Astronomy & Astrophysics manuscript no. main
August 6, 2019
©ESO 2019

VLBI imaging of black holes via second moment regularization

S. Issaoun[1,2], M. D. Johnson[2], L. Blackburn[2], M. Mościbrodzka[1], A. Chael[2], and H. Falcke[1]

[1] Department of Astrophysics/IMAPP, Radboud University, P.O. Box 9010, 6500 GL Nijmegen, The Netherlands
[2] Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA

August 6, 2019

ABSTRACT

The imaging fidelity of the Event Horizon Telescope (EHT) is currently determined by its sparse baseline coverage. In particular, EHT coverage is dominated by long baselines, and is highly sensitive to atmospheric conditions and loss of sites between experiments. The limited short/mid-range baselines especially affect the imaging process, hindering the recovery of more extended features in the image. We present an algorithmic contingency for the absence of well-constrained short baselines in the imaging of compact sources, such as the supermassive black holes observed with the EHT. This technique enforces a specific second moment on the reconstructed image in the form of a size constraint, which corresponds to the curvature of the measured visibility function at zero baseline. The method enables the recovery of information lost in gaps of the baseline coverage on short baselines and enables corrections of any systematic amplitude offsets for the stations giving short-baseline measurements present in the observation. The regularization can use historical source size measurements to constrain the second moment of the reconstructed image to match the observed size. We additionally show that a characteristic size can be derived from available short-baseline measurements, extrapolated from other wavelengths, or estimated without complementary size constraints with parameter searches. We demonstrate the capabilities of this method for both static and movie reconstructions of variable sources.

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

arXiv:1908.01296v1 [astro-ph.IM] 4 Aug 2019

1. Introduction

Very-long-baseline interferometry (VLBI) is a technique able to achieve high angular resolution imaging through the use of widely separated antennas. Unfortunately, as the observing frequency is increased, the availability of suitable sites on Earth is greatly reduced, leading to sparse arrays with a high angular resolution but a low spatial dynamic range. In particular, a simple inverse Fourier transform of the visibilities measured by an interferometer, or 'dirty image', is dominated by artifacts introduced by sparse sampling of the Fourier plane. Short baselines are particularly important in imaging, as they anchor the flux distribution and provide a crucial link between high-resolution small-scale features and the large-scale extent and morphology of the target. The sparser the array, the more challenging it is to reconstruct images from interferometric measurements. Additionally, weather and technical issues at sites that provide short/mid-range baselines can greatly degrade the ability to image a given data set.

Array sparsity and station-based errors can have dramatic effects on reconstructed images. Thus, the imaging process requires further information and assumptions beyond the visibility measurements from the interferometer. The choice of imaging method imposes additional constraints on the reconstructed image. Here, we will focus on extending the method of regularized maximum likelihood (RML) that performs well under sparse sampling conditions and does not involve direct inverse Fourier transforms of the data in the imaging process.

In this paper we present an algorithmic contingency to array sparsity and site issues in the form of a second moment regularization function. That is, the compactness of the source can be expressed as the second moment of the source brightness distribution (Moffet 1962; Burn & Conway 1976), which can be constrained to match, for example, confident source size measurements from short baselines of previous experiments or epochs. Enforcing this source size constraint supplements limited short-baseline information while fitting to long-baseline smaller scale structure from newer observations.

The Event Horizon Telescope (EHT), observing at a frequency of 230 GHz (Event Horizon Telescope Collaboration et al. 2019a,b), is a prime example of a high-frequency VLBI imaging experiment with image uncertainties dominated by the effects of sparse coverage. The EHT currently has only a single short/mid-range VLBI baseline, joining the Large Millimeter Telescope Alfonso Serrano (LMT) in Mexico to the Submillimeter Telescope (SMT) in Arizona. Recent observations with the EHT have shown that the LMT is difficult to calibrate, giving baselines with large measurement uncertainties dominated by uncharacterized station behavior in 2017 (Event Horizon Telescope Collaboration et al. 2019c,d).

Although the EHT observes a number of non-horizon-scale sources in conjunction with the Atacama Large Millimeter/submillimeter Array (ALMA), its primary targets are the two supermassive black hole candidates in the Galactic Center, Sagittarius A* (Sgr A*), and at the center of the radio galaxy M87. At the frequency of the EHT, these two sources are very compact, with sizes on the sky histor-

ically measured with three stations, in California, Arizona, and Hawaii, in early EHT observations, and are thus ideal imaging targets for second moment regularization (Doeleman et al. 2008; Fish et al. 2011; Doeleman et al. 2012; Akiyama et al. 2015; Johnson et al. 2015; Lu et al. 2018). Near-zero closure phases on the California–Arizona–Hawaii triangle are indicative of the source compactness and symmetry on scales of a few tens of μas (Akiyama et al. 2015; Fish et al. 2016). The California–Arizona baseline provided the short-baseline measurements needed to constrain the compactness and size of the sources in the visibility domain. Recent observations of M87 in 2017 also found a source size of $\sim 40\,\mu$as consistent with previous measurements (Event Horizon Telescope Collaboration et al. 2019a,b,c,d,e,f).

For Sgr A*, the source size is also well-constrained at lower frequencies due to its compactness and dominant diffractive scattering (Shen et al. 2005; Bower et al. 2006; Lu et al. 2011; Johnson et al. 2018). VLBI observations at 86 GHz taken one month apart give fitted Gaussian source sizes for the scattered image of Sgr A*with $< 10\%$ difference (Ortiz-León et al. 2016; Brinkerink et al. 2019). At this frequency, while the small scale structure is expected to vary, the large-scale information, dominated by the size of the scattering kernel, should be stable from epoch to epoch (Johnson et al. 2018).

Second moment regularization merges the benefits of model-fitting with the flexibility of imaging: compared to self-calibration to a known model, it does not actually modify the measured visibilities used for the imaging process or enforce a model-dependent solution, but instead provides additional information to improve image quality. The regularization constrains the spread of flux density to a motivated region in the image, discouraging non-physical morphology driven by fits to long-baseline data and accelerating convergence toward a plausible image. It is a natural extension of imaging tools that add source information in the imaging process in RML methods: a total flux constraint is in fact the zeroth moment of the image; an image centroid specification corresponds to the first moment of the image; and a short-baseline source size completes the picture by constraining the image second moment. The implementation of second moment regularization can be done in conjunction with other tools and constraints in RML, for both static and movie reconstructions. Furthermore, as the constraint function acts on the image itself and does not modify the visibility data, it can be used with any choice of data product, including minimally-calibrated closure phases and amplitudes.

The paper is structured as follows. We present the mathematical background to motivate the regularization in Sect. 2. We outline the method, assumptions, and physical motivation in Sect. 3. In Sect. 4 we demonstrate the improvements in image quality and fidelity using the regularization with or without a priori knowledge of the source size. We present possible applications of the second moment regularization to more sophisticated imaging techniques for scattering mitigation and movie reconstructions in Sect. 5. A summary is given in Sect. 6.

2. Background

By the van Cittert-Zernike theorem, an interferometer samples complex visibilities corresponding to Fourier components of an image (van Cittert 1934; Zernike 1938). Consequently, $n^{\rm th}$ moments of an image correspond to $n^{\rm th}$ derivatives of the visibility function at the origin. Specifically, an interferometric visibility $V(\mathbf{u})$ on a baseline $\mathbf{u}$ can be written as (e.g., Thompson et al. 2017)

$$V(\mathbf{u}) = \int d^2\mathbf{x}\, I(\mathbf{x}) e^{-2\pi i \mathbf{u}\cdot\mathbf{x}}, \tag{1}$$

where $I(\mathbf{x})$ is the brightness distribution on the sky, and $\mathbf{x}$ is an angular unit.

From this expression, $V(\mathbf{0}) = \int d^2\mathbf{x}\, I(\mathbf{x}) \in \mathbb{R}$ gives the total flux density of the image (the $0^{\rm th}$ moment). Likewise, the phase gradient of the visibility function at zero baseline gives a vector proportional to the centroid of the image,

$$\begin{aligned}\nabla V(\mathbf{u})\big\rfloor_{\mathbf{u}=\mathbf{0}} &= -2\pi i \int d^2\mathbf{x}\,\mathbf{x} I(\mathbf{x}) \\ &= -2\pi i V(\mathbf{0})\boldsymbol{\mu},\end{aligned} \tag{2}$$

where $\boldsymbol{\mu}$ is the image centroid (the normalized $1^{\rm st}$ moment):

$$\boldsymbol{\mu} = (\bar{x}\hat{\mathbf{x}}, \bar{y}\hat{\mathbf{y}}) = \frac{\int d^2\mathbf{x} I(\mathbf{x})\mathbf{x}}{\int d^2\mathbf{x}\, I(\mathbf{x})}. \tag{3}$$

Because the image is real, the gradient $\nabla V(\mathbf{u})\big\rfloor_{\mathbf{u}=\mathbf{0}}$ is purely imaginary. For images that are positive (e.g., images in total intensity), the visibility function must take its maximum amplitude at the origin. More generally, the visibility function is Hermitian; thus, its amplitude must always have a vanishing gradient at the origin because of the conjugation symmetry $V(\mathbf{u}) = V^*(-\mathbf{u})$.

The second derivative, or Hessian, of the visibility amplitude function at zero baseline gives a matrix (see Appendix A.1):

$$\begin{aligned}\nabla\nabla^\intercal |V(\mathbf{u})|\big\rfloor_{\mathbf{u}=\mathbf{0}} &= -4\pi^2 \int d^2\mathbf{x}\, I(\mathbf{x})(\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})^\intercal \\ &= -4\pi^2 V(\mathbf{0})\boldsymbol{\Sigma},\end{aligned} \tag{4}$$

where $\boldsymbol{\Sigma}$ is the normalized second central moment (or covariance matrix) of the image. We show in Appendix A.1 that this expression is equivalent to the curvature of the centered complex visibility function (see also Moffet 1962; Burn & Conway 1976). The visibility amplitude function is a more natural data product to use for observations with non-astrometric VLBI arrays such as the EHT, where there is no absolute phase information due to strong differential atmospheric propagation effects between sources, and thus no directly measured full complex visibilities. Therefore it is useful for us to determine image moments directly from the visibility amplitude function, which is measured.

The image covariance matrix $\boldsymbol{\Sigma}$ can be more intuitively expressed in terms of its principal axes, corresponding to the perpendicular axes about which the second moment reaches its maximum (Hu 1962). The matrix has two eigenvalues $\lambda_{\rm min}$ and $\lambda_{\rm maj}$, and can be diagonalized as follows:

$$\boldsymbol{\Sigma} = \mathbf{R}_\phi \begin{pmatrix} \lambda_{\rm min} & 0 \\ 0 & \lambda_{\rm maj} \end{pmatrix} \mathbf{R}_\phi^\intercal, \tag{5}$$

where $\mathbf{R}_\phi$ is the rotation matrix based on the position angle east of north ϕ of the major principal axis (Appendix A.2). The eigenvalues of the covariance matrix are the variances of the normalized image projected along the principal (major and minor) axes. The correspondence between $\lambda_{\rm maj}$,

Table 1. Correspondence of the mass, center of mass and moment of inertia in the image and visibility domains.

