

Article

The ngEHT Analysis Challenges

Freerk Roelofs ^{1,2,*} , Lindy Blackburn ^{1,2} , Greg Lindahl ¹ , Sheperd S. Doeleman ^{1,2} , Michael D. Johnson ^{1,2} , Philipp Arras ^{3,4} , Koushik Chatterjee ^{1,2} , Razieh Emami ¹ , Christian Fromm ^{5,6,7} , Antonio Fuentes ⁸ , Jakob Knollmüller ^{4,9} , Nikita Kosogorov ^{10,11} , Hendrik Müller ⁷ , Nimesh Patel ¹ , Alexander Raymond ^{1,2} , Paul Tiede ^{1,2} , Efthalia Traianou ⁸  and Justin Vega ^{1,12} 

¹ Center for Astrophysics | Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA

² Black Hole Initiative, Harvard University, 20 Garden Street, Cambridge, MA 02138, USA

³ Technical University Munich (TUM), Boltzmannstr. 3, 85748 Garching, Germany

⁴ Max-Planck Institute for Astrophysics, Karl-Schwarzschild-Str. 1, 85748 Garching, Germany

⁵ Institut für Theoretische Physik und Astrophysik, Universität Würzburg, Emil-Fischer-Str. 31, 97074 Würzburg, Germany

⁶ Institut für Theoretische Physik, Goethe-Universität Frankfurt, Max-von-Laue-Straße 1, 60438 Frankfurt am Main, Germany

⁷ Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany

⁸ Instituto de Astrofísica de Andalucía-CSIC, Glorieta de la Astronomía s/n, E-18008 Granada, Spain

⁹ Excellence Cluster ORIGINS, Boltzmannstr. 2, 85748 Garching, Germany

¹⁰ Moscow Institute of Physics and Technology, Institutsky per. 9, 141700 Dolgoprudny, Russia

¹¹ Lebedev Physical Institute of the Russian Academy of Sciences, Leninsky Prospekt 53, 119991 Moscow, Russia

¹² Department of Physics, Northeastern University, 360 Huntington Ave, Boston, MA 02115, USA

* Correspondence: freerk.roelofs@cfa.harvard.edu

Abstract: The next-generation Event Horizon Telescope (ngEHT) will be a significant enhancement of the Event Horizon Telescope (EHT) array, with ~ 10 new antennas and instrumental upgrades of existing antennas. The increased uv -coverage, sensitivity, and frequency coverage allow a wide range of new science opportunities to be explored. The ngEHT Analysis Challenges have been launched to inform the development of the ngEHT array design, science objectives, and analysis pathways. For each challenge, synthetic EHT and ngEHT datasets are generated from theoretical source models and released to the challenge participants, who analyze the datasets using image reconstruction and other methods. The submitted analysis results are evaluated with quantitative metrics. In this work, we report on the first two ngEHT Analysis Challenges. These have focused on static and dynamical models of M87* and Sgr A* and shown that high-quality movies of the extended jet structure of M87* and near-horizon hourly timescale variability of Sgr A* can be reconstructed by the reference ngEHT array in realistic observing conditions using current analysis algorithms. We identify areas where there is still room for improvement of these algorithms and analysis strategies. Other science cases and arrays will be explored in future challenges.

Keywords: very long baseline interferometry; black holes; active galactic nuclei; radio astronomy; imaging; instrument design; telescopes; algorithms; data analysis



Citation: Roelofs, F.; Blackburn, L.; Lindahl, G.; Doeleman, S.S.; Johnson, M.D.; Arras, P.; Chatterjee, K.; Emami, R.; Fromm, C.; Fuentes, A.; et al. The ngEHT Analysis Challenges. *Galaxies* **2023**, *11*, 12. <https://doi.org/10.3390/galaxies11010012>

Academic Editor: Margo Aller

Received: 10 November 2022

Revised: 16 December 2022

Accepted: 21 December 2022

Published: 10 January 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. The ngEHT

The Next-Generation Event Horizon Telescope (ngEHT) [1,2] will build on the success of the Event Horizon Telescope (EHT), the mm VLBI array, which has imaged the black hole shadows of M87* and Sgr A* [3–16]. The array will be transformatively enhanced with the current design envisioning ~ 10 additional stations, a quadrupled bandwidth, and frequency coverage, including 86 [17], 230, and 345 GHz. Multiple operating modes will make it suitable for a wide array of science cases. The primary science goals will involve making movies of M87* and Sgr A* resolving the plasma dynamics on event horizon scales,

providing black hole photon ring measurements sufficiently accurate to put constraints on black hole spin, and increasing the sample of black hole shadows imaged [18,19].

1.2. Challenge Motivation

End-to-end science simulations, which cover the full source physics, observation, calibration, and analysis processes, are of great value for the design and optimization of new instrumentation in astrophysics. These simulations realistically predict what the capabilities of the new instrument will be and which science questions it will be able to answer and can help guide the instrument design and analysis algorithm development. The ngEHT Analysis Challenges aim to provide such end-to-end simulations, bringing together expertise in all relevant areas to be applied to a well-defined set of problems. The challenge concept was inspired by the EHT Imaging Challenges [20]. In these challenges, EHT imaging experts imaged synthetic EHT datasets of different source models, which led to the rapid development of imaging algorithms and strategies tailored to the specifics of EHT datasets. While the EHT Imaging Challenges were aimed at maximizing the image quality that can be obtained from a known instrument, the ngEHT Analysis Challenges aim to help guide the development of a new instrument. Additionally, the ngEHT concept allows for the expansion of the imaging into two new dimensions, which are frequency (the ngEHT will operate at 2–3 distinct frequency bands simultaneously) and time (movie making). While not the focus of the challenges reported in this work, we aim to extend the ngEHT Analysis Challenges to model fitting and parameter estimation as well.

1.3. Challenge Procedure

For each challenge, we generate synthetic datasets from a set of source models. The source models (see also [21]) are representative of a specific ngEHT science case, may be static or time-variable, and may be generated for different (potential) ngEHT frequencies (86, 230, and 345 GHz in this work). The synthetic datasets are generated for different arrays (here, the 2022 EHT array and an ngEHT reference array) and contain different levels of data complexities (e.g., systematic weather or instrument noise). For each challenge, the synthetic datasets and other information and instructions are released to the challenge participants through the ngEHT Analysis Challenge website¹. Participants then upload their analysis results through the same website before a pre-set deadline. Image and movie reconstructions are then uniformly plotted for visual comparison and evaluated using quantitative metrics (see Section 3). Participation is open to anyone, with access to the downloads provided upon request to the organizers.

1.4. Outline

In this work, we report on the first two ngEHT Analysis Challenges. Section 2 details the reconstruction algorithms used by the challenge participants, and Section 3 describes the submission evaluation metrics. Challenge 1 and 2 source models, synthetic data generation, and results are presented in Sections 4 and 5, respectively, and the conclusions and outlook are discussed in Section 6.

2. Reconstruction Methods

In radio interferometry, image reconstruction is an underconstrained problem, as the finite number of telescopes and baselines cause only a limited number of Fourier components of the image (visibilities) to be measured. Hence, an infinite number of images could fit the data, and additional assumptions need to be made in order to arrive at a unique image solution. Different image reconstruction algorithms tackle this problem in different ways. The algorithms can be divided into inverse modeling, regularized maximum likelihood (RML), and Bayesian methods. The section below describes the algorithms used for the image reconstructions in this work, separated into methods reconstructing static images and methods reconstructing movies.

An alternative method to reconstruct the sky brightness distribution is fitting (geometrical) models to the interferometric data. Since such reconstructions have not been submitted for the challenges described in this paper, we do not discuss them here. We aim to explore these methods in future challenges, aimed at measuring specific black holes and accretion parameters.

2.1. Static Imaging

2.1.1. CLEAN

The CLEAN algorithm is a well-known inverse modeling imaging technique. The basic algorithm was developed by Högbom [22], with other variants developed later. CLEAN deconvolves a sampling function (known as the dirty beam) from the measured brightness (or dirty map) of a radio source. The imaging procedure via CLEAN involves a number of iterations, where in each iteration, the algorithm creates a point-source component, the CLEAN component, at the position of the brightness peak in the dirty image. Then, it convolves the CLEAN component with the dirty beam, subtracting it from the dirty image and transferring it to the clean map [22]. The cleaning iterations continue until a specific cleaning halting requirement is met. In the case of noisy data, the user can steer the process by limiting the searching area with CLEAN windows. Finally, the generated set of CLEAN components is convolved with a Gaussian restoring beam. The image quality can be further enhanced via self-calibration and references thereafter [23], which corrects the amplitude and phase information using the current image estimate. The residual dirty image, representing the image noise level, may be added to the clean map as the final step.

During the last decades, this technique has been widely used for imaging astronomical targets, as well as for a broad range of other applications [24]. Together with *eht-imaging* and SMILI (Section 2.1.2), it was one of the methods used for reconstructing the first EHT images of M87* [3–8] and Sgr A* [11–16].

The strategy followed for the ngEHT Analysis Challenges used a semi-scripted approach, in a similar fashion to the one in [13], employing the CLEAN algorithm via the software DIFMAP and references thereafter [25].

2.1.2. RML Methods: EHT-Imaging and SMILI

RML methods calculate each pixel of the source image \mathbf{I} by fitting directly to the data \mathbf{D} , with the fidelity of the final image to be adjusted by specific regularization terms, e.g., [26–28]. The data \mathbf{D} consists of separate data products d . These are typically visibility amplitudes, closure phases, or (log) closure amplitudes, see, e.g., [29]. These regularizers R could entail the entropy, sparsity, smoothness, or other properties of the image. For more details on regularizer definitions, see Appendix A of Event Horizon Telescope Collaboration et al. [6]. RML methods find an image that minimizes a specified objective function,

$$J(\mathbf{I}) = \sum_d \alpha_d \chi_d^2(\mathbf{I}) - \sum_R \beta_R S_R(\mathbf{I}), \quad (1)$$

consisting of goodness-of-fit (χ_d^2) and regularization (S_R) terms, weighted by hyperparameters (α_d and β_R).

Both the *eht-imaging* [29,30] and SMILI [31,32] frameworks are suitable for directly using the closure phases and (closure) amplitudes, making them ideal for high-frequency interferometric imaging [33]. As for CLEAN, multiple rounds of self-calibration are often performed.

Various submitters used different regularizers and weights for producing reconstructions of Challenge 1 and 2 data. For the M87 datasets, an informed prior was often used consisting of a small Gaussian with most of the flux, corresponding to the core, and a large disk with little flux to capture the extended emission from the jet. For the Sgr A* images, a disk or Gaussian prior was often used, deblurring the data and, in some cases, applying a constant noise floor to mitigate the intergalactic scattering before imaging. *eht-imaging* was also used to produce multi-frequency images, regularizing the spectral index map [34].

2.2. Dynamical Imaging

2.2.1. EHT-Imaging

The dynamical imaging module of `eht-imaging` (abbreviated to `ehtim-di` in this work) generalizes static imaging using a regularized maximum likelihood approach to reconstruct movies of time-variable sources [35]. Specifically, the reconstruction consists of a series of N_t images (movie), $\mathbf{M} = \{\mathbf{I}_1, \mathbf{I}_2, \dots, \mathbf{I}_{N_t}\}$. These images are determined by minimizing an objective function,

$$J \equiv \sum_d \alpha_d \chi_d^2(\mathbf{M}) - \sum_R \beta_R \left[\frac{1}{N_t} \sum_{j=1}^{N_t} S_R(\mathbf{I}_j) \right] + \sum_x \gamma_x \mathcal{R}_x(\mathbf{M}). \quad (2)$$

The objective function consists of three components:

- Like for static RML imaging, a data term which defines the log-likelihood of the reconstruction with respect to whatever data products are fit.
- A spatial regularization term, where for each regularizer, we compute a weighted sum over individual image regularization terms, $S_R(\mathbf{I}_j)$.
- A dynamical regularization term with temporal regularizers $\mathcal{R}_x(\mathbf{M})$ with associated hyperparameters γ_x . This term computes a penalty function that can be used to favor reconstructions that evolve smoothly in time ($\mathcal{R}_{\Delta t}$), that have small variations relative to the mean ($\mathcal{R}_{\Delta I}$), or that evolve according to fluid motion with a steady flow ($\mathcal{R}_{\text{flow}}$).

For the dynamical imaging reconstructions in the analysis challenges, we first fit a simple geometrical model to the full dataset, i.e., a thick “ m -ring”; see [14,36]. We then used this model as both a prior (for relative entropy of individual images) and initialization of reconstructed movies, with a typical frame separation of 1 min for Sgr A*. We fit amplitudes and closure phases, with iterative self-calibration of the visibility amplitudes. The imaging was performed using gradient descent with the limited-memory Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm [37], as implemented in Scipy [38].

2.2.2. StarWarps

The `StarWarps` algorithm [39] reconstructs time-variable sources by simultaneously reconstructing both the image and its time evolution. `StarWarps` reconstructs N_t images $\mathbf{M} = \{\mathbf{I}_1, \mathbf{I}_2, \dots, \mathbf{I}_{N_t}\}$, using the observational data snapshots $\mathbf{D} = \{\mathbf{D}_1, \mathbf{D}_2, \dots, \mathbf{D}_{N_t}\}$ at the corresponding timestamps. It employs a dynamical imaging model φ at each timestamp j :

$$\varphi_{\mathbf{D}_j|\mathbf{I}_j} = \mathcal{N}_{\mathbf{D}_j}(f_j(\mathbf{I}_j), \mathbf{R}_j), \quad (3)$$

$$\varphi_{\mathbf{I}_j} = \mathcal{N}_{\mathbf{I}_j}(\boldsymbol{\mu}_j, \boldsymbol{\Lambda}_j), \quad (4)$$

$$\varphi_{\mathbf{I}_j|\mathbf{I}_{j-1}} = \mathcal{N}_{\mathbf{I}_j}(\mathbf{A}\mathbf{I}_{j-1}, \mathbf{Q}), \quad (5)$$

where $\mathcal{N}_{\mathbf{D}_j}(f_j(\mathbf{I}_j), \mathbf{R}_j)$ refers to the multivariate normal distribution of \mathbf{D}_j with mean $f_j(\mathbf{I}_j)$ and the covariance \mathbf{R}_j . $\boldsymbol{\Lambda}_j = \text{diag}[\boldsymbol{\mu}_j]^T \boldsymbol{\Lambda}' \text{diag}[\boldsymbol{\mu}_j]$, where $\boldsymbol{\Lambda}'$ is defined in terms of the priors; see Equation (13) of [39] for more details. $\boldsymbol{\mu}_j$ is the mean of a multivariate Gaussian distribution and $\boldsymbol{\Lambda}$ describes the covariance, setting the spatial regularization. $f_j(\mathbf{I}_j)$ describes the functional relationship between the source image \mathbf{I}_j and the contemporaneous observed data \mathbf{D}_j . The global time evolution of the source between timestamps $j-1$ and j is described by the evolution matrix \mathbf{A} , so that $\mathbf{I}_j \approx \mathbf{A}\mathbf{I}_{j-1}$, and any additional perturbations of the source are constrained by the covariance matrix \mathbf{Q} . The process hence reduces to static imaging for ($\mathbf{A} = \mathbf{1}, \mathbf{Q} = \mathbf{0}$). The joint probability distribution is then given by:

$$p(\mathbf{M}, \mathbf{D}; \mathbf{A}) \propto \prod_{j=1}^{N_t} \varphi_{\mathbf{D}_j|\mathbf{I}_j} \prod_{j=1}^{N_t} \varphi_{\mathbf{I}_j} \prod_{j=2}^{N_t} \varphi_{\mathbf{I}_j|\mathbf{I}_{j-1}}. \quad (6)$$

In StarWarps, we jointly solve for the image reconstructions as well as \mathbf{A} . First, we learn \mathbf{A} using the Expectation-Maximization (EM) algorithm and then reconstruct the images with that \mathbf{A} .

For the analysis challenge image reconstruction with StarWarps, we used the visibility amplitudes, log closure amplitudes, and the bispectrum (triple amplitudes and closure phases) as our data products, with 2% added systematic noise. We used the EHT 2017 image of Sgr A* [11] blurred with a 25 μ as Gaussian kernel as a prior; see also [40].

2.2.3. Resolve

The algorithm `resolve`² approaches the imaging task for the (ng)EHT from a probabilistic, Bayesian perspective. It is based on Bayes' theorem:

$$\mathcal{P}(\mathbf{M}|\mathbf{D}) = \frac{\mathcal{P}(\mathbf{D}|\mathbf{M})\mathcal{P}(\mathbf{M})}{\mathcal{P}(\mathbf{D})}, \quad (7)$$

where \mathbf{D} refers to the measured data, and \mathbf{M} denotes the time-varying sky brightness distribution. The quantity $\mathcal{P}(\mathbf{M}|\mathbf{D})$ is called posterior probability density and contains all information on \mathbf{M} after taking the information from the data \mathbf{D} into account. In the case of the dynamic Sgr A* model, we consider \mathbf{M} to be a discretized quantity with spatial dimensions 200×200 and a temporal axis of length 720—in total, $2.88 \cdot 10^7$ degrees of freedom. Therefore, the posterior can be considered to be a function $\mathbb{R}^{28,800,000} \rightarrow \mathbb{R}^{>0}$.

