Selective Dynamical Imaging of Interferometric Data


The Event Horizon Telescope Collaboration, Antxon Alberdi, Walter Alfer, Juan Carlos Algha, Richard Ananthu, Keiichi Asada, Rebecca Azulay, Anne-Kathrin Bazko, David Ball, Mislaw Balokovic, John Barrett, Bradford A. Benson, Dan Bintley, Lindy Blackburn, Raymond Blundell, Wilfred Boland, Hope Boyle, Michael Bremer.


Joseph Cruz-Osorio, Yuzhu Cui, Jie Qi, Jody Davelaar, Marilina De Laurentis.


Kazuhiro Hada, Mareki Honma, Geoffrey B. Crew, Jun Yi Koay, Makoto Inoue, Kaj Wiik.


Marko Jurkovic, Hideki Kojima, Steven J. Kounkel, Shoko Koyama.


The Event Horizon Telescope Collaboration, Antxon Alberdi, Walter Alfer, Juan Carlos Algha, Richard Ananthu, Keiichi Asada, Rebecca Azulay, Anne-Kathrin Bazko, David Ball, Mislaw Balokovic, John Barrett, Bradford A. Benson, Dan Bintley, Lindy Blackburn, Raymond Blundell, Wilfred Boland, Hope Boyle, Michael Bremer.


Joseph Cruz-Osorio, Yuzhu Cui, Jie Qi, Jody Davelaar, Marilina De Laurentis.


Kazuhiro Hada, Mareki Honma, Geoffrey B. Crew, Jun Yi Koay, Makoto Inoue, Kaj Wiik.


Marko Jurkovic, Hideki Kojima, Steven J. Kounkel, Shoko Koyama.

Abstract

Recent developments in very long baseline interferometry (VLBI) have made it possible for the Event Horizon Telescope (EHT) to resolve the innermost accretion flows of the largest supermassive black holes on the sky. The sparse nature of the EHT’S (u, v)-coverage presents a challenge when attempting to resolve highly time-variable sources. We demonstrate that the changing (u, v)-coverage of the EHT can contain regions of time over the course of a single observation that facilitate dynamical imaging. These optimal time regions typically have projected baseline distributions that are approximately angularly isotropic and radially homogeneous. We derive a metric of coverage quality based on baseline isotropy and density that is capable of ranking array configurations by their ability to produce accurate dynamical reconstructions. We compare this metric to existing metrics in the literature and use these results to make recommendations for imaging the 2017 EHT Sgr A° data set.

Unified Astrophyical Thesaurus concepts: Astrophysical black holes (98); Very long baseline interferometry (1769); Aperture synthesis (53); Interferometry (808)

1. Introduction

Interferometric astronomical observations offer much larger resolving power than do single telescopes, with the interferometric resolution depending on the distance between the elements rather than the diameters of the individual apertures. Since an interferometer probes the Fourier transform of an on-sky source (and not the source image itself), the placement, selection, and availability of baselines to maximize coverage of the (u, v)-plane is an important and open optimization problem inherent to the interferometric image synthesis. For very long baseline interferometry (VLBI), which represents some of the most extreme interferometric observations, the arrays are generally very sparse,
and optimized placement or selection of baselines becomes critical to recovering reliable information about the source.

For any ground-based interferometric observations, including VLBI, Earth rotation causes each individual baseline to trace an elliptical path in the \((u, v)\)-plane. Such “Earth rotation aperture synthesis” can be used to improve the \((u, v)\)-coverage of sparse arrays. Standard VLBI imaging uses this additional coverage under the assumption that the target source structure remains unchanged for the duration of the observation. For the cases of sources that vary on shorter timescales (e.g., the Galactic center, X-ray binaries, microquasars), “snapshot images” can be produced using only time intervals over which the source can be considered static (e.g., Martí-Vidal et al. 2011; Massi et al. 2012; Miller-Jones et al. 2019). In the most extreme scenarios, only instantaneous coverage is suitable for use in static image reconstructions, severely limiting the \((u, v)\)-coverage and resulting image quality. Such a case is the main concern of this paper. Moreover, this limitation will generally result in portions of an observation that are better than others for imaging because the snapshot coverage time is dependent.

In particular, directionally biased \((u, v)\)-coverage results in a nonisotropic resolution that complicates interpretation of the reconstructed source geometry. The \((u, v)\)-coverage of a given array varies with the decl. of the source, due to its dependence on the projected baseline lengths and orientations. The impact of source decl. on the \((u, v)\)-coverage varies throughout a night of observation as the target rises and sets. A network of baselines that produces an evenly distributed \((u, v)\)-coverage of an equatorial source will be east–west biased when targeting a northern source. Observations where the source morphology and orientation of the source are unknown cannot exploit a priori knowledge to determine whether a directionally biased \((u, v)\) configuration will be able to reproduce source structure reliably in an image. Therefore, for cases where the source morphology is unknown, the optimal \((u, v)\)-coverage for producing high-quality image reconstructions will be those that are approximately isotropic in angular distribution.

Periods of sustained isotropic \((u, v)\)-coverage may allow multiple high-quality image reconstructions on short timescales to be strung together, resulting in “dynamical” reconstructions. For instance, the most basic dynamical reconstruction can be produced by joining a series of snapshot images. The quality of such a reconstruction will be time dependent owing to the evolving snapshot coverage, and it may be necessary to exclude periods with poor \((u, v)\)-coverage. In periods of time where the \((u, v)\)-coverage is optimal, the coverage used to produce each snapshot of a dynamical reconstruction will be sufficiently isotropic to reproduce source features on snapshot timescales, allowing for the production of reconstructions that adequately recover the evolution of highly time-variable sources. By contrast, periods of suboptimal \((u, v)\)-coverage may be unable to provide high-quality reconstructions on snapshot timescales, making the algorithmic detection and subsequent flagging of these periods important for the production of meaningful dynamical reconstructions.

The challenges of observing a variable source with a sparse interferometer can be exacerbated by poor coverage geometry. In particular, given the shortage of available facilities supporting high-frequency radioastronomical observations, corresponding VLBI arrays exhibit substantial variance in coverage quality in time.

The Event Horizon Telescope (EHT) network of telescopes that exploits the full diameter of Earth and the performance at the challenging 1.3 mm (230 GHz) wavelength to achieve the required angular resolution for horizon-scale tests of general relativity for the largest black holes on the sky (see, e.g., Johannsen & Psaltis 2010; Psaltis et al. 2015; Psaltis 2019). The EHT operates in millimeter wavelengths, the optimal range that enables the resolution of the Sgr A* black hole shadow and reduces the impact of the interstellar medium scattering effects dominant for longer wavelengths, while still being observable and manageable in the radio interferometric framework. Operating at 230 GHz, the EHT achieved a resolution of \(\sim 20–25 \mu\text{as}\), which the EHT Collaboration (EHTC) used to produce the first images of a supermassive black hole (Event Horizon Telescope Collaboration et al. 2019a, 2019b, 2019c, 2019d, 2019e, 2019f). These data and the resulting images were used to estimate a mass of \(M \approx 6.5 \times 10^9 M_\odot\) for the supermassive black hole in M87 (Event Horizon Telescope Collaboration et al. 2019f), based on an observed angular shadow diameter of \(\sim 42 \mu\text{as}\) (Event Horizon Telescope Collaboration et al. 2019a, 2019d, 2019f).

During the 2017 campaign, the EHT also observed the radio source Sgr A* in the Galactic center (Event Horizon Telescope Collaboration et al. 2022a, 2022b, 2022c, 2022d, 2022e, 2022f), associated with a supermassive black hole with \(M \approx 4.1 \times 10^6 M_\odot\) (GRAVITY Collaboration et al. 2018a; Do et al. 2019; Event Horizon Telescope Collaboration et al. 2022f). The expected mass-to-distance ratio \((M/D)\) of Sgr A* yields a predicted angular shadow diameter of \(\sim 50 \mu\text{as}\) (GRAVITY Collaboration et al. 2018b) and a minimum variability timescale (light-crossing time) of \(GM/c^2 \approx 20\) s. The corresponding timescale for M87 is \(\sim 1600\) times longer owing to the larger mass. Indeed, structural variability of the M87 shadow has been reported on timescales from \(\sim 1\) week (Event Horizon Telescope Collaboration et al. 2019d) to several years (Wielgus et al. 2020b). The rapid minimum variability timescale of Sgr A*, combined with the extreme sparsity of the EHT, presents an urgent and unique need to characterize the effects of time-dependent instantaneous \((u, v)\)-coverage.

In this paper, we develop a procedure for selective imaging of highly variable sources. In Sections 2 and 3, we summarize the synthetic data generation and imaging methods used herein. In Section 4, we show the limitations of imaging in sparse and uneven coverage. In Section 5, we survey several metrics capable of ranking \((u, v)\)-coverage quality. Additionally, we derive a novel isotropy-based metric that addresses the limitations described in Section 4. In Section 6, we apply these metrics to the 2017 EHT coverage of Sgr A*, validate their ability to predict reconstruction quality from \((u, v)\)-coverage geometry, and make recommendations for selective dynamical imaging of the 2017 EHT Sgr A* data set. In Section 7, we briefly discuss the utility of coverage metrics in ranking and selecting between different available observing periods. Finally, in Section 8, we summarize our results.