Physical Analog	Image Domain		Visibility Domain	
Mass	Total Flux	$\int I(\mathbf{x})d^2\mathbf{x}$	Peak Visibility	$V(\mathbf{0})$
Center of Mass	Centroid ($\boldsymbol{\mu}$)	$V(\mathbf{0})^{-1}\int \mathbf{x}I(\mathbf{x})d^2\mathbf{x}$	Phase Gradient	$(2\pi i V(\mathbf{0}))^{-1}\,\nabla V(\mathbf{u})\rfloor_{\mathbf{u}=\mathbf{0}}$
Moment of Inertia	Covariance ($\boldsymbol{\Sigma}$)	$V(\mathbf{0})^{-1}\int \mathbf{x}\mathbf{x}^\intercal I(\mathbf{x})d^2\mathbf{x}$	Amplitude Curvature	$(-4\pi^2 V(\mathbf{0}))^{-1}\,\nabla\nabla^\intercal V(\mathbf{u})\rfloor_{\mathbf{u}=\mathbf{0}}$

$\lambda_{\rm min}$, ϕ and the individual terms of $\boldsymbol{\Sigma}$ is given in Appendix A.2.

Following Equation 1, we can fully express the visibility function as a Taylor expansion in its derivatives. Each $n+1^{\rm th}$ term of the Taylor expansion is proportional to the $n^{\rm th}$ moment of the visibility function (see Table 1). At zero baseline, only the zeroth moment remains. We choose the coordinate system such that the centroid of the image is at the origin, and the first moment of the visibility function (the second term of the Taylor expansion) vanishes. At short baseline, the centered complex visibility function is therefore dominated by the quadratic term. The Taylor expansion of the visibility function at short baseline becomes:

$$V(\mathbf{u}) \simeq V(\mathbf{0}) - 2\pi^2 \int d^2\mathbf{x}\,(\mathbf{u}\cdot\mathbf{x})^2 I(\mathbf{x})$$
$$\simeq V(\mathbf{0}) - 2\pi^2 V(\mathbf{0})\mathbf{u}^\intercal\boldsymbol{\Sigma}\mathbf{u}. \quad (6)$$

Equation 6 describes the visibility function behavior on short baselines entirely in terms of the total flux $V(\mathbf{0})$ and the second moment covariance matrix $\boldsymbol{\Sigma}$ projected along the baseline direction. These parameters also describe a unique visibility function of a Gaussian source with total flux $V(\mathbf{0})$, and major/minor axes sizes and orientation prescribed by the same covariance matrix. We show this by comparing the general complex visibility function to that for a Gaussian source. For the simplest case of an isotropic Gaussian source of standard deviation σ with the same total flux $V(\mathbf{0})$, we have the following intensity pattern on the sky and corresponding visibility function:

$$I_{\rm gauss}(\mathbf{x}) = \frac{V(\mathbf{0})}{2\pi\sigma^2} e^{-|\mathbf{x}|^2/2\sigma^2}, \quad (7)$$
$$V_{\rm gauss}(\mathbf{u}) = V(\mathbf{0}) e^{-2\pi^2|\mathbf{u}|^2\sigma^2}. \quad (8)$$

More generally, an anisotropic Gaussian with a covariance matrix $\boldsymbol{\Sigma}$ gives:

$$I_{\rm gauss}(\mathbf{x}) = \frac{V(\mathbf{0})}{2\pi\sqrt{|\boldsymbol{\Sigma}|}} e^{-\mathbf{x}^\intercal\boldsymbol{\Sigma}^{-1}\mathbf{x}}, \quad (9)$$
$$V_{\rm gauss}(\mathbf{u}) = V(\mathbf{0}) e^{-2\pi^2\mathbf{u}^\intercal\boldsymbol{\Sigma}\mathbf{u}}. \quad (10)$$

Taking the Taylor expansion of the anisotropic Gaussian visibility function at short baselines, the first two terms dominate:

$$V_{\rm gauss}(\mathbf{u}) \simeq V(\mathbf{0}) - 2\pi^2 V(\mathbf{0})\mathbf{u}^\intercal\boldsymbol{\Sigma}\mathbf{u}. \quad (11)$$

We thus obtain an equivalence of the behavior of the general visibility function (Equation 6) and the Gaussian visibility function (Equation 11) at short baselines. This relation allows us to translate the second moment covariance matrix of the general visibility function to the covariance matrix of an anisotropic Gaussian, which provides a simple parametrization to describe the second moment in terms of the characteristic source extent. The sizes of the major and minor axes $\theta_{\rm maj}$ and $\theta_{\rm min}$ are simply the full widths at half-maximum (FWHMs) of the equivalent Gaussian derived from the variances projected along each principal axis:

$$\theta_{\rm maj} = \sqrt{8\ln(2)\lambda_{\rm maj}}, \quad (12)$$
$$\theta_{\rm min} = \sqrt{8\ln(2)\lambda_{\rm min}}. \quad (13)$$

The equivalence to the Gaussian also gives a natural break-off point where the characteristic source size constraint from the second moment ceases to be a good approximation to the full visibility function: the $1/e$ point determining the resolvability of a Gaussian translates to the baseline length at which the visibility amplitude reaches $V(\mathbf{0})/e$. Baseline lengths longer than the $1/e$ point will lead to higher order terms of the Taylor expansion dominating the behavior and sampling finer structure in the image. We employ the $1/e$ point as a conceptual and visual limit for the source size constraint applied via the second moment regularization. It is not a hard cut-off enforced by the imaging process.

In Fig. 1, we demonstrate the behavior of the normalized visibility amplitudes sampled along the source major axis as a function of projected baseline length for three images with distinctly different structure but an identical second moment. The behavior on short baselines aligns well for all three images, the amplitudes start to diverge at longer baselines. We denote the $1/e$ limit, corresponding to the resolvability of the Gaussian image, with a magenta vertical

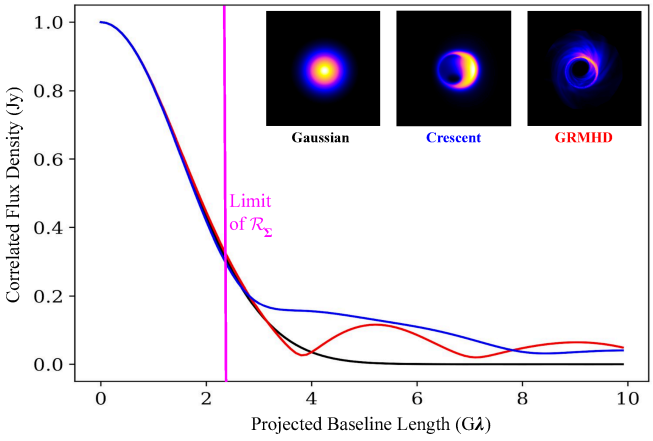

Fig. 1. Three images with equal extent along their respective major axis, from left to right: a Gaussian; a crescent model; a ray-traced image from a general relativistic magnetohydrodynamics (GRMHD) simulation of a black hole shadow and accretion disk. Model visibility amplitudes along the major axis of each source as a function of (u, v) distance, after flux and size normalization, show identical behavior at short baseline length but diverge at longer baseline length: the Gaussian in black; the crescent in blue; and the GRMHD simulation in red.

line. On baselines past this line, the amplitudes show very different behavior, dominated by the smaller-scale features in each image (or lack thereof). We can thus express the visibility amplitude function behavior on short baselines via the second moment of the image, defined by the total flux and just three Gaussian parameters: the principal axes FWHMs $\theta_{\rm maj}$ and $\theta_{\rm min}$ and the position angle ϕ of the major axis east of north. In the RML imaging process, there is an additional fifth input parameter, governing the weight of the second moment regularization, or hyperparameter $\beta_{\rm R}$, following Equation 14.

3. Method

RML focuses on pixelized reconstructions of the image, iteratively maximizing an "objective function", which is analogous to a log posterior probability function. This function is a weighted (via "hyperparameters") sum of both $\chi^2_{\rm D}$ goodness-of-fit data terms, and regularization functions $S_{\rm R}$, or "regularizers", governing specific image properties. In this paper, we use the RML method implemented in the `eht-imaging` Python library (Chael et al. 2016, 2018), where the objective function $J(I)$ is minimized via gradient descent, and can be written as:

$$J(I) = \sum_{\text{data terms}} \alpha_{\rm D} \chi^2_{\rm D}(I) - \sum_{\text{regularizers}} \beta_{\rm R} S_{\rm R}(I), \tag{14}$$

where $\alpha_{\rm D}$ and $\beta_{\rm R}$ are the input hyperparameters.

Using only five input parameters to the regularization ($V(\mathbf{0})$, $\theta_{\rm maj}$, $\theta_{\rm min}$, ϕ and $\beta_{\rm R}$) we can now constrain the second moment of the reconstructed image to match the size constraint provided by the user for RML imaging. In Sect. 3.1 we present our implementation of the second moment regularization function within the `eht-imaging` library minimization framework. In Sect. 3.2 we describe the assumptions and physical motivation for second moment regularization using historical observational measurements, known source properties and theoretical expectations.

3.1. Second moment regularization

Regularization functions in imaging enforce constraints on particular properties of the image, such as image entropy (e.g., Narayan & Nityananda 1986), smoothness (Bouman et al. 2016; Chael et al. 2016; Kuramochi et al. 2018) and/or sparsity (Wiaux et al. 2009a,b; Honma et al. 2014; Akiyama et al. 2017a,b). Simple constraints, such as image positivity, image total flux (zeroth moment) or image centering (first moment), are often applied to the image, utilizing known information on the behavior of the total intensity distribution of the source imaged. The implementation of a second moment regularization, constraining the size of the source, is thus a natural extension of common imaging tools that add source information to the imaging process.

We define a regularization function that is minimized when the covariance matrix of the reconstructed image $\mathbf{\Sigma}$ matches a user-specified covariance matrix $\mathbf{\Sigma}'$. In practice, this latter matrix is computed using user-specified principal axes FWHMs and position angle. We utilize the Frobenius norm to determine a penalty function that quantifies the difference between the user-specified and reconstructed covariance matrices:

$$\mathcal{R}_{\mathbf{\Sigma}} \equiv \mathrm{Tr}\left[\left(\mathbf{\Sigma} - \mathbf{\Sigma}'\right)^{\intercal}\left(\mathbf{\Sigma} - \mathbf{\Sigma}'\right)\right] \tag{15}$$

This regularizer is, by definition, simply the minimization of the difference between two covariance matrices. The procedure for the regularizer implementation in the `eht-imaging` library via gradient descent is presented in Appendix B.

3.2. Assumptions

The second moment regularization operates under a few key assumptions on the properties of the source observed. The main assumption of this method is the compactness of the source. In order to get a quadratic fall-off in the visibility function, as shown in Sect. 2, the source must be compact and resolved on longer baselines of the interferometer. This method would break down for point sources or sources with complex morphology and diffuse flux on large scales.

Another assumption concerns the stability of the source size across multiple epochs. The input axis sizes and position angle for the regularization will only be valid if the source does not vary significantly in size between observations. The source size input is typically derived from observations where weather conditions, coverage, and station performance on short baselines were adequate for higher precision model fitting. The source size can then be used for data sets with larger uncertainties to improve the fidelity and convergence of the imaging process. This assumption is well-motivated for the compact sources observed with the EHT:

- Sgr A* at 86 GHz, has been model-fitted with varying precision over two decades, with little variation in the obtained source size parameters, (Rogers et al. 1994; Krichbaum et al. 1998; Doeleman et al. 2001; Shen et al. 2005; Lu et al. 2011; Ortiz-León et al. 2016; Brinkerink et al. 2019)
- Sgr A* at 230 GHz has been measured to be compact and stable in size between 2007 and 2013 (Doeleman et al. 2008; Lu et al. 2018; Johnson et al. 2018),
- M87 at 230 GHz has been measured to be compact and stable in size over a decade (Doeleman et al. 2012; Akiyama et al. 2015; Event Horizon Telescope Collaboration et al. 2019a,b,c,d,e,f).