The prior probability density $\mathcal{P}(\mathbf{M})$ represents our knowledge of the source before the data are considered. Since the sky brightness distribution represents a flux density, we can safely assume that its values are non-negative. Additionally, we know a priori that the emission is correlated in both the spatial and temporal directions. Thus, we assume generic homogeneous and isotropic spatial and temporal correlation structures whose specific form is learned from the data alongside \mathbf{M} . For more details on the prior, refer to Arras et al. [41,42].

The likelihood $\mathcal{P}(\mathbf{D}|\mathbf{M})$ encodes our knowledge of the measurement process. In general, the calibration pipeline of the (ng)EHT provides visibilities whose phases suffer from temporally uncorrelated station-based effects. The amplitudes of the visibilities are approximately correct and only subject to small time-correlated station-based effects. Therefore, we use closure phases and self-calibrated (non-closure) amplitudes in the likelihood.

After combining prior and likelihood, our best guess for the time-variable behavior of the source is given by the expectation value of the posterior: $\int \mathbf{M}\mathcal{P}(\mathbf{M}|\mathbf{D}) d\mathbf{M}$. Since evaluating such high-dimensional integrals directly is virtually impossible, we use Metric Gaussian Variational Inference [43] that provides approximate solutions to Bayes' theorem efficiently. As a result, we obtain a collection of approximate posterior samples that can be averaged to obtain an approximate posterior mean. Additionally, the variability of the approximate posterior samples represents the uncertainty of the computed solution. This uncertainty could be propagated to downstream analyses of the time-variable reconstruction of the source.

For the ngEHT analysis challenges, we adopted the implementation of `resolve` of Arras et al. [42] where `resolve` has been applied to the 2017 EHT observation of M87*. For evaluating the interferometry measurement equation (i.e., the non-equidistant Fourier transform), we use the implementation presented in Arras et al. [44]. Variations of `resolve` have been verified and tested against standard methods in various contexts [41,42]. To enable comparisons to results of algorithms that do not quantify uncertainties, we will depict only the posterior mean of the sky brightness distribution s in the following.

2.2.4. DoG-HiT

DoG-HiT is a multi-scale RML imaging algorithm [45]. DoG-HiT models the image by a set of wavelets (a dictionary Γ) constructed by the difference of the Gaussian method: $\mathbf{I} = \Gamma \mathcal{J}$ [45,46]. With DoG-HiT, we aim to recover the array of wavelet coefficients \mathcal{J} that represents the true sky brightness distribution best. The wavelets define filters in the

Fourier domain that are ring-like. Hence, every wavelet compresses the spatial information from a specific band of baselines in the uv -coverage. For DoG-HiT, the scales are fitted to the uv -coverage, thus giving rise to wavelets most sensitive to gaps in the uv -coverage and wavelets most sensitive to Fourier coefficients sampled by baselines. In the spirit of compressed sensing, DoG-HiT utilizes a sparsity promoting penalization by a l_0 penalty term on the wavelet coefficients. In detail, we solve the following optimization problem consisting of data fidelity terms for the closure quantities (χ_{cp}^2 and χ_{camp}^2), the l_0 penalty term and a total flux constraint by an updated forward-backward splitting approach [45]:

$$\hat{\mathcal{J}} \in \operatorname{argmin}_{\mathcal{J}} \left[\chi_{\text{cp}}^2(\Gamma \mathcal{J}) + \chi_{\text{camp}}^2(\Gamma \mathcal{J}) + \alpha \cdot \|\mathcal{J}\|_{l_0} + R_{\text{flux}}(\mathcal{J}, f) \right], \quad (8)$$

where R_{flux} is a total flux indicator function with flux f , and Γ denotes the wavelet dictionary. DoG-HiT is a data-driven, automatic imaging pipeline that depends on only one hyper-parameter (the relative weighting of the penalization term α). It has been demonstrated to produce high-quality, super-resolved reconstructions for static sources in a relatively short time with minimal manual interaction and without the need for extensive parameter surveys. e.g., compare the Challenge 1 reconstructions with DoG-HiT in Section 6 of [45].

The dynamic reconstructions are based rather straightforwardly upon the success of this static imaging. We utilize the automatic static imaging pipeline to construct a mean image from the full length of the observation without taking the dynamics of the source into account. DoG-HiT computes a set of statistically significant wavelet coefficients from the mean image as a byproduct (the multiresolution support). Then the mean image (with a relatively bad fit to the data due to not respecting the source dynamics) is subtracted from the self-calibrated visibilities, and the observation is cut into frames of six minutes. The residuals are minimized frame by frame with StarWarps with implicit dynamic variability imposed by StarWarps. A small StarWarps internal regularization parameter is used, but, in contrast to StarWarps, the reconstruction is performed in a multiscale constrained minimization framework (multiresolution support constraint), i.e., only the wavelet coefficients classified as significant during the static image reconstruction are allowed to vary. This introduces a correlation between frames and consistency to the mean static image.

DoG-HiT is still under development and is currently extended to dynamic, polarimetric reconstructions [47] with promising first results on synthetic data (see upcoming Challenge 3 reconstructions). A finer set of directional-dependent wavelet functions [46] allows for dynamic reconstructions in a constrained minimization reconstruction on frames independently, thus replacing StarWarps during the current DoG-HiT dynamic imaging pipeline and relying on a completely unsupervised, automatic wavelet approach only. On the one hand, such an unsupervised, automatic imaging procedure is desired as it reduces the human bias in the reconstruction; on the other hand, driving by an astronomer could be crucial to address data issues, in particular for challenging data sets such as what will be produced by the ngEHT.

3. Submission Evaluation Metrics

Submitted reconstructions were evaluated with several quantitative quality metrics. These metrics, which all probe different aspects of what makes a high-quality reconstruction, are summarized below.

3.1. Data Fit Quality

The goodness-of-fit of the submitted reconstructions to the provided synthetic data was quantified by computing the reconstruction visibilities using the synthetic data uv -coverage and then calculating the χ^2 -metric on closure quantities, χ_{cphase}^2 and χ_{lcamp}^2 . These quantities are the closure phases, which are the sum of visibility phases measured simultaneously on a closed triangle of baselines, and the (log) closure amplitudes, respectively, which are

ratios of visibility amplitudes on a baseline quadrangle; see, e.g., [48]. Closure quantities are robust against station-based calibration errors.

3.2. Ground Truth Image Similarity

Apart from the goodness-of-fit to the synthetic data, another important quality metric is the similarity of the reconstruction to the ground truth source model. We quantify this similarity using the normalized cross-correlation. The normalized cross-correlation between two images X and Y is

$$\rho_{\text{NX}} = \frac{1}{N} \sum_{i=1}^N \frac{(X_i - \langle X \rangle)(Y_i - \langle Y \rangle)}{\sigma_X \sigma_Y}. \quad (9)$$

Here, N is the number of pixels in the images, X_i and Y_i are the pixel values of images X and Y , respectively, $\langle \dots \rangle$ denotes an average, and σ_X and σ_Y are the standard deviations of the pixel values of images X and Y , respectively. The value of ρ_{NX} will be equal to 1 for identical images (maximal correlation), 0 for completely uncorrelated images, and -1 for perfectly anticorrelated images. Using the implementation in `eht-imaging`, the two images are regridded to contain the same number of pixels with equal pixel size and aligned to maximize ρ_{NX} . For M87 reconstructions, we are often most interested in the arrays' ability to reconstruct the large-scale and low-surface brightness jet emission. Therefore, we also compute ρ_{NX} on the log pixel values ($\rho_{\text{NX},\log}$). In order to suppress the influence of low-surface brightness image noise, which may appear at a certain flux level depending on the particularities of the reconstruction algorithm, we limit the dynamic range of the ground truth and reconstructed images to 10^4 in this case.

3.3. Effective Resolution

ρ_{NX} also provides a way to compute the effective angular resolution obtained by the reconstructed image. Following Event Horizon Telescope Collaboration et al. [6], we blur the ground truth model images with a circular Gaussian with varying FWHM and calculate $\rho_{\text{NX},\text{FWHM}}$ with respect to the ground truth model images using Equation (9). For a submitted reconstruction with a particular $\rho_{\text{NX},\text{rec}}$ with respect to the ground truth, the effective resolution θ_{eff} is then the (interpolated) FWHM for which $\rho_{\text{NX},\text{rec}}$ is equal to $\rho_{\text{NX},\text{FWHM}}$.

3.4. Dynamic Range

Dynamic range is usually defined as the ratio between the brightest and dimmest pixel value in an image and has been frequently used in radio astronomy to assess the ability of an array to reconstruct low-surface brightness features. For images reconstructed with CLEAN algorithms, the dynamic range can be naturally calculated as the ratio between the brightest CLEAN component and the noise floor (Section 2.1.1). However, for images reconstructed with other algorithms (e.g., RML-based approaches), formally defining a dynamic range metric that works universally and reflects our intuitive sense of dynamic range is non-trivial. This difficulty has two main causes. First, not all imaging methods naturally produce a noise floor, such as CLEAN and have many (near)-zero pixel values. Second, many imaging algorithms produce spurious structures due to, e.g., sparse uv -coverage, so that the lowest reconstructed pixel brightness cannot be used to robustly define a dynamic range metric.

To evaluate the challenge reconstructions, we use a dynamic range proxy following Bustamante et al. [49]. This metric considers the ratios between the brightest pixel of the ground truth image $\mathbf{I}_{\text{groundtruth}}$ and the absolute pixel residuals of the reconstructed image $\mathbf{I}_{\text{reconstructed}}$ with respect to the ground truth,

$$\mathcal{D} = \frac{\max(\mathbf{I}_{\text{groundtruth}} * \mathcal{G}_{\theta_{\text{eff}}}^{2\text{D}})}{|\mathbf{I}_{\text{reconstructed}} - \mathbf{I}_{\text{groundtruth}} * \mathcal{G}_{\theta_{\text{eff}}}^{2\text{D}}|}. \quad (10)$$

Here, $*$ denotes convolution, $|\dots|$ indicates that we take the absolute values, and $\mathcal{G}_{\theta_{\text{eff}}}^{2\text{D}}$ represents a two-dimensional circular Gaussian with an FWHM equal to the effective resolution θ_{eff} of the reconstructed image. From \mathcal{D} , which has the form of an image, we can calculate a dynamic range proxy by selecting the q th quantile of the pixel values:

$$\mathcal{D}_q = \text{quantile}(\mathcal{D}, q). \quad (11)$$

By using the residuals, this metric penalizes spurious structures in the reconstructed image and does not rely on a noise floor being calculated as part of the imaging process. A disadvantage is that it requires the ground truth image and hence cannot be applied to real data. Setting q too low will make the metric dominated by outliers in the residual image, while setting it too high will not penalize high residuals strongly enough. We set $q = 0.1$ for our dynamic range proxy $\mathcal{D}_{0.1}$. Because of its sensitivity to q and the chosen blurring kernel, we emphasize that this metric should not be regarded as giving *the* dynamic range of the image but rather as a dynamic range proxy that can be used to compare different reconstructions of the same source model.

4. Challenge 1

4.1. Rationale and Charge

The primary objectives of the first challenge were to set up a framework for the generation of synthetic ngEHT data based on theoretical source models, to conduct the organized submission and cross-comparison of reconstruction results from multiple people, and to obtain a first idea of the benefits and challenges of ngEHT datasets as compared to the current EHT. The model and data properties were therefore kept relatively simple. Participants were asked to submit image reconstructions for each provided synthetic dataset. The challenge was not blind, i.e., the participants had access to the input source models and synthetic data generation script. The challenge was launched on 18 June 2021, and the submission deadline was 16 July 2021. It was advertised to the ngEHT simulations group. All information is available on the challenge website: <https://challenge.ngeht.org/challenge1/> (accessed on 19 December 2022).

4.2. Source Models

For Challenge 1, we used two static, unpolarized models of M87 and Sgr A*, respectively. Both models are displayed in Figure 1 and described below. More detailed descriptions and comparisons of the source models used for Challenge 1 and 2 can be found in Chatterjee et al. [21].

4.2.1. M87

The Challenge 1 M87 model is a magnetically arrested disk (MAD) general relativistic magnetohydrodynamics (GRMHD) frame from a rapid spinning black hole $a_* = 0.94$ with electron thermodynamics from reconnection heating; see [50] for details. The GRMHD simulation was performed with the BHAC code [51] using three levels of adaptive mesh refinement (AMR) in logarithm Kerr-Schild coordinates. The numerical grid covers $384 \times 4192 \times 192$ cells in radial, azimuthal and theta directions and extends up to 2500 gravitational radii (GM/c^2 , where G is Newton's gravitational constant, M is the black hole mass, and c is the speed of light) in the radial direction. The mass accretion rate and MAD parameter (see [52]) were monitored, and after obtaining a steady state, we performed the general relativistic radiative transfer (GRRT) calculations with the radiative transfer code BHOSS [53,54].

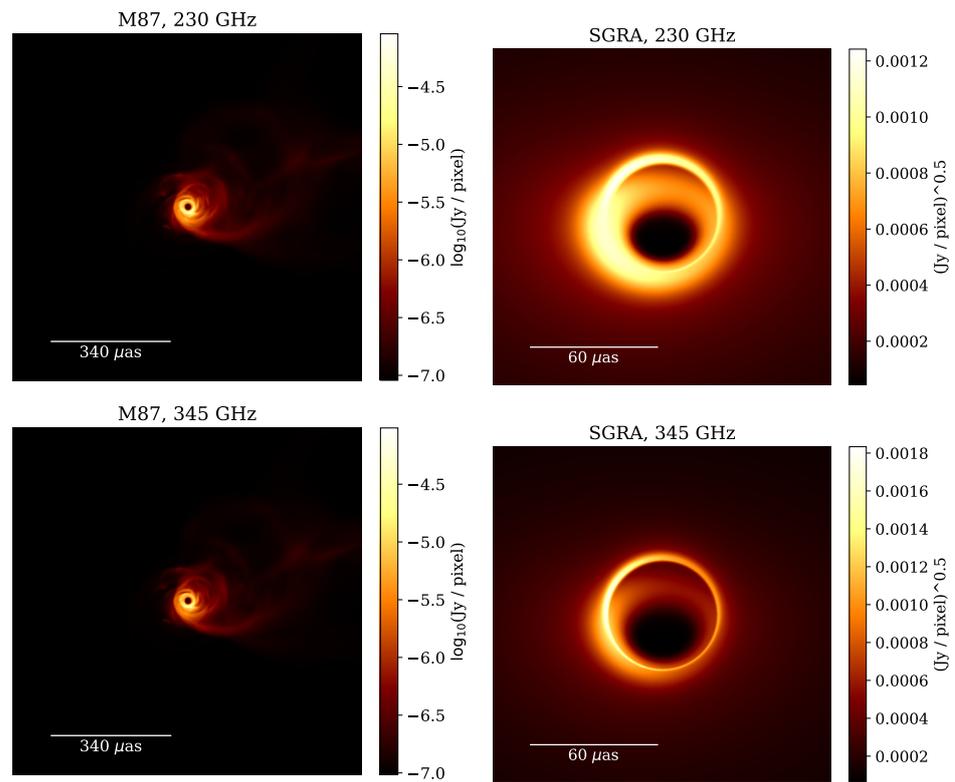


Figure 1. Source models used for Challenge 1.

During the radiative transport, we included non-thermal particles via the kappa electron distribution (see [55]) in the jet sheath while excluding the highly magnetized spine by using a cut in the magnetization at a value of 1 (typically referred to as a sigma cut). The power-law slope of the kappa distribution was set by a particle in cell (PIC)-motivated sub-grid model depending on the local magnetization and plasma-beta following Ball et al. [56]. In addition, we included a fraction of the magnetic energy density to accelerate the non-thermal particles (see [57,58]). In the jet wind and disk region, we used a thermal electron distribution, where the electron temperature is directly obtained from the GRMHD simulation. In order to guarantee capturing a small scale structure on the horizon scale and at the same time the large-scale jet structure, we used a field of view (FOV) of 1 mas using a resolution of 4096×4096 pixels. Since the GRMHD simulations are scale-free, we normalized our GRRT simulations by setting the mass ($6.5 \times 10^9 M_{\odot}$) and distance (16.9 Mpc) of the black hole in M87 and iterated the mass accretion rate until a compact flux density of 0.8 Jy at 230 GHz was obtained.

4.2.2. Sgr A*

The Challenge 1 Sgr A* model is a stationary semi-analytic radiatively inefficient accretion flow (RIAF) model, e.g., [59]. This model can be used to test the capabilities of next-generation arrays in precision modeling of black hole parameters. High resolution is needed to capture the unique signature from a subring structure. This model does not capture any variability due to turbulence in the system. The basic model has $a_* = 0$ (Schwarzschild) at an inclination of $i = 130$ deg and includes non-thermal particles. The model includes disk height following [60], sub-Keplerian flow properties ($\kappa = 0.5$, $\alpha = 0.5$), following the notation of Tiede et al. [61]—e.g., Equations (10) and (11), and fitted to the observed data of [62–65]. Images were ray-traced at 230 and 345 GHz with 4096×4096 pixels and a FOV of $128 GM/c^2$, using a distance of 8.178 kpc and mass of $4.14 M_{\odot}$ [66]. Finally, the 230 and 345 GHz images were scattered with the same realization of the Johnson et al. [35] interstellar scattering model before generating the synthetic data.

4.3. Synthetic Data

4.3.1. Station Locations

Two arrays were used to generate the Challenge 1 synthetic data. EHT2022 consists of the 11 stations that participated in the 2022 EHT observations. In the ngEHT reference array 1 (ngEHT1), 10 stations are added to this array. The station locations were chosen based on a uv -coverage analysis (A. Raymond, priv. comm.), investigating which combination of sites from Raymond et al. [67] provided optimal uv -coverage by performing a Monte Carlo simulation involving telescope dropouts due to weather conditions. The LMT, SPT, and KP were not included in the 345 GHz observations with EHT2022. The station locations are shown in Figure 2.

