The Astrophysical Journal Letters, 930:L14 (64pp), 2022 May 10 https://doi.org/10.3847/2041-8213/ac6429

First Sagittarius A* Event Horizon Telescope Results. III. Imaging of the Galactic Center Supermassive Black Hole

The Event Horizon Telescope Collaboration
(See the end matter for the full list of authors.)

Received 2022 February 28; revised 2022 March 25; accepted 2022 March 27; published 2022 May 12

Abstract

We present the first event-horizon-scale images and spatiotemporal analysis of Sgr A* taken with the Event Horizon Telescope in 2017 April at a wavelength of 1.3 mm. Imaging of Sgr A* has been conducted through surveys over a wide range of imaging assumptions using the classical CLEAN algorithm, regularized maximum likelihood methods, and a Bayesian posterior sampling method. Different prescriptions have been used to account for scattering effects by the interstellar medium toward the Galactic center. Mitigation of the rapid intraday variability that characterizes Sgr A* has been carried out through the addition of a "variability noise budget" in the observed visibilities, facilitating the reconstruction of static full-track images. Our static reconstructions of Sgr A* can be clustered into four representative morphologies that correspond to ring images with three different azimuthal brightness distributions and a small cluster that contains diverse nonring morphologies. Based on our extensive analysis of the effects of sparse (u, v)-coverage, source variability, and interstellar scattering, as well as studies of simulated visibility data, we conclude that the Event Horizon Telescope Sgr A* data show compelling evidence for an image that is dominated by a bright ring of emission with a ring diameter of $\sim$50 μas, consistent with the expected "shadow" of a $4 \times 10^6\ M_\odot$ black hole in the Galactic center located at a distance of 8 kpc.

Unified Astronomy Thesaurus concepts: Synthesis imaging (53); Black Holes (162); Galactic center (565); Radio astronomy (1338); Very long baseline interferometry (1769); High angular resolution (2167)

1. Introduction

At the center of our Galaxy is the nearest candidate supermassive black hole (SMBH), Sagittarius A* (Sgr A*). We present here the first horizon-scale images of it. Compared with M87*, Sgr A* is more challenging to image, mainly due to its rapid variability and the distortions of the intervening scattering medium. We develop methods to characterize and mitigate these two factors in order to reconstruct images that take us a step closer to establishing that Sgr A* is a black hole. This paper is the third in the Event Horizon Telescope's (EHT) series of six Sgr A* articles (Event Horizon Telescope Collaboration et al. 2022a, 2022b, 2022c, 2022d, and 2022e, hereafter Papers I, II, IV, V, and VI).

Since the first discovery of Sgr A* as a compact radio source with interferometric observations at centimeter (cm) wavelengths (Balick & Brown 1974), there have been many studies of this closest SMBH to Earth. Of particular importance is the study of stellar dynamics showing that the position of Sgr A* in the Galactic center coincides with the center of gravity of a dense cluster of young and old stars (Eckart & Genzel 1997; Menten et al. 1997; Ghez et al. 1998; Reid & Brunthaler 2004; Reid 2009). Moreover, the gravitational potential is dominated by a compact object of mass of $4 \times 10^6\ M_\odot$ contained within 120 au of Sgr A*, at a distance of 8 kpc from Earth (Ghez et al. 2008; Gillessen et al. 2009, 2017; Gravity Collaboration et al. 2018a; Do et al. 2019; Gravity Collaboration et al. 2019). Based on these facts, together with the continuous variability on characteristic timescales from minutes to hours, especially at near-infrared (NIR) wavelengths on angular scales of typically 150 μas (Gravity Collaboration et al. 2018b, 2020), the likely scenario of Sgr A* is that this compact object is an SMBH. The combination of mass and proximity makes Sgr A* the black hole subtending the largest angle on the sky with a Schwarzschild radius of 0.08 au $\sim$ 10 μas and an expected "shadow" angular size of $\sim$50 μas. Sgr A* was thus identified early on as a primary target for imaging a black hole "shadow" (Falcke et al. 2000), predicted by Einstein's theory of general relativity (Hilbert 1917; von Laue 1921; Bardeen 1973; Luminet 1979). Similar calculations put the shadow angular size of M87* at $\sim$40 μas ($6.5 \times 10^9\ M_\odot$, 16.4 Mpc from Earth), confirmed by the EHT through imaging and analysis (Event Horizon Telescope Collaboration et al. 2019a, 2019b, 2019c, 2019d, 2019e, 2019f, hereafter M87* Papers I–VI).

In very long baseline interferometry (VLBI) observations at cm wavelengths, the structure of Sgr A* is unresolved and dominated by scatter broadening caused by the ionized interstellar medium (ISM; see, e.g., Rickett 1990; Narayan 1992). As a result, the measured sizes are proportional to λ^2, where λ is the observing wavelength (Davies et al. 1976), with an asymmetric Gaussian shape elongated toward the east–west direction (i.e., stronger angular broadening; Lo et al. 1985; Alberdi et al. 1993; Krichbaum et al. 1993; Frail et al. 1994; Bower & Backer 1998; Lo et al. 1998). For several decades, many VLBI observations have attempted to reach smaller angular (and spatial) scales. These studies found that the observed size at millimeter (mm) wavelengths deviates from the λ^2-relation, indicating a larger intrinsic size than expected from scatter broadening of an intrinsically unresolved source (e.g., Krichbaum et al. 1998; Lo et al. 1998; Doeleman et al. 2008; Falcke et al. 2009).

After constraining the scattering effects (see Section 3.1), the intrinsic structure of Sgr A* at long radio wavelengths can be modeled with a single, nearly isotropic Gaussian source

(Bower et al. 2004; Shen et al. 2005; Lu et al. 2011; Bower et al. 2014; Johnson et al. 2018; Issaoun et al. 2021; Cho et al. 2022). Its size and orientation on the sky have remained fairly constant over timescales of days to years (e.g., Alberdi et al. 1993; Marcaide et al. 1999; Lu et al. 2011), but a marginal variation has also been suggested (Bower et al. 2004; Akiyama et al. 2013). At observing wavelengths >1 mm, some evidence for structure beyond a single Gaussian model has also been reported. While this is likely attributed to refractive scattering substructure (i.e., not intrinsic) at cm wavelengths (e.g., Gwinn et al. 2014), its cause is still unclear at mm wavelengths. For instance, at 7 mm, Rauch et al. (2016) reported a short-lived secondary component that could possibly be related to a preceding NIR flare. However, the detection of nonzero closure phases was only marginal and does not exclude a realization of thermal or other systematic errors. At 3.5 mm, several studies have found slight nonzero closure phases (Brinkerink et al. 2016; Ortiz-León et al. 2016) and asymmetric non-Gaussian structure along the minor axis (Issaoun et al. 2019, 2021), but its physical origin remained nonconclusive: it could be due to either scattering or an intrinsic asymmetry of Sgr A*.

Multiwavelength observations of Sgr A* show an inverted spectral energy distribution rising with frequency in the radio owing to synchrotron emission, with spectral break at THz frequencies (submillimeter wavelengths), where the accretion flow becomes optically thin (Falcke et al. 1998; Bower et al. 2015, 2019). Its bolometric luminosity was measured to be $\sim 5 \times 10^{35}$ erg s^{-1}, or $10^{-9} L_{\rm Edd}$ (Genzel et al. 2010; Bower et al. 2019). A more detailed description of the spectral properties of Sgr A* is presented in Paper II. At an observing frequency of 230 GHz (1.3 mm wavelength), the accretion flow is expected to be sufficiently optically thin to detect the black hole shadow in Sgr A* with an Earth-sized interferometric array, such as the EHT (Falcke et al. 2000; Doeleman et al. 2009; Broderick et al. 2016; M87* Paper II).

Sgr A* additionally exhibits variability across the entire electromagnetic spectrum (Genzel et al. 2003; Ghez et al. 2004; Neilsen et al. 2013; Bower et al. 2015; Neilsen et al. 2015; Boyce et al. 2019), with frequent flaring in the radio, infrared, and X-ray regimes. Variability and motion can be observed within a single observing night, with variability timescales of the order of seconds to hours, characteristic dynamic timescales for a $4 \times 10^6\ M_\odot$ black hole (Baganoff et al. 2003; Marrone et al. 2006; Meyer et al. 2008; Dexter et al. 2014; Hora et al. 2014; Bower et al. 2018; Witzel et al. 2018; Bower et al. 2019). In the radio and submillimeter, Sgr A* is constantly varying, with a variability level of <10% during quiescence (Macquart & Bower 2006; Paper II; Wielgus et al. 2022). A detailed description of the multiwavelength properties of Sgr A* is presented in Paper II.

At 1.3 mm, Sgr A* was detected for the first time with VLBI on a single baseline with the IRAM 30 m telescope and the Plateau de Bure Interferometer in 1995 (Krichbaum et al. 1998). In 2007, the first successful observations with the early EHT array, consisting of the Arizona Radio Observatory Submillimeter Telescope (SMT) in Arizona, the James Clerk Maxwell Telescope (JCMT) in Hawai'i, and the Combined Array for Research in Millimeter-wave Astronomy (CARMA) in California, offered the first line of evidence that the source size at 1.3 mm is comparable to the expected size of the shadow of an SMBH with the mass and position of Sgr A* (Doeleman et al. 2008; Fish et al. 2011). In 2013, 1.3 mm observations were carried out with an early subset of the present EHT array: five stations at four geographical sites (Arizona, California, Hawai'i, and Chile). An early processing of the US-only data resulted in a first measurement of relatively high linear polarization on the 50–100 μas scale by Johnson et al. (2015) and the detection of nonzero closure phases by Fish et al. (2016), indicative of asymmetric source structure on the Arizona–California–Hawai'i triangle. A final processing of the data with the addition of the Atacama Pathfinder Experiment telescope (APEX) in Chile by Lu et al. (2018), which provides a resolution of $\sim$30 μas (3 Schwarzschild radii for the estimated black hole mass) in the north–south direction, revealed the presence of compact structure residing within the scale of 50 μas and confirmed the previously reported asymmetry (nonzero closure phase) by Fish et al. (2016). A subsequent expansion effort of the EHT to increase array sensitivity and imaging ability culminated in the 2017 April EHT observing campaign (M87* Paper II).

The 2017 EHT observing campaign was scheduled over a 12-day time window in 2017 April, to minimize weather impact. The two primary targets, M87* (at the center of the giant elliptical galaxy M87) and Sgr A*, were observed for four and five nights, respectively. With a similar mass-to-distance ratio, the two targets are expected to exhibit similar angular sizes on the sky. Because M87* is about three orders of magnitude more massive compared to Sgr A*, its dynamical timescale is much longer, allowing for the straightforward use of standard aperture synthesis VLBI techniques over each observing night. The effects of scattering toward the M87 galaxy are also minimal. These factors render M87* the optimal first imaging target for the EHT. Based on the observations of M87*, the EHT Collaboration presented first direct images of an SMBH, showing a bright ring-like structure surrounding a central dark circular area (M87* Papers I–VI). Under the stellar dynamics mass measurement prior (Gebhardt et al. 2011) and with a scaling based on numerical simulations of the accretion flow (M87* Paper V), these images confirm the predictions of general relativity about the diameter of a black hole shadow (M87* Paper VI). Following these results, the analysis of the linear polarization observations of M87* produced the first polarized images of the M87* black hole and inferred a magnetic field strength and geometry in the immediate vicinity of the SMBH (Event Horizon Telescope Collaboration et al. 2021a, 2021b, hereafter M87* Papers VII and VIII). The technique and workflow developments for the analysis of M87* data serve as the basis for Sgr A* analysis, although new significant developments were introduced to address the challenges of interstellar scattering and short-timescale variability.

In this paper, we present the first imaging results of Sgr A* with the EHT for the 2017 April 6 and 7 observations. In Section 2, we describe the EHT observations of Sgr A* in 2017 and their properties. In Section 3, we estimate additional properties of two major effects anticipated for Sgr A*, interstellar scattering and intrinsic intraday brightness variations, from nonimaging analysis to aid the imaging process. In Section 4, we provide a brief review of the employed imaging techniques. In Section 5, we describe the process for synthetic data generation for the imaging parameter surveys outlined in Section 6. These surveys provide a set of imaging parameters that are used to produce Sgr A* images. In Section 7, we describe the resulting Sgr A* images for four different imaging

pipelines and assess their properties and uncertainties. In Section 8, we extract source parameters from Sgr A* images via ring parameter fitting. In Section 9, we utilize dynamical imaging and geometric modeling techniques to explore and characterize potential azimuthal time variations in the data. We summarize our results in Section 10.

2. Observations and Data Processing

In this section we describe the EHT observations of Sgr A* performed in 2017 April (Section 2.1), the data reduction (Section 2.2), and overall data properties (Section 2.3). A description of interferometric measurements and associated data products is provided in M87* Paper IV for reference.

2.1. EHT Observations

The EHT observed Sgr A* with eight stations at six geographic sites on 2017 April 5, 6, 7, 10, and 11. The participating radio observatories are the phased Atacama Large Millimeter/submillimeter Array (ALMA) and APEX in the Atacama Desert in Chile, the JCMT and the phased Submillimeter Array (SMA) on Maunakea in Hawai'i, the SMT on Mt. Graham in Arizona, the IRAM 30 m (PV) telescope on Pico Veleta in Spain, the Large Millimeter Telescope Alfonso Serrano (LMT) on the Sierra Negra in Mexico, and the South Pole Telescope (SPT) in Antarctica. The observations of Sgr A* were interleaved with two AGN calibrator sources, the quasars NRAO 530 and J1924–2914. Scientific analysis of the observations of calibrators will be presented in future publications (Issaoun et al. 2022; S. Jorstad et al. 2022, in preparation). The geocentric coordinates for each of the telescopes are presented in Table 2 of M87* Paper II.

The 2017 VLBI data were recorded in two polarizations and two frequency bands at a total data rate of 32 Gbps (for 2-bit sampling). All sites recorded two 2 GHz wide frequency windows centered at 227.1 and 229.1 GHz (low and high band, respectively). An extensive description of the EHT array setup, equipment, and station upgrades leading up to the 2017 observations is provided in M87* Paper II. All sites except ALMA and JCMT recorded dual circular polarization (RCP and LCP). ALMA recorded dual linear polarization, subsequently converted to a circular basis via the CASA-based software package `PolConvert` (Martí-Vidal et al. 2016; Matthews et al. 2018; Goddi et al. 2019), and JCMT recorded a single circular polarization (the recorded polarization component varied from day to day). Since JCMT recorded a single circular polarization, baselines to JCMT use the available parallel-hand component (RR or LL) visibilities to approximate Stokes $\mathcal{I}$ ("pseudo $\mathcal{I}$," $\mathcal{I} \equiv RR \equiv LL$). This is consistent with the Stokes $\mathcal{V} \equiv 0$ assumption taken for the data calibration, justified by the expected $|\mathcal{I}| \gg |\mathcal{V}|$ relation (Muñoz et al. 2012; Goddi et al. 2021). A small but detectable amount of intrinsic circular source polarization is present in our observations, which we account for in the systematic error analysis (Paper II).

2.2. Data Reduction

Following the correlation of the data recorded at different sites, instrumental bandpass effects and phase turbulence introduced by Earth's atmosphere were corrected using established fringe-fitting algorithms (M87* Paper III). We use two independent software packages, the CASA-based (McMullin et al. 2007) rPICARD pipeline (Janssen et al. 2019) and the HOPS-based (Whitney et al. 2004) EHT-HOPS pipeline (Blackburn et al. 2019). The mitigation of the atmospheric phase variation allows for coherent averaging of the data in order to build up signal-to-noise ratio (S/N) without substantial losses from decoherence. Instrumental RCP/LCP phase and delay offsets were corrected by referencing fringe solutions to ALMA, calibrated with `PolConvert` (Martí-Vidal et al. 2016). The assumption of Stokes $\mathcal{V} = 0$ on VLBI baselines is taken for the RCP/LCP gain calibration. Following the band averaging in frequency, data were amplitude-calibrated using station-specific measurements of the system equivalent flux density and time-averaged in 10 s segments. Stations with an intrasite partner (i.e., ALMA, APEX, SMA, and JCMT) were subsequently "network-calibrated" (M87* Paper III; Blackburn et al. 2019) to further improve the amplitude calibration accuracy and stability via constraints among redundant baselines. Polarimetric leakage is not corrected for this Stokes $\mathcal{I}$ analysis, but rather included as a source of systematic uncertainty in the parallel-hand visibilities (Paper II).

The data processing pipeline has been slightly updated with respect to the one described in M87* Paper III. Some notable changes include a recorrelation of the data following setting changes at ALMA and more accurate sky coordinates of Sgr A*; updated amplitude calibration (most notably for LMT and SMA) using more accurate measurements of the telescope aperture efficiency, found to be variable across the campaign; stronger polarimetric calibration assumptions ($\mathcal{V} = 0$); time-variable network calibration of Sgr A* using ALMA and SMA connected-element light curves (Wielgus et al. 2022); and a time-dependent transfer of the antenna gains to the visibility amplitudes, following the analysis of the data from the calibrators.

After the calibration using data reduction pipelines described by M87* Paper III and Paper II, additional steps were taken to mitigate specific data issues related to poorly constrained LMT gains and JCMT coherence losses. Following the source size constraints derived in Section 5.1.5 of Paper II, LMT amplitude gains have been pre-corrected assuming the 60 μas source size seen by the baselines shorter than 2 Gλ (only SMT–LMT). Visibility phases on JCMT baselines were stabilized by calibrating phase on an intrasite JCMT–SMA baseline to zero degrees, in agreement with the unresolved point-source visibility phases seen, for similar baseline lengths, in the intra-ALMA observations. A detailed description of the theoretical background from visibilities to images is presented in Thompson et al. (2017), M87* Paper IV, and Blackburn et al. (2020), as well as in the Appendix of Paper IV.

2.3. Data Properties

2.3.1. General Aspects of Sgr A* Data

The highly sensitive phased-ALMA array participated in three of the five observing days, 2017 April 6, 7, and 11. April 7 is the only day that additionally includes PV observations of Sgr A* and is therefore the day with the longest observation duration, the largest number of detections, and the best overall (u, v)-coverage, as shown in Figure 1. On April 11 an X-ray flare was reported shortly before the start of the EHT observations (Paper II). Strongly enhanced flux density variability is seen in the light curves on that day (Wielgus et al. 2022), possibly posing difficulties for the static imaging

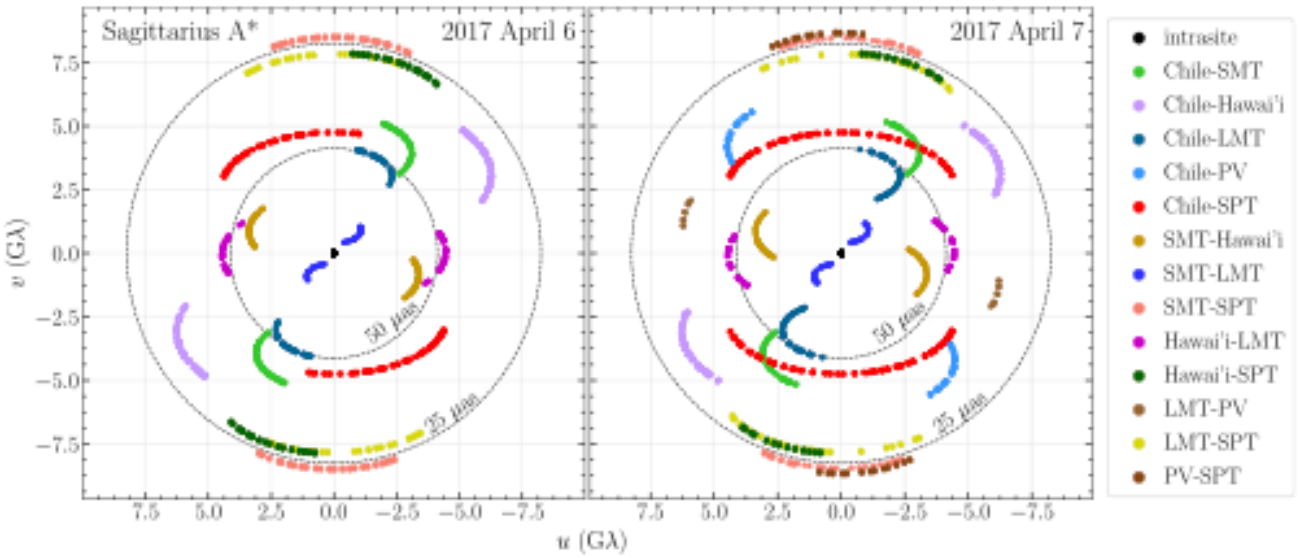


Figure 1. (u, v)-coverage of the EHT observations of Sgr A* on 2017 April 6 and 7, from the HOPS data set. Each point represents scan-averaged data; both bands are shown. "Chile" represents the stations ALMA and APEX. "Hawai'i" represents the stations SMA and JCMT. Dashed circles indicate the fringe spacing of 50 and 25 μas.

Table 1
Metrics of EHT Angular Resolution for the 2017 Observations of Sgr A* for the 229.1 GHz Band

	FWHM_{maj} (μas)	FWHM_{min} (μas)	P.A. (deg)
Minimum Fringe Spacing $1/\lvert\boldsymbol{u}\rvert_{\text{max}}$ (All Baselines)			
Apr. 6	24.2	...	...
Apr. 7	23.7	...	...
Minimum Fringe Spacing (ALMA Baselines)			
Apr. 6–7	28.6	...	...
CLEAN Beam (Uniform Weighting)			
Apr. 6	24.8	15.3	67.0
Apr. 7	23.0	15.3	66.6
CLEAN Restoring Beam (Used in This Paper)			
Apr. 6–7	20	20	...

Note. In order to avoid asymmetries introduced by restoring beams, and to homogenize the images among epochs, we adopt a circular Gaussian restoring beam with 20 μas FWHM for all CLEAN reconstructions.

of the April 11 data set. These constraints motivate utilizing the less variable 2017 April 7 data as the primary data set for static image reconstruction, with the April 6 observations as a secondary validation data set. Analysis of the remaining 2017 EHT observations of Sgr A* will be presented elsewhere.

In Figure 1, the (u, v)-coverage on 2017 April 7 is shown to be asymmetric, with the longest baselines along the north–south direction. The shortest baselines in the EHT are intrasite and sensitive to arcsecond-scale structure (i.e., the SMA and JCMT are separated by 0.16 km; ALMA and APEX are separated by 2.6 km). In contrast, the longest baselines are sensitive to microarcsecond-scale structure (see Table 1). The ~8.7 Gλ detections on PV–SPT and SMT–SPT baselines are among the longest published projected baseline lengths obtained with ground-based VLBI, alongside the recent EHT 3C 279 results of Kim et al. (2020), slightly longer than the longest baselines in the EHT observations of M87* (8.3 Gλ; M87* Paper IV). Sgr A* was detected on all baselines between stations with mutual visibility, leading to the April 7 (u, v)-coverage approaching the best one theoretically possible with the EHT 2017 array. Table 1 shows the angular resolutions derived from (u, v)-coverage on both April 6 and 7 data.