## 2. Model Definition and Synthetic Data Generation

In order to test the ability of various EHT array configurations to recover source variability in different observation periods, we designed and generated synthetic data for three different models. The models are chosen owing to their structural similarity (in image and visibility domain) to expected images of Sgr A*.
Similarity between the model and data was characterized as either displaying time variability or producing a Bessel function. The rotating elliptical Gaussian model is generated using a bivariate exponential with major-axis FWHM $\Gamma_a$, minor-axis FWHM $\Gamma_b$, and overall flux density $A$:

$$I(x, y) = \frac{4A \ln 2}{\pi \Gamma_a \Gamma_b} \exp \left[-4 \ln 2 \left(\frac{x^2}{\Gamma_a^2} + \frac{y^2}{\Gamma_b^2}\right)\right].$$

(1)

The image is padded and rotated by polar angle $\varphi$ after the model is generated. By default, the overall flux density $A$ was set to 1.0 Jy. Periods for the rotation ranged between 30 and 1000 minutes (longer than one night of observation).

2.1.2. Ring and Orbiting Hot Spot

The ring model is generated by the subtraction of two concentric uniform-brightness disks, equivalent to the crescent model described in Kamruddin & Dexter (2013), with the parameters $a = b = 0$. The positive disk has a radius of $25 \mu$as, and the subtracted disk has a radius of $18 \mu$as. All ring models in this paper use these parameters with the exception of the ring model in Figure 4, which has a subtracted disk radius of $20 \mu$as.

The ring model is used with two synthetic data tests: the static ring (with no hot spot) and the dynamic ring (a static ring plus an orbiting hot spot, referred to as ring+hs). The underlying ring has a diameter of $50 \mu$as and a flux density of 1.0 Jy. A hot spot total flux density of 0.25 Jy and an FWHM of 10 $\mu$as are added to the image, centered on the ring. After construction, the total flux density of each image is normalized to 1.0 Jy. In the dynamic ring model, the orbiting hot spot is centered on the ring and circularly orbiting at a distance of 21.5 $\mu$as with periods of 30 and 270 minutes. The static ring model is a special case of the ring+hot spot model with the flux density of the hot spot set to zero.

2.2. Synthetic Data Generation

Synthetic data were generated based on April 7 of the 2017 EHT coverage using the eht-imaging library (Chael et al. 2018). A simulated observation corresponded to approximately 11.5 hr of observing on all available baselines. Snapshot images from the simulated sources and the resulting amplitudes and closure phases can be seen in Figure 1 and Appendix A.

Observation parameters (e.g., R.A. and decl. of the source, observing frequency, bandwidth) were duplicated from the 2017 EHT survey of Sgr A*. The stations included in the simulated observations were the Atacama Large Millimeter/submillimeter Array (ALMA), the Atacama Pathfinder Experiment (APEX), the Large Millimeter Telescope (LMT), the James Clerk Maxwell Telescope (JCMT), the Submillimeter Array (SMA), the IRAM 30 m telescope on Pico Veleta (PV), and the South Pole Telescope (SPT). Weather complications were ignored, and the simulated observations assumed that all stations were observing the source for the chosen night.

All data were generated with thermal noise only. Realistic values for the thermal noise power were based on estimates from the real 2017 EHT data. No other noise, scattering, leakage, or simulated gain errors were applied to the simulated visibilities.

3. Imaging Approaches

Since the interferometric measurements are often incomplete in the Fourier domain, the inverse problem of reconstructing an image from the observed data set is usually underdetermined. Consequently, the image reconstruction requires prior information, assumptions, or constraints to derive a reasonable image from the infinite number of possibilities that can explain the measurements.

The two most popular categories of imaging methodologies are inverse modeling (e.g., CLEAN) and forward modeling (e.g., regularized maximum likelihood). See Event Horizon Telescope Collaboration et al. (2019d) for a general overview of the two methods. For time-variable sources, both approaches may allow for more effective reconstructions of dynamic structures than snapshot imaging by including assumptions or constraints on temporal variations of the source structure in addition to the spatial properties regularized in static imaging (e.g., Event Horizon Telescope Collaboration et al. 2019d). One such way imposes a temporal similarity constraint between

---

Figure 1. The three synthetic models detailed in Section 2 are displayed. The first row shows each model as seen on-sky at 6 UT, just after the observation begins. The second row shows the model convolved with an 18 $\mu$as diameter CLEAN beam. The white circle in the lower right corner shows the size of the beam. The third row shows the measured visibility amplitudes as a function of baseline length for the entire observation. A static imaging routine would fit to the full set of these amplitudes; however, a dynamical imaging routine only attempts to fit to small chunks of the full data set at any one time.
images at different times and between each time snapshot and the time-averaged structure. In the following subsections, we briefly describe each dynamical reconstruction method used in this paper. See Event Horizon Telescope Collaboration et al. (2022c) for more details.

3.1. Inverse-modeling Approaches

Imaging of radio interferometric data is traditionally carried out through CLEAN deconvolution algorithms (e.g., Högbom 1974; Clark 1980). These inverse-modeling approaches iteratively deconvolve the effects associated with the limited sampling of the $(u, v)$-plane, corresponding to the interferometer’s point-source response (the so-called “dirty beam”) to the inverse Fourier transform of the measured visibilities, commonly referred to as the “dirty image.” The source brightness distribution is modeled as a collection of point sources, which are extracted at the location of the peaks in the dirty image through an iterative process until some specified stopping criterion is reached. In observations with limited $(u, v)$ sampling, such as those obtained with the EHT, it is important to guide the CLEAN deconvolution process through the inclusion of the so-called “cleaning windows,” restricting the sky areas within which the point components are localized.

Mitigation of the a priori calibration uncertainties is commonly carried out through multiple rounds of CLEAN deconvolution followed by self-calibration, which solves for the station gains that maximize consistency between the current model and the measured visibilities (e.g., Wilkinson et al. 1977; Readhead et al. 1980; Cornwell & Wilkinson 1981; Pearson & Readhead 1984). Amplitude self-calibrations are necessarily limited to intervals of time larger than the expected variability in order to retain information about source variability. The final image is obtained by convolving the model components with a Gaussian CLEAN beam that approximates the central lobe of the point-spread function of the interferometer, with the addition of the last residual image, which represents some unconvolved additional structure and noise. In this paper we use the Difmap software (e.g., Shepherd 1997 and M. Shepherd 2011, private communication) for CLEAN imaging.

Once the imaging procedure converges based on a specified stopping criterion into an average static image, the time-averaged structure. In the following subsections, we briefly describe each dynamical reconstruction method used in this paper. See Event Horizon Telescope Collaboration et al. (2022c) for more details.

3.2. Forward-modeling Approaches

Unlike the inverse-modeling methods, which solve for a sparse image on the image domain from the dirty map transformed from the measurement sets, the forward-modeling methods solve for an image by evaluating the data likelihood derived from the consistency between actual measurements and the model data set forward-transformed from the image. It offers flexibility to the imaging through robust data products (e.g., closure quantities that are not affected by station-based calibration uncertainties) and incorporates various observing effects into the observational equation used in the forward transform.

Regularized maximum likelihood (RML) methods (see Event Horizon Telescope Collaboration et al. 2019d, for an overview) optimize a cost function composed of $\chi^2$ terms (proportional to log-likelihood terms) of visibility components and regularization terms that describe the prior assumptions for images. Each regularization term is described as a product of its relative weight (i.e., hyperparameter) and regularization functions. These regularization functions include, e.g., maximum entropy (e.g., Narayan & Nityananda 1986; Chael et al. 2016), total variation and its variants (e.g., Akiyama et al. 2017; Kuramochi et al. 2018), and sparsity priors (e.g., Honma et al. 2014). The cost function can be interpreted as a maximum a posteriori (MAP) estimation by considering the regularization terms as log prior distribution of the image, although regularization functions do not always have a probabilistic interpretation. The final reconstruction is convolved with the CLEAN beam of the interferometer to remove the effects of methodology-specific super-resolution.

The RML approach can be extended to dynamic reconstruction (henceforth RML dynamic imaging) in a conceptually simple way (Johnson et al. 2017). The likelihood term can be formulated by forward-transforming snapshots of a video, instead of a single image, to data. One can add temporal regularization terms that penalize temporal variations of the source structure by defining a metric for the “distance” between adjacent frames. A popular choice is a sum of squared pixel differences between two adjacent snapshots, assuming that snapshot-to-snapshot transition of the source brightness is piecewise smooth (e.g., the $R_\Delta$ regularizer in Johnson et al. 2017). Another widely used choice is a sum of squared differences between the time-averaged image and each snapshot, based on an assumption conceptually similar to dynamic CLEAN imaging (Section 3.1) that the deviations of each snapshot from the mean image are small and sparse (e.g., $R_\Delta$ regularizer in Johnson et al. 2017). The temporal regularization term necessarily suppresses intrinsic source variability if weighted too high; however, what constitutes “too high” varies depending on the source structure and variability timescale. Popular image distance metrics include the Euclidean norm or a relative entropy such as Kullback–Leibler divergence (Kullback & Leibler 1951).