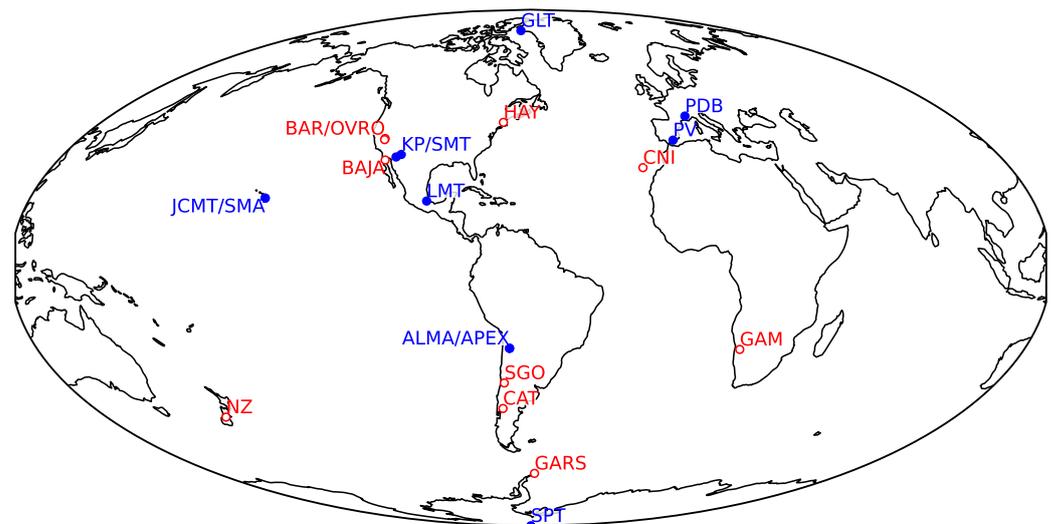


Figure 2. Station locations for Challenge 1 and 2. Stations in blue form the EHT2022 array, and stations in red are added to EHT2022 to form a reference array ngEHT1. The new station locations are near the National Astronomical Observatory in Baja California, Mexico (BAJA); Barcroft Field Station in California, USA (BAR); Cerro Catedral in R'io Negro in Argentina (CAT), La Palma, part of the Canary Islands, Spain (CNI); the Gamsberg in Namibia (GAM), the German Antarctic Receiving Station O'Higgins in Antarctica (GARS); Haystack Observatory in Westford, MA, USA; Canterbury, New Zealand (NZ); Owens Valley Radio Observatory in California, USA (OVRO); and Santiago, Chile (SGO). See Raymond et al. [67] for details, e.g., weather statistics for each site.

4.3.2. Data Properties

A 24-h observing track was simulated for each array, source, and frequency, resulting in eight separate datasets. Each track consists of 10-min scans interleaved with 10-min gaps and is identical for each dataset. A single frequency channel with a time resolution of 10 s was provided. Thermal noise expected from the receiver and atmospheric opacity were added to the complex visibilities calculated using `eht-imaging` [29,30]. The following assumptions were made for all sites:

- Receiver temperature: 60 K for 230 GHz; 100 K for 345 GHz
- Aperture efficiency: 0.68 for 230 GHz; 0.42 for 345 GHz
- Bandwidth: 8 GHz
- Quantization efficiency: 0.88
- Dish diameter: 6 m for new sites, actual diameter for existing sites
- Opacity: median values in April as extracted from the MERRA-2 data by Raymond et al. [67], at 30-degree elevation. The opacities were set constant throughout and across the different datasets but are frequency-dependent.

Visibility phases were scrambled but stabilized across scans. No further systematic errors were added. After data generation, data points with a signal-to-noise ratio less than 1 were flagged. Figure 3 shows the resulting uv -coverage for all datasets.

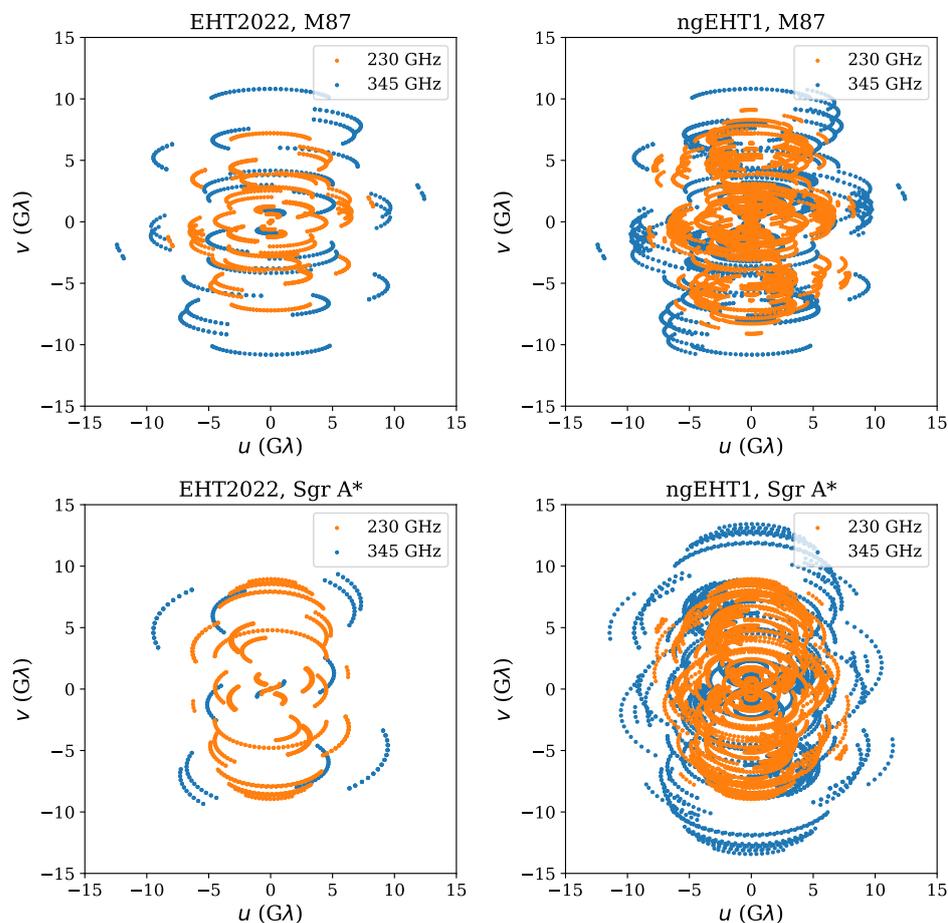


Figure 3. uv -coverage for the Challenge 1 datasets.

4.4. Results

Challenge 1 image reconstructions were submitted by three individual submitters and one team, using CLEAN, SMILI, and eht-imaging. For M87, one submitter performed a multi-frequency image reconstruction with eht-imaging, using information from the 230 GHz data to reconstruct the image at 345 GHz, and vice versa. All submitted reconstructions are displayed in Figures 4 and 5 for M87 and Sgr A*, respectively. Reconstruction quality metrics (Section 3) are shown in Table 1. While it should be kept in mind that the synthetic data was idealized in certain aspects (no systematic amplitude errors, and a static Sgr A* source model), these results show some interesting trends.

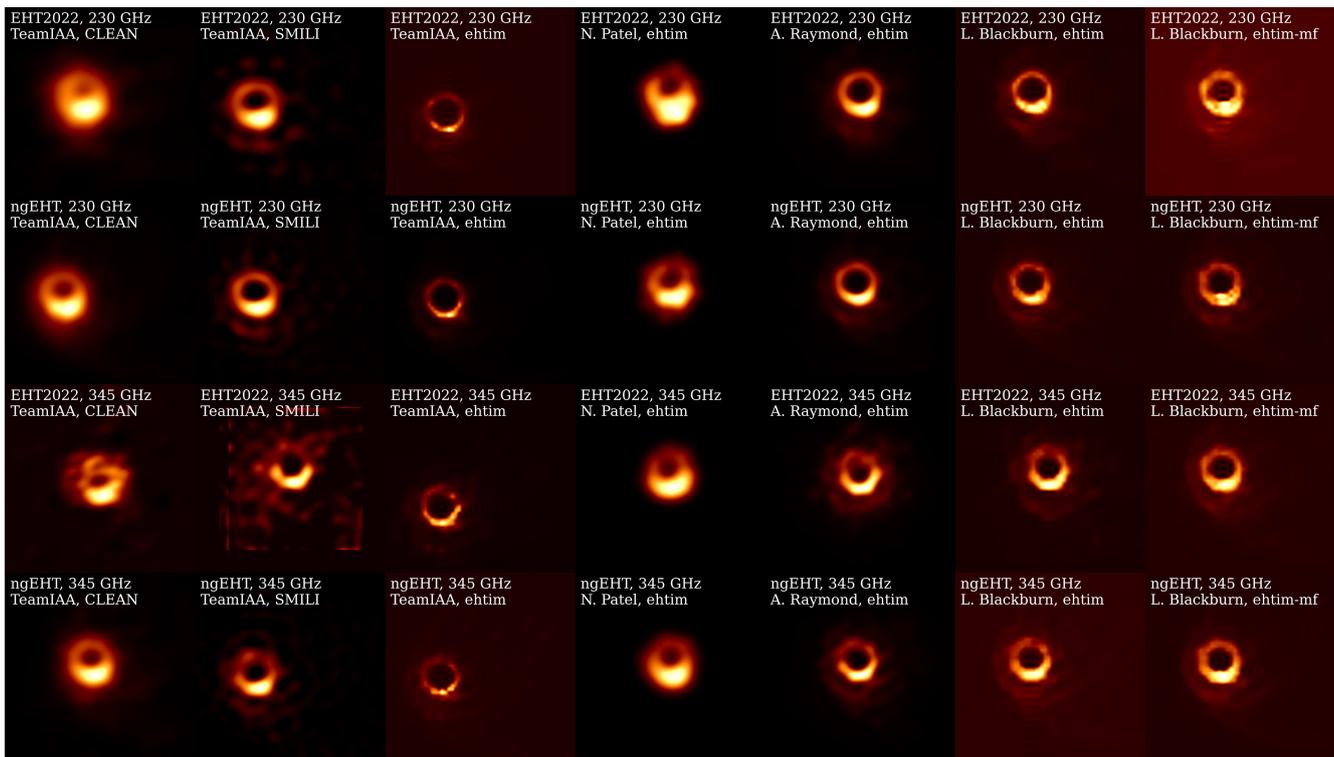
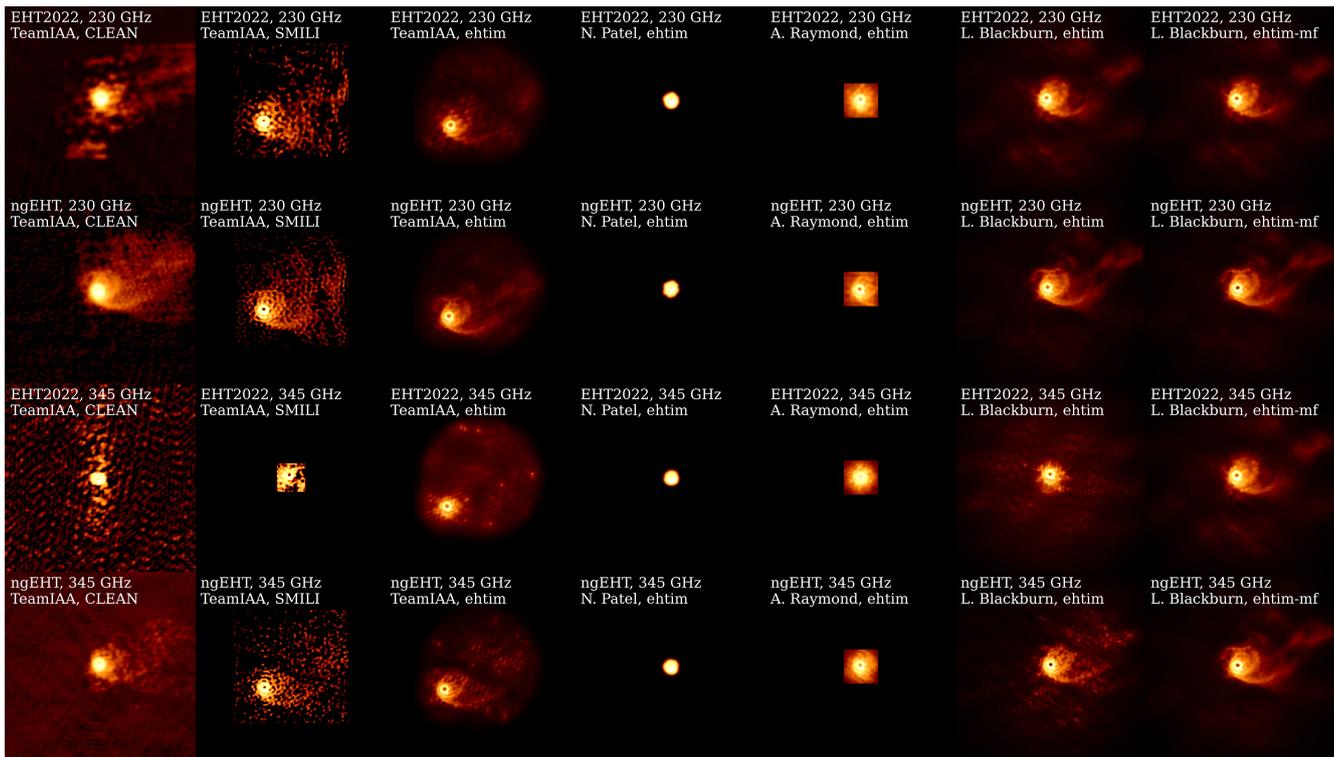


Figure 4. M87 reconstructions submitted for Challenge 1. Images are shown on a log scale with a 1 mas field of view in the top set of panels. The same images are shown on a linear scale with a 200 μ as field of view in the bottom set of panels.

All M87 reconstructions recover the black hole shadow, whereas the jet features are only recovered by some. The low surface brightness structure in the M87 jet is already visible in some EHT2022 reconstructions. The jet reconstructions improve significantly with ngEHT1 coverage, as attested by both visual inspection of the images and the quality

metrics. The *eht-imaging* submissions perform best, although the ρ_{NX} ranking of individual submissions changes depending on whether the linear or log-scale images are used for the comparison. $\rho_{\text{NX},\log}$ and $\mathcal{D}_{0.1}$ are more sensitive to the reconstruction of the extended jet structure and generally show a clearer improvement of ngEHT1 versus EHT2022 reconstructions. The CLEAN and SMILI reconstructions show poorer jet structure recovery than *eht-imaging*, although they may potentially be improved by adapting the specific scripts used for these reconstructions. The reconstruction quality is generally better for 230 GHz than for 345 GHz due to the higher flux and better *uv*-coverage at 230 GHz. Reconstructions with relatively high χ^2 values often have relatively low ρ_{NX} values. The multi-frequency analysis (*eht-im-mf*) is an exception with relatively good reconstruction quality (especially as shown by $\rho_{\text{NX},\log}$ and $\mathcal{D}_{0.1}$) for relatively high χ^2 values. The multi-frequency analysis is especially useful for reconstructing the jet features at 345 GHz, as these are reconstructed significantly more poorly with other methods.

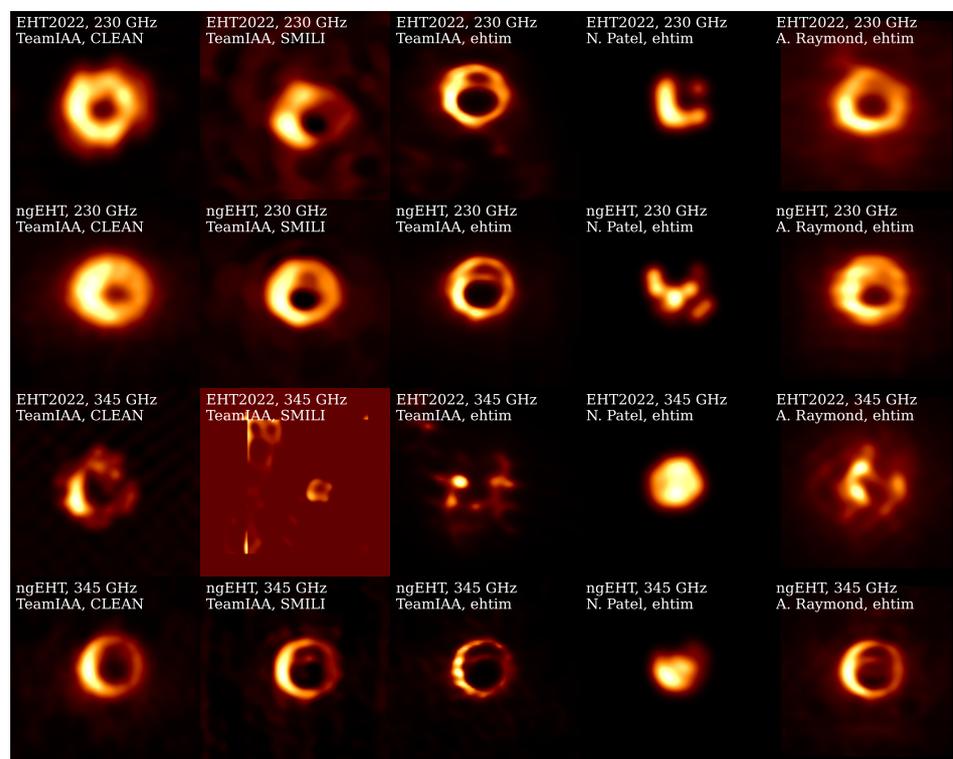


Figure 5. Sgr A* reconstructions submitted for Challenge 1. Images are shown on a linear scale with a $200 \mu\text{s}$ field of view.