In the top panel of Figure 2 we show the S/N of the Sgr A* observations as a function of projected baseline length, for the coherent averaging time of 120 s. The split in S/N distributions at various projected baseline lengths is due to the difference in sensitivity for the colocated Chile sites ALMA and APEX, with ALMA baselines yielding detections stronger by about an order of magnitude. In the bottom panel, we show the visibility amplitude (correlated flux density in units of Jy) for Sgr A* as a function of projected baseline length after applying full data calibration.

The fully calibrated visibility amplitudes exhibit a prominent secondary peak between two local minima. The first minimum is located at ~3.0 Gλ and is probed by the Chile–LMT north–south baselines on 2017 April 7. On April 6 recording started about 2 hr later, thus missing the relevant detections, as shown in Figure 3. The second minimum appears at ~6.5 Gλ, probed by the Chile–Hawai'i baselines on both 2017 April 6 and 7. Overall amplitude structure of the source appears to be consistent across both days, which is particularly well visible in the fully calibrated, light-curve-normalized data sets shown in Figure 3, as the light-curve normalization procedure strongly suppresses the large-scale source intrinsic variability (Broderick et al., 2022). The observed local visibility amplitude minima can be associated with the nulls of the Bessel function J_0, corresponding to the Fourier transform of an infinitely thin ring. For a ring that is 54 μas in diameter we would obtain local amplitude minima at 2.92 and 6.71 Gλ. This is illustrated in the bottom panel of Figure 2, where an analytic Fourier transform of an infinitely thin ring blurred with a 23 μas FWHM Gaussian kernel is shown with dashed lines.[149] While a blurred ring

[149] We note that plotting the visibilities from a thin ring over the measured visibility amplitudes is meant only to guide the eye in observing the double-null structure. Other geometries, such as a disk model, can also align with the double-null structure seen in visibility amplitudes (see Figure 7). Detailed fitting of different simple geometries to the visibilities is performed in Paper IV.

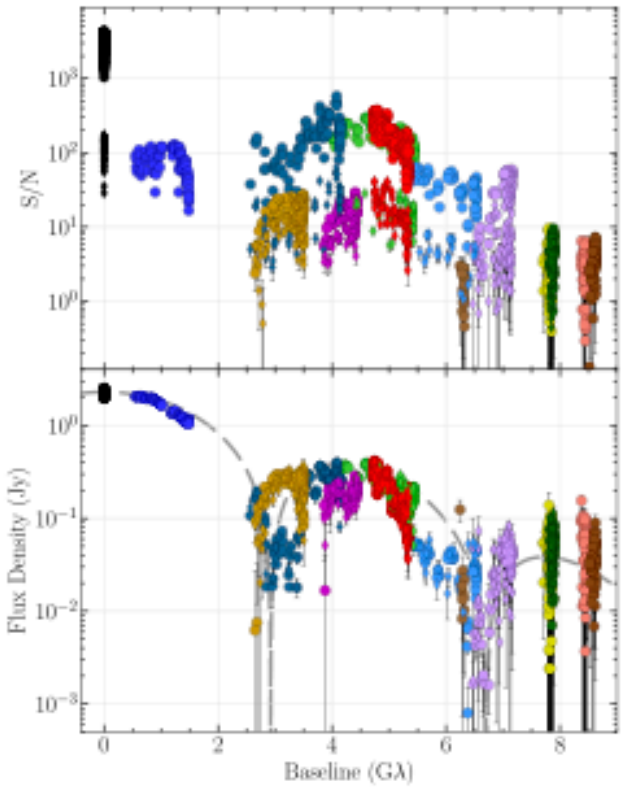


Figure 2. Top: S/N as a function of the (u, v)-distance (projected baseline length). Low-band HOPS data recorded on 2017 April 7, averaged in 120 s segments, are shown. The data points are color-coded with the baseline, following Figure 1. Bottom: fully calibrated visibility amplitude data. A model corresponding to a thin ring with a 54 μas diameter, blurred with a 23 μas FWHM circular Gaussian, is overplotted for a reference (dashed curve).

model roughly captures the dependence of visibility amplitudes on projected baseline length, there is also a clear indication of the source asymmetry, manifesting as amplitude differences between the Chile–LMT and SMT–Hawai'i baselines at the first minimum, probing the same range in projected baseline length (∼2.5–3.5 Gλ) in orthogonal directions. There is also a deficit of flux density with respect to the simple ring model at projected baseline lengths of ∼5–6 Gλ.

Finally, we detect very clear and unambiguous nonzero closure phases, which are indicative of source asymmetry. With multiple independent triangles and high S/N, these data sets offer insight into the source phase structure greatly surpassing that of any previous mm-wavelength observations (Fish et al. 2016; Lu et al. 2018). In Figure 4, we show examples of closure phases on several triangles exhibiting various degrees of probed asymmetry and inter/intraday time variability. Closure phases on ALMA–SMT–SMA and ALMA–LMT–SMA triangles immediately show interday variability of the source structure. In the case of Sgr A*, intrinsic source variability is expected also on timescales as short as minutes, adding to the closure phase intraday variability caused by the nontrivial average structure of the source (see Section 3.2). Very long baselines, such as LMT–SPT in the ALMA–LMT–SPT triangle, are additionally affected by the presence of refractive scattering structure (see Section 3.1).

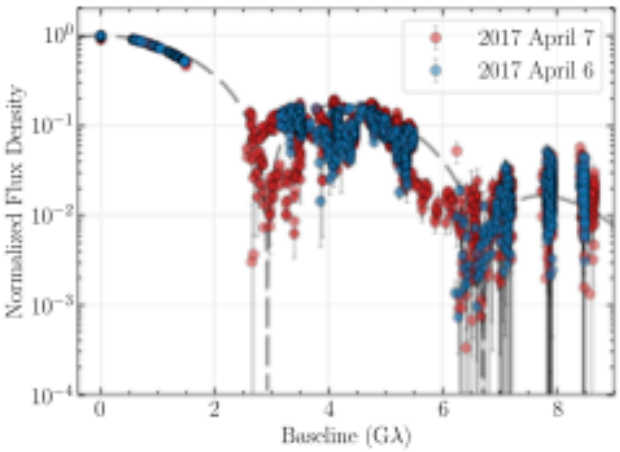


Figure 3. Comparison of the light-curve-normalized flux density measurements on 2017 April 6 and 7 in the fully calibrated HOPS data, averaged in 120 s segments and between low and high bands.

2.3.2. Station Gain Uncertainties and Nonclosing Errors

Typically, the antenna gains and sensitivities as a function of elevation are derived using polynomial fits to opacity-corrected antenna temperature measurements from quasars and solar system objects, tracked over a wide range of elevations. Residual errors in the characterization of these antenna gains lead to corruptions in the flux calibration of the visibilities. Quantifying these effects enables us to disentangle astrophysical variability of Sgr A* from apparent flux variations caused by the imperfect calibration. In this work, we mitigate the systematic gain errors in the Sgr A* data sets based on the analysis of the calibrator sources (J1924–2914 and NRAO 530), which remained stationary in their source structure and flux density on relevant timescales. The detailed procedure to estimate the antenna gains from the two calibrators is described in Section 5.1.3 in Paper II. In particular, it is shown that the a priori gains are 5%–15% for all baselines, except intrasite baselines (∼1%) and those including the LMT (∼35%).

In addition to the thermal noise and the antenna gain uncertainties, we also estimate the nonclosing errors in the data, based on deviations of the trivial closure quantities from zero, as well as from the observed inconsistencies in the distributions of closure quantities between bands and polarizations (M87* Paper III; Paper II). These nonclosing errors are expected to arise from the presence of a small circular polarization component, as well as from uncorrected polarimetric leakages, and other systematic errors, such as residual bandpass effects. For Sgr A*, the nonclosing errors are estimated to be 2° in closure phase and 8% in log closure amplitude (Paper II). Assuming that the errors are baseline independent, these translate to 1° systematic nonclosing uncertainties in visibility phases and 4% systematic nonclosing uncertainties in visibility amplitudes, on the top of the uncertainties related to the amplitude gain calibration. We found that the RR − LL discrepancies in closure quantities are more significant for Sgr A* than in the case of the calibrators. This hints at an intrinsic source property, possibly a contribution from a small circular polarization component. This is consistent with Goddi et al. (2021), reporting ∼1% circular polarization in the simultaneous ALMA-only data.

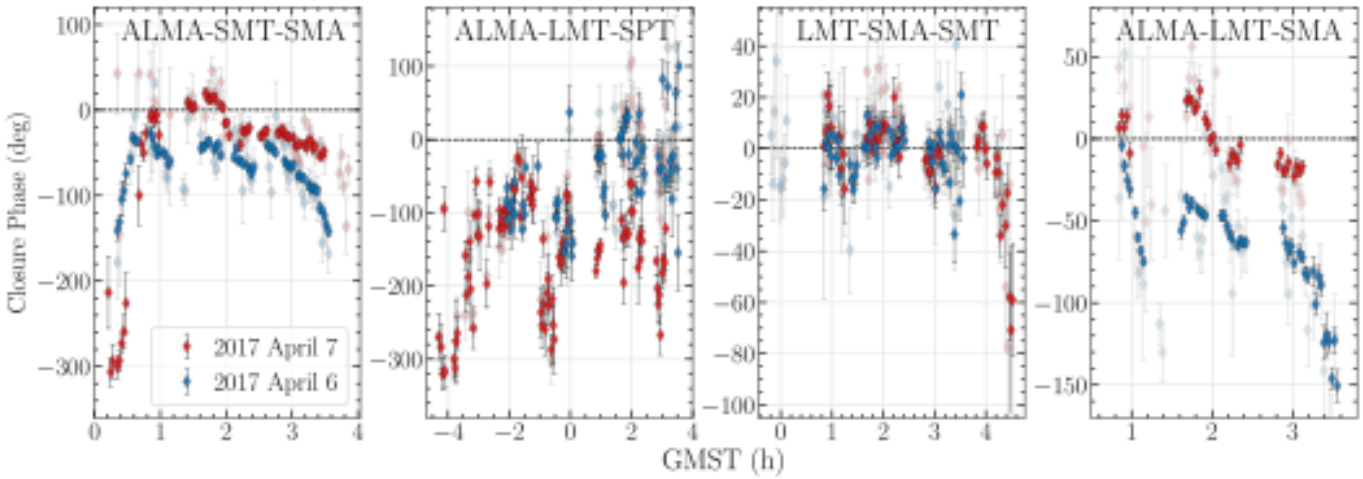


Figure 4. Examples of closure phases of Sgr A*, observed on 2017 April 6 and 7. Semitransparent points correspond to measurements on redundant APEX/JCMT triangles. Data reduced with the HOPS pipeline, integrated in 120 s segments, averaged between both bands, are shown. No significant differences are seen with respect to the CASA-based calibration.

3. Mitigation of Scattering and Time Variability

The imaging of Sgr A* at 230 GHz with the EHT is challenged by two important effects: interstellar scattering and short-timescale variability. In this section, we introduce strategies for mitigating the effects of scattering (Section 3.1) and intrinsic variability (Section 3.2) adopted in this work.

3.1. Effects of Interstellar Scattering

Fluctuations in the tenuous plasma's electron density along the line of sight cause scattering of the radio waves from Sgr A*. The scattering properties of Sgr A* can be well described by a single, thin, phase-changing screen $\phi(\boldsymbol{r})$, where $\boldsymbol{r}$ is a two-dimensional vector transverse to the line of sight. The electron density fluctuation on the phase screen is typically characterized by a single power-law spectral shape between the outer ($r_{\rm out}$) and inner ($r_{\rm in}$) scale as $Q(\boldsymbol{q}) \propto |\boldsymbol{q}|^{-\beta}$, where $\boldsymbol{q}$ is the wavevector of the propagating radio wave and a Kolmogorov spectrum of density fluctuations gives $\beta = 11/3$ (Goldreich & Sridhar 1995). The statistical effects of the scattering can then be related to a spatial structure function $D_\phi(\boldsymbol{r}) = \langle[\phi(\boldsymbol{r}+\boldsymbol{r}_0) - \phi(\boldsymbol{r}_0)]^2\rangle_{\boldsymbol{r}_0} \propto \lambda^2$, where $\langle\cdot\rangle_{\boldsymbol{r}_0}$ denotes the ensemble average over $\boldsymbol{r}_0$.

The interstellar scattering of radio waves from Sgr A* is in the regime of strong scattering, where scintillation is dominated by two distinct effects, diffraction and refraction, attributed to widely separated scales (see Narayan 1992; Johnson & Gwinn 2015). Diffractive scintillation arises from fluctuations on the scale of the phase coherence (or diffractive scale) given by $D_\phi(\boldsymbol{r}) \sim 1$. It causes rapid temporal variations on a timescale much shorter than 1 s for Sgr A*, which is also much shorter than the integration time of radio observations. As a result, radio observations measure ensemble averages of the diffractively scattered structure, appearing as the intrinsic structure blurred with the scattering kernel (see Section 3.1.1).

Refractive scintillation arises from fluctuations on the scale of the scattering kernel much larger than the phase coherence length in the strong scattering regime. For Sgr A*, the refractive scintillation causes temporal variations of the source images over a timescale of ~1 day at 1.3 mm (e.g., Johnson et al. 2018)—longer than the typical length of radio observations including our EHT observations. Consequently, a single realization of refractive scintillation will be observed by the EHT over an observing run; this will appear as an angular-broadened (i.e., diffractively scattered) source structure with compact substructure caused by refractive scintillation (Narayan & Goodman 1989; Goodman & Narayan 1989; Johnson & Gwinn 2015; Johnson & Narayan 2016).

A brief introduction of the expected scattering properties in the EHT 2017 data is described in Section 5.1 of Paper II. Here we describe the scattering mitigation strategy for the effects of angular broadening by diffractive scattering (Section 3.1.1) and substructure induced by refractive scattering (Section 3.1.2). To describe scattering effects on Sgr A*, we use a theoretical framework of these scattering effects developed by Psaltis et al. (2018), whose model parameters have been observationally studied by Johnson et al. (2018), Issaoun et al. (2019), Issaoun et al. (2021), and Cho et al. (2022). For general background and reviews on interstellar scattering, see Rickett (1990), Narayan (1992), or Thompson et al. (2017).

3.1.1. Mitigation of Angular Broadening

Angular broadening is described by a convolution of an unscattered image with a scattering kernel, or equivalently by a multiplication of unscattering interferometric visibilities by the appropriate Fourier-conjugate kernel. The Fourier-conjugate kernel is given by $\exp\left[-\frac{1}{2}D_\phi(\boldsymbol{b}/(1+\mathcal{M}))\right]$, where $\boldsymbol{b}$ is the baseline vector of the interferometer and $\mathcal{M}$ is the magnification of the scattering screen given by the ratio of the observer-to-screen distance to the screen-to-source distance. Interferometric measurements of Sgr A* with the EHT at the observing wavelength of 1.3 mm are primarily obtained on long baselines of $|\boldsymbol{b}| \gtrsim (1+\mathcal{M})r_{\rm in}$, or equivalently on angular scales of $\theta \lesssim \lambda/(1+\mathcal{M})r_{\rm in}$, where $r_{\rm in}$ is the inner scale of the fluctuations. In this regime, the angular broadening is affected by the power-law density fluctuations on scales between the inner and outer scales, giving the phase structure function of $D_\phi(\boldsymbol{r}) \propto \lambda^2|\boldsymbol{r}|^{-\alpha}$, where $\alpha = \beta - 2$.

In Figure 5, we show the scattering kernel in the visibility domains based on the scattering parameters in Johnson et al. (2018). Johnson et al. (2018) imply a near-Kolmogorov power-law spectral index $\beta \sim 3.38$ (or $\alpha \sim 1.38$), providing a non-Gaussian kernel more compact than the conventional Gaussian

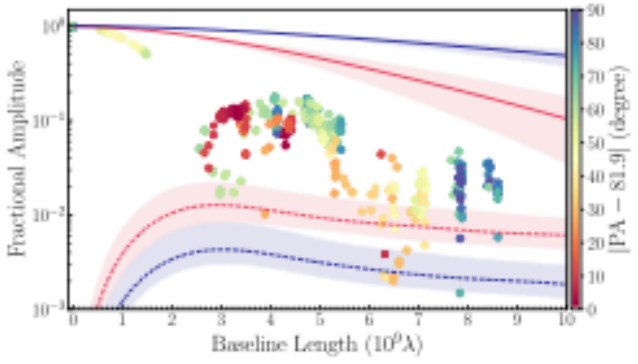


Figure 5. Projection of the diffractive scattering kernel (solid lines) and flux-normalized refractive noise amplitudes (dashed lines) at 1.3 mm based on the scattering model of Johnson et al. (2018), overlaid on light-curve-normalized Sgr A* data (low band on April 7). The red and blue lines correspond to the major and minor axes of the scattering, respectively. The associated shaded area indicates the 3σ uncertainty of the scattering model in Johnson et al. (2018). Sgr A* data are colored by their PA difference from the major axis of the scattering kernel—as the points change from red to blue, the (u, v) coordinates move from being closer to the major to minor scattering axis. Regardless of the PA, Sgr A* amplitudes appear to more rapidly decrease than the scattering kernel, indicating that the intrinsic structure is well resolved against the diffractive angular broadening effects. Additionally, most Sgr A* amplitudes are above the refractive floor. Thus, refractive effects should only dominate for a small amount of data above ~6 Gλ.

kernel. Consequently, the angular broadening effect, i.e., multiplication of the intrinsic visibilities with the Fourier-conjugate kernel of scattering, causes a slight decrease in visibility amplitudes and therefore also the S/N by a factor of a few at maximum.

Angular broadening provides deterministic and multiplicative effects on the observed visibility. Therefore, it is invertible—they can be mitigated by dividing the observed visibility and associated uncertainties by the diffractive kernel visibility (often called *deblurring*; Fish et al. 2014). However, the actual interferometric measurements of Sgr A* with the EHT have contributions from substructure that arises from the refractive scattering, often referred to as the "refractive noise." The refractive noise is stochastic and additive and therefore not invertible (Johnson & Narayan 2016). In fact, since refractive effects are included in the diffractively blurred image, the refractive noise will be amplified by simply deblurring with the scattering kernel and will likely create artifacts in the reconstructions if not accounted for (Johnson 2016). To avoid effects from refractive noise, we expand the noise budgets of the visibility data prior to deblurring, as described in Section 3.1.2.

3.1.2. Mitigation of Refractive Scattering

The contribution of the refractive noise to the observed visibility is anticipated to be not dominant, except for a small fraction of data beyond ~6 Gλ (Figure 5).[150] Signature of the refractive noise, namely, the long, flat tail of the visibility amplitude at long baselines found in recent longer-wavelength observations of Sgr A* (Johnson & Gwinn 2015; Johnson et al. 2018; Issaoun et al. 2019, 2021; Cho et al. 2022), is not clearly seen in the EHT data. In this EHT regime, where the refractive substructure is not unambiguously constrained from data, it is challenging to apply complex strategies that account for the stochastic properties or explicitly recover the refractive screen (e.g., Johnson 2016; Johnson et al. 2018; Issaoun et al. 2019; Broderick et al. 2020a).

We instead mitigate the effect of refractive substructure by introducing error budget models that approximate the anticipated refractive noise. The visibility error budget is increased based on these models prior to the mitigation of angular broadening via deblurring (Section 3.1.1). In this work, we consider four base models that approximate the refractive noise budgets: (i) `Const`: a constant noise floor (e.g., 10 mJy) for all baselines motivated by the fact that the refractive noise has a mostly flat profile as a function of the baseline length (see Figure 5). (ii) `J18model1`: (u, v)-dependent noise floor based on the scattering model and parameters described in Psaltis et al. (2018) and Johnson et al. (2018). Since the scattered image is not unique, we have simulated hundreds of scattering realizations and generated corresponding synthetic data that match the (u, v)-coverage of the actual April 7 observation of Sgr A*. The refractive noise values are then computed by taking the standard deviation of the complex visibilities across different realizations. Since the refractive noise is also dependent on the intrinsic source structure, in this case we consider a circular Gaussian model with the second moment that matches the pre-imaging size constraints (see Paper II). (iii) `J18model2`: Same as `J18model1`, but using the average refractive noise value of seven geometric models as possible intrinsic source structures (see Section 5). (iv) Considering not only the standard deviation of the refractive effects but also their correlations via a covariance matrix.

Note that using all the information encoded in the covariance matrix (not only in the variance of the refractive noise variables) will provide a better approximation of the refractive noise. However, the short time cadence and the redundant baselines in our data make this covariance matrix noninvertible and thus difficult to use in imaging. For the remaining three refractive noise models, we compute the complex visibility χ^2 for a suite of synthetic data (corrupted only by thermal noise along with scattering effects) based on seven geometric models of the intrinsic source structure (Section 5),

$$\chi^2(\sigma_{\mathrm{ref},i}) = \frac{1}{2N_{\mathrm{vis}}} \sum_{i=1}^{N_{\mathrm{vis}}} \frac{|V_i - V_{\mathrm{ea}}|^2}{\sigma_{\mathrm{th},i}^2 + \sigma_{\mathrm{ref},i}^2}, \quad (1)$$

where V_i are the data visibilities for each scattering realization, V_{ea} are the ensemble-averaged visibilities (i.e., corresponding to the image experiencing only diffractive scattering), $\sigma_{\mathrm{th},i}$ is the thermal noise, and $\sigma_{\mathrm{ref},i}$ is the corresponding refractive noise budget for each strategy. The χ^2 metric provides us with a statistic on how well the ensemble average image represents the synthetic data after taking into account the different modeled refractive noise budgets. Figure 6 shows a comparison of the different χ^2 distributions for 400 realizations of synthetic data for every scattering mitigation strategy. For `J18model1` we have derived a scaling factor to make the median of χ^2 of all models equal to 1 in order to overcome the dependence of the refractive noise level on the intrinsic source structure (see Appendix B).

Figure 6 demonstrates that the (`Const`, `J18model1`, and `J18model2`) refractive noise models result in reasonable χ^2

[150] The refractive noise in Figure 5 is estimated for a circular Gaussian source with an intrinsic FWHM of 45 μas under the same condition of interstellar scattering as constrained for Sgr A*. The size of the Gaussian is broadly consistent with the equivalent second moments of the geometric models in Section 5 that share similar visibility amplitude profiles with Sgr A* data.