StarWarpS (Bouman et al. 2018) is another forward-modeling method for dynamical imaging adopted in this work. StarWarpS, based on a probabilistic graphical model, solves for snapshots of a video by solving its posterior probability distribution defined as a product of three terms: data likelihood, multivariate Gaussian distributions for each snapshot, and transition probabilities between adjacent snapshots effectively working as spatial and temporal regularizations, respectively. StarWarpS allows for the exact inference of the
video by computing a mean and covariance of the image, which provides a complete description under Gaussian approximation. By contrast, the RML dynamic reconstruction derives only a MAP estimation. StarWarp requires an initial static image, which can be either a data-driven (e.g., a best-fitting static image of the entire data set being dynamically reconstructed) or prior driven (e.g., a synthetic image of a ring).

In this paper, we use the RML dynamic imaging algorithms implemented in eht-imaging (also referred to as ehtim) and SMILI (Sparse Modeling Imaging Library for Interferometry) and the StarWarp algorithm in eht-imaging. See Event Horizon Telescope Collaboration et al. 2022c for more details of regularization functions and other imaging parameters used in the reconstructions.

4. Limitations of Sparse $(u,v)$-coverage

In this section, we explore limitations of imaging with limited and directionally biased $(u,v)$-coverage. In particular, we will show that sparse $(u,v)$-coverage results in predictable limitations (i.e., deviations from the true source morphology), which are determined by the geometry of the $(u,v)$-coverage.

Figure 2 shows how imaging using directionally biased $(u,v)$-coverage (column (a)) fails to properly recover the orientation of an intrinsically noncircularly symmetric source when the $(u,v)$-coverage does not sufficiently sample the source structure in the relevant direction. By contrast, the same baselines—oriented in a more isotropic way (column B)—are capable of recovering the source profile in all directions. In addition, reconstructions of circularly symmetric sources by directionally biased $(u,v)$-coverage (column (a) of Figure 3) introduce a lack of circular symmetry that is not present in the underlying source or in reconstructions performed on nondirectionally biased coverage (column (b) of Figure 3). Due to the minimization of the artifacts introduced into a reconstruction via incomplete $(u,v)$-coverage, imaging algorithms work better when applied to isotropic $(u,v)$-coverages.

In addition to angular inhomogeneity, radially inhomogeneous coverage also leads to ambiguous image reconstruction. The Fourier transforms of various source types (e.g., rings, crescents, Gaussians) are approximately degenerate when observed using an interferometer that only marginally resolves the image (e.g., Thompson et al. 2017; Issaoun et al. 2019). More complex approximate degeneracies exist for radially inhomogeneous interferometers that only probe short and long but not intermediate baselines (e.g., Doeleman et al. 2008). For the EHT observing a $\sim 50 \mu$as diameter ring, as expected for Sgr A*, short baselines correspond to those with length $<2$ G$\lambda$ and long baselines correspond to those with length $>6$ G$\lambda$. Figure 4 demonstrates that a Gaussian model describes the simulated EHT observation of a 50 G$\lambda$ static ring nearly as well as the static ring model itself if only particular subsets of the data are fit. Even an infinitesimally thin ring model, when only fit to medium and long baselines, can provide a high-quality fit while misrepresenting the total intensity of the source. Without sufficient radial homogeneity (i.e., coverage of short, medium, and long baselines, as seen in column (c) of Figure 2), fitting and interpreting a model confidently can be difficult.

Periods of $(u,v)$-coverage where these limitations are more likely to occur can be identified by constructing a metric that scores directional bias and radial homogeneity (i.e., coverage of short, medium, and long baselines). The prevalence and severity of reconstruction artifacts that result from coverage limitations form a continuum that can be used to rank different $(u,v)$ configurations. A metric based on these limitations could be applied to a full observation to distinguish different observing periods (composed of many evolving $(u,v)$ configurations) by their ability to produce high-quality reconstructions.

5. Coverage Metrics

Multiple $(u,v)$-coverage metrics with different underlying considerations exist in the literature. Here we summarize several metrics and compare the way they score a given observation. In addition, we develop a novel “isotropy metric” that has been tailored to the specific vulnerabilities detailed in Section 4.

Our “selective dynamical imaging” approach uses such a metric to identify intervals during an observation where the coverage is optimally configured for imaging. Importantly, these intervals are chosen before image reconstruction is attempted. In contrast, methods such as lucky imaging (e.g., Fried 1978) identify particularly useful images (e.g., with
beam is enough to reproduce the salient artifacts.

The predictability of the image artifacts resulting from the CLEAN beams shown in the top row. Simple convolution with the CLEAN exemplifies how model misspecification corresponds to the CLEAN beam and is directly linked to the instantaneous (u, v)-coverage, shown as a set of white circles. The white ellipse corresponds to the CLEAN beam and is directly linked to the (u, v)-coverage. The predictability of the image artifacts resulting from (u, v)-coverage is exemplified by the bottom panels, which show the model ring convolved with the CLEAN beams shown in the top row. Simple convolution with the CLEAN beam is enough to reproduce the salient artifacts.

A single score computed directly from (u, v)-coverage is a minimal distortion) after and on the basis of the reconstruction. A single score computed directly from (u, v)-coverage known a priori is preferable to an empirical approach (e.g., performing a simulated observation of a synthetic model and comparing reconstructions with the model), as it provides source-structure-agnostic assessments with a substantially lower performance cost.

Any metric capable of scoring different periods of (u, v)-coverage would have demonstrable limitations. Within a single observation, a comparison of (u, v)-coverage at two different points in time is a reliable way of determining which time region of the observation will produce superior image reconstructions. However, certain reconstruction-impacting data quantities can vary independently of the (u, v)-coverage. Sensitivity, calibration, and systematic uncertainty can also be important factors but are not probed by coverage metrics.

5.1. Normalized Cross-correlation

The normalized cross-correlation between two images is a measure of their similarity. By performing a dynamical reconstruction on a simulated observation and comparing each image of the dynamical reconstruction to the model, we can heuristically identify which portions of the observation produced the best reconstruction (i.e., the portions of the reconstruction with the greatest similarity to the model). We define the normalized cross-correlation \( \rho_{NX}(X, Y) \) of two images \( X \) and \( Y \) in an identical fashion to Event Horizon Telescope Collaboration et al. (2019d),

\[
\rho_{NX}(X, Y) = \frac{1}{N} \sum_i \frac{(X_i - \bar{X})(Y_i - \bar{Y})}{\sigma_X \sigma_Y}.
\]

A normalized cross-correlation between a model and the associated reconstruction is the most straightforward way to identify trustworthy periods of an observation, assuming that the reconstruction on synthetic data will behave in a similar fashion to a reconstruction on real data. This assumption is only upheld if care is taken to ensure that the qualities of the synthetic data match those of the real data. In addition, heuristic tests such as the normalized cross-correlation can be biased depending on the structure and inherent variability of the model chosen. If the source randomly aligns with collimated (u, v)-coverage at some point in the observation, it can result in a misleadingly high normalized cross-correlation that cannot be replicated for a different source model.

5.2. (u, v) Filling Fraction

Palumbo et al. (2019) propose a geometric scoring procedure for the (u, v)-coverage based on the specification of a desired array resolution \( \theta_{res} \) and imaging field of view \( \theta_{FOV} \). \( \theta_{res} \) sets an outer boundary with radius \( 1/\theta_{res} \) in the (u, v)-plane within which a “filling fraction” is computed, and it is typically taken to be the nominal array resolution set by the longest baseline in an observation. \( \theta_{FOV} \) determines a convolution radius of 0.71/\( \theta_{FOV} \) corresponding to the scale in the (u, v)-plane over
which the Fourier response to a filled disk on the sky of diameter \( \theta_{\text{FOV}} \) would decay to half of its maximum amplitude; we use \( \theta_{\text{FOV}} = 100 \mu\text{s} \) for the filling fraction computation in Figure 6. Intuitively, the largest image feature considered in the optimization of coverage sets the smallest scale of interest in the \((u, v)\)-plane; thus, convolving a proposed set of \((u, v)\) points by \(0.71/\theta_{\text{FOV}}\) yields a measure of what region in the \((u, v)\)-plane is sampled by measured visibilities. The fraction of the bounding circle sampled by the convolved coverage is the filling fraction.