For the Sgr A* model, the black hole shadow is recovered by all arrays except the EHT2022 array at 345 GHz: the *uv*-coverage is too sparse for a high-fidelity image reconstruction in this case (Kitt Peak, the LMT, and the SPT cannot observe at 345 GHz yet). ngEHT reconstructions at 345 GHz are significantly sharper than EHT2022 and ngEHT reconstructions at 230 GHz. ngEHT reconstructions at 230 GHz are generally less noisy than EHT2022 reconstructions at the same frequency, but for Sgr A*, the real value of ngEHT coverage will be in dynamical reconstructions (Section 5). The χ^2_{icamp} are generally high for Sgr A* reconstructions, which is likely due to the comparison with the provided synthetic data, which includes interstellar scattering, while submitters may have deblurred the visibility amplitudes in the reconstruction process.

Table 1. Reconstruction quality metrics for Challenge 1.

Source	Array	ν (GHz)	Submitter	Method	χ_{cp}^2	χ_{lcamp}^2	ρ_{NX}	$\rho_{NX,log}$	θ_{eff}	$\mathcal{D}_{0.1}$
M87	EHT2022	230	L. Blackburn	ehtim	1.1	1.01	0.93	0.87	5.4	856
M87	EHT2022	230	L. Blackburn	ehtim-mf	5.17	4.36	0.88	0.9	9.8	797
M87	EHT2022	230	N. Patel	ehtim	3.66	1159.56	0.77	0.52	21.2	418
M87	EHT2022	230	TeamIAA	SMILI	0.99	1.06	0.83	0.79	14.6	409
M87	EHT2022	230	TeamIAA	CLEAN	2.94	879.77	0.8	0.8	17.7	529
M87	EHT2022	230	TeamIAA	ehtim	1.79	1.03	0.89	0.91	8.9	564
M87	EHT2022	230	A. Raymond	ehtim	2.28	1.77	0.9	0.72	8.0	291
M87	ngEHT	230	L. Blackburn	ehtim-mf	2.62	1.43	0.89	0.96	8.9	1681
M87	ngEHT	230	L. Blackburn	ehtim	1.07	1.01	0.93	0.95	5.4	1604
M87	ngEHT	230	N. Patel	ehtim	3.5	89.74	0.83	0.52	14.6	640
M87	ngEHT	230	TeamIAA	SMILI	1.01	1.03	0.87	0.85	10.8	708
M87	ngEHT	230	TeamIAA	CLEAN	1.32	138.45	0.84	0.91	13.6	1828
M87	ngEHT	230	TeamIAA	ehtim	1.08	1.01	0.91	0.97	7.1	1727
M87	ngEHT	230	A. Raymond	ehtim	1.65	2.14	0.92	0.73	6.2	532
M87	EHT2022	345	L. Blackburn	ehtim-mf	2.36	1.06	0.91	0.87	5.7	1403
M87	EHT2022	345	L. Blackburn	ehtim	1.19	0.62	0.91	0.72	5.7	984
M87	EHT2022	345	N. Patel	ehtim	1.2	7.29	0.79	0.53	16.7	734
M87	EHT2022	345	TeamIAA	SMILI	1.19	0.62	0.79	0.66	16.7	645
M87	EHT2022	345	TeamIAA	ehtim	1.22	0.62	0.88	0.81	8.2	700
M87	EHT2022	345	TeamIAA	CLEAN	3.34	2.77	0.82	0.38	13.7	320
M87	EHT2022	345	A. Raymond	ehtim	1.19	0.62	0.88	0.74	8.2	546
M87	ngEHT	345	L. Blackburn	ehtim	1.15	0.97	0.92	0.89	4.9	1570
M87	ngEHT	345	L. Blackburn	ehtim-mf	1.25	1.13	0.91	0.94	5.7	2244
M87	ngEHT	345	N. Patel	ehtim	1.2	9.99	0.79	0.54	16.7	853
M87	ngEHT	345	TeamIAA	CLEAN	1.31	4.39	0.84	0.75	11.8	651
M87	ngEHT	345	TeamIAA	SMILI	1.16	1.0	0.85	0.71	10.9	766
M87	ngEHT	345	TeamIAA	CLEAN	1.31	4.39	0.84	0.75	11.8	651
M87	ngEHT	345	TeamIAA	ehtim	1.16	0.98	0.9	0.92	6.5	1638
M87	ngEHT	345	A. Raymond	ehtim	1.17	1.0	0.91	0.75	5.7	782
Sgr A*	EHT2022	230	N. Patel	ehtim	6.08	347.88	0.8	-	45.5	-
Sgr A*	EHT2022	230	TeamIAA	ehtim	1.11	33.13	0.95	-	14.3	-
Sgr A*	EHT2022	230	TeamIAA	CLEAN	140.97	130.2	0.9	-	23.4	-
Sgr A*	EHT2022	230	TeamIAA	SMILI	1.47	23.19	0.85	-	32.6	-
Sgr A*	EHT2022	230	A. Raymond	ehtim	3.02	8.27	0.89	-	25.2	-
Sgr A*	ngEHT	230	N. Patel	ehtim	20.23	122.65	0.65	-	100.0	-
Sgr A*	ngEHT	230	TeamIAA	SMILI	1.4	8.81	0.95	-	14.3	-
Sgr A*	ngEHT	230	TeamIAA	CLEAN	2.3	23.3	0.9	-	23.4	-
Sgr A*	ngEHT	230	TeamIAA	ehtim	1.06	10.61	0.97	-	10.1	-
Sgr A*	ngEHT	230	A. Raymond	ehtim	1.14	1.87	0.93	-	18.1	-
Sgr A*	EHT2022	345	N. Patel	ehtim	1.03	20.32	0.64	-	61.9	-
Sgr A*	EHT2022	345	TeamIAA	CLEAN	71.44	66.33	0.79	-	24.5	-
Sgr A*	EHT2022	345	TeamIAA	ehtim	1.03	1.95	0.65	-	57.8	-
Sgr A*	EHT2022	345	TeamIAA	SMILI	1.63	1.7	0.34	-	100.0	-
Sgr A*	EHT2022	345	A. Raymond	ehtim	1.03	0.85	0.78	-	26.0	-
Sgr A*	ngEHT	345	N. Patel	ehtim	2.18	15.58	0.64	-	61.9	-
Sgr A*	ngEHT	345	TeamIAA	ehtim	1.14	1.19	0.93	-	7.5	-
Sgr A*	ngEHT	345	TeamIAA	CLEAN	2.24	4.48	0.87	-	14.0	-
Sgr A*	ngEHT	345	TeamIAA	SMILI	1.17	1.23	0.89	-	11.7	-
Sgr A*	ngEHT	345	A. Raymond	ehtim	1.14	1.15	0.9	-	10.6	-

5. Challenge 2

5.1. Rationale and Charge

With the challenge infrastructure and initial participant imaging efforts set up in the first challenge, the second challenge was more realistic and science-oriented and different from the first challenge in two aspects.

First, the ground truth source models were dynamic instead of static. For Sgr A*, with variability on timescales of \sim minutes, the charge to the participants was to reconstruct a movie of the source evolving across a single day of observations. We used two source models to test reconstruction capabilities for different variability properties: a GRMHD model with turbulent variability, and a shearing hotspot in a RIAF disk, exhibiting a more coherent variable structure. Hints of such coherent variability of Sgr A* at 230 GHz consistent with an orbiting hotspot have been observed by Wielgus et al. [68]. For M87*, with variability on timescales of \sim days, we used a bright jet GRMHD model like in the first challenge but evolved it over a period of five months, simulating a full-day observation every week. The charge was to reconstruct a movie of the large-scale and low-surface brightness jet emission, connecting it to the dynamics near the black hole shadow.

The second aspect in which this challenge differed from the previous one is that the synthetic observations included significantly more realistic effects. Contrary to the idealized data generated for Challenge 1, which only includes thermal noise, the Challenge 2 data sets have been generated under the assumption of realistic observing conditions and include data systematics originating from weather, instrumental, and calibration effects (see Section 5.3). The results of this challenge thus reflect what would actually be seen by an array built with the described specifics, using current reconstruction algorithms.

Challenge 2 was launched on 25 October 2021. It was advertised more broadly than the first challenge to the full ngEHT community. The first submission comparisons were completed in January 2022, and due to the complexity of the datasets and ongoing development of reconstruction algorithms, reconstructions were submitted until August 2022.

5.2. Source Models

This section describes the dynamical source models used for Challenge 2. See Chatterjee et al. [21] for more detailed model descriptions and comparisons.

5.2.1. M87

The Challenge 2 M87 model is a GRMHD movie with 20 frames that are spaced $20GM/c^3$ (~ 1 week) apart. The pixel resolution is 2048×2048 , with a field of view of 1 mas. The images were ray-traced from a H-AMR [69] simulation (MAD, spin 0.94) using ipole [70]. R_{high} was set to 160 and accelerated electron heating was included, setting $\kappa = 3.5$ [57]. We only use the Stokes I information from the model. The model is shown in Figure 6.

5.2.2. Sgr A*

For Challenge 2, two dynamic source models were used for Sgr A*: a GRMHD model exhibiting turbulent variability and a RIAF + shearing hotspot model with more ordered variability properties. Sample frames of the source models used for Challenge 2 are shown in Figure 7.

The GRMHD model is a MAD model with a spin of 0.5. The images were ray-traced in Stokes I with BHOSS [71], assuming thermal electrons. The 500 frames are spaced $10 GM/c^3$ (221 s) apart. The pixel resolution is 2048×2048 , with a field of view of $400 \mu\text{as}$.

The second Sgr A* model is a RIAF [72] plus shearing hotspot [61] semi-analytical model. The hotspot parameters are inspired by [73], and the black hole spin was set to 0.1. The pixel resolution is 313×313 , with a field of view of $315 \mu\text{as}$. The frames are spaced 30 s apart and form a 4-h movie of a hotspot shearing and falling in, which is repeated a few times over the course of the 24-h observation.

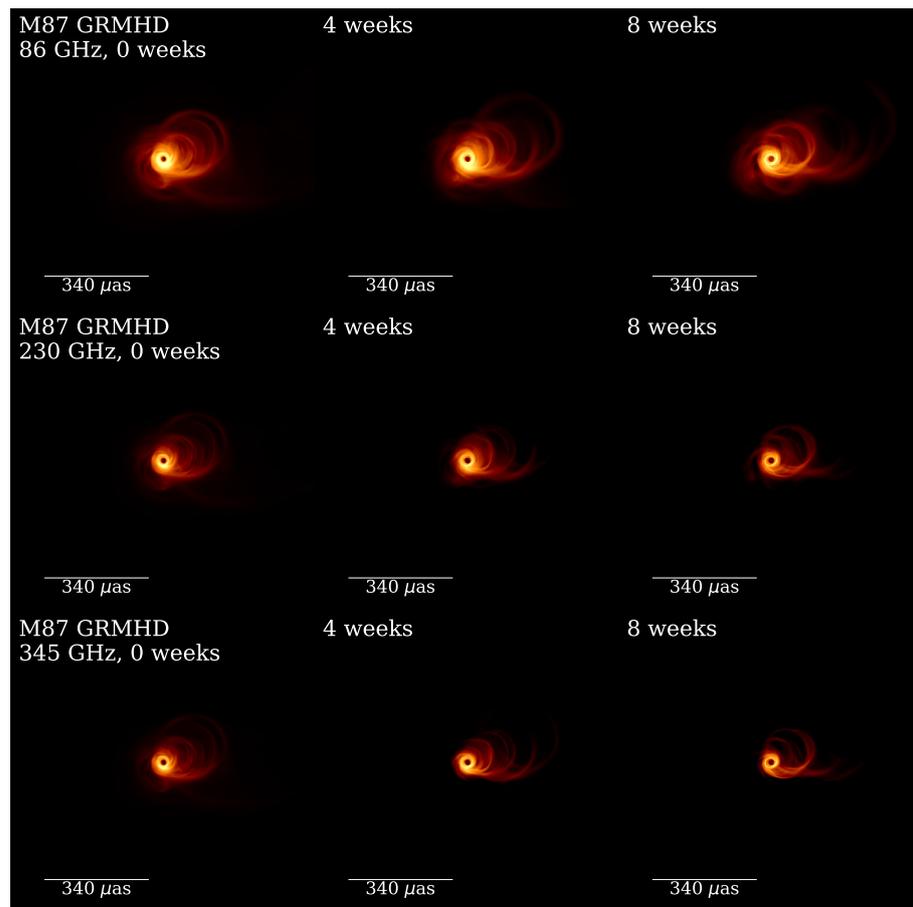


Figure 6. Ground truth M87 source models used for Challenge 2. For each frequency, three movie frames are shown. Images are shown on a log scale, which is normalized to the brightest pixel value across each set of three movie frames, with a dynamic range of $10^{3.5}$.

5.3. Synthetic Data

The synthetic data for Challenge 2 includes significantly more systematic effects than Challenge 1 (see also Section 5.1). In the SYMBA pipeline [74], atmospheric absorption, emission, delays and turbulence are simulated, and antenna pointing offsets are added to the simulated datasets with *MeqSilhouette* [75,76]. These are then calibrated by performing a fringe fit, a priori amplitude calibration with *rPICARD* [77], and network calibration with *eht-imaging*.

The station locations, dish sizes, aperture efficiencies, and receiver temperatures are identical to those used in Challenge 1 (Figure 2, Section 4.3). For each site, the input precipitable water vapor (PWV), ground temperature, and ground pressure were calculated from the Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2) from the NASA Goddard Earth Sciences Data and Information Services Center GES DISC, [78], processed with the *am* atmospheric model software [79] (see [74] for details). All weather quantities were based on median conditions on 1 April (2000–2020) as registered in the MERRA-2 climatological data. The atmospheric coherence time at 230 GHz was assumed to be 20 s for a PWV of 1 mm, 3 s for a PWV of 15 mm, interpolated linearly between these values for the different sites and scaled linearly with frequency. Pointing offsets were assumed to be stable across each 10-min scan and drawn randomly from a Gaussian distribution with an RMS of 2 arc-seconds.

5.4. Results

Seven submitters or teams provided dynamical reconstructions of the Challenge 2 datasets. Reconstruction quality metrics for all submissions are shown in Table 2.

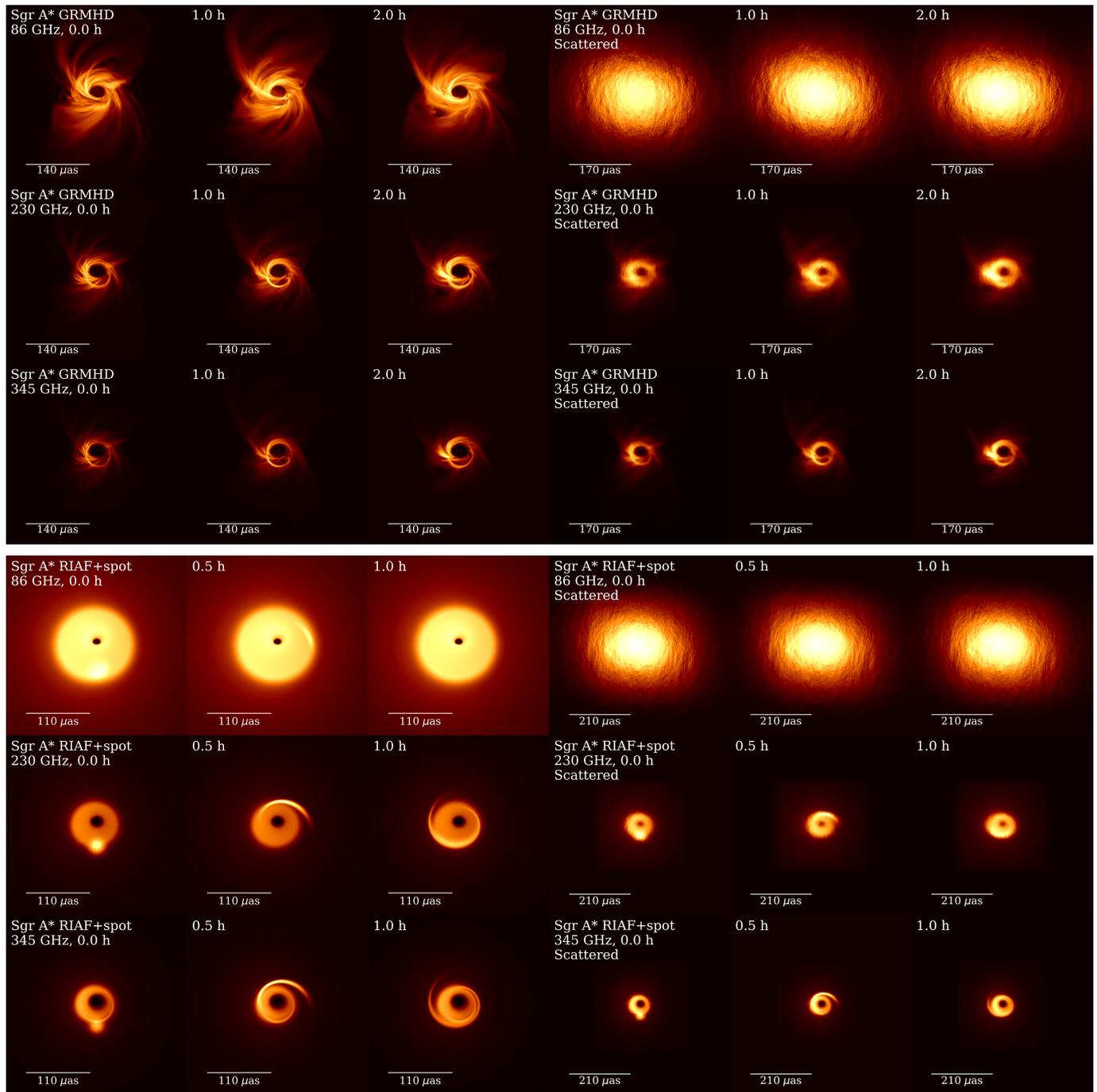


Figure 7. Ground truth Sgr A* source models used for Challenge 2. For each model and frequency (rows), three movie frames (columns) are shown, with interstellar scattering applied to the first frame in the rightmost column. Images are shown on a square root scale, which is normalized to the brightest pixel value across each set of three movie frames. The scattered movies were used as inputs for the Challenge 2 synthetic data generation.