It is worth noting that this assumption breaks down for sources with multiple bright components moving relative to each other, as is common for multi-epoch images of bright jet sources. An overall size measurement from a single epoch would not translate to other observations due to components appearing or moving outward, changing the source morphology significantly between observations. The quadratic fall-off approximation until the $1/e$ point would also not hold for two separated point sources, which do show a quadratic fall-off in the visibility amplitudes but the amplitudes would quickly evolve to more complex structure on longer baselines that could be identified as the behavior of two point sources interfering. The method is most effective whenever the emission is confined within a single compact region or on multiple scales that are substantially separated, and particularly if the scale of the emission in the image is comparable to the resolution of the array.

We also assume that the extent of the source does not significantly vary within a single epoch. For static imaging of slow-varying sources, it suffices to assume that the average size of the source matches the input, but this has further implications on reconstructions of variable sources within a single epoch. The structural variability on short timescales

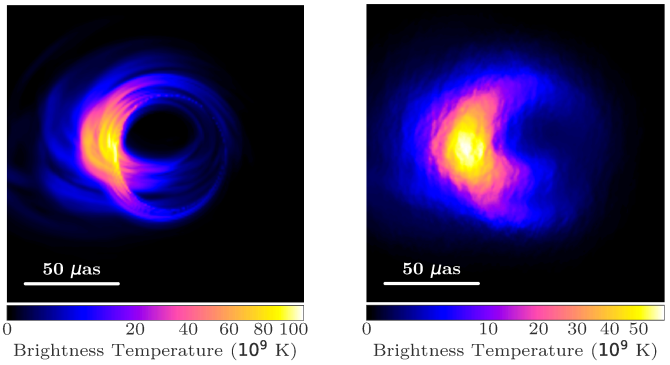

Fig. 2. Left: 230 GHz GRMHD simulation of Sgr A* (Mościbrodzka & Gammie 2018). Right: Same simulation including the effects of interstellar scattering (Johnson 2016; Johnson et al. 2018).

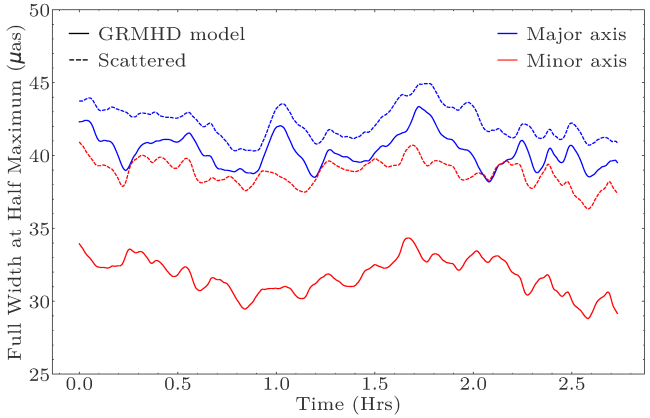

Fig. 3. Principal axes FWHMs as a function of time for the simulation of Sgr A* in Fig. 2 (Mościbrodzka & Gammie 2018). The solid lines show sizes for the simulation, the dotted lines show sizes for the simulation including the effects of interstellar scattering (Johnson 2016; Johnson et al. 2018). The scattering major axis is aligned with the source minor axis, and thus the scattering kernel slightly dominates the minor axis size, which stabilizes the minor axis FWHM time series. The sizes were obtained from measurements of the image second moment per frame. For all four size trends, the deviation about the mean size is $< 10\%$.

should be contained within the region constrained by the second moment. This is an issue particularly for imaging Sgr A*, as the source is known to vary on timescales of minutes, much shorter than the length of a single observing epoch. We assess the degree of variability of the source extent in quiescent (non-flaring) models of Sgr A* using general relativistic magnetohydrodynamic (GRMHD) simulations of variable emission on horizon scales (Fig. 2; Mościbrodzka & Gammie 2018). In Fig. 3, we show the variation in the principal axes FWHMs for a typical GRMHD simulation of the accretion flow of Sgr A* at 230 GHz, both excluding and including the effects of scattering due to the interstellar medium in our line of sight (Johnson 2016; Johnson et al. 2018). Although the simulation shows structural changes in the source morphology, deviations about the mean FWHM remain below 10% for both the model and scattered simulation principal axes.

Furthermore, the emitting gas around supermassive black holes in low-luminosity active galactic nuclei becomes optically thin as we increase the observing frequency. The

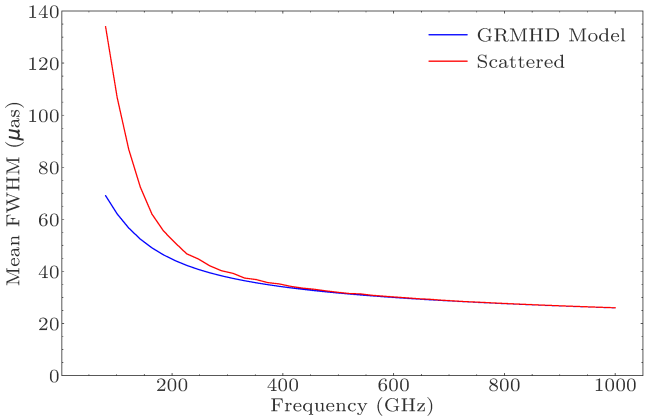

Fig. 4. Geometric mean FWHM of principal axes as a function of frequency for the ray-traced simulation of Sgr A* in Fig. 2 (Mościbrodzka & Gammie 2018). The blue curve shows size evolution for the simulation, the red curve shows size evolution for the simulation including the effects of interstellar scattering (Johnson 2016; Johnson et al. 2018). The sizes were obtained from measurements of the image second moment per frequency bin of 20 GHz. The change in size with increasing frequency becomes greatly reduced at frequencies above 300 GHz, where the size of the source is dominated by the achromatic black hole shadow and the Doppler boosted features (Falcke et al. 2000).

source extent is therefore dominated by the black hole shadow and Doppler-boosted features at higher frequencies (Falcke et al. 2000). This behavior is shown in Fig. 4 for the GRMHD simulation of the quiescent accretion flow of Sgr A* observed at frequencies from 80 GHz to 1 THz (Mościbrodzka & Gammie 2018). At frequencies of $\sim$300 GHz and above, the source size changes very little with increasing frequency. These achromatic properties motivate the extrapolation of a source size from lower-frequency observations with short baselines, such as the EHT at 230 GHz, to higher-frequency imaging experiments such as the upcoming EHT at 345 GHz (Event Horizon Telescope Collaboration et al. 2019b; Doeleman et al. 2019).

4. Demonstration

The second moment regularization can be used with informed size constraints from previous experiments, GRMHD simulations, or achromatic features from other observing frequencies. In this section, we demonstrate how the second moment regularization adds information to the imaging process if the data set to be imaged has no short baselines. For all following tests, we use a high $\beta_R = 10^5$, such that the input source size is strongly constrained in the imaging process. To put this value into perspective, $\beta_R = 10^5$ would cause a $\sim$10% difference between the input and reconstructed source sizes to be penalized equivalently to a change in reduced χ^2 of $\sim$1 in our imaging procedure. This regularization weight tends to drive the second moment of reconstructed images to be within 20% of the input values, therefore allowing some flexibility for the imaging process to deviate from the input second moment toward morphology favored by the available data.

In Sect. 4.1 we show improvements to the reconstructions when the source size is known. In Sect. 4.2 we study the image quality and fidelity dependence on the assumed size in the regularization. Finally in Sect. 4.3 we demon-

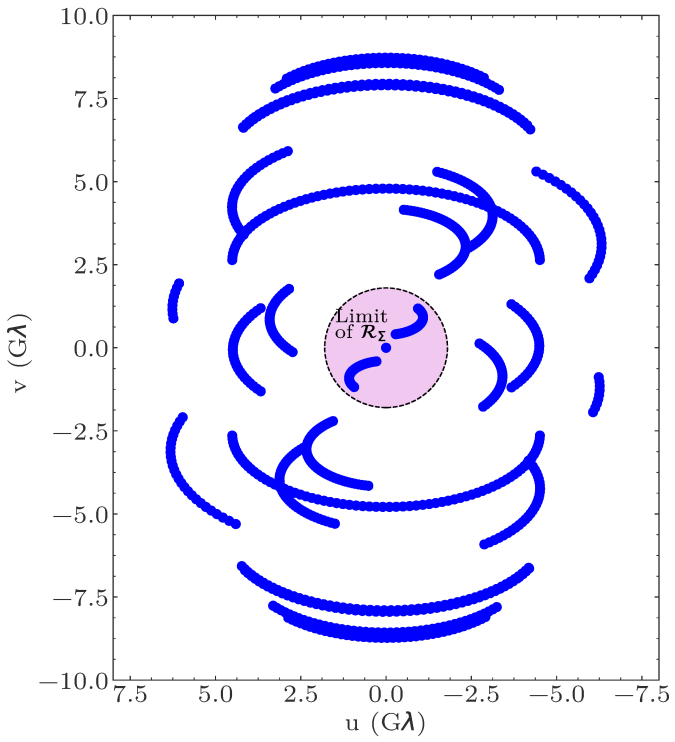

Fig. 5. (u, v) coverage for simulated observations of Sgr A* with the EHT 2017 array at 230 GHz. The magenta disk represents the range of (u, v) constrained by the second moment regularization, with the boundary at the $1/e$ point of the corresponding visibility amplitude function for Sgr A* assuming an isotropic source of 60 μas FWHM (Johnson et al. 2018).

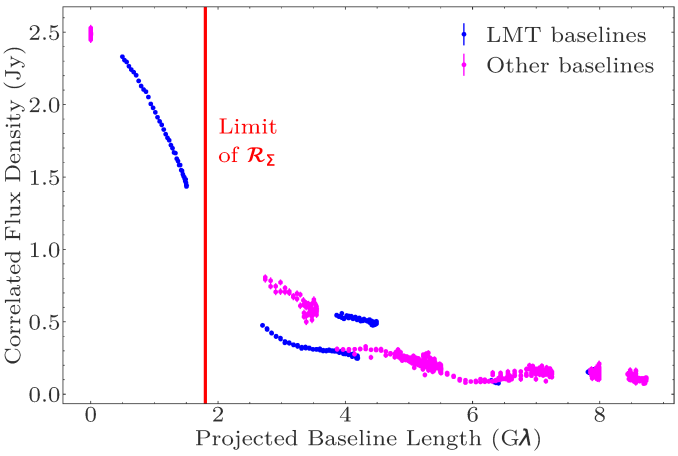

Fig. 6. Visibility amplitudes for a model image of a semi-analytic advection-dominated accretion flow (ADAF) model of Sgr A* (Broderick et al. 2011) with a FWHM of $\sim 60\,\mu$as as a function of (u, v) distance sampled by the EHT in 2017 with and without the LMT (affecting mid-range baselines). The regularizer $\mathcal{R}_\Sigma$ governs the visibility amplitude behavior at short baselines until the $1/e$ point. This allows us to constrain and correct limitations and uncertainties in LMT calibration based on the expected behavior of the LMT–SMT mid-range baseline.

strate that high fidelity images can be obtained without a priori knowledge of the source extent via input parameter searches.

4.1. Imaging with complementary size constraints

In Fig. 5, we illustrate the domain in which the second moment regularization ($\mathcal{R}_\Sigma$) operates. The (u, v) coverage is that of a typical observation of Sgr A* with the EHT at 230 GHz. Assuming a source extent of 60 μas from previous observations (Johnson et al. 2018), the $1/e$ boundary of the visibility function for a source with that characteristic size is shown as a disk on the (u, v) coverage. The only EHT baselines that lie within the $\mathcal{R}_\Sigma$ disk are intra-site baselines and the LMT–SMT short VLBI baseline. A single short VLBI baseline is very limited in constraining the overall extent of the source even assuming optimal performance of the telescopes.