Increasing the specified resolution (perhaps by increasing observation frequency) extends the bounding circle, decreasing the filling fraction unless \( \theta_{\text{FOV}} \) is correspondingly decreased. In this way, the filling fraction captures some features of the “spatial dynamic range” discussed in Lal & Lobanov (2007). As shown in Figure 7 of Palumbo et al. (2019), the filling fraction metric is a successful and nearly linear predictor of coverage. A coverage consisting of a single detection would demand that the CLEAN beamwidth in the case of the EHT is at approximately half of the diffraction-limited spatial dynamic range (Lal & Lobanov 2007). In this approach, we consider the coverage as a set of sampled \((u, v)\)-plane locations, to find the largest circle that can be drawn within the limits of the coverage that does not contain a coverage point in itself. Such a largest circular gap can be efficiently calculated with Delaunay triangulation of the coverage set (Barber et al. 1996). Then, the diameter of the gap \( d_{\text{max}} \) can be turned into a metric coefficient with

\[
m_{\text{LCG}} = 1 - d_{\text{max}}/\rho_{\text{max}},
\]

where \( \rho_{\text{max}} \) is the longest projected baseline length. If we demand that the \((u, v)\) distance corresponding to the center of the circle is less than \( \rho_{\text{max}} \), then we have \( 0 \leq m_{\text{LCG}} \leq 1 \), with \( m_{\text{LCG}} = 1 \) corresponding to the limit of a complete continuous coverage. A coverage consisting of a single detection would correspond to \( m_{\text{LCG}} = 0 \). Unlike the filling fraction metric, the LCG metric is independent of the assumed field of view.

5.3. Largest Coverage Gap

An alternative metric probing the coverage isotropy is based on identifying the largest gap in the \((u, v)\)-coverage; hence, we refer to it as the largest coverage gap (LCG) metric (Wielgus et al. 2020a). In this approach we consider the coverage as a set of sampled \((u, v)\)-plane locations, to find the largest circle that can be drawn within the limits of the coverage that does not contain a coverage point in itself. Such a largest circular gap can be efficiently calculated with Delaunay triangulation of the coverage set (Barber et al. 1996). Then, the diameter of the gap \( d_{\text{max}} \) can be turned into a metric coefficient with

\[
m_{\text{LCG}} = 1 - d_{\text{max}}/\rho_{\text{max}}.
\]

where \( \rho_{\text{max}} \) is the longest projected baseline length. If we demand that the \((u, v)\) distance corresponding to the center of the circle is less than \( \rho_{\text{max}} \), then we have \( 0 \leq m_{\text{LCG}} \leq 1 \), with \( m_{\text{LCG}} = 1 \) corresponding to the limit of a complete continuous coverage. A coverage consisting of a single detection would correspond to \( m_{\text{LCG}} = 0 \). Unlike the filling fraction metric, the LCG metric is independent of the assumed field of view.

5.4. Isotropy and Radial Homogeneity

We propose a novel metric of \((u, v)\)-coverage isotropy and radial homogeneity (hereafter referred to as the “isotropy metric”) based on the limitations described in Section 4. Similarly to the LCG metric, the isotropy metric penalizes anisotropy of the coverage, although the two approaches differ appreciably. In this approach, we treat the distribution of baselines in the Fourier plane as a mass distribution and quantify the radial and angular homogeneity using the second moments of inertia. We define the isotropy metric coverage parameter \( C \) for a given snapshot as

\[
C = I \left(1 - \frac{K}{K_{\text{max}}}ight),
\]

with

\[
I = 1 - \frac{\sqrt{(\langle u^2 \rangle - \langle v^2 \rangle)^2 + 4\langle uv \rangle^2}}{\langle u^2 \rangle + \langle v^2 \rangle},
\]

where \( \langle u^2 \rangle \), \( \langle v^2 \rangle \), and \( \langle uv \rangle \) are the second moments of the baseline distribution, \( K \) is the Kolmogorov–Smirnov (K-S) distance of the radial distribution of baseline lengths from uniform, and \( K_{\text{max}} \) is the maximum value of \( K \) at any point in time during the observation (or an arbitrary value for the purpose of cross-observation comparisons). The isotropy metric has the benefit of being fully analytic and automatically normalized between 0 and 1. A full derivation of the isotropy metric is presented in Appendix B.

5.5. Discussion of Metrics

Despite differences in methodology and implementation, the metrics we examined found similar fluctuations in \((u, v)\)-coverage quality and identify similar candidate time regions for high-quality imaging. A comparison of the metrics detailed in Sections 5.1–5.4 applied to the 2017 EHT \((u, v)\)-coverage of Sgr A* is shown in Figure 6. In general, the second half of the observation has superior \((u, v)\)-coverage as indicated by the metrics, and the period from ~01:00 GMST to ~03:30 GMST maximizes the various metrics.

The \((u, v)\) filling fraction and LCG metric generally produce results similar to the isotropy metric. Disagreements between the metrics can be seen especially at the beginning of the observation (e.g., 17–19 GMST) and in the middle (e.g., 21–23 GMST). These are periods where the \((u, v)\)-coverage is extremely sparse and not suitable for imaging, though the degree to which these periods are determined to be unsuitable varies depending on the specific considerations of the individual metrics.

The normalized cross-correlation metric, while the most direct measurement of time-varying reconstruction quality, lacks source structure agnosticism. If the source structure is known, the normalized cross-correlation metric can be a useful method of determining what periods of an observation are most advantageous for imaging that particular source structure. However, if the source structure is unknown, then a wide suite of representative source models must be tested to mitigate possible biases. Additionally, the normalized cross-correlation method will be tied to the particular imaging algorithm and hyperparameters used, making this metric less robust than the others considered.

The particular constructions of the metrics can lead directly to unintuitive or undesirable behavior. One example of undesirable behavior is a metric punishing a coverage for adding data points. Intuitively, more baselines lead to better coverage of the Fourier plane and therefore more information about the source. However, if these additional baselines are placed strategically, they can result in an unintuitive score assignment. A trivial example of this can be generated for the LCG metric. Consider a coverage with maximum baseline length less than \( \rho \) that achieves \( m_{\text{LCG}} \approx 1 \). By placing a single baseline of length \( L \) far outside the initial coverage (i.e., \( L \gg \rho \)), \( \rho_{\text{max}} = L \) and \( d_{\text{max}} \) goes as \( \sim L - \rho \). This drives \( m_{\text{LCG}} \) to zero and seems to indicate that the coverage has become demonstrably worse, when in reality the coverage quality has largely stayed the same, with the improvement of a single ultralong baseline. This type of array pathology does not occur
in the 2017 EHT coverage but may present an issue if the EHT goes to ultralong space baselines.

The isotropy metric exhibits similar misbehavior, as demonstrated in Figure 5. Given an isotropic coverage with a low number of baselines, adding just two baselines strategically can decrease the isotropy metric value by a substantial amount. With so few baselines, the addition of new baselines would intuitively be considered an improvement. However, the metric detects a decrease in isotropy and reports accordingly. This problem is only present for arrays with small numbers of baselines—it is difficult or impossible to significantly alter the isotropy of an array configuration for larger arrays using only a few baselines without resorting to ultra-long-baseline placement as in the LCG example.

An additional limitation that any metric based purely on coverage possesses results from unusual source structure. The metrics described above attempt to predict reconstruction quality by analyzing the coverage available, but this prediction is performed under the assumption of “reasonable” source structure (i.e., source structure with smooth, continuous Fourier representation). However, we can construct simple examples that would render these metrics unhelpful by violating the assumption of reasonable source structure. Consider a source whose Fourier transform has zero flux density everywhere an array has \((u, v)\)-coverage, and nonzero flux density everywhere else. Regardless of how good the coverage itself is (and therefore how well a given metric may score the coverage), there is no way to produce an accurate reconstruction of the source. This limitation is not likely to be an issue, as the restriction that source emission be nonnegative induces sufficient correlation in the Fourier domain that the possibility for arbitrary pathologies to “hide” in coverage-deficient swathes of the plane is severely limited.

The isotropy metric offers substantial performance (i.e., time and overhead cost) benefits over the other metrics considered, while reporting similar results. We test the performance of each metric by computing the per-snapshot score for a full 12 hr observation five times, resulting in \(~10,000\) snapshots available for scoring. We run the performance assessments on an i5-1038NG7 10th-generation Intel x86_64 processor with 16 GB of RAM. The normalized cross-correlation, which must reconstruct hundreds of images and compare them to model images, takes several hours to complete. The isotropy metric performs \(~10^5\) times faster than the full normalized cross-correlation, \(~20\) times faster than the \((u, v)\) filling fraction, and \(~10\) times faster than the LCG metric, while producing similar assessments of coverage. The substantial performance differences between the metrics will become more pronounced with larger arrays, such as the next-generation EHT (ngEHT) coverage. The low overhead generated by the isotropy metric in comparison to the other metrics examined makes the isotropy metric a more optimized method of scoring \((u, v)\)-coverages in important contexts, such as real-time track selection and long-term ngEHT site placement (Raymond et al. 2021), which is expanded on in Section 7.