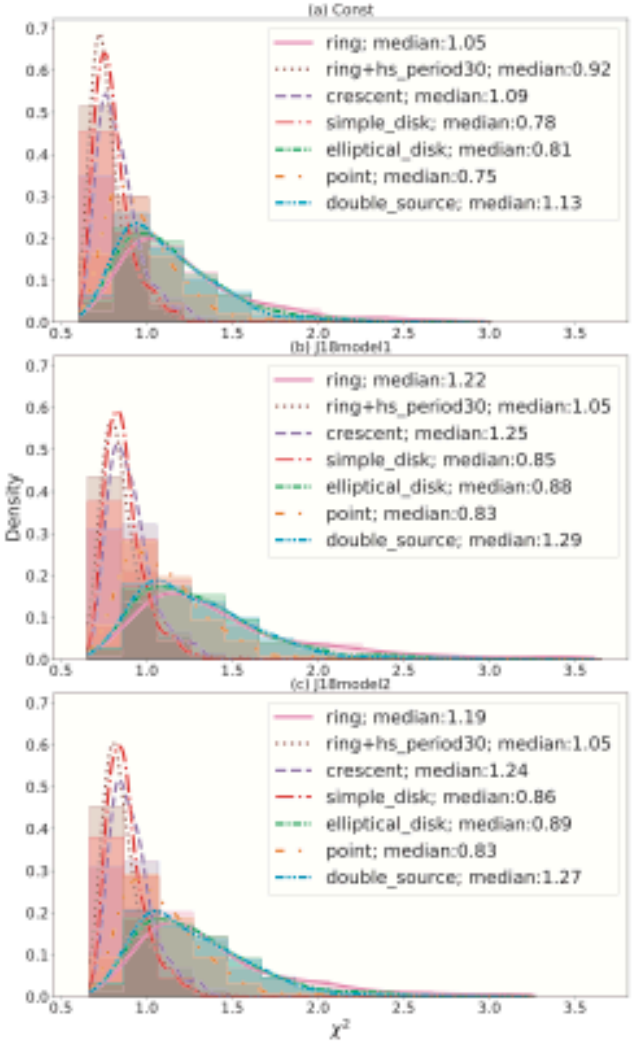


Figure 6. Reduced χ^2 distributions and their density function (in solid lines) for all geometric models of intrinsic source structure using the three different refractive noise models (`Const`, `J18model1`, and `J18model2`). For all noise models, the majority of the χ^2 values falls in the range between 0.5 and 2.0 (ideally $\chi^2 = 1.0$), which indicates that the ensemble average images provide reasonable fits to the simulated data after taking into account one of the proposed refractive noise budgets.

values for the simulated scattering realizations we tested, with a majority of the χ^2 values below 3.0 (ideally $\chi^2 = 1$). For this reason, in the rest of the paper we focus on the simpler `Const` and `J18model1` strategies.

3.2. Effects of Time Variability

With a gravitational timescale of only $GM/c^3 \approx 20$ s, Sgr A* is expected to be able to exhibit substantial changes in its emission structure on timescales of a few minutes or less. A single multihour observing track is thus sufficiently long for Sgr A* to significantly alter its appearance, potentially hundreds of times. Such structural variability violates a core assumption of Earth-rotation aperture synthesis—namely, that the source structure must remain static throughout the observation—and necessitates modifications to standard imaging practices.

3.2.1. Evidence for Source Variability in the Data

The EHT Sgr A* data contain unambiguous signatures of an evolving emission structure. On the largest spatial scales, the light curve varies at the ∼10% level on timescales of ∼hours (Wielgus et al. 2022). Variations in excess of those expected from thermal noise are seen on timescales as short as ∼1 minute. Over timescales typical of observation scans, ∼10 minutes, the degree of variation is on the order of 5% (Wielgus et al. 2022).

Direct evidence for short-timescale structural variations may be found in the evolution of closure quantities. Closure phases measured on certain triangles of baselines (e.g., ALMA–SMT–SMA) exhibit significant differences between April 6 and 7 (see Figure 4). The variations seen in the closure phases measured on multiple triangles show significant excesses, relative to thermal noise levels, as captured using the Q-metric statistic (Roelofs et al. 2017; Paper II).

Nonparametric estimates of the degree of variability as a function of baseline length may be generated by inspecting the visibility amplitudes directly. This is made possible by two fortuitous facts: first, the existence of crossing tracks, and second, that Sgr A* was observed on multiple days. As a result, many presumably independent realizations of the source structure may be compared. Practically, this is obtained by collecting visibility amplitudes in (u, v) bins, linearly detrending to remove the contribution from the static component of the image, and computing the mean and variance of the residuals. We average these estimates azimuthally to improve the significance of variability detection. For details on the procedure and validation examples, we direct the reader to Georgiev et al. (2022) and Paper IV.

In the case of Sgr A*, this nonparametric estimate produces a clear detection of variability that is significantly in excess of the expected thermal noise (Paper IV). Within the range of baseline lengths over which meaningful estimates can be produced, roughly $2\,\mathrm{G}\lambda < |u| < 6\,\mathrm{G}\lambda$, the observed excess variability is broadly consistent with that anticipated by GRMHD simulations, both in magnitude and in dependence on baseline length (Papers IV and V).

3.2.2. Strategies for Imaging Variable Data

Strategies for imaging in the face of source variability can be classified into one of three general categories:

1. Variability reconstruction, or "dynamic imaging," in which the evolution of the source emission structure is explicitly recovered during the imaging process. The output of this strategy is a movie of the source emission structure. We refer the reader to Section 4.4 for more discussion on methods for dynamic imaging.
2. Variability circumvention, or "snapshot imaging," in which standard image reconstruction is performed on segments of data ("snapshots") that are sufficiently short that the source may be approximated as static across them. The output of this strategy is a time series of static images.
3. Variability mitigation, or "variability noise modeling," in which the impact of structural changes in the visibilities is absorbed into an appropriately inflated error budget. The output of this strategy is a single static image of the source, indicative of the time-averaged image over this observation period.

In practice, for segments of data short enough that Sgr A* may be reasonably approximated as static, the (u, v)-coverage of the EHT is insufficient to support reliable snapshot image reconstruction (though more restrictive parameterizations of the source structure, such as permitted using geometric modeling, can still be applied; see Section 9 and Paper IV). However, because dynamic imaging enforces a degree of temporal continuity, it is able to leverage the information provided by densely covered intervals of time to augment the lack of information available during intervals of sparser coverage. Dynamic imaging can thus be thought of as a generalization of both standard (static) imaging and snapshot imaging, with the former being equivalent to dynamic imaging with maximal temporal continuity enforcement and the latter being equivalent to dynamic imaging with no temporal continuity enforcement at all. Because dynamic imaging falls in between these two extremes, it can potentially recover reliable source structure in regions where the data are both too variable for standard imaging and too sparse for snapshot imaging. Our efforts to perform dynamic imaging in the most densely (u, v)-covered regions of data are described in Section 9.

The third strategy listed above—the variability noise modeling approach—permits static images to be reconstructed even from time-variable data. Depending on the specifics of the implementation, the recovered image captures some representation of "typical" source structure. Imaging with variability noise modeling requires that the error budget of the data be first inflated in a way to capture the statistics—or "noise"—of the source variability. The specific form of the noise model we use in this work is a broken power law, for which the variance $\sigma^2_{\rm var}$, as a function of baseline length $|u|$, takes the form

$$\sigma^2_{\rm var}(|u|) = a^2\left(\frac{|u|}{4\ {\rm G}\lambda}\right)^c \frac{1 + [(4\ {\rm G}\lambda)/u_0]^{b+c}}{1 + (|u|/u_0)^{b+c}}. \quad (2)$$

Here $|u| \equiv \sqrt{u^2 + v^2}$ is the dimensionless length of the baseline located at (u, v), u_0 is the baseline length corresponding to the break in the power law, a is the variability noise amplitude at a baseline length of 4 Gλ, and b and c are the long- and short-baseline power-law indices, respectively. Equation (2) represents the variance that is associated with structural variability after removing the mean and normalizing by the light curve; see Paper IV for details.

By adding the variability noise given by Equation (2) in quadrature to the uncertainty of every visibility data point, the image becomes constrained to fit each data point to only within the tolerance permitted by the expected source variability. This parameterized variability noise model is generic and can explain well a wide range of source evolution, including complicated physical GRMHD simulations of Sgr A* (Georgiev et al. 2022).

Paper IV presents a nonparametric analysis of Sgr A*'s variability, which is further inspected to provide ranges of broken power-law model parameters that fit Sgr A* data (see Georgiev et al. 2022). As explained in Paper IV, given the baseline coverage of the 2017 Sgr A* campaign, little traction is found on the location of the break, u_0, and the short-baseline power law, c. However, the amplitude a is well constrained with an interquartile range from 1.9% to 2.1%. Similarly, the long-baseline power law, b, is modestly constrained, with interquartile range 2.2–3.2. These interquartile ranges are used to provide approximate priors on the variability noise that should be considered during static imaging reconstruction; values employed are listed in Table 2 for Sgr A*, as well as for a number of synthetic data sets described in Section 5. In the case of the latter, theoretical considerations imply that under very general conditions $c \approx 2$, and thus we adopt a general prior of [1.5, 2.5]. For the CLEAN (Section 4.1) and regularized maximum likelihood (RML; Section 4.2) imaging surveys described in Section 6, variability noise models in the identified Sgr A* range are added to the visibility noise budget before static imaging (for both synthetic and real Sgr A* data). For the THEMIS imaging method (Section 4.3), the parameters of the noise model are fit simultaneously with the image structure, subject to the data-set-specific values in Table 2 being used to define a uniform prior over each data set.

Table 2
Variability Noise Model Ranges Used for Static Imaging

Source[a]	a (%)	b[b]	u_0[a] (Gλ)
Sgr A*	[1.9,2.1]	[2.2,3.2]	[0.37,1.45]
Ring	[0.7,0.8]	[1.9,3.0]	[0.18,1.06]
Ring+hs	[0.6,0.7]	[3.0,4.1]	[0.33,0.90]
Crescent	[0.9,1.1]	[2.4,3.4]	[0.29,1.15]
Simple disk	[0.6,0.7]	[2.3,3.3]	[0.18,0.82]
Elliptical disk	[0.6,0.7]	[2.1,3.0]	[0.16,0.77]
Point	[0.8,0.9]	[1.8,3.1]	[0.32,2.67]
Double	[0.8,0.9]	[1.9,2.9]	[0.19,1.09]
GRMHD	[2.4,2.7]	[2.6,3.8]	[0.56,1.78]

Notes.
[a] All sources include high and low bands on observation days April 5, 6, 7, and 10.
[b] Interquartile (25th–75th percentile) ranges based on nonparametric analysis of suprathermal fluctuations in the visibility amplitudes on a per-scan basis.

4. Imaging Methods for Sgr A*

Recovering an image of Sgr A* from interferometric measurements amounts to solving an inverse problem. This inverse problem is challenging because of four primary reasons: (1) the interferometer incompletely samples the visibility domain, (2) there is significant structured noise included in the visibility measurements, (3) the source structure is evolving over the duration of the observation, and (4) the source is both diffractively and refractively scattered. The methods used in M87* Paper IV to recover an image of M87* from interferometric measurements had to address challenges 1 and 2 above; challenges 3 and 4 are unique to the rapidly evolving Sgr A* source, which we observe through the ISM. Strategies to mitigate the effects of scattering and time variability are discussed in detail in Section 3. In this section we assume that the data have already been modified by the appropriate descattering strategy and variability noise budget prior to imaging.[151]

To choose among the possible Sgr A* images, additional information, assumptions, or constraints must be included when solving the inverse problem. We broadly categorize imaging algorithms into three methodologies: CLEAN, RML, and Bayesian posterior sampling. We summarize these

[151] In THEMIS static imaging and dynamic imaging the variability noise is not included before imaging. In THEMIS the variability noise budget is estimated along with the image. In dynamic imaging variability noise is not included.

approaches here, but we refer the reader to M87* Paper IV for a more complete discussion of static imaging methods for EHT data. Additionally, we introduce the idea of dynamic imaging, which aims to reconstruct a movie rather than a single static image over an observation.

4.1. CLEAN Static Imaging

Traditionally, radio interferometric images have been made using nonlinear deconvolution algorithms of the CLEAN family (e.g., Högbom 1974; Schwarz 1978; Clark 1980; Schwab 1984; Cornwell et al. 1999; Cornwell 2008). These algorithms iteratively deconvolve the effects of the limited sampling of the (u, v)-plane, i.e., the interferometer's point-source response (also known as *dirty beam*), from the inverse Fourier transform of the visibilities (*dirty image*).

The classical CLEAN algorithm assumes that the sky brightness distribution can be represented as a collection of point sources. The imaging process consists of rounds of locating the brightness peak in the dirty image, generating a point source (CLEAN component) at this location with an intensity of some fraction of the map peak, and either convolving the CLEAN component with the dirty beam and subtracting it from the dirty image (Högbom 1974; Clark 1980) or subtracting the CLEAN components directly from the ungridded visibilities (Schwab 1984). This is continued until some specified cleaning stopping criterion is reached. One can supplement the process by restricting the area in which the peaks are searched (so-called CLEAN windows). This limits the parameter space in fitting and is especially important for data with sparse (u, v) sampling. The final image is made by convolving the obtained set of CLEAN components with a Gaussian restoring beam to smooth out the higher spatial frequencies and adding the last residual image to represent the remaining noise.

After image deconvolution, further improving of the image quality can be achieved using *self-calibration*, which uses the current image estimate to apply a correction to amplitude and phase information. Self-calibration is usually applied as part of an iterative process following each CLEAN iteration.

In this work we implement the CLEAN method using the `DIFMAP` (Shepherd et al. 1995; Shepherd 1997, 2011) pipeline described in Section 6 and Appendix D.1.

4.2. RML Static Imaging

The general approach in RML static imaging methods is to find an image, $\hat{I}$, that minimizes a specified objective function. As described further in M87* Paper IV, by using $\chi^2(I, V)$ as a measure of the inconsistency of the image, I, with the measurements, V, we can specify the objective function:

$$J(I) = \sum_{\text{data terms}} \alpha_D \chi^2_D(I, V) - \sum_{\text{regularizers}} \beta_R S_R(I). \qquad (3)$$

In this expression, the χ^2_D values are goodness-of-fit functions corresponding to the data product D, and the S_R values are regularization terms corresponding to the regularizer R. Maximum entropy (Narayan & Nityananda 1986; Chael et al. 2016), total variation, and sparsity priors (Wiaux et al. 2009a, 2009b; Honma et al. 2014; Akiyama et al. 2017b) have all been used to define $S_R(I)$ and have been demonstrated in the interferometic imaging of M87* (M87* Paper IV). The $\chi^2_D(I, V)$ and $S_R(I)$ terms often have different preferences for the "best" image and compete against each other in minimizing $J(I)$. Their relative impact in this minimization process is specified with the hyperparameters α_D and β_R.

This expression can be interpreted probabilistically when $\exp(-\sum \alpha_D \chi^2_D(I, V)) \propto p(V|I)$ and $\exp(\sum \beta_R S_R(I)) \propto p(I)$. In this case, minimizing the cost function $J(I)$ is equivalent to maximizing the log-posterior $\log p(I|V)$. Not all regularizer cost functions S_R correspond to a formal probability distribution. Nonetheless, while not all RML methods have a probabilistic interpretation, their formulation leads to a similar optimization setup.

For the EHT, RML methods have an advantage of being able to naturally constrain closure data products that are insensitive to atmospheric noise that corrupts EHT visibilities (Bouman et al. 2016; Chael et al. 2016; Akiyama et al. 2017b; Chael et al. 2018). In this work we implement RML methods using the `eht-imaging` (Chael et al. 2016, 2018, 2022) and `SMILI` (Akiyama et al. 2017a, 2017b; Moriyama et al. 2022) pipelines described in Section 6 and Appendices D.2 and D.3.

4.3. Bayesian Full Posterior Static Imaging

A fully Bayesian approach to imaging is a natural extension of the RML approach to image reconstruction. The primary output of Bayesian methods is an image posterior, i.e., not only a single "best-fit" image but also the family of images that are consistent with the underlying visibility data. In this way, the Bayesian image posterior encapsulates both the typical image reconstruction and its aleatoric (e.g., statistical) uncertainty, permitting quantitative analyses of the robustness of image features, array calibration quantities, and the relationships between each (see, e.g., Arras et al. 2019; Broderick et al. 2020b; Pesce 2021; Sun & Bouman 2021; M87* Paper VII).

We employ the general modeling framework THEMIS, developed specifically to compare parameterized models with the VLBI data produced by the EHT (Broderick et al. 2020a). The image model consists of three conceptually distinct components: a description of the brightness distribution on the sky, the variable complex gains at each station, and the additional "noise" associated with intraday structural variability in the source. Scan-specific complex station gains and variability "noise" parameters are recovered and marginalized over simultaneously with image exploration. Details on the model construction, adopted priors, sampling methods, and fidelity criteria are collected in Appendix A and are only briefly summarized here.

We make use of the adaptive splined raster models within THEMIS, consisting of a set of brightness control points that may vary in brightness on an adjustable rectilinear grid. In practice, only a handful of resolution elements are required (see Section 6.3 and Appendix D.4), and full images are produced via an approximate cubic spline (Broderick et al. 2020b). The dimensions of the raster, N_x and N_y, are selected based on the Bayesian evidence as discussed further in Section 6.3.

The combined parameter space, composed of the brightness control points, raster size and orientation, complex station gains, and noise model parameters, is sampled via a parallelly tempered, Hamiltonian Monte Carlo scheme, producing a chain of candidate images and ancillary quantities that are distributed according to their posterior probability, $p(I|V)$. In practice, the sampler must explore the parameter space sufficiently to produce an accurate reproduction of the posterior, often referred to as "convergence," which we assess via standard

convergence criteria. A fully converged Markov chain will have identified all available image modes that can be captured by the specified image representation and assessed their relative likelihoods.

4.4. Dynamic Imaging

As discussed in Section 3.2.2, the quickly evolving structure of Sgr A* poses significant challenges in reconstructing an image. Imaging techniques traditionally rely on Earth-rotation aperture synthesis, which is based on the fundamental assumption that the target being imaged remains stationary during the whole duration of the observation. This is no longer valid when the target source is expected to exhibit significant structural changes in timescales smaller than the observing run; thus, for static imaging we must incorporate an inflated "variability noise budget" to capture the "typical" source structure (refer to Section 3.2.2). If we instead wish to capture the evolving structure of Sgr A*, we can attempt to recover a full movie from the data, rather than just a single image.

Extensions of the CLEAN approach have been proposed to address time-variable sources (Stewart et al. 2011; Rau 2012; Farah et al. 2022). In Miller-Jones et al. (2019) evolution of the microquasar V404 Cygni was reconstructed using model fitting in `DIFMAP`. Arras et al. (2019) developed a variational inference approach for dynamic imaging that was used to simultaneously reconstruct images of M87* over four nights from EHT 2017 data. In this work we focus on methods that explicitly incorporate temporal regularization to allow for recovery of evolving sources with complex spatial structure in the presence of especially sparse (u, v)-coverage.

4.4.1. RML Dynamic Imaging

Extending the RML approach from static to dynamic imaging is simple conceptually. Rather than solving for a single image, $\hat{I}$, our new goal is to solve for a series of K images $\{\hat{I}_k\}$. Each of these images corresponds with small segments of data, which have been divided to have a time duration similar to the expected time variability of the target (typically tens of minutes for Sgr A*). Since the (u, v)-coverage of each data segment is severely limited, we must include an additional term that regularizes the images $\{I_k\}$ in time rather than just space. A general prescription in terms of the temporal regularizer S_Q can be written mathematically as

$$J(\{I_k\}) = \sum_{\text{data terms}} \alpha_D \chi^2_D(\{I_k\}, V) - \sum_{\text{spatial reg.}} \beta_R S_R(\{I_k\}) - \sum_{\text{temporal reg.}} \beta_Q S_Q(\{I_k\}). \quad (4)$$

The additional temporal regularization terms, S_Q, encourage smooth evolution of the target over the full observation. Descriptions of temporal regularizers and their application to EHT data are described in Johnson et al. (2017).

In Appendix G the RML dynamic imaging method is used to explore the structure of Sgr A* over the course of a night independent of the variability noise model introduced in Section 3.2.2.

4.4.2. `StarWarps` Dynamic Imaging

`StarWarps` is a dynamic imaging method based on a probabilistic graphical model (Bouman et al. 2018). Similar to RML dynamic imaging, `StarWarps` makes use of temporal regularization to solve for the frames of a movie $\{\hat{I}_k\}$ over an observation rather than a static image. In contrast to RML, `StarWarps` independently solves for the marginal posterior of each frame conditioned on all measurements in time; the reconstructed movie $\{\hat{I}_k\}$ is the mean of each marginal distribution. The advantage of `StarWarps` with respect to RML is that, when using a linearized measurement model, `StarWarps` can solve for the frames of a video $\{\hat{I}_k\}$ with exact inference—resulting in a better-behaved optimization problem that is less likely to get stuck in local minima when compared to RML dynamic imaging.

`StarWarps` defines a dynamic imaging model for observed data using the following potential functions:

$$\psi_{V_k|I_k} = \mathcal{N}_{V_k}(f_k(I_k),\, \sigma_k^2) \quad (5)$$

$$\psi_{I_0} = \mathcal{N}_{I_0}(\mu,\, \Lambda) \quad (6)$$

$$\psi_{I_k|I_{k-1}} = \mathcal{N}_{I_k}(I_{k-1},\, \beta_Q^{-1}\mathbb{I}), \quad (7)$$

for a normal distribution $\mathcal{N}(m, C)$ with mean m and matrix covariance or scalar standard deviation C. Each set of observed data V_k taken at time k is related to the underlying image, I_k, through the measurement model, $f_k(I_k)$ (e.g., visibility model, closure phase model). Spatial regularization is imposed through the second potential; I_k is encouraged to be a sample from a multivariate Gaussian distribution with mean μ and covariance Λ. In this work, we define Λ to encourage spatial smoothness with a spectral distribution profile $(u^2 + v^2)^{-a/2}$ controlled by hyperparameter a, as described in detail in Bouman et al. (2018). The third potential describes how images evolve over time; as β_Q increases, the temporal regularization increases, and vice versa. Although more complex evolution models are described in Bouman et al. (2018), in this paper we restrict ourselves to a simplified evolution model that encourages only small changes between adjacent frames I_{k-1} and I_k.

The joint probability distribution of this dynamic model can be written as

$$p(\{I_k\}, \{V_k\}) = p(I_1) \prod_{k=1}^{K} p(V_k|I_k) \prod_{k=2}^{K} p(I_k|I_{k-1}), \quad (8)$$

where $p(I_1) = \psi_{I_1}$, $p(V_k|I_k) = \psi_{V_k|I_k}$, and $p(V_k|I_k) \propto \psi_{I_k}\psi_{I_k|I_{k-1}}$. In the case of a linear measurement model, $f(I)$ (e.g., complex visibility model), the expected value of every I_k conditioned on all data $V = \{V_k\}$ can be solved in closed form efficiently using the elimination algorithm. However, in the case of complex gain errors the measurement model is no longer linear. By linearizing the model, we can solve in closed form for a linearized solution, $\{\hat{I}_k\}$. We then iterate between linearizing the measurement model around our current solution and solving the linearized solution in closed form until convergence.

The `StarWarps` method is used in Section 9 alongside snapshot geometric modeling methods to help analyze the short-timescale variations of Sgr A* over a ∼100-minute region of time on April 6 and 7.