5.4.1. M87 GRMHD

Six frames of the 86 and 230 GHz M87 reconstructions are shown in Figures 8 and 9, respectively. At 86 GHz, EHT2022 coverage allows reconstruction of the central component and overall shape of the extended jet emission, but all reconstructions contain spurious artifacts. With ngEHT1 coverage, these artifacts become far less severe or disappear completely, and the jet dynamics can be imaged as the jet features move outwards over the course of several weeks. The 230 GHz reconstructions provide significantly more detail, both in the

extended jet and in the visibility of the black hole shadow. The reconstructions again show a strong improvement of ngEHT1 compared to EHT2022, although the EHT2022 reconstructions from the `resolve` algorithm already show some jet dynamics with EHT2022 coverage. Figures 10 and 11 show spectral index and individual frequency image reconstructions, respectively, from `resolve` when solving for all frequencies simultaneously and imposing a prior on the spectral index map (`resolve-mf`) for the first movie frame only. This method leads to remarkably high-quality images for all frequencies and arrays, even showing the black hole's central brightness depression at 86 GHz. These results demonstrate that utilizing information from simultaneous multi-frequency observations can significantly boost the reconstruction quality; see also [17].

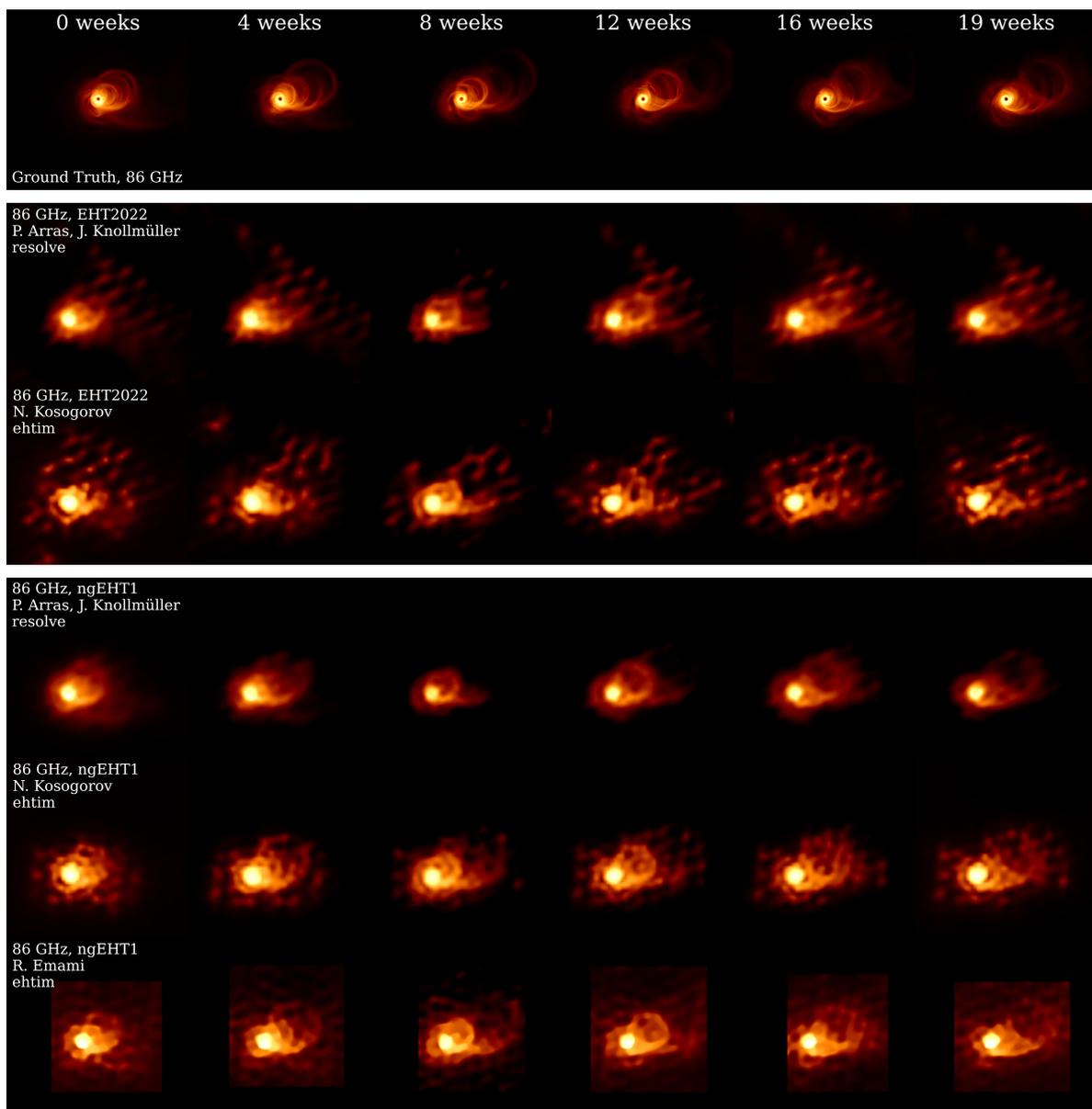


Figure 8. Selection of Challenge 2 M87 86 GHz submissions. Images are shown on a log scale, which is normalized to the brightest pixel value across each set of three movie frames, with a dynamic range of $10^{3.5}$ and field of view of 1 mas.

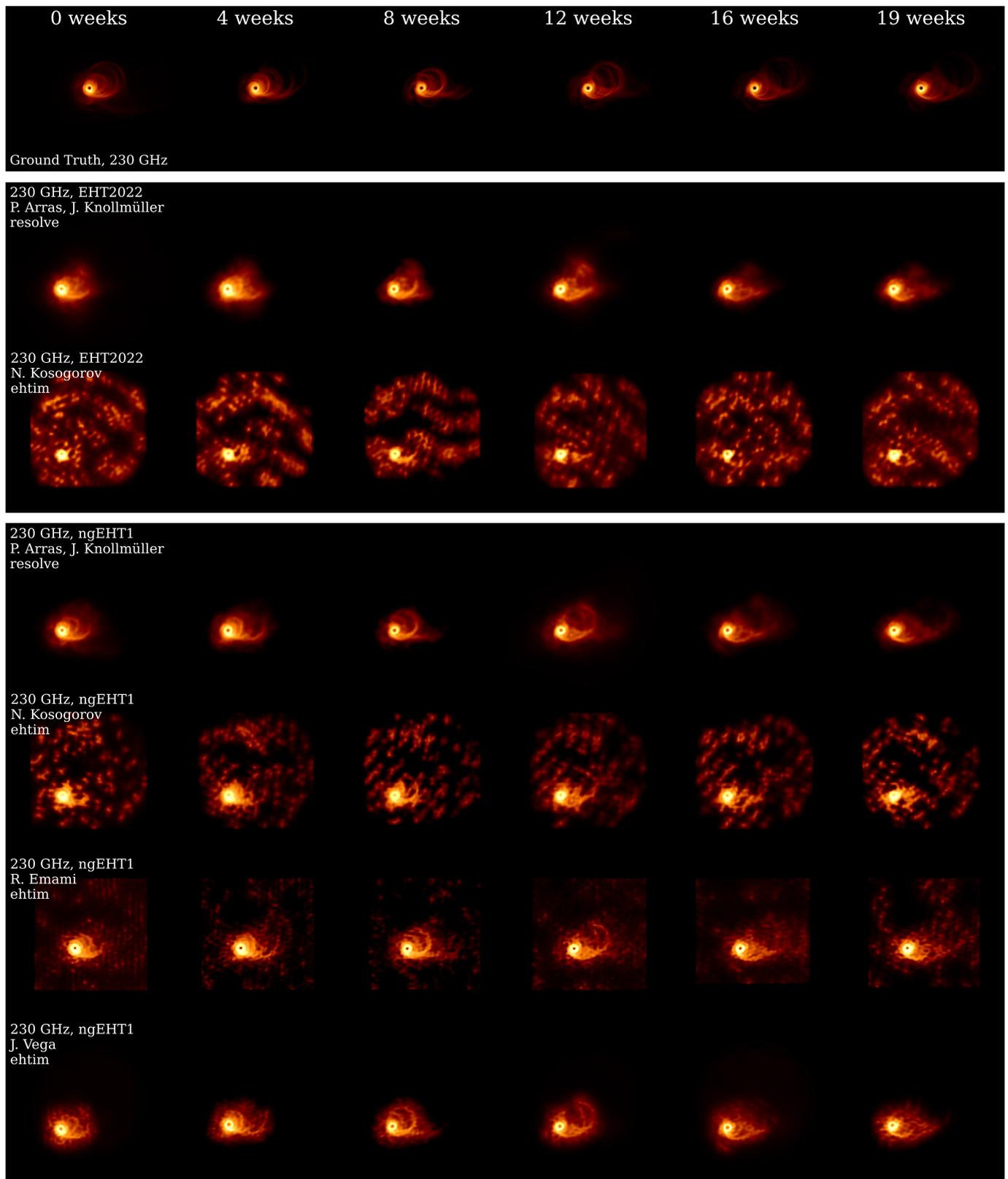


Figure 9. Selection of Challenge 2 M87 230 GHz submissions. Images are shown on a log scale, which is normalized to the brightest pixel value across each submitted set of movie frames, with a dynamic range of $10^{3.5}$, on a field of view of 1 mas.



Figure 10. Spectral index maps of the Challenge 2 M87 ground truth model (first frame) and *resolve* reconstructions of the spectral index map with the EHT2022 and ngEHT1 arrays. The ground truth spectral index map was blurred with a Gaussian with a FWHM of 9.4 μ as.

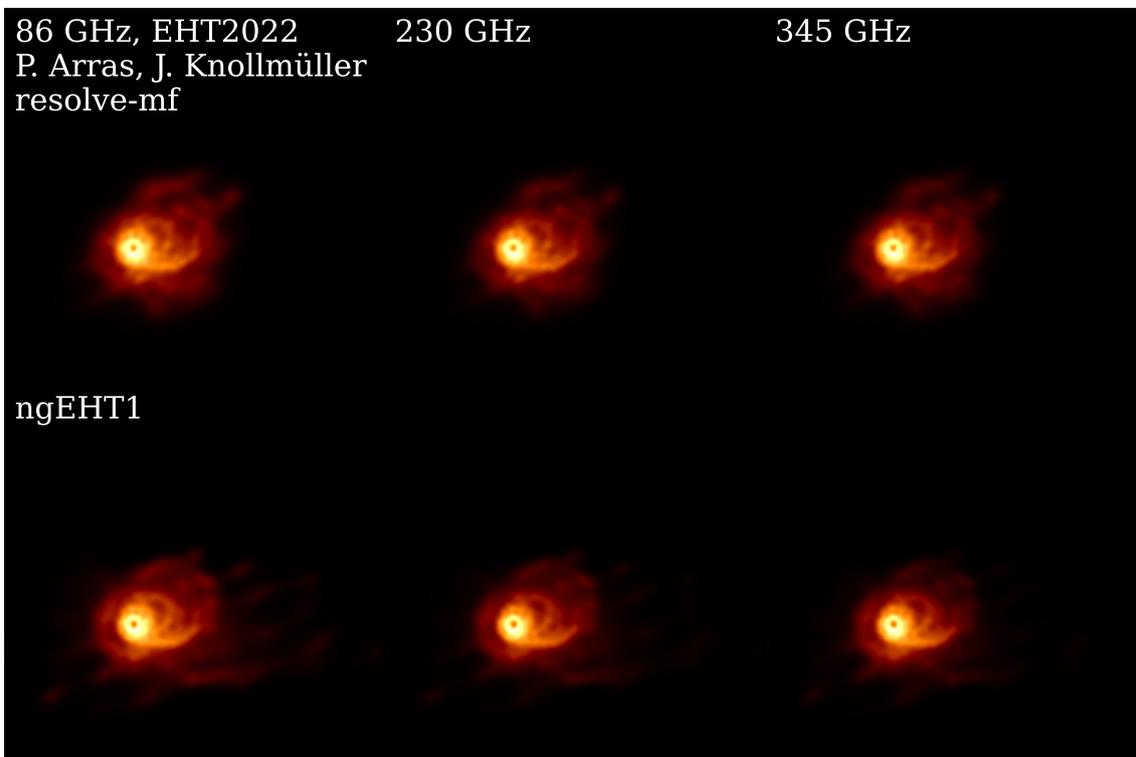


Figure 11. Multi-frequency *resolve* reconstructions of the Challenge 2 M87 model at 86, 230, and 345 GHz (first frame) with the EHT2022 and ngEHT1 arrays.

The reconstruction quality metrics show that images with low ρ_{NX} or $\rho_{\text{NX},\log}$ often have relative high χ^2 and low θ_{eff} and $\mathcal{D}_{0.1}$, with the *resolve* and especially the *resolve-mf* reconstructions performing best overall, with the caveat that the *resolve-mf* reconstructions were only performed for the first movie frame. For single-frequency reconstructions, θ_{eff} reaches 21.2 μ as at 86 GHz and 6.6 μ as at 230 GHz (median values across the 20 reconstructed frames); the super-resolution with respect to the nominal array resolution (60 and 23 μ as for 86 and 230 GHz, respectively) is significant (up to a factor 3.5) for most reconstructions. For multi-frequency reconstructions, the super-resolution factor increases even further, up to 8.6 at 86 GHz. χ^2 -values generally increase as a function of frequency, which is likely due to increased data complexity with more severe systematics. The 345 GHz reconstructions often showed difficulty in reconstructing the extended jet structure, which could be due to the sparser coverage and more severe corruptions and noise.

5.4.2. Sgr A* RIAF+hotspot

Figure 12 shows eight frames of all Sgr A* RIAF+hotspot submissions at 230 GHz. These frames span the 11–12 h UT window. This time window was chosen for the visual and metric submission comparisons as it corresponds to the first full rotation of the hotspot after it forms, and it is the hotspot rotation we aimed to reconstruct for this model. The 11–12 UT window also has a strong overlap with the ngEHT1 “best times”, e.g., [80] window for Sgr A*, with 14 stations observing Sgr A* simultaneously from 11.3 until 13.5 h UT. Furthermore, all submissions contained frames in this window, whereas the total UT ranges reconstructed varied strongly between submissions.