6. Application of Metric to 2017 EHT Array

We apply the coverage metrics discussed in Section 5 to the EHT \((u, v)\)-coverage of Sgr A* corresponding to 2017 April 7. The instantaneous metric values for each snapshot of the observation are shown in Figure 6. This method of scoring the observation clearly identifies distinct periods of varying coverage quality. The time region from \(~01:30\) GMST to \(~03:10\) GMST (denoted as “Region II” in Figure 6) has the highest overall isotropy and baseline density of the observation. We select this period as a candidate time region for high-quality imaging (a “good” time region, i.e., one where the typical metric score is high). In contrast, the time region from \(~19:45\) GMST to \(~21:00\) GMST (denoted as “Region I” in Figure 6), while relatively stable, displays substantially lower coverage quality. We select this time region to examine the behavior of reconstructions in periods of ambiguous coverage quality. Exact time stamps for these time regions are given in Table 1.

By performing reconstructions in these time periods, we can validate the capability of the \((u, v)\)-coverage metrics to predict reconstruction quality based on coverage alone. We reconstruct four configurations of the ring+hs toy model detailed in Section 2: a 270-minute orbital period clockwise and counterclockwise, and a 30-minute orbital period clockwise and counterclockwise. The reconstructions in each time region are produced according to the RML and CLEAN imaging methods in Section 3, and we perform feature extraction on each image using the REX module of eht-imaging (Event Horizon Telescope Collaboration et al. 2019d) to recover the position angle of the hot spot at each moment in time. The extracted hot spot angles and the model orbits for Regions I and II are shown in Figures 7 and 8, respectively. Images sampled from dynamical reconstructions generated by each method are displayed in Section C. The “success” of a reconstruction is...
determined by the successful extraction of the hot spot position angle.

We find that Region II produces reconstructions that facilitate accurate recovery of dynamical variability. The reconstructions show a ring of approximately $50 \mu$as in diameter with a distinct hot spot. The recovered hot spot orientations are shown in Figure 8, along with a comparison to the model values. To compare the $N$ recovered position angles $\varphi'_\gamma$ with the model angle $\varphi^m_\gamma$, we use a phase-adjusted rms $R$, defined as

$$R = \left[ \frac{1}{N} \sum_{\gamma=1}^{N} \left\{ \begin{array}{ll} (\varphi'_\gamma - \varphi^m_\gamma)^2, & \text{if } |\varphi'_\gamma - \varphi^m_\gamma| \leq \pi \\ (2\pi - \varphi'_\gamma + \varphi^m_\gamma)^2, & \text{if } |\varphi'_\gamma - \varphi^m_\gamma| > \pi \end{array} \right. \right]^{1/2}.$$  

Overall, the Region II reconstructions successfully recover the dynamical variability in the model, with rms of the recovered angles and model varying between 0.16 and 0.20 rad. Decreases in coverage quality as measured by the metrics (indicated pointwise in Figure 8 by increases in transparency, using the isotropy metric as an example) correlate with lower-quality recovery of the hot spot position angle. These low-quality angle recoveries are most obvious in the RML reconstructions of the counterclockwise $T = 30$-minute case. The CLEAN algorithm appears to be more resistant to the sudden loss of coverage and maintains reconstruction quality even through drops in metric score. Excluding these lapses in coverage quality, Region II clearly facilitates the recovery of source structure and at least one kind of dynamical variability, covering a wide range of periods and directions.

By contrast, a comparatively low scoring time region (Region I) does not produce reconstructions capable of accurately recovering dynamical variability. Dynamically reconstructed images show a ring-like feature, but the time variation of the brightness asymmetry does not match the model. Overall, the Region I reconstructions fail to recover the dynamical variability in the model, with rms of the recovered angles and model varying between 1.21 and 1.73 rad. The rms on the recovery in Region I is

---

Figure 6. The application of all metrics (top left: isotropy; top middle: LCG; top right: $(u, v)$ filling fraction) described in Section 5 to the 2017 EHT coverage of Sgr A*. The observation begins on April 7, and the 00:00 GMT indicates the day change to April 8. The red time region corresponds to Region I, and the blue time region to Region II, as described in Section 6. The normalized cross-correlation (green, representing the comparison between a snapshot reconstruction of a static ring and the model) is shown as a reference for all metrics, as it is the only direct measure of instantaneous reconstruction quality. All three $(u, v)$-coverage-based metrics show high-quality coverage in the time region from $\sim01:30$ GMT to $\sim03:10$ GMT (Region II). The bottom row of the figure shows the coverage of Region I (bottom right) and Region II (bottom left) along with a representative snapshot.

---

Table 1

| Time Stamps for the Observation and Beginning and End of the Time Regions of Interest in the 2017 April 7 EHT Coverage of Sgr A*. |
|-----------------|-----------------|-----------------|-----------------|
|                | Start Time      | End Time        |                |
|                | UT GMST         | UT GMST         |                |
| Observation    | 4.0458 17.087   | 15.598 4.6721 Apr 8 |
| Region II      | 12.618 16.634 Apr 8 | 14.046 3.1155 Apr 8 |
| LMT drop       | 13.435 2.5045 Apr 8 | 13.737 2.8058 Apr 8 |

Note. Regions I and II correspond to the time regions identified in red and blue, respectively, in Figure 6. The LMT dropout corresponds to the sudden loss of coverage that occurs partway through Region II. All time stamps correspond to 2017 April 7 unless otherwise noted. In UT, the observation begins and ends on April 7; however, when converted to GMST, Region I lies on 2017 April 7 while Region II lies on 2017 April 8.
between 7 and 10 times higher than in Region II. For all tests, the scatter around the model is substantially larger than the scatter in Region II, rendering the accurate extraction of a period difficult or impossible. Reconstructions in periods outside of Regions I and II cannot recover even basic source structure without significant a priori information.

Based on these results, we expect that reconstructions on real data will produce the most accurate and robust recoveries in Figure 7. SMILI (green circles), ehtim (yellow triangles), and CLEAN (red squares) position angle reconstructions in Region I of a ring+hs model for orbits of various periodicities and directions. Comparing the resulting hot spot position angle to the model angle (blue dashed line) clearly demonstrates that Region I provides a poor reconstruction of time variability.

Figure 8. SMILI (green circles), ehtim (yellow triangles), and CLEAN (red squares) position angle reconstructions in Region II of a ring+hs model for orbits of various periodicities and directions. The transparency of each data point corresponds to the isotropy of the snapshot observation used to reconstruct it. Comparing the resulting hot spot position angle to the model angle (blue dashed line) clearly demonstrates that Region II provides excellent conditions and coverage for meaningful recovery of periodicity given reasonable imaging assumptions.
orbital dynamical variability in Region II of April 7 of the 2017 EHT observations of Sgr A*. Reasonable imaging and feature extraction procedures failed to produce meaningful results in Region I. This ranking is consistent with the predictions of the coverage metrics described in Section 5 and Appendix B. A wide array of factors impact reconstructions, and the tests presented here do not compose a realistic synthetic data suite for assessing whether or not Region II can accurately recover dynamical variability for Sgr A*. The tests provided here solely demonstrate that Region II is the best time region for performing dynamical reconstructions based on coverage considerations. Additional testing on more complex source types with realistic data corruptions can be found in Event Horizon Telescope Collaboration et al. (2022c).

7. Metric-based Array Comparisons

Go/no-go decisions about whether to proceed with an observing run are often made with limited information about the readiness or weather conditions at particular sites. Simulating (and scoring) array configurations with different dropouts—and characterizing changes in the size and quality of the identified candidate temporal regions—can facilitate those go/no-go choices while incorporating uncertainties about station status. One observation configuration can be considered “better” than another if it provides a candidate time region with a higher metric score, or a similar metric score for a longer duration. Figure 9 shows an example of this kind of analysis performed with a hypothetical ngEHT array. The top panel shows the isotropy metric (see Section 5.4) score per snapshot for the full array during a night of observation in which every station is observing, which represents the ideal scenario. The middle and bottom panels reproduce the metric score per snapshot assuming that four sites are unable to observe. By identifying a candidate time region and characterizing its quality (based on, e.g., average metric score in the time region) and duration (denoted as $\Delta \tau$ in Figure 9), we can track these characterizations through different combinations of site dropouts and estimate how critical the sites are to the array. Based on the computations in Figure 9, ALMA, JCMT, LMT, and SMA dropping out would be catastrophic to the array performance, as the optimal time region for dynamical imaging reduces in duration by $\approx 80\%$ and reduces in quality by $\approx 50\%$. By comparison, the combined dropout of PV, PDB, CARMA, and LMT does not substantially change the duration or quality of the identified candidate time region.

A metric-based comparison additionally provides a natural and quantitative ranking for identifying which day of an observation campaign produced the most optimal coverage for dynamical imaging. We can rank observations on separate days in a similar fashion to the dropout scenarios visualized in Figure 9 by comparing the duration and quality of the identified candidate time regions for each day. An example of such a comparison between the April 7 and April 10 runs of the 2017 EHT observation campaign is shown in Figure 10. Instantaneous metric scores are computed for each scan of both days, and candidate time regions for dynamical imaging (green) are identified. While both candidate time regions are of approximately the same duration, the candidate time region associated with April 7 displays substantially better coverage, making April 7 a better choice for dynamical imaging than April 10.