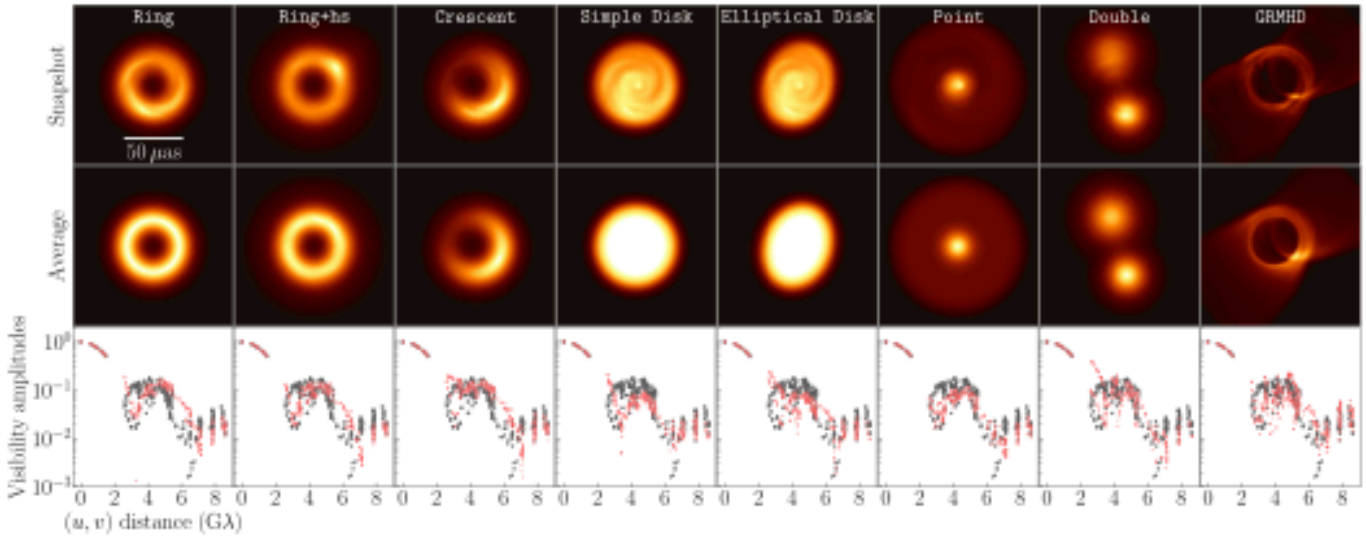


Figure 7. Eight synthetic models and corresponding visibility amplitudes. From left to right, we show the seven geometric models and the single GRMHD model. Top panels: a single frame of the eight synthetic movies highlighting the effect of temporal variability. Middle panels: time-averaged images illustrating the static component of the source structure. Bottom panels: a comparison of the simulated visibility amplitudes (red) and real Sgr A* measurements (black) as a function of projected baseline length.

5. Synthetic Data

While imaging is a powerful tool to identify the source morphology without a specific source model, reconstructed images obtained with the techniques described in Section 4 are sensitive to hyperparameter and optimization choices (in this paper, often referred to simply as parameter choices). For instance, in RML imaging methods, a common design choice is the type of regularizers and how much weight to assign the regularization terms relative to data-fitting terms. In CLEAN, common design choices include the location of CLEAN windows and the initial model used for self-calibration. Reconstructed images can be sensitive to these choices, especially when the data constraints are severely limited, as is the case in the sparse EHT measurements of Sgr A*.

In the second half of 2019, images of Sgr A* were initially reconstructed by five teams that worked independently of each other to identify the morphology of Sgr A* through imaging. As summarized in Appendix C, the five independent teams identified a ∼50 μas feature, but with a significant uncertainty in the detailed morphology. While many of the images contained a ring structure, some of the teams obtained nonring images that also reasonably fit the data. Furthermore, the flux distribution around the recovered rings showed large variation across different reconstructions. These initial images motivated a series of tests presented in this paper to systematically study the possible underlying source structure of Sgr A*.

To systematically explore and evaluate the imaging algorithms' design choices and their effects on the resulting image reconstructions, we generated a series of synthetic data sets. The synthetic data were carefully constructed to match properties of Sgr A* EHT measurements. The use of synthetic data enables quantitative evaluation of image reconstruction by comparison to the known ground truth. This in turn enables evaluation of the design choices and imaging algorithms' performance (Section 4). As summarized in Figure 7, two sets of time-variable synthetic data were generated for slightly different purposes. The first set are the geometric models (Section 5.1), which were used to both assess the capability of identifying and distinguishing different morphologies and select optimal design choices and parameters (for RML and CLEAN) that perform well across the entire data set. The second data set is the GRMHD model, which was used to evaluate imaging performance on physically motivated models of Sgr A* (see Section 6).

Data sets were generated using `eht-imaging`'s simulation library with a (u, v)-coverage identical to Sgr A* measurements. Prior to the synthetic observations, all movies were scattered based on the best-fit model of Johnson et al. (2018) (see Section 3.1 for details). The observed visibilities were further corrupted by thermal noise, amplitude gains, and polarization leakage, consistent with Sgr A* data (Paper II). Atmospheric phase fluctuations were simulated by randomizing the visibility phase gains on a scan-by-scan basis.

5.1. Geometric Models

To assess the capability of identifying source morphology, seven geometric models were used to generate synthetic data (Figure 7). As described in Section 5.1.1, the time-averaged morphology of these models was motivated by the first imaging results (Appendix C). Furthermore, the geometric model parameters were adjusted and selected to be qualitatively consistent with Sgr A* measurements. To assess the effects of temporal variability on the reconstructed images, a dynamic component is added to the time-averaged models. The static geometric models are modulated by an evolution-generated statistical model with parameters optimized to match metrics seen in Sgr A* data.

5.1.1. Time-averaged Morphology

We use the following three ring models motivated by the morphology identified in many "first images" presented in Appendix C: symmetric and asymmetric ring models (henceforth `Ring` and `Crescent`, respectively), and a symmetric ring model with a bright hot spot that rotates in the counterclockwise direction with a period of 30 minutes (henceforth `Ring+hs`). The first two models are designed to

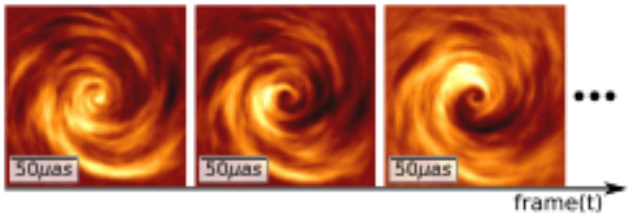


Figure 8. A sequence of three frames of statistical evolution $\exp\{\rho(t, x, y)\}$ that was sampled using `inoisy`. The random field was generated with correlations that mimic a disk rotating clockwise. Here the image sequence corresponds to ~10 minutes of observation time.

test whether our imaging methods can identify a symmetric versus asymmetric ring, while the latter hot spot model tests the effects of a fast-moving localized emission on the reconstructed images. Besides the ring models, we use four nonring images. To assess the robustness of the central depression seen in ring reconstructions, we adopt a uniform circular and an elliptical disk model (henceforth `Simple Disk` and `Elliptical Disk`, respectively). Finally, motivated by the nonring images recovered in Appendix C, we adopt a point-source and double-point-source model (henceforth `Point` and `Double`, respectively).

The parameters (e.g., diameter, width) of each geometric model are selected to be broadly consistent with representative properties of Sgr A*'s deblurred visibility amplitudes. We use the following four criteria: (1) the first null traced by Chile–LMT baselines is located at the baseline length of 3.25–3.65 Gλ and position angle (PA) of ~50°, with an amplitude of ~0.1 Jy; (2) the peak of the visibility amplitudes between the first two nulls has ~0.3 Jy; (3) the second null traced by Chile–Hawai'i and/or Chile–PV baselines is located at the (u, v)-distance of 6 Gλ, with an amplitude less than 0.1 Jy; and (4) for asymmetric models, visibility amplitudes on Chile–LMT baselines are ~1.5 times larger than on SMT–Hawai'i baselines. Figure 7 shows the comparison of visibility amplitudes between Sgr A* data on April 7 and corresponding synthetic data (after adding temporal variability—see Section 5.1.2), demonstrating qualitative agreement between the synthetic data and Sgr A* visibility amplitudes.

5.1.2. Characterization of the Time Variability

To mimic the temporal variability of Sgr A*, the geometric models, denoted by $I_{\rm geo}(x, y)$, are modulated by a temporal evolution sampled from a statistical model: `inoisy` (Lee & Gammie 2021). This model enables sampling random spatio-temporal fields, $\rho(t, x, y)$, according to specified local correlations. Lee & Gammie (2021) and Levis et al. (2021) showed that `inoisy` is able to generate random fields that capture the statistical properties of accretion disks (see Figure 8). Using this model, we modulate the static geometric models according to

$$I(t, x, y) = I_{\rm geo}(x, y)\exp\{\rho(t, x, y)\}. \quad (9)$$

Figure 7 shows both the time average and a single snapshot of each model highlighting the effect of temporal evolution.

The `inoisy` model parameters were selected to generate a similar degree of time variability to that of Sgr A* measurements. We impose the following conditions to match metrics of temporal variations: (1) the mean of each movie's total flux is 2.3 Jy, consistent with the ALMA light curve on April 7 (Section 2; Wielgus et al. 2022); (2) the standard deviation of the total flux is in the range of 0.09–0.28 Jy, as seen in the ALMA and SMA light curves and intrasite baselines; (3) the $\mathcal{Q}$-metric (Roelofs et al. 2017) of the intrinsic closure phase variability is comparable to Sgr A* data on all triangles.

Figure 9 shows temporal variations in the total flux density and closure phases with comparison to Sgr A* data. The synthetic data variability is able to capture the real data light-curve variablity. Moreover, the power spectrum density distributions of the light curves from synthetic models are broadly consistent with Sgr A* data. For closure phases, `inoisy` produces data with visible time variability seen in high-S/N triangles, such as the ALMA–SMT–LMT triangle. These synthetic movies also roughly match the Sgr A* variability amplitudes, averaged over all triangles, as evaluated using the $\mathcal{Q}$-metric. We note that while these synthetic data are in good agreement with the above aspects of the EHT data, their variability amplitudes in Fourier domain are slightly less than Sgr A* data (see Table 2 in Section 3.2).

5.2. GRMHD Model

In addition to the geometric models, we also generated synthetic data from GRMHD simulations to evaluate the performance of our imaging procedures on more complicated physically motivated models of Sgr A*. These GRMHD models are selected from the simulation library presented in Paper V and are in general agreement with Sgr A* data (Paper V, Section 3.1.2). Section 6.4.2 shows the result of applying our imaging procedure to a weakly magnetized "standard and normal evolution" (SANE) model, with dimensionless spin $a_* = -0.94$, electron temperature ratio $R_{\rm high} = 160$, and viewing inclination $i = 50°$. Although it failed in other constraints (see Paper V, Appendix A), this model satisfies the same criteria used for selection of the geometric models as seen by the resulting visibility amplitudes (Section 5.1.1) and temporal variability (see Section 5.1.2), as shown in Figure 9. Figure 7 shows a single snapshot frame of the GRMHD movie, along with a time-averaged structure. The GRMHD movie frames contain a sharp photon ring with a faint emission broadly extended over ~100 μas. Some of the frames contain notable spiral arm features surrounding the photon ring and extending beyond the compact structure (refer to the SANE frames shown in Figure 25). This spiral arm feature is smoothed out by averaging over the observational time.

In addition, Appendix H shows the result of applying the same imaging procedure on a strongly magnetized "magnetically arrested disk" (MAD) model with positive spin, which passes more observational constraints and is in the "best-bet region" considered by Paper V. By using these two GRMHD models, generated with different physical parameters, we demonstrate that our imaging procedure and the resulting performance are robust against the details of GRMHD models.

6. Imaging Surveys with Synthetic Data

We conducted surveys over a wide range of imaging assumptions with four scripted imaging pipelines using RML, CLEAN, and a Bayesian posterior sampling method. The surveys were performed on the synthetic data sets presented in Section 5, as well as on the real Sgr A* data. Reconstructing synthetic data with exactly the same procedure used on Sgr A*

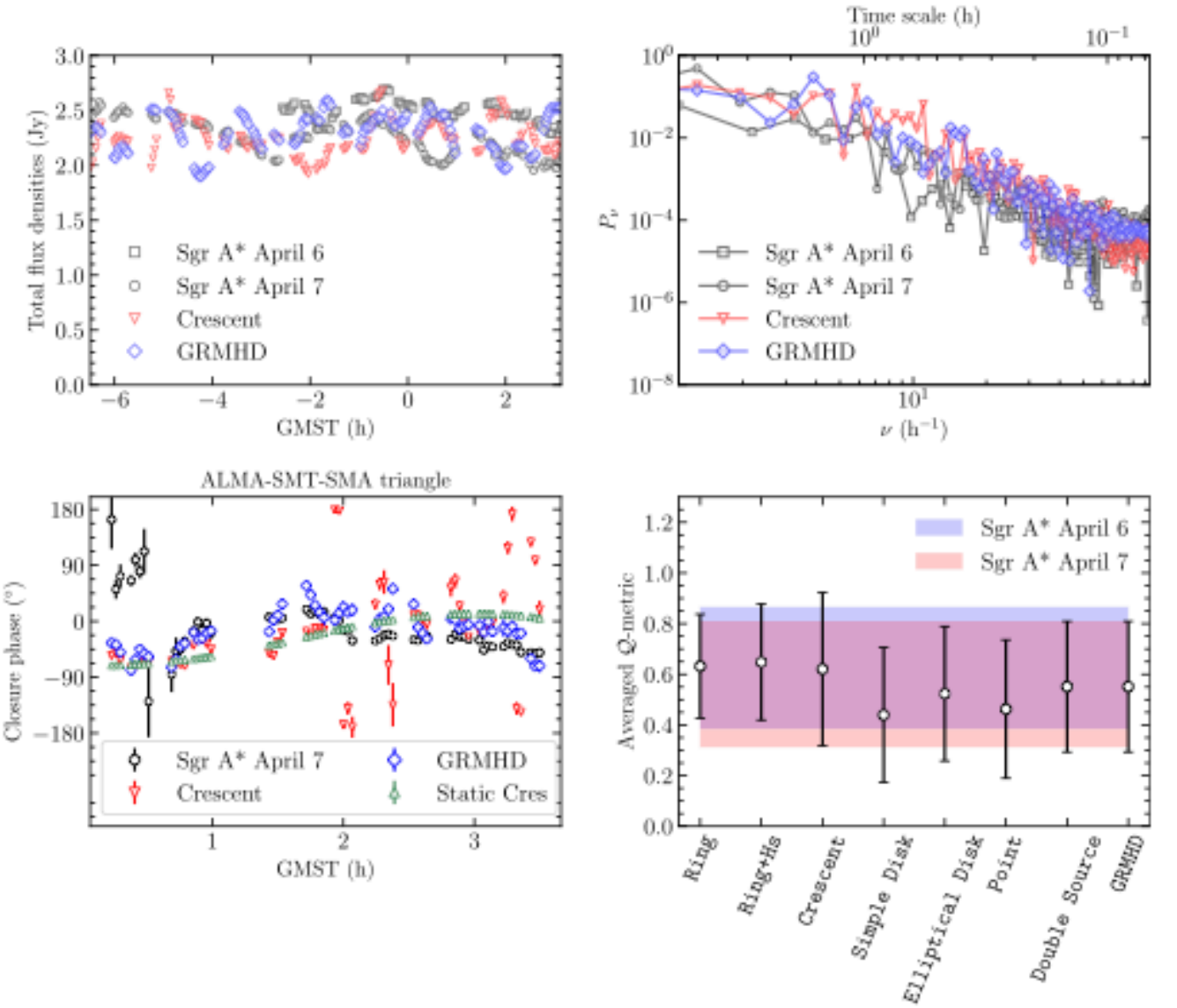


Figure 9. Temporal variability in Sgr A* and synthetic data. Top left: light curves of ALMA Sgr A* data (gray points), the crescent model (red points), and the GRMHD model (blue points). Top right: power spectrum density distributions of the light curves in the same color conventions. Bottom left: closure phases on ALMA–SMT–SMA triangle from the April 7 Sgr A* data (gray), the crescent model (red), its time-averaged static image (green), and movie of the GRMHD model (blue). Bottom right: $\mathcal{Q}$-metrics of closure phases from Sgr A* and synthetic data. We show the mean and standard deviation of $\mathcal{Q}$-metrics over all independent triangles for each data set, overlaid by two shaded areas indicating the corresponding ranges for Sgr A* data on April 6 and 7.

allows us to assess our ability to identify the true underlying time-averaged morphology. Both synthetic and real Sgr A* data sets were preprocessed with a common pipeline described in Section 6.1. We describe the RML and CLEAN imaging parameter surveys in Section 6.2 and imaging with a Bayesian posterior sampling method in Section 6.3. Images of synthetic data from imaging surveys are described in Section 6.4. We present images of the actual Sgr A* data in Section 7.

6.1. Common Pre-imaging Processing

To reconstruct a time-averaged image of Sgr A*, each pipeline used the original calibrated data sets described in Section 2 and/or data sets further normalized by the time-dependent total flux density of Sgr A*. Within each imaging pipeline, data sets were first time-averaged to enhance the S/N of visibilities; each pipeline adopted a single integration time or explored multiple choices of the integration time (see Tables 3–5). After time averaging, fractional errors of 0%, 2%, or 5% were added to the visibility error budget in quadrature to account for the nonclosing systematic errors (refer to Paper II).

As described in Sections 3.1 and 3.2, we employ additional strategies to mitigate extrinsic scattering and intraday variations. To assess these proposed strategies, and our ability to account for these two effects in the imaging process, we incorporate parameterized error budgets and systematically explore the various assumptions on these two effects in the RML and CLEAN surveys discussed in Section 6.2. In total, we potentially include up to three additional noise budgets that account for (1) systematic error, (2) interstellar scattering, and (3) time variability.

The second budget accounts for the substructure arising from refractive scattering. We added the anticipated refractive

Table 3
Parameters in the DIFMAP Pipeline Top Set

Apr. 7 (8,400 Param. Combinations; 1,626 in Top Set)							
Systematic error	0 25.6%		0.02 36.8%		0.05 37.5%		
Ref type	No 14.9%	Const 20.7%	2×Const 21.3%	J18 22.1%	2×J18 20.9%		
$a_{\rm psd}$	No 5.1%	0.015 28.5%	0.02 32.1%	0.025 34.3%			
$b_{\rm psd}$	No 5.1%	1 20.2%	3 35.5%	5 39.2%			
$\lvert u\rvert_0$	No 5.1%	2 94.9%					
Time average (s)	10 45.0%	60 55.0%					
ALMA weight	0.1 41.1%	0.5 58.9%					
UV weight	0 54.7%	2 45.3%					
Mask diameter (μas)	80 0.2%	85 2.5%	90 22.3%	95 25.0%	100 21.3%	105 20.0%	110 8.7%

Note. In each row, the upper line shows the surveyed parameter value corresponding to the parameter of left column, while the lower line shows the number fraction of each value in the Top Set. The total number of surveyed parameter combinations and the Top Set are shown in the first row.

Table 4
Parameters in the eht-imaging Pipeline Top Set

Apr. 7 (112,320 Param. Combinations; 5,594 in Top Set)					
Systematic error	0 21.4%	0.02 36.7%	0.05 41.8%		
Ref type	No 27.0%	Const 23.9%	2×Const 20.6%	J18 16.4%	2×J18 12.1%
$a_{\rm psd}$	No 11.4%	0.015 40.6%	0.02 26.6%	0.025 21.4%	
$b_{\rm psd}$	No 11.4%	1 24.8%	2 20.4%	3 21.5%	5 21.8%
$\lvert u\rvert_0$	No 11.4%	2 88.6%			
TV	0 13.2%	0.01 16.0%	0.1 36.1%	1 34.7%	
TSV	0 29.5%	0.01 32.3%	0.1 26.6%	1 11.5%	
Prior size (μas)	70 33.9%	80 34.7%	90 31.4%		
MEM	0 7.8%	0.01 18.4%	0.1 54.3%	1 19.6%	
Amplitude weight	0 0.9%	0.1 23.3%	1 75.8%		

Note. Same as Table 3.

noise level for the observed (i.e., interstellar-scattered) visibilities using two base models described in Section 3.1. The first model we explore is the `Const` model, which adds a constant noise budget across all visibilities prior to deblurring; we examined two noise levels: 0.4% and 0.9% of the total flux density at each time segment (i.e., `Const` and `2×Const`). The second model explored is the `J18model1 (J18)` model, which is no longer constant in the (u, v)-space; we adopt two scaling factors of 1.0 and 2.0 for this noise floor (i.e., `J18model1 (J18)` and `2×J18model1 (2×J18)`. The two noise levels adopted in each model reasonably cover differences in the noise levels caused by different potential intrinsic structures (see Section 3.1). After including one of the above budgets for the refractive noise, visibilities were divided by the diffractive scattering kernel based on the J18 model to mitigate diffractive scattering. In addition to the above four sets of the scattering mitigation schemes, we attempted imaging without any form of scattering mitigation to probe the interstellar-scattered source structure of Sgr A* (referred to as *on-sky* images).

The third noise budget explored accounts for the structural deviations from the time-averaged morphology due to the intraday variations. We further inflated the visibility error budget using the variability noise model described by Equation (2) in Section 3.2. This budget was added in quadrature to the visibility noise budget, after being normalized by the time-dependent total flux density. We systematically explored various sets of parameters in Equation (2), including the variability rms level at 4 Gλ (a), the break location (u_0), and the variability power-law spectra index at long baselines (b). Similar to scattering, we also attempted reconstructions without this error budget (i.e., assuming no intraday variation in data).

6.2. RML and CLEAN Imaging Parameter Surveys

In a manner similar to previous EHT imaging of M87* (M87* Papers IV and VII), we explore how recovered images are influenced by different imaging and optimization choices. In particular, we objectively evaluate each set of imaging parameters in scripted RML and CLEAN imaging pipelines using synthetic data with known ground-truth images. Each parameter survey leads to a *Top Set* of parameters: parameter combinations that each produce acceptable images on our entire suite of synthetic data. The distribution of Sgr A* images recovered with the Top Set parameter combinations reflects our uncertainty due to modeling and optimization choices made in imaging; thus, it is different from a Bayesian posterior and instead attempts to characterize what is sometimes referred to as epistemic uncertainty.

6.2.1. Imaging Pipelines

Similar to previous EHT work (M87* Papers IV and VII), we designed three scripted imaging pipelines utilizing the `DIFMAP`, `eht-imaging`, and `SMILI` software packages. After completing the common pre-imaging processing of data (Section 6.1), each pipeline reconstructs images using a broad parameter space (weights for the regularization functions, mask sizes, station gain constraints, variability noise budget parameters, etc.). We describe each pipeline in detail in Appendices D.1, D.2, and D.3.

Each pipeline explored on the order of 10^3–10^5 parameter combinations, as summarized in Tables 3, 4, and 5 for

Table 5
Parameters in the SMILI Pipeline Top Set

Apr. 7 (54,000 Param. Combinations; 2,763 in Top Set)					
Systematic error	0 33.9%		0.02 33.3%	0.05 32.8%	
Ref type	No 15.7%	Const 22.1%	2×Const 18.5%	J18 21.4%	2×J18 22.3%
$a_{\rm reg}$	No 7.5%		0.015 37.2%	0.02 26.7%	0.025 28.6%
$b_{\rm reg}$	No 7.5%	1 16.8%	2 27.7%	3 31.9%	5 16.1%
$\|w\|_0$	No 7.5%		1 40.7%	2 51.8%	
TV	10^2 8.6%		10^3 46.7%	10^4 44.4%	10^5 0.3%
TSV	10^2 38.7%		10^3 53.3%	10^4 8.0%	10^5 0.0%
Prior size (μas)	140 33.0%		160 33.6%	180 33.3%	
ℓ_1	0.1 44.8%		1 54.8%	10 0.3%	

Note. Same as Table 3.