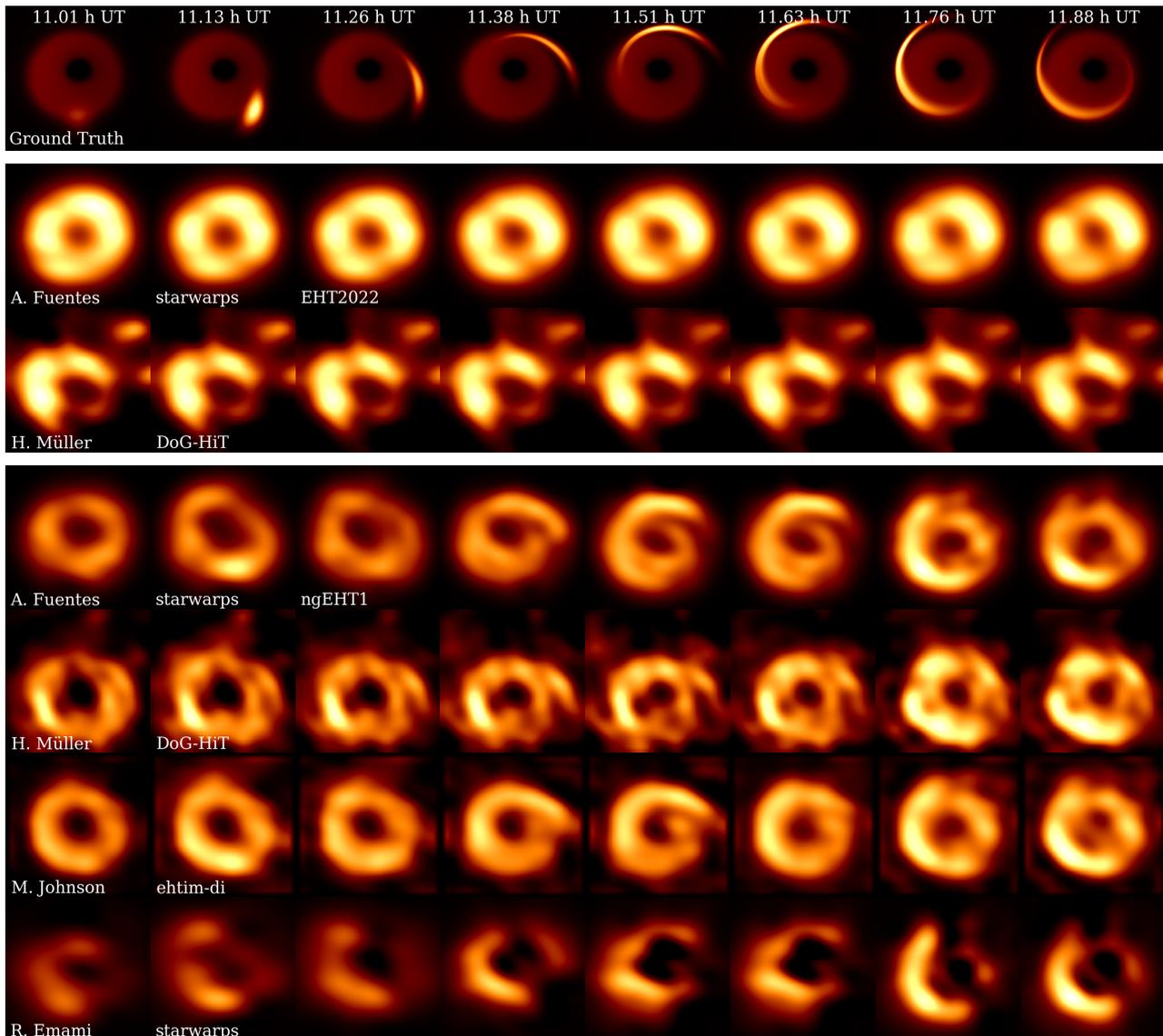


Figure 12. Challenge 2 Sgr A* RIAF+hotspot 230 GHz submissions. Images are shown on a linear scale, which is normalized to the brightest pixel value across each submitted set of movie frames on a field of view of $126 \mu\text{as}$.

None of the EHT2022 reconstructions show a significantly variable source structure, although a ring-like structure is recovered. The ngEHT1 reconstructions vary in quality, with especially the StarWarps and eht-imaging dynamical imaging algorithms recovering the shearing hotspot. This particular hotspot model was particularly challenging because of its quick shearing. Furthermore, the data sampling with 10-min scans interleaved with 10-min gaps was relatively sparse compared to the hotspot period, giving only ~ 3 scans

per full hotspot rotation. These features make the reconstruction quality obtained by some methods remarkable.

The χ^2 metrics (Table 2) are remarkably high for the Sgr A* reconstructions, which has several causes. First, in order to provide a uniform comparison, the metrics were calculated on the 11–12 h UT window only. The χ^2 are lower when considering the full UT ranges submitted (~ 3 – 4 for the highest-quality reconstructions). Considering that the source is particularly variable and hence more difficult to reconstruct during the 11–12 UT window, it is not surprising that the χ^2 are higher here. Second, the χ^2 were calculated with respect to the original synthetic data, which have a 10-s resolution within 10-min scans. Many submitters added systematic noise and time-averaged the data down to \sim minutes before imaging, which included averaging of rapidly variable structures in the visibility domain. For example, closure phases may swing by well over 120 degrees within a single 10-min scan. Combined with the high signal-to-noise ratios of the ngEHT visibilities, the fit quality to the original input data is then significantly poorer than seen during the imaging process. Finally, the submissions are compared to the data, which includes interstellar scattering, while many submitters deblurred the data before imaging. This process does not affect the closure phases, but the closure amplitudes are affected. The effective resolution θ_{eff} for the best reconstructions is comparable to the nominal array resolution, reflecting the increased difficulty of recovering intraday time-variable structures compared to static reconstructions, which often reached significant super-resolution.

Figure 13 shows the average image position angle as calculated by the Ring Extractor algorithm REx [14,81], which characterizes the properties of ring-like images for the ground truth and a few reconstructions in the 11–12 UT window. The ngEHT1 StarWarps reconstruction, in particular, shows excellent agreement with the ground truth, and the ngEHT1 ehtim-di reconstruction performs well after about 11.4 UT. The EHT2022 StarWarps reconstruction shows a stable position angle (note the 2π ambiguity) that is generally offset from the ground truth.

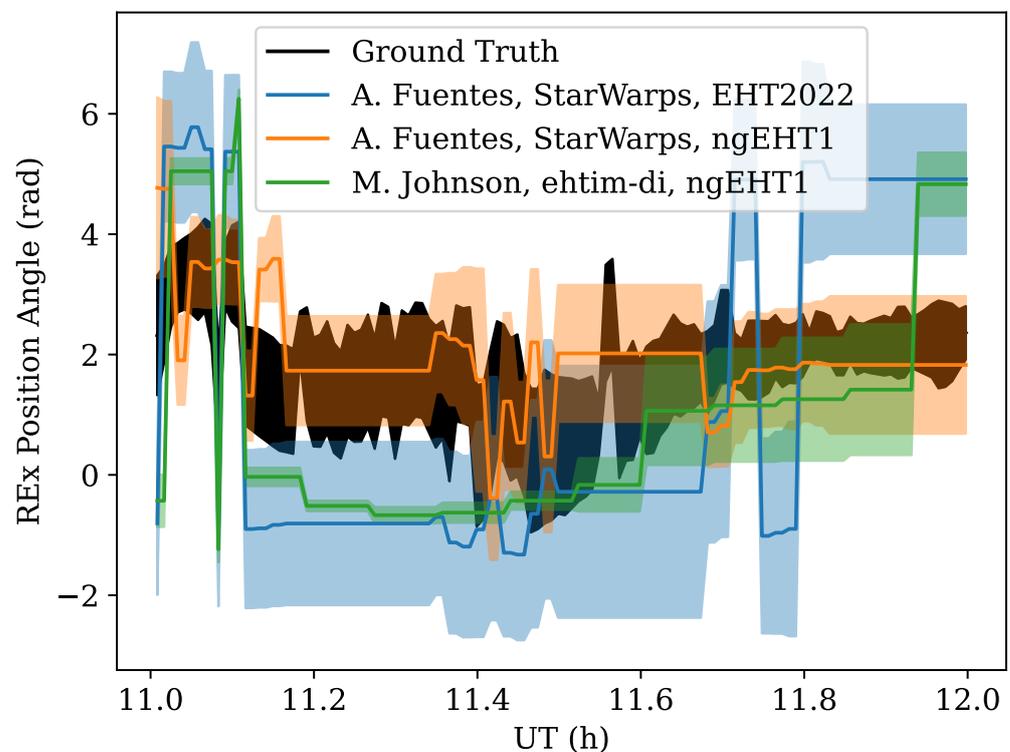


Figure 13. REx position angle fits with 1σ uncertainties for three Challenge 2 Sgr A* RIAF+hotspot 230 GHz submissions, compared to the ground truth.

Table 2. Reconstruction quality metrics for Challenge 2. Tabulated values are medians across the reconstructed frames except for the `resolve-mf` reconstructions, which were only done for the first movie frame. For the Sgr A* models, the metrics were evaluated on a common UT range for all reconstructions (see text for details).

Model	Array	ν (GHz)	Submitter	Method	χ^2_{cp}	χ^2_{lcamp}	ρ_{NX}	$\rho_{NX,log}$	θ_{eff}	$\mathcal{D}_{0.1}$
M87 GRMHD	EHT2022	86	P. Arras, J. Knollmüller	resolve	1.94	2.01	0.83	0.92	24.5	1156
M87 GRMHD	EHT2022	86	P. Arras, J. Knollmüller	resolve-mf	1.71	4.82	0.96	0.97	7.0	3970
M87 GRMHD	EHT2022	86	N. Kosogorov	ehtim	7.16	2.69	0.8	0.82	32.0	585
M87 GRMHD	ngEHT1	86	P. Arras, J. Knollmüller	resolve	1.45	1.4	0.85	0.96	21.2	3054
M87 GRMHD	ngEHT1	86	P. Arras, J. Knollmüller	resolve-mf	1.43	1.73	0.95	0.99	8.2	7248
M87 GRMHD	ngEHT1	86	R. Emami	ehtim	1.84	1.69	0.8	0.89	30.4	1315
M87 GRMHD	ngEHT1	86	N. Kosogorov	ehtim	2.06	1.58	0.8	0.93	30.4	919
M87 GRMHD	ngEHT1	86	N. Kosogorov	CLEAN	193.55	10,266.39	0.75	0.74	46.8	749
M87 GRMHD	EHT2022	230	P. Arras, J. Knollmüller	resolve	2.03	3.25	0.92	0.96	7.3	3881
M87 GRMHD	EHT2022	230	P. Arras, J. Knollmüller	resolve-mf	2.03	6.67	0.93	0.97	7.1	6424
M87 GRMHD	EHT2022	230	N. Kosogorov	ehtim	4.16	3.15	0.88	0.54	12.6	429
M87 GRMHD	ngEHT1	230	P. Arras, J. Knollmüller	resolve	2.53	2.35	0.92	0.98	6.6	8742
M87 GRMHD	ngEHT1	230	P. Arras, J. Knollmüller	resolve-mf	2.57	3.12	0.93	0.99	7.1	12,154
M87 GRMHD	ngEHT1	230	J. Vega	ehtim	2.55	2.3	0.91	0.97	8.1	4807
M87 GRMHD	ngEHT1	230	R. Emami	ehtim	2.55	2.47	0.89	0.83	10.9	2061
M87 GRMHD	ngEHT1	230	N. Kosogorov	ehtim	2.85	2.81	0.89	0.71	11.4	1060
M87 GRMHD	ngEHT1	230	N. Kosogorov	CLEAN	325.47	385.28	0.79	0.6	22.6	226
M87 GRMHD	EHT2022	345	P. Arras, J. Knollmüller	resolve-mf	5.26	5.79	0.93	0.97	7.2	6994
M87 GRMHD	ngEHT1	345	P. Arras, J. Knollmüller	resolve-mf	6.39	6.89	0.92	0.98	7.3	9732
M87 GRMHD	ngEHT1	345	R. Emami	ehtim	6.14	5.38	0.59	0.42	61.8	61
M87 GRMHD	ngEHT1	345	N. Kosogorov	ehtim	5.99	4.94	0.81	0.47	16.6	563
M87 GRMHD	ngEHT1	345	N. Kosogorov	CLEAN	12.41	16.28	0.83	0.67	14.4	1157
Sgr A* RIAFSPOT	EHT2022	230	A. Fuentes	StarWarps	1.85	1.78	0.83	-	37.3	-
Sgr A* RIAFSPOT	EHT2022	230	H. Müller	DoG-HiT	5.61	5.12	0.77	-	56.8	-
Sgr A* RIAFSPOT	ngEHT1	230	M. Johnson	ehtim-di	7.39	11.78	0.87	-	24.4	-
Sgr A* RIAFSPOT	ngEHT1	230	A. Fuentes	StarWarps	4.24	3.05	0.89	-	23.0	-
Sgr A* RIAFSPOT	ngEHT1	230	R. Emami	StarWarps	6.87	11.98	0.83	-	43.3	-
Sgr A* RIAFSPOT	ngEHT1	230	H. Müller	DoG-HiT	33.31	38.91	0.84	-	33.0	-
Sgr A* RIAFSPOT	ngEHT1	345	A. Fuentes	StarWarps	5.37	3.63	0.85	-	28.6	-
Sgr A* RIAFSPOT	ngEHT1	345	R. Emami	StarWarps	5.7	3.86	0.74	-	56.8	-
Sgr A* GRMHD	EHT2022	230	A. Fuentes	StarWarps	9.49	3.61	0.68	-	56.0	-
Sgr A* GRMHD	EHT2022	230	H. Müller	DoG-HiT	153.81	32.15	0.68	-	57.4	-
Sgr A* GRMHD	ngEHT1	230	M. Johnson	ehtim-di	3.99	7.14	0.87	-	18.4	-
Sgr A* GRMHD	ngEHT1	230	A. Fuentes	StarWarps	3.97	7.47	0.85	-	21.1	-
Sgr A* GRMHD	ngEHT1	230	R. Emami	StarWarps	4.0	6.91	0.87	-	17.5	-
Sgr A* GRMHD	ngEHT1	230	H. Müller	DoG-HiT	13.88	8.18	0.8	-	29.0	-
Sgr A* GRMHD	ngEHT1	230	P. Arras, J. Knollmüller	resolve	5.57	4.52	0.84	-	21.9	-
Sgr A* GRMHD	ngEHT1	345	R. Emami	StarWarps	4.94	4.19	0.61	-	56.9	-

5.4.3. Sgr A* GRMHD

Finally, Figure 14 shows eleven frames of all Sgr A* GRMHD submissions at 230 GHz, spanning the best times window (11.3–13.5 h UT). Like for the Sgr A* RIAF+hotspot model, the EHT2022 reconstructions are static. In the StarWarps reconstruction, the ring morphology is recovered, but the detailed emission along the ring is not. The ngEHT1 reconstructions are generally much sharper, and the azimuthal brightness variations are reconstructed accurately, with the `ehtim-di` and StarWarps submissions showing the best quality metric values (Table 2). Due to the relatively stable and turbulent nature of the variability in this model, the reconstruction of temporal variations is more difficult to assess than for the other source models.

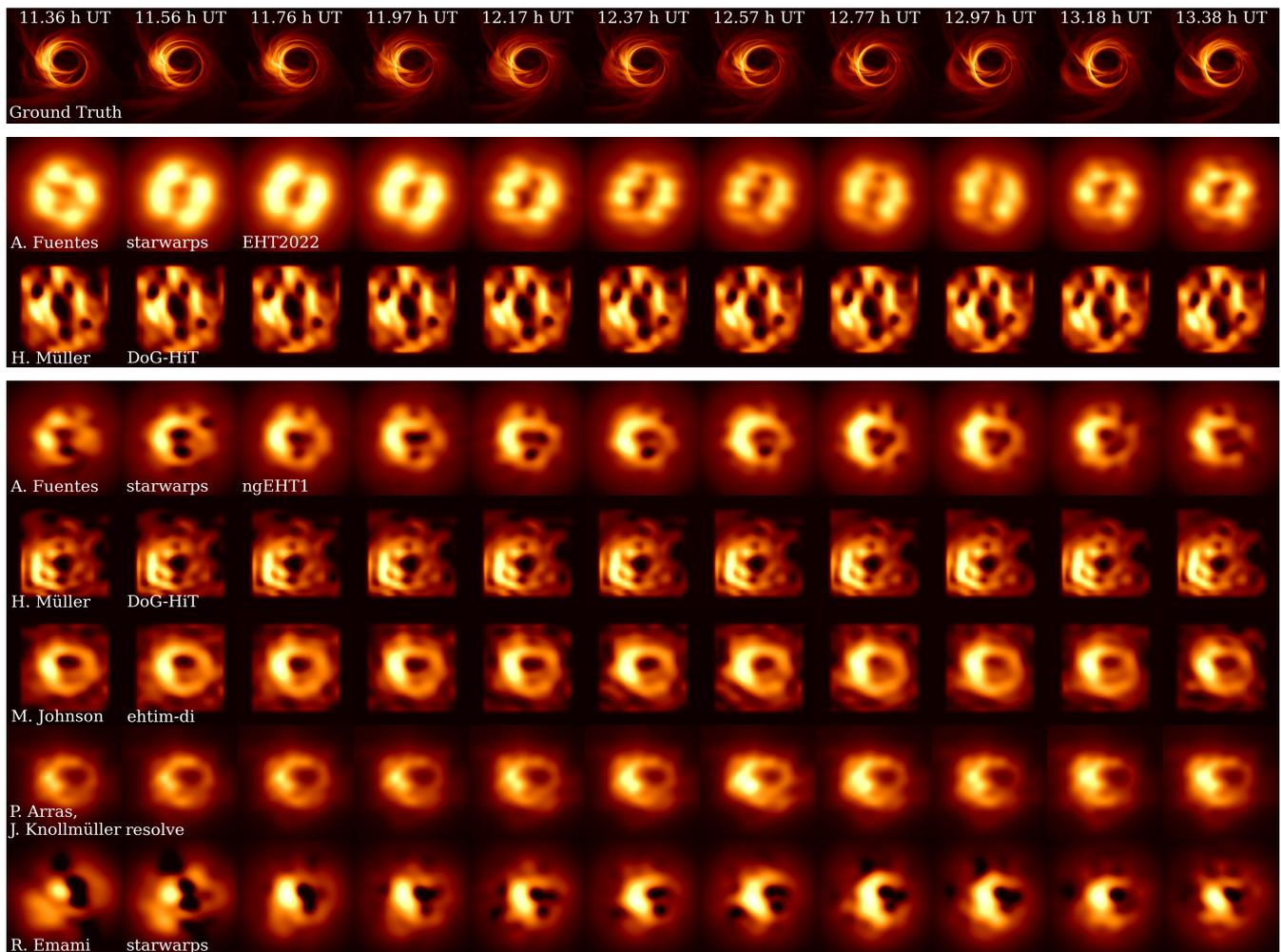


Figure 14. Challenge 2 Sgr A* GRMHD 230 GHz submissions. Images are shown on a square root scale, which is normalized to the brightest pixel value across each submitted set of movie frames on a field of view of 126 μs .