We selected a ray-traced image of a semi-analytic advection-dominated accretion flow (ADAF) model of Sgr A* (Broderick et al. 2011) with a similar characteristic size to the Sgr A* observations to assess the performance of the regularizer and to test the robustness of the imaging process as a function of the input parameters $\theta_{\rm maj}$, $\theta_{\rm min}$, and ϕ. We sample the image with EHT 2017 coverage (Fig. 5), where we have total flux density estimates from intra-site baselines and a valuable mid-range baseline (SMT–LMT) describing the extent of the source on the sky, as shown in Fig. 6. We chose to discard all LMT baselines to limit the coverage and remove the constraining mid-range baseline for the regularizer tests. The extent of the source will then solely be enforced by the user-defined $\theta_{\rm maj}$, $\theta_{\rm min}$, and ϕ input parameters for $\mathcal{R}_\Sigma$ in the imaging process. It should be noted that imaging without the LMT not only removes short-baseline information on source extent but also long-baseline information on finer features, creating further differences in reconstructed images. The LMT, due to its size and central location, holds a strong weight in triggering decisions, while the SMT is a smaller and well-exercised station and is fairly flexible to various observing conditions. The choice to discard the LMT is thus mainly motivated by the known difficulties, to date, for the station to observe in a wide range of observing conditions and obtain adequate calibration information (Event Horizon Telescope Collaboration et al. 2019c,d). Removing the SMT instead, for the purposes of these tests, would give similar results due to the lack of short-baseline information.

In Fig. 7, we show the model crescent image in the left panel, and example reconstructions for four different scenarios in the right panel. The first scenario is a reconstruction of the full EHT observations of the crescent, using closure quantities and visibility amplitudes, and maximizing simple image entropy. In that case, we obtain a good fit to the visibility amplitudes, and we recover an image very similar to the model image. Then, we reconstruct the same observations constraining the image to match the true second moment, as measured on the true image. With this method, we obtain a marginally improved fit to the amplitudes, but visibly less diffuse flux outside the crescent due to the constraint of $\mathcal{R}_\Sigma$. Once we remove the LMT however, the simple imaging with maximum entropy is not able to reconstruct the morphology of the source, although some compact features are reconstructed that enable a decent fit to the visibility amplitudes. When adding $\mathcal{R}_\Sigma$ to the process, the second moment constraint is able to offset the absence of short baselines and reconstructs an image of improved quality in terms of both image morphology and goodness-of-fit to the amplitudes. This demonstration

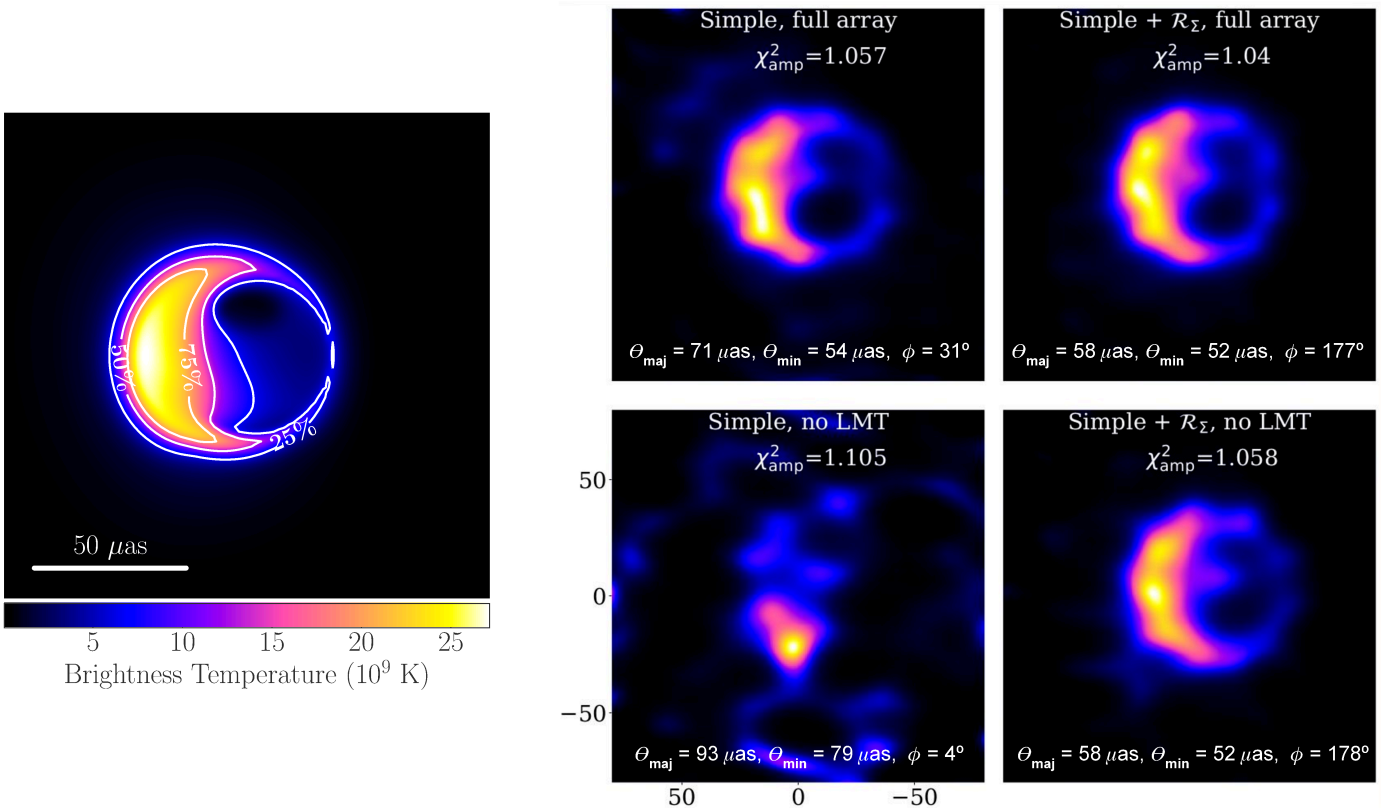

Fig. 7. *Left:* Model image of a semi-analytic ADAF model of Sgr A*(Broderick et al. 2011), contours of 25, 50, and 75% of the peak flux density are shown in white. *Right:* Tests of the second moment regularizer using the true image parameters as input ($\theta_{\rm maj} = 58\,\mu$as, $\theta_{\rm min} = 52\,\mu$as, $\phi = 177^\circ$ as measured directly from the model image), χ^2 values are calculated for the data set without the LMT. We additionally give the resulting source size parameters for each reconstruction. Imaging of the example data set with full EHT 2017 coverage shows little difference between the imaging process with and without the second moment regularizer. When the LMT is removed, and thus the mid-range baseline no longer constrains the source size, $\mathcal{R}_{\mathbf{\Sigma}}$ greatly improves the imaging. It should be noted that differences in finer features imaged with and without LMT are expected due to the loss of some long-baseline information from the removal of the LMT.

shows that $\mathcal{R}_{\mathbf{\Sigma}}$ successfully adds additional information to reconstruct a more physically plausible image even when mid-range baselines are lacking in the underlying data set. The improvement in the amplitude χ^2 also shows that $\mathcal{R}_{\mathbf{\Sigma}}$ is a useful tool to aid convergence in imaging.

4.2. Dependence of reconstructed images on assumed size

In the demonstration of $\mathcal{R}_{\mathbf{\Sigma}}$ we constrained the second moment to the true size of the source, to enable an accurate reconstruction of the image. However, in practice, the true size of the source is unknown, and is instead approximated from Gaussian model fitting to closure quantities and/or short-baseline visibility amplitudes and extrapolated from historical measurements. We therefore investigate the robustness of the image reconstructions when the input Gaussian parameters are strongly enforced in the imaging process, corresponding to a strong weight of the $\mathcal{R}_{\mathbf{\Sigma}}$ hyperparameter, while changing input principal axes FWHMs. We demonstrate this dependence by imaging the data set of the crescent model sampled by the EHT 2017 coverage without the LMT, such that the extent of the source is only enforced by the varying inputs to $\mathcal{R}_{\mathbf{\Sigma}}$. For simplicity, we use a single common imaging script varying only the input principal axes FWHMs. We assume an isotropic source size such that $\theta_{\rm maj} = \theta_{\rm min}$ and $\phi = 0^\circ$, and a range of input FWHMs of $5 - 90\,\mu$as.

We utilize two metrics to compare the quality of the reconstructed image to the true model image. The normalized root-mean-square error (NRMSE) of each image is given by:

$$\mathrm{NRMSE} = \sqrt{\frac{\sum_{\rm k}(I_{\rm i} - I'_{\rm i})^2}{\sum_{\rm k} I_{\rm i}^2}}\,, \tag{16}$$

where I' is the intensity of the reconstructed image and I is that of the true image (e.g., Chael et al. 2018). If the reconstructed image is identical to the true image, the NRMSE is zero. Therefore, the input FWHM for the reconstruction resulting in the minimum NRMSE in comparison to the true image gives the best fit.

The normalized cross-correlation (NXCORR) is a sliding inner-product of two normalized functions. For fast numerical computation, we determine the cross-correlation of the Fourier transforms of the normalized intensity patterns of the true image $I_{\rm norm}$ and the reconstructed image $I'_{\rm norm}$ at different relative shifts $\boldsymbol{\delta}$ across the extent of the images. For each pixel i in the image, we normalize the intensity via:

$$I_{\rm norm,i} = \frac{I_{\rm i} - \mu_I}{\sigma_I}\,, \tag{17}$$

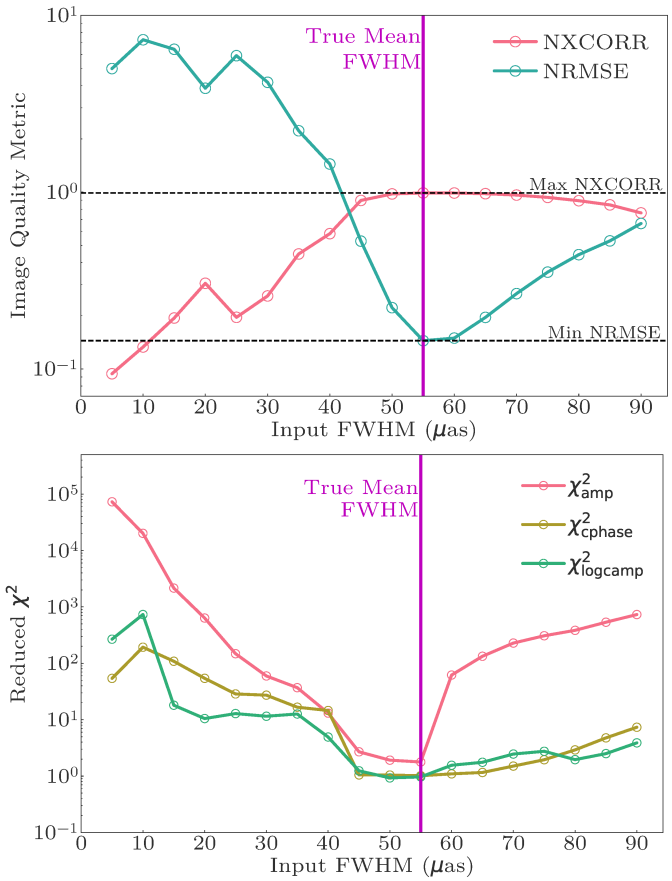

Fig. 8. Quality of the images obtained with different input FWHM (major and minor axes equal, position angle is zero). The image quality is measured in three ways: (1) the normalized cross-correlation against the true image, or NXCORR; (2) the normalized root-mean-square error against the true image, or NRMSE, shown in the top panel; and (3) reduced χ^2 goodness-of-fits to the three data products used in the reconstructions (visibility amplitudes, closure amplitudes and phases) shown in the bottom panel. NRMSE is more sensitive to subtle differences in the images than NXCORR due to the higher weight associated with large pixel-by-pixel errors and is minimized in a comparable range of input FWHMs to the reduced data χ^2. The narrow range of FWHMs encompasses the true mean source FWHM (magenta vertical line).