DIFMAP, eht-imaging, and SMILI, respectively. Each pipeline has some unique choices that are fixed (e.g., the pixel size, or the convergence criterion) and surveyed (e.g., the regularizer weights), while some parameters are commonly explored (e.g., parameters for the scattering and intraday variations in Section 6.1).

While all imaging pipelines adopt the common preprocessing of data described in Section 6.1, there are some differences in data processing. For instance, the noise budgets for refractive scattering and intraday variability are updated during self-calibration rounds in SMILI. RML imaging pipelines (eht-imaging and SMILI) adopt the same prior and initial images across all synthetic and real data sets. The DIFMAP pipeline uniformly explores a library of initial models for a first-phase self-calibration, selecting the one that provides the best fit to the closure phases after a first run of cleaning (see Appendix D.1). All three pipelines use combined low- and high-band data for imaging without any data flagging (including the intrasite baselines).

6.2.2. Top Sets of Imaging Parameters via Surveys on Synthetic Data

Large imaging surveys on synthetic data facilitate the evaluation of different potential parameter combinations. Following M87* Paper IV, the principal output from each parameter survey is a Top Set: a set of parameter combinations that produce acceptable images on the suite of synthetic data presented in Section 5.

The fidelity of synthetic image reproduction is measured using the normalized cross-correlation between the reconstructed images and ground-truth images. We define the normalized cross-correlation of two images X and Y made of N pixels as

$$\rho_{\rm NX} = \frac{1}{N}\sum_{i}^{N} \frac{(X_i - \langle X\rangle)(Y_i - \langle Y\rangle)}{\sigma_X \sigma_Y}. \quad (10)$$

Here X_i and Y_i denote the image intensity at the ith pixel, $\langle X\rangle$ and $\langle Y\rangle$ denote the mean pixel values of the images, and σ_X and σ_Y are the standard deviations of pixel values. The position offset between the frames is corrected by shifting one frame relative to the other along R.A. and decl. and identifying the shift coordinate that corresponds with the largest $\rho_{\rm NX}$.

In order to recover a Top Set, a threshold for $\rho_{\rm NX}$ is defined for each synthetic data set. Imaging parameter combinations that recover images that score above that threshold for all synthetic data sets are selected as Top Sets. The threshold values of $\rho_{\rm NX}$ are determined in a manner similar to M87* Paper IV. For each "ground-truth" image (obtained by time-averaging the synthetic movie), we evaluate $\rho_{\rm NX}$ between the ground-truth image and the same image blurred with a Gaussian beam of FWHM equivalent to the maximum fringe spacing of the Sgr A* observations, 24 μas. This value of $\rho_{\rm NX}$ quantifies the potential loss of the image fidelity due to the limited angular resolution. Figure 10 shows examples of the $\rho_{\rm NX}$ curves between unblurred and blurred ground-truth images as a function of the blurring size; the critical value corresponds to those at $\alpha = 24$ μas. Note that this value depends on the true source structure. Unlike in M87* Paper IV, we find that a relaxation of the $\rho_{\rm NX}$ threshold is required to account for the fact that we reconstruct static images from time-variable data sets. Hence, the critical $\rho_{\rm NX}$ values for all training data sets are multiplied by the relaxation factor of 0.95. In other words, the threshold for each synthetic data set is set to $0.95 \times \rho_{\rm NX}$ for $\rho_{\rm NX}$ evaluated at $\alpha = 24$ μas. This relaxed threshold allows for a large enough number of Top Set parameters to be identified for all epochs and imaging pipelines. We ensure that the relaxed threshold still reconstructs all representative ground-truth morphologies; in Appendix E, the worst $\rho_{\rm NX}$ images are shown for Top Set reconstructions of each model with each imaging pipeline to demonstrate that the representative features are recovered even in the worst-fidelity Top Set images.

In Tables 3, 4, and 5, we summarize the parameters and surveyed values in each pipeline's survey. These tables indicate the fraction of images corresponding to that parameter in each pipeline's Top Set for April 7 data. The results on April 6 data are summarized in Appendix E. The tables also provide the total number of surveyed parameter combinations, as well as the number of combinations selected for each Top Set. As seen in each table, there are more than 1000 parameter combinations in each pipeline's Top Set for April 7, which we find is sizable enough for downstream analysis. In contrast, we find that the April 6 Top Set sizes are much smaller, likely due to the poorer (u, v)-coverage.

6.3. Image Posteriors with THEMIS

The Bayesian imaging method employed in THEMIS differs from those described above in a number of respects. Most significantly, apart from sampler tuning—which affects only the efficiency with which the posterior is explored—the splined raster model has only two free parameters. Image resolution, raster orientation, the brightness at each control point, and noise model parameters are determined self-consistently by the

Table 6
Bayesian Imaging Priors

Parameter	Units	Prior[a]
Control points	Jy μas^{-2}	$\mathcal{L}(10^{-5}, 0.1)$
FOV$_x$	μas	$\mathcal{U}(50, 500)$
FOV$_y$	μas	$\mathcal{U}(50, 500)$
Raster rotation	rad	$\mathcal{U}(-0.25\pi, 0.25\pi)$
Shift in x	μas	$\mathcal{U}(-100, 100)$
Shift in y	μas	$\mathcal{U}(-100, 100)$
$\ln(\sigma_{\rm rat})$	···	$\mathcal{N}(-5.5, 1.0)$
$\ln(f)$	···	$\mathcal{N}(-4.6, 1.0)$
a	···	$\mathcal{L}(a_{25\%}, a_{75\%})$[b]
a_0	···	$\mathcal{L}(a_{0.25\%}, a_{0.75\%})$[b]
b	···	$\mathcal{L}(b_{25\%}, b_{75\%})$[b]
c	···	$\mathcal{L}(1.5, 2.5)$

[a] Linear priors from a to b are represented by $\mathcal{U}(a, b)$, logarithmic priors from a to b are represented by $\mathcal{L}(a, b)$, and normal priors with mean μ and standard deviation σ are represented by $\mathcal{N}(\mu, \sigma)$.
[b] $x_{25\%}$ and $x_{75\%}$ are the bottom and top quartile values of the quantity x given in Table 2.

fitting process (see Appendix A). This is achieved by replacing the hyperparameters associated with field of view (FOV), scattering threshold in the `Const` prescription, systematic error budget, and those that define the variability mitigation noise with fit parameters, precluding the need to survey over them. Priors for each quantity are listed in Table 6 and make use of the pre-imaging constraints described in Section 3 and listed explicitly in Table 2.

Importantly, eliminating the hyperparameters associated with the additional contributions to the visibility uncertainties eliminates the noise-related data preparation steps described in Section 6.1; we do not add any additional uncertainty prior to the THEMIS analysis. However, to reduce the data volume (and thus computational expense of the posterior sampling), we incoherently average the flux-normalized data over scans. To prevent significant coherence losses, prior to averaging we calibrate the phase gains of the JCMT using the intrasite baseline, JCMT–SMA, and assume that the source is unresolved at the corresponding spatial scales probed by that baseline.

The remaining two unspecified hyperparameters are the raster dimensions, N_x and N_y. Initial guesses for these are provided by the diffraction limit; for a typical source size of 80 μas, $N_x = N_y = 5$ is sufficient to marginally superresolve the source. This may then be refined via a modest survey over potential values, with the final values selected by maximizing the Bayesian evidence, computed in THEMIS via thermodynamic integration (Lartillot & Philippe 2006). In practice, due to the expense of such a survey, we restrict ourselves to $N_x = N_y = 5$ for the validation with synthetic data sets with a sole exception. We performed a raster dimension survey for the GRMHD data set presented in Section 6.4.2, finding that $N_x = N_y = 6$ is preferred.

For application to Sgr A*, we perform raster dimension surveys independently for April 6 and 7 as described in Appendix D.4, finding preferred dimensions of $N_x = N_y = 6$ and $N_x = N_y = 7$, respectively.

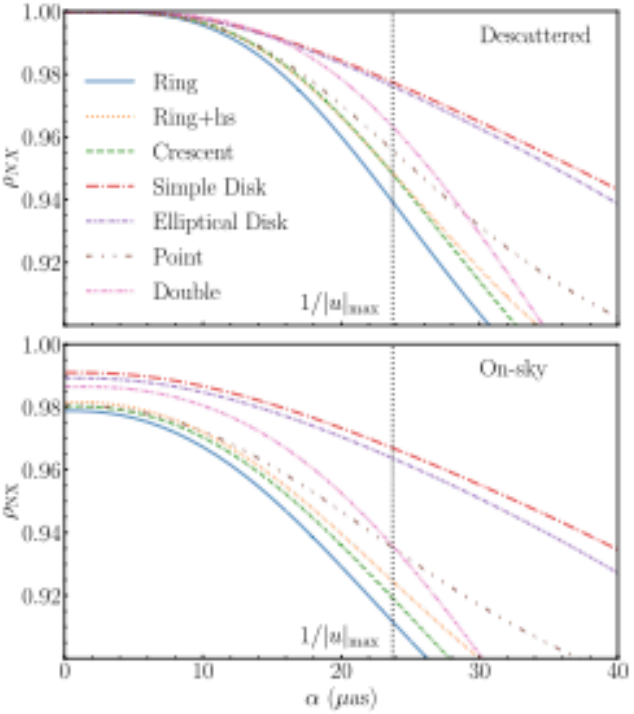


Figure 10. Normalized cross-correlation, $\rho_{\rm NX}$, between time-averaged ground-truth images and their corresponding blurred images, as a function of the blurring kernel size. These curves are shown for both the intrinsic images (top panels) and the scattered (i.e., on-sky) images (bottom panels). The dashed black line indicates the angular resolution equivalent to the maximum fringe spacing of Sgr A* observations, 24 μas; the corresponding value of $\rho_{\rm NX}$ at 24 μas is used to define a threshold that selects which `DIFMAP`, `eht-imaging`, and `SMILI` imaging parameter combination is applied to Sgr A* data.

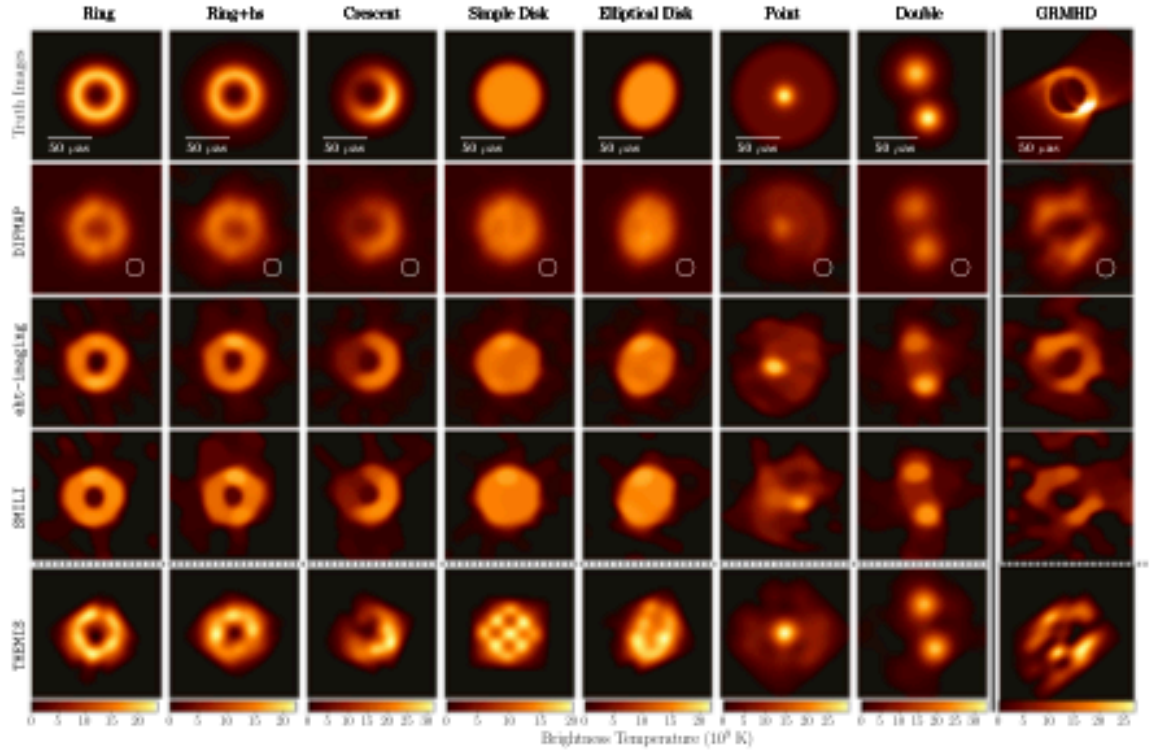
6.4. Synthetic Data Images

We first present Top Set and posterior images recovered from geometric model data sets in Section 6.4.1, followed by images recovered from the GRMHD data set in Section 6.4.2.

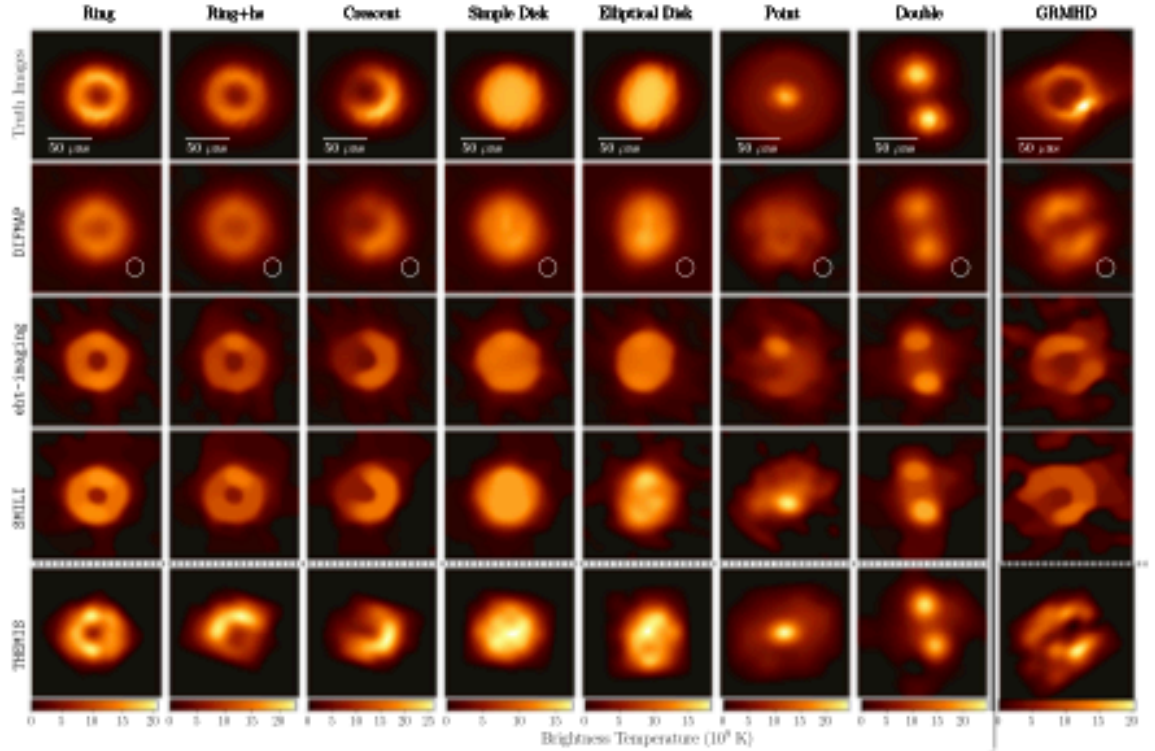
6.4.1. Geometric Model Images

In Figure 11, we show the time-averaged ground-truth images and corresponding image reconstructions for each of the synthetic data sets with April 7 (u, v)-coverage. Images from all pipelines are obtained with and without scattering mitigation (henceforth *descattered* and *on-sky* reconstructions). These descattered and on-sky reconstructions are compared to the time-averaged ground-truth images of the intrinsic and scattered structure, respectively.

`DIFMAP`, `eht-imaging`, and `SMILI` images for each geometric model in Figure 11 are obtained using cross-validation: the parameter combination that provides the best mean $\rho_{\rm NX}$ across other geometric models (i.e., except for the model being tested) is selected. The cross-validation images in Figure 11 (which are contained in each pipeline's Top Set) successfully recover the representative morphology of each geometric model, demonstrating the capability of a single imaging parameter combination to identify various source structures. The manifestation of structure significantly different from the ground-truth morphology is only seen in a small fraction of the cross-validated Top Set parameters. For instance, using the ring classification method described in



(a) Descattered Reconstructions



(b) On-sky Reconstructions

Figure 11. Reconstructed images of synthetic data sets on April 7 (a) with and (b) without scattering mitigation for seven geometric models and the GRMHD model. Reconstructions of each geometric model by `DIFMAP`, `eht-imaging`, and `SMILI` pipelines are made using a parameter combination identified via *cross-validation*: the imaging parameters that result in the best average $\rho_{\rm NX}$ across all other geometric models. These cross-validation results demonstrate the ability of the selected parameters to correctly reproduce novel source morphologies. GRMHD reconstructions for `DIFMAP`, `eht-imaging`, and `SMILI` are produced from an imaging parameter combination that performs best on all geometric models. In contrast, for THEMIS reconstructions the average posterior image is shown for each model; the average posterior image appears to correctly identify the true source morphology in all synthetic data sets tested.

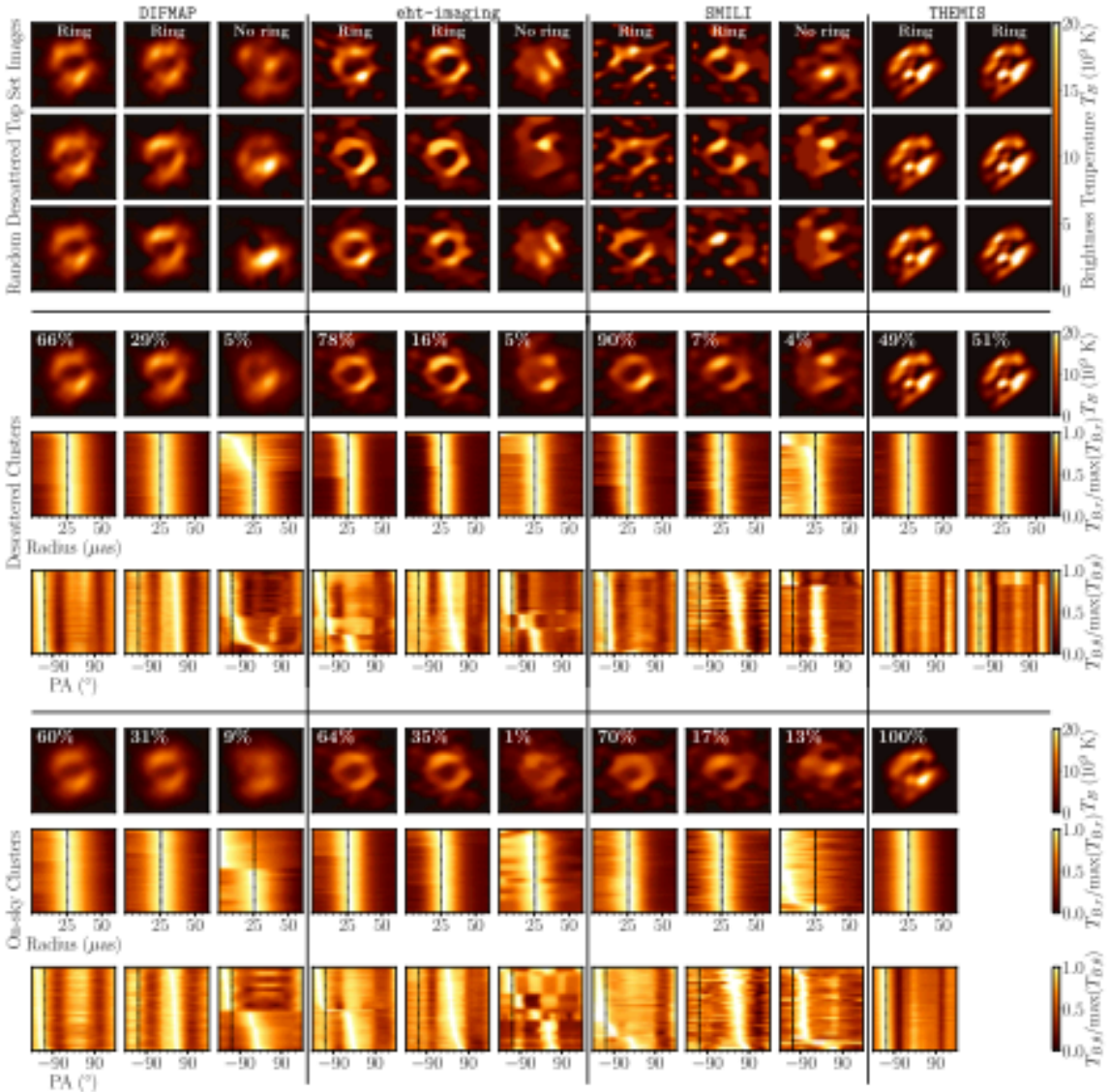


Figure 12. The distribution of images obtained from synthetic data provided using a GRMHD movie with April 7 (u, v)-coverage. From left to right (separated by vertical lines), we show the distributions of Top Set images from the `DIFMAP`, `eht-imaging`, and `SMILI` pipelines and posterior samples from the THEMIS pipeline; each vertical panel is further subdivided into clusters identifying common morphologies recovered by each pipeline. The figure is composed of three horizontal panels separated by horizontal lines. The top panel shows individual images randomly sampled from different clusters. The middle and bottom panels visualize the distributions of reconstructed descattered and on-sky images for each cluster, respectively. In each panel, from top to bottom, we show the average of each cluster, the distributions of the radial profiles, and the distributions of azimuthal intensity profiles. In the radial profiles, each horizontal slice corresponds to the azimuthally averaged intensity profile of an image, normalized by its peak value. Similarly, the azimuthal profiles show the azimuthal distribution of the radial peak intensity within a radius of 10–40 μas, also normalized by its peak value. The dotted lines in the profiles indicate the peak radius or PA of the ground-truth GRMHD movie. The vertical order of images in both profiles is independently sorted by the peak radius or PAs; therefore, the images are ordered differently in each profile distribution image. These results on synthetic data show that the Top Set and posterior samples from the imaging pipelines are able to recover ring images that resemble the true time-averaged structure of a GRMHD movie (see Figure 7). However, the imaging methods sometimes produce nonring images that still fit the data well.

Appendix F, only 2%, <1%, and <1% of the cross-validated Top Set reconstructions for the `Ring` model are identified as not having a ring feature for `DIFMAP`, `eht-imaging`, and `SMILI`, respectively; the reconstruction of a nonring morphology is also found to be limited to small fractions of 3%, 5%, and 1% for the `Crescent` model. Similarly, only a small fraction of the reconstructions from nonring models are reconstructed with a ring morphology—in particular, for the `DIFMAP`, `eht-imaging`, and `SMILI` pipelines, 7%, 11%, and 3% and <1%, <1%, and <1% of the cross-validated Top Sets reconstructed a ring feature from the `Point` and `Double` source models, respectively.

