D. Small, A. Schubert, and E. Meier. Estimation of atmospheric path delays in terrasar-x data using models vs. measurements. *Sensors*, 2008.
- [22] T. Johannsen and D. Psaltis. Testing the No-hair Theorem with Observations in the Electromagnetic Spectrum. II. Black Hole Images. *The Astrophysical Journal*, July 2010.
- [23] S. G. Jorstad, A. P. Marscher, M. L. Lister, A. M. Stirling, T. V. Cawthorne, W. K. Gear, J. L. Gómez, J. A. Stevens, P. S. Smith, J. R. Forster, et al. Polarimetric observations of 15 active galactic nuclei at high frequencies: jet kinematics from bimonthly monitoring with the very long baseline array. *The Astronomical Journal*, 2005.
- [24] T. Krichbaum, D. Graham, M. Bremer, W. Alef, A. Witzel, J. Zensus, and A. Eckart. Sub-millarcsecond imaging of sgr a\* and m 87. In *Journal of Physics: Conference Series*. IOP Publishing, 2006.
- [25] B. G. Kurien, Y. Rachlin, V. N. Shah, J. B. Ashcom, and V. Tarokh. Compressed sensing techniques for image reconstruction in optical interferometry. In

*Imaging and Applied Optics 2014*. Optical Society of America, 2014.

- [26] P. R. Lawson, W. D. Cotton, C. A. Hummel, J. D. Monnier, M. Zhao, J. S. Young, H. Thorsteinsson, S. C. Meimon, L. M. Mugnier, G. Le Besnerais, et al. An interferometry imaging beauty contest. In *Astronomical Telescopes and Instrumentation*. International Society for Optics and Photonics, 2004.
- [27] A. Levin, Y. Weiss, F. Durand, and W. T. Freeman. Efficient marginal likelihood optimization in blind deconvolution. In *Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2011.
- [28] R.-S. Lu, A. E. Broderick, F. Baron, J. D. Monnier, V. L. Fish, S. S. Doeleman, and V. Pankratius. Imaging the Supermassive Black Hole Shadow and Jet Base of M87 with the Event Horizon Telescope. *The Astrophysical Journal*, June 2014.
- [29] M. Lustig, D. Donoho, and J. M. Pauly. Sparse mri: The application of compressed sensing for rapid mr imaging. *Magnetic resonance in medicine*, 2007.
- [30] D. Martin, C. Fowlkes, D. Tal, and J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*. IEEE, 2001.
- [31] J. D. Monnier and R. J. Allen. Radio and optical interferometry: Basic observing techniques and data analysis. In *Planets, Stars and Stellar Systems*. Springer, 2013.
- [32] A. Moreira, P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek, and K. P. Papathanassiou. A tutorial on synthetic aperture radar. *IEEE Geoscience and Remote Sensing Magazine*, March 2013.
- [33] D. C. Munson, J. D. O'Brien, and W. K. Jenkins. A tomographic formulation of spotlight-mode synthetic aperture radar. *Proceedings of the IEEE*, Aug 1983.
- [34] A. V. Oppenheim and A. S. Willsky. *Signals and systems*. Prentice-Hall, 1997.
- [35] T. A. Pauls, J. S. Young, W. Cotton, and J. Monnier. A data exchange standard for optical (visible/ir) interferometry. *Publications of the Astronomical Society of the Pacific*, 2005.
- [36] S. Renard, E. Thiébaut, and F. Malbet. Image reconstruction in optical interferometry: benchmarking the regularization. *Astronomy & Astrophysics*, 2011.
- [37] J. Skilling. Quantified maximum entropy. In *Maximum Entropy and Bayesian Methods*. Springer, 1990.
- [38] G. B. Taylor, C. L. Carilli, and R. A. Perley. Synthesis imaging in radio astronomy ii. In *Synthesis Imaging in Radio Astronomy II*, 1999.
- [39] J.-B. Thibault, K. D. Sauer, C. A. Bouman, and J. Hsieh. A three-dimensional statistical approach to improved image quality for multislice helical ct. *Medical physics*, 2007.
- [40] E. Thiebaut and J.-F. Giovannelli. Image reconstruction in optical interferometry. *Signal Processing Magazine, IEEE*, Jan 2010.
- [41] A. R. Thompson, J. M. Moran, and G. W. Swenson Jr. *Interferometry and synthesis in radio astronomy*. John Wiley & Sons, 2008.
- [42] Z. Wang, E. P. Simoncelli, and A. C. Bovik. Multiscale structural similarity for image quality assessment. In *Signals, Systems and Computers, 2004. Conference Record of the Thirty-Seventh Asilomar Conference on*. Ieee, 2003.
- [43] D. Zoran and Y. Weiss. From learning models of natural image patches to whole image restoration. In *Computer Vision (ICCV), 2011 IEEE International Conference on*. IEEE, 2011.
- [44] D. Zoran and Y. Weiss. Natural images, gaussian mixtures and dead leaves. In *Advances in Neural Information Processing Systems*, 2012.