6. Conclusions and Outlook

The first two ngEHT Analysis Challenges have provided a number of useful insights. First, current imaging algorithms are capable of reconstructing high-resolution and high-fidelity movies of M87* and Sgr A*, revealing the jet dynamics of M87* on mas scales down to the event horizon and intra-hour dynamics of the Sgr A* accretion flow on event-horizon scales. This work provides high-quality reconstructions from synthetic data that include fully realistic observation and calibration effects under mediocre weather conditions (median conditions for April), showing excellent prospects for ngEHT performance.

The sources have been reconstructed with a breadth of traditional (EHT) imaging algorithms and newer algorithms that are under active development, such as *resolve* and *DoG-HiT*. Different algorithms showed different performances for different datasets. For example, the *StarWarps* and *eht-imaging* algorithms showed strong performance on dynamical reconstructions of Sgr A* (see [40,82] for other examples), while the *resolve* algorithm did remarkably well on recovering the extended jet structure of M87. Multi-frequency reconstructions gave the best M87 jet reconstruction results for both challenges, providing an opportunity to produce high-resolution 86 GHz images showing the central brightness depression related to the black hole (inner) shadow [83,84] and faithful reconstructions of the extended jet at 345 GHz. Both these features are difficult to reconstruct using the individual frequencies alone. *DoG-HiT* reconstructions of the Sgr A* RIAF+hotspot model resolved the hotspot orbit and shearing, albeit with a lower quality than with other

methods. Since DoG-HiT is the most recently developed algorithm used in the challenges and is completely automatic without special manual adaption to the data sets, these results are promising and can inform further development. SMILI and CLEAN have been applied to Challenge 1 data only, where they did not perform as well as eht-imaging in reconstructing the extended M87 jet.

The reconstructions from any algorithm do not necessarily show its maximum potential performance. Between algorithms, there are differences in the freedom that the user has to steer the reconstruction process. CLEAN traditionally requires significant user input and steering (e.g., defining CLEAN windows) but has been adapted to a more automated approach for EHT analysis [14]. RML methods such as eht-imaging and SMILI require setting regularizers and weights but also allow some input on the reconstruction procedure by setting, e.g., convergence criteria and blurring steps between image rounds. On the other hand, DoG-HiT depends on just one free parameter. The outcome of Bayesian methods generally depends on the set priors. From the results of these challenges, each method's dependence on user input is difficult to assess and would require dedicated parameter surveys. Based on, e.g., the eht-imaging submissions, the results can certainly depend strongly on the user. However, for submissions reconstructed with the same method but with strongly different resulting images, the χ^2 are a good indicator of the reconstruction quality. For the lower-quality image reconstructions, either the used parameters or the script setup often did not allow a good fit to the data.

Regarding data generation, one lesson learned is that the used schedule of 10-min scans and 10-min gaps makes reconstructing the rapid variability of Sgr A* challenging, and in fact, it is remarkable that dynamical imaging algorithms were able to reconstruct the one-hour period and rapidly shearing Sgr A* hotspot orbit with just three scans and a duty cycle of 50%. A denser schedule with shorter gaps could potentially improve these reconstructions significantly and also help in reconstructing the rapid variability from GRMHD simulations. Furthermore, the Challenge 2 345 GHz data have proven difficult to image, which is likely attributable to the severe atmospheric effects considering the weather parameters were medians for April at all sites. Since, in reality, 345 GHz observations would likely only be scheduled on days with excellent weather at suitable sites, a next challenge should be performed with more optimistic 345 GHz weather conditions. The experience from these challenges has shown that both eht-imaging and SYMBA are viable and well-performing pathways for generating synthetic ngEHT data. User-friendly tools to generate synthetic ngEHT data from a centralized repository of instrument and weather parameters using both pathways are under development [2].

ngEHT Analysis Challenge 3³ is an extension of Challenge 2 to full Stokes and will show how well various algorithms can reconstruct dynamics in polarization. Challenge 4 will focus on more specific ngEHT science goals, such as measuring the photon ring size and black hole spin, involving modeling methods as well, e.g., [85]. The merit of simultaneous multi-frequency observations allowing for frequency phase transfer, e.g., [86,87], will be explored in this challenge as well (see also [17]). Future challenges could also involve varying the number and locations of stations. While the impact of a single station's presence or location diminishes as the array grows and becomes more robust against station losses, the effect of using partial instead of the full array or using different sets of new sites could be tested in end-to-end simulations (see also [2]).

The ngEHT Analysis Challenges have brought together expertise in theoretical modeling, synthetic data generation, and image reconstruction, spurring development in all these areas. The continued challenges will involve polarization, model fitting, and science interpretation to form a complete and end-to-end process of ngEHT simulations, which helps in maximizing the science potential of the array.

Author Contributions: Conceptualization, L.B., S.S.D., M.D.J. and F.R.; methodology, P.A., L.B., K.C., R.E., C.F., A.F., M.D.J., J.K., N.K., H.M., N.P., A.R., F.R., P.T., E.T. and J.V.; software, G.L., P.A., K.C., C.F., J.K., H.M., M.D.J. and F.R.; validation, L.B. and F.R.; formal analysis, P.A., L.B., K.C., R.E., C.F., A.F., M.D.J., J.K., N.K., H.M., N.P., A.R., F.R., P.T., E.T. and J.V.; data curation, L.B., G.L. and F.R.; writing—original draft preparation, P.A., R.E., M.D.J., H.M., F.R. and E.T.; writing—review and editing, P.A., L.B., K.C., S.S.D., R.E., C.F., A.F., M.D.J., J.K., N.K., G.L., H.M., N.P., A.R., F.R., P.T., E.T. and J.V.; visualization, F.R.; supervision, S.S.D.; funding acquisition, S.S.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by NSF grants AST-1935980 and AST-2034306. This work was supported by the Black Hole Initiative at Harvard University, made possible through the support of grants from the Gordon and Betty Moore Foundation and the John Templeton Foundation. The opinions expressed in this publication are those of the author(s) and do not necessarily reflect the views of the Moore or Templeton Foundations. Hendrik Müller received financial support for this research from the International Max Planck Research School (IMPRS) for Astronomy and Astrophysics at the Universities of Bonn and Cologne. This research is supported by the DFG research grant “Jet physics on horizon scales and beyond” (Grant No. FR 4069/2-1), the ERC synergy grant “BlackHoleCam: Imaging the Event Horizon of Black Holes” (Grant No. 610058) and ERC advanced grant “JETSET: Launching, propagation and emission of relativistic jets from binary mergers and across mass scales” (Grant No. 884631). Jakob Knollmüller acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy—EXC 2094—390783311. Razieh Emami acknowledges the support by the Institute for Theory and Computation at the Center for Astrophysics as well as grant numbers 21-atp21-0077, NSF AST-1816420 and HST-GO-16173.001-A for very generous support.

Data Availability Statement: The synthetic data generated for this work can be downloaded from <https://challenge.ngeht.org/> (accessed on 19 December 2022), with access provided upon request to the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Notes

- ¹ <https://challenge.ngeht.org/> (accessed on 19 December 2022)
- ² <https://gitlab.mpcdf.mpg.de/ift/resolve> (accessed on 19 December 2022)
- ³ <https://challenge.ngeht.org/challenge3/> (accessed on 19 December 2022).

References

1. Doeleman, S.; Blackburn, L.; Dexter, J.; Gomez, J.L.; Johnson, M.D.; Palumbo, D.C.; Weintroub, J.; Farah, J.R.; Fish, V.; Loinard, L.; et al. Studying Black Holes on Horizon Scales with VLBI Ground Arrays. *Proc. Bull. Am. Astron. Soc.* **2019**, *51*, 256.
2. Doeleman, S. Reference Array and Design Consideration for the next-generation Event Horizon Telescope. *Galaxies* **2023**, *in prep.*
3. Event Horizon Telescope Collaboration; Akiyama, K.; Alberdi, A.; Alef, W.; Asada, K.; Azulay, R.; Baczkowski, A.K.; Ball, D.; Baloković, M.; Barrett, J.; et al. First M87 Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole. *Astrophys. J. Lett.* **2019**, *875*, L1. [[CrossRef](#)]
4. Event Horizon Telescope Collaboration; Akiyama, K.; Alberdi, A.; Alef, W.; Asada, K.; Azulay, R.; Baczkowski, A.K.; Ball, D.; Baloković, M.; Barrett, J.; et al. First M87 Event Horizon Telescope Results. II. Array and Instrumentation. *Astrophys. J. Lett.* **2019**, *875*, L2. [[CrossRef](#)]
5. Event Horizon Telescope Collaboration; Akiyama, K.; Alberdi, A.; Alef, W.; Asada, K.; Azulay, R.; Baczkowski, A.K.; Ball, D.; Baloković, M.; Barrett, J.; et al. First M87 Event Horizon Telescope Results. III. Data Processing and Calibration. *Astrophys. J. Lett.* **2019**, *875*, L3. [[CrossRef](#)]
6. Event Horizon Telescope Collaboration; Akiyama, K.; Alberdi, A.; Alef, W.; Asada, K.; Azulay, R.; Baczkowski, A.K.; Ball, D.; Baloković, M.; Barrett, J.; et al. First M87 Event Horizon Telescope Results. IV. Imaging the Central Supermassive Black Hole. *Astrophys. J. Lett.* **2019**, *875*, L4. [[CrossRef](#)]
7. Event Horizon Telescope Collaboration; Akiyama, K.; Alberdi, A.; Alef, W.; Asada, K.; Azulay, R.; Baczkowski, A.K.; Ball, D.; Baloković, M.; Barrett, J.; et al. First M87 Event Horizon Telescope Results. V. Physical Origin of the Asymmetric Ring. *Astrophys. J. Lett.* **2019**, *875*, L5. [[CrossRef](#)]

8. Event Horizon Telescope Collaboration; Akiyama, K.; Alberdi, A.; Alef, W.; Asada, K.; Azulay, R.; Baczko, A.K.; Ball, D.; Baloković, M.; Barrett, J.; et al. First M87 Event Horizon Telescope Results. VI. The Shadow and Mass of the Central Black Hole. *Astrophys. J. Lett.* **2019**, *875*, L6. [[CrossRef](#)]
9. Event Horizon Telescope Collaboration; Akiyama, K.; Algaba, J.C.; Alberdi, A.; Alef, W.; Anantua, R.; Asada, K.; Azulay, R.; Baczko, A.K.; Ball, D.; et al. First M87 Event Horizon Telescope Results. VII. Polarization of the Ring. *Astrophys. J. Lett.* **2021**, *910*, L12. [[CrossRef](#)]
10. Event Horizon Telescope Collaboration; Akiyama, K.; Algaba, J.C.; Alberdi, A.; Alef, W.; Anantua, R.; Asada, K.; Azulay, R.; Baczko, A.K.; Ball, D.; et al. First M87 Event Horizon Telescope Results. VIII. Magnetic Field Structure near The Event Horizon. *Astrophys. J. Lett.* **2021**, *910*, L13. [[CrossRef](#)]
11. Event Horizon Telescope Collaboration; Akiyama, K.; Alberdi, A.; Alef, W.; Algaba, J.C.; Anantua, R.; Asada, K.; Azulay, R.; Bach, U.; Baczko, A.K.; et al. First Sagittarius A* Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole in the Center of the Milky Way. *Astrophys. J. Lett.* **2022**, *930*, L12. [[CrossRef](#)]
12. Event Horizon Telescope Collaboration; Akiyama, K.; Alberdi, A.; Alef, W.; Algaba, J.C.; Anantua, R.; Asada, K.; Azulay, R.; Bach, U.; Baczko, A.K.; et al. First Sagittarius A* Event Horizon Telescope Results. II. EHT and Multiwavelength Observations, Data Processing, and Calibration. *Astrophys. J. Lett.* **2022**, *930*, L13. [[CrossRef](#)]
13. Event Horizon Telescope Collaboration; Akiyama, K.; Alberdi, A.; Alef, W.; Algaba, J.C.; Anantua, R.; Asada, K.; Azulay, R.; Bach, U.; Baczko, A.K.; et al. First Sagittarius A* Event Horizon Telescope Results. III. Imaging of the Galactic Center Supermassive Black Hole. *Astrophys. J. Lett.* **2022**, *930*, L14. [[CrossRef](#)]
14. Event Horizon Telescope Collaboration; Akiyama, K.; Alberdi, A.; Alef, W.; Algaba, J.C.; Anantua, R.; Asada, K.; Azulay, R.; Bach, U.; Baczko, A.K.; et al. First Sagittarius A* Event Horizon Telescope Results. IV. Variability, Morphology, and Black Hole Mass. *Astrophys. J. Lett.* **2022**, *930*, L15. [[CrossRef](#)]
15. Event Horizon Telescope Collaboration; Akiyama, K.; Alberdi, A.; Alef, W.; Algaba, J.C.; Anantua, R.; Asada, K.; Azulay, R.; Bach, U.; Baczko, A.K.; et al. First Sagittarius A* Event Horizon Telescope Results. V. Testing Astrophysical Models of the Galactic Center Black Hole. *Astrophys. J. Lett.* **2022**, *930*, L16. [[CrossRef](#)]
16. Event Horizon Telescope Collaboration; Akiyama, K.; Alberdi, A.; Alef, W.; Algaba, J.C.; Anantua, R.; Asada, K.; Azulay, R.; Bach, U.; Baczko, A.K.; et al. First Sagittarius A* Event Horizon Telescope Results. VI. Testing the Black Hole Metric. *Astrophys. J. Lett.* **2022**, *930*, L17. [[CrossRef](#)]
17. Issaoun, S.; Pesce, D.W.; Roelofs, F.; Chael, A.; Dodson, R.; Rioja, M.J.; Akiyama, K.; Aran, R.; Blackburn, L.; Doleman, S.S.; et al. Enabling transformational ngEHT science via the inclusion of 86 GHz capabilities. *Galaxies* **2023**, *submitted*.
18. Pesce, D.W.; Palumbo, D.C.M.; Narayan, R.; Blackburn, L.; Doleman, S.S.; Johnson, M.D.; Ma, C.P.; Nagar, N.M.; Natarajan, P.; Ricarte, A. Toward Determining the Number of Observable Supermassive Black Hole Shadows. *Astrophys. J. Lett.* **2021**, *923*, 260. [[CrossRef](#)]
19. Pesce, D.W.; Palumbo, D.C.M.; Ricarte, A.; Broderick, A.E.; Johnson, M.D.; Nagar, N.M.; Natarajan, P.; Gómez, J.L. Expectations for Horizon-Scale Supermassive Black Hole Population Studies with the ngEHT. *Galaxies* **2022**, *10*, 109. [[CrossRef](#)]
20. Bouman, K.L. Extreme Imaging via Physical Model Inversion: Seeing around Corners and Imaging Black Holes. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2017.
21. Chatterjee, K.; Chael, A.; Tiede, P.; Mizuno, Y.; Emami, R.; Fromm, C.; Ricarte, A.; Blackburn, L.; Roelofs, F.; Johnson, M.D.; et al. Comparing accretion flow morphology in numerical simulations of black holes from the ngEHT Model Library: The impact of radiation physics. *arXiv* **2022**, arXiv:2212.01804.
22. Högbom, J.A. Aperture Synthesis with a Non-Regular Distribution of Interferometer Baselines. *Astron. Astrophys. Suppl. Ser.* **1974**, *15*, 417.
23. Wilkinson, P.N.; Readhead, A.C.S.; Purcell, G.H.; Anderson, B. Radio structure of 3C 147 determined by multi-element very long baseline interferometry. *Nature* **1977**, *269*, 764–768. [[CrossRef](#)]
24. Cornwell, T.J. Hogbom's CLEAN algorithm. Impact on astronomy and beyond. Commentary on: Högbom J. A., 1974, A&AS, 15, 417. *Astron. Astrophys.* **2009**, *500*, 65–66. [[CrossRef](#)]
25. Shepherd, M.C.; Pearson, T.J.; Taylor, G.B. DIFMAP: An interactive program for synthesis imaging. *Proc. Bull. Am. Astron. Soc.* **1995**, *27*, 903.
26. Frieden, B.R. Restoring with Maximum Likelihood and Maximum Entropy. *J. Opt. Soc. Am.* **1972**, *62*, 511–518. [[CrossRef](#)] [[PubMed](#)]
27. Coughlan, C.P.; Gabuzda, D.C. High resolution VLBI polarization imaging of AGN with the maximum entropy method. *Mon. Not. R. Astron. Soc.* **2016**, *463*, 1980–2001. [[CrossRef](#)]
28. Narayan, R.; Nityananda, R. Maximum Entropy Image Restoration in Astronomy. *Annu. Rev. Astron. Astrophys.* **1986**, *24*, 127–170. [[CrossRef](#)]
29. Chael, A.A.; Johnson, M.D.; Bouman, K.L.; Blackburn, L.L.; Akiyama, K.; Narayan, R. Interferometric Imaging Directly with Closure Phases and Closure Amplitudes. *Astrophys. J.* **2018**, *857*, 23. [[CrossRef](#)]
30. Chael, A.A.; Johnson, M.D.; Narayan, R.; Doleman, S.S.; Wardle, J.F.C.; Bouman, K.L. High-resolution Linear Polarimetric Imaging for the Event Horizon Telescope. *Astrophys. J.* **2016**, *829*, 11. [[CrossRef](#)]