For THEMIS reconstructions, the mean posterior image is shown for each model. Figure 11 shows that the posterior images from THEMIS reconstructions identify the general morphology of each geometric model. Most of the THEMIS posterior images satisfy the criteria based on a minimum threshold of $\rho_{\rm NX}$ used in Top Set selections for `DIFMAP`, `eht-imaging`, and `SMILI` pipelines (see Section 6.2.2), with the exception of a few models discussed below. For April 7 images, all posterior images show higher $\rho_{\rm NX}$ than the threshold except for 1% and 7% of descattered reconstructions for `Double` source and `Ring+hs` models, respectively. For April 6, the images below the threshold are limited to 1% and

22% of descattered reconstructions for the Point and Double source models, respectively. However, 95% of on-sky reconstructions of the Point source model are above the $\rho_{\rm NX}$ threshold on April 6 for THEMIS. The high fractions of images beyond the $\rho_{\rm NX}$ threshold for all models on April 7 demonstrate the capability of the THEMIS pipeline to recover various representative morphologies at an acceptable fidelity.

Figure 11 also shows the resiliency of the reconstructed morphology to the scattering prescriptions. While the on-sky reconstructions without scattering mitigation tend to be slightly blurrier than those with descattering, there are not many other notable differences in their appearance. In particular, the refractive substructure, which adds spatial distortion of images on scales finer than the angular resolution of the EHT, is not well constrained in our EHT data and therefore does not strongly appear in any reconstructions. We further discuss the effects of scattering prescriptions for Sgr A* images in Section 7.5.2.

6.4.2. GRMHD Reconstructions

We show example GRMHD reconstructions on April 7 in the rightmost panels in Figure 11. This GRMHD simulation contains a ring with a diameter of ∼51 μas. GRMHD images from DIFMAP, eht-imaging, and SMILI pipelines are reconstructed with the Top Set parameter combinations that correspond with the largest average $\rho_{\rm NX}$ value across all seven geometric models. For THEMIS reconstructions, on the other hand, the expected (i.e., mean) posterior image is shown.[152]

Unlike with the simple geometric synthetic data sets, the distribution of GRMHD reconstructions shows wide variations in the image appearance.[153] Although GRMHD reconstructions in Figure 11 commonly show a ring-like morphology with a diameter of ∼50 μas, the azimuthal intensity distribution is not consistent across the Top Sets. In the top panels of Figure 12, we show images from GRMHD April 7 data from each pipeline, which are randomly selected from DIFMAP, eht-imaging, and SMILI Top Sets and posterior images from THEMIS. Most of the images have clear asymmetric ring features, but a few images show nonring structures. The diameters of ring features are broadly consistent across different reconstructions, comparable to the ground-truth image. On the other hand, the azimuthal distributions are not uniquely constrained by the Top Set images—different PAs are seen in the randomly sampled images from each Top Set.

To visualize the distribution of images with different morphology, we categorize all images into three major groups: ring images peaked at PAs within (a) the range of $-180^\circ \leqslant {\rm PA} < 0^\circ$, where the ground-truth value of -124° is located; (b) the range of $0^\circ \leqslant {\rm PA} < 180^\circ$; and (c) the remaining images comprising nonring or other ring-like images with much less consistency. The definition of a ring used in this paper to classify ring versus nonring images is described in Appendix F. Note that the particular definition of a ring will influence the quoted percentages of ring and nonring images recovered. We find that the particular definition chosen in this paper results in classification that largely aligns with human perception. However, the classification of images that are borderline between ring and nonring classification is sensitive to the exact criteria used. Therefore, ring definitions that make use of different criteria can lead to classification that still largely aligns with human perception but varies somewhat in the ring classification percentages quoted in this paper.

In the middle and bottom panels of Figure 12, we summarize the distribution of images from each pipeline with and without scattering mitigation, respectively. The GRMHD images within each pipeline's Top Set are clustered into image modes. The upper subpanel shows the mean image of each cluster, indicating a common or representative morphology recovered. The middle subpanel shows the distributions of the azimuthally averaged radial intensity profiles, where the vertical order of images is sorted by the peak radii of the profiles. The lower subpanel shows the azimuthal distribution of the radial peak intensity within a radius of 10–40 μas.

Figure 12 demonstrates that most of the Top Set or posterior images reconstruct ring images from the GRMHD data set. In particular, the radial profiles of the first two ring modes show a broad consistency of the peak radius around ∼25 μas, which implies a diameter of ∼50 μas consistent with that of ∼50 μas for the ground-truth model. The capability of identifying a ring with the consistent diameter does not appear to depend on the scattering mitigation. Similarly to geometric synthetic data (see Section 6.4.1), the scattering mitigation does not significantly affect the resulting source morphology in the reconstructed images, except that the images without scattering mitigation tend to be slightly broader. Refractive substructure does not appear in the reconstructed on-sky images, again indicating that the appearance of the refractive substructure is not strongly constrained with EHT 2017 data.

As seen in the azimuthal profiles, each pipeline provides at least one asymmetric ring mode with the peak PA roughly consistent with that of the mean ground-truth image of -124°. For this particular GRMHD example, this mode is found to be the most popular mode in most pipelines. However, Figure 12 indicates that the ring mode with the correct orientation is not always identified as the most popular mode among image samples—for instance, the correct orientation is not identified as the most popular ring mode for the eht-imaging on-sky pipeline. Therefore, caution should be taken, as the popularity of a mode in an RML or the CLEAN pipeline's Top Set does not necessarily always correspond with the true underlying structure.

We find that key takeaways from the GRMHD example are consistent with the "best-bet" GRMHD models presented in Paper V, identified based on various criteria using both EHT and non-EHT data. In Appendix H, we show example reconstructions of a "best-bet" GRMHD model. We find the identification of a ring morphology for the vast majority of the Top Set reconstructions; however, multiple ring modes and nonring images are still reconstructed. Note also that the peak PA of the ground-truth "best-bet" GRMHD model is not necessarily identified as the most popular mode reconstructed in each Top Set. These results indicate that the same trends are seen across multiple GRMHD models that are broadly consistent with EHT data.

7. Horizon-scale Images of Sgr A*

Having determined Top Set imaging parameters for RML and CLEAN and validated posterior estimation for THEMIS via tests on synthetic data (Section 6), we now show the result of

[152] Unlike for Sgr A*, for this particular GRMHD data set the THEMIS pipeline used a large-scale Gaussian to account for extended emission in the underlying source model.

[153] Note that this resembles the varied Sgr A* reconstructions in Section 7 and Appendix C.

these methods applied to Sgr A* data from the 2017 EHT observations. Unlike in the previous EHT imaging of M87* (M87* Papers IV and VII), Sgr A*'s recovered structure depends somewhat on the imaging strategy and parameter choices. Thus, this section presents our main imaging results and analyzes how the image structure is affected by different imaging choices.

We begin in Section 7.1 by giving an overview of the results, followed by a more detailed discussion of the image structures recovered in each pipeline's Top Set or posterior in Section 7.2. Average images across pipelines, calibrated data sets, and observing days are discussed in Section 7.3. In Section 7.4 we present Sgr A* imaging obtained by combining the data sets for April 6 and 7. In Section 7.5 we address the question of whether Sgr A* is a ring and explore how the recovered images are affected by the scattering and temporal variation mitigation strategy.

7.1. Overview of Recovered Sgr A* Structure

Figure 13 summarizes the common morphologies recovered from Sgr A* data by the four imaging pipelines for April 7. We find that the vast majority are rings that can be classified into three different clusters with varying azimuthal structures (Section 7.2), shown in the three bottom left panels of Figure 13, and a small fraction of nonring images that also fit the Sgr A* data well (bottom right panel of Figure 13). Since these nonring images are not as consistent in structure, they largely blur out when averaged together and primarily emphasize a double structure that is sometimes present. The representative Sgr A* image obtained by averaging reconstructions from all four clusters is shown in the large top panel of Figure 13, corresponding to that of a ring with a diameter of $\sim$50 μas (Section 8).

7.2. Clustering of Recovered Sgr A* Images

To effectively visualize the distributions of Sgr A* images, we cluster all reconstructed Top Set and posterior images using a similar criterion to that used in Section 6.4.2 and Figure 12 (see Appendix F for details). Images with a ring feature are grouped by the peak PAs in the southwest ($-180^\circ \leqslant \mathrm{PA} < -70^\circ$), northwest ($-70^\circ \leqslant \mathrm{PA} < 40^\circ$), and east ($40^\circ \leqslant \mathrm{PA} < 180^\circ$) directions. The Sgr A* image clustering results for each pipeline are summarized in Figure 14; here we focus on the observed date of April 7 with the HOPS data reduction pipeline (see Section 7.3 for April 6 and CASA-based reconstructions). The results of the DIFMAP, eht-imaging, SMILI, and THEMIS pipelines are displayed from left to right. Images are separately clustered within each pipeline.

The middle and bottom panels of Figure 14 show each cluster's average image for the descattered and on-sky images, respectively. To better visualize the properties of individual Sgr A* images, we present three randomly selected descattered images from each cluster in the columns of the top panel of Figure 14; within each cluster images appear to have largely consistent morphologies.

The percentages in each panel of Figure 14 show the fraction of Top Set or posterior images contained in that particular cluster. These percentages indicate that most of the Top Set and posterior images have ring structures. For instance, the fraction of nonring cluster images is $\leqslant$5% of the Top Set descattered images from DIFMAP, eht-imaging, and SMILI imaging pipelines. Although this is significant, we note that in the case of the Top Sets this does not constitute a likelihood, and therefore the reported fractions should not be considered as an exact measure of our degree of certainty. In addition, only ring images appear in THEMIS posterior estimation of descattered images.

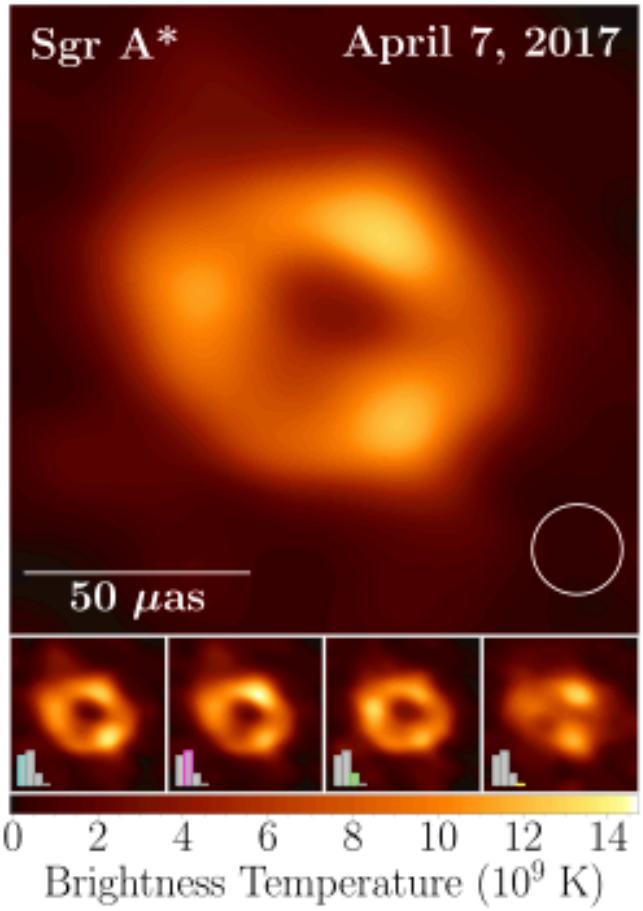


Figure 13. In the top main panel we show the representative image of Sgr A* obtained with the EHT from observations on 2017 April 7. This top image is obtained by averaging the bottom four images. On the bottom from left to right, we show the average images of three prominent ring clusters with different azimuthal structures and a nonring cluster. The height of the colored bar (lower left corner in each panel) represents the relative fraction of images in the Top Sets for each cluster. We note that the THEMIS posterior sample only includes ring images. In each cluster, the image is computed through a weighted average over the descattered reconstructions, including all Top Set images from the three imaging methods (DIFMAP, eht-imaging, and SMILI) and 1024 images randomly selected from descattered posteriors from THEMIS. Images are weighted by the inverse of the total number of Top Set or posterior images used from each pipeline, so that pipelines are represented equally in each image. Note that DIFMAP model images are convolved with a 20 μas beam (represented by the inset circle), while no blurring is applied to the rest of the images.

The bottom row of both the middle and bottom panels of Figure 14 shows the azimuthal distribution of the radial peak intensity within a radius of 10–40 μas. These profiles are sorted within each cluster by the location of peak brightness to best accentuate variations within a cluster. By inspecting the profiles in each cluster, it can be seen that three primary brightness distributions, with a peak brightness at $\sim -140^\circ$, -40°, or 70°, appear across the pipelines. Thus, even when we restrict our attention to ring-like morphology reconstructions, it is difficult

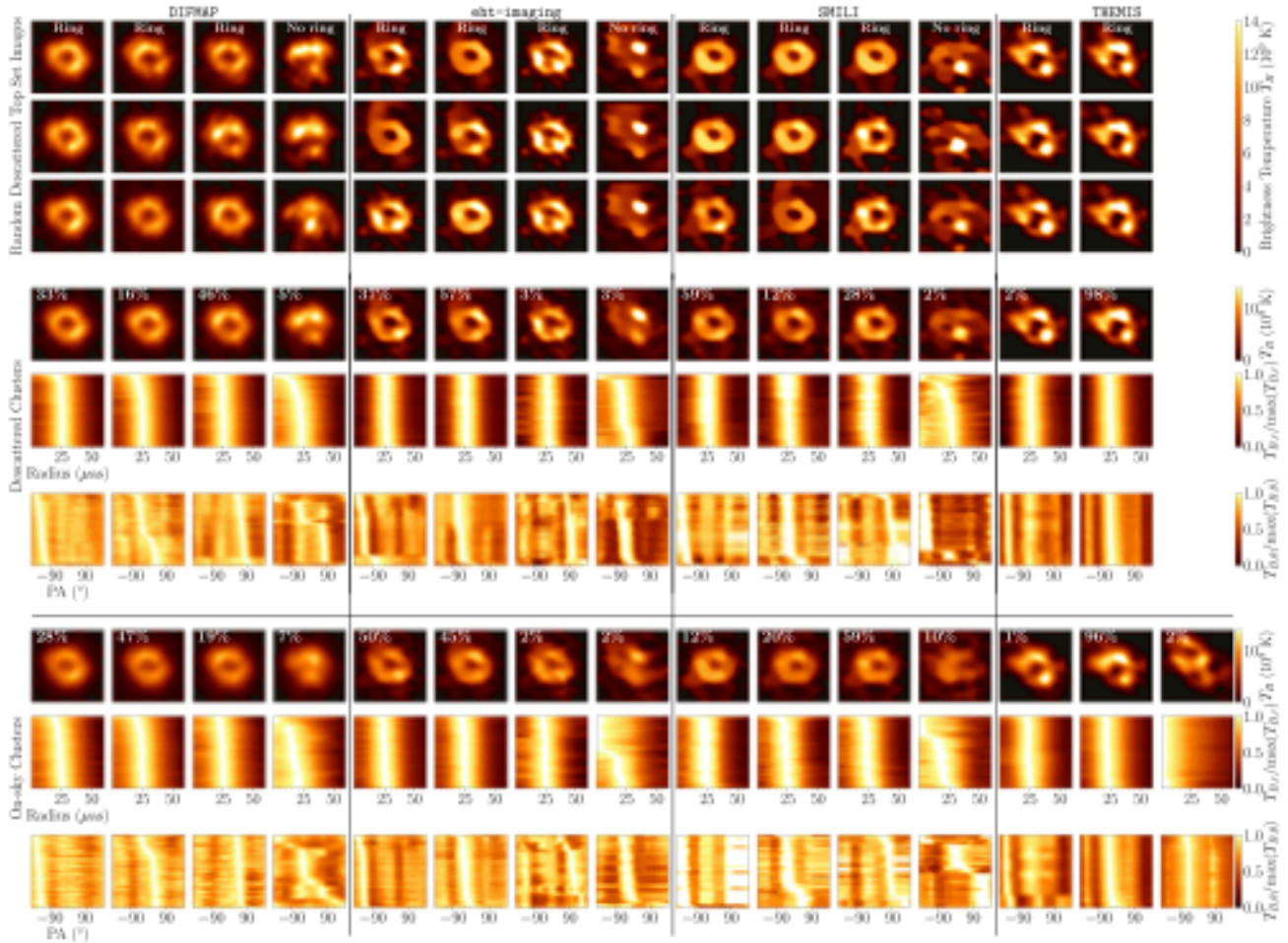


Figure 14. The distribution of reconstructed Sgr A* images on April 7. We show the distribution of images from each pipeline for each cluster with the same convention as Figure 12. From left to right (separated by vertical lines), we show the distributions of Top Set images from the DIFMAP, eht-imaging, and SMILI pipelines and posterior samples from the THEMIS pipeline; each vertical panel is further subdivided into clusters identifying common morphologies recovered by each pipeline. The figure is composed of three horizontal panels separated by horizontal lines. The top panel shows individual images randomly sampled from different clusters. The middle and bottom panels visualize the distributions of reconstructed descattered and on-sky Sgr A* images for each cluster, respectively. In each panel, from top to bottom, we show the average of each cluster, the distributions of the radial profiles, and the distributions of azimuthal intensity profiles. Note that THEMIS images have only three clusters for each of the descattered and on-sky reconstructions—their descattered posterior does not contain a nonring cluster, and their on-sky posteriors do not contain an east PA ring cluster.

to constrain the azimuthal profile around the ring via imaging. This azimuthal uncertainty could be due to data properties (e.g., sparsity or low-S/N data) or variation in the intrinsic azimuthal structure of Sgr A* due to intraday evolution.

To visualize the image-domain differences among each cluster, we compare in Figure 15 the relationship between the fractional central brightness (f_c, radial values) and the azimuthal peak brightness (PA, azimuthal values). The polar distribution of Top Sets and posterior images among all imaging pipelines and with and without scattering mitigation (left panel in Figure 15) shows that most of the ring images within each different cluster (indicated by the red, blue, and green points) have a smaller fractional central brightness ($f_c \lesssim 0.5$) than that for the images in the nonring cluster ($f_c \gtrsim 0.5$).

The histogram of the azimuthal peak brightness distributions shown in the left panel in Figure 15 provides a clear visualization of the clustering of ring images around PAs of $\sim -140°$, $-40°$, or $70°$, which correspond to the locations of the three knots that commonly appear in ring images. Changes in PA mostly reflect variations in the relative brightness of these knots. For the case of nonring images the PAs are, as expected, more randomly distributed. Figure 15 also confirms that the fractional central brightnesses are systematically larger for the on-sky images owing to the angular broadening produced by the interstellar scattering.

Both the ring and nonring morphologies identified in Top Set and posterior images show reasonable fits to Sgr A* data. In Appendix E, we show χ^2 distributions of Top Set images to Sgr A* data. After adding the budgets of nonclosing systematic errors, representative refractive noise, and time variability, all Top Set images result in a $\chi^2 < 2$ for both closure phases and log closure amplitudes—we refer the reader to Paper IV and M87* Paper IV for a discussion of these data products and χ^2 distributions. In Figure 16 we compare closure phases of Sgr A* to those of individual images from each cluster and pipeline for four selected closure triangles. As indicated by the χ^2 metrics, multiple ring and nonring images fit the observed closure phases within the range of deviations anticipated owing to temporal variability and refractive scattering effects.

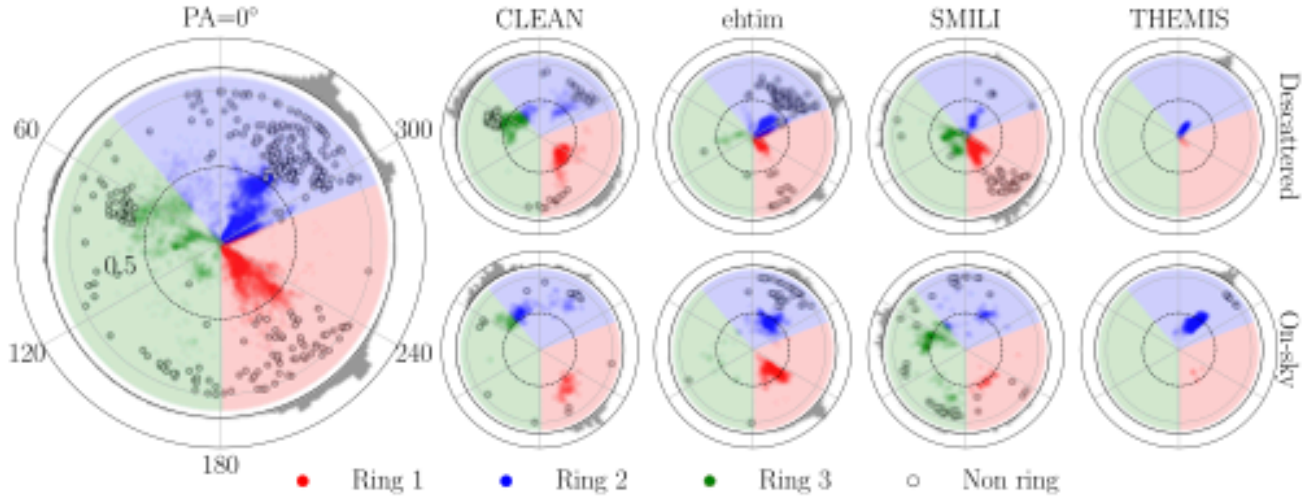


Figure 15. Comparison of image characteristics between each cluster. The radial and azimuthal values are the fractional central brightness and azimuthal peak brightness PA, respectively. The left panel shows the distribution of all Top Set and posterior images with all imaging pipelines (DIFMAP, eht-imaging, SMILI, and THEMIS) and scattering mitigation (descattered and on-sky), including the three ring clusters with different peak azimuthal brightness (red: $-180° \leqslant \mathrm{PA} < -70°$; blue: $-70° \leqslant \mathrm{PA} < 40°$; green: $40° \leqslant \mathrm{PA} < 180°$) and a nonring cluster (white). The right eight panels show distributions within a single imaging pipeline and scattering mitigation. The majority of azimuthal peak brightness among each ring cluster is shown in the outer gray histogram in each panel.