![Figure 9. Isotropy metric scores for a hypothetical ngEHT array with different dropout scenarios. The full array has all of the sites in the 2017 EHT array (see Section 2), with the addition of the Northern Extended Millimeter Array (PDB), the Combined Array for Research in Millimeter-wave Astronomy (CARMA), and the Kitt Peak National Observatory (KP). A universal value of $K_{\text{max}} = 0.513$ was used to perform the computation (see Appendix B). Generally, dropouts will impact the maximum coverage score achieved throughout the observation and the duration (shown as $\Delta \tau$) of the most optimal time region. Characterizing a site’s importance to the observation is a useful way of informing a go/no-go decision, which takes into account station readiness and probability of dropout. For the above observational scenario, the loss of, e.g., ALMA, JCMT, and SMA would likely motivate a “no-go” decision, whereas the loss of, e.g., CARMA and PV would not motivate canceling the night of observation.](image)

8. Conclusions

We have demonstrated that limitations of imaging associated with sparse baseline distribution can be inferred from the specific geometric properties of the $(u, v)$-coverage. Highly anisotropic coverage produces artifacts in reconstructions that distort the image in a direction consistent with stripes in the dirty beam. Additionally, an uneven radial distribution of the Fourier plane $(u, v)$-coverage tends to result in ambiguous image reconstruction. These limitations can be partially avoided by imaging in radially and angularly homogeneous coverage. The reconstruction issues associated with sparse coverage are exacerbated by rapid short-timescale variability, as seen in a wide array of astrophysical sources, including Sgr A* and the precessing jets of X-ray binaries.
Next, we surveyed and compared existing geometric measures of $(u, v)$-coverage, in addition to deriving a novel isotropy-based metric that addressed the specific limitations demonstrated in Section 4. The examined metrics included the normalized cross-correlation ($r$), the $(u, v)$ filling fraction (see Palumbo et al. 2019), and the LCG metric (Equation (3); see Wielgus et al. 2020a). The isotropy metric treats the distribution of baselines as a mass distribution and examines the second moment to rank coverages by homogeneity in the Fourier plane. The isotropy metric gives similar results to other $(u, v)$-coverage-based metrics while being more computationally efficient.

These metrics were applied to the April 7 data of the 2017 EHT observation of Sgr A*. Candidate time regions for dynamical imaging are highlighted in green. A universal value of $K_{max} \approx 0.513$ was used to perform the computation (see Appendix B). Dropouts can significantly change the available coverage, resulting in candidate time regions of different duration and quality. By sorting the coverage for the entire observation and directly comparing the predicted quality of candidate time regions, we can identify the best day to perform dynamical imaging. Although the candidate time regions identified on April 7 and April 10 are similar in duration, the candidate time region on April 7 presents $\approx 40\%$ better coverage.

Coverage metrics have additional utility for ranking interobservation comparisons based on their ability to recover dynamical variability, which has a variety of applications to the broader field of interferometry. These metrics provide the ability to make select observation time slots based on the capability of available antennas to recover particular dynamical evolution in the target. Such a scored assessment of coverage could well prove of use to other VLBI arrays both as they make go/no-go decisions about whether to observe on a particular night and then when identifying the periods of best coverage to perform static and dynamic imaging.

We thank the National Science Foundation (awards OISE-1743747, AST-1816420, AST-1716536, AST-1440254, AST-1935980) 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, which is funded by grants from the John Templeton Foundation and the Gordon and Betty Moore Foundation to Harvard University. Support for this work was also provided by the NASA Hubble Fellowship grant HST-HF2-51431.001-A awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555.