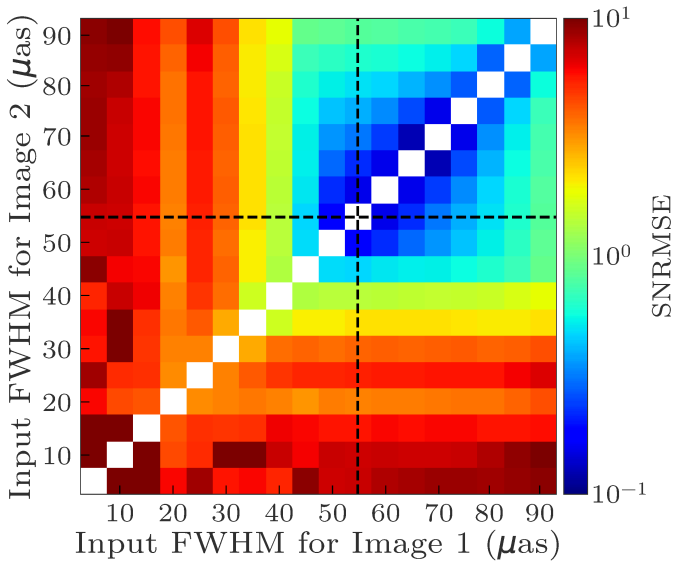

Fig. 9. Cross-comparisons of reconstructed images with varying isotropic input FWHMs using symmetrically normalized root-mean-square error (SNRMSE). The SNRMSE grid shows a greater correspondence of images with input FWHMs near the true mean FWHM of 55 μas, marked by the dashed black lines. The reconstructed images with varying input size (5–90 μas) are all compared to each other, where image 1 and image 2 are the two images to be compared (I'_1 and I'_2 respectively in Equation 19). The diagonal is each image compared to itself. The SNRMSE grid gives a range of plausible input FWHMs for $\mathcal{R}_\Sigma$ that result in high fidelity images even when the true source size is unknown.

where μ_I and σ_I are the mean and standard deviation of the intensity distribution in the image. The cross-correlation for a given shift $\boldsymbol{\delta}$ is then given by:

$$\mathrm{NXCORR}(\boldsymbol{\delta}) = |\mathcal{F}^{-1}\{\mathcal{F}\{I^*_{\mathrm{norm}}(\mathbf{x})\} \cdot \mathcal{F}\{I'_{\mathrm{norm}}(\mathbf{x}+\boldsymbol{\delta})\}\}|. \quad (18)$$

The shift at which the cross-correlation is maximized is then used to output the final NXCORR value for the two images. This method is less sensitive to individual features in the reconstructed image than NRMSE as it compares the bulks of each intensity pattern as opposed to the NRMSE pixel-to-pixel comparison. The χ^2 statistics follow the equations presented in Sect. 2.1 of Event Horizon Telescope Collaboration et al. (2019d).

In Fig. 8, we show the NRMSE and NXCORR metric scores for the reconstructed images compared against the true image (left panel of Fig. 7), and the reduced data χ^2 goodness-of-fits to the imaged data set (Fig. 6, no LMT). The NXCORR is maximized at an input FWHM of 55 μas, and the NRMSE is minimized at the same input FWHM. This value corresponds to the mean FWHM (average of θ_{maj}= 58 μas and θ_{min}= 52 μas) of the true image. With this test, we find an excellent correspondence between the reconstructed image with the highest quality (highest NXCORR, lowest NRMSE, and lowest reduced data χ^2) and the image with the input $\mathcal{R}_\Sigma$ FWHM closest to the true value. Images with input FWHMs close to the optimal value are of similarly good quality. We thus show a good performance of $\mathcal{R}_\Sigma$ in the imaging process even with input sizes inaccurate to within 20% of the true size. The reduction in data χ^2 values as we approach the true source size also indicates that $\mathcal{R}_\Sigma$ gives a convergence boost toward a higher fidelity image. This behavior is caused by $\mathcal{R}_\Sigma$ rapidly reducing the favored set of images to only those that constrain flux within a given region. The region limits that best represent the flux distribution in the true image allow the minimizing process to focus more quickly on the data terms and achieve better reduced χ^2 values within the given imaging conditions. This property also allows us to survey the response of the imaging process and goodness-of-fits to the available data via parameter searches over different favored second moments (and thus favored flux regions) and determine optimal parameters that best represent properties of the data set.

4.3. Imaging without complementary size constraints

The NRMSE metric proves to be more sensitive to differences in the image structure than NXCORR, as shown in

Fig. 8, due to the higher weight associated with large errors in the computation of the NRMSE. For that reason, we have selected NRMSE to score comparisons between the reconstructed images themselves. For this test, we assume that the true image and true FWHM are unknown, as is the case for real experiments. We instead focus on the morphological characteristics that appear in the images based on the underlying data, and how the inputs to $\mathcal{R}_{\mathbf{\Sigma}}$ affect the correspondence between reconstructed images. We restructure the metric into a symmetrically-normalized root-mean-square error (SNRMSE; Hanna et al. 1985; Mentaschi et al. 2013) to render the NRMSE independent of the input and comparison image choice:

$$\mathrm{SNRMSE} = \sqrt{\frac{\sum_{\mathrm{k}}(I'_{1,i} - I'_{2,i})^2}{\sum_{\mathrm{k}} I'_{1,i} I'_{2,i}}}. \quad (19)$$

Here I'_1 and I'_2 are the two reconstructed images to be compared. In Fig. 9, we show an SNRMSE grid comparing each reconstructed image to all others, where the diagonal squares correspond to each image compared with itself. We have marked with dashed lines where the mean FWHM of the true image lies. We find that images with input FWHMs near the true FWHM of the source have a better SNRMSE with each other than all other combinations of images. This test enables the user to find a range of characteristic sizes minimizing SNRMSE via a size parameter search. For compact sources that are distinctly elliptical, a one-dimensional size parameter search is useful to quickly sweep through a wide range of sizes and determine a range of plausible sizes for the source extent. A search within that range, varying parameters in two dimensions (θ_{maj}, θ_{min}, and ϕ), can then be carried out to refine the source size estimate for the imaging process.

We find that the use of the regularizer improves the quality of the resulting image even if the input parameters deviate by 20% from the true values. We also find that the strong use of the regularization, when combined with a size parameter search, is able to converge toward the true FWHM values, even when the true source dimensions are unknown. The use of SNRMSE and χ^2 statistics serve well to score individual images and parameters without a priori knowledge of the source extent.

5. Applications

In addition to simple static imaging, second moment regularization can easily be coupled to more sophisticated and complex imaging techniques. In Sect. 5.1 we present an example of the use of second moment regularization for scattering mitigation imaging of Sgr A* at longer wavelengths. In Sect. 5.2 we demonstrate how second moment regularization in individual sparse snapshots improves the quality of dynamical reconstructions of variable sources, such as a movie of an orbiting "hot spot" in Sgr A*'s accretion flow.

5.1. Scattering mitigation

The second moment constraint in imaging can both be used for data sets where short baselines are lacking, as demonstrated in Sect. 4, and for data sets where short-baseline measurements have large uncertainties due to difficult observing conditions. An example of the latter case is presented in Issaoun et al. (2019), where observations of Sgr A* at 86 GHz with the Global Millimeter VLBI Array and ALMA (project code MB007) yielded high signal-to-noise (SNR) detections on long baselines but bad weather at select Very Long Baseline Array (VLBA) stations led to poorly constrained short-baseline measurements. Imaging of the source with RML would not have been feasible with these measurements alone, as the large uncertainties in the short-baseline measurements caused flux to spread nonphysically across the reconstructed images. Since the size of Sgr A* on the sky is well studied and known to be affected by anisotropic scatter-broadening from the interstellar medium (Davies et al. 1976; van Langevelde et al. 1992; Frail et al. 1994; Bower et al. 2004; Shen et al. 2005; Bower et al. 2006; Psaltis et al. 2018; Johnson et al. 2018), previous size measurements (Ortiz-León et al. 2016; Brinkerink et al. 2019) were used to constrain the extent of Sgr A* in the imaging process with $\mathcal{R}_{\mathbf{\Sigma}}$. In this manner, we obtained an image that was able to fit new long-baseline detections to ALMA, likely refractive noise from scattering substructure.

The second moment regularization was also implemented in the scattering mitigation code *stochastic optics* developed by Johnson (2016). Stochastic optics aims to mitigate the effects of scattering to derive an intrinsic (unscattered) image of the source. The code solves for the unscattered image by separating and mitigating the two main components of the Sgr A* scattering screen: the diffractive scattering that causes the image to become a convolution of the true image and the scattering kernel; and the refractive scattering that introduces stochastic ripples that further distort the image. The stochastic optics framework therefore simultaneously solves for the unscattered image and the scattering screen assuming a given model for the diffractive blurring kernel and the time-averaged refractive properties. The model assumed here is the Johnson et al. (2018) scattering model, the best-fitting model to Sgr A* observations to date (Issaoun et al. 2019).

The implementation of $\mathcal{R}_{\mathbf{\Sigma}}$ in stochastic optics only constrains the size of the scattered source (Sgr A* as we see it on the sky) based on historical measurements from model fitting, such that the technique can more accurately mitigate the effects of interstellar scattering to obtain a physically motivated intrinsic image of the accretion flow of Sgr A* (for further details, see Issaoun et al. 2019). The intrinsic image itself is not directly constrained by the second moment regularization, but is derived from the combination of the constrained scattered image and knowledge of the interstellar scattering.