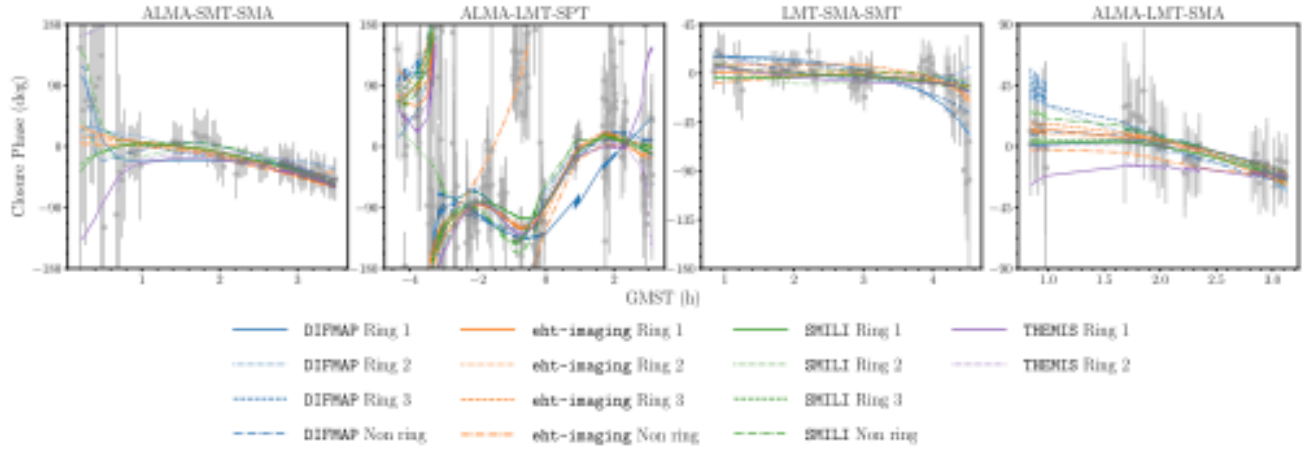


Figure 16. Closure phases plotted as a function of GMST on four selected triangles from April 7 observations. Each line indicates the corresponding closure phase curves from a single Top Set and posterior image randomly selected from each cluster. The error bars of Sgr A* data include the fractional 1% noise budget for systematic error and a representative budget for scattering and temporal variability—in particular, the `J18model` refractive noise model and a variability model with parameters $a = 0.02$, $u_0 = 2$, $b = 2.5$, and $c = 2$. These additional noise budgets are all added to the 60 s complex visibility noise budget prior to forming closure quantities. All images show reasonable fits to Sgr A* data, as they all are within two standard deviations of the observed data.

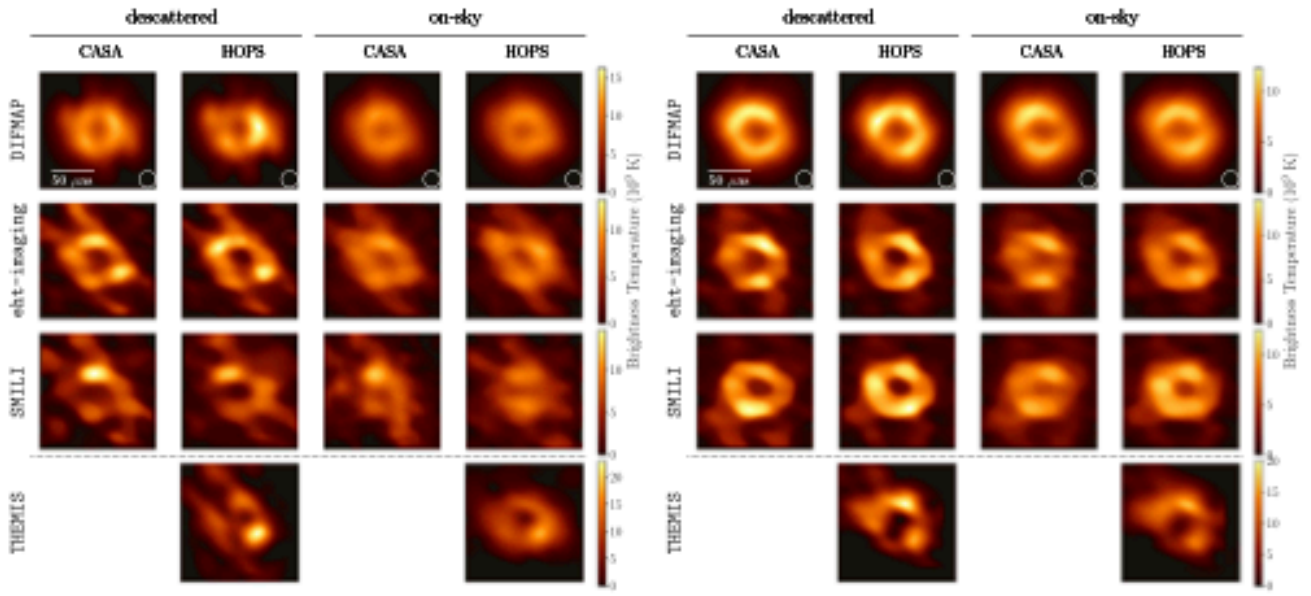
Therefore, none of the identified clusters can be excluded from possible Sgr A* morphologies in terms of the goodness of fit to Sgr A* data (or through synthetic data tests presented in Section 6).

7.3. Average Sgr A* Images across Pipelines

Figure 17 shows the average of Sgr A* images reconstructed by each of the four imaging pipelines (DIFMAP, eht-imaging, SMILI, and THEMIS) from each calibrated data set (CASA and HOPS) on each observing day (April 6 and 7). Only images from the HOPS data product have been reconstructed using the THEMIS pipeline. The images from DIFMAP, eht-imaging, and SMILI are obtained by averaging their respective Top Set images; these average images show the dominant features identified across different combinations of the selected imaging parameters in the Top Sets (refer to Section 6). For the THEMIS reconstructions we instead show the mean of each posterior obtained by averaging all the posterior samples. Figure 17 shows that the majority of images contain a ring-like structure. This ring morphology is common among all imaging pipelines and is resilient to the scattering mitigation strategy employed (as discussed in Section 7.5.2). Additionally, we recover largely consistent images between data calibrated by the HOPS and CASA calibration pipelines. Although we recover images with a ring morphology in the majority for all pipelines on April 7, these average images also highlight that the azimuthal brightness distribution is sensitive to small changes in the data and imaging strategy.

7.4. Imaging Combining April 6 and 7 Data Sets

In Figure 17 we show results from April 6 data, reconstructed using the same imaging procedure as April 7. We note that although a ring feature appears in most of these



(a) April 6 (b) April 7

Figure 17. The dominant recovered morphology in Sgr A* descattered and on-sky reconstructions identified from two VLBI data products (CASA and HOPS data) with all four imaging pipelines (DIFMAP, eht-imaging, SMILI, and THEMIS) for two observing days (April 6 and 7). Each panel shows the average image of the corresponding Top Set images for DIFMAP, eht-imaging, and SMILI pipelines and the average posterior image for the THEMIS pipeline. Only the HOPS data have been imaged using the THEMIS pipeline.

reconstructions, it is less prominent. The images also contain a diagonal rail-like feature going from northeast to southwest that corrupts the ring. This feature is especially prominent in the THEMIS April 6 descattered reconstruction. This corrupted or nonring mode is likely emphasized in THEMIS imaging compared to RML and CLEAN pipelines owing to the goal of THEMIS image samples to characterize the probability of an image rather than represent the variety of possible images that can fit the data.

Through an in-depth inspection of the April 6 data, documented in Paper IV, the Chile–LMT baselines (i.e., ALMA–LMT and APEX–LMT baselines) were identified as having large coherent visibility swings that are not effectively mitigated by the noise model on a single day, causing this particular feature to arise. Figure 18 shows the reconstructions using the DIFMAP, eht-imaging, and SMILI imaging pipelines when these particular baselines have been flagged, resulting in a cleaner ring structure.

The large visibility swings on the Chile–LMT baselines could be the cause of variability around Sgr A* that exceeds expectations set by the incorporated stationary noise model presented in Section 3.2. In particular, we believe that the variability noise model should capture Sgr A*'s stochastic evolution in expectation, but a single night may contain nonstochastic short-lived variability that can bias a single day's reconstruction. Correlated variability may be mitigated via multiday fits, which combine statistically independent structural fluctuations; this better matches the assumption within our variability mitigation scheme of an underlying stochastic process, though it does carry with it the additional assumption that Sgr A* is statistically stationary over the multiple days combined. Therefore, in addition to the single-day analyses performed by all imaging methods, the THEMIS image model was also fit to the combined April 6 and 7 scan-averaged data from high and low bands for the HOPS data product. The resulting static image shown in Figure 19 represents an image

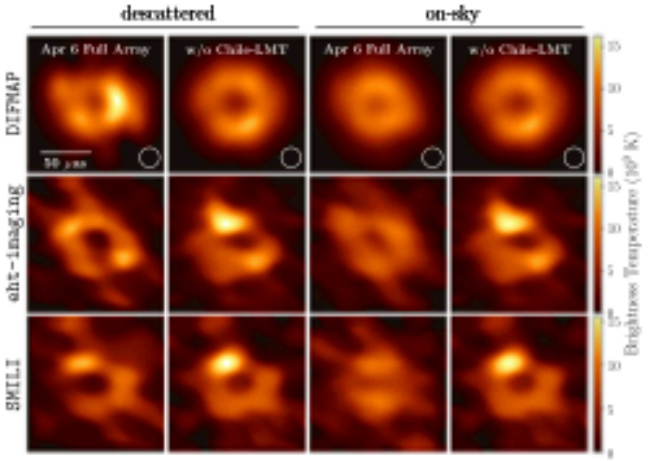


Figure 18. The effect of removing the Chile–LMT baseline from April 6 data reconstructions. Each panel shows the average image of the Top Set images for the DIFMAP, eht-imaging, and SMILI pipelines from April 6 and the HOPS data product for the descattered and on-sky reconstructions. For comparison we show the average images obtained from full data sets, as well as images obtained from data without the Chile–LMT baselines. This Chile–LMT baseline appears near the visibility null and appears to exhibit significant intraday variations on April 6 that are likely not captured by the variability noise model presented in Section 3.2.

recovered from the combined data sets. All images within the multiday THEMIS posterior exhibit a clear ring-like structure.

7.5. Is Sgr A* a Ring?

Our primary imaging goal is to answer the question, "Is Sgr A* a ring?" Although our reconstructions are overwhelmingly dominated by ring images, there are a small number of nonring images that fit the data well and cannot easily be excluded through additional tests. There are at least three

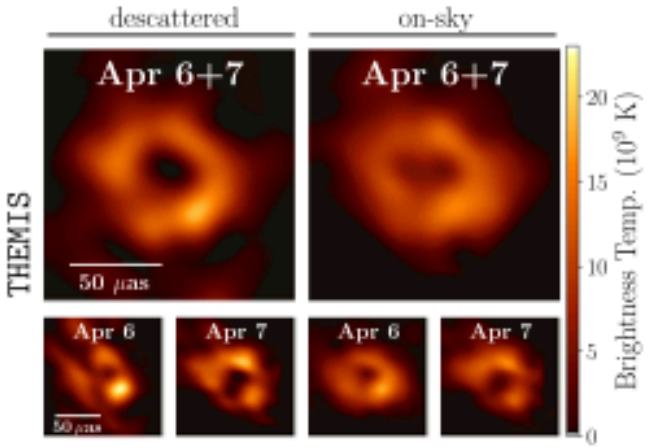


Figure 19. Mitigation of the apparent strong intraday variations seen in April 6 through THEMIS imaging of the combined April 6 and 7 data. For comparison we show the average posterior images from April 6 and 7 independently and that obtained by combining the data sets for April 6 and 7, exhibiting a clear ring for the descattered and on-sky reconstructions.

possible reasons for the recovery of nonring reconstructions of Sgr A*: (1) Sgr A*'s intrinsic structure is not ring shaped, (2) scattering causes a distortion of a ring morphology, resulting in a nonring image, and (3) imaging algorithms recover an incorrect source structure, aggravated by challenges of sparse (u, v)-coverage and Sgr A*'s intraday variation. In this section we will explore these three possible origins of nonring structure. Based on this exploration, we conclude that there is evidence that the nonring reconstructions are caused by our imaging algorithms resulting from the limited (u, v)-coverage and rapid variability, rather than being intrinsic to Sgr A*.

7.5.1. Manifestation of Rings from Intrinsic Nonrings

The first possible explanation for the small percentage of nonring reconstructions of Sgr A* is simply that Sgr A* does not possess a ring morphology. Although this possibility cannot be ruled out completely, we explore the possibility of a potential bias in our imaging approach toward recovering ring images from underlying nonring sources.

There is always a possibility that the parameters initially explored by the RML and CLEAN parameter surveys were inadvertently biased to produce mostly ring images. This would result in an artificially high percentage of ring reconstructions that could give overconfidence in an incorrect ring solution. To explore this hypothesis, we inspect the percentage of ring and nonring images of Sgr A* that are in the initial parameter survey versus the restricted Top Set on April 7. We find that although ring images make up a large percentage of the initial parameter surveys, 91%, 84%, and 60% for `DIFMAP`, `eht-imaging`, and `SMILI`, respectively, for the descattered reconstructions, this number becomes significantly larger after Top Set selection. In particular, the percentage of ring images rises to values of 95%, 97%, and 98%, respectively. Therefore, when we restrict to those parameters that can best disambiguate between the different ring and nonring source morphologies contained in our synthetic data sets (Section 5.1), the number of ring images increases. We also note that THEMIS posterior samples for April 7 only contain ring images.

To further test the possibility that we are biased to recover primarily rings from underlying nonring sources, we have explored the performance of our imaging methods on nonring synthetic data sets to see how they compare to Sgr A* results. In particular, we performed a Top Set analysis on the `Point` and `Double` source models. Two "cross-validation" Top Sets were identified by excluding the performance of either the `Point` or `Double` source model and only using the remaining six geometric models in selection of the imaging parameters. These two cross-validation Top Set parameter sets were then applied to the `Point` and `Double` synthetic data sets. Although each cross-validation Top Set does incorrectly reconstruct some ring images, each image set primarily contains reconstructions that look similar to the ground-truth `Point` and `Double` source structure. In particular, for descattered reconstructions, only 9% and <1% of the cross-validation Top Set reconstructions possess a ring morphology, compared to 97% in Sgr A* reconstructions of April 7. The THEMIS pipeline produces no ring images in the posterior samples for both the `Point` and `Double` data sets. Thus, in the nonring synthetic data sets we have explored we find that the number of incorrect ring reconstructions is much less than the number that we recover for Sgr A*. In summary, our imaging pipelines do not appear prone to artificially create a majority of ring structures in sources that do not possess an intrinsic ring morphology.

7.5.2. Scattering's Effect on Image Reconstruction

The second possible explanation for the nonring reconstructions of Sgr A* is that interstellar scattering causes a nonring morphology. This could take one of two forms: (a) Sgr A*'s ring morphology has been distorted to a nonring morphology by an interstellar scattering screen, or (b) our imaging algorithms have incorporated an incorrect scattering model that reconstructs a corrupted nonring morphology.

To address the possibility that interstellar scattering is causing Sgr A*'s intrinsic structure to be distorted to a nonring, we inspect differences in the descattered versus on-sky reconstructions. As outlined in Section 3.1, descattered reconstructions attempt to mitigate two primary effects of interstellar scattering: diffractive scattering that causes angular broadening, and refractive scattering that introduces small-scale structure to the on-sky image. Sgr A* reconstructions in Figure 14 show that on-sky images are systematically blurrier than descattered images, as expected owing to deblurring that is performed before recovering a descattered image (refer to Section 3.1.1). This systematic difference between on-sky and descattered images is also seen in synthetic data reconstructions presented in Section 6.4. This angular broadening causes the central dip of a ring to be less prominent. Nonetheless, we note that the vast majority of the on-sky (i.e., without any scattering mitigation prescription) images are still rings, with percentages of 93%, 98%, 90%, and 98% for CLEAN, `eht-imaging`, `SMILI`, and THEMIS, respectively. We note that it is highly unlikely that a nonring morphology was distorted into an on-sky ring morphology by interstellar scattering.

To address the possibility that our assumed model of the interstellar scattering is reconstructing a corrupted image of the descattered source, we explore the use of multiple refractive noise models to mitigate effects of refractive substructure before deblurring (Section 6.2). The contribution of the included refractive noise budget used to mitigate scattering (for descattered reconstructions) is nonnegligible for long-baseline data (see Section 3.1.2 and Figure 5). Nontheless, the RML and CLEAN imaging surveys show that the choice of the

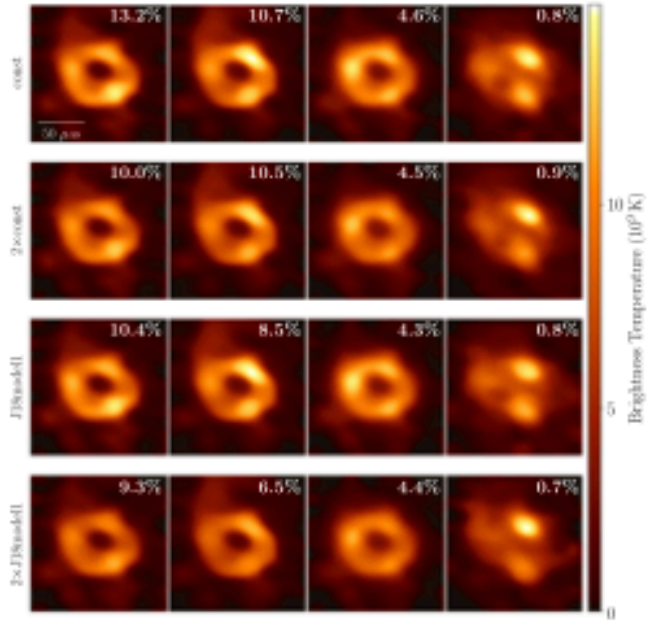


Figure 20. The range of descattered Sgr A* images recovered using different refractive scattering noise models. From left to right, images are shown clustered in the same order as Figure 15. Each panel shows an average of all descattered Top Set images of Sgr A* on April 7 from the DIFMAP, eht-imaging, and SMILI pipelines that were generated using the specified refractive noise model (from top to bottom: Const, 2×Const, J18model1, and 2×J18model1). The number on each panel shows the percentage of descattered Top Set images that were reconstructed using the specified refractive noise model. These percentages indicate that there is not a clear preference for a particular refractive noise model in Top Set selection. Additionally, the recovered image structure does not appear to correlate with the refractive noise model used.

refractive noise model does not significantly affect the resulting distributions of image structures. Figure 20 shows the average descattered Sgr A* images (clustered into the same four morphologies as are presented in Figure 13) for each refractive noise prescription explored. The comparable fractions of images in each cluster across the different descattering strategies indicate the resiliency of the recovered image structures to different refractive noise models. Another piece of evidence suggesting that the scattering prescription does not strongly affect results can be seen in the THEMIS results; in THEMIS a constant refractive noise floor is able to vary in posterior sampling of descattered images. Although an essentially unlimited refractive noise component is permitted by the THEMIS model, the posterior estimation instead typically chooses a noise level of <25 mJy, only 1% of the total flux of the source.

In summary, results indicate that for EHT measurements of Sgr A*, interstellar scattering does not significantly affect the recovered morphology. Both on-sky and descattered images contain a majority of ring morphologies. In addition, we find that the particular choice of scattering mitigation strategy only minimally affects the recovered image structures.

7.5.3. Variability's Effect on Image Reconstruction

The third possible origin for nonring Sgr A* images is poor reconstruction quality of the imaging methods. Imaging is solving an ill-posed inverse problem due to sparse (u, v)-coverage, which always will have the possibility of recovering an incorrect image. This is further exasperated in imaging Sgr A* by the challenges that come with recovering an evolving source. In this section we explore this possibility and find that our imaging methods often reconstruct nonring sources for variable ring sources with a comparable small percentage as found for Sgr A*.

The first natural question is how our imaging methods perform on reconstructing ring sources with similar data properties to Sgr A*. We find that our methods produce nonring images, even when the underlying source structure is ring-like. As an example, recall that for the variable GRMHD ring-like sources analyzed in Section 6 and Appendix H our imaging methods produced mostly rings, but also a very small fraction of nonring images. Therefore, we expect that for a variable ring source we would recover some nonrings that fit the data well. However, note that the number of nonrings was still fairly small in the case of the GRMHDs: 5% for the GRMHD presented in Figure 12, and 4% for the "best-bet" GRMHD presented in Figure 37. These values are comparable to the 3% of nonrings reconstructed for Sgr A* descattered images across all pipelines. We also find that cross-validated Top Set images reconstructed of the Crescent and Ring geometric sources produce a small fraction of nonring images. Note that these nonring percentages for the GRMHD, Crescent, and Ring sources are much less than the percentage of nonrings reconstructed for the variable Point and Double sources, as discussed in Section 7.5.1.

Due to Sgr A*'s intraday evolution, mitigation of temporal variability in the data is important for reconstructing a single static image of Sgr A*. We next explore how the variability mitigation approach affects the proportion of ring versus nonring images reconstructed. Section 3.2 introduced an approach to model the temporal variability as an additional noise budget that could be added in quadrature to the visibility thermal noise budget. The dependence of the temporal variability noise model parameters on resulting Sgr A* reconstructions for the RML and CLEAN pipelines was investigated. Although a variability noise budget generally helps with imaging (as evidenced by a higher percentage of Top Set parameters selected—89%–95% of the Top Set parameter combinations include a variability noise model on April 7), similar to scattering, we find that there are no significant differences between the images recovered under different variability noise parameters. We suspect that this is partly caused by the complex interplay between regularizers and different noise model parameters (e.g., variability, scattering, and systematic), although no significant trends were identified.

In Appendix G, we demonstrate how our results are insensitive to a different method of variability mitigation. In particular, we show time-averaged morphologies identified by a large survey over full-track RML dynamical imaging parameters (refer to Section 4.4.1) that do not rely on the variability noise model presented in Section 3.2.2. We again find the ring modes to be dominantly reconstructed, while a small fraction of nonring structures are also identified. The broad consistency indicates that our results are resilient to at least two different methods to recover time-averaged morphology.

In summary, we find that our imaging methods do reconstruct a small percentage of nonring images from ring sources with comparable variability to that seen in Sgr A* data. Similar behavior is observed across a variety of different

imaging approaches to mitigate variability. Therefore, we conclude that we would expect our methods to produce a small fraction of nonring images from an underlying variable ring morphology.

8. Image Analysis

8.1. Ring Parameter Fitting

To analyze the Sgr A* images, we use two tools, REx (Chael 2019) and VIDA[154] (Tiede et al. 2022), both of which are able to extract quantitative and pertinent information from the Top Set images. Here we briefly review the two image extraction techniques.

REx attempts to extract ring-like features by directly characterizing the features of the Top Set images. This is the same tool used in 2017 M87* analysis (M87* Paper IV). The detailed definitions of REx ring parameters follow those of the M87* analysis (M87* Paper IV).

In REx, the ring center (x_0, y_0) is determined so that the dispersion of intensity peak radii is minimized. Around the center, a polar intensity map $I(r, \theta|x_0, y_0)$ is constructed. The ring radius r_0 (or diameter $d = 2r_0$) is defined as the radius where azimuthally averaged radial intensity peaks. The ring width w is $\langle \mathrm{FWHM}[I(r, \theta|x_0, y_0) - I_{\mathrm{floor}}]\rangle_\theta$, where $I_{\mathrm{floor}} = I_{r_{\max}} = 60$ μas, $\theta|x_0, y_0\rangle_\theta$.[155] To characterize the azimuthal structure, we define the normalized first circular moment at radius r as

$$m_1(r) = \frac{\int_0^{2\pi} I(r, \theta|x_0, y_0)\cos(\theta)d\theta}{\int_0^{2\pi} I(r, \theta|x_0, y_0)d\theta}. \quad (11)$$

The ring PA η and asymmetry A are given by the radially averaged (from $r_0 - w/2$ to $r_0 + w/2$) argument and amplitude of $m_1(r)$, respectively. Finally, a central fractional brightness f_c is defined as a ratio of the mean brightness within 5 μas from the center to the azimuthally averaged brightness along the ring $(r = r_0)$.