The Event Horizon Telescope Collaboration thanks the following organizations and programs: the Academy of Finland (projects 274477, 284495, 312496, 315721); the Agencia Nacional de Investigación y Desarrollo (ANID), Chile via NCN19058 (TITANs) and Fondecyt 3190878, the Alexander von Humboldt Stiftung; an Alfred P. Sloan Research Fellowship; Allegro, the European ALMA Regional Centre node in the Netherlands, the NL astronomy research network NOVA and the astronomy institutes of the University of Amsterdam, Leiden University, and Radboud University; the Black Hole Initiative at Harvard University, through a grant (60477) from the John Templeton Foundation; the China Scholarship Council; Consejo Nacional de Ciencia y Tecnología (CONACYT, Mexico, projects U0004-246083, U0004-259839, F0003-272050, M0037-279006, F0003-281692, 104497, 275201, 263356); the Delaney Family via the Delaney Family John A. Wheeler Chair at Perimeter Institute; Direccin General de Asuntos del Personal Acadmico-Universidad Nacional Autnoma de Mxico (DGAPA-UNAM, projects IN112417 and IN112820); the European Research Council Synergy Grant “BlackHoleCam: Imaging the Event Horizon of Black Holes” (grant 610058); the Generalitat Valenciana postdoctoral grant APOSTD/2018/177 and GenT Program (project CIDEGENT/2018/021); MICINN Research Project PID2019-108995GB-C22; the European Research Council for advanced grant “JETSET: Launching, propagation and emission of relativistic jets from binary mergers and across mass scales” (grant No. 884631); the Gordon and Betty Moore Foundation (grant GBMF-3561); the Istituto Nazionale di Fisica Nucleare (INFN) sezione di Napoli, iniziative specifiche TEONGRAV; the two Dutch National Supercomputers, Cartesius and Snellius (NWO grant 2021.013); the International Max Planck Research School for Astronomy and Astrophysics at the Universities of Bonn and Cologne; Joint Princeton/Flatiron and Joint Columbia/Flatiron Postdoctoral Fellowships, with research at the Flatiron Institute being supported by the Simons Foundation; the Japanese Government (Monbukagakusho: MEXT) Scholarship; the Japan Society for the Promotion of Science (JSPS)
Grant-in-Aid for JSPS Research Fellowship (JP17J08829); the Key Research Program of Frontier Sciences, Chinese Academy of Sciences (CAS, grants QYZDJ-SSW-SLH057, QYZDJ-SSW-SYS008, ZDBS-LY-SLH011); the Leverhulme Trust Early Career Research Fellowship; the Max-Planck-Gesellschaft (MPG); the Max Planck Partner Group of the MPG and the CAS; the MEXT/JSPS KAKENHI (grants 18K00990, JP21H01137, JP18H03721, 18K03709, 18H01245, 25120007); the Malaysian Fundamental Research Grant Scheme (FRGS) FRGS/1/2019/STG02/UM/02/6; the MIT International Science and Technology Initiatives (MISTI) Funds; the Ministry of Science and Technology (MOST) of Taiwan (103-2119-M-001-010-MY2, 105-2112-M-001-025-MY3, 105-2119-M-001-042, 106-2112-M-001-011, 2019-2119-M-001-013, 106-2119-M-001-027, 106-2923-M-001-005, 107-2119-M-001-017, 107-2119-M-001-020, 107-2119-M-001-041, 107-2119-M-010-005, 107-2923-M-001-009, 108-2112-M-001-048, 108-2112-M-001-051, 108-2923-M-001-002, 109-2112-M-001-025, 109-2124-M-001-005, 109-2923-M-001-001, 110-2112-M-003-007-MY2, 110-2112-M-001-053, 110-2124-M-001-007, and 110-2923-M-001-001); the Ministry of Education (MoE) of Taiwan Yushan Young Scholar Program; the Physics Division, National Center for Theoretical Sciences of Taiwan; the National Aeronautics and Space Administration (NASA, Fermi Guest Investigator grant 80NSSC20K1567, NASA Astrophysics Theory Program grant 80NSSC20K0527, NASA NuSTAR award 80NSSC20K0645); the National Institute of Natural Sciences (NINS) of Japan; the National Key Research and Development Program of China (grant 2016YFA0400704, 2016YFA0400702); the National Science Foundation (NSF, grants AST-0906454, AST-0352953, AST-0521233, AST-0705062, AST-0905844, AST-0922984, AST-1126433, AST-1140030, DGE-1144085, AST-1207704, AST-1207730, AST-1207752, MRI-1228509, OPP-1248097, AST-1310896, AST-1555365, AST-1615796, AST-1715061, AST-1716327, AST-1903847, AST-2034306); the Natural Science and Engineering Research Council of Canada (NSERC, including a Discovery Grant and the NSERC Alexander Graham Bell Canada Graduate Scholarships-Doctoral Program); the National Youth Thousand Talents Program of China; the National Research Foundation of Korea (the Global PhD Fellowship Grant: grants NRF-2015H1A2A1033752, 2015-R1D1A1A01065607; the Korea Research Fellowship Program: NRF-2015H1D3A1066561, Basic Research Support Grant 2019R1F1A1059721); the Netherlands Organization for Scientific Research (NWO) VICI award (grant 639.043.513) and Spinoza Prize SPI 78-409; the New Scientific Frontiers with Precision Radio Interferometry Fellowship awarded by the South African Radio Astronomy Observatory (SARAO), which is a facility of the National Research Foundation (NRF), an agency of the Department of Science and Technology (DST) of South Africa; the Onsala Space Observatory (OSO) national infrastructure, for the provisioning of its facilities/observational support (OSO receives funding through the Swedish Research Council under grant 2017-00648); the Perimeter Institute for Theoretical Physics (research at the Perimeter Institute is supported by the Government of Canada through the Department of Innovation, Science and Economic Development and by the Province of Ontario through the Ministry of Research, Innovation (grant and the Spanish Ministerio de Ciencia e Innovacin (grants PGC2018-098915-B-C21, AYA2016-80889-P, PID2019-108995GB-C21, PID2020-117404GB-C21); the State Agency for Research of the Spanish MCIU through the “Center of Excellence Severo Ochoa” award for the Instituto de Astrofísica de Andalucia (SEV-2017-0709); the Toray Science Foundation; the Consejera de Economia, Conocimiento, Empresas and Universidade of the Junta de Andalucia (grant P18-FR-1769), the Consejo Superior de Investigaciones Cientificas (grant 2019AEP112); the M2FINDERS project, which has received funding by the European Research Council (ERC) under the European Unions Horizon 2020 Research and Innovation Programme (grant agreement No. 101018682); the US Department of Energy (USDOE) through the Los Alamos National Laboratory (operated by Triad National Security, LLC, for the National Nuclear Security Administration of the USDOE (contract 89233218CNA000001); the European Unions Horizon 2020 research and innovation program under grant agreement No. 730562 RadioNet; Shanghai Pilot Program for Basic Research, Chinese Academy of Science, Shanghai Branch (JCYJ-SHFY-2021-013); ALMA North America Development Fund; the Academia Sinica; Chandra DDT-18089X and TM6-17006X; and the GenT Program (Generalitat Valenciana) Project CIDEGENT/2018/021. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), supported by NSF grant ACI-1548562, and CyVerse, supported by NSF grants DBI-0735191, DBI-1265383, and DBI-1743442. The XSEDE Stampede2 resource at TACC was allocated through TG-AST170024 and TG-AST080026N. The XSEDE JetStream resource at PTI and TACC was allocated through AST170028. The simulations were performed in part on the SuperMUC cluster at the LRZ in Garching, on the LOEWE cluster in CSC in Frankfurt, and on the HazelHen cluster at the HLRS in Stuttgart. This research was enabled in part by support provided by Compute Ontario (http://computeontario.ca), Calcul Quebec (http://www.calculquebec.ca), and Compute Canada (http://www.computecanada.ca). C.C. acknowledges support from the Swedish Research Council (VR). We thank the staff at the participating observatories, correlation centers, and institutions for their enthusiastic support. This paper makes use of the following ALMA data: ADS/JAO.ALMA#2016.1.01154.V. ALMA is a partnership of the European Southern Observatory (ESO; Europe, representing its member states), NSF, and National Institutes of Natural Sciences of Japan, together with National Research Council (Canada), Ministry of Science and Technology (MOST; Taiwan), Academia Sinica Institute of Astronomy and Astrophysics (ASIAA; Taiwan), and Korea Astronomy and Space Science Institute (KASI; Republic of Korea), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, Associated Universities, Inc. (AUI)/NRAO, and the National Astronomical Observatory of Japan (NAOJ). The NRAO is a facility of the NSF operated under cooperative agreement by AUI. Hector Olivares and Gibwa Musoke were supported by Virtual Institute of Accretion (VIA) postdoctoral fellowships from the Netherlands Research School for Astronomy (NOVA). APEX is a collaboration between the Max-Planck-Institut für Radioastronomie (Germany), ESO, and the Onsala Space Observatory (Sweden). The SMA is a joint project between the SAO and ASIAA and is funded by the Smithsonian Institution and the Academia Sinica. The JCMT is operated by the East Asian Observatory on behalf of the NAOJ, ASIAA, and KASI, as well
as the Ministry of Finance of China, Chinese Academy of Sciences, and the National Key R&D Program (No. 2017YFA0402700) of China. Additional funding support for the JCMT is provided by the Science and Technologies Facility Council (UK) and participating universities in the UK and Canada. Simulations were performed in part on the SuperMUC cluster at the LRZ in Garching, on the LOEWE cluster in CSC in Frankfurt, on the HazelHep cluster at the HLRS in Stuttgart, and on the Pi2.0 and Siyuan Mark-I at Shanghai Jiao Tong University. The computer resources of the Finnish IT Center for Science (CSC) and the Finnish Computing Competence Infrastructure (FCCI) project are acknowledged. Junghwan Oh was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2021R1A6A3A01086420). We thank Martin Shepherd for the addition of extra features in the Difmap software that were used for the CLEAN imaging results presented in this paper. The computing cluster of Shanghai VLBI correlator supported by the Special Fund for Astronomy from the Ministry of Finance in China is acknowledged. The LMT is a project operated by the Instituto Nacional de Astrofísica, Óptica, y Electrónica (Mexico) and the University of Massachusetts at Amherst (USA). The IRAM 30 m telescope on Pico Veleta, Spain, is operated by IRAM and supported by CNRS (Centre National de la Recherche Scientifique, France), MPG (Max-Planck Gesellschaft, Germany), and IGN (Instituto Geográfico Nacional, Spain). The SMT is operated by the Arizona Radio Observatory, a part of the Steward Observatory of the University of Arizona, with financial support of operations from the State of Arizona and financial support for instrumentation development from the NSF. Support for SPT participation in the EHT is provided by the National Science Foundation through award OPP-1852617 to the University of Chicago. Partial support is also provided by the Kavli Institute of Cosmological Physics at the University of Chicago. The SPT hydrogen maser was provided on loan from the GLT, courtesy of ASIAss. Support for this work was provided by NASA through the NASA Hubble Fellowship grant No. HST-HF2-51494.001 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. The EHTC has received generous donations of FPGA chips from Xilinx Inc., under the Xilinx University Project. The EHTC has benefited from technology shared under open-source license by the Collaboration for Astronomy Signal Processing and Electronics Research (CASPER). The EHT project is grateful to T4Science and Microsemi for their assistance with Hydrogen Masers. This research has made use of NASA’s Astrophysics Data System. We gratefully acknowledge the support provided by the extended staff of the ALMA, both from the inception of the ALMA Phasing Project through the observational campaigns of 2017 and 2018. We would like to thank A. Deller and W. Brisken for EHT-specific support with the use of DiFX. We thank J. Delgado for helpful discussions and feedback. We acknowledge the significance that Maunakea, where the SMA and JCMT EHT stations are located, has for the indigenous Hawaiian people.

Appendix A
Sample Synthetic Data Products

Here we present representative samples of the visibility amplitudes and closure phases that are generated from the synthetic data described in Section 2.1. Sample amplitudes are shown in Figure 11, and sample closure phases are shown in Figure 12.

Figure 11. Example amplitudes for the three main synthetic data types described in Section 2.1. The time-variable models have periods of 270 minutes. The amplitudes recorded during Regions I and II (see Table 1) are shown in red and blue, respectively. Region II has higher radial homogeneity in $(u, v)$ distance, which contributes to its higher metric score and increased dynamical imaging capability. Error bars show $1\sigma$ thermal noise.
Derivation of Isotropy-based Coverage Metric

Section 4 demonstrated that quantifying the isotropy of a \((u, v)\)-coverage configuration can indicate whether it is suitable for producing accurate reconstructions of a dynamic source. We adopt a coverage metric of the form

\[
I = R(B) \rho(u, v),
\]

where \(\rho(u, v)\) is the set of \(2N\) baselines (including their Hermitian conjugates), \(I\) is a measure of the isotropy of the coverage, and \(R\) is a measure of the radial homogeneity of the coverage.

To estimate the radial homogeneity, we compare the cumulative distribution function (CDF) of the distribution of baseline lengths against the uniform CDF via the K-S test. The uniform distribution examined for this test ranges from 0 to the maximum baseline length achieved in the observation. This test returns a "distance" \(\rho\) between the distributions, which increases as the distribution becomes less radially homogeneous, making use of the test in this context a measure of radial inhomogeneity. To convert the result of the K-S test into a measure of radial homogeneity, we select an upper bound \(\rho_{\text{max}}\) corresponding to the maximum distance from uniform any individual baseline distribution obtains throughout the observation. We then subtract the result of the K-S test from this maximum, i.e.,

\[
\rho' = \rho_{\text{max}} - \rho
\]

where \(R(\rho) = \rho'\) is our new metric of homogeneity. This metric is conveniently bounded between 0 and 1. To make this metric absolute, a fixed value of \(\rho_{\text{max}}\) can be chosen arbitrarily and applied to multiple observations. For the absolute comparisons in this paper (see, e.g., Section 7) a value of \(\rho_{\text{max}} = 0.513\) is adopted. This value of \(\rho_{\text{max}}\) is chosen to be the maximum value of \(\rho\) achieved during the April 7 observation.

In order to measure the isotropy of the coverage, we examine the second moment (moment of inertia) of the distribution of baselines. As a spatial configuration of points with uniform weighting, the \((u, v)\)-coverage can be treated as a mass distribution.