Here we illustrate the use of $\mathcal{R}_{\mathbf{\Sigma}}$ within stochastic optics using a lower frequency data set. Observations of Sgr A* at 22 GHz with the VLBA+GBT (project code BG221A) showed clear long-baseline detections of refractive noise from interstellar scattering (Gwinn et al. 2014; Johnson et al. 2018). These long-baseline detections should translate to substructure in the image, distorting the intensity pattern seen for Sgr A* away from the scatter-broadened smooth elongated Gaussian-like morphology. While the scattering substructure is very apparent in the data set, it is a non-trivial task to successfully show its effects on the image itself and obtain an intrinsic image of the source. This is due to the imaging process being driven predominantly by the abundance of intra-VLBA short-baseline measurements in comparison to the few VLBA–GBT long-baseline detections. We therefore test the addition of $\mathcal{R}_{\mathbf{\Sigma}}$ on this

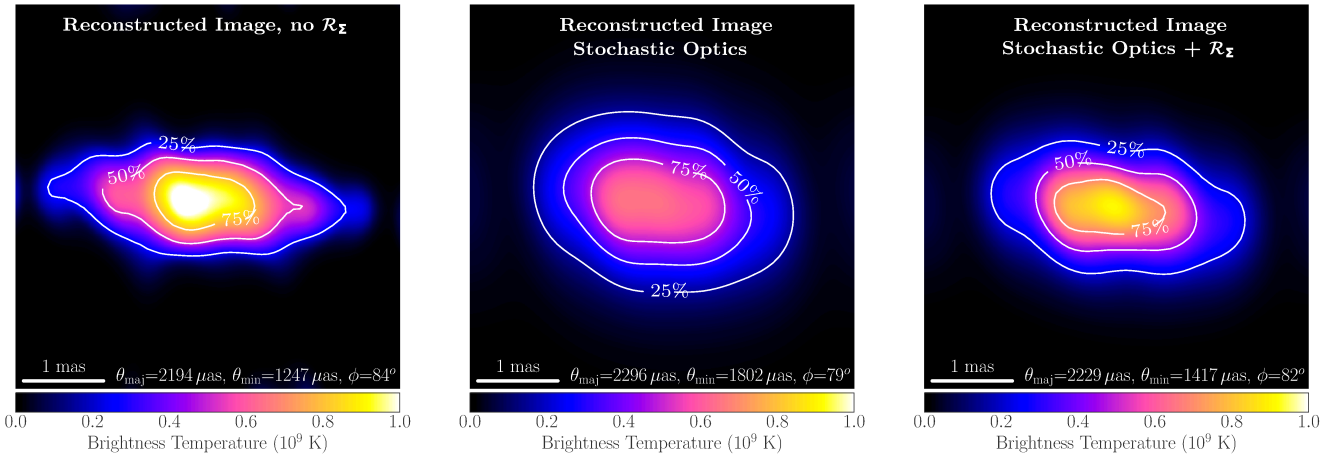

Fig. 10. Reconstructions of 22 GHz VLBA+GBT observations and their resulting source extents. MK and SC have no detections, and HN and NL are flagged due to their very low sensitivity in this experiment. *Left*: a simple reconstruction of the scattered image without $\mathcal{R}_\Sigma$. *Center*: a reconstruction of the scattered image via stochastic optics (Johnson 2016), using the scattering model by Johnson et al. 2018. *Right*: a reconstruction with stochastic optics, using $\mathcal{R}_\Sigma$ and the input source size as determined by Johnson et al. 2018 from high-precision Gaussian model fitting: $\theta_{\rm maj} = 2255 \pm 61\,\mu$as, $\theta_{\rm min} = 1243 \pm 39\,\mu$as, $\phi = 81.9 \pm 0.2°$. The reconstruction with $\mathcal{R}_\Sigma$ helps constrain the extent of the source in the north–south direction, where measurements are lacking due to the predominantly east–west configuration of the VLBA+GBT.

data set, using the source dimensions in Table 1 of Johnson et al. (2018) from elliptical Gaussian model fitting.

In Fig. 10, we show three separate reconstructions of the 22 GHz data set. A standard RML reconstruction of the data set (Fig. 10 left panel) shows some distortions in the scattered image, but the morphology remains fairly smooth and elongated. Standard RML imaging cannot solve for the scattering properties, therefore the procedure is solely focused on obtaining the highest fidelity scattered image possible from the data set. We will thus treat this image as our comparison image for this data set. When using stochastic optics however, the imaging process is more complex, as it is simultaneously imaging the scattered source and solving for the scattering properties to disentangle scattering from intrinsic source structure. This process derives a scattered image that is not well-constrained in the north–south direction due to the configuration of the VLBA+GBT, resulting in a large source image that is not fully converged to the image obtained from standard RML (Fig. 10 center panel). Since the scattered image does not match our expectations of the physical morphology of the source, the derived intrinsic image should also not be trusted. The challenge is then to improve the convergence of the imaging component of stochastic optics to quickly obtain a physically motivated scattered image and therefore undergo a higher-fidelity separation of the scattering and intrinsic structure. When using $\mathcal{R}_\Sigma$, where the scattered image is constrained to remain within the size obtained by Johnson et al. (2018) using elliptical Gaussian model fitting, the resulting scattered image is more elongated in the east–west direction (Fig. 10 right panel) and showing distortions similar to those of the standard RML reconstruction This shows that the use of $\mathcal{R}_\Sigma$ helps the convergence of the scattered image through stochastic optics to a more physically motivated reconstruction, and thus will give a more realistic underlying unscattered image of the source.

5.2. Dynamical imaging

There are additional applications for the second moment regularization in movie reconstructions of variable sources where single snapshots have very sparse coverage. We can test the robustness of movie reconstructions with the loss of short baselines using a simulated flare (model B of Doeleman et al. 2009) with an orbiting period of 27 minutes around the same crescent model as in Sect. 4. We reconstruct movies of the orbiting "hot spot" using *dynamical imaging*, enforcing temporal continuity between individual frames (for further details, see Johnson et al. 2017). We reconstruct a movie of the orbiting hot spot for four different scenarios: (1) we use the EHT 2017 array without the LMT, no short baselines are present in the individual snapshots to constrain the source extent; (2) we use the data set without the LMT, but constrain the extent of the source (the dimensions of the crescent model) with $\mathcal{R}_\Sigma$, (3) we use the full EHT 2017 to reconstruct the orbit; (4) we use the full EHT 2017 and $\mathcal{R}_\Sigma$ to reconstruct the orbit. In Fig. 11, we show individual frames of the true simulation and of the reconstructed movies for the four different scenarios. The reconstructions without $\mathcal{R}_\Sigma$ either yield unphysical source structure dominated by the dirty image (due to the lack of information without LMT) or contain imaging artifacts from flux spreading due to the sparse coverage of individual snapshots. In particular, even with the full EHT2017 array, dynamical imaging without $\mathcal{R}_\Sigma$ shows north–east and south–west artifacts from the dirty image that persist due to the sparse snapshot coverage. The reconstructions with $\mathcal{R}_\Sigma$, even without the LMT, are significantly cleaner and more accurately reconstruct the motion and morphology of the simulation, as shown by NXCORR results when compared to the truth simulated images.

6. Summary

In summary, we have developed a regularization function $\mathcal{R}_\Sigma$, for use in a regularized maximum likelihood framework for interferometric imaging, that constrains the spread

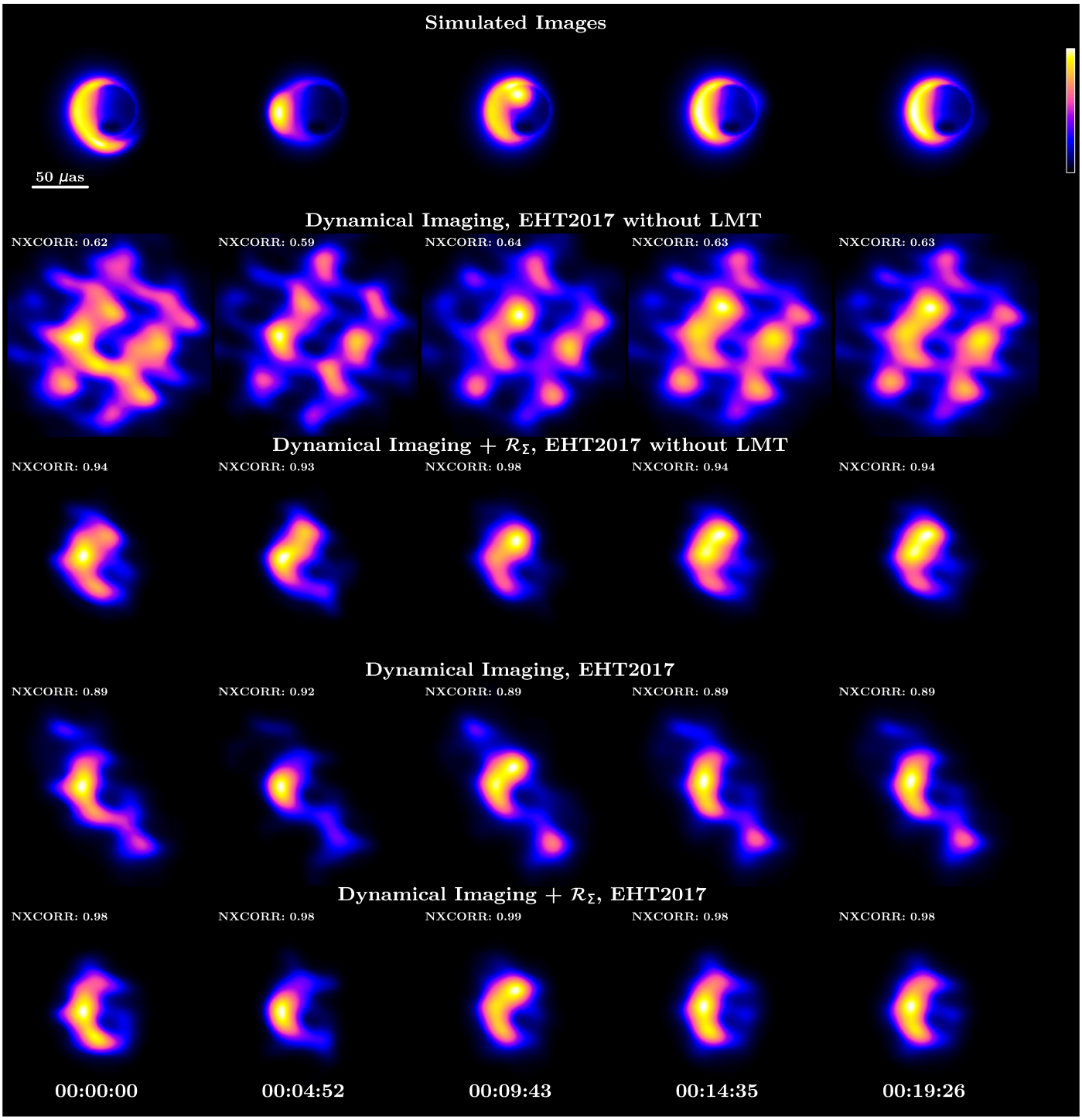

Fig. 11. Reconstruction of a simulated flare using dynamical imaging (Johnson et al. 2017). From top to bottom: simulated images of a flare with a period of 27 minutes (model B of Doeleman et al. 2009); simple dynamical imaging without the LMT (no short-baseline points constraining the source extent); dynamical imaging using $\mathcal{R}_{\mathbf{\Sigma}}$ without the LMT (the second moment regularization offsets the lack of short-baseline constraints); simple dynamical imaging with full EHT2017 sampling; dynamical imaging using $\mathcal{R}_{\mathbf{\Sigma}}$ with EHT2017 sampling. Using $\mathcal{R}_{\mathbf{\Sigma}}$ significantly improved the quality of dynamical reconstructions both with the full array and without the LMT. NXCORR values against the model images are shown in the top left corner for each reconstructed snapshot. The variations in the resulting FWHMs of the reconstructed images are visually evident.

of flux in reconstructed images to match input parameters defined by the user. The second moment regularization is a natural extension of common imaging tools, such as image total flux and image centroid constraints (zeroth and first moment respectively), that help to mitigate the missing information problem in high frequency VLBI. The regularization assumes that the source is compact, with a stable size, and is resolved on longer baselines of the interferometer. The validity of these assumptions for the EHT's primary targets, Sgr A* and M87, are well-motivated by state-of-the-

art GRMHD simulations and long-term observational studies. For well-studied sources, this method allows for contingency against weather, a major deterrent for high frequency VLBI, and gives more flexibility for triggering decisions if key short baselines yield poorly constrained measurements or become unavailable during or between observations.