VIDA takes a forward-modeling or template-matching approach for image analysis (Tiede et al. 2022). That is, we approximate the images with parametric families or templates f_Θ, such as rings, crescents, or Gaussians. The template used depends on both the observed image structure and the features of interest. VIDA's approach is therefore similar to geometric modeling presented in Paper IV, except that it is applied to the image reconstructions rather than in the visibility domain. The image features, such as diameter, are then given by the parameters of the optimal template. This differs from REx, which defines its quantities directly on the image.

To find the optimal template, we first renormalize the Top Set image to form a unit flux image $\hat{I}(x, y)$. We then take the L2 norm as our objective function:

$$J(\Theta) = \mathrm{LS}(f_\Theta||\hat{I}) = \int |f_\Theta(x, y) - \hat{I}(x, y)|^2 dxdy, \quad (12)$$

where Θ denotes the template parameters.

For analyzing Sgr A* Top Set images, we use VIDA's SymCosineRingwFloor template. This template is characterized by a ring center (x_0, y_0), diameter d, FHWM width w, and a cosine expansion to describe azimuthal brightness distribution S:

$$S(\theta) = 1 - 2\sum_{m=1}^{M} A_m \cos[m(\theta - \eta_m)]. \quad (13)$$

For this paper we take $m = 4$. Note that we also restrict $A_m < 0.5$ to restrict negative intensity in the template.[156] The PA of the image is taken as the phase of the first-order cosine expansion, i.e., $\eta_1 = \eta$. Similarly, we define the asymmetry $A = A_1$ to match REx's definition above. Note that this azimuthal structure is very similar to the m-ring model described in Section 9.

In addition, we add a central disk to constrain the central brightness depression of the ring. This disk is forced to have the same radius as the ring, and a Gaussian taper is included that matches the width of the ring. We then compute the central fractional brightness, f_c, of the optimal template, following the definition above.

In the next subsection, we provide the results of both REx and VIDA for five ring parameters, namely, ring diameter d, PA η, ring width w, fractional central brightness f_c, and asymmetry A. Note that while each method's parameter definitions are similar to each other, REx produces estimates on the image directly, while VIDA parameter values are defined from the optimal template. If the template provides a "good" approximation to the on-sky image, these estimates should be similar. Note that the negative pixels of THEMIS images are treated as zero in the REx and VIDA analyses.

8.2. Ring Fitting Results

In this subsection, we present ring parameter results only for images from ring morphology clusters (Figure 13). In Figure 21, we show the diameters measured for on-sky and descattered ring images for each pipeline. We find that the ring images are consistent with a diameter of ~50 μas, as shown for each ring morphology cluster separately in Section 7.2 and Figure 14.

For more detailed distributions of ring parameters, in Figure 22 we show the ring parameter fitting results of Sgr A* descattered Top Set or posterior images for April 7. From top to bottom, distributions of ring radius, PA, ring width, fractional central brightness, and asymmetry are presented for four pipelines. Results from both REx and VIDA are shown, and for comparison azimuthally or radially averaged brightness distributions are shown in the background of diameter d and PA η plots.

Radius distributions are clearly peaked at ~25 μas and mostly concentrated between 25 and 30 μas, being consistent across the three pipelines. Note that on some occasions REx and VIDA show discrepancy in the radius (and some other parameters), which is mostly due to the difference of the responses to an image with a salient feature. However, as seen in the kernel density distributions of the radius, contributions of such outliers are negligible for determinations of the mean ring radius. Meanwhile, the PA value is by far less consistent across the pipelines or even within a single pipeline. Multiple modes clearly appear in the Top Set images with various PA values, as already seen in Section 7. Note that the scatter of PAs tends to

[154] https://github.com/ptiede/VIDA.jl

[155] We set $r_{\max} = 60$ μas instead of the 50 μas used in M87* Paper IV owing to the larger ring size.

[156] It is still possible for negative intensity with this restriction. To prevent negative intensity, we further take $\max(0, S(\theta))$ when computing the template.

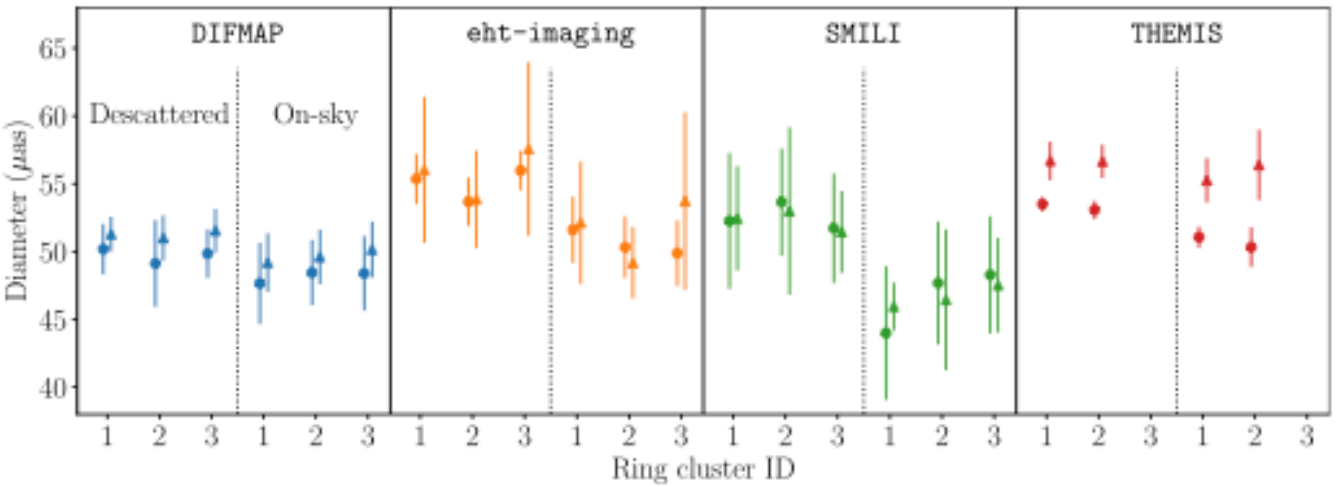


Figure 21. The ring diameters measured from April 7 images, shown separately for each cluster with a ring morphology, each pipeline, and descattered or on-sky reconstructions. Circles and triangles and associated error bars indicate the means and standard deviations of diameters measured with REx and VIDA, respectively. Note that THEMIS error bars appear significantly smaller than the error bars on DIFMAP, eht-imaging, and SMILI. This is partly due to the fact that THEMIS is primarily quantifying aleatoric (e.g., statistical) uncertainty, whereas the goal of the other imaging surveys is to also characterize epistemic (e.g., systematic) uncertainty.

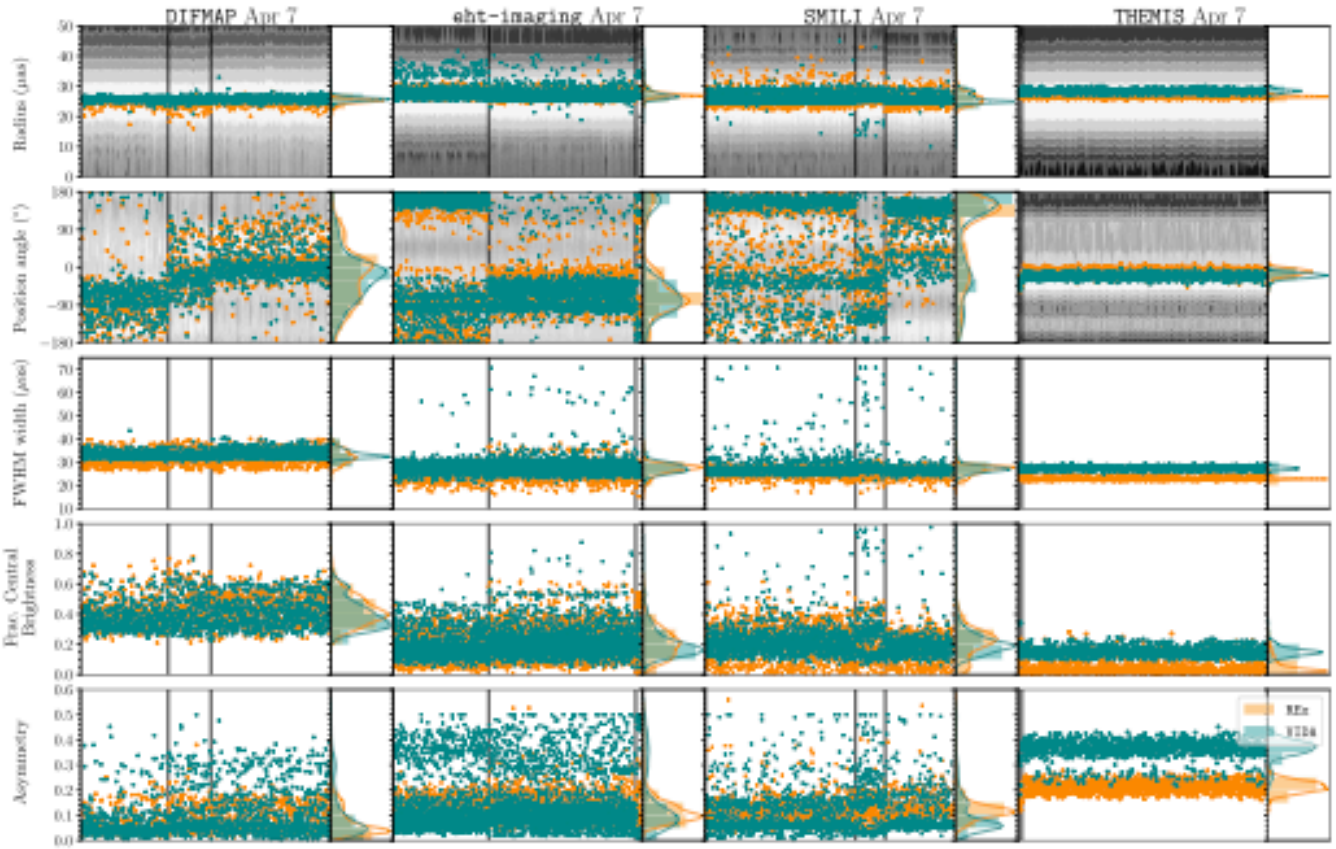


Figure 22. Ring fit results for Sgr A* Top Set descattered images in the ring clusters reconstructed from April 7 data. Each panel shows the distribution of ring parameters coresponding to the images resulting from a single pipeline. In each panel, the scatter plot on the left shows a ring parameter extracted from each image using REx (orange) and VIDA (green); the vertical histogram on the right shows the resulting kernel density estimation from the collection of extracted image parameters. Images are ordered by clusters described in Section 7.2, whose boundaries are shown with vertical solid lines. From left to right, the panels show the results for DIFMAP, eht-imaging, SMILI, and THEMIS. For the radius and PA, we also show the radial/azimuthal brightness distribution of each image in the background.

be larger when the ring has multiple bright spots that affect the resulting PA, or when the azimuthal profile is close to uniform without a clear peak. Ring width values are $\sim$30 μas, and this may come from the angular resolution of the observation. Fractional central brightness is mostly $\sim$0–0.3 for RML and THEMIS, while minor nonring modes tend to show somewhat

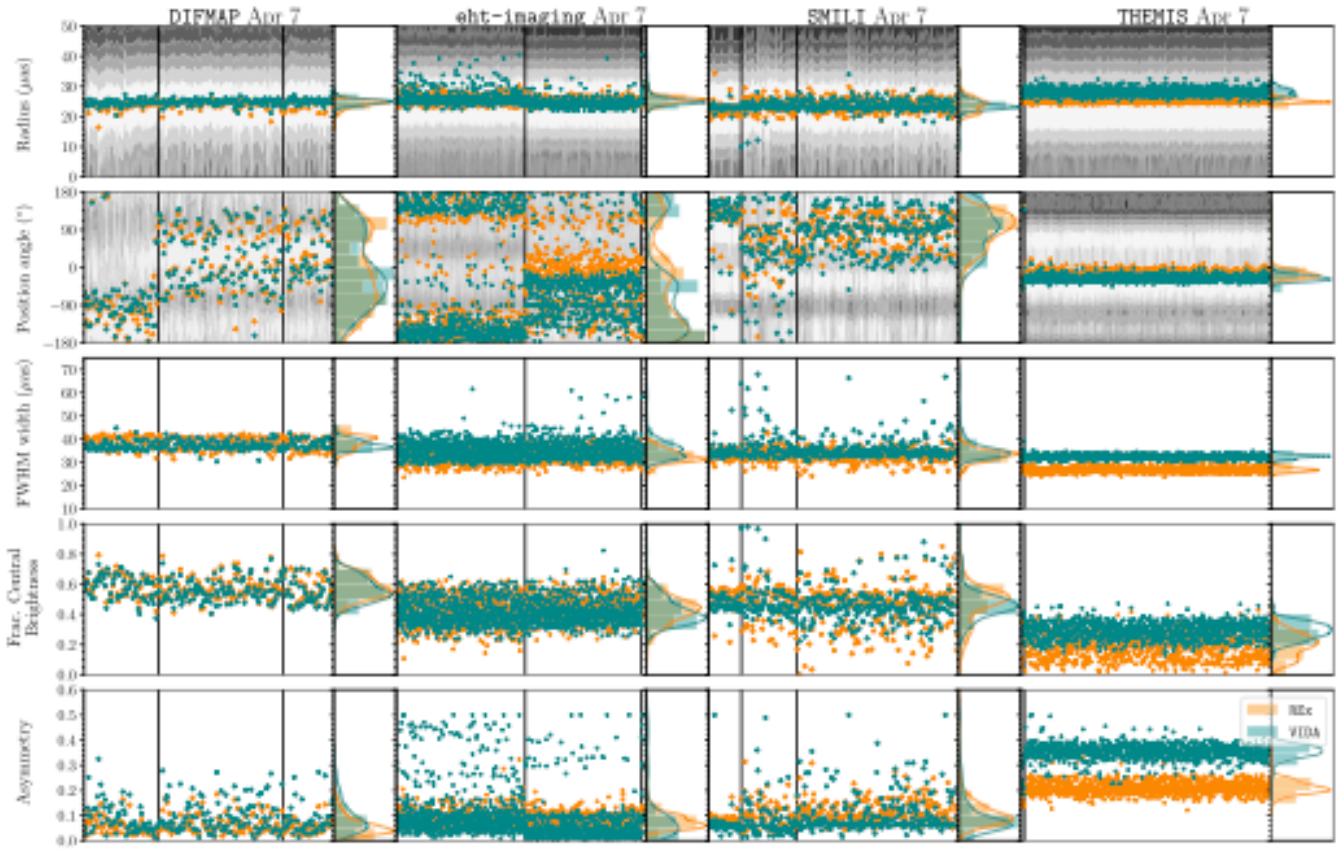


Figure 23. Ring fit results for Sgr A* Top Set on-sky images reconstructed from April 7 data. Refer to the caption of Figure 22 for more details.

higher values. These values confirm that a majority of the images show ring-like structures with clear central depression. CLEAN images tend to have a slightly larger value of width and central brightness than RML and THEMIS, as expected owing to the beam-convolution effect for CLEAN imaging. Asymmetry values are ∼0.1, indicating that most of the Sgr A* images have a nearly symmetric azimuthal intensity distribution on the ring. Apart from time variability, Sgr A*'s apparent symmetry could be one of the possible reasons for difficulty in constraining the PA.

Figure 23 shows the ring parameter fitting results of the on-sky Top Set or posterior images. Comparing the ring parameters with and without scattering mitigation, PA values are slightly different and FWHM and fractional central brightness values are larger for the on-sky images. These differences are reasonable when one considers the angular broadening effect due to scattering that remains in on-sky images. On the other hand, radius and asymmetry show similar distributions between descattered and on-sky images—we obtain a ring diameter of ∼50 μas for on-sky images, indicating that the ring size is robust with little dependence on the scattering correction.

In Table 7, we summarize the mean diameters and their standard deviations over the ring images reconstructed by each pipeline for both descattered and on-sky images and on both April 6 and 7. As seen in the table, the mean ring radii for on-sky images are slightly smaller than those of the descattered cases by a few microarcseconds. This small reduction of ring radii is mainly due to the difference of effective resolutions with/without scattering mitigation. Nevertheless, the ring radii are within their standard deviations and thus broadly consistent regardless of scattering mitigation.

Comparing the results of April 6 and 7 in Table 7, diameters derived from ring reconstructions are consistent within the computed standard deviations, regardless of the pipelines. The other parameters (except for PA) also give consistent values for both days. Again, PA values have a large scatter in April 6 images; the existence of large scatter in PA indicates that it is difficult to constrain the azimuthal brightness distributions along the ring. In general, most of the Top Set images show a ring morphology with a consistent diameter around ∼50 μas.

In Appendix I we list the fitted diameter, width, PA, asymmetry, and fractional central brightness measured for each one of the identified imaging clusters and different pipelines shown in Figures 22 and 23.

9. Short-timescale Dynamic Properties on Select Observation Window

The dynamical timescale at the location of the innermost stable circular orbit for Sgr A*, $t_g = 12\pi\sqrt{6}\,GM/c^3$ for zero spin, is approximately 30 minutes and can be smaller by a factor of ∼10 if the black hole is spinning rapidly. Variability at these timescales across the electromagnetic spectrum, including at 230 GHz, is one of Sgr A*'s salient features (see Wielgus et al. 2022 and references therein). As discussed in Section 3.2 and Paper II, a few EHT closure phase triangles show measurable variability across the 2017 observing campaign that can be attributed to intrinsic source variability. In this section, we explore the level and characteristics of

Table 7
Mean and Standard Deviation of Diameter d and Width w Measured from Top Set or Posterior Sgr A* Images for Each Pipeline

Descattered			
		d (μas)	w (μas)
DIPMAP			
Apr. 6 ring	REx	46 ± 4.1	33 ± 3.5
	VIDA	51 ± 3.1	33 ± 3.1
Apr. 7 ring		49 ± 2.1	32 ± 2.9
		51 ± 1.5	33 ± 1.7
eht-imaging			
Apr. 6 ring		56 ± 4.5	24 ± 2.4
		59 ± 11.3	30 ± 10.4
Apr. 7 ring		54 ± 2.0	26 ± 2.6
		54 ± 4.6	27 ± 3.5
SMILI			
Apr. 6 ring		57 ± 3.4	24 ± 1.9
		46 ± 12.0	50 ± 16.6
Apr. 7 ring		52 ± 4.7	26 ± 2.1
		52 ± 4.0	27 ± 4.7
THEMIS			
Apr. 6 ring		51 ± 3.9	25 ± 1.2
		54 ± 0.9	24 ± 0.9
Apr. 7 ring		53 ± 0.7	22 ± 0.5
		56 ± 1.2	27 ± 0.7
On-sky			
		d (μas)	w (μas)
DIPMAP			
Apr. 6 ring	REx	46 ± 3.0	34 ± 4.3
	VIDA	47 ± 1.9	39 ± 3.4
Apr. 7 ring		48 ± 2.7	38 ± 2.7
		49 ± 2.1	37 ± 2.0
eht-imaging			
Apr. 6 ring		49 ± 3.9	28 ± 3.6
		50 ± 5.6	41 ± 6.7
Apr. 7 ring		50 ± 2.5	32 ± 2.7
		50 ± 4.1	35 ± 3.8
SMILI			
Apr. 6 ring		43 ± 0.4	28 ± 3.1
		39 ± 3.9	46 ± 8.8
Apr. 7 ring		47 ± 4.7	33 ± 2.7
		47 ± 3.8	35 ± 5.2
THEMIS			
Apr. 6 ring		46 ± 2.1	30 ± 1.6
		47 ± 3.1	33 ± 2.0
Apr. 7 ring		50 ± 1.5	26 ± 1.1
		56 ± 2.6	32 ± 0.9

structural changes in the Sgr A* image that are consistent with the observed variability.

Recovering time-resolved structures on these short timescales is especially challenging owing to the sparse snapshot (u, v)-coverage for the EHT array. Indeed, without additional constraints, any observed change in the visibility domain can be interpreted as caused either by intrinsic variability or simply by the rotation of the baselines with Earth and their probing of different spatial structures—though fitting fast fluctuations in the visibilities with static emission requires larger fields of view. This is especially true for baselines that probe regions of the (u, v)-space in which the visibility amplitudes show deep minima (or nulls), across which the complex visibilities change by order unity over infinitesimal changes in baseline length.

In attempts to describe the EHT observations with a static image, we assign any observed variation to spatial structures and mitigate potential effects of variability by inflating the error budget. In this section, we instead attempt to fit the time-evolving data directly to produce spatially and temporally resolved images of Sgr A* on minute timescales. Our analysis of dynamic properties presupposes that the 230 GHz emission from Sgr A* is compact (see Section 2.3) and ring-like, such that the short-timescale variability we see can be attributed to changes in the image with time. We combine two independent analysis methods—dynamic imaging with temporal regularization between frames and snapshot geometric modeling—to identify trends in the spatial evolution of Sgr A*.

Our analysis shows that significant uncertainty exists in any attempt to characterize the spatially resolved dynamics of Sgr A* using EHT 2017 data. We expect that future observations with an expanded EHT array will yield significantly improved time-resolved and spatially resolved movies of Sgr A*.

9.1. Selecting an Observation Window

The rotation of Earth causes the EHT's snapshot (u, v)-coverage to change over time. Static imaging and modeling approaches assume that the source is unchanging in time, which allows these approaches to combine data from a full night of observations. However, recovery of short-timescale evolution requires that we only consider coverage synthesized on the variability timescale. This "snapshot" coverage is extremely sparse and introduces artifacts into image reconstructions. To minimize these artifacts, we constructed and evaluated metrics to assess the performance of the snapshot coverage and identify the most promising time windows for dynamic analysis. These metrics rely purely on the (u, v)-coverage rather than the properties of the underlying Sgr A* visibilities. The construction and validation of a suite of these metrics are reviewed in Farah et al. (2022).

We consider metrics that assess several attributes of the (u, v)-coverage, including the largest gap in coverage (Wielgus et al. 2020), the fraction of the (u, v)-plane covered (Palumbo et al. 2019), and the geometric properties of the coverage (Farah et al. 2022). We summarize the application of these metrics to the 2017 April 7 EHT Sgr A* data set in Appendix J. These three metrics identify a period from approximately 1.5 to 3.2 GMST on April 6 and 7 that maximally mitigates the EHT's snapshot coverage limitations. During this time window, all sites participate in observing Sgr A* except PV, though there is a notable dropout of the LMT between approximately 2.4 and 2.9 GMST on both days. All dynamic analyses discussed in the remainder of Section 9 are performed *only in this selected time window*.

Figure 24 shows the (u, v)-coverage for April 7 during the ~100-minute observation window selected for dynamic analysis, along with the coverage for a single 60 s "snapshot" integration. Closure phases from two informative triangles are overlaid for April 6 and 7 during this time window. These closure phases show distinct evolution in the resolved structure of Sgr A* during the same 100-minute window on April 6 and 7.