For a two-dimensional mass distribution, a disk is considered isotropic, a rod is considered anisotropic, and the spectrum between the two cases can be probed using the moment of inertia tensor. Given \(2N\) baselines and conjugate baselines with coordinates \(\rho(\{u_i, v_i\})\), we can compute the second moments of the distribution as

\[
\begin{align*}
\langle u^2 \rangle &= \frac{1}{2N} \sum_{i=1}^{2N} u_i^2, \\
\langle v^2 \rangle &= \frac{1}{2N} \sum_{i=1}^{2N} v_i^2, \\
\langle uv \rangle &= \frac{1}{2N} \sum_{i=1}^{2N} u_i v_i.
\end{align*}
\]

The moment of inertia tensor \(\mathcal{M}\) is

\[
\mathcal{M} = \begin{bmatrix} \langle u^2 \rangle & \langle uv \rangle \\ \langle uv \rangle & \langle v^2 \rangle \end{bmatrix}.
\]

The principal moments of inertia can be computed from the eigenvalues \(\lambda_1\) and \(\lambda_2\) of \(\mathcal{M}\). From these, we can derive the following orientation-independent measure of isotropy:

\[
\tilde{I}(\{\rho\}) = 1 - \frac{|\lambda_1 - \lambda_2|}{\lambda_1 + \lambda_2} = 1 - \frac{\sqrt{(\langle u^2 \rangle - \langle v^2 \rangle)^2 + 4\langle uv \rangle^2}}{\langle u^2 \rangle + \langle v^2 \rangle}.
\]

Note that \(\tilde{I}(\{\rho\})\) can also be written in terms of the FWHM of the dirty beam,

\[
\tilde{I}(\{\rho\}) = 1 - \frac{1/\theta_{\text{min}}^2 - 1/\theta_{\text{maj}}^2}{1/\theta_{\text{min}}^2 + 1/\theta_{\text{maj}}^2} = \frac{2\theta_{\text{min}}^2}{\theta_{\text{maj}}^2 + \theta_{\text{min}}^2}.
\]
quality metric:

\[ C = I(\{\bar{u}_i\})R(\{\bar{v}_i\}) \]

\[ = \left(1 - \frac{\sqrt{\langle u^2 \rangle - \langle v^2 \rangle^2 + 4\langle uv \rangle^2}}{\langle u^2 \rangle + \langle v^2 \rangle} \right) \left(1 - \frac{K}{K_{\text{max}}} \right). \]  

(B8)

This measure of coverage quality can be applied to partition an arbitrary VLBI observation into time regions ranked by their ability to accurately reconstruct dynamical sources.

Appendix C
Sampled Results from Synthetic Data Reconstructions

In Section 6, we test the coverage metrics examined in Section 5 by performing reconstructions of synthetic observations in selected time regions of the 2017 EHT observational coverage of Sgr A* . Here we provide representative snapshot images from each of the time regions. Sampled snapshot images from Region I are shown in Figure 13, and sampled snapshot images from Region II are shown in Figure 14.

Reconstructions in both Region I and Region II demonstrate clear recovery of a ring-like feature of approximately 50 \( \mu \text{as} \) in diameter, as both sufficiently probe the radial distribution of the source Fourier transform to constrain the overall size of the source. However, the directionally biased coverage of Region I produces incorrect reconstructions of hot spot location. By contrast, reconstructions in Region II repeatedly recover the correct hot spot location across all periods, directions, and imaging algorithms.

Though both time regions recover a ring-like feature of the approximately correct size, the Region II reconstructions provide a more accurate ring-to-central-depression flux density ratio. The increased accuracy is present in both the CLEAN and RML reconstructions. By contrast, reconstructions in Region I fail to consistently provide a visually distinctive depression and misrepresent the angular brightness profile of the source.

Figure 13. Sampled reconstructions from Region I. The top row shows the true model image for each configuration. All panels show model images and reconstructions at \( \sim 21:00 \) GMST on April 7. The white circle in the lower right corner of each panel corresponds to an 18 \( \mu \text{as} \) diameter CLEAN beam. Even with substantial prior assumptions that facilitate ring reconstruction, the hot spot is frequently placed incorrectly, rendering this time region unsuitable for recovery of orbital angular variability.

<table>
<thead>
<tr>
<th>30 min, CW</th>
<th>30 min, CCW</th>
<th>270 min, CW</th>
<th>270 min, CCW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Region I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CLEAN</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RML</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 14. Sampled reconstructions from Region II. The top row shows the true model image for each configuration. The white circle in the lower right corner of each panel corresponds to an 18 μas diameter CLEAN beam. All panels show model images and reconstructions at ~1:50 GMST on April 8.

ORCID iDs

Joseph Farah @ https://orcid.org/0000-0003-4914-5625
Peter Galison @ https://orcid.org/0000-0002-6429-3872
Kazunori Akiyama @ https://orcid.org/0000-0002-9475-4254
Katherine L. Bouman @ https://orcid.org/0000-0003-0077-4367
Geoffrey C. Bower @ https://orcid.org/0000-0003-4056-9982
Andrew Chael @ https://orcid.org/0000-0003-2966-6220
Antonio Fuentes @ https://orcid.org/0000-0002-8773-4933
José L. Gómez @ https://orcid.org/0000-0003-4190-7613
Mareki Honma @ https://orcid.org/0000-0003-4058-9000
Michael D. Johnson @ https://orcid.org/0000-0002-4120-3029
Daniel P. Marrone @ https://orcid.org/0000-0002-2367-1080
Kotaro Moriyama @ https://orcid.org/0000-0003-1364-3761
Ramesh Narayan @ https://orcid.org/0000-0002-1919-2730
Dominic W. Pesce @ https://orcid.org/0000-0002-5278-9221
Paul Tiede @ https://orcid.org/0000-0003-3826-5648
Maciek Wielgus @ https://orcid.org/0000-0002-8635-4242
Guang-Yao Zhao @ https://orcid.org/0000-0002-4417-1659
Antxon Alberdi @ https://orcid.org/0000-0002-9371-1033
Juan Carlos Algabe @ https://orcid.org/0000-0001-6993-1696
Richard Anantua @ https://orcid.org/0000-0003-3457-7660
Rebecca Azulay @ https://orcid.org/0000-0002-2200-5393
Anne-Kathrin Baczko @ https://orcid.org/0000-0003-3090-3975
Mislav Baloković @ https://orcid.org/0000-0003-0476-6647
John Barrett @ https://orcid.org/0000-0002-9290-0764
Bradford A. Benson @ https://orcid.org/0000-0002-5108-6823
Lindy Blackburn @ https://orcid.org/0000-0002-9030-642X
Raymond Blundell @ https://orcid.org/0000-0002-5929-5857
Hope Boyce @ https://orcid.org/0000-0002-6530-5783
Christiaan D. Brinkerink @ https://orcid.org/0000-0002-2322-0749
Roger Brissenden @ https://orcid.org/0000-0002-2556-0894
Silke Britzen @ https://orcid.org/0000-0001-9240-6734
Avery E. Broderick @ https://orcid.org/0000-0002-3351-760X
Thomas Bronzwaer @ https://orcid.org/0000-0003-1151-3971
Do-Young Byun @ https://orcid.org/0000-0003-1157-4109
Chi-kwan Chan @ https://orcid.org/0000-0001-6337-6126
Koushik Chatterjee @ https://orcid.org/0000-0002-2825-3590
Shami Chatterjee @ https://orcid.org/0000-0002-2878-1502
Ilje Cho @ https://orcid.org/0000-0001-6083-7521
Pierre Christian @ https://orcid.org/0000-0001-6820-9941
John E. Conway @ https://orcid.org/0000-0003-2448-9181
James M. Cordes @ https://orcid.org/0000-0002-4049-1882
Thomas M. Crawford @ https://orcid.org/0000-0001-9000-5013
Geoffrey B. Crew @ https://orcid.org/0000-0002-2079-3189
Alejandro Cruz-Osorio @ https://orcid.org/0000-0002-3945-6342
Yuzhu Cui @ https://orcid.org/0000-0001-6311-4345
Jordy Davelaar @ https://orcid.org/0000-0002-2685-2434
Mariafelicia De Laurentis @ https://orcid.org/0000-0002-9945-682X
Roger Deane @ https://orcid.org/0000-0003-1027-5043
Jessica Dempsey @ https://orcid.org/0000-0003-1269-9667
Gregory Desvignes @ https://orcid.org/0000-0003-3922-4055
Sheperd S. Doeleman @ https://orcid.org/0000-0002-9031-0904
Ralph P. Eatough @ https://orcid.org/0000-0001-6196-4135
Heino Falcke @ https://orcid.org/0000-0002-2526-6724
Vincent L. Fish @ https://orcid.org/0000-0002-7128-9345

References


Fried, D. L. 1978, JOSA, 68, 1651
Psaltis, D. 2019, GRGr, 51, 137
Shepherd, M. C. 1997, adass, 125, 77