We have shown that $\mathcal{R}_{\mathbf{\Sigma}}$ successfully informs the source behavior on short baselines and is defined only by three Gaussian parameters and the regularization hyperparameter. Imaging with $\mathcal{R}_{\mathbf{\Sigma}}$ is able to reconstruct high fidelity images fitting to the data products even if the input source dimensions deviate from the true values by up to 20%. The regularization therefore gives a larger flexibility than needed to account for changes in size from, for example, GRMHD simulations of highly variable sources such as Sgr A*. We have also shown that parameter searches over a range of isotropic FWHMs using $\mathcal{R}_{\mathbf{\Sigma}}$ in conjunction with goodness-of-fit statistics to data products and symmetrically-normalized root-mean-square error of image comparisons help determine high-fidelity source extent even if the exact size and morphology are unknown.

The regularization can be used to image with any choice of data products and any choice of feature-driven regularizers within the framework of the `eht-imaging` library (Chael et al. 2016, 2018) and is easily transferable to other tools or other RML imaging packages (e.g., `SMILI`; Akiyama et al. 2017a,b). We have shown that the $\mathcal{R}_{\mathbf{\Sigma}}$ implementation complements other techniques tackling source properties that add difficulty and complexity to the imaging process, such as time variability (via dynamical imaging; Johnson et al. 2017; Bouman et al. 2018) and interstellar scattering (Johnson 2016; Issaoun et al. 2019). Source parameter inputs can either be obtained from model fitting to abundant short-baseline measurements, historical measurements from observations with short baselines present, extrapolated from other wavelengths based on achromatic features, or estimated via parameter searches. The second moment regularization could prove particularly useful in future work with the EHT, both for dynamical reconstructions of variable sources such as Sgr A* and for upcoming imaging observations at 345 GHz (Event Horizon Telescope Collaboration et al. 2019b; Doeleman et al. 2019).

Acknowledgements. We thank John Wardle for his helpful comments and careful review. This work is supported by the ERC Synergy Grant "BlackHoleCam: Imaging the Event Horizon of Black Holes", Grant 610058. We thank the National Science Foundation (AST-1440254, AST-1716536, AST-1312651) and the Gordon and Betty Moore Foundation (GBMF-5278) for financial support of this work. This work was supported in part by the Black Hole Initiative at Harvard University, which is supported by a grant from the John Templeton Foundation.

References

Akiyama, K., Ikeda, S., Pleau, M., et al. 2017a, AJ, 153, 159
Akiyama, K., Kuramochi, K., Ikeda, S., et al. 2017b, ApJ, 838, 1
Akiyama, K., Lu, R.-S., Fish, V. L., et al. 2015, ApJ, 807, 150
Bouman, K. L., Johnson, M. D., Dalca, A. V., et al. 2018, IEEE Transactions on Computational Imaging, 4, 512
Bouman, K. L., Johnson, M. D., Zoran, D., et al. 2016, in IEEE (CVPR)
Bower, G. C., Falcke, H., Herrnstein, R. M., et al. 2004, Science, 304, 704
Bower, G. C., Goss, W. M., Falcke, H., Backer, D. C., & Lithwick, Y. 2006, ApJL, 648, L127
Brinkerink, C. D., Müller, C., Falcke, H. D., et al. 2019, A&A, 621, A119
Broderick, A. E., Fish, V. L., Doeleman, S. S., & Loeb, A. 2011, ApJ, 735, 110
Burn, B. J. & Conway, R. G. 1976, MNRAS, 175, 461
Chael, A. A., Johnson, M. D., Bouman, K. L., et al. 2018, ApJ, 857, 23
Chael, A. A., Johnson, M. D., Narayan, R., et al. 2016, ApJ, 829, 11
Davies, R. D., Walsh, D., & Booth, R. S. 1976, MNRAS, 177, 319
Doeleman, S., Akiyama, K., Blackburn, L., et al. 2019, in BAAS, Vol. 51, 537
Doeleman, S. S., Fish, V. L., Broderick, A. E., Loeb, A., & Rogers, A. E. E. 2009, ApJ, 695, 59
Doeleman, S. S., Fish, V. L., Schenck, D. E., et al. 2012, Science, 338, 355
Doeleman, S. S., Shen, Z.-Q., Rogers, A. E. E., et al. 2001, AJ, 121, 2610
Doeleman, S. S., Weintroub, J., Rogers, A. E. E., et al. 2008, Nature, 455, 78
Event Horizon Telescope Collaboration, Akiyama, K., Alberdi, A., et al. 2019a, ApJL, 875, L1
Event Horizon Telescope Collaboration, Akiyama, K., Alberdi, A., et al. 2019b, ApJL, 875, L2
Event Horizon Telescope Collaboration, Akiyama, K., Alberdi, A., et al. 2019c, ApJL, 875, L3
Event Horizon Telescope Collaboration, Akiyama, K., Alberdi, A., et al. 2019d, ApJL, 875, L4
Event Horizon Telescope Collaboration, Akiyama, K., Alberdi, A., et al. 2019e, ApJL, 875, L5
Event Horizon Telescope Collaboration, Akiyama, K., Alberdi, A., et al. 2019f, ApJL, 875, L6
Falcke, H., Melia, F., & Agol, E. 2000, ApJL, 528, L13
Fish, V. L., Doeleman, S. S., Beaudoin, C., et al. 2011, ApJL, 727, L36
Fish, V. L., Johnson, M. D., Doeleman, S. S., et al. 2016, ApJ, 820, 90
Frail, D. A., Diamond, P. J., Cordes, J. M., & van Langevelde, H. J. 1994, ApJ, 427, L43
Gwinn, C. R., Kovalev, Y. Y., Johnson, M. D., & Soglasnov, V. A. 2014, ApJL, 794, L14
Hanna, S., Heinold, D., Health, A. P. I., Dept, E. A., & Environmental Research & Technology, I. 1985, Development and application of a simple method for evaluating air quality models, API publication (American Petroleum Institute)
Honma, M., Akiyama, K., Uemura, M., & Ikeda, S. 2014, PASJ, 66, 95
Hu, M.-K. 1962, Information Theory, IRE Transactions on, 8, 179
Issaoun, S., Johnson, M. D., Blackburn, L., et al. 2019, ApJ, 871, 30
Johnson, M. D. 2016, ApJ, 833, 74
Johnson, M. D., Bouman, K. L., Blackburn, L., et al. 2017, ApJ, 850, 172
Johnson, M. D., Fish, V. L., Doeleman, S. S., et al. 2015, Science, 350, 1242
Johnson, M. D., Narayan, R., Psaltis, D., et al. 2018, ApJ, 865, 104
Krichbaum, T. P., Graham, D. A., Witzel, A., et al. 1998, A&A, 335, L106
Kuramochi, K., Akiyama, K., Ikeda, S., et al. 2018, ApJ, 858, 56
Lu, R.-S., Krichbaum, T. P., Eckart, A., et al. 2011, A&A, 525, A76
Lu, R.-S., Krichbaum, T. P., Roy, A. L., et al. 2018, ApJ, 859, 60
Mentaschi, L., Besio, G., Cassola, F., & Mazzino, A. 2013, Ocean Modelling, 72, 53
Moffet, A. T. 1962, ApJS, 7, 93
Mościbrodzka, M. & Gammie, C. F. 2018, MNRAS, 475, 43
Narayan, R. & Nityananda, R. 1986, ARA&A, 24, 127
Ortiz-León, G. N., Johnson, M. D., Doeleman, S. S., et al. 2016, ApJ, 824, 40
Psaltis, D. et al. 2018, ApJ, submitted
Rogers, A. E. E., Doeleman, S., Wright, M. C. H., et al. 1994, ApJL, 434, L59
Shen, Z.-Q., Lo, K. Y., Liang, M.-C., Ho, P. T. P., & Zhao, J.-H. 2005, Nature, 438, 62
Thompson, A. R., Moran, J. M., & Swenson, Jr., G. W. 2017, Interferometry and Synthesis in Radio Astronomy, 3rd Edition (Springer International Publishing)
van Cittert, P. H. 1934, Physica, 1, 201
van Langevelde, H. J., Frail, D. A., Cordes, J. M., & Diamond, P. J. 1992, ApJ, 396, 686
Wiaux, Y., Jacques, L., Puy, G., Scaife, A. M. M., & Vandergheynst, P. 2009a, MNRAS, 395, 1733
Wiaux, Y., Puy, G., Boursier, Y., & Vandergheynst, P. 2009b, MNRAS, 400, 1029
Zernike, F. 1938, Physica, 5, 785

Appendix A: Properties of the visibility function

Appendix A.1: Visibility derivatives and image moments

Non-astrometric VLBI experiments such as the EHT measure visibility amplitudes directly but do not provide absolute phase information. Nevertheless, the zeroth and second image moments are determined from visibility amplitudes alone (i.e., they do not depend on the measured phase; Moffet 1962; Burn & Conway 1976). For instance, the total flux density $\int I(\mathbf{x})d^2\mathbf{x} = V(\mathbf{0}) = |V(\mathbf{0})|$ because the zero-baseline visibility is real and positive, and therefore equal to its modulus.

More generally, we can express the visibility function as a Taylor expansion of its derivatives:

$$V(\mathbf{u}) = \int d^2\mathbf{x}\; I(\mathbf{x}) \left[1 - 2i\pi\mathbf{u}\cdot\mathbf{x} - \frac{(2\pi\mathbf{u}\cdot\mathbf{x})^2}{2} + \frac{i(2\pi\mathbf{u}\cdot\mathbf{x})^3}{6} + \frac{(2\pi\mathbf{u}\cdot\mathbf{x})^4}{24} + \cdots\right]. \tag{A.1}$$

The visibility amplitude function is image-translation invariant. To obtain a Taylor expansion for visibility amplitudes, we choose the image centroid to be at the origin. The first derivative of the visibility function (thus the second term of the Taylor expansion) then vanishes, giving

$$\begin{aligned} V(\mathbf{u}) &\simeq \int d^2\mathbf{x}\; I(\mathbf{x}) \left[1 - \frac{(2\pi\mathbf{u}\cdot\mathbf{x})^2}{2}\right] \\ &\simeq V(\mathbf{0}) - 2\pi^2 \int d^2\mathbf{x}\; (\mathbf{u}\cdot\mathbf{x})^2 I(\mathbf{x}). \end{aligned} \tag{A.2}$$

On short baselines (i.e., those with $\mathbf{u}\cdot\mathbf{x} \ll 1$), the visibility function is then positive and real, so $|V(\mathbf{u})| \simeq V(\mathbf{u})$. Since $\mathbf{u} = \begin{pmatrix} u \\ v \end{pmatrix}$ and $\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$, we can expand the inner product of the two vectors:

$$\begin{aligned} (\mathbf{u}\cdot\mathbf{x})^2 &= u^2x^2 + v^2y^2 + 2uvxy \\ &= \begin{pmatrix} u & v \end{pmatrix} \begin{pmatrix} x^2 & xy \\ xy & y^2 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}. \end{aligned} \tag{A.3}$$

Combining these results with the definition of the covariance matrix $\mathbf{\Sigma}$ (see Appendix A.2), we obtain:

$$\begin{aligned} |V(\mathbf{u})| &\simeq V(\mathbf{0}) - 2\pi^2 \int d^2\mathbf{x}\, (\mathbf{u}\cdot\mathbf{x})^2 I(\mathbf{x}) \\ &\simeq V(\mathbf{0}) - 2\pi^2 \begin{pmatrix} u & v \end{pmatrix} \int d^2\mathbf{x}\, I(\mathbf{x}) \begin{pmatrix} x^2 & xy \\ xy & y^2 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \\ &\simeq V(\mathbf{0}) - 2\pi^2 V(\mathbf{0}) \mathbf{u}^\intercal \mathbf{\Sigma} \mathbf{u}. \end{aligned} \tag{A.4}$$