31. Akiyama, K.; Kuramochi, K.; Ikeda, S.; Fish, V.L.; Tazaki, F.; Honma, M.; Doeleman, S.S.; Broderick, A.E.; Dexter, J.; Mościbrodzka, M.; et al. Imaging the Schwarzschild-radius-scale Structure of M87 with the Event Horizon Telescope Using Sparse Modeling. *Astrophys. J.* **2017**, *838*, 1. [[CrossRef](#)]
32. Akiyama, K.; Ikeda, S.; Pleau, M.; Fish, V.L.; Tazaki, F.; Kuramochi, K.; Broderick, A.E.; Dexter, J.; Mościbrodzka, M.; Gowanlock, M.; et al. Superresolution Full-polarimetric Imaging for Radio Interferometry with Sparse Modeling. *Astrophys. J.* **2017**, *153*, 159. [[CrossRef](#)]
33. Chael, A.A.; Bouman, K.L.; Johnson, M.D.; Narayan, R.; Doeleman, S.S.; Wardle, J.F.C.; Blackburn, L.L.; Akiyama, K.; Wielgus, M.; Chan, C.k.; et al. ehtim: Imaging, analysis, and simulation software for radio interferometry. *Astrophys. Source Code Libr.* **2019**, ascl-1904.
34. Chael, A.; Issaoun, S.; Pesce, D.W.; Johnson, M.D.; Ricarte, A.; Fromm, C.M.; Mizuno, Y. Multi-frequency Black Hole Imaging for the Next-Generation Event Horizon Telescope. *arXiv* **2022**, arXiv:2210.12226.
35. Johnson, M.D.; Narayan, R.; Psaltis, D.; Blackburn, L.; Kovalev, Y.Y.; Gwinn, C.R.; Zhao, G.Y.; Bower, G.C.; Moran, J.M.; Kino, M.; et al. The Scattering and Intrinsic Structure of Sagittarius A* at Radio Wavelengths. *Astrophys. J.* **2018**, *865*, 104. [[CrossRef](#)]
36. Johnson, M.D.; Lupsasca, A.; Strominger, A.; Wong, G.N.; Hadar, S.; Kapec, D.; Narayan, R.; Chael, A.; Gammie, C.F.; Galison, P.; et al. Universal interferometric signatures of a black hole's photon ring. *Sci. Adv.* **2020**, *6*, eaaz1310. [[CrossRef](#)] [[PubMed](#)]
37. Byrd, R.H.; Lu, P.; Nocedal, J. A Limited Memory Algorithm for Bound Constrained Optimization. *SIAM J. Sci. Stat. Comput.* **1995**, *16*, 1190–1208. [[CrossRef](#)]
38. Jones, E.; Oliphant, T.; Peterson, P. SciPy: Open Source Scientific Tools for Python. 2001. Available online: <https://github.com/takluyver/scipy.org-new/blob/master/www/scipylib/citing.rst> (accessed on 25 August 2015).
39. Bouman, K.L.; Johnson, M.D.; Dalca, A.V.; Chael, A.A.; Roelofs, F.; Doeleman, S.S.; Freeman, W.T. Reconstructing Video from Interferometric Measurements of Time-Varying Sources. *arXiv* **2017**, arXiv:1711.01357.
40. Emami, R.; Tiede, P.; Doeleman, S.S.; Roelofs, F.; Wielgus, M.; Blackburn, L.; Liska, M.; Chatterjee, K.; Ripperda, B.; Fuentes, A.; et al. Tracing the hot spot motion using the next generation Event Horizon Telescope (ngEHT). *arXiv* **2022**, arXiv:2211.06773.
41. Arras, P.; Bester, H.L.; Perley, R.A.; Leike, R.; Smirnov, O.; Westermann, R.; Enßlin, T.A. Comparison of classical and Bayesian imaging in radio interferometry. *Astron. Astrophys.* **2021**, *646*, A84. [[CrossRef](#)]
42. Arras, P.; Frank, P.; Haim, P.; Knollmüller, J.; Leike, R.; Reinecke, M.; Enßlin, T. Variable structures in M87* from space, time and frequency d interferometry. *Nat. Astron.* **2022**, *6*, 259–269. [[CrossRef](#)]
43. Knollmüller, J.; Enßlin, T.A. Metric Gaussian Variational Inference. *arXiv* **2019**, arXiv:1901.11033. [[CrossRef](#)]
44. Arras, P.; Reinecke, M.; Westermann, R.; Enßlin, T.A. Efficient wide-field radio interferometry response. *Astron. Astrophys.* **2021**, *646*, A58. [[CrossRef](#)]
45. Müller, H.; Lobanov, A.P. DoG-HiT: A novel VLBI multiscale imaging approach. *Astron. Astrophys.* **2022**, *666*, A137. [[CrossRef](#)]
46. Müller, H.; Lobanov, A. Multi-scale and Multi-directional VLBI Imaging with CLEAN. *Astron. Astrophys.* **2022**, submitted.
47. Müller, H.; Lobanov, A. Dynamic polarimetry with the multiresolution support. *Astron. Astrophys.* **2022**, submitted.
48. Thompson, A.R.; Moran, J.M.; Swenson, G.W., Jr. *Interferometry and Synthesis in Radio Astronomy*, 3rd ed.; Springer Nature: Berlin/Heidelberg, Germany, 2017. [[CrossRef](#)]
49. Bustamante, S.; Blackburn, L.; Narayanan, G.; Schloerb, F.P.; Hughes, D. The Role of the Large Millimeter Telescope in Black Hole Science with the Next-Generation Event Horizon Telescope. *Galaxies* **2023**, *11*, 2. [[CrossRef](#)]
50. Mizuno, Y.; Fromm, C.M.; Younsi, Z.; Porth, O.; Olivares, H.; Rezzolla, L. Comparison of the ion-to-electron temperature ratio prescription: GRMHD simulations with electron thermodynamics. *Mon. Not. R. Astron. Soc.* **2021**, *506*, 741–758. [[CrossRef](#)]
51. Porth, O.; Olivares, H.; Mizuno, Y.; Younsi, Z.; Rezzolla, L.; Moscibrodzka, M.; Falcke, H.; Kramer, M. The black hole accretion code. *Comput. Astrophys. Cosmol.* **2017**, *4*, 1. [[CrossRef](#)]
52. Tchekhovskoy, A.; McKinney, J.C. Prograde and retrograde black holes: Whose jet is more powerful? *Mon. Not. R. Astron. Soc. Lett.* **2012**, *423*, L55–L59. [[CrossRef](#)]
53. Younsi, Z.; Wu, K.; Fuerst, S.V. General relativistic radiative transfer: Formulation and emission from structured tori around black holes. *Astron. Astrophys.* **2012**, *545*, A13. [[CrossRef](#)]
54. Younsi, Z.; Porth, O.; Mizuno, Y.; Fromm, C.M.; Olivares, H. Modelling the polarised emission from black holes on event horizon-scales. In *Proceedings of the Perseus in Sicily: From Black Hole to Cluster Outskirts*; Asada, K., de Gouveia Dal Pino, E., Giroletti, M., Nagai, H., Nemmen, R., Eds.; Cambridge University Press: Cambridge, UK, 2020; Volume 342, pp. 9–12. [[CrossRef](#)]
55. Pandya, A.; Zhang, Z.; Chandra, M.; Gammie, C.F. Polarized Synchrotron Emissivities and Absorptivities for Relativistic Thermal, Power-law, and Kappa Distribution Functions. *Astrophys. J.* **2016**, *822*, 34. [[CrossRef](#)]
56. Ball, D.; Sironi, L.; Özel, F. Electron and Proton Acceleration in Trans-relativistic Magnetic Reconnection: Dependence on Plasma Beta and Magnetization. *Astrophys. J.* **2018**, *862*, 80. [[CrossRef](#)]
57. Davelaar, J.; Olivares, H.; Porth, O.; Bronzwaer, T.; Janssen, M.; Roelofs, F.; Mizuno, Y.; Fromm, C.M.; Falcke, H.; Rezzolla, L. Modeling non-thermal emission from the jet-launching region of M 87 with adaptive mesh refinement. *Astron. Astrophys.* **2019**, *632*, A2. [[CrossRef](#)]
58. Fromm, C.M.; Cruz-Osorio, A.; Mizuno, Y.; Nathanail, A.; Younsi, Z.; Porth, O.; Olivares, H.; Davelaar, J.; Falcke, H.; Kramer, M.; et al. Impact of non-thermal particles on the spectral and structural properties of M87. *Astron. Astrophys.* **2022**, *660*, A107. [[CrossRef](#)]

59. Broderick, A.E.; Loeb, A. Imaging optically-thin hotspots near the black hole horizon of Sgr A* at radio and near-infrared wavelengths. *Mon. Not. R. Astron. Soc.* **2006**, *367*, 905–916. [[CrossRef](#)]
60. Pu, H.Y.; Broderick, A.E. Probing the Innermost Accretion Flow Geometry of Sgr A* with Event Horizon Telescope. *Astrophys. J.* **2018**, *863*, 148. [[CrossRef](#)]
61. Tiede, P.; Pu, H.Y.; Broderick, A.E.; Gold, R.; Karami, M.; Preciado-López, J.A. Spacetime Tomography Using the Event Horizon Telescope. *Astrophys. J.* **2020**, *892*, 132. [[CrossRef](#)]
62. Zhao, J.H.; Young, K.H.; Herrnstein, R.M.; Ho, P.T.P.; Tsutsumi, T.; Lo, K.Y.; Goss, W.M.; Bower, G.C. Variability of Sagittarius A*: Flares at 1 Millimeter. *Astrophys. J.* **2003**, *586*, L29–L32. [[CrossRef](#)]
63. Bower, G.C.; Dexter, J.; Markoff, S.; Gurwell, M.A.; Rao, R.; McHardy, I. A Black Hole Mass-Variability Timescale Correlation at Submillimeter Wavelengths. *Astrophys. J. Lett.* **2015**, *811*, L6. [[CrossRef](#)]
64. Bower, G.C.; Dexter, J.; Asada, K.; Brinkerink, C.D.; Falcke, H.; Ho, P.; Inoue, M.; Markoff, S.; Marrone, D.P.; Matsushita, S.; et al. ALMA Observations of the Terahertz Spectrum of Sagittarius A*. *Astrophys. J. Lett.* **2019**, *881*, L2. [[CrossRef](#)]
65. Liu, H.B.; Wright, M.C.H.; Zhao, J.H.; Brinkerink, C.D.; Ho, P.T.P.; Mills, E.A.C.; Martín, S.; Falcke, H.; Matsushita, S.; Martí-Vidal, I. Linearly polarized millimeter and submillimeter continuum emission of Sgr A* constrained by ALMA. *Astron. Astrophys.* **2016**, *593*, A107. [[CrossRef](#)]
66. Gravity Collaboration; Abuter, R.; Amorim, A.; Bauböck, M.; Berger, J.P.; Bonnet, H.; Brandner, W.; Clénet, Y.; Coudé Du Foresto, V.; de Zeeuw, P.T.; et al. A geometric distance measurement to the Galactic center black hole with 0.3% uncertainty. *Astron. Astrophys.* **2019**, *625*, L10. [[CrossRef](#)]
67. Raymond, A.W.; Palumbo, D.; Paine, S.N.; Blackburn, L.; Córdoba Rosado, R.; Doleman, S.S.; Farah, J.R.; Johnson, M.D.; Roelofs, F.; Tilanus, R.P.J.; et al. Evaluation of New Submillimeter VLBI Sites for the Event Horizon Telescope. *Astrophys. J. Suppl. Ser.* **2021**, *253*, 5. [[CrossRef](#)]
68. Wielgus, M.; Moscibrodzka, M.; Vos, J.; Gelles, Z.; Martí-Vidal, I.; Farah, J.; Marchili, N.; Goddi, C.; Messias, H. Orbital motion near Sagittarius A*. Constraints from polarimetric ALMA observations. *Astron. Astrophys.* **2022**, *665*, L6. [[CrossRef](#)]
69. Liska, M.; Chatterjee, K.; Tchekhovskoy, A.; Yoon, D.; van Eijnatten, D.; Hesp, C.; Markoff, S.; Ingram, A.; van der Klis, M. H-AMR: A New GPU-accelerated GRMHD Code for Exascale Computing With 3D Adaptive Mesh Refinement and Local Adaptive Time-stepping. *arXiv* **2019**, arXiv:1912.10192.
70. Mościbrodzka, M.; Gammie, C.F. IPOLE—Semi-analytic scheme for relativistic polarized radiative transport. *Mon. Not. R. Astron. Soc.* **2018**, *475*, 43–54. [[CrossRef](#)]
71. Younsi, Z.; Psaltis, D.; Özel, F. Black Hole Images as Tests of General Relativity: Effects of Spacetime Geometry. *arXiv* **2021**, arXiv:2111.01752.
72. Broderick, A.E.; Fish, V.L.; Johnson, M.D.; Rosenfeld, K.; Wang, C.; Doleman, S.S.; Akiyama, K.; Johannsen, T.; Roy, A.L. Modeling Seven Years of Event Horizon Telescope Observations with Radiatively Inefficient Accretion Flow Models. *Astrophys. J.* **2016**, *820*, 137. [[CrossRef](#)]
73. Gravity Collaboration; Abuter, R.; Amorim, A.; Anugu, N.; Bauböck, M.; Benisty, M.; Berger, J.P.; Blind, N.; Bonnet, H.; Brandner, W.; et al. Detection of the gravitational redshift in the orbit of the star S2 near the Galactic centre massive black hole. *Astron. Astrophys.* **2018**, *615*, L15. [[CrossRef](#)]
74. Roelofs, F.; Janssen, M.; Natarajan, I.; Deane, R.; Davelaar, J.; Olivares, H.; Porth, O.; Paine, S.N.; Bouman, K.L.; Tilanus, R.P.J.; et al. SYMBA: An end-to-end VLBI synthetic data generation pipeline. Simulating Event Horizon Telescope observations of M 87. *Astron. Astrophys.* **2020**, *636*, A5. [[CrossRef](#)]
75. Blecher, T.; Deane, R.; Bernardi, G.; Smirnov, O. MEQSILHOUETTE: A mm-VLBI observation and signal corruption simulator. *Mon. Not. R. Astron. Soc.* **2017**, *464*, 143–151. [[CrossRef](#)]
76. Natarajan, I.; Deane, R.; Martí-Vidal, I.; Roelofs, F.; Janssen, M.; Wielgus, M.; Blackburn, L.; Blecher, T.; Perkins, S.; Smirnov, O.; et al. MeqSilhouette v2: Spectrally resolved polarimetric synthetic data generation for the event horizon telescope. *Mon. Not. R. Astron. Soc.* **2022**, *512*, 490–504. [[CrossRef](#)]
77. Janssen, M.; Goddi, C.; van Bemmelen, I.M.; Kettenis, M.; Small, D.; Liuzzo, E.; Rygl, K.; Martí-Vidal, I.; Blackburn, L.; Wielgus, M.; et al. rPICARD: A CASA-based calibration pipeline for VLBI data. Calibration and imaging of 7 mm VLBA observations of the AGN jet in M 87. *Astron. Astrophys.* **2019**, *626*, A75. [[CrossRef](#)]
78. Gelaro, R.; McCarty, W.; Suárez, M.J.; Todling, R.; Molod, A.; Takacs, L.; Randles, C.A.; Darmenov, A.; Bosilovich, M.G.; Reichle, R.; et al. The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *J. Clim.* **2017**, *30*, 5419–5454. [[CrossRef](#)]
79. Paine, S. The am Atmospheric Model. 2019. <https://zenodo.org/record/3406496#.Y7017nZByUk> (accessed on 19 December 2022).
80. Farah, J.; Galison, P.; Akiyama, K.; Bouman, K.L.; Bower, G.C.; Chael, A.; Fuentes, A.; Gómez, J.L.; Honma, M.; Johnson, M.D.; et al. Selective Dynamical Imaging of Interferometric Data. *Astrophys. J. Lett.* **2022**, *930*, L18. [[CrossRef](#)]
81. Chael, A.A. Simulating and Imaging Supermassive Black Hole Accretion Flows. Ph.D. Thesis, Harvard University, Cambridge, MA, USA, 2019.
82. Bella, N.L.; Issaoun, S.; Roelofs, F.; Fromm, C.; Falcke, H. Expanding Sgr A* dynamical imaging capabilities with an African extension to the Event Horizon Telescope. *Astron. Astrophys.* **2023**, submitted.
83. Chael, A.; Johnson, M.D.; Lupsasca, A. Observing the Inner Shadow of a Black Hole: A Direct View of the Event Horizon. *Astrophys. J.* **2021**, *918*, 6. [[CrossRef](#)]

84. Bronzwaer, T.; Falcke, H. The Nature of Black Hole Shadows. *Astrophys. J.* **2021**, *920*, 155. [[CrossRef](#)]
85. Tiede, P.; Johnson, M.D.; Pesce, D.W.; Palumbo, D.C.M.; Chang, D.O.; Galison, P. Measuring Photon Rings with the ngEHT. *arXiv* **2022**, arXiv:2210.13498.
86. Dodson, R.; Rioja, M.J.; Jung, T.; Gómez, J.L.; Bujarrabal, V.; Moscadelli, L.; Miller-Jones, J.C.A.; Tetarenko, A.J.; Sivakoff, G.R. The science case for simultaneous mm-wavelength receivers in radio astronomy. *New Astron. Rev.* **2017**, *79*, 85–102. [[CrossRef](#)]
87. Rioja, M.J.; Dodson, R. Precise radio astrometry and new developments for the next-generation of instruments. *Astron. Astrophys. Rev.* **2020**, *28*, 6. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.