The downward curvature of the amplitude function at zero baseline is thus related to the image covariance by:

$$\nabla\nabla^\intercal |V(\mathbf{u})|\big|_{\mathbf{u}=\mathbf{0}} = \nabla\nabla^\intercal V(\mathbf{u})\big|_{\mathbf{u}=\mathbf{0}} = -4\pi^2 V(\mathbf{0})\mathbf{\Sigma}. \tag{A.5}$$

Appendix A.2: Image principal axes and visibility curvature

From Equation A.5, the curvature of the visibility function on short baselines is proportional to the second central moment of the image projected along the baseline direction. The second central moment of the image is naturally expressed as a covariance matrix:

$$\mathbf{\Sigma} \equiv \frac{\int d^2\mathbf{x} I(\mathbf{x})(\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})^\intercal}{\int d^2\mathbf{x}\, I(\mathbf{x})} = \begin{pmatrix} \Sigma_{\rm xx} & \Sigma_{\rm xy} \\ \Sigma_{\rm yx} & \Sigma_{\rm yy} \end{pmatrix}, \tag{A.6}$$

$$\Sigma_{\rm xx} = \frac{\int d^2\mathbf{x}\, I(\mathbf{x})(x-\bar{x})^2}{\int d^2\mathbf{x}\, I(\mathbf{x})},$$

$$\Sigma_{\rm yy} = \frac{\int d^2\mathbf{x}\, I(\mathbf{x})(y-\bar{y})^2}{\int d^2\mathbf{x}\, I(\mathbf{x})},$$

$$\Sigma_{\rm xy} = \frac{\int d^2\mathbf{x}\, I(\mathbf{x})(x-\bar{x})(y-\bar{y})}{\int d^2\mathbf{x}\, I(\mathbf{x})} = \Sigma_{\rm yx}.$$

To put the covariance matrix in a more intuitive form, we express it in terms of its principal axes. The image covariance matrix has two eigenvalues, and can be diagonalized as follows:

$$\mathbf{\Sigma} = \mathbf{R}_\phi \begin{pmatrix} \lambda_{\rm min} & 0 \\ 0 & \lambda_{\rm maj} \end{pmatrix} \mathbf{R}_\phi^\intercal, \tag{A.7}$$

where the rotation matrix $\mathbf{R}_\phi$, based on the position angle ϕ (East of North) of the major principal axis, is given by:

$$\mathbf{R}_\phi = \begin{pmatrix} \cos(\phi) & \sin(\phi) \\ -\sin(\phi) & \cos(\phi) \end{pmatrix}. \tag{A.8}$$

The eigenvalues are derived from the quadratic equation:

$$\lambda_{\rm maj} = \frac{\Sigma_{\rm xx}+\Sigma_{\rm yy}}{2} + \frac{\sqrt{4(\Sigma_{\rm xy})^2 + (\Sigma_{\rm xx}-\Sigma_{\rm yy})^2}}{2}, \tag{A.9}$$

$$\lambda_{\rm min} = \frac{\Sigma_{\rm xx}+\Sigma_{\rm yy}}{2} - \frac{\sqrt{4(\Sigma_{\rm xy})^2 + (\Sigma_{\rm xx}-\Sigma_{\rm yy})^2}}{2}. \tag{A.10}$$

We can also express each term of the covariance matrix in terms of the eigenvalues and position angle ϕ:

$$\Sigma_{\rm xx} = \cos^2(\phi)\lambda_{\rm min} + \sin^2(\phi)\lambda_{\rm maj}, \tag{A.11}$$

$$\Sigma_{\rm yy} = \sin^2(\phi)\lambda_{\rm min} + \cos^2(\phi)\lambda_{\rm maj}, \tag{A.12}$$

$$\Sigma_{\rm xy} = (\lambda_{\rm maj} - \lambda_{\rm min})\cos(\phi)\sin(\phi). \tag{A.13}$$

The eigenvalues of the covariance matrix are the variances along the principal axes (major and minor axes).

Appendix B: Implementation via gradient descent

Appendix B.1: Pixel-based derivation of principal axes

Here we present the computation of the covariance matrix for the pixel-based reconstructions from RML. The centroid of an $n \times n$ pixel-based image is given by the following parameters:

$$\bar{x} = \frac{\sum\limits_{\rm k} x_{\rm i} I_{\rm i}}{\sum\limits_{\rm k} I_{\rm i}} \text{ and } \bar{y} = \frac{\sum\limits_{\rm k} y_{\rm i} I_{\rm i}}{\sum\limits_{\rm k} I_{\rm i}}, \tag{B.1}$$

where i is the pixel number (from 1 to k), $I_{\rm i}$ is the intensity at that pixel, $x_{\rm i}$ is the x-position and $y_{\rm i}$ is the y-position of

the pixel in the image. The second moment of the image is given by the covariance matrix

$$\mathbf{\Sigma} = \begin{pmatrix} \Sigma_{\rm xx} & \Sigma_{\rm xy} \\ \Sigma_{\rm xy} & \Sigma_{\rm yy} \end{pmatrix}, \tag{B.2}$$

where (B.3)

$$\Sigma_{\rm xx} = \frac{\sum\limits_{\rm k}(x_{\rm i}-\bar{x})^2 I_{\rm i}}{\sum\limits_{\rm k} I_{\rm i}}, \tag{B.4}$$

$$\Sigma_{\rm yy} = \frac{\sum\limits_{\rm k}(y_{\rm i}-\bar{y})^2 I_{\rm i}}{\sum\limits_{\rm k} I_{\rm i}}, \tag{B.5}$$

$$\Sigma_{\rm xy} = \frac{\sum\limits_{\rm k}(x_{\rm i}-\bar{x})(y_{\rm i}-\bar{y}) I_{\rm i}}{\sum\limits_{\rm k} I_{\rm i}}. \tag{B.6}$$

As in Appendix A.2, the image covariance matrix has two eigenvalues and can be diagonalized to obtain the principal axes FWHMs.

Appendix B.2: Gradient Descent Implementation

We have defined our regularization function via the Frobenius norm:

$$\mathcal{R}_{\mathbf{\Sigma}} = (\Sigma_{\rm xx}-\Sigma'_{\rm xx})^2 + (\Sigma_{\rm yy}-\Sigma'_{\rm yy})^2 + 2(\Sigma_{\rm xy}-\Sigma'_{\rm xy})^2. \tag{B.7}$$

Within the framework of the `eht-imaging` library, the objective function is minimized via gradient descent. Therefore, the regularization functions must also individually be minimized via gradient descent. The gradients for the quantities describing the properties of the image introduced thus far, for a given pixel j, are given below:

$$\frac{\delta\bar{x}}{\delta I_{\rm j}} = \frac{x_{\rm j}\sum\limits_{\rm k} I_{\rm i} - \sum\limits_{\rm k}(x_{\rm i} I_{\rm i})}{\left(\sum\limits_{\rm k} I_{\rm i}\right)^2} = \frac{x_{\rm j}-\bar{x}}{\left(\sum\limits_{\rm k} I_{\rm i}\right)},$$
$$\frac{\delta\bar{y}}{\delta I_{\rm j}} = \frac{y_{\rm j}\sum\limits_{\rm k} I_{\rm i} - \sum\limits_{\rm k}(y_{\rm i} I_{\rm i})}{\left(\sum\limits_{\rm k} I_{\rm i}\right)^2} = \frac{y_{\rm j}-\bar{y}}{\left(\sum\limits_{\rm k} I_{\rm i}\right)}, \tag{B.8}$$

$$\frac{\delta\Sigma_{\rm xx}}{\delta I_{\rm j}} = \frac{[(x_{\rm j}-\bar{x})^2 - 2(x_{\rm j}-\bar{x})\frac{\delta\bar{x}}{\delta I_{\rm j}} I_{\rm j}]\sum\limits_{\rm k} I_{\rm i} - \sum\limits_{\rm k}[(x_{\rm i}-\bar{x})^2 I_{\rm i}]}{\left(\sum\limits_{\rm k} I_{\rm i}\right)^2}$$
$$= \frac{[(x_{\rm j}-\bar{x})^2 - 2(x_{\rm j}-\bar{x})\frac{\delta\bar{x}}{\delta I_{\rm j}} I_{\rm j}] - \Sigma_{\rm xx}}{\sum\limits_{\rm k} I_{\rm i}}, \tag{B.9}$$

$$\frac{\delta\Sigma_{\rm yy}}{\delta I_{\rm j}} = \frac{[(y_{\rm j}-\bar{y})^2 - 2(y_{\rm j}-\bar{y})\frac{\delta\bar{y}}{\delta I_{\rm j}} I_{\rm j}]\sum\limits_{\rm k} I_{\rm i} - \sum\limits_{\rm k}[(y_{\rm i}-\bar{y})^2 I_{\rm i}]}{\left(\sum\limits_{\rm k} I_{\rm i}\right)^2}$$
$$= \frac{[(y_{\rm j}-\bar{y})^2 - 2(y_{\rm j}-\bar{y})\frac{\delta\bar{y}}{\delta I_{\rm j}} I_{\rm j}] - \Sigma_{\rm yy}}{\sum\limits_{\rm k} I_{\rm i}}, \tag{B.10}$$

$$\frac{\delta\Sigma_{\rm xy}}{\delta I_{\rm j}} = \frac{[(x_{\rm j}-\bar{x})(y_{\rm j}-\bar{y}) - (y_{\rm j}-\bar{y})\frac{\delta\bar{x}}{\delta I_{\rm j}} I_{\rm j}]\sum\limits_{\rm k} I_{\rm i}}{\left(\sum\limits_{\rm k} I_{\rm i}\right)^2}$$
$$- \frac{[(x_{\rm j}-\bar{x})\frac{\delta\bar{y}}{\delta I_{\rm j}} I_{\rm j}]\sum\limits_{\rm k} I_{\rm i}}{\left(\sum\limits_{\rm k} I_{\rm i}\right)^2} - \frac{\sum\limits_{\rm k}[(x_{\rm i}-\bar{x})(y_{\rm i}-\bar{y}) I_{\rm i}]}{\left(\sum\limits_{\rm k} I_{\rm i}\right)^2}$$
$$= \frac{[(x_{\rm j}-\bar{x})(y_{\rm j}-\bar{y}) - (y_{\rm j}-\bar{y})\frac{\delta\bar{x}}{\delta I_{\rm j}} I_{\rm j} - (x_{\rm j}-\bar{x})\frac{\delta\bar{y}}{\delta I_{\rm j}} I_{\rm j}] - \Sigma_{\rm xy}}{\sum\limits_{\rm k} I_{\rm i}} \tag{B.11}$$

We can now compute the gradient of the second moment regularization within the minimization framework of the `eht-imaging` library:

$$\frac{\delta\mathcal{R}_{\Sigma}}{\delta I_{\rm j}} = 2(\Sigma_{\rm xx}-\Sigma'_{\rm xx})\frac{\delta\Sigma_{\rm xx}}{\delta I_{\rm j}} + 2(\Sigma_{\rm yy}-\Sigma'_{\rm yy})\frac{\delta\Sigma_{\rm yy}}{\delta I_{\rm j}}$$
$$+4(\Sigma_{\rm xy}-\Sigma'_{\rm xy})\frac{\delta\Sigma_{\rm xy}}{\delta I_{\rm j}}. \tag{B.12}$$

Note that these equations correspond to regularization of the normalized second central moment of an image. In cases where the total flux density of an image is constrained or regularized, it would be advantageous to instead regularize the unnormalized second central moment, giving a substantially simplified and convex optimization problem.